# UNIVERSITYOF **BIRMINGHAM**

# University of Birmingham Research at Birmingham

# Particle identification in ALICE

**ALICE Collaboration** 

DOI:

10.1140/epjp/i2016-16168-5

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard): ALICE Collaboration 2016, 'Particle identification in ALICE: a Bayesian approach', European Physical Journal Plus, vol. 131, no. 5. https://doi.org/10.1140/epjp/i2016-16168-5

Link to publication on Research at Birmingham portal

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes

- •Users may freely distribute the URL that is used to identify this publication.
- •Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
  •User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- •Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 26. Apr. 2024

Regular Article

# Particle identification in ALICE: a Bayesian approach

The ALICE Collaboration<sup>a,b</sup>

European Organization for Nuclear Research (CERN), Geneva, Switzerland

Received: 8 February 2016 / Revised: 27 March 2016

Published online: 25 May 2016

© CERN for the benefit of the ALICE Collaboration 2016. This article is published with open access at

Springerlink.com

Abstract. We present a Bayesian approach to particle identification (PID) within the ALICE experiment. The aim is to more effectively combine the particle identification capabilities of its various detectors. After a brief explanation of the adopted methodology and formalism, the performance of the Bayesian PID approach for charged pions, kaons and protons in the central barrel of ALICE is studied. PID is performed via measurements of specific energy loss (dE/dx) and time of flight. PID efficiencies and misidentification probabilities are extracted and compared with Monte Carlo simulations using high-purity samples of identified particles in the decay channels  $K_S^0 \to \pi^-\pi^+$ ,  $\phi \to K^-K^+$ , and  $A \to p\pi^-$  in p-Pb collisions at  $\sqrt{s_{\rm NN}} = 5.02\,{\rm TeV}$ . In order to thoroughly assess the validity of the Bayesian approach, this methodology was used to obtain corrected  $p_{\rm T}$  spectra of pions, kaons, protons, and  $D^0$  mesons in pp collisions at  $\sqrt{s} = 7\,{\rm TeV}$ . In all cases, the results using Bayesian PID were found to be consistent with previous measurements performed by ALICE using a standard PID approach. For the measurement of  $D^0 \to K^-\pi^+$ , it was found that a Bayesian PID approach gave a higher signal-to-background ratio and a similar or larger statistical significance when compared with standard PID selections, despite a reduced identification efficiency. Finally, we present an exploratory study of the measurement of  $A_c^+ \to pK^-\pi^+$  in pp collisions at  $\sqrt{s} = 7\,{\rm TeV}$ , using the Bayesian approach for the identification of its decay products.

### 1 Introduction

Particle Identification (PID) provides information about the mass and flavour composition of particle production in high-energy physics experiments. In the context of ALICE (A Large Ion Collider Experiment) [1], identified particle yields and spectra give access to the properties of the state of matter formed at extremely high energy densities in ultra-relativistic heavy-ion collisions. Modern experiments usually consist of a variety of detectors featuring different PID techniques. Bayesian approaches to the problem of combining PID signals from different detectors have already been used by several experiments, e.g. NA27 [2], HADES [3] and BESII [4]. This technique was proposed for ALICE during its early planning stages [5] and then used extensively to prepare the ALICE Physics Performance Report [6,7]. An analogous method is also used to combine the PID signals from the different layers of the ALICE Transition Radiation Detector.

The ALICE detector system is composed of a central part that covers the mid-rapidity region  $|\eta| < 1$  (the "central barrel"), and a muon spectrometer that covers the forward rapidity region  $-4 < \eta < -2.5$ . The central barrel detectors that have full coverage in azimuth  $(\varphi)$  are, from small to large radii, the Inner Tracking System (ITS), the Time Projection Chamber (TPC), the Transition Radiation Detector (TRD)<sup>1</sup> and the Time Of Flight system (TOF). Further dedicated PID detectors with limited acceptance in  $\varphi$  and  $\eta$  are also located in the central barrel. These are the Electromagnetic Calorimeter (EMCal), the Photon Spectrometer (PHOS), and a Cherenkov system for High-Momentum Particle Identification (HMPID).

<sup>&</sup>lt;sup>a</sup> e-mail: alice-publications@cern.ch

<sup>&</sup>lt;sup>b</sup> See appendix A for the list of collaboration members.

<sup>&</sup>lt;sup>1</sup> The TRD was completed during the first Long Shutdown phase (2013–2015) and only had partial  $\varphi$  coverage during the Run 1 data taking period (2009–2013).

The central barrel detectors provide complementary PID information and the capability to separate particle species in different momentum intervals. At low momenta ( $p \lesssim 3\text{--}4\,\mathrm{GeV/}c$ ), a track-by-track separation of pions, kaons and protons is made possible by combining the PID signals from different detectors. At higher momenta, statistical unfolding based on the relativistic rise of the TPC signal can be performed for PID. Given the wide range of momenta covered, ALICE has the strongest PID capabilities of any of the LHC experiments. More details about the identification possibilities of the single detectors can be found in [1].

Several different PID methods were applied for analyses of data collected by ALICE during the Run 1 data taking period of the LHC (2009–2013). A non-exhaustive list of examples of PID in ALICE is given in the following. Results were published on the transverse momentum  $(p_{\rm T})$  distributions of charged pion, kaon and proton production [8–10] in different collision systems and centre-of-mass energies using the ITS, TPC, TOF and HMPID detectors. Electron measurements from semileptonic heavy-flavour hadron decays took advantage of the TPC, TOF, TRD and EMCal detectors [11,12]. Neutral pion production was studied via photon detection in the PHOS and the detection of  $e^+e^-$  pairs from gamma conversions in the TPC. PID detectors were also used extensively to improve the signal-to-background ratios when studying the production of certain particles based on the reconstruction of their decay products, such as D mesons [13],  $\phi$  and K\* resonances [14]; in studies of particle correlations, such as femtoscopy [15]; and to identify light nuclei [16]. In all analyses where PID was used, selections were applied based on individual detector signals and later combined.

In this paper we describe results obtained during Run 1 in pp collisions at  $\sqrt{s} = 7 \,\text{TeV}$ , Pb-Pb collisions at  $\sqrt{s_{\text{NN}}} = 2.76 \,\text{TeV}$  and p-Pb collisions at  $\sqrt{s_{\text{NN}}} = 5.02 \,\text{TeV}$ . In particular, we focus on the hadron identification capabilities of the central barrel detectors that had full azimuthal coverage during Run 1. These are the ITS, TPC and TOF detectors, which are described in more detail below. The other central barrel detectors are not discussed here.

The ITS is formed of six concentric cylindrical layers of silicon detectors: two layers each of Silicon Pixel (SPD), Silicon Drift (SDD) and Silicon Strip Detectors (SSD). The SDD and the SSD provide a read-out of the signal amplitude, and thus contribute to the PID by measuring the specific energy loss (dE/dx) of the traversing particle. A truncated mean of the (up to) four signals is calculated, resulting in a relative dE/dx resolution of approximately 12% [1].

The TPC [17] is the main tracking device of ALICE. Particle identification is performed by measuring the specific energy loss (dE/dx) in the detector gas<sup>2</sup> in up to 159 read-out pad rows. A truncated mean that rejects the 40% largest cluster charges is built, resulting in a Gaussian dE/dx response. The dE/dx resolution ranges between 5–8% depending on the track inclination angle and drift distance, the energy loss itself, and the centrality in p-Pb and Pb-Pb collisions due to the differing detector occupancy.

The TOF detector [18] is based on Multigap Resistive Plate Chamber technology. It measures the flight times of particles with an intrinsic resolution of  $\sim 80\,\mathrm{ps}$ . The expected flight time for each particle species is calculated during the reconstruction, and then PID is performed via a comparison between the measured and expected times.

Other forward-rapidity detectors are also relevant to the data analyses presented in this paper. The V0 plastic scintillator arrays, V0A (covering  $2.8 < \eta < 5.1$ ) and V0C ( $-3.7 < \eta < -1.7$ ), are required in the minimum-bias trigger. The V0 signals are also used to determine the centrality of p-Pb and Pb-Pb collisions. The T0 detector is a quartz Cherenkov detector that is used for start-time estimation with a resolution of  $\sim 40\,\mathrm{ps}$  in pp collisions and  $20-25\,\mathrm{ps}$  in Pb-Pb collisions.

A thorough understanding of the detector response is crucial for any particle identification method. In the case of dE/dx, this means a well parameterised description of the average energy loss (according to the Bethe-Bloch formula [19]) and a reliable estimate of the signal resolution. For the TOF measurement, the start-time information, start-time resolution, track reconstruction resolution, and intrinsic detector resolution need to be known [18].

Combining the PID signals of the individual detectors using a Bayesian approach [20] makes effective use of the full PID capabilities of ALICE. However, using combined probabilities to perform particle identification may result in unintuitive, non-trivial track selections. It is therefore important to benchmark the Bayesian PID method, compare efficiencies in data and Monte Carlo, and validate that this technique does not introduce a systematic bias with respect to previously published results. This paper focuses on the verification of the Bayesian PID approach on the basis of various data analyses.

Section 2 describes the Bayesian approach (2.1), the definitions of efficiency and contamination in the context of PID (2.2), the extraction and application of prior distributions (2.3), and different strategies for using the resulting probabilities (2.4). Section 3 presents benchmark analyses of high-purity samples of pions, kaons, and protons from the two-prong decays of  $K_S^0$  mesons,  $\phi$  mesons and  $\Lambda$  baryons. Section 4 presents validations of the Bayesian approach for two full analyses: the measurement of the transverse momentum spectra of pions, kaons and protons (4.1), and the analysis of  $D^0 \to K^-\pi^+$  (4.2). Section 5 illustrates the application of the Bayesian PID approach to maximise the statistical significance when analysing the production of  $\Lambda_c^+$  baryons. Finally, a conclusion and outlook are given in sect. 6.

 $<sup>^2\,</sup>$  A Ne-CO2-based gas mixture of  $\sim 90\,m^3$  in Run 1.

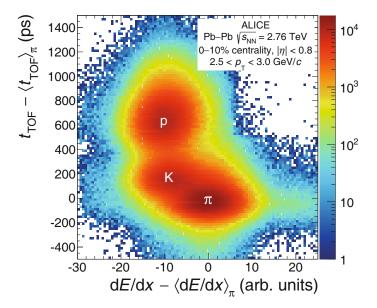


Fig. 1. Combined particle identification in the TPC and TOF for data from Pb-Pb collisions at  $\sqrt{s_{\rm NN}} = 2.76\,{\rm TeV}$ , shown as a two-dimensional plot. The PID signals are expressed in terms of the deviation from the expected response for pions in each detector.

## 2 Bayesian PID in ALICE

Simple selections based on the individual PID signals of each detector do not take full advantage of the PID capabilities of ALICE. An example of this is illustrated in fig. 1, which shows the separation of the expected TPC and TOF signals for pions, kaons and protons with transverse momenta in the range  $2.5 < p_T < 3 \, \mathrm{GeV/}c$ . Clearly, the separation in the two-dimensional plane (the peak-to-peak distance) of, e.g., pions and kaons is larger than the separation of each individual one-dimensional projection. A natural way of combining the information of independent detectors is to express the signals in terms of probabilities. An additional advantage of this method is that detectors with non-Gaussian responses can also be included in a straightforward way. A Bayesian approach makes use of the full PID capabilities by folding the probabilities with the expected abundances (priors) of each particle species. This section outlines the standard " $n_{\sigma}$ " PID approach, before presenting the method used to combine the signals from the different detectors when adopting a Bayesian PID approach.

## 2.1 PID signals in ALICE

The response of each detector can be expressed in terms of its raw signal, S. One of the simplest ways of performing PID is to directly select based on S. Examples of S include the flight-time information from the TOF detector,  $t_{\text{TOF}}$ , and the specific energy loss dE/dx in detector gas or silicon, measured by the TPC and the ITS, respectively. A more advantageous approach would be to use a discriminating variable,  $\xi$ , which makes use of the expected detector response R,

$$\xi = f(S, R),\tag{1}$$

where R can have functional dependences on the properties of the particle tracks, typically the momentum p, the charge Z, or the track length L. The detector response functions are usually complex parameterisations expressing an in-depth knowledge of subtle detector effects.

For a detector with a Gaussian response, R is given by the expected average signal  $\hat{S}(H_i)$  for a given particle species  $H_i$  and the expected signal resolution  $\sigma$ . The index i usually refers to electrons, muons, pions, kaons or protons, but may also include light nuclei such as deuterons, tritons, <sup>3</sup>He nuclei and <sup>4</sup>He nuclei. The most commonly used discriminating variable for PID is the  $n_{\sigma}$  variable, defined as the deviation of the measured signal from that expected for a species  $H_i$ , in terms of the detector resolution:

$$n_{\sigma_{\alpha}^{i}} = \frac{S_{\alpha} - \hat{S}(H_{i})_{\alpha}}{\sigma_{\alpha}^{i}}, \qquad (2)$$

where  $\alpha = (\text{ITS}, \text{TPC}, \ldots)$ . The resolution is given here as  $\sigma_{\alpha}^{i}$ , as it depends both on the detector and the species being measured. In the following,  $\sigma_{\alpha}^{i}$  is simply referred to as " $\sigma$ ".

The  $n_{\sigma}$  PID approach corresponds to a "true/false" decision on whether a particle belongs to a given species. A certain identity is assigned to a track if this value lies within a certain range around the expectation (typically 2 or  $3\sigma$ ). Depending on the detector separation power, a track can be compatible with more than one identity.

For a given detector  $\alpha$  with a Gaussian response, it is possible to define the conditional probability that a particle of species  $H_i$  will produce a signal S as

$$P(S|H_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}n_{\sigma}^2} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(S-\hat{S}(H_i))^2}{2\sigma^2}}.$$
 (3)

In the case of a non-Gaussian response, the probability is described by an alternative parameterisation appropriate to the detector. The advantage of using probabilities is that the probabilities from different detectors,  $P_{\alpha}$ , with and without Gaussian responses, can then be combined as a product:

$$P(\vec{S}|H_i) = \prod_{\alpha = \text{ITS,TPC,...}} P_{\alpha}(S_{\alpha}|H_i), \tag{4}$$

where  $\vec{S} = (S_{\text{ITS}}, S_{\text{TPC}}, \ldots)$ .

The probability estimate  $P(\vec{S}|H_i)$  (either for a single detector, or combined over many) can be interpreted as the conditional probability that the set of detector signals  $\vec{S}$  will be seen for a given particle species  $H_i$ . However, the variable of interest is the conditional probability that the particle is of species  $H_i$ , given some measured detector signal (i.e.  $P(H_i|\vec{S})$ ). The relation between the two for a combined set of detectors can be expressed using Bayes' theorem [20]:

$$P(H_i|\vec{S}) = \frac{P(\vec{S}|H_i)C(H_i)}{\sum_{k=e,\mu,\pi,\dots} P(\vec{S}|H_k)C(H_k)}.$$
 (5)

Here,  $C(H_i)$  is an a priori probability of measuring the particle species  $H_i$ , also known as the prior, and the conditional probability  $P(H_i|\vec{S})$  is known as the posterior probability.

The priors (which are discussed in more detail in sect. 2.3) serve as a "best guess" of the true particle yields per event. When such a definition is adopted for the priors, a selection based on the Bayesian probability calculated with eq. (5) then corresponds to a request on the purity (defined as the ratio between the number of correctly identified particles and the total selected). Additionally, priors can be used to reject certain particle species that are not relevant to a given analysis. Most commonly in the context of the ALICE central barrel, the prior for muons is set to zero. Due to the similarity between the pion and muon mass, the two species are almost indistinguishable over a broad momentum range; the efficiency of detecting a pion would thus be reduced if muons were not neglected. At the same time, this influences the number of particles wrongly identified as pions, since the true abundance of muons (roughly 2% of all particles, estimated using Monte Carlo simulations) is neglected. This case is further discussed in the following sections.

#### 2.2 PID efficiency and contamination

In order to obtain the physical quantity of interest (typically a cross section or a spectrum) from a raw yield, it is necessary to (a) compute the efficiency due to other selections applied before PID, and (b) compute the efficiency of the PID strategy. The PID efficiency is defined as the proportion of particles of a given species that are identified correctly by the PID selections. Both kinds of efficiency are usually estimated via Monte Carlo techniques. To precisely compute the efficiency of a given PID strategy, it is of utmost importance that an accurate description of the actual signals present in the data is provided by the Monte Carlo simulation. Special care must be given to the "tuning" of Monte Carlo simulations to reproduce all of the features and dependences observed in data for PID signals.

It is possible to define a  $\epsilon_{\text{PID}}$  matrix that contains the probability to identify a species i as a species j. If only pions, kaons and protons are considered, the  $3 \times 3$   $\epsilon_{\text{PID}}$  matrix is defined as

$$\epsilon_{\text{PID}} = \begin{pmatrix} \epsilon_{\pi\pi} & \epsilon_{\pi K} & \epsilon_{\pi p} \\ \epsilon_{K\pi} & \epsilon_{KK} & \epsilon_{Kp} \\ \epsilon_{p\pi} & \epsilon_{pK} & \epsilon_{pp} \end{pmatrix}, \tag{6}$$

where the diagonal elements  $\epsilon_{ii}$  are the PID efficiencies, and the non-diagonal elements  $(\epsilon_{ij}, i \neq j)$  represent the probability of misidentifying a species i as a different species j.

The abundance vectors for pions, kaons and protons are defined as

$$\vec{A}_{\text{meas}} = \begin{pmatrix} \pi_{\text{meas}} \\ K_{\text{meas}} \\ p_{\text{meas}} \end{pmatrix} \quad \text{and} \quad \vec{A}_{\text{true}} = \begin{pmatrix} \pi_{\text{true}} \\ K_{\text{true}} \\ p_{\text{true}} \end{pmatrix},$$
 (7)

where the elements of  $\vec{A}_{\text{meas}}$  ( $\vec{A}_{\text{true}}$ ) represent the measured (true) abundances of each species. The diagonal elements  $\epsilon_{ii}$  of the PID matrix are then defined as

$$\epsilon_{ii} = \frac{N_i \text{ identified as } i}{A_{\text{true}}^i} \,. \tag{8}$$

Techniques to estimate the matrix elements are discussed in sect. 3. The abundance vectors  $\vec{A}_{\text{meas}}$  and  $\vec{A}_{\text{true}}$  are linked by the following relation:

$$\begin{pmatrix}
\pi_{\text{meas}} \\
K_{\text{meas}} \\
p_{\text{meas}}
\end{pmatrix} = \begin{pmatrix}
\epsilon_{\pi\pi} & \epsilon_{\pi K} & \epsilon_{\pi p} \\
\epsilon_{K\pi} & \epsilon_{KK} & \epsilon_{Kp} \\
\epsilon_{p\pi} & \epsilon_{pK} & \epsilon_{pp}
\end{pmatrix}^{\mathsf{T}} \cdot \begin{pmatrix}
\pi_{\text{true}} \\
K_{\text{true}} \\
p_{\text{true}}
\end{pmatrix}.$$
(9)

Inverting the  $\epsilon_{\mathrm{PID}}^{\intercal}$  matrix, the physical quantities are then extracted via

$$\vec{A}_{\text{true}} = (\epsilon_{\text{PID}}^{\dagger})^{-1} \times \vec{A}_{\text{meas}}.$$
 (10)

The  $\epsilon_{\text{PID}}$  matrix elements have functional dependences on many variables, primarily  $p_{\text{T}}$  and collision system. Other second-order dependences (for example pseudorapidity, event multiplicity, and centrality) can also be studied depending on the specific track selections and PID strategies applied.

The contamination of the species j due to a different species i ( $c_{ji}$ ) is the number of particles belonging to species i that are wrongly identified as j ( $N_{i \text{ identified as } j}$ ), divided by the total number of identified j particles ( $A_{\text{meas}}^{j}$ ) [7],

$$c_{ji} = \frac{N_{i \text{ identified as } j}}{A_{\text{meas}}^{j}}, \quad i \neq j.$$
(11)

Contamination should not be confused with the misidentification probabilities defined in eq. (6), which do not depend on the real abundances. The connection between contamination and misidentification is

$$c_{ji} = \frac{\epsilon_{ij} A_{\text{true}}^i}{\epsilon_{jj} A_{\text{true}}^j + \sum_{i \neq k} \epsilon_{jk} A_{\text{true}}^k} \,. \tag{12}$$

There is usually a trade-off between efficiency and contamination. Both depend on the PID strategy (e.g. choice of detectors), the detector response, and the priors used, all of which are momentum-dependent. The contamination is additionally driven by the real abundances.

An accurate estimate of the contamination  $c_{ji}$  depends on the real abundances, and it must be determined with data-driven techniques (or with Monte Carlo simulations with abundances corresponding to what is found in the data).

Although the PID matrix elements are independent of the real abundances, they still depend on the choice of priors. They are therefore evaluated consistently as long as the same set of priors is used both in the data analysis and for the Monte Carlo simulations, provided that the detector responses are simulated correctly. In addition, a choice of priors that lies closer to the true abundances allows the best compromise to be found between the maximisation of the efficiency and the minimisation of the contamination probabilities. Considering that any systematic uncertainties in the detector response will be amplified if the efficiency is low or if the contamination probabilities are large, the method becomes more effective as the priors tend closer to reality. An extreme example of this would be the identification of pions and muons using equal priors for all species, as the two species are difficult to distinguish from each other ( $\varepsilon_{\pi\pi} \approx \varepsilon_{\mu\mu} \approx 0.5$  and  $\varepsilon_{\pi\mu} \approx \varepsilon_{\mu\pi} \approx 0.5$ ). In such a case, even a small discrepancy in the description of the detector response in Monte Carlo would cause a fluctuation in  $\varepsilon_{ij}$ , leading to a large uncertainty in the estimate of the pion yield; a choice of priors corresponding to the true abundances would prevent this from happening.

In cases where the detector responses are well separated between different species, the priors do not need to correspond to the true abundances, and could even be flat. The differences between different choices can then be used to provide an estimate of the systematic uncertainties depending on the current knowledge of the detector response. The influence of the choice of priors is further discussed as part of the analyses presented in sect. 4.

The combination of probabilities can also be used to identify and reduce the level of unphysical background that may arise in PID analyses due to track misassociation between detectors. For example, the fraction of TPC tracks that are not correctly associated with the corresponding TOF hit increases with the multiplicity of the event, and

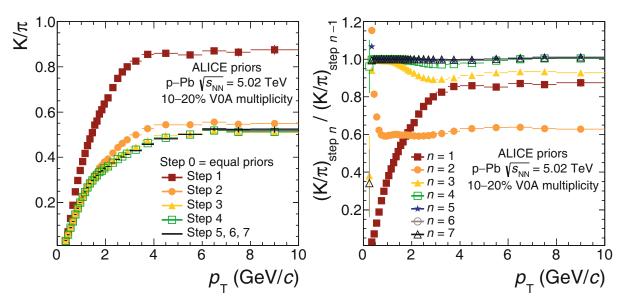


Fig. 2. An example of the iterative prior extraction procedure for p-Pb data (for the 10–20% V0A multiplicity class). The extracted  $K/\pi$  ratio of the priors is shown as a function of  $p_T$  at each step of the iteration (left) and as a ratio of the value between each successive step (right). Step 0 refers to the initial ratio, which is set to 1.

depends on the spatial matching window used in the reconstruction to associate a TOF hit to a track. In ALICE this effect only plays a role in Pb-Pb and p-Pb collisions, and even in the most central Pb-Pb collisions it remains below 10% for tracks with momentum above  $1\,\mathrm{GeV}/c$ . In these cases, a mismatch probability can be defined based on, e.g., the measured flight time being uncorrelated with the track reconstructed with the ITS and TPC due to the TOF hit being produced by another particle.

#### 2.3 Priors in ALICE

As discussed in sect. 2.1, an analysis using the Bayesian approach is expected to be moderately dependent on the choice of priors. Separate sets of priors were evaluated for each collision system. They were evaluated by means of an iterative procedure on data taken in 2010 and 2013 from pp, Pb-Pb, and p-Pb collisions at  $\sqrt{s} = 7 \,\text{TeV}$ ,  $\sqrt{s_{\text{NN}}} = 2.76 \,\text{TeV}$  and  $\sqrt{s_{\text{NN}}} = 5.02 \,\text{TeV}$ , respectively, and were computed as a function of transverse momentum. Priors were also determined as a function of centrality for Pb-Pb collisions, and of the multiplicity class (based on the signal in the V0A detector) for p-Pb data. All of the priors were obtained at mid-rapidity,  $|\eta| < 0.8$ .

The absolute normalisation of the priors is arbitrary, and was chosen so as to normalise all of the priors to the abundance of pions. The value of the priors for pions is thereby set to unity for all  $p_{\rm T}$ . Flat priors (i.e. 1 for all species) are applied at the beginning of the iterative procedure. Bayesian posterior probabilities  $P_n(H_i|S)$  are computed using the priors obtained in step n, as defined in eq. (5). These probabilities are then used in turn as weights to fill identified  $p_{\rm T}$  spectra  $Y(H_i, p_{\rm T})$  for the step n+1 starting from the inclusive (unidentified) measured  $p_{\rm T}$  spectra,

$$Y_{n+1}(H_i, p_T) = \sum_{S} P_n(H_i|S),$$
 (13)

where the summation is performed for all signals S induced by particles of a given  $p_T$  in the sample. It is then possible to obtain a new set of priors  $C_{n+1}$  from the relative ratios of the identified spectra according to

$$C_{n+1}(H_i, p_{\rm T}) = \frac{Y_{n+1}(H_i, p_{\rm T})}{Y_{n+1}(H_{\pi}, p_{\rm T})}.$$
(14)

The procedure is then iterated, and the extracted prior values converge progressively with each iteration. The values of the priors are shown in the left panel of fig. 2 for the  $K/\pi$  ratio obtained using data from p-Pb collisions. The convergence of this procedure is illustrated in the right-hand panel, as a ratio of the priors for successive steps. A satisfactory convergence is obtained after 6–7 iterations.

A set of priors was obtained for global tracks, defined as tracks reconstructed in both the ITS and TPC. This set is referred to in the following as  $C(H_i)_{TPC}$  (the "standard priors"). This set of priors is then propagated to other

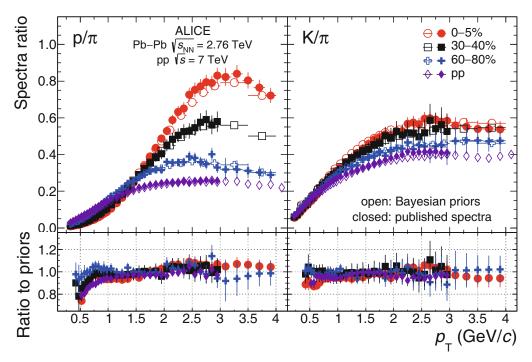


Fig. 3. The proton/pion ratio (left) and kaon/pion ratio (right), as measured by ALICE [8,10] using TPC and TOF (filled symbols), compared with the standard priors as described in the text (open symbols) for Pb-Pb and pp collisions. For Pb-Pb, the results are reported for different centrality classes. Particle ratios are calculated for mid-rapidity, |y| < 0.5. The double ratios (the measured abundances divided by the Bayesian priors) are shown in the lower panels.

detectors using "propagation factors"  $F_{\alpha}$ , which are detector-specific and dependent on transverse momentum. In some cases these multiplicative factors can also be charge-dependent (this is true for EMCal, for example).  $F_{\alpha}$ , which is obtained for each detector via Monte Carlo, takes into account the particles reaching the outer detectors, as well as the acceptances of the outer detectors and their corresponding energy thresholds. The abundances measured by TOF and TPC will differ due to these effects. The requirement of a given detector therefore changes the priors. For an outer detector  $\alpha$ , the priors for a track with momentum  $p_{\rm T}$  are determined as

$$C(H_i)_{\alpha}(p_{\mathrm{T}}) = F_{\alpha}(p_{\mathrm{T}}) \times C(H_i)_{\mathrm{TPC}}(p_{\mathrm{T}}). \tag{15}$$

Priors are currently propagated for TRD, TOF, EMCal and HMPID. Priors were also generated for tracks that are obtained using only ITS hits. These were used for the spectrum analysis at low momentum [9]. In the analyses presented in this paper, the propagation procedure of priors from TPC to TOF was tested extensively.

Since the priors correspond to relative particle abundances, it is possible to directly compare the priors obtained from the iterative procedure with the abundances measured by ALICE [8,10,21]. Comparisons between the priors and the measured p/ $\pi$  and K/ $\pi$  ratios are shown in fig. 3 for pp and Pb-Pb collisions at  $\sqrt{s}=7\,\mathrm{TeV}$  and  $\sqrt{s_\mathrm{NN}}=2.76\,\mathrm{TeV}$ , respectively, and in fig. 4 for p-Pb collisions at  $\sqrt{s_\mathrm{NN}}=5.02\,\mathrm{TeV}$ . In some cases the priors cover a larger range than the measurements made by ALICE.

In order to make these comparisons, the priors were corrected by several factors to take into account the different selections used in the physics analyses and in the priors computation. Since the standard priors are provided for the TPC case, only the tracking efficiency correction was applied. Additional conversions from pseudorapidity intervals (used for the priors) to rapidity intervals (used for the measurement) were also applied, as well as an average feed-down correction derived from the Pb-Pb analysis [10]. This feed-down mainly applies to protons from the decays of  $\Lambda$  baryons.

The priors (open symbols) and the measured abundances (filled symbols) are consistent with one another within roughly 10% over a wide momentum range for all centrality ranges (for Pb-Pb collisions) and V0A multiplicity classes (for p-Pb collisions). The choice to use feed-down corrections evaluated in a given system leads to better agreement at low momenta in Pb-Pb collisions than in p-Pb collisions. However, at high momenta, where the PID performance is better exploited, the results are independent of this correction. The overall level of agreement is satisfactory, as the priors represent a realistic description of the various particle abundances.

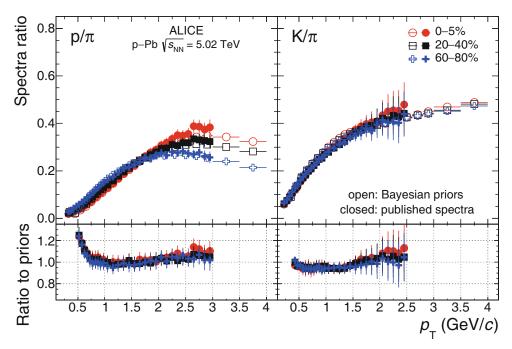


Fig. 4. The proton/pion ratio (left) and kaon/pion ratio (right), as measured by ALICE [21] using TPC and TOF (filled symbols), compared with the standard priors obtained with an iterative procedure (open symbols) for p-Pb collisions for different V0A multiplicity classes. Particle ratios are calculated for mid-rapidity, |y| < 0.5 with respect to the centre-of-mass system. The double ratios (the measured abundances divided by the Bayesian priors) are shown in the lower panels.

It is also important to test whether there is any dependence of the final result on the set of priors used. To perform these tests, it is possible to use alternative sets of priors as well as flat priors, which essentially combine the probabilities from the different detectors without weighting them with Bayesian priors. Slightly different sets of priors can be obtained by varying the track selection parameters. An example of a systematic check on the dependence on varying the set of priors is discussed in sect. 4.2 for the  $D^0$  case.

#### 2.4 Bayesian PID strategies

Once the Bayesian probability for each species (p, K,  $\pi$ , etc.) has been calculated for a given track, the PID selection may be applied with a variety of selection criteria. The three criteria applied in this paper are:

- Fixed threshold: The track is accepted as belonging to a species if the probability for this is greater than some pre-defined value. As an example, the choice of a 50% threshold means that a particle will only be accepted as a kaon if its Bayesian probability of being a kaon is greater than 50%. Note that this strategy is not necessarily exclusive, as a threshold of less than 50% could lead to multiple possible identities. As already discussed, a selection on the Bayesian probability corresponds to a purity requirement of the signal if the priors reflect the true particle abundances.
- Maximum probability: The track is accepted as the most likely species (i.e. the species with the highest probability).
- Weighted: All tracks reaching the PID step of the analysis are accepted, with a weight different from unity applied to their yield. The weight is defined as the product of the Bayesian probabilities obtained for the tracks involved (e.g. in D<sup>0</sup> → K<sup>-</sup> $\pi$ <sup>+</sup>, this is the kaon probability of the negative track multiplied by the pion probability of the positive track). The final result is corrected for the average weight determined in Monte Carlo simulations in the same way as is done for the PID efficiency in other methods.

The fixed threshold method is compared with  $n_{\sigma}$  PID as part of the benchmark analysis (see sect. 3). The maximum probability method is used in the single-particle spectrum analysis described in sect. 4.1 and for the  $\Lambda_c^+$  baryon analysis in sect. 5. Finally, all three of the aforementioned methods are tested for the analysis of  $D^0 \to K^-\pi^+$  in sect. 4.2.

## 3 Benchmark analysis on two-prong decay channels

High-purity samples of identified particles were selected via the study of specific decay channels. These samples served as a baseline for validating the Monte Carlo tools that are normally used to estimate the efficiencies and misidentifica-

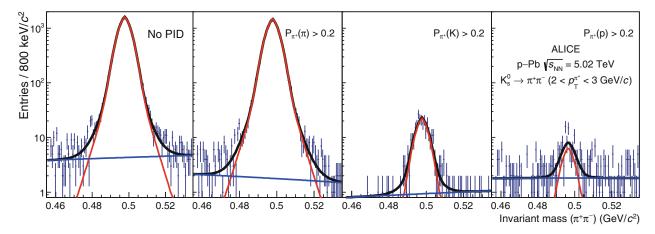


Fig. 5. V<sup>0</sup> fits to extract the yield and background for  $K_S^0 \to \pi^- \pi^+$  in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02 \, \text{TeV}$ . From left to right: no PID selections applied and selecting pions, kaons and protons using a specific PID strategy (here, Bayesian probability > 0.2). The yield estimated from the second plot from the left (compared with the no-PID yield result) gives a measure of the PID efficiency for the pions, while the remaining ones give information about misidentification.

tion probabilities (i.e. the  $\epsilon_{\text{PID}}$  matrix discussed in sect. 2) of the Bayesian PID approach. The methodology developed in this section was also applied to the  $n_{\sigma}$  PID approach, providing an important cross-check.

#### 3.1 Description of the method

The following decays were used to obtain high-purity samples of three different species:

- $K_S^0 \to \pi^-\pi^+$  to study charged pions;  $\Lambda \to p\pi^-$  (and respective charge conjugates) to study protons; and
- $-\phi \to K^-K^+$  to study charged kaons.

The first two cases are true  $V^0$  decays, comprising two charged prongs originating from a secondary vertex displaced from the interaction point, whereas the daughters of the  $\phi$  meson originate from the primary interaction vertex due to its shorter lifetime. For conciseness, all of these are referred to as  $V^0$  decays in this paper.

Here we report results based on the p-Pb data set, testing the PID method using both TPC and TOF. The p-Pb data sample is less affected by background uncertainties than the Pb-Pb sample when performing fits of the invariant mass spectra, due to the smaller amount of combinatorial background. This is especially true for the  $\phi$  analysis.

Different track selection criteria were applied in order to select daughter particles coming either from the primary vertex (for the  $\phi$  case) or from a secondary vertex (for the  $K_S^0$  and  $\Lambda$  cases). For a given  $V^0$ , a fit of the combinatorial invariant mass distribution allows the background to be subtracted and the yield of V<sup>0</sup> decays to be extracted. The estimated yield is considered to be a pure sample of a given species (a precise measurement of the total number of particles of a given species in a given data set). This estimation was done without applying any PID selections. Then the exercise was repeated applying PID selections on each of the two prongs, selecting between pions, kaons and protons. The comparison with the number of positively identified secondary prongs determines the efficiency and the misidentification with respect to the values estimated when not applying PID.

Figure 5 shows examples of the fitting procedure for the  $K_S^0$  invariant masses for  $2 < p_T^{\pi} < 3 \,\mathrm{GeV}/c$ . From left to right, the panels show the analysis without PID and then requiring the identification of a positive pion, kaon, or proton, respectively. The  $K_S^0$  signal (and background) fitted in the latter three cases, compared with the results without applying PID, allow the extraction of the identification efficiency and misidentification probabilities.

In order to reduce the background for the  $\phi$  analysis, before starting the procedure, one of the two decay tracks was "tagged" using a PID selection requiring compatibility with the kaon hypothesis and then the PID selection under study is applied on the other track. The tagging was performed with a  $2\sigma$  selection on a combination of the TPC and

The TPC and TOF signals are combined as  $|n_{\sigma}^{\text{Comb}}(j)| = \sqrt{(n_{\sigma}^{\text{TPC}}(j)^2 + n_{\sigma}^{\text{TOF}}(j)^2)/2}$ . The same method was also applied on the simulated sample in order to check the agreement between Monte Carlo and data on the estimated quantities. This serves as a validation for the Monte Carlo estimation of the efficiency of a given PID strategy, as well as for the subsequent corrections required in order to extract physics results.

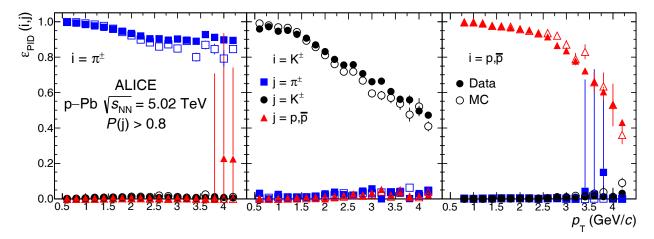


Fig. 6.  $\epsilon_{\text{PID}}$  matrix elements in p-Pb collisions after selection with a Bayesian probability greater than 80%. Comparisons with Monte Carlo (open symbols) are also shown.

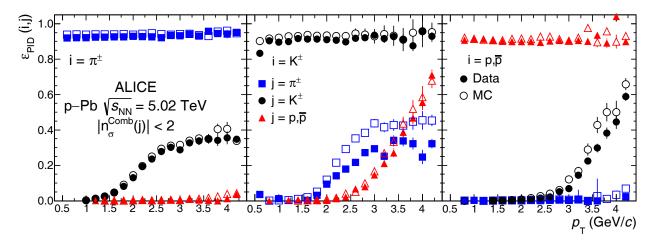


Fig. 7.  $\epsilon_{\text{PID}}$  matrix elements in p-Pb collisions after selection with a  $2\sigma$  selection on the combined TPC and TOF signal. Comparisons with Monte Carlo (open symbols) are also shown.

#### 3.2 Comparison of PID efficiencies between data and Monte Carlo

The PID matrix elements obtained for different Bayesian probability thresholds are presented here. An example for the highest purity case considered in this work (Bayesian probability greater than 80%) is shown in fig. 6. Each plot represents a row of the  $\epsilon_{\text{PID}}$  matrix for a given species i, with the i=j points corresponding to the PID efficiencies ( $\epsilon_{ii}$ ) and the  $i \neq j$  points corresponding to the misidentification probabilities ( $\epsilon_{ij}$ ). The matrix was evaluated separately for positively and negatively charged tracks. As no difference was found between the two cases, the results shown here were averaged over both charges. For Monte Carlo, the PID hypothesis was tested both via the true particle identity available in the simulation and by applying the same procedure as for the data.

As can be seen from fig. 6, the efficiencies and misidentification probabilities can be evaluated very precisely. The agreement between data and Monte Carlo is good, both in shape and absolute value. The general features of the PID strategies are also described well, and behave as expected: using a high threshold maximises the purity, but also sharply reduces the efficiency. Nevertheless, even for  $P \geq 80\%$ , where the efficiency estimate is more sensitive to the description of the detector responses in the simulations, the agreement with Monte Carlo remains within 5% below  $3 \, \text{GeV}/c$ .

The analysis was repeated using  $2\sigma$  and  $3\sigma$  selections on the combined TOF and TPC signals, as discussed above, to identify the three hadron species. The result is shown in fig. 7 for the  $2\sigma$  case. As expected, the  $\epsilon_{\text{PID}}$  matrix elements have different values from the 80% Bayesian probability threshold (given that the  $n_{\sigma}^{\text{Comb}}$  selection is somewhat more inclusive), and the probability of misidentification increases accordingly. The agreement between Monte Carlo and

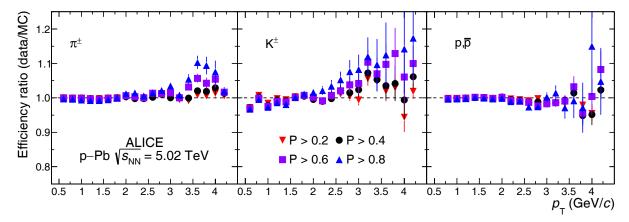


Fig. 8. Data/Monte Carlo ratios of PID efficiencies for pions, kaons and protons in p-Pb collisions, extracted using different Bayesian probability thresholds.

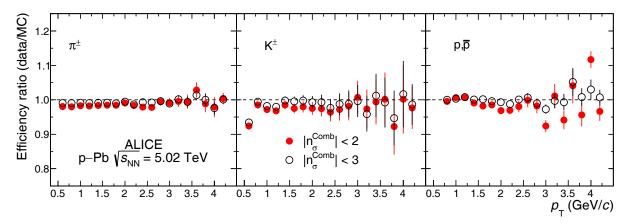


Fig. 9. Data/Monte Carlo ratios of PID efficiencies for pions, kaons and protons in p-Pb collisions, extracted using 2- and  $3\sigma$  selections on the combined TPC and TOF signal.

data for the misidentification probabilities is worse in the  $n_{\sigma}$  case than in the Bayesian case for kaons misidentified as pions. However, there remains a good agreement between Monte Carlo and data overall. The efficiencies are below the values expected from a perfectly Gaussian signal; a Monte Carlo or data-driven evaluation of the PID strategy efficiency is therefore mandatory. The non-Gaussian tail of the TOF signal [18], caused by charge induction on pairs of neighbouring readout pads, plays a significant role in this discrepancy. The mismatch fraction is also not negligible in p-Pb collisions, being  $\sim 2\%$  above  $1\,\mathrm{GeV}/c$ .

For a more detailed comparison, the ratios between data and Monte Carlo are presented in figs. 8 and 9 for efficiencies obtained using different Bayesian probability thresholds, and a  $2\sigma$  or  $3\sigma$  PID selection, respectively. The agreement is very similar (within  $\sim 5\%$ ) for both methods. A larger difference is seen for kaons when using a very high Bayesian probability threshold (corresponding to a "high purity" strategy). However, such an approach could still be beneficial in physics analyses where an efficiency correction is not needed (such as analyses investigating Bose-Einstein correlations [15] or the flow of identified particles [22]). The uncertainties reported in these plots are purely statistical. Apart from the case of kaons extracted from  $\phi$  decays, the statistical uncertainties on the Monte Carlo simulations give the largest contribution to the uncertainties shown here. The  $\phi$ -meson invariant mass plots are affected by a larger combinatorial background in the data sample when PID is not requested (i.e. in the denominator of the efficiency). As the ratios of data to Monte Carlo for both Bayesian and  $n_{\sigma}$  PID are close to unity in figs. 8 and 9, it can be concluded that the systematic uncertainties from the PID procedure are well under control in both cases.

In summary, the  $V^0$  analysis technique described in this section serves not only as a validation of the quality of the Monte Carlo description of the various detector responses, but can also be used in data and simulations to validate different PID strategies and track selections for any kind of analysis. Finally, it provides a tool to investigate the systematic uncertainties that arise due to the PID selection.

Hadron	TPC	TPC-TOF
$\pi$	$0.2 \le p_{\mathrm{T}} \le 0.5$	$0.5 \le p_{\mathrm{T}} \le 2.5$
K	$0.3 \le p_{\rm T} \le 0.45$	$0.45 \le p_{ m T} \le 2.5$
p	$0.5 \le p_{\mathrm{T}} \le 0.8$	$0.8 \le p_{\mathrm{T}} \le 2.5$

Table 1. PID detectors and transverse momentum ranges (in GeV/c) used for the analysis of identified hadron spectra.

## 4 Bayesian approach applied to physics analyses

In this section we present validations of the Bayesian PID approach for two analyses already published by the ALICE Collaboration in pp collisions at  $\sqrt{s} = 7$  TeV: identified pion, kaon and proton spectra [8] and  $D^0 \to K^-\pi^+$  [13]. While the previous papers remain the proper references for the extraction of the physical quantities such as cross sections, and for theory comparisons, this paper shows the results obtained by applying the Bayesian approach to the PID part of those analyses.

#### 4.1 Identified hadron spectra

The consistency of the Bayesian PID technique was tested using the tools described in the previous sections to obtain the  $p_{\rm T}$  spectra of pions, kaons and protons. This analysis used a data sample of  $1.2 \times 10^8$  inelastic pp collisions at  $\sqrt{s} = 7 \, {\rm TeV}$  that was collected in 2010.

The results are compared here with similar measurements that were already reported in [8]. In the quoted paper, different PID techniques were used depending on the detectors involved in different  $p_{\rm T}$  ranges ( $n_{\sigma}$  for ITS only and for the combined TPC and TOF signals, and unfolding techniques for TOF and HMPID separately). Charged kaon spectra were also measured via the identification of their decays (measurement of the kink topology). Full details of the original analysis, including the event selection criteria (which were also used here) and the procedure used to merge the various PID techniques and detectors, can be found in [8].

The efficiency, detector acceptance, and other correction factors, were estimated using Monte Carlo samples that simulated the detector conditions run by run. The simulation was based on the PYTHIA6.4 event generator [23] using the Perugia0 tune [24], with events propagated through the detector using GEANT3 [25].

Only the TPC and TOF detectors were used for PID. The particle species were identified using the maximum probability method outlined in sect. 2.4. The identification of charged hadrons used different detector combinations in different momentum ranges, as shown in table 1. Although a Bayesian approach does not necessarily require such a division, it was chosen in order to make a closer comparison with the analysis presented in [8]. In particular, the TOF efficiency drops very steeply at low momentum due to the acceptance, meaning that the systematic uncertainty on the TOF efficiency would become dominant and make the comparison difficult.

The PID efficiency is higher than 95% for pions and protons, while for kaons it begins to decrease from 100% at  $1\,\mathrm{GeV}/c$  to 75% at  $2.5\,\mathrm{GeV}/c$ . The misidentification percentage is below 5% for pions and protons for all  $p_\mathrm{T}$ , and reaches 20% for kaons at  $p_\mathrm{T}=2.5\,\mathrm{GeV}/c$ . In order to avoid the dependence of the corrections on the relative abundances of each hadron species in the event generator, the spectra were corrected for their respective PID efficiencies and for contamination using the  $\epsilon_\mathrm{PID}$  matrix method described in sect. 2.2. A  $4\times4$  matrix (also including electrons) was defined for each  $p_\mathrm{T}$  interval. These matrices were then inverted and used in eq. (10) in order to obtain the spectra. In addition, the spectra were corrected for the tracking, TPC-TOF matching and primary vertex determination efficiencies. The contributions from secondary particles that were not removed by a selection based on the distance of closest approach to the vertex were determined using a data-driven method, as explained in [8], and were subtracted from the final spectra.

Figure 10 compares the minimum-bias charged hadron spectra from pp collisions at  $\sqrt{s} = 7 \,\mathrm{TeV}$  obtained from this analysis with the published result. A very good agreement within the uncertainties can be observed between the  $p_{\mathrm{T}}$  spectra obtained using these different approaches. For this comparison, the statistical uncertainties are shown for both analyses, while the systematic uncertainties that are not only related to PID were only considered for the results published in [8]. The ratios of the spectra, presented in the lower panels of fig. 10, show an agreement within  $\pm 5\%$  for all species.

A further check was performed testing the stability of the method against the priors used. The analysis was repeated using flat priors, *i.e.* equal probabilities for the four particle species that are included in the PID matrix (electrons, pions, kaons and protons). The muon priors were set to zero, their percentage being negligible (with standard priors, they are estimated to be less than 2% with respect to all other particle species). Despite considering quite an extreme case here in terms of varying the priors, the pion and proton abundances were consistent with the result obtained

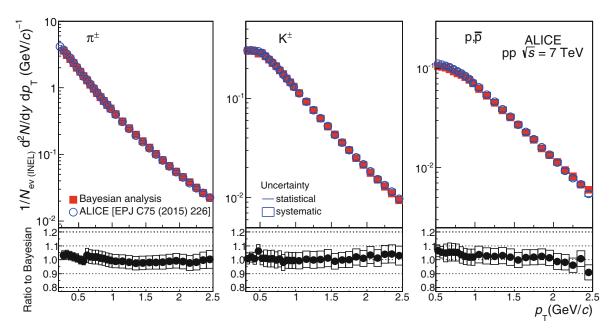


Fig. 10. Identified particle spectra from the Bayesian analysis, compared with the measurement reported by ALICE in pp collisions at 7 TeV [8].

with standard priors within 3%. A decrease in the estimated pion abundance at  $p_T > 1 \text{ GeV}/c$  resulted in a variation of up to 10% for kaons in some  $p_T$  intervals. The observed variations can be interpreted as being due to remaining uncertainties in the detector response; a naive set of priors, such as a set of flat priors, is generally expected to amplify such effects.

# 4.2 Analysis of $D^0 o K^-\pi^+$

This section presents a comprehensive overview of a variety of Bayesian PID strategies applied to the analysis of  $D^0 \to K^-\pi^+$  (and charge conjugates) in pp collisions at  $\sqrt{s}=7\,\mathrm{TeV}$ . The analysis was based on a data sample of roughly  $3\times 10^8$  events collected during Run 1. The geometrical selections on the displaced decay vertex topology matched those used in the ALICE measurement of  $D^0 \to K^-\pi^+$  in pp collisions reported in [13].

In order to make a detailed assessment of the Bayesian method, this analysis compares the results obtained using each of the PID strategies outlined in sect. 2.4. In each case, the PID method relied on selecting an oppositely charged pair of tracks corresponding to a kaon and a pion. The TPC and TOF detectors were used in conjunction with one another. For tracks without TOF information, the PID was based on information from the TPC only.

For the fixed-threshold method, probability thresholds of 40%, 50%, 70%, and 80% were tested. Note that in the case of the 40% threshold, there is the possibility that a daughter track may be compatible with both the kaon and pion hypothesis; in such cases, the track was accepted as possibly belonging to either species. For the maximum-probability and fixed-threshold methods, once the daughter tracks were analysed, the candidate was accepted or rejected according to the following criteria:

- if both daughters were identified as possible kaons, the candidate was accepted both as a  $\overline{D}^0$ ;
- if one daughter was identified as a kaon and the other as a pion, the candidate was accepted as a  $D^0$  if the negative track was a kaon and the positive track was a pion, and vice versa for  $\overline{D}^0$ ;
- if neither daughter was identified as a kaon, the candidate was rejected;
- if either daughter was not compatible with the kaon or the pion hypothesis, the candidate was rejected.

The weighted method was implemented as defined in sect. 2.4, whereby the invariant mass distributions were filled for each candidate with weights  $W_i$ . These weights are defined as

$$W_{\rm D^0} = P_{\rm K^-} \times P_{\pi^+},\tag{16}$$

and

$$W_{\overline{\mathbb{D}}^0} = P_{\mathcal{K}^+} \times P_{\pi^-},\tag{17}$$

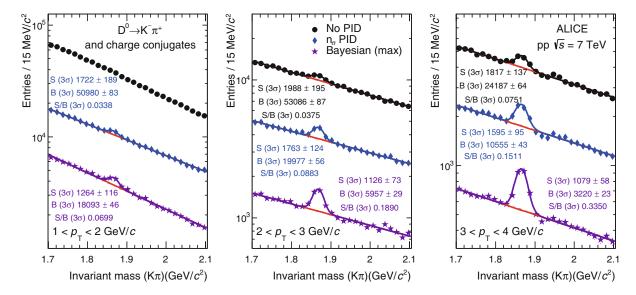


Fig. 11. A comparison of the invariant mass distributions in three  $p_{\rm T}$  intervals for D<sup>0</sup> candidates obtained without PID, with  $n_{\sigma}$  PID, and with Bayesian PID using the maximum probability condition. Due to the low statistical significance, it was not possible to extract a stable signal without PID for  $1 < p_{\rm T} < 2 \, {\rm GeV}/c$ , therefore this fit and its results are not shown.

where  $P_i$  corresponds to the Bayesian probability assigned to each track for a given species i. In this case, the PID efficiency from Monte Carlo corresponded to the average weight that was assigned to simulated  $D^0$  mesons in each  $p_T$  interval.

Once the invariant mass distributions in each  $p_{\rm T}$  interval had been obtained, the yields were fitted using a Gaussian function for the signal and an exponential function to estimate the background. The raw signal was extracted according to the integral of the fit to the signal distribution. In order to ensure the stability of the fitting procedure, the width and mean of the Gaussian functions were kept the same in each  $p_{\rm T}$  interval for all PID methods. To do this, the fit parameters were first determined for the  $n_{\sigma}$  PID method in each  $p_{\rm T}$  interval, and then fixed to those values in the fitting procedure for each other PID method.

Example invariant mass distributions obtained without PID, with  $n_{\sigma}$  PID, and with Bayesian PID using the maximum-probability strategy are shown in fig. 11, with a logarithmic scale on the vertical axis. There is a clearly visible increase in the signal-to-background ratio when using  $n_{\sigma}$  PID as compared to the analysis without PID, and a further increase when applying a Bayesian selection with the maximum-probability strategy. Due to the large background and low statistical significance, it was not possible to extract a stable yield for  $1 < p_T < 2 \,\text{GeV}/c$  without PID. For this reason, the results obtained without PID are excluded for this  $p_T$  interval in the following comparisons.

Full comparisons of the signal-to-background ratio and significance obtained for different PID strategies are shown in fig. 12. The statistical significance is defined as the signal divided by the square root of the sum of the signal and background. The signal-to-background ratio increases significantly for all  $p_{\rm T}$  when using the  $n_{\sigma}$  PID approach, as compared to not making a PID selection. A similar increase, of a factor of at least 2, was seen in each of the Bayesian methods when compared with  $n_{\sigma}$  PID.

As expected, the statistical significance of the signal was higher for the  $n_{\sigma}$  PID than when no PID selection was applied. At low  $p_{\rm T}$ , all but the strictest Bayesian selections (i.e. 80% probability threshold) gave a further increase in significance over the  $n_{\sigma}$  PID method, while at higher  $p_{\rm T}$  the significance of these methods became similar to that of  $n_{\sigma}$  PID. The 80% probability threshold yielded a lower statistical significance than  $n_{\sigma}$  PID for  $p_{\rm T} > 3\,{\rm GeV}/c$ , whereas the performance of the two methods is equivalent for  $p_{\rm T} \le 3\,{\rm GeV}/c$ .

The efficiencies for each particle identification method were obtained using charm-enriched Monte Carlo simulations, as described in [13]. The PID efficiency was defined as the proportion of simulated D mesons that passed the PID selection criteria, having already passed the other selections. In the case of the weighted Bayesian PID method, as previously mentioned, the PID efficiency corresponded to an average weight for true  $D^0$  candidates, and for the purposes of corrections was used in the same way as the other PID efficiencies. The PID efficiencies obtained for each method are shown in fig. 13.

When a stricter PID selection was applied (for example, an 80% probability threshold as compared to a 70% threshold), the PID efficiency fell correspondingly. All of the Bayesian methods gave a lower PID efficiency than the  $n_{\sigma}$  method. The observed significances and efficiencies highlight a benefit and a potential drawback in using a strict Bayesian PID strategy: while a tighter PID selection gives a yield of higher purity, it also requires a greater reliance on

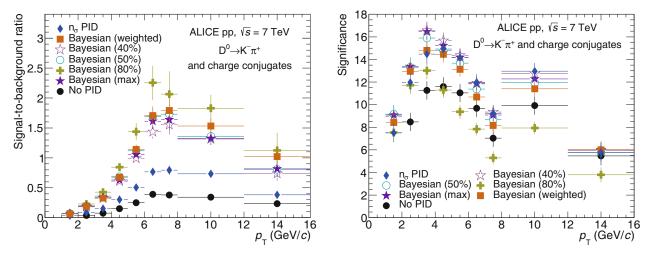


Fig. 12. (Left) Signal-to-background ratio and (right) statistical significance as a function of  $p_{\rm T}$  for various methods of particle identification. Note that the increase in significance at  $8 < p_{\rm T} < 12\,{\rm GeV}/c$  is an effect of the width of the  $p_{\rm T}$  interval increasing from 1 to  $4\,{\rm GeV}/c$ .

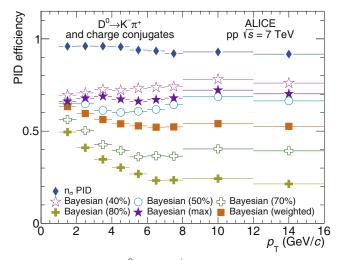


Fig. 13. A comparison of the PID efficiencies for  $D^0 \to K^-\pi^+$  obtained using various PID strategies, as a function of  $p_T$ .

the Monte Carlo simulation to give an accurate correction due to the reduced PID efficiency. However, it is important to note that this is not necessarily inherent to the Bayesian PID method. For example, a very narrow  $n_{\sigma}$  selection (e.g. at  $0.5\sigma$ ) would also yield a lower efficiency, leading to a similar degree of reliance on the Monte Carlo simulations; even a  $2\sigma$  selection reduces the PID efficiency to 80%.

The corrected yields obtained using each particle identification method are compared in fig. 14. In each case, the ratio shown is the corrected yield divided by that obtained using  $n_{\sigma}$  PID. As the candidates selected by each of the Bayesian methods are a subset of those selected by the  $n_{\sigma}$  approach, the uncertainties on the corrected yields are assumed to be fully correlated and so almost cancel out in the ratio.

It was found that there is a relatively good agreement between the Bayesian PID method and the  $n_{\sigma}$  PID method, even when very different probability thresholds are applied. Averaging over  $p_{\rm T}$ , the fixed probability threshold methods give a systematic difference between 5 and 9%, increasing for the lower probability thresholds. For the weighted and maximum-probability methods, the average deviations from  $n_{\sigma}$  are 3% and 7%, respectively. In every case this difference is of a similar order to the systematic uncertainty of the  $n_{\sigma}$  PID approach ( $\sim 5\%$ ) [13].

In effect, the difference of the corrected yield from the  $n_{\sigma}$  result represents the systematic uncertainty due to the choice of Bayesian PID method. Therefore it can be concluded that the PID systematic uncertainties for the Bayesian approach are not much larger than those from the  $n_{\sigma}$  approach. The observed deviations also contain systematic contributions related to the yield extraction procedure.

A final check was made to ensure that the results remained independent of the choice of the prior probability distribution used. Ideally, the requirement of an external input should guide the measurement, and fluctuations in

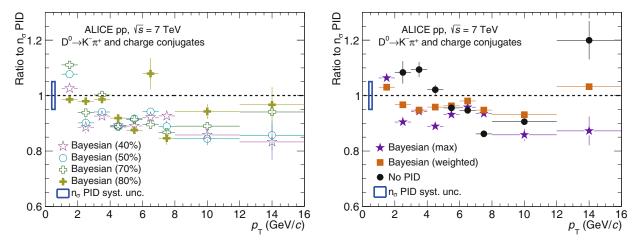


Fig. 14. Ratios of corrected yields obtained using various Bayesian PID methods to that obtained using  $n_{\sigma}$  PID, for (left) fixed probability thresholds, and (right) maximum probability and weighted Bayesian PID, and no PID. The 5% systematic uncertainty on the  $n_{\sigma}$  PID method is shown as a blue box at  $0.5 \,\mathrm{GeV}/c$ .

these priors should not have a substantial effect on the final result that is obtained. In order to test this, two modified sets of priors were prepared, with the proportion of kaons scaled up and down by 10% for all  $p_{\rm T}$ . An analysis using these modified sets of priors was then performed and compared to the analysis with the standard priors. The corrected yields using the modified priors were divided by the yields using the standard set of priors for each Bayesian analysis method. It was found that the average deviations from unity remained within roughly 2% for each set of alternative priors with all of the Bayesian PID strategies. The smallest deviations were observed for the maximum probability and weighted methods, and for lower fixed probability thresholds.

In conclusion, since the corrected yield is largely independent of the choice of PID method, and considering the improvement in signal-to-background ratio and (at low  $p_{\rm T}$ ) significance, we conclude that the Bayesian PID method is a valid choice for the analysis of  ${\rm D}^0 \to {\rm K}^-\pi^+$ . However, it must be noted that although the stricter Bayesian methods give a purer signal, they also create an increased reliance on the Monte Carlo simulations to correct for the lower efficiencies, as well as introducing a greater dependence on the choice of priors in those  $p_{\rm T}$  intervals that have lower statistical significance.

# 5 Application of Bayesian PID for $\Lambda_c^+$ analysis

In this section we present an exploratory study of  $\Lambda_c^+ \to p K^- \pi^+$  in pp collisions at  $\sqrt{s} = 7 \,\text{TeV}$  using a Bayesian PID approach. This decay channel suffers from a high level of combinatorial background, even in the relatively low-multiplicity environment of a pp collision. Due to the short decay length of the  $\Lambda_c^+$  ( $c\tau = 59.9 \,\mu\text{m}$  [26]) and current limitations in the spatial resolution of the ITS, topological and geometrical selections alone are not sufficient to reduce this background and extract a signal. This means that a robust PID strategy is required to select  $\Lambda_c^+$  candidates.

Here we compare analyses of  $\Lambda_c^+ \to p K^- \pi^+$  using the  $n_\sigma$  approach, a Bayesian approach using the maximum-probability criterion, and an alternative "minimum- $\sigma$ " approach that mimics the Bayesian strategy. This analysis used  $\sim 3 \times 10^8$  events from pp collisions at  $\sqrt{s} = 7 \, \text{TeV}$  collected during Run 1.

Both the TPC and TOF responses were used for the  $n_{\sigma}$  approach. A  $3\sigma$  selection was applied on the TOF signal for kaons, pions and protons at all  $p_{\rm T}$ , while in the TPC a  $p_{\rm T}$ -dependent  $n_{\sigma}$  selection (between 2 and  $3\sigma$ ) was applied for protons and kaons, and a  $3\sigma$  selection was applied for pions at all  $p_{\rm T}$ . If the track had a valid TOF signal, only this was used for the PID; otherwise the TPC response was used. For the Bayesian PID method, the TPC and TOF responses were combined; in cases where the TOF signal was not available, only the TPC response was used.

The minimum- $\sigma$  strategy serves as a middle ground between the  $n_{\sigma}$  analysis and the Bayesian approach. It is also based on the deviation from the expected detector signal, but in this case only the species resulting in the smallest  $n_{\sigma}$  value is chosen for each track, making this an exclusive selection. The TPC and TOF responses are combined by adding their respective  $n_{\sigma}$  values in quadrature. The potential benefit of this approach is that it uses similar logic to the Bayesian approach, without requiring the input of any priors.

In order to build  $\Lambda_c^+$  candidates, triplets of tracks were reconstructed and selected, based on PID and topological selections. Pairs of oppositely charged tracks were reconstructed from the set of selected tracks, and a secondary (decay) vertex was computed. Selections were applied to each pair based on the distance of closest approach between the two tracks, and the distance between the primary and the secondary vertex of the reconstructed pair. A third track was then associated to each selected pair to form a triplet. The secondary vertex of the triplet was calculated and the  $\Lambda_c^+$  candidates were selected using:

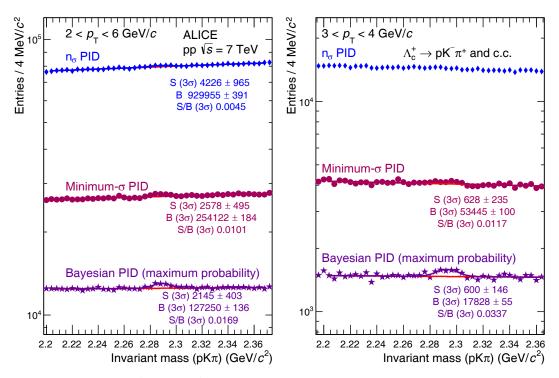


Fig. 15. Invariant mass spectra of  $\Lambda_c^+ \to p K^- \pi^+$  using  $n_\sigma$  PID, minimum- $\sigma$  PID and Bayesian PID for (left)  $2 < p_T < 6 \,\text{GeV}/c$  and (right)  $3 < p_T < 4 \,\text{GeV}/c$ . Due to the low statistical significance, it was not possible to extract a stable signal for  $n_\sigma$  PID for  $3 < p_T < 4 \,\text{GeV}/c$ , therefore this fit and its results are not shown.

- the quality of the reconstructed vertex, which was determined using the quadratic sum of the distances of the single tracks from the calculated secondary vertex;
- the distance between the primary and the secondary vertex of the triplet (decay length); and
- the cosine of the pointing angle. The pointing angle is the angle between the momentum vector of the reconstructed  $\Lambda_c^+$  candidate and the  $\Lambda_c^+$  flight line reconstructed from the line joining the primary and secondary vertices.

The invariant mass plots for pK $\pi$  candidates with  $2 < p_{\rm T} < 6\,{\rm GeV}/c$  are shown in the left-hand panel of fig. 15 for the standard  $n_{\sigma}$  PID approach (top), the alternative minimum- $\sigma$  approach (middle) and Bayesian PID (bottom). Over this  $p_{\rm T}$  interval, the statistical significance determined from the yield extraction (as defined in sect. 4.2) is  $4.4\pm1.0$  for  $n_{\sigma}$  PID,  $5.1\pm1.0$  for minimum- $\sigma$  PID and  $6.0\pm1.1$  for Bayesian PID. In addition, the signal-to-background ratio is three times higher in the Bayesian case than for  $n_{\sigma}$  PID, with a background that is reduced by a factor of approximately seven. On the other hand, the Bayesian PID method yields fewer  $\Lambda_{\rm c}^+$  candidates in the peak, indicating a lower selection efficiency.

The equivalent invariant mass distributions for  $3 < p_T < 4 \,\mathrm{GeV}/c$  are shown in the right-hand panel of fig. 15. In this case, the  $n_\sigma$  PID approach yields an insufficient statistical significance to extract a stable yield. As with the  $p_T$ -integrated case, the statistical significance improves when using the other two methods  $(2.7 \pm 1.0 \,\mathrm{for}$  minimum- $\sigma$  and  $4.4 \pm 1.1$  for Bayesian PID), and we again find a marked improvement in the signal-to-background ratio for these two approaches. From this we conclude that the Bayesian approach represents the best candidate PID method for future measurements of the  $\Lambda_c^+$  production cross section.

# 6 Conclusions and outlook

A Bayesian method for particle identification has been presented and validated for a variety of analyses. A comparison between different PID selection methods in ALICE was performed, with a focus on testing the suitability of Bayesian PID techniques. A selection based on the Bayesian probability can be interpreted as a request on the purity, and therefore the main benefit of these techniques is that they increase the purity of the extracted signal by combining information from different detectors via a relatively simple technique.

The iterative procedure used to extract the prior probabilities (corresponding to the relative particle abundances) was outlined for pp collisions at  $\sqrt{s} = 7 \,\text{TeV}$ , p-Pb collisions at  $\sqrt{s_{\text{NN}}} = 5.02 \,\text{TeV}$  and Pb-Pb collisions at  $\sqrt{s_{\text{NN}}} = 2.76 \,\text{TeV}$ . In each collision system, and for every centrality and multiplicity class studied in Pb-Pb and p-Pb collisions, the extracted priors were found to be consistent with the true particle abundances seen in data.

The ability of Monte Carlo simulations to compute efficiencies and misidentification probabilities was tested via high-purity samples of pions, kaons and protons from the two-prong decays  $K_S^0 \to \pi^-\pi^+$ ,  $\Lambda \to p\pi^-$ , and  $\phi \to K^-K^+$  in p-Pb collisions at  $\sqrt{s_{\rm NN}} = 5.02\,{\rm TeV}$ . Equivalent analyses were also performed using  $n_\sigma$  PID. It was found that the variation in the result when using Bayesian PID as opposed to  $n_\sigma$  PID was on the order of 5% in each case.

Comparisons between the Bayesian approach and PID methods used in measurements already published by ALICE in pp collisions at  $\sqrt{s}=7\,\mathrm{TeV}$  were performed for the analysis of the  $p_\mathrm{T}$ -differential yields of identified pions, kaons and protons. In this case, the previous analysis used a combination of  $n_\sigma$  PID, unfolding methods and spectra obtained with kinks. A simple maximum-probability Bayesian PID method was used in this analysis. When compared, the results from  $n_\sigma$  PID and Bayesian PID were found to be consistent with one another within uncertainties.

A detailed comparison of different Bayesian PID strategies was performed for the analysis of  $D^0 \to K^-\pi^+$  in pp collisions at  $\sqrt{s} = 7 \,\mathrm{TeV}$ . Through the use of different probability thresholds, the trade-off between selecting a purer sample and having a stronger dependence on the description of the detectors in Monte Carlo simulations was demonstrated. An increase in the signal-to-background ratio over the  $n_\sigma$  PID method was seen for all  $p_T$  for all of the tested Bayesian strategies. The statistical significance was found to be similar or greater than  $n_\sigma$  PID for all of the Bayesian methods other than a strict 80% probability threshold. In every case, the corrected yield was found to be stable against the choice of PID method when comparing with the  $n_\sigma$  PID method used for  $D^0$  mesons. The dependence on the choice of priors was also investigated, and it was found that the uncertainty on the corrected yield due to the choice of priors was within  $\sim 2\%$ . We conclude that the systematic uncertainties arising from the choice of Bayesian PID method are of a similar order to those of the  $n_\sigma$  approach.

In summary, a good level of consistency is seen in both of the full analyses considered in sects. 4.1 and 4.2 with respect to previously reported results. Because of this, we conclude that the validity of the Bayesian PID approach has been successfully assessed.

Furthermore, the improved performance of the Bayesian PID approach in certain scenarios will prove beneficial for future analyses. The case of the  $\Lambda_c^+$  baryon was presented as an example where a simple  $n_{\sigma}$  approach is unable to yield a stable signal, while the Bayesian approach is able to combine information from the different detectors effectively without the need to develop a complex variable- $\sigma$  selection. This presents a promising option for a more comprehensive study of the  $\Lambda_c^+$  production cross section in pp and p-Pb collisions.

The analyses presented here are currently limited to charged hadrons detected in the TPC and TOF detectors. However, the method can be extended to other particle species and detectors within the ALICE central barrel, and we expect to further test and refine these techniques in future. In particular, it will be possible to examine the performance of the Bayesian approach for electron identification with the TRD at full azimuthal coverage in Run 2.

The ALICE Collaboration would like to thank all its engineers and technicians for their invaluable contributions to the construction of the experiment and the CERN accelerator teams for the outstanding performance of the LHC complex. The ALICE Collaboration gratefully acknowledges the resources and support provided by all Grid centres and the Worldwide LHC Computing Grid (WLCG) collaboration. The ALICE Collaboration acknowledges the following funding agencies for their support in building and running the ALICE detector: State Committee of Science, World Federation of Scientists (WFS) and Swiss Fonds Kidagan, Armenia; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Financiadora de Estudos e Projetos (FINEP), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP); National Natural Science Foundation of China (NSFC), the Chinese Ministry of Education (CMOE) and the Ministry of Science and Technology of China (MSTC); Ministry of Education and Youth of the Czech Republic; Danish Natural Science Research Council, the Carlsberg Foundation and the Danish National Research Foundation; The European Research Council under the European Community's Seventh Framework Programme; Helsinki Institute of Physics and the Academy of Finland; French CNRS-IN2P3, the "Region Pays de Loire", "Region Alsace", "Region Auvergne" and CEA, France; German Bundesministerium fur Bildung, Wissenschaft, Forschung und Technologie (BMBF) and the Helmholtz Association; General Secretariat for Research and Technology, Ministry of Development, Greece; National Research, Development and Innovation Office (NKFIH), Hungary; Department of Atomic Energy and Department of Science and Technology of the Government of India; Istituto Nazionale di Fisica Nucleare (INFN) and Centro Fermi - Museo Storico della Fisica e Centro Studi e Ricerche "Enrico Fermi", Italy; Japan Society for the Promotion of Science (JSPS) KAKENHI and MEXT, Japan; Joint Institute for Nuclear Research, Dubna; National Research Foundation of Korea (NRF); Consejo Nacional de Cienca y Tecnologia (CONACYT), Direccion General de Asuntos del Personal Academico(DGAPA), México, Amerique Latine Formation academique - European Commission (ALFA-EC) and the EPLANET Program (European Particle Physics Latin American Network); Stichting voor Fundamenteel Onderzoek der Materie (FOM) and the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), Netherlands; Research Council of Norway (NFR); National Science Centre, Poland; Ministry of National Education/Institute for Atomic Physics and National Council of Scientific Research in Higher Education (CNCSI-UEFISCDI), Romania; Ministry of Education and Science of Russian Federation, Russian Academy of Sciences, Russian Federal Agency of Atomic Energy, Russian Federal Agency for Science and Innovations and The Russian Foundation for Basic Research; Ministry of Education of Slovakia; Department of Science and Technology, South Africa; Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas (CIEMAT), E-Infrastructure shared between Europe and Latin America (EELA), Ministerio de Economía y Competitividad (MINECO) of Spain, Xunta de Galicia (Consellería de Educación), Centro de Aplicaciones Tecnolgicas y Desarrollo Nuclear (CEADEN), Cubaenergía, Cuba, and IAEA (International Atomic Energy Agency); Swedish Research Council (VR) and Knut & Alice Wallenberg Foundation (KAW); Ukraine Ministry

of Education and Science; United Kingdom Science and Technology Facilities Council (STFC); The United States Department of Energy, the United States National Science Foundation, the State of Texas, and the State of Ohio; Ministry of Science, Education and Sports of Croatia and Unity through Knowledge Fund, Croatia; Council of Scientific and Industrial Research (CSIR), New Delhi, India; Pontificia Universidad Católica del Perú.

Open Access This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Appendix A. The ALICE Collaboration

Adamaya J. D. Adamova<sup>84</sup>. M.M. Agarwal<sup>88</sup>, G. Aglieri Rinella<sup>35</sup>, M. Agnello<sup>110</sup>, N. Agrawal<sup>17</sup>, Z. Ahammed<sup>132</sup>, S. Ahmad<sup>19</sup>, S.U. Ahn<sup>68</sup>, S. Alola<sup>168</sup>, A. Akindinov<sup>50</sup>, S.N. Alam<sup>123</sup>, D.S.D. Albuquerque<sup>213</sup>, D. Aleksandrov<sup>50</sup>, B. Alessandro<sup>100</sup>, D. Alexandre<sup>107</sup>, R. Alfaro Molina<sup>64</sup>, A. Alain<sup>124</sup>, A. Akini, J.R.M. Almaraz<sup>107</sup>, J. Alus<sup>27</sup>, T. Alti-18, A. Altipoliava<sup>13</sup>, L. Alyspocavilla, C. Alves Garcia Pradol<sup>202</sup>, C. Andrein<sup>17</sup>, A. Andronic<sup>17</sup>, V. Anguellov<sup>34</sup>, T. Anticic<sup>88</sup>, F. Antinori<sup>167</sup>, P. Antonioli<sup>168</sup>, L. Aphecetche<sup>130</sup>, H. Appelshäuser<sup>25</sup>, S. Arcelli<sup>27</sup>, R. Arnaldi<sup>110</sup>, O.W. Arnoldo<sup>369</sup>, I.C. Arsene<sup>22</sup>, M. Arslandok<sup>38</sup>, B. Augustimus<sup>35</sup>, R. Averbeck<sup>37</sup>, N. Bartalla<sup>110</sup>, A. Badalah<sup>100</sup>, W. W. Back<sup>77</sup>, S. Bagnasco<sup>110</sup>, R. Balibache<sup>35</sup>, R. Balas<sup>18</sup>, S. Balasubramanian<sup>150</sup>, A. Baldiscer<sup>15</sup>, R. C. Bartallin<sup>17</sup>, K. Bartha<sup>30</sup>, J. Bartkel<sup>117</sup>, E. Bartsch<sup>35</sup>, M. Basilc<sup>27</sup>, N. Bastlc<sup>70</sup>, S. Basil<sup>30</sup>, S. Bashenber<sup>30</sup>, D. Battwall<sup>30</sup>, P. C. Batzing<sup>21</sup>, I. Bellomont<sup>31</sup>, E. Belmont Moreno<sup>4</sup>, V. Belyear<sup>50</sup>, D. Bercani<sup>318</sup>, S. Recruella<sup>37</sup>, N. Berteno<sup>47</sup>, D. Bertano<sup>48</sup>, J. Belyear<sup>50</sup>, D. Bercani<sup>318</sup>, A. Berteno<sup>47</sup>, D. Bertano<sup>48</sup>, J. Belyear<sup>50</sup>, D. Bercani<sup>318</sup>, J. Bielchoont<sup>318</sup>, A. Bienohal<sup>318</sup>, J. Bellomont<sup>318</sup>, R. Belyear<sup>50</sup>, D. Bercani<sup>318</sup>, J. Bielchoont<sup>318</sup>, A. Bienohal<sup>318</sup>, J. Bielchoont<sup>318</sup>, J. Bielchoont<sup>318</sup>, J. Bielchoont<sup>318</sup>, J. Bielchoont<sup>318</sup>, J. Bielchoont<sup>318</sup>, J. J.F. Grosse-Oetringhaus<sup>35</sup>, R. Grosso<sup>97</sup>, F. Guber<sup>56</sup>, R. Guernane<sup>71</sup>, B. Guerzoni<sup>27</sup>, K. Gulbrandsen<sup>81</sup>, T. Gunji<sup>127</sup>,

<sup>&</sup>lt;sup>3</sup> e-mail:alice-publications@cern.ch

<sup>&</sup>lt;sup>4</sup> Also at: Georgia State University, Atlanta, Georgia, United States.

A. Gupta<sup>10</sup>, R. Gupta<sup>10</sup>, R. Haske<sup>20</sup>, O. Haslandi, C. Hadjidakis<sup>21</sup>, M. Haiduc<sup>12</sup>, H. Humagaki<sup>127</sup>, G. Hamar<sup>133</sup>, J.C. Hamar<sup>134</sup>, J.C. Hamar<sup>135</sup>, A. Hartoni, C. Hartoni, D. Hatzifotiadou<sup>13</sup>, S. Hayashi<sup>127</sup>, S. J. Hecke<sup>13</sup>, E. Hellbar<sup>13</sup>, D. Horak<sup>13</sup>, R. Hersolawa<sup>138</sup>, P. Histov<sup>13</sup>, T. J. Humanic<sup>138</sup>, N. Hussaini, T. Hussaini, D. Hutter<sup>13</sup>, D. Horak<sup>23</sup>, R. Hesolawa<sup>138</sup>, P. Histov<sup>13</sup>, T. J. Humanic<sup>138</sup>, N. Hussaini, T. Hussaini, D. Hutter<sup>13</sup>, D. Horak<sup>23</sup>, R. Haslov<sup>138</sup>, P. Histov<sup>138</sup>, T. J. Humanic<sup>138</sup>, N. Hussaini, T. Hussaini, D. Hutter<sup>138</sup>, D. Horak<sup>138</sup>, D. Horak<sup>138</sup>, D. Horak<sup>138</sup>, D. Horak<sup>138</sup>, D. Hartoni, D. Hatter<sup>138</sup>, D. Hartoni, D. Hutter<sup>138</sup>, D. Hartoni, D. Harton A. Gupta<sup>91</sup>, R. Gupta<sup>91</sup>, R. Haake<sup>35</sup>, Ø. Haaland<sup>18</sup>, C. Hadjidakis<sup>51</sup>, M. Haiduc<sup>62</sup>, H. Hamagaki<sup>127</sup>, G. Hamar<sup>135</sup>, J.C. Hamon<sup>55</sup>, J.W. Harris<sup>136</sup>, A. Harton<sup>13</sup>, D. Hatzifotiadou<sup>104</sup>, S. Hayashi<sup>127</sup>, S.T. Heckel<sup>53</sup>, E. Hellbär<sup>53</sup>, H. Helstrup<sup>37</sup>, A. Herghelegiu<sup>78</sup>, G. Herrera Corral<sup>11</sup>, B.A. Hess<sup>34</sup>, K.F. Hetland<sup>37</sup>, H. Hillemanns<sup>35</sup>, B. Hippolyte<sup>55</sup>,

 $<sup>{\</sup>it Also~at}$ : Department of Applied Physics, Aligarh Muslim University, Aligarh, India.

Also at: M.V. Lomonosov Moscow State University, D.V. Skobeltsyn Institute of Nuclear, Physics, Moscow, Russia.

V. Samsonov<sup>86</sup>, L. Šándor<sup>59</sup>, A. Sandoval<sup>64</sup>, M. Sano<sup>128</sup>, D. Sarkar<sup>132</sup>, P. Sarkar<sup>132</sup>, P. Sarma<sup>44</sup>, E. Scapparonel<sup>104</sup>, F. Scarlassara<sup>29</sup>, C. Schiukar<sup>38</sup>, R. Schicker<sup>94</sup>, C. Schmidt<sup>34</sup>, S. Schuchmann<sup>53</sup>, J. Schukrat<sup>155</sup>, M. Schuka<sup>39</sup>, Y. Schutz<sup>255,113</sup> K. Schwar<sup>97</sup>, K. Schweda<sup>97</sup>, G. Scioli<sup>27</sup>, E. Scomparin<sup>110</sup>, R. Scott<sup>123</sup>, M. Sefcik<sup>40</sup>, J.E. Seger<sup>87</sup>, Y. Sekiguchi<sup>127</sup>, D. Sekihata<sup>46</sup>, I. Selyuzhenkov<sup>97</sup>, K. Senosi<sup>65</sup>, S. Senyukov<sup>3,35</sup>, E. Serradilla<sup>10,64</sup>, A. Sevenero<sup>62</sup>, A. Shabanor<sup>56</sup>, A. Shabetai<sup>113</sup>, O. Shadura<sup>3</sup>, R. Shahoyan<sup>35</sup>, M.I. Shahzad<sup>16</sup>, A. Shargari, M. Sharma<sup>91</sup>, M. Sharma<sup>91</sup>, N. Sharma<sup>92</sup>, A.I. Shickhi<sup>132</sup>, K. Shigkai<sup>46</sup>, Q. Shon<sup>7</sup>, K. Shigeip<sup>40</sup>, Y. Sibiriak<sup>80</sup>, S. Siddhanta<sup>105</sup>, K.M. Sielewicz<sup>35</sup>, T. Siemiarczuk<sup>77</sup>, D. Silvermyr<sup>33</sup>, C. Silvestre<sup>71</sup>, G. Simatovic<sup>129</sup>, G. Simonetti<sup>35</sup>, R. Singaraju<sup>132</sup>, R. Singh<sup>79</sup>, S. Singha<sup>79,132</sup>, V. Singhal<sup>133</sup>, B.C. Sinhai<sup>132</sup>, T. Sinha<sup>100</sup>, B. Sitar<sup>35</sup>, M. Sitta<sup>31</sup>, T.B. Skaali<sup>22</sup>, M. Supceki<sup>123</sup>, N. Smirnov<sup>136</sup>, R.J.M. Snellings<sup>57</sup>, T.W. Snellman<sup>123</sup>, J. Song<sup>86</sup>, M. Song<sup>137</sup>, Z. Song<sup>7</sup>, F. Soramel<sup>29</sup>, S. Sorensen<sup>125</sup>, R.D.de Souza<sup>121</sup>, F. Sozzi<sup>97</sup>, M. Spacek<sup>39</sup>, E. Spiriti<sup>72</sup>, I. Sputowska<sup>117</sup>, M. Spyropoulou-Stassinaki<sup>39</sup>, J. Stachel<sup>34</sup>, I. Stan<sup>62</sup>, P. Stankus<sup>85</sup>, E. Stenlund<sup>33</sup>, G. Stevn<sup>65</sup>, J.H. Stiller<sup>94</sup>, D. Stocco<sup>113</sup>, P. Strmen<sup>38</sup>, A.A.P. Suaide<sup>120</sup>, T. Sugitate<sup>46</sup>, C. Suire<sup>51</sup>, M. Suleymanov<sup>16</sup>, M. Suljic<sup>25</sup>, c. R. Sultanov<sup>58</sup>, M. Sumbera<sup>84</sup>, S. Sumowidagdo<sup>49</sup>, A. Szabo<sup>38</sup>, A. Szanto de Toledo<sup>120</sup>, I. Szarka<sup>38</sup>, A. Szczepankiewicz<sup>55</sup>, M. Szymanski<sup>133</sup>, U. Tabassam<sup>16</sup>, J. Takahashi<sup>121</sup>, G.J. Tambu<sup>120</sup>, T. S. Tveter<sup>22</sup>, K. Ulaland<sup>18</sup>, M. Tariol<sup>19</sup>, M.G. Tarzila<sup>78</sup>, A. Tauro<sup>35</sup>, G. Tejeda Muñoz<sup>2</sup>, A. Telesca<sup>35</sup>, K. Terasaki<sup>127</sup>, C. Terrevoli<sup>29</sup>, B. Teyssier<sup>130</sup>, J. Thäder<sup>74</sup>, D. Thakur<sup>48</sup>, D. Thomas<sup>118</sup>, R. Tieulent<sup>130</sup>, A.R. Timminis<sup>122</sup>, A. Toiopo<sup>5</sup>, S. Trogolo<sup>50</sup>, G. Trombetta<sup>32</sup>, V. Trubnikov<sup>3</sup>, W.H. Trzaska<sup>123</sup>, T. Tsuji<sup>127</sup>

<sup>1</sup>A.I. Alikhanyan National Science Laboratory (Yerevan Physics Institute) Foundation, Yerevan, Armenia

<sup>&</sup>lt;sup>2</sup>Benemérita Universidad Autónoma de Puebla, Puebla, Mexico

<sup>&</sup>lt;sup>3</sup>Bogolyubov Institute for Theoretical Physics, Kiev, Ukraine

 $<sup>{\</sup>color{blue}^4} Bose\ Institute, Department\ of\ Physics\ and\ Centre\ for\ Astroparticle\ Physics\ and\ Space\ Science\ (CAPSS),\ Kolkata,\ India 1999 and\ Space\ Space\$ 

<sup>&</sup>lt;sup>5</sup>Budker Institute for Nuclear Physics, Novosibirsk, Russia

<sup>&</sup>lt;sup>6</sup>California Polytechnic State University, San Luis Obispo, California, USA

<sup>&</sup>lt;sup>7</sup>Central China Normal University, Wuhan, China

<sup>&</sup>lt;sup>8</sup>Centre de Calcul de l'IN2P3, Villeurbanne, France

<sup>&</sup>lt;sup>9</sup>Centro de Aplicaciones Tecnológicas y Desarrollo Nuclear (CEADEN), Havana, Cuba

<sup>&</sup>lt;sup>10</sup>Centro de Investigaciones Energéticas Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

<sup>&</sup>lt;sup>11</sup>Centro de Investigación y de Estudios Avanzados (CINVESTAV), Mexico City and Mérida, Mexico

<sup>&</sup>lt;sup>12</sup>Centro Fermi - Museo Storico della Fisica e Centro Studi e Ricerche "Enrico Fermi", Rome, Italy

<sup>&</sup>lt;sup>13</sup>Chicago State University, Chicago, Illinois, USA

<sup>&</sup>lt;sup>14</sup>China Institute of Atomic Energy, Beijing, China

<sup>&</sup>lt;sup>15</sup>Commissariat à l'Energie Atomique, IRFU, Saclay, France

<sup>&</sup>lt;sup>16</sup>COMSATS Institute of Information Technology (CIIT), Islamabad, Pakistan

<sup>&</sup>lt;sup>17</sup>Departamento de Física de Partículas and IGFAE, Universidad de Santiago de Compostela, Santiago de Compostela, Spain

<sup>&</sup>lt;sup>18</sup>Department of Physics and Technology, University of Bergen, Bergen, Norway

<sup>&</sup>lt;sup>19</sup>Department of Physics, Aligarh Muslim University, Aligarh, India

<sup>&</sup>lt;sup>20</sup>Department of Physics, Ohio State University, Columbus, Ohio, USA

<sup>&</sup>lt;sup>21</sup>Department of Physics, Sejong University, Seoul, South Korea

<sup>&</sup>lt;sup>22</sup>Department of Physics, University of Oslo, Oslo, Norway

<sup>&</sup>lt;sup>23</sup>Dipartimento di Fisica dell'Università "La Sapienza" and Sezione INFN Rome, Italy

<sup>&</sup>lt;sup>24</sup>Dipartimento di Fisica dell'Università and Sezione INFN, Cagliari, Italy

<sup>&</sup>lt;sup>25</sup>Dipartimento di Fisica dell'Università and Sezione INFN, Trieste, Italy

<sup>&</sup>lt;sup>26</sup>Dipartimento di Fisica dell'Università and Sezione INFN, Turin, Italy

- <sup>27</sup>Dipartimento di Fisica e Astronomia dell'Università and Sezione INFN, Bologna, Italy
- <sup>28</sup>Dipartimento di Fisica e Astronomia dell'Università and Sezione INFN, Catania, Italy
- <sup>29</sup>Dipartimento di Fisica e Astronomia dell'Università and Sezione INFN, Padova, Italy
- <sup>30</sup>Dipartimento di Fisica "E.R. Caianiello" dell'Università and Gruppo Collegato INFN, Salerno, Italy
- <sup>31</sup>Dipartimento di Scienze e Innovazione Tecnologica dell'Università del Piemonte Orientale and Gruppo Collegato INFN, Alessandria, Italy
- <sup>32</sup>Dipartimento Interateneo di Fisica "M. Merlin" and Sezione INFN, Bari, Italy
- <sup>33</sup>Division of Experimental High Energy Physics, University of Lund, Lund, Sweden
- <sup>34</sup>Eberhard Karls Universität Tübingen, Tübingen, Germany
- <sup>35</sup>European Organization for Nuclear Research (CERN), Geneva, Switzerland
- <sup>36</sup>Excellence Cluster Universe, Technische Universität München, Munich, Germany
- <sup>37</sup>Faculty of Engineering, Bergen University College, Bergen, Norway
- <sup>38</sup>Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovakia
- <sup>39</sup> Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Prague, Czech Republic
- <sup>40</sup>Faculty of Science, P.J. Šafárik University, Košice, Slovakia
- <sup>41</sup>Faculty of Technology, Buskerud and Vestfold University College, Vestfold, Norway
- <sup>42</sup>Frankfurt Institute for Advanced Studies, Johann Wolfgang Goethe-Universität Frankfurt, Frankfurt, Germany
- <sup>43</sup>Gangneung-Wonju National University, Gangneung, South Korea
- <sup>44</sup>Gauhati University, Department of Physics, Guwahati, India
- <sup>45</sup>Helsinki Institute of Physics (HIP), Helsinki, Finland
- $^{46}\mathrm{Hiroshima}$  University, Hiroshima, Japan
- <sup>47</sup>Indian Institute of Technology Bombay (IIT), Mumbai, India
- <sup>48</sup>Indian Institute of Technology Indore, Indore (IITI), India
- <sup>49</sup>Indonesian Institute of Sciences, Jakarta, Indonesia
- <sup>50</sup>Inha University, Incheon, South Korea
- <sup>51</sup>Institut de Physique Nucléaire d'Orsay (IPNO), Université Paris-Sud, CNRS-IN2P3, Orsay, France
- <sup>52</sup>Institut für Informatik, Johann Wolfgang Goethe-Universität Frankfurt, Frankfurt, Germany
- <sup>53</sup>Institut für Kernphysik, Johann Wolfgang Goethe-Universität Frankfurt, Frankfurt, Germany
- <sup>54</sup>Institut für Kernphysik, Westfälische Wilhelms-Universität Münster, Münster, Germany
- <sup>55</sup>Institut Pluridisciplinaire Hubert Curien (IPHC), Université de Strasbourg, CNRS-IN2P3, Strasbourg, France
- <sup>56</sup>Institute for Nuclear Research, Academy of Sciences, Moscow, Russia
- <sup>57</sup>Institute for Subatomic Physics of Utrecht University, Utrecht, the Netherlands
- <sup>58</sup>Institute for Theoretical and Experimental Physics, Moscow, Russia
- <sup>59</sup>Institute of Experimental Physics, Slovak Academy of Sciences, Košice, Slovakia
- <sup>60</sup>Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic
- <sup>61</sup>Institute of Physics, Bhubaneswar, India
- <sup>62</sup>Institute of Space Science (ISS), Bucharest, Romania
- <sup>63</sup>Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Mexico City, Mexico
- <sup>64</sup>Instituto de Física, Universidad Nacional Autónoma de México, Mexico City, Mexico
- <sup>65</sup>iThemba LABS, National Research Foundation, Somerset West, South Africa
- <sup>66</sup>Joint Institute for Nuclear Research (JINR), Dubna, Russia
- <sup>67</sup>Konkuk University, Seoul, South Korea
- <sup>68</sup>Korea Institute of Science and Technology Information, Daejeon, South Korea
- <sup>69</sup>KTO Karatay University, Konya, Turkey
- <sup>70</sup>Laboratoire de Physique Corpusculaire (LPC), Clermont Université, Université Blaise Pascal, CNRS-IN2P3, Clermont-Ferrand, France
- <sup>71</sup>Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS-IN2P3, Grenoble, France
- <sup>72</sup>Laboratori Nazionali di Frascati, INFN, Frascati, Italy
- <sup>73</sup>Laboratori Nazionali di Legnaro, INFN, Legnaro, Italy
- <sup>74</sup>Lawrence Berkeley National Laboratory, Berkeley, California, USA
- <sup>75</sup>Moscow Engineering Physics Institute, Moscow, Russia
- <sup>76</sup>Nagasaki Institute of Applied Science, Nagasaki, Japan
- <sup>77</sup>National Centre for Nuclear Studies, Warsaw, Poland
- <sup>78</sup>National Institute for Physics and Nuclear Engineering, Bucharest, Romania
- <sup>79</sup>National Institute of Science Education and Research, Bhubaneswar, India
- <sup>80</sup>National Research Centre Kurchatov Institute, Moscow, Russia
- <sup>81</sup>Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark
- <sup>82</sup>Nikhef, Nationaal instituut voor subatomaire fysica, Amsterdam, Netherlands
- <sup>83</sup>Nuclear Physics Group, STFC Daresbury Laboratory, Daresbury, UK

- <sup>84</sup>Nuclear Physics Institute, Academy of Sciences of the Czech Republic, Řež u Prahy, Czech Republic
   <sup>85</sup>Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
   <sup>86</sup>Petersburg Nuclear Physics Institute, Gatchina, Russia
   <sup>87</sup>Physics Department, Creighton University, Omaha, Nebraska, USA
- <sup>88</sup>Physics Department, Panjab University, Chandigarh, India
- <sup>89</sup>Physics Department, University of Athens, Athens, Greece
- <sup>90</sup>Physics Department, University of Cape Town, Cape Town, South Africa
- <sup>91</sup>Physics Department, University of Jammu, Jammu, India
   <sup>92</sup>Physics Department, University of Rajasthan, Jaipur, India
- 93 Physik Department, Technische Universität München, Munich, Germany
- <sup>94</sup>Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany
- <sup>95</sup>Purdue University, West Lafayette, Indiana, USA
- <sup>96</sup>Pusan National University, Pusan, South Korea
- <sup>97</sup>Research Division and ExtreMe Matter Institute EMMI, GSI Helmholtzzentrum für Schwerionenforschung, Darmstadt, Germany
- <sup>98</sup>Rudjer Bošković Institute, Zagreb, Croatia
- <sup>99</sup>Russian Federal Nuclear Center (VNIIEF), Sarov, Russia
- $^{100}\mathrm{Saha}$ Institute of Nuclear Physics, Kolkata, India
- $^{101}\mathrm{School}$  of Physics and Astronomy, University of Birmingham, Birmingham, UK
- 102 Sección Física, Departamento de Ciencias, Pontificia Universidad Católica del Perú, Lima, Peru
- <sup>103</sup>Sezione INFN, Bari, Italy
- $^{104}\mathrm{Sezione}$  INFN, Bologna, Italy
- $^{105}\mathrm{Sezione}$ INFN, Cagliari, Italy
- $^{106}\mathrm{Sezione}$  INFN, Catania, Italy
- <sup>107</sup>Sezione INFN, Padova, Italy
- <sup>108</sup>Sezione INFN, Rome, Italy
- <sup>109</sup>Sezione INFN, Trieste, Italy
- $^{110}$ Sezione INFN, Turin, Italy
- <sup>111</sup>SSC IHEP of NRC Kurchatov institute, Protvino, Russia
- $^{112}\mathrm{Stefan}$  Meyer Institut für Subatomare Physik (SMI), Vienna, Austria
- <sup>113</sup>SUBATECH, Ecole des Mines de Nantes, Université de Nantes, CNRS-IN2P3, Nantes, France
- $^{114}\mathrm{Suranaree}$  University of Technology, Nakhon Ratchasima, Thailand
- $^{115}\mathrm{Technical}$  University of Košice, Košice, Slovakia
- <sup>116</sup>Technical University of Split FESB, Split, Croatia
- <sup>117</sup>The Henryk Niewodniczanski Institute of Nuclear Physics, Polish Academy of Sciences, Cracow, Poland
- <sup>118</sup>The University of Texas at Austin, Physics Department, Austin, Texas, USA
- <sup>119</sup>Universidad Autónoma de Sinaloa, Culiacán, Mexico
- <sup>120</sup>Universidade de São Paulo (USP), São Paulo, Brazil
- <sup>121</sup>Universidade Estadual de Campinas (UNICAMP), Campinas, Brazil
- $^{122}\mathrm{University}$  of Houston, Houston, Texas, USA
- <sup>123</sup>University of Jyväskylä, Jyväskylä, Finland
- <sup>124</sup>University of Liverpool, Liverpool, UK
- <sup>125</sup>University of Tennessee, Knoxville, Tennessee, USA
- <sup>126</sup>University of the Witwatersrand, Johannesburg, South Africa
- <sup>127</sup>University of Tokyo, Tokyo, Japan
- <sup>128</sup>University of Tsukuba, Tsukuba, Japan
- <sup>129</sup>University of Zagreb, Zagreb, Croatia
- <sup>130</sup>Université de Lyon, Université Lyon 1, CNRS/IN2P3, IPN-Lyon, Villeurbanne, France
- <sup>131</sup>V. Fock Institute for Physics, St. Petersburg State University, St. Petersburg, Russia
- <sup>132</sup>Variable Energy Cyclotron Centre, Kolkata, India
- <sup>133</sup>Warsaw University of Technology, Warsaw, Poland
- <sup>134</sup>Wayne State University, Detroit, Michigan, USA
- <sup>135</sup>Wigner Research Centre for Physics, Hungarian Academy of Sciences, Budapest, Hungary
- <sup>136</sup>Yale University, New Haven, Connecticut, USA
- <sup>137</sup>Yonsei University, Seoul, South Korea
- <sup>138</sup>Zentrum für Technologietransfer und Telekommunikation (ZTT), Fachhochschule Worms, Worms, Germany

#### References

- 1. ALICE Collaboration (B. Abelev et al.), Int. J. Mod. Phys. A 29, 1430044 (2014) arXiv:1402.4476 [nucl-ex].
- 2. LEBC-EHS Collaboration (M. Aguilar-Benitez et al.), Z. Phys. C 50, 405 (1991).
- 3. HADES Collaboration (G. Agakishiev et al.), Eur. Phys. J. A 40, 45 (2009) arXiv:0902.4377 [nucl-ex].
- 4. X. Ye, H. Jian, G. Kai-En, Chin. Phys. C 32, 201 (2008).
- I. Belikov, P. Hristov, M. Ivanov, T. Kuhr, K. Šafařik, Bayesian approach for particle identification in ALICE experiment at LHC, in Computing in High Energy Physics and Nuclear Physics. Proceedings of CHEP04 Conference, Interlaken, Switzerland, 2004 (2005).
- 6. ALICE Collaboration (F. Carminati et al.), J. Phys. G 30, 1517 (2004).
- 7. ALICE Collaboration (B. Alessandro et al.), J. Phys. G 32, 1295 (2006).
- 8. ALICE Collaboration (B. Abelev et al.), Eur. Phys. J. C 75, 226 (2015) arXiv:1504.00024 [nucl-ex].
- 9. ALICE Collaboration (B. Abelev et al.), Phys. Lett. B 736, 196 (2014) arXiv:1401.1250 [nucl-ex].
- 10. ALICE Collaboration (B. Abelev et al.), Phys. Rev. Lett. 109, 252301 (2012) arXiv:1208.1974 [hep-ex].
- 11. ALICE Collaboration (B. Abelev et al.), Phys. Lett. B 721, 13 (2013) arXiv:1208.1902 [hep-ex].
- 12. ALICE Collaboration (B. Abelev et al.), Phys. Rev. D 91, 012001 (2015) arXiv:1405.4117 [nucl-ex].
- 13. ALICE Collaboration (B. Abelev *et al.*), JHEP **01**, 128 (2012) arXiv:1111.1553 [hep-ex].
- 14. ALICE Collaboration (B. Abelev et al.), Phys. Rev. C 91, 024609 (2015) arXiv:1404.0495 [nucl-ex].
- 15. ALICE Collaboration (J. Adam et al.), Phys. Rev. C 91, 034906 (2015) arXiv:1502.00559 [nucl-ex].
- 16. ALICE Collaboration (J. Adam et al.), Nature Phys. 11, 811 (2015) arXiv:1508.03986 [nucl-ex].
- 17. ALICE Collaboration, Technical Design Report of the Time Projection Chamber, CERN/LHCC 2000-001, 2000, on-line at: https://edms.cern.ch/document/398930/1.
- 18. A. Akindinov et al., Eur. Phys. J. Plus 128, 44 (2013).
- 19. W. Blum, L. Rolandi, Particle Detection with Drift Chambers (Springer-Verlag, 1994).
- 20. P. Gregory, Bayesian Logical Data Analysis for the Physical Sciences (Cambridge University Press, 2005).
- 21. ALICE Collaboration (B. Abelev et al.), Phys. Lett. B **728**, 25 (2014) arXiv:1307.6796 [nucl-ex].
- 22. ALICE Collaboration (B. Abelev et al.), JHEP 06, 190 (2015) arXiv:1405.4632 [nucl-ex].
- 23. T. Sjöstrand, S. Mrenna, P.Z. Skands, JHEP 05, 026 (2006) arXiv:hep-ph/0603175 [hep-ex].
- 24. P.Z. Skands, Phys. Rev. D 82, 074018 (2010) arXiv:1005.3457 [hep-ex].
- R. Brun, F. Carminati, S. Giani, GEANT Detector Description and Simulation Tool, Tech. Rep. CERN-W5013, CERN, 1994.
- 26. Particle Data Group Collaboration (K.A. Olive et al.), Chin. Phys. C 38, 090001 (2014).