New evaluation of the antiproton production cross section
for cosmic ray studies

Mattia di Mauro,[1,2] Fiorenza Donato,[1] Andreas Goudelis,[2] and Pasquale Dario Serpico[2]

[1]*Physics Department, Torino University, and Istituto Nazionale di Fisica Nucleare,*
Sezione di Torino, via Giuria 1, 10125 Torino, Italy
[2]*LAPTh, Université de Savoie, CNRS, B.P.110, Annecy-le-Vieux F-74941, France*

Theoretical predictions for the cosmic antiproton spectrum currently fall short of the corresponding experimental level of accuracy. Among the main sources of uncertainty are the antiproton production cross sections in cosmic ray inelastic interactions. We analyze existing data on antiproton production in pp scattering, including for the first time the measurements performed by the NA49 Collaboration. We compute the antiproton spectrum finding that in the energy range where data are available (antiproton energies of about 4–550 GeV) different approaches lead to almost equivalent results, with an uncertainty of 10%–20%. Extrapolations outside this region lead to different estimates, with the uncertainties reaching the 50% level around 1 TeV, degrading the diagnostic power of the antiproton channel at those energies. We also comment on the uncertainties in the antiproton production source term coming from nuclei heavier than protons and from neutrons produced in pp scatterings, and point out the need for dedicated experimental campaigns for all processes involving antiproton production in collisions of light nuclei.

I. INTRODUCTION

Cosmic ray (CR) antiprotons are a remarkable diagnostic tool for astroparticle physics. The bulk of the measured flux is certainly consistent with a purely secondary origin in CR collisions onto interstellar medium gas, but additional *primary* components are not excluded, either of astrophysical origin (see for instance [1]) or of exotic nature, such as dark matter annihilation or decay [2]. At the very least, antiprotons provide a consistency check for the current understanding of galactic CR modeling and can narrow down propagation parameters (see e.g. [3–5]).

This tool is however only as sharp as the uncertainties entering the background (i.e. the secondary component) and signal (i.e. the primary component) computations are robust. Statistical and systematic errors reported by the PAMELA Collaboration [6] are already at the 10% level up to the 10 GeV scale, below the theoretical error. In a short time, AMS-02 [7] is expected to provide significantly higher precision, calling for a reassessment of the theoretical predictions.

The contribution of different processes to the $\bar{p}$ secondary yield has been studied in the past; see e.g. [8–10]. In [9], for instance, the uncertainties on the production cross sections were estimated to be ∼25%, and already identified as the limiting factor in theoretical predictions (see also [10] for similar considerations). In practice, nuclei heavier than protons and helium only contribute at a few percent level (see e.g. [8]), thus playing a very marginal role, either as projectiles or targets, in the antiproton production. Reactions involving helium (p-He, He-p, He-He) represent

a sizable fraction of the total yield, easily reaching ∼50% at low energies [9].

While for processes involving helium nuclei no data are available, the situation is different for the proton-proton case, where there are several experimental studies. The latest reevaluation of the antiproton production yield in pp collisions was reported in [11], while the Tan and Ng parametrization [12] is still largely used, despite being more than 30 years old. The reason is that, until recently, the available data set was limited to data only collected in the 1960s and 1970s. In the last decade, however, two more experimental data sets have become available: the BRAHMS data [13] and—more important for the energies of interest for AMS-02 applications—the NA49 results collected at the CERN Super Proton Synchrotron [14].

Given the importance of these nuclear data for new measurements in astroparticle physics, it seems thus timely to reevaluate the antiproton production cross section in pp collisions in light of this newly available information. In this paper we engage ourselves in this task, in order to provide the community with a parametrization for the inclusive antiproton production cross section as well as with a reliable assessment of the corresponding uncertainties that should be taken into account in CR studies.

The outline of the paper goes as follows: in Sec. II we set up the relevant formalism, present the experimental data that will be used subsequently and describe the analysis methods. In Sec. III we present our results. We begin by validating our analysis framework by reproducing existing results in the literature and then move on to evaluating the

inclusive antiproton production cross section, first relying solely on the novel NA49 data and then on the full set of available measurements. In Sec. IV we briefly comment on the impact of other contributions entering the secondary antiproton source term, namely of antineutrons and helium nuclei. Finally, in Sec. V we discuss our key results and present our conclusions. Two appendixes follow, where we present some—standard but useful—kinematics used in our analysis as well as a new evaluation of the total, elastic and inelastic pp scattering cross sections that we performed for the energy range of interest to our work.

II. FRAMEWORK, DATA AND METHODS

A. Theoretical framework

CR protons interact with the interstellar medium (ISM) and may produce secondary antiprotons. Different channels are involved, with the dominant one being the CR proton flux collisions with the target hydrogen gas (pp). The corresponding source term is the convolution of the antiproton production cross section $\frac{d\sigma_{pp\to\bar{p}}}{dE_{\bar{p}}}(E_p, E_{\bar{p}})$ and the interstellar CR proton energy spectrum

$$q_{\bar{p}}^{pp}(E_{\bar{p}}) = \int_{E_{\rm th}}^{+\infty} \frac{d\sigma_{pp\to\bar{p}}}{dE_{\bar{p}}}(E_p, E_{\bar{p}})n_H(4\pi\Phi_p(E_p))dE_p, \tag{1}$$

where n_H is the ISM hydrogen density, Φ_p is the CR proton flux, E_p and $E_{\bar{p}}$ are the CR proton and antiproton energies, and $E_{\rm th}$ the production threshold energy equal to $7m_p$. The overall ISM composition is H:He:C = $1:0.1:5 \times 10^{-4}$ cm^{-3} [15]. Whenever needed for illustrative purposes, we will fix Φ_p to the fit to the preliminary AMS-02 data [16] reported in [17].

The differential cross section $d\sigma_{pp\to\bar{p}}/dE_{\bar{p}}$ is in turn the integral over the angle ϑ between the incoming proton and the final state antiproton momenta

$$\frac{d\sigma_{pp\to\bar{p}}}{dE_{\bar{p}}}(E_p, E_{\bar{p}}) = 2\pi p_{\bar{p}}\int_{\vartheta_{\rm min}}^{\vartheta_{\rm max}} E_{\bar{p}}\frac{d^3\sigma}{dp_{\bar{p}}^3} d(-\cos\vartheta), \tag{2}$$

where $p_{\bar{p}} = \sqrt{E_{\bar{p}}^2 - m_{\bar{p}}^2}$, $\vartheta_{\rm min} = 0°$ and the expression for $\vartheta_{\rm max}$ is given by Eq. (A15) in Appendix A. The integral in Eq. (2) is computed in the galactic frame at fixed antiproton energy $E_{\bar{p}}$. Its integrand represents the "Lorentz-invariant distribution function"[1] for the process $p + p \to \bar{p} + X$, i.e. the inclusive antiproton production. Inclusive cross sections for processes of the form $a + b \to c + X$ can be described in terms of the Lorentz-invariant distribution function

---

[1]For simplicity, in the rest of this work we will interchange this term with "cross section."

$$f(a + b \to c + X) = E_c\frac{d^3\sigma}{dp_c^3} = \frac{E_c}{\pi}\frac{d^2\sigma}{dp_L dp_T^2} = \frac{d^2\sigma}{\pi dy dp_T^2}, \tag{3}$$

where p_L, p_T and y are respectively the longitudinal and transverse momentum and the rapidity of particle c. Traditionally, the independent variables most frequently used to parametrize this quantity are

(i) the center-of-mass (CM) energy $\sqrt{s} = \sqrt{2m_p(E_p + m_p)}$, which is uniquely fixed by the total incident proton energy in the lab frame, E_p;

(ii) p_T, the antiproton transverse momentum;

(iii) the so-called "radial scaling" variable x_R, defined as

$$x_R = \frac{E_{\bar{p}}^*}{E_{\bar{p},\max}^*}, \tag{4}$$

where $E_{\bar{p}}^*$ is the antiproton energy and $E_{\bar{p},\max}^*$ is the maximal energy it can acquire, both defined in the CM frame. The maximal antiproton energy is (see Appendix A for details)

$$E_{\bar{p},\max}^* = \frac{s - 8m_p^2}{2\sqrt{s}}, \tag{5}$$

which, from the condition $E_{\bar{p},\max}^* \geq m_p$, also implies the threshold energy for the incident proton in the lab frame, $E_p \geq 7m_p$.

The inclusive antiproton production cross section cannot be computed from first principles. Our primary goal in this work is to provide reliable estimates for the magnitude and the uncertainties of the invariant distribution (3). Our results will be presented mostly in the form of suitable fitting functions. However, we also want to test how reasonable the ansatz of the chosen functional form(s) is. To that purpose, we will also compare the fitted functions to an "agnostic" spline interpolation of the data, which only requires a smooth, piecewise functional dependence. We will mainly focus on antiprotons with energies ranging from a few GeV up to $\mathcal{O}(1)$ TeV, with the upper limit of this interval corresponding roughly to the highest energy that can be probed by AMS-02 and the lower one to the point where astrophysical uncertainties become so large that they constitute the dominant limiting factor in CR studies, a point which we will also briefly comment upon in Sec. IV.

B. Experimental data

In order to estimate the inclusive antiproton production cross section, we consider the data sets [13,14,18–23], reported in Table I in terms of the center-of-mass energy $\sqrt{s}$ and (p_T, x_R) regions (often not rectangular) covered by each experiment. Note that not all experiments provide data in terms of these kinematic variables; in those cases, the data were first converted in terms of $\{s, p_T, x_R\}$. We report

TABLE I.   Data sets used in our analysis along with their corresponding $\sqrt{s}$ values and (p_T, x_R) regions.

Experiment	$\sqrt{s}$ (GeV)	p_T (GeV)	x_R
Dekkers et al., CERN 1965 [18]	6.1, 6.7	(0.00, 0.79)	(0.34, 0.65)
Allaby et al., CERN 1970 [19]	6.15	(0.05, 0.90)	(0.40, 0.94)
Capiluppi et al., CERN 1974 [20]	23.3, 30.6, 44.6, 53.0, 62.7	(0.18, 1.29)	(0.06, 0.43)
Guettler et al., CERN 1976 [21]	23.0, 31.0, 45.0, 53.0, 63.0	(0.12, 0.47)	(0.036, 0.092)
Johnson et al., FNAL 1978 [22]	13.8, 19.4, 27.4	(0.25, 0.75)	(0.31, 0.55)
Antreasyan et al., FNAL 1979 [23]	19.4, 23.8, 27.4	(0.77, 6.15)	(0.08, 0.58)
BRAHMS, BNL 2008 [13]	200	(0.82, 3.97)	(0.11, 0.39)
NA49, CERN 2010 [14]	17.3	(0.10, 1.50)	(0.11, 0.44)

in Appendix A the straightforward but somewhat lengthy derivation of the transformation equations. The data are graphically illustrated in Fig. 1. In the left panel, the cross section is shown as a function of $E_{\bar{p}}^{\rm LAB}$, for different combinations of p_T and x_R. In the right panel, the same data are seen in the $p_T - x_R$ plane. The NA49 data cover wide ranges in both p_T and x_R, and describe lab antiproton energies from about 8 GeV up to 70 GeV.

Compared to the previous works [11,12], the analyses of the NA49 [14] and BRAHMS [13] data sets are new to this paper. Note that the BRAHMS center-of-mass energy corresponds to an incident proton energy of roughly 21 TeV in the lab frame, which lies somewhat beyond the energy region of interest for our work. Given the absence of data for incident proton energies above ∼200 GeV, however, we have included this data set since it can help in guiding the high-energy extrapolation of the fit to physical values. It is worth stressing that in the more interesting tens of GeV region for the antiproton laboratory energy, the major impact will be provided by far by the NA49 data.

Another important conceptual issue concerns the possibility to combine data—whose quality and robustness of

error assessment is very diverse—in a global fit. There is no simple answer to this question: on one hand there are some systematic effects that are certainly present in the old data and hard to estimate and correct for. A known example is provided by the *feed-down* effect. A significant fraction of antiproton production [easily of $\mathcal{O}(20\%)$] comes from strange hyperon (Λ, Σ) decays, whose decay lengths are comparable to or larger than length scales of current microvertex detections or precision tracking. This effect was taken into account in the NA49 data analysis, where the contribution from hyperons has been subtracted from the measured yields [24]. For older experiments, no such correction was performed: while in some cases—as for the CERN ISR—it may be argued that reasonable estimates make the expected correction negligible, for fixed-target experiments covering an extended range of lab momenta the situation is somewhat more complicated. No *a priori* correction has been applied in the following for this effect, especially since ex-novo simulations of trajectories through the detectors and the collimators would be needed for robust estimates. For a semiquantitative discussion, we address the reader to [14]. However, in deriving global fits, we shall allow for experiment-dependent renormalizations,
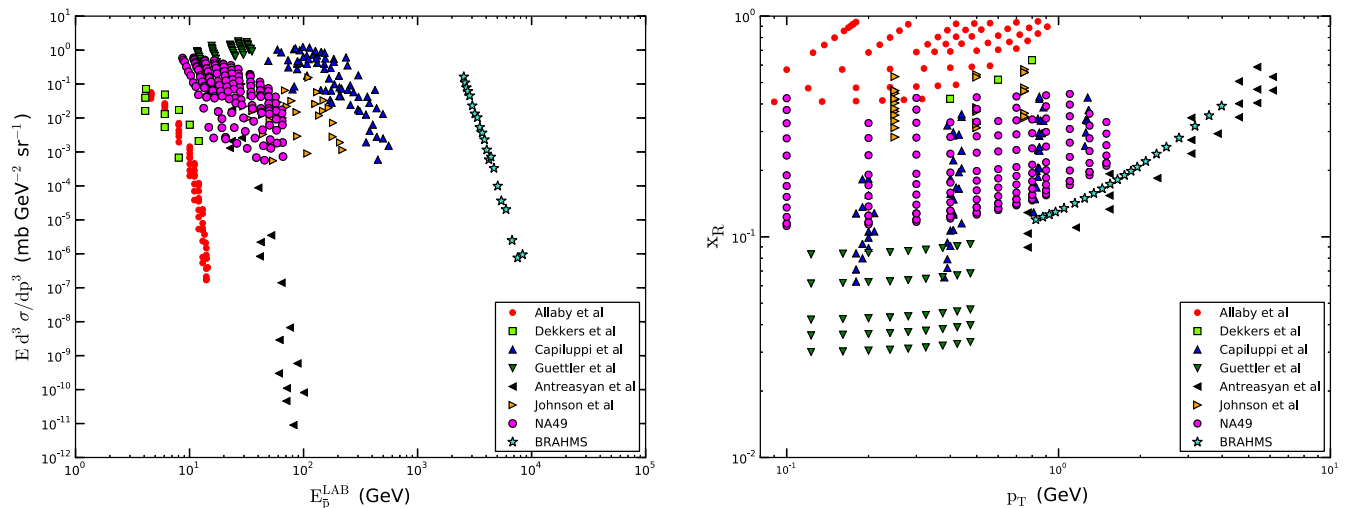

FIG. 1 (color online).   The data on $d^3\sigma_{pp\to\bar{p}}/dp^3$ employed in our analysis are reported as a function of $E_{\bar{p}}^{\rm LAB}$ (left panel) and in the $p_T - x_R$ space (right panel). For the data details, see Table I.

which may account (at least in an averaged way) for such a correction; see below.

On the other hand, relying only on contemporary data, notably NA49, means having the invariant cross section at only one point in $\sqrt{s}$, i.e. at one beam energy. Then, in order to obtain the general cross section, one has to supplement the data with some additional theoretical assumption, such as the *scaling* hypothesis [25], namely that the cross section only depends on p_T and x_R. While this behavior is expected to be approximately respected, notably at high $\sqrt{s}$, its quantitative accuracy can only be gauged by comparison with experimental data. For this reason we decided to apply both strategies and to use either fits or interpolations, to all data sets or to NA49 only, with or without the scaling hypothesis, to assess the importance of these effects.

C. Method

Our fits were performed with the MINUIT minimization package. Let us denote with $k = 1, ..., L$ the different experimental data sets, with i_k the ith point of the data set k, and let $\mathbf{C}$ be the vector of the cross section parameters. The fitting procedure consists of varying the values of the cross section parameters $\mathbf{C}$, comparing the theoretical cross section $F(s_{i_k}, x_{i_k}, p_{i_k}; \mathbf{C})$ with the data $f_{i_k}(s_{i_k}, x_{i_k}, p_{i_k})$ and finally finding the minimum of $\chi^2(\mathbf{C})$ function defined below. This procedure gives the best-fit configuration $\mathbf{C}_{\text{best}}$ with the corresponding 1σ errors $\sigma_{\mathbf{C}}$. We define the $\chi^2(\mathbf{C})$ function to be minimized in the fitting procedure in the following way:

$$\chi^2(\mathbf{C}) = \chi^2_{\text{stat}}(\mathbf{C}) + \chi^2_{\text{sys}} \tag{6}$$

where

$$\chi^2_{\text{sys}} = \sum_{k=1}^{L} \frac{(\omega_k - 1)^2}{\epsilon_k^2}, \tag{7}$$

and

$$\chi^2_{\text{stat}}(\mathbf{C}) = \sum_{k=1}^{L} \sum_{i_k} \frac{(\omega_k f_{i_k} - F(s_{i_k}, x_{R_{i_k}}, p_{T_{i_k}}, \mathbf{C}))^2}{\omega_k^2 \sigma_{i_k}^2}. \tag{8}$$

In the equations above, ϵ_k is a systematic overall scale error of the data set k (either quoted in the experimental paper, or assumed conservatively to be of the same order of the statistical one if this information is not available, notably for older references); the parameter ω_k renormalizes the data set k and is determined consistently by the global fit: of course, large renormalizations with respect to ϵ_k are disfavored by the large penalties to be paid in the global analysis; σ_{i_k} is the statistical error on the data point i_k, while the factor $(\omega_k f_{i_k} - F)/\omega_k$ is the difference between experimental values f (accounting for a possible renormalization

ω_k, unique for each data set) and the fitting function F, which depends on the independent variables described above, and on the vector of fitting parameters $\mathbf{C}$.

This method is the most natural generalization of the *unbiased* one presented in [26] [see therein Eq. (3) and Sec. IV] and it has already been successfully used in the past for other astroparticle physics analyses involving combinations of different data sets, as for instance nuclear reaction rates in primordial nucleosynthesis [27].

Passing on to the data interpolations against which we will be comparing our fitting procedure results, one difficulty lies with the fact that 3-dimensional interpolation of scattered data is a nontrivial problem in contemporary numerical analysis, with very few (if any) relevant publicly available tools. In order to tackle this issue, in our interpolations we will be making the assumption that the invariant distribution (3) scales with $\sqrt{s}$ only through an overall multiplicative dependence on the inelastic pp scattering cross section σ_{in}. Under this assumption, by dividing the experimental data with σ_{in} we obtain a $\sqrt{s}$-invariant set of points for which only a 2-dimensional interpolation is needed. Besides, as a by-product of our analysis we have reevaluated the inelastic cross section as described in Appendix B.

The interpolations were performed by means of the Python routine SmoothBivariateSpline contained in the scipy library, choosing piecewise cubic polynomials as interpolating functions. Note that the routine does not actually perform an exact interpolation, but rather finds a compromise between the smoothness of the interpolating function and the closeness to the experimental data.

Estimating a statistically meaningful uncertainty band in this approach is fairly tricky. What we did was to consider each experimental best determination and error as the mean and standard deviation of a Gaussian probability distribution of the cross section at that point. We then sampled these distributions accordingly, thus creating a large number of pseudoexperimental points. Each set of points is then interpolated (and, depending on the variable one is interested in, eventually integrated over $\cos\vartheta$ and the proton incident energy/proton flux) to obtain an "envelope band" for the quantity of interest. The average between maximum and minimum of the envelope at each point then defines an "average interpolation curve."

III. RESULTS

A. Validation of fitting method

As a preliminary exercise, and in order to validate our kinematical data conversion and fitting routines, we checked if the fit of Eq. (6) in [11] is reproduced, of course restricting ourselves to the data sets available at the time [18–23]. The parametrization of the invariant cross section is in this case

TABLE II.   Best-fit values and 1σ errors for the parameters D_i in Eq. (9) resulting from a fit to the [11] data set.

D_1 (error)	D_2 (error)	D_3 (error)	D_4 (error)	D_5 (error)	D_6 (error)	D_7 (error)
4.22(0.66)	3.435(0.016)	0.0067(0.0014)	0.0510(0.050)	3.609(0.021)	0.0209(0.0010)	3.086(0.083)

$$E\frac{d^3\sigma}{dp^3} = \sigma_{\text{in}}(s)(1 - x_R)^{D_1} e^{-D_2 x_R}$$

$$\times [D_3(\sqrt{s})^{D_4} e^{-D_5 p_T} + D_6 e^{-D_7 p_T^2}], \qquad (9)$$

where σ_{in} is the total inelastic cross section for pp collisions for which here, and only here, we used the parametrization adopted in [11]

$$\sigma_{\text{in}}(s) = \sigma_0 \left[1 - 0.62 e^{-\frac{E_{\text{inc}}(s)}{0.2}} \sin\left(\frac{10.9}{E_{\text{inc}}^{0.28}(s)}\right)\right], \qquad (10)$$

where $E_{\text{inc}}(s)$ is the incident kinetic energy in GeV defined as $E_{\text{inc}}(s) = s/(2m_p) - 2m_p$ and $\sigma_0 = 44.40$ mbarn. We show in Table II our best-fit values and 1σ errors for the cross section parameters D_i in Eq. (9). A modest disagreement with the results of [11] was initially found only for D_1 and D_2, which eventually we could attribute to a typo in the fitting parameter values reported in their Table V. If we invert D_1 with D_2, not only do we obtain a very good agreement with our results, but also a reduced chi-squared $\chi_\nu^2 \simeq 3.6$, the same value the authors quote in their paper. By insisting in interpreting literally the values of their Table V, we would find $\chi_\nu^2 \simeq 9.9$, clearly inconsistent with the value of 3.6. In Fig. 2 we display the comparison of the best fit and 3σ uncertainty band of the source term derived with our best-fit values of the parameters in Table II, and the best-fit source term with the results reported in [11] with D_1 and D_2 inverted. The two results are essentially indistinguishable.
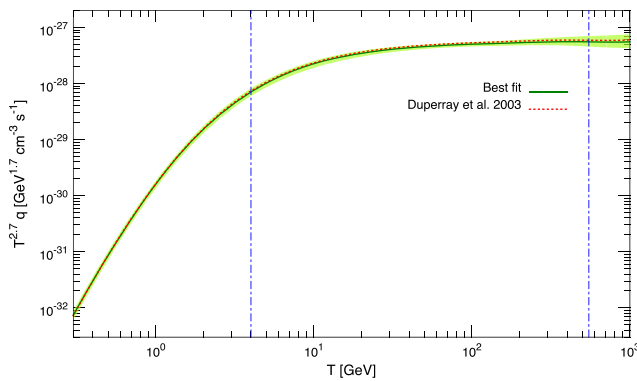

FIG. 2 (color online).   Comparison of the source term for antiproton production in pp collisions as obtained in [11] (see however text for a correction in their table) and in this work, by refitting the same data sets with the same functional form, with our nominal 3σ statistical error band. Vertical dot-dashed lines show the domain of energy actually covered by the experiments analyzed.

B. Analysis of NA49 data

Once our routines are validated, we proceed first with fits to the NA49 data set only. We use the functional form

$$E\frac{d^3\sigma}{dp^3} = \sigma_{\text{in}}(s)(1 - x_R)^{C_1} e^{-C_2 x_R}$$

$$\times [C_3 e^{-C_4 p_T} + C_5 e^{-C_6 p_T^2}], \qquad (11)$$

where $\sigma_{\text{in}}(s)$ is defined in Appendix B, Eq. (B2). This functional form is a simplified version of the standard parametrization proposed in [28] (it has four parameters less), which we found to provide an accurate and more compact description of the data. Note that we implicitly assume some form of scaling, in that the only dependence on s is given by the overall multiplication with the inelastic cross section. The best-fit values and the 1σ errors are reported in Table III, with the corresponding fit having a reduced chi-square $\chi_\nu^2 = 1.3$ for 137 degrees of freedom. The comparison between data and fitted function (along with the corresponding 3σ bands) is presented in Fig. 3. We see that the data are well represented by the fitting function, Eq. (11), for all the p_T and x_R values.

Next, we checked that the chosen fitting formula does not impose too strong a theoretical bias. To that purpose, as described in Sec. II C, we performed an "educated" interpolation of the data by dividing the data points by $\sigma_{\text{in}}(s_{\text{NA49}})$ and assuming that the resulting function is independent of s. The final function which is obtained by remultiplying by $\sigma_{\text{in}}(s)$ thus still has a dependence on s, albeit a trivial one, via the overall factor $\sigma_{\text{in}}(s)$. The comparison between our fitting and interpolating procedures is shown in yellow in Fig. 4. The vertical lines correspond to the equivalent antiproton energy spanned by the NA49 experiment, where an interpolation is meaningful. In order to obtain a reasonable interpolation outside this interval, we supplemented the data sets with "fake" points at the boundaries of the interpolation grid with very large errors that do not artificially influence the curve, yet are sufficient to prevent the numerical routine from being driven to extreme functional form interpolations (for example, negative cross sections). No reasonable error can be however assigned outside the region covered by the data, apart for a lower limit that should be at least as large as the maximum relative width of the yellow band. The fact that the average interpolation curve is always within $\sim 3\sigma$ of the best fit previously obtained suggests that this 3σ band is roughly representing the maximum uncertainty (at least where data exist), accounting not only for

TABLE III.   Best-fit parameters and 1σ errors to the NA49 data [14] with Eq. (11).

C_1 (error)	C_2 (error)	C_3 (error)	C_4 (error)	C_5 (error)	C_6 (error)
7.56(1.15)	0.245(0.148)	0.0164(0.0025)	2.37(0.13)	0.0352(0.0020)	2.902(0.059)

statistical errors, but also for possible theoretical biases, acting as additional systematics, related to the choice of the fitting function.

C. Global analysis

Finally, we proceed to the global analysis of all available data on $pp \rightarrow \bar{p} + X$ listed in Table I. In this case, given that we wish to describe data referring to quite different $\sqrt{s}$ values and covering different (p_T, x_R) regions, it is expected that we will have to introduce some complication with respect to the previous paragraph. In this spirit, we tried numerous different functional forms, essentially variations of the standard parametrization proposed in [28]. We present here results on our two most successful attempts, which also provide interesting insights on the extrapolation to regions where data are either scarce or altogether unavailable, a point that we shall discuss in more detail in Sec. V A.

As a first step, we used an improved version of Eq. (11), introducing an explicit dependence on s, namely

$$E\frac{d^3\sigma}{dp^3} = \sigma_{\text{in}}(s)(1-x_R)^{C_1}e^{-C_2 x_R}$$
$$\times [C_3(\sqrt{s})^{C_4}e^{-C_5 p_T} + C_6(\sqrt{s})^{C_7}e^{-C_8 p_T^2}]. \quad (12)$$

This parametrization of the cross section is similar to the one proposed in [28] except for the absence of a $\sqrt{s}$ exponent in the $(1-x_R)^{C_1}$ term. The fit gives a reduced chi-square of $\chi_\nu^2 = 4.16$, with a number of degrees of freedom of 385. The best-fit values and uncertainties are

reported in Table IV. We have also checked that considering the exact form as in [28] we obtain an even worse fit to the data ($\chi_\nu^2 = 5.6$ with 385 degrees of freedom).

Motivated by the relatively poor quality of the fit, we tried an extended version of Eq. (12), namely

$$E\frac{d^3\sigma}{dp^3} = \sigma_{\text{in}}(s)(1-x_R)^{C_1}e^{-C_2 x_R}|[C_3(\sqrt{s})^{C_4}e^{-C_5 p_T}$$
$$+ C_6(\sqrt{s})^{C_7}e^{-C_8 p_T^2} + C_9(\sqrt{s})^{C_{10}}e^{-C_{11}p_T^3}]|, \quad (13)$$

where the absolute value simply prevents the function from becoming negative in some corners of parameter space. Compared to the previous function, this one further contains the exponential of a cubic function of p_T and an additional s dependence. The best-fit parameters for Eq. (13) are reported in Table V. This parametrization yields a somewhat better $\chi_\nu^2 = 3.30$ for 382 degrees of freedom. In order to test the validity of the scaling hypothesis, we obtain that the fit to the same data with Eq. (13), containing no dependence on s but not in $\sigma_{\text{in}}(s)$, gives a reduced $\chi_\nu^2 = 4.82$.

The improved value of χ_ν^2 is obtained at the expense of some rescaling of the data sets. With respect to our best result given by Eq. (13), the measurements reported in [13,14,18–23] are renormalized respectively by factors ω_k of {0.87, 1.04, 1.16, 0.98, 0.95, 1.13, 1.02, 1.16}. Therefore the NA49 data [14], which represent the bulk of the fitting procedure, are renormalized by a negligible ~2% while [13,18,19,23] by more than 10%. Interestingly, the largest renormalization value is 16% for the BRAHMS data set [13]

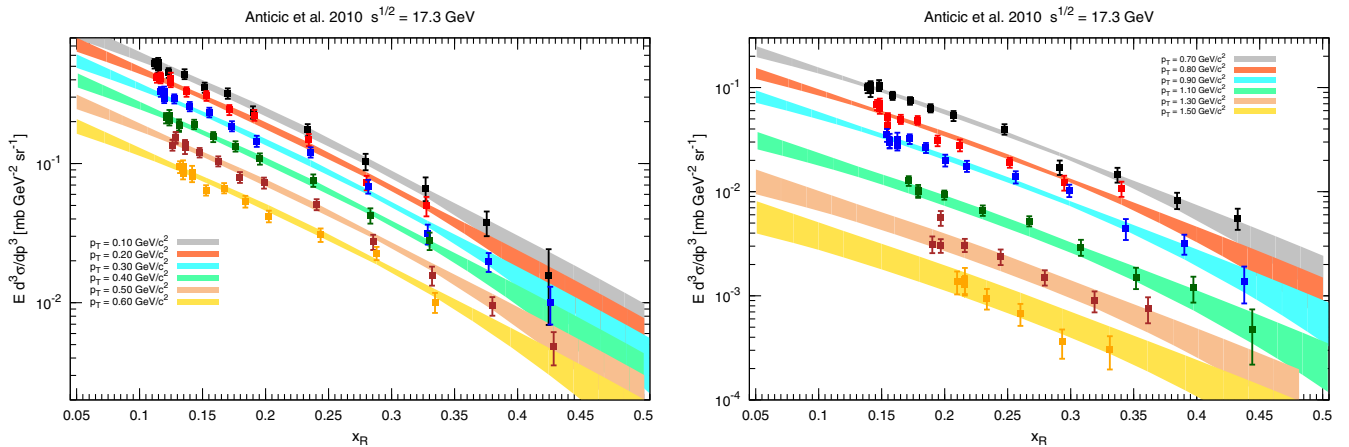FIG. 3 (color online).   Comparison between NA49 data with the fitting function of Eq. (11) (see Table III), with 3σ error bands. For clarity, the data and the theoretical curves at each p_T value have been multiplied by a factor of $0.9^{n_{p_T}}$, where n_{p_T} is the integer counting the p_T, from lower to higher (i.e. for $p_T = 0.60$ GeV/c^2 the rescaling is 0.9^6).
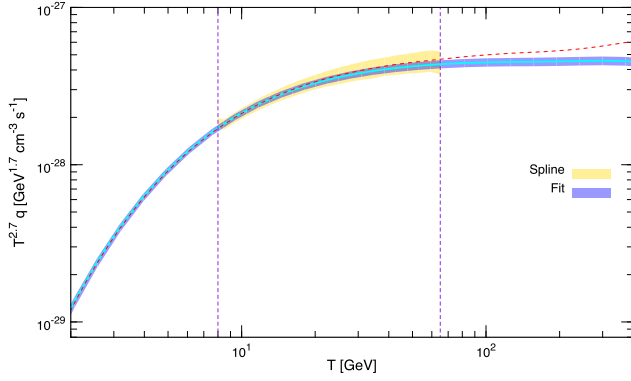
FIG. 4 (color online). Comparison between the function Eq. (11) fitted to NA49 data, with 3σ error band (solid curve with cyan/blue shaded band), and interpolated curve (dashed red), with the interpolation envelope band, yellow/light shading. The dashed vertical lines correspond to the equivalent antiproton energy sampled by the NA49 experiment, where an interpolation is actually justified.

giving a factor of 1.16, which is still not particularly significant given the statistical errors, yet perhaps indicative of some "theoretical error" effects which become more prominent when an agreement over a large energy range is demanded. We display in Fig. 5 the comparison of the cross section best fit and 3σ uncertainty band according to Eq. (13) with the data sets [14,18–23]. We omit the comparison with the BRAHMS results, because in this case the cross section has only been measured along a line in the (p_T, x_R) space (see Fig. 1). Nevertheless, the difference between our best-fit cross section and the data in [13] is at most ∼30%. We see that most of the data are well reproduced by the fitting function of Eq. (13) at all p_T values. This is true in particular for the NA49 data, except for a slight overestimation at the lowest p_T value. We have however checked that a 20% shift in the differential cross section for $p_T < 0.15$ has a negligible effect on the antiproton source term (less than 5%).

We then repeated the interpolation analysis, previously only performed for NA49, for the entire data set. In this case, the parameter space coverage is such that there is no need to supplement the data set with "fake" points, as previously done for the NA49 data alone. The spline method results in this case are, thus, fully data driven, modulo our implicit assumption concerning the cross section $\sqrt{s}$ scaling according to an overall factor $\sigma_{in}(s)$.

In Fig. 6 we compare the results obtained for the antiproton source term through our fits according to Eqs. (13) and (12), as well as the estimate based on our spline interpolation. The energy range where pp data (except for BRAHMS) are available is bracketed by the vertical lines. We see that above 10 GeV, and within the region where experimental data are available, all three methods yield compatible results. At lower and higher energies, however, there is a significant departure of the three estimates. We will discuss the implications of these results in much more detail in Sec. VA.

In order to compare the results derived in this section to previous published proton-proton cross section estimates, we show in Fig. 7 the best-fit and 3σ uncertainty band source term calculated with our results and the best-fit source term derived with the parametrizations adopted in [11,12]. In the range of antiproton kinetic energy where data exist, our 3σ band is marginally compatible with the parametrization in [11,12], which is overestimated (underestimated) below (above) about 20 GeV.

IV. CONTRIBUTIONS FROM NEUTRONS AND NUCLEI

In order to obtain the total antiproton source term, two more effects should be taken into account: the effects of nuclear projectiles and targets in the collisions, and the yield coming from antineutron production. An exhaustive treatment of both subjects goes beyond our current purposes. For completeness, however, in the following we summarise the rescalings of the yield from the $pp \to \bar{p}$ process that are usually adopted in the literature to account for both the processes, and some of the issues involved.

Concerning nuclear enhancements (effects of proton-nucleus, nucleus-proton, and nucleus-nucleus collisions),

TABLE IV. Best-fit values and 1σ errors for the parameters C_i in Eq. (12) derived with a fit to all data sets.

C_1 (error)	C_2 (error)	C_3 (error)	C_4 (error)	C_5 (error)	C_6 (error)	C_7 (error)	C_8 (error)
4.499(0.040)	3.41(0.11)	0.00942(0.00083)	0.445(0.027)	3.502(0.018)	0.0622(0.0086)	−0.247(0.049)	2.576(0.027)

TABLE V. Best-fit values and 1σ errors for the parameters C_i in Eq. (13) derived with a fit to all data sets.

C_1 (error)	C_2 (error)	C_3 (error)	C_4 (error)	C_5 (error)	C_6 (error)
4.448(0.035)	3.735(0.094)	0.00502(0.00036)	0.708(0.019)	3.527(0.014)	0.236(0.024)

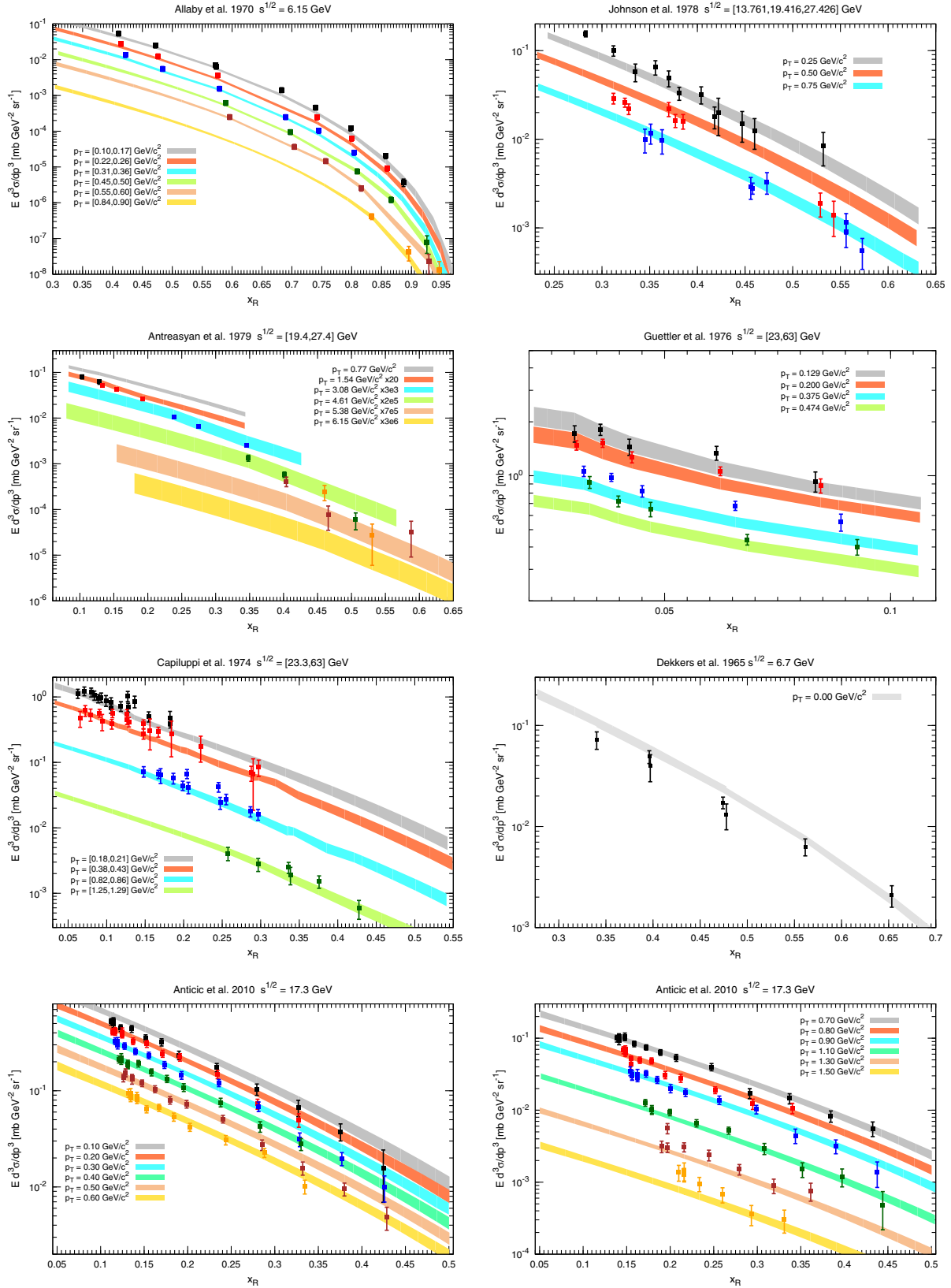| C_7 (error)	C_8 (error)	C_9 (error)	C_{10} (error)	C_{11} (error)
−0.729(0.036)	2.517(0.027)	$-1.822(0.009)10^{-11}$	3.527(0.022)	0.384(0.021)

FIG. 5 (color online).   Differential cross section for antiproton production in pp scattering, as a function of x_R, for different p_T values. The curves refer to the 3σ uncertainty band around the best fit obtained with a fit of Eq. (13) to the data sets in Table I. The data, from top left to bottom right, are from [14,18–23]. For the sake of clarity, the data from [19] and [14] and the relevant theoretical curves at each p_T value have been rescaled by a factor $0.6^{n_{p_T}}$ and $0.9^{n_{p_T}}$, respectively, as described in Fig. 3.
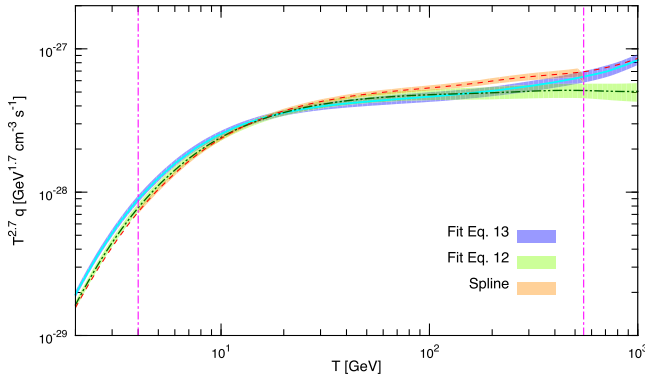
FIG. 6 (color online).   Comparison between fitted function of Eq. (13) with 3σ band (solid curve with cyan/blue shaded band), of Eq. (12) with 3σ band (dot-dashed curve with green/light green shaded band) and interpolated curve (dashed red), with the interpolation envelope band (red/orange shading). The dashed vertical lines correspond to the equivalent antiproton energy sampled by the global data set, where an interpolation is in principle meaningful.
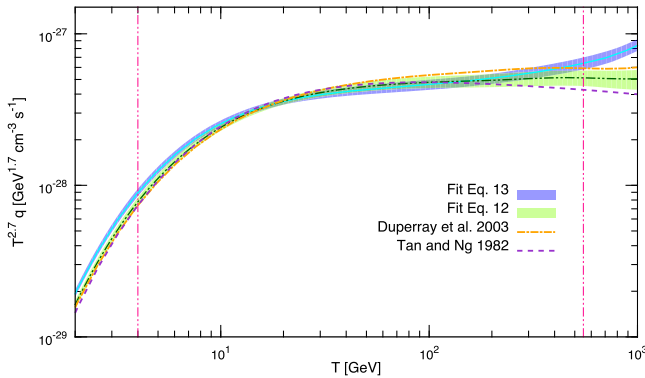

FIG. 7 (color online).   The best-fit and 3σ uncertainty band source term derived with the fit of Eq. (12) and Eq. (13) to all data sets in Table I is shown together with the source term obtained using [11,12] cross section parametrizations.

unfortunately very few data are present, notably none for the most important channels which are the ones involving helium. One possible strategy to deduce cross sections for reactions involving helium is to constrain those of nuclear species for which some data are available, and extrapolate from heavier species to lighter ones; see e.g. [11]. Given that helium is quite light, however, it has often been considered reliable to deduce the relevant cross sections from rescaling the pp ones either with semiempirical formulas or via hadronic models; see e.g. [9]. The most recent dedicated studies were performed on the basis of the Monte Carlo (MC) model DTUNUC in [29] and in [9]. The models implemented in the software are based on the dual parton model [30] and the Gribov-Glauber approach for a unified treatment of soft and hard scattering processes. The former are parametrized according to Regge

phenomenology whereas the latter rely on lowest order perturbative QCD. Eventually, DTUNUC formed the basis of/merged into DPMJET-III (see [31] and references therein for further details).[2]

A fit of the nuclei enhancement yield of antiprotons found in [29] compared to the one in pp collision is given in [10]:

$$Q_{\text{tot}}/Q_{\text{pp}} = 0.12(T_{\bar{p}}/\text{GeV})^{-1.67} + 1.78, \qquad (14)$$

with the above expression assuming a 10% density ratio of He to H nuclei. Note that this ignores different spectral indices and species-dependent spectral breaks, which have been reported by some experiments but have not been confirmed by preliminary results of AMS-02. To gauge their possible effect at high energies, we address the reader to the brief discussion in [33].

Lacking empirical information for the most relevant channels involving helium, it is hard to assess the accuracy of the previous models. The overall uncertainty (on the total source term yield Q_{tot}) was estimated in [29] to reach 40%, from the dispersion of predictions based on different prescriptions, but this conclusion is overly pessimistic, since not all the models/evaluations have the same reliability (some were based on obsolete prescriptions, for instance). In [9] the error estimate was closer to 20%–25%, provided that the pp cross section does not depart from the Tan and Ng parametrization [12] by more than 10%, which seems to be only marginally compatible with our results.

Since all these prescriptions do not include subtle nuclear effects, it is also likely that the uncertainty at low energies (where nuclear effects are expected to play a larger role) is significantly larger than at high energy. If, in addition, one considers the more complicated astrophysical propagation effects at low energies (reacceleration, convection, solar modulation)[3] and the need to correctly account for (catastrophic and noncatastrophic) energy losses, themselves affected by errors, it is clear that below a few GeV's the lower the energy the less reliable is the theoretical prediction. Most likely, this energy window cannot be used (but very crudely) for astroparticle physics constraints.

Another correction which is needed to infer the total antiproton flux from $\sigma_{pp\to\bar{p}}$ consists in accounting for the antiproton flux coming from antineutron production. Traditionally, it has been assumed that $\sigma_{pp\to\bar{n}} = \sigma_{pp\to\bar{p}}$, so that the previous results have been simply multiplied by a factor 2. However, the NA49 Collaboration itself [36] has reported an isospin dependence in studies of secondary yields in np and pp collisions: in pp reactions, there is a significant preference of the positively charged $p\bar{n}$

---

[2]Recently, a new theoretical evaluation appeared in the Appendix of [32].

[3]It has been clearly shown as different propagation setups can be responsible for a ~50% min-max dispersion in the flux expected at low energy [34,35].

combination over $\bar{p}n$ (the opposite being true for neutron projectiles). This results in $\sigma_{pp\to\bar{n}} = \kappa\sigma_{pp\to\bar{p}}$ with $\kappa \simeq 1.5$ around $x_F \sim 0$ (see also Fig. 3 in [36]; x_F is defined in Appendix A), although the effect depends on x_F to some extent. Given the still rudimentary knowledge of these effects, an energy-independent rescaling $\kappa \simeq 1.3 \pm 0.2$ should encompass the data and be a better approximation than the usually assumed $\kappa = 1$. It is clear that addressing these issues is of paramount importance for further reducing the uncertainties in the antiproton source term.

V. DISCUSSION AND CONCLUSIONS

A. Discussion

We now discuss our findings focusing on the global analysis outlined in Sec. III C. As we can see in Fig. 7, for antiproton energies lying roughly within the interval (10, 300) GeV, we find that our results on the antiproton source term from proton-proton scattering are consistent with previous estimates. They are moreover stable with respect to reasonable changes in the parametrization choice and in agreement with data-driven methods. These findings can be understood considering that the majority of the data lie in the $T \in (10, 300)$ GeV range, where the most reliable estimates of the distribution in Eq. (3) can be obtained and which, even prior to the NA49 and BRAHMS measurements, were already discreetly populated with data. In this sense, given that the NA49 data are not in contradiction with previous experimental results, it is expected (and verified) that the estimates presented in Tan and Ng [12] and Duperray *et al.* [11] are in good agreement with our findings for this energy range. Moreover, as long as a reasonable functional form is adopted for the invariant distribution, it is more or less bound to predict a comparable source term within this energy range. The small discrepancies of our spline interpolation and fitting approaches could be likely attributed to the fact that the interpolation essentially neglects the *scaling violation*, while the fits do allow for some flexibility (extra dependence on s) to accommodate it.

On the other hand, at low and high energies, the relatively small amount of available data essentially implies extrapolations of the fits performed principally for T between 8 and 300 GeV. Consequently, moderately different assumptions can yield significantly different results.

This is demonstrated by the fact that adopting two slightly different parametrizations while using the same data set changes the high-energy source term prediction quite dramatically. Moreover, these findings are insensitive to the inclusion or not of the BRAHMS data in the analysis, which means that the results in [13] are not sufficient to constrain the high-energy behavior of the invariant distribution and, hence, the antiproton source function. This is due to the fact that the data of [13] only cover the exponentially suppressed high-p_T region (similarly to the ones of [23]); see Fig. 1. In this sense, both the low-energy and high-energy behaviors of the invariant distribution remain highly uncertain. Given that both the spline method and the fit with Eq. (13) demonstrate a similar trend at high energies, we believe that making any conclusive statement concerning the high-energy behavior of the antiproton inclusive cross section would be risky. This is all the more the case since spline interpolations can be notoriously misleading when extrapolated outside data-covered regions.

Whereas in the low-energy regime this point is not very important, given that in any case the secondary antiproton flux is dominated by huge uncertainties coming from astrophysical sources (solar modulation, propagation parameters, antiproton scattering cross sections), it is plausible that in the region of several hundreds of GeV and higher the main uncertainty is still due to the antiproton production cross section.

We summarize in Table VI the pp-induced source term along with the associated percentage uncertainties resulting from our analysis of the NA49 data according to Eq. (11), our global analysis according to Eqs. (12) and (13), our spline interpolation method of the full data set, and the previous estimates in [12] and [11], for a few representative values of the antiproton energy. This table simply illustrates the results reported in Figs. 6 and 7: with increasing energy, the different approaches turn from marginally compatible (at the lowest energies, few GeV) to fully compatible until, towards the end of the region for which experimental data are available, they yield very different results.

Concerning the error estimates, we also point out that the nominal 1σ error band obtained from the χ^2 minimization procedure is underestimated, for several reasons. In some case where $\chi^2/\text{d.o.f.}$ is close to 1, as in the fit to NA49 data only with a simple fitting formula, we showed how the agreement with an interpolation method is only meaningful

TABLE VI. Best-fit values and corresponding percentage relative errors for the pp-induced source term (in GeV^{-1} cm^{-3} s^{-1}), for some representative antiproton energies and different approaches in the data analysis.

| T (GeV)	Eq. (11) (% error)	Eq. (12) (% error)	Eq. (13) (% error)	Spline (% error)	Tan and Ng	Duperray *et al.*
5	$1.23 \times 10^{-30}(4.9)$	$1.47 \times 10^{-30}(6.1)$	$1.67 \times 10^{-30}(5.4)$	$1.38 \times 10^{-30}(2.7)$	1.42×10^{-30}	1.40×10^{-30}
10	$4.31 \times 10^{-31}(4.2)$	$4.87 \times 10^{-31}(3.0)$	$5.17 \times 10^{-31}(4.8)$	$4.34 \times 10^{-31}(2.5)$	4.96×10^{-31}	4.74×10^{-31}
100	$1.70 \times 10^{-33}(5.9)$	$1.82 \times 10^{-33}(8.7)$	$1.77 \times 10^{-33}(6.8)$	$2.03 \times 10^{-33}(3.2)$	1.82×10^{-33}	2.04×10^{-33}
500	$2.42 \times 10^{-35}(6.2)$	$2.82 \times 10^{-35}(9.5)$	$3.39 \times 10^{-35}(8.8)$	$3.26 \times 10^{-35}(5.2)$	2.38×10^{-35}	3.27×10^{-35}
1000	$3.13 \times 10^{-36}(6.9)$	$4.16 \times 10^{-36}(11)$	$6.83 \times 10^{-36}(10)$	$7.02 \times 10^{-36}(5.8)$	3.29×10^{-36}	4.93×10^{-36}
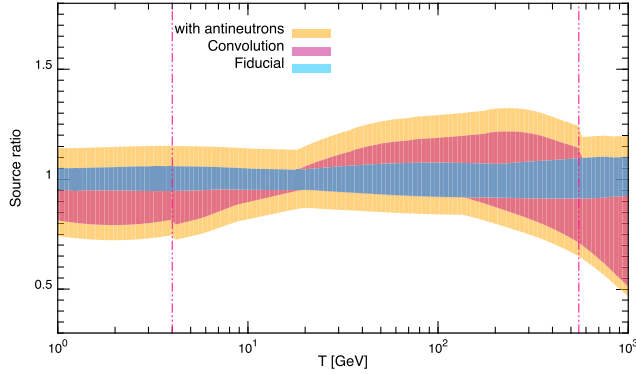
FIG. 8 (color online). Estimate of the uncertainties in the antiproton source term from inelastic pp scattering. The blue band indicates the 3σ uncertainty band due to the global fit with Eq. (13), while the red band corresponds to the convolution of the uncertainties brought by fits to the data with Eq. (13), Eq. (12) and with the spline interpolation (see Fig. 6). The orange band takes into account the contribution from decays of antineutrons produced in the same reactions. Vertical bands as in Fig. 6. See text for details.

if roughly a 3σ band is used as a typical estimate of the error. This is the choice we presented in our plots. A similar prescription was found to be more indicative of the real uncertainty, once global fits were performed. In this case, the inadequacy of the nominal 1σ error band was already hinted at by the relatively large reduced χ^2, never smaller than $\chi_\nu^2 = 3.30$. We attribute these results to a combination of factors: (i) underestimated experimental errors, notably in (some of) the older data sets, due to effects that were neglected as the feed-down we mentioned; (ii) inadequacy of any simple functional form tested to describe faithfully the data, especially on a large dynamic range; (iii) some sort of more or less implicit analytical extrapolation assumption in order to achieve coverage of the 3-dimensional space $(\sqrt{s}, p_T, x_R)$ starting from a discrete set of points. Note that this also applies to interpolation techniques, which for instance rely on some theoretical assumptions such as scaling. The situation may be certainly improved if high-quality measurements such as the ones provided by NA49 could be extended to a broader dynamic range.

We also stress that outside the regions where data are available, there is no compelling reason for either one of our results according to Eqs. (12) and (13) to be more realistic than the other. Whereas the agreement of all of our computations at intermediate energies hints that the error estimates there is fairly reliable, this is not at all the case at very low and high energies. A more conservative approach is to assume that in this case the error is dominated by the extrapolation uncertainty, for which a proxy is given by the region spanned by the ensemble of our approaches, amounting to about 50% at 1 TeV.

As a practical summary of our analysis, we report in Fig. 8 an estimate for the uncertainties inherent to the

production of antiprotons from inelastic pp scatterings. The results are expressed as the ratio of the antiproton source term in Eq. (1) to a reference value. For the blue and the red bands, this reference value has been fixed to the source term obtained by setting the pp production cross section to the best fit to all the data obtained with Eq. (13) (parameters as in Table V). Outside the vertical bands— delimiting the energy range in which data are available— we extrapolate the production cross section by means of the same formula.

The blue band corresponds to considering parametrization (13) alone. By simple inspection we can clearly see that the relevant uncertainty is maximally of the order of 10%. The red band is obtained by convoluting the uncertainty bands resulting from fits through Eqs. (13) and (12) and (within the vertical bands) the spline interpolation. This more conservative approach sizes the uncertainties from 20% at the lowest energies to the extrapolated 50% at 1 TeV.

The most conservative estimate is shown by the orange band, where the additional uncertainty on the antineutron production has been taken into account. In this case, the normalization has been fixed to a source term in which the antineutrons produced in pp scatterings contribute with an energy-independent rescaling factor $\kappa = 1.3$ (with respect to 1). The relevant uncertainty band has been derived by shifting the (red) previous convolution by an additional factor to account for the antineutron decay, $\kappa \simeq 1.3 \pm 0.2$, as discussed in Sec. IV. The orange band indicates that the antiproton source term may vary by 30% at 1 GeV, and by up to more than 50% at 1 TeV. In the energy range where scattering data are available (between 4 GeV and 550 GeV), the uncertainty is of $\mathcal{O}(20\%–30\%)$ nonsymmetrically around the reference source term, with the ambiguity increasing with energy.

Finally, although we did not include an analysis of the uncertainties due to the contribution of nuclei to the yield, it is obvious that the total relative uncertainty cannot be smaller than the one showed above.

B. Conclusions

In this work we have performed a reevaluation of the Lorentz-invariant distribution for inclusive antiproton production in pp collisions in light of the recent results from the NA49 and BRAHMS experiments. We have combined these observations with older measurements at different center-of-mass energies in order to extract a reliable estimate both for the average value of the cross section and for the current status of the corresponding uncertainties. Our main results have been presented in Secs. III B, III C and IV and summed up as handy fitting functions.

We have paid special attention in quantifying the extent to which the functional form we adopt introduces theoretical bias, by comparing our primary results to those obtained by using different parametrizations, data-driven

estimates or different subsets of the data. We are therefore confident that the uncertainties quoted throughout our paper are robust (cf. also the discussion in Sec. V A).

Our findings show that, despite the experimental progress in measuring the inclusive antiproton production cross section, uncertainties persist. In most of the well-constrained intermediate antiproton energy range (few GeV to few hundreds GeV), these uncertainties can be as low as 10%–15%. At higher energies, we have shown that our knowledge is much worse, with extrapolations leading to errors larger than ∼50% at 1 TeV.

A complete determination of the antiproton yield from pp scattering must include the antineutron decay contribution, usually assumed to be identical to the antiproton one. A recent measurement has reported a significant isospin dependence, which could amount to a slight increase of the antiproton source term and be accompanied by a non-negligible uncertainty in the overall flux.

Last but not least, we remind the reader that a significant contribution to the cosmic antiproton flux is due to the reactions involving helium nuclei, both in the incoming radiation and in the ISM. The relevant cross sections have *never* been measured, and the corresponding antiproton source term can be estimated only by relying on some theoretical models and extrapolations, introducing an additional uncertainty which is hard to quantify.

We expect that for the few years to come, these uncertainties will continue imposing non-negligible limitations on the interpretation of cosmic antiproton data, expected to be measured by AMS-02 with an unprecedented accuracy. As a side effect, it appears unlikely that any definite conclusion for dark matter indirect detection could be drawn from a relatively featureless "excess" in the antiproton yield, expected to be below a few tens of percent of the overall flux, unless perhaps correlated features are found in several different cosmic ray channels. At very low energies, the situation is further aggravated by the poor knowledge of astrophysical parameters, plus additional nuclear and particle physics uncertainties related e.g. to noncatastrophic energy losses of antiprotons.

Despite current limitations, many of these sources of error are not irreducible, but could be addressed with dedicated experimental campaigns. As we stressed, the key advances would come from planning experiments with an extended coverage in the (p_T, x_R) plane, like NA49, with higher energies, with helium target/projectile, and by providing systematic analyses of antineutron yields. Another very useful enterprise would be a reanalysis of existing data sets with the aim of assessing the feed-down corrections associated to different experiments. Indeed, the CR antiprotons must include the contributions from hyperon decays, while the current databases are not uniform in that respect.

Cosmic ray antiprotons offer an important tool to address a number of astrophysical and astroparticle questions, as certified by the resources invested in balloon-borne and space-borne detectors in recent years. We believe that a nuclear and particle physics commitment to measure many of these missing ingredients should be identified as a strategic task, to provide astroparticle missions with the crucial laboratory input to fully exploit their results.

Note added.—During the completion of this work, we became aware of a study by R. Kappl and M. W. Winkler [37]. On one side, it focuses on NA49 data interpolation rather than performing a global data set reanalysis with fitting formulas, like ours. On the other hand, it extends to propagated fluxes including other channels as well. As far as a comparison is possible, our results agree with theirs within quoted errors.

APPENDIX A: USEFUL KINEMATICS

In order to extract the invariant distribution of Eq. (3), we have transformed all data (when needed) in the CM frame, expressed them in the form $(p_T, x_R, f(\sqrt{s}, p_T, x_R))$ and then performed our fit and interpolation as described in the main text. Here we collect for convenience some useful formulas by means of which these conversions can be done. In what follows, all kinematic variables carrying stars are defined in the CM frame whereas those carrying LAB superscripts are defined in the lab frame.

Note that experiments [18,19,23] and [14] are fixed-target experiments, whereas [20–22] and [13] are colliding beam ones. With the exception of NA49 [14], fixed-target experiments give their results in the lab (or, equivalently in this case, target) frame. All other data sets are given in the CM frame.

The experiments [18] and [19] do not provide results for the quantity defined in (3), but rather for $d^2\sigma/d\Omega dp$ which can be recast as [38]

$$\frac{d^2\sigma}{d\Omega dp} = p_{\bar{p}}^{\text{LAB},2} \frac{d^3\sigma}{dp^3}. \tag{A1}$$

This leads to

$$E_{\bar{p}}^{\text{LAB}} \frac{d^3\sigma}{dp_{\bar{p}}^3} = \frac{E_{\bar{p}}^{\text{LAB}}}{p_{\bar{p}}^{\text{LAB},2}} \frac{d^2\sigma}{d\Omega dp} \qquad (A2)$$

which can be computed straightforwardly from the available data using $E_{\bar{p}}^{\text{LAB}} = \sqrt{p_{\bar{p}}^{\text{LAB},2} + m_p^2}$.

The results are given as a function of the transverse momentum component p_T and ϑ^{LAB}, the $\bar{p}$ emission angle with respect to the incoming proton beam in the lab frame. Since p_T remains unchanged in the two reference frames, one needs to compute the x_R values which the measurements correspond to. We have that

$$E_{\bar{p}}^* = \gamma^{\text{CM}}(E_{\bar{p}}^{\text{LAB}} - v^{\text{CM}} p_{\bar{p}}^{\text{LAB}} \cos \vartheta^{\text{LAB}}) \qquad (A3)$$

where v^{CM} and γ^{CM} are the velocity of the CM reference frame with respect to the lab one and the corresponding Lorentz boost. These are given by

$$v^{\text{CM}} = p_p^{\text{LAB}}/(E_p^{\text{LAB}} + m_p)$$
$$= p_p^{\text{LAB}}/\left(\sqrt{p_p^{\text{LAB}2} + m_p^2} + m_p\right)$$
$$\gamma^{\text{CM}} = (E_p^{\text{LAB}} p + m_p)/\sqrt{s}. \qquad (A4)$$

Then, direct use of the definition (4) allows us to compute x_R.

Experiments [20,21] and [14] provide results for $f(p + p \to \bar{p} + X)$ in the CM frame, as a function of p_T, and the alternative Feynman scaling variable x_F is defined as

$$x_F = \frac{2p_L^*}{\sqrt{s}} \simeq \frac{p_L^*}{p_{L,\text{max}}^*}, \qquad (A5)$$

where p_L^* is the antiproton longitudinal momentum and $p_{L,\text{max}}^*$ is its maximum possible longitudinal momentum. It is thus necessary to find the correspondence between the x_R and x_F scaling variables. First, from $p_L^{*2} = p^{*2} - p_T^{*2}$ and $p^2 = E^2 - m^2$ we get

$$E_{\bar{p}}^{*2} = p_L^{*2} + m_p^2 + p_T^{*2}, \qquad (A6)$$

whereas $E_{\bar{p}\,\text{max}}^*$ is given by (5). Direct use of (A5) and (4) then gives

$$x_R = \frac{\sqrt{x_F^2(s/4) + m_p^2 + p_T^2}}{E_{\bar{p},\text{max}}^*}. \qquad (A7)$$

Reference [23] provides results for $f(p + p \to \bar{p} + X)$ in the CM frame as a function of p_T and ϑ^{LAB}, also providing the values ϑ^* of the angle in the CM frame. The p_T values remain unchanged in the CM frame. So all that is left is to calculate x_R. Since $p_{\bar{p}}^* = p_T/\sin\vartheta^*$, by using $E = \sqrt{\vec{p}^2 + m_p^2}$ and Eq. (A13), we can calculate x_R through (4).

The BRAHMS experiment [13] chooses to give its results for the invariant cross section

$$\frac{1}{2\pi p_T} \frac{d^2\sigma}{dp_T dy} = E_{\bar{p}} \frac{d^3\sigma}{dp_{\bar{p}}^3} \qquad (A8)$$

as a function of the antiproton transverse momentum p_T and the antiproton rapidity in the CM frame, y^*, defined as

$$y^* = \frac{1}{2} \ln\left(\frac{E^* + p_L^*}{E^* - p_L^*}\right). \qquad (A9)$$

From this definition and $E_{\bar{p}}^{*2} = p_L^{*2} + p_T^2 + m_{\bar{p}}^2$ we get

$$E_{\bar{p}}^* = \sqrt{m_{\bar{p}}^2 + p_T^2} \cosh y. \qquad (A10)$$

Then, x_R can be calculated by means of this relation and (4).

It is maybe useful to report the computation of the maximal antiproton energy $E_{\bar{p}\,\text{max}}^*$ introduced in Eq. (5). For a general inclusive reaction $a + b \to c + X$ in the CM frame we have

$$\sqrt{s} = E_c^* + E_X^*. \qquad (A11)$$

By replacing E_c^*, E_X^* through $E = \sqrt{p^2 + m^2}$, squaring the resulting relation, reintroducing the energies in the crossed terms, eliminating E_X in favor of $\sqrt{s}$ and E_c^*, using $p^{*2} - E^{*2} = -m^2$ to eliminate p^{*2} and solving for E_c^* we get

$$E_c^* = \frac{s + m_c^2 - m_X^2}{2\sqrt{s}}, \qquad (A12)$$

where we introduced for compactness a slight abuse of notation for the X system by assigning it a mass variable m_X. In reality m_X simply refers to internal energy of the X system.

Now, for s and m_p fixed, we see that $E_{\bar{p},\text{max}}^*$ is obtained through (A12) once the energy m_X of the X system becomes minimal. Conservation of baryon number fixes $m_{X,\text{min}} = 3m_p$ (production of 3 additional protons at rest). So we get

$$E_{\bar{p},\text{max}}^* = \frac{s - 8m_p^2}{2\sqrt{s}}. \qquad (A13)$$

Finally we derive the maximum antiproton angle with respect to the incoming proton in the laboratory frame $\vartheta_{\text{max}}^{\text{LAB}}$. This quantity can be derived using the condition $x_R \leq 1$ and the definition of x_R in Eq. (4). Then writing $E_{\bar{p}}^*$ in the laboratory frame [see Eq. (A3)] the condition $x_R \leq 1$ can be written as

$$x_R = \frac{\gamma^{\text{CM}} 2\sqrt{s}(E_{\bar{p}}^{\text{LAB}} - v^{\text{CM}} p_{\bar{p}}^{\text{LAB}} \cos \vartheta^{\text{LAB}})}{s - 8m_p^2} \leq 1. \qquad (A14)$$

By using the definition of v^{CM} and γ^{CM} given in Eq. (A4) we get

$$\cos\vartheta_{max}^{LAB} = \frac{E_{\bar{p}}^{LAB}(E_p^{LAB} + m_p) - m_p(E_p^{LAB} - 3m_p)}{\sqrt{E_{\bar{p}}^{LAB\,2} - m_p^2}\sqrt{E_p^{LAB\,2} - m_p^2}}.$$

$$(A15)$$

TABLE VII. Fit results for the total and elastic proton scattering cross sections according to Eq. (B2).

Parameter	Total	Elastic
M	2.06	3.06
Z^{pp}	33.44	144.98
Y_1^{pp}	13.53	2.64
Y_2^{pp}	6.38	137.27
η_1	0.324	1.57
η_2	0.324	-4.65×10^{-3}

APPENDIX B: INELASTIC CROSS SECTION PARAMETRIZATION

The inelastic proton cross section is defined as the difference between the total pp scattering cross section σ_{tot}^{pp} and its elastic counterpart σ_{el}^{pp}:

$$\sigma_{in}^{pp} = \sigma_{tot}^{pp} - \sigma_{el}^{pp}. \quad (B1)$$

In order to estimate σ_{in}^{pp} for our energy region of interest, we employ the experimental data provided by the Particle Data Group (PDG) on the total and elastic pp cross sections [39]. We fit this data by means of the highest-ranking parametrization of the total proton cross section suggested by the PDG itself, which reads

$$\sigma_{tot}^{pp} = Z^{pp} + B^{pp}\log^2(s/s_M) + Y_1^{pp}(s_M/s)^{\eta_1} - Y_2^{pp}(s_M/s)^{\eta_2}$$

$$(B2)$$

where $B^{pp} = \pi(\hbar c)^2/M^2$, $s_M = (2m_p + M)^2$, all energies are given in GeV and all cross sections in mbarn.

Although this parametrization is given for the total proton cross section, noticing the resemblance in the $\sqrt{s}$ dependence of both the total and the elastic cross section, we employ the same functional form in order to fit the data

on both σ_{tot}^{pp} and σ_{el}^{pp}. In both cases, we include in the fitting procedure only data referring to $\sqrt{s} > 5$ GeV. Moreover, given the large amount of relevant data available for low $\sqrt{s}$ values as opposed to the relative scarcity of measurements for $\sqrt{s} > 80$ GeV, we include the highest-energy data contained in the sample in order to capture the cross section behavior over our entire energy range of interest in a reliable manner.

The fitting parameters for both σ_{tot}^{pp} and σ_{el}^{pp} are shown in Table VII. The resulting functions provide very good fits to the available data, with a χ^2 per degree of freedom of 0.78 for the total cross section and 1.5 for the elastic case, where statistical and systematic uncertainties have been added in quadrature. With these two parametrizations at hand, we then compute the inelastic cross section through its definition (B1). We have verified that, modulo small differences due to the specific goals of our analysis described before, our results are in very good agreement not only with the PDG results on σ_{tot}^{pp} but also with the function for σ_{el}^{pp} quoted by all LHC experiments (see e.g. [40,41]) as well as with the results on both σ_{tot}^{pp} and σ_{el}^{pp} by Block and Halzen [42].

[1] P. Blasi and P. D. Serpico, Phys. Rev. Lett. **103**, 081103 (2009).

[2] P. Salati, F. Donato, and N. Fornengo, arXiv:1003.4124.

[3] D. Maurin, R. Taillet, F. Donato, P. Salati, A. Barrau *et al.*, arXiv:astro-ph/0212111.

[4] C. Evoli, D. Gaggero, D. Grasso, and L. Maccione, J. Cosmol. Astropart. Phys. 10 (2008) 018.

[5] F. Donato, D. Maurin, P. Brun, T. Delahaye, and P. Salati, Phys. Rev. Lett. **102**, 071301 (2009).

[6] O. Adriani *et al.* (PAMELA Collaboration), Phys. Rev. Lett. **105**, 121101 (2010).

[7] http://www.ams02.org/.

[8] I. Moskalenko, A. W. Strong, and O. Reimer, Astron. Astrophys. **338**, L75 (1998).

[9] F. Donato, D. Maurin, P. Salati, A. Barrau, G. Boudoul, and R. Taillet, Astrophys. J. **563**, 172 (2001).

[10] I. V. Moskalenko, A. W. Strong, J. F. Ormes, and M. S. Potgieter, Astrophys. J. **565**, 280 (2002).

[11] R. Duperray, C.-Y. Huang, K. Protasov, and M. Buenerd, Phys. Rev. D **68**, 094017 (2003).

[12] L. Tan and L. Ng, Phys. Rev. D **26**, 1179 (1982).

[13] I. Arsene *et al.* (BRAHMS Collaboration), Phys. Rev. Lett. **98**, 252001 (2007).

[14] T. Anticic *et al.* (NA49 Collaboration), Eur. Phys. J. C **65**, 9 (2010).

[15] J.-P. Meyer, Astrophys. J. Suppl. Ser. **57**, 173 (1985).

[16] H. Haino for the Ams-02 Collaboration, in *Proceedings of the 33rd International Cosmic Ray Conference, Rio de Janeiro, 2013*.

[17] M. Di Mauro, F. Donato, N. Fornengo, R. Lineros, and A. Vittino, J. Cosmol. Astropart. Phys. 04 (2014) 006.

[18] D. Dekkers, J. Geibel, R. Mermod, G. Weber, T. Willitts, K. Winter, B. Jordan, M. Vivargent, N. King, and E. Wilson, Phys. Rev. **137**, B962 (1965).

[19] J. V. Allaby, F. G. Binon, A. N. Diddens, P. Duteil, A. Klovning et al., Report No. CERN-70-12, 1970.

[20] P. Capiluppi, G. Giacomelli, A. M. Rossi, G. Vannini, A. Bertin, A. Bussiere, and R. J. Ellis, Nucl. Phys. **B79**, 189 (1974).

[21] K. Guettler et al. (British-Scandinavian-MIT Collaboration), Nucl. Phys. **B116**, 77 (1976).

[22] J. Johnson, R. Kammerud, T. Ohsugi, D. Ritchie, R. Shafer, D. Theriot, J. Walker, and F. Taylor, Phys. Rev. Lett. **39**, 1173 (1977).

[23] D. Antreasyan, J. Cronin, H. Frisch, M. Shochet, L. Kluberg, P. Piroué, and R. Sumner, Phys. Rev. D **19**, 764 (1979).

[24] H. G. Fischer (private communication).

[25] F. Taylor, D. Carey, J. Johnson, R. Kammerud, D. Ritchie, A. Roberts, J. Sauer, R. Shafer, D. Theriot, and J. Walker, Phys. Rev. D **14**, 1217 (1976).

[26] G. D'Agostini, Nucl. Instrum. Methods Phys. Res., Sect. A **346**, 306 (1994).

[27] P. D. Serpico, S. Esposito, F. Iocco, G. Mangano, G. Miele, and O. Pisanti, J. Cosmol. Astropart. Phys. 12 (2004) 010.

[28] A. N. Kalinovskii, M. V. Mokhov, and Y. Nikitin, Passage of High-Energy Particles Through Matter (AIP, New York, 1989) p. 262.

[29] M. Simon, A. Molnar, and S. Roesler, Astrophys. J. **499**, 250 (1998).

[30] A. Capella, U. Sukhatme, C.-I. Tan, and J. Tran Thanh Van, Phys. Rep. **236**, 225 (1994).

[31] S. Roesler, R. Engel, and J. Ranft, Proceedings of International Conference on Advanced Monte Carlo for Radiation Physics, Lisbon, 2000, p. 1033.

[32] M. Kachelriess and S. Ostapchenko, Phys. Rev. D **86**, 043004 (2012).

[33] F. Donato and P. D. Serpico, Phys. Rev. D **83**, 023014 (2011).

[34] C. Evoli, I. Cholis, D. Grasso, L. Maccione, and P. Ullio, Phys. Rev. D **85**, 123511 (2012).

[35] M. Cirelli, D. Gaggero, G. Giesen, M. Taoso, and A. Urbano, arXiv:1407.2173.

[36] H. Fischer (NA49 Collaboration), Heavy Ion Phys. **17**, 369 (2003).

[37] R. Kappl and M. W. Winkler, arXiv:1408.0299.

[38] E. Byckling and K. Kajantie, Particle Kinematics (John Wiley and Sons, New York, 1971).

[39] http://pdg.lbl.gov/2013/hadronic-xsections/hadron.html.

[40] G. Antchev et al. (TOTEM Collaboration), Europhys. Lett. **101**, 21004 (2013).

[41] G. Antchev et al. (TOTEM Collaboration), Phys. Rev. Lett. **111**, 012001 (2013).

[42] M. M. Block and F. Halzen, Phys. Rev. D **86**, 014006 (2012).