

CONSEIL EUROPÉEN POUR LA RECHERCHE NUCLÉAIRE

SUMMER STUDENT PROJECT REPORT

Determining the Spin of the Hypertriton

Elli Jobst

Technical University of Munich

03. July 2023 - 29. September 2023



ALICE

Supervisors: Maximiliano Puccio, Sourav Kundu

Abstract

This Report presents the results of my Summer Student Project. The goal was to determine the spin of the Hypertriton by analyzing ALICE Run2 Pb-Pb at $\sqrt{s_{NN}} = 5.02\text{TeV}$ Data. Firstly some theoretical background information is established and the strategy explained. The taken steps are described and the obtained results are presented. The Hypertriton invariant mass signal was successfully extracted from the data and an analysis chain was developed. The observed effect is not statistically significant enough to determine the spin, but with Run3 data we will be able to find it.

Acknowledgements

Many thanks to my supervisors Maximiliano and Sourav, who supported me throughout my time at CERN and to the ALICE group who welcomed me as a Summer Student. Many thanks also to CERN for giving me the opportunity to come and work here for 3 months, and to my professors Laura Fabbietti and Elisa Resconi for helping with my application. Lastly I want to thank all the amazing people and friends I met at CERN. I had an amazing time.

Contents

1	Introduction on CERN and the ALICE Experiment	3
2	Motivation for Studying the Hypertriton	3
3	How to Measure the Spin	4
3.1	Angular Distributions for Spin 1/2 and Spin 3/2	4
3.2	Exploiting Heavy Ion Collisions	6
4	Data Analysis	7
4.1	Choosing the Right Data	7
4.1.1	Detector Acceptance	7
4.1.2	Training a Model Using Boosted Decision Trees	9
4.1.3	Applying the Model to the Data	10
4.2	Extracting the Hypertriton Spin	12
4.2.1	Fitting the Invariant Mass Spectrum	12
4.2.2	Extracting the Angular Distribution	13
4.2.3	Selecting on Event Plane Angle and Centrality	14
4.3	Systematic Uncertainties	16
5	Outlook	18

1 Introduction on CERN and the ALICE Experiment

The European Organization for Nuclear Research (CERN) is one of the biggest laboratories for nuclear research in the world. Located near the city of Geneva, CERN has been founded in the 1950s and has been contributing to science since then. CERN is home to the biggest particle collider in the world, the Large Hadron Collider (LHC), which was started up in 2008 and was a key element in the discovery of the Higgs Boson [1] in 2012. The LHC lies roughly 100m underground and has a circumference of 27km. There hadrons, like protons and heavier ions, can be accelerated to almost the speed of light, and collided with one another. In these collisions the boundaries of physics can be probed by analyzing the particles that get created under the extreme collision conditions. Right now the LHC has started Run3 of its data taking, which will last until 2025.

Along the LHC there are 8 points, where collisions happen - 4 of them house giant detectors that collect the data from the particle collisions. At point 2 across the french border the Detector of A Large Ion Collider Experiment (ALICE) is located. Here mostly heavy ion - like Lead - collisions are probed. In these highly energetic collisions a new state of matter the so called Quark Gluon Plasma (QGP) is formed, where the single quarks are no longer bound in nuclei, and can form rare, exotic particles. By looking at these collisions - which resemble the state our universe was in just after the Big Bang - new and exciting physics can be discovered and studied.[2]

2 Motivation for Studying the Hypertriton

The Hypertriton, is the simplest hypernucleus one can imagine. Its constituents are a closely bound proton and neutron and a lambda baryon, whose quark content is (uds). Compared to the binding energy of a deuteron ($B_D = 2.2\text{MeV}$), the binding energy of the Λ is very small ($B_\Lambda = 100\text{keV}$), which almost makes it seem like a bound state of a deuteron(pn) and a Λ . Up till now the ALICE Collaboration has measured the lifetime and mass of the Hypertriton successfully [3]. Now we want to determine its spin.

The reason this nucleus is studied is that it can help giving the answer to a very fundamental question of physics: Nucleosynthesis, how nuclei firstly formed in the beginning of the universe. It is a very convenient tool, since the Λ does not follow Pauli's Exclusion Principle since it has strangeness and the proton and neutron do not, and secondly since the Λ is very loosely bound, the Hypertriton's wave function is very broad, spanning almost 10fm [4]. This makes it very easy to distinguish between two competing theories: the Coalescence Model [5] and Statistical Hadronization [6]. The two models predict very different Hypertriton yields in p-Pb collisions, since the wave function is much larger than the hadron emission radius [4] and the ratio of nucleus size to source size enters the yield in the Coalescence model but not in the Statistical Hadronization model. Thus the production of Hypertritons would be suppressed in the Coalescence model. The spin directly enters the calculations for the different models and is therefore important for the analysis. Studying the Hypertriton spin is also relevant for understanding Nucleon-Hyperon Interaction at the center of neutron stars . It is theorized that in the core of neutron stars, due to the high densities, the nucleon chemical potential is large enough to result in the production of Hyperons. It is believed that there are forces acting between the nuclei and hyperons, that can alter the properties of the neutron stars and explain the puzzling results that are measured [7].

The Hypertriton spin is lastly also very important for dark matter search in space, where the fluxes of nuclei and antinuclei are measured, for example in the AMS-02 Experiment [8]. Possible dark matter candidates can annihilate and produce (anti)nuclei in this process. Since antiparticles are very rarely produced in secondary processes, the increased yield of antiparticles due to dark matter, would result in a measurable effect. Depending on the spin the Hypertriton and Antihypertriton yields will be very different, thus it is crucial to determine its spin. Furthermore also nucleosynthesis plays a huge role in determining the expected (anti)nuclei yields.

3 How to Measure the Spin

3.1 Angular Distributions for Spin 1/2 and Spin 3/2

In order to determine the spin of the Hypertriton Run2 data from 2018 is analyzed. The Hypertriton itself has a lifetime of only around 260ps [3], thus we need to look at the decay daughters to be able to reconstruct the Hypertritons. We are analyzing the weak, two-body, mesonic decay channel

$$\left({}^3_{\Lambda}\text{H} \rightarrow {}^3\text{He} + \pi^{-}\right) + c.c. \quad (1)$$

in this decay parity is violated, and the Helium-daughter is preferably emitted in the same direction as the spin of the Hypertriton in the Hypertriton rest frame.

Since the Hypertriton is composite of 3 spin 1/2 particles, its spin can either be 1/2 or 3/2 for the ground state. Using simple quantum mechanics it can be derived that depending on its spin the angular distribution of the Helium daughter will look differently [9].

In the following the used cartesian coordinate system used will have the z-direction parallel to the beam and the $\hat{x}$ unit vector pointing towards the center of the LHC.[10]

Spin 1/2 Case

If the spin of the Hypertriton is 1/2, all possibilities for the final state can be listed. Since the total angular momentum J should be conserved $J_{initial} = J_{final}$ and the Hypertriton itself has no angular momentum $J_{initial} = S_{\Lambda\text{H}}$ one can derive all the possible combinations. The Pion's spin is 0, while the Helium has a spin of 1/2, therefore in this case the angular Momentum $L({}^3\text{He}+\pi^{-})$ can be either 0 or 1. The elements of the decay matrix are

$$\begin{aligned} T(1/2 \rightarrow 1/2, 0)_{11} &= T_s Y_{00} + \frac{1}{\sqrt{3}} T_p Y_{10} \\ T(1/2 \rightarrow 1/2, 0)_{-11} &= -\sqrt{\frac{2}{3}} T_p Y_{11} \\ T(1/2 \rightarrow 1/2, 0)_{1-1} &= \sqrt{\frac{2}{3}} T_p Y_{1-1} \\ T(1/2 \rightarrow 1/2, 0)_{-1-1} &= T_s Y_{00} - \frac{1}{\sqrt{3}} T_p Y_{10} \end{aligned} \quad (2)$$

where the respective amplitudes for the s-wave and p-wave state were considered if $L=0,1$ was possible, and the Clebsch-Gordan Coefficients were multiplied by the respective spherical harmonic Y_{lm} . The index of the matrix element describes the spin alignment of the Helium and the Hypertriton $2m_{\text{He}}2m_{\Lambda\text{H}}$. Inserting the explicit form of the spherical harmonics, the full decay matrix T can be obtained

$$T(1/2 \rightarrow 1/2, 0) = \frac{1}{\sqrt{4\pi}} \begin{bmatrix} T_s + T_p \cos \theta^* & T_p \sin \theta^* e^{-i\varphi} \\ T_p \sin \theta^* e^{i\varphi} & T_s - T_p \cos \theta^* \end{bmatrix} \quad (3)$$

The angular distribution of the decay daughters is given by

$$I = \text{tr}(T \rho_{initial} T^\dagger) = \text{tr}(\rho_{initial} T^\dagger T), \quad (4)$$

thus as a first step the matrix $A = T^\dagger T$ is calculated and the parameter $\alpha = 2\text{Re}(T_s \cdot T_p)$ is introduced [9].

$$A(1/2 \rightarrow 1/2, 0) = \frac{1}{4\pi} \begin{bmatrix} 1 + \alpha \cos \theta^* & \alpha \sin \theta^* e^{-i\varphi} \\ \alpha \sin \theta^* e^{i\varphi} & 1 - \alpha \cos \theta^* \end{bmatrix} \quad (5)$$

The angular distribution can then be calculated using the density matrix derived in [9] and after integrating over the azimuthal angle φ one obtains:

$$I(\theta^*) \propto 1 + \alpha P_H \cos \theta^* \quad (6)$$

where P_H is the polarization, α is the Hypertriton decay parameter and θ^* denotes the angle between the polarization axis and the momentum of the decay daughter, in this case ${}^3\text{He}$, in the Hypertriton rest frame (see Fig. 1).

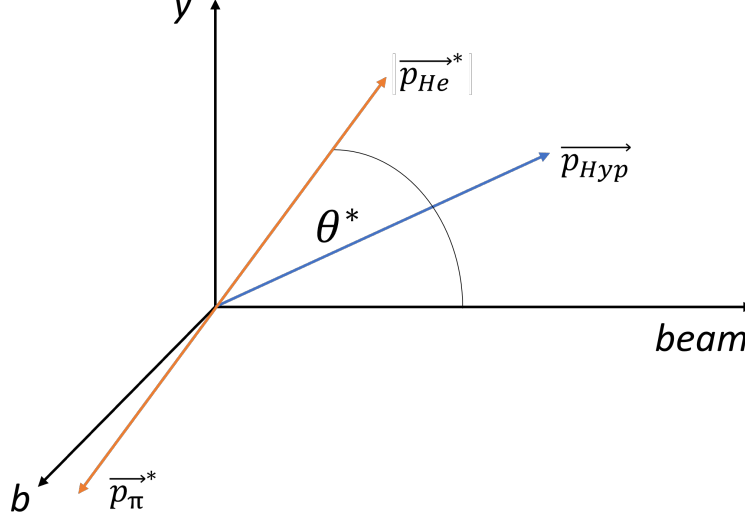


Figure 1: θ^* is defined as the angle between the beam axis and the ${}^3\text{He}$ momentum in the Hypertriton rest frame.

Spin 3/2 Case

Following the same steps as for the spin 1/2 case, one can obtain the angular distribution for spin 3/2. The angular momentum L can have a value of 1 or 2, thus in the final states p- and d-waves need to be considered. Furthermore the spin quantum number for the Hypertriton $m_{\Lambda\text{H}}$ can have values ranging from $-\frac{3}{2}$ to $\frac{3}{2}$, which results in 8 possible final states.

$$\begin{aligned}
T(3/2 \rightarrow 1/2, 0)_{13} &= T_p Y_{11} + \frac{1}{\sqrt{5}} T_d Y_{21} \\
T(3/2 \rightarrow 1/2, 0)_{-13} &= -\frac{2}{\sqrt{5}} T_d Y_{22} \\
T(3/2 \rightarrow 1/2, 0)_{11} &= \sqrt{\frac{2}{3}} T_p Y_{10} + \sqrt{\frac{2}{5}} T_d Y_{20} \\
T(3/2 \rightarrow 1/2, 0)_{-11} &= \frac{1}{\sqrt{3}} T_p Y_{11} - \sqrt{\frac{3}{5}} T_d Y_{21} \\
T(3/2 \rightarrow 1/2, 0)_{1-1} &= \frac{1}{\sqrt{3}} T_p Y_{1-1} + \sqrt{\frac{3}{5}} T_d Y_{2-1} \\
T(3/2 \rightarrow 1/2, 0)_{-1-1} &= \sqrt{\frac{2}{3}} T_p Y_{10} - \sqrt{\frac{2}{5}} T_d Y_{20} \\
T(3/2 \rightarrow 1/2, 0)_{1-3} &= \frac{2}{\sqrt{5}} T_d Y_{2-2} \\
T(3/2 \rightarrow 1/2, 0)_{-1-3} &= T_p Y_{1-1} - \frac{1}{\sqrt{5}} T_d Y_{2-1}
\end{aligned} \tag{7}$$

The following steps are exactly as for the spin 1/2 case, and are described in detail in [9]. In the end the

angular distribution after integrating over φ is given by

$$I(\theta^*) \propto \rho_{33}(1 - \cos^2(\theta^*)) + \rho_{11} \left(\frac{1}{3} + \cos^2(\theta^*) \right) \quad (8)$$

In contrast to the spin 1/2 case, where the angular distribution depends linearly on $\cos(\theta^*)$, the dependency in this case is quadratic.

Obviously these two distributions exhibit very different behaviours, and thus by measuring the angular distribution of the decay daughter ^3He , one can distinguish between spin 1/2 and spin 3/2 and also extract the decay parameter α . In order to obtain this angular distribution we need to measure the Hypertriton polarization. Since Hypertritons are very rare, it is experimentally very challenging to produce them in a laboratory. Therefore we use Heavy Ion collisions to produce polarized Hypertritons and exploit the unique geometry of Heavy Ion collisions to measure it.

3.2 Exploiting Heavy Ion Collisions

The LHC is the only machine world wide that can create Hypertritons in a large sample whilst still having a clean enough signal. In the extreme conditions of the QGP hypernuclei can be formed, and be detected by ALICE. To measure the spin we first need to find $\cos(\theta^*)$.

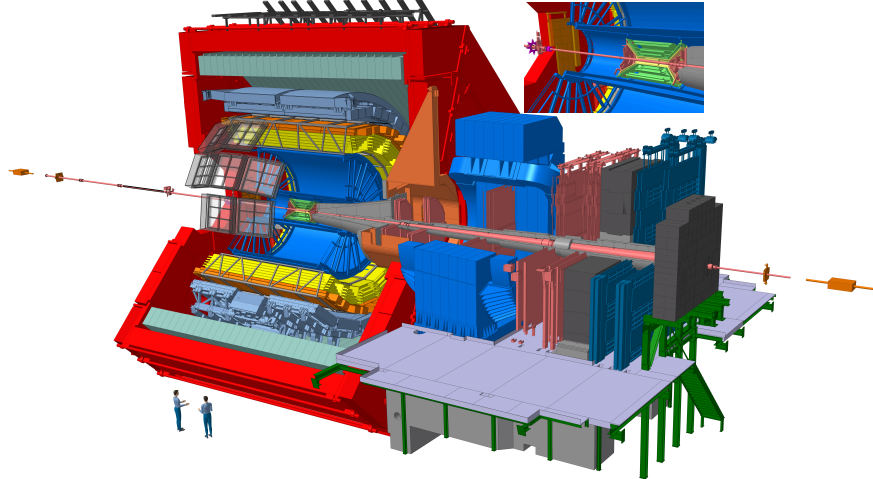


Figure 2: The Alice Detector as it was during Run 2. [11]

The parameter $\cos(\theta^*)$ can be derived from the collisions. In this work only mid-rapidity collision with rapidity $y \in [-0.5, 0.5]$ were selected. In collisions where the hadrons do not collide head-on, it results in an elliptically shaped overlap region. This asymmetry creates a pressure gradient, since the density along the x-axis is higher, than along the y-axis, which results in a velocity gradient when expanding. A velocity gradient will introduce an angular momentum

$$\omega = \frac{1}{2} (\nabla \times v) \quad (9)$$

and since the rapidity is small, this means the particles get mostly emitted in the x-y plane, which means v_z is small, so we can approximate

$$\omega \approx \omega_z = \frac{1}{2} \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) \quad (10)$$

This results in an angular momentum along the beam direction, and because of Spin-Angular Momentum Interaction the spin of the Hypertriton will align to the beam axis. Thus the polarization axis is the beam direction and $\cos(\theta^*)$ describes the angle between the beam axis and the momentum of the Helium daughter in the Hypertriton rest frame.

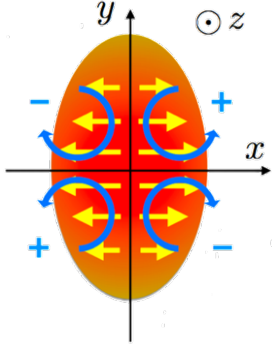


Figure 3: An angular momentum in beam direction is introduced due to the elliptical overlap region. [12]

Looking at Fig. 3 it is clear that due to the geometrical properties of the cross product, the angular momentum will point in different directions (as indicated by the blue arrows and signs). This quadrupole like structure in the transverse plane (x-y plane) will lead to a sine like behaviour of the polarization in $\hat{z}$ direction, which is characterized by the second harmonic sine component in the Fourier decomposition of the polarization [13].

$$P_{z,s2} = \langle P_z \sin(2\varphi - 2\Psi_2) \rangle \quad (11)$$

where φ is the azimuthal angle and Ψ_2 the event plane angle, measured by ALICE's V0 detectors. This behaviour was previously seen in the Λ , $\bar{\Lambda}$ Polarization measured by the STAR Collaboration [14], and is expected to be seen for the Hypertriton, since it behaves similarly due to the small binding energy of the Λ .

We can expect to measure a more significant effect if we split our data sample into two intervals: $|\varphi - \Psi_2| \in [0, \pi/2)$ and $|\varphi - \Psi_2| \in [\pi/2, \pi)$. This will lead to an increased effect by 1/2 of the amplitude, compared with the average value, which will be obtained if we consider the whole data sample.

The decay daughters are measured in the inner part of the ALICE detector in the Time Projection Chamber (TPC) which is mainly used for tracking and identifying particles. The TPC covers the full range of the azimuthal angle φ and up to pseudorapidity $|\eta| < 0.9$. [15]. The energy loss dE/dx is measured in the TPC and using the Bethe-Bloch formula the different kinds of particles can be distinguished (see Fig. 4).

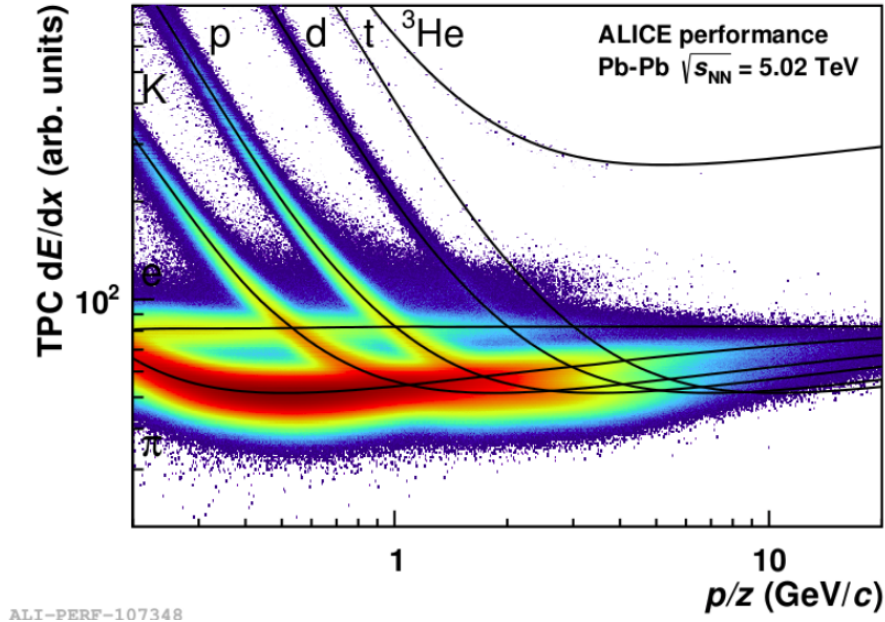


Figure 4: Particle Identification in the ALICE TPC[16]. The Helium(top) and pion(bottom) daughters are clearly visible.

4 Data Analysis

4.1 Choosing the Right Data

4.1.1 Detector Acceptance

The first thing that needs to be considered is what transverse momentum p_T range to choose. Since the ALICE detector does not cover the whole phase space the detector acceptance will not be constant throughout the

whole p_T range from 2GeV to 9GeV. In the Monte Carlo data the Hypertriton candidates are injected without any angular distribution, thus all angular dependencies are due to the detector geometry. Therefore we try to find the best range, where the detector acceptance is nearly constant, so that it doesn't introduce a bias to our data.

To do this the detector acceptance has to be calculated for Monte Carlo data. Firstly the energy of the Hypertriton is obtained by adding the Pion and Helium momenta and using $m_{^3\text{He}} = 2.80839\text{GeV}$ and $m_\pi = 0.139570\text{GeV}$. As energy is conserved the Hypertriton energy is also given by the sum of the daughter energies. The mass of the Hypertriton can then be obtained from the relativistic Energy-Momentum Relation

$$E^2 = m^2 + p^2 \quad (12)$$

using the Hypertriton momentum in the LAB frame, which is simply the added momenta from the decay daughters. Once the energy of the Hypertriton is obtained, the decay daughters can be boosted into the rest frame of the Hypertriton using a simple Lorentz Boost. As established above $\cos\theta^*$ is defined as the cosine of the angle between the boosted Helium and the beam direction, which is represented by the unit vector in z direction in this coordinate frame. For each generated Hypertriton $\cos\theta^*_{w.r.t.beam}$ is calculated using this method. Then the p_T range is split up in 1GeV intervals. For each bin the amount of generated Hypertritons is compared with the number of reconstructed Hypertritons, by plotting the ratio in a $\cos\theta^*$ histogram that ranges from -1 to 1. The goal is then to determine the p_T range where this ratio is relatively constant, as the Hypertritons in the Monte Carlos are inserted with a flat angular distribution and thus all angular dependencies are introduced by the detector geometry. This is a very necessary step, since later the shape angular distribution will be analyzed to determine the spin, and any biases introduced by the detector can influence this distribution and lead to a false result.

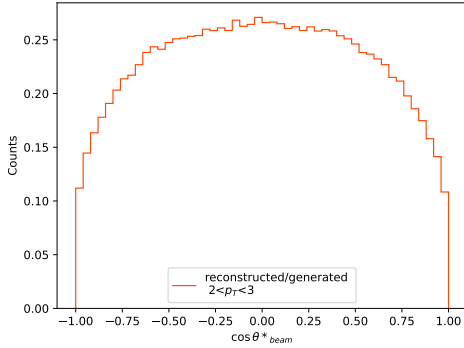


Figure 5: $\cos\theta^*$ dependency for $2\text{GeV} < p_T < 3\text{GeV}$. Clearly there is a strong bias introduced, therefore this interval was excluded from the analysis.

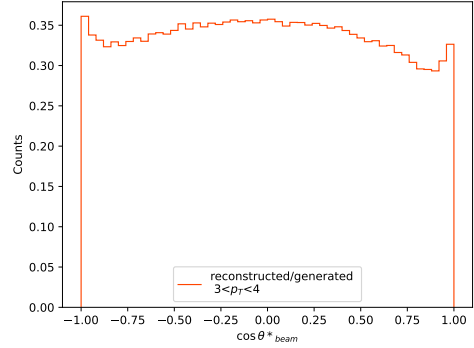


Figure 6: $\cos\theta^*$ dependency for $3\text{GeV} < p_T < 4\text{GeV}$. The ratio is almost constant.

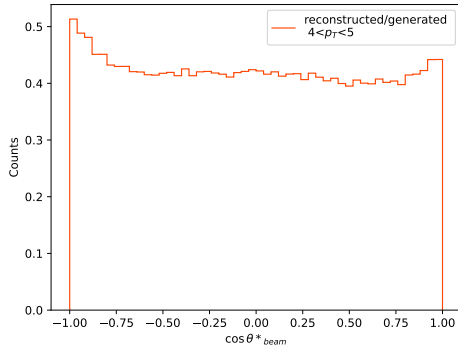


Figure 7: $\cos\theta^*$ dependency for $4\text{GeV} < p_T < 5\text{GeV}$.

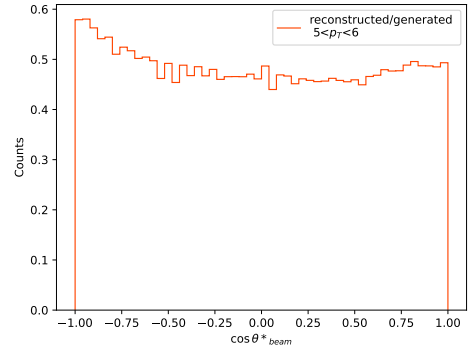


Figure 8: $\cos\theta^*$ dependency for $5\text{GeV} < p_T < 6\text{GeV}$.

It was decided to choose a p_T interval ranging from around 3GeV to 6GeV (see Figures 5 - 8).

4.1.2 Training a Model Using Boosted Decision Trees

To classify the data into background and signal candidates Boosted Decision Trees (BDTs)[17] were used as a machine learning tool. In this work we used a python package called `hipec4ml`[18], which was developed for ion physics as a minimal environment for machine learning. Once the right p_T range was established, the model was then trained on Monte Carlo data and background data. Since in the decay that we are interested in both decay daughters have opposite sign charges, for the background it was decided to use real data where the decay daughters had the same sign, the so called like-sign background. This way is much more precise than taking real data and just cutting out the respective invariant mass region where we expect the Hypertriton peak.

Firstly a testing and training sample is generated from 50% Monte Carlo and 50% background candidates and randomly shuffled. The model is then trained on multiple variables at the same time. These include the cosine of the pointing angle, which is defined as the angle between the vector that connects the primary vertex with the decay vertex and the reconstructed Hypertriton momentum, the Distance of Closest Approach (DCA) with respect to the primary vertex for both the Helium and Pion track, the DCA to the primary vertex in the x-y plane again for both Helium and Pion and the DCA between the two decay daughter tracks. Furthermore the number of clusters in the TPC and the ITS and the $n_{sigma,TPC}$, which is defined as the deviation of the measured signal from that expected for a species in terms of the detector resolution, were included (see Fig. 9).

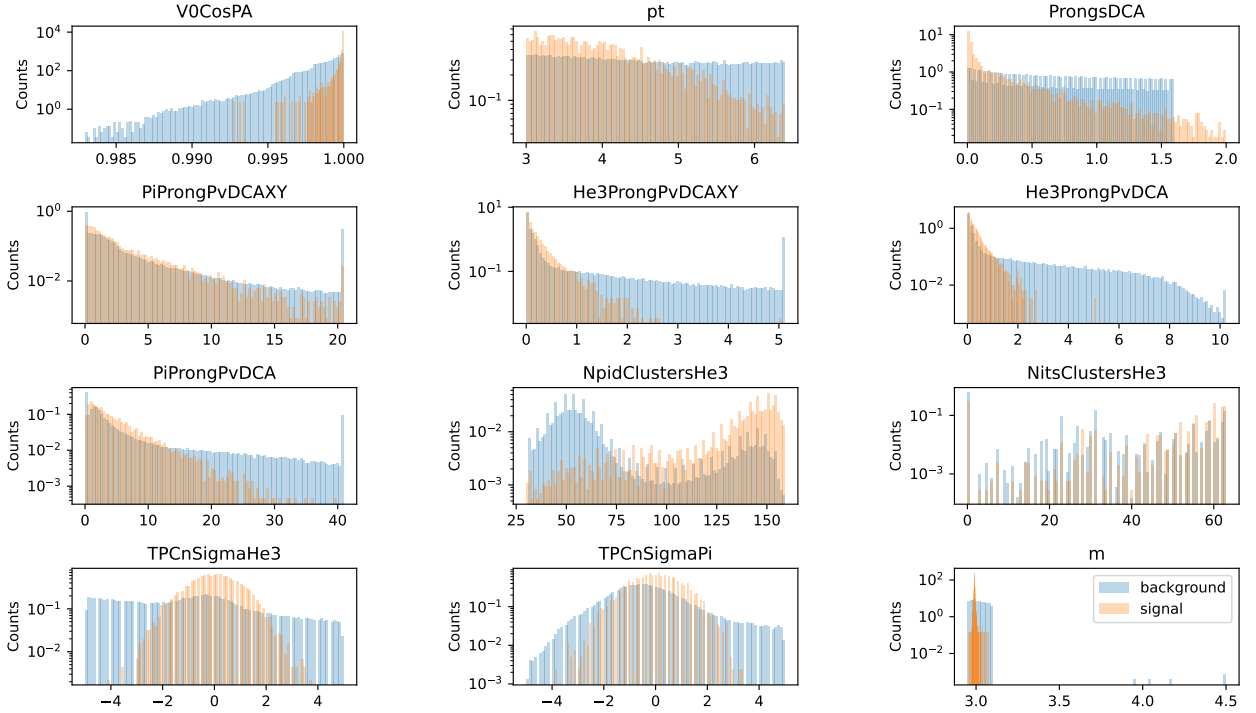


Figure 9: Distribution of the different training variables for the signal and background sample. The invariant mass and p_T are excluded from the training.

Before applying the model to the data it needs to be checked that there aren't any correlations in the training variables to the invariant mass. We do not want to train on the invariant mass, but on the variable features, since we aim to measure the invariant mass distribution as a function of p_T , thus both these features are not included in the training process and we need to remove any variables that show a high correlation. Luckily, as seen in Figs. 10, 11, there is no strong correlation with the invariant mass for the training variables, therefore all of them are included.

The Model is then trained on the training sample using `xgboost` [19]. To further improve the training the hyperparameters of the decision tree are also optimized with the `optuna`[20] package, using the same hyperparameter ranges as in the ALICE Analysis Note [21]. The model is then tested on the testing sample, to check how good the classification worked. The degree of goodness of the model is measured using a so called

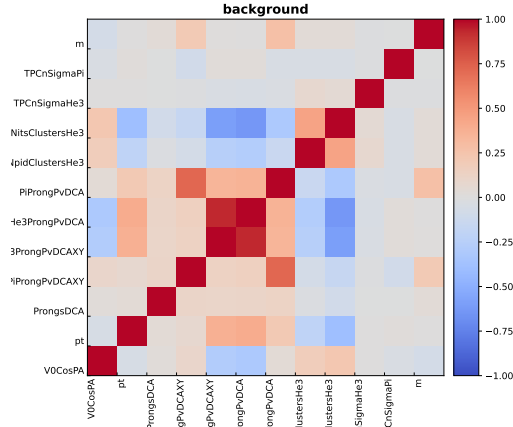


Figure 10: Correlation of the training variables for the background sample

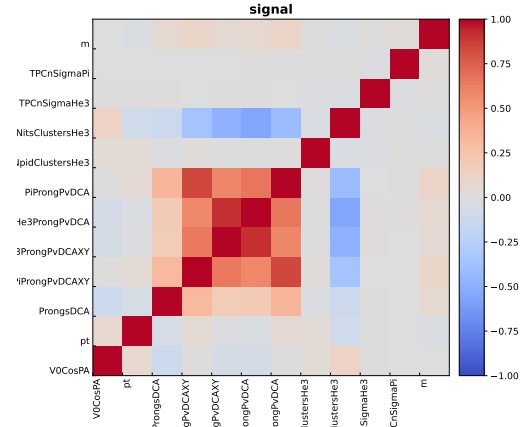


Figure 11: Correlation of the training variables for the signal Monte Carlo sample

ROC AUC plot (see Fig. 12), which correlates the area under the curve with the percentage of candidates that have been correctly classified. A relatively high score is achieved and points towards a correctly trained model, which will then be used for further analysis.

4.1.3 Applying the Model to the Data

After training the model it is applied to the real data. The Run 2 data set that was used consisted of 278971416 collisions, and contained 28405280 Hypertriton candidates. Each candidate gets a BDT score assigned, which tells how likely this candidate is signal or background. To distinguish the signal from the background data a BDT score cut is applied.

In order to find the cut value that gives the best results, the following steps are taken. Firstly the efficiency as a function of the BDT cut needs to be calculated. In this project two different approaches were taken.

The efficiency is calculated by looking at the Monte Carlo. For 100 different BDT Scores the ratio of surviving reconstructed Hypertriton candidates to generated Hypertritons was calculated and plotted as a function of the BDT scores. This approach is very simple, but takes quite a long time. Therefore at a later stage in the project another, less time-consuming method was developed. Here a ROOT histogram of all assigned BDT scores is created and converted into the survival function of its cumulative distribution function. The resulting curve shows the BDT efficiency as a function of BDT score. This way it only has to be calculated once, for all further calculations the efficiency can just be extracted from the histogram (see Fig. 13, 14). Note that the reconstruction efficiency has not been taken into account yet, which will result in an additional factor at a later stage.

The next step is to calculate the significance as a function of the BDT score that is used for the cut. Alternatively the significance can also be calculated as a function of the BDT efficiency, which was done in the second approach. The significance is defined as

$$Significance = \frac{N_{Signal}}{\sqrt{N_{Signal} + N_{Background}}} \quad (13)$$

where N_{Signal} is the number of selected signal candidates and $N_{Background}$ the number of selected background candidates. To obtain the number of background candidates a fit is performed on the Monte Carlo data, which only contains signal. The peak is fitted with a crystal ball function, and the parameters μ and σ , which describe the mean and standard deviation of the peak respectively, are extracted. Then the number of candidates in the background file, for which the invariant mass is in the 3σ region of the Hypertriton invariant mass and with the BDT Score cut applied, is extracted.

To obtain the number of expected signal candidates we multiply the number of total events, which is about 300 million, with the yield, and a factor 2 for considering matter and antimatter and 0.25 for the branching ratio of the decay. This needs to be multiplied with the BDT Efficiency Eff_{BDT} and in the case of the second approach to calculating the efficiency also the reconstruction efficiency Eff_{Reco} . Lastly also the relative

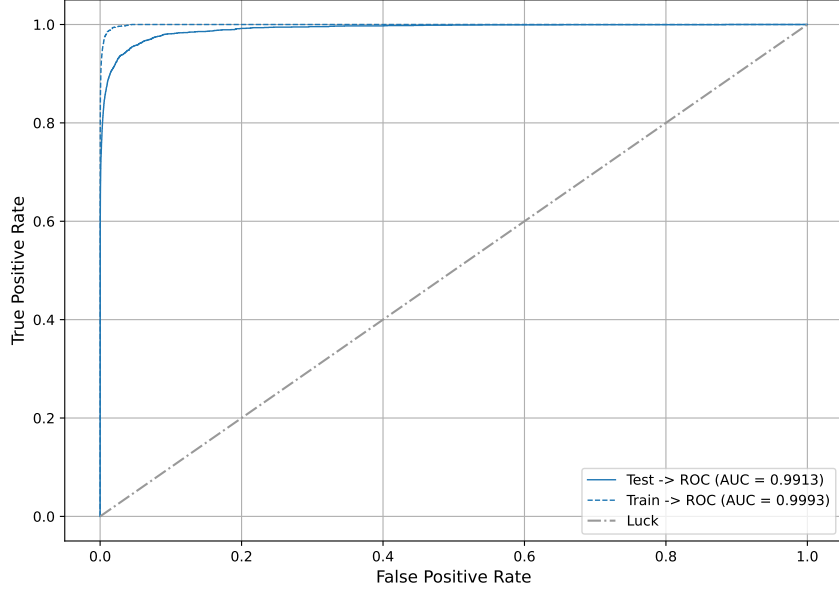


Figure 12: ROC AUC Plot

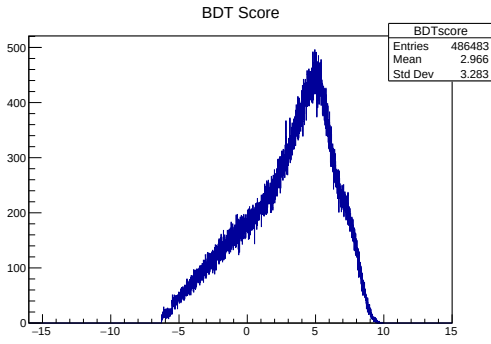


Figure 13: BDT score distribution of the Monte Carlo data

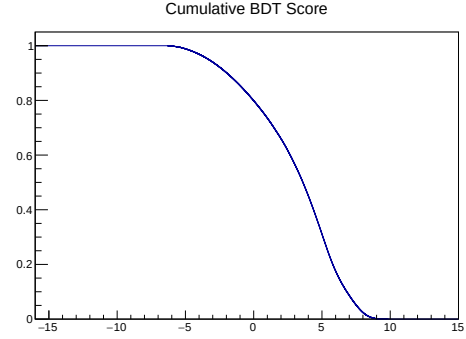


Figure 14: survival function of the cumulative distribution function of the BDT score histogram. This function represents the BDT efficiency.

amount of generated Hypertritons in the desired p_T range compared with the total number of generated Hypertritons n_{Hyp,p_T} needs to be considered, this we end up with this formula for the number of signal candidates:

$$N_{Signal} = (M + AM) \cdot N_{events} \cdot Yield \cdot BR \cdot Eff_{BDT} \cdot Eff_{reco} \cdot n_{Hyp,p_T} \quad (14)$$

$$= 2 \cdot 278971416 \cdot 2.6 \cdot 10^{-5} \cdot 0.25 \cdot Eff_{BDT} \cdot Eff_{reco} \cdot n_{Hyp,p_T}$$

Once the significance and BDT efficiency are both calculated we need to look at their product. Since choosing a high efficiency will result in a higher amount of background events and the significance is very low. On the other hand choosing a high significance, will result in very few events surviving the BDT score cut. Thus it is crucial to find the right interplay between efficiency and significance. Finally the BDT score is chosen, that results in the maximum value for Significance $\cdot$ Efficiency (Fig. 15).

It was then decided to try and optimize the used p_T range, which up to this point had been $p_T \in [3, 6]$ GeV. The interval ranging from 2GeV to 9GeV was considered. The interval size was decreased from 7GeV to 1.6GeV in steps of 200MeV; for each size the whole range was covered in steps of again 200MeV, i.e. the

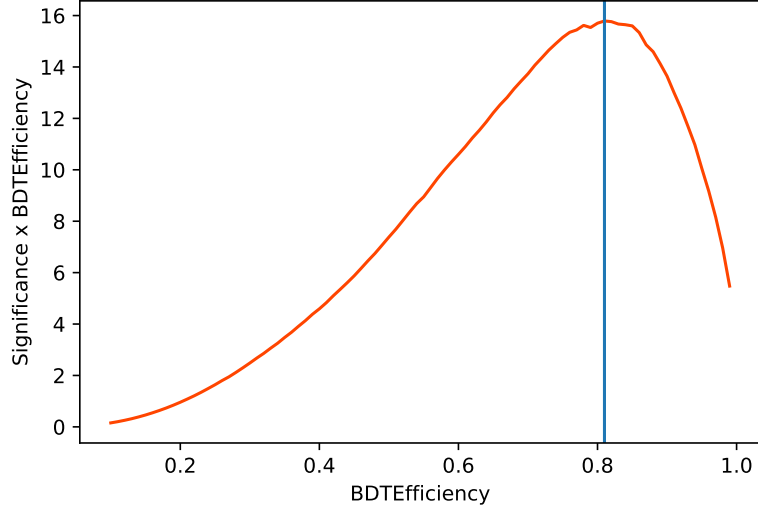


Figure 15: Significance $\cdot Eff_{BDT}$ as a function of Eff_{BDT} . The Maximum is marked by the blue line at a BDTEfficiency of 81%.

first range that was considered ranged from 2-9GeV, the next from 2-8.8GeV and 2.2-9GeV, the next from 2-8.6GeV, 2.2-8.8GeV and 2.4-9GeV and so on. For each p_T range, the optimal BDT score cut value was determined according to section 4.1.3, and the significance corresponding to the chosen BDT cut (or the chosen BDT Efficiency) was calculated. Then the significances for all p_T ranges were compared and the one that resulted in the highest significance was chosen. The selected p_T range was the interval $[3.0, 6.4]$ GeV, which leads to the best significance, while still being in an interval where the detector acceptance $\cdot$ efficiency is flat. This would then decrease the uncertainties on the extracted number of Hypertritons to a minimum, which in turn would result in a more significant effect - if there was one.

The final data will then be cut on

- candidate BDT score $>$ selected BDT score cut
or candidate BDT Efficiency $<$ BDT Efficiency cut
- $p_T \in$ selected p_T range
- $1 < ct < 35$, where the most Hypertritons are expected.

Once all these cuts are applied, a clear peak is visible in the invariant mass spectrum, which resembles the Hypertriton signal.

4.2 Extracting the Hypertriton Spin

4.2.1 Fitting the Invariant Mass Spectrum

Since we want to investigate the angular distribution, in the next step $\cos \theta^*$ was calculated for the surviving candidates. The whole range from -1 to 1 was then divided into equidistant bins, here it was chosen to use 7 bins in total. For each bin, we can then look at the invariant mass distribution of the selected candidates.

Once this invariant mass spectrum for the data is obtained, it needs to be fitted with a suitable function. The cuts that are applied do not filter out all the background, so this needs to be taken care of now, by using different functions for the background and the signal component of the spectrum. Here the background was fitted using an exponential and the signal was fitted using a crystal ball function. A Gaussian fit was tried before, but the crystal ball provides much better results. Since the crystal ball shape has very versatile tails it can actually approximate the background as well, and when fitting this is what will happen. This needs to

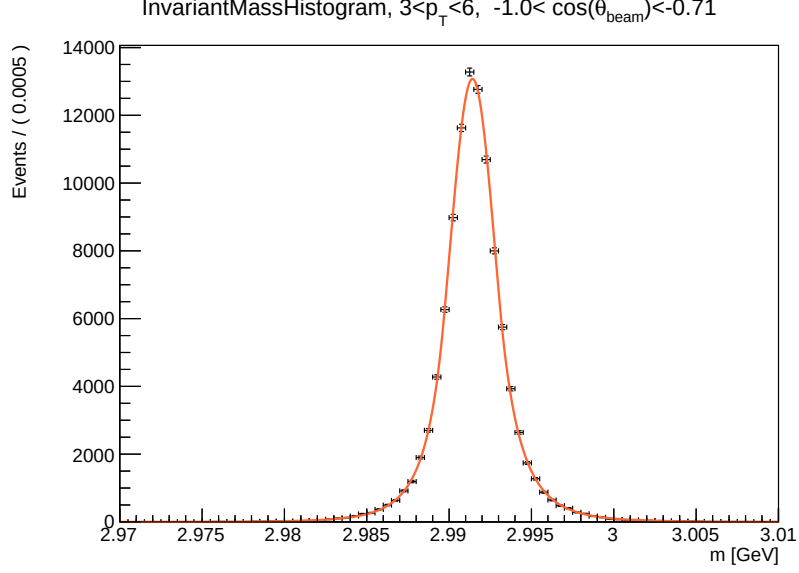


Figure 16: Fitting the Monte Carlo data set in the first $\cos \theta^*$ bin with a crystal ball function to fix the tail parameters before fitting the whole invariant mass spectrum.

be avoided, therefore the parameters that specify the shape of the tails were fixed by firstly fitting only the crystal ball shape to the Monte Carlo data set, which again just contains signal events and no background, and then extracting the fit parameters. The crystal ball function is defined as

$$f_{CB}(m; m_0, \sigma, \alpha_L, \alpha_R, n_L, n_R) = \begin{cases} A_L \cdot \left(B_L - \frac{m-m_0}{\sigma_L} \right)^{-n_L} & \text{for } \frac{m-m_0}{\sigma_L} < -\alpha_L \\ \exp \left(-\frac{1}{2} \left(\frac{m-m_0}{\sigma_L} \right)^2 \right) & \text{for } \frac{m-m_0}{\sigma_L} \leq 0 \\ \exp \left(-\frac{1}{2} \left(\frac{m-m_0}{\sigma_R} \right)^2 \right) & \text{for } \frac{m-m_0}{\sigma_R} \leq \alpha_R \\ A_R \cdot \left(B_R + \frac{m-m_0}{\sigma_R} \right)^{-n_L} & \text{otherwise,} \end{cases}$$

with

$$\begin{aligned} A_i &= \left(\frac{n_i}{|\alpha_i|} \right)^{n_i} \cdot \exp \left(-\frac{|\alpha_i|^2}{2} \right) \\ B_i &= \frac{n_i}{|\alpha_i|} - |\alpha_i| \end{aligned} \quad (15)$$

where $\alpha_L, \alpha_R, n_L, n_R$ are extracted from the MC fit and σ, m_0 are left as free fitting parameters. With these the spectrum is then fitted with

$$f(m; a, m_0, \sigma, \lambda) = a \cdot f_{CB}(m; m_0, \sigma) + (1-a) \cdot \exp(\lambda \cdot m) \quad (16)$$

using RooFit.

4.2.2 Extracting the Angular Distribution

The number of Hypertriton candidates is then extracted using

$$N_{Hyp} = a \cdot N_{Candidates} \quad (17)$$

where $N_{Candidates}$ is the total number of candidates in the invariant mass spectrum. This number then needs to be corrected with the detector acceptance in this $\cos \theta^*$ bin, that was calculated in the beginning (see Section 4.1.1). Once the number of Hypertritons is divided by the detector acceptance, the number is plotted for each bin, which then gives the angular distribution of Hypertritons (see Fig. 18).

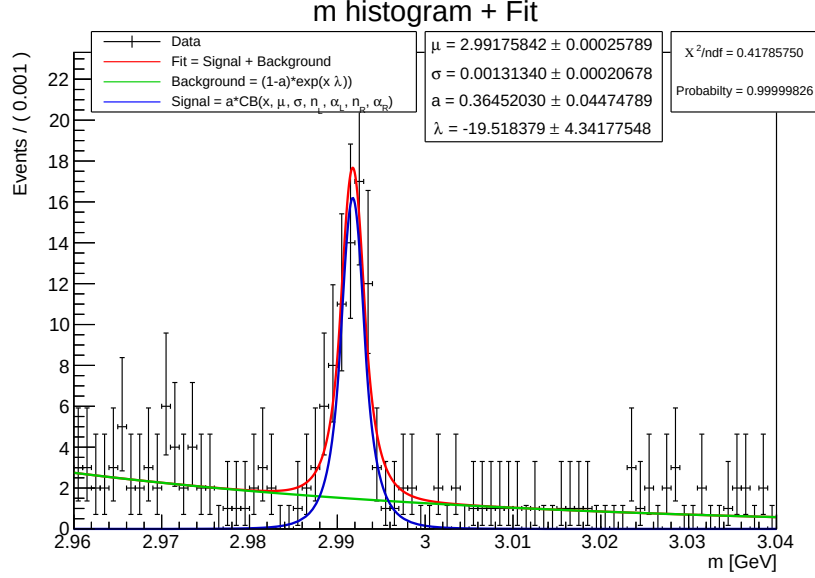


Figure 17: Using the fit parameters determined from fitting the MC data, the invariant mass spectrum is fitted.

The angular distribution depends on the parameter α as discussed in section 3.1. This parameter is expected to have a different sign for matter and antimatter, therefore it is important to split the data sets into matter and antimatter, so ${}^3_{\Lambda}\text{H}$ and ${}^3_{\Lambda}\bar{\text{H}}$ respectively.

A simultaneous fit was then performed on both the matter and antimatter angular distribution for a first and second order polynomial, which would correspond to spin 1/2 and spin 3/2. It was then decided to instead look at the difference between matter and antimatter. The effect would be increased by a factor of 2 and the uncertainties only by a factor of $\sqrt{2}$, thus the overall effect would result in a more significant result and also give smaller systematic errors.

The whole analysis was repeated for the newly chosen p_T range. The cuts that were applied are

- $p_T \in [3.0, 6.4]\text{GeV}$
- $1 < ct < 35$
- $Eff_{BDT} < 0.81$.

A F-Test, which is used to determine whether a more complex model, i.e. with more degrees of freedom (in this case the polynomial) fits the data set significantly better than a simpler model, which would be the case of a constant angular distribution, i.e. no polarization, showed that actually pol2 was preferred, which would correspond to a spin of 3/2, but both fits, pol1 and pol2, resulted in a less than 1σ effect, so it could not be decided, which distribution fit better.

4.2.3 Selecting on Event Plane Angle and Centrality

To further improve the results two more cuts were applied to the data. The first one was a cut on the Event Plane.

In the Pb-Pb collisions the nuclei do not necessarily collide head on. The plane that is spanned by the impact parameter and the beam direction ($\hat{z}$) is called reaction plane. This plane is not known, but can be reconstructed by the V0 detectors that are on the end caps of the ITS. On the ATLAS side there is V0A, while V0C is located on the CMS side of the LHC. Both these detectors measure the so called event plane angle, which describes the orientation of the reconstructed reaction plane, the so called event plane.

In principle when considering the event plane angle Ψ_2 one would consider the average of the two angles measured by the V0 detectors. When looking at the distribution of the difference ($\Psi_{V0A} - \Psi_{V0C}$) it is clear

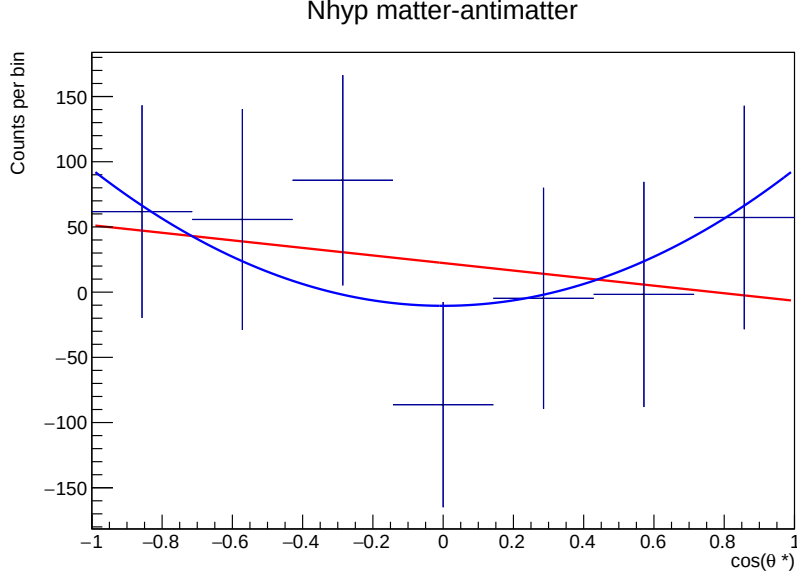


Figure 18: Difference in the number of measured Hypertritons for matter and antimatter in the 3.0GeV to 6.4GeV p_T range. Two fits were tried: a linear fit (red) and a quadratic one (blue). For both fits the effect was less significant than 1σ .

that the tails are quite pronounced. Looking at the correlation between the difference and the centrality, it becomes evident, that the tails are mainly for low centrality events.

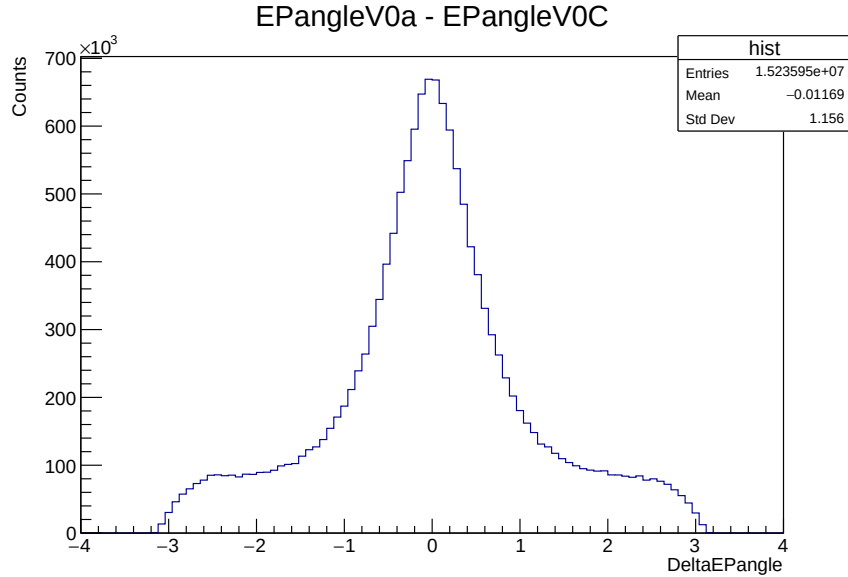


Figure 19: Histogram of the difference between the two measured event plane angles.

In the final analysis these candidates would ideally not be considered, since the nuclei collide almost head-on in these events and the polarization effect will be less. Unfortunately we are right now very limited in statistics, therefore it was decided to only take $\Psi_2 = \Psi_{V0A}$ and not cut on centrality.

The difference between $\Psi_2 \in [0, \pi)$ (Fig. 22) and $\varphi \in [0, 2\pi)$ (Fig. 21) was calculated and folded into the interval $[0, \pi)$ (Fig. 23). Then the dataset was split up into two intervals $|\varphi - \Psi_2| \in [0, \pi/2)$ and $[\pi/2, \pi)$.

Since the number of candidates decreased, 5 bins instead of 7 were used and the σ parameter of the signal fit was fixed to the value extracted from the Monte Carlo fit, to ensure a successful fit of the invariant mass spectrum. [14]. We were able to observe different slopes for the two intervals, that were in the same order

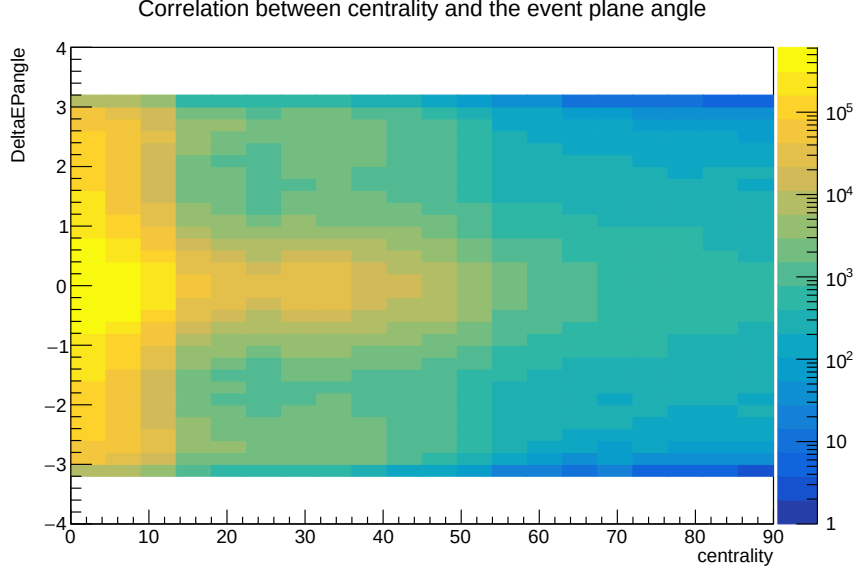


Figure 20: Correlation between the centrality and the difference in event plane angles. It was decided to cut away candidates with a centrality of 0-10%.

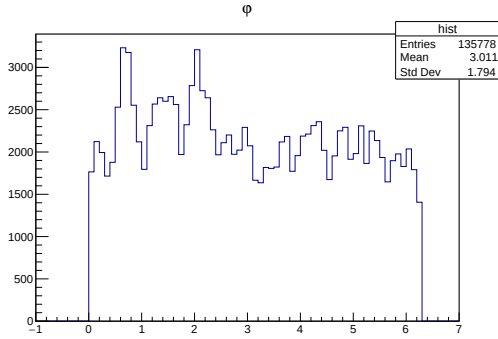


Figure 21: Distribution of the azimuthal angle in the data set.

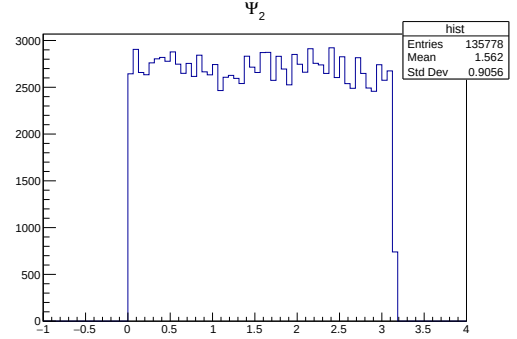


Figure 22: Distribution of the event plane angle measured by V0A in the data set.

of magnitude, and were leading to a polarization effect that was in the same order of magnitude as the one measured for Λ polarization [14]. A simultaneous fit was successfully done on the two intervals (see Figs. 24, 25) which lead to a slope of 19.59 ± 17.41 . Unfortunately the effect was not notable enough, to conclude anything.

4.3 Systematic Uncertainties

As a last step of the analysis the systematic uncertainties were calculated. By choosing a certain BDT score to cut and deciding on the fit model (see Eqn.16) a bias was introduced into the data, which needs to be adjusted by adding the systematic uncertainties.

The way these are calculated is by variation of the cut and the model. The BDT score was varied by $\pm 10\%$, thus in total 21 different BDT cuts were considered. For each of these the analysis chain was done again. Instead of just using the model defined in 16, the model was varied as well: the parameter σ of the crystal ball function was once fixed to the one determined by fitting the Monte Carlos, once left as a free parameter for the fit. Furthermore the function that is approximating the background was varied to be exponential, a first and a second order polynomial. For each cut all six models were fitted and the extracted number of Hypertritons per $\cos(\theta^*)$ bin stored. Then a random sample of size 10.000 was drawn from each bin and the matter-antimatter slope was fitted for each combination. All the slopes were then stored in a histogram,

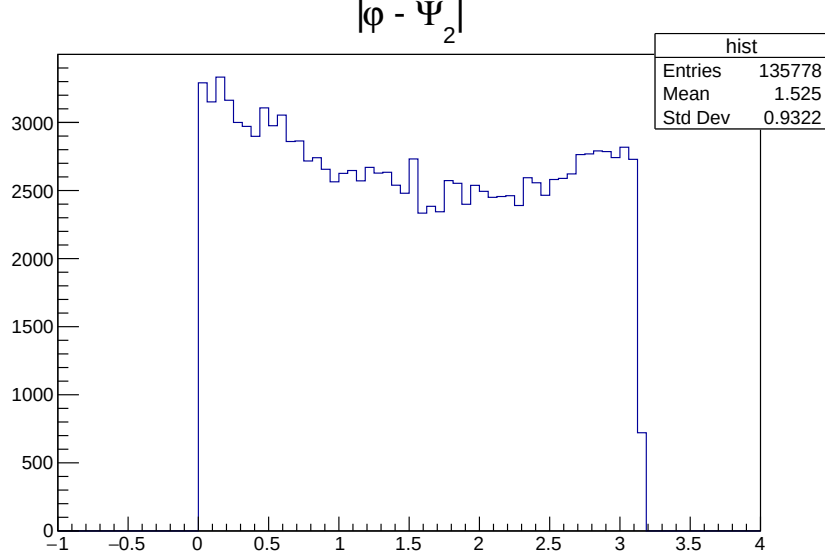


Figure 23: Difference between φ and Ψ_2 folded to the interval $[0, \pi)$

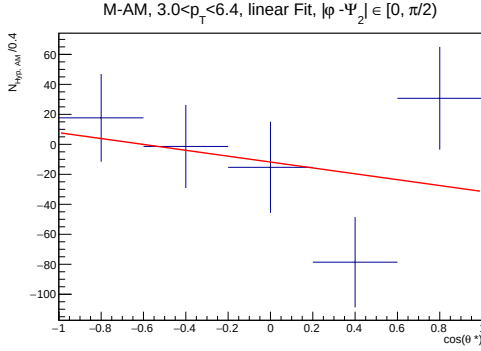


Figure 24: Matter-Antimatter difference observed in the first interval. The slope is negative, but since we do not know the decay parameter this doesn't contradict the observations made for the Λ [14].

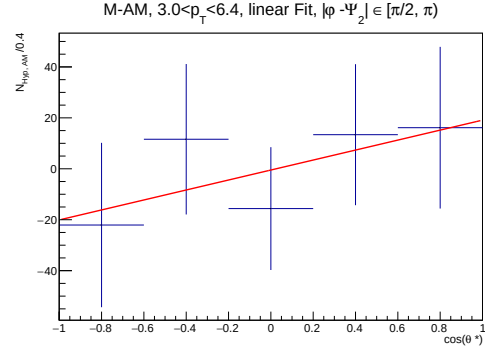


Figure 25: Matter-Antimatter difference observed in the second interval. The linear fit was done simultaneously for both intervals.

which was fitted with a Gaussian distribution. The standard deviation of this Gaussian then resembles the systematic uncertainties. Further sources of systematic errors were investigated.

Firstly the Monte Carlos do not contain infinite candidates, thus fitting them will result in a small but non-zero uncertainty of the fit parameters, that are used in the analysis of the data. To calculate the systematic uncertainties introduced by this the parameters value and uncertainty was stored. A random value was drawn from a Gaussian distribution with μ equal to the parameter value and σ equal to its uncertainty, and then used for the further analysis steps. The systematic errors were calculated again, the effect was negligible.

Another test was conducted, to check how strongly the real signal data deviated from the Monte Carlo data, since the MC does not represent reality to 100%. Since the Crystal Ball shape will fit the background as well, only two of the four fixed parameters can be set to free fitting parameters at the same time. It was done for fixed n_L, n_R and free α_L, α_R and vice versa. The systematic errors were calculated again using the same procedure as described above, and the effect was again negligible.

Lastly the fact that the MCs were fitted with a crystal ball, also introduces a small bias. To examine the systematic uncertainties introduced by this a RooKeysPdf Kernel Density Estimation (KDE) fit was applied to the Monte Carlo Data instead of the Crystal Ball fit. The KDE will fit the shape of the MC data distribution with a linear combination of multiple Gaussian distributions. Then the KDE is used in the analysis instead of the Crystal Ball. Since the invariant mass of the Hypertriton, which corresponds to the

mean of the peak, differs in the MCs and the real data, a free shift parameter was left to shift the KDE along the x-axis of the spectrum. The systematic uncertainties were calculated for this modification as well, and no significant change was found.

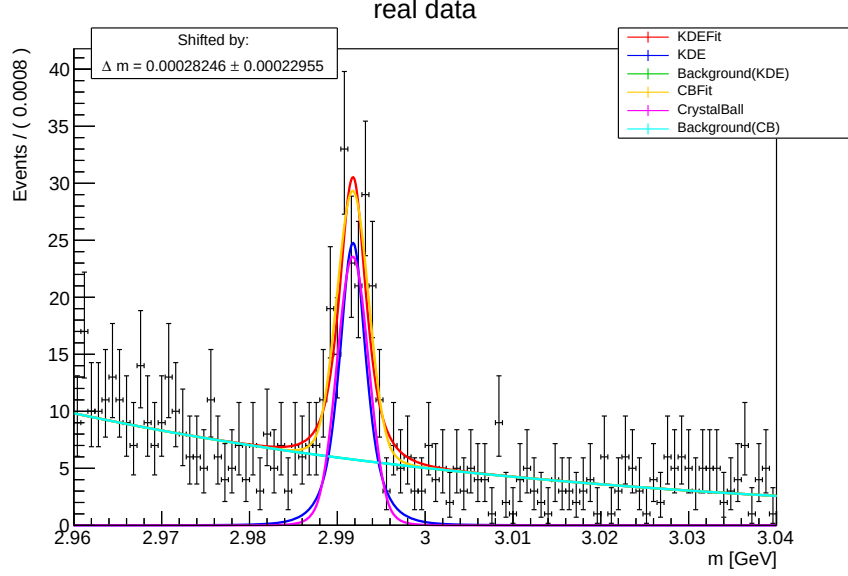


Figure 26: Comparing the fit of a real data set using the crystal ball fit versus using the KDE; the background was approximated by an exponential for both cases. The shift parameter Δm was extracted from the KDE fit. No significant change in the extracted number of Hypertritons was found.

These tests conclude the analysis steps that were taken. Since the statistical uncertainties alone were very high for the results, it was decided to not calculate the systematic uncertainties.

5 Outlook

In the end unfortunately the Run2 Data was not enough to find a significant polarization, therefore the spin of the Hypertriton could not be determined. Still the scripts and algorithms for the analysis chain were developed and can be used to analyze Run3 data, which will contain more Hypertritons. Furthermore with the described methods we were able to successfully extract the Hypertriton invariant mass signal for each $\cos(\theta^*)$ bin, matter and antimatter, and each event plane angle region. With the Pb-Pb data that will be taken in 2023 alone, the significance will be increased by a factor of 6. Hopefully then we will be able to find a significant effect and solve the question about the Hypertriton spin.

Acronyms

ALICE A Large Ion Collider Experiment. 3, 6, 7, 9, 20

ATLAS A Toroidal LHC Apparatus. 14, 20

BDT Boosted Decision Tree. 10–12, 16

BDTs Boosted Decision Trees. 9

CERN European Organization for Nuclear Research. 3

CMS Compact Muon Solenoid. 14

DCA Distance of Closest Approach. 9

ITS Inner Tracking System. 9, 14

KDE Kernel Density Estimation. 17, 18

LHC Large Hadron Collider. 3, 4, 6, 14

MC Monte Carlo. 13, 14, 17, 18

QGP Quark Gluon Plasma. 3, 6

STAR Solenoidal Tracker at RHIC. 7, 20

TPC Time Projection Chamber. 7, 9

References

- [1] ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. 2012. URL: <https://arxiv.org/abs/1207.7214>.
- [2] Official ALICE Webpage. Accessed 11th September 2023. URL: <https://www.home.cern/science/experiments/alice>.
- [3] ALICE Collaboration. “Measurement of the lifetime and Λ separation energy of ${}^3_\Lambda H$ ”. 2022. URL: <https://arxiv.org/abs/2209.07360>.
- [4] ALICE Collaboration. “Hypertriton production in p-Pb collisions at $\sqrt{s_{NN}} = 5.02\text{TeV}$ ”. 2021. URL: <https://arxiv.org/abs/2107.10627>.
- [5] U. Heinz R. Scheibl. “Coalescence and flow in ultra-relativistic heavy ion collision”s. 1999. URL: <https://arxiv.org/abs/nuc1-th/9809092>.
- [6] H. Stoecker V. Vovchenko B. Donigus. “Multiplicity dependence of light nuclei production at LHC energies in the canonical statistical model”. 2018. URL: <http://arxiv.org/abs/1808.05245>.
- [7] L. Fabbietti L. Tolos. “Strangeness in Nuclei and Neutron Stars”. 2020. URL: <https://arxiv.org/abs/2002.09223>.
- [8] N. Fornengo M. Korsmeier F. Donato. “Prospects to verify a possible dark matter hint in cosmic antiprotons with antideuterons and antihelium”. 2018. URL: <https://arxiv.org/abs/1711.08465>.
- [9] E. Thomé. “Multi-Strange and Charmed Antihyperon Hyperon Physics for PANDA”. PhD thesis. Uppsala University, 2012.
- [10] “Definition of the ALICE Coordinate System and Basic Rules for Sub-detector Components Numbering”. URL: https://alice-servotech.web.cern.ch/HELP_DCDB-SVT/Help_Files/ALICE-INT-2003-038.pdf.
- [11] CERN Document Server. Accessed 25th September 2023. URL: <https://cds.cern.ch/record/2263642>.
- [12] D. Sarkar. “Spin polarization measurements in heavy-ion collisions”. 2022. URL: <https://indico.cern.ch/event/895086/contributions/4615144/>.
- [13] S. Acharya *et al.* “Polarization of Λ and $\bar{\Lambda}$ Hyperons along the Beam Direction in Pb-Pb Collisions at $\sqrt{s_{NN}} = 5.02\text{TeV}$ ”. 2022. URL: <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.128.172005>.
- [14] STAR Collaboration. “Polarization of $\Lambda(\bar{\Lambda})$ Hyperons along the Beam Direction in Au+Au Collisions at $\sqrt{s_{NN}} = 200\text{GeV}$ ”. 2019. URL: <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.123.132301>.
- [15] J. Alme *et al.* “The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events”. 2010. URL: <https://arxiv.org/abs/1001.1950>.
- [16] ALICE Analysis Tutorial. Accessed 25th September 2023. URL: <https://alice-doc.github.io/alice-analysis-tutorial/analysis/pid.html>.
- [17] Y. Coadou. “Boosted decision trees”. 2022. URL: <https://arxiv.org/abs/2206.09645>.
- [18] hipe4ml python package documentation. Accessed September 26th 2023. URL: <https://hipe4ml.github.io/hipe4ml/index.html>.
- [19] XGBoost Python Package Documentation. Accessed September 26th 2023. URL: <https://xgboost.readthedocs.io/en/stable/python/index.html>.
- [20] Optuna Documentation. “Optuna: A hyperparameter optimization framework Documentation”. URL: <https://optuna.readthedocs.io/en/stable/index.html>.
- [21] G. Alocco *et al.* “Hypertriton lifetime and Λ -separation energy measurement in the 2-body decay channel with a Machine Learning approach in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02\text{TeV}$ ”. 2021. URL: https://alice-notes.web.cern.ch/system/files/notes/analysis/999/2021-12-11-hypertriton_ML_2018PbPb5.02TeV_V4.pdf.