Material budget studies in the ALICE central barrel using proton production asymmetry

Naufan Aurezan Mulyawan[1]
Supervisors: *Mesut Arslandok*[2,3], *Sandro Christian Wenzel*[2]
[1]Bandung Institute of Technology (ITB), Bandung, Indonesia
[2]CERN, Geneva, Switzerland
[3]Yale University, New Haven, CT 06520, USA

Abstract
A differential analysis is performed to investigate an observed overabundance of protons relative to antiprotons in ALICE Time Projection Chamber (TPC) data when compared to Monte Carlo (MC) simulations. Two event generators, PYTHIA and EPOS, are employed to model Pb–Pb collisions and to provide reconstructed MC samples for direct comparison with experimental data. Proton identification in both data and MC is based on TPC specific energy loss ($\mathrm{d}E/\mathrm{d}x$) and track-quality selections. The first stage of the analysis focuses on pseudorapidity distributions and yield ratios of protons to antiprotons. The second stage identifies the spatial and material origin of protons and antiprotons in the MC simulations using detailed detector geometry information. The comparison indicates a pronounced proton-antiproton imbalance in data that is not reproduced by the simulations.

Keywords
ALICE, TPC, specific energy loss ($\mathrm{d}E/\mathrm{d}x$), proton-antiproton asymmetry, material budget, Monte Carlo, PYTHIA, EPOS.

1 Introduction

It is known that the proton to antiproton ratio ($p/\bar{p}$) of produced particles in Pb–Pb collisions is unity and symmetric with respect to the collision vertex [1]. However, when protons originating from particle interactions with the detector material are included, this ratio exceeds one. A precise understanding of this contribution is essential for efficiency corrections and for controlling secondary-particle contamination in the majority of physics analyses. In earlier studies, an asymmetry of the $p/\bar{p}$ between the two sides of the detector was observed. This effect can be attributed to differences in the amount of material located upstream of the TPC, which were not accurately reproduced in Monte Carlo simulations. To trace the origin of this asymmetry, a detailed study of the material budget was carried out, comparing two models (PYTHIA and EPOS) against real data.

1.1 The ALICE experiment

The A Large Ion Collider Experiment (ALICE) at CERN's Large Hadron Collider (LHC) is a dedicated heavy-ion experiment designed to study strongly interacting matter at extreme temperatures and energy densities, where the formation of a quark-gluon plasma (QGP) is expected. Although ALICE's primary physics program is centered on heavy-ion (Pb–Pb) collisions, it also records other collision systems—such as proton–proton (pp), proton–nucleus (p–A), and oxygen–oxygen (O–O)—to provide essential baselines and to study collective phenomena in smaller systems. The central barrel hosts the Inner Tracking System (ITS) and the Time Projection Chamber (TPC) for precise tracking and vertexing. For this study, the TPC provides the principal particle identification via $\mathrm{d}E/\mathrm{d}x$ [2, 3].

1.2 Monte Carlo simulation framework

The workflow comprises (i) event generation using a physics model (PYTHIA, EPOS); (ii) detector simulation with the exact ALICE geometry and material budget; and (iii) event reconstruction using the same algorithms for data. We then compare detector-level η spectra and $p/\bar{p}$ ratios under identical reconstruction of both models to separate generator effects from geometry/transport effects.

PYTHIA is a general-purpose event generator for high-energy hadron collisions. It models hard QCD $2 \rightarrow 2$ scatterings, multiple parton interactions (MPI), initial and final-state radiation, and hadronization via the Lund string model. Soft diffraction is implemented using a Regge-based Pomeron model with extensions to hard diffraction. Despite extensive LHC tunings, PYTHIA is known to underpredict baryon yields, notably protons, at low transverse momentum in minimum-bias regimes, providing a baseline for identifying MC data deviations [4, 5].

EPOS is a minimum-bias event generator designed for pp, pA, and AA systems. It is based on a parton-based Gribov-Regge multiple-scattering framework and features a core-corona picture: dense regions form a collectively hadronizing core (with parametrized longitudinal and radial flow), while dilute regions hadronize via string fragmentation. The collective component tends to enhance baryon and strangeness production in high-density events, which can improve the description of identified-particle spectra in small systems [4]. EPOS provides an alternative baseline to PYTHIA for baryon production.

2 ALICE detector overview

Figure 1 shows the principal subsystems of the ALICE apparatus. The central barrel (approximately $|\eta| < 0.9$) provides charged-particle tracking and particle identification (PID), while forward detectors enable some specific capabilities (e.g. muon identification).

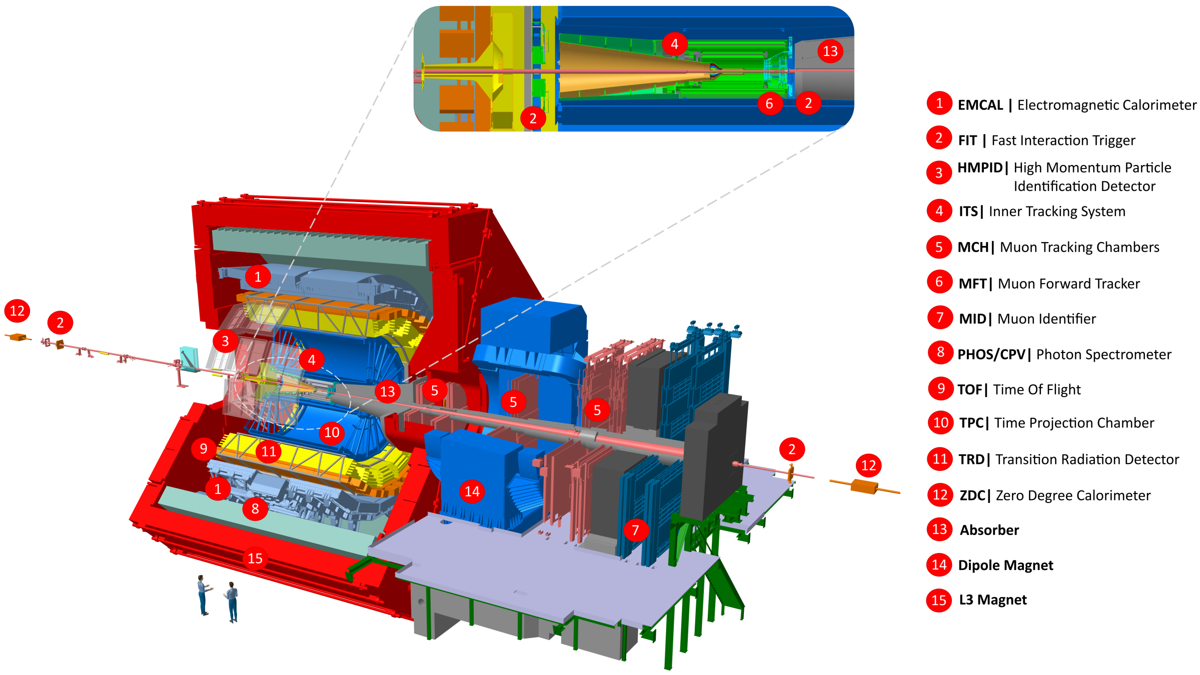

Figure 1: Schematic of the ALICE detector and its subsystems [6].

The two main subdetectors in the central barrel are the Inner Tracking System (ITS) and the Time Projection Chamber (TPC). The ITS is a silicon detector that improves primary and secondary vertex resolution as well as low-momentum tracking in the central barrel. It provides the innermost space points for global tracking and significantly enhances the impact-parameter resolution [3]. The TPC, a cylindrical gaseous detector, serves as the primary tracking and particle-identification detector of the central barrel. Following LS2, its GEM-based readout suppresses ion backflow to control space-charge distortions, enabling continuous readout at interaction rates of up to ∼50kHz in Pb–Pb collisions, while maintaining tracking and $\mathrm{d}E/\mathrm{d}x$ performance [3].

3 Analysis methodology

The analysis primarily relies on the ALICE TPC that delivers continuous 3D tracking and high-granularity $\mathrm{d}E/\mathrm{d}x$ measurements across a broad momentum range. In this work, proton and antiproton identification is achieved solely via TPC $\mathrm{d}E/\mathrm{d}x$ with additional track-quality selections.

3.1 Identification of protons and antiprotons

In total, about 1600 Pb–Pb simulated events were analyzed for each of the PYTHIA and EPOS models. For the experimental data, a comparable event statistics was selected to ensure consistency with the simulations. First, short and low-quality tracks are discarded to enhance the purity of the (anti-)proton identification. As shown in Figure 2, tracks are required to meet the following criteria: (i) $N_{\mathrm{cl}}^{\mathrm{TPC}} > 60$ TPC clusters, and (ii) $\chi^2\mathrm{TPC}/N\mathrm{cl}^{\mathrm{TPC}} < 2.5$. Since the dominant material contribution to protons occurs at momenta below 1 GeV, a graphical $\mathrm{d}E/\mathrm{d}x$ band cut is applied, as illustrated in the right panel of Figure 2, to isolate proton and antiproton candidates.

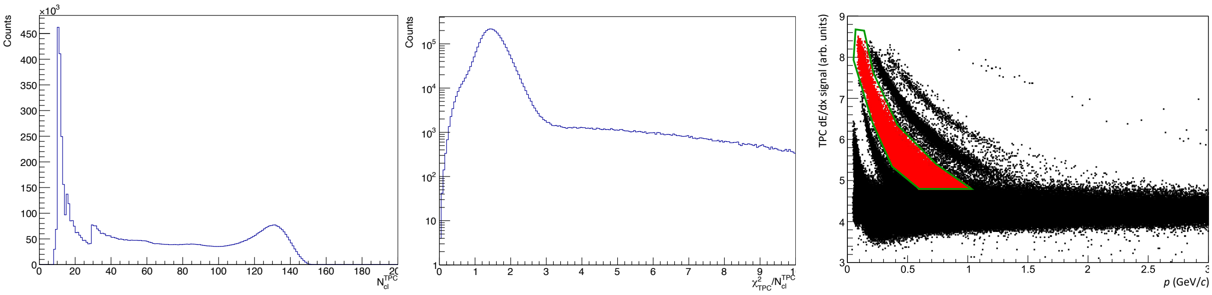

Figure 2: Distributions of $N_{\mathrm{cl}}^{\mathrm{TPC}}$ (left) and $\chi^2_{\mathrm{TPC}}/N_{\mathrm{cl}}^{\mathrm{TPC}}$ (middle) on experimental data; the graphical band cut applied on the energy loss distribution of protons below 1 GeV/C (right).

3.2 Material origin of protons and antiprotons in MC

To locate the spatial/material origin of protons and antiprotons in MC, we use the kinematics file to read truth records and fetch the production vertex (x, y, z). The ALICE simulation geometry (under the file `o2sim_geometry.root`) is imported into a `TGeoManager`. Then for each vertex we query the node and material via

```
gGeoManager->FindNode(x,y,z)->GetVolume()->GetMaterial()
```

and histogram yields by material name/ID. Contrasting PYTHIA and EPOS maps helps disentangle physics versus material-driven contributions to any observed excess.

4 Results

4.1 Imbalance of the $p/\bar{p}$ ratio in specific η regions

To investigate the origin of the aforementioned asymmetry, pseudorapidity (η) distributions for identified protons and antiprotons, as well as their ratios, were studied. Figure 3 compares $\mathrm{d}E/\mathrm{d}x$ vs. $q \cdot p$, where q denotes the electric charge of the particle, for all tracks, selected tracks, and (anti-)proton-band candidates, shown side-by-side for PYTHIA, EPOS, and data with the same PID selection criteria.

For clarity, the number of entries on each of these distributions are tabulated in Table B.1. An evident imbalance between protons and antiprotons can be observed in the histograms shown in the bottom row. Moreover, as shown in Figure 4, the normalized η distributions of protons (red) and antiprotons (blue) in data are significantly different from those in PYTHIA and EPOS. The difference between the two models can be attributed to the different particle spectra and underlying physics mechanisms.

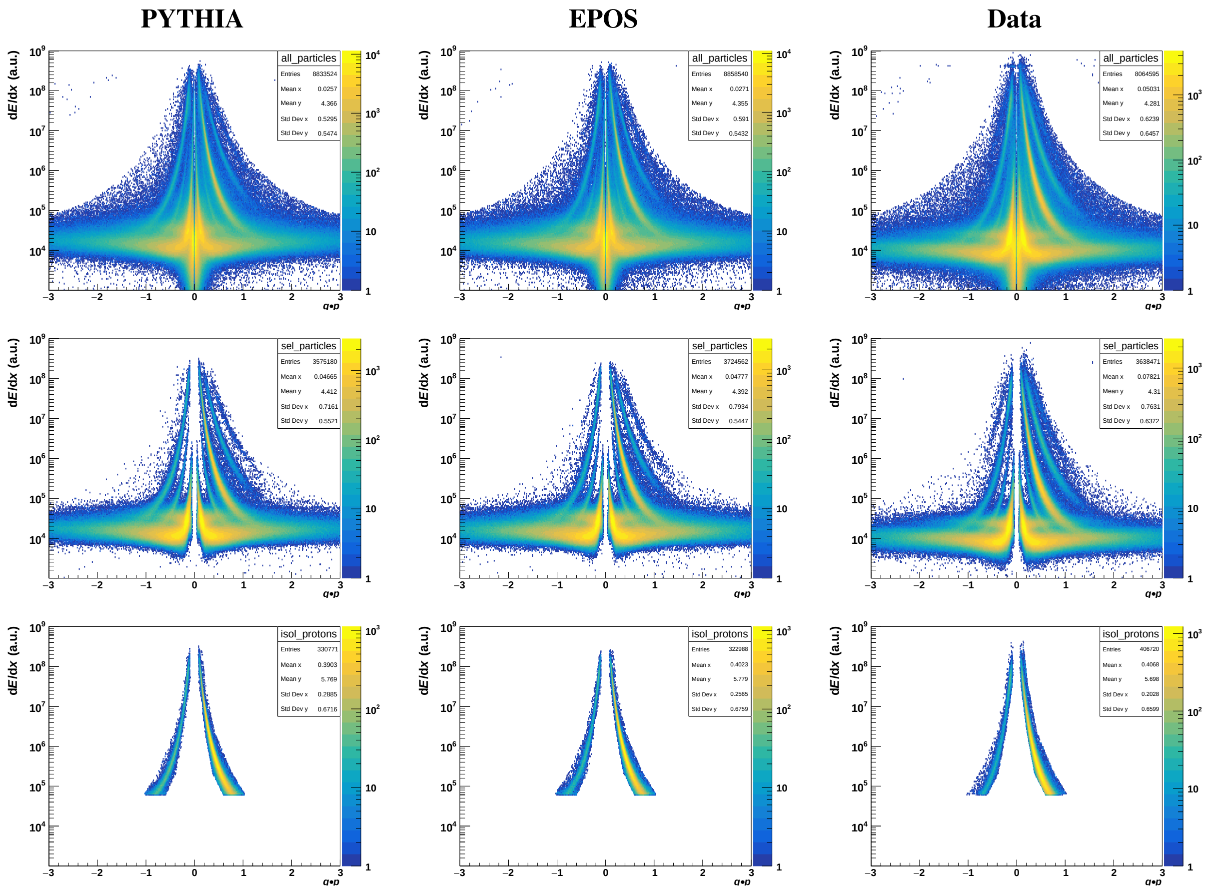

Figure 3: $\mathrm{d}E/\mathrm{d}x$ vs. $q \cdot p$ for (rows) all tracks, selected tracks, and proton-band candidates, shown for (columns) PYTHIA, EPOS, and data.

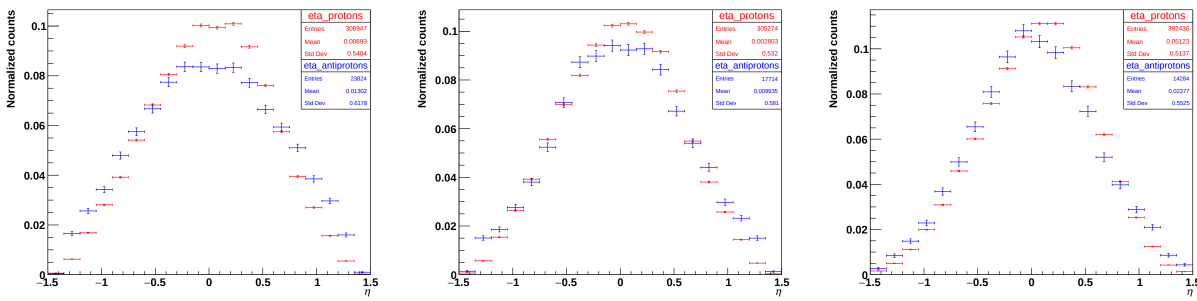

Figure 4: Normalized η distributions for protons (red) and antiprotons (blue) in PYTHIA (left), EPOS (center), and data (right).

Figure 5 presents the proton-to-antiproton ratio (($p/\bar{p}$ vs. η) in raw (left) and normalized (right) form. From the *raw* ratios, it is observed that the experimental data are significantly higher than the models, with the asymmetry at positive η values around 0.25–0.75 being rather pronounced. The *area-normalized* comparison (right) brings the shapes into closer agreement, indicating that the discrepancy arises mainly from a *yield-scale offset* rather than from a genuine η-shape difference. This may indicate a missing simulation of the relative material imbalance in those specific η regions.

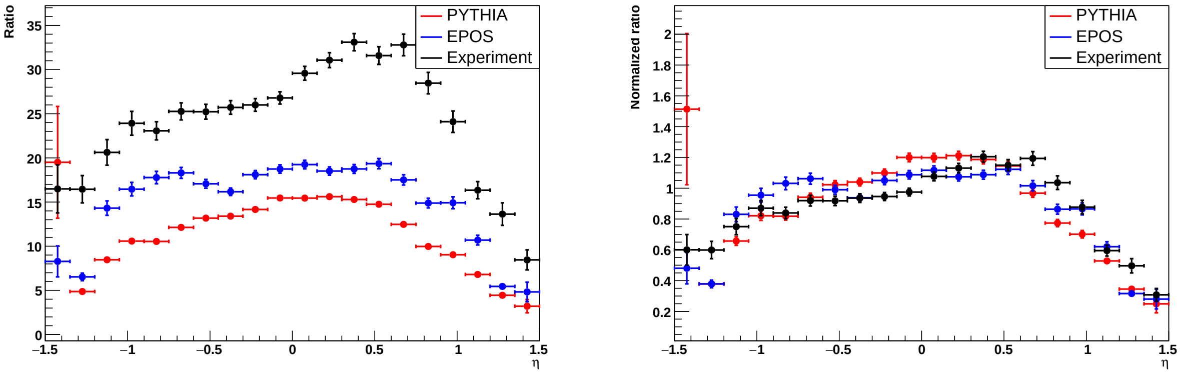

Figure 5: $p/\bar{p}$ ratio vs. η. Left: raw ratios for PYTHIA (red), EPOS (blue), and data (black). Right: ratios normalized to unit area.

4.2 Material-origin analysis in MC

The observed pattern in $p/\bar{p}$ points to three mechanisms: (i) additional *secondary proton production* in material (raising proton counts), (ii) enhanced *antiproton loss* by annihilation in matter (lowering antiprotons), and (iii) *generator limitations*. To test mechanisms (i)–(ii), we localize (anti-)proton vertices by querying the `TGeoManager` at the production point. In this study, we are focusing more on vertices before the TPC as shown by the cross-section in Figure 1.

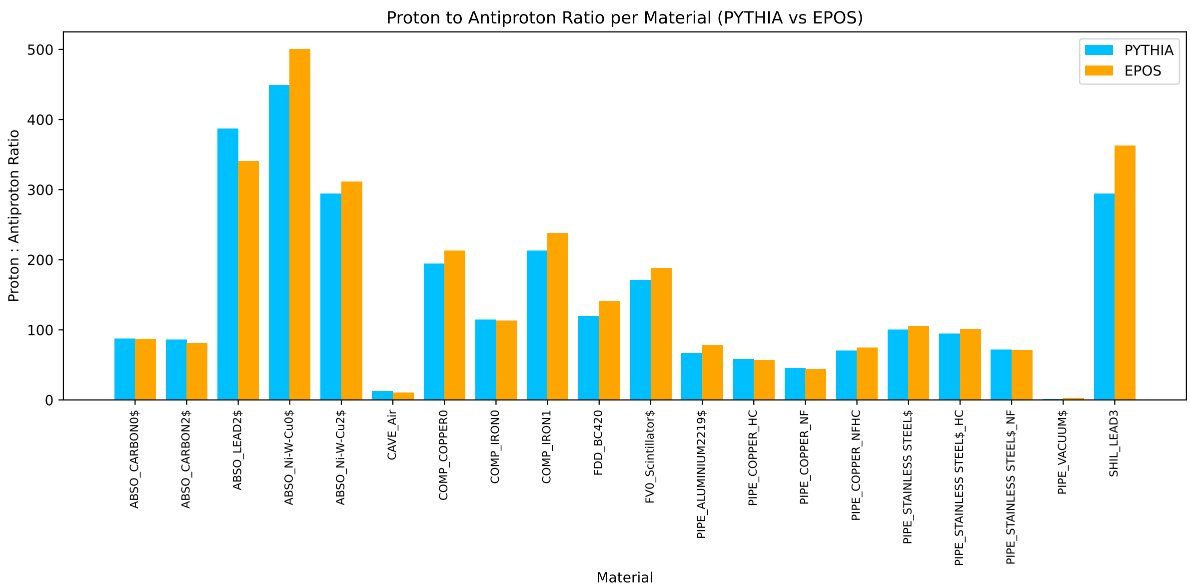

Figure 6: Proton to antiproton ratio for different materials in MC truth. Categories correspond to TGeo-Material names.

Figure 6 compares $p/\bar{p}$ ratios by material category. This shows that dense materials can be a factor in proton production hot-spots. Categories associated with absorbers, beam-pipe alloys, supports, and local shielding (e.g. stainless steel, Al/Cu, W/Pb composites) contribute disproportionately to proton vertices. In addition, antiproton vertices are comparatively more frequent in or near vacuum volumes. Those produced inside dense regions would preferentially annihilate before reconstruction and thus lowering their contribution. Figure 7 shows the distribution of (anti-)proton vertices on PYTHIA and EPOS with radius as the y-axis and z-cross-section as the x-axis. Both models produce the same leading categories (top 20 by proton count) with EPOS tending to populate outer volumes slightly more strongly.

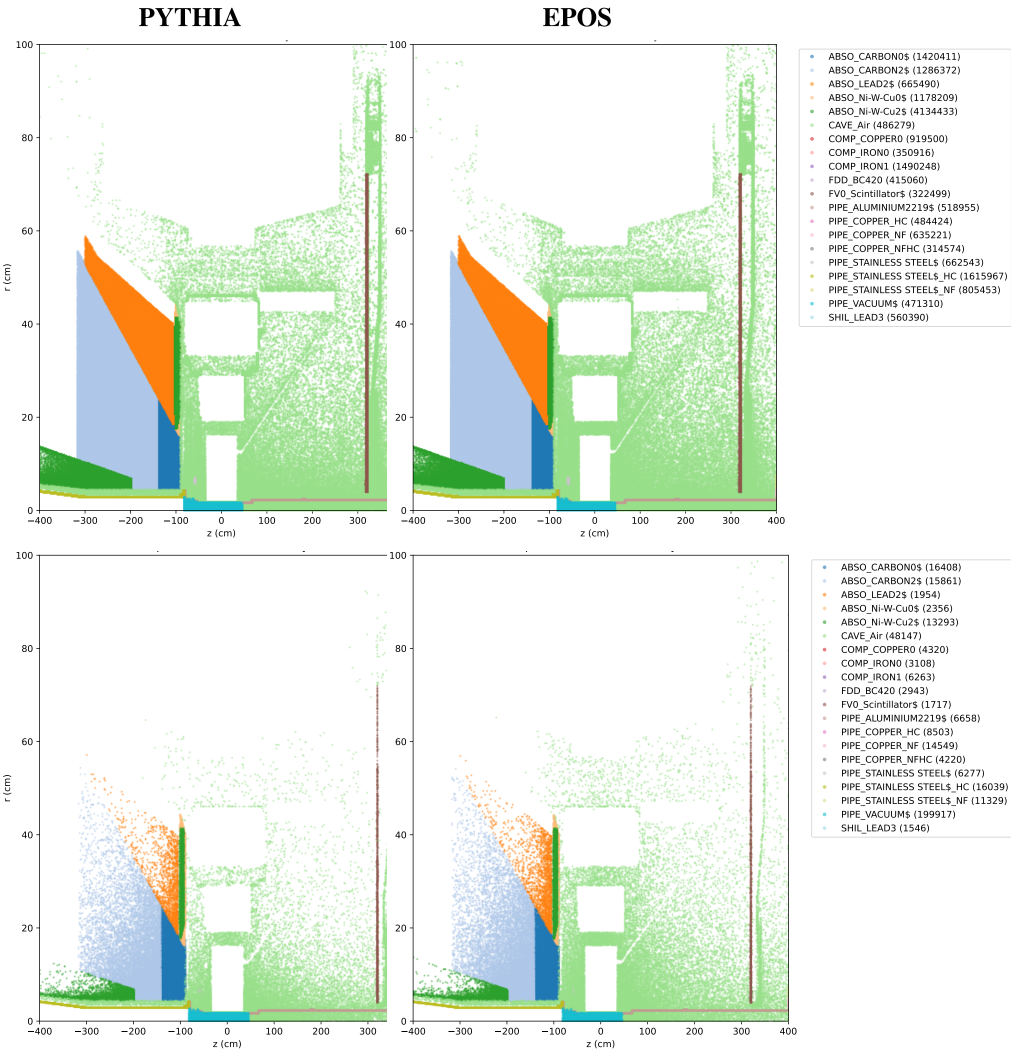

Figure 7: Distributions for protons (top) and antiprotons (bottom) of top 20 materials in PYTHIA and EPOS. TPC area is located on radius $\gtrapprox 85$.

5 Future directions

While the present analysis provides strong evidence for a data–simulation discrepancy in proton-to-antiproton yields, several aspects remain to be systematically explored. The following steps are proposed to refine the interpretation of the observed asymmetry and to isolate its possible origins:

1. Separate primary and secondary protons in MC truth and repeat the material analysis for secondaries only to isolate material effects.
2. Bin vertices in certain radius, project to η, and correlate with known structures.
3. Scan hadronic lists/parameters (especially $\bar{p}$ annihilation and inelastic cross sections) and compare category-wise ratios in Figure 6.
4. Vary thicknesses or densities in suspect volumes to bracket the observed data/MC offset.

6 Conclusions

A comparative study of proton and antiproton yields in ALICE TPC data and reconstructed MC indicates a pronounced proton excess on the A-side in data not reproduced by PYTHIA and EPOS. The material-tagging analysis localizes much of the imbalance to specific detector materials and categories. These observations point to: (i) refining geometry/material descriptions and validations using the identified volumes; (ii) scans of hadronic-interaction settings in the transport, including $\bar{p}$ annihilation/absorption systematics; and (iii) updated generator tunes or models targeting soft baryon production.

Acknowledgements

I thank Mesut Arslandok and Sandro Christian Wenzel for their supportive guidance and constructive discussions, the team behind ROOT software, and CERN's LXPLUS service for the computing resources used throughout this work.

Bibliography

[1] ALICE Collaboration, "Measurements of Chemical Potentials in Pb-Pb Collisions at 5 TeV", Phys. Rev. Lett. 133 (2024) 092301.

[2] ALICE Collaboration, "ALICE Experiment," CERN, `https://home.cern/science/experiments/alice`.

[3] ALICE Collaboration, *ALICE LS2 upgrades and detector systems* (detector overview and subsystem roles); internal/overview materials used in ALICE education and upgrade notes.

[4] T. Pierog, Iu. Karpenko, J. M. Katzy, E. Yatsenko and K. Werner, "EPOS LHC: Test of collective hadronization with LHC data," *JHEP* **06** (2013) 015, arXiv:1306.0121.

[5] T. Sjöstrand *et al.*, "An introduction to PYTHIA 8.2," *Comput. Phys. Commun.* **191** (2015) 159–177.

[6] CERN JIRA ATO-619, "Material budget investigation – O2 geometry (data volume studies ATO-609)," goal and notes on excess secondary protons/electrons, related simulation paths and utilities, `https://its.cern.ch/jira/browse/ATO-619`.

Appendix

A Reproducibility and computing environment

Analyses are executed on a CERN Linux node (LXPLUS) with the ALICE software environment and GRID access. Event generation and full simulation are run with the O2 stack (e.g. `O2sim` with `TGeant4` and `Pythia8` for Pb-Pb) writing `o2sim_Kine.root`, detector hit files (e.g. `o2sim_HitsTPC.root`), and geometry (`o2sim_geometry.root`). Visualizations are made using Jupyter on the server with a web browser client via SSH tunneling.

B Supplementaries

Table B.1 summarizes the number of entries that is shown in Figure 3. Additional figures of individual raw and normalized η ratios and protons-antiprotons material distributions can also be found in Figure B.1–B.2.

Table B.1: Entry counts shown in Figure 3.

Sample	All	Sel.	p	$\bar{p}$	$p/\bar{p}$
PYTHIA	8,833,524	3,575,180	306,947	23,824	12.88
EPOS	8,858,540	3,724,562	305,274	17,714	17.23
Data	8,064,595	3,638,471	392,436	14,284	27.47

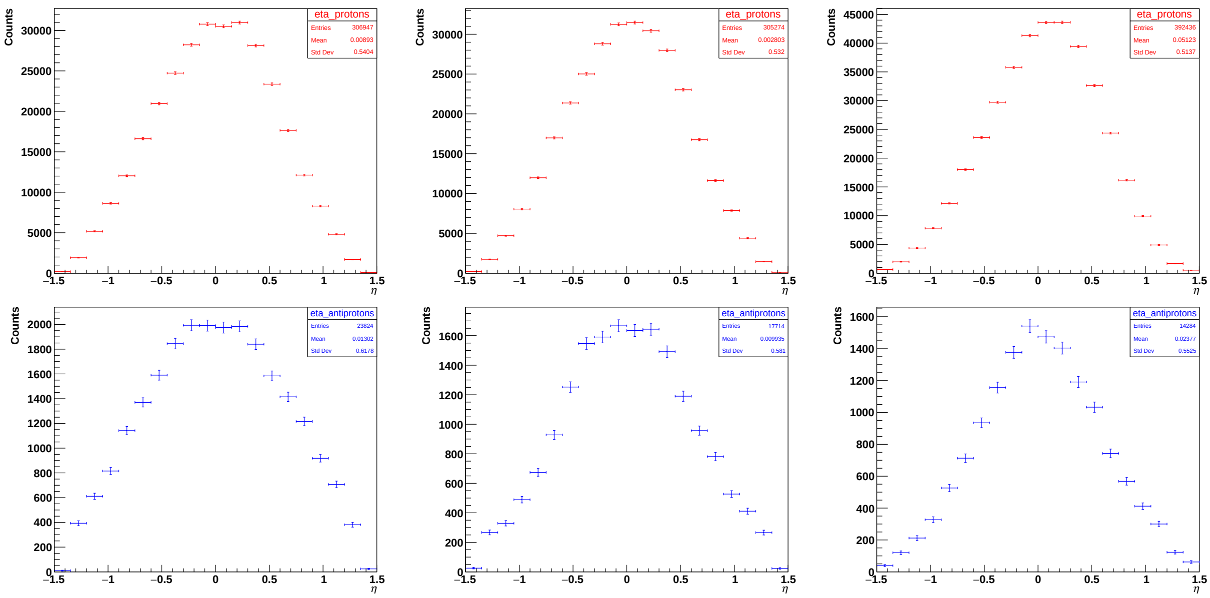

Figure B.1: η distributions for protons (top, red) and antiprotons (bottom, blue) in PYTHIA (left), EPOS (center), and data (right).

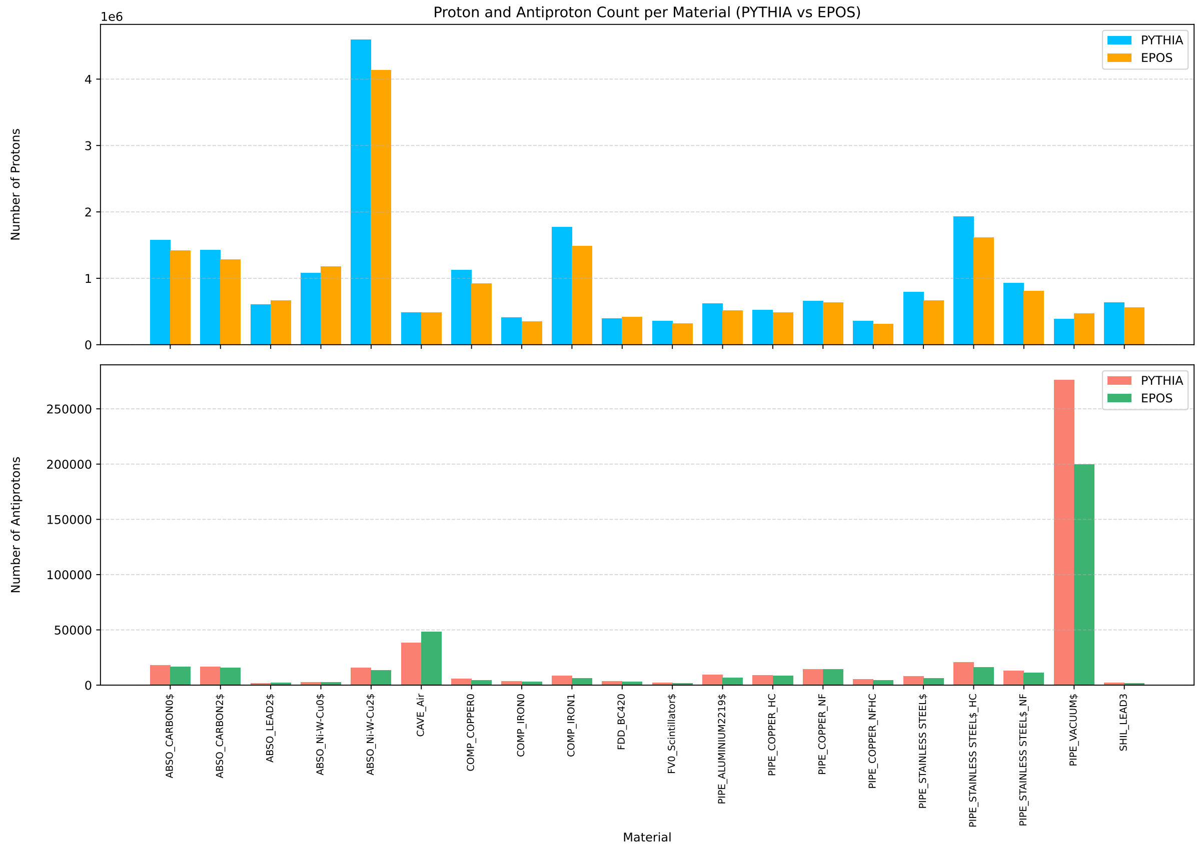

Figure B.2: Material-category yields for protons (top) and antiprotons (bottom) in MC truth. Categories correspond to `TGeoMaterial` names.