Electron identification and hadron contamination studies in proton-proton collisions with ALICE

Anisa Dashi on behalf of the ALICE Collaboration[*]

Technical University Munich, Excellence Cluster Universe, Boltzmanstr. 2, 85748 Garching
E-mail: anisa.dashi@cern.ch

The continuum of electron-positron pairs, produced in heavy ion collisions, provides an excellent probe of quark-gluon plasma formation in nucleus-nucleus collisions. Dielectrons are also studied in proton-proton collisions to provide a crucial reference. To extract possible signatures, it is important to work with an electron sample of high purity. This can be achieved optimizing the electron identification and reducing the hadron contamination. In this proceeding, the analysis of a dataset of pp collisions at $\sqrt{s}$ = 7 TeV measured with the ALICE detector at the Large Hadron Collider is described. The invariant mass spectra of electron-positron pairs based on three different PID sets are compared and the hadron contamination of the dielectron sample is estimated as a function of invariant mass. The analysis shows that one can achieve good electron identification efficiencies and a single electron purity higher than 90% by combining the signals of the ITS, TPC and TOF detectors in ALICE.

---

[*]Speaker.

1. Introduction

In heavy ions collisions at the Large Hadron Collider (LHC), where very high energy densities are reached [1], a hot and dense matter is produced, the quark-gluon plasma (QGP) [2]. Along with this phase transition of the matter, chiral symmetry restoration (CSR) is predicted [3], which affects the hadron spectral properties. This phenomenon was already investigated by experiments at the SPS and RHIC [4, 5, 6, 7]. Dileptons coming from vector meson decays, like the ρ meson, carry information about these properties since their invariant masses (m_{ee}) correspond to the masses of the vector mesons at the time of the decay and, as they undergo only electromagnetic interactions, they do not strongly interact with the hadronic medium after emission. Dileptons are produced by different processes at all stages of the collision, hence their invariant mass spectrum contains various contributions. To extract possible hints for CSR or QGP from the continuum of e^+e^- pairs, it is essential to work with a high purity sample of electrons and positrons, i.e. not contaminated by hadrons. Besides in heavy-ion collisions, it is important to study dielectrons also in proton-proton collisions, where QGP and CSR are not expected due to the insufficient energy density, in order to have reference data which are crucial for the interpretation of the heavy-ion data. In this work, a data set of 200 million proton-proton collisions at $\sqrt{s} = 7$ TeV measured with ALICE (A Large Ion Collider Experiment) at the LHC in 2010 is analyzed with the aim of obtaining a high purity electron sample.

2. Particle Identification with ALICE

The ALICE apparatus is shown in Fig. 1 and comprises several subdetectors [8]. This analysis
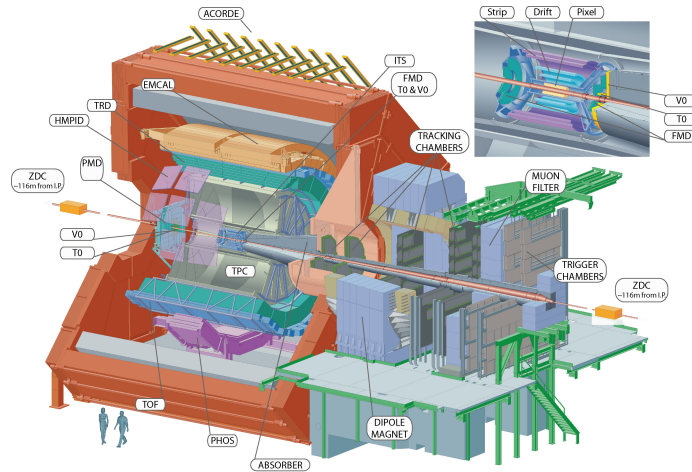

Figure 1: Setup of the ALICE detector system.

makes use of the Inner Tracking System (ITS), Time Projection Chamber (TPC) and Time of Flight (TOF) detectors. All three are enclosed in a solenoid with a magnetic field of 0.5 T that allows the transverse momentum (p_T) of charged particles to be determined. The ITS and TPC are both used for tracking and particle identification (PID) via their specific energy loss (dE/dx), which is parametrized by the Bethe-Bloch equation. The TOF provides information about the time of

flight of the particles, hence of their velocity, helping with the PID in the low and intermediate momentum range. If the measured dE/dx distribution is combined with the momentum, the mass of the particles can be derived, i.e. particles with different masses can be distinguished from each other. Projecting this distribution for a given momentum range, one gets a approximately Gaussian distribution for each particle type: e, π, K, p, d. Using a Gaussian fit, the dE/dx spectra can be parametrized around the mean energy loss value expected for each particle type. For that, the mean energy loss is subtracted from the measured dE/dx value and then the result is divided by the width (σ) of the fit. The energy loss is than expressed in number of sigmas for each particle, with the considered particle type centered at zero. Figure 2 shows the $n\sigma_e$ distribution as a function of momentum in the TPC. The ITS energy loss distribution and the TOF velocity distribution are parametrized in the same way.
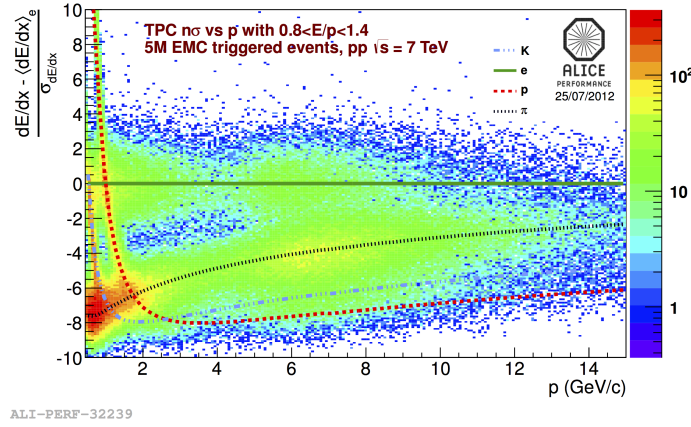

ALI–PERF–32239

Figure 2: The dE/dx distribution parametrized around the expected electron signal as a function of momentum measured in pp collisions at $\sqrt{s} = 7$ TeV.

3. Analysis

3.1 PID

To select electron candidates, tracks with dE/dx in a certain $n\sigma_e$ range have to be selected. The biggest contamination in the electron sample comes from pions, since they are the particles most copiously produced in a collision. Since the TPC has a good electron/pion separation, a first attempt to identify electrons is to require electrons candidates to fall into the $n\sigma_e^{\text{TPC}}$ range of -1.5 to 3 for the whole momentum range ($0.2 < p < 100$ GeV/c). Besides the electron selection, a pion rejection is needed since the separation power between electrons and pions is momentum dependent. Hence, pions are rejected for $n\sigma_\pi^{\text{TPC}} < 4$. In this first cut combination only the TPC is used to select electron candidates, therefore it is called TPC-only. With this kind of selection the electron sample still contains kaons and protons, as the bands of these particles cross the electron band for $p \approx 0.6$ GeV/c and $p \approx 1$ GeV/c, respectively. Since the TOF has a good separation of electrons from kaons and protons, the effect of adding the TOF information to the TPC cut can be studied by applying a TOF cut along with the TPC PID selection, in a cut combination called TPC/TOF. For that, the $n\sigma_e^{\text{TOF}}$ range from -3 to 3 is selected in the momentum range 0.4

2

to 5 GeV/*c*, while the TPC selection remains the same. As the TOF matching efficiency is only $\approx$ 50%, it is useful to use the ITS information whenever no TOF information is available. The ITS can be used to improve the particle identification since the kaon and proton bands cross the electron band at other momenta than in the TPC due to a different detector material, which leads to different energy loss curves. Hence, in the third cut combination (ITS/TPC/TOFif), a cut is applied on the ITS in the $n\sigma_e^{\text{ITS}}$ range from -5 to 1 for the whole momentum range, while the TOF signal is used only if it is available (TOFif).

3.2 Electron Pair Analysis

Since in electromagnetic decay electrons are always produced in pairs, a pair analysis aims to identify the origin of electrons. As in a single event several e^+e^- pairs are produced from different decays, the source of each electron is a priori unknown. Therefore, each electron has to be combined with all other electrons in the same event. This way, one obtains three different distributions: all e^+e^- pairs, N_{+-}, and two like-sign combinations, N_{++} and N_{--}. The N_{+-} distribution is then composed of physical and combinatorial pairs, while the like-sign spectra contain no physics information but only combinatorial background. For each PID set introduced in section 3.1, the N_{+-} distribution is studied as a function of the invariant mass of the pair, integrated over all pair p_{T}. The unlike-sign background (B_{+-}) contained in these spectra is estimated from the geometric mean of the like-sign pairs distributions: $2\sqrt{N_{++}N_{--}}$. Because of the different acceptance for like- and unlike-sign pairs in the central-barrel detectors, a correction factor $R = \frac{M_{+-}}{2\sqrt{M_{++}M_{--}}}$, calculated with mixed events, is used. Hence, the complete description of the background is given by:

$$B = B_{+-} \cdot R = \sqrt{N_{++}N_{--}} \cdot \frac{M_{+-}}{\sqrt{M_{++}M_{--}}}. \tag{3.1}$$

In this way the physical signal S can be obtained:

$$S = N_{+-} - B \tag{3.2}$$

For a proper comparison of the different PID cuts (TPC-only, TPC/TOF, ITS/TPC/TOFif), the significance of the signal, which is the inverse of the relative statistical uncertainty on the signal (σ_S/S), is studied as a function of invariant mass. Since the N_{+-} and B mass spectra contain count rates that are described by Poisson distributions, their statistical uncertainties are given by the square-root of the yields. The signal statistical uncertainty is then given by the quadratic sum of the errors on N_{+-} and B, so that for the statistical significance one obtains:

$$\frac{S}{\sigma_S} = \frac{S}{\sqrt{S+2B}}. \tag{3.3}$$

The significance as a function of invariant mass is shown for all three PID cuts in Fig. 3. The distributions clearly show that the TPC-only cut yields the largest S/σ_S ratio in almost all invariant mass bins. Hypothetically, the PID set with the higher significance is the most appropriate to identify electrons. However, it is important to take into account the contamination of the signal that comes from hadrons misidentified as electrons, since the like-sign mass spectra do not completely eliminate these hadrons from the unlike-sign spectrum. This happens if hadrons are produced in resonance decays as unlike-sign pairs, e.g. $\phi \rightarrow K^+K^-$. Therefore, a contamination study is performed for the TPC and ITS/TPC/TOFif cuts.
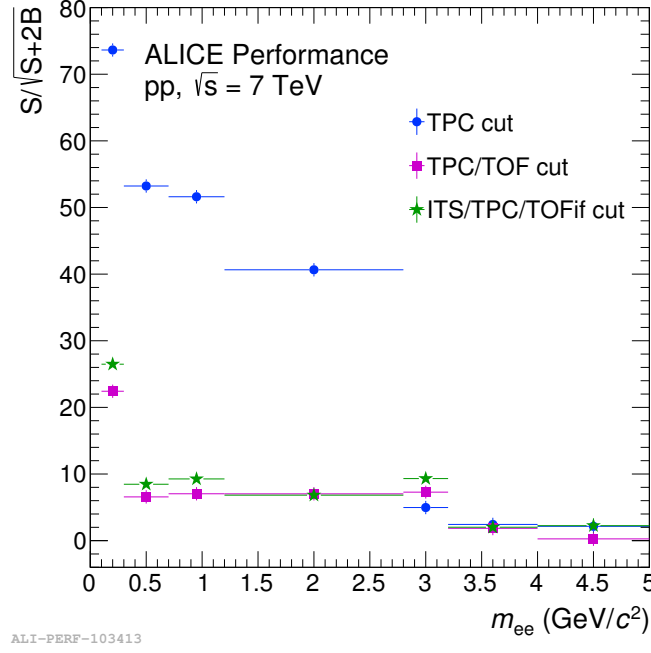
ALI-PERF-103413

Figure 3: Statistical significance $S/\sqrt{S+2B}$ as a function of invariant mass for all three PID cuts.

3.3 Signal Purity Studies

In order to determine the purity of the ITS/TPC/TOFif cut, the electron selection is applied only in the ITS and TOF, while on the TPC no cut is performed. In this way, the hadrons surviving the ITS and TOF selections are still present in the $n\sigma_e^{\mathrm{TPC}}$ vs p distribution, as well as the hadrons that can be rejected with a TPC cut. Therefore, fitting this distribution in momentum intervals of 0.1 GeV/c allows the purity of the ITS/TPC/TOFif PID set to be estimated as a function of p. For each momentum interval, the distribution is projected onto the $n\sigma_e^{\mathrm{TPC}}$ axis and fitted. The signal (electrons) and background (hadrons) are first fitted separately with two different functions, i.e. a Gaussian and the sum of several Gaussians or of a Gaussian and a polynomial, respectively. Then, the whole distribution is fitted with the sum of the signal and background functions maximizing the loglikelihood. The fits allow the purity of the electron sample to be determined by dividing the electron fit function by the total fit function, both integrated in the range $-1.5 < n\sigma_e^{\mathrm{TPC}} < 3$. A fit example is shown in Fig. 4 for the momentum range $0.8 < p < 0.9$ GeV/c. The contamination of the TPC-only cut is determined in a similar way, but fitting the TOF $n\sigma_e$ vs p distribution.

4. Results

The single electron purity of both TPC and ITS/TPC/TOFif PID sets as a function of momentum is shown in Fig. 5. Since the purity of the TPC-only cut is significantly lower than with the ITS/TPC/TOFif PID set, one can conclude that in the TPC cut a fake signal from correlated hadrons is present, which artificially increases the statistical significance of the PID set. Hence, a combined electron selection in the ITS, TPC and TOF detectors proves best to identify electrons in ALICE.
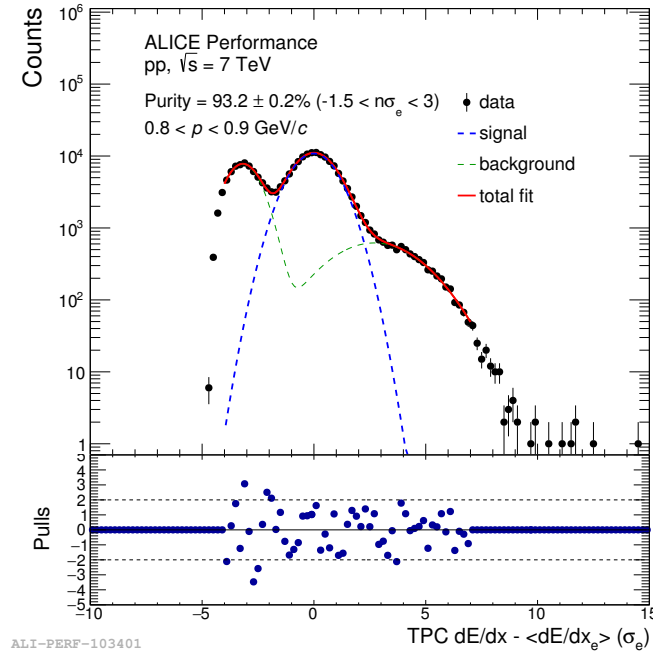
Figure 4: $n\sigma_e^{\mathrm{TPC}}$ ditribution in the momentum range $0.8 < p < 0.9$ GeV/c.
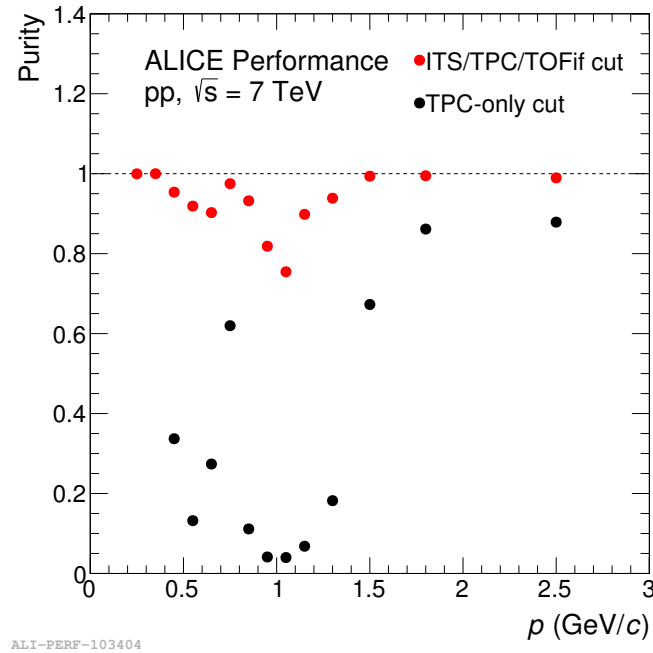

Figure 5: Single electron purity of TPC and ITS/TPC/TOFif PID sets as a function of momentum.

4.1 Signal Purity

To determine the e^+e^- signal purity, one has to calculate first the purity of the dielectron pairs (like-sign and unlike-sign), which is given by the product of the two single electrons purities, assigned to the electrons composing the pair depending on their momentum:

$$P_{ee}(m_{ee}) = P_{e_1}(p_1) \cdot P_{e_2}(p_2) \tag{4.1}$$

Figure 6 shows the pair purity as a function of invariant mass for all three charge combinations: $P_{N_{+-}}$, $P_{N_{++}}$ and $P_{N_{--}}$. As it can be seen, for all three distributions the purity is always higher than
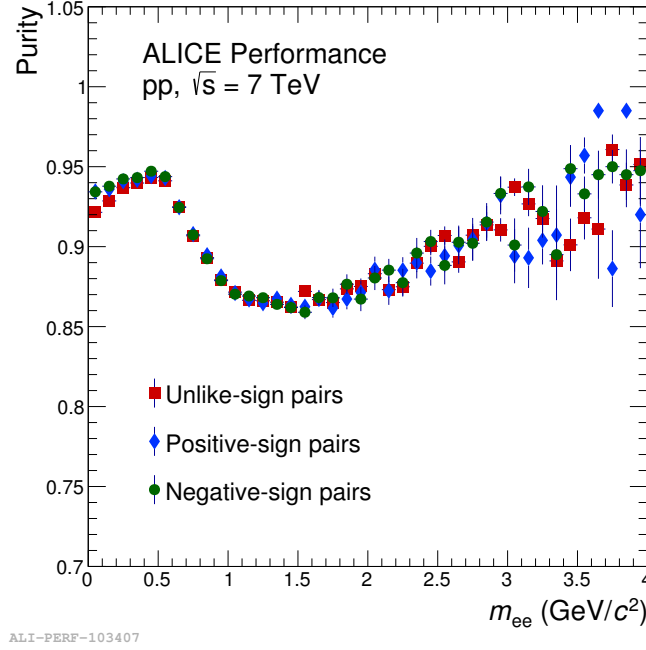

ALI-PERF-103407

Figure 6: Dielectron purity for unlike- and like-sign pairs as a function of invariant mass in pp collisions at $\sqrt{s} = 7$ TeV.

85%. The trend of the distributions reflects the purity at the average momentum of the single electrons composing a pair with a certain invariant mass.

The signal purity of the ITS/TPC/TOFif cut can be calculated from the purity of the unlike-sign pairs, $P_{N_{+-}}$, which is equal to the weighted sum of the signal and background purities, P_S and P_B, since $N_{+-} = S + B$:

$$P_{N_{+-}} = P_S \cdot \frac{S}{N_{+-}} + P_B \cdot \frac{B}{N_{+-}}. \tag{4.2}$$

From this equation, one obtains the purity of the signal:

$$
\begin{aligned}
P_S &= (P_{N_{+-}} \cdot N_{+-} - P_B \cdot B)/S \\
&= P_{N_{+-}} \cdot \left(\frac{B}{S} + 1\right) - P_B \cdot \frac{B}{S},
\end{aligned} \tag{4.3}
$$

with P_B the weighted sum of the like-sign purities:

$$P_B = P_{N_{++}} \cdot \frac{N_{++}}{B} + P_{N_{--}} \cdot \frac{N_{--}}{B}. \tag{4.4}$$

6

The signal purity of the ITS/TPC/TOFif cut is shown in Fig. 7 as a function of invariant mass. As it
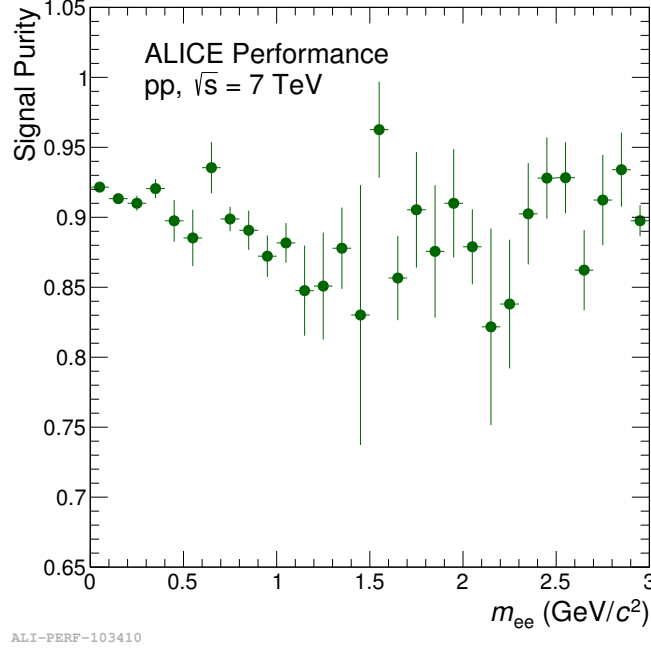

ALI-PERF-103410

Figure 7: Signal purity of the ITS/TPC/TOFif PID set as a function of invariant mass in pp collisions at $\sqrt{s} = 7$ TeV.

can be seen, the signal shows a purity higher than $\approx 82\%$ over the $0 < m_{ee} < 3$ GeV/c^2 invariant mass range. For some invariant mass values, the purity presents large statistical errors, which are due to small S/B ratios. The signal purity distribution reflects the fact that the dielectron signal contains a variety of physics sources, therefore it varies more than the purity of the N_{+-} distribution, which is dominated by combinatorial background. However, a trend roughly similar to the unlike- and like-sign purities in Fig. 6 can be recognized: for low and high invariant mass values the signal purity is higher, while the lowest purity values appear in the intermediate mass region, where the average single electron momentum has a higher probability to fall into the ranges where kaons and protons cross the electron band.

5. Summary and Outlook

In this work, three different PID sets were analyzed and compared regarding their single electron purity and the statistical significance S/σ_S of their dielectron signal, with the aim of optimizing the identification of electrons produced in pp collisions at $\sqrt{s} = 7$ TeV. A combined electron selection in the $n\sigma_e$ vs. p distributions of the ITS, TPC and TOF detectors proved to be the most appropriate to identify electrons among the three analyzed here. For this set, the dielectron purity and the purity of the signal were determined as a function of invariant mass. The signal shows a purity higher than 82%. In the future, one could optimize the model used to fit the data in order to extract the single electron purity, e.g. instead of parametrizations of signal and background shapes, templates obtained via Monte Carlo simulations could be used. A study of the systematic uncertainties of the fits should be performed too. The main result of this analysis, the signal purity of the

ITS/TPC/TOFif PID set, will be used in the dielectron analysis of the LHC 2010 pp data. Furthermore, the framework developed for this work can be used to improve the electron identification in the analysis of the LHC Run-2 pp and Pb-Pb collision data.

References

[1] ALICE Collaboration, arXiv:1603.04775 [nucl-ex].

[2] E. V. Shuryak, Sov. Phys. JETP **47** (1978) 212 [Zh. Eksp. Teor. Fiz. **74** (1978) 408].

[3] F. Karsch and E. Laermann, "Quark Gluon Plasma", Hwa, RC (ed) et al, 1-59 [hep-lat/0305025].

[4] CERES Collaboration, Eur. Phys. J. C **41** (2005) 475 [nucl-ex/0506002].

[5] NA60 Collaboration, Phys. Rev. Lett. **96** (2006) 162302

[6] PHENIX Collaboration, Phys. Rev. C **93** (2016), 014904

[7] STAR Collaboration, Phys. Rev. Lett. **113** (2014), 022301 Addendum: [Phys. Rev. Lett. **113** (2014) no.4, 049903]

[8] ALICE Collaboration, JINST **3** (2008) S08002.

Acknowledgements