

**Copyright
by
Ryan Patrick Hannigan
2023**

The Dissertation Committee for Ryan Patrick Hannigan certifies that
this is the approved version of the following dissertation:

Stranger Things at the LHC

Committee:

Christina Markert, Supervisor

Peter Onyisi

Richard Hazeltine

Chandrajit Bajaj

Stranger Things at the LHC

by

Ryan Patrick Hannigan, B.S.

Dissertation

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the degree of
Doctor of Philosophy

The University of Texas at Austin
May, 2023

To Jaynee, who unquestionably is the reason why this document exists.

Stranger Things at the LHC

Publication No. _____

Ryan Patrick Hannigan, Ph.D.
The University of Texas at Austin, 2023

Supervisor: Christina Markert

Among the most mysterious of the four fundamental forces is the strong nuclear force. Responsible for both the binding together of nucleons within an atom, as well as the

Contents

List of Figures

List of Tables

Chapter One: Introduction

This initial section of this chapter will mostly serve as a historical overview of the field of particle physics, leading to the development of the **Standard Model**—the theory that describes all of the fundamental¹ particles and the way in which they interact with each other. An emphasis will be made on the discovery of quarks, as the research presented in this thesis is centered around these particles. This first section will contain little-to-no mathematics, as historical lessons generally do not.

The second section will more thoroughly introduce the Standard Model, and how it was developed from a theoretical perspective. Words like “fundamental representation” and “gauge symmetry” will be thrown around, but the goal is to provide a high-level mathematical overview of the theory and how it was developed without getting bogged down in the details.

As this thesis is focused nearly entirely on the strong nuclear force, the remaining sections of the chapter will provide a thorough discussion of Quantum ChromoDynamics (QCD), which is the component of the Standard Model which describes the interactions between **quarks** and **gluons**—the constituent particles of the more familiar protons and neutrons. Unfortunately, QCD is enormously complicated and the full theory is not yet fully understood. While this may be disheartening for theorists, it is a boon for experimentalists as it provides a wealth of opportunities to probe the theory in regimes where it is both understood and not understood.

To this end, the remainder of this chapter will focus on the ways in which QCD can be investigated using heavy-ion collisions, with an emphasis on the **Quark-Gluon Plasma** (QGP)—a novel state of nuclear matter that QCD predicts should exist at the extreme temperatures and densities that are achieved in these collisions. The experimental signatures of QGP formation will also be discussed, with a particular focus on **strangeness enhancement**—the phenomenon where the production of strange quarks is enhanced in heavy-ion collisions relative to proton-proton collisions.

¹As of the year 2023, but reading this chapter will hopefully illustrate why this may be subject to change in the (likely very distant) future.

1.1 What is fundamental?

The answer to the question “What are the fundamental building blocks of our universe?” has changed drastically over the course of human history. The idea that all matter is composed of smaller, uncuttable pieces has been around since 5th century BCE when Greek philosophers Democritus and Leucippus first introduced the concept of an atom **GreekAtom**. While this idea was mostly motivated by philosophical reasoning, it was later adopted by the English scientist John Dalton in the 19th century to explain the results of his chemical experiments, where he found that chemical elements always combined with each other by discrete units of mass **Dalton**. As scientists discovered more and more of these elements, the number of “fundamental” building blocks grew as well. By the late 1800s, over 70 unique chemical elements had been discovered, though they would often be grouped together due to similar chemical properties using what chemist Dimitri Mendeleev dubbed the *periodic table of elements* **PeriodicTable**. An example of the periodic table from the time of Mendeleev can be seen in Figure ???. While this grouping was useful for chemists, it also served as a hint to physicists that perhaps these elements were not actually fundamental, but rather composed of even smaller pieces.

		Ti = 50	Zr = 90	? = 180
		V = 51	Nb = 94	Ta = 182
		Cr = 52	Mo = 96	W = 186
		Mn = 55	Rh = 104,4	Pt = 197,4
		Fe = 56	Ru = 104,4	Ir = 198
H = 1		Ni = 59	Pd = 106,6	Os = 199
		Cu = 63,4	Ag = 108	Hg = 200
	Be = 9,4	Mg = 24	Zn = 65,2	Cd = 112
	B = 11	Al = 27,4	? = 68	Ur = 116
	C = 12	Si = 28	? = 70	Sn = 118
	N = 14	P = 31	As = 75	Sb = 122
	O = 16	S = 32	Se = 79,4	Te = 128?
	F = 19	Cl = 35,5	Br = 80	J = 127
Li = 7	Na = 23	K = 39	Rb = 85,4	Ca = 133
		Ca = 40	Sr = 87,6	Tl = 204
		? = 45	Ce = 92	Pb = 207
		?Er = 56	La = 94	
		?Yt = 60	Di = 95	
		?In = 75,6	Th = 118?	

Figure 1.1: Dimitri Mendeleev’s periodic table of elements from the late 1800s, taken from [MendeleevPaper](#). The elements are grouped by similar chemical properties, and the gaps in the table are where Mendeleev predicted that new elements would be discovered.

Things changed quite a bit around the turn of the 20th century, with scientists

like Rutherford and Chadwick determining that the supposedly indivisible atom was composed of even smaller sub-atomic particles, eventually named electrons, protons and neutrons **Electrons, Protons, Neutrons**. Thus the number of fundamental blocks of matter had decreased substantially from nearly 100 to just three, but only very briefly. Only months after the discovery of the neutron, the fundamental anti-particle of the electron—known as the positron—was discovered in 1932 by Carl Anderson **Positron**. In the next two decades, the number of known fundamental particles would skyrocket. In 1947, the muon was discovered **Muon**, followed by the discovery of a laundry list of particles that participate in the same interaction that holds oppositely charged protons together in the nucleus of an atom—the so-called **strong nuclear force**. These “fundamental” particles were collectively called **hadrons**, which were further separated into lighter and heavier categories dubbed **mesons** and **baryons**, respectively **MesonBaryon**. By the late 1960s, the number of known hadrons had grown to well over 100 **ParticleDiscoveries**, which is ironically much higher than the number of “fundamental” chemical elements that were known to exist in the 1800s.

In the same way that Mendeleev tried to group the elements by their similar chemical properties, physicists attempted to group the hadrons together based on their known sub-atomic properties at the time. The first successful attempt at such a grouping was the **Eightfold Way**, which was independently proposed by Murray Gell-Mann and Yuval Ne’eman in 1961 **GellMann, Neeman**. This grouping was found by examining the following properties of the hadrons:

1. **Isotopic spin:** a quantum number introduced by Werner Heisenberg in 1932 to try to explain the apparent symmetries between the proton and neutron with respect to the strong nuclear force **IsotopicSpin** (i.e. although the proton and neutron have different electric charges, the strong interaction does not seem to distinguish between the two)
2. **Strangeness:** another quantum number introduced by Gell-Mann and Nishijima in 1953 to explain why some hadrons decayed much more slowly than expected, but such particles appeared to be created in pairs **Strangeness**. In other words, the strong interaction responsible for the creation of these particles appeared to conserve strangeness, but the weak interaction responsible for

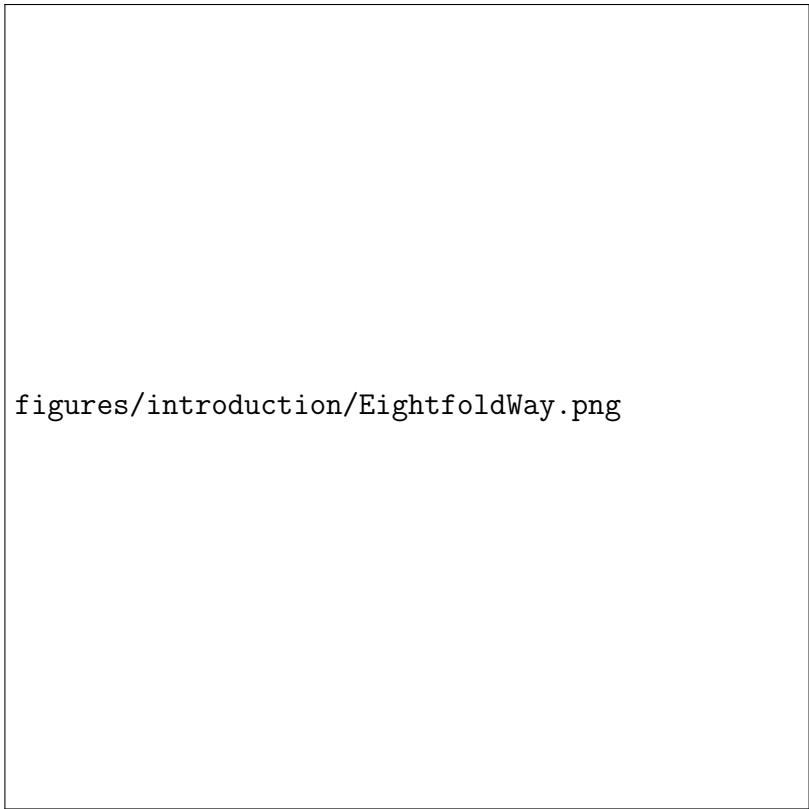
the slower decay of these particles did not. This² quantity is of utmost importance to this thesis, and will be discussed in much greater detail in the coming sections.

Plotting the baryons and mesons in a two-dimensional space based on these two properties revealed striking patterns, as shown in Figure ???. Similar to Mendeleev, GellMann also left a blank space³ where he believed a new particle—the Ω^- —would be discovered. The patterns in these diagrams hinted at an underlying symmetry governing the strong nuclear force, and ultimately led to the invention of the very first quark model by Gell-Mann and Zweig in 1964 **QuarkModel**. This model proposed that all of the hadrons were actually composed of even smaller particles, which Gell-Mann dubbed “quarks”. The quark model was able to explain the patterns seen in Figure ?? by introducing three different types of fermionic quarks—up, down and strange—along with their corresponding anti-quarks. Baryons would then be composed of three such quarks, whereas mesons would be composed of quark and anti-quark pairs. If the quark model were correct, the number of fundamental building blocks of matter would again decrease from over 100 to just 14: electrons, muons, electron neutrinos, muon neutrinos, up quarks, down quarks, strange quarks, and all of their corresponding anti-particles.

Initially, many physicists believed that the quarks from this model were just a mathematical abstraction **QuarkAbstraction**. This possibility did not stop Sheldon Glashow and James Bjorken from extending the quark model in less than a year after its inception by introducing a fourth quark: the charm **CharmQuark**. This new quark was primarily introduced to equalize the number of leptons (four at the time: electron, muon, and their respective neutrinos) with the number of quarks. The theory was mostly aesthetic **AestheticCharm** as the charm quark was not explicitly required by any known mechanisms. It was only after the Glashow-Iliopoulos-Maiani (GIM) mechanism was introduced in 1970 **GIM** that the existence of the charm quark became “necessary”. This mechanism helped explain why neutral kaons decayed into

²Strangeness was introduced a few years before the very first quark model, but it now has the modern interpretation which is directly related to the number of strange and anti-strange quarks within a hadron.

³The original paper on the Eightfold Way does not contain any of these diagrams, but there are discussions about the properties of particles that should exist if the theory were correct, but had not been observed.



figures/introduction/EightfoldWay.png

Figure 1.2: The Eightfold Way, which was the first successful attempt at grouping the known hadrons based on their properties. The baryons are shown in the top row, and the mesons are shown in the bottom row. The isotopic spin is shown on the horizontal axis, and the strangeness is shown on the vertical axis. The baryons are grouped by their spin, and the mesons are grouped by their parity.

two muons at a much lower rate than expected, but it required the existence of a quark with the same charge as the up quark but with a much larger mass.

On the experimental side of things, the notion that protons and neutrons were fundamental particles was also being challenged. The deep inelastic scattering experiments at the Standford Linear Accelerator Center (SLAC) performed by Kendall, Friedman and Taylor in 1968 **Kendall, Friedman, Taylor** revealed unexpected⁴ behavior when probing the structure of the proton: it appeared to be composed of point-like particles. These experiments were performed by firing electrons at stationary protons and measuring the energy distributions of the scattered electrons at

⁴Depending on who you asked at the time, both the three and four quark models were not universally accepted.

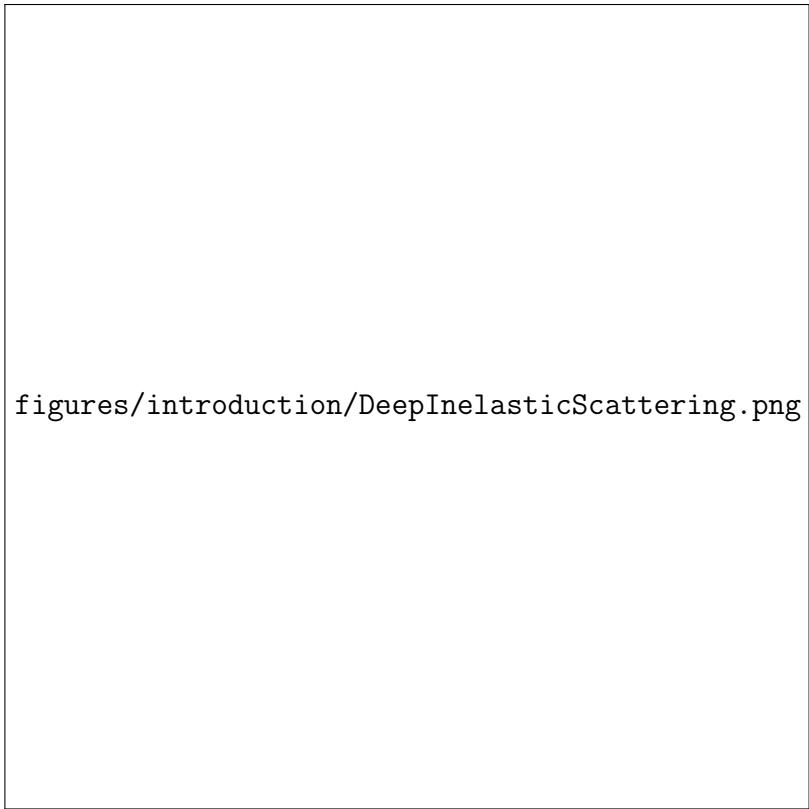
different scattering angles. An example such a distribution for electrons with initial energies of 10 GeV scattered at 6 degrees can be seen in Figure ???. The large spike on the left side of the distribution corresponds to the elastic scattering of the electron off the proton, which was well understood at the time **ElasticScattering**. The “bumps” observed at lower values of the scattered electron energy were also well understood **Resonances**, and they correspond to the “shallow” inelastic scattering of the electron off the proton, where the proton gets excited into a so-called *resonance* state (like the Δ baryon). However, the “background” underneath the bumps and the apparent continuum of events at even lower values of the scattered electron energy correspond to a mess of unknown particles being produced. This mess of particles appeared to grow with increasing scattering angle and decreasing scattered electron energy, which ultimately led to the conclusion that the proton was composed of point-like particles that were being “knocked out” of the proton by the incoming electron **Kendall**.

While many physicists were perfectly happy to interpret these point-like particles as the very same quarks from the aforementioned quark model(s), they received the much more noncommittal name **partons** after Richard Feynman’s parton model of hadrons **Partons**. The association of these partons with quarks was not universally accepted⁵ until the discovery of the J/ψ meson in 1974 **Jpsi**. In the meantime, the theoretical description of the strong nuclear force was closing in on its final form.

It turns out that Feynman’s noncommittal name was the correct choice, as in the coming years it was discovered that partons were actually quarks *and* **gluons**—the mediating bosons of the stong interaction **QCDMGM**. The gluons themselves were part of the newly formulated theory of Quantum ChromoDynamics (QCD) by AUTHORS, which addressed many of the issues that were present in the initial quark model(s).

which was developed by David Gross, Frank Wilczek and David Politzer in 1973 **QCDGross**, **QCDWilczek**, **QCDPolitzer**. This theory was able to explain the apparent “point-like” nature of the partons by introducing the idea of **asymptotic freedom**—the notion that the strong interaction becomes weaker at higher energies and shorter distances **AsymptoticFreedom**.

⁵No acceptance of any model is a step function, but the discovery of J/ψ seems to be a turning point in literature.



figures/introduction/DeepInelasticScattering.png

Figure 1.3: The energy distribution of electrons scattered off of protons at an initial electron energy of 10 GeV and a scattering angle of 6 degrees. The large spike on the left side of the distribution corresponds to the elastic scattering of the electron off the proton, and the “bumps” correspond to the inelastic scattering of the electron off the proton. The “background” underneath the bumps and the apparent continuum of events at even lower values of the scattered electron energy correspond to a mess of unknown particles being produced.

1.2 The Standard Model

During this time, the theory that describes how light and matter interact (known as Quantum ElectroDynamics, or QED) was being developed by the likes of Feynman, Schwinger, Tomaga and Dyson **QEDFeymnan**, **QEDSchwinger**, **QEDTomaga**, **QEDDyson**.

The notion that protons and neutrons were unbreakable was relatively short lived, as not even half a century later the deep inelastic scattering experiments performed by Kendall, Friedman and Taylor **Kendall**, **Friedman**, **Taylor** revealed that pro-

Standard Model of Elementary Particles

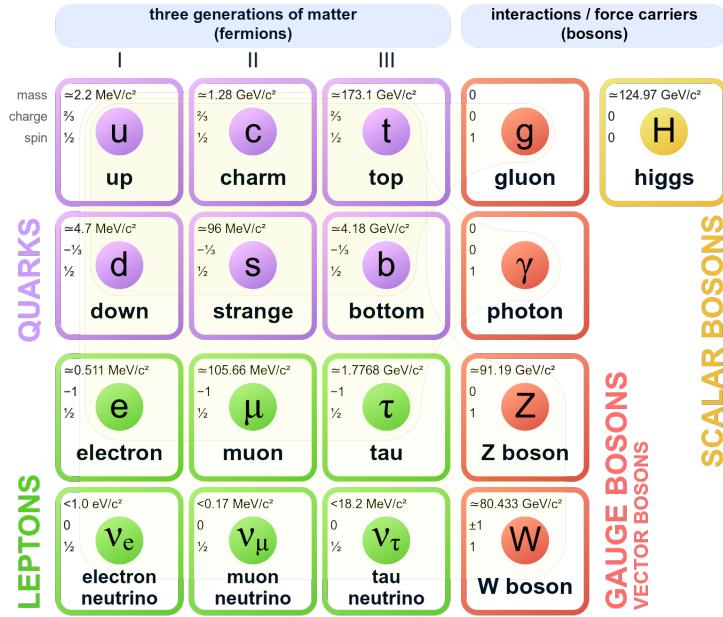


Figure 1.4: A diagram depicting the particles we currently believe are fundamental within the so-called “Standard Model” of particle physics.

tons (and subsequently neutrons) were actually composed of even smaller particles, eventually dubbed “partons”. **Partons**. This discovery was one of the largest contributing factors to the creation of the so-called Standard Model of particle physics, a theory which describes all of the fundamental particles and the way in which they interact with each other. A diagram of those fundamental particles can be seen in Figure ???. It should be noted that all of the particles labeled as quarks and leptons – collectively as “fermions” – have corresponding anti-particles with opposite electric charge. The equation that describes all of these particles and their interactions, often incorrectly¹ referred to as the “Standard Model Lagrangian”, can be compactified into a relatively palatable form that can easily fit on a coffee cup like the one shown in Figure ???.

While this equation may appear brief⁶, it can be used to completely describe three

¹It is “incorrect” because this is technically a Lagrangian density (i.e. Lagrangian per unit volume), but as it is usually integrated over all space the distinction is mostly irrelevant.

⁶Here “brief” is in the eye of the beholder, but certainly its brevity is misleading as even in the first line the $F_{\mu\nu}$ refers to three completely different gauge field tensors with their indices fully

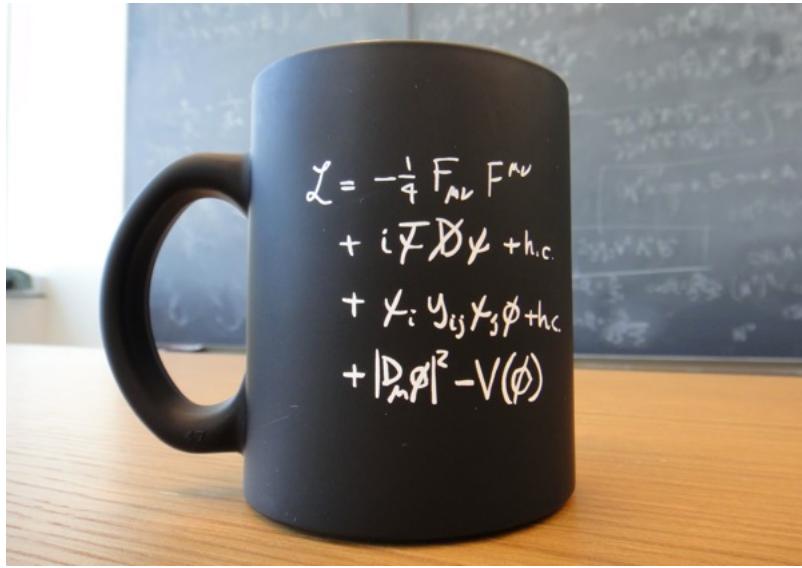


Figure 1.5: A coffee cup with the Standard Model Lagrangian density printed on its side. Please ignore the “+ h.c.” term following the $i\bar{\psi}\not{D}\psi$, it is the result of a small lapse in judgement from the mug makers.

of the four fundamental forces of nature:

1. The Electromagnetic Force, which is responsible for the electrons pushing against each other to keep you from falling through your chair,
2. The Weak Nuclear Force, which is responsible for initiating the nuclear fusion reactions that fuel our sun, and
3. The Strong Nuclear Force, which is responsible for holding quarks and gluons together in bound states known as hadrons, like the protons and neutrons that make up everyday matter.

The only fundamental force missing from this list is the Gravitational Force, which is described by a completely separate set of equations⁷

Each of the three forces that are described within the Standard Model are mediated by different gauge bosons. For example, the electromagnetic force is mediated by the boson known as the photon, the weak nuclear force is mediated by the W and contracted...

⁷Specifically, the Einstein Field Equations, $G_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu}$, but this is the thesis of a particle physicist so gravity is taboo.

Z bosons, and the strong nuclear force is mediated by bosons known as gluons. In this thesis we will be primarily focusing on the Strong Nuclear Force, which acts solely on particles with color charge – an intrinsic property of quarks and gluons. The “color” charges are red, green, and blue with antio

Even though each of the electromagnetic, weak and strong forces can be described using the Standard Model Lagrangian, the way in which they appear within the equation is not easy to determine. For example, the electromagnetic force actually corresponds to line 1

Chapter Two: Experimental Apparatus

As this thesis is focused on the physics of heavy-ion collisions, it stands to reason that the data analyzed in this thesis was gathered using the only detector along the LHC dedicated to studying such collisions: the ALICE detector. In this chapter, a brief synopsis of the LHC will be provided, followed by a much more detailed overview of the ALICE detector and its corresponding sub-detectors.

2.1 The LHC

Located along the Swiss-French border near Geneva, Switzerland, the Large Hadron Collider (LHC) is the largest particle accelerator on the planet. At a circumference of 27 kilometers, its tunnels lie almost 200 meters beneath the surface of the earth. Inside the tunnels are two high-energy particle beams pointing in opposite directions, with the beam pipes being kept inside of an ultra-high vacuum. The particles inside the beam are guided by a multitude of superconducting magnets: 393 quadrupole magnets keep the beam focused, while 1232 dipole magnets bend the particles along the circular path. The beams are designed to collide at four intersection points along the LHC, each with a corresponding detector surrounding the collision points: (1) ALICE, which specializes in heavy-ion collisions; (2) ATLAS, which specializes in studying high- p_T particles produced in pp collisions, (3) CMS, which TODO and (4) LHCb, which is designed to study CP violations through measurements of B mesons at forward rapidity. A diagram of the LHC with these four intersection points can be seen in Figure ??.

Currently, the highest center of mass energies achieved for each of the main collision systems are $\sqrt{s} = 13$ TeV for pp, $\sqrt{s} = 7$ TeV for p-Pb and $\sqrt{s} = 5.02$ TeV for Pb-Pb. The LHC underwent a long shutdown from XXXX to YYYY, in order to upgrade the beam luminosity and COM energies. The projected final COM energies for each collision system will be ...

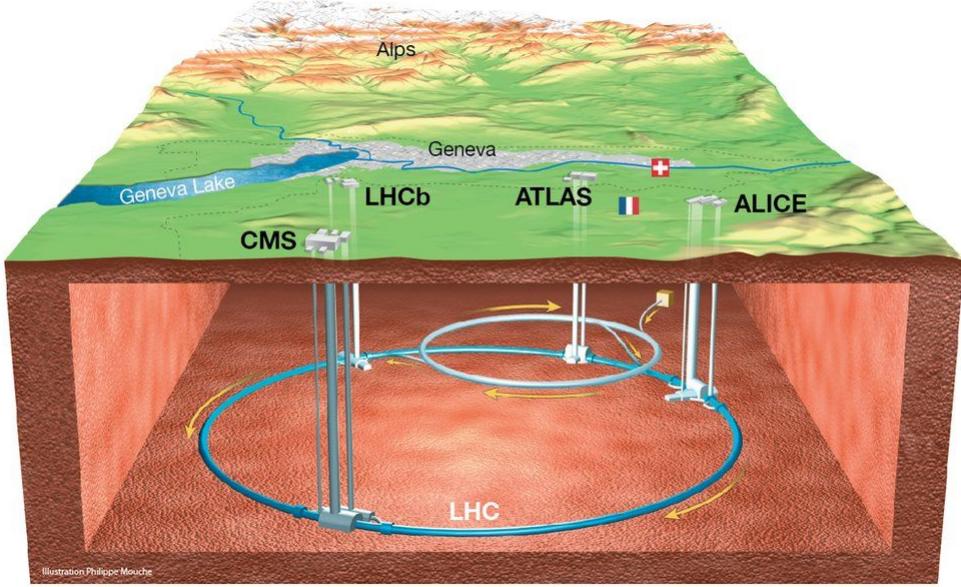


Figure 2.1: A diagram depicting the LHC with its various main detectors shown underground. Illustration by Phillippe Mouche, from BBC News.

2.2 The ALICE Detector

The detector used by the ALICE collaboration, unsurprisingly known as the ALICE detector, has the primary focus of investigating the physical properties of the strongly interacting quark-gluon plasma created during heavy-ion collisions. Building the detector was a massive effort, requiring the help from over 1000 people from 105 institutes in 30 different countries. The detector itself is also massive, weighing in at around 10,000 tons and spanning 26 meters in length with a 16-meter height and width. It is composed of 18 sub-detector systems, all of which work together to help reconstruct the event. A diagram of the detector with its corresponding sub-detector systems can be seen in in Figure ???. As the primary focus of the ALICE detector is to study heavy-ion collisions, all of its components must work together to reconstruct very high multiplicity events.

2.2.1 The Inner Tracking System

The Inner Tracking System (ITS) is the inner-most component of the ALICE detector, lying closest to the beam pipe. It is composed of six cylindrical layers of silicon

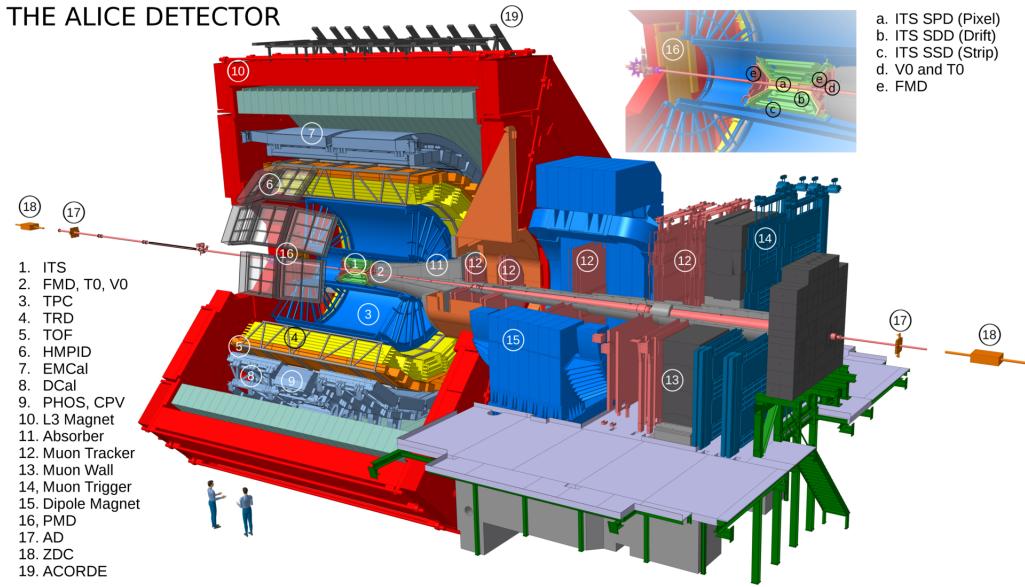


Figure 2.2: A 3-D schematic of the ALICE detector, with labels for all of the sub-detectors. Note the humans-for-scale in the bottom left of the diagram.

detectors that are coaxial with the beam pipe and cover the pseudorapidity range $|\eta| \leq 0.9$. The distance from the beam line varies from 3.9 cm for the first layer to 43 cm for the sixth layer. The ITS uses different types of silicon detectors for each layer: Silicon Pixel Detectors (SPD) for the first and second layers, Silicon Drift Detectors (SDD) for the third and fourth layers, and double sided Silicon Strip Detectors (SSD) for the fifth and sixth layers. Because of its proximity to the interaction point, the ITS is invaluable for reconstructing both primary and secondary vertices and enhancing the tracking capabilities of the ALICE detector near the interaction point. Moreover, the ITS can also track particles that are not detected or missed by the external barrel detector due to acceptance limitations and momentum cutoff.

2.2.1.1 ITS Upgrade

The LHC underwent a fairly substantial upgrade to the beam luminosity from X to Y. In order to utilize all of the The writer of this thesis was the main reason why the ITS Upgrade actually finished, as he is the smartest person of all time and actually hand built the entire thing.

2.2.2 The V0 Detector

V0 is made of two arrays of scintillator counters set on both sides of the ALICE interaction point, and called V0-A and V0-C. The V0-C counter is located upstream of the dimuon arm absorber and cover the spectrometer acceptance while the V0-A counter will be located at around 3.5 m away from the collision vertex, on the other side. It is used to estimate the centrality of the collision by summing up the energy deposited in the two disks of V0. This observable scales directly with the number of primary particles generated in the collision and therefore to the centrality. V0 is also used as reference in Van Der Meer scans that give the size and shape of colliding beams and therefore the luminosity delivered to the experiment.

2.2.3 The Time Projection Chamber

The largest component of the ALICE detector is known as the Time Projection Chamber (TPC). The TPC is a gas-filled volume with The ALICE Time Projection Chamber (TPC) is a large volume filled with a gas as detection medium and is the main particle tracking device in ALICE.[19][20] Charged particles crossing the gas of the TPC ionize the gas atoms along their path, liberating electrons that drift towards the end plates of the detector. The characteristics of the ionization process caused by fast charged particles passing through a medium can be used for particle identification. The velocity dependence of the ionization strength is connected to the well-known Bethe-Bloch formula, which describes the average energy loss of charged particles through inelastic Coulomb collisions with the atomic electrons of the medium. Multiwire proportional counters or solid-state counters are often used as detection medium, because they provide signals with pulse heights proportional to the ionization strength. An avalanche effect in the vicinity of the anode wires strung in the readout chambers, gives the necessary signal amplification. The positive ions created in the avalanche induce a positive current signal on the pad plane. The readout is performed by the 557 568 pads that form the cathode plane of the multi-wire proportional chambers (MWPC) located at the end plates. This gives the radial distance to the beam and the azimuth. The last coordinate, z along the beam direction, is given by the drift time. Since energy-loss fluctuations can be considerable, in general many pulse-height measurements are performed along the particle track in order to optimize the resolution of the ionization measurement. Almost all

of the TPC’s volume is sensitive to the traversing charged particles, but it features a minimum material budget. The straightforward pattern recognition (continuous tracks) make TPCs the perfect choice for high-multiplicity environments, such as in heavy-ion collisions, where thousands of particles have to be tracked simultaneously. Inside the ALICE TPC, the ionization strength of all tracks is sampled up to 159 times, resulting in a resolution of the ionization measurement as good as 5

2.2.4 The Electromagnetic Calorimeter

The EMCAL is a lead-scintillator sampling calorimeter comprising almost 13,000 individual towers that are grouped into ten super-modules. The towers are read out by wavelength-shifting optical fibers in a shashlik geometry coupled to an avalanche photodiode. The complete EMCAL will contain 100,000 individual scintillator tiles and 185 kilometers of optical fiber, weighing in total about 100 tons. The EMCAL covers almost the full length of the ALICE Time Projection Chamber and central detector, and a third of its azimuth placed back-to-back with the ALICE Photon Spectrometer – a smaller, highly granular lead-tungstate calorimeter. The super-modules are inserted into an independent support frame situated within the ALICE magnet, between the time-of-flight counters and the magnet coil. The support frame itself is a complex structure: it weighs 20 tons and must support five times its own weight, with a maximum deflection between being empty and being fully loaded of only a couple of centimeters. Installation of the eight-ton super-modules requires a system of rails with a sophisticated insertion device to bridge across to the support structure. The Electro-Magnetic Calorimeter (EM-Cal) will add greatly to the high momentum particle measurement capabilities of ALICE.[26] It will extend ALICE’s reach to study jets and other hard processes.

Chapter Three: Analysis Details

This chapter builds upon the analysis overview presented in the previous chapter by providing a much more detailed description of each component of the analysis. These components can be summarized as follows. First, a high-quality data sample of p–Pb collisions is selected, with events further differentiated by their multiplicity. Then, quality tracks are selected for the trigger and associated charged hadrons, and the Λ baryons are reconstructed from lower quality tracks using their characteristic decay topology. These Λ daughter tracks are identified as protons or pions using information from the TPC and TOF detectors. Within a given event, the trigger hadrons are then combined with either the associated charged hadrons or the Λ candidates to form pairs, where a distribution of their relative azimuthal angle ($\Delta\varphi \equiv \varphi_{trig.} - \varphi_{assoc.}$) and pseudorapidity ($\Delta\eta \equiv \eta_{trig.} - \eta_{assoc.}$) is filled for each pair. These h– Λ and h–h angular distributions are then corrected for a laundry list of detector effects using both data- and MonteCaro-driven methods. Further corrections are applied to the h– Λ distributions to account for effects like the combinatorial background associated with the Λ reconstruction and the two-track merging effect, whereby one of the daughter tracks gets merged with the trigger hadron track, causing a h– Λ pair deficit at small angles. Once all corrections are applied, the h– Λ and h–h distributions are finalized and ready for the extracting of the many observables discussed at the end of the previous chapter.

3.1 Dataset and event selection

3.1.1 Dataset

Every event in this analysis was a p–Pb collision at $\sqrt{s_{NN}} = 5.02$ TeV with data collected by the ALICE detector during the 2016 LHC run. This analysis uses the data from these runs with the “FAST” reconstruction, meaning the data was taken without the ITS’s SDD subdetector due to issues with readout during this period. The total number of events (prior to any selection) is roughly 400 million. For the efficiency studies, the analysis was performed using a standard purpose MC-generated

production anchored to the dataset using the DPMJET **DPMJET** event generator. This production consists of around 400 million minimum bias events, which is roughly equivalent to data.

3.1.2 Event Selection

Events are selected by requiring the location of the primary collision interaction point (called the “primary vertex” or PV) to be no more than 10 cm from the center of the detector along the beam axis or “z”-direction. Furthermore, every event is required to have at least three reconstructed tracks that contributed to the reconstruction of the PV. This reduces the total number of events considered to approximately 350 million events, and a summary of the effects of these selection criteria can be seen in Table ???. The events are further separated into three charged particle multiplicity classes (0-20%, 20-50% and 50-80%) based off event activity in the forward-rapidity V0A detector.

Table 3.1: Number of events passing our criteria for each multiplicity bin considered. Here Z_{vtx} refers to the position of the PV along the beam (z) axis.

Multiplicity	Total evts.	Has 3 tracks	$ Z_{vtx} < 10\text{cm} + 3 \text{ tracks}$	% Pass
0-20%	1.0E08	1.0E08	0.8E08	87%
20-50%	1.6E08	1.6E08	1.3E08	86%
50-80%	1.6E08	1.6E08	1.3E08	86%

3.2 Charged hadron track selection

3.2.1 Trigger track cuts

For any two-particle correlation analysis, the selection criteria of the trigger hadron is of utmost importance as any geometric biases introduced by the trigger selection could be reflected in the final correlation distributions. However, correlation analyses generally require large statistics, thus the selection criteria shown in Table ?? are applied to ensure the quality of the trigger hadron track while maximizing the statistics of the analysis. Furthermore, the trigger hadron tracks are required to be

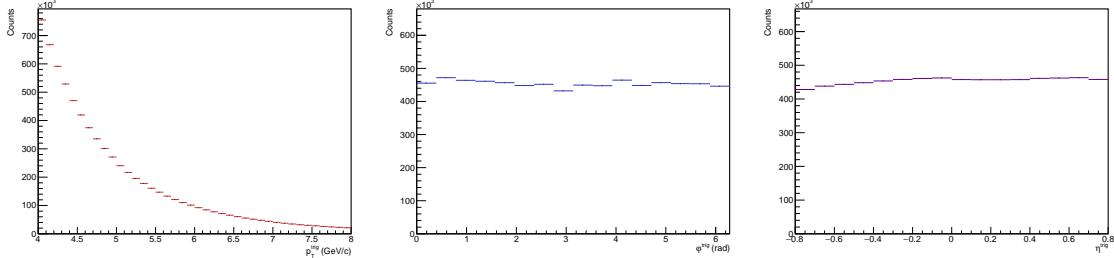


Figure 3.1: The p_T (left), φ (middle) and η (right) distributions for the trigger hadrons in the multiplicity range 0-20%.

at midrapidity ($|\eta| < 0.8$) and have high¹ momentum with $4.0 < p_T^{\text{trig.}} < 8.0 \text{ GeV}/c$, as the trigger is meant to serve as a proxy for a jet axis. Plots of the p_T , φ and η distributions for the trigger hadrons that pass these cuts in the 0-20% multiplicity bin can be seen in Figure ??.

Table 3.2: The track quality cuts applied to the trigger hadrons in this analysis.

Selection criterion	Value
TPC clusters	≥ 50
χ^2 per TPC cluster	< 4
Fraction of shared TPC clusters	< 0.4
DCA _{xy}	< 2.4 cm
DCA _z	< 3.2 cm
Accept kink daughters	No

3.2.2 Associated hadron track cuts

To keep the results of this analysis more comparable to previous measurements of the $\Lambda/\pi \approx \Lambda/h$ ratio, the selection criteria for the associated hadrons are more strict than those for the trigger hadrons as the associated hadrons are meant to be “primary”, meaning they did not originate from a weak decay. All associated hadrons are required to meet the ALICE standard track quality cuts for primary charged hadrons described in Table ???. Furthermore, the associated hadrons are selected only at midrapidity ($|\eta| < 0.8$) in the momentum region $1.0 < p_T < 4.0 \text{ GeV}/c$, with further binning

¹“High” in this case means high enough to guarantee the hadron is produced close (in $\Delta\varphi\Delta\eta$ -space) to a jet axis.

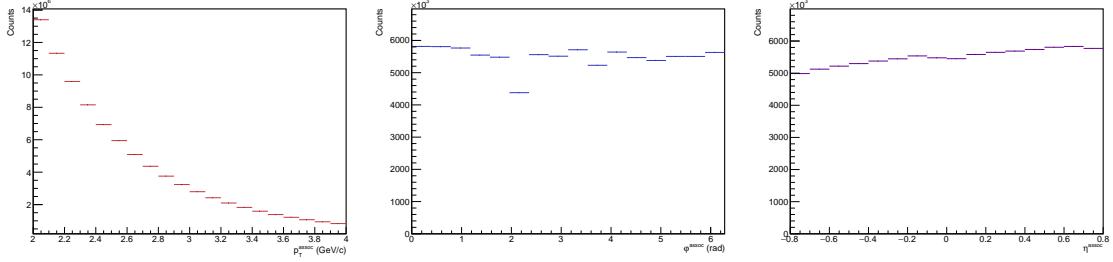


Figure 3.2: The p_T (left), φ (middle) and η (right) distributions for the associated hadrons in the multiplicity range 0-20%. The dips observed in the φ distribution are due to the TPC sector boundaries.

performed offline. The p_T , φ and η distributions for the associated hadrons that pass these cuts in the 0-20% multiplicity bin can be seen in Figure ??.

Table 3.3: The ALICE standard track quality cuts for primary charged hadrons, used for the selection of the associated hadrons in this analysis.

Selection criterion	Value
Crossed rows in TPC	≥ 80
Crossed rows/findable clusters in TPC	> 0.8
TPC clusters	≥ 80
ITS clusters	≥ 3
χ^2 per TPC cluster	< 4
χ^2 per ITS cluster	< 36
TPC and ITS refit required	Yes
DCA _{xy}	$< 0.0105 + 0.0350/p_T^{1.1}$ cm
DCA _z	< 2 cm

3.3 Λ reconstruction

3.3.1 Characteristic V⁰ decay topology

The Λ candidates in this analysis are reconstructed using their characteristic “V”-shaped decay topology, which is seen in the detector as two oppositely charged tracks originating from a common vertex which is sufficiently displaced from the PV (called the “secondary vertex” or SV). Such particles capable of being reconstructed via this topology are called “V⁰’s: the V describing the decay shape and the 0 indicating that

the particle is neutral. A diagram depicting a typical V^0 decay is shown in Figure ??, with labels given for the most relevant kinematic variables.

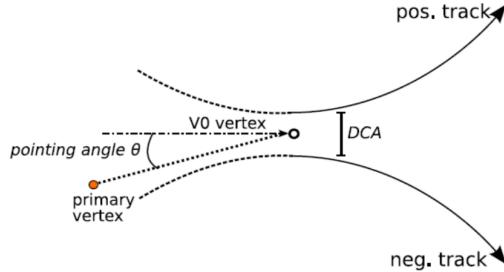


Figure 3.3: A diagram depicting a typical V^0 decay with labels for the most important kinematic variables. The diagram was taken from **V0Decay**.

The first and most important of these variables is the distance of closest approach (DCA) between the two tracks. This DCA needs to be small enough (relative to the tracking resolution) to ensure that the tracks originated from a common vertex. Another important variable is the transverse decay length of the V^0 , which is the distance between the PV and the SV measured in the xy -plane. The importance of this variable is twofold: if the decay length is too small, then it may not even be possible to resolve the SV from the PV, plus it allows for the distinction between V^0 s of differing decay lengths. The final relevant variable is the cosine of the pointing angle, which is the angle between the momentum vector of the V^0 and the vector pointing from the PV to the SV. As V^0 candidates are generally required to be sufficiently collimated to ensure that the V^0 originated from the PV, the cosine of the pointing angle is usually close to unity.

Using these variables, a list of likely V^0 candidates is generated for each event, from which further cuts are applied to maximize the likelihood of the candidate being a true Λ baryon. These cuts are summarized in the following section. There is also another technique for Λ reconstruction whereby all oppositely charged proton-pion pairs are combined to form Λ candidates, which is explored in more detail in Chapter ?. However, due to the large combinatorial background associated with this technique, the V^0 method described above is nominal for this analysis.

3.3.2 Λ daughter proton and pion track cuts

Because of the longer decay length of the Λ ($c\tau \approx 10$ cm), the corresponding daughter proton and pion tracks generally have fewer hits in both the ITS and TPC, resulting in “lower quality” track parameters. Because of this, the cuts applied to the daughter tracks used to reconstruct Λ candidates are the least strict of all the track quality cuts in this analysis and are summarized in Table ???. The daughter proton and pion are also required to be at midrapidity ($|\eta| < 0.8$) and have a minimum p_T of $p_T > 0.15$ GeV/ c .

Table 3.4: The track quality cuts applied to both the daughter proton and pion tracks used to reconstruct Λ candidates. These cuts are intentionally less strict than those applied to the trigger and associated hadrons as the daughter tracks are reconstructed from secondary particles.

Selection criterion	Value
TPC refit required	Yes
Crossed rows in TPC	≥ 70
Crossed rows/findable clusters in TPC	> 0.8

Following the particle identification procedure outlined in Sections ?? and ??, the daughter proton and pion tracks are required to pass the following PID cuts using both the TPC and TOF detectors:

- $|n\sigma_{\text{TPC},p}| < 2$
- $|n\sigma_{\text{TPC},\pi}| < 3$
- $|n\sigma_{\text{TOF},p}| < 2$ (if signal exists)
- $|n\sigma_{\text{TOF},\pi}| < 3$ (if signal exists)

The values of these cuts were chosen to maximize the Λ signal while avoiding contamination from other particle species. The parenthetical “if signal exists” means that the TOF PID cut is only applied if the track has a TOF signal. Due to the large distance between the TOF detector and the PV, many lower momentum tracks are deflected by the magnetic field before reaching the TOF detector, resulting in no signal. Excluding such tracks results in a more pure sample of protons and pions, at the cost of a much lower number of Λ candidates. While such a cost is not

acceptable for the nominal analysis, the effect of excluding these tracks is investigated in Chapter ???. The $n\sigma$ distributions for both the TPC and TOF detectors of the daughter proton and pion tracks that pass the aforementioned quality cuts are shown in Figure ?? and Figure ??, respectively. To check for contamination from other particle species, the TOF and TPC information is combined to form a $n\sigma_{\text{TOF}}$ vs $n\sigma_{\text{TPC}}$ plot, which is shown for both the protons and pions in Figure ???. No contamination is observed for either the proton or pion tracks.

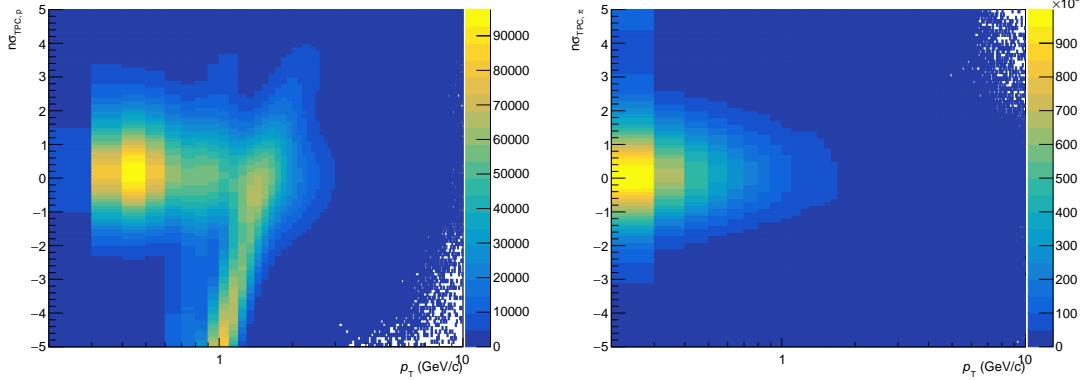


Figure 3.4: $n\sigma$ for protons (left) and pions (right) in the TPC detector as a function of p_T .

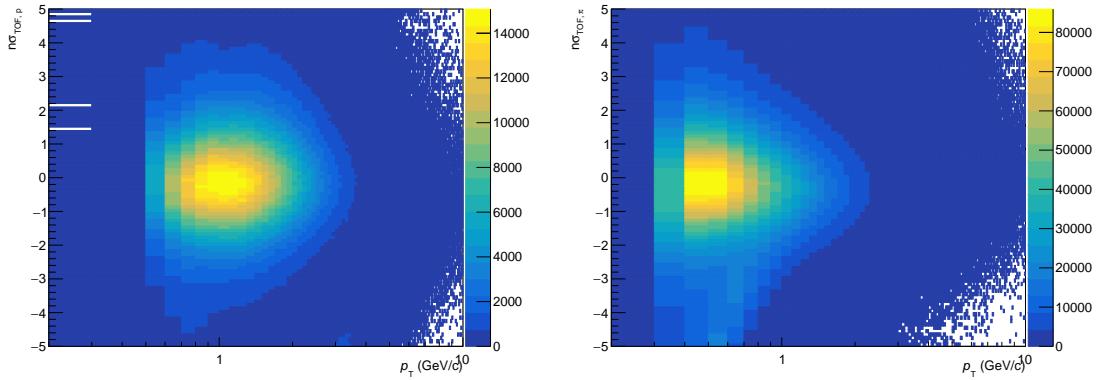


Figure 3.5: $n\sigma$ for protons (left) and pions (right) in the TOF detector as a function of p_T .

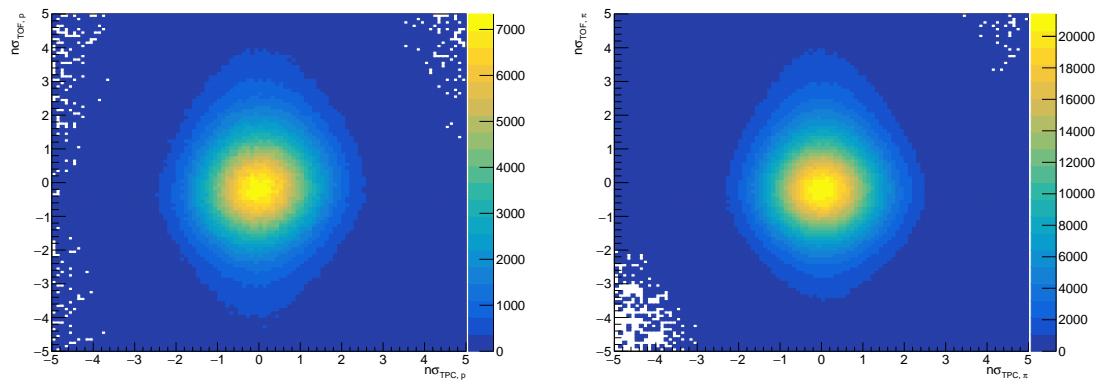


Figure 3.6: $n\sigma$ in TOF vs $n\sigma$ in TPC for protons (left) and pions (right). No contamination is observed for both of the particle species.

3.3.3 Λ candidate selection

With the daughter proton and pion tracks selected, the Λ candidates are generated by combining all oppositely charged proton-pion pairs into V^0 s which meet the topological selection criteria described in Table ??.

Table 3.5: Topological selection criteria applied to Λ candidates.

Selection criterion	Value
$ \eta $	< 0.8
Decay radius (cm)	> 0.2
DCA _{xy} of pion track to PV (cm)	> 0.06
DCA _{xy} of proton track to PV (cm)	> 0.06
DCA _{xy} between daughter tracks ($n\sigma$)	< 1.5
$\cos(\theta_{\text{pointing}})$	> 0.9
Invariant mass (GeV/c^2)	$1.102 < M_{p\pi} < 1.130$

The invariant mass $M_{p\pi}$ is calculated using

$$M_{p\pi} = \sqrt{(E_p + E_\pi)^2 - (\vec{p}_p + \vec{p}_\pi)^2}, \quad (3.1)$$

where $E_x = \sqrt{m_x^2 + p_x^2}$ is the energy of the particle of species x . The $M_{p\pi}$ distributions for the Λ candidates for all multiplicity and momentum bins are shown in Figure ???. The distributions are also fit with a Voigtian function (convolution of Breit-Wigner and Gaussian **Voigt**) plus a straight line to describe the background. Note that despite our selection criteria, there is still a non-negligible background due to the presence of misidentified Λ candidates. As this background inevitably makes its way into the final h- Λ correlation distributions, it is removed using the technique described in Section ??.

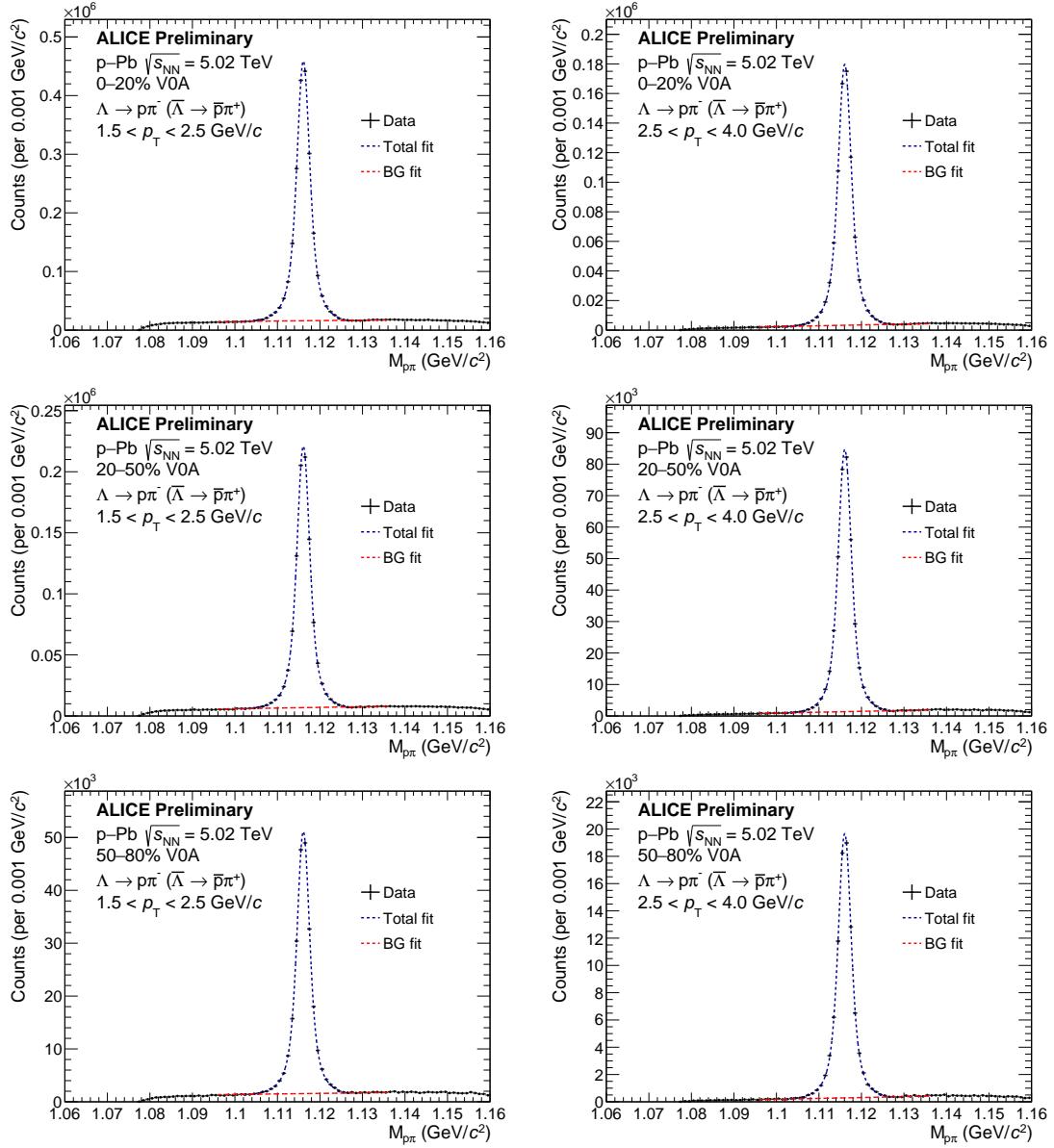


Figure 3.7: Invariant mass distributions in the 0-20% (top), 20-50% (middle), and 50-80% (bottom) multiplicity bins for the Λ candidates which pass the selection criteria with $1.5 < p_T < 2.5 \text{ GeV}/c$ (left) and $2.5 < p_T < 4.0 \text{ GeV}/c$ (right). A Voigtian signal + straight-line background fit to the data is shown in blue, with just the background fit shown in red. For these plots, the Λ s were only reconstructed in events with a trigger hadron.

3.4 Reconstruction efficiency

In an ideal world, the number of reconstructed particles of interest would be equal to the number of particles produced in the collision. Unfortunately this is not the case, as there are a number of detector effects which can cause particles to be “lost” during reconstruction. To correct for these effects, the reconstruction efficiency

$$\epsilon(x_1, x_2, \dots, x_n) \equiv P(f(x_1, x_2, \dots, x_n) | g(x_1, x_2, \dots, x_n)), \quad (3.2)$$

is used. Here x_i are the kinematic variables of the particle of interest (e.g. p_T , η , φ), $f(x_1, x_2, \dots, x_n)$ is the probability that a particle is reconstructed (“found”) with kinematic variables x_i , and $g(x_1, x_2, \dots, x_n)$ is the probability that a particle is produced (“generated”) with the same variables. While the distributions f and g are inaccessible within a given event, the efficiency can be calculated using Monte Carlo simulation techniques via the equation

$$\epsilon(x_1, x_2, \dots, x_n) = \frac{N_{\text{reco.}}(x_1, x_2, \dots, x_n)}{N_{\text{gen.}}(x_1, x_2, \dots, x_n)}, \quad (3.3)$$

where $N_{\text{reco.}}$ and $N_{\text{gen.}}$ are the reconstructed and generated particle distributions, respectively, usually taken across a large number of simulated events. In this analysis, these distributions are calculated as a function of p_T and η for each multiplicity class using 30 million events generated by the Monte Carlo event generator DPM-JET **DPMJET** with particle propagation through the ALICE detector performed by the GEANT3 **GEANT** detector simulation software. These efficiency distributions are then used to correct the h- Λ and h-h correlation distributions using the procedure described in Section ??.

3.4.1 Charged hadron reconstruction efficiency

The trigger and associated hadron track reconstruction efficiencies are calculated using Equation ??, where the trigger and associated hadrons from $N_{\text{reco.}}$ are subject to the following:

- The track passes the quality cuts outlined in Tables ?? (trigger) or ?? (associated)
- The track has a corresponding generated particle

- That generated particle is either a pion, proton, kaon, electron or muon
- $|\eta_{\text{track}}| \leq 0.8$,

and the trigger and associated hadrons from N_{gen} . are subject to:

- $|\eta_{\text{track}}| \leq 0.8$
- The particle is either a pion, proton, kaon, electron or muon
- The particle is primary (i.e. did not originate from a weak decay)

The trigger and associated track reconstruction efficiencies are shown for each multiplicity class as a function of p_T in Figure ???. While these efficiencies exhibit relatively flat behavior as a function of p_T and multiplicity, they are still treated as p_T and multiplicity dependent during the correction procedure.

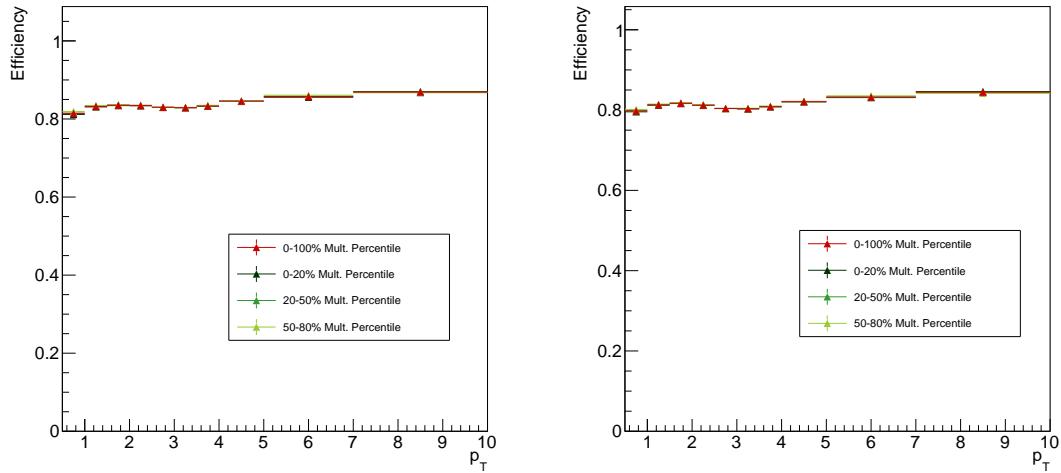


Figure 3.8: Efficiency vs. p_T for trigger (left) and associated (right) hadrons. While they may look identical, the associated hadron efficiency is slightly lower due to the stricter selection criteria.

3.4.2 Λ reconstruction efficiency

The Λ reconstruction efficiency is calculated as a function of p_T and η using Equation ??, where the Λ s from N_{reco} . are subject to the following:

- They pass the topological selection criteria from Table ??
- The reconstructed daughter p, π tracks pass the quality cuts from Table ??
- The daughter p, π tracks have corresponding generated p, π particles
- Those generated p, π daughters come from the same mother Λ
- $|\eta_\Lambda| \leq 0.8$,

and the Λ s from $N_{\text{gen.}}$ are subject to:

- $|\eta_\Lambda| \leq 0.8$
- The Λ decays to $p\pi$.

The requirement that the generated Λ s decay into $p\pi$ means the branching ratio is not included in the efficiency calculation as it is corrected for separately (see Section ??). The Λ reconstruction efficiency can be seen for each multiplicity class as a function of p_T and η in Figure ???. Note that the efficiency is no longer flat as a function of η due to the $|\eta| < 0.8$ requirement for the daughter tracks, which kinematically restricts the Λ reconstruction to a smaller η range.

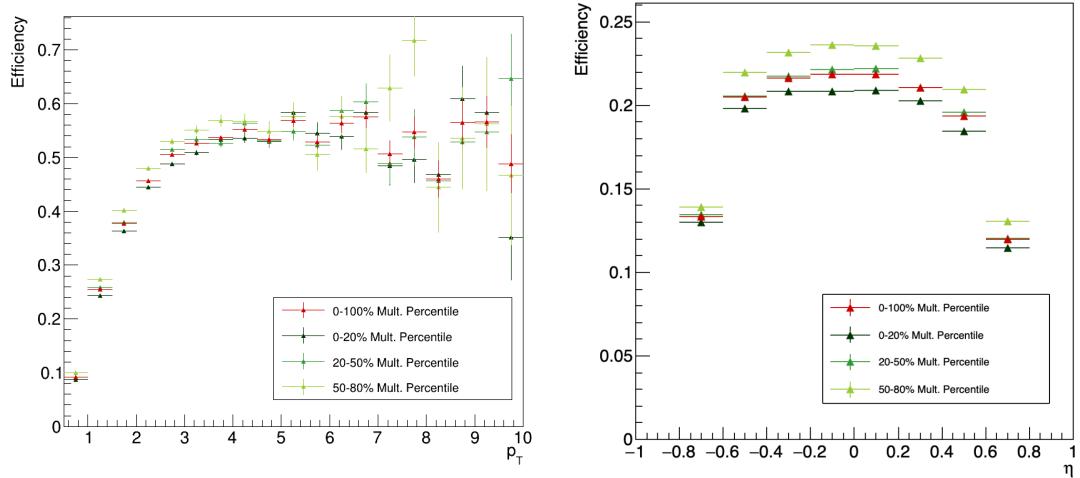


Figure 3.9: Efficiency vs. p_T (left) and η (right) for Λ reconstruction in each multiplicity bin, along with an integrated 0-100% point in red.

3.5 Corrections to the correlation distributions

Once the trigger and associated particles are selected, the two-particle h- Λ and h-h correlation distributions are generated. As mentioned in the previous chapter, the corrected two-particle correlation function is given by

$$\frac{1}{N_{trig}} \frac{d^2 N_{pair}}{d\Delta\varphi d\Delta\eta} = \frac{1}{N_{trig}^{corr}} \frac{1}{\epsilon_{trig} \times \epsilon_{assoc}} B(0, 0) \frac{S(\Delta\varphi, \Delta\eta)}{B(\Delta\varphi, \Delta\eta)}. \quad (3.4)$$

which contains a number of explicit correction terms (in the form of ϵ s) along with some implicit corrections. These corrections are described in this section, and are presented in the order in which they are applied to the data.

3.5.1 Single-particle efficiency corrections

As both the trigger and the associated particles have their own independent reconstruction efficiencies, the trigger-associated pair reconstruction efficiency should be

$$\epsilon_{trig,assoc} = \epsilon_{trig} \times \epsilon_{assoc}, \quad (3.5)$$

meaning the single-particle efficiency distributions from Section ?? can be used to calculate the weight $1/(\epsilon_{trig} \times \epsilon_{assoc})$. This weight is applied for each h- Λ and h-h pair in the two-dimensional correlation distribution. However, the assumption that the reconstruction efficiencies are independent is slightly incorrect in the case of the h- Λ distributions due to track merging effects, thus an additional ϵ_{pair} correction is required (discussed in detail in Section ??).

The trigger efficiency weight $1/\epsilon_{trig}$ is also applied to the single-particle trigger hadron distribution in data to obtain N_{trig}^{corr} .

3.5.2 Mixed-event acceptance correction

As mentioned in Section ??, the $B(0, 0)/B(\Delta\varphi, \Delta\eta)$ term in Equation ?? corrects for the finite acceptance along η as both our trigger and associated particles are required to be within $|\eta| < 0.8$. The mixed-event distribution $B(\Delta\varphi, \Delta\eta)$ shown in Figure ?? has a characteristic triangular shape along $\Delta\eta$, which is purely due to detector geometry as no physical correlations are present. When scaled by $1/B(0, 0)$, the mixed event distribution becomes the probability that a particle pair is found given

that the trigger particle is within $|\eta| < 0.8$, which is unity at $\Delta\varphi, \Delta\eta = 0, 0$. Thus correcting the same-event distribution $S(\Delta\varphi, \Delta\eta)$ by $B(0, 0)/B(\Delta\varphi, \Delta\eta)$ removes this acceptance effect and allows for a more accurate determination of the pair-wise yields.

While the generation of the mixed event distribution $B(\Delta\varphi, \Delta\eta)$ was discussed briefly in Section ??, the specific details are as follows. First, in order to ensure that the mixed-event pairs are coming from similar events, the events in the mixing pool are separated by both multiplicity percentile and Z_{vtx} position. The categorizing of events based off of Z_{vtx} position is an integral part of the acceptance correction: events with a Z_{vtx} at one edge of the detector have a completely different (and nearly inverted) η acceptance than those on the opposite edge. The multiplicity bins are the same as they are for the same-event distributions (namely 0-20%, 20-50% and 50-80%), and the ten Z_{vtx} bins are split evenly from -10 cm to 10 cm. For each multiplicity and Z_{vtx} bin, the acceptance correction

$$S_{\text{corr.}}(\Delta\varphi, \Delta\eta) = \frac{S(\Delta\varphi, \Delta\eta)}{B(\Delta\varphi, \Delta\eta)/B(0, 0)} \quad (3.6)$$

is performed, and the results for each multiplicity bin are then merged across all Z_{vtx} bins. The same-event distributions are also split into Z_{vtx} bins during this correction procedure. The uncorrected distributions $S(\Delta\varphi, \Delta\eta)$ and the mixed-event distributions $B(\Delta\varphi, \Delta\eta)$ are shown for both the h- Λ and h-h cases for all multiplicity and associated momentum bins in Figures ?? through ??.

This mixed-event correction is the final correction applied to the h-h distributions. However, the h- Λ distributions require additional corrections that are not present in the dihadron case.

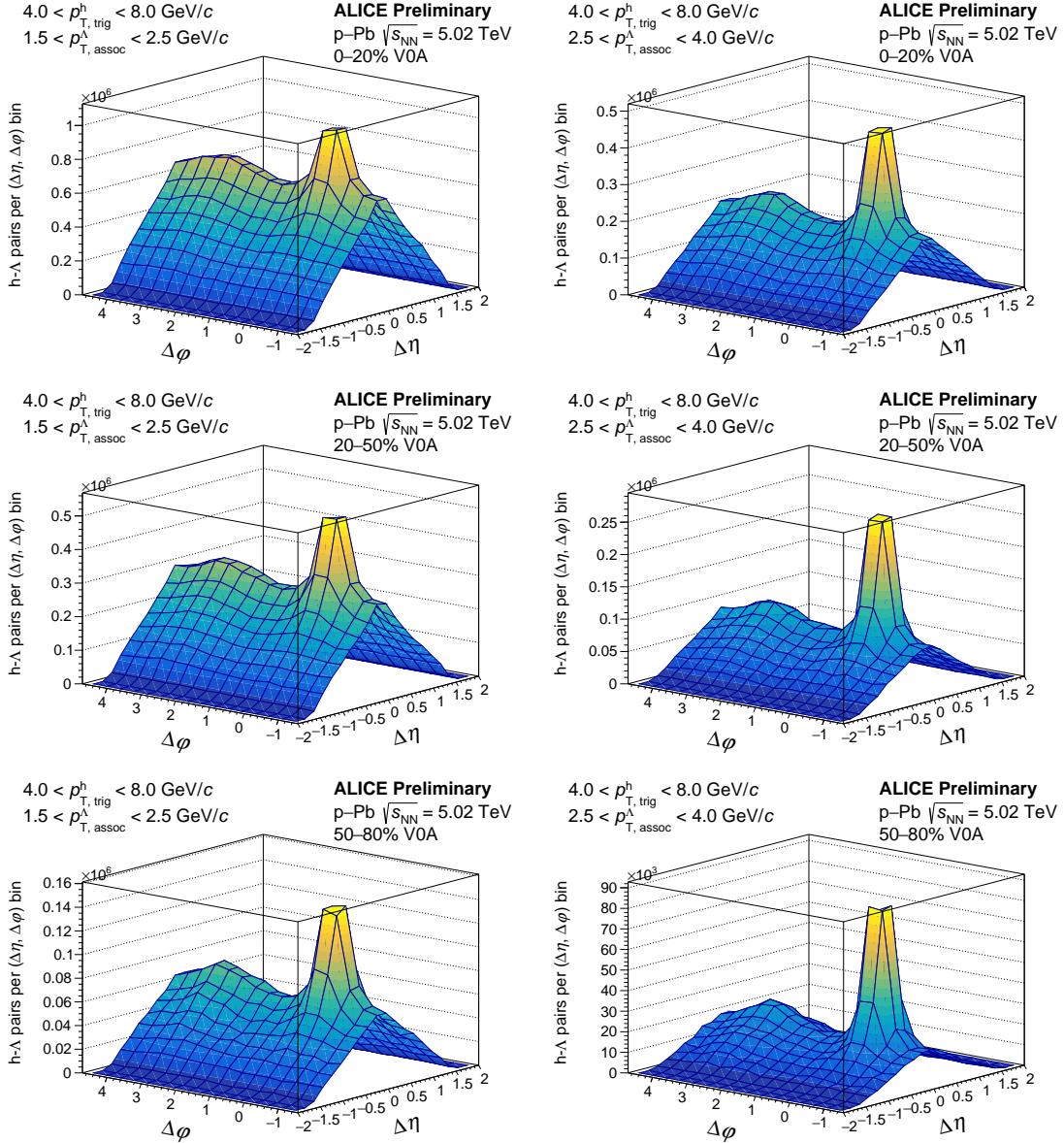


Figure 3.10: 2-D non-acceptance corrected $h\bar{\Lambda}$ angular correlations for the 0–20% (top), 20–50% (middle), and 50–80% (bottom) multiplicity bins for $1.5 < p_T < 2.5 \text{ GeV}/c$ (left) and $2.5 < p_T < 4.0 \text{ GeV}/c$ (right).

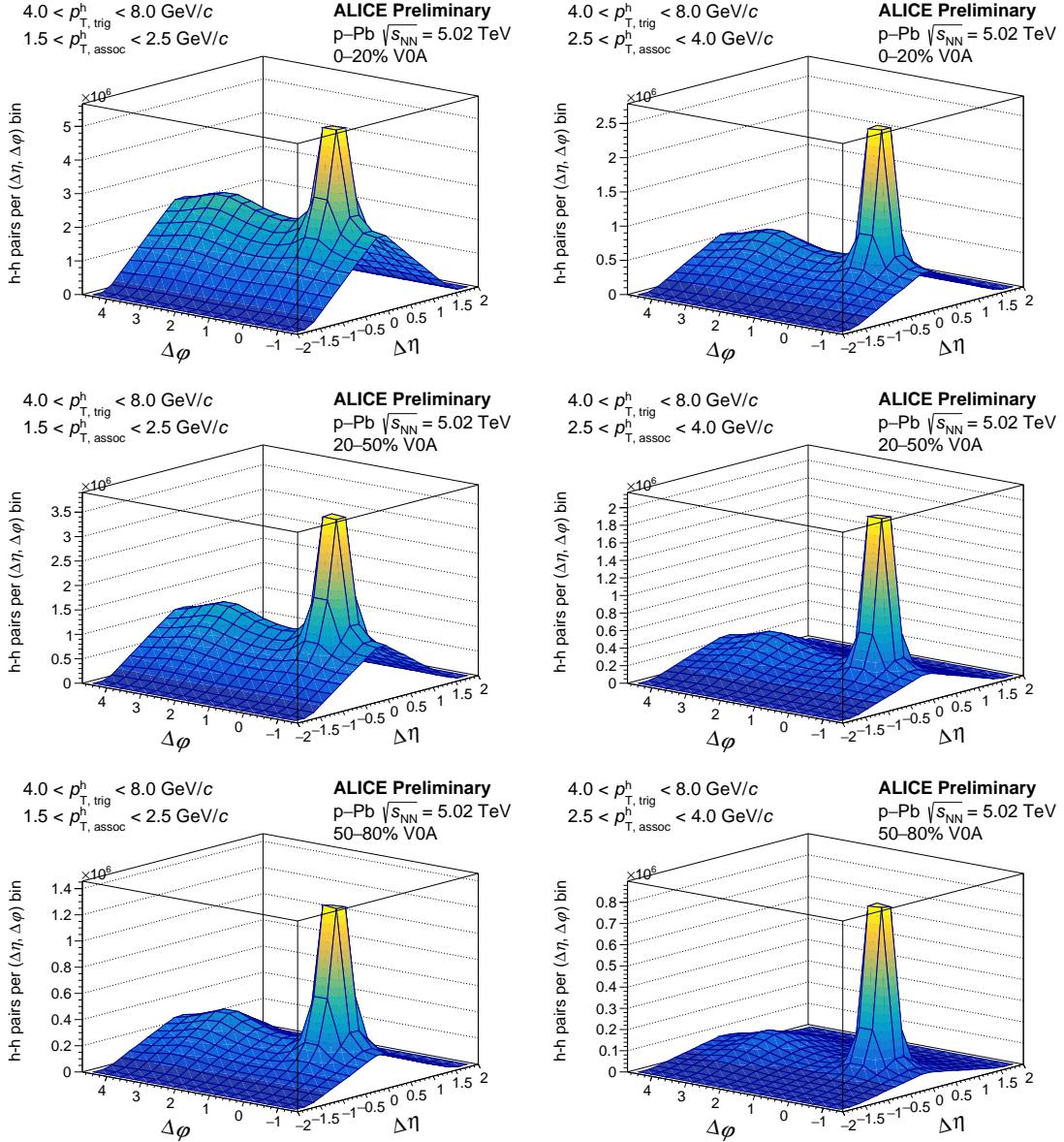


Figure 3.11: 2-D non-acceptance corrected h-h angular correlations for the 0-20% (top), 20-50% (middle), and 50-80% (bottom) multiplicity bins for $1.5 < p_T < 2.5$ GeV/c (left) and $2.5 < p_T < 4.0$ GeV/c (right).

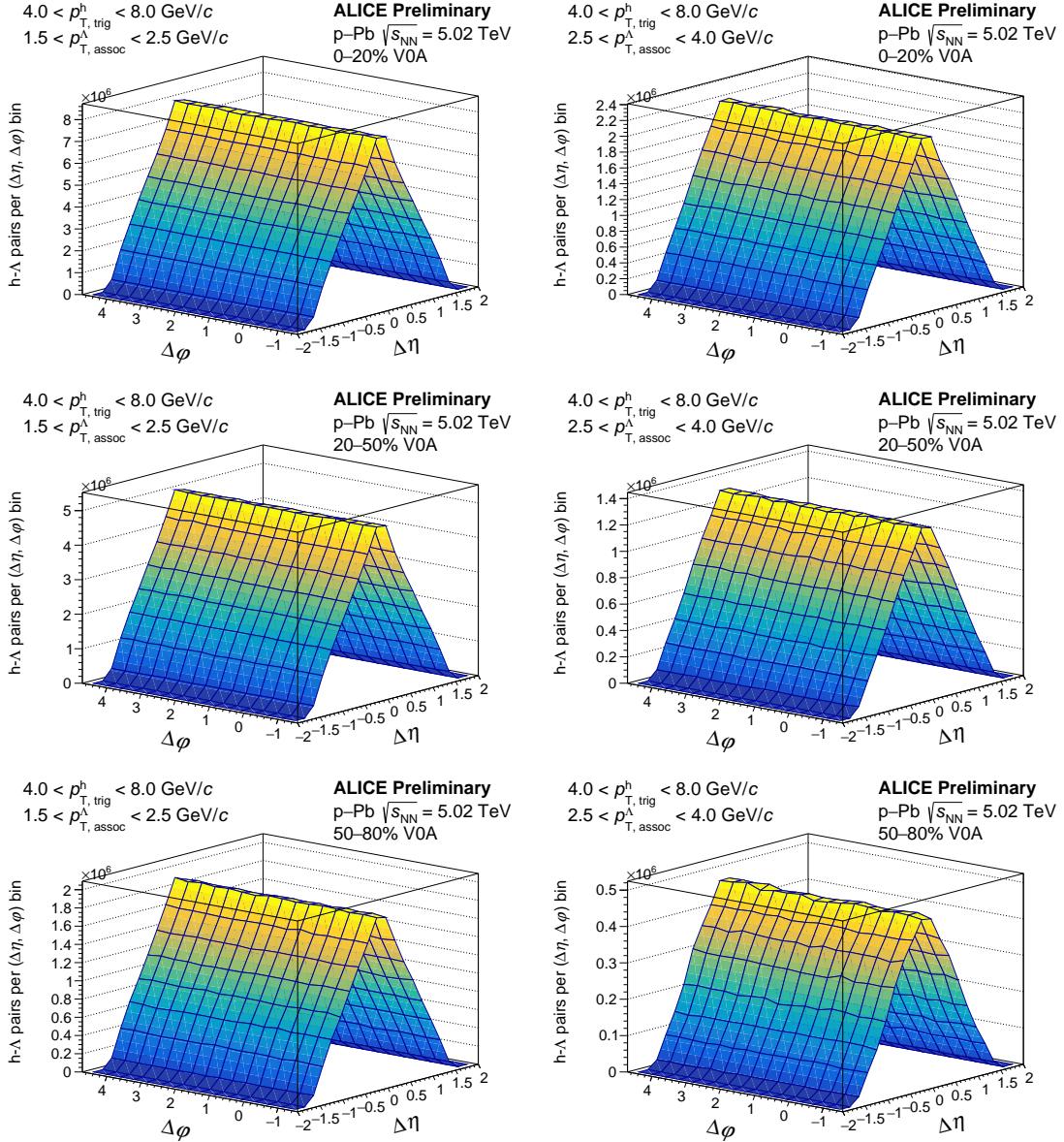


Figure 3.12: 2-D mixed-event h- Λ angular correlations for the 0-20% (top), 20-50% (middle), and 50-80% (bottom) multiplicity bins for $1.5 < p_T < 2.5 \text{ GeV}/c$ (left) and $2.5 < p_T < 4.0 \text{ GeV}/c$ (right). The $Z_{\text{vtx.}}$ bins are merged together for these plots.

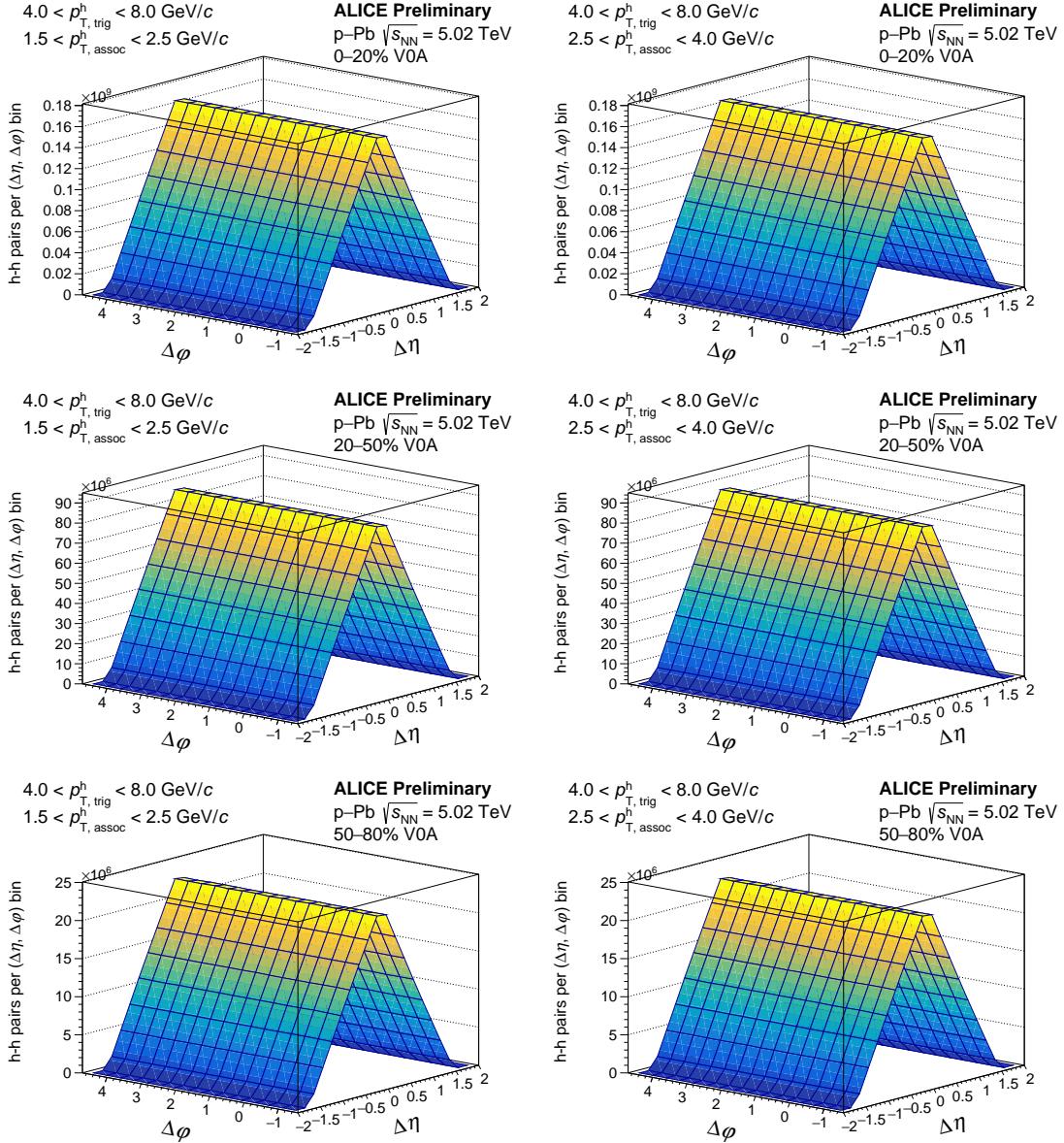


Figure 3.13: 2-D mixed-event h-h angular correlations for the 0-20% (top), 20-50% (middle), and 50-80% (bottom) multiplicity bins for $1.5 < p_T < 2.5 \text{ GeV}/c$ (left) and $2.5 < p_T < 4.0 \text{ GeV}/c$ (right). The Z_{vtx} bins are merged together for these plots.

3.5.3 Additional corrections for the h- Λ distributions

While the corrected correlation function from Equation ?? is generally true for two-particle correlations, there are a few additional corrections that must be applied to the h- Λ distributions due to the Λ reconstruction procedure and the presence of track merging effects. To formalize this, the corrected h- Λ correlation function can be written as

$$C_{\text{corr.}}^{\text{h-}\Lambda}(\Delta\varphi, \Delta\eta) = \frac{r_{\text{signal}} \times r_{\text{branch}}}{\epsilon_{\text{pair}}(\Delta\varphi, \Delta\eta)} \left(C_{\text{signal}}^{\text{h-p}\pi}(\Delta\varphi, \Delta\eta) - r_{\text{comb.}} \times C_{\text{sideband}}^{\text{h-p}\pi}(\Delta\varphi, \Delta\eta) \right), \quad (3.7)$$

where $C_{\text{corr.}}^{\text{h-}\Lambda}$ is the final corrected h- Λ distribution. Each term on the RHS of the equation will be described in detail in the following sections, and they are presented in the order in which they are applied to the distributions.

3.5.3.1 Combinatorial background removal

The term

$$C_{\text{signal}}^{\text{h-p}\pi}(\Delta\varphi, \Delta\eta) - r_{\text{comb.}} \times C_{\text{SB, norm.}}^{\text{h-p}\pi}(\Delta\varphi, \Delta\eta) \quad (3.8)$$

describes the removal of the combinatorial background resulting from the Λ reconstruction procedure from Section ?? using the **sideband subtraction** technique. The $C_{\text{signal}}^{\text{h-p}\pi}$ distribution corresponds to Λ candidates where the invariant mass of the p π pair falls within the range specified in Table ??, and the self-normalized $C_{\text{SB, norm.}}^{\text{h-p}\pi}$ distribution corresponds to candidates where the mass of the p π pair falls within the so-called “sideband” region. An invariant mass plot highlighting these different regions can be seen in Figure ???. Both of the C_{signal} and C_{sideband} distributions are corrected for acceptance and efficiency using the techniques described in the previous sections. The sideband region is chosen such that it is far enough away from the signal region to be free of any Λ signal, but close enough to ensure that the background p π pairs in the signal region are kinematically similar to the pairs in the sideband region as to not introduce any biases in the correlations. The underlying assumption of this technique is that the correlation shape of h-p π pairs from the sideband region is the same as the shape from the background h-p π pairs in the signal region. For this analysis, the nominal sideband region was chosen to be $1.135 < M_{p\pi} < 1.150$

GeV/c^2 , but the effects of varying this region are studied in detail in the next chapter. The $r_{\text{comb.}}$ is the integral of the combinatorial background in the $C_{\text{signal}}^{\text{h-p}\pi}$ distribution, obtained by

$$r_{\text{Comb}} \equiv \frac{B}{S + B} \int \int C_{\text{signal}}^{\text{h-p}\pi}(\Delta\varphi, \Delta\eta) d\Delta\varphi d\Delta\eta, \quad (3.9)$$

where S and B are the signal and background obtained from the fits to the Λ invariant mass distributions in Figure ???. As the S/B ratio is the same for the Λ invariant mass distributions in events with a trigger hadron as it is for the h- Λ distributions, the scale $B/(S + B)$ can be used to give only the background contribution from the integral $\int \int C_{\text{signal}}^{\text{h-p}\pi}(\Delta\varphi, \Delta\eta) d\Delta\varphi d\Delta\eta$. The self-normalized $C_{\text{SB, norm.}}^{\text{h-p}\pi}$ distribution is then scaled by $r_{\text{comb.}}$ and subtracted from $C_{\text{signal}}^{\text{h-p}\pi}$ to remove the combinatorial background.

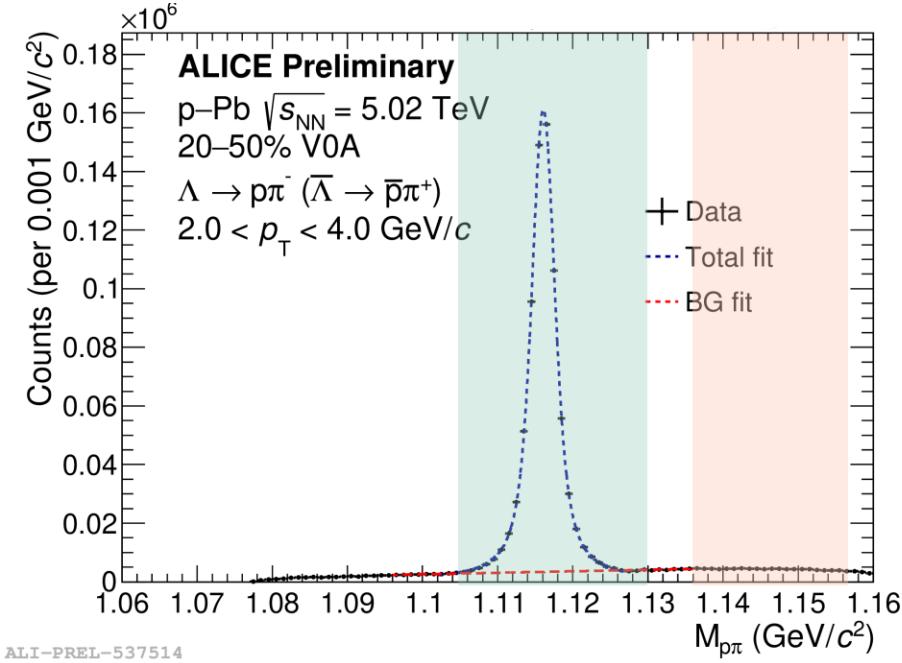


Figure 3.14: Invariant mass distribution of $p\pi$ pairs in the 20-50% multiplicity class. The signal region is shown in light green, and the sideband region is shown in light pink. The correlation distribution in the sideband region is used to remove the combinatorial background from the signal region.

While the above procedure describes the background removal in a more technical manner, it can be condensed into the following steps:

1. Generate the correlation distribution using Λ candidates in the signal invariant mass region

2. Do the same thing for Λ candidates in the sideband invariant mass region
3. Scale the sideband distribution to match the background in the signal region
4. Subtract the sideband distribution from the signal distribution

Examples of the signal and sideband distributions $C_{\text{signal}}^{\text{h-p}\pi}$ and $C_{\text{SB}}^{\text{h-p}\pi}$ are shown for the 0-20% multiplicity bin in Figure ??.

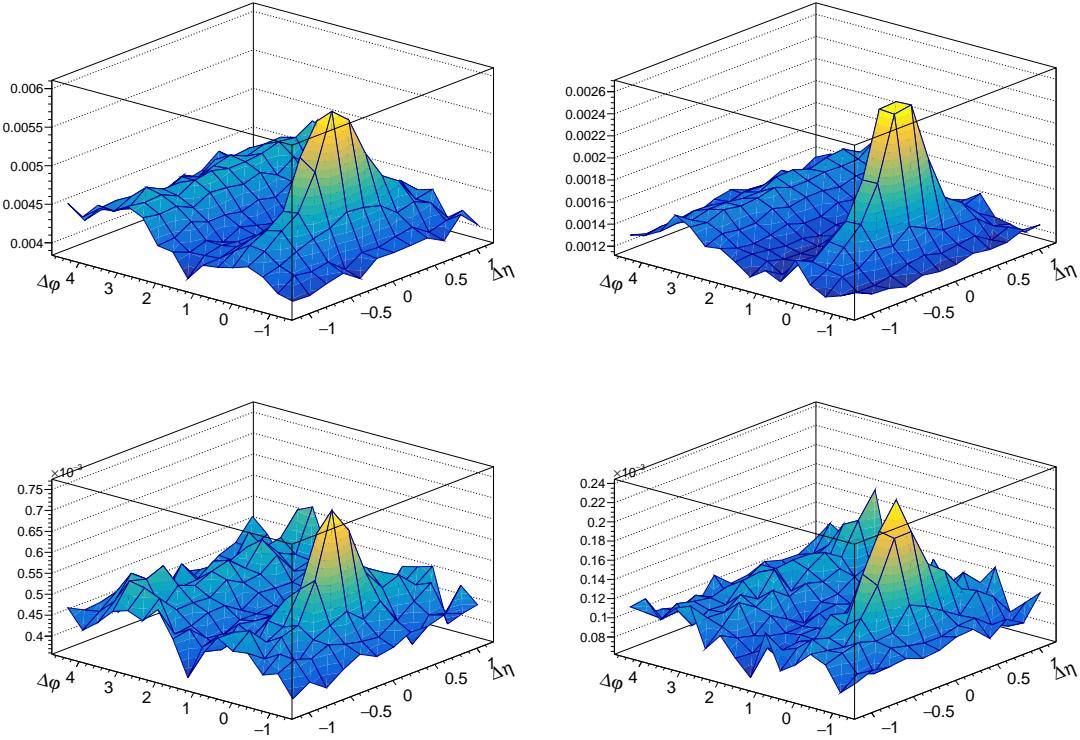


Figure 3.15: The signal (top) and sideband (bottom) distributions $C_{\text{signal}}^{\text{h-p}\pi}$ and $C_{\text{SB}}^{\text{h-p}\pi}$ for the lower (left) and higher (right) associated p_T bins. All plots were generated in the 0-20% multiplicity class.

3.5.3.2 Signal scaling

As the Λ candidate invariant mass signal region is finite, the fraction of the Λ signal that is missed in the tails of the invariant mass distribution must be corrected for.

This is handled by the r_{signal} term in Equation ??, which is calculated by

$$r_{\text{signal}} \equiv \left(\frac{\text{Integral of residual in signal region}}{\text{Integral of residual between 1.098 and 1.134}} \right)^{-1} \quad (3.10)$$

where “residual” refers to the invariant mass distributions from Figure ?? after subtracting the straight-line background fit. 1.098 and 1.134 are the points in which there is effectively zero signal, verified in Monte Carlo. Due to the width of the signal region, r_{signal} is usually near unity. However, to study the effects of narrowing the signal region, this term must be included in the analysis.

3.5.3.3 Branching ratio correction

The most simple correction from Equation ?? comes from the branching ratio term, namely

$$r_{\text{branch}} \equiv \frac{1}{BR(\Lambda \rightarrow p\pi)} = \frac{1}{0.639}. \quad (3.11)$$

As not all Λ s decay into $p\pi$ pairs, this term corrects for the fraction of Λ s that decided to decay into something else. In many analyses, this term is not required as it is already included in the efficiency computation $\epsilon_{\text{assoc.}}$. As the Λ reconstruction efficiency from this analysis is calculated using only Λ s that decay into $p\pi$ pairs, this term must be included separately.

3.5.3.4 Pair efficiency correction

The ϵ_{pair} term in Equation ?? is the h- Λ “pair” efficiency, which is used to correct for track merging effects. Many correlation studies are susceptible to track merging inefficiencies **TwoTrack1**, **TwoTrack2**, whereby either the trigger or associated particle gets merged over by the other during the track reconstruction. This results in a dip at small angles in the angular correlation distribution when compared to a similar distribution with no instances of track merging. As this effect cannot be seen directly in data due to the missing reconstructed tracks, it is investigated using the Monte Carlo sample, where the reconstructed tracks are compared to the MC-generated particles they were reconstructed from. While this effect is usually negligible and only relevant at extremely small angles ($\Delta\varphi < 0.01, \Delta\eta < 0.1$), in this analysis this effect is more severe and occurs at larger angles ($\Delta\varphi < 1, \Delta\eta < 0.6$), shown in Figure ??.

The severity of this effect for the h- Λ distributions is likely due to two factors:

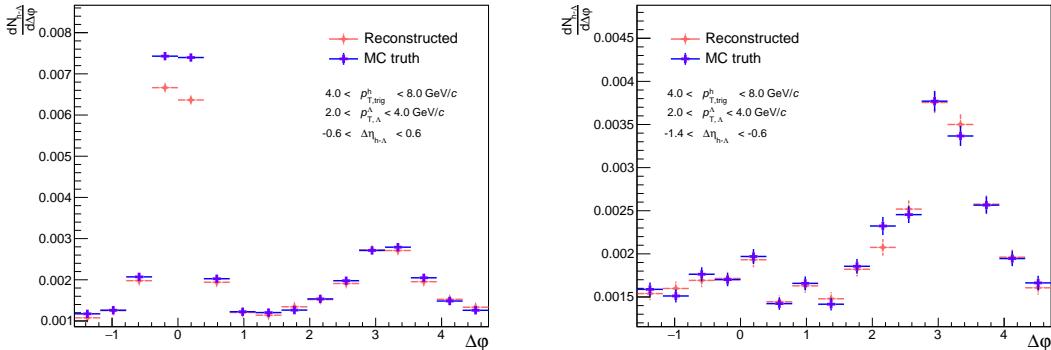


Figure 3.16: Demonstration of the track merging effect for h- Λ pairs, whereby we see a dip in the reconstructed distribution at small $\Delta\varphi$ and $\Delta\eta$ when compared to the MC ground truth (left). This dip is not present at large $\Delta\eta$ (right), but we also lose nearly the entirety of our near-side peak.

- The Λ decay length is large ($c\tau \approx 10$ cm), meaning the daughter particles will have less hits in the detector than the trigger particle (which is produced at the primary vertex). As Kalman filtering (track reconstruction) favors the track with more hits, the Λ daughter track is “merged” over by the trigger track.
- The Λ decay is asymmetric ($m_p/m_\pi \approx 7$), so the Λ and daughter proton end up with similar momenta (and thus φ and η). This means that whenever a proton from a Λ decay is “merged” over by a trigger track, a h- Λ pair with small $\Delta\varphi, \Delta\eta$ is lost.

To see how the decay length can affect the track merging, the (reconstructed)/(MC ground-truth) $C(\Delta\varphi, \Delta\eta)$ distribution ratio for h-pion pairs in our MonteCarlo sample where the pion is **secondary**—meaning it came from a weak decay with decay length > 2 cm—is measured. Pions are chosen for this demonstration as they are more abundantly produced than protons, and charged track reconstruction is particle species agnostic. Any “dips” from unity present in this ratio are indicative of pairs being lost during reconstruction. This is then compared to the same ratio for h-pion pairs where the pion is **primary**, and the results are shown in Figure ???. All reconstructed triggers and pions pass the trigger hadron and Λ daughter cuts from Tables ?? and ??, respectively. Furthermore, all distributions have been fully corrected for single-particle efficiencies and detector acceptance using the procedures

from Sections ?? and ??, respectively. A large suppression at small $(\Delta\varphi, \Delta\eta)$ is observed for the h-secondary pion case, but the h-primary pion case exhibits no such suppression. As such, it stands to reason that this suppression is at least in part due to the decay length of the Λ , as all particles that come from Λ s are secondaries (by a long shot).

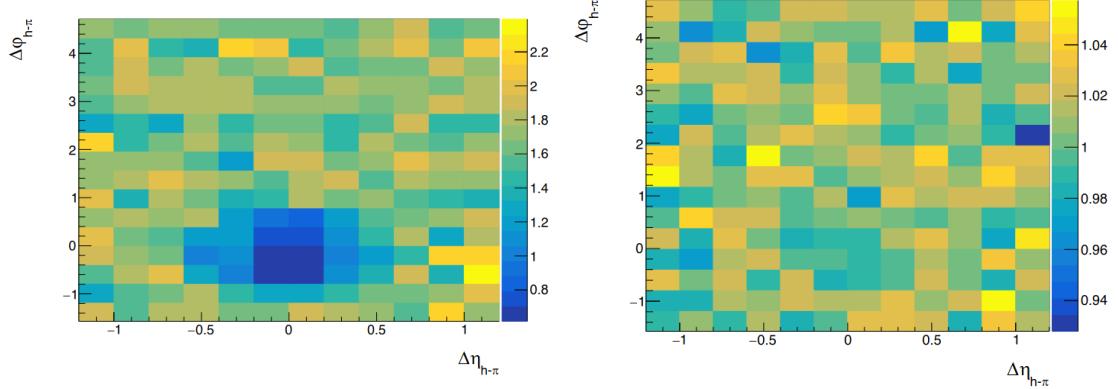


Figure 3.17: The reconstructed/ground truth ratios of the 2D $C(\Delta\varphi, \Delta\eta)$ distributions for h-(secondary pions) (left) and h-(primary pions) (right). The suppression at smaller $\Delta\varphi, \Delta\eta$ is clearly seen in the secondary case, but is not observable in the primary case, indicating a decay-length dependence.

The p_T dependence of this effect can also be studied by measuring the reconstructed and ground truth h -(secondary pion) $\Delta\varphi$ distributions at low ($1.0 < p_T < 2.0$ GeV/c) and high ($2.0 < p_T < 4.0$ GeV/c) associated momentum. The result is shown in Figure ???. Note that the distributions were projected onto $\Delta\varphi$ with $|\Delta\eta| < 1.2$. A suppression relative to MC ground-truth is observed in the near-side of the reconstructed distribution in the higher p_T range, which is not seen in the low p_T bin. This is also consistent with the decay length dependence shown in the previous figures, as decay length is roughly proportional to p_T .

The p_T dependence of this inefficiency demonstrates why this effect is so severe in the h- Λ case: due to the asymmetry of the Λ decay ($m_p/m_\pi \approx 7$), the daughter proton receives most of the momentum. Therefore when investigating h- Λ correlations within a given associated p_T range, any inefficiencies present in the corresponding h-(daughter proton) distribution with the same associated momentum would also be present in our final h- Λ distribution within a similar $\Delta\varphi, \Delta\eta$ range. As demonstrated in Figure ???, secondary charged particles with $2 < p_T < 4$ GeV/c see a large ineffi-

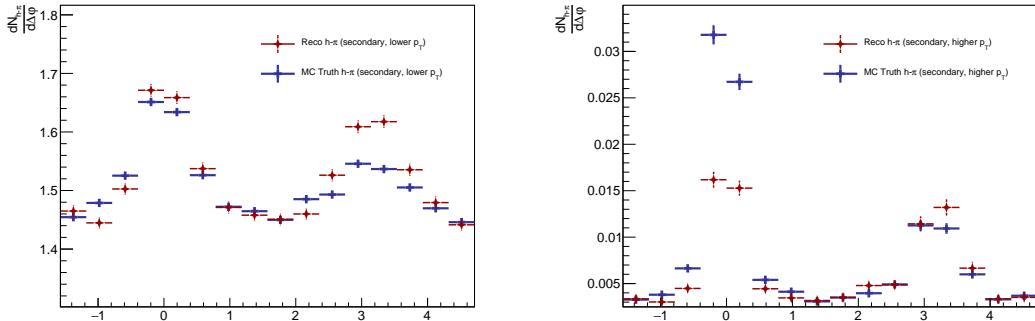


Figure 3.18: The reconstructed and ground truth $\Delta\varphi$ distributions in the $-1.2 < \Delta\eta < 1.2$ region for h-(secondary pions) with $0.15 < p_T < 2$ (left) and $2 < p_T < 4$ (right). The suppression at smaller $\Delta\eta, \Delta\varphi$ is clearly seen in the higher momentum bin, but not present in the lower one.

ciency, and therefore we would expect a similar inefficiency to be present in our h- Λ distribution (which was shown in Figure ??). In similar analyses using the K_S^0 in leiu if the Λ **Lambda1**, such an effect is not as present, both because the decay length is much shorter (2 cm vs 10 cm), and the K_S^0 decay is symmetric, meaning the daughter pions will have momenta that are no longer similar to the mother kaon.

The following techniques have been investigated to correct for this effect:

- Applying a $\Delta\varphi^*$ correction, described in **DphiStar**: While $\Delta\varphi$ and $\Delta\varphi^*$ are different, they are correlated enough that in order to remove this effect, a $|\Delta\varphi^*| < 0.7$ cut is required, which removes a significant amount of the near-side yield in the corresponding $\Delta\varphi$ distribution.
- Applying a cut on the minimum distance between the fully reconstructed helices of the trigger and Λ daughter proton (varied between 0.1 cm and 10 cm): Again, this cut removes roughly the same amount of near-side yield as the $\Delta\varphi^*$ cut, as this minimum distance is also highly correlated with $\Delta\varphi$.
- Using the resonance technique for Λ reconstruction (more details in Section ??): this moderately reduces the severity of this effect, but the statistical fluctuations introduced by the smaller S/B ratio make it difficult to gauge how effective this correction is.

- Only correlating h- Λ pairs where the charge of the Λ daughter proton (or antiproton) is opposite to the trigger, as oppositely charged tracks bend in opposite directions in the detector magnet: This reduces the effect by a considerable amount, but reduces our overall correlation statistics by a factor of 2.
- Selecting “lower quality” trigger tracks by loosening the cuts from Table ?? so they are less likely to be merged over the low-quality daughter tracks: this reduces the effect, but introduces a large amount of secondary contamination. Furthermore, we would like this analysis to be directly compared with other analyses, and therefore want to maintain the same cuts on the trigger hadron.
- Selecting “higher quality” Λ daughter tracks by tightening the cuts from Table ?? (and adding additional selection criteria): this again reduces the effect but heavily cuts into the Λ signal

As each of these techniques reduces the statistics of the h- Λ correlation distribution beyond the realm of acceptability, the two-track inefficiencies are instead corrected for using a MC-generated template method, similar to the one used in **TwoTrack2**. For this method, the pair efficiency is given by

$$\epsilon_{pair}(\Delta\varphi, \Delta\eta) \equiv \frac{C_{\text{reco}}^{\text{tag}}(\Delta\varphi, \Delta\eta)}{C_{\text{gen}}(\Delta\varphi, \Delta\eta)}, \quad (3.12)$$

where $C_{\text{reco}}^{\text{tag}}$ is the efficiency-corrected correlation distribution calculated in MC using reconstructed trigger hadrons and Λ candidates with the same selection criteria as described in Section ??, with the additional requirement that the Λ candidate has a corresponding generated Λ which is used for all calculations involving kinematic quantities. This removes the need to perform any of the additional corrections from the previous sections (e.g. background subtraction, signal scaling) as the invariant mass of generated lambdas is exact. C_{gen} is the correlation distribution calculated in MC using only generated trigger hadrons and Λ candidates. The template $\epsilon_{pair}(\Delta\varphi, \Delta\eta)$ is applied for each associated p_T bin in this analysis, but it is independent of multiplicity and event generator. The templates for each associated p_T bin are shown in Figure ???. This correction is applied to the h- Λ distributions after side-band subtraction, signal scaling and the branching ratio correction.

After these corrections, both the h- Λ and h-h 2D distributions are finalized and ready for projection onto $\Delta\varphi$ to extract the yields and widths of interest from the

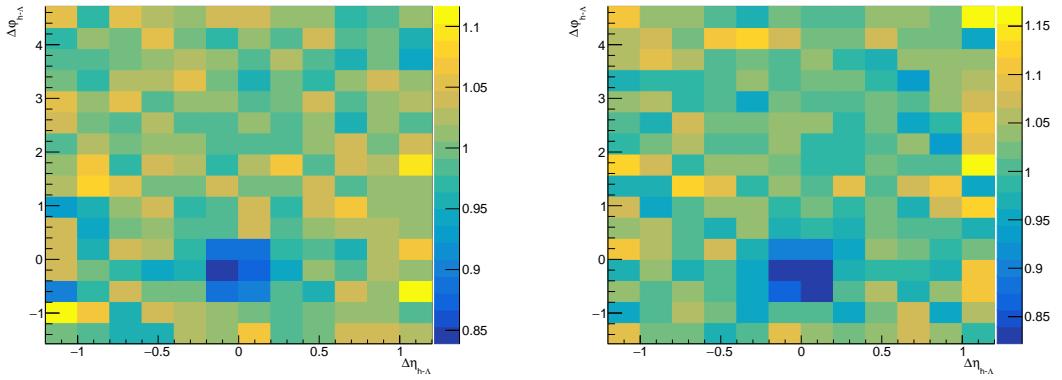


Figure 3.19: The $\epsilon_{\text{pair}}(\Delta\varphi, \Delta\eta)$ templates for the track merging correction in the lower ($1.5 < p_T < 2.5$ GeV/ c , left) and higher ($2.5 < p_T < 4.0$ GeV/ c , right) associated momentum bins. While it may be difficult to observe, the lower p_T bin has a minimum dip of around 0.84, whereas the higher p_T bin has a minimum dip of around 0.81, reflecting the p_T dependence discussed in this section.

previous chapter. However, there are still a number of systematic uncertainties to investigate and cross-checks required to ensure the validity of the final results.

Chapter Four: Systematic uncertainties and cross-checks

While it would be nice to produce exact results for the measurements presented in this thesis, uncertainties are inevitable. These uncertainties are usually grouped into two categories: **statistical uncertainties** and **systematic uncertainties**. Statistical uncertainties are those which arise from things like detector imprecisions or even the inherent statistical nature of the measurement itself. Calculating these uncertainties is *usually* straightforward, though it often involves well-motivated assumptions about the underlying probability distributions of the data. Determining systematic uncertainties, on the other hand, is a much more complicated process. To that end, the first section of this chapter goes into a large amount of detail about the calculation of systematic uncertainties for this analysis, starting with clearly defining what “systematic uncertainty” means and how it is quantified in the context of this thesis, followed by comprehensive investigations into each of the sources of uncertainty considered in this research, and ending with a summary of the final systematic uncertainties for each observable.

The second section of this chapter is dedicated to the cross-checks performed to ensure the validity of the analysis procedure. There are a variety of such checks, ranging from a general Monte Carlo closure test to more specific checks into physical biases that may be introduced by the analysis procedure.

4.1 Systematic uncertainties

“The treatment of systematic errors is often mishandled. This is due to lack of understanding and education, based on a fundamental ambiguity as to what is meant by the term”

—Roger Barlow, *Systematic Errors: Facts and Fictions* **BarlowCheck**

In every experimental analysis, choices must be made. These choices can be as simple as the selection of a particular data set, or as complex as choosing a fit function of fifteen parameters instead of three. In either case, these choices can have an effect

on the final results, which is usually quantified by the **systematic uncertainty**. Broadly speaking, the systematic uncertainty is a measure of the sensitivity of the final results to the choices made during the analysis procedure.

In this analysis, the procedure for estimating the systematic uncertainty from a certain choice is as follows:

1. Vary the choice in a way that is reasonable and justifiable
2. Measure the observable of interest after the variation
3. Quantify the effect of the variation on the observable by calculating the percent change from the original value of the observable
4. Vary the choice in a slightly different way, and repeat steps 2 and 3
5. Repeat steps 2 through 4 until all reasonable variations have been considered
6. Calculate the systematic uncertainty as the root mean square (RMS) of the percentages from step 3.

Note the usage of the words “reasonable” and “justifiable”, which seem to indicate that the systematic uncertainty is a subjective quantity. Indeed the process of obtaining the systematic uncertainty involves even more choices, like the choice of which choices to consider, the choices of how to vary the choices, and even the choice on how to quantify the uncertainty itself. In this sense, systematic uncertainty calculations are an art form, where the artist (analyzer) must use their creativity and best judgement to determine which choices are indeed reasonable. To that end, the list of sources of systematic uncertainties considered for this analysis is not exhaustive, but it is the best attempt at a comprehensive list of reasonable choices that affect the final results.

To provide more structure to this section, the analysis procedure is broken into the following components:

1. The generation of the $h-\Lambda$ and $h-h$ $\Delta\varphi$ distributions,
2. The extraction of the pairwise yields from the $\Delta\varphi$ distributions, and
3. The extraction of the near- and away-side widths from the fits of the $\Delta\varphi$ distributions,

which are used to separate this section into three subsections, one for each of these components. In each section, the sources of systematic uncertainties are described, followed by a Barlow analysis **BarlowCheck** to ensure the variations result in a statistically significant deviation from the nominal values. Finally, a summary of the final systematic uncertainties is provided.

4.1.1 $\Delta\varphi$ distribution generation

The sources of systematic uncertainties that affect the $\Delta\varphi$ distribution considered for this analysis are the following, in order of decreasing magnitude:

- Λ topological selection (h- Λ distribution only)
- Material budget
- Tracking efficiency
- Λ daughter PID cuts (h- Λ distribution only)
- Λ invariant mass signal region selection (h- Λ distribution only)
- Λ invariant mass sideband region selection (h- Λ distribution only)

As the Λ topological selection, material budget, and tracking efficiencies have been studied in detail in previous analyses using the same particle species and collision system **V0Reco1**, **V0Reco1**, **Lambda1**, **ALICEMatBud**, the systematic uncertainty associated with these sources is taken directly from these analyses and presented in Table ???. The uncertainties from these sources exhibit no multiplicity dependence, and a very small dependence on p_T . These uncertainties are also assumed to be independent of $\Delta\varphi$, although this is studied more thoroughly in Section ??.

Each of the other sources of systematic uncertainty is described in detail in the following sections.

4.1.1.1 Signal region selection

The nominal signal region for the Λ invariant mass is fairly wide, accounting for nearly 97% of the total Λ signal. However, the final result should not be heavily influenced by the choice of signal region so long as it is centered about the true Λ

Table 4.1: The systematic uncertainties for the $\Delta\varphi$ distributions which are not directly calculated in this thesis, instead taken from previous analyses using the same particle species and collision system **V0Reco1**, **V0Reco2**, **Lambda1**, **ALICEMatBud**. Each source of uncertainty is verified to be independent of multiplicity, but the Λ material budget and topological selection uncertainties exhibit a small dependence on p_T .

Source name	Lower p_T %	Higher p_T %
Λ topological selection	3.2%	3.0%
Λ material budget	1.1%	0.6%
Charged h tracking efficiency	3.5%	3.5%
Charged h material budget	negl.	negl.

mass. Furthermore, altering the signal region tests the validity of the signal scaling procedure outlined in Section ???. To investigate this, the signal region is varied in the ways presented in Table ???. The resulting $\Delta\varphi$ distributions and ratios to the nominal distribution for each signal region variation in each multiplicity and associated p_T bin are shown in Figures ?? (lower p_T) and ?? (higher p_T). The average deviation from the nominal distribution is around 2%, with no individual variation exceeding 5%. As no significant dependence on $\Delta\varphi$ is observed, the systematic uncertainty is calculated as the RMS of the percent change from each variation across the entire $\Delta\varphi$ range as opposed to calculating the RMS in each bin.

Table 4.2: The variations of the Λ invariant mass signal region considered for this analysis.

Variation name	Signal range (GeV/c^2)
Narrow	$1.108 < M_{p\pi} < 1.124$
Narrower	$1.112 < M_{p\pi} < 1.120$
Wide	$1.100 < M_{p\pi} < 1.132$
Wider	$1.096 < M_{p\pi} < 1.136$

4.1.1.2 Sideband region selection

The choice of sideband region also leaves a lot of room for reasonable variation: all that is required is that the region is 1) large enough to produce a smooth h-p π distribution with minimal statistical fluctuations and 2) close enough to the signal region that the p π pairs are kinematically similar to those in the background of the

signal region. As long as these requirements are met, the final result should not be very dependent on the choice of sideband region. To investigate the effects of changing the sideband region, the variations presented in Table ?? are considered. The measured $\Delta\varphi$ distributions and variation/nominal ratios for each sideband region variation in each multiplicity and associated p_T bin are shown in Figures ?? (lower p_T) and ?? (higher p_T). The result is even less deviation from the nominal distribution than the signal region variations, with the average deviation being closer to 1%. Again, no significant dependence is observed on $\Delta\varphi$, so the systematic uncertainty is calculated as the RMS of the percent change from each variation across every $\Delta\varphi$ bin.

Table 4.3: The variations of the Λ invariant mass sideband region considered for this analysis. Note that the “shifted left” sideband falls on the opposite (left) side of the signal region.

Variation name	Sideband range (GeV/c^2)
Narrow	$1.135 < M_{p\pi} < 1.145$
Wide	$1.135 < M_{p\pi} < 1.16$
Shifted left	$1.086 < M_{p\pi} < 1.098$
Shifted right	$1.14 < M_{p\pi} < 1.155$

4.1.1.3 Λ daughter particle identification

The Λ daughter particle identification (PID) cuts are chosen to be wide enough to ensure a high efficiency, but narrow enough to ensure a high purity. As the requirement for a higher purity should be offset by the subtraction of the combinatorial background, altering the PID cuts should only minimally effect the final $\Delta\varphi$ distributions. To study this, the PID cuts are varied in the ways presented in Table ???. The final $\Delta\varphi$ distributions and ratios to the nominal distribution for each PID cut variation in each multiplicity and associated p_T bin are shown in Figures ?? (lower p_T) and ?? (higher p_T). Requiring a signal in the TOF detector drastically reduces the Λ signal as the daughter pions are often heavily deflected by the magnetic field due to their lower p_T (again, $m_p/m_\pi \approx 7$, so most of the mother momentum belongs to the proton). This causes a large amount of statistical fluctuations in the corresponding $\Delta\varphi$ distribution, which is why the “require TOF” variation is inevitably

excluded after the Barlow check presented in Section ???. The other variations result in only around a 2% deviation from the nominal $\Delta\varphi$ distribution, on average.

Table 4.4: The variations of the Λ daughter PID cuts considered for this analysis. The “require TOF” variation requires a TOF hit for both the proton and pion, but maintains the nominal values for $|n\sigma_{\text{TPC, TOF}}|$.

Variation name	$ n\sigma_{\text{TPC, TOF}}^\pi $	$ n\sigma_{\text{TPC, TOF}}^p $
Narrow	< 1.8	< 1.2
Wide	< 4.2	< 2.8
Require TOF	< 3.0	< 2.0

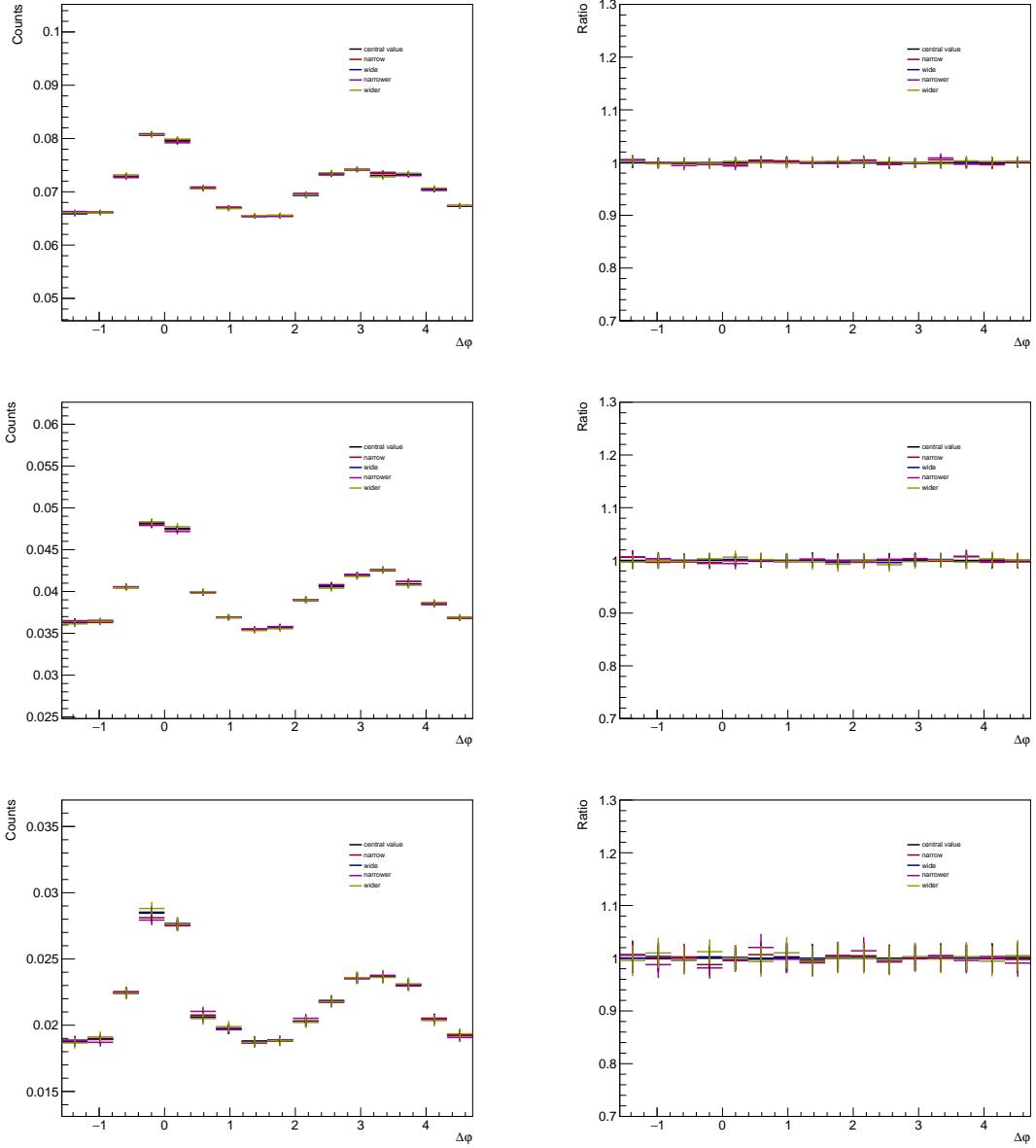


Figure 4.1: The $h-\Lambda$ $\Delta\varphi$ distributions within the 0-20% (top), 20-50% (middle) and 50-80% (bottom) multiplicity bins in the lower associated p_T bin for each of the signal region variations (left) with the ratios to the nominal distribution (right).

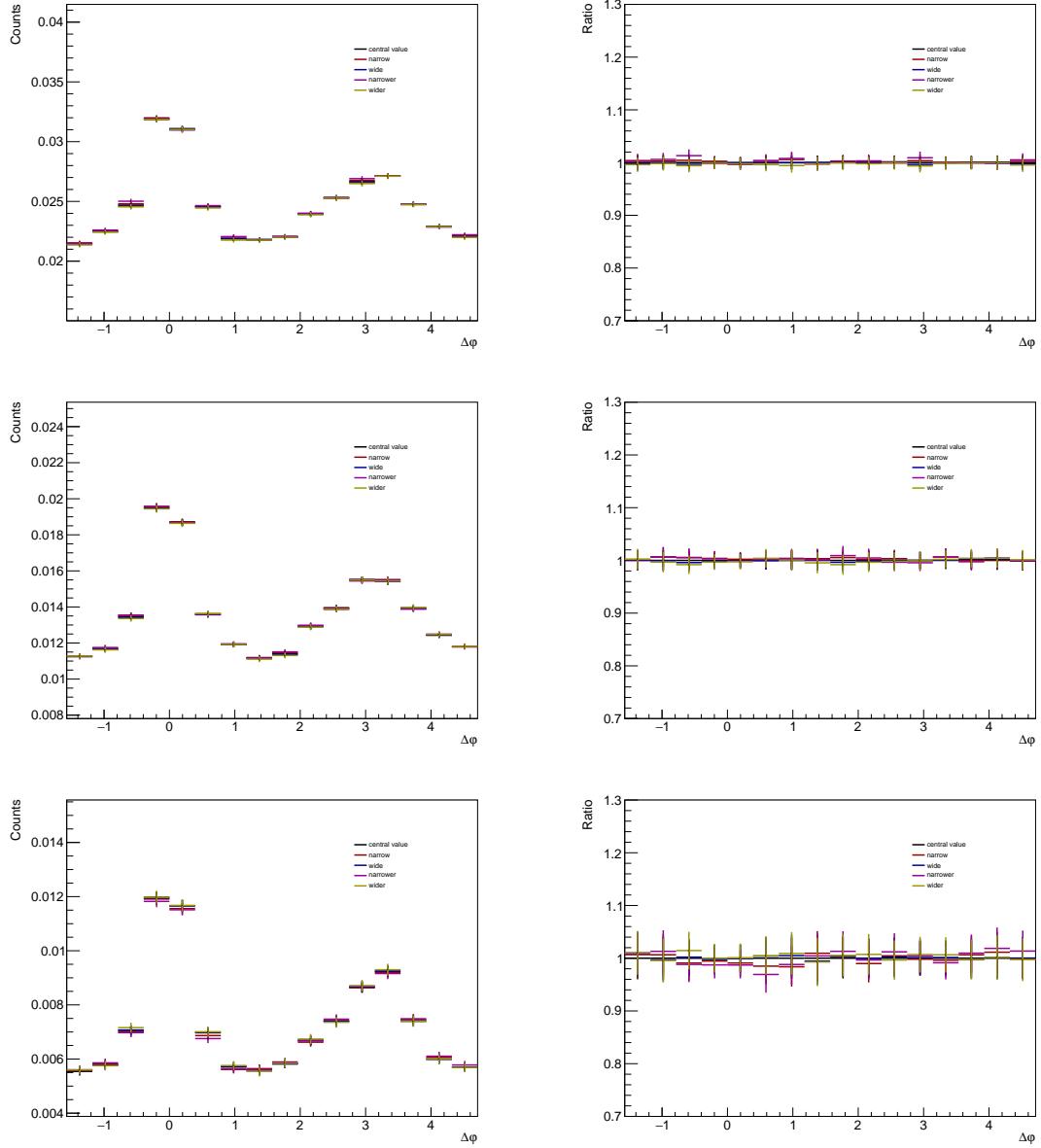


Figure 4.2: The $h-\Lambda$ $\Delta\varphi$ distributions within the 0-20% (top), 20-50% (middle) and 50-80% (bottom) multiplicity bins in the higher associated p_T bin for each of the signal region variations (left) with the ratios to the nominal distribution (right).

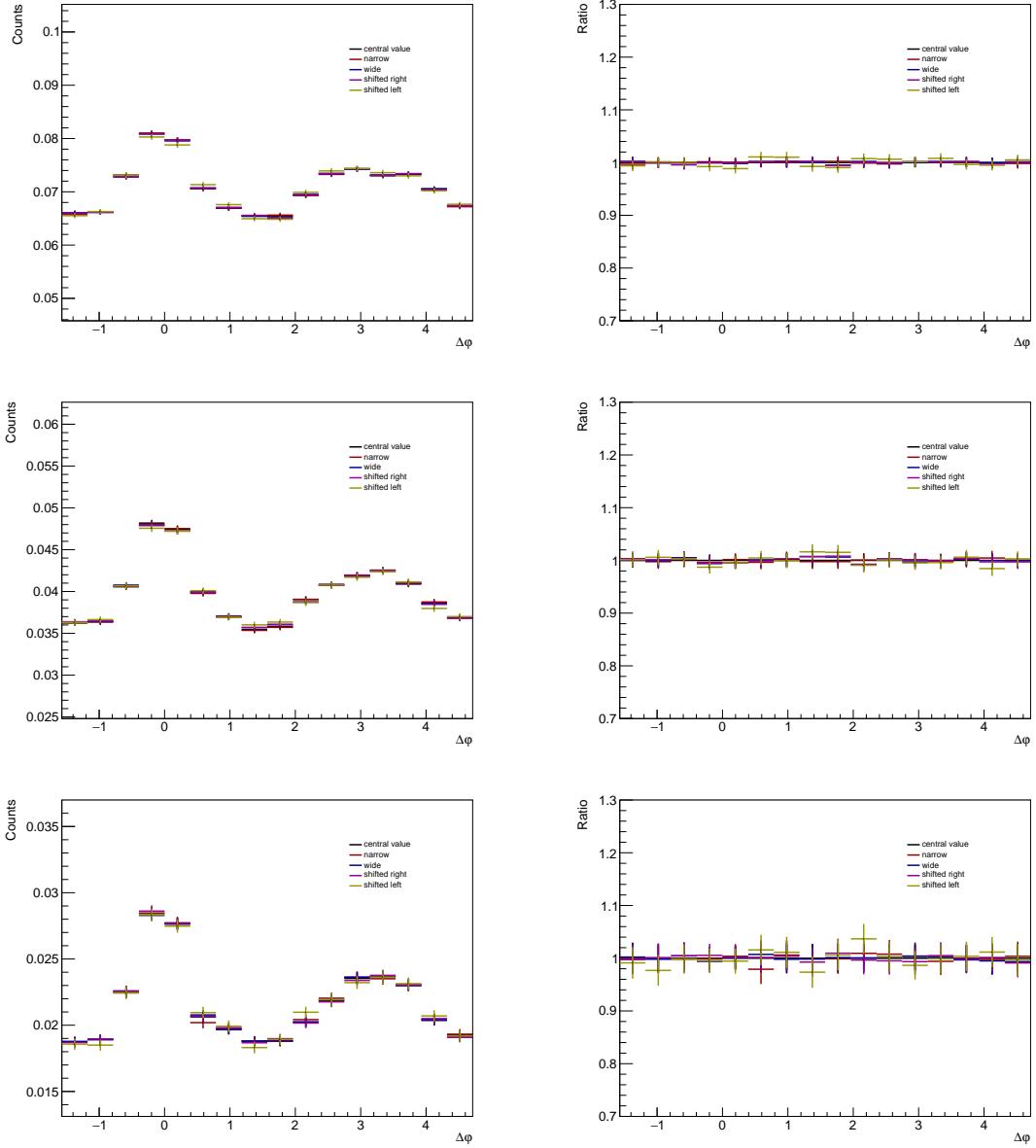


Figure 4.3: The $h-\Lambda$ $\Delta\varphi$ distributions within the 0-20% (top), 20-50% (middle) and 50-80% (bottom) multiplicity bins in the lower associated p_T bin for each of the sideband region variations (left) with the ratios to the nominal distribution (right).

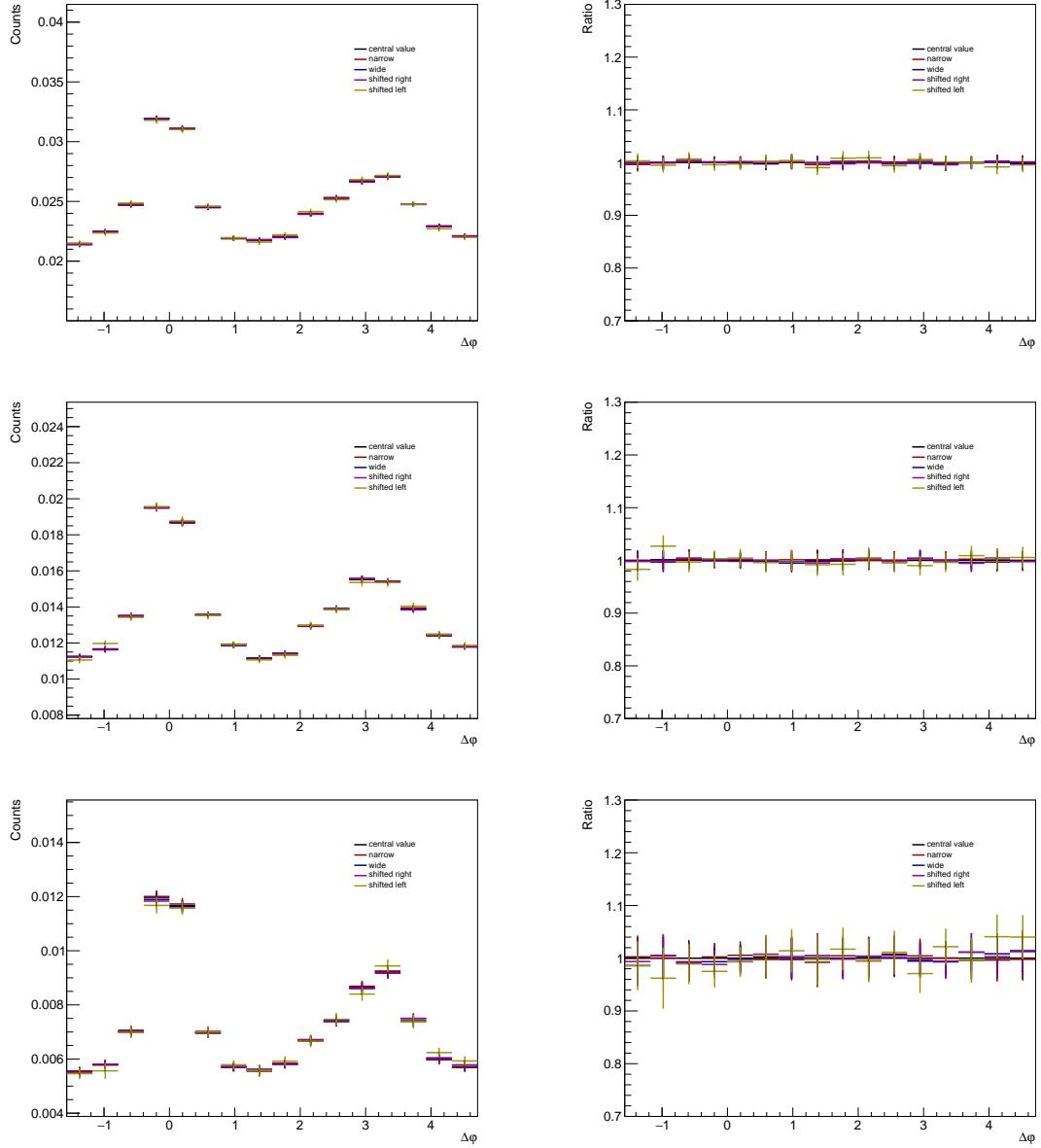


Figure 4.4: The $h-\Lambda$ $\Delta\varphi$ distributions within the 0-20% (top), 20-50% (middle) and 50-80% (bottom) multiplicity bins in the higher associated p_T bin for each of the sideband region variations (left) with the ratios to the nominal distribution (right).

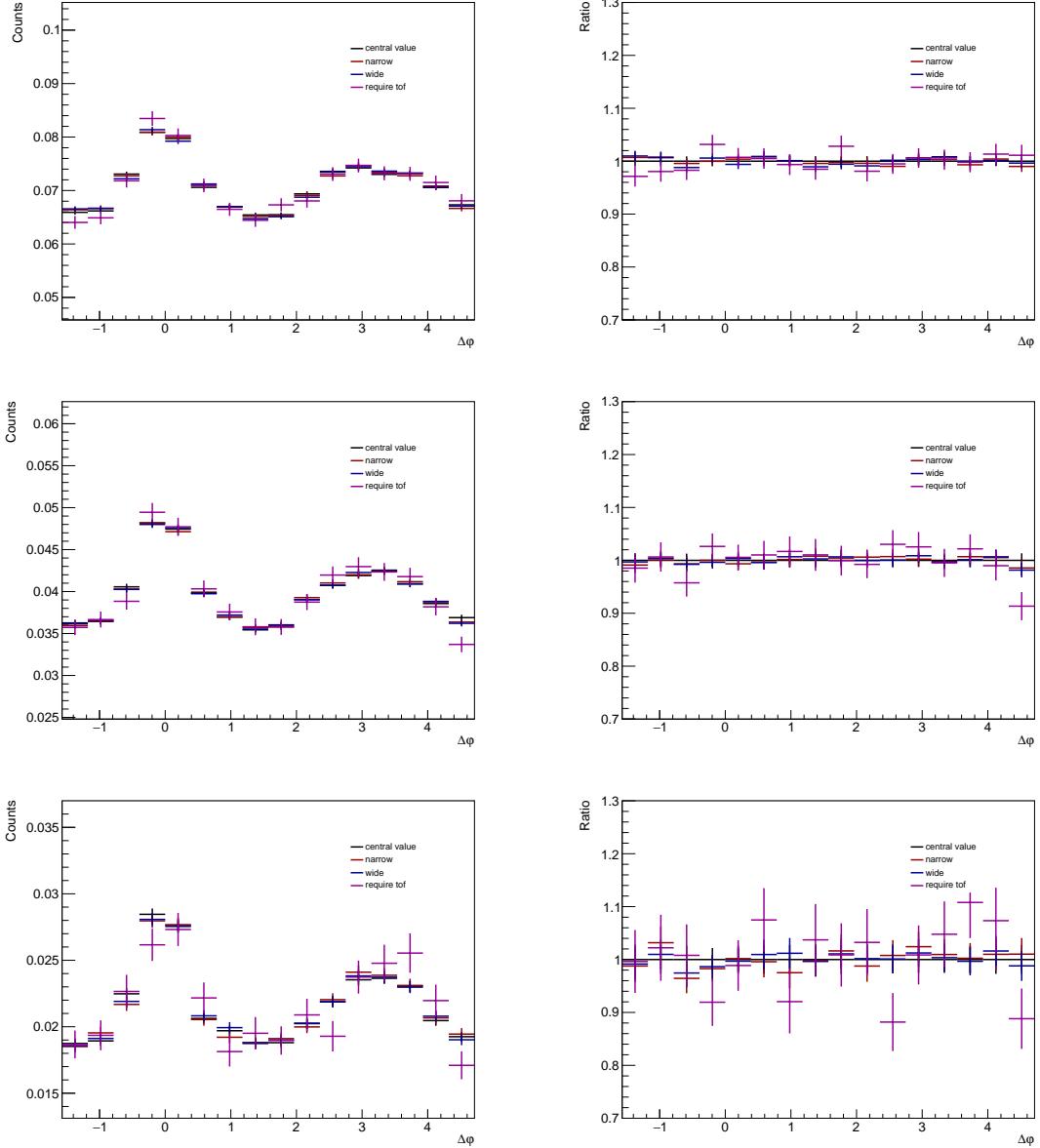


Figure 4.5: The $h-\Lambda$ $\Delta\varphi$ distributions within the 0-20% (top), 20-50% (middle) and 50-80% (bottom) multiplicity bins in the lower associated p_T bin for each of the PID cut variations (left) with the ratios to the nominal distribution (right).

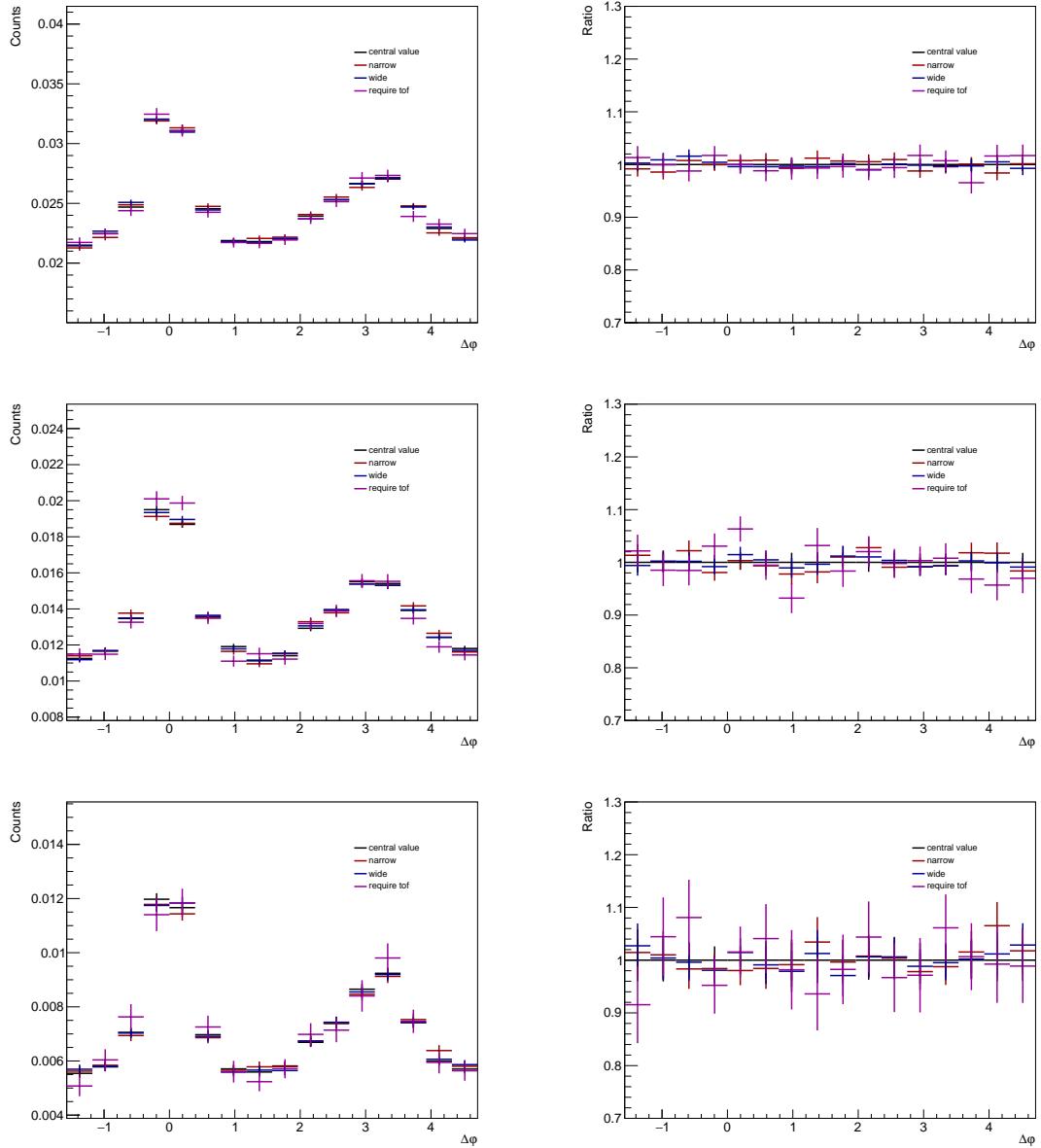


Figure 4.6: The $h-\Lambda$ $\Delta\varphi$ distributions within the 0-20% (top), 20-50% (middle) and 50-80% (bottom) multiplicity bins in the higher associated p_T bin for each of the PID cut variations (left) with the ratios to the nominal distribution (right).

4.1.1.4 Barlow check for $\Delta\varphi$ distribution generation

Due to statistical fluctuations, it may be the case that some variations result in a statistically insignificant deviation from the nominal $\Delta\varphi$ distribution. In such cases, the variation should not be considered in the final systematic uncertainty calculation. To determine which variations give statistically significant deviations, a Barlow check **BarlowCheck** is performed. For each $\Delta\varphi$ bin, the following quantity is calculated:

$$N\sigma_{RB} := \frac{y_{\text{var.}} - y_{\text{nom.}}}{\sqrt{|\sigma_{\text{var.}}^2 - \sigma_{\text{nom.}}^2|}}, \quad (4.1)$$

where $y_{\text{var.}}$ and $\sigma_{\text{var.}}$ are the measured yield and statistical uncertainty for the variation, and $y_{\text{nom.}}$ and $\sigma_{\text{nom.}}$ are the yield and statistical uncertainty for the nominal value.

To determine whether a given variation should be excluded, the number $\Delta\varphi$ bins that have $|N\sigma_{RB}| < 1$ is counted. If this is the majority of the bins (across all multiplicity and associated p_T ranges), the variation is excluded from the systematic calculation. Example plots of $N\sigma_{RB}$ for each variation of the signal, sideband and PID cuts are shown in Figure ???. The red lines represent $N\sigma_{RB} = \pm 1$.

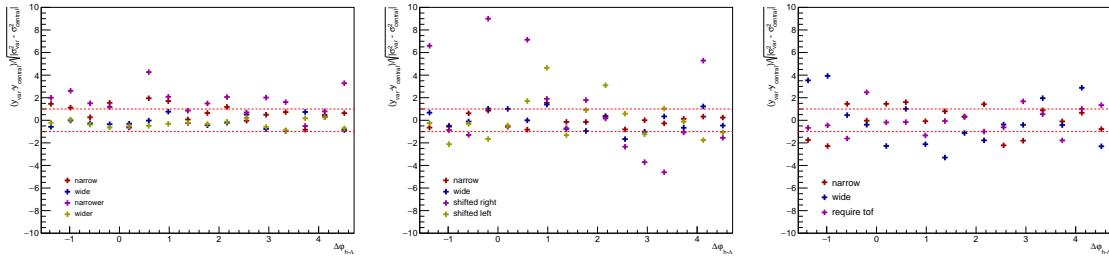


Figure 4.7: Barlow check for the signal (left), sideband (middle), and PID (right) variations in the 0-20% multiplicity bin. The red lines represent $N\sigma_{RB} = \pm 1$, and if the majority of the points fall within the red lines (across all $\Delta\varphi$, multiplicity and p_T bins), they are excluded from the systematic uncertainty calculation.

As a result of the check, the following variations are excluded from the final systematic uncertainty calculation:

- Signal: wide, wider
- Sideband: wide, narrow

- PID: require TOF

These exclusions are not so surprising. As the nominal signal region is already fairly wide, making it wider does not significantly change the $\Delta\varphi$ distribution. Similarly, the initial sideband region falls fairly close to the signal region. So long as there are enough statistics in the corresponding sideband $h\text{-}\Lambda$ distribution, changing its width should not affect the $\Delta\varphi$ distribution in a meaningful way. It also appears that requiring a TOF hit introduces large statistical errors, which dominate the denominator in Equation ??.

4.1.1.5 $\Delta\varphi$ distribution systematics, summarized

The final systematic errors (after the Barlow check) from the $h\text{-}\Lambda$ $\Delta\varphi$ distribution generation for each multiplicity bin and p_T bin are shown in Table ???. The total systematic uncertainty is calculated by adding each systematic error in quadrature. This table is consolidated into plots showing the systematic errors for each multiplicity bin and p_T bin, which are presented in Figure ???. As the systematic uncertainties associated with the generation of the dihadron $\Delta\varphi$ distributions are only from the tracking efficiency presented in Table ???, they are not plotted in this section.

Table 4.5: The final systematic uncertainties (in percentages) from the $h\text{-}\Lambda$ $\Delta\varphi$ distribution generation for each multiplicity and associated p_T bin.

Mult. and p_T bin	Sig.	Sideband	PID	Topo. sel.	Mat. bud.	Total
0-20%, low p_T	0.36	0.53	0.64	3.2	1.1	3.3
20-50%, low p_T	0.35	0.67	0.65	3.2	1.1	3.4
50-80%, low p_T	0.76	1.1	1.4	3.2	1.1	3.8
0-20%, high p_T	0.42	0.42	0.76	3.0	0.6	3.2
20-50%, high p_T	0.4	0.71	1.2	3.0	0.6	3.3
50-80%, high p_T	1.1	1.6	2.0	3.0	0.6	4.1

The total systematic error is observed to be mostly p_T -independent. However, there appears to be a slight correlation between the systematic uncertainty and multiplicity, with the 0-20% bin exhibiting lower uncertainties than the 50-80% bin across both p_T ranges. This can become problematic when investigating the multiplicity dependence of observables extracted from the $\Delta\varphi$ distributions, as the fraction of the systematic uncertainty which is directly correlated with multiplicity should not be considered when measuring multiplicity-dependent trends like slopes and percent

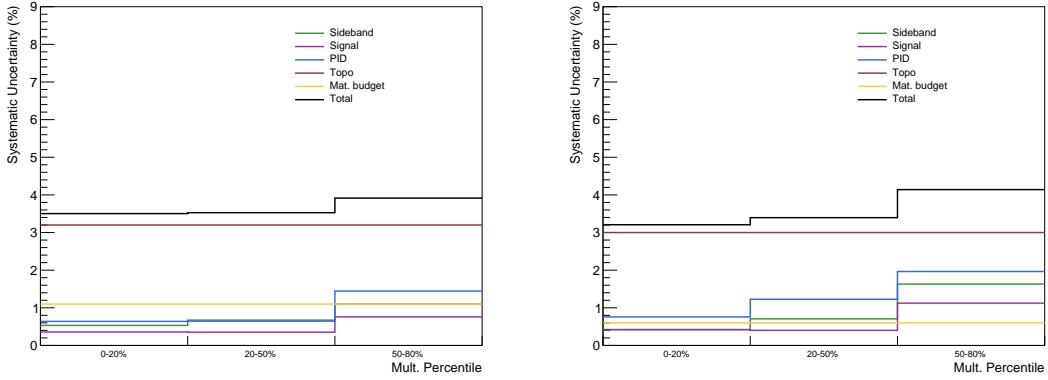


Figure 4.8: A visual depiction of the final systematic errors for the $h\text{-}\Lambda \Delta\varphi$ distributions for each multiplicity bin in the low (left) and high (right) associated p_T bins. The total systematic error is shown in black.

changes. Because of this, the fraction of the systematic uncertainty which is uncorrelated with multiplicity is approximated using

$$\sigma_{\text{uncor},i}^2 = \sum_{\text{vars}} (R_{\text{var},i} - 1)^2, \quad (4.2)$$

where

$$R_{\text{var},i} = \left(\frac{y_{\text{var},i}}{y_{\text{nom},i}} \right) / \left(\frac{y_{\text{var}}^{\text{MB}}}{y_{\text{nom}}^{\text{MB}}} \right), \quad (4.3)$$

where “i” refers to the i th multiplicity bin, and “MB” refers to the min-bias (multiplicity-integrated) results. The deviations of $R_{\text{var},i}$ from unity quantify how the deviations in multiplicity bin i differ from those in the MB sample. $\sigma_{\text{uncor},i}$ is computed for each $\Delta\varphi$ bin, then the RMS is taken across all $\Delta\varphi$ bins to obtain the final multiplicity-uncorrelated portion of the systematic errors. The results for each p_T bin are shown in Figure ???. These systematic errors are only used when quantifying the multiplicity dependence of an observable extracted from the $\Delta\varphi$ distributions.

4.1.2 Yield extraction

One of the largest sources of systematic uncertainty of this analysis corresponds to the different techniques that can be used to extract the yields in the near-side jet, away-side jet and underlying event from the $\Delta\varphi$ distributions. As mentioned in Section ??,

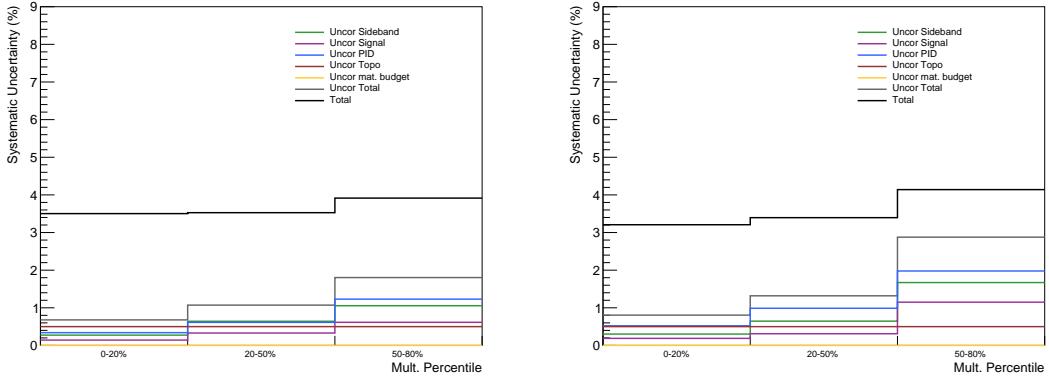


Figure 4.9: Visual depiction of the multiplicity-uncorrelated systematic errors for the $h\Lambda$ $\Delta\varphi$ distributions for each multiplicity bin in the low (left) and high (right) associated p_T bins, along with the total systematic error shown in black.

the equations for extracting these yields are

$$Y_{near} = \int_{-\pi/2}^{\pi/2} \left(\frac{dN}{d\Delta\varphi} - U(\Delta\varphi) \right) d\Delta\varphi, \quad Y_{away} = \int_{\pi/2}^{3\pi/2} \left(\frac{dN}{d\Delta\varphi} - U(\Delta\varphi) \right) d\Delta\varphi \quad (4.4)$$

$$Y_{UE} = \int_{-\pi/2}^{3\pi/2} U(\Delta\varphi) d\Delta\varphi, \quad (4.5)$$

where $\frac{dN}{d\Delta\varphi}$ is the $\Delta\varphi$ distribution and $U(\Delta\varphi)$ is the underlying event fit. As the $\Delta\varphi$ distribution is present in these equations, all of the previous variations concerning the generation of this distribution must be considered. However, these equations also naturally introduce two new categories of systematic uncertainty: those associated with the underlying event fit, and those associated with the integration of the $\Delta\varphi$ distribution. Both of these categories will be discussed in detail in the following sections.

4.1.2.1 Underlying event fit techniques

As the underlying event term $U(\Delta\varphi)$ is present in every yield extraction equation above, any changes in the underlying event fitting procedure will affect the final yield measurements. To maintain compatibility with previous analyses (specifically for the dihadron correlations), the nominal underlying event fit is a straight line to the average of the $\Delta\varphi$ distribution in the ranges $[-\frac{\pi}{2}, -\frac{\pi}{4}] \cup [\frac{\pi}{4}, \frac{5\pi}{8}] \cup [\frac{11\pi}{8}, \frac{3\pi}{2}]$. These ranges were initially chosen as there is expected to be little-to-no contamination from the jet

components in each range. However, to investigate the effect the UE fitting procedure may have on the final yields, the following alternative methods were considered:

1. Straight line fit in a more restricted range, specifically $[-\frac{\pi}{2}, -\frac{3\pi}{8}) \cup [\frac{3\pi}{8}, \frac{5\pi}{8}) \cup [\frac{11\pi}{8}, \frac{3\pi}{2})$
2. Straight line fit using the Zero Yield At Minimum (ZYAM) technique, where the underlying event line is set to the minimum of the $\Delta\varphi$ distribution
3. Sinusoidal fit which includes a non-zero v_2 contribution

The first two techniques are similar enough to the nominal technique that they will not be explicitly shown in this section. Restricting the range of the flat fit region results in deviations from the nominal procedure of around 2%, whereas the ZYAM technique gives much larger deviations at about 15%. Ultimately the ZYAM procedure is not included in the final systematics calculation due to a physical incompatibility, whereby the presence of v_2 in the higher multiplicity $\Delta\varphi$ distributions causes the ZYAM procedure to massively underestimate the underlying event contribution.

Including a non-zero v_2 contribution is a much more involved procedure and requires new machinery to be developed, so it will be described in detail in the next section.

4.1.2.2 Including a non-zero v_2 contribution

All of the “straight line” UE fitting techniques are based on the flatness assumption of the non-jet part of the correlation in $\Delta\varphi$. This means that the dijet axis direction does not affect the non-jet particle distribution’s overall shape within an event. However, as mentioned in Section ??, previous Pb–Pb, p–Pb and even pp collision studies have shown that the QGP’s collective flow components (v_1 , v_2 , etc.) influence the phase-space distribution of particles within an event. Using Fourier decomposition, the $\Delta\varphi$ distributions on an event-by-event basis can be written as

$$\frac{dN}{d\Delta\varphi} = a_0 + \sum_{n=1}^{\infty} 2a_n \cos(n\Delta\varphi), \quad (4.6)$$

where a_n are the Fourier coefficients. Surprisingly, these coefficients have been shown **Justin108**, **Justin109**, **Justin110** to be related to the collective flow coefficients v_n via

$$v_n = \frac{a_n}{a_0}. \quad (4.7)$$

Table 4.6: v_2 values used in this analysis for each associated p_T bin. The values were calculated as the weighted average of published p_T -differential v_2 measurements with the published p_T spectra, taken across the entire associated p_T range.

$p_T^{\text{assoc.}}$	$v_2^{\text{trig.}}$	$v_2^{\text{assoc. h}}$	$v_2^{\text{assoc.}\Lambda}$
1.5 - 2.5	0.092	0.100	0.075
2.5 - 4.0	0.092	0.119	0.137

This means that even without reconstructing the reaction plane within a specified event, the effects of collective flow are present in the $\Delta\varphi$ distributions. This manifests in the correlation distributions as an underlying event which is not flat with respect to $\Delta\varphi$, but rather sinusoidal. While this is in direct conflict to the initial assumption of a flat underlying event, this nominal choice was made to maintain compatibility with previous measurements of dihadron yields using correlation techniques, which also assume a flat UE in $\Delta\varphi$.

As the v_2 or “elliptic flow” coefficient is the most dominant of the collective flow coefficients measured in p–Pb collisions **Justin111** in the p_T ranges for this analysis, it is the only one considered. Furthermore, the v_2 coefficients are exceedingly difficult to determine, with fully published papers solely dedicated to measuring the v_2 for different particle species and collision systems. Luckily, these coefficients have been measured by ALICE in p–Pb collisions for both charged hadrons and Λ baryons across a wide range of p_T **ALICEv2·1**, **ALICEv2·2**. As the p_T binning in this analysis is much wider, the weighted average

$$v_2^{\text{avg}} = \frac{\int_{p_{T,\min}}^{p_{T,\max}} v_2(p_T) \frac{dN}{dp_T} dp_T}{\int_{p_{T,\min}}^{p_{T,\max}} \frac{dN}{dp_T} dp_T}, \quad (4.8)$$

is used, where $p_{T,\min}$ and $p_{T,\max}$ are the minimum and maximum values of p_T in the bins from this analysis (namely $1.5 - 2.5$ and $2.5 - 4.0$ GeV/c). The $v_2(p_T)$ values for charged hadrons and Λ baryons are taken from **ALICEv2·1**, and dN/dp_T is taken from the published p_T spectra for charged hadrons and Λ baryons from **PtSpectra**, plots of which can be seen in Figures ?? (v_2) and ?? (p_T spectra). The values of v_2^{avg} for each p_T bin are shown in Table ???. Note that the trigger v_2 remains the same, as the trigger p_T range is fixed for this analysis. The Λ v_2 is markedly higher than the charged hadron v_2 , which ultimately manifests itself as a larger deviation from the nominal UE fit when compared to the dihadron case. Unfortunately there are few multiplicity-dependent measurements of the v_2 coefficients for identified particle

species. Because of this, the v_2 values from Table ?? are used only in the 0-20% multiplicity bin, with the v_2 values for the 20-50% and 50-80% multiplicity bins taken as 0.85 and 0.50 times the 0-20% value, respectively.

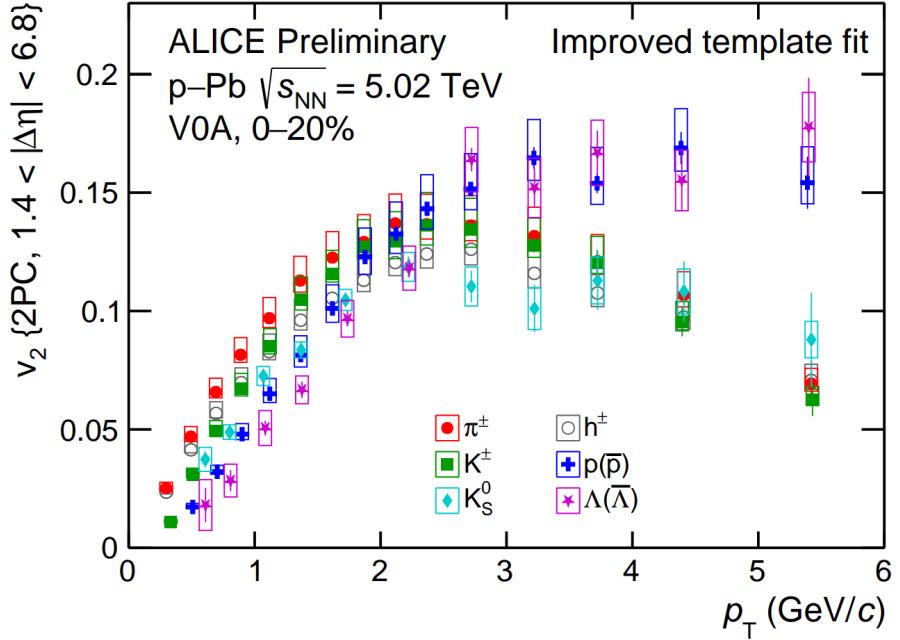


Figure 4.10: The v_2 values for identified hadrons as a function of p_T , taken from **ALICEv2.1**.

Using these v_2 values, the underlying event is estimated by fitting the function

$$U_{v_2}(\Delta\varphi) = A \times (1 + 2v_2^{\text{trig.}} v_2^{\text{assoc.}} \cos(2\Delta\varphi)) \quad (4.9)$$

in the ranges $-\pi/2 < \Delta\varphi < -1$ and $1 < \Delta\varphi < +\pi/2$, where little jet contribution is expected. The underlying event **pedestal** A is allowed to vary during the fit, but the v_2 values are fixed. Examples of h- Λ and h-h $\Delta\varphi$ distributions with the UE fit using this procedure are shown in Figure ??.

The validity of this procedure can be tested by examining the $\Delta\varphi$ distributions at large $\Delta\eta$, where the near-side jet component is minimal, leaving just the UE at small $\Delta\varphi$ ¹. In fact, this procedure is often used to determine the v_2 coefficients in the first place. In this case, however, it will just be used to serve as a sanity check for both the fitting procedure and the fixed v_2 coefficients from ??.

¹At large $\Delta\varphi$, the away-side ridge is still present.

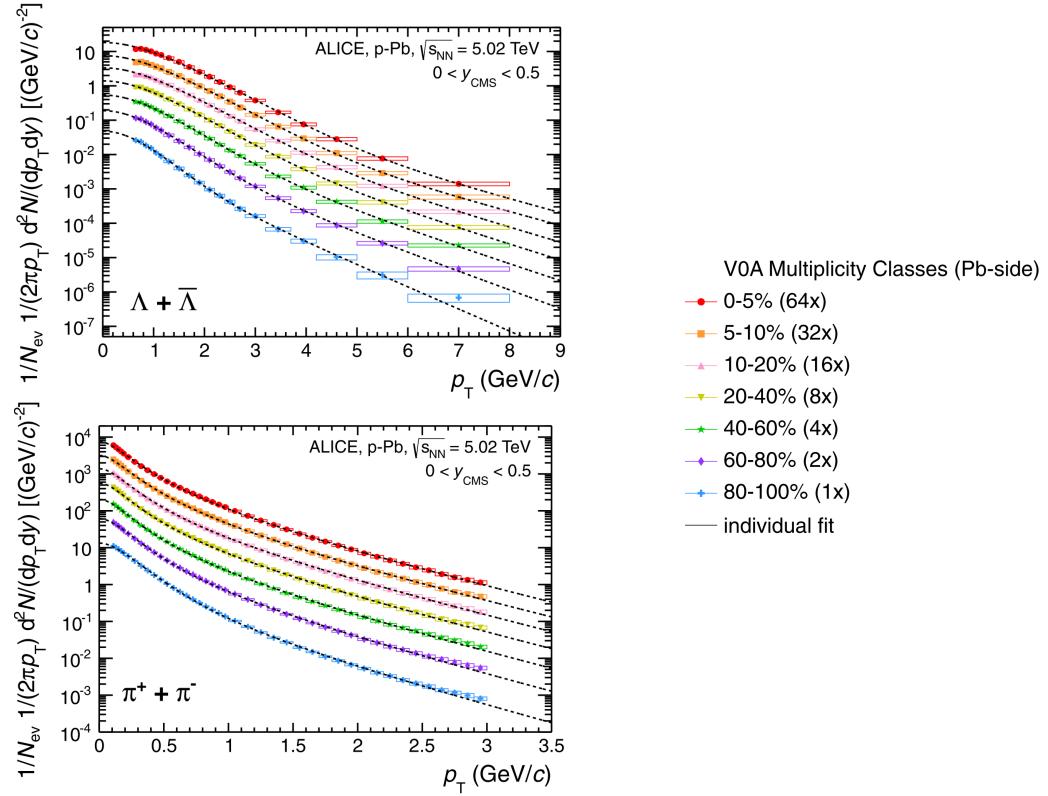


Figure 4.11: The published **ALICEMatBud** p_T spectra for Λ baryons (top) and charged hadrons (\approx pions) (bottom), used to compute the weighted average of the v_2 coefficients across the wide momentum bins used in this analysis.

matches the near-side of the $\Delta\varphi$ distribution at large $\Delta\eta$, then the v_2 coefficients and fitting procedure are likely valid. Examples of the $\Delta\varphi$ distributions with $|\Delta\eta| > 1.4$ and $|\Delta\eta| < 1.2$ showing the v_2 -based UE fit can be seen in Figure ???. These are generated in the highest multiplicity and momentum bins, where the effects of the v_2 contribution are maximal. Both the h- Λ and h-h $\Delta\varphi$ distributions show good agreement between the UE fit and the data at small $\Delta\varphi$ and large $\Delta\eta$, where the near-side jet peak and away-side ridge are no longer present, pointing to the validity of the v_2 -based UE fitting procedure.

The effects of including v_2 has on the extracted h- Λ and h-h yields in each region is not at all obvious at first glance. For the most central collisions, the inclusion of v_2 results in nearly a 5% decrease for the jet-like yields when compared to the nominal technique. This can mostly be seen in Figure ??, where the peaks of the sinusoidal fit achieve their maxima within the near- and away-side components of

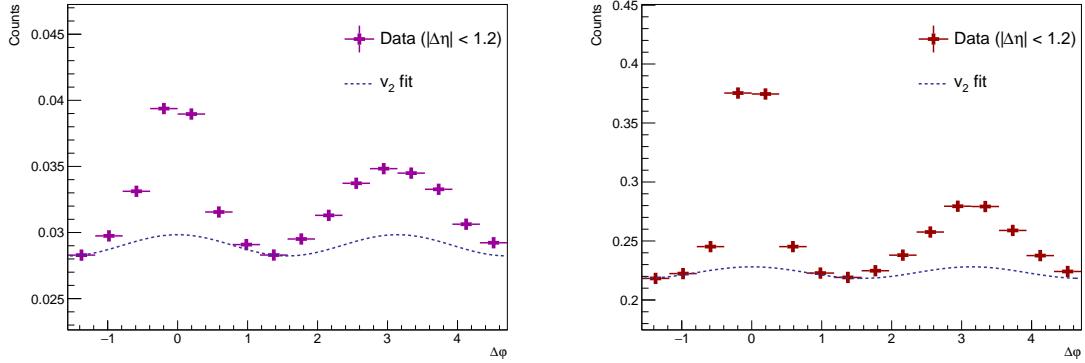


Figure 4.12: Examples of the underlying event fit using the v_2 -based procedure for the $h-\Lambda$ (left) and $h-h$ (right) $\Delta\varphi$ distributions in the 0-20% multiplicity bin in the higher associated p_T bin.

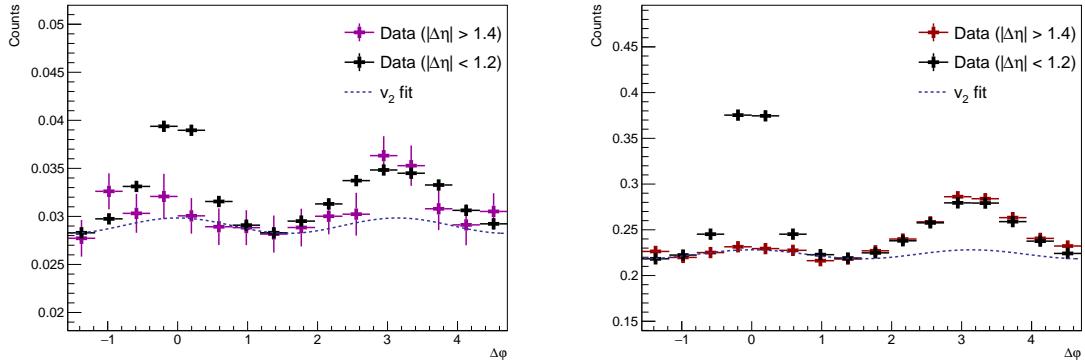


Figure 4.13: The $h-\Lambda$ (left) and $h-h$ (right) $\Delta\varphi$ distributions in the 0-20% multiplicity bin and higher p_T bin at small and large values of $\Delta\eta$, with the UE fit using the v_2 -based procedure shown in blue. The fits are in good agreement with data in both cases.

the jet, causing the overall yields to be lower than those obtained by the flat UE assumption. However, at lower multiplicities (20-50%, 50-80%), the extracted $h-\Lambda$ and $h-h$ jet-like yields actually exhibit a slight increase of around 5% in their extracted yields when compared to those measured using the nominal UE fit. This is due to the variation of the pedestal A in Equation ?? during the fit, which results in a smaller pedestal value than the nominal fit in these multiplicity ranges. The extracted underlying event yield never deviates by more than 3% from the yield obtained using

the nominal procedure for all multiplicity and momentum bins for both the h- Λ and h-h cases.

4.1.2.3 Integration procedures

The general yield-extraction equation

$$Y_{\Delta\varphi} = \int_{\Delta\varphi_1}^{\Delta\varphi_2} \left(\frac{dN}{d\Delta\varphi} - U(\Delta\varphi) \right) d\Delta\varphi \quad (4.10)$$

leaves some room for interpretation. Obviously the $\Delta\varphi$ distributions shown thus far are in some way related to $dN/d\Delta\varphi$, but integrals prefer continuous integrands, which the $\Delta\varphi$ distributions are clearly not as they have finitely many (16) bins. Furthermore, there is nothing explicitly preventing

$$\frac{dN}{d\Delta\varphi} < U(\Delta\varphi), \quad (4.11)$$

possibly resulting in a *negative* yield, which is clearly unphysical. There are a few ways to alleviate these issues, which are discussed in this section.

For all of the yield extraction procedures discussed thus far, the usage of the integration symbol in Equation ?? is *slightly* misleading: the yields are actually calculated by summing the bin contents of the $\Delta\varphi$ distribution in the specified range, and subtracting off the value of $U(\Delta\varphi)$ at the center of each bin. To be more explicit, the yields are calculated as

$$Y_{\Delta\varphi} = \sum_{i=L}^U \left(\frac{dN}{d\Delta\varphi_i} - U(\Delta\varphi_i) \right), \quad (4.12)$$

where L and U are the bin numbers of the Lower and Upper $\Delta\varphi$ bins in the specified range, $dN/d\Delta\varphi_i$ is the value of the correlation distribution in the i th $\Delta\varphi$ bin, and $U(\Delta\varphi)$ is the value of U in the center of the i th $\Delta\varphi$ bin.

Equation ?? provides an easy way to deal with the negative yield issue: if the value of $U(\Delta\varphi_i)$ is greater than the value of $dN/d\Delta\varphi_i$ in a given bin, the yield in that bin is set to zero. While this is not done for the nominal yield extraction procedure in this analysis, it is a completely reasonable technique and is therefore explored in the systematic uncertainty analysis. Using the flat UE assumption with the UE average taken in the nominal range, the results are relatively unsurprising: the yields

extracted using this procedure are strictly higher than those extracted using the main procedure where negative contributions are allowed. However, the deviations never exceed more than 3.5%, with the average deviation being around 2% for both the h- Λ and h-h cases.

Another way to address the lack of continuity in the $\Delta\varphi$ distributions is to fit these distributions with continuous functions, then use the corresponding fit for the integration in Equation ???. This thesis considers two such functions, which are presented in the following two sections.

4.1.2.4 The double Gaussian fit

There are a number of functions that may appear suitable to fit the $\Delta\varphi$ distributions, but given the Gaussian-like appearance of the near- and away-side jet components, a double Gaussian fit is a natural choice. The double Gaussian fit function is given by

$$f(\Delta\varphi) = U + A_{\text{NS}} e^{\frac{(\Delta\varphi - \mu_{\text{NS}})^2}{2\sigma_{\text{NS}}^2}} + A_{\text{AS}} e^{\frac{(\Delta\varphi - \mu_{\text{AS}})^2}{2\sigma_{\text{AS}}^2}} + A_{\text{NS}}^{\text{mirror}} e^{\frac{(\Delta\varphi - \mu_{\text{NS}} + 2\pi)^2}{2\sigma_{\text{NS}}^2}} + A_{\text{AS}}^{\text{mirror}} e^{\frac{(\Delta\varphi - \mu_{\text{AS}} - 2\pi)^2}{2\sigma_{\text{AS}}^2}}, \quad (4.13)$$

where A and μ are the amplitude and means of the Gaussian components, and the subscript “NS” (“AS”) refers to the near-side (away-side) jet component. The “mirror” terms are added to account for the 2π periodicity of the $\Delta\varphi$ distribution, and are required to obtain a convergent fit. The U term describes a flat underlying event, and is fixed to the average of the $\Delta\varphi$ distribution in the regions $[-\frac{\pi}{2}, -\frac{\pi}{4}] \cup [\frac{\pi}{4}, \frac{5\pi}{8}] \cup [\frac{11\pi}{8}, \frac{3\pi}{2}]$, as is done for the nominal UE determination procedure. Furthermore, the mean in the near-side gaussian (and corresponding mirror term) is fixed to 0 (2π) and the mean in the away-side guassian (and corresponding mirror term) is fixed to π ($-\pi$), leaving only the amplitudes and widths to vary freely. The double Gaussian fits to both the h- Λ and h-h correlation distributions for every multiplicity and momentum bin are shown in Figures ?? (h- Λ) and ???. The fits generally describe the data quite well, though an extreme amount of effort went in to ensuring the convergence of each fit due to an inordinate amount of instability.

The yields extracted using the double Gaussian fit are nearly identical to those extracted using the nominal bin-wise integration procedure, with deviations from the nominal procedure never exceeding 1% for either the h- Λ or h-h cases. This indicates two things:

1. The fits describe the data quite well, and
2. The differences between Equations ?? and ?? are mostly aesthetic, so long as (1) holds and the choice of $U(\Delta\varphi)$ is consistent.

As the deviations are so small, this procedure ends up excluded from the final systematic uncertainty calculation after the Barlow check in Section ??, but the fits are still used for the systematic studies involving the near- and away-side jet widths discussed in Section ??.

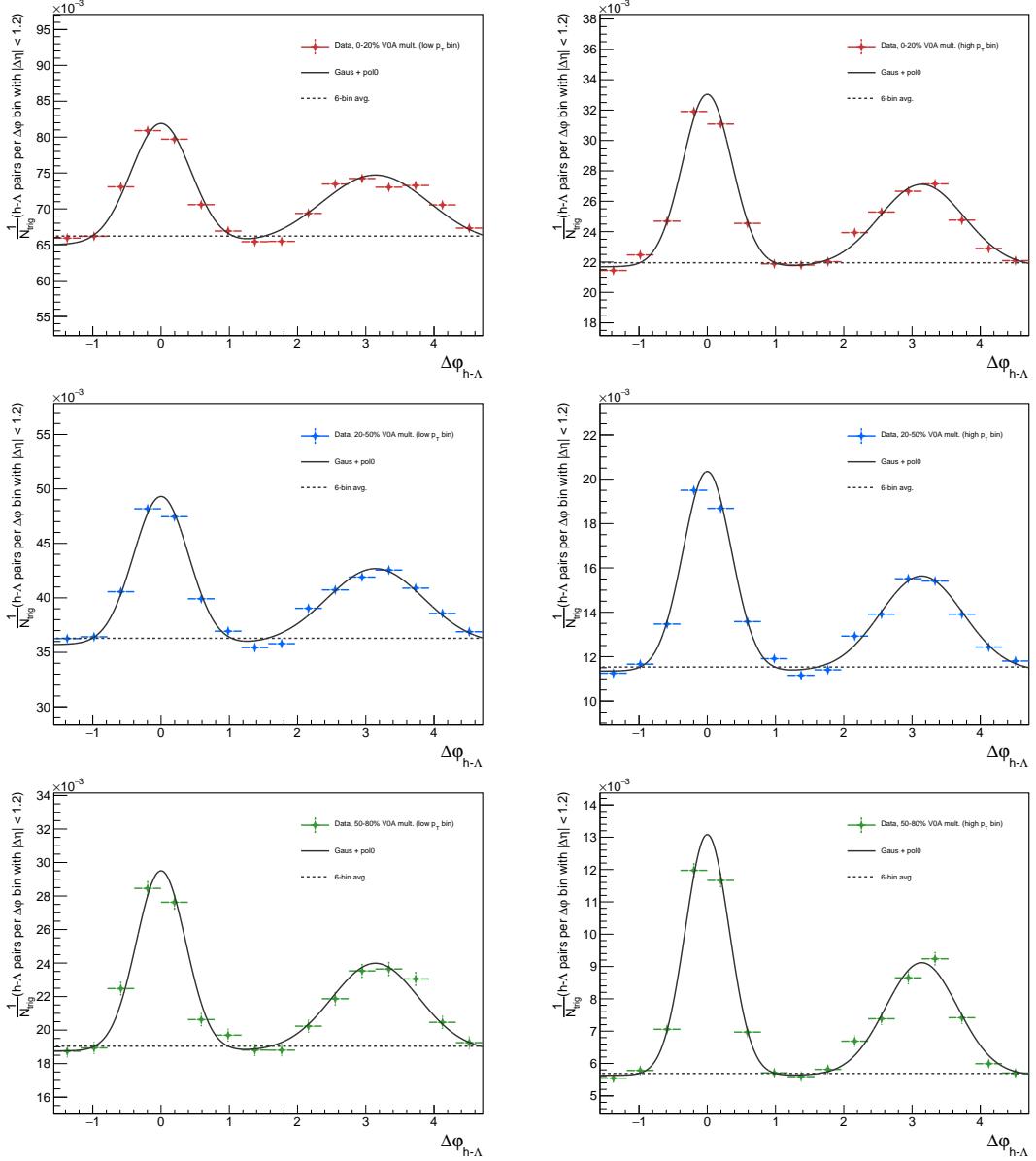


Figure 4.14: The final per-trigger h- Λ $\Delta\varphi$ correlation distributions with the full double Gaussian fit for the 0-20% (top), 20-50% (middle) and 50-80% (bottom) multiplicity bins in the lower (left) and higher (right) associated p_T bins. The UE fit is also shown as a dashed line.

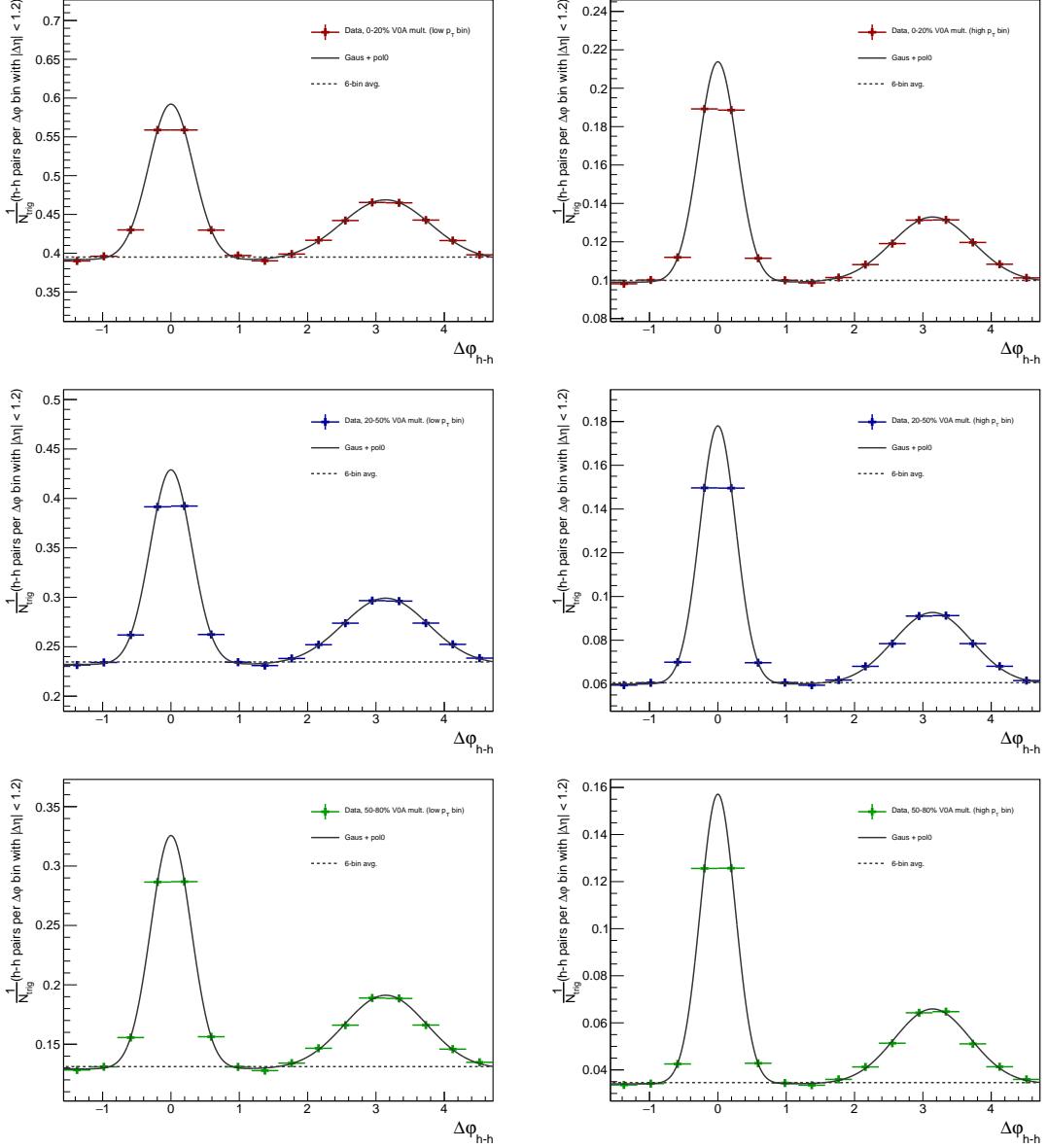


Figure 4.15: The final per-trigger h-h $\Delta\varphi$ correlation distributions with the full double Gaussian fit for the 0-20% (top), 20-50% (middle) and 50-80% (bottom) multiplicity bins in the lower (left) and higher (right) associated p_T bins. The UE fit is also shown as a dashed line.

4.1.2.5 The von Mises fit

While only briefly mentioned in Section ?? in the context of extracting the near- and away-side jet widths, the von Mises distribution is another natural choice for fitting the $\Delta\varphi$ distributions due to its combined Gaussian-like behaviour while naturally exhibiting 2π -periodicity (which was “forced” onto the double Gaussian fit via the mirror terms). As a reminder, the von Mises fit function is given by

$$f(\Delta\varphi) = U(\Delta\varphi) + \frac{A_{\text{NS}}}{2\pi I_0(k_{\text{NS}})} e^{k_{\text{NS}} \cos(\Delta\varphi - \mu_{\text{NS}})} + \frac{A_{\text{AS}}}{2\pi I_0(k_{\text{AS}})} e^{k_{\text{AS}} \cos(\Delta\varphi - \mu_{\text{AS}})}, \quad (4.14)$$

where A and μ are as they were in Equation ??, and k is a measure of the collimation of the distribution, which is inversely related to the width through

$$\sigma = \sqrt{-2 \ln \frac{I_1(\kappa)}{I_0(\kappa)}}. \quad (4.15)$$

In these equations, I_n refers to the modified Bessel function of the n th kind. Note that the $U(\Delta\varphi)$ has been given explicit $\Delta\varphi$ dependence, as this fit function does not require the UE component to be flat with respect to $\Delta\varphi$ in order to converge nicely.

During the fitting procedure, the means μ_{NS} and μ_{AS} are again fixed to 0 and π , respectively, and the $U(\Delta\varphi)$ term is fixed to the function obtained by fitting the UE which includes a non-zero v_2 contribution, as described in Section ???. While the fitting of this function is much more stable than the aforementioned double Gaussian function—possibly allowing for the variation of U during the fit—the term is ultimately fixed due to an interesting feature of the von Mises distribution, which is discussed in ???. The von Mises fits to both the h- Λ and h-h correlation distributions for every multiplicity and momentum bin are shown in Figures ?? (h- Λ) and ???. Again, the fits describe the data very well. Furthermore, the fits are extremely stable, which is a welcome change from the double Gaussian fits and was the initial motivation for the width analysis presented in this thesis.

Given these fits are taken with a different choice of $U(\Delta\varphi, \Delta\eta)$, the extracted yields from the von Mises function deviate from the nominal extraction procedure by a relatively large amount, with both the h- Λ and h-h jet-like yields seeing a decrease of around 5%. This percentage is familiarly the same as the percent deviation seen in Section ?? from the h- Λ and h-h yields when using the v_2 -based UE fit while still using bin-wise summation to extract the yields. In fact, all of the yields extracted

using the von Mises fitting procedure are nearly identical to the yields extracted using the v_2 -based UE fit with bin-wise summation, again indicating that the data are well described by the fits. As these two procedures are nearly identical in their results, the von Mises fitting procedure is also excluded from the final systematic uncertainty calculation for the yield extraction after the Barlow check in Section ???. However, the fits are so incredibly well-behaved that they spawned the initial investigation into the near- and away-side jet widths, which eventually became a major topic of interest in this thesis².

4.1.2.6 Barlow check for yield extraction

Following the same procedure as outline in Section ??, a Barlow check is performed for the different techniques for extracting the per-trigger yields from Equations ?? and ???. If the majority of the measured h- Λ yields for a given variation have $|N\sigma_{RB}| < 1$, that variation is excluded from the systematic uncertainty calculation. This majority is calculated across all kinematic regions (near-side jet, away-side jet, underlying event), multiplicity bins and associated momentum bins. While the h-h yields were initially considered for this check, their statistical errors are so small that the denominator in Equation ?? is close to zero, resulting in erratic $N\sigma_{RB}$ values. Thus any technique which gets excluded for the h- Λ yields will also be excluded for the h-h yields. Examples of the Barlow check for the yield extraction are shown in Figure ??.

As a result of the Barlow check, the following variations are excluded from the systematic uncertainty calculation:

- The double Gaussian fit procedure
- The von Mises fit procedure

These exclusions were foreshadowed in the previous sections, as the fits describe the data well enough that there are no statistically significant differences between the yields extracted using these fit functions and the yields extracted using bin-wise summation.

²Systematic uncertainty calculations involving misbehaving fits are a nightmare.

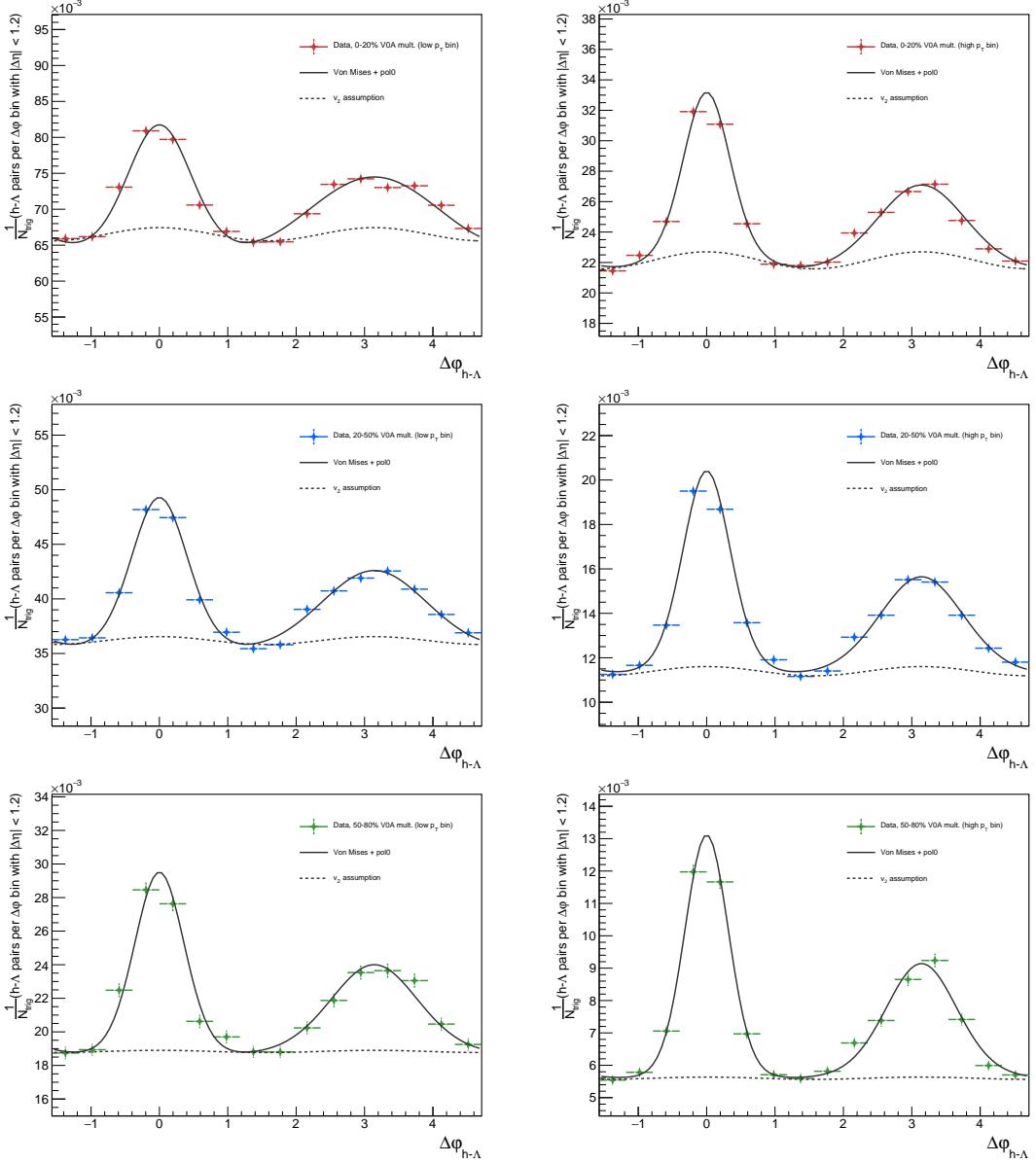


Figure 4.16: The final per-trigger h- Λ $\Delta\phi$ correlation distributions with the von Mises fit for the 0-20% (top), 20-50% (middle) and 50-80% (bottom) multiplicity bins in the lower (left) and higher (right) associated p_T bins. The UE fit with v_2 contribution is also shown as a dashed line.

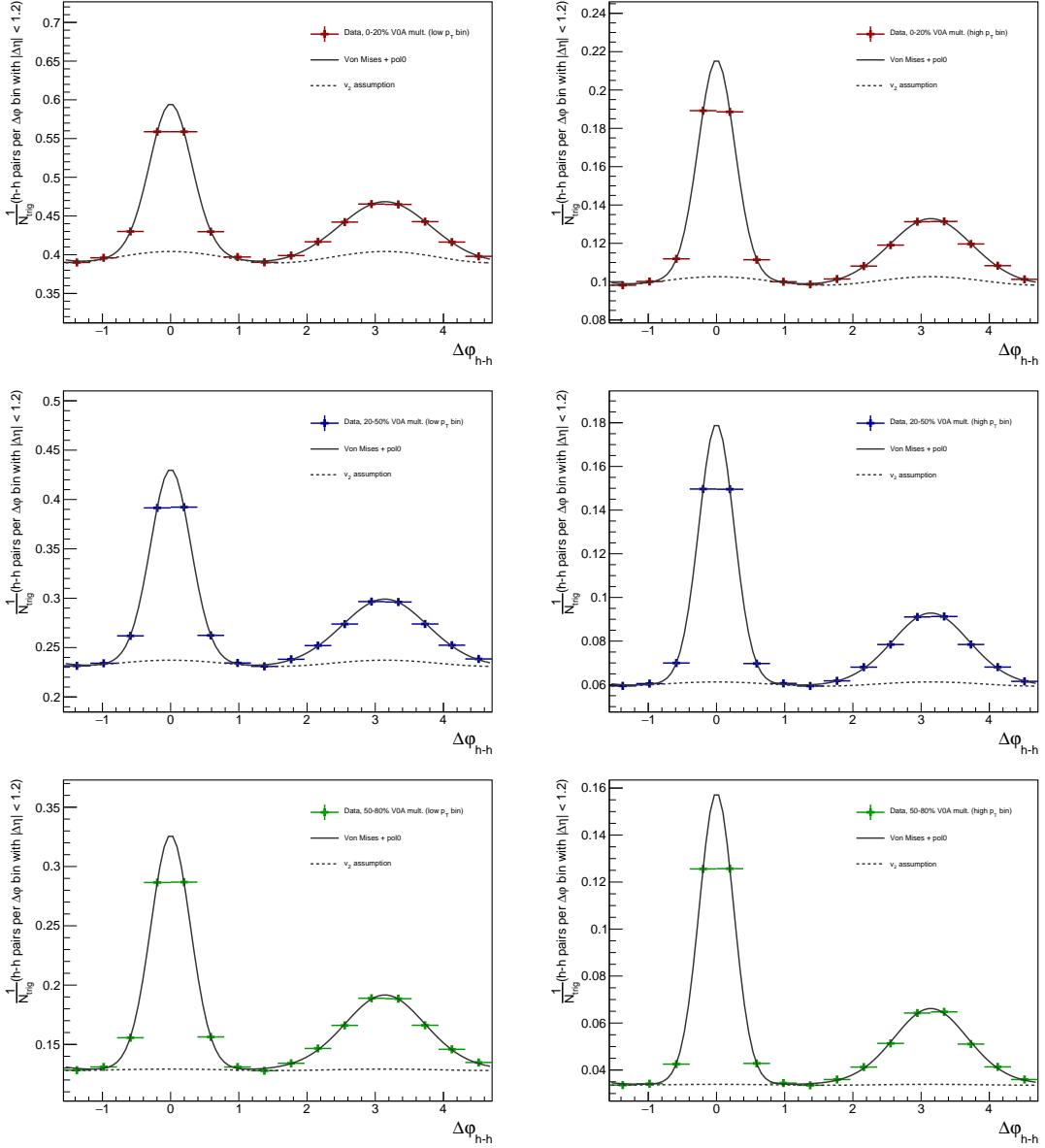


Figure 4.17: The final per-trigger h-h $\Delta\phi$ correlation distributions with the von Mises fit for the 0-20% (top), 20-50% (middle) and 50-80% (bottom) multiplicity bins in the lower (left) and higher (right) associated p_T bins. The UE fit with v_2 contribution is also shown as a dashed line.

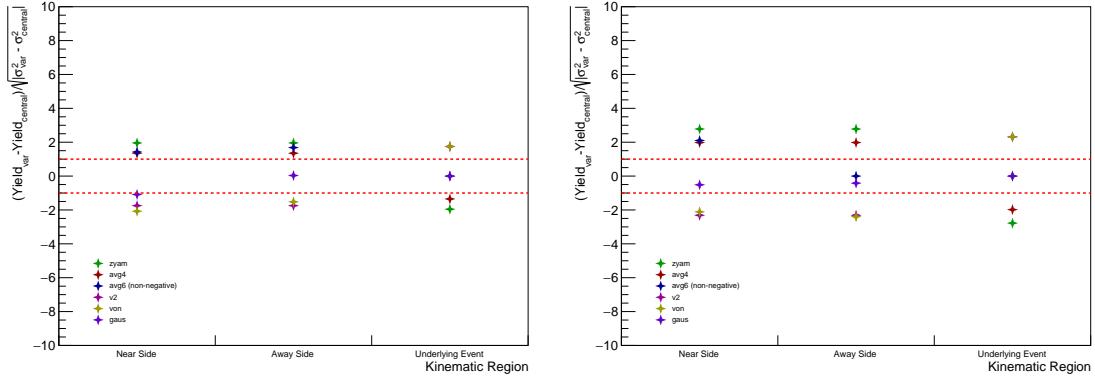


Figure 4.18: The Barlow check for the yield extraction procedure in the 0-20% multiplicity bin for the lower (left) and higher (right) associated p_T bins. The red lines represent $N_{\sigma_{RB}} = \pm 1$, and if the majority of the points from a given procedure fall within these lines (across all multiplicity and momentum bins), the procedure is excluded from the systematic uncertainty calculation.

4.1.2.7 Yield extraction systematics, summarized

As most of the main results of this thesis involve the h- Λ and h-h per-trigger yields in the various kinematic regions, the systematic uncertainties associated with the extraction of these yields deserve to be consolidated to concise tables and plots from the overly detailed descriptions of the previous sections. To that end, plots demonstrating the resulting yield deviations from the nominal technique for each of the aforementioned yield extraction procedure variations (post Barlow check) can be seen in Figures ?? (h- Λ) and ?? (h-h). Furthermore, tables containing the final systematic uncertainties for both the h- Λ and h-h per-trigger yields in each kinematic region for every multiplicity and associated p_T bin can be seen in Tables ?? (h- Λ) and ?? (h-h). Note that included in these systematic uncertainties are both the technique variations associated with the yield extraction procedure, as well as the variations in the yields due to the variations in the $\Delta\varphi$ distributions themselves (after the Barlow check), as discussed in Section ?? . The UE yield systematic uncertainties are generally much lower than the jet-like yields, averaging around 3.5% for both the h- Λ and h-h cases. The larger uncertainties for the jet-like yields (4-7%) are primarily due to the fact that the jet-like yield extraction techniques rely both on the integration procedure and the choice of $U(\Delta\varphi)$, where both including v_2 and excluding negative yield con-

tributions can have a large effect on these extracted yields. Both of those choices have little effect (or no effect in the case of the non-negative yield requirement) on the integral of the $U(\Delta\varphi)$ across the entire azimuthal range (i.e. the UE yield).

As was the case in Section ??, the portion of systematic uncertainty which is uncorrelated with multiplicity is computed using Equations ?? and ?? and presented in Tables ?? (h- Λ) and ???. Whenever the trends of these yields with respect to multiplicity are measured (either by taking slopes or looking at percent differences), the errors are always calculated using these multiplicity-uncorrelated systematic uncertainties.

Table 4.7: Final systematic errors (in %) for the per-trigger h- Λ yields in each kinematic region, multiplicity and momentum bin.

Mult. and p_T bin	Near-side (jet)	Away-side (jet)	UE
0-20%, low	5.50e+00	5.57e+00	3.12e+00
20-50%, low	4.94e+00	5.24e+00	3.22e+00
50-80%, low	6.34e+00	7.19e+00	3.68e+00
0-20%, high	5.47e+00	6.10e+00	3.15e+00
20-50%, high	5.88e+00	6.54e+00	3.33e+00
50-80%, high	4.75e+00	5.25e+00	3.72e+00

Table 4.8: Final systematic errors (in %) for the h-h per-trigger yields in each kinematic region, multiplicity and momentum bin.

Mult. and p_T bin	Near-side (jet)	Away-side (jet)	UE
0-20%, low	4.46e+00	4.75e+00	3.51e+00
20-50%, low	3.75e+00	3.63e+00	3.50e+00
50-80%, low	4.94e+00	6.35e+00	3.67e+00
0-20%, high	3.80e+00	4.00e+00	3.51e+00
20-50%, high	3.61e+00	3.63e+00	3.51e+00
50-80%, high	4.17e+00	5.16e+00	3.81e+00

Table 4.9: Final multiplicity-uncorrelated systematic errors (in %) for the per-trigger h- Λ yields in each kinematic region, multiplicity and momentum bin, used for calculating errors associated with quantities describing trends versus multiplicity (slopes and percent changes).

Mult. and p_T bin	Near-side (jet)	Away-side (jet)	UE
0-20%, low	2.73e+00	2.83e+00	6.66e-01
20-50%, low	3.09e+00	3.57e+00	1.00e+00
50-80%, low	5.80e+00	6.95e+00	2.23e+00
0-20%, high	2.91e+00	3.65e+00	7.40e-01
20-50%, high	4.32e+00	5.41e+00	1.28e+00
50-80%, high	4.47e+00	5.45e+00	2.35e+00

Table 4.10: Final multiplicity-uncorrelated systematic errors (in %) for the h-h per-trigger yields in each kinematic region, multiplicity and momentum bin, used for calculated errors associated with quantities describing trends versus multiplicity (slopes and percent changes).

Mult. and p_T bin	Near-side (jet)	Away-side (jet)	UE
0-20%, low	1.72e+00	2.31e+00	1.86e-01
20-50%, low	1.31e+00	1.92e+00	2.29e-01
50-80%, low	3.95e+00	6.23e+00	1.21e+00
0-20%, high	1.03e+00	1.68e+00	2.42e-01
20-50%, high	7.24e-01	1.22e+00	2.76e-01
50-80%, high	2.28e+00	4.08e+00	1.54e+00

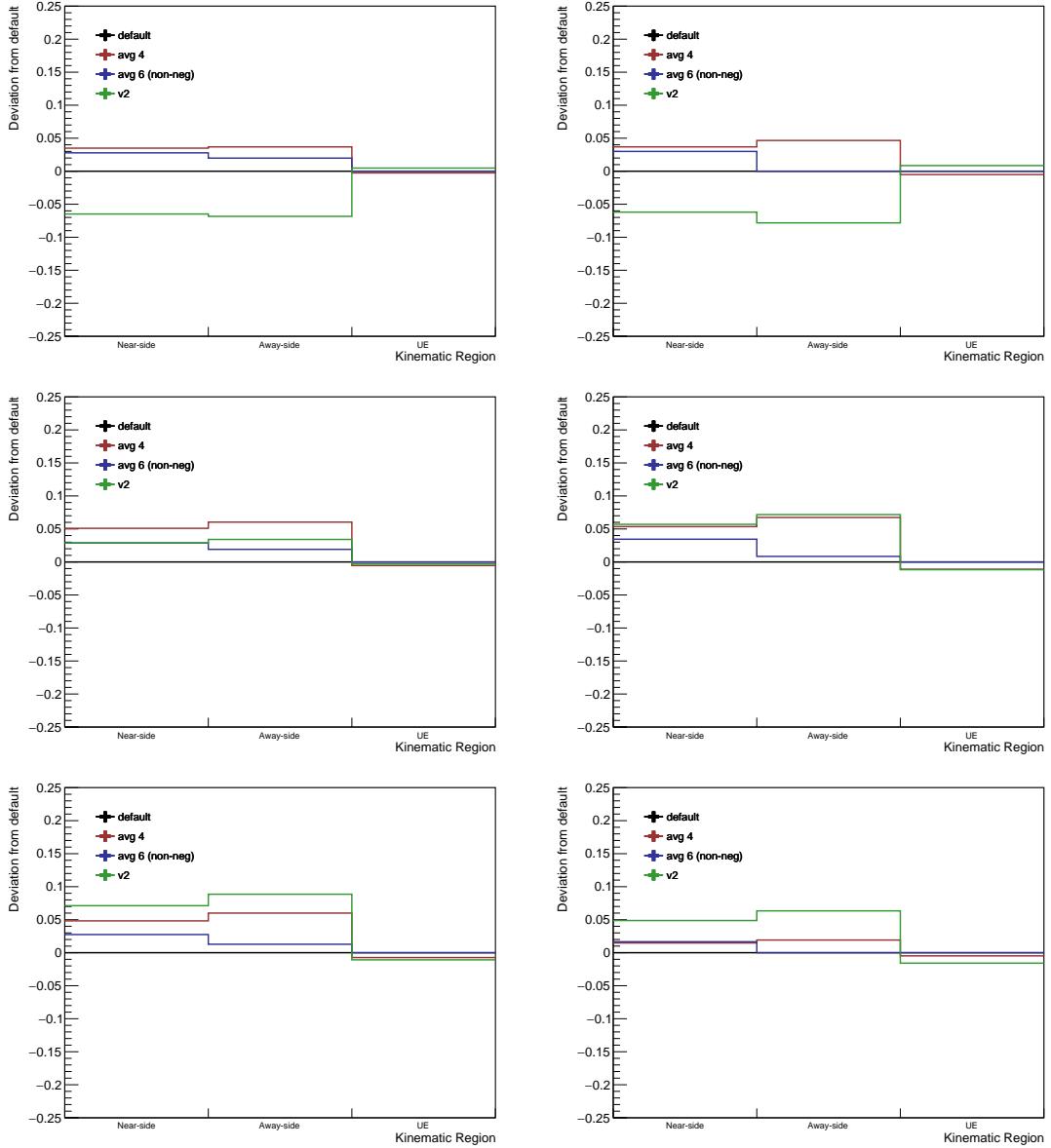


Figure 4.19: The deviation from the nominal $h\text{-}\Lambda$ per-trigger yields for each yield extraction procedure variation (post Barlow check) in each kinematic region for the 0-20% (top), 20-50% (middle) and 50-80% (bottom) multiplicity bins in the lower (left) and higher (right) associated p_T bins. The systematic uncertainty for the yield extraction procedure is calculated by taking the RMS of these deviations for each kinematic region.

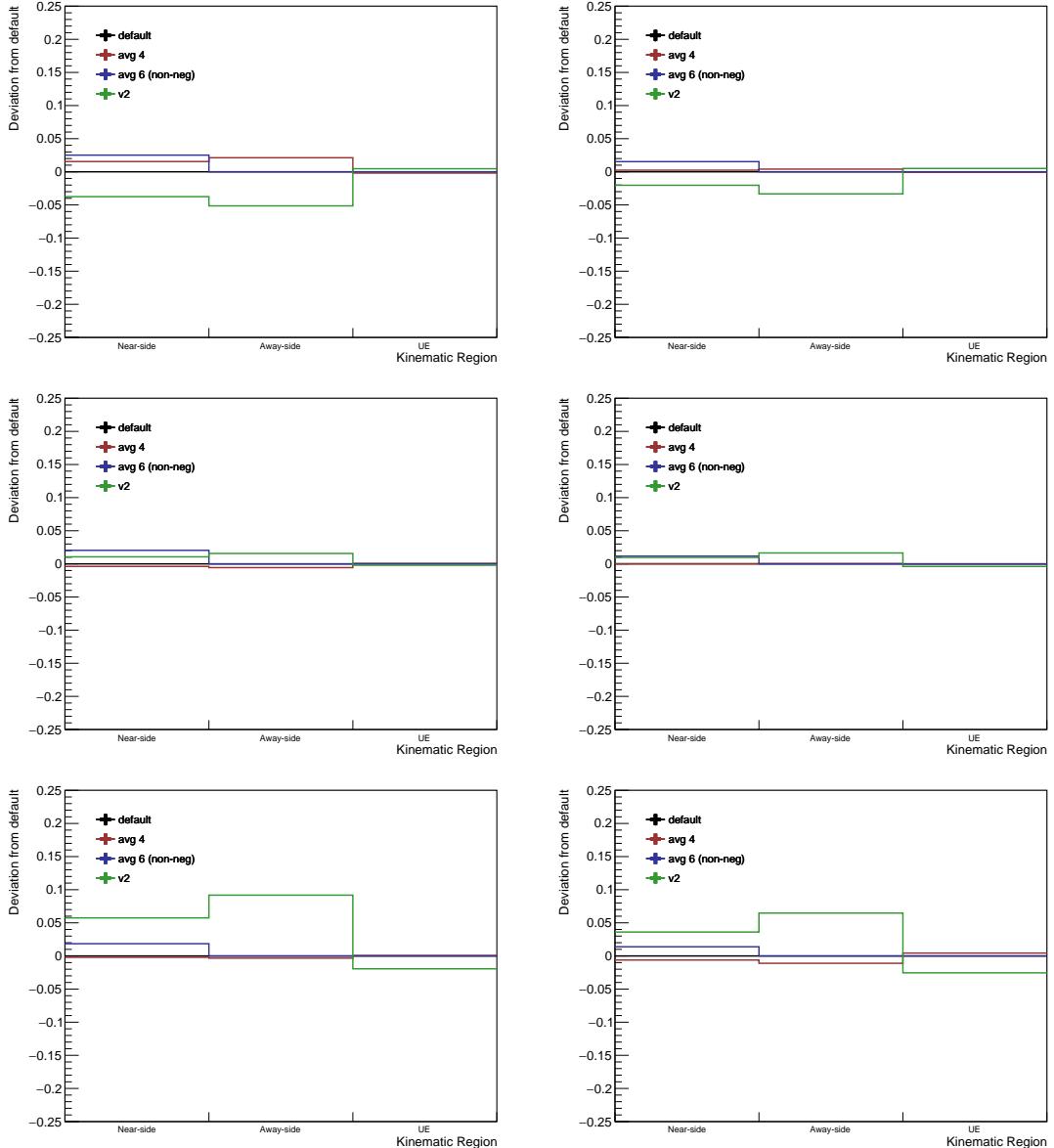


Figure 4.20: The deviation from the nominal h-h per-trigger yields for each yield extraction procedure variation (post Barlow check) in each kinematic region for the 0-20% (top), 20-50% (middle) and 50-80% (bottom) multiplicity bins in the lower (left) and higher (right) associated p_T bins. The systematic uncertainty for the yield extraction procedure is calculated by taking the RMS of these deviations for each kinematic region.

4.1.3 Near- and away-side width extraction

The extracted widths of the near- and away-side jet components from the $\Delta\varphi$ distributions are also subject to a fair amount of systematic uncertainty, both from the variations of the $\Delta\varphi$ distributions themselves as well as the techniques used to extract said widths. None of the $\Delta\varphi$ distributions variations are “new”, in that they have been discussed in some way within the previous sections. However, the effects of these variations on the extracted widths are not as straightforward as the effects on the yields, and thus deserve their own section. Furthermore, the fitting techniques used to extract the widths are different enough than the fitting techniques for yield extraction (Sections ?? and ??), that they will be discussed separately.

As a brief reminder, the nominal procedure for extracting the jet widths is by fitting the $\Delta\varphi$ distribution to the von Mises-based fit function,

$$f(\Delta\varphi) = U(\Delta\varphi) + \frac{A_{\text{NS}}}{2\pi I_0(k_{\text{NS}})} e^{k_{\text{NS}} \cos(\Delta\varphi - \mu_{\text{NS}})} + \frac{A_{\text{AS}}}{2\pi I_0(k_{\text{AS}})} e^{k_{\text{AS}} \cos(\Delta\varphi - \mu_{\text{AS}})}, \quad (4.16)$$

where A and μ are the amplitudes and means of the von Mises components, and k is a measure of the collimation of the distribution, which is inversely related to the width through

$$\sigma = \sqrt{-2 \ln \frac{I_1(\kappa)}{I_0(\kappa)}}. \quad (4.17)$$

In these equations, I_n refers to the modified Bessel function of the nth kind. During the fitting procedure, the means μ_{NS} and μ_{AS} are again fixed to 0 and π , respectively, and the $U(\Delta\varphi)$ term is fixed to the function obtained by fitting the UE which includes a non-zero v_2 contribution, as described in Section ??.

The reason for fixing the U component during fitting is subtle, as the von Mises distributions describe the data extremely well and generally allow for the variation of this component while still obtaining a convergent fit. However, the form of the Von Mises function,

$$f(x) = e^{k \cos(x)}, \quad (4.18)$$

presents a unique issue: if the width is sufficiently large, meaning k is sufficiently small (≈ 1), there is an “offset” from the U term that never tapers off. This is fundamentally different than a Gaussian, which will always converge to zero (or U in this case) at large enough x . A visual depiction of this effect can be seen in Figure ???. In most cases, this never presents an issue as the widths are usually such that $k > 2$.

However, in the lowest momentum bin for the more central $h\text{-}\Lambda$ $\Delta\varphi$ distributions, allowing the U term to vary during the fitting procedure has a very large effect on the corresponding k value, as the fitting software tries to “absorb” this offset into U . Because of this, the U term is fixed during all fitting procedures, and instead the techniques for determining U are varied.

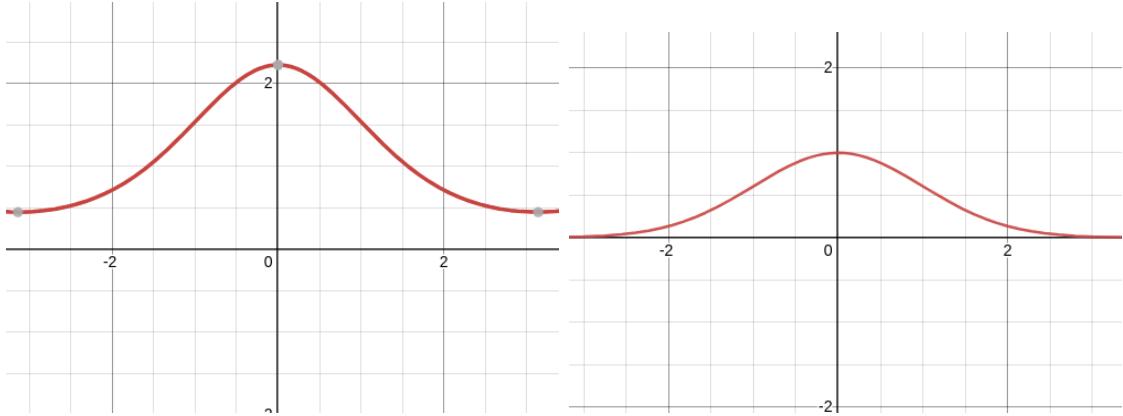


Figure 4.21: The Von Mises (left) and Gaussian (right) dist with $\kappa = 1$ and $\sigma = 1$. Note that the von Mises distribution does not approach zero at large x , while the Gaussian does.

4.1.3.1 Signal, sideband and PID cut variations

Each of the signal, sideband and PID cut variations that affect the $\Delta\varphi$ distributions (from Section ??) can also affect the corresponding extracted widths. To that end, the $h\text{-}\Lambda$ $\Delta\varphi$ von Mises-based fits and extracted widths for each of these variations for all multiplicity and momentum bins can be seen in Figure ???. Note that these widths were extracted using the nominal procedure described at the beginning of this section. Deviations from the nominal near-side widths never exceed 3.5% in most cases, whereas the away-side width deviations are much larger, with some variations resulting in over a 10% change from the nominal value, even after removing statistically insignificant variations via the Barlow procedure in Section ??.

Again, problematic behavior arises from the PID cut variation which requires of a TOF hit for both of the Λ daughters, with both the near- and away-side widths being particularly sensitive to this cut. As the lower momentum daughter pions are usually deflected by the detector’s magnetic field before reaching the TOF, in cases

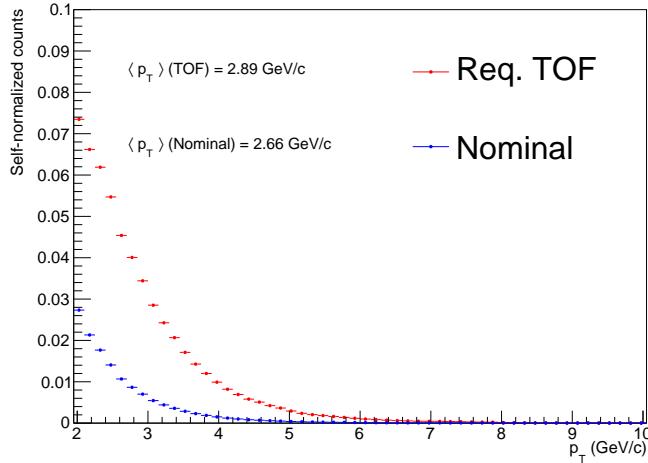


Figure 4.22: The p_T distributions of Λ candidates with and without the TOF signal requirement. The mean p_T of the Λ candidates with daughters that generate a TOF signal is nearly 10% higher than those without this requirement, which indicates that the requiring a TOF signal for the Λ daughters introduces a physical bias.

when the pion actually generates a TOF signal, its corresponding momentum is likely higher than usual. This would result in a higher-than-usual Λ momentum, which in turn would give lower-than-usual jet widths in the corresponding h- Λ correlations as jets become less collimated as their constituent momentum decreases. To test this, the p_T distributions of Λ candidates with and without the TOF signal are compared in Figure ???. The mean p_T of the Λ candidates with daughters that generate a TOF signal is around 10% higher than those without this requirement, which indicates that the TOF signal requirement introduces a physical bias into the h- Λ $\Delta\varphi$ distributions, manifesting as unusually low jet widths. Luckily, this requirement also happens to reduce the overall yield of Λ candidates by huge margin, causing a large amount of statistical fluctuations in the corresponding h- Λ correlation distributions. Because of this, the TOF signal requirement ends up being excluded after the Barlow check. However, it is important to distinguish between excluding variations because they are statistically insignificant and excluding them because they introduce *physical* biases into the data. Had the data sample been larger, the TOF signal requirement would have likely survived the Barlow check, leaving no choice but to rely on the aforementioned argument.

4.1.3.2 Fitting procedure variations

Both of the fitting functions discussed in this thesis are of the form

$$f(\Delta\varphi) = U(\Delta\varphi) + f_{\text{NS}}(\Delta\varphi) + f_{\text{AS}}(\Delta\varphi), \quad (4.19)$$

where U is the underlying event function, and f_{NS} (f_{AS}) is the distribution that describes the near-side (away-side) jet. To estimate the systematic uncertainty associated with the fitting procedure, the following variations are considered:

1. **Varying $U(\Delta\varphi)$:** The $U(\Delta\varphi)$ term is varied by replacing the nominal v_2 -based UE function with a flat line equal to the average of the $\Delta\varphi$ distribution in the ranges $[-\frac{\pi}{2}, -\frac{\pi}{4}) \cup [\frac{\pi}{4}, \frac{5\pi}{8}) \cup [\frac{11\pi}{8}, \frac{3\pi}{2})$ and $[-\frac{\pi}{2}, -\frac{3\pi}{8}) \cup [\frac{3\pi}{8}, \frac{5\pi}{8}) \cup [\frac{11\pi}{8}, \frac{3\pi}{2})$ (i.e. the nominal and restricted-range UE variations from the yield extraction procedure)
2. **Varying the f_{NS} and f_{AS} functions:** The f_{NS} and f_{AS} functions are varied by replacing the nominal von Mises distributions with the Gaussian ones, as described in Section ??, and the U term is fixed to the average of the $\Delta\varphi$ distribution in the ranges $[-\frac{\pi}{2}, -\frac{\pi}{4}) \cup [\frac{\pi}{4}, \frac{5\pi}{8}) \cup [\frac{11\pi}{8}, \frac{3\pi}{2})$ (again, the nominal UE determination technique from the yield extraction procedure)

More variations were initially considered—namely using the v_2 -based UE with the Gaussian functions from (2) and even trying generalized Gaussians **GeneralizedGaus** to describe the jet components—but they were discarded as many of the fits did not converge for all multiplicity and momentum ranges³ despite a large amount of effort.

The choice of making the v_2 -based UE determination procedure the nominal one was not made lightly, as it breaks the symmetry with the nominal yield extraction technique. In the presence of non-zero elliptic flow (as is likely the case at higher multiplicities), the v_2 -based UE determination procedure is the most physically motivated, as it is the only one that takes into account the underlying event’s azimuthal anisotropy. The only reason this procedure was not chosen as the nominal technique for yield extraction is extremely specific to this analysis: at the time of writing this thesis, the only available $h\text{-}\phi(1020)$ correlation results in p–Pb at $\sqrt{s_{\text{NN}}} = 5.02$ TeV come in the form of the $h\text{-}\phi/h\text{-}h$ per-trigger yield ratios, where the yields are extracted assuming a flat UE. As mentioned in Section ??, one of the topics of this analysis

³This is a strict requirement, there are twelve fits in total, each with at least four free parameters, and the fitting software is not particularly robust.

involves investigating open versus hidden strangeness in the form of the h- Λ /h- ϕ per-trigger yield ratios in the different kinematic regions. This is done by taking a ratio of ratios, namely

$$\frac{(h-\Lambda)}{(h-h)_1} / \frac{(h-\phi)}{(h-h)_2} \quad (4.20)$$

where the subscripts 1 and 2 are used to differentiate between this research and the ϕ analysis, respectively. This only reduces to the h- Λ /h- ϕ ratios if two conditions are met: the first is that the h-h distributions are the same (which is investigated more thoroughly in Section ??), and the second is that the yields are extracted from these dihadron distributions using the exact same procedure. While this is not required in the case of the h- Λ yields, the same procedure is applied for the sake of consistency.

The resulting fits and extracted widths for each of the variations listed above in all multiplicity and momentum ranges for both the h- Λ and h-h cases can be seen in Figure ???. Again, only small deviations from the central values are observed in the near-side widths across all variations, with the largest percent difference being around 3% across all multiplicity and momentum bins for both the h- Λ and h-h distributions. Interestingly, the away-side widths appear to be much more sensitive to the inclusion of v_2 , as all variations from the nominal technique—again, each variation assumes a flat UE—result in widths which are systematically lower than the central values by around 5-10%. This is a strange result, as the v_2 -based UE is completely symmetric about $\Delta\varphi = \pi/2$, and thus should affect the near-side widths in the same way as the away-side. As mentioned above, including v_2 in the UE is the more *physically* motivated choice, and therefore is chosen to be nominal despite these large deviations. Note that the Gaussian and von Mises widths are more similar in the cases where the flat UE is used, indicating that the observed deviations are not due to the choice of fitting function.

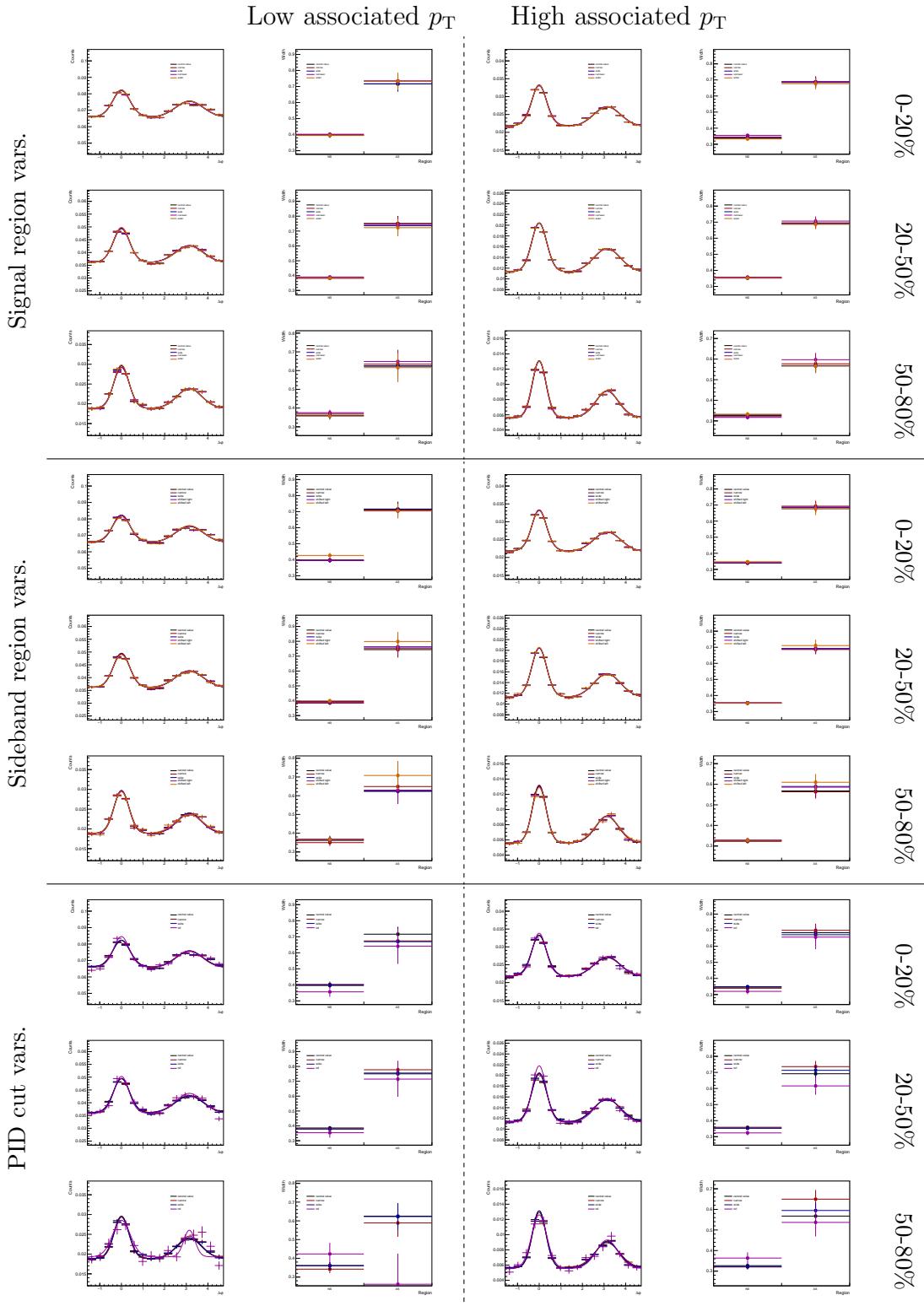


Figure 4.23: The resulting von Mises fits and extracted jet widths after the signal, sideband and PID cut variations are applied to the $h-\Lambda \Delta\varphi$ distributions.

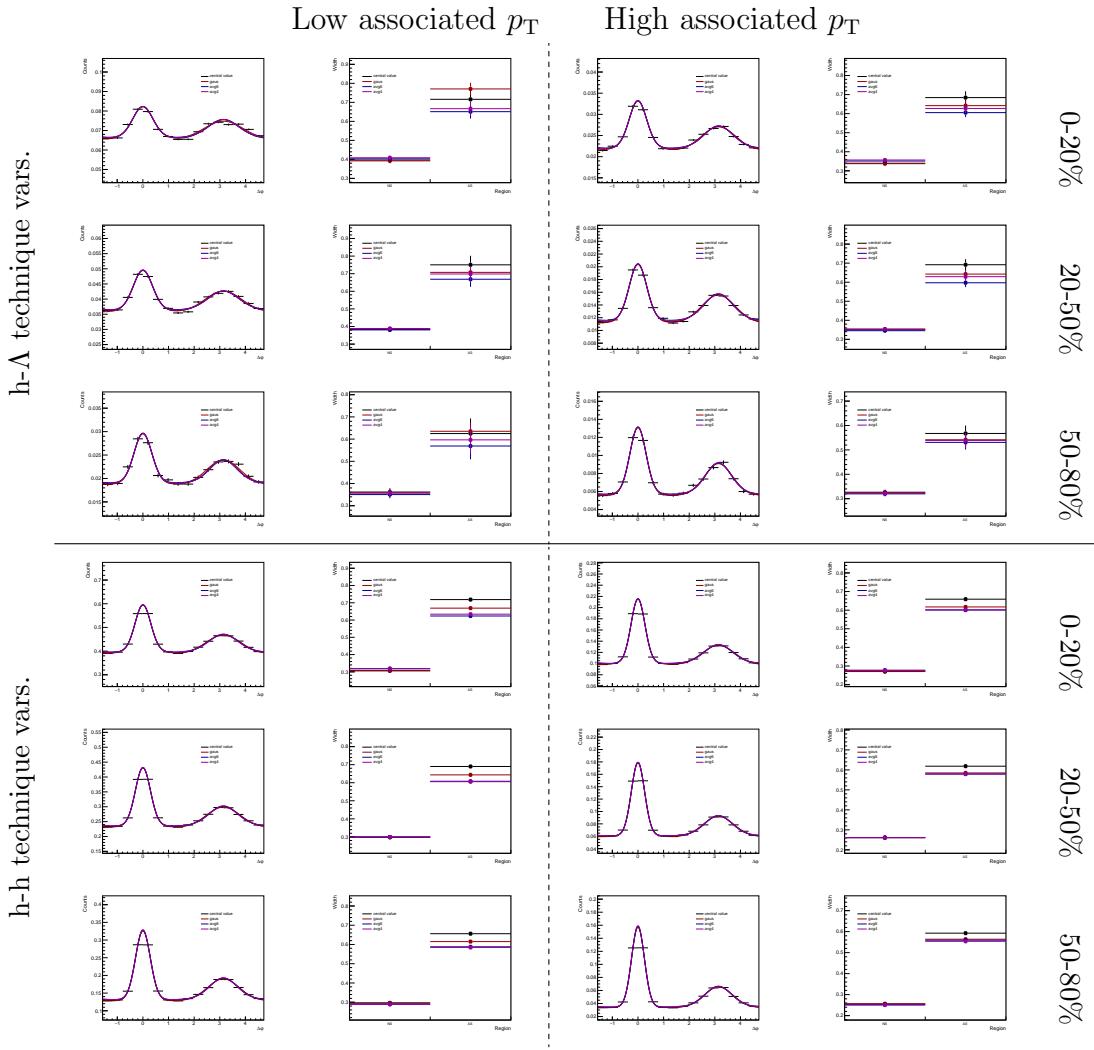


Figure 4.24: The resulting $h\text{-}\Lambda$ (top) and $h\text{-}h$ (bottom) von Mises fits and extracted jet widths in each multiplicity and momentum bin after each variation of the fitting procedure.

4.1.3.3 Barlow check for width extraction

Again, following the same techniques as outlined in Sections ?? and ??, a Barlow check is performed for the variations presented in this section that affect the extracted near- and away-side widths. As was before, if the majority of the extracted h- Λ widths for a given variation have $|N\sigma_{RB}| < 1$, that variation is excluded from the systematic uncertainty calculation. This majority is calculated using both the near- and away-side jet components across all multiplicity bins and associated momentum bins. Again, the dihadron widths are not considered for this procedure⁴, and any fitting technique variation that is excluded from the uncertainty calculation as the result of this check for the h- Λ widths will also be excluded for the h-h case. Visual depictions of the Barlow procedure can be seen in Figure ??.

As a result of the check, the following variations were excluded from the systematic uncertainty calculation for the jet widths:

- Signal: Wide, Wider
- Sideband: Wide, Narrow
- PID: Require TOF

Curiously, these are the same variations that were excluded for the $\Delta\varphi$ distributions, which is *mostly*⁵ a coincidence.

4.1.3.4 Width extraction systematics, summarized

The systematic uncertainties associated with the h- Λ and h-h jet width extraction for each multiplicity and associated momentum bin can be seen in Tables ?? (h- Λ) and ?? (h-h), with visual depictions shown in Figures ?? (h- Λ) and ???. The total systematic uncertainties are obtained by adding the individual uncertainties in quadrature. Note that the away-side width uncertainties are much larger than the near-side, indicating that constraining the away-side jet width is a much more difficult procedure, especially

⁴Due to extremely small statistical errors, see Section ??

⁵There is obviously some correlation between variations in the individual $\Delta\varphi$ bins and the widths. Consider a hypothetical variation that causes an unusually large spike in a $\Delta\varphi$ bin near zero: this would certainly cause the near-side width to be considerably smaller than normal. This is an extreme example, however, and generally these variations affect the individual $\Delta\varphi$ bins in similar ways.

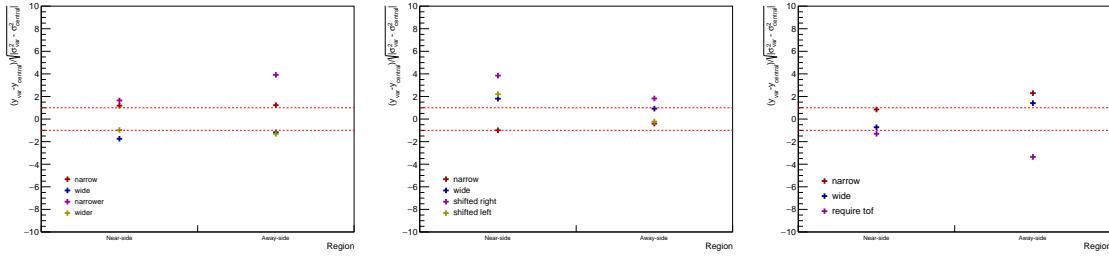


Figure 4.25: Barlow check for the width extraction procedure for the signal (left), sideband (middle), and PID (right) variations in the 20-50% multiplicity bin. The red lines represent $N\sigma_{RB} = \pm 1$, and if the majority of the points fall within the red lines (across all p_T and multiplicity bins), they are excluded from the width systematic uncertainty calculation.

in the presence of non-zero elliptic flow (as discussed in Section ??). As the width uncertainties do not exhibit any significant multiplicity dependence, the multiplicity-uncorrelated portion of these uncertainties is not reported.

Both the topological selection and associated hadron efficiency uncertainties are given in the tables and figures, which were obtained by randomly varying the $\Delta\varphi$ distributions within their respective systematic uncertainties (Table ??) and extracting the widths using the nominal procedure. As the topological selection and tracking efficiency uncertainties were not directly calculated in this analysis, these results serve as a “best guess” for how the widths would be affected by variations in the corresponding selection criteria. The results of this procedure for the h- Λ and h-h distributions in the 20-50% multiplicity bin for both associated momentum ranges can be seen in Figure ??.

Mult.	p_T	Peak	Signal	Sideband	PID	Fit proc.	Topo.	Total
0-20%	low	NS	1.19	5.41	1.19	2.09	3.40	6.93
20-50%	low	NS	1.13	2.70	1.74	0.93	3.40	4.90
50-80%	low	NS	3.20	1.19	3.82	2.05	3.40	6.48
0-20%	low	AS	2.53	1.64	6.21	7.91	2.40	10.77
20-50%	low	AS	0.57	4.68	2.62	8.16	2.40	10.07
50-80%	low	AS	2.88	9.43	4.11	5.88	2.40	12.43
0-20%	high	NS	3.44	1.66	2.19	3.51	3.10	6.43
20-50%	high	NS	1.10	0.48	1.09	1.35	3.10	3.75
50-80%	high	NS	2.46	0.56	1.68	1.54	3.10	4.60
0-20%	high	AS	0.51	1.00	1.99	8.95	6.10	11.07
20-50%	high	AS	1.56	1.93	5.02	10.33	6.10	13.24
50-80%	high	AS	3.93	5.74	10.80	5.40	6.10	15.21

Table 4.11: The final systematic errors from the h- Λ $\Delta\varphi$ near-side (NS) and away-side (AS) width extraction for each multiplicity bin in the lower (top) and higher (bottom) associated p_T bins. The total systematic error is calculated by adding each systematic error in quadrature.

Mult.	p_T	Peak	Fit proc.	Trk. eff.	Total
0-20%	low	NS	2.21	1.00	2.42
20-50%	low	NS	0.21	1.00	1.02
50-80%	low	NS	1.74	1.00	2.01
0-20%	low	AS	8.75	1.50	8.87
20-50%	low	AS	6.66	1.50	6.82
50-80%	low	AS	6.91	1.50	7.07
0-20%	high	NS	1.86	1.00	2.11
20-50%	high	NS	0.12	1.00	1.01
50-80%	high	NS	1.47	1.00	1.78
0-20%	high	AS	8.02	1.50	8.16
20-50%	high	AS	6.06	1.50	6.24
50-80%	high	AS	5.73	1.50	5.93

Table 4.12: The final systematic errors from the h-h $\Delta\varphi$ near-side (NS) and away-side (AS) width extraction for each multiplicity bin in the lower (top) and higher (bottom) associated p_T bins. The total systematic error is calculated by adding each systematic error in quadrature.

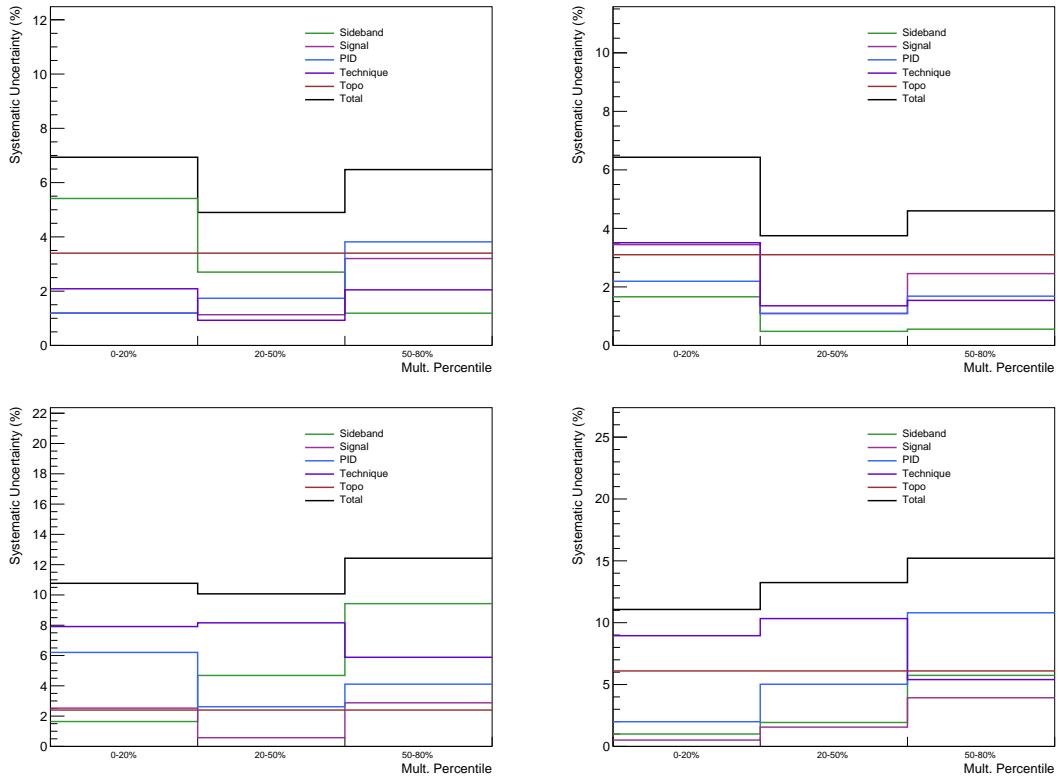


Figure 4.26: Final systematic errors for the $h\Lambda \Delta\varphi$ near-side (top) and away-side (bottom) width extraction for each multiplicity bin in the lower (left) and higher (right) associated p_T bins.

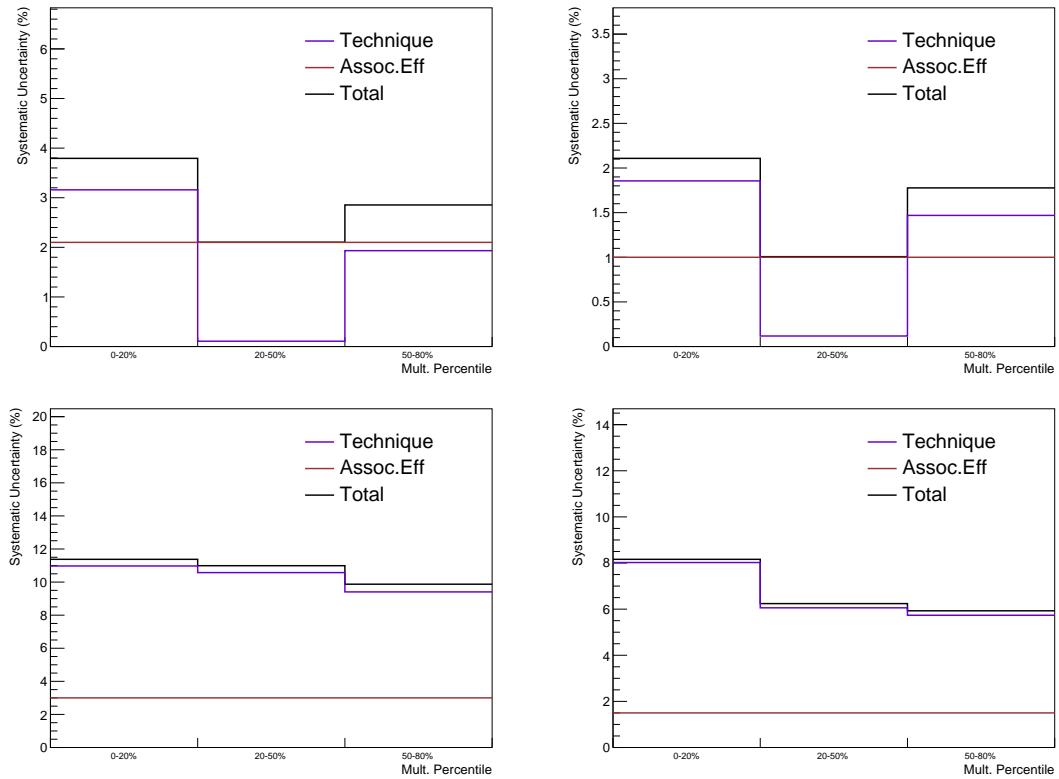


Figure 4.27: Final systematic errors for the h-h $\Delta\varphi$ near-side (top) and away-side (bottom) width extraction for each multiplicity bin in the lower (left) and higher (right) associated p_T bins.

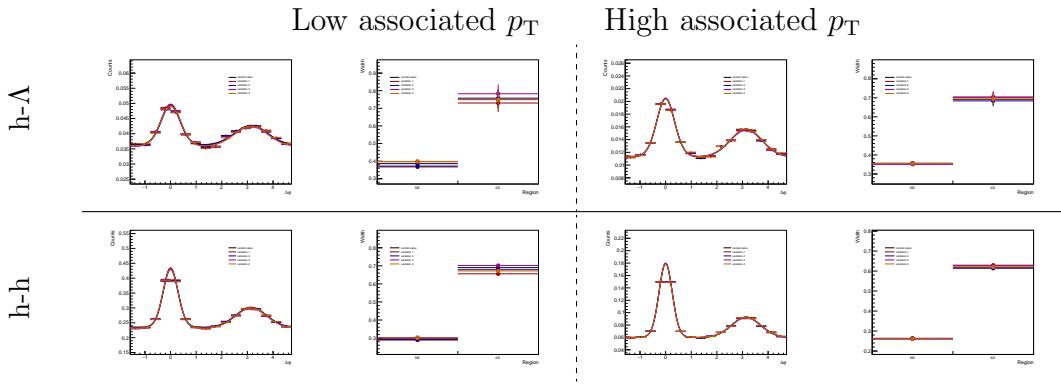


Figure 4.28: The resulting h- Λ (top) and h-h (bottom) von Mises fits and extracted jet widths in the 20-50% multiplicity bin for each momentum bin after random variations of the individual $\Delta\varphi$ bins within the topological selection (h- Λ) and tracking efficiency (h-h) uncertainties from Table ??.

4.1.4 Final systematics tables

For ease of reference, all of the final systematic uncertainties discussed in the previous sections have been consolidated into Tables ?? (h- Λ) and ?? (h-h).

Table 4.13: The final systematic uncertainties for the h- Λ $\Delta\varphi$ distributions, the per-trigger jet and UE yields, and the jet widths in all multiplicity and associated momentum bins. The fraction of the systematic uncertainty which is uncorrelated with multiplicity is reported separately in parentheses for the extracted yields. The low (high) momentum bin is such that $1.5 < p_T^{\text{assoc.}} < 2.5 \text{ GeV}/c$ ($2.5 < p_T^{\text{assoc.}} < 4.0 \text{ GeV}/c$).

Mult. and p_T bin	$\Delta\varphi$ dist.	Y_{near}	Y_{away}	Y_{UE}	σ_{NS}	σ_{away}
0-20%, low	3.3%	5.5(2.7)%	5.6(2.8)%	3.1(0.7)%	6.9%	10.8%
20-50%, low	3.4%	4.9(3.1)%	5.2(3.6)%	3.2(1.0)%	4.9%	10.1%
50-80%, low	3.8%	6.3(5.8)%	7.2(7.0)%	3.7(2.2)%	6.5%	12.4%
0-20%, high	3.2%	5.5(2.9)%	6.1(3.7)%	3.2(0.7)%	6.4%	11.1%
20-50%, high	3.3%	5.9(4.3)%	6.5(5.4)%	3.3(1.3)%	3.8%	13.2%
50-80%, high	4.1%	4.8(4.5)%	5.3(5.5)%	3.7(2.4)%	4.6%	15.2%

Table 4.14: The final systematic uncertainties for the h-h $\Delta\varphi$ distributions, the per-trigger jet and UE yields, and the jet widths in all multiplicity and associated momentum bins. The fraction of the systematic uncertainty which is uncorrelated with multiplicity is reported separately in parentheses for the extracted yields. The low (high) momentum bin is such that $1.5 < p_T^{\text{assoc.}} < 2.5 \text{ GeV}/c$ ($2.5 < p_T^{\text{assoc.}} < 4.0 \text{ GeV}/c$).

Mult. and p_T bin	$\Delta\varphi$ dist.	Y_{near}	Y_{away}	Y_{UE}	σ_{NS}	σ_{away}
0-20%, low	3.5%	4.5(1.7)%	4.8(2.3)%	3.5(0.2)%	2.4%	8.9%
20-50%, low	3.5%	3.8(1.3)%	3.6(1.9)%	3.5(0.2)%	1.0%	6.8%
50-80%, low	3.5%	4.9(3.9)%	6.4(6.2)%	3.7(1.2)%	2.0%	7.1%
0-20%, high	3.5%	3.8(1.0)%	4.0(1.7)%	3.5(0.2)%	2.1%	8.2%
20-50%, high	3.5%	3.6(0.7)%	3.6(1.2)%	3.5(0.3)%	1.0%	6.2%
50-80%, high	3.5%	4.2(2.3)%	5.2(4.1)%	3.8(1.5)%	1.8%	5.9%

4.2 Cross-checks

As mentioned in the introduction to this chapter, this section presents a number of cross-checks that were performed to verify the validity and robustness of the analysis procedure.

4.2.1 Monte Carlo closure tests

One of the most common procedures for verifying the validity of an experimental analysis comes in the form of a **Monte Carlo closure test MCClosure**. The basic steps of any MC closure test are as follows:

1. Generate a large sample of simulated events using a Monte Carlo generator, which models both the physics of the collision and the detector response
2. Apply the same analysis method and selection criteria to the detector-simulated events as to the experimental data to obtain a **reconstructed** observable
3. Compare the reconstructed observable with the **ground-truth** observable, which is obtained directly from the generator-level particles

Generally the third step is done by taking a ratio of the reconstructed observable to the ground-truth observable, which should be consistent with unity if the analysis procedure is valid. Significant deviations from unity would indicate that the experimental procedure introduces a non-physical bias to the measurement, which should be addressed.

For the Monte Carlo closure tests presented in this section, the p–Pb collision events are simulated using the DPMJet generator **DPMJet**, with the ALICE detector response to the simulated particles handled by the GEANT **GEANT** software package. The reconstructed h- Λ and h-h 2D correlation distributions are generated using the same procedure as described in ???. In particular, this means that:

- Both the reconstructed trigger and associated Λ (h) pass the track cuts described in Tables ?? and ?? (??), respectively

- To avoid issues with PID using GEANT⁶, the Λ daughter tracks are verified to be from a proton/pion by checking their corresponding generator-level particles
- The efficiency and acceptance corrections are applied to the h-h and h- Λ distributions in the same was as described in Section ??
- The additional corrections (sideband subtraction, signal scaling, two-track template) are applied to the h- Λ distributions as they are in Section ??

For the ground-truth distributions, the trigger and associated Λ (h) are taken directly from the generator-level particles. The same kinematic cuts are also applied, namely the trigger is required to have momentum $4.0 < p_T < 8.0$ GeV/ c , the associated Λ (h) have the same momentum ranges as the reconstructed case, and all particles are required to fall within $|\eta| < 0.8$. Applying these pseudorapidity cuts on the trigger and associated particles ensures that the underlying physics⁷ is the same as the reconstructed case. Unfortunately this requirement also introduces the same triangular shape along $\Delta\eta$ in the correlation distributions as seen in data, which is corrected for using the mixed-event technique described in Section ??.

Both the reconstructed and ground-truth 2D correlation distributions are projected onto $\Delta\varphi$ in the range $|\Delta\eta| < 1.2$, and the results for the h- Λ and h-h $\Delta\varphi$ distributions are shown for each associated p_T bin in Figures ?? (h- Λ) and ?? (h-h), along with the corresponding (reconstructed)/(ground-truth) distribution ratios. A fit to the ratio is also shown, which is consistent with unity in all cases. This indicates that the analysis procedure is valid, and that the corrections applied to the h- Λ and h-h distributions are not introducing any non-physical biases to the $\Delta\varphi$ distribution measurements. As all of the observables are derived directly from these distributions, it is safe to assume the reconstructed and ground-truth versions of these observables are also consistent with each other.

While the MC closure tests provide a general framework for checking the validity of an experimental analysis, there are a few more specific checks that are performed to ensure the robustness of this analysis, which are described in the following sections.

⁶The $n\sigma_{\text{TPC,TOF}}$ values predicted by GEANT differ from experimental data by a large margin, so using GEANT for PID is generally avoided whenever possible

⁷Particle production is rapidity-dependent, so choosing particles in a different rapidity range could alter the spectrum.

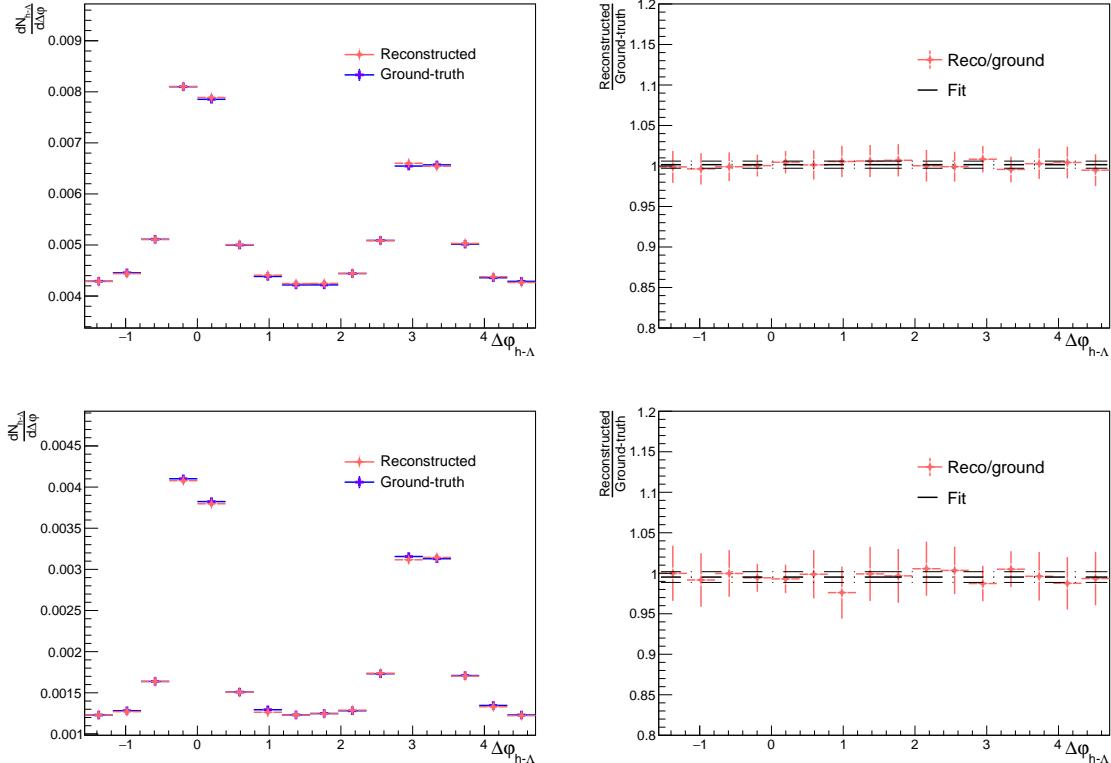


Figure 4.29: The reconstructed (pink) and ground-truth (blue) $h\text{-}\Lambda$ $\Delta\varphi$ distributions in the lower (top) and higher (bottom) associated p_T bins, along with a (reconstructed)/(ground-truth) ratio and straight-line fit with errors shown as dashed lines. The ratio is consistent with unity, and thus the corrections applied to the $h\text{-}\Lambda$ distributions are valid.

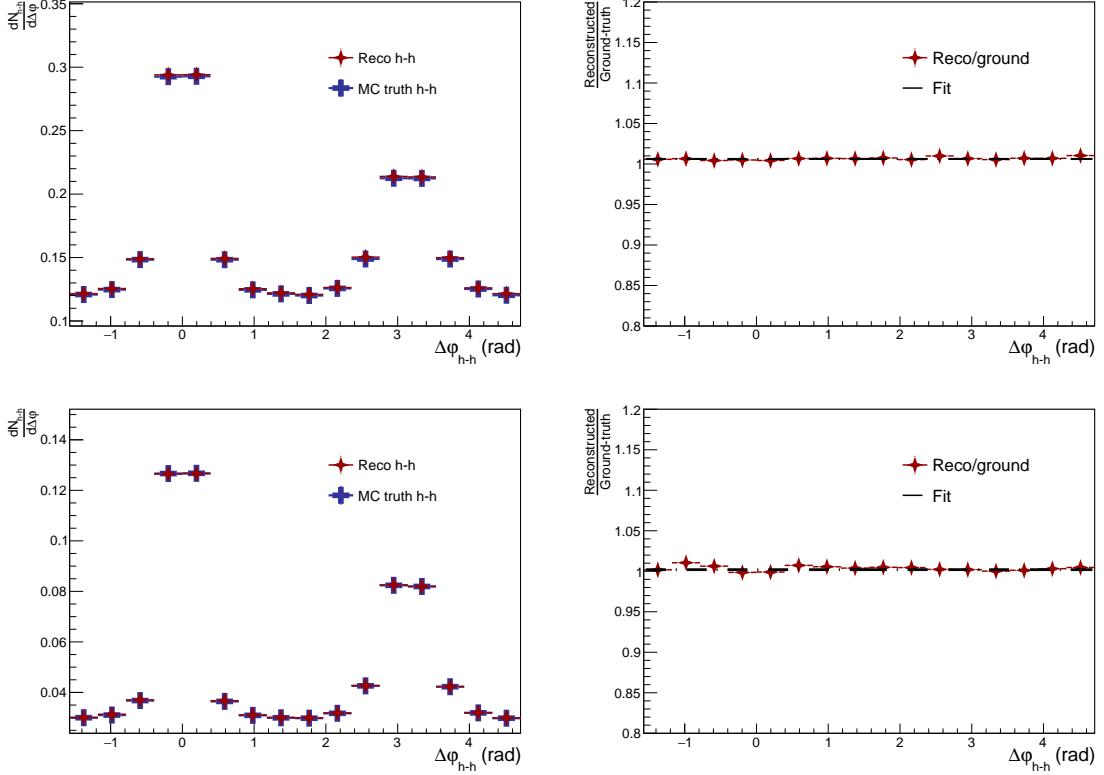


Figure 4.30: The reconstructed (pink) and ground-truth (blue) h-h $\Delta\varphi$ distributions in the lower (top) and higher (bottom) associated p_T bins, along with a (reconstructed)/(ground-truth) ratio and straight-line fit with errors shown as dashed lines. Again, the ratio is consistent with unity, and thus the corrections applied to the h-h distributions are valid.

4.2.2 Correlations with a single trigger

A central feature of this analysis relies on the assumption that the per-trigger pairwise yields from the $h-\Lambda$ or $h-h$ distributions are roughly equal to the per-trigger associated yields for Λ baryons or charged hadrons in events that have a trigger hadron, so that the $(h-\Lambda)/(h-h)$ yield results can be interpreted as the Λ/h ratio in each region. This is only true if there is only a single trigger hadron in each event, which is obviously not the case across a large event sample. To see why this is required, consider two events: the first has a single trigger with three associated particles, while the second has two triggers with three associated particles. In the first event, there are three trigger-associated pairs, which corresponds exactly to the number of associated particles in the event. In the second, there are twice as many pairs as there are associated particles, meaning that the pair-wise yield “double counts” the associated particles. A diagram of this effect is shown in Figure ???. This effect scales with the number of triggers: if there are N triggers in an event, there would be N times as many pairs as associated particles. This diagram also introduces another subtle effect of multiple triggers: if triggers t_1 and t_2 belong to separate jets within a single event, then the associated particles that fall into the jet-like regions when correlated with t_1 could be mistakenly placed in the “underlying event” when correlated with t_2 , artificially inflating the UE yields.

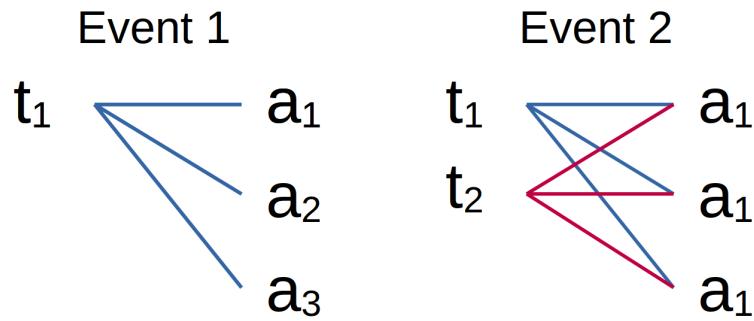


Figure 4.31: A diagram showing the effect of multiple triggers on the per-trigger pair-wise yield. In the first event, there is a single trigger with three associated particles, so the pair-wise yield is equal to the number of associated particles. In the second event, there are two triggers with three associated particles, so the pair-wise yield is twice the number of associated particles.

Luckily, less than 1% of all events have more than a single trigger, as shown in Figure ???. Even still, the effect of multiple triggers on the per-trigger pair-wise yield is investigated by repeating the same analysis procedure as described in the previous chapter, with one change:

- If an event has more than one trigger, only the trigger with the highest momentum within the $4.0 < p_T < 8.0 \text{ GeV}/c$ range is used

This guarantees that no associated particles will be counted more than once, at the expense of a slightly modified trigger p_T spectrum⁸. The final single trigger h- Λ and h-h $\Delta\varphi$ distributions for each multiplicity bin are compared with the original method (all trigger) distributions in Figure ???. A flat deviation of around 5% is observed for both the h- Λ and h-h cases, which is consistent across all multiplicity bins.

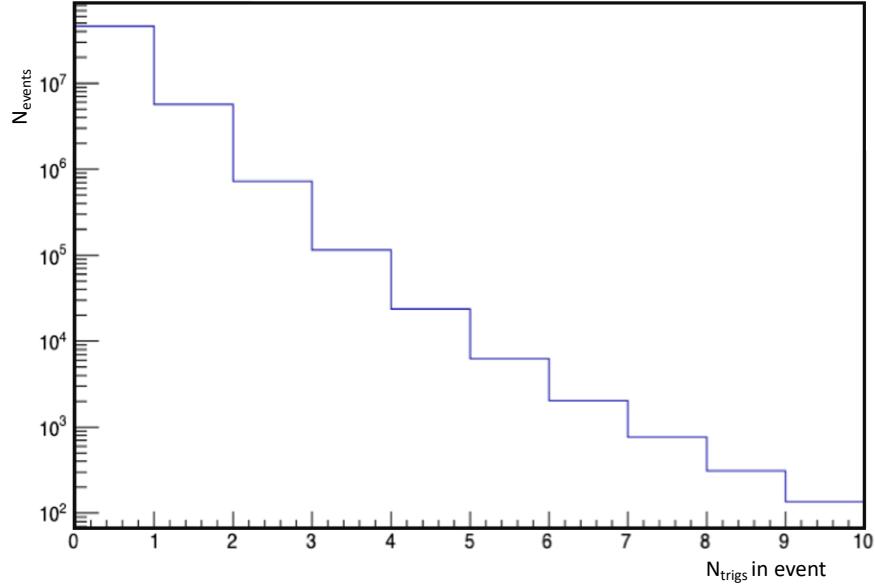


Figure 4.32: A log-plot showing the number of triggers per event across the entire data sample. Only a small fraction of events have at least a single trigger, and of those events, only a small fraction have more than one trigger.

⁸Selecting the highest p_T trigger in the 4-8 GeV/c range usually still amounts to selecting a $\approx 4 \text{ GeV}/c$ particle.

As the (single trigger)/(all trigger) distribution ratios are mostly flat for both the h- Λ and h-h distributions across all multiplicity bins, the effect on the final yield measurements would only be a $\approx 5\%$ change of scale along the y-axis, and the ratio measurements would remain unchanged. Furthermore, this scale factor is the same across all multiplicity bins, indicating that using the standard per-trigger correlation method accurately captures the multiplicity-dependent behaviour of the Λ/h ratios in events with a trigger. Because the conclusions drawn from the results of this thesis only rely on the relative y-axis behavior of the measured observables, the effect of multiple triggers on the final distributions is negligible.

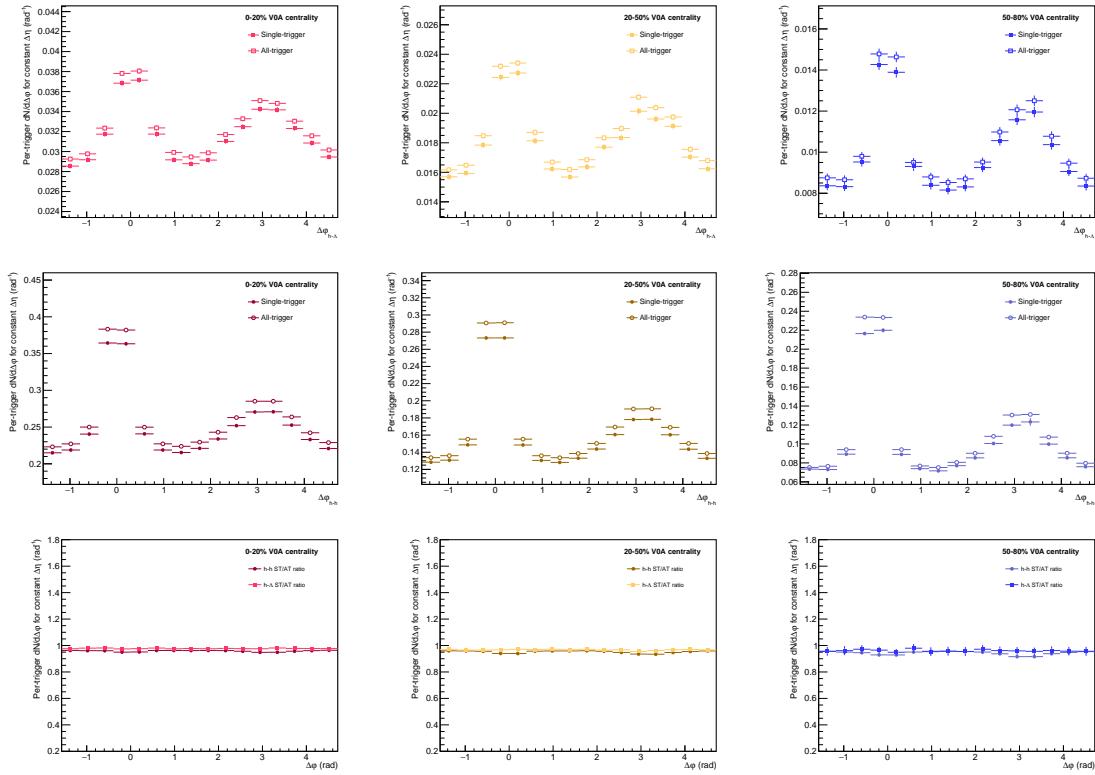


Figure 4.33: The per-trigger h- Λ (top) and h-h (middle) $\Delta\varphi$ distribution comparison between using a single trigger (closed points) and all triggers (open points) in a given event, along with the (single trigger)/(all trigger) distribution ratio (bottom), for the 0-20% (left), 20-50% (center) and 50-80% (bottom) multiplicity bins in the associated momentum range $2.0 < p_T < 4.0 \text{ GeV}/c$. The distributions are nearly identical up to a ≈ 0.95 scale factor, which is constant as a function of multiplicity.

4.2.3 Dihadron comparison with ϕ analysis

As mentioned in Section ??, the differences between open and hidden strangeness are investigated by taking the $\Lambda/\phi(1020)$ per-trigger yield ratios in the different kinematic regions (e.g. near-side jet, away-side jet, underlying event). This is done by taking a ratio of ratios, namely

$$\frac{(h-\Lambda)}{(h-h)_1} / \frac{(h-\phi)}{(h-h)_2}, \quad (4.21)$$

where the $(h-\phi)/(h-h)_2$ ratio in the denominator is taken from previously published results in p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV which use the same angular correlation techniques to extract the per-trigger yields in the different regions **JustinPaper**. This only reduces to the per-trigger $(h-\Lambda)/(h-\phi) \approx \Lambda/\phi$ yield ratios if two conditions are met:

1. The $(h-h)_1$ and $(h-h)_2$ yields are identical
2. The yields are extracted using the same procedure

Condition 2 has been met by design, as both analyses assume a flat UE taken as the average of the correlation distributions in the same $\Delta\varphi$ regions, and they both use bin-wise summation to extract the per-trigger yields. Condition 1, however, needs to be checked more thoroughly. While all of the selection criteria is the same between the two, the analyses were performed by two different humans with vastly different coding styles.

To this end, a direct comparison between the dihadron per-trigger $\Delta\varphi$ distributions from this analysis and the ϕ analysis is shown in Figure ???. The distributions are nearly identical across all multiplicity bins, which is indicated by the ratios at the bottom of the figure never deviating unity. As such, the cancellation of the $(h-h)_1$ and $(h-h)_2$ yields in Equation ?? is valid.

4.2.4 Resonance technique for Λ reconstruction

The Λ baryons used in this analysis are reconstructed by exploiting their characteristic V^0 decay topology, as discussed in Section ???. However, this method introduces a small physical bias in the Λ sample: only Λ s which decay far enough from the primary vertex to be reconstructed as V^0 s (i.e. they have a detector-resolvable secondary

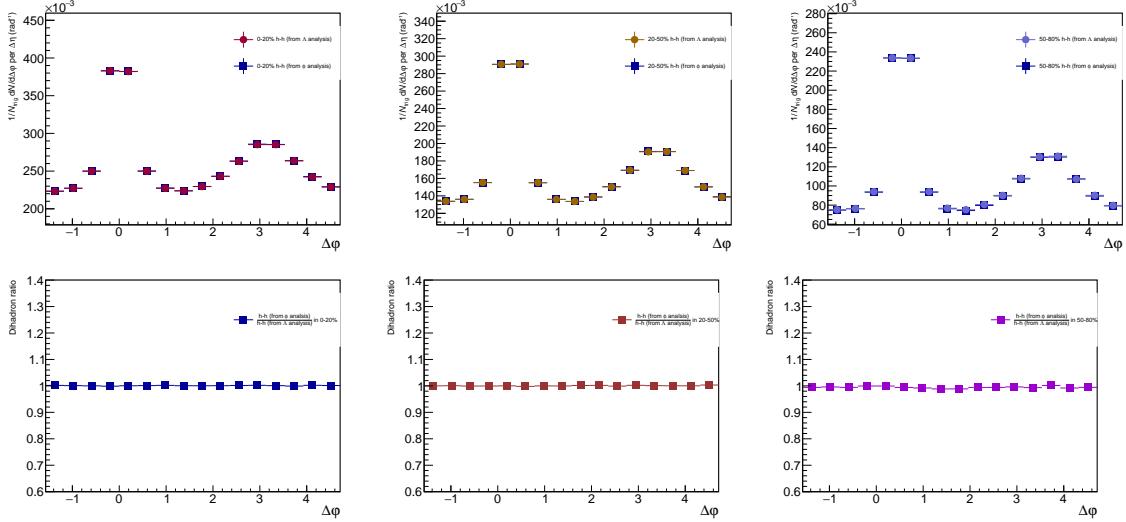


Figure 4.34: Comparison of the dihadron $\Delta\phi$ correlations between the ϕ and Λ analyses for the 0-20% (left), 20-50% (middle) and 50-80% (right) multiplicity bins, taken in the $2.0 < p_T < 4.0 \text{ GeV}/c$ associated momentum range. They are functionally identical for all multiplicity bins.

vertex) are considered. While the average decay length of the Λ is quite large ($c\tau \approx 7.89 \text{ cm}$ where τ is the average lifetime), decay length distributions are exponentials of the form $e^{-t/\tau}$. This means that there is a small fraction of Λ s that decay too quickly to be resolved using their V^0 topology, and thus are not included in the analysis. While the single particle Λ efficiency is used to correct for the overall Λ yield, it may be possible that the $h-\Lambda$ correlation shape is influenced by the exclusion of these short-lived Λ baryons.

To investigate this possible bias, a new set of $h-\Lambda$ distributions are produced, where the associated Λ s are reconstructed using the **resonance technique**: all oppositely charged proton-pion pairs from the event are combined to form Λ candidates. This is the only way to reconstruct extremely short-lived particles—like resonances—as they will never have a resolvable secondary vertex. This technique gives access to all Λ baryons that *can* be reconstructed in the event, regardless of their decay length. However, the large combinatorial background associated with the resonance technique presents many difficulties, which are discussed thoroughly in Appendix ???. Even still, it is possible to repeat the same analysis procedure from the previous chapter with Λ s reconstructed using this technique, and the resulting $h-\Lambda$ $\Delta\phi$ distributions are

shown in Figure ??, along with the original distributions for comparison. It should be noted that neither of the distributions are corrected for the two-track inefficiency or the branching ratio, as this investigation was performed prior to those tools being developed in this analysis. However, the comparison should be mostly⁹ unaffected by these corrections, as they are the same in both cases.

The distributions are very similar across all multiplicity bins, although the resonance technique distributions exhibit much larger statistical fluctuations due to the subtraction of the large combinatorial background using the sideband procedure. This provides solid evidence that the V^0 reconstruction technique is not introducing a significant bias in the shape of the $h\text{-}\Lambda \Delta\varphi$ distributions, and that the exclusion of short-lived Λ baryons has little effect on the final results.

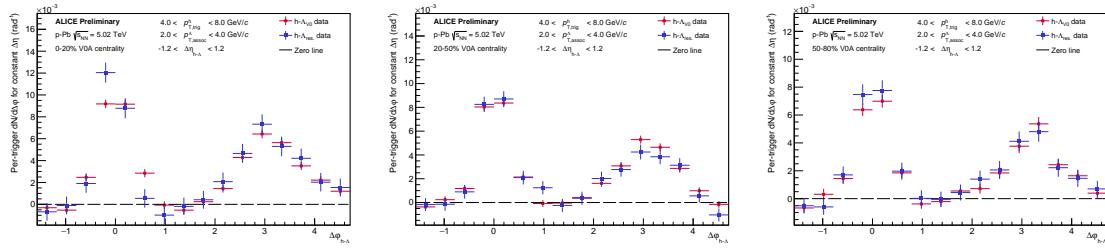


Figure 4.35: The final per-trigger $h\text{-}\Lambda \Delta\varphi$ correlations for Λ s reconstructed using the resonance technique (blue) and the V^0 -based technique (red) in the 0-20% (left), 20-50% (middle) and 50-80% (right) multiplicity bins, taken in the associated momentum range $2.0 < p_T < 4.0$ GeV/c, after the subtraction of the UE. The distributions show good agreement across all multiplicity bins, indicating that the V^0 -based reconstruction technique is not introducing a bias in the correlation shape.

After calculating the systematic uncertainties and performing the various cross-checks described in this chapter, the final results of this thesis are ready to be presented.

⁹The two-track inefficiency is actually very slightly less for the resonance technique as the shorter-lived Λ s usually have daughter tracks with higher quality track parameters due to more hits in the detector, making them less likely to be merged with the trigger.

Chapter Five: Results and discussion

The time has come to present the final results of this thesis. As there are quite a few exciting measurements to be presented, the structure of this chapter is as follows. In the first section, a brief summary of the *why* and *how* are given to provide context for the results. The next section presents the final results as they are, with lengthy discussions about the trends for each observable. The final section will compare these results to previous measurements and theoretical models, and will discuss the implications of these measurements on the current understanding of strangeness production in heavy-ion collisions.

5.0.1 Per-trigger $\Delta\varphi$ distributions

The final per-trigger h- Λ and h-h $\Delta\varphi$ distributions are shown for each multiplicity class for the lower ($1.5 < p_T < 2.5$ GeV/ c) and higher ($2.5 < p_T < 4.0$ GeV/ c) associated p_T bins in Figures ?? and ??, respectively. The entire range along the y-axis is shown to emphasize the relative contribution to each distribution from the UE. As discussed in Section ??, the UE line is calculated using the average of the distribution in the regions $[-\frac{\pi}{2}, -\frac{\pi}{4}] \cup [\frac{\pi}{4}, \frac{5\pi}{8}] \cup [\frac{11\pi}{8}, \frac{3\pi}{2}]$. The relative fraction of h-h and h- Λ pairs in the UE region is maximized in the highest multiplicity class, and subsequently decreases with decreasing multiplicity. The UE contribution is also seen to decrease substantially with increasing associated p_T . Conversely, the near- and away-side jet regions decrease with increasing multiplicity and decreasing associated p_T . These observations suggest that production in the UE region truly is “softer” than production in the near- and away-side jet regions, as expected.

The per-trigger jet-like yields (Y_{near} , Y_{away}) are shown in each associated p_T bin as a function of multiplicity for both the h- Λ and dihadron correlations in Figure ???. To improve compatibility with previous results, the multiplicity classes have been converted to charged particle multiplicity by computing $\langle dN_{\text{ch}}/d\eta \rangle$ in each class for all charged hadrons with $|\eta| < 0.5$ and $p_T > 0.15$ GeV/ c . Straight line fits of the data are shown as dashed lines. The same yields obtained using DPMJET are also shown, with a ratio to the data presented in the bottom panel. A dashed line at unity is

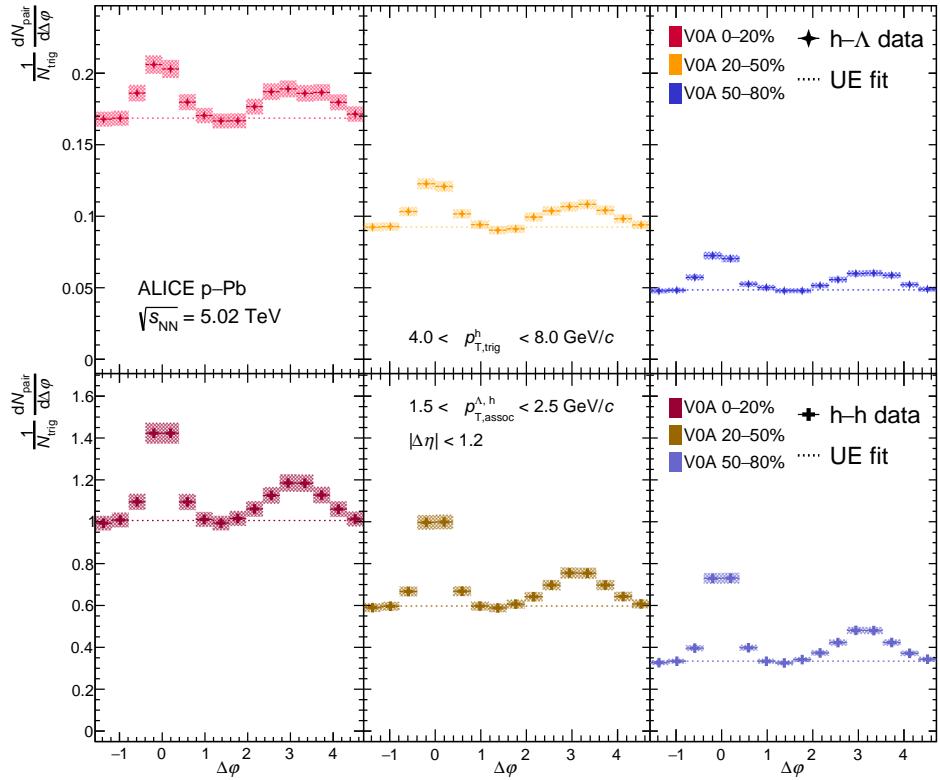


Figure 5.1: The h- Λ (top) and h-h (bottom) $\Delta\varphi$ distributions for each multiplicity class with $1.5 < p_{T,\text{assoc}} < 2.5 \text{ GeV}/c$, with statistical (systematic) uncertainties shown as vertical lines (shaded boxes). The multiplicity classes are plotted from most central (left) to least central (right). The UE estimate is shown as a dashed line, and is taken as the average of the distribution in the regions $[-\frac{\pi}{2}, -\frac{\pi}{4}) \cup [\frac{\pi}{4}, \frac{5\pi}{8}) \cup [\frac{11\pi}{8}, \frac{3\pi}{2})$.

drawn to help better visualize the deviations between data and the model.

Across both associated p_T bins, the h- Λ per-trigger yields see a substantial increase with respect to multiplicity for both the near- and away-side jet regions. This is in stark contrast to the dihadron yields, which see only a small increase as a function of multiplicity in the lower p_T bin and effectively zero increase across the entire multiplicity range in the higher p_T bin. The increase can be quantified by calculating the percent change in the per-trigger yields from the lowest to highest multiplicity class, which is shown for each momentum bin in Table ???. The errors reported are calculated using only the fraction of systematic uncertainty that is uncorrelated with multiplicity (shown in Table ??), and the p-value is obtained by testing against the

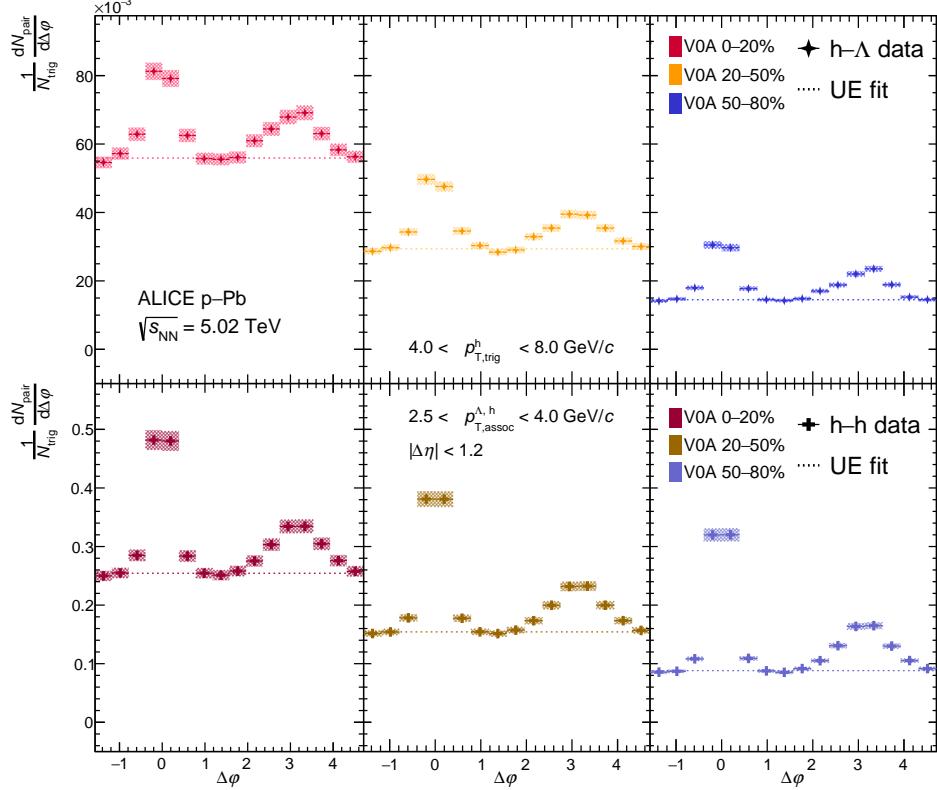


Figure 5.2: The h- Λ (top) and h-h (bottom) $\Delta\varphi$ distributions for each multiplicity class with $2.5 < p_{T,\text{assoc}} < 4.0 \text{ GeV}/c$, with statistical (systematic) uncertainties shown as vertical lines (shaded boxes). The multiplicity classes are plotted from most central (left) to least central (right). The UE estimate is shown as a dashed line, and is taken as the average of the distribution in the regions $[-\frac{\pi}{2}, -\frac{\pi}{4}) \cup [\frac{\pi}{4}, \frac{5\pi}{8}) \cup [\frac{11\pi}{8}, \frac{3\pi}{2})$.

null hypothesis of zero increase. The p-values obtained from the h- Λ yields across both regions and p_T ranges are all $< 1 \times 10^{-5}$, indicating that the increase is statistically significant at $p < 0.01$. However, the dihadron yields see no statistically significant increase in the near-side for both momentum ranges, and only the lower momentum range shows a significant increase in the away-side. The lower associated p_T range also exhibits a much larger increase in the away-side yields when compared to the near-side yields for both the h- Λ and h-h cases. These differences between the near- and away-side yields' behavior as a function of multiplicity hint at a possible modification of the away-side jet production due to soft scattering.

The per-trigger near- and away-side yields predicted by DPMJET are mostly

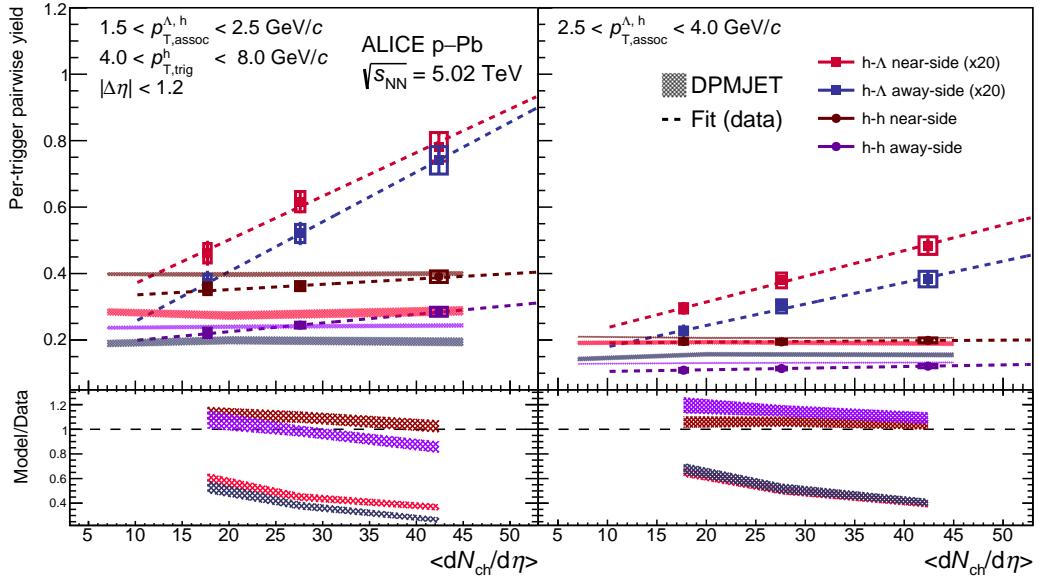


Figure 5.3: The per-trigger pair-wise yields Y_{near} , Y_{away} as a function of charged particle multiplicity for the h- Λ (square markers) and h-h (circle markers) correlations in the lower (left) and higher (right) associated p_T bins. The statistical (systematic) uncertainties are shown as vertical lines (boxes), and a first order polynomial fit to the data is shown as a dashed line. The same yields predicted by DPMJET are also shown as shaded bands, with the width of the band representing the systematic uncertainty on the model. The ratio of the model to the data is shown in the bottom panel, along with a dashed line drawn at unity.

consistent with data in the dihadron case. This can be seen in the model/data ratio, with both the near- and away-side ratios remaining close to unity across the entire multiplicity range. The h- Λ yields, however, are not well described by the model. Both the near- and away-side h- Λ yields predicted by DPMJET are lower than data by around a factor of two across the entire multiplicity range in both momentum bins, and there is no significant increase in these yields as a function of multiplicity.

To gain more insight to the underlying mechanisms responsible for strangeness production in jets, the widths of the near- and away-side jet regions are extracted from the h- Λ and h-h $\Delta\varphi$ distributions using Equations ?? and ???. Plots of these widths as a function of multiplicity for both associated momentum ranges are shown in Figure ??, along with the same widths predicted by DPMJET. A ratio of the model to the data is also presented in the bottom panel of the figure.

Expectedly, the near-side widths exhibit a significant decrease ($>15\%$) from the

Table 5.1: The percent change in the per-trigger yields from the lowest to highest multiplicity class in the lower ($1.5 < p_T < 2.5$ GeV/c) and higher ($2.5 < p_T < 4.0$ GeV/c) associated momentum bins. The errors reported are calculated using only the fraction of systematic uncertainty that is uncorrelated with multiplicity, and the p-values are obtained by testing against the null hypothesis of zero increase.

Region	Percent Change for lower (higher) p_T	Lower (higher) p_T p-value
h- Λ near-side	+69.4 \pm 12.3 (+63.9 \pm 9.6)	< 1×10^{-5} (< 1×10^{-5})
h- Λ away-side	+99.8 \pm 16.1 (+68.7 \pm 11.5)	< 1×10^{-5} (< 1×10^{-5})
h-h near-side	+10.3 \pm 4.9 (+1.7 \pm 2.7)	0.04(0.53)
h-h away-side	+28.5 \pm 8.0 (+11.2 \pm 4.7)	4 $\times 10^{-4}$ (0.02)

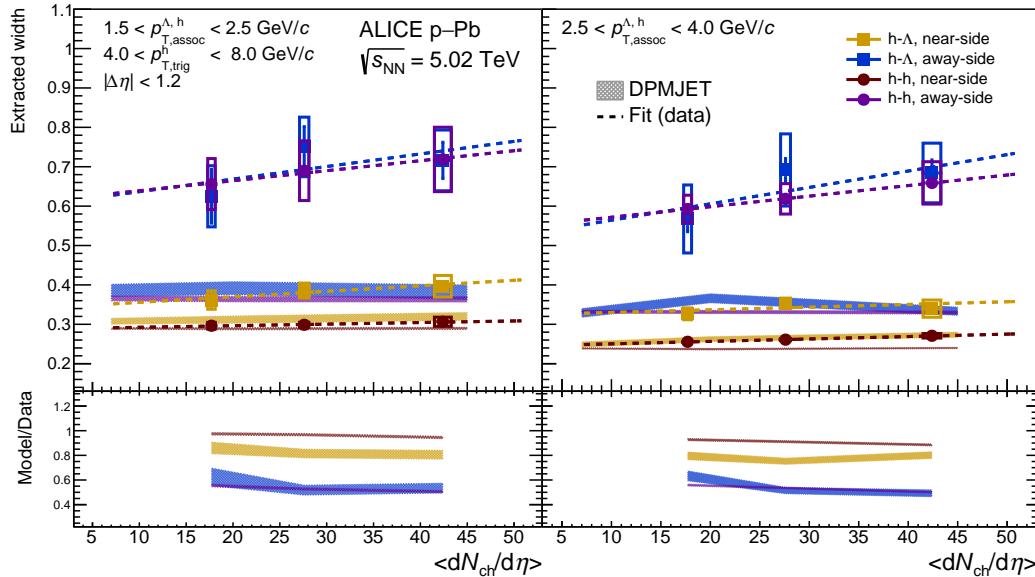


Figure 5.4: The h- Λ and h-h jet widths shown as a function of multiplicity for both associated momentum ranges, along with a straight-line fit to the data. The statistical (systematic) uncertainties are shown as vertical lines (boxes). The ratios predicted by DPMJET are also presented as shaded bands, with the width of the band representing the systematic uncertainty on the model. The ratio of the model to the data is presented in the bottom panel.

lower momentum bin to the higher for both the h- Λ and h-h cases, indicating that the production within the jet becomes more collimated at higher momentums. A more surprising feature comes from comparing the h- Λ and h-h away-side jet widths, which are found to be the same within systematic uncertainties across all multiplicity and momentum ranges. This contrasts with the h- Λ near-side widths, which are

around 40% (2σ) than the h-h widths across the entire multiplicity range for both momentum bins. For the h-h near-side widths, DPMJET describes the data well across both momentum ranges, with a $< 5(10)\%$ deviation from data seen in the lower (higher) momentum bin. DPMJET also predicts the h- Λ near-side width to be larger than the h-h width, though the values of the h- Λ widths are much lower than they are in data. One explanation for these differences between the near-side widths could be due the presence of gluon jets, which are generally more wide than quark jets **GluonJet1** and exhibit an increased production of Λ baryons **GluonJet2**. As DPMJET includes both quark and gluon jets, it is possible that the predicted differences between the h- Λ and h-h jet widths are due to this effect.

DPMJET also under-predicts both the h- Λ and h-h away-side widths by around 40% across both momentum ranges. As the DPMJET model does not include any medium effects, this suggests that the away-side jet widths in data are possibly “broadened” by jet-medium interactions. However, the larger uncertainties on the away-side widths prevent the exclusion of flat behavior with respect to multiplicity (i.e. increasing medium size), as the slopes are all consistent with zero within uncertainties.

5.0.2 Per-trigger yield ratios

To better understand the differences between Λ and charged hadron production both in and out-of jets, the per-trigger yield ratios $R_i^{\Lambda/h} \equiv Y_i^{h-\Lambda}/Y_i^{h-h}$ (i = near-side jet, away-side jet, UE) are measured as a function of multiplicity in both associated momentum bins. These ratios serve as a proxy for the Λ/π ratio in each region, and are shown in Figure ???. Straight line fits to the data are shown as dashed lines, with slopes and corresponding errors reported in Table ???. The same ratios predicted by DPMJET are again shown, with a ratio to the data presented in the bottom panel.

One surprising feature of these results is the clear separation between the ratios in each region across the entire multiplicity range in both momentum bins, with the UE ratio being the largest, followed by the away-side ratio, and finally the near-side. This indicates that most of the relative Λ production is occurring in the UE, which is consistent with the idea that s -quark production is maximized in the medium. This is further supported by the fact that the away-side ratio is larger than the near-side, as Λ production on the away-side is likely due to both the fragmentation of the away-

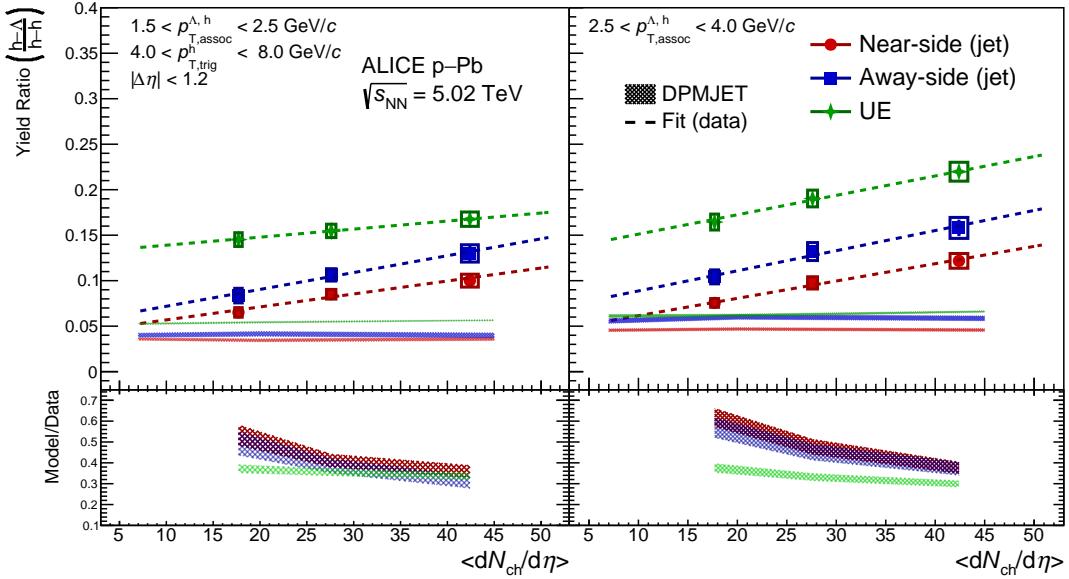


Figure 5.5: The per-trigger pair-wise yield ratios $R_i^{\Lambda/h} \equiv Y_i^{h-\Lambda}/Y_i^{h-h}$ ($i = \text{near-side jet, away-side jet, UE}$) as a function of multiplicity in the lower (left) and higher (right) associated momentum bins. The statistical (systematic) uncertainties are shown as vertical lines (boxes). The ratios predicted by DPMJET are also presented as shaded bands, with the width of the band representing the systematic uncertainty on the model. The ratio of the model to the data is presented in the bottom panel.

Table 5.2: The slopes obtained from the straight-line fits to the per-trigger pair-wise $(h-\Lambda)/(h-h)$ yield ratios as a function of multiplicity in both associated momentum bins. The fits are made using only the fraction of systematic uncertainty that is uncorrelated with multiplicity. All fits are such that $\chi^2/\text{ndf} < 1$.

Region	Lower p_T slope ($\times 10^{-3}$)	Higher p_T slope ($\times 10^{-3}$)
Near-side	1.4 ± 0.3	1.9 ± 0.3
Away-side	1.9 ± 0.4	2.2 ± 0.4
UE	0.9 ± 0.1	2.1 ± 0.2

side jet coupled with the production of strange quarks in the medium. Interestingly, DPMJET is able to produce this ordering in the ratios, but the magnitude of the ratios is not well described.

The near- and away-side slopes reported in Table ?? are not compatible with zero, indicating that there is an enhancement of relative Λ production in jets as a function of multiplicity. This result is unexpected, as it was previously believed that the multiplicity-dependent enhancement of strange hadron production was strictly

due to soft s -quark production in the QGP medium. The away-side slopes are also systematically larger than the near-side slopes in both momentum bins, again hinting at possible modification of the away-side s -quark production due to jet-medium interactions. Similarly, the UE slopes are not compatible with zero, but the value is smaller than the near- and away-side slopes by about 2σ in the lower momentum bin. However, the larger values of the UE ratios overall still suggest that a significant portion of the observed enhancement in the Λ/π ratio is due to softer production from the UE. The slopes calculated using the ratios obtained from DPMJET are all nearly exactly zero, and are thus not shown in the table.

5.0.3 Comparison with the $\phi(1020)$

While the $\phi(1020)$ meson's net strangeness $|S| = 0$, it has been observed to exhibit a similar enhancement in production as a function of multiplicity as other hadrons with non-zero strangeness **PhiEnhancement**. Due to their similar masses ($\Delta M < 100 \text{ MeV}/c^2$), the $\phi(1020)$ is an excellent candidate to compare directly with the Λ in order to better understand the differences between open ($|S| \neq 0$) and hidden ($|S| = 0$) strange hadron production. Using previously published results on $\phi(1020)$ production in and out-of jets in p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$ **JustinPaper**, the per-trigger pair-wise yield ratios $R_i^{\Lambda/\phi} \equiv Y_i^{h-\Lambda}/Y_i^{h-\phi}$ ($i = \text{near-side jet, away-side jet, UE}$) can be measured as a function of multiplicity, which are shown in Figure ?? for both the lower and higher associated p_{T} bins. Again, straight line fits to the data are shown as a dashed lines, with the slopes and corresponding errors reported in Table ???. The same ratios predicted by DPMJET are also presented, with a ratio to the data presented in the bottom panel.

Interestingly, Λ/ϕ near-side ratios are systematically higher than the ratios in all other regions across the entire multiplicity range and both momentum bins. This indicates relative enhancement (suppression) of Λ (ϕ) production along the jet axis. As strangeness is always produced in the form of $s\bar{s}$ pairs, one possible explanation of this effect is that when these pairs are produced from jet fragmentation, the s and \bar{s} are less likely to hadronize into the same ϕ due to their separation in phase-space. This effect would be diminished in the away-side, as the jet-fragmentation-produced s (\bar{s}) could potentially hadronize with an \bar{s} (s) produced in the medium, which is observed as a systematic decrease in the Λ/ϕ ratio in the away-side when compared

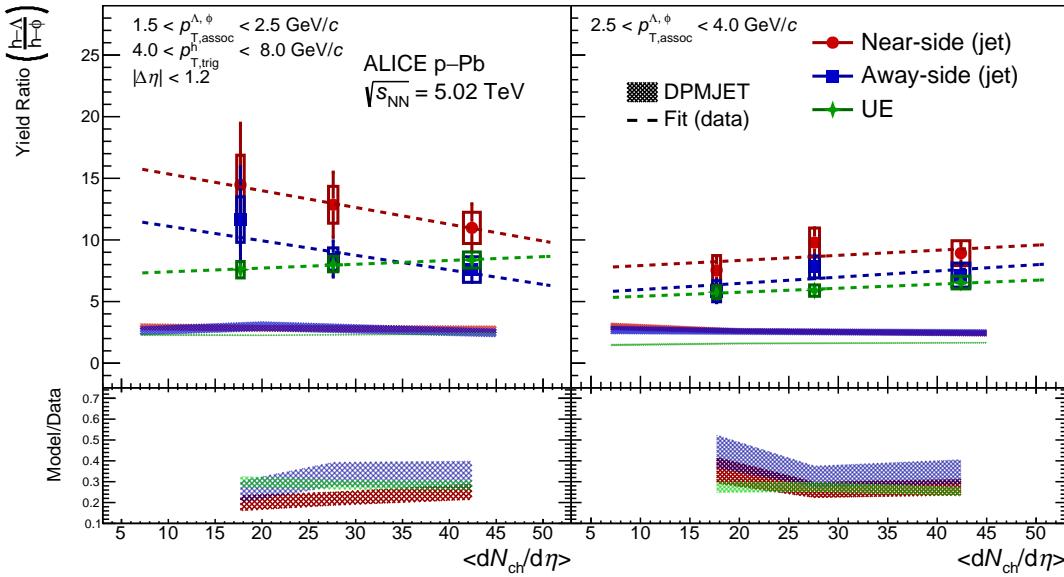


Figure 5.6: The per-trigger pair-wise yield ratios $R_i^{\Lambda/\phi} \equiv Y_i^{h-\Lambda}/Y_i^{h-\phi}$ ($i = \text{near-side jet, away-side jet, UE}$) as a function of multiplicity in the lower (left) and higher (right) associated momentum bins. The statistical (systematic) uncertainties are shown as vertical lines (boxes). The ratios predicted by DPMJET are also presented as shaded bands, with the width of the band representing the systematic uncertainty on the model. The ratio of the model to the data is presented in the bottom panel.

Table 5.3: The slopes obtained from the straight-line fits to the per-trigger pair-wise $(h-\Lambda)/(h-\phi)$ yield ratios as a function of multiplicity in both associated momentum bins. The fits are made using only the fraction of systematic uncertainty that is uncorrelated with multiplicity (in the case of the numerator). All fits are such that $\chi^2/\text{ndf} < 1$.

Region	Lower p_T slope ($\times 10^{-1}$)	Higher p_T slope ($\times 10^{-1}$)
Near-side	-1.4 ± 0.9	$+0.4 \pm 0.6$
Away-side	-1.2 ± 0.7	$+0.5 \pm 0.5$
UE	$+0.3 \pm 0.4$	$+0.3 \pm 0.3$

to the near-side. The ratios predicted by DPMJET provide further evidence for this explanation, as the model shows the jet-like ratios are systematically higher than the UE ratios across the entire multiplicity range. However, DPMJET predicts no differences between the near- and away-side Λ/ϕ ratios, possibly due to the missing medium effects in the model.

Most of the slopes presented in Table ?? are compatible with zero, indicating

no dependence on multiplicity for the Λ/ϕ ratios. However, the near- and away-side slopes in the lower momentum bin exhibit a small ($< 2\sigma$) tension with zero, with both slopes being negative. This again supports the $s\bar{s}$ phase-space argument presented above, as at lower momentum the s and \bar{s} are more likely to be less “separated” in phase-space than at higher momentum, making it easier for them to hadronize into the same ϕ . The slopes obtained from DPMJET are again all compatible with zero, and are thus not shown in the table.

Chapter Six: Summary and outlook

This thesis presents the first results using angular correlations to measure jet and non-jet Λ and charged hadron production in p–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. Both the integrated per-trigger yields and the widths on the near- and away-side of the jet are extracted from the azimuthal correlation functions and studied as a function of associated momentum and event multiplicity. A strong dependence on multiplicity is observed for both the near- and away-side yields in the case of the Λ , whereas the charged hadron associated yields exhibit a much smaller (nearly statistically insignificant) dependence. The away-side yields also show a systematically larger increase with multiplicity than the near-side yields for both cases, hinting at modification of the away-side production due to jet-medium interactions. The h- Λ and h-h near-side jet widths reveal a large dependence on p_{T} , becoming more collimated as momentum increases. The widths of the away-side jets are found to be independent of both p_{T} and multiplicity, however the larger systematic uncertainties introduce difficulties with excluding flat behavior. Comparing width values of the h- Λ and h-h correlations, the h- Λ near-side widths are found to be significantly ($> 2\sigma$) larger than the dihadron widths, whereas the away-side widths are consistent within uncertainties.

The per-trigger pair-wise yield ratios $R_i^{\Lambda/h}$ and $R_i^{\Lambda/\phi}$ (i = near-side jet, away-side jet, UE) are also studied as a function of associated p_{T} and multiplicity. The Λ/h ratios exhibit a clear ordering in each region for the entire multiplicity range in both p_{T} bins, with the UE ratios being larger than the away-side ratios, which are larger than the near-side ratios. The Λ/h ratios in each region also reveal a strong dependence on multiplicity, with slopes that are greater than zero by nearly 5σ for both momentum bins. Furthermore, the away-side slopes are found to be systematically higher than the near-side slopes, indicating that the away-side Λ production is more strongly enhanced than the near-side Λ production with increasing multiplicity. The Λ/ϕ ratios in the near-side jet region are measured to be systematically higher than both the away-side and UE ratios, hinting at a suppression of ϕ mesons along the jet axis. While the Λ/ϕ ratios show no significant dependence on multiplicity, the slopes of these ratios in the near- and away-side jet regions in the lower momentum bin show

a small ($< 2\sigma$) deviation from zero, indicating that the ϕ mesons may become less suppressed with increasing multiplicity at low p_T .

All measured observables are compared with predictions from the DPMJET model. The predicted near- and away-side yields are found to be in relatively good agreement with data in the dihadron case, but the h- Λ yield predictions deviate from data by a large ($> 40\%$) margin. DPMJET also fails to predict any of the observed multiplicity dependence for both the h- Λ and h-h jet yields. However, the model is able to closely predict the near-side widths of the dihadron distributions across all multiplicity and momentum ranges, although it underpredicts both the h- Λ near-side widths and the away-side widths for both (Λ , h) cases. The model also predicts a difference between the h- Λ and h-h near-side widths, which is observed in data as well. The per-trigger Λ/h and Λ/ϕ yield ratios are consistently underpredicted by DPMJET, and exhibit no multiplicity dependence. Even still, DPMJET manages to predict the ordering of the Λ/h ratios in each region (UE > away-side jet > near-side jet) and the enhancement of the jet-like Λ/ϕ ratio when compared to the UE region.

Appendix A: Resonance technique for Λ reconstruction

A.1 Introduction

This portion of the appendix is dedicated to describing the analysis procedure for generating the $h\text{-}\Lambda$ correlation distributions using lambdas which are reconstructed using the **resonance technique**, where all proton-pion pairs in an event are combined to form Λ candidates. All of the proton and pion daughter tracks meet the same selection criteria as the tracks used in the V^0 technique, described in Table ???. All in all, the procedure is very similar to the one described in Chapter ??, but with a few key differences that will be highlighted in the following sections.

A.2 Combinatorial background estimation

As Λ baryons reconstructed using the resonance technique will have a much larger combinatorial background than those from the nominal procedure, the final correlation will contain a higher fraction of $h\text{-}(p\pi)$ pairs that need to be removed. While the sideband subtraction technique provides a general procedure for removing these pairs, the signal S and the background B of the Λ invariant mass distribution must be well described. To estimate these quantities, the following techniques were explored:

- **Like-sign $p\pi$ pairs** - Reconstruct the invariant mass of like-sign (LS) $p\pi$ pairs, and scale the like-sign $p\pi$ distribution to the unlike-sign (US) $p\pi$ distribution in a region outside of the Λ signal region.
- **Rotated $p\pi$ pairs** - Reconstruct the invariant mass of US $p\pi$ pairs, but rotate either the pion or proton around the z-axis by π radians, and scale the rotated $p\pi$ distribution to the original US sign $p\pi$ distribution in a region outside of the Λ signal region.

- **Voigtian + polynomial fit** - Perform a standard fitting procedure using a Voigtian distribution for the signal along with a second-order polynomial for the background.

The last technique will be addressed first, as it fails to properly estimate the signal and background in data. To illustrate this, the best possible fits in data are found and the corresponding signal shape is extracted and compared with the signal shape in Monte Carlo using full track reconstruction via GEANT. This comparison is done for the 20-50% multiplicity bin in Figure ???. Note that the background shown in the MC plot is the true combinatorial background, as the $p\pi$ pairs are accessed directly at the generator level to confirm they did not come from a Λ decay.

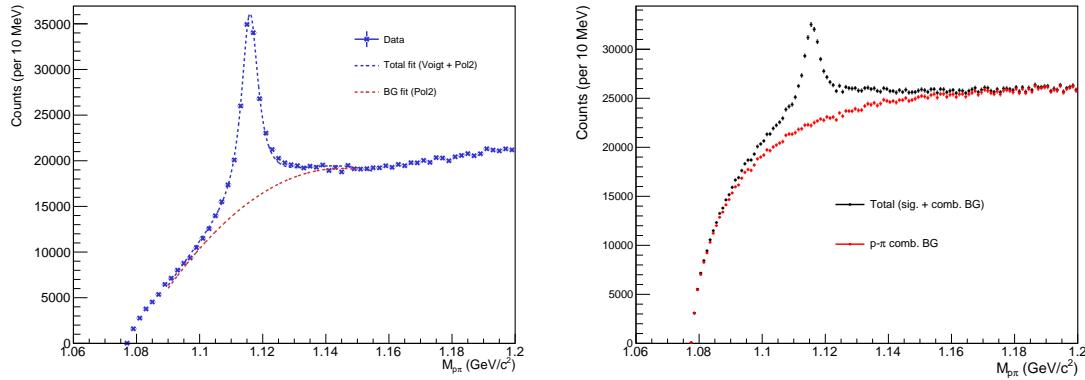


Figure A.1: Left: Invariant mass distribution with corresponding Voigt + Polynomial fit in the 20-50% multiplicity bin (data). Right: The signal and background shapes in MonteCarlo (MC). Note that even though MC appears to have a completely different S/B, the signal shapes should be similar. The fit in data appears to be massively underestimating the Λ signal, as the MC sample indicates there is Λ signal where the total data fit converges with the BG fit.

This plot shows the main issue with reconstructing Λ baryons using the resonance technique: the tails of the signal distribution are much wider than the signal distribution obtained using the V^0 method. This is due to the fact that the kinematics of the corresponding daughter tracks are calculated assuming they originated from the primary vertex, which is only approximately true in the cases where the Λ is short-lived. This is different than the V^0 method, which calculates the kinematics for the daughter tracks assuming they originated from the secondary vertex. The wider tails of the distribution make it extremely difficult to describe using any common

distribution, thus all techniques that rely on fitting the signal shape are not viable. Because of this, only the first two techniques (like-sign and rotated p π pairs) will be considered for the rest of this analysis.

To determine which of the two remaining techniques is more effective, the background shape of the Λ invariant mass distribution for both techniques in MonteCarlo is compared to the ground-truth background shape. The resulting invariant mass distributions from like-sign and rotated p π pairs are shown in Figure ??, along with a comparison of the extracted signal shapes. The LS and rotated p π distributions are scaled to match the US distribution in the sideband region, which will be discussed in the next section. The LS p π pairs match the background shape of the Λ invariant mass distribution more closely than the rotated p π pairs, so they are used to estimate the combinatorial background in the Λ invariant mass distribution in data.

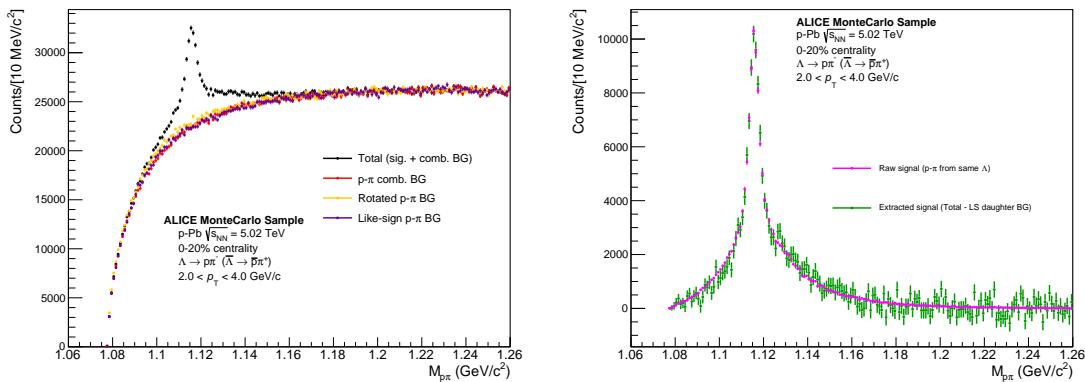


Figure A.2: Left: Invariant mass distribution for reconstructed unlike-sign p π pairs (black) in the MonteCarlo sample. The like-sign p π pair mass distribution (purple) and unlike-sign rotated p π distributions are scaled to match the unlike-sign distribution outside of the Λ signal range. The true combinatorial background (red) matches most closely with the like-sign pairs. Right: The actual Λ signal (magenta) compared with the result of subtracting the like-sign from the total unlike-sign p π distribution (green). The two distributions show good agreement.

A.3 Signal and sideband regions

As the invariant mass distributions from lambdas reconstructed using the resonance technique are very different from those reconstructed using the V⁰ technique, so

too must the signal and sideband regions be different. The signal region was again chosen to maximize significance across all multiplicity bins, and is defined as the range $1.014 < M_{p\pi} < 1.026 \text{ GeV}/c^2$. Choosing the sideband region is a more complicated procedure, as there is no obvious region in the invariant mass distribution where the signal vanishes. Instead, the sideband region is chosen to minimize the difference between the extracted signal in data and the signal shape in MonteCarlo, which can be seen in Figure ???. The resulting sideband region is $1.160 < M_{p\pi} < 1.180 \text{ GeV}/c^2$.

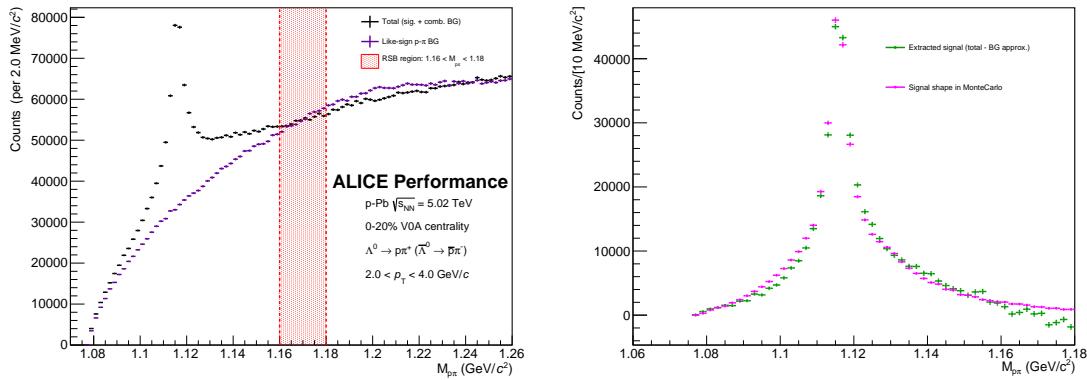


Figure A.3: Left: Invariant Mass distribution for unlike-sign $p\pi$ pairs (black) along with the like-sign $p\pi$ background (purple) and the sideband region (red) in the 0-20% multiplicity bin. Right: The extracted signal (green) compared with the resonance-technique reconstructed signal shape in MonteCarlo (magenta). The sideband region was chosen to minimize the differences between these distributions.

A.3.1 Efficiency correction

Again, the resonance technique-based Λ reconstruction efficiency is calculated in a similar manner as the V^0 technique, using the formula

$$\epsilon(x_1, x_2, \dots, x_n) = \frac{N_{\text{reco.}}(x_1, x_2, \dots, x_n)}{N_{\text{gen.}}(x_1, x_2, \dots, x_n)}, \quad (\text{A.1})$$

where $N_{\text{reco.}}$ and $N_{\text{gen.}}$ are the reconstructed and generated particle distributions with kinematic variables x_1, x_2, \dots, x_n . The main difference from the V^0 efficiency computation comes from $N_{\text{reco.}}$, where each Λ candidate is generated using the following procedure:

- Find all protons and pions within the track list that pass the daughter selection criteria

- For each proton in the list, determine if it came from a Λ (at generator level)
- If the proton came from a Λ , loop through the pion list until the pion that came from the same Λ is found (again, verified at the generator level)
- Reconstruct the Λ using the daughter tracks found in the previous two steps
- Only keep the Λ if $|\eta| < 0.8$

The denominator $N_{\text{gen.}}$ is calculated in the same way as the V^0 technique. The resulting efficiency is shown as a function of p_T for each multiplicity bin in Figure ???. As expected, the efficiency is higher than the V^0 technique, as every Λ reconstructed using the resonance technique would also be reconstructed using the V^0 technique.

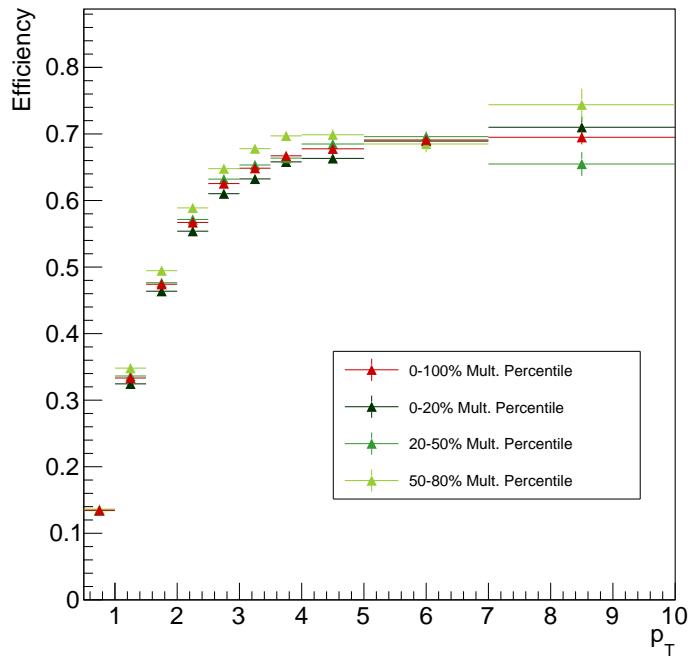


Figure A.4: Efficiency vs. p_T for Λ reconstruction using resonance technique for each multiplicity bin, along with an integrated 0-100% point in red. There does not appear to be any significant dependence on multiplicity. Also worth nothing that the efficiency is higher for this technique when compared to the V^0 technique, as expected (all AOD tracks from V^0 finder daughters are also in total AOD track list).

A.4 Corrections to the h- Λ distributions

All of the efficiency and acceptance corrections are applied to the resonance technique-based h- Λ distribution in the same way as the V⁰ technique. The only difference comes from the removal of the combinatorial background, as:

1. The signal S and background B are calculated in a slightly different manner, and
2. The sideband region is vastly different.

For the first point, the signal and background are calculated via bin-wise summation of the invariant mass distribution using the LS p π pairs as an estimate for the background, scaled to the US distribution in the sideband region.

The second point is mostly inconsequential as the h-p π distributions are very similar in a wide range of sideband regions, as shown in Figure ???. The nominal sideband region was chosen to be $1.160 < M_{\text{SSB}} < 1.180 \text{ GeV}/c^2$, but any region with a lower bound greater than $1.160 \text{ GeV}/c^2$ and upper bound less than $1.22 \text{ GeV}/c^2$ should produce similar results.

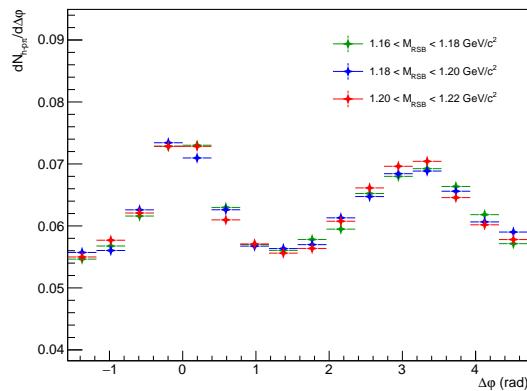


Figure A.5: The projected $\Delta\varphi$ distributions for different choices of sideband, taken within the $-1.2 < \Delta\eta < 1.2$ region. The correlation shapes are identical within the statistical errors.

The signal scaling factor is calculated in the same way as it is in Equation ??, but with the residual now generated by subtracting the sideband-scaled LS p π pairs from the US distribution. The two-track efficiency correction is not applied, as the tools

used to calculate the $\epsilon_{\text{pair}}(\Delta\varphi, \Delta\eta)$ template were not developed before the resonance technique-based analysis was completed.

A.5 MC closure test

An MC closure test was also performed for the resonance technique-based analysis, and the results are shown in Figure ???. The ratio is consistent with unity, but the statistical fluctuations make it difficult to draw any meaningful conclusions.

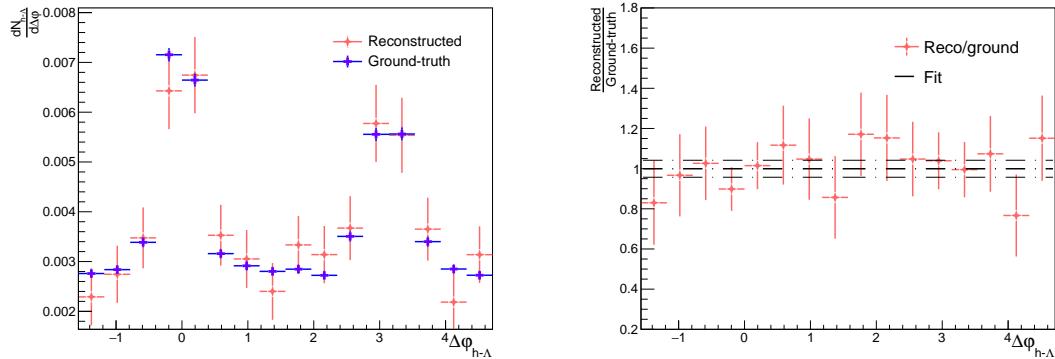


Figure A.6: The reconstructed (red) and ground-truth (blue) $h\text{-}\Lambda_{\text{res}}$ $\Delta\varphi$ distributions along with a (reconstructed)/(ground-truth) ratio and straight-line fit. The fit is technically consistent with unity, but the statistical fluctuations are quite large.

As the reconstructed distribution has not been corrected for the two-track merging effect, it is surprising that the ratio does not exhibit a significant deviation from unity at small $\Delta\varphi$. This is likely due to two factors:

1. The resonance technique has a much lower S/B, and therefore the sideband subtraction introduces a large amount of statistical fluctuations making such deviations difficult to observe, and
2. The reconstructed daughter tracks have a larger fraction of higher quality tracks when compared to the V^0 technique, and those tracks are less likely to be merged over by the trigger during reconstruction.

To elaborate on the second point, while the resonance and V^0 techniques use the same loose quality cuts, the daughter tracks coming from the V^0 technique must

have a resolvable secondary vertex, which biases the corresponding Λ sample to those with a higher decay length. As discussed in ??, the two-track merging effect is more pronounced at larger decay lengths, thus the $h\text{-}\Lambda$ distributions using the V^0 reconstruction technique will have a larger fraction of merged tracks when compared to the resonance technique-based distributions.

To further investigate this surprising closure of the resonance technique-based $h\text{-}\Lambda$ $\Delta\varphi$ distributions, the same closure test is performed, but for the reconstructed $h\text{-}\Lambda$ distribution, the Λ candidate is required to have have a corresponding particle at the generator-level, making the combinatorial background exactly zero and removing the need for sideband subtraction. The results of this test are shown in Figure ???. The ratio is no longer consistent with unity at small $\Delta\varphi$, as expected¹

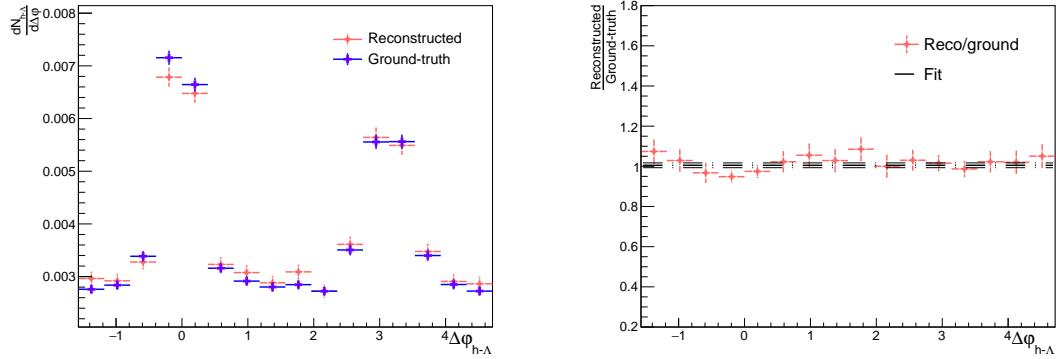


Figure A.7: The reconstructed (red) and ground-truth (blue) $h\text{-}\Lambda_{\text{res}}$ $\Delta\varphi$ distributions along with a (reconstructed)/(ground-truth) ratio and straight-line fit, but instead requiring the reconstructed Λ to have a corresponding particle at the generator level to make sideband subtraction unnecessary. The result is no longer consistent with unity at small $\Delta\varphi$ due to the track merging effect, but the non-closure is much smaller than the V^0 technique.

A.6 Some additional results

A comparison of the final per-trigger $h\text{-}\Lambda$ $\Delta\varphi$ correlation structure from the resonance and V^0 -based techniques was shown in Chapter ??, but it can be seen again in

¹It is strange to *want* non-closure, but it would be even stranger if the track merging effect were somehow not present in the resonance technique-based analysis.

Figure ???. As mentioned previously, the correlation shapes are nearly identical, with the resonance technique having slightly larger uncertainties due to the combinatorial background subtraction.

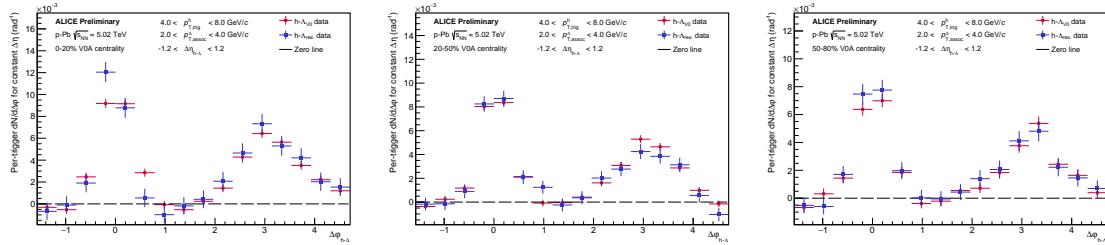


Figure A.8: The final per-trigger $h\text{-}\Lambda$ $\Delta\varphi$ correlations for Λ s reconstructed using the resonance technique (blue) and the V^0 -based technique (red) in the 0-20% (left), 20-50% (middle) and 50-80% (right) multiplicity bins, taken in the associated momentum range $2.0 < p_T < 4.0 \text{ GeV}/c$, after the subtraction of the UE. The distributions show good agreement across all multiplicity bins, indicating that the V^0 -based reconstruction technique is not introducing a bias in the correlation shape.

Additionally, the per-trigger near- and away-side pairwise yields and the $(h\text{-}\Lambda)/(h\text{-}h)$ ratios with Λ s reconstructed using the resonance technique are shown in Figure ???. The results are qualitatively very similar to the nominal results, indicating that the resonance technique is a reasonably viable alternative to the V^0 technique. However, due to the larger combinatorial background (and likely very large systematic uncertainties), the V^0 technique is still the preferred method for Λ reconstruction.

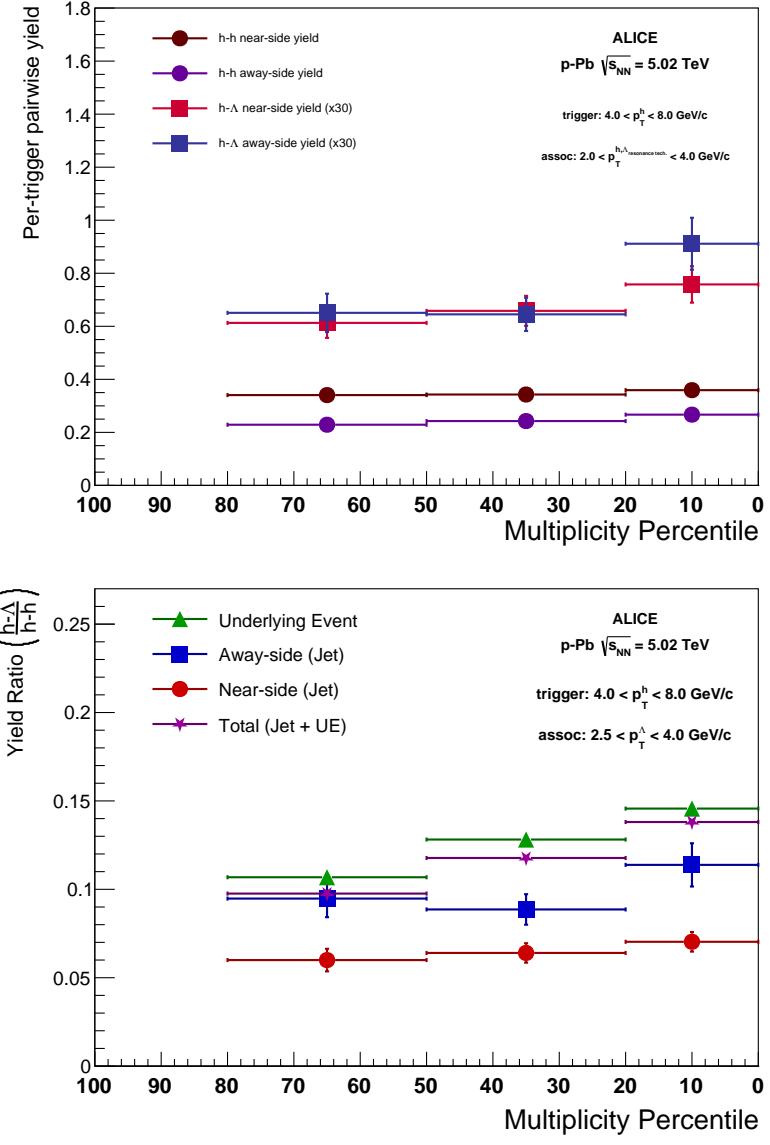


Figure A.9: The final $h\text{-}\Lambda$ and $h\text{-}h$ per-trigger pairwise jet yields (top) and $(h\text{-}\Lambda)/(h\text{-}h)$ yield ratios (bottom) vs. multiplicity in the associated momentum range $2.0 < p_T < 4.0$ GeV/c for Λ s reconstructed using the resonance technique. The general trends are similar to the nominal V^0 technique-based procedure, with larger statistical uncertainties.