Astroparticle Physics 182 (2026) 103265

Contents lists available at ScienceDirect

Astroparticle Physics

journal homepage: www.elsevier.com/locate/astropartphys

CORSIKA 8: A general framework for particle cascade simulations

Jean-Marco Alameddine [a,b], Johannes Albrecht [a,b], Antonio Augusto Alves Jr. [c,d], Juan Ammerman-Yebra [e,f], Luisa Arrabito [g], Dominik Baack [a,b], Alan Coleman [h,i], Cosmin Deaconu [j], Hans Dembinski [a,b], Dominik Elsässer [a,b], Ralph Engel [c], Alice Faure [g], Alfredo Ferrari [c], Chloé Gaudu [k], Christian Glaser [h,a,b], Marvin Gottowik [c,*], Dieter Heck [c], Tim Huege [c,l], Karl-Heinz Kampert [k], Nikolaos Karastathis [c], Jeffrey Lazar [m], Lukas Nellen [n], David Parello [o,p], Tanguy Pierog [c], Remy Prechelt [q], Radek Privara [r,s,t], Maximilian Reininghaus [u,i], Wolfgang Rhode [a,b], Felix Riehn [a,v,f], Maximilian Sackel [a,b], Pranav Sampathkumar [c], Alexander Sandrock [k], André Schmidt [c], Jan Soedingrekso [a,b], Ralf Ulrich [c], Philipp Windischhofer [j], Baobiao Yue [k]

[a] Technische Universität Dortmund (TU), Department of Physics, Dortmund, Germany
[b] Lamarr Institute for Machine Learning and Artificial Intelligence, Dortmund, Germany
[c] Karlsruhe Institute of Technology (KIT), Institute for Astroparticle Physics (IAP), Karlsruhe, Germany
[d] University of Cincinnati, Cincinnati, OH, United States
[e] IMAPP, Radboud University Nijmegen, Nijmegen, The Netherlands
[f] Universidade de Santiago de Compostela, Instituto Galego de Física de Altas Enerxías (IGFAE), Santiago de Compostela, Spain
[g] Laboratoire Univers & Particules de Montpellier, CNRS & Université de Montpellier (UMR-5299), 34095 Montpellier, France
[h] Uppsala University, Department of Physics and Astronomy, Uppsala, Sweden
[i] Independent researcher
[j] Department of Physics, Enrico Fermi Institute, Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
[k] Bergische Universität Wuppertal, Department of Physics, Wuppertal, Germany
[l] Vrije Universiteit Brussel, Astrophysical Institute, Brussels, Belgium
[m] UCLouvain, Centre for Cosmology, Particle Physics and Phenomenology, CP3, Chemin du Cyclotron 2, 1348 Louvain-la-Neuve, Belgium
[n] Universidad Nacional Autónoma de México (UNAM), Instituto de Ciencias Nucleares, Mexico, Mexico
[o] DALI, Univ Perpignan, Perpignan, France
[p] LIRMM Univ Montpellier, CNRS, Montpellier, France
[q] University of Hawai'i at Manoa, Department of Physics and Astronomy, Honolulu, USA
[r] European Southern Observatory (ESO), Garching, Germany
[s] Palacký University in Olomouc, Faculty of Science, Joint Laboratory of Optics, Olomouc, Czech Republic
[t] Institute of Physics of the Academy of Sciences of the Czech Republic, Joint Laboratory of Optics, Olomouc, Czech Republic
[u] Karlsruhe Institute of Technology (KIT), Institut für Experimentelle Teilchenphysik (ETP), Karlsruhe, Germany
[v] Laboratório de Instrumentação e Física Experimental de Partículas (LIP), Lisboa, Portugal

ARTICLE INFO

Keywords:
Air shower simulation
Monte Carlo simulations
Cosmic rays
Extensive air showers

ABSTRACT

The simulation of extensive air showers and particle cascades in general is a cornerstone of modern astroparticle physics. For more than two decades, CORSIKA, currently in version 7, has been one of the most widely used tools for this purpose. However, its architecture reflects design constraints of an earlier computing era, as well as increasingly limiting extensibility, maintainability, and adaptability to modern experimental requirements. CORSIKA 8 is a complete redesign of the original CORSIKA code, implemented in modern C++ and based on contemporary software engineering principles. It introduces a modular and extensible simulation framework with explicit handling of units, flexible geometry, and environment descriptions. In this paper, we present the design philosophy and core architecture of CORSIKA 8, describe the implementation of electromagnetic and hadronic shower physics, and validate air shower simulations against CORSIKA 7. The results demonstrate good agreement at the ~5 – 10% level for key observables, confirming the physics fidelity of CORSIKA 8. The observed differences reflect the current level of intrinsic systematic uncertainties

* Corresponding author.
E-mail address: corsika8@kit.edu (M. Gottowik).

https://doi.org/10.1016/j.astropartphys.2026.103265
Received 2 April 2026; Received in revised form 18 May 2026; Accepted 1 June 2026
Available online 9 June 2026

in state-of-the-art air-shower simulations. We also showcase new use cases that were beyond the capabilities of version 7, such as the simulation of cross-media showers and particle cascades in ice, including radio-signal propagation.

1. Introduction

The simulation of extensive air showers (EAS), as well as particle cascades in other media, is a cornerstone of modern astroparticle physics. Ground-based experiments rely on detailed Monte Carlo simulations both when air showers constitute the primary signal, as in cosmic-ray and gamma-ray observatories, and when they form an irreducible background, as in neutrino and multi-messenger experiments. Accurate air-shower simulations are indispensable for detector design, event reconstruction, energy and mass composition measurements, and for the interpretation and comparison of results across experiments.

For more than two decades, several mature Monte Carlo simulation programs have been used to simulate extensive air showers in the atmosphere. Among the most widely used are CORSIKA [1] and AIRES [2]. These codes have been adopted by a broad range of cosmic-ray, gamma-ray, and neutrino detection experiments and have played a central role in the interpretation of air-shower measurements, as illustrated, for example, by the more than 1000 citations of the original CORSIKA publication [1] as counted by inspirehep.net and the collection of "Mentions" and "Testimonials" in Ref. [3].

The current version of CORSIKA, commonly referred to as CORSIKA 7, is still being actively maintained and continues to deliver excellent physics performance. However, its software architecture reflects design decisions rooted in the constraints of FORTRAN 77 and subsequent extensions. Fortran *common blocks*, roughly equivalent to global variables in other programming languages and used ubiquitously in CORSIKA 7, have been declared obsolescent in the Fortran 2018 standard [4]. Even though only minor adaptations were needed during the 30-year history of CORSIKA 7 to ensure its buildability on ever-evolving platforms, the obsolescence of legacy FORTRAN may render the code unusable in the (admittedly distant) future.

At the same time, the requirements placed on air-shower simulations by current and future experiments are evolving rapidly. Modern observatories demand increased flexibility in the choice and combination of physics models, support for heterogeneous detector geometries and media, improved interfaces to analysis frameworks, and efficient exploitation of contemporary computing architectures. While the highly optimized and monolithic structure of CORSIKA 7 delivers excellent performance for the well-established use case of particle cascades in air, it significantly complicates the implementation, validation, and long-term maintenance of new features. Furthermore, the pool of developers with deep expertise in large FORTRAN codebases is steadily shrinking, posing an additional risk to the sustainability of the software ecosystem.

Motivated by these considerations, CORSIKA 8 was developed as a complete redesign of the CORSIKA core, based on modern C++ and contemporary software engineering principles [5]. Rather than a line-by-line translation, CORSIKA 8 provides a modular, extensible, and transparent simulation framework. It preserves the well-tested physics concepts of its predecessor, enables new use cases that are challenging to address with CORSIKA 7, and ensures long-term maintainability. The design emphasizes a modular structure in which each component – such as physics models, geometry definitions, particle transport, and output handling – is implemented independently through well-defined interfaces. This allows components to be developed, tested, or replaced without affecting the rest of the framework, and enables flexible combinations of physics models, geometries, and output formats.

CORSIKA 8 is being developed as a fully open-source project [6] and is intended as a long-term community effort. Contributions in the form of code, physics models, validation studies, documentation, and user feedback are greatly appreciated and actively encouraged. At the time of writing, CORSIKA 8 is considered "physics complete" for the simulation of extensive air showers, i.e., all relevant physical processes and interaction models required for such simulations have been implemented and are available within the framework. Stable, versioned releases are being published on Zenodo [7] in the form of Apptainer containers [8], enabling straightforward out-of-the-box usage on a wide range of computing platforms.

In this article, we present an overview of the motivation, design philosophy, and current status of CORSIKA 8. We describe the architectural concepts and core components of the framework, summarize the implemented physics and technical features, and present validation studies comparing CORSIKA 8 air-shower simulations with CORSIKA 7 (version 7.75). We also illustrate the capabilities of CORSIKA 8 beyond traditional air-shower applications by presenting selected example results. The results presented in the following sections were obtained using the "ICRC 2025 release" version of CORSIKA 8 [9].

2. Architecture

In a very simplified manner, CORSIKA 8 can be described as performing mainly two tasks: propagation of individual particles under the influence of electromagnetic forces and energy losses, and the handling of stochastic events, i.e., decays and interactions, that involve the generation of new particles.

The first task involves solving the respective equations of motion, typically via numerical methods, taking into account constraints that limit the range of propagation such as geometrical and energy cuts. Moreover, the range of propagation is limited by the occurrence of stochastic events. Care must be taken to ensure that these stochastic ranges of propagation are correctly sampled from their respective distributions, which often do not have a closed-form analytic expression (e.g., a particle with an energy-dependent interaction cross-section propagating in a medium with inhomogeneous density while suffering from continuous energy losses).

The physically accurate event generation is itself a huge task on its own. It requires careful modeling of the relevant physics and is therefore outsourced to specialized codes, such as hadronic or electromagnetic interaction models that provide on the one hand cross sections of the respective processes and on the other hand the machinery to sample such an exclusive event with the full kinematic state of the produced secondary particles.

However, these two aspects alone would be pointless to implement without proper *bookkeeping* facilities that gather and finally store the data users are interested in, such as particle distributions on the ground and longitudinal profiles. Fig. 1 shows how these components are arranged and interact within the overall structure of CORSIKA 8. In the following, we will explain the components in CORSIKA 8 in a "bottom-up" manner. Further details are given in Ref. [10].

2.1. Fundamentals

First, we discuss several fundamental cornerstones used in the design of CORSIKA 8.

2.1.1. Unit system

Incorrect usage of physical units of measure is regarded one of the three most common errors in scientific computing [11].

Especially in a collaborative environment, such as CORSIKA 8, a strategy to avoid these kinds of error is a key asset, also because their presence may render the simulation wrong in an

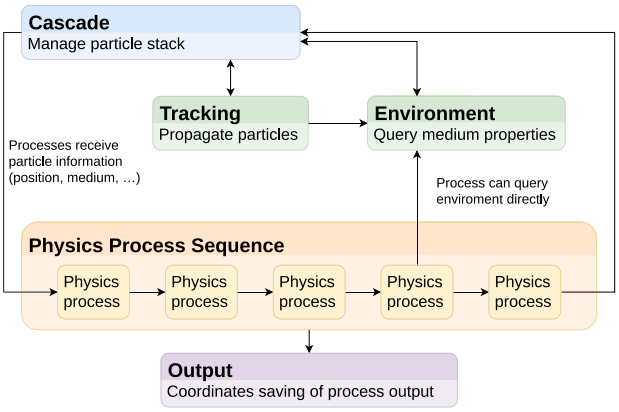

Fig. 1. General structure of the CORSIKA 8 code. The Cascade manages the particle stack and main loop, invoking tracking and a configurable sequence of physics processes. Processes may query environmental properties directly, and output modules collect and store simulation results.

inconspicuous manner. One possible solution is to annotate data types with units. This solution was proposed more than 40 years ago [12]. However, hardly any programming language today provides built-in support, so custom libraries are usually required [13]. In C++ the technique of template metaprogramming can be used to implement a compile-time dimensional analysis [14]. In CORSIKA 8, we make use of the *PhysUnits C++11* library [15], which employs this technique, with some custom modifications.

The basic idea is to introduce individual data types for each dimensionful quantity occurring in the code (length, time, density, etc.) and enforce their usage instead of plain floating-point numbers. The multiplication and division of dimensionful quantities yields quantities with different dimensions, so that an arbitrary number of such data types must be defined. In C++ this can readily be achieved with template metaprogramming:

A templated type `dimensions<N1, N2, N3, N4, N5, N6, N7, N8 >` is introduced to keep track of the dimensions of a quantity. The first seven integers $N_1, \ldots, N_7$ represent the exponents of the basic physical dimensions *length*, *mass*, *time*, *electric current*, *absolute temperature*, *amount of substance*, and *luminous intensity*. The final integer represents the exponent of a custom *HEP energy dimension* whose purpose will be explained below.

The actual type for quantities, `quantity<Dim, Rep=double>`, is likewise templated. The first template argument is a `dimensions` type, the second argument selects the underlying floating point type that is used to store the numerical value of the quantity, by default `double`, in its base unit. For better readability, type aliases like `LengthType`, `TimeType` or `GrammageType` are defined so that one usually does not need to handle the dimension indices explicitly.

The laws of quantity calculus (see, e.g., Ref. [16]) are encoded in C++ code. For instance, only quantities of the same dimension can be added and multiplying a `quantity<dimensions<N1, N2, N3 , ...>>` with a `quantity<dimensions<M1, M2, M3, ...>>` returns a `quantity<dimensions<N1+M1, N2+M2, N3+M3, ...>>`. This way, dimensional analysis of all calculations involving quantities is conducted during compilation and any violation of the laws of quantity calculus results in a compiler error.

Besides the dimensional analysis, also the conversion of units to common base units is performed. A number of predefined constants are provided, as well as *user-defined literals* for convenience to initialize a quantity. For example, writing `height = 1.450_km` (where `height` is a `quantity` for length) or `B = 0.48 * gauss + 5_uT` (where `B` is a `quantity` for magnetic flux density) converts the quantity from the value in the stated unit to the base unit.

The strict dimensional analysis based on the SI becomes inconvenient when switching to the natural units of particle physics, where the convention $c = \hbar = 1$ is employed, which reduces the number of dimensions by two. The unit system as described so far prevents the use of intentionally "sloppy" expressions, such as $E^2 = p^2 + m^2$, in code. This is remedied by the introduction of the spurious eighth HEP energy dimension, which we treat as independent. Quantities of this type (`HEPEnergyType`) are intended for particle masses, energies and momenta, expressed in multiples of electronvolts. Only a few situations in CORSIKA 8 exist where these microscopic units come into contact with the macroscopic SI units. For these cases, conversion functions are provided to convert back and forth between SI and natural units when possible, which mainly relies on the identity $\hbar c = 197.327\,\mathrm{MeV\,fm}$, which relates energy to length. To convert a quantity q whose dimensions are $\mathrm{length}^l \times \mathrm{time}^t \times \mathrm{mass}^m$ from SI to natural units, the conversion reads

$$\frac{q}{\mathrm{eV}^{m-t-l}} = \left(\frac{\hbar c}{\mathrm{eV\,m}}\right)^{m-t-l} \cdot \left(\frac{\hbar}{\mathrm{kg\,m^2\,s^{-1}}}\right)^{-m} \cdot \left(\frac{c}{\mathrm{m\,s^{-1}}}\right)^{m+t} \times \frac{q}{\mathrm{m}^l\,\mathrm{s}^t\,\mathrm{kg}^m}, \quad (1)$$

so that the final dimensions are energy^{m-t-l}. The inverse conversion of a quantity in natural units with dimension energy^e reads

$$\frac{q}{\mathrm{m}^l\,\mathrm{s}^t\,\mathrm{kg}^m} = \left(\frac{\hbar c}{\mathrm{eV\,m}}\right)^{-e} \cdot \left(\frac{\hbar}{\mathrm{kg\,m^2\,s^{-1}}}\right)^{m} \cdot \left(\frac{c}{\mathrm{m\,s^{-1}}}\right)^{t-m} \times \frac{q}{\mathrm{eV}^e}, \quad (2)$$

with the constraint $m - l - t = e$.

2.1.2. Geometry

The second cornerstone of CORSIKA 8 is its geometry classes that deal with points, vectors, and coordinate systems. The design borrows ideas from the Off line software framework of the Pierre Auger Collaboration [17]. Points and vectors are modeled not just as 3-dim. tuples of their coordinates/components, but are always defined with respect to a specific coordinate system (CS), of which there can be multiple. Having several CSs is useful for instance when interfacing event generators, which usually follow the convention to align the momenta of the projectile particle along the z-axis. One unique root CS is predefined and new CSs are defined in terms of a reference CS and a transformation matrix that relates the new CS with its reference. Supported transformations are elements of the special Euclidean group $\mathrm{SE}(3)$, i.e., translations and rotations, represented by 4×4 transformation matrices. These come into play when operations involve two vectors or points defined in different CSs. In this case, the components/coordinates of one of the objects are temporarily transformed into the CS of the other object. Since the CSs form a tree structure, as illustrated in Fig. 2, this is achieved by multiplying the transformation matrices along the shortest path connecting the two CSs. For example, the transformation required to go from CS_3 to CS_1 is $T_{3\to1} = T_{3\to2}T_{2\to0}T^{-1}_{1\to0}$. These transformations happen completely transparently to the developer and the explicit usage of CS is barely necessary. Most often, expressions can be very close to the symbolic notation used in equations as well. Explicit usage of coordinates is required mostly during the initialization phase of a simulation, when the geometry is set up. For the actual linear algebra computations, the geometry system relies on the Eigen3 library [18], which is highly optimized.

We clearly distinguish between points and vectors as they appear in *affine space* (see e.g. Ref. [19]): Points are subject to rotations and translations. Vectors can be thought of as displacements between two points (or multiples thereof) and are not affected by translations. Moreover, vector components can also carry arbitrary dimensions ("units") as described in Section 2.1.1, while point coordinates are necessarily lengths. Allowed operations follow the rules of the affine structure of Euclidean space: Points and length vectors can be added to return another point. Subtracting two points yields a length vector. Scalar and cross-products are defined for vectors, respecting their dimensions.

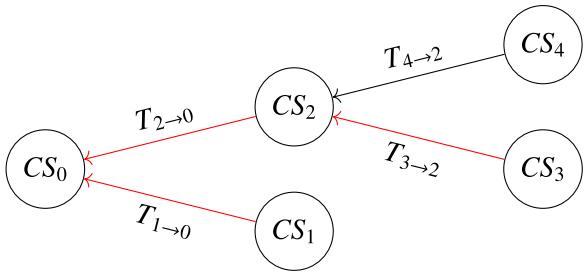

Fig. 2. Example of a coordinate system tree. CS_0 is the root coordinate system, new coordinate systems are always defined with respect to an existing coordinate system. Directed edges represent transformation matrices, T, between coordinate systems. The red path highlights the chain from CS_3 to CS_1 used as an example in the text.

2.1.3. Particle ID and physical properties

To distinguish different particle species, an integer *particle ID* is introduced. It is used not only to label a certain particle, but it also serves as an index to lookup tables containing information such as masses, lifetimes, electric charges, etc. These data are obtained at build time from the `particle` [20] Python package, which provides a curated and up-to-date collection of particle properties derived from the PDG tables. At the same time, the CORSIKA-8-internal particle IDs are generated for each particle by simply incrementing a counter, resulting in numbers currently in the range 1 to 231. The numeric values never need to be used directly, though. Instead, particle IDs are exposed via human-readable `enum class`es, like `Code::MuPlus` or `Code::SigmaMinusBar` for higher expressiveness.

Conversion functions between CORSIKA 8 IDs and the *Monte Carlo Particle Numbering Scheme* of the Particle Data Group (PDG codes) [21] are provided for interoperability with other software. These cannot conveniently be used directly as internal particle IDs themselves because they sparsely cover a large range of numbers, making them unsuitable as lookup indices.

2.1.4. Random-number generation

Being a Monte Carlo code, CORSIKA 8 relies heavily on (pseudo-)random numbers, which are used in large quantities. We have adopted the counter-based random number generator (CBRNG) architecture, and choose the Philox [22] as the default [23] algorithm. The defining property of CBRNG is that their internal state is implemented as incrementable and decrementable counter, which allows advancing the state by an arbitrary number of steps, at constant computational cost, to jump directly to a specific position in the sequence without passing through all intermediate numbers. This makes CBRNGs attractive for use in parallel algorithms that require random numbers, since a CBRNG-based iterable stream of random numbers can be broken into several independent sub-streams by dividing the counter range into many distinct nonoverlapping intervals, each still containing more numbers than the overall simulation needs.

When interfacing FORTRAN-based hadronic interaction models, CORSIKA 8 follows the same idea as CORSIKA 7 (which itself relies on the RNG RANMAR [24,25]) of employing a distinct stream of random numbers for each module, each with its own seed. In particular, individual streams are forwarded to the hadronic interaction models. Technically, random numbers for the FORTRAN-based hadronic interaction models are generated in batches: A buffer is filled with 16 standard random numbers, to be consumed one at a time. Once the buffer is depleted, it is refilled to full capacity at once. This procedure drastically reduces the overhead of function calls from the FORTRAN code into our framework, which would otherwise occur abundantly and negatively impact performance.

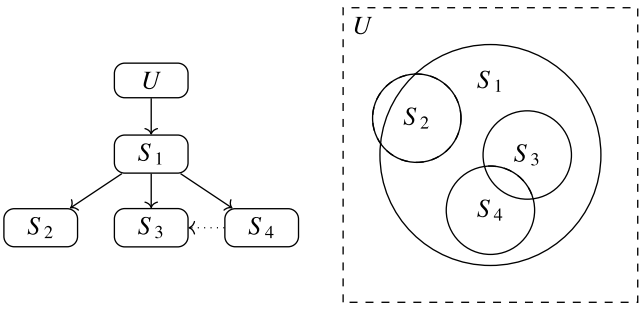

Fig. 3. Example volume tree and corresponding spatial interpretation. The left panel shows a hierarchical structure of volumes, where the root node U (Universe) contains all geometric primitives S_i. Child nodes represent containment relations, while intersections between sibling nodes are resolved using explicit exclusion lists, as indicated by the dotted arrow between S_4 and S_3. The right panel illustrates the corresponding geometry, where S_1 contains S_2, S_3, and S_4, and the overlap between S_3 and S_4 is resolved by excluding S_3 from S_4.

2.2. Environment

CORSIKA 8 is designed to allow the propagation of particle cascades in user-defined environments, providing a great deal of flexibility in three aspects:

- The medium of propagation can be freely selected, considering the limitations set by the physics modules.
- Worlds consisting of different media can be modeled. Example use-cases include air showers that continue their propagation below ground, e.g. in soil, water, or ice, or ν_τ-induced showers emerging from mountains. Transitions between two media are also considered.
- The environment/medium properties can be customized. Depending on the use-case, simulations may, e.g., require querying (electro-)magnetic field strengths or the medium's refractive index as a function of position.

2.2.1. Worldbuilding with the volume tree

To simulate particle cascades in multiple media, a data structure and the corresponding algorithms must be set up. This structure must map spatial regions to media in a flexible way that does not significantly impact performance. In CORSIKA 8 this is achieved with the *volume tree*. The nodes of this tree structure are geometric primitives (volumes), such as spheres, and their placement in the tree represents geometric containment, i.e., a parent node contains its children. This arrangement makes it is possible to determine to which node a certain point belongs. More importantly, the search space for calculating the possible entry and exit points of particle trajectories that cross volume boundaries is small because only a subset of all nodes needs to be considered.

Fig. 3 depicts an example world consisting of a number of spheres, S_i, with its corresponding volume tree. The root node U, called *Universe*, is equivalent to a sphere with infinite radius. It exists to make sure each point in space can be mapped to a well-defined volume node. Complications arise when volumes intersect so that the idea of containment must be reconsidered. We distinguish between two classes of such intersections:

- An intersection of a child volume with its parent ($S_1 - S_2$ in Fig. 3). In this case, the outer volume (parent) clips the inner volume (child).
- An intersection of two child nodes having the same parent requires additional information. For these cases, nodes contain a (by default empty) list of pointers to *excluded* nodes. A point does not belong to a node if it is contained by any of the excluded nodes.

In Fig. 3 this is the case for the overlap between S_3 and S_4. S_4 has S_3 in its exclusion list (indicated by the dotted arrow) so that points in the overlapping region belong to S_3.

For applications requiring observation surfaces of arbitrary shape, such as mountain-based observatories like the Tau Air-Shower Mountain-Based Observatory [26] or muography studies with complex geometries, CORSIKA 8 will soon provide a triangular mesh geometry framework. This framework supports geometries loaded from standard *Wavefront OBJ* [27] or binary *PLY* [28] files, which must be oriented and watertight to be effectively integrated into the volume tree. To maintain performance for large meshes, ray–triangle intersections are accelerated using a bounding volume hierarchy that recursively partitions the triangles into a binary tree of axis-aligned bounding boxes, reducing the average intersection query cost from linear to logarithmic in the number of triangles.

2.2.2. Dressing volumes with models

The individual volumes alone do not constitute any description of the media and their physical properties. Therefore, they have to be furnished with models of these properties. For this purpose, we employ dynamic polymorphism: An abstract class defines the interface and serves as base class. Implementations of this interface occur in classes that inherit from the interface class. A special challenge arises in CORSIKA 8 because the medium interface is not fixed and depends on the physics modules one wants to use for a particular simulation: For example, Cherenkov light and radio emission modules need to query the index of refraction while electromagnetic interaction models require material constants like the radiation length, and for studies of EAS development in thunderclouds, electric fields need to be modeled. To accommodate these requirements, we make use of *mixin inheritance*, which is a template class that inherits from its template argument (see e.g. Ref. [29]). This way, several interface classes, each one responsible for a single aspect, can be chained together to form the complete interface. A simplified example of how different properties are combined to form a single abstract base class is provided in Appendix B.

The actual implementations of the interfaces, in the end required to inherit from their respective interface class, follow the same pattern so that they can be freely combined while orthogonal aspects stay independent. An example implementation corresponding to the interface composition discussed above, together with its usage via dynamic polymorphism, is provided in Appendix B. Similarly, volume nodes are linked to models only via their interfaces.

The most fundamental medium properties needed in CORSIKA 8 are density and the fractions of the isotopes. They always need to be specified even for the most basic cascade simulations.

2.2.3. Propagation medium

The propagation medium in CORSIKA 8 is in general arbitrary, provided that its density profile $\rho(h)$, nuclear composition, and consistent conversions between geometrical length and grammage ($X(\ell, h_0)$ and $\ell(X, h_0)$) are specified. This enables simulations in heterogeneous environments and across medium boundaries, as demonstrated later for air–ice cross-media showers in Section 5. A particularly important and frequently used special case is the atmosphere, which is therefore discussed in detail in the following.

A key component of all air shower simulations are models of the atmosphere, particularly its density, $\rho(h)$, as a function of altitude, h, which often consists of multiple layers. One well-known model is Linsley's parameterization of the U.S. Standard Atmosphere [30], which consists of four isothermal-barotropic layers (i.e., exponential decrease of density with altitude with a fixed scale height) and one topmost layer with a constant density. Such layered parameterizations provide a compact and analytically tractable description of the average atmospheric density profile and have therefore been widely used in air-shower simulations for a variety of applications.

CORSIKA 8 follows this general approach, but implements the atmosphere within a modular and extensible framework. Rather than using Linsley's original parameterization, described e.g. in Ref. [31], we have adopted Keilhauer's parameterization [32] as the default option. This parameterization consists of the same number of layers but the scale heights and reference densities are different. Linsley's original parameterization exhibits discontinuities in density at the layer boundaries, which lead to unphysical artifacts seen in muon production profiles [10]. For convenience, the coefficients of all atmosphere parameterizations available in CORSIKA 7 are also provided. Furthermore, users are free to specify an arbitrary number of exponential and homogeneous layers.

For the simulation, it is technically more important to consider the integrated density along a particle's trajectory, called grammage, $X = \int \rho \mathrm{d}l$, and its inverse rather than to consider the density $\rho(h)$ itself. Since the relationship between altitude and path length in a spherical atmosphere is nonlinear, the integration cannot be performed analytically using closed-form expressions. The mathematical treatment of this integral in an isothermal-barotropic density model leads to the definition of the Chapman function [33], on which a substantial number of approximations have been developed [34,35]. The implementation in CORSIKA 8 follows the *sliding planar atmosphere* approximation, which was originally developed for AIRES. It is also used in CORSIKA 7 when the *curved* option [36] is enabled, where it is applied locally during particle transport with a limitation on the step length to ensure the validity of the approximation. For the precalculation of the slant depth along the shower axis, however, CORSIKA 7 employs a separate numerical integration. In CORSIKA 8, this approximation is used to compute the grammage along particle trajectories, while the particle propagation itself is performed in the full spherical geometry. In contrast to CORSIKA 7, no explicit step-length limitation is currently imposed, such that very long and highly inclined trajectories can lead to a reduced accuracy. The main assumption is that the atmosphere is considered locally flat with an isothermal-barotropic density model having a scale height h_{sc}. In this case, the grammage along a rectilinear trajectory starting at point $\boldsymbol{P}_0$, pointing in the (normalized) direction $\vec{u}$ and having length L can be expressed as

$$X = \varrho(\boldsymbol{P}_0)\frac{h_{\mathrm{sc}}}{\vec{a}\cdot\vec{u}}\left(\exp\left(\vec{a}\cdot\vec{u}\frac{L}{h_{\mathrm{sc}}}\right) - 1\right). \tag{3}$$

Here, $\vec{a}$ denotes a unit vector pointing in direction of the gradient, i.e., vertically downwards, that gets recalculated for each trajectory as pointing from the starting point towards the center of the Earth. In contrast, in a globally flat model, $\vec{a}$ would be a fixed constant.

In contrast to CORSIKA 7, the atmosphere in CORSIKA 8 is treated as a generic propagation medium rather than a hard-coded special case. This allows for the implementation of different atmospheric descriptions in a uniform and consistent manner. This flexibility is part of CORSIKA 8's extensible medium framework, allowing more general density profiles to be implemented when desired. The exponential model remains a commonly used baseline due to its simplicity and analytical tractability.

2.3. Processes

Processes are the entities that act on particles in various ways. They are grouped in six categories:

InteractionProcess This class of processes models interactions and related functionality. Typical examples are wrappers around hadronic interaction models. Two methods must be provided: `CrossSectionType getCrossSection(TParticle const&, Code const, FourMomentum const&)` needs to return the (possibly infinite) interaction length of the modeled physical process for the particle being propagated in its current medium. The actual event generation has to be performed in the `doInteraction(TSecondaryView&, Code const, FourMomentum const&)` method, which typically adds secondary particles via the `SecondaryView` object.

DecayProcess This class of processes models decays. Methods to be provided are: `TimeType getInverseLifeTime(Particle const&)`, which returns the lab-frame decay time of the particle being propagated. Analogous to `InteractionProcess`, the decay event is generated in the `doDecay(SecondaryView&)` method, which typically fills the decay products into the `SecondaryView`.

BoundaryCrossingProcess This class of processes is relevant if any action shall be performed when a particle exits its current volume to enter another. In that case, the method `doBoundaryCrossing(Particle&, VolumeNode &, VolumeNode&)` is called. The last two parameters refer to the current and new volume nodes, respectively.

SecondariesProcess After an interaction or decay event has been performed, the newly generated secondaries can be processed further, e.g. filtered and/or modified. This happens in the `doSecondaries(StackView&)` method, in which the secondaries are typically iterated over.

ContinuousProcess In this class of processes, aspects concerning the continuous movement of a particle along its trajectory are handled. Before the actual propagation takes place, its `LengthType getMaxStepLength(Particle const&, Trajectory const &)` method is called, which returns a maximum (possibly infinite) step length. It serves as a hint to the propagation to limit the step length to that value if necessary. This functionality is provided to ensure numerical accuracy. For instance, a process implementing energy losses may limit the step length to make sure that the energy of the particle does not change too much so that the decay time stays approximately valid. `ProcessReturn doContinuous(Step<TParticle>&, bool const)` is executed after the length of the trajectory has been determined. The particle properties may be modified, e.g. the energy reduced. The boolean input parameter indicates whether the previous step-length limitation had been caused by the `getMaxStepLength()` method of the same process.

The process can indicate whether the particle shall be regarded as absorbed via the return value, which is an `enum` called `ProcessReturn`.

StackProcess Primarily for statistical purposes, it is beneficial to execute code periodically after each N cascade steps. This functionality is provided with this class of processes, which require a `doStack(Stack&)` method. An example use-case is the estimation of the remaining runtime of the simulation by checking how much of the initial energy is still stored in particles on the stack. This quantity decreases linearly with time and reaches zero at the end of the run.

An arbitrary number of processes can be combined to constitute the *process sequence*.

2.4. Particle stack

The particle stack is the central object that manages the particles in memory. The current implementation uses a *structure of arrays*-like memory layout for the particles, meaning that each particle property is stored in a separate, contiguous array. This also means that no independent, compact "particle object" that one could create anywhere exists. Instead, a "particle" is a mere reference to the actual data in the individual arrays. This reference object provides methods like `particle.getEnergy()`, `particle.setPID(Code)`, etc., so that the developer can be oblivious to the underlying memory layout.

The default particle properties stored on the stack are:

- the particle ID (`Code`, integer, also containing nuclear isotope data (A, Z)),
- its kinetic energy (`HEPEnergyType`),
- its position (`Point`),
- its direction vector (dimensionless, normalized `Vector`),
- the time (`TimeType`),
- a pointer to current volume,
- a boolean flag indicating whether a particle is deleted.

To avoid searching the volume tree each time the current volume is needed, a pointer is stored. It is updated each time a boundary is crossed, as described below.

The *is-deleted* flag allows a particle to be marked as "removable" at any position on the stack. A marked particle is only removed when it is read again on top of the stack. This procedure is useful e.g. for thinning and similar filtering purposes in-place, without the need for a (costly) reordering of the particles. It is also a cornerstone of the cascade history described in Section 4.3.

2.5. Program flow

We are now in a position to discuss how the individual building blocks gear into each other to process a full particle cascade. The basic principle is not different from the standard prescription of Monte Carlo cascade codes: particles from the stack are propagated and if they produce any secondaries these are pushed onto the stack. In CORSIKA 8, this main loop is implemented in the `Cascade` class, which has access to the stack, the environment and the process sequence. Furthermore, one may select a certain tracking algorithm, i.e., the procedure by which the trajectory is determined. Propagation in magnetic fields is much more complex than propagation without magnetic fields. A propagation step for a single particle consists of four parts: first, the trajectory is calculated, followed by determining the step length to be taken. Then, the particle is propagated by that length. Finally, the action corresponding to the chosen step length is executed. Let us consider each of these substeps in more detail.

2.5.1. Trajectory determination

Without a Lorentz force, the trajectory is a straight line (ray). The only calculation necessary is determining its maximum length, which is determined by its intersection with boundaries of the volume. In the volume tree, only a subset of the existing volumes needs to be considered: the current volume itself, its child volumes, and its excluded volumes. At the time of writing, the only implemented geometric primitive is a sphere. For the intersection of a straight line parameterized by a real number, ℓ, such that $\ell = 0$ corresponds to the current position and $\ell > 0$ ($\ell < 0$) corresponds to points in front of (behind) the current position, with a sphere, a quadratic equation must be solved. This equation yields up to two real solutions for ℓ. Care must be taken to choose the correct solution: Negative solutions are always excluded because they refer to the past. If both solutions are positive, the smaller (greater) value must be chosen when entering (exiting) a sphere.

A complication arises when a particle has moved to the boundary in the previous step. Due to limited numerical precision, the particle appears to be slightly before or after the boundary. If the intersection is recalculated in the current step and the particle appears to be still before the boundary, the same crossing can be proposed again. This results in a stuck particle and an infinite loop. This can be avoided by keeping a reference to the current node in memory and updating it during a boundary crossing. This ensures that it is always clear whether the considered volume is being entered or exited.

The situation with magnetic fields is more involved. In this case, we employ the leapfrog algorithm as implemented in AIRES and described in Ref. [37]. Updating the position from $\mathbf{P}_0$ to $\mathbf{P}_1$ consists of two half steps, between which the direction $\vec{u}$ is updated as shown in Fig. 4:

$$\mathbf{P}' = \mathbf{P}_0 + \frac{\ell}{2}\vec{u}_0, \tag{4}$$

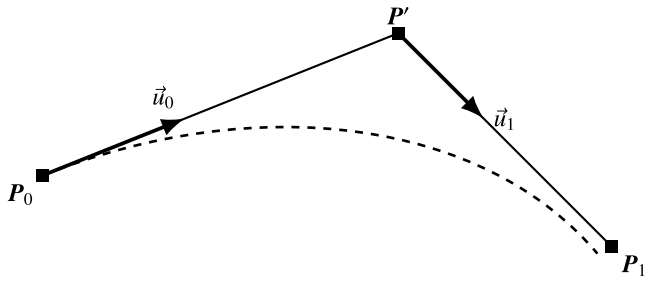

Fig. 4. Schematic of the leapfrog algorithm used for charged particle propagation in a magnetic field. The update from $\mathbf{P}_0$ to $\mathbf{P}_1$ is split into two half steps via an intermediate point $\mathbf{P}'$. The direction $\vec{u}$ is updated between the steps due to the Lorentz force. The dashed curve represents the corresponding true continuous trajectory.

$$\vec{u}_1 = \vec{u}_0 + \vec{u}_0 \times \ell \frac{q}{p} \vec{B}, \tag{5}$$

$$\mathbf{P}_1 = \mathbf{P}' + \frac{\ell}{2} \vec{u}_1. \tag{6}$$

If we consider the magnetic field $\vec{B}$ constant during the propagation (we set $\vec{B} = \vec{B}(\mathbf{P}_0)$) and do not normalize u_1, then $\mathbf{P}_1(\ell)$ is a parabola. The intersection with a sphere is described by a quartic equation that can be solved analytically, albeit with greater computational effort. The correct solution to select is the one with the smallest positive value. Additionally, several safeguards are introduced to handle situations in which the procedure proves to be numerically unstable.

For charged leptons, multiple Coulomb scattering is taken into account following the treatment by Molière [38] as implemented in PROPOSAL, a C++ lepton and photon propagation library originally developed for the propagation of charged leptons [39,40]. It was restructured and expanded to also include photon propagation and serves as a complete electromagnetic interaction model in CORSIKA 8 [41,42]. The effect of multiple scattering is implemented as an update of the direction vector; the additional translation of the position vector is currently neglected. In PROPOSAL, also stochastic deflections in electromagnetic interactions are implemented [43], but this subdominant effect is currently ignored for CORSIKA 8.

2.5.2. Step-length determination

After the trajectory is proposed, the distance that the particle will propagate is determined. To do so, a number of candidate step lengths are considered, of which the minimum is chosen:

- The maximum step length to reach a volume boundary as described in the previous section.
- In case of magnetic fields, the angular deflection per step is limited to a user-configurable value. The default value is 0.2 rad.
- A candidate time of decay is sampled from the current (lab-frame) lifetime and converted to a length using the current velocity. The current lifetime is queried by summing the contributions (branching ratios) of all `DecayProcesses` in the process sequence:

$$\frac{1}{\tau_{\text{tot}}} = \sum_i \frac{1}{\tau^{(i)}}. \tag{7}$$

- A candidate interaction point is sampled: First, a grammage is sampled from the exponential distribution with the current interaction length λ_{int} as parameter, which is determined from the contributions of all (competing) `InteractionProcesses` in the process sequence:

$$\frac{1}{\lambda_{\text{int,tot}}} = \sum_i \frac{1}{\lambda_{\text{int}}^{(i)}}. \tag{8}$$

  To accomplish this, the individual `InteractionProcesses` themselves typically calculate the sum of the cross-sections for each isotope in the medium of the current volume, weighted by its corresponding fractional abundance. Then, the grammage is converted to length using a function provided by the density model. Except for the trivial case of homogeneous density, the density models usually assume a rectilinear trajectory.
- The minimum of the values obtained from the `getMaxStepLength()` methods of the `ContinuousProcesses`.

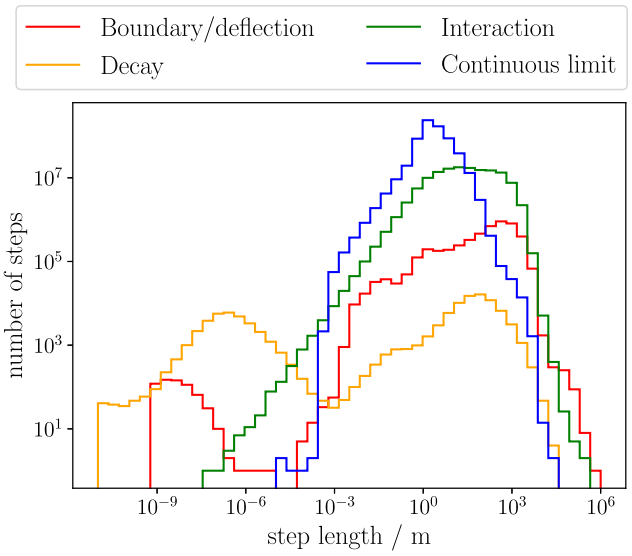

Fig. 5. Distributions of step lengths, i.e. distances between two consecutive positions of a propagated particle. Step lengths are additionally categorized by the process which provides the minimum length – a volume boundary hit or magnetic field deflection, a particle decay, a discrete interaction or reaching a continuous process limit. The distributions are accumulated by logging the propagation distance for all particles across a sample of 100 proton-induced vertical air showers with a primary energy of 10 TeV.

Each step can be classified according to the candidate step length that determined the minimum. Example distributions of step lengths are shown in Fig. 5. These distributions offer useful diagnostic insight into the propagation algorithm.

2.5.3. Propagation

Once the step length has been selected, the initially proposed trajectory is shortened to that length. Then, continuous processes are applied to the particle along the trajectory, which may modify its properties. Purely "observing" continuous processes, which do not modify the particle properties, are also possible. Examples include calculating the radio emission and recording the longitudinal profile.

The step length is additionally constrained by energy-loss considerations such that at most 10% of the particle's energy is lost over a single step or such that the particle energy drops just below the energy cut at the end of the step. In the latter case, it is absorbed and marked for deletion from the stack.

2.5.4. Final action

In case a particle survives the continuous processes, the action corresponding to the selected minimum step length is performed:

- If the step length was set by a length limitation of a continuous process, nothing happens. All steps described in Sections 2.5.1–2.5.3 will be repeated.
- In case of an interaction, the interaction length is recalculated based on the particle properties after the propagation, denoted $\lambda'_{\text{int,tot}}$. This may differ from the initial value $\lambda_{\text{int,tot}}$, which was also used for the sampling. One of the competing

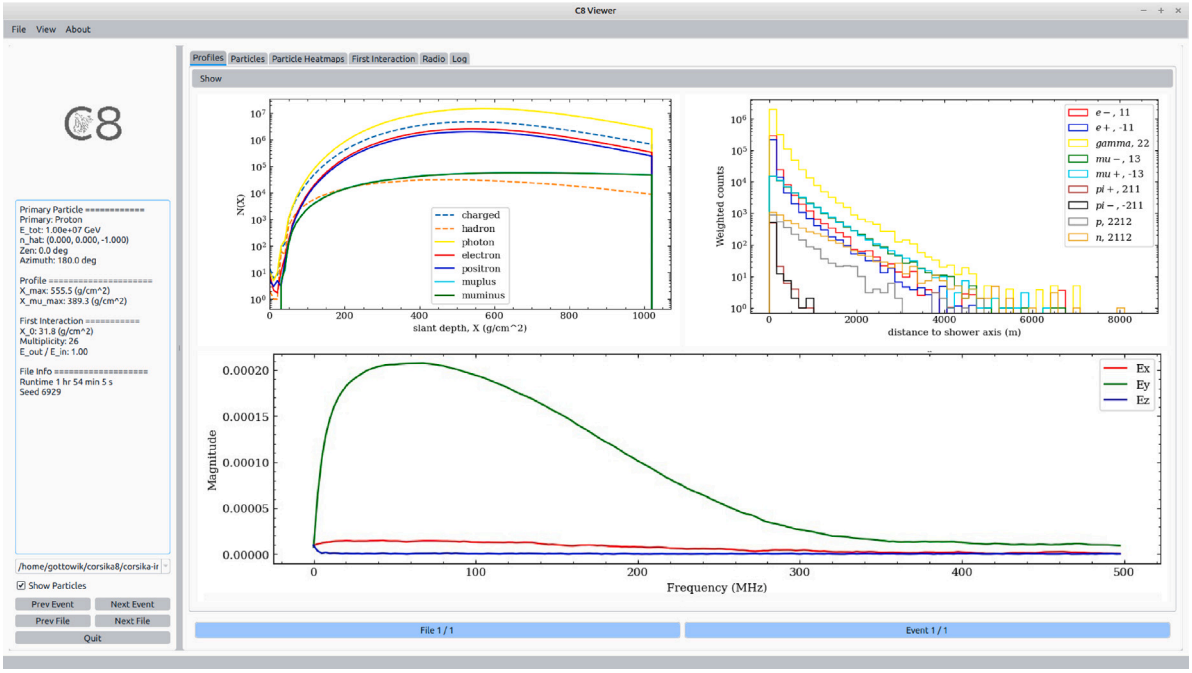

Fig. 6. Collection of visualizations from the *CORSIKA 8 Viewer*. The top-left panel shows the longitudinal particle profile of the shower, while the top-right panel displays the particle distribution at ground level. The bottom panels present the radio frequency spectrum evaluated at a selected observer position.

`InteractionProcesses` needs to be selected to perform the interaction. This is done by randomly selecting the ith process with probability $p_i = \max(\lambda_{\text{int,tot}}, \lambda'_{\text{int,tot}})/\lambda^{(i)}_{\text{int}}$. If $\lambda_{\text{int,tot}} < \lambda'_{\text{int,tot}}$, the probabilities will not sum up to one. The remaining fraction corresponds to a rejection of the interaction.
The sampling in this case is still exact despite the change of $\lambda_{\text{int,tot}}$. If, on the other hand, $\lambda_{\text{int,tot}} > \lambda'_{\text{int,tot}}$, the sampling is in principle not exact, but for sufficiently small changes in energy the incurred error is negligible.

- For decays, the procedure is analogous.
- In case of a volume boundary transition, the `doBoundaryCrossing()` methods of all `BoundaryCrossingProcesses` are called, in which the particle may be marked for deleted. Additionally, the pointer to the current volume is updated to the new volume.

Interactions and decays create new secondaries, which are pushed onto the stack via the `SecondaryView` that offers an `addSecondary()` method, which is called with the properties of the secondary. After the `SecondaryView` has been filled, it is passed subsequently to all `SecondariesProcesses`. These may iterate over the new secondaries and alter or even mark as delete some of them again. Finally, the projectile particle, now no longer on top of the stack, is marked as deleted.

2.6. Output infrastructure

CORSIKA 8 produces a wide range of output data, from compact datasets such as longitudinal profiles to very large particle samples recorded at the observation level. Correspondingly, typical use cases range from single ultra-high-energy cosmic-ray showers to large ensembles of gamma-ray showers with comparatively smaller per-shower output. For reference, the simulations used in Section 4 typically produce output files of about 100 MB per shower.

At the time of writing, the simulation output is written using a combination of the Apache *Parquet* [44] format and accompanying *YAML* files. Parquet is used to store the binary event data in a columnar layout that enables efficient I/O and is widely supported across programming languages. The YAML files provide a human-readable representation of the simulation configuration and a description of the output content. Each output module writes its data independently, and all files are organized in a hierarchical directory structure. A special case in the current implementation are the interaction histograms. They are implemented using the *Boost.Histogram* [45] library, which supports multi-dimensional histograms with flexible axis definitions. Since Boost.Histogram does not define a native persistent storage format, the histograms are currently saved using a lightweight *NumPy*-based file format [46].

We provide a Python library for data access and analysis, that allows for the convenient reading and processing of the generated output [47], thereby making integration into user codes and analysis workflows transparent and independent of the underlying CORSIKA 8 output file format. We strongly recommend that users build on the provided read-in library because the underlying file format might change in the future if needed. However, should lower-level integration be strictly necessary, Parquet readers are readily available for many programming languages.

Furthermore, the *CORSIKA 8 Viewer* [48], shown in Fig. 6, is provided as a graphical user interface (GUI) for interactive inspection and visualization of simulated CORSIKA 8 showers. The viewer allows users to explore the particle content, longitudinal profiles, and radio emission at ground level directly from the simulation output. This offers an intuitive way to validate simulations and inspect individual events.

3. Electromagnetic showers

Having discussed the cornerstones of the CORSIKA 8 design, we will now showcase the simulation of electromagnetic particle showers in air.

3.1. Implementation and results

Earlier versions of CORSIKA simulated the electromagnetic component of air showers using a customized version of EGS4 [49], which was deeply integrated into the source code of CORSIKA. Alternatively, the electromagnetic shower component could be described by the analytic NKG formula [50,51]. In CORSIKA 8, the electromagnetic component

is currently being simulated by the lepton propagator PROPOSAL [39–41], which is also used to describe muon and tau lepton interactions. To ensure comparability with the legacy version of CORSIKA, the cross sections from EGS4 were implemented in PROPOSAL. The standard parameterization of the bremsstrahlung cross section is taken from Ref. [52]; above 50 MeV, the extreme relativistic cross section with Coulomb correction is used, while at lower energies an empirical correction factor is added, which thus rescales the total cross section. The cross section for electron-positron pair production by a photon uses the same screening functions as the bremsstrahlung cross section with a low-energy empirical correction factor taken from Ref. [53]. The photoelectric effect parameterization was taken from Refs. [54,55] for the Kα shell, along with an empirical correction factor from Ref. [56]. The cross section for photohadronic interactions is identical to its description in CORSIKA 7 [57], while interfaces to SOPHIA [58] and SIBYLL 2.3d [59] were implemented for the actual event generation. The implemented cross sections differ at most by a few percent, and only at the lowest or highest energies or in regions where the contribution of the process is very small, cf. Fig. 7 for exemplary comparisons of positron and photon cross sections between CORSIKA 7 and CORSIKA 8. Lastly, PROPOSAL simulates electron-positron pair production by an ingoing electron or positron – a process which was not taken into account by CORSIKA 7.

Unlike the highly integrated usage of EGS4 in legacy CORSIKA versions, PROPOSAL within CORSIKA 8 is integrated analogously to the hadronic interaction models. This allows for the addition of other electromagnetic interaction models alongside PROPOSAL in the future. Furthermore, the modular structure of PROPOSAL allows for disabling, interchanging, or adapting of individual interaction parameterizations.

Comparisons of longitudinal and lateral profiles, as well as the charge excess, defined as $(N_{e^-} - N_{e^+})/(N_{e^-} + N_{e^+})$, of electromagnetic showers at a primary energy of 100 TeV with CORSIKA 7 are shown in Figs. 8 to 11. Further information on the exact configuration is provided in Appendix A. In all profile plots, the solid lines show the median value and the shaded bands indicate the interquartile range (25th to 75th percentile), representing shower-to-shower fluctuations. For this comparison, a total of 2500 photon-induced showers were simulated with CORSIKA 8 and CORSIKA 7. The agreement between CORSIKA 7 and CORSIKA 8 is on the few-percent level, except for very low particle numbers at the earliest stage of the shower development and very small distances to the shower axis.

The Landau–Pomeranchuk–Migdal (LPM) effect [60,61] changes the behavior of electromagnetic showers at extremely high energies and at large atmosphere densities. Under these circumstances, the approximation of an interaction with a single isolated atom breaks down and the coherence of the initial and final wave functions is lost due to scattering on other atoms. This suppresses the emission of bremsstrahlung photons, particularly those with a low energy compared to the initial lepton energy, as well the production of electron-positron pairs, particularly those with a symmetric energy distribution between the produced leptons. This effect slows down the development of the cascade and shifts the shower maximum to deeper depths. For earlier work on electromagnetic cascades in water and lead and the formulas used to calculate the LPM effect, we refer the reader to Ref. [62]. The LPM effect is taken into account in CORSIKA 8 using a rejection sampling technique, where each interaction is rejected with a probability based on the difference between the interaction cross section with and without LPM suppression. This approach is similar to the method used in CORSIKA 7 [63], but it can now be used in air as well as in dense media. The effect of the LPM suppression on 100 EeV electromagnetic showers in air is shown in Fig. 12, based on a total of 5000 photon-induced showers simulated with and without the LPM effect in both CORSIKA 7 and CORSIKA 8.

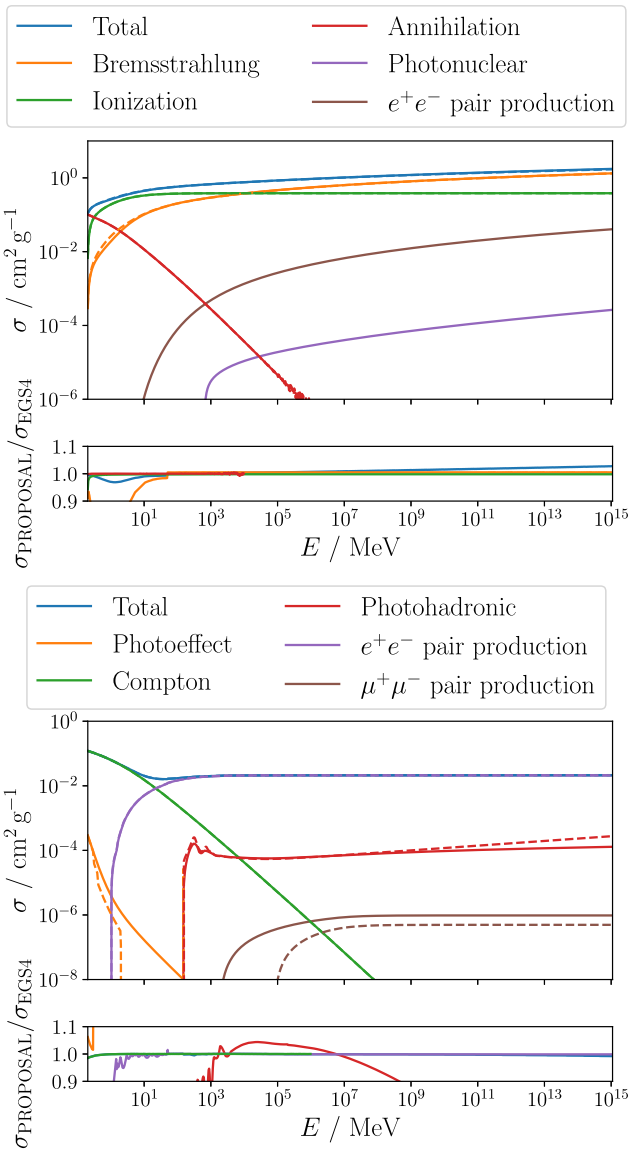

Fig. 7. Total stochastic cross-sections for positrons (top) and photons (bottom) in air. The losses smaller than 0.2 MeV are treated continuously to obtain a finite cross-section. The dashed lines refer to cross-sections from CORSIKA 7, the solid lines show the implementation in CORSIKA 8.

3.2. Thinning

Thinning, introduced by Hillas [64], is a technique that reduces the runtime of shower simulations, especially at the highest energies, to a manageable timeframe. It is an example of a more general class of methods called *Russian roulette* that has been used in other fields where radiation transport is relevant [65]. The general idea is to reduce the number of individually tracked particles in the shower by *pruning* the shower randomly, i.e., by discarding a large fraction of the secondaries at each vertex. The discarded particles are accounted for by assigning statistical weights to the retained particles. This ensures that the expectation values of observables calculated from the pruned shower match those of the full shower. However, the cost of applying thinning is the introduction of *artificial* statistical fluctuations in addition to the physical shower-to-shower fluctuations. It has been shown that these artificial fluctuations can be reduced by introducing a weight limitation [66], which prevents particles from obtaining weights larger than a prescribed value.

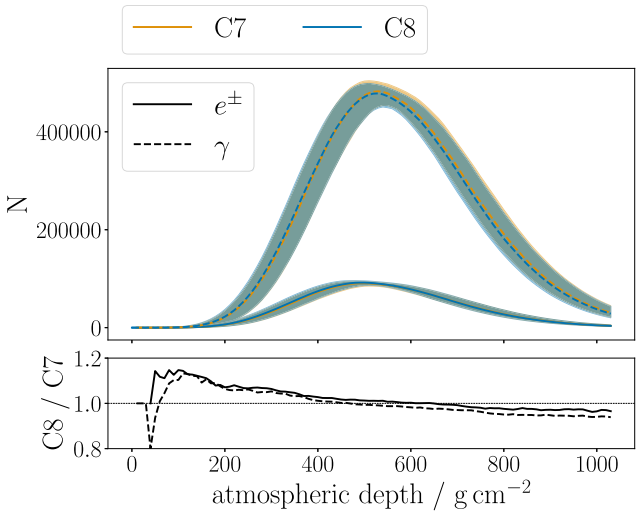

Fig. 8. Median longitudinal profile of electromagnetic particles with energies above 1 MeV in 100 TeV photon-induced showers in CORSIKA 7 and CORSIKA 8. The shaded bands indicate the interquartile range.

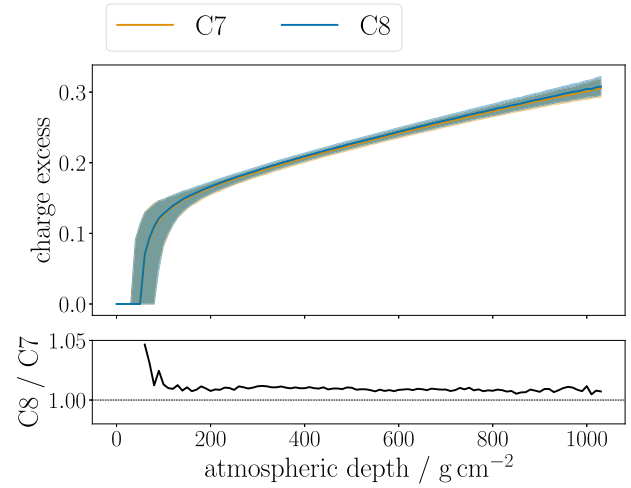

Fig. 9. Median charge excess of 100 TeV photon-induced showers in CORSIKA 7 and CORSIKA 8.

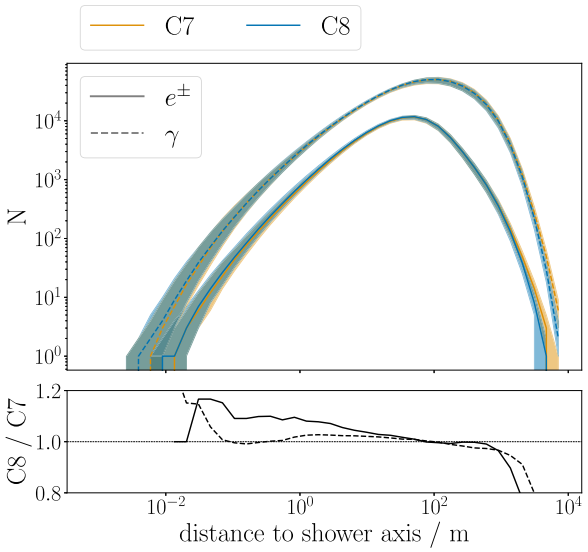

Fig. 10. Median lateral profile of electromagnetic particles in 100 TeV photon-induced showers at a height of 5.8 km, typical depth of shower maximum, in CORSIKA 7 and CORSIKA 8.

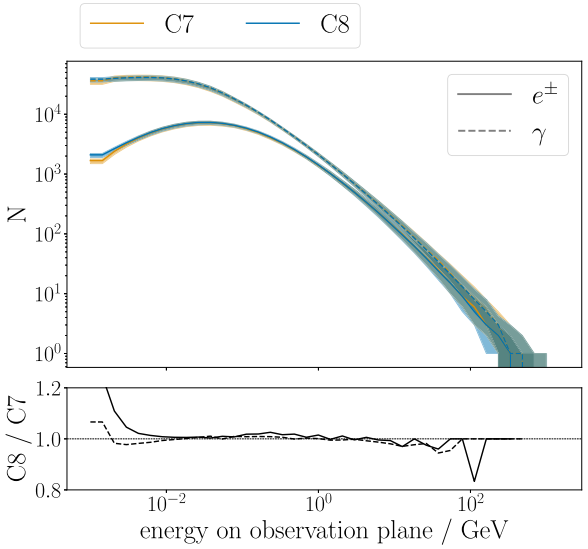

Fig. 11. Median energy distribution of electromagnetic particles in 100 TeV photon-induced showers at a height of 5.8 km, typical depth of shower maximum, in CORSIKA 7 and CORSIKA 8.

The implementation in CORSIKA 8 is as follows[1]: two parameters determine the phase space of particles subject to thinning, the *threshold energy*, E_{th} (usually expressed as a fraction $\epsilon = E_{\mathrm{th}}/E_{\mathrm{CR}}$ of the primary energy), and the *maximum weight*, w_{max}. If the energy of the incoming particle, E_0, is above the threshold, $E_0 > E_{\mathrm{th}}$, no thinning is applied, i.e., the secondaries are propagated further as usual. The early stage of the shower development, consisting of only a small fraction of the particles, is the dominant source of shower-to-shower fluctuations. This stage largely determines the "fate" of the later development. In this regime, application of thinning would be detrimental. Secondaries with $E_0 \leq E_{\mathrm{th}}$ are subject to the thinning procedure. When weight limitation have not yet set in (to be defined later), one of the two secondaries is randomly chosen to be retained while the other is discarded. The selection probability of the ith secondary is defined to be proportional to its energy, where E_1 and E_2 denote the energies of the two secondary particles produced in the interaction,

$$p_i = \frac{E_i}{E_1 + E_2}. \tag{9}$$

The weight of the retained particle is increased by a factor $f_i = 1/p_i$, i.e.

$$w_i = w_0 f_i = \frac{w_0}{p_i}, \tag{10}$$

where w_0 denotes the weight of the incoming particle (= 1 for particles that have not yet been subject to thinning yet).

Weight limitation is considered as follows: if, at any point, the potential new weight of any secondary, w_i, exceeds the maximum weight, w_{max}, we resort to *statistical thinning* [66], in which each secondary is considered for retention or discarding independently. In this setting, we have more freedom to alter the retention probabilities as desired, since the sum of the retention probabilities does not need to equal one. Here,

[1] Currently, thinning is only applied to electromagnetic interactions with a $1 \rightarrow 2$ splitting.

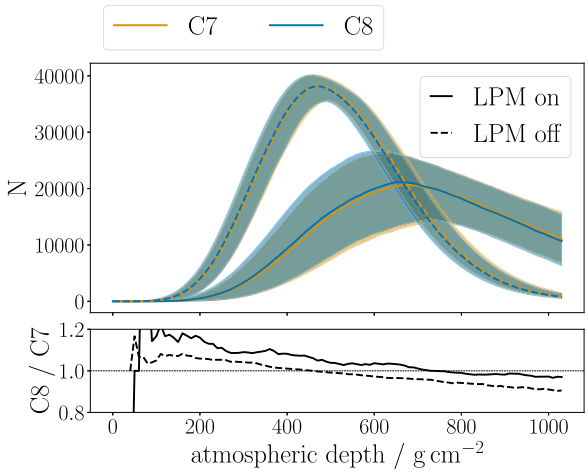

Fig. 12. Median longitudinal profile of charged particles above 100 TeV in 100 EeV photon-induced showers with and without the LPM effect in CORSIKA 7 and CORSIKA 8.

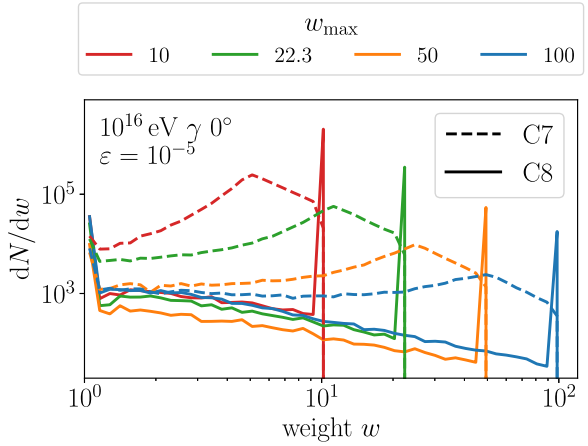

Fig. 13. Weight distributions of photons due to thinning for different settings of the allowed maximum weight in comparison for CORSIKA 8 and CORSIKA 7.

we set

$$p_i = \max\left(\frac{E_i}{E_1 + E_2}, \frac{w_0}{w_{\max}}\right), \tag{11}$$

so that

$$w_i = \frac{w_0}{p_i} = \min\left(w_0 \frac{E_1 + E_2}{E_i}, w_{\max}\right). \tag{12}$$

As soon as the maximum weight is reached in a particular branch of the shower, all particles descending from that vertex are tracked again with the same weight, $w_{\max}$.

Fig. 13 shows a comparison of the weight distributions of photons obtained in 10^{16} eV photon showers using CORSIKA 8 and CORSIKA 7 with a thinning level $\varepsilon = 10^{-5}$ and several maximum weight factors. In the high-weight range, CORSIKA 8 features a narrow peak, while CORSIKA 7 features a broad peak at a value of $w_{\max}/2$. This difference is due to the different implementations of weight limitations. The current CORSIKA 8 implementation switches to statistical thinning, which provides the freedom to adjust the retention probability to reach $w_{\max}$ exactly. CORSIKA 7, on the other hand, has no such freedom. The weight limitation effectively kicks in earlier. When p_i of Eq. (9) is too small to keep $w_i \leq w_{\max}$, no thinning is applied anymore, so that particles are "stuck" around $w \simeq w_{\max}/2$.

When comparing thinned and unthinned showers, it is important to note that enabling thinning alters the shower development. Even when using the same random number seed, a thinned shower will not be a pruned version of the shower obtained by disabling thinning. Since some of the particles and their interactions are skipped, the random numbers that would have been used there are used elsewhere, resulting in a different outcome.

To directly compare a full shower with one or more thinned realizations of the same shower, we offer the option to enable thinning *without discarding* particles (inspired by the multithinning method of CORSIKA 7 [67]). With this option, particles are assigned a weight of zero instead of being discarded, and are retained and further propagated. Thus, the random-number sequence remains undisturbed (the random numbers used for thinning are drawn from an independent sequence). By taking into account all particles with their weight $w \geq 1$, one can calculate observables of the thinned shower, as weight-zero particles do not contribute. Disregarding their weights and treating all particles as having a weight of one allows calculating observables of the full shower. Furthermore, different thinned realizations of the same shower can be obtained by running the simulation again with the same primary seed but a different one for the random number sequence used for thinning. A sufficiently large sample of these realizations is useful for studying the artificial fluctuations induced by thinning. For example, one could use it to test biases or sophisticated reconstruction algorithms that attempt to recover the real particle distribution given a thinned shower, as described in Refs. [68,69].

4. Hadronic showers

In a next step, we discuss and benchmark the simulation of hadronic showers in air with CORSIKA 8.

4.1. Implementation

In contrast to electromagnetic interactions, there is no closed theory for hadronic interactions. In other words, there is no unambiguous way to calculate predictions about particle production from first principles (QCD). In CORSIKA, phenomenological models are used to simulate hadronic showers. To understand the uncertainty resulting from imperfect modeling of hadronic interactions, it is necessary to support a variety of models. Currently, CORSIKA 8 provides interfaces to several widely used hadronic interaction models in cosmic-ray and air-shower physics, including EPOS-LHC [70], QGSJet-II.04 [71], SIBYLL 2.3d [59], and, more recently, QGSJet-III [72], EPOS-LHC-R [73] (not yet available in the "ICRC 2025 release"), and Pythia 8/Angantyr [74,75].[2] Additional hadronic interaction models may be included in future releases. Unlike models used in high-energy physics (HEP), which predominantly focus on high-p_T processes (e.g. Herwig++ [76]), the models used for air shower calculations must be fully inclusive. To facilitate interaction with the HEP community, the interface between CORSIKA 8 and the hadronic models is designed to be generic. The two main interface routines `getCrossSection` and `doInteraction` only require the identification codes and four-momenta of the projectile and target particles in an arbitrary frame of reference. The transformation to the frame of reference of the respective model is done inside the interface in such a way that the overhead due to the transformations is minimal in the air shower context. Consequently, the same interface used in CORSIKA 8 can be used to provide predictions for accelerator experiments, facilitating the testing of the models against experimental data (Rivet [77], MCplots [78]) without the need for additional code packages like CRMC [79].

The models listed above are high-energy (HE) hadronic interaction models specifically designed to describe interactions of hadrons in the multi-particle production regime ($\sqrt{s} \gtrsim 10$ GeV). At lower energies, particle production is dominated by discrete hadron resonances. Since the

[2] v8.315

targets in air showers are nuclei, additional nuclear effects (e.g. Fermi motion of nucleons, intranuclear cascade, and fragmentation) become important when interaction energies reach the GeV scale and below. Not all HE models include these effects. Therefore, in CORSIKA high- and low-energy (LE) interactions are distinguished. Currently, CORSIKA 8 provides an interface for FLUKA [80,81] to handle LE hadronic interactions. The energy at which to switch between the HE and LE models varies among the different HE models and can be configured by the user. The default transition energy is 79.4 GeV for all supported HE models, except for Pythia 8/Angantyr for which a higher threshold of 100 GeV is used for technical reasons.

4.2. Hadronic decays

After their production, hadrons either interact again or decay. For the modeling of decays Pythia 8 is used in CORSIKA 8 and the lifetimes of hadrons are defined according to the Review of particle physics 2024 [20,82]. The probability whether a decay or an interaction occurs is determined by the interaction and decay lengths. In CORSIKA 7, which is strictly designed for cascades in air, the list of hadrons that interact before they can decay is short. With the exception of neutral pions, any hadron more short-lived than K^0_{short} is not tracked in CORSIKA 7 but forced to decay inside the generator. This may have an effect at energies beyond 100 EeV. In CORSIKA 8, which is designed to simulate cascades in any medium, in principle any hadron could interact and so all hadrons should be tracked and no generator-internal decays should be allowed. However, for the moment only interaction cross sections for long-lived hadrons are implemented and short-lived resonances are set to decay immediately after their production. This configuration was chosen because parameterizations of the interaction cross sections for short-lived hadrons are not easily available for all the interaction models. Since this configuration is very similar to CORSIKA 7, this choice also makes the comparison of shower observables much easier. A notable exception is QGSJet-II.04. For Λ hyperons produced in showers simulated with this model, no interaction is implemented in CORSIKA 7, while in CORSIKA 8 the Λs are replaced with neutrons. We have verified that the different treatments introduce an effect of less than $3\,\mathrm{g\,cm^{-2}}$ on the average X_{max}.

4.3. Particle history

The uncertainty in modeling hadronic interactions, as mentioned above, has real consequences for cosmic-ray air shower experiments. It is an established fact that air shower simulations do not consistently describe the data of cosmic ray observatories [83,84]. The prevailing view is that this discrepancy stems from the incomplete modeling of HE hadronic interactions [85]. However, it is not yet clear which aspect of hadronic interactions is misrepresented. In fact, the exact mapping of hadronic interaction properties to air shower observables is not well understood. This is due to the immense increase in the number of degrees of freedom during the development of an air shower. In CORSIKA 7, all of these degrees of freedom are represented in the simulation, but they cannot be read out without significantly restructuring the code. CORSIKA 8 was designed specifically to lift this limitation. For the first time, it is now possible to inspect the entire history of every particle that reaches the ground [86,87]. This allows for much more detailed studies of the air shower development where the causal connection of each particle to the primary particle can be traced.

4.4. Validation

To validate the hadronic interaction framework implemented in CORSIKA 8, we compare key air-shower observables obtained with CORSIKA 8 to those obtained with the well-established CORSIKA 7 code. Since both frameworks provide interfaces to the same set of high- and low-energy hadronic interaction models, such a comparison allows us to isolate effects arising from the new simulation architecture and interfaces rather than from differences in the underlying physics models. Therefore, the focus of this validation is not to test agreement with experimental data; rather, it constitutes a consistency study between two independent implementations that nominally rely on the same underlying physics models.

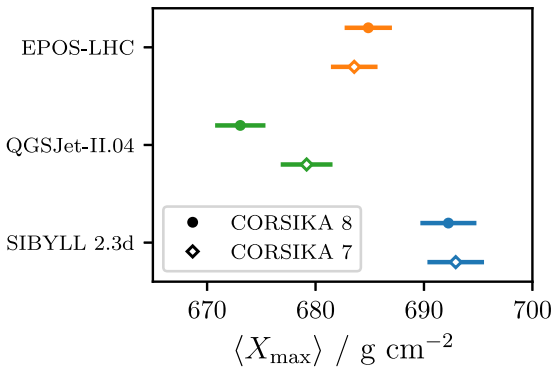

Fig. 14. Average X_{max} for proton-induced showers at 10^{17} eV simulated with CORSIKA 7 and CORSIKA 8 for different high-energy hadronic interaction models. Error bars indicate the statistical uncertainty of the mean. Note that while hadronic models are the same between CORSIKA 7 and CORSIKA 8, they are interfaced to different transport and electromagnetic interaction codes which may affect the overall results. See text for details.

For this purpose, we study vertical proton-induced air showers at a primary energy of 10^{17} eV. We analyze the resulting showers, including the electromagnetic cascade generated by hadronic interactions, in terms of longitudinal and lateral observables for different particle components. High-energy hadronic interactions are modeled using EPOS-LHC, QGSJet-II.04, and SIBYLL-2.3d, while FLUKA is used as the LE interaction model. Identical thinning levels, energy cuts, and atmospheric conditions are applied in both simulation frameworks. We use a thinning level of 10^{-6}, with energy cuts of 10 MeV for electromagnetic particles and 300 MeV for hadrons and muons. Note that in CORSIKA 8, thinning is applied only to the electromagnetic component, whereas in CORSIKA 7 it is applied to the electromagnetic and hadronic components. Further information on the exact configuration is provided in Appendix A. Statistical uncertainties are estimated using bootstrap resampling of the simulated shower sets.

First, we focus on the depth of the shower maximum, X_{max}. We compare the average X_{max} obtained from 1000 CORSIKA 8 simulations with the result of 1000 CORSIKA 7 simulations. The longitudinal shower profile, i.e., the energy deposited as a function of atmospheric depth, is used to determine X_{max} by fitting a parabola to the peak using the five neighboring bins on either side of the maximum bin. The results are shown in Fig. 14. Note that while hadronic models are the same between CORSIKA 7 and CORSIKA 8, they are interfaced and embedded among different transport and electromagnetic interaction codes (see previous sections). For EPOS-LHC and SIBYLL-2.3d, the results from both frameworks are nevertheless consistent within statistical uncertainties. A small but significant deviation of $\approx 6\,\mathrm{g\,cm^{-2}}$ is observed for QGSJet-II.04, the origin of which is still under investigation.

We now compare the longitudinal profiles of various particle types p, considering electrons and positrons, photons, muons, and hadrons separately. For each shower, the longitudinal profile is expressed relative to its maximum, i.e., as a function of $X - X^p_{\mathrm{max}}$. This allows for a direct comparison of the shower shapes, independent of a potential systematic shift between CORSIKA 7 and CORSIKA 8. The median longitudinal profiles of the different particle types, expressed in depth with respect to X^p_{max}, are shown in Fig. 15. Throughout most of the shower development, the ratio C8/C7 remains relatively flat, deviating by less than ∼5 %. Larger deviations primarily appear in the very early

Table 1
Mean $X_{\max}{}^{p}$ [g cm^{-2}] for different particle types and hadronic models. The quoted uncertainties are the uncertainties of the mean.

Model	EAS Code	$e^{\pm}$	$\mu^{\pm}$	Hadrons	γ
SIBYLL-2.3d	C7	692 ± 2	869 ± 2	639 ± 2	725 ± 2
	C8	692 ± 2	872 ± 2	644 ± 2	725 ± 2
QGSJet-II.04	C7	677 ± 2	853 ± 2	622 ± 2	710 ± 2
	C8	670 ± 2	845 ± 2	617 ± 2	703 ± 2
EPOS-LHC	C7	681 ± 2	853 ± 2	641 ± 2	714 ± 2
	C8	682 ± 2	857 ± 2	647 ± 2	714 ± 2

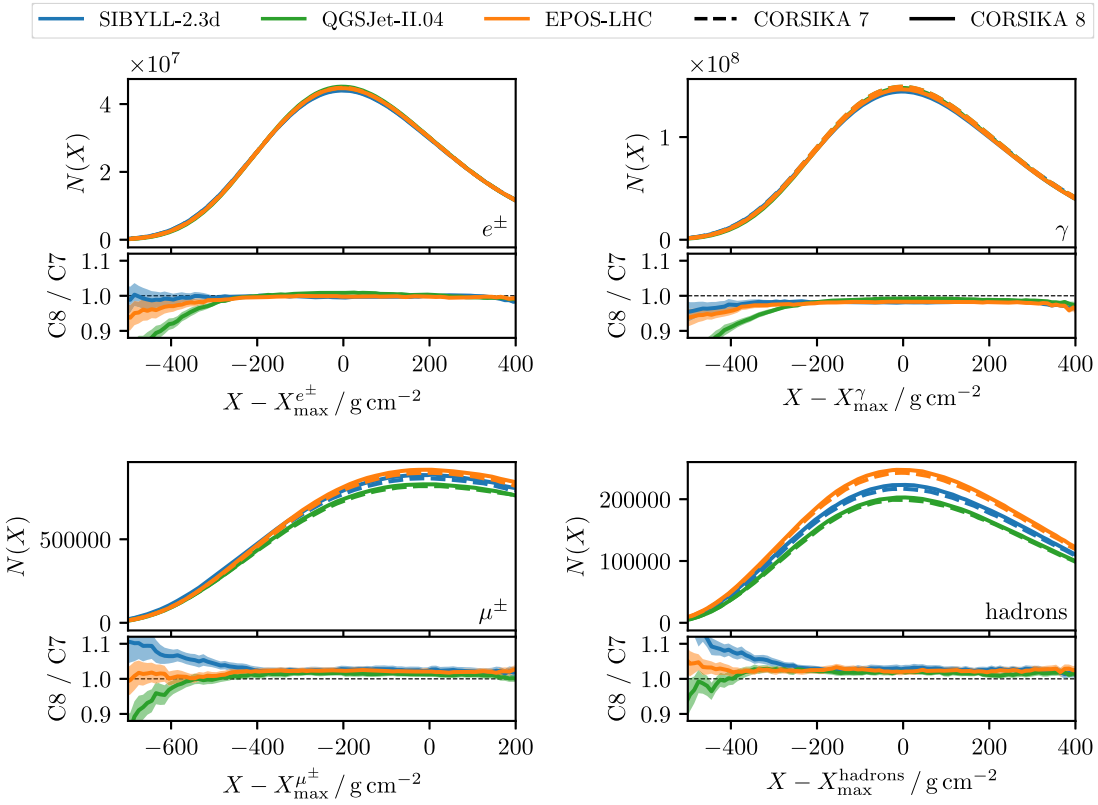

Fig. 15. Median longitudinal profiles of 1000 proton-induced vertical air showers with primary energy of 10^{17} eV for various particle types (see indicator in plots) arising from hadronic cascades simulated with various HE hadronic interaction models and FLUKA as LE interaction model, both with CORSIKA 7 and CORSIKA 8. The shaded area shows the uncertainty of the median estimated via bootstrapping.

stages of the shower and near the end of the profile close to the ground level. To the best of our knowledge, the very early stage of the shower development is not directly constrained by experimental observables and therefore has limited practical relevance for most analyses. We nevertheless report these differences explicitly and refrain from attributing them to either implementation, as it is not evident which code is more accurate. These results demonstrate that the overall longitudinal evolution of the showers is very similar between CORSIKA 7 and CORSIKA 8. We also inspected the distribution of $X^{p}_{\max}$. A comparison of the values of the mean between the two codes is summarized in Table 1. The width of the $X^{p}_{\max}$ distributions is identical within the statistical uncertainties.

Next, we examine shower observables at ground level, focusing on the lateral distribution of particles as a function of distance from the shower axis and their energy spectra. Fig. 16 shows the median lateral distributions and energy spectra for electrons and positrons, and muons from the same set of showers. Shaded areas indicate the statistical uncertainty of the median, estimated via bootstrap resampling. In general, the lateral distributions agree at the ∼10 % level for most particle species. The observed differences appear to depend on the HE interaction model. Although uncertainties at the 10% level are non-negligible and warrant further investigation, we consider reaching agreement at this level between completely independent codes already a notable accomplishment.

5. Cross-media showers

The development of high-energy neutrino experiments in ice and water has created a need for a Monte Carlo code capable of simulating cross-media showers. One of the largest sources of background for these experiments are secondary high-energy muons generated by high-energy cosmic-ray-induced air showers [88,89]. When these muons lose a significant amount of energy, they tend to mimic the signal of an upward-going neutrino. Thus, simulations of showers from one medium to another are necessary for background estimation. Additionally, radio emission from cross-media showers have been observed in radio-based neutrino detectors in ice [90], constituting an important background and valuable calibration source, which requires detailed simulation.

Until now, most simulation programs, such as AIRES and CORSIKA 7, could only run particle showers in one medium. Geant4 [91], on the other hand, could run them in multiple media, but each of them only with constant density. Since none of the codes can handle both air

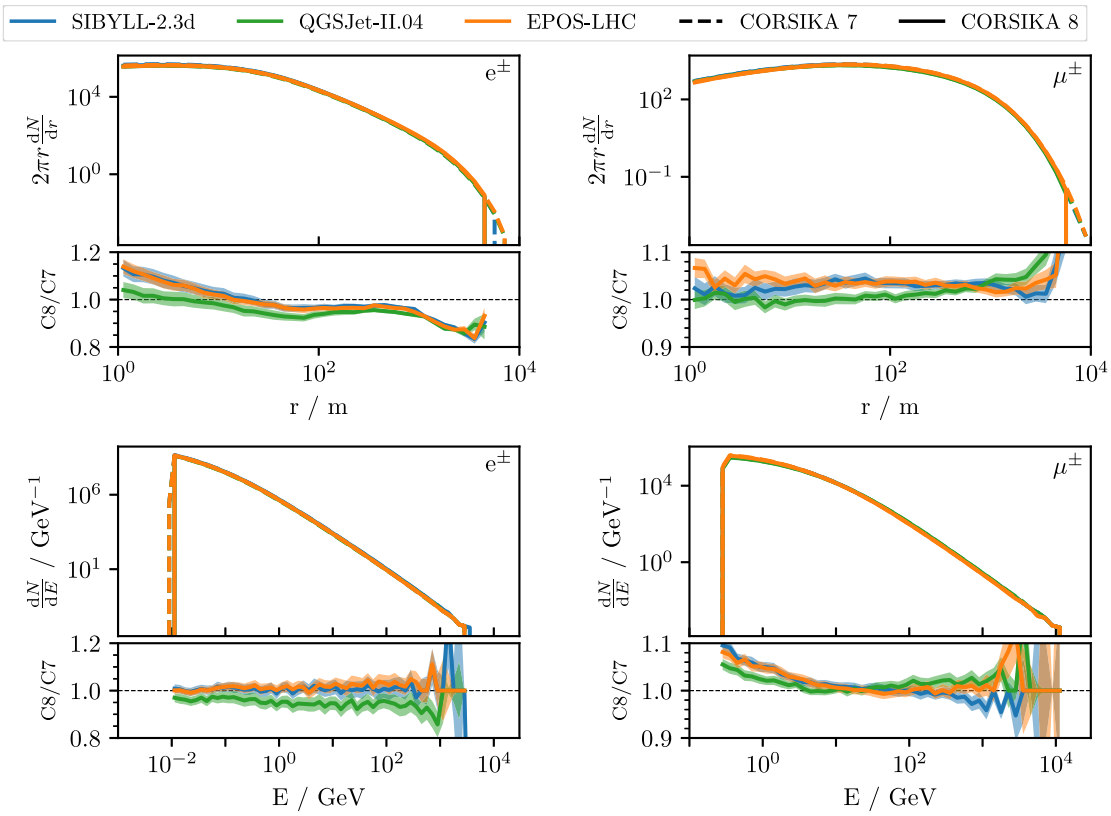

Fig. 16. Top: Median lateral distributions of electrons and positrons (left) and muons (right) at ground level for the same showers as shown in Fig. 15. Bottom: Median energy spectra of electrons and positrons (left) and muons (right). The shaded areas show the uncertainty of the median estimated via bootstrapping.

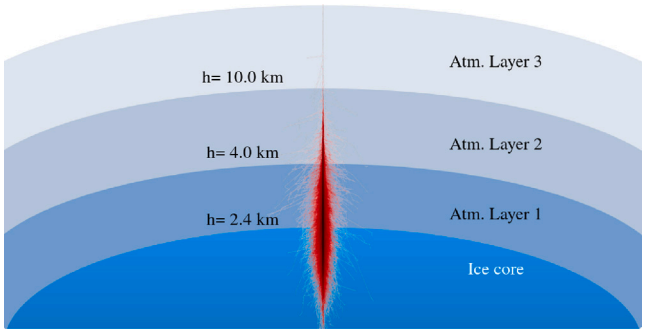

Fig. 17. Simulation geometry for the cross-media shower example. The setup includes a five-layer exponential atmosphere and an ice volume starting at 2.4 km above sea level. Only the first three atmospheric layers are shown.

and dense media simultaneously, most experiments opted to run the showers in two separate codes. For example, they combined Geant4 and CORSIKA 7 [92], saving particle data at the interface boundary and propagating it from one code to the other. While this approach has enabled detailed simulations of cross-media showers and their radio emission, a fully integrated solution would certainly be preferable.

CORSIKA 8 solves this issue with its new environment setup. It is now possible to run showers in environments with multiple types of media of various shapes, each with different properties, all in one go. For testing purposes, we built a realistic environment found in Antarctica consisting of a five-layer atmosphere and an ice layer starting at 2.4 km above sea level, as shown in Fig. 17.

We simulated a vertical shower induced by a 100 PeV proton primary injected at the top of the atmosphere, with a magnetic field of (9.07, 0, 61.80) μT intersecting the air–ice interface at an altitude of 2.4 km. The shower was simulated using a thinning level of 10^{-6}, a maximum weight of 50, an electromagnetic energy cut at 0.2 MeV, and a muon and hadronic energy cut at 0.3 GeV. The longitudinal profile of the different particles in the shower is shown in Fig. 18. While the profile from an electromagnetic shower would appear identical when plotted against grammage, even when transitioning from one medium to another, the hadronic component is sensitive to changes in density due to the interplay between the decay and interaction lengths. When the hadrons start propagating through ice, their interaction probability increases proportionally to the density increase (three orders of magnitude between air and ice). Consequently, hadrons start interacting with nuclei rather than decaying, producing the bump observed in Fig. 18. This also causes a slight increase in the electromagnetic component of the shower as π^0 particles decay into photons.

Lastly, the energy deposition of the shower in ice is shown in Fig. 19. The change in scale is evident, as the energy is primarily deposited in a cylinder measuring 10 m in length and 1 m in radius. This setup has also been used to test the results of CORSIKA 8 against those of CORSIKA 7 plus GEANT4 [93]. A good agreement is found between the frameworks [94]. Work on implementing the calculation of radio emission from cross-media showers with CORSIKA 8 is currently ongoing.

6. Electromagnetic emissions of showers

The astroparticle physics community requires not only Monte Carlo simulations of particle cascades in air and dense media, but also high-fidelity simulations of the electromagnetic emissions of these showers.

6.1. Cherenkov light emission

For several decades, the detection of extensive air showers via optical light, particularly Cherenkov and fluorescence light emission,

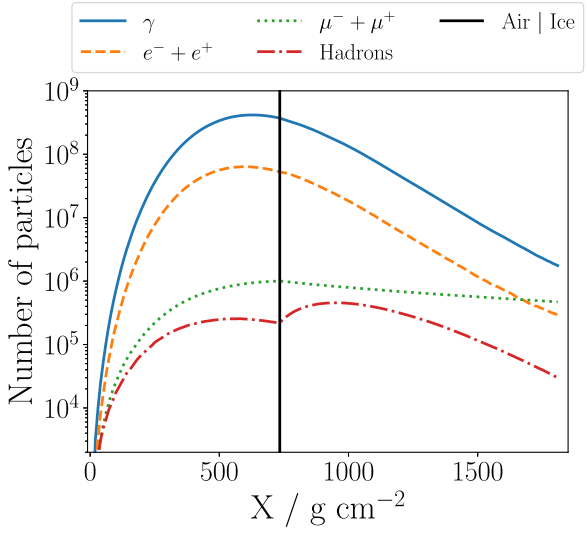

Fig. 18. Longitudinal profile of a vertical 100 PeV proton cross-media shower starting at the top of the atmosphere and intersecting an ice layer at an altitude of 2.4 km above sea level, as shown by the black line.

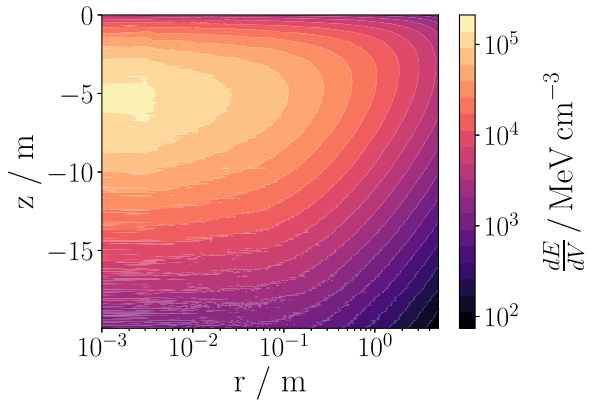

Fig. 19. Radial energy deposit of a 100 PeV proton vertical cross-media shower starting at the top of the atmosphere and intersecting an ice layer at an altitude of 2.4 km above sea level.

has been a cornerstone of high-energy cosmic-ray and gamma-ray astronomy. Advances in photon detection technologies, optical systems, and atmospheric monitoring have enabled increasingly precise air shower measurements using imaging Cherenkov and fluorescence light telescopes. These developments require simulation tools that can accurately and flexibly model the production and propagation of optical light in realistic experimental environments. Earlier generations of air-shower simulation codes tightly coupled optical light production to the shower simulation itself, limiting extensibility and making it difficult to adapt the implementation to novel detector concepts or complex geometries. Therefore, in CORSIKA 8, the simulation of optical light is being redesigned as a fully modular, first-class component of the framework, following the same design principles as other observational processes.

Simulating optical light can account for a significant portion of the total runtime. To address this issue, the optical light module currently being developed in CORSIKA 8 provides dedicated interfaces that allow for the early rejection of particle tracks and optical photons that do not contribute to measurable signals. This feature is especially important for high-energy cascades and for simulating of edge cases, in which only a portion of the photon field is within the sensitive area.

These interfaces allow the emission and propagation components to apply experiment-specific criteria, such as geometric visibility, wavelength acceptance, and detector thresholds, before incurring the full computational cost of photon generation and transport. Therefore, particle tracks that cannot produce Cherenkov light under the local refractive conditions or whose fluorescence emission cannot reach any active detector element can be discarded early on. Similarly, optical photons can be removed during propagation once it is determined that they no longer contribute to the detector response.

This approach maintains the purely observational nature of the optical processes while significantly reducing runtime for realistic detector geometries. It enables the efficient, large-scale production of simulated events and systematic studies without compromising physical accuracy in the measurable phase space.

From a computational perspective, a new approach is taken that enables easier compiler-based vectorization by disconnecting photons from the main simulation loop. This approach also lends itself to an asynchronous processing model for photons. For example, a GPU-based implementation is available [95].

The optical light module currently exists as a fork of the CORSIKA 8 project and has been validated against the established implementations used in CORSIKA 7 and in experiment-specific simulation chains. Benchmarks demonstrate good agreement in photon yields, ground distributions, and longitudinal profiles for both Cherenkov and fluorescence light. This confirms the physical consistency of the new implementation. The decoupled design also substantially improves maintainability and extensibility, providing a solid foundation for future developments [96]. The upcoming Cherenkov Telescope Array Observatory (CTAO) [97], in particular, will require simulations with a fidelity that is not currently achievable with CORSIKA 7, as CORSIKA 7 is limited by the concept of five-layer exponential atmospheres. Similarly, legacy codes do not currently support Cherenkov light simulations from upward or atmosphere-skimming air showers.

Work is underway to fully integrate the optical light module in the mainline branch of the project. The current implementation is primarily optimized for atmospheric Cherenkov and fluorescence observations. Photon propagation in ice requires dedicated treatments of additional effects related to the crystal structure of the ice, and is therefore left for future work.

6.2. Radio emission

Technological advancements in digital signal processing, combined with the excellent understanding of the radio emission phenomena have led to a significant progress in radio detection techniques for air showers over the past two decades [98]. As a result, experiments have been proposed to study the radio signals emitted from "non-standard" air showers. Examples include air showers crossing media from air to ice, refracting and/or reflecting radio signals at a boundary crossings, and experiments with an ever-increasing number of antennas. The community's leading standards for the simulation of the radio emission are the CORSIKA 7 code with its CoREAS extension [99] and the ZHAireS code [100]. However, both of these solutions suffer from the flexibility limitations of the underlying CORSIKA 7 and AIRES simulation codes.

To address these limitations, a radio module has been designed and developed as an integral part of CORSIKA 8. Following the modular design principles of the framework, the radio emission simulation is decomposed into four independent, user-configurable components: optional filtering of particle tracks, a formalism for calculating the radio emission, propagation of radio signals through complex media (potentially with multiple paths), and signal reception by antennas. Currently, the ZHS [101,102] and CoREAS ("endpoints") [103] algorithms in the time domain along with propagators that use a straight-ray approximation and time domain "ideal antennas", are fully implemented. One of the main advantages of the module's architecture is its ability to include

custom-made propagators designed for specific experimental setups. For example, it can handle air–ice transitions in cross-media showers, as discussed in the previous section and demonstrated in Ref. [94,104]. The radio process is implemented as a purely observational component and therefore does not influence the development of the air shower.

The radio module has been extensively validated as both a standalone feature and a part of CORSIKA 8. A comprehensive description of the module design and implemented algorithms, as well as detailed validation and benchmark studies, can be found in Ref. [105]. When sufficiently small step lengths are enforced, the total radio emission from CORSIKA 8 agrees with CORSIKA 7 to better than 10 % in the 30 MHz to 80 MHz band and is negligible in the 50 MHz to 350 MHz band. Therefore, we conclude that the radio module within CORSIKA 8 is ready for use in physics applications.

While the standard *radio process* is sufficient for most air shower applications, CORSIKA 8 is currently being extended to also handle radio emission and propagation in complex media. For smoothly-inhomogeneous environments in which media properties change on length scales that are large compared to the wavelength, geometric optics is an adequate description and radiation can be thought of as propagating along curved rays. CORSIKA 8 is being equipped with a general numerical raytracing propagator that can be used as part of the *radio process* [106]. This approach will allow for the accurate simulation of radio signals from neutrino-induced in-ice showers. To handle even more complicated situations where inhomogeneities occur on the scale of the wavelength, CORSIKA 8 is being interfaced with the external Eisvogel package [107,108]. Eisvogel provides a full-electrodynamics treatment of radio emission and propagation based on Green's functions [106]. This allows for accurate simulations in complex geometries and highly inhomogeneous media, such as radiation propagating in shallow ice, where wave-optics phenomena become relevant. Recently, CORSIKA 8 and Eisvogel have been used to model radio signals from air shower cores impacting polar ice sheets, enabling direct comparisons to experimental observations of in-ice Askaryan radiation [90].

7. Performance benchmarks

We benchmarked not only the simulation predictions of CORSIKA 8, but also its runtime performance against CORSIKA 7. Here,[3] we showcase its performance on a dedicated cluster node running AlmaLinux 9.6. The processor was a 12-core Intel Xeon running at 2.40 GHz. Only AVX vector extensions were available. Both codes were compiled with GCC 11.5 and the following optimization flags: `-O3 -mavx -march=native -mtune=native`, which ensured exploitation of the available vector instruction set and architecture-specific optimizations. To minimize hardware-induced variability in runtime measurements, Intel Turbo Boost and hyperthreading were disabled on the cluster node. Similarly, I/O operations were performed on a local disk to minimize latency.

To reduce statistical uncertainty on the measured runtimes, 10 runs were performed for each benchmark configuration. Each run consisted of 10, or 100 showers for short runtimes, to minimize fixed overheads, such as initialization, while keeping the overall computational cost reasonable. To get a more representative performance measurement, a distinct random number generator seed was assigned to each shower, so that showers in different runs but in the same position in the sequence start with the same seeds. Our benchmarks focus primarily on pure hadronic showers, with all electromagnetic particles removed from the simulation. A brief qualitative comment on electromagnetic shower performance and associated challenges is provided in a separate subsection. Identical physics configurations and primary showers were used in both CORSIKA 7 and CORSIKA 8. All benchmarks correspond to vertical showers.

[3] For these benchmarks, a slightly optimized version of the ICRC2025 release was used.

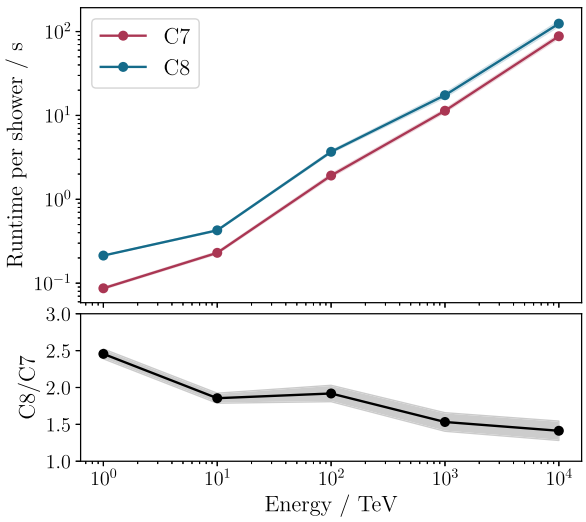

Fig. 20. Runtime per shower for proton-induced hadronic showers from 1 TeV to 10 PeV with all electromagnetic particles cut. Each point represents the mean runtime per shower, computed from 10 runs of either 10 or 100 showers. The shaded bands indicate the standard deviation over 10 runs.

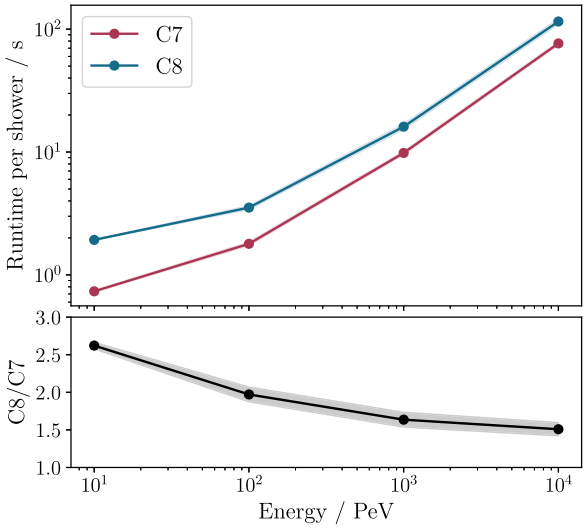

Fig. 21. Same as Fig. 20, but for energies ranging from 10 PeV to 10 EeV. An additional cut was added for hadronic particles below 1 TeV to reduce the overall runtimes for CORSIKA 7 and CORSIKA 8.

7.1. Hadronic showers

Fig. 20 shows the runtime per shower for proton-induced hadronic showers ranging from 1 TeV to 10 PeV using the SIBYLL 2.3d model for high-energy hadronic interactions. In this energy range, CORSIKA 8 is between 1.5 and 2.5 times slower than CORSIKA 7. For shorter showers (lasting only a few seconds), the ratio is closer to 2.5, likely due to larger initialization overheads in CORSIKA 8. For more energetic showers, the ratio decreases to about 1.5.

The results for higher-energy showers, ranging from 10 PeV to 10 EeV, are presented in Fig. 21. An additional cut on hadrons with energies below 1 TeV was applied here to keep runtimes at a manageable level. The same qualitative behavior is observed, with comparable runtime ratios between CORSIKA 7 and CORSIKA 8, and a similar dependence on the primary energy.

It is expected that CORSIKA 8 is somewhat slower than CORSIKA 7, as CORSIKA 7 has already undergone decades worth of optimization,

whereas CORSIKA 8 has not yet been systematically optimized. Therefore, the results shown here only constitute an initial performance baseline. Furthermore, it should be kept in mind that CORSIKA 7 is heavily hand-optimized for a limited number of particular use cases. This makes it very performant but also limits its flexibility. Thus, it is expected that the fully flexible approach of CORSIKA 8 cannot entirely reach the performance of CORSIKA 7. We consider a factor of 1.5 to be very acceptable, demonstrating that the chosen code design is fully adequate.

The benchmarks were repeated using QGSJet-II.04 and EPOS-LHC as hadronic interaction models. For QGSJet-II.04, runtime ratios similar to those obtained with SIBYLL 2.3d were observed. However, for EPOS-LHC, CORSIKA 8 was approximately 3.5 times slower than CORSIKA 7, a difference that will be investigated further in the future.

7.2. Electromagnetic showers

While hadronic showers in CORSIKA 8 perform only slightly worse than in CORSIKA 7, electromagnetic showers currently experience a roughly one order of magnitude slowdown in CORSIKA 8 compared to CORSIKA 7. The main difference between the approaches of CORSIKA 8 and CORSIKA 7 for simulating the electromagnetic cascade is that CORSIKA 8 uses PROPOSAL, which is interfaced by the CORSIKA 8 core functionality in a generic way. In contrast, CORSIKA 7 uses a strongly integrated custom version of EGS4 that handles a much larger part of the simulation than PROPOSAL does in CORSIKA 8. We are investigating the exact source of the slowdown, which could reside in the generic interface between CORSIKA 8 and PROPOSAL, the performance of PROPOSAL, or the tracking of electromagnetic particles in magnetic fields. As these investigations are still ongoing, no dedicated electromagnetic benchmark results are presented here.

8. Discussion

We have performed a detailed comparison of CORSIKA 8 with CORSIKA 7 for both electromagnetic and hadronic air showers. As shown in the preceding sections, the agreement between the two codes is generally within $\sim 10\,\%$. These validation studies demonstrate that the core physics implementations of CORSIKA 8 reproduce the well-established results of CORSIKA 7 for standard air shower scenarios and provide a solid baseline for further developments. For this purpose, CORSIKA 8 includes a standard application called `c8_air_shower` for air-shower simulations that closely mimics the traditional CORSIKA 7 setup. This application was used for the comparisons presented in this work.

Importantly, this level of agreement should not be interpreted as a trivial cross-check. For the first time in decades, these results provide an independent comparison of air-shower simulations based on two largely separate implementations that strive to use the same algorithms and physical models. At present, the origin of these differences is not fully understood. This comparison does not imply that one code is more correct than the other. Rather, it highlights the need for continued scrutiny of simulation tools that form the basis for the interpretation of experimental data. The observed differences at the level of $\sim 5 - 10\%$ therefore indicate the current level of intrinsic systematic uncertainty in state-of-the-art air-shower simulations.

At the same time, CORSIKA 8 is still subject to performance limitations and further optimization is ongoing. For this reason, the current recommendation is to continue using CORSIKA 7 for standard air shower simulations in the atmosphere of the Earth. However, for applications that are difficult or impractical to address within the constraints of the CORSIKA 7 code base, CORSIKA 8 is now the recommended choice. New investments in the legacy code base of CORSIKA 7 might not yield a favorable return in the long run.

To demonstrate the expanded applicability of CORSIKA 8, we present example applications involving particle cascades in media other than air, such as in the Martian atmosphere and water. These examples demonstrate the flexibility of the CORSIKA 8 framework with respect to atmospheric profiles, material properties, and geometry, and as its ability to handle cross-media showers in a unified and consistent way. Additionally, CORSIKA 8 provides the foundation for detailed studies of radio signal generation and propagation in dense media, such as ice, which is particularly relevant for neutrino detection experiments. An example of radio propagation in ice based on CORSIKA 8 can be found in Ref. [104].

At the current stage of development, some physical aspects are treated approximately. In particular, muon decays are modeled without polarization, which alters the angular distribution of decay electrons and positrons [109]. This induces a small effect on the resulting $e^{\pm}$ energy spectra, cf. Ref. [110]. However, we emphasize that similar limitations exist in CORSIKA 7. The treatment of the electromagnetic cascade in CORSIKA 7 relies on EGS4, subsequent developments and refinements, such as EGS5 [111] and EGSnrc [112], were never backported.

9. Conclusions and outlook

We have presented CORSIKA 8, a modern and modular framework for simulating extensive air showers and particle cascades in arbitrary media. Motivated by the long-term sustainability challenges and architectural limitations of CORSIKA 7, CORSIKA 8 was developed as a complete redesign rather than an incremental update. Its modular architecture has well-defined interfaces between physics processes, environment models, and geometry descriptions. This enables maintainable, extensible, and flexible simulations across a wide range of use cases.

The implemented electromagnetic and hadronic shower simulations have been validated through detailed comparisons with CORSIKA 7. For a broad set of observables, including longitudinal and lateral particle distributions and energy spectra, CORSIKA 8 reproduces the established results of its predecessor to a degree better than $10\,\%$. This is a significant achievement but also warrants further investigation. These differences should be understood as reflecting the current level of intrinsic systematic uncertainties in state-of-the-art air-shower simulations, rather than indicating a discrepancy between a "correct" or "incorrect" implementation.

Beyond reproducing traditional air-shower simulations, CORSIKA 8 is designed to support a broader range of applications. Its flexible environment model enables simulations in heterogeneous and non-standard media as well as arbitrarily complex geometries. Its modular process structure facilitates the integration of new interaction models. Modern data formats and accompanying Python analysis tools simplify integration into contemporary analysis pipelines.

Future work will focus on expanding the range of supported hadronic interaction models and enhancing the performance of the electromagnetic cascade simulation in particular. Thanks to its modular design and modern infrastructure, CORSIKA 8 provides a robust and future-proof foundation for air shower and cascade simulations in current and next-generation astroparticle physics experiments. We invite the astroparticle physics community to contribute to its further development and to provide feedback from user applications. All relevant information, including source code, issue tracking, and contribution guidelines, is available at the GitLab [6].

CRediT authorship contribution statement

Jean-Marco Alameddine: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Johannes Albrecht:** Supervision, Resources, Investigation, Funding acquisition. **Antonio Augusto Alves:** Supervision, Software, Project administration, Methodology, Formal analysis, Conceptualization. **Juan Ammerman-Yebra:** Writing – original draft, Validation, Software, Methodology, Conceptualization. **Luisa Arrabito:**

Supervision. **Dominik Baack:** Validation, Software, Resources, Project administration, Methodology, Investigation, Conceptualization. **Alan Coleman:** Validation, Software, Methodology, Conceptualization. **Cosmin Deaconu:** Resources, Funding acquisition. **Hans Dembinski:** Supervision, Software, Methodology, Investigation, Conceptualization. **Dominik Elsässer:** Supervision. **Ralph Engel:** Supervision, Funding acquisition, Conceptualization. **Alice Faure:** Writing – original draft, Visualization, Software, Formal analysis. **Alfredo Ferrari:** Writing – review & editing. **Chloé Gaudu:** Writing – review & editing, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Christian Glaser:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Marvin Gottowik:** Writing – original draft, Visualization, Validation, Software, Investigation, Formal analysis. **Dieter Heck:** Writing – review & editing. **Tim Huege:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Karl-Heinz Kampert:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Nikolaos Karastathis:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Jeffrey Lazar:** Writing – original draft, Validation, Software. **Lukas Nellen:** Validation, Resources, Project administration. **David Parello:** Supervision. **Tanguy Pierog:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization. **Remy Prechelt:** Writing – review & editing. **Radek Privara:** Validation, Software, Investigation. **Maximilian Reininghaus:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Wolfgang Rhode:** Supervision, Resources, Investigation, Funding acquisition. **Felix Riehn:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Maximilian Sackel:** Validation, Software, Conceptualization. **Pranav Sampathkumar:** Validation, Investigation. **Alexander Sandrock:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Conceptualization. **André Schmidt:** Software. **Jan Soedingrekso:** Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Ralf Ulrich:** Supervision, Software, Project administration, Methodology, Conceptualization. **Philipp Windischhofer:** Writing – review & editing, Validation, Software, Investigation. **Baobiao Yue:** Validation, Software, Investigation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Given his role as journal editor, Tim Huege had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank T. Sjöstrand, L. Lönnblad, and the Pythia 8 collaborators for their support in the implementation of Pythia 8/Angantyr in CORSIKA 8.

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 445154105 and Collaborative Research Center SFB1491 "Cosmic Interacting Matters - From Source to Signal". This research has been partially funded by the German Federal Ministry of Research, Technology and Space (BMFTR) and the state of North Rhine-Westphalia through the Lamarr Institute for Machine Learning and Artificial Intelligence. We acknowledge support through project UNAM-PAPIIT IN114924.

This work has also received financial support from Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación (PRE2020-09 2276). A. Coleman is supported by the Swedish Research Council (Vetenskapsrådet) under project no. 2021-05449. C. Glaser is supported by the Swedish Research Council (Vetenskapsrådet) under project no. 2021-05449, and the European Union (ERC, NuRadioOpt, 101116890). D. Baack is supported by the European Union (ERC, NuRadioOpt, 101116890). F. Riehn received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101065027. P. Windischhofer and C. Deaconu thank the National Science Foundation (NSF) for Award 2411662. R. Privara's work on this paper was financially supported by the Ministry of Education of the Czech Republic and the project MEYS Infra Auger LM2023032. J. Lazar is supported by the Belgian Fonds de la Recherche Scientifique (FRS-FNRS). The authors acknowledge support by the High Performance and Cloud Computing Group at the Zentrum für Datenverarbeitung of the University of Tübingen, the state of Baden-Württemberg through bwHPC and the DFG through grant no. INST 37/935-1 FUGG. The computations were partially carried out on the PLEIADES cluster at the University of Wuppertal, which was supported by the DFG (grant no. INST 218/78-1 FUGG) and the BMFTR.

Appendix A. Simulation setup

The CORSIKA 8 simulations are performed using the `c8_air_shower` standard application, which closely replicates the traditional CORSIKA 7 setup. The simulation uses a five-layer atmosphere based on Keilhauer's parameterization of the US standard atmosphere, as well as north-pointing magnetic field of $50\,\mu$T. All particles are propagated until they reach the observer level at sea level.

Unless stated otherwise, a thinning level of 10^{-6} is adopted, with a maximum allowed weight set to half of the optimal value (Kobal's optimum) to match the typical behavior of CORSIKA 7. Depending on the simulation setup, particle energy cuts vary and are specified below. In particle tracking, a maximum deflection of 0.2 radians is permitted.

For electromagnetic showers with a primary energy of 100 TeV, the energy cuts are 1 MeV for electromagnetic particles, 500 MeV for hadrons and muons, and 300 MeV for tau leptons. A typical command to run CORSIKA 8 simulations is:

```
c8_air_shower --pdg 22 --energy 1e5 --emcut 0.001
    --hadcut 0.5 --mucut 0.5 --nevent 1
    --seed {seed} --filename {filename}
    --observation-level 0e3
```

For electromagnetic showers with a primary energy of 100 EeV, the energy cuts are 100 GeV for electromagnetic particles, 100 GeV for hadrons and muons, and 300 MeV for tau leptons. A typical command is:

```
c8_air_shower --pdg 22 --energy 1e11
    --emcut 100e3 --hadcut 100e3 --mucut 100e3
    --nevent 1 --seed {seed}
    --filename {filename} --observation-level 0e3
```

For hadronic showers, the energy cuts are 10 MeV for electromagnetic particles, and 300 MeV for hadrons, muons, and tau leptons. A typical command is:

```
c8_air_shower -p 2212 -E 1e8 -z 0 -N 1
    --emcut 0.01 --mucut 0.3 --hadcut 0.3
    --max-weight 0 -s {seed} -M SIBYLL-2.3d
    -f {filename}
    --disable-interaction-histograms
```

CORSIKA 7 input cards for the electromagnetic showers have the following form:

```
RUNNR   {run_idx}
EVTNR   {first_event_idx}
NSHOW   {n_show}
PRMPAR  {primary}
ERANGE  100.E9  100.E9
THETAP  0.     0.
PHIP    -180.  180.
SEED    {seed_1}   0   0
SEED    {seed_2}   0   0
OBSLEV  0.E0
MAGNET  50  0
ELMFLG  F   T
ECUTS   100.E3 100.E3 100.E3  100.E3
LONGI   T  10.  F  T
ATMOD   17
EXIT
```

and for the hadronic showers the following form:

```
RUNNR   {run_idx}
EVTNR   {first_event_idx}
NSHOW   1
PRMPAR  14
ERANGE  1.0E8  1.0E8
THETAP  0.  0.
PHIP    0.  0.
SEED    {seed_1}   0   0
SEED    {seed_2}   0   0
OBSLEV  0
MAGNET  50 0
ECUTS   0.3 0.3 0.01  0.01
ELMFLG  F   T
LONGI   T  10.  T  T
THIN    1.0e-06 100. 0.
THINH   1.00 1.00
ATMOD   17
EXIT
```

Appendix B. Code listings

Here, we collect selected code listings that illustrate key design concepts discussed in the main text. In particular, they provide simplified examples of the mixin-based composition of medium interfaces and their corresponding implementations. The listings are not intended to be exhaustive, but rather to highlight the underlying patterns and demonstrate how different physical properties can be combined in a modular way.

Two objects, `modelA` and `modelB`, are created in Listing 2 each containing a different implementation of the `getMassDensity()` interface. The function `printMassDensity()` handles both objects via their interface, the `MediumInterface` class from Listing 1, and is agnostic about the implementation, which is selected only at runtime.

```
struct IMediumModel {
  virtual DensityType getMassDensity(Point const &) const = 0;
};

template <typename T> struct IRefractiveIndexModel : public T {
  virtual double getRefractiveIndex(Point const &) const = 0;
};

template <typename T> struct IMagneticFieldModel : public T {
  virtual MagneticFieldVector getMagneticField(Point const &) const = 0;
};

using MediumInterface = IMagneticFieldModel<IRefractiveIndexModel<IMediumModel>>;
```

Listing 1: Mixin-based environment interface composition (simplified). Separate interface components, each describing a specific physical property, are combined using template-based inheritance.

```
template <typename T> struct FlatExponentialDensity : public T {
  // [...]

  virtual DensityType getMassDensity(Point const &P) const override {
    return rho0 * exp(axis.dot(P - Pref) / scaleHeight);
  }
};

template <typename T> struct HomogeneousDensity : public T {
  // [...]

  virtual DensityType getMassDensity(Point const &p) const override {
    return rho0;
  }
};

template <typename T> struct ExponentialRefractiveIndex : public T {
  // [...]

  virtual double getRefractiveIndex(Point const &p) const override {
    // [...] some implementation
  }
};

template <typename T> struct UniformMagneticField : public T {
  // [...]

  virtual MagneticFieldVector getMagneticField(Point const &p) const override {
    // [...] some implementation
  }
};

void printMassDensity(MediumInterface const& medium) {
  Point const p = make_some_point(); // obtain a Point

  // query density at point p
  // concrete implementation selected via dynamic dispatch
  DensityType const rho = medium.getMassDensity(p);
  std::cout << "density at p = " << rho << std::endl;
}

int main() {
  ExponentialRefractiveIndex<
      FlatExponentialDensity<UniformMagneticField<MediumInterface>>>
      modelA;
```

```
  ExponentialRefractiveIndex<
      HomogeneousDensity<UniformMagneticField<MediumInterface>>>
      modelB;

  printMassDensity(modelA); // calls FlatExponentialDensity::getMassDensity()
  printMassDensity(modelB); // calls HomogeneousDensity::getMassDensity()

  return 0;
}
```

Listing 2: Mixin-based composition of implementations (simplified). Concrete classes implement individual physical properties and are combined via template-based inheritance. The resulting models are accessed via dynamic polymorphism.

Appendix C. Dependencies

For proper attribution and reproducibility, we list the main software libraries used by CORSIKA 8 and in the analysis presented in this work.

C.1. Corsika 8

- `Apache Arrow` [113]
- `Boost::histogram` [45,114,115]
- `CLI11` [116]
- `cnpy` [117]
- `Eigen` [18]
- `particle` (v0.25.1) [20]
- `spdlog` [118]
- `yaml-cpp` [119]

C.2. Software used for this paper

- `boost-histogram` (python)
- `matplotlib` [120,121]
- `mplhep` [122]
- `numpy` [46]
- `pandas` [123]

Data availability

Data will be made available on request.

References

[1] D. Heck, et al., CORSIKA: A Monte Carlo code to simulate extensive air showers, Tech. Rep. FZKA-6019, Forschungszentrum Karlsruhe, 1998, http://dx.doi.org/10.5445/IR/270043064.

[2] J.S. Sciutto, AIRES: A system for air shower simulations. User's guide and reference manual, 2019, http://dx.doi.org/10.13140/RG.2.2.12566.40002.

[3] CORSIKA – Helmholtz Research Software Directory, 2026, URL https://helmholtz.software/software/corsika. (Accessed: 29 April 2026).

[4] J. Reid, The new features of Fortran 2018, SIGPLAN Fortran Forum 37 (1) (2018) 5–43, http://dx.doi.org/10.1145/3206214.3206215.

[5] R. Engel, et al., Towards a Next Generation of CORSIKA: A Framework for the Simulation of Particle Cascades in Astroparticle Physics, Comput. Softw. Big Sci. 3 (1) (2019) 2, http://dx.doi.org/10.1007/s41781-018-0013-0.

[6] CORSIKA 8 Collaboration, CORSIKA 8, 2025, https://gitlab.iap.kit.edu/AirShowerPhysics/corsika.

[7] CORSIKA 8 Collaboration, CORSIKA 8, 2024, http://dx.doi.org/10.5281/zenodo.15525791, Concept DOI representing all released versions.

[8] G.M. Kurtzer, V. Sochat, M.W. Bauer, Singularity: Scientific containers for mobility of compute, PLoS One 12 (5) (2017) 1–20, http://dx.doi.org/10.1371/journal.pone.0177459.

[9] CORSIKA 8 Collaboration, CORSIKA 8, 2025, http://dx.doi.org/10.5281/zenodo.18195386, Zenodo. version: ICRC2025.

[10] M. Reininghaus, The air shower simulation framework CORSIKA 8: Development and first applications to muon production (Ph.D. thesis), Karlsruher Institut für Technologie (KIT) & Universidad Nacional de General San Martín, 2021, http://dx.doi.org/10.5445/IR/1000152097.

[11] ISO/IEC JTC1/SC22/WG5, Information Technology — Programming languages — Fortran — Units of measure for numerical quantities, Tech. Rep. N2113, 2016.

[12] N. Gehani, Units of measure as a data attribute, Comput. Lang. 2 (3) (1977) 93–111, http://dx.doi.org/10.1016/0096-0551(77)90010-8.

[13] S. McKeever, O. Bennich-Björkman, O.-A. Salah, Unit of measurement libraries, their popularity and suitability, Softw. Pr. Exper. 51 (4) (2021) 711–734, http://dx.doi.org/10.1002/spe.2926.

[14] Z.D. Umrigar, Fully Static Dimensional Analysis with C++, SIGPLAN Not. 29 (9) (1994) 135–139, http://dx.doi.org/10.1145/185009.185036.

[15] M. Moene, PhysUnits C++11. URL https://github.com/martinmoene/PhysUnits-CT-Cpp11.

[16] A.P. Raposo, The algebraic structure of quantity calculus, Measur. Sci. Rev. 18 (4) (2018) 147–157, http://dx.doi.org/10.1515/msr-2017-0021.

[17] S. Argirò, et al., The Offline Software Framework of the Pierre Auger Observatory, Nucl. Instrum. Meth. A 580 (2007) 1485–1496, http://dx.doi.org/10.1016/j.nima.2007.07.010.

[18] G. Guennebaud, B. Jacob, et al., Eigen, 2010, https://libeigen.gitlab.io.

[19] I.R. Shafarevich, A.O. Remizov, Linear algebra and geometry, Springer, 2013, pp. 289–317, http://dx.doi.org/10.1007/978-3-642-30994-6.

[20] E. Rodrigues, H. Schreiner, Particle. http://dx.doi.org/10.5281/zenodo.2552429 URL https://github.com/scikit-hep/particle.

[21] R.L. Workman, et al., Particle Data Group Collaboration Collaboration, Review of Particle Physics, PTEP 2022 (2022) 083C01, http://dx.doi.org/10.1093/ptep/ptac097.

[22] J.K. Salmon, et al., Parallel random numbers: as easy as 1, 2, 3, in: Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis, SC '11, 2011, http://dx.doi.org/10.1145/2063384.2063405.

[23] A.A. Alves, A. Poctarev, R. Ulrich, Counter-based pseudorandom number generators for CORSIKA 8 - A multi-thread friendly approach, EPJ Web Conf. 251 (2021) 03039, http://dx.doi.org/10.1051/epjconf/202125103039.

[24] F. James, A Review of Pseudorandom Number Generators, Comput. Phys. Commun. 60 (1990) 329–344, http://dx.doi.org/10.1016/0010-4655(90)90032-V.

[25] G. Marsaglia, A. Zaman, W. Wan Tsang, Toward a universal random number generator, Statist. Probab. Lett. 9 (1990) 35–39, http://dx.doi.org/10.1016/0167-7152(90)90092-L.

[26] C.A. Argüelles, et al., TAMBO Collaboration Collaboration, TAMBO: A Deep-Valley Neutrino Observatory, 2025, arXiv:2507.08070.

[27] Wavefront Technologies, Wavefront OBJ File Format. Format specification (Appendix B1, Advanced Visualizer Manual). URL https://www.loc.gov/preservation/digital/formats/fdd/fdd000507.shtml.

[28] G. Turk, The PLY polygon file format, 1994, Stanford University Computer Graphics Laboratory. URL https://paulbourke.net/dataformats/ply/.

[29] D. Nesteruk, Design Patterns in Modern C++: Reusable Approaches for Object-Oriented Software Design, A Press, 2018, http://dx.doi.org/10.1007/978-1-4842-3603-1.

[30] National Oceanic and Atmospheric Administration, National Aeronautics and Space Administration, United States Air Force, U.S. Standard Atmosphere, 1976, Tech. Rep NOAA-SIT-76-1562, National Oceanic and Atmospheric Administration, Washington, D.C., 1976.

[31] J. Cruz Moreno, S. Sciutto, Characterization of the atmospheric depth profile using the ground-level temperature: The case of Malargüe, Argentina, Eur. Phys. J. Plus 128 (2013) 104, http://dx.doi.org/10.1140/epjp/i2013-13104-3.

[32] B. Keilhauer, et al., Impact of varying atmospheric profiles on extensive air shower observation: Atmospheric density and primary mass reconstruction, Astropart. Phys. 22 (2004) 249–261, http://dx.doi.org/10.1016/j.astropartphys.2004.08.004.

[33] S. Chapman, The absorption and dissociative or ionizing effect of monochromatic radiation in an atmosphere on a rotating earth part II. Grazing incidence, Proc. Phys. Soc. 43 (5) (1931) 483–501, http://dx.doi.org/10.1088/0959-5309/43/5/302.

[34] D.L. Huestis, Accurate evaluation of the Chapman function for atmospheric attenuation, J. Quant. Spectrosc. Radiat. Trans. 69 (6) (2001) 709–721, http://dx.doi.org/10.1016/S0022-4073(00)00107-2.

[35] D. Vasylyev, Accurate analytic approximation for the Chapman grazing incidence function, Earth Planets Space 73 (2021) 112, http://dx.doi.org/10.1186/s40623-021-01435-y.

[36] D. Heck, The CURVED version of the air shower simulation program CORSIKA, Tech. Rep. FZKA-6954, Forschungszentrum Karlsruhe, 2004, http://dx.doi.org/10.5445/IR/270057207.

[37] A. Cillis, S. Sciutto, Geomagnetic field and air shower simulations, 1997, arXiv:astro-ph/9712345.

[38] G. Moliere, Theory of the scattering of fast charged particles. 2. Repeated and multiple scattering, Z. Naturforsch. A 3 (1948) 78–97.

[39] J.H. Koehne, et al., PROPOSAL: A tool for propagation of charged leptons, Comput. Phys. Commun. 184 (2013) 2070–2090, http://dx.doi.org/10.1016/j.cpc.2013.04.001.

[40] M. Dunsch, et al., Recent Improvements for the Lepton Propagator PROPOSAL, Comput. Phys. Commun. 242 (2019) 132–144, http://dx.doi.org/10.1016/j.cpc.2019.03.021.

[41] J.-M. Alameddine, et al., PROPOSAL: A library to propagate leptons and high energy photons, in: P. Teterin (Ed.), J. Phys. Conf. Ser. 1690 (1) (2020) 012021, http://dx.doi.org/10.1088/1742-6596/1690/1/012021.

[42] J.-M. Alameddine, et al., Improvements in charged lepton and photon propagation for the software PROPOSAL, Comput. Phys. Commun. 302 (2024) 109243, http://dx.doi.org/10.1016/j.cpc.2024.109243.

[43] P. Gutjahr, et al., Simulation of deflection uncertainties on directional reconstructions of muons using PROPOSAL, Eur. Phys. J. C 82 (12) (2022) 1143, http://dx.doi.org/10.1140/epjc/s10052-022-11102-5, [Erratum: Eur.Phys.J.C 83, 12 (2023)].

[44] Apache Software Foundation, Apache parquet, 2013, https://parquet.apache.org.

[45] H. Schreiner, et al., Boost-histogram: High-Performance Histograms as Objects, in: M. Agarwal, et al. (Eds.), Proceedings of the 19th Python in Science Conference, 2020, pp. 63–69, http://dx.doi.org/10.25080/Majora-342d178e-009.

[46] C.R. Harris, et al., Array programming with NumPy, Nature 585 (7825) (2020) 357–362, http://dx.doi.org/10.1038/s41586-020-2649-2.

[47] CORSIKA 8 Collaboration, CORSIKA 8 - Python Library, 2026, https://gitlab.iap.kit.edu/AirShowerPhysics/corsika/-/tree/main/python.

[48] CORSIKA 8 Collaboration, https://gitlab.iap.kit.edu/AirShowerPhysics/corsika-tools/-/tree/main/gui?ref_type=heads,

[49] W.R. Nelson, H. Hirayama, D.W.O. Rogers, The EGS4 Code System, (SLAC-265) Stanford Linear Accelerator Center, 1985, http://dx.doi.org/10.2172/1453993.

[50] K. Kamata, J. Nishimura, The Lateral and the Angular Structure Functions of Electron Showers, Prog. Theor. Phys. Suppl. 6 (1958) 93–155, http://dx.doi.org/10.1143/PTPS.6.93.

[51] K. Greisen, The Extensive Air Showers, in: J. Wilson (Ed.), in: Progress in Cosmic Ray Physics, vol. 3, North-Holland, Amsterdam, 1956, pp. 8–141.

[52] H.W. Koch, J.W. Motz, Bremsstrahlung Cross-Section Formulas and Related Data, Rev. Mod. Phys. 31 (1959) 920–955, http://dx.doi.org/10.1103/RevModPhys.31.920.

[53] L. Storm, H.I. Israel, Photon cross sections from 1keV to 100MeV for elements $Z = 1$ to $Z = 100$, Atom. Data Nucl. Data Tabl. 7 (1970) 565–681, http://dx.doi.org/10.1016/S0092-640X(70)80017-1.

[54] F. Sauter, Über den atomaren Photoeffekt bei großer Härte der anregenden Strahlung, Annalen Phys. 401 (1931) 217, http://dx.doi.org/10.1002/andp.19314010205.

[55] F. Sauter, Über den atomaren Photoeffekt in der K-Schale nach der relativistischen Wellenmechanik Diracs, Annalen Phys. 403 (1931) 454, http://dx.doi.org/10.1002/andp.19314030406.

[56] J.H. Hubbell, Photon cross-sections, attenuation coefficients, and energy absorption coefficients from 10keV to 100GeV, (NSRDS-NBS-29) National Bureau of Standards, 1969, http://dx.doi.org/10.6028/nbs.nsrds.29.

[57] D. Heck, Production of Vector Mesons by Photonuclear Interactions in the Program CORSIKA, KIT Sci. Work. Papers 6 (2012) 1, http://dx.doi.org/10.5445/IR/1000034381.

[58] A. Mücke, et al., SOPHIA: Monte Carlo simulations of photohadronic processes in astrophysics, Comput. Phys. Commun. 124 (2000) 290–314, http://dx.doi.org/10.1016/S0010-4655(99)00446-4.

[59] F. Riehn, et al., Hadronic interaction model SIBYLL 2.3d and extensive air showers, Phys. Rev. D 102 (6) (2020) 063002, http://dx.doi.org/10.1103/PhysRevD.102.063002.

[60] L.D. Landau, I. Pomeranchuk, Predely primenimosti teorii tormoznogo izlučenija élektronov i obrazovanija par pri bol'ših énergijah (Limits of applicability of the theory of electron bremsstrahlung and pair production at large energies), Dokl. Akad. Nauk. Ser. Fiz. 92 (1953) 535–536, In Russian.

[61] A.B. Migdal, Bremsstrahlung and pair production in condensed media at high-energies, Phys. Rev. 103 (1956) 1811–1820, http://dx.doi.org/10.1103/PhysRev.103.1811.

[62] T. Stanev, et al., Development of ultrahigh-energy electromagnetic cascades in water and lead including the Landau-pomeranchuk-migdal effect, Phys. Rev. D 25 (1982) 1291–1304, http://dx.doi.org/10.1103/PhysRevD.25.1291.

[63] D. Heck, J. Knapp, Upgrade of the Monte Carlo code CORSIKA to simulate extensive air showers with energies $> 10^{20}$ eV, Tech. Rep. FZKA-6097, Forschungszentrum Karlsruhe, 1998, http://dx.doi.org/10.5445/IR/270043705.

[64] A.M. Hillas, Two interesting techniques for Monte-Carlo simulation of very high energy hadron cascades, in: Proc. 17th Int. Cosmic Ray Conf., vol. 8, 1981, p. 193.

[65] S. García-Pareja, A.M. Lallena, F. Salvat, Variance-Reduction Methods for Monte Carlo Simulation of Radiation Transport, Front. Phys. 9 (2021) http://dx.doi.org/10.3389/fphy.2021.718873.

[66] M. Kobal, Pierre Auger Collaboration Collaboration, A thinning method using weight limitation for air-shower simulations, Astropart. Phys. 15 (2001) 259–273, http://dx.doi.org/10.1016/S0927-6505(00)00158-4.

[67] D. Heck, The MULTITHIN Option of the Air Shower Simulation Program CORSIKA, KIT Sci. Work. Papers 17 (2014) http://dx.doi.org/10.5445/IR/1000039436.

[68] B.T. Stokes, et al., Dethinning Extensive Air Shower Simulations, Astropart. Phys. 35 (2012) 759–766, http://dx.doi.org/10.1016/j.astropartphys.2012.03.004.

[69] P. Billoir, A sampling procedure to regenerate particles in a ground detector from a 'thinned' air shower simulation output, Astropart. Phys. 30 (2008) 270–285, http://dx.doi.org/10.1016/j.astropartphys.2008.10.002.

[70] T. Pierog, et al., EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider, Phys. Rev. C 92 (3) (2015) 034906, http://dx.doi.org/10.1103/PhysRevC.92.034906.

[71] S. Ostapchenko, Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJET-II model, Phys. Rev. D 83 (2011) 014018, http://dx.doi.org/10.1103/PhysRevD.83.014018.

[72] S. Ostapchenko, QGSJET-III model of high energy hadronic interactions: The formalism, Phys. Rev. D 109 (3) (2024) 034002, http://dx.doi.org/10.1103/PhysRevD.109.034002.

[73] T. Pierog, K. Werner, EPOS LHC-R : a global approach to solve the muon puzzle, PoS ICRC2025 (2025) 358, http://dx.doi.org/10.22323/1.501.0358.

[74] C. Bierlich, et al., A comprehensive guide to the physics and usage of PYTHIA 8.3, SciPost Phys. Codebases (2022) 8, http://dx.doi.org/10.21468/SciPostPhysCodeb.8.

[75] C. Gaudu, M. Reininghaus, F. Riehn, CORSIKA Collaboration Collaboration, From collider to cosmic rays: Pythia 8/Angantyr for air shower simulations in CORSIKA 8, PoS ICRC2025 (2025) 267, http://dx.doi.org/10.22323/1.501.0267.

[76] M. Bahr, et al., Herwig++ Physics and Manual, Eur. Phys. J. C 58 (2008) 639–707, http://dx.doi.org/10.1140/epjc/s10052-008-0798-9.

[77] C. Bierlich, et al., Robust Independent Validation of Experiment and Theory: Rivet version 3, SciPost Phys. 8 (2020) 026, http://dx.doi.org/10.21468/SciPostPhys.8.2.026.

[78] A. Karneyeu, et al., MCPLOTS: a particle physics resource based on volunteer computing, Eur. Phys. J. C 74 (2014) 2714, http://dx.doi.org/10.1140/epjc/s10052-014-2714-9.

[79] R. Ulrich, T. Pierog, C. Baus, Cosmic Ray Monte Carlo Package, CRMC, 2021, http://dx.doi.org/10.5281/zenodo.5270381.

[80] F. Ballarini, et al., The FLUKA code: Overview and new developments, EPJ Nucl. Sci. Technol. 10 (2024) 16, http://dx.doi.org/10.1051/epjn/2024015.

[81] A. Ferrari, et al., FLUKA: A multi-particle transport code (Program version 2005), Tech. Rep. CERN-2005-010, SLAC-R-773, INFN-TC-05-11, CERN-2005-10, CERN, 2005, http://dx.doi.org/10.2172/877507.
[82] S. Navas, et al., Particle Data Group Collaboration Collaboration, Review of particle physics, Phys. Rev. D 110 (3) (2024) 030001, http://dx.doi.org/10.1103/PhysRevD.110.030001.
[83] J.C. Arteaga Velazquez, A report by the WHISP working group on the combined analysis of muon data at cosmic-ray energies above 1 PeV, PoS ICRC2023 (2023) 466, http://dx.doi.org/10.22323/1.444.0466.
[84] J. Albrecht, et al., The Muon Puzzle in cosmic-ray induced air showers and its connection to the Large Hadron Collider, Astrophys. Space Sci. 367 (3) (2022) 27, http://dx.doi.org/10.1007/s10509-022-04054-5.
[85] J. Albrecht, et al., Global tuning of hadronic interaction models with accelerator-based and astroparticle data, Nature Rev. Phys. 8 (2026) 98–114, http://dx.doi.org/10.1038/s42254-025-00897-3.
[86] A.A. Alves, et al., CORSIKA Collaboration Collaboration, CORSIKA 8 - A novel high-performance computing tool for particle cascade Monte Carlo simulations, EPJ Web Conf. 251 (2021) 03038, http://dx.doi.org/10.1051/epjconf/202125103038.
[87] M. Reininghaus, R. Ulrich, T. Pierog, Air shower genealogy for muon production, PoS ICRC2021 (2021) 463, http://dx.doi.org/10.22323/1.395.0463.
[88] S.W. Barwick, C. Glaser, Radio Detection of High Energy Neutrinos in Ice, in: the Encyclopedia of Cosmology, World Scientific, 2023, pp. 237–302, http://dx.doi.org/10.1142/9789811282645_0006.
[89] P. Bagley, et al., KM3NeT Collaboration Collaboration, KM3NeT: Technical Design Report for a Deep-Sea Research Infrastructure in the Mediterranean Sea Incorporating a Very Large Volume Neutrino Telescope, 2009.
[90] N. Alden, et al., Observation of in-ice askaryan radiation from high-energy cosmic rays, Phys. Rev. Lett. (2026) http://dx.doi.org/10.1103/xwqy-yzrk.
[91] S. Agostinelli, et al., GEANT4 Collaboration Collaboration, GEANT4 - A Simulation Toolkit, Nucl. Instrum. Meth. A 506 (2003) 250–303, http://dx.doi.org/10.1016/S0168-9002(03)01368-8.
[92] S. De Kockere, et al., Simulation of radio signals from cosmic-ray cascades in air and ice as observed by in-ice Askaryan radio detectors, Phys. Rev. D 110 (2) (2024) 023010, http://dx.doi.org/10.1103/PhysRevD.110.023010.
[93] S. De Kockere, et al., Simulation of in-ice cosmic ray air shower induced particle cascades, Phys. Rev. D 106 (4) (2022) 043023, http://dx.doi.org/10.1103/PhysRevD.106.043023.
[94] J.-M. Alameddine, et al., CORSIKA Collaboration Collaboration, Simulations of cross media showers with CORSIKA 8, PoS ICRC2023 (2023) 442, http://dx.doi.org/10.22323/1.444.0442.
[95] D. Baack, W. Rhode, CORSIKA 8 Collaboration Collaboration, GPU Accelerated optical light propagation in CORSIKA 8, PoS ICRC2021 (2021) 705, http://dx.doi.org/10.22323/1.395.0705.
[96] D. Baack, J.-M. Alameddine, CORSIKA 8 Collaboration Collaboration, Comparison and efficiency of GPU accelerated optical light propagation in CORSIKA 8, PoS ICRC2023 (2023) 417, http://dx.doi.org/10.22323/1.444.0417.
[97] The Cherenkov Telescope Array Consortium, Science with the Cherenkov Telescope Array, World Scientific, 2019, http://dx.doi.org/10.1142/10986.
[98] T. Huege, Radio detection of cosmic ray air showers in the digital era, Phys. Rept. 620 (2016) 1–52, http://dx.doi.org/10.1016/j.physrep.2016.02.001.
[99] T. Huege, M. Ludwig, C.W. James, Simulating radio emission from air showers with CoREAS, in: R. Lahmann, T. Eberl, K. Graf, C. James, T. Huege, T. Karg, R. Nahnhauer (Eds.), AIP Conf. Proc. 1535 (1) (2013) 128, http://dx.doi.org/10.1063/1.4807534.
[100] J. Alvarez-Muniz, W.R. Carvalho Jr., E. Zas, Monte Carlo simulations of radio pulses in atmospheric showers using ZHAireS, Astropart. Phys. 35 (2012) 325–341, http://dx.doi.org/10.1016/j.astropartphys.2011.10.005.
[101] E. Zas, F. Halzen, T. Stanev, Electromagnetic pulses from high-energy showers: Implications for neutrino detection, Phys. Rev. D 45 (1992) 362–376, http://dx.doi.org/10.1103/PhysRevD.45.362.
[102] J. Alvarez-Muniz, A. Romero-Wolf, E. Zas, Cherenkov radio pulses from electromagnetic showers in the time-domain, Phys. Rev. D 81 (2010) 123009, http://dx.doi.org/10.1103/PhysRevD.81.123009.
[103] C.W. James, et al., General description of electromagnetic radiation processes based on instantaneous charge acceleration in 'endpoints', Phys. Rev. E 84 (2011) 056602, http://dx.doi.org/10.1103/PhysRevE.84.056602.
[104] A. Coleman, et al., In-ice Askaryan emission from air showers: Implications for radio neutrino detectors, Astropart. Phys. 172 (2025) 103136, http://dx.doi.org/10.1016/j.astropartphys.2025.103136.
[105] J.M. Alameddine, et al., CORSIKA 8 Collaboration Collaboration, Simulating radio emission from particle cascades with CORSIKA 8, Astropart. Phys. 166 (2025) 103072, http://dx.doi.org/10.1016/j.astropartphys.2024.103072.
[106] P. Windischhofer, A. Coleman, CORSIKA 8 Collaboration Collaboration, Simulating radio emissions from particle showers in complex media with CORSIKA 8, PoS ICRC2025 (2025) 1214, http://dx.doi.org/10.22323/1.501.1214.
[107] P. Windischhofer, C. Welling, C. Deaconu, Eisvogel: Exact and efficient calculations of radio emissions from in-ice neutrino showers, PoS ICRC2023 (2023) 1157, http://dx.doi.org/10.22323/1.444.1157.
[108] P. Windischhofer, C. Deaconu, C. Welling, Fully-electrodynamic radio simulations with Eisvogel, PoS ARENA2024 (2024) 051, http://dx.doi.org/10.22323/1.470.0051.
[109] L. Michel, Interaction between four half spin particles and the decay of the μ meson, Proc. Phys. Soc. A 63 (1950) 514–531, http://dx.doi.org/10.1088/0370-1298/63/5/311.
[110] P. Lipari, Lepton spectra in the earth's atmosphere, Astropart. Phys. 1 (1993) 195–227, http://dx.doi.org/10.1016/0927-6505(93)90022-6.
[111] H. Hirayama, et al., The EGS5 Code System, Tech. Rep. SLAC-R-730, Stanford Linear Accelerator Center, 2005, http://dx.doi.org/10.2172/877459.
[112] National Research Council Canada, EGSnrc: software for Monte Carlo simulation of ionizing radiation, 2021, http://dx.doi.org/10.4224/40001303.
[113] W. McKinney, et al., Apache Arrow: A Cross-Language Development Platform for In-Memory Data, Proc. the VLDB Endow. 11 (12) (2018) 1927–1930, http://dx.doi.org/10.14778/3229863.3229881.
[114] H.P. Dembinski, J. Pivarski, H. Schreiner, Recent developments in histogram libraries, in: C. Doglioni, D. Kim, G.A. Stewart, L. Silvestris, P. Jackson, W. Kamleh (Eds.), EPJ Web Conf. 245 (2020) 05014, http://dx.doi.org/10.1051/epjconf/202024505014.
[115] H. Schreiner, et al., Scikit-hep/boost-histogram, 2022, http://dx.doi.org/10.5281/zenodo.7094304.
[116] H. Schreiner, et al., Cliutils/CLI11, 2023, http://dx.doi.org/10.5281/zenodo.7502818.
[117] A. Kloeckner, Cnpy: Library to read/write numpy .npy and .npz files in C/C++, 2014, https://github.com/rogersce/cnpy.
[118] G. Melman, spdlog: Fast C++ logging library, 2024, https://github.com/gabime/spdlog.
[119] J. Beder, et al., yaml-cpp: A YAML parser and emitter in C++, 2024, https://github.com/jbeder/yaml-cpp.
[120] J.D. Hunter, Matplotlib: A 2D Graphics Environment, Comput. Sci. Eng. 9 (3) (2007) 90–95, http://dx.doi.org/10.1109/MCSE.2007.55.
[121] The Matplotlib Development Team, Matplotlib: Visualization with Python. http://dx.doi.org/10.5281/zenodo.592536, Zenodo.
[122] A. Novak, et al., Scikit-hep/mplhep, 2022, http://dx.doi.org/10.5281/zenodo.6807166.
[123] W. McKinney, et al., Data structures for statistical computing in python, Proc. the 9th Python Sci. Conf. (2010) 56–61, http://dx.doi.org/10.25080/Majora-92bf1922-00a.