Décrypter le couplage top-Higgs avec l'expérience ATLAS au LHC

Neelam Kumari

Aix*Marseille université
Socialement engagée

NNT/NL : 2022AIXM0404/047ED352

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 16 novembre 2022 par

Neelam KUMARI

Unravelling the top-Higgs coupling with the ATLAS experiment at LHC

Discipline
Physique et Sciences de la Matière

Spécialité
Physiques des Particules et Astroparticules

École doctorale
ED 352 Physique et Sciences de la Matière

Laboratoire/Partenaires de recherche
Centre de Physique des Particules de Marseille (CPPM),
Aix-Marseille Université - CNRS/IN2P3

Composition du jury

Carlo SCHIAVI INFN/Université de Genova, Italie	Rapporteur
Julie MALCLES CEA, Paris-Saclay, France	Rapporteure
Caroline COLLARD IPHC, Strasbourg, France	Examinatrice
Aurelio JUSTE ROZAS IFAE/ICREA, Barcelone, Espagne	Examinateur
Arnaud DUPERRIN CPPM, Marseille, France	Directeur de thèse
Thomas STREBLER CPPM, Marseille, France	Co-directeur de thèse
Cristinel DIACONU CPPM, Marseille, France	Président du jury

ACKNOWLEDGEMENTS

The work presented is the result of the support of many people, to whom I am profoundly grateful.

To my supervisors, Arnaud Duperrin and Thomas Strebler, for giving me the opportunity to pursue my PhD under their supervision and work on this thesis topic. Sincere gratitude for their day-to-day supervision, guidance, encouragement, constructive feedback, and kind support throughout my PhD. Special mention should be made of the times during covid, when they provided constant discussions to help me keep up with my work. Thank you for giving me many opportunities to develop as a researcher. A special thanks to Yann Coadou for providing valuable guidance, and feedback throughout my analysis work.

To the reviewers of this document, Carlo Schiavi and Julie Malcles, as well as the other jury members, Cristinel Diaconu, Aurelio Juste Rozas and Caroline Collard, for taking the time to read my thesis work and sharing valuable comments and discussing wider prospects of my work.

To my colleagues and friends for sharing a variety of knowledge and experiences. Thank you to my seniors: Nihal, Nguyen, and Khanh, for their advice and help during my initial year. Thanks to Nemer for his constant help, humour, and fun and to other CPPM members: Resmi, Chenliang, Lauri, Nairit, and Maxime for having shared insightful discussions. Special thanks to Brigitte for her prompt assistance in resolving administrative issues to the best of her ability and making life easier in a foreign land. Thanks to my great friends: Anagha and Kevissen for constantly motivating me and sharing my weekly retreats.

To many members in the ATLAS groups who provided their technical expertise and valuable feedback.

To my loving family: Abhishek, my mom, and dad, for their love, support, and encouragement throughout my life.

Lastly, to a special person, Hiren, for always providing me with constant support and strength.

RÉSUMÉ

La découverte du boson de Higgs en 2012 est un succès majeur du Large Hadron Collider (LHC) au CERN. Grâce au jeu de données important collecté au LHC depuis, des mesures de précision des propriétés du boson de Higgs (H) sont possibles. La production du boson H en association avec une paire de quarks top, suivie d'une désintégration du boson H en paires de quarks b, ttHbb, permet une mesure directe du couplage de Yukawa du top. Le processus ttHbb présente un état final complexe avec au moins quatre jets de b et requiert une méthode d'identification des jets de b sophistiquée, appelée b-tagging. Le trajectographe d'ATLAS sera mis à niveau pour maintenir ses performances dans les conditions du LHC à haute luminosité (HL-LHC). Cette thèse présente les performances des algorithmes de b-tagging, en se focalisant sur les algorithmes basés sur les paramètres d'impact ou les vertex secondaires, avec des simulations du Inner Tracker d'ATLAS pour le HL-LHC. L'analyse $t\bar{t}H(H \to b\bar{b})$ est effectuée sur la base de 139 fb^{-1} de données collectées par ATLAS au Run 2 à $\sqrt{s}$ = 13 TeV et tire parti des plus récents algorithmes de reconstruction et d'identification. La grande multiplicité du nombre de jets de b due aux produits de désintégration supplémentaires du quark top rend nécessaires des stratégies d'analyse dédiées basées sur le machine learning. Des Deep Neural Networks (DNNs) ont été développés pour améliorer la sensibilité de l'analyse en contraignant les sous-composantes du bruit de fond $t\bar{t}$+jets dominant. Les performances du DNN et des Boosted Decision Trees (BDTs) utilisés précédemment sont présentées dans cette thèse. L'intensité du signal est mesurée de manière inclusive et différentielle par rapport à l'impulsion transverse du boson de H en utilisant l'approche Simplified Template Cross-Section (STXS). La sensibilité au signal $t\bar{t}H$ par rapport au bruit de fond attendu en utilisant les algorithmes DNN est de 2.71σ, comparée à 2.54σ.

Mots-clés: LHC, ATLAS, Higgs boson, MVA, $t\bar{t}H$, $t\bar{t}$ + jets, b-tagging, LHC à haute luminosité.

ABSTRACT

The discovery of the Higgs boson in 2012 was a major success of the Large Hadron Collider (LHC) at CERN. With larger data-sets collected at the LHC since then, precise measurements of the Higgs boson properties are possible. The Higgs boson production in association with a pair of top quarks, where Higgs decays into a pair of *b*-quarks, $t\bar{t}H(H \to b\bar{b})$, allows direct measurement of the top-Yukawa coupling. The $t\bar{t}H(H \to b\bar{b})$ process has a challenging final state with at least four *b*-jets, which requires a *b*-jet identification method known as *b*-tagging. In view of the operation of the ATLAS detector under High-Luminosity LHC conditions, the central tracking system will be upgraded to maintain high levels of performance. This thesis presents the performance of the b-tagging algorithms focusing on impact parameter-based and secondary vertex-based taggers using the updated ATLAS Inner Tracker (ITk) simulation. Some re-optimization of the *b*-taggers is performed, and the impact on performance is also presented. The $t\bar{t}H(H \to b\bar{b})$ analysis is performed using 139 fb^{-1} of ATLAS Run 2 data at $\sqrt{s}$ = 13 TeV and takes advantage of the most recent object performance algorithms. The high *b*-jet multiplicity due to additional top quark decay products requires dedicated analysis strategies based on machine learning. The multivariate analysis approaches using Deep Neural Networks (DNNs) were developed to improve the search sensitivity while constraining large $t\bar{t} + b\bar{b}$ background sub-components. This thesis compares DNN performance to that of previously used Boosted Decision Trees (BDTs). The signal strength is measured inclusively and differentially with respect to the Higgs boson transverse momentum using the Simplified Template Cross-Section formalism. The expected significance of the $t\bar{t}H$ signal over the expected SM background using DNN is 2.71 σ, compared to 2.54 σ using BDTs.

Keywords: LHC, ATLAS, Higgs boson, MVA, $t\bar{t}H$, $t\bar{t}$ + jets, *b*-tagging, High Luminosity LHC.

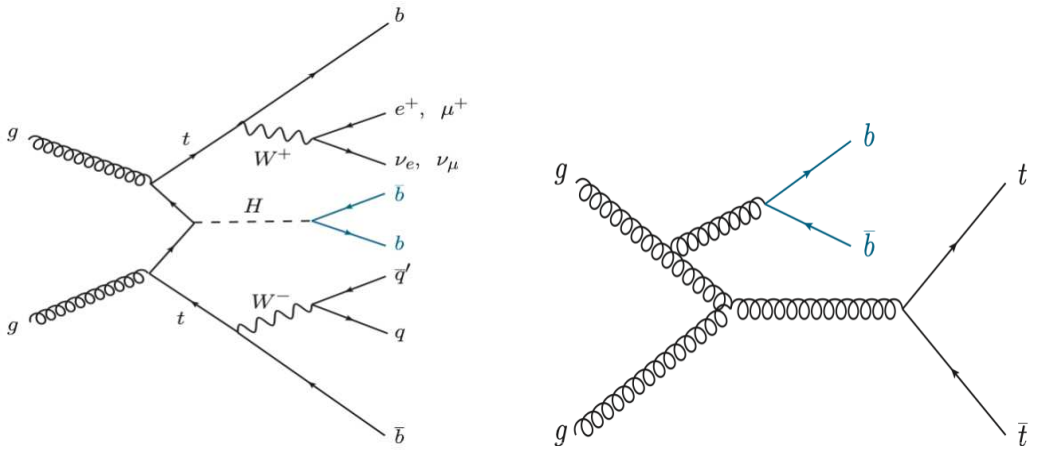
TABLE OF CONTENTS

FRENCH SUMMARY

Quels sont les constituants fondamentaux de l'Univers ? Quelles sont les règles qui les gouvernent ? De telles questions ont longtemps laissé l'Homme perplexe. Le Modèle Standard (MS) de la physique des particules est le modèle théorique correspondant à notre compréhension actuelle des particules élémentaires et décrivant les forces fondamentales qui régissent leurs interactions. Depuis des décennies, les prédictions du MS ont été testées et validées de plus en plus précisément. Le boson de Higgs, dernière pièce manquante du MS, a été découvert en 2012 par les collaborations ATLAS et CMS au *Large Hadron Collider* (LHC). Cette découverte a été une avancée majeure en physique des particules, puisqu'elle a étendu notre compréhension des mécanismes par lesquels les particules élémentaires, acquièrent leur masse, au travers de leur interaction avec le boson de Higgs. Grâce aux nombreuses collisions proton-proton enregistrées au LHC, les propriétés du boson de Higgs peuvent aujourd'hui être mesurées très précisément. Une de ces propriétés est la façon dont le boson de Higgs interagit avec les fermions. L'intensité de cette interaction est proportionnelle à la masse de la particule et le quark top, la plus massive d'entre elles, interagit donc fortement avec le boson de Higgs. La mesure précise de cette interaction est une sonde puissante pour mettre en lumière des phénomènes de nouvelle physique, tels que de nouvelles sources de violation charge-parité qui peuvent exister au-delà du MS. Le Chapitre 1 introduit le MS, en se focalisant notamment sur le mécanisme de Higgs et les recherches du boson de Higgs au LHC.

Le LHC est l'accélérateur de particules circulaire le plus complexe jamais construit. Le détecteur ATLAS, le plus grand détecteur de particules généraliste au LHC, est conçu pour tester le MS de la physique des particules tout en permettant l'exploration de la physique au-delà du MS. Le Chapitre 2 décrit le LHC et l'expérience ATLAS. La production du boson de Higgs en association avec une paire de quarks top, suivie d'une désintégration du boson de Higgs en paire que quarks b, $t\bar{t}H(H \to b\bar{b})$, permet une mesure directe du couplage de Yukawa du quark top (Figure 1 a). Le processus $t\bar{t}H(H \to b\bar{b})$ présente un état final complexe avec au moins quatre jets de b, issus de l'hadronisation de quarks b. Une méthode d'identification complexe, appelée b-tagging, est employée au LHC pour identifier les jets de b. Le principal défi de l'analyse $t\bar{t}H(H \to b\bar{b})$ est la production importante du bruit de fond irréductible $t\bar{t}$ accompagnée des jets supplémentaires, notamment lorsque ceux-ci sont des jets de b (Figure 1 b). Le b-tagging est non seulement crucial pour le processus $t\bar{t}H(H \to b\bar{b})$, mais aussi pour un grand nombre d'analyses de physique au LHC. Celui-ci sera également essentiel au HL-LHC, pour étudier l'auto-couplage du boson de Higgs et

maximiser la sensibilité des mesures de précision et des recherches de nouvelle physique. Durant ma thèse, j'ai principalement contribué à des développements liés au b-tagging, notamment pour l'upgrade de Phase 2 du détecteur ATLAS pour la phase de Haute Luminosité du LHC (HL-LHC), ainsi qu'à l'analyse $t\bar{t}H(H \to b\bar{b})$ basée sur l'ensemble des données collectées par ATLAS durant le Run 2 du LHC. Mes contributions à chacun de ces projets, détaillées dans les Chapitres 3, 4 et 5, sont résumées ci-dessous.



(a) (b)

Figure 1 – Diagrammes de Feynman (a) du processus $t\bar{t}H(H \to b\bar{b})$ dans le canal à un lepton, comprenant au moins six jets dans l'état final, dont au moins quatre jets de b et (b) du bruit de fond principal $t\bar{t}$ + $b\bar{b}$.

Performances de b-tagging avec le détecteur ATLAS de Phase 2

Le HL-LHC est un upgrade du LHC qui vise à augmenter la luminosité instantanée jusqu'à 7×10^{34} cm^{-2}s^{-1}, résultant en une luminosité intégrée de 4ab^{-1} d'ici la fin de la prise de données. Cela aura également pour conséquence d'augmenter le nombre de collisions proton-proton par croisement de faisceau (*pile-up*) jusqu'à $\langle\mu\rangle = 200$. Pour maintenir des performances équivalentes de reconstruction des traces au HL-LHC, l'*Inner Detector* (ID) d'ATLAS sera remplacé par l'*Inner Tracker* (ITk). L'ITk est un trajectographe entièrement en silicium dont la couverture s'étend jusqu'à $|\eta| = 4$ (au lieu de 2.5 pour l'ID). Il est constitué de deux sous-systèmes: le détecteur à pixels interne et le détecteur à pistes externe. Les algorithmes de b-tagging développés dans ATLAS se basent sur l'information des traces reconstruites à partir de l'ID ou de l'ITk et leurs performances représentent une référence importante en termes de reconstruction et d'identification des objets. La Figure 2 présente un schéma de la géométrie de l'ITk, qui a été utilisée pour les études de b-tagging présentées dans cette thèse. Le Chapitre 3 détaille mon travail concernant les algorithmes de b-tagging et leur optimisation avec l'ITk.

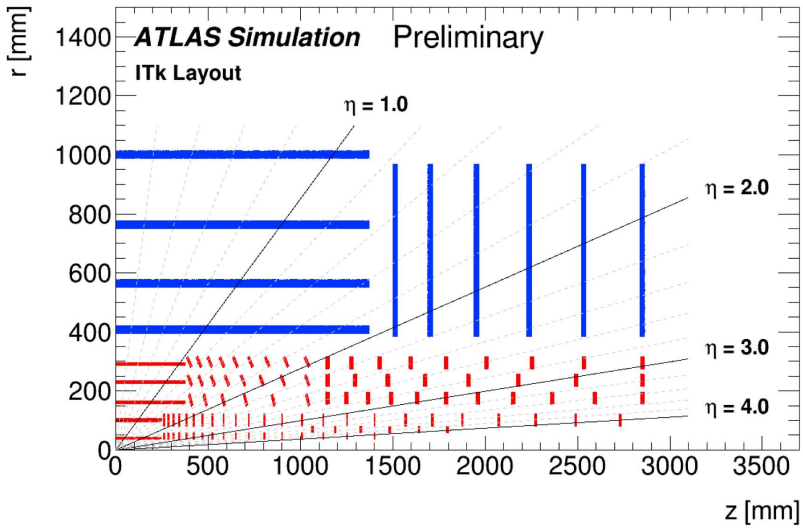
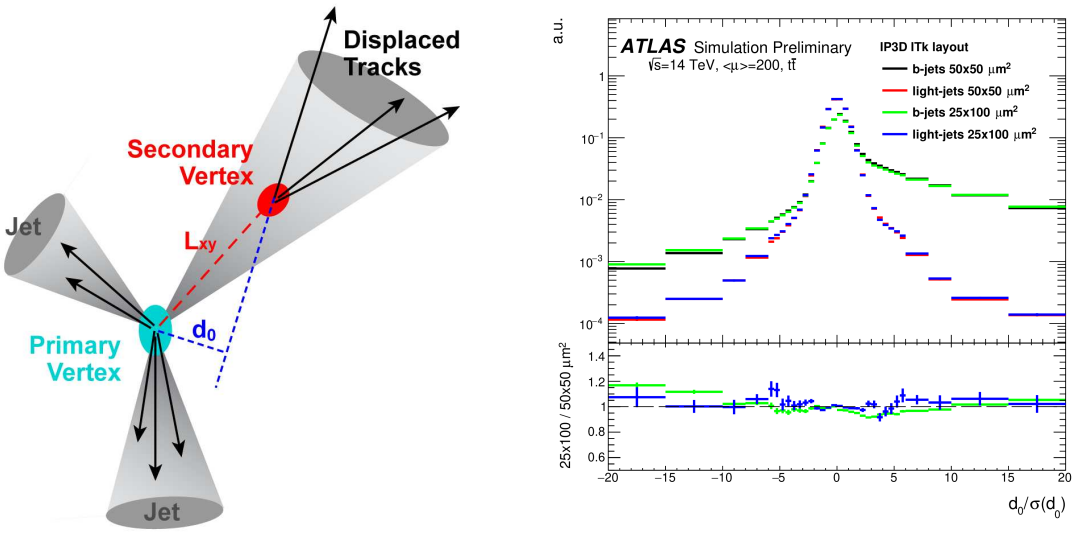


Figure 2 – Schéma de la géométrie de l'ITk considérée pour la majorité des études présentées dans cette thèse. [1].

Les caractéristiques uniques des hadrons b sont à la base de l'identification des jets de b. Les hadrons b ont une longue durée de vie ($\tau \approx 1.5$ ps). Cela offre aux algorithmes de b-tagging deux signatures exploitables. D'une part, la désintégration d'un hadron b crée un vertex secondaire déplacé par rapport au vertex primaire, où la collision de protons s'est produite. D'autre part, les hadrons b ont une masse élevée ($\sim$5 GeV) par rapport aux hadrons c ou légers. Les particules chargées produites dans la désintégration du hadron b au niveau du vertex déplacé seront alors reconstruites comme des traces avec de grands paramètres d'impact, correspondant à la distance d'approche minimale de la trace par rapport au vertex primaire. La Figure 3 (a) représente les propriétés susmentionnées. Les algorithmes de b-tagging utilisés dans les analyses de physique d'ATLAS pour distinguer les jets de b des jets de c et légers (associés aux gluons ou aux quarks légers), sont construits en deux étapes impliquant des algorithmes de bas et de haut niveau. Les algorithmes de bas niveau, basés sur les paramètres d'impact (IP3D) ou les vertex secondaires (SV1), sont directement liés aux propriétés des hadrons b. IP3D est un algorithme utilisant un formalisme de *likelihood* et exploitant une catégorisation des traces basée sur des informations liées aux *hits* du détecteur à pixels. Des distributions de référence, correspondant aux significances des paramètres d'impact d_0 et $z_0 \sin\theta$, sont construites à partir de cette catégorisation, pour les traces de jets de b, de c et légers. La Figure 3 (b) présente la distribution de la significance de d_0 pour des traces de jets de b et de jets légers. SV1 est un autre algorithme basé sur une *likelihood* exploitant la reconstruction des vertex secondaires et leurs propriétés. Puisque les performances de ces algorithmes de bas niveau sont décorrélées, ils peuvent être combinés et utilisés comme variables d'entrée pour les algorithmes de haut niveau, tels que MV2, pour améliorer les performances globales de b-tagging.



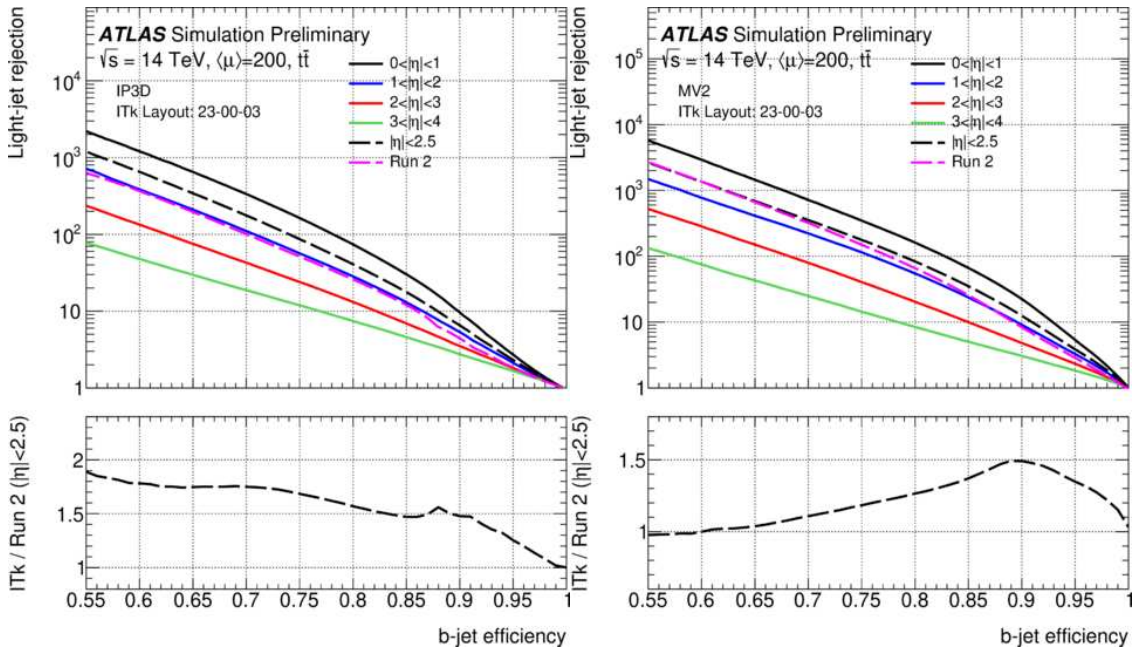
(a) (b)

Figure 3 – (a) Schéma illustrant un vertex secondaire dans un jet et les paramètres d'impact importants (d_0) des traces produites au niveau du vertex déplacé. (b) Distribution de la significance de d_0 pour des traces issues de jets de b et légers dans la région $|\eta| < 1$, pour différentes tailles de pixels de l'ITk.

Les développements en termes de b-tagging auxquels j'ai contribué se sont focalisés sur les algorithmes de bas niveau IP3D et SV1. Un des buts de mon travail a notamment été de valider l'implémentation des algorithmes de b-tagging dans la version la plus récente du software d'ATLAS dédiée aux études upgrade (r21.9) en vue du remplacement de la version précédente (r20.20). J'ai également mené des optimisations de IP3D et SV1 pour améliorer leurs performances. Les catégories de traces de l'algorithme IP3D ont été ainsi mises à jour pour mieux exploiter la dépendance en p_T de la résolution des paramètres d'impact dans la région centrale ($|\eta| < 2$) et la dépendance au contenu en *hits* dans la région avant ($|\eta| > 2$) de l'ITk. Une réjection des vertex secondaires issus d'interactions avec les matériaux du détecteur a également été implémentée en tenant compte de la géométrie complexe du détecteur à pixels de l'ITk.

Ces implémentations des algorithmes de b-tagging en r21.9, incluant ces nouvelles optimisations, ont été utilisées pour évaluer les performances de b-tagging avec la géométrie mise à jour de l'ITk. Les performances attendues ont été comparées avec celles obtenues avec l'ID du Run 2. La Figure 4 (a) montre ainsi les meilleures performances d'IP3D, liées à la meilleure résolution sur le paramètre d'impact avec l'ITk ainsi qu'à la nouvelle catégorisation réoptimisée des traces. La Figure 4 (b) présente les performances améliorées de MV2. Comme illustration, pour un point de fonctionnement à 77% d'efficacité de sélection des jets de b, une réjection des jets légers 20% plus élevée que celle obtenue avec le détecteur du Run 2 est observée, en lien avec les performances

améliorées d'IP3D. Pour tous les algorithmes, les performances se dégradent dans la région à plus grand $|\eta|$, dû à l'augmentation des effets de diffusion multiple.



(a) (b)

Figure 4 – Réjection des jets légers en fonction de l'efficacité de sélection des jets de *b* pour les algorithmes (a) IP3D et (b) MV2, dans différents intervalles en $|\eta|$, évaluée pour des événements $t\bar{t}$ avec un *pile-up* de 200 et la géométrie mise à jour de l'ITk. Pour comparaison, les performances obtenues avec l'ID du Run 2 et un *pile-up* moyen de 38 sont également représentées.

Analyse $t\bar{t}H(H \to b\bar{b})$ utilisant une nouvelle technique d'analyse multivariée basée sur des réseaux de neurones profonds

Les études présentées dans cette thèse sont effectuées en vue de la prochaine version de l'analyse $t\bar{t}H(H \to b\bar{b})$ basée sur l'ensemble des données collectées par ATLAS durant le Run 2, dénommée analyse *legacy*. Cette analyse vise à améliorer la version précédente grâce à de nouvelles méthodologies, une meilleure modélisation du bruit de fond simulé tt+bb et des algorithmes améliorés de reconstruction et d'identification des objets, principalement basés sur les jets reconstruits grâce au *Particle-Flow* et à l'algorithme de *b*-tagging DL1r. Cet algorithme est un des plus récemment développés pour les analyses d'ATLAS et présente des performances améliorées par rapport à MV2. L'analyse $t\bar{t}H(H \to b\bar{b})$ est effectuée dans le canal avec un lepton, où un des bosons W se désintègre en état final hadronique et l'autre en état final avec un lepton.

L'analyse utilise le formalisme de *Simplified Template Cross-Section* (STXS) pour mesurer la section efficace du signal en fonction de l'impulsion transverse du boson de Higgs $\left(p_T^H\right)$, c'est pourquoi il est important de reconstruire de façon précise cette variable. Six classes STXS différentes

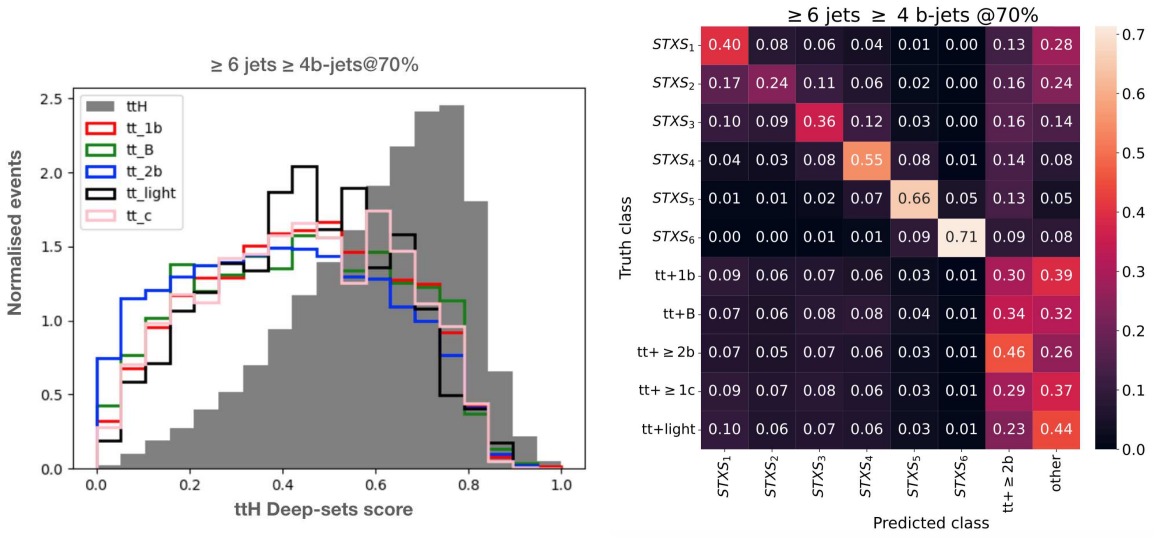
correspondant aux intervalles 0-60 GeV, 60-120 GeV, 120-200 GeV, 200-300 GeV, 300-450 GeV et 450-∞ GeV, dénotées $\text{STXS}_{1...6}$, sont définies pour la mesure STXS. Des sélections d'événements dédiées sont utilisées pour améliorer le rapport signal sur bruit. Les événements sont répartis entre régions enrichies en signal (*Signal Regions*, SR) et appauvries en signal (*Control Regions*, CR). Des techniques d'analyse multivariée (*Multivariate Analysis*, MVA) améliorent davantage la sensibilité dans les régions d'analyse. Une nouvelle technique de MVA basée sur des réseaux de neurones profonds (*Deep Neural Networks*, DNN) a été développée afin de remplacer les arbres de décision boostés (*Boosted Decision Trees*, BDT) utilisés précédemment et d'améliorer la sensibilité globale. Un ré-entraînement des BDT a été effectué pour l'analyse *legacy* et a servi de point de référence pour les nouveaux développements MVA utilisant des DNN. Un *profile-likelihood fit* a également été effectué, en combinant les SR et CR, pour extraire la section efficace $t\bar{t}H$.

Les techniques MVA examinent de façon simultanée plusieurs variables pour identifier des motifs récurrents et des corrélations. Il s'agit d'un outil utilisé en physique des hautes énergies pour reconstruire et classifier de façon efficace des processus physiques. Dans cette analyse, les techniques de MVA se basent sur les combinaisons de jets présents dans l'état final et leurs corrélations, exploitant notamment des masses invariantes et des distances angulaires. Chaque combinaison de jets est une permutation correspondant à une interprétation de l'événement associant les objets reconstruits aux produits de désintégration du boson de Higgs et des quarks top. Les DNN Deep-sets ont été exploités pour développer un nouveau modèle de MVA, car ils fournissent une invariance par permutation en traitant les combinaisons de jets en tant qu'ensemble non-ordonné pour leur entraînement. Le modèle Deep-sets combine les étapes suivantes:

- une régression sur la variable cinématique p_T^H, utilisée pour définir des classes STXS pour la mesure différentielle STXS,
- une multi-classification pour distinguer le signal $t\bar{t}H$ et les sous-composantes du bruit de fond $t\bar{t}$+jets, tels que $t\bar{t}$+1b, $t\bar{t}$+B, $t\bar{t}+ \geq 2b$ ou $t\bar{t}+ \geq 1c$, permettant de construire des CR utiles pour contraindre certaines systématiques et améliorer ainsi la sensibilité.

Le modèle Deep-sets produit dix scores discriminants: six scores STXS associés au signal $t\bar{t}H$ et quatre scores associés au bruit de fond. Chaque score est interprété comme la probabilité pour un événement d'être issu de la catégorie correspondante. Des régions d'analyse peuvent ainsi être définies en se basant sur le score le plus élevé parmi les dix. La somme des six scores STXS correspond à la probabilité inclusive d'être un événement de signal $t\bar{t}H$ et est appelé score $t\bar{t}H$. La Figure 5 (a) présente la distribution du score $t\bar{t}H$ pour les échantillons $t\bar{t}H$ et les sous-composantes tt+jets. La discrimination la plus importante est observée entre $t\bar{t}H$ et $t\bar{t}+ \geq 2$b,

indiquant une séparation performante. La Figure 5 (b) présente les performances du Deep-sets au travers d'une matrice de migration, qui indique la répartition des événements issus d'un processus donné entre les catégories prédites par le classificateur Deep-sets. La diagonale couvrant les classes STXS correspond aux fractions d'événements correctement assignés à chaque catégorie STXS. La catégorie prédite *other* correspond à la fusion des catégories $t\bar{t}+1b$, $t\bar{t}+B$ et $t\bar{t}+\geq 1c$, regroupées suite aux performances similaires des discriminants spécifiques à ces catégories. Les scores Deep-sets sont exploités par la suite pour le *profile-likelihood fit* utilisé pour l'extraction du signal.

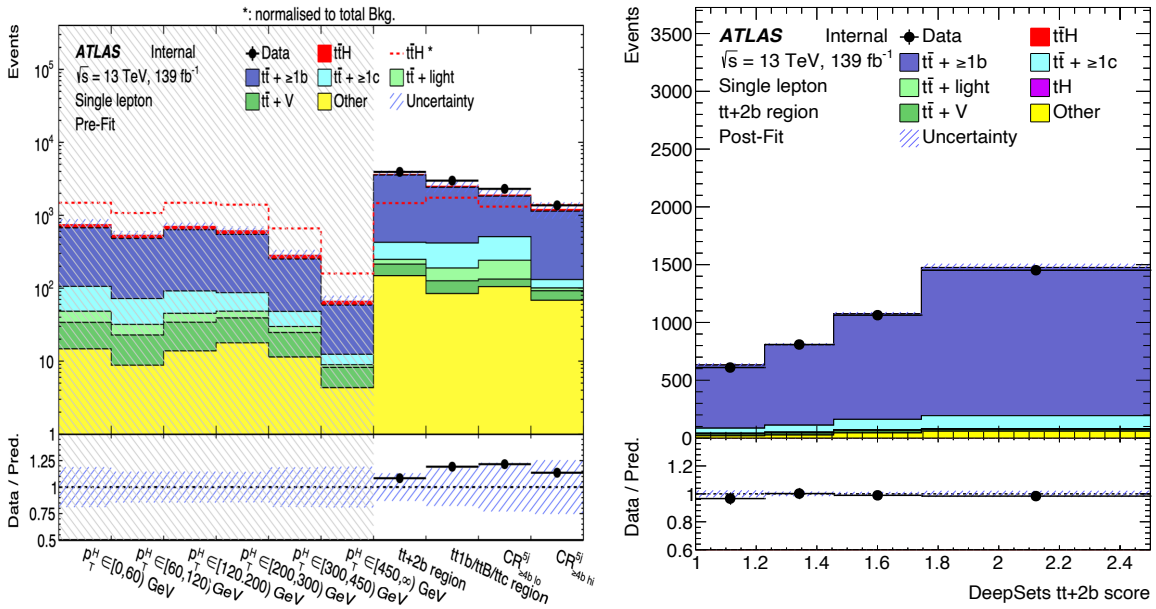


(a) (b)

Figure 5 – (a) Distribution du score $t\bar{t}H$ pour le signal $t\bar{t}H$ et les différentes sous-composantes du bruit de fond. (b) Matrice de migration, normalisée suivant les rangées, présentant les performances du modèle Deep-sets.

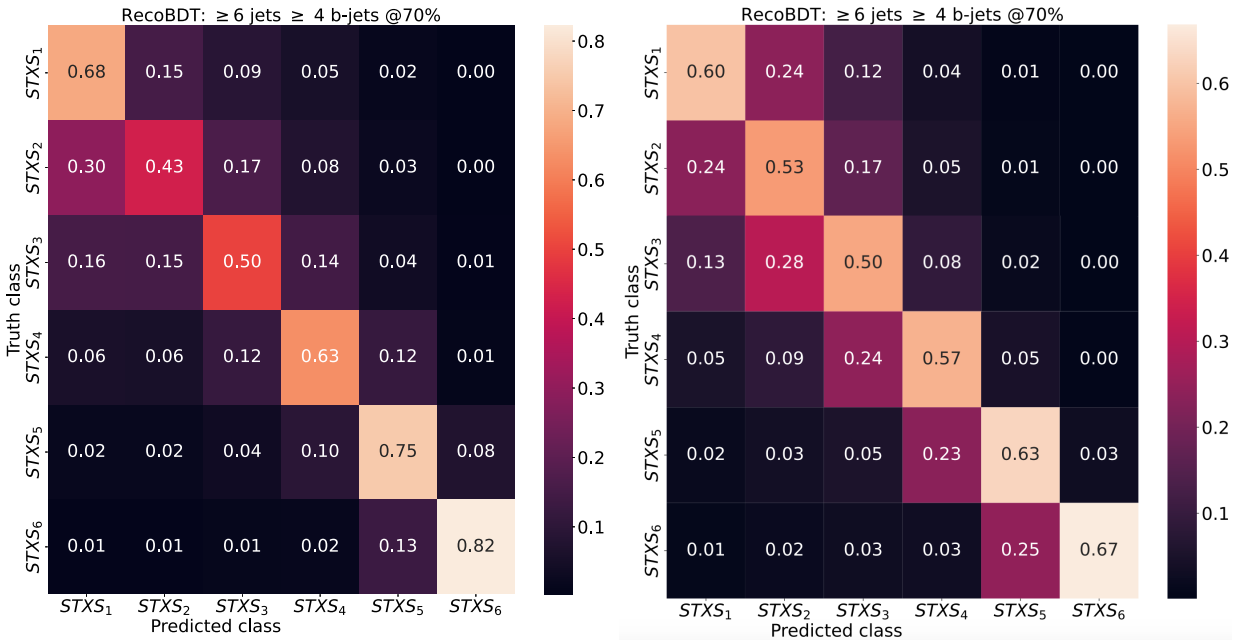
La Figure 6 (a) présente l'accord données-Monte Carlo obtenu pré-fit pour la distribution des événements dans les différentes catégories d'analyses.

La version précédente de l'analyse utilisait une stratégie basée sur deux BDT pour définir les régions STXS (recoBDT) et pour séparer le signal $t\bar{t}H$ du bruit de fond $t\bar{t}$+jets (classBDT). La variable classBDT était utilisée pour effectuer le *profile-likelihood fit* utilisé pour l'extraction du signal. Les performances de prédiction des catégories STXS pour l'ensemble des événements sont comparées entre Deep-sets et le recoBDT en Figure 7. Deep-sets présente des performances améliorées de façon significative pour les régions avec $p_T^H > 120$ GeV ($STXS_{3...6}$). La Figure 8 présente les courbes ROC (*Receiver Operating Characteristic*) illustrant les performances de classification de Deep-sets et du classBDT. La séparation obtenue entre le signal $t\bar{t}H$ et le bruit de fond tt+jets est similaire entre les deux méthodes mais Deep-sets a l'avantage d'offrir une discrimination supplémentaire entre les sous-composantes du bruit de fond.



(a) (b)

Figure 6 – (a) Distribution du nombre dévénements attendus et observés dans chaque SR et CR. Les distributions des données sont masquées dans les régions avec un rapport signal sur bruit S/B > 7.7%. (b) Distribution du score $t\bar{t}+ \geq 2b$ obtenue après un fit sans signal dans la CR $t\bar{t}+ \geq 2b$.



(a) (b)

Figure 7 – Matrice de migration des classes STXS prédites pour (a) Deep-sets et (b) le RecoBDT. La matrice est normalisée suivant chaque rangée.

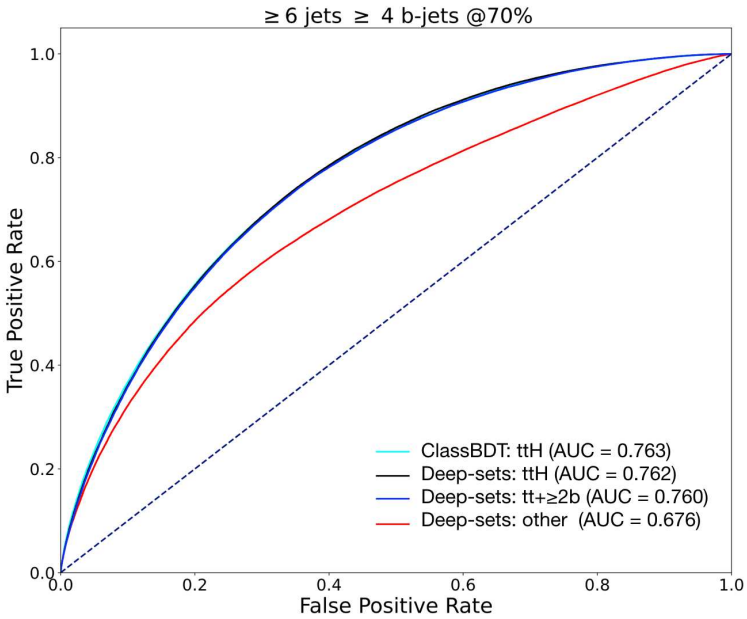
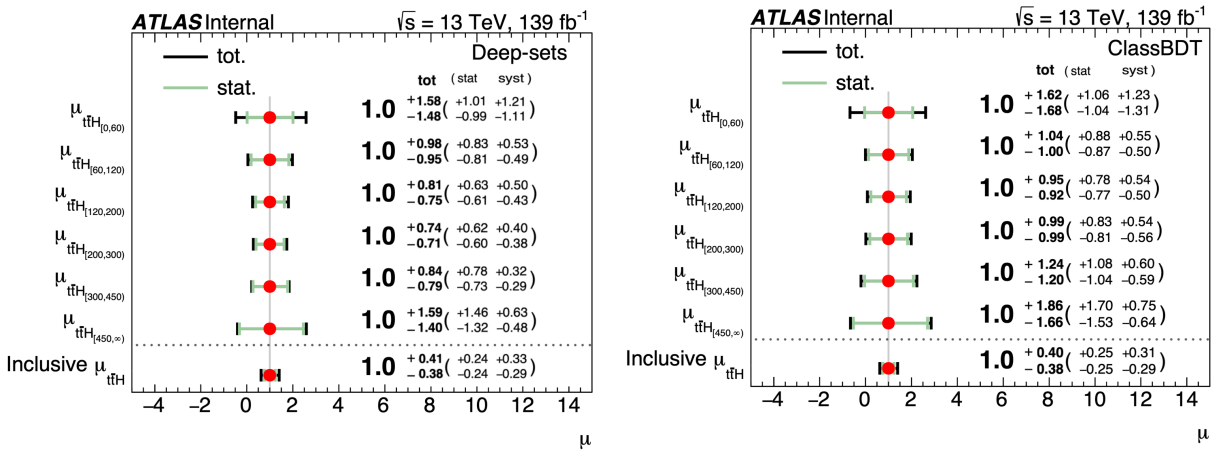


Figure 8 – Courbes ROC illustrant les performances de classification du multi-classificateur Deep-sets. Les courbes ClassBDT et Deep-sets $t\bar{t}H$ correspondent à la séparation entre $t\bar{t}H$ et les événements de bruit de fond inclusif. Les autres courbes Deep-sets correspondent à la séparation entre $t\bar{t}H$ et les sous-composantes du bruit de fond.

Les performances attendues pour l'analyse $t\bar{t}H(H \to b\bar{b})$ sont évaluées avec un *Asimov fit* incluant les incertitudes systématiques, sous l'hypothèse signal et bruit de fond. Un facteur de normalisation du bruit de fond $t\bar{t}+ \geq 1b$ flottant, sans contrainte pré-fit, est utilisé. Les résultats du fit sont analysés à la fois pour les mesures inclusive et STXS d'intensité du signal, correspondant au ratio entre la section-efficace mesurée et prédite par le MS. Pour les mesures STXS, le fit est effectué en utilisant plusieurs paramètres d'intensité du signal décorrélés, associés aux six régions STXS. La performance attendue obtenue avec les nouveaux discriminants Deep-sets est comparée avec celle des BDT utilisés précédemment, dans des configurations équivalentes. Les précisions attendues pour les mesures d'intensité du signal STXS et inclusive sont présentées en Figure 9 pour Deep-sets et les BDT. Dans chacune des région STXS, la précision attendue de la mesure est améliorée entre 5 et 35%, grâce à la réduction des corrélations entre les paramètres d'intensités du signal et les systématiques dominantes. Les incertitudes liées à la modélisation du bruit de fond ont le plus large impact sur la mesure inclusive. La significance attendue pour la mesure inclusive en utilisant Deep-sets est de 2.71σ, comparée à 2.54σ avec les BDTs.

Étant donné que l'analyse doit encore être soumise à la revue de la Collaboration ATLAS au moment de l'écriture de cette thèse, les résultats de la mesure ne sont pas présentés dans ce document, afin de ne pas biaiser la revue des résultats. Une configuration du fit dans l'hypothèse sans signal est utilisée pour effectuer un fit des données avec Deep-sets, fournissant des informations précieuses pour la



(a) (b)

Figure 9 – Précision attendue pour les mesures d'intensité du signal inclusive et STXS avec (a) Deep-sets et (b) les BDT.

validation de la modélisation des bruits de fond. Un bon accord post-fit entre données et Monte Carlo est observé dans les CR. Un bon accord données-Monte Carlo est obtenu post-fit, comme illustré dans la CR $t\bar{t}+ \geq 2b$ en Figure 6 (b). Le résultat de la mesure du facteur de normalisation du bruit de fond $t\bar{t}+ \geq 1b$ est :

$$k(t\bar{t}+ \geq 1b) = 1.12^{+0.06}_{-0.06}$$

Pour résumer le travail accompli durant cette thèse, les algorithmes de *b*-tagging étudiés ont bénéficié d'une optimisation dédiée afin d'exploiter la géométrie la plus récente du détecteur ITk. Les améliorations que j'ai développées pour les algorithmes de bas niveau IP3D et SV1 pourront servir pour la prochaine génération d'algorithmes de *b*-tagging basés sur des DNN en cours d'adaptation pour le détecteur ATLAS au HL-LHC. Le nouveau modèle Deep-sets que j'ai mis en place présente une amélioration significative des performances en termes de classification STXS, en comparaison des BDT, et permet également l'implémentation d'une CR optimisée pour les composantes non-$t\bar{t}+ \geq 2b$ du bruit de fond. Le modèle Deep-sets a été intégré dans le software central d'analyse utilisé par le groupe $t\bar{t}H(H \rightarrow b\bar{b})$ et pourra ainsi servir de méthode de référence pour l'analyse *legacy* en cours de finalisation.

INTRODUCTION

What are the fundamental building blocks of the universe? What are the rules that govern them ? Questions like these have perplexed humans for a long time. The Standard Model (SM) of fundamental particle physics is the theoretical framework describing our current understanding of elementary particles and explaining the fundamental forces that govern their interactions. For decades, the predictions of the SM have been tested and validated. The Higgs boson, which was the last missing piece of the SM, was discovered in 2012 by both ATLAS and CMS collaborations at the Large Hadron Collider (LHC). The breakthrough was a significant step forward in particle physics because it expanded our understanding of how elementary particles like quarks and leptons gain mass through interactions with the Higgs field. With many proton-proton collisions now recorded at the LHC, the properties of the Higgs boson can now be measured very precisely. One such property is how Higgs interacts with fermions. This interaction strength is proportional to the mass of the fermion, and the top quark, being the heaviest, interacts strongly with the Higgs. Its precise measurement is a powerful probe to unravel new physics phenomena like the new sources of Charge-Parity violation that might exist beyond the SM. Chapter 1 provides an overview of the SM with a focus on the Higgs mechanism and the Higgs boson searches at the LHC.

The LHC is the most complex circular particle accelerator ever built. The ATLAS detector, which is the largest general-purpose particle detector at the LHC, is designed to test the SM of particle physics while also assisting in the exploration of the physics beyond the SM. Chapter 2 describes the LHC and ATLAS experiment setups. Since the Higgs boson cannot be directly detected by the detector, different Higgs decay modes are investigated. The production of the Higgs boson in association with a pair of top quarks, in which the Higgs decays into a pair of bottom quarks, $t\bar{t}H(H \rightarrow b\bar{b})$, allows direct measurement of the top-Yukawa coupling. This process suffers from the large irreducible background production of $t\bar{t}$ with additional jets ($t\bar{t}$ + jets) and mainly when these jets are b-jets, which originate from the hadronization of b-quark. The $t\bar{t}H(H \rightarrow b\bar{b})$ process has a challenging final state with multiple jets where at least four jets are b-jets. Thus, a sophisticated b-jet identification method known as b-tagging is employed in the ATLAS experiments. The b-tagging is crucial not only for the $t\bar{t}H(H \rightarrow b\bar{b})$ process, but also for a large fraction of physics analyses at the LHC. It will also play a key role at the upcoming High-luminosity LHC (HL-LHC) phase to investigate Higgs boson self-coupling and maximise sensitivity for physics analyses.

During my thesis, I mainly contributed to addressing the challenges and performing developments in the areas of *b*-tagging, mainly for the Phase 2 upgrade of the ATLAS detector at the HL-LHC, and the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis using full Run 2 ATLAS data-set at the LHC. My contributions to each of these projects are detailed in Chapters 3, 4 and 5.

CHAPTER

1

THEORETICAL INTRODUCTION AND HIGGS BOSON SEARCHES AT THE LHC

The Standard Model (SM) of fundamental particle physics [2–4] is the theoretical framework describing our current understanding of elementary particles and explaining the fundamental forces that govern their interactions. The SM describes three of the four fundamental interactions in nature i.e. the electromagnetic, the strong, and the weak forces, and only gravity remains unexplained. This theory has been probed over the last decades with enormous precision, although there are also some limitations in explaining many phenomena such as baryon asymmetry, the origin of mass, gravity, dark mass, dark energy, etc. This chapter presents the relevant theory for describing the SM along with an overview of the experimental searches for the Higgs boson at the LHC experiments.

Section 1.1 gives a brief overview of the SM and its particle composition. Section 1.2 and Section 1.3 provide an introduction to the Quantum Electrodynamics (QED) and electroweak unification, respectively. The Higgs boson mechanism and its role in electroweak symmetry breaking are discussed in Section 1.4 followed by an overview of Quantum Chromodynamics (QCD), presented in Section 1.5. The characteristics of the proton-proton collisions at the LHC involving Parton Distribution Functions (PDFs) are discussed in Section 1.6. The searches for the Higgs boson at LHC are described in Section 1.7 followed by the latest results of $t\bar{t}H$ measurements at the LHC experiments, presented in Section 1.8.

1.1 The Standard Model overview

The SM provides a theoretical framework that captures the relationship between the fundamental interactions and the elementary particles to our best of knowledge . Figure 10 shows the summary

of known fundamental particles whose dynamics are described by the SM. The elementary matter constituents are the *fermions* of spin 1/2, obeying the Fermi-Dirac statistics, and their interactions are mediated by the exchange of spin 1 particles called *bosons*, governed by the Bose-Einstein statistics. The fermions are classified between leptons and quarks. These fermions are classified into three generations, each consisting of one charged lepton, one neutrino, and two quarks. The particles of a different generation have identical quantum numbers with the exception of their mass, which increases as we go from the first to the third generation. The first-generation fermions are the ones that make up ordinary matter. In addition, each and every fermion particle is paired with an anti-particle that carries the opposite charge.

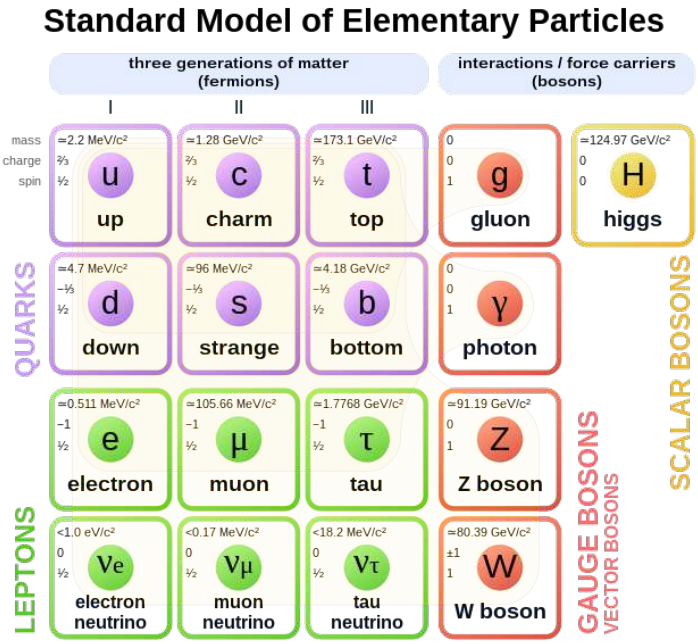


Figure 10 – Summary of the SM elementary particles and their properties.

The quarks carry an electric and a colour charge, and hence, they can interact weakly, electromagnetically as well strongly. Each generation has an up-type quark with an electric charge of e = + 2/3 and a down-type quark with charge e = - 1/3. The first generation consists of up (u) and down (d) quarks, the second of charm (c) and strange (s) quarks, and the third of top (t) and bottom (b) quarks. In nature, only particles without colour charge are observed, and free quarks have never been observed. This is explained by colour confinement, where the quarks exist in bound states called *hadrons* and cannot propagate as free particles. These bound states can be composed of a quark-anti-quark pair, in which case they are called *mesons*, or can be composed of three quarks in a formation called *baryons*. Hadrons with four or five quarks have also been observed. These multi-quark hadrons are called *exotics hadrons*.

In regards to leptons in each generation, there are the electron (e), the muons (μ) and the taus (τ), and the associated neutrinos are the electron-neutrino (ν_e), the muon-neutrino (ν_μ) and the tau-neutrino (ν_τ), respectively. The leptons (e, μ, τ) carry electric charge of e = - 1 whereas the neutrinos have no electric charge. Both charged leptons and neutrinos have no colour charge and interact through weak forces, as well as electromagnetic forces in the case of charged leptons. Table 1 provides a summary of the masses of the quarks and leptons as well as their respective electric charges.

The gauge bosons are the vector bosons, which are the carrier particles for the three fundamental forces described by the SM. The photon (γ), which is massless, is the mediator of electromagnetic force. The $W^{\pm}$ and Z bosons are massive and are mediators of weak forces. The gluons (g) are the force carriers of the strong interactions. There are a total of eight types of gluons. The characteristics of the gauge bosons are outlined in Table 2.

In addition to fermions and bosons, a scalar spin-zero particle also exists—the Higgs boson. Section 1.4 describes the Higgs mechanism and concept of electroweak symmetry breaking (EWSB) in detail.

	Quarks			Leptons		
Generation	Flavour	Mass	Charge (e)	Flavour	Mass	Charge (e)
1	u	$2.16^{+0.49}_{-0.26}$ MeV	+ 2/3	e	0.511 ± 1.5 x 10^{-10} MeV	-1
	d	$4.67^{+0.48}_{-0.17}$ MeV	- 1/3	ν_e	< 2 eV	0
2	c	1.27 ± 0.02 GeV	+ 2/3	μ	105.7 ± 2.3 x 10^{-6} MeV	-1
	s	$93.4^{+8.6}_{-3.4}$ MeV	- 1/3	nu_μ	< 0.19 MeV	0
3	t	172.89 ± 0.30 GeV	+ 2/3	τ	1776.8 ± 0.12 MeV	-1
	b	$4.18^{+0.03}_{-0.02}$ GeV	- 1/3	nu_τ	< 18.2 MeV	0

Table 1 – The summary of mass and electric charge of the quarks and leptons in SM [5].

Bosons	Interactions	Mass	Charge (e)
g	strong	0	0
γ	electromagnetic	0	0
$W^{\pm}$	weak	80.38 ± 0.012 GeV	± 1
Z	weak	91.19 ± 0.0021 GeV	0

Table 2 – A summary of the properties of the gauge boson in SM [5].

The SM is a Quantum Field Theory (QFT), postulating a set of symmetries. The QFT describes fundamental particles in the SM as an excitation in a quantum field defined at all points in space-time.

The QFT provides the mathematical framework of the SM, in which a Lagrangian controls the dynamics of the theory and is renormalizable and invariant under local gauge transformation. The SM is a non-abelian gauge theory and is invariant under the following gauge symmetry group:

$$G_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{1.1}$$

where C is colour, L represents that the weak isopsin group couples to the left-handed fermions, and Y corresponds to the hypercharge defined by the Gell-Mann-Nishijima formula:

$$\mathrm{e} = \mathrm{I}_3 + \frac{\mathrm{Y}}{2} \tag{1.2}$$

where, I_3 is the third component of the isospin and e is the electric charge. The $\mathrm{SU(3)_C}$ group is the strong interaction symmetry group. The $\mathrm{SU(2)_L \otimes U(1)_Y}$ gauge group corresponds to the gauge group of the unified weak and electromagnetic interactions. A more detailed discussion of the gauge groups, described above, will be provided in the coming sections.

1.2 Quantum Electrodynamics

The Quantum Electrodynamics, commonly named as QED is the theoretical description of the electromagnetic interactions between the charged fermions and the photons and is invariant under local U(1) transformation. The Dirac Lagrangian describes a freely propagating fermion field:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{1.3}$$

where, Ψ is the free spinor field, m is the mass of the particle, γ^μ are Dirac gamma matrices, ∂_μ is $\partial/\partial x_\mu$ derivative and $\bar{\Psi} = \Psi^\dagger \gamma^0$ denotes the Dirac adjoint of the spinor. The spinor field transforms as follows under the U(1) local space transformation:

$$\psi(x) \longrightarrow \psi' = \exp(ie\theta(x))\psi(x) \tag{1.4}$$

where, e is the electric charge, $\theta(x)$ corresponds to a generic function that describes the local phase. Applying this transformation to the free spin 1/2 particle, the Lagrangian gives:

$$\mathcal{L} \longrightarrow \mathcal{L}' = \mathcal{L} - q\bar{\psi}\gamma^\mu(\partial_\mu \theta(x))\psi \tag{1.5}$$

For the QED Lagrangian to be invariant under this gauge transformation, the derivative ∂_μ can be replaced by a covariant derivative (D_μ) which denotes the coupling between the Dirac fermion and the vector field A_μ:

$$\partial_\mu \longrightarrow D_\mu = \partial_\mu + ieA_\mu \tag{1.6}$$

For the addition of the new field A_μ to provide cancellation of the additional term i.e $q\bar{\psi}\gamma^\mu(\partial_\mu\theta(x))\psi$ and as result, A_μ transforms under gauge transformation as:

$$A_\mu \longrightarrow A'_\mu = A_\mu - \partial_\mu\theta(x) \tag{1.7}$$

The final QED Lagrangian which is invariant under a local U(1) gauge transformation, is:

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.8}$$

where, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength tensor and $F_{\mu\nu}F^{\mu\nu}$ describes the kinetic energy term of the photon.

1.3 Electroweak unification

The theory developed by Glashow, Salam and Weinberg [2–4], unifies the weak and electromagnetic forces under $SU(2)_L \otimes U(1)_Y$ gauge group. $SU(2)_L$ group is associated to the weak isospin group, and its generators in the spinor representation are the Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix};\ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix};\ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.9}$$

I_3 is the eigenvalue of the operator σ_3. $U(1)_Y$ is the Abelian hypercharge group. The generator of $U(1)_Y$ is simply denoted 1 and its eigenvalue by Y. A scalar field ϕ is introduced that transforms as a doublet under $SU(2)_L$ and is symmetric under gauge transformations:

$$\phi = e^{\left(i\alpha^a\sigma^a/2\right)}e^{(i\beta/2)}\phi; \qquad a \in 1,2,3, \tag{1.10}$$

where, β is the gauge coupling constant and α represents three functions specifying the local phase in space-time. A kinetic term and a potential term ϕ are used to build the electroweak Lagrangian:

$$\mathcal{L} = |D_\mu^\phi|^2 + V(\phi) \tag{1.11}$$

where D_μ is the electroweak covariant derivative. For the kinetic term to be invariant under the gauge transformation, D_μ must take the form:

$$D_\mu \phi = \left(\partial_\mu - i\frac{g'_\mu}{2} B_\mu - i\frac{g}{2} W_\mu^j \sigma_j \right) \phi \tag{1.12}$$

where $W^{1,2,3}$, B are the field strengths and g, g' are the coupling constants associated to the gauge groups $\mathrm{SU(2)_L}$ and $\mathrm{U(1)_Y}$, respectively. The electroweak mixing angle θ_W, quantifies the relative strengths of electromagnetism and the weak force and can be expressed in terms of coupling constants:

$$\cos(\theta_W) = \frac{g}{\sqrt{g^2 + g'^2}}, \quad \sin(\theta_W) = \frac{g'}{\sqrt{g^2 + g'^2}} \tag{1.13}$$

which also allows to represent the electric charge in terms of the gauge couplings as $e = g sin\theta_W$. By expanding the Lagrangian in terms of field tensors and employing Pauli matrix identities, the Lagrangian is transformed into:

$$\mathcal{L} = |\partial_\mu \phi|^2 + \frac{g'^2}{2} B_\mu B^\mu |\phi|^2 + \frac{g^2}{2} W_\mu^j W_j^\mu |\phi|^2 + \frac{g' g}{4} B_\mu W^{k\mu} \left(\phi^\dagger \sigma_k \phi + (\sigma_k) \phi^\dagger \right) + V(\phi) \tag{1.14}$$

for the simplicity following substitutions are applied:

$$A_\mu = \frac{g' B_\mu + g W_\mu^3}{\sqrt{g'^2 + g^2}} \tag{1.15}$$

$$W_\mu^\pm = \frac{W_\mu^1 \pm i W_\mu^2}{\sqrt{2}} \tag{1.16}$$

$$Z_\mu = \frac{g' B_\mu - g W_\mu^3}{\sqrt{g'^2 + g^2}} \tag{1.17}$$

where, the transformed field strength A_μ corresponds to the electromagnetic field and has eigenvalue $e = I_3 + Y$, and the transformed field strengths W_μ^+, W_μ^- and Z_μ correspond to the W and Z bosons, respectively. The final Lagrangian becomes:

$$\mathcal{L} = |\partial_\mu \phi|^2 + \frac{g'^2 + g^2}{2} Z_\mu Z^\mu |\phi|^2 + \frac{g^2}{2} W_\mu^+ W^{-\mu} |\phi|^2 + V(\phi) \tag{1.18}$$

When postulating the SM based on symmetries, the particles in the model should be massless because introducing mass term in the Lagrangian for bosons and fermions would violate the local gauge invariance. Experiments, on the other hand, have provided clear evidence that particles such as electroweak gauge bosons and fermions are massive. Both gauge and fermion mass terms are introduced with the mechanism of spontaneous electroweak symmetry breaking described in the next section.

1.4 The Higgs mechanism

The mass generation of the fermions and the electroweak gauge bosons is introduced through the Brout-Englert-Higgs (BEH) mechanism [6], with the spontaneous electroweak symmetry breaking (EWSB). The potential for the doublet $V(\phi)$ is taken to be:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \frac{\lambda}{2} \left(\phi^\dagger \phi \right)^2 \tag{1.19}$$

where, μ and λ are scalar constants and, in particular, λ describes the quadratic self-interaction among the scalar fields. The spontaneous symmetry breaking is based on the non-invariance of the vacuum state with respect to the SU(2) symmetry. When μ^2 and λ are both positive, the minimum of the potential is found in the unique configuration $\phi = 0$. If $\mu^2 < 0$ and $\lambda > 0$, the minimum of the potential $\mathrm{V}(\phi)$ is described by an infinite number of solutions satisfying:

$$\phi^\dagger \phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda} \tag{1.20}$$

where, $\phi_{1,2,3,4}$ are real scalar fields. Figure 11 shows the shape of the Higgs field potential $\mathrm{V}(\phi)$. The symmetry is spontaneously broken, choosing the system to fall into one of the multiple ground states, for example:

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.21}$$

Then, expanding around the ground state, the Higgs field itself is given by:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+H \end{pmatrix} \tag{1.22}$$

where H is the real scalar field corresponding to the Higgs boson. Performing this substitution, the Lagrangian becomes:

$$\mathcal{L} = (\partial_\mu H)^2 + \frac{g'^2 + g^2}{2} Z_\mu Z^\mu (v+H)^2 + g^2 W_\mu^+ W^{-\mu} (v+H)^2 + V(v+H) \tag{1.23}$$

The mass terms for W and Z bosons have been obtained in addition to the interactions between these particles and the Higgs field H:

$$m_W = \frac{1}{2} v g_w, \quad m_Z = \frac{1}{2} v \sqrt{g_w^2 + g'^2}, \quad m_\gamma = 0 \tag{1.24}$$

where the photon remains massless and does not interact with the Higgs boson at tree level. The last term in eq 1.23 gives the Higgs mass term ($m_H = v\sqrt{2\lambda}$) as well as the interactions between the Higgs boson and itself. Self-coupling of the Higgs boson is a parameter of significant interest, and the experimental studies have been performed [7, 8].

Fermion masses and interactions:

The mechanism used to explain how fermions acquire mass is based on the same principles as the vector boson masses, but the formalism is slightly different. The Lagrangian for a fermion field ψ can be written as:

$$\mathcal{L}_\psi = \bar{\psi} \gamma^\mu D_\mu \psi = \bar{\psi} i \not{D}_\mu \psi \tag{1.25}$$

where the covariant derivative D_μ is same as eq 1.12. The term in the Lagrangian can be broken into two: left- and right-handed components of the field:

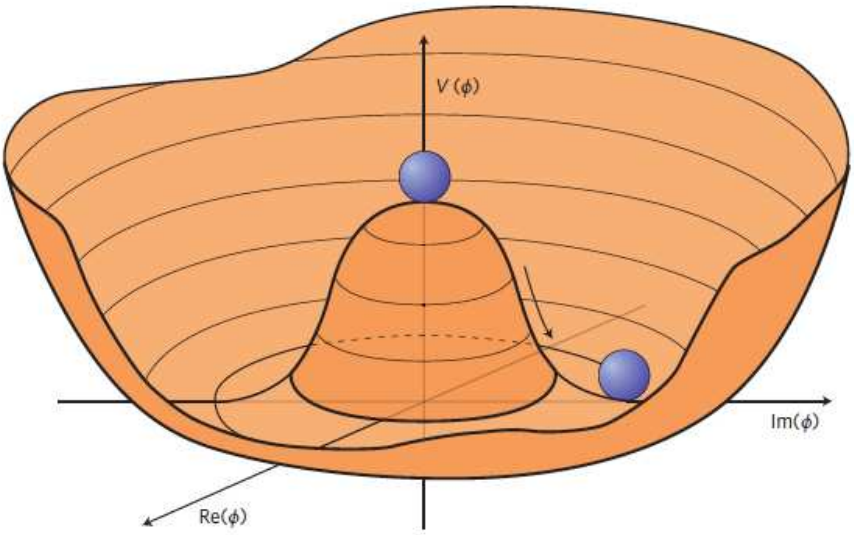


Figure 11 – The shape of the Higgs field potential V(ϕ) for a single complex field in the simplified case of the U(1) rotational symmetry.

$$\bar{\psi} i \not{D} \psi = \bar{\psi}_L \not{D} \psi_L + \bar{\psi}_R i \not{D} \psi_R \tag{1.26}$$

Since the W-boson only couples to the left-handed components of a fermion field, left-handed fermions in the $SU(2)_L$ gauge group are represented as doublets, while right-handed fermions are represented as singlets. As a result, right-handed fermions have $I_3 = 0$ and $Y = e$, whereas left-handed fermions have $I_3 = 1/2$ and $Y = e \pm 1/2$.

One can write the Lagrangian for a single generation of fermions as:

$$\mathcal{L} = \bar{E}_L i \not{D} E_L + \bar{e}_R u \not{D} e_R + \bar{Q}_L i \not{D} Q_L + \bar{u}_R i \not{D} u_R + \bar{d}_R i \not{D} d_R \tag{1.27}$$

where, E_L and Q_L are left-handed lepton doublets and the left-handed quark doublets, respectively. Expansion of the covariant derivative yields kinetic terms for every fermion as well as fermion-gauge boson interactions. To generate the masses of the fermions, an additional Yukawa term is introduced in the Lagrangian:

$$\mathcal{L}_e = \lambda_e \bar{E}_{Li} \phi_{Rj} \tag{1.28}$$

where ϕ is the Higgs doublet and λ_e is the electron Yukawa coupling. The Lagrangian for the first generation leptons (here e) is given by:

$$\mathcal{L}_e = \frac{-v\lambda_e}{\sqrt{2}}\bar{e}_L e_R + \frac{\lambda_e}{\sqrt{2}} H \bar{e}_L e_R \tag{1.29}$$

The mass of e obtained is $m_e = \lambda_e v/\sqrt{2}$, and the strength of the electron Higgs coupling is proportional to m_e In a similar way, it is possible to generate mass terms and Higgs boson coupling terms for other charged leptons and down-type quarks. The mechanism to explain the mass of the up-type quarks requires the gauge-invariant transformation $\phi \to i\sigma_2\phi^*$ which gives Yukawa terms, a slightly different form:

$$\mathcal{L}_u = -i\lambda_\mu \bar{Q}_L \sigma_2 \phi^* u_R \tag{1.30}$$

In the case of the top quark, which is the focus of this thesis work, the mass and coupling terms are identical to those of the electron:

$$\mathcal{L}_t = \frac{-v\lambda_t}{\sqrt{2}}\bar{t}_L t_R + \frac{\lambda_t}{\sqrt{2}} H \bar{t}_L t_R \tag{1.31}$$

with the top mass proportional to the Yukawa coupling via $m_t = \lambda_t v/\sqrt{2}$.

The Pontecorvo–Maki–Nakagawa–Sakata matrix (PMNS matrix), represented by a 3×3 unitary mixing matrix, describes the mixing between the neutrino mass and three flavour eigenstates [9, 10]. The lepton mass matrix G^{ℓ}_{ij} has no off-diagonal entries since the lepton number is conserved for each generation in the SM and the neutrinos are assumed to be massless. However, this is not the case with the quark mass matrix $G^{u,d}_{ij}$, where off-diagonal entries are present and can be diagonalised via four unitary matrices $V^{u,d}_{L,R}$, which results in the mass eigenstates:

$$\bar{u}^i_{L,R} = \left(V^u_{L,R}\right)_{ik} u^k_{L,R}; \qquad \bar{d}^i_{L,R} = \left(V^d_{L,R}\right)_{ik} u^k_{L,R} \tag{1.32}$$

By introducing mass eigenstates into the Lagrange density, its structure is preserved with the exception of the flavor-changing quark interactions that are mediated by the charged vector bosons. The quark mixing matrix, also known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix, is as follows:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.33}$$

$$V_L^u V_L^{d\dagger} = V_{CKM} \tag{1.34}$$

where, the diagonal elements of the CKM matrix are close to one ($|V_{tb}| = 0.999105 \pm 0.000032$) [5].

1.5 Quantum Chromodynamics

The theory that describes strong interactions based on $SU(3)_C$ colour charge is preserved is called Quantum Chromodynamics (QCD). The QCD Lagrangian governs the dynamics of quarks and gluons. The Lagrangian can be obtained in the same way as QED. The quark field can be expressed as colour triplets $\bar{q_k} = (q_{\bar{red}}, q_{\bar{blue}}, q_{\bar{green}})$ that undergo a local gauge transformation as:

$$q_k(x) \longrightarrow \exp(i\alpha_s\lambda_a/2(x))\, q_k(x), \qquad \alpha \in \mathbb{R}, \qquad a \in 1,...,8 \tag{1.35}$$

where, k is the flavour index, α_a are the coupling constant of strong interactions, λ_a are the generators of the SU(3) group called Gell-Mann matrices [11] and a is the colour index. The coupling between quarks and gluons is added as the appropriate covariant derivative D_μ to maintain the local gauge invariance of the Lagrangian:

$$\mathrm{D}_\mu = \partial_\mu - ig_s\frac{\lambda_a}{2}G_\mu^a(x) \tag{1.36}$$

where, G_μ^a are the eight gluon field strength tensors and g_s is the coupling constant, which can be expressed as the function of the coupling constant of the strong interaction as:

$$\alpha_s = \frac{g_s^2}{4\pi} \tag{1.37}$$

A gluon field transforms under $\mathrm{SU(3)_C}$ according to:

$$G_\mu^a \longrightarrow \bar{G}_\mu^a = G_\mu^a - \frac{1}{g_s}\partial_\mu\theta^a(x) + f^{abc}\theta_b(x)G_{\mu c}(x) \tag{1.38}$$

in order to guarantee gauge invariance of the QCD Lagrangian under colour charge transformations, where f^{abc} are structure constants of $\mathrm{SU(3)_C}$ group. The final QCD Lagrangian which is SU(3) gauge invariant is:

$$\mathcal{L}_{QCD} = \Sigma_k \bar{q}_k \left(i\gamma^\mu D_\mu - m_k \right) q_k - \frac{1}{4} G_a^{\mu\nu} G^a_{\mu\nu} \tag{1.39}$$

with the gluon field strength tensor:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f^{abc} G^b_\mu G^c_\nu \tag{1.40}$$

The additional term $\frac{1}{4} \mathrm{G}_{\mathrm{a}}^{\mu\nu} \mathrm{G}^{\mathrm{a}}_{\mu\nu}$ is needed in order to maintain local gauge invariance of the Lagrangian and results in the gluon self interactions. The gluons couple to themselves via three-gluon or four-gluon interactions.

The rate or cross-section of the strong interactions can be expressed as a power series in α_s:

$$\sigma_{QCD} = \Sigma_i^\infty c_i \alpha_s^i \tag{1.41}$$

When α_s is small, this expansion is perturbative, meaning that the contribution of each subsequent term is small compared to the contribution of the preceding terms. If the highest order term in a σ_{QCD} calculation is linear in α_s, the calculation is referred to as *leading order* (LO). If the sum is completed up to the quadratic or cubic term in α_s, the calculation is referred to as *next-to-leading order* (NLO) or *next-to-next-to-leading order* (NNLO), respectively.

When compared to QED, self-coupling causes a different energy scaling behaviour. The dependence of α_s on the energy-scale (renormalization scale μ_{R}) is known as *running of the strong coupling* [12], given by:

$$\alpha_s(\mu_R^2) = \frac{12\pi}{(33 - 2n_f) \log\left(\frac{\mu_R^2}{\Lambda^2_{QCD}} \right)} \tag{1.42}$$

where, Λ_{QCD} is the non-perturbative constant indicating the scale at which the coupling diverges, and n_f is the number of light quark flavours. The running of the strong coupling is shown in Figure 12 for $\mu_{\mathrm{R}} = \mathrm{Q}^2$, where Q^2 is the momentum transfer.

As the μ_{R} is varied, each term in the perturbative expansion in 1.41 also changes. However, the infinite sum remains unchanged. Generally, the μ_{R} values employed in pertubative QCD calculations are set to $\mathcal{O}(\mathrm{Q}^2)$, ensuring that higher order terms remain small. The application of perturbation

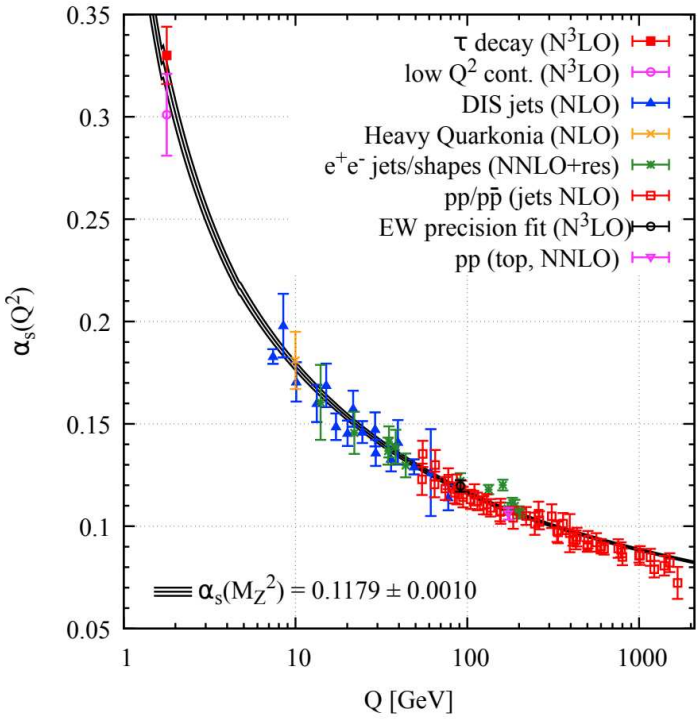


Figure 12 – Measurements of the running of the strong coupling collected by the Particle Data Group [5].

theory to compute scattering amplitudes is only possible for scales where $Q^2 \gg \Lambda_{QCD}$ where, $\alpha_s(Q^2) \ll 1$. At low energy scales, the effective coupling between two coloured particles increases, and thus coloured objects cannot freely exist and form colourless bound states (hadrons). This phenomenon is known as *colour confinement*. However, at high Q^2 the α_s is very small and quarks are not confined to bound states. This phenomena is characterised as *asymptotic freedom* [13] in QCD. This behaviour contrasts strikingly with QED, for example, where interaction rates fall nearly to zero for particles that are separated by very long distances or decreasing Q^2. The calculations of collider phenomena, discussed in Section 1.6, are divided into two regimes [14, 15].

1. Physics at short distance where Q^2 is high Q^2 (perturbative in QCD)
2. Physics at long distance where Q^2 is low (non-perturbative in QCD)

The separation of the two regimes can be done via factorization scale μ_F. The μ_F value is set such that α_s is sufficiently small and the cross-section can be calculated at fixed order in perturbation theory. Below this value, calculations are non-perturbative in QCD and predictions rely on other phenomenological models of low-energy. μ_F is commonly set to $\mathcal{O}(Q^2)$, like μ_R. The calculations at higher order terms are considered to be independent of the renormalization and factorization

scales. However, some small residual dependence on μ_F and μ_R are expected from finite perturbative expansions and can be covered as systematic uncertainties (see Section 5.5).

1.6 Characteristics of proton–proton collisions

In order to investigate physics of the elementary particles at the highest energies and the smallest distance scales, collider experiments focus on *hard-scatter* (HS) interactions. A large momentum transfer Q^2 occurs between two incident particles in hard-scatter interactions. For $Q^2 \gg \Lambda_{QCD}$, $\alpha_s(Q^2) \ll 1$, hadrons are considered to be made of weakly interacting constituent called *partons* (quark, anti-quark and gluons).

One constituent parton of each incoming proton participates in the hard scatter interaction in the proton-proton (pp) collisions. The remaining partons can contribute to the sum of all the processes that build up the final hadronic state in a collision, called *underlying event* (UE). Additionally, the UE consists of particles from initial- and final-state radiation resulting from the hard interaction. Figure 13 shows a diagram of a proton–proton event containing a HS interaction of cross section $\hat{\sigma}$.

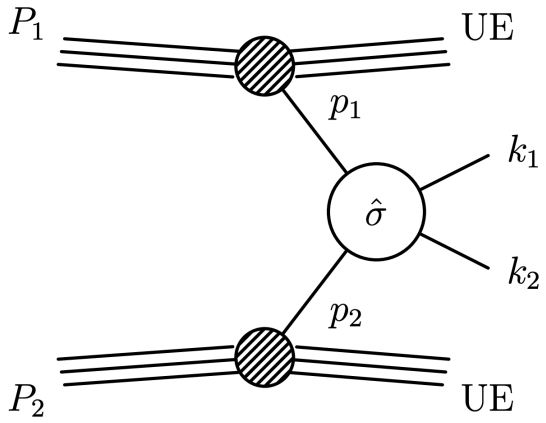


Figure 13 – Feynman diagram of an example proton–proton event containing a hard-scatter interaction of cross section $\hat{\sigma}$.

The distribution of the momentum fraction of the partons are described by the Parton Distribution Functions (PDFs). The PDFs are convolved with parton-parton HS cross sections $\hat{\sigma}$ to obtain the full pp cross section σ:

$$\sigma(P_1,P_2) = \Sigma_{ij} \int dx_1 dx_2 f_i(x_1,\mu_F^2) f_j(x_2,\mu_F^2) \hat{\sigma}_{ij}(p_1,p_2,\alpha_s(\mu_F^2),Q^2/\mu_F^2) \tag{1.43}$$

where, $\hat{\sigma}$ is calculated to fixed order in perturbation theory for all the combinations of incoming partons, denoted by i and j. $f_{i,j}$ represents their corresponding PDFs.

Parton Distribution Functions: As mentioned before, the PDFs describes the probability density of a parton inside a proton to carry a certain momentum fraction x (Bjrken x) = $\overrightarrow{p}/\overrightarrow{P}$, where p is the parton momentum and P is the proton momentum. The PDF depends on μ_F. Measurements of PDFs are primarily based on data from Deep Inelastic Scattering, with additional input data from neutrino scattering and hadron collider measurements [14, 15]. An example of PDFs at two values of $\mu_F^2 = Q^2$ are shown in Figure 14. At high values of Q^2, the gluon is the dominant proton constituent.

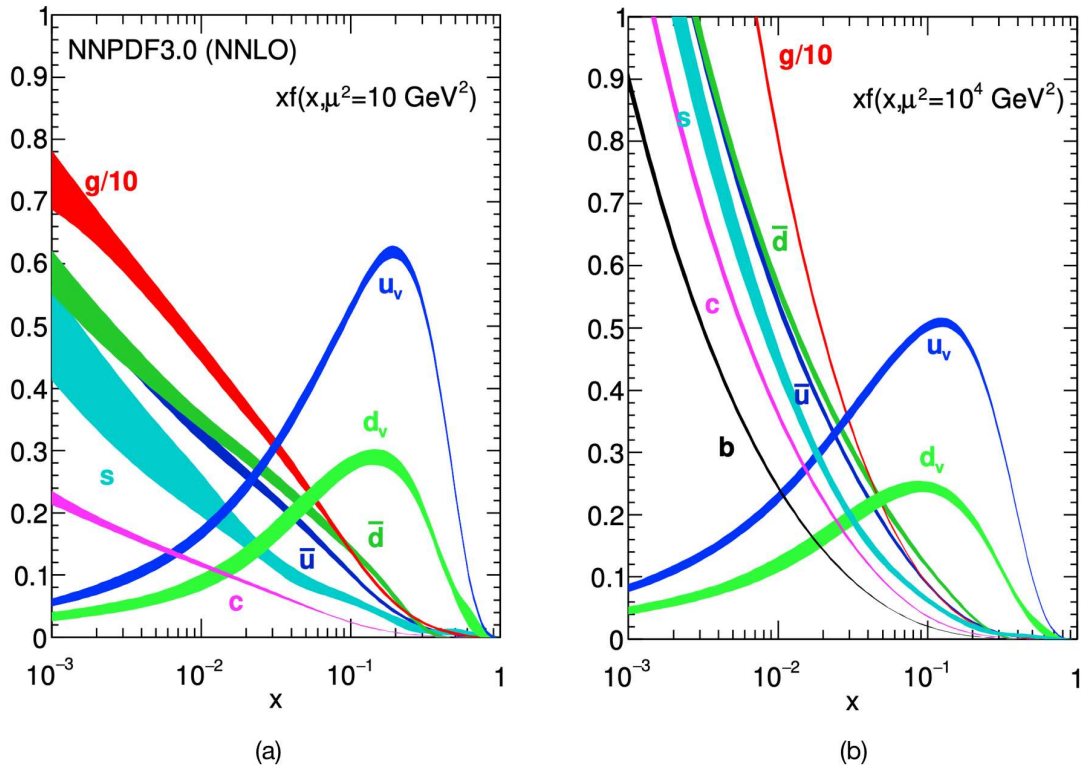


Figure 14 – The Parton Distribution Functions (PDFs), shown for two different factorisation scales $\mu_F^2 = Q^2$ = (a) 10 GeV^2 and (b) 10^4 GeV^2 [5].

Four flavour scheme and five flavour scheme: The processes involving b-quarks at HS level can be treated in two distinct ways in QCD using two different factorisation schemes based on the b-quark mass $\Lambda_{QCD} < m_b \ll v$: the four flavour scheme (4FS) and the five flavour scheme (5FS). In 5FS, a massless b-quark is included among the constituents of the proton. The b-quarks are treated in the same manner as the other light quarks. It comprises a b-quark PDF and the number of light flavour quarks is set to $n_f = 5$ in eq 1.42. In 4FS, the b-quarks are considered to be massive and they appear only in the final state. The presence of b-quarks in the proton can happen only through initial-state gluon splitting. The b-quarks PDF is set to zero, decoupling it from the QCD

perturbative evolution, which includes only the four lightest flavours and the gluon. Thus also decouples from the running of α_s and the number of light flavour quarks is set to $n_f = 4$. For high scales, mass effects become negligible, collinear logarithms $\left(\log(Q^2/m_b)\right)$ associated with b-quark radiation are large and must be resummed, and the 5FS is always more accurate. In contrast, mass effects are significant near the production threshold, while collinear logs are small, and the 4FS is more precise.

The QCD radiations below μ_F are modelled by parton showering (PS) algorithms. After the calculation of the HS cross section, additional QCD radiation below the factorization scale is modeled by parton showering algorithms. Parton showering algorithms use many different models, such as angular ordered [16] and dipole [17] showering, to simulate chains of low-energy and small-angle QCD radiation. To combine a parton showering algorithm with an HS matrix element calculation, matching and merging algorithms are used to make sure the calculations are the same and there is over-counting.

During the parton showering process, the quarks and gluons lose energy and move from a high Q^2 regime to a low Q^2 regime. During this transition, partons undergo hadronisation, resulting in a collimated *jet* of hadrons. There are several models that can be used to simulate hadronization and colour flow in hadronic final states. The string [18] and cluster [19] models are the most common ones. The underlying event, parton shower and hadronization, together referred as UEPS, are non-perturbative in QCD. Generally, the parameters of these models are adjusted (tuned) [20] based on data rather than being predicted a priori. These tunes are used a lot in simulations and predictions, and the differences in their predictions are covered by taking into account the systematic uncertainties.

1.7 Search for the Higgs Boson at LHC

The Higgs boson, which was the last missing piece of the SM, was discovered in 2012 by both ATLAS [21] and CMS [22] collaborations at the LHC. This scalar particle associated to the Higgs field, has a mass of (125.10 ± 0.14) GeV [5]. With mass parameter as fixed, its couplings and production rates can be calculated. Since its discovery, with much larger data-set available at the LHC, its properties have been investigated and found to be consistent with the SM predictions. The precision measurements of the Higgs boson properties and couplings is crucial to probe the nature of the EWSB and to unravel new physics phenomena [23].

Different Higgs production modes in various Higgs decay channels are exploited by the LHC experiments to perform measurements of its properties. These production modes are listed here in decreasing order of their cross-sections:

- **Gluon-gluon fusion (ggH):** It is the dominant production mode, mediated by a virtual quark loop, with the dominant contribution from the top due to its large Yukawa coupling.
- **Vector boson fusion (VBF):** In this production mode, the fusion of two vector boson (either $W^{\pm}$ or Z) produces a Higgs boson. This mode gives direct access to the coupling of heavy gauge bosons to the Higgs. The initial quarks that are radiated are only slightly deviated, and travel along their initial paths. The jets are mainly the remnants of the protons so they show up two forward jets in the detector with large transverse momenta. Due to no QCD activity between them, these jets are distinct and forms the characteristic signature of the process used in the experimental analyses to exploit this production mode.
- **Associated production with a vector boson (VH):** This production mode is also named as Higgs Strahlung, where the Higgs boson is produced in association with a $W^{\pm}$ or a Z boson. The presence of $W^{\pm}$ / Z bosons in the final states is widely used experimentally as leptonic decay is possible which helps to better identify the events and to reduce the background contributions.
- **Associated production with top-quark pair ($t\bar{t}H$):** This production mode allows a direct measurement of the top Yukawa coupling, by producing the Higgs in the fusion of a top quark-antiquark pair or through radiation from a top quark. This occurs at the tree-level instead of virtual loop and allows the measurement in a model-independent way, unlike ggH.

Figure 15 shows the leading order Feynman diagrams of the *ggH*, *VBH* and *VH* processes. The leading order Feynman diagrams for $t\bar{t}H$ production are shown in Figure 16.

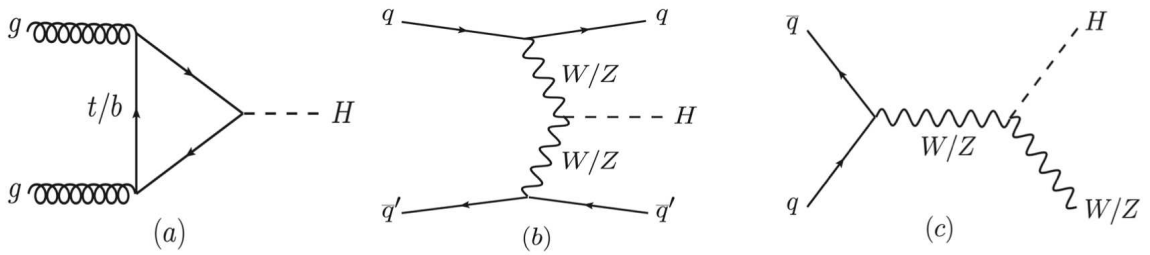


Figure 15 – Representative leading order Feynman diagrams of the four main production modes of the Higgs boson at the LHC [24]: (a) gluon-gluon fusion (ggH), (b) vector boson fusion (VBF), (c) Higgs strahlung (VH).

Figure 17 illustrates the dependence of the cross-sections of the main Higgs boson production modes on the center-of-mass energy in pp collisions. These cross-sections increase as the center-of-mass energy increases. When compared to Run 1 (2010-2013), where the center-of-mass energy was 7 and 8 TeV, the $t\bar{t}H$ cross-section increased 13 TeV i.e by a factor of four in Run 2 (2015-2018).

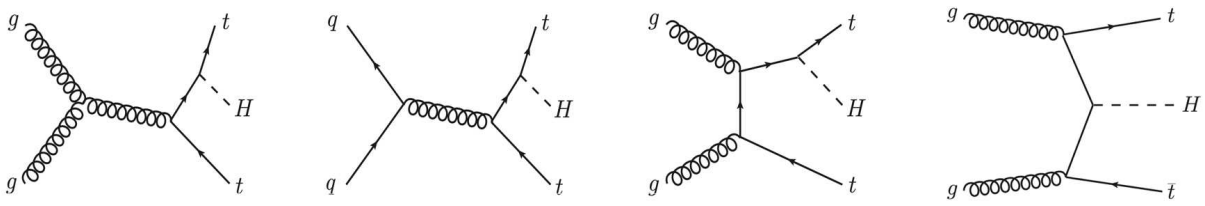


Figure 16 – Leading order Feynman diagrams of the $t\bar{t}H$ process.

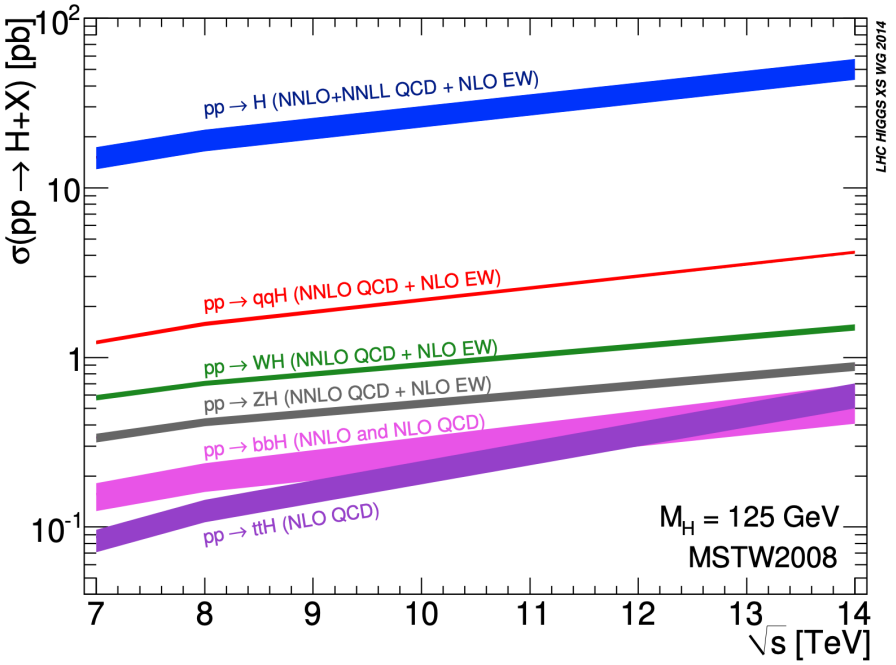


Figure 17 – Cross-section production of Higgs boson with m_H = 125 GeV as a function of the center-of-mass energy $\sqrt{s}$ [25].

Figure 18 illustrates the branching ratios of the Higgs boson decay modes as a function of its mass. With a mass of m_H = 125 Gev, the Higgs boson decays primarily to a $b\bar{b}$ pair. The branching ratio values of different decay modes with a mass of m_H = 125 Gev is shown in Figure 19. The $H \to b\bar{b}$ mode has the largest branching ratio of around 58.2%. The $H \to WW^*$ has the second highest branching ratio of 21.4%. Even though both of these decay modes have a high branching ratio, they are difficult to access due to difficulties distinguishing them from background processes. The purest decay modes are the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ decay modes. These Higgs decay modes to $\gamma\gamma$ and ZZ ($ZZ^* \to 4\ell$) decay modes have very low branching ratios of 0.2 % and 2.6 % (0.01 %) respectively. These decay channels have high purity and and offers very clean signatures so they are often referred to as the *golden channels*.

In this thesis, the $t\bar{t}H$ production is explored in $H \to b\bar{b}$ decay mode and more detailed overview and motivation for this analysis is presented in Chapter 5.

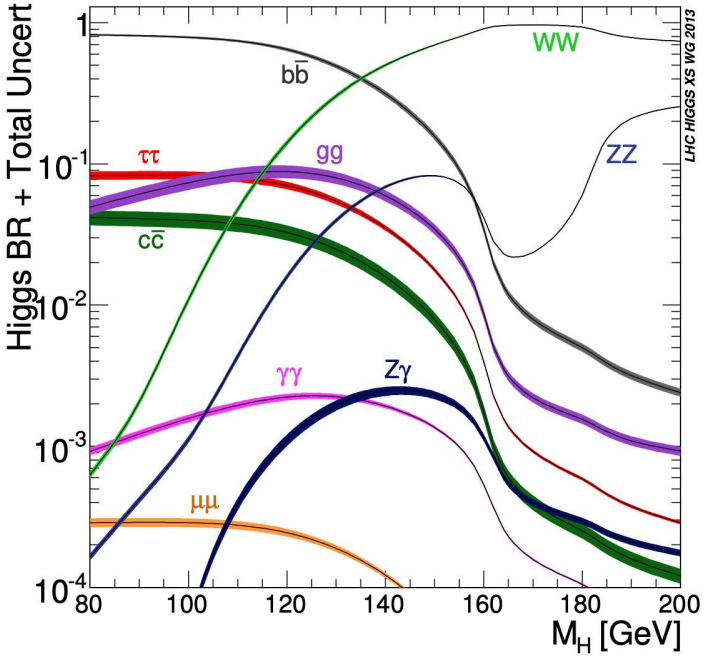


Figure 18 – The branching ratios of the various decay modes of the SM Higgs boson as a function of its mass [25].

The discovery of Higgs boson in 2012 was performed in these golden channels by exploiting the data collected during Run 1 at 7 and 8 TeV center-of-mass energies at the LHC. The local significant excess at ≈ 125 GeV was observed by ATLAS experiment as shown in Figure 20 in $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels. The significance of $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ was combined and was measured to be 5.9 σ by ATLAS [26] and 5.8 σ by CMS [27], thus exceeding the 5 σ discovery threshold. Over the next years, other main Higgs production modes in different Higgs decay channels were also observed by both ATLAS and CMS collaborations [28–30]. However, this thesis will only focus on the $t\bar{t}H(H \to b\bar{b})$ analysis, presented in Chapter 5.

1.8 Status of ttH measurements at the LHC

The $t\bar{t}H$ cross-section accounts for about 1% of the total Higgs boson production cross section. To compensate this, the $t\bar{t}H$ measurement targets as many decay modes as possible. The first observation of $t\bar{t}H$ production was reported by ATLAS [31] and CMS [32] in 2018. It was a very important achievement since the discovery of Higgs boson in 2012 as it provides a direct way to probe top-Higgs coupling, which would not be possible via Higgs decays. Different Higgs decay channels:

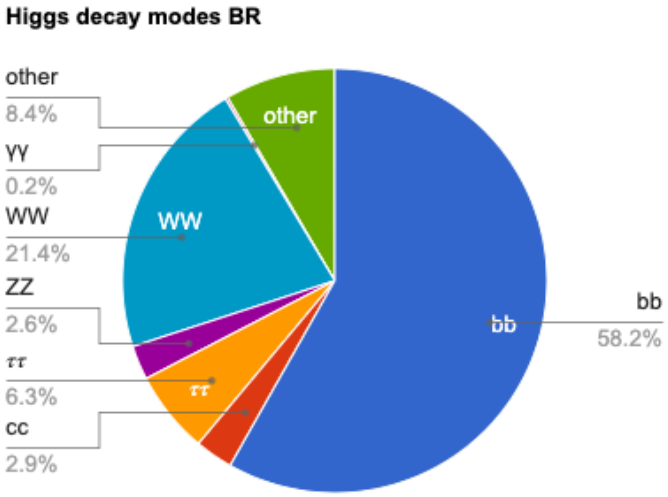
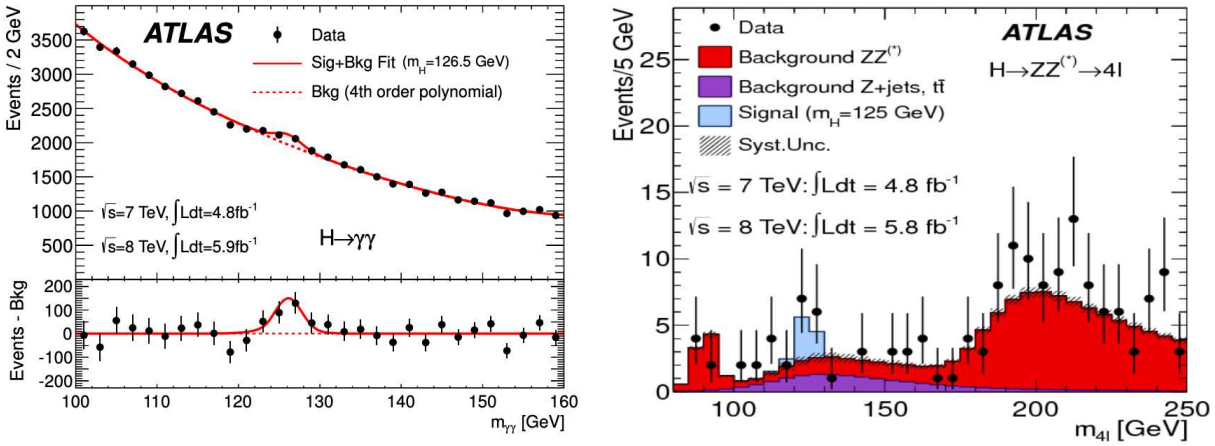


Figure 19 – Branching ratio of the Higgs boson decays with a Higgs mass of m_H = 125 GeV.



(a) (b)

Figure 20 – The invariant mass distribution of di-photon system in (a) $H \to \gamma\gamma$ and (b) $H \to ZZ^* \to 4l$ channel, compared to the background expectation for the combination of the $\sqrt{s}$ = 7 TeV and $\sqrt{s}$ = 8 TeV data-sets [21].

$H \to b\bar{b}/WW^*/\tau\tau/ZZ^*/\gamma\gamma$ were considered for the ttH measurement using Run 1 and partial Run 2 data-set. Based on their branching ratios and signal purity, each decay mode has its own set of merits and challenges to exploit. Table 3, lists the sensitivities of the individual channels for the ATLAS $t\bar{t}H$ combination. The luminosity varies among the various ttH analyses due to different levels of partial data utilization. The $H \to \gamma\gamma$ and multilepton ($H \to WW^*/\tau\tau/ZZ^*$) decay channels have the largest significance. $H \to \gamma\gamma$ decay channel is mainly dominated by statistical uncertainties. The $H \to b\bar{b}$ channel has a significantly lower sensitivity and is dominated by systematic uncertainties, primarily due to large uncertainties in the $t\bar{t}$ + $b\bar{b}$ background modelling. The overall observed (expected) significance was reported to be 6.3 (5.1) σ by ATLAS and 5.2 (4.2) σ by CMS. The signal strength, which is the ratio of the observed production rate to the SM prediction, was reported to be $1.32^{+0.28}_{-0.26}$ by the ATLAS and $1.26^{+0.31}_{-0.26}$ by the CMS. Figure 21 and Figure 22 shows the distribution of events in all analysis regions as a function of $\log_{10}(S/B)$ in ATLAS and CMS, respectively.

Analysis	Integrated luminosity (fb^{-1})	$t\bar{t}H$ cross-section (fb)	Obs. sig.	Exp. sig.
$H \to \gamma\gamma$	79.8	710^{+210}_{-190} (stat.) $^{+120}_{-90}$ (syst.)	4.1 σ	3.7 σ
$H \to$ Multilepton	36.1	790^{+150}_{-140} (stat.) $^{+150}_{-140}$ (syst.)	4.1 σ	2.8 σ
$H \to b\bar{b}$	36.1	400^{+150}_{-140} (stat.) ± 270 (syst.)	1.4 σ	1.6 σ
$H \to ZZ^* \to 4\ell$	79.8	< 900 (68% CL)	0	1.2 σ
Combined (13 TeV)	36.1 - 79.8	670 ± 90 (stat.)$^{+\ 110}_{-\ 100}$(syst.)	5.8 σ	4.9 σ
Combined (7, 8, 13 TeV)	4.5, 20.3, 36.1 - 79.8	-	6.3 σ	5.1 σ

Table 3 – List of the ATLAS measurement of the total $t\bar{t}H$ production cross sections at 13 TeV along with the observed (Obs.) and expected (Exp.) significances (sig.) relative to the background-only hypothesis. The results of both the individual and combined analyses are shown. An observed upper limit is set to 68% confidence level (CL) in the $H \to ZZ^* \to 4\ell$ decay channel using pseudo-experiments, since no event is observed in this channel [31].

1.8.1 $t\bar{t}H(H \to b\bar{b})$ measurement results

The measurement of $t\bar{t}H$ production where Higgs boson decaying into a pair of b-quarks is performed both by ATLAS and CMS. The latest $t\bar{t}H(H \to b\bar{b})$ measurement [33] by ATLAS was performed using full Run 2 data corresponding to an integrated luminosity of 139 fb^{-1} [34]. This was the first differential measurement of $t\bar{t}H(H \to b\bar{b})$, explored through Simplified Template Cross Sections (STXS) formalism, described in Section 1.8.2. The ATLAS $t\bar{t}H(H \to b\bar{b})$ analysis, discussed in detail in Chapter 5, focused on the events where at least one W-boson from either of the two top-quarks

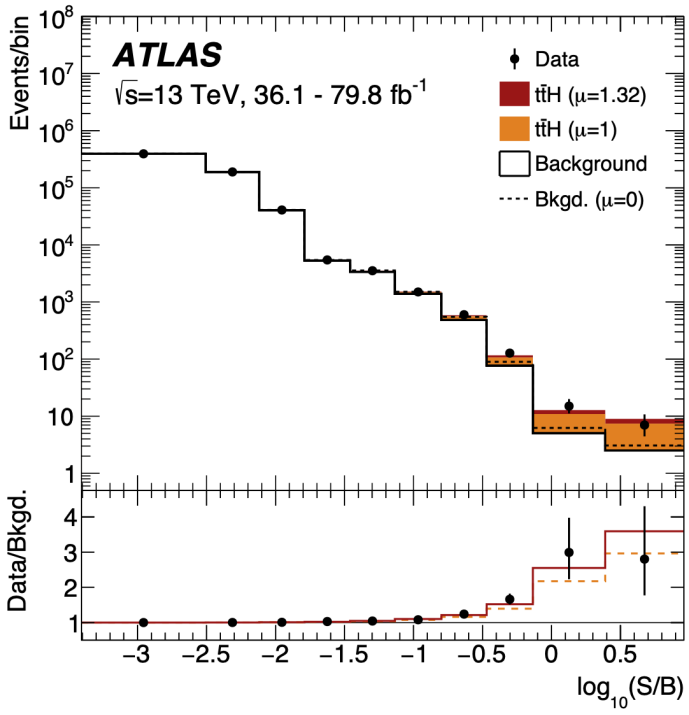


Figure 21 – Observed combined event yields with the ATLAS experiment in all analysis categories as a function of $\log_{10}(S/B)$, where S is the expected signal yield and B the background yield extracted from the fit with freely floating signal. A clear $t\bar{t}H$ signal-like excess over the background is visible for high $\log_{10}(S/B)$ [31].

decay leptonically, referred to as the single-lepton channel. It used complex event categorisation based on number of leptons, number of jets, number of b-jets at different b-tagging efficiencies and number of boosted Higgs boson candidates to better extract the information of the signal and background processes in dedicated phase space. The analysis was optimised using machine learning techniques such as Boosted Decision Trees (BDTs) for reconstruction and classification, as discussed in detail in Chapter 4. In addition to that, dedicated control regions were also defined to constrain the irreducible $t\bar{t}$ + $b\bar{b}$ background. The signal strength was measured with an observed (expected) significance of 1.0 (2.7) standard deviations and measured to be $\mu = 0.35^{+0.36}_{-0.34}$, shown in Figure 23. The measurement uncertainty was dominated by systematic uncertainties, especially from $t\bar{t}$+$\geq$1b modelling.

The latest CMS measurement of $t\bar{t}H(H \rightarrow b\bar{b})$ [35] was performed by combining two analyses, using the 2016 data-set with 35.9 fb^{-1} and 2017 data-set with 41.5 fb^{-1}. The CMS analysis cover three channels i.e single-lepton, dilepton and fully hadronic channels (see Section 5.1). The analysis employs multivariate techniques using BDTs (in dilepton channel) and Deep Neural Networks (in single-lepton channel) and a matrix element method (in hadronic channel) for the signal extraction fit. Figure 24 shows the value of the measured signal strength, $\mu = 1.15 \pm 0.15(\text{stat.})^{+0.28}_{-0.25}$ (syst.)

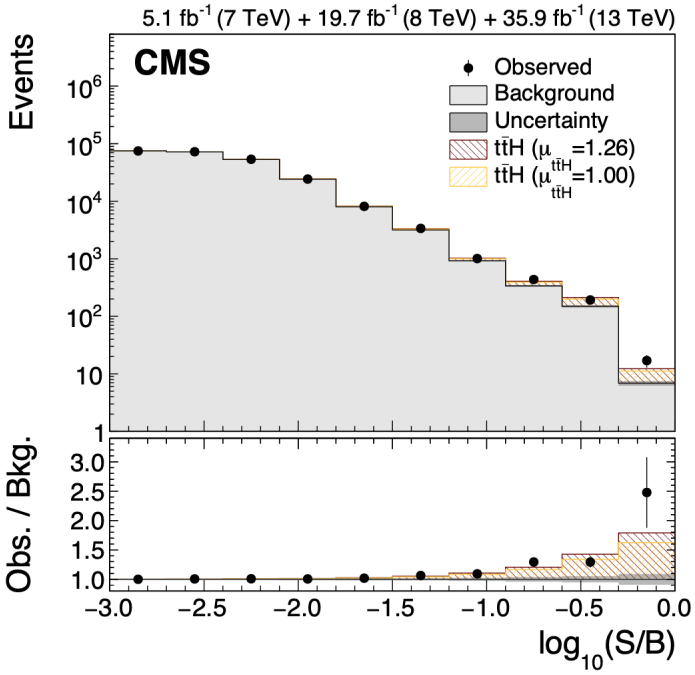


Figure 22 – Observed combined event yields with the CMS experiment in all analysis categories as a function of the $\log_{10}(S/B)$, where S is the expected post-fit signal and B is expected post-fit background yields. A clear $t\bar{t}H$ signal-like excess over the background is visible for high $\log_{10}(S/B)$ [32].

with an observed (expected) significance of 3.9 (3.5) standard deviations.

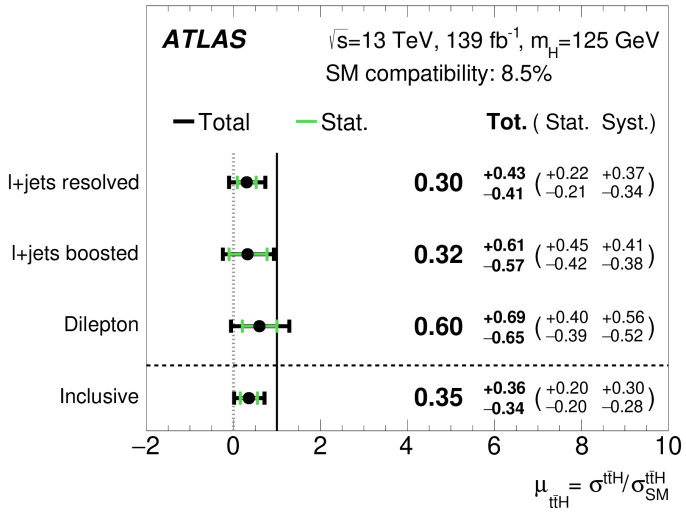


Figure 23 – Fitted values of the $t\bar{t}H$ signal-strength parameter in the individual channels and in the inclusive signal strength measurement with ATLAS experiment [33].

The systematic uncertainties of the $ttH(H \to b\bar{b})$ CMS result are considerably lower than the ATLAS result. The systematic model configuration used in ATLAS and CMS are relatively dif-

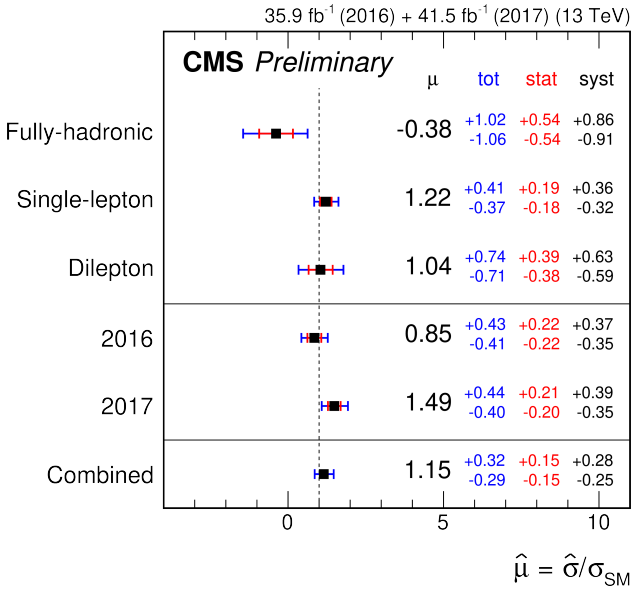


Figure 24 – Best fit values of the $t\bar{t}H$ signal-strength in the combined fit of the 2016 and 2017 data-sets per channel with CMS experiment [35].

ferent especially for $t\bar{t}$ + $\geq$1b modelling. As mentioned before, the uncertainty associated to the $t\bar{t}$ + $\geq$1b modelling is dominant in the ATLAS analysis. In CMS, the combined $t\bar{t}$ + $b\bar{b}$ and $t\bar{t}$ + $c\bar{c}$ ($t\bar{t}$ + heavy flavour) is around the same order as for the signal modelling uncertainty. This is explained by the fact that ATLAS takes into account more $t\bar{t}$ + heavy flavour modelling components than CMS. While the CMS result considers fewer shape uncertainties on $t\bar{t}$ + jets. In addition, the differences between the 4FS and 5FS (see Section 2.4.4) in the $t\bar{t}$ + $b\bar{b}$ modelling is fairly large whereas this uncertainty is not considered in the CMS measurement.

1.8.2 Simplified Template Cross-section measurements

The Simplified Template Cross-Section (STXS) formalism [25,36] has been developed as a common effort of different LHC experiments and theorists to perform signal strength measurements in different exclusive kinematic phase space regions referred to as STXS bins. It allows much easier combination of the results in different decay channels as well as different experiments. The kinematic bins in the STXS formalisation are optimised to reduce theory model dependence, that are folded into the measurements, to the greatest extent possible. A separate signal template, which presents the signal MC prediction in the targeted kinematic region at the truth level, is defined for each STXS bin. Most $t\bar{t}H$ production measurements, in prominent Higgs decay channels, allow for STXS measurements where cross-section is measured as a function of Higgs transverse momentum $\left(p_T^H\right)$. The $t\bar{t}H(H \rightarrow b\bar{b})$ process is ideal for probing differential cross-sections measurements, due to its high production rate. The precise measurements in the higher p_T regimes can also be

exploited, which is otherwise challenging in other decay channels like $t\bar{t}H(H \to \gamma\gamma)$, which lacks statistics [37]. This process also allows the kinematic reconstruction of the candidate Higgs boson, which is not possible in other Higgs decay channels, such as Higgs decay to multilepton. The STXS bins definitions: 0–120 GeV, 120–200 GeV, 200–300 GeV, 300–450 GeV and 450-∞ GeV, were used in the previous round of $t\bar{t}H(H \to b\bar{b})$ analysis using full Run 2 ATLAS experiment data. The analysis presented in this thesis, several improvements were performed of which the first STXS class (0–120 GeV) has been split into two i.e 0–60 GeV and 60–120 GeV. This was done to further improve the measurement sensitivity in the lower p_T^H regions, discussed in Chapter 5. Figure 25 shows the first differential measurement in $t\bar{t}H(H \to b\bar{b})$ channel performed in different truth p_T^H bins, labelled as $\hat{p}_T^H$, following the STXS formalism.

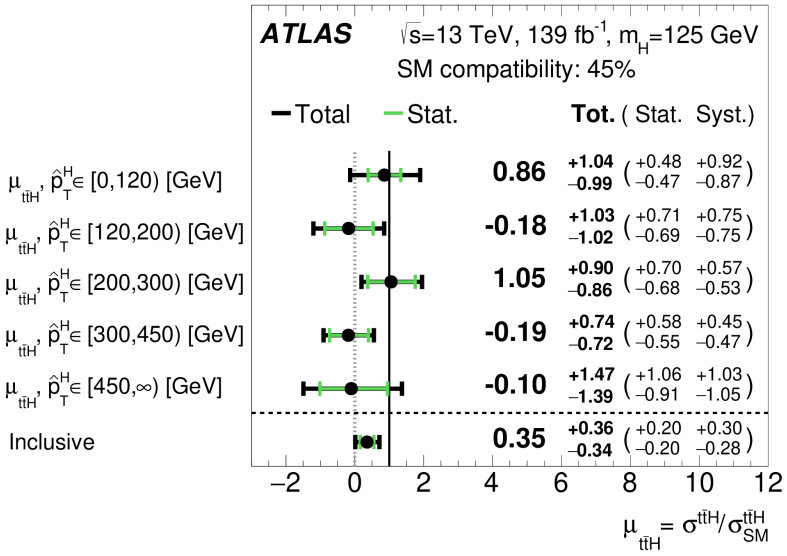


Figure 25 – Signal-strength measurements in the individual STXS truth p_T^H bins along with inclusive signal strength in $t\bar{t}H \to b\bar{b}$ channel [33].

The STXS bins defined based on p_T^H are also sensitive to the Charge-Parity (CP) structure of the Higgs boson [38] and to the Higgs self-coupling [39].

CHAPTER

2

THE ATLAS EXPERIMENT

The Large Hadron Collider (LHC) [40], discussed in Section 2.1, is the most complex circular particle accelerator ever built. The purpose of the LHC is to test the SM of particle physics while also assisting in the exploration of new physics not covered by the SM. This machine produces high-energy proton-proton (pp) collisions at a high rate, making it an ideal environment for studying the elementary particles and their interactions. The ATLAS detector, which is the largest general-purpose particle detector at the LHC is discussed in Section 2.2. The data collected by this detector is used to perform the studies, presented in this thesis. Section 2.3 gives a summary of the simulation chain set up for physics analyses. Section 2.4 describes the reconstruction and identification methods used in ATLAS detector, which targets the objects produced during the collisions. In view of the High-Luminosity upgrade of the LHC, the ATLAS detector will undergo a major upgrade, discussed in Section 2.5.

2.1 The Large Hadron Collider

The LHC consists of a 26.7 km ring, placed at a depth of around 100 m underground at the Franco-Swiss border near Geneva. It is operated by the European Organization for Nuclear Research (CERN[1]). Currently more than 6,000 users, over half of the planet's high-energy physicists, carry out fundamental research at CERN. The LHC was approved by the CERN Council in 1994 and built from 1998 to 2008 to replace the Large Electron-Positron (LEP) collider [41] decommissioned in 2000. The Large Hadron Collider (LHC) is a particle particle accelerator that collides charged particles (protons and heavy ions) with incredibly high energy and velocity. Bunches of approximately 10^{11} protons collide with a gap of 25 ns between two consecutive LHC bunches, known as *bunch-spacing*.

[1] CERN is the acronym of the organization's name in French, Conseil Européen pour la Recherche Nucléaire

The collisions occur at specific points where particle detectors are installed. This makes the LHC a key tool for modern particle physics research, allowing researchers to investigate not only numerous Standard Model (SM) processes, but also rare processes and the new physics scenarios that may arise at high energies. The first beam was circulated in 2008, and the first operational run (Run 1) began in March 2010 with successful collisions of proton beams at a center-of-mass energy of 7 TeV, and lasted until early 2013 with the collision energy boosted successfully to 8 TeV in 2012. The LHC ring and detectors were upgraded from 2013 to 2015 to operate at higher center-of-mass energy and luminosity. The Run 2 phase, which lasted from mid-2015 to the end of 2018, had a 13 TeV energy. Between 2019 and 2022, the LHC and the detectors has undergone another upgrade (LS2), now followed by the Run 3 where a 13.6 TeV energy is recorded. The LHC will be upgraded to the High Luminosity-LHC (HL-LHC) and Run 4 is planned to start in 2029 (see Section 2.5).

2.1.1 The accelerator design

The LHC is mainly designed to accelerate protons, as they are easy to obtain, stable and have low energy loss through synchrotron radiation compared to electrons or positrons. The protons used for collisions at the LHC are obtained by ionizing hydrogen gas using a strong electric field. These protons undergo consecutive accelerating stages before reaching the LHC which is the final element in the CERN accelerator chain depicted in Figure 26.

Linear accelerator 4 (Linac4) [42] accelerates the protons to an energy of 160 MeV which are then passed through three successive circular accelerators: proton synchrotron BOOSTER, Proton Synchrotron (PS) and Super Proton Synchrotron (SPS), increasing their energy up to around 1.4 GeV , 25 GeV and 450 GeV respectively. Finally, the protons are transferred to the two beam pipes of the LHC, where they circulate in two opposite-direction beams. After around 20 minutes, the proton beams at the LHC are stabilized and given their final shape while their energy is increased up to 6.8 TeV in Run 3.

The LHC ring is made up of eight straight sections and arcs. The radiofrequency (RF) cavities in the LHC, which are placed in the straight sections, are used to accelerate the particles. The protons are accelerated to the necessary energy by an electric field of 5 MV/m oscillating at 400 MHz generated by the RF cavities. The proton beams inside the LHC are maintained in a curved trajectory by a variety of superconducting magnets that are cooled to a temperature of 1.9 K using superfluid Helium. Two primary types of magnets are used: dipoles, which bend proton beams in arcs, and quadrupoles, which focus the beams and keep the protons away from the pipe walls. A magnetic field of 8.3 T is generated by 1232 dipole magnets, each measuring roughly 15 m in length. A cross-section of an LHC dipole element is shown in the Figure 27. It consists of two beam pipes around which two opposite-sign magnetic fields are formed to curve the trajectory of the counter-

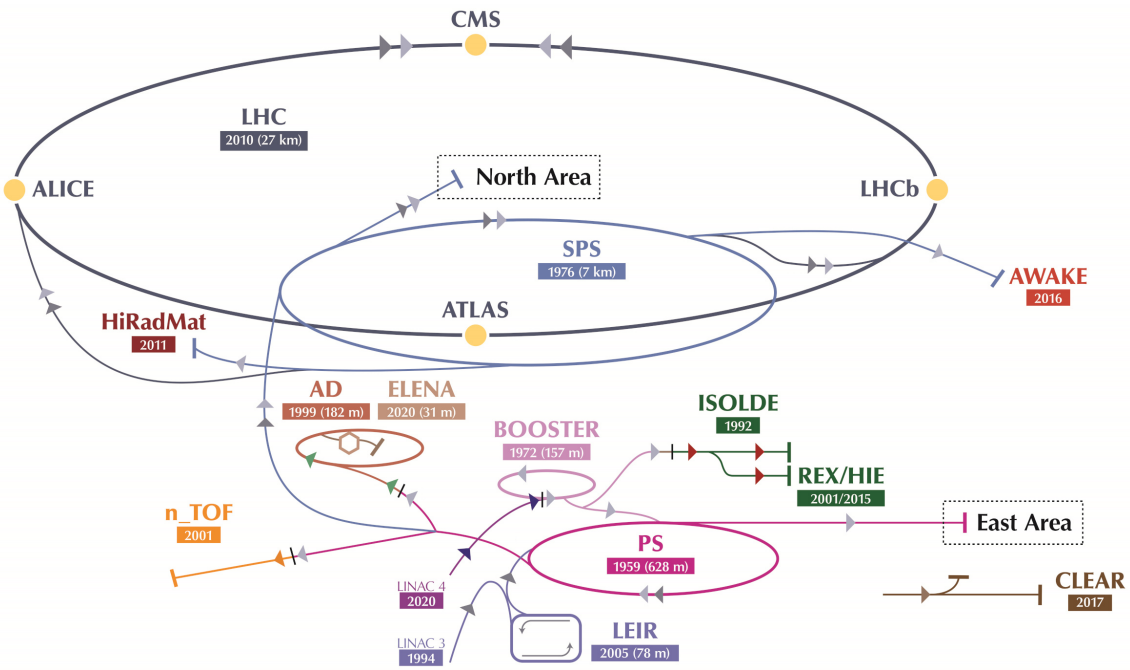


Figure 26 – Schematic view of the CERN hadron accelerator chain [43]. The LHC is the last element of this system, in which the beams reach their highest energies.

rotating proton beams. There are 858 quadrupoles in all, with nominal gradients of 223 T/m and 241 T/m. The breadth and height of the beams are controlled by these magnets, which are positioned in pairs along the LHC ring. In the LHC, additional magnetic multipoles (sextupoles, octupoles, and decapoles) are used to counteract divergence-inducing effects that result in proton beam defects.

The instantaneous luminosity $\mathcal{L}$ can be used to describe the collision rate at a collider, which is an important value for characterizing the performance of the accelerator. It is defined as:

$$\mathcal{L} = \frac{N_p^2 n_b f_{rev}}{4\pi\sigma_x\sigma_y} F \tag{2.1}$$

where N_p is the number of protons per bunch, n_b the number of bunches per beam, f_{rev} the beam revolution frequency (around 11 kHz), $\sigma_{x(y)}$ the beam dispersion in the transverse plane corresponding to the size of the beamspot and F a geometric correction factor encoding the effect of the beams crossing angle at the interaction point. The peak instantaneous luminosity achieved in Run 2 was around 2 x 10^{34} $cm^{-2}s^{-1}$. The integrated luminosity given by $\mathscr{L} = \int \mathcal{L} dt$, is a measurement of the collected data size. The ATLAS Run 2 data sample corresponded to an integrated luminosity of 139 fb^{-1} is used to perform $t\bar{t}H(H \to b\bar{b})$ analyses, presented in Chapter 5.

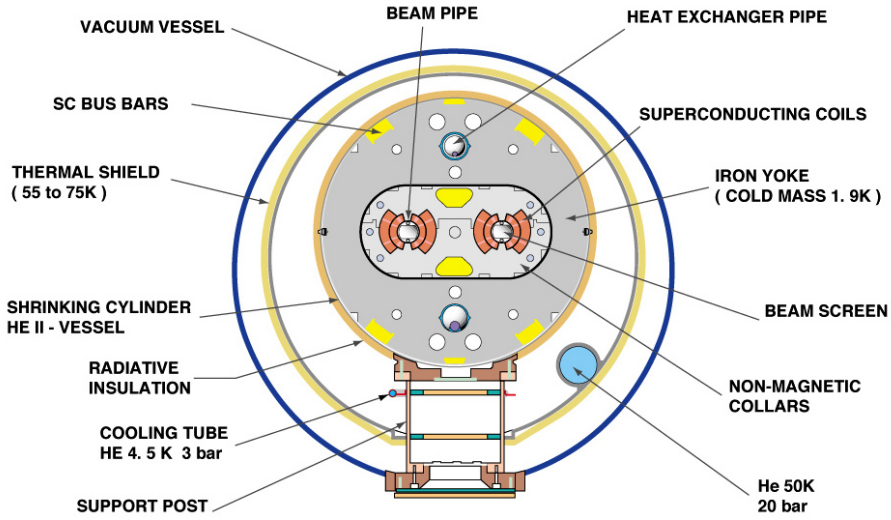


Figure 27 – The cross section of the LHC main dipole [44].

During a data collection period, the total number of events collected for a certain process considered by an analysis, via

$$N = \mathscr{L}\sigma\varepsilon \tag{2.2}$$

where σ is the process cross-section and ε is the event selection efficiency convoluting trigger, offline reconstruction and analysis phase space selection efficiencies. The large number of protons contained inside a bunch can lead to multiple proton-proton collisions per bunch crossing, called *pile-up*. The mean number of interactions per crossing is denoted as $<\mu>$. The hard-scatter interactions are of interest for the physics analyses and the pile-up effects needs to be suppressed. The Run 2 operation at the LHC, gives an average of $<\mu> = 33.7$ for the ATLAS experiment in the years from 2015 to 2018, as shown in Figure 28.

2.1.2 The LHC experiments

The four key experiments at the LHC are discussed here:

- The ATLAS (A Toroidal LHC ApparatuS) experiment [46]: It is a multipurpose experiment intended to detect the majority of pp collision products over a wide energy range. The ATLAS detector is discussed in detail in Section 2.2.

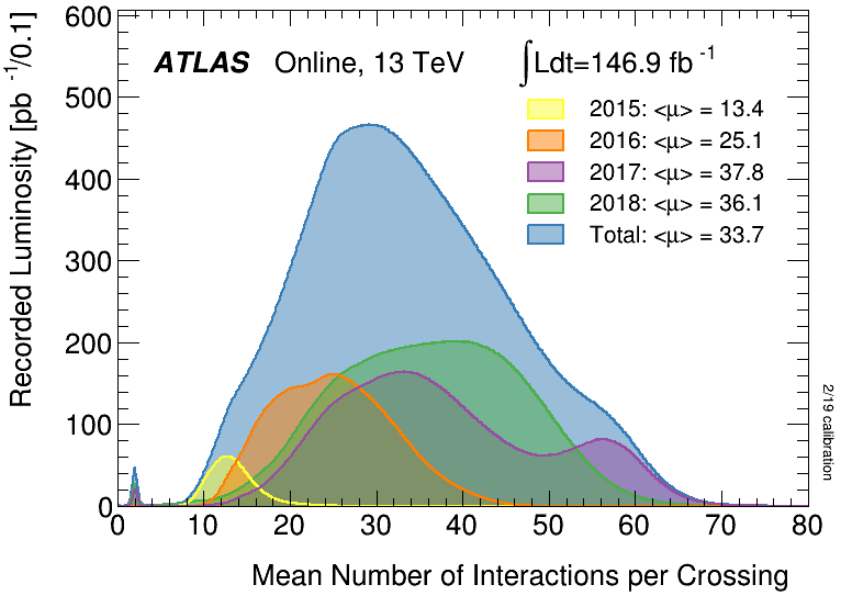


Figure 28 – Luminosity-weighted distribution of the mean number of interactions per crossing $< \mu >$ for the Run 2 pp collision data at $\sqrt{s}$ = 13 TeV [45].

- The CMS (Compact Muon Solenoid) experiment [47]: It is a general-purpose experiment with a physics program similar to ATLAS, although it employs different detector technologies. It is advantageous to have two independent experiments targeting the same physics analyses since it allows them to cross-check their results and increase the precision.

- The LHCb (LHC beauty) experiment [48]: It is an experiment dedicated to heavy flavour physics at the LHC. Its main goal is to look for indirect evidence of new physics via precision measurements of CP violation and rare decays of beauty and charm hadrons. LHCb operates at a lower luminosity compared to ATLAS and CMS (around 10^{32} cm^{-2}s^{-1}).

- The ALICE (A Large Ion Collider Experiment) experiment [49]: It is a heavy-ion experiment that studies the physics of strongly interacting matter and quark-gluon plasma at high energy density and temperature.

Besides the above main detectors, there are other small experiments [50–53], installed around the LHC designed to perform dedicated physics analyses.

2.2 The ATLAS detector

The ATLAS detector [46] is a general-purpose detector located at one of the LHC interaction points, to collect the data used in the study presented in this thesis. It measures 25 m in height, 44 m in

length, and weighs approximately 7000 tons. Figure 29 shows a computer-generated view of the ATLAS detector. It has a cylindrical structure made up of several sub-detector layers and offers nearly full solid angle coverage of 4π. The detector consists of a Inner Detector (ID) surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and an outer Muon Spectrometer (MS) situated on three large superconducting air-core toroidal magnets. These are designed to achieve high granularity, fast electronics readout, efficient object reconstruction and resolution.

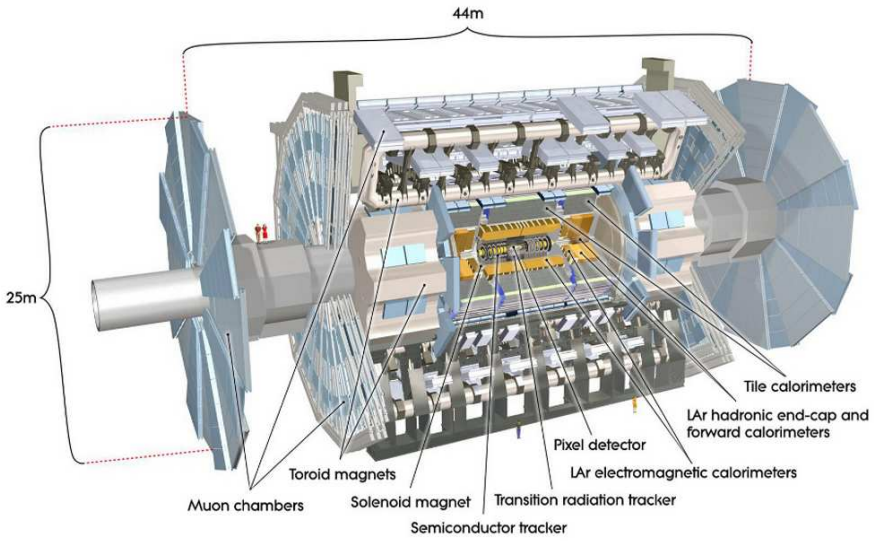


Figure 29 – A schematic view of the full ATLAS detector [54].

2.2.1 The ATLAS coordinate system

Throughout ATLAS, a right-handed coordinate system is used, with the origin defined at the nominal interaction point. Figure 30 shows the sketch of this coordinate system where the longitudinal z-axis runs parallel to the beam direction, whereas the transverse xy-plane runs perpendicular to it. The positive x-axis points from the origin to the center of the LHC ring, the positive y-axis points to the ground surface. A spherical polar coordinate system is also commonly used, where the polar angle θ is measured from the positive z-axis, the azimuthal angle ϕ is measured from the positive x-axis, and the radial coordinate $\mathrm{R} = \sqrt{\mathrm{x}^2 + \mathrm{y}^2}$ is the measured distance from the beamline.

The momentum of a collision product in the transverse direction is referred to as p_T, where p_T = p $\sin\theta$. The rapidity y is defined as:

$$y = \frac{1}{2} - \ln\left(\frac{E + p_z}{E - p_z}\right) \tag{2.3}$$

where, E is the energy and p_z is the z component of the momentum vector. When the mass of the particle is negligible with respect to its energy, as if often true for particles at the LHC, the rapidity reduces to the pseudorapidity (η):

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{2.4}$$

Using pseudorapidity is beneficial as its difference $\Delta\eta$ is invariant under Lorentz boost along the beam axis. The distance of two points in the η-ϕ plane is given by:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{2.5}$$

where $\Delta\eta$ and $\Delta\phi$ are differences of η and ϕ, respectively.

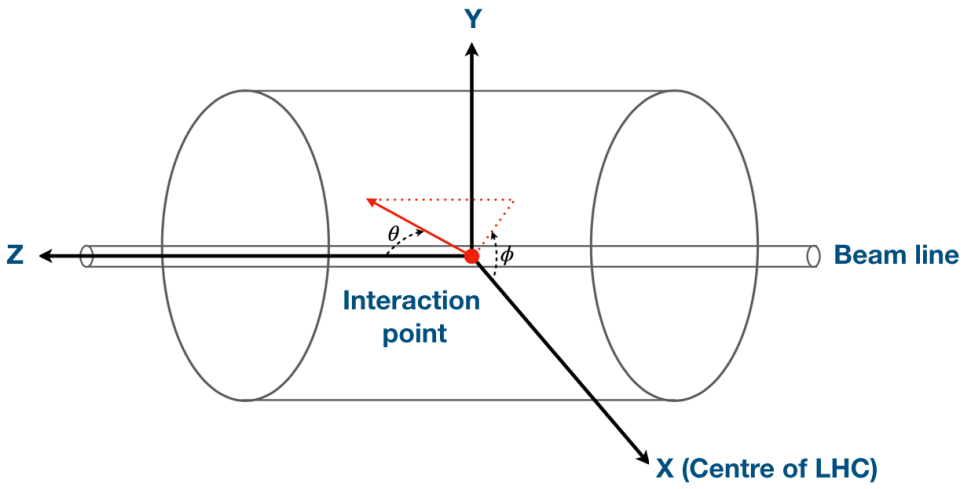


Figure 30 – A sketch of the ATLAS coordinate System.

2.2.2 Inner Detector

The Inner Detector (ID) [55, 56] is the innermost component of the ATLAS detector, shown in Figure 31. The ID is crucial in measuring the momentum of charged particles using the 2T solenoid magnetic field and reconstructing the primary vertex. The highly energetic particles usually do not get absorbed in this part of the detector. The path of charged particles described by an object is known as a *track*. The precise tracking of particles is required for momentum measurements, particle and vertex identification, and physics object reconstruction.

The ID has three sub-components: the Pixel Detector, the Semi-Conductor Tracker (SCT), and Transition Radiation Tracker (TRT). These three sub-components deliver precise measurements and tracking within the pseudorapidity range $|\eta| < 2.5$.

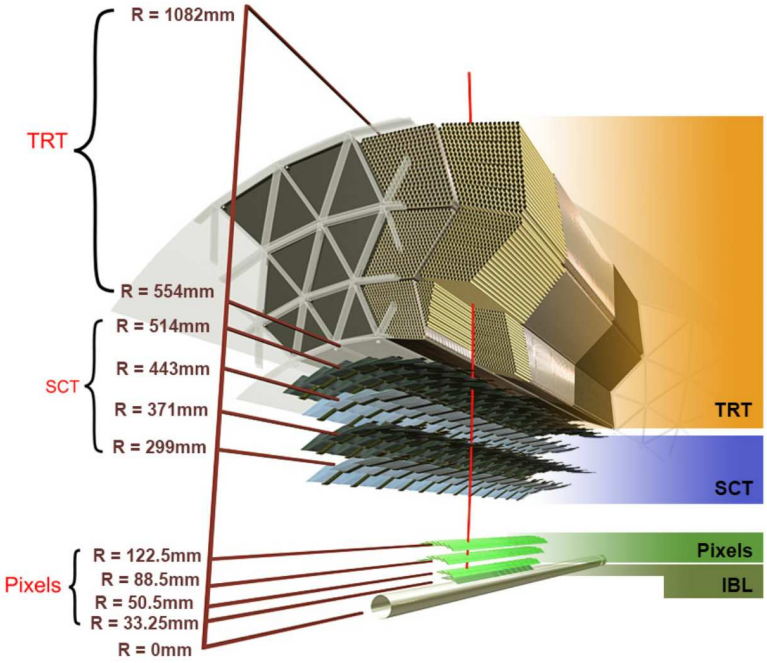


Figure 31 – Overview of the of ATLAS Inner Detector, showing the three primary sub-detectors: Pixels (inclusing IBL), SCT, and TRT [57].

Pixel detector:

The Pixel Detector [58] is at the innermost part of the ID. By providing the tracking information so close to the interaction point, accurate track reconstruction can be achieved. The innermost layer is called IBL [59], and was installed in-between Run 1 and Run 2 at $R = 33.25\ mm$ in the barrel region. On the outer side, three pixel layers in the barrel region are installed at R = 50.5, 88.5, 122.5 *mm*. It provides 3D measurements with an excellent resolution of 10 μm in r-ϕ and 115 μm in z direction. The resolution improves even more with the smaller pixels of the IBL: 8 μm in

r-ϕ and 40 μm in z direction. This enables accurate reconstruction of charged tracks, interaction vertices, and secondary vertices, which is crucial for b-tagging (Chapter 3). There are six pixel end-cap layers, three on either side of the interaction points. The pixel detector also includes pixel sensors that works through the creation of electron-hole pairs when a particle passes through them. In total, the pixel detector contains 92 M readout channels, which accounts for 50% of the total readout channels of the ATLAS detector.

Semi-Conductor Tracker:

The SCT [60] works in a similar way to the pixel detector but is made up of narrow and long strips rather than small pixels. The strips cover up to $|\eta| < 2.5$ and has a size of 12 cm in length and 80 μm in width. It comprises four barrel layers and 2 end-caps with nine disks each. The SCT system achieves a hit resolution of 17 μm in r-phi and 580 μm in z direction. The SCT is crucial for tracking in the plane perpendicular to the beam and it measures particles across a significantly broader area than the Pixel Detector. It is composed of four double layers of silicon strips, with a total of 6.3 M readout channels.

Transition Radiation Tracker:

The last and outermost component of the ID is the Transition Radiation Tracker (TRT) which has an effective range of $|\eta| < 2.0$. The TRT is a gaseous detector comprised of straw tubes with a diameter of 4 mm and with a maximum length of 150 cm. The straw tubes contain gold-plated tungsten wires, and are filled with a gas mixture of Xe (70 %), CO_2 (27 %) and O_2 (3 %). In Run 2, Xe was replaced with argon for the straw tubes belonging to modules with large gas leakage [61]. In the barrel and end-cap sections, the hit resolution is 120 μm and 130 μm, respectively. Particles that pass through the straws of the TRT ionize the gas within the tube resulting in a current along the wires leading to a detection. The size of the TRT active area allows for many measurements throughout the path of charged particle, improving momentum measurements. TRT has a total of 350,000 read-out channels. Electrons can be identified by their larger in ionization deposited in each straw in comparison to other particles.

2.2.3 Calorimeters

The ATLAS calorimeter system is located outside of the ID with $|\eta| < 4.9$ and full ϕ coverage around the beam axis. Figure 32 shows the design of entire calorimeter system. It is capable of precisely measuring the particle energy by absorbing them, as well as measuring shower parameters to allow particle identification. It is composed of electromagnetic and hadronic calorimeters . Both are sampling calorimeters, which absorb energy in high-density metal and periodically sample the shape of the subsequent particle shower, inferring the energy deposit of the particles.

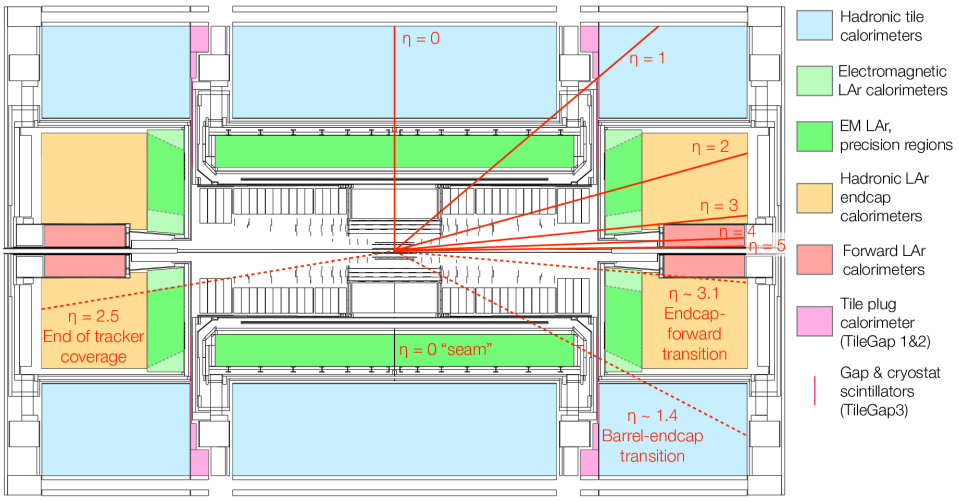


Figure 32 – Longitudinal cut of the calorimeter system showing the Liquid Argon and tile calorimeters [62, 63].

Electromagnetic calorimeter:

The Electromagnetic calorimeter (ECal), collects energy from electromagnetically interacting particles such as charged particles and photons. It has two half-barrel sections separated at the center of the detector and two end-cap wheels. The barrel sections cover the region $|\eta| < 1.475$ and the end-caps extend over $1.375 < |\eta| < 3.2$. It is a high granularity sampling calorimeter that is composed out of layers of lead plate absorbers that are interleaved with Kapton electrodes and immersed in a liquid argon (LAr) active material. An accordion-shaped geometry was adopted for the absorbers and electrodes to ensure a full coverage in ϕ without gaps. The end-cap wheels sit on both sides of the barrel and are made up of an inner and outer coaxial wheel.

Hadronic calorimeter:

The Hadronic calorimeter (HCal) is located outside the ECal, and is composed of three different components. First, the tile calorimeter extends over $|\eta| < 1.7$. It is made up of three sections, the central one with a range from $|\eta| < 1.0$ and two outer sections over $0.8 < |\eta| < 1.7$. The tile calorimeter are made up of steel and plastic and measure the energy deposition of charged and neutral hadrons. The two sides of the tiles have wavelength shifting fibers that direct the photons from the scintillator into photomultiplier tubes. Second, the end-cap calorimeters, which are based on the the LAr technology, are directly outside the EM calorimeter, covering the range $1.5 < |\eta| < 3.2$. Finally the Forward calorimeter (FCal) finishes out the LAr system in the range $3.1 < |\eta| < 4.9$. The FCal has three modules, one made of copper designed for EM measurements and two made of tungsten designed for hadronic measurements. The modules consist of longitudinal

channels parallel to the beam axis. These longitudinal channels contain concentric rods and tubes separated by a LAr gap, and the metal of the modules serves as the sensitive medium.

2.2.4 Muon Spectrometer

The Muon Spectrometer (MS) [64], shown in Figure 33 is placed at the outermost part of the ATLAS since muons punch through the calorimeters due to their long lifetime and small energy loss. The MS is made up of precision tracking concentric cylindrical shell chambers in the barrel region at radii of 5 m, 7.5 m, and 10 m. In the two end-cap regions the muon chambers are made up of large wheels at |z| $\sim$ 7.5 m, 10.8 m, 14 m, and 21.5 m all perpendicular to the z-axis. The forward muon-tracking region (called the muon Small Wheel) has been upgraded with two New Small Wheels (NSW) [65] for Run 3.The NSW will have two chamber technologies, one for the Level-1 trigger function and one for precision tracking. The powerful toroidal magnetic field produced by the magnetic system, covered in the next section deflects muons exiting the calorimetry system and the MS tracks provide precision measurements of their momentum in complement to the ID. The MS is made up of four different kind of chambers that serve either precision tracking or triggering functionality. The two high precision chambers are Monitored Drift Tubes (MDT) and Cathode Strip Chambers (CSC). The MDT chambers are installed in the barrel and end-cap region covering covering $|\eta| < 2.7$ region, except for the part where CSC chambers are installed covering $2 < |\eta| < 2.7$ region. It delivers spatial resolution of 80 μm per tube and a resolution of 35 μm per chamber. The other tracking chamber, the CSC, operates from $2.0 < |\eta| < 2.7$. Each chamber contains four planes which provides four independent η, ϕ measurements. It delivers a spatial resolution of 40 μm in the η-plane and 5 mm in the ϕ-plane. The two trigger systems work together to cover the range $|\eta| < 2.4$ with the Resistive Plate Chambers (RPC) operating at $|\eta| < 1.05$. Thin Gap Chambers (TGC) covering the remainder of the range. Both these trigger chambers primarily provide bunch crossing identification and coarse muon tracking information.

2.2.5 Magnet system

The magnet system [67] of the ATLAS detector is designed to generate a magnetic field that allows for precise measurement of charged particle momentum. The magnet system is made up of four superconducting magnets: the central solenoid, the barrel toroid, and the two end-cap toroids. The four components produce different magnetic fields with the central solenoid having a peak field strength of 2.6 T and the barrel and end-cap toroids having 3.9 T and 4.1 T strengths respectively. The central solenoid is aligned on the beam axis in order to provide an axial magnetic field for ID.

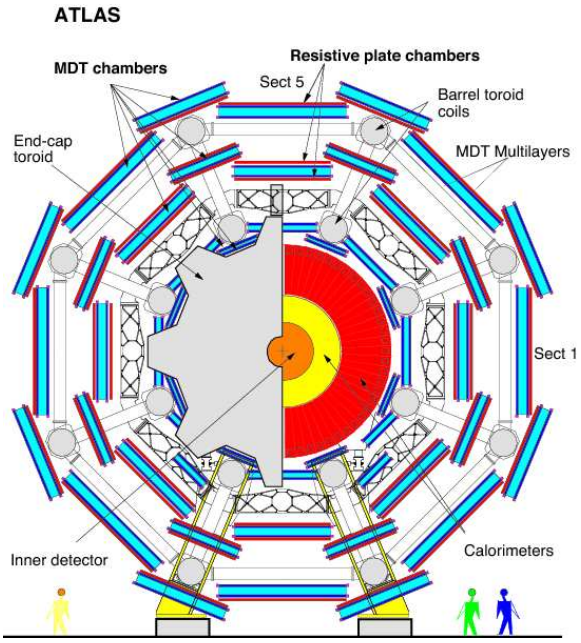


Figure 33 – Transverse cut view of the central muon system [66].

2.2.6 Trigger system and Data Acquisition

Around 40 million bunch crossings per second takes place at the LHC, which creates a massive data acquisition and storage challenge. The vast majority of these collisions involve only minor scattering of protons and are thus less important to the physics analyses. At a size of around 1.5 MB per event there is no feasible way to store all this data for later analysis anyway. In addition to this, the read-out system is not capable to extract data at 40 MHz and the CPU time needed to reconstruct 40 MHz of events is not manageable. To solve this, ATLAS uses a trigger [68] and data acquisition system, shown in Figure 34, to filter the data in quasi real-time and decide which events to keep. The first level of filtering is accomplished with a system of custom hardware electronics called the Level-1 hardware trigger. The Level 1 trigger uses information from the calorimeter system and the MS to decide whether or not to keep an event. The incoming data rate of 40 MHz is reduced to 100 kHz before reaching the the high level trigger (HLT). The HLT is a software based trigger running on a computing system where reconstruction algorithms are run on either full event readout data, or by looking at particular isolated regions of the detector. This software-based trigger reduces the L1 rate of 100 kHz to an average output rate of 1.5 kHz. The data rate is reduced by a factor of about 40,000 in total, allowing the full record of all remaining events to be written to permanent storage. Since 2015, ATLAS has continued to record increasing luminosities every year, necessitating an increase in the trigger thresholds to keep the rate within limits. However, several improvements to the trigger have been made in order to avoid raising the thresholds as much as possible.

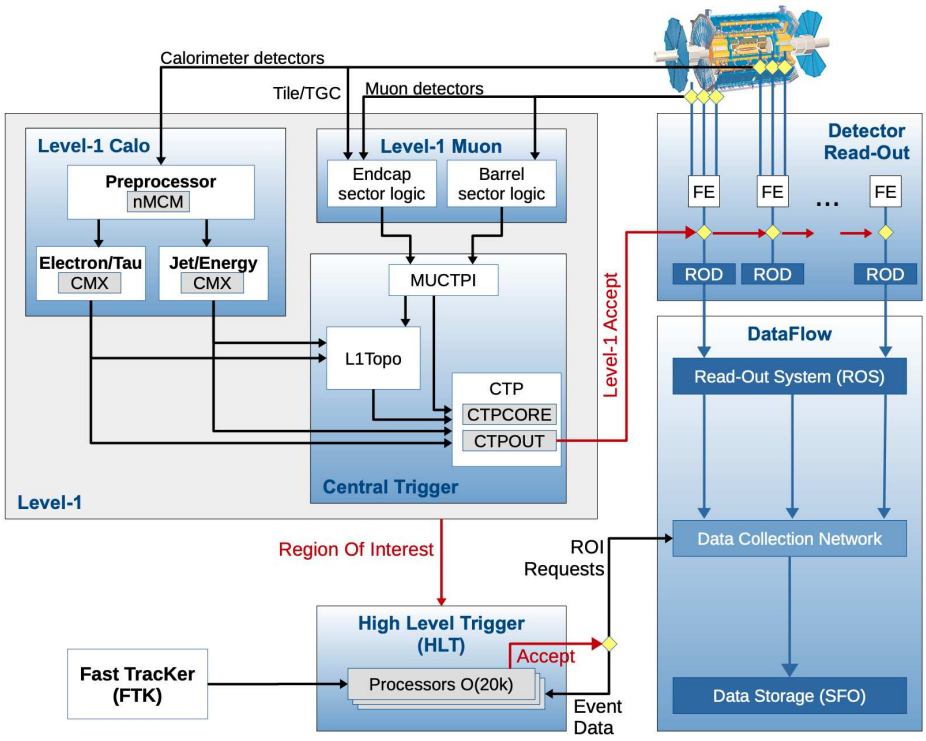


Figure 34 – Schematic view of the ATLAS trigger and data acquisition system in Run 2 [68].

2.3 Simulating Collision Events

The connection between the theory of SM, described in Chapter 1.1 and the data collected by the ATLAS experiment hinges on the ability to accurately predict the rate and properties of physics events. The simulation of pp collision events, as outlined in Section 1.6, is required to perform ATLAS physics analyses.

Event Generation:

The event generation is split into two parts: the matrix element (ME) generation, which describes the hard scattering, and the parton shower (PS) evolution and hadronisation modelling, which includes initial state radiation (ISR) and final state radiation (FSR). The factorization of pp collisions and the event simulation using Parton Distribution Functions (PDFs) are covered in Section 1.6. The generation of simulated events corresponding to various physics processes begins with producing what is known as event-level Monte-Carlo (MC). This is accomplished by using a generator to calculate the ME. The ME generators used in the studies presented in this thesis are POWHEGBOX [69], GRAPH5_aMC@NLO [70] and SHERPA [71]. These ME generators are interfaced with the parton shower generators: PYTHIA8 [72] and HERWIG7 [73]. On the other hand, SHERPA has its own showering and hadronization models and thus does not require interfacing. In addition, the param-

eters of the models used to describe non-perturbative processes can be adjusted using collision data. The ATLAS experiment commonly uses A14 parameters [20] for PYTHIA8 or the H7UE set of tuned parameters [74] for HERWIG7. The output produced via the event generator sequence is known as the MC history. It is a list of four-vectors of all the particles produced in the event after hadronization and decay of the majority of the intermediate unstable particles. The particles in the MC history record are commonly known as *truth particles*.

Detector Simulation:

The simulated MC event depicts the physics that occurs during a collision, but it does not represent the real signal produced by the ATLAS detector. The simulation of the detector is performed using the GEometry ANd Tracking (GEANT4) toolkit [75]. GEANT4 simulates an event by propagating the particles from the event through the detector and evaluating the interactions, energy depositions, and momentum transfers that occur as the particles interact with the different materials. The detector simulation phase consumes the majority of the computational power required for the entire event simulation. This is primarily due to the formation of electromagnetic PS in the ATLAS calorimeter system, which takes the longest to simulate. As an illustration, a single $t\bar{t}$t event simulation using a highly detailed detector description, known as the full detector simulation (FULLSIM), requires approximately 15 mn of CPU time [76]. To reduce the amount of CPU time required to process an event, a faster and less refined alternative simulation known as ATLFAST-II (AFII) [76] was implemented. This is accomplished primarily by using pre-simulated electromagnetic showers of low energy particles rather than developing them from scratch. FULLSIM samples have higher precision and are used as the nonimal samples for physics analyses. AFII samples are used for optimization studies and to evaluate theoretical systematic uncertainties.

Digitization:

When collision data is collected, the detector electronics responds to the physical objects, resulting in digital signals. Digitization is accomplished by converting the simulated energy deposits from the GEANT 4 toolkit into electrical responses. Noise is introduced to better reflect the hardware setup and finally the data readout structure of the ATLAS electronics is recreated. The HLT and reconstruction are then applied to the simulated outputs in the same way that they are applied to experimental data.

The final step, known as reconstruction, is performed and is common to both simulated and real data events.

2.4 Object reconstruction and identification in ATLAS

The ATLAS experiment employs a number of specialised algorithms designed to determine and reconstruct the properties of the various physics objects present in an event. These objects are jets, muons, photons, electrons, taus, and missing transverse energy. Figure 35, depicts the path that different particles take through the detector. The charged particles leave a track in the ID, electrons and photons shower in the ECal and hadrons shower in the HCal. The $t\bar{t}H(H \rightarrow b\bar{b})$ analysis, presented in Chapter 5, makes use of electrons, muons, jets, and missing transverse momentum and are discussed in Section 2.4.1, 2.4.2, 2.4.3, 2.4.4 and 2.4.5, respectively.

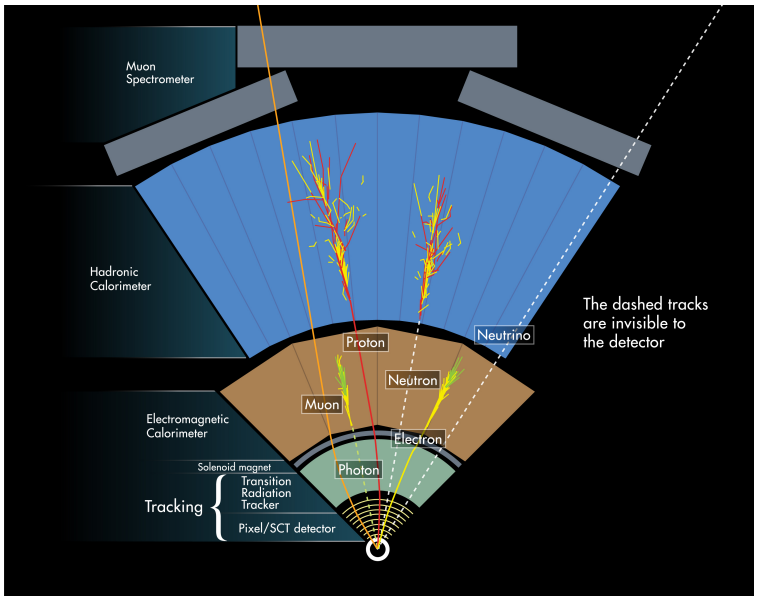


Figure 35 – Diagram of particle paths in the ATLAS detector [77].

2.4.1 Tracks

Tracking algorithm:

Tracks are the reconstructed trajectories of charged particles in the ATLAS ID.A charged particle hitting a sensor of the ID are first translated into a space-point called hit. The reconstruction of the tracks [78] is based on the combination of several algorithms [71]. In the first step, the track finding employs an inside-out algorithm in which track seeds are initially constructed from the hits in the pixel and SCT detectors. These seeds are then extended by iteratively adding hits from the subsequent SCT layers to form a track candidate. This is performed by a combinatorial Kalman filter [79]. At this stage, several track candidates are obtained per seed if there are multiple compatible track extrapolations on the same layer. A procedure for resolving ambiguity is used to identify the best track candidates and reduce the number of *fake tracks*. The procedure involves a

score-based ranking scheme for tracks which takes into account several inputs. The number and type of hits on the track are important factors considered when ranking. When a hit is associated to one track, a bonus is given to the score of this track. Sometimes a hit can also be associated to several tracks and is called *shared-hit*. This could be due to the presence of a fake track or because the detector granularity is insufficient to resolve nearby particles. A neural network is trained to distinguish between these two cases [80]. The tracks which pass these quality requirements are then extended towards the TRT, where further hits are added. The final stage of track reconstruction is based on outside-in algorithm. The algorithm takes into account hits that were not chosen by the inside-out algorithm. It starts by finding track segments in the TRT and then extrapolates them back to the silicon detectors. This strategy is well suited for the tracks stemming from long-lived particles.

Vertexing algorithm:

A vertex is the point of origin of tracks. The primary vertex (PV) is of particular interest because it represents the HS interaction of partons in colliding protons. The PV within an event are reconstructed by applying an iterative vertex finding algorithm on tracks with $p_T > 400$ MeV. To find a vertex seed, the algorithm first searches for the global maximum in the distribution of the z-positions of the selected tracks. In the next step, the optimal vertex position is determined and the associated uncertainties are estimated. To achieve this, the seed position and nearby tracks are fed into the χ^2 fitting algorithm as input, which iteratively removes tracks that are incompatible with the vertex. Tracks which are displaced by more than 7σ are used as a new vertex seed and the procedure is repeated until no more seeds are found in the event. The vertex with the largest $\sum p_T^2$ of the associated tracks forms the PV, while the remaining vertices are considered to be *pile-up* interactions.

Track parameters:

After the tracks are reconstructed, a fit is performed to estimate the track parameters. The track candidates are defined by five variables, depicted in Figure 36:

- Transverse impact parameter (d_0): It is the distance of closest approach of the track to the PV in the transverse plane.
- Longitudinal impact parameter (z_0): It is the distance between the PV and the point of closest approach of the track to it, along the beam axis.
- Azimuthal angle (ϕ): It is the angle between the transverse momentum vector $\overrightarrow{p_T}$ and the x-axis in the transverse plane.
- Polar angle θ: It is the angle between the momentum vector $\overrightarrow{p}$ and the z-axis.

- Electric charge (e) over the transverse momentum (p_T): The ratio is calculated using the magnetic field strength B and the track curvature R_{curv}.

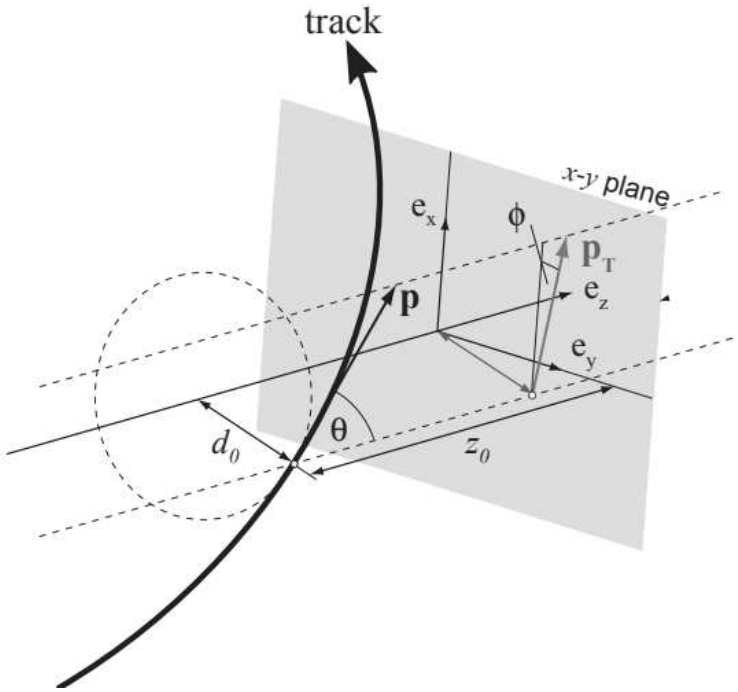


Figure 36 – A schematic diagram of a track depicting the track impact parameters d_0 and z_0.

2.4.2 Electrons

The reconstruction of electrons [81] is performed using the ID and calorimeter system information. The electron, being the charged particle, leaves a track in the ID and is then absorbed in the electromagnetic calorimeter, where it leaves an electromagnetic shower.

Reconstruction:

The electrons are reconstructed in the region $|\eta| < 2.47$ excluding $1.37 < |\eta| < 1.52$ region, which is the transition region of barrel and end-cap. First, electron candidates are selected by matching reconstructed tracks to clusters formed by the energy deposited in the electromagnetic calorimeter [82]. These tracks are obtained after applying the Gaussian-sum filter [83] (GSF) method which takes into account the energy loss through related to bremsstrahlung. Second, the clusters are built around the seeds using an extended window of size 3×7 in the barrel region or 5 x 5 in the end-cap region. It is done by expanding the cluster size on either side of the original seed in ϕ or η. In the transition region ($1.37 < |\eta| < 1.52$), a method employing both elements of the extended-window size is utilised. Ultimately, the energy of the clusters must be calibrated

to match the energy of the original electrons based on data and simulated samples [84, 85]. The calibration is performed only after selecting electron candidates. The energy scale and resolution are calibrated using Z→ee decays. The energy of the final candidate electron is calculated from the calibrated energy of the extended-window cluster, while the ϕ and η directions are derived from the corresponding track parameters.

Identification and isolation criteria:

Once the electron reconstruction is complete, identification algorithms are used to remove the misidentified electrons. These could be from non-prompt electrons produced in photon conversions or from QCD jets and heavy-flavoured hadron decays. The identification is done using a likelihood discriminant using variables measured in the electromagnetic calorimeter and the ID. The discriminating variables are defined based on the electromagnetic shower shape and the track quality information. Depending on the likelihood discriminant distribution, three working points are defined in increasing order of background rejection: Loose, Medium, and Tight. Figure 37 depicts the electron identification efficiency for the aforementioned working points in Z→ee decays.

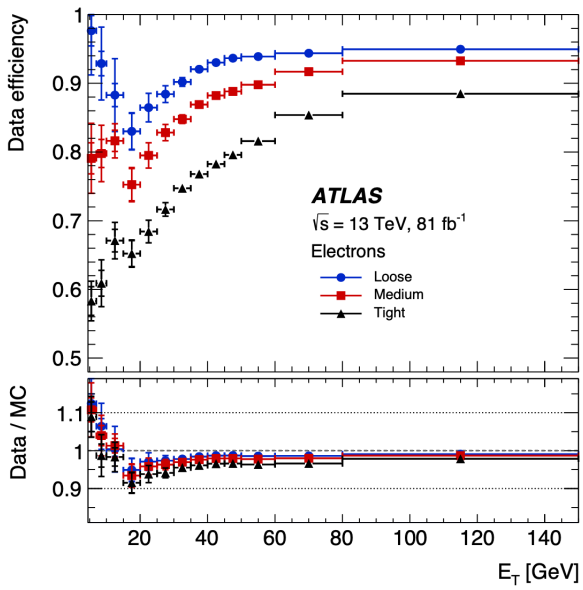


Figure 37 – Electrons identification efficiency in Z→ee decays. The efficiencies are calculated with respect to reconstructed electrons [81].

The isolation criteria are applied to suppress background from non-prompt electrons produced in the heavy-flavour hadron decays or from light hadrons misidentified as electrons. These requirements uses a track-based and a calorimeter-based variable, which are associated with the p_T of the tracks and the energy deposited in the calorimeters within a cone centred around the electron track, respectively.

2.4.3 Muons

Muons pass through the calorimeter system with minimum energy loss. The information from the ID and the MS can be combined for muon reconstruction [86].

Reconstruction:

The reconstruction of muons starts with independent reconstructions in the ID and MS. Then information from both reconstruction algorithms is combined to form muon tracks. The reconstruction in the ID is the same as for other charged particles, described in Section 2.4.1. In the MS, the reconstruction starts with the formation of local track segments in each muon chamber. These segments are then combined into track candidates, which require at least two matching segments per track, with the exception of the barrel end-cap transition region, where one segment can be used. The track candidates are built by fitting these segments using a global χ^2 fit. The combined muon reconstruction is carried out using various algorithms that take the information provided by the ID, MS, and calorimeters as inputs. Muons are divided into Combined (CB), Segment-tagged (ST), Calorimeter-tagged (CT) and Extrapolated (ET) muons, based on which sub-detector is used in the reconstruction.

- *Combined muons*: These muon candidates can be reconstructed by fitting hits from matched ID and MS tracks.
- *Segment-tagged muons*: These muons are those that do not fall into the combined muons category. They are reconstructed by an ID muon track being associated with one local track segment from the MS.
- *Calorimeter-tagged muons*: These muons are reconstructed by an ID track matched to an energy deposit in the calorimeter. This muon type has the lowest purity and is optimised for the region $|\eta| < 0.1$.
- *Extrapolated muons*: These muons are only reconstructed in the MS by extending the acceptance to $2.5 < |\eta| < 2.7$, which is not covered by the ID.

Identification and isolation criteria:

The non-prompt muons which mainly comes from pions and kaons decays, can be suppressed via identification procedure similar to electrons. Reconstructing prompt muons with a high efficiency and a good momentum resolution requires particular amount of hits in both the ID and MS. The combined muons are required to have at least three hits ((one hit and at most one hole) in at least two MDT layers except for the $|\eta| < 0.1$ region where only one MDT layer is sufficient. In physics

analyses, different working points can be used based in the analyses. In this thesis, Loose and Medium are used. The reconstruction efficiency at medium WP with $p_T > 20$ GeV is 96.1%. The medium WP aims to reduce systematic reconstruction and calibration uncertainties associated with muons. In case of loose WP, the reconstruction efficiency reaches up to 98.1%, where all the muon types are used. Figure 38 shows the reconstruction efficiency measured in data from $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events for Medium WPs.

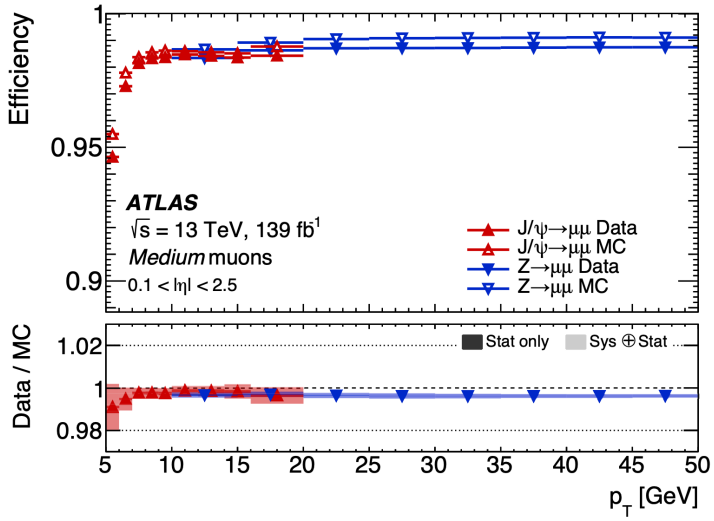


Figure 38 – The reconstruction efficiency as a function of the muon p_T for $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events for Medium WPs [86].

The muon isolation is used to reject heavy flavour semi-leptonic decays taking place inside the jets. The isolation criterias are similar to the ones used for the electrons. The track-based and calorimeter-based isolation criteria are used, similar to electron.

2.4.4 Jets

Due to the colour confinement, discussed in Section 1.5, gluons and quarks produced as final state partons cannot exist as free particles and are not directly observed in the detector. Instead, they hadronize by producing collimated showers, called *jets*. The jets are observed as clusters of energy deposits in the calorimeter system, which can be associated to the charged particle tracks in the ID.

Jet reconstruction:

The reconstruction of jets depends on the chosen clustering algorithm which enables to match the physics objects to the initial partons from which they originated. The charged particles in the jet, deposit energy in the ECal and HCal after leaving tracks in the ID. The energy is also deposited in these calorimeters by neutral objects. These form topoclusters, which are topologically adjacent clus-

ters of calorimeter. These are constructed with the help of a specific clustering method [87]. These clusters are used to perform standard ATLAS jet reconstruction using the *anti*-k_t *algorithm [88]*. It is a clustering algorithm that combines sequentially the four-vector objects into final jets, which is a cone-like object, characterised by a radius parameter R. The distances d_{ij} between objects i and j, d_{iB} between object i and the beam[2] line B, is introduced:

$$d_{ij} = min(p_{Ti}^{2p}, p_{Tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \tag{2.6}$$

$$d_{iB} = \frac{1}{p_{Ti}^2} \tag{2.7}$$

where, $\Delta_{ij}^2 = (y_i - yj)^2 + (\phi_i - \phi j)^2$ and $p_{T(i)}$, y_i, ϕ_i are the transverse momentum, rapidity and azimuthal angle of particle $i(j)$. The clustering process, involves the identification of the minimum between d_{ij} and d_{iB}. When d_{iB} is the smallest, particle i is termed as jet and is removed from the list of objects. This recursive method is followed until no objects are left in the list. This algorithm has the advantage of producing reconstructed jets with stable shape and properties that are not easily influenced when one particle is replaced with two collinear particles and when soft particles are added. In this thesis, the jets used for the analyses are reconstructed via the anti-k_t algorithm with R=0.4.

Other than topo-clusters, the anti $-k_t$ algorithm takes a variety of objects as inputs. The jet that can be reconstructed using tracks from ID is referred to as *track-jet*. Both the calorimeter jets and the track jets are reconstructed based on the detector's response. These jets are referred to as *reco-jets*. The jets that are reconstructed in generator-level simulated events are known as *truth-jets*. These are not reconstructed from topo-clusters found in the calorimeter, but rather from stable particles generated in MC samples.

Jet calibration:

These reconstructed jets are calibrated to match the energy of the truth-jets. Figure 39, describes different steps involved in jet calibration in ATLAS. The energy scale of the electromagnetic showers is used for the initial calibration of the calorimeter clusters. The four-moment of jets are reconstructed from the clusters, which go through several stages of calibration. First the jet is corrected to point back to the PV, without affecting the measured energy of the jet. Second, the pile-up corrections are applied to reduce the contribution from the pile-up. This is carried out in

[2] not to be confused with a real beam

two stages: the area-based correction followed by the residual pile-up correction. After removing the pile-up contribution, the MC samples without pile-up are used to calculate the absolute jet energy scale (JES). It allows the correction of the jet four-momenta to the particle-level energy scale of truth-jets. Following that, the global sequential calibration is implemented to account for multiplicative corrections to the jet energy measurement. These corrections are based on calorimeter, MS, and track-based variables. The tracking information is used to reduce the flavour dependence of the calorimeter response to jets, i.e. the observed difference in calorimeter response to jets initiated by different parton flavours. The calorimeter energy deposit and MS information are utilised to enhance the jet energy resolution. Finally the in situ energy calibration is derived from data using well-measured physics objects such as photons and Z-bosons, to account for mismodelings in simulations and their calibration steps.

Pile-up jet corrections:

Many physics analyses require the suppression of jets originating from pile-up interactions [89] in order to properly measure the targeted HS process. Jets are filtered using a Jet Vertex Tagging (JVT) algorithm [90] to select HS jets. The JVT provides a discriminant based on likelihood function of two variables. First, the fraction of the total momentum of tracks in the jet which is associated with the hard-scatter PV. Second, the ratio of the sum of the p_T of tracks within the jet which originate from the hard-scatter PV to the fully calibrated jet p_T. The JVT score ranges from 0 (pileup-like) to 1 (HS-like). After the jet is calibrated and reconstructed, JVT is used to reject pileup for jets by requiring JVT > 0.59 for jets with $p_T < 60$ GeV and $|\eta| < 2.4$.

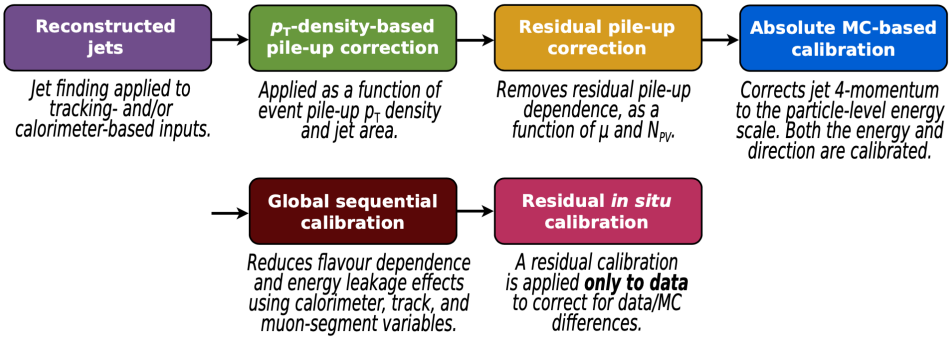


Figure 39 – Jet calibration chain used in ATLAS [91].

Electromagnetic Topocluster and Particle Flow jets: ATLAS employs a variety of jet collections based on the anti $-$ k_t algorithm. *EMTopo* and *Particle Flow jets* are two of these which are employed in the $t\bar{t}H(H \to b\bar{b})$ analysis, presented in Chapter 5. EMTopo jets are reconstructed using only calorimeter-based energy information [91]. The topoclusters are calibrated at the EM energy scale.

The jets are fully calibrated using the jet calibration technique including a correction for pileup, as described above. Until recently, EMTopo jets were the main jet collection used for ATLAS physics analyses. The ATLAS Particle Flow (PFlow) [92] reconstruction method combines the measurements from the calorimeter and the ID to improve jet measurements as opposed to simply using track based corrections in the jet calibration. First, the tracks from charged particles in the ID are matched to the topoclusters from calorimeters. The energy deposit of the topoclusters matched to tracks is replaced with the corresponding track momentum. The addition of track measurements is beneficial because the momentum resolution of tracking detectors exceeds the energy resolution of calorimeters for low-energy charged particles. In addition, the reconstruction is extended to softer particles whose energy deposits do not surpass the threshold for topocluster formation. The tracking detectors have finer granularity than calorimeters and thus, provide much better angular resolution for single charged particles. As a result of these advantages, PFlow jets have improved angular and energy resolution compared to EMTopo jets. After proper jet calibration and reconstruction, the PFlow jets are associated to the HS interaction by requiring JVT > 0.2 for jets with $p_T < 60$ GeV and $|\eta| < 2.4$. This selection criteria has the same pileup jet rejection as the tighter JVT selection for EMTopo jets with a higher HS jet efficiency.

***b*-tagging working points**:

The *b*-jets coming from the hadronization of *b*-quarks can be distinguished from other kinds of jets by exploiting the distinctive features of *b*-hadrons. The identification of *b*-jets, known as *b*-tagging, is important for many physics analyses and is detailed in Chapter. 3.

In an ideal situation, the full spectrum of the final *b*-tagging discriminant would be calibrated and utilised in the physics analyses. This continuous calibration would necessitate a separate calibration in very fine efficiency bins, resulting in an enormous complexity and required workload that is not feasible within required time scales in ATLAS. Therefore, four distinct *b*-tagging *working points* (WPs) covering various physics analysis requirements are defined. The efficiency with which a jet of specific flavour j (b, c or light) is identified by a *b*-tagging algorithm is given by:

$$\varepsilon^j = \frac{N^j_{pass}(\mathcal{D} > C_f)}{N^j_{total}} \tag{2.8}$$

where $N^j_{pass}(\mathcal{D} > C_f)$ are the number of jets of flavour j passing the cut C_f on the tagger discriminant $\mathcal{D}$ and N^j_{total} are number of jets of flavour j before the cut. The WPs are defined using the *b*-jet efficiency ε^b evaluated on a $t\bar{t}$ sample. The b-tagging is calibrated independently for each WP in order to correct the b-tagging efficiency in simulation to data. To do this, data-MC scale factors (SF) are estimated as the ratio between data and simulated efficiencies which are applied on the MC

result. Table 4 shows the list of WPs used in ATLAS. In addition to the *b*-jet efficiency, each WP is distinguished by its *c*- and light-jet rejection, which is the inverse of the *c*- and light-jet efficiency, respectively.

Selection	*b*-jet efficiency
loose	85%
medium	77 %
tight	70 %
very-tight	60 %

Table 4 – List of the single cut WPs for *b*-tagging in ATLAS.

The misidentification improves with a decrease in signal efficiency and consequently rejects more background in exchange for a lower signal statistics. Every jet satisfying the WP criteria is classified as a *b*-jet. In addition to WPs, five tag score bins, known as the pseudo-continuous *b*-tagging scores , are defined. These scores are retrieved based on the single cut WPs, listed in Table 5.

b-jet efficiency	tag score
(85,100)%	1
(77,85)%	2
(70,77)%	3
(60,70)%	4
(0,60)%	5

Table 5 – List of the pseudo-continuous *b*-tagging scores in ATLAS.

Most of the *c*- and light -jets end up in first bin [85, 100]%. The last bin [0, 60]% mainly consists of *b*-jets which passes the tightest WP (60%). Compared to single cut WPs, the pseudo-continuous b-tagging scores provides significant additional information. The requirement for *b*-tagging can be eased, and by including the the pseudo-continuous *b*-tagging score as an additional input, the information is retained while gaining more statistics. It is a measure of the quality of the *b*-jet that is particularly useful for the multivariate analysis methods, presented in Chapter 4.

2.4.5 Missing transverse momentum

The measurement of missing transverse momentum [93], referred as MET or E_T^{miss}, is critical for processes involving decays into neutrino which majorly escapes from the detector without any interaction. In this thesis, the E_T^{miss} is of interest because the leptonic final states in the $t\bar{t}H(H \rightarrow b\bar{b})$ analyses contain neutrinos. Practically, it is not possible to measure the full missing energy due to the inaccessibility of initial parton momenta. These however can be assumed to be negligible in the

transverse plane given that pp collisions occur along the z-axis. E_T^{miss} is the negative vector sum of the transverse momenta of all reconstructed and calibrated electrons, photons, jets, hadronically decaying τ-leptons, and muons, denoted as hard term. To account for the contribution from underlying events and soft radiations, an additional term is considered. It is computed using the momenta of tracks from the ID that match the HS vertex but are not associated with the reconstructed objects. This method is more resistant to pile-up interaction contamination than the calorimeter-based approach used in Run 1, which used energy deposits in calorimeters that were not associated with the hard objects.

2.5 High-Luminosity LHC upgrade and ATLAS Phase II upgrades

The High Luminosity Large Hadron Collider (HL-LHC) is an upgrade of the LHC which aims at maximize the sensitivity for physics analyses by boosting the instantaneous luminosity. Figure 40 shows the timeline of the LHC baseline programme including the HL-LHC phase. The HL-LHC upgrade [94] is expected to increase the instantaneous luminosity up to $7\times10^{34}\text{cm}^{-2}\text{s}^{-1}$, resulting in an accumulated integrated luminosity of around 4 ab^{-1} by the end of HL-LHC data-taking. This will also lead to an increase of up to 200 proton-proton collisions per bunch crossing.

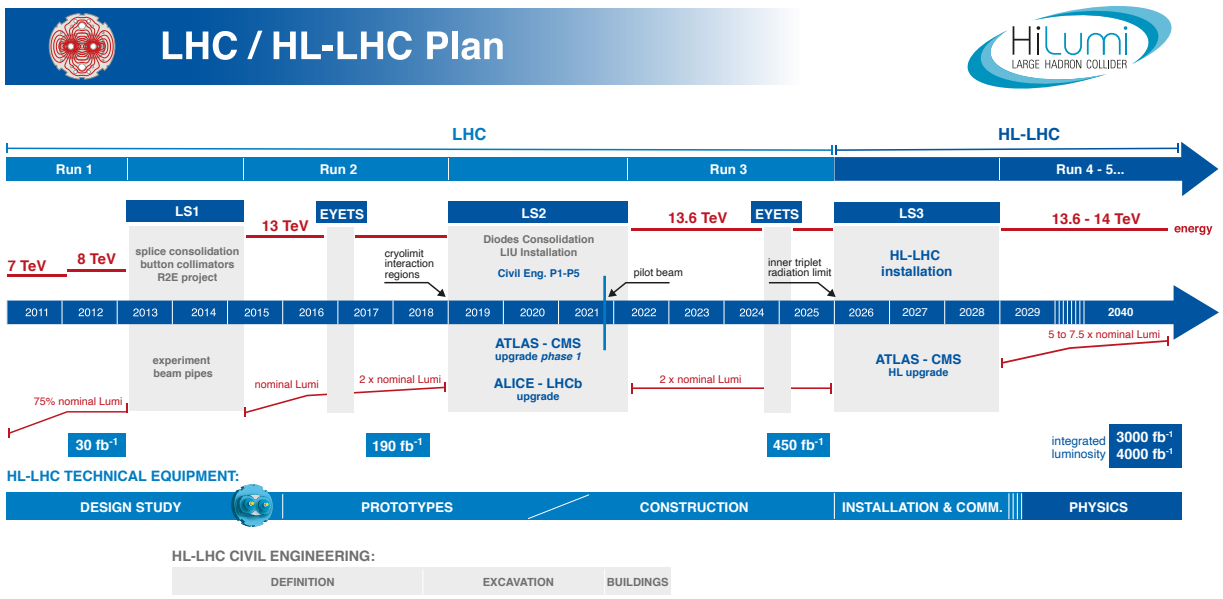


Figure 40 – Timeline of the LHC baseline programme including the HL-LHC phase [95].

The large amount of data collected during HL-LHC phase will greatly improve the experimental precision of SM measurements, and also allow to investigate Higgs boson self-coupling and rare phenomenas. The collider upgrade is complemented by experiment upgrades allowing the detectors to fully exploit the HL-LHC potential while also dealing with very high pile-up of up to 200 in the ultimate scenario. Due to these challenging conditions the current ATLAS ID [46] must be replaced with ATLAS Inner Tracker (ITk) [96, 97]. In addition, a High-Granularity Timing Detector [98] will also be installed covering the η region between 2.4 and 4 to help mitigate pile-up effects by using timing information to distinguish between collisions that are close in space but far apart in time. Major upgrades are also planned for the calorimeters [99] [100] and MS [86] read-out electronics. The trigger and data acquisition system will undergo upgrade [101], with the purpose of a faster and more efficient online selection of collision events.

2.5.1 ATLAS Inner Tracker

The ATLAS Inner Tracker (ITk) is an all-silicon tracker which extends the tracking coverage up to $|\eta| = 4$ (was 2.5 for the ID). It consists of two subsystems: the inner Pixel detector [97] and the outer Strip detector [96].

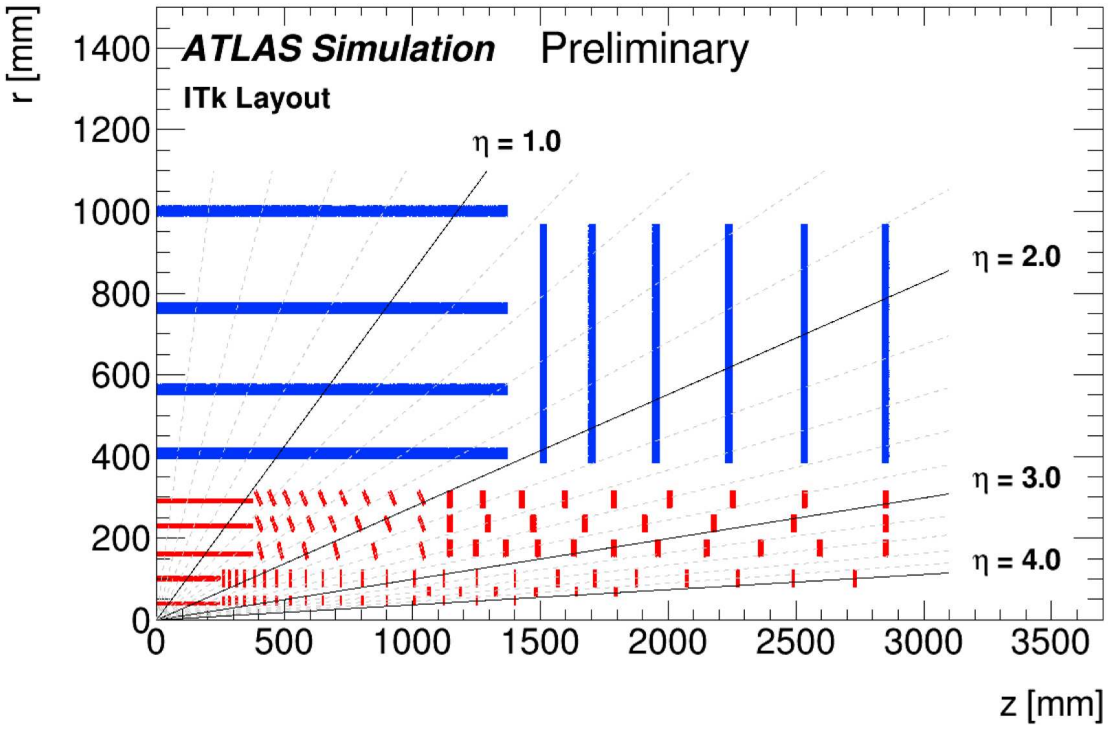
The combination of these detectors are designed to provide at least nine precision measurements per track for the full expected beam spot of size of 50 mm, for each charged particle with $p_T > 1$ GeV passing through the detector within $|\eta| < 4$. The ITk layout, considered for performing the *b*-tagging studies, presented in Chapter 4, is shown in Figure 41. The radius at which the barrel layers are placed and the z-positions of the end-cap disks, have been selected to optimize the number of hits on a track and the p_T resolution. Figure 42 shows a display of the ATLAS ITK detector.

ITk Strip detector

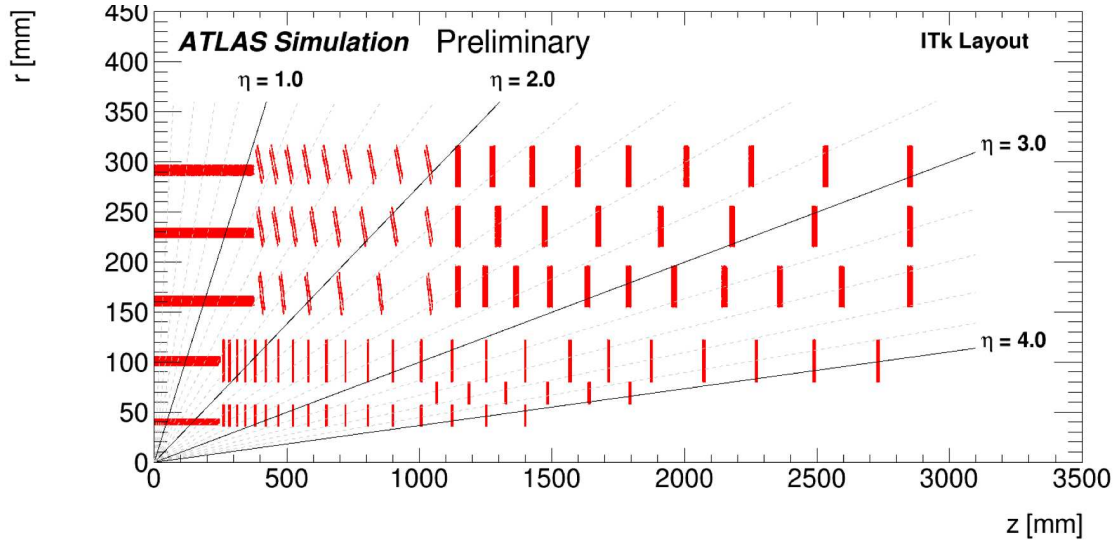
The Strip Detector that extends up to $|\eta| = 2.7$ has a four-layer barrel section and a six-disk end-cap on each side (cf. Figure 41). Each strip barrel layer extends along the z-axis from -1.4 m to +1.4 m. The two inner (outer) barrel layers have 300 (320) μm thick and 24.1 mm (48.2 mm) long strip sensors whereas the end-caps have strips ranging in length from 19 to 60.1 mm depending on the radius. Each barrel layer (end-cap disk) consists of pairs of detectors mounted back-to-back.

ITk Pixel detector

The Pixel detector provides tracking coverage up to $|\eta| = 4$ and has five pixel layers and multiple inclined or vertical ring-shaped end-cap disks (see Figure 41). Different pixel pitches and pixel sensors are used in the ITk layers. The pixel pitch of 50×50 μm^2 and 25×100 μm^2 have been considered for the pixel sensors. The ITk layout considered for the studies shown in Chapter 3, uses a pixel pitch of 50×50 μm^2 throughout the whole pixel detector. Due to the extreme radiation



(a)



(b)

Figure 41 – (a) A schematic depiction of the ITk layout considered in most studies presented in this thesis. (b) A zoomed-in view of the Pixel detector. In each case, only one quadrant and only active detector elements are shown. The active elements of the strip detector are shown in blue, and those of the pixel detector are shown in red. The horizontal axis is along the beam line with zero being the interaction point. The vertical axis is the radius measured from the interaction region [102].

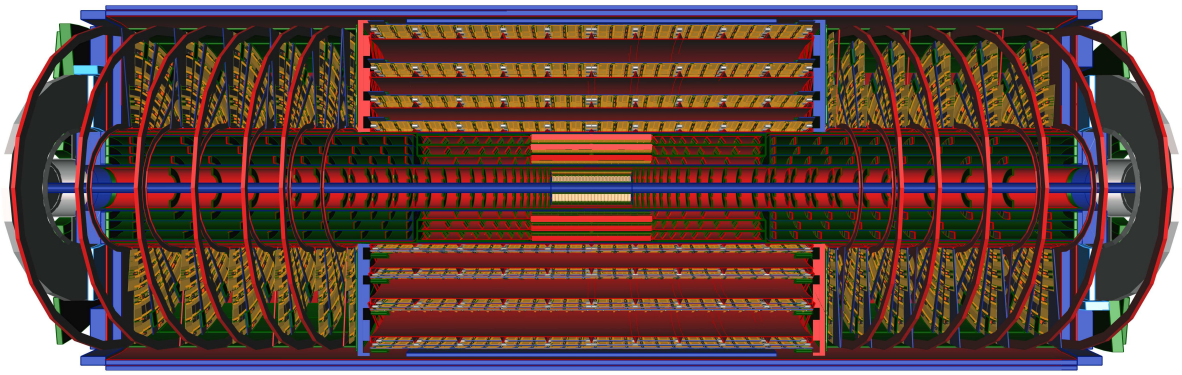


Figure 42 – Display of the latest Inner Tracker layout [1].

environment expected at the HL-LHC, the two innermost barrel and ring layers are replaceable. The remaining outer layers and rings are expected to last the entire lifetime of the ITk.

At the time of writing, optimizations of the ITk layour have been carried out and a new updated ITk layout has been adopted by the ATLAS [1]. The radius of the innermost pixel layer (L_0) is reduced from 39 to 34 mm and the pixel sensor pitch to be used in the barrel in L_0 has also been fixed to $25 \times 100\ \mu m^2$. The rest of the pixel detector will still be based on the original pixel pitch of $50 \times 50\ \mu m^2$. These changes significantly improved the transverse impact parameter resolution of the reconstructed tracks, which is strongly correlated with the *b*-tagging performance. The accuracy of the simulation has also been improved, reflecting the most recent ITk detector engineering developments.

CHAPTER

3

EXPECTED *B*-TAGGING PERFORMANCE WITH THE ATLAS PHASE 2 DETECTOR

Many physics analyses at the LHC rely on *b*-jet identification, such as top physics and measuring Higgs boson couplings to heavy quarks like $t\bar{t}H(H \rightarrow b\bar{b})$ process, which is explored in Chapter 5. With a much larger data-set at the upcoming HL-LHC, *b*-jet identification will be crucial for investigating Higgs boson self-coupling and maximising sensitivity for precision measurements and new physics searches. This is why several algorithms are developed within ATLAS for the identification of *b*-jets, commonly called *b*-tagging. The intrinsic properties of *b*-jets, introduced in Section 3.1, are used to define the *b*-tagging algorithms, described in Section 3.2. The tracking information from ATLAS ID is critical for *b*-tagging. As mentioned before, the current ID will be replaced by the ATLAS ITk, to maintain tracking performance in extreme conditions during the HL-LHC phase. The robust tracking performance against HL-LHC conditions is essential in order to maintain and possibly improve the current Run 2 *b*-tagging performance for the ITk. The *b*-tagging performance is thus an important benchmark for the tracking performance of the ITk detector. After a general introduction to *b*-tagging, this chapter will primarily investigate the *b*-tagging performance with the ITk, as presented in Section 3.3. One of the main contributions of this thesis are some re-optimisation of *b*-taggers, discussed in Section 3.4. The overall developments in *b*-tagging algorithms for the ITk have been used for performance studies based on the evolved ITk layout, as described in 3.5.

3.1 *b*-hadron properties

The identification of the *b*-jets originating from the hadronization of *b*-quarks is made possible by the distinctive properties of the b-hadrons produced in the process. These are the fundamental elements upon which *b*-tagging algorithms are built. One of the most distinctive *b*-hadrons property is their relatively long lifetime ($\tau \approx 1.5$ ps). As an example, *b*-hadrons with a transverse momentum of $p_T = 50$ GeV and a mass of about 5 GeV, will have a significant transverse flight path length of $L_{xy} = \beta\gamma c\tau = 4$ mm. This provides two signatures that the *b*-tagging algorithms can directly exploit. Firstly, the *b*-hadron decay gives rise to at least one secondary vertex (SV) displaced with respect to the primary vertex (PV) where the hard-scatter collision occurred. Secondly, the charged decay products of the *b*-hadron coming from the displaced vertex, will be reconstructed as tracks having large Impact Parameters (IP). The transverse IP (d_0), is the distance of closest approach of the track to the PV point, in the $r-\phi$ projections. The longitudinal IP (z_0) is defined as the difference between the coordinates of the PV and of the track at this point of closest approach. The Figure 43 depicts the aforementioned properties.

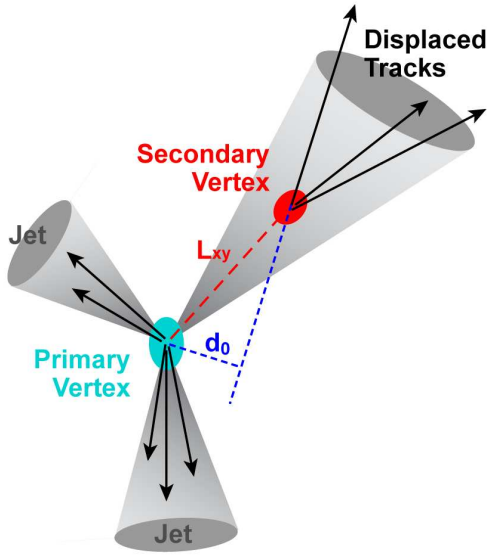


Figure 43 – A sketch illustrating the distinctive properties of a *b*-jet, namely the presence of a secondary vertex within the jet and the large transverse impact parameters (d_0) of the tracks originating from the displaced vertex.

Another feature of *b*-hadrons that is useful for *b*-tagging algorithms is their relatively high mass ($\approx$ 5 GeV) in comparison to *c*- and light-hadrons. As a result, *b*-hadron decays generate a large number of charged particles, resulting in high track multiplicity and large opening angles within *b*-jets. Similar to *b*-hadrons, *c*-hadrons have a lower decay multiplicity and have shorter lifetime, resulting in a smaller IP. This results in similar but not identical jet topologies of *c*- and *b*-jets, making them

difficult to distinguish. The majority of the tracks within the light-jets originate directly from the fragmentation of quarks.

3.2 *b*-tagging algorithms

The *b*-tagging algorithms used by ATLAS physics analyses to distinguish *b*-jets from *c*- and light-jets are built in two steps. First, the low-level b-taggers are developed by directly utilising the properties of b-hadrons (see Section 3.1). These low-level b-tagging algorithms, are either IP- or SV-based algorithms, as detailed in Section 3.2.1 and Section 3.2.2, respectively. The development of low-level *b*-tagging algorithms is the main focus of the expected *b*-tagging performance studies with the ITk detector. Second, since their performances are largely uncorrelated, they can be combined into a single powerful discriminant, referred to as high-level *b*-taggers. Until recently, ATLAS employed the BDT-based *b*-tagger MV2, described in Section3.2.3, which has been replaced by the Deep Neural Network-based *b*-tagger DL1, described in Section3.2.4. The developments in low-level *b*-taggers, presented in Section 3.3 and Section 3.4, were utilised in performance studies based on the updated ITk layout, which also investigated MV2 performance. Overall, the final *b*-tagging quantities are computed after applying a specific cut on the *b*-tagging discriminant distribution, commonly known as the working point (WP), introduced in Section 2.4.4.

Figures of merit for *b*-tagging performance: Many figures of merit for *b*-tagging performance are based on certain key quantities. The first, is the efficiency with which a *b*-jet is identified by a *b*-tagging algorithm, known as the *b*-jet tagging efficiency. The second, is the probability of mistakenly tagging a light-jet as a *b*-jet, known as mis-tag rate. This can also be expressed as rejection factor, which is defined as the inverse of the mis-tag rates. The classification power can be measured by plotting the receiver operating characteristic (ROC) curve for each tagger. The ROC curve is essentially the background rejection as a function of the signal efficiency. Both the signal efficiency (here *b*-jet efficiency) and the background rejection are varied by the WP, which defines the signal/background cut on the tagger output. As an illustration, when the WP is set to 70%, any jet with a *b*-tagger output greater than 0.7 is classified as a *b*-jet, while any jet with a *b*-tagger output less than 0.7 is classified as light-jet.

3.2.1 Impact parameter-based algorithm

The IP-based algorithms utilise the long decay path of the *b*-hadron, resulting in a displaced vertex. The IP of tracks is computed with respect to the event PV candidate. It is possible to assign a sign to the IP (d_0 and z_0), defined relatively to the jet direction. The sign is based on whether the point of closest approach of the track to the PV is in front of or behind the PV, as illustrated in Figure 44.

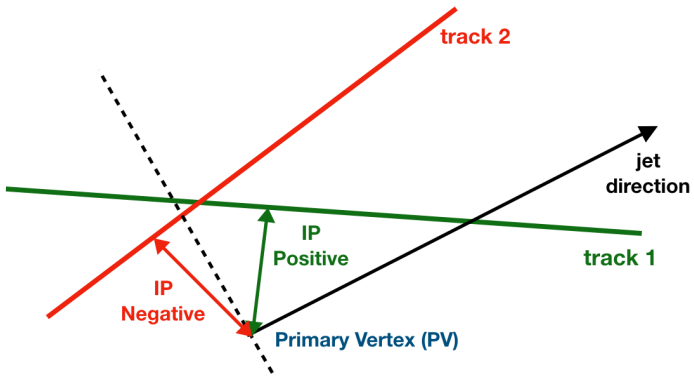


Figure 44 – A sketch illustrating the definition of positive and negative impact parameters signs based on the angle between the jet and the line between the primary vertex and the point of closest approach of the track.

The tracks associated with *b*-hadron decays are more likely to have a positive sign, while tracks associated with particles produced directly at the PV have equal probabilities of getting a positive or negative sign. The tracks with $p_T > 1$ GeV and fulfilling the requirements of $|d_0| < 1$ mm and $|z_0 \sin\theta| < 1.5$ mm are considered to reduce the contamination from pile-up interactions. IP3D tagger makes use of these d_0 and z_0 significances in two-dimensional template

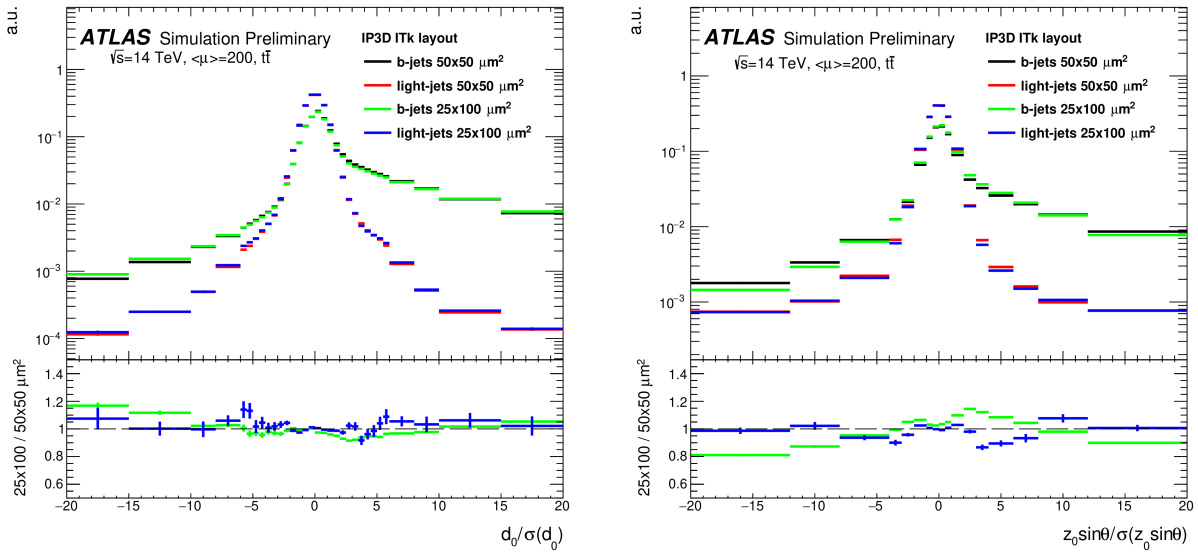
$$S \equiv \left(\frac{d_0}{\sigma_{d_0}}, \frac{z_0 \sin\theta}{\sigma_{z_0 \sin\theta}} \right) \tag{3.1}$$

where, σ_{d_0} and $\sigma_{z_0 \sin\theta}$ are the uncertainties on the reconstructed d_0 and $z_0 \sin\theta$ respectively. Figure 45 shows the distributions for the d_0 and $z_0 \sin\theta$ significances for the tracks from *b*- and light- jets.

Probability density functions (PDFs) obtained from reference histograms for d_0 and $z_0 \sin\theta$ significances are derived from MC simulation. They are separated into exclusive categories based on the hit-based categories associated with different track quality[1] and track kinematic. In order to increase the discriminating power, PDFs are used for the various track categories and for each jet-flavour hypothesis to calculate the ratios of the *b*- and light-jet probabilities. The IP3D is based on the log-likelihood ratio (LLR) method, which is computed as the sum of the per-track contributions:

$$\text{IP3D} = \sum_{Tracks} \log \frac{p_b}{p_u} \tag{3.2}$$

[1] Quality of tracks is determined by the resolution with which its impact parameters are measured



(a) (b)

Figure 45 – The distributions of (a) d_0 and (b) $z_0 \sin\theta$ significances in the category of central tracks from *b*- and light- jets with $|\eta| < 1$. The distributions are obtained with "50x50 μm^2" and "25x100 μm^2" pixel pitches of the ITK layout [103].

where, p_b and p_u are the PDFs for the *b*- and light-flavour jet flavour hypotheses, respectively. Figure 46 shows the IP3D LLR distribution for *b*-jets, *c*-jets and light-jets.

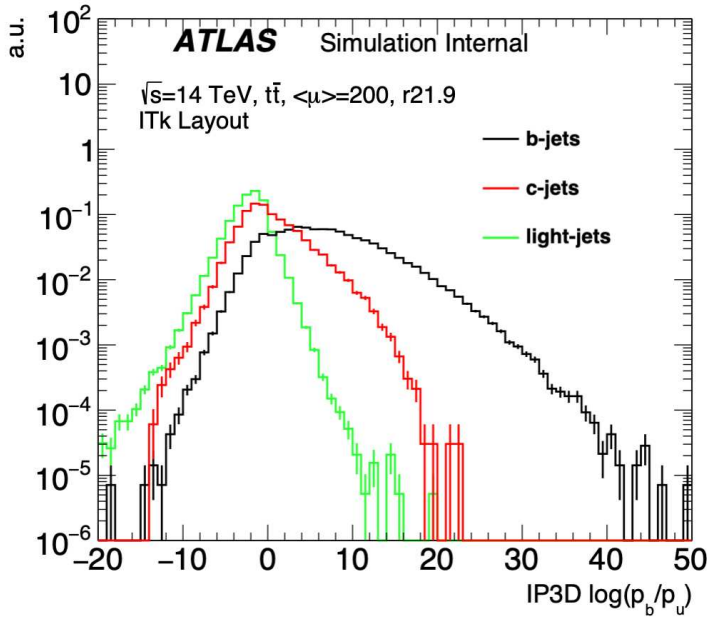


Figure 46 – IP3D log-likelihood ratio distribution for *b*-jets, *c*-jets and light-jets obtained in $t\bar{t}$ events simulated with the ITk detector.

Different LLRs are potentially built to separate either b- from light-jets, as well as b- from c-jets, and c-jets from light-jets. All three discriminants can be used by high-level flavour taggers when fed as inputs. The scope of this study focuses only on IP3D algorithm separating b- and light-flavour jets and the other two are not used hereafter. The IP3D b-tagging performance with the ITk detector is presented in Section 3.3.

3.2.2 Secondary vertex-based algorithms

The reconstruction of displaced SV within the jet is exploited by the SV-based algorithms. The Single Secondary Vertex Finder (SSVF) [104] is one such algorithm that aims to reconstruct an inclusive SV per jet. The reconstruction begins with the identification of all possible two-track vertices built with all tracks associated with the jet, while rejecting tracks compatible with long-lived particle decay (Kaons or Λ), photon conversions, or hadronic interactions with the detector material. The description is precise for ITk and very similar to what is done for recent Run 3, where some of the features involving track reconstruction are updated in Run 3. [105, 106]

If at least a two-track vertex remains, a new vertex is fitted with all tracks from the accepted two-track vertices, with outliers removed iteratively from this collection of tracks. In case a b-jet contains both b- and c-hadron decay vertices, the SSVF merges these vertices into a common single vertex if they are close, or reconstructs the vertex with the largest track multiplicity if they are far apart. The SSVF algorithm has a small probability of reconstructing a (fake) vertex when there are no real SV in the jet. This feature is crucial for the rejection of light jets in the b-tagging. Several track selections are applied before forming all possible two-track vertices. Tracks are ordered according to their p_T and at most 25 tracks with largest p_T are considered, in order to reduce the number of fake vertices. Their transverse impact parameter must fulfil $|d_0| < 3.5$ mm and $|z_0 \sin\theta| < 25$ mm selection. In order to mitigate the higher hadronic interactions between detector materials, further track cleaning is done such as implementing material rejection to discard SV close to the layers of the pixel detector.

The properties of this SV are then used as inputs to a likelihood ratio formalism similar to the one used for the IP3D tagger (cf. Section 3.2.1), known as the SV1 algorithm. SV1 exploits the following properties of the reconstructed secondary vertex:

- Vertex mass (invariant mass of all charged-particle tracks used to reconstruct the vertex)
- Ratio of the sum of the energies of these tracks to the sum of the energies of all tracks in the jet (charged energy fraction)
- Number of two-track vertices (N2T)

- Angular distance ΔR between the jet axis and the direction of the line joining PV and SV

SV1 uses a two-dimensional distribution of the first two variables and two one-dimensional distributions of the latter variables. Figure 47 illustrates some of the properties of SVs reconstructed by the SSVF algorithm. The SV1 tagger performance in context to the default ITk layout is presented in Section 3.3.

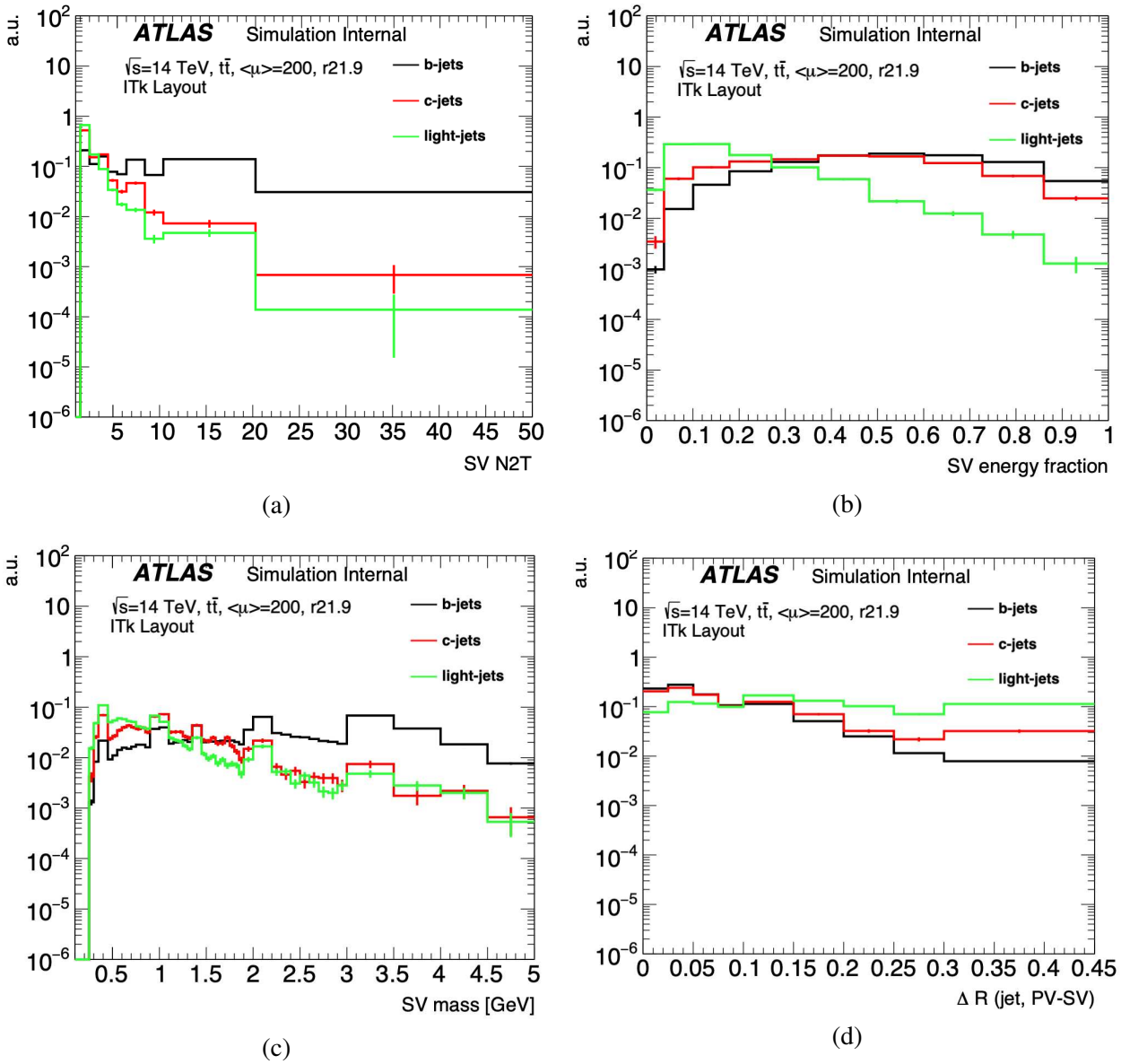


Figure 47 – The properties associated with the secondary vertices reconstructed by the SSVF algorithm for *b*-, *c*- and light-jets in simulated tt̄ events: (a) number of two-track vertices reconstructed within the jet, (b) invariant mass and (c) the energy fraction defined as the ratio of the energy of tracks associated with the SV relative to the energy of all tracks reconstructed within the jet and (d) ΔR between the jet axis and the direction of the line joining the PV and the SV.

JetFitter [107, 108] is a topological multi-vertex algorithm that attempts to reconstruct the full *b* - hadron decay chain by exploiting the *b*- and *c*-hadrons decay topology inside the jet. The *b*-tagging studies presented in this thesis do not cover JetFitter.

3.2.3 MV2

MV2 [109] is a multivariate algorithm based on a Boosted Decision Tree (see Section 4.2) that combines inputs from the low-level taggers, IP3D, SV1, and JetFitter, to improve discrimination between jet flavours. To take advantage of the correlations with the other input variables, the kinematic properties of the jets, like p_T and $|\eta|$, are also included in the BDT training. The BDT employs *b*-jets as the signal class and *c*- and light-jets as a single background class. A *c*-jet fraction of 7% and a light-jet fraction of 93% was found to be best for balancing *c* vs light-jet performance, at least for a wide range of analyses. This tagger is known as *MV2c10*, which was used as the recommended flavour-tagger for the EMTopo jets (see Section 2.4.4) for ATLAS physics analyses using Run 2 data. In this thesis, the following tagger will be referred to as MV2.

3.2.4 DL1

DL1 [110] is a multivariate algorithm based on a Deep Neural Networks (DNNs see Section 4.4) that combines inputs used for MV2 with the addition of JetFitter *c*-tagging variables. The DL1 has a multi-class DNN output corresponding to the probabilities for a jet to be a *b*-jet, a *c*-jet or a light-jet. Figure 48 shows the performance of high level tagger (MV2 and DL1) and low-level taggers (IP3D, SV1, JetFitter) in terms of the background light-jets rejection as a function of the *b*-jet efficiency. When compared to low-level algorithm performances, high-level tagger shows increase in light-jet rejection by factor of around 10 at 70% *b*-tagging efficiency. When compared between MV2 and DL1, the light-jet rejection with DL1 is improved by 30% at 70% *b*-tagging efficiency. A *c*-jet fraction of 1.8% and a light-jet fraction of 98.2% were found to be the best values for balancing *c* vs light - jet performance. The DL1r *b*-tagger, which is a version of DL1, also combines the result of the RNNIP algorithm [111], which is based on a recurrent neural network. RNNIP uses the same information as the IP-based algorithms and also exploits the correlation between the IP of the tracks. Due to the improved b-tagging performance of DL1r when compared to MV2c10, it is used as the recommended flavour-tagger for the PFlow jets (see Section 2.4.4). The latest $t\bar{t}H(H) \rightarrow b\bar{b}$ analysis, documented in Chapter 5, uses PFlow-jets using DL1r tagger and aims to exploit the benefits from improvements in *b*-tagging algorithm DL1r to improve the sensitivity.

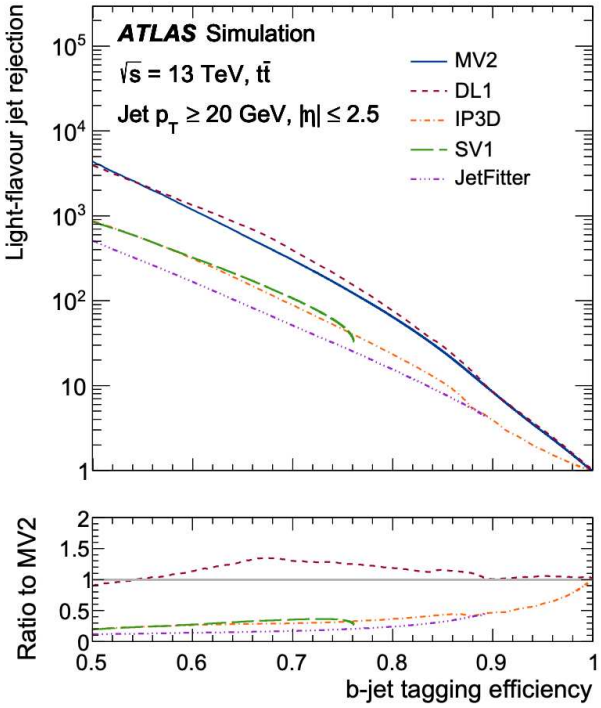


Figure 48 – Light-jet rejection vs the *b*-tagging efficiency for the IP3D, SV1, JetFitter, MV2 and DL1 *b*-tagging algorithms evaluated in $t\bar{t}$ events [109].

3.3 *b*-tagging performance in ITk

b-tagging performance studies in ITk aim to have at-least comparable performance to current Run-2, in spite of the more challenging pile-up conditions. The *b*-tagging developments related to the ITk at the ATLAS software level have been carried out for a while, and the previous developments were incorporated in ATLAS software release 20.20 (will be referred as r20.20). This software release was decommissioned after the HGTD TDR [98], in early 2020. Therefore, there is a need to re-establish r20.20 functionality and performance in new ATLAS upgrade release 21.9 (will be referred as r21.9). My work primarily focuses on migrating the previous developments from r20.20 to r21.9, particularly low-level *b*-taggers. The reference histograms templates used for the IP3D and SV1 LLR formalism in the following study are derived from simulations in r20.20 with an ITk layout similar to the one used in this study for r21.9. In this section, the overall *b*-tagging performances are compared and analyzed between r20.20 and r21.9. The primary goal is to achieve similar performance between the releases. Additional developments are also done in r21.9, discussed in Section 3.4.

3.3.1 Samples and physics objects selections

The studies described in this section, are done using simulated Monte Carlo samples produced in proton-proton collisions with a center-of-mass energy of $\sqrt{s}$ = 14 TeV. This sample was generated using POWHEG BOX [69, 112] interfaced to the PYTHIA 8 [72] parton shower model with the A14 set of tuned parameters. The EVTGEN 1.2.0 [113] was employed for the simulation of the *b*- and *c*-hadron decays. In order to assess the *b*-tagging performance in HL-LHC conditions, samples are generated with pile-up value of μ = 200. Samples without pile-up have also generated in order to perform cross-checks. The tracks reconstructed in the ITk, the primary vertex, and the jet direction acquired from calorimeter jets are the most relevant objects for *b*-tagging. The primary and pile-up interaction vertices are reconstructed using the Adaptive Multi-Vertex Finder(AMVF) [114] algorithm. The primary vertex is chosen as the one with the largest sum of squared transverse momenta (Σp_T^2) of the associated tracks. Simulation studies in $t\bar{t}$ events illustrate that using these criteria, the probability of identifying the proper signal primary vertex is around 95%.

The tracks are reconstructed, following the procedure described in Section 2.4.1.Tracks are then ranked accordingly to their hit content and track fit quality, where ambiguities between multiple track candidates are solved by comparing the corresponding scores and keeping the track with the highest score. The final reconstructed tracks are obtained after applying the η dependent tracking selections, summarized in Table 6. Depending on the *b*-tagging algorithm used, further track selections are applied subsequently. The *b*-tagging algorithms also rely on the association of tracks with jets. The association of jets to tracks is realized through a geometric matching by exploiting ΔR between the track and the jet axis. Since decay products from energetic heavy-flavoured hadrons are more collimated, the R requirement varies as a function of jet p_T, being wider for low p_T values and narrower for high p_T values [115].

Selections	$\lvert\eta\rvert > 2$	$2 < \lvert\eta\rvert < 2.6$	$2.6 < \lvert\eta\rvert < 4$
pixel + strip hits	≥ 9	≥ 8	≥ 7
pixel hits	≥ 1	≥ 1	≥ 1
holes	≤ 2	≤ 2	≤ 2
p_T [MeV]	> 900	> 400	> 400
$\lvert d_0\rvert$ mm	≤ 2	≤ 2	≤ 10
$\lvert z_0\rvert$ cm	≤ 20	≤ 20	≤ 20

Table 6 – Summary of the tracking selection criteria applied to reconstructed tracks in different η regions in the ITk detector [97]. A hole is an intersection of the predicted particle's trajectory with an active sensor element from which no measurement is assigned to the track (inactive sensors are not taken into account). Additional cuts are also applied for *b*-tagging quality selection.

The jets are reconstructed by clustering energy deposits in the calorimeter with the anti-k_t algorithm [88] with a radius parameter of 0.4. The key objects, *b*-, *c*- or light-jets in the simulation are

labeled, first by searching for *b*-hadrons within the cone of radius $\Delta R < 0.3$ around the jet-axis. If no *b*-hadron is found, the algorithm is repeated for *c*-hadrons, and for τ leptons. Finally, the remaining jets that do not fit into the previous categories are classified as light-flavor jets. The transverse momenta of the *b*- and *c*-hadrons are required to be larger than 5 GeV. In this document, only jets with $p_T > 20$ GeV and $|\eta| < 4$ are selected. In order to reject jets originating from pile-up interactions, truth-matching is applied to these selected jets. Truth-matching refers to the matching criteria to match the reconstructed jets to the particle-level jets or truth jets within $\Delta R < 0.3$ around the jet-axis. *Truth-jets* are reconstructed in simulated events at generator-level and are not reconstructed from topo-clusters found in the calorimeter, but rather from stable particles generated in MC samples. Also, the events where the selected PV is not within 0.1 mm from the true position of the HS vertex along the beam axis are discarded. Furthermore, generator-level jet filter for light-jets is also implemented, to reject light-jets matched to isolated electrons. This selection is defined as *electron veto* and is used as default matching scheme for light-jets in the following studies. Apart from this, another light-jet matching scheme to match light-quarks coming from top decays, referred as *top-matching* is used for the additional checks.

3.3.2 Track variable and track category studies

The *b*-tagging algorithm implementation in r21.9 has been synchronised with the developments in r20.20, particularly in IP3D and SV1 tagging algorithms. Thus, it is necessary to check the functionality and performance of the latest software implementation, and compare those with the one obtained with r20.20. To begin with, the track-related variables used in the definition of the track categories are analysed between r21.9 and r20.20. These track categories, which can depend on the quality of tracks are used in computation of IP3D tagger performance. The quality of tracks is determined by the resolution with which its impact parameters are measured and is inherently dependent on the hit content. Figure 49 shows the number of pixel and strip hits as a function of η.

In contrast to r20.20, the tracks with zero pixel hits are observed in r21.9. This might come from the track conversions, primarily the proportion of photons converting as a result of interactions with the detector material, which reaches 20% at $\eta = 0$ [1]. For accurate extrapolation of a trajectory track back to the PV, it is essential for a track to have a first measured hit at a small radius in the detector. If tracks lack this measurement in the first pixel layer i.e innermost or the second pixel layer i.e next-to-innermost layers, the resolution of their impact parameters will be low. Knowing whether the hit in the innermost and next-to-innermost layers is expected or not is also a quality criterion. If the extrapolation of the helix in that layer crosses an active part of the silicon sensor, a hit is *expected* in that layer. Figure 50 shows the number of innermost pixel hits and next-to innermost pixel hits as a function of η.

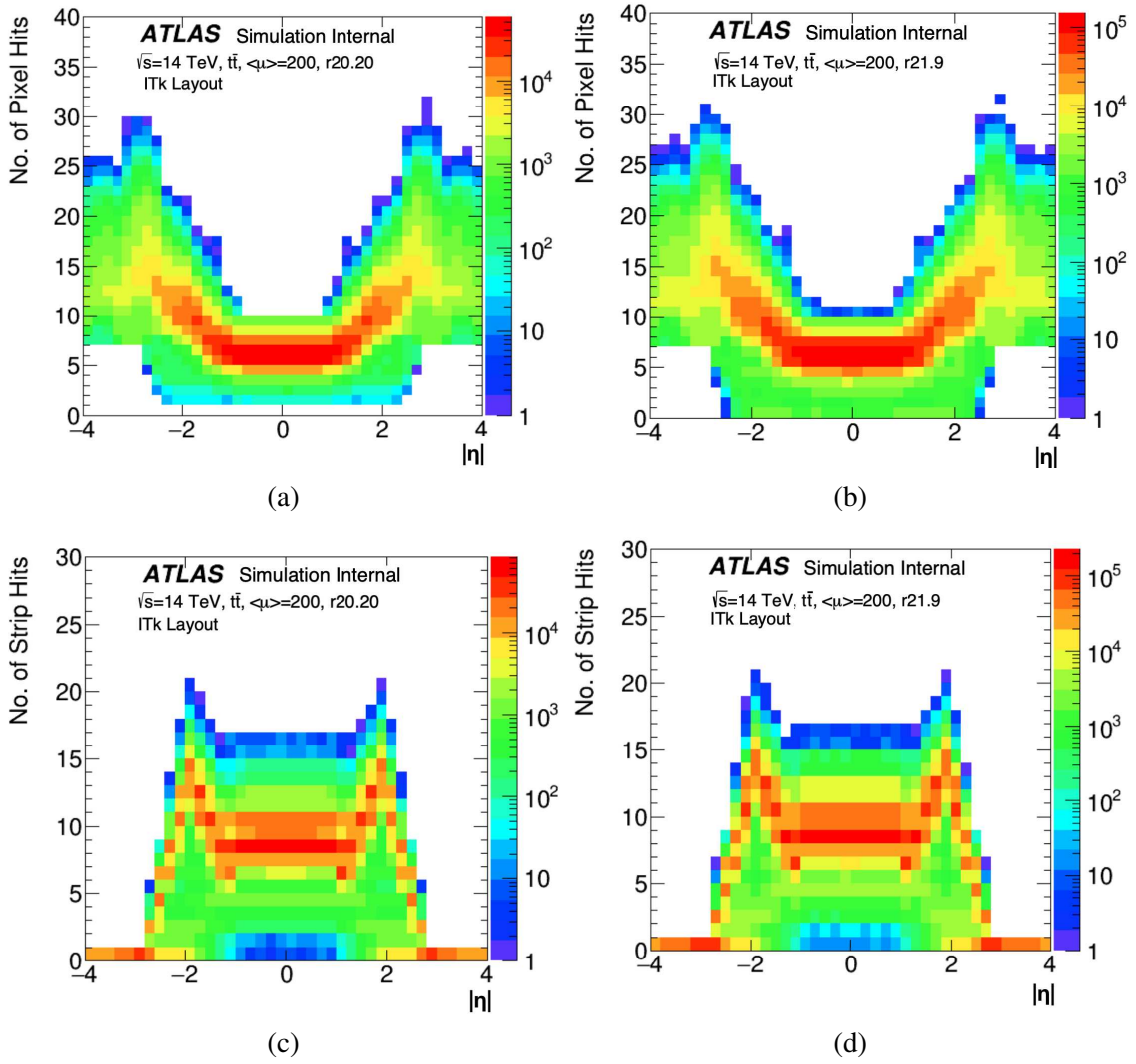


Figure 49 – Distribution of the number of pixel hits as function of η in (a) r20.20 and (b) r21.9 and distribution of the number of strip hits as function of η in (c) r20.20 and db) r21.9 in $t\bar{t}$ events with an average pile-up of 200.

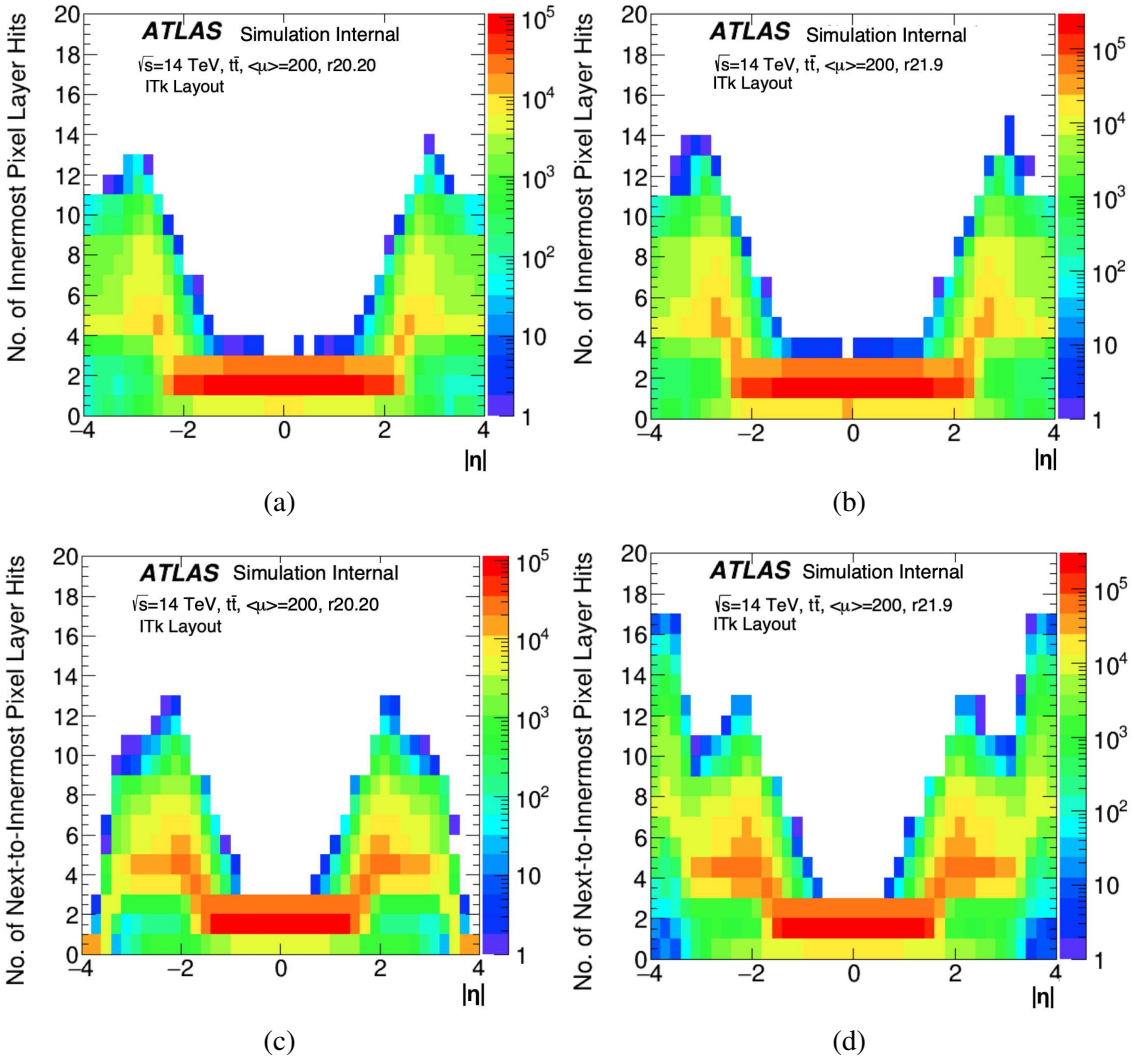


Figure 50 – Distribution of the number of hits in the innermost pixel layer as function of η in (a) r20.20 and (b) r21.9 and Distribution of the number of hits in the next-to-innermost pixel layer as function of η in (c) r20.20 and (d) r21.9 in $t\bar{t}$ events with an average pile-up of 200.

Differences are observed in next-to-innermost pixel layer hits distribution between the releases, highlighting a different convention used in r20.20, where only pixel rings up to $z = 1000$ mm were counted. The tracks which have no hit in in neither the first nor the second pixel layers, while they were expected, contains poor quality tracks. Other hit content parameters also have an impact on the quality of the tracks. A hit is considered *split* if it was identified as being created by several charged particles. If a pixel or strip hit is connected with more than one track and is not previously marked as split, it is labelled as *shared*, which is frequently a sign of a pattern-recognition issue caused by the high track density. The requirement that a track has a low number of shared measurements is a very powerful tool for preventing duplicate and fake[2] tracks.

The track categories used in the IP3D studies, aims to handle good and poor quality of tracks separately to improve the overall performance. These hit content parameters are used for defining track categories in central region $\eta < 1$, referred to as *region A*. The tracks of the worst quality are in category 01, while the best quality tracks belong to category 14. Similar criteria, based on hit information, are used for $1 < \eta < 2$ region, referred to as *region B*. This region have only five sub-categories as compared to region A. In general, the final track categories are defined after some merging of the categories with low statistics. In the forward region $\eta > 2$, the track categories are not defined based on the hit content but from an additional parameter γ. The tracks in this region suffers from a larger rate of material interaction and are indeed dominated by multiple scattering effects. Thus, the IP resolutions tends to decrease at higher η regions. The multiple scattering effect is proportional to the amount of material crossed by the particle. Figure 51 shows that the resolution of the IP can be set to be correlated with γ value given by: $\gamma \equiv \frac{1}{p_T \sqrt{sin\theta}}$

where, γ is inversely proportional to the p_T of the particle and θ is the polar angle. The low momentum tracks are likely to undergo large scattering effects. The resolution worsens as gamma increases, which is to be expected because tracks with a high gamma are expected to suffer the most from multiple scattering.

Finally, the IP3D PDF templates are obtained separately in each region, for the track categories just described, as summarized in Table 7. One way to assess the quality of the tracks is to examine their impact parameter resolution, mostly in d_0. The d_0 resolution is based on the width of the distribution of d_0 residual, which is the difference between the reconstructed and true values of d_0. The d_0 resolution has been computed per track category in the inclusive track p_T and compared between the releases, shown in Figure 53. Figure 52 shows the track categories distribution as function of η, which is consistent between r20.20 and r21.9. The d_0 resolution for each category is estimated by iterative convergence (IterConv) parameter, which relates to the number of iterations required to

[2] Tracks which are not associated with real particles.

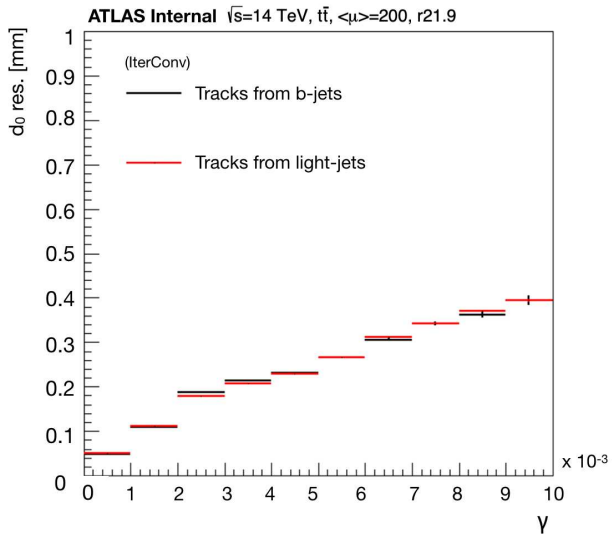


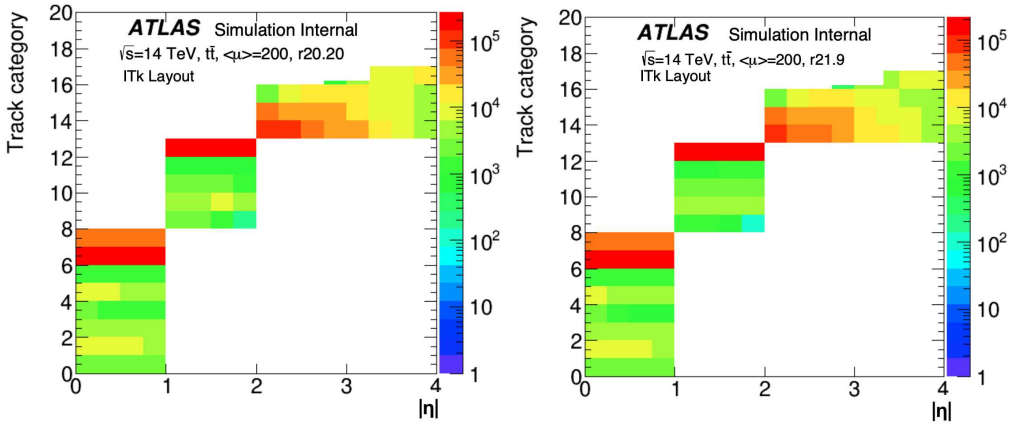
Figure 51 – γ parameter as a function of d_0 resolution for the tracks from b- and light-jets.

obtain a stable RMS[3] width of the residual about mean. On comparing the estimated d_0 resolution values between r21.9 and r20.20, differences are observed in A01, A05, B01 and B05 categories.

η region	Category	Description
$\lvert\eta\rvert < 1$	A01	No hit in 1st pixel layer; no hit in 2^{nd} pixel layer
$\lvert\eta\rvert < 1$	A05	No hit in 1^{st} pixel layer but expected
$\lvert\eta\rvert < 1$	A06	No hit in 1^{st} pixel layer and not expected
$\lvert\eta\rvert < 1$	A07	No hit in 2^{nd} pixel layer but expected
$\lvert\eta\rvert < 1$	A08	No hit in 2^{nd} pixel layer and not expected
$\lvert\eta\rvert < 1$	A09+10	Total number of shared hits in pixel layers > 0
$\lvert\eta\rvert < 1$	A14A	All other tracks; number of hits in 1st pixel layer = 1
$\lvert\eta\rvert < 1$	A14B	All other tracks; number of hits in 1st pixel layer ≥ 2
$1 < \lvert\eta\rvert < 2$	B01	No hit in 1^{st} pixel layer; no hit in 2^{nd} pixel layer
$1 < \lvert\eta\rvert < 2$	B05	No hit in 1st pixel layer but expected
$1 < \lvert\eta\rvert < 2$	B09+10	Total number of shared hits in pixel layers > 0
$1 < \lvert\eta\rvert < 2$	B11	Total number of shared hits in strip layers > 1
$1 < \lvert\eta\rvert < 2$	B14	All other tracks
$\lvert\eta\rvert > 2$	Zone 1	$\gamma \leq 0.001$ MeV^{-1}
$\lvert\eta\rvert > 2$	Zone 2	$0.001 < \gamma \leq 0.002$ MeV^{-1}
$\lvert\eta\rvert > 2$	Zone 3	$0.002 < \gamma \leq 0.003$ MeV^{-1}
$\lvert\eta\rvert > 2$	Zone 4	$\gamma \geq 0.003$ MeV^{-1}

Table 7 – Description of the track categories used by IP3D algorithms for the ITk layout.

[3] The standard deviation of the distribution



(a) (b)

Figure 52 – Distributions of 17 Track categories as function of η for (a) r20.20 and (b) r21.9 in $t\bar{t}$ events with an average pile-up of 200.

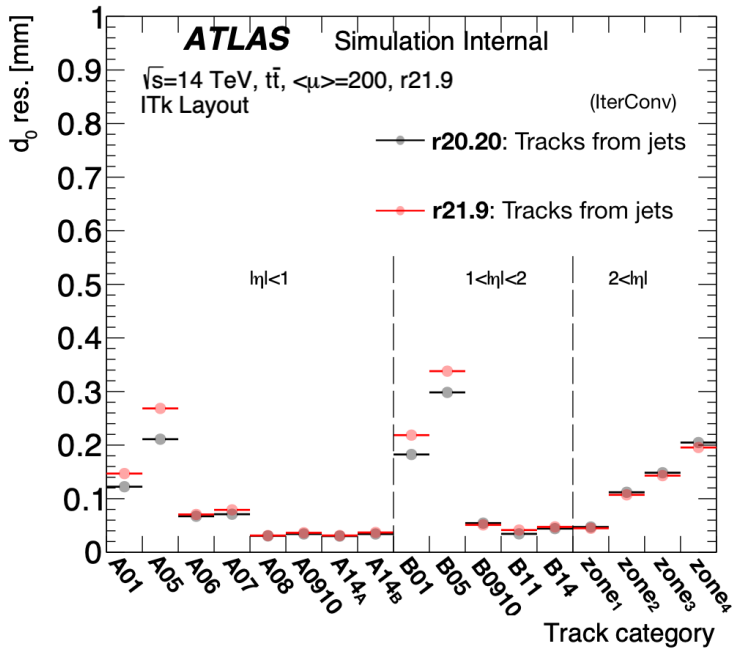


Figure 53 – d_0 resolution as function of the track category in the inclusive track p_T, compared between r20.20 (black) and r21.9 (red) in $t\bar{t}$ events with an average pile-up of 200.

Additional checks are done to study the d_0 resolution for each of those categories, shown in Figure 54. Overall, the differences in the IterConv estimates of the d_0 resolution are found to be compatible with the low statistics in r21.9, and better consistency is observed with RMS and Gauss[4] values. In order to ensure proper statistical coverage, uncertainty is assigned to the IterConv method.

[4] The width of a Gaussian fit

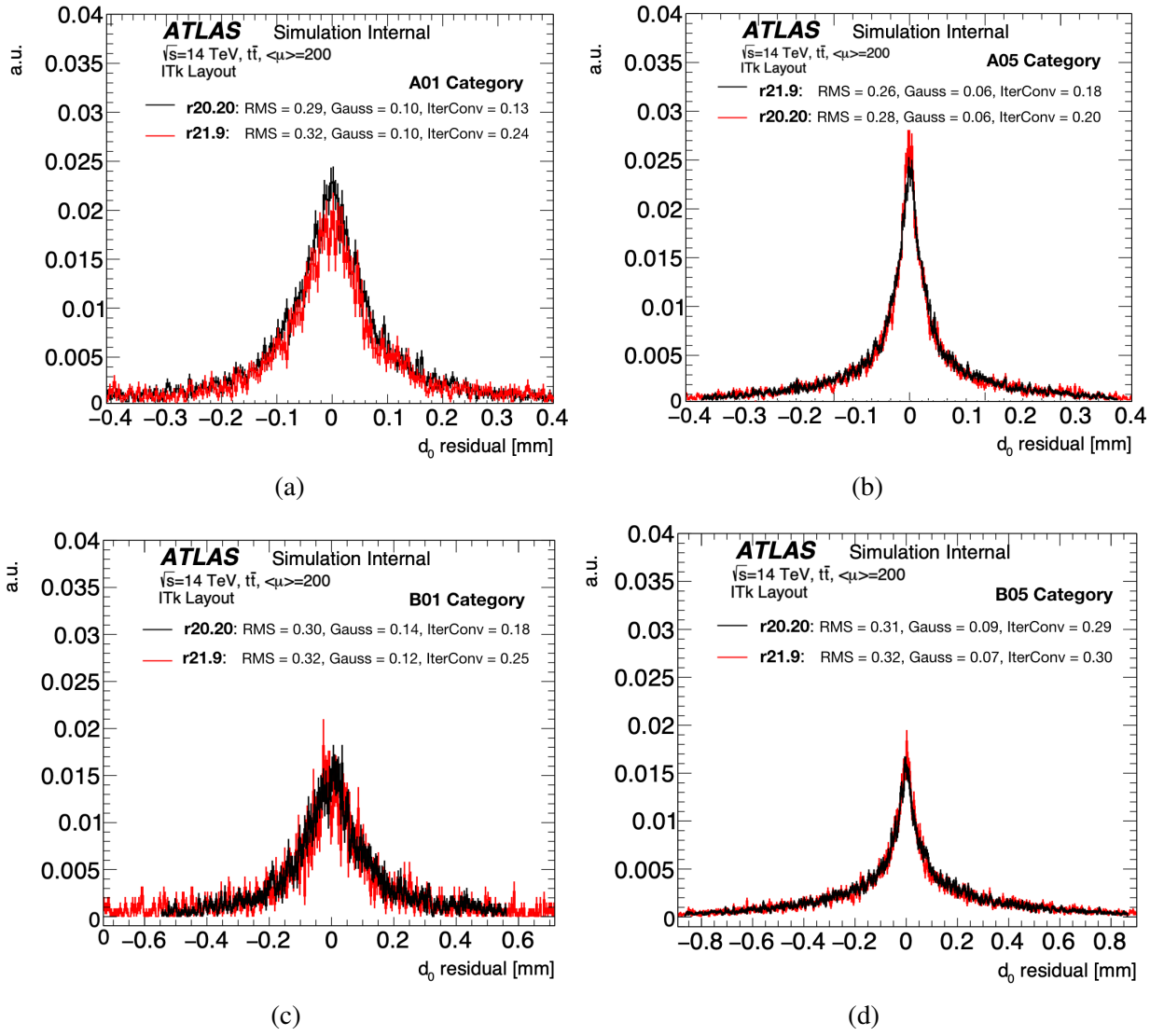
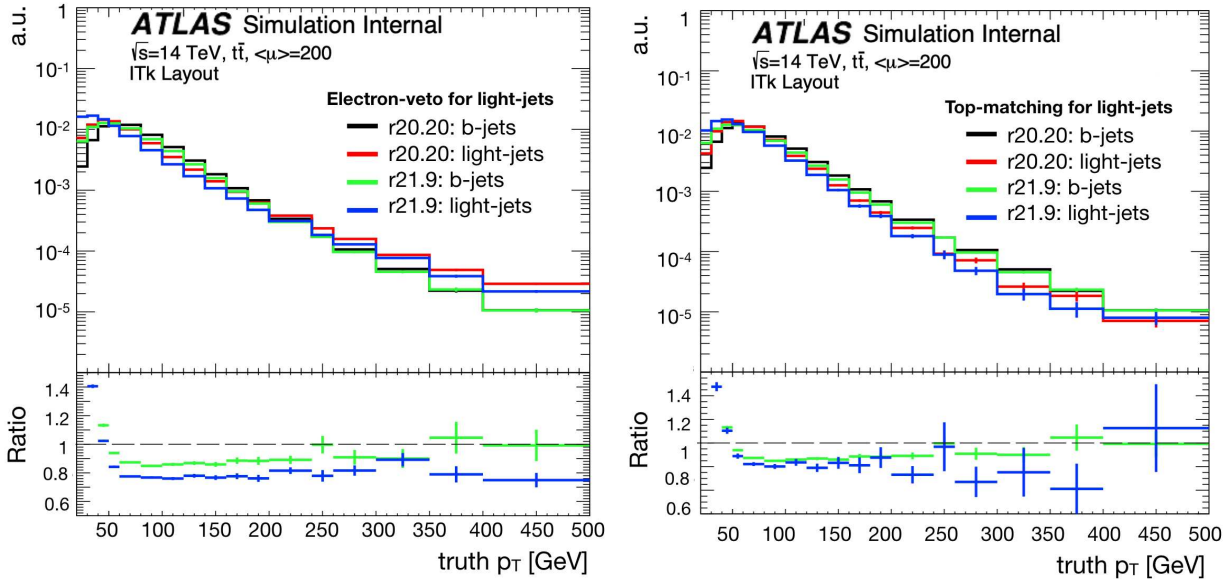


Figure 54 – d_0 residual compared between r20.20 (black) and r21.9 (red) for (a) A01 (b) A05, (c) B01 and (d) B05 track categories in $t\bar{t}$ events with an average pile-up of 200. The IterConv , RMS and Gauss values are used to estimate the d_0 resolution for the track category.

3.3.3 Jet kinematics studies

The performance of *b*-tagging is impacted by jet kinematics, particularly by jet p_T and η. To have a similar baseline for the *b*-tagging performance, it is necessary to analyze and compare these jet kinematic distributions between r20.20 and r21.9 at both truth and reconstructed (reco) levels. Deeper investigations on jet kinematics were carried on by the jet performance group.

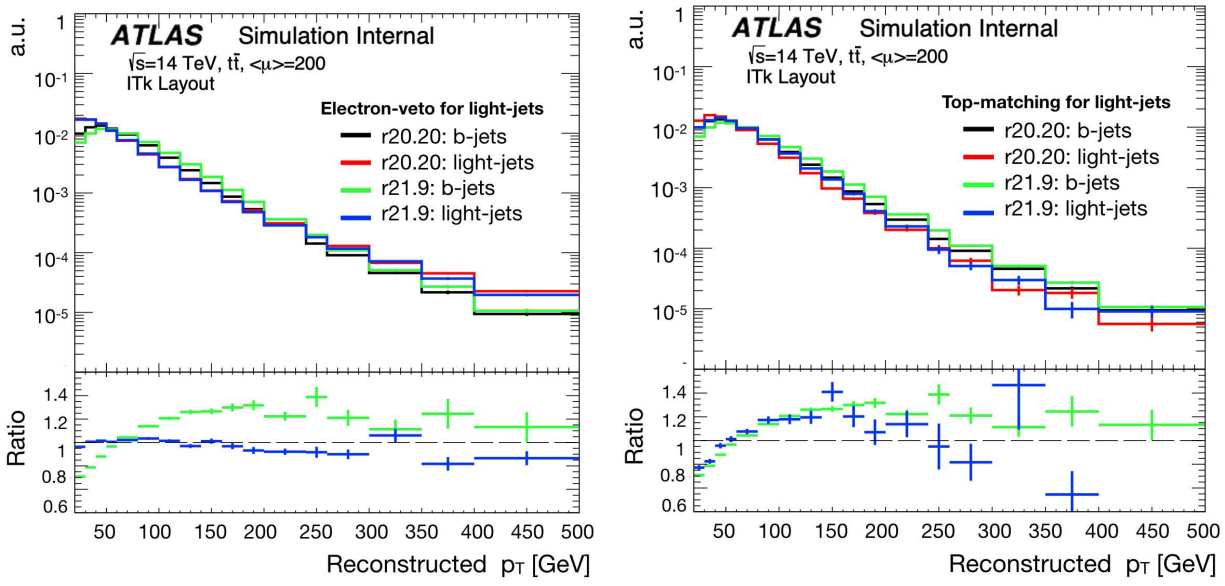
After applying the selections described in Section 3.3.1, the jet distributions are studied separately for *b*-jets and light-jets, with electron veto as the default matching scheme for light-jets. Similar studies using the top-matching scheme for the light-jets are also carried out for additional checks. Figure 55, shows the truth p_T spectrum after matching with reconstructed jets, where light-jets are matched using electron veto and top-matching. Similarly, the reconstructed jet p_T and jet η spectrum are shown in Figure 56 and Figure 57, respectively.



(a) (b)

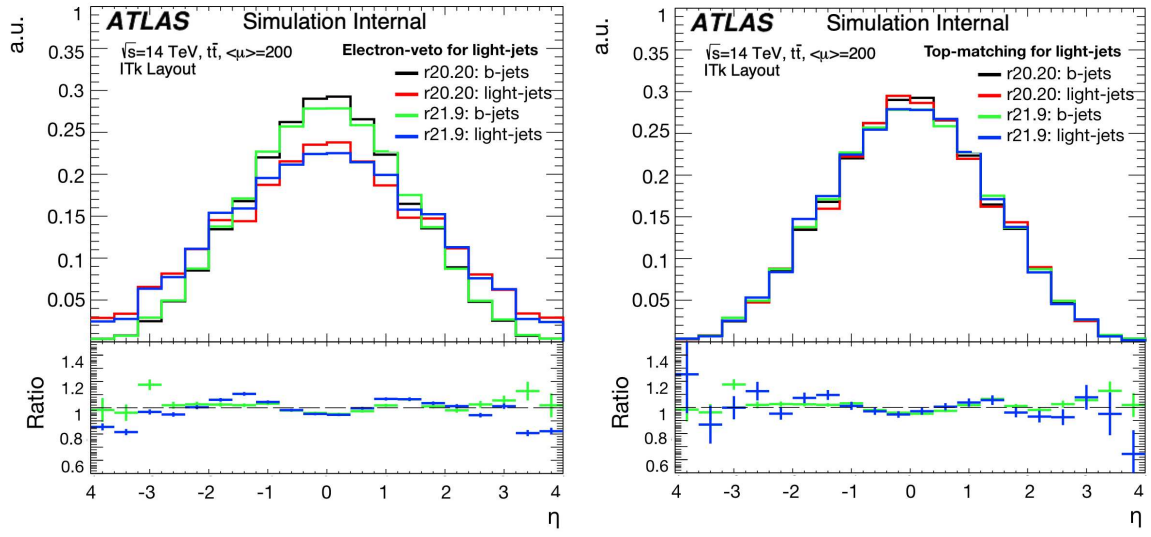
Figure 55 – Truth-jet p_T distribution after matching with reconstructed jets, compared between r20.20 and r21.9, where light-jets are selected using (a) default matching and (b) top-matching in $t\bar{t}$ events with an average pile-up of 200. The green (blue) ratio distribution corresponds to the r20.20 and r21.9 ratios of the *b*-jets (light-jets) distribution.

Between the releases, some inconsistencies are observed in truth p_T and reconstructed jet p_T distributions, while the jet η distribution appears to be consistent to some extent, especially with top-matching selection. In truth p_T distribution, the ratio between r20.20 and r21.9 for truth $p_T > 50$ Gev appears flat, indicating a possible threshold effect associated with the reconstructed p_T selection cut of $p_T > 20$ GeV. The fact that the discrepancy in the reconstructed p_T between the releases appears different suggests a discrepancy in the jet calibration between the releases. Some of the differences



(a) (b)

Figure 56 – Reconstructed jet p_T distribution compared between r20.20 and r21.9, where light-jets are selected using (a) default matching and (b) top-matching in $t\bar{t}$ events with an average pile-up of 200. The green (blue) ratio distribution corresponds to the r20.20 and r21.9 ratios of the *b*-jets (light-jets) distribution.



(a) (b)

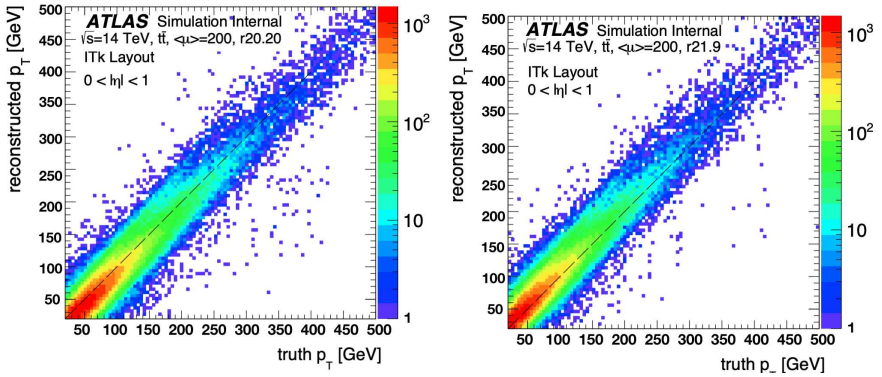
Figure 57 – Reconstructed jet p_T distribution compared between r20.20 and r21.9, where light-jets are selected using (a) default matching and (b) top-matching in $t\bar{t}$ events with an average pile-up of 200. The green (blue) ratio distribution corresponds to the r20.20 and r21.9 ratios of the *b*-jets (light-jets) distribution.

in the selections, in particular due to the $p_T > 20$ GeV requirement combined with different p_T response, could also have impacted the truth p_T distribution.

The correlation between truth p_T and reconstructed p_T is compared between r20.20 and r21.9 in different η bins, shown in Figure 58. A slight under-calibration is observed in r20.20, consistent in all the η bins. The jet energy response is also investigated between the releases in different η bins, as shown in Figure 59. Visible differences in jet energy response are observed. Due to the different jet selection efficiency associated with the reconstructed $p_T > 20$ GeV threshold, the truth p_T spectrum is expected to be impacted by the different jet energy response.

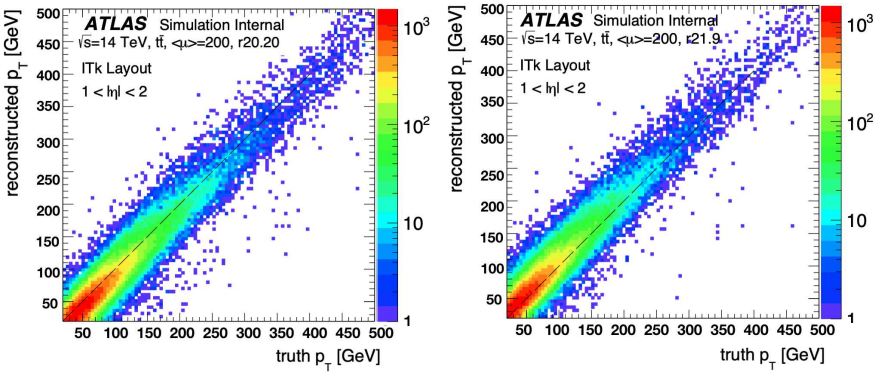
The observed differences in jet kinematics are significant and must be mitigated in order to disentangle between releases from the jet kinematic reconstruction and b-tagging performance.. Based on the feedback received from the jet calibration group, the implementation of event variables used for jet pile-up energy subtraction has evolved between r20.20 and r21.9, making the jet calibration derived in r20.20 obsolete. In r21.9, new calibrations were to be re-derived, but were not yet available for the studies shown in this thesis.

Given the expected p_T dependence in *b*-tagging performance, it is best not to rely on reconstructed p_T for the investigation of the *b*-tagging performance shown here. Instead, truth p_T will be used after re-weighting in r21.9, shown in Figure 60, to match the truth p_T distribution in r20.20. The default jet selection defined using reconstructed p_T is replaced with the corresponding truth p_T (re-weighted) cut of truth $p_T > 20$ GeV. It was observed that, the truth p_T cut showed similar performance when using reconstructed p_T selection cut. This re-weighting may not account for all differences between r20.20 and r21.9 in jet reconstruction. The p_T-dependent track-jet association may still be affected and influence the subsequent *b*-tagging performance. Some proper comparisons would ideally be possible after a new 21.9 jet calibration is released.



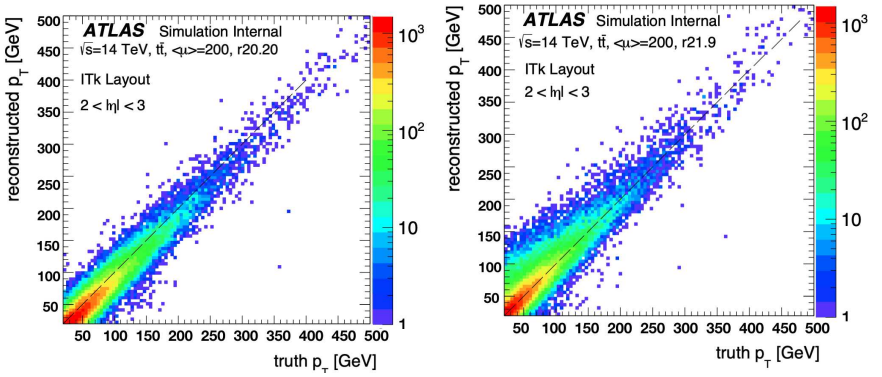
(a)

(b)



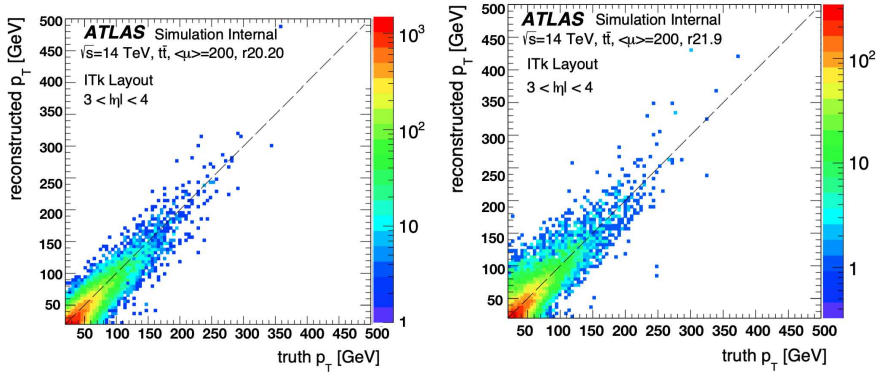
(c)

(d)



(e)

(f)



(g)

(h)

Figure 58 – Reconstructed p_T vs truth p_T correlation distribution in r20.20 and r21.9 for jets in $0 < |\eta| < 1$, $1 < |\eta| < 2$, $2 < |\eta| < 3$, and $3 < |\eta| < 4$ regions in $t\bar{t}$ events with an average pile-up of 200.

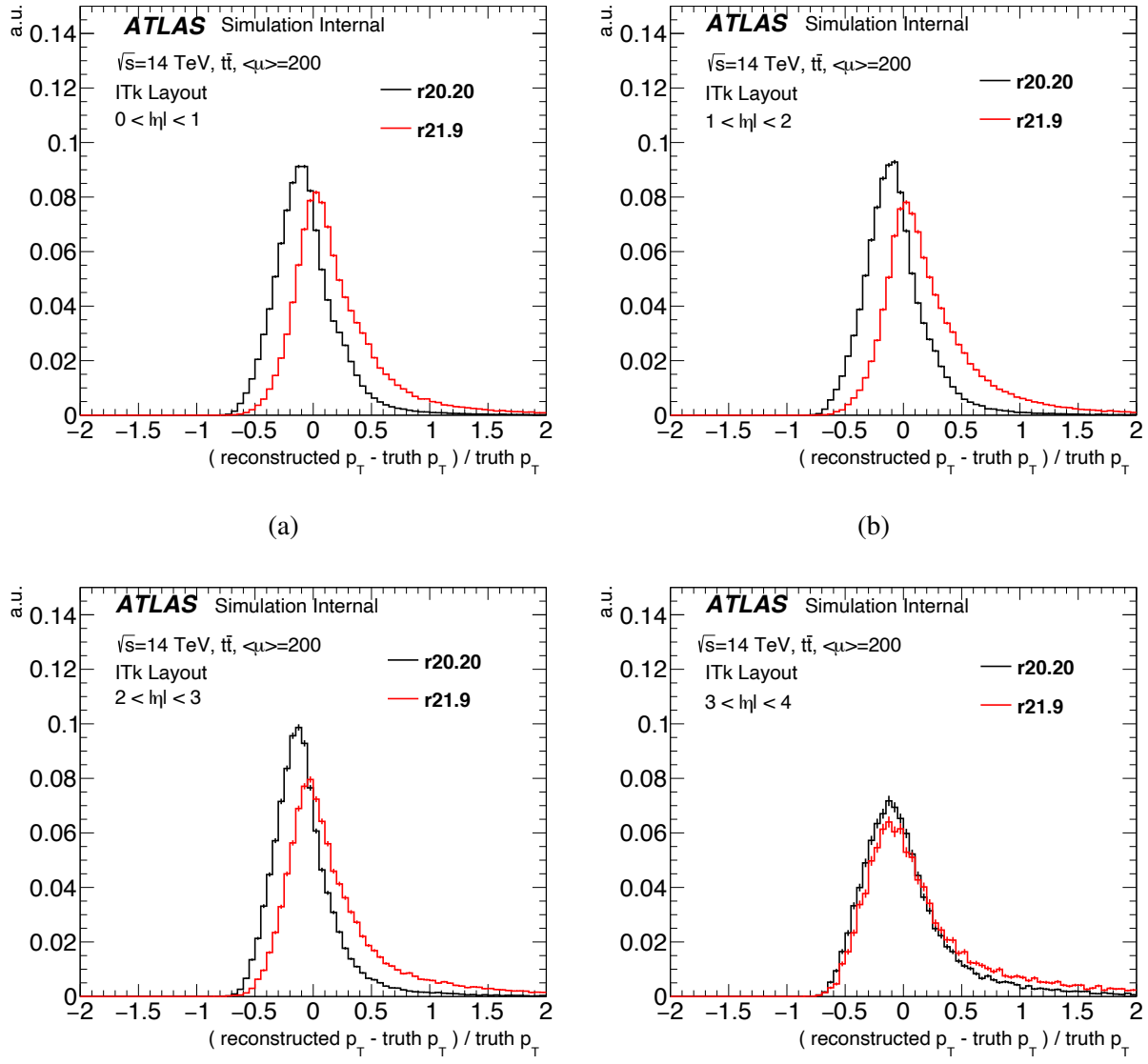
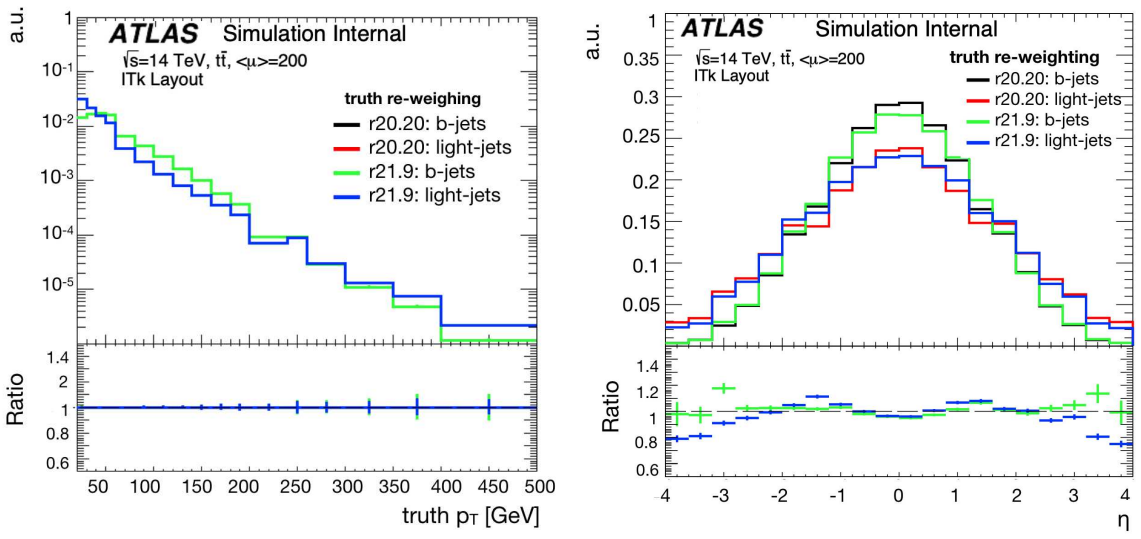


Figure 59 – Jet energy response distribution for the jets compared between r20.20 (black) and r21.9 (red) in (a) $0 < |\eta| < 1$, (a) $1 < |\eta| < 2$, (a) $2 < |\eta| < 3$ and (a) $3 < |\eta| < 4$ region in $t\bar{t}$ events with an average pile-up of 200.



(a) (b)

Figure 60 – The distribution of (a) truth p_T and (b) η for *b*- and light-jets, compared between release 20.20 and 21.9 after truth p_T re-weighting in r21.9 in $t\bar{t}$ events with an average pile-up of 200.

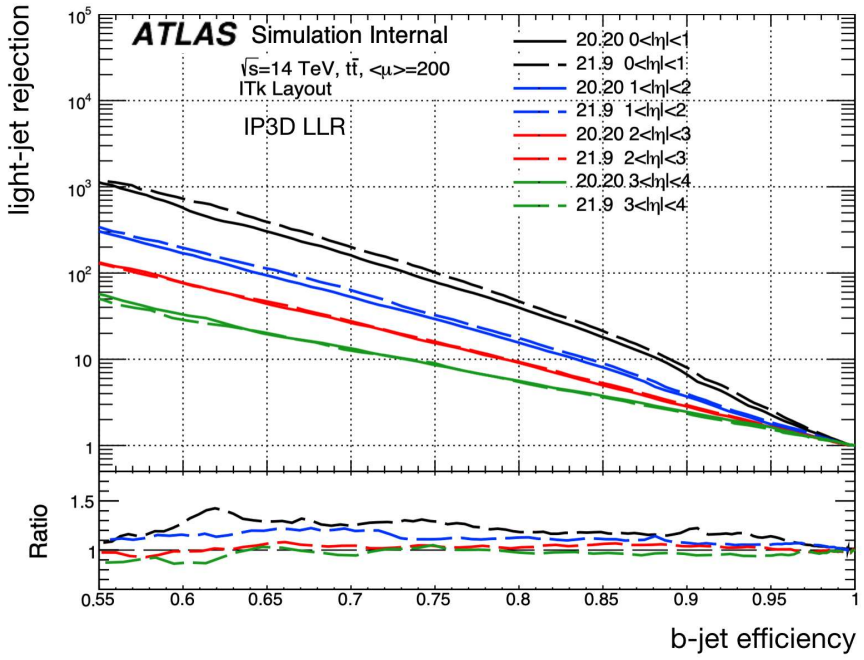
3.3.4 IP3D performance

The expected *b*-tagging performance in r21.9 in terms of ROC curves for the IP3D tagger is shown in Figure 61, compared with r20.20 in exclusive η and truth p_T regions.

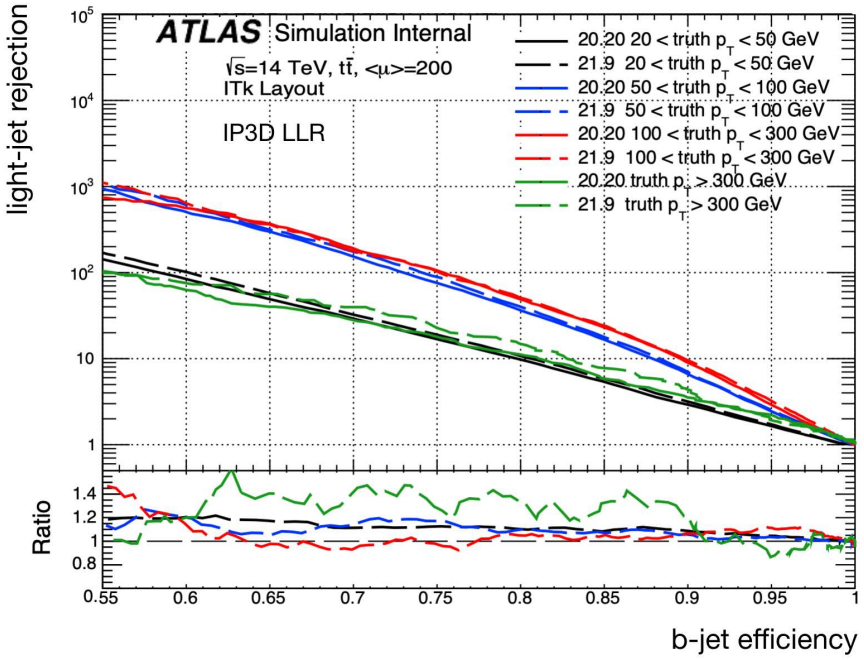
These results show that the *b*-tagging implementation for the IP3D algorithm developed in r21.9 restores the developments in r20.20 within 20-30%, which is considered reasonable, given the differences that may arise from different jet calibrations which are not corrected by simple truth p_T re-weighting. Largest discrepancies are observed in the central region and at high p_T.

3.3.5 SV1 performance

The expected *b*-tagging performance in terms of ROC curves for the SV1 tagger is shown in Figure 62, comparing the performance obtained with r21.9 to r20.20 in exclusive η and truth p_T regions. The reference histograms templates used for the SV1 LLR formalism is also derived from simulations in r20.20, similar to IP3D. The performance looks consistent between releases, within 20-30% for $|\eta| < 3$ region, for working points away from the ROC curve endpoints. Those endpoints correspond to the bare SV reconstruction rate, with no SV1 discriminant selection applied. The SV reconstruction rate is increased by a few percents for *b*-jets in the $|\eta| < 2$ region and the light jet rejection at the endpoint is identical. There are more significant differences observed in the region $|\eta| > 3$. Such differences were not observed in the IP3D performance, which uses a tighter $|z_0 \sin\theta|$ selection than SV1 (< 1.5 mm instead of < 25 mm). As a result, the SV1 performance in the higher



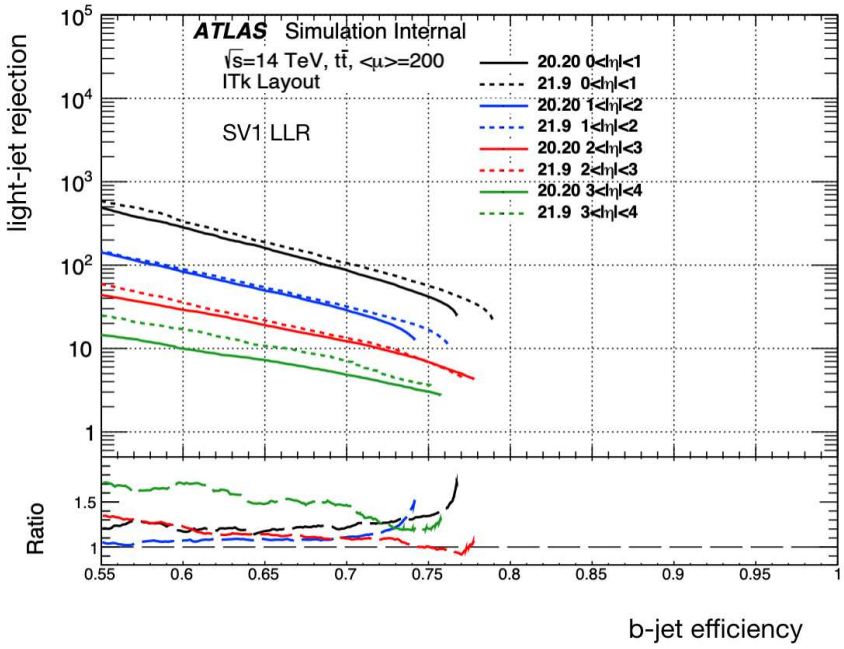
(a)



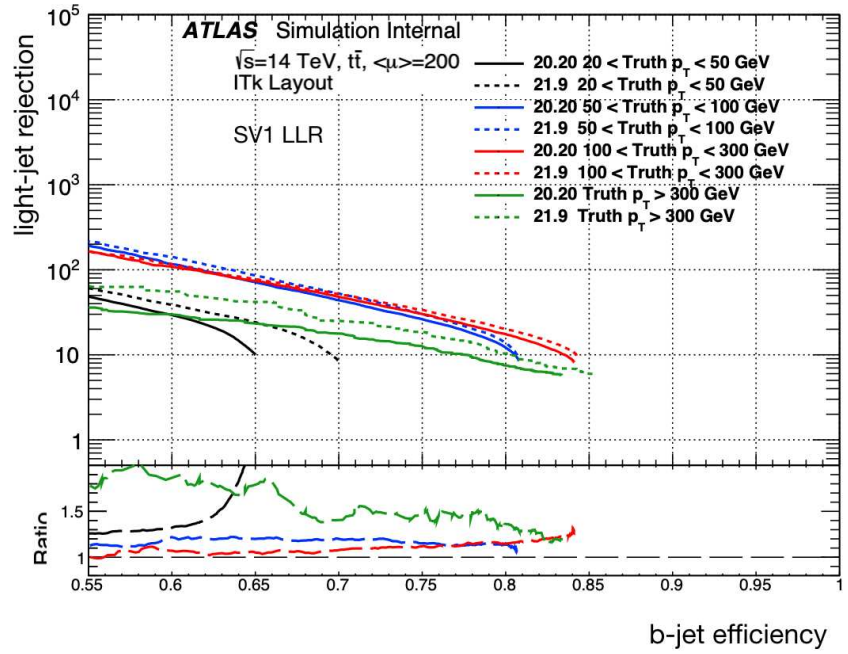
(b)

Figure 61 – Light-jet rejection vs b-tagging efficiency in exclusive (a) η and (b) truth p_T regions, compared between r20.20 and r21.9 for the IP3D tagging algorithm evaluated in $t\bar{t}$ events with an average pile-up of 200.

η region required further investigation.



(a)



(b)

Figure 62 – Light-jet rejection vs b-tagging efficiency in exclusive (a) η and (b) truth p_T regions, compared between r20.20 and r21.9 for the SV1 tagging algorithm evaluated in $t\bar{t}$ events with an average pile-up of 200.

Further studies are done in order to understand the source of improved *b*-tagging performance for the SV1 algorithm at larger η. To begin with, studies are performed between releases to determine whether the differences are due to different SVs reconstruction rate, is defined as the fraction of jets with a reconstructed SV. Figure 63a shows the SV reconstruction rate as a function of η compared between releases separately for *b*- and light-jets, where a large difference is observed for light-jets. Figure 63b and Figure 63c shows SV reconstruction rate as function of truth p_T inclusive in η and specifically for $3 < |\eta| < 4$ region, respectively. Overall, the difference in r21.9 is mostly due to the lower SV reconstruction rate for light jets, which leads to better rejection at the SV1 ROC curve endpoint. The differences observed in the SV reconstruction rate is also investigated further.

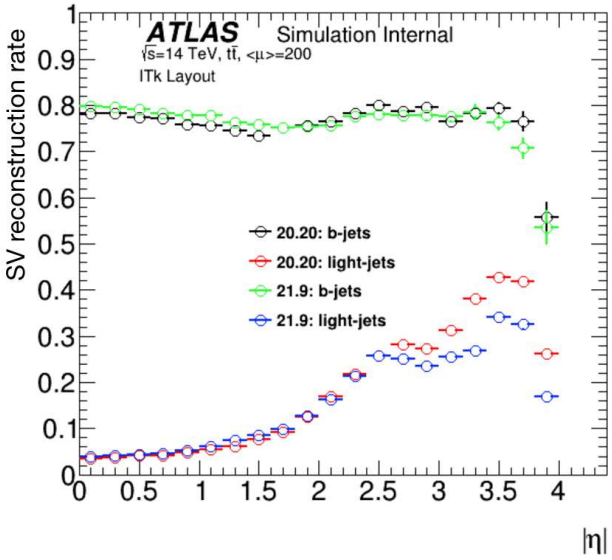
Figure 64 shows the SV radius ($= \sqrt{(SV_x)^2 + (SV_y)^2}$) as a function of SV_z for the light-jets, highlighting the position of SVs reconstructed close to the beam pipe (R = 24 mm) and the first ITk pixel layer (R = 39 mm), where no significant difference is observed between the releases. The difference in the forward region was still being investigated at the time of this study. The difference in detailed ITk simulation in the forward region, as well as the difference in detailed implementation of the SSVF reconstruction algorithm in r21.9, could have been the potential causes. As previously indicated, the different conventions used in r20.20, highlighted in the next-to-innermost pixel layer hits distribution between the releases, could also contribute to the differences in the forward region. The observed differences in jet kinematics are significant and must be mitigated in order to disentangle between releases from the jet kinematic reconstruction and *b*-tagging performance.

3.4 *b*-tagging re-optimisation studies in ITk

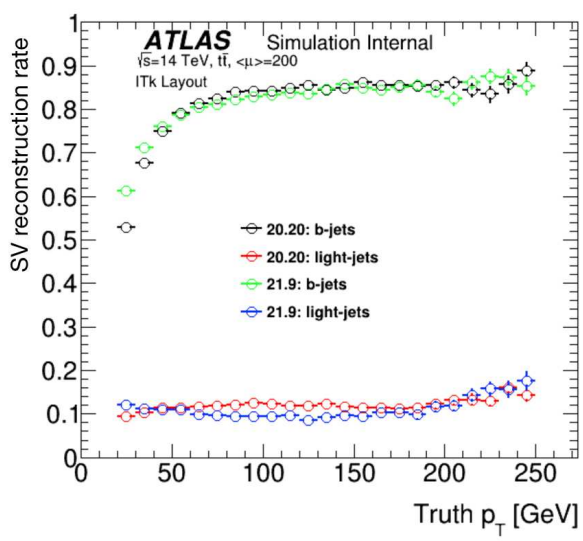
This section will focus on the studies I conducted to further optimise the low-level *b*-taggers, particularly in developing new track categories for IP3D and implementing material rejection for ITk to improve SV1 performance, discussed in Section 3.4.1 and Section 3.4.2, respectively. The IP3D and SV1 performances are then compared to those obtained with the default implementations, presented in Section 3.3. Due to the larger statistical samples available in r20.20, the references used here are still derived from r20.20. As a result, the final results will be sub-optimal, demonstrating the lower limit of the achievable performance.

3.4.1 IP3D performance with re-optimised categorisation

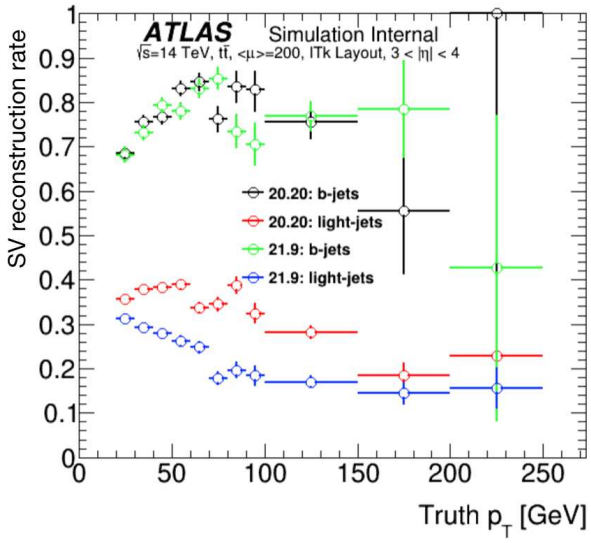
The re-optimisation of the IP3D track categorises are performed to better exploit the expected p_T-dependence of the IP resolution in the central region with $|\eta| < 2$ and the hit-content dependence in the forward region with $|\eta| > 2$. Table 8 describes the updated track categories. The default



(a)

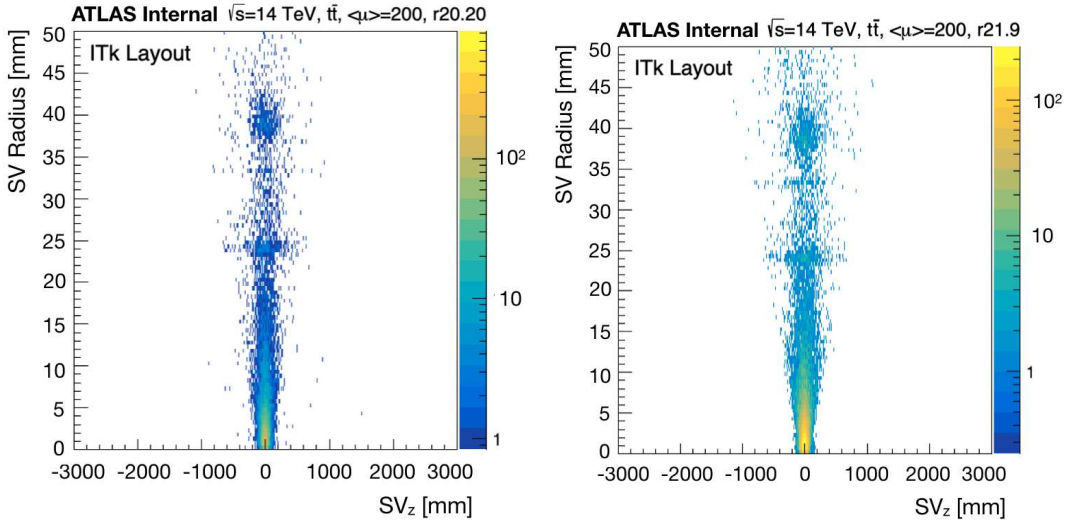


(b)



(c)

Figure 63 – SV reconstruction rate for *b*- and light-jets, as a function of (a) η (b) truth p_T in inclusive η and (c) truth p_T in $3 < |\eta| < 4$ in $t\bar{t}$ events with an average pile-up of 200.



(a) (b)

Figure 64 – Distribution of SV radius as function of SV$_z$ for light-jets for (a) r20.20 and (b) r21.9 in $t\bar{t}$ events with an average pile-up of 200.

categories, listed in Table 7, $|\eta| < 1$ region (region A) already strongly exploits the detailed hit content of the tracks.

Hence, the idea is to make categories in $1 < |\eta| < 2$ (region B) and $|\eta| > 2$ (referred as region C here) regions, similarly detailed with some merging of the categories where statistics are very low, especially in the forward region. As previously stated, there is a correlation between the IP resolution and the γ parameter, shown in Figure 51. The default track categories in the forward region already exploited this dependence by splitting the track categories by γ values. This motivated to similarly split the sensitive categories in region A and region B, i.e. A14 and B14, using γ values. Figure 65 shows the distribution of the tracks from b- and light-jets in these updated track categories. Figure 66 illustrates the d_0 resolution as a function of new track categories, where the updated track categorization makes it easier to isolate tracks with poor d_0 resolution and avoid contamination in the most sensitive categories impacting in the IP3D performance improvement.

Figure 67 shows the b-tagging performance in terms of ROC curves for the IP3D tagger in exclusive η bins, comparing the performance with updated and default track categories. In $|\eta| < 2$ region, a significant improvement of up to 50% is seen with the new track categorization.

3.4.2 SV1 performance with material rejection

SVs can possibly be reconstructed when the particles interact with the detector material. Figure 68 shows the reconstructed SV radius distribution for b- and light-jets. Visible peaks are observed at

η region	Category	Description
$\lvert\eta\rvert < 1$	A01	No hit in 1st pixel layer, no hit in 2$^{\text{nd}}$ pixel layer
$\lvert\eta\rvert < 1$	A02	No hit in 1$^{\text{st}}$ pixel layer but expected
$\lvert\eta\rvert < 1$	A03	No hit in 1$^{\text{st}}$ pixel layer and not expected
$\lvert\eta\rvert < 1$	A04	No hit in 2$^{\text{nd}}$ pixel layer but expected
$\lvert\eta\rvert < 1$	A05	No hit in 2$^{\text{nd}}$ pixel layer and not expected
$\lvert\eta\rvert < 1$	A06	Total number of shared hits in pixel layers > 0
$\lvert\eta\rvert < 1$	A07	Number of outliers[5] in 1$^{\text{st}}$ and 2$^{\text{nd}}$ pixel layer > 0
$\lvert\eta\rvert < 1$	A08	Number of spoilt[6] hits ≥ 2
$\lvert\eta\rvert < 1$	A14$_1$	$0.75{\cdot}10^{-3}$ MeV$^{-1} \leq \gamma$
$\lvert\eta\rvert < 1$	A14$_2$	$0.5{\cdot}10^{-3} < \gamma \leq 0.75{\cdot}10^{-3}$ MeV^{-1}
$\lvert\eta\rvert < 1$	A14$_3$	$0.25{\cdot}10^{-3} < \gamma \leq 0.5{\cdot}10^{-3}$ MeV^{-1}
$\lvert\eta\rvert < 1$	A14$_4$	$\gamma \leq 0.25{\cdot}10^{-3}$ MeV^{-1}
$1 < \lvert\eta\rvert < 2$	B01	No hit in 1$^{\text{st}}$ pixel layer and hit 2$^{\text{nd}}$ pixel layer
$1 < \lvert\eta\rvert < 2$	B02	No hit in 1$^{\text{st}}$ pixel layer but expected
$1 < \lvert\eta\rvert < 2$	B03	No hit in 1$^{\text{st}}$ pixel layer and not expected
$1 < \lvert\eta\rvert < 2$	B04	No hit in 2$^{\text{nd}}$ pixel layer but expected
$1 < \lvert\eta\rvert < 2$	B05	No hit in 2$^{\text{nd}}$ pixel layer and not expected
$1 < \lvert\eta\rvert < 2$	B06	Total number of shared pixel hits > 0
$1 < \lvert\eta\rvert < 2$	B07	Number of outliers in 1$^{\text{st}}$ and 2$^{\text{nd}}$ pixel layer > 0
$1 < \lvert\eta\rvert < 2$	B08	Number of pixel spoilt hits $\geq$2
$1 < \lvert\eta\rvert < 2$	B14$_1$	$1.5{\cdot}10^{-3}$ MeV$^{-1} \leq \gamma$
$1 < \lvert\eta\rvert < 2$	B14$_2$	$1{\cdot}10^{-3} < \gamma \leq 1.5{\cdot}10^{-3}$ MeV^{-1}
$1 < \lvert\eta\rvert < 2$	B14$_3$	$0.5{\cdot}10^{-3} < \gamma \leq {\cdot}10^{-3}$ MeV^{-1}
$1 < \lvert\eta\rvert < 2$	B14$_4$	$\gamma \leq 0.5{\cdot}10^{-3}$ MeV^{-1}
$\lvert\eta\rvert > 2$	C01	No hit in 1$^{\text{st}}$ and 2$^{\text{nd}}$ pixel layer
$\lvert\eta\rvert > 2$	C020304	No hit in 1$^{\text{st}}$ pixel layer or no hit 2$^{\text{nd}}$ pixel layer but expected
$\lvert\eta\rvert > 2$	C05	No hit 2$^{\text{nd}}$ pixel layer and not expected
$\lvert\eta\rvert > 2$	C06	Total number of shared pixel hits > 0
$\lvert\eta\rvert > 2$	C07	Number of outliers in 1$^{\text{st}}$ and 2$^{\text{nd}}$ pixel layer > 0
$\lvert\eta\rvert > 2$	C0809	Number of pixel spoilt hits $\geq$2
$\lvert\eta\rvert > 2$	C14$_1$	$3{\cdot}10^{-3}$ MeV$^{-1} \leq \gamma$
$\lvert\eta\rvert > 2$	C14$_2$	$2{\cdot}10^{-3} < \gamma \leq 3{\cdot}10^{-3}$ MeV^{-1}
$\lvert\eta\rvert > 2$	C14$_3$	${\cdot}10^{-3} < \gamma \leq 2{\cdot}10^{-3}$ MeV^{-1}
$\lvert\eta\rvert > 2$	C14$_4$	$\gamma \leq 1{\cdot}10^{-3}$ MeV^{-1}

Table 8 – Description of the re-optimised track categories for the IP3D algorithms with the ITk layout.

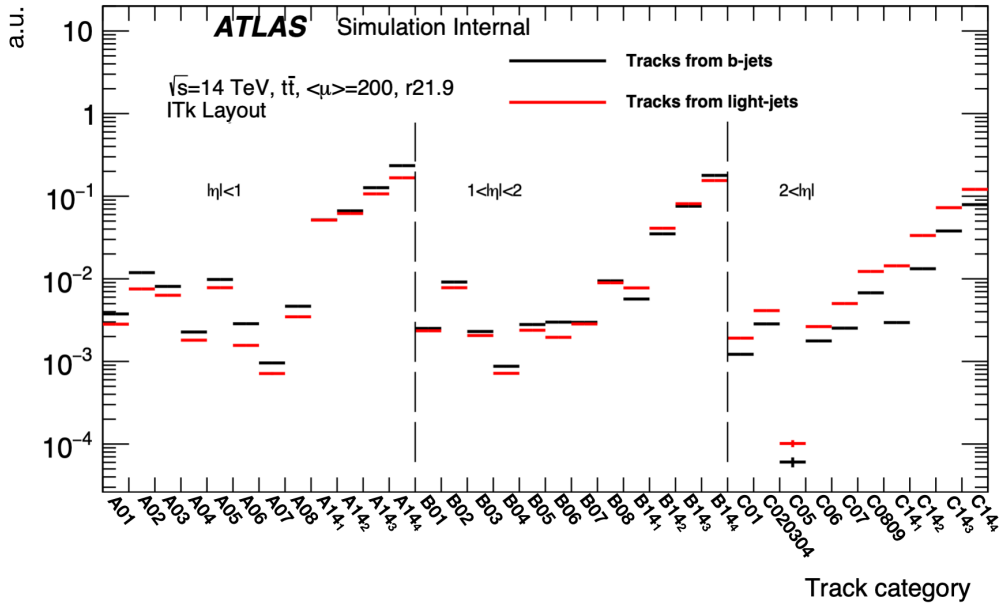


Figure 65 – Distribution of updated track categories for the tracks from *b*- and light-jets in $t\bar{t}$ events with an average pile-up of 200, in r21.9, obtained with the ITk layout.

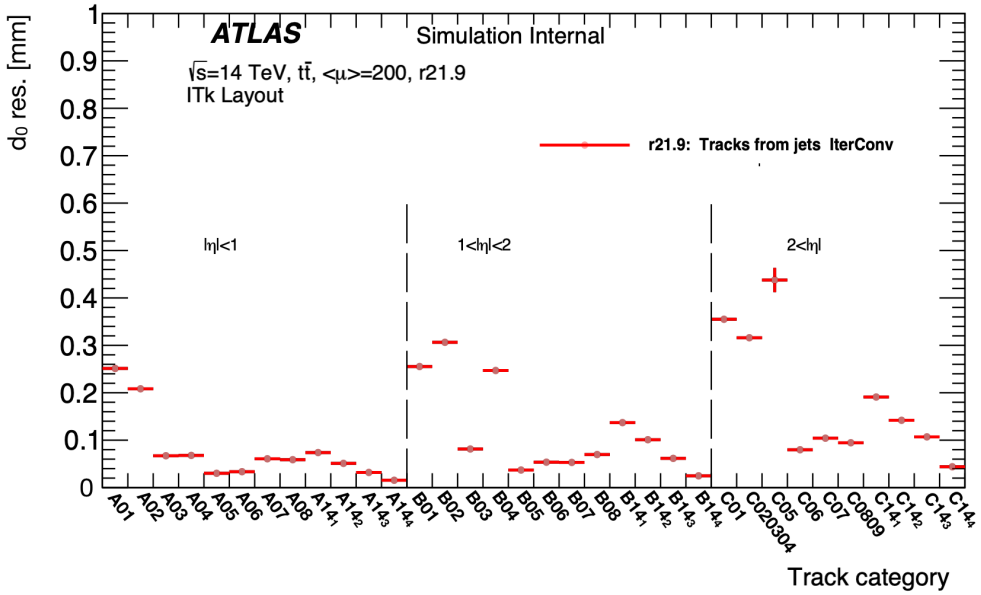


Figure 66 – d_0 resolution as function of the updated track categories in $t\bar{t}$ events with an average pile-up of 200 in r21.9, obtained with the ITk layout.

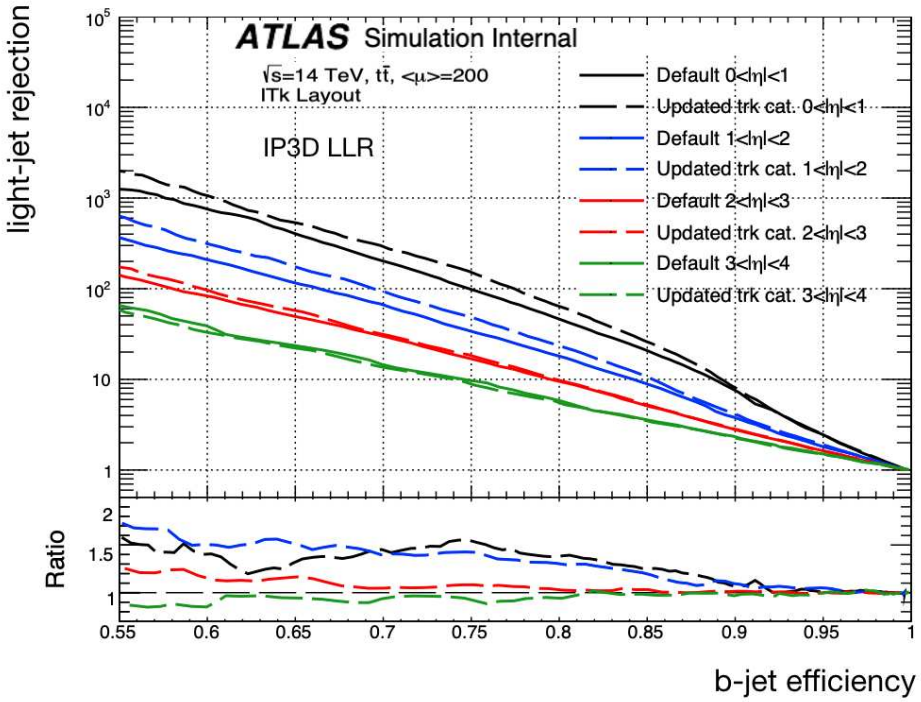


Figure 67 – Light-jet rejection vs b-tagging efficiency in exclusive η regions, with the default and the updated track categories, for the IP3D tagging algorithm evaluated in $t\bar{t}$ events with an average pile-up of 200 in r21.9, obtained with the ITk layout.

certain SV radii in light jets, whereas such peaks are not observed or are less significant in *b*-jets. The peaks in light-jets correspond to material interactions near the pixel layers and represent a non-negligible source of reconstructed SV for light-jets. Thus, it is needed to reject these SVs reconstructed near pixel layers or other mechanical structures, to reduce the source of background and improve the SV1 performance. Some material rejection is already hard-coded in the Run 2 detector software, but the implementation is incompatible with the more complex ITk geometry. I performed studies using a *tracking material map* to implement SV rejection compatible with the ITk layout. This map is a 2D histogram distribution of beam pipe and pixel sensor positions generated directly from the ITk detector layout as implemented in the simulation software. Figure 69 shows the the reconstructed SV radius distribution of tracks from light-jets after material rejection is applied, where peaks observed previously at certain SV radii in light-jets, are no longer observed.

Figure 70 depicts the SV reconstruction efficiency as function of $|\eta|$ for *b*- and light-jets, with light-jets (*b*-jets) showing a $\sim$5% ($\sim$1.5%) reduction in SV reconstruction efficiency in the central region. The overall impact on the SV1 performance with material rejection is illustrated in Figure 71, for exclusive η regions, compared with default performance (without material rejection). The light-jet rejection is increased by 5 to 20% using material rejection with the ITk detector.

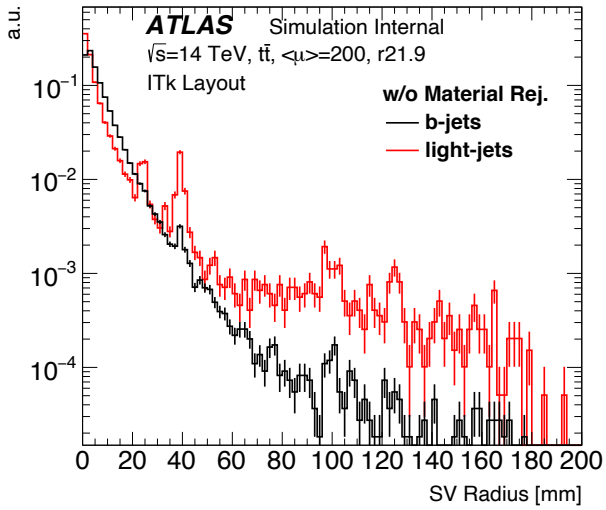


Figure 68 – SV radius distribution of tracks from *b*- and light-jets in $t\bar{t}$ events with an average pile-up of 200 in r21.9, obtained with the ITk layout.

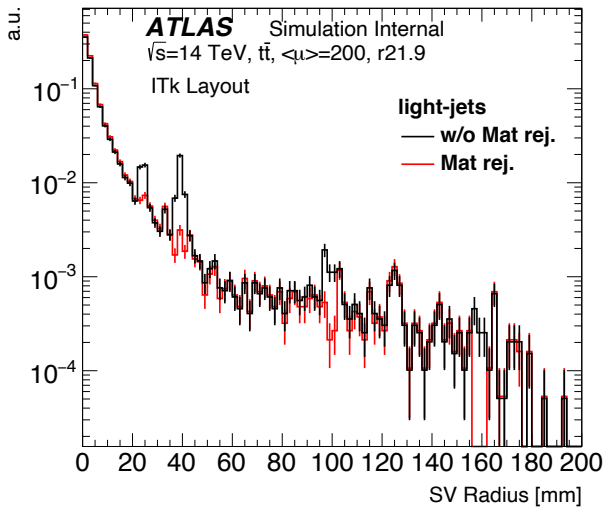


Figure 69 – SV radius distribution of tracks from light-jets when material rejection is applied in $t\bar{t}$ events with an average pile-up of 200 in r21.9, obtained with the ITk layout.

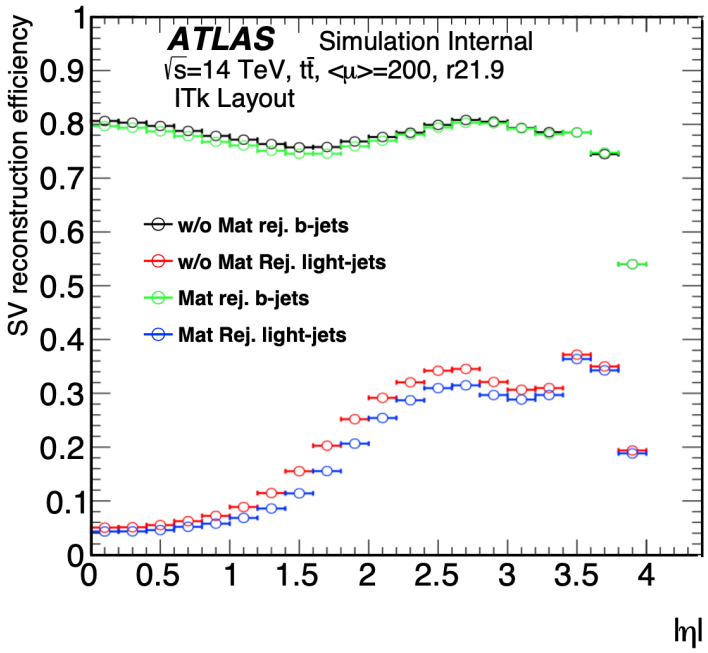


Figure 70 – SV Reconstruction efficiency as a function of η, with and without material rejection, for *b*- and light-jets in $t\bar{t}$ events with an average pile-up of 200, obtained with the ITk layout.

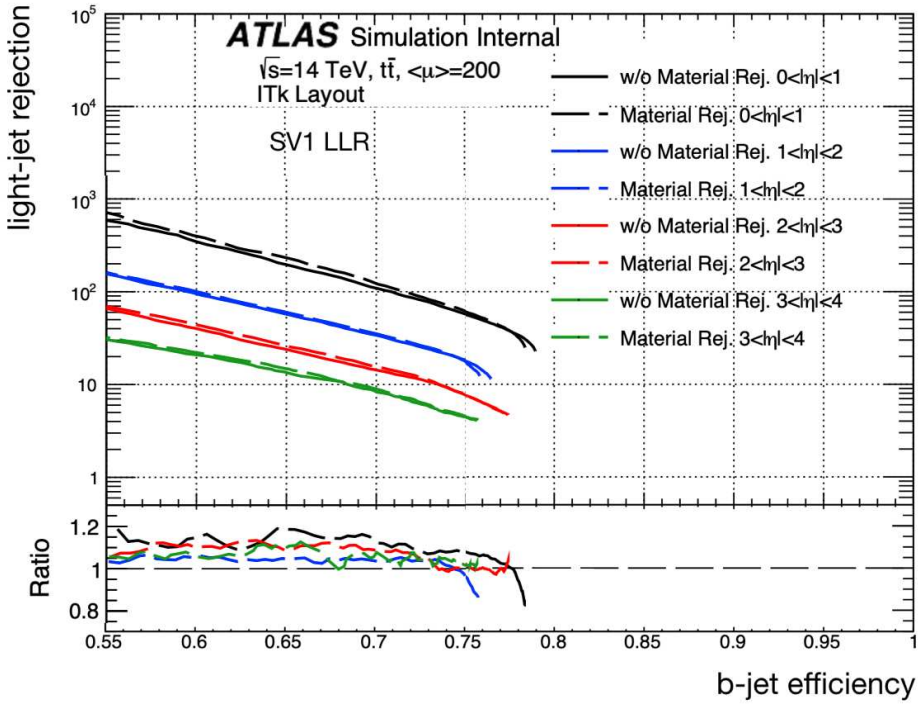
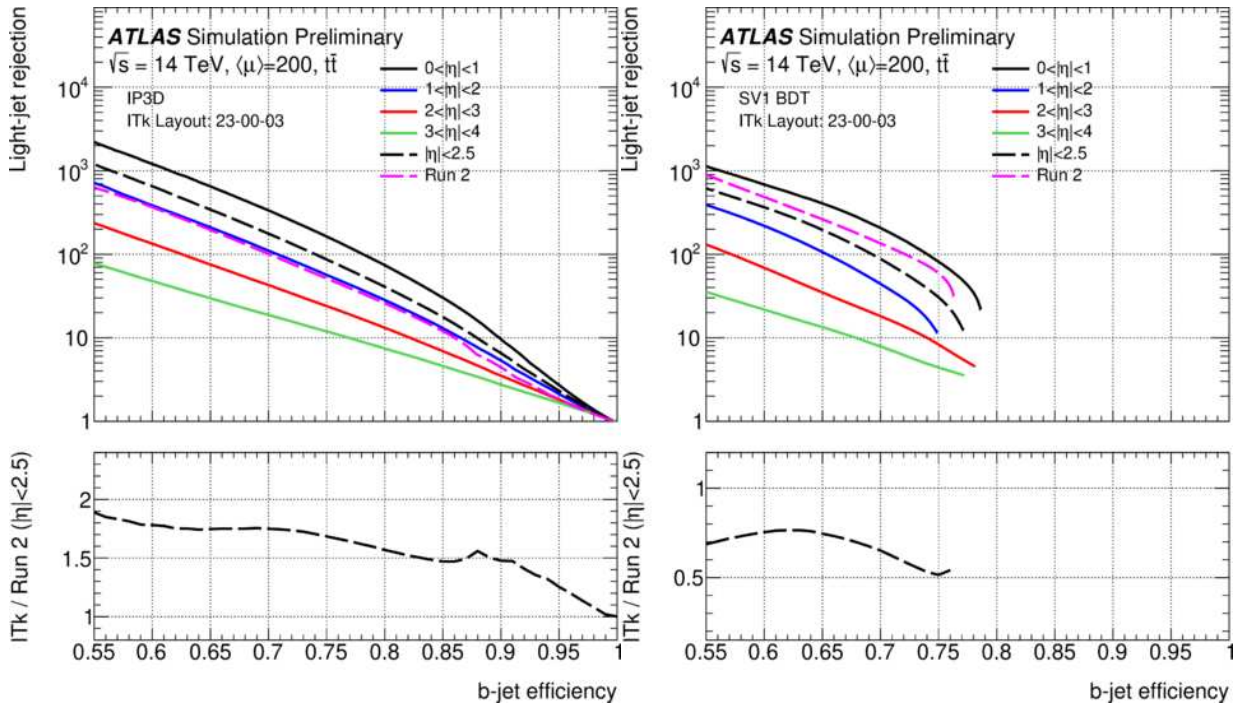


Figure 71 – Light-jet rejection vs *b*-tagging efficiency, with and without material rejection, in exclusive η regions for the SV1 tagging algorithm evaluated in $t\bar{t}$ events with 200 pile-up event in r21.9, obtained with the ITk layout.

3.5 *b*-tagging performance with updated ITk layout

The implementation of the *b*-tagging algorithms in r21.9, including the new optimizations presented in Section 3.4, was used to evaluate *b*-tagging performance with the updated ITk layout, described in Section 2.5.1. The references used for the low-level *b*-taggers are re-derived from r21.9. The expected performance with the updated ITK is compared with the performance obtained with the Run 2 ID. These results are detailed in [1]. The comparison with Run 2 is presented to provide a rough comparison and ensure that the order of magnitude of performance is similar. Figure 72(a) depicts the improved IP3D performance as a result of the improved IP resolution expected with ITk, as well as the new, re-optimized track categorization. The performance of the SV1 algorithms with ITk, shown in Figure 72(b), is evaluated using a BDT with several kinematic and topological vertex properties as inputs to define standalone discriminants. The SV1 algorithm studied here is different from the SV1 results shown previously, which were likelihood-based. The SV1 performance is degraded due to the larger rate of reconstructed SVs (real and fake) in light jets, associated with the larger density of pileup tracks and the different amount of material in pixel layers close to the interaction point. However, SV1 still retains some intrinsic discrimination that could be exploited for *b*-tagging.

The MV2 algorithm, which is based on another BDT, combines IP3D, SV1 and JetFitter algorithm outputs. The re-training of MV2 is performed in r21.9. Figure 73 shows the improved MV2 performance. As an illustration, for a 77% *b*-jet efficiency working point, a light-jet rejection 20% higher than what has been obtained with the Run 2 detector, driven by the IP3D improved performance. For all algorithms, the performance degrades at higher $|\eta|$ region. The DL1 tagger is not tested here as it was not available in r21.9 at the time of writing.



(a) (b)

Figure 72 – Light-jet rejection vs b-tagging efficiency for the (a) IP3D and (b) SV1 *b*-tagging algorithms for different η ranges, evaluated in $t\bar{t}$ events with 200 pileup events, obtained with the updated ITk layout. For comparison, the performance obtained with the Run 2 Inner Detector with an average pileup of 38 is also shown [1].

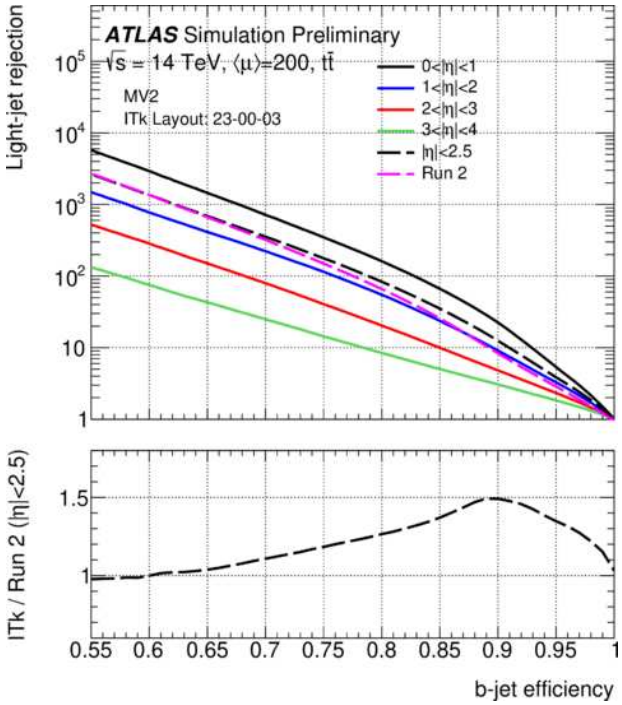


Figure 73 – Light-jet rejection vs b-tagging efficiency for the MV2 *b*-tagging algorithms for different η ranges, evaluated in $t\bar{t}$ events with 200 pileup events, obtained with the updated ITk layout. For comparison, the performance obtained with the Run 2 Inner Detector with an average pileup of 38 is also shown [1].

CHAPTER

4

MULTIVARIATE ANALYSIS USING MACHINE LEARNING TECHNIQUES IN $t\bar{t}H(H \to b\bar{b})$ ANALYSIS

Multivariate analysis (MVA) is a statistical method that analyses more than two variables simultaneously to identify patterns and relationships between them. This is a commonly used tool in experimental high-energy physics for efficiently reconstructing and classifying physics topologies. This chapter will focus on MVA techniques used in $t\bar{t}H(H \to b\bar{b})$ analysis, which allows the kinematic reconstruction of the Higgs boson candidate and the rejection of the $t\bar{t}$+jets backgrounds. The high *b*-jet multiplicity due to additional top quark decay products requires dedicated MVA strategies based on machine learning. Section 4.1 provides a general introduction to machine learning and the terminology used in the MVA studies. The previous round of the $t\bar{t}H(H \to b\bar{b})$ analysis used two different Boosted Decision Trees (BDTs) to identify the Higgs boson candidate and distinguish the signal ($t\bar{t}H$) from the main backgrounds ($t\bar{t}$+jet), discussed in Section 4.2. This chapter focuses on the MVA studies performed in the next round of $t\bar{t}H(H \to b\bar{b})$ analysis, referred to as *legacy analysis*, using full Run 2 data. The analysis aims to take advantage of the most recent object performance algorithms, mainly the Particle-Flow jets (see Section 2.4.4) and the newly optimised *b*-tagging algorithm DL1r (see Section 3.2). In addition, different simulated background samples are used in the legacy analysis, specifically different $t\bar{t}$+ $b\bar{b}$ samples with improved modelling (see Section. 5.3). The BDTs are retrained with the updated input samples, and their performance is compared to that of the previous round of BDT performance, as discussed in Section 4.3. The legacy round also opened new possibilities to develop and optimise different tools used in the previous rounds, particularly the MVA methods. One of the main contributions of this thesis is the development of a novel MVA approach based on Deep Neural Networks (DNNs). The DNN model

is employed to combine reconstruction of the Higgs p_T and classification to distinguish $t\bar{t}H$ from the background sub-components in a single step. Section 4.4 describes the Deep-sets DNN model, used in the development of the new MVA technique, which aims to replace the BDTs and improve overall sensitivity. Section 4.5 discusses the overall performance of Deep-sets MVA, including comparisons with previously employed BDTs.

4.1 Machine learning in physics

Machine learning (ML) is a broad field that encompasses a variety of algorithms that are optimised for specific tasks. These methods are adaptable to different problem sets depending on the model's architecture and by tuning their parameters. ML is widely used in high energy physics experiments at the LHC. It is deployed in FPGAs for the trigger [116], object identification like τ identification [117] and *b*-tagging [109,118], where MV2 and DL1r-based b-taggers are used in the studies in this chapter. In physics analyses, sophisticated ML methods aid in the reconstruction and discrimination of signal processes, such as in $t\bar{t}H(H \to b\bar{b})$ analysis, which is one of the main applications discussed in this thesis. ML methods are now also being further optimised to cope with the increasing luminosity during Run 3 of the LHC and the HL-LHC. In general, there are two types of ML: *supervised learning*, which requires fully labelled training data, and *unsupervised learning*, which does not require any labelled data. In this thesis, supervised learning approaches based on Boosted Decision Trees (BDTs) and Deep Neural Networks (DNNs) are used. This section introduces common ML terminologies that are used in the MVA studies, presented in this chapter.

Data preparation: In order to ensure an unbiased training process, at least three orthogonal data-sets are required. The *training sample* is utilised for the actual algorithm training. The *validation sample* is used to optimise the model further by tuning the hyper-parameters of the model. The validation set provides an unbiased evaluation of the trained model. The *testing sample* is used solely to evaluate the final performance and is not incorporated into the training procedure. In the context of the BDTs and DNNs studies, the *event number*[1] variable is used to split the data-set into the training and testing set. This way, one can know which events were used in the training. During the training, 20% of the training data-set is used as the validation data. In order to cover full statistics, the model trained on even events is applied to events with an odd event number, and vice versa. This is referred to as *cross-training* in this thesis. The training data-set of length N can be represented as $\overrightarrow{X} = (x_1, x_2 ... x_N)$, where each event i has a feature set $\overrightarrow{x_i} = (x_i^1, x_i^2 ... x_i^M)$, with M features.

[1] The event number is a unique integer number associated to each event in the MC samples.

Prediction in ML: The true outputs or labels on a given data-set are referred to as *target outputs*. The function that maps the input to its correct labels is called the *target function*. Many ML methods aim to produce a function that is as close to the target function as possible without sacrificing generalizability. The target output can be used to compare and determine the accuracy of a model prediction.

Regression and classification problems: ML models are commonly used to perform regression and classification problems. A regression technique is used to determine the relationship between features and the target. The goal is to predict the value of a continuous variable. In contrast, the goal of the classification problem is to predict a class from a set of classes.

Under-fitting and over-fitting: During the training, the ML network can become over specialised on the training events and loose the generalization ability when doing predictions on independent testing samples. This phenomenon is called *over-fitting*. This usually happens when the model is too complex. To mitigate this, a *stopping condition* parameter is used to terminate the training process when the model starts over-fitting. Another phenomenon can occur, where the ML model is unable to fit even the training set due to its low complexity. This phenomenon is called *under-fitting*, which can be mitigated by increasing the complexity of the model, the number of training features, or the training duration.

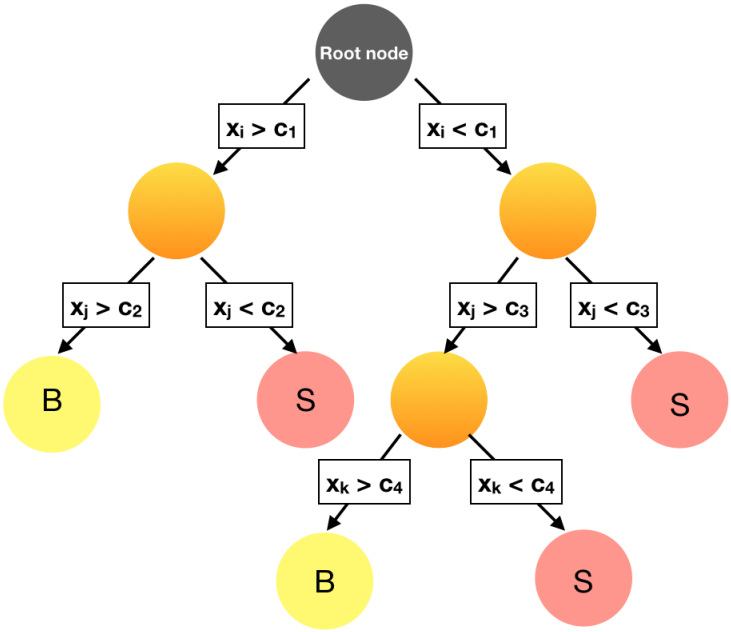
Loss function: The loss function is a measure of the difference between the model prediction $\hat{Y}$ and the target values Y. The model is trained to minimise the loss function. Different loss function choices could result in different optimal solutions. The ML algorithm finds this mapping by solving an optimisation task, measuring the quality of the prediction at each event i with a loss function $L(Y_i, \hat{Y}_i)$, which should be minimised to find the best model according to some metric.

4.2 MVA using Boosted Decision Trees

This section will focus on the MVA technique using two different BDT tools, employed in the previous round of $t\bar{t}H(H \to b\bar{b})$ analysis [33]. A general introduction to BDTs is discussed in Section 4.2.1. The two types of BDTs employed are the reconstruction BDT and the classification BDT. The reconstruction BDT, which is used to find the correct jet-parton assignment per event and reconstruct the Higgs boson and top-quark candidates, is described in Section 4.2.2. The classification BDT, which is used to distinguish $t\bar{t}H$ from $t\bar{t}$ + jets backgrounds, is discussed in Section 4.2.3.

4.2.1 Introduction to BDTs

A decision tree is a powerful MVA technique used to perform regression and classification tasks by learning a set of cut-based decisions inferred from the training data-set. It uses an algorithmic approach, where events that fail a criterion are not immediately rejected but rather re-examined with other features or the same features in a subset of events. Figure 74 shows the basic structure of a decision tree with only two possible classes: signal and background. The BDTs employed in the previous round were based on this binary class approach. In general, a decision tree can also have multiple output classes, where each branch is split into several sub-branches.



(a)

Figure 74 – An illustration of a binary decision tree showing the splitting algorithm from the initial node using variables x_i, x_j and x_k with c_1, c_2, c_3 and c_4 as their corresponding cut values. $S(B)$ denotes signal (background).

The decision tree, as its name suggests, is a tree-like structure with branches connected by nodes. A *node* is a decision point where a variable and cut value are provided. The first node is called the *root node*, where the training starts. The training sample consists of signal and background events, described by a set of variables $\vec{x}$. Each event i is assigned an initial weight w_i (usually random). The root node then splits into two daughter branches by applying a selection of the form $x_i > c_i$, on a single variable. The optimal variable x_i and the cut value c_i are chosen as the ones providing the best separation between signal and background. At the root node, the events passing this cut will enter a signal-rich node, and those failing will enter a background-rich node. Similarly,

each node can be split recursively into two branches. In this way, the events that did not satisfy a certain criterion are not immediately discarded but rather re-examined with the same or other features in a subset of events. The last nodes where the splitting of the trees stops are called *leaf nodes*. Each leaf node is assigned a *purity value* $p = S/S+B$, where $S(B)$ corresponds to the sum of weights of signal (background) events that ended up in the leaf during the training. The value is bound between 0 and 1, with a value close to 1 indicating a signal-like event and a value close to 0 indicating a background-like event. The result of the decision tree for an event tested on a single tree is equal to the purity of the leaf on which the testing terminates. There are several stopping condition parameters used in the training. The *minimum leaf size* is the threshold number of events in a node, below which splitting is not allowed, converting the node into a leaf node. The splitting also stops if perfect separation is achieved, which means that all of the events in the node are of the same class. The maximum number of nodes in the training is given by the *maximal tree depth*, after which the training is terminated.

Although a single decision tree can improve upon a simple cut-based analysis when used on its own, techniques based on the plain averages of multiple decision trees could be more reliable. The BDT works on the principle of the *boosting method* [119], which is expected to be more powerful than single or multiple decision trees without boosting, where all the events are given the same weight. In the boosting method, the mis-classified events (signal as background, or vice versa) are given larger weights, and a new tree is built using these new weights. The new weights α_m are assigned to a single tree T_m (where m is the tree number), given by:

$$\alpha_m = \beta \cdot \ln \frac{1-\varepsilon_m}{\varepsilon_m} \tag{4.1}$$

where, β is a free parameter to adjust the strength of boosting factor and ε_m is the mis-classification rate of the m^{th} tree. The training events for the current tree T_{m+1} are then assigned a new weight, given by

$$w_j^m \to w_j^{m+1} = w_j^m \cdot e^{\alpha_m} \tag{4.2}$$

The classified events are unchanged from T_m to T_{m+1}, while weights of the mis-classified events are increased by a factor e^{α_m}. This way, the mis-classified events are given more importance during the training of the next tree. The boosted result of event i is a weighted average of all the decision trees into single discriminant, and is defined as:

$$\mathcal{T}(i) = \frac{1}{\Sigma_{m=1}^{N_m} \alpha_m} \Sigma_{m=1}^{N_m} \alpha_m T_m(i) \tag{4.3}$$

The Adaptive Boost (AdaBoost) [120] is used to perform BDT studies, presented in this thesis. AdaBoost employs an exponential loss function to minimise the loss function: $\mathcal{J}_{\mathrm{BDT}} = \mathrm{e}^{-\mathrm{p(x)y}}$, where p denotes the model prediction and y is the truth label, and x represents the input features. In general, BDTs can be easily incorporated into a variety of analysis software workflows due to the fact that they are accessible within the data analysis package called ROOT [121] via TMVA package [122].

4.2.2 Reconstruction BDT

The reconstruction BDT, referred to as recoBDT, is used to extract the information from the final state of $\mathrm{t\bar{t}H(H \to b\bar{b})}$ process in the single-lepton channel (see Section 5.1). Each possible combination of jets in the final state is a permutation, which will end up with a corresponding reconstruction of objects, e.g. the Higgs boson, the hadronic top etc., as shown in Figure 75.

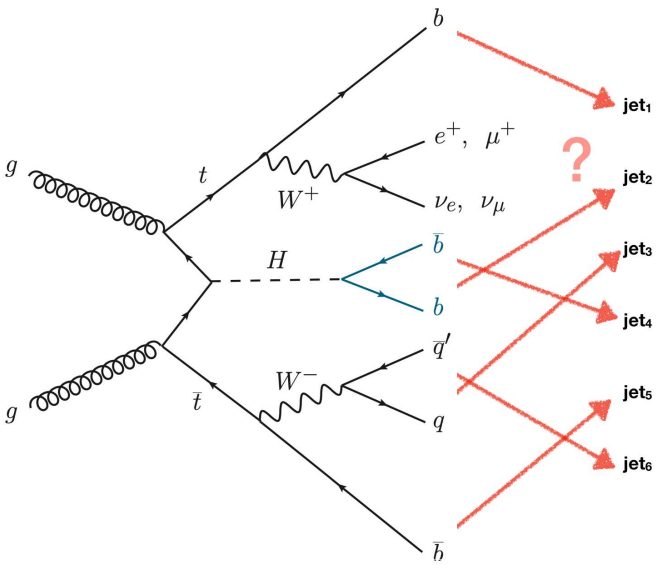


Figure 75 – A Feynman diagram depicting the ambiguity that exists when associating several jets to the partons in the final state of $\mathrm{t\bar{t}H(H \to b\bar{b})}$ single-lepton channel.

The reconstruction algorithms for different objects in the final states for an event containing at least six jets, of which four jets are b-tagged jets, are described here:

- **Hadronic W-boson:** The reconstruction of hadronic W-boson is done using all the possible combinations of jets that are not *b*-tagged jets. If an event contains less than two non-*b*-tagged jets, then in that case a *b*-tagged jet is allowed to be used for hadronic W-boson reconstruction.

- **Leptonic W-boson:** The reconstruction of leptonic W-boson requires lepton and full neutrino momentum reconstruction. The transverse component of neutrino is given by E^T_{miss} whereas the longitudinal component ($p_{z\nu}$) is not measurable. It is possible to deduce this by presuming that the lepton and E^T_{miss} are the sole products of a W-boson decay and constraining the invariant mass to that of the W-boson. $p_{z\nu}$ is inferred by solving:

$$m_W^2 = (p_\ell + p_\nu)^2 \tag{4.4}$$

  where, m_W corresponds to the W-boson mass. If there are two solutions to this quadratic equation, they are both considered, resulting in two distinct interpretations for the leptonic W-boson.

- **Top quarks:** The combination of a W-boson and a *b*-tagged jet is used for the reconstruction of top quarks.

- **Higgs boson:** The remaining *b*-tagged jets are used to reconstruct the Higgs boson candidate.

A number of different quantities, including the invariant mass of the object candidates and their angular distances, are computed for each possible combination of jets. The complete list of inputs used for performing recoBDT is shown in Table 9. The list includes a total of 19 variables among which, 15 variables are based on the topological information of the $t\bar{t}$ system and four variables are related to the Higgs system. When applied to $t\bar{t}$ + jets background events, the jet combination selected by recoBDT with Higgs information may be more $t\bar{t}H$-like, resulting in less discriminating kinematics against $t\bar{t}H$ when fed into the classification BDT. Thus, in order to prevent a bias in the background processes, additional training is carried out without the inclusion of the topological Higgs information.

In order to decrease the number of possible combinations of jets for the training, additional *b*-tagging information is exploited. The jets are sorted based on their pseudo-continuous *b*-tagging score (see Section 2.4.4), and if the scores are the same, the jets are sorted based on their p_T. Only the first six sorted leading jets are considered, where *b*-quarks can only be associated with *b*-jets and other jets are associated with W decay products. The number of combinations are further reduced by assuming that one of the two final partons of the Higgs and hadronic W decays carries the higher p_T. This p_T ordering does not affect the physics, as the important information is the association of

Variables
Topological information from $t\bar{t}$
Mass of top_{lep}
Mass of top_{had}
Mass of W_{had}
Mass of W_{had} and b from top_{lep}
Mass of W_{lep} and b from top_{had}
$\Delta R(W_{\text{had}}\ b \text{ from } \text{top}_{\text{had}})$
$\Delta R(W_{\text{had}}\ b \text{ from } \text{top}_{\text{lep}})$
$\Delta R(\ell, b \text{ from } \text{top}_{\text{lep}})$
$\Delta R(\ell, b \text{ from } \text{top}_{\text{had}})$
$\Delta R(b \text{ from } \text{top}_{\text{lep}}, b \text{ from } \text{top}_{\text{had}})$
$\Delta R(q_1 \text{ from } W_{\text{had}}, q_2 \text{ from } W_{\text{had}})$
$\Delta R(b \text{ from } t_{\text{had}}, q_1 \text{ from } W_{\text{had}})$
$\Delta R(b \text{ from } t_{\text{had}}, q_2 \text{ from } W_{\text{had}})$
Min. $\Delta R(b \text{ from } \text{top}_{\text{had}}, q_i \text{ from } W_{\text{had}})$
$\Delta R(\text{lep}, b \text{ from } \text{top}_{\text{lep}})$ - min. $\Delta R(b \text{ from } \text{top}_{\text{had}}, q_i \text{ from } W_{\text{had}})$
Topological information from the Higgs boson candidate
Mass of Higgs
Mass of Higgs and q_1 from W_{had}
$\Delta R(b_1 \text{ from Higgs}, b_2 \text{ from Higgs})$
$\Delta R(b_1 \text{ from Higgs}, \text{lepton})$

Table 9 – List of input variables for the reconstruction BDTs in the single-lepton channel. The subscript had (lep) indicates the hadronically (leptonically) decaying W-boson or the corresponding top-quark candidates. The symbol b_i refers to b-tagged jets from the Higgs-boson candidate decay, sorted by p_T. The symbol q_i refers to jets from the W-boson decay, sorted by p_T [33].

the jets to the Higgs boson and the W. As an illustration, for an event with exactly six jets (events with one solution for $p_{z\nu}$, given by eq 4.4) in the final state, a total of 720 different permutations would be needed to be tested for an event. With the additional *b*-tagging requirements, this number is reduced to 48, and without an attempt to distinguish the Higgs and hadronic W decay products, only 12 combinations are required to be tested. In the case of two solutions for $p_{z\nu}$, the number of combinations doubles i.e 24 combinations are tested.

The recoBDT is trained only on $t\bar{t}H$ events containing at least six jets where four jets are *b*-tagged jets at @85% WP, referred as $6ji4bi@85\%$ hereafter. Cross-training is performed to cover training on all the ttH events. For the recoBDT training, *signal* refers to all of the correct combinations in the $t\bar{t}H$ sample, while *background* refers to all other combinations. The jet combination with the highest recoBDT score is considered as the right interpretation of the event and is used to define high-level observables, such as the reconstructed Higgs p_T or the top mass. The reconstruction of Higgs p_T candidate is crucial, since the cross-section measurement is performed in exclusive STXS bins (see Section 1.8.2). Table 10 lists the fraction of $t\bar{t}H$ events in the single-lepton channel, where the *b*-jets from the Higgs decay products are correctly assigned in a specific STXS bin using the recoBDT in previous round.

Higgs p_T	$SR^{\geq 6j}_{\geq 4b}$
Inclusive	43%
[0-120) GeV	35%
[120-200) GeV	45%
[200-300) GeV	57%
[300-450) GeV	59%
[450-∞) GeV	

Table 10 – List of the fraction of ttH events where the *b*-jets from the Higgs decay products are properly assigned in a specific STXS bin using the recoBDT in single-lepton channel [33].

4.2.3 Classification BDT

The classification BDT, referred to here as classBDT, is used to separate $t\bar{t}H$ from the $t\bar{t}$ + jets background in the single-lepton channel. The classBDT training is performed on $t\bar{t}H$ and inclusive $t\bar{t}$ + jets background samples. A total of 18 discriminating variables are used for the classBDT training, shown in Table 11.

Variable	Definition
General kinematic variables	
ΔR_{bb}^{avg}	Average ΔR for all b-tagged jet pairs
$\Delta R_{bb}^{maxp_T}$	ΔR between the two b-tagged jets with largest vector sum p_T
$\Delta\eta_{jj}^{max}$	Maximum $\Delta\eta$ between any two jets
$m_{bb}^{min\Delta R}$	Mass of the combination of two b-tagged jets with the smallest ΔR
$N_{bb}^{Higgs_{30}}$	Number of b-tagged jet pairs with invariant mass within 30 GeV of the Higgs-boson mass
Aplanarity	$1.5\lambda_2$, where λ_2 is the second eigenvalue of the momentum tensor built with all jets
H_1	Second Fox–Wolfram moment computed using all jets and the lepton
Variables from reconstruction BDT	
BDT output	Output of the reconstruction BDT †
m_{bb}^{Higgs}	Higgs candidate mass
$m_{H,b_{leptop}}$	Mass of Higgs candidate and b-jet from leptonic top candidate
ΔR_{bb}^{Higgs}	ΔR between b-jets from the Higgs candidate
$\Delta R_{H,t\bar{t}}$	ΔR between Higgs candidate and $t\bar{t}$ candidate system †
$\Delta R_{H,leptop}$	ΔR between Higgs candidate and leptonic top candidate
Variables from likelihood calculations	
LHD	Likelihood discriminant
Variables from b-tagging	
w_{b-tag}^{Higgs}	Sum of b-tagging discriminants of jets from best Higgs candidate from the recoBDT
B_{jet}^{3}	3^{rd} largest jet b-tagging discriminant
B_{jet}^{4}	4^{th} largest jet b-tagging discriminant
B_{jet}^{5}	5^{th} largest jet b-tagging discriminant

Table 11 – List of input variables to the classBDT in the single-lepton channel. For variables from the recoBDT, those with a † are from the BDT using Higgs boson information, those with no † are from the BDT without Higgs boson information [33]

.

The training learns from various information listed here:

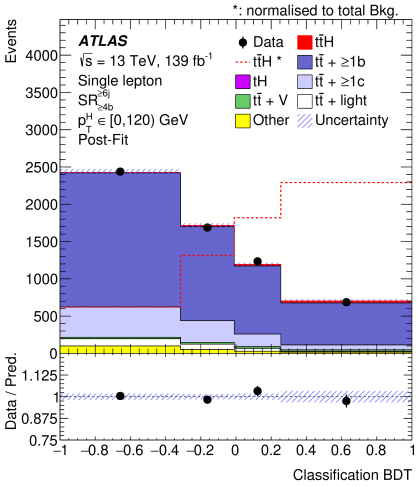
- The recoBDT scores (with or without using Higgs-related features) and the resulting kinematic variables of the Higgs and $t\bar{t}$ systems.
- The combination–independent kinematics denoted as global variables.
- The pseudo-continuous *b*-tagging information.
- Likelihood discriminant (LHD) [33]: It is an intermediate MVA output and is one of the dominant discriminants of the classBDT. LHD computation uses 1D PDFs of the invariant masses and angular variables from reconstructed objects. The signal hypotheses (p_{sig}) and background hypotheses (p_{bkg}) are derived from the product of the PDFs for a single variable. These are then averaged over all possible parton combinations while being weighted by the *b*-tagging information. The ratio $p_{sig}/(p_{sig}+p_{bkg})$ is used as input for the classBDT per event.

The classBDT cross-training is performed in $6ji4bi@85\%$ region. During the training, the level of importance of the input features can be evaluated. It is determined by the frequency with which the classBDT makes use of that variable in the different nodes. In the previous round, the classBDT outputs were used as the discriminants in each STXS region and as inputs for the profile likelihood fit (see Section 5.6) to extract the signal. Figure 76 shows the classBDT output distribution for the $t\bar{t}H$ signal and all the background components for different STXS regions.

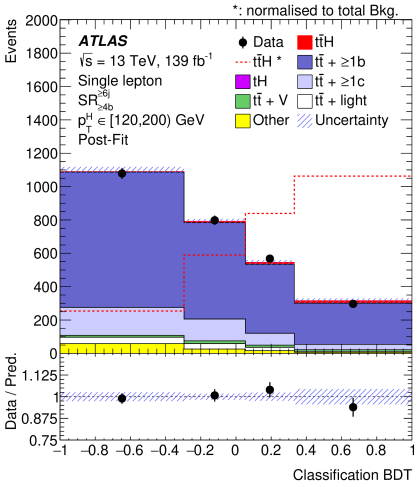
4.3 BDT performance with legacy $t\bar{t}H(H \to b\bar{b})$ analysis

As mentioned before, the legacy analysis aims to benefit from the latest object performance algorithms. First, the PFlow-jets are used instead of EMTopo-jets, described in detail in Section 2.4.4. Second, the DL1r *b*-tagger is used instead of MV2c10 *b*-tagger, in order to benefit from better *b*-tagging performance. For simplicity, the jets used in legacy round will be referred as *PFlow/DL1r jets* and the one used in the previous round will be referred as *EMTopo/MV2c10* jets.

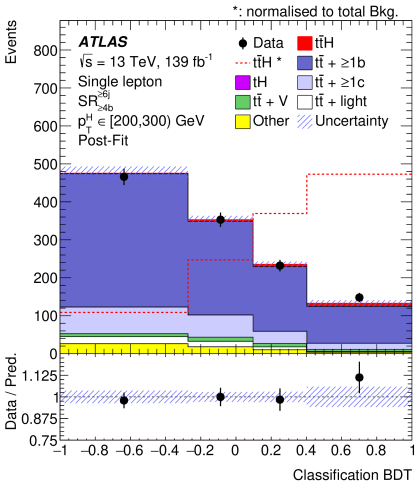
To begin with, the old weights obtained from the previous training (using the EMTopo/MV2c10 jets) are used to study the recoBDTs and the classBDT performance with the PFlow/DL1r jets. This is done to understand the overall BDT configurations used in the previous round. Then later, retraining using the PFlow/DL1r jet samples are also performed. The overall performance of the BDTs using the PFlow/DL1r jets are compared with the performance obtained with the EMTopo/MV2c10 jets. The recoBDT and the classBDT performances, as described above, are presented in Section 4.3.1 and Section 4.3.2, respectively.



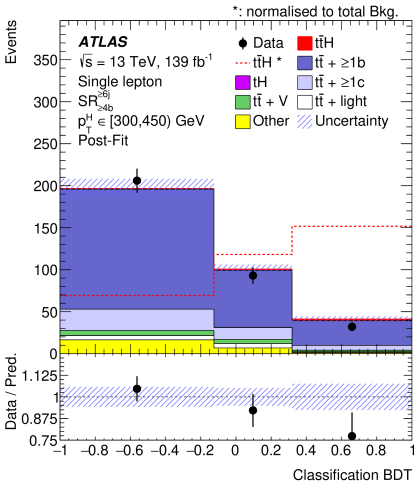
(a)



(b)



(c)



(d)

Figure 76 – The classification BDT output distribution in (a) $0 \leq p_T^H < 120$ GeV, (b) $120 \leq p_T^H < 200$ GeV, (c) $200 \leq p_T^H < 300$ GeV and (d) $300 \leq p_T^H < 450$ GeV STXS regions in the single-lepton channel. The dashed line shows the $t\bar{t}H$ signal distribution normalised to the total background prediction [33].

4.3.1 RecoBDT performance

When evaluating the performance, the recoBDT takes into account how well they are able to successfully reconstruct certain objects. The performances are analysed using all the events in the 6*ji*4*bi*@85%, which corresponds to the training region, where the dedicated BDTs were trained or retrained. Before examining the performance based on the recoBDT score, *event fraction* is evaluated. This corresponds to the fraction of events where the truth objects in the final state are reconstructed. The event fraction value of a matched object, represents the maximum reconstruction performance attainable for that object. The *b*-tagging requirement, discussed in Section 4.2.2, is not taken into account for evaluating the event fraction. This means that *b*-quarks can be associated to any jet. Figure 77, shows the event fraction as a function of the matched objects. The matched objects definitions are described as followed:

- all: six selected jets, all matched to the six partons in the final state
- b+1W: four *b*-quarks from top and Higgs decays and one quark from W decay
- all_b: four *b*-quarks from the top and Higgs decays
- H: Higgs boson
- b_{top}: two *b*-quarks from $t\bar{t}$ decays.
- W: hadronic W-boson
- H_{b1}: leading p_T *b*-quark from Higgs decay
- H_{b2}: sub-leading p_T *b*-quark from Higgs decay
- bl_t: *b*-quark from leptonic top decay
- bh_t: *b*-quark from hadronic top decay
- wj_1: leading p_T quark from hadronic W-boson decay
- wj_2: sub-leading p_T quark from hadronic W-boson decay

For illustration, comparing the "all" category, around 39%(36%) of the events have six selected jets all truth matched to the six partons in the final state using PFlow/DL1r jets (EMTopo/MV2c10 jets). When compared to EMTopo/MV2c10 jets, slightly higher event fraction values are observed for all the matched objects with PFlow/DL1r jets. This might be due to different reconstruction of a given event, shown later in this section. The measure of the reconstruction performance can

be analysed via *matching fraction*. This refers to the fraction of events where the truth objects are reconstructed from the jet combinations with highest recoBDT score. Figure 78 shows the matching fraction as a function of different matched objects in $6ji4bi@85\%$ region. The matching efficiency performance using PFlow/DL1r jets is slightly better (around 2-3%) than EMTopo/MV2c10 jets performance, whereas, retraining shows negligible impact on the performance. Thus, switching from EMTopo/MV2c10 jets to PFlow/DL1r jets has very slight effect on the variables entering the recoBDT training.

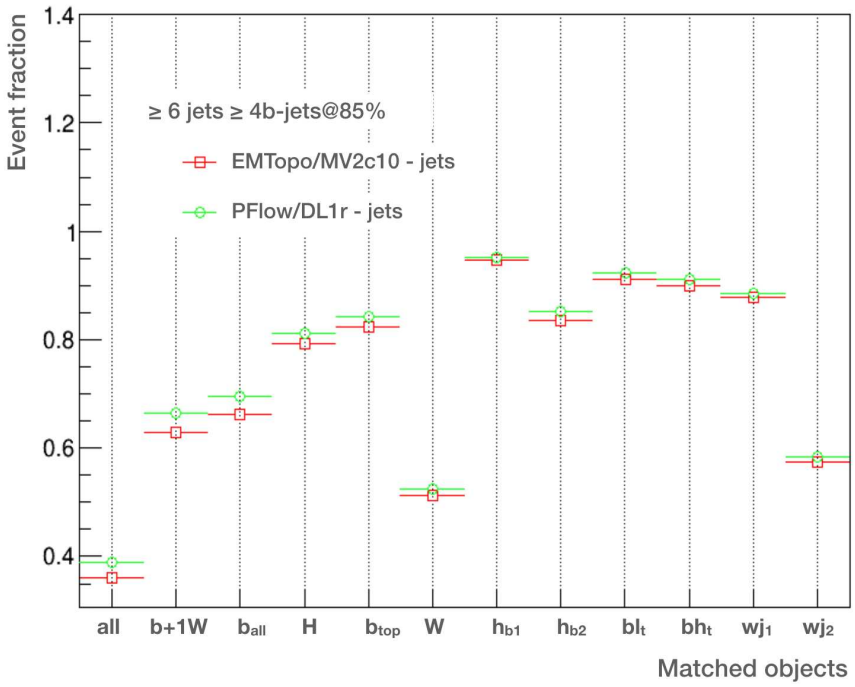


Figure 77 – The fraction of selected events for which their corresponding objects are truth matched.

Further studies are also performed to compare event by event reconstruction performance using PFlow/DL1r and EMTopo/MV2c10 jets. Figure 79 and Figure 80 depict the event fraction and matching fraction performance for events with and without identical event numbers, respectively. The fraction values for all the matched objects are compared between PFlow/DL1r and EMTopo/MV2c10 jets samples. Around 82% of events are selected for both sets of jet reconstruction/*b*-tagging algorithms. Even after selecting the identical events, differences in event and matching fractions are still observed. Compared to events that are identical, the difference in matching fraction performance between samples is larger for uncommon events. Similar patterns can be observed in the event fraction performance. The slight overall improvement with PFlow/DL1r jets could be attributed to the different reconstruction of a given event in the two samples.

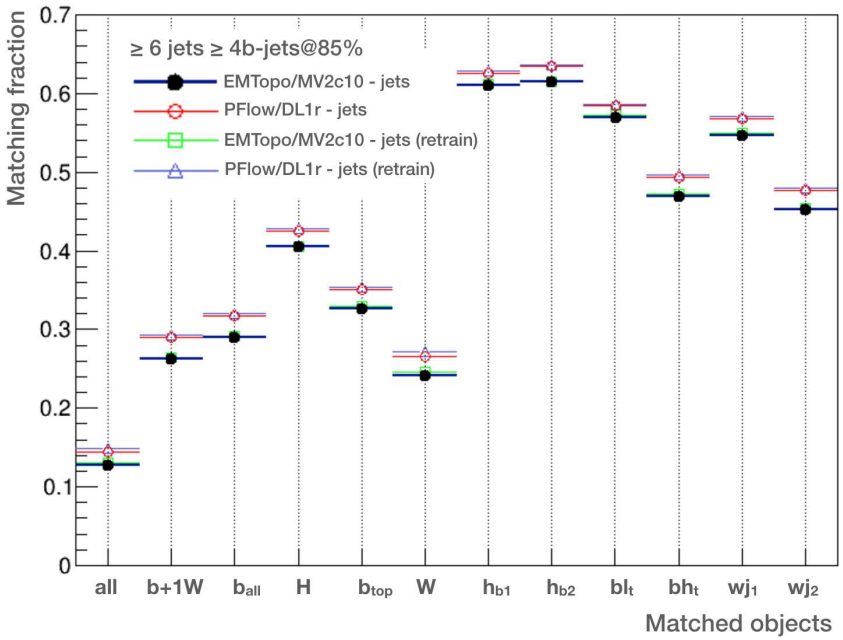
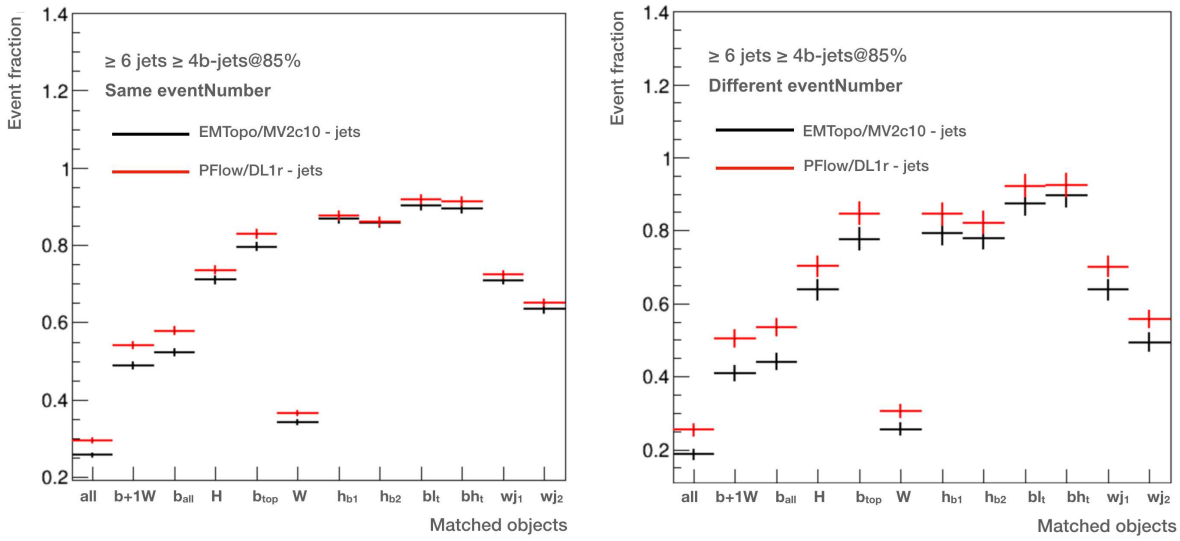
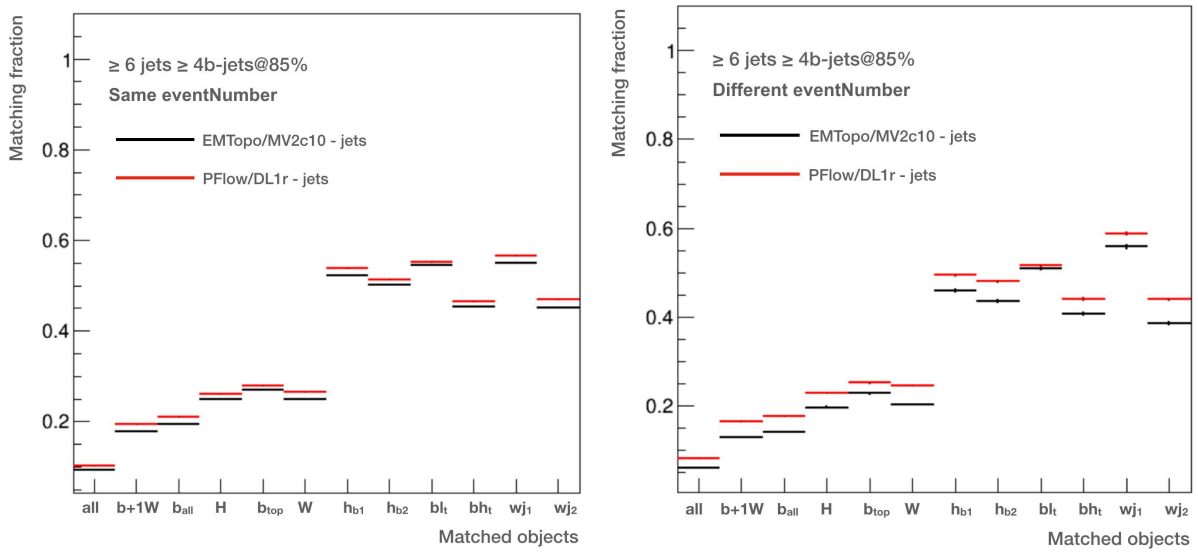


Figure 78 – The fraction of selected events for which their corresponding objects are truth matched only for the jet combinations with the highest recoBDT score. Here, tag "retrain" refers to the performance obtained after retraining with the corresponding jet samples.



(a) (b)

Figure 79 – Event fraction of the final state objects for events with (a) same and (b) different event number, compared between PFlow/DL1r and EMTopo/MV2c10 jets samples.



(a) (b)

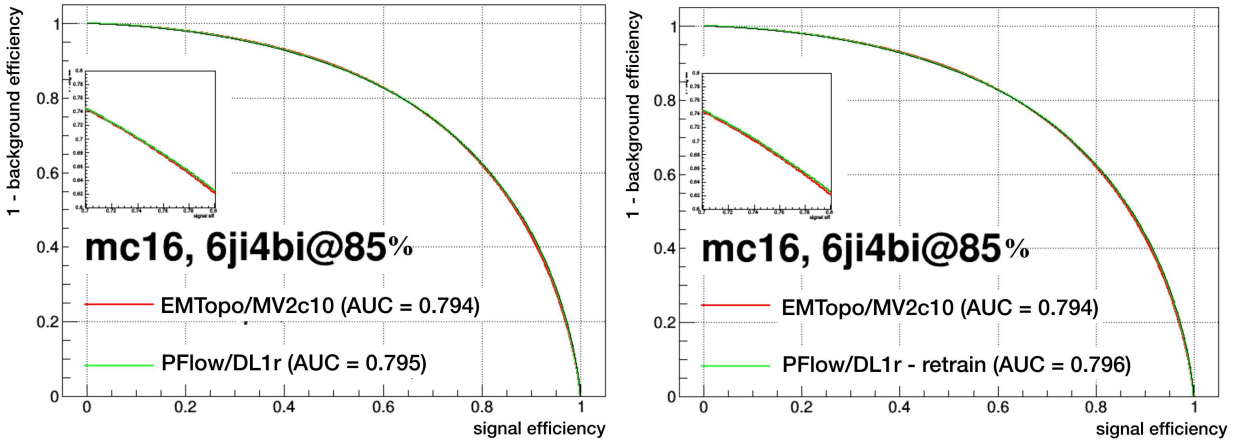
Figure 80 – Matching fraction of the final state objects for events with (a) same and (b) different event number, compared between PFlow/DL1r and EMTopo/MV2c10 jets samples.

4.3.2 ClassBDT performance

The ClassBDT performance can be measured using Receiver Operating Characteristics (ROC). The signal efficiency, or the ratio of correctly identified signal events to total number of signal events, is calculated for each decision cut on the classifier output against how often the background is not mis-classified (1 - background efficiency). The integrated Area Under Curve (AUC) is calculated to quantify the performance.

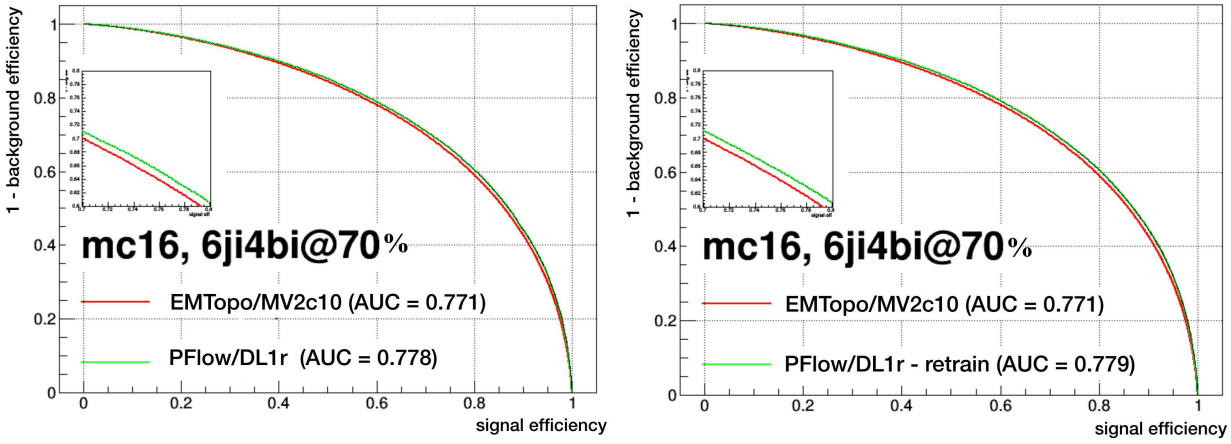
The classBDT performances using PFlow/DL1r and EMtopo/MV2c10 jets are analysed using all the events in the $6ji4bi@85\%$ region. Figure 81(a) shows the performance of the classBDT when the old weights are used. The classBDT performance with PFlow/DL1r jets is similar to EMTopo/MV2c10 performance. The retraining with the new samples shows negligible improvements, shown in Figure 81(b). Similar to the recoBDT case, the overall classBDT variables have a very low dependence on the change from EMtopo/MV2c10 to PFlow/DL1r samples. The classBDT performance is also checked in the region with tighter *b*-tagging requirement i.e $6ji4bi@70\%$. Figure 82 depicts the performance of classBDT in the $6ji4bi@70\%$ region, where the same trend is observed as in the $6ji4bi@85\%$ region.

To conclude, the retraining of the BDTs using PFlow/DL1r jets will serve as a baseline for new MVA developments using Deep Neural Networks, which will be discussed in Section 4.5. The previously trained BDT weights in the central $t\bar{t}H(H \to b\bar{b})$ analysis framework were also replaced by the retrained weights using PFlow/DL1r jets.



(a) (b)

Figure 81 – Performance of classification BDT in 6*ji*4*bi*@85% region, labeled as 6ji4bi@85%. "retrain" refers to the performance obtained using retrained weights from PFlow/DL1r jets.



(a) (b)

Figure 82 – Performance of classification BDT in 6*ji*4*bi*@70% region. "retrain" refers to the performance obtained using retrained weights from PFlow/DL1r jets.

4.4 MVA developments using Deep Neural Networks

Different DNN models [123] are already extensively used in various experimental analyses using LHC data and have provided improvements with respect to the previously used methods. The studies presented in this section will focus on new MVA developments using DNNs for the legacy $t\bar{t}H(H \to b\bar{b})$ analysis using full Run 2 ATLAS data. The goal is to replace the two-step BDT model, as described in Section 4.2, with the single-step DNN model. The general idea is to exploit the flexibility of DNN architecture to perform the following MVA steps in a single tool:

- Perform regression on Higgs kinematic variables like p_T^H, in order to define different STXS bins for differential STXS measurements in the $t\bar{t}H(H \to b\bar{b})$ analysis.
- Perform multi-classification to distinguish $t\bar{t}H$ from $t\bar{t}$+jets sub-background components i.e $t\bar{t}$+1b, $t\bar{t}$+B, $t\bar{t}+ \geq 2b$, $t\bar{t}+ \geq 1c$ and $t\bar{t}$ + light, introduced in Section 5.3. This also allows to build discriminant with a probability associated with each sub-background components. These discriminants can be used to build dedicated control regions in the analysis to constrain the uncertainties of the different processes and increase the overall sensitivity. The multi-class DNN approach is already employed in the $t\bar{t}H(H \to b\bar{b})$ analysis by CMS [35], which also served as the motivation to investigate the MVA strategy using DNNs.

Section 4.4.1 provides an overview of DNNs, followed by Section 4.4.2, which introduces the specific DNN model used in the MVA studies, namely Deep-sets. Section 4.4.3 details on the computation of the high-level inputs used in the DNN training and testing. Section 4.4.4 describes the Higgs kinematic regression performed using simple DNN and Deep-set networks. The strategies, the design and the optimisations involved in the development of the final Deep-sets MVA model is discussed in Section 4.4.5.

4.4.1 Introduction to DNNs

Deep learning [124] is a machine-learning method that aims to model data using complex structures that combine several non-linear transformations. The Neural Networks (NNs), which are the fundamental building blocks of deep learning, are composed of the elementary units known as *artificial neurons*. Figure 83 shows a sketch of the operation of the artificial neuron. The artificial neurons receives the input x, which are then multiplied by their assigned weights W_i and then a *bias term* b is added to the result. This forms the weighted sum which is then passed through the *activation function* f. The activation function is chosen to introduce non-linearity into the model (described in detail later in the section). The bias term is a constant value that is used to produce an

offset such that the output y takes the linear form $b + W \cdot x$. This helps the model in fitting best for the given data that do not pass through the origin by enabling the activation function to be shifted accordingly.

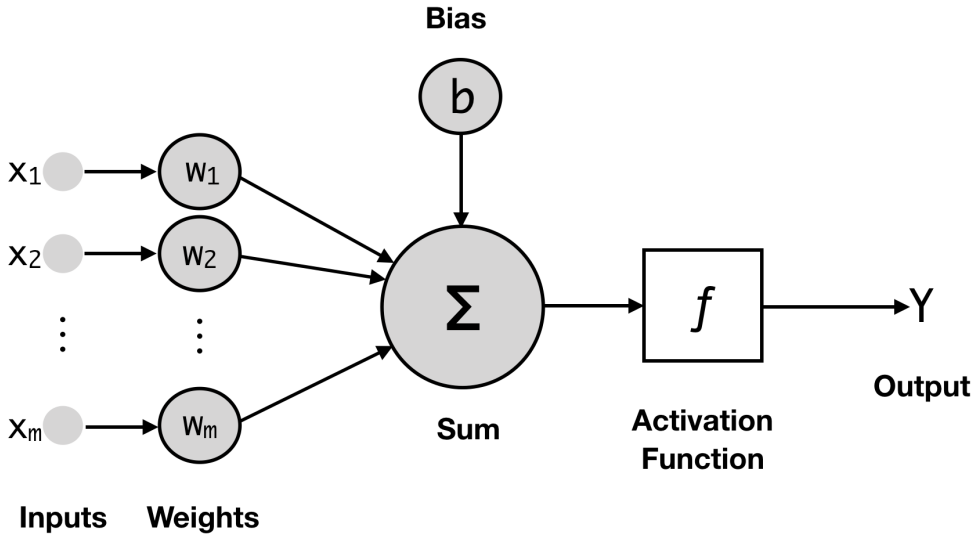


Figure 83 – Sketch showing the operation of the artificial neuron in neural networks.

One of the simplest type of network is called *feed-forward network*, shown in Figure 84. It consists of one input layer, one or more hidden layers, and one output node. The hidden layer is where the transformations are performed on the features, fed into it via the input layer, and are then passed to the output layer. The mathematical point of view of a feed-forward network can be represented as:

$$\mathcal{P}^{NN} = f_2\left(W_2 f_1(W_1 x + b_1) + b_2\right) \tag{4.5}$$

The feed forward network with one hidden layer is fairly simple and computationally limited, but it can provide insight into more complex neuron structures. By varying the amount of hidden layers and nodes, a NNs can approximate any arbitrary function and allows the network sufficient freedom. For instance, W and b values are free parameters that can be optimised. Deep Neural Networks (DNNs) are built by combining NNs with multiple hidden layers. A very complex DNNs can easily reach several ten-thousands of free parameters.

During the training, the data-set is shuffled randomly and is divided into equal-sized segments, known as batches. The size of the batches is determined by *batch-size* parameter, which is a free parameter and can be optimised during the training. At each iteration of the training of a batch, the model tries to extract patterns from the sample by adjusting the set of weights. Each weight

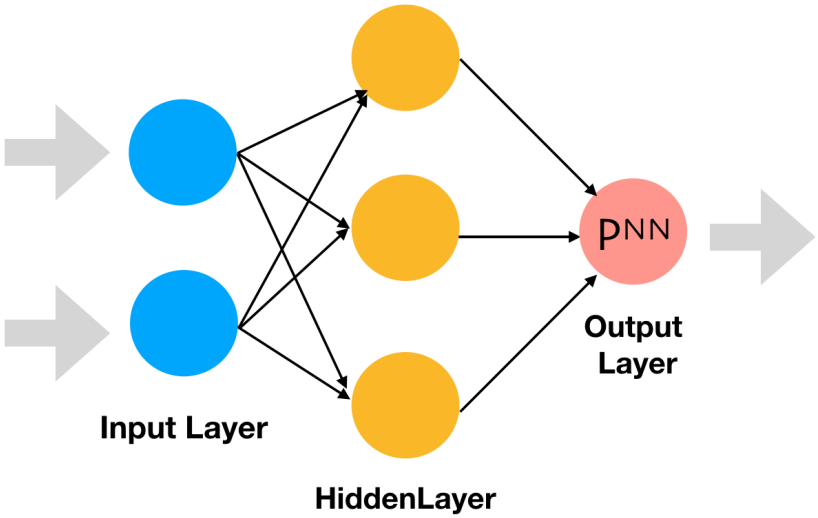


Figure 84 – The sketch of feed forward network with one hidden layer.

update is called a *batch training*. The process is repeated for all the training samples, and a round of updating over the full training data-set is called an *epoch*. For each epoch, the total training data-set is randomly shuffled and therefore the batch sampling are different from the previous epoch. In general, DNNs have non-trainable parameters that fix the architecture and training process of a DNN, known as *hyper-parameters*. The most important hyper-parameters are described here:

- **Loss function**: For the regression problems, the most commonly used loss function is the *mean square error (MSE)*, defined as:

$$\mathcal{J}_{\mathrm{MSE}} = \frac{1}{\mathrm{N}}\Sigma_{i=1}^{N}(\mathrm{Y_i} - \hat{\mathrm{Y}}_\mathrm{i})^2 \tag{4.6}$$

  where, N is the length of the training data-set $\overrightarrow{x}$, Y is the target value and $\hat{Y}$ is the model prediction for each event i. For the binary classification problem, the negative log-likelihood of a Bernoulli distribution is used, called *binary cross-entropy*:

$$\mathcal{J}_{\mathrm{binary-cross-entropy}} = -\frac{1}{\mathrm{N}}\Sigma_{i=1}^{N}\mathrm{Y_i}\cdot\log(\hat{\mathrm{Y}}_\mathrm{i}) + (1-\mathrm{Y_i})\log(1-\hat{\mathrm{Y}}_\mathrm{i}) \tag{4.7}$$

  In case of multi-classification problem with C classes, *categorical cross-entropy* is used:

$$\mathcal{J}_{\mathrm{categorical-cross-entropy}} = -\frac{1}{\mathrm{N}}\Sigma_{i=1}^{N}\Sigma_{c=1}^{C}\mathrm{Y_{ic}}\log(\hat{\mathrm{Y}}_\mathrm{ic}) \tag{4.8}$$

- **Optimizer**: The training of the model involves a process of updating the set of weights (described above) to minimise the loss function. The optimizer is used to determine how these weights are optimised in each layer l, based on the gradient of the loss function. The method

to fully compute the loss gradients is called *back-propagation*. After a first initialisation of the weights W_0, the weights are updated iteratively using the formula:

$$W_{l+1} = W_l - \lambda \bigtriangledown_{W_l} \mathcal{J} \tag{4.9}$$

where, λ is the *learning rate* which is a tunable parameter which has an impact on the convergence speed. The learning rate is defined as the magnitude of model weight update/change during the back-propagation training procedure. If the learning rate is set too high, convergence may not occur because the optimization will skip over the minimum. A low learning rate, on the other hand, can cause the optimization process to move more slowly, and the optimizer to become stuck in a local minimum. The most commonly used optimiser called, ADAM optimiser [125] is used for the studies presented in this thesis.

- **Activation functions**: The activation functions are used to allow the model to learn the non-linear patterns (complex data) to to make precise predictions. Non-linear activation functions have the property of being differentiable; otherwise, they cannot work during DNN back-propagation. Typically, the activation functions used in output nodes differ from those used in hidden layers. The DNN studies shown in this chapter employs the rectified linear unit (ReLU) in the hidden layers. ReLU function is defined as:

$$f_{ReLU}(x) = \begin{cases} 0, & \text{for } x < 0 \\ x, & \text{for } x \geq 0 \end{cases} \tag{4.10}$$

In case of regression studies, shown in Section 4.4.4, a linear activation function is used in the output layer. For the multi-classification studies, shown in Section 4.4.5, softmax [124] activation function is used in the output layer, defined as:

$$f_{softmax}(x_C) = \frac{e^{x_l}}{\Sigma_{c=1}^{C} e^{x_c}} \tag{4.11}$$

fulfilling,

$$\Sigma_{c=1}^{C} f_{softmax}(x_c) = 1 \tag{4.12}$$

The output for each class have value between 0 and 1 which can be interpreted as probabilities.

- **Regularisation**: In order to keep the model robust and avoid-over-fitting, regularisations are used. This includes mainly *batch-normalization* [126], *dropout* [127] and *early-stopping*. Batch normalisation re-normalises and re-scales the output of the previous layer of the batch.

Early stopping terminates the training process when certain criterias are met, such as when the loss has not decreased over a predetermined number of epochs. In this chapter, the training is early stopped if the loss on the validation data-set does not improve in 35 epochs before reaching the specified training epoch, and the model with the minimum validation loss is finally saved for application. Other parameters like AUC on validation data-set is also considered as the alternative for early stopping criteria. Dropout is used to randomly drop a percentage of neuron connections, called dropout rate. During the first iteration of training of a batch, only the un-dropped neuron weights are updated and the dropped ones remaining unchanged. Then, in a new batch training, another set of neurons is chosen to be dropped from among all the neurons. As a result, the overall result is similar to an average of various networks. And each neuron is compelled not to rely too heavily on neurons from the previous layer.

- **Sample weights**: A set of weights, of length equivalent to the training data-set, can be passed during the training to add more importance to some events than others. These weights are called *sample weights*. The sample weight affects how the loss is calculated and can help to maximize the decrease in the loss function.

Python and its related libraries are used to implement all DNN training, testing, and performance checks before sending them to the ATLAS framework. Uproot [128] is used to read branches from ROOT files directly into the Python environment as Numpy arrays. DNNs can be accessed through a variety of sophisticated packages and user-friendly software. Popular software packages are KERAS [129], TensorFlow, and scikit-learn [130]. An open-source format known as ONNX [131] is being used to facilitate the exchange of models between various tools and to allow the incorporation of DNN models into ATLAS software architectures. In this thesis, the ONNX package is used to deploy KERAS models in the central ATLAS $t\bar{t}H(H \to b\bar{b})$ framework. The Deep-sets network, described in next section, is used in the development of the new MVA model.

4.4.2 Deep-sets

Deep-sets [132] is a permutation-invariant DNN model where the input features are treated as *sets* and no specific ordering of the elements is considered. Deep-sets architecture was first used in particle physics to identify jet types [133] and was also recently used to develop the ATLAS flavour-tagging algorithm called DIPS [118]. As described in Section 4.4.3, the inputs features used for the training of any DNN model are computed for 12/24 different combinations per event. In $t\bar{t}H(H \to b\bar{b})$ analysis, the deep-sets network can offer permutation invariance by treating the jet combinations per event as unordered during the training. The basic structure can be similar

to feed-forward NNs connected with multiple hidden layers. The Deep-set computation can be represented in the form of:

$$\rho\left(\sum_{i=1}^{n}\phi(x_i)\right) \tag{4.13}$$

where, n is total number of combinations and ϕ and ρ corresponds to by fully connected dense layers with each dense layer using ReLU activation function. The two networks i.e ϕ and ρ deals with different operations. The combination network ϕ network deals with the operations on the inputs, where it extracts the relevant combination features whereas the ρ network focuses on the operations over the sets which deals with the correlation among the combinations of jets. The sum operation is employed to ensure permutation invariance. This summed up layer encodes the information about the number of combinations in each event which helps in the identification of the best combination in an event. Thus, the Deep-sets architecture was considered for the development of the MVA model using DNNs. The strategy and final model adapted for the new MVA model in $t\bar{t}H(H \to b\bar{b})$ analysis is described in detail in Section 4.4.5.

4.4.3 Input preparation for DNNs

As already described in Section 4.2, the final state of $t\bar{t}H(H \to b\bar{b})$ is quite complex involving many jets and b-jets. The combined list of inputs, listed in Table 9 and Table 11 for the events in $\geq$ 6-jets and $\geq$4 b-jets@70% region are used for performing training on the DNN networks. A technique similar to BDTs is used to limit the number of combinations of jets entering the network, described in Section 4.2.2. A total of 34 different features computed for each combinations in $\geq$ 6-jets and $\geq$4 b-jets@70% region are used, exploiting both the multiple jet combinations as well as the correlation between the features. Out of this total, 19 variables are the ones from recoBDT inputs, complemented by the Higgs p_T i.e. the p_T of the di-jet system assumed to come from the Higgs boson decay products. The remaining 14 variables are taken from the classBDT variables. LHD was not included as it did not improve the final performance and would have in principle required to re-derive new likelihood templates as part of the training process.

The input features used for the DNN training have different range of values. The normalisation of the inputs can make the model more stable and aids in the convergence and can generalize appropriately. The input features used in the training of the DNN model is normalised to to have a mean zero and standard deviation one. In order to make the inputs fixed-size, *zero padding* is performed on events with fewer than 24 combinations, resizing them to 24 by adding zeroes.

The training and testing data-sets are prepared by dividing the full statistics based on their event numbers and 20% of training data-set is used as the validation data.

4.4.4 Higgs p_T regression

The first step in the effort to develop new MVA using DNNs was to perform Higgs kinematic regression on $t\bar{t}H$ signal events to reconstruct the Higgs p_T, to be used for differential STXS measurements. This was achieved by using the DNN network to perform regression on Higgs p_T in $t\bar{t}H$ events. Since the $t\bar{t}H$ events are statistically limited and high-level[2] inputs are used for the training, a simple DNN model was used, as illustrated in Figure 85.

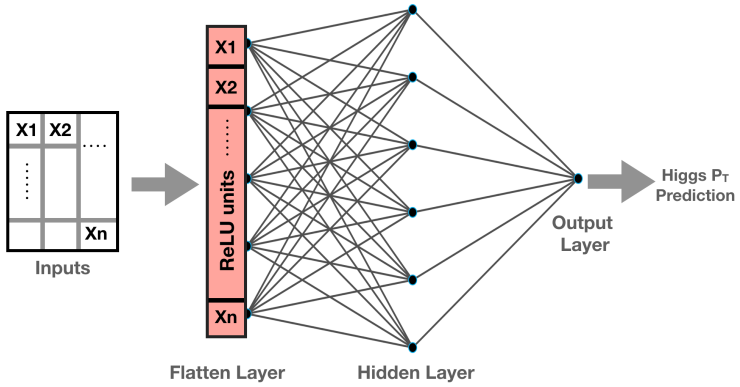


Figure 85 – The architecture of the DNN model for performing Higgs p_T regression.

The DNN model consists of an input layer, where the inputs are flattened into a long single-array. The random ordering of jet combinations per event entering the input layer is maintained in the flattening layer. The dense layers are incompatible with multi-dimensional array data processing. Since, the trained inputs are a three-dimensional array of size = (event, combination, feature), the data is converted to a single one-dimensional vector before being fed to the next dense layer. Using the ReLU activation function, the network has two fully connected hidden layers with dropout-rate value of 0.2 applied between each layers. To provide linear continuous predictions, the output layer uses a linear activation function. As, the Higgs p_T regression accounts for the reconstruction part, only 20 features discussed in Section 4.4.3 are used. The target value is the truth level Higgs p_T distribution. The Table 12 shows the list of hyper-parameter values for the DNN network to perform Higgs p_T regression.

The output of the DNN model described above is used to predict the reconstructed Higgs p_T spectrum. The performance of the DNN model is compared to the recoBDT, which uses the reconstructed Higgs p_T defined based on the jet permutation with the highest BDT score.

[2] High-level variables are functions of low-level features like p_T of a final state particle

Hyper-parameters	Values
N_{Nodes}	[240, 64, 8, 1]
Dropouts	0.2
Batch-size	256
Learning rate	0.01

Table 12 – The list of hyper-parameter values for the DNN model used to perform Higgs p_T regression.

Figure 86 depicts the the DNN prediction spectrum of Higgs p_T, and RecoBDT output as well as the target spectrum, i.e. the truth Higgs p_T. Clearly, the shape of the predicted Higgs p_T using DNN is quite different from the truth Higgs p_T shape. Figure 87 depicts the 2D distribution of the truth vs prediction, which demonstrates a clear offset from the diagonal, which is not observed in the reconstructed Higgs p_T. Figure 88 shows the difference between prediction and truth Higgs p_T (prediction - truth) as a function of truth Higgs p_T. When compared to recoBDT in truth $p_T < 200$ GeV region, DNN output shows less deviation (prediction - truth) from zero. In $p_T > 200$ GeV region, DNN shows more deviation between 10% and 50%, compared to recoBDT. Thus, the degraded performance of the simplified DNNs, motivated to perform regression with more complex and optimised DNN models like Deep-sets.

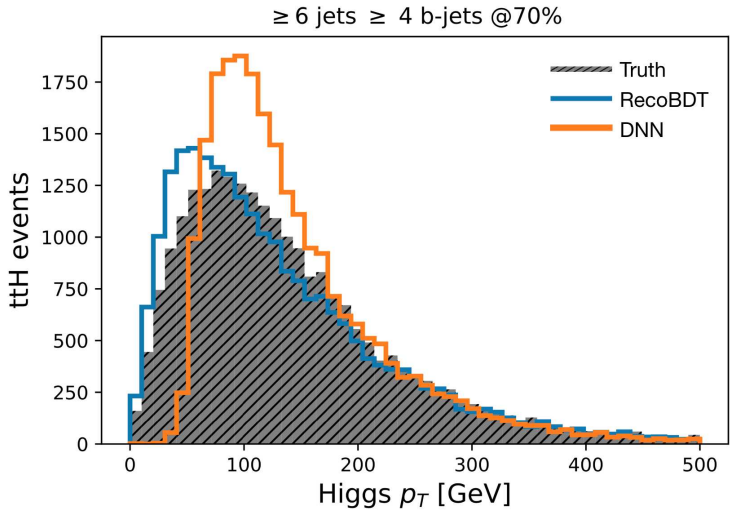
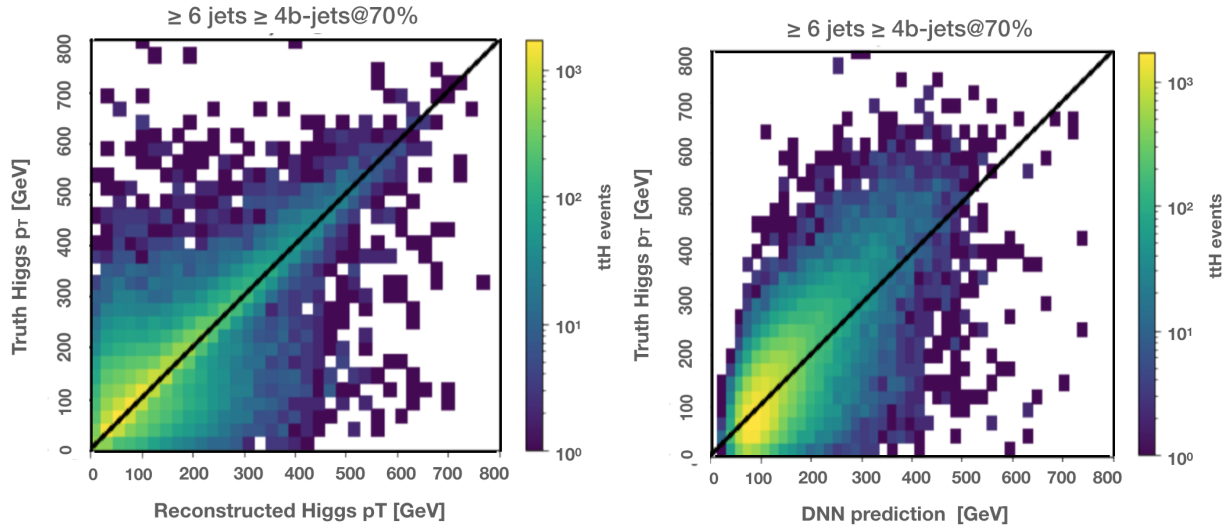


Figure 86 – The Higgs p_T distribution [GeV] where (solid gray) shows the truth Higgs p_T spectrum, (blue) shows the reconstructed Higgs p_T spectrum with highest recoBDT score and (orange) shows the predicted Higgs p_T spectrum from the DNN network

The Deep-sets model, which can handle the inputs better without flattening them is used to perform regression by replacing the simple dense layers with a Deep-set network (see Section 4.4.2). The Deep-sets model used for this regression task is described in Section 4.4.2. The hyper-parameters



(a) (b)

Figure 87 – Scatter plots showing correlation between (a) truth vs reconstructed Higgs p_T and (b) truth vs DNN prediction.

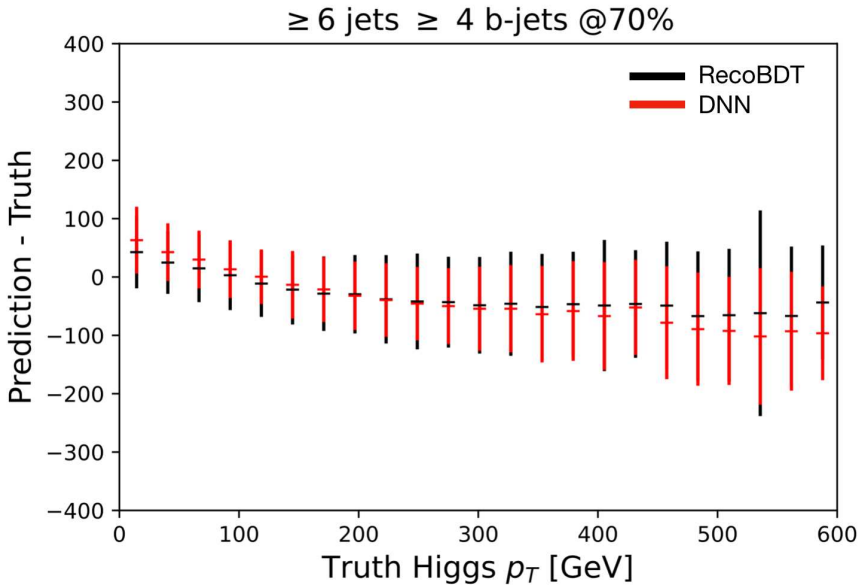


Figure 88 – The profile distribution showing (prediction - truth) as a function of truth Higgs p_T using reconstructed Higgs p_T (black) and Deep-sets prediction(red). The dashed line is the mean value and the vertical error bar corresponds to the standard deviation.

like the number of layers in the ϕ and ρ networks, the number of nodes in the ϕ and ρ, batch-size and learning-rate of the Deep-sets network are optimised by hand at first. Table 13 shows the list of the hyper-parameters used for the Deep-sets regression model. Figure 89 shows the Deep-sets regression prediction spectrum of predicted Higgs p_T compared with recoBDT and truth Higgs p_T distribution, where the predicted shape is comparable. Figure 90 shows the truth vs Deep-sets scatter plot, where no visible offset from the diagonal is observed. Figure 91, shows the deviation of (prediction - truth) as a function of truth Higgs p_T. Comparing Deep-sets performance with the recoBDT, Deep-sets shows lower deviation between 10% and 60% across truth Higgs p_T range.

Hyper-parameters	Values
Number of ϕ layers	3
Number of ρ layers	2
ϕ N_{Nodes}	[240, 240, 128]
ρ N_{Nodes}	[240, 128]
Dropouts	0.2
Batch-size	128
Learning rate	0.001

Table 13 – The list of hyper-parameter values for the Deep-sets model used to perform Higgs p_T regression.

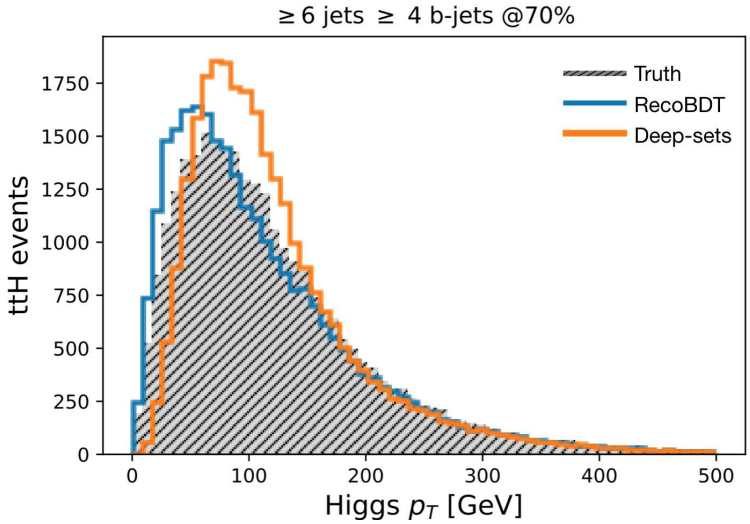


Figure 89 – The Higgs p_T distribution [GeV] where (solid gray) shows the truth Higgs p_T spectrum, (blue) shows the reconstructed Higgs p_T spectrum with highest recoBDT score and (orange) shows the predicted Higgs p_T spectrum from the Deep-sets regression network

The preliminary DNN setup using deep-set produced positive results, which formed the basis for further developments in MVA. Since the cross-section measurements are performed per STXS

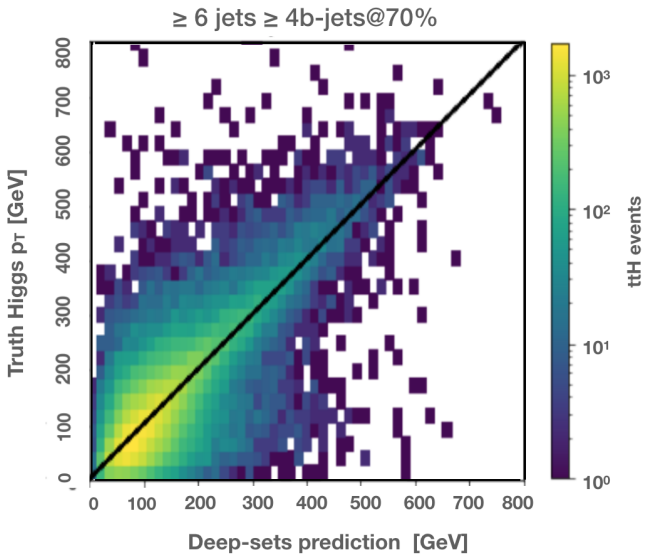


Figure 90 – Scatter plots showing truth Higgs p_T vs Deep-sets prediction p_T correlations.

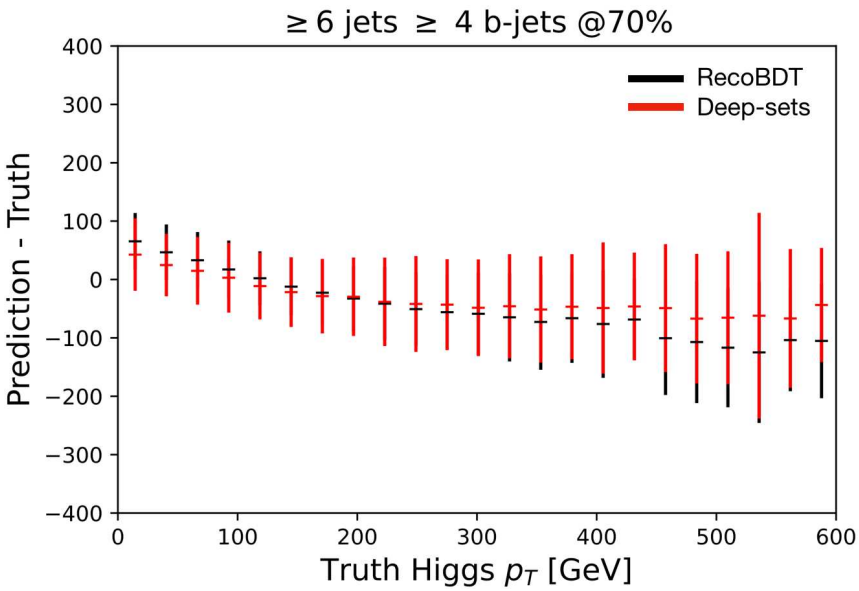


Figure 91 – The profile distribution showing (prediction - truth) as a function of truth Higgs p_T using reconstructed Higgs p_T (black) and Deep-sets prediction p_T (red). The dashed line is the mean value and the vertical error bar corresponds to the standard deviation.

bin, the Deep-sets performance was also investigated by converting the Deep-sets regression model to a multi-classifier, with STXS p_T bins as different classes. Figure 92 shows the Migration matrix compared between the Deep-sets regression model and the multi-classifier model. In case of regression model, the STXS bins are chosen based on the predicted p_T distribution while in multi-classifier, predicted STXS class outputs are directly used.

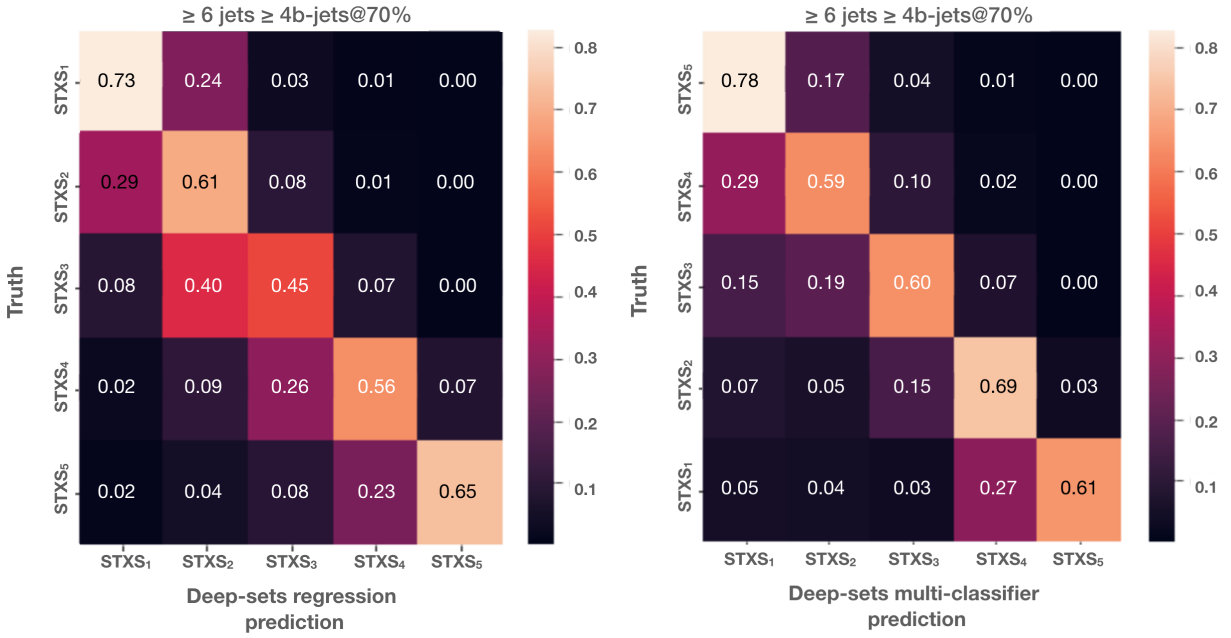


Figure 92 – Migration matrix showing (a) Deep-sets regression and (b) Deep-sets multi-classifier predictions. The matrix is normalised along the row.

The training was performed using the similar hyper-parameters, listed in Table 13. Deep-sets multi-classifier performance shows better performance in most of the STXS bins. In addition to the performance, multi-classifier approach was also beneficial in terms of the easier extension to background classes. Thus, it was reasonable to use the Deep-sets multi-class model as the baseline DNN configuration. The final single-step deep-set MVA model, with dedicated STXS and background classes, is discussed in detail in the following Section 4.4.5.

4.4.5 DeepSet-based MVA design and strategy

The Deep-sets multi-classifier model, as described in the previous section, which used to classify events into different STXS bins, showed overall better performance when compared to recoBDT. This multi-classifier model was extended to the background classes in order to simultaneously perform multi-classification to assign events to different STXS bins and to separate $t\bar{t}H$ signal

from the $t\bar{t}$ + jets background sub-components in a single-step. As mentioned before, the Deep-sets multi-classifier model had six different STXS classes defined. The model was extended by further adding four different corresponding to $t\bar{t}$+1b, $t\bar{t}$+B, $t\bar{t}+ \geq 2b$ and $t\bar{t}+ \geq 1c$ backgrounds. $t\bar{t}$+light was initially considered, but was later dropped due to the low fraction of $t\bar{t}$+light events in the training (1% of total background events used in the training). The idea to develop a new MVA model which performs the reconstruction of Higgs p_T and classification of ttH and sub-background components in a single step was effectively achieved by multi-class model using Deep-sets multi-classifier model (referred as Deep-sets hereafter). The final deep-set multi-class model used for the new MVA studies is shown in Figure 93. The final Deep-sets model has a total of 10 output classes

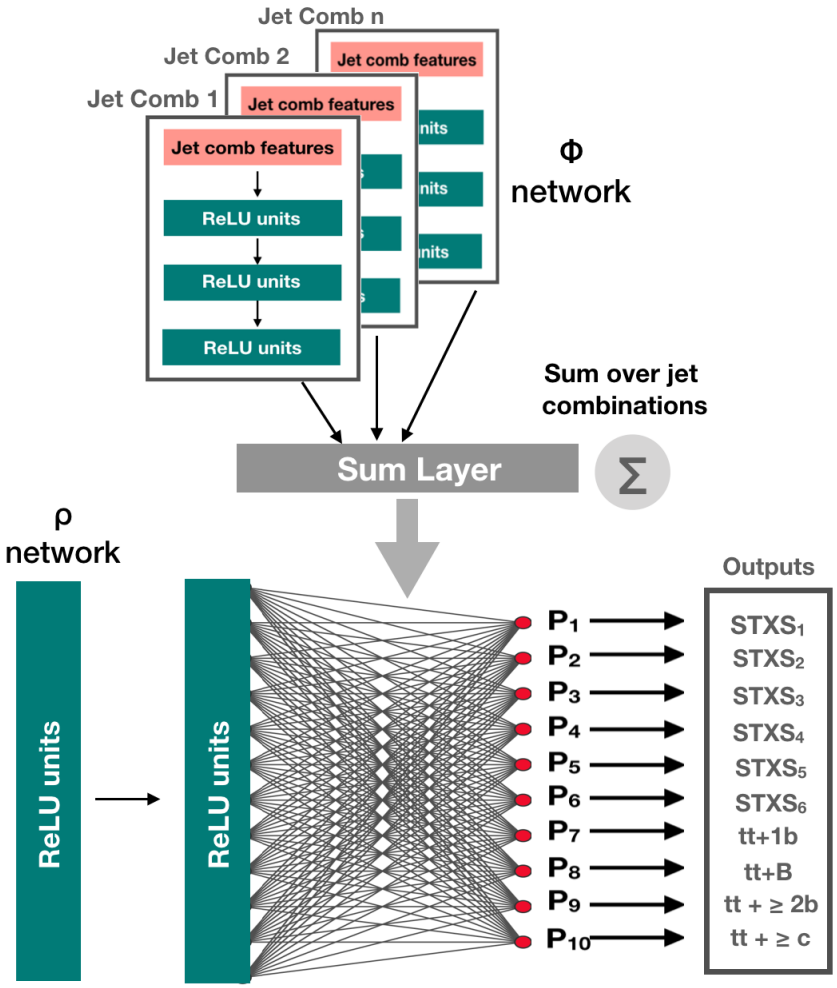


Figure 93 – Deep-set multi-class architecture showing 10 output classes with a probability assigned to them, which add up to 1. The first six multi-classifier outputs corresponds to the STXS classes while other four corresponds to the $t\bar{t}$+1b, $t\bar{t}$+B, $t\bar{t}+ \geq 2b$ and $t\bar{t}+ \geq 1c$ background classes.

Event weights[3] are used during the training and passed as the sample weights (see Section 4.4.1).

[3] Event weights takes into account different factors like cross-section of the process, collected luminosity

MC event generators used to produce $t\bar{t}H$ and $t\bar{t}$ + jets backgrounds processes can have events with negative event weights. Negative weights are not supported by the DNN architecture in general, so events with negative event weight (quite small fraction) must be removed from training. To avoid losing a large number of training events, the absolute value of the event weights are passed as sample weights.

The events are weighted in order to give equivalent training statistics for signal and total background. Since the Higgs p_T spectrum is not present in background samples, up sampling is not considered. Due to the limited statistics involved, performing down sampling is also not preferred. Table 14 lists the fraction of training events after applying event weights in each output class. The weighted number of events in the backgrounds classes i.e $t\bar{t}$+1b, $t\bar{t}+ \geq$2b, $t\bar{t}$+B, $t\bar{t}+ \geq$c are also very different. Among all the background sub-components, $t\bar{t}+ \geq$2b jets is the major background. Hence, it was better for the network to focus more on the $t\bar{t}H$ vs $t\bar{t}+ \geq$2b discrimination rather than the $t\bar{t}H$ vs $t\bar{t}$+B discrimination because the fraction of $t\bar{t}$+B events is 2% of the total backgrounds. Thus, the background classes are not relatively balanced in the training.

Both $t\bar{t}H$ and $t\bar{t}$+jets samples are used for the Deep-sets training. Cross-training is performed to cover full MC event statistics. The training is performed for 120 epochs to reach convergence. The training is terminated if no new validation minimum is achieved after training for 25 consecutive epochs. Since the deep-set performs multi-classification, the loss function is categorical cross-entropy, and the output layer employs a softmax activation function. The final output has 10 output classes, which assign a probability per event to each of the 10 classes. The class with the highest probability is used as the predicted class. The Deep-sets model was initially optimised by hand and further improved by performing hyper-parameter optimizations using grid search. In grid search, the set of candidate values for each variable parameter is first defined. The models are then trained for all possible parameter combinations, and the one with the best testing performance is chosen. Table 15 shows the list of grid-searched hyper-parameters for the Deep-sets. The optimised hyper-parameter values used in the final training are shown in bold.

4.5 MVA performance using DeepSet-sets

This section will present the performance of the Deep-sets MVA method that was developed in context to the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis. Since, the newly developed MVA model is intended to replace the recoBDT and the classBDT, the performance of the Deep-sets is compared to the BDTs. The measure of the performance is done via migration matrix, which can tell how the model behaves on various classes. It is an *N X M* matrix representation of predicted and actual classes, where *N* is

and various correction factors.

Deep-sets classes	Fraction of events
[0, 60)	0.09
[60, 120)	0.15
[120, 200)	0.14
[200, 300)	0.08
[300, 450)	0.03
[450,∞)	0.01
$t\bar{t}$+1b	0.05
$t\bar{t}$+B	0.02
$t\bar{t}+ \geq$2b	0.40
$t\bar{t}+ \geq$c	0.03

Table 14 – List of the fraction of training events in each output class. The $t\bar{t}H$ events are re-weighted to have an an equivalent training statistics to that of the total background.

Hyper-parameters	Values
Number of ϕ layers	**3**
Number of ρ layers	**2**
ϕ N_{Nodes}	128, 256, **512**
ρ N_{Nodes}	64, **128**, 256
Dropouts	0.2, 0.1, **0.5**
Batch-size	64, **128**, 256
Learning rate	0.01, **0.001**, 0.0001

Table 15 – Grid-searched hyper-parameters for the Deep-sets model. The optimized values used in the final training is shown in bold.

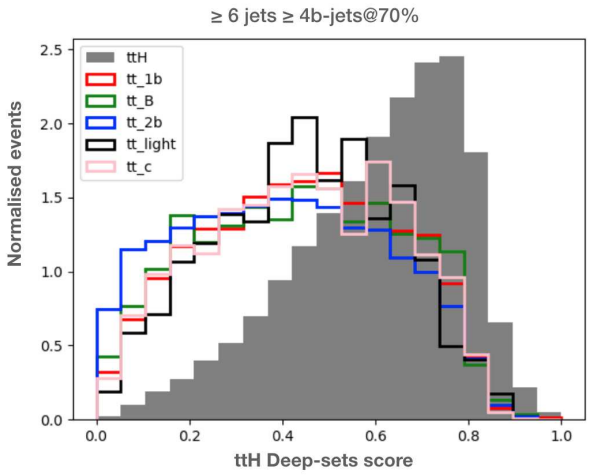
the number of truth classes and M is the number of predicted classes in the Deep-sets model. The migration matrix shown in this section has $N > M$ as $t\bar{t}$+light truth class is also explored in order to observe the Deep-sets model's performance for $t\bar{t}$+light events. The x-axis of the matrix represents the predicted class value, and the y-axis represents the truth class value. This provides a method for examining how effectively the model identified specific class, and it also enables to discover which class the model is not correctly classifying. Two kinds of migration matrices are shown in this section. The migration matrix can be normalised along either the predicted or true classes. When normalised along the predicted classes, the matrix gives information about the composition of each predicted class. When normalised along the truth class, it instead reflects the repartition of the truth class events between the predicted classes. ROC curves are employed to evaluate classification performance using Deep-sets. The ROC curves shown in this section have a Y axis for *true positive rate* and an X axis for *false positive rate*. This means that the *ideal* point is in the top left corner of the plot, with a false positive rate of zero and a true positive rate of one.

The first look at the raw Deep-sets outputs is discussed in Section 4.5.1. The Deep-sets MVA performance is presented in the Section 4.5.2. Comparison studies of the reconstruction and classification performance of the Deep-sets model to those of the BDTs are described in Section 4.5.3.

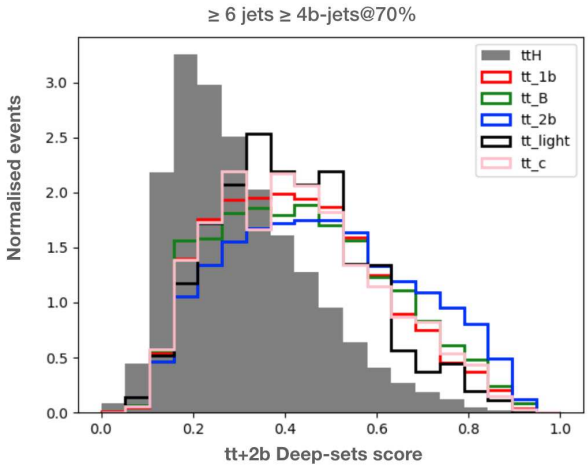
4.5.1 First look at the Deep-sets outputs

The Deep-sets MVA model produces 10 different discriminant scores, of which the first six are the STXS scores and the rest four are the different background scores, i.e $t\bar{t}$+1b, $t\bar{t}$+B, $t\bar{t}+ \geq 2b$, $t\bar{t}+ \geq 1c$. The sum of the first six STXS class scores gives the total probability of an event being a signal i.e $t\bar{t}H$ event. This is referred to as the $t\bar{t}H$ score. Figure 94 shows the $t\bar{t}H$ score and $t\bar{t}+ \geq 2b$ score distribution obtained from Deep-sets model, shown for $t\bar{t}H$ and different $t\bar{t}$+jets component samples. The highest discrimination is observed between $t\bar{t}H$ and $t\bar{t}+ \geq 2b$, showing good separation. Additionally, $t\bar{t}$+1b, $t\bar{t}$+B, and $t\bar{t}+ \geq 1c$ scores have consistently lower values.

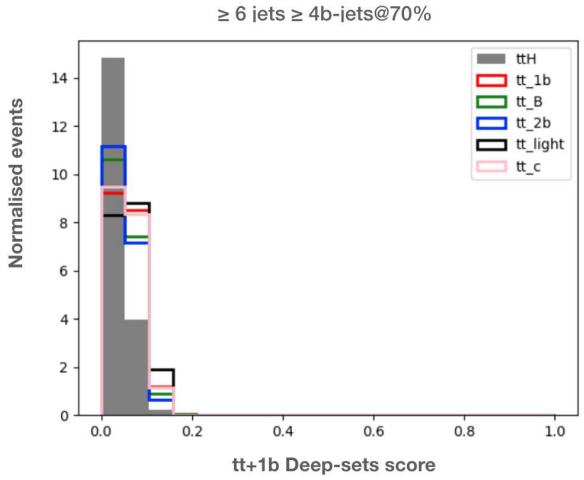
In case of $t\bar{t}$+1b and $t\bar{t}+ \geq 1c$, some level of some separation is also observed between the background sub-components. However, the separation among $t\bar{t}+1b$, $t\bar{t}+B$ and $t\bar{t}+ \geq 1c$ does not look significant, and merging of these non-$t\bar{t}+ \geq 2b$ classes can be considered. Figure 95 shows a migration matrix with 10 classes, normalised along the row. The matrix shows the fraction of events that enter the actual category after the prediction of the classifier. Here, $t\bar{t}+ \geq 2b$ truth events are distributed with 77% in predicted $t\bar{t}+ \geq 2b$ category, 3 % in the first STXS bin (0-60 GeV), 7% in the second STXS bin (60-120 GeV) and so on. Clearly, the output probability for $t\bar{t}+1b$, $t\bar{t}+B$ and $t\bar{t}+ \geq 1c$ is consistently low, mainly due to the lower fraction of these events entering the training. The current network only classifies background-like events into $t\bar{t}+ \geq 2b$ category which had large fraction of events in the training (around 40%) compared to other sub-background



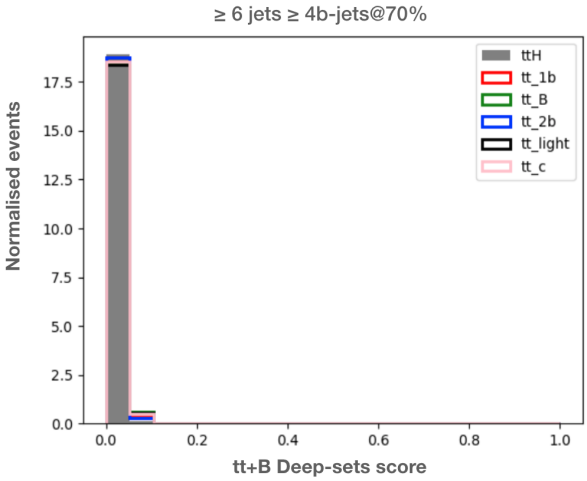
(a)



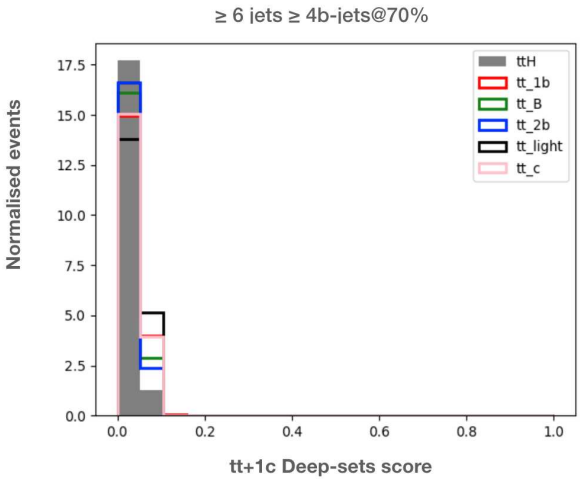
(b)



(c)



(d)



(e)

Figure 94 – The distribution of (a) $t\bar{t}H$, (b) $t\bar{t}+ \geq 2b$, (c) $t\bar{t}+1b$, (d) $t\bar{t}+B$ and (e) $t\bar{t}+ \geq 1c$ score scores for ttH signal and different background sub-components.

categories (around 10%). In order to fix this bias towards the $t\bar{t}+ \geq 2b$ events, the raw output discriminant scores are re-scaled by their corresponding fraction of events in the training, listed in Table 14. The Deep-sets model was consequently integrated in the central analysis framework. s performed on unbalanced background classes with unnormalized loss function.

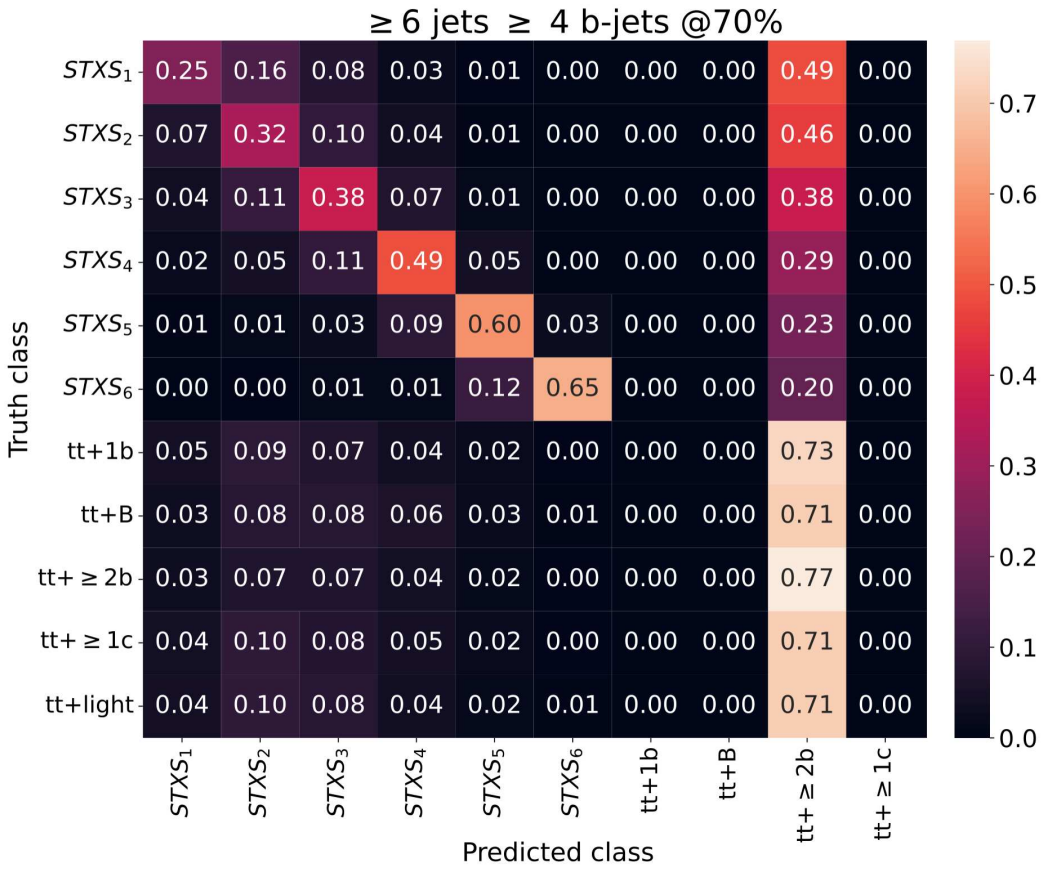


Figure 95 – Migration matrix of Deep-sets multi-classifier with 10 classes which shows the fraction of events that enter the actual category after the prediction of the classifier. The matrix is normalised along the truth classes.

4.5.2 Deep-sets performance

The two-step categorization is implemented to analyse the overall Deep-sets MVA performance using the re-scaled Deep-sets outputs. The aim is to give more emphasis to $t\bar{t}H$ vs background discrimination and then deal with the STXS assignment of the predicted $t\bar{t}H$ events. These steps are described as follows:

1. Step I: The step involves classification using inclusive $t\bar{t}H$ as a single class. The events can be assigned to either of the $t\bar{t}H$/ $t\bar{t}+ \geq 2b$/ $t\bar{t}$ +1b/ $t\bar{t}$+B/ $t\bar{t}+ \geq 1c$ categories, where the events go to the class which has highest output probability after re-scaling.

2. Step II: All the events, including $t\bar{t}H$ and background, are assigned to STXS classes using re-scaled STXS scores. The assignment of the STXS classes in the background categories is done to study the p_T shape of the dominant background in the STXS region, discussed in Section 5.5

Figure 96 shows how an event is assigned to one of the predicted classes out of the 10 classes based on the two-step categorization using the re-scaled outputs. Figure 97 shows the performance after implementing step I, as described above.

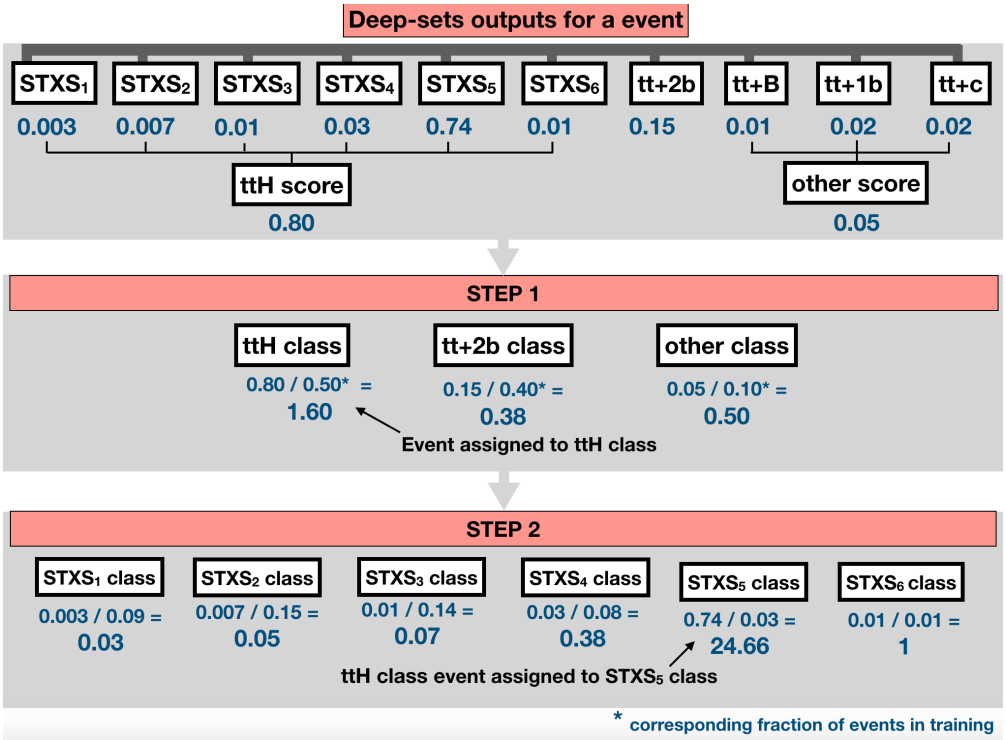
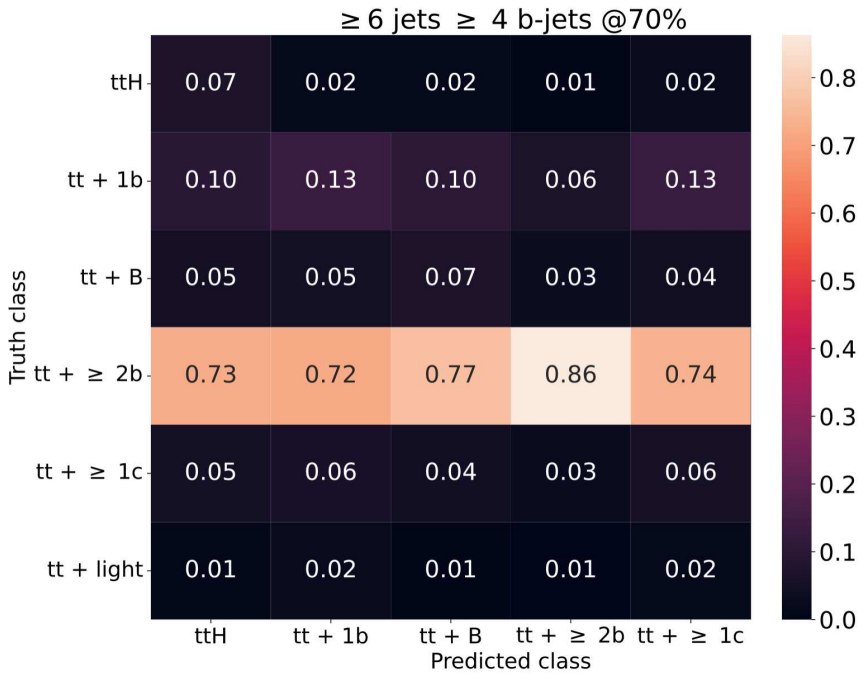
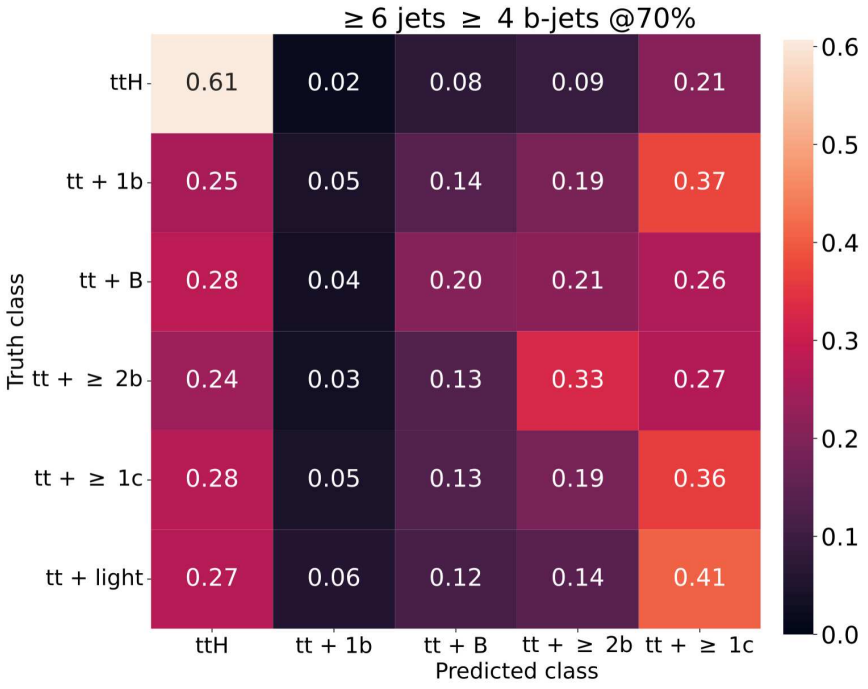


Figure 96 – An illustration of performing Step I and Step II categorization. The probability values of a given event are arbitrary.



(a)



(b)

Figure 97 – Migration matrix of Deep-sets multi-classifier after Step I categorization, normalised along the (a) predicted classes and (b) truth classes.

As a result of re-scaling, the bias towards $t\bar{t}+ \geq 2b$ is no longer observed as few events now fall in $t\bar{t}+1b$, $t\bar{t}+B$ and $t\bar{t}+ \geq 1c$ categories as well. Figure 98 shows the overall Deep-sets performance via the migration matrix normalised along the predicted class after implementing step I and step II. As hinted before, the performance in $t\bar{t}+1b$, $t\bar{t}+B$ and $t\bar{t}+ \geq 1c$ are similar and do not vary significantly. Thus, the output probabilities are merged, and the combined discriminant score will be used for these sub-background categories in the analysis. Figure 99 shows the similar migration matrix normalised along the column but here the classes $t\bar{t}+1b$, $t\bar{t}+B$ and $t\bar{t}+ \geq 1b$ are represented as merged referred as *other* predicted class. A dedicated training was done using the same deep-sets MVA network, but reducing the total number of classes from 10 to eight classes where $t\bar{t}+1b$, $t\bar{t}+B$ and $t\bar{t}+ \geq 1c$ classes are merged into one at the time of training. The overall performance was found to be similar to the performance when $t\bar{t}+1b$, $t\bar{t}+B$ and $t\bar{t}+ \geq 1c$ scores are treated separately during the training but merged during the evaluation. Keeping the initial setup provides flexibility to refer to their individual output scores and will be used subsequently. Figure 100 shows the performance via migration matrix normalised along the row obtained after step I and then Step II followed by merging $t\bar{t}+1b$, $t\bar{t}+B$ and $t\bar{t}+ \geq 1c$ categories.

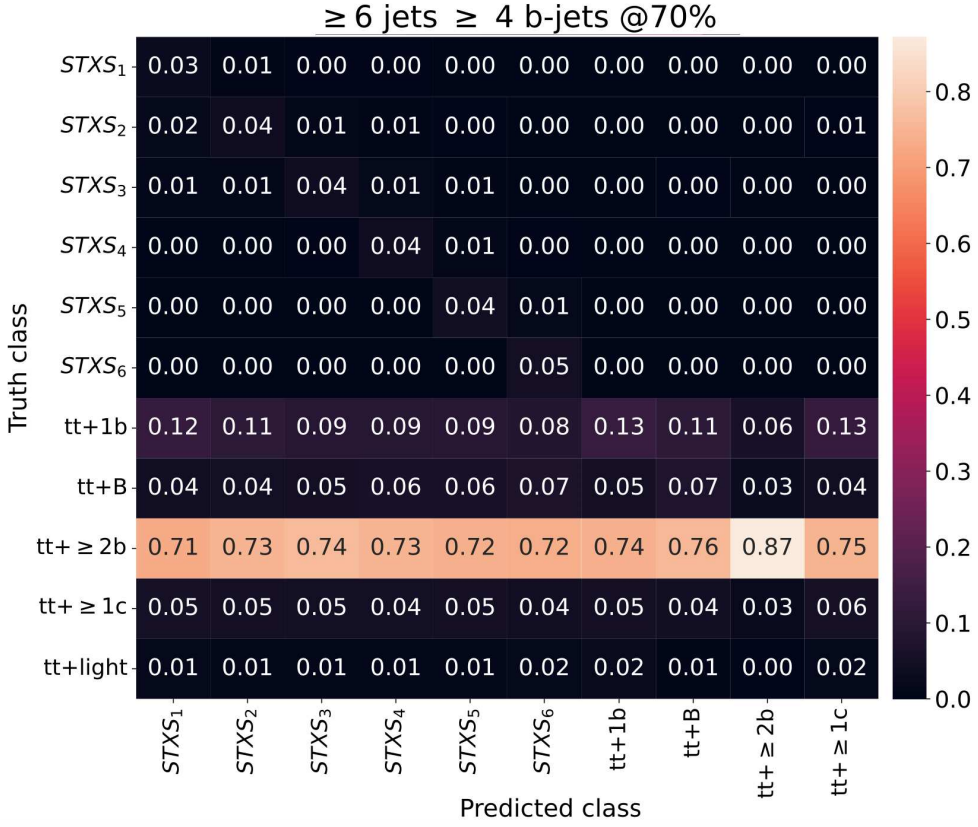


Figure 98 – Migration matrix of Deep-sets multi-classifier after Step I and Step II categorization. The matrix is normalised along the predicted classes.

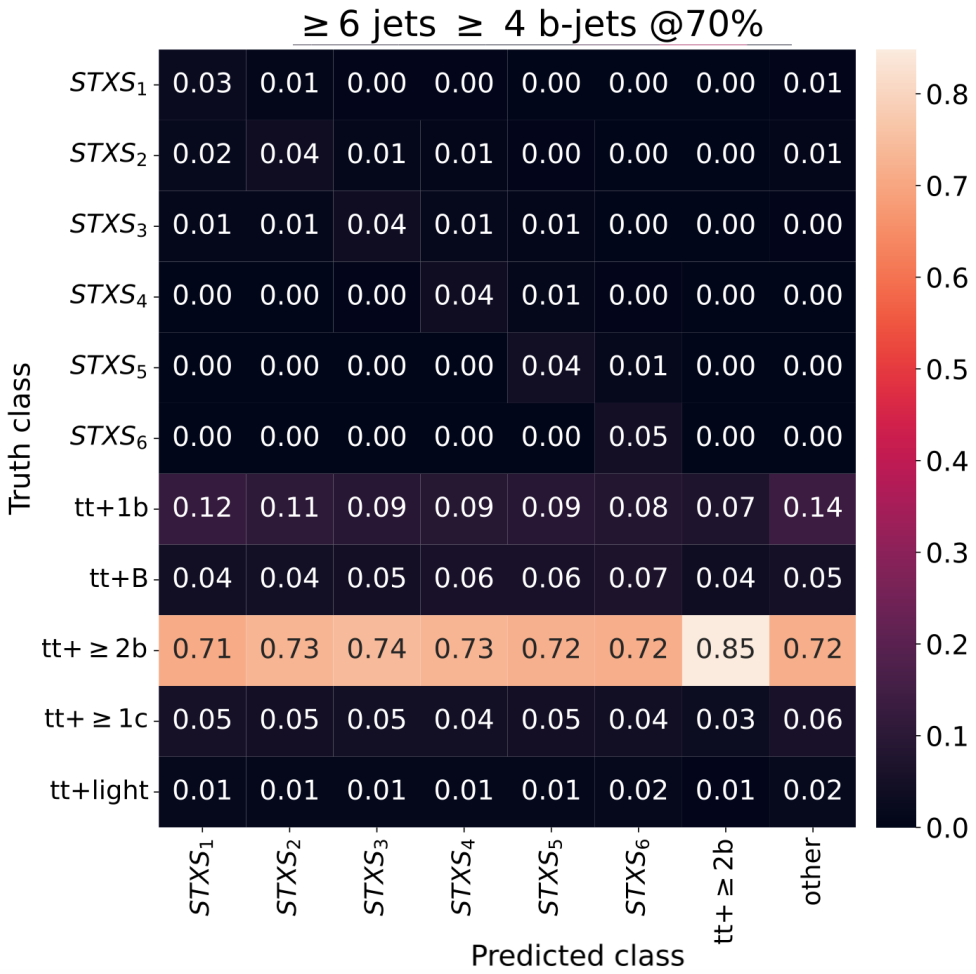


Figure 99 – Migration matrix of Deep-sets multi-classifier after Step I and Step II categorization and merging $t\bar{t}+1b$, $t\bar{t}+B$ and $t\bar{t}+\geq 1c$ categories into "other" category. The matrix is normalised along the predicted classes.

4.5.3 Deep-sets vs BDTs comparisons

The performance of the predicted STXS classes for all the events is compared between Deep-sets and recoBDT. In case of recoBDT, the reconstructed Higgs p_T for the highest recoBDT score is used to define the STXS classes. Figure 101 shows the migration matrix of all the predicted STXS classes, normalised along the row. The diagonal covering the STXS classes represents the fraction of all the events that are correctly predicted in each STXS bin. Deep-sets shows significantly improved performance for regions with p_T^H ($STXS_{3..6}$). This improvement can majorly impact overall analysis due to the better STXS bin assignment in the the higher p_T regions.

The classification performance is illustrated via ROC curves, as shown in Figure 102. The separation between the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background is comparable between the two setups with AUC value of around 76%. In addition, Deep-sets offer the advantage of discrimination between

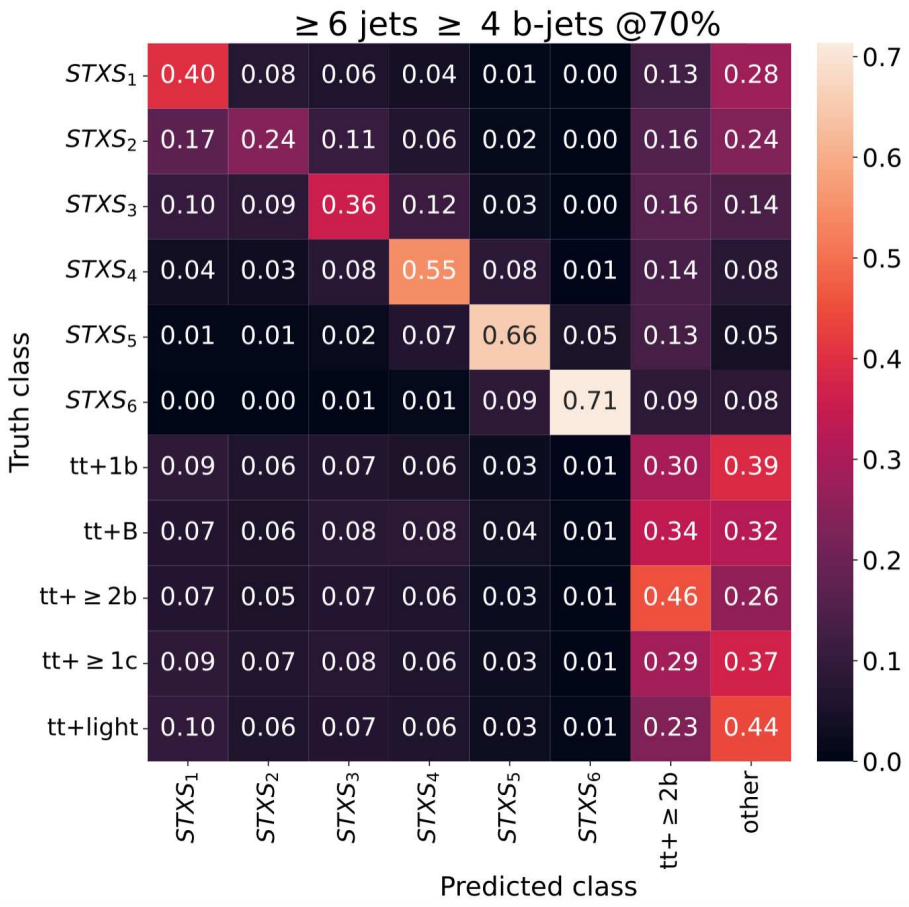
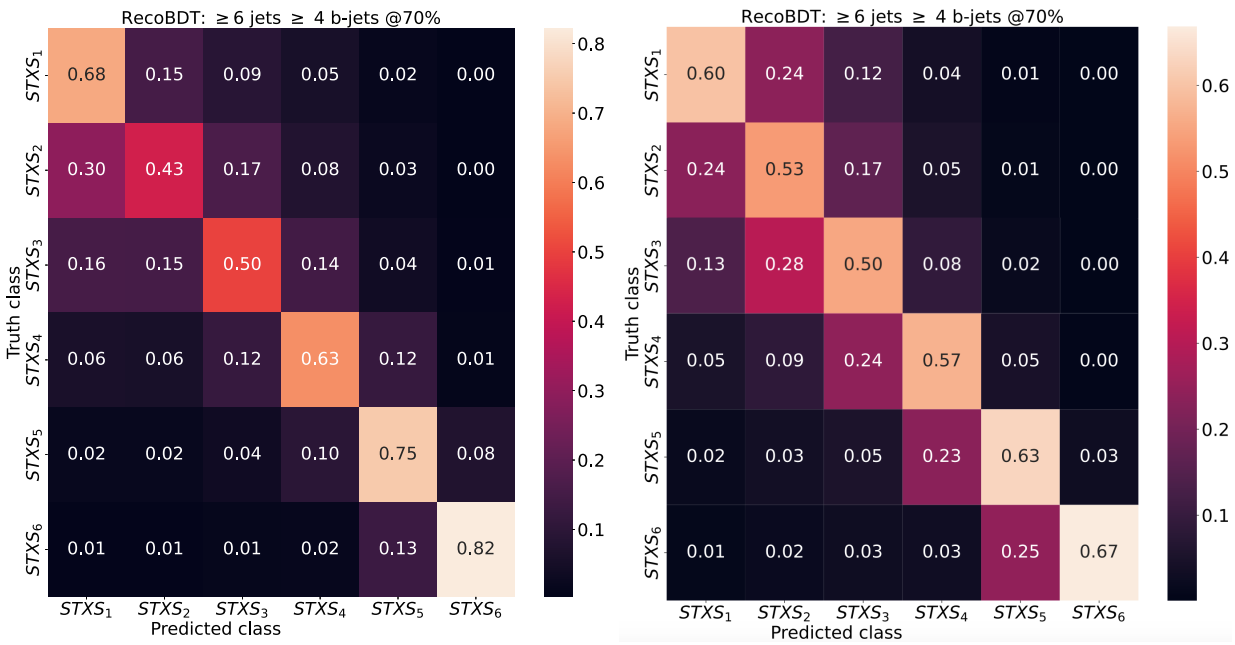


Figure 100 – Migration matrix of Deep-sets multi-classifier after Step I and Step II categorization and merging $t\bar{t}+1b$, $t\bar{t}+B$ and $t\bar{t}+\geq 1c$ categories into "Other" category. The matrix is normalised along the truth classes.

additional background sub-components, i.e $t\bar{t}+\geq 2b$ and other. The discrimination between the $t\bar{t}H$ signal and the $t\bar{t}+\geq 2b$ background is lower but still similar to classBDT performance. The AUC value for classifying other background against $t\bar{t}H$ is 67%.

The performance of the new Deep-sets model seems promising when compared to BDTs. In addition, the Deep-sets model benefits from the dedicated background node for the non-$t\bar{t}+\geq 2b$ background. The Deep-sets model was consequently integrated into the central analysis framework. The integration into a full analysis strategy is going to be discussed in the next Chapter 5.



(a) (b)

Figure 101 – Migration matrix of STXS predicted classes for (a) Deep-sets and (b) RecoBDT. The matrix is normalised along the truth classes.

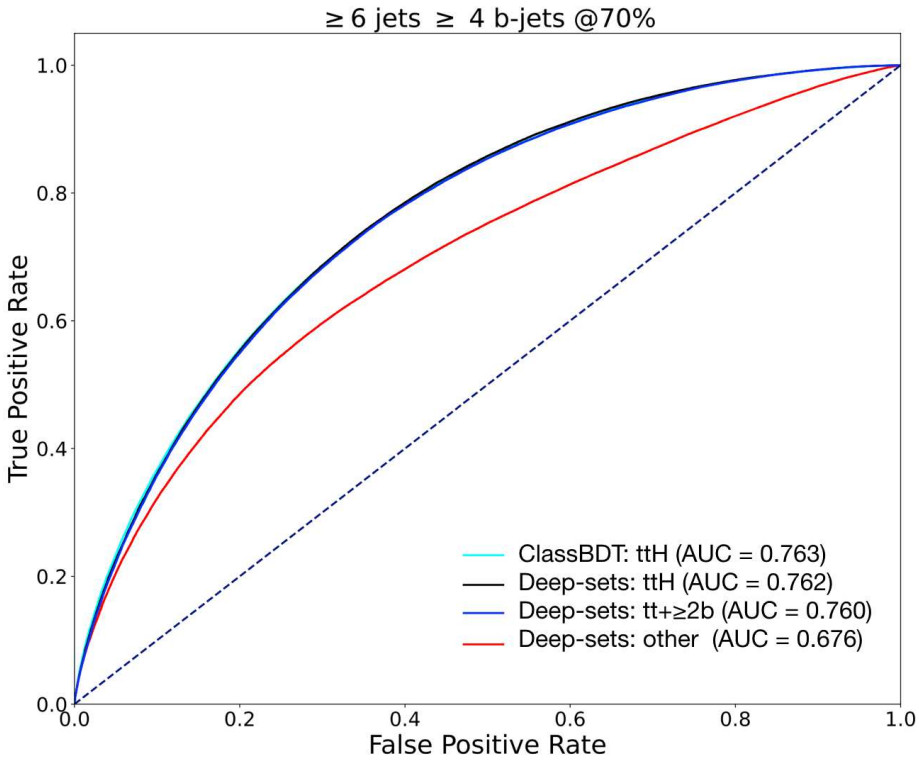


Figure 102 – ROC curves showing the classification performance of Deep-Sets multi-classifier and classBDT.

CHAPTER

5

SEARCH FOR THE HIGGS BOSON IN THE SINGLE LEPTON $t\bar{t}H(H \to b\bar{b})$ CHANNEL

The associated production of a Higgs boson with a top quark pair ($t\bar{t}H$) allows for a direct measurement of the top-Yukawa coupling. Its precise measurement is a powerful probe to unravel new physics phenomena [23] like the new sources of CP violation that might exist beyond the SM. This thesis explores the measurement of $t\bar{t}H$ production, where the Higgs boson decays into a *b*-quark pair. The $t\bar{t}H(H \to b\bar{b})$ analysis presented in this chapter accounts for the legacy round, which is built on the previous round [33] using full Run 2 ATLAS data. The objective is to further optimise the analysis by employing new methodologies and improving object definitions. As mentioned in Chapter 4, the legacy round will employ new MVA methods developed using Deep-sets, as the discriminant for the signal extraction fit in the single-lepton channel. The analysis uses dedicated event selections to enhance signal contribution and suppress background. The events are then divided into signal-regions (SRs) and control regions (CRs), which are signal-enriched and signal-depleted categories, respectively. Deep-sets outputs are used as discriminants to assign events to STXS bins and to separate signal and background processes. The combined SRs and CRs are used to perform a profile-likelihood fit while accounting for the systematic uncertainties. Section 5.1 provides an overview of the $t\bar{t}H(H \to b\bar{b})$ analysis. The selection criteria for reconstructed objects and events used in this analysis are described in Section 5.2 followed by an overview of the modelling of the signal and background processes in Section 5.3. The analysis regions are discussed in Section 5.4. Section 5.5 provides an overview of the sources of the systematic uncertainties used in this analysis, followed by an introduction to the profile-likelihood fit in Section 5.6. The final results of the analysis are presented in Section 5.7.

5.1 The $t\bar{t}H(H \to b\bar{b})$ analysis in a nutshell

In $t\bar{t}H(H \to b\bar{b})$ analysis, there are three different channels based on the decay mode of the W-bosons that come from the decay of the $t\bar{t}$ pair:

- **All-hadronic channel**: This channel has the highest branching ratio of about 45.7%, where both W-bosons decay hadronically. Since there are no leptons in the final state, the multi-jet background is difficult to suppress.
- **Dilepton channel**: This channel has the smallest branching ratio of 10.5%, where both W-bosons decay leptonically. Due to the presence of two neutrinos from W-boson decays, the reconstruction of the final states becomes less precise. However, the leptons in the final states offer the cleanest signature to suppress the multi-jet background.
- **Single-lepton channel**: This channel corresponds to the events where one W-boson decays hadronically and the other one leptonically. The branching ratio is about 43.8%, which is very close to the hadronic channel but offers a relatively clean topology with a lepton in the final state to suppress the multi-jet background. In addition, only one neutrino is there in the final state and, thus can be fully determined from the event E_T^{miss}. Figure 103, shows the Feynman diagram of $t\bar{t}H(H \to b\bar{b})$ production in the single-lepton channel. Since the branching ratio is fairly high, even the higher p_T regions have reasonable statistics, which is essential for the STXS measurements. In the previous analysis, the single-lepton channel was further divided into *resolved* and *boosted* channels [33]. Resolved refers to the events where Higgs boson decay products are reconstructed as two separate jets, whereas boosted refers to the events where the decay products are reconstructed into a single large-R jet.

The analysis presented here is carried out in the resolved single-lepton channel. Both the boosted channel and the dilepton are analysed separately and are not covered in this thesis. The boosted channel may overlap with the resolved channel, particularly in the $p_T > 300$ GeV region. This is not addressed in this thesis, but will be addressed in the final combination fit, which includes all three channels to extract the $t\bar{t}H$ signal cross-section.

In the single-lepton channel, six jets are expected in the final state, with four of them being *b*-jets. Furthermore, it includes exactly one isolated lepton in the final state. The term lepton (ℓ) can refer to an electron or a muon. The leptonic decay of τ leptons into electrons and muons is also taken into account. The complex final state topology of this channel induces several challenges, the most significant of which is the assignment of reconstructed objects to the decay products of the top quarks and the Higgs boson. Another major challenge for this process is the large $t\bar{t}$ + jets

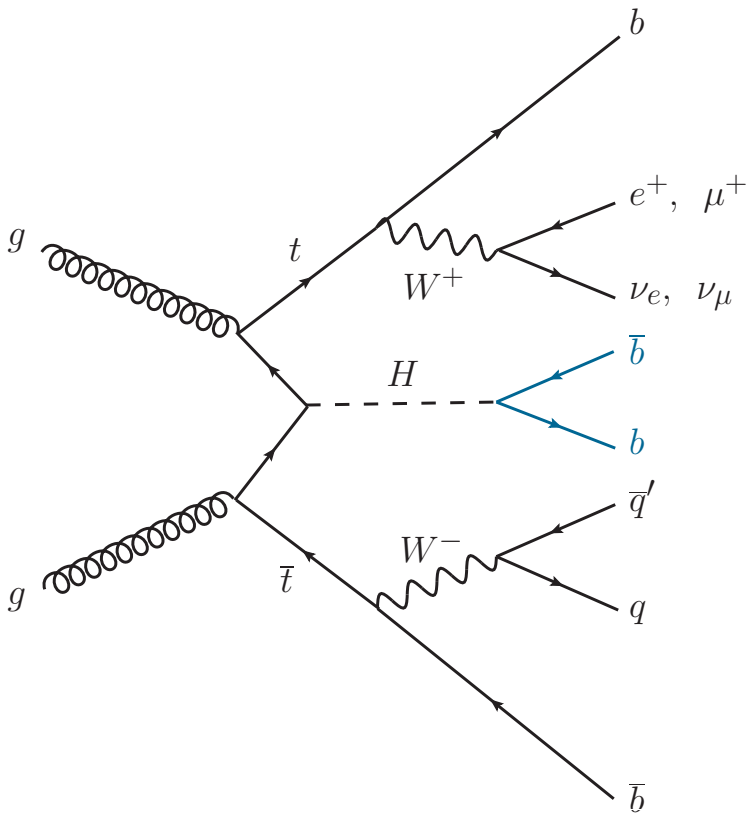


Figure 103 – Feynman diagram of the Higgs boson production with associated top quarks, where the Higgs boson decays into a *b*-quark pair. The final state of this process corresponds to the the single-lepton channel.

background, in particular when these jets originate from *b*-quarks. This background has a much larger production cross-section than the $t\bar{t}H$ signal. Figure 104 shows a Feynman diagram for the $t\bar{t}$ + $b\bar{b}$ production process. This process suffers from large theoretical uncertainties and is poorly constrained by data measurements, which limits the sensitivity of this analysis. The analysis will explore the STXS formalism, described in Section 1.8.2, where cross-sections are measured as a function of the STXS bins defined using reconstructed Higgs p_T $\left(p_T^H\right)$. In the previous round, five different STXS classes corresponding to the 0-120 GeV, 120-200 GeV, 200-300 GeV, 300-450 GeV, and 450-∞ GeV regions were used. For the legacy round, the first STXS class (0–120 GeV) has been split into two i.e 0–60 GeV and 60–120 GeV, to further improve the measurement sensitivity in the lower p_H^T regions. Thus, a total of six STXS bins will be used for performing the STXS cross-section measurement of the $t\bar{t}H$ signal.

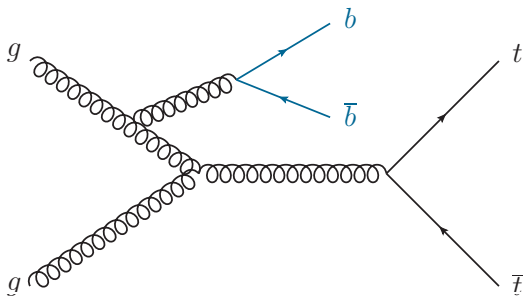


Figure 104 – A Feynman diagram of the main $t\bar{t}$ + $b\bar{b}$ background.

5.2 Event selection

The events used for performing this analysis are selected from the proton-proton collisions at $\sqrt{s}$ = 13 TeV recorded by the ATLAS experiment during Run 2 at the LHC between 2015 and 2018, with an integrated luminosity of 139 fb^{-1}. The events are recorded using single-lepton triggers that have different p_T thresholds. These triggers are combined in a logical "OR" to ensure higher selection efficiency. The events must either satisfy a low p_T threshold and lepton isolation, or a higher trigger threshold with a looser identification criterion and without any isolation requirement. For data collected in 2015 (2016–2018), the lowest p_T threshold for the single-muon trigger was 20 (26) GeV, whereas for the single-electron trigger, it was 24 (26) GeV. Table 16 shows the summary of single-lepton triggers used in this analysis. The reconstruction methods for the physics objects used in this analysis are described in Section 2.4.

lepton	p_T (GeV)		Identification		Isolation	
	2015	2016-2018	2015	2016-2018	2015	2016-2018
electrons	24	26	medium	tight	-	loose
	60	60	medium	medium	-	-
	120	140	loose	loose	-	-
muons	20	26	medium	medium	loose	medium
	50	50	medium	medium	-	-

Table 16 – Single-lepton trigger requirements used for $t\bar{t}H(H \to b\bar{b})$ analysis, shown separately for the data taken in 2015 and 2016–2018 [134].

Electrons are required to have $p_T > 10$ GeV and $|\eta| < 2.47$ (excluding $1.37 < |\eta| < 1.52$). Muons need to satisfy $p_T > 10$ GeV and $|\eta| < 2.5$ selection. Electrons and muons need to fulfil the Medium [82] and Loose [86] identification criteria, respectively. In order to match the lepton tracks to the PV of the event, they have to fulfill $|z_0 \sin(\theta)| < 0.5$ mm and $d_0/\sigma(d_0) < 5$ (3) criteria for

electrons (muons). The efficiency of the trigger chain used to record electrons and muons, as well as their reconstruction and identification, differs between data and MC. To account for this effect, appropriate lepton scale factors are applied to electrons and muons in the MC simulation.

The PFlow jets, as described in Section 2.4.4, are built using the anti-k_T algorithm with radius parameter 0.4. The jets are required to pass $p_T > 25$ GeV and $|\eta| < 2.5$. In order to reduce the pile-up effects, jets with low p_T have to pass through the medium WP of the jet vertex tagger (JVT).

The jets consisting of *b*-hadrons are identified with the high-level *b*-tagging algorithm DL1r. Four working points (WPs): 85%, 77%, 70% and 60% are used, defined based on the different thresholds for the DL1r discriminant. A pseudo-continuous *b*-tagging score is assigned to the jet, as described in Section 2.4.4. The correction factors are applied to the simulated events to compensate for differences between data and simulation in the *b*-jet efficiency, *c*- and light-jet mis-tag rate separately.

An overlap removal procedure between jets and leptons is used, so that a single detector response is not counted as more than one physics object in an event. It follows an iterative process, and only the objects that pass through all the previous steps are used in the next steps. If a muon shares a track with an electron, the electron is discarded. To prevent double-counting of electron energy deposits as jets, the jets within $\Delta R < 0.2$ of an electron are rejected. To reduce the impact of non-prompt electrons, the electron is rejected if a jet is present within $0.2 < \Delta R < 0.4$. In case a jet has fewer than three tracks and is within $\Delta R < 0.4$ of a muon, the jet is removed. However, if this jet is associated with at least three tracks, the muon is removed instead. This helps in the suppression of the muons from heavy-flavour decays inside the jets.

The selected events are required to have at least five jets, where at least four jets are tagged with the 70% *b*-tagging WP. Moreover, there should be exactly one lepton with $p_T > 27$ GeV in the event that satisfies the Tight and Medium identification criterion for electrons and muons, respectively. In addition, events with any additional lepton with a $p_T > 10$ GeV that satisfies the Medium (Loose) identification operation point for electrons (muons) are discarded. In order to maintain the orthogonality with other $t\bar{t}H$ channels [135], events with two or more hadronic τ candidates are removed.

5.3 Signal and background modelling

The $t\bar{t}H$ signal and the background processes are modelled using simulated event samples. These are generated using ATLAS simulation software, as described in Section 2.3. The MC samples used for nominal predictions of all the considered processes are generated using the full ATLAS detector simulation (FULLSIM) based on GEANT4. For some of the alternative samples used to

define modelling systematic uncertainties, the faster simulation (ATLFAST) is also used. The two simulations gives similar modelling for the observables used in the analysis. PYTHIA8 is used to simulate the pile-up interactions using a specific set of parameters, called *A*3 tune [136]. The simulated samples are also re-weighted to match the pile-up conditions observed in the full Run 2 data-set. Table 17 shows the list of the MC samples used in this analysis. It summarises both the nominal samples that are used for the baseline modelling and the alternative samples that are used to estimate systematic uncertainties. The PS and other multi-parton interactions (MPI), beam remnant and hadronisation processes are generated using either PYTHIA8 or HERWIG7 [73, 137], except the case where the samples are produced with SHERPA [71]. The *b*- and *c* hadron decays are simulated using EVTGEN v1.6.0 [113]. For PYTHIA 8, the A14 [138] set of tuned parameters with the NNPDF2.3LO PDF set [20] are used. For HERWIG 7, the H7UE tune is used with the MMHT2014LO PDF set [74]. Top quarks, *Z*- and *W*- bosons are decayed at leading order (LO) using MadSpin [139, 140] to preserve all spin correlations in all the samples generated using MADGRAPH5_aMC@NLO [70] for the ME, and in the single-top (t-channel) POWHEG samples. For most of the samples, the ME generator is used at NLO precision in QCD. The top-quark mass is set to m_{top} = 172.5 GeV and mass of *b*-quark coming from the top decay is set to m_b = 4.95 GeV when the decay is modelled by POWHEG BOX or MADSPIN, and to m_b = 4.75 GeV for the samples produced with SHERPA. To improve the precision, some samples are normalised to account for the QCD corrections up to next-to-next-to-leading order (NNLO) and the electroweak (EW) corrections up to next-to-leading order (NLO).

5.3.1 $t\bar{t}H$ signal modelling

The $t\bar{t}H$ signal events are modelled in the 5FS scheme (see Section 1.6) using POWHEG Box [69, 112, 147, 148] generator and PS and hadronisation processes with PYTHIA8. The ME generator is used at NLO in QCD with NNPDF3.0NLO PDF set [142]. The functional form of renormalisation and factorisation scales are $\mu_R = \mu_F = \sqrt[3]{m_T(t) \cdot m_T(\bar{t}) \cdot m_T(H)}$, where $m_T = \sqrt{m^2 + p_T^2}$ is the transverse mass of a particle. The h_{damp} [1] parameter is set to $0.75 \cdot (m_t + m_{\bar{t}} + m_H)$ = 325.5 GeV. The samples are normalised to the $t\bar{t}H$ cross-section of $\sigma_{t\bar{t}H} = 507^{+35}_{-50}$ fb, which includes NLO QCD and EW corrections for m_H = 125 GeV. All the Higgs boson decay modes are considered but the overall optimisation of the analysis is focused on $H \to b\bar{b}$ as other decay modes account for a negligible fraction of events after the single-lepton event selections. Two alternative samples: POWHEG+HERWIG7 and POWHEG+MADGRAPH5_aMC@NLO, are used to assess the uncertainty associated to the $t\bar{t}H$ modelling.

[1] The h_{damp} parameter controls the p_T of the first additional emission beyond the leading-order in the PS and therefore regulates the high-p_T emission against which the $t\bar{t}$ system recoils.

Process	ME generator	ME PDF	PS	Normalisation
Higgs boson				
$t\bar{t}H$	POWHEG Box v2	NNPDF3.0NLO	PYTHIA 8.230	NLO+NLO (EW)
	POWHEG Box v2	NNPDF3.0NLO	HERWIG 7.04	NLO+NLO (EW)
	MADGRAPH5_aMC@NLO 2.6.0	NNPDF3.0NLO	PYTHIA 8.230	NLO+NLO (EW)
tHjb	MADGRAPH5_aMC@NLO 2.6.2	NNPDF3.0NLO nf4	PYTHIA 8.230	-
tWH	MADGRAPH5_aMC@NLO 2.6.2 [DR]	NNPDF3.0NLO	PYTHIA 8.235	-
$t\bar{t}$+*jets backgrounds*				
$t\bar{t}$	POWHEG BOX v2	NNPDF3.0NLO	PYTHIA 8.230	NNLO+NNLL
	POWHEG BOX v2	NNPDF3.0NLO	HERWIG 7.04	NNLO+NNLL
	SHERPA 2.2.1	NNPDF3.1NLO nf4	SHERPA	NNLO+NNLL
$t\bar{t}$+bb	POWHEG Box Res	NNPDF3.1NLO nf4	PYTHIA 8.230	-
	POWHEG BOX v2	NNPDF3.1NLO nf4	HERWIG 7.04	NNLO+NNLL
	SHERPA 2.2.1	NNPDF3.1NLO nf4	SHERPA	-
Other backgrounds				
tW	POWHEG Box v2 [DR]	NNPDF3.0NLO	PYTHIA 8.230	NNLO+NNLL
	POWHEG BOX v2 [DS]	NNPDF3.0NLO	PYTHIA 8.230	NNLO+NNLL
	POWHEG Box v2 [DS]	NNPDF3.0NLO	HERWIG 7.04	NNLO+NNLL
	MADGRAPH5_aMC@NLO 2.6.2 [DR]	CT10NLO	PYTHIA 8.230	NNLO+NNLL
t-channel	POWHEG Box v2	NNPDF3.0NLO nf4	PYTHIA 8.230	NLO
	POWHEG Box v2	NNPDF3.0NLO nf4	HERWIG 7.04	NLO
	MADGRAPH5_aMC@NLO 2.6.2	NNPDF3.0NLO nf4	PYTHIA 8.230	NLO
s-channel	POWHEG BOX v2	NNPDF3.0NLO	PYTHIA 8.230	NLO
	POWHEG BOX v2	NNPDF3.0NLO	HERWIG 7.04	NLO
	MADGRAPH5_aMC@NLO 2.6.2	NNPDF3.0NLO	PYTHIA 8.230	NLO
W + jets	SHERPA 2.2.1 (NLO [2j], LO [4j])	NNPDF3.0NNLO	SHERPA	NNLO
Z + jets	SHERPA 2.2.1 (NLO [2j], LO [4j])	NNPDF3.0NNLO	SHERPA	NNLO
VV (had)	SHERPA 2.2.1	NNPDF3.0NNLO	SHERPA	-
VV (lep.)	SHERPA 2.2.2	NNPDF3.0NNLO	SHERPA	-
VV(lep.)+jj	SHERPA 2.2.2 (LO [EW])	NNPDF3.0NNLO	SHERPA	-
$t\bar{t}W$	MADGRAPH5_aMC@NLO 2.3.3	NNPDF3.0NLO	PYTHIA 8.210	NLO+NLO (EW)
	SHERPA 2.0.0 (LO [2j])	NNPDF3.0NNLO	SHERPA	NLO+NLO (EW)
$t\bar{t}\ell\ell$	MADGRAPH5_aMC@NLO 2.3.3	NNPDF3.0NLO	PYTHIA 8.210	NLO+NLO (EW)
	SHERPA 2.0.0 (LO [1j])	NNPDF3.0NNLO	SHERPA	NLO+NLO (EW)
$t\bar{t}Z$ (qq,$\nu\nu$)	MADGRAPH5_aMC@NLO 2.3.3	NNPDF3.0NLO	PYTHIA 8.210	NLO+NLO (EW)
	SHERPA 2.0.0 (LO [2j])	NNPDF3.0NNLO	SHERPA	NLO+NLO (EW)
tttt	MADGRAPH5_aMC@NLO 2.3.3	NNPDF3.0NLO	PYTHIA 8.210	NLO+NLO (EW)
tZq	MADGRAPH5_aMC@NLO 2.3.3 (LO)	CTEQ6L1	PYTHIA 8.212	-
tWZ	MADGRAPH5_aMC@NLO 2.3.3 [DR]	NNPDF3.0NLO	PYTHIA 8.230	-

Table 17 – List of the MC samples used in this analysis with their corresponding generator set-ups. The first row of each sample correspond to the nominal sample, and the rest are used to estimate systematic uncertainties. [33, 141]. '$nf4$' corresponds to the $N_f = 4$ versions of the NNPDF3.0 sets [142]. To handle interference with tt production, the diagram removal (DR) [143] scheme is used for tW production. To assess an uncertainty in the modelling of this interference, an additional sample using the diagram subtraction (DS) [143,144] scheme is used. In W+jets and Z+jets, NLO-accurate MEs are calculated for up to two partons ($2j$) and LO accurate MEs for up to four partons ($4j$). They were matched with the SHERPA PS using the MEPS@NLO [145, 146] prescription and the NNPDF3.0NNLO PDFs set. For the VV Sherpa samples, "*had*" ("*lep*") denotes that one decays leptonically and one hadronically (both bosons decay leptonically).

5.3.2 $t\bar{t}$+jets background modelling

Different MC samples are used to model the nominal $t\bar{t}$ + jets background as well as associated systematic uncertainties, each with a different PS and with a different level of precision in the ME generator. The $t\bar{t}$ + jets samples are divided into three main components based on the truth-flavour of the jets produced in addition to the top quark decays. The truth-flavour of the jets is defined using so-called truth-jets, formed from stable particles, and determined based on a matching criterion similar to the matching done for b-tagging, as described in Section 3.3.1. Events are labelled as $t\bar{t}+ \geq 1b$ if at least one *b*-jet is identified, or else as $t\bar{t}+ \geq 1c$ if at least one *c*-flavour jet is identified, and otherwise as tt+light. The $t\bar{t}+ \geq 1b$ events can be further split into $t\bar{t}+1b$ (exactly one *b*-jet), $t\bar{t}+B$ (exactly one *B*-jet[2]), and $t\bar{t}+ \geq 2b$ (at-least two *b*-jet).

The modelling of the $t\bar{t}+ \geq 1c$ and $t\bar{t}$+light events is performed in the 5FS using POWHEG BOX v2 with $\mu_R = \mu_R = \sqrt{m_{top}^2 + p_T^2} = m_{top}$ [3] and $h_{damp} = 1.5 \cdot \cdot m_{top}$ [144]. In order to precisely model the main $t\bar{t}+ \geq 1b$ background, a sample with $t\bar{t}+b\bar{b}$ MEs is produced at NLO QCD accuracy in 4FS scheme using POWHEG Box Res [149] generator and OpenLoops [150]. The *b*-quarks are treated as massive with a mass of $m_b = 4.75$ GeV. The setting of the POWHEG+PYTHIA8 samples are optimised [141] with respect to the $t\bar{t}+b\bar{b}$ samples used in the previous round of the $t\bar{t}H(H \to b\bar{b})$ analysis. This was done to improve $t\bar{t}+b\bar{b}$ modelling and optimise background uncertainty model. The optimization process involves lowering the renormalization scale in order to get a better match with the data. This resulted in the reduction of the NLO K-factor from 1.62 to 1.54, obtained by comparing the $t\bar{t}$ production cross-section with at least two additional *b*-jet or at least one at NLO to LO. These optimisations resulted in the reduction of events with negative weight by factor of two . The decay handling and spin correlation have also been changed, using the default method in POWHEG. In comparison to the ATLAS measurement, the fiducial cross-section of $t\bar{t}+b\bar{b}$ by POWHEG+PYTHIA8 predictions using different scale choices is shown in Figure 105. Different scale choices are described as follows:

- Scale 1: $\mu_R = \mu_R^{def}$, $\mu_F = \mu_F^{def}$
- Scale 2: $\mu_R = 0.5 \cdot \mu_R^{def}$, $\mu_F = \mu_F^{def}$
- Scale 3: $\mu_R = 0.5 \cdot \mu_R^{def}$, $\mu_F = 0.5 \cdot \mu_F^{def}$

[2] *B*-jets are the jets matched to more than one *b*-hadron, where the sub-leading *b*-hadrons have no p_T requirement.

[3] Since this scale is computed in the $t\bar{t}$ rest-frame, it follows that the p_T values of the top quark and the top anti-quark are identical.

where, $\mu_R^{def} = \sqrt[4]{m_T(t) \cdot m_T(\bar{t}) \cdot m_T(b) \cdot m_T(\bar{b})}$ and $\mu_F^{def} = 0.5$ x $\Sigma_{i=t,\bar{t},b,\bar{b}} m_T(i)$. The predictions made by using *scale 1* are consistent with the one used in the previous $t\bar{t}H(H \to b\bar{b})$ analysis. The *scale 2* setting shows the improved data/MC agreement compared to *scale 1* and *scale 3* , and is used for generating $t\bar{t}$+$b\bar{b}$ nominal and the alternative samples for the current analysis round. Two alternative samples are used to assess the uncertainty, simulated using POWHEG+HERWIG7 and using SHERPA 2.2.10 which implements an MC@NLO [151] type matching and generates its own PS and particle decays.

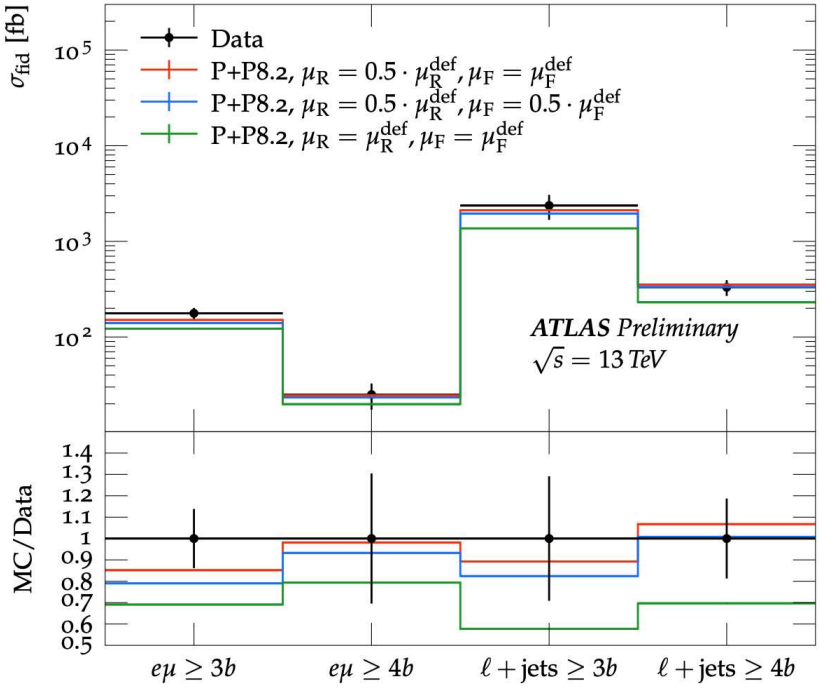


Figure 105 – Fiducial cross-section of $t\bar{t}$+$b\bar{b}$ as measured by POWHEG+PYTHIA8 (P+P8.2) predictions with different scale settings compared to the ATLAS measurement [152].

5.3.3 Other background modelling

Table 17 lists the simulated samples used for the other background process involved in the analysis.

The $t\bar{t}W$ and $t\bar{t}Z$ samples are generated using the MADGRAPH5_aMC@NLO 2.3.3 generator at NLO in QCD with $\mu_R = \mu_F = 0.5\ \Sigma_i m_T(i)$. $t\bar{t}W$ and $t\bar{t}Z$ samples will be collectively referred as $t\bar{t}V$ in the following studies.

The tH events, corresponding to the tHjb and tWH production, are treated as background. These are generated using the MADGRAPH5_aMC@NLO 2.6.2 generator at NLO in QCD with the function form of μ_R and μ_F similar to $t\bar{t}V$.

The QCD V + jets processes (i.e. W + jets and Z + jets) were simulated with the SHERPA 2.2.1 generator. The Comix [145] and OpenLoops libraries are used to calculate NLO-accurate MEs for up to two partons and LO accurate MEs for up to four partons. They are matched with the SHERPA PS using the MEPS@NLO [145, 146] prescription and the NNPDF3.0NNLO PDFs set.

The diboson event simulation is performed with the SHERPA 2.2.1 and 2.2.2 generators. Multiple MEs are matched with Sherpa PS based on Catani–Seymour dipole factorisation [153] using the MEPS@NLO prescription and the NNPDF3.0NNLO PDFs set.

The Single-top in *s*-, *t*-channel, and *tW* production are modelled using the POWHEG BOX v2 at NLO precision in QCD. The NNPDF3.0NNLO nf4 PDF set [142] in 4FS is used to generate t-channel samples, with $\mu_R = \mu_F = m_T(b)$. The s-channel and tW samples are generated in 5 FS with the NNPDF3.0NLO PDF set, and $\mu_R = \mu_F = m_{top}$. To handle interference with $t\bar{t}$ production, the diagram removal (DR) [143] scheme is used for tW production. To assess an uncertainty in the modelling of this interference, an additional sample using the diagram subtraction (DS) [143, 144] scheme is used.

The four top quarks samples are generated using the MADGRAPH5_aMC@NLO 2.3.3 generator at NLO in QCD with the NNPDF3.1NLO PDF set, and $\mu_R = \mu_F = 0.25 \cdot \Sigma_i m_T(i)$.

The tZq samples are generated using the MADGRAPH5_aMC@NLO 2.3.3 generator in the 4FS at LO in QCD with the CTEQ6L1 [154] PDF set, and $\mu_R = \mu_F = 4 \cdot m_T(b)$, where the b-quark is the one coming from the gluon splitting.

The tWZ sample samples are generated using the MADGRAPH5_aMC@NLO 2.3.3 generator at NLO in QCD with the NNPDF3.0NLO PDF set, and $\mu_R = \mu_F = m_{top}$. The DR scheme is used to handle the interference between tWZ and $t\bar{t}Z$.

Events that involve a jet or a photon that is incorrectly identified as a lepton, as well as events that involve a non-prompt lepton, are found to contribute a negligible amount to the single-lepton channel. Except the $t\bar{t}V$ background, all other non-$t\bar{t}$ + jets background processes will be referred as the "Other" background in the following studies.

5.4 Analysis regions

The events passing the event selection, described in Section 5.2, are further categorised into mutually exclusive regions. The sensitivity of the analysis regions is further improved by employing Multivariate Analysis (MVA) techniques, as described in Chapter 4. The MVAs employed in the previous round were based on the BDTs, whereas in the legacy round Deep-sets are used. The analysis regions are defined based on the total number of jets as well as on the number of *b*-jets at different *b*-tagging

WPs: 60% and 70%. The regions with higher signal-to-background (S/B) ratio are classified as signal regions (SRs). The analysis regions which are signal-depleted are called control region (CRs), which provides stringent constraints on the normalisation of the backgrounds and on the systematic uncertainties when used in a fit with SRs. The analysis regions definitions in the single-lepton channel are summarised in Table 18.

The Data/MC comparisons of distributions in the SRs and the CRs in the single-lepton channel, prior to any profile likelihood fit, are studied to identify mis-modelling. In the analysis presented here, there are a total of 10 analysis regions defined, with six SRs and four CRs. The events are first categorised based on the number of jets i.e the ones with at least six jets and the ones with exactly five jets. In each of these regions, events are further sub-classified.

Region		$SR^{\geq 6j}_{\geq 4b}$						$CR^{\geq 6j}_{\geq 4b}$		$CR^{5j}_{\geq 4b}$	
sub-region		STXS 1	STXS 2	STXS 3	STXS 4	STXS 5	STXS 6	$t\bar{t}+\geq 2b$	other	hi	lo
pT [GeV]		[0, 60)	[60, 120)	[120, 200)	[200, 300)	[300, 450)	[450, ∞)	Inclusive			
#leptons		= 1									
#jets		>=6								= 5	
#b-tag	@70%	>=4									
	@60%	-								>=4	<4
Fit input		DNN output								ΔR^{bb}_{avg}	

Table 18 – List of definitions of the analysis regions in single-lepton channel. The last row represents the inputs used in each analysis region for the signal extraction fit. "other" corresponds to $t\bar{t}$+1b/ $t\bar{t}$+B/ $t\bar{t}+\geq c$ region.

The events in the six-jet region are split into six SRs and two CRs based based on the multi-class Deep-sets scores, described in Section 4.4. The STXS multi-classifier Deep-sets scores serve as discriminants for defining the six SRs, which corresponds to the six STXS classes. The two CRs in six-jet region (6jCRs) are defined based on $t\bar{t}+\geq 2b$ and other ($t\bar{t}+1b$/ $t\bar{t}+B$/ $t\bar{t}+\geq 1c$) Deep-sets scores. The Deep-sets outputs are used as the inputs for the SRs and the 6jCRs to the signal extraction fit. The events in the five-jet region are split between the CRs in five-jet. The variable ΔR^{avg}_{bb}, which is the average ΔR for all the possible combinations of b-jet pairs, is used as the inputs for the 5jCRs to the signal extraction fit. The dedicated 6jCRs and 5jCRs help to constrain the different systematic uncertainties associated to sub-background processes. Figure 106 depicts the composition of $t\bar{t}H$ simulated events in each STXS regions. The expected event yield corresponding to the $t\bar{t}H$ components in each STXS regions is shown in Table 19. As expected, the STXS event category is dominated by its corresponding $t\bar{t}H$ STXS component. Figure 107 shows the background contributions in SRs, 6jCRs and 5jCRs, where $t\bar{t}$ + jets background is dominant and

only small contributions come from the Other background process. The largest fraction of $t\bar{t}$ + jets corresponds to the $t\bar{t}+ \geq 2b$ background, which accounts for more than 75% of the total background in all the SRs, followed by $t\bar{t}+1b$,$t\bar{t}+B$ and $t\bar{t}+ \geq c$, which are around 10-20% across all the SRs. The fraction of $t\bar{t}+ \geq 2b$ events is dominant in the dedicated CR($t\bar{t}+ \geq 2b$), while the fraction of $t\bar{t}+1b$,$t\bar{t}+B$ and $t\bar{t}+ \geq c$ background events are higher in the dedicated CR(other), compared to other regions.

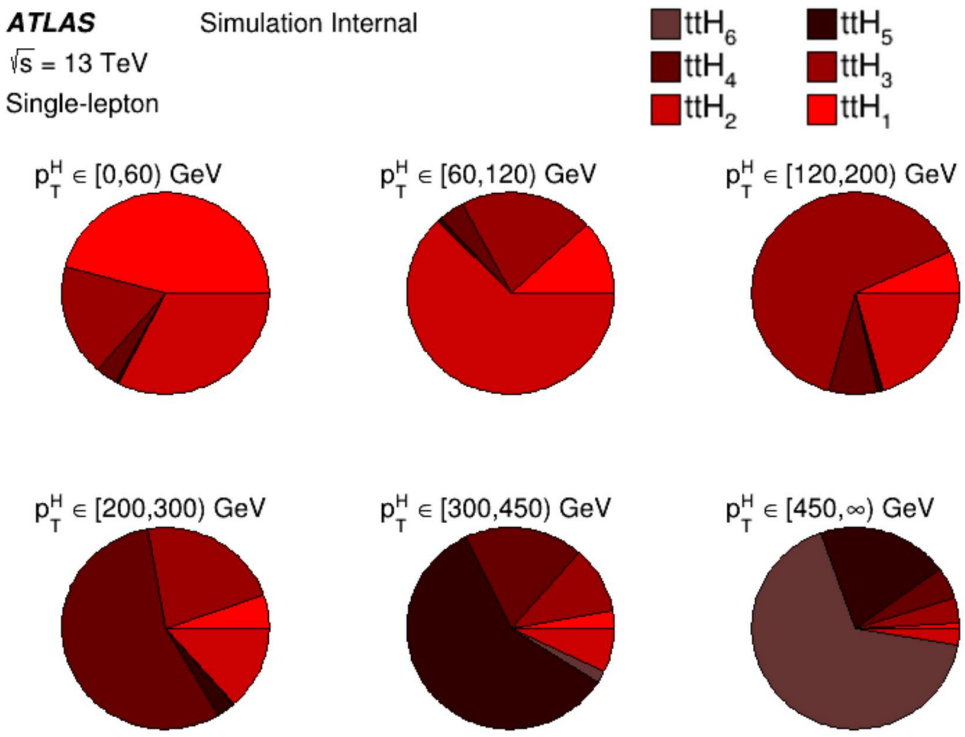


Figure 106 – Pie chart showing the $t\bar{t}H$ signal composition in different STXS regions (SRs) defined based on $t\bar{t}H$ STXS multi-classifier scores.

$t\bar{t}H$ events	$STXS_1$ region	$STXS_2$ region	$STXS_3$ region	$STXS_4$ region	$STXS_5$ region	$STXS_6$ region
	$p_T^H \in [0,60)$ GeV	$p_T^H \in [60,120)$ GeV	$p_T^H \in [120,200)$ GeV	$p_T^H \in [200,300)$ GeV	$p_T^H \in [300,450)$ GeV	$p_T^H \in [450,\infty)$ GeV
$t\bar{t}H_1$	33.13	6.23	4.99	3.66	0.91	0.07
$t\bar{t}H_2$	23.54	32.88	15.04	9.17	2.29	0.22
$t\bar{t}H_3$	13.01	10.92	46.26	15.29	3.42	0.31
$t\bar{t}H_4$	2.47	2.16	5.47	37.97	6.015	0.40
$t\bar{t}H_5$	0.304	0.27	0.60	2.01	18.96	1.59
$t\bar{t}H_6$	0.017	0.02	0.06	0.08	0.63	5.17

Table 19 – Expected pre-fit yields for $t\bar{t}H(i)$ components in all the signal regions in the single-lepton channel, where $i_{[1-6]}$ corresponds to the six STXS regions split based on truth p_T^H.

Figure 108 shows the pre-fit Data/MC distributions over the different event categories including the systematic uncertainties band, discussed in Section 5.5. The expected yields in different analysis

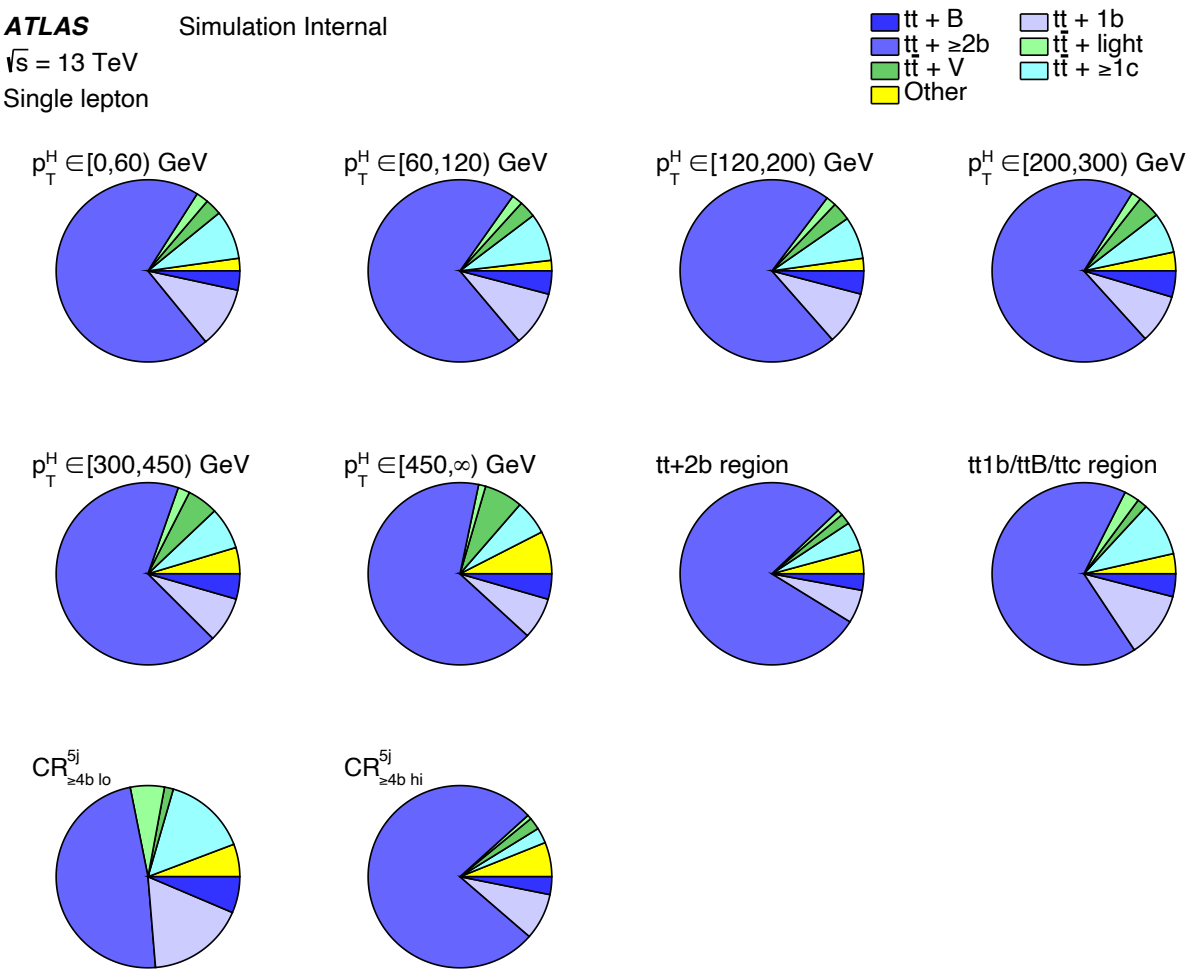


Figure 107 – Pie chart showing all the background composition in the analysis regions.

regions are shown in Table 20. In order to perform the analysis optimisation without looking at the data in regions sensitive to $t\bar{t}H(H \to b\bar{b})$, a blinding threshold is applied based on the S/B ratio. In this analysis the bins are blinded if S/B is larger than 7.7%. Figure 109 shows the pre-fit distributions of the $t\bar{t}H$, $t\bar{t}+ \geq 2b$ and other scores in the inclusive six-jet region (before splitting between STXS SRs and $t\bar{t}+ \geq 2b$ and other CRs). The data/MC agreement is observed to be around $\sim$ 10-20%.

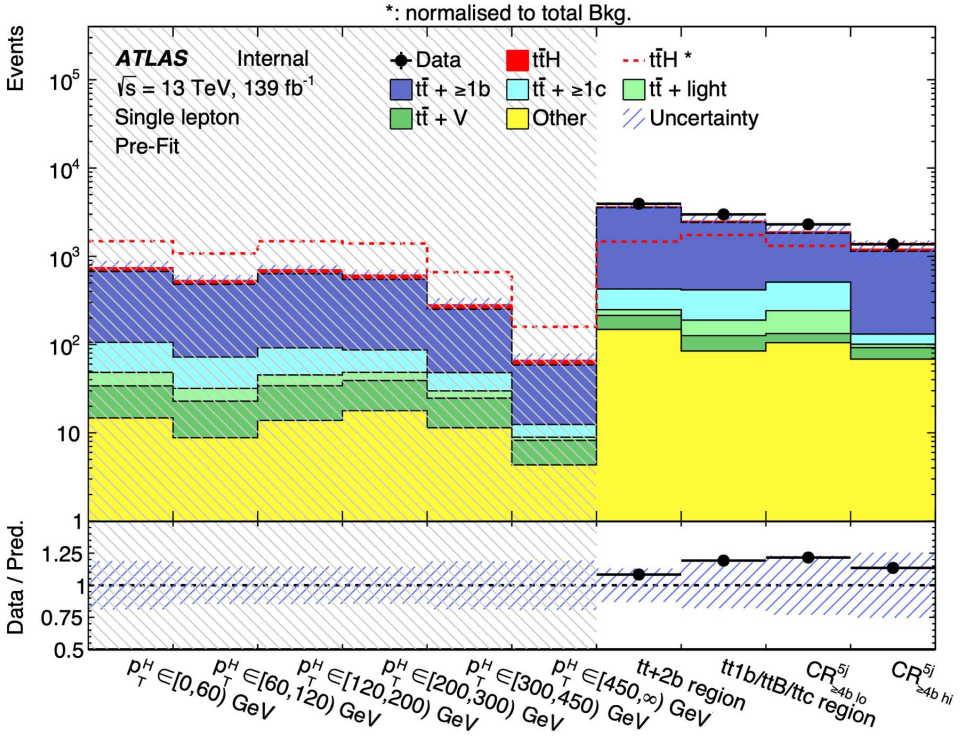
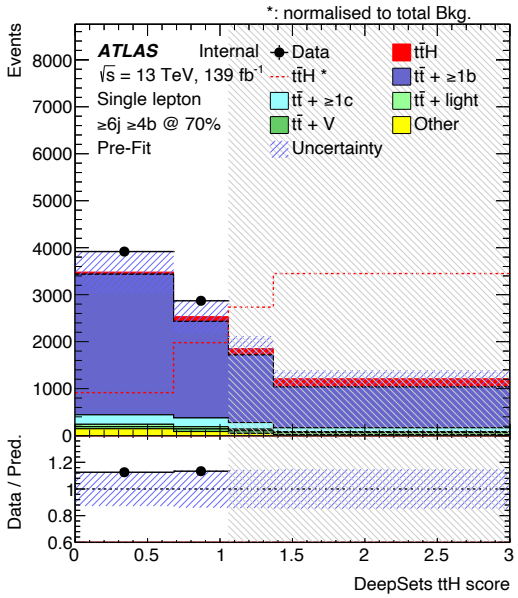


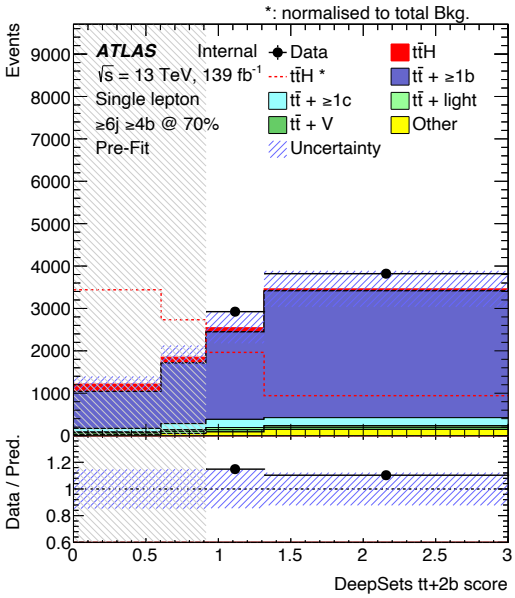
Figure 108 – Pre-fit summary showing the event yields in each of SRs and CRs in the single-lepton channel. The uncertainty band consists the uncertainties described in Section 5.5, except for the normalisation factor k($t\bar{t}+ \geq 1b$)The pre-fit distributions are blinded where S/B > 7.7%.

	$p_T^H \in [0,60)$ GeV	$p_T^H \in [60,120)$ GeV	$p_T^H \in [120,200)$ GeV	$p_T^H \in [200,300)$ GeV	$p_T^H \in [300,450)$ GeV	$p_T^H \in [450,\infty)$ GeV	$t\bar{t}+ \geq 2b$ region	other region	$CR^{5j}_{\geq 4b lo}$	$CR^{5j}_{\geq 4b hi}$
$t\bar{t}H$	72.5 ± 9.8	52.5 ± 6.9	72.4 ± 11.01	68.2 ± 9.5	32.2 ± 4.8	7.8 ± 1.5	71.7 ± 12.9	85.07 ± 9.8	64.1 ± 9.4	67.9 ± 10.8
$t\bar{t}+ \geq 1b$	566.01 ± 122.5	408.1 ± 63.9	540.7 ± 85.9	459.3 ± 73.4	202.3 ± 47.8	46.1 ± 11.6	3133.2 ± 430.9	2001.3 ± 367.1	1320.4 ± 328.4	1005.8 ± 302.9
$t\bar{t}+W$	0.9 ± 0.2	0.7 ± 0.1	1.3 ± 0.5	1.07 ± 0.2	0.7 ± 0.1	0.2 ± 0.07	5.2 ± 0.9	3.1 ± 0.5	1.4 ± 0.3	0.4 ± 0.08
$t\bar{t}+Z$	18.4 ± 2.4	13.3 ± 1.7	18.1 ± 2.5	20.2 ± 3.9	12.5 ± 2.3	3.6 ± 1.2	60.05 ± 8.8	38.7 ± 5.2	26.8 ± 3.3	23.4 ± 3.2
Wt channel	3.4 ± 2.5	2.4 ± 1.3	3.5 ± 0.8	3.7 ± 1.1	2.8 ± 1.1	1.3 ± 0.8	38.3 ± 8.5	22.2 ± 15.2	24.3 ± 16.7	10.05 ± 8.9
tH	0.7 ± 0.1	0.4 ± 0.1	0.6 ± 0.1	1.04 ± 0.2	0.6 ± 0.1	0.1 ± 0.06	1.4 ± 0.2	2.6 ± 0.32	3.5 ± 0.4	3.6 ± 0.4
Other top sources	4.3 ± 0.5	3.1 ± 0.4	4.1 ± 0.5	5.4 ± 0.8	2.9 ± 0.4	0.7 ± 0.2	41.2 ± 2.9	24.3 ± 2.01	34.5 ± 2.6	27.4 ± 1.9
VV & V + jets	6.4 ± 2.6	2.9 ± 1.05	5.5 ± 2.2	7.7 ± 3.2	4.1 ± 1.8	2.2 ± 0.9	67.3 ± 25.9	35.2 ± 12.9	42.6 ± 15.8	27.1 ± 9.5
Total	744.3 ± 143.6	532.6 ± 83.7	704.8 ± 111.4	614.3 ± 89.5	282.5 ± 53.4	66.3 ± 13.9	3630.5 ± 477.8	2502.8 ± 446.4	1893 ± 436.2	1205 ± 306.3

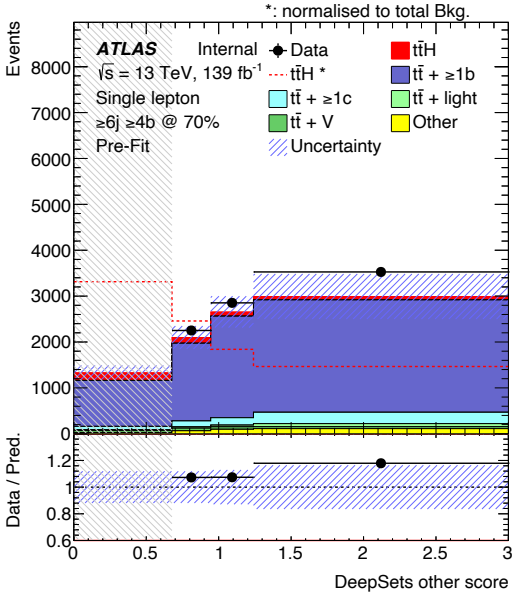
Table 20 – The expected pre-fit yields in all the analysis regions in the single-lepton channel. The errors include systematic uncertainties, discussed in Section 5.5.



(a)



(b)



(c)

Figure 109 – The pre-fit distribution of Deep-sets (a) $t\bar{t}H$ score (b) $t\bar{t}+ \geq 2b$ score and (c) other score in the inclusive six-jet region. The probability scores are not within [0,1] range as the Deep-sets scores are rescaled (cf. Section 4.5). The dashed red line shows the $t\bar{t}H$ signal normalised to the total background prediction. The uncertainty band consists of all the uncertainties except for the normalisation factor k($t\bar{t}+ \geq 1b$). The pre-fit distributions are blinded where S/B > 7.7%.

In the SRs, two possible discriminants could be used as the inputs to the signal extraction fit.

- **Inclusive** $t\bar{t}H$ **scores**: These are the Deep-sets score obtained by adding the individual STXS scores per event. This gives the probability for an event to be $t\bar{t}H$ signal.
- **Individual STXS scores**: These scores corresponds to the assigned probabilities per event for being in each STXS regions.

Figure 110, shows the pre-fit distribution of $t\bar{t}H$ scores in different STXS regions. S/B is observed to be 10-12% in STXS regions except the first bin, which is the only one unblinded. Similarly, Figure 111, shows the individual STXS scores in different STXS regions before the fit. All the STXS bins are blinded due to high S/B ratios. Both of these DNN outputs have the potential to be a strong discriminant; therefore, each is evaluated separately in the statistical fit to determine which one gives better performance. Thus, the choice of the discriminant to be used as inputs in the SRs is done via fit results, presented in Section 5.7.2. Figure 112 and Figure 113 shows the pre-fit distributions of 6jCRs and 5jCRS, respectively. The data/MC agreement was analysed and is found to be uniform and within systematic uncertainties.

In the previous round of the $t\bar{t}H(H \to b\bar{b})$ analysis, a total of eight analysis regions, with six SRs[4] and two CRs were defined in the single-lepton channel. Six STXS categories are defined in the six-jet region, similar to Deep-sets, but the categories are defined based on the reconstructed p_T^H computed via recoBDT score, as described in Section 4.2, rather than Deep-sets score. The dedicated BDT discriminants for defining CRs to constrain $t\bar{t}+ \geq 2b$ and other backgrounds are not available and thus, is one of the main difference in the event categorisation between Deep-sets and the BDTs. In addition, 5jCRs are defined similar to the Deep-sets. The classBDT outputs distributions are used in the SRs as the input to the fit, whereas ΔR_{bb}^{avg} distribution is used in the CRs. Figure 114 shows the S/B and $S/\sqrt{B}$ ratios in the STXS regions defined using the Deep-sets. The S/B ratios in STXS regions defined using recoBDTs cannot be directly compared with Deep-sets, since it does not exclude the contribution from the background. In order to have a fair comparison of $t\bar{t}H$ signal-enriched regions, a classBDT > 0.38 cut is applied. The cut gives similar signal efficiency to the fraction of events in Deep-sets in the STXS regions altogether. Figure 115, shows the S/B and $S/\sqrt{B}$ ratios after the classBDT cut in the STXS regions. The S/B in Deep-sets is around 10-13 % across the SXTS regions, comparable to the BDT regions.

[4] The previous round used five SRs corresponding to five STXS regions, but first SR is split into two in the studies shown in this thesis.

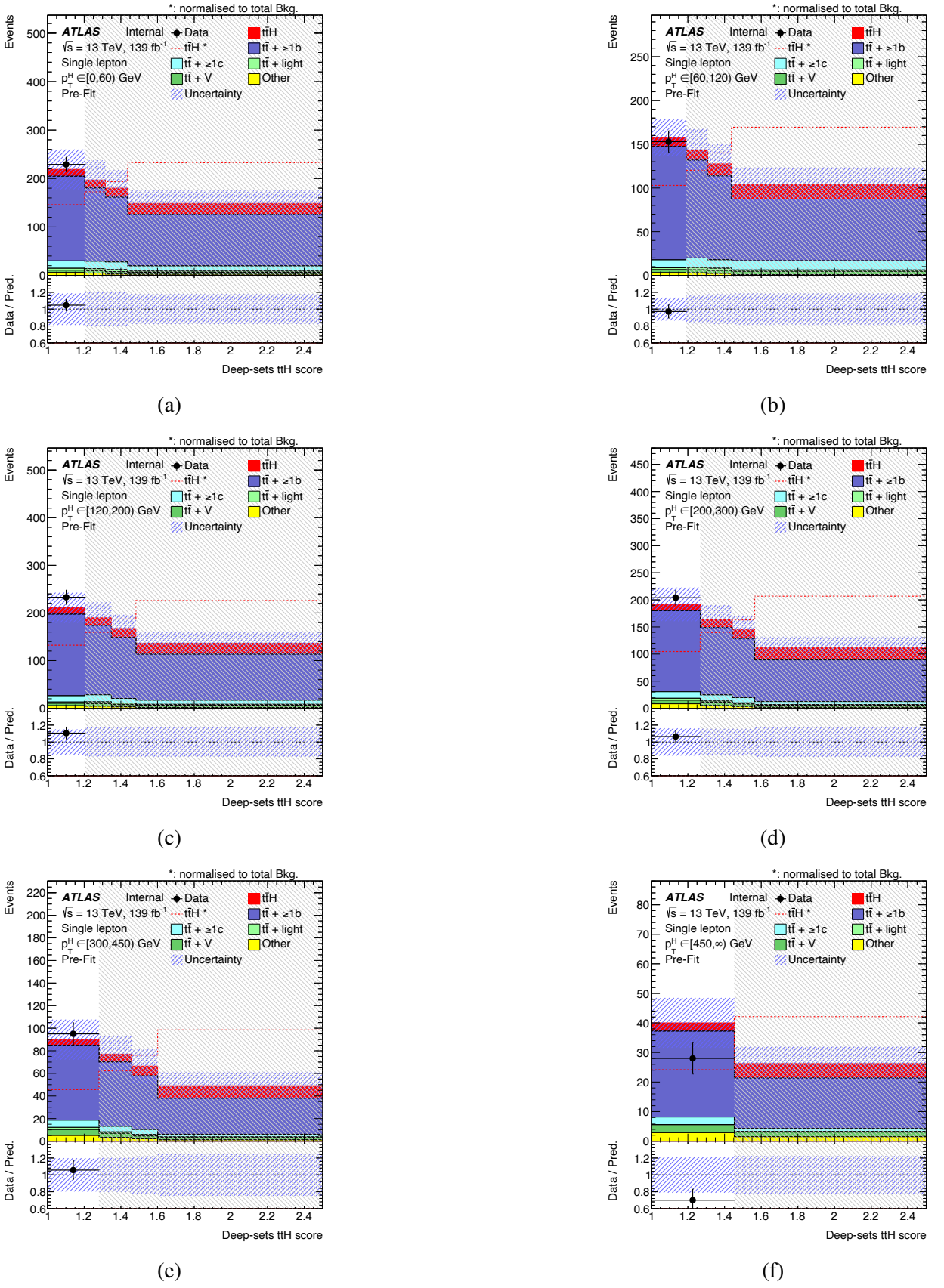
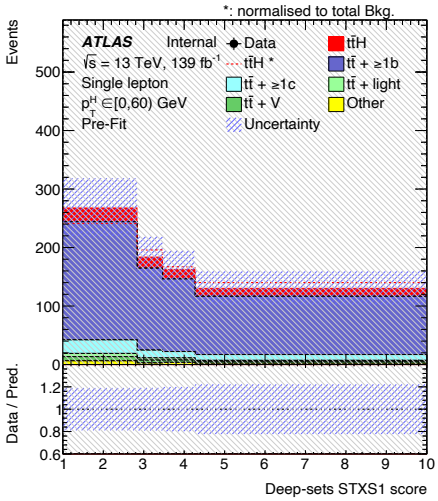
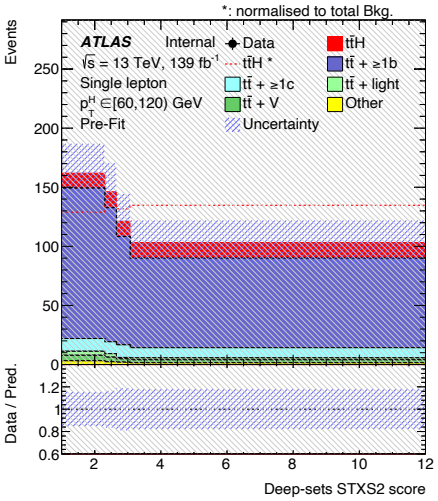


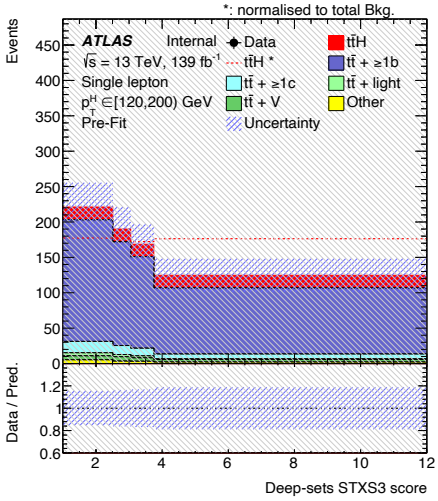
Figure 110 – Pre-fit distribution of the inclusive ttH DNN scores in the (a) $p_T^H \in [0,60)$ GeV, (b) $p_T^H \in [60,120)$ GeV, (c) $p_T^H \in [120,200)$ GeV, (d) $p_T^H \in [200,300)$ GeV, (e) $p_T^H \in [300,450)$ GeV, (f) $p_T^H \in [450,\infty)$ GeV regions. The uncertainty band consists of all the uncertainties except for the normalisation factor k($t\bar{t}+ \geq 1b$). The pre-fit distributions are blinded where S/B > 7.7%.



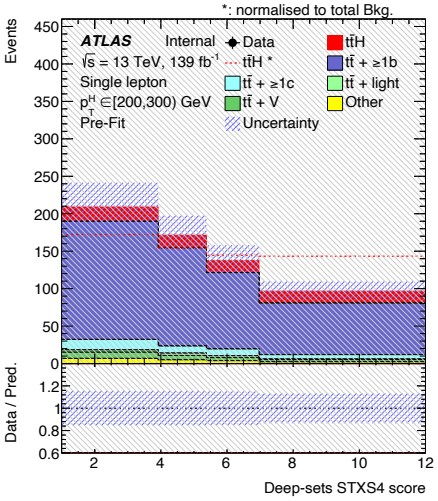
(a)



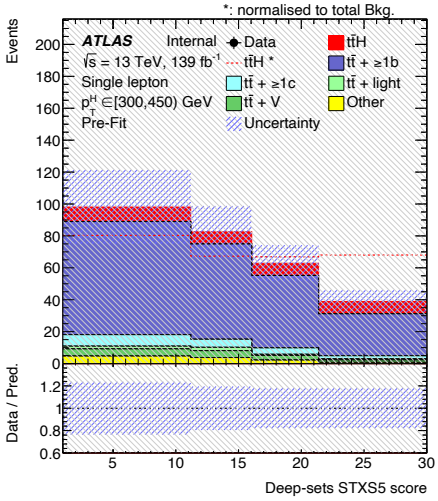
(b)



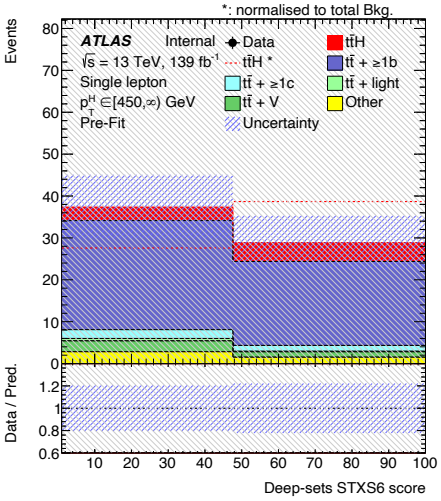
(c)



(d)

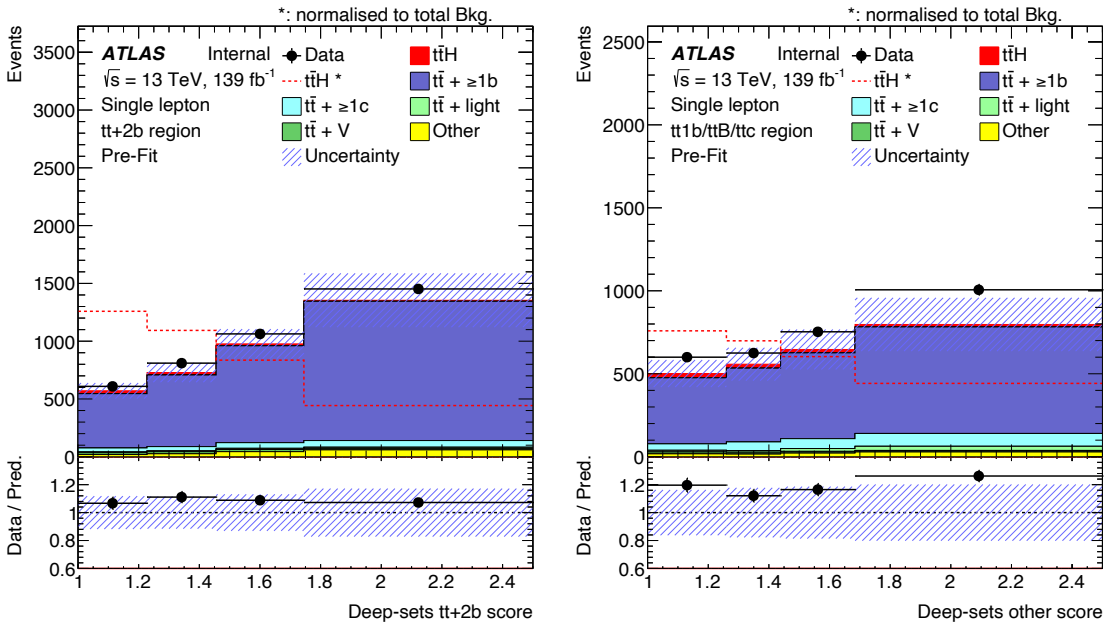


(e)



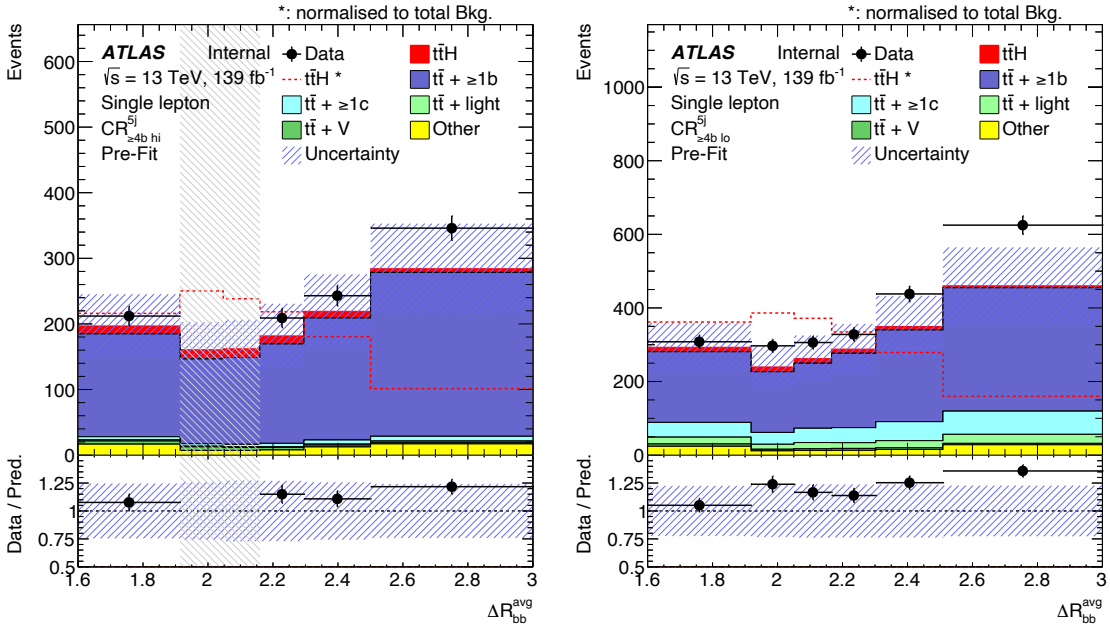
(f)

Figure 111 – Pre-fit distribution of the STXS DNN scores in the (a) $p_T^H \in [0,60)$ GeV, (b) $p_T^H \in [60,120)$ GeV, (c) $p_T^H \in [120,200)$ GeV, (d) $p_T^H \in [200,300)$ GeV, (e) $p_T^H \in [300,450)$ GeV, (f) $p_T^H \in [450,\infty)$ GeV regions. The uncertainty band consists of all the uncertainties except for the normalisation factor k($t\bar{t}+ \geq 1b$). The pre-fit distributions are blinded where S/B > 7.7%.



(a) (b)

Figure 112 – Pre-fit distribution of the (a) $t\bar{t}+\geq 2b$ DNN score in the $t\bar{t}+\geq 2b$ CR, (b) $t\bar{t}$+1b / $t\bar{t}$+B / $t\bar{t}+\geq c$ DNN score in the other CR. The uncertainty band consists of all the uncertainties except for the normalisation factor k($t\bar{t}+\geq 1b$). The pre-fit distributions are blinded where S/B > 7.7%.



(a) (b)

Figure 113 – Pre-fit ΔR_{bb}^{avg} distribution in (a) $CR^{5j}_{\geq 4b\ hi}$ and (b) $CR^{5j}_{\geq 4b\ lo}$. The uncertainty band consists of all the uncertainties except for the normalisation factor k($t\bar{t}+\geq 1b$). The pre-fit distributions are blinded where S/B > 7.7%.

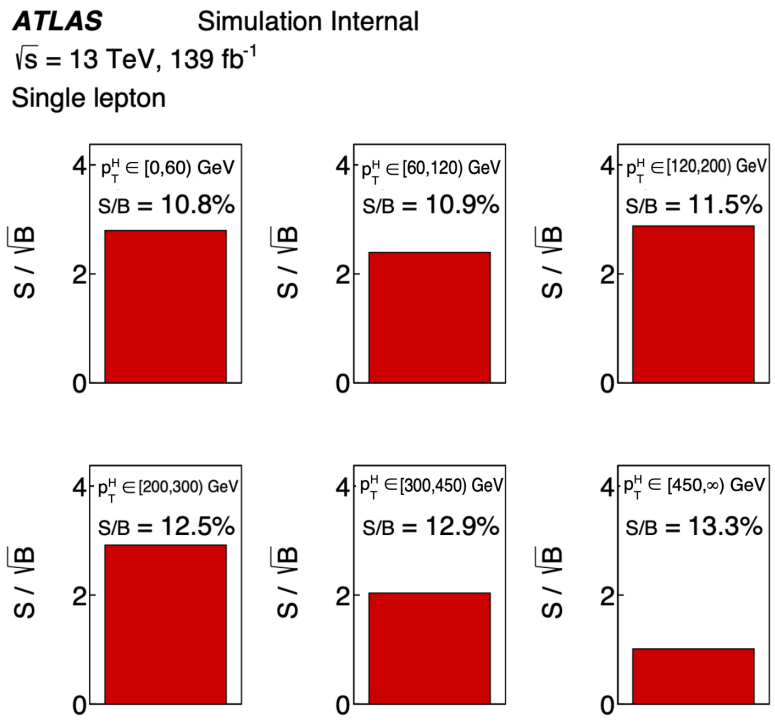


Figure 114 – The $S/\sqrt{B}$ ratio distribution along with the S/B ratio values for different STXS regions defined using Deep-sets scores.

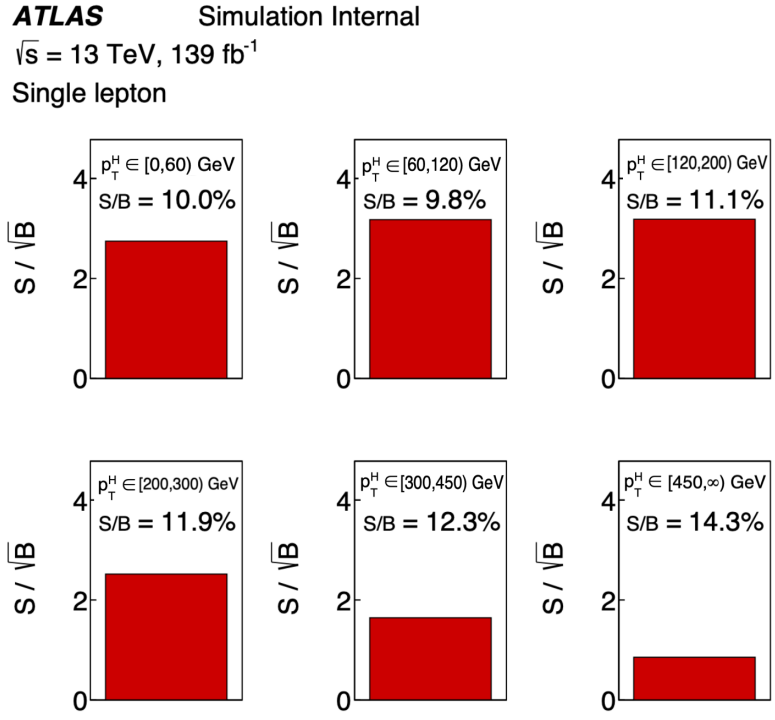


Figure 115 – The $S/\sqrt{B}$ ratio distribution along with the S/B ratio values for different STXS regions defined using recoBDTs.

5.5 Systematic uncertainties

The systematic uncertainties, which come from a variety of sources that are discussed in this section, have a significant impact on the $t\bar{t}H(H \to b\bar{b})$ analysis. These can not only affect the overall normalisation (N) of the samples but also can affect both the shape of the various distribution in the analysis regions and the normalisation (SN). The uncertainties affecting the normalisation change the relative fraction of the different samples, resulting in a shape uncertainty in the distribution of the final discriminant for the total prediction in the different analysis categories. The systematic model used in this analysis is adapted from the one used in the previous round [33]. Table 21 shows a list of all systematic uncertainties considered in the analysis, and indicates the number of independent components and also whether they affect only the normalisation or both the shapes and the normalisation. However, a few of the experimental uncertainties, discussed later in this section, are not taken into account in the analysis results.

The systematic uncertainties can be categorized into two main parts. The first one corresponds to the experimental uncertainties, coming from the calibration and reconstruction of physics objects. With the exception of uncertainty in the luminosity, the experimental uncertainties affect both the shape and normalisation of all MC samples. The experimental uncertainties are discussed in Section 5.5.1. The second part includes uncertainties from the modelling of the signal and backgrounds processes, described in Section 5.5.2 and Section 5.5.3, respectively. Few of the systematic uncertainties are also broken down into multiple independent sources, where each individual source has an effect that is correlated across all of the analysis categories.

5.5.1 Experimental uncertainties

The experimental uncertainties are described here, which are correlated across all of the analysis regions. The uncertainties associated to *b*-tagging and jets, have a significant impact in the profile-likelihood fit, discussed in Section 5.6.

Luminosity and pile-up modelling: The uncertainty associated with the integrated luminosity for the combined 2015-2108 Run-2 data is 1.7% [34]. The uncertainty on the modelling of the pile-up is also considered, to account for the difference between the predicted and measured values for the inelastic cross-section [155].

Jets and *b*-tagging: The experimental uncertainties are dominated by the uncertainties associated with jets. Even if the single components have relative uncertainties in the range of 1-5 %, the huge number of jets in the final state magnifies their effect. Combining the information gained from the test-beam data, the LHC data, and simulation allows to extract the jet energy scale (JES) uncertainty [156]. The uncertainty on the jet energy resolution (JER) is derived by combining

Systematic uncertainties	Type	Components
Experimental uncertainties		
Luminosity	N	1
Pileup modelling	SN	1
Physics Objects		
Electrons	SN	7
Muons	SN	15
Jet energy scale	SN	31
Jet energy resolution	SN	9
E_T^{miss}	SN	1
Jet vertex tagger	SN	3
b-tagging		
Efficiency	SN	45
Mis-tag rate (c)	SN	20
Mis-tag rate (light)	SN	20
Signal modelling		
$t\bar{t}H$ cross-section	N	2
$t\bar{t}H$ modelling	SN	4
Higgs branching fractions	N	3
Background modelling		
tt background		
$t\bar{t}$ cross-section	N	1
$t\bar{t}+ \geq 1c$ normalisation	N	1
$t\bar{t}+ \geq 1b$ normalisation (free-floating)	N	1
$t\bar{t}$+light normalisation	SN	4
$t\bar{t}+ \geq 1c$ modelling	SN	4
$t\bar{t}+ \geq 1b$ modelling	SN	6
Other background		
$t\bar{t}W$ cross-section	N	2
$t\bar{t}Z$ cross-section	N	2
$t\bar{t}W$ modelling	SN	1
$t\bar{t}Z$ modelling	SN	1
Single-top cross-section	N	3
Single-top modelling	SN	7
W+jets normalisation	N	3
Z+jets normalisation	N	3
Diboson normalisation	N	1
four-tops cross-section	N	1
tZq, tWZ, tHjb, and tWH cross-sections	N	3

Table 21 – Summary of all the systematic uncertainties used in the analysis. "N" corresponds to the uncertainty is affected by the normalisation-only. "SN" denotes that both the normalisation and the shapes are considered. For a more accurate treatment, some of the systematic uncertainties are divided into several components! [33].

information from simulation and di-jet p_T balance measurements [156]. There are in total 31 and 9 independent sources of uncertainties considered on JES and JER, respectively. An uncertainty associated to the jet vertex tagger (JVT) calibration is also considered to account for the differences between data and simulations [157]. Additional uncertainty sources are also taken into account such as related to the jet flavour, pile-up corrections and jet kinematics as well as the full and fast simulation differences. The *b*-tagging calibration, which accounts for the correction of *b*-tagging efficiencies in simulated samples to match those in data, is used to extract the associated uncertainties as a function of the different *b*-tagging WPs and the jet p_T. A total of 85 uncorrelated uncertainties for *b*-, *c*- and light-jets are taken into account. Depending on the WPs, the uncertainty on the *b*-tagging jets calibration is around 2-10 %. The mis-tag calibration uncertainty for *c*- and light-jets is 10-25 % and 15-50 %.

Leptons: The uncertainty on the leptons mainly comes from the the trigger, reconstruction, identification and isolation efficiencies as well as the lepton momentum scale and resolution for the muons and the electrons. These independent sources have small effect on the final fit.

Missing Transverse Momentum: The estimation of the missing transverse momentum is subject to all of the uncertainties that are associated with the energy scales or resolution of the reconstructed objects. In this analysis, three additional uncertainties associated with the scale and resolution of the soft term are considered. These uncertainties have a relatively minor effect in the fit, given that the missing transverse momentum is only used as input of the Deep-sets discriminants used for the event categorisation and the distributions used in the profile-likelihood fit and not for the selection of the event.

Only the luminosity, pileup and *b*-tagging experimental uncertainties are taken into account in the analysis results. The other experimental uncertainties described above required dedicated sample processing to estimate their impact on the sample acceptance, and was still ongoing at the time of the writing the thesis. However, the estimated impact of on the uncertainty on the final results based on the previous round results are discussed in Section 5.7.2.

5.5.2 Signal modelling uncertainties

To estimate the signal cross-section uncertainty, an uncertainty of $\pm$ 3.6 % is applied by varying the PDF and α_s in the fixed-order calculation [158–160]. The PDF variations have a negligible effect on the shapes of the distributions used in the analysis. Uncertainties in the Higgs boson branching fractions are also taken into account. The uncertainty coming from the branching fraction of the Higgs boson into $b\bar{b}$ amounts to 2.2 % [158]. The ISR uncertainty is estimated by simultaneously varying μ_R and μ_F up (down) by a factor of 0.5 (2) in the ME and α_s^{ISR} in the PS, while an

uncertainty due to FSR is estimated by varying α_s^{FSR} in the PS by a factor of 0.5 (2) [161]. Event weights are used to assess the impact of these systematic uncertainties in the analysis. *Two-point systematics* are used in order to estimate some modelling uncertainties. This method retrieves the uncertainty, by comparing two different MC generator setups and deriving an uncertainty from the differences between them. The impact of the PS and hadronisation model is evaluated by comparing the nominal generator setup with a sample produced with POWHEG+HERWIG 7 in 5FS. To assess the uncertainty due to the choice of the matching scheme, the nominal sample is compared to a generator setup of MADGRAPGH5_aMC@NLO+PYTHIA8. Additional uncertainty is applied, estimated due to missing higher order terms in the perturbative QCD calculations on the total cross-section. This amounts to an scale uncertainty of 9.2% for the total cross-section retrieved using the Stewart-Tackmann procedure [162]. All the signal modelling uncertainties are correlated across all the STXS bins.

5.5.3 Background modelling uncertainties

$t\bar{t}$+jets backgrounds:

The background modelling uncertainties of $t\bar{t}$+jets used in this analysis are listed in Table 22. The different $t\bar{t}$+jets components i.e $t\bar{t}+\geq 1b$, $t\bar{t}+\geq 1c$ and $t\bar{t}$+light are impacted by different types of uncertainties. The $t\bar{t}$+light processes benefits from relatively precise measurements in data. $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ processes are fairly sensitive to the precision of the ME and the flavour scheme used for the PDF. Thus, separate uncertainties are assigned for the $t\bar{t}+\geq 1b$, $t\bar{t}+\geq 1c$ and $t\bar{t}$+light processes. Across these three components, all systematic uncertainty sources are treated as uncorrelated. The impact of the PDF uncertainties is found to be negligible.

Uncertainty source	Description		Components
$t\bar{t}$ cross-section	$\pm$ 6%		$t\bar{t}$+light
$t\bar{t}+\geq 1b$ normalisation	free-floating		$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1c$ normalisation	$\pm$100%		$t\bar{t}+\geq 1c$
ISR	Varying μ_R^{ISR} (PS), μ_R and μ_F (ME)	in POWHEGBox+PYTHIA8 ttbb (4FS)	$t\bar{t}+\geq 1b$
		in POWHEGBox+PYTHIA8 $t\bar{t}$ (5FS)	$t\bar{t}+\geq 1c$, $t\bar{t}$+light
FSR	Varying μ_R^{FSR} (PS)	in POWHEGBox+PYTHIA8 ttbb (4FS)	$t\bar{t}+\geq 1b$
		in POWHEGBox+PYTHIA8 $t\bar{t}$ (5FS)	$t\bar{t}+\geq 1c$, $t\bar{t}$+light
NLO matching	SHERPA 2.210 $t\bar{t}$+$b\bar{b}$ (4FS)	vs. POWHEGBox+PYTHIA8 ttbb (4FS)	$t\bar{t}+\geq 1b$
	SHERPA 2.210 $t\bar{t}$ (5FS)	vs. POWHEGBox+PYTHIA8 tt (5FS)	$t\bar{t}+\geq 1c$, $t\bar{t}$+light
PS and hadronisation	POWHEGBox+HERWIG7 $t\bar{t}$+$b\bar{b}$ (4FS)	vs. POWHEGBox+PYTHIA8 $t\bar{t}$+$b\bar{b}$ (4FS)	$t\bar{t}+\geq 1b$
	POWHEGBox+HERWIG7 tt (5FS)	vs. POWHEGBox+PYTHIA8 $t\bar{t}$ (5FS)	$t\bar{t}+\geq 1c$, $t\bar{t}$+light
$t\bar{t}+\geq 1b$ fractions	Variation of the relative fractions of $t\bar{t}+\geq 2b$ and $t\bar{t}$+1b/1B		$t\bar{t}+\geq 1b$
p_T^{bb} shape	Shape mismodelling measured from data		$t\bar{t}+\geq 1b$

Table 22 – List of the sources of systematic uncertainty for $t\bar{t}$+jets modelling. The last column of the table shows the $t\bar{t}$ + jets components to which a systematic uncertainty is assigned. Across the three components, all systematic uncertainty sources are treated as uncorrelated.

When calculating the inclusive $t\bar{t}$ cross-section (NNLO+NNLL), an uncertainty of ± 6 % is only applied to the $t\bar{t}$+light samples, because it predominates in the inclusive phase space. This uncertainty accounts for various effects from varying the normalisation and the factorisation scales, α_s, the PDFs and the top-quark mass. An uncertainty of 100 % is applied to the normalisation of the $t\bar{t}+ \geq 1c$ sample. As mentioned before, the $t\bar{t}+ \geq 1b$ normalisation is kept free-floating in the fit.

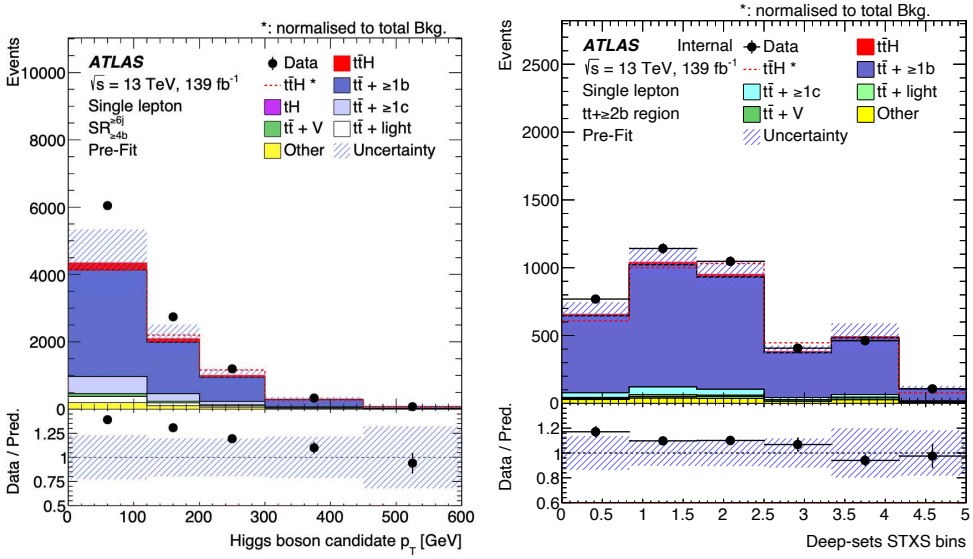
The ISR and FSR uncertainties for $t\bar{t}$+jets components are estimated using the same procedure as for $t\bar{t}H$. The variations for both the systematic uncertainties are performed on the respective nominal $t\bar{t}$+jets components samples. Two-point systematics are also used in order to estimate $t\bar{t}+ \geq 1b$, $t\bar{t}+ \geq 1c$ and $t\bar{t}$+light modelling uncertainties. Table 22 shows the list of the nominal and the alternative samples for each of the $t\bar{t}$+jets component, used for estimating the impact of the PS and hadronisation and the choice of matching scheme.

The fraction of $t\bar{t}+ \geq 1b$ events in the selected phase-space (here $\geq$ 5j $\geq$4b@70%) in all the alternative samples is re-weighted to match the fraction in the nominal sample. This is to allow the normalisation of $t\bar{t}+ \geq 1b$ to be driven solely by the free-floating parameter in the signal extraction fit to data. $t\bar{t}+ \geq 1b$ NLO matching uncertainty is decorrelated between SRs and CRs. To account for variations in the $t\bar{t}+ \geq 1b$ sub-component fractions found in different predictions, an additional NP is introduced to cover the largest discrepancy between two models for the fraction of $t\bar{t}+1b$ and $t\bar{t}+ \geq 2b$. The 1σ variation of this NP corresponds to reducing the amount of $t\bar{t}+ \geq 2b$ by 3% and increasing the amount of $t\bar{t}+1b$ by 11%. This uncertainty is correlated across all regions, and impacts each region differently due to the varying compositions of $t\bar{t}+ \geq 1b$.

The p_T^{bb} shape uncertainty was introduced in the previous round of the analysis to account for the data/MC mis-modelling of $t\bar{t}+ \geq 1b$ events in the pre-fit distributions of the reconstructed p_T^H in the signal-lepton channel. Figure 116 (a) shows a clear slope in data/MC ratio, where in the first bin the ratio is more than 25 %. Figure 116 (b) shows a similar distribution in the legacy round, where slight slope is observed but less prominent than the previous round. Since there is no STXS bin prediction for CRs, an indirect method is used to study p_T^{bb} modelling in CRs. Similarly to the SRs, the STXS scores are used for the STXS bin assignment. This uncertainty in the previous round was derived from the data-MC difference observed in the six-jet SR, whereas the uncertainty in this round is derived from the dedicated $t\bar{t}+ \geq 2b$ CR, is orthogonal to the SR, which is expected to reduce the impact of this systematic in the analysis results.

Other backgrounds:

Table 23 lists the sources of the systematic uncertainties Other background modelling processes. These uncertainties have much lower impact in the final fit, compared to the $t\bar{t}$+jets uncertainties.



(a) (b)

Figure 116 – Pre-fit distribution of the reconstructed Higgs boson p_T in the SR ($\geq$ 6j $\geq$ 4b@70%) in the single-lepton channel for the (a) the previous [33] and the (b) legacy round of $t\bar{t}H(H \to b\bar{b})$ analysis.

	Uncertainty source	Description
$t\bar{t}V$ [163]	cross-section	±15%
	NLO matching	MADGRAPH5_aMC@NLO +PYTHIA8 vs SHERPA
	PS and hadronisation	MADGRAPH5_aMC@NLO +PYTHIA8 vs SHERPA
Single-top [164–166]	NLO cross-section	± 5%
	NLO matching	MADGRAPH5_aMC@NLO +PYTHIA8 vs POWHEGBox+PYTHIA8
	PS and hadronisation	POWHEGBox+HERWIG7 vs POWHEGBox+PYTHIA8
	tW and tt interference [167]	POWHEGBox+PYTHIA8 (Diagram removal (DR) vs Diagram subtraction (DS) scheme)
W+jets	cross-section	±40%
	normalisation	± 30% (events with = 2 and > 2 Heavy flavour jets are uncorrelated)
Z+jets	normalisation	±35%
Diboson [168]	cross-section and additional jet production	±50%
tZq [169]	cross-section (PDFs)	±0.9%
	factorisation and normalisation	± 7.9%
tWZ [169]	cross-section	±50%
tHjb [25]	cross-section (PDFs)	±3.7%
	factorisation and normalisation	+6.5 / -14.9%
tWH [25]	cross-section (PDFs)	±6.3%
	factorisation and normalisation	+6.5 / -6.7%
four-tops	normalisation	±50%

Table 23 – List of the sources of systematic uncertainty for Other background modelling processes [33].

5.6 Profile-likelihood fit

The statistical test of signal presence is based on Neyman-Pearson lemma [170], which states that to reject a hypothesis H_0 in favour of hypothesis H_1 the most powerful test is the ratio of their likelihoods. The general form of a likelihood function can be expresses as $\mathcal{L}(\theta) = \prod_{i=1}^{n} f(x_i, \Theta)$, where $x_1 ... x_n$ are the measurement of random variable x repeated n times and $f(x_i, \Theta)$ is the probability density function of x known a priori except Θ which corresponds to a set of parameters with unknown values. In this analysis, the set of signal strengths μ, which is ratio of the expected or observed signal cross section to the SM cross section, the background normalisation factor k, and the set of Nuisance Parameters (NPs) θ, are the unknown parameters. The set of NPs gives extra degrees of freedom to the fit and characterises the effects of systematic uncertainties in the signal and background expectations. A binned data distribution with number of events N_i in bin i, is compared via the fit procedure to its expectation value of the number of events:

$$N_i^{exp}(\mu, k, \theta) = \sum_{\mu} \mu_s \cdot sig_{s,i}(\theta) + \sum_{k} k_b \cdot bkg_{b,i}(\theta) \tag{5.1}$$

where, $sig_{s,i}$ are the expected $t\bar{t}H$ events of component s and $bkg_{b,i}$ are the expected background events of component b in bin i. The parameter of interest is defined as:

$$\mu_s = \frac{\sigma^s}{\sigma^s_{SM}} \tag{5.2}$$

where, μ_s is one element in the set of signal strength parameters μ, σ^s is the cross-section of the signal and σ^s_{SM} is the expected SM cross-section. In this analysis, only one signal-strength parameter is used for an inclusive cross-section measurement, whereas six signal-strength parameters are used for the STXS measurement. k_b acts only on the $t\bar{t}+ \geq 1b$ background and is kept freely floating. On the other hand, all of the other background processes are normalised to their projected cross-sections, and k_b is set to one. A central value of $\theta = 0$ is used, which corresponds to the best estimate of a given parameter associated with a systematic uncertainty, and a $\pm$ shift reflects the distribution variation resulting from its 1σ uncertainty. The continuous NPs are defined by extrapolation for $|\theta| > 1$ and interpolation for $|\theta| < 1$. NPs are implemented in the likelihood as Poissonian or Gaussian priors, given by:

$$\mathcal{L}(\mu, k, \theta) = \prod_i \frac{N_i^{exp}(\mu, k, \theta)^{N_i}}{N_i!} e^{-N_i^{exp}(\mu,k,\theta)} \prod_i^M \frac{1}{\sqrt{2\pi}} e^{\frac{\theta_\mu^2}{2}} \tag{5.3}$$

where M is the total number of NPs. The best estimates of the parameters in the fit: μ, k, θ are achieved by maximizing the likelihood function. It is more convenient to minimize the negative log-likelihood, as it is numerically more stable. The test statistics λ_μ is defined as:

$$\lambda_\mu = -2\ln\left(\frac{\mathcal{L}(\mu, \hat{\hat{k}}_\mu, \hat{\hat{\theta}}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{k}, \hat{\theta})}\right) \tag{5.4}$$

where the parameters denoted with a single-hat are the values that maximise the likelihood function, whereas the parameters denoted with a double-hat correspond to the NPs that maximise the likelihood function for a given μ value [171]. The compatibility of the observed data with the background-only hypothesis ($\mu = 0$) is measured using the test statistic λ_0. The uncertainty associated to the best-fit μ value is determined by finding the values of μ that correspond to changing λ_μ by one unit. The statistical tool for further implementing the statistical inferences on μ is provided by RooFit package [172]. One of them is the significance $S = \sqrt{\lambda_0}$, which represents the significance of a deviation from the background-only hypothesis.

In the $t\bar{t}H(H \to b\bar{b})$ analysis, the distribution of the DNN discriminants in the SRs and 6jCRs, and the ΔR_{bb}^{avg} distribution in 5jCRs, are combined in a profile-likelihood fit to test for the presence of the $t\bar{t}H$ signal. One NP is assigned for each of the bins in the analysis regions, to account for the limited statistics of the simulated samples. It incorporates the predicted yields and the uncertainties in each of the bins of the analysis regions to fit them to the data. The binning of these distributions is optimised in such a way as to maximise the sensitivity. This is done by simultaneously maintaining the total MC statistical uncertainty in each bin at a level that is adjusted to prevent biases caused by the fluctuations in the predicted yields.

A simultaneous fit to the data is performed in the single-lepton channel under the background-only hypothesis, where the normalisation factor for $t\bar{t}+ \geq 1b$ background is kept freely floating. Since the legacy analysis still needs to go through ATLAS unblinding approval at the time of writing, thus, full (S+B) fits to the data will not be presented in this thesis. In addition to the fit to the data, the Asimov data-set can also be used instead of the real data in all the regions. The Asimov data-sets are binned data-sets with the event count in each bin set to the expected event yield for the selected model parameters. This is mainly used as a statistical test to obtain the expected uncertainties on the signal strength and the background normalisation as well as the expected significance. The profile-likelihood fit studies are performed for both inclusive and STXS cross-section measurements on the Asimov data-set under the (S+B) hypothesis.

5.7 Analysis results

This section presents the results of the analysis performed in the single-lepton channel using the profile-likelihood fit, described in Section 5.6, where all 10 analysis regions were used in the signal extraction fit. Both the inclusive and the STXS cross-section measurements are performed where similar event categorization is used. For the STXS measurements, the fit is performed with multiple signal strengths corresponding to the six $\hat{p}_T^H$ bins.

The Asimov fit (statistics-only) which is performed using different choices of DNN discriminants in the SRs is discussed in Section 5.7.1 followed by the expected performance of the fit under the signal-plus-background (S+B) assumption, including the systematic uncertainties, discussed in Section 5.7.2. The background-only fit is also performed, presented in Section 5.7.3. The expected performance using the newly developed DNN discriminants are compared with the previously used BDT discriminants, using the same configuration for signal extraction fit in the single-lepton channel.

5.7.1 Statistics-only Asimov fit

The fit is first carried out using the Asimov data set, with only the statistical uncertainties being taken into consideration. The signal strength and normalisation factors for the backgrounds measured with that configuration are 1 as expected, and their corresponding statistical uncertainties can be extracted from the fit. For the results presented here, two independent normalisation factors are considered corresponding to the dedicated CR regions for $t\bar{t}+ \geq 2b$ and other ($t\bar{t}$+1b, $t\bar{t}$+B and $t\bar{t}+ \geq c$) background sub-components. These are denoted as k($t\bar{t}+ \geq 2b$) and k(other) in the fit and are kept freely-floating. These have been replaced by an inclusive normalisation factor covering the inclusive $t\bar{t}+ \geq 1b$ background component in the final fit configuration including the systematic uncertainties used to produce the results presented in this thesis. However, the availability of an exclusive CR region focused on different background sub-components helps to constrain the uncertainties specifically associated with those by reducing their correlation.

As mentioned before, two choices of discriminant in the SRs are available, i.e., Deep-sets $t\bar{t}H$ scores and Deep-sets STXS scores. In order to determine which discriminant offers the better fit performance, the statistics-only fit is carried out independently for both of these discriminants. Figure 117 and Figure 118 show the expected statistical uncertainties on the signal strengths fitted simultaneously in all the p_T^H bins as well as on the inclusive signal strength, using the STXS and the $t\bar{t}H$ scores as the discriminants in the SR, respectively. Figure 119 shows the expected statistical uncertainties on both inclusive and STXS fitted signal strengths using classBDT as the discriminant in the STXS categories built based on the RecoBDT, as described in Section 4.2. Figure 121 shows

the normalisation factor for the $t\bar{t}+\geq 2b$ and other backgrounds. When using the STXS scores as the SR discriminants, the expected uncertainties on the fitted signal strengths in each p_T^H bin are the lowest when compared to both the $t\bar{t}H$ score and the classBDT. The expected statistical uncertainties on the inclusive signal strength ($\mu_{t\bar{t}H}$) is comparable among all the three discriminants, however the $t\bar{t}H$ score performs slightly better.

First, when comparing $t\bar{t}H$ and STXS discriminant performance, STXS discriminant is slightly better than $t\bar{t}H$, showing a reduction in the expected statistical uncertainty between 1% and 10% across the STXS bins. This is consistent with the reduced correlation between the STXS bins, as shown in Figure 120. In both cases, the uncertainty on $k(t\bar{t}+\geq 2b)$ and k(other) is consistent. As a result of the better performance for the STXS measurement, the Deep-sets STXS scores are considered as the default choice of discriminants (further referred to as Deep-sets) in the SR for the signal extraction fit and will be used as such in the studies shown from now on.

Second, when comparing the performance of Deep-sets and classBDT discriminants, the Deep-sets discriminant outperforms the classBDT, showing a reduction in the expected statistical uncertainty between 5% and 30% across the STXS bins. The improved performance is much more pronounced in the higher p_T bins. Again this is consistent with the significantly reduced correlation between the STXS bins, as shown in Figure 120.

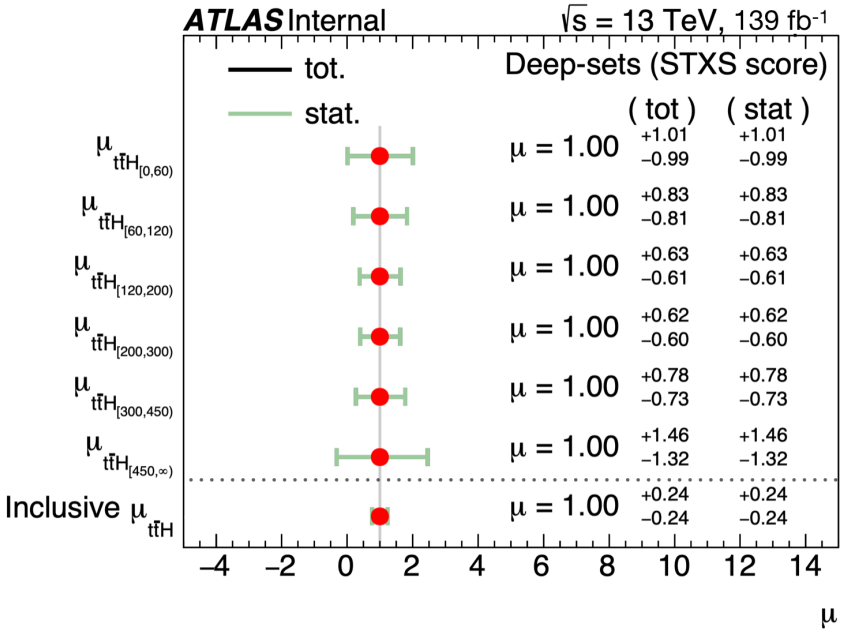


Figure 117 – The inclusive and STXS signal-strength measurements, including the expected statistical uncertainty, using the STXS Deep-sets scores as the signal region discriminant in the Asimov (S+B) fit in the single-lepton channel.

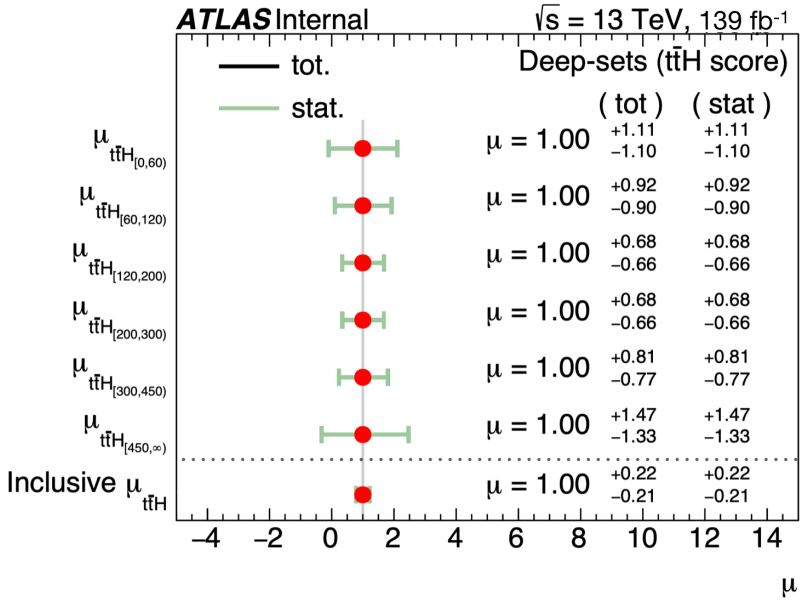


Figure 118 – The inclusive and STXS signal-strength measurements, including the expected statistical uncertainty, using the $t\bar{t}H$ Deep-sets scores as the signal region discriminant in the Asimov (S+B) fit in the single-lepton channel.

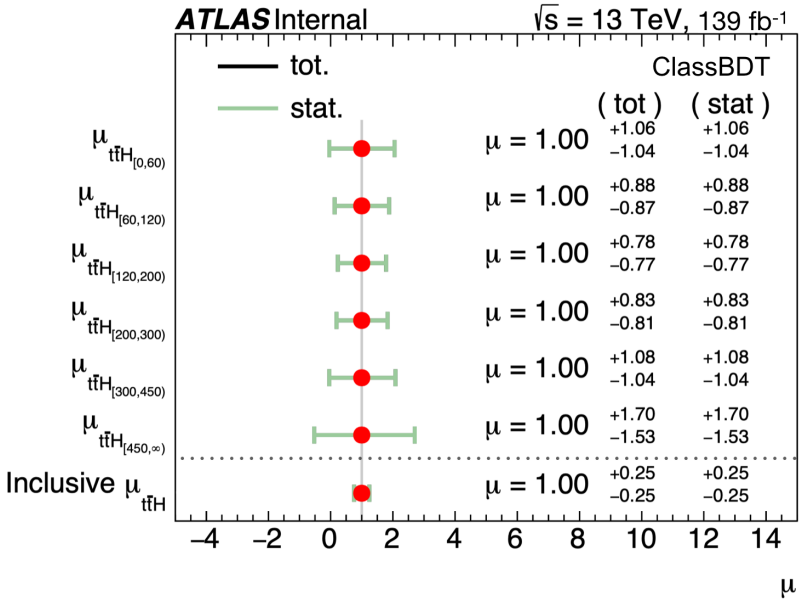


Figure 119 – The inclusive and STXS signal-strength measurements, including the expected statistical uncertainty, using the classBDT outputs as the signal region discriminant in the Asimov (S+B) fit in the single-lepton channel.

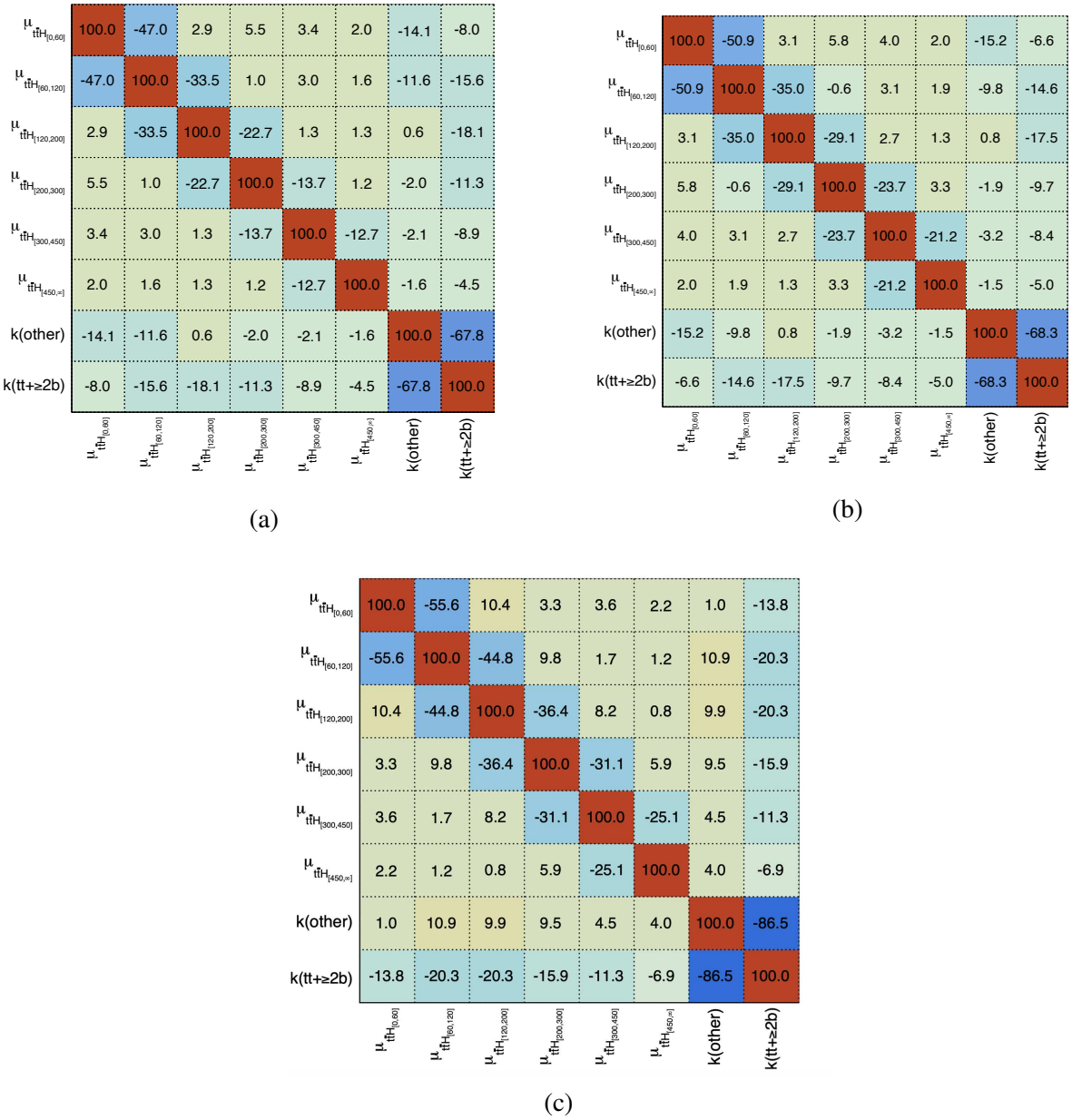


Figure 120 – The correlation matrix between the nuisance parameters in the Asimov (S+B) fit in the single-lepton channel using (a) the STXS scores (b) the $t\bar{t}H$ scores and (c) the classBDT scores at the signal region discriminant. Each NP has to have at least one correlation above 20% to be included here.

The overall lower correlation between STXS signal strengths and background normalisation factors account for a part of the improvement in STXS uncertainties and background normalisations, shown in Figure 121.

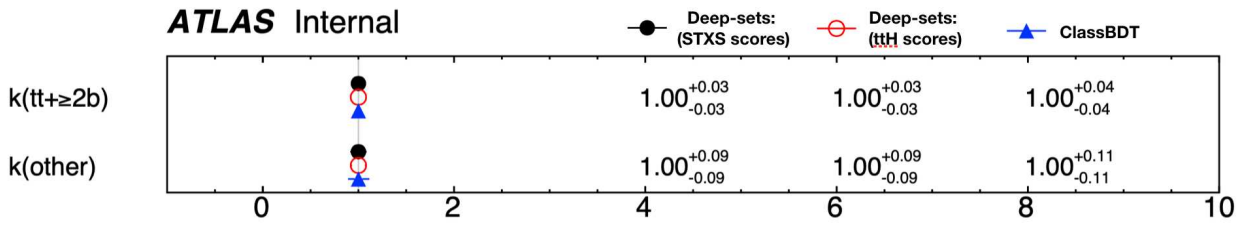


Figure 121 – Fitted normalisation factors for the $t\bar{t}+ \geq 1b$ and other backgrounds, including the expected statistical uncertainty, in the Asimov (S+B) fit in the single-lepton channel.

5.7.2 Expected performance

The expected uncertainties on the signal strengths and normalisation factor, as well as the expected significance, are extracted from the Asimov fit. The systematic uncertainties, discussed in Section 5.5 are also included as NPs in the fit. The studies presented here are performed considering $k(t\bar{t}+ \geq 1b)$ as the only background normalisation factor and kept freely floating. In order to speed up the convergence of the fit, certain optimizations are done on the systematic uncertainties. The shape and normalisation of systematic uncertainties are pruned, separately per region and per sample, with the pruning threshold set to 1%. This means that all the shape and normalisation systematic variations below this threshold will be removed from the fit. It helps to reduce the computation time of the fit and at the same time does not change the NP constraints nor affect the final signal strength. Moreover, a smoothing method is also applied to reduce the impact from the statistical fluctuations in simulated samples.

Figure 122 shows the normalisation factor for the $k(t\bar{t}+ \geq 1b)$ background, where the expected total uncertainty on the normalisation factor for $t\bar{t}+ \geq 1b$ background with Deep-sets is improved by 25% in comparison to the classBDT. The total uncertainties includes the systematic as well as the statistical uncertainty.

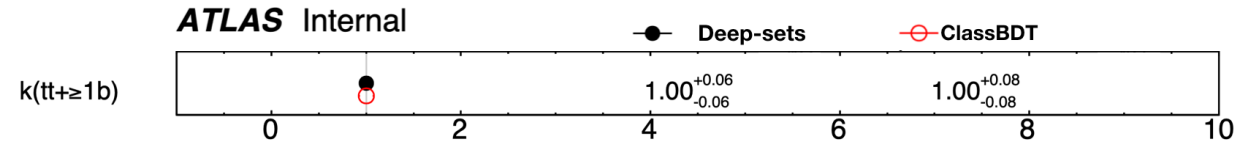


Figure 122 – Fitted normalisation factors for the $t\bar{t}+ \geq 1b$ background, including the total expected uncertainty, in the Asimov (S+B) fit in the single-lepton channel.

Figure 123 and Figure 124 show the inclusive signal strength and the STXS measurements along with the total expected uncertainties, using the Deep-sets and the classBDT, respectively. The value of the systematic uncertainty is derived by subtracting the statistical uncertainty quadratically from the total uncertainty i.e $\sigma_{syst} = \sqrt{\sigma^2_{tot} - \sigma^2_{stat}}$. The expected $\mu_{t\bar{t}H}$ is comparable between the Deep-sets and classBDT. Similar to the statistics-only Asimov fit, the Deep-sets performs better than the classBDT in each of the p_T^H bins. The expected signal strength measurements in the SXTS bins show an improvement between 3% and 35%. This is consistent with the reduced correlation between the dominant NPs, shown in Figure 125 and Figure 126 for the Deep-sets and classBDT in the $\mu_{t\bar{t}H}$ measurement, respectively. The NPs which have dominant impact on fit are discussed later in this section. Only the NPs which have at least one correlation above 20% are included in the correlation matrix. Therefore, the total number of NPs shown in the correlation matrix differs between the two setups.

The expected significance from the S+B fit using an inclusive signal strength using Deep-sets is **2.71 standard deviations**, compared to **2.54 standard deviations** using BDTs.

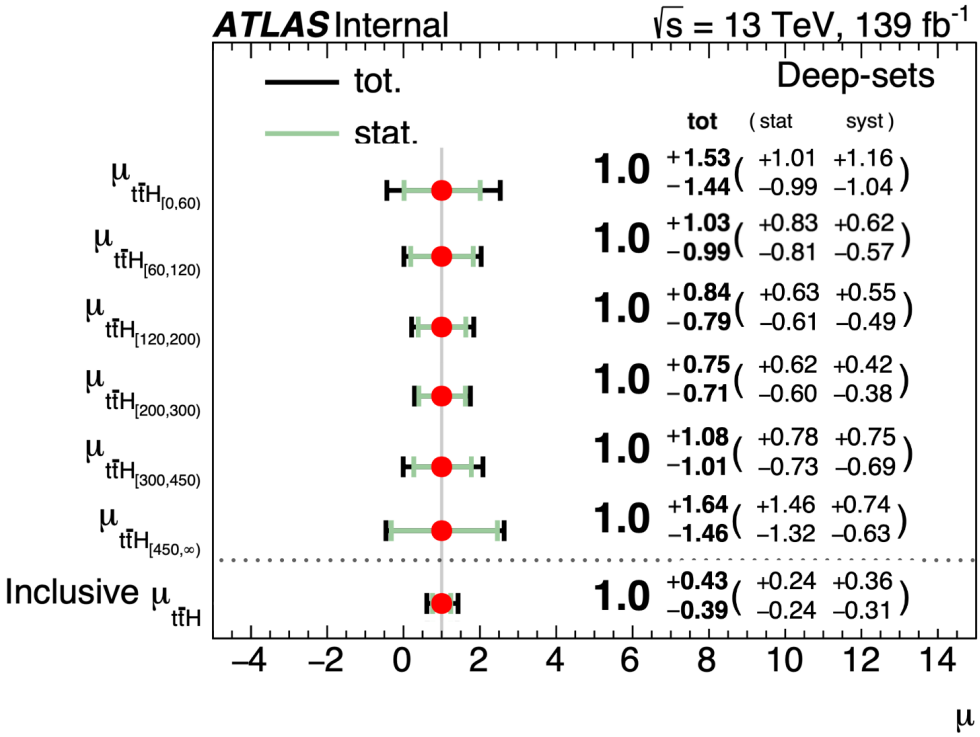


Figure 123 – The inclusive and STXS signal-strength measurements, including the total expected uncertainty, in the Asimov (S+B) fit using the Deep-sets discriminants in the single-lepton channel.

Figure 127 presents the fitted NPs in the S+B Asimov fit in the single-lepton channel. In the Deep-sets, $t\bar{t}+\geq 1b$ ISR, $t\bar{t}+\geq 1b$ PS and hadronisation, $t\bar{t}+\geq 1b$ NLO (CR and SR) show high constraints. In the classBDT, $t\bar{t}+\geq 1b$ NLO have much higher constraint in SR and relatively

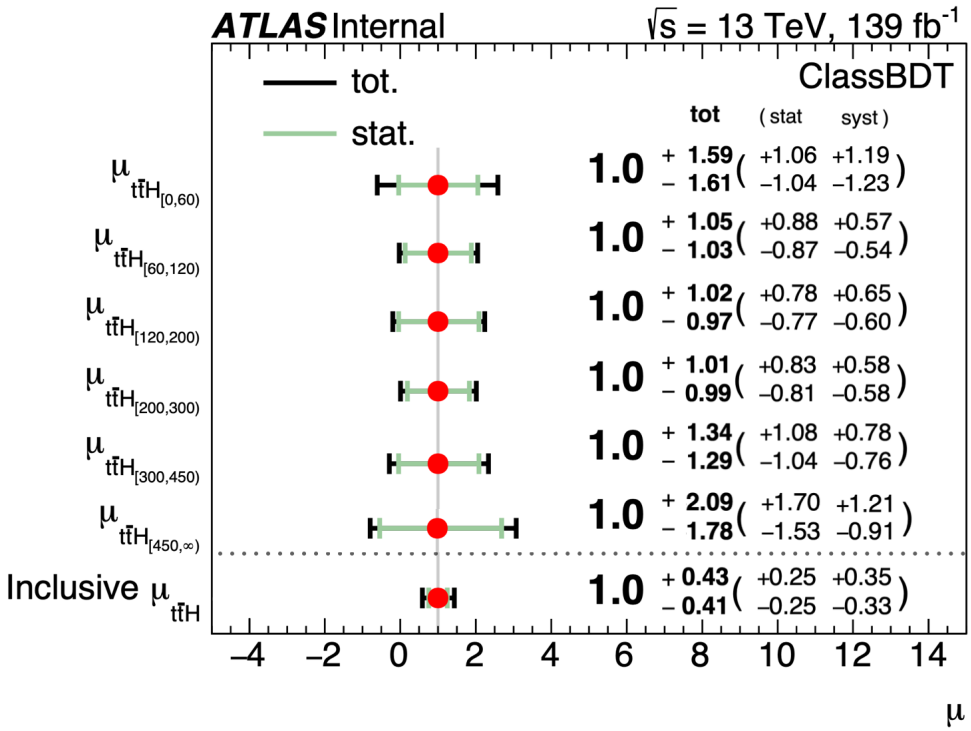


Figure 124 – The inclusive and STXS signal-strength measurements, including the total expected uncertainty, in the Asimov (S+B) fit using the classBDT discriminants in the single-lepton channel.

relaxed constraint in CR compared to the Deep-sets. As mentioned in Section 5.4, the categorization of CRs differs between Deep-sets and classBDT, where Deep-sets have additional 6jCRs, from which the constraints are derived in $t\bar{t}+\geq 1b$ NLO CRs. The constraints on p_T^{bb} shape and $t\bar{t}+\geq 1b$ PS and hadronisation are comparable between the two setups. The $t\bar{t}+\geq 1b$ FSR show lower constraint in Deep-sets while constraint on $t\bar{t}+\geq 1b$ ISR is higher, which might due to the higher correlation between $t\bar{t}+\geq 1b$ ISR and PS and hadronisation in Deep-sets. The remaining fitted NPs show comparable expected constraints on the NPs between the two setups.

Further studies on systematic uncertainties have been performed in this analysis using Deep-sets discriminant. Figure 128 shows the 20 leading NPs whose associated systematic uncertainties contribute the most to the overall uncertainty in the inclusive signal strength measurement. The impact of each NP is denoted by $\Delta\mu$. It is calculated by comparing the nominal best-fit value of μ, to the result of the fit when fixing the considered NP to its best-fit value, $\hat{\theta}$, shifted by its pre-fit (post-fit) $\pm\triangle\theta$ ($\pm\triangle\hat{\theta}$). The pulls on the NP correspond to: $(\hat{\theta}-\theta_0)/\triangle\theta$. The NLO matching affecting the $t\bar{t}+\geq 1b$ background in CRs has the largest impact on the inclusive signal strength. The leading systematic is followed by three uncertainties related also to the modelling of the $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq c$ background components. The p_T^{bb} shape uncertainty is also among the five dominant uncertainties impacting the fit. As previously stated, there are a few experimental uncertainties that

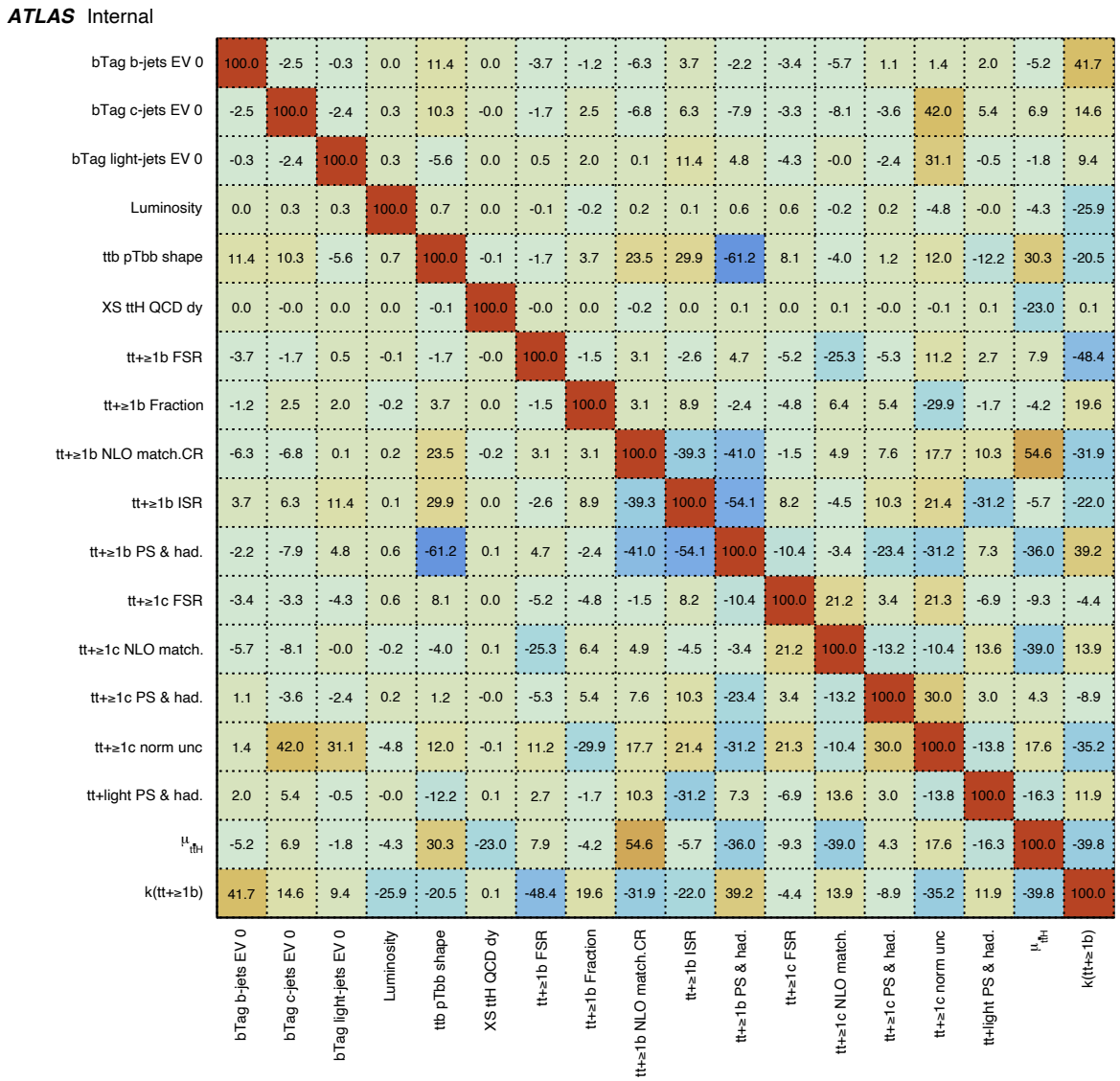


Figure 125 – The correlation matrix between the nuisance parameters for the inclusive signal strength measurement in the Asimov (S+B) fit using the Deep-sets discriminants. Each NP has to have at least one correlation above 20% to be included here. "QCD dy" denotes the scale uncertainty on the total $t\bar{t}H$ cross-section. "EV 0" refers to the first component of the *b*-tagging uncertainty source.

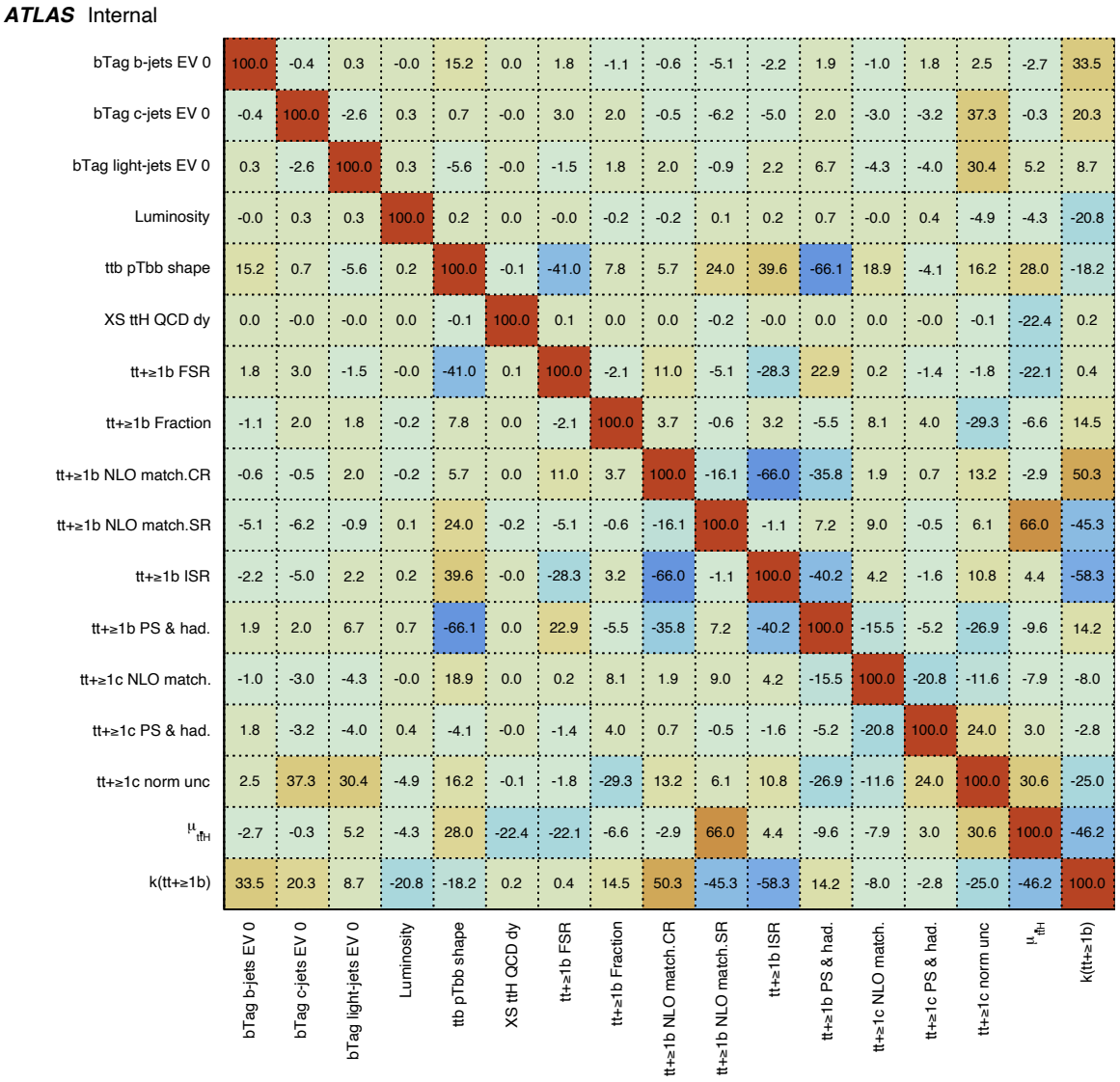


Figure 126 – he correlation matrix between the nuisance parameters for the inclusive signal strength measurement in the Asimov (S+B) fit using the classBDT discriminants. Each NP has to have at least one correlation above 20% to be included here. "QCD dy" denotes the scale uncertainty on the total $t\bar{t}H$ cross-section. "EV 0" refers to the first component of the *b*-tagging uncertainty source.

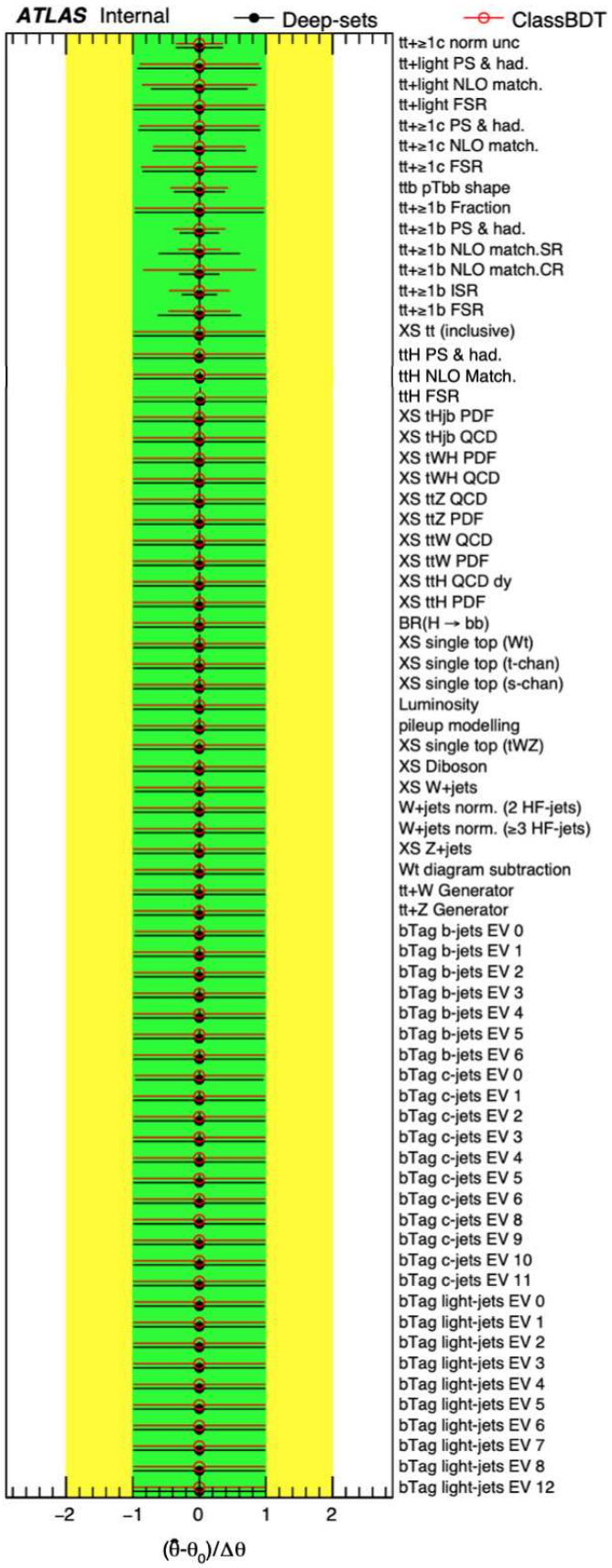


Figure 127 – Fitted nuisance parameters in the S+B fit on the Asimov data-set in the single-lepton channel, performed using Deep-sets (solid black point) and using classBDT (red circles).

are not accounted for in the fit. These include jet energy scale and resolution uncertainties, which had a less significant impact on the fit in the previous round and were not among the top 20 NPs in the ranking plot [33]. Among the experimental uncertainties, the uncertainty originating from the flavour-tagging calibration has the highest impact and is included in the fit.

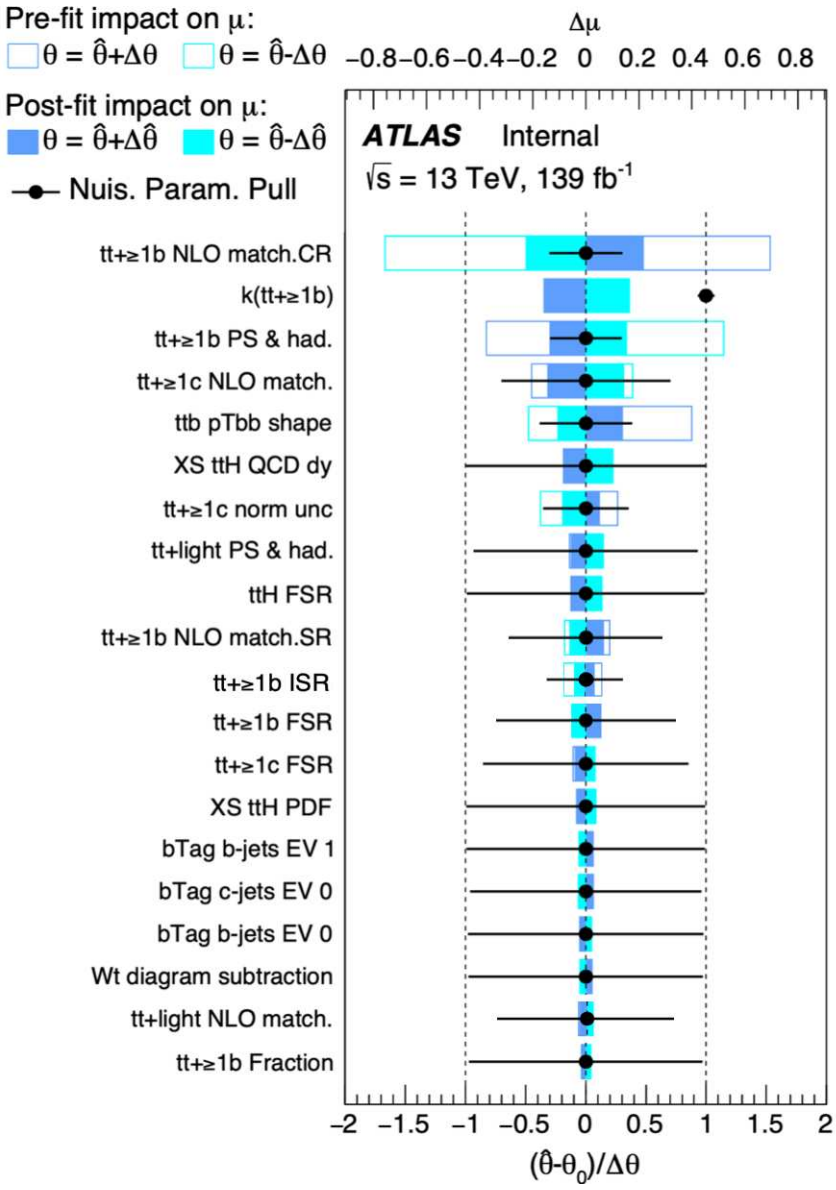


Figure 128 – Ranking of the 20 nuisance parameters with the largest post-fit impact on μ in the fit. The empty blue rectangles correspond to the pre-fit impact on μ and the filled blue ones to the post-fit impact on μ. The black points show the pulls of the nuisance parameters. k($t\bar{t}+\geq 1b$) is the freely floating normalisation factor for which pre-fit impact on μ is not defined and values of θ_0 and $\Delta\theta$ are set to 1.

5.7.3 Fit results on data

The background (B)-only fit to the data is performed using Deep-sets discriminants in the single-lepton channel. This fit provides valuable information for the validation of the background modelling. Since this is a B-only fit, the results are valid for both inclusive and STXS cross-section measurements, as they only provide information about the backgrounds. The normalisation factor for $t\bar{t}+\geq 1b$ background after the fit to data is $k(t\bar{t}+\geq 1b) = 1.12^{+0.06}_{-0.06}$. Figure 129 shows the fitted NPs for the $t\bar{t}+\geq 1b$ background modelling, mainly where the largest impact on the pulls are observed.

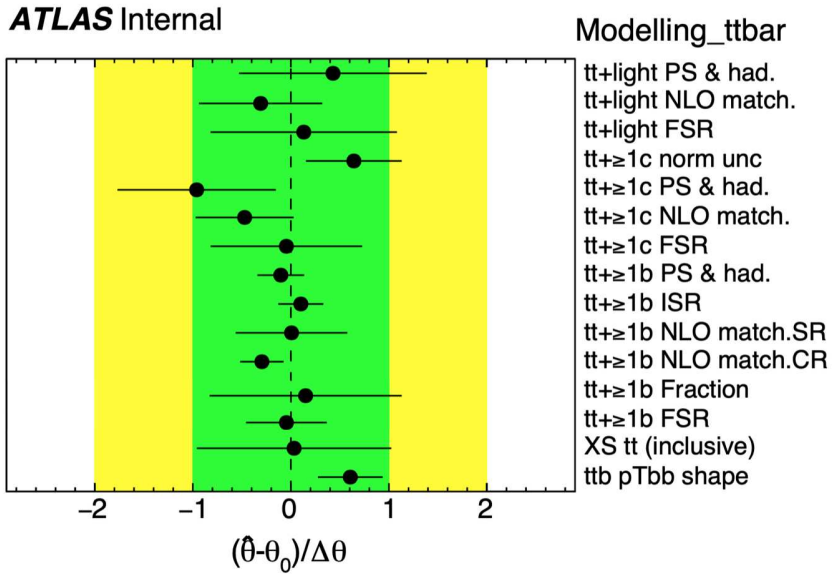


Figure 129 – Fitted $t\bar{t}+\geq 1b$ nuisance parameters in the B-only fit on data in the single-lepton channel.

A large pull is observed on $t\bar{t}+\geq 1c$ PS and hadronisation, which is pulled down by 0.9 σ with respect to the nominal fit. $t\bar{t}$+light PS and hadronisation is pulled up by factor of 0.6. The p_T^{bb} shape uncertainty is also pulled up 0.6 σ, which is expected from the pre-fit mismodelling, shown in Figure 116. The normalisation of the $t\bar{t}+\geq 1c$ background is pulled up by factor of 0.7, which is consistent with the previous round of analysis, which observed 0.6σ pull [33]. A pull of 0.3 σ is observed in $t\bar{t}+\geq 1b$ NLO CR, which ranks highest in terms of the impact on the fit and also showed large constraints in the (S+B) Asimov fit (see Figure 127). The other constrained NPs which come from $t\bar{t}+\geq 1b$ ISR and PS and hadronisation exhibit much smaller pulls. The B-only post fit input variables in 6jCRs and 5jCRs are shown in Figure 130. Overall, there is good post-fit Data/MC agreement across all distributions. The uncertainty is reduced post-fit due to the fit introducing constraints and correlations among NPs. The SR bins are not shown since, the bins are blinded. Instead, the STXS p_T^H shape distribution after the post fit for all the events in six-jet region are

shown in Figure 131. The p_T^{bb} shape uncertainty contribute to correcting the p_T^{bb} shape in six-jet region, resulting in no slope and good post-fit data/MC agreement.

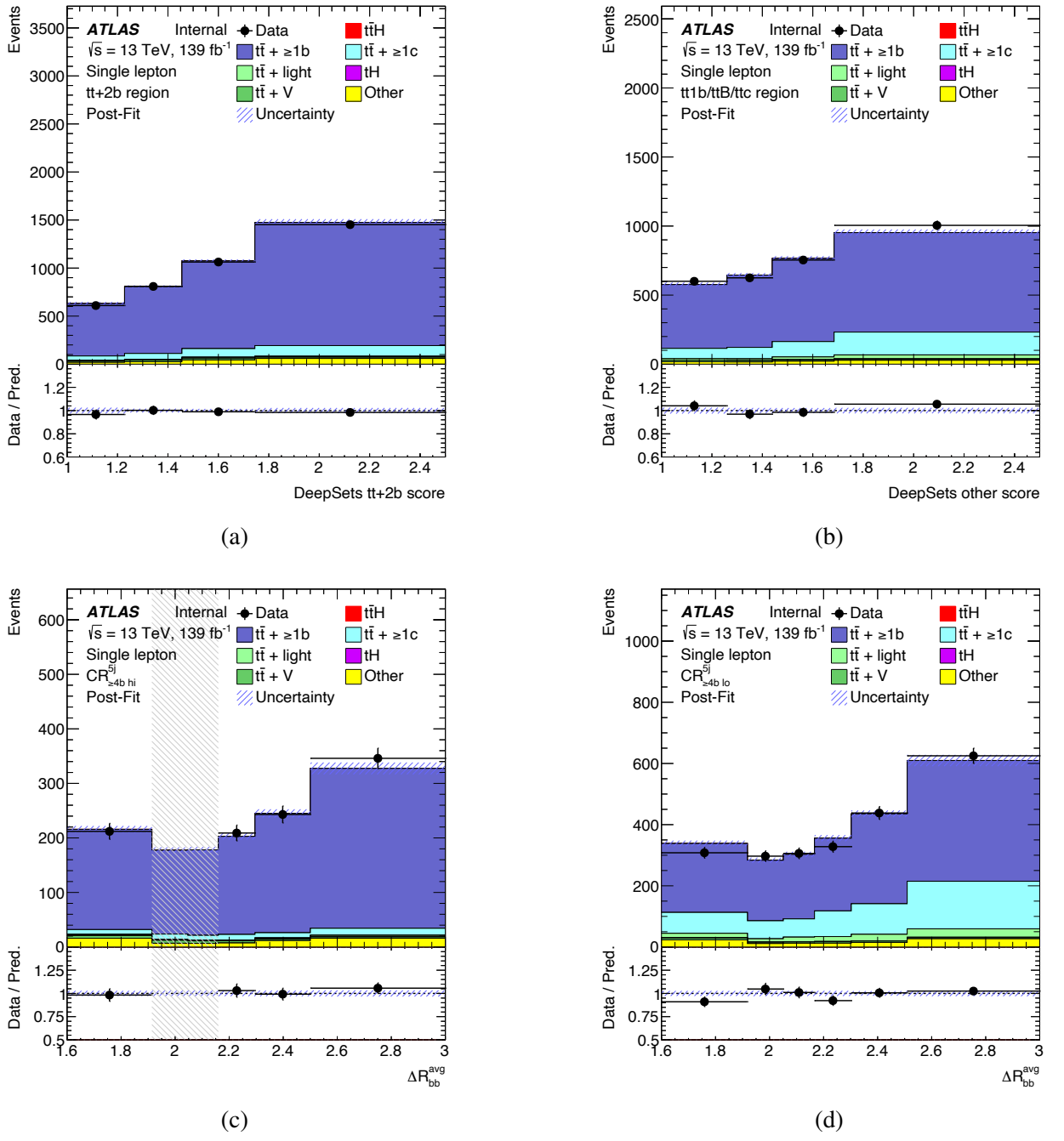


Figure 130 – Post-fit distributions of the Deep-sets determinant in the 6jCRs: (a) ($t\bar{t}+ \geq 2b$) and (b) other regions; and the ΔR_{bb}^{avg} distribution in 5jCRs: (c) CR(hi) and (d) CR(lo) in the single-lepton channel. The bins which were blinded (where $S/B > 7.7\%$) in the corresponding pre-fit distributions are kept blinded

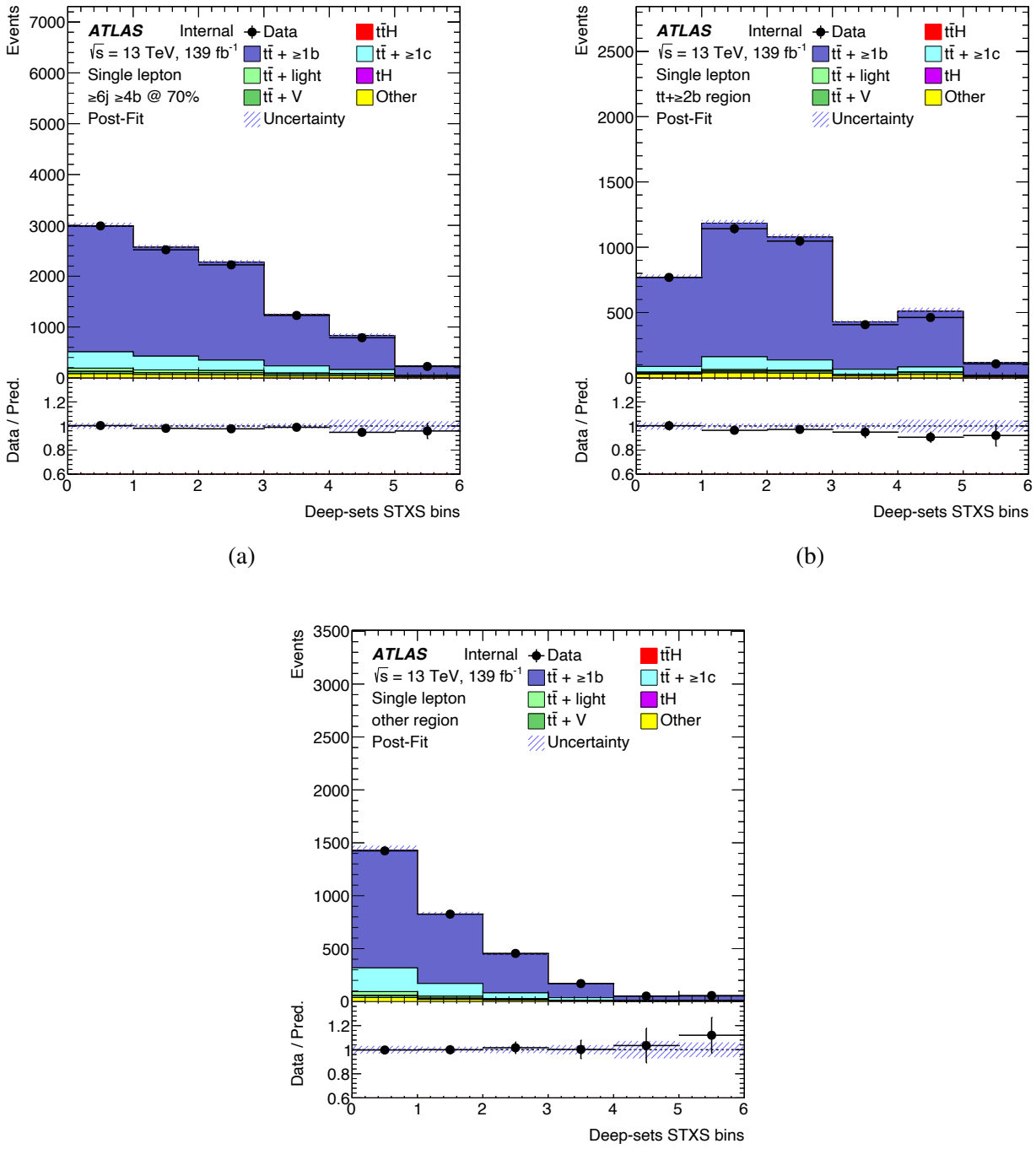


Figure 131 – Post-fit distributions of the Deep-sets STXS bins in (a) inclusive $\geq$ 6j $\geq$ 4b@70%, (b) $t\bar{t}+ \geq$ 2b and (c) other regions. The uncertainty band includes all uncertainties and their correlations. The bins which were blinded (where S/B > 7.7%) in the corresponding pre-fit distributions are kept blinded.

SUMMARY AND CONCLUSION

This thesis focuses on addressing the challenges and making developments in two closely related fields of experimental high-energy physics with the ATLAS experiment at the Large Hadron Collider (LHC). The first one is related to the *b*-tagging algorithms, mainly for the Phase 2 upgrade of the ATLAS detector at the High Luminosity-LHC (HL-LHC). The *b*-tagging algorithms rely on the tracking information from the ATLAS Inner Detector (ID), which will be replaced by the ATLAS Inner Tracker (ITk) to maintain tracking performance during HL-LHC operation. To maintain or possibly improve the current Run 2 *b*-tagging performance, *b*-tagging developments with the ITk were investigated. The next research topic presented in this thesis connects to the Higgs boson searches in association with a pair of top quarks, in which the Higgs decays into a pair of bottom quarks ($t\bar{t}H(H \to b\bar{b})$), using the full Run 2 ATLAS data-set. The analysis explores the Simplified Template Cross-Section (STXS) formalism to measure the signal cross-section as a function of the Higgs transverse momentum (p_T^H). The high multiplicity of the number of *b*-jets due to the additional decay products of the top quark, requires dedicated multivariate analysis techniques based on machine learning, which were explored in this thesis.

The *b*-tagging developments mainly focused on low-level *b*-taggers such as IP3D and SV1. In preparation for the phase-out of the previous upgrade release (r20.20), the *b*-tagging algorithm developments with ITk were synchronised to the most recent ATLAS upgrade software release (r21.9). To improve the performance of IP3D and SV1, some optimizations were also performed. The IP3D algorithm track categories were updated to better exploit the expected p_T-dependence of the impact parameter resolution in the central region ($|\eta| < 2$) and the hit-content dependence in the forward region ($|\eta| > 2$) of the ITk. The new track categorization resulted in a significant improvement in the central region, with up to 50% of light-jet rejection at 70% b-tagging efficiency. Material rejection was also implemented by taking the ITk pixel geometry into account, which 18% improvement in the light-jet rejection at 70% b-tagging efficiency. Those developments were used to perform *b*-tagging performance studies with the updated ITk layout. These *b*-taggers have already been exploited with MV2 and are also expected to benefit future DNN based ITk taggers. The $t\bar{t}H(H \to b\bar{b})$ analysis also benefits from the advanced *b*-tagging methods, due to its challenging final state with at least four *b*-jets.

The legacy round of $t\bar{t}H(H \to b\bar{b})$ analysis was performed in the single-lepton channel. The analysis was carried using Run 2 data-set of pp collision data collected at $\sqrt{s}$ = 13 TeV by the ATLAS detector at the LHC, corresponding to an integrated luminosity 139 fb^{-1}. The purpose of the legacy round was to improve upon the previous round by implementing new methodologies, improving the modelling of dominant background ($t\bar{t} + b\bar{b}$) simulated samples, and improving object definitions, primarily based on Particle-Flow reconstructed jets and the DL1r b-tagging algorithm. The legacy round provided new opportunities to develop and optimise various tools used in the previous version of the analysis, especially the Multivariate analysis (MVA) techniques. One of the main contributions of this thesis is the development of a novel MVA approach based on DNNs. This was used to combine reconstruction of the Higgs p_T and classification to distinguish $t\bar{t}H$ from the background sub-components in a single step. The Deep-sets DNN model was used, which aimed to replace the previously used Boosted Decision Trees (BDTs) and improve the overall sensitivity. Under the signal and background hypothesis, the expected performance of the $t\bar{t}H(H \to b\bar{b})$ analysis was evaluated using an Asimov fit including systematic uncertainties. The fit results were analysed for both STXS and inclusive signal strength measurements. The expected significance for the inclusive measure using Deep-sets is 2.71σ, compared to 2.54σ with BDTs. Similar to the previous round, $t\bar{t}+ \geq 1b$ modelling uncertainty had the greatest impact on the inclusive signal strength measurement. The new Deep-sets MVA model showed a significant improvement in performance in terms of STXS classification, in comparison to BDTs. It also allowed the implementation of a dedicated control regions (CR) for the non-$t\bar{t}+ \geq 2b$ background. A background-only fit was also performed with the data and offered valuable insights for the validation of the background modelling. A good post-fit agreement was observed between the data and Monte Carlo simulation in the CRs. Owing to the more efficient MVA model showing improved STXS performance, the Deeps-sets MVA model can therefore serve as the potential baseline for the legacy $t\bar{t}H(H \to b\bar{b})$ analysis, to be finalised in ATLAS over the coming months, to improve the sensitive of the search.

BIBLIOGRAPHY

[1] ATLAS Collaboration, *Expected tracking and related performance with the updated ATLAS Inner Tracker layout at the High-Luminosity LHC*, ATL-PHYS-PUB-2021-024, 2021, http://cds.cern.ch/record/2776651. Citations on pages 3, 65, 76, 99, 100, and 101.

[2] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588. Citations on pages 13 and 17.

[3] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. B **13** (1964) 168–171. Citations on pages 13 and 17.

[4] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266. Citations on pages 13 and 17.

[5] P. D. Group, *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001. Citations on pages 15, 23, 25, 27, and 28.

[6] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323. Citation on page 19.

[7] ATLAS Collaboration, *Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector*, Phys. Lett. B **800** (2020) 135103, arXiv:1906.02025 [hep-ex]. Citation on page 20.

[8] CMS Collaboration, *Combination of searches for Higgs boson pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Cms-pas-hig-17-030, 2018, https://cds.cern.ch/record/2628486. Citation on page 20.

[9] B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of Leptonic Charge*, Zh. Eksp. Teor. Fiz. **53** (1967) 1717–1725. Citation on page 22.

[10] V. N. Gribov and B. Pontecorvo, *Neutrino astronomy and lepton charge*, Phys. Lett. B **28** (1969) 493. Citation on page 22.

[11] M. Gell-Mann, *Symmetries of baryons and mesons*, Phys. Rev. **125** (1962) 1067–1084. Citation on page 23.

[12] G. P. Salam, *Elements of QCD for hadron colliders*, in *2009 European School of High-Energy Physics*. 11, 2010. arXiv:1011.5131 [hep-ph]. Citation on page 24.

[13] D. J. Gross and F. Wilczek, *Asymptotically Free Gauge Theories - I*, Phys. Rev. D **8** (1973) 3633–3652. Citation on page 25.

[14] R. K. Ellis and W. J. Stirling, *QCD and collider physics*, in *1989 CERN-JINR School of Physics*. 8, 1990. Citations on pages 25 and 27.

[15] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*, Eur. Phys. J. C **75** (2015) 132, arXiv:1412.7420 [hep-ph]. Citations on pages 25 and 27.

[16] J. Alwall et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, Eur. Phys. J. C **53** (2008) 473–500, arXiv:0706.2569 [hep-ph]. Citation on page 28.

[17] S. Catani, S. Dittmaier, M. H. Seymour, and Z. Trocsanyi, *The Dipole formalism for next-to-leading order QCD calculations with massive partons*, Nucl. Phys. B **627** (2002) 189–265, arXiv:hep-ph/0201036. Citation on page 28.

[18] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31–145. Citation on page 28.

[19] K. Geiger, *Particle production in high-energy nuclear collisions: A Parton cascade cluster hadronization model*, Phys. Rev. D **47** (1993) 133–159. Citation on page 28.

[20] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244–289, arXiv:1207.1303 [hep-ph]. Citations on pages 28, 51, and 148.

[21] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1–29, arXiv:1207.7214 [hep-ex]. Citations on pages 28 and 32.

[22] CMS Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30–61, arXiv:1207.7235 [hep-ex]. Citation on page 28.

[23] C. Englert et al., *Precision Measurements of Higgs Couplings: Implications for New Physics Scales*, J. Phys. G **41** (2014) 113001, arXiv:1403.7191 [hep-ph]. Citations on pages 28 and 143.

[24] ATLAS Collaboration, *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment*, Eur. Phys. J. C **76** (2016) 6, arXiv:1507.04548 [hep-ex]. Citation on page 29.

[25] L. H. C. S. W. Group, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, arXiv:1610.07922 [hep-ph]. Citations on pages 30, 31, 36, and 168.

[26] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1–29, arXiv:1207.7214 [hep-ex]. Citation on page 31.

[27] CMS Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30–61, arXiv:1207.7235 [hep-ex]. Citation on page 31.

[28] ATLAS and CMS Collaboration, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, JHEP **08** (2016) 045, arXiv:1606.02266 [hep-ex]. Citation on page 31.

[29] CMS Collaboration, *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector*, Phys. Lett. B **786** (2018) 59–86, arXiv:1808.08238 [hep-ex]. Citation on page 31.

[30] CMS Collaboration, *Observation of Higgs boson decay to bottom quarks*, Phys. Rev. Lett. **121** (2018) 121801, arXiv:1808.08242 [hep-ex]. Citation on page 31.

[31] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173–191, arXiv:1806.00425 [hep-ex]. Citations on pages 31, 33, and 34.

[32] CMS Collaboration, *Observation of $t\bar{t}H$ production*, Phys. Rev. Lett. **120** (2018) 231801, arXiv:1804.02610 [hep-ex]. Citations on pages 31 and 35.

[33] ATLAS Collaboration, *Measurement of Higgs boson decay into b-quarks in associated production with a top-quark pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **06** (2022) 097, arXiv:2111.06712 [hep-ex]. Citations on pages 33, 35, 37, 104, 109, 110, 111, 112, 113, 143, 144, 149, 163, 164, 168, 181, and 182.

[34] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, https://cds.cern.ch/record/2677054. Citations on pages 33 and 163.

[35] CMS Collaboration, *Measurement of* $\mathrm{t\bar{t}H}$ *production in the* $\mathrm{H} \to \mathrm{b\bar{b}}$ *decay channel in* $41.5\,\mathrm{fb}^{-1}$ *of proton-proton collision data at* $\sqrt{s} = 13\,\mathrm{TeV}$, CMS-PAS-HIG-18-030, 2019, https://cds.cern.ch/record/2675023. Citations on pages 34, 36, and 119.

[36] N. Berger et al., *Simplified Template Cross Sections - Stage 1.1*, arXiv:1906.02754 [hep-ph]. Citation on page 36.

[37] ATLAS Collaboration, *Measurement of Higgs boson production in association with a* $t\bar{t}$ *pair in the diphoton decay channel using 139fb*$^{-1}$ *of LHC data collected at* $\sqrt{s} = 13$*~TeV by the ATLAS experiment*, ATLAS-CONF-2019-004, 2019, https://cds.cern.ch/record/2668103. Citation on page 37.

[38] F. Boudjema, R. M. Godbole, D. Guadagnoli, and K. A. Mohan, *Lab-frame observables for probing the top-Higgs interaction*, Phys. Rev. D **92** (2015) 015019, arXiv:1501.03157 [hep-ph]. Citation on page 37.

[39] F. Maltoni, D. Pagani, A. Shivaji, and X. Zhao, *Trilinear Higgs coupling determination via single-Higgs differential measurements at the LHC*, Eur. Phys. J. C **77** (2017) 887, arXiv:1709.08649 [hep-ph]. Citation on page 37.

[40] O. S. Brüning et al., *LHC Design Report*, 2004, https://cds.cern.ch/record/782076. Citation on page 38.

[41] LEP Injector Study Group, *LEP Design Report Vol.1: The LEP Injector Chain*, 1983, http://cds.cern.ch/record/98881. Citation on page 38.

[42] Vollaire, J. and others, *Linac4 design report*, 9, 2020, https://cds.cern.ch/record/2736208. Citation on page 39.

[43] F. Marcastel, *CERN's Accelerator Complex*, 2013. https://cds.cern.ch/record/1621583. Citation on page 40.

[44] J.-L. Caron, *Cross section of LHC dipole.*, 1998. https://cds.cern.ch/record/841539. Citation on page 41.

[45] ATLAS Collaboration, *Luminosity Public Results for Run 2*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2. Citation on page 42.

[46] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003. Citations on pages 41, 42, and 63.

[47] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004. Citation on page 42.

[48] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005. Citation on page 42.

[49] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002. Citation on page 42.

[50] LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006. Citation on page 42.

[51] TOTEM Collaboration, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007. Citation on page 42.

[52] FASER Collaboration, *Technical Proposal for FASER: ForwArd Search ExpeRiment at the LHC*, arXiv:1812.09139 [physics.ins-det]. Citation on page 42.

[53] MoEDAL Collaboration, *The Physics Programme Of The MoEDAL Experiment At The LHC*, Int. J. Mod. Phys. A **29** (2014) 1430050, arXiv:1405.7662 [hep-ph]. Citation on page 42.

[54] J. Pequenao, *Computer generated image of the whole ATLAS detector*, 2008. https://cds.cern.ch/record/1095924. Citation on page 43.

[55] ATLAS Collaboration, *ATLAS inner detector: Technical design report. Vol. 1*, atlas-tdr-4; cern-lhcc-97-17, 1997. https://cds.cern.ch/record/331063. Citation on page 45.

[56] ATLAS Collaboration, *ATLAS inner detector: Technical design report. Vol. 2*, cern-lhcc-97-17, 1997. https://cds.cern.ch/record/331064. Citation on page 45.

[57] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector in Run-2*, Eur. Phys. J. C **80** (2020) 1194, arXiv:2007.07624 [hep-ex]. Citation on page 45.

[58] ATLAS Collaboration, *ATLAS Pixel Detector: Technical Design Report*, atlas-tdr-11; cern-lhcc-98-013, 1998. https://cds.cern.ch/record/2285585. Citation on page 45.

[59] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, atlas-tdr-19; cern-lhcc-2010-013, 2010. https://cds.cern.ch/record/1291633. Citation on page 45.

[60] ATLAS Collaboration, *Operation and performance of the ATLAS semiconductor tracker*, JINST **9** (2014) P08009, arXiv:1404.7473 [hep-ex]. Citation on page 46.

[61] ATLAS Collaboration, *Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties*, JINST **12** (2017) P05002, arXiv:1702.06473 [hep-ex]. Citation on page 46.

[62] ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report*, cern-lhcc-96-42, 1996. https://cds.cern.ch/record/331062. Citation on page 47.

[63] ATLAS Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*, cern-lhcc-96-41, 1996. https://cds.cern.ch/record/331061. Citation on page 47.

[64] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*, tech. rep., 1997. https://cds.cern.ch/record/331068. Citation on page 48.

[65] ATLAS Collaboration, *New Small Wheel Technical Design Report*, tech. rep., 2013. https://cds.cern.ch/record/1552862. Citation on page 48.

[66] ATLAS Collaboration, *Commissioning of the ATLAS Muon Spectrometer with Cosmic Rays*, Eur. Phys. J. C **70** (2010) 875–916, arXiv:1006.4384 [physics.ins-det]. Citation on page 49.

[67] ATLAS Collaboration, *ATLAS magnet system: Technical Design Report, 1*, cern-lhcc-97-18, 1997. https://cds.cern.ch/record/338080. Citation on page 48.

[68] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004, arXiv:2007.12539 [physics.ins-det]. Citations on pages 49 and 50.

[69] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv:1002.2581 [hep-ph]. Citations on pages 50, 75, and 148.

[70] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv:1405.0301 [hep-ph]. Citations on pages 50 and 148.

[71] Sherpa collaboration, *Event Generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv:1905.09127 [hep-ph]. Citations on pages 50 and 148.

[72] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv:1410.3012 [hep-ph]. Citations on pages 50 and 75.

[73] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. C **58** (2008) 639–707, arXiv:0803.0883 [hep-ph]. Citations on pages 50 and 148.

[74] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C **75** (2015) 204, arXiv:1412.3989 [hep-ph]. Citations on pages 51 and 148.

[75] GEANT4 Collaboration, GEANT4- *a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250. Citation on page 51.

[76] W. Lukas, *Fast Simulation for ATLAS: Atlfast-II and ISF*, tech. rep., CERN, Geneva, 2012. https://cds.cern.ch/record/1458503. Citation on page 51.

[77] J. Pequenao and P. Schaffner, *How ATLAS detects particles: diagram of particle paths in the detector*, 2013. https://cds.cern.ch/record/1505342. Citation on page 52.

[78] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, and A. Salzburger, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, ATL-SOFT-PUB-2007-007, ATL-COM-SOFT-2007-002, 2007, https://cds.cern.ch/record/1020106. Citation on page 52.

[79] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444–450. Citation on page 52.

[80] ATLAS Collaboration, *A neural network clustering algorithm for the ATLAS silicon pixel detector*, JINST **9** (2014) P09009, arXiv:1406.7690 [hep-ex]. Citation on page 53.

[81] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv:1908.00005 [hep-ex]. Citations on pages 54 and 55.

[82] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, ATLAS-CONF-2016-024, 2016, https://cds.cern.ch/record/2157687. Citations on pages 54 and 146.

[83] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012, https://cds.cern.ch/record/1449796. Citation on page 54.

[84] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74** (2014) 3071, arXiv:1407.5063 [hep-ex]. Citation on page 55.

[85] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015-2016 LHC proton-proton collision data*, JINST **14** (2019) P03017, arXiv:1812.03848 [hep-ex]. Citation on page 55.

[86] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at* $\sqrt{s}$ *= 13 TeV)*, Eur. Phys. J. C **81** (2021) 578, arXiv:2012.00578 [hep-ex]. Citations on pages 56, 57, 63, and 146.

[87] W. Lampl et al., *Calorimeter clustering algorithms: Description and performance*, ATL-LARG-PUB-2008-002, 2008, https://cds.cern.ch/record/1099735. Citation on page 58.

[88] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph]. Citations on pages 58 and 75.

[89] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, https://cds.cern.ch/record/1700870. Citation on page 59.

[90] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in collisions at* $\sqrt{s}$ *= 8 TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv:1510.03823 [hep-ex]. Citation on page 59.

[91] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 689, arXiv:2007.02645 [hep-ex]. Citation on page 59.

[92] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv:1703.10485 [hep-ex]. Citation on page 60.

[93] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at* $\sqrt{s}$ *= 13 TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv:1802.08168 [hep-ex]. Citation on page 61.

[94] G. Apollinari, O. Brüning, T. Nakamoto, and L. Rossi, *High Luminosity Large Hadron Collider HL-LHC*, CERN Yellow Rep. (2015) 1–19, arXiv:1705.08830 [physics.acc-ph]. Citation on page 62.

[95] CERN, *HL-LHC Project Schedule*, 2022. https://project-hl-lhc-industry.web.cern.ch/content/project-schedule. Citation on page 62.

[96] ATLAS Collaboration, *ATLAS Inner Tracker Strip Detector: Technical Design Report*, atlas-tdr-025; cern-lhcc-2017-005, 2017. https://cds.cern.ch/record/2257755. Citation on page 63.

[97] ATLAS Collaboration, *ATLAS Inner Tracker Pixel Detector: Technical Design Report*, atlas-tdr-030; cern-lhcc-2017-021, 2017. https://cds.cern.ch/record/2285585. Citations on pages 63 and 75.

[98] ATLAS Collaboration, *A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade: Technical Design Report*, atlas-tdr-031; cern-lhcc-2020-007, 2020. https://cds.cern.ch/record/2719855. Citations on pages 63 and 74.

[99] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*, atlas-tdr-028; cern-lhcc-2017-019, 2017. https://cds.cern.ch/record/2285583. Citation on page 63.

[100] ATLAS Collaboration, *ATLAS Liquid Argon Calorimeter Phase-II Upgrade: Technical Design Report*, atlas-tdr-027; cern-lhcc-2017-018, 2017. https://cds.cern.ch/record/2285582. Citation on page 63.

[101] ATLAS Collaboration, *ATLAS TDAQ Phase-II Upgrade: Technical Design Report*, atlas-tdr-029; cern-lhcc-2017-020, 2017. https://cds.cern.ch/record/2285584. Citation on page 63.

[102] ATLAS Collaboration, *Expected Tracking Performance of the ATLAS Inner Tracker at the HL-LHC*, tech. rep., 2019. http://cds.cern.ch/record/2669540. Citation on page 64.

[103] ATLAS Collaboration, *Expected b-tagging performance with the upgraded ATLAS Inner Tracker detector at the High-Luminosity LHC*, ATL-PHYS-PUB-2020-005, 2020, http://cds.cern.ch/record/2713377. Citation on page 70.

[104] ATLAS Collaboration, *Secondary vertex finding for jet flavour identification with the ATLAS detector*, ATL-PHYS-PUB-2017-011, 2017, https://cds.cern.ch/record/2270366. Citation on page 71.

[105] ATLAS Collaboration, *Development of ATLAS Primary Vertex Reconstruction for LHC Run 3*, ATL-PHYS-PUB-2019-015, 2019, https://cds.cern.ch/record/2670380. Citation on page 71.

[106] ATLAS Collaboration, *Software Performance of the ATLAS Track Reconstruction for LHC Run 3*, ATL-PHYS-PUB-2021-012, 2019, https://cds.cern.ch/record/2766886. Citation on page 71.

[107] ATLAS Collaboration, *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*, ATL-PHYS-PUB-2018-025, 2018, https://cds.cern.ch/record/2645405. Citation on page 73.

[108] ATLAS Collaboration, *Performance and Calibration of the JetFitterCharm Algorithm for c-Jet Identification*, ATL-PHYS-PUB-2015-001, 2015, https://cds.cern.ch/record/1980463. Citation on page 73.

[109] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with* $\mathrm{t\bar{t}}$ *events in pp collisions at* $\sqrt{s}$ *= 13 TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv:1907.05120 [hep-ex]. Citations on pages 73, 74, and 103.

[110] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, 2017, https://cds.cern.ch/record/2273281. Citation on page 73.

[111] ATLAS Collaboration, *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*, ATL-PHYS-PUB-2017-003, 2017, https://cds.cern.ch/record/2255226. Citation on page 73.

[112] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv:hep-ph/0409146. Citations on pages 75 and 148.

[113] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152. Citations on pages 75 and 148.

[114] ATLAS Collaboration, *Development of ATLAS Primary Vertex Reconstruction for LHC Run 3*, ATL-PHYS-PUB-2019-015, 2019, https://cds.cern.ch/record/2670380. Citation on page 75.

[115] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11** (2016) P04008, arXiv:1512.01094 [hep-ex]. Citation on page 75.

[116] J. Duarte et al., *Fast inference of deep neural networks in FPGAs for particle physics*, JINST **13** (2018) P07027, arXiv:1804.06913 [physics.ins-det]. Citation on page 103.

[117] ATLAS Collaboration, *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, ATL-PHYS-PUB-2019-033, 2019. Citation on page 103.

[118] ATLAS Collaboration, *Deep Sets based Neural Networks for Impact Parameter Flavour Tagging in ATLAS*, ATL-PHYS-PUB-2020-014, 2020, https://cds.cern.ch/record/2718948. Citations on pages 103 and 123.

[119] Y. Coadou, *Boosted decision trees*, arXiv:2206.09645 [physics.data-an]. Citation on page 106.

[120] Y. Freund and R. E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, J. Comput. Syst. Sci. **55** (1997) 119–139. Citation on page 107.

[121] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81–86. Citation on page 107.

[122] A. Hocker et al., *TMVA - Toolkit for Multivariate Data Analysis*, arXiv:physics/0703039. Citation on page 107.

[123] B. Nachman, *A guide for deploying Deep Learning in LHC searches: How to achieve optimality and account for uncertainty*, SciPost Phys. **8** (2020) 090, arXiv:1909.03081 [hep-ph]. Citation on page 119.

[124] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org. Citations on pages 119 and 122.

[125] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, arXiv:1412.6980. Citation on page 122.

[126] S. Ioffe and C. Szegedy, *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*, arXiv:1502.03167. Citation on page 122.

[127] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, *Dropout: A Simple Way to Prevent Neural Networks from Overfitting*, Journal of Machine Learning Research **15** (2014) 1929–1958. Citation on page 122.

[128] Uproot, *Uproot*, 2019. https://uproot.readthedocs.io/en/latest/index.html. Citation on page 123.

[129] F. Chollet et al., *KERAS*, 2015. https://keras.io/. Citation on page 123.

[130] F. Pedregosa et al., *Scikit-learn: Machine Learning in Python*, Journal of Machine Learning Research **12** (2011) 2825–2830. Citation on page 123.

[131] J. Bai et al., *ONNX: Open Neural Network Exchange*, 2021. https://onnx.ai/. Citation on page 123.

[132] M. Zaheer et al., *Deep Sets*, arXiv:1703.06114. Citation on page 123.

[133] P. T. Komiske, E. M. Metodiev, and J. Thaler, *Energy Flow Networks: Deep Sets for Particle Jets*, JHEP **01** (2019) 121, arXiv:1810.05165 [hep-ph]. Citation on page 123.

[134] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv:1611.09661 [hep-ex]. Citation on page 146.

[135] ATLAS Collaboration, *Analysis of $t\bar{t}H$ and $t\bar{t}W$ production in multilepton final states with the ATLAS detector*, ATLAS-CONF-2019-045, 2019, https://cds.cern.ch/record/2693930. Citation on page 147.

[136] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*, ATL-PHYS-PUB-2016-017, 2016, https://cds.cern.ch/record/2206965. Citation on page 148.

[137] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv:1512.01178 [hep-ph]. Citation on page 148.

[138] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, https://cds.cern.ch/record/1966419. Citation on page 148.

[139] S. Frixione et al., *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **04** (2007) 081, arXiv:hep-ph/0702198. Citation on page 148.

[140] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **03** (2013) 015, arXiv:1212.3460 [hep-ph]. Citation on page 148.

[141] ATLAS Collaboration, *Studies of Monte Carlo predictions for the $t\bar{t}b\bar{b}$ process*, ATL-PHYS-PUB-2022-006, 2022, https://cds.cern.ch/record/2802806. Citations on pages 149 and 150.

[142] NNPDF collaboration, *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv:1410.8849 [hep-ph]. Citations on pages 148, 149, and 152.

[143] S. Frixione et al., *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv:0805.3067 [hep-ph]. Citations on pages 149 and 152.

[144] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, ATL-PHYS-PUB-2016-020, 2016, https://cds.cern.ch/record/2216168. Citations on pages 149, 150, and 152.

[145] E. Bothmann, M. Schönherr, and S. Schumann, *Reweighting QCD matrix-element and parton-shower calculations*, Eur. Phys. J. C **76** (2016) 590, arXiv:1606.08753 [hep-ph]. Citations on pages 149 and 152.

[146] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027, arXiv:1207.5030 [hep-ph]. Citations on pages 149 and 152.

[147] S. Frixione et al., *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv:0709.2092 [hep-ph]. Citation on page 148.

[148] H. B. Hartanto et al., *Higgs boson production in association with top quarks in the POWHEG BOX*, Phys. Rev. D **91** (2015) 094003, arXiv:1501.04498 [hep-ph]. Citation on page 148.

[149] T. Ježo et al., *New NLOPS predictions for $t\bar{t}+b$ -jet production at the LHC*, Eur. Phys. J. C **78** (2018) 502, arXiv:1802.00426 [hep-ph]. Citation on page 150.

[150] F. Cascioli et al., *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv:1111.5206 [hep-ph]. Citation on page 150.

[151] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv:1111.1220 [hep-ph]. Citation on page 151.

[152] ATLAS Collaboration, *Study of ttbb and ttW background modelling for ttH analyses*, ATL-PHYS-PUB-2022-026, 2022, ttp://cds.cern.ch/record/2810864. Citation on page 151.

[153] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv:0709.1027 [hep-ph]. Citation on page 152.

[154] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv:hep-ph/0201195. Citation on page 152.

[155] ATLAS Collaboration, *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (2016) 182002, arXiv:1606.02625 [hep-ex]. Citation on page 163.

[156] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 689, arXiv:2007.02645 [hep-ex]. Citations on pages 163 and 165.

[157] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv:1510.03823 [hep-ex]. Citation on page 165.

[158] R. Raitio and W. W. Wada, *Higgs-boson production at large transverse momentum in quantum chromodynamics*, Phys. Rev. D **19** (1979) 941. Citation on page 165.

[159] W. Beenakker et al., *NLO QCD corrections to $t\bar{t}H$ production in hadron collisions*, Nucl. Phys. B **653** (2003) 151–203, arXiv:hep-ph/0211352. Citation on page 165.

[160] Y. Zhang et al., *QCD NLO and EW NLO corrections to $t\bar{t}H$ production with top quark decays at hadron collider*, Phys. Lett. B **738** (2014) 1–5, arXiv:1407.1110 [hep-ph]. Citation on page 165.

[161] ATLAS Collaboration, *Study of top-quark pair modelling and uncertainties using ATLAS measurements at $\sqrt{s}$=13 TeV*, ATL-PHYS-PUB-2020-023, 2020, https://cds.cern.ch/record/2730443. Citation on page 166.

[162] I. W. Stewart and F. J. Tackmann, *Theory Uncertainties for Higgs and Other Searches Using Jet Bins*, Phys. Rev. D **85** (2012) 034011, arXiv:1107.2117 [hep-ph]. Citation on page 166.

[163] J. M. Campbell and R. K. Ellis, *$t\bar{t}W^{\pm}$ production and decay at NLO*, JHEP **07** (2012) 052, arXiv:1204.5678 [hep-ph]. Citation on page 168.

[164] M. Aliev et al., *HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR*, Comput. Phys. Commun. **182** (2011) 1034–1046, arXiv:1007.1327 [hep-ph]. Citation on page 168.

[165] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz, P. Rieck, and P. Uwer, *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, Comput. Phys. Commun. **191** (2015) 74–89, arXiv:1406.4403 [hep-ph]. Citation on page 168.

[166] A. D. Martin et al., *Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections*, Eur. Phys. J. C **64** (2009) 653–680, arXiv:0905.3531 [hep-ph]. Citation on page 168.

[167] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber, and C. D. White, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv:0805.3067 [hep-ph]. Citation on page 168.

[168] ATLAS Collaboration, *Multi-Boson Simulation for 13 TeV ATLAS Analyses*, ATL-PHYS-PUB-2016-002, 2016, https://cds.cern.ch/record/2119986. Citation on page 168.

[169] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv:1405.0301 [hep-ph]. Citation on page 168.

[170] J. Neyman and E. S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character **231** (1933) 289–337, http://www.jstor.org/stable/91247. Citation on page 169.

[171] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv:1007.1727 [physics.data-an]. Citation on page 170.

[172] L. Moneta et al., *The RooStats project*, PoS **ACAT2010** (2011) 057, arXiv:1009.1003 [physics.data-an]. Citation on page 170.