T 70: Experimental Methods (general) 3

Time: Wednesday 16:15–18:30 Location: T-H29

T 70.1 Wed 16:15 T-H29

Calibration of the b-tagging mis-tag rate for charm jets based on W+c events at $\sqrt{s} = 13$ TeV with the ATLAS experiment — Johannes Erdmann and •Benedikt Gocke — TU Dortmund University, Department of Physics

Identifying jets containing heavy-flavour hadrons is needed for many analyses in the ATLAS experiment. This is done using flavour tagging algorithms. Since these algorithms need to be calibrated, it is important to also calibrate the mis-tag rate for non-heavy flavour jets. These calibrations are done by matching their performance in simulation to data.

The mis-tag rate calibration for c-jets using W+c events is presented, which allows for a relatively pure sample of c-jets. For this process, a prompt muon and a muon inside a jet (soft muon) are expected in the final state. Therefore, the usage of an event selection which selects events in regions defined by the electric charge of the expected muon pair in the final state is shown. Furthermore, the construction of a signal region to reduce the background contributions, which is based on the expected opposite-sign-charged lepton pair for the signal process is illustrated. The calculation of the efficiencies for the used heavy flavour tagger is shown. Finally, the resulting data-to-simulation scale factors are discussed.

T 70.2 Wed 16:30 T-H29

Convolutional Networks and Deep Learning at the Belle II Experiment — •Johannes Bilk, Sören Lange, Katharina Dort, Stephanie Käs, and Timo Schellhaas — Justus-Liebig-Universität, Gießen,Germany

The Belle II pixeldetector (PXD) has a trigger rate of up to 30 kHz for 8 M pixels. Its proximity to the interaction point allows it to detect exotic highly ionizing particles such as antideuterons, magnetic monopoles, stable tetraquarks or pions with small transverse momenta < 100 MeV. Those particles leave no tracks in the outer parts of the Belle II detector, and thus their pixel data may be deleted online as part of background suppression. In this contribution, we evaluate the performance of a machine learning algorithm to identify slow pions only on the basis of pattern recognition of pixel cluster structures. We employ convolutional neural networks with different kernel configurations and use images of 9x9 pixel matrices as input. On the long term such image recognition techniques could provide a rescue mechanism for the pixel data before they are erased. Results on accuracy and sensitivity are presented.

T 70.3 Wed 16:45 T-H29

b-Tagging studies for the ATLAS experiment — •Eleonora Loiacono — Deutsches Elektronen-Synchrotron DESY, Hamburg and Zeuthen, Germany

At the LHC, among the different Higgs boson production mechanisms, the associated production with a vector boson $VH \to \bar{b}b$ is considered as the golden channel for the measurement of the Higgs coupling to b-quarks, due to its high sensitivity. In the VH(bb) analysis, the available statistics of the simulated Monte Carlo (MC) samples is largely reduced by the b-tagging algorithms, especially for background processes, resulting in large statistical fluctuations in the MC templates. The first study that will be presented, is related to the reduction of this statistical uncertainty. In order to avoid being dominated by this uncertainty a technique based on Truth Tagging is performed, exploiting a Graph Neural Network to make an optimal use of multidimensional information associated to jets. The second study relies on improving of the connection between Flavour Tagging scale factors measured on data and those corresponding to MC to MC variations. b- tagging is very important for VH(bb) and the calibration procedures provide the use of the Scale Factors (SF). The so-called smoothing procedure is currently applied to these SF, which allows to have SF that are not dependent on the binning of the histograms used to do the calibration. In particular, there will be a focus on reviewing the currently smoothing procedure and also on including more statistical tests on procedure currently adopted for data SF.

T 70.4 Wed 17:00 T-H29

Extrapolation of flavour tagging calibrations to high transverse momenta — Arnulf Quadt, Elizaveta Shabalina, and •Sreelakshmi Sindhu — II. Physikalisches Institut, Georg-August-Universität Göttingen

Identifying jets containing heavy flavoured hadrons can be very beneficial for a variety of analyses and this can be done using flavour tagging algorithms. Currently, these algorithms are calibrated by matching their performance on data to simulation. However, for jets with transverse momentum greater than a few hundred GeV these calibrations do not exist due to statistical limitations. Flavour-tagging information for jets with high transverse momenta can be quite useful for analysis such as search for heavy resonances. Hence, a Monte Carlo based extrapolation of the data based flavour tagging calibrations is done to extend these up to 3 TeV. The extrapolation uncertainties are calculated by propagating the relevant modelling, tracking and jet uncertainties through the tagging algorithm. In this talk, the extrapolation procedure will be explained and the results from my work on the extrapolation uncertainties for charm and light jets will be presented for the b-tagging algorithms (DL1, DL1r) that are used for the analysis of the Run 2 data by the ATLAS Collaboration.

T 70.5 Wed 17:15 T-H29

Utilizing muons to tag b-jets in ATLAS — •Frederic Renner, Clara Elisabeth Leitgeb, and Cigdem Issever — DESY, Zeuthen, Germany

Several Beyond Standard Model theories predict Higgs boson pair production with much larger cross sections compared to the Standard Model. While Higgs boson pair production has not yet been observed at the LHC due to the rarity of the process, it may come within reach with more data in upcoming runs with improved statistics and improved particle reconstruction. A Higgs boson pair decays predominantly into four b-quarks, which makes the identification of jets originating from b-quarks particularly important. B-quarks have a longer lifetime than lighter quarks, leading to a second displaced vertex separated from the primary vertex which is mainly exploited in b-tagging algorithms. However, when using the results in a neural network to distinguish the flavor origin of jets, the network still misidentifies a lot of jets originating from lighter quarks as b-quark jets. About 20% of b-quarks decay semileptonically with a so-called soft muon in the final state. They are soft because they are generally less energetic than particles originating from the primary vertex. Utilizing the muon information, the rejection of fake b-jets can be improved substantially at a given selection efficiency.

T 70.6 Wed 17:30 T-H29

Signal efficiency corrections for boosted $X \to b\bar{b}$ tagger using $Z \to b\bar{b}$ events with the ATLAS experiment — •Daariimaa Battulga, Arely Cortes Gonzalez, and Cigdem Issever — Institut für Physik, Humboldt-Universität zu Berlin

It is of utmost interest to efficiently identify a new heavy resonant particle or to study boosted Higgs boson decaying into a pair of b-quarks in the ATLAS experiment. Particles produced with high transverse momentum (p_T) will have very collimated decay products in the final state. This boosted topology makes it particularly challenging to distinguish two b-jets in the calorimeter. To overcome this, these b-jet fragmentations are clustered within a large radius R=1.0 jet, and its associated track jets are b-tagged. However, even with this approach, the double b-tagging efficiency decreases at the large p_T. The boosted $X \to b\bar{b}$ tagger improves the efficiency of the b-tagging at the higher p_T region. This double b-tagger is based on the neural network that uses the kinematic distributions of the large-R jet and flavour information of variable radius track jets. In order to apply this $X \to b\bar{b}$ tagger to the physics analysis, it needs a dedicated calibration.

Hence, this talk presents the in situ signal efficiency calibrations of a new $X \to b\bar{b}$ tagger using $Z \to b\bar{b}$ events. We have derived the data-to-simulation scale factors using full Run 2 pp collision data collected by the ATLAS experiment at the center of mass energy of $\sqrt{s} = 13$ TeV with an integrated luminosity of $\mathcal{L} = 139\,\text{fb}^{-1}$. The signal efficiency corrections covering the soft (hard) p_T region are derived using $Z \to b\bar{b}$ events with a recoiling photon (jets).

T 70.7 Wed 17:45 T-H29

A machine-learning based method to improve isolation variables for photon identification with the ATLAS detector —

JOHANNES ERDMANN, OLAF NACKENHORST, and •MICHAEL WINDAU — TU Dortmund University, Department of Physics

The study of photons is crucial for finding and measuring many processes at colliders. Predominantly, prompt photons, which are created during the collisions, play an important role and have to be distinguished from hadrons decaying into photons. Different methods are used to distinguish this signal from the background. One of these is the use of isolation variables. These are based on track measurements and information from the calorimeters, where they are defined by the activity in a cone around the candidate object. They are currently built in ATLAS by discrete cuts.

In this talk, studies on improving isolation variables using deep neural networks will be presented.

T 70.8 Wed 18:00 T-H29

Data-driven corrections to shower shape variables for photon identification at the ATLAS experiment with 13 TeV *pp* collision data — •JAN LUKAS SPÄH, BJÖRN WENDLAND, and JOHANNES ERDMANN — Technische Universität Dortmund, Fakultät Physik

Measurements of Standard Model processes, searches for new particles or processes forbidden in the Standard Model with photons in the final state play an important role in the physics programme of the ATLAS experiment. At hadron colliders, studies of photons are particularly challenging, as large background contributions arise from jets that can be misidentified as photons. This requires an identification algorithm that provides high efficiency for genuine photons while ensuring an excellent background rejection for misreconstructed objects.

Currently, this method relies on rectangular cuts on so-called shower shape variables, which capture relevant information about the shape and evolution of the electromagnetic shower and the possible leakage into the hadronic calorimeter. While the longitudal shower development through the calorimeter layers is modelled well, residual mismodelling is observed for lateral shower shape distributions. Therefore, the simulated distributions are corrected with a data-driven approach.

In this talk, studies of univariate first- and second-order corrections obtained from the full Run 2 dataset are discussed and recent improvements are highlighted.

T 70.9 Wed 18:15 T-H29

Towards tuning electromagnetic shower properties to data with AtlFast3 — •JOSHUA BEIRER[1,2], MICHAEL DUEHRSSEN[1], and STAN LAI[2] — [1]CERN — [2]Georg-August-Universität Göttingen

AtlFast3 is the next generation of high precision fast simulation in ATLAS and encompasses a parametrised and a machine-learning approach based on Generative Adversarial Networks (GANs). With respect to its predecessor, AtlFast3 significantly improves in physics performance while retaining the benefit of a considerably faster simulation in comparison to Geant4.

A precise simulation of electromagnetic (EM) shower properties in the ATLAS calorimeter is crucial for the identification of particle showers originating from electrons and photons. While AtlFast3 precisely simulates the properties of EM showers, it inevitably inherits any mismodelling of the full Geant4 simulation, upon which its parametrization is based. Differences between the Geant4 simulation and data collected by the ATLAS detector are well known but insufficiently understood. Traditionally, these discrepancies are corrected using ad hoc methods such as the applications of shifts to the central values of the corresponding distributions, a procedure known as fudging.

In this talk, a brief overview of fast simulation in ATLAS is given. Furthermore, the development of different models directly embedded within the simulation framework used to tune EM shower properties directly to data are described and it is shown that AtlFast3 can be modified in a way that the shower shapes observed in data are accurately reproduced by the simulation.