

## Modèle de rapport du CSI

### Fiche signalétique du doctorant ou de la doctorante

<b>NOM et Prénom</b>	LEONARDO Splendori
<b>Titre de la thèse</b>	Probing the Higgs boson pair production through Vector Boson Fusion at the LHC in the ATLAS experiment
<b>Date de 1<sup>ère</sup> inscription en doctorat</b>	01/10/2023
<b>Année d'inscription demandée</b>	2
<b>Condition financière de préparation de la thèse</b>	Thèse avec financement dédié : <input checked="" type="checkbox"/> Thèse préparée en parallèle d'une activité professionnelle : <input type="checkbox"/> Thèse en « autofinancement » : <input type="checkbox"/> Autre : <input type="checkbox"/> Su autre, préciser :
<b>Pour les thèses avec financement dédié, nature et période de financement</b>	Thèse AMU, 3 ans
<b>Thèse préparée à temps</b>	Complet : <input checked="" type="checkbox"/> Partiel : <input type="checkbox"/> Si temps partiel, % de temps consacré à la thèse :
<b>Unité de recherche d'accueil</b>	CPPM
<b>NOM-Prénom-email du directeur ou de la directrice de thèse</b>	ARNAUD Duperrin arnaud.duperrin@cppm.in2p3.fr
<b>NOM-Prénom-email du codirecteur ou de la co-directrice de thèse, des co-encadrants ou des co-encadrantes</b>	THOMAS Strebler thomas.strebler@cppm.in2p3.fr
<b>Précisions particulières (FTLV<sup>1</sup>, situation de handicap, sportif de haut niveau, ...)</b>	

<sup>1</sup> Formation tout au long de la vie.

**Date de la réunion du CSI :**

**Composition du comité de suivi**

	<b>Prénom, Nom, Titre</b>	<b>Adresse email</b>	<b>Unité de recherche, affiliation<sup>2</sup></b>
<b>Membre spécialiste du domaine de la thèse</b>	Camacho Toro Reina CR	reina.camacho@lphne.in2p3.fr	LPNHE
<b>Membre non spécialiste externe au domaine de la thèse</b>	Mancinelli Giampiero DR	giampi@cppm.in2p3.fr	CPPM
<b>Le cas échéant, Autre membre<sup>3</sup></b>			

<sup>2</sup> Si un membre du CSI est externe à AMU, merci d'indiquer le nom de l'établissement auquel il ou elle est rattaché.e.

<sup>3</sup> Merci d'ajouter autant de ligne que de membres supplémentaires.

## Rapport d'activités rédigé par le doctorant

Présenter succinctement (2 à 3 pages maximum) les travaux et les activités réalisés dans l'année, l'état d'avancement de votre recherche, les conditions de réalisations de votre thèse et les éventuelles difficultés que vous rencontrez.

During this first year the research work has been divided in two parts, the Authorship Qualification Project (AQP) and work on the di-Higgs analysis in the bbtatau channel. The former was the main focus of the year.

The Authorship Qualification Project consisted of the training and study of a graph based flavour tagging neural network for use in High Luminosity LHC. The particular neural network this study is based on is the GN2 architecture, the tagger to be used for ATLAS Run 3 analyses. This study involved the preparation of a training dataset from simulated ATLAS data, the training of the network itself, analysing the behaviour and performance of the model and producing documentation to ease future work involving these programs.

The training of a complex model required the pre-processing of large amounts of simulated data and significant computing resources for the training itself. Some technical difficulties were encountered during the latter step as special hardware (specifically GPUs) is required to complete the training in a reasonable amount of time (~5 days per training run). This is an intrinsic limitation given by the computational intensity of the training. The main issue being incompatibility of a proprietary and closed source software required to utilize the GPUs available at CPPM. This problem caused a delay in the project as it made it virtually impossible to conduct training runs, effectively halting all new trainings for approximately one month. After working in close contact with a software engineer the issue has been fixed thanks to an update to the CPPM GPU clusters.

Currently the AQP is entering its conclusive stage. With completion expected for October 2024. The behaviour of the flavour tagging software was studied extensively with different metrics and modification to the training data and the model itself to assert the impact of different conditions we expect to encounter during the High Luminosity LHC phase. These studies involved analysing the change in performance of the network when modifying the composition of the training data, the variables available as an input to the network, and kinematic cuts applied on the simulated data before training.

The performance of the model and its dependence on the jet kinematics and detector geometry was studied extensively. This was done by producing numerous performance plots showing the jet tagging efficiency and background rejection as functions of a jet's transverse momentum and pseudorapidity. Similar studies were performed to evaluate the performance of the network in the identification of track origins, one of the auxiliary tasks performed by the GN2 model.

In the coming months this work will be documented further, with the potential publishing of a PUB note covering it.

Given the increased pile-up expected during High Luminosity LHC particular care was also given to analysing and understanding the pile-up robustness of the model.

These results were presented to the ATLAS flavour tagging working group in multiple meetings. These meetings provided feedback which proved useful in improving the understanding of this model as well as providing direction in which to take its development. Furthermore, a website has been created to document the full preprocessing and training pipeline.

A significant part of the AQP, related to the hyperparameter optimization of the model, had to be modified due to the unavailability of the computing resources necessary to perform such a task. The problem has been reported to the ATLAS flavour tagging group and has yet to be resolved. Given the limited time remaining for the completion of the AQP the ATLAS flavour tagging group has agreed to modify the direction of this last part, aiming instead at completing other objectives originally considered to be secondary and optional.

Work on the diHiggs to bbtatau analysis has started with contributions to the necessary software frameworks. These contributions included solving issues encountered by the analysis teams during the development of the necessary frameworks, the implementation of observables and Boosted Decision Trees used in previous analyses into the current frameworks, and the analysis of a software bug affecting the muon trigger efficiency in simulated samples, including the implementation of a temporary workaround to sanitize the affected samples.

Concurrently, a number of workshops and seminars have been attended, including : 1st DIVE (Di-Higgs In VBF Events, a group of French ATLAS teams collaborating on vector boson fusion diHiggs analyses) workshop at LPNHE, the ATLAS Software tutorial, the ATLAS HH workshop, the ATLAS HH→bbtau hackathon, the Journées HH ATLAS France, where the work on flavour tagging was presented, and the IN2P3 School of Statistics 2024.

The Doctoral courses attended so far are : Awareness to Open Science (3h), French as a foreign language (30h), the ITER seminar & visit (2h), Research Integrity in scientific professions (10h) and the IN2P3 School of Statistics (21h).

The plan for the next year of PhD studies involves shifting the full focus on the HH→bbtau analysis and related frameworks, due to the completion of the AQP. With the main focus of the analysis work being the study of the VBF categories and the integration of the boosted H→tau tagger in the analysis, which is itself scheduled to be published during the second year, in summer 2025.

The completion of the mandatory training hours needed for the full PhD course (50 hours of scientific training + 50 hours of professional training) is also planned for the second year.

## Compte rendu et avis du CSI

### Compte rendu synthétique des échanges

**État d'avancement des travaux de recherche** (appropriation du sujet, principaux résultats, respect du calendrier prévisionnel, qualité de la présentation, réponse aux questions etc.)

--

<p>Avis :</p> <p>Points de vigilance :</p> <p>Recommandations :</p>
---

### Production scientifique, valorisation, expérience internationale

<input type="checkbox"/> article(s), <input type="checkbox"/> communication(s) lors de congrès, conférence(s) , colloque(s) <input type="checkbox"/> brevet(s), transfert <input type="checkbox"/> autre(s) production(s) scientifique(s)	<input type="checkbox"/> dont co-publication(s) internationale(s) <input type="checkbox"/> dont communication(s) internationale(s) <input type="checkbox"/> mobilité internationale courte (< à 3 mois) <input type="checkbox"/> mobilité internationale longue (> à 3 mois)
--	---

<p>Avis :</p> <p>Points de vigilance :</p> <p>Recommandations :</p>
---

**Conditions de la formation doctorale** (Intégration dans l'unité de recherche, conditions d'encadrement, , développement des compétences et de la culture scientifique, préparation du devenir professionnel, le cas échéant, **aménagements de parcours de formation doctorale en FTLV** etc.)

Avis :

Points de vigilance :

Recommandations :

### Difficultés et/ou dysfonctionnements

<p><b>Le doctorant rencontre-t-il des difficultés ?</b></p>	<p><input type="checkbox"/> oui : <span style="float: right;"><input type="checkbox"/> non</span> Si oui, lesquelles ?</p>
<p><b>Le CSI alerte l'école doctorale et/ou fait un signalement</b></p>	<p><input type="checkbox"/> oui <input type="checkbox"/> non</p>

**En cas de dernière réinscription en vue de la soutenance de thèse**

**Indiquer l'échéancier de fin de thèse fixé en accord avec le doctorant et le directeur de thèse :**

**Avis du CSI en vue de la réinscription**

**Avis du CSI sur la réinscription en doctorat**

*Favorable* :

*Réservé* :

*Défavorable* :

**Date, Noms, Prénoms et signatures des membres du CSI**