Third Year PhD Research Project

Lucrezia Rambelli
Supervisors: Andrea Coccaro, Francesco Armando Di Bello

XVIII PhD Cycle

During the final year of my PhD, I focused on two main projects: one carried out within the ATLAS Collaboration, and an independent research on Machine Learning for tracking.

Work within the ATLAS Collaboration

Within ATLAS, I contributed to the Flavour Tagging group, and more specifically to the Xbb subgroup, which is responsible for the development of the new Xbb/cc boosted tagger. This tagger was introduced and developed with the aim of having an optimized selection of events in which Higgs boson decay in a couple of b or c quarks when the transverse momentum of the decaying particle is above 250 GeV (i.e. when is 'boosted'), as it can improve the sensitivity of searches for new resonances in beyond Standard Model scenarios and can help in having precise measurements of the Higgs boson pT spectrum, also allowing to improve the experimental sensitivity to the H→cc decay. As in this regime the final state objects result highly collimated, the tagging algorithms developed for smaller pT values don't show optimal performance, the new GN2X tagger was developed looking for a performance improvement both on signal classification and background rejection. The GN2X tagger operates by processing track-level information within a large-R jets with R = 1.0, classifying then the identified large-R jets into categories such as H→bb, H→cc, top quark decays, or generic QCD backgrounds, assigning to each classified jet a score for each category. The model output scores for each category are then combined into a discriminant, which, resulting in different final shapes for different topologies, allows to define on it cut values which correspond to known signal efficiency values (called 'working points'). Since the model is trained on Monte Carlo (MC) samples and since these samples are also employed in the working point definition, it requires calibration through scale factors to ensure consistent performance between data and simulation. In this context, my primary task concerned the calibration of the ATLAS GN2X tagger. Contribution to GN2X Calibration

Calibrations of the GN2X tagger relies on $\rightarrow bb$+jets events, which is the main channel for calibrating bb events, as they share similar kinematic properties with Higgs decays but offer significantly higher statistics.

They consist in the computation of the so-called 'Scale Factors' (SF) values, which are defined as the ratio of the tagging efficiency measured in data to that obtained on Monte Carlo samples, and which can be equivalently expressed as the ratio of the signal strengths measured after and before applying the GN2X tagger selection.

$$SF = \epsilon_{data}/\epsilon_{MC} = \mu_{post-tag}/\mu_{pre-tag} \tag{1}$$

Since before applying the GN2X selection the overwhelming background contamination makes it unfeasible to directly measure the Zbb contribution, the pre-tag signal strength is extracted from fits to Zll+jets distribution. The leptonic preselection in Zll events provides a clean and unbiased measurement, independent of the GN2X performance. The post-tag signal strengths are then measured on Zbb+jets events by fitting data, for different transverse momentum and working point configurations, with a signal+background template and computing the signal strength for each case. In the fit procedure, while the signal template can be easily extracted from the $Z \rightarrow bb$ post-tag mass distribution for each working point and transverse momentum value, the background one has to be studied more in depth. Sensitivity studies show that in this environment the main background contribution is the one from events with QCD dijets which are mis-identified as bb or cc large-R jets, and, after a first reweighting of these MC samples for a better overall alignment of all the MC contributions to the data, I worked on the background modelling procedure for the final SF computation. My main contribution was the spurious signal test, designed to quantify biases from background modelling in the fit. Since the invariant mass of background events is described with analytic functions (polynomials or exponential-polynomials), the test evaluates the bias each choice introduces. It is performed by fitting background-only samples (MC dijets after GN2X tagging) with a signal-plus-background model, where the signal template is taken from the $\rightarrow bb$ mass distribution at a given efficiency. The resulting spurious signal strength is the fitted-over-expected signal yield, which should be zero on background-only data. For each p_T bin and working point, the function minimising this bias is chosen as the background model for calibration. In addition, I worked on the signal injection test, in which a known distribution consisting of signal plus background MC (with the signal taken from $\rightarrow bb$ and the background from the appropriate MC component) is fitted using a signal-plus-background template. The signal template is always derived from $\rightarrow bb$ MC samples, while the background component corresponds to the candidate function chosen by the spurious signal test. The difference between the expected injected signal strength ($\mu = 1$) and the fitted one is interpreted as a systematic uncertainty, which is then included in the final uncertainty assigned to the scale factor values. Another contribution was the fit-range study, aimed at assessing the stability of the results with respect to the chosen fit window. In this case, the dependence of the fitted signal strength on the invariant mass range used in the fit was investigated. The standard mass window of 50–150 GeV was varied to include several alternative ranges, and the largest deviation observed between the nominal signal strength and those obtained with the alternative windows was taken as an additional systematic uncertainty. After accounting for the systematic contributions from the various studies, I derived the final scale factors. The resulting calibrated scale factors are crucial for aligning the GN2X tagger's response between data and simulation. They enable the rescaling of Monte Carlo event weights to more accurately reproduce the observed data, thereby ensuring the reliability of physics measurements carried out within ATLAS. Contribution to the VHbb/cc Analysis Group: Together with my work on the calibration effort, I also contributed to the VHbb/cc analysis group, which develops and optimizes the methodologies and framework for studying events where a Higgs boson is produced in association with a vector boson (W or Z) and subsequently decays into a pair of b or c quarks. Although the production cross-section of the VH process is smaller than that of gluon–gluon fusion and vector-boson fusion, this channel plays an important role in the ATLAS physics program. It provides direct sensitivity to the couplings of the Higgs boson to both vector bosons and quarks, while at the same time offering relatively clean experimental signatures. This is due to the fact that the final states of the main processes include leptons, which can be efficiently identified and which significantly reduce the contamination from QCD multijet backgrounds. The analysis is divided into three leptonic channels, defined by the vector boson decay: 0-lepton ($Z \to \nu\nu$), 1-lepton ($W \to l\nu$), and 2-lepton ($Z \to ll$). Focusing on these channels suppresses QCD backgrounds and enhances signal sensitivity. The search is carried out in both the resolved and the boosted topologies. In the resolved regime, the two jets originating from the Higgs boson decay can be reconstructed separately, whereas in the boosted regime the decay products are too collimated to resolve individually, and a single large-R jet must be used to represent the entire Higgs decay which will be classified using the GN2X boosted tagger . Building on my previous work with the GN2X calibration, my contribution to the VHbb/cc analysis focused on the boosted topology, which is defined in this context for Higgs bosons produced with transverse momentum above 400 GeV. In this context, I worked on incorporating the GN2X information into the analysis framework, and in particular into the Multi-Variate Analysis (MVA) algorithm used to discriminate between signal and background events. The MVA relies on boosted decision trees (BDTs), which classify events by applying a sequence of binary splits on the input variables, progressively dividing the phase space into regions enriched in either signal or background. While a single decision tree is often too simple, the boosting procedure combines many trees, with each new tree focusing on correcting the mistakes of the previous ones. The final classifier is therefore much more powerful, as it exploits correlations between multiple input features to achieve strong discrimination. My contribution focused on exploring different strategies to include the GN2X output within the MVA. Specifically, I tested the use of the GN2X score, assigned to each large-R jet based on the track information within it, as an additional input feature to the BDT optimisation procedure. For each configuration, I studied the resulting sensitivity improvement, comparing the performance with respect to two reference scenarios: the case where only the mass information is used for computing the sensitivity, and the case where the MVA output score distribution is used. At present, I am working on extending this effort by incorporating the scale factor information obtained from the GN2X calibration into the analysis framework for each working point and transverse momentum bin, ensuring that the MVA fully accounts for the calibrated tagger response.

ARIDAJE: A differentiable Machine Learning model for tracking

In addition to my work within the ATLAS Collaboration, during the final year of my PhD I completed an independent project focused on the application of machine learning and differentiable programming techniques to track reconstruction. Track reconstruction is a fundamental task in particle-physics experiments, as it provides the essential observables needed for physics analyses. Traditional algorithms, such as the Hough Transform used in the ATLAS muon spectrometer, are effective but computationally demanding and rely on a factorized approach, where pattern recognition and track fitting are treated as separate steps. With the increase in luminosity and data complexity at the LHC, developing faster and more flexible methods has become an important research direction.

In this context, I designed and implemented a novel Machine Learning model named ARIDAJE, which represents an end-to-end, differentiable approach to muon track reconstruction. Unlike traditional methods, ARIDAJE, thanks to the definition of differentiable blocks which allows to have a physics-informed weight backpropagation, directly maps raw detector hits to the particle's transverse momentum within a single integrated pipeline. The model is based on a Graph Attention Network that interprets detector hits as nodes of a graph and exploits their spatial correlations. Clustering and fitting operations are embedded as differentiable modules, allowing the entire process, from hit classification to momentum estimation, to be jointly optimized through backpropagation. This design makes the reconstruction inherently physics-informed and capable of learning the track curvature directly from the data. The model was developed in a JAX-based environment, which provides automatic differentiation, just-in-time compilation, and vectorization, ensuring both flexibility and computational efficiency. Applied to a simplified ATLAS muon-spectrometer geometry, the model has demonstrated promising results in both hit classification and transverse momentum estimation, outperforming a sequential baseline approach. These achievements show the potential of differentiable, end-to-end methods to become a scalable alternative to traditional tracking algorithms. This project constitutes a proof of concept for physics-informed track reconstruction with machine learning and opens the way for further developments applicable to different detectors and experimental setups. Based on the results obtained, I am currently preparing a publication to present this work to the community.