Lucio Derin

lucio.derin@cern.ch lucio.derin@ge.infn.it

Dottorato in Fisica e Nanoscienze - Ciclo XL - First Year Report

Tutors

Carlo Schiavi: Associate Professor at *University of Genova* - Physics Department
Fabrizio Parodi: Full Professor at *University of Genova* - Physics Department
Francesco Armando Di Bello: Reseacher at *University of Pisa* - Physics Department

Research activity

My research activity takes place within the Genoa group of the ATLAS experiment at the Large Hadron Collider (CERN). The ATLAS experiment is one of the biggest Particle Physics experiments and its efforts focus on the understanding of the fundamental building blocks of our Universe, studying high-energy proton-proton collisions. My work focuses on two areas: the development of a unified trigger for b quarks and τ leptons, and the use of Differentiable Programming as a tool for injecting physical priors in Machine Learning models. The following sections summarize my contributions to these two areas.

Unified b and τ trigger

Beauty quarks and tau leptons are very different in their nature, as described by the Standard Model (SM) of Particle Physics: the former is a second generation quark, while the latter is a third generation lepton. Nevertheless, their experimental signatures in the detectors can share some similarities: they can both give rise to *jets* with displaced vertices. Indeed, b-quarks hadronize in B mesons, which have a lifetime long enough to fly a few millimeters before decaying and creating a displaced vertex. The same can be said for τ leptons: they can fly measurable distances before decaying into hadrons (mainly pions). Nevertheless, their flight lengths differ as the tau lepton tends to decay sooner: the $c\tau$ of the B meson[1] is $\sim 0.45mm$ while that of the tau lepton is $\sim 0.09mm$.

This similarity is not the only thing that binds b and τ particles: they play a central role in the detection of Higgs bosons pairs, which is a challenging hunt set to take place in the High Luminosity LHC (HL-LHC) campaign and a powerful way to test the SM predictions. A channel that plays an important role in this measure is the $HH \to bb\tau\tau$ one, as it is a good trade-off between probability of the final state (i.e. the Branching Ratio -BR) and a clean signature in the detector. This channel gives rise to the necessity of triggering beauty quarks and taus with a single machinery, which would allow to compute shared systematics and orthogonal decisions. This is an alternative to the present state-of-the-art approach, which merges the decisions of two separate triggers and requires a complex overlap removal scheme to avoid double counting of τ and b candidates.

My contribution in this year

To help push for the development of such a trigger, I joined the Tau Trigger Group in the scope of my ATLAS Authorship Qualification Project[2]. I contributed significantly to the first deployment of a new τ identification algorithm, which is based on a Transformer architecture and has replaced the previous DeepSet-based trigger in this year RUN 3 data acquisition. Then, I worked on overcoming some of its limitations and on harmonizing the training and testing pipeline with the present b-jet trigger, which is Transformer-based as well. The last point required extensive cross-checks of the equivalence of Tau group and b-jet group software tools, in order to allow for the possibility of a common codebase.

Finally, I started to implement and validate a trigger that is able to simultaneously identify the presence of a τ or b-jet. I am currently involved in making sure that this trigger will be deployed online, as a test chain in the ATLAS experiment for the 2026 data-taking period.

[1] Here τ indicates the lifetime of the particle.
[2] i.e. a one-year-long qualification task, ending in October, that will allow me to sign ATLAS papers.

Differentiable Programming for Physics-informed Machine Learning

Machine Learning (ML) models are widely deployed in many Physics areas; in High Energy Physics (HEP), Flavour Tagging (FTag) and trigger algorithms deeply rely on ML. Oftentimes, there is a need for an algorithm that can simultaneously solve different, though related, problems: for instance, predicting the flavour of a jet while simultaneously fitting the secondary vertex position. In order to merge these problems, a joint minimization of various complicated cost functions is needed, and can be done exploiting *Differentiable Programming* (DP) techniques.

Moreover, although ML algorithms have showcased unmatchable performances, they tend to be "black-box" models: this means that it is difficult to understand which inputs led to the model's response -a problem that is known as *explainability*. In Physics applications, it is important to understand the inner mechanisms of the model, and to be sure that it learned the underlying laws of Nature. One way to do so is to inject physical priors (i.e. constraints that are known a priori) within the training pipeline of the ML model, enforcing its consistency with the theoretical foundations of the problem.

One of the techniques that enables physics-informed ML is to penalize the models with functionals that encode the given physical constraints and then, exploiting Automatic Differentiation (AD) and DP, enable the model to learn from them.

My contribution in this year

In my work I am applying this programming paradigm to two relevant problems: jet tagging with explicit vertex reconstruction and end-to-end model calibration (i.e. estimating its performance on real data after training on synthetic ones).

The first application, a spin-off of my Master Thesis, is carried out in collaboration with F.A. Di Bello, C. Schiavi and F. Parodi, focusing on the reconstruction of the secondary vertex of the D^* decay. The goal of this work is to see if the physical prior (enforced by a differentiable χ^2 vertex fit during the model's training) can help a ML model to better recognize c-jets and improve the mass peak resolution of the D meson.

The second application, carried on in a broader group with N.M. Heartman and L. Heinrich (TUM) among others, focuses on the possibility of training a *self-calibrating* FTag algorithm: in this case, as an application of *Simulation-Based Inference*, a differentiable constrain is added in the training of the model to enforce its performance to be similar on synthetic and real data. This could enable the training of an algorithm that is already calibrated on real data, possibly reducing the calibration systematics that play an important role in the final uncertainties of many analyses.

PHD COURSES

Machine Learning for Particle Physics (course from the Master's degree in Physics): *A. Coccaro, F.A. Di Bello, R. Torre*
QCD and Collider Physics: *S. Marzani*
Advanced Statistics for Data Analysis:*F. Badaracco, F.A. Di Bello, F. Parodi*
High Performance Computing for heterogeneous accelerator architectures: *D. D'Agostino (DIBRIS)*

PUBLICATIONS

The ATLAS collaboration, **Transforming jet flavour tagging at ATLAS**, CERN cds , I am part of the Analysis team, though I am not yet eligible to sign the paper as my qualification task is still ongoing.

CONFERENCES AND WORKSHOPS

ATLAS Collaboration Week, CERN, 17 - 21 February 2025.
ATLAS TDAQ Week, CERN, 18 - 19 March 2025. **Talk** on behalf of the Tau Trigger Group: *GNTau: GNN-based Tau Trigger. Current State and Future Prospects.* Slides here.
ATLAS Di-Higgs Workshop, CERN, 10 - 12 June 2025.
XVIII ATLAS Italia Workshop, Udine, 01 - 03 July 2025. **Talk**: *Transformer-Based Tau Identification at the High-Level Trigger*[3]
ATLAS Hadronic Identification Workshop, CERN, 8 - 10 September 2025. **Contribution**: my developments of the Tau Trigger's algorithm *GNTau* were presented by the Tau Group speaker in the overview and future prospects talk.

OTHER ACTIVITIES

Tutor for CERN International Masterclasses: supervised high school students participating in the CERN Masterclasses program. Facilitated data analysis exercises with actual ATLAS and LHCb datasets and clarified essential principles of high-energy physics.

[3] I could not give the talk due to an injury; my work was included in the talk given by F.A. Di Bello.

Tutor – Stage for High School Students: led high school students through introductory activities in particle physics as part of an outreach program. Taught core concepts and oversaw practical exercises using Jupyter notebooks connected to the ATLAS and LHCb experiments at CERN.
Tutor for General Physics in Bioengineering: guided exercises and reviewed theoretical concepts in Mechanics, Electromagnetism, and Thermodynamics for students in the General Physics course.