The Astrophysical Journal, 959:139 (9pp), 2023 December 20 https://doi.org/10.3847/1538-4357/ad0b09

The Implication of the Helium-to-proton Flux Ratio in Galactic Cosmic Rays

Xiaojian Song and Xi Luo
Shandong Institute of Advanced Technology (SDIAT), 250100 Jinan, Shandong, People's Republic of China; xi.luo@iat.cn
Received 2023 September 20; revised 2023 November 6; accepted 2023 November 7; published 2023 December 15

Abstract

The measurement of daily proton and helium fluxes by AMS-02 shows that the helium-to-proton flux ratio (He/H) is negatively correlated with the solar activity at rigidity 1.7–7 GV. What is the behavior of He/H at rigidity lower than 1.7 GV? How can this phenomenon be described quantitatively and linked to the underlying mechanisms? In this work, based on a theoretical derivation, we find that the slope (S) of ln(He/H) versus ln(H) is a good indicator of the variation of He/H. $S > 0$ means the variation of He/H is anticorrelated with the solar activity, and vice versa. Furthermore, the value of S is quantitatively related to the difference in protons and helium nuclei in the local interstellar spectrum (LIS) and in solar modulation. A 3D time-dependent numerical model is used to reproduce the observed proton and helium data. The results of the simulation show that, with increasing rigidity, the parameter S increases from negative values to positive values at around 1 GV, reaches a maximum value at around 2 GV, and then decreases until it approaches 0. Interestingly, S is equal to 0 at around 1 GV. This means that He/H at this rigidity is almost unchanged with the variation in solar activity. Finally, the expression for S is validated by changing the LIS used in the numerical procedure. This shows that the overall profile of S as a function of rigidity is mainly determined by the mass-to-charge ratio, and the specific value of S is strongly affected by the difference in the LIS between protons and helium nuclei.

Unified Astronomy Thesaurus concepts: Galactic cosmic rays (567); Heliosphere (711); Solar activity (1475)

1. Introduction

Galactic cosmic rays (GCRs) originate from high-energy astrophysical processes (such as in supernova remnants) and possibly exotic processes (such as dark matter annihilation/decays) in the Galaxy. Their flux outside the heliosphere, usually called the local interstellar spectrum (LIS), remains constant over long timescales, because the Sun's motion relative to the sources in the Milky Way is too slow to have a noticeable effect (Büsching et al. 2005). Each GCR species has its unique LIS, and can be classified into three groups based on the power-law index of the LIS at the high-energy ($\gtrsim$100 GV) end (Aguilar et al. 2021a), which characterizes their different origins. With the help of PAMELA, AMS-02, and Voyager, the LISs at high ($\gtrsim$50 GV) and low ($\lesssim$1 GV) energies are measured precisely. However, due to the solar modulation, the LIS at middling energy ([$\sim$1, $\sim$50 GV]) can only be obtained by numerical models, which calculate the propagation of GCRs in the Galaxy, such as GALPROP (Vladimirov et al. 2011) and so on. Without direct measurements, there still exist some uncertainties in the LIS (e.g., Corti et al. 2019; Boschini et al. 2020; Zhu et al. 2021), especially for heavy ions.

GCRs are modulated by the disturbed solar wind and its embedded magnetic field when passing through the heliosphere, resulting in changes in spectral shape and flux intensity compared to the original LIS. The transport of particles in the heliosphere is governed by the well known Parker transport equation (TPE; Parker 1965):

$$\frac{\partial f}{\partial t} = -(\boldsymbol{V}_{\rm sw} + \boldsymbol{V}_d)\cdot\nabla f + \nabla\cdot(\overleftrightarrow{K}_s\cdot\nabla f) + \frac{1}{3}(\nabla\cdot\boldsymbol{V}_{\rm sw})\frac{\partial f}{\partial \ln p}, \quad (1)$$

where $f(\boldsymbol{r}, p, t)$ is the GCR phase space density, and f is related to the differential intensity by $j(T) = p^2 f$, where T is a particle's kinetic energy per nucleon and p is the momentum; $\boldsymbol{V}_{\rm sw}$ is the solar wind velocity, $\boldsymbol{V}_d$ is the pitch angle-averaged drift velocity, and $\overleftrightarrow{K}_s$ denotes the diffusion tensor. The terms on the right-hand side of Equation (1) describe the four main physical processes: (1) convection caused by the outward flowing solar wind; (2) adiabatic energy loss due to the expansion of solar wind; (3) gradient, curvature, and drift of the heliospheric current sheet caused by the large-scale magnetic field structure; (4) diffusion resulting from the small-scale magnetic field irregularities. The first process has the same effect for all species, the effect of the second process depends on the spectral index of the LIS, and the effects of the last two processes are different for particle with different mass-to-charge ratios (A/Z), resulting in the dependence of the solar modulation on A/Z.

Based on observations from the ACE-CRIS, Zhao & Zhang (2015) found that the abundance ratios of most relatively abundant species, except carbon, exhibit considerable solar-cycle variations, which are obviously positively correlated with the sunspot numbers with a time lag of about one year. The observation of AMS-02 shows that the monthly proton-to-helium flux ratio (H/He) is time-dependent under 3 GV (Aguilar et al. 2018), and it also seems correlated with the solar activity, i.e., H/He decreases from 2014 (solar maximum) to 2017 (solar declining phase). This behavior is reproduced by Corti et al. (2019), and they claim that this time variation of H/He may result from the A/Z dependence of the solar modulation. Note that the lowest rigidity of H/He measured by AMS-02 in Aguilar et al. (2018) is 1.92 GV. In the work of Ngobeni et al. (2020), based on the observation of PAMELA from 2006 to 2009, the lowest rigidity of simulation is extended to nearly 0.1 GV. The result of the simulation shows that H/He at rigidity lower than 0.6 GV increases from 2007

(solar declining phase) to 2009 (solar minimum), which is contrary to the behavior of H/He at higher rigidity. They claim that this is mainly caused by the different spectral indices of the LIS for protons and helium. Recently, Marcelli et al. (2022) reported the latest measurement of helium flux by PAMELA. Through a weighted average over nine months (which reduces the time resolution of the data), the observation time is extended to 2014 from 2009. This shows that the time profiles of H/He at lower ($\lesssim$1 GV) and higher ($\gtrsim$2 GV) rigidity seem opposite to each other. Limited by insufficient observational data, previous research using simulations is dominated by qualitative analysis. The precise measurements of daily helium and proton fluxes from 2011 to 2019 (2824 days) (Aguilar et al. 2022) by AMS-02 provide us with a large amount of data to quantitatively research the variation of He/H over time. Some preliminary analyses are conducted by Aguilar et al. (2022): for the first time they find that He/H seems proportional to the proton flux at rigidity 1.7–7 GV.

How to characterize the time variation of He/H quantitatively? Is the behavior of He/H at low rigidity ($\lesssim$1 GV) really opposite to that at high rigidity ($\gtrsim$2 GV)? What is the underlying mechanism? Based on the theory of solar modulation, the variation of He/H as a function of solar activity may come from (1) the difference in LIS between protons and helium, and (2) the difference in solar modulation on particles with different *A*/*Z*. What are the respective contributions of these two factors to the time variation of He/H? In this work, we make an effort to answer these questions. We start off with a theoretical derivation to quantitatively describe the change in flux ratio of any GCR particle to protons in Section 2. In Section 3, the modulation parameters in the numerical model tuned by Song et al. (2021) are used to reproduce the measurement of AMS-02, and the lowest rigidity of simulation is extended to 0.1 GV. Furthermore, according to the quantitative description, the relative contributions of the two factors are explored using this numerical simulation. The summary and discussion are placed in Section 4.

2. Quantitative Description

The first step of this work is to find a way to describe the variation of He/H as a function of solar activity. It is well known that the proton flux is anticorrelated with the solar activity (e.g., Potgieter et al. 2015), i.e., high solar activity will hinder GCRs from entering the heliosphere, resulting in a low GCR flux. So, similar to Figure 4 in Aguilar et al. (2022), the proton flux can be used to characterize the strength of the solar activity. However, is He/H versus H a good description in theory? Some derivations are needed to find the appropriate expression.

According to Green's theorem (e.g., Webb & Gleeson 1977; Pei et al. 2010), the modulated GCR differential intensity is determined by

$$j(R_0) \propto f(R_0) = \int_0^\infty G(R|R_0, \boldsymbol{Q}(t)) f_b(R) dR, \tag{2}$$

where G is the normalized transition probability function of observing a particle with rigidity R_0 that has rigidity R at the boundary; $\boldsymbol{Q}(t)$ is a vector of time-dependent parameters describing the transport coefficients. The variables with subscripts 0 and b represent values at the position of the Earth and at the outer boundary of the heliosphere, respectively.

For the sake of simplicity and without loss of generality, f_b is taken to be proportional to a power law of R, and the power index is set to k, i.e.,

$$f_b(R) \propto R^k. \tag{3}$$

Equation (2) can be rewritten as

$$j(R_0) \propto \int_0^\infty G(R|R_0, \boldsymbol{Q}(t)) R^k \, dR. \tag{4}$$

The relative rate of change of differential intensity can be expressed as

$$\frac{d\ln j}{dt} = \frac{dj/dt}{j} = \frac{\int_0^\infty (dG/dt) R^k \, dR}{\int_0^\infty G R^k \, dR} = \sum_i \frac{\int_0^\infty (dG/dQ_i) R^k \, dR}{\int_0^\infty G R^k \, dR} \frac{dQ_i}{dt}, \tag{5}$$

$$d\ln j = \sum_{n=0}^{\infty} k^n a^{(n)}. \tag{6}$$

Equation (6) is the Taylor expansion about k at $k = 0$, where $a^{(n)}$ is only related to the normalized transition probability function G, which is determined by the four main physical processes (convection, drift, diffusion, and adiabatic cooling). Note that, when $k = 0$ (in other words, f_b is independent of the rigidity), j is just the integral of G (as shown in Equation (4)), which is a constant that does not change over time ($dj = 0$). So, $a^{(0)} = 0$.

Let X represent any nuclei other than protons; the rate of variation of the flux ratio of X to protons can be described by

$$S = \frac{d\ln(X/\mathrm{H})}{d\ln(\mathrm{H})} = \frac{k_X}{k_\mathrm{H}} \frac{a_X^{(1)} + \sum_{n=2}^\infty k_X^{n-1} a_X^{(n)}}{a_\mathrm{H}^{(1)} + \sum_{n=2}^\infty k_\mathrm{H}^{n-1} a_\mathrm{H}^{(n)}} - 1, = \frac{k_X}{k_\mathrm{H}} S_r - 1. \tag{7}$$

The variables with subscripts X and H represent values for species X and protons, respectively. The parameter S provides a quantitative description of the variation of flux ratio, and the relative contributions of the differences in LIS and solar modulation are expressed explicitly. The expansion term of S_r depends on the form of $G(R|R_0)$, which is different in different models, and it will be discussed in the next subsection.

2.1. Force-field Approximation

The force-field approximation has become the most commonly used modulation parameter in the literature (e.g., Gleeson & Axford 1968; Moraal 2013; Shen et al. 2021). It originates from several simplifications to the transport equation (TPE), including assuming spherical symmetry, neglecting particle drifts, strictly setting the diffusion coefficient to be proportional to rigidity to the power one, and so on. These simplifications make the model result inevitably deviate from the measurement, but experience shows that the error is within an acceptable range when considering only one GCR species (Herbst et al. 2010; Usoskin et al. 2011). In the force-field approximation, the normalized transition probability function G in Equation (5) is a delta function, i.e., the energy losses for all particles are a definite value, which only changes with the solar

activity. So, in Equation (5), $G(R|R_0, \boldsymbol{Q}(t)) = \delta(R - R_b)$, where

$$R_b = \sqrt{R_0^2 + \phi^2 + 2\phi\sqrt{R_0^2 + (AE_0/Z)^2}}\,, \tag{8}$$

E_0 is the nucleon rest energy, A and Z are the particle's mass and charge numbers, and ϕ denotes the solar modulation potential. So, dG in Equation (5) can be expressed as

$$dG = \frac{d\delta(R - R_b)}{dR_b}\frac{dR_b}{d\phi}d\phi, \tag{9}$$

$$= \frac{d\delta(R - R_b)}{dR_b}\frac{1}{\beta_b}d\phi. \tag{10}$$

The expansion terms in Equation (6) can be given by

$$a^{(n)} = \begin{cases} 1/(\beta_b R_b)\, d\phi & n = 1, \\ 0 & n > 1. \end{cases} \tag{11}$$

Substituting this variable into Equation (7) yields

$$S_r = \frac{\beta_{b,\mathrm{H}} R_{b,\mathrm{H}}}{\beta_{b,X} R_{b,X}}. \tag{12}$$

It can be seen that, in the force-field approximation, S_r is independent of the LIS and is determined by the solar activity (ϕ).

2.2. Comprehensive Potential Model

In our previous work (Song et al. 2022), unlike the force-field approximation, the full TPE without any simplification is used to derive the modulation potential theoretically. It is found that the modulation potential for Galactic cosmic particles with a given rigidity at the Earth is a random variable, which can be fit by an inverse Gaussian (IG) distribution in the following form:

$$IG(\phi, \lambda, \mu) = \sqrt{\frac{\lambda}{2\pi\phi^3}}\exp\left\{\frac{-\lambda(\phi - \mu)^2}{2\mu^2\phi}\right\}, \tag{13}$$

where μ is the average value of potential and its value is determined by three main physical processes (diffusion, convection, drift); λ controls the dispersion of ϕ on both sides of μ and relies solely on the diffusion process. μ and λ depend on time (each AMS-02 Bartels rotation has its own $\mu(R_0)$ and $\lambda(R_0)$ due to the different best-fit parameters for the transport coefficients) and particle species (refer to Song et al. 2022).

By substituting Equation (13) into Equation (5), the explicit expression for $a^{(n)}$ can be obtained (due to the complexity, the specific derivation process is given in the Appendix). Surprisingly, as shown in the Appendix, the values of expansion terms with order greater than 2 are small enough that they can be ignored. Correspondingly, the parameter S_r can be written approximately in the following form:

$$S_r = \frac{a_X^{(1)} + a_X^{(2)} k_X}{a_{\mathrm{H}}^{(1)} + a_{\mathrm{H}}^{(2)} k_{\mathrm{H}}}. \tag{14}$$

This equation is in sharp contrast to the result of the force-field approximation shown in Equation (12). In this model, S_r is still related to the LIS, but the relationship is very simple. As shown in Figure 6, the first term $a^{(1)}$ is much larger than the second term $a^{(2)}k$. This means that the variation of S_r is not significant as a function of k_X and k_{H}. Furthermore, since the change in S_r is small, it can be obtained that the isolines of S_r in the k_X–k_{H} plane should be a series of nearly parallel straight lines.

3. Numerical Simulation

The value of He/H is subject to the dual effects of the LIS and solar modulation, and their respective contributions cannot be distinguished through observation. The theoretical derivation is conducted in the previous section. However, as some approximations are made in the derivation and different approximations bring about diverse results, which one is the most appropriate to describe the variation of He/H needs to be determined through numerical simulation.

In recent years, solving the GCR transport equation by means of time-backward stochastic differential equations (SDEs) has become quite popular. Unlike the finite difference method, the SDE method is unconditionally numerically stable and capable of parallel computation (e.g., Zhang 1999; Kopp et al. 2012; Shen et al. 2019; Luo et al. 2020), which can maximize the power of supercomputers. The SDE equivalent to the TPE has the form

$$d\boldsymbol{r} = (\nabla \cdot \overleftrightarrow{K}_s - \boldsymbol{V}_{\mathrm{sw}} - \boldsymbol{V}_d)ds + \overleftrightarrow{\sigma} \cdot d\boldsymbol{W},$$
$$dp = \frac{1}{3}p(\nabla \cdot \boldsymbol{V}_{\mathrm{sw}})ds,$$

where s is the backward time and $ds = -dt$, $\overleftrightarrow{\sigma} \cdot \overleftrightarrow{\sigma} = 2\overleftrightarrow{K}_s$, and each element in $d\boldsymbol{W}$ satisfies a Wiener process given by the standard normal distribution.

One important task in numerical simulation is to get a suitable model to describe the state of the whole heliosphere, including the solar wind velocity, heliospheric magnetic field, and current sheet, based on observations from a single point. In this work, the same model and methodology are used as before, which are extensively reported by Song et al. (2021) and will not be repeated in this article. In the drift and diffusion model, some parameters are related to the solar activity or time and can be obtained from the observed GCR flux using the Markov Chain Monte Carlo method. In the work of Song et al. (2021), these parameters between 2006 and 2017 are obtained by comparing numerical results with the measurements from PAMELA and AMS-02. In this work, the same method is used to extend the simulation time to 2019 (the latest observation time of AMS-02).

Figure 1 shows the comparison between the numerical result (colored lines) and the monthly averaged daily flux (solid circles) from Aguilar et al. (2021b, 2022). The vertical band denotes the time interval for solar polarity reversal (SPR). As the centers of the lowest rigidity bins of AMS-02 published for proton and helium are 1.08 and 1.815 GV, respectively, some numerical result at lower rigidity cannot be compared or verified by observation. It can be seen that, in Figures 1(a) and (b), our simulation of proton and helium fluxes fits the observations well. The GCR intensity decreases as the solar activity rises, reaching the minimum value after the SPR time, and increases again until the latest observation time. In particular, the sharp dip in 2017 is also reproduced well. Figure 1(c) shows the normalized helium-to-proton flux ratio (divided by the averaged value) for rigidities 0.52 and 1.08 GV (left axis scale), and 2.035 and 4.225 GV (right axis scale). In particular, the ratio has larger variations at rigidities $\lesssim$1 GV, where there are no observations, than at rigidities $\gtrsim$2 GV. It can be seen that, though the numerical result of He/H deviates

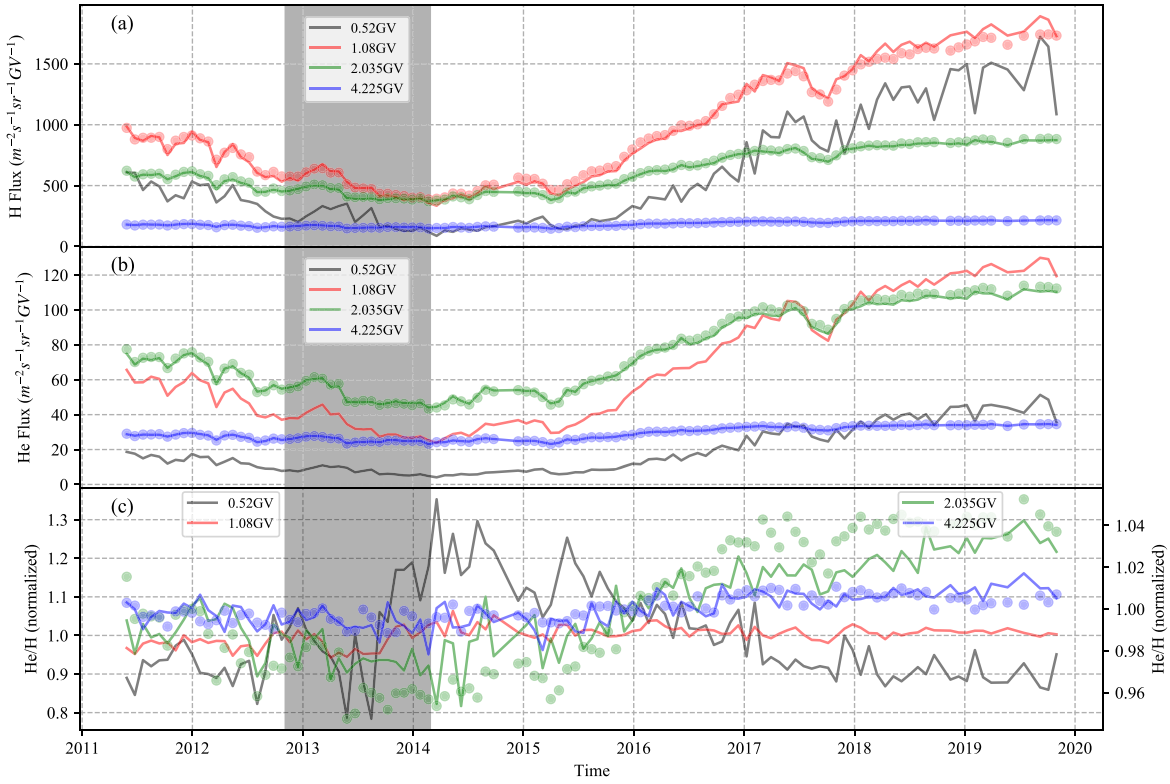

Figure 1. Comparison between simulation (lines) and observation (filled circles) at four selected rigidities (marked by different colors). (a) The proton flux; (b) the helium flux; (c) the helium-to-proton flux ratio, which is normalized to the average value. Note that, restricted by the geomagnetic cutoff and instrumental effects, the centers of the lowest rigidity bins of proton and helium measured by AMS-02 are 1.08 and 1.815 GV, respectively.

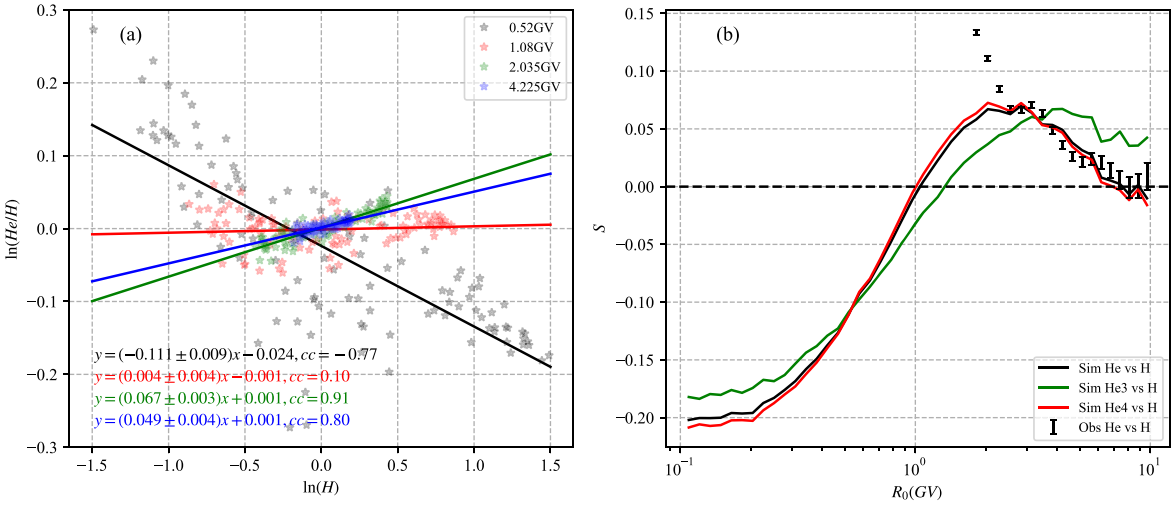

Figure 2. (a) The variation of ln(He/H) as a function of ln(H) for particles at different rigidities (marked by different colors). The average values are subtracted from both the x and y axes. The data points are fitted by the straight line; the fit parameters and the linear correlation coefficient are given at the bottom. (b) The dependence of the parameter S on rigidity for different helium isotopes; the observed result is marked by error bars.

slightly from the observation, the overall trend of the variation with time is obtained. The variation of He/H at 2.035 and 4.225 GV is similar to that of the proton flux; the dip in 2017 also appears on the profile of He/H. However, He/H at 0.52 GV shows a completely opposite trend, i.e., it increases before 2014 then decreases until 2020. Besides, He/H at 1.08 GV does not change over time though some fluctuations exist on the profile. So, one question emerged: what is the tendency of He/H over the full energy range?

According to Equation (7), Figure 2(a) shows the variation of ln(He/H) as a function of ln(H) at different rigidities. It can be seen that these two variables, no matter at what rigidity, have a good linear relationship (though the linear relationship at 0.52 GV is not so significant as that at 2.035 and 4.225 GV; this will be discussed in Section 4). Then a linear fit is performed and the parameters of fitted lines are shown at the bottom of the figure. cc denotes the linear correlation coefficient, which quantitatively characterizes the degree of linear correlation between the two variables. A bigger $|cc|$ means a stronger correlation. It can be seen that the slope of the fitted line or S is negative at 0.52 GV, positive at 2.035 and 4.225 GV, and nearly 0 at 1.08 GV. This result is consistent

with what is shown in Figure 1(c). On the other hand, this means that the parameter S is truly an effective characterization of the variation of He/H. However, in Aguilar et al. (2022), He/H versus H also seems pretty linear. This is due to the small relative variation of H and He during this period (2011–2019). According to Equation (7), $d(\mathrm{He}/\mathrm{H})/d\mathrm{H}$ can be expressed as

$$\frac{d(\mathrm{He}/\mathrm{H})}{d\mathrm{H}} = \frac{\mathrm{H}^2}{\mathrm{He}}S.$$

This means that the slope of He/H versus H is related to the fluxes of protons and helium, which will increase the complexity of analysis. So, S is used to quantify the variation of He/H in this paper.

Then, the S value for the full rigidity range is computed by fitting the modeled ln(He/H) versus ln(H) with a straight line, and the result is shown in Figure 2(b). It can be seen that S in the simulation (black line) starts from −0.2 at 0.1 GV, then it increases with increasing rigidity, crosses 0 at around 1 GV, reaches its maximum value near 2 GV, and then gradually decreases toward 0 at around 10 GV. Helium in cosmic rays is mainly composed of ^{3}He and ^{4}He, which have distinct LISs and A/Z. In order to find the impact of isotopes on the result, ^{3}He (green line) and ^{4}He (red line) are simulated using the same transport parameters. It can be seen that the overall trend of ^{3}He (green line) is similar to that of ^{4}He (red line). They only differ in some details: the rigidity when crossing the 0 line and the rates of ascent and descent before and after the maximum value. Meanwhile, the black line almost follows the red line, because the most abundant isotope of helium is ^{4}He (Aguilar et al. 2019). So, in the following part we focus on ^{4}He. Error bars denote the result calculated using observational data from AMS-02. The simulation fits the observation well when $R_0 > 2.5$ GV, but underestimates the value at lower rigidity. Since the same transport parameters are used in the simulation of proton and helium fluxes, the impact of their difference in A/Z should have been reflected in this picture. Consequently, as discussed by Corti et al. (2019), this deviation may comes from the uncertainty of the LIS. On the basis of Equation (7), a small change in the power index of the phase space density may result in a large change in S, and this effect will be discussed in detail in the next section. Meanwhile, the differences between data and model in this plot seem larger than the differences between data and model in He/H versus time. So, this method could be more sensitive to the LIS than directly comparing the observed and modeled flux ratios. Conversely, this method can be used to correct the LIS until the modeled S matches the observed S.

3.1. Dependence of S on the LIS

As shown in Equation (7), the parameter S is expected to depend on the power index k of the phase space density of the LIS. However, in the real LIS, or the LIS proposed in the literature, k is a not a definite value but changes as a function of rigidity. To investigate the dependence of S on the LIS, the phase space density of the LIS in our numerical procedure is set to follow Equation (3). When the values of k_{He} and k_{H} are determined, S is computed by first solving the SDE with the transport parameters from Song et al. (2021) and a single power-law LIS, and then by fitting S to ln(He/H) versus ln(H).

Figure 3 shows the contour of S in the k_{He}–k_{H} plane for different rigidities. The four panels are very similar to each other, i.e., the isoline of S is almost along the line with constant $k_{\mathrm{He}}/k_{\mathrm{H}}$; this result is consistent with the prediction as shown in Equation (7). However, some small deviations still exist between these two types of line, which may come from the dependence of S_r on k as shown in Equation (14). Different combinations of k_{He} and k_{H} may result in S being negative, positive, or equal to zero. The overall trend is that bigger k_{H} and smaller k_{He} result in larger S. The zero isoline almost coincides with the line $k_{\mathrm{He}}/k_{\mathrm{H}} = 1$ at $R_0 = 0.52$ GV; it rotates clockwise toward the line of $k_{\mathrm{He}}/k_{\mathrm{H}} = 0.9$ at 1.08 and 2.035 GV, and then it rotates in the opposite direction and toward the line of $k_{\mathrm{He}}/k_{\mathrm{H}} = 1$ again at 4.225 GV. This rotation of the zero isoline may come from the dependence of S_r on rigidity as shown in Equation (7). In each simulation of the Bartels rotation, we have 3000 pseudoparticles and 3000 values of R_b are obtained when they cross the heliopause. The mode of the 3000 values of R_b is used to calculate the corresponding k value of the proposed LIS by Corti et al. (2019). The k_{He}, k_{H} pair is plotted in Figure 3 as an asterisk. Though R_0 is a definite value in each panel, the mode of R_b changes with the solar activity, resulting in the dispersion of asterisks. The value of S in Figure 2(b) can be regarded as the average of S at the positions of these asterisks. When $R_0 = 0.52$ GV, the asterisks are mainly distributed in the area above the zero isoline, corresponding to S in Figure 2(b) being less than 0; the asterisks cross the zero isoline when $R_0 = 1.08$ GV, corresponding to S in Figure 2(b) being nearly 0. In Figures 3(c) and (d), the asterisks are mainly distributed in the area below the zero isoline, corresponding to S in Figure 2(b) being greater than 0. As shown in Figure 2(b), our simulation underestimates the observation; the asterisks in Figure 3 should move to the lower right corner (larger k_{H} and/or smaller k_{He}) to match the observation. If the proton LIS is assumed to be accurate, the k value of the helium LIS can be determined from Figure 3.

From the preceding description, we find that the deviation of the isoline with the line of constant $k_{\mathrm{He}}/k_{\mathrm{H}}$ and the rotation of the zero isoline may result from the dependence of S_r on k (contrary to the prediction of the force-field approximation) and R_0. The format of Figure 4 is the same as that of Figure 3, but it is depicted for S_r. It can be seen that S_r varies for different combinations of k_{He} and k_{H}; bigger k_{H} and smaller k_{He} will result in a smaller S_r. The isolines in Figure 4 appear parallel to each other, which is consistent with the prediction of the comprehensive potential model.

Let us focus on the black line, which represents the isoline of $S_r = 1$. In the prediction, its slope is determined by $a_{\mathrm{H}}^{(2)}/a_{\mathrm{He}}^{(2)}$ and its intercept is controlled by $(a_{\mathrm{H}}^{(1)} - a_{\mathrm{He}}^{(1)})/a_{\mathrm{He}}^{(2)}$. So, if the protons and helium nuclei have the same A/Z, the black line should follow the diagonal line no matter what the rigidity. However, it moves to the lower right corner when R_0 changes from 0.52 to 2.035 GV, and moves a little back to the upper left corner when R_0 continues to increase. The asterisks are also marked in this figure. It can be seen that, in all four panels, the asterisks are mainly located above the line $k_{\mathrm{He}}/k_{\mathrm{H}} = 1$ but not far apart. Due to the motion of the black line, the value of S_r at the position of the asterisks increases from less than 1 to more than 1 when R_0 changes from 0.52 to 2.035 GV, and then decreases slightly when R_0 continues to increase. It is precisely this effect that shapes the overall profile of S in Figure 2(b). Meanwhile, the change in S_r in the k_{He}–k_{H} plane is much smaller than the change in $k_{\mathrm{He}}/k_{\mathrm{H}}$, so the specific value of S is mainly determined by $k_{\mathrm{He}}/k_{\mathrm{H}}$.

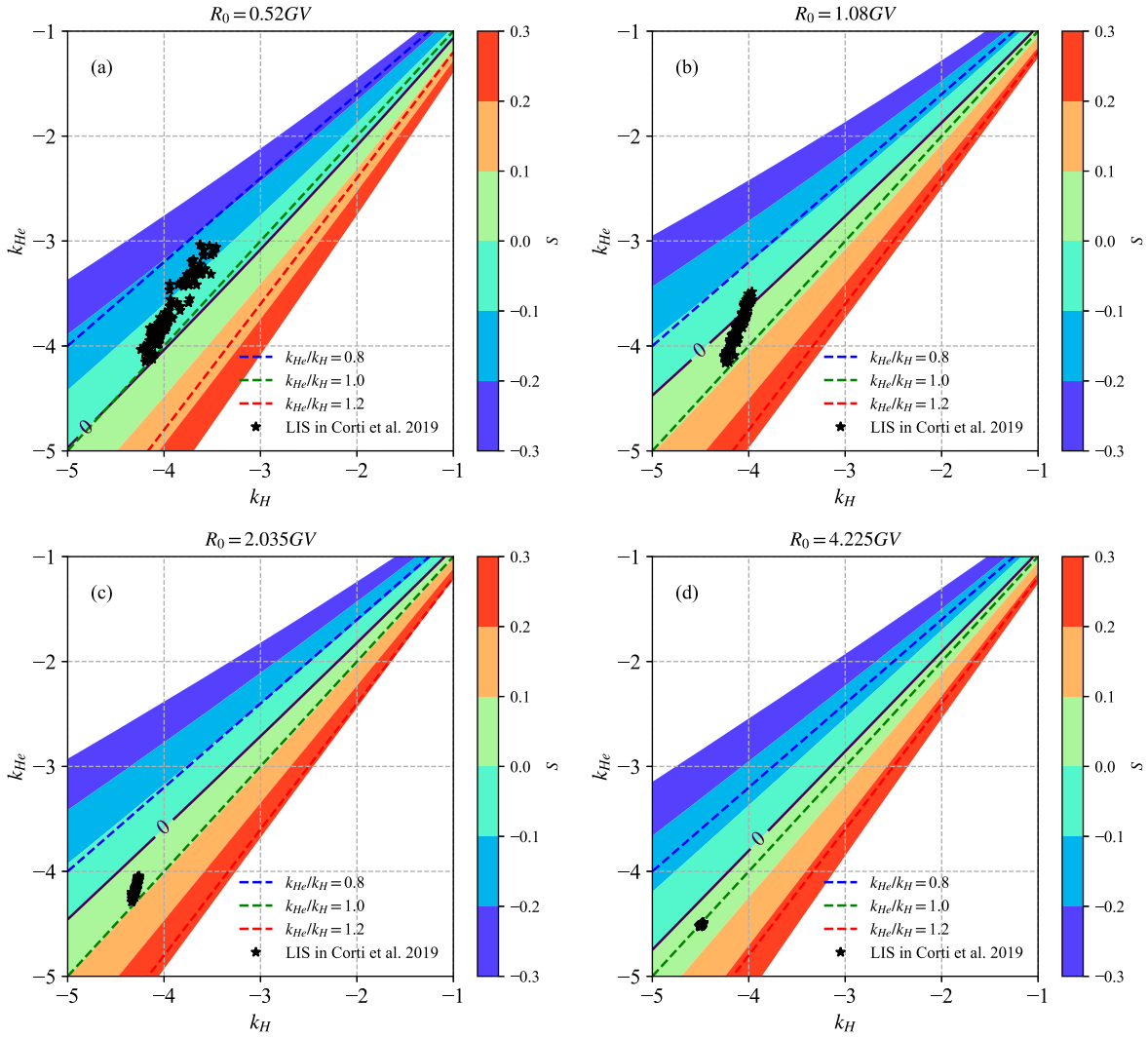

Figure 3. The dependence of S on k. The x and y axes show the power index of proton and helium phase space densities of the LIS, respectively. The asterisks denote the values of $k_{\rm He}$ and $k_{\rm H}$ used to fit the observation.

Figures 3 and 4 show that the change in ln(He/H) as a function of ln(H) indeed matches the prediction of the comprehensive potential model very well. The differences in the LIS and in solar modulation for distinct species can both affect the variation of the flux ratio, and their contribution can be quantitatively described by Equation (14). As shown in Figure 3, the difference in A/Z causes the deviation between the 0 isoline and the diagonal line, the and the difference in the LIS causes the deviation between the asterisks and the diagonal line. In Figure 3(c), the 0 isoline is further away from the diagonal line than the asterisks. So, the difference in A/Z is dominant (this conclusion is the same as in Corti et al. 2019). Moreover, Figure 3(c) shows how large a difference in LIS is needed to outweigh the influence of the difference in A/Z, resulting the decrease in He/H after 2015 March (opposite to the observation), which cannot be obtained from previous works.

4. Summary and Discussion

Protons and helium nuclei undergo the same modulation process after entering the heliosphere, but their flux ratio at the Earth varies as a function of solar activity. The 2824 days of precise observations by AMS-02 unveil for the first time the detailed evolution of He/H, which motivates us to explore its implication. In this context, based on a theoretical derivation, the parameter S in Equation (7) is used to quantify the variation of He/H and links to the underlying micromechanisms, i.e., the difference in the LIS and A/Z dependences of solar modulation. Then the validity of this equation is tested by numerical simulation. By calculating S_r for different combinations of the power index k of the proton and helium phase space densities of the LIS in a numerical procedure, the contour of S_r in the $k_{\rm He}$–$k_{\rm H}$ plane is obtained, and the pattern matches the prediction of the comprehensive potential model very well.

It should be noted that, as shown in Figure 5, when $k_{\rm He}$ and $k_{\rm H}$ are set to a fixed value, the variation of ln(He/H) as a function of ln(H) is not along a strictly straight line. The slope becomes smaller and smaller with increasing proton flux when $R_0 = 0.52$ GV, and the change in slope can be ignored at larger rigidity. This behavior can be explained by the dependence of $a^{(n)}$ (more strictly speaking, the parameters μ and λ, as shown in Equation (A4) in the Appendix) on solar activity. Nonetheless, as the variation of the slope is not significant, and for simplicity, S in Figure 3 is approximately given by the fitted straight line. This processing method makes us focus on the

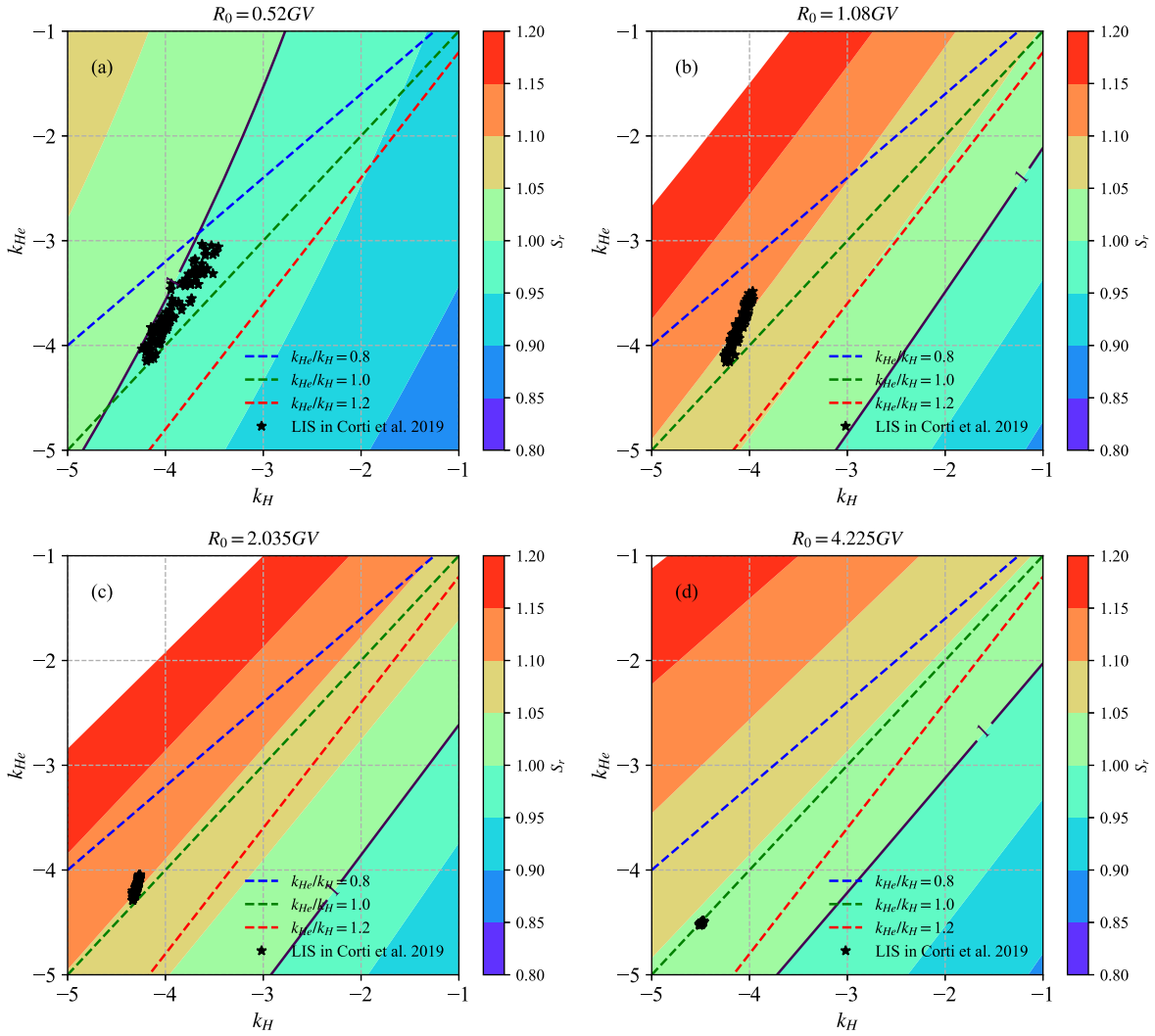

Figure 4. The dependence of S_r on k. The format is the same as in Figure 3.

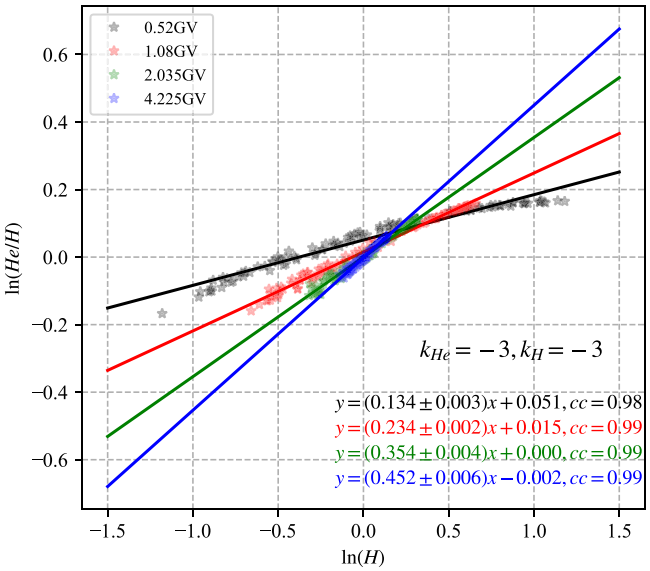

Figure 5. The format is the same as in Figure 2(a), but the data points are calculated using the customized LIS with $k_{\rm He} = -3$ and $k_{\rm H} = -3$, which differ from those for the real LIS.

overall trend and overlook the nuances for different solar activity.

Another noteworthy point is that, compared to Figure 5, the distribution of data points in Figure 2 is very scattered, although it generally follows a straight line. The only difference in plotting the two graphs is the LIS used in numerical simulation: Figure 5 corresponds to an arbitrary LIS with fixed k, whereas the LIS used in Figure 2 is close to the real case (Corti et al. 2019), i.e., k decreases from around -1 at 0.1 GV to close to -5 at 10 GV, and the value and rate of change of k are different for protons and helium. Note that, in Equation (7), k is the power index at rigidity R_b, which changes with the solar activity. So, the combination of $k_{\rm He}$ and $k_{\rm H}$ changes over a large range, which deviates from the assumption of fixed k value in Equation (7), resulting in the weak linearity of data points at low ($\lesssim$1 GV) rigidity. When R_0 is big enough, the variation of R_b as a function of solar activity is smaller, causing a low degree of dispersion for data points at high rigidity.

As shown by Figure 4 in the recent publication of AMS-02 (Aguilar et al. 2022), a hysteresis between He/H and H is observed before and after the solar maximum in 2014, and this hysteresis gradually disappears with increasing rigidity. This

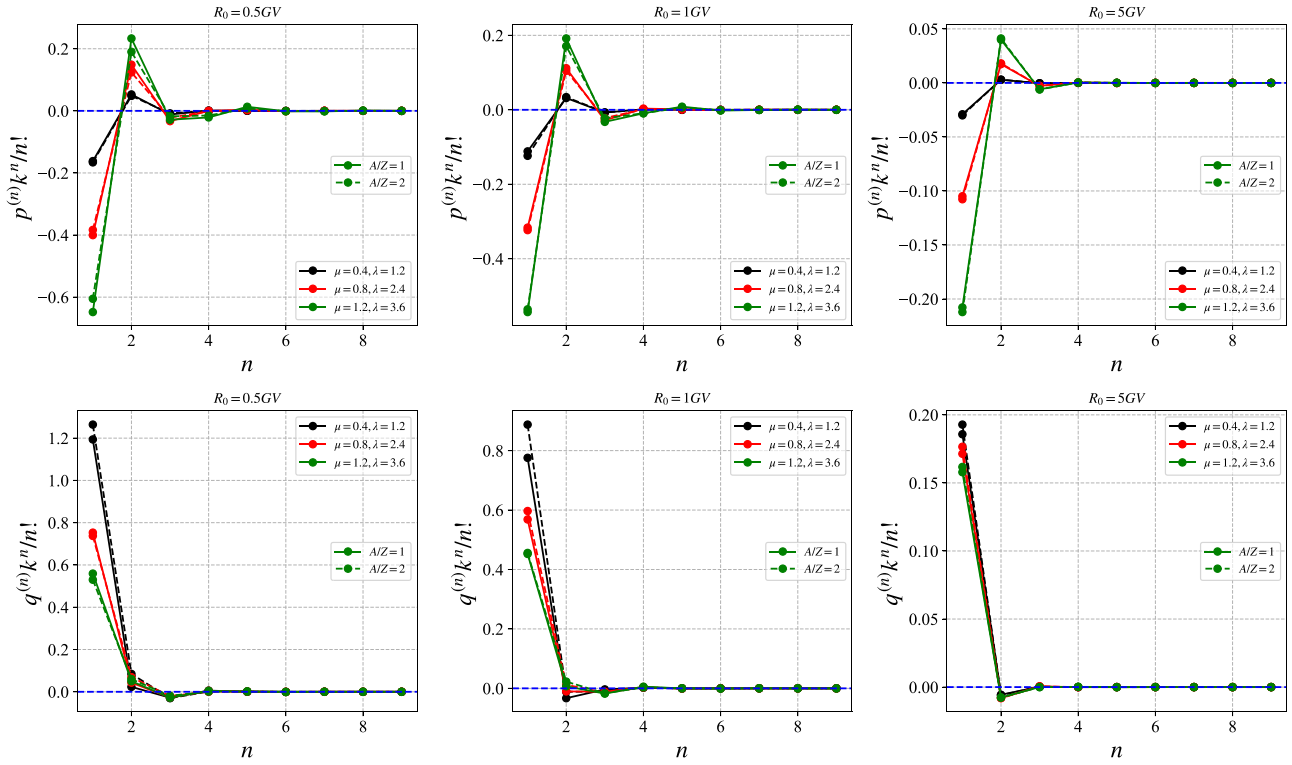

Figure 6. The numerical results of $p^{(n)}k^n/n!$ (top) and $q^{(n)}k^n/n!$ (bottom) for different rigidities when k is set to the proper value -3. Their values are determined by the combination of λ, μ (represent by different colors), A/Z (represented by different line styles), and R_0 (marked in the title).

phenomenon should also exist in our plot of ln(He/H) and ln(H). However, limited by the computing resources, only the monthly averaged data are fitted. So, we do not have as many data as in the AMS-02 observations to carry out detailed research about the hysteresis. Nonetheless, this work is focused on the overall trend of the time variation of He/H, so all the data from 2011 to 2019 are used together to get one parameter S. In our opinion, the hysteresis may result from the difference in the transport time of protons and helium nuclei in the heliosphere (e.g., O'Gallagher 1975; Bazilevskaya et al. 1995; Zhuravlev & Tretyakova 2009) or the change in drift pattern before and after solar maximum (Aguilar et al. 2023). A detailed investigation is beyond the scope of this work; the solution of this problem requires a more accurate numerical model. This will be our next research project.

In the process of deriving the formulae in Section 2, Equation (7) holds unconditionally. Whether Equation (14) holds depends on the normalized transition probability function G. Short-term GCR modulations, such as Forbush decreases or recurrent variations every 27 days, which result from the localized magnetic structure (coronal mass ejection, corotating interaction region, and so on), will disturb the form of G. Whether it can still be matched by an inverse Gaussian distribution is an open question. Does the variation pattern of the GCR flux ratio in the short-term modulation deviate from that in the long-term modulation? This is a question worth further research.

Acknowledgments

The present work is jointly supported by NSFC Grants No. U2106201, No. 42150105, the startup funding of the Shandong Institute of Advanced Technology No. 2020106R01, Taishan Scholar Project of Shandong Province No. 202103143, and Natural Science Foundation of Shandong Province: ZR202111150048.

Appendix

In the comprehensive potential model, $G(R|R_0, \boldsymbol{Q})dR = IG(\phi, \lambda, \mu)d\phi$. So, in Equation (5), $dGdR = dIGd\phi$. Based on the definition of the inverse Gaussian distribution, its derivative can be written in the following form:

$$\frac{dIG}{IG} = \frac{\lambda}{\mu^3}(\phi - \mu)d\mu + \frac{1}{2}\left[\frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\phi} - \frac{1}{\mu^2}(\phi - \mu)\right]d\lambda. \tag{A1}$$

So, the relative rate of change of the observed GCR flux is

$$d\ln j = \frac{\lambda}{\mu^3}\left(\frac{f_2}{f} - \mu\right)d\mu + \frac{1}{2}\left[\frac{1}{\lambda} + \frac{1}{\mu} - \frac{f_1}{f} - \frac{1}{\mu^2}\left(\frac{f_2}{f} - \mu\right)\right]d\lambda \tag{A2}$$

$$= \sum_{n=1}^{\infty} k^n a^{(n)}, \tag{A3}$$

where

$$f = \int_0^\infty IG(\phi, \lambda, \mu)R_b^k d\phi = E(R_b^k),$$

$$f_1 = \int_0^\infty IG(\phi, \lambda, \mu)R_b^k/\phi d\phi = E(R_b^k/\phi),$$

$$f_2 = \int_0^\infty IG(\phi, \lambda, \mu)R_b^k \phi d\phi = E(R_b^k \phi).$$

By comparing Equations (A3) and (A2), the expansion term can be expressed as

$$a^{(n)} = \frac{1}{n!}\left\{\frac{\lambda}{\mu^3}p^{(n)}d\mu - \frac{p^{(n)}/\mu^2 + q^{(n)}}{2}d\lambda\right\}\Bigg|_{k=0}, \quad \text{(A4)}$$

where $p^{(n)}$ is the nth derivative of f_2/f, $q^{(n)}$ is the nth derivative of f_1/f, and they are all independent of k. With increasing n, the expressions for $p^{(n)}$ and $q^{(n)}$ become increasingly complex. However, based on the numerical result (as shown in Figure 6), when k is set to a commonly used value, -3 (which corresponds to $j(T) \sim R(T)^{-1}$ or $j(R) \sim \beta R^{-1}$), the terms with order higher than 2 are small enough that they can be ignored. In this work, only the first two terms are considered, and their analytical expressions are listed below:

$$\begin{aligned}
p^{(1)} &= E(\phi \ln R_b) - E(\phi)E(\ln R_b),\\
p^{(2)} &= E(\phi \ln^2 R_b) - E(\phi)E(\ln^2 R_b)\\
&\quad - 2E(\phi \ln R_b)E(\ln R_b) + 2E(\phi)E^2(\ln R_b),\\
q^{(1)} &= E(\ln R_b/\phi) - E(1/\phi)E(\ln R_b),\\
q^{(2)} &= E(\ln^2 R_b/\phi) - E(1/\phi)E(\ln^2 R_b)\\
&\quad - 2E(\ln R_b/\phi)E(\ln R_b) + 2E(1/\phi)E^2(\ln R_b).
\end{aligned}$$

Meanwhile, it can be seen that in Figure 6 there exists a small difference between the values of $p^{(n)}$ or $q^{(n)}$ for particles with $A/Z = 1$ and $A/Z = 2$. It is precisely these small differences that lead to the black line ($S_r = 1$) in Figure 4 away from the green dotted line ($k_{\rm He}/k_{\rm H} = 1$).

ORCID iDs

Xiaojian Song https://orcid.org/0000-0002-7723-5743
Xi Luo https://orcid.org/0000-0002-4508-6042

References

Aguilar, M., Ali Cavasonza, L., Alpat, B., et al. 2018, PhRvL, 121, 051101
Aguilar, M., Ali Cavasonza, L., Ambrosi, G., et al. 2019, PhRvL, 123, 181102
Aguilar, M., Cavasonza, L. A., Alpat, B., et al. 2021a, PhRvL, 127, 021101
Aguilar, M., Cavasonza, L. A., Ambrosi, G., et al. 2021b, PhRvL, 127, 271102
Aguilar, M., Cavasonza, L. A., Ambrosi, G., et al. 2022, PhRvL, 128, 231102
Aguilar, M., Cavasonza, L. A., Ambrosi, G., et al. 2023, PhRvL, 130, 161001
Bazilevskaya, G. A., Krainev, M. B., Makhmutov, V. S., et al. 1995, AdSpR, 16, 227
Boschini, M. J., Torre, S. D., Gervasi, M., et al. 2020, ApJS, 250, 27
Büsching, I., Kopp, A., Pohl, M., et al. 2005, ApJ, 619, 314
Corti, C., Potgieter, M. S., Bindi, V., et al. 2019, ApJ, 871, 253
Gleeson, L. J., & Axford, W. I. 1968, ApJ, 154, 1011
Herbst, K., Kopp, A., Heber, B., et al. 2010, JGR, 115, D00I20
Kopp, A., Büsching, I., Strauss, R. D., & Potgieter, M. S. 2012, CoPhC, 183, 530
Luo, X., Zhang, M., Feng, X., et al. 2020, ApJ, 899, 90
Marcelli, N., Boezio, M., Lenni, A., et al. 2022, ApJL, 925, L24
Moraal, H. 2013, SSRv, 176, 299
Ngobeni, M. D., Aslam, O. P. M., Bisschoff, D., et al. 2020, Ap&SS, 365, 182
O'Gallagher, J. J. 1975, ApJ, 197, 495
Parker, E. N. 1965, P&SS, 13, 9
Pei, C., Bieber, J. W., Burger, R. A., & Clem, J. 2010, JGRA, 115, A12107
Potgieter, M. S., Vos, E. E., Munini, R., Boezio, M., & Felice, V. D. 2015, ApJ, 810, 141810
Shen, Z., Yang, H., Zuo, P., et al. 2021, ApJ, 921, 109
Shen, Z. N., Qin, G., Zuo, P., & Wei, F. 2019, ApJ, 887, 132
Song, X., Luo, X., Potgieter, M. S., Liu, X., & Geng, Z. 2021, ApJS, 257, 48
Song, X., Luo, X., Potgieter, M. S., & Zhang, M. 2022, PhRvD, 106, 123004
Usoskin, I. G., Bazilevskaya, G. A., & Kovaltsov, G. A. 2011, JGRA, 116, A02104
Vladimirov, A. E., Digel, S. W., Jóhannesson, G., et al. 2011, CoPhC, 182, 1156
Webb, G. M., & Gleeson, L. J. 1977, Ap&SS, 50, 205
Zhang, M. 1999, ApJ, 513, 409
Zhao, L. L., & Zhang, H. 2015, ApJ, 805, 6
Zhu, C.-R., Yuan, Q., & Wei, D.-M. 2021, APh, 124, 102495
Zhuravlev, D. A., & Tretyakova, C. A. 2009, CosRe, 47, 75