PhD/MSc Project in ALICE 2025 with the South African team

Title: Utilization of machine learning techniques in ALICE Run 3 data analysis for heavy-flavour hadron production at the LHC

Principal Investigators:

A/Prof. Zinhle Buthelezi (zinhle@tlabs.ac.za), iThemba LABS/WITS School of Physics

Dr. SV Förtsch (forstch@tlabs.ac.za),

Project description

ALICE (A Large Ion Collider Experiment) [1] is one of the 4 major experiments at the Large Hadron Collider (LHC). It is a dedicated heavy-ion detector designed to study the physics of the strongly interacting primordial state of matter known as the quark-gluon plasma (QGP) at the highest energy densities reached in the laboratory so far. The QGP is studied using heavy-ion collisions (A-A) at ultra-relativistic energies provided by the LHC. ALICE studies the QGP through a comprehensive study of particles (hadrons, leptons, photons, etc.) produced in these collisions. In the LHC Run 3, ALICE has already collected a substantial amount of data both in proton-proton (pp) and in lead-lead (Pb-Pb) collisions at √s=13.6 TeV and 5.36 TeV, respectively. This provides a good platform for identifying key physics measurements to further our understanding of the hot and dense QGP created in Pb-Pb collisions and to test perturbative Quantum Chromodynamics (pQCD) based models.

One such interesting key observable is the production of hadrons containing heavy quarks: charm (1.29 GeV/c^2) and beauty (4.18 GeV/c^2) produced in the collision via initial hard scattering processes. The characteristic flavour of heavy quarks is conserved throughout the evolution of the medium formed in A-A collisions. Hence, their measurement provides insight into the mechanisms of in-medium energy loss, propagation and hadronization. Therefore, heavy quarks are regarded as excellent probes of the QGP. It is therefore crucial to understand their interaction with and their evolution in the underlying medium. More experimental data of refined heavy-flavour decay are therefore needed to improve the existing theoretical models. In ALICE, heavy-quark production so far was studied by exploiting various experimental techniques, via the hadronic (kaon, pion, etc.) [2] and leptonic (electron or muon) [3] decay channels. This project is looking, for the first time, into the production of these heavy quarks and their evolution in the QGP by utilizing traditional and machine learning (ML) techniques for the Run 3 (2022 - 2026) data analysis.

In this project, a feasibility study will be performed using commonly used ML techniques and artificial Neural Networks for a comparative study in Run 3 data anchored Monte Carlo production and online-offline (O2) data [4]. The interest is to look at heavy flavour hadrons produced in proton-proton (pp) production to train the machinery and then in Pb-Pb data collected in 2023. The challenge in these analyses is the signal-to-background ratio, thus, a good quality analysis is needed to distinguish the signal from the background.

Scope:

The project scope covers the initial study using Monte Carlo data including realistic detector configuration and will then be extended to the analysis of dedicated pp/Pb-Pb data collected by ALICE in Run 3. The

tools and data samples are available on the ALICE GRID as per the requirement of the ALICE Collaboration. Analysis Tutorials are given 3 times a year by the ALICE O2 team. The candidate will be expected to follow the tutorials to get familiarized with the analysis techniques. The initial study of generator-level simulations will be followed by the development of the analysis software in the O2 framework. The analysis will first be performed on simulated data and switch to recorded collisions once the reconstruction of pp and/or Pb-Pb collected in 2023 during Run 3 is ready. The data analysis, presentation of results, and regular updates will be done in line with the ALICE Physics Working Group Heavy Flavour (PWG-HF), the O2 project.

Tools and Infrastructure:

In this project, the candidate will learn how to work in a modern high-energy physics data analysis framework, written in C++ and based on ROOT and ALICE Online-offline (O2). The work will also include Monte-Carlo simulations within this framework and the usage of the worldwide LHC ALICE computing grid.

References:

[1] "*The ALICE experiment at the CERN LHC*", The ALICE Collaboration, et al, 2008 JINST 3 S08002, 10.1088/1748-0221/3/08/S08002.

[2] Measurement of beauty and charm production in pp collisions at √s=5.02 TeV via non-prompt and prompt D mesons, ALICE Collaboration, JHEP 05 (2021) 220 https://doi.org/10.1007/JHEP05(2021)220

[3] Production of muons from heavy-flavour hadron decays in pp collisions at √s = 5.02 TeV, ALICE Collaboration, JHEP 09 (2019)008, https://doi.org/10.1007/JHEP09(2019)008.