

# Modeling Cosmic-Ray propagation for Lithium, Beryllium, Boron, Nitrogen, and Helium with AMS-02 data

Master Thesis of

**Nathanael Weinrich**

At the Department of Physics  
Institute of Experimental Particle Physics (ETP)

In cooperation with  
Laboratoire de la Physique subatomique et Cosmologie (LPSC),  
CNRS/IN2P3, Grenoble, France  
and  
Université Grenoble-Alpes (UGA)

Reviewer:	Prof. Dr. Wim de Boer
Second reviewer:	Prof. Dr. Ulrich Husemann
Advisor:	Dr. David Maurin (LPSC) and Dr. Iris Gebauer (KIT)

Karlsruhe, 10.03.2020



---

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	About Galactic Cosmic radiation (CR)	3
1.2	Galaxy properties	4
1.3	Diffusion model for CR	6
1.3.1	Basic concept	6
1.3.2	One-dimensional thin-disc approach	8
1.3.3	The BIG, SLIM, and QUAINT benchmark models	14
1.4	The Alpha Magnetic Spectrometer 2 (AMS-02) experiment	16
1.4.1	Detector description	16
1.4.2	Flux measurements	18
<b>2</b>	<b>Studying Cosmic-Ray model parameters</b>	<b>21</b>
2.1	Techniques	21
2.1.1	About USINE	21
2.1.2	$\chi^2$ definition	21
2.1.3	MINUIT2 and MINOS	22
2.2	Nuisance parameters for nuclear cross sections	22
2.2.1	The NSS parametrization	23
2.2.2	Choosing the important cross sections to treat	23
2.2.3	Determining the nuisance parameter prior values	25
2.3	Data uncertainties	25
2.3.1	Possible data type uncertainties in USINE	25
2.3.2	Creation of covariance matrices	25
2.3.3	Impact on parameter constraints	26
2.3.4	Inspecting residuals of correlated data with the $\tilde{z}$ -score	27
2.4	Parameter space studies	29
2.4.1	Minimizers	29
2.4.2	Monte Carlo methods	31
<b>3</b>	<b>Universality of CR transport for Lithium, Beryllium, Boron, Nitrogen, and Helium</b>	<b>35</b>
3.1	Current CR transport studies with AMS-02 high-precision data	35
3.2	AMS data and other low-energy experiments	36
3.3	Compatibility of Lithium, Beryllium, and Boron	38
3.4	Helium and Nitrogen	41
3.5	MCMC application in the SLIM model for LiBeB/C	44
3.6	Perspective: Moving from flux ratios to direct fluxes	46

<b>4</b>	<b>Study of the galactic halo size with Beryllium data</b>	<b>49</b>
4.1	Status of halo size analysis and relation to Dark Matter . . . . .	50
4.2	Existing data sensitive to $^{10}\text{Be}$ . . . . .	50
4.3	Expected constraining power . . . . .	51
4.4	Direct determination or $L$ . . . . .	52
4.4.1	Data fit combinations . . . . .	52
4.4.2	Benchmark constraint with all data . . . . .	54
4.4.3	Comparison of $^{10}\text{Be}/^9\text{Be}$ data between experiments . . . . .	55
4.5	MCMC application in the SLIM model with all Beryllium data . . . . .	55
4.6	Perspective: Isotopic AMS-02 Beryllium data and other heavy radioactive nuclei . . . . .	57
<b>5</b>	<b>Conclusion</b>	<b>59</b>
<b>A</b>	<b>PyUSINE</b>	<b>61</b>
<b>B</b>	<b>Minimization techniques</b>	<b>65</b>
B.1	Minimizer algorithms . . . . .	65
B.1.1	Minuit with Migrad . . . . .	65
B.1.2	Powell's method . . . . .	65
B.1.3	Conjugate Gradient Method (CG) . . . . .	66
B.1.4	Naive Genetic approach . . . . .	66
B.2	Merging multiple MINUIT2 fit results . . . . .	67
B.3	Brief introduction to Markov Chain Monte Carlo (MCMC) . . . . .	68
B.3.1	Sampling a Markov Chain . . . . .	69
B.3.2	Metropolis Hastings . . . . .	70
B.3.3	Hamiltonian Monte Carlo . . . . .	70
<b>C</b>	<b>Rescaling parameters with <math>L</math></b>	<b>73</b>
<b>D</b>	<b>Other additional figures</b>	<b>77</b>
<b>E</b>	<b>Cross section nuisance parameter plots</b>	<b>83</b>

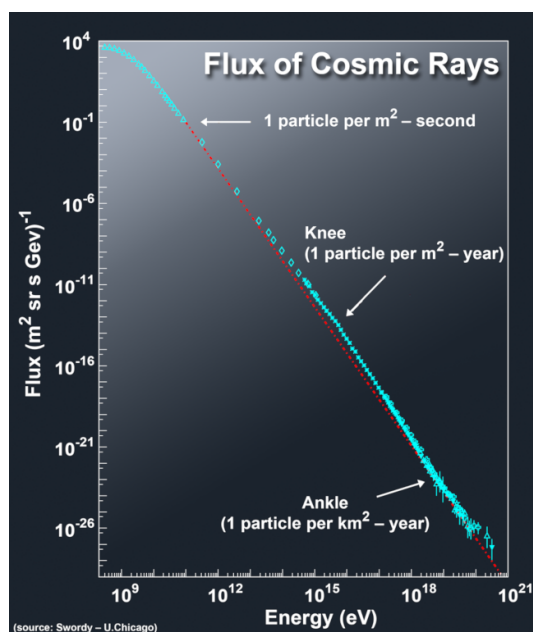
# Introduction

Cosmic Rays (CRs) are elementary particles of high energies, propagating through space. The spectrum is power-law decreasing with energy, but reaching up to  $\mathcal{O}(10^{20} \text{ eV})$ , presented in figure 1.1. The CR reaching the Earth can come from the sun, outside the solar system or even outside the Galaxy. The first one, also called solar radiation, is usually dominant at lower energies  $< 10^9 \text{ eV}$ . Above a certain break in the CR spectrum at  $\sim 10^{15} \text{ eV}$ , also known as the *Knee*, the sources of CR are suspected to be extra-galactic. In the intermediate regime we talk about Galactic Cosmic Ray, on which we will focus in the latter and refer to as GCR or simply CR. Between their galactic sources and an eventual collision with one of the rare, hard objects in space, they cover galactic scale distances during their  $\mathcal{O}(10 \text{ Myr})$  travel time. This connection of GCR to astrophysics on a galactic scale, makes them a useful asset of research and motivates their understanding.

This thesis is about the attempt of describing their transport mechanism in order to predict the CR fluxes as potentially measured above the Earth's atmosphere. In particular, the subjects studied are the fluxes of light and medium ions of atomic charge  $Z = 2 - 8$ . Most of the data discussed in this thesis come from the Alpha Magnetic Spectrometer 2 (AMS-02) experiment, operated on the ISS since 2011. The first data were released less than 4 years ago and are of unprecedented precision. This chapter gives a brief overview of the physics background of CR, the Galaxy and the important mechanisms in section 1.1 and 1.2. The phenomenological theory of CR modeling with a simplistic diffusion model is given in section 1.3. An overview of the AMS-02 detector is given in section 1.4.

## 1.1 About Galactic Cosmic radiation (CR)

This section gives an overview of the physics behind CR, based on the Particle Data Group review<sup>1</sup>. An incident CR particle hitting the terrestrial atmosphere interacts with the air and produces a cascade of shower particles. This breaks down the high energetic CR particles into a distribution of lower energetic products in the atmosphere, which can then be detected with air or surface based experiments. These studies can explore the energy spectrum of CR up to very high energies  $< 10^{12} \text{ GeV}$  (see figure 1.1) by measuring the lower energetic shower products (most prominently done by the Pierre Auger Observatory). But one of the disadvantages of this



**Figure 1.1:** Energy spectrum of the total CR flux above the atmosphere. The turquoise points are measurements of CR and the red dotted line is an interpolation with constant spectral index of 2.8.

<sup>1</sup><http://pdg.lbl.gov/2019/reviews/rpp2019-rev-cosmic-rays.pdf>

approach is, that the information of the CR compositions is mostly lost. The latter is crucial for CR research, which is why the experiments and CR modeling need to take place in free space outside, or rather above the Earth’s atmosphere, as well. This is the motivation of running the AMS-02 experiment in Earth’s orbit, which enables a direct and precise measurement of CR. The focus of this is the intermediate energy spectrum  $10^{-1} - 10^4 \text{ GeV}$ , which has a significantly higher abundance than the high energy spectrum, shown in figure 1.1.

Cosmic Rays can be classified into two main categories, primary and secondary particles. The *primaries* have direct astrophysical sources, in which they are produced and accelerated up to very high energies. This includes e.g. electrons, protons, Helium-4, Carbon, Oxygen, and Ion, as synthesized in stars. Due to nuclear interaction with intergalactic gas particles (mostly Hydrogen), *secondary* particles are produced as a spallation product. Lithium, Beryllium, and Boron are just three examples of such secondary species, that have a significant abundance in the CR spectrum without being (yet) linked to direct cosmic sources. Some species, like Nitrogen for instance, are also mixed, which means that they are both produced in direct sources and by interaction of heavier species with gas. Antiprotons and positrons are suspected to be mostly secondary particles, whereas a possible primary content is a subject of current debates.

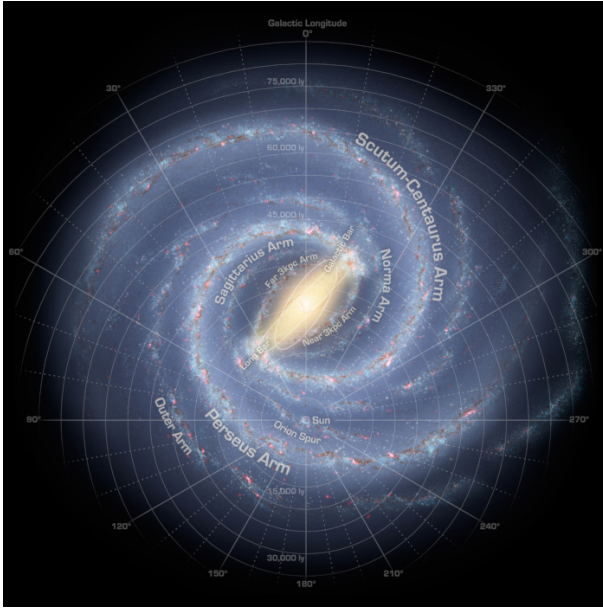
The absence of dense matter in most parts of the Galaxy allows particles of all kind to move freely through space with very little nuclear interaction. Unlike in the atmosphere with dense air, even heavy ions can propagate unhindered for a long time and eventually be measured above the atmosphere. The CR spectrum around the Earth thus includes all charged particles with a sufficient lifetime of  $\gtrsim 10^6 \text{ Myr}$ . This is why their transport mechanism is described on a galactic range of kilo-parsec ( $\text{kpc} = 30.9 \times 10^{15} \text{ m}$ ), linking them to principle properties of the Galaxy itself.

All the atoms found in CR spectrum are ions, as the energies of  $> \text{MeV}$  exceed the respective ionization energies by far. After their production in cosmic sources, they are deflected by galactic magnetic fields. Due to their fluctuations, the particle paths are mostly chaotic. This leads to an isotropic distribution of CR ions. Unlike uncharged  $\gamma$  particles, neutrinos, and neutrons, most of their direction information is lost. The diffusive propagation of CR can thus be modeled with the diffusion model in section 1.3.1, simplifying the transport description by neglecting direction information or individual particles.

The CR from outside the solar system is affected by solar wind from the sun, i.e. by a magnetized plasma expanding from it. Charged CR gets deviated and slowed down by this. The impact of this effect on CR is relatively high for lower energetic particles ( $< 10 \text{ GeV}$ ), hindering them from being detected at Earth to some extent, eventually modulating the measured low energy flux spectrum. The way this effect is treated in the model is described in section 1.3.2. Another obstacle for measuring low energy CR is the Earth’s magnetic field. If the energy is too low, the gyro-radius of the particle is smaller than the size of the Earth. This gyromagnetic cut-off effect introduces a lower-energy limit for measuring the particle above atmosphere, and explains the lower limit of AMS-02 data.

## 1.2 Galaxy properties

The Galaxy we live in, also known as ‘Milky Way’ is a regular spiral galaxy. The make-up and the relation to CR are explained in Maurin (2001). The visible mass, grouped to spiral arms, is mostly situated within a flat disk, rotating around its core center. Figure 1.2 shows an estimation of what this may look like from the outside (top). The solar system, and thus we as an observer, are situated inside the disk with a distance to the galactic core of  $R_{\odot} \sim 8 \text{ kpc}$ .



**Figure 1.2:** Top-view illustration of the Galaxy by NASA<sup>2</sup>.

<https://solarsystem.nasa.gov/resources/285/the-milky-way-galaxy>

Alongside big objects like stars, galactic clouds between the latter contribute to the make-up of the Galaxy as well, consisting of light particle gas. More than 90% of this interstellar material are neutral Hydrogen and Helium atoms. Most of this gas is grouped together to clouds, with varying densities of  $10\text{--}10^6\text{ cm}^{-3}$ , temperatures and coverage. The distribution of gas is expected to decrease in vertical and radial dimensions (Grenier et al. 2015). But instead of describing this heterogeneous distribution in details, an averaging approach is taken in the following model, omitting the radial and simplifying the vertical distribution. An average density of  $n \sim 1\text{ cm}^{-3}$  inside of a discrete disc (of half-height  $h = 100\text{ pc}$ ) throughout the whole Galaxy is taken. This approximation will be elaborated in the section 1.3.2, using the assumption that only knowing the surface density value  $\Sigma_{\text{ISM}} = \int dz n(z) = 2hn$  is sufficient.

As Supernovae are dying stars, they have the same distribution inside the galactic disc as the visible matter. This is important, as they are considered to be the main source of CR in the Galaxy. Inside its shock-fronts, particles are accelerated up to high energies of of  $\lesssim 10^3\text{ TeV}$  with first order Fermi acceleration. In reality, Supernovae are point like objects with a discrete distribution  $q(\vec{r}, t)$  in time and space. In the context of (mostly) stable nuclei with sufficiently large propagation times and path lengths, the assumption of the sources to be equally distributed in time and space is possible, given that the sources will be integrated over a large scale later in equation (1.8) and following. Just like the interstellar gas, the radial distribution is omitted later and only the spatial integral  $\Sigma_{\text{Source}} = \int dz q(z) = 2hq$  is kept, as described in section 1.3.2.

Inside and around the disc is a magnetic field with multiple phenomenological explanations. Important for CR propagation are not only the homogeneous global form but also the turbulent substructures. The fluctuations of the scale of  $\mathcal{O}(100\text{ pc})$  are minor to the galactic dimensions of  $\mathcal{O}(10\text{ kpc})$ . Angular distributions of CR become isotropic due to this distortion. This chaotic component is responsible for the diffusion of charged particles, motivating the CR model below. Another diffusive phenomenon, whose impact is subject to ongoing discussions, is the re-acceleration or charged particles by hydro-magnetic shock waves, propagation through the interstellar plasma.

These simplistic assumptions about the magnetic field, the interstellar gas, and the supernovae distribution are by no means a claim that the reality is that simple and generic. It is rather based

on the idea that a much simplified galaxy description can produce CR predictions in a sufficient way, demonstrated in this thesis.

### 1.3 Diffusion model for CR

The huge size of the Galaxy, the enormous quantity of CR and the lack of detailed information about individual sources, intergalactic gas and magnetic fields do not favor modeling CR via individual particle simulations. Instead, a more macroscopic approach is used here. Subject of this model are the energetic and spatial densities of the elements. In this thesis, a special version of the diffusion model will be used for modeling CR, the one-dimensional galactic disk approach (Jones et al. 2001), which is taken from Putze et al. (2009) and is also applied in Génolini et al. (2019) and Derome et al. (2019).

This is a comparably lightweight mathematical framework. However, it has proven over the last decades to be still very powerful in terms of CR modeling and keep up with modern CR data in terms of predicting capability. This thesis will thus not use the most complex, but rather the simplest sufficiently precise model and test its capabilities.

#### 1.3.1 Basic concept

The abundance of each ion in the CR spectrum is described as a spatial and energy density  $N_i(\vec{r}, E)$ , with  $i$  referring to a specific isotope. The justification of omitting the momentum direction of the particle density will be given later. Note that the abundance  $N_i$  is directly proportional to the flux  $\Phi_i$ . There are two different energy units used alongside each other. The two dominant ones are rigidity and kinetic energy per nucleon,  $R = \frac{p}{Z}$  and  $E_{k/n} = E_{\text{kin}}/A$  respectively, who are more or less conserved during nuclear interactions. Their conversion depends on the particle mass and charge which is discussed later in details.

**The diffusion equation** All the densities and the whole Galaxy state is taken as an equilibrium in time. Thus, one can express the system of differential equations in a reduced form as a diffusion equation (Berezinskii et al. 1990):

$$\begin{aligned} \frac{\partial N_i}{\partial t} = & \underbrace{+\vec{\nabla}(K_i\vec{\nabla}N_i)}_{\text{diffusion}} + \underbrace{q_i}_{\text{source}} - \underbrace{nv\sigma_i N_i}_{\text{inel. reac.}} + \underbrace{\sum_{j>i} nv\sigma_{ij} N_j}_{\text{production}} \\ & + \underbrace{+\vec{\nabla}(\vec{V}_c N_i)}_{\text{Gal. wind}} + \underbrace{\frac{\partial}{\partial E_{\text{kin}}} A_i}_{\text{energy changes}} - \underbrace{\frac{1}{\tau_i} N_i + \sum_{j>i} \frac{1}{\tau_{ij}} N_j}_{\text{radioactive decays}} = 0. \end{aligned} \quad (1.1)$$

In this equation  $v(E_{k/n})$  is the particle velocity,  $\sigma_i(E_{k/n})$  the cross section for inelastic reactions, and  $\sigma_{ij}(E_{k/n})$  the cross section for production (break-up, spallation, and fragmentation) of a nucleus  $i$  from a heavier one  $j$  by interaction with the interstellar gas of density  $n(\vec{r})$ . The galactic wind and energy changes are discussed later in 1.3.2 and 1.3.1 respectively. The (spatial) diffusion tensor  $K(\vec{r}, R) \equiv K(R)$  is assumed to be isotropic and location independent. It plays a major role in the diffusion model, which is why its parametrization is not trivial and further discussed in 1.3.1.

The mean lifetime  $\tau_i$  and  $\tau_{ij}$  describe the radioactive decay of species  $i$  or species  $j$  to  $i$  respectively. For stable nuclei with  $\tau = \infty$ , as the case for most species, the terms vanish. The only non-stable species relevant in this thesis is  $^{10}\text{Be}$  which will be discussed in chapter 4.



### Injection of CR

The main physical source of CR are supernovae remnants, accelerating particles in their shock fronts via first order Fermi acceleration, which motivates the power-law description. For each isotope  $i$  the source term  $q_i(\vec{r}, R)$  describes the resulting energy spectrum of this injection. The form

$$q_i(R) = \hat{q}_i \beta (R/(1 \text{ GV}))^{-\alpha_i} \quad (1.2)$$

is assumed, with a spectral index  $\alpha_i$  and a normalization  $\hat{q}_i$ , which are isotope specific. For pure secondaries is  $\hat{q}_i = 0$ . In this thesis, the spectral index is fixed to a common value  $\alpha$  for all the species. An exception to this is Helium-4, for which the formula (1.2) is modified as

$$q_{\text{He}}(R) = \hat{q}_{\text{He}} \beta^{\eta_s} (R/(1 \text{ GV}))^{-\alpha} \left(1 + (R/R_{qL})^{-10}\right)^{\Delta\alpha_{\text{He}}}, \quad (1.3)$$

using 3 additionally low-rigidity parameters  $R_{qL}$ ,  $\Delta\alpha_{\text{He}}$  and  $\eta_s$ . It this last bracket introduces a low-rigidity break at  $R_{qL}$  below which the slope is changed. This is for the reason, that  ${}^4\text{He}$ , and  ${}^3\text{He}$  (mostly produced via  ${}^4\text{He}$ ) are the only directly fitted fluxes in this thesis and that preliminary test showed discrepancies in the lower energy spectrum.

### Energy losses and gains

A definition of the term  $A_i$  from equation (1.1) is given in Putze et al. (2009) as

$$A_i = \underbrace{-(b_i N_i)}_{\text{losses}} + \underbrace{\frac{(\vec{\nabla} \vec{V}_c)}{3} \left( \frac{2E_0 + E_{\text{kin}}}{E_0 + E_{\text{kin}}} E_{\text{kin}} N_i \right)}_{\text{adiabatic expansion}} + \underbrace{K_p \left[ -\frac{(1 + \beta^2)}{E_{\text{kin}} + E_0} N_i + \beta^2 \frac{\partial N_i}{\partial E_{\text{kin}}} \right]}_{\text{re-acceleration}} \quad (1.4)$$

$$\text{with } K_p(p) = \frac{4}{3} \frac{V_a^2 \beta^2 p^2}{\delta(4 - \delta^2)(4 - \delta) \cdot K_0}, \quad (1.5)$$

whereas this form of energy diffusion coefficient  $K_p(p(R))$  is given in Maurin (2018).  $E_0 \sim A \cdot m_p$  and  $E_{\text{kin}} = E - E_0 = A \cdot E_{k/n}$  are the mass and kinetic energy respectively, with proton mass  $m_p$ .  $\vec{V}_a$  is the Alfvén velocity. The parameters  $\delta$ , and  $K_0$  are introduced later in the definition of  $K(R)$  in 1.3.1. The first term in (1.4) describes the continuous energy losses via ionization and coulomb interaction with  $b_i = dE_{\text{kin}}/dt$ . The second term describes the energy changes related to adiabatic expansion of the CR from the dense disc to the outer regions due to the galactic wind. The third term introduces the re-acceleration due to magnetic waves propagating with speed  $V_a$  through intergalactic plasma. In analogy to  $K(R)$  for spatial dimensions,  $K_p(p)$  describes the diffusion in energy.

### Parametrization of the diffusion coefficient $K(R)$

The diffusion coefficient  $K(R)$  plays a major role in the diffusion model as demonstrated in equation (1.1) and the illustrative solution later in (1.11). Its main form is a power-law

$$K(R) = K_0 \cdot (R/1 \text{ GV})^\delta \quad (1.6)$$

with normalization  $K_0$  and slope  $\delta$ .

However, the publication of more precise CR data in the last decade showed a spectral change of CR fluxes, which could be explained with a spectral break of the diffusion coefficient at

$\sim 300$  GV (Vladimirov et al. 2012). This may be due to different origin mechanisms of the magnetic turbulence. Taking this idea further, Génolini et al. (2019) introduced a more potent description

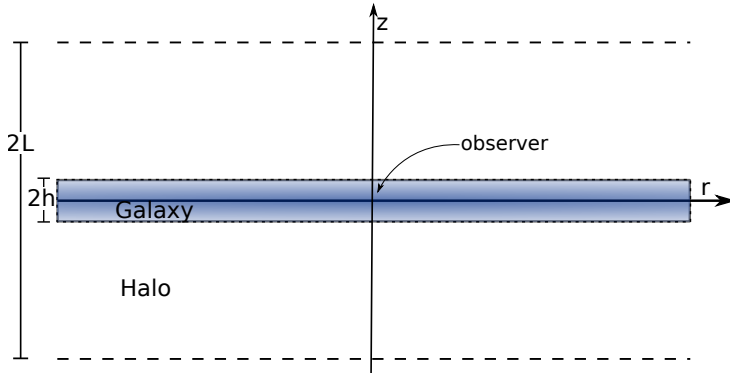
$$K(R) = K_0 \cdot \underbrace{\beta^\eta}_{\text{LE adjust.}} \cdot \underbrace{\left\{ 1 + \left( \frac{R}{R_l} \right)^{\frac{\delta_l - \delta}{s_l}} \right\}^{s_l}}_{\text{Low Rigidity break}} \cdot \left( \frac{R}{1 \text{ GV}} \right)^\delta \cdot \underbrace{\left\{ 1 + \left( \frac{R}{R_h} \right)^{\frac{\delta - \delta_h}{s_h}} \right\}^{-s_h}}_{\text{High Rigidity break}}, \quad (1.7)$$

including a low as well as a high rigidity break. They are described with a threshold rigidity  $R_{l,h}$ , a transition smoothness  $s_{l,h}$  an alternative slope  $\delta_{l,h}$  respectively. The parameter  $\eta$  adjusts the low-rigidity behavior, as the intermediate and high rigidity regions above several GV with  $\beta = 1$  are unaffected. This parametrization is further discussed in section 1.3.3.

In the energy region  $R_l < R < R_h$  the left and right brace-bracket in (1.7) are  $\sim 1$ , leaving  $K(R)$  in its initial form (1.6) (assuming that  $\beta|_{R_l} = 1$ ). Below  $R_l$  the left brace-bracket takes the form  $(R/R_l)^{\delta_l - \delta}$ , exchanging the dominant slope  $K(R) \sim R^\delta \rightarrow K(R) \sim R^{\delta_l}$ . In analogy for high rigidity  $R > R_h$ , the right brace-bracket takes the form  $(R/R_h)^{\delta - \delta_h}$ , changing the slope  $\delta \rightarrow \delta_h$ . The transition between these low, mid, and high-rigidity regimes at  $R_l$  and  $R_h$  respectively is tweaked by the smoothness factors  $s_l, s_h$  respectively, where a smaller value for the latter means a more rapid and a bigger value a smother transition. An illustration of  $K(R)$  is plotted in figure 1.6.

### 1.3.2 One-dimensional thin-disc approach

Solving the equation (1.1) is a non-trivial task, motivating a simplification of the spatial problem, without losing the capability of predicting fluxes correctly. One of the possible ways to go is the one-dimensional thin-disc approximation of CR transport, which is used throughout the thesis and explained in this section. Note that all the CR-propagation parameter values discussed and the studies carried out in this thesis are extremely biased by this geometric model and its fixed parameters.



**Figure 1.3:** Sketch of the Galaxy model. In the  $r$  dimension, the Galaxy has infinite range, introducing full isotropy in radial dimensions.

### Geometric assumptions

We suppose that the galaxy is a thin and infinite disc with half height  $h = 100$  pc with an evenly distributed density of sources  $q_i$  and intergalactic gas inside of it. A illustration is given in figure 1.3. Outside of this disc ( $|z| > h$ ) is a halo of height  $L \gg h$ . In this halo the CRs are only diffusing or decaying. Sources  $q_i$  and intergalactic gas  $n$  are not present outside the Galactic

disc. Particles leaving the halo also leave the magnetic field and do not return. In the following,  $n = 1 \text{ cm}^{-3}$  is taken, which is an approximation of the average distribution of gas in the galactic disc (Maurin 2001).

This is of course a very generic simplification. The distribution of gas (and thus of stars and supernovae) is rather estimated as  $\propto \exp[-(z/h_m)^2]$  with  $h_m = \mathcal{O}(100 \text{ pc})$  (Ferrière 2001). With the idea of a diffusive halo size  $L \gg h_m$  however, the exact form of the vertical distribution does not matter. The exact inner distribution becomes negligible in favor of only looking at vertically integrated values of the latter. The idea behind that is, that only the integrated density of ISM which the CR passes through is important, but not its distribution along the particle path, since the CR particle paths covered exceed  $h$  by far.

In the light of a rather small CR anisotropy for charged particles of  $\mathcal{O}(1\%)$ , the isotropy assumptions taken in the following should not have a big impact on local CR abundances. Following this, radial distributions are omitted, demanding full isotropy in the  $x$  and  $y$  dimensions. The latter assumption eliminates all the gradients in radial dimensions in equation (1.1), reducing the whole problem to the  $z$ -axis, eventually being a one-dimensional problem. Figure 1.3 shows a sketch the galactic geometry, on which of this model is based.

### Motivation and context

The results presented throughout this thesis used up to  $\lesssim 10^8$  CR-propagation calculations in total in order to minimize or sample flux predictions. This is only feasible (in limited resources) because the calculation time of the model could be reduced to  $\mathcal{O}(1 \text{ s})$ .

However, this does not mean that these are unreasonable simplifications. Describing the galaxy in such a manner is used throughout CR ion studies for decades (Jones et al. 2001) and nowadays still applied by different groups (Evoli et al. 2020; Génolini et al. 2019). This is for the reason that it is comparably easy to calculate but still sufficiently precise for CR ion flux prediction above earths atmosphere.

But there are still many limits of this approach, which are however not significant for this thesis. If for instance the diffusion or source spectrum is highly radial depended, the averaging *infinite disc* might struggle to describe this. Uncharged particles like neutrinos or  $\gamma$ -rays, coming from primary, secondary or tertiary sources, can also not be described with this approach. The same applies for very high energetic particles as soon as their gyro-radius exceeds the size of the halo.

More realistic and/or sophisticated descriptions of the galaxy and its geometry are indeed possible, which is shown by the introduction of a local bubble and the diffusion equation solution in two dimensions in Putze et al. (2010). It was briefly tested in Génolini et al. (2019) how compatible the 1D and 2D flat-disc models are, using the assumption of a galactic radius of 20 kpc and earths position at  $R = 8.5 \text{ kpc}$  from center and a halo-half size  $L = 10 \text{ kpc}$ . The differences for the parameters found were mostly inside their respective uncertainties.

But for this thesis, the one dimension disc model holds sufficient precision. The focus of this thesis is on species universality, the possibility of fitting the vertical halo size, and general methodological questions. Thus, a simple but yet sufficient model is needed. When attempting to study new questions with no direct, precedent example, being in the more simple environment is often an advantage. More-over, the 1D thin disc model used in this thesis was used with success and sufficient precision in (Génolini et al. 2019; Derome et al. 2019).

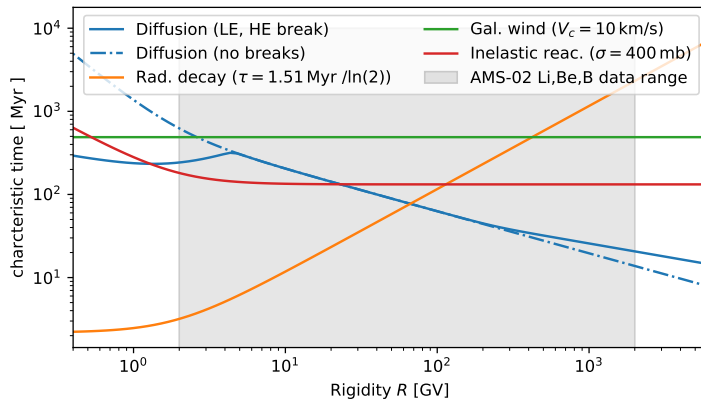
For nuclei, the prediction power of the 1D flat disc model as used in Génolini et al. (2019) is similar to that of the 3D propagation approach using Galprop in Boschini et al. (2020). This thesis thus follows the general idea of using the simplest model possible and testing how far one

can go with it.

If all the new phenomenons and methods in this thesis are well understood and under control, they can also be applied to more sophisticated, detailed, and calculation intensive Galaxy descriptions in future.

### Timescales of the physical processes involved

Being in a simplified regime allows to compare the magnitude of impact for the different terms in (1.1). This is done here by calculating a time scale estimation for the diffusion, galactic wind, radioactive decay, and nuclear production in the one-dimensional model. Figure 1.4 shows these time scale estimation and explains their respective calculations. The process with the smallest one is expected to dominate the transport mechanisms in the respective energy region.



**Figure 1.4:** Characteristic time for the basic physical impacts on CR, based on preliminary assumptions and only for illustration purpose. These estimations only apply for the 1D-disc model. For each term, the characteristic time is calculated as the inverse of the factors in front of  $N_i$  in equation (1.1). The gradients were evaluated as  $\vec{\nabla} \sim 1/\mathcal{O}(\vec{r}) \sim 1/L$ ,  $L = 5$  kpc. For the nuclear reactions, which only occurs inside of the disc, the time is re-scaled by  $L/h = 50$ , which is the volume ratio of halo to disc. One of the few, but important unstable elements is  $^{10}\text{Be}$  whose mean lifetime  $\tau = \tau_0 \cdot \gamma(R)$  is used here. The diffusion coefficient  $K(R)$  is calculated with the LiBeB/C best fit values for SLIM (see chapter 3), whereas the no-break line simply omits the low- and high-rigidity break factors in (1.7).

In this example of  $^{10}\text{Be}$ , it is found that for  $R \lesssim 100$  GV the radioactive decay clearly dominates any other process by far. The increase of the characteristic time for the decay, falling behind diffusion above 100 GV, can be explained by the fact, that the lifetime is multiplied by the relativistic Lorentz factor, increasing with energy. For stable species with infinitive lifetime, the orange graph in figure 1.4 would not exist (or be infinitively high). Above  $\sim 50$  GV, the diffusion dominates the other effects. Inelastic reactions and galactic wind only have an impact in low rigidity regions  $\lesssim 20$  GV and 8 GV respectively. In a low-rigidity break scenario for  $K(R)$  however (see section 1.3.1 for details), the latter remains important for this low-energy region as well.

The goal of this exercise is to illustrate which parameters are the most sensitive at which energies. In the case of one effect dominating the other, like diffusion at high energies, the respective parameters can thus be adjusted with data in the respective energy region. At low energies however, galactic wind, inelastic reactions and diffusion all have similar characteristic time scales. This could lead to a degeneracy between these parameters.

### Solution for diffusive transport

Without the galactic wind  $\vec{V}_c = 0$ , energy losses and gains  $A_i = 0$  and stable nuclei  $\tau_i = \infty$ , the one-dimensional approach provides an analytical solution, as shown in this part. A more detailed derivation of the following can be found in [Putze et al. \(2009\)](#) from where this is taken. As energy changes are not included here, the equation is solved for each energy independently. For the sake of readability the intrinsic energy dependencies  $N(R), Q(R), K(R), v(R), \sigma_i(E_{k/n})$  are dropped in the following. Due to the isotropy in radial dimensions  $\vec{\nabla}_{x,y} N = 0$  and the diffusion coefficient being totally isotropic, the whole equation (1.1) crumbles down to the  $z$  axis, as shown below. This does not mean that the diffusion itself takes place in one dimension, but rather that the amount of particles moving in the radial dimensions are in an equilibrium and do not change the densities. With the resulting flat disc and the isotropy simplifications

$$\int dz \{n(z), q(z), N(z)\} = \int dz 2h\delta(z) \{n(z), q(z), N(z)\}$$

$$\text{and } \vec{\nabla} K \vec{\nabla} = \frac{d}{dz} K \frac{d}{dz} = K \frac{d^2}{dz^2}$$

respectively, the equation (1.1) reduces to

$$\left( -K \frac{d^2}{dz^2} + 2h n v \sigma_i \delta(z) \right) N_i = 2h \delta(z) \left\{ Q_i = q_i + \sum_{j>i} n v \sigma_{ij} N_j \right\}. \quad (1.8)$$

Outside of the disc ( $|z| > h$ ), the equation breaks down to

$$\frac{d^2}{dz^2} N_i = 0 \quad \Rightarrow \quad N_i = a \cdot z + b, \quad a, b \in \mathbb{R}.$$

Given the condition  $N_i(z = \pm L) = 0$  and  $N_i$  being continuous within  $-L$  and  $L$  we get

$$N_i(z > h) = N_0 \left(1 - \frac{z}{L}\right), \quad N_i(z < h) = N_0 \left(1 + \frac{z}{L}\right) \quad (1.9)$$

with  $N_i^0 \equiv N_i(|z| \leq h)$ . The integration of (1.8) between  $-h$  and  $+h$  yields together with (1.9)

$$2h [Q_i - n v \sigma_i N_i^0] = -K \left. \frac{dN_i}{dz} \right|_{z=-h}^{z=+h} = +2K N_i^0 \frac{1}{L} \quad (1.10)$$

$$N_i^0 \equiv N_i^0(R) = \frac{2h Q_i(R)}{\frac{2K(R)}{L} + 2h n(R) \sigma_i(R)}. \quad (1.11)$$

The calculation of any abundance  $N_i$  relies on the calculation of all  $N_{j>i}$  in an iteration downwards from heavy to light nuclei using a cross-section network. Note that we, the observer, are at  $z \approx 0$ , leaving  $N_i(0) \equiv N_i^0$  as sufficient information for predictions later.

The complexity of the diffusion equation (1.1) with  $A_i \neq 0$  does not allow an analytical solution like this. It rather must be solved with a semi-analytical approach, as implemented the CR framework USINE. But the solution (1.11) remains a good approximation, that allows to understand the impact and relation of certain parameters.

### Galactic wind

The Galactic wind  $\vec{V}_c$  describes a general movement of the particles and is directed outwards of the Galaxy. An example for the several possible physical reasons are supernovae explosions or CR pressure. The calculation for  $\vec{V}_c \neq 0$  is analogue to the previous case. Instead it is started with

$$\left(-K \frac{d^2}{dz^2} + V_c(z) \frac{d}{dz} + 2h\nu\sigma_i\delta(z)\right) N_i = 2h\delta(z) \left\{ Q_i = q_i + \sum_{j>i} \nu\sigma_{ij} N_j \right\}. \quad (1.12)$$

The galactic wind is taken to point outwards and to be constant with  $V_c(z) = -V_c$  below and  $V_c(z) = +V_c$  above the disc. In a similar manner we start by solving the equation outside the disc yielding (1.13). Solving the equation inside of the disc completes the solution with (1.14) as

$$N_i(z) = N_i^0 \left[ 1 - e^{\frac{V_c}{K}(|z|-L)} \right] \quad (1.13)$$

$$\text{with } N_i^0 = \frac{hQ_i(0)}{V_c + h\nu\sigma_i \left( 1 - e^{-\frac{V_c}{K}L} \right)}. \quad (1.14)$$

Equation (1.14) finally takes a form that is similar to (1.11). For consistency one can check the case of  $V_c \rightarrow 0$  and find equation (1.11):

$$\lim_{V_c \rightarrow 0} \left[ 1 - e^{\frac{V_c}{K}x} \right] = -\frac{V_c}{K}x$$

$$N_i(z) \stackrel{V_c \rightarrow 0}{=} \frac{hQ_i \frac{V_c}{K}(-|z| + L)}{V_c - h\nu\sigma_i \frac{V_c}{K}L} = \frac{hQ_i}{\frac{K}{L} - h\nu\sigma_i} \left( 1 - \frac{|z|}{L} \right)$$

### The flux ratio approach

In this thesis only flux ratios of mainly secondary to mainly primary species are analyzed, for which the reason shall be given now. It is demonstrated with equation (1.11) or (1.14), that factorize into

$$N_i = T_i Q_i; \quad T_i = \frac{1}{\frac{K}{hL} + \nu\sigma_i} \text{ or } T_i = \frac{1}{\frac{V_c}{h} + \nu\sigma_i \left( 1 - e^{-\frac{V_c}{K}L} \right)} \quad (1.15)$$

respectively with their source terms  $Q_i$  (see eq. (1.8)) and the source independent transport terms  $T_i$ . As we only care about local abundances,  $z$  dependency is dropped here  $N_i = N_i^0$ . For a pure primary  $k$  it is now assumed that  $\sigma_{kj} = 0, j < k$  reflecting the assumption that  $q_k$  from (1.2) is the dominating source contribution:

$$N_k = T_k Q_k = T_k \hat{q}_k [R/(1\text{GV})]^\alpha. \quad (1.16)$$

For  $i$  being a pure secondary with  $\hat{q}_i = 0$ ,

$$N_i = T_i Q_i = T_i \sum_{j>i}^{\text{primary}} \nu\sigma_{ij} N_j \quad (1.17)$$

$$\frac{N_i}{N_k} = T_i \left[ \nu\sigma_{ik} + \sum_{j>i, j \neq k}^{\text{primary}} \nu\sigma_{ij} \frac{T_j \hat{q}_j}{T_k \hat{q}_k} \right]. \quad (1.18)$$

The simplified flux ratio (1.18) does no longer include the source parameter  $\alpha$ . Instead, the transport term  $T_i$  is directly sensitive to the flux ratio with some interdependence to other transport terms  $T_{k,j}$  from other species. Example of species justifying the assumption of being pure secondaries (i.e. having no direct source) are Lithium, Beryllium and Boron. In the assumption of  $k$  being the sole parent of  $i$ , the second term vanishes, making the flux ratio independent of source parameters all together. For this reason, secondary-to-primary flux ratios are a powerful way of studying the transport parameters.

### Energy scales

Instead of using the total energy for the description of densities, it is better to use either  $R = \frac{p}{Z}$  or  $E_{k/n}$ , the rigidity and kinetic energy per nucleon respectively. These two have the advantage of being more or less conserved during nuclear interactions, even if the total energy is not due to mass changes. The conversion between  $R$  and  $E_{k/n}$  relies on the particles mass to charge ratio  $A/Z$  with the relation

$$R \approx \frac{A}{Z} \sqrt{E_{k/n}(E_{k/n} + 2m_p)}, \quad (1.19)$$

with  $m_p$  being the proton mass (which is similar to the neutron mass and a reasonable simplification for this illustration purpose). The charge  $Z$  is unique for each element and can often be measured with a comparably good precision, illustrated in plot 1.8. In the AMS-02 regime however, the mass number  $A$  can only be determined with comparably bad precision, prohibiting an isotopic separation with cuts. A species with unique charge, which includes several isotopes of distinct mass, does not have a unique definition of  $A/Z$ . As an example, the conversion  $R \leftrightarrow E_{k/n}$  is exact for  ${}^6\text{Li}$  but not for  $\text{Li}$ , including  ${}^6\text{Li}$  and  ${}^7\text{Li}$  in mostly unknown isotopic fractions. In order to avoid uncertainties from energy scale conversions, each data is supposed to be analyzed in its respective experiment energy scale. The AMS-02 detector for example measures the particle rigidity, which is why  $R$  will be the dominant energy scale for data analysis in this thesis.

The theoretical propagation model above is mostly based on  $E_{k/n}$  and  $R$  in interdependence. This is however no problem, as the abundance of each isotope is calculated as a separate quantity, like  ${}^6\text{Li}$  and  ${}^7\text{Li}$ , for which conversion is unique. The summed species abundance  $\text{Li}$  is then yielded in the end, after all the iterative calculations are done. A detailed study about the impact of the energy scale choice for the data analysis with combined species is carried out in Derome et al. (2019).

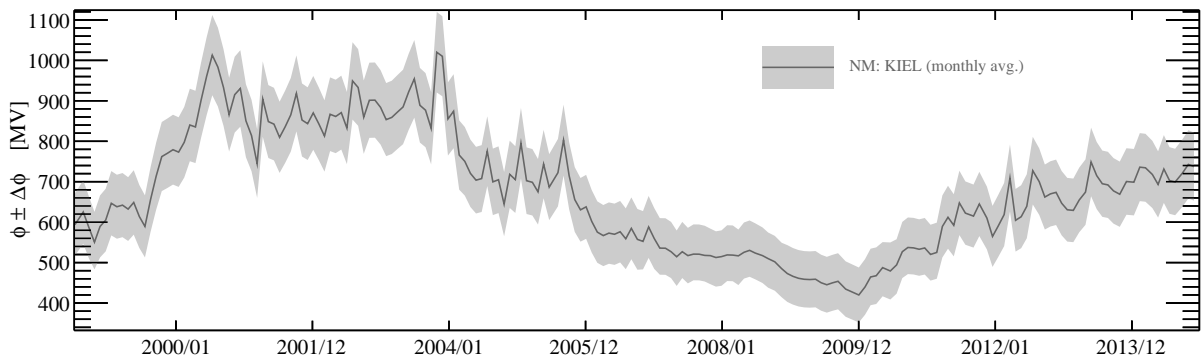
### Solar modulation

The impact of solar activity which de-accelerates CR that enters the solar system has already been mentioned in section 1.1. It mainly affects lower energetic particles with  $R < 10 \text{ GeV}$ . The data discussed in this thesis are to some extent in that area, which is why the impact and the way it is coped with needs to be briefly addressed. The flux calculations in this section all take place in the interstellar regime (outside of the solar cavity), eventually yielding  $\Phi_{\text{IS}} = A/ZR_{\text{IS}}^2 N(R_{\text{IS}})$  using the rigidity  $R_{\text{IS}}$  and CR density  $N(R_{\text{IS}})$ . As we, the observer, are situated inside of the solar system, the results of interstellar CR propagation need to be transferred into the local *top-of-atmosphere* (TOA) regime with flux  $\Phi_{\text{TOA}}$  and rigidity  $R_{\text{TOA}}$ , which our detectors eventually measure. The impact of solar modulation on the fluxes in this thesis is introduced via the force

field approximation (Gleeson & Axford 1968), which is used as

$$\Phi_{\text{TOA}}(R_{\text{TOA}}) = \frac{R_{\text{TOA}}^2}{R_{\text{IS}}^2} \cdot \Phi_{\text{IS}}(E_{\text{IS}}) \quad \text{with} \quad R_{\text{TOA}} = R_{\text{IS}} - \phi, \quad (1.20)$$

taken from (Caballero-Lopez & Moraal 2004). In practice, the whole propagation calculations of  $\Phi_{\text{IS}}$  are carried out in the USINE code and only the final results converted into the local  $\Phi_{\text{TOA}}$  observable. The formula (1.20) uses one additional parameter  $\phi$ [MV] which describes the time dependent solar flare activity and is a simple negative rigidity shift. Fluxes in the range of  $R \lesssim \phi < 1$  GV are affected by this, whereas the impact of  $\phi$  becomes negligible with increasing energy. The values are  $\phi \sim 400 - 1000$  MV and variate in time, demonstrated in plot 1.5. This leaves every data with a different value for  $\phi$ , respective to its acquisition period. Typically, these values are the integrated averages mean over the whole duration. In this thesis, the solar modulation  $\phi$  is introduced as a individual nuisance parameter (see section 2.1.2) for each experiment. The bias value is the integrated mean over the experiments data taking period with an uncertainty 100 MV. The data available only reaches until the end of 2014, which is why the solar modulation values after that time are linearly extrapolated using the last two data points, shown in figure 1.5. For instance, the AMS-02 data for Li, Be, B, C, N, and O, which was measured from 2011/05 to 2016/05, corresponds to a mean of  $\phi = 676$  MV.



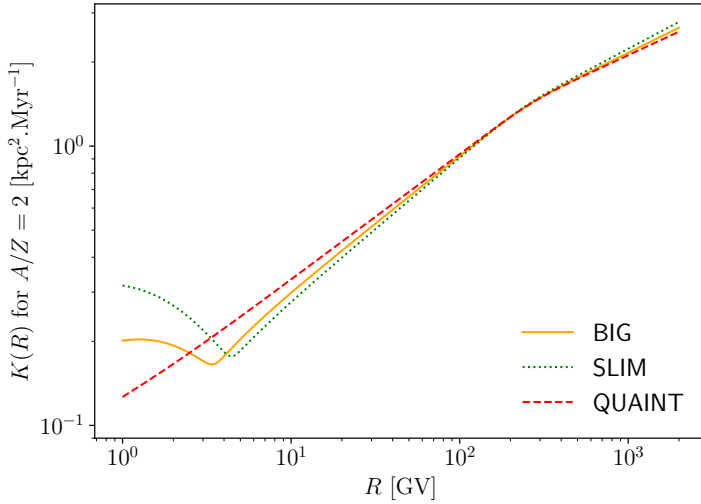
**Figure 1.5:** Solar modulation parameter  $\phi$  measured by the Kiel Neutron Monitor. Only a small fraction of the existing data period is shown.

### 1.3.3 The BIG, SLIM, and QUAINT benchmark models

Even though the diffusion model in equation (1.1) has been applied for years in this form, the parametrization of some parts has changed. The most recent contribution from Génolini et al. (2019) are the three models BIG, SLIM, and QUAINT applied throughout this thesis. Whereas the QUAINT model existed before, the new form of the diffusion coefficient (1.7) is the key for the SLIM and BIG model. Each of them represents a different approach to the diffusion model.

Defining these three models is very useful when studying different data and geometries over a longer period. The fact that this thesis now re-uses the same definitions for BIG, SLIM, and QUAINT from Génolini et al. (2019) allows to reuse fixed parameters like e.g. the injection normalization  $q_{C,O}$ , as they were fitted with the same models. This avoids redoing redundant analysis and allows to start with new studies more quickly. Furthermore, the results for the free parameters can also be compared within this thesis, and between this thesis and other papers in case of steady model definitions. However, as we still have three different models, following different physical ideas, the possibilities of studying and interpreting CR are not biased. It rather





**Figure 1.6:** Plot of the diffusion tensor  $K(R)$  for the three models BIG, SLIM, and QUAINT. The respective parameter values are taken from Génolini et al. (2019). The ratio of nuclei mass to charge  $A/Z = 2$  is taken as an example here. Note that these are just example best-fit values stemming from B/C adjustments.

helps to comment on the comparability of physical processes and CR data simply by looking at the differences in model fit goodness or parameter fit result reliability.

The fact that not one but three equally reasonable models are tested throughout this thesis in parallel is why best-fit parameters are always given for all three models individually throughout this thesis. Further studies can then use either one of them for model predictions.

The details for the parameters for the models BIG, SLIM, and QUAINT used in this thesis are in table 1.1, following Génolini et al. (2019). The impact of this parameter choice on the diffusion coefficient is demonstrated in figure 1.6.

**SLIM** is a diffusion-only approach, without any re-acceleration or galactic wind ( $V_a, V_c = 0$ ). Instead, both the high and low rigidity diffusion break of  $K(R)$  are used, leaving  $R_l, \delta_l$  for the latter as free parameters. As demonstrated in figure 1.4, this low-rigidity break gravely boosts the importance of diffusion for the description of low-rigidity fluxes, whereas the diffusive transport would be subdominant to other effects in the absence of a LE break. The idea of using a low-rigidity break at  $\sim 4$  GV for the sake of omitting re-acceleration is not a new idea and has already shown in Ptuskin et al. (2006) to be competitive with re-acceleration scenarios. As the definition (1.4) breaks down to the first term only, the remaining diffusion equation (1.1) can be solved completely analytically. With only 4 free parameters and no sophisticated energy diffusion, this model is usually the most stable and reliable when it comes to parameter minimization.

**BIG** has all the same features as SLIM. However, it includes re-acceleration and galactic wind. The latter two, and the low-rigidity break in the diffusion spectrum lead to a possible degeneracy or over-parametrization of the low-energy regime. Fitting BIG is thus an interesting thing to do since it has the possibility to combine all of these physical effects or go for either one only. With 6 free parameters, it is the *biggest* of the free models. On the downside, this model will show to be less reliant in terms of fitting it to complicated data combinations.

**QUAINT** also includes re-acceleration and galactic wind. But contrarily to the other two models, it does not include the low-rigidity diffusion break, omitting the first brace-bracket in equation (1.7). The flexibility in low-rigidity regime is instead provided by  $\eta$  being a free parameter. This model was already used in a similar form in Maurin et al. (2001); Ptuskin et al.

(2006). However, it usually needs a large re-acceleration (via  $V_a$ ) in order to describe the low-rigidity regime. This correlates the CR spectra over a large energy range, increasing parameter correlations (Génolini et al. 2019) and eventually challenging the fitting algorithm. Although it has only 5 free parameters and thus less than BIG, it is equally or even more difficult to fit in complicated configurations, due to its usually higher  $V_a$  and  $V_c$  values.

Parameter	Unit	BIG	SLIM	QUAINT
Intermediate-rigidity parameters				
$K_0$	[kpc <sup>2</sup> Myr <sup>-1</sup> ]	✓	✓	✓
$\delta$	[-]	✓	✓	✓
Low-rigidity parameters				
$V_c$	[km s <sup>-1</sup> ]	✓	n/a	✓
$V_a$	[km s <sup>-1</sup> ]	✓	n/a	✓
$\eta$	[-]	1	1	✓
Low-rigidity break				
$\delta_l$	[-]	✓	✓	n/a
$R_l$	[GV]	✓	✓	n/a
$s_l$	[-]	0.04	0.04	n/a
High-rigidity break				
$\delta_h$	[-]	0.18	0.19	0.17
$R_h$	[GV]	247.2	237.1	269.8
$s_l$	[-]	0.04	0.04	0.04
Geometry parameter				
L	[kpc]	5	5	5
$\sum_{\text{free}}$		6	4	5

**Table 1.1:** Table of important model parameters. The ✓ show the free parameters for the respective models. The total number of free parameters is given in the last row. Numerical values indicate that this parameter is fixed. The values for the rigidity break parameters  $s_l$ ,  $\delta_h$ ,  $R_h$  and  $s_h$  where taken from Génolini et al. (2019). The halo size  $L$  is chosen after some preliminary studies and will be justified later in chapter 4 of this thesis.

## 1.4 The Alpha Magnetic Spectrometer 2 (AMS-02) experiment

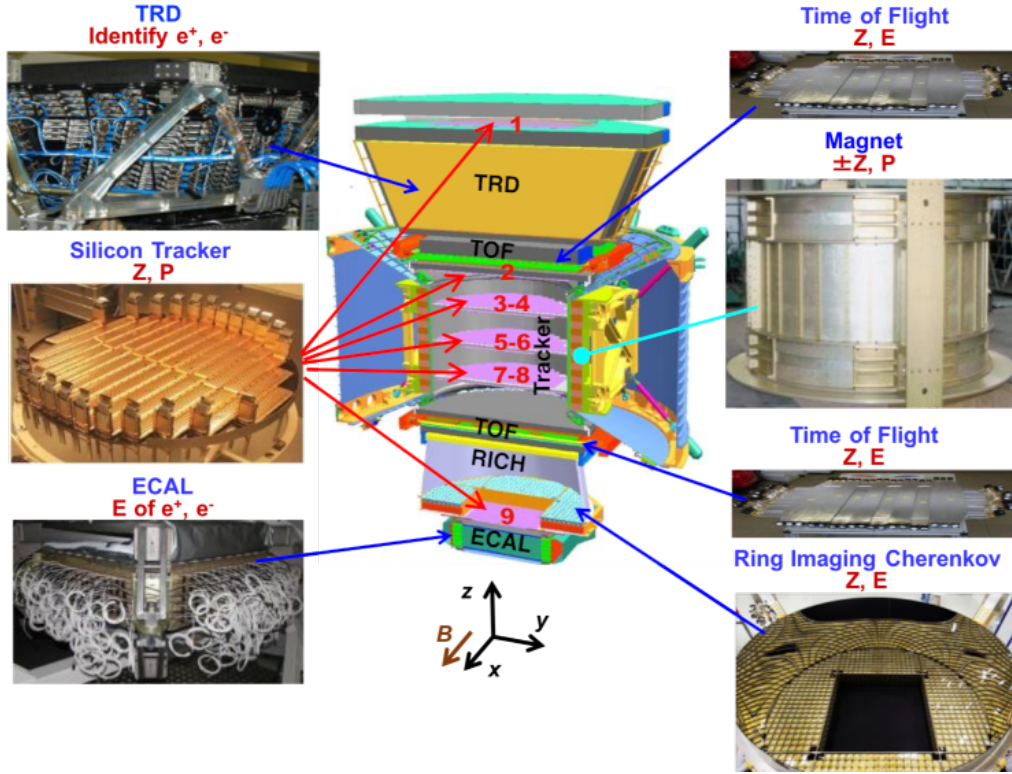
The Alpha Magnetic Spectrometer 2 is a state-of-the-art particle detector, operated as an external module mounted on the International Space Station (ISS). Due to its design, its extra-terrestrial environment and the long data-taking period (from 2011 ongoing), AMS-02 has an unprecedented precision of  $\mathcal{O}(1\%)$  for above-atmosphere CR flux measurements in the GV to TV rigidity range.

### 1.4.1 Detector description

As shown in figure 1.7, the detector consists of five independent sub-detectors, that are briefly described in the following. Some important physical quantities are thus measured redundantly. This subsection is based on the official website for the experiment<sup>3</sup> and the research group at the RTWH Aachen<sup>4</sup>.

<sup>3</sup><https://ams02.space/detector>

<sup>4</sup><https://www1b.physik.rwth-aachen.de/~schael/AMS.html>



**Figure 1.7:** Overview of the AMS-02 detector and its sub-detectors from <https://ams02.space/detector>.

**Transition Radiation Detector (TRD)** By measuring the transition radiation of penetrating particles, the TRD can distinguish those of the same charge by their mass. The main purpose is the separation of electrons/positrons and (anti-)protons, who share the same absolute charge but have a different mass. The probability for the particle in each layer to produce transition radiation is highly dependent on the Lorentz factor  $\gamma$ . Due to their mass difference, the latter differs between electrons and protons of the same energy by several orders of magnitude.

The TRD is built up by 20 similar, stacked layers, each consisting of PE fibers and proportional gas tubes to catch the emitted transition radiation. The sum of all the gas tube signals thus allows a binary segregation of light electrons/positrons from heavy (anti-)protons.

**Permanent Magnet and Silicon Tracker** The experiment is in a permanent magnet of 1.4 kG in horizontal direction. This strong field allows rigidity measurements via the curvature of particles paths. The latter is tracked within the detector with a combination of nine layers of high-resolution silicon detectors (horizontal resolution of  $\mathcal{O}(10\mu\text{m})$ ). The first and last layers are placed above the TRD and below the RICH, providing the highest possible detection range within the detector. The other layers are inside the magnet, where the Lorentz force is the strongest and the particles are deflected the most.

The signals of all the sub-layers allow a global path reconstruction and thus a recalculation of the particle rigidity. In addition to its position, the tracker layers also collect a relatively precise charge information out of its detector signal strengths. This can be used to distinguish elements.

**Time Of Flight (TOF)** The TOF counter consists of four planes of scintillating paddles, of which two are placed above and two below the inner tracker. The four planes together work as a trigger for charged particles. The multiple measurements of the ionization energy loss yield a precise charge information. Using the precise time measurement of the time difference between the signals in the top and bottom planes, the particle velocity  $\beta = v/c$  is found.

**Ring Imaging Cherenkov Detector (RICH)** The velocity  $\beta$  and charge  $Z$  of the particle are measured in the RICH. The top of the RICH is a radiator plane, consisting of either sodium-fluoride in the center or silica aerogel in the outer regions. Below the empty expansion volume is a grid of photo-multipliers (PMTs), covering the bottom of the volume except for an inner square sparing the ECAL below.

Particles with a  $\beta$  higher than the inverse of the dielectric constant produce a cone of Cherenkov radiation when penetrating the radiator material. The PMT signals are then used to geometrically reconstruct the cone and thus its opening angle. The latter is directly related to  $\beta$ , yielding a velocity measurement. With the proportionality of the PMTs signal amplitude to  $Z^2$ , another charge estimation is given.

**Electromagnetic Calorimeter (ECAL)** With a depth of 17 radiation lengths for electrons and positrons, their respective shower shape and energy can accurately be reconstructed in the ECAL. It consists of many layers of lead foils followed by scintillating fibers. The geometric information is given because the fibers alternate their orientation in x and y.

### 1.4.2 Flux measurements

This section gives a brief idea of how the fluxes for different elements are obtained using the AMS-02 detector signals. In order to illustrate this, the recent analysis for Lithium, Beryllium, and Boron ([Aguilar et al. 2018a](#)) is taken as an example.

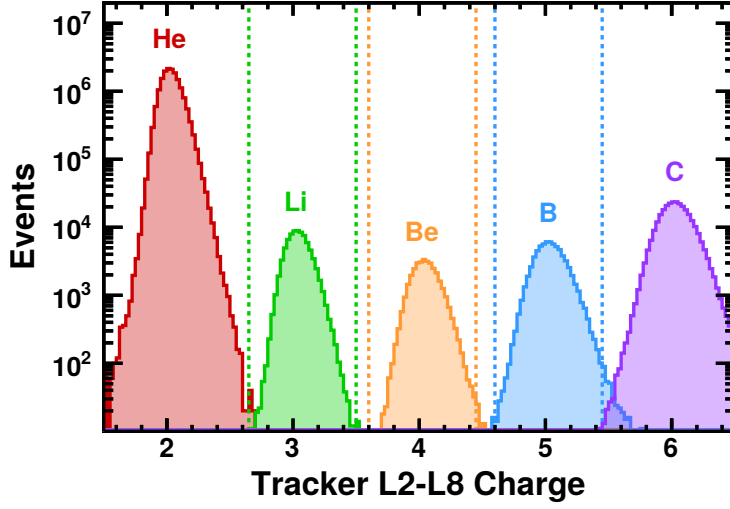
#### Species abundances

Before taking a particle event into account, it must fulfill several criteria. A good path reconstruction in the tracker, its flight direction, and the incident angle are three important examples. The next selection is that the charge measurements for the top tracker plate, the inner tracker plates and the TOF are compatible with the respective charge  $Z = 3, 4, 5$  for Li, Be, B respectively. In the figure 1.8 this selection is illustrated for the inner tracker plates. The multiple charge measurements function not only as an iterative selection chain, but also as a mean of cross validation and interdependent calibration. As shown in figure 1.8, the pre-selection of events from top-, bottom-tracker, and TOF can help to pick the selection criteria for the inner tracker.

All these steps are not only applied on the real CR data measurements, but also for a great sample of Monte Carlo CR events, using a simulation of the AMS-02 detector to estimate the event signals. The definition of the isotopic flux  $\Phi_i$  for the rigidity bin  $i$  with  $R = [R_i, R_i + \Delta R_i]$  is

$$\Phi_i = \frac{N_i}{A_i \epsilon_i T_i \Delta R_i} \quad (1.21)$$

where  $N_i$  is the quantity of total events selected to be in the rigidity bin  $i$  (by the tracker) after all the selection cuts. The effective acceptance  $A_i$  is the ratio of the simulated particles to the ones remaining after the respective selection.  $\epsilon_i$  is the trigger efficiency and  $T_i$  the exposure time



**Figure 1.8:** Charge distribution measured by the inner tracker plates (labeled as 2-8 in figure 1.7). The colors for the events correspond to the charge selection of the top and bottom tracker and the lower TOF. The vertical lines are the charge selection cuts values that are used to further select the particles with the inner tracker. This figure is from the Supplemental material of Aguilar et al. (2018a).

(effective data taking period). Flux ratios like B/C, which are mainly used in this context, are simply the ratio  $F_i^{\text{Be/B}} = \Phi_i^{\text{Be}} / \Phi_i^{\text{B}}$ .

### Isotopic fractions of species

Isotopic fractions, like  $^{10}\text{Be}/^9\text{Be}$  or  $^{10}\text{Be}/\text{Be}$ , are another difficult thing to reconstruct. Getting a precise measurement of Be flux is feasible, because it can be selected by charge  $Z$ , as described above. The selection of  $\text{Be} \rightarrow ^7\text{Be}, ^9\text{Be}, ^{10}\text{Be}$  depends on the mass selection capabilities of the detector. It can only be indirectly found in AMS-02 via  $\beta$ , using the relation  $A = \frac{RZ}{m_p} \sqrt{\beta^{-2} - 1}$ . The uncertainties for the  $\beta$  measurements however make the mass estimations overlap, which does not allow a simple cutting method like for  $Z$ . Instead, approaches like mass template fits are a possibility. The difficulty of obtaining isotopic fluxes (and fractions) are reflected in the big uncertainties of the respective data, whereas species fractions like Be/B are much more precise, which can be observed in section 4.3.



## Studying Cosmic-Ray model parameters

With a theoretical model prediction for CR and experimental data, those two elements can now be brought together in order to constrain the model parameters. For this non-trivial task a lot of technical details concerning the data, the model, and the fitting process must be taken into account. A brief introduction to the general analysis setup, mainly relying on the USINE framework (Maurin 2018), is given in section 2.1. All studies in this thesis are done with USINE v3.5. In section 2.2 the uncertainties of nuclear cross sections are discussed, introducing additional nuisance parameters. Data uncertainties are then discussed in section 2.3. Section 2.4 shows studies of the performance of several analysis techniques. In addition to this, a newly implemented Monte Carlo method is introduced.

### 2.1 Techniques

#### 2.1.1 About USINE

USINE<sup>1</sup> is a library for Galactic CR propagation models. It contains several semi-analytical propagation models, including the one-dimensional diffusion model used in this thesis. It provides routines for the calculation of cosmic ray fluxes or flux ratios for any nuclei ( $Z < 30$ ). This is done by solving the differential equations, presented in the previous chapter, with partly numerical methods. A cross-section reaction network is then used to simulate the secondary production in an iteration from heavy to light nuclei. The advantages are among other things the performance as it is written in pure C++ and the flexibility.

Any parameter within the model can potentially be fitted, fixed to a certain value, or set as a nuisance parameter with bias. The minimization is done with MINUIT2 and MINOS from the ROOT framework, that were integrated into USINE. Later in this thesis, the newly Python-interface PYUSINE allows the combination the CR propagation prediction capabilities of USINE with any Python routine, such as custom minimizers or Monte Carlo methods.

#### 2.1.2 $\chi^2$ definition

All CR model fits are performed with the  $\chi^2$  minimization method. Its definition is described in Maurin (2018) and implemented in the USINE-code as

$$\chi^2(\vec{\theta}, \vec{\eta}) = \chi_{\text{data}}^2(\vec{\theta}, \vec{\eta}) + \chi_{\text{nucl.}}^2(\vec{\eta}) \quad (2.1)$$

$$\chi_{\text{data}}^2(\vec{\theta}, \vec{\eta}) = \sum_k^{\text{datasets}} (\vec{y}_k^{\text{data}} - \vec{y}_k^{\text{model}}(\vec{\theta}, \vec{\eta})) C_k^{-1} (\vec{y}_k^{\text{data}} - \vec{y}_k^{\text{model}}(\vec{\theta}, \vec{\eta})) \quad (2.2)$$

$$\chi_{\text{nucl.}}^2(\vec{\eta}) = \sum_i^{n_{\text{nucl.}}} \left( \frac{\eta_i - \eta_i^*}{\sigma_{\eta_i}^*} \right)^2. \quad (2.3)$$

$C_k$  is the covariance matrix of each data set, whereas  $\vec{y}_k$  contains the values of the energy bins for the data set  $k$ .  $n_{\text{nucl.}}$  is the amount of nuisance parameters. The case of uncorrelated data

<sup>1</sup><https://dmaurin.gitlab.io/USINE/>

corresponds to a diagonal matrix  $C_k$  that breaks up the product into a sum over the energy bins.  $y_k$  may be either a flux or a flux ratio of either species, like Be, or isotopes like  $^{10}\text{Be}$ .  $\vec{\theta}$  and  $\vec{\eta}$  are the tested model and nuisance parameters respectively. The nuisance parameters have a Gaussian prior  $(\eta_i^*, \sigma_{\eta_i}^*)$ , with the latter being individual choices for every nuisance parameter  $i$ .

The two ratios

$$\chi^2/n_{\text{dof}} \quad \text{and} \quad \chi_{\text{nui.}}^2/n_{\text{nui.}} \quad (2.4)$$

are indicators of the goodness of the model fit to the data with the degrees of freedom being defined as  $n_{\text{dof}} = n_{\text{datapoints}} - n_{\text{freepar.}}$ . Note that each nuisance parameter counts as one additional free parameter, whereas its bias counts as an additional pseudo-data-point as well. This is why  $n_{\text{dof}}$  is always independent of the number of nuisance parameters involved in the fit. The fraction  $\chi^2/n_{\text{dof}}$ , which is often referred to as the reduced  $\chi^2$ , shows how well the model describes the data. If it is greater than 1, the model can not sufficiently describe the data. If it is below 1, the data uncertainties are over-estimated. The fraction indicates the usage of nuisance parameters for the fit, showing up to which proportion the penalty of picking a value next to the bias is accepted in order to improve the overall fit. If the values for the nuisance parameters are not changed during the fit, this ratio stays 0. If it is above 1, the nuisance parameters have deviated more than their respective bias range  $\sigma_{\eta}^*$  predicted.

### 2.1.3 Minuit2 and Minos

The MINUIT2 package provides routines for the minimization of the  $\chi^2$  function in the parameter space. It uses the Migrad algorithm for minimization, further explained in B.1.1. Alongside the best fit, this also provides a linear parameter uncertainty estimations, using the Hesse matrix. The latter ones will be referred to as *migrad errors* in this thesis and are not considered to be sufficiently reliable or precise for our analysis.

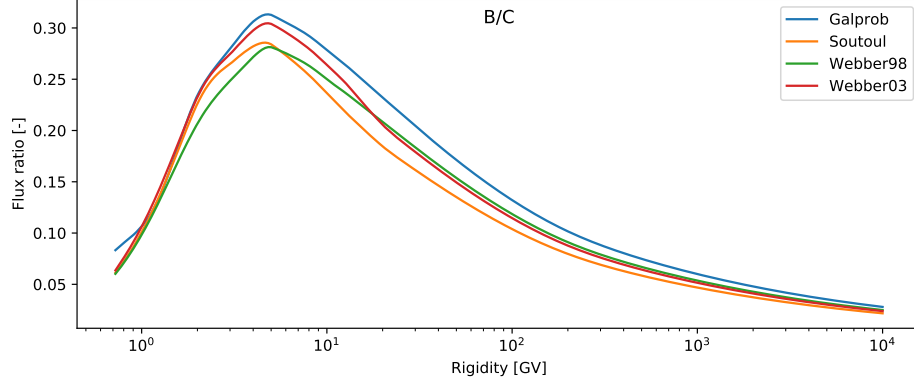
The MINOS algorithm, also implemented in the MINUIT2 package, can be used to better estimate the parameter uncertainties for a previously found best-fit result, as described in James & Winkler (2018). Its approach is searching for the upper and lower tolerance value  $\theta_{t,b}^k$  respectively for each parameter  $\theta^k$  with  $\tilde{\chi}^2(\theta_{t,b}^k) = \chi_{\text{min}}^2 + 1$ . While varying the parameter  $\theta^k$ , all the other free parameters are refitted each time to find the corresponding minimal  $\tilde{\chi}^2(\theta^k)$ . In order to accelerate this procedure, the other parameters are predicted using the covariance matrix. The latter is provided by the Migrad algorithm, as described in B.1.1. Contrarily to *migrad errors*, these uncertainties  $\sigma_{t,b}^k = \theta_{t,b}^k - \theta_{\text{bestfit}}^k$  can be asymmetric. Parameter uncertainty estimations found by this method are referred to as *minos errors* and are chose over *migrad errors* as the standard in this thesis. The reconstruction of a minos like covariance matrix by combination of minos errors and the Migrad covariance matrix is explained in B.2.

## 2.2 Nuisance parameters for nuclear cross sections

The nuclear cross-section (XS) data, which are used in order to calculate CR fluxes, have large uncertainties. Figure 2.1 demonstrates that using different parametrization from the literature can cause significant deviations in flux calculations (here  $\sim 10\%$  for B/C). This is in some rigidity regions larger than the data uncertainties and thus a problem.

For a full discussion of nuclear cross sections, see Génolini et al. (2018). In order to propagate the influence of the XS uncertainty to the fluxes of interest, nuisance parameters for those XS are imposed. This allows the cross sections to differ from their given values, with the chance of a better agreement of our models to the data.





**Figure 2.1:** B/C flux ratios calculated using USINE and the SLIM model. The parameters are the same for all curves with a different set of production XS for each, namely Galprob (Moskalenko et al. 2001; Moskalenko & Mashnik 2003), Soutoul (private communication between A. Soutoul and D. Maurin), Webber 98 (Webber et al. 1998a,b,c), Webber 03 (Webber et al. 2003).

### 2.2.1 The NSS parametrization

Following Derome et al. (2019), the *NSS* system is used throughout this thesis, modifying the nuclear cross sections via a **N**ormalization, an energy-**S**caling and a low-energy **S**lope with

$$\sigma_{\text{new}}^{\text{Norm.}}(E_{k/n}) = \text{Norm} \times \sigma_{\text{ref}}(E_{k/n}) \quad (2.5)$$

$$\sigma_{\text{new}}^{\text{Scale}}(E_{k/n}) = \sigma_{\text{ref}}(E_{k/n} \times \text{Scale}) \quad (2.6)$$

$$\sigma_{\text{new}}^{\text{Slope}}(E_{k/n}) = \begin{cases} \sigma_{\text{ref}}(E_{k/n}) & \text{if } E_{k/n} \geq E_{k/n}^{\text{thresh.}}; \\ \sigma_{\text{ref}}(E_{k/n}) \times \left( \frac{E_{k/n}}{E_{k/n}^{\text{thresh.}}} \right)^{\text{Slope}} & \text{otherwise.} \end{cases} \quad (2.7)$$

This channels the XS uncertainties to few parameters that can be fitted (free or as nuisance), modified and inspected.

### 2.2.2 Choosing the important cross sections to treat

There are multiple sets of cross section data for the production and inelastic destruction of nuclear isotopes. Even though there are thousands of different nuclear reactions involved in CR transport and its reaction network, only the few with direct impact for the species studied are taken into account in the latter. For each secondary species of interest (Li, Be, B, N,  $^3\text{He}$ ) the dominant reactions, according to Génolini et al. (2018), are studied using the SLIM model. The criteria is not the impact on the flux itself, but rather on the flux ratios, Li/C, Be/C, B/C, N/O and  $^3\text{He}/^4\text{He}$  respectively. Following Derome et al. (2019) the approach for finding the most critical reactions is:

1. For each reaction type (production, inelastic), pick one XS data set as reference.
2. For a given element, look up the contributions of the reactions in Génolini et al. (2018). Pick the most important ones each for production and inelastic destruction.
3. For each of these reactions, exchange the specific cross section data in the reference data set with those from another data set.

**Table 2.1:** Nuclear reactions for which nuisance parameters introduced. The values in the brackets are the maximum discrepancies (to the reference) found in the analysis above for the flux ratios (Li/C, Be/C, B/C,  $^3\text{He}/^4\text{He}$ , N/O). The percentages in bold indicate the reactions with the biggest impact per flux ratio and reaction type (Inelastic, Production); these mark the choice of nuisance parameters when the respective secondary species is fitted.

Element	Reaction	Prod. contrib.	Max. impact from $\Delta\sigma/\sigma$	Norm. $\mu \sigma$	Scale $\mu \sigma$	Slope $\mu \sigma$
<i>Li</i>	$^{16}\text{O}+\text{H}$		(1.2%)	1.03   0.04	0.7   0.5	n/a
	$^{12}\text{C}+\text{H}$		( <b>1.3%</b> )	1.01   0.04	0.8   0.5	n/a
	$^6\text{Li}+\text{H}$		(0.8%)	1.02   0.04	0.7   0.4	n/a
	$^{16}\text{O}+\text{H}\rightarrow^6\text{Li}$	14%	( <b>6.8%</b> )	0.89   0.28	n/a	0.00   0.15
	$^{12}\text{C}+\text{H}\rightarrow^7\text{Li}$	12%	(3.9%)	0.90   0.12	n/a	0.03   0.15
	$^{12}\text{C}+\text{H}\rightarrow^6\text{Li}$	14%	(4.7%)	0.87   0.15	n/a	0.00   0.15
<i>Be</i>	$^{16}\text{O}+\text{H}$		(0.9%)	1.03   0.04	0.7   0.5	n/a
	$^{12}\text{C}+\text{H}$		(1.4%)	1.01   0.04	0.8   0.5	n/a
	$^9\text{Be}+\text{H}$		(1.1%)	0.95   0.06	0.7   0.4	n/a
	$^7\text{Be}+\text{H}$		( <b>2.7%</b> )	1.10   0.10	0.7   0.4	n/a
	$^{16}\text{O}+\text{H}\rightarrow^9\text{Be}$	5%	(3.2%)	1.00   0.30	n/a	0.00   0.15
	$^{16}\text{O}+\text{H}\rightarrow^7\text{Be}$	18%	( <b>7.2%</b> )	0.85   0.15	n/a	0.00   0.15
	$^{12}\text{C}+\text{H}\rightarrow^9\text{Be}$	9%	(5.9%)	0.87   0.20	n/a	0.03   0.15
	$^{12}\text{C}+\text{H}\rightarrow^7\text{Be}$	16%	(4.0%)	1.00   0.25	n/a	0.00   0.15
<i>B</i>	$^{16}\text{O}+\text{H}$		(0.8%)	1.03   0.04	0.7   0.5	n/a
	$^{12}\text{C}+\text{H}$		(1.0%)	1.01   0.04	0.8   0.5	n/a
	$^{11}\text{B}+\text{H}$		( <b>1.7%</b> )	0.98   0.04	0.7   0.4	n/a
	$^{16}\text{O}+\text{H}\rightarrow^{11}\text{B}$	18%	(4.0%)	0.96   0.18	n/a	0.00   0.15
	$^{12}\text{C}+\text{H}\rightarrow^{11}\text{B}$	34%	( <b>7.1%</b> )	1.10   0.12	n/a	0.03   0.15
	$^{12}\text{C}+\text{H}\rightarrow^{10}\text{B}$	7%	(2.5%)	1.07   0.15	n/a	0.00   0.15
$^3\text{He}$	$^3\text{He}+\text{H}$		(1.8%)	1.00   0.15	1.2   0.5	n/a
	$^4\text{He}+\text{H}$		( <b>5.0%</b> )	1.00   0.10	1.00   0.25	n/a
	$^{16}\text{O}+\text{H}\rightarrow^3\text{He}$	5%	(2.1%)	1.10   0.30	n/a	0.10   0.10
	$^{12}\text{C}+\text{H}\rightarrow^3\text{He}$	5%	(1.5%)	1.10   0.30	n/a	0.05   0.15
	$^4\text{He}+\text{H}\rightarrow^3\text{He}$	80%	( <b>7.3%</b> )	1.00   0.10	n/a	0.00   0.025
<i>N</i>	$^{16}\text{O}+\text{H}$		( <b>1.8%</b> )	1.03   0.04	0.7   0.5	n/a
	$^{15}\text{N}+\text{H}$		(1.0%)	1.00   0.05	0.7   0.5	n/a
	$^{14}\text{N}+\text{H}$		(1.6%)	1.02   0.07	0.7   0.5	n/a
	$^{16}\text{O}+\text{H}\rightarrow^{15}\text{N}$	50%	( <b>5.9%</b> )	0.90   0.15	n/a	0.05   0.10
	$^{16}\text{O}+\text{H}\rightarrow^{14}\text{N}$	40%	(1.7%)	1.00   0.15	n/a	0.00   0.05

- For each reaction, calculate the flux with the various XS data set. Draw the flux ratio residuals for modified and reference XS data set.
- Inspect the relative differences. Large deviations indicate, that XS discrepancies between the data sets have an impact on the flux and that a nuisance parameter is thus needed for this element and reaction. If the deviation for two reactions resemble each other, it is numerically sufficient to introduce nuisance parameters to only one of the reactions, as it will cope for both deviations together.

The results for B/C are taken from [Derome et al. \(2019\)](#). For the other species Li/C, Be/C, N/O and  $^3\text{He}/^4\text{He}$  this procedure is reapplied for this thesis. The reactions found to be worth studying for all the species are shown in table 2.1. The plots created in step 4 are shown in the appendix E for the reactions in table 2.1. The nuisance parameters are meant to improve the

fit of the respective secondary flux. Even though they were examined using a specific flux ratio, they are applied when fitting any flux ratio or direct flux involving the respective secondary species Li, Be, B, N, or  $^3\text{He}$ .

### 2.2.3 Determining the nuisance parameter prior values

Following [Derome et al. \(2019\)](#), for each inelastic reaction cross section a normalization and an energy-scale factor is introduced. For production cross sections a normalization and a slope for low energies below a fixed threshold of  $\mathcal{O}(1 \text{ GV})$  is used. In addition, each nuisance parameter  $\eta$  needs a Gaussian uncertainty  $\sigma_\eta$  in order to account for its deviations in the  $\chi^2_{\text{nuisance}}$ .

The associated values and uncertainties for each of the nuisance parameters are determined subjectively. This is done by selecting the respective  $\eta^*$  and  $\sigma_\eta^*$  by hand, such that the statistical distribution of the modified reference cross sections matches the other data sets within a  $1\sigma$  range. Table 2.1 shows the picked values. The plots used for this analysis are in appendix E. The values for B/C are kept the same as in [Derome et al. \(2019\)](#) while extending them onto the other species. In this thesis however, cross-section nuisance parameters were used for fitting, only for the most critical inelastic and production reaction, as indicated by the bold percentages in table 2.1.

## 2.3 Data uncertainties

### 2.3.1 Possible data type uncertainties in USINE

There are 3 different options on how data uncertainties can be taken into account when fitting CR models with USINE:

**Statistical errors** are directly taken from the data distribution itself, using the number of events in each rigidity bin of the initial data. They are bin-uncorrelated and describe the data uncertainties as if the experiment had no systematics.

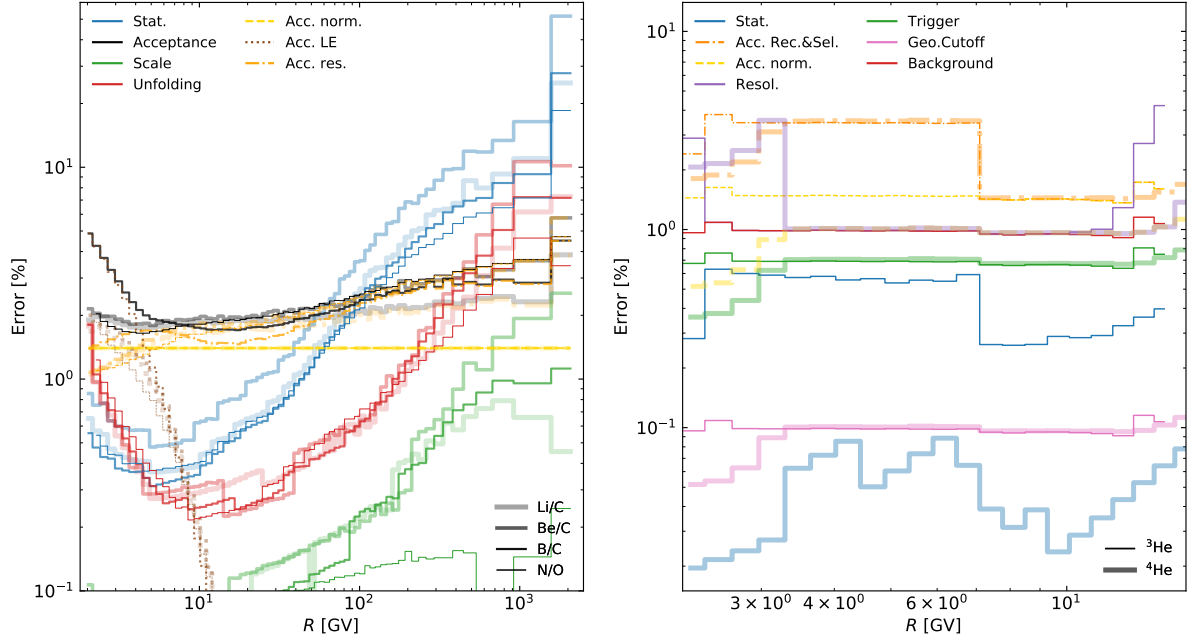
**Total errors** are the quadratic sum of statistic and systematic error estimations from the respective experimental setup. This is the commonly used error type in almost all CR analysis in the literature. But it does not take into account errors of the measured rigidity and thus treats energy bins as uncorrelated.

**Covariance matrices** introduce a correlation between different energy bins and are combination of both statistical and systematical error sources. Such a correlation can have multiple reasons, for example a general miscalculation of the energies, uncertainties in the unfolding of bins, or interdependence of re-normalization due to the acceptance.

### 2.3.2 Creation of covariance matrices

For the published AMS-02 data, no official covariance matrices are available. The (relative) covariance matrices for the analysis were thus all estimated by L. Derome, as described in [Derome et al. \(2019\)](#). For each type of error  $\alpha$  (e.g. statistic, unfolding, acceptance), each matrix element, describing the covariance between the rigidity-bin  $R_i$  and  $R_j$ , is calculated as

$$(C_{\text{rel}}^\alpha)_{ij} = \sigma_i^\alpha \sigma_j^\alpha \exp\left(-\frac{1}{2} \frac{(\log(R_i/R_j))^2}{(l_\rho^\alpha)^2}\right), \quad (2.8)$$



**Figure 2.2:** Relative uncertainties for the AMS-02 data, used to construct the covariance matrices. The black curves are the quadratic sum of all the *Acc.* (i.e. acceptance) error types for each respective species.

with the relative errors  $\sigma_i^\alpha, \sigma_j^\alpha$  for bin  $i, j$  respectively and error specific correlation length  $l_p^\alpha$ . This leads to one covariance matrix per species per error type  $\alpha$ . The values for  $l_p^\alpha$  and the resulting correlation for B/C are plotted in figure 2.3. For each error type  $\alpha$  the relative data uncertainties  $\sigma_i^\alpha$  are shown in the plots 2.2, which are the diagonal elements of their respective covariance matrix.

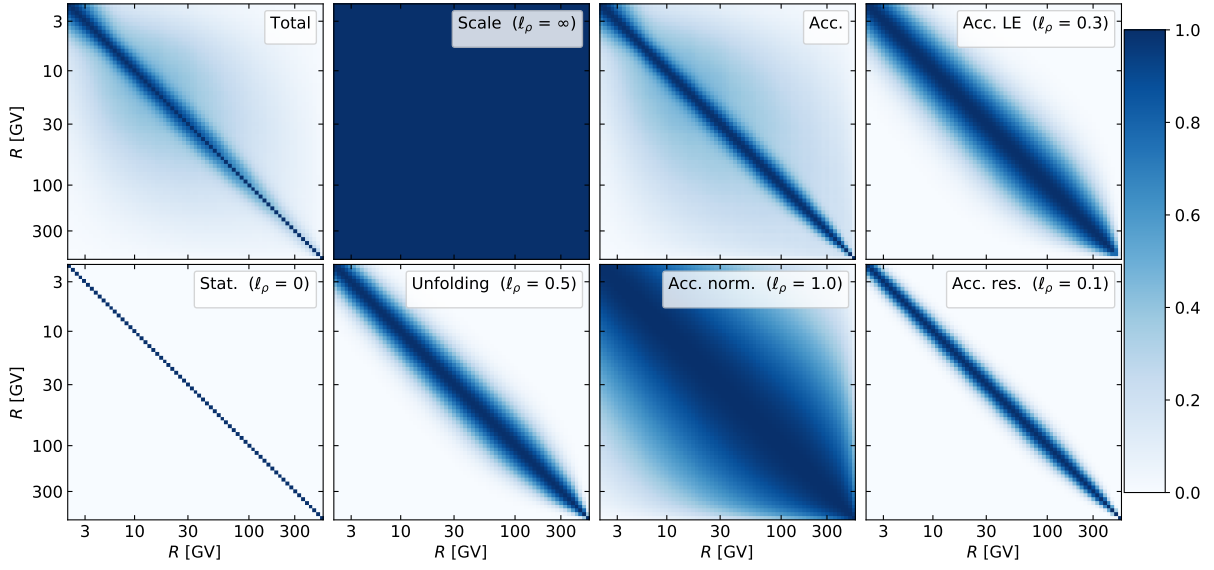
For the secondary-to-primary ratios involving Li, Be, B, and N, the correlation lengths are kept the same, using the values from Derome et al. (2019). The values used come from mostly phenomenological or technical assumptions. The correlation length  $l_q^{\text{acc.res.}}$  however is not very well defined experimentally and is difficult to characterize. For this reason, it was adjusted at that time in order to ensure  $\chi^2 = 1$  for B/C AMS-02 data. Reusing this value for Li, Be, and N, is for the sake of consistency. This can also be justified by their similarities in terms of production parents and data energy range. Note that thus the individual  $\chi^2$  values for e.g. Li/C or Be/C may differ from 1. This must not necessarily have a further meaning, hence the correlation lengths still remain an estimation.

Because of a different error distribution and energy range however, the acceptance correlation length is readjusted to  $l_q^{\text{Acc.Rec.\&Sel.}} = 0.05$  ( $\equiv l_q^{\text{acc.res.}}$ ) for all Helium covariance matrices. An overview of the impact of  $l_q^{\text{Acc.Rec.\&Sel.}}$  for Helium is shown in the appendix figure D.1.

### 2.3.3 Impact on parameter constraints

In a first preliminary test, all 3 error types are tested using the SLIM model with either no or the main cross section nuisance parameters according to table 2.1. The SLIM model is chosen because of its simplicity in this case.

The best-fit parameters are shown in figure 2.4. The results in the first two columns underline the fact, that statistical errors only are an underestimation of the data uncertainties and thus lead to an extremely high  $\chi^2/n_{\text{dof}}$ . The third and fourth column use the total errors and fare



**Figure 2.3:** Correlation from uncorrelated (white) to fully correlated (blue) for the relative covariance matrices for each error type. The subplot legends provide the specific correlation length used for the respective error types.

much better in terms of  $\chi^2/n_{\text{dof}}$ . The same applies to the last two columns using the covariance matrix. Looking at the first, third and fifth columns and comparing their results (especially those for  $\delta$  and  $K_0$ ) with the second, fourth and sixth respectively, the nuisance parameters do indeed add a flexibility to the fit and have a strong impact on the parameter determination. In order to get the unbiased results, realistic uncertainty estimations are needed, which favors the last of the six columns.

The goal of this first test was not to comment on the transport parameter results or their species universality, which is revisited in the next chapter. It is rather shown to justify the choice made for this thesis of using always a minimal set of nuisance parameters for any species involved and a data covariance matrix if possible. If no covariance matrix is available, as the case for non AMS-02 data, total errors are used.

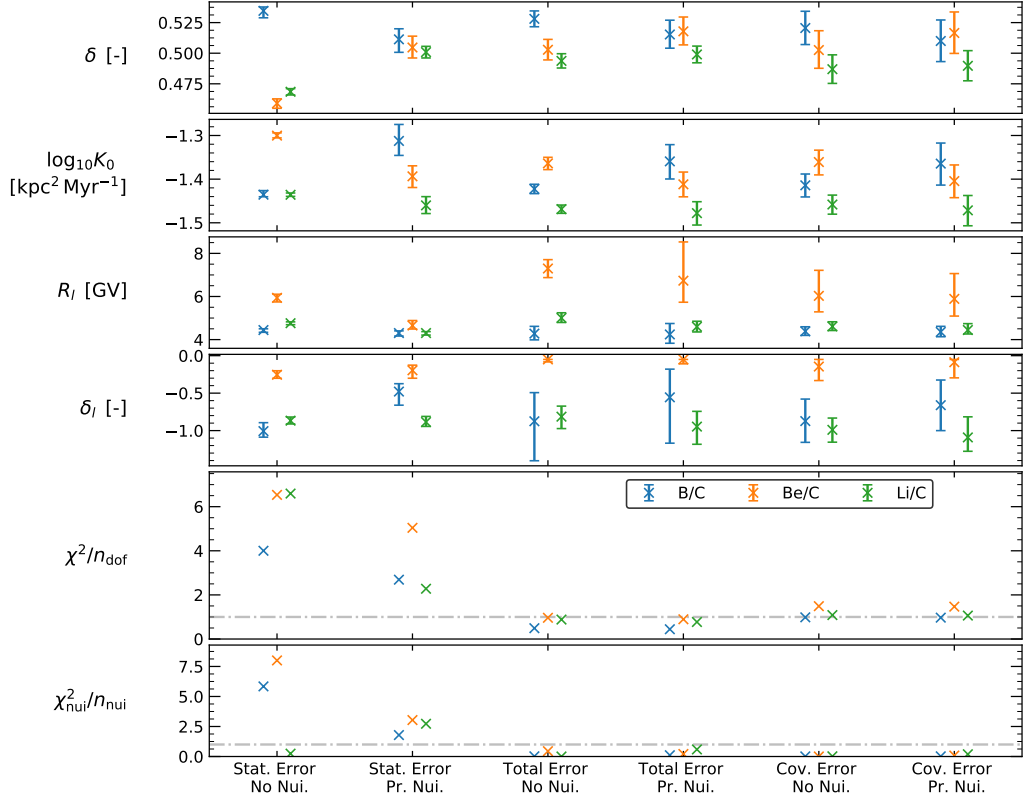
### 2.3.4 Inspecting residuals of correlated data with the $\tilde{z}$ -score

The so called  $z$ -score for rigidity bin  $R_i$ , which defined as

$$z_i = \frac{x_i}{\sigma_i} \quad \text{with } x_i = y_i^{\text{data}} - y_i^{\text{model}}, \quad (2.9)$$

is an important tool of investigation of model to data distances. In the correlated case however,  $\sigma_i = \sqrt{C_{ii}}$  does not reflect the full data error, as all the off-diagonal elements of the correlation matrix  $C$  are not taken into account. This is why the  $\tilde{z}$ -score was introduced in Boudaud et al. (2019), resolving this problem and giving a uncorrelated residual estimation. In order to scale  $\sigma$  correctly, the covariance matrix is rotated into a new base in which it is diagonal, using the orthogonal matrix  $U$ . This yields

$$\tilde{x}_i = U_{ij}x_j; \quad \tilde{C} = U C U^T; \quad \tilde{z}_i = \tilde{x}_i / \sqrt{\tilde{C}_{ii}} \quad (2.10)$$



**Figure 2.4:** Fit results for B/C, Be/C, and Li/C to AMS-02 data with the SLIM model. Each column corresponds to a different use of error handling and set of cross-section nuisance parameters. *No Nui* describes the absence of any cross section nuisance parameters whereas *Pr.Nui* means that the minimal cross section nuisance parameter set for the respective species according to table 2.1 is used. Note that  $\phi$  is an always involved nuisance parameter, which is why  $\chi^2_{\text{nui}}/n_{\text{nui}}$  is not 0 in the left-most column.

whereas  $\tilde{C}$  is a diagonal covariance matrix and  $\tilde{z}$ -score the rotated  $z$ -score. This re-simplifies

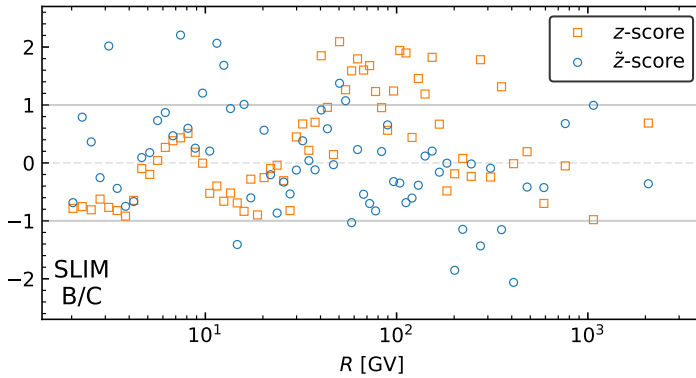
$$\chi^2 = \vec{x}^T C \vec{x} = \vec{x}^T \tilde{C} \vec{x} = \sum_i \tilde{z}_i^2. \quad (2.11)$$

The (pseudo) rigidity

$$\tilde{R}_i = \sum_j U_{ji}^2 R_j, \quad (2.12)$$

which does not differ a lot from the physical value  $R_i$ , can then be used to draw graphs of  $\tilde{z}$  over  $\tilde{R}$ , which show the correct distance between the model and the data, unlike  $z$  over  $R$  in the correlated case. An example of a direct comparison of  $z$ -score and  $\tilde{z}$ -score is shown in figure 2.5, demonstrating that the correlation between neighboring data-model discrepancies in the  $z$ -score is resolved for the  $\tilde{z}$ -score.

The advantages of  $\tilde{z}$  over  $z$  are, that differences between model and data can be classified as statistical fluctuation or feature more easily. Neighboring model-data-discrepancies due to data correlation are visible in the  $z$ -score but are filtered out in the  $\tilde{z}$ -score.



**Figure 2.5:** Comparison of the  $z$ -score in orange and the  $\tilde{z}$ -score in blue for B/C AMS-02 data respective to the SLIM model fitted in section 3.3.

## 2.4 Parameter space studies

Most of the work in this thesis relies on the numerical exploration of the parameter space in order to define best-fit values and boundaries of the model parameters. The state-of-the-art method is using MINUIT2 to minimize the  $\chi^2$  function and thus extract the parameter values and uncertainties. But as this is not the only possible technique, this section will show different options and try to address their usability. To make this easier, a Python port PYUSINE was developed, that allows to access USINEs  $\chi^2(\vec{\theta})$  function with any parameter choice  $\vec{\theta}$ . This way any Python code can be used to explore the  $\chi^2$  parameter space, leaving the CR model evaluation against the data within the USINE code untouched.

### 2.4.1 Minimizers

There are a lot of different ways to find the minimum of a scalar function in a (multidimensional) parameter space. In order to judge performance of the common MINUIT2, it is experimentally compared to different minimizer algorithms. The python package IMINUIT was used as reference instead of the implementation inside USINE. This is for the sake of consistency with the other python packages in terms of common conditions and means of monitoring their performance.

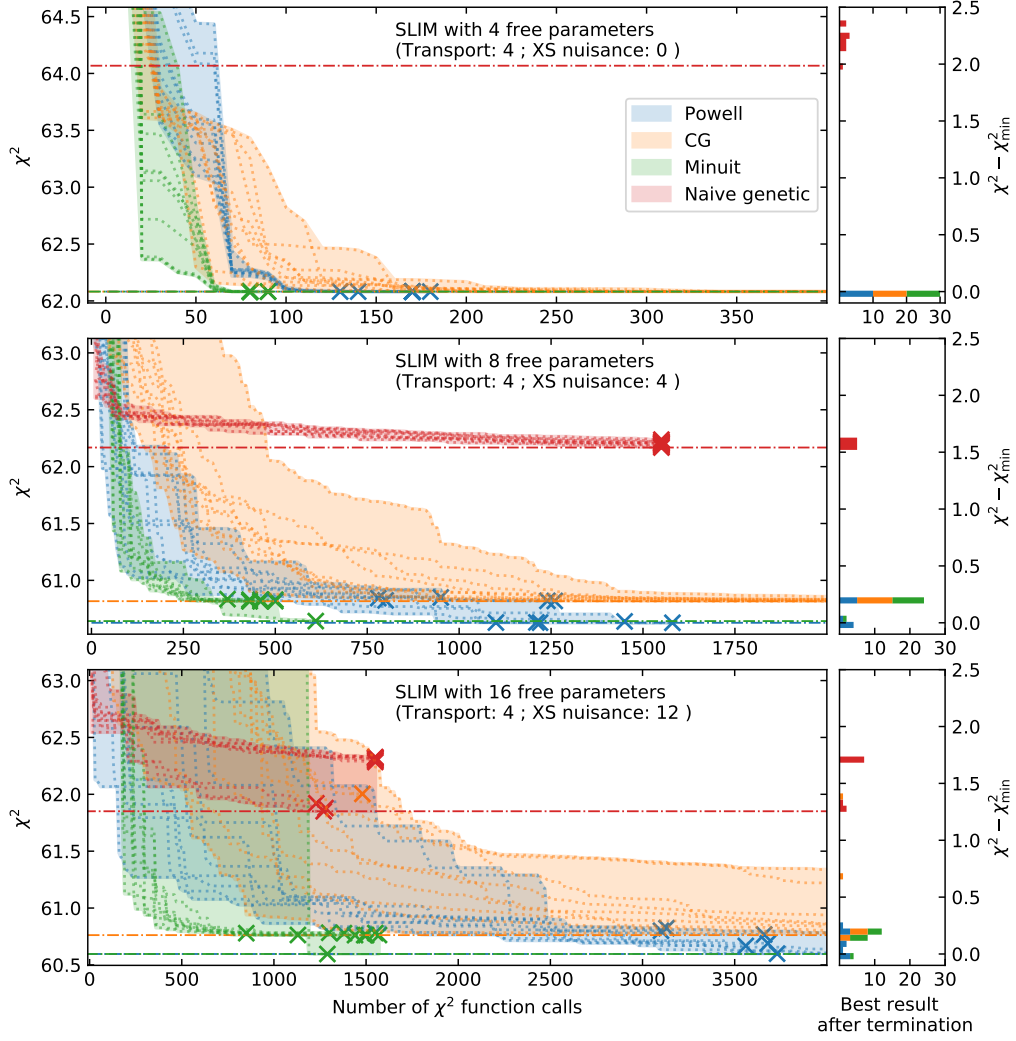
The other algorithms tested are the Conjugate Gradient method (implemented in the SciPy python package), Powell's method (implemented in the SciPy python package) and a naive genetic algorithm (self written). A short description of all of these can be found in Appendix B.1.

#### The test setup

The reference task for minimizer evaluation is fitting B/C AMS-02 data in the SLIM model. This is repeated with 3 different cases of nuisance parameters for cross sections: One without, one with only the main nuisance parameters and one case with many nuisance parameters (see table 2.1). For each case the process was repeated from 10 different initial points (same ones for all minimizers). All the other settings are kept identical. The behavior of the different methods over time are shown in figure 2.6.

#### The speed of convergence

One necessary condition for a minimizer is to converge at all. However, it may still be unusable if it takes too long to converge. The algorithms to minimize the  $\chi^2$  function originate from different packages, authors, and even programming languages (mostly pure python). But regardless of



**Figure 2.6:** This figure shows the performance of the Powell, Conjugate gradient(CG), Minuit and Naive Genetic fit methods. The dash-dotted horizontal lines in the left-hand pictures are the global  $\chi^2$ -minimums for each method. The degrees of freedom ( $n_{\text{dof}}$ ) are always 63, hence there are 67 data points and 4 free parameters involved, whereas the nuisance parameters leave it unchanged. Each dotted line corresponds to different starting points and the colored area illustrates the range of different curves for each method. In the top left figure the red curves are not visible because they are out of the image boundaries. The right-hand side pictures are histograms of the  $\chi^2$  values, at which the minimizers terminated. The left- and right-hand side pictures share the same y-axis. Powell, Conjugate Gradient and Migrad have their own termination criteria. For the Naive genetic however, the maximal number of 1600  $\chi^2$  calls was enforced.

this, their execution time is almost negligible compared to the cost of evaluating the  $\chi^2$  function in USINE. This is why the number of  $\chi^2(\vec{\theta})$  calls is an easy and reliable estimator instead of the machine-dependent execution time.

Figure 2.6 shows that iMinuit is the overall quickest method in any tested case. The Powell algorithm from SciPy is the second best one. The genetic, random-based approach fares worse than any other tested method. From a subjective point of view it is even unlikely that it would ever converge to the global minimum in a realistic amount of iterations. This is why this method



was studied no further than up to 1600 calls.

### True minimum reliability

As figure 2.6 shows, the minimum found by all of the 3 converging methods (iMinuit, Powell, and CG) is always the same in the absence of nuisance parameters for this model. It can thus be believed to be the true global minimum. This is no longer the case if the parameter space gets bigger and eventually more complicated with additional (nuisance) parameters. Note that *true minimum* here means the best  $\chi^2$  ever to be encountered by any of the 40 minimization runs for each parameter-setup (column) respectively.

It is shown in the right-hand side histograms that CG and iMinuit have sometimes terminated in a wrong  $\chi^2$  minimum for the second and third configuration. Even the Powell method has trouble finding the true minimum in the big parameter space, but it is in this setup more likely to really find it.

The mid-right and bottom-right histograms in 2.6 indicate that the best-fit parameters in our studies underlie a undeniable possibility of being incorrect when being determined with Minuit. Although the deviation of  $\sim 0.2$  seems to be quite small, it imposes a  $\sqrt{0.2}\sigma = 0.447\sigma$  bias on the best-fit parameters.

### Minimization strategy for this Thesis

The decision is to keep on using the MINUIT2 (Migrad) implementation in USINE for this thesis, as it also allows to use MINOS and is relatively fast compared to its competitors. However, each individual model fit configuration is repeated from  $\mathcal{O}(100)$  different, random starting points in order to cancel out the risk of not finding the true minimum because of a wrong starting point. Out of these several independent fit results for the same fit configuration one final result is then extracted. The heuristic procedure for the latter is explained in B.2.

#### 2.4.2 Monte Carlo methods

Whereas minimizers look for the single best fit parameter vector as the minimum of the  $\chi^2$  function in the parameter space, Monte Carlo (MC) methods follow a different approach: The posterior probability density function (PDF) for the true parameters is sampled in the multidimensional parameter space. This PDF  $p(\theta|\text{data})$  resembles the probability for each  $\theta$  to be the true parameter vector according to the data. An estimate for the best-fit parameter vector and corresponding errors can easily be obtained by calculating the mean, covariance and  $1\sigma$  quantiles from the sampled posterior distribution. In theory, the sampled values lie in a global minimum area of the  $\chi^2$  function. An efficient way to obtain a sample of parameter vectors according to the  $\chi^2$  distribution is the Markov Chain Monte Carlo (MCMC) method.

The parameter distribution obtained by MCMC may differ from a Gaussian form and resembles the true range of what the correct parameters may be. This is an important tool for understanding the parameters and especially their relations. MINUIT2 and MINOS may yield covariance matrices of fit parameters and asymmetric errors, but may come to their limits in some cases of unorthodox parameter space phenomena. In the best case however, MCMC will just reveal Gaussian distributions and elliptic quantile contours. This way, MCMC can be seen as a validation check for the MINUIT2 + MINOS approach and their respective results.

A description of the algorithms mentioned is given in Appendix B.3. For all the MCMC calculations the PyMC3 python package was used together with the USINE wrapper described in Appendix A. All techniques, mentioned in the following, are implemented in this library.

**Table 2.2:** Auto-correlation estimations of the Markov Chain for a preliminary test setup, using B/C AMS-02 data and the SLIM model. This test run does not reflect a real analysis case, because a gaussian prior  $\pi(\theta)$  was imposed for all parameters. The correlation lengths are the average from several different chains and points of departure, calculating formula (B.7) individually for each chain and parameter. The data is not statistical significant and can only give a general idea about the order of magnitude. The bold numbers are the highest transport parameter correlation lengths for each configuration (column), which would represent the overall correlation length as used to thin down the chains. The (estimated) average number of  $\chi^2$  calls per true step, is given in the last line. It is the total auto-correlation length for the Metropolis samplers. In this experiment, the points between two steps for H-MC are uncorrelated, which is why the number of  $\chi^2$  calls is simply the the number of calls needed between two steps. The latter is setup dependent, deriving from the number of sub-steps and the number of free parameters, all being fixed in this test run.

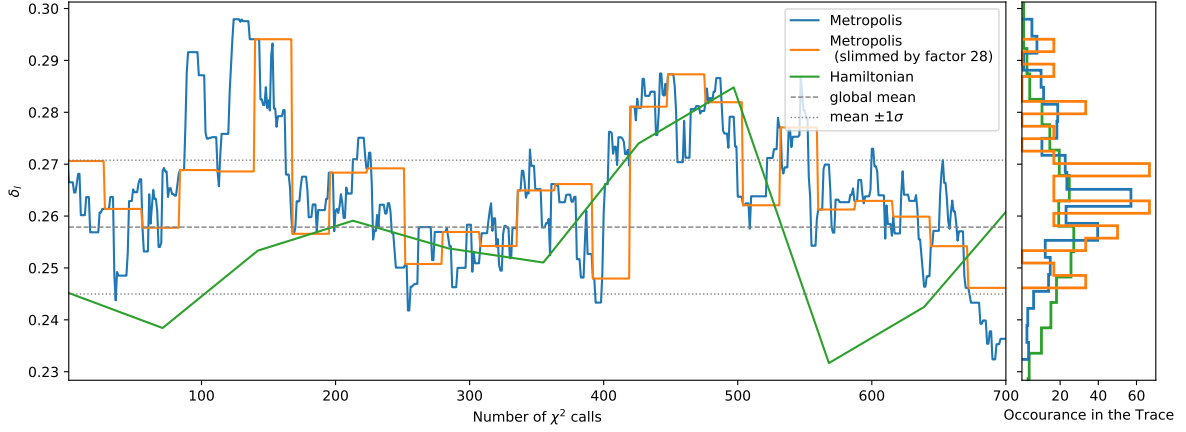
	parameters	Hamiltonian		Metropolis	
		no nui.	w/ nui.	no nui.	w/ nui.
Transport	$\delta$	1.0	1.2	8.2	21.2
	$\log_{10} K_0$	1.0	1.0	20.2	29.8
	$R_l$	1.0	1.0	25.3	<b>32.6</b>
	$\delta_l$	1.0	1.0	<b>28.5</b>	30.7
Nuisance	Norm $\sigma(^{11}\text{B}+\text{H})$		1.0		23.1
	EAxisScale $\sigma(^{11}\text{B}+\text{H})$		31.0		23.3
	Norm $\sigma(^{12}\text{C}+\text{H} \rightarrow ^{11}\text{B})$		3.8		27.9
	SlopeLE $\sigma(^{12}\text{C}+\text{H} \rightarrow ^{11}\text{B})$		2.5		15.5
Av. No. calls / un-corr. step		71	242	28	32

### Hamiltonian vs. Metropolis Hastings MCMC

As described in the appendix B.3, there are two possible sampling techniques for the MCMC taken into account. Due to the different algorithm concepts, both have different expenditures and auto-correlations in practice. In order to understand the amount of calculations needed and decide what sampler to use, a preliminary test is done with both Hamiltonian (H-MC) and Metropolis Hastings, using either no or minimal nuisance parameters and a SLIM model. The results for the correlation lengths are given in the table 2.2. It shows not only that auto-correlation is a huge problem, but also that this scales with the complexity (free parameters) of the model.

An illustration example of the different chains is given figure 2.7. It shows, that thinning out the chain by the correlation estimation can resolve the problem to some extent. The chains eventually reproduce a parameter distribution, as shown in the right-hand histogram. However, long chains with a lot of elements, and even more  $\chi^2$  calls, are needed, which makes this a costly procedure.

At this point, it is still difficult to prove the superiority of one MCMC sampling method over the other. The choice for Metropolis Hastings (MH) for this thesis is thus quite subjective. One of the reasons is, that it needs fewer settings to be tuned. H-MC requires much more tweaking, like defining a step and sub-step length. This is why MH MCMC is easier to understand and to apply in new conditions.



**Figure 2.7:** Illustration of the Markov Chain for the preliminary test without nuisance parameters, further explained in table 2.2. This chart shows the chain element values for  $\delta_l$  over the number of  $\chi^2$  calls, that were needed to calculate it. Note that Hamiltonian-MC uses a lot of sub-step calculations, whereas Metropolis-Hastings yields a priori one chain element per  $\chi^2$  call. The second, orange curve is obtained by picking only each 28<sup>th</sup> element from the chain in order to remove the auto-correlation (see appendix B.3.1 and table 2.2). The gray dotted lines show the symmetric standard deviation of the posterior distribution of the Hamiltonian-MC Chain, as shown in the right-hand side histogram. The Hamiltonian-MC trace does not end at 700 calls.

### Application of MCMC in CR studies

The role of MCMC is not yet to replace minimizers, for which it is not fast and stable enough. It is rather used as a result cross check for important studies by redoing them with MCMC and comparing the results with the MINUIT2+MINOS standard approach. This will be done for two important fit configurations in sections 3.5 and sections 4.5. In order to reduce the burn-in time (see Putze et al. (2009) for details), the starting points and sampling covariance matrices are drawn from the MINUIT2 best-fit.

Because of the large computation time needed,  $\mathcal{O}(100)$  parallel Markov-Chains are sampled. Each chain is then thinned down by its respective auto-correlation length. The sum of all the trimmed chains is then used for the actual parameter evaluation.

From the final parameter sample, the two dimensional  $1\sigma$  contour are drawn as the limit in which 68% of the points lie. The parameter value  $\theta^\alpha$  for parameter  $\alpha$  is the mean of the sample. Additionally, the lower and upper  $1\sigma$  quantiles  $\theta_{l,h}^\alpha$  define the respective errors  $\sigma_{l,h}^\alpha = \theta_{l,h}^\alpha - \theta^\alpha$ .



## Universality of CR transport for Lithium, Beryllium, Boron, Nitrogen, and Helium

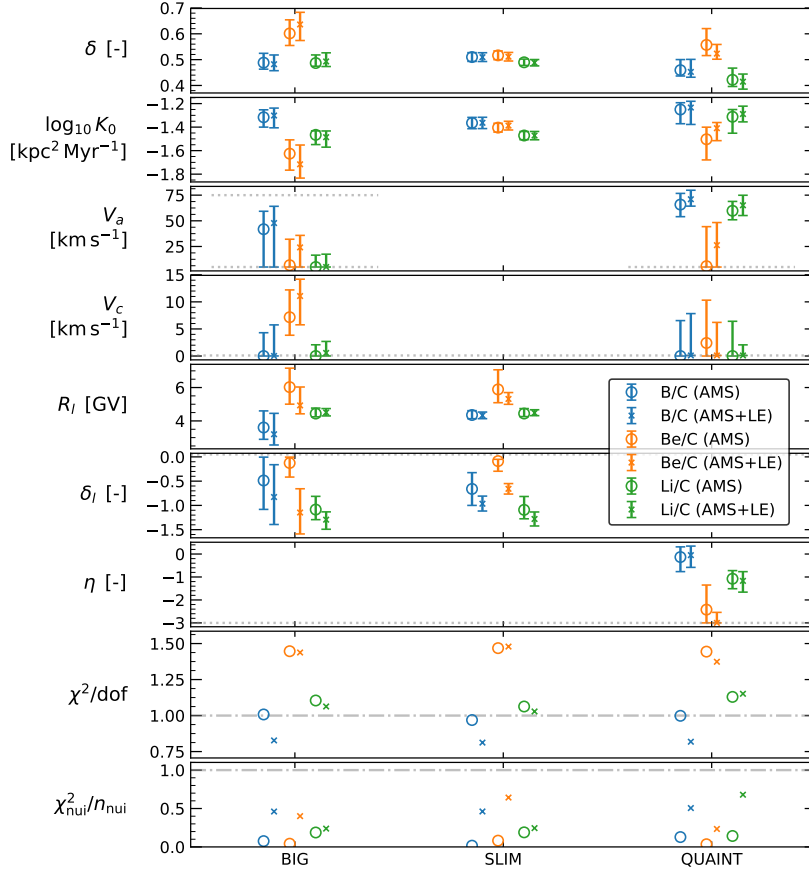
Combining the theory from Chapter 1 and the techniques from Chapter 2, the CR transport parameters for the presented models can now be examined with recent AMS-02 data. This Chapter gives a follow up to recent studies of the AMS-02 B/C data (Génolini et al. 2019) (Derome et al. 2019) extending their work to Lithium and Beryllium. The motivation of these studies is further given in section 3.1. In order to avoid source dependencies and focus on transport mechanisms only, flux ratios are used for the most part instead of direct flux fitting, motivated in section 1.3.2 and successfully tested before. The studies and results presented in this chapter are also submitted in a paper (Weinrich et al. 2020).

The first thing studied in this chapter is the consistency of AMS-02 Li/C, Be/C and B/C data with data from low-energy experiments in section 3.2. In 3.3, the main analysis of finding transport parameters for the CR models is carried out using Li, Be, and B and their combination. As a validation test, the analysis is expanded onto Helium and Nitrogen in 3.4. The configuration using Li/C, Be/C, and B/C in the SLIM model is then redone with MH-MCMC instead of MINUIT2 in section 3.5 as a cross-check and proof-of-concept.

### 3.1 Current CR transport studies with AMS-02 high-precision data

The precision of the recent AMS-02 flux data of a few percent over a wide energy range has been a step-up for CR research, allowing better studies of its spectrum and spectral features. This is further improved with the use of uncertainty covariance matrices for the AMS-02 data (Derome et al. 2019) as explained in section 2.3.2, that introduce correlations between the data points, further distinguishing fluctuations from real spectral phenomenons. In this new conditions, the B/C flux ratio (Aguilar et al. 2016) data has lead to a robust prediction of the diffusion slope  $\delta$  and the coefficient  $K_0$  in the 1D models used in this thesis (Génolini et al. 2019). In reverse, this AMS-02 data could eventually be reproduced with these models in a very good agreement. It was further possible to constrain a high rigidity break at  $R \sim 250$  GV in the diffusion spectrum with strong hints for a second low rigidity break at  $R \sim 5$  GV (Génolini et al. 2019). With the same data, the need for a high-rigidity spectral break at  $\sim 300$  GV has also been found by Evoli et al. (2019), using a similar 1D flat-disc model.

The recent publications of AMS-02 secondary-to-primary flux ratio data for Lithium and Beryllium (Aguilar et al. 2018a) give the opportunity to extend these studies and reinforce or probe the results for B/C (Génolini et al. 2019). This is important in the context of the debated universality of CR transport for light and intermediate species, testing it in the charge regime of  $Z = 3, 4, 5$ . Reproducing the precisely known AMS-02 data for all three species with the same, previously successful, and basic 1D models is the goal behind this. The motivation comes from the fact that all of these three species are considered to be pure secondaries and that their respective AMS-02 data covers the same rigidity range. They can thus be expected to underlie similar mechanisms of CR transport and lead to a better model constraint via combined analysis.

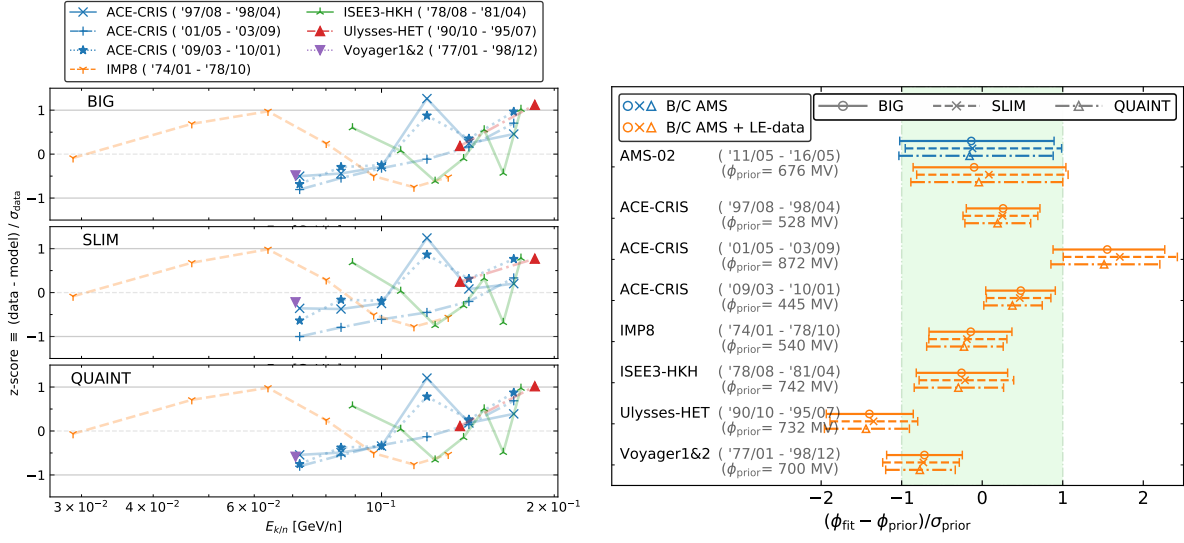


**Figure 3.1:** Fit results for B/C, Be/C and Li/C from either only AMS-02 or AMS-02 and several low-energy experiments combined. Each of the major three columns correspond to one of the tested models BIG, SLIM, and QUAINT. The experiments corresponding to B/C LE data are given in the legend of figure 3.2. Be/C LE data comes from ACE-CRIS ('98/01-'99/01), IMP8 ('74/01-'78/10), Ulysses-HET ('90/10-'95/07), and Voyager1&2 ('77/01-'98/12), whereas Li/C LE data comes from ACE-CRIS ('98/01-'99/01), IMP8 ('74/01-'78/10), and Voyager1&2 ('77/01-'98/12). The gray dotted lines are the parameter bounds for the minimizer. The dashed-dotted lines for the two lower panels are  $\chi^2$  related guidelines explained in section 2.1.2.

AMS-02 data for  $^3\text{He}/^4\text{He}$ ,  $^3\text{He}$ ,  $^4\text{He}$ , and N/O have recently been published as well (Aguilar et al. 2018b, 2019), leaving the opportunity of even further analysis, as both N (to some extent) and  $^3\text{He}$  are secondary species as well. However, these data are only used as a cross-check in this thesis, because they introduce additional difficulties, which is further elaborated in section 3.4.

### 3.2 AMS data and other low-energy experiments

As mentioned in the previous section, the recent AMS-02 data provide precision of few percent over a wide rigidity range. But there are other experiments that measured the secondary to primary CR flux ratios in the past, mostly having a lower energy range than AMS-02 and less precision. This is why the first thing to do is an agreement cross-check for transport parameter results for AMS-02 data with those from previous experiments in the low-energy regime, later referenced as *LE data* in this thesis. Focusing on B/C, Be/C, and Li/C, the fits to these secondary to primary species are repeated for the two cases of AMS-02 data only and a combination of



**Figure 3.2:** Left panels: Z-scores for the B/C LE data relative to the models BIG, SLIM, and QUAINT, fitted to AMS-02 and LE B/C data. Data points from the same experiment are connected with lines for visibility. The dates in brackets behind the experiments name is the data taking period in the format 'YY/MM.

Right panel: Fit results for the solar modulation parameter for each experiment relative to their priors (see section 1.3.2). The gray brackets next to the experiment names show the data taking period and the respective parameter prior, whereas  $\sigma_{prior} = 100$  MV is the same for all the experiments. The green band shows the  $1\sigma$  range.

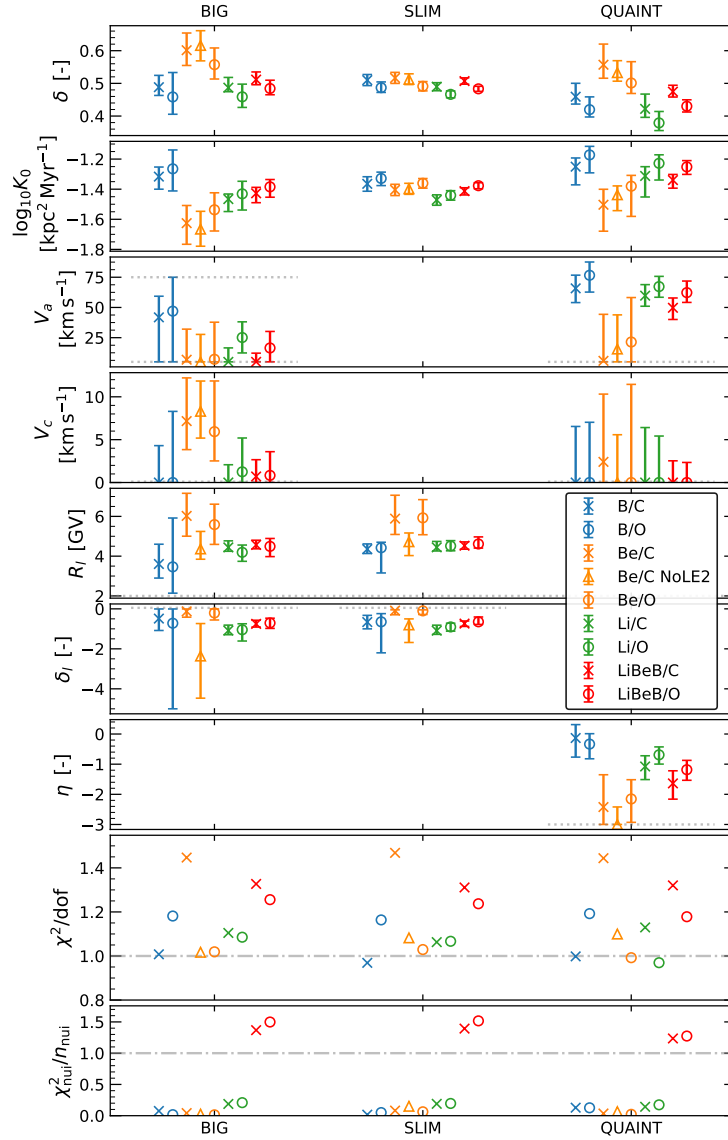
AMS-02 and LE data for each respective ratio. A fit to LE data only would not make much sense, since the models are optimized for intermediate and high rigidity regimes. The data included in the LE set for Be/C and Li/C is explained in the caption of figures 3.1 and for B/C in the figures 3.2.

Figure 3.1 shows the respective parameter fit results. For most of the parameters, models and species, the difference when adding LE data to the fit is minor and within their respective uncertainty estimation. The most important transport parameters  $K_0$  and  $\delta$  remain mostly unaffected for BIG and SLIM. However, the low energy parameters  $V_a$ ,  $V_c$ ,  $R_l$ ,  $\delta_l$ , and  $\eta$  are mostly constraint to different values when adding LE data. The latter is not surprising, given that they lie in a rigidity region which is highly sensitive to those parameters. A comparison of the diffusion coefficient  $K(R)$  for several different fits (including results from the next section), is given in figure D.3.

The figures 3.2 show the behavior of the LE solar modulation nuisance parameter and the z-score of the LE data in details. Note that this plot could have been made for Li/C, Be/C as well, whereas B/C is chosen as an example due to larger amount of LE data. It simply illustrates that fitting this LE data in addition to AMS-02 data works without notable model-data discrepancies.

Given the image that LE data does not change the outcome dramatically, it is proceeded without it for this chapter, focusing in the following on AMS-02 data only. This is for the sake of AMS-02 only consistency. But as mean of better constraining the low rigidity regime, the LE data for B/C and Li/C is picked up later in chapter 4.

### 3.3 Compatibility of Lithium, Beryllium, and Boron



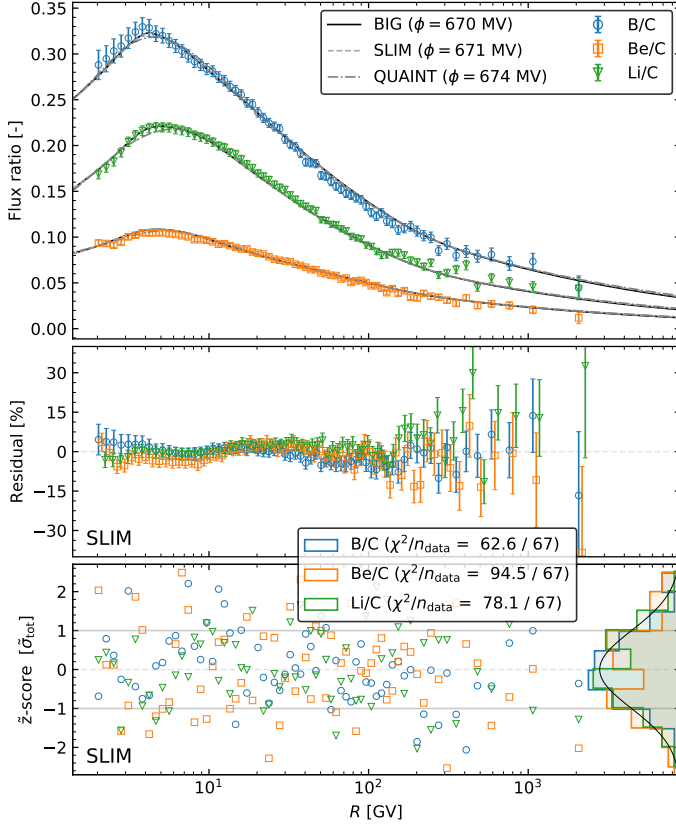
**Figure 3.3:** Fit results for several AMS-02 secondary over primary flux ratios, alone or in combination. The orange triangle for 'Be/C NoLE2' corresponds to the published Be/C AMS-02 data (Aguilar et al. 2018a) without its two lowest rigidity data points (see text for discussion). LiBeB/C or LiBeB/O correspond to a combination of Lithium, Beryllium and Boron, over Carbon or Oxygen flux ratio combination respectively. The gray dotted lines are the parameter bounds for the minimizer. The dashed-dotted lines for the two lower panels are  $\chi^2$  related guidelines explained in section 2.1.2

There are three different questions to be addressed in this section. First, how similar the parameter results for the three secondary species Li, Be, and B are. Second, if they can be combined, performing a simultaneous fit to get better parameter constraint. And third, if the use of Carbon or Oxygen as a primary flux is decisive.

To do so, the flux ratios Li/C, Be/C, and B/C were fitted to their respective AMS-02 data individually and simultaneously as well. The same is then repeated, but with Oxygen instead of Carbon as the respective primary species. The transport parameter fit results are shown in figure 3.3.

In this figure, the  $\chi^2/\text{dof}$  of Be/C is surprisingly and inconsistently higher than the one from Be/O. A closer look at the post-fit model and data plot 3.4 for Be/C reveals, that the first two AMS-02 data points with the lowest rigidity do not align with the rest of the data and model from a subjective point of view. This is why the Be/C fit is redone without these respective two data points as a preliminary test. Ignoring the first two data points for Be/C improves the





**Figure 3.4:** *Top panel:* AMS-02 data for Li/C, B/C, and Be/C in color. The lines are the model predictions for the three species for the models BIG, SLIM, and QUAINT after fitting them to all three species simultaneously.

*Middle panel:* Residuals of the data points in the top panel to their respective SLIM model prediction in the top panel.

*Bottom panel points:*  $\tilde{z}$ -score for the data, further explained in section 2.3.4.

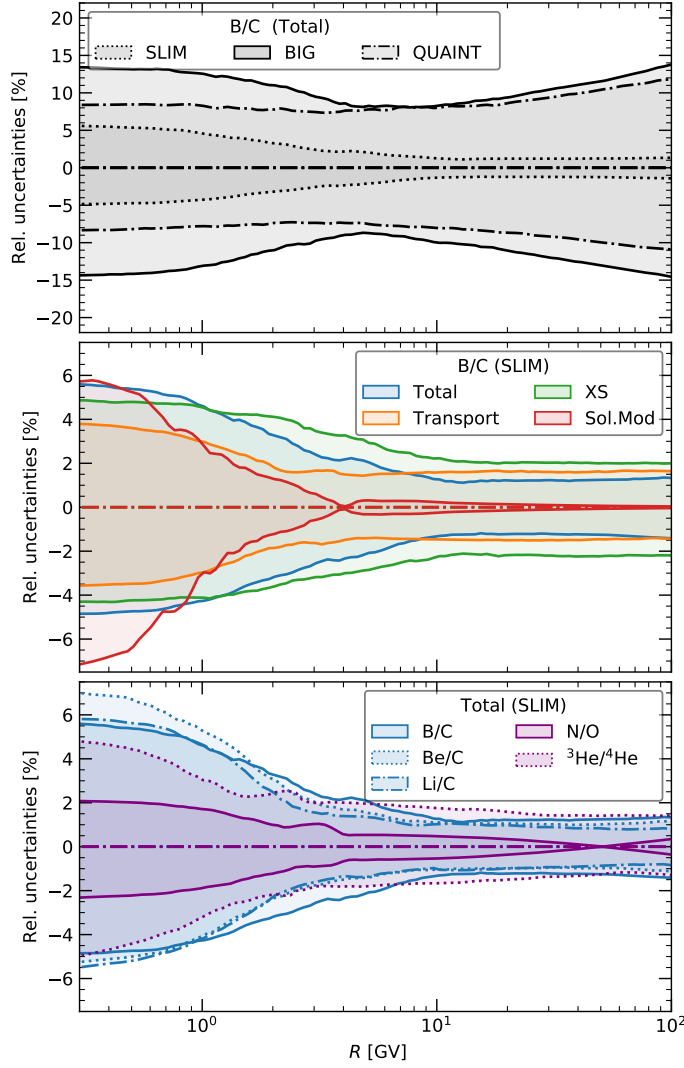
*Bottom right histograms:* histogram of the  $\tilde{z}$ -score values shown in this panel for all three species in their respective color. For reference, the normal distribution is drawn as a black solid line.

by far  $\chi^2$  and changes the fit results for low-energy parameters like  $R_l$  and  $\delta_l$ . Whether this is a statistical fluctuation or an unknown spectral feature of Beryllium shall not be further discussed at this point. For the sake of consistency, this is just left as a side-note and proceeded with the complete AMS-02 Be/C data.

Looking at the figure 3.3 one can now give an answer to the three initial questions of this section. The three species do lead to different transport parameter results, but do not differ far from their respective uncertainty ranges. It would be unjustified to claim an incompatibility. When comparing the red points, which show the simultaneous fit results to the others, one can see a slight improvement over fitting species individually. The combination fit results lie in between the ones from separate fits and have uncertainties of the same size or smaller. The comparison of Carbon and Oxygen as primary species denominator shows, that they generally lead to different results for all species and combinations, being however within their respective tolerances. Both of them are equally usable, especially when combining Li, Be, and B secondary data.

It is thus chosen to take the results from the fits using Li/C, Be/C, and B/C (a.k.a. LiBeB/C) as the reference results, staying in line with the previous B/C work (Génolini et al. 2019). These LiBeB/C parameter results for BIG, SLIM, and QUAINT are presented in table 3.1 and are from now on the most reliable and universal estimate for our model. As figure 3.4 demonstrates, all three ratios can be reproduced with a common model parameters respectively. This is so to say an update to the benchmark models from Génolini et al. (2019). Figure 3.5 gives an overview of the model uncertainties for LiBeB/C.

Comparing B/C to LiBeB/C, this is an improvement in terms of parameter result precision in all three models. On the downside, the  $\chi^2$  is worse. One of the reasons for this are the higher  $\chi^2$  values for the species Be/C and Li/C by default, as they use the correlation lengths for their



**Figure 3.5:** Model uncertainties dependent on the post-fit parameter uncertainties. The bands illustrate the  $1\sigma$  model quantiles for 1000 parameters samples using the post-fit covariance matrix estimation. *Top panel:* Comparison on the three models BIG, SLIM, and QUAINT when varying all the free parameters. *Middle panel:* Using the post-fit cross-section uncertainties (due to nuisance parameters), the green band is sampled. Varying the solar modulation according to its post-fit results yields the red band. The other free transport parameters yield the orange band. The blue band comes from varying all of them together. *Bottom panel:* Bands for SLIM when varying of all parameters for different species. Note that the N/O and  $^3\text{He}/^4\text{He}$  in purple where not used for fitting the parameters. The overall Nitrogen flux is normalized during the model estimations, which is why its variation is zero at  $R \sim 7$  GV.

respective covariance matrices that were optimized and normalized for B/C (Derome et al. 2019). As mentioned in section 2.3.2, this does not necessarily mean a systematic discrepancy between model and data. Reusing the same correlation lengths is rather for the sake of consistency between Li, Be, and B. Another reason for a bigger  $\chi^2$  is the exploitation of the freedom for the cross section nuisance parameters, as indicated in the last panel of 3.3, in order to achieve a compromise between the species.

Although LiBeB/O, i.e. Li/O, Be/O, and B/O, has shown to be possible to fit, Carbon is chosen over Oxygen as parent, as it is more abundant and thus has less statistical error. Using LiBeB/C + LiBeB/O is not suspected to give any improvement since this would not contribute new secondary data, but rather introduce an additional dependency to C/O. From a technical point of view, the combinations Li/C, Be/C, B/C, Li/O, Be/O, and B/O carry the same information as Li/C, Be/C, B/C, and C/O, as all flux ratios rely on the same AMS-02 measurements of Li, Be, B, C, and O.

Parameter	[ Unit ]	BIG	SLIM	QUAINT
$\delta$	[ - ]	$0.51^{+0.02}_{-0.02}$	$0.51^{+0.01}_{-0.01}$	$0.47^{+0.02}_{-0.02}$
$\log_{10} K_0$	[kpc <sup>2</sup> Myr <sup>-1</sup> ]	$-1.43^{+0.04}_{-0.06}$	$-1.41^{+0.02}_{-0.03}$	$-1.34^{+0.04}_{-0.05}$
$V_a$	[km s <sup>-1</sup> ]	$0^{+12.2}$	n/a	$49.7^{+8.1}_{-9.7}$
$V_c$	[km s <sup>-1</sup> ]	$0.7^{+2.0}_{-0.7}$	n/a	$0.0^{+2.5}$
$R_l$	[ GV ]	$4.57^{+0.23}_{-0.20}$	$4.53^{+0.19}_{-0.18}$	n/a
$\delta_l$	[ - ]	$-0.75^{+0.18}_{-0.16}$	$-0.74^{+0.13}_{-0.14}$	n/a
$\eta$	[ - ]	n/a	n/a	$-1.64^{+0.42}_{-0.52}$
$\chi^2_{\min} / n_{\text{dof}}$		258.8 / 195	258.3 / 197	258.8 / 196
$\chi^2_{\text{nui}} / n_{\text{nui}}$		21.9 / 16	22.3 / 16	19.8 / 16

**Table 3.1:** Free parameter results for the fit of the BIG, SLIM, and QUAINT models to Li/C, Be/C, and B/C (aka. LiBeB/C) simultaneously. For  $\log_{10} K_0$ , which is technically unit-less, the unit for  $K_0$  is given.

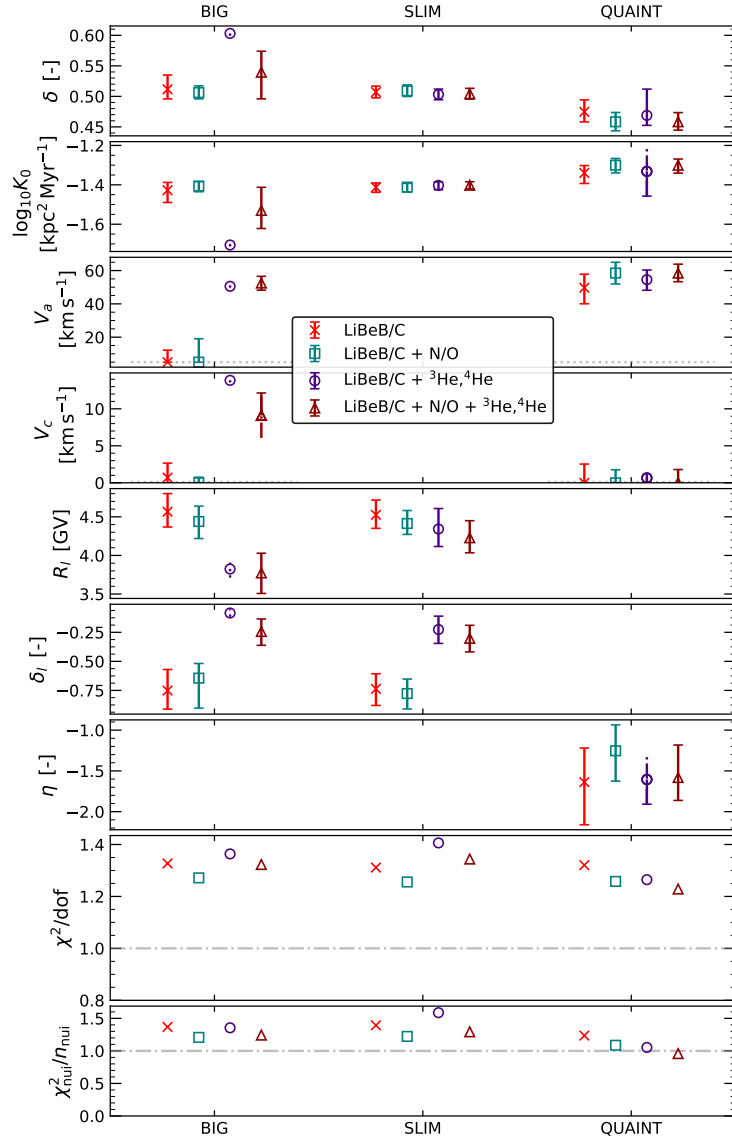
### 3.4 Helium and Nitrogen

The promising studies of comparing and combining different secondary species are now tried to be taken further, including the flux ratio N/O (Aguilar et al. 2018b) and the fluxes  $^3\text{He}$ ,  $^4\text{He}$  (Aguilar et al. 2019) into the analysis. As explained later, Helium and Nitrogen underlie different production mechanisms than Li, Be, or B and are thus only a validation cross-check. For this, the fluxes  $^3\text{He}$  and  $^4\text{He}$  or N/O are fitted, combining them either individually or together with the previously tested Li/C, Be/C, B/C combination. Figure 3.6 shows the fit result overview. Individually fits of N/O or  $^3\text{He}$  and  $^4\text{He}$  only to AMS-02 data are not further tried, as they are not sufficient to constrain the model parameters.

Nitrogen is a mixed species, meaning that it is not purely secondary but includes an abundance proportion from direct sources. Fitting its flux ratio thus re-introduces a dependency to source parameters unlike Li, Be and B. As it is heavier than Carbon, it also does not share the latter as a parent species, setting it apart from Li, Be, and B.

Around 80% of  $^3\text{He}$  comes from  $^4\text{He}$  (Coste et al. 2012), making it influenced by different source uncertainties than Li, Be, and B, who mainly rely on C and O. With  $A/Z = 1.5$ ,  $^3\text{He}$  it is also slightly different to Li, Be, B with  $A/Z \in [1.75 - 2.5]$ . Although a flux ratio  $^3\text{He}/^4\text{He}$  has been published (Aguilar et al. 2019) as well, it is chosen to rather fit the fluxes directly. According to formula (1.19), the conversion between these two energy scales would rely on very distinct values 1.5 and 2 for  $^3\text{He}$  and  $^4\text{He}$  respectively. The production cross sections in this model follow the straight-ahead approach, i.e. the  $E_{k/n}$  for the parent is the same as the produced secondary particle. The solar modulation and diffusion mechanism however is described in the rigidity scale. Moreover the source spectrum of  $^4\text{He}$  is also described in rigidity. With the different  $R \leftrightarrow E_{k/n}$  conversion for the two helium isotopes, the flux ratio  $^3\text{He}/^4\text{He}$  does not eliminate source dependency in any of the energy scales. With a unavoidable dependency on source parameters, a direct flux is fitted instead of a flux ratio.

The necessity for a source spectrum for  $^4\text{He}$  different than for heavier particles has also recently been confirmed by (Evoli et al. 2019). Thus the source spectrum (1.3) needs to be used in order to enable a more detailed description and freedom. This results in the source spectrum to taken as a broken power-law. Additionally to the three new parameters from (1.3),  $q_{^4\text{He}}$  and  $\alpha_{^4\text{He}}$  are added as free parameters as well, whereas the latter is independent from the fixed  $\alpha$  for the other species. The post-fit source spectrum for  $^4\text{He}$ , stemming from a combined fit carried out in this

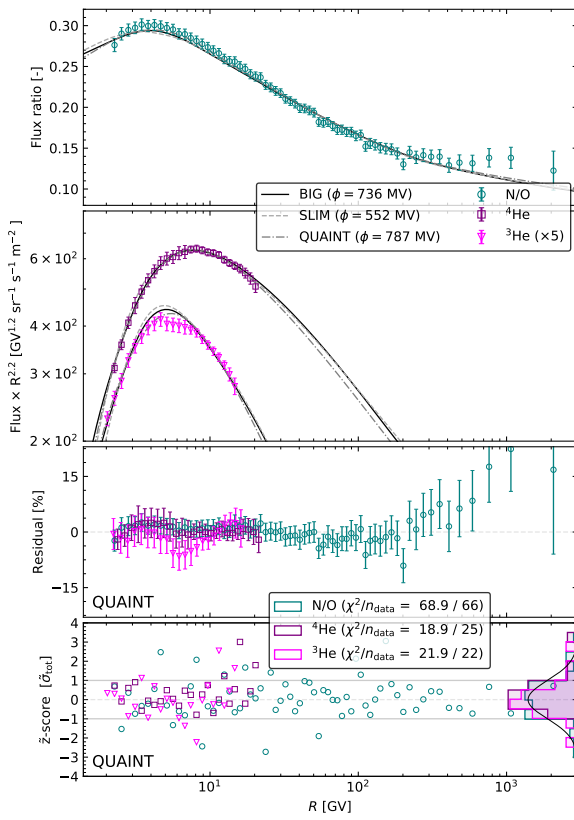


**Figure 3.6:** Fit results for several AMS-02 secondary over primary flux ratio combinations. The gray dotted lines are the parameter bounds for the minimizer. Error bars with caps correspond to an error estimation from MINOS whereas dotted error bars are an insecure estimation based on either a mirrored MINOS estimation or the migrad covariance matrix. LiBeB/C is short for Li/C, Be/C, B/C combined.

section, is shown in figure D.2.

Just like Li, Be, and B, the fitted fluxes N/O,  $^3\text{He}$ , and  $^4\text{He}$  could be predicted by the models, shown by the post-fit flux and data comparison in figure 3.7. Looking at the top left panel, all three models BIG, SLIM, and QUAINT seem to reproduce N/O equally well, with a good  $\chi^2/n_{\text{dof}}$  for QUAINT. The data of  $^4\text{He}$  is also in very good agreement with the three models. This may be the result of having three additional free sources parameters dedicated to this flux only and an individually adjusted covariance matrix. However, models start to differ for  $^3\text{He}$ , which can be observed in the second top panel on the left side in 3.7. QUAINT still aligns with the data, whereas BIG and SLIM have a clear offset at  $R \sim 5$  GV.

The results for the transport parameters are summarized in figure 3.6. When comparing the LiBeB/C benchmark and the addition of N/O, all three models show a better  $\chi^2/n_{\text{dof}}$  and better constrained transport parameters, being consistent with the LiBeB/C results. This shows, that the idea of transport universality and thus one model to describe multiple species ratios could be extended to Nitrogen.



**Figure 3.7:** Model predictions for BIG, SLIM, and QUAINT after their respective fit to Li/C, Be/C, B/C, N/O,  $^3\text{He}$ , and  $^4\text{He}$  AMS-02 data (see text for details) together with N/O,  $^3\text{He}$ ,  $^4\text{He}$ , and  $^3\text{He}/^4\text{He}$  data. *Left panels:* Flux ratio or fluxes, residuals to QUAINT, and  $\tilde{z}$ -score for QUAINT for N/O,  $^3\text{He}$ , and  $^4\text{He}$  AMS-02 data. See figure 2.3.4 for more details.

*Right top panel:* model prediction for  $^3\text{He}/^4\text{He}$  for BIG, SLIM, and QUAINT and respective data from the PAMELA and AMS-02 experiment. Purple lines use the post-fit solar modulation given in the bottom right legend, whereas green lines use the pre-fit PAMELA modulation given in the top legend. Note that PAMELA contains two data sets (based on TOF or CAL) that are shown alongside each other.

The consistency for LiBeB/C transport parameters and those from fits including the helium fluxes is different in each model. QUAINT is the model with the best agreement between the benchmark results from the previous section and the results including Helium. The global  $\chi^2/n_{\text{dof}}$  drops down, and even the parameters are constrained better when using both Nitrogen and Helium data. This is why this model is chosen to be used for residuals and  $\tilde{z}$ -score plots in 3.7. For SLIM, the inclusion of Helium shifts the  $\delta_l$  result and worsens the  $\chi^2/n_{\text{dof}}$  a lot. In the BIG model the two settings with and without Helium have interesting differences. Like SLIM, the low-rigidity parameters do change a lot, being constraint to values outside of the respective LiBeB/C parameter uncertainty ranges. The  $\chi^2/n_{\text{dof}}$  also indicates difficulties of fitting  $^3\text{He}$  and  $^4\text{He}$ . Unlike for LiBeB/C, the BIG model starts to reuse the re-acceleration and galactic wind, which has been dumped before. One could argue at this point, that the BIG model pushes more towards a QUAINT style parameter choice for Helium, whereas in the LiBeB/C benchmark case a SLIM style model description, abandoning  $V_a$  and  $V_c$ , was preferred, given the results in plot 3.3 and table 3.1. The fact that no valid MINOS parameter uncertainty estimations could be found for the parameters in the BIG model may reflect these phenomenological discrepancies, ultimately confirming the concerns about instability mentioned in 1.3.3. Even if the low-rigidity break has shown to be universal for Li, Be, B, and N, its validity for light species may be challenged by these results for Helium.

Although the fluxes of  $^3\text{He}$  and  $^4\text{He}$  fit well in the combination with other flux ratios, figure 3.7 shows that their respective flux ratio  $^3\text{He}/^4\text{He}$  is reproduced with less agreement by the models BIG and SLIM, with QUAINT still being the best. The latter aligns quite well with the data, whereas the other two have a clear discrepancy at  $R \sim 5$  GV, where both the  $K(R)$  and  $q(R)$  (for SLIM) show a low-rigidity break.

Alongside the AMS-02  $^3\text{He}/^4\text{He}$ , there is also precise data from PAMELA available with a lower energy spectrum, presented in 3.7. For the reason of consistency, this data is not fitted in this chapter. Although the PAMELA and AMS-02 data seem to overlap with each other, only the SLIM model (as fitted in this section) seems to reproduce the PAMELA  $^3\text{He}/^4\text{He}$  flux ratio.

### 3.5 MCMC application in the SLIM model for LiBeB/C

It is shown in the previous section 3.3 that the SLIM model can fit the Li/C, Be/C, and B/C flux ratios (LiBeB/C) simultaneously in a good manner, using the MINUIT2 code. This fitting task is now repeated with the Markov-Chain Monte-Carlo approach with a Metropolis-Hastings sampling technique (random walk), introduced in section 2.4.2 and B.3.

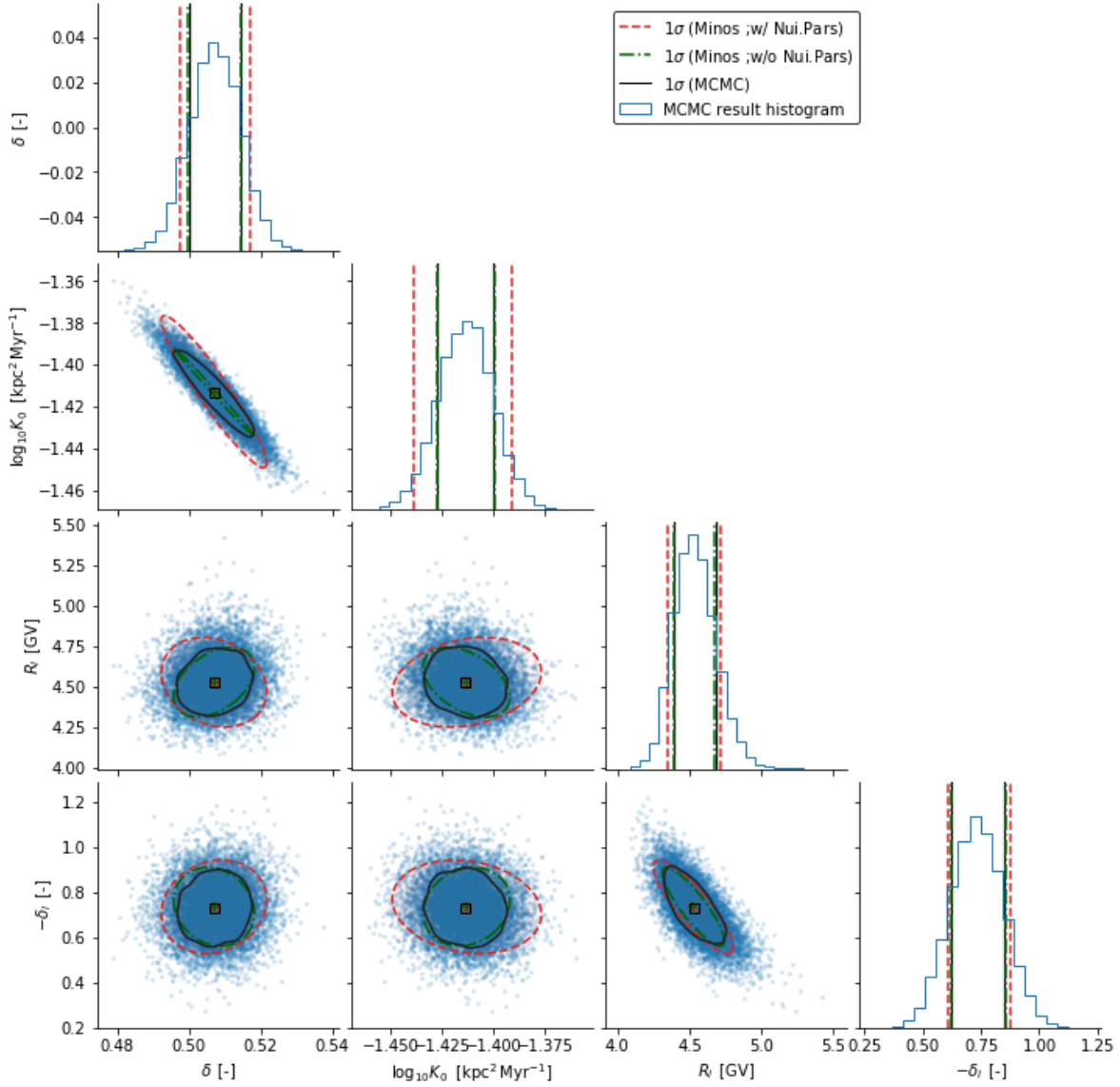
MCMC is a calculation intensive method. For this reason the fitting was done in a simplified regime, in which all nuisance parameters are fixed to their respective best-fit values, found in section 3.3. This leaves a model with only 4 free parameters ( $\log_{10} K_0, \delta, R_l, \delta_l$ ). This simplified model setup is additionally fitted with minos and migrad as well, such that the model setup to be fitted in MCMC has a direct counter-test.

In order to ensure sufficient statistics, 250 independent Markov-Chains were started. However, some runs crashed due to technical issues, leaving 237 valid, independent Markov Chains with 6 468 000 entries in total. After slimming down each chain individually by its correlation length, 10 932 sampled chain elements are left in the final, combined Markov-Chain. The resulting sample is shown in figure 3.8. The histograms and scatter distributions of the Markov Chain elements are in Gaussian and elliptic shape respectively. This indicates that the parameter sampling worked successfully. Using quantiles, contours and medians as a mean of parameter and uncertainty estimations thus seems valid.

Judging by the red 'x', the green '+', and the small black squares in the off-diagonal bars, both migrad fits and the MCMC sampling yield the same best-fit parameter results. Both the green dashed-dotted lines and the black lines correspond to the same fit setup with fixed nuisance parameters and only 4 free parameters. The minos error bars in green align almost perfectly with the  $1\sigma$  quantiles from the MCMC sample. Given the fact that the black and green ellipses also align in a good manner, migrad can be estimated to correctly estimate the parameter correlations. This MCMC sample thus validates the approach of using minos and migrad for correlation, best-fit and uncertainty estimations.

However, the red dashed lines and ellipses differ from the green and black ones in size and shape. Again, they illustrate the best-fit results from the fit in section 3.3, using not only the 4 shown transport parameters, but also 13 cross-section and solar-modulation nuisance parameters. The wider range between the vertical lines in the diagonal bars shows, that the uncertainties estimated by minos are higher, most prominently for the parameter  $\log_{10} K_0$ . Also, the correlation takes different values, e.g. the correlation between  $\log_{10} K_0$  and  $R_l$  is positive instead of negative. This shows, that the presence of 13 additional non-fixed parameters worsens the precision. As a recap, the parameter  $\log_{10} K_0$  is an indirect flux normalization (see equation (1.11) and (1.7)). The much bigger uncertainty in the case with free nuisance parameters (red) compared to the case with fixed nuisance parameters (green and black) is possibly due to a high degeneracy with the production and inelastic cross-sections normalization nuisance parameters (see section 2.2).

To conclude this, the MCMC method used here gives reliable parameter and uncertainty estimations and validates the best-fit values of migrad. In identical settings, MCMC and migrad+minos yield almost identical estimations, which vastly improves the credibility of all fit results throughout this thesis. Cross section uncertainties, which are channeled by nuisance pa-



**Figure 3.8:** Pairplot of the parameter samples when fitting Li/C, Be/C, and B/C in the SLIM model. The elements come from 237 independent Markov Chains, giving 10932 (uncorrelated) parameter vector elements in total. The off-diagonal panels show the scattered chain element parameter values in blue. The black contours and red and green ellipses all correspond to the 68% confidence contours. In the diagonal bars, the single-parameter histograms are shown. The black, red, and green lines also correspond to the 68% confidence interval. The black contours and vertical bars for the histogram or scatter plots are calculated in a way, that 68% of the sampled elements lie inside them. The red dashed lines correspond to the migrad + minos fit results from 3.3. The green dashed-dotted lines correspond to the migrad + minos fits with reduced parameters in section 3.5. The vertical lines in red and green are the respective minos uncertainty range estimations for the individual parameters. The red and green ellipses are generic ellipses calculated from the respective migrad correlation estimations, scaled in x- and y-direction by the respective mean minos-uncertainty estimation times a factor  $\sqrt{2.3}$ . In the off-diagonal bars, the black squares show the best-fit result, estimated via the sample median. The green '+' and red 'x' show the migrad best-fit estimation. In theory black, green and red lines and ellipses all represent the same quantities based on different estimation methods.

rameters here, worsen transport parameter estimations. Improving the first would automatically improve the second as well.

### 3.6 Perspective: Moving from flux ratios to direct fluxes

This chapter has mainly focused on fitting flux ratios, with only a few exceptions. With the successful fits of Li, Be, and B to Carbon or Oxygen flux ratios in section 3.3, this attempt of simultaneous analysis could be taken further by fitting the fluxes directly. Recent examples have shown, that Carbon and Oxygen can be well matched by our models (Génolini et al. 2019). Using the source parameters found in this paper and the LiBeB/C benchmark transport parameters found in this thesis in section 3.3, a flux prediction is made and shown in figure D.5. The latter shows good agreement and underlines the idea of flux fitting. A possible future analysis could be fitting the direct fluxes Li, Be, B, C, and O in different combinations or all together. The respective data with percentage precision already exists (Aguilar et al. 2018a). In the light of having yielded consistent transport parameters for the flux ratios, such an analysis should be feasible. This would not only address the universality of transport parameters, but also a possible universality of source parameters. The fact that the lightweight 1D flat-disc model was sufficient for the studies so far gives hope that it may hold for these future analysis as well, saving a lot of calculation resources.

When analyzing direct fluxes, the models are again fully dependent on source parameters. The spectral form of the injection (1.2) with global slope  $\alpha$ , which holds so far for flux ratios as explained in section 1.3.2, would be probed in that case. In this thesis, the source parameters  $\hat{q}_i$  and  $\alpha$  have still been fixed (except for  ${}^4\text{He}$ ). Moving on to direct flux analysis would thus enlarge the amount of sensitive parameters, which could be involved as nuisance or free parameters.

A preliminary test of such a direct flux analysis was carried out with an older version of the SLIM model. The focus was on Boron, Oxygen, and Carbon data, fitting them either directly or as primary to secondary flux ratios in several combinations. The result of the free parameters used are given in the appendix figure D.4. This first attempt was promising in the way that the major transport parameters  $\delta$ ,  $K_0$  and also the source parameters  $\alpha$ ,  $q_{12\text{C}}$ ,  $q_{16\text{O}}$  are sensitive to the species involved, without showing major inconsistencies for the tested combinations.

The fact that Nitrogen to Oxygen flux ratio was in good alignment with the LiBeB/C benchmark results in all three models, as presented in section 3.4, justifies to also take the Nitrogen flux (Aguilar et al. 2018b) into account as well. One could thus think of a global flux fit for  $Z = 3 - 8$ , reaching from Lithium to Oxygen. In the QUAINT model, even  ${}^3\text{He}$ ,  ${}^4\text{He}$  AMS-02 data could be reproduced, which is why for this model a global combined analysis of  $Z = 2 - 8$  may be thinkable.

Another direction of updating the benchmark models would be including the low-energy data, discussed in section 3.2. It has proven to be consistent with AMS-02 and impact on the low-rigidity parameter constrains. Another important analysis to be done is thus a combined fit of all LE AMS-02 data for Li/C, Be/C, and B/C.

Possible publications of AMS-02 data for even heavier nuclei up to Iron will be a chance to further test the CR models used. Possible primary species that are heavier than Oxygen would be Neon, Magnesium, and Silicon, whereas the next heavier (partly) secondary species would be Fluorine, Sodium and Aluminum (Maurin 2001). For the secondary species Li, Be, B discussed so far, Carbon (except Nitrogen) and Oxygen remain the major production parents (Génolini et al. 2018). This may be one of the reasons for the success of simultaneous fits of their flux ratios. As shown in section 3.4 however, moving to other species like Helium, using different parent nuclei



for the production in ISM, can challenge the transport and source model descriptions. This is why analyzing heavier nuclei beyond Oxygen, who will not be produced via the latter, would be an interesting test of these models in the future.



## Study of the galactic halo size with Beryllium data

The CR models used in this thesis were shown in the last chapter to reproduce the flux ratios including Li, Be, and B very well. In this condition, another critical 1D model parameter is tried to be determined in this chapter, the galactic halo size  $L$ . In the previous studies it has been fixed to 5 kpc, taking ahead results from this chapter. This parameter is of major interest for Dark Matter studies using antiprotons, as explained later in section 4.1.

The main effect of  $L$  on CR transport is scaling the volume in which the diffusive propagation takes place (see figure 1.3). This also scales the average propagation time of particles, as a larger halo favors larger particle path lengths through the halo. Note that  $K_0/L$  is hard-constraint due to the formula 1.11, which is why  $K_0$  is assumed to automatically scale with  $L$  in the following discussion. For stable particles with unlimited lifetime, changing  $L$  (and  $K_0$ ) should not make any major difference, hence we only look at diffusion of intermediate ions. As most of the nuclear interaction, changing the species composition of CR, takes place in the thin galactic disc, the outer diffusive halo size itself leaves this mechanism unaffected. The abundance of radioactive nuclei however is highly dependent on the mean travel time and thus on the size of the diffusive halo. Comparing the abundance of radioactive nuclei to stable ones by fitting their flux ratio could thus be a way of constraining  $L$  as tested in this chapter.

The isotope  $^{10}\text{Be}$  with its half-life of 1.51 Myr is chosen as radioactive candidate. Its advantages are that with  $\sim 10\%$  isotopic fraction of Be it provides a sufficient abundance and that its life span is in the same order of magnitude as the CR transport, demonstrated in figure 1.4. This means that not only  $^{10}\text{Be}$ , but also Be data itself are sensitive to  $L$ , leaving several flux ratios as fitting candidates. In this thesis,  $^{10}\text{Be}/^9\text{Be}$ ,  $^{10}\text{Be}/\text{Be}$ , and  $\text{Be}/\text{B}$  are the main ones used, as justified later in section 4.3. The advantage of secondary-to-secondary flux ratios is, that not only source dependencies, but also transport parameter dependencies are reduced, if both secondaries underlie similar means of transport and production. Because of this, isotopic fractions like  $^{10}\text{Be}/^9\text{Be}$ , that cancel out most transport and source systematics, are the most sensitive to  $L$ . But unfortunately they remain very difficult to measure, which is explained in section 1.4.2 in the context of AMS-02, because separation by mass is often more difficult than separation by charge.

In section 4.1, the current state of halo size studies and their importance for dark matter research is revisited. The next section 4.2 then gives an overview of the CR data for Beryllium, that is used for the later analysis. Before attempting a direct fit of  $L$ , a first test of the realistic dependency on  $L$  for the studied flux ratios is carried out in section 4.3, showing the constraints that can be expected by the given data. The actual fit of  $L$  to the selected CR data is then carried out in section 4.4. One of the most important fits is then repeated with MCMC sampling method in section 4.5. What the yielded results for the halo size mean for future research and how they can be further improved is then discussed in the perspective 4.6.

The studies of the halo size using  $^{10}\text{Be}$ , as presented in this chapter, will contribute to a paper (Weinrich et al. 2020b). This paper will also include  $L$  constraints using the positron analysis by colleagues in LAPTh and LUPM (not covered in this thesis).

## 4.1 Status of halo size analysis and relation to Dark Matter

The previous studies of the diffusive halo size have all yielded values in a comparable region of few kpc, using different approaches. Positrons, who also receive an unavoidable secondary contribution, have also been used to set a lower limit of  $L \gtrsim 3$  kpc (Lavallo et al. 2014). Before the publications of AMS-02 Be/B data, this secondary to secondary flux ratio was predicted to enable a precise constraint of  $L$  with  $\sigma_L \sim 0.5$  kpc, if the nuclear cross sections were to be exact (Tomassetti 2015). With the now existing Be/B AMS-02 data (Aguilar et al. 2018a), a best fit of  $L \sim 7$  kpc was found (Evoli et al. 2020), but the previously expected precision was not met. A huge problem however are the uncertainties of the nuclear cross sections of Beryllium, that dramatically reduce the precision of possible  $L$  constrains. This problem is addressed in this thesis with the use of nuisance parameters (see section 2.2), that parameterize these uncertainties.

The knowledge of the halo size is important for Dark Matter (DM) searches. This is because antiprotons are suspected to be possible annihilation products of WIMPs from a potential DM halo. Although the DM halo ( $r \sim 200$  kpc) is expected to be far bigger than the diffusive halo ( $L \sim 5$  kpc), only the volume in the latter contributes to the local antiproton flux from DM sources (Barrau et al. 2002). In a simplistic model (Maurin et al. 2006), the flux of DM antiprotons is estimated to have a dependency of

$$N_{\bar{p}}(R) \propto L^2/K(R), \quad (4.1)$$

at first order. Because of its degeneracy, given in equation (1.11), only  $L/K_0$  can be directly determined with data from stable CR. With that ratio being fixed however, the degree of freedom  $L$  will leave a direct dependency  $N_{\bar{p}} \sim L$ . In the presence of precise AMS-02 antiproton data, knowing the halo size, and thus the theoretical contribution of antiprotons from WIMPs, can constrain DM models.

More serious DM signal studies are usually carried out in a 2D or 3D regime. However, having a good constraint for  $L$  in the 1D flat halo model could still be a major improvement for studies that previously had to use loose assumptions of  $L$  lying between 1 and 15 kpc (Lavallo & Salati 2012).

## 4.2 Existing data sensitive to $^{10}\text{Be}$

In this thesis, the flux ratios  $^{10}\text{Be}/^9\text{Be}$ ,  $^{10}\text{Be}/\text{Be}$ , and  $\text{Be}/\text{B}$  are taken as the main candidates for constraining  $L$ . For the first two isotopic fractions there is no AMS-02 data yet. Instead, the data used come from several other low- and intermediate-rigidity experiments.

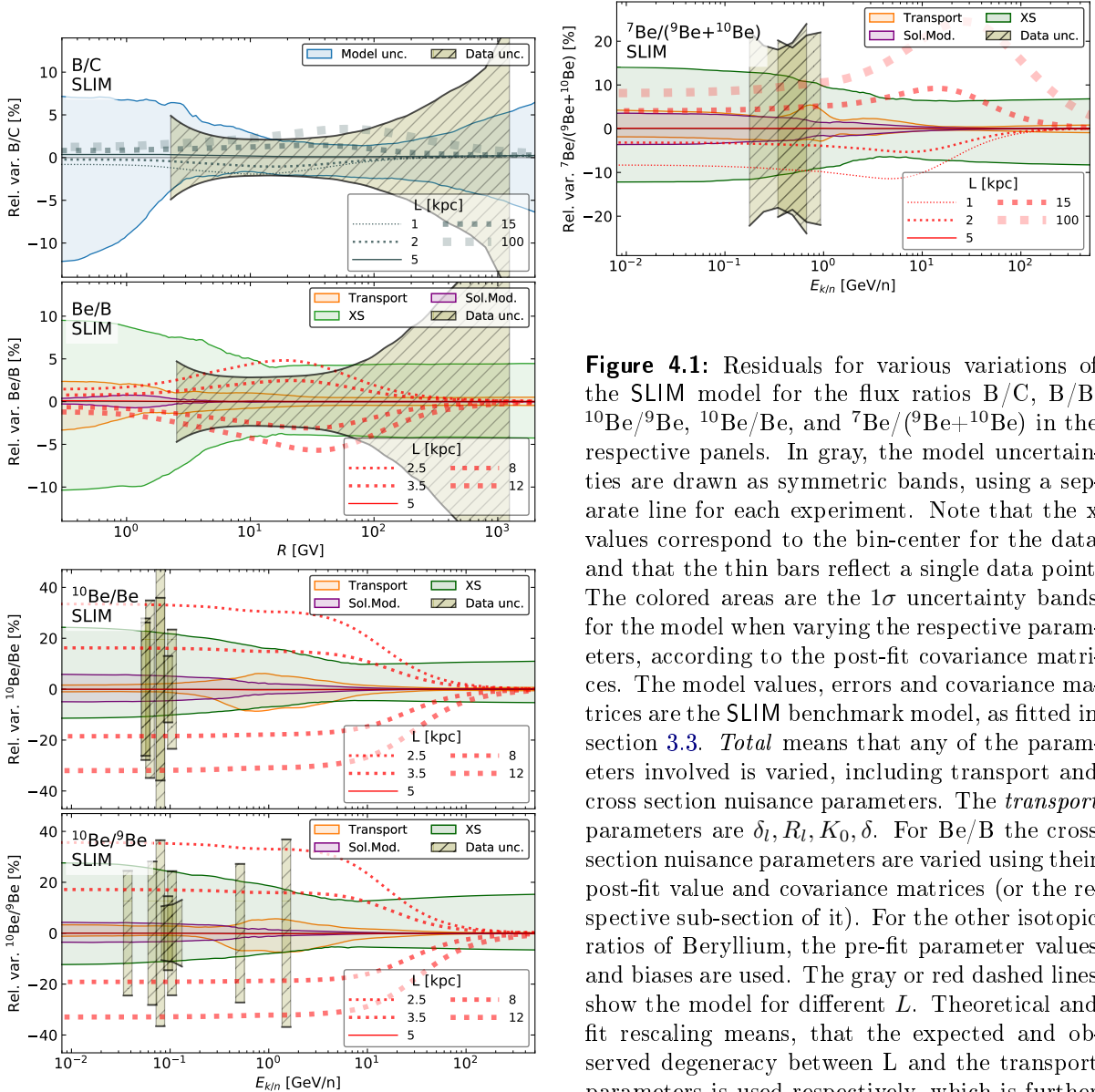
For  $^{10}\text{Be}/\text{Be}$ , the data come from IMP7&8 (Garcia-Munoz et al. 1977, 1981), ISEE3-HKH (Wiedenbeck & Greiner 1980), Ulysses-HET (Connell 1998), and Voyager1&2 (Lukasiak 1999), all being grouped together at  $\sim 0.1$  GeV. The data for  $^{10}\text{Be}/^9\text{Be}$  comes from the same experiments as  $^{10}\text{Be}/\text{Be}$ , but also from ACE-CRIS (Yanasak et al. 2001), ACE-SIS (Yanasak et al. 2001), and ISOMAX (Hams et al. 2004), covering a range of data points of 0.03 – 1.1 GeV. With the recent publications (Aguilar et al. 2018a), Be/B AMS-02 data is available with few percentage errors in the intermediate rigidity regime of 2 - 100 GV. Another possible flux ratio is  $^7\text{Be}/(^9\text{Be}+^{10}\text{Be})$ , as being recently published by the PAMELA experiment (Menn et al. 2018). Although this quantity is relatively easy to measure, it does not provide significant constraining power and is only taken into account in the next section for completeness.

The prominent flux ratios B/C (Aguilar et al. 2016) is also included in this analysis. The latter should in theory not be sensitive to  $L$  as both Boron and Carbon are stable. But because of the nuclear decay  $^{10}\text{Be} \rightarrow ^{10}\text{B}$ , a minor, indirect dependency should be present.

In order to address the cross section uncertainties, nuisance parameters for the main production reactions for the three isotopes  $^7\text{Be}$ ,  $^9\text{Be}$ , and  $^{10}\text{Be}$  are included whenever Beryllium data are fitted in this chapter. The values for  $^{10}\text{Be}$  are generic, whereas the values for  $^9\text{Be}$  and  $^7\text{Be}$  were found in previous analyses in section 2.2. Fitting Beryllium in the following will always include

$$\begin{aligned} (\mu, \sigma)^{12\text{C}+\text{H}\rightarrow^{10}\text{Be}} &= (1.00, 0.25)^{\text{Norm}} \quad \text{and} \quad (0.00, 0.15)^{\text{Slope}}, \\ (\mu, \sigma)^{12\text{C}+\text{H}\rightarrow^9\text{Be}} &= (0.87, 0.20)^{\text{Norm.}} \quad \text{and} \quad (0.03, 0.15)^{\text{Slope}}, \\ (\mu, \sigma)^{16\text{O}+\text{H}\rightarrow^7\text{Be}} &= (0.85, 0.15)^{\text{Norm.}} \quad \text{and} \quad (0.00, 0.15)^{\text{Slope}}. \end{aligned}$$

### 4.3 Expected constraining power



**Figure 4.1:** Residuals for various variations of the SLIM model for the flux ratios B/C, B/B,  $^{10}\text{Be}/^9\text{Be}$ ,  $^{10}\text{Be}/\text{Be}$ , and  $^7\text{Be}/(^9\text{Be}+^{10}\text{Be})$  in the respective panels. In gray, the model uncertainties are drawn as symmetric bands, using a separate line for each experiment. Note that the x values correspond to the bin-center for the data and that the thin bars reflect a single data point. The colored areas are the  $1\sigma$  uncertainty bands for the model when varying the respective parameters, according to the post-fit covariance matrices. The model values, errors and covariance matrices are the SLIM benchmark model, as fitted in section 3.3. *Total* means that any of the parameters involved is varied, including transport and cross section nuisance parameters. The *transport* parameters are  $\delta_l, R_l, K_0, \delta$ . For Be/B the cross section nuisance parameters are varied using their post-fit value and covariance matrices (or the respective sub-section of it). For the other isotopic ratios of Beryllium, the pre-fit parameter values and biases are used. The gray or red dashed lines show the model for different  $L$ . Theoretical and fit rescaling means, that the expected and observed degeneracy between  $L$  and the transport parameters is used respectively, which is further explained in appendix C.

Until now the flux ratios that include a proportion of  $^{10}\text{Be}$  are only suspected to have a dependency on  $L$ , which is now examined in this section. Before attempting a direct fit of  $L$ ,

this is an important check of its potential. The method used in this section is varying the fixed halo size  $L$  for a given model and studying the effect on the flux ratio. The residuals can then be compared to the uncertainties of the respective data existing and also to the uncertainties based on model transport parameter covariance matrices. The resulting residuals, relative to the reference with  $L = 5$  kpc, are plotted in figure 4.1.

For B/C, even the very drastic changes of  $L = 1 - 100$  kpc cause no variation of the model, that would be bigger than the existing data uncertainties. This means that a fit of  $L$  to B/C directly would probably not yield any precise result. The  ${}^7\text{Be}/({}^9\text{Be}+{}^{10}\text{Be})$  flux ratio also has a comparably low dependency on  $L$  that stays below the data uncertainty and thus excludes the given Pamela data as too weak in terms of precision.

For the other three species ratios in figure 4.1 however, their respective data uncertainties are mostly smaller than the variations caused by intermediate variations of  $L = 2.5 - 12$  kpc. For the tested halo sizes, Be/B shows a variation of  $\sim 5\%$  in the intermediate rigidity regime. Although this is not very much, the high AMS-02 data precision of  $\sim 4\%$  is still superior at  $R \sim 20$  GV. This hints on a possible constrain of  $L$  to at  $\pm 3$  kpc with Be/B. The model of the isotopic fraction  ${}^{10}\text{Be}/\text{Be}$  on the other hand shows a very high variation of over  $\pm 30\%$  for the tested  $L = 2.5 - 12$  kpc, overshooting most of the data uncertainties. Even  $L = 3.5 - 8$  kpc cause bigger variations than most of the parameter uncertainties, which is why a  $\pm 3$  kpc constraint can be expected due to  ${}^{10}\text{Be}/\text{Be}$  as well. The variations for  ${}^{10}\text{Be}/{}^9\text{Be}$  are similar to those from  ${}^{10}\text{Be}/\text{Be}$  and thus very promising in terms of constraining power. With the additional data for  ${}^{10}\text{Be}/{}^9\text{Be}$ , including the very precise ACE-CRIS data of 10% uncertainty at 0.1 GeV, and the bigger data range, this flux ratio is suspected to be the most potent one to fit the galactic halo size  $L$  to a potential precision of  $\pm 2$  kpc.

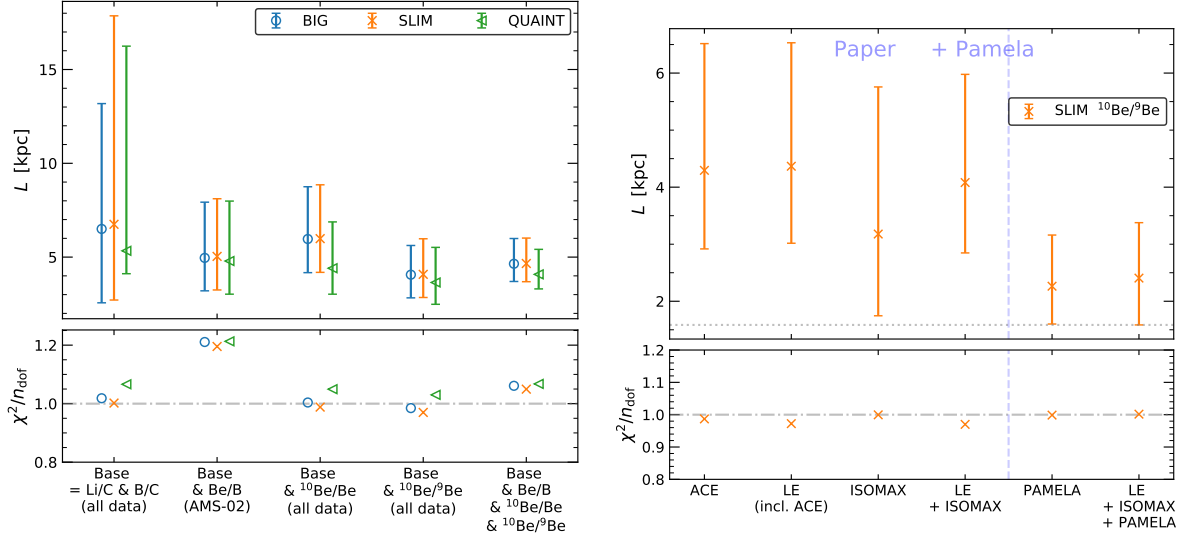
However, all sensitive species show a very big model uncertainty deriving from the cross-sections. The order of magnitude for the pre- or post-fit (called Bias and BestFit respectively) is the same as the data precision. This will most probably enlarge the uncertainty range of  $L$ . Having more precise cross-section data for the Beryllium production would thus improve the  $L$  fit value precision.

## 4.4 Direct determination of $L$

This section describes the procedure and results of finding an estimation and limits of the halo size  $L$  by fitting it directly as a free parameter. For this, several different combinations of Beryllium data, as introduced in the section before, are used.

### 4.4.1 Data fit combinations

As mentioned before, and elaborated in appendix C,  $L$  is more or less correlated with all the other transport parameters. A realistic and good constraint of  $L$  can thus only be found if the other transport parameters are also free (or nuisance parameters). In order to assure a robust transport parameter determination, the stable B/C and Li/C from both AMS-02 and the LE experiments (see section 3.2 for details) are always included. This also includes their respective minimal nuisance parameter setup from table 2.1. Using LE data is also for the sake of consistency, since it comes from similar experiments as  ${}^{10}\text{Be}/{}^9\text{Be}$  and  ${}^{10}\text{Be}/\text{Be}$ . This Li/C+B/C fitting task will be called *Base* fit-configuration in the following, and is supposed to constrain all the non-geometric parameters. The models in the *Base* configuration are thus identical to the ones in 3.2, combining B/C and Li/C. Note that all respective transport parameters are refitted in every case in order to cope for the inter-dependency of  $L$  to other transport parameters.



(a) Each column uses a different combination of species (see text for details). The first column includes only B/C and Li/C from both AMS-02 and LE, which is the base configuration applied in any other column. For QUAINT and Be/B, no lower minos error is found. The dotted lower errorbar is simply the mirrored upper error-bar for illustration purposes.

(b) This figure shows only fits of the base configuration with  $^{10}\text{Be}/^9\text{Be}$  data in the SLIM model. The fourth column corresponds to the fit in the fourth column on the picture 4.2a. The vertical blue-ish line separates the columns including Pamela data, which is not further studied in this thesis. The vertical purple line is only an illustration of the fixed value  $L = 5$  kpc used for all the studies in chapter 2 and 3.

**Figure 4.2:** Halo size  $L$  parameter results for a direct fit. Each symbol represents one of the models BIG, SLIM, and QUAINT. The columns show different species and/or data combinations. For every experiment that has at least one data point involved in the fit, a solar modulation nuisance parameter  $\phi$  was imposed (see section 1.3.2). The other free parameters included in the fit (see table 1.1) are not shown. Errorbars with caps correspond to a minos uncertainty estimations. Dotted error bars show that now valid error estimation was found.

The fit configurations for determining  $L$  now only adds two things. The first addition is obviously  $L$  as a free parameter. The second addition is the use of isotopic cross section nuisance parameters for Beryllium, as explained in 4.1.

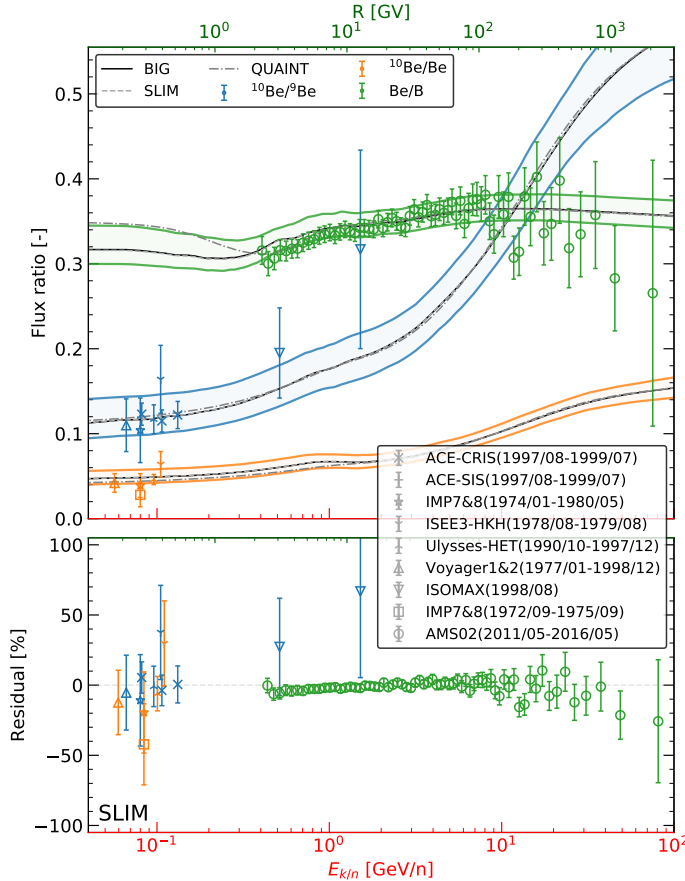
As discussed in section 4.1, there are mainly three ratios tested, namely Be/B,  $^{10}\text{Be}/\text{Be}$  and  $^{10}\text{Be}/^9\text{Be}$ . In order to assure consistency and cross-check with the studies in section 4.3, each species is combined individually with the base configuration in order to fit  $L$ . A fourth fitting attempt uses the base configuration and all the species. A fifth cross-check fit is the absence of any Beryllium data, only fitting the base configuration. However, the Base + Be/B fit is a small exception here, since it is a Li/C, B/C, and Be/B AMS-02 only fit, excluding Li/C and B/C LE data.

The results for the free halo size parameter  $L$  for all three models BIG, SLIM, and QUAINT are given in figure 4.2a. The first column reflects the expectation of not really constraining  $L$  to a precise value in the absence of a radioactive species. However, as demonstrated in figure 4.1, a minor dependency of Boron to the halo size is existing. However, given the other constraints in section 4.1,  $L$  is already much better constrained than this. For other fit combinations including Beryllium, the results for  $L$  are much more precise (note that  $L$  was fitted on log scale). The  $\chi^2/n_{\text{dof}}$  is still good and the lower panel in figure 4.2a shows that the cross-section nuisance parameters were used in a moderate extent.

Combining the base configuration with AMS-02 Be/B yielded  $L \sim 5_{-2}^{+3}$  kpc, which is exactly the constrain that corresponds to the errors presented in figure 4.1 for Be/B. As this fit did not use any LE or ISOMAX data in favor of a being AMS-02 only fit, its exact results are presented in table 4.1, representing the AMS-only response.

The fit including  $^{10}\text{Be}/\text{Be}$  has a similar precision of  $\sigma_L \sim_{-2}^{+3}$  kpc, again validating the predictions from the preliminary tests in section 4.3. However, this species is much more sensitive to the low-energy regime, which can be seen in figure 4.1. This is the region where the three models BIG, SLIM, and QUAINT have most of their discrepancies. As the sensitivity to  $L$  and the data for  $^{10}\text{Be}/\text{Be}$  are only in lower energies, the QUAINT model and its different parametrization yield a different  $\chi^2$  and result for  $L$ . The same effects apply for  $^{10}\text{Be}/^9\text{Be}$  in an analogue way, as it is also using the low-energy region. As already seen in figure 4.1 however, this species covers a larger energy range and has more data with better precision than  $^{10}\text{Be}/\text{Be}$ . This explains the relatively higher precision of the  $L \sim 4_{-1.5}^{+2}$  constrains when fitting  $^{10}\text{Be}/^9\text{Be}$ , as presented in figure 4.2a.

#### 4.4.2 Benchmark constraint with all data



**Figure 4.3:** Post-fit comparison of the model and the data, which was included in the fit. The BIG, SLIM, and QUAINT models were fitted to all data shown simultaneously, corresponding to the fifth column in fig 4.2a. Note that Be/B is plotted in rigidity (x-ticks given in the top), whereas  $^{10}\text{Be}/^9\text{Be}$  and  $^{10}\text{Be}/\text{Be}$  are plotted in kinetic energy per nucleon (x-ticks in the bottom), which is their respective data energy scale. The symbols for each experiment are shown in gray in the lower right legend, whereas the species colors are defined in the top.

*Top panel:* Flux ratio model prediction and the corresponding fit data. The colored bands are the SLIM model uncertainties. They are calculated using the  $1\sigma$  quantiles for all the flux calculations when sampling all fitted parameters with the post-fit covariance matrix.

*Bottom panel:* Residuals of the data points in the top panel to their respective SLIM model prediction in the top panel.

As all three Beryllium flux ratios have given consistent constraints, it is not surprising that the combination of all three together with the base configuration yields the best constrain of less than  $\sigma_L \sim_{-1}^{+1.5}$ . Again, their results are model dependent, with QUAINT preferring a slightly lower, but still consistent value, probably due to its different low-energy propagation description. This data combination of using Be/B,  $^{10}\text{Be}/^9\text{Be}$ , and  $^{10}\text{Be}/\text{Be}$  gives the best constraint and is



thus chosen to be the benchmark result, presented in the table 4.1. The post-fit comparison of all the data involved in this global fit and the resulting model uncertainties are presented in figure 4.3.

	BIG	SLIM	QUAINT
Be/B (AMS-02 only)			
$L$ [kpc]	$4.96^{+2.97}_{-1.76}$	$5.04^{+3.07}_{-1.79}$	$4.79^{+3.19}_{-1.77}$
$\chi^2 / n_{\text{dof}}$	233.7 / 193	233.1 / 195	235.3 / 194
$\chi^2_{\text{nui}} / n_{\text{nui}}$	17.4 / 20	17.4 / 20	15.8 / 20
Be/B, $^{10}\text{Be}/^9\text{Be}$ , $^{10}\text{Be}/\text{Be}$ (all experiments)			
$L$ [kpc]	$4.64^{+1.35}_{-0.94}$	$4.66^{+1.35}_{-0.97}$	$4.08^{+1.33}_{-0.78}$
$\chi^2 / n_{\text{dof}}$	266.3 / 251	265.6 / 253	269.0 / 252
$\chi^2_{\text{nui}} / n_{\text{nui}}$	25.6 / 35	25.4 / 35	25.6 / 35

**Table 4.1:** Halo size fit results for the combined analysis. The first block corresponds to Li/C, B/C,  $^{10}\text{Be}/^9\text{Be}$ , and  $^{10}\text{Be}/\text{Be}$  from AMS-02, LE, and ISOMAX. The second block shows the results when using only the AMS-02 data for Li/C, B/C, and Be/B. See text for details.

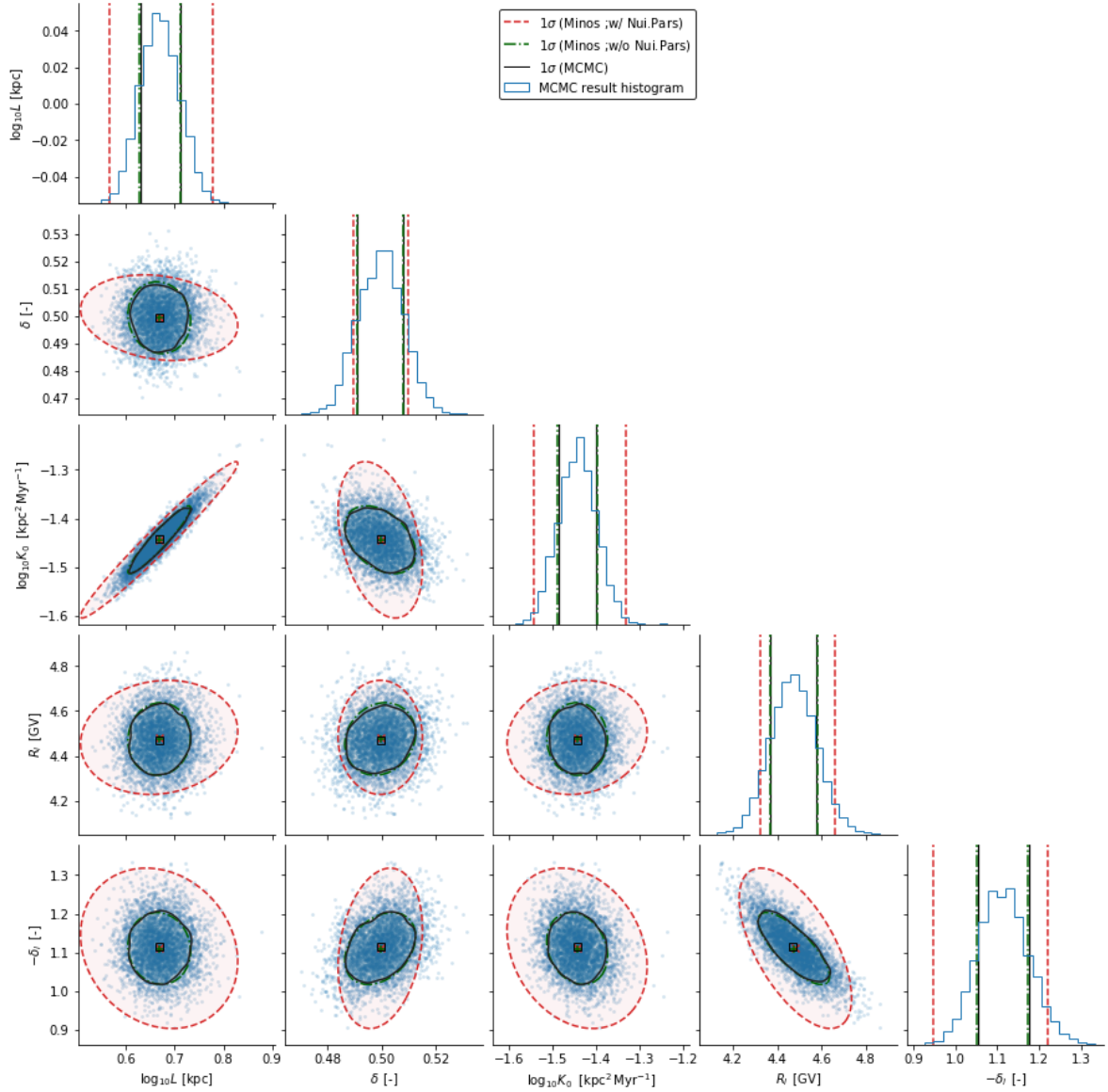
#### 4.4.3 Comparison of $^{10}\text{Be}/^9\text{Be}$ data between experiments

As mentioned in the introduction of this chapter, predicted in section 4.3, and proven in figure 4.2a,  $^{10}\text{Be}/^9\text{Be}$  is the most potent species for constraining the galactic halo size. Fortunately there are several different experiments, which provide respective data. In another test, the constraints of the different  $^{10}\text{Be}/^9\text{Be}$  data sets are compared. For this, the SLIM model is used to fit the base configuration with the respective sub-selection of  $^{10}\text{Be}/^9\text{Be}$  data. Low-Energy (LE) corresponds to all the experiments mentioned in section 4.2 without ISOMAX. ACE is simply short for ACE-CRIS and ACE-SIS together, which is tested individually, although it is included within LE by default. The preliminary PAMELA data for  $^{10}\text{Be}/^9\text{Be}$  was not used in this thesis so far and is now compared in their impact on the  $L$  result as well.

The results for this study are given in figure 4.2b. For ISOMAX and LE, the results for  $L$  are slightly different but well within their respective uncertainty estimations. Only testing ACE without the other experiments in the LE set gives almost the exact same constraint as using the whole LE set. This is probably due to the fact that ACE-CRIS and ACE-SIS exceed all other LE experiments in terms of precision. The fourth column, which corresponds to the standard data combination for  $^{10}\text{Be}/^9\text{Be}$  as used in figure 4.2a, has a value between the one for LE and ISOMAX and slightly reduces the error bar. The fits using preliminary data from PAMELA totally change the results, even when being combined with the other data. The lower error bar touches the parameter limit of 1.6 kpc. The best fit values for the fit with and without Pamela are outside the  $1\sigma$  uncertainties of the respective other result. Using the base configuration and all  $^{10}\text{Be}/^9\text{Be}$  data, including PAMELA, yields  $L = 2.4^{+1.2}_{-0.7}\text{kpc}$ . Because of these huge differences compared to table 4.1, we will be very interested in the final PAMELA analyses and also AMS-02 data.

## 4.5 MCMC application in the SLIM model with all Beryllium data

Just like in section 3.5, the benchmark fit of SLIM to Li/C, B/C, Be/B,  $^{10}\text{Be}/\text{Be}$ , and  $^{10}\text{Be}/^9\text{Be}$  data in the section 4.4.2 is repeated in with the MCMC method. Again, all nuisance parameters are fixed to their best-fit values from the fit carried out in 4.4.2. The original fit from section



**Figure 4.4:** Pairplot of the parameter samples when fitting Li/C, B/C, Be/B,  $^{10}\text{Be}/\text{Be}$ , and  $^{10}\text{Be}/^9\text{Be}$  in the SLIM model. The elements come from 99 independent Markov Chains, giving 4767 (uncorrelated) parameter vector elements in total. For a detailed plot-type description see figure 3.8, which is created with identical methods.

4.4.2 used 16 solar-modulation and 14 cross-section nuisance parameters, with 6 describing only the production of Beryllium isotopes in details (see section 4.2). This leaves a regime of 5 free parameters ( $\log_{10} L, \log_{10} K_0, \delta, R_l, \delta_l$ ), which helps with speed and numerical stability. In order to have a direct mean of comparing MCMC to migrad+minos methods, this reduced model fit is carried out with the latter as well.

From the 250 attempted Markov Chains, 99 went successfully with a combined total length of 4950000. Thinning down each chain by its respective auto-correlation length leaves a final parameter sample with 4767 elements. The sample is shown in figure 4.4.

The things to conclude from this plot are quite similar to the MCMC studies in section

3.5. First, all the parameter histograms in 4.4 are of Gaussian shape and the 2D parameter distributions are elliptic, reflecting the naive assumption about the parameter space. Second, for the identical model setups, MCMC and minos+migrad yield almost identical best-fit, correlation and uncertainty estimations, judging by the alignment of green and black lines in plot 4.4. Third, having nuclear cross sections as nuisance (modifiable) parameters vastly distorts constraints of other parameters. Most prominently, the uncertainty range of  $\log_{10} L$  and  $\log_{10} K_0$  is almost three times higher in the model with, compared to the reduced one without nuisance parameters. This can be explained with the presence of many cross-section normalization nuisance parameters, being degenerated with the normalization parameters  $K_0, L$ . The additional uncertainty of cross-sections due to nuisance parameters thus enlarges the uncertainties of  $K_0$  and  $L$ . This is another clear demonstration of the need for better and more precise nuclear cross sections.

## 4.6 Perspective: Isotopic AMS-02 Beryllium data and other heavy radioactive nuclei

Although the existing data for Be/B,  $^{10}\text{Be}/^9\text{Be}$  and  $^{10}\text{Be}/\text{Be}$  has lead to a precise constraint of  $L$  in the 1D model, a certain part of the remaining uncertainty could be eliminated with more precise cross-section data for Beryllium production. Before its publication, the assumptions about the AMS-02 Be/B data precision where predicted to give a  $\sigma_L^{\text{Be/B}} = 0.5$  constrain in Tomassetti (2015). The fact that this expectation was not met in this thesis ( $\sigma_L^{\text{Be/B}} \gtrsim 2 \text{ kpc}$ ), is at least to some extent due to cross-section uncertainties.

Out of all the three species fitted so far,  $^{10}\text{Be}/^9\text{Be}$  is the most powerfull. Unfortunately, there is no isotopic data from AMS-02 yet available. But the AMS-02 data publications have proven several times to outperform other experiments in terms of data range and precision. If an eventual publication AMS-02 of  $^{10}\text{Be}/^9\text{Be}$  will have similar of better precision than the existing data, the fit of the halo size  $L$  would be further improved.

But Beryllium is not the only radioactive species, that is suitable for geometric analysis. The criteria for a radioactive isotope in order to be usefull for geometric analysis is mostly its decay half-life  $t_{1/2}$ . If the latter is too small, the isotope will decay to rapidly after production and not reach the Earth in a significant abundance fraction. If the half-life of the isotope is however to big, its abundance is less sensitive to the halo size. Recalling figure 1.4, the characteristic times scales are 10 – 100 Myr (in this assumption). If the species nucleus lives longer than that, its decay is not the dominant physical effect, making it act like a stable particle.

Looking at the radioactive isotopes heavier than  $^{10}\text{Be}$ , the next two who are suitable in this regard are  $^{26}\text{Al}$  and  $^{36}\text{Cl}$  (Maurin 2001). With their half-life times for  $\beta$  decay of  $t_{1/2} = 0.91 \text{ Myr}$  and  $t_{1/2} = 0.31 \text{ Myr}$  respectively, they decay in a similar time scale as  $^{10}\text{Be}$  with  $t_{1/2} = 1.51 \text{ Myr}$ . The the decay products for these heavy isotopes are  $^{26}\text{Mg}$  and  $^{36}\text{Ar}$  respectively. Given the charge measurement precision (see section 1.4), AMS-02 should be able to provide elemental fluxes or flux ratios for Aluminum, Chlorine, Magnesium, and Argon. However, a charge separation of e.g.  $^{26}\text{Al}$  from the stable  $^{27}\text{Mg}$  is not possible with the precision of  $\beta$  measurements in AMS-02. In analogy to Be/B, the flux ratios to study would be Al/Mg and Cl/Ar.

Before analyzing these two radioactive species, there are two major obstacles to be solved. First, data with a sufficient precision must be available. Second, the CR transport for Aluminum and Chlorine must be understood. The studies in this Chapter with  $^{10}\text{Be}$  were only possible because the transport for Beryllium could already be sufficiently described in section 3.3. Using  $^{26}\text{Al}$  and  $^{36}\text{Cl}$  for studies of the galactic geometry would first need to expand the CR transport studies to  $Z = 13$  and  $Z = 17$  respectively. But this is rather a question of time and not of feasibility.



## Conclusion

In this thesis, we looked at the transport mechanism of CR, in particular light and medium ions. Although this is a complex phenomenon at galactic scales, the description could be broken down by several assumptions to a one dimensional simplistic diffusion model. This model allowed to compute model predictions for different species abundances and their fractions with the USINE framework in agreement with the data. Before starting to use the existing minimization techniques, they were first cross-checked, which revealed that fit results from the MINUIT2 are not always secure and must be taken with caution. This problem was solved by always using hundreds of independent retries of each fitting task, eventually finding proper results for almost any fit configuration.

The models and techniques in this thesis had previously been used to study AMS-02 B/C data and yield benchmark models. The additional, recent publications for Lithium and Boron high precision data now gave the opportunity to look at a wider range of species and test the universality of transport for  $Z = 3, 4, 5$ . The existing uncertainties for nuclear cross sections, which are an important factor in CR propagation, were treated with nuisance parameters. The necessary configurations for this, which had already been determined for Boron, were then also worked out for Lithium, Beryllium, Nitrogen, and Helium. With an analysis set up inspired by two recent papers (Génolini et al. 2019; Derome et al. 2019), the diffusion CR models could successfully be constrained with the AMS-02 secondary-to-primary flux ratio data. Combining Lithium, Beryllium, and Boron data has ultimately yielded universal CR propagation models holding for  $Z = 3, 4, 5$ , without sacrificing the precision. This moves forward the idea of universal CR transport model for sub-Oxygen ions. Since this thesis has shown that the 1D flat-disc approach is sufficient to describe the discussed AMS-02 data, its viability as a tool of research is approved. These new and solid benchmark models can now be used to predict other, yet unstudied CR regimes, or be a point of departure for analyzing secondary or primary fluxes directly.

Coming from these successful combined analysis, the description of secondary species was tried to be taken further by fitting Nitrogen to Oxygen flux ratio and Helium fluxes, published by AMS-02 as well. Despite revealing new technical and phenomenological difficulties, first attempts were promising concerning a potential transport universality study for  $Z = 2 - 8$ . This is the first time in literature that all these species are well matched with one universal CR propagation model.

With this solid CR models that can describe the precise AMS-02 Beryllium data, the radioactive species  $^{10}\text{Be}$  and its impacts became observable. The diffusive halo size of the galaxy and its impacts on the CR propagation time was studied with the species fractions  $\text{Be}/\text{B}$ ,  $^{10}\text{Be}/\text{Be}$ , and  $^{10}\text{Be}/^9\text{Be}$  from different experiments. The sensitivity of these ratios to the halo size was examined in respect of other model and data uncertainties. Fitting the combination of all the data has yielded a unprecedented constraint for the halo size of  $L \sim [3.3 - 6.0] \text{ kpc}$  in  $1\sigma$  CL. This can give better constraints for DM candidates, e.g. via antiproton analyses. Although this geometric fit was carried out in a one-dimensional model, the halo size  $L$  constraints can be applied in a 2D or 3D galactic disc model as well. Given the fit reliability of the 1D geometry one could even think of analyzing extended geometry with the Beryllium data. With this knowledge of the sensitivity of several species fractions to the galactic halo size and the given benchmark values, we have expectations for possible future  $^{10}\text{Be}/\text{Be}$  AMS-02 data releases.

After implementing a Python bridge of the USINE framework, the Markov-Chain Monte-Carlo method could be used successfully to find parameter distributions for the models. These MCMC parameter results are not only well defined, but also in very good agreement to the respective MINUIT2 results, eventually validating the fitting methods throughout the thesis.

The 1D flat disc model calculated in the USINE framework, which was used in this thesis and keeps up with the published AMS-02 CR ion data so far, is a powerful tool for CR research. As demonstrated with the halo size constraint, it still captures most of the relevant processes. So far, it does not stand behind 2D or 3D approaches, so that it will be interesting to further test it and see how it behaves for heavier nuclei. The implementation of cross-section nuisance parameters and a covariance matrix of systematics, not used by other models, make a difference here. This shows that, given the precision of AMS-02 data, cross section and data systematics uncertainties must be accounted for, when aiming for meaningful results. We can conclude that the AMS-02 data for heavier nuclei will further boost the models and shed light on source and transport of cosmic rays in the Galaxy.

# PyUSINE

USINE has proven to be a very powerful and effective framework for CR model calculations. The idea behind the PYUSINE extension is to enable the combination of this existing CR calculation framework with custom python code, giving a flexibility in the numerical post-processing like minimization. This is achieved without changes to the framework itself and a C++-Python bridge that interfaces only the class containing the  $\chi^2$  function.

## Interfacing the $\chi^2$ function

Any minimizer or MCMC routine needs to be able to call the  $\chi^2(\vec{\theta}, \vec{\eta}) \rightarrow \mathbb{R}$  function with values for the free and nuisance parameters  $\vec{\theta}, \vec{\eta}$  respectively as input and get back the scalar value that can be further interpreted. The choice of which parameters to try and what to do with the evaluated  $\chi^2$  result depends on the algorithm. It is not necessary for the minimizer of MCMC code to interact or know the internal procedure of the  $\chi^2$  function. The objective is to construct a way of calling this  $\chi^2$  function within Python with custom parameter values but leaving its evaluation untouched within the C++ USINE code.

Within USINE is the `TURunPropagation` class<sup>1</sup>, which includes the necessary `Chi2_TOAFluxes` function. When being called with a parameter vector  $(\vec{\theta}, \vec{\eta})$ , it calculates internally the CR flux and evaluates the  $\chi^2$  respectively to the data. But before being able to do so, the class needs to be once initialized with an input file using its `SetClass` and `UpdateFitParsAndFitData` functions. The model settings and fixed, nuisance, and free parameters are defined in the USINE ASCII input file. The data as well as the cross section sets are provided in external files.

## Wrapping C++ functions in Python with Pybind11

In order to make a C++ function callable in Python, one needs a so-called wrapper code, acting as a bridge. It contains the desired C++ functions to be executed in Python and is dependent on the whole USINE code. Additionally one needs to add the `PyBind11` package<sup>2</sup> and declare the functions, classes, and variables in a special syntax. This wrapper code is then compiled as C++ code, creating a Python package.

The latest code version is shown:

```
// Wrapper_Code.cpp
#include <pybind11/pybind11.h> // Class containing the wrapper functionality
#include <pybind11/stl.h>
#include <vector>

#include <TURunPropagation.h> // Class to be partly interfaced
#include <TUMath.h>

// New Python class containing the desired functions
class PyRunPropagation: public TURunPropagation{
```

<sup>1</sup><https://dmaurin.gitlab.io/USINE/doxygen/classTURunPropagation.html>

<sup>2</sup><https://pybind11.readthedocs.io/>

```
private:
    FILE *f_log;          // this is the log file that the super-class demands
    std::vector<int> fixed_indices;

    // routine to recombine fixed and free parameters to one vector
    std::vector<double> RecombinePars (std::vector<double> const v_free_pars){
        // Merges the given Free and Nuisance parameters 'v_free_pars' with the fixed ones
        std::vector<double> v_all_pars (v_free_pars);
        for (signed int i = 0; i < static_cast<int>( fixed_indices.size() ); i++){
            int FI = fixed_indices[i];
            double val = GetFitPars()->GetParEntry(FI)->GetFitInit() ;
            v_all_pars.insert(v_all_pars.begin()+ FI, val);
        }
        return v_all_pars;
    }

public:
    PyRunPropagation():TUNRunPropagation(){
        // this sets the log file demanded by the super class
        f_log = fopen("run.log", "w");
    };

    // The super-class TUNRunPropagation needs to be initialized with the model setting
    // provided in the ASCII usine_initfile
    void PySetClass(string const &usine_initfile, Bool_t is_verbose){
        string const output_dir = "Usine-Out";
        SetClass(usine_initfile, is_verbose, output_dir, f_log);
        UpdateFitParsAndFitData(usine_initfile, is_verbose, is_verbose);
        fixed_indices = GetFitPars()->IndicesPars( gENUM_FREEPARTYPE(1) );
    }

    // The Chi^2 function to be called with a vector of free parameters v_pars
    Double_t PyChi2(std::vector<double> const v_pars){
        std::vector<double> v_input = RecombinePars(v_pars);
        return Chi2_TOAFluxes(&v_input[0]);
    }

    // Function to get a list of the names of all non-fixed parameters
    std::vector<string> PyGetFreeParNames(){
        std::vector<string> par_names;
        for (Int_t l = 0; l < GetFitPars()->GetNPars(); ++l){
            if (GetFitPars()->GetParEntry(l)->GetFitType() != gENUM_FREEPARTYPE(1)){
                par_names.push_back(GetFitPars()->GetParEntry(l)->GetFitName());
            }
        }
        return par_names;
    }

    // Function to get a list of the names of all fixed parameters
    std::vector<string> PyGetFixedParNames(){
        std::vector<string> par_names;
        for (Int_t l = 0; l < GetFitPars()->GetNPars(); ++l){
            if (GetFitPars()->GetParEntry(l)->GetFitType() == gENUM_FREEPARTYPE(1)){
                par_names.push_back(GetFitPars()->GetParEntry(l)->GetFitName());
            }
        }
        return par_names;
    }
}
```



```

// initial free parameter values from the init file
std::vector<double> PyGetInitTheta(){
    std::vector<double> theta;
    for (Int_t l = 0; l < GetFitPars()->GetNPars(); ++l){
        if (GetFitPars()->GetParEntry(l)->GetFitType() != gENUM_FREEPARTYPE(1)){
            theta.push_back(GetFitPars()->GetParEntry(l)->GetFitInit());
        }
    }
    return theta;
}

// initial free parameter uncertainties from the init file
std::vector<double> PyGetInitSigma(){
    std::vector<double> sigma;
    for (Int_t l = 0; l < GetFitPars()->GetNPars(); ++l){
        if (GetFitPars()->GetParEntry(l)->GetFitType() != gENUM_FREEPARTYPE(1)){
            sigma.push_back(GetFitPars()->GetParEntry(l)->GetFitInitSigma());
        }
    }
    return sigma;
}

// parameter bounds from the init file
std::vector<std::vector<double>> PyGetInitBound(){
    std::vector<std::vector<double>> bound;
    for (Int_t l = 0; l < GetFitPars()->GetNPars(); ++l){
        if (GetFitPars()->GetParEntry(l)->GetFitType() != gENUM_FREEPARTYPE(1)){
            std::vector<double> b;
            b.push_back(GetFitPars()->GetParEntry(l)->GetFitInitMin());
            b.push_back(GetFitPars()->GetParEntry(l)->GetFitInitMax());
            bound.push_back(b);
        }
    }
    return bound;
}
};

namespace py = pybind11;
namespace pybind11{

// This defines the Python package 'PyProp'
PYBIND11_MODULE(PyProp, m){
    // Interface for the Usine super-class; functions are also callable by 'PyRunPropagation'
    class_<TURunPropagation>(m, "TURunPropagation")
        .def(init<>(), "Initializer, not doing anything")
        .def("SetOutputDir", &TURunPropagation::SetOutputDir)
        .def("GetOutputDir", &TURunPropagation::GetOutputDir)
        .def("GetNdof", &TURunPropagation::GetNdof)
        ;

    // Interface for the inherited class with the custom functions
    class_<PyRunPropagation, TURunPropagation>(m, "PyRunPropagation")
        .def(init<>(), "Initializer, not doing anything")
        .def("PySetClass", &PyRunPropagation::PySetClass)
        .def("PyChi2", &PyRunPropagation::PyChi2)
        .def("PyGetFreeParNames", &PyRunPropagation::PyGetFreeParNames)
        .def("PyGetFixedParNames", &PyRunPropagation::PyGetFixedParNames)

        .def("PyGetInitTheta", &PyRunPropagation::PyGetInitTheta)

```

```
.def("PyGetInitSigma", &PyRunPropagation::PyGetInitSigma)
.def("PyGetInitBound", &PyRunPropagation::PyGetInitBound)
;
}
```

## The basic Python usage

The following code is a minimal example of how to use the python package created by Pybind11.

```
import PyProp as PP          # import of the Python package

PRP = PP.PyRunPropagation()  # Setting up the TURunPropagation class interface
PRP.PySetClass("init_file.par", 1, "Usine_OutDir") # "init_file.par" is the ASCII input file

# Call of the Chi^2 with the free parameter values in list Theta
Chi2 = PRP.PyChi2(Theta)
```

The `PRP.PyChi2` function can be called with any list `Theta` of the length of `PRP.PyGetInitVals()` or `PRP.PyGetFreeParNames()`, which is the number of the free (and nuisance) parameters. Although the USINE code makes a difference between free and nuisance parameters, in the python code this difference no longer exists. Both of them are free to be modified and treated in one vector  $(\vec{\theta}, \vec{\eta}) \rightarrow \vec{\theta}$ . Because the `TURunPropagation` super-class only executes the initialization and  $\chi^2$  calls, the directory `Usine_OutDir` will remain empty. Having this usability of the  $\chi^2(\theta)$  function in place, it can now be given to any minimizer or MCMC algorithm to be called with any free parameter values.

## PyUsine Python class

Based on the `PyProp` class, the `PyUsine` class is written, allowing more flexibility and features. It has a `PyProp` instance as a member and is the base for all the MCMC and minimization classes. Alongside more verbosity, the additional interface avoids recalculation of  $\chi^2$  for identical  $\vec{\theta}$  with caching, logs the function calls in a dataframe and enforces parameter-boundaries.

```
import PyUsine
PU = PyUsine.PyUsine('init_file.par') # Setting up the class
PU.Chi2(Theta)                        # Chi^2 call with parameter-list Theta
```

As an example, the following lines show the basic implementation of minimizing the  $\chi^2$  with the python package `scipy.optimize` after execution of the three lines above. The list `PU.Theta0` contains the starting parameter values, declared in `init_file.par`.

```
from scipy.optimize import minimize
result = minimize(PU.Chi2, PU.Theta0, method="Powell")
```

As all the  $\chi^2$  and the respective  $\vec{\theta}$  are stored in a data frame automatically for each call, the performance of e.g. minimizers can be observed.

## Minimization techniques

### B.1 Minimizer algorithms

This section briefly describes the algorithms for minimizing a  $\chi^2$  function in this thesis. The Naive Genetic algorithm, explained in B.1.4, is implemented for this thesis and only one example of a genetic minimization approach.

#### B.1.1 Minuit with Migrad

Minuit was always used with the Migrad minimizer algorithm in this study. The Migrad algorithm is based on the Davidon-Fletcher-Powell Method<sup>1</sup>. It uses the gradients to find the minimum of a scalar function. The parameter covariance matrix  $C$  can be calculated from the Hessian matrix  $H(\chi^2_{\min}) = C^{-1}$ , resulting as a by-product of the minimization. Let  $f(x) \equiv \chi^2(x \equiv \theta)$  be the function to minimize and  $N = \dim(\theta)$ .  $H(f(x)) \in \mathbb{R}^{N \times N}$  is the Hessian matrix of  $f(x)$ . With a given starting vector  $x_0$  the algorithm works as follows:

0.  $H_0 = \mathbb{I}_N$ ,  $k = 0$
1. calculate the deviation  $d_k = H_k \cdot \nabla f(x_k)$
2. minimize  $f(x_k + \delta_k)$ ,  $\delta_k \propto d_k$  with line search along axis  $d_k$   
set  $x_{k+1} = x_k + \delta_k$
3. calculate  $g_k = \nabla f(x_{k+1}) - \nabla f(x_k)$
4. set  $H_{k+1} = H_k + \frac{\delta_k \delta_k^T}{\delta_k^T g_k} - \frac{H_k g_k g_k^T H_k}{g_k^T H_k g_k}$  as new Hessian matrix
5. restart step 1. with  $k \rightarrow k + 1$

Possible breaking conditions are  $|d_k| \sim 0$  or  $|\delta_k| \sim 0$ .

#### B.1.2 Powell's method

The Powell's method does not depend on any gradient. Depending on the function to be minimized, this may be of an advantage. Sometimes (like in this thesis) no analytical definition of the function gradient is available. Instead it needs to be estimated with calculations of the function in small offsets in each direction. This may add numerical instability. Powell's method however is based on the line-search approach only. In several iterations, the minimum along different axes in the parameter space is searched. These axes are continually updated after some steps.

Let  $\theta_0$  be the starting points and  $U_i = \hat{e}_i, i \leq N$  the search vectors with  $N = \dim(\theta)$ . The basic algorithm concept, as described in Mathews & Fink (2004), works the following way:

1. for all  $i \leq N$  find  $P_i \propto U_i$  that minimizes  $\chi^2(\theta^* + P_i)$ ;  $\theta^* \rightarrow \theta^* + P_i$
2.  $U_i = U_{i+1}$  for all  $i < N$ ,  $U_N = \theta^* - \theta_0$

<sup>1</sup>[http://www.mymathlib.com/optimization/nonlinear/unconstrained/fletcher\\_powell\\_davidon.html](http://www.mymathlib.com/optimization/nonlinear/unconstrained/fletcher_powell_davidon.html)

3. find  $P \propto U_N$  that minimizes  $\chi^2(\theta^* + P)$  ;  $\theta^* \rightarrow \theta^* + P$
4.  $\theta_0 \rightarrow \theta^*$ , restart step 1.

A possible stopping criteria is  $|\theta_0 - \theta^*| \approx 0$  after step 3.

### B.1.3 Conjugate Gradient Method (CG)

The CG algorithm tries to solve the minimization problem with numerical linear algebra:

$$\begin{aligned}\chi^2(\mathbf{x} \equiv \theta) &\equiv f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}^T \mathbf{b} \\ \nabla f(\mathbf{x}) &= \mathbf{A} \mathbf{x} - \mathbf{b} \stackrel{!}{=} \mathbf{0} \\ \nabla^2 f(\mathbf{x}) &= \mathbf{A}\end{aligned}$$

The Algorithm is taken from Nocedal & Wright (1999). From a starting vector  $x_0$  it works as follows:

0.  $p_0 = -\nabla f(x_0)$ ,  $k = 0$
1. find  $\alpha_k \in \mathbb{R}$  that minimizes  $f(x_{k+1})$  with  $x_{k+1} = x_k + \alpha_k p_k$
2.  $\beta_{k+1} = \frac{(\nabla f(x_{k+1}))^2}{(\nabla f(x_k))^2}$
3.  $p_{k+1} = -\nabla f(x_{k+1}) + \beta_{k+1} p_k$
4. restart step 1. with  $k \rightarrow k + 1$

The loop is stopped once that  $\nabla f(x_k)$  has reached a sufficiently small value  $\approx \vec{0}$ .

### B.1.4 Naive Genetic approach

The genetic approach of finding a minimum is random based. It does not analyze the structure of the  $\chi^2$  hyper-surface but rather follows the *try and error* principle. The algorithm works in generations. In each generation  $G_k$  the new parameter vectors are randomly drawn with respect to those that had the best  $\chi^2$  value in the previous generation  $G_{k-1}$ . There are different ways how to sample the new generation. In the following,  $\mathcal{N}(\theta, \sigma_\theta)$  is a random vector drawn from a multidimensional Gaussian distribution, centered around  $\theta$  with scaling  $\sigma_\theta$ .  $(\theta_0, \sigma_{\theta,0})$  are the starting parameters. In this studies, the following algorithm was used:

0. Generate  $G_0$  with elements  $\theta_i^{(0)} = \mathcal{N}(\theta_0, \sigma_{\theta,0} \cdot 0.1)$ ,  $i < N_{\text{Pop}}$  and calculate  $\chi^2(\theta_i^{(0)})$ ,  $k \rightarrow 1$
1.  $F = \frac{5}{5+k} \cdot 0.1 \sigma_{\theta,0}$  as new, smaller sampling-scale.
2. add  $N_{\text{Parent}}$  elements of  $G_0$  with highest  $\chi^2$  from  $G_{k-1}$  to  $P_k$
3. Calculate mean  $\bar{\theta}$  from all the elements from  $P_k$   
Add elements  $\theta_i^{(k)} = \mathcal{N}(\bar{\theta}, F \cdot 2)$ ;  $i \leq N_{\text{new}}$  to  $G_k$
4. for any  $N_{\text{new}} < j \leq N_{\text{Pop}}$ , pick two different elements  $\theta_{a,b}$  from  $P_k$  and create vector  $\theta_j^{(k)}$  with each vector element being taken randomly from the respective position of either  $\theta_a$  or  $\theta_b$ .  
Modify the element  $\theta_j^{(k)} \rightarrow \theta_j^{(k)} + \mathcal{N}(\theta_0, F)$  and add it to  $G_k$

5. Calculate  $\chi^2(\theta_i^{(k)})$  for any  $\theta_i^{(k)} \in G_k$
6.  $\theta_{\min} = \begin{cases} \theta_i^{(k)}, & \chi^2(\theta_i^{(k)}) = \min(\chi^2(\theta \in G_k), \chi^2(\theta_{\min})) \\ \theta_{\min} & \text{else} \end{cases}$
7. restart from step 1. with  $k \rightarrow k + 1$

The scaling  $F$  introduces a randomness, that starts at  $0.1\sigma_{\theta,0}$  and decreases with every generation. For the studies in section 2.4, the numbers  $N_{\text{Pop}}, N_{\text{Parent}}, N_{\text{new}}$  were set to be 50, 15 and 5 respectively. Step 3. introduces  $N_{\text{new}}$  new random elements, whereas step 4. *crossbreeds*  $N_{\text{Pop}} - N_{\text{new}}$  new elements from the  $N_{\text{Parent}}$  parents.

A possible break-up condition for the loop would be, if the best-fit parameter  $\theta_{\min}$  was not updated in step 6. for several generations.

## B.2 Merging multiple Minuit2 fit results

This section describes the heuristic procedure of combining the results of many independent MINUIT2 runs of the same fitting task (due to different points of departure) into one final result. It is developed for this thesis and used to extract every fitted parameter result shown. The process consists of two steps, selection and merge.

### Selection

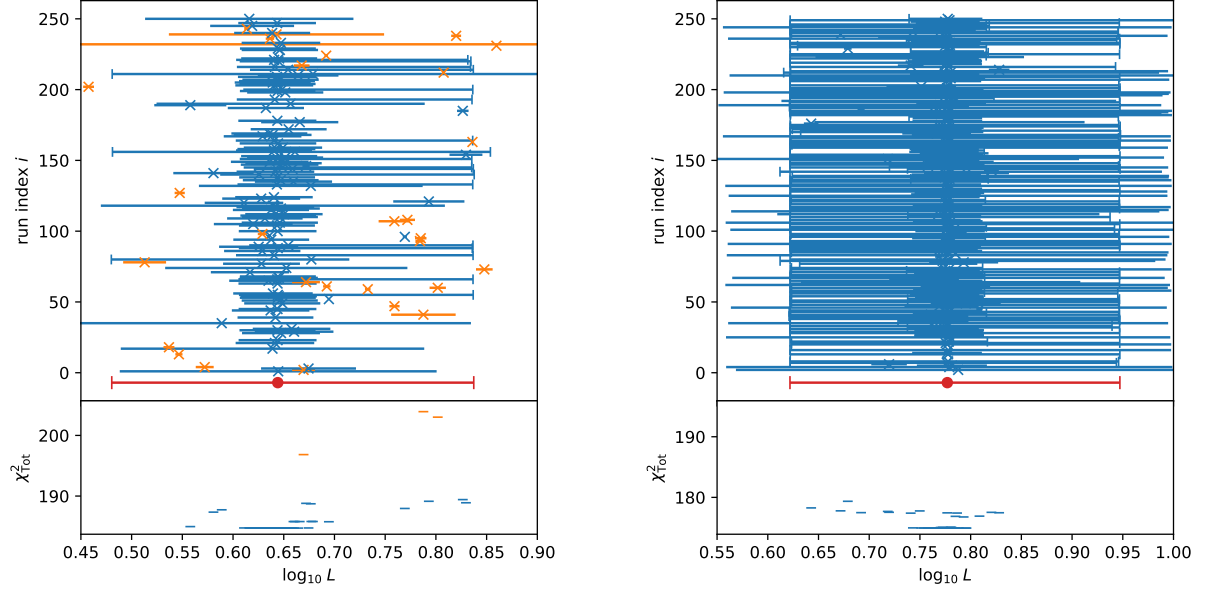
Not all the runs to be launched eventually terminated successfully, leaving only those with index  $i \in \mathcal{R}$  that did.

1. (Optional) Define post-fit parameter limits for any parameter  $\alpha$ , remove  $i$  from  $\mathcal{R}$  if  $\theta_{\alpha}^i$  does not lie inside of them.
2. Get  $\chi_{\min}^2 = \min\{(\chi_{\min}^2)^i, \forall i \in \mathcal{R}\}$ .
3. Remove  $i$  from  $\mathcal{R}$  if  $(\chi_{\min}^2)^i > \chi_{\min}^2 + \Delta\chi^2$ .

The results still remaining in  $\mathcal{R}$  are considered to be a sample of representative *best-fit* results. Setting the cutoff tolerance  $\Delta\chi^2$  to low could to drastically reduce the amount of results and thus their credibility, whereas setting it too high bears the risk of including inconsistent parameter results. In practice  $\Delta\chi^2 = 1$  or 5 is used for simple and complicated fitting tasks respectively. Benchmark-fits that are eventually presented in tables like 4.1 or 3.1 were extracted with a  $\Delta\chi^2 = 0.5$ .

### Merging parameters and uncertainties

For each parameter  $\alpha$ , the best fit result are then calculated using  $\theta_{\alpha}^i, i \in \mathcal{R}$ . The lower edge  $\theta_{l,\alpha} = \theta_{\alpha} - |\sigma_{l,\alpha}|$  is determined as the lower 15% quantile from all the  $\theta_{l,\alpha}^i$ . Similarly, the upper edge  $\theta_{h,\alpha} = \theta_{\alpha} + \sigma_{h,\alpha}$  is the 85% quantile of all  $\theta_{h,\alpha}^i$ . If the respective run  $i$  did not yield an asymmetric minos error  $\sigma_{l,\alpha}^i$  or  $\sigma_{h,\alpha}^i$ , the respective edge  $\theta_{l,\alpha}^i$  or  $\theta_{h,\alpha}^i$  is left out of the quantile calculation. The parameter best-fit estimation  $\theta_{\alpha}$  is then taken as the median of all  $\theta_{\alpha}^i$ . An illustration of this value merging is given in the figures B.1a and B.1b, whereas the first one is a slightly ambiguous and the second one a very explicit case. In addition, the mean minos uncertainty  $\sigma_{m,\alpha} = 0.5 \cdot [\theta_{h,\alpha} - \theta_{l,\alpha}] = 0.5 \cdot [\sigma_{h,\alpha} + |\sigma_{l,\alpha}|]$  is determined for every parameter.



(a) Fitting  $^{10}\text{Be}/^9\text{Be}$ , Li/C and B/C to the QUAINT model in 250 attempts and  $\chi^2$  tolerance of 5.

(b) Fitting  $^{10}\text{Be}/\text{Be}$ , Li/C and B/C to the SLIM model in 250 attempts and  $\chi^2$  tolerance of 5.

**Figure B.1:** Overview of all the individual fit results for the parameter  $\log_{10}(L)$ . Blue bars correspond to runs whose  $\chi^2$  is within the tolerance range above the global best  $\chi^2$ , whereas orange correspond to the other runs. The top panels show the respective fit values on the x-axis with its error-bars. The crosses are the parameter best-fit estimation and endcaps indicate a valid MINOS error bound estimation. Error-bars without endcap are migrad error estimations, that do not contribute to the global parameter bound estimations. The red point and error bar is the merged parameter result and error estimation. The lower panel shows the  $\chi^2$  value for the respective best-fit values in the upper panel.

### Covariance and correlation matrices

At this point, a further sub-selection is done with  $\mathcal{C} = (i \in \mathcal{R} \mid (\chi_{\min}^2)^i > \chi_{\min}^2 + 0.01)$ , leaving only the very best fit results. The covariance matrix elements are then averaged as

$$C_{kl} = \overline{C_{kl}^i} \quad i \in \mathcal{C}. \quad (\text{B.1})$$

From  $C$ , the correlation matrix  $A$  is calculated. Because the migrad errors in the covariance matrix are suspected to be less realistic than the minos errors, a minos-covariance matrix  $M$  is reconstructed with the use of the mean minos error vector  $\vec{\sigma}_m$  via

$$M_{kl} = A_{kl} \cdot \sigma_{m,k} \sigma_{m,l}. \quad (\text{B.2})$$

This artificial covariance matrix  $M$  is used instead of  $C$  if possible. An example for this is sampling parameters in order to determine model uncertainties.

## B.3 Brief introduction to Markov Chain Monte Carlo (MCMC)

The objective of MCMC is to sample a probability distribution  $p(\theta)$  of the model parameters. This is done in iterations, creating a chain of data points  $\theta_i$ , each depending on its ancestor  $\theta_{i-1}$ .  $\theta_i \equiv \vec{\theta}_i$  can either be a scalar or vector with  $p(\theta)$  being a multidimensional PDF. If the Markov-Chain is sufficiently long and contains enough elements, it can in reverse be taken as

an estimate of  $p(\theta)$ . The advantage of MCMC is that even without an analytical definition of  $p(\theta)$ , a sample and thus an estimation can be drawn. More details about MCMC can be found in [Putze et al. \(2009\)](#).

### B.3.1 Sampling a Markov Chain

It is started with a point of departure (a vector in the parameter space to be sampled)  $\theta_0$  as the first element of the Markov Chain. One proceeds then for the  $i^{\text{th}}$  step like this:

1. Sample  $\theta_{\text{probe}}$  from the sampling function using  $\theta_{i-1}$
2. Calculate  $p(\theta_{\text{probe}}), \pi(\theta_{\text{probe}})$
3.  $p_{\text{accept}} = \min(1, \frac{p(\theta_{\text{probe}})\pi(\theta_{\text{probe}})}{p(\theta_{i-1})\pi(\theta_{i-1})})$
4.  $\theta_i = \begin{cases} \theta_{\text{probe}} & \text{with probability } p_{\text{accept}} \\ \theta_{i-1} & \text{otherwise} \end{cases}$
5. proceed with step 1. for  $i \rightarrow i + 1$

This procedure has no limit in length and is similar for any MCMC technique. The decisive components are the posterior probability  $p(\theta)$ , its prior  $\pi$  and the sampling function. The latter define the performance and the outcome of the chain.

### The posterior probability for data analysis

In the case of data-analysis, the posterior probability is not previously known and one can not take  $p(\theta)$  from a discrete, well-defined, function. Instead it is the object of interest. The definition can be obtained from the Bayes Theorem:

$$p(\theta|\text{data}) = \frac{p(\text{data}|\theta) \pi(\theta)}{\pi(\text{data})} \quad (\text{B.3})$$

$$\frac{p(\theta_{\text{probe}}|\text{data})}{p(\theta_{i-1}|\text{data})} = \frac{p(\text{data}|\theta_{\text{probe}}) \pi(\theta_{\text{probe}}|\theta_{i-1})}{p(\text{data}|\theta_{i-1}) \pi(\theta_{i-1}|\theta_{\text{probe}})} \quad (\text{B.4})$$

$$p(\text{data}|\theta) = \exp(-0.5 \cdot \chi^2(\theta)) \quad (\text{B.5})$$

$\pi(\theta)$  is the prior probability for the parameter and  $\pi(\text{data})$  the one for the data, whereas the latter vanishes anyways. The  $p(\text{data}|\theta)$  is defined only by the  $\chi^2$ . In most cases  $\pi(\theta)$  is either 1 or symmetric, which would make the second fraction of priors vanish in equation (B.4), leading to

$$\frac{p(\theta_{\text{probe}}|\text{data})}{p(\theta_{i-1}|\text{data})} = \exp\{0.5 [\chi^2(\theta_{i-1}) - \chi^2(\theta_{\text{probe}})]\}. \quad (\text{B.6})$$

### Sampling functions

The sampling function, or rather algorithm, proposes a  $\theta_{\text{probe}}$  with respect to a given ancestor  $\theta_{i-1}$ . The outcome of such a sampler shall be as random as possible but on the other hand somewhat corresponding to the acceptance region (which will be the posterior distribution). In this thesis two totally different sampling techniques are tried with Metropolis Hastings and Hamiltonian MC being very simply and quite sophisticated respectively.

### Correlation length

The value of each sample vector is calculated using the previous vector in the chain as an input in a random process as explained above. It is even likely for them to be identical. In a bad case, the posterior probability for the new point is strongly constrained by the previous point, making the chain not explore the whole acceptance region. But this relation of each point to another contradicts the idea of having independent stochastic samples.

A mean to measure the correlation within the final chain is the mean correlation length  $l$ . In the uncorrelated case it is  $l = 1$  and  $l > 1$  if not. For any parameter  $\alpha$  it is defined as

$$l^\alpha = \min(j \mid c_j^\alpha < 0.5) \quad \text{with} \quad c_j^\alpha = \frac{E[\theta_i^\alpha \theta_{j+i}^\alpha] - (E[\theta_i^\alpha])^2}{E[(\theta_i^\alpha)^2]}.$$

Whereas the expectation value  $E$  is the statistic mean in this case. Different parameters  $\theta^\alpha$  may have different lengths  $l^\alpha$ , so that one needs to take the highest value of all the parameters  $l = \max(l^\alpha)$ . If  $l$  is greater than 1, the chain needs to be thinned out by only picking each  $l^{\text{th}}$  element. This way, the remaining elements are supposed to be mostly uncorrelated. A preliminary test of the correlation length is represented in table 2.2.

### B.3.2 Metropolis Hastings

The Metropolis Hastings algorithm samples in each iteration step a proposed parameter  $\theta_{\text{probe}}$  from a single probability distribution. That is why it is also sometimes referred to as *Random Walk* approach. In this thesis the new parameter  $\theta_{\text{probe}}$  is drawn from a multidimensional normal distribution, centered around  $\theta_{i-1}$ . The scale or covariance matrix is defined by the user. A more sophisticated version would be to calculate the covariance matrix from previous results and update it during the run.

Hence  $\pi(\theta) = 1$ , the probability fraction follows (B.6) yielding

$$p_{\text{accept}} = \min(1, \exp\{0.5 [\chi^2(\theta_{i-1}) - \chi^2(\theta_{\text{probe}})]\}) \quad (\text{B.7})$$

### B.3.3 Hamiltonian Monte Carlo

The Hamiltonian MC algorithm, as described in Neal (2012), is much more complex than the Metropolis one, but uses the same basic concept described in B.3.1. The sampling is a *deplacement* from the  $\theta_{i-1}$  to  $\theta_{\text{probe}}$  in multiple sub-steps.

The idea is to simulate the physical movement of an object in the parameter-hyper-sphere with a numerical step based on the solution of Hamiltonian differential equations. On its *journey* through the valley of minimal  $\chi^2$ , the algorithm will accept and reject parameters *along the way*.

**Displacement** At the beginning, a momentum vector  $k_{i-1}$  is drawn from a normal distribution with the dimensions of  $\theta_{i-1}$ . Starting with this momentum  $k_{i-1}$  and the position  $\theta_{i-1}$  the movement is calculated as a series of numerical steps, solving the Hamiltonian mechanics in iterations. In each step the next position  $\theta_{\text{current}}$  is calculated and the momentum updated to  $k_{\text{current}}$ . The therefore necessary gradient  $\nabla \chi^2(\theta)$  is determined by evaluating the  $\chi^2$  at points directly next to  $\theta$  in each dimension.

After  $N_i$  of these steps, the resulting point is tested with  $\theta_{\text{probe}} = \theta_{\text{current}}$ . No matter the outcome of the acceptance, the algorithm continues its *movement* from  $\theta_{\text{current}}$  to compute an amount of  $N_{i+1}$  steps to the next proposed position. In this particular implementation used,  $N_i$



is taken randomly between 0 and  $N^{\max}$ .  $N^{\max} = \mathcal{O}(10 - 100)$  is either user-set or auto-adjusted. If it is too low, the *voyage* between  $\theta_i$  and  $\theta_{i+1}$  did not cross a significant part of the valley and the two following chain elements will be correlated. If it is too high, calculation time is wasted.

**Acceptance** The acceptance is also different than the one of Metropolis Hastings, as it contains a probability-factor for the momentum vector  $k$ , which can be expressed as a prior  $\pi(\theta_i) := p(k_i)$ . Note that the latter is not dependent on the values of  $\theta_i$  but on  $k_i$ , as explained below. After the displacement from  $\theta_{i-1}$  to  $\theta_{\text{current}}$ , the accepting-rejecting probability can be calculated with

$$p(\theta) = \exp \{-0.5 U(\theta)\} \quad p(k) = \exp \{-0.5 K(k)\} \quad (\text{B.8})$$

$$\text{with } K(k) = k^T M^{-1} k, \quad U(\theta) \equiv \chi^2(\theta) \quad (\text{B.9})$$

$$\rightarrow p_{\text{accept}} = \min \left[ 1, \frac{p(\theta_{\text{probe}})p(k_{\text{probe}})}{p(\theta_{i-1})p(k_{i-1})} \right] \quad (\text{B.10})$$

$$= \min [1, \exp \{ 0.5 \cdot [U(\theta_{i-1}) - U(\theta_{\text{probe}}) + K(k_{i-1}) - K(k_{\text{probe}})] \} ]. \quad (\text{B.11})$$

The term  $K(p)$  corresponds to the kinetic energy and  $U + K = H$  are the total energy.  $M$  is the mass matrix. Just like in the Metropolis algorithm, the new element  $\theta_{\text{probe}}$  is accepted with the probability  $p_{\text{accept}}$  as  $\theta_i$ . Because  $p(k_{i-1})$  does not depend on the values of  $\theta$ , the resulting distribution of accepted  $\theta$  will resemble  $p(\theta)$ .

### Costs and Benefits

The execution of one of these sub-steps during the displacement needs one evaluation of the  $\chi^2$  in order to get the corresponding potential  $U$ , and  $\dim(\theta)$  more evaluations in order to get the gradient of it. Between two elements in the Markov Chain lie thus  $N_i \cdot (\dim(\theta) + 1)$  calls of the  $\chi^2$  function.

On the other hand, the resulting propositional  $\theta_{\text{probe}}$  after each step depends directly on the  $\chi^2$ . With a sufficient  $N^{\max}$  the Markov Chain elements are fairly uncorrelated to each other.



## Rescaling parameters with L

Some parameters have a strong dependency to the halo size  $L$ , notably  $K_0$  and  $V_a$ . These degeneracies can partly be found in the model definitions like (1.11) and (1.5). When setting  $L$  to a fixed value and fitting the others, this does not need to be addressed, as every parameter re-adjusts itself during the fitting process. In reverse however, when changing the value of  $L$ , the other transport parameters must be re-scaled in order to stay in the best-fit regime. A formula is thus needed to modify the best-fit transport parameters, when using them for another  $L$  than they were fitted with.

For this, the flux B/C was minimized against AMS-02 data for several different, fixed values of  $L$  between 1 and 20 kpc for all three models BIG, SLIM, and QUAINT. For each parameter  $P$  the function

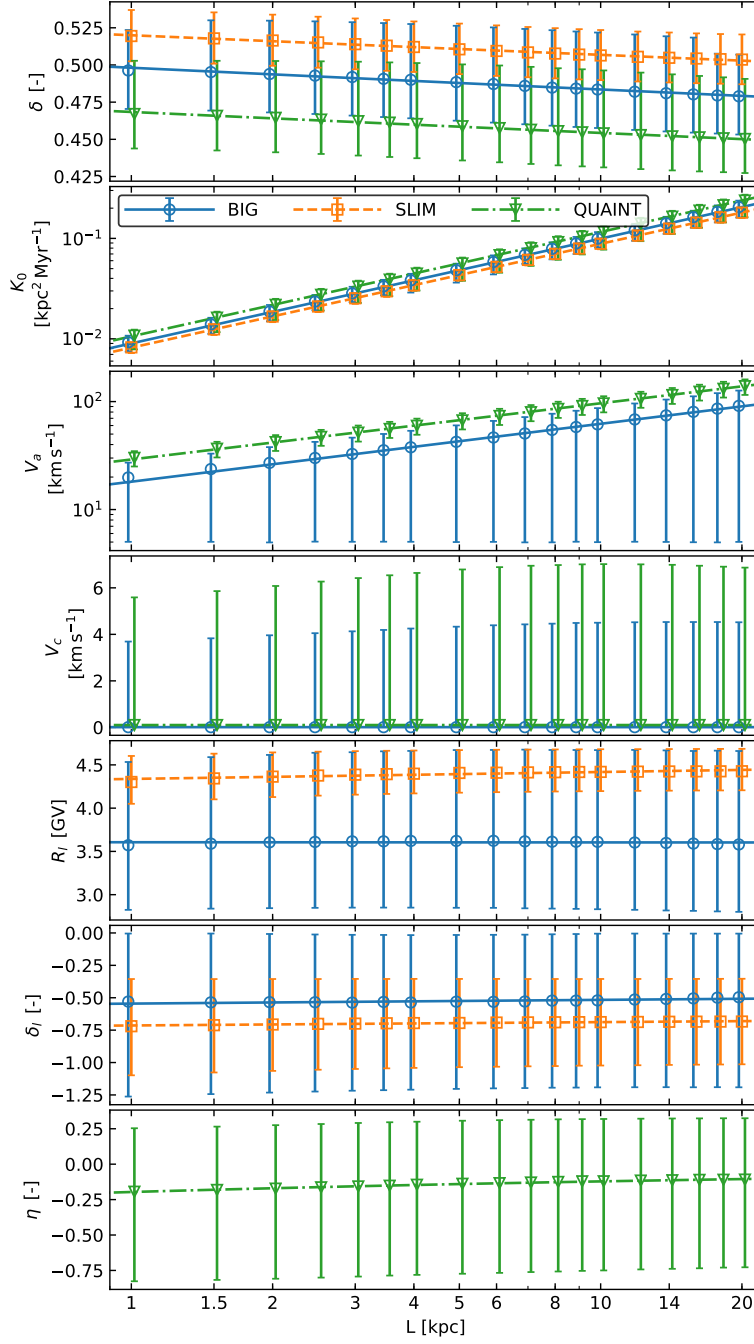
$$P(L) = A \cdot \left( \frac{L}{5 \text{ kpc}} \right)^B \quad (\text{C.1})$$

was adjusted to the respective best-fit parameters.  $A$  is the parameter value for  $L = 5 \text{ kpc}$  and  $B$  the power-law dependency of  $L$  for the respective parameter. The fit results and the fitted function are shown in figure C.1.

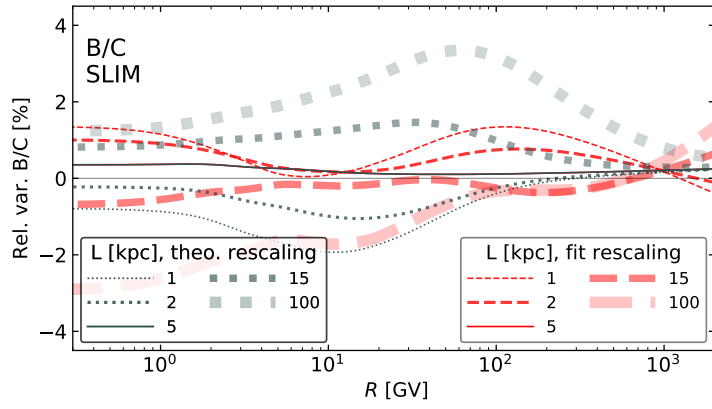
Parameter [Unit]	Coef.	Expect.	BIG	SLIM	QUAINT
$\delta$ [-]	A		0.488	0.511	0.458
	B	?	-0.013	-0.011	-0.013
$K_0$ [kpc <sup>2</sup> Myr <sup>-1</sup> ]	A		0.048	0.043	0.056
	B	1	1.043	1.034	1.040
$V_a$ [km s <sup>-1</sup> ]	A		42.94	n/a	67.24
	B	0.5	0.536	n/a	0.520
$V_c$ [km s <sup>-1</sup> ]	A		0.010	n/a	0.100
	B	?	-0.008	n/a	0.000
$R_l$ [GV]	A		3.605	4.393	n/a
	B	?	-0.000	0.008	n/a
$\delta_l$ [-]	A		-0.525	-0.696	n/a
	B	?	-0.024	-0.016	n/a
$\eta$ [-]	A		n/a	n/a	-0.140
	B	?	n/a	n/a	-0.208

**Table C.1:** Coefficients for parameter rescaling for formula (C.1), derived from fitting it to the respective values shown in figure C.1. The third column are the naive expectation values, using their respective relations in (1.11) and (1.5). A '?' marks that no simple, direct dependency between  $L$  and the respective parameter could be found in the formula, suspecting  $\sim 0$ .

As the table C.1 shows, there are two different choices for the parameter B for each parameter when re-scaling it by  $L$ , the Naive expectation and the value fitted by B/C. The latter is the more realistic case that is used if possible. For predicting the impact on B/C however, these can not be used, as they are strongly biased. For this reason, the parameter are rescaled with the expectation value of B when studying B/C. A comparison of both cases is given in figure C.2.



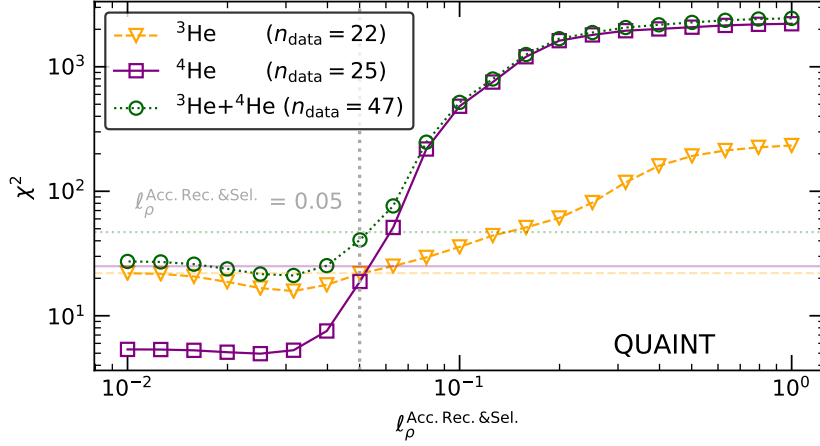
**Figure C.1:** Overview of the MINUIT2 values for the models BIG, SLIM, and QUAINT for different, fixed halo-sizes  $L$ . The points and error bars are the parameter values and MINOS uncertainty estimations. The curves correspond to the function (C.1) fitted to the respective parameter results, yielding the coefficients in table C.1. Note that the second and third panel for  $K_0$  and  $V_a$  are in logarithmic y-scale.



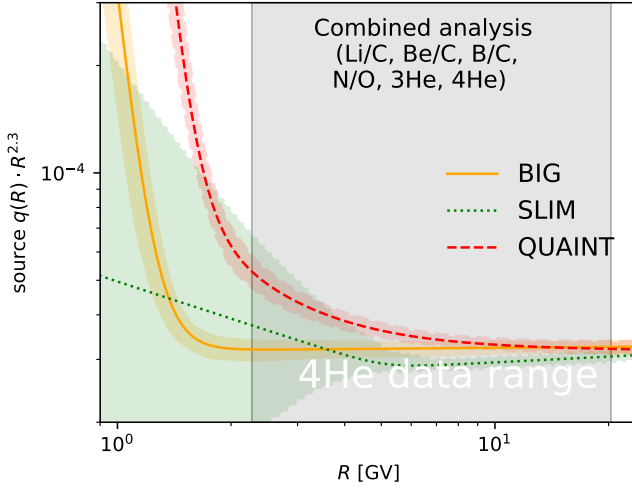
**Figure C.2:** Flux ratio variations for B/C in the SLIM model when varying  $L$  and rescaling the other parameters using formula (C.1). In gray are the variations, based on the expected dependency according to table C.1, whereas the variations based on the fitted dependency are in red.



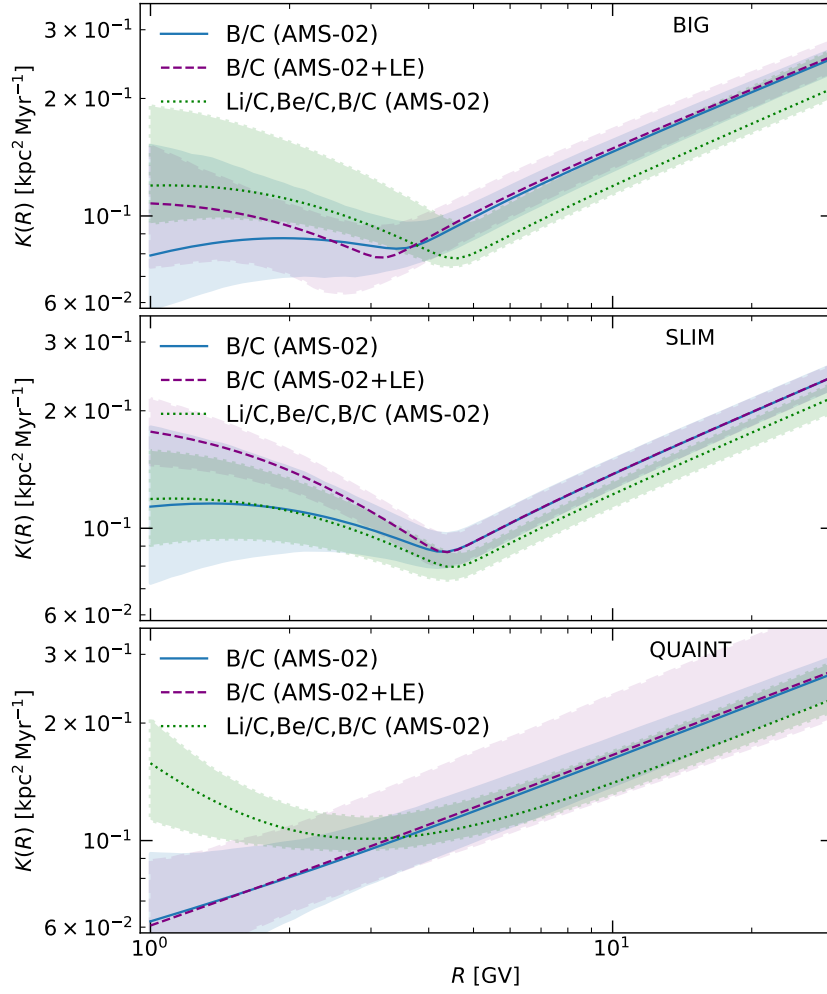
## Other additional figures



**Figure D.1:**  $\chi^2$  values for  $^3\text{He}$  and  $^4\text{He}$  AMS-02 data in the QUAINT model when using different covariance matrices. The model parameters are constant and were fitted to LiBeB/C,N/O, $^3\text{He}$ , $^4\text{He}$ , using  $l_q^{\text{acc.}} = 0.05$ , as described in section 3.4. This value for the correlation length is indicated as the vertical gray dotted line. The horizontal dashed lines at 22, 25, and 47 indicate the  $n_{\text{dof}}$  for  $^3\text{He}$ ,  $^4\text{He}$ , and their sum respectively.

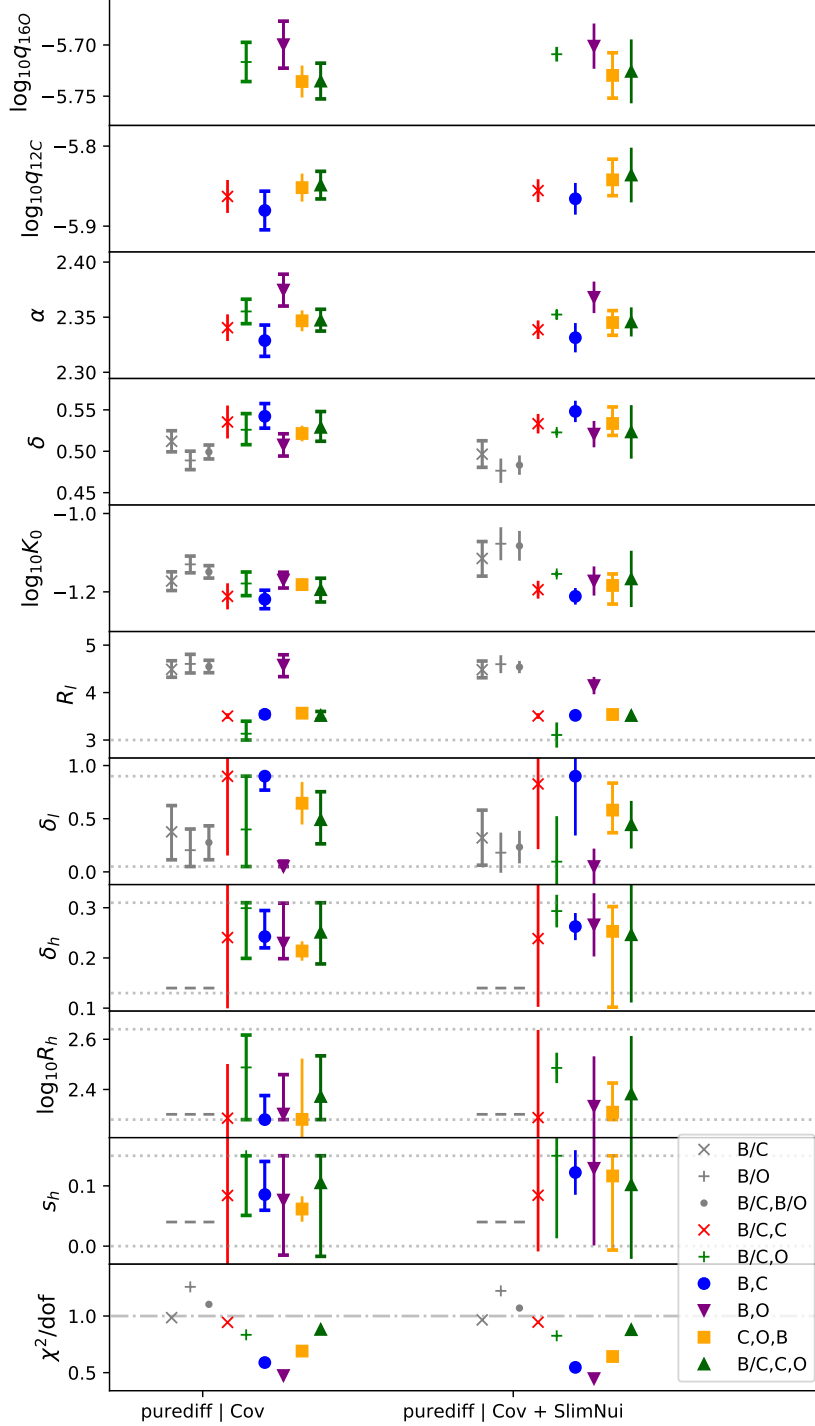


**Figure D.2:** Source injection spectrum for  $^4\text{He}$  with the post fit source parameter values and uncertainties using the spectral formula (1.3). The fit included to Li/C, Be/C, B/C, N/O,  $^4\text{He}$ , and  $^3\text{He}$  AMS-02 data. The gray rectangle illustrates the range between the lowest and highest rigidity bin center in the  $^4\text{He}$  data set.

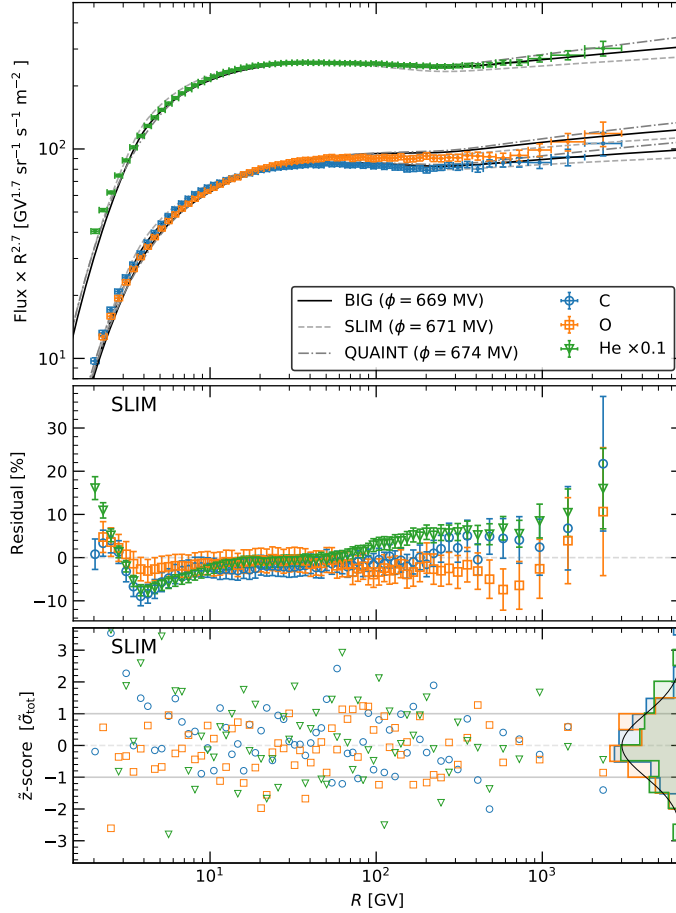


**Figure D.3:** Plot of the diffusion tensor  $K(R)$  and its  $1\sigma$  quantil for the three models BIG, SLIM, and QUAINT fitted to three different data combinations. The respective parameter values and covariance matrices for sampling the models stem from the analyses done in sections 3.2 and 3.3. The ratio of nuclei mass to charge  $A/Z = 2$  is taken as an example here.





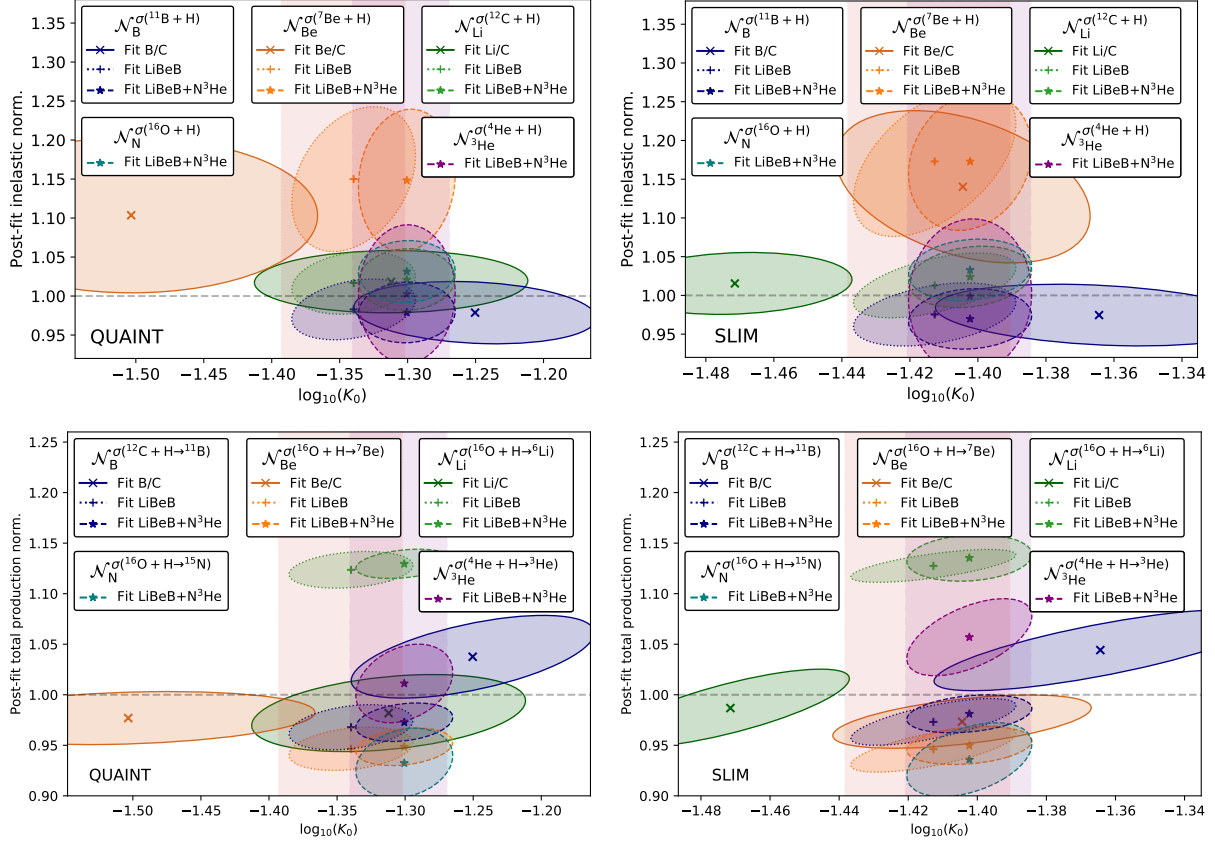
**Figure D.4:** Best-fit parameter results for the free parameters when fitting different flux and flux ratio combinations of Boron, Carbon, and Oxygen AMS-02 data. The *purediff* model used is an older version of the SLIM model discussed in this thesis. Note that  $\delta_l$  in this plot corresponds to  $-\delta_l$  in this thesis. The fits presented in the first column did not involve any cross section nuisance parameters, whereas the fits in the second column involved the minimal nuisance parameter set for Boron, presented in table 2.1. Gray points correspond to combinations of flux ratios only. The "-" markers in the three bottom panels show that these values were fixed and not fitted. The galactic halo size  $L$  was fixed at 10 kpc in that case.



**Figure D.5:** *Top panel:* AMS-02 data for Carbon, Oxygen and Helium. The lines are the model predictions for the three species for the benchmark models BIG, SLIM, and QUAINT fitted to LiBeB/C (see section 3.3). All the data and model values and errors for Helium are multiplied by 0.1 for visibility.

*Middle panel:* Residuals of the data points in the top panel to their respective SLIM model prediction in the top panel.

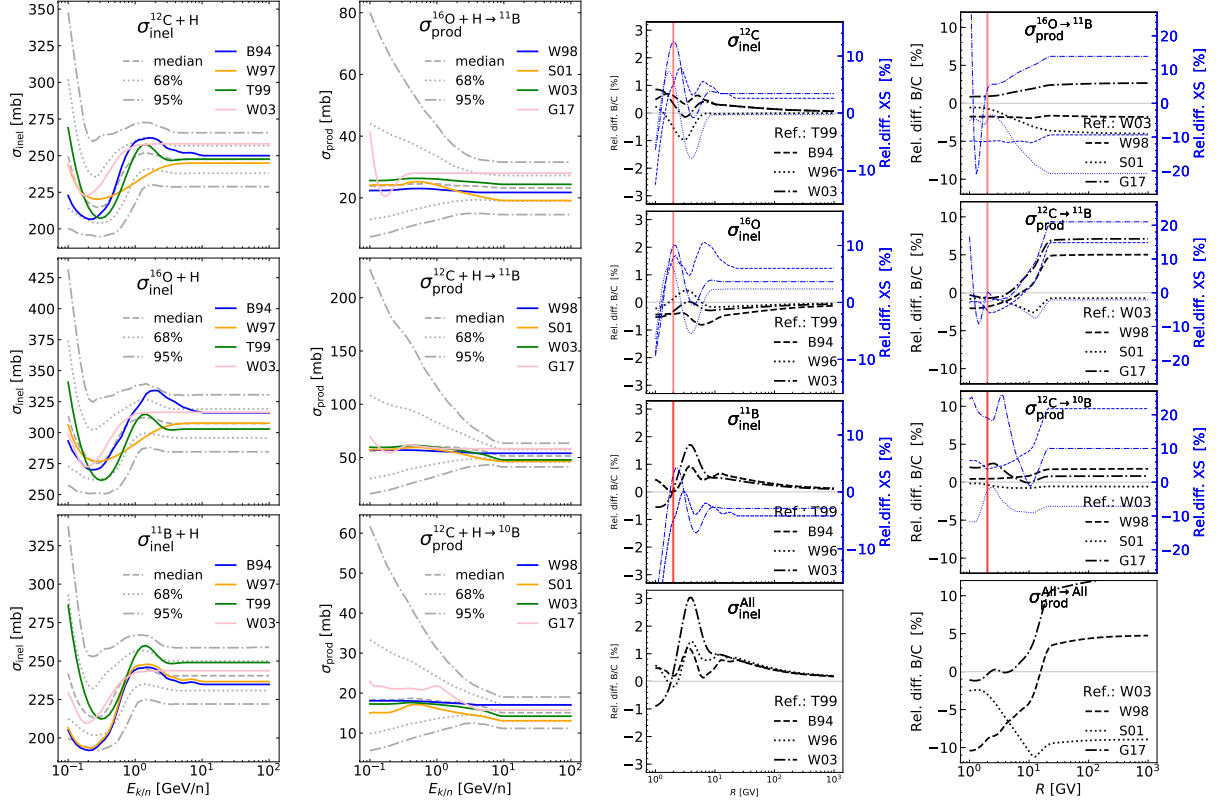
*Bottom panel points:*  $\tilde{z}$ -score for the data points in the top panel to their respective SLIM model prediction in the top panel. See caption of figure 3.4 for details. *Bottom right histograms:* histogram of the  $\tilde{z}$ -score values shown in this panel for all three species in their respective color. For reference, the normal distribution is drawn as a black solid line.



**Figure D.6:** Correlation and uncertainty illustration for cross section nuisance parameter and  $K_0$  fit results in the QUAINT and SLIM model (left and right column respectively). The vertical and horizontal sizes of the ellipses are the mean MINOS error estimation (see B.2) of the cross section nuisance parameters and  $\log_{10}(K_0)$  respectively. The ellipses shape is calculated from the correlation estimation of MINUIT2 (extracted from the covariance matrix). Each color corresponds to a different cross section and each contour line style to a different combination of fitted species. LiBeB is short for Li/C, Be/C, and B/C fitted simultaneously. LiBeB+N<sup>3</sup>He is short for Li/C, Be/C, B/C, N/O, <sup>4</sup>He, and <sup>3</sup>He fitted simultaneously. Note that every species shows a different respective nuisance parameter in its respective color, which can be found in table 2.1. The upper row shows the normalization of the inelastic cross section on the vertical axes. In an analogue manner the lower row shows the total production and its error  $P = 1 + (\mathcal{N} - 1) \cdot c, \sigma_P = c \cdot \sigma_{\mathcal{N}}$  for each species, where  $N$  is the normalization of the production cross section result with mean MINOS error  $\sigma_{\mathcal{N}}$  and  $c$  the contribution of the respective reaction. Which production cross section is used and what contribution is linked to it can be found in table 2.1. The vertical pinkish bands with dotted and slashed contours show the asymmetric MINOS uncertainty range for  $\log_{10}(K_0)$  for the LiBeB and LiBeB+N<sup>3</sup>He fit respectively.



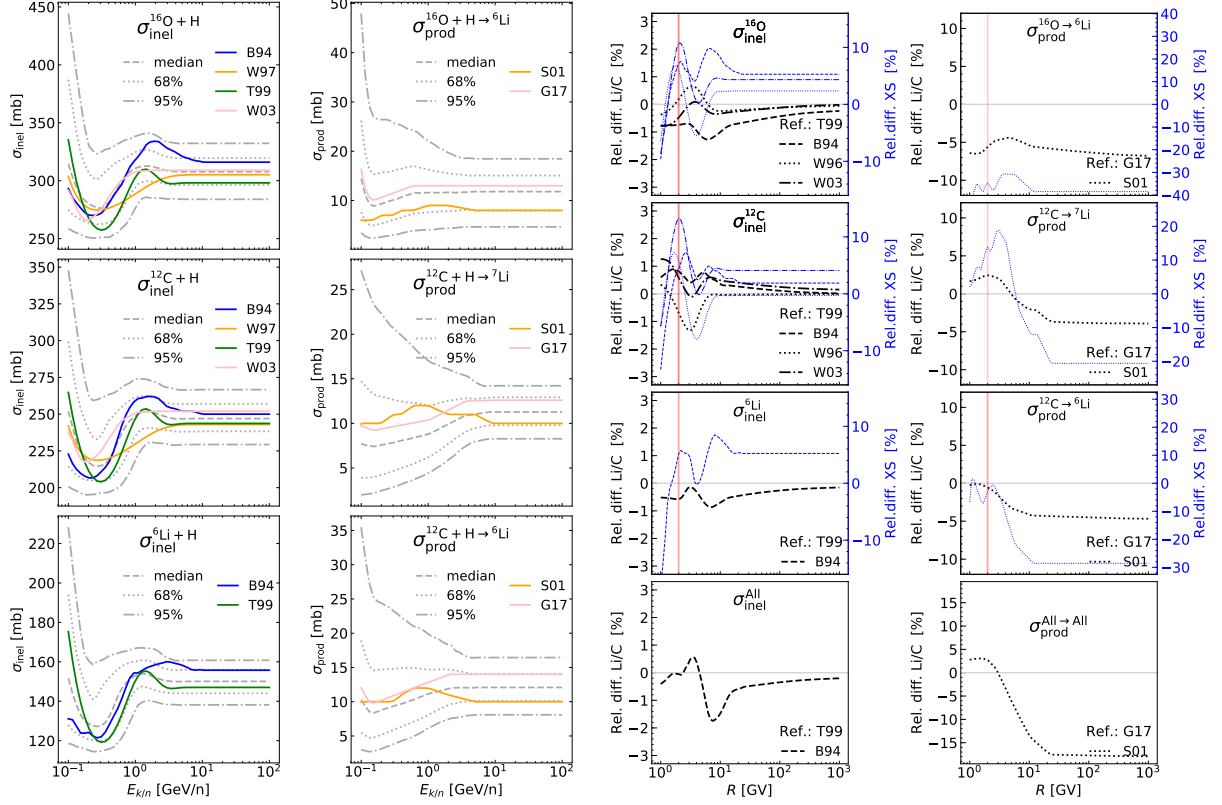
# Cross section nuisance parameter plots



**Figure E.1:** Plots for the nuclear cross section nuisance parameter study for B/C.

*First & second column:* Comparison of existing cross section data for inelastic reactions and production. The gray lines show statistical range when modifying and sampling the reference cross-section using the values from table 2.1.

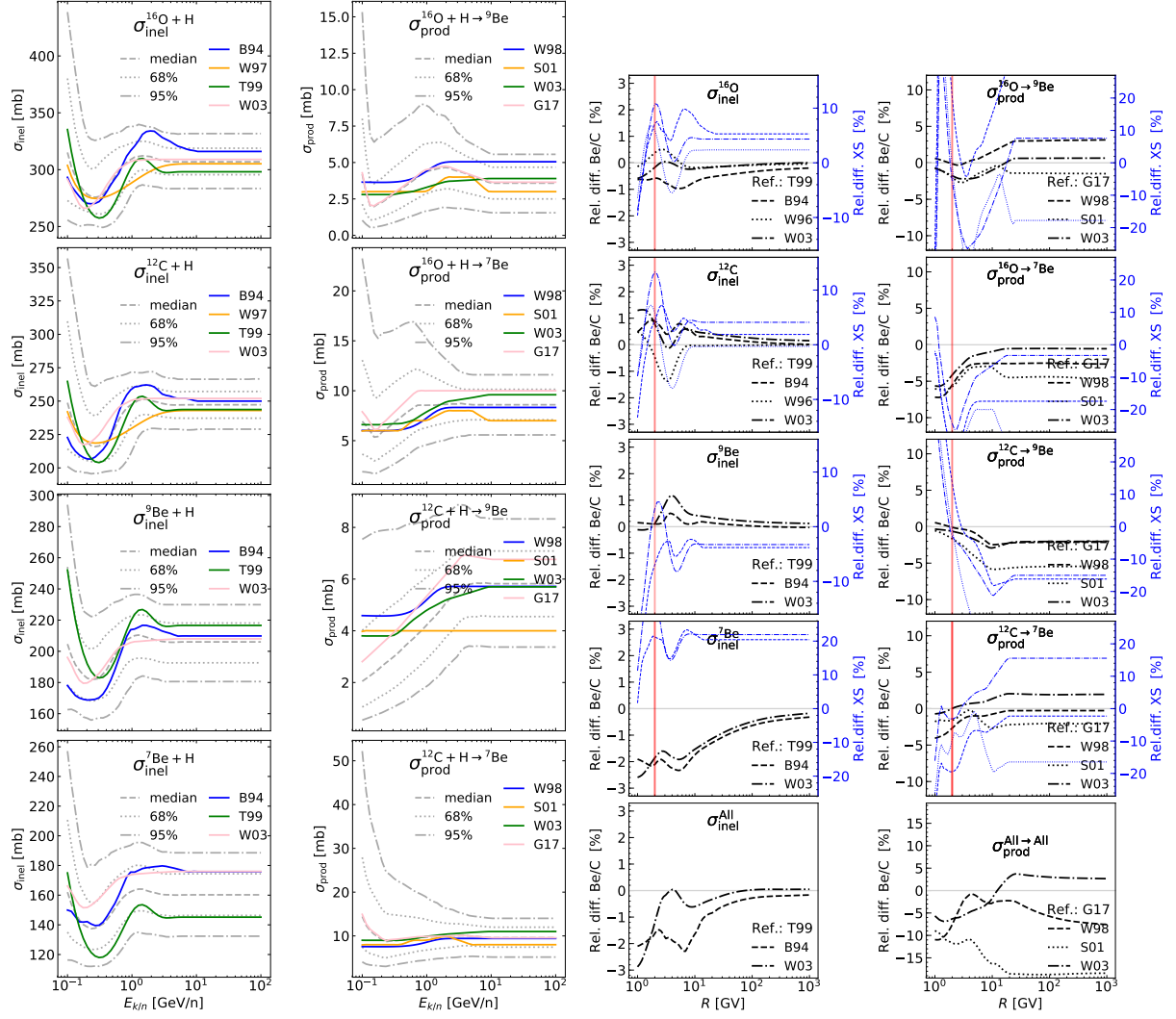
*Third & fourth column:* Relative flux ratio differences in black, when exchanging the reference cross section data for the reaction by another data set. The cross section differences for inelastic reactions and production, relative to their respective references, are shown in blue.



**Figure E.2:** Plots for the nuclear cross section nuisance parameter study for Li/C.

*First & second column:* Comparison of existing cross section data for inelastic reactions and production. The gray lines show statistical range when modifying and sampling the reference cross-section using the values from table 2.1.

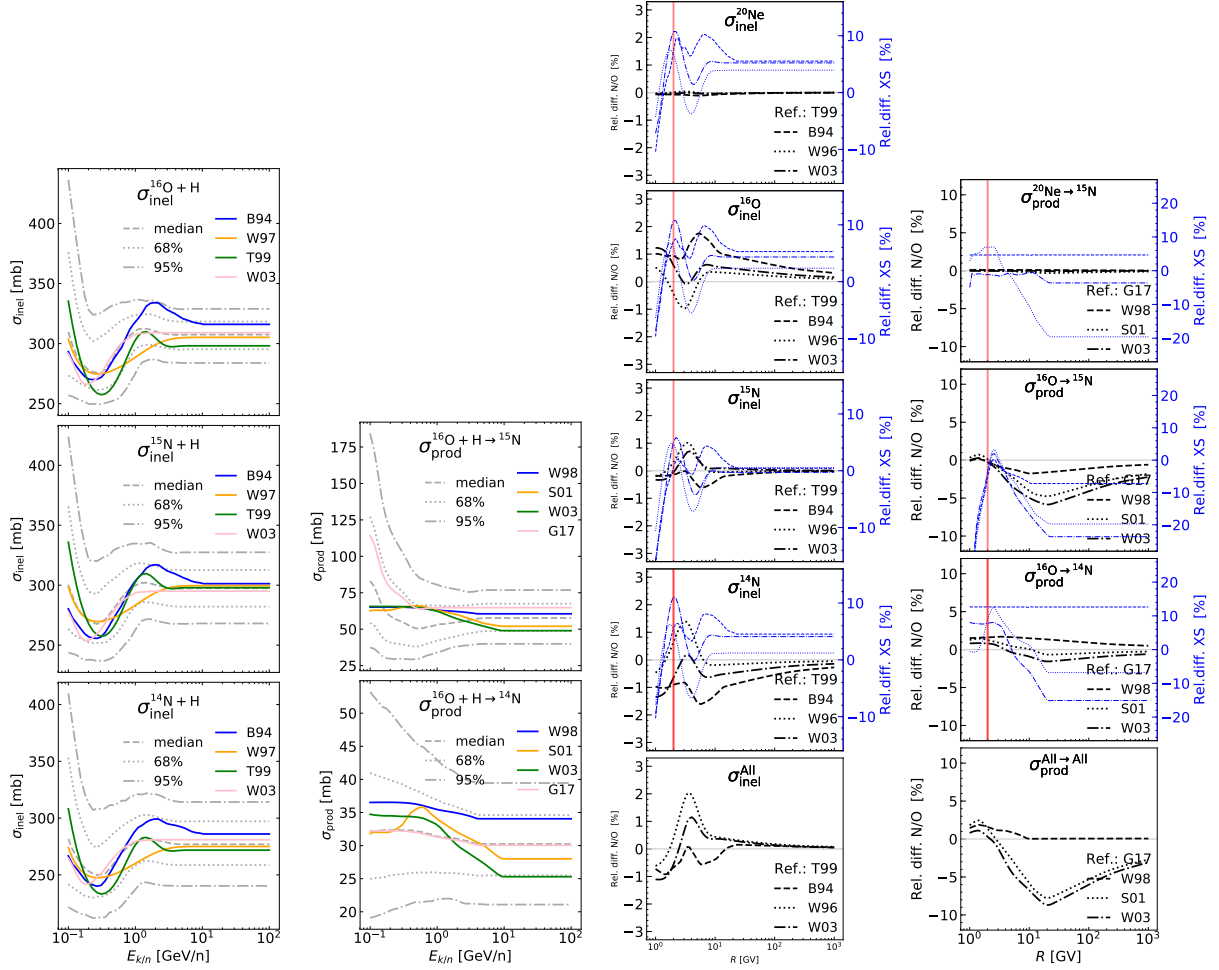
*Third & fourth column:* Relative flux ratio differences in black, when exchanging the reference cross section data for the reaction by another data set. The cross section differences for inelastic reactions and production, relative to their respective references, are shown in blue.



**Figure E.3:** Plots for the nuclear cross section nuisance parameter study for Be/C.

*First & second column:* Comparison of existing cross section data for inelastic reactions and production. The gray lines show statistical range when modifying and sampling the reference cross-section using the values from table 2.1.

*Third & fourth column:* Relative flux ratio differences in black, when exchanging the reference cross section data for the reaction by another data set. The cross section differences for inelastic reactions and production, relative to their respective references, are shown in blue.

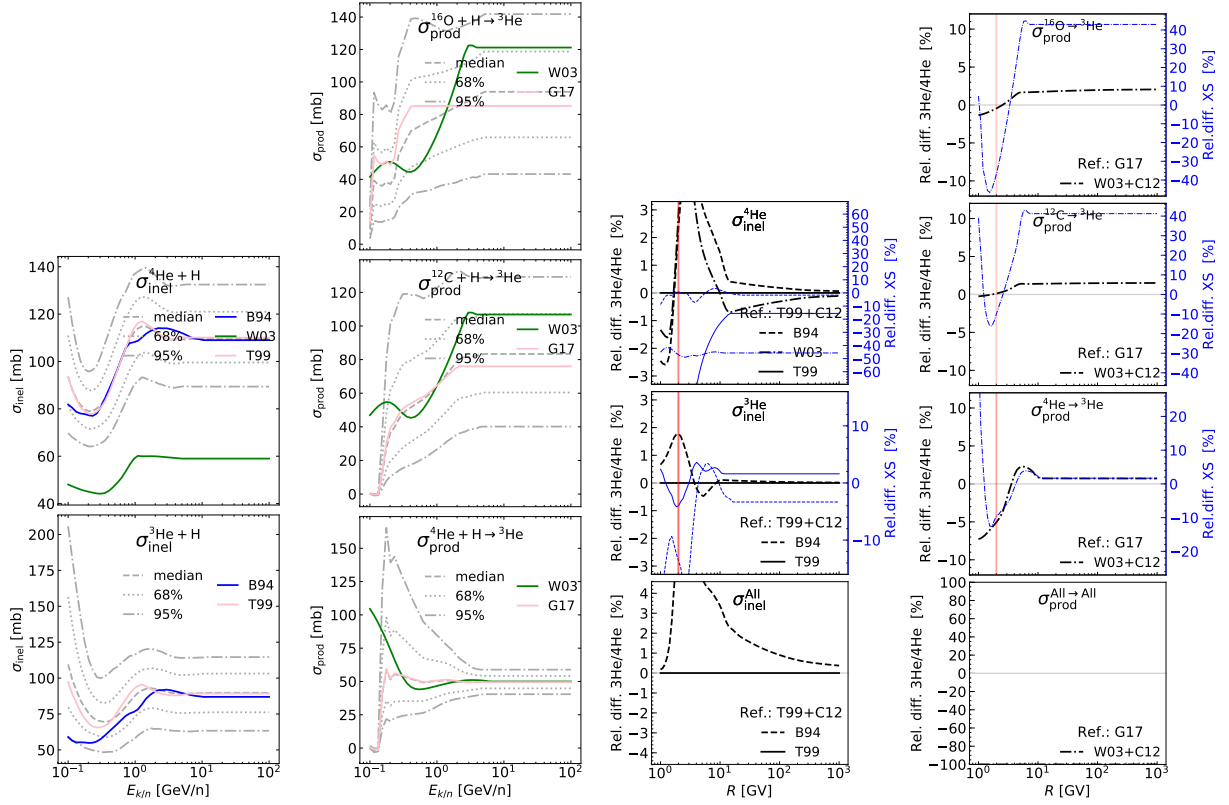


**Figure E.4:** Plots for the nuclear cross section nuisance parameter study for N/O.

*First & second column:* Comparison of existing cross section data for inelastic reactions and production. The gray lines show statistical range when modifying and sampling the reference cross-section using the values from table 2.1.

*Third & fourth column:* Relative flux ratio differences in black, when exchanging the reference cross section data for the reaction by another data set. The cross section differences for inelastic reactions and production, relative to their respective references, are shown in blue.





**Figure E.5:** 3He/4He

**Figure E.6:** Plots for the nuclear cross section nuisance parameter study for  ${}^3\text{He}/{}^4\text{He}$ .

*First & second column:* Comparison of existing cross section data for inelastic reactions and production. The gray lines show statistical range when modifying and sampling the reference cross-section using the values from table 2.1.

*Third & fourth column:* Relative flux ratio differences in black, when exchanging the reference cross section data for the reaction by another data set. The cross section differences for inelastic reactions and production, relative to their respective references, are shown in blue.



---

## Bibliography

- Aguilar, M., Ali Cavasonza, L., Alpat, B., et al. 2018b, *Phys. Rev. Lett.*, 121, 051103
- Aguilar, M., Ali Cavasonza, L., Ambrosi, G., et al. 2016, *Physical Review Letters*, 117, 231102
- Aguilar, M., Ali Cavasonza, L., Ambrosi, G., et al. 2018a, *Phys. Rev. Lett.*, 120, 021101
- Aguilar, M., Ali Cavasonza, L., Ambrosi, G., et al. 2019, *Phys. Rev. Lett.*, 123, 181102
- Barrau, A., Boudoul, G., Donato, F., et al. 2002, *A&A*, 388, 676
- Berezinskii, V. S., Bulanov, S. V., Dogiel, V. A., & Ptuskin, V. S. 1990, *Astrophysics of cosmic rays*
- Boschini, M. J., Torre, S. D., Gervasi, M., et al. 2020, *ApJ*, 889, 167
- Boudaud, M., Genolini, Y., Derome, L., et al. 2019, in *International Cosmic Ray Conference*, Vol. 36, 36th International Cosmic Ray Conference (ICRC2019), 38
- Caballero-Lopez, R. A. & Moraal, H. 2004, *Journal of Geophysical Research (Space Physics)*, 109, A01101
- Connell, J. J. 1998, *ApJ*, 501, L59
- Coste, B., Derome, L., Maurin, D., & Putze, A. 2012, *A&A*, 539, A88
- Derome, L., Maurin, D., Salati, P., et al. 2019, *A&A*, 627, A158
- Evoli, C., Aloisio, R., & Blasi, P. 2019, *Phys. Rev. D*, 99, 103023
- Evoli, C., Morlino, G., Blasi, P., & Aloisio, R. 2020, *Phys. Rev. D*, 101, 023013
- Ferrière, K. M. 2001, *Reviews of Modern Physics*, 73, 1031
- Garcia-Munoz, M., Mason, G. M., & Simpson, J. A. 1977, *ApJ*, 217, 859
- Garcia-Munoz, M., Simpson, J. A., & Wefel, J. P. 1981, in *International Cosmic Ray Conference*, Vol. 2, International Cosmic Ray Conference, 72–75
- Génolini, Y., Boudaud, M., Batista, P. I., et al. 2019, *Phys. Rev. D*, 99, 123028
- Génolini, Y., Maurin, D., Moskalenko, I. V., & Unger, M. 2018, *Phys. Rev. C*, 98, 034611
- Gleeson, L. J. & Axford, W. I. 1968, *ApJ*, 154, 1011
- Grenier, I. A., Black, J. H., & Strong, A. W. 2015, *Annual Review of Astronomy and Astrophysics*, 53, 199

- Hams, T., Barbier, L. M., Bremerich, M., et al. 2004, *ApJ*, 611, 892
- James, F. & Winkler, M. 2018, *Minuit2 manual*
- Jones, F. C., Lukasiak, A., Ptuskin, V., & Webber, W. 2001, *ApJ*, 547, 264
- Lavalle, J., Maurin, D., & Putze, A. 2014, *Phys. Rev. D*, 90, 081301
- Lavalle, J. & Salati, P. 2012, *Comptes Rendus Physique*, 13, 740
- Lukasiak, A. 1999, in *International Cosmic Ray Conference*, Vol. 3, *International Cosmic Ray Conference*, 41
- Mathews, J. H. & Fink, K. K. 2004, *Numerical Methods Using Matlab* (Pearson)
- Maurin, D. 2001, *Theses, Université de Savoie*
- Maurin, D. 2018, *ArXiv e-prints 1807.02968*
- Maurin, D., Donato, F., Taillet, R., & Salati, P. 2001, *ApJ*, 555, 585
- Maurin, D., Taillet, R., & Combet, C. 2006, *Transport of exotic anti-nuclei: I- Fast formulae for antiproton fluxes*
- Menn, W., Bogomolov, E. A., Simon, M., Vasilyev, G., & et al. 2018, *ApJ*, 862, 141
- Moskalenko, I. V. & Mashnik, S. G. 2003, *ICRC*, 4, 1969
- Moskalenko, I. V., Mashnik, S. G., & Strong, A. W. 2001, *ICRC*, 5, 1836
- Neal, R. M. 2012, *arXiv e-prints, arXiv:1206.1901*
- Nocedal, J. & Wright, S. J. 1999, *Numerical Optimization* (New York, NY, USA: Springer)
- Ptuskin, V. S., Moskalenko, I. V., Jones, F. C., Strong, A. W., & Zirakashvili, V. N. 2006, *ApJ*, 642, 902
- Putze, A., Derome, L., & Maurin, D. 2010, *A&A*, 516, A66
- Putze, A., Derome, L., Maurin, D., Perotto, L., & Taillet, R. 2009, *A&A*, 497, 991
- Tomassetti, N. 2015, *Phys. Rev. C*, 92, 045808
- Vladimirov, A. E., Jóhannesson, G., Moskalenko, I. V., & Porter, T. A. 2012, *ApJ*, 752, 68
- Webber, W. R., Kish, J. C., Rockstroh, J. M., et al. 1998a, *ApJ*, 508, 940
- Webber, W. R., Kish, J. C., Rockstroh, J. M., et al. 1998b, *ApJ*, 508, 949
- Webber, W. R., Soutoul, A., Kish, J. C., & Rockstroh, J. M. 2003, *ApJS*, 144, 153
- Webber, W. R., Soutoul, A., Kish, J. C., et al. 1998c, *Phys. Rev. C*, 58, 3539
- Weinrich, N., Génolini, Y., Boudaud, M., Derome, L., & Maurin, D. 2020, *arXiv e-prints, arXiv:2002.11406*
- Weinrich, N. et al. 2020b, *ArXiv Astrophysics e-prints*
- Wiedenbeck, M. E. & Greiner, D. E. 1980, *ApJ*, 239, L139
- Yanasak, N. E., Wiedenbeck, M. E., Mewaldt, R. A., et al. 2001, *ApJ*, 563, 768