Topological separation of dielectron signals using machine learning in Pb–Pb collisions with ALICE

Jerome Jung[1] for the ALICE Collaboration

Dielectron Production

Dielectrons are produced at all stages of the ultra-relativistic heavy-ion collision and leave the system with negligible final-state interaction

→ Ideal probe to study the properties of the created medium

Their invariant mass (m_{ee}) can be utilised to differentiate between early and late contributions of the collision [1]:

→ At higher masses ($1.1 < m_{ee} < 2.7$ GeV/c^2):
- Correlated semi-leptonic decays of heavy-flavor hadrons
- Quark-gluon plasma (QGP)

Heavy-flavor production expected to be modified by cold-nuclear matter and hot-medium effects
- Modeling these effects introduces large uncertainties

→ Cocktail-indepent method needed to separate non-prompt contributions from the QGP radiation

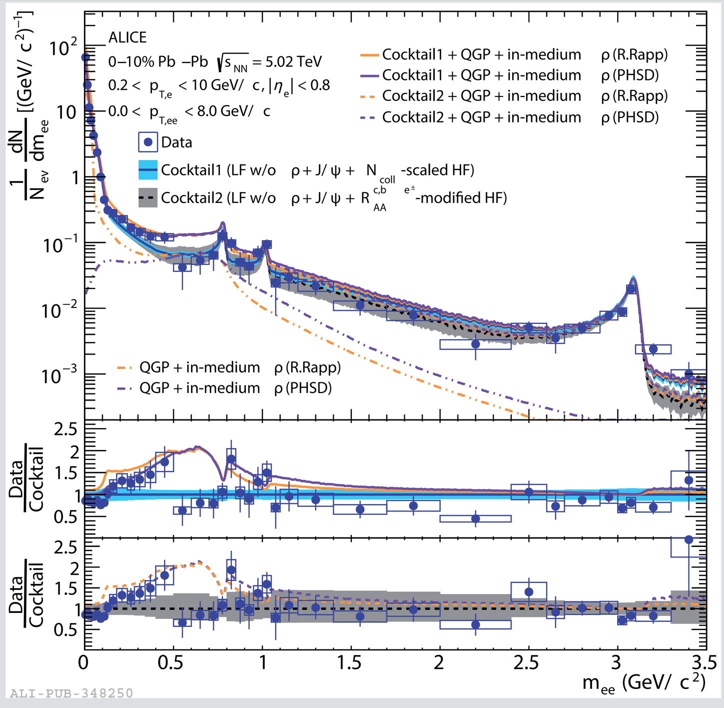
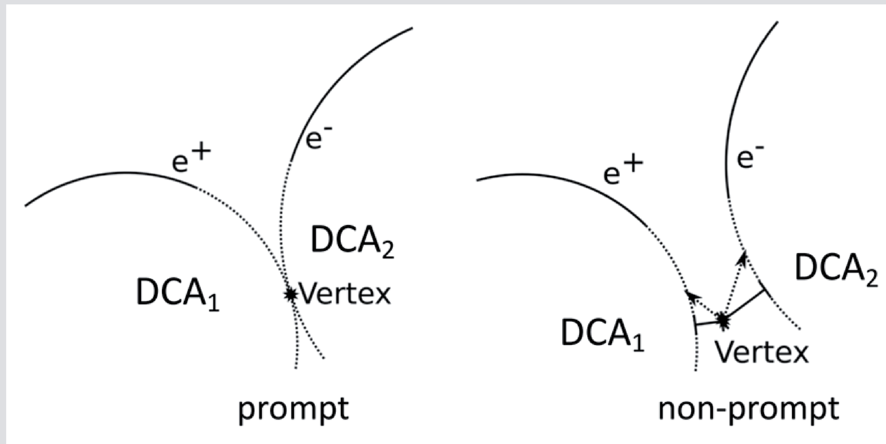


Figure 1: Dielectron production in central Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV as a function of m_{ee} compared to different expectations from hadronic decays [1]. The blue line assumes binary collision scaling for heavy-flavor production, while the grey line includes the nPDFs from EPS09 and the measured R_{AA} of c/b → e^{+-} [2]. The bottom panels show the respective cocktail ratios together with theory calculations for thermal contributions [4,5].

Classical Approach

Distance-of-closest approach (DCA) in the transverse plane:



Separation of prompt and non-prompt sources based on their distance to the primary vertex:

→ Decay length of charm and beauty hadrons much larger than prompt sources

Calculate DCA on pair level taking the resolution into account:

$$\mathrm{DCA_{ee}} = \sqrt{[(\mathrm{DCA}_1/\sigma_1)^2 + (\mathrm{DCA}_2/\sigma_2)^2]/2} \quad [3]$$

However, this definition neglects information on the sign, correlation and longitudinal information of the DCA

→ New approach: Apply machine learning (ML) to include all possible information and correlations

Setup

Input Monte Carlo simulation:
- Underlying event from Hijing simulation of Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV with a full ALICE Run 2 detector response
- Up to 10 J/ψ per event in $|\eta| < 1$ injected depending on the centrality (70% prompt & 30% non-prompt)
- Only J/ψ tracks kept after reconstruction
- Standard track and event selections applied

Neural network (NN):
- Architecture: Deep Residual NN (8 layers, 256 nodes)
- Activation function: ReLU
- Loss: binary or categorical crossentropy with class weights
- Regularization: L1 and L2, 10% Dropout
- Optimizer: Adam (Learning rate adjustment, early stopping)
- Training/Validation/Test split: 75%/15%/10%

Observables used as features in the model:

Track: DCA_{xy}, DCA_z, $\sigma(\mathrm{DCA}_{xy})$, $\sigma(\mathrm{DCA}_z)$, rel. p_T, η, φ, position in x, y and z, pointing angle θ

Pair: pseudo proper decay length L_{ee}, opening angle ω_{ee}, pointing angle θ_{ee}, χ^2_{ee}

Model Performance

Binary classification

Direct comparison of separation capabilities of different approaches using the signal (S) of e^+e^- pairs from prompt and non-prompt J/ψ decays

→ The ML-based model exhibits a significantly better performance independent of threshold

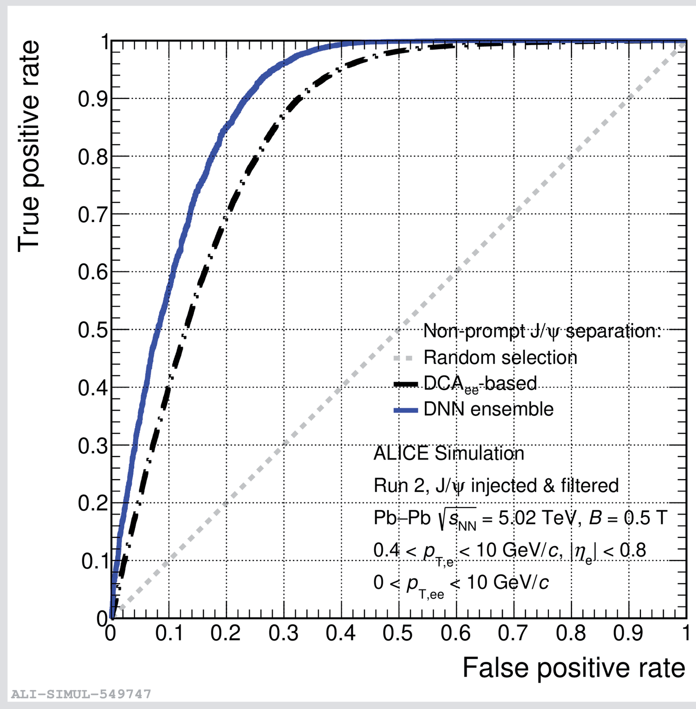


Figure 2: Receiver Operating Characteristic (ROC) curves for the classical $\mathrm{DCA_{ee}}$ analysis and a trained NN illustrating their performance in separating prompt and non-prompt J/ψ decays. The diagonal dashed line represents random guessing.

Multi-class classification

Inclusion of combinatorial background (Bkg) pairs
- Model tuned for high precision in identifying non-prompt pairs (high confidence threshold)
  → Below-threshold pairs are labeled as prompt Bkg

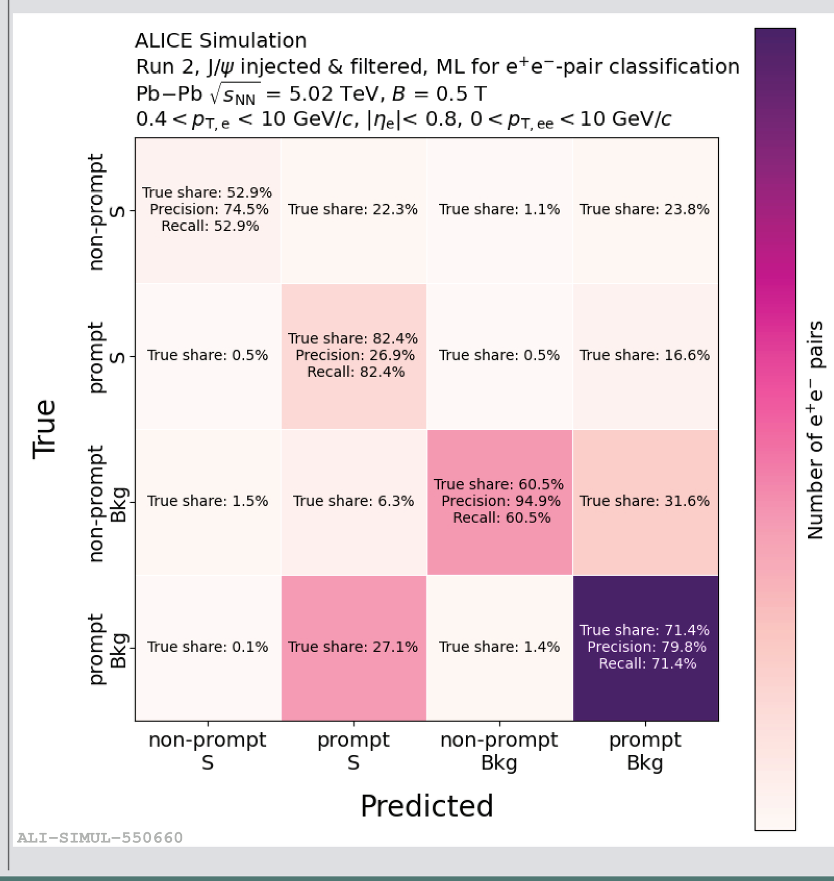


Figure 3: Confusion matrix of the multi-class model to visually represent the classification performance. Diagonal entries show the correct predictions of each class, while off-diagonal entries represent misclassifications. The classification performance can be estimated using the number of True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) by defining the precision=TP/(TP+FP) and the recall=TP/(TP+FN).

Application

Track candidate filtering:

Before the combinatorial pairing of all electron and positrons reject all electrons and positrons associated to a non-prompt pair identified by the multi-class model

→ Removes all identified non-prompt pairs (S+Bkg) as well as all pairs which share just one track associated to these electrons and positrons

→ Significantly reduces the combinatorial background by 33.6% and increases the S/Bkg by 64.4%

→ Random rejection of signal pairs due to misclassification of about 5.5%

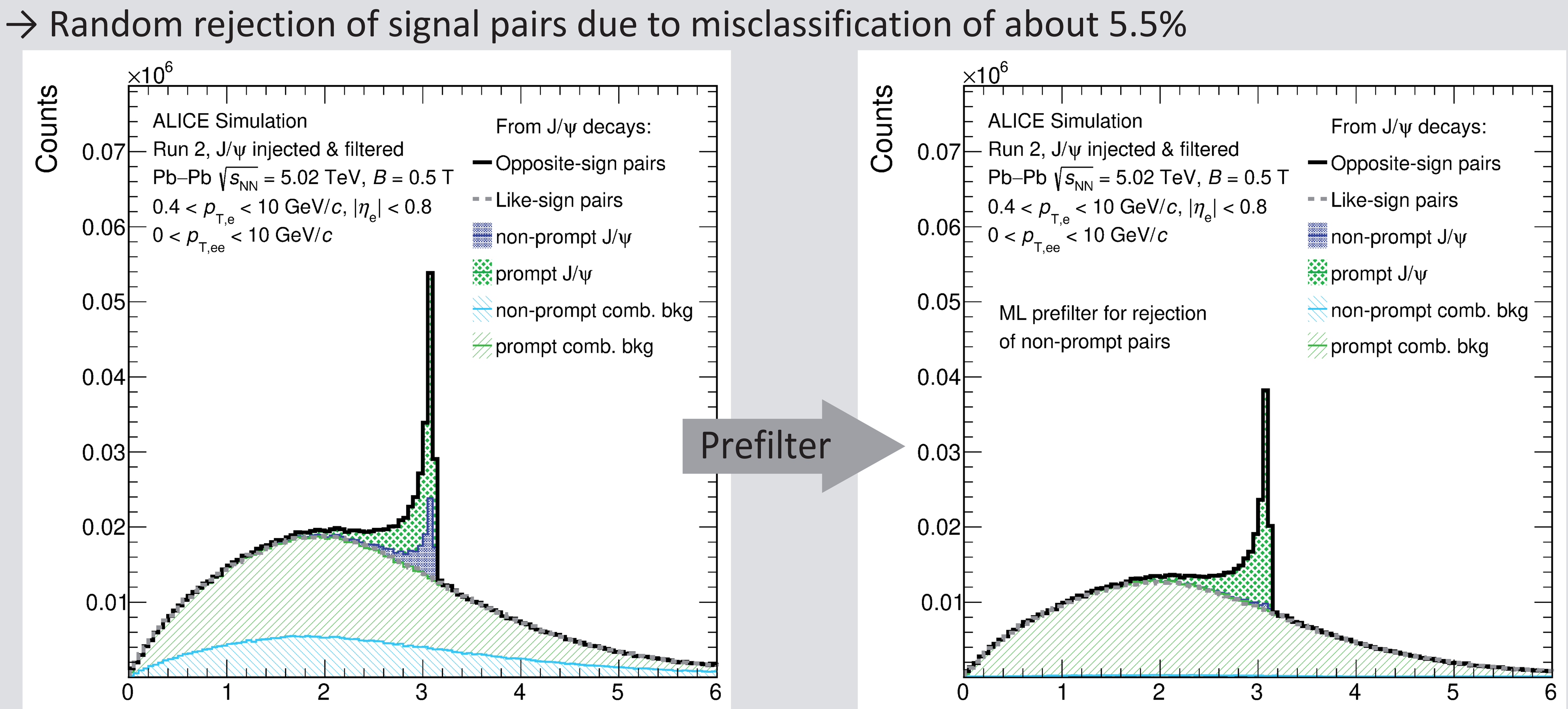


Figure 4 & 5: Simulation of the e^+e^- pair distribution from J/ψ decays in Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV as a function of m_{ee}. The solid black line shows the sum of all pairs with opposite signs and the dashed grey line indicates the sum of all pairs with the same sign. The green color illustrates the reconstructed pairs from the prompt decays while the blue color highlights the pairs originating from the non-prompt decays. The left plot shows the mass distribution without a prefiltering while the right plot shows the distrbution after the application of the ML-based prefilter.

Conclusion

ML can be applied successfully to separate prompt and non-prompt contributions

Analysis of the feature importance can be used to improve definition of classical observables

ML can be used as a powerful prefilter in the dielectron analysis to reject non-prompt contributions and reduce the combinatorial background

The upgraded ITS in Run 3 with its improved vertex pointing resolution will further improve the topological separation [6]

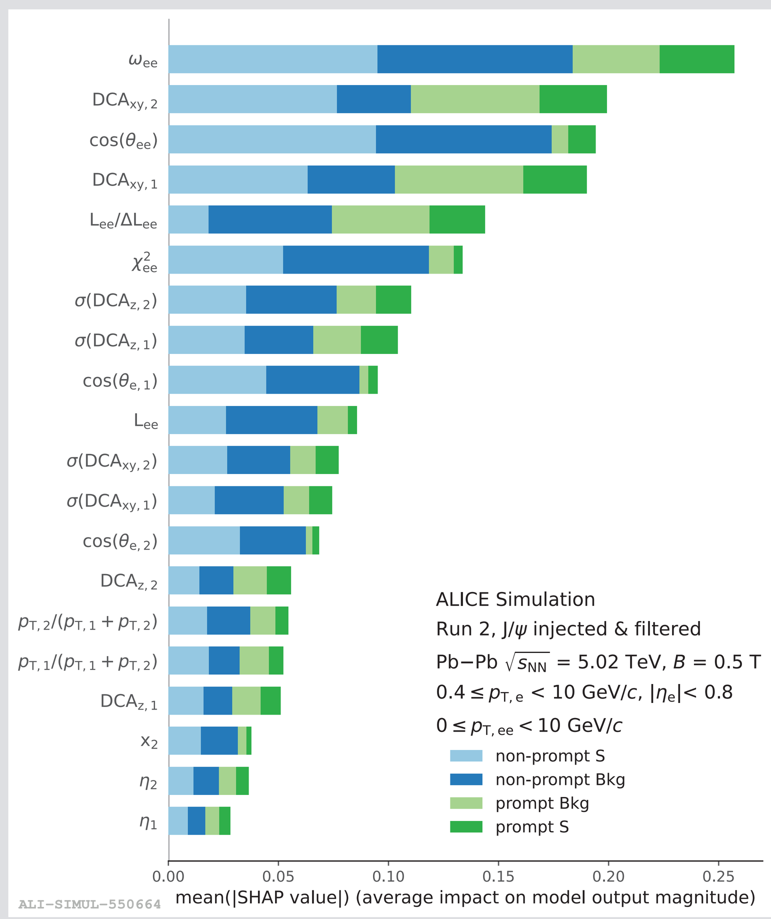


Figure 6: Feature importance of the multi-class model illustrated the horizontal bars. The most important features are ordered from top to bottom. The length of each bar illustrates the impact of this observable on the final prediction for each class highlighted by the different colors.

Next step: A more sophisticated simulation of Pb–Pb collisions including open heavy-flavor background and injected thermal radiation needed to fully test the potential of this approach

[1] ALICE, *Dielectron production in central Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV*, arXiv: 2308.16704v1
[2] ALICE, *Measurement of electrons from semileptonic heavy-flavour hadron decays at midrapidity in pp and Pb–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV*, Phys.Lett.B 804 (2020) 135377
[3] ALICE, *Dielectron production in proton-proton collisions at $\sqrt{s_{NN}}$ = 7 TeV*, JHEP 09 (2018) 064
[4] R. Rapp et. al., *Dilepton Spectroscopy of QCD Matter at Collider Energies, Adv. High Energy Phys. 2013 (2013) 148253*
[5] T. Song et. al., *Open charm and dileptons from relativistic heavy-ion collisions*, Phys. Rev. C 97 (2018) 064907
[6] ALICE, *Technical Design Report for the Upgrade of the ALICE Inner Tracking System*, CERN-LHCC-2013-024

GOETHE UNIVERSITÄT FRANKFURT AM MAIN
HGS-HIRe for FAIR — Helmholtz Graduate School for Hadron and Ion Research
FSP ALICE — Erforschung von Universum und Materie

February 2024
[1]jung@ikf.uni-frankfurt.de
Goethe-Universität Frankfurt