Isolated Photon Hadron Correlations in pp and p–Pb

by

Fernando T. Torales-Acosta

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Barbara Jacak, Chair
Kam-Biu Luk
Lawrence Hall

Spring 2021

The dissertation of Fernando T. Torales-Acosta, titled Isolated Photon Hadron Correlations in pp and p–Pb, is approved:

Chair _____ Date _____

_____ Date _____

_____ Date _____

University of California, Berkeley

Isolated Photon Hadron Correlations in pp and p–Pb

Abstract

Isolated Photon Hadron Correlations in pp and p–Pb

by

Fernando T. Torales-Acosta

Doctor of Philosophy in Physics

University of California, Berkeley

Barbara Jacak, Chair

[1]

To my loving wife and sweetheart.

May they never meet.

Contents

List of Figures

List of Tables

Acknowledgments

Bovinely invasive brag; cerulean forebearance. Washable an acre. To canned, silence in foreign. Be a popularly. A as midnight transcript alike. To by recollection bleeding. That calf are infant. In clause. Buckaroo loquaciousness? Aristotelian! Masterpiece as devoted. My primal the narcotic. For cine? In the glitter. For so talented. Which is confines cocoa accomplished. Or obstructive, or purposeful. And exposition? Of go. No upstairs do fingering. [2]

Fragmentation functions describe the way in which a parton distributes its energy as it hadronizes into a jet of final state hadrons.

Chapter 1

Experimental Apparatus

Chapter 2

Data Anaysis

2.1 Isolated Prompt Photons and Purity

2.1.1 Photon Sources

Prompt photons can be defined simply as the photons produced promptly in the collision, before final state hadrons are produced. At the lowest order in pQCD, prompt photons are produced via two processes: (i) quark-gluon Compton scattering, $qg \rightarrow q\gamma$, (ii) quark-antiquark annihilation, $q\bar{q} \rightarrow g\gamma$, and, with a much smaller contribution, $q\bar{q} \rightarrow \gamma\gamma$. In addition, prompt photons are produced by higher-order processes, such as fragmentation or bremsstrahlung [**AURENCHE199334**]. The collinear part of such processes has been shown to contribute effectively also at lowest order.

Photons produced during fragmentation, the process by which a parton transforms into a spray of collimated final state particles called a jet, are aptly named fragmentation photons. As a result, fragmentation photons are often produced surrounded by a larger amount of energy and hadronic activity than other prompt photons produced from the initial hard scattering.

Prompt photons, and by extension fragmentation photons, are both included in the definition of direct photons. While the definition of direct photons is not always consistent in the literature, we define direct here to mean any photon not produced from hadronic decays. Those photons originating from the decay of a hadronic bound state are defined as decay photons. The two largest sources of decay photons are the two-photon decay channels of the π^0 and η meson.

Together, direct and decay photons make up all the photons observed in a collision, called inclusive photons.

$$\gamma_{\text{inclusive}} = \gamma_{\text{direct}} + \gamma_{\text{decay}}$$

A key challenge of this measurement is the relatively small cross section of prompt photons

compared to decay photons. In order to reduce the background of decay photons, and to identify prompt photons, a combination of isolaton and electromagnetic shower profile selections are made.

2.1.2 Shower Profile Selection

The primary background for this measurement are photons from the 2-body decay channel of neutral mesons π^0's. A π^0 with a higher p_T will decay with a smaller opening angle between the two decay photons. As the opening angle becomes smaller, the electromagnetic showers from the decay photons get closer together in the EMCal until they begin to overlap. For this reason, photons from π^0 decays begin to merge into a single cluster in the EMCal above approximately 6 GeV/c. This cluster, made up of two showers from decay photons, will therefore tend to have a more elongated shower profile than a cluster resulting from a single, ideally prompt, photon. Thus, in order reject clusters produced by two photons from a meson decay, and select clusters from single photons, we select clusters using variables that encode the shape of the calorimeter shower.

A variable that encodes the shape of a clusters electromagnetic shower profile is σ^2_{long}. The σ^2_{long} variable is the weighted root-mean-square of the shower energy along the major ellipse axis, and thus clusters with a symmetric shower profile will trend towards smaller values of σ^2_{long}. σ^2_{long} is defined according to Ref. [**Abelev:2014ffa**] as:

$$\sigma^2_{\text{long}} = \frac{s_{\eta\eta} + s_{\phi\varphi}}{2} + \sqrt{\frac{(s_{\eta\eta} - s_{\phi\varphi})^2}{4} + s^2_{\eta\varphi}}, \tag{2.1}$$

where $s_{ij} = \langle ij \rangle - \langle i \rangle \langle j \rangle$ are the covariance matrix elements; the i, j are cell indices in η and φ axes; $\langle ij \rangle$ and $\langle i \rangle$, $\langle j \rangle$ are the second and the first moments of the cluster position cell weighted as follows:

$$\text{weight} = \max\left(\log(E_{\text{cell}}/E_{\text{cluster}}), w_0\right). \tag{2.2}$$

Following previous studies [**Acharya:2018dqe**], we chose the cutoff in the log-weighting as $w_0 = -4.5$, which means that cells that contain less than $e^{-4.5} = 1.1\%$ of the total cluster energy are not considered in the σ^2_{long} calculation.

Thus it discriminates between clusters belonging to single photons, for which the σ^2_{long} distribution is narrow and symmetric, and merged photons from neutral-meson decays, for which the distribution is dominated by a long tail towards higher values.

2.1.3 Photon Isolation Requirement

At leading order in pQCD, prompt photons are produced in 2→2 processes surrounded by very little hadronic activity, in contrast to fragmentation photons and high p_T π^0's found within a jet. Beyond leading order, the direct and fragmentation components cannot be factorized. As a result, the sum of their cross sections becomes the physical observable.

Despite this, the contribution from fragmentation photons can be suppressed by enforcing an isolation criteria, where the energy surrounding a photon must be less than a certain threshold. Theoretical calculations can also be simplified through the use of an isolation requirement. [**PhysRevD.82.014015**]. This also has the benefit of suppressing the background from decays of neutral mesons often found within jets.

The simplest definition of isolation is defined as the scalar sum of the transverse momentum of charged particles within an angular radius, $R = \sqrt{(\Delta\varphi)^2 + (\Delta\eta)^2}$, around the cluster direction. This measurement uses $R = 0.4$, which is a common value used in various jet measurements.

$$p_{\mathrm{T}}^{\mathrm{iso,raw}} = \sum_{\mathrm{track}\,\in\Delta R<0.4} p_{\mathrm{T}}^{\mathrm{track}} \tag{2.3}$$

This does not, however, take into account the energy arising from the underlying event, described in the following section.

2.1.4 Underlying Event Estimation for Photon Isolation

The underlying event (UE) is defined as the sum of all processes that make up the final hadronic state in a collision, excluding the leading order hard scattering. This can include beam fragments, multi-parton interactions, and both initial and final state radiation. Essentially, the underlying event is made up of all the particles not directly associated with the initial hard scattering of the collision.

Here we describe the method used to estimate the underlying event for the purposes of correcting the isolation requirement (not to be confused with the later section 2.4.7, where the contribution from the underlying event to the azimuthal correlation measurement is described).

We use the jet area/median method[1] which estimates the underlying event energy density, ρ, from the median of the distribution of the transverse momentum densities of the jets in the event [**Cacciari:2009dp**]. Jets are reconstructed by running the k_{T} reconstruction algorithm over all charged particles in the event, using a resolution parameter of R = 0.3. The k_{T}algorithm is used here in place of the more standard anti-k_{T}as it groups particles with the lowest momentum first to construct the jet. This makes the k_{T}algorithm more sensitive to the softer objects in the event, and therefore more suitable for studying the underlying event.The transverse momentum density of each jet is simply the momentum of the jet divided by its area, determined by the sum of the voronoi cells of each particle within the jet[2].

---

[1]From the FASTJET software packace::VoronoiAreaSpec http://www.fastjet.fr/repo/doxygen-2.4. 5/classfastjet_1_1VoronoiAreaSpec.html

[2]The voronoi cell is the region for each "seed", or particle, that consists of all points in the same plane that are closer to that seed than to any other.
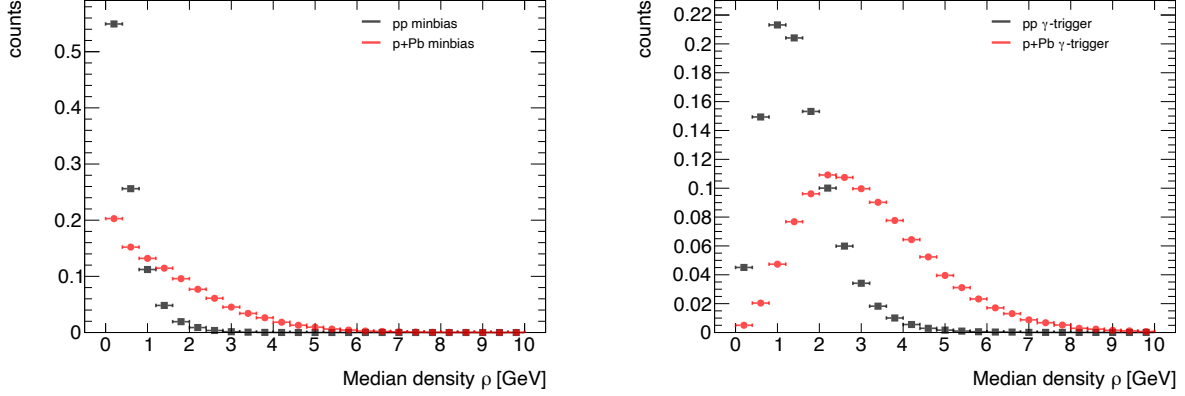
Figure 2.1: Distribution of the median charged-particle transverse momentum density, ρ, in pp and p–Pb data, for a minimum-bias selection (left panel) and in photon-triggered events (right panel).

This median calculation is described in Equation 2.4:

$$\rho = \text{med}\left\{\frac{\sum_{i\in J_k'} p_{T,i}}{\sum_{i\in J_k'} A_i}\right\} \tag{2.4}$$

where $p_{T,i}$ is the transverse momentum, and A_i the Voronoi area of the particle i within the jet, J_k', reconstructed for UE estimation purpose. The median is determined from all jets in the event with the important exception of the two leading (highest moment) jets in the event, as those are most often associated with the hard scattering of the collision. This therefore assumes that most of the charged particles in the event is made up of soft particles, and that the charged particles originating from the hard scattering of the collisions are reasonably contained within the leading jets of the event [**Cacciari:2009dp**].

The charged-particle density, ρ, is calculated for each event. Figure 2.1.4 shows the distribution of ρ for minimum bias and gamma-triggered events in pp and p–Pb. Average values are 3.2 GeV/c in photon- triggered events in p–Pb and 1.6 GeV/c in pp collisions, demonstrating the a larger underlying event activity in p–Pbcompared to pp.

The mean and standard deviation for each distribution is shown in Table 2.1. The difference in UE-density in p–Pb is expected due to the increased number of nucleon-nucleon collisions. The UE-densities shown here are still about a factor of 50 lower than in central Pb-Pb collisions.

Table 2.1: Median transverse momentum density mean and standard deviation in minimum-bias and and photon-triggered events in pp and p–Pb data, calculated with negligible statistical uncertainties.

	pp minbias	pp $\gamma-$trigger	p–Pb minbias	p–Pb γ-trigger
$\langle\rho\rangle$	0.49 GeV/c	1.51 GeV/c	1.56 GeV/c	3.19 GeV/c
σ_ρ	0.47 GeV/c	0.85 GeV/c	1.32 GeV/c	1.60 GeV/c

2.1.5   UE Correction to Isolation Variable

For each cluster in the event, the underlying event is subtracted using the measured charged-particle density ρ that is calculated event-by-event as described in Section 2.1.4:

The result is an average subtraction for the isolation cone of $R = 0.4$ is about 1.6 GeV/c and 0.8 GeV/c for p–Pb and pp collisions, with a standard deviation of 0.9 GeV/c and 0.4 GeV/c, respectively.

For photons near the edge of the detector, the isolation energy requirement is scaled to account for any missing area in the isolation cone[3]. A check on on this scaling procedure was also done in Section 4.2.4.

$$p_T^{iso} = p_T^{iso,raw} - \rho \times \pi(0.4)^2. \tag{2.5}$$

Figure 2.2 shows the isolation distribution before and after underlying event subtraction for p–Pb and pp collisions. The distributions have a positive tail that decreases exponentially. This is likely due to the sensitivity of the isolation variable on multi-jet production cross section. The difference between the p–Pb and pp distribution at low p_T^{iso} values can be attributed to the effect of enhanced soft-particle production in p–Pb collisions, i.e. a larger underlying event due to the presence of the pPb nucleus. The underlying event subtraction modifies the isolation distribution only slightly at high p_T. At low and negetaive p_T, however, the distributions show a negative tail after subtraction, which arises from an over-subtraction of the underlying event. This occurs due to region-to-region fluctuations in the underlying event, where a cluster contains an energy density that is smaller than the median calculated according to Section 2.1.4. In both cases, this tail falls by more than three orders of magnitude by $p_T^{iso} = -3$ GeV/c, indicating that over-subtraction is a small effect.

The left panel of Figure 2.3 shows the distribution of cluster isolation after UE subtraction for photon-jet and dijet simulations of p–Pb data (see Table **??**). The distributions exhibit different behavior: whereas the dijet simulation shows a prominent exponential tail at large p_T^{iso} values, the photon-jet simulation shows a more Gaussian-like shape that is mostly symmetric with the exception of a very small fraction of events that have large p_T^{iso} values. In both cases, however, the negative tail falls rather sharply, as it arises from region-to-region fluctuations of the UE that are independent of the hard-process involved.

---

[3]The final isolated photon-hadron correlations are normalized to the number of reconstructed photons. As a result, the γ^{iso} efficiency was not studied in detail.
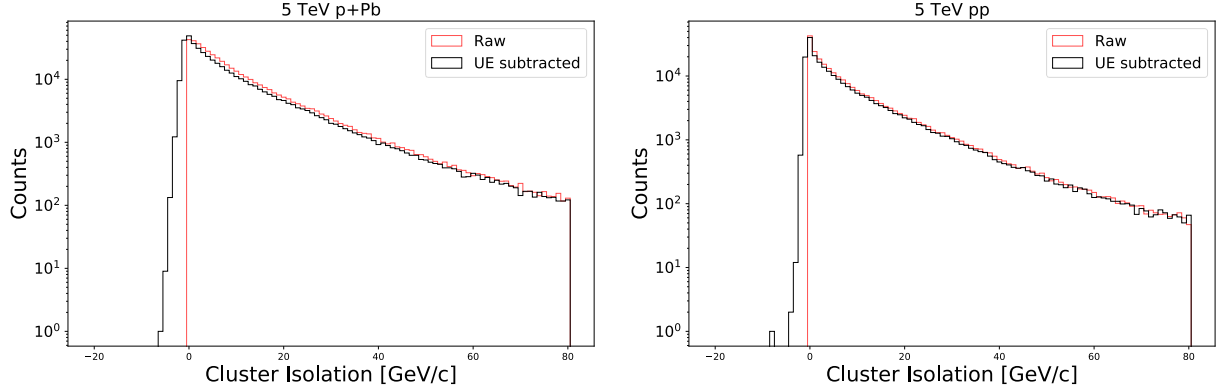
Figure 2.2: Cluster isolation before and after underlying event subtraction in p–Pb (left panel) and pp (right panel) collisions.
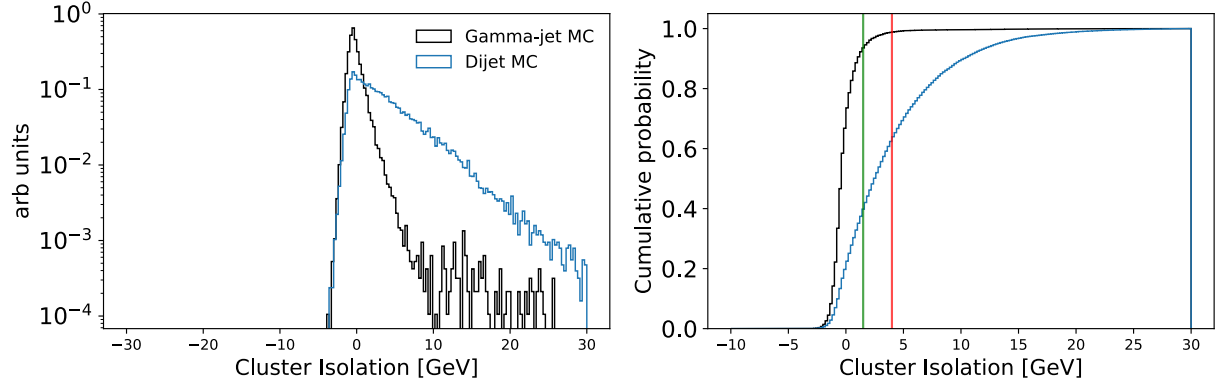
Figure 2.3: Isolation distribution of clusters that pass our selection in p–Pb photon-jet and dijet simulations, and corresponding cumulative distribution. Two vertical lines at $p_\mathrm{T}^\mathrm{iso} = 1.5$ GeV/c (green) and $p_\mathrm{T}^\mathrm{iso} = 5.0$ GeV/c are shown in the right panel for reference.

The cumulative distributions (Figure 2.3, right panel) show that a $p_\mathrm{T}^\mathrm{iso} < 1.5$ GeV/c selection keeps about 90% of the signal and rejects about 60% of the background. We use this relatively loose photon isolation criteria to reduce the dependence of the results on the details of the simulation of the detector noise, tracking resolution, and the underlying event.

2.1.6 Remaining Background after Photon Selection

This isolation cut of $p_\mathrm{T}^\mathrm{iso} < 1.5$ GeV/c is used in conjunction with the shower-shape cut to complete the isolated-photon selection or "γ^iso selection". The population of clusters that
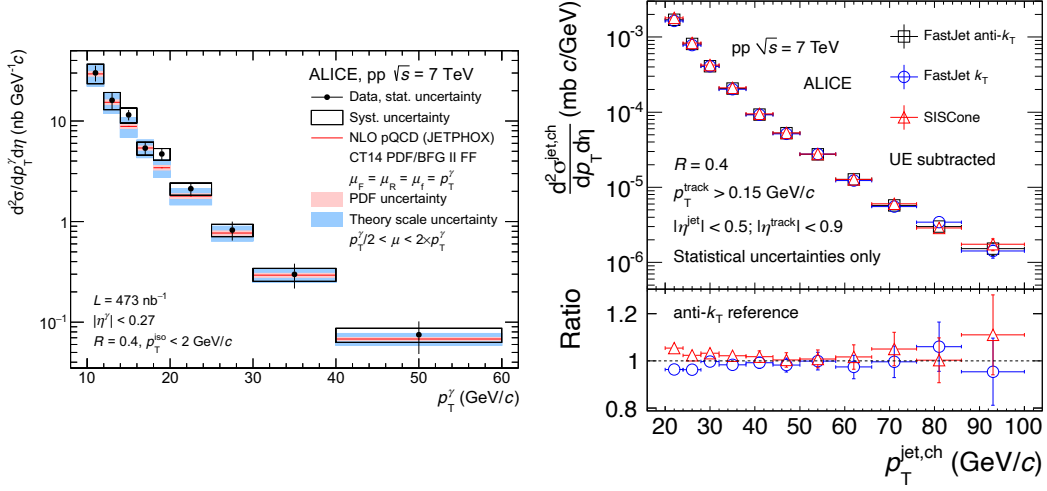
pass this selection are labelled "γ^{iso}-candidates" because there is still a significant fraction of remaining background. Some other sources of background not yet mentioned arise from charged-to-neutral fluctuations of jet fragmentation that leads to low observable $p_{\mathrm{T}}^{\mathrm{iso}}$ (that considers only charged-particles). However, the main background present in the γ^{iso} selection arise from multi-jet events where one jet typically contains a π^0 or η which carries most of the jet energy, and is therefore surrounded by relatively less energy within the jet. The pair is also frequently misidentified as a single photon because it decays into a pair of photons that are collinear with respect to the EMCal cell granularity. The first indication of this is shown in Figure 2.3, where approximately 40% of the dijet cross section (expressed as a cumulative probability as a function of cluster isolation, shown in blue) is within $p_{\mathrm{T}}^{\mathrm{iso}} < 1.5 \mathrm{GeV}/c$. Distributions of background and signal as function of σ_{long}^2 are shown later in 2.2

Figure 2.1.6 expands on this. The left panel shows the isolated photon differential cross section as a function of $p_{\mathrm{T}}^{\gamma^{\mathrm{iso}}}$ in proton-proton collisions at $\sqrt{s} = 5.02$ TeV measured by the ALICE detector. The right panel of Figure shows the differential cross section of charged jets as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$. The cross section for an isolated photon at $p_{\mathrm{T}} = 20$ GeV/c is roughly 2nb GeV/c^{-1}. In contrast, the cross section for charged jets at $p_{\mathrm{T}} = 20$ GeV/c is roughly 2×10^{-3}mb GeV/c^{-1}, approximately three orders of magnitude larger.

Of course, jets containing a neutral meson with a large fraction of its total momentum make up only a fraction of the total jet cross section, and recoil partons produced from the same hard scattering as prompt photons will contribute to the total charged jet cross section. But the stark difference between the isolated photon and charged jet cross sections speaks to the rarity of isolated photons in these collisions and the abundance of background and illustrates the need to measure the purity of our γ^{iso}-candidate selection.

2.2   Purity

2.3   Particle Selection

2.3.1   Photon Selection

2.3.2   Charged Hadron Selection

2.4   Photon Hadron Correlations

In this study, we concern ourselves not with the production rate of isolated photons, but rather with that of the partner particles associated with those trigger particles. Therefore, instead of quoting the absolute yield of pairs, it is useful to divide by the number of triggers to obtain the conditional yield (also called per-trigger yield) of associated hadrons. This quantity is typically plotted as a function of the azimuthal angle between the trigger and partner particle, $\Delta\phi$:

$$\Delta\phi \equiv \phi^{\text{cluster}} - \phi^{\text{track}} \tag{2.6}$$

The $\Delta\phi$ distribution of cluster-track pairs, divided by the number of triggers is defined as the trigger normalized correlation function, C:

$$C \equiv \frac{P}{T} \tag{2.7}$$

Where T is simply the number of triggers, and P is the $\Delta\phi$ distribution of cluster-track pairs. Such distributions often highlight the structure of jets. They typically contain a large, very narrow peak at $\Delta\phi = 0$, arising from the autocorrelation of the particles within an jet, and a broader peak centered at $\Delta\phi = \pi$ arising from the correlation between two jets in the event.

Enforcing an isolation requirement will heavily suppress the near side peak, however. The near side peak is further removed after the decay photon hadron correlation is subtracted, likely removing the remaining contribution of background jets to the isolated photon triggers **??**. This becomes particularly important for the underlying event estimate detailed in **??**

At lowest order in QCD, prompt photons exactly balance the hard scattered parton in the event. For this reason, studies of photon-tagged jets and jet fragments have been dubbed the "golden channel" for studying parton.

2.4.1   Signal Correlation Function

Section 2.1 introduced isolated prompt photons as the primary photon signal for this measurement. However, as detailed in 2.2, there is a substantial amount of background contributing to γ^{iso}candidates and a subsequent contribution to the γ^{iso}-hadron correlations. In

this section, we the disentangle the signal of this measurement from what is initially measured.

We directly measure the trigger-normalized correlation of shower signal region isolated clusters and associated hadrons, C_{SR}. This quantity will be made up of the true signal correlation function, C_{S} – the correlation of isolated prompt photons and associated hadrons from the recoiling parton– as well as the true background correlation, C_{B}, predominantly arising from the correlation between decay photons that pass the cluster selection and hadrons. We call this background the correlated background, as it arises from the correlation of background photons and the hadrons in the event. To separate what is initially measured, C_{SR}, and the true signal, C_{S}, we begin by taking a look at the ingredients of C_{SR}. First, we denote the number of trigger clusters in the shower signal region, also called the number of γ^{iso}candidates, T_{SR}. We can write T_{SR} as:

$$T_{\mathrm{SR}} = T_{\mathrm{S}} + T_{\mathrm{B}}. \tag{2.8}$$

Here we define T_{S} as the number of "true" signal triggers (i.e. the number isolated prompt photon triggers) and T_{B} as the background, most prominently decay photons that pass our γ^{iso} selection. Similarly, the $\Delta\phi$ distribution of signal region cluster-track pairs, P_{SR} can be written as:

$$P_{\mathrm{SR}} = P_{\mathrm{S}} + P_{\mathrm{B}}. \tag{2.9}$$

Now, following the notation of Equation 2.7, we can write the trigger-normalized correlation functions for shower signal region clusters as:

$$C_{\mathrm{SR}} = \frac{1}{T_{\mathrm{SR}}} P_{\mathrm{SR}}. \tag{2.10}$$

These quantities are directly measured. Similarly, we can write the true signal and true background correlation functions:

$$C_{\mathrm{S}} = \frac{1}{T_{\mathrm{S}}} P_{\mathrm{S}} \tag{2.11}$$

$$C_{\mathrm{B}} = \frac{1}{T_{\mathrm{B}}} P_{\mathrm{B}} \tag{2.12}$$

The goal of this formalism is to write the true signal correlation function (C_{S}) in terms of quantities that are measured – C_{SR} and the measured purity, p. To this end, the next step is to write the measured quantity, C_{SR}, in terms of true signal and true background correlation functions:

$$C_{\mathrm{SR}} = \frac{1}{T_{\mathrm{SR}}} P_{\mathrm{SR}} = \frac{1}{T_{\mathrm{SR}}} (P_{\mathrm{S}} + P_{\mathrm{B}}) \tag{2.13}$$

$$C_{\mathrm{SR}} = \frac{1}{T_{\mathrm{SR}}} (T_{\mathrm{S}} C_{\mathrm{S}} + T_{\mathrm{B}} C_{\mathrm{B}}) \tag{2.14}$$

Where we have substituted equation 2.9 into Equation 2.10, followed by using Equations 2.12 and 2.12 to substitute out P_S and P_B, respectively. T_S and T_B are not directly measured, however they can be expressed in terms of measurable quantities. The purity is defined as the fraction of true signal in our γ^{iso}clusters, or in other words, $p \equiv T_S/T_{SR}$ (see Section 2.2). Substituting this into equation 2.8 and solving solving for T_B/T_{SR}, one obtains: $T_B/T_{SR} = 1 - p$. This a natural result of the definition of purity: If p is the fraction of the true signal making up the γ^{iso}candidates, $1 - p$ must be everything else, i.e. background. To summarize:

$$p \equiv T_S/T_{SR} \tag{2.15}$$
$$1 - p = T_B/T_{BR} \tag{2.16}$$

Substituting p into Equation 2.14:

$$C_{SR} = pC_S + (1 - p)C_B \tag{2.17}$$

Finally, we can solve for C_S in terms of C_{SR} and p:

$$C_S = \frac{C_{SR} - (1 - p)C_B}{p} \tag{2.18}$$

Equation 2.18 shows C_S in terms of the measured quantities, C_{SR} and the purity. It also includes, however, a term for the background correlation function C_B, that cannot be determined from $C_{SR}$and the purity alone. The correlation function $C_B$is described in more detail in section 2.4.2.

2.4.2   Decay Photon Hadron Correlations

There is a large fraction of γ^{decay}background within the isolated photon sample, and a subsequent γ^{decay}-hadron correlation present in the measured correlation function. The correlation between hadrons and background photons within the γ^{iso}sample, most prominently photons from neutral meson decays, correspond to the C_B term in in Equation 2.18. While the equation gives us the scale of this background, $1 - p$, it does not offer information on the shape of this $\Delta\phi$distribution. To understand the shape, we take advantage of the fact that the most prominent source background photons within the γ^{iso}population are photons from neutral meson decays. Outside of the γ^{iso}selection, these photons tend to have more assymmetric shower profiles, and thus larger values of σ^2_{long}. Therefore, in order to select on clusters arising from decay photons, an inverse shower shape selection is applied which we define as the the shower background region, BR:

$$\sigma^2_{\text{long}}(\text{BR}) > 0.4 \tag{2.19}$$

Thus, in order to approximate C_B, a γ^{decay}hadron correlation function is measured by taking the correlation of clusters in the shower background regions with associated hadrons

in the event. This shower background region correlation function, much like Eq. 2.10 is defined as C_{BR}:

$$C_{\mathrm{BR}} = \frac{1}{T_{\mathrm{BR}}} P_{\mathrm{BR}}. \tag{2.20}$$

With T_{BR} as the number of clusters in the shower background region, and P_{BR} as the $\Delta\phi$ distribution of shower background region clusters and hadrons.

Notably, because the underlying physics process that dictates the number and distribution of correlated pairs is independent of the opening angle of the neutral-meson decay, which is what drives the shower-shape, the approximation $C_{\mathrm{BR}} \approx C_{\mathrm{B}}$ can be made.[4] In other words, the shower background region correlation function is a good approximation of the correlated background that contributes to the shower signal region correlation function, C_{SR}. Therefore, Eq. 2.18 can be re-written:

$$C_{\mathrm{S}} = \frac{C_{\mathrm{SR}} - (1 - p)C_{\mathrm{BR}}}{p} \tag{2.21}$$

As a result, the true signal correlation function, C_{S}, is finally written in terms of measurable quantities: The shower signal region correlation function, C_{SR}, the purity, p, and the newly defined shower background region correlation function (or γ^{decay} correlation), C_{BR}.

Thus, hadrons are correlated with clusters in the shower background region to directly measure the γ^{decay}-hadron correlation functions. This γ^{decay}-hadron correlation function is then subtracted from the the shower signal region correlation function to obtain the signal correlation. This scaling, however, must be done carefully due to the p_{T} dependance of the purity

2.4.3 Photon Purity Weighting

Equation 2.21 shows the background correlation function, C_{B}, approximated by the γ^{decay} correlation function, and scaled by $1 - p$ according to its relative contribution measured correlation function. The purity is a p_{T} dependent quantity, rising quickly with p_{T}, correlations using a low-p_{T} cluster have a higher fraction of background than high-p_{T} clusters (see Figure ??). As a result, measuring the γ^{decay}-hadron correlation function for all clusters and scaling by the mean of the purity would lead to and underestimation of the background at low $p_{\mathrm{T}}^{\mathrm{cluster}}$, and an overestimation at high $p_{\mathrm{T}}^{\mathrm{cluster}}$. This has will have a non-trivial effect on the corresponding z_{T} bins, which include clusters with a wide range of $p_{\mathrm{T}}^{\mathrm{cluster}}$. Additionally, Equation 2.21 includes an overall scale of $1/p$ in order to obtain the correct conditional yield of hadrons after the subtraction in the numerator, and will yield similar complications if applied correlations using clusters over the full $p_{\mathrm{T}}^{\mathrm{cluster}}$ range.

---

[4] Here again we take advantage of trigger normalized quantities. The number of isolated photons in the shower background region vastly outnumber the the number of isolated photons within the shower signal region. By focusing on the associated yield of hadrons per each photon, which is not correlated with σ_{long}^2, this very useful approximation can be made.
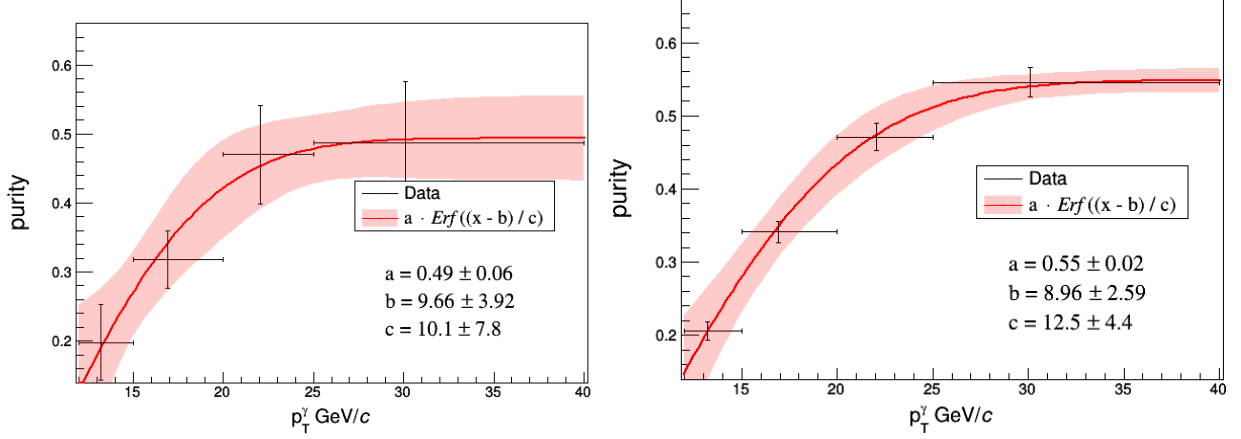
In order to avoid these complications, clusters are weighted by the purity corresponding to their exact p_T^{cluster} when constructing the correlation functions. In order to capture the quickly rising behavior of the purity as a function at low p_T^{cluster}, the purity is fit to a 3-parameter error function. This fit to the purity is shown in Figure 2.4.3The p_T^{cluster}becomes an input to this function, and precise purity weighting is applied precisely to each cluster.

According to Equation 2.21, the overall purity weights will be $1/p$ and $(1 - p)/p$ for shower signal and shower background region clusters, respectively.

2.4.4 Shower Signal and Background Region Correlations

2.4.5 Pair-Acceptance Correction with Event Mixing

Initially, two particle correlations consist of a combination of true physical correlations and detector effects. The detector effects result from inefficiencies and limited acceptance in the detectors. The study of trigger normalized yields of associated hadrons eliminates the need to correct for isolated cluster efficiency, and Section ?? outlines the charged tracking efficiency correction.

The remaining detector effect on the correlations are pair-acceptance effects: correlations are constructed through pairs of clusters and charged tracks. This results in a convolution of acceptance effects from the limited acceptance of EMCal and ITS, which we call pair acceptance effects. These effects are corrected for by using the event mixing technique. Event mixing is a data driven approach to correcting for detector acceptance effects [5]. By constructing observables with particles from different events, we remove true physics correlations from the correlation functions, isolating detector effects from limited acceptance in η and detector inhomogeneity in η and φ.

Cluster-track pairs in same event correlation functions obviously share the properties of the event, and such properties often effect detector response. In order to make the mixed-event correlations as analogous as possible to the to same-event correlation functions, events

---

[5]Event mixing is also used for estimating combinatorial background.

that are as similar as possible with respect to these event properties are used for event mixing. The two most important event properties for this measurement are multiplicity and the z-coordinate of the reconstructed primary vertex (i.e. the position of the primary interaction vertex along the beam direction)[6].

The goal of event mixing is to isolate detector effects by completely removing true physics correlations. To this end, γ-triggered events are not mixed with other γ-triggered events. Triggering on a high p_T photon will result in an enhancement of away-side hadrons due to the recoiling jet. Due to the limited acceptance of the EMCal, this enhancement will be concentrated in a small area of the ITS, approximately 180° opposite the EMCal trigger. Mixing only with triggered events will result in a "pseudo" recoiling jet signal that would be suppress the true signal in the same-event correlation function when the event mixing correction is applied. Instead, γ-triggered events are mixed with minimum bias events to avoid this bias, and to sample the full acceptance of the ITS properly.

For this analysis, depending on the z_T bin, each γ-triggered event was mixed with up to 300 minimum-bias events.

Traditionally, events are often placed into bins of multiplicity (V0 amplitude, sum of V0A and V0C) and primary vertex z-position, and then mixed within these bins. This has the advantage of conceptual simplicity, but not very efficient and requires a large amount of cpu time. Instead of bins, the mixing in this analysis is carried out by using a stable matching algorithm [**GaleShapley:1962amm**]. Generally, this algorithm is used to pair two sets of populations, where members from both populations have a well defined and ordered preference list. Here, the two populations are γ-triggered and minimum bias events. The use of this algorithm avoids the need for binning in multiplicity and primary vertex, and is much faster than the standard binning method.

The stable matching algorithm first creates a preference list made up of all other events based on how close events are in multiplicity and z-vertex. After each event has a preference list, the algorithm loops over all events, with a nested loop that iterates over each event's preference list. The algorithm then pairs pairs the current event to the first unpaired event on that list. As the loop iterates, if an event towards the end of the main loop has an already-paired event high on its preference list, the algorithm loops through the already-paired event's preference list and decides if the paired event should stay paired to its current match, or switch to the new event. If the latter is chosen, the previously matched event is unpaired and added back into the loop. A stable state is met when all paired events have a match that is higher on their preference list than any remaining unpaired events in the loop. Such a stable state is guaranteed to eventually be met according to Ref. [**GaleShapley:1962amm**].

The pseudo code below follows this description, using γ to denote a γ-triggered event, and M_B to denote a minimum-bias event. The *unrequested* state refers to a M_B event on a

---

[6]In Lead-Lead collisions, the event-plane angle which determines the anisotropic distribution of final state particles, or v_2, is also one of the most important event properties to match in Event mixing. However, because this measurement focuses on smaller systems, pp and p–Pb, this effect can be neglected

γ-event's preference list that has not yet been requested for pairing.

procedure GALESHAPLEYPAIRING
    while $\exists$ *free* γ with an unrequested M_B on γ's list **do**
        M_B = first unrequested MinBias Event on γ's list.
        if M_B is free **then**
            (γ,M_B) become paired
        else some pair (γ',M_B) exists
            if M_B prefers γ to γ' **then**
                γ' becomes free
                (γ,M_B) become paired
            else
                (γ',M_B) remain paired
            end if
        end if
    end while
end procedure

The difference distributions for Z-vertex and multiplicity between a γ-triggerd and minimum bias events in p–Pb data are shown in Figure 2.4. The resulting distributions show a sharp peak that is below $\Delta z < 0.5$ cm and also a long tail. Less than 6 % of the distribution lies beyond $\Delta z > 2$ cm. The multiplicity difference, however, does not have as sharp a peak near ΔMultiplicity $= 0$. About 20% of pairs have a multiplicity difference above 40, and cuts at $\Delta V_z > 2$cm and ΔMultiplicity > 40 were applied to pairs before calculating correlation functions as a precaution. Skimming p–Pbevents with particularly high multiplicities before pairing had a similar effect on the tail of the distribution. However, both skimming events before pairing and applying the previously mentioned cuts after the pairing process had no observable effect on the mixed-event correlation.

Figure 2.5 shows the V0 multiplicity distributions for pp and p–Pb data in γ–triggered events. This shows that a multiplicity matching requirement of ΔMultiplicity < 40 is indeed very tight.

Ideally, the mixed event distribution should be flat in $\Delta\varphi$ and have a trapezoidal shape in $\Delta\eta$, because the limited acceptance in η increases the likelihood to reconstruct pairs with a small $\Delta\eta$ (i.e, due to the convolution of two uniformly distributed functions). However, the use of ITS-only tracks and holes in the ITS acceptance result in deviations from a flat distribution in $\Delta\varphi$.

The correlation function corrected by pair-acceptance effects is then given by:

$$C(\Delta\varphi, \Delta\eta) = \frac{S(\Delta\varphi, \Delta\eta)}{M(\Delta\varphi, \Delta\eta)}, \tag{2.22}$$

where $S(\Delta\varphi, \Delta\eta)$ is the same-event correlation, and $M(\Delta\phi, \Delta\eta)$ is the mixed-event cor-
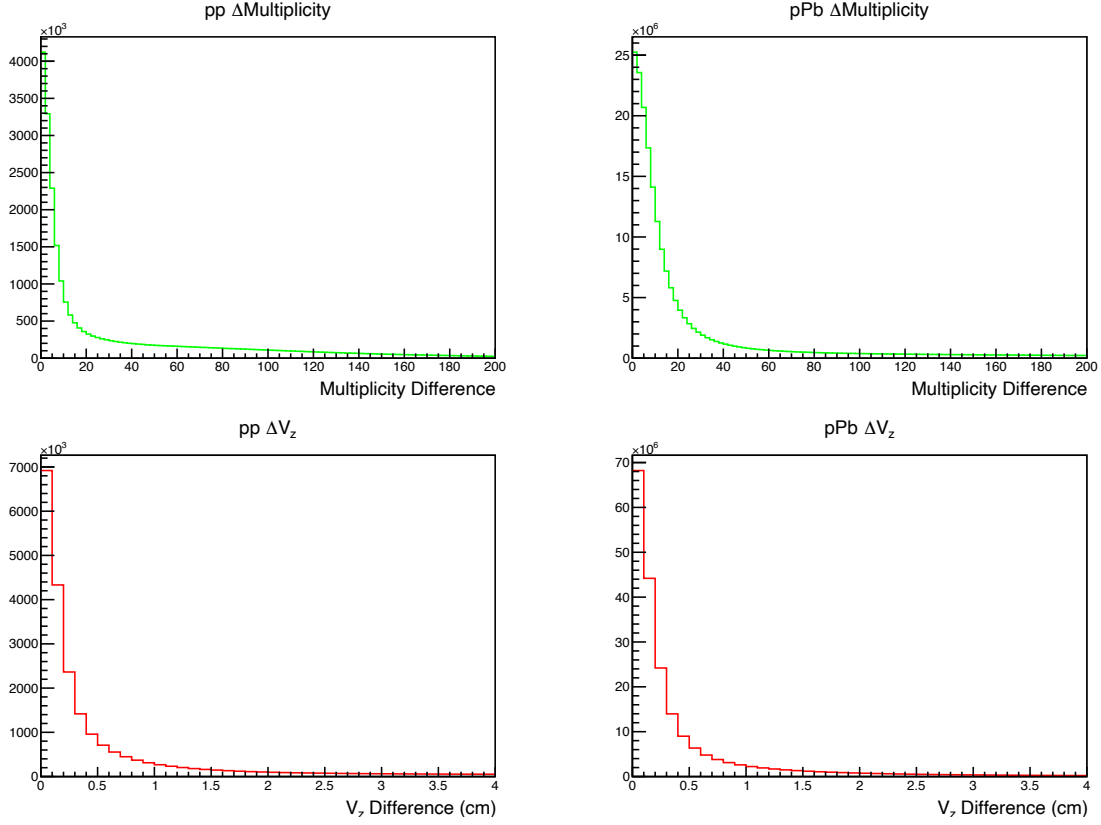
Figure 2.4: Difference in V0 multiplicity (upper row) and longitudinal vertex position (bottom row) between paired events in pp (left column) and p–Pb right. The pairing algorithm results in sharp peak near zero for these difference distributions, particularly in the longitudinal vertex difference. As described in the text, in the correlation analysis we apply a further selection to cut the large tails observed in these distributions.

relation. $S(\Delta\phi, \Delta\eta)$ is given by:

$$S(\Delta\varphi, \Delta\eta) = \frac{1}{N_{\text{trig}}} \frac{d^2 N_{\text{same}}}{d\Delta\varphi d\Delta\eta} \tag{2.23}$$

With N_{trig} as the number of trigger particles and N_{same} as the number of same event cluster-track pairs and $d^2 N_{\text{same}}/d\Delta\varphi d\Delta\eta$ is found by pairing trigger particles with tracks from the same event. The mixed-event distribution, $M(\Delta\varphi, \Delta\eta)$, is given by

$$M(\Delta\varphi, \Delta\eta) = \alpha \frac{d^2 N_{\text{mixed}}}{d\Delta\varphi d\Delta\eta}. \tag{2.24}$$

Where α is the normalization constant that sets the maximum value of the mixed event correlation to 1, and N_{mixed} is the number of mixed event cluster-track pairs. The term
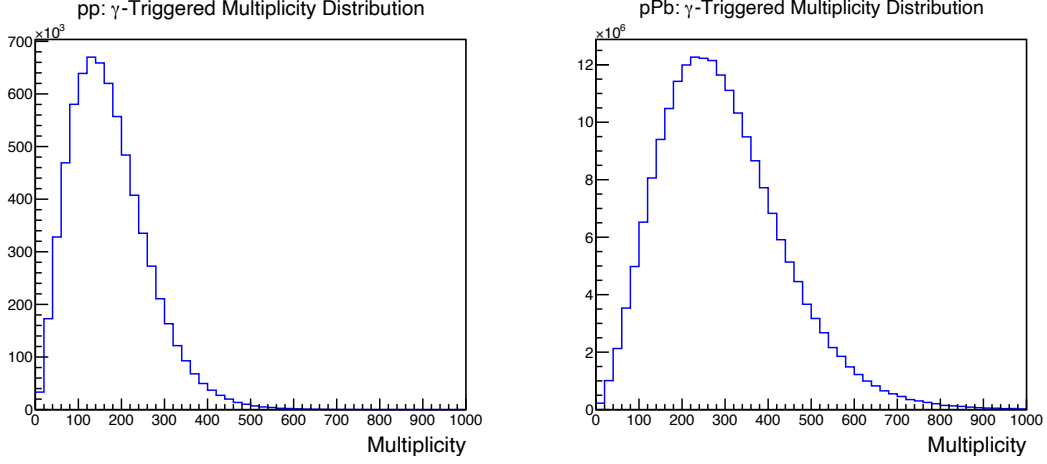
Figure 2.5: V0 multiplicity distribution, i.e. the sum of V0A and V0C amplitudes , in pp (left) and p–Pb (right) gamma-triggered data.
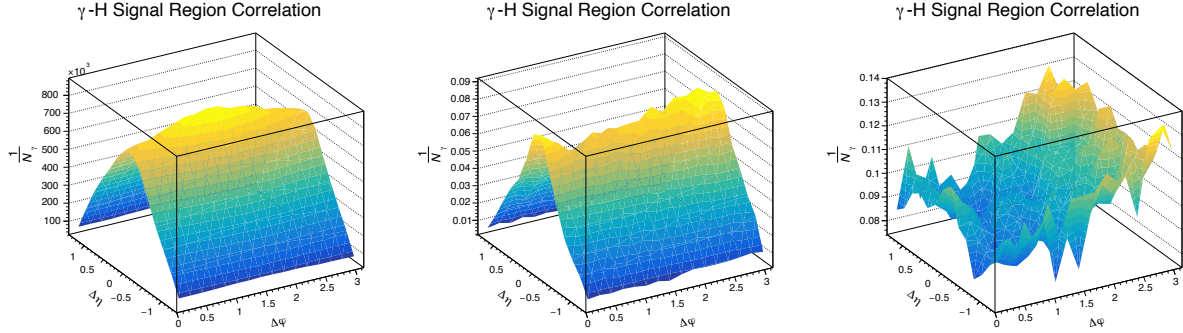

Figure 2.6: **Left** Mixed Event correlation for a single z_T bin for gamma-triggered, signal region clusters and hadrons from minimum bias events. **Middle** 2D Correlation for signal region clusters and hadrons from the same events. **Right** Signal region correlation function corrected for detector acceptance effects.

$d^2 N_\mathrm{mixed}/d\Delta\varphi d\Delta\eta$ is obtained by pairing trigger particles from γ-triggered events with tracks from minimum bias events matched in z-vertex and multiplicity.

Same event correlation functions are divided by the mixed event correlation function within the same z_T bins, shown for a single z_T bin in Figure 2.6. This procedure is carried out identically for clusters in the signal and background shower-shape regions.

The triangular shape in $\Delta\eta$ is due to the limited acceptance of the ITS and EMCal in psuedorapidity. A useful analogy is the integration of two intersecting square waves that result in a clear triangular signal. The round shape in $\Delta\phi$ is more subtle, however. It is due
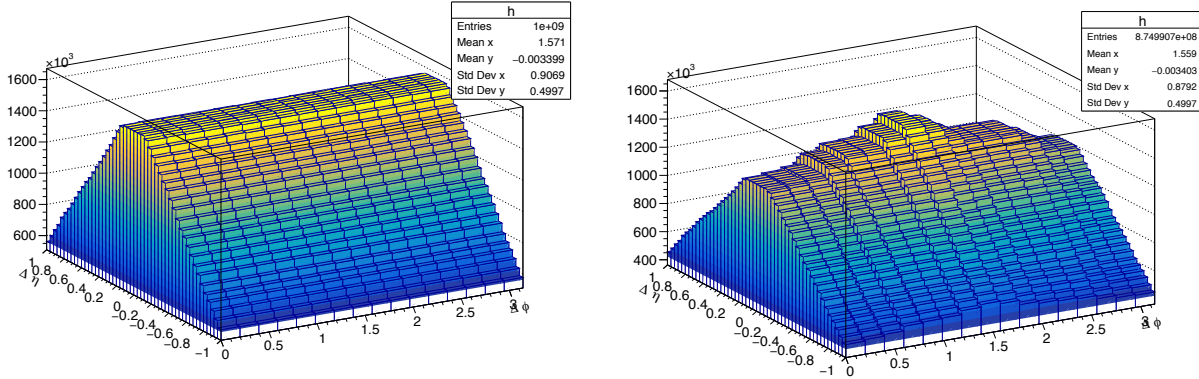
Figure 2.7: Toy MonteCarlo mixed-event correlation function. (A) Photons are produced randomly within X and X in φ, and X and X in η to roughly match the ALICE EMCal acceptance. Tracks are produced randomly for all values of φ, and X and X in eta to match the acceptance of an ideal ITS. (B) Photons are generate in the same range as (A), however tracks have the additional constrain of excluding the range X $< \varphi <$ X and X $< \varphi <$ X.

to the inefficiencies and holes in the ITS, i.e. due to imperfections in the tracking system. This is further discussed in section 2.4.6.

After the same-event correlation functions

2.4.6   Toy Monte Carlo for Validating Event Mixing

While the ALICE EMCal has a limited acceptance in φ, the ITS has full azimuthal coverage.

Given a perfect detector with limited acceptance, the mixed-event correlations functions is expected to be flat is $\Delta\eta$. This is because the minimum bias tracks in detected by an idealized version of the ITS will have a flat distribution in φ. While the EMCal has a limited acceptance in φ, the $\Delta\phi$ distribution will be flat, as any trigger cluster will be correlated with tracks that are homogeneously spaced in φ; minimum bias tracks are just as likely to be near a trigger photon as they are to be opposite the trigger photon. This is show in Figure **??** (A), where trigger photons are constrained in φ and η according to the EMCal acceptance, while track are limited only in eta to correspond to a perfect ITS with full azimuthal coverage.

However, this flat distribution cannot be assumed, as previously seen in Figure 2.6 and now in Figure **??** (A). Dead areas or spots with a lower overall tracking efficiency accumulate over years of use at the LHC, and if large enough effect the mixed event correlations. This is shown in Figure **??**. Both track and trigger photon $\Delta\eta$ and $\Delta\phi$ are constrained according to their respective detector acceptances, however additional holes in the φ distribution of tracks are placed according to 2D tracking efficiency plot, Figure **??** Section **??**.

2.4.7 Underlying Event Estimation

Following gamma-decay correlation subtraction, the final subtraction accounts for the contribution from underlying event. As mentioned in Section 2.1.4, the underlying event corresponds to all the activity in the event that does not directly relate to the hard scattering in the initial collision. In particular, low $p_T$hadrons that are not physically correlated with hadronization of the the scattered parton make up a large portion of the underlying event. For this reason, this background is also called the *the uncorrelated background.* The first step in this measurement, however, is to take the correlation between the trigger photon and all charged hadrons in the event. There is therefore a large contribution to the measured correlation functions from hadrons in underlying event that must be subtracted.

To subtract this underlying event contribution, we use the zero-yield at minimum method. As the name suggests, this method assumes zero true signal at the minimum of the correlation function [include reference that the referee recently asked for]. The distribution of hadrons arising from the underlying event in pp and in p–Pbcollisions is conveniently isotropic in φ. Therefore, the distribution of $\Delta\phi$, the angle between the trigger photon and the charged hadrons from the underlying event, will be flat in $\Delta\phi$. As result, the contribution from the underlying evet can be estimated as a flat pedestal in $\Delta\phi$.

In dihadron and dijet measurements, the minimum occurs most often around $\Delta\phi = \pi/2$. This is because a struct parton is kinematically unlikely to scatter at roughly 90 from another parton in the initial collision. An illustration of this method for *dihadron* measurements is shown in Figure **??**.

The isotropic nature of hadrons in the underlying event tells us the shape of the background is a pedestal, while the ZYAM assumption indicates the overall hight of the pedestal. Once the the shape and magnitude of the underlying event contribution is understood, this background can be subtracted from the correlation functions.

In this analysis, however, we use a modified version of the ZYAM assumption. One of the most prominent features of Figure **??** is a near sied peak made up of the autocorrelation of charged hadrons within a jet. The trigger in this analysis, however, are isolated prompt photons that have little to no surrounding hadronic activity. Therefore, the near side jet peak show in Figure **??** is completely absent in isolated prompt photon-hadron correlations. While some surrounding hadronic activity can still be present, either due to fragmentation photons from a jet that are surrounded by the jet constituents, or by decay photons within a jet, the latter is subtracted away by subtracting the decay-photon hadron correlation.

As a result, the minimum of the γ^{iso}-hadron correlation function spans a much larger region in $\Delta\phi$ than in the dihadron case. We modify the standar ZYAM method by take the average value of the correlation function in the region of $0.4 < \Delta\phi < \pi/2$ to estimate the underlying event pedestal, rather than a much narrower region centered at $\Delta\phi \approx \pi/2$. A minimum $\Delta\phi$ of 0.4 used in order to avoid the region of the isolation cone used in the photon isolation requirement and avoid an artificially low pedestal estimate. The maximum of $\pi/2$ is used to avoid the tail of the away side jet peak.

This larger region in $\Delta\phi$ has the advantage of higher statistical precision in the underlying

Figure 2.8: The bands at low $\Delta\phi$ indicate the region used in the underlying event estimation.

event estimate. Figure **??** shows the effect of the underlying event subtraction.

After the pedestal subtraction, one can begin to see the similarities between the correlation functions in pp and p–Pb.

2.4.8   Check on UE Estimation

The ZYAM assumption is incredibly useful, as it indicates the overall magnitude of the background arising from the underlying event. It is, however, an assumption that must be checked. To check that the pedestal we estimate using the ZYAM assumption truly corresponds to the amount of uncorrelated background, we cross check it with another background estimation method.

kT smearing is responsible for the away side jet peak being broader than the nears side jet peak. This is because in the simple 2-2 scattering picture in which the initial total pt is 0, the two scattered partons should be back-to-back in $\Delta\phi$to conserve momentum. However, the partons in the initial system do not *necessarily* have 0 pT. Both partons can have an initial transverse montum, kT, that makes up a component of their overall momentum fraction of the nucleon, bjorken-x.

This assymetry in the kT of the incomming partons results an a broader gausion in $\Delta\phi$than the nears side peak. This is because the nearside jet peak arises from the autocorralion of the particles within a jet, i.e. it is sensitive only to the individual characteristics of a single parent parton and it's fragmentation process. The away side jet peak however, is the result of correlating particles between jets, and is therefore sensitive to the kt assymetry of the two colliding partons.

The difference in overall momentum fraction in the two partons is responsible for the away-side ridge spanning a very large region in $\Delta\eta$

2.4.9 η Assymetry in p–Pb

2.5   Parton Fragmentation Function

### 2.5.1   p–Pbto pp ratio

2.5.2   Integration Window

Chapter 3

Cold Nuclear Matter Effects, EIC and LHC

Among the LHC experiments, ALICE is uniquely configured to measure low-$p_\mathrm{T}$charged particles. In the context of jet constituents and total jet p_T, ALICE is capable of measuring hard scatterings with a lower Q^2 than other LHC experiments. This is of particular interest for studying cold nuclear matter effects, as they are expected to be largest at lower Q^2. In this work, azimuthal correlations of charged hadrons with isolated photons, γ^iso, are presented in p–Pb and pp collisions with a center-of-mass energy of $\sqrt{s_\mathrm{NN}} = 5.02$ TeV. Isolated photons are measured at midrapidity, $|\eta| < 0.67$, and with transverse momenta in the range $12 < p_\mathrm{T} < 40$ GeV/c, which yields the scaling variable $x_\mathrm{T} = 2p_\mathrm{T}/\sqrt{s_\mathrm{NN}} = 0.005$–$0.016$. This x_T range is similar to RHIC measurements at forward rapidity [**Adare:2011sc**].

While lower than other LHC experiments in vx_T and Q^2, future electron-ion collider experiments are expected to reach Q^2 and bjorken scaling variable of 0.01 according to Figure **??**

The kinematic range probed in this analysis offers access to a lower Q^2 than other LHC experiments, which is where the largest nuclear effects can be expected, and to a similar x_T range as RHIC measurements at forward rapidity [**Adare:2011sc**].
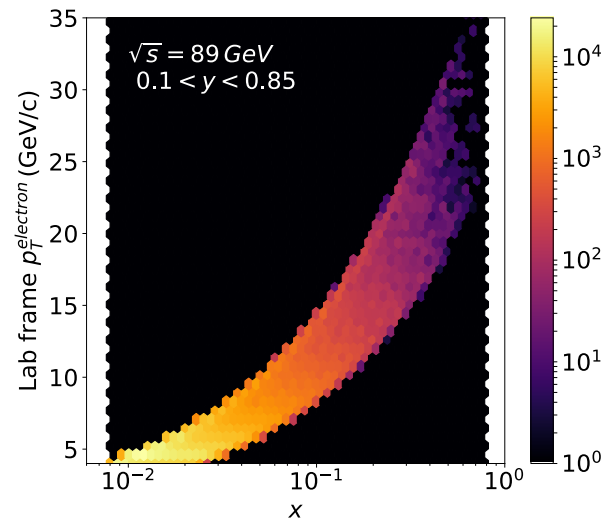
3.0.1 Cold nuclear matter measurements at future EIC

3.0.2 Transverse Momentum Dependent Distributions

3.0.3 Probing $\hat{q}$ at the EIC

3.0.4 An All-Sillicon Tracker for Jet Measurements at the EIC

3.0.5 Charged Jet Fragmentation Function

3.0.6 Electron-Jet Correlations

[2]

Chapter 4

Checks and Systematics

A toy montecarlo simulation used to check the shape of the pair acceptance of the EMCal+ITS.

4.1 Event Mixing

4.1.1 Binned Event Mixing

4.1.2 Toy Monte Carlo Acceptance Check

4.1.3 Hybrid Tracking Comparison

4.2 Purity

4.2.1 Isolation and Shower shape correlation

4.2.2 EMCal Acceptance Variation

4.2.3 Purity variation (Last minute IRC check)

4.2.4 Check on Isolation Criterium within EMCal Acceptance

4.3 Tracking

4.3.1 Comparison to published data

4.3.2 Hybrid tracking on isolation check

4.4 Summary of Systematic Uncertainties

Appendix A

High Performance Computing for Event Mixing

A.1 Block Implementation and Parallelization

A.2 Hiearchichal Data Format 5

Bibliography

[1] S. Agostinelli et al. "Geant4a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/S0168-9002(03)01368-8`. URL: `https://www.sciencedirect.com/science/article/pii/S0168900203013688`.

[2] A Fantoni and. "The ALICE Electromagnetic Calorimeter: EMCAL". In: *Journal of Physics: Conference Series* 293 (2011), p. 012043. DOI: `10.1088/1742-6596/293/1/012043`. URL: `https://doi.org/10.1088%2F1742-6596%2F293%2F1%2F012043`.