

Toward universal coalescence models for antideuteron production

Mattia Di Mauro^{1,*}, Jordan Koechler^{1,†}, Lorenzo Stefanuto^{2,1,‡}, Francesca Bellini^{3,4,§},
Fiorenza Donato^{2,1,||} and Nicolao Fornengo^{2,1,¶}

¹*Istituto Nazionale di Fisica Nucleare, Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy*

²*Department of Physics, University of Torino, Via P. Giuria, 1, 10125 Torino, Italy*

³*Dipartimento di Fisica e Astronomia, Università di Bologna, Via Irnerio 46, 40126 Bologna, Italy*

⁴*Istituto Nazionale di Fisica Nucleare, Sezione di Bologna, Viale Berti Pichat 6/2, 40127, Bologna, Italy*



(Received 30 March 2026; accepted 11 May 2026; published 3 June 2026)

Cosmic-ray (CR) antinuclei, especially antideuteron $\bar{D}$ and antihelium-3 nuclei ${}^3\bar{\text{He}}$, are among the most promising messengers for indirect dark matter (DM) searches. This is because secondary production in CR interactions with the interstellar medium is strongly suppressed at kinetic energies $K \simeq (0.1\text{--}1)$ GeV/ n , typically one to two orders of magnitude below fluxes expected in standard DM scenarios. From the theoretical side, the formation of $\bar{D}$ and ${}^3\bar{\text{He}}$ is governed by coalescence, whose dynamics cannot yet be reliably derived from first principles. Phenomenological approaches therefore introduce effective coalescence parameters, possibly dependent on collision energy and production environment (hadronic versus electro-weak). In this work we show that a common set of physically motivated coalescence models can simultaneously reproduce collider data in two qualitatively different regimes: ALICE measurements of (anti)deuteron production in pp collisions at $\sqrt{s} = (0.9\text{--}13)$ TeV and the ALEPH $\bar{D}$ multiplicity in hadronic Z decays at $\sqrt{s} = m_Z$. We test both simple event-by-event prescriptions based on a relative-momentum cutoff, finding a preferred coalescence scale $p_{\text{coal}} \simeq 0.2$ GeV, and quantum-mechanical models in the Wigner formalism. In the latter, a Gaussian bound-state wave function gives a best-fit momentum width, corresponding to $\delta \simeq 1.7$ fm, while a parameter-free implementation using the Argonne v_{18} wave function (constrained by proton-neutron scattering data) agrees with ALICE spectra at the $\sim 25\%$ level. Overall, our results support an approximately universal coalescence description across energies and production environments, strengthening the theoretical basis for interpreting upcoming CR antinuclei searches.

DOI: [10.1103/pytl-mmbf](https://doi.org/10.1103/pytl-mmbf)

I. INTRODUCTION

Cosmic-ray (CR) light antinuclei, especially antideuterons ($\bar{D}$) and antihelium nuclei (${}^3\bar{\text{He}}$), are considered among the cleanest messengers for indirect dark matter (DM) searches. The key reason is the pronounced suppression, at kinetic energy per nucleon $K \lesssim 1$ GeV/ n , of the secondary production associated with interactions of primary CRs with the interstellar medium (ISM). In fact, baryon-number conservation implies a large kinematic threshold for

producing $\bar{D}$ in pp collisions: the projectile kinetic energy must satisfy $K \gtrsim 16m_p$, such that secondary $\bar{D}$ fluxes are strongly depleted below ~ 1 GeV/ n . Instead, DM annihilation or decay in the Galactic halo can populate the sub-GeV/ n regime more efficiently, since they are not subject to the same threshold suppression. Theoretical predictions for the $\bar{D}$ flux from standard weakly interacting massive particles (WIMPs) can typically be up to two orders of magnitude higher than the secondary production expectations at low K [1–6]. Consequently, even a small number of detected events at these energies would be difficult to reconcile with standard astrophysical backgrounds and could provide a distinctive DM signature.

Despite sustained experimental efforts over the past decades, no conclusive detection of cosmic antinuclei has been established [7–12]. The strongest published constraint on the antideuteron flux is currently provided by the BESS-Polar II experiment [12], which sets an upper limit of $6.7 \times 10^{-5} \text{ (m}^2 \text{ s sr GeV}/n\text{)}^{-1}$ in the range $K \in (0.163, 1.100)$ GeV/ n . In the near future, the situation is expected to improve substantially: AMS-02 on the International Space Station [13] has been taking data

*Contact author: dimauro.mattia@gmail.com

†Contact author: jordan.koechler@gmail.com

‡Contact author: lorenzo.stefanuto@unito.it

§Contact author: f.bellini@unibo.it

||Contact author: fiorenza.donato@unito.com

¶Contact author: nicolao.fornengo@unito.com

since 2011 and is expected to reach, by about 2030, a sensitivity of $(2\text{--}8) \times 10^{-7} \text{ (m}^2 \text{ sr GeV/n)}^{-1}$ in the interval $K \in (0.2, 4) \text{ GeV/n}$. Moreover, the balloon-borne GAPS program [10,11,14] is designed to reach peak sensitivity comparable to AMS-02 precisely in the low-energy window $K \in (0.1, 1) \text{ GeV/n}$ where the secondary background is minimal [15], and has just completed its first flight (25 days) [16]. These searches therefore have the potential to access flux levels where DM signals can emerge over astrophysical production.

Even though antinuclei may provide one of the cleanest channels for DM indirect detection among cosmic messengers, the theoretical description of their production is affected by significant uncertainties. Interpreting current limits and translating future sensitivities into robust constraints therefore requires accurate predictions for the antinucleus source term. A widely used framework is provided by *coalescence models* [17–19], in which antinuclei form when the constituent antinucleons (antiprotons and antineutrons) are sufficiently close in phase space to fuse into a bound state.

One of the most relevant implications of coalescence, together with Standard Model hadronization, is that the antinuclei formation probability is strongly suppressed with increasing mass number. For instance, in a representative Monte Carlo setup, the relative abundances from DM annihilation scale steeply from antiprotons to $\bar{D}$ and to ${}^3\bar{\text{He}}$ [20],

$$\bar{p}:\bar{D}:{}^3\bar{\text{He}} \sim 1:1.4 \times 10^{-4}: 3.4 \times 10^{-8}. \quad (1)$$

This means that the formation probability drops drastically, by about a factor of 10^4 , for each additional antinucleon in the nucleus. Standard WIMP models that would be able to produce observable fluxes of $\bar{D}$ and ${}^3\bar{\text{He}}$ would also predict antiproton fluxes from DM that should already have been detected by AMS-02, on top of the known secondary astrophysical component [3].

In the simplest implementation—often referred to as the *spherical coalescence* approach—a bound state is produced if, in the pair center-of-mass frame, the relative momentum satisfies $\Delta p < p_{\text{coal}}$, where p_{coal} is the coalescence momentum. This parameter cannot be computed reliably from first principles and is instead calibrated on accelerator measurements when available [2,4]. In phenomenological applications it is also common to allow for different effective values of p_{coal} depending on the production environment, e.g. DM annihilation/decay versus secondary production in hadronic CR interactions. For example, Ref. [3] adopts $p_{\text{coal}} \sim (200\text{--}260) \text{ MeV}$ for secondary production and $p_{\text{coal}} \sim 160 \text{ MeV}$ for DM annihilation (consistent with the calibration in Ref. [2]).

More refined coalescence models, based on a full quantum-mechanical treatment with Wigner formalism, account for the dependence on the momentum distribution of the nucleons, the nucleus wave function, and the

characteristics of the nucleon-emitting source [4,21–24]. This approach is the foundation of recent developments that calculate the production of (anti)nuclei in hadronic interactions event-by-event by employing Monte Carlo simulations [24–26].

In this context, collider measurements provide crucial external anchors to tune the coalescence model. The ALICE Collaboration has measured light (anti)nuclei production in pp collisions over a wide range of center-of-mass energies and multiplicities, delivering multidifferential information in transverse momentum and event activity [27–32]. On the other hand, ALEPH has measured the $\bar{D}$ multiplicity in hadronic Z decays [33,34], a final state often taken as a close proxy for antinucleon production from DM annihilation into quarks. These datasets probe distinct underlying environments—hadronic pp interactions versus hadronic fragmentation in e^+e^- collisions—and are therefore ideally suited to stress test the robustness of coalescence modeling.

Recently, in Ref. [35] some of the authors used the ALEPH $\bar{D}$ measurement [33,34] to calibrate several coalescence prescriptions and, in turn, to significantly reduce the systematic uncertainty affecting the predicted DM-induced $\bar{D}$ source spectra. In particular, they employed a model that implements the Argonne nucleon-nucleon potential, with explicit charge dependence and charge asymmetry [36], in a Wigner formalism (hereafter Argonne) for a full quantum-mechanical treatment of the coalescence process. The Argonne potential is tuned directly to p - p and n - p inelastic scattering data, to low-energy n - n scattering, and to the deuteron binding energy [36]. This approach is more predictive than purely phenomenological prescriptions since the nucleon-nucleon potential is fixed by scattering data and, in this implementation, the coalescence mechanism does not introduce additional free parameters, while providing excellent agreement with the ALEPH measurement [34] of the $\bar{D}$ multiplicity. They also show that, once the antinucleon yield is properly calibrated and the coalescence model is tuned on a common dataset, the residual theoretical systematics become negligible compared to other uncertainties, such as those from QCD modeling, the DM density distribution, or CR propagation.

Building on this progress, the purpose of this work is twofold. On the one hand, we apply the analysis method of Ref. [35] to the secondary production of antinuclei. On the other hand, we directly test the *universality* of the coalescence mechanism across production processes and energies. Concretely, we investigate whether a single, common coalescence framework can simultaneously reproduce (i) the light (anti)nuclei measurements in pp collisions by ALICE at TeV center-of-mass energies [27–32] and (ii) the $\bar{D}$ multiplicity in hadronic Z decays measured by ALEPH [33,34]. Establishing (or falsifying) such universality is critical for consolidating collider-calibrated antinucleus

production models into a coherent input for cosmic-ray antinuclei searches with AMS-02 and GAPS [13–15].

The paper is organized as follows: In Sec. II we describe the coalescence model we employ. Section III describes the tuning of the PYTHIA Monte Carlo code that we use to calibrate the production of antinucleons. Section IV contains the description of the assumptions adopted for the antiproton/antineutron production vertices, and we investigate their production from resonance-baryon decays and from prompt hadronization. In Sec. V we present the results, in Sec. VI we compute the source spectrum in the Galaxy, and we conclude in Sec. VII.

II. COALESCENCE MODELS

Light (anti)nuclei formation is commonly described in terms of *coalescence*: a bound state is produced if the antinucleons are sufficiently close in phase space. In practice, this idea is implemented at different levels of sophistication, ranging from analytic “spherical” prescriptions to event-by-event Monte-Carlo afterburners, up to fully quantum-mechanical formulations based on Wigner functions. In the following we summarize the approaches relevant for this work, emphasizing their assumptions and domains of applicability.

A. Spherical approach

The most elementary coalescence implementation assumes that (anti)protons and (anti)neutrons are emitted independently and isotropically, and that correlations between them can be neglected [1,3]. Under these hypotheses, the formation probability depends on a single parameter, the *coalescence momentum* p_{coal} , and the (anti)nucleus spectrum can be related analytically to the single-(anti) nucleon spectra. For a nucleus with mass number A and proton/neutron numbers (Z, N) , one writes

$$E_A \frac{d^3 N_A}{dp_A^3} = B_A \left(E_p \frac{d^3 N_p}{dp_p^3} \right)^Z \left(E_n \frac{d^3 N_n}{dp_n^3} \right)^N, \quad (2)$$

where the nucleon momenta are evaluated at $p_{p,n} = p_A/A$ and $E_{p,n}$ denote the corresponding energies. In practice, the neutron spectrum is often taken to be equal to the proton one (and analogously for antiparticles), so that for antideuterons one obtains

$$E_{\bar{D}} \frac{d^3 N_{\bar{D}}}{dp_{\bar{D}}^3} = B_2 \left(E_{\bar{p}} \frac{d^3 N_{\bar{p}}}{dp_{\bar{p}}^3} \right)^2, \quad (3)$$

with $B_2 \equiv B_{A=2}$.

In the spherical coalescence approximation, B_A can be expressed in terms of the coalescence momentum p_{coal} as [1,3]

$$B_A = S_A \frac{m_A}{m_p^Z m_n^N} \left(\frac{4\pi}{3} p_{\text{coal}}^3 \right)^{A-1}, \quad (4)$$

where m_A is the nucleus mass and $m_{p,n}$ are the proton and neutron masses. The factor S_A accounts for spin/isospin combinatorics (for $\bar{D}$ one typically has $S_2 = 3/8$).¹ Owing to its simplicity, this prescription has been widely adopted in astroparticle applications, since once p_{coal} is fixed one can compute D or $\bar{D}$ spectra analytically from the (anti) nucleon spectra.

From the experimental side, the coalescence parameter is extracted to be of order $B_2 \sim \mathcal{O}(10^{-2})$ GeV^2/c^3 in pp collisions at LHC energies in the transverse momentum range $p_T/A \sim 1$ GeV/c , and it exhibits a rise with p_T/A in multiplicity-integrated samples [30,37]. In larger collision systems and/or at lower beam energies (e.g. heavy-ion data at LHC/RHIC/SPS), smaller values are commonly found, consistent with the interpretation of B_A as being inversely related to an effective emission volume [38–40]. A residual p_T dependence of B_2 can also arise from spectral-shape effects: if the (anti)proton spectrum is harder than the (anti) deuteron one (i.e., the shapes are not self-similar), then a constant- B_2 ansatz cannot reproduce both slopes simultaneously and the extracted $B_2(p_T)$ acquires an effective p_T dependence [30,41].

Despite its simplicity, neglecting correlations in momentum and configuration space can lead to biased predictions in realistic hadronic environments. In particular, the spherical approximation is known to fail in regions where correlations are most important (e.g. at high- p_T) [2,35]. These limitations motivate the use of event-by-event Monte Carlo implementations discussed in the next sections.

B. Monte-Carlo coalescence with phase-space cutoffs

Monte-Carlo afterburners provide an event-by-event realization of coalescence and, by construction, avoid assuming isotropy or uncorrelated particle production [2,35]. In their simplest form, they implement coalescence through geometric cuts in phase space.

A standard choice is to require that, in the pair center-of-mass frame, the relative momentum of a proton-neutron pair lies within a sphere of radius p_{coal} ,

$$|\vec{p}_p - \vec{p}_n| < p_{\text{coal}}. \quad (5)$$

When applied to each generated event, this criterion selects candidate pairs solely based on their kinematics. For this reason we refer to it as the Δp coalescence model. Owing to its minimal assumptions and computational simplicity, the Δp prescription has long been a default tool for modeling light antinuclei formation in high-energy processes [2,42].

A straightforward extension incorporates spatial correlations by supplementing the momentum cut with a

¹Different conventions for p_{coal} [e.g. using the relative momentum $\vec{q} = (\vec{p}_p - \vec{p}_n)/2$] change only the overall numerical prefactor in Eq. (4).

coordinate-space requirement,

$$|\vec{r}_p - \vec{r}_n| < r_{\text{coal}}, \quad (6)$$

where r_{coal} parametrizes the effective spatial proximity needed for binding. The scale of r_{coal} is expected to be of the order of the physical size of the $\bar{D}$ nucleus [43], i.e. $r_{\text{coal}} \sim 2\text{--}3$ fm. We denote this combined criterion as the $\Delta p + \Delta r$ coalescence model. While still phenomenological, it accounts at least qualitatively for the fact that coalescence is sensitive to both relative momentum and emission geometry. In particular, it naturally suppresses the coalescence of antinucleons originating from displaced vertices (e.g. from weakly decaying hadrons) with promptly produced antinucleons, since such pairs typically fail the spatial cut [2,20].

The Δp and $\Delta p + \Delta r$ prescriptions are historically the most widely used event-by-event coalescence models and were introduced to overcome the main limitations of the spherical approach by retaining kinematic (and, in the second case, geometric) correlations from the underlying event generator [2,42].

C. Monte-Carlo coalescence in the Wigner formalism

To address the shortcomings of purely geometric prescriptions—notably the lack of a first-principles determination of p_{coal} and r_{coal} , their possible dependence on process and energy, and the incomplete treatment of the space-time structure of the emitting source—one can formulate coalescence within a quantum-mechanical framework based on Wigner functions [19,21–24,44]. In this approach, antinucleons are described by quantum states with nontrivial distributions in both coordinate and momentum space, and the formation probability of a bound state is obtained by projecting the two-particle state onto the deuteron bound-state density matrix.

The (anti)deuteron differential spectrum can be written as

$$\frac{d^3 N_D}{dp_D^3} = \text{Tr}(\rho_D \rho_{pn}), \quad (7)$$

with $\rho_D = |\phi_D\rangle\langle\phi_D|$ and $\rho_{pn} = |\psi_{pn}\rangle\langle\psi_{pn}|$, where ϕ_D is the (anti)deuteron wave function and ψ_{pn} the two-particle p – n wave function.

Introducing the relative and center-of-mass coordinates,

$$\vec{r} \equiv \vec{r}_p - \vec{r}_n, \quad \vec{R} \equiv \frac{\vec{r}_p + \vec{r}_n}{2} \equiv \vec{r}_D, \quad (8)$$

one may factorize the bound-state wave function into internal and center-of-mass parts as

$$\phi_D(\vec{r}, \vec{R}; \vec{p}_D) \propto \varphi_D(\vec{r}) e^{i\vec{p}_D \cdot \vec{R}}, \quad (9)$$

where $\varphi_D(\vec{r})$ denotes the *internal* wave function (depending on the relative coordinate $\vec{r}$), and $\vec{p}_D$ is the deuteron

three-momentum. The differential spectrum can then be expressed as [19,24]

$$\frac{d^3 N_D}{dp_D^3} = S \int \frac{d^3 r_D d^3 r d^3 q}{(2\pi)^6} \mathcal{D}(\vec{r}, \vec{q}) \times W_{pn} \left(\frac{\vec{p}_D}{2} + \vec{q}, \frac{\vec{p}_D}{2} - \vec{q}, \vec{r}_p, \vec{r}_n \right), \quad (10)$$

where S accounts for spin/isospin statistics (for D , $S = 3/8$), $\vec{q} \equiv (\vec{p}_p - \vec{p}_n)/2$ (so that $\Delta\vec{p} = 2\vec{q}$), and $\vec{r} \equiv \vec{r}_p - \vec{r}_n$ as above. In the rest of the paper we will use both Δr and r to define the modulus of the vector $\vec{r} \equiv \vec{r}_p - \vec{r}_n$, i.e. the distance between p and n . The function W_{pn} is the Wigner function of the p – n pair, while $\mathcal{D}$ is the deuteron Wigner function, defined by [19]

$$\mathcal{D}(\vec{r}, \vec{q}) = \int d^3 \xi e^{-i\vec{q} \cdot \vec{\xi}} \varphi_D(\vec{r} + \vec{\xi}/2) \varphi_D^*(\vec{r} - \vec{\xi}/2), \quad (11)$$

and normalized as

$$\int d^3 r \int \frac{d^3 q}{(2\pi)^3} \mathcal{D}(\vec{r}, \vec{q}) = 1. \quad (12)$$

The functional form of $\mathcal{D}(\vec{r}, \vec{q})$ is set by the choice of φ_D ; below we consider two standard options.

For the Wigner function of the p – n pair we adopt the factorized form

$$W_{pn} = H_{pn}(\vec{r}_p, \vec{r}_n) G_{pn} \left(\frac{\vec{p}_D}{2} + \vec{q}, \frac{\vec{p}_D}{2} - \vec{q} \right), \quad (13)$$

where G_{pn} is the two-particle momentum distribution obtained from the Monte-Carlo generator (including correlations). For the spatial part, we approximate

$$H_{pn}(\vec{r}_p, \vec{r}_n) = h(\vec{r}_p) h(\vec{r}_n), \quad (14)$$

with h extracted from the same generator. We stress that this approximation neglects genuine two-particle spatial correlations beyond those implicitly induced through the common event geometry. The corresponding impact is explored through variations of the vertex modeling discussed in Sec. IV. With these ingredients specified, the remaining model dependence resides primarily in the bound-state wave function and in the effective source size.

1. Gaussian wave function

As a first benchmark (labeled as Gaussian), we adopt a Gaussian ansatz for the internal deuteron wave function,

$$\varphi_D(r) = (\pi d^2)^{-3/4} \exp \left(-\frac{r^2}{2d^2} \right), \quad (15)$$

where $r \equiv |\vec{r}|$ denotes the modulus of the relative separation $\vec{r} = \vec{r}_p - \vec{r}_n$. The size parameter d can be related

to the rms charge radius [defined through $r_{\text{RMS}}^2 = \int d^3r (r/2)^2 |\varphi_D(r)|^2$]; in particular, $d \simeq 3.2$ fm reproduces the measured deuteron rms charge radius [43]. We further assume Gaussian single-particle source profiles,

$$h(r_{p,n}) = \frac{1}{(2\pi\sigma^2)^{3/2}} \exp\left(-\frac{r_{p,n}^2}{2\sigma^2}\right), \quad (16)$$

where σ characterizes the spatial extent of the emitting region. For a Gaussian bound state, the corresponding Wigner function takes the simple form [4]

$$\mathcal{D}(r, q) = 8 \exp\left(-\frac{r^2}{d^2}\right) \exp(-q^2 d^2). \quad (17)$$

Following the common procedure in [24,26], one may integrate out the explicit coordinate dependence in Eq. (10). Under the Gaussian assumptions above, this yields

$$\frac{d^3 N_D}{dp_D^3} = \frac{3}{(2\pi)^6} \left(\frac{d^2}{d^2 + 4\sigma^2}\right)^{3/2} \int d^3q e^{-q^2 d^2} G_{pn}(\vec{p}_D, \vec{q}), \quad (18)$$

so that the spectrum is obtained numerically from the Monte Carlo estimate of G_{pn} , while the prefactor depends only on the deuteron size d and the source size σ .

This “integrated” implementation is the one considered for example in [4,24,25], which takes from the Monte Carlo only the momentum distribution of the antinucleons. Instead, our goal is to exploit the full $\vec{p}-\vec{n}$ phase space information available in the event generator including their spatial distribution. To this end, we interpret the Wigner kernel as a probability density in $(\Delta\vec{r}, \Delta\vec{p})$ and adopt a Gaussian probability distribution function (PDF),

$$\text{PDF}(r, q) = N \exp\left(-\frac{r^2}{2\sigma^2}\right) \exp\left(-\frac{q^2 \delta^2}{2}\right), \quad (19)$$

with N fixed by normalization. In the strict Gaussian Wigner limit one expects σ and δ to be, respectively, of order $d\sqrt{2} \simeq 4$ fm and $d/\sqrt{2} \simeq 2$ fm [cf. Eqs. (17) and (19)]. However Eq. (19) should be interpreted as a generalization of Eq. (17), thus we first assume that σ and δ are uncorrelated. The cumulative coalescence probability for a pair with separation r and relative momentum q can then be written as

$$\mathcal{P}(r, q) = \int_0^r dr' \int_0^q dq' \mathcal{D}(r', q'), \quad (20)$$

which is consistent with the normalization in Eq. (12).

2. Argonne wave function

As a second (labeled as Argonne), more realistic choice for the bound-state dynamics, we consider the deuteron

wavefunction associated with the Argonne v_{18} potential, which is constrained by nucleon-nucleon scattering data and the deuteron binding energy [36]. In this case one may write

$$\varphi_D(\vec{r}) = \frac{1}{\sqrt{4\pi r}} \left[u(r) + \frac{1}{\sqrt{8}} w(r) S_{12}(\hat{r}) \right] \chi_{1m}, \quad (21)$$

where $S_{12}(\hat{r}) = 3(\vec{\sigma}_1 \cdot \hat{r})(\vec{\sigma}_2 \cdot \hat{r}) - (\vec{\sigma}_1 \cdot \vec{\sigma}_2)$ is the tensor operator, σ_i are Pauli matrices, χ_{1m} is the spinor, and $u(r)$ and $w(r)$ are the radial S- and D-wave components. The normalization condition reads

$$\int d^3r |\varphi_D(\vec{r})|^2 = \int dr [u^2(r) + w^2(r)] = 1. \quad (22)$$

For this wave function the Wigner kernel $\mathcal{D}(r, q)$ does not admit a simple analytic expression; in our implementation we use a tabulated representation of $\mathcal{D}(r, q)$ as a function of (r, q) based on the fit provided in [26].

In our event-by-event implementation, we interpret $\mathcal{D}(\Delta r, \Delta p)$ as an acceptance weight in $(\Delta r, \Delta p)$ space. For every $\vec{p}-\vec{n}$ pair we evaluate $\mathcal{D}(\Delta r, \Delta p)$, draw a random number $u \in [0, 1]$, and form a $\bar{D}$ if u is smaller than the (properly normalized) weight. This is the same method employed in Ref. [35].

III. PYTHIA TUNING FOR (ANTI)NUCLEON PRODUCTION

A reliable prediction of light (anti)nuclei yields from secondary production requires, as a first step, an accurate Monte Carlo description of the underlying (anti)nucleon production in pp collisions. This is particularly important in our setup, where (anti)nuclei are formed by applying the coalescence prescription directly to the event record. Throughout this work we use PYTHIA 8.315 [45] as our baseline generator and implement the coalescence procedure at the Monte-Carlo level. In Ref. [35] we followed an analogous strategy for DM annihilation, by first tuning PYTHIA to hadron production data at Large Electron Positron collider (LEP) in e^+e^- collisions, which are commonly used as a proxy for the hadronization environment of DM annihilation into quarks.

The tuning strategy adopted here focuses on reproducing antiproton observables over the widest possible energy range. In principle, a complete validation would also require antineutron data. However, (anti)neutrons are notoriously difficult to measure in collider environments and no dedicated LHC measurements exist. A possible hint of an isospin asymmetry, with $\bar{n}/\bar{p} \simeq 1.2-1.3$ in fixed-target pp at $\sqrt{s} \simeq 17$ GeV, was reported by NA49 [46], but it has not been firmly established by subsequent publications. At high energies, inclusive multiplicities of p and $\bar{p}$ are comparable, as are those of charged and neutral pions [47], suggesting approximate isospin symmetry in soft production. For these reasons, and consistent with the default PYTHIA treatment (i.e. the Monash tune [48]),

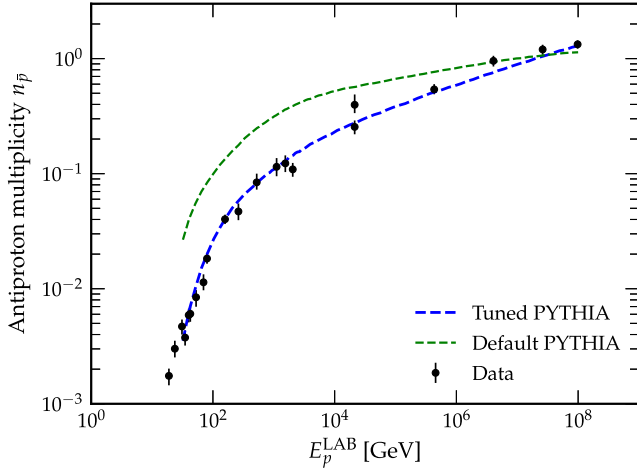


FIG. 1. Antiproton multiplicity $n_{\bar{p}} = N_{\bar{p}}/N_{\text{ev}}$. Antiproton multiplicity in pp collisions as a function of the (equivalent fixed-target) proton beam energy E_p^{LAB} , compiled from Refs. [49–54]. Black dots show the experimental measurements. The blue dashed curve is the prediction obtained with our energy-dependent PYTHIA 8.315 tune (see text), while the green dashed curve corresponds to the default PYTHIA 8.315 configuration.

we assume equal $\bar{p}$ and $\bar{n}$ yields in pp collisions in the remainder of this work.

The default PYTHIA configuration does not reproduce satisfactorily the low-energy $\bar{p}$ production measured in fixed-target experiments (e.g. NA61 and NA49), as illustrated by the green curve in Fig. 1. In particular, at $E_p^{\text{LAB}} \sim (30\text{--}200)$ GeV the default setup overpredicts $n_{\bar{p}}$ by up to about an order of magnitude, which would directly bias the predicted antinuclei yields in the region of low kinetic energy per nucleon K most relevant for DM searches. This motivates an energy-dependent tune that remains compatible with low-energy fixed-target data and high-energy collider measurements simultaneously.

We build our setup starting from the Tune 2M baseline [55], together with the prescriptions denoted as Mode 2 in Ref. [56]. On top of that, we introduce an explicit energy dependence for a small set of fragmentation parameters that are known to affect baryon production and the transverse-momentum spectra:

- (i) `StringPT:sigma`. This parameter controls the width of the Gaussian transverse kicks assigned at each string breaking. PYTHIA draws independent Gaussians in p_x and p_y with variance σ^2 , such that $\langle p_T^2 \rangle = 2\sigma^2$. We use

$$\begin{aligned} &\text{StringPT:sigma} \\ &= \begin{cases} 0.40, & E_p^{\text{LAB}} \leq 1 \text{ TeV}, \\ 0.40 + 0.035 \log_{10} \left(\frac{E_p^{\text{LAB}}}{1 \text{ TeV}} \right), & E_p^{\text{LAB}} > 1 \text{ TeV}. \end{cases} \end{aligned} \quad (23)$$

- (ii) `StringFlav:probQQtoQ`. This parameter sets the suppression of diquark production relative to quark production, and thus regulates baryon versus meson production. We adopt

$$\begin{aligned} &\text{StringFlav:probQQtoQ} \\ &= \begin{cases} 0.055 - 0.017 \log_{10}^2 \left(\frac{E_p^{\text{LAB}}}{1 \text{ TeV}} \right), & E_p^{\text{LAB}} \leq 1 \text{ TeV}, \\ 0.030 + 0.002 \log_{10} \left(\frac{E_p^{\text{LAB}}}{1 \text{ TeV}} \right), & E_p^{\text{LAB}} > 1 \text{ TeV}. \end{cases} \end{aligned} \quad (24)$$

- (iii) `StringFlav:probStoUD`. This parameter controls the suppression of strange-quark production relative to the up and down ones. We set it to 0.217 for $E_p^{\text{LAB}} < 1$ TeV and to 0.200 at higher energies.

Although the tune is energy dependent, we verified that the resulting hadron multiplicities (including pions and baryons) evolve smoothly with energy and do not show visible discontinuities across the transition at $E_p^{\text{LAB}} = 1$ TeV.

We validate the tune using two complementary classes of measurements: (i) the *multiplicity* $n_{\bar{p}} = N_{\bar{p}}/N_{\text{ev}}$ (average number of antiprotons per inelastic pp collision, integrated over transverse momentum and rapidity), which robustly constrains the overall normalization and its energy evolution; and (ii) differential multiplicity distributions (hereafter referred to as *spectra*), which probe the kinematic shapes relevant for coalescence (in particular as a function of transverse momentum p_T , rapidity y , or x_F).

Figure 1 shows $n_{\bar{p}}$ as a function of the incoming proton energy for both the default PYTHIA configuration and our tune. For collider measurements quoted at a center-of-mass energy $\sqrt{s}$, we report the equivalent fixed-target beam energy

$$E_p^{\text{LAB}} = \frac{s - 2m_p^2}{2m_p}, \quad (25)$$

so that fixed-target and collider results can be displayed on the same axis. We consider data from NA61, NA49, PHENIX, ALICE and older measurements [49–54], spanning many orders of magnitude in E_p^{LAB} .² The agreement between the data and our energy-dependent tune indicates that the global $\bar{p}$ yield is under control across the full energy range of interest.

A validation on a differential quantity is presented in Fig. 2. The top panel compares our prediction for $d^2n/(dp_T dy)$ to NA61 data at $\sqrt{s} = 7.74$ GeV [50] in several rapidity slices. The middle panel compares the

²ALICE measures $dn_{\bar{p}}/dy$ at midrapidity, $|y| < 0.5$. To compare with full-phase-space multiplicities, we rescale the ALICE results using the rapidity distribution predicted by our tuned PYTHIA configuration.

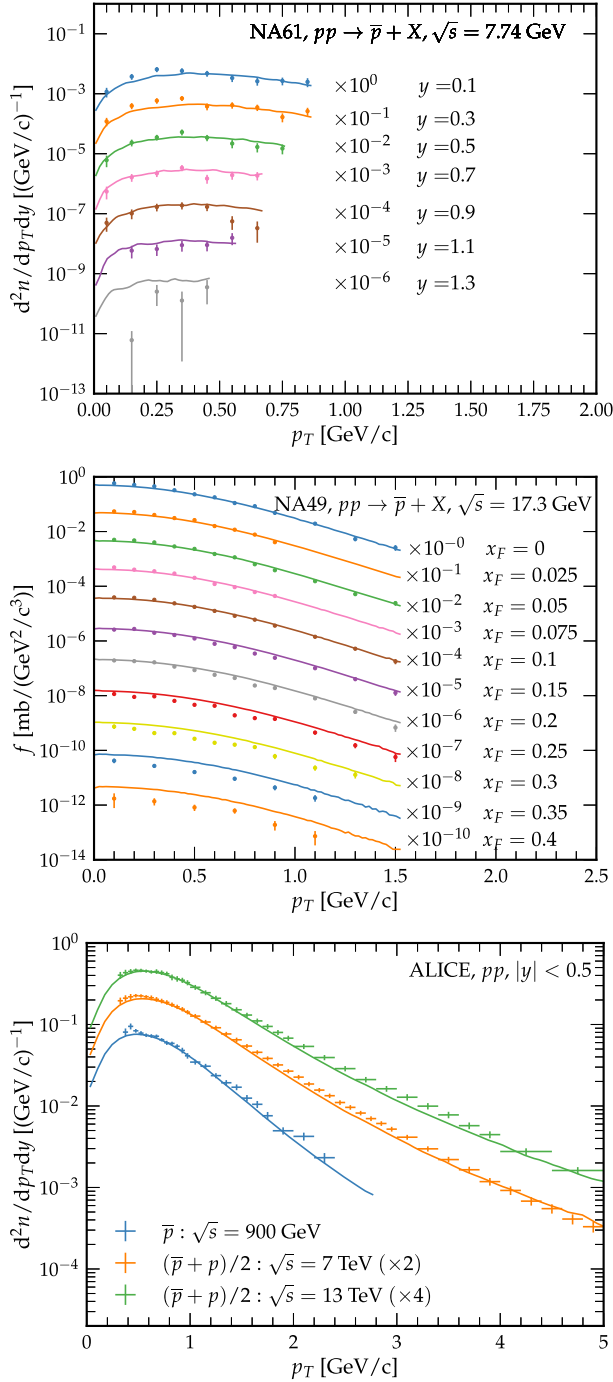


FIG. 2. Validation of our PYTHIA tune with antiproton spectra. Top: $d^2n/(dp_T dy)$ in fixed-target pp collisions at $\sqrt{s} = 7.74$ GeV (NA61) for several rapidity bins y ; spectra are rescaled by the factors indicated in the panel for readability [50]. Middle: invariant cross section $f = (1/\pi)d^3\sigma/(dp_T^2 dx_F)$ in fixed-target pp at $\sqrt{s} = 17.3$ GeV (NA49) for several x_F bins, again rescaled by the indicated powers of ten [51]. Bottom: midrapidity ($|y| < 0.5$) p_T spectra in pp collisions measured by ALICE at $\sqrt{s} = 0.9, 7$, and 13 TeV; the 7 and 13 TeV datasets are multiplied by factors 2 and 4 , respectively, as indicated in the legend [53,54,57]. In all panels, the points denote data and solid curves the predictions from our tuned PYTHIA configuration.

Lorentz-invariant cross section $f = (1/\pi)d^3\sigma/(dp_T^2 dx_F)$ to NA49 data at $\sqrt{s} = 17.3$ GeV [51] for a range of x_F values. Here y and x_F trace the longitudinal momentum p_L in the center-of-mass frame,

$$y \equiv \frac{1}{2} \ln \left(\frac{E + p_L}{E - p_L} \right), \quad x_F \equiv \frac{2p_L}{\sqrt{s}}. \quad (26)$$

We observe that the tune reproduces the overall shapes and normalizations well; residual discrepancies become more visible in the most forward bins (large y or x_F), where the experimental uncertainties also increase. Finally, the bottom panel compares our prediction to ALICE midrapidity spectra at $\sqrt{s} = 0.9, 7$, and 13 TeV [53,54,57], showing good agreement also in the high-energy regime.

In Appendix A, we provide additional comparisons for other kinematic projections. Overall, our tune yields a good description of the differential $\bar{p}$ spectra, especially at low $\sqrt{s}$ and low rapidity y . We also include comparisons of the $(p + \bar{p})/(\pi^+ + \pi^-)$ ratio, which show some tension at low p_T . This discrepancy suggests that the tune does not fully capture the NA61 measurements of the charged pion spectra. Even so, our tuning strategy is chiefly intended to model $\bar{p}$ production accurately, so as to provide robust predictions for $\bar{D}$ production in pp collisions.

IV. PYTHIA SETUP FOR ANTINUCLEI PRODUCTION

A. Antinuclei spatial and momentum distributions

In PYTHIA, the transition from partons to hadrons is described by the Lund string model: color confinement is modeled as a relativistic flux tube (a “string”) stretched between color-connected partons, with an (approximately) constant tension $\kappa \sim \mathcal{O}(1)$ GeV/fm. As the partons separate, potential energy builds up in the string until it becomes energetically favorable to break via nonperturbative $q\bar{q}$ pair creation (often interpreted as tunneling in the string field), producing a sequence of string breakups that are iterated from the endpoints toward the interior. Hadrons are formed by combining adjacent string pieces (and, in baryon production, diquark mechanisms), with longitudinal momentum sharing governed by the Lund fragmentation function and transverse momenta generated from a (roughly) Gaussian kick at each breakup [58,59]. Historically, this framework has been implemented primarily as a momentum-space model, while the space-time structure of hadron formation is necessarily more model dependent; in particular, production “vertices” should be understood as effective points inferred from the string worldsheet rather than directly observable formation points of extended bound states [60]. A concrete space-time prescription for multiparton strings, including practical vertex definitions for hadrons formed between two adjacent string breaks, is provided in PYTHIA following Ref. [60].

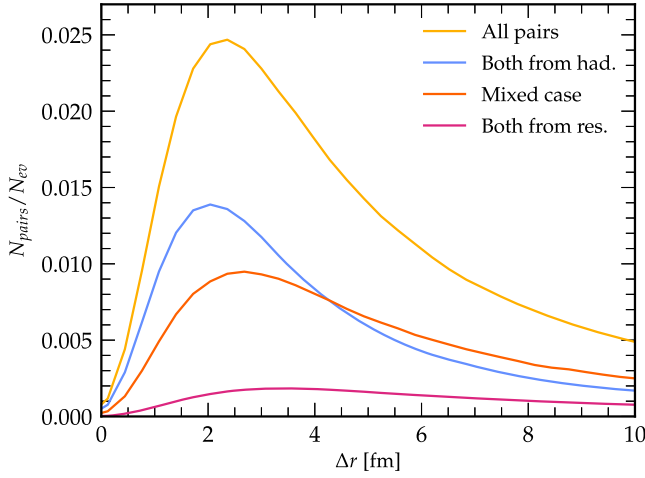


FIG. 3. Normalized distribution of the $\bar{p}$ - $\bar{n}$ pair separation Δr , in the pair reference frame, for antinucleons generated with our tuned PYTHIA setup in pp collisions at $\sqrt{s} = 7$ TeV. Contributions are shown separately for pairs in which both antinucleons originate from resonance decays (magenta), both originate from hadronization (blue), one originates from hadronization and the other from a resonance decay (orange), and for the inclusive sample of all $\bar{p}$ - $\bar{n}$ pairs (yellow).

The event record can be augmented with space-time information through two conceptually distinct switches:

- (i) `PartonVertex:setVertex = on` assigns production vertices to partons from the hard process and the parton shower, using formation-time/virtuality arguments such that these vertices typically lie on femtometer scales. If this option is enabled *without* enabling hadron-vertex generation, the produced hadrons inherit the space-time positions of their parent partons or string endpoints, and thus remain at $\mathcal{O}(1 \text{ fm})$ scales.
- (ii) `Fragmentation:setVertices = on` activates the Lund-model space-time construction of hadron production points from the string fragmentation process itself, i.e. hadron vertices are computed from the space-time locations of the relevant adjacent string breaks as in Ref. [60].

Because long strings in boosted topologies can extend significantly along the beam direction, hadron production vertices can reach much larger separations in the pair reference frame (especially in z), as illustrated, e.g., by the distributions shown in Fig. 4 of Ref. [61]. Since our goal is to use *hadron* production vertices (rather than inherited parent positions), we enable both switches.

Figure 3 shows the distribution of spatial separations $|\Delta \vec{r}|$ between $\bar{p}$ - $\bar{n}$ pairs, evaluated in the pair rest frame.³ Pairs produced directly at hadronization peak at $\Delta r \simeq 2 \text{ fm}$, whereas pairs in which both antinucleons originate from

³Here $\Delta \vec{r} \equiv \vec{r}_{\bar{p}} - \vec{r}_{\bar{n}}$ is constructed from the production vertices stored in the PYTHIA event record.

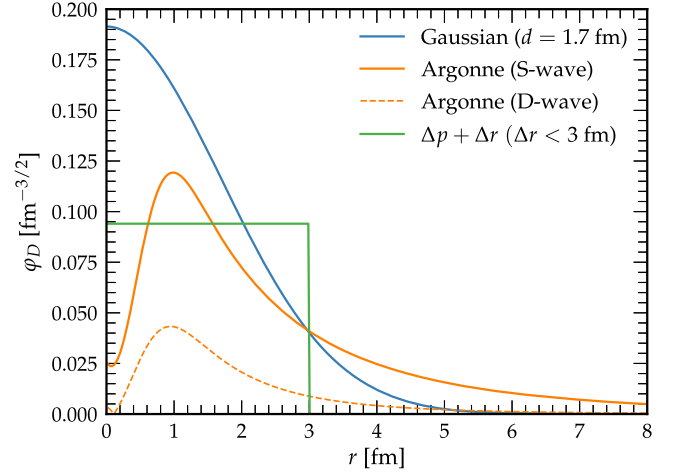


FIG. 4. Deuteron wave functions $\phi_D(r)$ used in the Wigner-based coalescence models: Gaussian ($d = 1.7 \text{ fm}$) S-wave and D-wave components (solid and dashed orange, respectively). The steplike profile corresponding to the $\Delta p + \Delta r$ model is also shown (green) for illustrative purposes, although it is not a Wigner-based prescription.

resonance decays peak at larger separations, $\Delta r \simeq 3.5 \text{ fm}$. The hadronization component is also noticeably narrower than the resonance-decay component. This implies that coalescence is generically less efficient for antinucleons produced via resonance decays, since they are typically emitted farther apart. The inclusive distribution peaks at $\Delta r \simeq 2.5 \text{ fm}$. We also note that the resonance-decay contribution is subdominant in PYTHIA compared to direct hadronization (see Sec. IV below).

The antinucleon separation distribution can be compared to the shapes of the deuteron wave functions entering the Wigner kernels. The Gaussian and Argonne wave functions, shown in Fig. 4, are concentrated within the inner few femtometers; the Gaussian is broader than the Argonne one. In particular, about 60% of the Argonne probability density is contained within $r < 3 \text{ fm}$. As a result, the Argonne kernel preferentially selects pairs produced at small separations, which in PYTHIA are more often associated with prompt hadronization, whereas the broader Gaussian kernel has a comparatively larger overlap with the resonance-decay tail.

B. Resonance and hadronic production of antinuclei

Recently, the ALICE Collaboration reported that approximately 80% of (anti)deuterons are formed through the coalescence of at least one (anti)nucleon originating from the decay of short-lived resonances [62]. In contrast, PYTHIA produces a significantly smaller fraction of (anti)nucleons from resonance decays. Using both the default PYTHIA settings and our $\bar{p}$ -tuned configuration (which does not modify the resonance production mechanism), we find that about 34% (22%) of antiprotons originate from resonance decays in pp collisions at $\sqrt{s} = 7 \text{ TeV}$ (17.3 GeV).

This discrepancy can be partially attributed to the fact that PYTHIA underestimates the multiplicities of strongly

TABLE I. Comparison between DELPHI and OPAL measured [68] and predicted multiplicities of selected resonances in e^+e^- collisions at the Z pole.

Resonance	DELPHI and OPAL	PYTHIA
$\Delta(1232)^{++}$	0.087 ± 0.033	0.059
$\Xi(1530)^0$	0.0059 ± 0.0011	0.002
$\Sigma(1385)^+$	0.0239 ± 0.0014	0.012

decaying resonances and does not include the full set of resonances listed by the PDG. This effect can be exacerbated when multiparton interactions and color reconnection are enabled, as in our setup [63]. In Table I we compare the predicted multiplicities of several resonances— $\Delta(1232)^{++}$, $\Xi(1530)^0$, and $\Sigma(1385)^+$ —with measurements performed by the DELPHI [64,65] and OPAL [66,67] Collaborations in e^+e^- collisions at the Z pole, as reported by the PDG [68]. Similarly, in Table II we compare the predicted multiplicities of $\Xi(1530)^0$, $\Sigma(1385)^\pm$, and their antiparticles with mid-rapidity measurements performed by ALICE in pp collisions at $\sqrt{s} = 7$ TeV [69]. In both cases, the PYTHIA predictions are systematically lower than the data, which at least partially explains why PYTHIA underestimates the resonance-decay contribution to (anti)nucleon production.

Several models of hadron production in high-energy collisions, in particular hadron resonance gas (HRG) models that treat hadrons as an ideal gas, predict a substantially larger fraction of nucleons originating from resonance decays, in closer agreement with the ALICE results than PYTHIA. Using the Thermal-FIST package [70], we find that in pp collisions at $\sqrt{s} = 13$ TeV, approximately 73% of nucleons are produced via resonance decays.

Since this result is directly constrained by experimental data only at LHC energies, its extension to lower collision energies remains speculative in the absence of additional measurements. Nevertheless, if HRG models provide an accurate description of hadron production in high-energy collisions, one would expect the fraction of nucleons originating from resonance decays to be only weakly dependent on the collision energy. Consistent with this expectation, Thermal-FIST yields a similar ratio for pp collisions at $\sqrt{s} = 900$ GeV.

In an attempt to reproduce a comparable fraction of nucleons from resonance decays within PYTHIA, we explored

TABLE II. Comparison between ALICE measured [62] and predicted multiplicities of selected resonances in pp collisions at midrapidity ($|y| < 0.5$).

Resonance	ALICE	PYTHIA
$\Sigma(1385)^+$	$(1.00^{+0.15}_{-0.14}) \times 10^{-2}$	$(3.30 \pm 0.06) \times 10^{-3}$
$\bar{\Sigma}(1385)^+$	$(1.03^{+0.17}_{-0.15}) \times 10^{-2}$	$(3.16 \pm 0.06) \times 10^{-3}$
$\Sigma(1385)^-$	$(1.08^{+0.17}_{-0.16}) \times 10^{-2}$	$(3.04 \pm 0.06) \times 10^{-3}$
$\bar{\Sigma}(1385)^-$	$(0.91^{+0.15}_{-0.14}) \times 10^{-2}$	$(2.91 \pm 0.06) \times 10^{-3}$
$\Xi(1530)^0$	$(2.56^{+0.41}_{-0.38}) \times 10^{-3}$	$(6.7 \pm 0.3) \times 10^{-4}$

discarding a large fraction of pp collision events in which antiprotons are produced directly during hadronization. This procedure naturally reduces the total $\bar{p}$ multiplicity, since PYTHIA predominantly produces nucleons through hadronization rather than resonance decays. To compensate for this effect, we increased the overall baryon production by tuning the `StringFlav:probQQtoQ` parameter. However, reproducing the experimentally measured $\bar{p}$ multiplicity would require setting this parameter beyond its maximum value allowed in the code. Consequently, within the current version of PYTHIA, it is not possible to reproduce the resonance-decay fraction reported by ALICE without introducing tensions with measurements of the total $\bar{p}$ multiplicity. For this reason, we retain our $\bar{p}$ -tuned PYTHIA configuration as the benchmark setup for the remainder of this paper.

V. DETERMINING THE COALESCENCE PARAMETERS

A. Fitting ALICE (anti)deuteron spectra

To calibrate the coalescence-model parameters we perform a simultaneous fit to the ALICE p_T -differential spectra of deuterons and antideuterons measured at $\sqrt{s} = 0.9, 2.76, 7$ TeV [28] and $\sqrt{s} = 13$ TeV [30]. All spectra are provided at midrapidity, $|y| < 0.5$. We then compare the resulting coalescence parameters with those extracted from the ALEPH $\bar{D}$ multiplicity in hadronic Z decays, as reported in Ref. [35], to test the universality of the coalescence description. We apply the analysis to the Δp , $\Delta p + \Delta r$ and Gaussian implementations. Since the Argonne setup does not contain any free parameters, as its wave function is fully fixed by nucleon-nucleon scattering data and the deuteron binding energy, we therefore only show its predictions.

As discussed in Ref. [35], coalescence prescriptions that include one parameter controlling the spatial distribution and one controlling the momentum dependence can exhibit a strong degeneracy. In the Gaussian setup, for instance, changes in the effective source size σ can be partially compensated by changes in the bound-state scale (here denoted by δ), leading to families of parameter pairs with similar goodness of fit. To reduce this degeneracy and keep the comparison across datasets and collision systems transparent, we fix the spatial scale to a benchmark value of $\sigma = 3$ fm, comparable to the deuteron size inferred from its rms charge radius [43]. Concretely, we adopt: (i) a single free parameter p_{coal} for the Δp model; (ii) a single free parameter p_{coal} for the $\Delta p + \Delta r$ model, fixing $r_{\text{coal}} \equiv \Delta r = 3$ fm; and (iii) a single free parameter δ for the Gaussian model, fixing the source size $\sigma = 3$ fm. The Argonne implementation is treated as parameter-free and is therefore not tuned to the data.

The baryon-emission source radius inferred from ALICE femtoscopia measurements in high-multiplicity pp collisions provides an important qualitative reference for the space-time structure relevant to antinuclei formation;

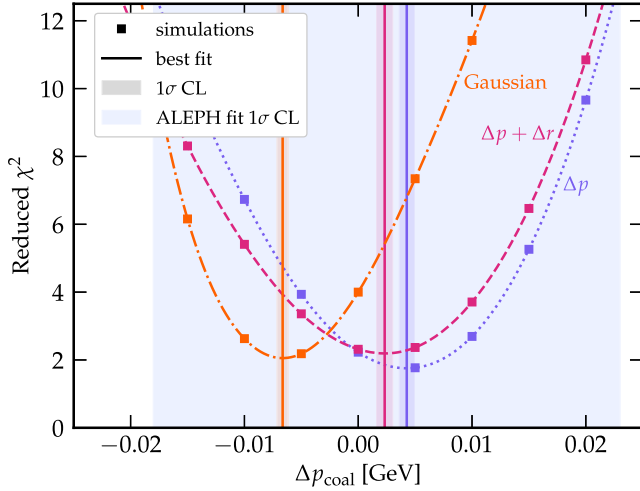


FIG. 5. Reduced χ^2 profiles as a function of $\Delta p_{\text{coal}} \equiv p_{\text{coal}} - p_{\text{coal}}^{\text{ALEPH}}$, where $p_{\text{coal}}^{\text{ALEPH}}$ is the best fit for the coalescence parameter found by fitting ALEPH data as reported in Ref. [35], while p_{coal} is the best fit obtained by fitting ALICE deuteron and antideuteron spectra measured at $\sqrt{s} = 0.9, 2.76, 7$, and 13 TeV. The squared markers indicate the χ^2 obtained when simulating the D and $\bar{D}$ data at specific values of p_{coal} . The curves represent the interpolations of the single χ^2 points. We show the results for the following models: Δp (dotted purple), $\Delta p + \Delta r$ (dashed magenta), and Gaussian (dash-dotted orange). In the Gaussian model we rescaled the δ parameter by a factor 0.11 GeV/1.80 fm. For each model, the vertical solid line and the correspondingly colored shaded band denote the extracted best-fit value and its 1σ interval, respectively. Additionally, the blue shaded band represents the 1σ interval obtained from the ALEPH fit for the Δp case, plotted for visual comparison.

however, since its extraction depends on the assumed source function, the treatment of final-state interactions, and corrections for resonance decays, it cannot be directly interpreted as a universal coalescence radius applicable to all event-by-event coalescence models [71].

We generate PYTHIA simulations of pp collisions to obtain a pool of antinucleons. For each coalescence model, we then scan a discrete grid of the relevant free parameter, applying the coalescence criteria to the same set of events. This procedure minimizes relative statistical fluctuations and isolates the dependence on the coalescence parameter. We compute the predicted spectra in the same kinematic bins as ALICE and evaluate a global χ^2 over all energies and over both D and $\bar{D}$ datasets. We then interpolate χ^2 as a

function of the scanned parameter to extract the best-fit value and its 1σ uncertainty. We simulated 100, 150, 450, and 450 million events for $\sqrt{s} = 0.9, 2.76, 7$, and 13 TeV, respectively. These sample sizes ensure that the Monte-Carlo relative statistical uncertainty remains below 5% for each experimental data point in p_T , for the parameter point closest to the best fit. For example, for the 7 and 13 TeV cases, the relative statistical uncertainty is $\approx 1\%$ in the lowest p_T bins and stays below 5% even in the highest p_T bins included in our analysis.

Figure 5 shows the χ^2 profiles resulting from the ALICE global fit. The derived best-fit values for the free parameter of each model are reported in Table III, which compares them to the corresponding values extracted from ALEPH e^+e^- data in Ref. [35]. Thanks to the small experimental uncertainties of the ALICE spectra and the large Monte-Carlo statistics, the ALICE-derived parameters p_{coal} and δ are determined at the subpercent level (relative uncertainties $\sim 0.5\%$). By contrast, the ALEPH-based calibration in Ref. [35] yields $\mathcal{O}(10\%)$ relative uncertainties, driven primarily by the larger experimental error on the measured $\bar{D}$ multiplicity ($\sim 15\%–20\%$).

We find good overall consistency between the parameters inferred from ALICE pp data and from hadronic Z decays. While the global-fit $\chi^2/N_{\text{d.o.f.}}$ values in Table III are of order 1.7–2.2, Table IV shows that excluding the $\sqrt{s} = 13$ TeV data significantly improves the fit quality ($\chi^2/N_{\text{d.o.f.}} \sim 0.73–0.91$) while shifting the best-fit parameters only mildly. This indicates that the extracted coalescence scales are stable, and that the residual tension is mainly associated with the highest-energy dataset (or, equivalently, with the generator-level description of the input (anti)nucleon kinematics at 13 TeV rather than with the coalescence prescription itself).

For the momentum-cut model we obtain $p_{\text{coal}} \simeq 0.20$ GeV, fully compatible with the ALEPH determination within uncertainties. When adding the spatial requirement ($\Delta p + \Delta r$), the best fit shifts to a slightly larger value, $p_{\text{coal}} \simeq 0.21$ GeV. This is expected: imposing $\Delta r < 3$ fm removes candidate $\bar{p}-\bar{n}$ pairs that would otherwise satisfy the momentum criterion only, and a larger p_{coal} compensates this reduction to reproduce the same $\bar{D}$ yield.

The agreement of the extracted parameters across (i) different production environments (hadronic pp at the LHC versus hadronic fragmentation in e^+e^- at the Z pole) and (ii) widely separated energies ($\sqrt{s} = 91$ GeV versus

TABLE III. Comparison of coalescence parameters extracted from pp and e^+e^- collisions. The ALICE results correspond to a simultaneous global fit of deuteron and antideuteron spectra across four center-of-mass energies ($\sqrt{s} = 0.9, 2.76, 7, 13$ TeV). The ALEPH results refer to antideuteron production in e^+e^- collisions at the Z resonance ($\sqrt{s} = 91$ GeV) and are taken from Ref. [35].

Model	ALICE (Global fit) pp @ 0.9–13 TeV	$\chi^2/N_{\text{d.o.f.}}$ (ALICE)	ALEPH (Reference) e^+e^- @ 91 GeV
Δp	$p_{\text{coal}} = 0.201 \pm 0.001$ GeV	1.75	$p_{\text{coal}} = 0.196^{+0.018}_{-0.023}$ GeV
$\Delta p + \Delta r$	$p_{\text{coal}} = 0.214 \pm 0.001$ GeV	2.19	$p_{\text{coal}} = 0.212^{+0.019}_{-0.024}$ GeV
Gaussian	$\delta = 1.69 \pm 0.01$ fm	2.05	$\delta = 1.80^{+0.2}_{-0.3}$ fm

TABLE IV. Fit quality excluding the highest-energy dataset. Same comparison as in Table III, but performing the ALICE global fit only for $\sqrt{s} = 0.9, 2.76$, and 7 TeV (thus excluding 13 TeV).

Model	ALICE (Lower $\sqrt{s}$ only) pp @ 0.9–7 TeV	$\chi^2/N_{\text{d.o.f.}}$ (ALICE)	ALEPH (Reference) e^+e^- @ 91 GeV
Δp	$p_{\text{coal}} = 0.198 \pm 0.001$ GeV	0.73	$p_{\text{coal}} = 0.196^{+0.018}_{-0.023}$ GeV
$\Delta p + \Delta r$	$p_{\text{coal}} = 0.211 \pm 0.001$ GeV	0.91	$p_{\text{coal}} = 0.212^{+0.019}_{-0.024}$ GeV
Gaussian	$\delta = 1.72 \pm 0.01$ fm	0.75	$\delta = 1.80^{+0.2}_{-0.3}$ fm

$\sqrt{s} \gtrsim 0.9$ TeV) supports a *universal* coalescence description at the level probed by current collider data. This universality is a key ingredient to reduce the source-model uncertainty in cosmic antinuclei predictions.

In Figs. 6 and 7 we compare the ALICE data with the best-fit predictions for $\bar{D}$ and D , respectively.

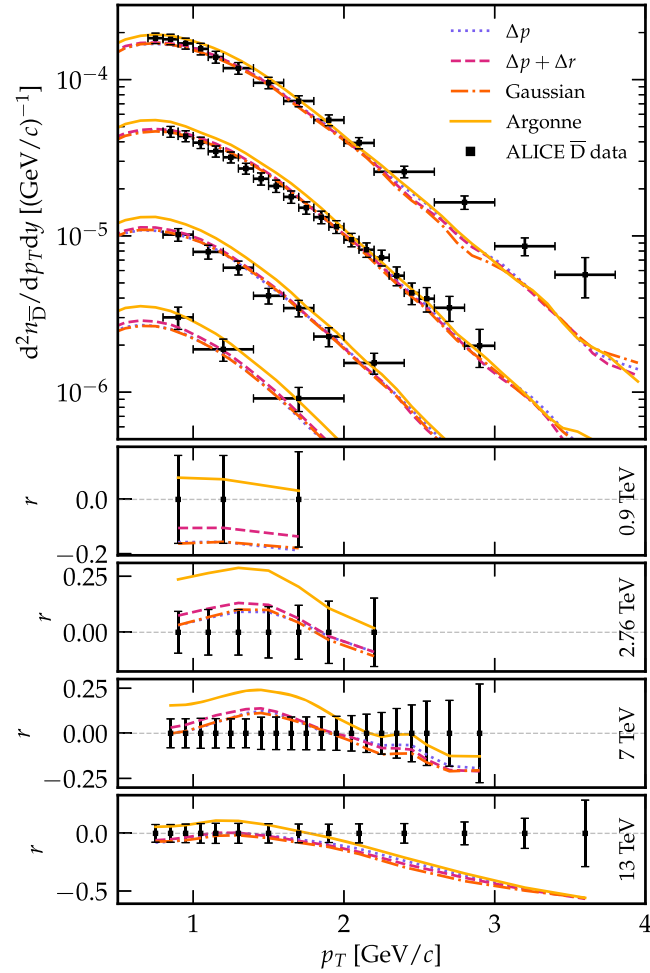


FIG. 6. Antideuteron p_T -differential spectra. The upper panel shows the ALICE $\bar{D}$ spectra at midrapidity ($|y| < 0.5$) for $\sqrt{s} = 13, 7, 2.76, 0.9$ TeV (from top to bottom, compared to best-fit predictions of the tuned coalescence models [Δp , $\Delta p + \Delta r$, Gaussian]) and to the parameter-free Argonne implementation (lines). For readability, the spectra (data and curves) are rescaled by factors 1, 1/3, 1/10, and 1/28, respectively. The four lower panels show the relative deviation $r = (\text{model} - \text{data})/\text{data}$ for each energy.

The three tuned setups (Δp , $\Delta p + \Delta r$, Gaussian) yield very similar differential multiplicities, typically differing by $\lesssim 10\%$ over the kinematic range covered by the data. This mirrors what was observed in the ALEPH-based calibration [35]: once tuned to a reference dataset, different reasonable coalescence implementations converge to nearly indistinguishable $\bar{D}$ spectra. For completeness, in Table V we provide the differential multiplicities dn/dy of (anti)deuterons at midrapidity obtained in our simulations compared to ALICE minimum-bias data.

The Argonne prediction, although not tuned, remains close to the tuned models, with deviations at the level of

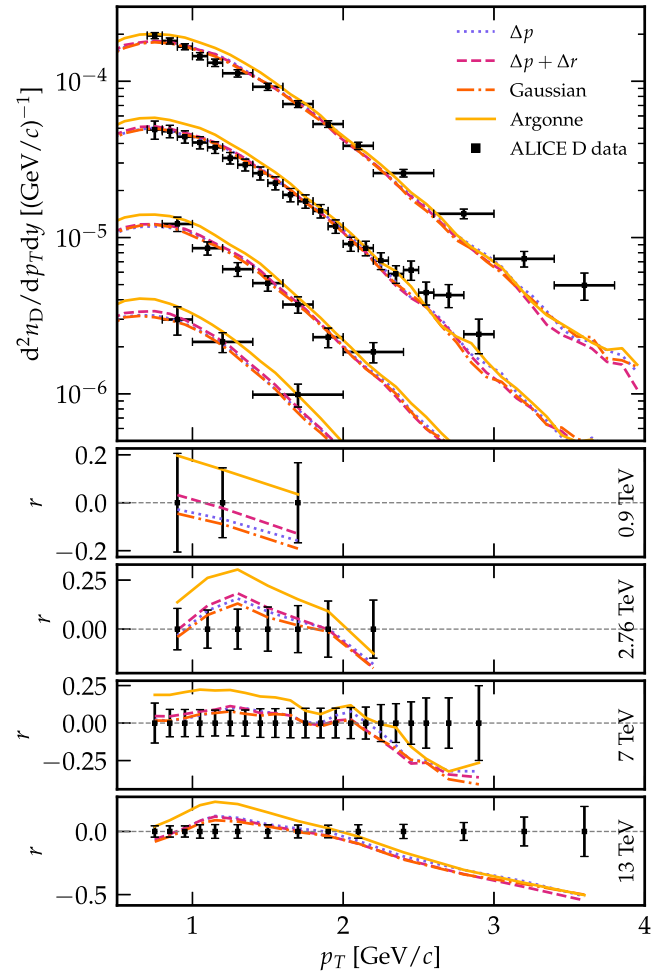


FIG. 7. Deuteron p_T -differential spectra. Same as Fig. 6, but for the ALICE D spectra at midrapidity.

TABLE V. Midrapidity yields of deuterons and antideuterons. Differential multiplicities dn/dy at midrapidity ($|y| < 0.5$) extracted from the simulations for the considered coalescence models, compared to ALICE minimum-bias data at $\sqrt{s} = 0.9, 2.76, 7$ [28], and 13 TeV [30]. Values are scaled by a factor 10^4 for readability.

Data and models	0.9 TeV	2.76 TeV	7 TeV	13 TeV
<i>Antideuterons ($\bar{D}$), $dn/dy \times 10^4$</i>				
ALICE	$1.11 \pm 0.10 \pm 0.09$	$1.37 \pm 0.04 \pm 0.12$	$1.92 \pm 0.02 \pm 0.15$	$2.54 \pm 0.03 \pm 0.30$
Δp	0.933 ± 0.005	1.434 ± 0.009	1.980 ± 0.008	2.46 ± 0.01
$\Delta p + \Delta r$	0.994 ± 0.005	1.480 ± 0.009	2.020 ± 0.008	2.48 ± 0.01
Gaussian	0.928 ± 0.005	1.437 ± 0.008	1.959 ± 0.008	2.42 ± 0.01
Argonne	1.208 ± 0.005	1.713 ± 0.009	2.252 ± 0.009	2.74 ± 0.01
<i>Deuterons (D), $dn/dy \times 10^4$</i>				
ALICE	$1.12 \pm 0.09 \pm 0.09$	$1.53 \pm 0.05 \pm 0.13$	$2.02 \pm 0.02 \pm 0.17$	$2.54 \pm 0.03 \pm 0.30$
Δp	1.097 ± 0.005	1.564 ± 0.009	2.084 ± 0.009	2.55 ± 0.01
$\Delta p + \Delta r$	1.155 ± 0.005	1.602 ± 0.009	2.116 ± 0.009	2.56 ± 0.01
Gaussian	1.077 ± 0.005	1.560 ± 0.009	2.054 ± 0.009	2.51 ± 0.01
Argonne	1.366 ± 0.006	1.818 ± 0.009	2.372 ± 0.009	2.83 ± 0.01

$\sim 20\%$ – 25% in the measured p_T window. We interpret this spread as a conservative estimate of the residual model uncertainty associated with the choice of deuteron wave function/Wigner kernel in our framework. Finally, the relative-deviation panels show that the largest discrepancies for all models tend to appear at the highest p_T , where the spectra contribute less to the total multiplicity than the low- p_T region, which dominates the integrated yield. Notably, the residual structure is largely correlated among models, suggesting that remaining differences are driven primarily by the underlying (anti)nucleon production in the generator rather than by the details of the coalescence prescription. This interpretation is consistent with the fact that our tuned PYTHIA configuration slightly underestimates the ALICE antiproton spectra in the range

$p_T \sim 1.5\text{--}4$ GeV/ c , which overlaps with the region where we also tend to underestimate the (anti)deuteron spectra.

B. Model predictions for antihelion-3

In addition to the (anti)deuteron spectra, we also examine the predictions of our PYTHIA tune for the ${}^3\bar{\text{He}}$ spectrum. In our model, a ${}^3\bar{\text{He}}$ nucleus is formed when the coalescence criteria are simultaneously satisfied for all three antinucleon pairs in a $\bar{p}\bar{p}\bar{n}$ configuration.

The ALICE Collaboration has released transverse-momentum-differential spectra of ${}^3\bar{\text{He}}$ at midrapidity for pp collisions at $\sqrt{s} = 7$ TeV [62] and 13 TeV [31]. Figure 8 shows a comparison between the ALICE measurement at $\sqrt{s} = 7$ TeV [62] and our predictions for all coalescence models considered, obtained from simulations of 5×10^9 events. Within the statistical uncertainties of both the experimental data and the Monte Carlo simulations, the predictions are consistent with the ALICE results. At large p_T the Monte-Carlo statistics are not sufficient to obtain a smooth theoretical spectrum; however, in this region the experimental uncertainties are also large.

VI. SECONDARY ANTIDEUTERON SOURCE TERM

Secondary antideuterons are produced in inelastic interactions of primary CRs with the atoms of the ISM. The dominant channels are CR p and He projectiles on H and He targets. The differential source term at Galactic position $\vec{x}$ (production rate per unit volume, time, and kinetic energy per n) is

$$q_{\bar{D}}^{\text{sec}}(K_{\bar{D}}, \vec{x}) = \sum_{i,j} \int_{K_{i,\text{th}}}^{\infty} dK_i 4\pi n_{\text{ISM},j}(\vec{x}) \phi_i(K_i, \vec{x}) \times \frac{d\sigma_{ij \rightarrow \bar{D}}}{dK_{\bar{D}}}(K_i, K_{\bar{D}}), \quad (27)$$

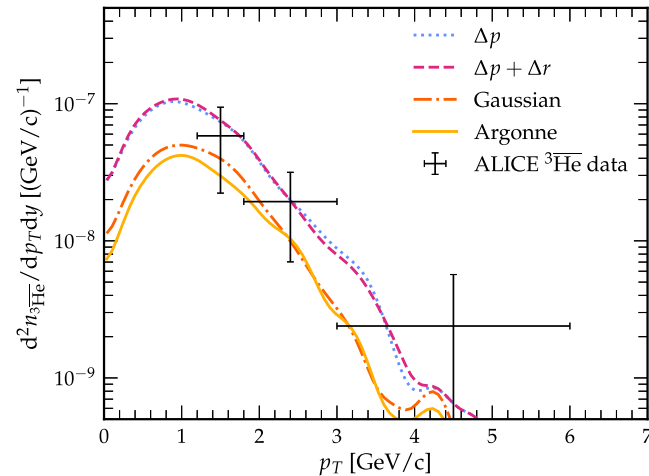


FIG. 8. Antihelion-3 p_T -differential multiplicity. ALICE ${}^3\bar{\text{He}}$ spectrum at midrapidity ($|y| < 0.5$) for $\sqrt{s} = 7$ TeV, compared to the best-fit predictions of the tuned coalescence models (Δp , $\Delta p + \Delta r$, Gaussian) and to the parameter-free Argonne implementation.

where K_i and $K_{\bar{D}}$ denote the kinetic energies per n (in GeV/ n) of the projectile i and of the produced antideuteron, respectively; $n_{\text{ISM},j}$ is the number density of the ISM target species j , and ϕ_i is the interstellar (demodulated) differential CR flux per unit area, time, solid angle, and kinetic energy per n . The sums run over the relevant projectiles and targets; in practice $i \in \{p, \text{He}\}$ and $j \in \{\text{H}, \text{He}\}$ dominate, while channels induced by secondary $\bar{p}$ are subleading. In a disk-averaged treatment one may adopt constant densities $n_{\text{ISM,H}} \simeq 1 \text{ cm}^{-3}$ and $n_{\text{ISM,He}} \simeq 0.1 \text{ cm}^{-3}$, while more refined calculations can use spatially dependent gas maps.

The lower limit $K_{i,\text{th}}$ in Eq. (27) accounts for the kinematic threshold for antideuteron production implied by baryon-number conservation. For fixed-target pp collisions, antideuteron production requires $\sqrt{s} \gtrsim 6m_p$, which corresponds to a threshold projectile total energy $E_{p,\text{th}} \simeq (s - 2m_p^2)/(2m_p) \gtrsim 17m_p$ thus a threshold kinetic energy per n $K_{p,\text{th}} \equiv E_{p,\text{th}} - m_p \gtrsim 16m_p$. This energy is substantially higher than for antiproton production and is the main reason why the secondary $\bar{D}$ spectrum is strongly suppressed at low $K_{\bar{D}}$, i.e. in the energy window most relevant for DM searches. In addition, coalescence selects $\bar{p}$ - $\bar{n}$ pairs with small relative momentum, so that the production of very energetic antinucleons further suppresses the probability to form a $\bar{D}$.

In our framework, the differential production cross section $d\sigma_{ij \rightarrow \bar{D}}/dK_{\bar{D}}$ is obtained from a Monte Carlo description of the underlying antinucleon production followed by an event-by-event coalescence prescription. It is convenient to express it as

$$\frac{d\sigma_{ij \rightarrow \bar{D}}}{dK_{\bar{D}}}(K_i, K_{\bar{D}}) = \sigma_{ij}^{\text{inel}}(K_i) \frac{dN_{\bar{D}}}{dK_{\bar{D}}}(K_i, K_{\bar{D}}), \quad (28)$$

where $dN_{\bar{D}}/dK_{\bar{D}}$ denotes the $\bar{D}$ multiplicity per inelastic ij interaction and $\sigma_{ij}^{\text{inel}}$ is the inelastic cross section for the CR projectile i colliding against the ISM target j . The quantity $dN_{\bar{D}}/dK_{\bar{D}}$ is obtained from our Monte-Carlo implementation of the coalescence process, evaluated on a grid of 200 incoming proton energies between ~ 25 and 10^7 GeV .⁴

For the inelastic cross sections entering Eq. (28), we approximate the proton-nucleus case using an empirical mass-scaling law,

$$\sigma_{pX}^{\text{inel}}(K_i) = A_X^{0.8} \sigma_{pp}^{\text{inel}}(K_i), \quad (29)$$

motivated by parametrizations of the form $\sigma_{A_{\text{proj}}, A_{\text{targ}}}^{\text{inel}} \simeq \sigma_{pp}^{\text{inel}}(A_{\text{proj}} A_{\text{targ}})^{0.8}$ (see, e.g., Refs. [72,73]). We use the

⁴In practice, for the He-induced channels we use the same multiplicity tables as a function of the projectile kinetic energy per n , consistently with the superposition approximation adopted for the hadronic interaction.

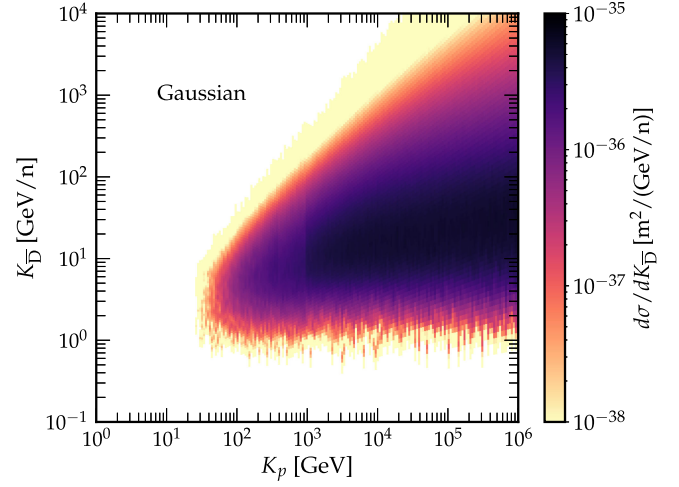


FIG. 9. Differential antideuteron production cross section $d\sigma/dK_{\bar{D}}$ in pp collisions (color scale) in the laboratory frame as a function of the projectile and produced kinetic energies per nucleon ($K_p, K_{\bar{D}}$). The result is computed with the Gaussian Wigner coalescence model.

parameterization of $\sigma_{pp}^{\text{inel}}$ reported in Appendix A of Ref. [73], and we rewrite it in Appendix B for completeness.

Figure 9 shows the differential $\bar{D}$ production cross section in pp collisions, computed with the Gaussian Wigner coalescence model, as a function of the projectile and produced kinetic energies per n , ($K_p, K_{\bar{D}}$). The cross section vanishes below the kinematic threshold, visible as the sharp cutoff for $K_p \lesssim 20 \text{ GeV}/n$, as expected from baryon-number conservation. This kinematic threshold actually almost coincides with the projectile kinetic energy associated to the lowest center-of-mass energy allowed by PYTHIA. The implementation of these cross sections in the calculation of the antideuteron flux at Earth is discussed in Ref. [74]. The differential cross-sections were computed for $\bar{p}$ and $\bar{D}$ (for the different coalescence models), and are publicly available on the Zenodo repository [75].

Although the production cross section increases with the projectile energy, the CR flux decreases approximately as a power law, $\phi_p \propto K_p^{-2.7}$, implying a suppression by about $10^{2.7}$ per decade in energy. Therefore, the convolution in Eq. (27) enhances the relative importance of projectile energies below the multi-TeV region. The resulting secondary source term is shown in Fig. 10 for the different collision channels (p -H, p -He, He-H, He-He) and for the different coalescence prescriptions. The source term peaks at $K_{\bar{D}} \sim$ a few GeV/ n , and the p -H contribution dominates because of the larger H target density, while the He-He contribution is suppressed [at the level of $\mathcal{O}(30)$] by both the smaller He abundance and the smaller He flux.

Figure 10 also shows that the Δp model predicts a larger number of secondary $\bar{D}$ compared to the other coalescence models. In the ALICE analyses, antiprotons originating from weakly decaying hadrons (e.g., Λ baryons) are removed from

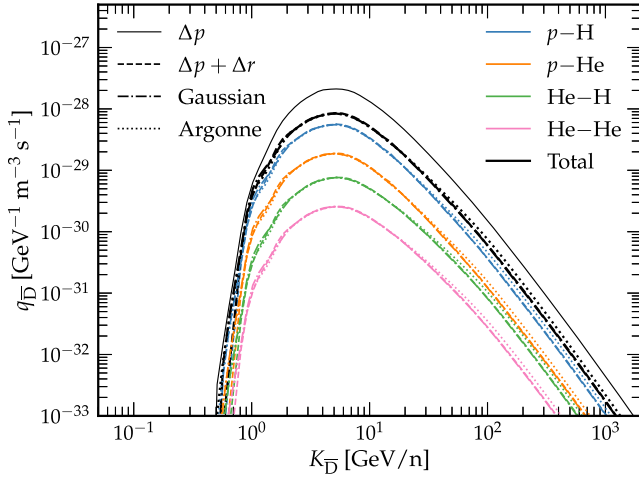


FIG. 10. Secondary antideuteron source term $q_{\bar{D}}(K_{\bar{D}})$ for secondary $\bar{D}$ production, split by collision channel (p -H, p -He, He-H, He-He) and total (black). For each channel, different line styles correspond to different coalescence prescriptions: $\Delta p + \Delta r$ (dashed), Gaussian (dot-dashed), Argonne (dotted). For the Δp model, we show only the total source term (solid).

the sample. This is typically called feed-down correction. In the simulations used for the source term evaluation, however, such antinucleons must be retained, as they can contribute to total secondary $\bar{D}$ production in the Galaxy. Antinucleons produced in the decays of weakly decaying hadrons typically originate from displaced vertices and are therefore much more spatially separated from other antinucleons than those produced directly through hadronization or strong resonance decays. In particular, there is basically no chance to make $\bar{D}$ if one antinucleon comes from a weakly decaying hadron and the other is generated from another particle decay or is primordial. As a result, coalescence models that include a spatial criterion, such as the $\Delta p + \Delta r$ model, effectively remove this possible contribution. In contrast, the Δp model applies only a momentum-space criterion, so this possibility remains, leading to an enhanced, biased secondary $\bar{D}$ yield. Consequently, special care must be taken when applying the Δp model to $\bar{D}$ production in astrophysical environments (e.g., DM annihilation or secondary production in cosmic-ray interactions), as it may lead to the formation of a non-negligible fraction of nonphysical $\bar{D}$.

VII. DISCUSSION AND CONCLUSIONS

In this work we studied the production of light antinuclei in hadronic collisions with the goal of reducing the dominant source-term uncertainty in CR antideuteron (and, prospectively, antihelium) searches for DM. The key difficulty is that antinucleus formation is intrinsically nonperturbative: it depends both on the underlying antinucleon production in QCD hadronization and on the subsequent coalescence of $\bar{p}$ and $\bar{n}$ into a bound state. Our strategy was therefore to (i) calibrate the antinucleon

production with collider data over a wide energy range and (ii) test a set of increasingly realistic coalescence prescriptions, anchoring their parameters to accelerator measurements in different production environments.

On the generator side, we employed PYTHIA 8.315 as baseline and implemented antinucleus formation directly on the event record. Since coalescence is sensitive to antinucleon correlations and, for models with spatial information, to production vertices, it is crucial that the generator reproduces the relevant single-particle observables. For this reason, the tune was validated against a compilation of antiproton multiplicities across fixed-target and collider energies, as well as against differential p_T multiplicities (and invariant cross sections) from NA61, NA49, and ALICE. The outcome is a generator setup that reproduces both the global antiproton multiplicity and spectra, and the key kinematic shapes over more than three orders of magnitude in the equivalent beam energy, which is essential for reliable predictions of low- K antinuclei yields relevant for DM searches. In the absence of dedicated LHC antineutron measurements, we adopted isospin symmetry for $\bar{p}$ and $\bar{n}$ production at high energy, consistent with inclusive hadron data and with the default PYTHIA treatment.

We then compared several coalescence implementations that span the range of common approaches in the literature. First, we considered the standard event-by-event “phase-space cutoff” prescriptions: the Δp model, where a $\bar{p}$ - $\bar{n}$ pair forms an antideuteron if its relative momentum in the pair center-of-mass frame satisfies $|\Delta\vec{p}| < p_{\text{coal}}$, and the $\Delta p + \Delta r$ extension, where one additionally requires a spatial proximity cut $|\Delta\vec{r}| < r_{\text{coal}}$. These models retain dynamical correlations from the generator and, in the Δr case, suppress coalescence between promptly produced antinucleons and those originating from displaced vertices (e.g., weak decays), which are naturally separated in space-time. Second, we implemented quantum-mechanical coalescence in the Wigner formalism, where the formation probability is obtained by projecting the two-particle phase-space distribution onto the bound-state Wigner function. We studied both a Gaussian benchmark (controlled by an effective momentum/size scale δ) and a more predictive Argonne implementation based on the Argonne v_{18} deuteron wavefunction, constrained by nucleon-nucleon scattering data and the deuteron binding energy.

The core calibration step of this paper is a simultaneous fit to the ALICE p_T -differential multiplicities of deuterons and antideuterons measured at midrapidity ($|y| < 0.5$) at $\sqrt{s} = 0.9, 2.76, 7$ and 13 TeV. We performed global χ^2 fits on a grid of the relevant coalescence parameter for each model, comparing the Monte Carlo predictions in the same kinematic bins as the data. To avoid parameter degeneracies that are known to arise when both a source-size parameter and a bound-state parameter are left free (especially in Gaussian/Wigner setups), we fixed the spatial scale to a benchmark 3 fm, comparable to the deuteron size inferred from its charge radius, and fitted only a single parameter per model: p_{coal} for Δp and $\Delta p + \Delta r$ (with $r_{\text{coal}} = 3$ fm),

and δ for the Gaussian/Wigner setup (with $\sigma = 3$ fm). In addition to the ALICE calibration, we compared the extracted parameters to those obtained in Ref. [35] from the ALEPH measurement of the $\bar{D}$ multiplicity in hadronic Z decays at $\sqrt{s} = m_Z$, an environment often used as a proxy for DM annihilation into quarks.

A central result is that the coalescence scales extracted from ALICE pp data are consistent with those inferred from ALEPH within uncertainties, despite the very different underlying environments and center-of-mass energies. Quantitatively, we find $p_{\text{coal}} \simeq 0.20$ GeV for the Δp model and a slightly larger value, $p_{\text{coal}} \simeq 0.21$ GeV, when a spatial cut is imposed, as expected because the Δr requirement removes a fraction of pairs that would otherwise coalesce. In the Gaussian setup, the best-fit scale is $\delta \simeq 1.7$ fm, close to the value obtained from ALEPH within the larger uncertainties of the LEP measurement. The ALICE-based determination achieves subpercent statistical precision in the fitted parameter (driven by the small experimental errors and large Monte-Carlo samples), while the ALEPH calibration remains at the $\mathcal{O}(10\%)$ level because of the larger uncertainty on the measured multiplicity. Importantly, once tuned, the Δp , $\Delta p + \Delta r$, and Gaussian prescriptions lead to very similar antideuteron spectra, typically differing by $\lesssim 10\%$ over the measured p_T range. The Argonne prediction, although parameter-free and not tuned to ALICE, remains compatible at the $\sim 20\%$ – 25% level. We interpret this spread as a conservative estimate of the residual model dependence associated with the choice of bound-state wavefunction/Wigner kernel, once the antinucleon production is properly calibrated.

We also find that the global fit quality improves significantly when excluding the $\sqrt{s} = 13$ TeV dataset, while the best-fit parameters shift only mildly. This pattern indicates that the coalescence scales are stable and that the remaining tension is more plausibly linked to the generator-level description of (anti)nucleon kinematics at the highest energy (notably at higher p_T) than to the coalescence prescription itself. The correlated residual structure across coalescence models supports this interpretation and motivates future work combining (i) further refinements of the high-energy baryon-production tune and (ii) additional differential constraints (e.g., multiplicity- or event-activity-dependent light-nuclei measurements) to stress-test the universality of the coalescence description beyond the inclusive midrapidity spectra used here.

Reference [4] likewise found that a conventional coalescence model implemented within the Wigner formalism can describe the ALICE and ALEPH data for $\bar{D}$. Nevertheless, that analysis did not make use of the full event-by-event spatial and momentum information of the antinucleon pairs produced by the event generator. Instead, it employed an integrated version of the Wigner formalism, see Eq. (10), based on the Monte Carlo antinucleon momentum distributions and an effective treatment of the source size, thereby averaging over the detailed spatial structure of the $\bar{p}$ – $\bar{n}$ emission process.

Finally, we outlined how the calibrated event-by-event framework connects to the secondary antideuteron source term in the Galaxy: the production rate is obtained by convolving the interstellar CR fluxes with ISM gas densities and the differential production cross sections, which in our Monte-Carlo approach can be written as $d\sigma_{ij \rightarrow \bar{D}}/dK_{\bar{D}} = \sigma_{ij}^{\text{inel}} dN_{\bar{D}}/dK_{\bar{D}}$. The collider-calibrated coalescence parameters derived in this work therefore provide a controlled and transferable input for the antideuteron production multiplicities entering astrophysical calculations. We however stress that the Δp model would generate a non-negligible fraction of nonphysical secondary $\bar{D}$.

In conclusion, by jointly confronting coalescence models with ALICE pp spectra and ALEPH hadronic- Z data, we provide quantitative evidence that a common coalescence description can be applied across widely different energies and production mechanisms. This substantially strengthens the physical basis of antinucleus source-term predictions and improves the robustness of interpreting forthcoming searches with AMS-02 and GAPS, where even a handful of low- K antideuteron candidates could carry decisive information about particle DM.

ACKNOWLEDGMENTS

M. D. M., F. D., N. F., J. K., and L. S. acknowledge support from the research grant *Theoretical Astroparticle Physics (TAsP)* funded by Istituto Nazionale di Fisica Nucleare (INFN). M. D. M. and J. K. acknowledge support from the Italian Ministry of University and Research (MUR), PRIN 2022 “EXSKALIBUR—Euclid-Cross-SKA: Likelihood Inference Building for Universe’s Research,” Grant No. 20222BBYB9, CUP I53D23000610 0006, and from the European Union—Next Generation EU. J. K. acknowledges support from the Italian Space Agency through the ASI INFN Grant Agreement No. 2018-28-HH.0: “Partecipazione italiana al GAPS—General AntiParticle Spectrometer.” N. F. is supported by the Italian Ministry of University and Research (MUR) via the PRIN 2022 Project No. 20228WHTYC—CUP: D53C24003550006 and from the European Union—Next Generation EU. F. B. acknowledges support from the European Research Council under the European Union’s Horizon 2020 research and innovation program through the ERC-H2020-STG CosmicAntiNuclei project (Grant Agreement No. 950692).

DATA AVAILABILITY

The data that support the findings of this article are openly available [75].

APPENDIX A: ADDITIONAL PYTHIA TUNING DETAILS

In the left panels of Fig. 11 we compare the predicted antiproton transverse momentum p_T spectra for a range of rapidities y with the measurements carried by NA61, while in the right panels we do the same for the p_T spectra of the

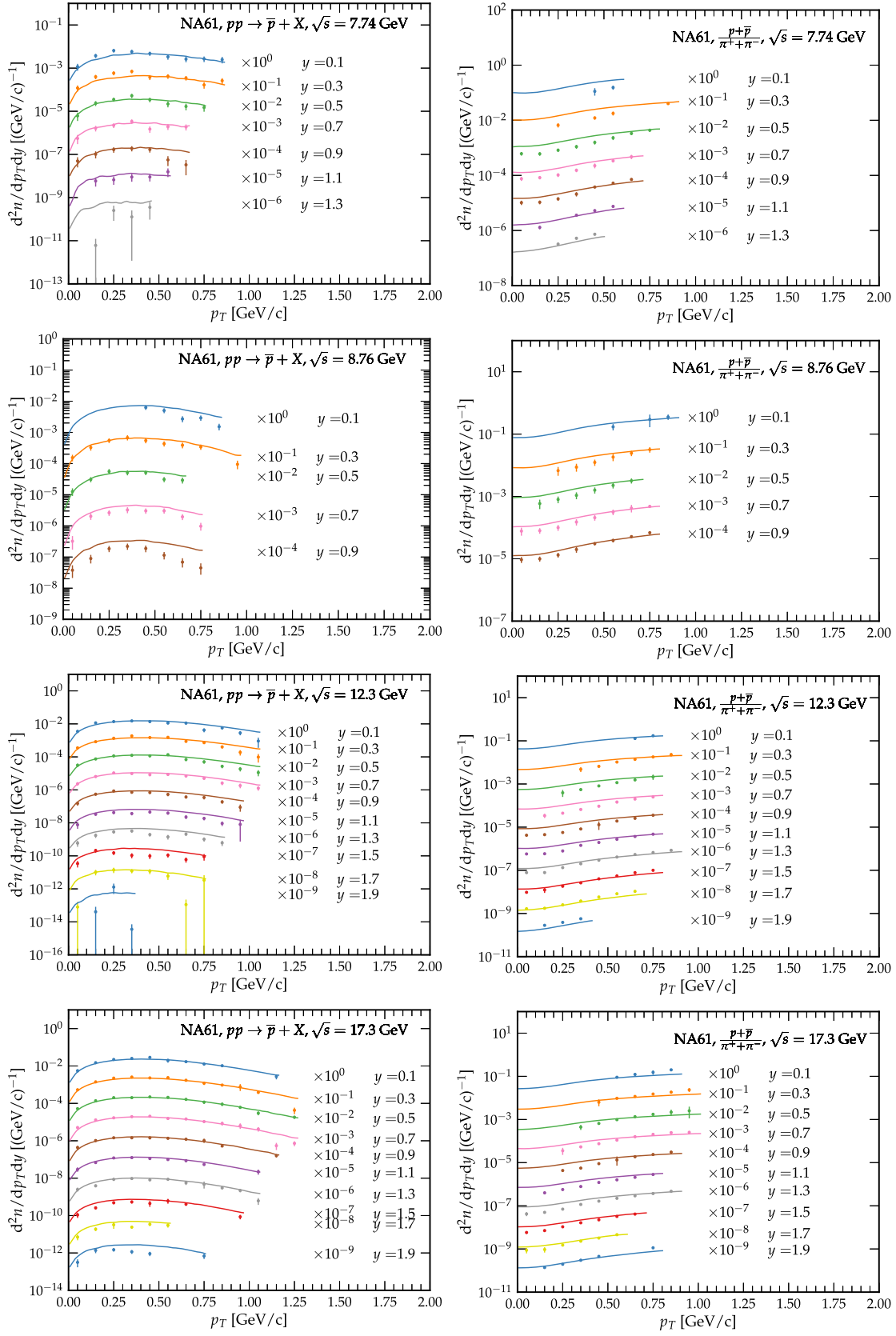


FIG. 11. Transverse momentum spectra of antiprotons (left panels) and proton-over-pion ratios (right panels) generated with our tuning of PYTHIA, compared to NA61 data [50] at different y in fixed-target pp collisions at $\sqrt{s} = 7.74$ GeV (top panels), 8.76 GeV (middle top panels), $\sqrt{s} = 12.3$ GeV (middle bottom panels), and 17.3 GeV (bottom panels).

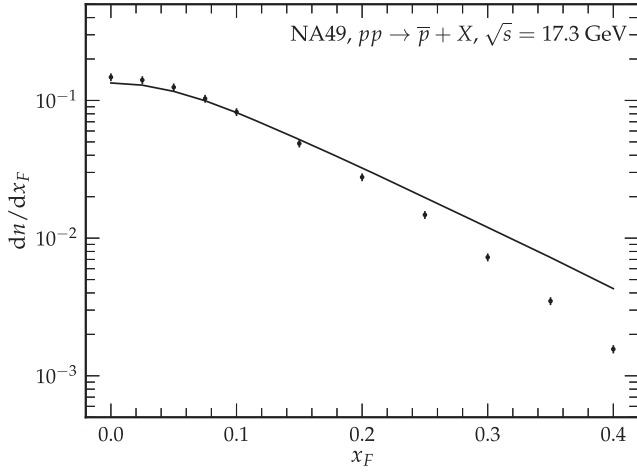


FIG. 12. p_T -integrated x_F -spectrum dn/dx_F of antiprotons generated with our tuning of PYTHIA, compared to NA49 data [51], in fixed-target pp collisions at $\sqrt{s} = 17.3$ GeV.

ratio between protons and pions $(p + \bar{p})/(\pi^+ + \pi^-)$. The panels show, from top to bottom, results at $\sqrt{s} = 7.74$ GeV ($E_p^{\text{LAB}} = 31$ GeV), $\sqrt{s} = 8.76$ GeV ($E_p^{\text{LAB}} = 40$ GeV), $\sqrt{s} = 12.3$ GeV ($E_p^{\text{LAB}} = 80$ GeV), and $\sqrt{s} = 17.3$ GeV ($E_p^{\text{LAB}} = 158$ GeV). In Fig. 12, we compare the predicted p_T -integrated antiproton x_F spectrum dn/dx_F with the measurements reported by NA49 at $\sqrt{s} = 17.3$ GeV ($E_p^{\text{LAB}} = 158$ GeV).

These plots confirm what was already stated in Sec. III, which is that our PYTHIA tuning predicts well the measured properties of antiprotons produced in pp collisions across a broad range of input proton energies E_p^{LAB} .

TABLE VI. Parameters for the total and elastic proton scattering cross sections according to Eq. (B1).

Parameter	Total	Elastic
M	1.589	3.094
Z^{pp}	59.58	21.34
Y_1^{pp}	0.890	2.667
Y_2^{pp}	19.35	14.21
η_1	2.543	1.003
η_2	-0.0895	-0.0327

APPENDIX B: ADDITIONAL CROSS SECTION DETAILS

In this appendix, we report the parametrization of the total inelastic cross section for pp collisions, as calculated in Ref. [73]. This is done by first parametrizing the total collision (σ_{tot}^{pp}) and elastic (σ_{el}^{pp}) cross sections where the following function is used, which depends on the Mandelstam variable s ,

$$\sigma_{\text{tot/el}}^{pp}(s) = Z^{pp} + B^{pp} \log^2(s/s_M) + Y_1^{pp}(s_M/s)^{\eta_1} - Y_2^{pp}(s_M/s)^{\eta_2}, \quad (\text{B1})$$

where $B^{pp} = \pi(\hbar c)^2/M^2$, $s_M = (2m_p + M)^2$ and the other parameters are summarized in Table VI. The resulting inelastic cross section is then simply $\sigma_{\text{inel}}^{pp}(s) = \sigma_{\text{tot}}^{pp}(s) - \sigma_{\text{el}}^{pp}(s)$.

-
- [1] F. Donato, N. Fornengo, and P. Salati, *Phys. Rev. D* **62**, 043003 (2000).
 - [2] N. Fornengo, L. Maccione, and A. Vittino, *J. Cosmol. Astropart. Phys.* **09** (2013) 031.
 - [3] M. Korsmeier, F. Donato, and N. Fornengo, *Phys. Rev. D* **97**, 103011 (2018).
 - [4] M. Kachelrieß, S. Ostapchenko, and J. Tjemsland, *J. Cosmol. Astropart. Phys.* **08** (2020) 048.
 - [5] L. Šerkšnytė *et al.*, *Phys. Rev. D* **105**, 083021 (2022).
 - [6] P. De La Torre Luque, M. W. Winkler, and T. Linden, *J. Cosmol. Astropart. Phys.* **10** (2024) 017.
 - [7] K. Abe *et al.*, *Phys. Rev. Lett.* **108**, 131301 (2012).
 - [8] O. Adriani *et al.* (PAMELA Collaboration), *Riv. Nuovo Cimento* **40**, 473 (2017).
 - [9] M. Aguilar *et al.* (AMS Collaboration), *Phys. Rep.* **894**, 1 (2021).
 - [10] G. Osteria (GAPS Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **958**, 162201 (2020).
 - [11] P. von Doetinchem *et al.*, *J. Cosmol. Astropart. Phys.* **08** (2020) 035.
 - [12] K. Sakai *et al.* (BESS Collaboration), *Phys. Rev. Lett.* **132**, 131001 (2024).
 - [13] V. Choutko and F. Giovacchini, in *International Cosmic Ray Conference*, International Cosmic Ray Conference (2008), Vol. 4, pp. 765–768.
 - [14] T. Aramaki, C. J. Hailey, S. E. Boggs, P. von Doetinchem, H. Fuke, S. I. Mognet, R. A. Ong, K. Perez, and J. Zweerink (GAPS Collaboration), *Astropart. Phys.* **74**, 6 (2016).
 - [15] P. von Doetinchem *et al.*, *Proc. Sci. ICRC2015* (2016) 1218 [arXiv:1507.02712].
 - [16] C. Hailey (GAPS Collaboration), in *Invited Talk at the Workshop Overview* (, 2026), <https://indico.cern.ch/event/1480110/>.
 - [17] J. I. Kapusta, *Phys. Rev. C* **21**, 1301 (1980).
 - [18] S. T. Butler and C. A. Pearson, *Phys. Rev.* **129**, 836 (1963).
 - [19] R. Scheibl and U. W. Heinz, *Phys. Rev. C* **59**, 1585 (1999).

- [20] M. Di Mauro, A. Jueid, J. Koechler, and R. R. de Austri, *Phys. Rev. D* **112**, 083017 (2025).
- [21] F. Bellini and A. P. Kalweit, *Phys. Rev. C* **99**, 054905 (2019).
- [22] K. Blum, K. C. Y. Ng, R. Sato, and M. Takimoto, *Phys. Rev. D* **96**, 103021 (2017).
- [23] F. Bellini, K. Blum, A. P. Kalweit, and M. Puccio, *Phys. Rev. C* **103**, 014907 (2021).
- [24] M. Kachelrieß, S. Ostapchenko, and J. Tjemsland, *Eur. Phys. J. A* **56**, 4 (2020).
- [25] M. Kachelrieß, S. Ostapchenko, and J. Tjemsland, *Phys. Rev. C* **108**, 024903 (2023).
- [26] M. Horst, L. Barioglio, F. Bellini, L. Fabbietti, C. Pinto, B. Singh, and S. Tripathy, *Eur. Phys. J. C* **83**, 804 (2023).
- [27] J. Adam *et al.* (ALICE Collaboration), *Phys. Rev. C* **93**, 024917 (2016).
- [28] S. Acharya *et al.* (ALICE Collaboration), *Phys. Rev. C* **97**, 024615 (2018).
- [29] S. Acharya *et al.* (ALICE Collaboration), *Phys. Lett. B* **794**, 50 (2019).
- [30] S. Acharya *et al.* (ALICE Collaboration), *Eur. Phys. J. C* **80**, 889 (2020).
- [31] S. Acharya *et al.* (ALICE Collaboration), *J. High Energy Phys.* 01 (2022) 106.
- [32] S. Acharya *et al.* (ALICE Collaboration), *Eur. Phys. J. C* **82**, 289 (2022).
- [33] S. Schael, R. Barate, R. Brunelière *et al.*, *Phys. Lett. B* **639**, 192 (2006).
- [34] S. Schael *et al.* (ALEPH Collaboration), *Phys. Lett. B* **639**, 192 (2006).
- [35] M. Di Mauro, N. Fornengo, A. Jueid, R. R. de Austri, and F. Bellini, *Phys. Rev. Lett.* **135**, 131002 (2025).
- [36] R. B. Wiringa, V. G. J. Stoks, and R. Schiavilla, *Phys. Rev. C* **51**, 38 (1995).
- [37] ALICE Collaboration, *Phys. Lett. B* **860**, 139191 (2025).
- [38] S. Acharya *et al.* (ALICE Collaboration), *Phys. Rev. C* **107**, 064904 (2023).
- [39] J. Adam *et al.* (STAR Collaboration), *Phys. Rev. C* **99**, 064905 (2019).
- [40] T. Anticic *et al.* (NA49 Collaboration), *Phys. Rev. C* **85**, 044913 (2012).
- [41] ALICE Collaboration, *Phys. Lett. B* **800**, 135043 (2019).
- [42] A. Ibarra and S. Wild, *J. Cosmol. Astropart. Phys.* 02 (2013) 021.
- [43] R. Pohl *et al.* (CREMA Collaboration), *Science* **353**, 669 (2016).
- [44] M. Mahlein, L. Barioglio, F. Bellini, L. Fabbietti, C. Pinto, B. Singh, and S. Tripathy, *Eur. Phys. J. C* **83**, 804 (2023).
- [45] C. Bierlich *et al.*, *SciPost Phys. Codebases* **2022**, 8 (2022).
- [46] H. G. Fischer (NA49 Collaboration), *Nucl. Phys. A* **715**, 118 (2003).
- [47] E. Abbas *et al.* (ALICE Collaboration), *Eur. Phys. J. C* **73**, 2496 (2013).
- [48] P. Skands, S. Carrazza, and J. Rojo, *Eur. Phys. J. C* **74**, 3024 (2014).
- [49] M. Antinucci, A. Bertin, P. Capiluppi, M. D'Agostino-Bruno, A. M. Rossi, G. Vannini, G. Giacomelli, and A. Bussiere, *Lett. Nuovo Cimento* **6**, 121 (1973).
- [50] A. Aduszkiewicz *et al.* (NA61/SHINE Collaborations), *Eur. Phys. J. C* **77**, 671 (2017).
- [51] T. Anticic *et al.* (NA49 Collaboration), *Eur. Phys. J. C* **65**, 9 (2010).
- [52] A. Adare *et al.* (PHENIX Collaboration), *Phys. Rev. C* **83**, 064903 (2011).
- [53] K. Aamodt *et al.* (ALICE Collaboration), *Eur. Phys. J. C* **71**, 1655 (2011).
- [54] J. Adam *et al.* (ALICE Collaboration), *Eur. Phys. J. C* **75**, 226 (2015).
- [55] R. Corke and T. Sjostrand, *J. High Energy Phys.* 03 (2011) 032.
- [56] J. R. Christiansen and P. Z. Skands, *J. High Energy Phys.* 08 (2015) 003.
- [57] S. Acharya *et al.* (ALICE Collaboration), *Eur. Phys. J. C* **81**, 256 (2021).
- [58] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Phys. Rep.* **97**, 31 (1983).
- [59] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *Comput. Phys. Commun.* **191**, 159 (2015).
- [60] S. Ferreres-Solé and T. Sjöstrand, *Eur. Phys. J. C* **78**, 983 (2018).
- [61] C. Bierlich, T. Sjöstrand, and M. Utheim, *Eur. Phys. J. A* **57**, 227 (2021).
- [62] S. Acharya *et al.* (ALICE Collaboration), *Nature (London)* **648**, 306 (2025).
- [63] R. Acconcia, D. D. Chinellato, R. Derradi de Souza, J. Takahashi, G. Torrieri, and C. Markert, *Phys. Rev. D* **97**, 036010 (2018).
- [64] P. Abreu *et al.* (DELPHI Collaboration), *Phys. Lett. B* **361**, 207 (1995).
- [65] J. Abdallah *et al.* (DELPHI Collaboration), *Eur. Phys. J. C* **44**, 299 (2005).
- [66] G. Alexander *et al.* (OPAL Collaboration), *Phys. Lett. B* **358**, 162 (1995).
- [67] G. Alexander *et al.* (OPAL Collaboration), *Z. Phys. C* **73**, 569 (1997).
- [68] S. Navas *et al.* (Particle Data Group), *Phys. Rev. D* **110**, 030001 (2024).
- [69] B. B. Abelev *et al.* (ALICE Collaboration), *Eur. Phys. J. C* **75**, 1 (2015).
- [70] V. Vovchenko and H. Stoecker, *Comput. Phys. Commun.* **244**, 295 (2019).
- [71] S. Acharya *et al.* (ALICE Collaboration), *Phys. Lett. B* **811**, 135849 (2020).
- [72] M. Korsmeier, F. Donato, and M. Di Mauro, *Phys. Rev. D* **97**, 103019 (2018).
- [73] L. Orusa, M. Di Mauro, F. Donato, and M. Korsmeier, *Phys. Rev. D* **105**, 123021 (2022).
- [74] L. Stefanuto, M. Di Mauro, F. Donato, N. Fornengo, J. Koechler, and D. Maurin, *Phys. Rev. D* (2026).
- [75] M. Di Mauro, J. Koechler, L. Stefanuto, F. Bellini, F. Donato, and N. Fornengo, Differential cross-sections for secondary antiproton and antideuteron production (2026), [10.5281/zenodo.19099608](https://doi.org/10.5281/zenodo.19099608).