



**HAL**  
open science

# Measurement of the differential cross-section of the $t\bar{t}H$ process and its CP properties in the multilepton final state using deep learning with the ATLAS experiment

Vera Maiboroda

## ► To cite this version:

Vera Maiboroda. Measurement of the differential cross-section of the  $t\bar{t}H$  process and its CP properties in the multilepton final state using deep learning with the ATLAS experiment. High Energy Physics - Experiment [hep-ex]. Université Paris-Saclay, 2024. English. NNT : 2024UPASP062 . tel-04744298

**HAL Id: tel-04744298**

**<https://theses.hal.science/tel-04744298v1>**

Submitted on 18 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Measurement of the differential  
cross-section of the  $t\bar{t}H$  process and its  
 $\mathcal{CP}$  properties in the multilepton final  
state using deep learning with the  
ATLAS experiment

*Mesure de la section efficace différentielle du processus  $t\bar{t}H$  et de  
ses propriétés  $\mathcal{CP}$  dans l'état final multileptons en utilisant  
l'apprentissage profond avec l'expérience ATLAS*

**Thèse de doctorat de l'université Paris-Saclay**

École doctorale n°576 particules hadrons énergie et noyau :  
instrumentation, imagerie, cosmos et simulation (PHENIICS)

Spécialité de doctorat: Physique des particules

Graduate School : Physique. Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche **Département de Physique des  
Particules IRFU (Université Paris-Saclay, CEA)**, sous la direction de  
**Frédéric DELIOT**, directeur de recherche

**Soutenue à Paris-Saclay, le 19 septembre 2024**

**Vera MAIBORODA**

**Composition du jury**

Membres du jury avec voix délibérative

**Jeremy ANDREA**

Directeur de recherche, IPHC

**Lucia DI CIACCIO**

Professeur des Universités, LAPP

**Yann COADOU**

Chargé de recherche, CPPM

**Jean-Roch VLIMANT**

Chargé de recherche,

CalTech

Rapporteur & Examineur, Président de jury

Rapporteur & Examinatrice

Examineur

Examineur

**Titre:** Mesure de la section efficace différentielle du processus  $t\bar{t}H$  et de ses propriétés  $\mathcal{CP}$  dans l'état final multileptons en utilisant l'apprentissage profond avec l'expérience ATLAS.

**Mots clés:** couplage de Yukawa, l'apprentissage profond, boson de Higgs, ITk.

**Résumé:** Cette thèse mesure la section efficace de production du processus  $t\bar{t}H$  et les propriétés  $\mathcal{CP}$  du couplage de Yukawa du quark top dans le cadre des Sections Efficaces Simplifiées (STXS). Cette étude est réalisée dans l'état final multilepton (utilisant deux leptons de même signe et trois leptons) avec l'ensemble des données enregistrées par le détecteur ATLAS au LHC, correspondant à  $140\text{fb}^{-1}$ . La mise en œuvre des mesures STXS nécessite la reconstruction de l'impulsion transverse du boson de Higgs, ce qui est difficile en raison de la présence de plusieurs neutrinos dans les états finaux. Une approche basée sur un réseau de neurones graphiques (GNN) a été explorée pour cette reconstruction. Comme l'impulsion transverse du boson de Higgs est également sensible aux propriétés  $\mathcal{CP}$  du couplage de

Yukawa du quark top, une telle reconstruction aide également à étudier la violation  $\mathcal{CP}$  dans le processus  $t\bar{t}H$ . La section efficace STXS du processus  $t\bar{t}H$  est mesurée avec une incertitude attendue de  $-0.85$  à  $+1.15$ . La limite attendue sur l'angle de mélange  $\mathcal{CP}$ , qui détermine l'ampleur de la violation  $\mathcal{CP}$  dans le couplage de Yukawa du quark top, est :  $|\alpha/\pi| < 0.31$  à 95% CL. Pour la phase haute luminosité du LHC, le trajectographe d'ATLAS sera remplacé par un détecteur amélioré (ITk). La thèse présente également le développement d'un outil pour l'inspection visuelle automatique des modules de pixels ITk. Cet outil, basé sur un algorithme de détection d'anomalies par apprentissage automatique, est conçu pour aider les opérations pendant la production.

**Title:** Measurement of the differential cross-section of the  $t\bar{t}H$  process and its  $\mathcal{CP}$  properties in the multilepton final state using deep learning with the ATLAS experiment.

**Keywords:** Yukawa coupling, deep learning, Higgs boson, ITk.

**Abstract:** This thesis measures the  $t\bar{t}H$  production cross-section and the  $\mathcal{CP}$  properties of the top Yukawa coupling within the Simplified Template Cross-Sections (STXS) framework. This study is performed in the multilepton final state (using two leptons with same sign and three leptons) with the full  $140\text{fb}^{-1}$  dataset recorded by the ATLAS detector at the LHC. Implementing the STXS framework requires reconstructing the Higgs transverse momentum, which is challenging due to the presence of several neutrinos in the final states. A Graph Neural Network approach was explored for this reconstruction. As the Higgs transverse momentum is also sensitive to the  $\mathcal{CP}$  properties

of the top Yukawa coupling, such reconstruction helps to study  $\mathcal{CP}$  violation in the  $t\bar{t}H$  process. The  $t\bar{t}H$  STXS cross-section is measured with an expected uncertainty ranging from:  $-0.85$  to  $1.15$ . The expected limit on the  $\mathcal{CP}$ -mixing angle that drives the amount  $\mathcal{CP}$  violation in the top Yukawa coupling is:  $|\alpha/\pi| < 0.31$  à 95% CL. For the high-luminosity phase of the LHC, the ATLAS tracker will be replaced with an upgraded detector (ITk). The thesis also presents the development of a tool for automatic visual inspection of the ITk pixel modules. This tool is based on a machine learning anomaly detection algorithm, and is designed to help the operations during production.

# Acknowledgments

The work presented in this manuscript is the result of a long journey that would not be possible without support from many people, to whom I am deeply grateful.

First, I would like to thank my supervisor Frédéric for allowing me to pursue my PhD under his guidance and for his invaluable support. I also would like to express my gratitude to Matthias for his guidance during my qualification task. Frédéric's and Matthias' supervision and advice have been instrumental in my progress. I also appreciate the support from my senior colleagues, Laurent, Julien, and Emilien, who patiently answered all my questions, helping me expand my knowledge. Thanks to Jérôme and Julia for being open to new ideas, sharing their expertise and providing helpful insights. I am happy to have worked with Benjamin and Tanguy, brilliant students, showing great dedication and providing their valuable assistance. I am grateful to the many members of the ATLAS group and beyond who offered their technical expertise and constructive feedback.

I would like to thank the reviewers of this document, Jeremy Andrea, Lucia Di Ciaccio, Yann Coadou and Jean-Roch Vlimant for taking the time to read my thesis and for their useful comments and discussions about the broader implications of my work.

I also want to acknowledge my PhD and postdoc colleagues-friends for bringing joy to my daily lab experience. Special thanks to Tristan, Tanguy, Marianna, Alberto and Peter. I am very glad to share my PhD journey with you, inside and outside the lab.

I am thankful to my family who managed to provide their unwavering support throughout this journey despite the long distance between us.

In the loving memory of Valera.

# Contents

<b>1</b>	<b>Theoretical introduction and Higgs boson searches at the LHC</b>	<b>15</b>
1.1	Standard Model general overview . . . . .	15
1.1.1	Elementary particles . . . . .	15
1.1.2	Lagrangian and symmetries . . . . .	16
1.1.3	Chirality . . . . .	17
1.2	Gauge principle in SM interactions . . . . .	18
1.2.1	Concept of the Gauge principle . . . . .	18
1.2.2	Gauge principle in QED and QCD . . . . .	19
1.2.3	Gauge principle in electroweak interaction . . . . .	21
1.3	Spontaneous symmetry breaking and the Brout-Englert-Higgs mechanism . . . . .	22
1.3.1	The Brout-Englert-Higgs mechanism . . . . .	22
1.3.2	Fermion masses via the BEH mechanism . . . . .	25
1.3.3	Higgs boson couplings . . . . .	25
1.4	Higgs boson measurements at the Large Hadron Collider . . . . .	26
1.4.1	Higgs production modes . . . . .	27
1.4.2	Higgs decay modes . . . . .	28
1.5	Higgs-top Yukawa coupling . . . . .	29
1.5.1	Top-quark physics . . . . .	29
1.5.2	The top Yukawa coupling . . . . .	30
1.6	$\mathcal{CP}$ -violation in the SM . . . . .	32
1.7	Probing the $\mathcal{CP}$ nature of the top-Higgs coupling . . . . .	33
1.8	Simplified Template Cross-section framework . . . . .	36
1.9	$t\bar{t}H$ measurement at the LHC . . . . .	36
<b>2</b>	<b>Experimental setup</b>	<b>38</b>
2.1	The Large Hadron Collider . . . . .	38
2.1.1	LHC setup . . . . .	39
2.1.1.1	Basic collider parameters . . . . .	39
2.1.1.2	The accelerator chain . . . . .	40
2.1.1.3	LHC experiments . . . . .	42
2.2	The ATLAS detector . . . . .	43
2.2.1	ATLAS coordinate system and detector requirements . . . . .	44
2.2.2	The magnet system . . . . .	44
2.2.3	The inner detector . . . . .	45
2.2.3.1	Pixel detector . . . . .	45
2.2.3.2	Semiconductor Tracker . . . . .	46
2.2.3.3	Transition Radiation Tracker . . . . .	46

2.2.4	Calorimeters . . . . .	47
2.2.4.1	Electromagnetic calorimeter . . . . .	47
2.2.4.2	Hadronic calorimeter . . . . .	48
2.2.4.3	Forward Calorimeter . . . . .	48
2.2.5	Muon systems . . . . .	49
2.2.6	Trigger and Data Acquisition system and data quality . . . . .	49
2.3	Event simulation . . . . .	51
2.3.1	Cross-section computation in proton-proton collisions . . . . .	52
2.3.2	Event generation . . . . .	53
2.3.3	Detector simulation . . . . .	54
2.3.4	Digitization . . . . .	55
2.4	Object reconstruction and identification in ATLAS . . . . .	55
2.4.1	Tracks and vertices . . . . .	55
2.4.2	Electrons and photons . . . . .	56
2.4.3	Muons . . . . .	57
2.4.4	Jets . . . . .	59
2.4.4.1	Reconstruction . . . . .	59
2.4.4.2	Calibration . . . . .	60
2.4.4.3	<i>b</i> -tagging . . . . .	61
2.4.5	Missing Transverse Energy . . . . .	62
2.4.6	Hadronically decaying taus . . . . .	63
<b>3</b>	<b>Automatic visual inspection of ITk pixel modules</b>	<b>64</b>
3.1	All-silicon Inner Tracker and Pixel modules . . . . .	64
3.1.1	Inner Detector upgrade . . . . .	64
3.1.2	ITk Pixel modules and its production chain . . . . .	65
3.2	Choosing the automatic visual inspection algorithm . . . . .	67
3.2.1	Photo-taking procedure . . . . .	68
3.2.2	Building datasets . . . . .	70
3.2.2.1	MVTec dataset structure . . . . .	70
3.2.2.2	Annotations . . . . .	71
3.2.2.3	Preprocessing . . . . .	72
3.2.2.4	Photos collections and datasets . . . . .	73
3.2.3	Anomaly detection unsupervised models . . . . .	74
3.2.3.1	PatchCore . . . . .	75
3.2.3.2	CFA . . . . .	76
3.2.3.3	FastFlow . . . . .	77
3.2.3.4	CFLOW-AD . . . . .	77
3.2.3.5	DRAEM . . . . .	77
3.2.4	Metrics . . . . .	78
3.2.5	Comparison of unsupervised models performance . . . . .	80
3.3	Performance and limitations studies . . . . .	81

3.3.1	Augmentation and preprocessing studies on selected unsupervised models .	81
3.3.1.1	Tile size . . . . .	81
3.3.1.2	Impact of filtering for complexity reduction . . . . .	82
3.3.1.3	Augmentations: basic spatial . . . . .	83
3.3.1.4	Colour preprocessing . . . . .	83
3.3.1.5	Augmentations: custom spatial . . . . .	85
3.3.1.6	Dataset complexity: MVTEC tile subset . . . . .	85
3.3.1.7	Dataset complexity: denoising . . . . .	86
3.3.1.8	Augmentations: noise generation . . . . .	87
3.3.2	Adaptable solution for the visual inspection . . . . .	89
3.4	Graphical User Interface . . . . .	90
<b>4</b>	<b><math>t\bar{t}H</math> cross-section measurement and <math>\mathcal{CP}</math>-violation studies in the top Yukawa coupling</b>	<b>93</b>
4.1	Object identification and event selection . . . . .	94
4.1.1	Event preselection . . . . .	94
4.1.2	Object definitions . . . . .	95
4.1.2.1	Loose leptons . . . . .	95
4.1.2.2	Jets . . . . .	96
4.1.2.3	Removing ambiguities between overlapping reconstructed objects	96
4.1.3	Region selection . . . . .	97
4.2	Signal and background modelling . . . . .	97
4.2.1	$t\bar{t}H$ signal modelling . . . . .	98
4.2.2	$t\bar{t}H$ and $tH$ modelling for $\mathcal{CP}$ -studies . . . . .	99
4.2.3	$t\bar{t}W$ background modelling . . . . .	99
4.2.4	Other backgrounds modelling . . . . .	100
4.3	Regions definition . . . . .	101
4.3.1	Cut-based control regions . . . . .	101
4.3.2	Signal regions and MVA-based control regions . . . . .	102
4.3.2.1	Pre-MVA selection . . . . .	103
4.3.2.2	$2lSS$ channel MVA regions . . . . .	105
4.3.2.3	$3l$ channel MVA regions . . . . .	107
4.4	Systematic uncertainties . . . . .	108
4.4.1	Instrumental uncertainties . . . . .	109
4.4.2	Signal modelling uncertainties . . . . .	111
4.4.3	Background uncertainties . . . . .	111
4.5	Statistical analysis techniques . . . . .	112
4.5.1	Binned profile likelihood fit . . . . .	113
4.5.2	Test statistics . . . . .	113
4.5.3	Uncertainties treatment . . . . .	115
4.5.4	Asimov and hybrid fit setups . . . . .	115
4.6	STXS setup . . . . .	116

---

4.6.1	Event kinematics in the $2lSS$ and $3l$ channels . . . . .	116
4.6.1.1	$2lSS$ channel . . . . .	116
4.6.1.2	$3l$ channel . . . . .	117
4.6.2	Higgs $p_T$ reconstruction . . . . .	117
4.6.3	GNN to reconstruct the STXS bins . . . . .	118
4.6.4	STXS signal region definition . . . . .	119
4.6.5	Checks on the GNN performance . . . . .	121
4.6.6	Higgs $p_T$ sensitivity to the $\mathcal{CP}$ -mixing angle . . . . .	123
4.7	Study of the $\mathcal{CP}$ -properties of the top Yukawa coupling . . . . .	124
4.7.1	$t\bar{t}H$ parametrization . . . . .	125
4.7.2	$tH$ parametrization . . . . .	125
4.8	Results . . . . .	126
4.8.1	Statistical-only Asimov fit results . . . . .	128
4.8.2	$\mathcal{CP}$ results with systematic uncertainties . . . . .	128
4.8.2.1	Statistical-only and systematic comparison in an Asimov fit . . . . .	129
4.8.2.2	Hybrid fit results . . . . .	132
4.8.3	Differential cross-section measurement within STXS framework . . . . .	133
4.8.4	Conclusion . . . . .	133

## French summary

Le modèle standard de la physique des particules est une théorie bien établie qui décrit les particules fondamentales et leurs interactions. Ces particules sont divisées en quarks et leptons, qui interagissent par le biais de trois forces fondamentales : la force forte, la force faible et la force électromagnétique. Le modèle standard repose sur le concept de symétrie de jauge et est représenté par un Lagrangien reflétant ces symétries et les interactions entre particules. La théorie prédit également l'existence de bosons de jauge — des porteurs de force comme le photon pour les interactions électromagnétiques, les bosons  $W$  et  $Z$  pour les interactions faibles, et le gluon pour les interactions fortes. De plus, le boson de Higgs confère une masse aux autres particules par le biais du mécanisme de Higgs. Cependant, le modèle standard présente plusieurs limitations. Il ne prend pas en compte la gravité, ni n'explique la matière noire et l'énergie noire. Il existe aussi des lacunes dans la compréhension de l'asymétrie matière-antimatière de l'univers. Le Big Bang aurait dû créer des quantités égales de matière et d'antimatière, pourtant nous observons un univers dominé par la matière. Une explication possible réside dans la violation de la symétrie  $\mathcal{CP}$ , où certains processus traitent différemment les particules et les antiparticules. Bien que le modèle standard prévoit une certaine violation de  $\mathcal{CP}$  dans les interactions faibles et que cette violation ait été observée, elle est insuffisante pour expliquer la différence que nous observons aujourd'hui, suggérant une physique au-delà du modèle standard. Il existe différentes approches pour rechercher cette nouvelle physique. Les recherches directes consistent à chercher de nouvelles particules. Une autre approche est de faire des mesures précises des processus connus du modèle standard. Toute déviation par rapport aux résultats prévus peut donner des indices sur une physique au-delà du modèle standard.

Le grand collisionneur de hadrons (LHC) est un collisionneur proton-proton circulaire de 27 km fonctionnant avec des énergies dans le centre de masse à l'échelle du TeV. L'une des réalisations les plus significatives du LHC a été la découverte du boson de Higgs en 2012. Aujourd'hui, le LHC a plusieurs objectifs importants: étudier en détail les propriétés du Higgs, y compris comment il interagit avec les autres particules; mesurer avec une grande précision les paramètres du modèle standard, tels que les masses et les couplages des particules connues, pour valider le modèle standard; rechercher de signes de nouvelle physique. Pour atteindre ces objectifs, le LHC s'appuie sur plusieurs expériences dont les expériences ATLAS et CMS, qui sont tous deux des détecteurs à usage général, conçus pour mesurer la plus large gamme possible de signaux. Les périodes de prise de données au LHC sont suivies de longues interruptions, durant lesquelles le matériel est mis à niveau. Jusqu'à présent, deux périodes complètes de prise de données, Run-1 et Run-2, ont été effectuées, alors que le Run-3 est actuellement en cours. Les mesures présentées dans cette thèse sont basées sur l'ensemble de données du Run-2 collectées de 2015 à 2018 à une énergie de centre de masse de 13 TeV par l'expérience ATLAS.

## Outil d'inspection visuelle automatique des modules pixel d'ITk pour la mise à niveau du LHC à haute luminosité

Après le Run-3, un long arrêt débutera, durant laquelle le LHC subira sa mise à niveau pour fonctionner à haute luminosité, augmentant considérablement la luminosité. Le détecteur ATLAS sera également mis à niveau pour répondre à ces nouvelles exigences. Une partie clé de cette mise à niveau est le remplacement de l'actuel détecteur interne par un nouveau trajectographe tout en silicium (ITk), conçu pour supporter l'augmentation de la luminosité et du flux de données. ITk se compose de deux composants principaux : le détecteur à bandes, couvrant les couches externes, et le détecteur à pixels, plus proche du point d'interaction. Le détecteur à pixels est composé de modules de pixels, qui sont des unités individuelles détectant les impacts de particules avec une grande précision. ITk offrira une couverture beaucoup plus large en pseudorapidité que le détecteur interne actuel, allant jusqu'à  $\eta = 4$ . Pour répondre aux exigences du LHC à haute luminosité, environ 10000 modules de pixels devront être assemblés pour former le détecteur à pixels. Le bloc de construction de base du détecteur à pixels ITk est un module. Un module se compose d'un capteur en silicium connecté à une puce frontale de lecture. La partie nue du module est collée à un circuit imprimé flexible (PCB). La connexion entre le PCB et le module se fait à l'aide de fils fins soudés.

L'assemblage du module de pixels ITk implique plusieurs étapes, chacune suivie de procédures de contrôle qualité rigoureuses, telles que des tests électriques et des inspections visuelles. Le travail de cette thèse se concentre sur l'étape de soudure des fils. Environ 700 fils pour connecter les pads du PCB avec la puce frontale de lecture sont placés de chaque côté d'un module, qui mesure seulement 4 par 4 cm. Cependant, des défauts sur la surface du pad, comme une contamination chimique, des rayures ou de la poussière, peuvent empêcher de réussir ces soudures. Dans ce cas, l'équipe de soudure a besoin d'informations détaillées sur l'emplacement des défauts pour ajuster la procédure en conséquence. Actuellement, la surface des pads est inspectée manuellement sous microscope. L'objectif du projet présenté dans la thèse est d'automatiser le processus d'inspection visuelle.

Pour résoudre ce problème, on a choisi d'utiliser des algorithmes d'intelligence artificielle par apprentissage automatique. Les techniques d'apprentissage automatique sont largement utilisées pour le contrôle qualité dans l'industrie. En considérant les défauts présents sur les pads comme des anomalies, on peut utiliser des algorithmes de segmentation d'anomalies visant à localiser les défauts. Les algorithmes de segmentation d'anomalies non supervisés sont entraînés sur des images de pads ne contenant aucun défaut. Cela permet au modèle d'apprendre une distribution des caractéristiques normales. Lors de la phase d'inférence, les caractéristiques en dehors de la distribution apprise sont considérées comme celles des zones anormales contenant des défauts. Pour évaluer un modèle, le masque de défaut de chaque image test est comparé au masque prédit, où les pixels blancs indiquent des défauts.

Plusieurs étapes ont été nécessaires pour constituer un ensemble de données. Tout d'abord, les photos des pads des modules ont été prises avec un microscope, en choisissant les meilleurs paramètres d'éclairage pour rendre les défauts les plus visibles. Ensuite, les images collectées des pads ont subi une procédure de prétraitement pour la correction des couleurs et le retrait

de l'arrière-plan. À des fins d'évaluation, pour chaque image de pad, le masque correspondant a été créé manuellement, mettant en évidence les zones présentant des défauts. Les images et leurs masques ont ensuite été découpés en tuiles plus petites. Les tuiles séparées en sous-ensembles d'entraînement, de validation et de test forment un ensemble de données. Plusieurs ensembles de données correspondant à différentes versions de modules de pixels, à différentes tailles de tuiles, ont été construits. Cinq modèles de segmentation d'anomalies non supervisés pré-sélectionnés ont été entraînés sur ces ensembles de données, ce qui a abouti à la sélection de deux modèles les plus performants pour des études détaillées. J'ai essayé d'améliorer la performance des deux modèles sélectionnés sur nos ensembles de données en augmentant le nombre d'images d'entraînement par retournement, découpage, amélioration du contraste, génération d'images de pads synthétiques, etc. En combinant les meilleures approches, la configuration finale a été fixée, fournissant l'ensemble des augmentations et des étapes de traitement pour garantir la meilleure performance des modèles.

Pour rendre l'outil d'inspection visuelle utilisable et accessible sur différents sites de production, deux fonctionnalités supplémentaires ont été fournies. Tout d'abord, supposant que différents sites de production ont un matériel photo différent, nous ne pouvons pas nous attendre à ce que notre modèle fonctionne avec la performance attendue sur des données de différents sites. La solution proposée est la suivante : avec un ensemble spécifique d'augmentations, il suffit de disposer de seulement 10 tuiles contenant des photos de pads pour réentraîner un modèle et atteindre une performance comparable à la solution finale. Cela facilite l'adaptation de l'outil d'inspection visuelle automatique. Deuxièmement, pour rendre les algorithmes accessibles aux équipes de soudure sur les différents sites de production, nous avons développé une interface utilisateur graphique (GUI) avec une interface intuitive pour interagir avec l'algorithme. Ce GUI élimine le besoin pour les utilisateurs d'interagir avec le code source de l'algorithme, fournissant un wrapper pour l'algorithme et toutes les connexions pertinentes aux bases de données pour le stockage des données.

## Mesure de la section efficace différentielle du processus $t\bar{t}H$ et de ses propriétés $\mathcal{CP}$ dans l'état final multilepton

Cette thèse recherche aussi des effets non-standard dans le couplage de Yukawa Higgs-top. La production de bosons de Higgs en association avec des quarks top-antitop offre une méthode directe pour mesurer ce couplage. Le processus  $t\bar{t}H$  représente environ 1% du taux total de production de Higgs boson. En raison des nombreux modes de désintégration du boson de Higgs, et dans une certaine mesure, des bosons  $W$  dans les désintégrations des paires de quarks top-antitop, il existe de nombreux états finaux  $t\bar{t}H$  qui peuvent être ciblés. Les analyses sont largement regroupées en fonction des différents modes de désintégration du boson de Higgs ciblés. Les études présentées dans cette thèse se concentrent sur les états finaux où le boson de Higgs se désintègre en plusieurs leptons chargés,  $H \rightarrow WW/ZZ/\tau\tau \rightarrow$  multileptons, offrant un état final propre avec des leptons et un fond irréductible modéré. Avec une section efficace de production faible, le processus  $t\bar{t}H$  a été observé pour la première fois lors de Run-2 par les collaborations ATLAS et CMS.

Les mesures sont effectuées en analysant plusieurs états finaux qui sont catégorisés en un total

de six canaux selon leur nombre de leptons tau se désintégrant hadroniquement, et de leptons légers (électrons et muons). Les régions de signal (SR) sont définies comme des régions enrichies en processus d'intérêt. Les événements dans les SR issus de processus différents de celui ciblé sont appelés événements de fond. Les principales sources de fond ( $t\bar{t}W$ ,  $t\bar{t}Z$  et  $VV$ ) ainsi que les fonds difficiles à modéliser (contenant des leptons non-prompt et fakes) reçoivent des corrections basées sur les données. La normalisation de ces bruits de fond est ajustée simultanément avec l'intensité du signal en effectuant un ajustement de vraisemblance sur les données. L'ajustement couvre à la fois les régions de signal et de contrôle, où les régions de contrôle sont des régions dédiées enrichies en bruit de fond spécifique. Les bruits de fond ayant une cinématique similaire au signal ont des régions de contrôle dédiées, définies à l'aide d'un classificateur d'arbres de décision boostés (BDT). D'autres régions de contrôle sont basées sur des coupures.

Mes contributions personnelles au sein de ces analyses se sont concentrées sur le développement de l'algorithme de reconstruction de l'impulsion transverse ( $p_T$ ) du boson de Higgs et sur l'étude des propriétés  $\mathcal{CP}$  du couplage de Yukawa du quark top dans le cadre du modèle de *Simplified Template Cross Section* (STXS).

Le modèle STXS a été introduit pour améliorer les mesures de précision de la force du signal de production du boson de Higgs. Les intervalles STXS sont définis comme des sections efficaces physiques dans des régions de phase d'espace mutuellement exclusives, appelées bins STXS. Dans le contexte des mesures de production  $t\bar{t}H$ , le cadre STXS est utilisé pour déterminer la section efficace en fonction de l'impulsion transverse du boson de Higgs. Les bins sont définis comme:  $p_0 \in [0, 60]$  GeV,  $p_1 \in [60, 120]$  GeV,  $p_2 \in [120, 200]$  GeV,  $p_3 \in [200, 300]$  GeV,  $p_4 \in [300, 450]$  GeV, and  $p_5 \in [450, +\infty]$  GeV.

La mise en œuvre du modèle STXS implique la reconstruction du  $p_T$  du boson de Higgs. La présence de plusieurs leptons dans l'état final implique également celle de neutrinos, produisant une énergie transverse significativement manquante. Cela rend la reconstruction du  $p_T$  du boson de Higgs complexe. Une approche utilisant des réseaux de neurones graphiques a été explorée pour cette reconstruction. La mesure de la section efficace différentielle a été réalisée pour la force du signal dans les différents bins STXS. La force du signal mesurée est :  $\mu_{p_T^H \in [0, 120]} \text{ GeV} = 1.05_{-0.85}^{+0.95}$ ,  $\mu_{p_T^H \in [120, 200]} \text{ GeV} = 0.85_{-0.85}^{+1.15}$  and  $\mu_{p_T^H \in [200, +\infty]} \text{ GeV} = 1.31_{-0.78}^{+0.69}$ . Certains des intervalles ont été considérés ensemble pour obtenir une mesure plus précise.

Étant donné que le  $p_T$  du boson de Higgs est sensible à la violation de  $\mathcal{CP}$ , les bins STXS correspondants sont également intéressants pour explorer la structure  $\mathcal{CP}$  du couplage de Yukawa du quark top. La modification potentielle du couplage du boson de Higgs avec le quark top peut être paramétrée en introduisant dans le lagrangien les paramètres  $\kappa'_t$  et  $\alpha$  dans le terme de couplage. La valeur du paramètre  $\kappa'_t$  correspond à l'amplitude du couplage de Yukawa du quark top et est égale à 1 dans le modèle standard. La modification  $\mathcal{CP}$  pourrait être paramétrée par une phase complexe  $\alpha$ , avec  $\alpha = 0^\circ$  correspondant au modèle standard,  $\alpha = 90^\circ$  à un couplage pur  $\mathcal{CP}$ -impair et  $\alpha = 45^\circ$  au scénario de mélange maximal. La section efficace du processus  $t\bar{t}H$  varie avec  $\alpha$ . La section efficace du processus  $tH$  varie également avec  $\alpha$ , mais avec une dépendance différente. La section efficace du processus  $tWH$  est également affectée par la variation de  $\alpha$ . Ainsi, lors de l'étude de la modification  $\mathcal{CP}$  du couplage de Yukawa du quark top, il est important de considérer les variations des processus  $t\bar{t}H$ ,  $tH$  et  $tWH$  ensemble. Les intervalles de confiance

---

à 95% résultants sont pour  $\alpha/\pi$  : [0.31, +0.31] et pour  $\kappa'_t$  : [0.94, 1.47] (avec une valeur centrale de 1.14). Nous avons observé que l'utilisation du cadre STXS améliore la sensibilité de la recherche de violation de  $\mathcal{CP}$ .

# Introduction

The Standard Model (SM) of particle physics is the theoretical framework that encapsulates our current understanding of elementary particles and the fundamental forces that govern their interactions. Over the past decades, the predictions of the SM have been rigorously tested and validated with increasing precision. Knowing the the SM has its limitations, we also investigate potential Beyond Standard Model (BSM) effects. This thesis explores the potential non-SM effect of a non-zero  $\mathcal{CP}$ -mixing angle in the Higgs-top Yukawa coupling. The production of Higgs bosons in association with top-antitop quarks offers a direct method to measure this coupling. Chapter 1 introduces the SM, with a particular focus on the Higgs mechanism and Higgs couplings, the search for the Higgs boson at the Large Hadron Collider (LHC), and the  $t\bar{t}H$  process.

The LHC and the ATLAS detector, the largest general-purpose particle detector at the LHC, are designed to test the SM and explore physics beyond it. The experimental setup is described in detail in Chapter 2.

The  $t\bar{t}H$  process can be studied in various final states. The analysis presented in this thesis focuses on the multilepton final state, characterized by the presence of electrons and muons. In particular, we require the final state to contain either two same-charge or three light leptons (electrons or muons) and no hadronically decaying  $\tau$ -leptons. This thesis measures the  $t\bar{t}H$  production differential cross-section and the  $\mathcal{CP}$  properties of the top Yukawa coupling using the Simplified Template Cross-Sections (STXS) framework with the complete  $140 \text{ fb}^{-1}$  dataset recorded by ATLAS during Run-2 at a center of mass energy of  $\sqrt{s} = 13 \text{ TeV}$ . Implementing the STXS framework involves reconstructing the Higgs transverse momentum, which is challenging due to the presence of several neutrinos in the final states. A Graph Neural Network approach was explored for this reconstruction. Since the Higgs transverse momentum is also sensitive to the  $\mathcal{CP}$  properties of the top Yukawa coupling, the Higgs  $p_T$  reconstruction aids in studying  $\mathcal{CP}$  violation in the  $t\bar{t}H$  process. The analysis setup and expected results of the differential cross-section and the mixing angle measurements are presented in Chapter 4.

For the High-Luminosity phase of the LHC, the ATLAS inner tracker will be upgraded with a new ITk detector. Thousands of ITk pixel modules need to be assembled at various production sites. The assembly process consists of multiple steps, each requiring quality control checks. This thesis also discusses the development of a tool for automatic visual inspection of the ITk pixel modules before the wire-bonding assembly step. Designed to help quality control during production, this tool uses a machine learning anomaly detection algorithm to localize defects on the ITk pixel module pads. The automatic visual inspection tool and its performance are described in Chapter 3.

# 1 Theoretical introduction and Higgs boson searches at the LHC

In particle physics, the Standard Model (SM) serves as a foundational framework, offering an instrument to understand the behaviour of elementary particles and their interactions. This chapter gives an overview of the Standard Model in order to provide an understanding of the Higgs-top Yukawa coupling. From the experimental point of view, the current status of Higgs-related measurements, particularly related to the  $t\bar{t}H$  process, is discussed here. Apart from the cross-section measurement, there is a specific interest in measuring the  $\mathcal{CP}$  properties of the Higgs-top Yukawa coupling. The theoretical and experimental aspects of this subject are also discussed in this chapter.

## 1.1 Standard Model general overview

### 1.1.1 Elementary particles

The SM summarizes our present knowledge of fundamental particles and their interactions, describing them via the notion of fields using quantum field theory (QFT). Within the QFT framework, particles are represented as excited states (quanta) of quantum fields. The SM includes fields with spins 0,  $\frac{1}{2}$  and 1.

Matter consists of fermions (spin- $\frac{1}{2}$  particles) that follow the Fermi-Dirac statistics. Fermions (quarks and leptons) are divided into three generations. Each generation includes two quarks and two leptons: one charged lepton and one neutrino (see Tab. 1.1.1).

Generation	Quarks			Leptons		
	Flavour	Mass	Charge (e)	Flavour	Mass	Charge (e)
1	$u$	$2.16^{+0.49}_{-0.26}$ MeV	$+\frac{2}{3}$	$e$	$0.511 \pm 1.5$ MeV	$-1$
	$d$	$4.67^{+0.48}_{-0.17}$ MeV	$-\frac{1}{3}$	$\nu_e$	$< 2$ eV	$0$
2	$c$	$1.27^{+0.02}_{-0.02}$ GeV	$+\frac{2}{3}$	$\mu$	$105.7 \pm 2.3$ MeV	$-1$
	$s$	$93.4^{+8.6}_{-3.4}$ MeV	$-\frac{1}{3}$	$\nu_\mu$	$< 0.19$ MeV	$0$
3	$t$	$172.89^{+0.30}_{-0.30}$ GeV	$+\frac{2}{3}$	$\tau$	$1776.8 \pm 0.12$ MeV	$-1$
	$b$	$4.18^{+0.03}_{-0.02}$ GeV	$-\frac{1}{3}$	$\nu_\tau$	$< 18.2$ MeV	$0$

Table 1.1.1: Summary of the properties of the quarks and leptons in the SM [1].

The particle interactions happen via three SM-described forces: the electromagnetic, the weak and the strong interactions. The SM interactions are mediated by gauge bosons (spin-1 particles) listed in Tab. 1.1.2. Gravity, the fourth known force, is not included in the SM. One of the reasons is that the mediating particle, the graviton, needs to be of spin-2 and thus does not belong to

the quantum fields easily encoded in the SM. Moreover, there is no experimental evidence for the existence of the graviton.

Bosons	Interaction	Mass	Charge ( $e$ )	Spin
$g$	strong	0	0	0
$\gamma$	electromagnetic	0	0	0
$W^\pm$	weak	$80.38 \pm 0.0012$ GeV	$\pm 1$	0
$Z$	weak	$91.19 \pm 0.0021$ GeV	0	0
$H$		$125.25 \pm 0.17$ GeV	0	1

Table 1.1.2: Summary of the properties of the bosons, gauge bosons and Higgs boson, in the SM [1].

Each matter elementary particle has a corresponding antiparticle with the same mass but opposite charge. Quarks carry electric and colour charges (and appear in 3 different colours) and thus participate in all three SM-described interactions. Leptons participate in the weak interactions, but not in the strong one since they do not carry colour charge. Unlike neutrinos, electrons, muons and  $\tau$ -leptons carry electric charges, and, thus, can interact also electromagnetically.

### 1.1.2 Lagrangian and symmetries

In quantum field theory, particles are considered fluctuations within fields typically denoted as  $\psi(x)$ , where  $x$  corresponds to the position four-vector. Under such representation, particles can be studied using the formalism of Lagrangians. The Lagrangian density  $\mathcal{L}$ , a function of the fields  $\psi$  and their derivatives, encodes essential information about fields' properties and interactions. The kinetic terms correspond to the fields' motion, coupling terms to fermions-bosons interactions, etc.

$$\mathcal{L}(x) = \mathcal{L}(\psi, \partial_\mu \psi). \quad (1.1.1)$$

The concept of symmetry is important in the SM formulation: symmetry represents a group of transformations applied to the fields that maintains the Lagrangian invariant. According to Noether's theorem [2], the Lagrangian density's invariance under continuous transformation implies the existence of a conserved quantity. Conversely, this implies that each conserved quantity corresponds to an underlying symmetry. The SM quantum fields preserve local gauge symmetry, Poincaré symmetries, and some accidental symmetries. The local gauge symmetry makes the SM an invariant theory under transformations of the direct product of (special) unitary symmetry groups:

$$SU_C(3) \otimes SU_L(2) \otimes U_Y(1). \quad (1.1.2)$$

This group structure describes strong, weak and electromagnetic interactions, respectively, via the exchange of the corresponding gauge fields: eight massless gluons and one massless photon, respectively, for the strong and electromagnetic interactions, and three massive bosons,  $W^\pm$  and  $Z$ , for the weak interaction. The number of gauge bosons corresponding to each interaction field is given by that the number of symmetry group generators. On top of this structure, the gauge symmetry  $SU_L(2) \otimes U_Y(1)$  is broken by Spontaneous Symmetry Breaking (SSB) from the

electroweak group to the electromagnetic and weak subgroups. The SSB mechanism generates the masses of the weak gauge bosons, as well as fermion masses and mixing.

### 1.1.3 Chirality

The notion of chirality is crucial to formulate the electroweak part of the SM. Particle chirality is determined by whether the particle transforms in a right- or left-handed representation of the Poincaré group. The helicity of a particle is defined as the projection of the direction of its spin on the direction of motion. For massless particles, chirality coincides with helicity. For massive particles, a chirality +1 state only coincides with a +1 helicity state up to terms suppressed by powers of  $m/E$ .

Within Weyl representation, fermionic fields  $\psi$  can be described through their left-handed  $\psi_L$  and right-handed  $\psi_R$  spinors:

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \quad (1.1.3)$$

with chiralities  $-1$  and  $+1$  correspondingly. To get the left- and right-handed components of the field, one can use the helicity projectors  $P_{L,R}$ :

$$P_{L,R} = \frac{(1 \pm \gamma_5)}{2}, \quad (1.1.4)$$

$$\psi_{L,R} = P_{L,R}\psi = [(1 \mp \gamma_5)/2]\psi, \quad \bar{\psi}_{L,R} = P_{L,R}\bar{\psi} = \bar{\psi}[(1 \pm \gamma_5)/2], \quad (1.1.5)$$

defined in terms of the fifth Dirac's matrix:

$$\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3, \quad (1.1.6)$$

and the other  $\gamma$  matrices as:

$$\gamma_\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \tilde{\sigma}^\mu & 0 \end{pmatrix}, \quad (1.1.7)$$

where

$$\sigma^\mu = (1, \vec{\sigma}) = \left[ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right], \quad \tilde{\sigma}^\mu = (1, -\vec{\sigma}). \quad (1.1.8)$$

The  $\sigma^\mu$  here are the Pauli matrices forming a basis for the  $SU(2)$  group.

Experiments show that weak interactions violate charge-parity symmetry and thus differentiate between left- and right-handed chiral particles (see Section 1.6). The left-handed fields transform as  $SU(2)_L$  doublets, while their right-handed partners transform as  $SU(2)_L$  singlets. Fermionic

matter content in the SM can be represented then as a three-fold family structure:

$$\begin{bmatrix} \nu_e & u \\ e^- & d \end{bmatrix}, \quad \begin{bmatrix} \nu_\mu & c \\ \mu^- & s \end{bmatrix}, \quad \begin{bmatrix} \nu_\tau & t \\ \tau^- & b \end{bmatrix}, \quad (1.1.9)$$

with the notation:

$$\begin{bmatrix} \nu_l & u \\ l^- & d \end{bmatrix} \equiv \begin{pmatrix} \nu_l \\ l^- \end{pmatrix}_L, \quad \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad l_R^-, \quad u_R, \quad d_R, \quad (1.1.10)$$

and their corresponding antiparticles. Lepton and quark fields have left- and right-handed components, while there is no right-handed component for the neutrino since neutrinos are purely left-handed particles within the SM.

## 1.2 Gauge principle in SM interactions

### 1.2.1 Concept of the Gauge principle

Consider a Lagrangian density  $\mathcal{L}(\psi, \partial_\mu \psi)$ , depending on a field  $\psi$  and its first derivatives, which are invariant under a  $D$ -dimensional continuous group  $\Gamma$  of transformations:

$$\psi'(x) = U(\theta^A)\psi(x) \quad (A = 1, 2, \dots, D), \quad (1.2.1)$$

with:

$$U(\theta^A) = \exp \left\{ i \sum_A \theta^A t_A \right\} \sim 1 + i \sum_A \theta^A t_A + \dots \quad (1.2.2)$$

Here the quantities  $\theta^A$  are numerical parameters, and the approximate expression on the right is valid for  $\theta^A$  infinitesimal. The matrices  $t_A$  are the generators of the group  $\Gamma$  of transformations (see Eq. 1.2.1) in the appropriate representation of the fields  $\psi$ . For  $D = 2$ , the above expression corresponds to the group  $SU(2)$  whose 3 generators are the Pauli matrices. For  $D = 3$ , the 8 generators of  $SU(3)$  correspond to the Gell-Mann  $\lambda_\alpha$  matrices [3].

In Eq. 1.2.1 we considered only the case of local symmetries so that  $t_A$  are matrices that are independent of the space-time coordinates as well as of the fields  $\psi$  and  $\psi'$ . The gauge principle arises when one makes the parameters  $\theta^A$  depend on the space-time coordinates  $\theta^A = \theta^A(x_\mu)$  going from global to local transformation. In this case  $\mathcal{L}(\psi, \partial_\mu \psi)$  is in general no longer invariant under the gauge transformations  $U[\theta^A(x_\mu)]$  because of the derivative terms:  $\partial_\mu \psi' = \partial_\mu (U\psi) \neq U\partial_\mu \psi$ . The gauge invariance is then recovered only if the ordinary derivative is replaced by the covariant derivative:

$$D_\mu = \partial_\mu + igV_\mu^A t_A, \quad (1.2.3)$$

where  $g$  is a coupling constant and  $V_\mu^A$  are a set of gauge fields (in one-to-one correspondence with the group generators). For the SM gauge groups  $SU_C(3) \otimes SU_L(2) \otimes U_Y(1)$  the corresponding coupling constants are noted as  $g_s$ ,  $g_2$  and  $g_1$ , and their gauge fields as  $G_\mu^a$ ,  $W_\mu^a$  and  $B_\mu$ . An

abstract set of gauge fields  $V_\mu^A$  follows the transformation law:

$$V_\mu'^A = UV_\mu^A t_A U^{-1} - \frac{1}{ig} (\partial_\mu U) U^{-1}. \quad (1.2.4)$$

This ensures that:

$$(D_\mu \psi)' = U(D_\mu \psi). \quad (1.2.5)$$

Thus  $\mathcal{L}(\psi, D_\mu \psi)$  is invariant under gauge transformations. At this stage, the gauge fields  $V_\mu^A$  appear as external fields that do not propagate. In order to construct a gauge-invariant kinetic Lagrangian density for the gauge fields  $V_\mu^A$ , we consider:

$$[D_\mu, D_\nu] \psi = ig \{ \partial_\mu V_\nu^A t_A - \partial_\nu V_\mu^A t_A + ig [V_\mu^B t_B, V_\nu^C t_C] \} \psi \equiv ig F_{\mu\nu}^A t_A \psi, \quad (1.2.6)$$

with:

$$F_{\mu\nu}^A = \partial_\mu V_\nu^A - \partial_\nu V_\mu^A - gf_{ABC} V_\mu^B V_\nu^C. \quad (1.2.7)$$

From Eq. (1.2.1), (1.2.5) and (1.2.6) it follows that the Lagrangian density:

$$\mathcal{L}_{\text{YM}} = -\frac{1}{4} \sum_A F_{\mu\nu}^A F^{A\mu\nu} + \mathcal{L}[\psi, D_\mu \psi] \quad (1.2.8)$$

is invariant under the local gauge transformations [4]. This is how the gauge principle acts to generate the interaction terms  $F_{\mu\nu} F^{\mu\nu}$  in the Lagrangian via gauge fields, which can also have self-interaction through the kinetic part of the Lagrangian density.

## 1.2.2 Gauge principle in QED and QCD

Applying the gauge invariance principle and using particle classification, we can determine the Lagrangian densities for QED (quantum electrodynamics) and QCD (quantum chromodynamics). For simplicity, let us consider a single family of quarks and the notations following Eq. 1.1.10:

$$\psi_1(x) = \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad \psi_2(x) = u_R, \quad \psi_3(x) = d_R. \quad (1.2.9)$$

The same is valid for the lepton sector, with the notation

$$\psi_1(x) = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \quad \psi_2(x) = \nu_{eR}, \quad \psi_3(x) = e_R^-. \quad (1.2.10)$$

The propagation of free fermions is given by the Dirac Lagrangian, which is obtained by writing the Lorentz invariant operators acting on  $\Psi$ :

$$\mathcal{L}_{\text{free}} = \bar{\Psi}(i \not{\partial} - m)\Psi, \quad (1.2.11)$$

where  $\bar{\Psi} = \Psi^\dagger \gamma^0$  and  $\not{\partial} = \partial_\mu \gamma^\mu$ . For the massless particles it gives:

$$\mathcal{L}_{free} = \sum_{j=1}^3 i \bar{\psi}_j(x) \gamma^\mu \partial_\mu \psi_j(x). \quad (1.2.12)$$

$\mathcal{L}_{free}$  is invariant under the following global transformations:

$$\psi_1(x) \xrightarrow{U(1)} \psi'_1(x) \equiv \exp\{i y_1 \beta\} U_L \psi_1(x), \quad (1.2.13)$$

$$\psi_2(x) \xrightarrow{U(1)} \psi'_2(x) \equiv \exp\{i y_2 \beta\} \psi_2(x), \quad (1.2.14)$$

$$\psi_3(x) \xrightarrow{U(1)} \psi'_3(x) \equiv \exp\{i y_3 \beta\} \psi_3(x), \quad (1.2.15)$$

where the  $SU(2)_L$  transformation:

$$U_L \equiv e^{i t_A \alpha^A} \quad (i = 1, 2, 3) \quad (1.2.16)$$

only acts on the doublet field  $\psi_1$  and  $y_i$  are the hypercharges.

To make the Lagrangian density from Eq. 1.2.12 gauge invariant, the invariance under local  $SU(2)_L \otimes U(1)_Y$  gauge transformations, i.e., with  $\alpha^A = \alpha^A(x)$  and  $\beta = \beta(x)$  is required. To satisfy this symmetry requirement, we introduce covariant derivatives and thus 4 gauge bosons with 2 coupling constants,  $g_1$  and  $g_2$ :

$$D_\mu \psi_1(x) \equiv [\partial_\mu + i g_2 t_A W_\mu^A(x) + i g_1 y_1 B_\mu(x)] \psi_1(x), \quad (1.2.17)$$

$$D_\mu \psi_2(x) \equiv [\partial_\mu + i g_1 y_2 B_\mu(x)] \psi_2(x), \quad (1.2.18)$$

$$D_\mu \psi_3(x) \equiv [\partial_\mu + i g_1 y_3 B_\mu(x)] \psi_3(x). \quad (1.2.19)$$

Using these definitions, the Lagrangian density for QED (fermions and kinetic terms for the gauge fields) can be rewritten as follows:

$$\mathcal{L}_{QED} = \sum_{j=1}^3 i \bar{\psi}_j(x) \gamma^\mu D_\mu \psi_j(x) - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^A W_A^{\mu\nu}. \quad (1.2.20)$$

For the strong interaction, the procedure is similar, except that the right and left-handed fermions interact in the same way. The final quantum chromodynamics Lagrangian density reads:

$$\begin{aligned} \mathcal{L}_{QCD} &= \bar{\psi}(x) (i \gamma^\mu D_\mu - m) \psi(x) - \frac{1}{4} G_{\mu\nu}^\alpha G^{\alpha,\mu\nu} \\ &= \bar{\psi}(x) (i \gamma^\mu \partial_\mu - m) \psi(x) - g_s \bar{\psi} \gamma^m u \frac{\lambda_\alpha}{2} \psi G_\mu^\alpha - \frac{1}{4} G_{\mu\nu}^\alpha G^{\alpha,\mu\nu}, \end{aligned} \quad (1.2.21)$$

where  $G_{\mu\nu}^\alpha$  is the color fields tensor.

The kinetic term of the free Lagrangian in Eq. 1.2.11 has the form  $i \bar{\psi} \gamma_\mu \partial^\mu \psi = i \bar{\psi} \not{\partial} \psi$ . The covariant derivative necessary to make the Lagrangian gauge invariant is given by

$$D_\mu \psi = \left( \partial_\mu + ig_s \frac{\lambda^\alpha}{2} G_\mu^\alpha + ig_2 \frac{\sigma^j}{2} W_\mu^j + ig_1 \frac{Y_q}{2} B_\mu \right) \psi, \quad (1.2.22)$$

where  $Y_q$  is the hypercharge, defined via the electric charge  $Q$  and the third component of the weak isospin  $T_3$  by the Gell-Mann-Nishijima formula [5, 6]:

$$Q = T_3 + \frac{Y_q}{2}. \quad (1.2.23)$$

The free kinetic term for a fermion is then

$$\mathcal{L}_{fermion} = i\bar{\psi} \not{D}\psi. \quad (1.2.24)$$

### 1.2.3 Gauge principle in electroweak interaction

In the SM, the fermion kinetic terms and the interactions of gauge bosons with vector and axial vector fermion currents all conserve chirality while the fermion mass terms flip chirality. For example, if an electron emits a photon, the electron chirality is unchanged. In the ultrarelativistic limit, when the electron mass can be neglected, chirality and helicity are approximately the same, and we can state that the helicity of the electron is unchanged by the photon emission. In a massless gauge theory, the  $\psi_{L,R}$  fermion components are uncoupled and can be transformed separately. If, in a gauge theory, the  $\psi_{L,R}$  components transform as different representations of the gauge group, one speaks of a chiral gauge theory, while if they have the same gauge transformations, one has a vector gauge theory. QED and QCD are vector gauge theories: for each given fermion,  $\psi_L$  and  $\psi_R$  have the same electric charge and the same colour. Instead, the Electroweak (EW) interaction is a chiral theory, in the sense that  $\psi_L$  and  $\psi_R$  behave differently under the gauge group.

The free Dirac Lagrangian expanded in terms of Weyl spinors (making use of the identity  $\psi = \psi_L + \psi_R$ ) takes the form:

$$\begin{aligned} \mathcal{L}_{free} &= \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x) \\ &= \bar{\psi}_L i\gamma^\mu \partial_\mu \psi_L + \bar{\psi}_R i\gamma^\mu \partial_\mu \psi_R - m\bar{\psi}_L \psi_R - m\bar{\psi}_R \psi_L. \end{aligned} \quad (1.2.25)$$

The fermion mass terms are not invariant under  $SU(2)_L$  transformation, as left fermions are grouped into  $SU(2)_L$  doublets while right fermions are  $SU(2)_L$  singlets. The term of the form  $\frac{1}{2}M_V^2 W_\mu W^\mu$  in the gauge field Lagrangian part would not preserve local gauge invariance. Thus, mass terms for fermions (of the form  $\bar{\psi}_L \psi_R + \text{h.c.}$ ) are forbidden in the EW gauge-symmetric limit. Moreover, while the strong interaction mediators, gluons, are massless, the masses of weak interaction mediators,  $W$  and  $Z$  bosons, are non-zero. The SM was extended to explain why  $W$  and  $Z$  bosons (and fermions) have masses by introducing the Brout-Englert-Higgs (BEH) mechanism (see Section 1.3).

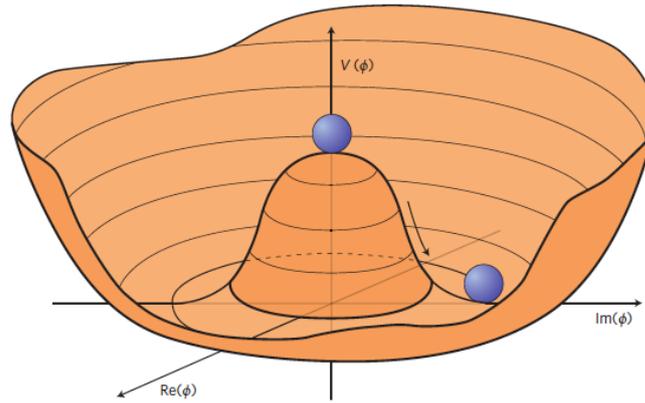


Figure 1.3.1: An illustration of the Higgs potential in the case  $\mu^2 < 0$  with the minimum at  $|\phi|^2 = -\mu^2/(2\lambda)$  [7].

## 1.3 Spontaneous symmetry breaking and the Brout-Englert-Higgs mechanism

### 1.3.1 The Brout-Englert-Higgs mechanism

The Brout-Englert-Higgs (BEH) mechanism allows massive gauge bosons without EW gauge invariance breaking. The idea of the BEH mechanism is to introduce a scalar field with a potential that contains mass-term and self-coupling terms, combining local gauge invariance with spontaneous symmetry breaking (SSB). SSB happens in a system when its stable state is not symmetric under a symmetry of its Lagrangian. For the rotation-invariant Higgs field, the ground state acquires a non-zero value (vacuum expectation value) that breaks the symmetry.

Let us see how this works for a simplified electromagnetic Lagrangian density including a scalar field  $\phi$ :

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^2 + |(\partial_\mu - ieA_\mu)\phi|^2 + \mu^2\phi^*\phi - \frac{\lambda}{2}(\phi^*\phi)^2. \quad (1.3.1)$$

The sign in front of the mass term for the scalar field  $\phi$  is necessary for the spontaneous symmetry breaking to take place. Since we consider the electromagnetic interaction, the Lagrangian (Eq. 1.3.1) is invariant under  $U(1)$ :

$$A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu\theta(x), \quad \phi \rightarrow \phi' = \exp[-ie\theta(x)] \phi. \quad (1.3.2)$$

Let  $\phi^0 = v \neq 0$ , with  $v$  real, be the ground state that minimizes the potential and induces spontaneous symmetry breaking. In our case,  $v$  is given by  $v^2 = \mu^2/\lambda$ . Using the gauge invariance, we can do the change of variables

$$\begin{aligned} \phi(x) &\rightarrow \left[v + \frac{h(x)}{\sqrt{2}}\right] \exp\left[-i\frac{\zeta(x)}{v\sqrt{2}}\right], \\ A_\mu(x) &\rightarrow A_\mu - \partial_\mu\frac{\zeta(x)}{ev\sqrt{2}}. \end{aligned} \quad (1.3.3)$$

The field  $\zeta(x)$  is called the Goldstone boson [8]. The position of the minimum at  $\phi^0 = v$  corre-

sponds to  $h = 0$ , and the Lagrangian density becomes:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + e^2v^2A_\mu^2 + \frac{1}{2}e^2h^2A_\mu^2 + \sqrt{2}e^2hvA_\mu^2 + \mathcal{L}(h). \quad (1.3.4)$$

Here, the term  $e^2v^2A_\mu^2$  is a mass term for the field  $A_\mu$ , which means that the gauge boson acquires a mass  $M = \sqrt{2}ev$ . Leaving a constant term aside, the last term in Eq. 1.3.4 is given by:

$$\mathcal{L}(h) = \frac{1}{2}\partial_\mu h \partial^\mu h - h^2\mu^2 + \dots, \quad (1.3.5)$$

where the dots stand for cubic and quartic terms in  $h$ . We see that the  $h$  mass term has the "right" sign due to the combination of the quadratic terms in  $h$  that, after the shift, arise from the quadratic and quartic terms in  $\phi$ . The  $h$  mass is then given by  $m_h^2 = 2\mu^2$ .

This mechanism can be applied to the full Lagrangian density of the SM after breaking the gauge symmetry  $SU_L(2) \otimes U_Y(1)$  into massive  $W$  and  $Z$  bosons and a massless photon, thus with a preserved  $U_{QED}(1)$  gauge symmetry. Then, according to the Goldstone's theorem, we are left with a scalar BEH boson that is also charge neutral and  $\mathcal{CP}$ -even (see Section 1.6). SSB also allows the BEH field to give mass to fermions through the Yukawa couplings (see Section 1.3.2).

Let us introduce an  $SU(2)_L$  doublet of complex scalar fields:

$$\phi(x) \equiv \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix}. \quad (1.3.6)$$

The Lagrangian density for this field can be written as:

$$\mathcal{L}_S = (D_\mu\phi)^\dagger D^\mu\phi - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \quad (\lambda > 0, \mu^2 < 0), \quad (1.3.7)$$

which is invariant under local  $SU(2)_L \otimes U(1)_Y$  transformations. There is an infinite set of degenerate states with minimum energy, satisfying:

$$|\langle 0|\phi_2|0\rangle| = \sqrt{\frac{-\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}}. \quad (1.3.8)$$

Once a particular ground state is chosen, the  $SU(2)_L \otimes U(1)_Y$  symmetry gets spontaneously broken to the electromagnetic subgroup  $U(1)_{QED}$ , which by construction still remains a true symmetry of the vacuum. According to the Goldstone theorem, 3 massless states should appear. However, we have the freedom to parameterize the scalar doublet in the general form:

$$\phi(x) = \exp\{\{it_A\theta^A(x)\}\} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (1.3.9)$$

with four real fields:  $\theta^i(x)$  and  $h(x)$ . The crucial point is that the local  $SU(2)_L$  symmetry allows us to rotate away any dependence on  $\theta^A(x)$ : instead of  $\phi(x)$  from Eq. 1.3.9, we can consider

equivalently:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (1.3.10)$$

where the 3 other fields ( $\theta^A(x)$ ) are the massless Goldstone bosons associated with the SSB mechanism. Developing Eq. 1.3.7 will generate masses to the gauge bosons  $W^\pm$  and  $Z$ . Using the covariant derivatives defined in Eq. 1.2.19, we find (with a unit hypercharge for the BEH scalar field):

$$\begin{aligned} (D_\mu \phi)^\dagger D^\mu \phi &= \left| (\partial_\mu + ig_2 t_A W_\mu^A + ig_1 B_\mu) \frac{(v+h)}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|^2 \\ &= \left| \begin{pmatrix} 0 \\ \partial_\mu h / \sqrt{2} \end{pmatrix} + i \frac{g_2}{2} (v+h) \begin{pmatrix} W_\mu^+ \\ (-1/\sqrt{2} \cos \theta_W) Z_\mu \end{pmatrix} \right|^2 \\ &= \frac{1}{2} \partial_\mu h \partial^\mu h + \frac{g_2^2}{4} (v+h)^2 \left( W_\mu^+ W^{-\mu} + \frac{1}{2 \cos^2 \theta_W} Z_\mu Z^\mu \right). \end{aligned} \quad (1.3.11)$$

Eq. 1.3.11 leads to terms proportional to  $(\partial h)^2$ ,  $h$  and  $h^2$  and to a term (independent of  $h$ ) of the form:  $M_W^2 W^{+\mu} W_\mu^- + \frac{M_Z^2}{2} Z^\mu Z_\mu$ . That's why SSB generates mass terms for the  $W$  and  $Z$  gauge bosons:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}} (W_\mu^1 \mp i W_\mu^2) \\ Z_\mu &= -\sin \theta_W B_\mu + \cos \theta_W W_\mu^3, \end{aligned} \quad (1.3.12)$$

and the field associated to the photon:

$$A_\mu = \cos \theta_W B_\mu + \sin \theta_W W_\mu^3 \quad (1.3.13)$$

remains massless. From Eq. 1.3.11 and 1.3.12 the masses are found to be  $M_W = \frac{gv}{2}$  and  $M_Z = \frac{M_W}{\cos \theta_W}$  where the weak angle is defined by  $\tan \theta_W \equiv g_1/g_2$ . The generation of masses can be thought of as due to the fact that the  $W$  and  $Z$  bosons interact constantly with the condensate of scalar fields and, therefore, acquire masses. The Goldstone boson has disappeared from the theory but has reemerged as the longitudinal degree of freedom of a massive vector particle.

To summarize, with SSB, there are 3 broken generators giving rise to 3 massless Goldstone bosons, which, owing to the underlying local gauge symmetry, can be eliminated from the Lagrangian density. It is instructive to count the number of degrees of freedom (d.o.f.). Before the SSB mechanism, the Lagrangian density contains massless  $W^\pm$  and  $Z$  bosons, 6 d.o.f., due to the two possible polarizations of a massless spin-1 field, and 4 real scalar fields: a total of 10. After SSB, the 3 Goldstone modes can be gauged by the weak bosons, which become massive and, therefore, acquire one additional longitudinal polarization. Then, we have 9 d.o.f. in the gauge sector, plus the remaining scalar particle  $h$ , a total of 10, that corresponds to the number before SSB.

### 1.3.2 Fermion masses via the BEH mechanism

In the SM, a fermionic mass term  $\mathcal{L}_m = -m\bar{\psi}\psi = -m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)$  is not allowed because it breaks the gauge symmetry. However, with the notion introduced in Section 1.3.1, namely the additional scalar doublet, we can write the following gauge-invariant fermion-scalar coupling Lagrangian density:

$$\mathcal{L}_Y = -c_1 (\bar{u}, \bar{d})_L \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} d_R - c_2 (\bar{u}, \bar{d})_L \begin{pmatrix} \phi_2^* \\ -\phi_1^* \end{pmatrix} u_R - c_3 (\bar{\nu}_e, \bar{e})_L \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} e_R + \text{h.c.}, \quad (1.3.14)$$

where the notation  $\phi_{1,2}^*$  stands for the complex conjugate of the fields  $\phi_{1,2}$ . The interaction between a fermionic and a scalar field is called the Yukawa interaction [9].

Replacing the  $\phi$  with its expression in terms of  $h$  (Eq. 1.3.10), this Lagrangian density can be rewritten as:

$$\mathcal{L}_Y = -\frac{1}{\sqrt{2}}(v+h) \{c_1 \bar{d}d + c_2 \bar{u}u + c_3 \bar{e}e\}, \quad (1.3.15)$$

which shows that the SSB mechanism generates fermion masses:

$$m_d = c_1 \frac{v}{\sqrt{2}}, \quad m_u = c_2 \frac{v}{\sqrt{2}}, \quad m_e = c_3 \frac{v}{\sqrt{2}}. \quad (1.3.16)$$

These masses are expressed in terms of the couplings of the fermions to the Higgs doublet.

We considered only one fermion generation. Returning to the general case would imply that  $m_f$  and  $c_f$  could be represented as  $3 \otimes 3$  matrices mixing the different generations. For example, for the charged leptons:

$$M^e = U_L^{e\dagger} \begin{pmatrix} m_e & 0 & 0 \\ 0 & m_\mu & 0 \\ 0 & 0 & m_\tau \end{pmatrix} U_L^e, \quad (1.3.17)$$

where  $U$  is a unitary matrix.

In practice, fermionic fields are chosen so that these mass matrices are diagonal, which has no effect on most of the other terms of the Lagrangian, except for the coupling of fermions to the  $W_\mu^\pm$  fields: these are the only fields that simultaneously involve the two components of an  $SU(2)_L$  doublet, all the others being diagonal. Consequently, we denote the Cabibbo-Kobayashi-Maskawa (CKM) matrix as  $V^q = U_L^q U_L^{d\dagger}$  [10, 11], and its equivalent in the lepton sector, the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix  $V^l = U_L^\nu U_L^{e\dagger}$  [12, 13].

### 1.3.3 Higgs boson couplings

Let us note that after symmetry breaking the BEH potential becomes:

$$V(\phi) \simeq -\frac{\mu^4}{4\lambda} - \mu^2 h^2 + \lambda v h^3 + \frac{\lambda}{4} h^4. \quad (1.3.18)$$

The second term represents a (tree-level) mass:

$$M_H = \sqrt{-2\mu^2} = \sqrt{2\lambda}v. \quad (1.3.19)$$

The Higgs Lagrangian describes the Higgs boson couplings to fermions (with a coupling constant proportional to their mass), to gauge bosons (with a coupling constant proportional to their mass squared) and to itself (with a coupling constant proportional to its mass squared). Figure 1.3.2 summarizes these couplings.

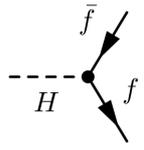
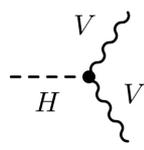
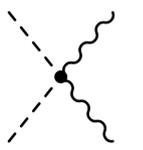
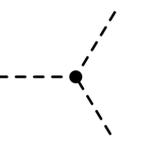
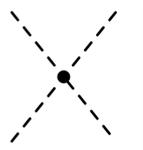
Diag.					
Coupl.	$\frac{m_f}{v}$	$2\frac{M_V^2}{v}$	$2\frac{M_V^2}{v^2}$	$3\frac{M_H^2}{v}$	$3\frac{M_H^2}{v^2}$

Figure 1.3.2: Higgs boson couplings to fermions, gauge bosons and itself [14].

## 1.4 Higgs boson measurements at the Large Hadron Collider

In 2012, the ATLAS [15] and CMS [16] collaborations at the LHC discovered the Higgs boson. This scalar particle, associated with the Higgs field, has a mass of  $125.25 \pm 0.17$  GeV [1]. Using the measured mass, the couplings, production and decay rates of the Higgs boson can be calculated. Its properties are found to be consistent with the SM predictions. Nevertheless, the precise measurements of the Higgs boson properties are still important for unravelling new physics phenomena and studying the electroweak symmetry breaking. This section discusses Higgs production and decay modes in proton-proton collisions at the LHC. A detailed description of the Large Hadron Collider and ATLAS experiment is given in the Chapter 2.

The production of a Higgs boson involves its coupling to different SM particles: either during production (quarks and vector bosons) or during decay (into leptons, quarks, and vector bosons). For any given production ( $ii$ ) and decay ( $ff$ ) configurations, the number of Higgs bosons can be calculated using the following formula [17]:

$$N(ii \rightarrow H \rightarrow ff) \approx \sigma(ii \rightarrow H) \cdot B(H \rightarrow ff) = \sigma_i \times \text{BR}^f \approx \frac{\sigma_i \Gamma_f}{\Gamma_{\text{tot}}}. \quad (1.4.1)$$

Here  $\Gamma_f$  is the partial decay width of the Higgs boson to a pair of  $f$  particles, and  $\Gamma_{\text{tot}}$  is the total Higgs boson decay width. The production and decay mechanism can be considered independent and can be factorized due to the small predicted width of the Higgs boson. To parameterize eventual deviations from the SM, one can introduce the signal strength modifiers as  $\mu_i^f = \frac{\sigma_i \times \text{BR}^f}{(\sigma_i \times \text{BR}^f)_{\text{SM}}}$ . By definition, in the SM  $\mu_i^f = 1$  for all the allowed decays. Any substantial deviation from unity would suggest the existence of physics beyond the standard model (BSM).

### 1.4.1 Higgs production modes

The main production mechanisms at  $pp$  colliders are gluon fusion ( $ggF$ ), weak-boson fusion ( $VBF$ ), associated production with a gauge boson ( $VH$ ), and associated production with a  $t\bar{t}$  pair ( $t\bar{t}H$ ) or with a single top quark ( $tHq$ ). Figure 1.4.1 shows Feynman diagrams for these dominant Higgs boson production processes.

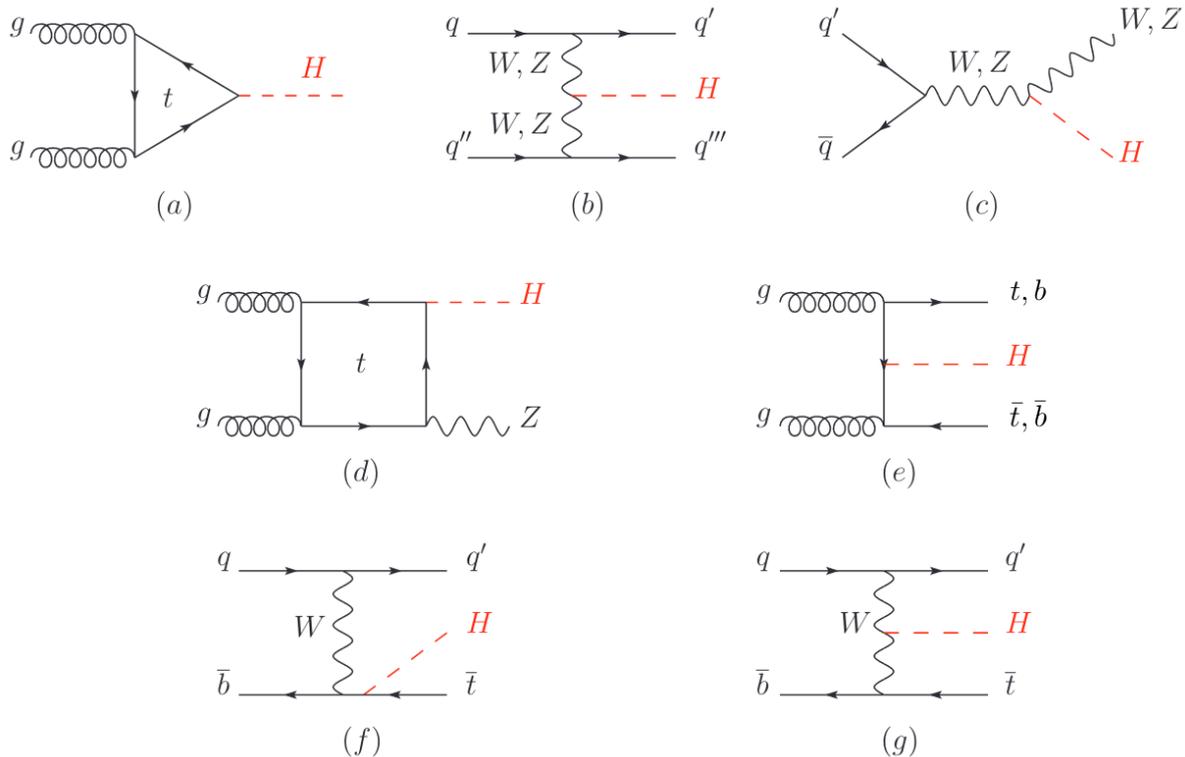


Figure 1.4.1: Main leading order Feynman diagrams contributing to single Higgs boson production in (a) gluon fusion, (b) vector-boson fusion, (c) associated production with a gauge boson at tree level from a quark-quark interaction, (d) associated production with a gauge boson (at loop level from a gluon-gluon interaction), (e) associated production with a pair of top/bottom quarks, (f-g) production in association with a single top quark [1]

The cross sections for the production of SM Higgs boson as a function of the centre of mass energy for  $pp$  collisions are depicted in Figure 1.4.2. With the increase of the centre of mass energy, the production cross-sections also increase. For example, the  $t\bar{t}H$  production cross-section within Run-2 at  $\sqrt{s} = 13$  TeV is around four times higher compared to the one within Run-1 at  $\sqrt{s} = 7$  and 8 TeV.

We will discuss in the following the main Higgs production processes:

- Gluon fusion production mechanism ( $ggH$ ). The Higgs boson production mechanism with the largest cross-section at LHC is the gluon-fusion process,  $gg \rightarrow H + X$  (see Figure 1.4.1 (a)). It occurs when two gluons from the colliding protons interact by exchanging quarks. The quarks themselves can then emit a Higgs boson. At leading order, the Higgs boson coupling to gluons is induced by a one-loop process in which the Higgs boson couples to a virtual  $t\bar{t}$  pair (with minor contributions from the other lighter quarks) [19].

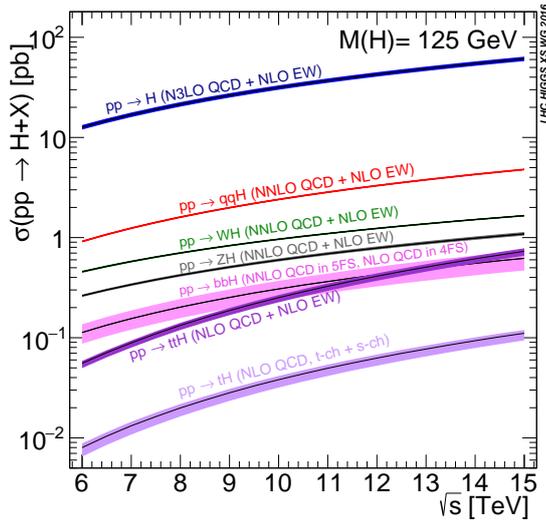


Figure 1.4.2: The SM Higgs boson production cross-sections as a function of the centre of mass energy,  $\sqrt{s}$ , for  $pp$  collisions [18]. The theoretical uncertainties are indicated as bands.

Decay channel	Branching ratio	Rel. uncertainty
$H \rightarrow \gamma\gamma$	$2.27 \times 10^{-3}$	2.1%
$H \rightarrow ZZ$	$2.62 \times 10^{-2}$	$\pm 1.5\%$
$H \rightarrow W^+W^-$	$2.14 \times 10^{-1}$	$\pm 1.5\%$
$H \rightarrow \tau^+\tau^-$	$6.27 \times 10^{-2}$	$\pm 1.6\%$
$H \rightarrow b\bar{b}$	$5.82 \times 10^{-1}$	+1.2% -1.3%
$H \rightarrow c\bar{c}$	$2.89 \times 10^{-2}$	+5.5% -2.0%
$H \rightarrow Z\gamma$	$1.53 \times 10^{-3}$	$\pm 5.8\%$
$H \rightarrow \mu^+\mu^-$	$2.18 \times 10^{-4}$	$\pm 1.7\%$

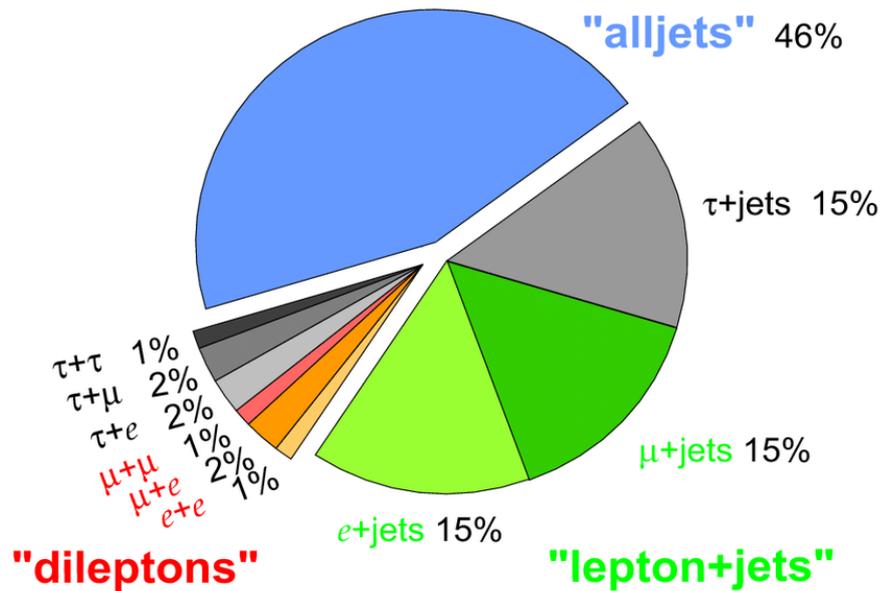
Table 1.4.1: The branching ratios and the relative uncertainty for an SM Higgs boson with  $m_H = 125$  GeV [1].

- Vector boson fusion ( $VBF$ ). The vector boson fusion,  $qq \rightarrow qqH$ , is more than 10 times rarer than  $ggH$ . This production mode proceeds by the scattering of two quarks, mediated by the exchange of a  $W$  or  $Z$  boson (see Figure 1.4.1 (b)).  $VBF$  gives direct access to the Higgs coupling to the heavy gauge bosons. The scattered quarks give rise to two hard jets in the forward and backward regions of the detector with large transverse momentum. These characteristic features of  $VBF$  processes is exploited to distinguish them from QCD backgrounds.
- Associated production with a vector boson ( $VH$ ). The next largest Higgs boson production mechanisms are associated production with  $W$  and  $Z$  gauge bosons,  $pp \rightarrow VH + X$  (see Figure 1.4.1 (c, d)). The  $WH$  and  $ZH$  production modes may provide a relatively clean environment for studying the decay of the Higgs boson into bottom quarks.
- Higgs boson production in association with top quarks ( $t\bar{t}H$ ,  $tH$ ) has a relatively small cross-section: production with a pair of quarks accounts for around 1% and with a single top-quark 0.05% of the total production cross-section. Within the production mode  $pp \rightarrow t\bar{t}H$ , the Higgs boson is produced through the fusion of a top quark-antiquark pair or through radiation from a top quark (see Figure 1.4.1 (e, f, g)).

## 1.4.2 Higgs decay modes

With a Higgs boson mass of approximately 125 GeV, extensive exploration of its couplings to various SM particles is feasible. Given its short lifetime of about  $1.6 \times 10^{-22}$  seconds, the Higgs boson decays almost immediately after being produced, so experiments, not being able to directly observe the Higgs boson, have to reconstruct it from its decay products. The Higgs boson decays

Figure 1.5.1: Top-quark branching ratios [20].



preferably to the heaviest kinematically accessible particles. Notably, the primary decay modes include  $H \rightarrow b\bar{b}$  and  $H \rightarrow W^+W^-$  (see Table 1.4.1).

Although in the SM, direct coupling of the Higgs boson to massless photons is absent, it can yield two photons in the final state via loops. The loop-induced decays into gluons, diphotons, and  $Z\gamma$  provide indirect insights into the Higgs boson's couplings to  $WW$ ,  $ZZ$ , and  $t\bar{t}$ .

## 1.5 Higgs-top Yukawa coupling

### 1.5.1 Top-quark physics

The top quark, which belongs to the third generation of quarks along with the bottom quark, is the heaviest particle in the SM with a mass of approximately  $172.69 \pm 0.30$  GeV [1]. Accurate measurements of top-quark production and decay are essential for verifying QCD calculations at high energy scales and for potentially uncovering indirect signs of new physics. At the LHC, the primary mode of top-quark production is through the strong interaction, resulting in the creation of top-antitop pairs. Single top-quark production also occurs, though it is less common and arises from the electroweak interaction, involving bottom quarks or antiquarks in the initial state.

The top quark has a substantial width ( $1.41 \pm 0.17$  GeV [1]) and a very short lifetime, decaying before it can undergo hadronization. The top quark almost always decays into a bottom quark and an on-shell  $W$  boson. The bottom quark typically hadronizes, while the  $W$  boson can decay either into hadrons (68% of the time) or into leptons and neutrinos (32% of the time). Consequently, the final state of a  $t\bar{t}$  event is characterized by the type and number of leptons produced by the two  $W$  bosons in the event (see Figure 1.5.1). The decays might be fully hadronic with both  $W$  decaying into hadrons, leading to a final state with at least six jets. This mode has the largest branching ratio at approximately 46%. In semi-leptonic decays one  $W$  boson decays into leptons, while the other decays into hadrons, accounting for 35% of the total decays. The dilepton final state originates from both  $W$  bosons decaying into leptons, resulting in a branching ratio of about 6.4%.

## 1.5.2 The top Yukawa coupling

As discussed before, the strength of the Higgs coupling to fermions is proportional to the fermion mass. The top quark is the heaviest particle in the SM, and the top quark Yukawa coupling is expected to be close to unity. This coupling is important for the stability of the SM and is sensitive to potential BSM effects. For example, the top Yukawa coupling and Higgs self-coupling are important parameters to understand the vacuum stability of the Universe (see Fig. 1.5.2).

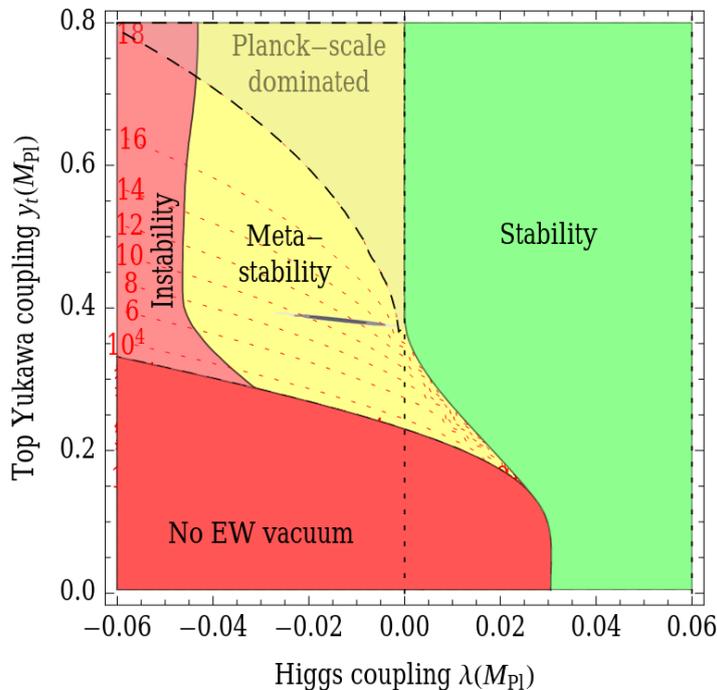
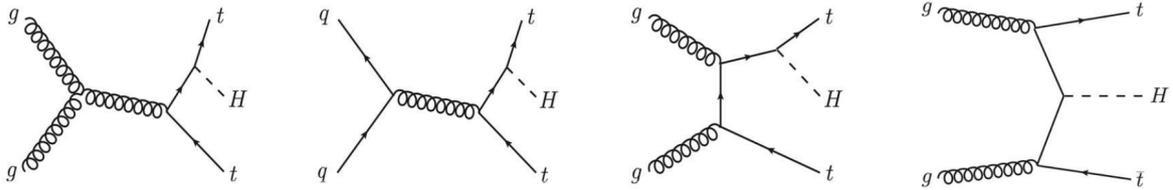


Figure 1.5.2: Electroweak vacuum stability as a function of the top Yukawa and the Higgs self-coupling [21].

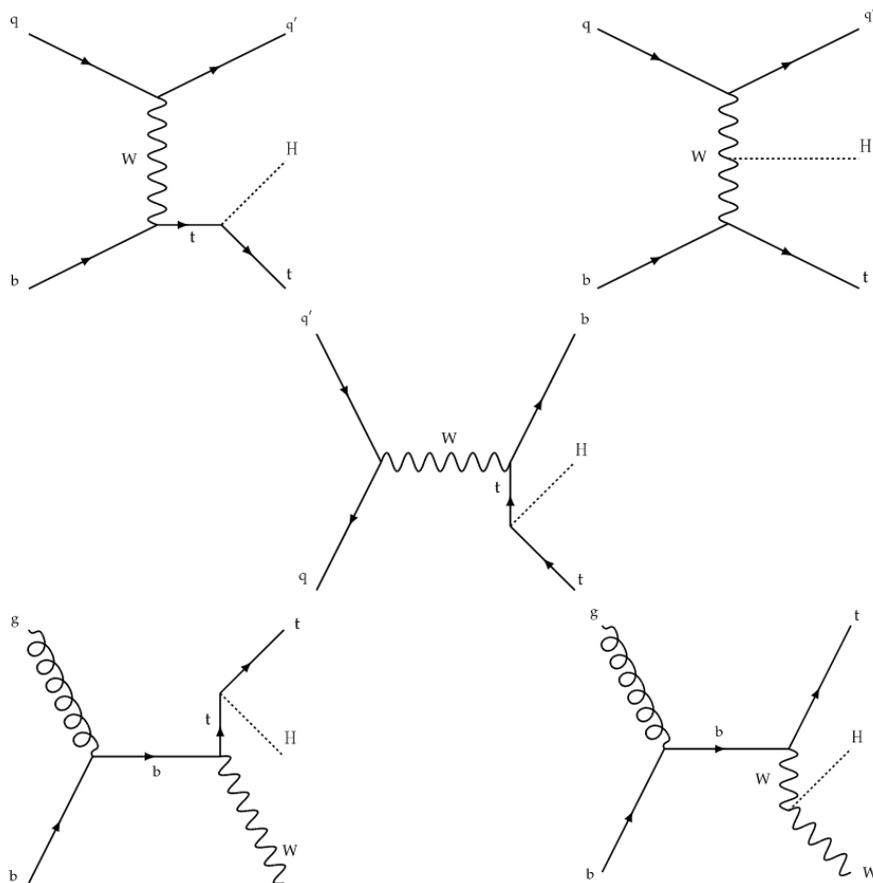
Since the top quark is heavier than the Higgs boson, the  $H \rightarrow t\bar{t}$  decay is not allowed, and, unlike it is for other fermions, the Higgs to top quark coupling can not be measured via Higgs decay rate to the quark-antiquark pair. The top Yukawa coupling can then be measured in a direct or indirect way. Indirect constraints to the top Yukawa coupling are possible through  $ggH$  production as well as the  $H \rightarrow \gamma\gamma$  decay process. Even if the gluon fusion loop is dominated by virtual top quarks, it may contain particles other than  $t$ , which forces us to make assumptions about the influence of BSM effects, making such measurements model-dependent. However, the top Yukawa coupling is directly accessible through the associated production of a Higgs boson with a pair of top quarks or a single top quark. Since these processes are realized at the lowest order in perturbation theory, no assumptions on potential BSM physics are needed. Moreover, comparing the directly measured top-Higgs coupling with the one inferred by other cross-section measurements can constrain contributions from new physics to the gluon fusion loop.

The associated production of a Higgs boson with a pair of top quarks can proceed through gluon-gluon fusion or quark-anti-quark annihilation (see Fig. 1.5.3). The inclusive cross-section of this process is calculated at NNLO in QCD [22]:  $506.5^{+5.8\%}_{-9.2\%}$  (scale)  $\pm 3.6\%$  (PDF +  $\alpha_S$ ) [23], for a center-of-mass energy of  $\sqrt{s} = 13$  TeV.

The associated production of a Higgs boson with a single of top quark is another Higgs boson production mode with a top Yukawa coupling contribution. This production mode can be classified into three main channels:  $t$ -channel,  $s$ -channel and  $tW(H)$  associated production (see Fig. 1.5.4).

Figure 1.5.3: Leading order Feynman diagrams for the  $t\bar{t}H$  process.

This Higgs production mode is always mediated by a  $tWb$  vertex leading to the presence of a  $b$ -quark in either the initial ( $t$ -channel and associated production with a  $W$ -boson) or final state ( $s$ -channel). Among the three mentioned channels, the contribution from the  $t$ -channel is the largest at the LHC. In the  $t$ -channel the Higgs boson is radiated from a top quark or a  $W$ -boson. The opposite signs of these two couplings implies that the corresponding diagrams interfere, leading to a predicted NLO cross-section of 18 fb. At the Run-2 LHC, the total cross-section for  $tH$  production is calculated at NLO QCD:  $74.26^{+6.5\%}_{-14.7\%}$  (scale +FS)  $\pm 3.7\%$  (PDF +  $\alpha_S$ ) fb [23].

Figure 1.5.4: Feynman diagrams at LO for  $tH$  production via the  $t$ -channel (top row) and  $s$ -channel (middle row) processes, and for associated production of a Higgs boson with a single top quark and a  $W$  boson (bottom row) [24].

## 1.6 $\mathcal{CP}$ -violation in the SM

The discrete  $\mathcal{CP}$  transformation merges charge conjugation ( $\mathcal{C}$ ) with parity ( $\mathcal{P}$ ). Under the transformation  $\mathcal{C}$ , a particle is transformed into an antiparticle, involving the conjugation of all internal quantum numbers, such as the electromagnetic charge  $Q$  being transformed to  $-Q$ . Parity reverses the sign of the spacial components:  $\vec{x} \rightarrow -\vec{x}$ . Consequently, for instance, a left-handed electron  $e_L^-$  transforms under  $\mathcal{CP}$  into a right-handed positron  $e_R^+$ . Additionally, there is another operation  $T$  applying the time reversal:  $t \rightarrow -t$ .

Empirical evidence, like the measurements of mass difference between the neutral kaon and antikaon [1, 25], suggests that the combination of these three operations,  $\mathcal{CPT}$ , is a fundamental symmetry of nature [1]. Moreover, it is impossible to construct a locally Lorentz-invariant quantum field theory with a Hermitian Hamiltonian that violates  $\mathcal{CPT}$ .

If  $\mathcal{CP}$  symmetry was exact, the laws of nature would treat matter and antimatter identically. Empirical observations demonstrate that most phenomena exhibit  $\mathcal{C}$  and  $\mathcal{P}$  symmetry, hence also  $\mathcal{CP}$ -symmetry. Although these symmetries hold true for electromagnetic and strong interactions, weak interactions display a violation of both  $\mathcal{C}$  and  $\mathcal{P}$ :  $\psi_L$  and  $\psi_R$  behave differently under the gauge group so that parity and charge conjugation non-conservation is made possible in principle. Within the SM framework,  $\mathcal{CP}$  symmetry is broken by complex phases present in the Yukawa couplings, which represent the interactions of the Higgs scalar with quarks. Despite exhaustive attempts to eliminate unphysical phases in this model, a sole  $\mathcal{CP}$ -violating parameter persists. In the context of mass eigenstates, this parameter manifests as a single phase within the  $3 \times 3$  unitary matrix dictating the  $W$ -boson couplings between an up-type antiquark and a down-type quark. The Standard Model's explanation of  $\mathcal{CP}$  violation via Yukawa couplings, known as the Kobayashi-Maskawa (KM) mechanism, aligns seamlessly with all experimental measurements thus far.

The first evidence of  $\mathcal{CP}$  violation was established in neutral  $K$  meson decays: the longer-lived  $K_L^0$ , which typically decays to three pions and would therefore exhibit odd parity, was observed also decaying to a pair of pions (a  $\mathcal{CP}$ -even state) [26]. This phenomenon can be explained by considering that short- and long-lived neutral kaons are not  $\mathcal{CP}$  eigenstates but rather a combination of  $\mathcal{CP}$ -even and  $\mathcal{CP}$ -odd states. Another observed  $\mathcal{CP}$  violating phenomenon is neutrino oscillations, which means that neutrinos possess mass, implying the arise of a potential mixing between flavor and mass states in the lepton sector. This mixing could be governed by a complex mixing matrix analogous to the CKM matrix, thereby introducing the prospect of  $\mathcal{CP}$  violation in the lepton sector. Detecting  $\mathcal{CP}$  violation in the lepton sector stands as one of the primary objectives for ongoing and near-future experiments. As of now,  $\mathcal{CP}$  violation has not been observed in processes involving the top quark, nor in flavor-conserving processes like electric dipole moments. Should any significant observations arise in these realms, it would unequivocally signal the presence of physics extending beyond the Standard Model.

The current  $\mathcal{CP}$ -violating effects observed are insufficient to account for the matter-antimatter imbalance in the Universe. The Big Bang baryogenesis hypothesis [27] predicts baryon asymmetry (i.e. the difference in the number of baryons and anti-baryons per unit volume) that is measurable by two independent experimental approaches: the abundance of light elements in the intergalac-

tic medium and from the power spectrum of the thermal fluctuations in the cosmic microwave background. Being in good agreement between each other, the measured values of the baryon asymmetry ( $\sim 10^{-10}$  [28]) considerably differs from the predicted one ( $\sim 10^{-18}$  [29]).

To reconcile the observed baryon asymmetry in the present Universe, additional  $\mathcal{CP}$ -violating effects are necessary. We can introduce extra  $\mathcal{CP}$  violation in the scalar sector, for instance through the Two-Higgs-Doublet Model (2HDM [30]). This model incorporates an additional Higgs doublet with identical quantum numbers to the Standard Model's Higgs. Within the 2HDM, the mass spectrum includes one charged Higgs field and three neutral fields (two scalars and one pseudoscalar). Notably, the pseudoscalar field's coupling to fermions is  $\mathcal{CP}$ -odd. Detecting a non-zero value for this coupling would serve as a compelling indicator of  $\mathcal{CP}$  violation in the Higgs sector, thereby suggesting physics beyond the Standard Model. Instead of relying on a specific model, the  $t\bar{t}H$  analysis presented in this work parametrizes the potential  $\mathcal{CP}$ -odd component of the top Yukawa coupling using an effective model, as elaborated in the subsequent section.

## 1.7 Probing the $\mathcal{CP}$ nature of the top-Higgs coupling

The SM predicts the Higgs boson to be a scalar particle with quantum numbers  $J^{CP} = 0^{++}$ , and no foreseen  $\mathcal{CP}$ -violating interactions are predicted in the Higgs sector. Experiments demonstrated consistency with the Higgs  $\mathcal{CP}$ -even quantum state predicted by the SM, while still allowing for small admixtures of non-SM  $\mathcal{CP}$ -even or  $\mathcal{CP}$ -odd states.

Search for a Higgs  $\mathcal{CP}$ -violating component could proceed via studying Higgs bosonic or fermionic couplings. In bosonic couplings,  $\mathcal{CP}$ -odd contributions enter via non-renormalizable higher-order operators that are suppressed by powers of  $\frac{1}{\Lambda^2}$  [31], where  $\Lambda$  is the scale of the physics beyond the SM in an effective field theory. Consequently, the examination of the Higgs boson's interactions with fermions holds promise for heightened sensitivity and increased model independence. This is attributed to the fact that  $\mathcal{CP}$ -odd terms can emerge at comparable magnitudes to  $\mathcal{CP}$ -even terms within the Standard Model, without being constrained by  $\frac{1}{\Lambda^2}$  suppression factors. Given the fact that among coupling with other fermions, the top-Higgs coupling is the strongest one, studying the  $t\bar{t}H$  and  $tH$  productions can contribute to determining whether such coupling is  $\mathcal{CP}$  violating.

As mentioned in Section 1.5, the top Yukawa coupling and its  $\mathcal{CP}$  properties are accessible to be measured directly and indirectly. Some bounds might be set by indirect measurements [32]. If the  $\mathcal{CP}$  component is present in the top-Yukawa coupling, the rates of gluon-fusion production and  $H \rightarrow \gamma\gamma$  decay will be modified in comparison with pure  $\mathcal{CP}$ -even scenario. Studies of Higgs interactions with vector bosons and analyses of associated production with a pair of top quarks in the  $H \rightarrow \gamma\gamma$  channel have excluded a pure  $\mathcal{CP}$ -odd coupling at more than  $3\sigma$  significance [33] (see Section 1.9).

At the LHC the  $\mathcal{CP}$  properties of the top Yukawa coupling can be directly probed through Higgs boson production in association with a pair ( $t\bar{t}H$ ) or a single top quark ( $tH$ :  $tHjb$  and  $tWH$ ). The presence of  $\mathcal{CP}$ -odd component in these processes would impact not only the production rates, but also affect the kinematic distributions. For example, due to the interference happening in the  $tH$   $t$ -channel production between the Higgs boson radiated from a top quark or a  $W$ -boson, the

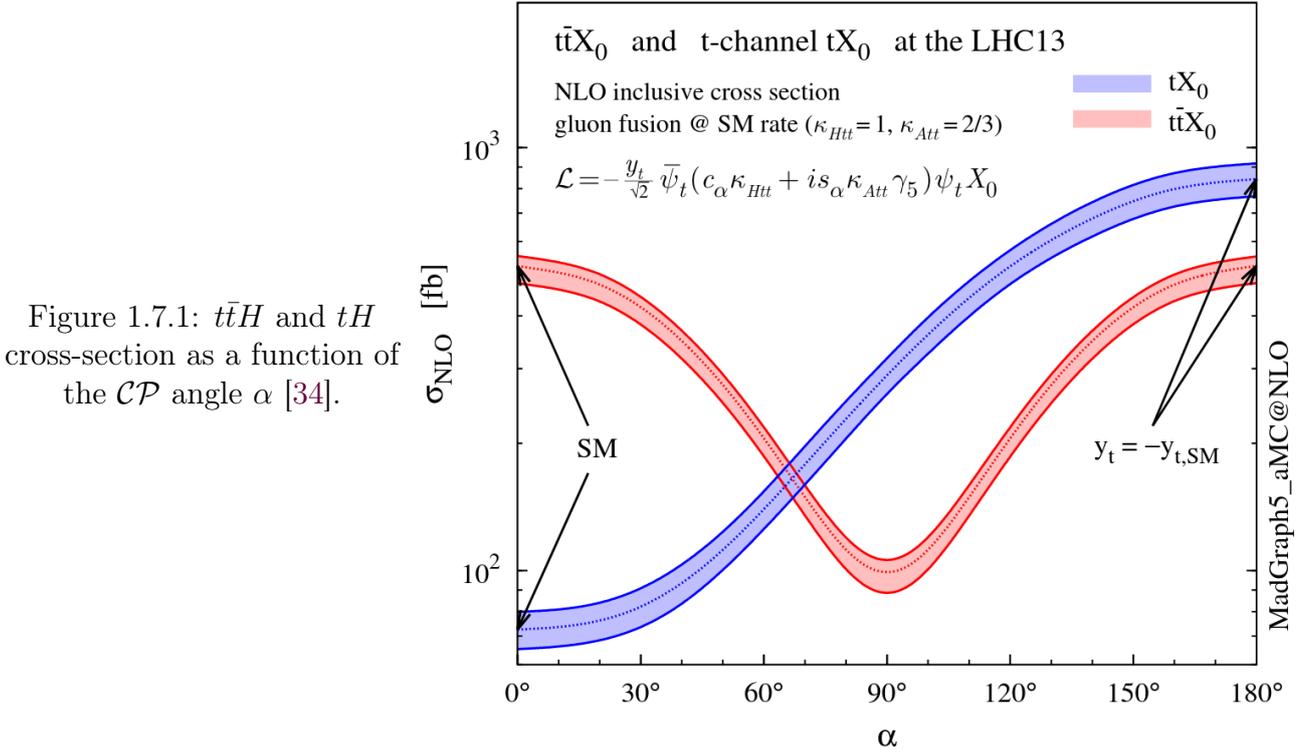


Figure 1.7.1:  $t\bar{t}H$  and  $tH$  cross-section as a function of the  $\mathcal{CP}$  angle  $\alpha$  [34].

$t$ -channel cross-section has a component that depends linearly on the top Yukawa coupling. Thus the  $tH$  production cross-section is sensitive to the sign of the top Yukawa coupling, while  $t\bar{t}H$  is more sensitive to the amplitude. The dependency of the NLO cross sections for  $t\bar{t}X_0$  and  $t$ -channel  $tX_0$  production at  $\sqrt{s} = 13$  TeV on the  $\mathcal{CP}$ -mixing angle  $\alpha$  is illustrated on the Figure 1.7.1. Here,  $X_0$  stands for a boson with  $J^P = 0^+$ .

One of the approaches to measure the possible  $\mathcal{CP}$ -odd component of the top Yukawa coupling is by using a parameterization of an effective model. Effective field theories (EFTs) [35] provide a consistent set of perturbations of the SM Lagrangian under the assumption that any new physics occurs at energy levels much higher than those the LHC can reach and that there are no new light particles. EFTs modify the couplings of the SM Lagrangian by adding higher-dimensional operators, allowing the symmetries of the SM to be mostly preserved. Within the Higgs Effective Lagrangian (HEL) model [36, 37], the SM Lagrangian is extended by adding 39 flavor-independent dimension-6 operators, denoted as  $\mathcal{O}_j$ :

$$\mathcal{L}_{HEL} = \mathcal{L}_{SM} + \sum_j \mathcal{O}_j \frac{f_j}{\Lambda^2}. \quad (1.7.1)$$

Evidence of new physics would appear as non-zero values in the Wilson coefficients,  $\frac{f_j}{\Lambda^2}$ . The Lagrangian describing the interaction between the Higgs boson (represented by field  $\phi$ ) and the top quark ( $\psi_t$ ) can then be represented as:

$$\mathcal{L}_{t\bar{t}H} = -k'_t y_t \phi \bar{\psi}_t (\cos \alpha + i \gamma_5 \sin \alpha) \psi_t. \quad (1.7.2)$$

Here  $y_t$  is the SM top Yukawa coupling strength, adjusted by the coupling modifier  $k'_t$ , and  $\alpha$  is the  $\mathcal{CP}$ -mixing angle. The SM setting is defined by  $k'_t = 1$  and  $\alpha = 0$ . The term proportional to  $\gamma^5$

in the Lagrangian corresponds to the possible  $\mathcal{CP}$ -odd pseudoscalar component of the coupling. A measured value of  $\alpha$  different from zero would produce an admixture with a pseudoscalar coupling ( $J^{CP} = 0^{+-}$ ) and would change the differential cross-section of the  $t\bar{t}H$  process compared to the SM expectation. A variation of  $k'_t$  would change the total cross-section.

The introduced Lagrangian modification impacts the computation of matrix element for the  $t\bar{t}H$  process. The main diagram contributing to the  $t\bar{t}H$  production is shown in Figure 1.7.2. The

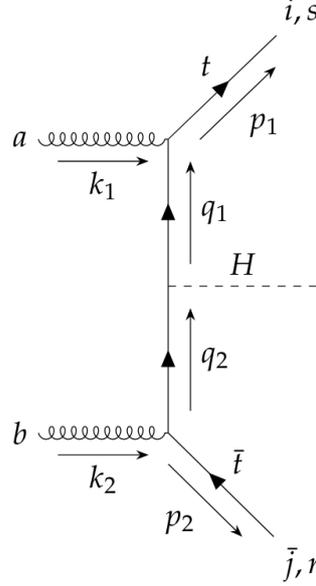


Figure 1.7.2: Main diagram for  $t\bar{t}H$  production

matrix element for this diagram can be written using the Feynman rules:

$$\mathcal{M} = \sum_l \bar{u}^s(p_1) \left(-i\frac{g_s}{2}\right) \lambda_{ik}^a \not{\epsilon}^\sigma(k_1) i\delta_{lk} \frac{q_1 + m}{q_1^2 - m^2} \frac{y_t}{\sqrt{2}} e^{i\alpha\gamma^5} i\delta_{lp} \frac{q_2 + m}{q_2^2 - m^2} \left(-i\frac{g_s}{2}\right) \lambda_{jp}^b \not{\epsilon}^\rho(k_2) v^r(p_2), \quad (1.7.3)$$

where  $\not{\epsilon} = \gamma^\mu a_\mu$ ,  $m$  refers to the mass of the top quark,  $g_s$  is the strong coupling constant,  $\lambda^a$  are the Gell-Mann matrices, and  $\epsilon$  are the gluon polarization vectors. The sum over  $l$  accounts for the colour of the virtual top quark. The difference with respect to the SM matrix element lies in the term corresponding to the top-Higgs interaction,  $\frac{y_t}{\sqrt{2}} e^{i\alpha\gamma^5}$ , that contains the  $\mathcal{CP}$ -mixing angle in our case.

By denoting

$$A = \frac{g_s^2 y_t}{4\sqrt{2}} \frac{1}{q_1^2 - m^2} \frac{1}{q_2^2 - m^2}, \quad \Gamma^{(1)\mu} = \gamma^\mu (q_1 + m), \quad \Gamma^{(2)\nu} = (q_2 + m) \gamma^\nu, \quad (1.7.4)$$

the matrix element expression becomes

$$\mathcal{M} = A \sum_l \bar{u}^s(p_1) \lambda_{il}^a \not{\epsilon}_\mu^\rho(k_1) \Gamma^{(1)\mu} e^{i\alpha\gamma^5} \Gamma^{(2)\nu} \lambda_{jl}^b \not{\epsilon}_\nu^\rho(k_2) v^r(p_2). \quad (1.7.5)$$

The squared matrix element, after summing over every possible outgoing colour of the  $t\bar{t}$  pair

and its helicity states, gluon colour and polarization, photon polarization, takes the form

$$\langle |\mathcal{M}|^2 \rangle = \frac{2|A|^2}{3} \text{Tr} \left[ (\not{p}_1 + m) \Gamma^{(1)\mu} e^{i\alpha\gamma^5} \Gamma^{(2)\nu} (\not{p}_2 - m) \Gamma_\nu^{(2)} e^{i\alpha\gamma^5} \Gamma_\mu^{(1)} \right]. \quad (1.7.6)$$

Considering pure  $\mathcal{CP}$ -even and  $\mathcal{CP}$ -odd cases, one can write the corresponding squared matrix elements:

$$\langle |\mathcal{M}|^2 \rangle_{\mathcal{CP}\text{-odd}} = \frac{2|A|^2}{3} \text{Tr} \left[ (\not{p}_1 + m) \Gamma^{(1)\mu} (-\not{q}_2 + m) \gamma^\nu (\not{p}_2 + m) (\not{q}_2 - m) \gamma_\nu \Gamma_\mu^{(1)} \right], \quad (1.7.7)$$

$$\langle |\mathcal{M}|^2 \rangle_{\mathcal{CP}\text{-even}} = \frac{2|A|^2}{3} \text{Tr} \left[ (\not{p}_1 + m) \Gamma^{(1)\mu} (\not{q}_2 + m) \gamma^\nu (\not{p}_2 - m) (\not{q}_2 + m) \gamma_\nu \Gamma_\mu^{(1)} \right], \quad (1.7.8)$$

implying the different  $\mathcal{CP}$ -even and  $\mathcal{CP}$ -odd cross-sections.

## 1.8 Simplified Template Cross-section framework

The Simplified Template Cross Sections (STXS) framework has been introduced to enhance the precision measurements of the Higgs production signal strength [23]. STXS are defined as physical cross-sections in mutually exclusive regions of phase space, referred to as STXS bins. This framework serves two main purposes: providing more fine-grained measurements for Higgs production modes in different kinematic regions, and reducing theoretical uncertainties inherent in the measurements. Measurements are unfolded to the STXS bins, allowing for a global combination of results from different analyses and experiments. The STXS measurements, both individual and combined, serve as inputs for further interpretations within and beyond the SM.

In the context of  $t\bar{t}H$  production measurements, the STXS framework is utilized to determine the cross-section as a function of the Higgs transverse momentum  $p_T^H$ . The bins are defined as following:  $p_0 \in [0, 60]$  GeV,  $p_1 \in [60, 120]$  GeV,  $p_2 \in [120, 200]$  GeV,  $p_3 \in [200, 300]$  GeV,  $p_4 \in [300, 450]$  GeV, and  $p_5 \in [450, +\infty]$  GeV. Given that  $p_T^H$  is sensitive to  $\mathcal{CP}$ -violation [38], the corresponding STXS bins are also instrumental in probing the  $\mathcal{CP}$  structure of the Higgs boson.

## 1.9 $t\bar{t}H$ measurement at the LHC

Several  $t\bar{t}H$  measurements were performed with both ATLAS and CMS experiments using Run-1 and Run-2 data [39, 40]. The production cross-section for a 125 GeV Higgs boson associated with pair of top quarks in  $pp$  collisions at  $\sqrt{s} = 8$  TeV is about 130 fb, while at  $\sqrt{s} = 13$  TeV it reaches 500 fb.

The  $t\bar{t}H$  analyses can be grouped into classes based on the Higgs decays, while considering most of the decay final states of the top quarks: fully hadronic, semi-leptonic and dilepton decay final states. The  $H \rightarrow \gamma\gamma$  [41, 42] and  $H \rightarrow 4l$  [43, 44] both have very clean signatures, but limited statistics. The opposite is observed for  $H \rightarrow b\bar{b}$  [45, 46]: this search is complicated due to the large backgrounds, both physical and combinatorial, but has a high branching ratio (see Figure 4.0.2). The last class, so-called "multilepton", considers the  $W^+W^-$ ,  $\tau^+\tau^-$  (decaying leptonically) and  $ZZ^*$  Higgs final states inclusively in multilepton event topologies. The measurements are listed

in Table 1.9.1.

$t\bar{t}H$	ATLAS Run-1	CMS Run-1	ATLAS Run-2	CMS Run-2
$H \rightarrow \gamma\gamma$	$1.3^{+2.6+2.5}_{-1.7-1.7}$	$1.2^{+2.5+2.6}_{-1.7-1.8}$	$1.38^{+0.33+0.26}_{-0.31-0.18}$	$2.27^{+0.86}_{-0.74}$
$H \rightarrow 4\ell$			$1.2^{+1.4}_{-0.8}$	$1.05^{+0.19}_{-0.17}$
$H \rightarrow WW/\tau\tau/ZZ$	$1.4 \pm 0.6 \pm 1.0$	$3.3 \pm 1.4$	$1.56^{+0.30+0.30}_{-0.29-0.27}$	$0.92^{+0.26}_{-0.23}$
$H \rightarrow b\bar{b}$	$1.4 \pm 1.0 \pm 0.6$	$1.6^{+1.6}_{-1.5}$	$0.81 \pm 0.11^{+0.20}_{-0.16}$	$0.33 \pm 0.16 \pm 0.21$
Combination	$1.7 \pm 0.5 \pm 0.8$	$2.6^{+1.0}_{-0.9}$	$0.97 \pm 0.14^{+0.14}_{-0.13}$	$0.94 \pm 0.20^{+0.09}_{-0.14}$

Table 1.9.1: Summary of the results for the measured cross-section measurements of Higgs boson in association with top quark pair by ATLAS and CMS. The results are given in terms of the measured signal strength [1].

Both the CMS and ATLAS experiments utilize a multilepton approach, incorporating various channels including electron, muon, and  $\tau$  lepton in the final states, to examine the  $W^+W^-$ ,  $\tau^+\tau^-$  and  $ZZ^*$  Higgs decay modes [24, 47]. While CMS has analyzed the entire Run-2 dataset, ATLAS has analyzed only a portion of it. In a preliminary analysis using a subset of the Run-2 dataset (approximately  $80 \text{ fb}^{-1}$  integrated luminosity), ATLAS noted discrepancies in the normalization of the  $t\bar{t}W$  background, which exceeded SM predictions by factors ranging from 1.3 to 1.7. Additionally, issues with modeling were observed in regions where the  $t\bar{t}W$  process dominates. In this thesis the result of the analysis in the multilepton final state with full Run-2 dataset is presented. The differential cross-section measurement within STXS framework is performed in  $H \rightarrow b\bar{b}$  by both ATLAS and CMS. STXS measurements were never performed for the multilepton channel yet.

The  $\mathcal{CP}$  properties of the coupling between the Higgs boson and the top quark were also investigated in  $H \rightarrow b\bar{b}$  [45, 46] and  $H \rightarrow \gamma\gamma$  [41, 42] by the ATLAS and CMS experiments, and in  $t\bar{t}H$  in multilepton state by CMS [48]. The  $H \rightarrow \gamma\gamma$   $\mathcal{CP}$ -measurement performed using the parametrization described in Section 1.7 resulted in a pure  $\mathcal{CP}$ -odd coupling exclusion at  $3.9\sigma$ , and a limit  $|\alpha| > 43^\circ$  at 95% confidence level. A comparable study from the CMS experiment excluded  $\alpha = 90^\circ$  at  $3.2\sigma$ . In the  $H \rightarrow b\bar{b}$  ATLAS measured the mixing angle between  $\mathcal{CP}$ -even and  $\mathcal{CP}$ -odd couplings to be  $\alpha = 11^{+52^\circ}_{-73^\circ}$ , with  $\mathcal{CP}$ -odd exclusion significance observed at  $1.2\sigma$ . The  $\mathcal{CP}$ -odd  $\mathcal{CP}$ -even discrimination is based on  $\mathcal{CP}$ -sensitive variables defined using momentum three-vectors of the two top quarks.

## 2 Experimental setup

The  $t\bar{t}H$  multilepton measurement presented in this work was conducted using the ATLAS detector at the Large Hadron Collider (LHC). The LHC, a circular particle accelerator, generates high-energy proton-proton and heavy ion collisions at a high rate, offering an environment to probe rare processes and investigating BSM scenarios. Section 2.1 provides an overview of the key characteristics of the LHC. Positioned at one of the four collision points of the LHC, the ATLAS experiment is detailed in Section 2.2, including its sub-detectors. Section 2.3 outlines the simulation of collision data used for physics analyses within the ATLAS experiment. The procedures for reconstructing objects generated in these collisions are elaborated in Section 2.4.

### 2.1 The Large Hadron Collider

In collider experiments, typically, two beams of particles are first accelerated to high energies  $E_1$  and  $E_2$  and then collide under a small or zero crossing angle, providing access to a centre-of-mass energy of  $\sqrt{s} = 2\sqrt{E_1 E_2}$ . The LHC was designed to study physics at the TeV energy scale. The technical choice of using proton collisions was made to ensure high energies at collision points. In electron colliders, reaching high targeted energy is complicated because of synchrotron radiation [49]. The process of interest,  $e^+e^- \rightarrow t\bar{t}H$ , would, for example, become important for  $\sqrt{s} \geq 500$  GeV [50] that is much higher than the centre-of-mass energy of the electron-positron colliders built so far. For example, the Large Electron-Positron Collider (LEP) operated at a maximum centre-of-mass energy of around 200 GeV. Currently the  $e^+e^-$  colliders are mainly target to study  $b$ -physics and run at much smaller centre-of-mass energy. A proton-antiproton collider would not allow reaching the desired collision rate due to the difficulty of producing antiprotons [51]. Proton-proton collisions create a setup satisfying the above criteria but also raise some challenges discussed in Section 2.3.1. Apart from proton-proton collisions, the LHC can also perform heavy ion collisions, implying the creation of quark-gluon plasma.

The LHC is a circular 26.7 km particle accelerator. The LHC accelerator design is detailed in the Section 2.1.1. The collisions take place at specific points where particle detectors are positioned. The main LHC experiments are discussed in Section 2.1.1.3.

The LHC is located at the French-Swiss border and is operated by the European Organization for Nuclear Research (CERN). The collider is around 100 m underground, occupying the same tunnel previously used for the Large Electron Positron collider (LEP). Since its construction (1998 to 2008) and circulation of the first beam in 2008, the LHC performed several operational data-taking runs, separated by hardware upgrades (see Figure 2.1.1). The Run-1 [52] marked by the Higgs boson discovery lasted from 2010 to 2013, starting with the  $pp$   $\sqrt{s} = 7$  TeV and boosted to  $\sqrt{s} = 8$  TeV in 2012. During the following two years, 2013-2015, the LHC ring and detectors underwent upgrades during the Long Shutdown 1 (LS1), after which the Run-2 [53] between 2015-2018 operated at higher luminosity, and centre-of-mass energy of 13 TeV. The  $t\bar{t}H$  measurement

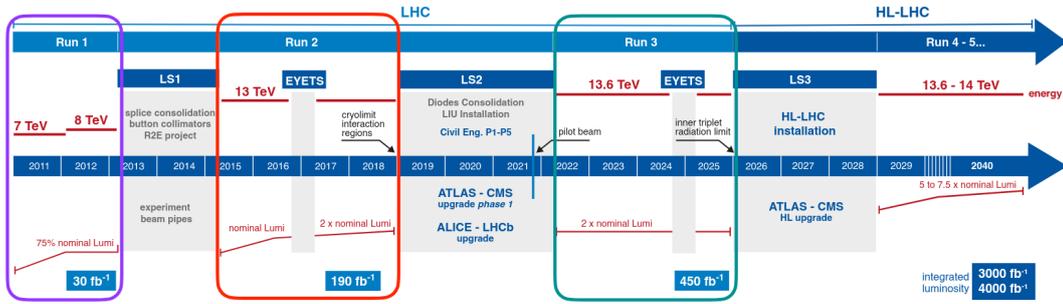


Figure 2.1.1: LHC runs and upgrades [55].

Parameter	Design	2015	2016	2017	2018
Centre-of-mass energy ( $\sqrt{s}$ ) [GeV]	14000	13000	13000	13000	13000
Number of colliding bunches ( $n_b$ )	2808	1000	2208	1866	2556
Bunch charge ( $N_p$ ) [1e11 charges]	1.5	1.0	1.2	1.2	1.1
Revolution frequency ( $f_{\text{rev}}$ ) [Hz]	11245.5	11245.5	11245.5	11245.5	11245.5
Normalised beam emittance ( $\epsilon_n$ ) [ $\mu\text{m rad}$ ]	3.75	3.47	3.72	2.22	1.99
$\beta^*$ [m]	0.55	0.80	0.40	0.37	0.30
Peak instantaneous luminosity [ $1\text{e}30 \text{ cm}^{-2}\text{s}^{-1}$ ]	10000	6309.06	13751.86	20571.53	21700.84
Mean interactions per crossing ( $\langle\mu\rangle$ )	23	13.4	25.1	37.8	36.1
Delivered integrated luminosity [ $\text{fb}^{-1}$ ]	–	4.2	38.5	50.2	63.3

Table 2.1.1: LHC operating parameters for the design conditions and through Run-2 [56].

presented in this work is based on data recorded during Run-2. Between 2019 and 2022, the LS2 upgrade was performed, followed by the start of the Run-3 [54] collecting data at  $\sqrt{s} = 13.6$  TeV. The next upgrade, the High-Luminosity LHC (HL-LHC), will happen between 2026-2028 during the LS3.

## 2.1.1 LHC setup

### 2.1.1.1 Basic collider parameters

A particle collider is characterized by numerous parameters, a few mains of which are discussed in the following with a focus on the Run-2 LHC setup. The LHC operating parameters for the design conditions and through Run-2 are summarized in Table 2.1.1.

Apart from the already mentioned centre-of-mass energy, another key parameter is instantaneous luminosity. It gives the number of collisions per second per unit area:

$$L = \frac{n_b \cdot f_{\text{rev}} \cdot N_p^2 \cdot \gamma_r}{4 \cdot \pi \cdot \sigma_x \sigma_y} \cdot F(\theta_c, \sigma_x, \sigma_z). \quad (2.1.1)$$

This formula assumes a Gaussian beam distribution. It is defined using the following parameters:

- $n_b$  is the number of colliding bunches. For Run-2, the typical number of bunches in each LHC ring was 2544.
- $f_{\text{rev}}$  stands for the per-bunch revolution frequency. For the LHC, it is equal to 11245.5 Hz, based on the LHC ring circumference and the value of the speed of the light.
- $N_p$  is the number of protons per bunch (around  $10^{11}$ ), referred to as the beam intensity.

- $\gamma_r$  is the relativistic gamma factor for protons.
- $\sigma_{x(y)}$  is the transverse beam dispersion corresponding to the beamspot size. At LHC in 2016 it was measured around  $10 \mu m$ .
- $F(\theta_c, \sigma_x, \sigma_z)$  is the factor accounting for the reduction in the luminosity due to non-zero beam crossing angle  $\theta_c$ .

The number of produced events of a given process with a cross-section  $\sigma$  is obtained by the integrated luminosity  $L$ :

$$N_{\text{events}}^{\text{obs}} = \sigma \cdot \epsilon \int dt L, \quad (2.1.2)$$

where  $\epsilon$  is the detection efficiency combining trigger, offline reconstruction and analysis selection efficiencies.

A high luminosity provides the opportunity to observe rare processes having small cross-sections. However, high instantaneous luminosity reached with increasing the density of beams leads to the presence of pile-up events happening at the same time as the process of interest or at the adjacent bunch crossing. The average value of the number of pile-up events at a given luminosity is defined by:

$$\langle \mu \rangle = \frac{\sigma_{\text{inel}} \cdot L}{f_{\text{rev}} \cdot n_b}, \quad (2.1.3)$$

where  $\sigma_{\text{inel}}$  is the inelastic cross-section.

### 2.1.1.2 The accelerator chain

The CERN's accelerator complex is represented in Figure 2.1.2. Protons are supplied to the LHC by an injector chain. First, the protons are sourced from gaseous hydrogen ionised by a strong electric field. Before reaching the LHC, protons energy gets increased by a chain of accelerators: a linear accelerator (LINAC), Proton Synchrotron Booster (PSB), Proton Synchrotron (PS) and Super Proton Synchrotron (SPS). Within the LHC setup at the time of collecting Run-2 data (before LS2), protons are first accelerated up to 50 MeV by LINAC-2 using radio-frequency cavities. The following acceleration happens via circular accelerators PSB (up to  $\sim 1.4$  GeV), PS (up to  $\sim 25$  GeV) and SPS (up to  $\sim 450$  GeV) and lasts around 4.5 minutes in total. After going through auxiliary accelerators, protons get transferred to the LHC, where they experience the final acceleration and stabilization, reaching the energy of 6.5 TeV in around 20 minutes. The LHC is a two-ring accelerator and collider that gets filled with two proton beams running in clockwise and anticlockwise directions in separate beam pipes. There are eight straight sections and arcs in the LHC ring. The beam pipes have separate magnetic fields and vacuum chambers in the main arcs, and have common straight sections at the collision points where detectors are located. The straight section's main purpose is the protons' acceleration that, analogically to LINAC, is reached via radiofrequency chambers (RFs). The 16 superconducting RFs generate an electric field of 5 MV/m with 400 MHz oscillations. The frequency is adjusted in a way to be an integer multiple of the  $f_{\text{rev}}$ , so that the proton experience only the accelerating voltage. Once a proton's revolution frequency corresponds exactly to the RF frequency, no additional acceleration is possible, and such a proton is called a synchronous particle. Other protons group around

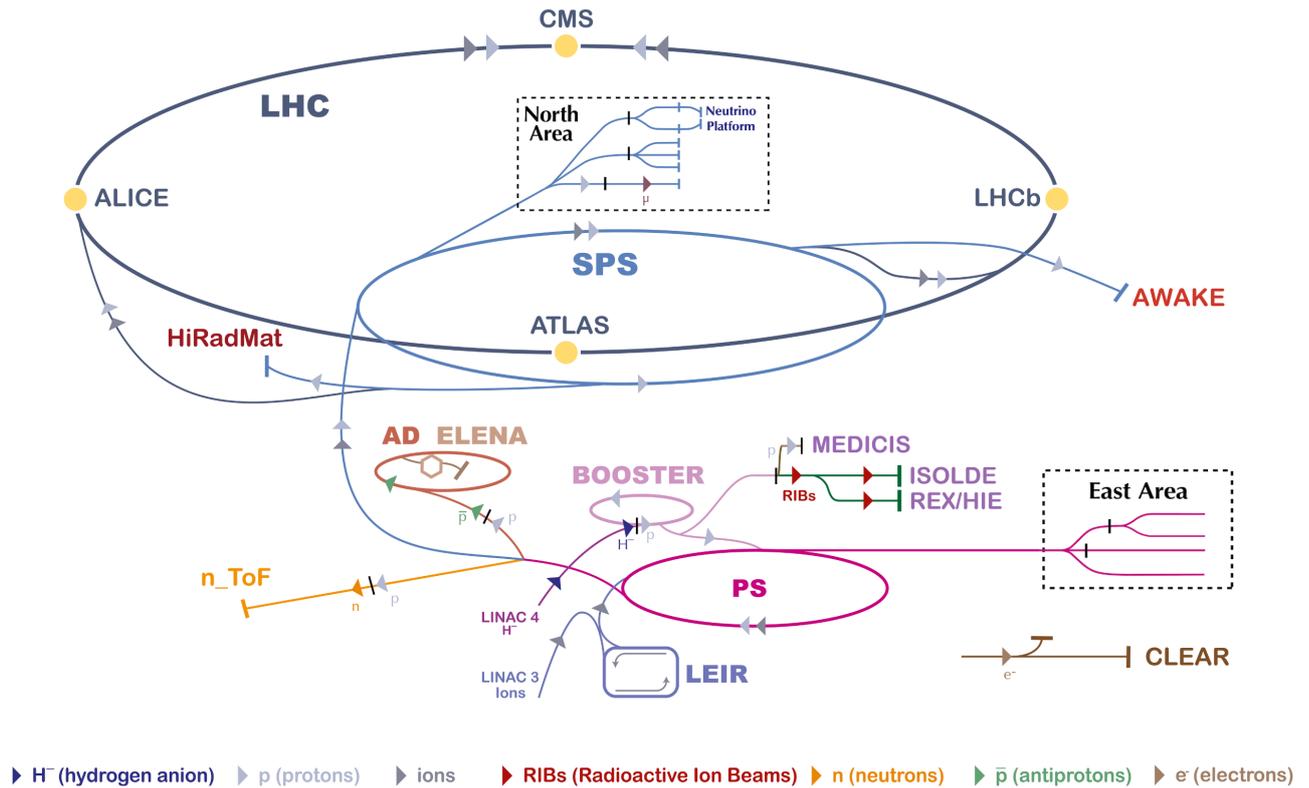


Figure 2.1.2: CERN's accelerator complex [57].

synchronous particles longitudinally oscillating, forming a bunch. There are 2808 bunches in the LHC beams, each containing approximately  $10^{11}$  protons. The distance between two consecutive bunches is called bunch spacing and is amount to to 50 or 25 ns in Run-2, depending on the operating year.

Within the beam pipes of the LHC, proton beams are guided along curved trajectories by an array of superconducting magnets, cooled to a temperature of 1.9 K using superfluid Helium. Two primary types of magnets are employed: dipoles, responsible for bending the proton beams within the arcs, and quadrupoles, which focus the beams while ensuring their separation from the pipe walls. The LHC comprises a total of 1232 dipole magnets, each approximately 15 meters in length, generating a magnetic field of 8.3 T. In contrast, the 858 quadrupoles deliver nominal gradients of 223 T/m and 241 T/m, strategically positioned in pairs along the LHC ring to regulate both beam width and height. Additional magnetic multipoles, including sextupoles, octupoles, and decapoles, are also integrated into the LHC system to mitigate divergence-induced effects leading to imperfections in the proton beams.

The operation of the LHC is organized into fills: a fill starts upon injecting the desired number of bunches into the machine and concludes with beam dumping. When beam conditions reach a state suitable for physics data collection, stable beams are declared, initiating data acquisition for physics analyses by the experiments. Throughout a fill, the instantaneous luminosity declines as protons are utilized in collisions or lost outside the interaction regions. After approximately 10 – 20 hours, the collision rate diminishes to a level where it becomes more efficient to dump the beam and dedicate a few hours to reinjecting a new beam.

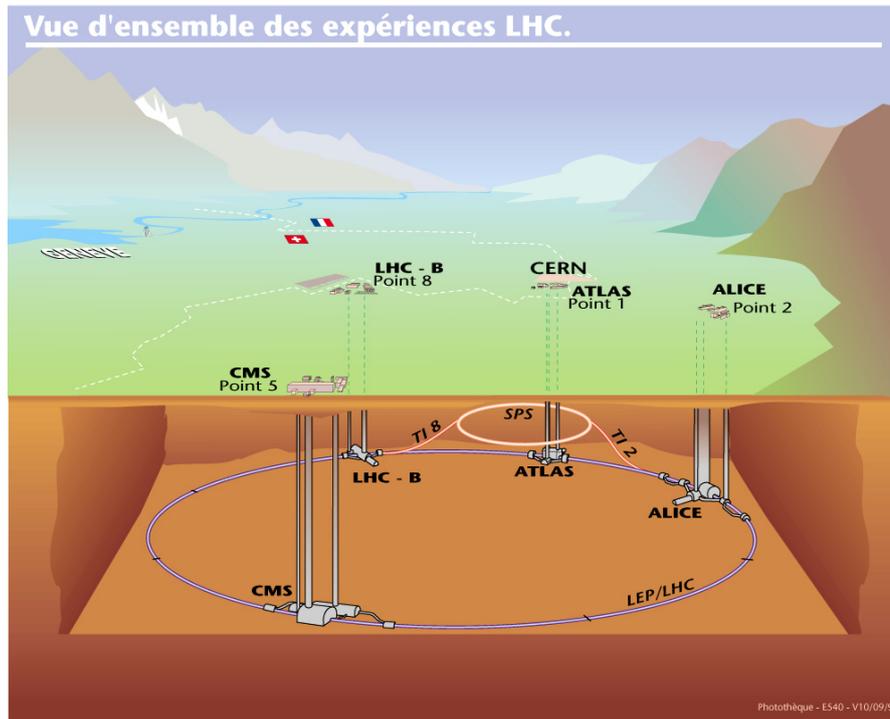


Figure 2.1.3: The LHC experiments locations [65].

### 2.1.1.3 LHC experiments

The LHC hosts several experiments that have a rich physics program: precise measurements of SM parameters, search for new physics, studying the nature of dark matter and matter-antimatter asymmetry, etc. In total, there are seven experiments installed at the LHC: A Large Ion Collider Experiment (ALICE) [58], A Toroidal LHC ApparatuS (ATLAS) [59], the Compact Muon Solenoid (CMS) [60], the Large Hadron Collider beauty (LHCb) experiment [61], the Large Hadron Collider forward (LHCf) experiment [62], the TOTal Elastic and diffractive cross section Measurement (TOTEM) experiment [63] and Monopole and Exotics Detector at the LHC (MoEDAL) [64]. The four biggest experiments, ALICE, ATLAS, CMS and LHCb, are located around the four collision points (see Figure 2.1.3), while others are located near the biggest ones:

- ATLAS and CMS experiments are general-purpose detectors with a broad physics program that were designed and optimized to study proton-proton collisions over a wide energy spectrum. Both detectors employ distinct technologies for tracking and calorimetry. Using two independent experiments aiming at the same physics analyses proves advantageous, as it facilitates cross-checking their findings. Additionally, combining the datasets that the two detectors recorded enhances the analysis's precision.
- LHCb experiment targets studying matter-antimatter asymmetry via  $b$  and  $c$  quarks.
- ALICE experiment is dedicated to quark-gluon plasma studies via high-energy heavy-ion collisions.

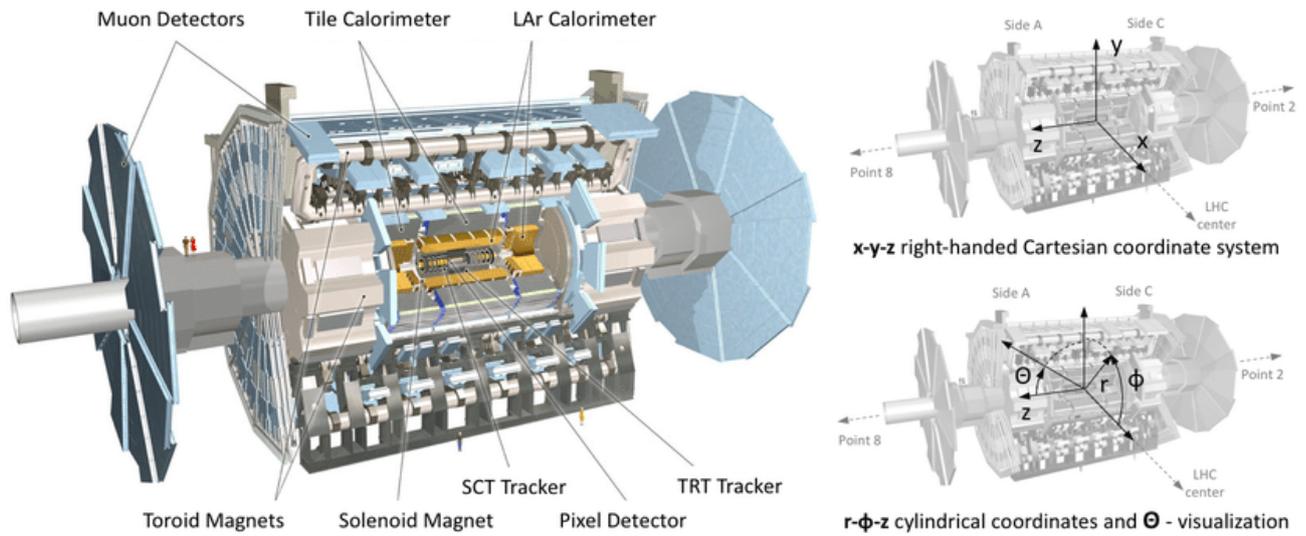


Figure 2.2.1: A schematic view of the full ATLAS detector [66].

- TOTEM aims for a regime inaccessible to other LHC experiments: the protons that emerge at narrow angles relative to the beam pipe.
- LHCf comprises two detectors on opposite sides of the ATLAS experiment, positioned 140 meters from the collision point. It utilizes particles emitted in the forward direction to simulate cosmic rays under controlled laboratory conditions.
- MoEDAL, located around the same interaction point as LHCb, conducts direct searches for magnetic monopoles, which are hypothetical particles possessing magnetic charge.

## 2.2 The ATLAS detector

The ATLAS experiment was designed to allow high-resolution measurements of LHC proton-proton and heavy-ion collisions. A computer-generated view of the ATLAS detector is depicted in Figure 2.2.1 (left), illustrating its concentric cylindrical structure composed of multiple sub-detector layers. Each of the sub-detectors is designed to measure distinct properties of particles or objects generated from collisions.

The ATLAS sub-detectors can be categorized into three main groups: the inner tracker, calorimeters, and the muon spectrometer. The inner detector (see Section 2.2.3), situated closest to the beam pipe within the solenoid magnet, comprises silicon pixel, silicon microstrip, and straw tube tracking detectors, facilitating high-precision reconstruction of particle trajectories and momentum measurements. Electromagnetic and hadronic calorimeters (see Section 2.2.4), located outside the solenoid magnet, absorb energy from photons, electrons, and hadrons. The muon spectrometer (see Section 2.2.5), situated outside the calorimeters and supported by a barrel toroidal magnet (see Section 2.2.2), utilizes gaseous tracking detectors to measure particles trajectories and momenta. Interesting events are detected by trigger systems that lead to detector measurements being read out and saved for further processing (see Section 2.2.5). The stored data are then used to reconstruct objects such as electrons, muons, and jets for physics analyses (Section 2.4).

Detector component	Required resolution	$\eta$ coverage	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$	$\pm 2.5$	
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	$\pm 3.2$	$\pm 2.5$
Hadronic calorimetry (jets)			
barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	$\pm 3.2$	$\pm 3.2$
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 <  \eta  < 4.9$	$3.1 <  \eta  < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T=10\%$ at $p_T = 1$ TeV	$\pm 2.7$	$\pm 2.4$

Table 2.2.1: General performance goals of the ATLAS detector. Note that, for high- $p_T$  muons, the muon-spectrometer performance is independent of the inner-detector system. The units for  $E$  and  $p_T$  are in GeV [67].

## 2.2.1 ATLAS coordinate system and detector requirements

The ATLAS experiment uses a right-handed coordinate system along with a spherical polar one, with the origin set at the nominal interaction point (see Figure 2.2.1 (right)). Within the cartesian system, the longitudinal  $z$ -axis aligns with the beam direction, the positive  $x$ -axis points from the origin toward the center of the LHC ring, and the positive  $y$ -axis points upward toward the ground surface. The polar system defines the polar angle  $\theta$  measured from the positive  $z$ -axis, the azimuthal angle  $\phi$  measured from the positive  $x$ -axis, and the radial coordinate  $R = \sqrt{x^2 + y^2}$  representing the distance from the beamline. The spherical coordinate system naturally complements the cylindrical geometry of the detector. It allows for the definition of several variables that are invariant under boosts along the  $z$ -axis, given that the  $z$ -component of the initial parton momentum is unknown. For example, the rapidity  $y$  (or pseudo-rapidity  $\eta$  in the ultra-relativistic approximation) replaces the distance in polar angle  $\theta$ :

$$y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right) \xrightarrow{E \gg mc^2} \eta = -\ln \left( \tan \frac{\theta}{2} \right). \quad (2.2.1)$$

Here  $E$  is the particle energy and  $p_z$  the  $z$  component of its momentum vector. The transverse component of the particle momentum  $p$  is referred to as  $p_T = p \sin\theta$ . Parameter  $\Delta\eta$  is invariant under Lorentz boost along the beamline, and with such a replacement the radial distance becomes  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ .

The key design considerations for ATLAS include the necessity for precise electromagnetic calorimetry for the identification of electrons and photons, complemented by hadronic calorimetry for accurate measurements of jets and missing transverse energy. A high-resolution muon spectrometer ensures precise determination of muon energy and trajectory, while high-efficiency tracking enables particle identification and event reconstruction. The general performances of the ATLAS detector are summarised in Table 2.2.1.

## 2.2.2 The magnet system

The ATLAS detector's magnet system provides the magnetic field necessary for precise measurement of charged particle momentum and particle identification. It consists of four superconducting magnets: the Central Solenoid, the Barrel Toroid, and two End-Cap Toroids [68].

- The central solenoid provides an axial magnetic field of 2 T within the inner detector and is

mounted in the vacuum vessel of the Liquid Argon calorimeter.

- The barrel toroid comprises eight flat coils in a racetrack configuration, each housed in individual cryostats, producing a magnetic field of 3.9 T for the muon spectrometer.
- The end-cap toroids each consists of eight superconducting coils in a single cryostat, generating a 4.1 T magnetic field and can be retracted to allow access to the central detector.

The magnets use a composite conductor made of NbTi/Cu cable within an aluminum stabilizer and are cooled to 4.5 K using superfluid helium.

### 2.2.3 The inner detector

The Inner Detector (ID) is the innermost part of the ATLAS detector designed to reconstruct charged particle tracks, determine their momenta, and identify interaction vertices. The ID is surrounded by a central solenoid and comprises three subsystems: the Pixel detector and Semiconductor Tracker (SCT) near the interaction point and the Transition Radiation Tracker (TRT) further out. These subsystems are divided into two regions: the barrel section, which covers low pseudorapidities and consists of concentric cylinders around the beam axis, and the end-caps, positioned at high pseudorapidities, arranged in disks perpendicular to the beam axis. This configuration optimizes pseudorapidity coverage, reaching  $|\eta| = 2.5$  for the pixels and SCT systems and up to  $|\eta| = 2$  for the TRT. The measurement is possible for particles with  $p_T > 0.4$  GeV. Table 2.2.2 summarizes the specifications of the tracking system subcomponents. Hits here refer to the energy deposits that a particle leaves in active ID elements.

Detector	#channels	#hits per track	$R$ - $\phi$ resolution	$R$ - $z$ resolution
Pixel	80M	$\sim 3$	$< 10 \mu\text{m}$	$\sim 100 \mu\text{m}$
SCT	6.3M	$\sim 4$ (modules)	$17 \mu\text{m}$	$\sim 800 \mu\text{m}$
TRT	320k	$\sim 30$	$130 \mu\text{m}$	-

Table 2.2.2: Specifications of the Inner Detector.

#### 2.2.3.1 Pixel detector

The Pixel detector is the innermost component of the ATLAS ID, that provides precise tracking near the interaction point for accurate track reconstruction. The Run-2 Pixel detector consists of four barrel layers and six disk layers (see Figure 2.2.2 (*left*)). The outer barrel region contains three pixel layers at radii of 50.5, 88.5, and 122.5 mm, and built out of 1744 identical modules [69]. The innermost layer, the Insertable B-Layer (IBL), added during LS1, is situated at a radius of 33.25 mm in the barrel region. The Pixel detector employs silicon pixel sensors that generate electron-hole pairs when particles pass through. Each pixel sensor has around  $47 \cdot 10^3$  pixels bump-bonded to 16 frontend chips with 2880 readout channels each. These components, along with power cables and services, form a module (see Figure 2.2.2 (*right*)). Modules are mounted on barrel staves and disk sectors, ensuring detector hermeticity.

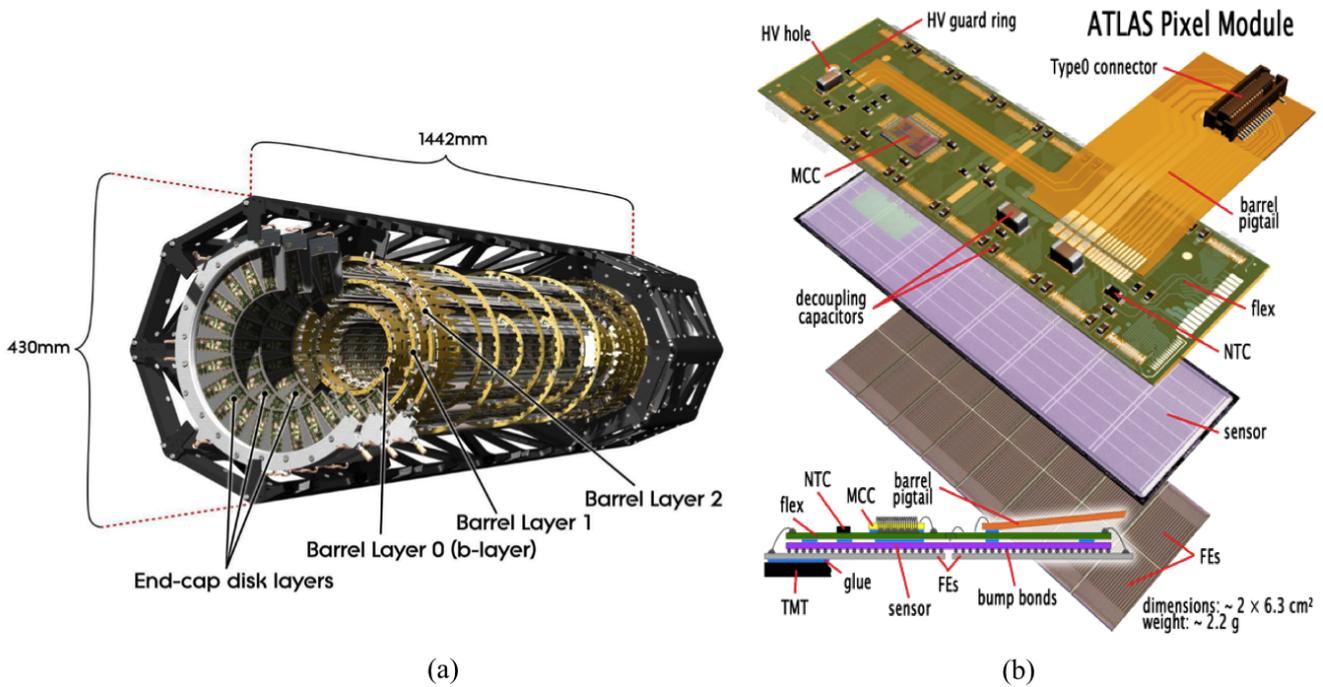


Figure 2.2.2: (a) Schematic drawing of the ATLAS Pixel Detector. (b) Assembly view and cross-section of an ATLAS Pixel Detector module [69].

### 2.2.3.2 Semiconductor Tracker

The SCT enhances the ID's tracking capabilities at radial distances between 299 mm and 560 mm [70]. The SCT is essential for tracking particles in the plane perpendicular to the beam, across a larger area than the Pixel detector. It consists of four layers of silicon microstrip sensors in the barrel and nine disks in each end-cap. Each barrel layer contains 12 rings of modules arranged along the  $z$ -axis, resulting in 2112 modules, while the end-caps consists of 1976 modules. Each module is made up of four rectangular silicon sensors. The SCT operates similarly to the pixel detector but uses narrow, long strips instead of small pixels. These strips are 12 cm in length and 80  $\mu\text{m}$  in width. The SCT spatial resolutions are 17  $\mu\text{m}$  ( $R - \phi$ ) and 580  $\mu\text{m}$  ( $z/R$ ) in both the barrel and end-cap regions.

### 2.2.3.3 Transition Radiation Tracker

The outermost component of the ID, the TRT, is a gaseous detector made up of straw tubes, each 4 mm in diameter and up to 150 cm in length [71]. These straw tubes contain gold-plated tungsten wires and are filled with a xenon-based gas mixture. As particles pass through the straws, they ionize the gas, creating on average 5–6 ionisation clusters per mm of path length. The free electrons drifting to the central wire create a detectable current along the wires. The TRT has a total of  $35 \cdot 10^4$  readout channels and contributes to electrons identification separating them based on their ionization level. The barrel section includes three concentric layers, spanning a radial range of 560 mm to 1080 mm and a longitudinal range of  $|z| < 712$  mm. Each end-cap section contains two independent wheels positioned perpendicular to the beam pipe, covering radial distances from 644 mm to 1004 mm and longitudinal distances from  $|z| = 827$  mm to  $|z| = 2744$  mm.

## 2.2.4 Calorimeters

The ATLAS calorimeter system is situated outside the ID and provides coverage up to  $|\eta| < 4.9$  with full  $\phi$  coverage around the beam axis. This system consists of three main components: the electromagnetic calorimeter, the hadronic calorimeter, and the forward calorimeter. It measures particle energy by absorbing particles and analyzes shower parameters for particle identification. The sampling calorimeter system uses alternating layers of active material and dense absorber material. When particles pass through the absorbers, they create electromagnetic or hadronic showers, which then ionize the active material in the electromagnetic calorimeter or produce scintillation photons in the hadronic one. The generated signal in the active medium is proportional to the deposited energy. Calorimeters are designed to be large enough to fully contain the showers, minimizing energy leakage into the muon spectrometer. Electromagnetic calorimeters use high atomic number materials characterized by the material radiation length  $X_0$ , while hadronic calorimeters are designed based on the nuclear interaction length  $\lambda$ . Muons pass through the calorimeters, neutrinos do not interact with the calorimeters, appearing as missing energy in the detector (see Section 2.4.5). The calorimeter energy resolution mentioned in Figure ?? can be written as

$$\frac{\sigma(E)}{E} = \frac{S}{\sqrt{E}} \otimes \frac{N}{E} \otimes C, \quad (2.2.2)$$

where  $\frac{S}{\sqrt{E}}$  is related to the fluctuations in shower development,  $\frac{N}{E}$  is coming from the electronic noises and pile-up, and  $C$  is the constant term addressing the nonuniformities effects.

### 2.2.4.1 Electromagnetic calorimeter

The Electromagnetic calorimeter (ECal) of the ATLAS detector is designed to measure the energy of electrons and photons with high precision. ECal is a sampling calorimeter that uses lead plates as absorbers and Kapton electrodes immersed in Liquid Argon (LAr) as active material. The calorimeter consists of a barrel section and two endcaps. Each component of the ECal is housed in its own cryostat to maintain the Argon in a liquid state at 89 K. The ECal employs an accordion geometry for the absorbers and electrodes to ensure full coverage in the  $\phi$  direction without gaps and a faster readout. Figure 2.2.3 illustrated the ECal setup (left) and a cut-out view of the calorimeter in the barrel including dimensions and coordinates (right). The barrel covers the region up to  $|\eta| = 1.475$ , while the endcaps consist of two wheels each, covering the pseudorapidity range of  $1.375 < |\eta| < 3.2$ . The barrel and outer endcap wheels in the region dedicated to precision physics ( $|\eta| < 2.5$ ) are longitudinally segmented into three layers: the strip layer, the middle layer, and the back layer. The strip layer is highly granular in the  $\eta$  direction, while the middle and back layers offer improved granularity in the  $\phi$  direction. The endcaps' inner wheels,  $|\eta| > 2.5$ , have two layers with coarser granularity in both  $\eta$  and  $\phi$ . The barrel calorimeter is located outside the solenoid magnet and has a minimal thickness of 22 radiation lengths  $X_0$ . The endcap calorimeters have a minimal depth of 24  $X_0$ .

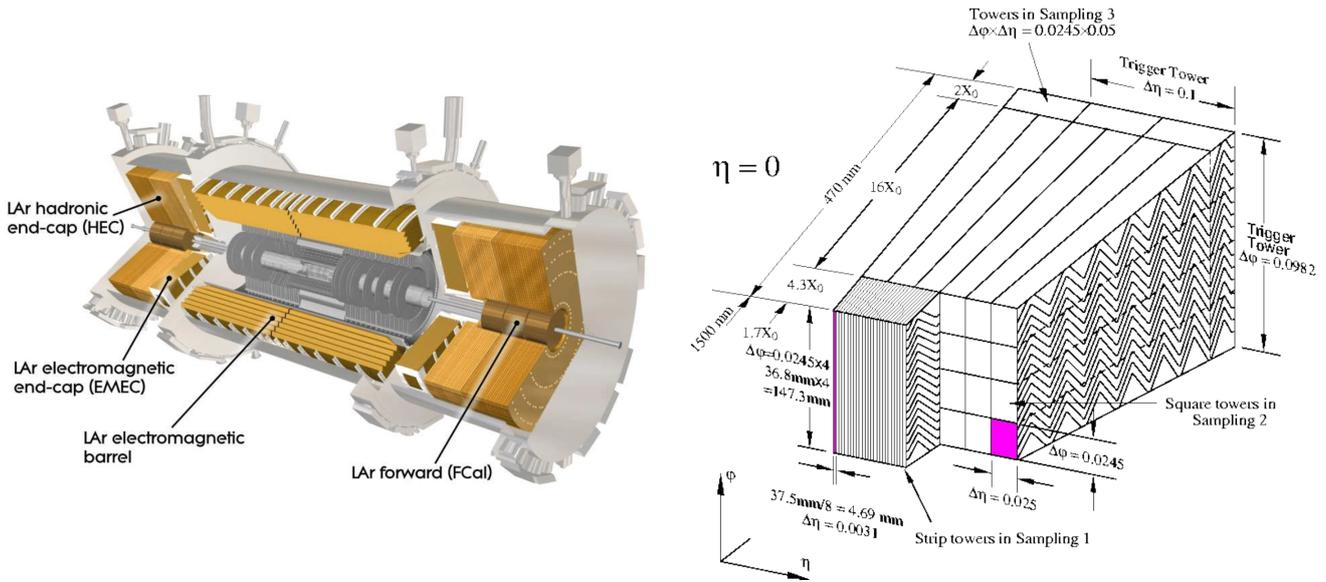


Figure 2.2.3: Schematic view of the different components of the ATLAS LAr calorimeter system (*left*), and a cut-out view of the calorimeter in the barrel including dimensions and coordinates (*right*) [72].

### 2.2.4.2 Hadronic calorimeter

The hadronic calorimeter system [73] is designed to measure the energy of hadronic showers, called jets. The barrel part consists of plastic polystyrene scintillating tiles as active material, interleaved with steel plate absorbers. It measures the energy deposition of charged and neutral hadrons. Charged particles passing through the scintillating tiles produce ultraviolet light, which is collected and converted into visible light via wavelength shifting fibers before reaching photomultiplier tubes (PMTs) that output an amplified electrical signal. The barrel is divided into three parts: a central part covering up to  $|\eta| = 1$ , and two extensions on each side covering the range of  $0.8 < |\eta| < 1.7$ . The hadronic endcap calorimeters use liquid argon (LAr) technology with flat copper plates as absorbers instead of lead. These endcaps are located behind the electromagnetic calorimeter endcaps and are housed in the same cryostats. Each endcap consists of two independent wheels, covering the pseudorapidity range of  $1.5 < |\eta| < 3.2$ .

### 2.2.4.3 Forward Calorimeter

Two forward calorimeters are positioned on each side of the ATLAS detector to measure forward jets and improve the estimation of missing energy. They also help reduce backgrounds reaching the muon system. These calorimeters cover the pseudorapidity range of  $3.1 < |\eta| < 4.9$ . Each of its endcaps is divided into three longitudinal modules. The first module, closest to the interaction point, uses copper as an absorber and is optimized for detecting electromagnetic showers. The other two modules use tungsten as an absorber and are optimized for measuring hadronic showers. All three modules are housed within the same LAr cryostat as the other endcap calorimeters.

### 2.2.5 Muon systems

The muon spectrometer (MS) [74], the outermost subdetector of the ATLAS experiment, is dedicated to detecting muons exiting the calorimeters, reconstructing their tracks and precisely measuring their momenta over the pseudorapidity range of  $|\eta| < 2.7$ . It also provides triggering capabilities for  $|\eta| < 2.4$ . The momentum resolution of the muon spectrometer can be expressed for a given  $\eta$  as a function of  $p_T$ :

$$\frac{\sigma(p)}{p} = \frac{p_0^{MS}}{p_T} \otimes p_1^{MS} \otimes p_2^{MS} \cdot p_T, \quad (2.2.3)$$

where  $p^{MS}$  are the coefficients related to energy loss in the calorimeter material ( $p_0^{MS}$ ), multiple scattering ( $p_1^{MS}$ ) and intrinsic resolution ( $p_2^{MS}$ ) terms [75]. The MS relies on the deflection of muon tracks in the magnetic field generated by the ATLAS toroidal magnets system (see Section 2.2.2) to measure the muon momentum. The MS is structured with three concentric layers in the barrel region and three endcap wheels (see Figure 2.2.4). It comprises four subsystems: Monitored Drift Tubes (MDT), Cathode Strip Chambers (CSC), Resistive Plate Chambers (RPC), and Thin Gap Chambers (TGC). MDT and CSC provide precision measurements, while RPC and TGC are primarily used for triggering. MDT chambers cover the entire  $|\eta| < 2.7$  range, except for the innermost endcap layer ( $2 < |\eta| < 2.7$ ), where CSC chambers are used due to their higher radiation tolerance and rate capability. MDT chambers consist of three to eight layers of 30 mm diameter drift tubes filled with a gas mixture with a central tungsten-rhenium wire. Each tube achieves a spatial resolution of 80  $\mu\text{m}$ , improved to about 35  $\mu\text{m}$  for a chamber. CSC chambers, used in the highest particle flux region, are multi-wire proportional chambers with radial wires and segmented cathodes, achieving a spatial resolution of 40  $\mu\text{m}$  in the bending  $\eta$ -plane and 5 mm in the non-bending  $\phi$ -plane. The RPCs provide fast and coarse muon tracking information and bunch-crossing identification in the barrel region ( $|\eta| < 1.05$ ). They consist of three concentric layers around the beam axis, each with two parallel resistive plates filled with a gas mixture, achieving a timing resolution of around 2 ns. The TGCs are installed in the endcap wheels, covering the range  $1.05 < |\eta| < 2.7$  (2.4 for triggering). These multi-wire proportional chambers, filled with a gas mixture, achieve a timing resolution of around 4 ns.

### 2.2.6 Trigger and Data Acquisition system and data quality

The ATLAS experiment faces the challenge of managing a data influx at a bunch crossing rate of 40 MHz, which far exceeds storage capabilities. To handle this, the trigger and data acquisition system (TDAQ) selectively stores events for analysis through a two-tiered selection process: the hardware-based Level-1 (L1) trigger and the software-based High-Level Trigger (HLT). The TDAQ system must ensure synchronization across all detectors, which operate on a 40 MHz clock matching the LHC collision frequency. Clock signals are distributed via optical cables and adjusted to maintain phase alignment, preventing timing discrepancies that could affect data accuracy.

The L1 trigger, operating within a 2.5  $\mu\text{s}$  latency, uses fast electronics and field-programmable gate arrays (FPGAs) to analyze low-granularity data from the calorimeter and muon detectors. It identifies high-energy leptons, photons, jets, and significant transverse energy imbalances, reducing

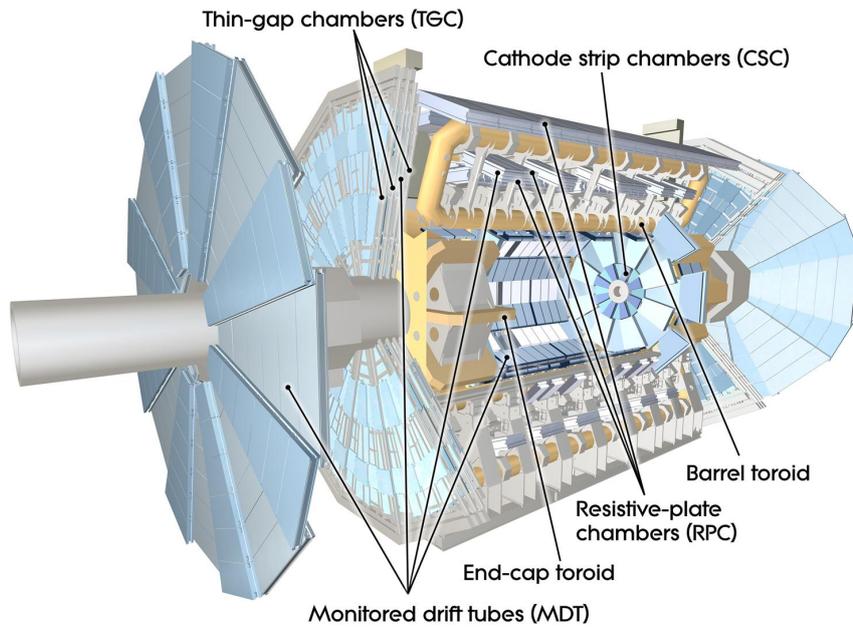


Figure 2.2.4: Computer generated image of the ATLAS Muon subsystem [76].

the event rate from 40 MHz to 100 kHz. The L1 trigger defines regions of interest, the  $\eta \times \phi \times z$  regions, and involves subsystems like the L1Calo and L1Muon. The L1Calo processes the event information in parallel by the Cluster Processor (CP) that identifies electron, photon and  $\tau$ -lepton candidates and the Jet/Energy-sum processor (JEP) identifying jet candidates and produces global sums of total and missing transverse energy. The L1Muon processes signal from RPCs in the barrel ( $|\eta| < 1.05$ ) and TGCs in the end-caps ( $1.05 < |\eta| < 2.4$ ). Combined information is processed by the L1Topo system for event-level computations, and the Central Trigger Processor (CTP) finalizes the L1 trigger decision.

Accepted events from the L1 trigger are forwarded to the high level trigger (HLT), where detailed reconstruction algorithms, similar to those used in offline analyses, process the data using full detector granularity. The HLT operates within a few hundred milliseconds, leveraging approximately  $4 \cdot 10^4$  processing units to reduce the event rate further to around 1 kHz. It uses sophisticated techniques to enhance particle identification, such as discriminating electrons and photons from hadrons, ensuring high selection efficiency.

Events passing through the HLT are fully assembled into event records and stored for offline analysis. Data is periodically reprocessed to incorporate software updates and better understand detector conditions. Quality checks ensure that events from periods of detector malfunction are flagged and excluded from analyses. A schematic view of the ATLAS trigger and data acquisition system is shown in Figure 2.2.5.

During Run-2, ATLAS recorded a total integrated luminosity at  $\sqrt{s} = 13$  TeV of  $147 \text{ fb}^{-1}$ , of which  $139 \text{ fb}^{-1}$  was classified as “good for physics” reflecting an overall data quality efficiency of approximately 95%. The recorded data is examined to exclude any events with detector-related problems, such as missing or corrupt data, high voltage (HV) trips, etc., which could result in poorly reconstructed objects and complications in data analysis. Figure 2.2.6 shows the data delivered by the LHC, recorded by ATLAS, and deemed as “good for physics” during Run-2.

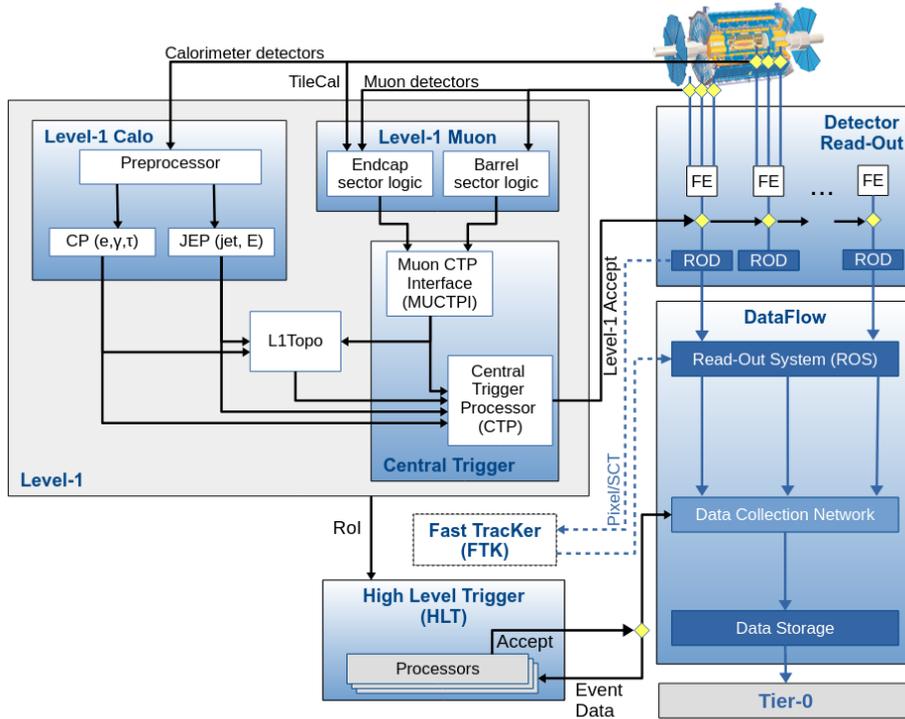


Figure 2.2.5: Schematic view of the ATLAS trigger and data acquisition system in Run 2 [77].

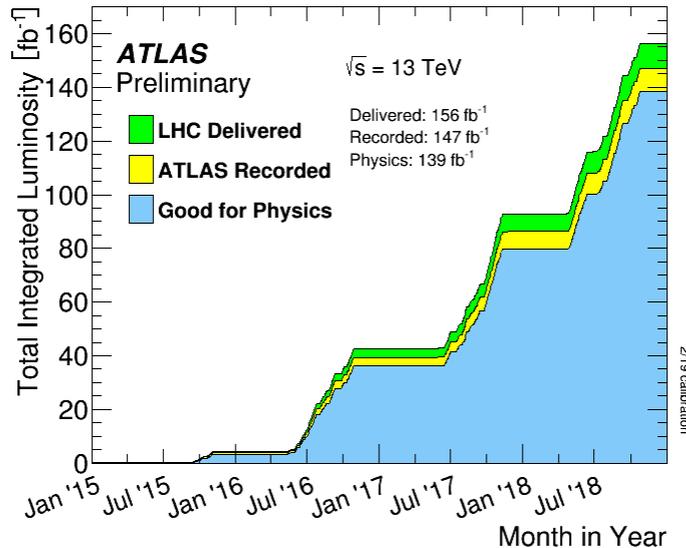


Figure 2.2.6: Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for  $pp$  collisions at 13 TeV centre-of-mass energy in 2015-2018 [78].

## 2.3 Event simulation

Understanding and interpreting results from particle colliders requires comparing them with theoretical predictions. To conduct ATLAS physics analyses, it is important to accurately simulate  $pp$  collision events. The simulation procedure relies on the Monte Carlo (MC) method, which involves repeated random sampling of variables defined by complex probability density functions to model various physics processes. The simulation progresses through a series of Markov Chain Monte

Carlo integration steps, where each step depends solely on the previous one. Section 2.3.2 details the successive stages of the event generation procedure. The following step (see Section 2.3.3) simulates the interactions of the generated particles with the detector material and the subsequent building of the electronic detector response (see Section 2.3.4). This approach enables direct comparisons between real and simulated data, both stored in the same format.

### 2.3.1 Cross-section computation in proton-proton collisions

Protons are composite particles containing quarks ( $q$ ) and gluons ( $g$ ). A proton is composed of two up quarks and one down quark being in constant interaction with each other mediated by gluons. Gluons can radiate quickly annihilating pairs of quark-antiquark within the protons. Together, these composites, referred to as partons, create a so-called sea of quarks and gluons. The parton interaction inside the proton is complicated to model at the considered energy regime since it cannot be described within perturbative QCD. When referring to the center-of-mass energy in proton-proton collisions ( $\sqrt{s}$ ), one defines it based on the proton beam energy. However, since the real interaction occurs between the proton constituents that carry only a fraction of the proton energy, the hard collisions happen at these reduced partonic center-of-mass energies. This fact complicates the modelling of the processes, introducing a source of uncertainty. The  $pp$  interaction can in fact be divided into three components: the hard-scattering, the underlying event and the hadronization. Parton's hard scatter, the interaction that happens with the highest energy, generally gives access to study the process of interest. The underlying event comprises all the activity in the collision not directly related to the hard scatter. It includes initial and final state radiation (ISR/FSR), multiple parton interactions, and beam remnants. These components interact mainly via QCD processes, involving additional quarks and gluons. After the hard scatter and the underlying event processes, partons (quarks and gluons) that are not free particles undergo hadronization. This process converts them into collimated sprays of hadrons, known as jets, which are observed in the detector.

The energy fraction carried by a parton inside a proton is defined via empirical parton distribution functions (PDFs). The production cross-section of a given particle  $X$  at a  $pp$  collider is then defined by the formula [79]:

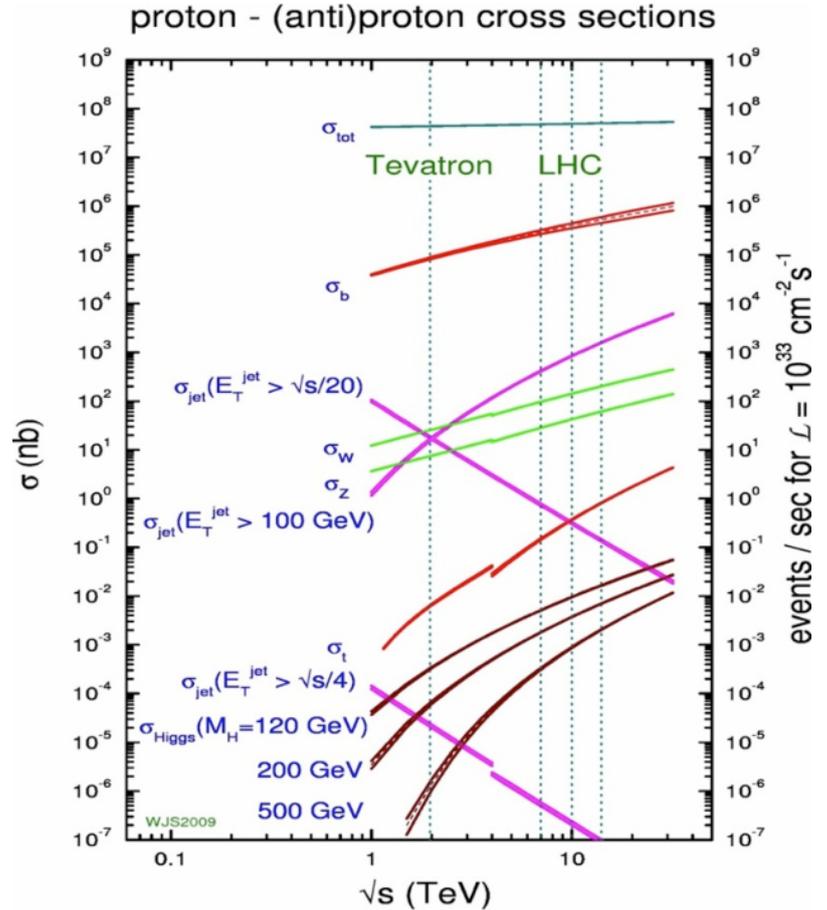
$$\sigma = \int dx_1 f_{a/p}(x_1, \mu_F) \int dx_2 f_{b/p}(x_2, \mu_F) \hat{\sigma}_{ab \rightarrow X}(x_1, x_2, s). \quad (2.3.1)$$

The protons interaction is represented via a parton-level cross-section (here,  $a$  and  $b$  are the partons participating in the production process) weighted by the corresponding PDFs ( $f_{a/p}$  and  $f_{b/p}$ ). At the LHC,  $gg$  initiated processes are more favoured than  $qg$  or  $q\bar{q}$ . The cross-section of the hard-scattering,  $\hat{\sigma}_{ab \rightarrow X}(x_1, x_2, s)$ , is calculable in perturbative QCD (where the strong coupling constant is small enough to allow for a series expansion in terms of Feynman diagrams). It is a function of the squared partonic center-of-mass energy  $\hat{s} = x_1 x_2 s$ , with  $x_1$  and  $x_2$  being the fractions of the total proton momentum carrying by the interacting partons. The parameter  $\mu_F$  is the factorization scale defining the energy below which any emission from the initial state parton is absorbed by the PDF. The introduction of  $\mu_F$  is necessary to deal with the divergencies in

theoretical calculations related to initial state partons.

Figure 2.3.1 shows the cross-sections of several SM processes in  $pp$  collisions as a function of the center-of-mass energy. Unlike Tevatron, the Higgs production cross-section is high enough at the LHC so that it is possible to measure it. The b-quark and jet productions are the dominant contributions to the total cross-section.

Figure 2.3.1: Predicted cross-sections of several SM processes in proton-(anti)proton collisions as a function of center-of-mass energy [80]. The discontinuities around  $\sqrt{s} = 4$  TeV are due to the transition between  $p\bar{p}$  on the left to  $pp$  collisions on the right. The dotted lines mark the  $\sqrt{s}$  of Tevatron and LHC.



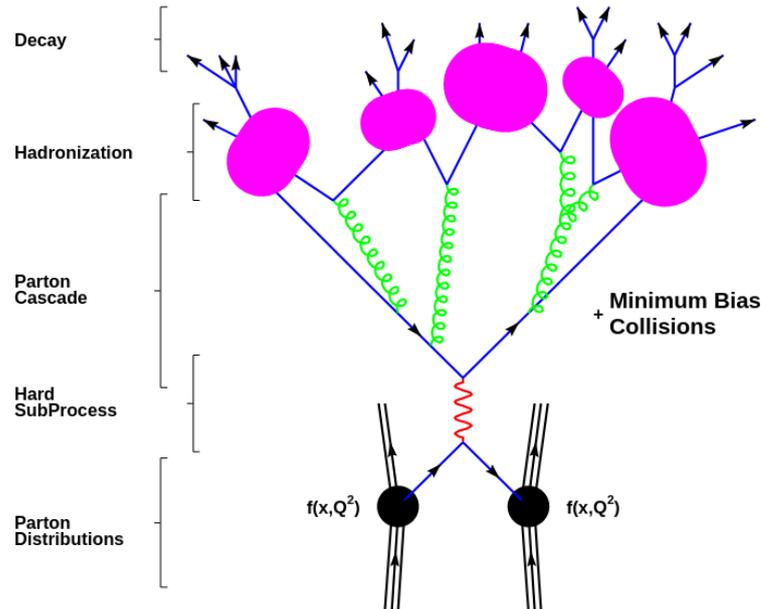
### 2.3.2 Event generation

The first stage in the ATLAS simulation chain is event generation. The simulation steps are the following (see Figure 2.3.2): partons extraction from the incoming protons, simulation of the hard scattering, parton shower simulation, and then hadronization, followed by the subsequent decay of hadrons.

The hard-scatter step, is based on the matrix element (ME) calculation, and uses the kinematics and flavours of the extracted partons to compute the cross-section within perturbative QCD. This involves evaluating the relevant Feynman diagrams for the process of interest, at a given level of precision. This phase is handled by parton-level generators (ME generators) like MadGraph [82], Sherpa [83], and Powheg [84], which calculate the probabilities and kinematics of particle interactions, incorporating electroweak and strong interactions.

The output from the matrix element generators serves as input for the subsequent parton showering (PS) process. During parton showering, gluons are emitted, carrying color charge and undergoing further radiation. This process is modelled by MC generators such as Pythia [85],

Figure 2.3.2: Schematic overview of a  $pp$  collision event generation [81].



Herwig [86], and Sherpa and accounts for the higher-order QCD corrections approximations. It continues until reaching the non-perturbative regime around 1 GeV. When partons reach such reduced energy level, they undergo hadronization, a non-perturbative process modeled using phenomenological models tuned to data. This stage involves partons binding together to form colorless baryons and mesons, which then decay into more stable particles. This is also performed using Pythia or Herwig.

### 2.3.3 Detector simulation

After the hadronization process, the generation provides a complete description of the event, independent of the detector. The next step involves simulating how generated particles interact with the ATLAS detector. This is achieved using a component-level model of the ATLAS detector implemented in Geant4 [87]. Particles are propagated and stochastic calculations determine the energy deposition of the generated particles in each detector component, resulting in a collection of hits.

The detector simulation phase is the most computationally intensive part of the full event simulation. The primary reason for this is the simulation of electromagnetic particle showers within the ATLAS calorimeter system, which takes the longest time. For instance, simulating a single event with the full detail of the detector can take up to 15 minutes of CPU time. Due to this high demand, relying solely on this detailed simulation would be impractical for generating the large number of events needed for various physics studies. To address this, a faster alternative simulation method, called ATLFAST-II (AFII) [88], was developed. AFII reduces CPU time by using pre-simulated electromagnetic showers for low-energy particles instead of generating them anew for each event. Although full simulation samples are more precise and are used for detailed physics analyses, AFII samples are employed for optimization studies and to evaluate theoretical uncertainties because they can be produced more quickly.

### 2.3.4 Digitization

The next simulation step is aiming at building detector electronics response. This process, called digitization, involves converting the simulated energy deposits from the Geant4 toolkit into electrical responses. To more accurately reflect the hardware setup, noise is introduced during digitization, and the data readout structure of the ATLAS electronics is replicated. The High-Level Trigger (HLT) and reconstruction processes are then applied to the simulated outputs in the same manner as they are to experimental data.

The final step, known as reconstruction, is performed identically for both simulated and real data events, ensuring consistency in the analysis.

## 2.4 Object reconstruction and identification in ATLAS

The real or simulated output from the detector undergoes a reconstruction process to identify the type and properties of the particles present in an event. The  $t\bar{t}H$  analysis in the multilepton channels, presented in this thesis, utilizes electrons, muons, jets (and  $b$ -jets), and missing transverse momentum.

### 2.4.1 Tracks and vertices

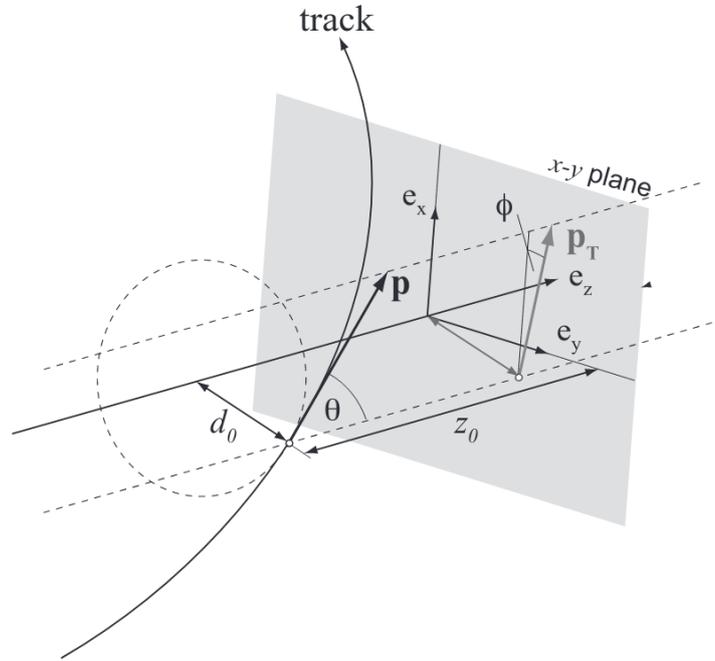
The reconstruction of charged particle trajectories in the ATLAS Inner Detector are based on tracking algorithms [89]. When a charged particle hits a sensor in the ID, it generates a hit, i.e. a space-point. First, an inside-out algorithm constructs track seeds from hits in the pixel and SCT detectors. These seeds are extended by adding hits from subsequent SCT layers using a combinatorial Kalman filter [90]. This process generates several track candidates per seed, necessitating an ambiguity resolution procedure that employs a score-based ranking system considering factors like the number and type of hits, number of holes (missing hits), goodness of fit, and track momentum. A neural network helps distinguishing between shared hits, which may arise from insufficient detector granularity or fake tracks. In the final stage, an outside-in algorithm identifies hits not chosen by the inside-out method, starting with track segments in the TRT and extrapolating them back to the silicon detectors. This approach is particularly effective for tracks from long-lived particles. Tracks passing ambiguity resolution and TRT extension are subjected to a final fit before inclusion in the final track collection. The fit estimates the track parameters (see Figure 2.4.1):

- Transverse Impact Parameter ( $d_0$ ): the shortest distance between the track and the primary vertex (PV) in the transverse plane.
- Longitudinal Impact Parameter ( $z_0$ ): the distance along the beam axis between the PV and the track's closest approach point.
- Azimuthal Angle ( $\phi$ ): the angle between the transverse momentum vector  $p_T$  and the  $x$ -axis in the transverse plane.

- Polar Angle ( $\theta$ ): the angle formed between the momentum vector  $\vec{p}$  and the  $z$ -axis.
- $e/p_T$ , the electric charge over the transverse momentum ratio, is determined by the magnetic field strength and the track curvature.

The track resolution of  $d_0$  ( $z_0$ ) ranges between  $99 - 25 \mu\text{m}$  ( $170 - 75 \mu\text{m}$ ) for  $0.5 < p_T < 10 \text{ GeV}$  in  $0.2 < |\eta| < 0.4$ .

Figure 2.4.1: Schematic diagram of the track parameters [91]



Vertex reconstruction identifies the origin points of tracks. In particular, the primary vertex (PV) is the point in space where the initial collision between protons. The vertexing algorithm starts by identifying a global maximum in the  $z$ -position distribution of tracks selected based on their  $p_T$ . The optimal vertex position is then refined using a  $\chi^2$  fitting algorithm, which iteratively excludes tracks incompatible with the vertex. Displaced tracks are used to seed new vertices, and this process repeats until no further seeds are found. The vertex with the highest sum of the squared transverse momenta of associated tracks is identified as the PV, while others are considered as coming from pile-up interactions.

## 2.4.2 Electrons and photons

Electrons and photons are reconstructed within  $|\eta| < 2.47$  from electromagnetic topoclusters and tracks identified in the inner detector, with the barrel-endcap transition region ( $1.37 < |\eta| < 1.52$ ) excluded. Figure 2.4.2 shows the path of an electron through the detector.

The reconstruction begins with seed-cluster formation in the EM calorimeter [93], split into towers of  $0.025 \times 0.025$  in  $\eta \times \phi$  space. A sliding-window algorithm with a  $3 \times 5$  tower window scans for seeds with transverse energy above 2.5 GeV. Clusters are then built around these seeds and matched to tracks by applying the Gaussian-sum filter [94], which considers possible energy loss due to bremsstrahlung. Successful matches are tagged as converted photons if the track is not from the primary vertex, otherwise, they are electron candidates. The cluster is labelled

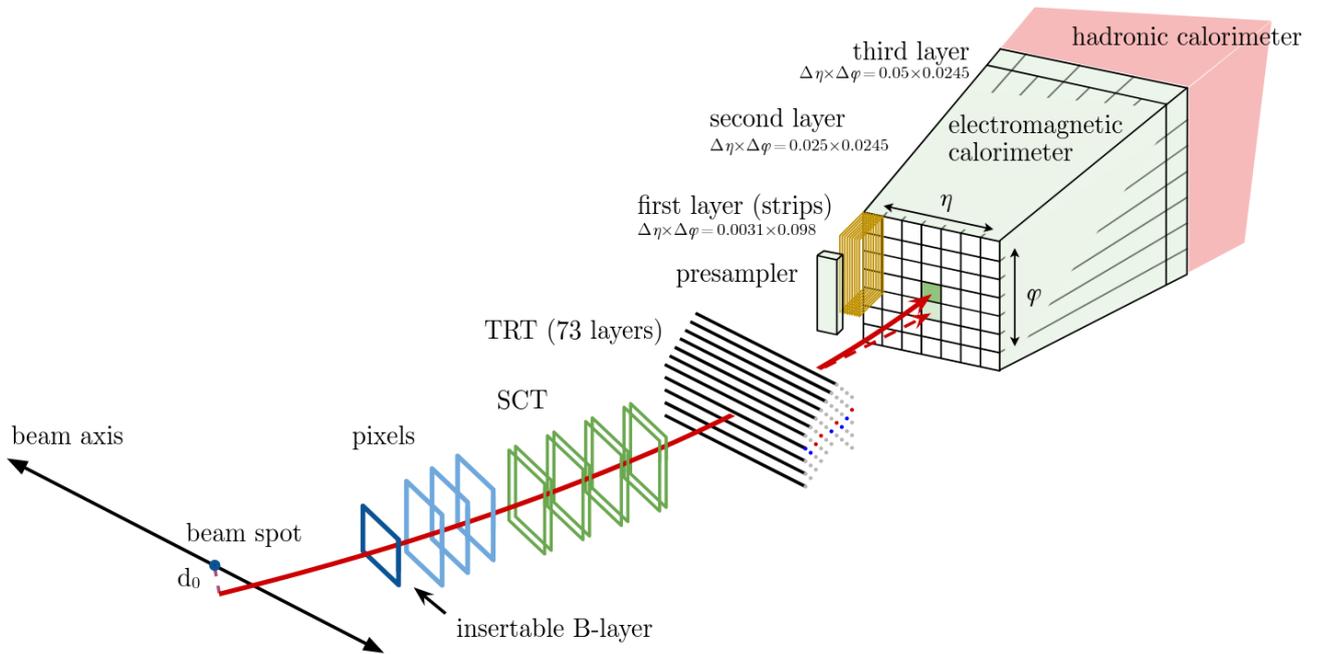


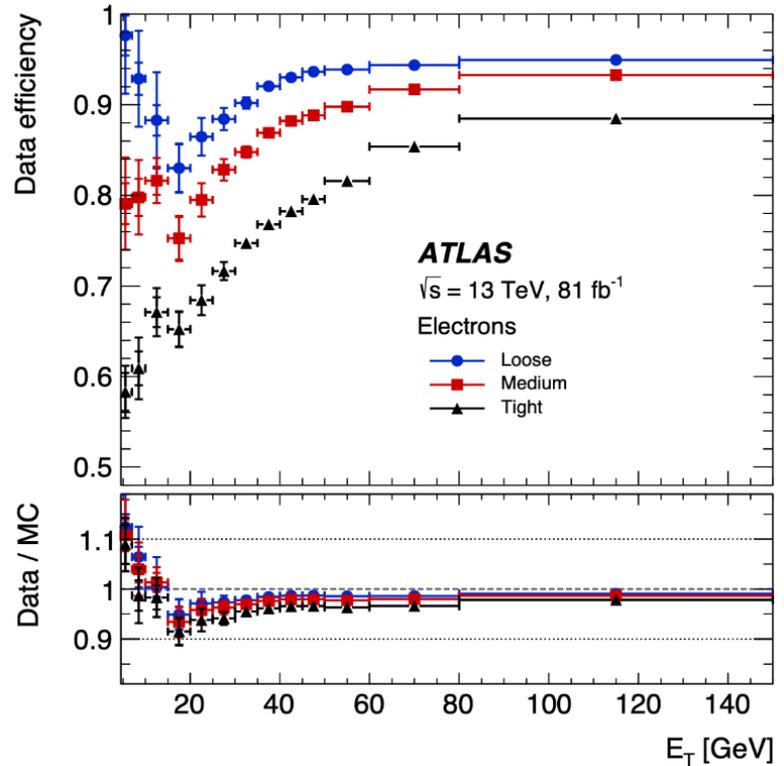
Figure 2.4.2: A schematic illustration of the path of an electron through the detector. The red trajectory shows the hypothetical path of an electron, which first traverses the tracking system (pixel detectors, then silicon-strip detectors and lastly the TRT) and then enters the electromagnetic calorimeter. The dashed red trajectory indicates the path of a photon produced by the interaction of the electron with the material in the tracking system [92].

as an unconverted photon when no match is established. Matched clusters are merged with nearby EM clusters to form superclusters, improving energy resolution. Calorimeter cells from specific layers are associated with superclusters, with energy and position corrections applied. Corrections are made for intercalibration of calorimeter layers, pile-up-induced energy shifts, and energy uniformity. The energy scale and resolution are calibrated using  $Z \rightarrow ee$  decays using the so-called tag and probe method as for muons (see Section 2.4.3). For electrons, identification algorithms reject background objects like QCD jets or non-prompt electrons. This uses calorimeter and track-based variables. Likelihood-based identification is used, with working points defined as: Loose, Medium, and Tight. Figure 2.4.3 shows the electron identification efficiencies at these working points. Photons are identified through a cut-based approach, with separate criteria for unconverted and converted photons. Isolation variables are defined to quantify activity around electron candidates, involving the energy or momentum in a cone around the candidate.

### 2.4.3 Muons

The muon reconstruction process is initially conducted independently in both the ID and the MS. These measurements are later combined to form the muon tracks used in physics analyses. The muon reconstruction in the ID follows the same procedure as other charged particles (see Section 2.4.1). The muon reconstruction in the muon spectrometer starts by forming local track segments from hits in each MS chamber. These segments are combined into track candidates, requiring at least two matching segments per track, except in the barrel-endcap transition region

Figure 2.4.3: Electrons identification efficiency in  $Z \rightarrow ee$  data events as a function of the electron transverse energy  $E_T$  [93].



where one segment suffices. A global  $\chi^2$  fit of the hits associated with each track candidate is performed, and tracks failing the selection criteria are discarded. Muons are classified into different categories based on which subdetectors are involved in their reconstruction:

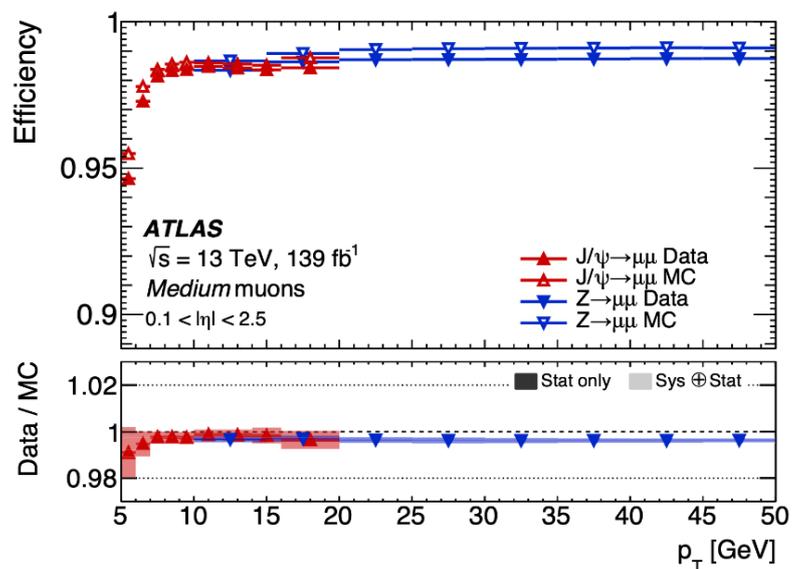
- Combined Muons: results from a global refit using hits from both ID and MS, providing the best quality in terms of fakes rejection and momentum resolution.
- Segment-tagged Muons: obtained from the association of an ID muon track with one local track segment from the MS, for muons crossing only one MS layer due to low  $p_T$  or reduced acceptance regions.
- Calorimeter-tagged Muons: Obtained by matching an ID muon track to an energy deposit in the calorimeter compatible with a minimum-ionizing particle, used to recover acceptance in the  $|\eta| < 0.1$  region.
- Extrapolated Muons: Uses only MS information with a loose requirement compatible with the interaction point, extending muon reconstruction acceptance into the  $2.5 < |\eta| < 2.7$  region not covered by the ID.

After reconstructing muon candidates, an identification procedure is implemented to reduce the contamination from background, primarily coming from pion and kaon decays, and to enhance the selection efficiency of prompt muons. There are five identification levels available to meet the diverse requirements of physics analyses: Medium, Loose, Tight, Low  $p_T$ , and High  $p_T$ . This section focuses on the Medium working point that is used in the analyses presented in this thesis.

Efficiencies for reconstruction, identification, and isolation are measured in both data and simulations using the tag-and-probe method. This method uses a well-identified "tag" particle

and a less strictly identified "probe" particle within the same event to determine the efficiency of the probe particle's identification and reconstruction. The evaluation is performed using  $Z \rightarrow \mu\mu$  events for muons with  $p_T > 15$  GeV and  $J/\psi \rightarrow \mu\mu$  events for muons with  $5 < p_T < 15$  GeV. Similarly, these events are used to study muon momentum scale and resolution. This process allows for the derivation of correction factors to apply to simulations to match data measurements. Figure 2.4.4 shows the muon identification efficiency for the Medium working point. Additionally, muons must be isolated from other detector activities to effectively reject those from heavy flavor semileptonic decays occurring within jets. To achieve this, nine isolation selections, or working points, are provided, each optimized for specific analyses. These criteria use a track-based isolation variable that depends on the momenta of the muon track and surrounding tracks, along with a calorimeter-based variable that utilizes energy deposits in a cone around the muon track.

Figure 2.4.4: The reconstruction efficiency as a function of the muon  $p_T$  for  $Z \rightarrow \mu\mu$  and  $J/\psi \rightarrow \mu\mu$  events for the Medium working point [95]



## 2.4.4 Jets

Color confinement prevents gluons and quarks produced as final state partons from existing as free particles. Instead, they undergo hadronization, producing collimated showers known as jets. These jets are observed as clusters of energy deposits in the calorimeter system, which can be linked to charged particle tracks in the Inner Detector.

### 2.4.4.1 Reconstruction

The tracker's ability to reconstruct charged particles is enhanced by the calorimeter's capability to reconstruct both charged and neutral particles. At high energies, the calorimeter offers better energy resolution compared to the tracker's momentum resolution, making the combination of these two subsystems optimal for the jet reconstruction. Outside the tracker's geometrical acceptance, only calorimeter information is available. Jets reconstructed from tracks in the ID are called track jets, while those reconstructed from calorimeter data are calorimeter jets. ATLAS employs various jet collections based on different reconstruction algorithms. EMTopo jets are reconstructed using only calorimeter-based energy information. Particle Flow (PFlow) jets [96],

based on the anti- $k_t$  algorithm [97], combines measurements from the calorimeter and ID. PFlow jets benefit from superior momentum resolution of the tracking detectors for low-energy particles and better angular resolution, leading to improved jet energy and angular resolution compared to EMTopo jets. The analysis presented in this thesis uses PFlow jets.

The calorimeter information comes from the energy deposits in the electromagnetic and hadronic calorimeters from charged particles and neutral objects. These deposits should be topologically adjacent clusters of calorimeter cells, or topoclusters defined by a clustering algorithm [98]. The purpose of the clustering algorithm is to group cells with significant energy deposits, effectively suppressing noise. The total noise in calorimeter cells ( $\sigma$ ) is determined by combining the measured electronics noise and pile-up noise quadratically. The clustering algorithm is initiated with seed cells that have energy deposits greater than  $4\sigma$ . It then iteratively adds neighboring cells to the topocluster if their energy exceeds  $2\sigma$ . Following this, all adjacent cells are incorporated. In the final step, a cluster splitting algorithm divides the produced topoclusters based on local energy maxima to avoid overlap.

The anti- $k_t$  algorithm uses the four-vectors that may be truth stable particles from MC, charged-particles tracks or energy deposits from the calorimeter. The PFlow anti- $k_t$  algorithm uses a radius parameter  $R$  set to  $R = 0.4$  and processes inputs from two sources: a list of tracks and a list of topoclusters. The process begins by selecting well-measured tracks. For each of these tracks, the algorithm attempts to find a corresponding topocluster in the calorimeter. It calculates the expected energy deposit in the calorimeter from the particle that created the track, using the track's momentum and the position of the topocluster. Next, the algorithm evaluates whether additional topoclusters need to be included to account for the entire shower energy. It then subtracts the expected energy deposit, calculated on a cell-by-cell basis, from the matched topoclusters. If the remaining energy in the system aligns with the expected shower fluctuations of a single particle's signal, the leftover topocluster remnants are discarded. This process is applied in order of descending track  $p_T$ , starting with cases where a track is matched to a single topocluster, and subsequently extending to other selected tracks.

#### 2.4.4.2 Calibration

Calibrating reconstructed jets to match the energy of truth-jets, formed from stable final-state particles and reconstructed by the anti- $k_t$  algorithm, involves several steps (see Figure 2.4.5). The calibration process combines simulation and data-driven techniques to account for various factors affecting jet response, such as pile-up, biases from the clustering algorithm, detector geometry effects, and data-simulation differences.

Pile-up causes the presence of additional energy deposits. To mitigate this contribution, a pile-up subtraction is performed in two steps [100]. The first step is called area-based subtraction and is using a per-event pile-up estimation with a method based on  $p_T$ -density. In a second step, a residual correction is applied to jet  $p_T$  to account for the difference between the reconstructed jet  $p_T$  and the true jet  $p_T$ . This residual dependence is observed as a function of both the number of reconstructed primary vertices in the event and the number of interactions per bunch crossing, which are sensitive to in-time and out-of-time pile-up, respectively. Jets originating from pile-

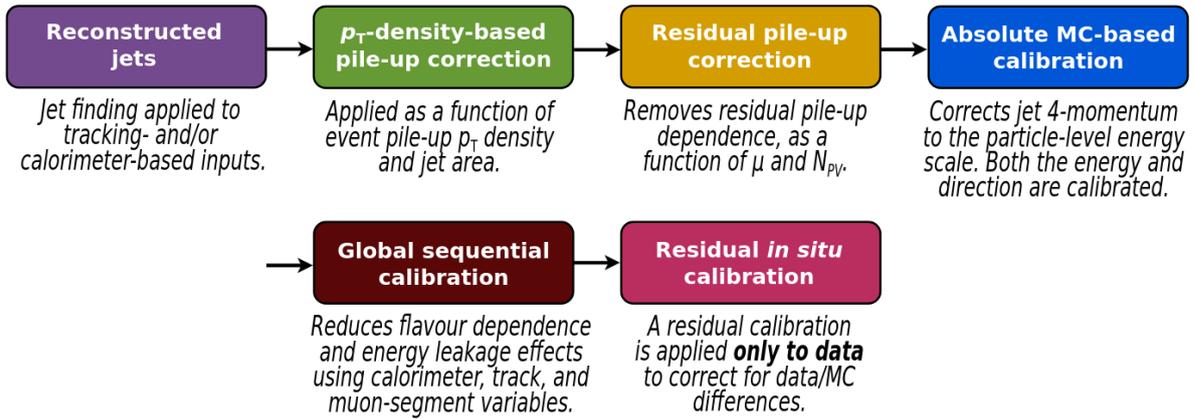


Figure 2.4.5: Jet calibration chain [99].

up interactions are also filtered using the Jet Vertex Tagging (JVT) algorithm [101]. The JVT algorithm associates jets to the primary vertex of the event based on the sum of the transverse momenta of tracks within the jet that originate from the primary vertex using a 2-dimensional likelihood variable. A JVT score ranges from 0 (pile-up-like) to 1 (hard-scatter-like). For jets with  $p_T < 60$  GeV and  $|\eta| < 2.4$ , a  $JVT > 0.2$  is required to associate them with the hard-scatter interaction.

The absolute jet energy scale (JES) and  $\eta$  calibrations adjust the reconstructed jet four-momentum to match the particle-level energy scale. JES measures the average ratio between reconstructed and true jet energy. These calibrations account for the non-compensating response of the calorimeter, energy losses in passive materials, out-of-cone effects, and biases in the jet  $\eta$  reconstruction and are determined using pile-up-free MC samples.

The Global Sequential Calibration (GSC) adjusts for various factors, including: non-compensation (accounting for different responses to hadrons, leptons, and photons), flavor dependence (accounting for different responses to quarks and gluons) and punch-through (accounting for jets extending beyond the calorimeters). The GSC applies a series of multiplicative corrections to mitigate these fluctuations and enhance jet resolution without altering the average jet energy response. Jet Energy Resolution (JER) indicates the width of the fitted jet response distribution, which should be minimized for optimal resolution.

The final calibration step addresses discrepancies between the jet response in data and simulation. These discrepancies arise due to imperfect simulation of detector materials and physics processes, including the hard scatter, underlying event, jet formation, pile-up, and particle interactions with the detector. This final in situ calibration involves measuring the jet response separately in data and MC simulations, then applying the ratio of these measurements as an additional correction to the data.

### 2.4.4.3 $b$ -tagging

$b$ -tagging refers to identifying hadronic jets that originate from  $b$ -quarks. In simulated data, the true flavour of a jet is established through spatial  $\Delta R$  matching between stable hadrons and reconstructed jets, employing a hierarchical matching process that prioritizes  $b$ -hadrons first,

followed by charm hadrons and then  $\tau$ -leptons.

Distinct properties of  $b$ -hadrons, such as their longer flight path which results in a displaced secondary vertex, and their heavy hadronic decay signature with higher charged track multiplicity, are exploited to identify them. ATLAS utilizes low- and high-level  $b$ -tagging algorithms where the latter ones are multivariate classifiers combining low-level algorithm results. The multilepton  $t\bar{t}H$  measurements presented in this thesis relies on the DL1 high-level algorithm [102] for the  $b$ -jets identification. The DL1 algorithm is based on a deep feed-forward neural network. As an input, it takes some kinematic variables and the output of low-level impact parameter-based algorithms (IP2D and IP3D [102]), secondary vertex reconstruction algorithm (SV1 [103]), and decay chain reconstruction algorithm (JETFITTER [104]). The impact parameter algorithms focus on the signed impact parameter significance of tracks. Secondary vertex algorithm reconstructs displaced secondary vertices from selected tracks within a jet. JETFITTER uses the topological structure of weak  $b$ - and  $c$ -hadron decays within the jet to attempt a reconstruction of the entire  $b$ -hadron decay chain. The final output tagging discriminant is defined by combining the probabilities of  $b$ -jet,  $c$ -jet and light-flavour jets given by the neural network output. The DL1 discriminant distribution is divided into five, so-called, pseudo-continuous bins, based on selections. Based on them, five working points are defined: Loose 85%, Medium 77%, Tight 70%, Very-Tight 60%, where the percentage corresponds to the efficiency evaluated on a  $t\bar{t}$  sample. Each WP undergoes independent calibration to align the  $b$ -tagging efficiency in simulation with that in data. This is achieved by estimating data-MC scale factors, which are the ratios of efficiencies between data and simulation, and applying these scale factors to the simulated results.

## 2.4.5 Missing Transverse Energy

The presence of particles escaping detection in the final state can be inferred from an imbalance in the measured momentum of detected particles, referred to as missing energy. This quantity is crucial for processes involving decays to neutrinos and for searches for beyond the Standard Model phenomena, where additional weakly-interacting particles are predicted. Measuring the full missing energy is impossible because the initial momenta of partons are inaccessible. However, assuming that the initial momenta are negligible in the transverse plane, given that  $pp$  collisions occur along the  $z$ -axis, allows access to the Missing Transverse Energy (MET), denoted as  $E_{miss}^T$ .

The reconstruction of MET is based on energy deposits in the calorimeters, and tracks reconstructed by the Inner Detector and Muon Spectrometer [105]. MET is derived from the negative vector sum of the transverse momenta of all reconstructed and calibrated physics objects, including electrons, photons, hadronically decaying  $\tau$ -leptons, jets, and muons. Additionally, a term accounts for soft energy contributions not associated with these objects, primarily resulting from underlying events and soft radiation. The soft term is calculated using the momenta of tracks from the ID that match the hard scatter primary vertex and are not associated with the reconstructed objects.

### 2.4.6 Hadronically decaying taus

A hadronically decaying  $\tau$  lepton ( $\tau_{had}$ ) produces a neutrino and a set of charged and neutral pions. To reconstruct the visible products, jets are selected as described in Section 2.4.4, and a boosted decision trees are then employed to assess whether tracks near a cluster come from a  $\tau_{had}$ .

# 3 Automatic visual inspection of ITk pixel modules

The High-Luminosity Large Hadron Collider (HL-LHC) is an upgrade of the current LHC, aiming to increase its nominal instantaneous luminosity by 5 to 7 times [106], resulting in  $7.5 \cdot 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , nearly three times that of LHC Run-2. This boost will lead to a significantly higher number of collisions, allowing scientists to make more precise measurements of the Standard Model and explore Beyond Standard Model physics, potentially leading to new discoveries. The upgrade is planned for 2026-2028 during Long Shutdown 3. The ATLAS detector will undergo significant upgrades to cope with these enhancements, notably replacing its current Inner Detector with an all-silicon Inner Tracker (ITk), described in Section 3.1. ITk Pixel modules are one of the basic building blocks of the ITk. About  $10^4$  ITk pixel modules need to be assembled. The production chain involves many sites, and each production step includes quality control procedures, such as visual inspection, metrology, and electrical testing at different temperatures. This chapter presents work on creating a tool for automatic visual inspection of the ITk Pixel module pads necessary for the wire-bonding assembly step. The developed tool is based on the anomaly detection (AD) deep learning algorithm. Section 3.2 introduces the visual inspection procedure and the preparation steps followed to test and optimise the performance of existing AD algorithms. The study of the limitations of the chosen algorithms' performances and the proposed solution are presented in Section 3.3.1. Section 3.4 presents a Graphical User Interface (GUI) created to make the algorithms accessible to the wire-bonding teams across the different production sites.

## 3.1 All-silicon Inner Tracker and Pixel modules

### 3.1.1 Inner Detector upgrade

The main goals of ITk are to enhance radiation hardness, improve granularity, boost timing capabilities, and reduce the material budget. ITk must withstand the higher radiation levels from the increased collision rate, which will be addressed using new radiation-tolerant materials and electronics. To achieve improved granularity, ITk will feature more modules and additional tracking layers with smaller pixel sizes and strip pitches compared to the current ID. Enhanced timing capabilities will be achieved by reducing the timing uncertainty of the detector modules. Reducing the material budget is crucial to minimize multiple scattering and improve particle trajectory measurements. This will be accomplished using lightweight materials and optimizing the detector layout, lowering material density and reducing the scattering and energy loss of particles passing through the detector, thus improving measurement accuracy. These can be quantified with the following requirements [107]:

- Full detector coverage for tracks with transverse momentum  $p_T > 1 \text{ GeV}$  originating from a

region within  $|z| < 15$  cm and  $r < 2$  mm;

- Reconstruction of tracks from the primary vertex up to  $|\eta| < 4$ ;
- Efficient tracking performance comparable to or better than the current Inner Detector under increased HL-LHC pile-up conditions;
- Radiation resistance for the expected 10 years of HL-LHC operation with improved instantaneous and integrated luminosities (for example, radiation hardness in the inner system up to  $8 \text{ MGy} / 10^{16} \text{ n.cm}^{-2}$  at  $2000 \text{ fb}^{-1}$  is expected).

The ATLAS all-silicon ITk consists of a strip detector in the outermost layers and a pixel detector closer to the interaction point (see Figure 3.1.1).

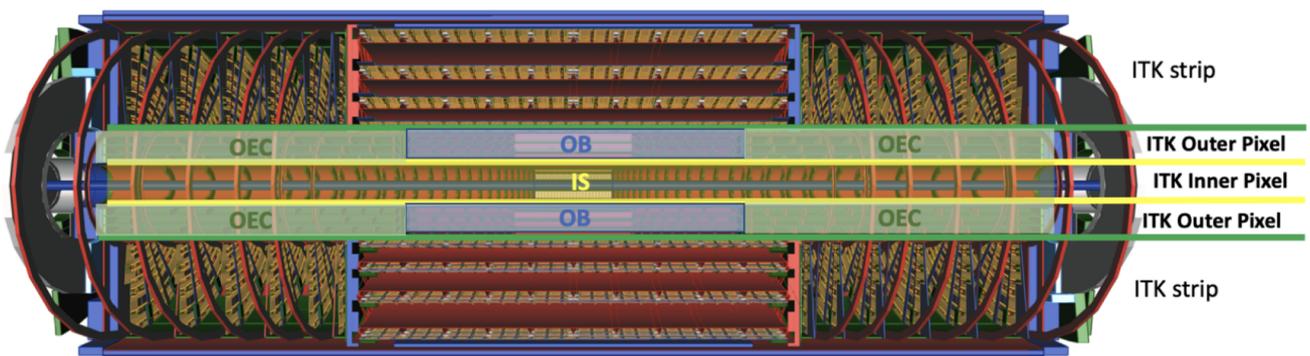


Figure 3.1.1: The full ATLAS ITk layout [108]. OB stands for the Outer Barrel subsystem, IS for Inner System and OEC for Outer Endcaps.

The strip detector features four barrel layers, and six end-cap rings on each side of the barrel, spanning from 385 mm to 1000 mm in radius and up to  $|z| < 2850$  mm [107]. It covers the pseudorapidity region of  $|\eta| < 2.7$ .

The pixel detector of ITk is divided into three subsystems [109]: the Inner System (IS), the Outer Barrel (OB), and two Outer Endcaps (OECs). Figure 3.1.2 shows the layout of the upper-right quadrant of ITk and the corresponding pseudorapidity coverage. The IS comprises a barrel and an endcap section within  $|z| < 250$  mm, covering up to  $|\eta| < 4$ . Being the closest to the interaction point, the Inner System endures the highest radiation levels and is designed to be replaceable once its radiation tolerance is exceeded, which occurs around an integrated luminosity of  $2000 \text{ fb}^{-1}$ . The OB layout consists of flat and inclined sections within  $|z| < 400$  mm, ensuring the hermeticity of ITk, reducing the number of required modules, and minimizing the material needed.

### 3.1.2 ITk Pixel modules and its production chain

The ITk Pixel detector's basic building block is a module. A bare module comprises a silicon sensor connected to a readout front-end chip. When charged particles pass through the sensor, an electrical current is induced by the motion of charges after incident radiation, creating a signal that the readout chip collects and transmits for further processing. Each sensor pixel is connected

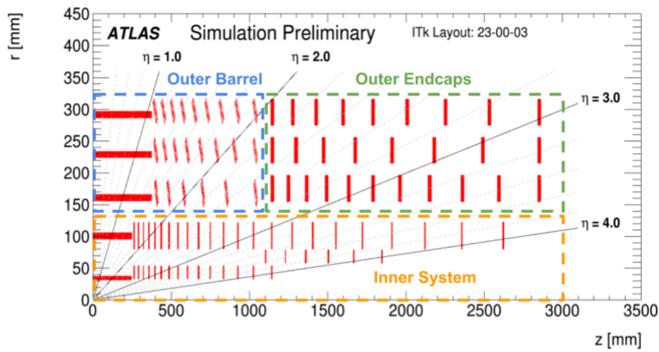


Figure 3.1.2: Schematic layout of one quadrant of the ITk pixel system [110].

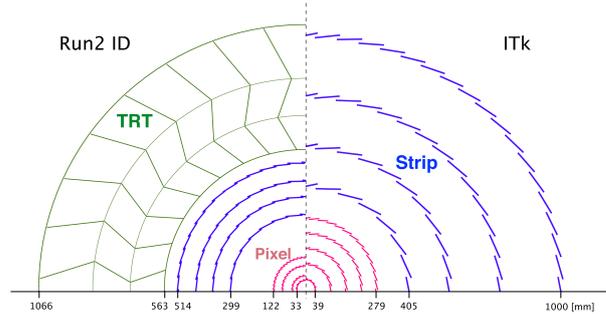


Figure 3.1.3: The radial layouts of the current ATLAS tracker and the ITk. [111].

to the corresponding analogue circuitry on the front-end chip by bump bonding. The ITk Pixel module is a bare module glued to a flexible Printed Circuit Board (PCB). The PCB and bare module are connected using microwires that are wired-bonded after glueing. Figure 3.1.4 (left) shows the layer structure of an ITk module.

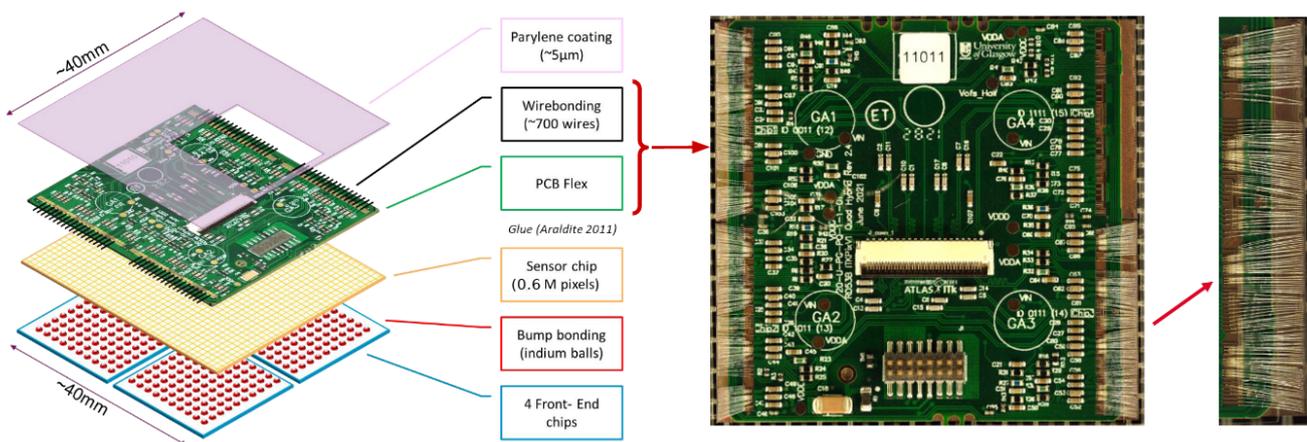


Figure 3.1.4: *Left*: structure of an ITk pixel module; *Middle*: photo of a partially wire-bonded module; *Right*: a zoomed-in part of a wire-bonded pad.

The modules considered in this work contain 4 readout chips and are called quad modules. The size of each module is about  $4 \times 4$  cm. There were several versions of the readout chip. First, RD53A is a prototype chip that features a matrix of  $192 \times 400$  pixels, each  $50 \mu\text{m} \times 50 \mu\text{m}$ . The next chip version, the ATLAS ITkPixV1, is more stable and closer to the final design. It has a larger matrix of  $384 \times 400$  pixels and the differential front-end circuit. The final design implemented in ITkPixV2 will be used for the ITk assembly. Apart from the frontend chip versions, there are also different flex PCB versions, RD53A-style and ITkPix-style, with different sub-versions. The flex PCB versions have distinct pad geometries.

The module assembly process involves several crucial steps: visual inspection, metrology, flex attach, parylene coating, and mechanical wire-bond protection assembly. First, digital microscopes are used to visually inspect the bare components upon arrival to ensure good quality. Metrology involves measuring components before and after assembly to confirm they meet specified dimensions. These measurements are repeated after each significant modification in the assembly process

to consistently verify the module's integrity and adherence to specifications. Although a complete visual inspection of the modules is performed regularly during the quality control chain, the work presented in this thesis specifically focuses on the visual inspection of the wire-bonding pads.

Wire-bonding is a method of making connections between an integrated circuit (IC) and other semiconductor devices or between two printed circuit boards (PCBs). It is the most cost-effective and flexible interconnect technology and is used to assemble the majority of semiconductor packages. The wire-bonding machine used at Saclay is of the F&S BondTec 58 series. This machine has a working area of  $200 \times 150$  mm, a programmable resolution of  $1 \mu\text{m}$  and a programmable  $z$ -axis with a 100 mm stroke and a placement accuracy of  $\pm 5 \mu\text{m}$ .

Around 700 wires are used to link the PCB with the bare module. The outcome of this wire-bonding depends heavily on the quality of the PCB pads, located on the PCB's sides (see Figure 3.1.4 *middle* and *right*). Defects such as traces of chemical contamination, scratches and dust can be present on the module pads (see Figure 3.1.5). Having the knowledge of these defects and their locations can help the wire-bonding teams to optimise the process by avoiding placing wires on the defected surfaces.

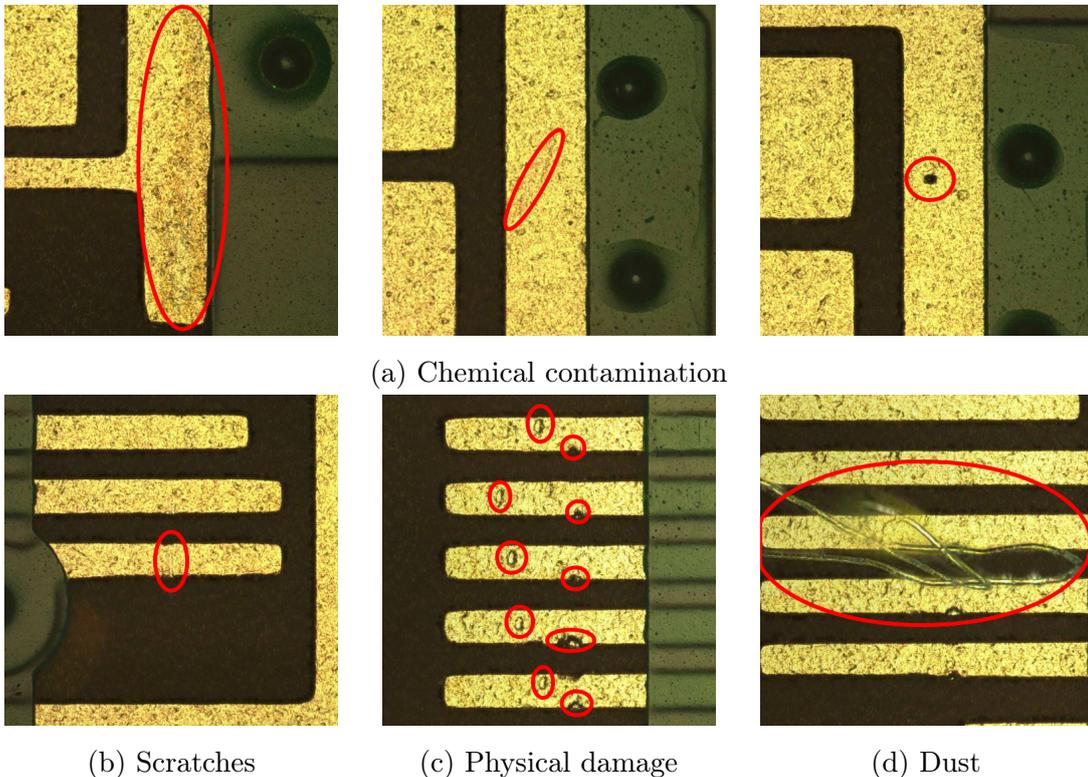


Figure 3.1.5: Examples of ITkPix pads surface defects (circled in red).

## 3.2 Choosing the automatic visual inspection algorithm

The approach to perform the automatic visual inspection of the ITk pixel modules pads suggested in this work is based on machine learning. Machine learning (ML) is an area of artificial intelligence (AI) that focuses on training computer systems to perform tasks without explicit instructions. Machine learning algorithms are used in a wide range of applications. In computer vision, a domain

studying the algorithms able to retrieve information from digitalized visual input, machine learning methods often outperform other approaches. For example, a computer vision ML algorithm can be trained to recognize specific objects in a dataset of images by extracting statistical properties. Anomaly detection algorithms can detect unusual events or patterns in a given dataset. Anomaly detection on images can be applied for quality control on industrial production sites to automatise the visual inspection of the products. There are several possible ways of using the anomaly detection algorithm on images:

- classification: yields binary output telling if a defect is presented on an input image;
- detection: gives a rectangular box within which the defect is located;
- segmentation: provides an output mask of the same size as the input image with black pixels corresponding to non-anomalous areas and white pixels located at the defective areas.

To efficiently perform visual inspection before wire-bonding, the desired output of the automatic algorithm is a map highlighting the defects' location, size and shape. Thus, we proceeded with the anomaly segmentation type of ML model.

This section first describes the procedure of collecting pad photos (Section 3.2.1) and then explains how the datasets for the training of machine learning algorithm are built from the collected photos (Section 3.2.2). Our main focus are unsupervised anomaly detection algorithms, the descriptions of the studied models can be found in Section 3.2.3.

### 3.2.1 Photo-taking procedure

To create a dataset of images to train an ML algorithm, photos of module pads were collected. Given the availability of two types of microscopes (LEICA and Keyence), the Keyence VHX-7000 was chosen to take the photos. Indeed, this instrument offers images of superior resolution, thus allowing us to test the concept of automated anomaly detection with the best data quality. The Keyence microscope provides numerous options for lighting, magnification, colour correction, etc... (see Figure 3.2.1).

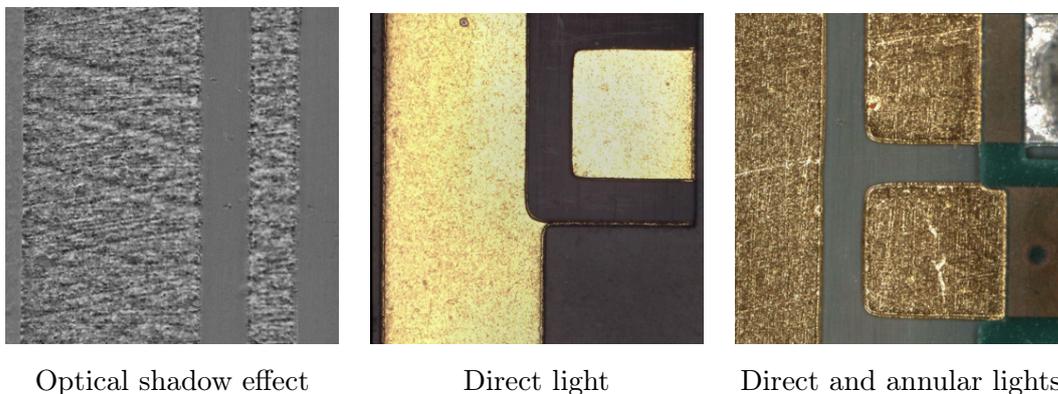


Figure 3.2.1: Examples of lightning modes available on the Keyence microscope, photos of RD53A modules.

It is important to find a setup that provides the photos on which most of the pads defects are visible. The photos of the whole modules that are taken at earlier assembly stages for the

visual inspection of, for example, electronic components placement, are not convenient for pads inspection. The resolution of these photos does not allow us to see the pad's defects. For the pad visual inspection, new photos need to be collected, limited to the area of interest: two pad areas on the left and right side of a module, around 4 cm height and 20 mm width. At the same time, there are some technical constraints to the pad photos and photo-taking procedure:

1. Being part of visual inspection before wire-bonding, the photo-taking procedure should be as simple as possible to save time for the wire-bonding team;
2. The pad photos are supposed to be stored in the production database. Thus, there is a strong preference to limit the size to 50 Mb per module. However, it is possible to store only parts of the pad photo containing significant defects.

The direct light and the mix of direct and annular lights were considered as the most convenient lighting options. While side light and optical shadow modes allow to better see relief connected problems, the chemical contamination, which is the dominating kind of defects, is hardly visible. The optical shadow effect captures multiple frames of an image in varying lighting angles, resulting in photos about 4 times bigger in memory than those taken without this mode and do not pass the size requirement.

The photo-taking procedure was slightly modified for the different types of modules: changing the vendor the pads granularity changed too, as well as the PCB colour connected to the coverlay material. The different modules types also have differences in the pads shape. This led to the creation of separate collections of photos for each module version. The detailed description of these photo sets is provided in the Section 3.2.2.

The photos of RD53A modules (having RD53A version of frontend chip and PCB) were taken with x150 magnification and direct light. This allowed us to make visible even the smallest defects, but the photo-taking procedure and the size of the photos per module were at the edge of the acceptable values. Because the microscope could not take a photo of more than 20000 pixels size, to cover one module, 4 photos had to be taken. Each photo is about 25 Mb on average. For the ITkPixV1.1 modules, we decided to decrease the magnification to x100. Due to this choice, only two 27 Mb photos per module are needed. However, the magnification change also induced a change in the light setting: with this new magnification value, the best visibility was achieved with the mix of direct and annular lights. There is also a trade-off between details and size, but in our case, most of the defects were still visible even with a smaller magnification.

To make a high-resolution photo, the Keyence microscope collects several photos of the pad from different viewpoints and combines them together. Unfortunately, such an approach leads to an uneven luminosity distribution over the resulting photos taken under the mix of direct and annular lights (see Figure 3.2.2). This issue was fixed by using a polarization filter, and the fourth set of pads photos was collected this way.

A 3D-printed support was created by the engineering team to fix the module position under the microscope (see Figure 3.2.3). The Keyence microscope software provides a functionality to create user profiles containing pre-saved settings such as lightning modes, position of an area needed to be photographed, etc... The photo-taking procedure is described in a document shared with other

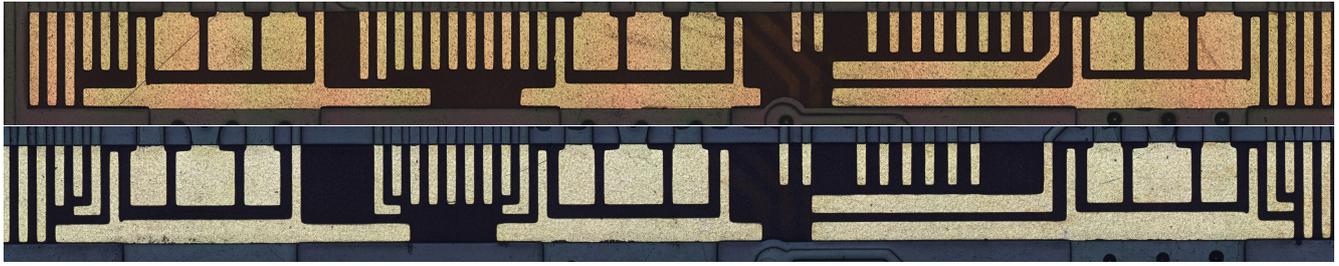


Figure 3.2.2: Top: example of pad photo taken without polarization filter. There are vertical stripes denoting different luminosity level. Bottom: example of pad photo taken with polarization filter.

ITk Pixel module production sites (see Figure 3.2.4). This simplified the photo-taking procedure and the required photos could be taken in less than 5 minutes per module.



Figure 3.2.3: The 3D-printed jig for the Pixel module fixation under the microscope.

#### Profil selection (if applicable)

1. Click on "Option" and select "User Setting...".
2. Select the profile "VI\_left\_x100\_ann" and click "Load".
3. Click "Close".

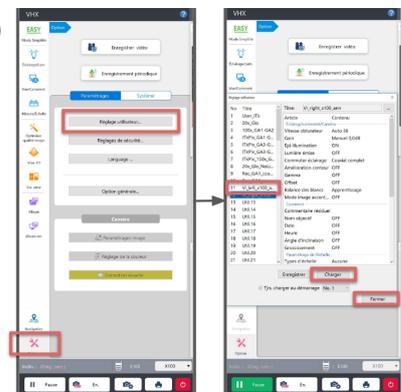


Figure 3.2.4: An example of a microscope configuration setting step described in the photo-taking procedure document.

## 3.2.2 Building datasets

The collected photos need to be processed in order to create datasets suitable for use by ML algorithms. The machine learning algorithms usually treat images part by part due to the high computational and memory resources needed to process images as a whole. We thus split the photos into smaller pieces or tiles. To organise these pieces, we adopt the MVTec dataset structure (see Section 3.2.2.1). For algorithm evaluation purposes, for each photo, a corresponding black mask that highlights the presented on the pad defects, if any, (marked with colour pixels) is needed (see Section 3.2.2.2). To reduce the complexity of the datasets, the PCB background was removed, leaving only the pad surface on the photos (see Section 3.2.2.3). The resulting sets of photos are then discussed at the end of this section.

### 3.2.2.1 MVTec dataset structure

MVTec AD [112] is a benchmark dataset for anomaly detection on images in industrial applications. It serves as a reference point for comparing the performance of many modern anomaly detection methods. The datasets for this project adopted the structure of the MVTec dataset.

The MVTec dataset is divided into defect-free training images and into a test set of images with different types of defects, as well as images that do not have any defects. The dataset has the structure presented in Figure 3.2.5. Here class1, class2,.. refer to the MVTec subsets corresponding to different categories: tile, leather, bottle, etc. Additionally, pixel-precise annotations of all anomalies are included (see Figure 3.2.6).

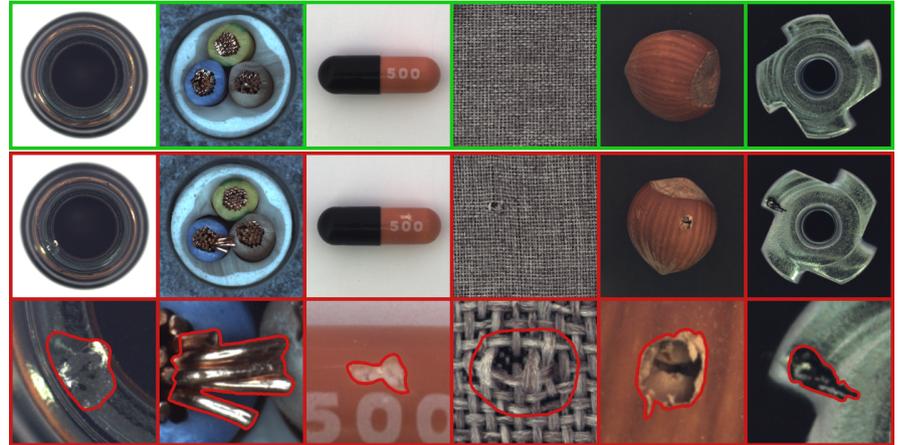
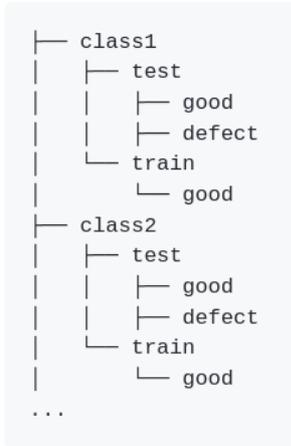


Figure 3.2.5: Folder structure of the MVTec dataset.

Figure 3.2.6: Example of the MVTec AD dataset content [112]: first row - anomaly-free samples, second row - samples with defects, third row - pixel-wise anomalies annotation.

Anomalies annotations are provided as binary masks (see Figure 3.2.7). Black mask pixels correspond to the anomaly-free areas, while white mask pixels correspond to the defects.

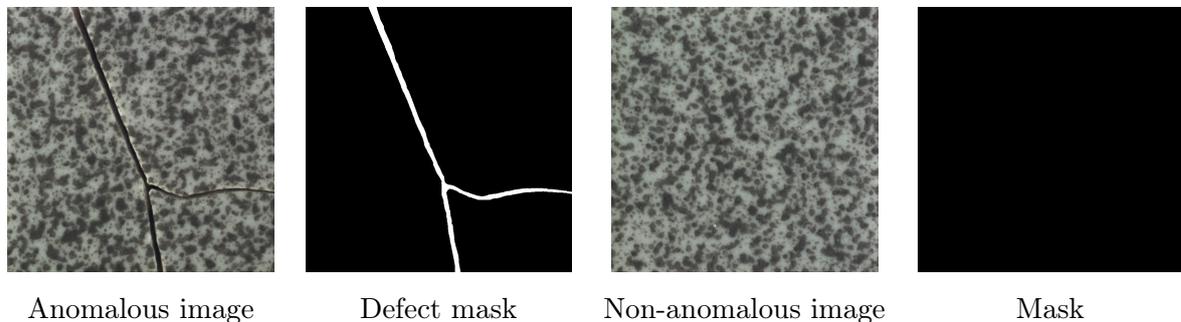


Figure 3.2.7: Examples of anomalous and non-anomalous images and their masks from *tile* subset of the MVTec dataset.

### 3.2.2.2 Annotations

The MVTec AD dataset contains pixel-precise segmentation masks for all the defects presented in the images. This allows us to evaluate the performance of anomaly segmentation algorithms.

The pad defects can be classified into three categories: chemical contaminations, scratches and dust/other. The first two groups represent the majority of the observed defects. The segmentation masks are created using the GIMP software [113]: for every photo a mask with black pixels on non-anomalous areas, red pixels on traces of chemical contamination, blue for scratches, and green for dust and other kind of defects is drawn. This procedure is done manually and is time-consuming. Figure 3.2.8 shows an example of the resulting mask.

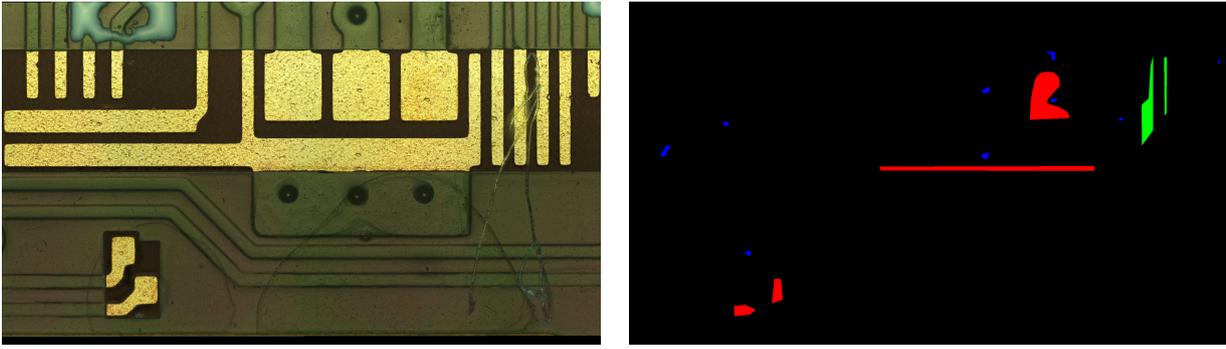


Figure 3.2.8: Example of a part of a pad photo (*left*) and its mask highlighting different types of pad’s defects (*right*). Red mask pixels correspond to traces of chemical contamination, blue to scratches, and green to dust and other kind of defects.

### 3.2.2.3 Preprocessing

The module photos and their corresponding masks were splitted into  $500 \times 500$  pixel images called tiles and compared with the MVTec AD object and textures subdatasets (see Figure 3.2.9). One can see that pads dataset differs from the MVTec in terms of complexity: the objects of interest (pads) on tiles are of various forms, and their background is not homogeneous. While the issue of uniformity cannot be addressed, simplification of the distribution that should be learned by the algorithm can be done by removing the PCB background from the pad photos. This is achievable by using the so-called Otsu’s method [114].

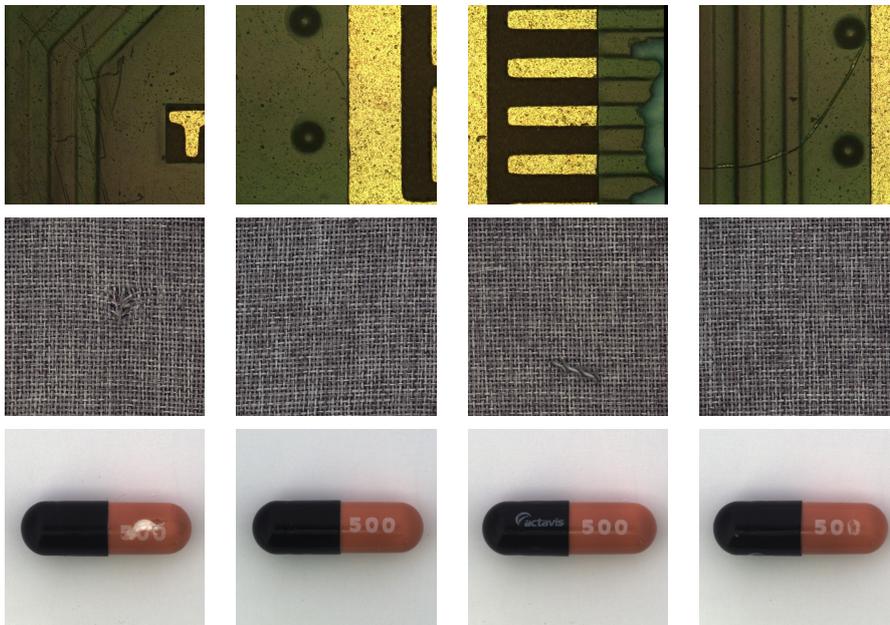


Figure 3.2.9: Example of tiles from ITkPixV1.1 photos (first row), from tile subset of MVTec (second row) and from pill subset of MVTec (third row). Unlike the tile subset, pads texture doesn’t cover the whole image, and the objects of interest are not located on a monochrome background as it is for pill subset. Thus the complexity of our dataset is higher than of MVTec.

The Otsu’s method is a widely used thresholding algorithm created for binary pixel classification to distinguish foreground from background. The algorithm is part of the classical Python image processing libraries, such as openCV [115]. The threshold is determined automatically by

maximizing inter-class variance. As both foreground and background are noised in our case, to obtain a binary mask, a Gaussian kernel needs to be applied first to remove the noise. Considering the photo size, the kernel size was chosen to be  $35 \times 35$ . The Otsu's threshold yields a binary mask, corrected by the erode and dilate operations (that remove/add pixels on the segmentation boundaries) in order to correct the sharpness of the mask borders. The last step is to apply the binary mask to the original picture.

Due to the difference in the collected photos, not all datasets were preprocessed in the same way. The photos of RD53A and ITkPixV1.1 (where the latter has a PCB version of the same ITkPix-type) went through an additional preprocessing step for background removal: threshold in HSV colour space, detailed in the next paragraph. The latest ITkPixV1.1 (with a slightly updated PCB version) photos didn't need this second thresholding step to achieve good background removal. However, a per-colour channel normalization was applied to mitigate luminosity differences in some photos.

The Otsu's threshold allows for efficiently extracting pads from photos, but it does not work perfectly on some photos: a few "artefacts" are still presented on the preprocessed photos (see Figure 3.2.10 (a.1) and (b.1)). Thus a second preprocessing step is required aiming to filter the rest of the background. One can notice that the value, or brightness, parameter in HSV colour space (Hue, Saturation, Value) differs for the pad and for part of the background. So, these entities could be separated by setting a threshold. The threshold value, however, varies for different module types (see Figure 3.2.10). The threshold value was decided to be floating and was left as a parameter available for the user to change in the developed tool. The parameter itself is defined as a quantile value on all non-black pixels presented on the image after applying Otsu's mask and setting to zero all pixels whose value is smaller than the mean one. As in the previous step, the resulting mask is then corrected by simple erode/dilate operations.

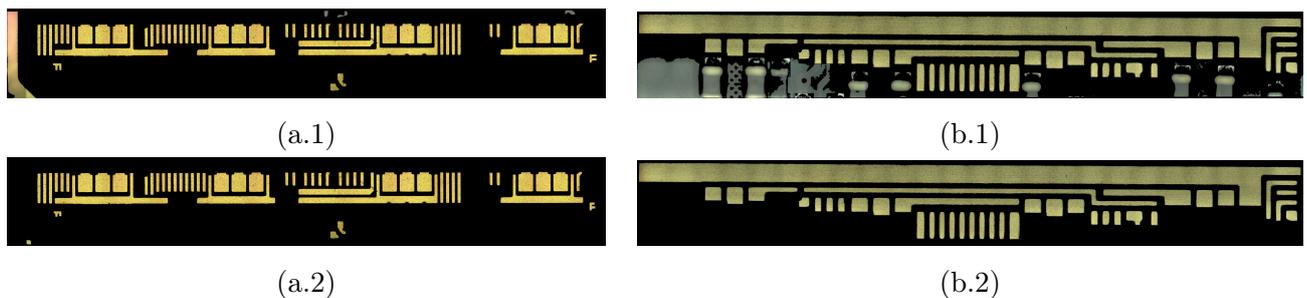


Figure 3.2.10: Examples of ITkPixV1.1 and RD53A pads preprocessing steps. (a.1) - ITkPixV1.1 module after Otsu's threshold, (a.2) - ITkPixV1.1 module after value threshold with quantile 0.7. (b.1) - RD53A module after Otsu's threshold, (b.2) - RD53A module after value threshold with quantile 0.5.

#### 3.2.2.4 Photos collections and datasets

Throughout the work on this project, we collected 4 sets of ITk pixel module pads photos.

- The first set is called RD53A and contains photos of five RD53A modules collected during the preproduction phase. The photos were taken using direct light and x150 magnifica-

tion, resulting in 4 photos per module. The background removal procedure follows 2 steps: application of Otsu's threshold and a user-defined cut on value in HSV colour space.

- The photo set named ITkPix\_Green has 6 photos collected from 3 ITkPixV1.1 modules with ITkPix-type flex. The photos were taken with the magnification x100 and the mix of direct and annular lights. Due to the change in the photo-taking procedure, we decided to separate this set from the RD53A one. Similar to the previous dataset, the ITkPix\_Green photos were preprocessed with Otsu's and value thresholds.
- The third set, ITkPix\_Feb, was collected from 5 modules of the same type as ITkPix\_Green. However, their pad surface has notably coarser granularity in comparison with previous modules, resulting in the necessity of putting them as a separate dataset. The photo-taking and preprocessing procedures were not modified.
- ITkPix\_Polar photo set contains photos from 6 ITkPixV1.1 modules with a slightly updated flex version. The pad surface on these modules has very fine granularity similar to the RD53A modules. To mitigate the issue with uneven distribution of luminosity, we used a polarisation filter to take photos for this set. This resulted in the necessity to modify the preprocessing procedure: the cut on value in HSV space was removed, while adding the per colour channel normalization. This modules version is considered to be the one used for the production and we tuned our final result on this dataset.

To create datasets for ML algorithms training from collected photos of modules' pads, we cut these photos and their corresponding masks into smaller pieces called tiles and structured them according to the reference dataset guidelines. Based on each photo set, we create several datasets varying the tile size and complexity level. The so-called filtered versions of datasets contain only the tiles with clearly visible defects for the testing part, and clean images without any imperfections for the training. The filtered datasets are supposed to have reduced complexity in comparison with the original ones: excluding the images with minor defects that could not impact wire-bonding provides more accurate non-anomalous data and simplifies the testing step. The datasets are then described by the name of the photo set, tile size, and image selection level: original or filtered.

In the following, we describe the study of different computer vision algorithms on these datasets. Since not all photo sets were available simultaneously, some experiments were conducted on certain datasets, and the conclusions made were extrapolated to other datasets.

### 3.2.3 Anomaly detection unsupervised models

Supervised and unsupervised learning are two ML paradigms. Supervised learning algorithms are trained on the examples of an input objects and corresponding desired output values. In the case of pad defects detection a supervised algorithm would receive anomalous and non-anomalous images and their corresponding masks and, based on image-mask pair, learn how to identify the presence of defects. The supervised model is required to have an ability to generalize from the training data to unseen cases, e.g. be able to segment a defect of a shape or a colour different from those

of the training examples. Unsupervised algorithms learn patterns exclusively from unlabeled data meaning receiving only objects and not pairs object-value as an input for training. For example, to perform an anomaly detection, unsupervised algorithm retrieves statistical properties from non-anomalous examples and during the inference marks any areas with statistical properties different from ones seen during the training as anomalous. The anomaly detection problems in computer vision are usually addressed with unsupervised algorithms, since this algorithms type tends to demonstrate greater performance in comparison with supervised one [116]. Indeed, we tested some supervised models (Unet [117], and a model from Ref. [118]), but did not succeed to reach good performance. Thus unsupervised anomaly detection algorithms was chosen as a main focus.

Anomalib ([119], version 0.3.7) is a deep-learning open-source library for unsupervised anomaly detection algorithms. It provides several ready-to-use implementations of anomaly detection algorithms described in the recent literature, as well as a set of tools that facilitate the development and implementation of custom models, written in Python and powered by PyTorch [120]. We use this library as a framework for our studies.

Several models showing the best performances on MVTEC dataset were investigated. The models we considered can be roughly categorised into three groups detailed below: memory-bank based (CFA [121] and Patchcore [122]), normalizing flow based [123] (FastFlow [124] and CFlow [125]) and generative adversarial network based (DRAEM [126]). All listed algorithms except of DRAEM use a pretrained network as a feature extractor. Feature extraction is a part of the dimensionality reduction process allowing to represent a given image via, often, one-dimensional vector keeping the information about distinctive structures of the image. Networks trained on a large corpora of images are usually used as the feature extractors, in our case it was ResNet50 [127] trained on a large ImageNet dataset [128]. An important property of this approach is that semantically close images usually have close representations in terms of euclidean distances between their feature vectors, allowing the comparison of images based on their feature vectors.

The idea behind memory-bank models is that multi-dimensional space of the image features of non-anomalous images is simplified to make the comparison between learnt feature distribution and the input image features faster and less memory-demanding. The normalizing flows are functions that transform the non-anomalous images features into a normal distribution, and yield outliers after receiving anomalous images as an input. The generative network can learn to extract image features via its encoder part, and to reconstruct an original input image based on this feature-vector via its decoder part. Since the anomalous images are not seen during the training, the model fails to output the same image that was given as an input, and by input-output comparison it is possible to make a conclusion about the presence of a defect. The considered models are briefly described below.

### 3.2.3.1 PatchCore

The PatchCore method [122] uses a memory bank of patch-level features extracted from pre-trained ResNet. This memory bank consists of intermediate or mid-level feature representations from ImageNet. By avoiding features that are too generic or overly biased towards ImageNet classification, Patchcore maximizes the nominal information available at testing time. To reduce

the memory bank, the Patchcore uses a coreset algorithm, that creates smaller, representative subset, or a coreset, from a large memory bank set.

During inference, the nearest neighbour search is executed on the memory bank. The image-level anomaly score is estimated by the maximum distance score between test patch-features in its patch collection and each respective nearest neighbour in the memory bank.

The overall architecture of PatchCore is depicted in Figure 3.2.11.

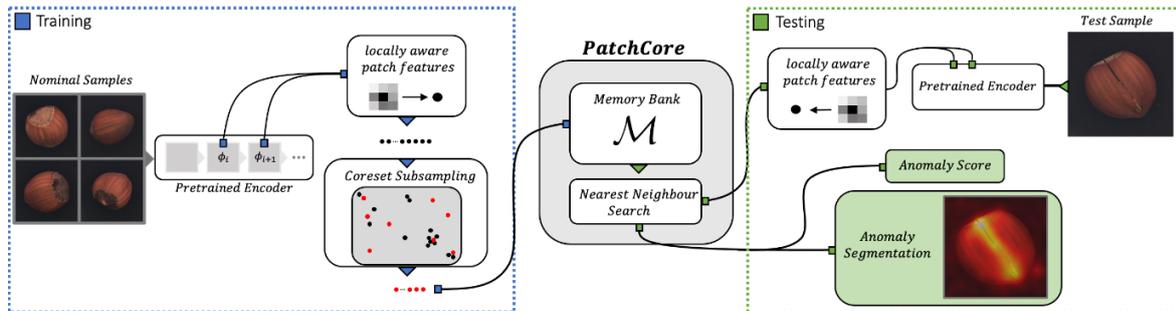


Figure 3.2.11: Schematic structure of the PatchCore algorithm [122].

### 3.2.3.2 CFA

CFA (Coupled-hypersphere-based Feature Adaptation) [121], similar to PatchCore, works by extracting features from pre-trained ResNet, and storing them in a memory bank (see Fig. 3.2.12). In the training phase, CFA employs a novel loss function to define hyperspheres with a minimum radius that contains normal features, thus clustering them. These hyperspheres are defined by their centre stored as memorized feature vectors in the memory bank. The abnormal features then can be distinguished from the learnt normal ones by their distance to hyperspheres' centres. In the test phase, CFA matches the input sample with the nearest neighbour in the memory bank and generates heatmaps indicating the degree of abnormality for every image patch based on the distance between its representation vector to hyperspheres' centres.

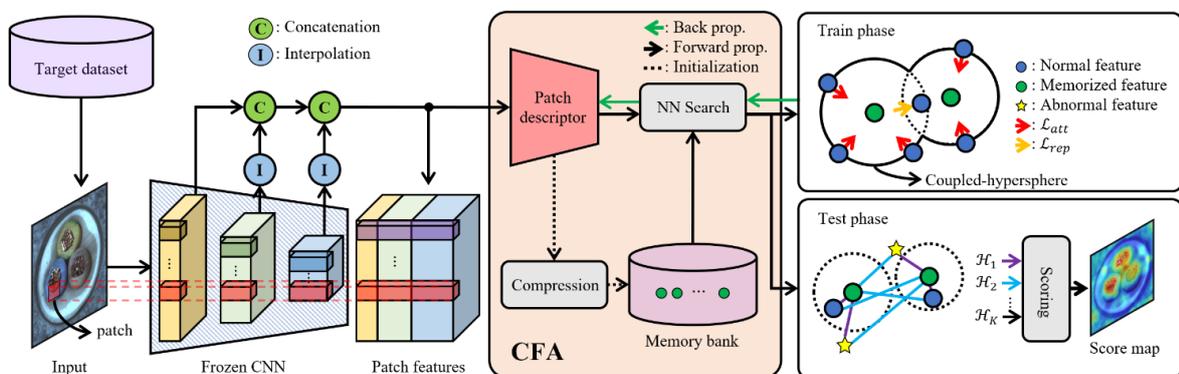


Figure 3.2.12: Schematic structure of the CFA algorithm [121].

### 3.2.3.3 FastFlow

FastFlow [124], a two-dimensional normalizing flow algorithm, detects and localizes anomalies with a two-dimensional loss function to accurately capture global and local distributions (on the levels of image details and image as a whole). During the training phase, FastFlow learns how to transform the original distribution of non-anomalous image features into a 2-dimensional standard normal distribution. In inference, the probability value of each location on the two-dimensional feature is used as the anomaly score. The lower the likelihood the image has, the higher the anomaly score it will have.

The FastFlow algorithm architecture is shown in Figure 3.2.13.

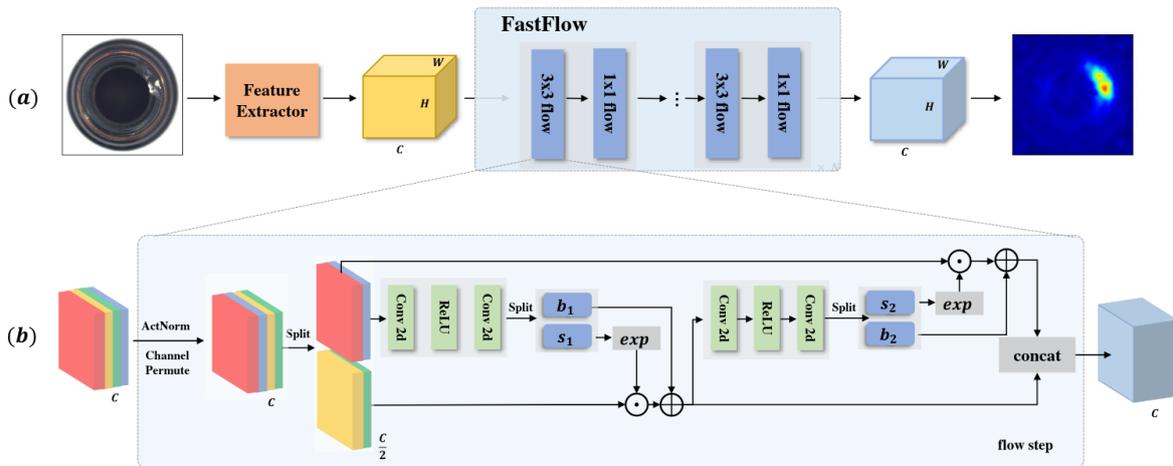


Figure 2: (a) the whole pipeline for unsupervised anomaly detection and localization in our method, which consists of a feature extractor and our FastFlow model. We can use an arbitrary network as the feature extractor such as CNN or vision transformer. FastFlow is alternately stacked by the “ $3 \times 3$ ” and “ $1 \times 1$ ” flow. (b) one flow step for our FastFlow, the “Conv 2d” can be  $3 \times 3$  or  $1 \times 1$  convolution layer for  $3 \times 3$  or  $1 \times 1$  flow, respectively.

Figure 3.2.13: Schematic structure of the Fastflow algorithm [124].

### 3.2.3.4 CFLOW-AD

CFLOW-AD is a normalizing flow model for anomaly detection with localization. Its structure is presented in Figure 3.2.14. The feature vectors extracted via ResNet are inputted into a generative decoder, which consists of multiple pooling layers. Each layer extracts features of different scales from the feature vector, resulting in a multi-scale feature pyramid. A distribution of anomaly-free image patches and their corresponding probabilities are learnt from this feature pyramid using a normalizing flow. This distribution is then transformed into a Gaussian distribution with a corresponding probability. Finally, a threshold is computed as the Euclidean distance from the distribution mean, which is used to separate in-distribution patches from out-of-distribution patches and assign the anomaly-score.

### 3.2.3.5 DRAEM

DRAEM [126] is a deep neural network that consists of two sub-networks: a reconstructive and a discriminative (see Figure 3.2.15). The reconstructive sub-network is trained to learn how to erase the defects from an image by reconstructing an image without including any anomalies.

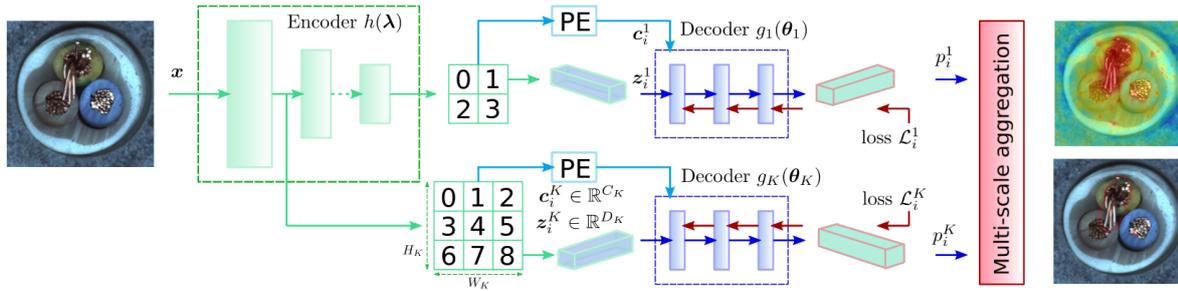


Figure 3.2.14: Schematic structure of the CFLOW-AD algorithm [125].

The discriminative sub-network is trained on both anomaly-free and synthetic anomalous images to learn how to detect and segment anomalies from the joint appearance of the original and reconstructed images. The synthetic anomalous images are generated by a conceptually simple process, providing an arbitrary amount of anomalous samples as well as pixel-perfect anomaly segmentation maps which can be used to train the proposed method without real anomalous samples. Once trained, DRAEM model produces a per-pixel anomaly detection map for anomalies localization.

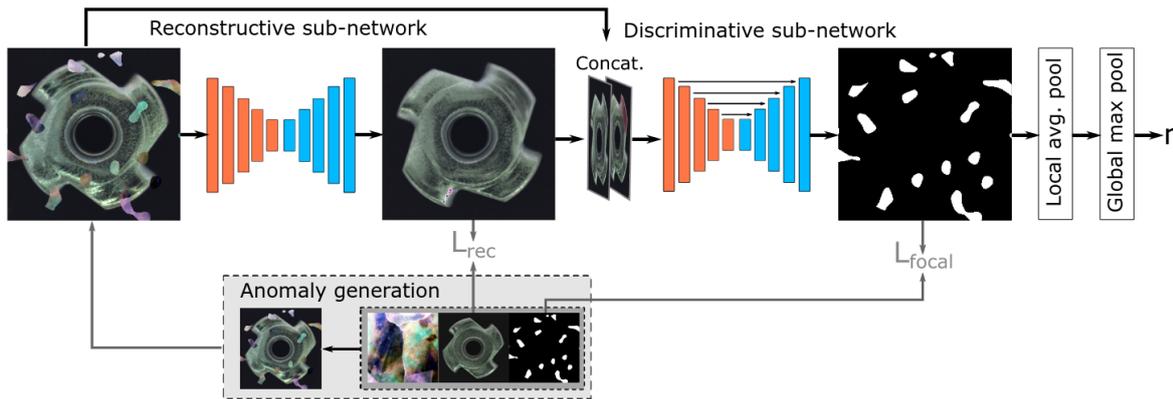


Figure 3.2.15: Schematic structure of the DRAEM algorithm [126].

### 3.2.4 Metrics

Before moving to more sophisticated measures such as area under the receiver operating characteristic (AUROC) and  $F_1$ -score that will be described below, some basic metrics are first introduced. The so-called Precision, the Recall, true positive rate (TPR) and false positive rate (FPR) are all important metrics to evaluate the performance of a machine learning model. In the context of binary classification in machine learning and statistical analysis, the terms true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) are used to describe the outcomes of a classifier's predictions compared to the actual ground truth. True positives (TP) is the number of instances where the classifier correctly predicts the positive class, while true negatives (TN) is the number of correctly classified objects of negative class. False positive (FP) gives the number of negative class object classified as belonging to positive class, and false negative (FN) is the number of instances where the classifier incorrectly predicts the negative class.

The Precision is defined as the ratio of true positives to the total number of predicted positives (TP+FP), and is typically written as

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (3.2.1)$$

The Recall, also known as true positive rate (TPR), is the ratio of true positives to the total number of actual positives (TP+FN), and is written as

$$\text{Recall} = \text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (3.2.2)$$

FPR is the ratio of false positives to the total number of actual negatives, and is written as

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}. \quad (3.2.3)$$

The Area Under the Curve (AUC) is a measure of the performance of an anomaly segmentation model, which can be calculated at both the image-level and pixel-level. At the image-level, the AUROC is calculated as the average of the true positive rate (TPR) and the false positive rate (FPR) over the  $N$  images in the dataset:

$$\text{AUROC}_{\text{image}} = \frac{1}{N} \sum_{i=1}^N (\text{TPR}_i + \text{FPR}_i). \quad (3.2.4)$$

At the pixel-level, the AUROC is similarly calculated but over the  $M$  pixels in the dataset:

$$\text{AUROC}_{\text{pixel}} = \frac{1}{M} \sum_{i=1}^M (\text{TPR}_i + \text{FPR}_i). \quad (3.2.5)$$

The AUROC metric varies between 0 and 1, with 0.5 indicating a random classifier, and 1 indicating a perfect classifier. It is advantageous because it is independent of the operating point of the segmentation algorithm, meaning that it is not affected by the threshold used to classify the two classes. The main drawback of AUROC includes its sensitivity to class imbalance, which might be important in evaluating the pixel-level performance.

The image-level  $F_1$ -score is calculated as the harmonic mean of the Precision and the Recall for a given image, as expressed in Eq. 3.2.6:

$$F_{1,\text{image}} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (3.2.6)$$

Analogously, the pixel-level  $F_1$ -score is calculated over all the pixels in a given image:

$$F_{1,\text{pixel}} = 2 \cdot \frac{\frac{1}{L} \sum_{i=1}^L \text{Precision}_i \cdot \frac{1}{L} \sum_{i=1}^L \text{Recall}_i}{\frac{1}{L} \sum_{i=1}^L \text{Precision}_i + \frac{1}{L} \sum_{i=1}^L \text{Recall}_i} \quad (3.2.7)$$

where  $L$  is the number of pixels in the image and the  $\text{Precision}_i$ ,  $\text{Recall}_i$  are the precision and recall of the  $i$ -th pixel.

The  $F_1$ -score varies in the interval  $[0, 1]$ , with higher values indicating better performance. It

is good for evaluating models with imbalanced class distributions, but can be affected by outliers or rare classes.

In the following, we evaluated the models' performances using AUROC and the  $F_1$ -score both on image and pixel levels.

### 3.2.5 Comparison of unsupervised models performance

To narrow down the list of unsupervised models to study, we tested the models performances on different datasets. The models mentioned above were trained with the RD53A and ITkPix\_Green datasets with tile size  $500 \times 500$  like in the MVTec dataset and  $256 \times 256$  like the default image size in Anomalib. The original and filtered datasets versions are considered. Table 3.2.1 presents the overview of considered datasets, showing the number of images in each dataset category.

	Train	Test				
		Total	Good	Chem.cont.	Scratch	Dust
RD53A, 500, original	268	1062	194	646	185	37
RD53A, 500, filtered	187	147	57	75	15	0
RD53A, 256, original	913	1471	269	946	218	38
RD53A, 256, filtered	167	132	84	43	5	0
ITkPix_Green, 500, filtered	167	189	57	84	43	5
ITkPix_Green, 256, filtered	203	78	29	27	21	1
ITkPix_Polar, 256, filtered	167	78	57	56	60	46
MVTec AD subset, average	242	115	23	23	23	23

Table 3.2.1: Overview of the datasets with the number of images in all categories.

The performances of unsupervised anomaly detection models on the described above datasets are summarised in Table 3.2.2. Due to time and computational resource limitations, the results presented correspond to one training. The results may vary with each subsequent training, given the random nature of neural networks. The performance of FastFlow on the RD53A original dataset with tile size of 500 is not presented in the table since its training required too many computation resources. For the same reason, the CFA model was trained with the  $500 \times 500$  px tiles rescaled to  $400 \times 400$ . The studied models don't have many hyperparameters to tune, and after several experiments that were not reported in this work, we concluded that the default hyperparameter values seem to work best. CFA and Patchcore consistently demonstrate the best performance, with CFA having better results on pixel-level predictions and Patchcore on image-level. These two models are selected for further studies on the RD53A, ITkPix\_Green and ITkPix\_Feb datasets. Patchcore showed the best performance on the ITkPix\_Polar dataset.

We note that a lower performance at the image level usually implies more false positive predictions, that is to say, the model classifies more non-anomalous images as anomalous ones. This type of error is preferable to false negatives, as we would prefer to detect false anomalies rather than to miss one.

		CFA	Patchcore	FastFlow	CFlow	DRAEM
		image/pixel	image/pixel	image/pixel	image/pixel	image/pixel
RD53A, 500x500, original	AUROC	0.54/0.77	0.65/0.79		0.68/0.70	0.60/0.72
	$F_1$	0.90/0.22	0.90/0.13		0.91/0.09	0.90/0.14
RD53A, 500x500, filtered	AUROC	0.52/0.89	0.80/0.89	0.72/0.80	0.68/0.79	0.59/0.72
	$F_1$	0.78/0.25	0.86/0.24	0.78/0.21	0.77/0.17	0.77/0.10
RD53A, 256x256, original	AUROC	0.54/0.77	0.57/0.76	0.63/0.72	0.53/0.69	0.52/0.79
	$F_1$	0.90/0.22	0.90/0.21	0.90/0.19	0.90/0.17	0.90/0.26
RD53A, 256x256, filtered	AUROC	0.48/0.79	0.55/0.77	0.67/0.72	0.47/0.72	0.55/0.56
	$F_1$	0.89/0.22	0.89/0.21	0.89/0.20	0.89/0.19	0.89/0.15
ITkPix_Green, 500x500, filtered	AUROC	0.71/0.88	0.71/0.86	0.55/0.87	0.46/0.82	0.61/0.81
	$F_1$	0.85/0.11	0.85/0.11	0.84/0.12	0.84/0.08	0.79/0.18
ITkPix_Green, 256x256, filtered	AUROC	0.70/0.91	0.75/0.74	0.45/0.67	0.49/0.63	0.66/0.80
	$F_1$	0.75/0.39	0.82/0.14	0.78/0.12	0.77/0.10	0.77/0.17
ITkPix_Polar, 256x256, filtered	AUROC	0.83/0.90	0.92/0.86			
	$F_1$	0.87/0.36	0.92/0.36			

Table 3.2.2: Summary of the performances of several models for the different datasets. In each row, the AUROC is presented in the top line (image-level AUROC/ pixel-level AUROC), and the  $F_1$ -score in the bottom line (image-level  $F_1$ -score/ pixel-level  $F_1$ -score). For each dataset, the model best scores at the image-level are highlighted in red and at the pixel-level in blue.

### 3.3 Performance and limitations studies

#### 3.3.1 Augmentation and preprocessing studies on selected unsupervised models

After selecting the promising models, we studied the impact of the different dataset parameters and modifications on models' performances. Since both CFA and Patchcore have similar performances and the concept of work, we assume that the conclusions made about their performances are interchangeable. We use CFA or Patchcore for tests on a dataset, depending on which model showed the best performance on the filtered version of the considered dataset. We mainly used CFA for RD53A and ITkPix\_Green datasets, and Patchcore for the rest.

##### 3.3.1.1 Tile size

The results from Table 3.2.2 do not clearly explain which tile size works best for the CFA and Patchcore models. The ITkPix\_Green photos were split into  $512 \times 512$  tiles and filtered, then these tiles were split to  $256 \times 256$  and  $128 \times 128$  px. This ensures that the datasets contain the same amount of information differing only in tile size. Some all-black tiles appeared while splitting bigger tiles into smaller ones and were removed from the datasets. The summary of the number of images in each category is shown in Table 3.3.1.

The CFA performance is evaluated on the created datasets. The tiles of size 512 were rescaled into tiles of  $400 \times 400$  and  $256 \times 256$ , due to the large amount of computational resources required to train on the original size tiles. The results are summarised in Table 3.3.2. We observe the

	Train	Test				
		Total	Good	Chem.cont.	Scratch	Dust
ITkPix_Green, 512, filtered	167	189	57	84	43	5
The above split into 256	393	646	156	298	172	20
The above split into 128	1202	2539	427	1344	688	80

Table 3.3.1: Overview of the datasets for checking the tile size influence on the CFA performance.

best performance on the datasets with tiles of size 512 rescaled to 400 and with  $256 \times 256$  pixels tiles. We choose to work with tiles of size 256: not loosing in performance in comparison with tiles of 512 pixels, this setting requires considerably less computational resources. The rescaling of  $512 \times 512$  to  $256 \times 256$  might lead to losing some information and, thus, to weaker performance in comparison to model trained on bigger images. The setting with tiles size 128 contains a lot of false positive predictions, that, along with a bigger test set, gives a drop in the AUCROC metric at the image-level.

	512 to 400	512 to 256	256	128
	image/pixel	image/pixel	image/pixel	image/pixel
AUROC	0.65/0.90	0.60/0.79	0.65/0.89	0.38/0.86
$F_1$	0.84/0.14	0.83/0.07	0.84/0.13	0.91/0.10

Table 3.3.2: CFA model performance on datasets with different tile sizes based on the ITkPix\_Green photo set.

### 3.3.1.2 Impact of filtering for complexity reduction

From Table 3.2.2 we observe that the pixel-level performances are better for the filtered datasets, while the image-level performances are higher for the original datasets. This could be attributed to the fact that the original datasets include many images mostly containing the background without much pads areas and, thus, are easily classified as non-anomalous by a model.

To understand the impact of filtering described in Section 3.2.2, we compared the performance of models trained on filtered and on original training sets from ITkPix\_Green photos split to tiles of 256 pixels. We run a model trained on the original datasets, tested against both the original and filtered test subsets, and a model trained on filtered datasets, tested against both the original and filtered datasets. The results of these tests are presented in Table 3.3.3.

The model trained on a filtered dataset provides the best results on both versions of the testing set, especially at pixel-level. The filtered set used for training might provide less varying distribution of non-anomalous images features resulting in more accurate distinction between anomalous and non-anomalous areas on the testing step. Therefore, we choose to use the filtered datasets for training.

			Test	
			original	filtered
			image/pixel	image/pixel
Train	original	AUROC	0.63/0.79	0.61/0.80
		$F_1$	0.64/0.05	0.80/0.27
	filtered	AUROC	0.64/0.82	0.70/0.91
		$F_1$	0.66/0.12	0.78/0.39

Table 3.3.3: CFA model performance on datasets with different selectivity levels, original and filtered, on ITkPix\_Green photos and tile size 256.

### 3.3.1.3 Augmentations: basic spatial

The performance of a model is expected to depend on the size of the training dataset. To check this, a fixed test dataset of filtered ITkPix\_Green tiles of 256 px was used while varying the size of the training dataset by deleting random samples or applying spatial augmentations such as cropping, padding, and flipping (we used the Albumentation [129] library to apply augmentations). The result of these trials is presented in Table 3.3.4. It demonstrates that the larger the training dataset is, the better the model performs on the test dataset. However, augmentations can only be applied up to a certain point until the model can not extract any more additional information from the augmentations: the performance of models trained on a doubled training set is the same as on the one increased by a factor of 1.5.

	-50%	-25%	nominal	+25%	+50%	+100%
	image/pixel	image/pixel	image/pixel	image/pixel	image/pixel	image/pixel
AUROC	0.74/0.90	0.76/0.91	0.70/0.91	0.79/0.92	0.77/0.93	0.76/0.93
$F_1$	0.82/0.35	0.83/0.38	0.78/0.39	0.83/0.42	0.83/0.45	0.81/0.44

Table 3.3.4: CFA model performance on ITkPix\_Green 256 datasets with different training set sizes. Here  $\pm N\%$  refers to the decreasing/augmentation of the nominal dataset on  $N\%$  of its initial size.

### 3.3.1.4 Colour preprocessing

In order to investigate if the colour preprocessing could improve the anomaly detection performance, we studied two preprocessing types. The experiments were performed using the ITkPix\_Polar dataset with filtered tiles of size  $256 \times 256$  pixels and CFA model.

The first preprocessing approach tested was contrast enhancement. Colours can be accentuated by applying a histogram equalization per colour channel, increasing the intensity difference between the light and dark areas. This procedure helps to make chemical contamination defects more visually prominent (see Figure 3.3.1). Table 3.3.5 shows that contrast enhancement indeed influences the CFA performance resulting in more accurate results.

The second tested approach is to try to code as much information about the defects as possible into colour channels. The images were processed as follows:

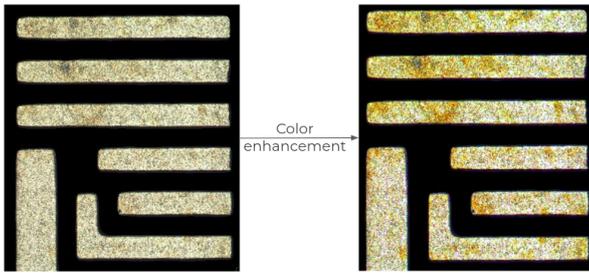


Figure 3.3.1: Example of how contrast enhancement makes chemical contamination defects more visible (on ITkPix\_Polar photo).

		CFA
		image/pixel
ITkPix_Polar, 256, filtered	AUROC	0.50/0.78
	$F_1$	0.95/0.38
Contrast enhancement	AUROC	0.61/0.83
	$F_1$	0.95/0.41

Table 3.3.5: Performance of the CFA model on filtered ITkPix\_Polar tiles of size 256 and its contrast enhanced version.

- The first colour channel contains the grayscale image of the described above colour enhanced version of the original tile.
- The original image under a colour threshold making dust more noticable.
- The Sobel edge-detection filter [130] applied to the original image highlighting the scratches

An example of a per channel view and the resulting tile used for training is shown in Figure 3.3.2. Table 3.3.6 shows the influence of applying such preprocessing on the model performance. We observed an improvement on the image-level performance, but the drop of the performance in term of both evaluated metrics on pixel-level. This preprocessing approach was excluded from the consideration.

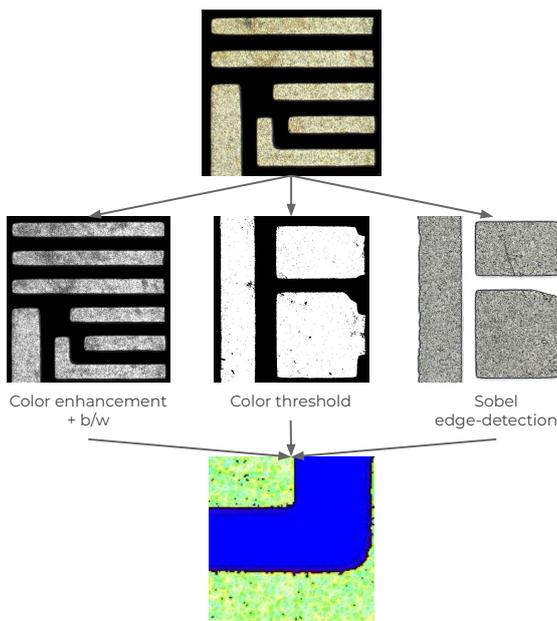


Figure 3.3.2: The structure of three-channel encoding (on ITkPix\_Polar photo).

		CFA
		image/pixel
ITkPix_Polar, 256, filtered	AUROC	0.50/0.78
	$F_1$	0.95/0.38
Per channel preprocessing	AUROC	0.78/0.81
	$F_1$	0.87/0.32

Table 3.3.6: Performance of the CFA model on filtered ITkPix\_Polar tiles of size 256 and its per-channel processed version.

### 3.3.1.5 Augmentations: custom spatial

Apart from the standard augmentations, we implemented a custom one. The idea is to extract the pad parts from the tiles from the whole dataset and randomly combine them putting in black image ensuring there's no overlapping and that the pad parts are separated by at least 10 pixels wide gap. The pad parts are also randomly flipped and 90 degrees rotated before being placed into a new image. This augmentation doesn't bring new information about pads surface but augments the diversity in pad shapes. An example of the such augmentation is shown on Figure 3.3.3. As shown in Table 3.3.7, the augmentation doesn't contribute much to the performance increase. Creating too many training augmented images drops the performance, that is in agreement with the observation made for CFA model and discussed above.

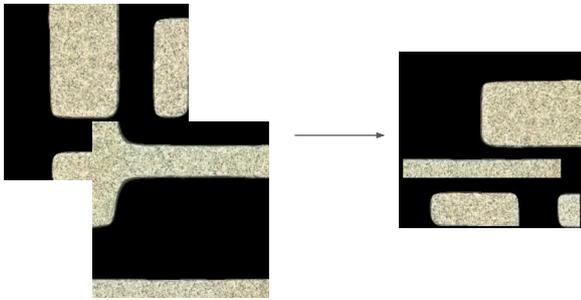


Figure 3.3.3: Example of how the pads from dataset tiles are recombined on a new image.

		Patchcore image/pixel
ITkPix_Polar, 256, filtered	AUROC	0.92/0.86
	$F_1$	0.92/0.36
Pads: 100	AUROC $F_1$	0.93/0.86 0.92/0.37
Crop: 50		
Flip: 217		
Default: 167		
Pads: 1700	AUROC $F_1$	0.56/0.83 0.73/0.12
Crop: 500		
Flip: 1300		
Default: 167		

Table 3.3.7: Performance of the Patchcore model on filtered ITkPix\_Polar tiles of size 256 and augmented by different number of images versions.

### 3.3.1.6 Dataset complexity: MVTec tile subset

The performances of the CFA and Patchcore models are not as good as for the subdatasets of MVTec. The assumption is that pad datasets have higher complexity than the MVTec, since on the pad tiles not only the pad surface is presented, but also a black background, and the pads are all of different shape. Since it is not possible to create a dataset with tiles containing only pad surface with tile size 256, the opposite direction was chosen. We can artificially introduce the complexity coming from pad shapes to the MVTec tile dataset (the one the most similar to pads dataset) to check if the difference in performance is caused by non-uniform pad shapes. Mimicking the MVTec tile dataset to pad dataset can be done randomly placing black rectangles on the MVTec tile training and testing images (see Figure 3.3.4). If the performance on such modified dataset drops, it would mean that the reason we do not have a high performance on the pads is the pads shapes. We tested the CFA model on the original MVTec tile dataset, and its modified training dataset versus the original and modified test sets. Table 3.3.8 summarises the results. We can conclude that the performance is not affected by the pad shapes. So the reason

of not reaching the performance on the pad datasets as on MVTec dataset should come from the pad texture itself.

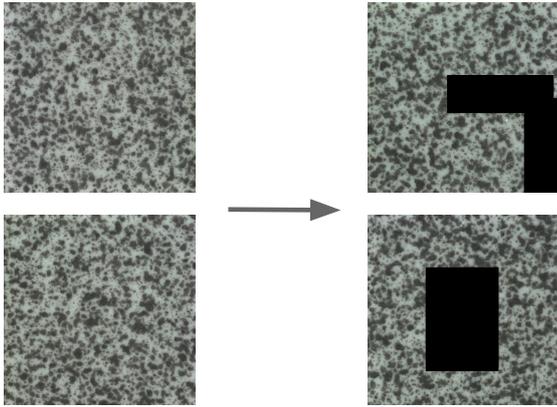


Figure 3.3.4: Example of increasing MVTec tile dataset by mimicking the pad shapes.

		CFA
		image/pixel
MVTec tile	AUROC	1.00/0.98
	$F_1$	1.00/0.76
MVTec tile train on imitating pad shape	AUROC	1.00/0.99
	$F_1$	0.98/0.79
MVTec tile train on pad shape imitations test on pad shape imitations	AUROC	1.00/0.99
	$F_1$	0.97/0.79

Table 3.3.8: Performance of the CFA model on MVTec tile dataset and its modifications.

### 3.3.1.7 Dataset complexity: denoising

To address the question of the pad texture impact on model performances, we tried to simplify the texture by a denoising algorithm. The idea is that the pad surface can be represented as a kind of a noise that could be filtered out making the defects obviously visible. The denoising is done via Isotropic Undecimated Wavelet Transform also called the starlet transform [131]. The starlet transform is a special wavelet transform: isotropic, preserving the translation invariance and allowing for the simultaneous processing of data at different scales. These properties make the starlet transform widely used in astronomical image processing to separate signal sources from background noises. The idea is to represent a two-dimensional array, or an image, denoted  $I$ , as a sum of a coarse resolution image  $c_J$  and wavelet coefficient images  $w_j$ :

$$I(x, y) = c_J(x, y) + \sum_{j=0}^{j_{max}} w_j(x, y). \quad (3.3.1)$$

Each wavelet image  $w_j$  carries the details of  $I$  at the scale of  $2^j$  pixels, while the coarse image  $c_J$  - at scale  $J = j_{max} + 1$ . The idea of the decomposition procedure is the iterative application of the starlet filter:

1. Start with initial image  $I$  and  $j = 0$ .
2. Convolve with the special filter  $h$  to get the next scale:  $c_j = h * I = h * c_{j-1}$ .
3. Compute wavelet coefficients:  $w_j = c_{j-1} - c_j$ .

At the end the residual coarse image is computed at final scale:  $c_J = h * c_{J-1}$ .

An example of a decomposed tile with a defect from the ITkPix\_Polar dataset is presented on Figure 3.3.5. Conceptually, the smaller scales correspond to finer details, and the first scales correspond to the pad surface noise. Larger scales can catch bigger objects, such as pad defects.

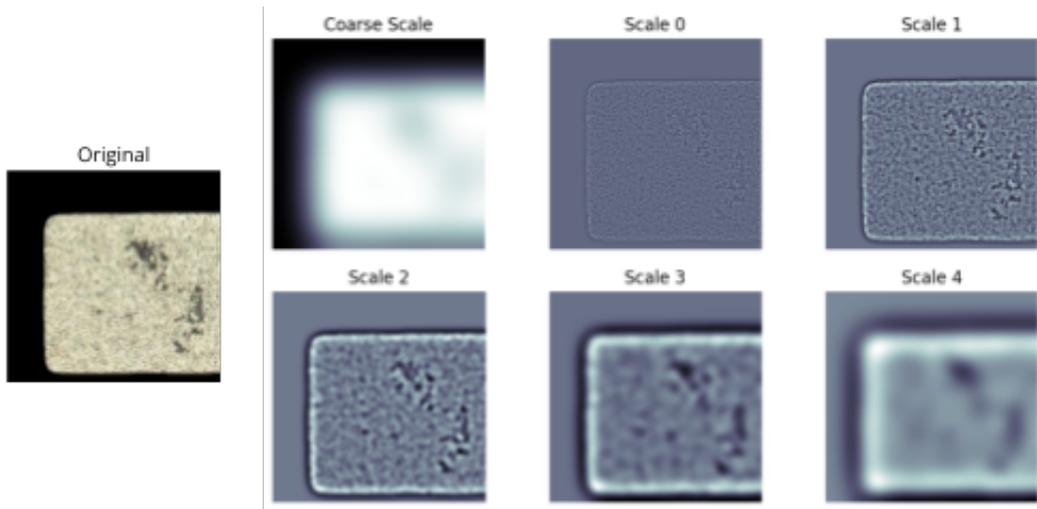


Figure 3.3.5: Starlet decomposition of a pad image with a defect at five scales ( $J = 5$ ). The smaller scales catch the pad surface structure considered as noise while the higher scales make the defect visible.

The denoised image can be obtained from this decomposition by summing only a significant wavelet coefficient. To distinguish the significant wavelet coefficient from those corresponding to noise, a  $w_j$  is compared to a threshold  $K\sigma$  with  $\sigma$  being the noise standard deviation. The wavelet coefficients not passing the threshold are put to 0 enforcing the  $w_j$  to be sparse. The parameters of the denoising algorithm are varied like the number of scales  $J$  and the factor of the sparsity threshold  $K$ . Higher values of the parameter  $K$  smooth the denoised image more.

We consider three denoising settings:  $J = 2, K = 1$ , then  $J = 4, K = 1$ , and  $J = 4, K = 3$ . Examples of the denoised tiles for each setting are shown on Figure 3.3.6. Visually, it seems that  $J = 4, K = 1$  provides the optimal result filtering out most of the non-anomalous pad surface while leaving the defects visible.

Table 3.3.9 shows the results of the Patchcore model on the filtered ITkPix\_Polar 256 dataset referred to as default, and its modifications. We tested the denoised version of the default dataset, a denoised version of the dataset after colour preprocessing and a denoised version of the default dataset after spatial pad augmentation. The denoising with  $J = 4, K = 3$  caused considerable drop in image-level performance and was excluded from consideration. The  $J = 2, K = 1$  and  $J = 4, K = 1$  give a slight improvement:  $J = 2, K = 1$  at image level and  $J = 4, K = 1$  at pixel-level.

### 3.3.1.8 Augmentations: noise generation

Instead of modifying the existing pad texture, the texture can be simulated producing more non-anomalous images for algorithm training. The wavelet decomposition can be used for texture synthesis. With the help of Wavelet Phase Harmonics (WPH), we generated non-anomalous pad images to augment the ITkPix\_Polar training dataset. Similarly to the decomposition based on the starlet transformation, a two-dimensional array can be represented as a sum of wavelet coefficients, defined by the filter  $h$  (different from value used in the starlet transformation) catching the phase-related features. WPH is used for image pattern recognition, texture analysis, and

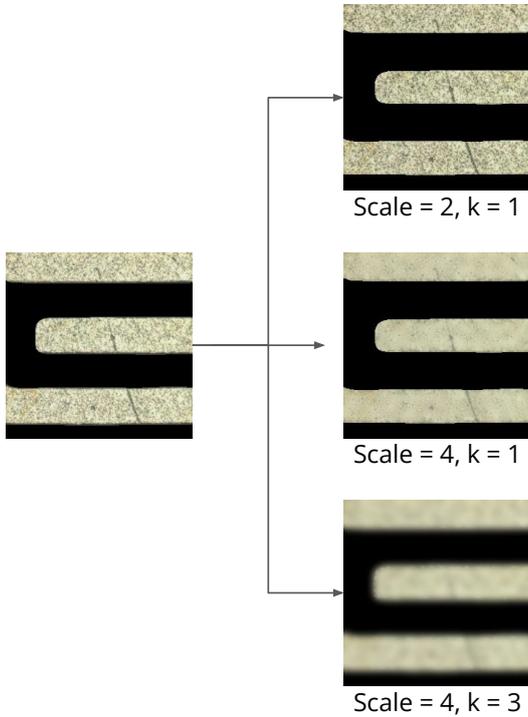


Figure 3.3.6: Example of denoising results with different parameters (with an ITkPix\_Polar 256 px tile).

		Patchcore
		image/pixel
ITkPix_Polar, 256, filtered	AUROC	0.92/0.86
	$F_1$	0.92/0.36
Denoised	AUROC	0.94/0.86
Scale = 2, k = 1	$F_1$	0.94/0.35
Denoised	AUROC	0.92/0.87
Scale = 4, k = 1	$F_1$	0.92/0.38
Denoised	AUROC	0.84/0.86
Scale = 4, k = 3	$F_1$	0.87/0.35
ITkPix_Polar, 256, filtered + contrast enhancement	AUROC	0.93/0.86
	$F_1$	0.93/0.36
Denoised	AUROC	0.93/0.86
Scale = 2, k = 1	$F_1$	0.93/0.36
Denoised	AUROC	0.92/0.87
Scale = 4, k = 1	$F_1$	0.91/0.38
ITkPix_Polar, 256, filtered + spatial pad augmentations	AUROC	0.93/0.86
	$F_1$	0.92/0.37
Denoised	AUROC	<b>0.95/0.86</b>
Scale = 2, k = 1	$F_1$	<b>0.94/0.36</b>
Denoised	AUROC	<b>0.93/0.87</b>
Scale = 4, k = 1	$F_1$	<b>0.93/0.39</b>

Table 3.3.9: Performance of the Patchcore model on the filtered ITkPix\_Polar 256 dataset referred to as default, and its modifications. The best scores at the image-level are highlighted in red and at the pixel-level in blue.

feature extraction.

The idea of texture synthesis with WHP is that two images with matching statistical properties of WPH coefficients  $w$  have similar textures. We start with a reference image  $I_{ref}$  and an image generated with Gaussian noise  $I_{syn}$ . The WPH coefficients are extracted via the PyWHP library. Iteratively, the  $w_{syn}$  coefficients of  $I_{syn}$  are adjusted via backpropagation in a way to minimise the objective function:

$$L = \frac{\sum_i |w_{ref,i} - w_{syn,i}|^2}{\sum_i |w_{ref,i}|^2}. \quad (3.3.2)$$

The Limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm (L-BFGS-B [132]) algorithm is used as an optimizer, and the process is repeated for 50 iterations. The generated texture is then combined with one of the tiles from the dataset to mimic the shape of the pad. Then spatial augmentations, such as cropping, padding and flipping, are applied to modify the pad segment position. An example of generated images is shown in Figure 3.3.7. Note that the WHP-based synthesis can generate a two-dimensional array corresponding to a grayscale image. Both CFA and Patchcore algorithms work with grayscale images, so this is not a limitation. The synthesis

of 3-colour channel images is not straightforward: it is not possible to consider the 3 channels independently since they carry information about the same texture. It might be possible to use some dimension reduction techniques such as PCA [133], principal components analysis, to first transform 3-channel image into 1-channel, perform the synthesis based on it, and then apply the inverse dimensional transformation onto the generated images. This approach was not studied in this work.

Table 3.3.10 shows the Patchcore performance on the ITkPix\_Polar dataset augmented with the generated images. This performance does not differ from the performance on the unaugmented version of the considered dataset. A model trained only on the synthetic images still shows a non-random performance on the test subset of the filtered ITkPix\_Polar. We can conclude that the image generation procedure partially captures the feature distribution of non-anomalous images but cannot bring any new information for the training necessary to improve performance.

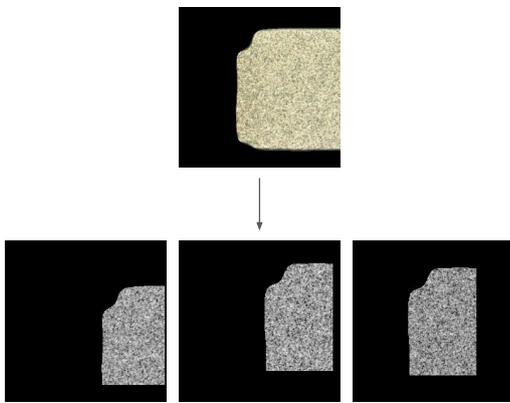


Figure 3.3.7: An example of synthetic texture augmentation, where texture of pads from bottom images is generated, and pad shape is taken from original images and modified.

		Patchcore
		image/pixel
ITkPix_Polar, 256, filtered	AUROC	0.92/0.86
	$F_1$	0.92/0.36
ITkPix_Polar, 256, filtered + 230 synthetic images	AUROC	0.92/0.86
	$F_1$	0.92/0.36
250 synthetic images	AUROC	0.74/0.77
	$F_1$	0.86/0.22

Table 3.3.10: Performance of the Patchcore model on the filtered ITkPix\_Polar 256 dataset, its version with training set augmented by synthetic images, and a dataset containing only synthetic images as train set and test set from filtered ITkPix\_Polar.

### 3.3.2 Adaptable solution for the visual inspection

A CFA or Patchcore model trained on one dataset does not perform well on another: images taken with different photo-taking conditions or showing pads with different granularity have distinct representations of non-anomalous texture. To make the automatic visual inspection tool usable across production sites, where identical conditions cannot be guaranteed, the final solution should be able to handle these differences. Creating a model capable of yielding accurate predictions for different image sets is a complex problem we haven't found a solution for. Instead, we propose a pipeline for easy model retraining.

The Patchcore model performs well within the few-shot learning setting [134]. Few-shot learning is a framework in which a model is trained on a very small number of labelled examples. We show that Patchcore model trained with 10 non-anomalous images augmented with spatial augmentations reaches a performance comparable with the performance of Patchcore model trained on a filtered dataset containing more than training 100 images. This approach is tested on datasets

based on ITkPix\_Polar and ITkPix\_Feb, the results are shown in Table 3.3.11.

		Patchcore image/pixel
ITkPix_Polar, 256, filtered	AUROC	0.92/0.86
	$F_1$	0.92/0.36
ITkPix_Polar, 256, filtered Spatial pad augmentations Denoised: Scale= 4, k= 1	AUROC	0.93/0.87
	$F_1$	0.93/0.39
ITkPix_Polar, 256, only 20 training images	AUROC	0.91/0.85
	$F_1$	0.93/0.36
ITkPix_Polar, 256, only 10 training images	AUROC	0.74/0.80
	$F_1$	0.88/0.30
ITkPix_Polar, 256, only 10 training images Spatial pad augmentations	AUROC	0.90/0.84
	$F_1$	0.91/0.34
ITkPix_Polar, 256, only 10 training images Spatial pad augmentations Denoised: Scale= 4, k= 1	AUROC	0.85/0.85
	$F_1$	0.87/0.34
ITkPix_Feb, 256, filtered	AUROC	0.63/0.84
	$F_1$	0.73/0.15
ITkPix_Feb, 256, only 10 training images	AUROC	0.54/0.80
	$F_1$	0.73/0.11
ITkPix_Feb, 256, only 10 training images Spatial pad augmentations	AUROC	0.56/0.83
	$F_1$	0.73/0.12

Table 3.3.11: Performance of Patchcore model in few-shot learning setting on ITkPix\_Polar and ITkPix\_Feb based datasets with tile size 256.

To adapt the automatic visual inspection tool in case of differences in photo-taking procedure or pads' granularity, a production site needs to collect 10 images of size  $256 \times 256$  containing no anomalies. These images are then used to retrain a Patchcore model. This approach allows the use of the developed instruments for visual inspection without spending much time on collecting and labelling datasets and is the solution suggested for the automatic visual inspection of pads.

### 3.4 Graphical User Interface

To make the algorithms accessible to the wire-bonding teams across the different production sites, we developed a Graphical User Interface (GUI) with an intuitive interface for interaction with the algorithm. This GUI eliminates the need for users to interact with the algorithm source code, providing a wrapper for the algorithm and all relevant connections to databases for data storage.

Among the different forms of GUI, a web application was chosen for this purpose for several reasons. Its portability and accessibility from any internet-connected device ensures easy

maintenance and easy updating process that doesn't require any additional actions from users. Furthermore, the central access to the local GPU farm, mongoDB and EOS space at CERN's lxplus can be set up.

A GUI is usually divided into two parts, the frontend and the backend (see Figure 3.4.1). The frontend is the user interface responsible for taking user input and displaying information. The backend is the application that processes the user input.

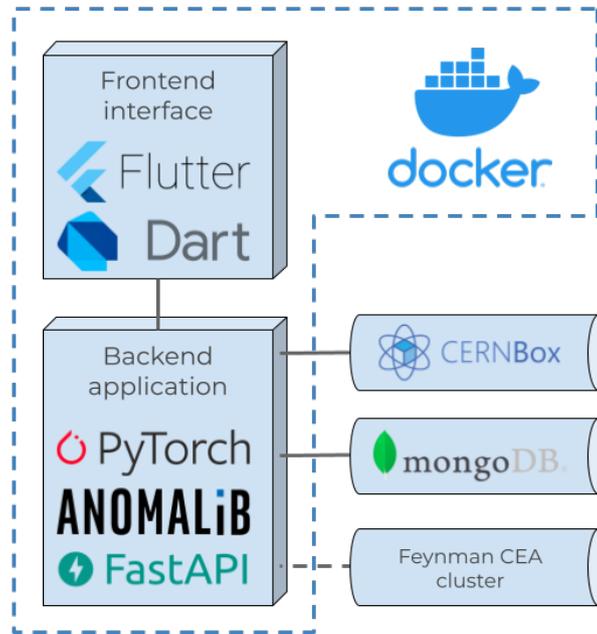


Figure 3.4.1: Schematic structure of the GUI.

The backend of the system takes two photos of pads per module, one for each side, as input. These photos are then processed individually, starting with the background removal procedure described in Section 3.2.2. The photos are padded with black pixels until their size can be evenly divided by the size of the input image of the neural network. Each photo is then split into tiles and given to the model inferencer, which returns a binary mask for each image, highlighting all the detected defects. The individual masks are then assembled into a single mask that corresponds to the backend input photo. The input photos and their anomaly masks are saved to a local MongoDB (to be eventually replaced by a production database in the future). The ML processing algorithm is packed into a docker image which is pulled by the web application backend and start it inside a container either on the server directly or, optional, on a computing cluster node, and can be run on GPU.

FastAPI [135] is used to organize the backend into an application, allowing to test its functionalities (even without the frontend). This includes preprocessing quality testing and threshold tuning, as well as model testing on photos. The pretrained model weights are stored on a lxplus eos space, which is accessed via SFTP (SSH File Transfer Protocol). The inference procedure is launched on GPUs in our institute cluster.

The user interface of the system is written in Dart [136], using the Flutter library [137]. The GUI offers four main functionalities:

1. Look at the photos: the user can load the photos of the module taken by the microscope, look at them with a zoom, and automatically remove the PCB background.
2. Get anomaly masks: the model can be run on the uploaded photos, and a mask highlighting the detected defects can be returned.
3. Correct the anomaly masks: the user can delete parts of the mask that are not considered as important defects, or add defects if the algorithm missed some, using the provided drawing/erasing tools.
4. Load the results to the database: the photos and the obtained masks can be saved to the database under a unique name.

The functionality to process the few-shot framework is under development and not currently available in the web application. For now this approach is available outside the web application and the integration is left for future work.

## 4 $t\bar{t}H$ cross-section measurement and $\mathcal{CP}$ -violation studies in the top Yukawa coupling

This chapter presents the measurement of the  $t\bar{t}H$  production cross-section and the study of the  $\mathcal{CP}$ -properties of the top Yukawa coupling using the STXS framework in the multilepton channel. It uses the full  $140 \text{ fb}^{-1}$  Run-2 dataset. The personal contributions within these analyses were focused on the development of the Higgs  $p_T$  reconstruction algorithm and on the study of the  $\mathcal{CP}$ -properties of the top Yukawa coupling within the STXS framework.

The measurements are performed by analysing several final states depending on the number of hadronically decaying  $\tau$  leptons and light leptons:  $e$ ,  $\mu$ . The resulting six channels are (see Figure 4.0.1):

- $2lSS$ : two same-charge light leptons and no hadronically decaying  $\tau$ -lepton candidates,
- $3l$ : three light leptons and no hadronically decaying  $\tau$ -lepton candidates,
- $4l$ : four light leptons,
- $2lSS + 1\tau_{had}$ : two same-charge light leptons and one hadronically decaying  $\tau$ -lepton candidate,
- $1l + 2\tau_{had}$ : one light lepton and two hadronically decaying  $\tau$ -lepton candidates,
- $2lSS + 2\tau_{had}$ : two same-charge light leptons and two hadronically decaying  $\tau$ -lepton candidates.

The orthogonality of the channels, meaning that an event can belong only to one channel, is ensured. This thesis focuses on two channels without  $\tau$  leptons: the channel with two leptons of the same charge named  $2lSS$ , and the channel with three leptons referred to as  $3l$ . An example of a decay chain corresponding to the  $3l$  final state is shown in Figure 4.0.2. The selection of the objects of interest, namely the electrons and muons, jets and missing transverse energy, is detailed in Section 4.1.

The signal regions (SRs) are defined as kinematic phase space containing the process of interest. Section 4.1 presents the object-level selection criteria. The events in the SRs originating from processes different from the target one are called background events. The strategy to estimate the background contribution in the signal regions mainly relies on using simulated event samples discussed in Section 4.2. The dominant background sources ( $t\bar{t}W$ ,  $t\bar{t}Z$  and  $VV$ ) and the known difficult-to-model backgrounds (containing non-prompt and fake leptons) get data-driven corrections. The yields of these backgrounds, represented as normalisation factors, are adjusted simultaneously with the yields of the signal processes by performing a likelihood fit to data, referred

Number of $\tau_{had}$	2	$1\ell + 2\tau$	$2\ell + 2\tau$		4
	1		$2\ell SS + 1\tau$		
	0		$2\ell SS + 0\tau$	$3\ell + 0\tau$	
		1	2	3	
		Number of light $\ell$			

Figure 4.0.1: The  $t\bar{t}H$  multilepton channels in terms of light lepton and  $\tau_{had}$  multiplicities.

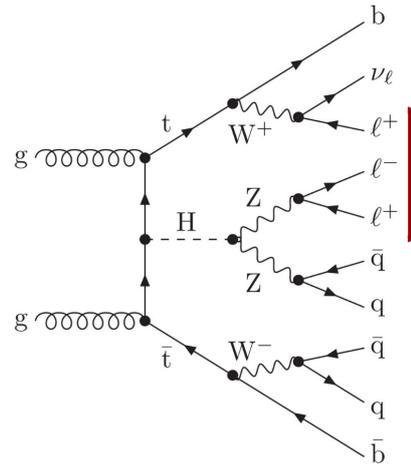


Figure 4.0.2: Examples of decay chain corresponding to the  $3l$  channel.

to as a template fit (see Section 4.5). The fit covers both signal and control regions, where control regions are dedicated regions enriched with specific background processes. Backgrounds having similar kinematic phase space and not easily separable from the signal have control regions defined with the help of a BDT classifier and are called MVA-based. Other control regions are cut-based. The detailed region definitions and the backgrounds are presented in Section 4.3. The analysis strategy and the main control region definitions follow the  $t\bar{t}W$  measurement [138] incorporating the improved understanding of fakes and  $t\bar{t}W$  modelling derived in that analysis. The systematic uncertainties taken into account are described in Section 4.4. In this work, we mainly concentrate on the part of the analysis covering the STXS definition for both the  $t\bar{t}H$  cross-section measurement and the  $\mathcal{CP}$ -studies (see Section 4.6 and Section 4.7). The results of the measurements are reported in Section 4.8.

## 4.1 Object identification and event selection

The events analysed in the  $2lSS$  and  $3l$  channels are selected using several steps. The reconstructed events passed the preselection criteria (see Section 4.1.1) then undergo an overlap removal to ensure no double counting of the same object is happening. The resulting events are split into different analysis channels. This categorisation happens based on the multiplicity light leptons and hadronically decaying  $\tau$  leptons (see Section 4.1.2). Afterwards, as discussed in Section 4.1.3, each channel uses tighter lepton definitions to increase the purity in their dedicated signal regions.

### 4.1.1 Event preselection

The preselection cuts are the first to be applied to data or MC events:

- at least one reconstructed vertex with  $\geq 2$  associated tracks having  $p_T > 500$  MeV,
- no significant noise in the calorimeters and no data corruption,
- response from single- and di-lepton triggers combined by a logical OR.

## 4.1.2 Object definitions

Reconstructed objects from preselected events undergo additional selection to define final state particles, leptons and jets in particular. The overlap removal procedure is applied to ensure the absence of double counting an object as a reconstructed particle.

### 4.1.2.1 Loose leptons

The loose lepton definition, denoted by  $L$ , serves to categorise the events into different analysis channels. The lepton candidates, reconstructed as described in Section 2.4, should pass the following criteria:

- An electron candidate must have  $p_T > 10$  GeV and  $|\eta_{cluster}| < 2.47$ .
- A muon candidate is required to satisfy  $p_T > 10$  GeV and  $|\eta| < 2.5$ .

	$e$					$\mu$			
	L	L'	M	$M_{ex}$	T	L	M	$M_{ex}$	T
LooseVar_Rad isolation	Yes					Yes			
Non-prompt lepton BDT (PLIV)	No		<i>Tight</i>	<i>Tight-not-VeryTight</i>	<i>VeryTight</i>	No	<i>Tight</i>	<i>Tight-not-VeryTight</i>	<i>VeryTight</i>
Identification	Loose	Tight				Loose	Medium		
Charge mis-assignment veto (ECIDS)	No	Yes				N/A			
Material/internal conversion veto	No	Yes				N/A			
Transverse impact parameter significance $ d_0 /\sigma_{d_0}$	$< 5$					$< 3$			
Longitudinal impact parameter $ z_0 \sin \theta $	$< 0.5$ mm								

Table 4.1.1: Lepton definitions in the  $2lSS$  and  $3l$  channels: Loose ( $L$  or  $L'$ ), Medium ( $M$ ), Medium exclusive ( $M_{ex}$ ), and Tight ( $T$ ).

The definition criteria are summarized in Table 4.1.1 for electrons and muons. The criteria used in the table are defined below:

- LooseVar\_Rad: lepton isolation working point defined in Ref. [139] separately for electrons and muons based on tracks and clusters isolation.
- PLIV (Prompt Lepton Isolation Veto) [140]: tagger based on a BDT algorithm designed for a better non-prompt leptons rejection. The algorithm takes as input particle tracks and energy deposits in a cone around the lepton direction and the kinematics of the lepton candidate.
- Identification: working points described in Sections 2.4.2 and 2.4.3 for electrons and muons.
- ECIDS (Electron Charge Identification and Separation) [92]: a BDT-based tagger for electron charge determination. Electrons (or positrons) might be reconstructed with an incorrect

charge, for example when an original  $e^+$  emits a hard bremsstrahlung photon that further converts into a  $e^+e^-$  pair where the  $e^-$  has a momentum high enough to be reconstructed instead of the initial  $e^+$ . For muons, charge misidentification is negligible due to its rareness.

- Material/internal conversion veto [141]: electron candidates might originate from internal or material conversions (see Section 4.1.3). Material conversions refer to photon conversions that occur when a photon interacts with the material of the detector and converts into an electron-positron pair. This typically happens in the inner detector layers or calorimeters. Material conversions are identified by tracing back the  $e^-e^+$  pairs to their conversion points in the detector material. The material conversion candidates are selected based on requirements on displaced vertex and invariant mass (between the associated track and the closest track with the opposite charge). Internal conversions are processes where an emitted virtual photon leads to an  $e^-e^+$  pair where only one of the electrons is reconstructed. Criteria used to select electrons coming from internal conversion are based on the di-track invariant calculated at the primary vertex. Both Material and internal conversion candidates are vetoed.
- $|d_0|/\sigma_{d_0}$  and  $|z_0\sin\theta|$  : requirements on the transverse ( $d_0$ ) and longitudinal ( $z_0$ ) impact parameters (see Section 2.4.1) used to select particles originating from the primary vertex.

#### 4.1.2.2 Jets

Jets, reconstructed via the procedure described in Section 2.4.4, are required to satisfy  $p_T > 25$  GeV and  $|\eta| < 2.5$  with a "Tight" working point of Jet Vertex Tagger. Forward jets follow the same requirements apart from the pseudorapidity one:  $|\eta| < 4.5$ .

Jets containing  $b$ -hadrons must have  $p_T > 20$  GeV and  $|\eta| < 2.5$ . A jet is identified as a  $b$ -jet based on the  $DL1r$  deep learning discriminant [142] using a working point of 85% efficiency.

#### 4.1.2.3 Removing ambiguities between overlapping reconstructed objects

To avoid a single detector response being counted as multiple physics objects in an event, the overlap removal procedure is applied to the preselected loose leptons and jets.

Among electrons with shared track or overlapping cluster, only the one with the highest  $p_T$  is kept. To remove light leptons and  $\tau_{had}$  overlap, any  $\tau$  candidate within  $\Delta R < 0.2$  of a light lepton is neglected. Loose electrons/muons overlap is resolved by removing the electron lying within  $\Delta R = 0.1$  of the muon. In case of electron/jet overlap within  $\Delta R < 0.2$ , an object is considered as an electron if the jet is not  $b$ -tagged or if the jet has  $p_T > 200$  GeV. A jet overlapping with a loose muon is removed if this jet is not  $b$ -tagged and has less than 3 tracks with  $p_T > 500$  MeV. In case of doubt between  $\tau_{had}$  and jet found within  $\Delta R < 0.2$ , the object is interpreted as a  $\tau$  lepton. To solve the overlap between loose leptons and jets within a cone of a radius  $0.004 + 10\text{GeV}/p_{T,lep}$  up to 0.4, not covered in the previous steps, the lepton is removed.

The association of the event to one of the analysis channels happens based on the number of the loose leptons and  $\tau$  leptons.

### 4.1.3 Region selection

Tighter in-channel lepton definitions are necessary to reduce (in SRs) or select (in CRs) fake and non-prompt leptons. Fake leptons are objects misidentified as light leptons: it might be, for example, a misidentified jet or photon converted into electron-positron pair. The two types of photon conversions are also considered as sources of fake leptons: material and internal conversions. Non-prompt leptons are real leptons that do not originate directly from the primary interaction vertex but from the secondary decays, e.g. from leptonically decaying heavy-flavour (HF) hadrons or taus.

The tight  $T$  working point (see Table 4.1.1) serves to reject material and internal conversions. Leptons satisfying the tight  $T$  selection are used to define signal regions in both the  $2lSS$  and  $3l$  channels (see Table 4.1.2). Exclusive medium  $M_{ex}$  lepton selection uses criteria identical to the tight  $T$  definition, with the exception of the PLIV working point. The  $M_{ex}$  definition is specially formulated for the fakes background estimation strategy: tight and exclusive medium leptons are used to define the regions dedicated to non-prompt HF leptons in the  $2lSS$  channel. More details about region definitions are provided in Section 4.3.

The medium  $M$  working point is needed to reduce the number of misidentified objects, allowing to select prompt-like leptons after PLIV cut. A looser tight definition  $L'$  doesn't use any PLIV selection, but uses stricter identification and charge mis-assignment requirements as well as conversions vetos.  $M$ ,  $L/L'$  selections are applied in the  $3l$  channel to have well populated non-prompt lepton control regions.

Channel	Cut-based Control Regions	Signal and MVA-based Control regions
$2lSS$	$TM_{ex}, M_{ex}T, M_{ex}M_{ex}$	$TT$
$3l$	$L'MM/LMM$ ( $L'$ for $\mu$ and $L$ for $e$ )	$LTT$ ( $L$ for $\ell_0$ )

Table 4.1.2: Summary of the lepton definitions used for each of the leptons selected in a given channel. For instance,  $TT$  corresponds to the definition of the two leptons in the SR of the  $2lSS$  channel. The leptons of the same sign are ordered by decreasing  $p_T$ , and the first lepton in  $3l$  corresponds to the lepton having the charge opposite to two other leptons.

## 4.2 Signal and background modelling

To model the signal and background processes, the analysis uses MC simulated samples produced with the ATLAS simulation software (see Section 2.3).

The majority of the MC samples are generated via full ATLAS detector simulation based on GEANT4, although some samples of  $tWH$ ,  $tHqb$  and  $t\bar{t}\bar{t}$  processes are available as ATLAS fast simulation only. Pile-up is modelled by using low-momentum strong-interaction processes simulated with PYTHIA8, which are then superimposed onto the simulated hard-scatter events based on the luminosity profile of the collected data. The simulated events are reconstructed following the same procedure used for the data. Corrections are applied to ensure that the reconstruction and identification efficiencies, energy scales, and energy resolutions of particle candidates in

the simulations align with those determined from data control samples. Additionally, the simulated events are reweighted based on the number of generated events and the applied corrections. The MC samples are normalised to the corresponding processes cross-sections. The uncertainties arising from the data-driven corrections and the modelling limitations are discussed in Section 4.4.

The overview of the generators and their configurations used to simulate the considered processes is presented in Table 4.2.1. Samples used to estimate the systematic uncertainties are indicated in grey.

Process	Generator	ME order	Parton shower	PDF	Tune
$t\bar{t}H$	POWHEG-BOX	NLO	PYTHIA8	NNPDF3.0 NLO	A14
	(POWHEG-BOX)	(NLO)	(HERWIG7.0.4)	(NNPDF3.0 NLO)	(H7-UE-MMHT)
	(MG5_aMC)	(NLO)	(PYTHIA8)	(NNPDF3.0 NLO)	(A14)
$t\bar{t}W$	SHERPA 2.2.10	MEPs@NLO	SHERPA	NNPDF3.0 NNLO	SHERPA default
	(MG5_aMC)	(FxFx NLO)	(PYTHIA8)	(NNPDF3.0 NLO)	(A14)
	(POWHEG)	(NLO)	(PYTHIA8)	(NNPDF3.0 NLO)	(A14)
	(POWHEG)	(NLO)	(HERWIG7)	(NNPDF3.0 NLO)	(H7-UE-MMHT)
$t\bar{t}W$ (EW)	SHERPA 2.2.10	LO	SHERPA	NNPDF3.0 NNLO	SHERPA default
	(MG5_aMC)	(LO)	(PYTHIA8)	(NNPDF3.0 NLO)	(A14)
$t\bar{t}l$	MG5_aMC	NLO	PYTHIA8	NNPDF3.0 NLO	A14
	(MG5_aMC)	(NLO)	(HERWIG7)	(NNPDF3.0 NLO)	(H7-UE-MMHT)
	(MG5_aMC)	(NLO)	(PYTHIA8)	(NNPDF3.0 NLO)	(A14 VAR3C)
$t\bar{t} \rightarrow W^+bW^-\bar{b}l^+l^-$	MG5_aMC	LO	PYTHIA8	NNPDF2.3 LO	A14
$t\bar{t}t\bar{t}$	MG5_aMC	NLO	PYTHIA8	NNPDF3.1 NLO	A14
$t\bar{t}$	POWHEG-BOX	NLO	PYTHIA8	NNPDF3.0 NLO	A14
	(POWHEG-BOX)	(NLO)	(HERWIG7.1.3)	(NNPDF3.0 NLO)	(H7-UE-MMHT)
$t\bar{t}t$	MG5_aMC	LO	PYTHIA8	NNPDF2.3 LO	A14
Single top ( $t$ -, $Wt$ -, $s$ -channel)	POWHEG-BOX	NLO	PYTHIA8	NNPDF3.0 NLO	A14
$VV, qqVV, VVV$	SHERPA 2.2.2(1)	MEPs@NLO	SHERPA	NNPDF3.0 NNLO	SHERPA default
$Z \rightarrow l^+l^-$	SHERPA 2.2.1	MEPs@NLO	SHERPA	NNPDF3.0 NNLO	SHERPA default
$Z \rightarrow l^+l^-(\gamma \rightarrow e^+e^-)$	POWHEG-BOX	NLO	PYTHIA8	CTEQ6L1 NLO	A14
$Z \rightarrow l^+l^-(\gamma^* \rightarrow e^+e^-)$	POWHEG-BOX	NLO	PYTHIA8	CTEQ6L1 NLO	A14

Table 4.2.1: Summary of the configurations used for event generation of the signal and background processes. Samples in grey are the samples used to estimate the systematic uncertainties.

### 4.2.1 $t\bar{t}H$ signal modelling

The signal  $t\bar{t}H$  events are modelled via the POWHEGBOXv2 generator ([143]). The matrix element is generated at next-leading-order (NLO) in QCD with NNPDF3.0 NLO [144] parton distribution function set. The parameter  $h_{damp}$ , that matches the POWHEG ME to the parton shower and regulates the high- $p_T$  radiation against which the  $t\bar{t}$  system recoils, is set to  $0.75 \times (m_t + m_{\bar{t}} + m_H) = 352.5$  GeV. The PS and hadronisation processes are generated with PYTHIA8. The heavy flavour

hadrons decays are simulated by EVTGEN1.6.0 [145]. The MC sample is normalised using a cross-section  $507_{-50}^{+35}$  fb calculated at NLO in QCD with the leading NLO electroweak corrections (i.e.  $\mathcal{O}(\alpha_s^2\alpha^2)$  where  $\alpha_s$  is the strong coupling). The cross-section uncertainties include PDF+ $\alpha_s$  uncertainties and variations of the renormalisation and factorisation scales (see Section 4.4.2).

### 4.2.2 $t\bar{t}H$ and $tH$ modelling for $\mathcal{CP}$ -studies

To perform the  $\mathcal{CP}$ -measurement, the  $t\bar{t}H$  samples corresponding to  $\mathcal{CP}$ -odd and  $\mathcal{CP}$ -even hypotheses are needed, as well as a sample with  $\mathcal{CP}$ -mixing angle  $\alpha = 45^\circ$  for validation purposes. Matrix element calculation in these samples is done via the MADGRAPH5\_AMC generator (followed by PYTHIA8 to model the showering and hadronisation) instead of the POWHEG one used to produce signal samples for  $t\bar{t}H$  cross-section measurements. The MADGRAPH5\_AMC generator allows generating samples with non-zero  $\alpha$  assuming a Yukawa coupling Lagrangian extension (see Eq. 1.7.2) that includes a  $\mathcal{CP}$ -odd part to the top–Higgs interaction. These non-SM samples are generated with the NLO Higgs Characterisation (HC) [146] UFO model [147] implemented in Feynrules [148].

The Higgs production in association with single-top is considered in two sub-processes:  $tHjb$  and  $tWH$ . The hard-scattering part is generated using MADGRAPH5\_AMC@NLO at NLO in QCD with NNPDF3.0 NLO PDF set. The showering is modelled with PYTHIA8. In analogy with the  $t\bar{t}H$   $\mathcal{CP}$  samples, the non-SM  $tH$  samples are generated using the NLO HC model assuming a  $\mathcal{CP}$ -odd part in the top–Higgs interaction.  $tH$  samples with different values of  $\kappa'_t$  and  $\alpha$  are needed to derive the  $tH$  parameterization described in Section 4.7.3. In total 11 samples corresponding to points in  $(\kappa'_t, \alpha)$  phase space were generated for both sub-processes:

- $\alpha = 0$  with  $\kappa'_t \in -1, 0.5, 2$ ,
- $\alpha$  from  $15^\circ$  to  $90^\circ$  in  $15^\circ$  steps with  $\kappa'_t = 1$ ,
- $\alpha = 45^\circ$  with  $\kappa'_t = 2$ .

The list of generated samples with corresponding their cross-sections is presented in Table 4.2.2.

### 4.2.3 $t\bar{t}W$ background modelling

The nominal  $t\bar{t}W$  sample is simulated with SHERPA using the NNPDF3.0 NLO PDF set. The ME is calculated up to one additional parton at NLO and up to two partons at LO and merged with the SHERPA PS. The functional form of renormalisation and factorisation scales is  $\mu_R = \mu_F = H_T/2$ , with  $H_T$  being the scalar sum of transverse masses  $\sqrt{p_T^2 + m^2}$  of all final state particles. Additionally, higher-order corrections corresponding to EWK contributions were included through correction factors and an independent LO QCD SHERPA sample. The combination of the described contributions results in a total cross-section of  $\sigma(t\bar{t}W) = 722_{-78}^{+70}$  (scale)  $\pm 7$  (PDF) fb after normalisation to the value from Ref. [149].

<b>k</b>	$\alpha$	$\sigma \times \text{k-factor (fb}^{-1}\text{)}$			
		$t\bar{t}H$	$tHjb$	$tWH$	
1	0°	507.44	60.30	16.72	
1	15°		63.57	18.32	
1	30°		75.12	23.11	
1	45°		361.57	99.01	31.04
1	60°		138.16	42.12	
1	75°		195.96	56.22	
1	90°		218.8	271.47	72.90
1	180°		706.75	149.91	
0.5	0°		95.27	14.80	
2	0°		240.07	9.15	
2	45°		207.56	109.42	

Table 4.2.2: Summary of the  $t\bar{t}H$ ,  $tH$  and  $tWH$  MC samples at different  $\alpha$  angles and different  $\kappa_t'$ .

#### 4.2.4 Other backgrounds modelling

The  $t\bar{t}Z/\gamma^*(\rightarrow l^+l^-)$  background process was generated using MADGRAPH5\_AMC@NLO + PYTHIA8, using a ME simulated at NLO in  $\alpha_s$  with the NNPDF3.0 NLO PDF set. The top quarks decays are modelled with MADSPIN [150] at LO. The inclusive  $t\bar{t}l^+l^-$  ME is computed at NLO, and takes into account off-shell  $Z$  and  $\gamma^*$  contributions with  $m(l^+l^-) > 1$  GeV. The cross-section is normalised to the calculation at NLO QCD and NLO EWK accuracy reported in Ref. [151].

The di-boson event simulation is performed with SHERPA2.2.1 for semi-leptonic decays and SHERPA2.2.2 for fully-leptonic decays with up to one additional parton at NLO and three additional partons at LO. A mismodelling of the number of jets distribution of the diboson process was reported in SHERPA2.2 generated samples [138], leading to a necessity of deriving a corresponding correction using data. The single boson production in association with multiple jets is simulated with the SHERPA2.2.1 generator using NLO-accurate matrix elements for up to two jets, and LO-accurate ones for up to four jets with the NNPDF3.0 NLO set of PDFs.  $V$ +jets samples are normalized to next-to-next-to-leading order (NNLO) predictions.

The  $t\bar{t}$  and single top quark samples are generated with POWHEGBOXV2 + PYTHIA8 at NLO with the NNPDF3.0 NLO PDF set and  $h_{damp} = 1.5m_t$ , while the parton shower, hadronisation, and underlying event are modelled using NNPDF2.3 L0 PDF set. Heavy flavour decays are performed by EVTGEN 1.6. For single-top  $s$ - and  $t$ -channel productions, the POWHEGBOXV2 at NLO in QCD with the NNPDF3.0 NLO PDF set is used. The  $t\bar{t}$  sample was normalised to the cross-section prediction at next-to-next-to-leading order (NNLO) in QCD [152], and single top quark inclusive cross-section was adjusted to the theory prediction calculated at NLO in QCD [153].

Higgs production associated with a  $W$  or  $Z$  boson is simulated using PYTHIA8 with the NNPDF2.3 L0 PDF set. Triple top-quark production ( $t\bar{t}t$ ) and  $t\bar{t}$  production with two  $W$  bosons ( $t\bar{t}WW$ ) are generated at leading order using MADGRAPH5\_AMC@NLO+PYTHIA8. For four top-quark production ( $t\bar{t}t\bar{t}$ ), simulations are performed at NLO with the MADGRAPH5\_AMC@NLO

generator, interfaced with PYTHIA8 with the NNPDF3.0 NLO PDF set. Processes involving three heavy gauge bosons ( $WWW$ ,  $WWZ$ ,  $WZZ$ ,  $ZZZ$ ) with up to six leptons in the final states are simulated using SHERPA2.2.2 and the NNPDF3.0 NLO PDF set. Final states without additional partons are calculated at NLO, while those with one, two, or three additional partons are calculated at LO.

## 4.3 Regions definition

Figure 4.3.1 shows an overview of the analysis regions in the  $2lSS$  and  $3l$  channels. There are three signal regions: the  $t\bar{t}H$  SRs with two same-sign leptons, positive and negative charged, and  $t\bar{t}H$  SR in three lepton final state. In the  $\mathcal{CP}$ -studies, the  $tH$  process is also considered as a signal, this is detailed in Section 4.7. Background processes might be either reducible or irreducible. The categorisation is done based on the presence of non-prompt or misidentified leptons.

The reducible backgrounds contain at least one prompt charge-flip electron or one non-prompt light lepton originating from material conversions, hadron decays, or the improper reconstruction of other particles as leptons. More precisely, the main source of non-prompt leptons is  $t\bar{t}$  production, followed by  $V$ +jets and single-top processes. In the MC samples, the non-prompt leptons are labelled according to their origin: heavy-flavour (HF) hadron decays of bottom and charm quarks, light-flavour (LH) hadron decays or material conversion. The LH contribution in the  $2lSS$  and  $3l$  channels is negligible. The cut-based control regions, described in Section 4.3.1, are mainly dedicated to the reducible backgrounds.

Irreducible background processes have only prompt leptons in their final state, making it challenging to distinguish them from the signal. Such backgrounds are produced in  $W/Z$  or  $\tau$  decays, as well as in internal conversions. The main sources of irreducible backgrounds are  $t\bar{t}W$ ,  $t\bar{t}Z/\gamma^*$  and  $VV$ , with smaller contributions from  $tZ$ ,  $tW$ ,  $tWZ$ ,  $t\bar{t}WW$ ,  $VVV$ ,  $t\bar{t}t$  and  $t\bar{t}t$ . The control regions corresponding to some irreducible backgrounds are defined via MVA-optimised selection (see Section 4.3.2). As shown in Figure 4.3, the  $t\bar{t}Z$  and  $VV$  control regions are present both as cut- and MVA-based regions. The  $t\bar{t}Z$  and  $VV$  cut-based regions are inherited from the  $t\bar{t}W$  analysis and are needed to improve the prediction of the background contribution from the  $VV$  and  $t\bar{t}Z/\gamma^*$  processes. The strategy for estimating the backgrounds consists of using MC simulation to model their shapes (see Section 4.2). The following background yields are adjusted via normalisation factors in the template fit (see Section 4.5):  $t\bar{t}W$ ,  $t\bar{t}Z$ ,  $VV$  and non-prompt leptons from HF decays.

### 4.3.1 Cut-based control regions

Two control regions are built to estimate the contribution from conversions with  $m(e^+e^-) < 1$  GeV in the  $2lSS$  and  $3l$  channels: one corresponding to internal conversion and one corresponding to material conversion. The events in these regions are required not to have  $b$ -jets, and not to have an opposite sign di-lepton mass close to  $Z$  mass. For both contributions, the normalisation factor is fitted together with the signal. Figure 4.3.2 shows the composition of both regions.

There are six CRs enriched in contributions from HF non-prompt leptons. These regions differ

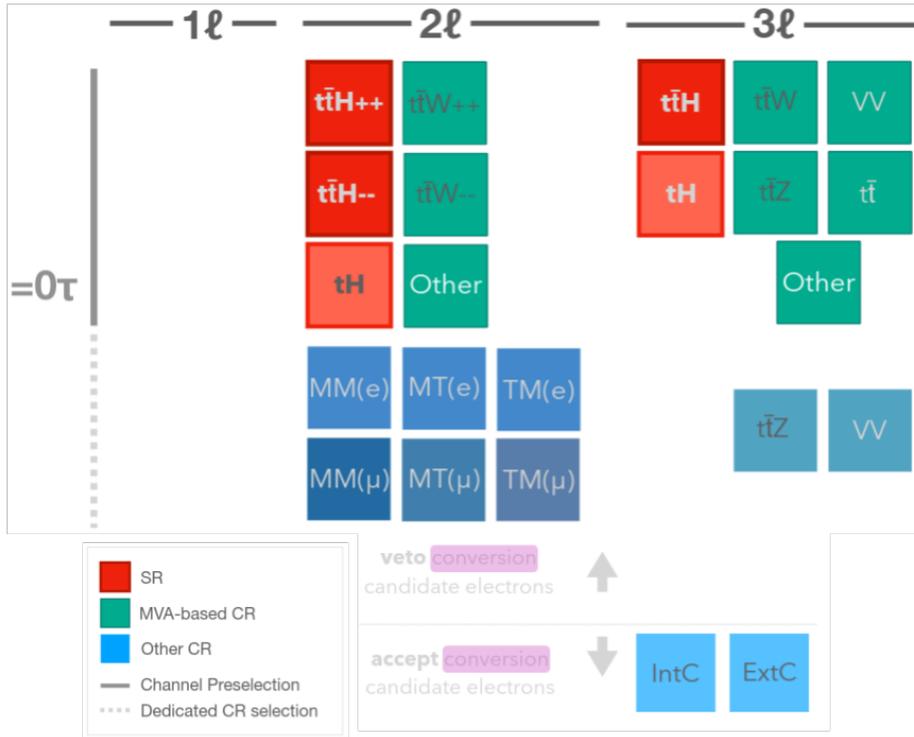


Figure 4.3.1:  $t\bar{t}H$  multilepton regions in  $2lSS$  and  $3l$  channels split into signal regions, MVA-optimised control regions, and cut-based (or other) control regions.

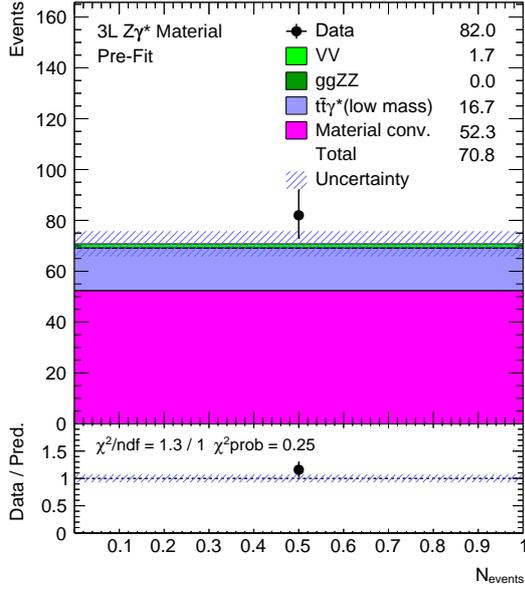
by the lepton type,  $e$  or  $\mu$ , and working points used to define the leptons. The requirement to have at least one  $M_{ex}$  lepton allows to separate the non-prompt lepton processes from processes with prompt leptons. Two normalization factors are estimated from the simultaneous likelihood fit to data: HF non-prompt lepton contribution from electrons and muons. The HF regions composition is illustrated in Figure 4.3.3.

The MC sample of the  $VV$  events used to build the corresponding control region in the  $3l$  channel is adjusted before the fit using data-driven correction since the simulated  $VV$  sample does not model well the jet multiplicity spectrum in data. This correction is based on  $N_{jets}$  and adopted from the  $t\bar{t}W$  analysis [138]. The region is composed of events with 2 or 3 jets. The  $t\bar{t}Z$  region has events with  $\geq 4$  jets. Both regions require to include events with one  $b$ -jet, have the opposite sign leptons mass should not be far away from the  $Z$  mass, but the overall lepton mass should significantly differ from  $Z$  mass. The  $VV$  and  $t\bar{t}Z$  regions are shown in Figure 4.3.4.

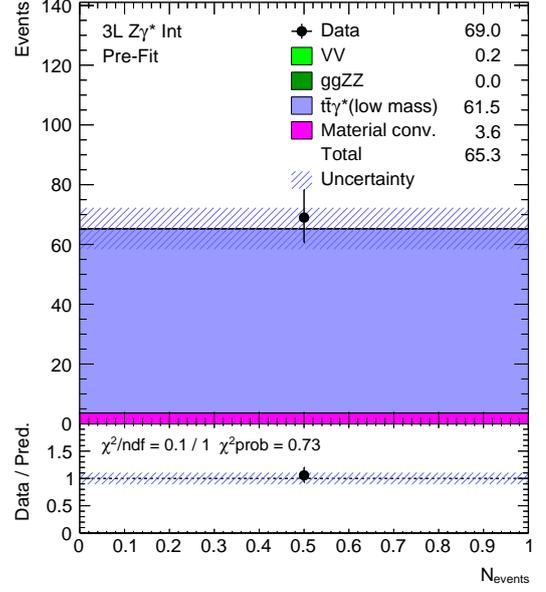
The definitions of the cut-based control regions are summarised in Table 4.3.1.

### 4.3.2 Signal regions and MVA-based control regions

The procedure to optimize the signal region definition is similar between the  $2lSS$  and  $3l$  channels. First, the requirements are applied to select simulated events (see Section 4.3.2.1) that are further used to train a BDT classifier that categorises events into different classes. The BDT training uses the MVA-trainer framework [154].



(a) Material conversion CR.



(b) Internal conversion CR.

Figure 4.3.2: Distribution of data and MC prediction in the control regions corresponding to material and internal conversions.

Control regions for:	Diboson	$t\bar{t}Z$	Conversions	HF non-prompt
$N_{jets}$	2 or 3	$\geq 4$	$\geq 0$	$\geq 2$
$N_{b-jets}$	1 $b^{85\%}$		0 $b^{85\%}$	1 $b^{85\%}$
Lepton requirement	3l		$\mu\mu e^*$	2lSS
Lepton definition		(L, M, M)		(T, $M_{ex}$ )    ( $M_{ex}$ , T)    ( $M_{ex}$ , $M_{ex}$ )
Lepton $p_T$ [GeV]		(10, 15, 15)		(15, 15)
$ m_{l+l}^{SF} - m_Z $ [GeV]	$< 10$		$> 10$	–
$ m_{ll} - m_Z $ [GeV]	$> 10$		$< 10$	–
$m_T(\ell_0, E_T^{miss})$ [GeV]		–		$< 250$ for $TM_{ex}$ and $M_{ex}T$ pairs
$\tau_{had}$ candidates (Medium)		0		0
Region split	–	–	internal / material	subleading $e/\mu \times (TM_{ex}, M_{ex}T, M_{ex}M_{ex})$
Region naming	3lVV	3l $t\bar{t}Z$	3lIntC 3lMatC	2lSS $t\bar{t}(e)_{TM_{ex}}$ , 2lSS $t\bar{t}(e)_{M_{ex}T}$ , 2lSS $t\bar{t}(e)_{M_{ex}M_{ex}}$ 2lSS $t\bar{t}(\mu)_{TM_{ex}}$ , 2lSS $t\bar{t}(\mu)_{M_{ex}T}$ , 2lSS $t\bar{t}(\mu)_{M_{ex}M_{ex}}$

Table 4.3.1: Event selection summary for the background control regions.

#### 4.3.2.1 Pre-MVA selection

The selection criteria applied after preselection are summarized in Table 4.3.2 and are referred to as pre-MVA selection. In both channels, events with  $\tau_{had}$  candidates are excluded from consideration.

The events considered to be in the SRs or in the MVA-based CRs in the 2lSS channel must have exactly two reconstructed same-sign light leptons satisfying the  $T$  working point criteria and having  $p_T > 15$  GeV. To suppress the contribution from the main irreducible backgrounds, only the events with at least three reconstructed jets, with at least one of them being  $b$ -tagged, are selected.

For the 3l channel, the events with three reconstructed leptons with a total charge equal to  $\pm 1$

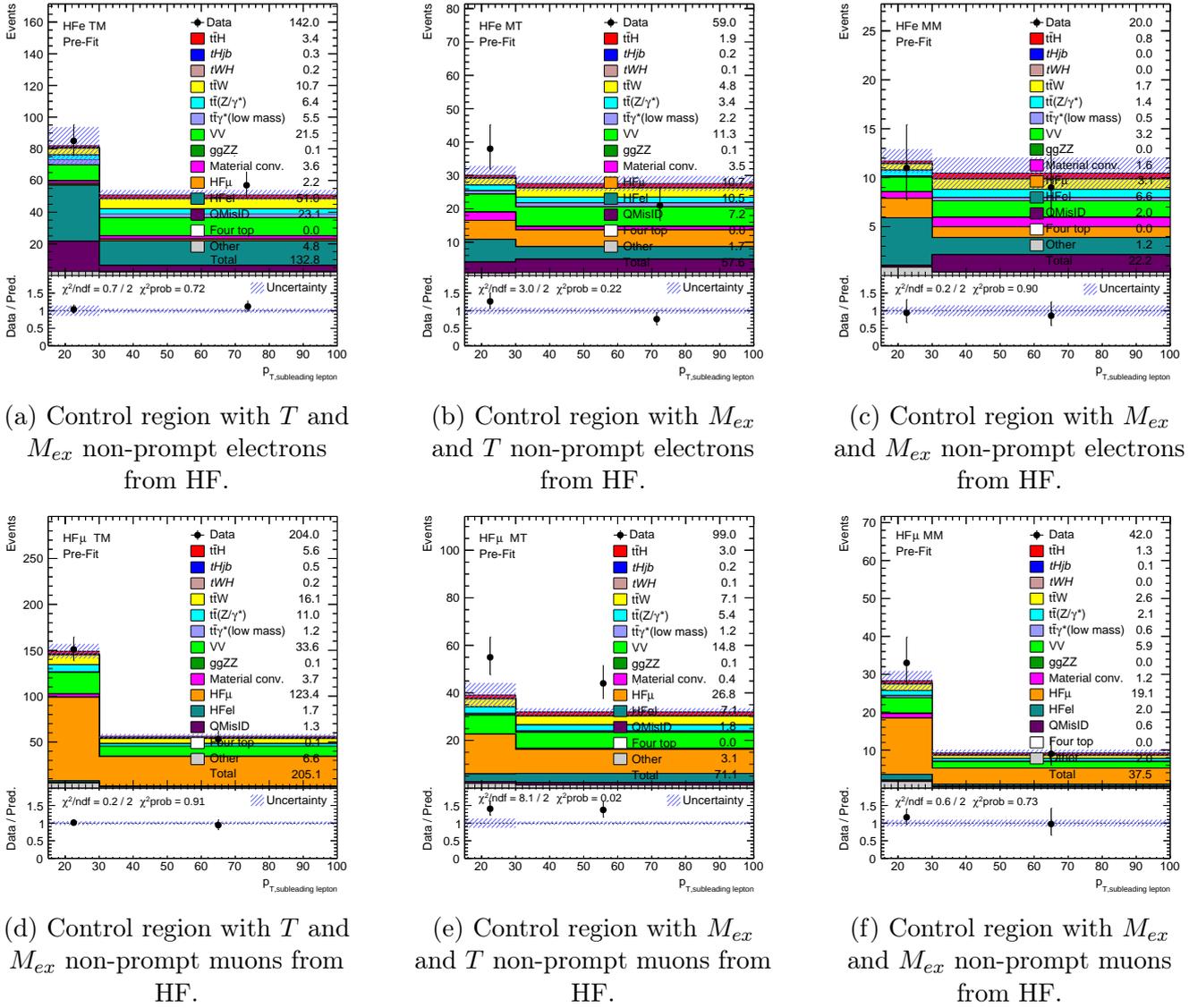
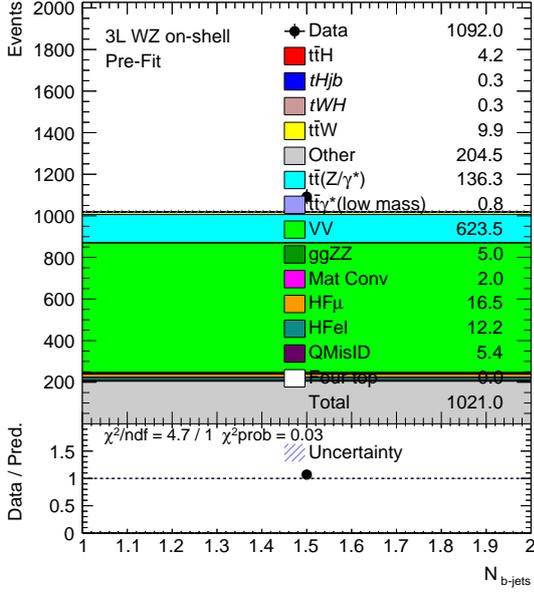
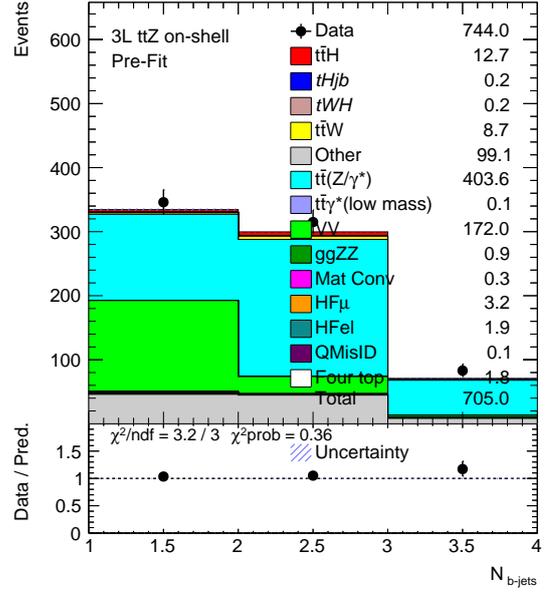


Figure 4.3.3: Distribution of data and MC prediction in the control regions corresponding to differently defined non-prompt leptons from HF decays.

are selected, where the lepton of the opposite charge satisfies the  $L$  lepton definition and has  $p_T > 10$  GeV, and the same-sign leptons are  $T$  defined with  $p_T > 15$  GeV. Additional requirements are applied to suppress the contributions from  $t\bar{t}$ ,  $t\bar{t}W$ ,  $VV$ ,  $t\bar{t}Z$  and HF events. The background from hadronic resonances decaying to light same-flavoured leptons is rejected by requiring  $m(l^+l^-) > 12$  GeV. The  $Z$  veto,  $|m(l^+l^-) - m(Z)| > 10$  GeV, serves to decrease the contribution from  $t\bar{t}Z$  events. Other background processes are suppressed by requiring an event to have at least two reconstructed jets, with at least one of these being  $b$ -tagged. The samples used to model the  $t\bar{t}Z$  and  $t\bar{t}$  processes in the  $3l$  channel are different from those described in Section 4.2. The nominal  $t\bar{t}Z$  sample led to some instabilities in the BDT training due to a large number of negative weights. Thus, this sample, generated with MADGRAPH, was replaced with one generated with SHERPA. The non-all-hadronic  $t\bar{t}$  sample lacked sufficient statistics and was replaced with the dileptonic sample. To further increase the number of  $t\bar{t}$  events in the pre-MVA  $3l$  region, the selection criteria on leptons were loosened, and the sample underwent reweighting to restore the shape of the original  $t\bar{t}$  distribution.

(a) Di-boson  $VV$  CR.(b)  $t\bar{t}Z$  CR.Figure 4.3.4: Distribution of data and MC prediction in the control regions enriched with  $VV$  and  $t\bar{t}Z$  events.

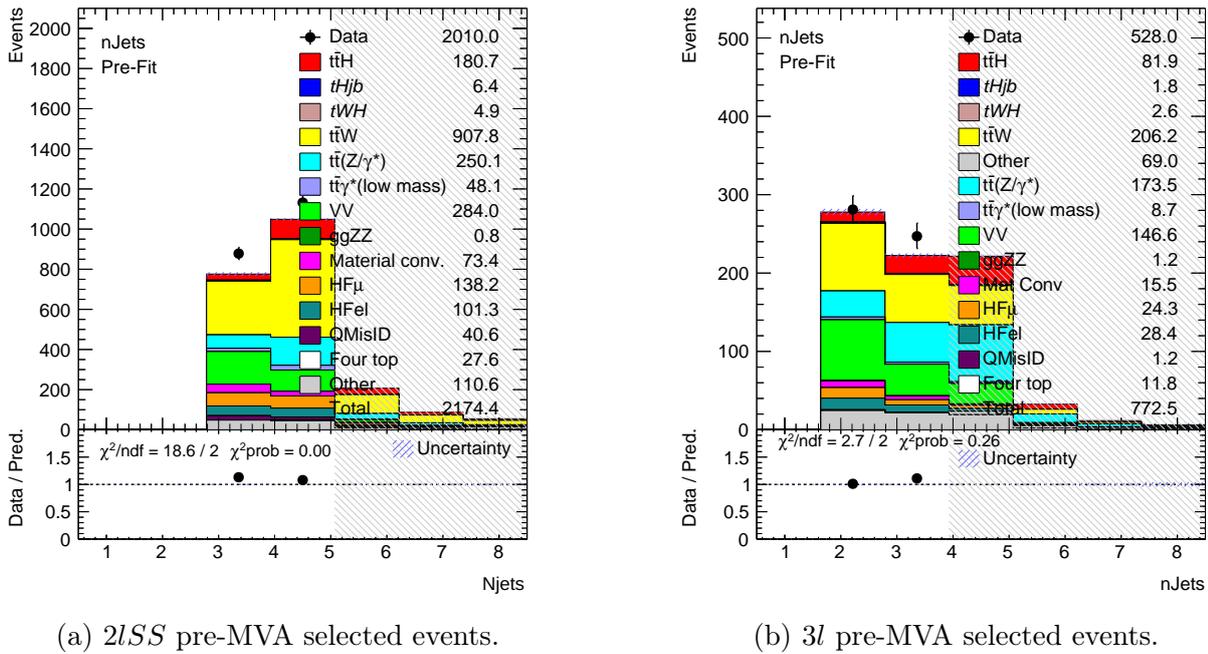
	$2lSS$	$3l$
$\tau_{had}$ candidates	$==0 M$	$==0 M$
Leptons counting	$==2 T: p_T > 15 \text{ GeV}$	$==3 (T,T,L): p_T > 15, 15, 10 \text{ GeV}$
Lepton details	SS	OS (to others): $L, p_T > 10 \text{ GeV}$ SS pair: $T, p_T > 15 \text{ GeV}$ OS pair: $ m(ll) - m_Z  > 10 \text{ GeV}$ and $m(ll) > 12 \text{ GeV}$
$N_{jets}$	$\geq 3$	$\geq 2$
$N_{b-jets}$ (@ 85% WP)	$\geq 1$	$\geq 1$

Table 4.3.2: Pre-MVA selection criteria for the  $2lSS$  and  $3l$  channels.

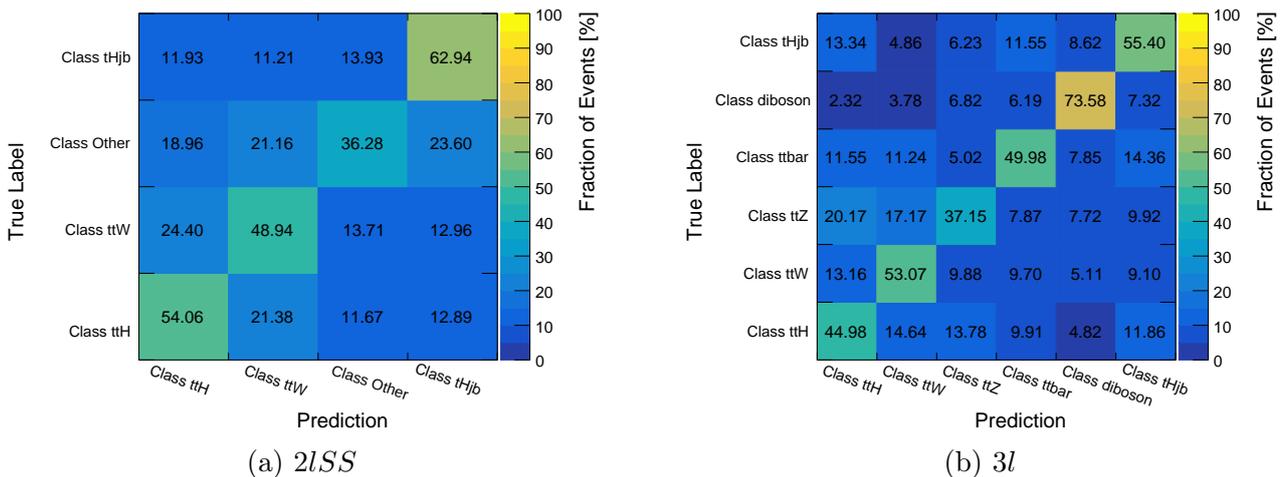
The samples composition after pre-MVA selection is shown in Figure 4.3.5. The  $2lSS$  pre-MVA region is largely dominated by the  $t\bar{t}W$  background, followed by  $t\bar{t}Z/\gamma^*$ ,  $VV$  and the  $t\bar{t}H$  signal. There is a smaller but comparable contribution from HF non-prompt lepton background in the  $2lSS$  channel, while other processes contribute to both channels much less. The pre-MVA set of events serves as training set for a BDT classifier.

#### 4.3.2.2 $2lSS$ channel MVA regions

In the  $2lSS$  channel, the events are categorised by a BDT into four classes:  $t\bar{t}H$ ,  $tH$ ,  $t\bar{t}W$  and *Other*, where *Other* stands for the background processes such as  $t\bar{t}Z$ ,  $t\bar{t}$ ,  $VV$ , single top,  $t\bar{t}t\bar{t}$  and fake/non-prompt leptons. The training is done based on 20 kinematic variables, such as the number of jets, angular distances between leptons and jets, jets transverse momenta and pseudo-


 Figure 4.3.5: Distribution of MC prediction in the pre-MVA sets in the  $2lSS$  and  $3l$  channels.

rapidities, invariant masses of leptons, MET and jets combinations. The trained model yields a probability of a given event to belong to the four classes. The final prediction for a given event is defined by the class with the highest probability score. The confusion matrix representing the predictions for each class given the true labels is shown in Figure 4.3.6 (a). While the  $tH$ ,  $t\bar{t}H$  and  $t\bar{t}W$  classes are fairly well predicted (having the correct prediction rate at 63%, 54% and 49%), there is still considerable confusion of around 20% between events from *Other* category and the rest of the categories, as well as between  $t\bar{t}H$  and  $t\bar{t}W$  events.


 Figure 4.3.6: Confusion matrices for the  $2lSS$  and  $3l$  channels BDTs.

The output of the BDT is used to define the signal and the MVA-based control regions. The region definitions are summarized in Table 4.3.3. The  $t\bar{t}H$  and  $t\bar{t}W$  regions are further divided into two regions corresponding to the sign of the total leptons charge to improve the precision of the signal strength. Figure 4.3.7 shows the composition of the resulting regions. The data in a region is not shown due to blinding: the signal of interest is expected to be seen only after all

analysis methods are fully developed and validated to avoid introducing a bias. If a bin has a signal over background ratio larger than 0.15, it gets blinded.

Regions	Selections
Signal Regions	
$t\bar{t}H$ ++	$2lSS$ pre-MVA selection total charge $> 0$ $t\bar{t}t\bar{t}$ veto (exclude events with $N_{jets} \geq 6$ & $N_{b-jets}(@ 77\% \text{ WP}) \geq 3$ ) $\text{Max}(\text{BDT}_{t\bar{t}H}, \text{BDT}_{tH}, \text{BDT}_{t\bar{t}W}, \text{BDT}_{\text{Other}}) = \text{BDT}_{t\bar{t}H}$
$t\bar{t}H$ --	$2lSS$ pre-MVA selection total charge $< 0$ $t\bar{t}t\bar{t}$ veto $\text{Max}(\text{BDT}_{t\bar{t}H}, \text{BDT}_{tH}, \text{BDT}_{t\bar{t}W}, \text{BDT}_{\text{Other}}) = \text{BDT}_{t\bar{t}H}$
$tH$	$2lSS$ pre-MVA selection $\text{Max}(\text{BDT}_{t\bar{t}H}, \text{BDT}_{tH}, \text{BDT}_{t\bar{t}W}, \text{BDT}_{\text{Other}}) = \text{BDT}_{tH}$
Control Regions	
$t\bar{t}W$ ++	$2lSS$ pre-MVA selection total charge $> 0$ $\text{Max}(\text{BDT}_{t\bar{t}H}, \text{BDT}_{tH}, \text{BDT}_{t\bar{t}W}, \text{BDT}_{\text{Other}}) = \text{BDT}_{t\bar{t}W}$
$t\bar{t}W$ --	$2lSS$ pre-MVA selection total charge $< 0$ $\text{Max}(\text{BDT}_{t\bar{t}H}, \text{BDT}_{tH}, \text{BDT}_{t\bar{t}W}, \text{BDT}_{\text{Other}}) = \text{BDT}_{t\bar{t}W}$
Other	$2lSS$ pre-MVA selection $\text{Max}(\text{BDT}_{t\bar{t}H}, \text{BDT}_{tH}, \text{BDT}_{t\bar{t}W}, \text{BDT}_{\text{Other}}) = \text{BDT}_{\text{Other}}$

Table 4.3.3: Definitions of the MVA-based CRs and SRs used in the  $2lSS$  channel.

#### 4.3.2.3 $3l$ channel MVA regions

The  $3l$  BDT classifies the events into six categories:  $t\bar{t}H$ ,  $tH$ ,  $t\bar{t}W$ ,  $t\bar{t}Z$ ,  $t\bar{t}$  and  $VV$ . The training is based on 13 kinematic variables, including the number of jets and  $b$ -jets, MET, leptons and jets transverse momentum, angular separation between leptons and jets and their pairs, invariant masses of lepton pairs, azimuthal separation between leading jet and MET. The performance of the trained BDT is shown as a confusion matrix in Figure 4.3.6 (b). The  $t\bar{t}H$  and  $tH$  events are correctly classified in 45% and 55% of the cases. Around 40% of the  $t\bar{t}Z$  events are incorrectly classified as  $t\bar{t}H$  or  $t\bar{t}W$ . The regions are defined iteratively based on the BDT score cuts. First, the  $t\bar{t}H$  process is considered: the Poisson counting significance  $Z$  (see Section 4.5) is maximised over all other backgrounds by adjusting the requirements on the set of BDT scores. The same procedure is repeated for the rest of the classes in the following order:  $tH$ ,  $t\bar{t}W$ ,  $t\bar{t}Z$ ,  $VV$  and  $t\bar{t}$ . Once an event is assigned to a region, it gets excluded from the consideration of the subsequent

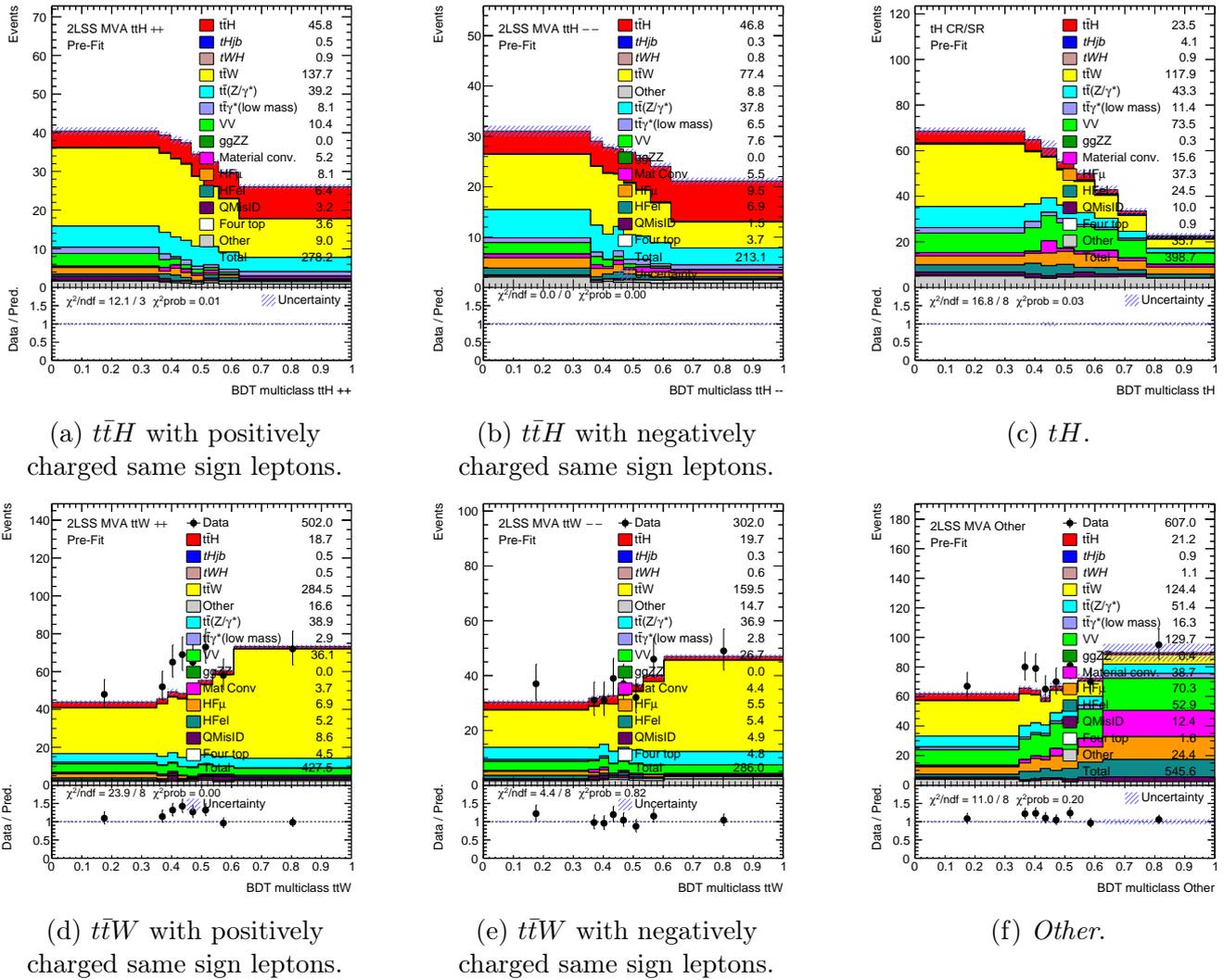


Figure 4.3.7: Distribution of data and MC prediction in the MVA-based signal and control regions in the  $2lSS$  channel.

region definitions. The events not included in any of these six regions are assigned to the *Other* region. The summary of the region definition requirements is shown in Table 4.3.4. Figure 4.3.8 illustrates the composition of signal and MVA-based regions in  $3l$  channel. The data in the SR is blinded.

## 4.4 Systematic uncertainties

The systematic uncertainties affect the measurement of the parameters of interest: the signal strength or the  $\mathcal{CP}$ -mixing angle via the normalisation of signal and background samples and the shape of the discriminates. There are three main sources of systematic uncertainties: instrumental uncertainties associated with the reconstruction, identification and calibration of final state particles (Section 4.4.1), theoretical uncertainties on the signal process modelling (Section 4.4.2) and uncertainties on the background estimations (Section 4.4.3).

An independent nuisance parameter is associated with each systematic source and is included in the likelihood function in Eq. 4.5.2 (see Section 4.5). A single nuisance parameter corresponding

Regions	Selections
$t\bar{t}H$ SR	$t\bar{t}H > 0.2$ .
$tH$ SR	$tH_{jb} > 0.25, t\bar{t}H < 0.2$ .
$t\bar{t}W$ CR	$t\bar{t}W > 0.3, t\bar{t}H < 0.2, tH_{jb} < 0.25$ .
$t\bar{t}Z$ CR	$t\bar{t}Z > 0.45, t\bar{t}H < 0.2, tH_{jb} < 0.25, t\bar{t}W < 0.3$ .
$VV$ CR	$VV > 0.65, t\bar{t}H < 0.2, tH_{jb} < 0.25, t\bar{t}W < 0.3, t\bar{t}Z < 0.45$ .
$t\bar{t}$ Region	$t\bar{t} > 0.25, t\bar{t}H < 0.2, tH_{jb} < 0.25, t\bar{t}W < 0.3, t\bar{t}Z < 0.45, VV < 0.65$ .
<i>Other</i> Region	$t\bar{t} < 0.25, t\bar{t}H < 0.2, tH_{jb} < 0.25, t\bar{t}W < 0.3, t\bar{t}Z < 0.45, VV < 0.65$ .

Table 4.3.4: Definitions of the MVA-based CRs and SRs used in the  $3l$  channel.

to a given uncertainty is used across channels and samples.

#### 4.4.1 Instrumental uncertainties

Systematic uncertainties related to the lepton reconstruction and identification, non-prompt lepton BDT working point calibrations, jet calibration,  $b$ -tagging and trigger efficiencies are referred to as instrumental uncertainties. The scale factors used to adjust the simulations to the data come with associated uncertainties and are applied as a weight to the event.

The uncertainty associated with the integrated luminosity for the full Run 2 data impacts each process with an overall 0.83% normalisation uncertainty [155].

The pile-up reweighting scale factor is applied to simulated events to account for the difference between the predicted and measured pile-up distribution. The uncertainty on the reweighting is estimated by varying the nominal scale factor with the data statistics.

The uncertainties coming from the trigger, reconstruction, identification, and isolation efficiencies are considered for both electrons and muons. Additionally, a separate uncertainty from the low  $p_T$  muon identification and an uncertainty from track to vertex association are taken into account. These uncertainties are applied as event weights. The resolution (combined momentum resolution and sagitta corrections) and  $p_T$  scale uncertainties are implemented as  $p_T$  corrections on muons. The uncertainties originating from the electron energy scale and energy resolution are applied as energy corrections.

The jet uncertainties originate from the jet vertex tagger, jet energy scale and jet energy resolution. The uncertainty coming from the JVT accounts for the efficiency of passing the JVT selection and originates from the potential mismodelling that could impact the efficiency estimation [156] and is applied as a reweighting. The 30 uncertainties considered for JES uncertainties arise from

- the jet calibration method;
- $\eta$  inter-calibration, accounting for physics mismodeling, statistical uncertainty component and non-closure;

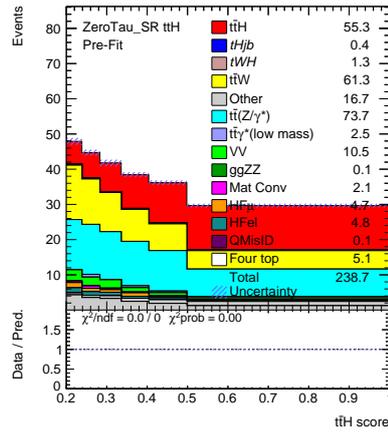
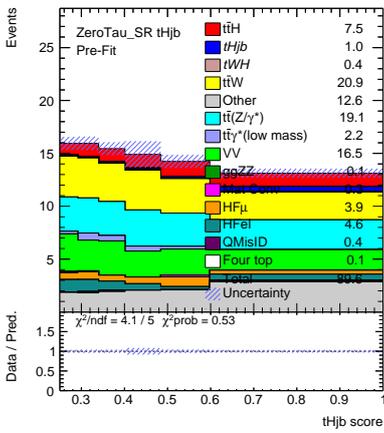
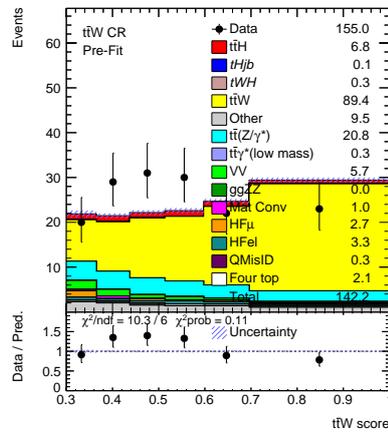
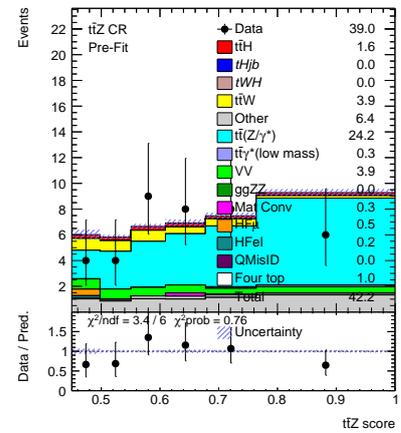
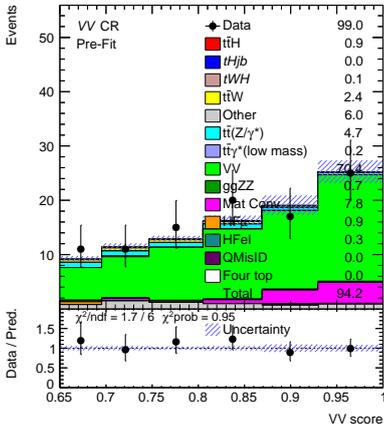
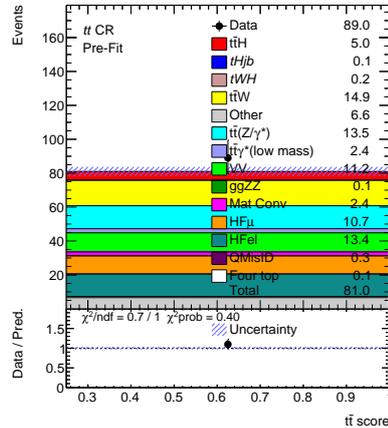
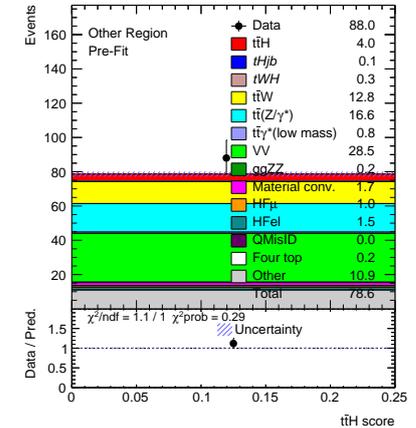

 (a) Signal MVA-based region with  $3l t\bar{t}H$ .

 (b) Signal/control MVA-based region with  $3l tH$ .

 (c) Control MVA-based region enriched with  $t\bar{t}W$  events.

 (d) Control MVA-based region enriched with  $t\bar{t}Z$  events.

 (e) Control MVA-based region enriched with  $VV$  events.

 (f) Control MVA-based region enriched with  $t\bar{t}$  events.

 (g) Control MVA-based region with events categorised to *Other* class.

 Figure 4.3.8: Distribution of data and MC prediction in the MVA-based signal and control regions in the  $3l$  channel.

- high jet  $p_T$  extrapolation;
- pile-up comprising uncertainties in the  $\mu$ ,  $N_{PV}$  and per-event  $p_T$  density modelling and

uncertainty in the residual  $p_T$  dependence;

- jet flavour where the uncertainties are coming from the differences in the response of the detector to jets originating from different types of quarks or from gluons;
- punch-through uncertainty accounting for mismodelling of the global sequential calibration.

The JER uncertainties are characterized using 13 separate sources of uncertainty. These uncertainties account for the differences between the nominal and smeared MC results. In this process, a smearing factor is applied as a scale factor to the MC jets, enhancing their resolution to match the jet resolution observed in the data, including its associated error [157]. Both JES and JER uncertainties are applied as  $p_T$  corrections.

A total of 83 uncorrelated uncertainties for  $b$ -,  $c$ - and light-jets are considered in the  $b$ -tagging calibration. These uncertainties originate from statistical sources, uncertainties related to the detector calibration and physics modelling. The  $b$ -tagging correction factors and their uncertainties for  $b$ -jets are estimated from  $t\bar{t}$  data events [158]. In the case of  $c$ -jets, the estimation is based on jets originating from  $W$  bosons in  $t\bar{t}$  events [159]. The di-jets are used to derive the corrections for light-flavour jets [160]. The  $b$ -tagging related uncertainties are accounted for as reweighting.

The uncertainties on the MET are propagated from the energy scale and resolution uncertainties of the detected objects, leptons and jets. Uncertainties from both sources are applied as  $p_T$  corrections.

#### 4.4.2 Signal modelling uncertainties

There are multiple sources of uncertainties impacting the  $t\bar{t}H$  process. The modelling uncertainties are applied as the shape uncertainties. The initial-state radiation uncertainty is estimated using weights in the matrix element calculation and in the parton shower: the factorisation  $\mu_F$  and renormalisation  $\mu_R$  scales are simultaneously varied up (down) by a factor of 0.5 (2) in the ME and  $\alpha_s^{ISR}$  is varied in the PS. The choice of a generator might introduce additional uncertainties. To assess the uncertainty due to PS and hadronization, a comparison is made between the POWHEG-BOX+PYTHIA8 and POWHEGBOX+HERWIG7 generators, and to estimate the ME uncertainty, the POWHEGBOX+PYTHIA8 is compared with MADGRAPH5\_AMC@NLO+PYTHIA8.

#### 4.4.3 Background uncertainties

One of the biggest sources of systematic uncertainties in this analysis is coming from the shape uncertainty of the  $t\bar{t}W$  process. Similarly to  $t\bar{t}H$ , the uncertainty on the  $t\bar{t}W$  predictions due to missing higher-order QCD correction is derived by varying  $\mu_F$  and  $\mu_R$ . The uncertainty originated from ambiguities in the ME and PS algorithms, and its parameter choices are estimated based on comparing the nominal Sherpa prediction with the prediction from MADGRAPH5\_AMC@NLO+PYTHIA8FXFX (see Table 4.2.1). The PS model uncertainty is evaluated by comparing the POWHEG+PYTHIA8 and POWHEG+HERWIG7 generators. The ME and PS-related uncertainties are derived separately for the QCD and EW components, keeping the

correlation among samples. An uncertainty on the relative yields of the EW sample within the QCD+EW  $t\bar{t}W$  templates (varying the EW fraction between the values of the SHERPA and the MADGRAPH5\_AMC@NLO prediction) is also applied. The uncertainties related to the PDF modelling are estimated by varying the  $\alpha_S$  and the PDF set.

The uncertainties due to missing higher orders in QCD in  $t\bar{t}Z$  modelling are estimated following the  $t\bar{t}H$  procedure by varying the nominal  $\mu_F$  and  $\mu_R$  by factors of 0.5 and 2. Uncertainties in the additional jets modelling are derived by performing ISR  $\alpha_S$  variations. The rest of the uncertainties, related to the PS, UE and hadronisation modelling, were evaluated by comparing the nominal MADGRAPH5\_AMC@NLO+PYTHIA8 MC simulation with MADGRAPH5\_AMC@NLO+HERWIG7.

The di-boson  $VV$  modelling uncertainties originate from missing higher orders in QCD (with the same estimation procedure as for previously discussed samples) and from data-driven  $N_{jets}$ -dependent corrections (see Section 4.2). The latter is derived based on the fit function parameter uncertainties.

The uncertainties in higher-order cross-sections are considered for the rare background processes, except for those whose normalisation is measured in the fit. The following uncertainties are applied to the theory cross-sections:  $\pm 30\%$  on  $VH$  and  $VVV$ ,  $\pm 35\%$  on three-top,  $\pm 50\%$  on  $t\bar{t}WW$  and  $t\bar{t}WZ$  and  $\pm 5\%$  on  $t\bar{t}Z$ . The four-top normalisation uncertainty is set to  $^{+70\%}_{-15\%}$  where the upper variation comes from the central value of the ATLAS measurement [161] and the down variation comes from the theoretical four-top uncertainty.

For the non-prompt lepton backgrounds from heavy-flavour decays and conversions, the uncertainties in the modelling of the PLIV input variables are included as variations affecting the shape of distributions in each region without affecting the normalisation. The  $t\bar{t}$  PS uncertainty is separately assessed for electrons and muons by comparing POWHEG+HERWIG7 and POWHEG+PYTHIA8 simulations. To account for the  $t\bar{t}$  production in association with HF jets modelling uncertainties, an uncorrelated 50% uncertainty is assigned to the  $t\bar{t} + b$  and  $t\bar{t} + c$  processes. Uncertainties on internal (50%) and material (10%) conversion extrapolations are applied. This estimation is derived from the extrapolation of the electron conversion background from  $Z$ -enriched to  $t\bar{t}$ -enriched regions and is derived from the residual data/simulation mismodelling in validation regions. The MC-based background from electrons with a misidentified charge has a 20% systematic uncertainty.

## 4.5 Statistical analysis techniques

To evaluate the agreement between the MC prediction and the observed data, we make use of a template fit. The template fit method is a statistical technique to extract signal events from data that contain both signal and background contributions. The key idea is to use templates, which are histograms in this analysis, that represent the expected distributions of some observable quantities for both signal and background processes. The templates are defined for MC event or data events. A binned profile likelihood fit, described in Section 4.5.1, adjusts the templates simultaneously in the SRs and the CRs, providing the signal and background yields best matching the target distributions. The statistical tests used to define the confidence intervals are detailed

in Section 4.5.2. The uncertainties treatment is discussed in Section 4.5.3 and Section 4.5.4. The statistical analysis is performed within the TREXFITTER framework [162].

### 4.5.1 Binned profile likelihood fit

A binned profile likelihood function is defined over all analysis bins. The expected number of events in the  $i$ -th bin is given by:

$$\lambda_i = \mu \cdot S_i + \sum_b^{\text{bkg}} k_b \cdot B_{bi}. \quad (4.5.1)$$

Here  $S_i$  and  $B_{bi}$  are expected signal and background yields in the  $i$ -th bin. The parameters  $\mu$  and  $k_b$  rescale the corresponding yields:  $\mu$  is the signal strength defined in Section 1.9, and  $k_b$  are the normalisation factors (NFs) of the fitted backgrounds. The probability of observing  $N$  events given the expected number of events  $\lambda$  is expressed via the Poisson probability distribution  $\mathcal{P}(N | \lambda)$ . The likelihood function can then be defined as a product of Poisson distributions over the analysis bins. Such a function depends on the parameter(s) of interest (POI). In this example the POI is the signal strength  $\mu$ . However, with the presence of systematic uncertainties, the likelihood function must encode the information about them through nuisance parameters (NPs):  $\theta = \{\theta_a, \theta_b, \dots\}$ . With the introduction of  $\theta$ , the final general form of the likelihood function is:

$$\mathcal{L}(\mu, \theta) = \mathcal{L}(\mu)\mathcal{L}(\theta) = \prod_i^{\text{bins}} \mathcal{P}(N_i | \lambda_i) \times \prod_j^{\text{NPs}} \mathcal{F}_j(\theta_j). \quad (4.5.2)$$

The first term describes the analysis regions bins where  $S_i$  and  $B_{bi}$  acquire a dependence on  $\theta$ :

$$S(\theta) = S_0 \times \prod_k^{\text{NPs}} \nu(\theta_k), \quad B(\theta) = B_0 \times \prod_k^{\text{NPs}} \nu(\theta_k), \quad (4.5.3)$$

so that  $S_0$  and  $B_0$  are the nominal expected signal and background yield values, and  $\nu(\theta_k)$  are the functions parametrizing dependency of the yields on the nuisance parameters  $\theta_k$ . These affect both the shape and normalisation of the distributions. The NPs are also incorporated into the likelihood with the functions  $\mathcal{F}_j$  (Eq. 4.5.2), taken to be Gaussian or log-normal distributions.

The fit is performed by adjusting the free-floating parameters (POI  $\mu$  and the backgrounds normalisation factors  $k_b$ ) and the NPs in a way the likelihood function reaches its maximum. The statistical bin-by-bin uncertainties of adjusted distributions are also considered in the fits through the Beeston-Barlow method [163] using dedicated parameters  $\gamma_i$ . Note that the likelihood function doesn't contain the constraints  $\mathcal{F}$  corresponding to the dominant backgrounds NFs left as free fit parameters. The contribution of smaller backgrounds is controlled by nuisance parameters.

### 4.5.2 Test statistics

The best-fit values are then used to build a statistical test function  $q$  of the type  $\Delta \ln(\mathcal{L})$  for the definition of confidence intervals and hypothesis testing. Two likelihood functions are considered:

- conditional maximum likelihood estimator for a given POI value,
- unconditional maximum likelihood estimator for a POI and NPs values maximizing the likelihood function.

For example, when measuring the signal strength, the test statistics function takes the form shown in Eq. 4.5.4:  $\hat{\theta}(\mu)$  is the conditional maximum likelihood estimator, or profiled value, of  $\theta$  for a given value of  $\mu$  ( $\mu = 0$  for the background-only hypothesis), and  $\hat{\mu}$ ,  $\hat{\theta}$  are the fitted parameters. With  $\mu = \hat{\mu}$ ,  $\hat{\theta}(\hat{\mu}) = \hat{\theta}$  and the test statistics function becomes  $q(\hat{\mu}) = 0$ , with:

$$q(\mu) = -\ln \frac{\mathcal{L}(\mu, \hat{\theta}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \quad (4.5.4)$$

As discussed in Section 1.9, the  $\mathcal{CP}$ -properties of the  $t\bar{t}H$  are parametrized via two parameters:  $\alpha$ : the  $\mathcal{CP}$ -mixing angle, and  $\kappa_t$ : the coupling modifier. In this case, the test statistic might be expressed in two ways:

$$q(\alpha) = -\ln \frac{\mathcal{L}(\alpha, \hat{\kappa}'_t(\alpha), \hat{\theta}(\alpha))}{\mathcal{L}(\hat{\alpha}, \hat{\kappa}'_t, \hat{\theta})}, \quad (4.5.5)$$

$$q(\alpha, \kappa'_t) = -\ln \frac{\mathcal{L}(\alpha, \kappa'_t, \hat{\theta}(\alpha, \kappa'_t))}{\mathcal{L}(\hat{\alpha}, \hat{\kappa}'_t, \hat{\theta})} \quad (4.5.6)$$

Here  $\hat{\alpha}$ ,  $\hat{\kappa}'_t$  and  $\hat{\theta}$  are the values maximising the likelihood function. In Eq. 4.5.5,  $\hat{\kappa}'_t(\alpha)$  and  $\hat{\theta}(\alpha)$  are the values of  $\kappa'_t$  and  $\theta$ , that maximise the likelihood function for a fixed value of  $\alpha$ . Considering the asymptotic approximation,  $q_\alpha$  is distributed as a  $\chi^2$  distribution with one degree-of-freedom for the true value of  $\alpha$ . The  $\hat{\theta}(\alpha, \kappa'_t)$  in Eq. 4.5.6 is the conditional maximum likelihood estimator for each pair of values  $(\alpha, \kappa'_t)$ .

The  $p$ -value is defined as the probability for the test statistic  $q(\mu)$  to be larger than the observed one  $q_{obs}$ , under the null hypothesis (considering  $\mu$  as a POI in the following examples):

$$p_\mu = \int_{q_{obs}}^{\infty} f(q(\mu)|\mu) dq(\mu), \quad (4.5.7)$$

where  $f$  is the probability density function of  $q(\mu)$ . A  $p$ -value indicates the degree of consistency between the observed data and the hypothesis. The obtained  $p$ -value is then reported in terms of the significance  $Z$ : the number of standard deviations for a normal distribution  $\phi$  producing a one-sided tail integral equal to the  $p$ -value,  $\phi^{-1}(1 - p)$ . Measurement with  $Z = 3$  is called an evidence for a processes, while  $Z = 5$  is required for an observation. The maximum possible signal strength that can be excluded within a certain confidence interval is defined using the notion of  $CL_s$ :

$$CL_s = \frac{p_{s+b}}{p_b}, \quad (4.5.8)$$

with  $p_{s+b}$  and  $p_b$  being the  $p$ -values in the signal+background hypothesis and the background-only hypothesis respectively. To get the limit value  $\mu_{lim}$  at 95% confidence level, the argument of the

$p$ -value is varied until  $CL_s = 0.05$ . So the values of  $\mu$  corresponding to  $CL_s < 0.05$  are excluded at 95%. Another way to set confidence intervals is by likelihood intervals. Since by the Wilks' theorem [164] the  $q(\mu)$  asymptotically follows a  $\chi^2$  law with the number of degrees of freedom being the dimensionality of  $\mu$ , the likelihood  $2\Delta\ln\mathcal{L} = \phi^{-1}(1 - p)$  intervals are approximate confidence intervals with confidence level of  $1 - p$ .

For the  $\mathcal{CP}$ -measurement, to get the confidence interval on  $\alpha$  at a give confidence level  $X$ , one-dimensional scans of  $q(\alpha)$  (see Eq. 4.5.5) are performed along the  $\alpha$  values range, while profiling all other parameters. Profiling NPs means minimising the log-likelihood with respect to them for each point in the  $\alpha$  range. The resulting curve is called a profile likelihood scan. It is used to set up the POI's exclusion limits and estimate the NPs uncertainties when considered a 68% confidence level. The  $q(\alpha, \kappa'_i)$  (see Eq. 4.5.6) follows a  $\chi^2$  distribution, and requires a 2-dimensional likelihood scans on the  $(\alpha, \kappa'_i)$  coordinates.

### 4.5.3 Uncertainties treatment

The uncertainties impact on the fit is studied via nuisance parameter pulls, rankings and correlation matrix. The correlation matrix is defined as the inverse of the Hessian matrix  $H^{-1}$ , where the matrix element of  $i$ -th row and  $j$ -th column is defined as

$$(H_{\mathcal{L}(\theta)})_{i,j} = \frac{\delta^2\mathcal{L}}{\delta\theta_i\delta\theta_j}, \quad (4.5.9)$$

with  $\mathcal{L}$  being the likelihood function defined in Eq. 4.5.2. A nuisance parameter pull is given by the difference between the pre- and post-fit values of the nuisance parameter, divided by its pre-fit uncertainty. A nuisance parameter is called constrained if the post-fit uncertainty is smaller than the pre-fit one, meaning that the fit extracted some additional information from the data about the considered NP. In case of difference between the pre- and post-fit values, a NP is referred to as pulled. Large pulls mean that data favor a NP value different from its initial estimate.

The pre-fit impact of a nuisance parameter is defined as the difference in POI between two fits: one with the NP fixed to its nominal value  $\theta$  and another with the NP fixed to  $\theta \pm \Delta\theta$ . The post-fit impact is calculated in a similar way by varying the best-fit  $\hat{\theta}$  value. The NPs considered in the fit can be ordered by their impacts to get the NP ranking. In the  $\mathcal{CP}$ -analysis, the NP impact is assessed based on the likelihood scans. The  $1\sigma$  intervals of POI in the nominal model and in the model with the uncertainty source removed (i.e. the corresponding NP fixed) are compared: the impact is given by subtracting in quadrature the lengths of these intervals.

### 4.5.4 Asimov and hybrid fit setups

Before getting the results using data collected by the ATLAS detector, fits on Asimov and hybrid datasets are performed to obtain the expected uncertainties on the measured parameter.

In the Asimov dataset, each bin's value is set to the expected event yield, excluding statistical uncertainties. This dataset is created by replacing the data with the sum of the signal and background predictions. Fitting this dataset using the nominal values of the nuisance parameters from simulations recovers the input values, providing insights into expected constraints and correlations

between NPs. The expected results from an Asimov dataset fit give the median outcome under the signal-plus-background hypothesis.

In the hybrid approach, Asimov datasets are used in signal regions, while real data is used in control regions. A hybrid fit is a conditional fit to the Asimov dataset performed using background NP values determined from an unconditional fit to data. This setting provides a more realistic background model.

## 4.6 STXS setup

To perform the measurements of the  $t\bar{t}H$  signal strength or of the  $\mathcal{CP}$ -mixing angle in the STXS setup, a reconstruction of the Higgs boson transverse momentum  $p_T^H$  is needed.

### 4.6.1 Event kinematics in the $2lSS$ and $3l$ channels

To understand better the challenges of the Higgs  $p_T$  reconstruction, the events kinematics and reconstruction efficiencies were studied in the analysis phase space of the  $2lSS$  and  $3l$  channels.

#### 4.6.1.1 $2lSS$ channel

In the  $2lSS$  channel, two same-sign leptons have to come from different subsystems: one from the Higgs boson decay and another from the top-antitop system. The Higgs decay should then be a semi-leptonic  $H \rightarrow WW$  or  $H \rightarrow \tau\tau \rightarrow lX$  decay. One of the top quarks should decay semi-leptonically, and another hadronically. Figure 4.6.1 shows that after the SR selection at the reconstruction level, around 81% of the selected events come from the  $H \rightarrow WW$  decay, and around 16% from  $H \rightarrow \tau\tau$  decay. The presence of multiple leptons in the final state also infers the presence of neutrinos, resulting in around 2 neutrinos per event (see Figure 4.6.2), producing significant missing transverse energy.

Comparing the reconstructed events in the  $2lSS$  SR with the truth-level event information, it can be seen that 25% of the selected events are not actually events with two same-sign leptons. Instead, such events have 3 or more leptons at the truth level, but some of these leptons did not pass the selection criteria or were out of acceptance. The number of matched truth- and reconstructed objects (both leptons and jets) can be computed to estimate the goodness of the reconstructed objects selected within the SRs. Objects are considered matched if they are close enough, i.e. within  $\Delta R < 0.3$ . If multiple candidates satisfy this requirement, the one with the smaller  $\Delta R$  is chosen. Around 85% of the truth leptons are matched to Higgs decay products, and 88% to the top-antitop decays. Unlike leptons, jets are less often matched, resulting in the matching rate of 43% in Higgs decays and 69% in top-antitop decays, where the  $b$ -jets are matched in 85% of the events. The difference in matching rate can be understood by the fact that jets coming from Higgs decays tend to have lower  $p_T$  than the ones from the top decays, and so they are failing the selection criteria more often.

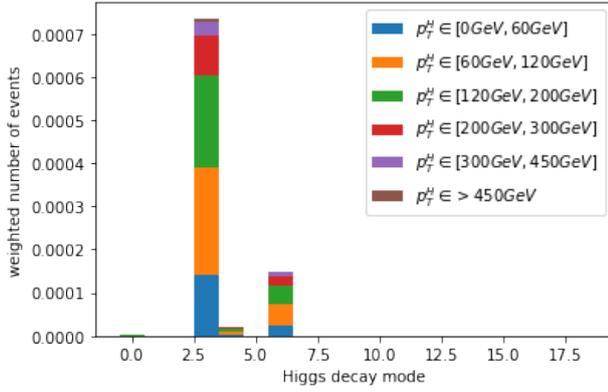


Figure 4.6.1: Number of events per Higgs decay mode in the  $2lSS$  SRs, where 3 corresponds to  $H \rightarrow WW$  and 6 to  $H \rightarrow \tau\tau$ .

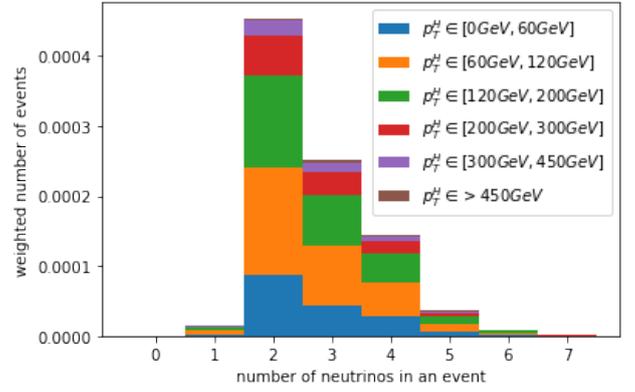


Figure 4.6.2: Number of neutrinos in the final state in  $2lSS$  SRs.

#### 4.6.1.2 $3l$ channel

The possible composition of the decay chain in the  $3l$  channel is more diverse than in the  $2lSS$  channel. To get three leptons in the final state, one or two leptons should come from the Higgs decay and two or one leptons from the top-antitop quark decays. Therefore, the Higgs boson can decay leptonically while one of the top quark can decay leptonically and another hadronically, or the Higgs boson can decay semi-leptonically while both top quarks decay leptonically. As shown in Figure 4.6.3, in 75% of the  $3l$  events, the Higgs decays to a pair of  $W$  bosons. The next contribution comes from  $H \rightarrow \tau\tau$  decay (14%) followed by  $H \rightarrow ZZ$  (7%) and  $H \rightarrow \mu\mu$  (3%). The average number of neutrinos in the  $3l$  channel final state is around 3 (see Figure 4.6.4).

Some events in the  $3l$  signal region do not actually contain three leptons in the final state at truth level. However, almost 94% of the events have the correct number of reconstructed leptons. There are also around 5% of the events with two or fewer truth leptons and 1% with four or more truth leptons. The matching rate between truth- and reconstructed leptons reaches 92% for Higgs decay and 96% for top-antitop decays, which is significantly higher than in the  $2lSS$  channel. The jet matching rate is similar to the one observed in the  $2lSS$  channel with 44% matched jets in Higgs decays and 71% in top quark decays.

### 4.6.2 Higgs $p_T$ reconstruction

The reconstruction of  $p_T^H$  in  $t\bar{t}H$  multilepton events is complicated due to the presence of several neutrinos in the final state. The neutrinos come from both Higgs and top-quark decays, making the exact reconstruction of the Higgs  $p_T$  impossible. Several approaches were tested to address this problem.

First, variables highly correlated to the  $p_T^H$  were searched. Three low-level kinematic parameters were found to be correlated with the Higgs  $p_T$ : the scalar sum of the two leading leptons, the scalar sum of the two leading jets and the missing transverse energy. The first one shows the highest correlation, while the correlation of the latter two with  $p_T^H$  is around 50%.

Another way to access the Higgs  $p_T$  is to reconstruct the decay chain directly. Here, three

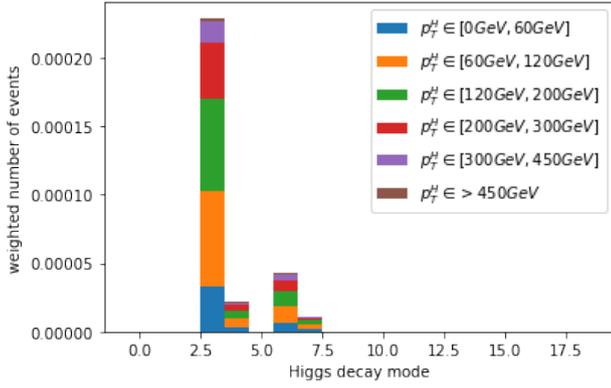


Figure 4.6.3: Number of events per Higgs decay mode in the  $3l$  SR, where 3 corresponds to  $H \rightarrow WW$  and 6 to  $H \rightarrow \tau\tau$ .

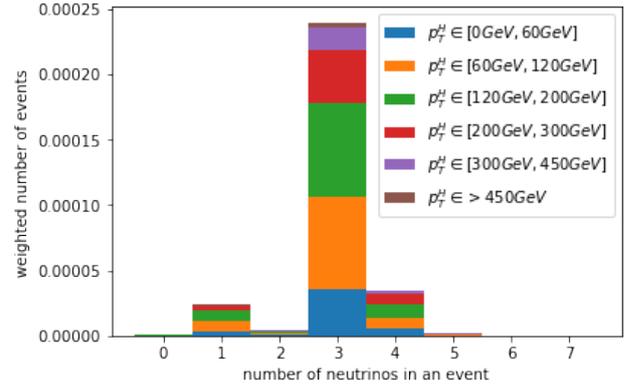


Figure 4.6.4: Number of neutrinos in final state in the  $3l$  SR.

options were considered: partial reconstruction by considering either the Higgs or the top anti-top system and full event consideration. The Higgs reconstruction based on its visible final state particles shows a 75% correlation with the  $p_T^H$  at the truth level. Potentially, this approach might outperform the one based on the low-level variables. However, given the low truth-reconstruction matching rate for jets in the Higgs decay, there is not enough information in the reconstructed final state particles to proceed with partial reconstruction of the Higgs boson. Partial reconstruction of the top anti-top system does not look promising already at the truth level due to a relatively low correlation to the  $p_T^H$ . At the reconstruction level, the correlation reaches 64%. The full reconstruction of the top anti-top system is complicated due to the presence of neutrinos. Moreover, the  $t\bar{t}$  system is sensitive to the extra jet radiation.

In the tentative to make use of the information from both the Higgs and the top anti-top system for the Higgs  $p_T$  reconstruction, different neural network algorithms were tested: regression DNN and GNN predicting continuous value of  $p_T^H$ , and DNN and GNN classifiers, categorising events into an STXS bin. The highest performance was achieved using a GNN-based classifier, described in Section 4.6.3.

### 4.6.3 GNN to reconstruct the STXS bins

To predict the Higgs  $p_T$  STXS bins, we use a Graph Neural Network (GNN) [165] based on the results presented in [166]. Each event is represented as a graph with nodes corresponding to the final state particles, edges related to kinematics between particles and global attributes with event-level information. The graphs are then processed by the GNN model, resulting in the prediction of the Higgs  $p_T$  STXS bins.

To fill a graph, we use the following information independently for events with two same-sign leptons and three leptons in the final state:

1. Nodes:  $p_T$ ,  $\eta$ ,  $\phi$ ,  $p$ , particle type (lepton, jet or MET),  $b$ -tagging score and charge;
2. Edges:  $\Delta\eta$ ,  $\Delta\phi$ ,  $\Delta R$  between all different pairs of objects (lepton, jet or MET),  $p_T$ ,  $p$  and invariant mass for the sum of the 4-vectors of each different object pairs;

- Global: number of jets  $N_{jets}$ , number of forward jets  $N_{jets, fwd}$ , number of pile-up  $\mu$ ,  $H_T(jets)$ ,  $H_T(leptons)$ ,  $H_T(jets + leptons)$ ; all leptons  $p_T$ ; the mass  $M$ ,  $p_T$  of the sum of all pairs of leptons and their  $\Delta\phi$  and  $\Delta\eta$ ;  $M$ ,  $p_T$  of the sum of all pairs of lepton and MET and their  $\Delta\phi$ ;  $M$ ,  $p_T$  of the sum of all combinations of two leptons and MET, and the sum of the three leptons and MET (in the  $3l$  cases).

The graphs are fully connected: all nodes are connected with all other nodes. For training, we also create target graphs containing only true STXS bin as a global attribute.

The model architecture includes sequences of multilayer perceptrons with LeakyReLU activation functions between them:  $256 \times 256$  to update edges feature vectors,  $256 \times 256$  for nodes feature vectors,  $512 \times 256 \times 128 \times 64 \times 32$  for global feature vector, with normalized sum aggregation function. The GNN is trained on the  $t\bar{t}H$  samples split into training and testing sets in the proportion of 75/25. We used the Adam optimizer with  $1e^{-3}$  learning rate for a categorical cross-entropy loss with sample weights calculated from the training data. The model for events with  $2lSS$  was trained with 450 epochs, the  $3l$  model on 200 epochs.

The confusion matrices evaluated on the testing sets for the  $2lSS$  and  $3l$  cases are presented in Fig. 4.6.6. The diagonal elements are usually dominant as expected, except in some bins in the  $3l$  channel. We also observe significant migration between neighbour bins for both final states, especially between the two first STXS bins in  $2lSS$  case, and the first and last two in  $3l$  case.

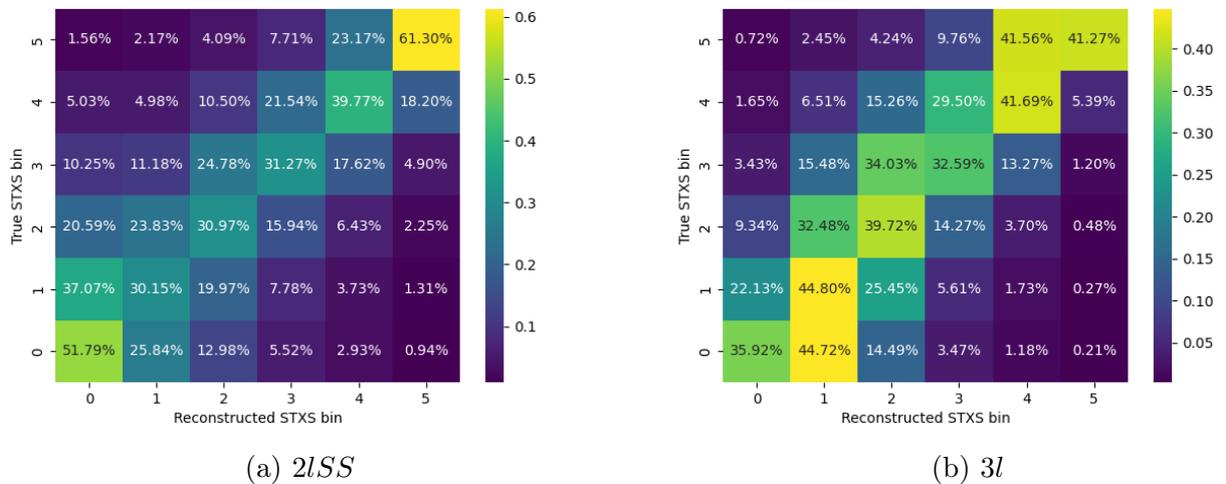


Figure 4.6.6: Confusion matrices of Higgs  $p_T$  STXS bin classification for the  $2lSS$  (a) and  $3l$  (b) channels.

Multiple variations of the described GNN architecture were tested, but the chosen setup provided the best performance.

#### 4.6.4 STXS signal region definition

There are two approaches to incorporating the STXS bin predictions into the signal regions definitions: bin the regions in Higgs  $p_t$  or split the regions by Higgs  $p_t$ .

In the first approach, the signal regions are the same as in the inclusive setup, with the difference that the region is binned based on the reconstructed Higgs  $p_T$  STXS bin instead of the

MVA score. The yield distributions in the rebinned SRs are shown in Figure 4.6.8. Since the highest  $p_T$  bin corresponding to the Higgs  $p_T$  values higher than 450 GeV has very low statistics, the two highest bins were merged, resulting in 5 final bins:  $p_0 \in [0, 60]$  GeV,  $p_1 \in [60, 120]$  GeV,  $p_2 \in [120, 200]$  GeV,  $p_3 \in [200, 300]$  GeV and  $p_{45} \in [300, +\infty]$  GeV.

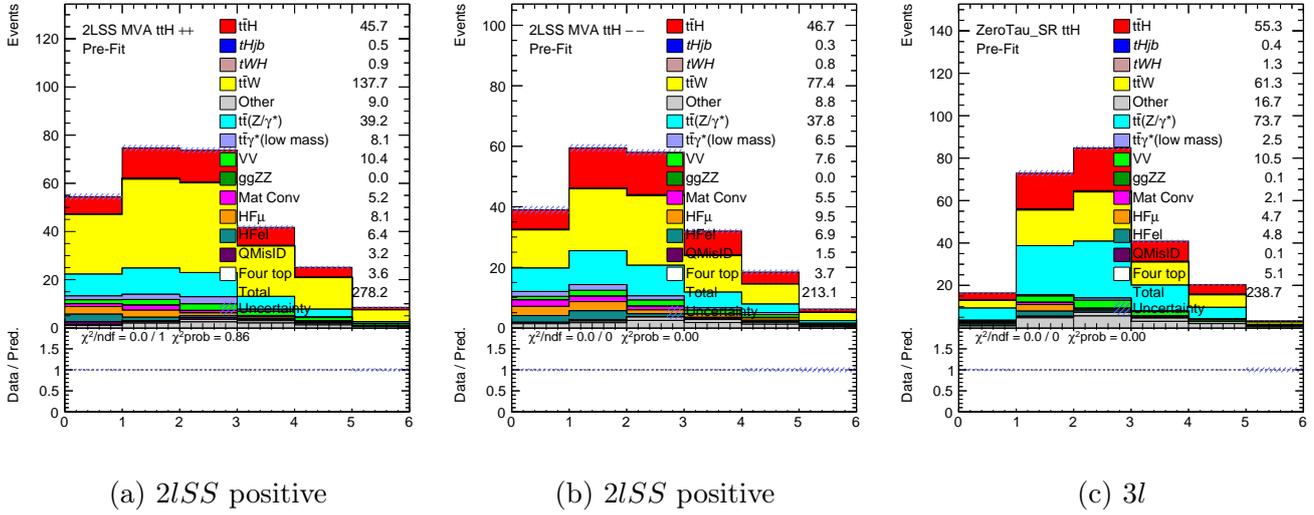


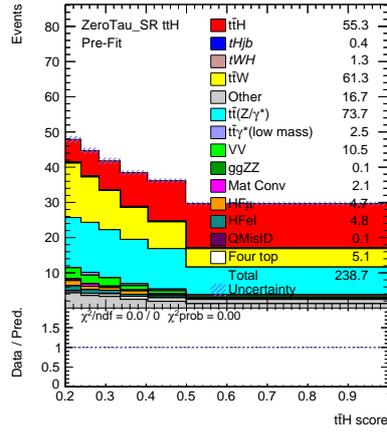
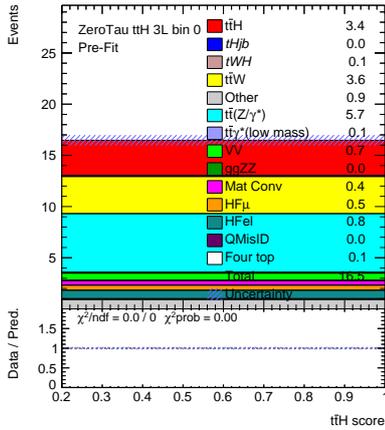
Figure 4.6.8: Signal regions in  $2lSS$  and  $3l$  channels binned in Higgs  $p_T$  STXS bins.

The second approach suggests splitting the SRs from the inclusive fit by STXS reconstructed bins, keeping the signal to background discrimination MVA score distribution in each split region. Figure 4.6.9 shows the split of the  $3l$  signal region.

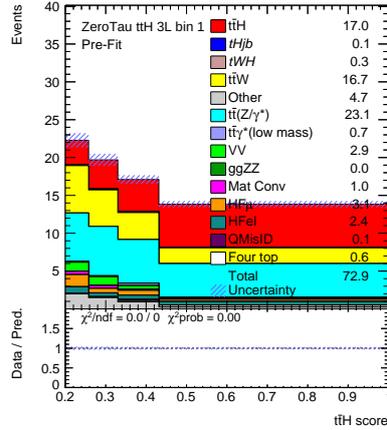
To compare these approaches, two checks were performed: an Asimov statistics-only inclusive fit measuring the signal strength and an Asimov statistics-only  $\mathcal{CP}$  fit measuring the  $\mathcal{CP}$ -mixing angle. The results of the first check are shown in Figure 4.6.10 and compared with the inclusive fit where no Higgs  $p_T$  reconstruction is used. The sensitivity to the inclusive  $t\bar{t}H$  cross-section stays the same in the fit with SRs split by  $p_T^H$ , while the fit with SRs binned in reconstructed Higgs  $p_T$  is less sensitive.

The second check compares the results of the  $\mathcal{CP}$  fits (the  $\mathcal{CP}$  fit setting is detailed in Section 4.7). In this case, the binning of the split SRs is chosen so no empty bins correspond to any process appearing in an SR. Figure 4.6.11 shows the  $\alpha$  likelihood scans and the corresponding exclusion limits. The  $\mathcal{CP}$ -odd exclusion limit is slightly better in the  $p_T^H$  split SR setup, and the confidence interval is slightly smaller in the SR  $p_T^H$  binning setting.

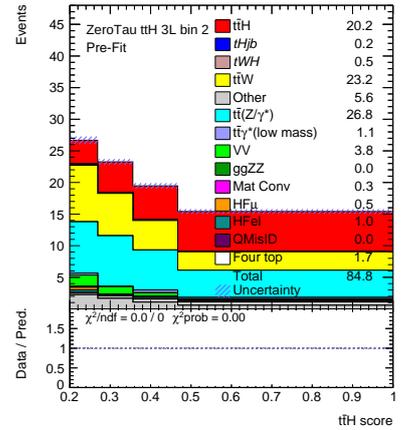
Given these performances, the second approach corresponding to splitting the SR based on the reconstructed Higgs  $p_T$  and using the MVA score in each of them was chosen as the default for the STXS studies. Multiple tests were performed to investigate the impact of the split regions binning on the fit sensitivity. The results show that for both the  $2lSS$  and  $3l$  channels, the binning optimisation procedure does not impact the fit result. The final SR binning was optimised to exclude any empty bin corresponding to the signal or background process while maximizing the number of bins.

(a)  $3l$   $t\bar{t}H$  SR.

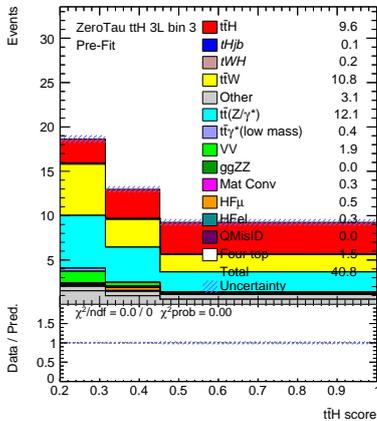
(b) STXS bin 0 SR.



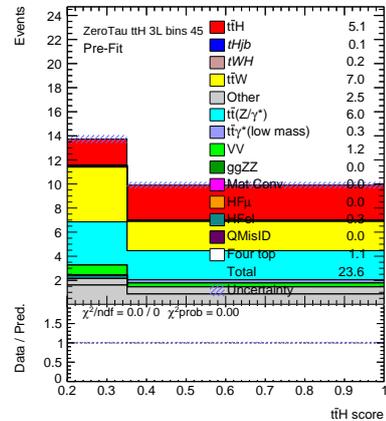
(c) STXS bin 1 SR.



(d) STXS bin 2 SR.



(e) STXS bin 3 SR.



(f) Merged STXS bins 4 and 5 SR.

Figure 4.6.9: MC prediction in the  $3l$  signal region and in  $3l$  signal regions split by STXS bins.

### 4.6.5 Checks on the GNN performance

Since the signal regions within the STXS framework use both the GNN for  $p_T^H$  reconstruction and the MVA for background-signal discrimination, it is important to understand the interplay between these two classifiers. Figure 4.6.12 shows the MVA behaviour as a function of the Higgs  $p_T$ . The plots illustrate the distribution of the MVA score for different truth STXS bins in the  $2lSS$  and  $3l$  signal regions. The tables on the bottom of Figure 4.6.12 contain information about

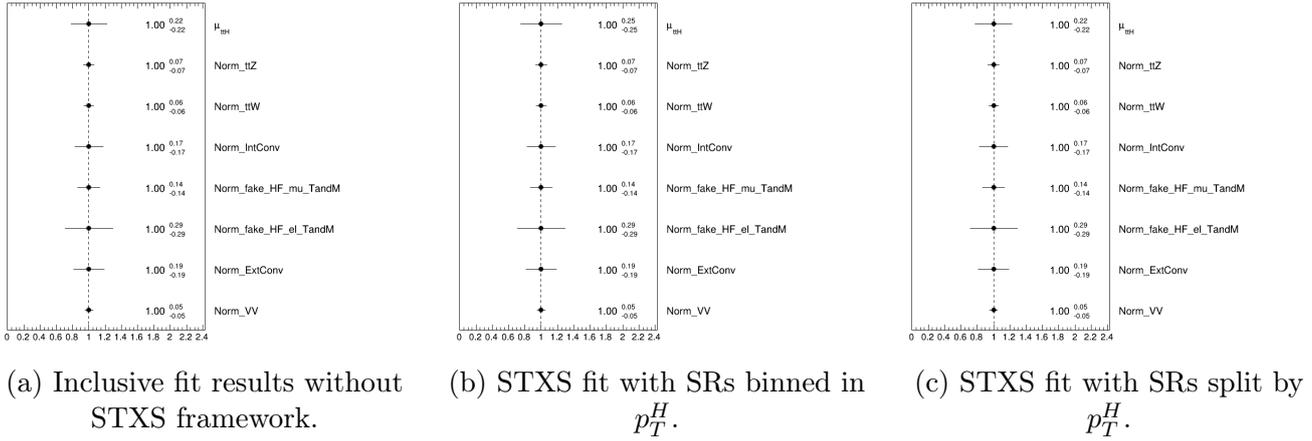


Figure 4.6.10: Fitted signal strength and dominant backgrounds normalization factors in different STXS settings.

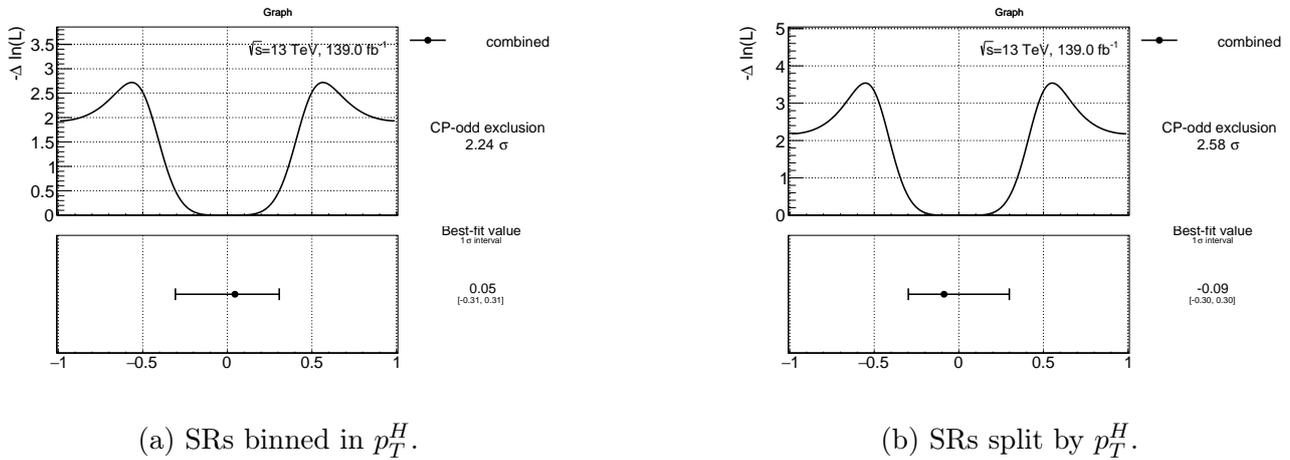


Figure 4.6.11: Results of likelihood scan over  $\alpha$  with different STXS settings in the combined  $2lSS$  and  $3l$  channels.

which percentage of events in the truth STXS bins are classified into the SR and into the CRs. A slight loss in sensitivity might be caused by events with low or high Higgs  $p_T$  being less likely attributed to the SRs than events with  $p_T^H$  in the central  $p_T$  region.

To ensure that no additional uncertainties are introduced by reconstructed STXS bins, the GNN behaviour on unblinded CRs ( $t\bar{t}W$  and *Others*) was studied. These control regions were binning according to the GNN outputs (see Figure 4.6.13). No significant differences were observed between the GNN shapes in data and MC in any CR regions.

Since the GNN classifier uses only  $t\bar{t}H$ -even events in the training, its performance might differ for  $t\bar{t}H$ -odd events used for the  $t\bar{t}H$  parametrization (see Section 4.7). Potential difference might come from the difference of kinematic properties of the  $t\bar{t}H$ -odd events. Figure 4.6.15 shows the confusion matrices between truth and predicted STXS bins for  $t\bar{t}H$ -odd sample in the  $2lSS$  and  $3l$  channels. No significant difference in the performance was observed.

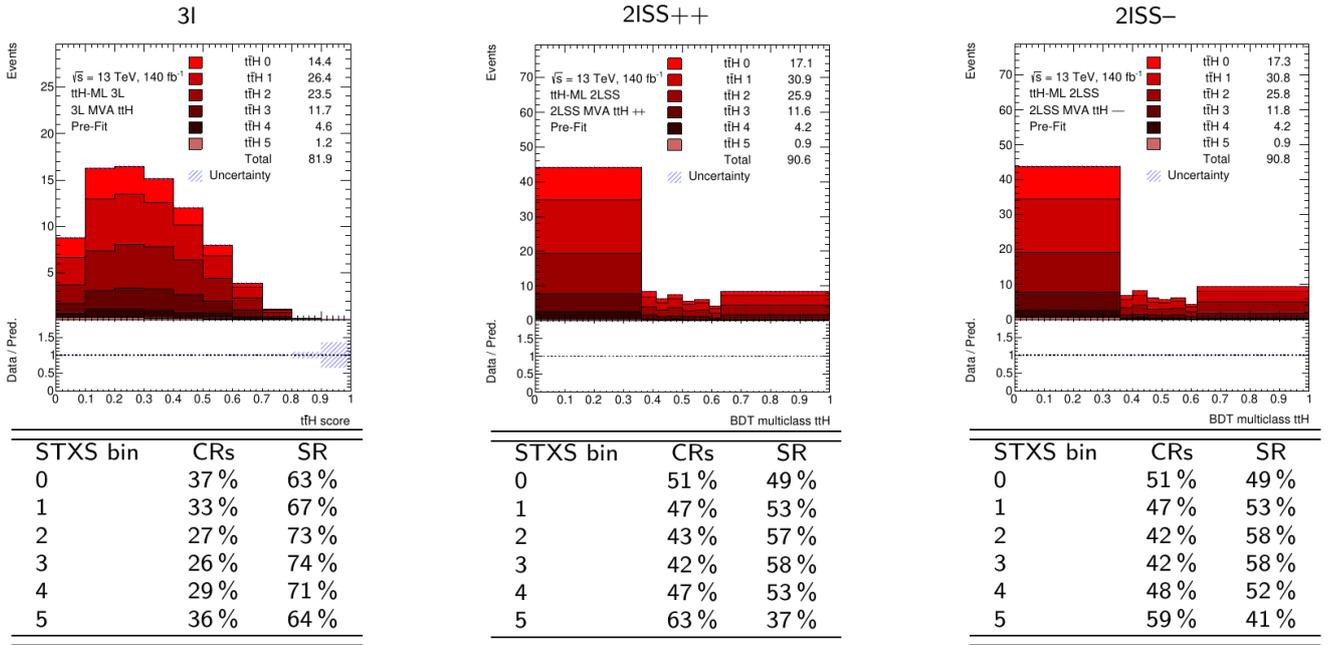


Figure 4.6.12: Distribution of the  $t\bar{t}H$  MVA score for different STXS bins in  $2lSS$  and  $3l$  (top) and percentage of the  $t\bar{t}H$  events classified into the SR and CRs for the different truth STXS bins (bottom).

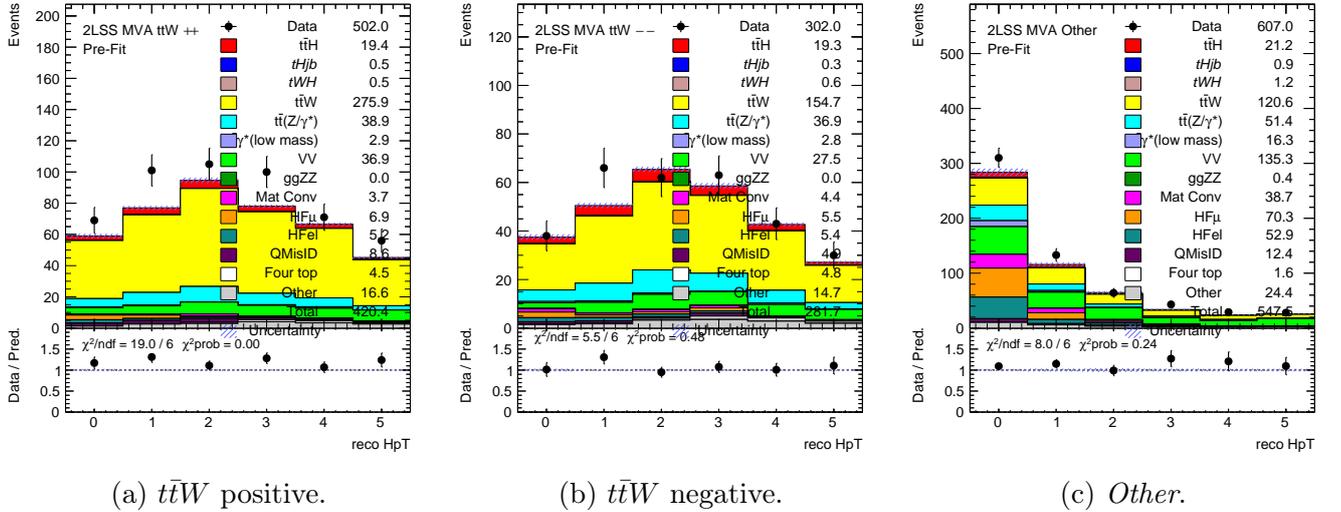


Figure 4.6.13: CRs binned according to STXS GNN outputs in the  $2lSS$  channel.

#### 4.6.6 Higgs $p_T$ sensitivity to the $CP$ -mixing angle

As mentioned in Section 1.8, the Higgs transverse momentum  $p_T^H$  is sensitive to  $CP$ -violation [38] that makes it possible to use the STXS bins as instruments in probing the  $CP$  structure of the Higgs boson. As shown in Figure 4.6.16, the  $t\bar{t}H$ -even and  $t\bar{t}H$ -odd events have distinguishable yield distributions, especially in higher Higgs  $p_T$  STXS bins.

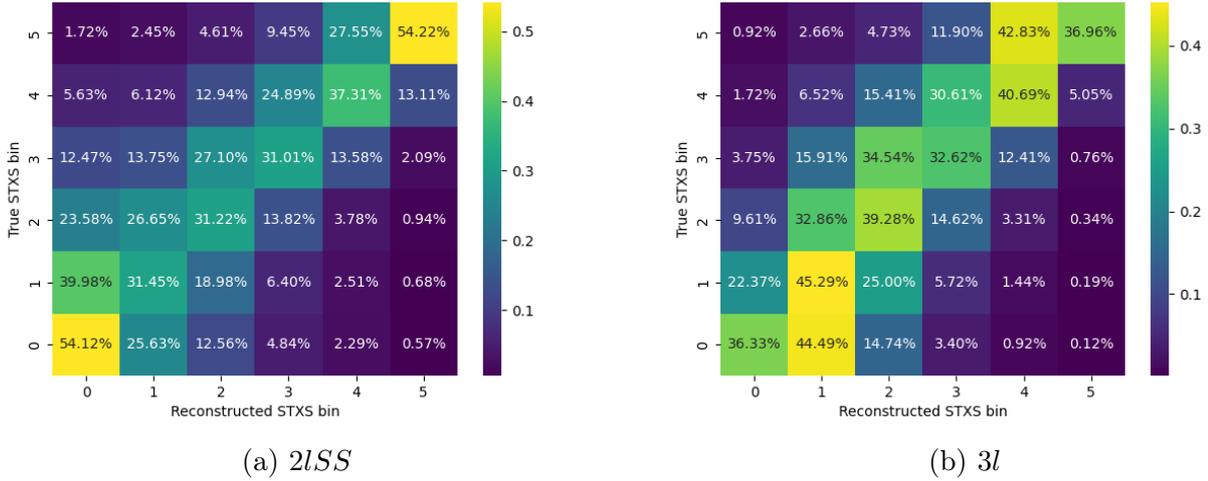


Figure 4.6.15: Confusion matrices of Higgs  $p_T$  STXS bin classification for the  $2lSS$  (a) and  $3l$  (b) channels using  $t\bar{t}H$ -odd events.

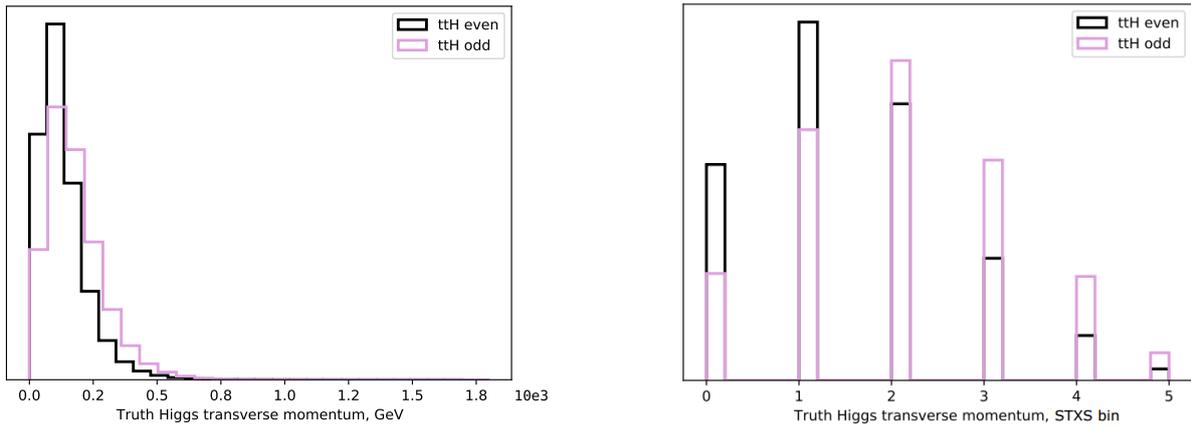


Figure 4.6.16: Distribution of  $t\bar{t}H$ -even and  $t\bar{t}H$ -odd events over the Higgs  $p_T$  (left), and over Higgs  $p_T$  STXS bins (right).

## 4.7 Study of the $\mathcal{CP}$ -properties of the top Yukawa coupling

As discussed in Section 1.5, the potential modification of the Higgs boson coupling to the top quark can be parametrized by introducing in the Lagrangian the parameters  $\kappa'_t$  and  $\alpha$  to the coupling term (see Equation 4.7.1). The value of parameter  $\kappa'_t$  corresponds to the top Yukawa coupling magnitude and is equal to 1 in the SM. The  $\mathcal{CP}$  modification could be parameterised by a complex phase  $\alpha$  with  $\alpha = 0$  corresponding to the SM,  $\alpha = 90^\circ$  to pure  $\mathcal{CP}$ -odd coupling and  $\alpha = 45^\circ$  to the maximum mixing scenario. The notions of  $\kappa_t = \kappa'_t \cos(\alpha)$  and  $\tilde{\kappa}_t = \kappa'_t \sin(\alpha)$  are used through the rest of this chapter.  $m_t$  in Eq. 4.7.1 stands for the top-quark mass,  $v$  for the electroweak symmetry vacuum expectation value,  $\psi_t$  for the top-quark fermionic field, and  $H$  for

the scalar Higgs boson field:

$$\begin{aligned}
\mathcal{L} &= -\frac{m_t}{v} \left\{ \bar{\psi}_t [\kappa_t + i\gamma_5 \tilde{\kappa}_t] \psi_t \right\} H \\
&= -\frac{m_t}{v} \left\{ \bar{\psi}_t \kappa'_t [\cos(\alpha) + i \sin(\alpha) \gamma_5] \psi_t \right\} H \\
&= -\frac{m_t}{v} \left\{ \bar{\psi}_t [\kappa_t + i \tilde{\kappa}_t \gamma_5] \psi_t \right\} H.
\end{aligned} \tag{4.7.1}$$

Figure 1.7.1 illustrates how the  $t\bar{t}H$  cross-section changes with  $\alpha$ . The minimum value occurs at  $\alpha = 90^\circ$ , where it decreases to about one-fifth of the SM prediction. The  $tH$  cross-section changes also with  $\alpha$ , increasing by a factor 5 at  $\alpha = 45^\circ$  and by a factor of 10 at  $\alpha = 45^\circ$ . The cross-section of the  $tWH$  process is also affected by variation of  $\alpha$ . Therefore, when studying the  $\mathcal{CP}$  modification of the top Yukawa coupling, it is important to consider the variations in  $t\bar{t}H$ ,  $tH$ , and  $tWH$  processes together.

The  $\mathcal{CP}$  interpretation follows the same overall analysis strategy that is applied to measure the  $t\bar{t}H$  cross-section, except that in the  $\mathcal{CP}$  fit, the  $tH$  and  $tWH$  processes are also treated as signal. The MC samples used to model the  $t\bar{t}H$  and  $tH$  processes are described in Section 4.2.

### 4.7.1 $t\bar{t}H$ parametrization

The parametrization of the expected  $t\bar{t}H$  yield can be calculated in the following way:

$$y_{t\bar{t}H}(\alpha, \kappa'_t) = \kappa_t^2 \cdot y_{\text{even}} + \tilde{\kappa}_t^2 \cdot y_{\text{odd}}, \tag{4.7.2}$$

where  $y_{\text{even}}$  and  $y_{\text{odd}}$  correspond to the expected yield from the pure  $\mathcal{CP}$ -even and  $\mathcal{CP}$ -odd  $t\bar{t}H$  samples, respectively, in a given bin. There is no interference between the  $\mathcal{CP}$ -even and  $\mathcal{CP}$ -odd components in this parametrization since it is expected to be zero. To validate the parametrization, the MC  $t\bar{t}H$  sample corresponding to  $\alpha = 45^\circ$  was compared with distributions obtained using the parametrization for the maximal mixing scenario. Figure 4.7.1 shows these comparisons for the  $3l$  channel for the Higgs  $p_T$  at truth level and for the MVA scores and for the jet multiplicity at the reconstruction level after pre-selection. The parametrization is found to be a good approximation both in rate and shape.

### 4.7.2 $tH$ parametrization

The  $tH$  parameterization has a more complex form (see Equation 4.7.3) since the interference between amplitudes proportional to the top-Higgs coupling and those proportional to the W-Higgs coupling can not be neglected as it was done in the  $t\bar{t}H$  case:

$$\begin{aligned}
\frac{y(\kappa_t, \tilde{\kappa}_t)}{y_{\text{even}}} &= A\kappa_t^2 + B\tilde{\kappa}_t^2 \\
&+ C\kappa_t + D\tilde{\kappa}_t \\
&+ E\kappa_t\tilde{\kappa}_t \\
&+ F.
\end{aligned} \tag{4.7.3}$$

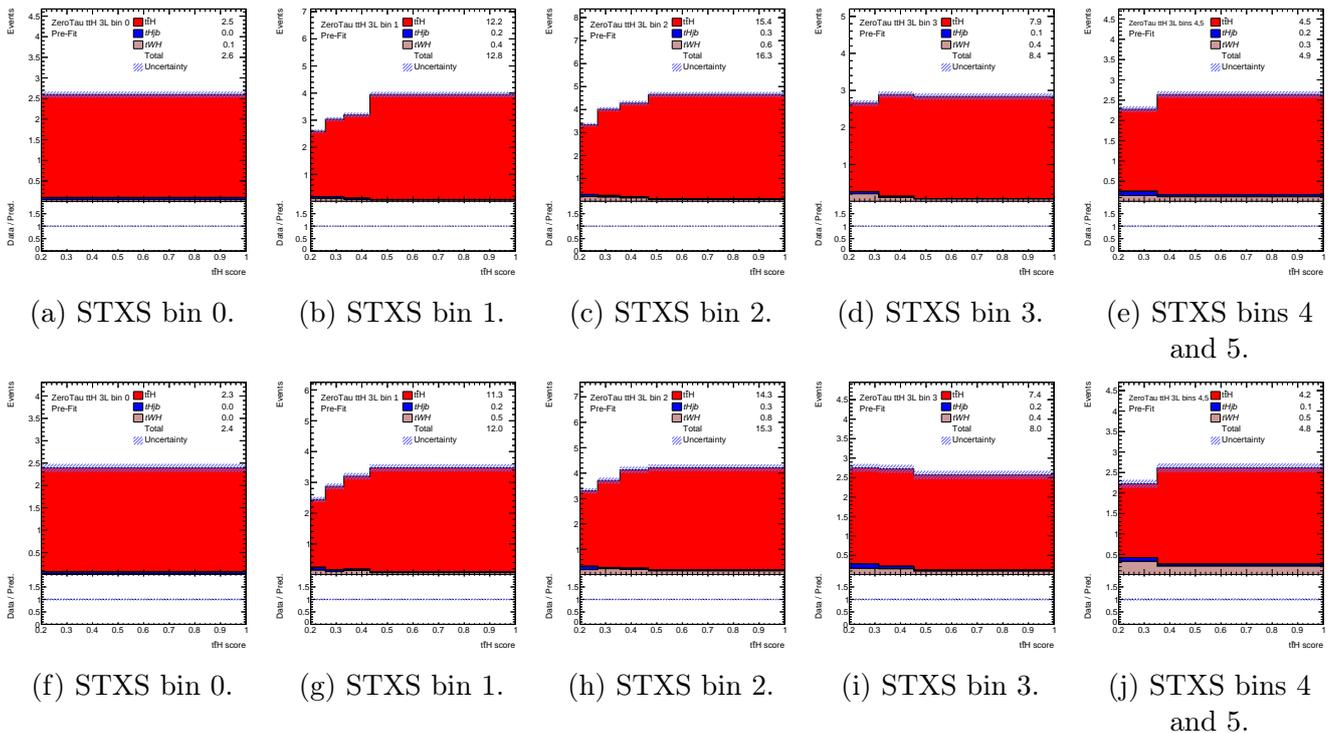


Figure 4.7.1: Comparison between yields predicted by the  $t\bar{t}H$  parametrization for  $\alpha = 45^\circ$  (*top*) and yields in a sample corresponding to  $\alpha = 45^\circ$  (*bottom*) for the STXS SRs in the  $3l$  channel.

This results in the additional terms: while coefficients  $A$  and  $B$  stand for the  $\mathcal{CP}$ -even and  $\mathcal{CP}$ -odd contributions,  $C$  ( $D$ ) corresponds to the interference between the  $\mathcal{CP}$ -even ( $\mathcal{CP}$ -odd)  $tH$  and the  $WH$  coupling,  $E$  to the interference between the  $\mathcal{CP}$ -even and  $\mathcal{CP}$ -odd  $tH$  couplings and  $F$  to the pure  $WH$  coupling contribution. Following theoretical studies described in Ref. [167] for phase space close to ours, the coefficients  $D$  and  $E$  in Equation 4.7.3 have been set to zero. This allows us to always predict positive yields even for negative  $\alpha$ , which was not always the case otherwise due to statistical fluctuations in the fit. The parametrization coefficients are obtained by fitting the polynomial of Equation 4.7.3 to the  $tH$  and  $tWH$  samples generated for different coupling parameters. The  $tH$  and  $tWH$  parameterizations are applied independently for  $tH$  and  $tWH$  in each region's bin yield. For the bins where the fit was shown to be unreliable (i.e.  $\chi^2/\text{ndf} < 7$ ), the parametrization has also been rederived using the neighbouring bin(s). In such cases, for the  $\mathcal{CP}$  measurements, the same parametrization is used for these different bins. To validate the parametrization, the yields from different MC samples are compared with the expected fitted yield value in different regions and bins. The comparison plots for  $tH$  are shown in Figure 4.7.2.

## 4.8 Results

This section presents the results of the  $t\bar{t}H$  differential cross-section measurement and the study of the  $\mathcal{CP}$  properties of the top-Higgs Yukawa coupling within the STXS framework performed in the  $2lSS$  and  $3l$  channels separately and combined. The measurement is performed via a profile-likelihood fit described in Section 4.5 using the analysis regions defined in Section 4.3.

First, the Asimov statistical only  $\mathcal{CP}$  fit is performed separately for the  $2lSS$  and  $3l$  channels

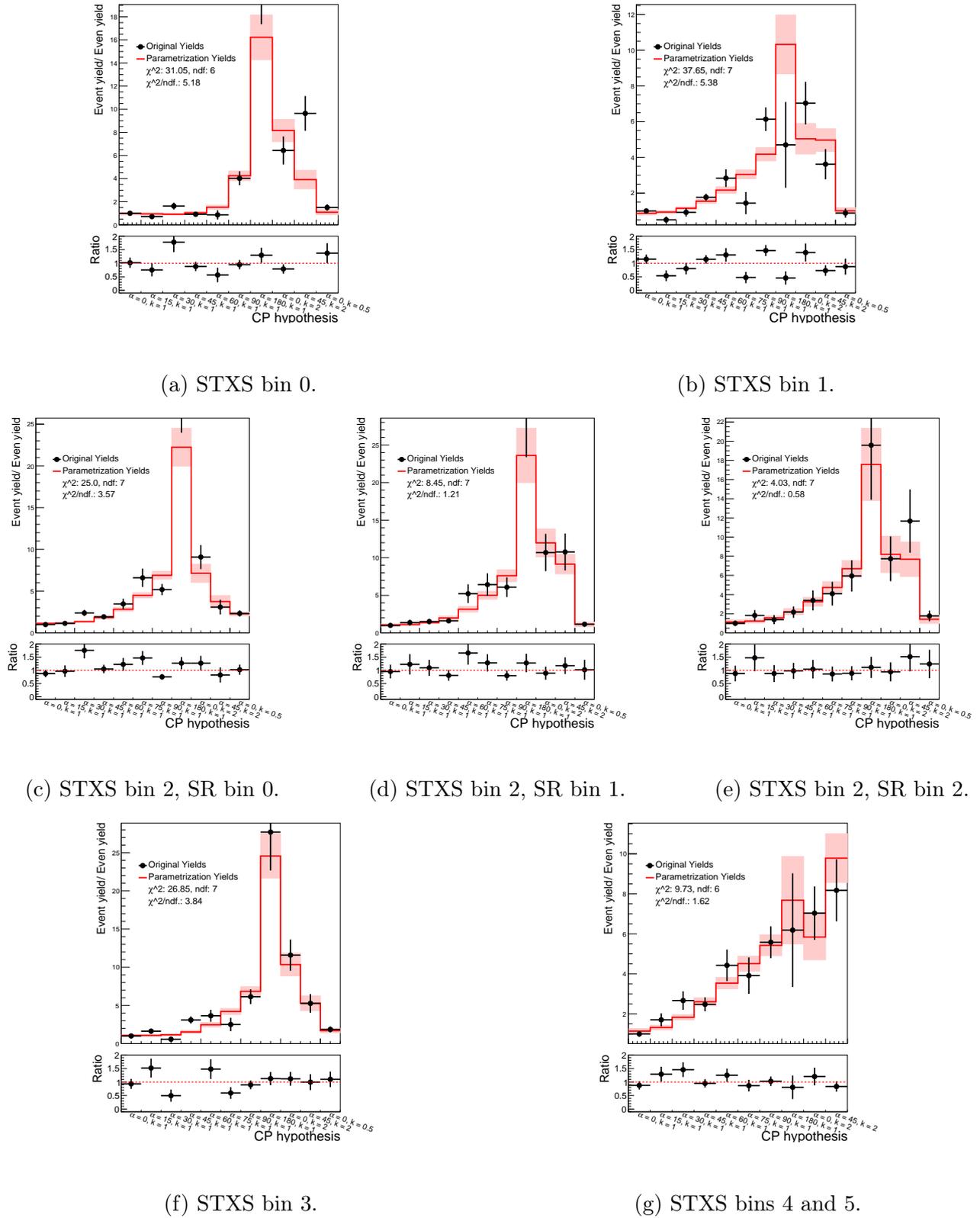


Figure 4.7.2: Yields predicted for different  $(\alpha, \kappa'_t)$  parameters (red histograms) compared to the yields obtained from the MC samples (black dots) for the different bins of the  $3l$  STXS  $tH$  SRs.

and their combination in Section 4.8.1. The expected performance of the final  $\mathcal{CP}$  fit in the combined  $2lSS$  and  $3l$  channels is presented in Section 4.4 including systematic uncertainties. The STXS  $t\bar{t}H$  cross section measurement with systematic uncertainties is presented in Section 4.8.3.

### 4.8.1 Statistical-only Asimov fit results

The first fit is performed using the Asimov setting. This section presents the statistical-only results, omitting the systematic uncertainties. It is useful to compare the sensitivities brought by both the  $2lSS$  and  $3l$  channels.

The fitted  $\mathcal{CP}$ -parameters,  $\kappa'$  and  $\alpha$ , and background normalisation factors are shown in Figure 4.8.1 separately for the  $2lSS$  and  $3l$  channels and their combination. In the Asimov case, the  $\kappa'$  and NFs are measured at 1 and  $\alpha$  at 0, as expected. As expected, the combination of the two channels has smaller errors on the background NFs estimations compared to the separate channel fits. For example, the error on  $t\bar{t}W$  NF being  $\pm 0.08$  and  $\pm 0.15$  in  $2lSS$  and  $3l$  channels correspondingly decreased up to  $\pm 0.06$  in the combination. Overall, the errors look coherent and stable across the channels and their combination. However, the parameters  $\alpha$  and  $\kappa'_t$  are highly correlated (see Figure 4.8.2). This happens due to the flatness of the likelihood near the  $\alpha = 0$ . Such correlation influences the uncertainty in the measurement of the POIs, resulting in a large uncertainty on  $\alpha$ . Thus, the sensitivity comparison on POIs is not possible via the NF plots and should be compared based on the likelihood scans. The likelihood scans over  $\alpha$ , where the  $\alpha$  value is given in units of  $\pi$ , are presented in Figure 4.8.3. Again, the combination of two channels profits from the increased statistics, resulting in higher  $\mathcal{CP}$ -odd exclusion limits and tighter  $1\sigma$  region:  $2.73\sigma$  in comparison with  $2.14\sigma$  in  $2lSS$  and  $1.68\sigma$  in  $3l$ ; and  $[-0.29, 0.29]$  in comparison with  $[-0.36, 0.36]$  in  $2lSS$  and  $[-0.32, 0.32]$  in  $3l$ .

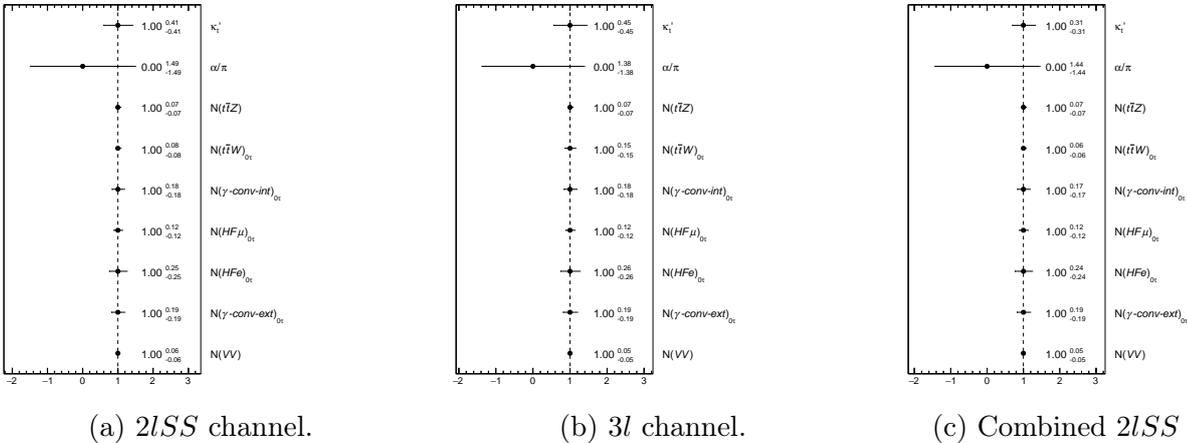


Figure 4.8.1: Fitted values of the POI and main background normalization factors after an Asimov statistical-only STXS fit.

Apart from the correlation with each other, the POIs correlate with the  $t\bar{t}W$  NF (around 50% on  $\kappa'$  and around 40% on  $\alpha$ ). Parameter  $\kappa'$  also has a smaller correlation with  $t\bar{t}Z$  NF. Other POIs correlations with nuisance parameters are insignificant.

### 4.8.2 $\mathcal{CP}$ results with systematic uncertainties

After considering the statistical-only results, the expected results are estimated by performing the fit including both statistical and systematic uncertainties. Uncertainty computation considers the pulls, constraints and correlations between nuisance parameters.

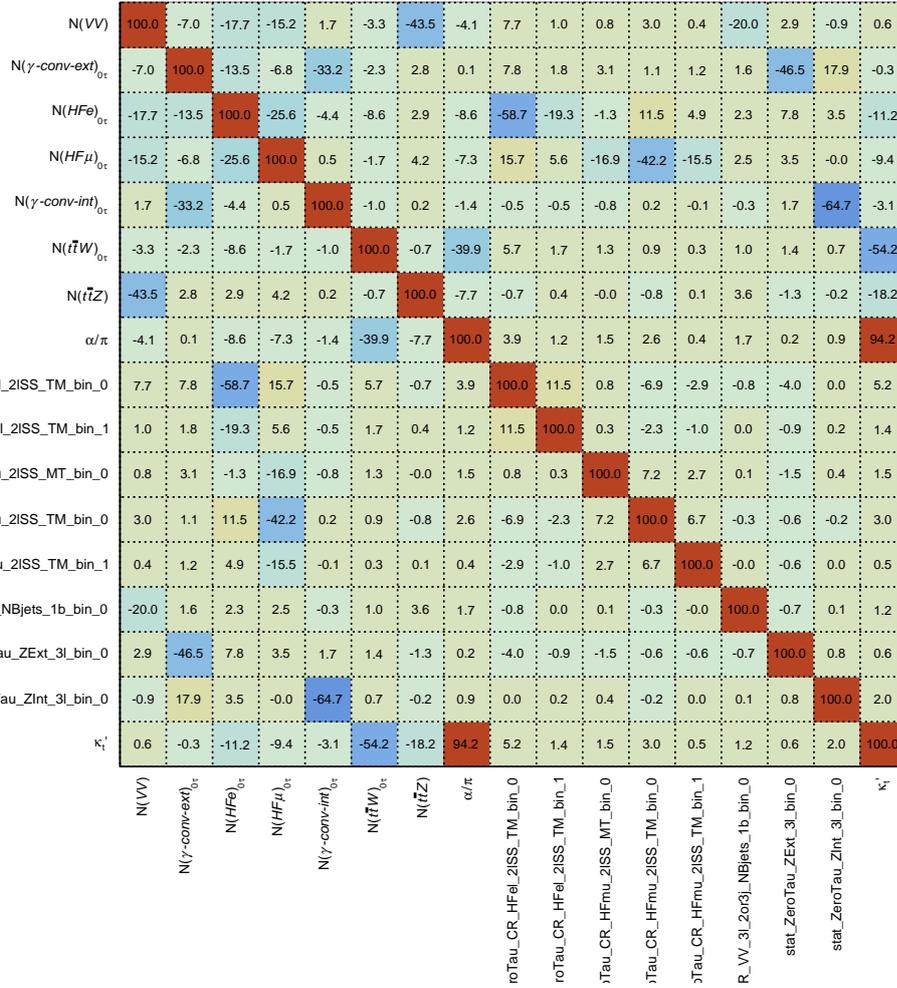


Figure 4.8.2: Correlation matrix showing the correlations between nuisance parameters in an Asimov statistical-only STXS fit.

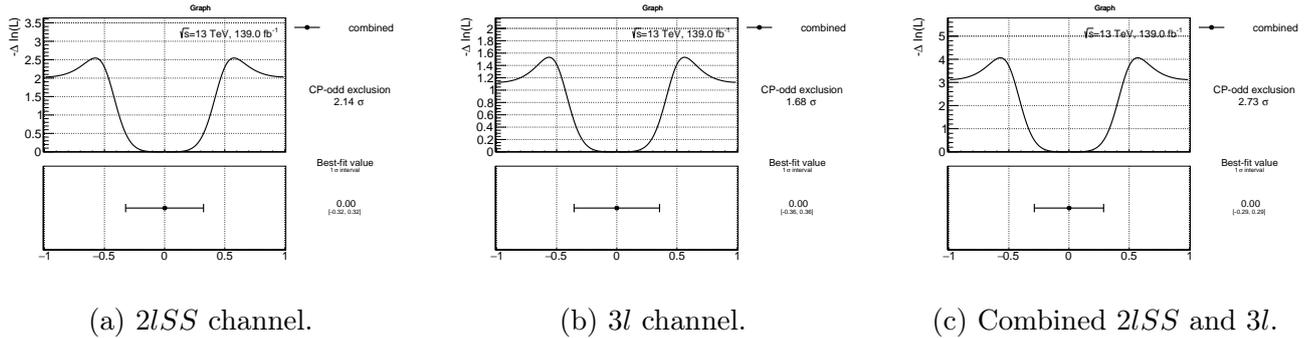
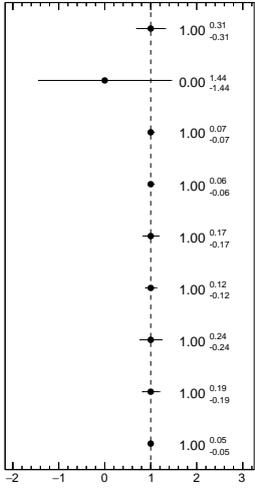


Figure 4.8.3: Likelihood scans over  $\alpha$  in an Asimov statistical-only STXS fit.

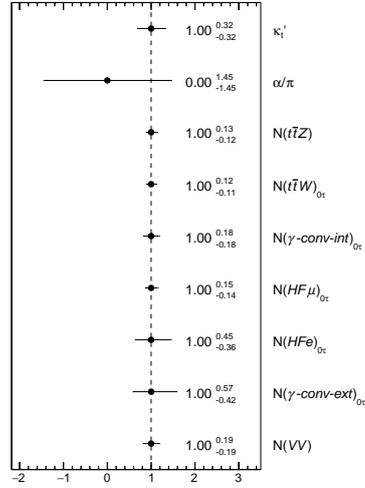
#### 4.8.2.1 Statistical-only and systematic comparison in an Asimov fit

The drop in sensitivity introduced by systematic uncertainties in comparison with the statistics-only results is shown in Figure 4.8.4 and Figure 4.8.5 for an Asimov fit. The errors on the fitted nuisance parameters and POIs are considerably larger when including all systematic uncertainties. Taking into account all the systematic uncertainties causes a decrease in the  $CP$ -odd exclusion limit by around  $0.5\sigma$ , from  $2.73\sigma$  to  $2.26\sigma$ , while the  $1\sigma$  confidence interval for  $\alpha/\pi$  does not

change drastically ( $[-0.29, 0.29]$  compared with  $[-0.31, 0.31]$ ). The  $\kappa_t'$  95% confidence interval is  $[0.86, 1.34]$ .

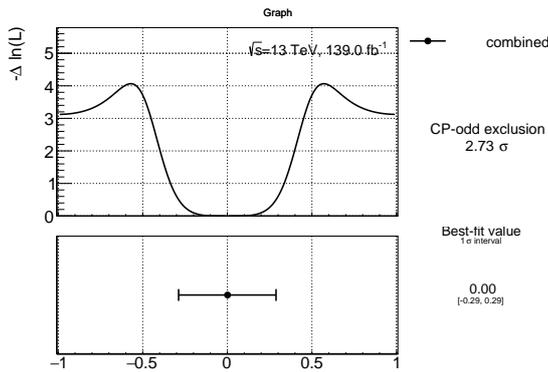


(a) Statistical-only.

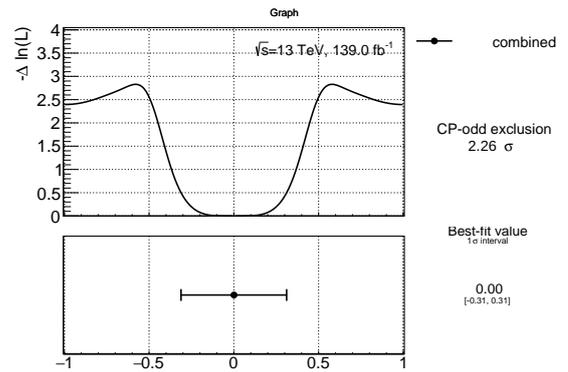


(b) With systematic uncertainties.

Figure 4.8.4: Fitted values of the POI and main background normalization factors in an Asimov STXS  $\mathcal{CP}$  fit for the combined  $2lSS$  and  $3l$  channels.



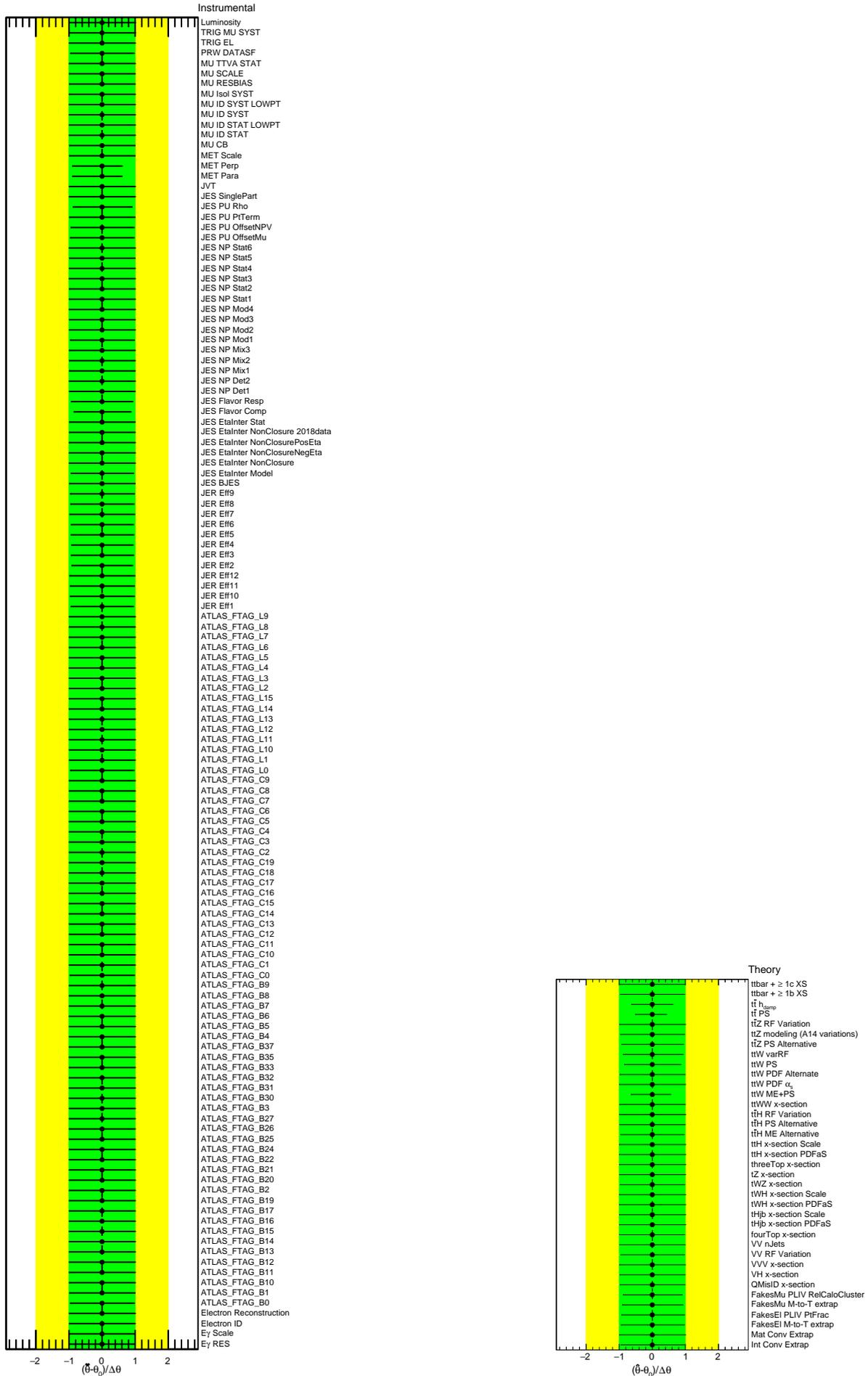
(a) Statistics-only.



(b) With systematic uncertainties.

Figure 4.8.5: Likelihood scans over  $\alpha$  in an Asimov STXS fit for the combined  $2lSS$  and  $3l$  channels.

The pull plot for the systematic uncertainties is shown in Figure 4.8.6. No pulls out of one sigma interval is observed. However, some NPs are constrained: for example the  $t\bar{t}W$  modelling uncertainties (on ME and PS) is constrained, which is coming from the large differences between generators used to evaluate these uncertainties, or the  $t\bar{t}$  PS and  $t\bar{t}_{hdamp}$  where the constraints is coming from large statistic uncertainties HF fake samples in HF CRs.



(a) Instrumental uncertainties.

(b) Theoretical uncertainties.

Figure 4.8.6: Pull plots for the systematics in Asimov STXS setup.

### 4.8.2.2 Hybrid fit results

The results of the hybrid fit with all systematic uncertainties included are presented. Figure 4.8.7 shows the fitted normalization factors, Figure 4.8.8 shows the likelihood scan over  $\alpha$  and the 2D likelihood scan, and Figure 4.8.9 shows the correlation matrix. The exclusion  $\mathcal{CP}$ -odd is measured at  $2.15\sigma$  with the 95% confidence interval on  $\alpha/\pi$  being  $[-0.31, +0.31]$ , and on  $\kappa_t'$  being  $[0.94, 1.47]$  where  $\kappa_t'$  is fitted to 1.14. The correlations between NFs and statistics-related NPs are similar to those observed in the Asimov statistics-only fit. The POIs in a fit with systematics get decorrelated (both in hybrid and Asimov setups). High correlations appear between related NFs and the modelling uncertainties of the corresponding processes: for example, for di-boson and  $t\bar{t}W$  process. Correlation matrices for the Asimov and hybrid setups with systematic uncertainties for the STXS  $\mathcal{CP}$  fits look similar.

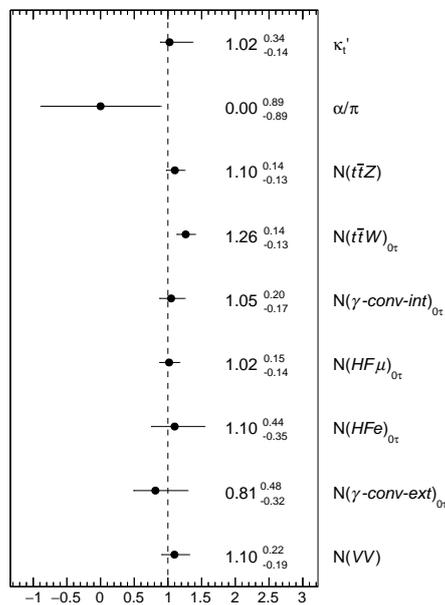


Figure 4.8.7: Fitted values of the POI and main background normalization factors in an hybrid STXS  $\mathcal{CP}$  fit for the combined  $2lSS$  and  $3l$  channels.

The distribution of the post-fit yields over the signal regions is shown in Figure 4.8.10 and over the control regions in Figure 4.8.11.

The statistical uncertainties are not dominant in this measurement. The pulls on the systematic uncertainties is shown in Figure 4.8.12. Even though some systematic uncertainties are pulled, all of them stay within  $1\sigma$  interval.

The impact of the nuisance parameters on the measurement is estimated following the procedure described in Section 4.5.3. The dominant contributions (also in the Asimov fit setup) come from the  $t\bar{t}H$  and  $t\bar{t}W$  modelling related uncertainties (see Figure 4.8.14).

Since the Higgs  $p_T$  is a  $\mathcal{CP}$ -sensitive parameter, the  $\mathcal{CP}$  measurement within the STXS framework is expected to improve the sensitivity on the  $\mathcal{CP}$  parameters. Figure 4.8.15 shows a comparison of the likelihood scans between the inclusive and STXS setups. Introducing the STXS binning indeed improves both the  $\mathcal{CP}$ -odd exclusion limit (from  $1.64\sigma$  without STXS to  $2.15$  with it) and the confidence interval (from  $[-0.35, +0.35]$  to  $[-0.31, +0.31]$ ).

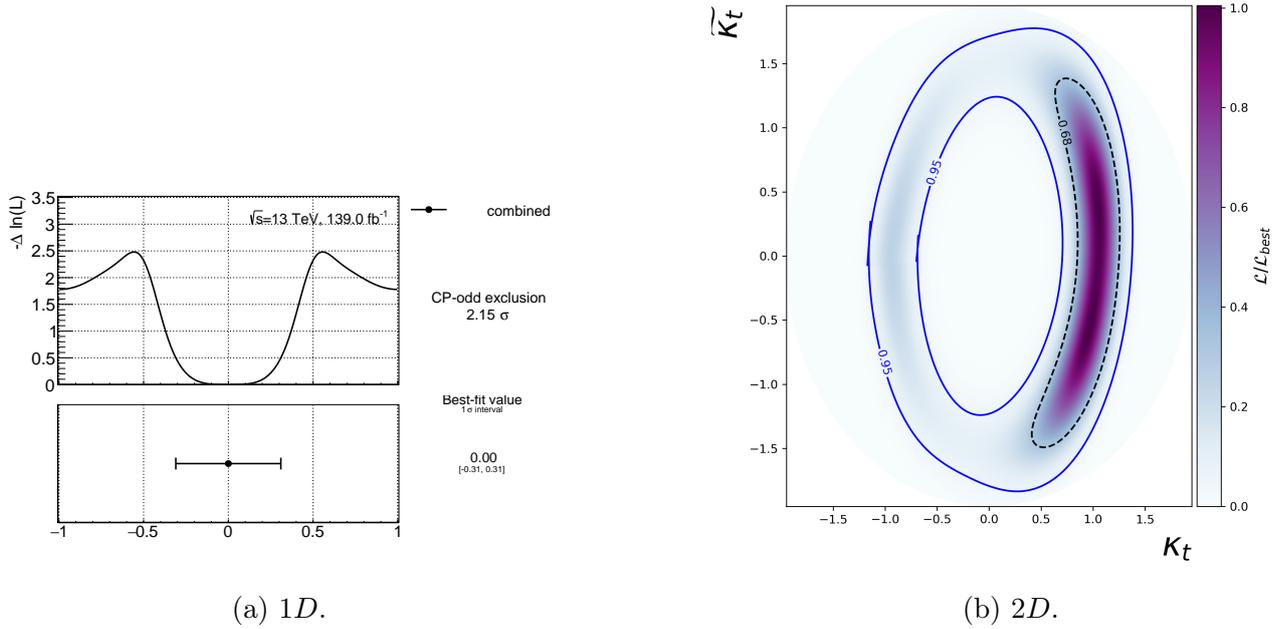


Figure 4.8.8: Likelihood scans over  $\alpha$  (left) and over both  $\alpha$  and  $\kappa'_t$  (right) in a hybrid STXS fit with systematic uncertainties for the combined  $2lSS$  and  $3l$  channels.

### 4.8.3 Differential cross-section measurement within STXS framework

The differential cross-section measurement was performed for the signal strength over 0 and 1 STXS bins together, 2 STXS bin, and over 3, 4 and 5 STXS bins. The result is shown in Figure 4.8.16. The measured signal strengths in the different STXS bins are:  $\mu_{01} = 1.05^{+0.95}_{-0.85}$ ,  $\mu_2 = 0.85^{+1.15}_{-0.85}$  and  $\mu_{345} = 1.31^{+0.69}_{-0.78}$ . Some of the bins were considered together to get a more accurate measurement. The large uncertainties in the measured values are caused by the not perfect reconstruction of the Higgs  $p_T$  bins. The correlations are similar to those presented in the  $\mathcal{CP}$  fit, with additional correlations between the POIs. The pulls (see Figure 4.8.13) also closely follows the results of the  $\mathcal{CP}$  STXS fit.

### 4.8.4 Conclusion

In this work, we performed a study of the  $\mathcal{CP}$  structure of the top Yukawa coupling using the STXS framework. Some limits were set on parameters introducing  $\mathcal{CP}$ -violation in this coupling. The resulting expected 95% confidence intervals are for  $\alpha$ :  $[-0.56^\circ, +0.56^\circ]$  and for  $\kappa'_t$ :  $[0.94, 1.47]$  (with a central value of 1.14). We observed, that using the STXS framework improves the sensitivity. This measurement will be combined with the  $t\bar{t}H$  multilepton with  $2\mathcal{T}_{had}$  in the final state channel. Moreover, the combination with other  $t\bar{t}H$  measurements with comparable sensitivity,  $H \rightarrow \gamma\gamma$  (that excluded  $|\alpha| < 43^\circ$  at 95% confidence level [41]) and  $H \rightarrow bb$  (that measured  $\alpha = 11^{+52^\circ}_{-73^\circ}$  [45]) is also planned.

The differential  $t\bar{t}H$  cross-section expected values in the  $2lSS$  and  $3l$  channels were also measured within the STXS framework. The signal strength was measured in three phase space regions defined by the STXS Higgs  $p_T$  bins:  $\mu_{p_T^H \in [0,120]} \text{ GeV} = 1.05^{+0.95}_{-0.85}$ ,  $\mu_{p_T^H \in [120,200]} \text{ GeV} = 0.85^{+1.15}_{-0.85}$

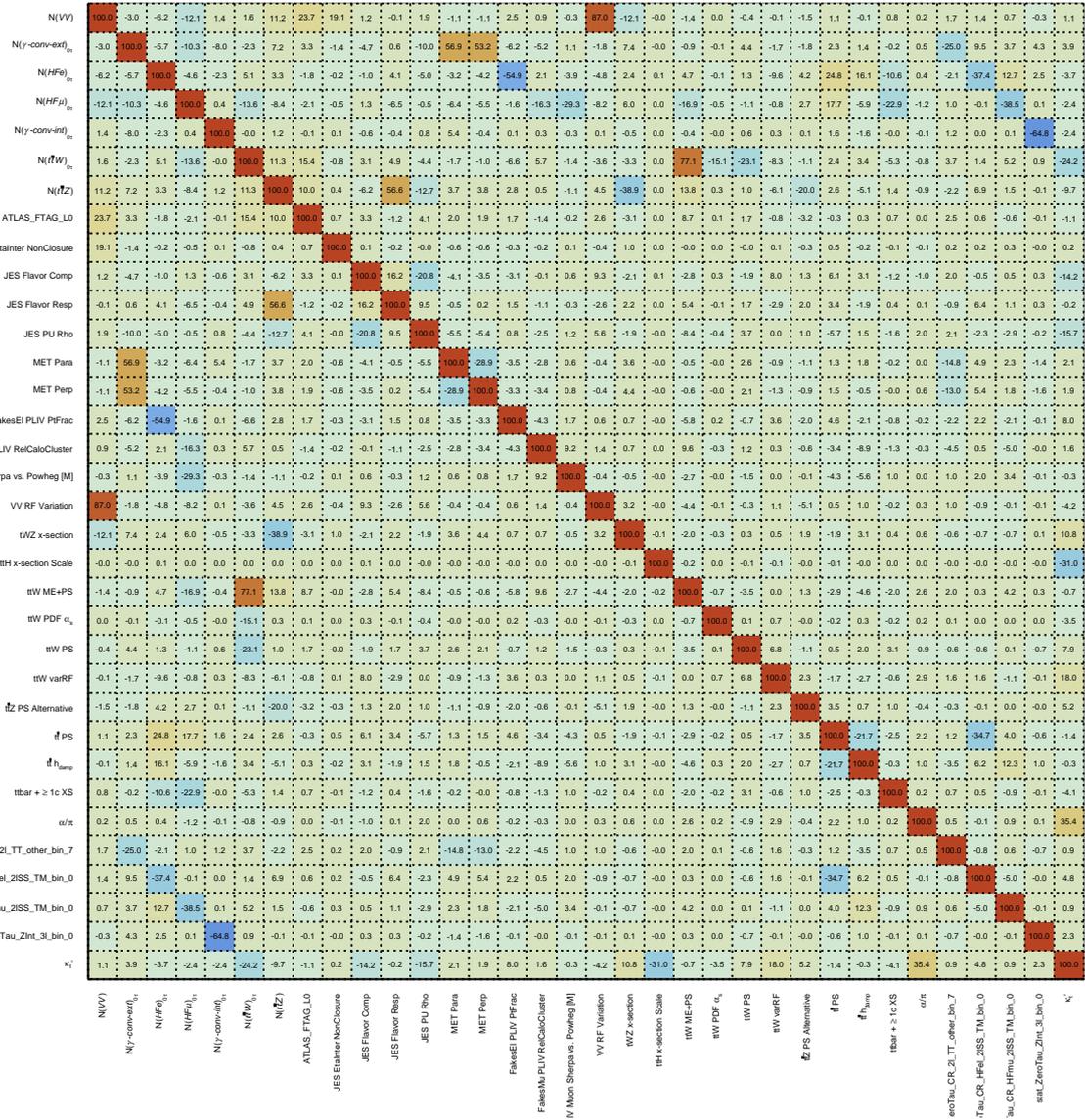
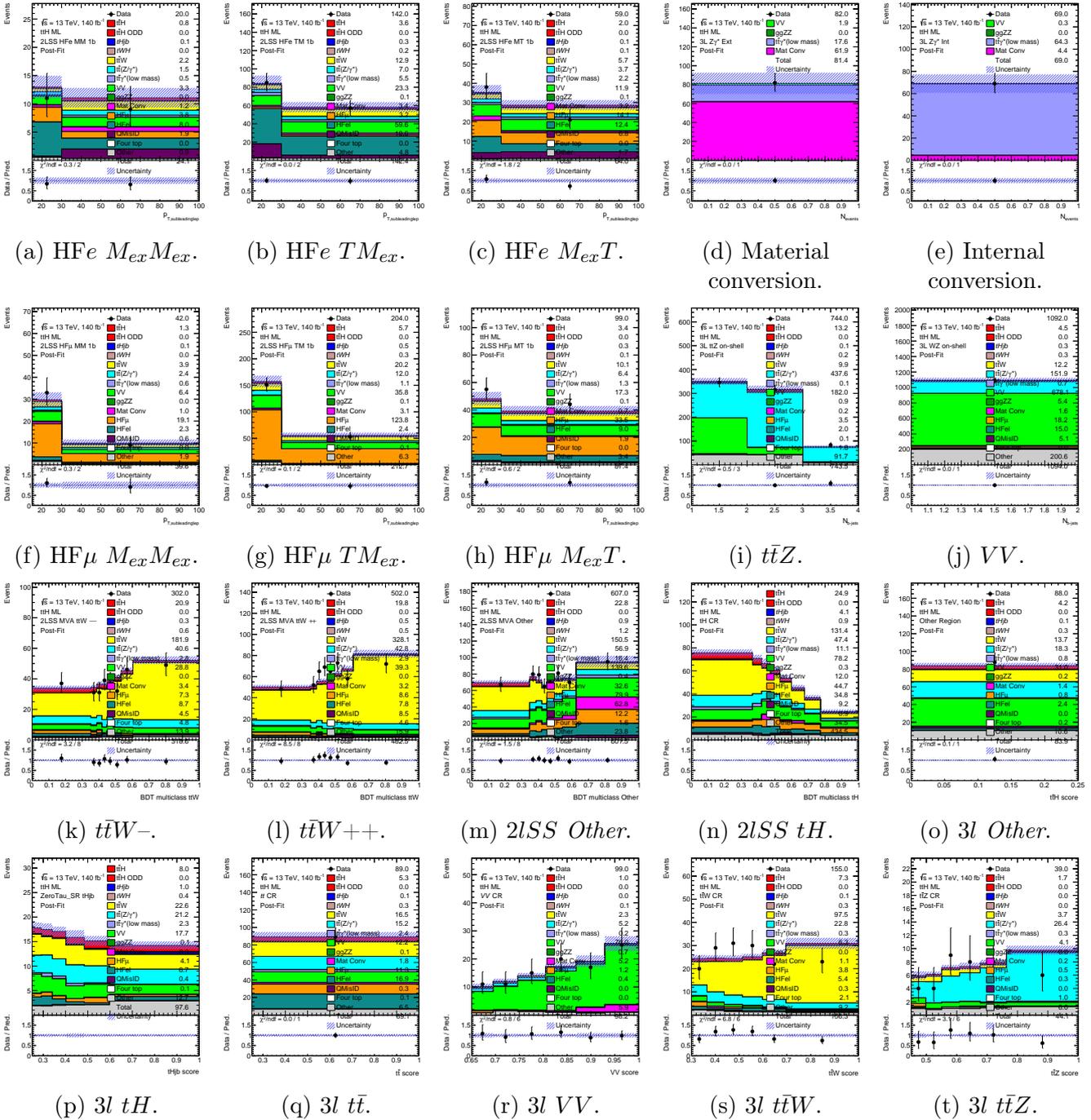


Figure 4.8.9: Correlation matrix showing the correlations between nuisance parameters in an hybrid STXS fit with systematic uncertainties.

and  $\mu_{p_T^H \in [200, +\infty]} \text{ GeV} = 1.31_{-0.78}^{+0.69}$ . For comparison, in the channel with  $H \rightarrow bb$ , the differential  $t\bar{t}H$  cross-section was measured as follows:  $\mu_{p_T^H \in [0,60]} \text{ GeV} = 1.25_{-0.65}^{+0.69}$ ,  $\mu_{p_T^H \in [60,120]} \text{ GeV} = 0.77_{-0.52}^{+0.54}$ ,  $\mu_{p_T^H \in [120,200]} \text{ GeV} = 0.88_{-0.43}^{+0.46}$ ,  $\mu_{p_T^H \in [200,300]} \text{ GeV} = 0.77_{-0.42}^{+0.44}$ ,  $\mu_{p_T^H \in [300,450]} \text{ GeV} = 0.27_{-0.54}^{+0.55}$  and  $\mu_{p_T^H \in [450, +\infty]} \text{ GeV} = 0.63_{-0.83}^{+0.89}$  [168]. The number of bins used in the  $t\bar{t}H$  multilepton channel is lower than in the  $H \rightarrow bb$  channel, the uncertainties on the signal strength are comparable in the two analyses.




 Figure 4.8.11: Post-fit distributions in the CRs after an hybrid  $\mathcal{CP}$  fit with systematic uncertainties.



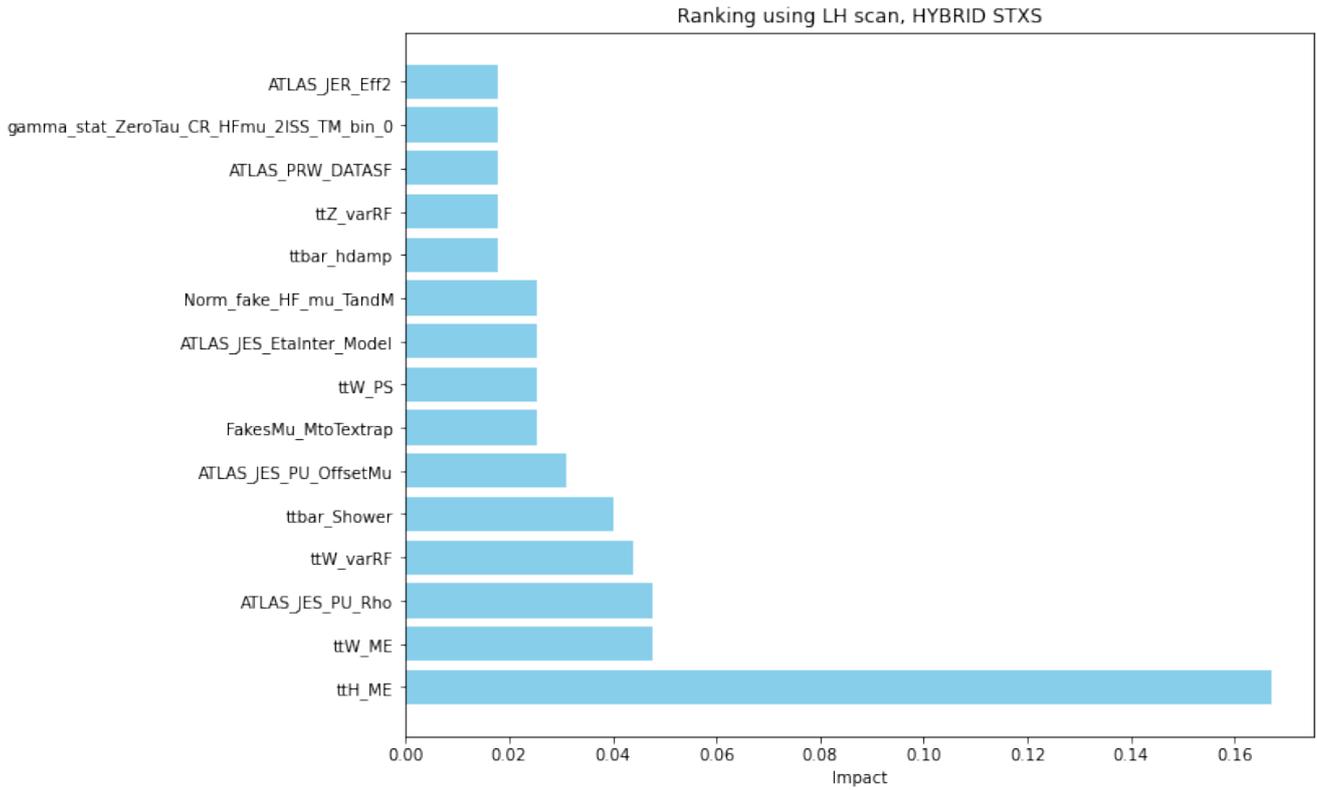


Figure 4.8.14: Ranking of the systematic uncertainties for a hybrid STXS fit for the combined  $2lSS$  and  $3l$  channels.

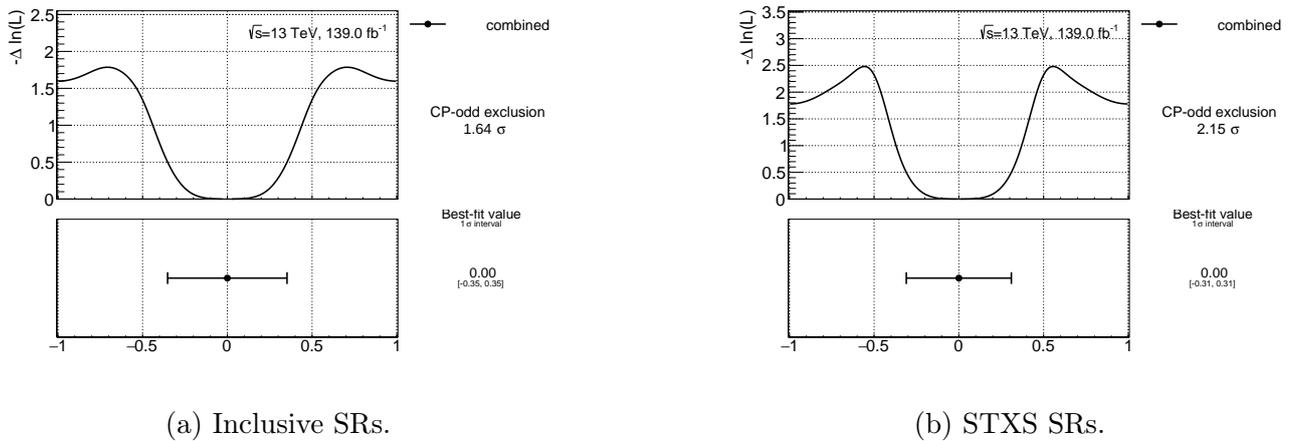


Figure 4.8.15: The likelihood scans over  $\alpha$  in an hybrid fit with systematic uncertainties for the combined  $2lSS$  and  $3l$  channels.

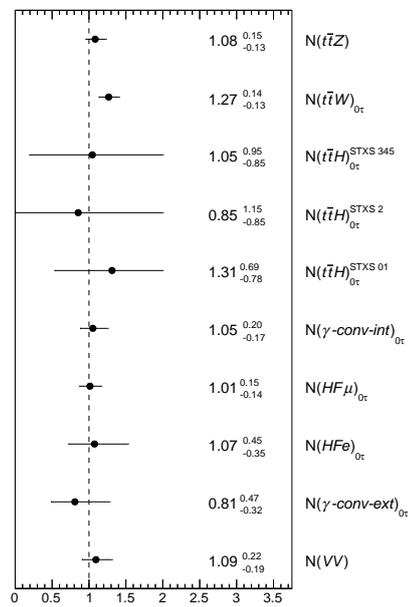


Figure 4.8.16: Fitted values of the POI and main background normalization factors in a hybrid STXS cross-differential section fit for the combined  $2lSS$  and  $3l$  channels.

# Conclusions

This thesis presented two contributions: the development of a tool for automatic visual inspection of ITk Pixel modules for the HL-LHC upgrade and the measurement of the differential cross-section and  $\mathcal{CP}$ -properties of the  $t\bar{t}H$  process with  $140 \text{ fb}^{-1}$  at  $\sqrt{s} = 13 \text{ TeV}$  recorded by the ATLAS detector at the LHC.

Several thousands of ITk pixel modules must be assembled on different production sites during the High-Luminosity LHC upgrade. The assembling procedure includes multiple steps, each followed by quality control checks. This thesis addressed the automatization of the visual inspection before the wire-bonding assembly step. A tool for automatic visual inspection of ITk Pixel modules has been developed. The images of modules are collected with the help of a microscope, and the photo-taking procedure is optimised and described, providing instructions on how to collect photos suitable for further processing. Collected photos are further processed by a machine learning algorithm that yields the defects' location on the module pads. Multiple approaches based on different machine learning algorithms were studied to process the photos. Several unsupervised anomaly detections were compared, and ways to improve their performances were studied. A graphical user interface was developed to make the tool accessible to the technical teams.

The  $t\bar{t}H$  process in the multilepton final state was studied within the STXS framework to scrutinize the top Yukawa coupling. Two channels were considered: final states with two same charged light leptons ( $e$  or  $\mu$ ) and final states with three light leptons, both without hadronically decaying  $\tau$ . The reconstruction of the Higgs boson  $p_T$ , necessary for the STXS framework implementation, was performed based on a Graph Neural Network algorithm. The following STXS bins in Higgs boson  $p_T$  were considered:  $p_0 \in [0, 60] \text{ GeV}$ ,  $p_1 \in [60, 120] \text{ GeV}$ ,  $p_2 \in [120, 200] \text{ GeV}$ ,  $p_3 \in [200, 300] \text{ GeV}$ ,  $p_4 \in [300, 450] \text{ GeV}$ , and  $p_5 \in [450, +\infty] \text{ GeV}$ . The differential cross-section was measured with Asimov events in the analysis signal regions using data in control regions, resulting in the following expected signal strengths:  $\mu_{p_0+p_1} = 1.05_{-0.85}^{+0.95}$ ,  $\mu_2 = 0.85_{-0.85}^{+1.15}$  and  $\mu_{p_3+p_4+p_5} = 1.31_{-0.78}^{+0.69}$ . Some of the bins were considered together to get a more accurate measurement. The sensitivity of the Higgs boson  $p_T$  to  $\mathcal{CP}$  effects was exploited to measure the overall coupling modifier  $\kappa$  and  $\mathcal{CP}$ -mixing angle  $\alpha$  parameters within the STXS framework. The expected 95% confidence intervals found to be:  $\alpha \in [-56^\circ, +56^\circ]$  and  $\kappa'_t \in [0.94, 1.47]$  (with a central value of  $\kappa'_t = 1.14$ ).

# Bibliography

1. Particle Data Group. *Progress of Theoretical and Experimental Physics* **2022**, 083C01. ISSN: 2050-3911. eprint: <https://academic.oup.com/ptep/article-pdf/2022/8/083C01/49175539/ptac097.pdf>. <https://doi.org/10.1093/ptep/ptac097> (Aug. 2022).
2. Noether, E. Invariante Variationsprobleme. ger. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1918**, 235–257. <http://eudml.org/doc/59024> (1918).
3. Gell-Mann, M. Symmetries of Baryons and Mesons. *Phys. Rev.* **125**, 1067–1084. <https://link.aps.org/doi/10.1103/PhysRev.125.1067> (3 Feb. 1962).
4. Yang, C. N. & Mills, R. L. Conservation of Isotopic Spin and Isotopic Gauge Invariance. *Phys. Rev.* **96**, 191–195. <https://link.aps.org/doi/10.1103/PhysRev.96.191> (1 Oct. 1954).
5. Nakano, T. & Nishijima, K. Charge Independence for V-particles. *Progress of Theoretical Physics* **10**, 581–582. <https://api.semanticscholar.org/CorpusID:15391977> (1953).
6. Gell-Mann, M. & Other. The Interpretation of the New Particles as Displaced Charge Multiplets. *Il Nuovo Cimento*. <https://doi.org/10.1007/BF02748000> (Apr. 1956).
7. Ellis, J. Higgs Physics, 117–168. arXiv: 1312.5672. <https://cds.cern.ch/record/1638469> (2015).
8. Goldstone, J., Salam, A. & Weinberg, S. Broken Symmetries. *Phys. Rev.* **127**, 965–970. <https://link.aps.org/doi/10.1103/PhysRev.127.965> (3 Aug. 1962).
9. Yukawa, H. On the Interaction of Elementary Particles I. *Progress of Theoretical Physics Supplement* **1**, 48–57. <https://api.semanticscholar.org/CorpusID:122787304> (1955).
10. Cabibbo, N. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.* **10**, 531–533. <https://link.aps.org/doi/10.1103/PhysRevLett.10.531> (12 June 1963).
11. Kobayashi, M. & Maskawa, T. CP Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics* **49**, 652–657. <https://api.semanticscholar.org/CorpusID:14006603> (1973).
12. Pontecorvo, B. Mesonium and Antimesonium. *Journal of Experimental and Theoretical Physics* **33**, 429. <https://api.semanticscholar.org/CorpusID:222515134> (1957).
13. Maki, Z., Nakagawa, M. & Sakata, S. Remarks on the Unified Model of Elementary Particles. *Progress of Theoretical Physics* **28**, 870–880. ISSN: 0033-068X. eprint: <https://academic.oup.com/ptp/article-pdf/28/5/870/5258750/28-5-870.pdf>. <https://doi.org/10.1143/PTP.28.870> (Nov. 1962).

14. Chapon, E. *Recherche du boson de Higgs et de couplages de jauge quartiques anormaux dans le canal WW en électrons dans l'expérience D<sub>0</sub> au Tevatron* PhD thesis (2013). <https://theses.fr/2013PA066070>.
15. Aad, G. *et al.* Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B* **716**, 1–29. ISSN: 0370-2693. <http://dx.doi.org/10.1016/j.physletb.2012.08.020> (Sept. 2012).
16. Chatrchyan *et al.* Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B* **716**, 30–61. ISSN: 0370-2693. <http://dx.doi.org/10.1016/j.physletb.2012.08.021> (Sept. 2012).
17. Ortona, G. The Higgs boson couplings: past, present, and future. The relationships between Higgs boson and other known particles as measured by current and future experiments. *Frontiers in Physics* **11**. ISSN: 2296-424X. <https://www.frontiersin.org/articles/10.3389/fphy.2023.1230737> (2023).
18. Grojean, C. Higgs Physics, 143–158. arXiv: 1708.00794. <https://cds.cern.ch/record/2243593> (2016).
19. Georgi, H. M., Glashow, S. L., Machacek, M. E. & Nanopoulos, D. V. Higgs Bosons from Two-Gluon Annihilation in Proton-Proton Collisions. *Phys. Rev. Lett.* **40**, 692–694. <https://link.aps.org/doi/10.1103/PhysRevLett.40.692> (11 Mar. 1978).
20. Monini, C. *Single-top s channel cross-section measurement with the ATLAS detector* PhD thesis (Sept. 2014). <https://cds.cern.ch/record/1709700>.
21. Buttazzo, D. *et al.* Investigating the near-criticality of the Higgs boson. *Journal of High Energy Physics* **2013**. ISSN: 1029-8479. [http://dx.doi.org/10.1007/JHEP12\(2013\)089](http://dx.doi.org/10.1007/JHEP12(2013)089) (Dec. 2013).
22. Catani, S. *et al.* Higgs Boson Production in Association with a Top-Antitop Quark Pair in Next-to-Next-to-Leading Order QCD. *Physical Review Letters* **130**. ISSN: 1079-7114. <http://dx.doi.org/10.1103/PhysRevLett.130.111902> (Mar. 2023).
23. CERN. *CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector* en. 2017. <https://e-publishing.cern.ch/index.php/CYRM/issue/view/32>.
24. Sirunyan, A. M. *et al.* Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at  $\sqrt{s} = 13$  TeV. *The European Physical Journal C* **81**. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-021-09014-x> (Apr. 2021).
25. Belyaev, A., Cerrito, L., Lunghi, E., Moretti, S. & Sherrill, N. *Probing CPT invariance with top quarks at the LHC* 2024. arXiv: 2405.12162.
26. Christenson, J. H., Cronin, J. W., Fitch, V. L. & Turlay, R. Evidence for the  $2\pi$  Decay of the  $K_2^0$  Meson. *Phys. Rev. Lett.* **13**, 138–140. <https://link.aps.org/doi/10.1103/PhysRevLett.13.138> (4 July 1964).

27. Olive, K. A. in *Lecture Notes in Physics* 1–37 (Springer Berlin Heidelberg, 1994). ISBN: 9783540490425. [http://dx.doi.org/10.1007/3-540-58689-X\\_40](http://dx.doi.org/10.1007/3-540-58689-X_40).
28. Aghanim, N. e. a. Planck2018 results: VI. Cosmological parameters. *Astronomy, Astrophysics* **641**, A6. ISSN: 1432-0746. <http://dx.doi.org/10.1051/0004-6361/201833910> (Sept. 2020).
29. Thomson, M. Modern Particle Physics. <https://www.hep.phy.cam.ac.uk/~thomson/MPP/ModernParticlePhysics.html> (2013).
30. Branchina, V., Contino, F. & Ferreira, P. M. Electroweak vacuum lifetime in two Higgs doublet models. *Journal of High Energy Physics* **2018**. ISSN: 1029-8479. [http://dx.doi.org/10.1007/JHEP11\(2018\)107](http://dx.doi.org/10.1007/JHEP11(2018)107) (Nov. 2018).
31. Zhang, C. & Willenbrock, S. Effective-field-theory approach to top-quark production and decay. *Physical Review D* **83**. ISSN: 1550-2368. <http://dx.doi.org/10.1103/PhysRevD.83.034006> (Feb. 2011).
32. Brod, J., Haisch, U. & Zupan, J. Constraints on CP-violating Higgs couplings to the third generation. *Journal of High Energy Physics* **2013**. ISSN: 1029-8479. [http://dx.doi.org/10.1007/JHEP11\(2013\)180](http://dx.doi.org/10.1007/JHEP11(2013)180) (Nov. 2013).
33. Aad, G. *et al.* CP Properties of Higgs Boson Interactions with Top Quarks in the ttH and tH Processes Using  $H \rightarrow \gamma\gamma$  with the ATLAS Detector. *Physical Review Letters* **125**. ISSN: 1079-7114. <http://dx.doi.org/10.1103/PhysRevLett.125.061802> (Aug. 2020).
34. Demartin, F., Maltoni, F., Mawatari, K. & Zaro, M. *Higgs production in association with a single top quark at the LHC* 2015. arXiv: [1504.00611](https://arxiv.org/abs/1504.00611) [hep-ph].
35. Brivio, I. & Trott, M. The standard model as an effective field theory. *Physics Reports* **793**, 1–98. ISSN: 0370-1573. <http://dx.doi.org/10.1016/j.physrep.2018.11.002> (Feb. 2019).
36. Contino, R., Ghezzi, M., Grojean, C., Mühlleitner, M. & Spira, M. Effective Lagrangian for a light Higgs-like scalar. *Journal of High Energy Physics* **2013**. ISSN: 1029-8479. [http://dx.doi.org/10.1007/JHEP07\(2013\)035](http://dx.doi.org/10.1007/JHEP07(2013)035) (July 2013).
37. Alloul, A., Fuks, B. & Sanz, V. Phenomenology of the Higgs effective Lagrangian via Feyn-Rules. *Journal of High Energy Physics* **2014**. ISSN: 1029-8479. [http://dx.doi.org/10.1007/JHEP04\(2014\)110](http://dx.doi.org/10.1007/JHEP04(2014)110) (Apr. 2014).
38. Boudjema, F., Guadagnoli, D., Godbole, R. M. & Mohan, K. A. Laboratory-frame observables for probing the top-Higgs boson interaction. *Physical Review D* **92**. ISSN: 1550-2368. <http://dx.doi.org/10.1103/PhysRevD.92.015019> (July 2015).
39. Aaboud, M. *et al.* Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector. *Physics Letters B* **784**, 173–191. ISSN: 0370-2693. <http://dx.doi.org/10.1016/j.physletb.2018.07.035> (Sept. 2018).
40. Sirunyan, A. M. *et al.* Observation of ttH Production. *Physical Review Letters* **120**. ISSN: 1079-7114. <http://dx.doi.org/10.1103/PhysRevLett.120.231801> (June 2018).

41. Aad, G. *et al.* Properties of Higgs boson interactions with top quarks in the  $t\bar{t}H$  and  $tH$  processes using  $H \rightarrow \gamma\gamma$  with the ATLAS Detector. *Physical Review Letters* **125**. ISSN: 1079-7114. <http://dx.doi.org/10.1103/PhysRevLett.125.061802> (Aug. 2020).
42. Sirunyan, A. *et al.* Measurements of  $t\bar{t}H$  production and the CP structure of the Yukawa interaction between the Higgs Boson and Top Quark in the Diphoton Decay Channel. *Physical Review Letters* **125**. ISSN: 1079-7114. <http://dx.doi.org/10.1103/PhysRevLett.125.061801> (Aug. 2020).
43. Aad, G. *et al.* Higgs boson production cross-section measurements and their EFT interpretation in the  $4l$  decay channel at  $\sqrt{s} = 13$  TeV with the ATLAS detector. *The European Physical Journal C* **80**. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-020-8227-9> (Oct. 2020).
44. Sirunyan, A. M. *et al.* Constraints on anomalous Higgs boson couplings to vector bosons and fermions in its production and decay using the four-lepton final state. *Physical Review D* **104**. ISSN: 2470-0029. <http://dx.doi.org/10.1103/PhysRevD.104.052004> (Sept. 2021).
45. Aad, G. *et al.* Probing the CP nature of the top-Higgs Yukawa coupling in  $t\bar{t}H$  and  $tH$  events with  $H \rightarrow b\bar{b}$  decays using the ATLAS detector at the LHC. *Physics Letters B* **849**, 138469. ISSN: 0370-2693. <http://dx.doi.org/10.1016/j.physletb.2024.138469> (Feb. 2024).
46. *Measurement of the  $t\bar{t}H$  and  $tH$  production rates in the  $H \rightarrow b\bar{b}$  decay channel with  $138 \text{ fb}^{-1}$  of proton-proton collision data at  $\sqrt{s} = 13$  TeV* tech. rep. (CERN, Geneva, 2023). <https://cds.cern.ch/record/2868175>.
47. *Analysis of  $t\bar{t}H$  and  $t\bar{t}W$  production in multilepton final states with the ATLAS detector* tech. rep. (CERN, Geneva, 2019). <http://cds.cern.ch/record/2693930>.
48. Tumasyan, A. *et al.* Search for CP violation in  $t\bar{t}H$  and  $tH$  production in multilepton channels in proton-proton collisions at  $\sqrt{s} = 13$  TeV. *Journal of High Energy Physics* **2023**. ISSN: 1029-8479. [http://dx.doi.org/10.1007/JHEP07\(2023\)092](http://dx.doi.org/10.1007/JHEP07(2023)092) (July 2023).
49. Walker, R. P. Synchrotron radiation. <https://cds.cern.ch/record/398429> (1994).
50. Dittmaier, S., Krämer, M., Liao, Y., Spira, M. & Zerwas, P. Higgs radiation off top quarks in  $e^+e^-$  collisions. *Physics Letters B* **441**, 383–388. ISSN: 0370-2693. [http://dx.doi.org/10.1016/S0370-2693\(98\)01192-7](http://dx.doi.org/10.1016/S0370-2693(98)01192-7) (Nov. 1998).
51. *CAS - CERN Accelerator School : Antiprotons for Colliding-beam Facilities: CERN, Geneva, Switzerland 11 - 21 Oct 1983. CAS - CERN Accelerator School : Antiprotons for Colliding-beam Facilities* CERN (CERN, Geneva, 1984). <https://cds.cern.ch/record/157137>.
52. Alemany-Fernandez, R. *et al.* Operation and Configuration of the LHC in Run 1. <https://cds.cern.ch/record/1631030> (2013).
53. Wenninger, J. Operation and Configuration of the LHC in Run 2. <https://cds.cern.ch/record/2668326> (2019).

54. Fartoukh, S. e. a. *LHC Configuration and Operational Scenario for Run 3* tech. rep. (CERN, Geneva, 2021). <http://cds.cern.ch/record/2790409>.
55. Marchiori, G. Highlights on Higgs boson physics from ATLAS. <https://cds.cern.ch/record/2856190> (2023).
56. Pascuzzi, V. *Looking for Beyond the Standard Model Physics in Dijet-Plus-Lepton Final-State Events Collected with the ATLAS Detector* (Toronto U., 2019). <https://cds.cern.ch/record/2705515>.
57. Niemi, A. *et al.* Availability modeling approach for future circular colliders based on the LHC operation experience. *Physical Review Special Topics - Accelerators and Beams* **19**, 121003. <https://cds.cern.ch/record/2281376> (Dec. 2016).
58. Aamodt, K. e. a. The ALICE experiment at the CERN LHC. A Large Ion Collider Experiment. *JINST* **3**, S08002. <http://cds.cern.ch/record/1129812> (2008).
59. Aad, G. e. a. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST* **3**, S08003. <https://cds.cern.ch/record/1129811> (2008).
60. Chatrchyan, S. e. a. The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment. *JINST* **3**, S08004. <https://cds.cern.ch/record/1129810> (2008).
61. Alves, A. A. e. a. The LHCb Detector at the LHC. *JINST* **3**, S08005. <https://cds.cern.ch/record/1129809> (2008).
62. Adriani, O. *et al.* The LHCf detector at the CERN Large Hadron Collider. *JINST* **3**, S08006. <https://cds.cern.ch/record/1129808> (2008).
63. Anelli, G. e. a. The TOTEM Experiment at the CERN Large Hadron Collider. *JINST* **3**, S08007. <https://cds.cern.ch/record/1129807> (2008).
64. Pinfeld, J. L. The MoEDAL experiment at the LHC. Searching beyond the standard model. *EPJ Web Conf.* **126**, 02024. <https://cds.cern.ch/record/2282506> (2016).
65. Caron, J.-L. *Overall view of LHC experiments.. Vue d'ensemble des experiences du LHC.* 1998. <https://cds.cern.ch/record/841555>.
66. Pequenaio, J. *Computer generated image of the whole ATLAS detector* 2008. <https://cds.cern.ch/record/1095924>.
67. Collaboration, T. A. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation* **3**, S08003. <https://dx.doi.org/10.1088/1748-0221/3/08/S08003> (Aug. 2008).
68. *ATLAS magnet system: Technical Design Report, 1* <https://cds.cern.ch/record/338080> (CERN, Geneva, 1997).
69. La Rosa, A. *ATLAS Pixel Detector: Operational Experience and Run-1 to Run-2 Transition* tech. rep. (CERN, Geneva, 2015). arXiv: 1410.6347. <https://cds.cern.ch/record/1956433>.

70. collaboration, T. A. Operation and performance of the ATLAS semiconductor tracker. *Journal of Instrumentation* **9**, P08009–P08009. ISSN: 1748-0221. <http://dx.doi.org/10.1088/1748-0221/9/08/P08009> (Aug. 2014).
71. Abat, E. e. a. The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance. *JINST* **3**, P02013. <http://cds.cern.ch/record/1094549> (2008).
72. Gonski, J. L. *Development of the ATLAS Liquid Argon Calorimeter Readout Electronics and Machine Learning for the HL-LHC* tech. rep. 3 (CERN, Geneva, 2022). <http://cds.cern.ch/record/2816454>.
73. *ATLAS tile calorimeter: Technical Design Report* <https://cds.cern.ch/record/331062> (CERN, Geneva, 1996).
74. *ATLAS magnet system: Technical Design Report, 1* <https://cds.cern.ch/record/338080> (CERN, Geneva, 1997).
75. Aad, G. & Other. Muon reconstruction efficiency and momentum resolution of the ATLAS experiment in proton–proton collisions at  $\sqrt{s} = 7$  TeV in 2010. *The European Physical Journal C* **74**. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-014-3034-9> (Sept. 2014).
76. Pequeno, J. *Computer generated image of the ATLAS Muons subsystem* 2008. <https://cds.cern.ch/record/1095929>.
77. collaboration, T. A. Operation of the ATLAS trigger system in Run 2. *Journal of Instrumentation* **15**, P10004–P10004. ISSN: 1748-0221. <http://dx.doi.org/10.1088/1748-0221/15/10/P10004> (Oct. 2020).
78. Aad, G. e. a. ATLAS data quality operations and performance for 2015-2018 data-taking. *JINST* **15**, P04003. arXiv: [1911.04632](https://arxiv.org/abs/1911.04632). <https://cds.cern.ch/record/2700249> (2020).
79. Salam, G. P. *Elements of QCD for hadron colliders* 2011. arXiv: [1011.5131](https://arxiv.org/abs/1011.5131) [hep-ph].
80. Cartiglia, N. *Measurement of the proton-proton total, elastic, inelastic and diffractive cross sections at 2, 7, 8 and 57 TeV* 2013. arXiv: [1305.6131](https://arxiv.org/abs/1305.6131) [hep-ex].
81. Dobbs, M. & Hansen, J. B. *The HepMC C++ Monte Carlo Event Record for High Energy Physics* tech. rep. (CERN, Geneva, 2000). <https://cds.cern.ch/record/684090>.
82. Frixione, S. & Webber, B. R. Matching NLO QCD computations and parton shower simulations. *Journal of High Energy Physics* **2002**, 029–029. ISSN: 1029-8479. <http://dx.doi.org/10.1088/1126-6708/2002/06/029> (June 2002).
83. Höche, S., Krauss, F., Schönherr, M. & Siegert, F. QCD matrix elements + parton showers. The NLO case. *Journal of High Energy Physics* **2013**. ISSN: 1029-8479. [http://dx.doi.org/10.1007/JHEP04\(2013\)027](http://dx.doi.org/10.1007/JHEP04(2013)027) (Apr. 2013).
84. Alioli, S., Nason, P., Oleari, C. & Re, E. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *Journal of High Energy Physics* **2010**. ISSN: 1029-8479. [http://dx.doi.org/10.1007/JHEP06\(2010\)043](http://dx.doi.org/10.1007/JHEP06(2010)043) (June 2010).

85. Sjöstrand, T., Mrenna, S. & Skands, P. PYTHIA 6.4 physics and manual. *Journal of High Energy Physics* **2006**, 026–026. ISSN: 1029-8479. <http://dx.doi.org/10.1088/1126-6708/2006/05/026> (May 2006).
86. Bähr, M. *et al.* Herwig++ physics and manual. *The European Physical Journal C* **58**, 639–707. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-008-0798-9> (Nov. 2008).
87. Agostinelli, S. *e. a.* GEANT4—a simulation toolkit. GEANT4. A Simulation toolkit. *Nucl. Instrum. Methods Phys. Res., A* **506**, 250–303. <https://cds.cern.ch/record/602040> (2003).
88. Lukas, W. *Fast Simulation for ATLAS: Atfast-II and ISF* tech. rep. (CERN, Geneva, 2012). <https://cds.cern.ch/record/1458503>.
89. Cornelissen, T. *et al.* *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)* tech. rep. (CERN, Geneva, 2007). <https://cds.cern.ch/record/1020106>.
90. Frühwirth, R. Application of Kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262**, 444–450. ISSN: 0168-9002. <https://www.sciencedirect.com/science/article/pii/0168900287908874> (1987).
91. Cornelissen, T. G. *et al.* *Updates of the ATLAS Tracking Event Data Model (Release 13)* tech. rep. (CERN, Geneva, 2007). <https://cds.cern.ch/record/1038095>.
92. Aaboud, M. *et al.* Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at  $\sqrt{s} = 13$  TeV. *The European Physical Journal C* **79**. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-019-7140-6> (Aug. 2019).
93. Aad, G. *e. a.* Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. *Journal of Instrumentation* **14**, P12006–P12006. ISSN: 1748-0221. <http://dx.doi.org/10.1088/1748-0221/14/12/P12006> (Dec. 2019).
94. Kotecha, J. & Djuric, P. *Gaussian sum particle filtering for dynamic state space models in IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221)* **6** (2001), 3465–3468 vol.6. <https://ieeexplore.ieee.org/document/940587>.
95. Aad, G. *e. a.* Muon reconstruction and identification efficiency in ATLAS using the full Run-2 *pp* collision data set at  $\sqrt{s} = 13$  TeV. *The European Physical Journal C* **81**. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-021-09233-2> (July 2021).
96. Aaboud, M. *e. a.* Jet reconstruction and performance using particle flow with the ATLAS Detector. *The European Physical Journal C* **77**. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-017-5031-2> (July 2017).

97. Cacciari, M., Salam, G. P. & Soyez, G. The anti- $k_t$  jet clustering algorithm. *Journal of High Energy Physics* **2008**, 063–063. ISSN: 1029-8479. <http://dx.doi.org/10.1088/1126-6708/2008/04/063> (Apr. 2008).
98. Aad, G. *et al.* Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *The European Physical Journal C* **77**. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-017-5004-5> (July 2017).
99. Aad, G. *e. a.* Jet energy scale and resolution measured in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector. *The European Physical Journal C* **81**. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-021-09402-3> (Aug. 2021).
100. *Tagging and suppression of pileup jets with the ATLAS detector* tech. rep. (CERN, Geneva, 2014). <https://cds.cern.ch/record/1700870>.
101. Aad, G. *e. a.* Performance of pile-up mitigation techniques for jets in  $pp$  collisions at  $\sqrt{s} = 8$  TeV using the ATLAS detector. *The European Physical Journal C* **76**. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-016-4395-z> (Oct. 2016).
102. *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run* tech. rep. (CERN, Geneva, 2017). <https://cds.cern.ch/record/2273281>.
103. *Secondary vertex finding for jet flavour identification with the ATLAS detector* tech. rep. (CERN, Geneva, 2017). <https://cds.cern.ch/record/2270366>.
104. *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC* tech. rep. (CERN, Geneva, 2018). <https://cds.cern.ch/record/2645405>.
105. Aaboud, M. *e. a.* Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at  $\sqrt{s} = 13$  TeV. *The European Physical Journal C* **78**. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-018-6288-9> (Nov. 2018).
106. Brüning, O. & Rossi, L. *The High Luminosity Large Hadron Collider* eprint: <https://www.worldscientific.com/doi/pdf/10.1142/9581>. <https://www.worldscientific.com/doi/abs/10.1142/9581> (WORLD SCIENTIFIC, 2015).
107. *Technical Design Report for the ATLAS Inner Tracker Strip Detector* tech. rep. (CERN, Geneva, 2017). <https://cds.cern.ch/record/2257755>.
108. Nechaeva, S. *The ATLAS ITk Pixel Detector. The biggest challenges from design to construction.* tech. rep. (CERN, Geneva, 2023). <http://cds.cern.ch/record/2883654>.
109. *Technical Design Report for the ATLAS Inner Tracker Pixel Detector* tech. rep. (CERN, Geneva, 2017). <https://cds.cern.ch/record/2285585>.
110. Heggelund, A. L. *Overview of the ATLAS ITk Pixel Detector* tech. rep. 02 (CERN, Geneva, 2023). <http://cds.cern.ch/record/2834491>.
111. *Sensitivity of the ATLAS experiment to long-lived particles with a displaced vertex and  $E_T^{miss}$  signature at the HL-LHC* tech. rep. (CERN, Geneva, 2018). <http://cds.cern.ch/record/2647992>.

112. Bergmann, P., Fauser, M., Sattlegger, D. & Steger, C. MVTEC AD - A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9584–9592. <https://ieeexplore.ieee.org/document/8954181> (2019).
113. The GIMP Development Team. *GIMP* <https://www.gimp.org>.
114. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Sys. Man. Cyber. 9 (1): 62–66*. <https://ieeexplore.ieee.org/document/4310076> (1979).
115. Bradski, G. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*. <https://opencv.org/> (2000).
116. Wilmet, V., Verma, S., Redl, T., Sandaker, H. & Li, Z. *A Comparison of Supervised and Unsupervised Deep Learning Methods for Anomaly Detection in Images* 2021. <https://arxiv.org/abs/2107.09204>.
117. Ronneberger, O., Fischer, P. & Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation* 2015. arXiv: 1505.04597 [cs.CV]. <https://arxiv.org/abs/1505.04597>.
118. Božič, J., Tabernik, D. & Skočaj, D. Mixed supervision for surface-defect detection: From weakly to fully supervised learning. *Computers in Industry* **129**, 103459. ISSN: 0166-3615. <http://dx.doi.org/10.1016/j.compind.2021.103459> (Aug. 2021).
119. Akcay, S. *et al. Anomalib: A Deep Learning Library for Anomaly Detection* 2022. arXiv: 2202.08341 [cs.CV]. <https://github.com/openvinotoolkit/anomalib>.
120. Paszke, A. *et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library* 2019. arXiv: 1912.01703 [cs.LG].
121. Lee, S., Lee, S. & Song, B. C. CFA: Coupled-hypersphere-based Feature Adaptation for Target-Oriented Anomaly Localization. arXiv: 2206.04325 (2022).
122. Roth, K. *et al. Towards Total Recall in Industrial Anomaly Detection* 2021. arXiv: 2106.08265 [cs.CV].
123. Rezende, D. J. & Mohamed, S. *Variational Inference with Normalizing Flows* 2016. arXiv: 1505.05770 [stat.ML]. <https://arxiv.org/abs/1505.05770>.
124. Yu, J. *et al. FastFlow: Unsupervised Anomaly Detection and Localization via 2D Normalizing Flows* 2021. <https://arxiv.org/abs/2111.07677>.
125. Gudovskiy, D., Ishizaka, S. & Kozuka, K. *CFLOW-AD: Real-Time Unsupervised Anomaly Detection with Localization via Conditional Normalizing Flows* 2021. <https://arxiv.org/abs/2107.12571>.
126. Zavrtnik, V., Kristan, M. & Skočaj, D. *DRAEM – A discriminatively trained reconstruction embedding for surface anomaly detection* 2021. <https://arxiv.org/abs/2108.07610>.
127. He, K., Zhang, X., Ren, S. & Sun, J. *Deep Residual Learning for Image Recognition* 2015. <https://arxiv.org/abs/1512.03385>.

128. Deng, J. *et al.* *ImageNet: A large-scale hierarchical image database* in (2009), 248–255. <https://ieeexplore.ieee.org/document/5206848>.
129. Buslaev, A. *et al.* *Albumentations: Fast and Flexible Image Augmentations*. *Information* **11**. ISSN: 2078-2489. <https://www.mdpi.com/2078-2489/11/2/125> (2020).
130. Sobel, I. *An Isotropic 3x3 Image Gradient Operator*. *Presentation at Stanford A.I. Project 1968*. [https://www.researchgate.net/publication/239398674\\_An\\_Isotropic\\_3x3\\_Image\\_Gradient\\_Operator](https://www.researchgate.net/publication/239398674_An_Isotropic_3x3_Image_Gradient_Operator) (Feb. 2014).
131. Starck, J.-L., Murtagh, F. & Bertero, M. *The Starlet Transform in Astronomical Data Processing : Application to Source Detection and Image Deconvolution* in (2011). <https://api.semanticscholar.org/CorpusID:6483692>.
132. Liu, D. C. & Nocedal, J. *On the limited memory BFGS method for large scale optimization*. *Mathematical Programming* **45**, 503–528. <https://api.semanticscholar.org/CorpusID:5681609> (1989).
133. Maćkiewicz, A. & Ratajczak, W. *Principal components analysis (PCA)*. *Computers and Geosciences* **19**, 303–342. ISSN: 0098-3004. <https://www.sciencedirect.com/science/article/pii/009830049390090R> (1993).
134. Santos, J., Tran, T. & Rippel, O. *Optimizing PatchCore for Few/many-shot Anomaly Detection* 2023. arXiv: 2307.10792 [cs.CV]. <https://arxiv.org/abs/2307.10792>.
135. Ramirez, S. *FastAPI* <https://fastapi.tiangolo.com>.
136. *Dart* <https://dart.dev>.
137. *Flutter* <https://flutter.dev>.
138. Agaras, M. N. *et al.* *Measurement of the total and differential cross sections of a top-quark-antiquark pair in association with a W boson in proton-proton collisions at a centre-of-mass energy of 13 TeV with ATLAS detector at the Large Hadron Collider* tech. rep. (CERN, Geneva, 2020). <https://cds.cern.ch/record/2712986>.
139. <https://atlas-iff.docs.cern.ch/rel22recommendedisowps/index.html>.
140. Aaboud, M. *et al.* *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*. *Physical Review D* **97**. ISSN: 2470-0029. <http://dx.doi.org/10.1103/PhysRevD.97.072003> (Apr. 2018).
141. *Analysis of  $t\bar{t}H$  and  $t\bar{t}W$  production in multilepton final states with the ATLAS detector* tech. rep. (CERN, Geneva, 2019). <https://cds.cern.ch/record/2693930>.
142. *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run* tech. rep. (CERN, Geneva, 2017). <https://cds.cern.ch/record/2273281>.
143. Hartanto, H., Jäger, B., Reina, L. & Wackerroth, D. *Higgs boson production in association with top quarks in the POWHEG BOX*. *Physical Review D* **91**. ISSN: 1550-2368. <http://dx.doi.org/10.1103/PhysRevD.91.094003> (May 2015).
144. Ball, R. D. *et al.* *Parton distributions for the LHC run II*. *Journal of High Energy Physics* **2015**. ISSN: 1029-8479. [http://dx.doi.org/10.1007/JHEP04\(2015\)040](http://dx.doi.org/10.1007/JHEP04(2015)040) (Apr. 2015).

145. Ryd, A. *et al.* *EvtGen: A Monte Carlo Generator for B-Physics* in (2005). <https://api.semanticscholar.org/CorpusID:126229755>.
146. CERN. *CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector* en. 2017. <https://e-publishing.cern.ch/index.php/CYRM/issue/view/32>.
147. Degrande, C. *et al.* UFO – The Universal FeynRules Output. *Computer Physics Communications* **183**, 1201–1214. ISSN: 0010-4655. <http://dx.doi.org/10.1016/j.cpc.2012.01.022> (June 2012).
148. Alloul, A., Christensen, N. D., Degrande, C., Duhr, C. & Fuks, B. FeynRules 2.0 — A complete toolbox for tree-level phenomenology. *Computer Physics Communications* **185**, 2250–2300. ISSN: 0010-4655. <http://dx.doi.org/10.1016/j.cpc.2014.04.012> (Aug. 2014).
149. Frederix, R. & Tsinikos, I. On improving NLO merging for  $t\bar{t}W$  production. *Journal of High Energy Physics* **2021**. ISSN: 1029-8479. [http://dx.doi.org/10.1007/JHEP11\(2021\)029](http://dx.doi.org/10.1007/JHEP11(2021)029) (Nov. 2021).
150. Frixione, S., Laenen, E., Motylinski, P. & Webber, B. R. Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations. *Journal of High Energy Physics* **2007**, 081–081. ISSN: 1029-8479. <http://dx.doi.org/10.1088/1126-6708/2007/04/081> (Apr. 2007).
151. CERN. *CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector* en. 2017. <https://e-publishing.cern.ch/index.php/CYRM/issue/view/32>.
152. Kidonakis, N. *Top Quark Production* 2013. arXiv: [1311.0283](https://arxiv.org/abs/1311.0283).
153. Czakon, M. & Mitov, A. NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels. *Journal of High Energy Physics* **2012**. ISSN: 1029-8479. [http://dx.doi.org/10.1007/JHEP12\(2012\)054](http://dx.doi.org/10.1007/JHEP12(2012)054) (Dec. 2012).
154. Korn, S. *MVA-trainer* <https://mva-trainer-docs-site.docs.cern.ch/>.
155. Aad, G. *et al.* Luminosity determination in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC. *The European Physical Journal C* **83**. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-023-11747-w> (Oct. 2023).
156. *Tagging and suppression of pileup jets with the ATLAS detector* tech. rep. (CERN, Geneva, 2014). <https://cds.cern.ch/record/1700870>.
157. Aad, G. *et al.* Jet energy measurement and its systematic uncertainty in proton-proton collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector. *European Physical Journal C: Particles and Fields* **75**, 17. <https://hal.in2p3.fr/in2p3-01003959> (2015).
158. Aad, G. *et al.* ATLAS b-jet identification performance and efficiency measurement with  $t\bar{t}$  events in  $pp$  collisions at  $\sqrt{s} = 13$  TeV. *The European Physical Journal C* **79**. ISSN: 1434-6052. <http://dx.doi.org/10.1140/epjc/s10052-019-7450-8> (Nov. 2019).

159. *Measurement of b-tagging Efficiency of c-jets in  $t\bar{t}$  Events Using a Likelihood Approach with the ATLAS Detector* tech. rep. (CERN, Geneva, 2018). <https://cds.cern.ch/record/2306649>.
160. *Calibration of light-flavour b-jet mistagging rates using ATLAS proton-proton collision data at  $\sqrt{s} = 13$  TeV* tech. rep. (CERN, Geneva, 2018). <https://cds.cern.ch/record/2314418>.
161. Aad, G. *et al.* Measurement of the  $t\bar{t}\bar{t}$  production cross section in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector. *Journal of High Energy Physics* **2021**. ISSN: 1029-8479. [http://dx.doi.org/10.1007/JHEP11\(2021\)118](http://dx.doi.org/10.1007/JHEP11(2021)118) (Nov. 2021).
162. ATLAS collaboration. *TRExFitter* <https://trexfitter-docs.web.cern.ch/trexfitter-docs/>.
163. Barlow, R. & Beeston, C. Fitting using finite Monte Carlo samples. *Computer Physics Communications* **77**, 219–228. ISSN: 0010-4655. <https://www.sciencedirect.com/science/article/pii/001046559390005W> (1993).
164. Wilks, S. S. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics* **9**, 60–62. <https://doi.org/10.1214/aoms/1177732360> (1938).
165. Battaglia, P. W. *et al.* Relational inductive biases, deep learning, and graph networks. arXiv: [1806.01261](https://arxiv.org/abs/1806.01261) [cs.LG] (2018).
166. Falke, P. Higgs pT reconstruction in  $t\bar{t}H$  ML with graph neural networks. [https://indico.cern.ch/event/991736/contributions/4184609/attachments/2171184/3665677/GNN\\_STXS\\_Peter.pdf](https://indico.cern.ch/event/991736/contributions/4184609/attachments/2171184/3665677/GNN_STXS_Peter.pdf) (2021).
167. Kraus, M., Martini, T., Peitzsch, S. & Uwer, P. *Exploring BSM Higgs couplings in single top-quark production* 2019. arXiv: [1908.09100](https://arxiv.org/abs/1908.09100) [hep-ph]. <https://arxiv.org/abs/1908.09100>.
168. Collaboration, A. *Measurement of the associated production of a top-antitop-quark pair and a Higgs boson decaying into a  $b\bar{b}$  pair in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC* 2024. arXiv: [2407.10904](https://arxiv.org/abs/2407.10904) [hep-ex]. <https://arxiv.org/abs/2407.10904>.