Study of the Higgs production in association with tt quarks

Ana Elena Dumitriu

CPPM-T-2018-02

AIX-MARSEILLE UNIVERSITÉ
Ecole Doctorale 352
Faculté des Sciences
Centre de Physique des Particules de Marseille

UNIVERSITÉ DE BUCAREST
Faculté de Physique
Horia Hulubei Institut national de recherche et développement en physique et en génie nucléaire (IFIN-HH)

Thèse présentée pour obtenir le grade universitaire de docteur

Discipline : Physique et Sciences de la Matière
Spécialité : Physique des Particules et Astroparticules

Ana Elena DUMITRIU

Étude de la production de Higgs en association avec tt quarks
Study of the Higgs production in association with tt quarks

Soutenue le 18/09/2018 devant le jury :

Dr. Cristinel DIACONU	**CPPM (France)**	**Directeur de thèse**
Dr. Emmanuel MONNIER	**CPPM (France)**	**Directeur de thèse**
Dr. Calin ALEXA	**IFIN-HH (Ro)**	**Directeur de thèse**
Prof. Virgil BARAN	**Université de Bucarest**	**Rapporteur**
Dr. Anne-catherine LE BIHAN	**Institut Pluridisciplinaire Hubert Curien**	**Rapporteur**
Dr. Jean Baptiste DE VIVIE DE REGIE	**Laboratoire de l'Accélérateur Linéaire**	**Examinateur**
Prof. Daniela DRAGOMAN	**Université de Bucarest**	**Examinateur**

Numéro national de thèse/suffixe local : 2018AIXM0304/031ED352

CPPM-T-2018-02

UNIVERSITY OF BUCHAREST
Faculty of Physics
Horia Hulubei National Institute
for R&D in Physics and Nuclear Engineering
(IFIN-HH)

AIX-MARSEILLE UNIVERSITY
Doctorale School 352
Faculty of Sciences
Centre de Physique des Particules de
Marseille

Thesis presented to obtain the degree of Doctor in Physics of the University of Bucharest and Aix-Marseille University

Ana Elena DUMITRIU

Studiul producerii bozonului Higgs in asociere cu o pereche de cuarci tt in experimentul ATLAS de la LHC
Study of the Higgs production in association with tt quarks

Sustained on 18/09/2018 in front of a jury composed of:

Dr. Cristinel DIACONU	**CPPM (France)**	**Thesis supervisor**
Dr. Emmanuel MONNIER	**CPPM (France)**	**Thesis supervisor**
Dr. Calin ALEXA	**IFIN-HH (Ro)**	**Thesis supervisor**
Prof. Virgil BARAN	**University of Bucharest**	**Reviewer**
Dr. Anne-catherine LE BIHAN	**Institut Pluridisciplinaire Hubert Curien**	**Reviewer**
Dr. Jean Baptiste DE VIVIE DE REGIE	**Laboratoire de l'Accélérateur Linéaire**	**Examiner**
Prof. Daniela DRAGOMAN	**University of Bucharest**	**Examiner**

Numéro national de thèse/suffixe local : 2018AIXM0304/031ED352

Abstract

Résumé

Cette thèse de doctorat se déroule dans le cadre de la collaboration ATLAS, opérant le détecteur éponyme du même nom au Large Hadron Collider (LHC).

ATLAS est l'un des deux détecteurs à usage général, avec CMS, fonctionnant auprès du grand collisioneur de Hadron, (LHC) au CERN. Il a pour objectif, outre l'étude des origines de la masse, de rechercher des signes de nouvelle physique.

Le 4 juillet 2012, les expériences ATLAS et CMS ont annoncé avoir observé une nouvelle particule dans la région de masse autour de 126 GeV, conforme au boson de Higgs prévu par le modèle standard. Ce modèle standard de la physique des particules est l'une des théories les plus réussies en physique. Il a été validé expérimentalement avec beaucoup de succès jusqu'à présent, même si certaines questions restent sans réponse, telles que la hiérarchie des masses, la matière noire et l'énergie sombre, le graviton, etc. Il décrit les interactions fondamentales entre les particules élémentaires et le formalisme de l'acquisition des masses à travers leur interaction au champ de Higgs.

Le premier chapitre de cette thèse introduit de manière exhaustive le cadre mathématique du modèle standard en mettant l'accent sur les concepts utilisés par la suite. Le LHC, le complexe d'accélérateurs qui lui est associé et le détecteur ATLAS sont ensuite présentés dans le deuxième chapitre définissant ainsi la partie expérimentale des recherches présentée dans cette thèse. La définition et la sélection des objets physique y est également donnée.

Dans ce chapitre, un accent particulier est mis sur le système calorimétrique car il a une importance significative pour ce travail de thèse notamment sur les performances d'identification des électrons étudiées à l'aide de la méthode W et présentées au troisième chapitre. Cette étude correspond également à la première année de qualification, nécessaire pour tout étudiant travaillant sur ATLAS pour devenir auteur de la production scientifique de l'expérience.

La campagne de prise de données d'ATLAS, appelée Run 2, débutée en 2015 et qui s'achèvera fin 2018, offre par rapport à la précédente campagne, dite Run 1, de nouveaux défis expérimentaux en raison de l'augmentation de l'énergie et de la luminosité des faisceaux. Cela a pour conséquence que les performances du détecteur sont modifiées et par conséquent doivent être optimisées et mesurées au mieux. C'est l'objet de l'analyse originale,

basée sur des événements de désintégration semi-leptoniques de boson W, présentée au troisième chapitre et comparée à l'analyse de référence utilisant des événement de désintégration leptonique de boson Z. Ces études ont ensuite été étendues en étudiant puis proposant une nouvelle configuration de déclenchement des W leptoniques décrite également dans ce chapitre. L'analyse a ensuite été poursuivie sur les données 2017-2018 grâce à une étroite collaboration avec les experts des performances basés au CERN.

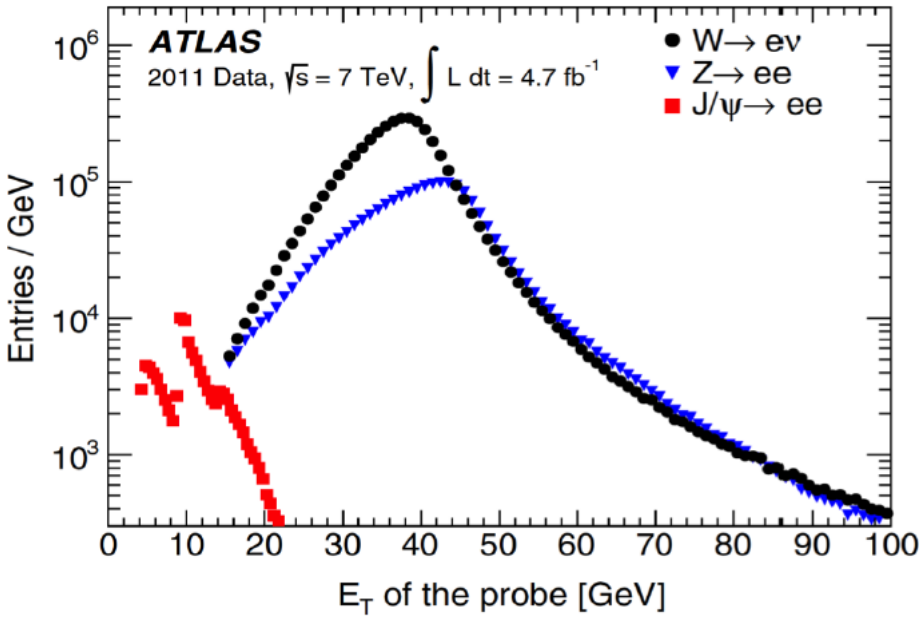


Fig. 1 Distributions d'énergie transverse des électrons sondes satisfaisant aux critères d'identification dur pour des $W \to e\nu$ (cercles complets) et des $Z \to ee$ (triangles bleus) et $J/\Psi \to ee$ (carrés rouges) provenant d'une sélection utilisant l'ensemble des données 2011.

L'échantillon de données collecté de 2015 à 2017 par ATLAS a été utilisé pour mesurer les efficacités d'identification des électrons en utilisant une méthode d'étiquetage sonde. Les résultats d'efficacité d'identification des électrons présentés dans cette thèse utilise les désintégrations $W \to e\nu$. De nombreux avantages sont apparus pour utiliser ces événements $W \to e\nu$ dans la mesure des efficacités d'identification électronique: les bosons W et Z ont une grande plage cinématique commune, comme on peut le voir sur la Figure 1, et donc les événements $W \to e\nu$ peuvent contribuer à comprendre et améliorer les incertitudes systématiques; les W ont une statistique importante pour la gamme de E_T basse (15-25 GeV), ce qui permet de compléter le canal J/Ψ et offre des signatures expérimentales intéressantes en raison de ces grandes statistiques. Les algorithmes d'identification utilisés sont basés sur un discriminateur de vraisemblance multivarié. Les différences entre les données et la

simulation Monte Carlo ont été prises en compte dans facteur d'échelle SF=données/MC, obtenus en fonction de l'énergie transversale et de la pseudo-rapidité des électrons.

La principale difficulté de ce canal de désintégration réside dans le control des bruits de fond. Pour avoir un signal aussi propre que possible, des études d'optimisation et d'estimation de ces bruits de fond ont été effectuées en utilisant des méthodes basées sur des coupures et TMVA comme discuté précédemment. Elles sont présentée dans ce chapitre trois. Les facteurs d'échelle utilisant les W ont été comparés avec ceux issus de $Z \to ee$. Un accord de 10% entre les deux méthodes a été obtenu validant cette méthode. Ceci représente la première mesure d'efficacité d'identification des électrons avec les données Run 2 utilisant les événements $W \to e\nu$. En outre, pour la période 2017-2018, un nouvel ensemble de configurations de triggers a été émulé puis proposé et validé permettant une augmentation importante de la statistique disponible sur l'ensemble du Run 2.

L'ensemble de ce travail est important pour la collaboration ATLAS car, après une importante étude d'optimisation du control des fonds, il a permis une première validation complète des performances sur l'identification des électrons et les facteurs d'échelle s'appuyant sur des événements $W \to e\nu$ dans les données Run 2. Maintenant, après les études d'émulation de déclenchement et la modification des conditions de déclenchement, ce canal est très intéressant pour contribuer avec les événements $Z \to ee$ et J/Ψ à améliorer l'identification des électrons sur une large plage en impulsion transverse et surtout à diminuer les erreurs systématiques.

Un des premiers objectifs du Large Hadron Collider était la découverte du boson de Higgs puis l'étude de ses propriétés. Après la découverte du boson de Higgs, réalisation majeure du Run I, l'effort du Run II est maintenant centré sur la mesure des propriétés du boson de Higgs, en particulier de son couplage aux quarks. Le quark top est la particule élémentaire la plus lourde connue à ce jour, il a donc le plus grand couplage au boson de Higgs. Pour évaluer directement la mesure de ce couplage au LHC, on peut rechercher la production associée d'un boson de Higgs avec une paire de quarks top ($t\bar{t}H$).

Cette thèse présente, dans le quatrième chapitre, une recherche du mode de production $t\bar{t}H$ dans l'état final multi-leptonique basée sur 36.1 fb^{-1} de données de collision pp produites pas le LHC à $\sqrt{s}$ = 13 TeV et collectées en 2015 et 2016 par le détecteur ATLAS. Sept états finaux (comme on peut le voir sur la Figure 2), ciblant les désintégrations du boson de Higgs vers WW*, $\tau\tau$ et ZZ*, sont étudiés pour rechercher la présence du boson Higgs du

modèle standard (SM). L'analyse $t\bar{t}H \rightarrow$ multileptons est ensuite combinée avec d'autres états finaux $t\bar{t}H$ pour fournir un résultat final sur la force de signal:

- Multilepton
- $H \rightarrow b\bar{b}$
- $H \rightarrow \gamma\gamma$
- $H \rightarrow 4l$

Le résultat combiné donne la valeur du meilleur ajustement de la force du signal $\mu_{t\bar{t}H} = 1.17 \pm 0,19(stat)^{+0.27}_{-0.23}$, avec un valeur correspondante de la section efficace de $\sigma_{t\bar{t}H} = 590^{+160}_{-150} fb$ et une signification de 4.2σ pour une valeur attendue de 3.8σ.

Cette thèse, se concentre sur le canal 4l (Figure 3), présentant toutes les étapes de la procédure d'analyse: études d'optimisation, définition des régions de signal et des régions de contrôle, estimation des bruits de fond de fausse identification et procédure d'ajustement. Un ajustement, dit de "maximum-de vraisemblance ", est effectué simultanément sur tous les canaux multilepton, y compris le canal à 4l dans l'état final, pour extraire la force du signal, μ, en supposant une masse de boson de Higgs de 125 GeV. La force du signal est définie comme le rapport entre le signal $t\bar{t}H$ mesuré et celui attendu du modèle standard. Dans le cas d'un boson SM Higgs μ = 1, mais s'il est absent, alors μ = 0. La valeur de meilleur ajustement observée (attendue) de $\mu_{t\bar{t}H}$, combinant tous les canaux, est de $1.6^{+0.3}_{-0.3}$ (stat) $^{+0.4}_{-0.3}$ (syst), comme indiqué dans la Figure 4 pour chaque catégorie individuelle d'état final et la combinaison de tous les canaux. La signification observée (attendue) par rapport à l'hypothèse sans $t\bar{t}H$ est de 4.1 σ (2.8 σ) avec une extrapolation dépendante du modèle effectuée dans l'espace de phase inclusif. Cela se traduit par une section efficace de production observée de $\sigma_{t\bar{t}H} = 790^{+230}_{-210} fb$ pour une section efficace théorique de $\sigma_{t\bar{t}H} = 507^{+35}_{-50} fb$.

La Figure 5 montre une comparaison entre les données et la simulation après ajustement pour les huit régions de signal et les quatre régions de contrôle.

Profitant de la similitude des états finaux des recherches de $t\bar{t}H$ multi-lepton et de boson de Higgs doublement chargé en $W^{\pm}W^{\pm}$, et en s'appuyant sur l'expertise obtenue sur le canal en quatre leptons dans la première analyse, une recherche de boson de Higgs doublement chargé a été faites et combinée à l'analyse semileptonique complète. L'ensemble de cette

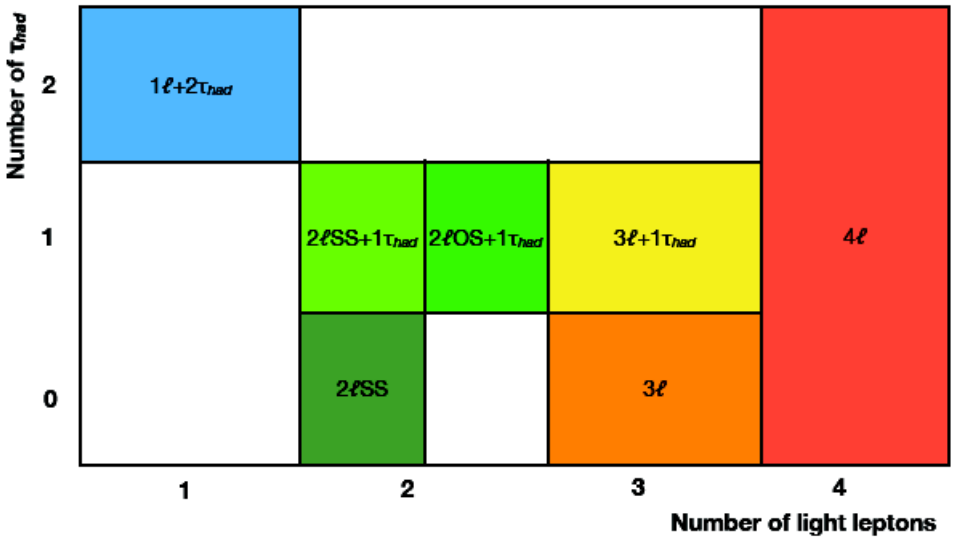


Fig. 2 Les états finaux du canal $t\bar{t}H$ multileptons.

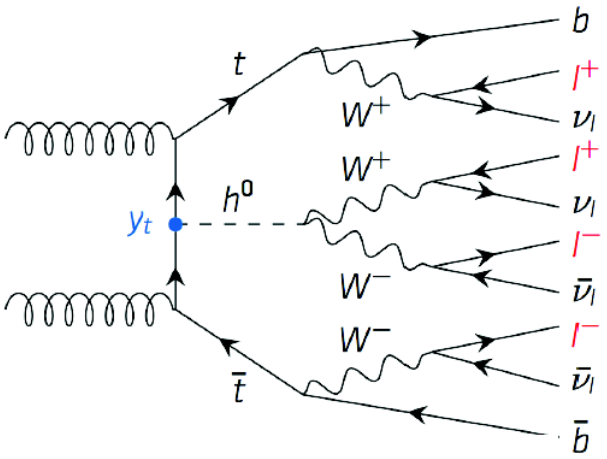


Fig. 3 La topologie des états finaux à 4 leptons.

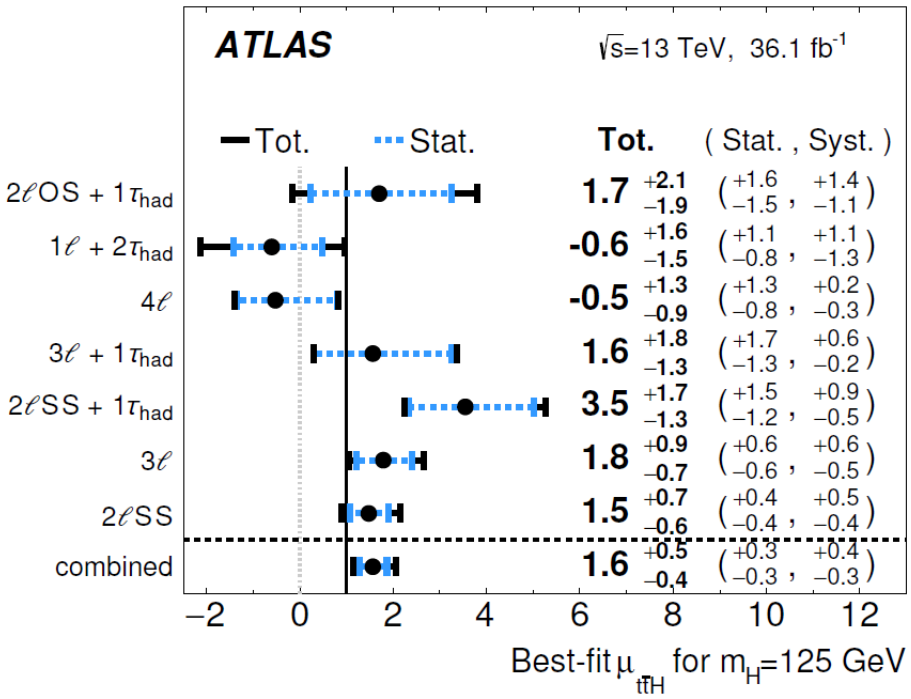


Fig. 4 Les valeurs ajustées observées pour la force du signal $t\bar{t}H$ mu$t\bar{t}H$ et leurs incertitudes par catégorie d'état finals et combinées. La prédiction SM est $\mu = 1$.

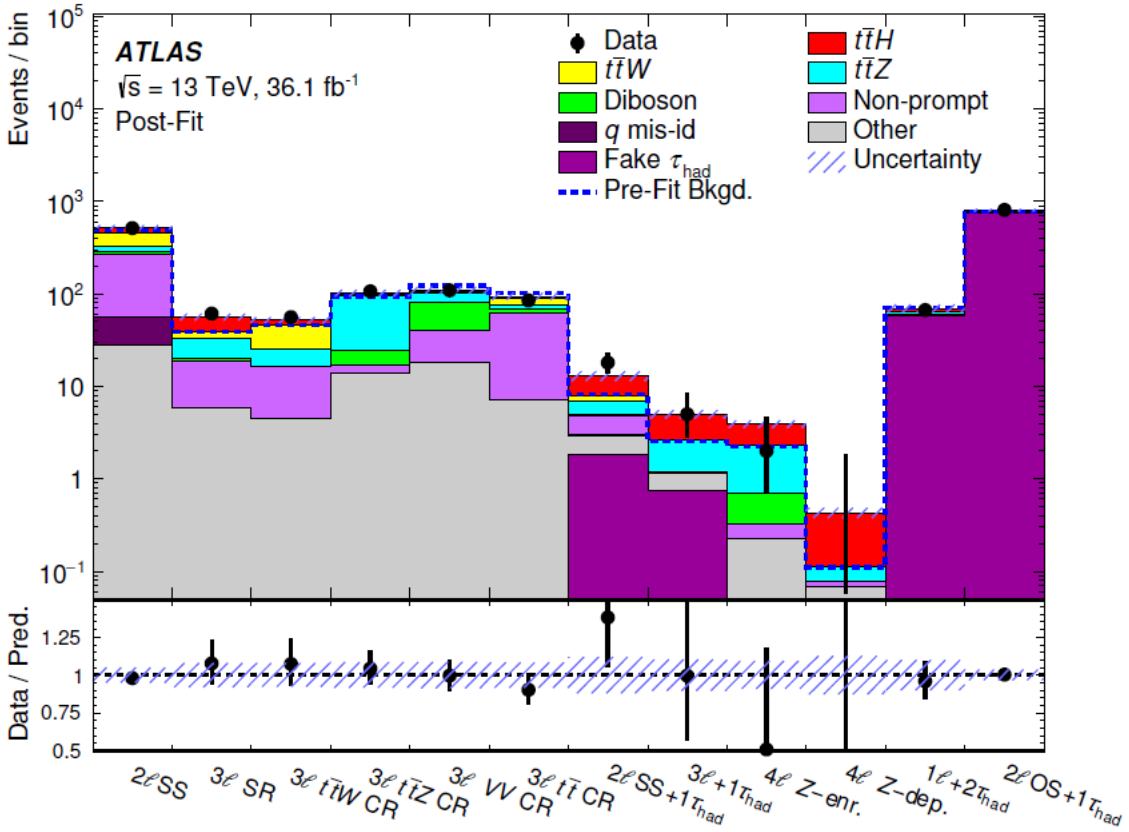


Fig. 5 Comparaison de la prédiction aux données après l'ajustement des huit régions de signal et des quatre régions de contrôle. Les incertitudes systématiques sur les rendements prédits sont indiquées par les bandes bleues hachées.

étude et de ses résultats est présentée au cinquième chapitre.

L'étude se concentre sur la recherche de bosons scalaires doublement chargés $H^{\pm\pm}$, dans le cas d'un couplage aux leptons est faible et où la désintégrations en bosons vecteurs chargés du SM est le canal de désintégration dominant $H^{\pm\pm} \to W^{\pm}W^{\pm}$.

L'addition d'un triplet scalaire est mentionnée dans le contexte des modèles "see-saw", un modèle générique destiné à expliquer les tailles relatives des masses de neutrinos observées (de l'ordre de l'eV), inférieur par rapport aux masses de quarks et de leptons chargés. En dépit de sa beauté théorique, le mécanisme est difficile à tester expérimentalement et contient généralement plus de paramètres que les observables physiques.

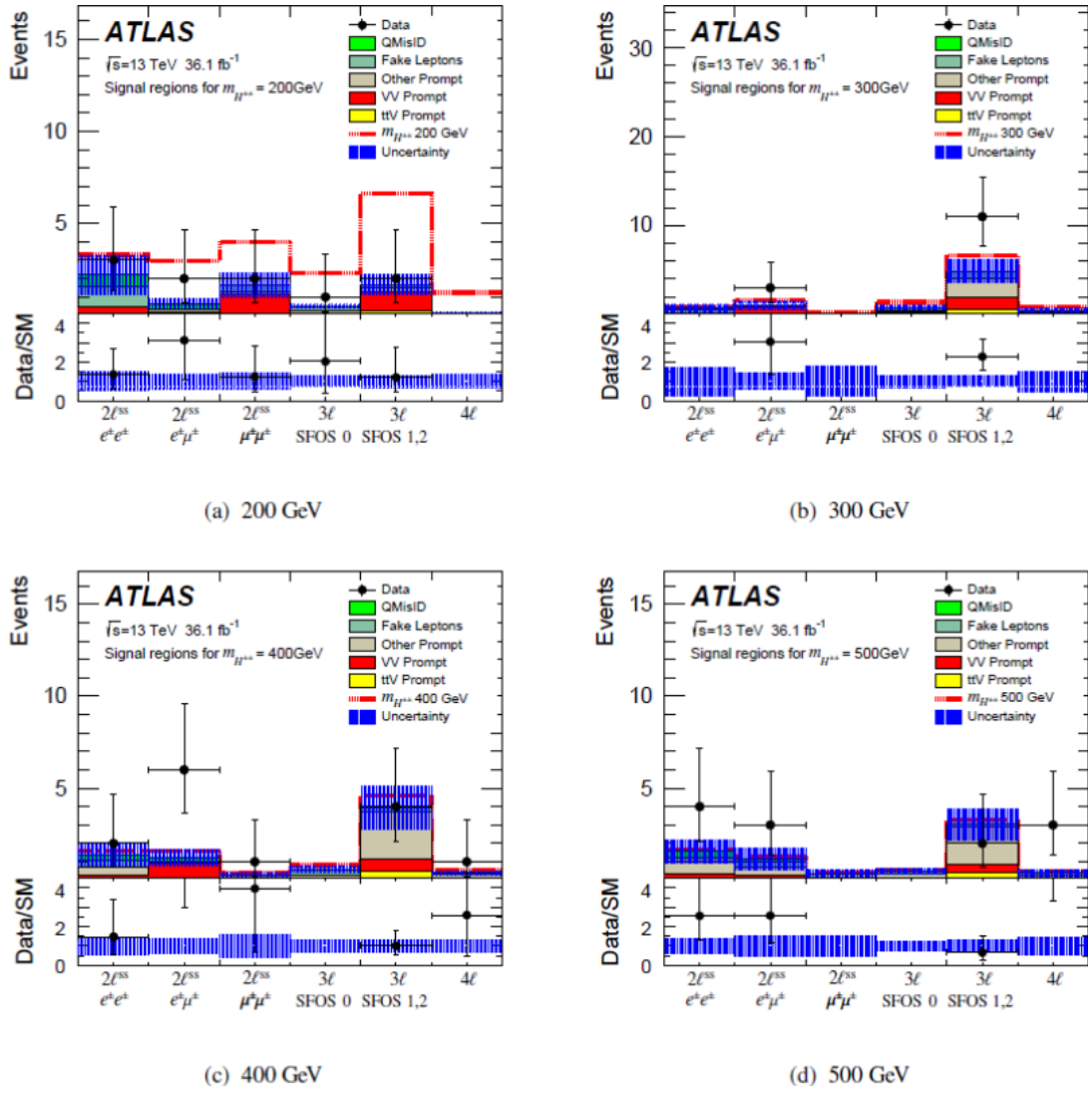
Le modèle utilisé dans cette analyse est basé sur le mécanisme de type 1 qui est une extension du SM, le doublet scalaire SM étant complété par un triplet scalaire hypercharge Y= 2, Δ. Cinq couplages décrivent l'interaction des champs tandis que la rupture de symétrie électrofaible est définie par deux valeurs d'espacement du vide: v_d = 246 GeV et v_t , contraint à la plage sub-GeV. La rupture de symétrie conduit à une phénoménologie riche, comprenant sept scalaires, $H^{\pm\pm}, H^{\pm}, A^0$ (CP-impar), H^0 (CP-par), h^0 (CP-par). Un des scalaires neutres h^0 (scalaire) est identifié au boson Higgs SM.
Les rapports de branchement des différents états finaux contenant des leptons légers (élec-

$H^{\pm\pm}H^{\mp\mp} \to 4W$	**BR (in %)**
$\to$ 0l + 8 jets	20.1
$\to$ 1l + E_T^{miss} + 6 jets	39.7
$\to$ 2l + E_T^{miss} + 4 jets	29.3
$\to$ 3l + E_T^{miss} + 2 jets	9.6
$\to$ 4l + E_T^{miss} + no jets	1.2

Table 1 Production de paires de $H^{\pm\pm}$

trons et muons) sont présentés dans le Tableau 1. Les échantillons de signaux ont été simulés avec $m_{H^{++}}$ = 200, 300, 400, 500, 600 et 700 GeV respectivement.

Une description générale du canal 4l est données, puis des études d'optimisation, la définition des régions de signal et des régions de contrôle, les estimations des bruits de fond de fausse identification et la procédure d'ajustement sont ensuite développées. Ce canal est enfin combiné avec les canaux 2lSS et 3l pour fournir un résultat final. Pour augmenter la signification du signal, les deux autres canaux sont divisés en fonction du nombre de SFOS



(a) 200 GeV

(b) 300 GeV

(c) 400 GeV

(d) 500 GeV

Fig. 6 Rendements obtenus dans tous les canaux à $m_{H^{++}}$ = 200 (en haut à gauche), 300 (en haut à droite), 400 (en bas à gauche) et 500 (en bas à droite) GeV. Les barres d'erreur représentent l'erreur complète (statistique et systématique).

(même saveur, signes opposés). La Figure 6 montre les rendements attendus et observés dans les régions de signal pour les points de masse de 200 à 500 GeV. Les barres d'erreur représentent l'erreur complète (statistique et systématique). Aucun excès significatif n'a été observé dans les données. Par conséquent, nous pouvons fixer une limite au modèle considéré.

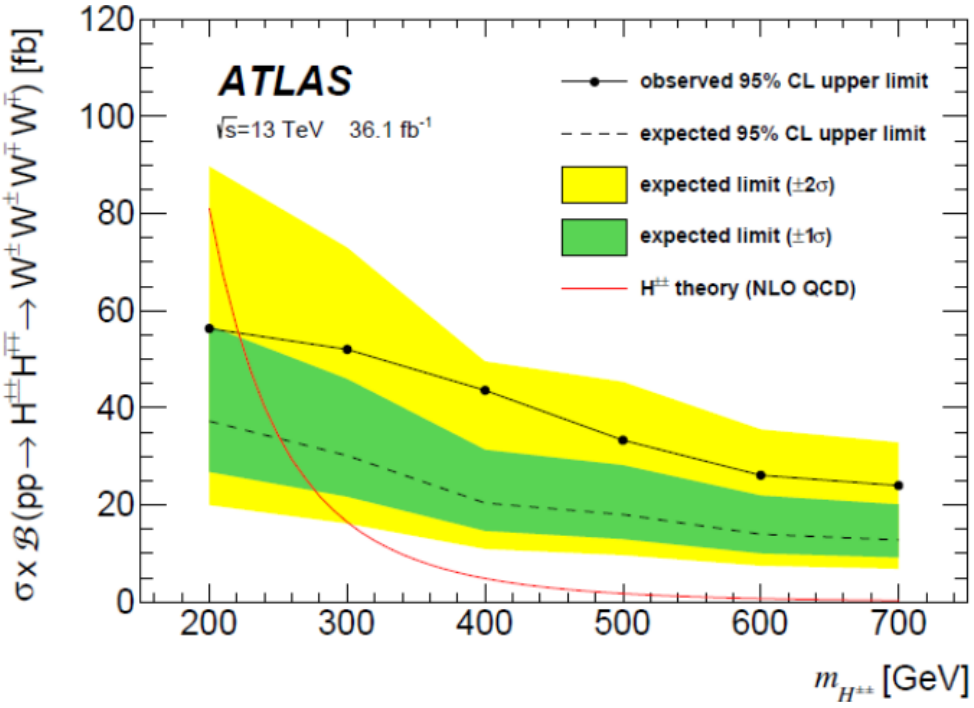


Fig. 7 Limites prévues et observées pour la combinaison des canaux 2LSS, 3L et 4L.

Les significations attendues et observées en fonction de la masse de H^{++} sont présentées dans la Figure 7. En combinant tous les 3 canaux, la valeur attendue dépasse 3 écarts-types pour le point de masse de 200 GeV. Cependant, aucun excès significatif n'est observé.
Les limites attendues et observées de la section transversale fois la fraction de ramification à 95% C.L. La combinaison des canaux 2LSS, 3L et 4L est illustrée à la Figure 8. La théorie est représentée par la ligne rouge, tandis que l'exclusion à 2 σ et 1 σ est représentée en jaune et en vert. Aucun excès significatif n'est observé dans les données. En combinant ces canaux, le modèle est exclu à 95% C.L. pour des masses comprises entre 200 et 220 GeV.

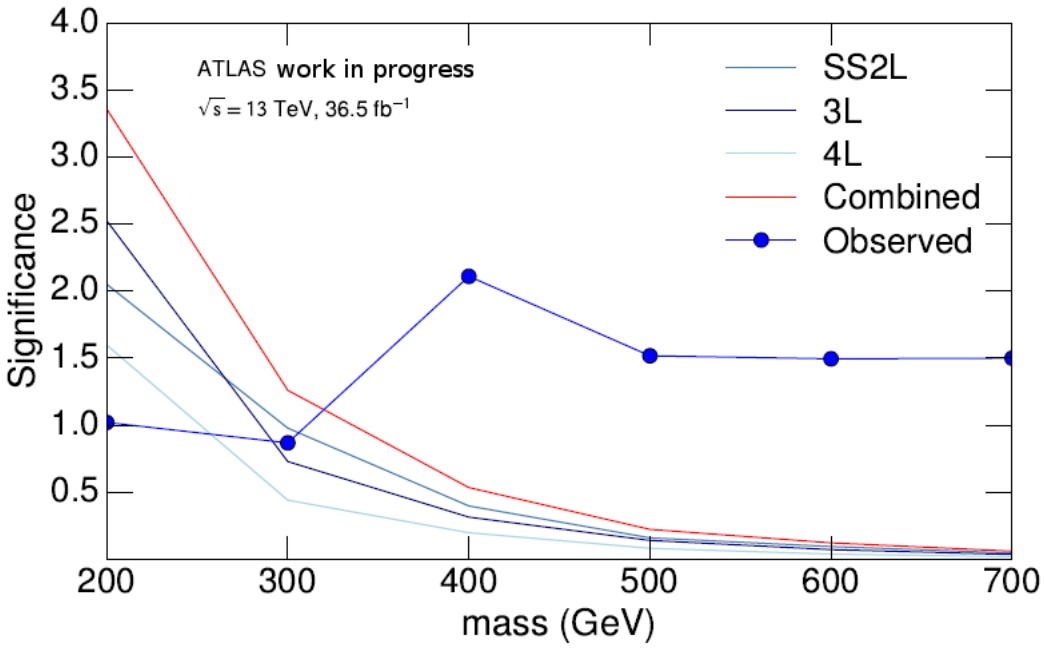


Fig. 8 Significations attendues et observées en fonction de la masse de H^{++} pour les 3 canaux.

Rezumat

Teza mea de doctorat are loc in cadrul colaborarii ATLAS, operand detectorul cu acelasi nume, in cadrul experimentului Large Hadron Collider (LHC).

ATLAS este unul dintre cei doi detectori de uz general impreuna cu CMS la Large Hadron Collider (LHC), utilizat pentru a cauta semne de fizica noua, inclusiv raspunsuri la intrebari precum originile dimensiunilor de masa etc.

La data de 4 iulie 2012, experimentele ATLAS si CMS au raportat ca au observat o noua particula in regiunea de masa in jur de 126 GeV, in conformitate cu bosonul Higgs furnizat de modelul standard. Modelul standard al fizicii particulelor este una dintre cele mai de semnificative teorii in fizica, care a fost testata cu succes pana acum, desi unele intrebari raman fara raspuns. Dintre acestea pot fi enumerate ierarhia maselor, materia intunecata si energia intunecata, prezenta gravitonului etc. Modelul Standard descrie interactiunile fundamentale dintre particulele elementare si formalismul de achizitie de masa prin interactiunea cu campul Higgs.

Dupa prezentarea cadrului matematic al modelului standard, voi continua cu o descriere a experimentului LHC la CERN, acordand o atentie deosebita detectorului ATLAS, definind astfel partea experimentala a cercetarii mele.

Voi oferi o descriere completa a detectorului ATLAS si a selectiei obiectelor, cu accent pe sistemul calorimetric, care are o importanta deosebita in activitatea mea privind performanta detectorului (studiul eficientei identificarii electronilor utilizand metoda W Tag and Probe).

In experimentul ATLAS, pentru a fi autor oficial, trebuie realizata o perioada de calificare in studii privind performanta detectorului. Run 2 ofera mai multe provocari experimentale datorita energiei crescute si a numarului de coliziuni de particule, asadar, masuratorile de performanta sunt esentiale pentru a fi imbunatatite si aliniate cu procedurile de actualizare a detectorului. Pentru asta, am fost si continu sa fiu principalul analist al identificarii electronilor folosind W dezintegrandu-se in electroni si neutri. In plus, sunt, de asemenea, implicata in prelucrarea datelor ce vor fi luate in perioada 2017-2018 prin dezvoltarea unei noi configuratii de triggeri pentru analiza proceselor $W \to e\nu$, fiind in stransa colaborare cu personalul de la CERN si experti.

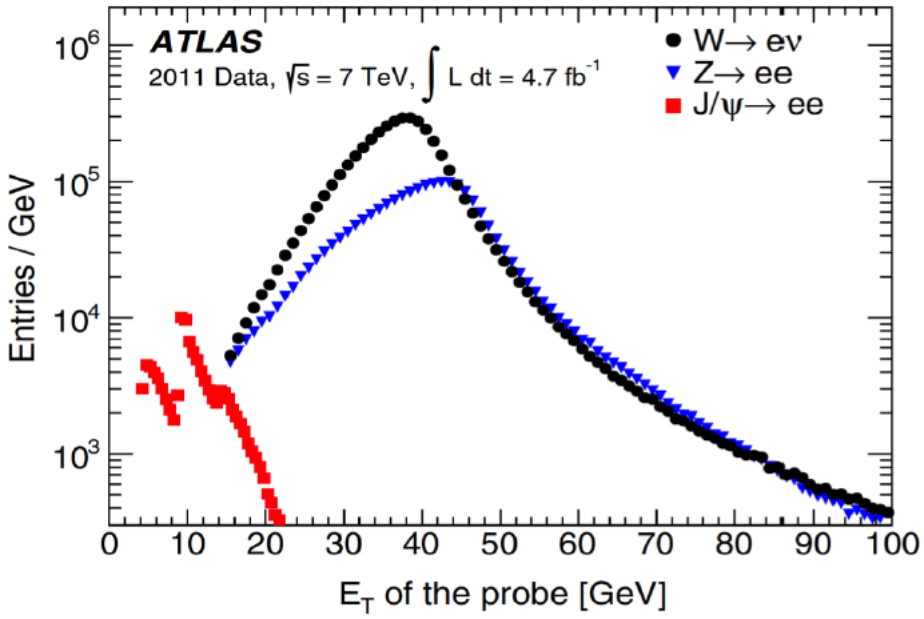


Fig. 9 Distributia energiei transversale a electronilor candidati care indeplinesc criteriul de identificare Tight in $W \to e\nu$ (cercuri complete), $Z \to ee$ (triunghiuri albastre) si $J/\Psi \to ee$ (patrate rosii) utilizand setul de date 2011.

Setul de date colectat in 2015-2017 de catre ATLAS a fost utilizat pentru masurarea eficientei identificarii electronilor utilizand metoda numita "Tag and Probe". Algoritmii de identificare a leptonilor se bazeaza pe un discriminator de probabilitate multivariat. Diferentele dintre date si generarile simulate Monte-Carlo (MC) au fost integrate in rapoartele SF = data / MC (factorii de scala), obtinute in functie de energia transversala si de pseudorapiditatea electronilor.

Am prezentat in aceasta teza eficienta identificarii electronilor folosind dezintegrarea $W \to e\nu$. Numeroase avantaje pentru a folosi evenimentele $W \to e\nu$ pentru a masura eficienta identificarii electronilor pot fi precizate: bozonele W si Z acopera o gama larga de cinematica (asa cum se poate vedea in Figura 9 dar evenimentele $W \to e\nu$ pot contribui la intelegerea incertitudinii sistematice de la Z. In plus, W ofera statistica suplimentara pentru gama de energii joase (15$< E_T <$25 GeV) care poate completa canalul J/Ψ. De asemenea ofera semnaturi experimentale interesante din cauza statisticii sale extraordinare.

Principala dificultate a acestui canal de dezintegrare a fost zgomotul de fond. Pentru a avea un semnal adecvat, s-au efectuat studii de optimizare folosind metode bazate pe algoritmul "cut-based" si pe tehnici multivariate TMVA. In plus, pentru perioada 2017-2018, a fost

propus si validat un nou set de configuratii de triggeri, facand ca acest canal $W \to e\nu$ sa fie interesant pentru furnizarea unei verificari a factorilor de scara obtinuti folosind evenimentele $Z \to ee$.

S-au prezentat factorii de scara folosind W T&P precum si o comparatie cu rezultatele $Z \to ee$ T&P, observandu-se un acord de 10% intre W si Ziso. Aceasta reprezinta prima masurare a datelor Run 2 a eficientei de indentificare a electronilor utilizand evenimentele $W \to e\nu$.

Munca mea a fost valoroasa pentru colaborarea ATLAS, datorita procedurilor inovative: studii de optimizare si de fond, prima masurare a factorilor de eficienta si la scara folosind $W \to e\nu$ in datele Run 2, studile de emulari si configurare a unui noi set de triggeri ce vor fi utilizati peentru setul de date colectate in perioada 2017-2018.

Unul din principalele obiective ale Large Hadron Collider a fost descoperirea bosonului Higgs, o realizare majora a Run I. Dupa descoperirea unui boson Higgs de 125 GeV, un efort major este dedicat in Run II, in masurarea proprietatilor sale, in special cuplajului constituientilor materiei cu boson Higgs. Deoarece cuarcul top este cea mai grea particula elementara cunoscuta, ea are cel mai mare cuplaj cu bosonul Higgs. Pentru a evalua in mod direct amploarea acestui cuplare la experimentul LHC, se poate analiza productia asociata a bosonului Higgs cu o pereche de cuarci top ($t\bar{t}H$).

Aceasta teza prezinta studiul productiei $t\bar{t}$ H in stare finala multileptonica, bazat pe datele colectate in 2015 si 2016 la o luminozitate integrala de $36.1 lafb^{-1}$ de catre detectorul ATLAS la LHC, la o energie a centrului de masa de $\sqrt{s}$ = 13 TeV.

S-au examinat sapte stari finale (asa cum se poate vedea in Figura 10), care vizeaza bosonul Higgs dezintegrand in WW *, $\tau\tau$ si ZZ *, in cadrul teoriei modelului standard (SM).

Analiza multileptonilor $t\bar{t}H \to$ este apoi combinata cu alte stari finale $t\bar{t}H$ pentru a furniza un rezultat final al semnalului:

- Multilepton
- $H \to b\bar{b}$
- $H \to \gamma\gamma$
- $H \to 4l$

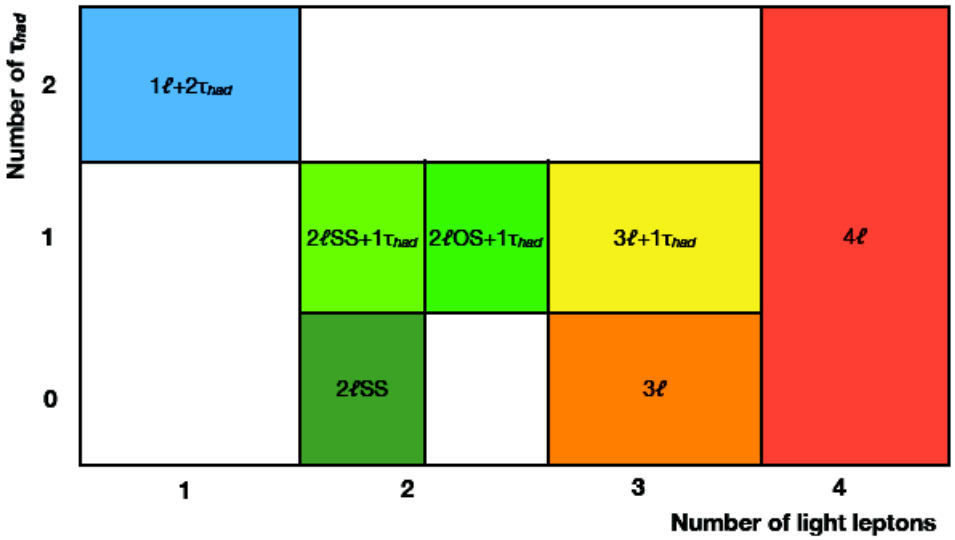


Fig. 10 Canalul $t\bar{t}$ H dezintegrand in starea finala de multileptoni.

Rezultatul combinat ofera cea mai buna valoare a puterii semnalului este $\mu_{t\bar{t}H} = 1,17 \pm 0,19(stat)^{+0.27}_{-0.23}$ cu o valoare corespunzatoare a sectiunii transversale de $\sigma_{\bar{t}H} = 590^{+160}_{150} fb$ si semnificatie de 4.2σ la o valoare estimata de 3.8σ.

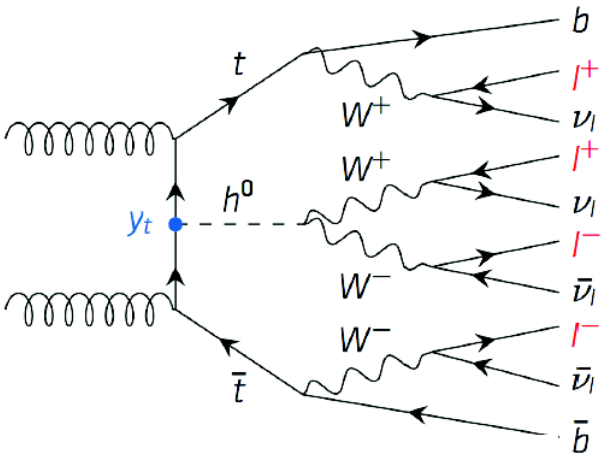


Fig. 11 Topologia starii finale de 4 leptoni.

In aceasta teza, m-am axat pe canalul 41 (Figura 11), prezentand toti pasii procedurii de analiza: studii de optimizare, definirea regiunilor de semnal si a regiunilor de control, estimarea candidatilor "falsi" si procedura de fit.

Se realizeaza o "potrivire a probabilitatii maxime" pe toate canalele multilepton, inclusiv a canalulului $4l$, simultan, pentru a extrage puterea semnalului, presupunand o masa a bosonului Higgs de 125 GeV. Puterea semnalului este definita ca raportul dintre randamentul semnalului masurat $t\bar{t}H$ si asteptarile modelului standard. In cazul unei semnaturi a productiei de cuarci top asociata cu bosonul SM Higgs atunci $\mu = 1$, dar daca acesta este absent, atunci $\mu = 0$.

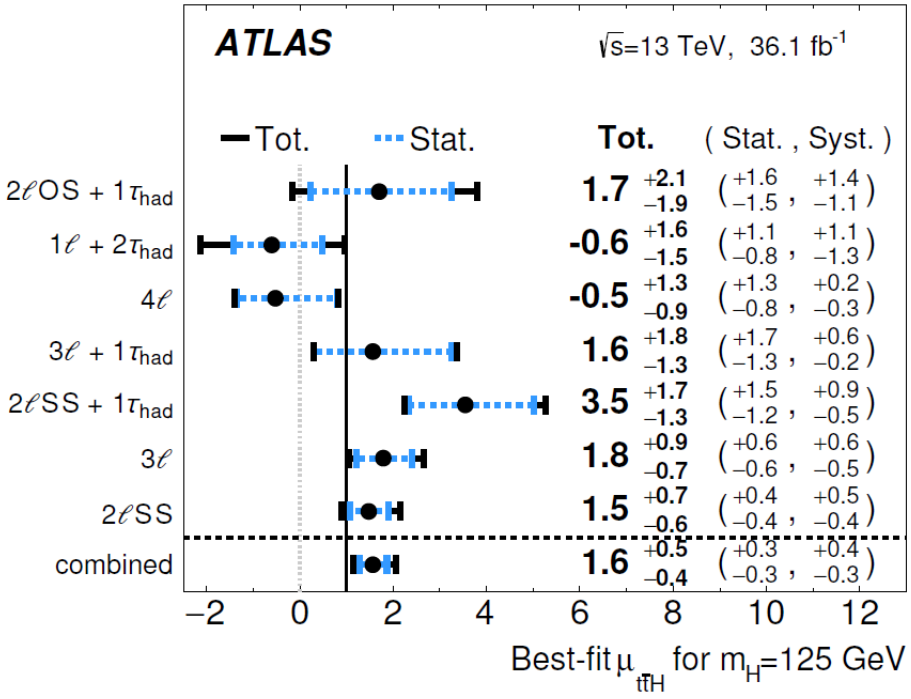


Fig. 12 Puterea semnalului in productia $t\bar{t}H$ $\mu_{t\bar{t}H}$ si incertitudinile aferente.

Cea mai buna valoare observata de $\mu_{t\bar{t}H}$, care combina toate canalele, este $1.6^{0.3}_{+0.3}$ (stat) $^{+0.4}_{-0.3}$ (syst), dupa cum se arata in Figura 12 pentru fiecare categorie individuala de stare finala si combinatia tuturor canalelor. Semnificatia observata (asteptata) in ceea ce priveste ipoteza no- $t\bar{t}H$ este 4,1 σ (2,8 σ), valoare ce corespunde unei sectiuni eficace de σ ($t\bar{t}H$) = 790 $^{150}_{-150}$ (stat) $^{+170}_{-150}$ (syst) fb. Sectiunea transversala teoretica asteptata este σ ($t\bar{t}H$) = 507 $^{+35}_{-50}$ fb.

Figura 13 prezinta o comparatie in cele opt regiuni de semnal si patru regiuni post-fit de control a datelor cu evenimentele simulate.

Am fost, de asemenea, analistul responsabil de canalul $4l$ in articol ce prezinta studiul dezintegrarii bosonului Higgs dublu incarcat in $W^{\pm}W^{\pm}$ care se concentreaza asupra fenomenolo-

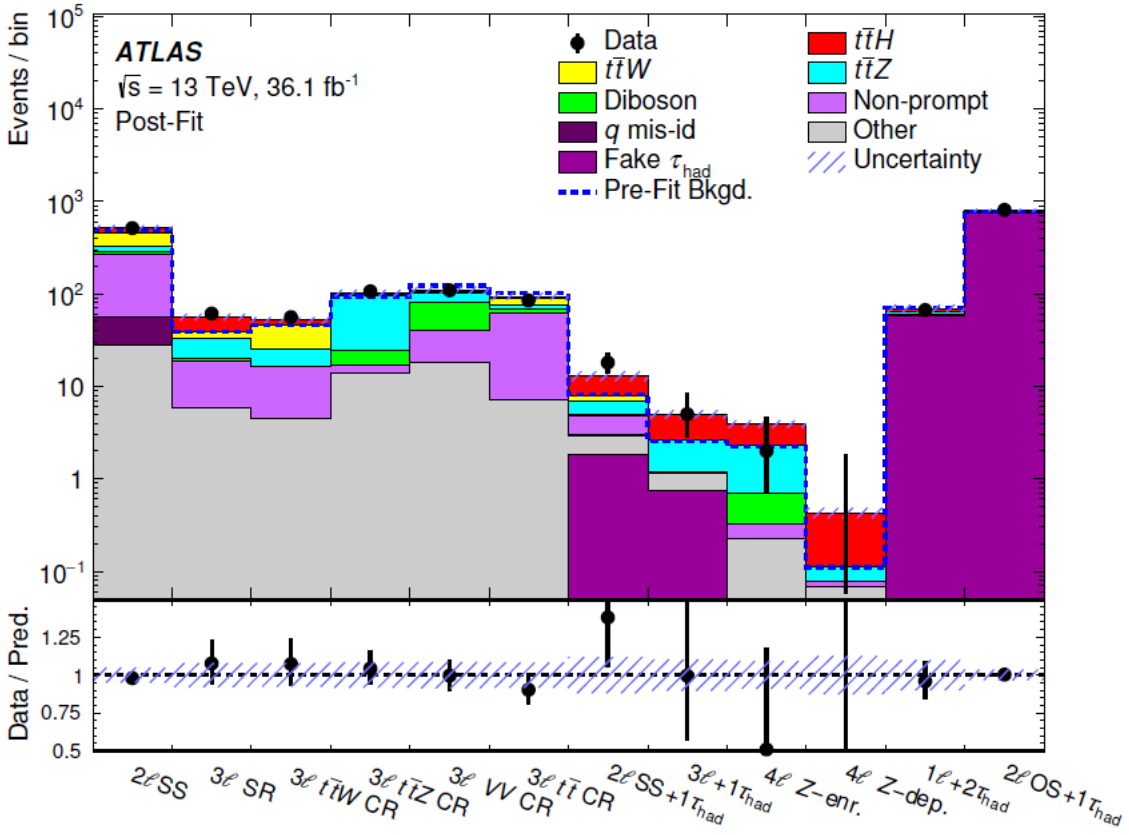


Fig. 13 Comparatia intre date si evenimentele generate in configuratia "post-fit" pentru toate regiunile considerate in analiza.

giei bosonilor $H^{\pm\pm}$, in cazul in care magnitudinea cuplajului cu leptoni este scazuta si dezintegreaza bozonilor vector este dominant de canalul $H^{\pm\pm} \to W^{\pm}W^{\pm}$.

Adaugarea unui triplet scalar este mentionata in contextul modelelor „see-saw“, un model generic pentru a explica dimensiunile relative ale maselor de neutrini observate (de ordinul eV), mai mici comparativ cu masele cuarcurile si leptonilor. In ciuda frumusetii sale teoretice, mecanismul este dificil de testat experimental si de obicei contine mai multi parametri decat observabile fizice.

Modelul utilizat in analiza mea se bazeaza pe tipul de mecanism „see-saw“ 1, care este o extensie a modelului standard, dubletul scalar fiind completat de un triplet hypercharge scalar cu Y = 2, Δ. Cinci cuplaje descriu interactiunea dintre campuri, in timp ce simetria electroslaba este definita prin doua valori ale parametrilor de vid: v_d = 246 GeV si v_t, constransa la gama sub-GeV. Simetria de rupere conduce la o fenomenologie bogata care cuprinde sapte scalari, $H^{\pm\pm}, H^{\pm}, A^0$ (CP-impar), H^0 (CP-par), h^0 (CP-par). Unul dintre scalarii neutru h^0 (scalar) este identificat cu bosonul Higgs definit de modelul standard.

$H^{\pm\pm}H^{\mp\mp} \to 4W$	**BR (in %)**
$\to$ 0l + 8 jets	20.1
$\to$ 1l + E_T^{miss} + 6 jets	39.7
$\to$ 2l + E_T^{miss} + 4 jets	29.3
$\to$ 3l + E_T^{miss} + 2 jets	9.6
$\to$ 4l + E_T^{miss} + no jets	1.2

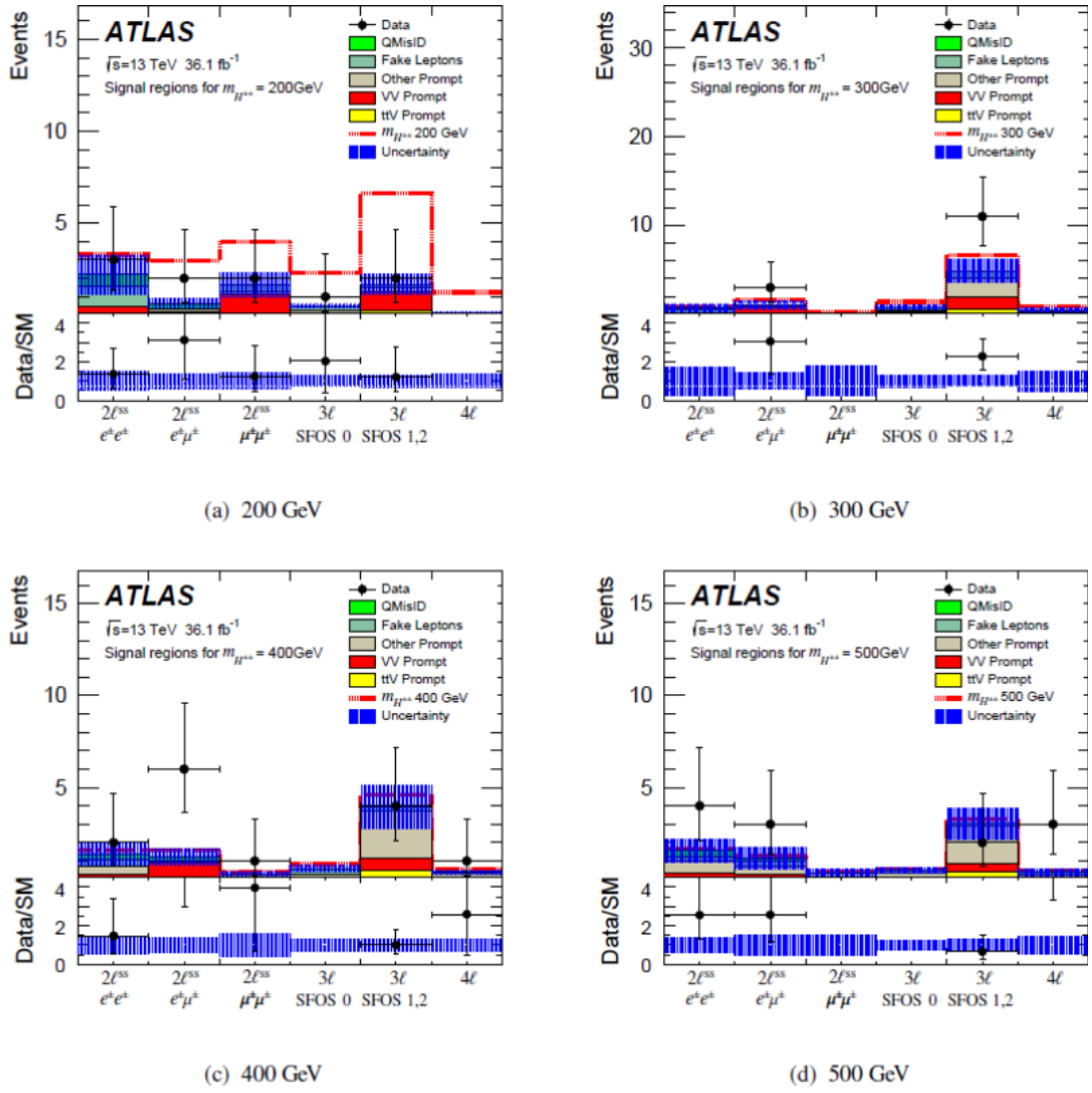
Table 2 Canalele de producere a $H^{\pm\pm}$

Ratele de productie ale diferitelor stari finale care contin leptoni in starea finala (electroni si muoni) sunt prezentate in Tabelul 2.

Probele de semnal au fost simulate cu $m_{H^{++}}$ = 200, 300, 400, 500, 600 si respectiv 700 GeV.

In aceasta teza, am dat o descriere generala a canalului 4*l* pentru care eram principalul conducator. Este un canal combinat cu canale de stare finala cu 2*l* si 3*l* pentru a oferi un rezultat final.

Voi prezinta analiza completa efectuata pe canalul de $4l$, si anume: studii de optimizare, definirea regiunilor de semnal si regiuni de control, procedura de estimare si ajustare a candidatilor falsi.



(a) 200 GeV

(b) 300 GeV

(c) 400 GeV

(d) 500 GeV

Fig. 14 Randamentele obtinute in toate canalele pentru masele bozonului dublu incarcat de m_{H++} = 200 (stanga sus), 300 (dreapta sus), 400 (inferior stanga) si 500 (in dreapta jos) GeV. Barele de eroare reprezinta eroarea completa (statistica si sistematics).

Canalul de $4l$ este combinat cu canalele $2lSS$ si $3l$ pentru a furniza rezultatul final. Pentru a mari semnificatia semnalului, celelalte doua canale sunt impartite in functie de SFOS (pereche de leptoni de aceeasi aroma dar cu sarcina opusa). Figura 14 prezinta randamentele asteptate si observate in regiunile de semnal pentru punctele de masa de la 200 la 500 GeV.

Barele de eroare reprezinta eroarea completa (statistica si sistematica). Nu a fost observat un exces semnificativ in date. Prin urmare, putem stabili o limita a modelului considerat.

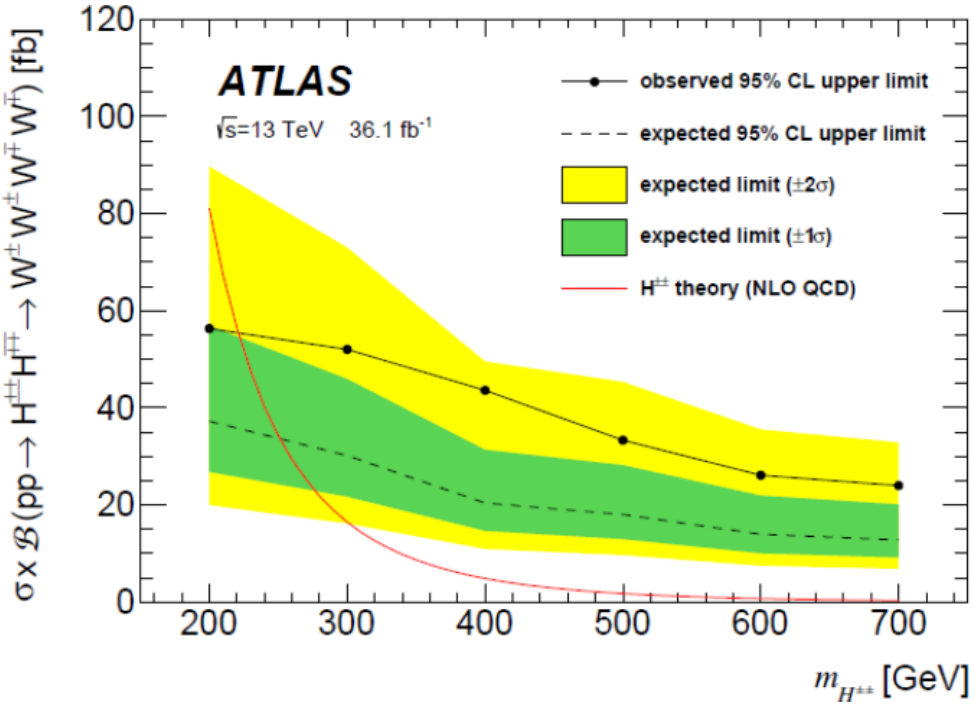


Fig. 15 Limitele asteptate si observate pentru combinarea canalelor 2LSS, 3L si 4L.

Semnificatia semnalului asteptata si observata in functie de masa lui H^{++} este prezentata in Figura 16. Prin combinarea tuturor celor 3 canale, valoarea asteptata depaseste 3 deviatii standard pentru punctul de masa 200 GeV. Cu toate acestea, nu se observa un exces semnificativ.

Limitele asteptate si observate ale sectiunii transversale pentru un nivel de incredere de 95% este ilustrata in Figura 15. Teoria este reprezentata de linia rosie, in timp ce excluderea la 2 σ si 1 σ este afisata in galben si verde. Nu se observa nici un exces semnificativ in date. Prin combinarea acestor canale, modelul este exclus la un nivel de incredere de 95% pentru mase intre 200 si 220 GeV.

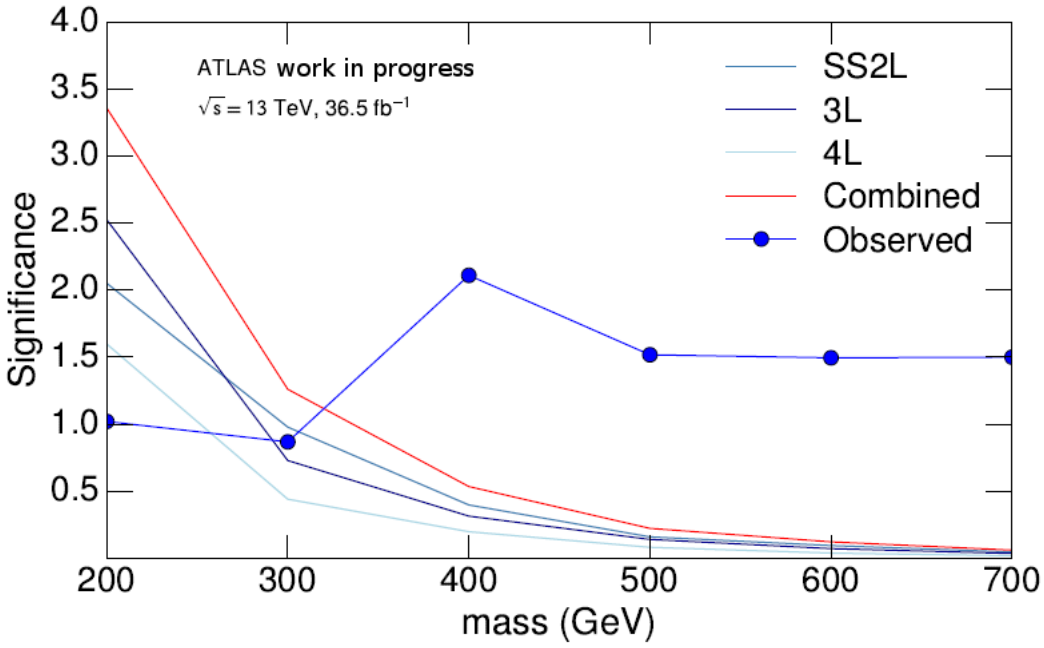


Fig. 16 Semnificatia semnalului asteptat si observat in functie de masa lui H^{++} pentru cele 3 canale combinate.

Acknowledgements

I want to take this opportunity to express my gratitude to the Ph.D. supervisors: Cristinel Diaconu, Calin Alexa, and Emmanuel Monnier. Thank you for all your help, your support, patience, motivation, and immense knowledge, for managing to correct me and guide me to become a better scientist and a better person.

These three years in Marseille have been one of the hardest but most fulfilling periods of my life. I have learned how to set a goal and do anything to achieve it. My goal is physics, and physics is what defines me. For this, I want to thank all the people who helped me all the way through this career starting from my high-school physics professor, to the Faculty of Physics professors who lead me all the way. They are ones of the greatest and inspiring people I have met.

Special "Va multumesc enorm!" for Virgil Baran, my professor of theoretical particle physics who made me choose this field a few years ago when I was in Bachelor degree. Special "Va multumesc enorm!" to Daniela Dragoman, who believed in me when I went to her as a first year Bachelor student and I told her I want to do more than just go to courses. Any many, many thanks to all my professors, I was honored to be your student, and I bow in front of you thanking you for putting a brick in what I am now.

Special "Va multumesc enorm,enorm,enorm!" to my Romanian Ph.D. supervisor and the head of Romanian ATLAS team, Calin. I know I am not an easy person to handle, but you trusted me to be a part of your team. I will always be grateful for the opportunity you gave me to be part of the ATLAS team and I can promise that I am doing my best not to disappoint you. You are the best supervisor/boss I could have, and I admire you for that!

Special "Merci beaucoup!" to my CPPM Ph.D. supervisors, Emmanuel and Cristi, who made my life more comfortable and more pleasant at CPPM. You are the best supervisors I could get and models for me to follow professionally as well as personally.

Besides my supervisors, I would like to thank the rest of my thesis committee: Dr. Anne-catherine LE BIHAN and Dr. Jean Baptiste DE VIVIE DE REGIE, for their insightful comments and encouragement.

Many thanks,multumesc, merci to all my colleagues (Grigore, my twins Valentina and Alexandra, Kevin, Venu and many more) for your help, support and nice times outside the office. Many thanks to all my friends who offered me a great time during all these years (Mada Miloi, Age for our poker games, Ionut etc). Special gratefulness to Max for offering me equilibrium, for being patient and accepting all me.

In the end thanks to my family for supporting me and accepting that my career comes first :).

Table of contents

List of figures

List of tables

Chapter 1

Standard Model

1.1 Introduction to Standard Model (SM)

The history of the Standard Model came as an interplay between some “good ideas” and some “misunderstandings” [38]. However, as optimists may say: "Failure is the key to success; each mistake teaches us something".

In 1964, the idea of the quark model was proposed, independently by Gell-Mann and Zweig. The hypothesis that hadrons are made out of three quarks (baryons), and quark-antiquarks (mesons) allowed one to understand their quantum numbers and mass spectrum in terms of an approximate $SU(3)$ flavor symmetry.

In 1954, the invention of non-Abelian gauge theories by Yang and Mills led to the identification of the V-A structure of the weak interactions by Bludman, Glashow, Salam and Ward and others developed theories of the weak interactions with intermediate vector bosons. One feature of the non-Abelian gauge theories is the mechanism of breaking the local gauge symmetry. Consequently, theorists developed the idea of spontaneous symmetry breaking: There can be symmetries of the Lagrangian that are not symmetries of the vacuum. According to the Goldstone theorem, there must be a massless spinless particle for every spontaneously broken global symmetry.

In 1964 Higgs and Englert and Brout found a way to circumvent Goldstone’s theorem: The theorem does not apply if the symmetry is a gauge symmetry. In this case, the Goldstone boson becomes the helicity-zero part of the gauge boson, which thereby acquires a mass.

In 1967 Weinberg applied the idea of spontaneous symmetry breaking to the weak interactions of the leptons of the first family, (ν_L, e_L) and e_R, finding the bosons associated to the gauge group $SU(2) \times U(1)$ and the symmetry breaking: the massive W and Z bosons, a massless photon and the Higgs boson.

Since 1973 many important experiments proved that the Standard Model is, up to present days, the correct theory of elementary particles:

- 1973: discovery of neutral currents [21]
- 1979: discovery of the gluon at DESY by TASSO and the other experiments at the PETRA collider [22]
- 1983: discovery of the W and Z bosons at CERN SppS collider, at $\sqrt{s}$=540 GeV, UA-1/2 experiments [23]
- 1975 - 2000: discovery of the third family, τ , b, t and ν_τ [24]
- 4 July 2012: the ATLAS and CMS experiments at CERN's Large Hadron Collider announced they had each observed a new particle in the mass region around 126 GeV, consistent with the Higgs boson predicted by the Standard Model. [25] [26]

The Standard Model of particle physics is one of the most successful theory in physics which has been extensively tested experimentally with great success up to now even though some unanswered questions remain such as mass hierarchy, dark matter and dark energy, inclusion of gravity, etc. This model describes the fundamental interactions between elementary particles and the formalism of mass acquirement of the particles through the Higgs field interaction [10].

The main principles which paved the way for the SM construction are:

- The generalized correspondence to various existing theories and models like Quantum Mechanics, QED, the Fermi model, etc.;
- QCD describes successfully the strong interactions;
- The weak and electromagnetic forces are "unified" in the so-called electroweak interaction;
- The renormalizability is necessary for the derivation of finite predictions for observable quantities at the quantum level;
- the gauge principle for the introduction of interactions.

The SM possesses several different symmetries:

- the Lorentz (and Poincaré) symmetry,
- the CPT symmetry,
- three gauge symmetries $SU(3)_C \times SU(2)_L \times U(1)_Y$,
- the global $SU(2)_L \times SU(2)_R$ symmetry in the Higgs sector,
- some other symmetries, etc.

In the following sections, I will describe the theory of electroweak symmetry breaking through global and local symmetries as well as the mathematical description of the mechanism of giving mass to known particles. The content of this chapter is based on "An introduction to the Standard Model of particle physics", W. N. Cottingham and D. A. Greenwood, Cambridge University Press, 2007 [10].

1.2 Global symmetry breaking and Goldstone bosons

We start from a complex scalar field of the form:

$$\Phi = (\phi_1 + i\phi_2)/\sqrt{2} \tag{1.1}$$

satisfying the Klein–Gordon equation (See Annex A.3). If the field Φ carries charge q, then the field Φ^* carries charge $-q$. The Klein–Gordon equation for a complex field Φ is obtained from the (real) Lagrangian density:

$$\mathscr{L} = \partial_\mu \Phi^\dagger \partial^\mu \Phi - m^2 \Phi^\dagger \Phi \tag{1.2}$$

If Φ is constant and independent of space and time, the only contribution to the energy is the term $m^2\Phi^\dagger\Phi$. The minimum of this expression is when $\Phi_1 = \Phi_2 = 0$ (since m^2>0), corresponding to the vacuum state.

If we take now a Lagrangian density of the form:

$$\mathscr{L} = \partial_\mu \Phi^\dagger \partial^\mu \Phi + m^2 \Phi^\dagger \Phi \tag{1.3}$$

this would be unstable. Stability can be restored by introducing a new term, taken in the form of $m^2\phi_0^2/2(\phi^\dagger\phi)^2$, with ϕ_0^2 another real parameter, so that now the Lagrangian density can be written as:

$$\mathscr{L} = \partial_\mu \Phi^\dagger \partial^\mu \Phi - V(\Phi^\dagger \Phi) \tag{1.4}$$

where

$$V(\Phi^\dagger \Phi) = \frac{m^2}{2\phi_0^2}[\Phi^\dagger \Phi - \phi_0^2]^2 \tag{1.5}$$

The minimum field energy is now obtained with Φ constant independent of space and time, but such that $\Phi^\dagger\Phi = |\Phi|^2 = \phi_0^2$, with an infinite possible vacuum states. In addition, in this case the Lagrangian density has a global U(1) symmetry: $\Phi \to \Phi' = e^{-i\alpha}\Phi$ for any real α.

Expanding about the ground state $(\phi_0, 0)$, and put $\Phi = \phi_0 + (1/\sqrt{2})(\chi + i\Psi)$, the Lagrangian density becomes:

$$\mathscr{L} = \frac{1}{2}\partial_\mu \chi \partial^\mu \chi + \frac{1}{2}\partial_\mu \Psi \partial^\mu \Psi - \frac{m^2}{2\phi_0^2}[\sqrt{2}\phi_0\chi + \frac{\chi^2}{2} + \frac{\Psi^2}{2}]^2 \tag{1.6}$$

This Lagrangian density contains now a free term of form:

$$\mathscr{L}_{free} = \frac{1}{2}\partial_\mu \chi \partial^\mu \chi + \frac{1}{2}\partial_\mu \Psi \partial^\mu \Psi - m^2\chi^2 \tag{1.7}$$

which represents free particle fields, and contains all the terms in $\mathscr{L}$ that are quadratic in the fields.

The rest of the Lagrangian density, denoted $\mathscr{L}_{int}$, corresponds to interactions between the free particles including the higher order corrections to their motion.

The quadratic term $-m^2\chi^2$ corresponds to a scalar spin-zero particle of mass $\sqrt{2}m$. However, in the case of the Ψ field, there is no such quadratic term: the corresponding scalar spin-zero particle is therefore massless. The massless particles that always arise as a result of global symmetry breaking are called **Goldstone bosons**.

1.3 Local symmetry breaking and the Higgs boson

We now generalise further, and construct a Lagrangian density that is invariant under a local $U(1)$ gauge transformation:

$$\Phi \to \Phi' = e^{iq\theta}\Phi \tag{1.8}$$

with $\theta = \theta(x)$ space and time dependent.

This further requires the introduction of a (massless) gauge field A_μ:

$$\mathscr{L} = [(\partial_\mu - iqA_\mu)\Phi^\dagger][(\partial^\mu + iqA^\mu)\Phi] - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - V(\Phi^\dagger\Phi) \tag{1.9}$$

with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and

$$V(\Phi^\dagger\Phi) = \frac{m^2}{2\phi_0^2}[\Phi^\dagger\Phi - \phi_0^2]^2 \tag{1.10}$$

We can observe now that the Lagrangian density is invariant under the local gauge transformation:

$$\begin{aligned} \Phi(x) &\to \Phi'(x) = e^{iq\theta}\Phi(x) \\ A_\mu(x) &\to A'_\mu(x) = A_\mu(x) + \partial_\mu\theta(x) \end{aligned} \tag{1.11}$$

Given $\Phi(x)$, we can always choose $\theta(x)$ so that the field $e^{iq\theta}\Phi(x)$ is real. This breaks the symmetry since we are no longer free to make further gauge transformations.

Putting $\Phi'(x) = \phi_0 + h(x)/\sqrt{2}$ with $h(x)$ real, we can further rewrite the Lagrangian density in terms of free and interaction terms:

$$\mathscr{L}_{free} = \frac{1}{2}\partial_\mu h \partial^\mu h - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - m^2h^2 + q^2\phi_0^2 A_\mu A^\mu \tag{1.12}$$

$$\mathscr{L}_{int} = q^2 A_\mu A^\mu(\sqrt{2}\phi_0 h + \frac{1}{2}h^2) - \frac{m^2h^2}{2\phi_0^2}(\sqrt{2}\phi_0 h + \frac{1}{4}h^2) \tag{1.13}$$

In $\mathscr{L}_{free}$ we have a single scalar field $h(x)$ corresponding to a spinless boson of mass $\sqrt{2}m$, and a vector field A_μ, corresponding to a vector boson of mass $\sqrt{2}q\phi_0$. The particle corresponding to the field $h(x)$ is called a Higgs boson. As a consequence of local symmetry breaking, the gauge field acquires a mass, in contrast to the massless spin-zero Goldstone

boson.

Breaking the local symmetry results in the appearance of a vector field carrying mass, together with a scalar Higgs field also carrying mass. In the Weinberg and Salam electroweak theory [33], the masses of the $W^{\pm}$ and Z particles arise as a result of symmetry breaking, as will be shown in the following sections.

1.4 The standard model $SU(2)_W \times U(1)_Y$

The symmetries play an essential role in the Standard Model by implying specific nontrivial relations between masses and mixing angles that have been successfully tested experimentally.

The minimal way to introduce electromagnetic and weak interactions as gauge ones is to take the group $SU(2) \times U(1)$. The abelian group $U(1)$ is the same as the one that gives conservation of the electric charge in QED. Instead of the electric charge Q we introduce now the hypercharge Y. The $U(1)$ gauge symmetry provides interactions of fermions with a massless vector (photon-like) field. The non-abelian group $SU(2)$ is the same as the one used for the description of spinors in Quantum mechanics. Instead of spin, we use here the notion of weak isospin I.

The choice of the Higgs potential $V(\phi^{\dagger}\phi)$ is performed in such a way to give rise to spontaneous symmetry breaking 1.5:

$$V(\phi^{\dagger}\phi) = \frac{m^2}{2\phi_0^2}[(\phi^{\dagger}\phi) - \phi_0^2]^2 \tag{1.14}$$

where ϕ_0 is a fixed parameter that is the analogue of 1.5. With this expression for V, the vacuum state of our system is degenerate in the four-dimensional space of the scalar fields. We now break the SU(2) symmetry.

Introducing a two-component field, following the idea introduced by Yang and Mills in 1954, of the form:

$$\Phi = \begin{pmatrix} \Phi_A \\ \Phi_B \end{pmatrix} \tag{1.15}$$

it leads to a Lagrangian invariant under a local $SU(2) \times U(1)$ transformation.
Taking $\Phi_A = 0$ and Φ_B is real , the ground state is then:

$$\Phi_{ground} = \begin{pmatrix} 0 \\ \phi_0 \end{pmatrix} \tag{1.16}$$

and excited states as:

$$\Phi = \begin{pmatrix} 0 \\ \phi_0 + h(x)/\sqrt{2} \end{pmatrix} \tag{1.17}$$

with the real field $h(x)$.

In terms of $h(x)$ field, the $V(\phi^\dagger \phi)$ is written as:

$$V(\phi^\dagger \phi) = V(h) = m^2 h^2 + \frac{m^2 h^3}{\sqrt{2}\phi_0} + \frac{m^2 h^4}{8\phi_0^2} \tag{1.18}$$

Hence the locally gauge invariant Lagrangian density corresponding to is:

$$\mathscr{L}_\Phi = (D_\mu \Phi)^\dagger D^\mu \Phi - V(\phi^\dagger \phi) \tag{1.19}$$

Using the gauge field vector gauge field $B_\mu(x)$ with the transformation law:

$$B_\mu(x) \to B'_\mu(x) = B_\mu(x) + (2/g_1)\partial_\mu \theta \tag{1.20}$$

with the constant g_1 is a dimensionless parameter of the theory and for the global $SU(2)$ symmetry to be made into a local $SU(2)$ symmetry, with $U = U(x)$ dependent on space and time coordinates, we must introduce a vector gauge field $W_\mu^k(x)$ for each generator τ^k (identical to the Pauli spin matrices).
The transformation law for the matrices $W_\mu(x) = W_\mu^k(x)\tau^k$ is:

$$W_\mu(x) \to W'_\mu(x) = U(x)W_\mu(x)U^\dagger(x) + (2i/g_2)(\partial_\mu U(x))U^\dagger(x) \tag{1.21}$$

Here g_2 is another dimensionless parameter of the theory.

It is convenient to define the complex combinations:

$$W_\mu^+ = (W_\mu^1 - iW_\mu^2)/\sqrt{2} \tag{1.22}$$

and the complex conjugate:

$$W_\mu^- = (W_\mu^1 + iW_\mu^2)/\sqrt{2} \tag{1.23}$$

Now, going back to the Lagrangian density:

$$\mathscr{L}_\Phi = \frac{1}{2}\partial_\mu h \partial^\mu h + \frac{g_2^2}{2} W_\mu^- W^{+\mu}(\phi_0 + h/\sqrt{2})^2 + \frac{1}{4}(g_1^2 + g_2^2) Z_\mu Z^\mu (\phi_0 + h/\sqrt{2})^2 - V(h) \tag{1.24}$$

We have written

$$Z_\mu = W_\mu^3 cos\theta_w - B_\mu sin\theta_w \tag{1.25}$$

where

$$cos\theta_w = \frac{g_2}{(g_1^2 + g_2^2)^{1/2}} \tag{1.26}$$

$$sin\theta_w = \frac{g_1}{(g_1^2 + g_2^2)^{1/2}} \tag{1.27}$$

θ_w is called the Weinberg angle.

We can also define a field in the form:

$$A_\mu = W_\mu^3 sin\theta_w + B_\mu cos\theta_w \tag{1.28}$$

Therefore:

$$B_{\mu\nu} = A_{\mu\nu} cos\theta_w - Z_{\mu\nu} sin\theta_w \tag{1.29}$$

$$W_{\mu\nu}^3 = A_{\mu\nu} sin\theta_w + Z_{\mu\nu} cos\theta_w - ig_2(W_\mu^- W_\nu^+ - W_\nu^- W_\mu^+) \tag{1.30}$$

where $A_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, and $Z_{\mu\nu} = \partial_\mu Z_\nu - \partial_\nu Z_\mu$.

We can now identify the fields by writing the full Lagrangian density in terms of the fields fields A_μ and Z_μ: $\mathscr{L} = \mathscr{L}_1 + \mathscr{L}_2$ with:

$$\begin{aligned} \mathscr{L}_1 = \frac{1}{2}(\partial_\mu h \partial^\mu h) - m^2 h^2 - \frac{1}{4}[Z_{\mu\nu} Z^{\mu\nu}] \\ + \frac{1}{4}\phi_0^2 (g_1^2 + g_2^2) Z_\mu Z^\mu \\ - \frac{1}{4}[A_{\mu\nu} A^{\mu\nu}] \\ - \frac{1}{2}[(D_\mu W_\nu^+) * - (D_\nu W_\mu^+) *][D^\mu W^{+\nu} - D^\nu W^{+\mu}] + \frac{1}{2} g_2^2 \phi_0^2 W_\mu^- W^{+\mu} \end{aligned} \tag{1.31}$$

and $D_\mu W_\nu^+ = (\partial_\mu + ig_2 sin\theta_w A_\mu)W_\nu^+$.

$\mathscr{L}_1$ is the Lagrangian density for a free massive neutral scalar boson field $h(x)$, a free massive neutral vector boson field $Z_\mu(x)$, and a pair of massive charged vector boson fields $W_\mu^+(x)$ and $W_\mu^-(x)$ interacting with the massless electromagnetic field $A_\mu(x)$.

$\mathscr{L}_2$ is the sum of the remaining interaction terms:

$$\begin{aligned}\mathscr{L}_2 = &(\frac{1}{4}h^2 + \frac{1}{\sqrt{2}}h\phi_0)(g_2^2 W_\mu^- W^{+\mu} + \frac{1}{2}(g_1^2+g_2^2)Z_\mu Z^\mu) \\ &- \frac{m^2h^3}{\sqrt{2}\phi_0} - \frac{m^2h^4}{8\phi_0^2} + \frac{g_2^2}{4}(W_\mu^- W_\nu^+ - W_\nu^- W_\mu^+)(W^{-\mu}W^{+\nu} - W^{-\nu}W^{+\mu}) \\ &+ \frac{ig_2}{2}(A_{\mu\nu}sin\theta_w + Z_{\mu\nu}cos\theta_w)(W^{-\mu}W^{+\nu} - W^{-\nu}W^{+\mu}) \\ &- g_2^2 cos^2\theta_w(Z_\mu Z^\mu W_\nu^- W^{+\nu} - Z_\mu Z^\nu W_\nu^- W^{+\mu}) \\ &+ \frac{ig_2}{2}cos\theta_w[(Z_\mu W_\nu^- - Z_\nu W_\mu^-)(D^\mu W^{+\nu} - D^\nu W^{+\mu}) \\ &- (Z_\mu W_\nu^+ - Z_\nu W_\mu^+)(D^\mu W^{+\nu}) * - (D^\nu W^{+\mu}) *]\end{aligned} \tag{1.32}$$

In the Standard Model, the symmetry breaking mechanism is realized linearly by a non-zero vacuum expectation value scalar field. The resulting physical spectrum contains the massive intermediate vector bosons and fermionic matter fields, on top of the Higgs particle.

Most of the $U(1) \times SU(2)$ symmetry with which we began has been lost on symmetry breaking. Now we identify the three vector fields, W_μ^+, W_μ^-, Z_μ, with the mediators of the weak interaction, the W^+, W^- and Z particles, which, subsequent to the theory, were discovered experimentally:

$$M_W = \phi_0 g_2/\sqrt{2} \tag{1.33}$$

$$M_Z = \phi_0(g_1^2+g_2^2)^{1/2}/\sqrt{2} \tag{1.34}$$

Then $cos\theta_w = M_w/M_z = 0.8810 \pm 0.0016$.

The $W^\pm$ bosons are found experimentally to carry charge $\pm e$, with $e = g_2 sin\theta_w = g_1 cos\theta_w$.

Experimentally these quantities were measured at the values:

- $e^2/4\pi \approx 1/137.035$ (an energy scale Q=0),

Interaction field	Boson	Spin	Strength at the scale of quarks
Strong field	Gluons	1	1
Electromagnetic field	Photons	1	1/137
Weak field	W+ , W- , Z particles	1	10^{-6}
Gravitational field	'Gravitons' postulated	2	10^{-41}

Table 1.1 Types of interaction field.

- $sin^2\theta_W = 0.23120 \pm 0.00015$,
- $M_W = 80.379 \pm 0.012$ GeV
- $M_Z = 91.1876 \pm 0.0021$ GeV.

As a quantum field W_μ^+ destroys W^+ bosons and creates W^- bosons; W_μ^- destroys W^- bosons and creates W^+ bosons.

There remains the scalar Higgs field $h(x)$. The vacuum state expectation value ϕ_0 of the Higgs field is:

$$\phi_0 = \frac{\sqrt{2}M_W}{g_2} = 180\,GeV \tag{1.35}$$

To conclude, the requirements of $U(1)$ and $SU(2)$ symmetry, followed by $SU(2)$ symmetry breaking, generated the electromagnetic field, the massive vector $W^\pm$ and Z boson fields, and the scalar Higgs field.

At the quantum level, the structure of the Standard Model includes elementary fermions interacting through fields mediated by the exchange of particles, named bosons. Four types of interaction field are distinguished and described in Table 1.1 along with their intermediate particles, their spin property and the strength at the scale of quarks.

1.5 The composition of the SM

Atoms are the bound states due to electrostatic attraction composed of negatively charged electrons (e^-) that orbit around a central nucleus (formed by positively charged protons (p) and electrically neutral neutrons (n)). The Quantum Electrodynamics (QED) [41] is the quantum field theory of the interactions of charged particles with the electromagnetic field

Quarks	Mass ($\times c^{-2}$)	Electric charge
Up u	1.5 to 4 MeV	2/3
Down d	4 to 8 MeV	-1/3
Charmed c	1.15 to 1.35 GeV	2/3
Strange s	80 to 130 MeV	-1/3
Top t	172.44 GeV	2/3
Bottom b	4.1 to 4.4 GeV	-1/3

Table 1.2 Characteristics of the known quarks of the SM.

through emission and absorption of massless photons.

In the atomic nucleus, the protons and neutrons are bound together by the strong nuclear force, a manifestation of the fundamental theory of Quantum Chromodynamics (QCD) [41], mediated by massless gluons. Unlike the electromagnetic field, the gluon fields have the property of confinement, the phenomenon that describes color charged particles (such as quarks and gluons) can not be isolated.

The weak force completes the fundamental interactions of particle physics, mediated by charged W^+ and W^- bosons and the neutral Z^0 boson, responsible for the nuclear β-decays of specific radioactive isotopes and the nuclear fusion processes that fuel the Sun.

In the SM all the currently elementary particles are successfully organized as building blocks of three generations of fermions divided into quarks and leptons. All have spin 1/2, in units of $\hbar$. Leptons interact only through the electromagnetic interaction (when charged) and the weak interaction while quarks interact through the electromagnetic and weak interactions and also through the strong interaction.

The dynamics of the fermions are described by the Dirac equation of relativistic quantum mechanics (see Annex A.2). In this theory, for each of the twelve fermions there exists a state with the same mass but opposite charge named antiparticle state denoted either by their charge or by a bar over the corresponding particle symbol. For example, the anti-electron (or positron) is denoted by e^+, and the anti-strange-quark is written $\bar{s}$.

The known leptons and quarks along with their properties are listed in Figure 1.1 and in Table 1.2.

A property of the quarks that is mandatory to mention is the fact that they are confined in compound systems that extend over distances of about 1 fm, making their experimental observation very difficult. Via high-energy collisions, we can easily separate electrons from atoms, while, in the case of hadronic physics, there is no "ionization" type mechanism since the hadron spectrum found in nature consists of color singlet combinations of color nonsinglet objects namely the quarks and gluons, property of QCD theory referred to as “color confinement” [66].

The most elementary quark systems are **mesons** (quark + antiquark) which have net quark number zero and **baryons** number one.

The proton, the only stable baryon, basically contains two up quarks and one down quark (*uud*). The neutron contains two down quarks and one up (*udd*), and is therefore, a little more massive than the proton (by about 1.3 MeV/c^2). In the case of mesons, all are unstable. The lightest mesons are the $\pi^{\pm}$ mesons, the charged ‘pions’.

	Leptons				Quarks			
	Particle		Q	mass/GeV	Particle		Q	mass/GeV
First	electron	(e^-)	−1	0.0005	down	(d)	−1/3	0.003
generation	neutrino	(ν_e)	0	$< 10^{-9}$	up	(u)	+2/3	0.005
Second	muon	(μ^-)	−1	0.106	strange	(s)	−1/3	0.1
generation	neutrino	(ν_μ)	0	$< 10^{-9}$	charm	(c)	+2/3	1.3
Third	tau	(τ^-)	−1	1.78	bottom	(b)	−1/3	4.5
generation	neutrino	(ν_τ)	0	$< 10^{-9}$	top	(t)	+2/3	174

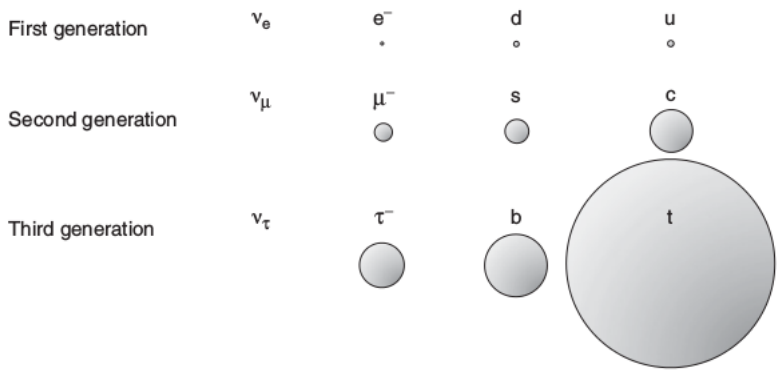


Fig. 1.1 Summary of the SM leptons and quarks.

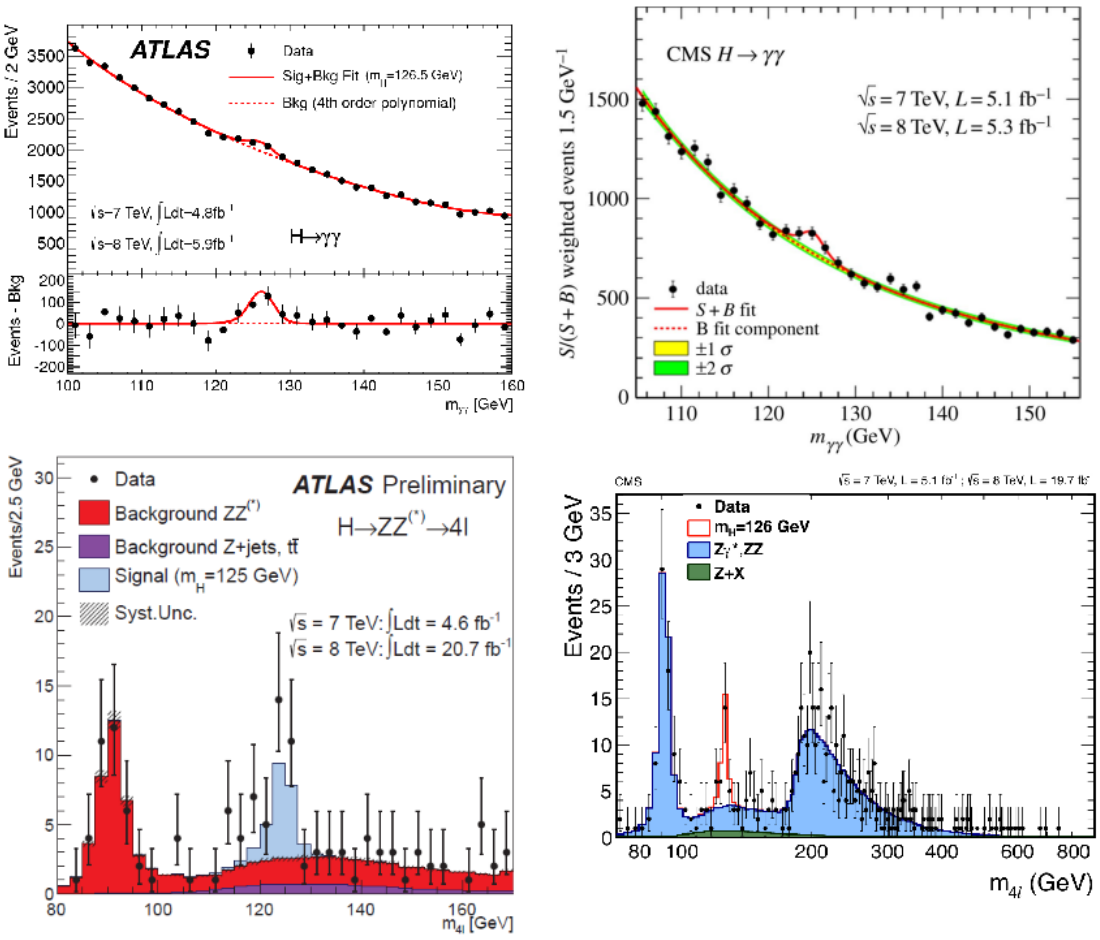


Fig. 1.2 The discovery of the Higgs boson by the ATLAS and CMS experiments

The next sections are dedicated to the Higgs boson and Top quark properties, the two primaries "ingredients" of my analysis.

1.6 The Higgs boson

In the year of 2012, a missing piece of the Standard Model puzzle was added by the discovery of the Higgs boson by the ATLAS and CMS experiments at the Large Hadron Collider (LHC) [46] (Figure 1.2). On 8 October 2013 the Nobel prize in physics was awarded jointly to François Englert and Peter Higgs " *for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider* ".

The Higgs boson is a spin-0 scalar particle. The mass measurement of Higgs boson performed by ATLAS and CMS experiments rely on the two high mass resolution and sensitive channels, $\gamma\gamma$ and ZZ. The ATLAS and CMS combined mass measurement puts a value on the Higgs mass [42] [43] at $m_H = 125.09 \pm 0.21(stat.) \pm 0.11(syst.)$ GeV with a precision of 0.2%, dominated by statistical uncertainties.

1.6.1 Higgs production and decays at the LHC

The main production mechanisms at the LHC (see the Feynman diagrams in Figure 1.3) are gluon fusion, weak-boson fusion, associated production with a gauge boson and associated production with a pair of top/antitop quarks.

The cross sections for the production of a SM Higgs boson as function of the center of mass energy in a pp collision ($\sqrt{s}$) are summarized in Figure 1.4 (left) while in the (right) part we can see the SM predictions for the decay fractions of the Higgs boson (branching ratios as a function of Higgs boson mass). The theoretical uncertainties are indicated as bands.

Since the Higgs couplings to other particles are proportional to their masses, the heaviest particles that are kinematically accessible will be preferred (Figure 1.4 on the right, which gives an overview of Higgs decay branching ratios for a mass range near the SM Higgs masses). The largest H production mechanism, $gg \to H$, and one of the cleanest H decay channels, $H \to \gamma\gamma$, are both loop-induced processes. Thus, LHC data already give us access to quantum aspects of Higgs physics, including the possible existence of new heavy particles beyond the Standard Model.

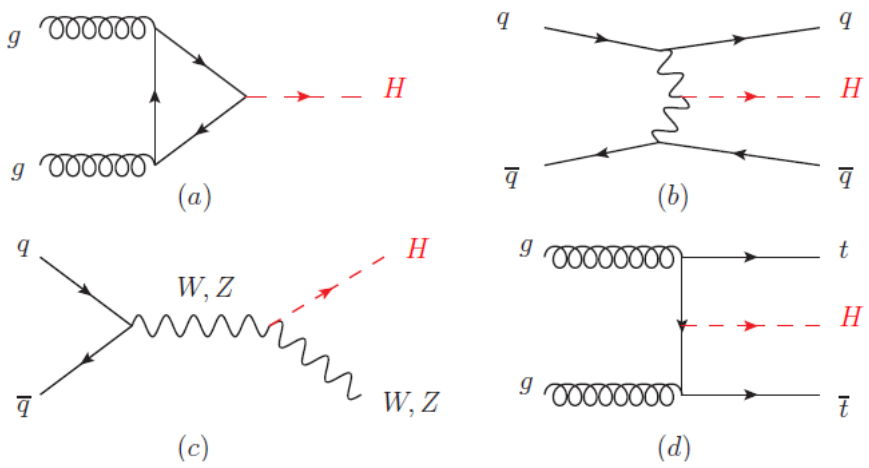


Fig. 1.3 Generic Feynman diagrams contributing to the Higgs production in (a) gluon fusion, (b) weak-boson fusion, (c) Higgs-strahlung (or associated production with a gauge boson) and (d) associated production with top quarks.

1.6.2 The mass of the Higgs boson

An accurate measurement of m_H is an essential parameter for precision tests of the Standard Model since it enters in the Higgs production cross section and decay branching ratios.

The discovery of the Higgs boson in 2012 was primarily based on the observation of excesses of events in the $\gamma\gamma$ and $2l^+2l^-$ channels (l = muon or electrons). The final combined results from ATLAS and CMS LHC Run 1 data leads to:

$$m_H = 125.09 \pm 0.21(statistical) \pm 0.11(systematic) GeV \tag{1.36}$$

1.6.3 The measurement of μ

The signal strengths μ_i^f are defined as the ratios of cross sections and branching fractions to the corresponding SM predictions such that:

$$\mu_i^f = \frac{\sigma_i \cdot BR^f}{(\sigma_i)_{SM} \cdot (BR^f)_{SM}} = \mu_i \cdot \mu^f \tag{1.37}$$

where the subscript i and superscript f indicate the production mode and decay channel, respectively. By definition, μ_i^f equal to 1 indicates the cpmpatibility of the data with the SM Higgs boson hypothesis.

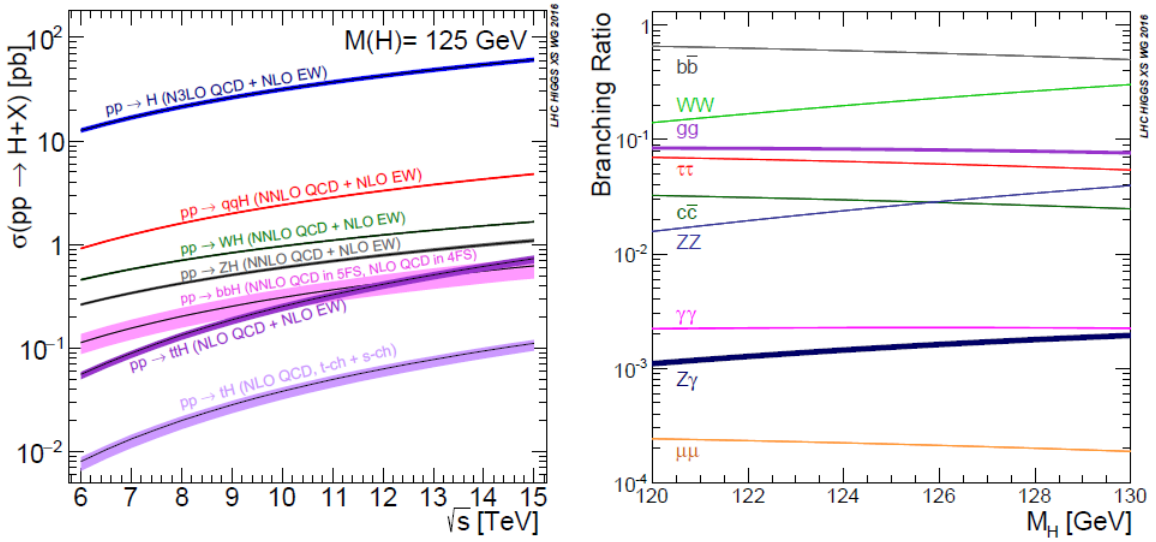


Fig. 1.4 (Left) The SM Higgs boson production cross sections as a function of the center of mass energy. (Right) The branching ratios for the main decays of the SM Higgs boson near m_H = 125 GeV.

A fit to the combined ATLAS and CMS data at $E_{CM} = 7,8$ TeV (Run I) results in the best-fit value:

$$\mu = 1.09^{+0.11}_{-0.10} \tag{1.38}$$

This result is consistent with the SM expectation within less than 1σ, and the p-value of the compatibility between the data and the SM prediction is 34%.

Individual signal strengths for each production mode and by the decay process can be seen in Figure 1.5.Comparing likelihood of the best-fit with likelihood assuming μ_{prod}=0 or μ^{decay}=0 we see the VBF production and $H \to \tau\tau$ are established at over 5 σ (Table 1.3).

1.6.4 A summary plot on Higgs property

The relation between the couplings and the SM predictions depends on the mass of the particles to which the H boson is coupling. In a case of a vector boson of mass m_V, the coupling to the Higgs expressed as a function on κ_V is:

$$y_{V,i} = \sqrt{\kappa_{V,i}} \frac{m_{V,i}}{v} \tag{1.39}$$

where v = 246 GeV is the vacuum expectation value of the Higgs field.

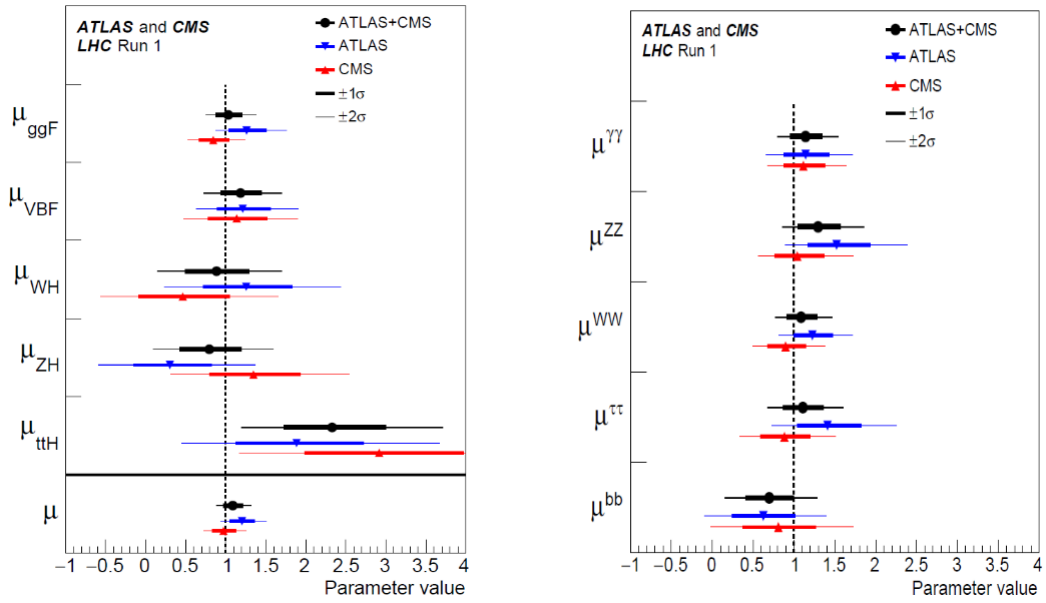


Fig. 1.5 Best fit results for the production signal strengths (left) and for the decay signal strengths (right), for the combination of ATLAS and CMS data. The error bars indicate the 1σ (thick lines) and 2σ (thin lines) intervals.

In the case where we have a fermion of mass m_f, the coupling of the Higgs as a function on κ_f is:

$$y_{f,i} = \kappa_{f,i}\sqrt{2}\frac{m_{f,i}}{v} \tag{1.40}$$

The linear scaling of the reduced coupling modifiers as a function of the particle masses is shown in Figure 1.6 and indicates the consistency of the measurements with the SM.

The top quark [63] is unique among the known quarks due to its high mass. Therefore top Yukawa coupling is roughly 30 times the bottom Yukawa coupling, the next largest.

The measurement of the $t\bar{t}$H production rate is directly linked to the top Yukawa coupling. It provides a direct test of the Higgs couplings but also serves a test for the new physics beyond SM, making it an essential parameter of the SM.

Run I [70] results on $t\bar{t}$H production signal strengths are shown in Table 1.4.

The $t\bar{t}$H production within the SM framework will be presented in more details in Chapter 4. I will cover the search of $t\bar{t}H \rightarrow$ multileptons, focusing on the $4l$ channel.

Production process	Observed Significance(σ)	Expected Significance (σ)
VBF	5.4	4.7
WH	2.4	2.7
ZH	2.3	2.9
VH	3.5	4.2
ttH	4.4	2.0
Decay channel		
$H \to \tau\tau$	5.5	5.0
$H \to bb$	2.6	3.7

Table 1.3 The observed an expected significance of combined observations.

production process	ATLAS	CMS	ATLAS&CMS
$\mu_{t\bar{t}H}$	$1.9^{+0.8}_{-0.7}$	$2.9^{+1.0}_{-0.9}$	$2.3^{+0.7}_{-0.6}$

Table 1.4 Run 1 [70] $t\bar{t}$H production signal strengths.

In the next subsection, I will give an overall description of the ATLAS searches for Beyond the Standard Model Higgs Bosons. A more detailed analysis of BSM search using doubly charged Higgs model will be presented in Chapter 5.

1.6.5 ATLAS Searches for Beyond the Standard Model Higgs Bosons

The discovery of a new 125 GeV boson by CMS and ATLAS at the Large Hadron Collider (LHC) was a great success for the Standard Model (SM) mechanism for electroweak symmetry breaking. However, there is still room for physics beyond SM since this theory does not explain gravity, dark matter, dark energy, baryon asymmetry or the hierarchy problem.

These arguments suggest that the SM must be extended with more complete theoretical models, beyond SM (BSM) [81], which sometimes predict the existence of more than one Higgs boson.

In a generic Two-Higgs-Doublet Model (2HDM) there are two electroweak doublets, in contrast to the one doublet of the SM. The 2HDM can be of Types I, II or III depending on how the doublets couple to fermions.

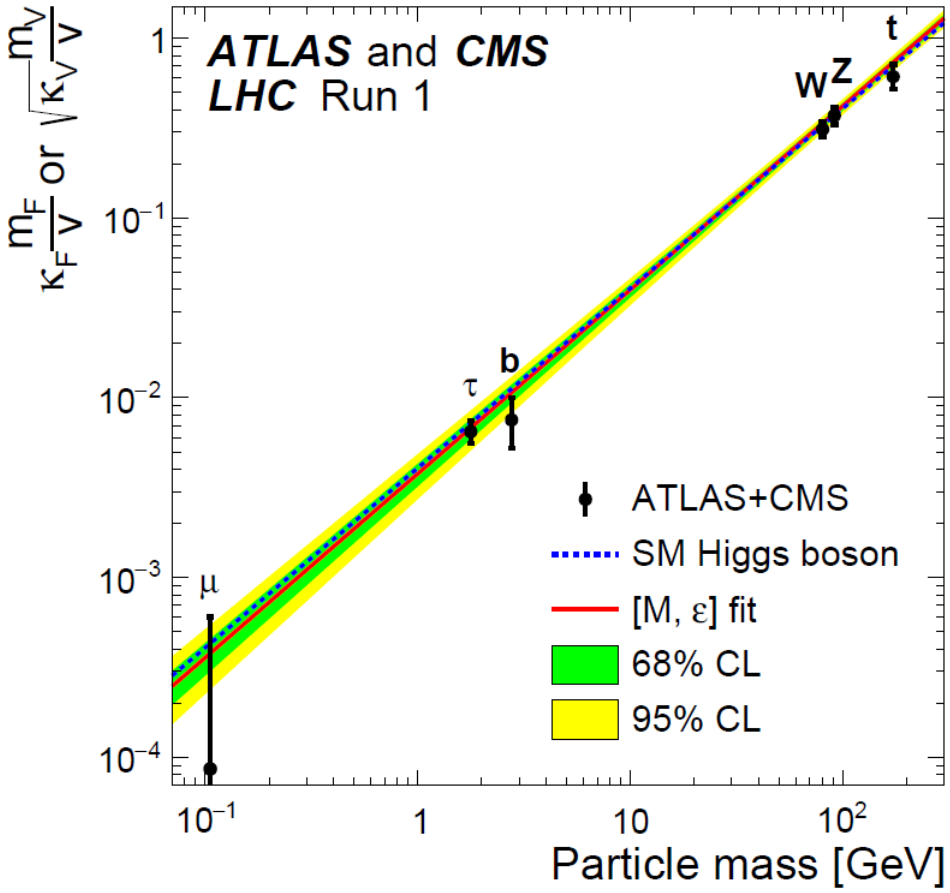


Fig. 1.6 Fit results as a function of the particle mass for the combination of ATLAS and CMS. The dashed line indicates the predicted dependence on the particle mass for the SM Higgs boson. The solid (red) line indicates the best fit result the corresponding 68% and 95% CL bands in green and yellow.

In Type I one doublet does not couple to fermions, in Type II one doublet (H_u) couples to up-type fermions while the other doublet (H_d) couples to down-type fermions, and in Type III both doublets couple to both up- and down-type fermions.

Another intriguing theoretical model is supersymmetry (SUSY) [82], which brings along super-partners of the SM particles. The Minimal Supersymmetric Standard Model (MSSM), whose Higgs sector is equivalent to the one of a constrained 2HDM of type II and the next-to MSSM (NMSSM) are among the experimentally most studied models, because they provide good benchmarks for SUSY, reachable at LHC in Run II.

ATLAS and CMS also perform general searches for new heavy particles that can decay to a pair of Higgs bosons, which are motivated by (warped) extra dimension models.

In the SM, the invisible decay of the Higgs boson ($H \to ZZ \to \nu\nu\nu\nu$) has a branching ratio BH $\to$ inv of 1.06 $\times 10^{-3}$ for m_H = 125 GeV. A larger BH $\to$ inv can exist in many extensions of the SM, giving a strong indication for physics beyond the SM (BSM). Models with a sizable BH $\to$ inv often involve a Higgs boson decaying into weakly interacting massive particles (WIMPs) [83], a possible candidate of dark matter, and thus, studying BH $\to$ inv gives a unique probe into DM through its coupling to the Higgs boson.

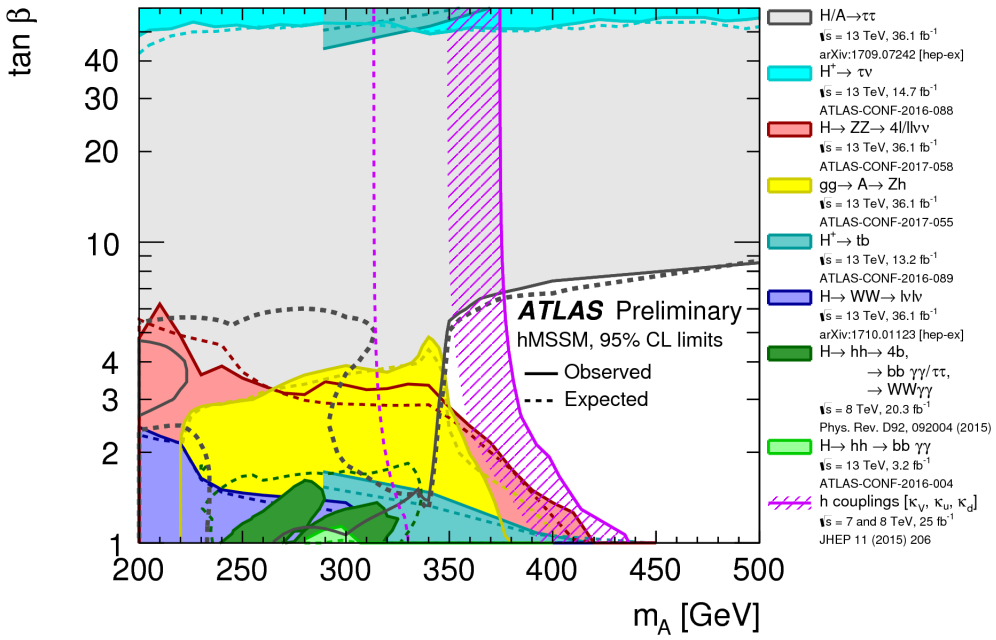


Fig. 1.7 hMSSM overlaid studies in ATLAS.

The 2 Higgs doublet model (2HDM) is a simple extension of the SM which can exhibit large resonance effects. The 2HDM has 5 physical Higgs bosons: 2 neutral scalars, h^0, H^0, a pseudo-scalar, A, and a charged Higgs boson pair $H^{\pm}$.
The couplings of the Higgs bosons to fermions are described by two free parameters: the ratio of vacuum expectation values of the two Higgs doublets, $\tan\beta = v_2/v_1$, and the mixing angle α which diagonalizes the neutral scalar mass matrix. 2HDMs are significantly limited by experimental data. Higgs boson coupling measurements restrict $\cos(\alpha - \beta)$ to be close to the

SM limit, $\cos(\alpha-\beta)\sim 0$, while heavy Higgs searches restrict M_H^0 as a function of $\cos(\alpha-\beta)$.

A summary plot of ATLAS BSM searches in presented in Figure 1.7. This includes indirect searches namely looking for deviations from the SM Higgs boson or its properties (spin, couplings) or direct searches by looking for additional Higgs-like scalars, neutral or charged, decays to SM particles or to H bosons or by looking for BSM Higgs decays and couplings, searching for new resonances, long lived particles, flavor violating couplings, invisible decays.

With the appetite" opened for BSM searches, I will present a study beyond Standard Model search namely a model of doubly charged scalar bosons decaying to $W^{\pm}$ $W^{\pm}$ in Chapter 5.

1.7 Top Quark Physics

Due to its large mass, the top quark is an excellent probe of electroweak gauge symmetry breaking mechanism and should, therefore, play a key role in clarifying the nature of the force(s)/particle(s) responsible for this phenomenon. The top quark is also a proper probe for possible new parity-violating and/or non-SM CP violating interactions which could be induced, for instance, by non-standard Higgs bosons. An extension of the SM is also considered by introduction new top-quark decay modes, for instance to supersymmetric particles. However, so far, experimental data are consistent with the SM prediction, $t\rightarrow W^{+}b$ being the dominant mode – but its branching ratio and the structure of the tbW vertex is not yet measured directly with high accuracy.

The top quark [63] is unique among the known quarks due to its high mass, decaying before it can form hadronic bound states. It offers the possibility to explore the interactions of a bare quark at energies of a few hundred GeV to several TeV.

The top quark pair and single production can be seen in Table 1.5 both for Tevatron (at a peak luminosity of $L=4\times 10^{32}cm^{-2}s^{-1}$) and LHC (at a peak luminosity reached in 2018 of $L=1.58\times 10^{34}cm^{-2}s^{-1}$)colliders.

$t\bar{t}$ pairs	dominant reaction	$N_{t\bar{t}}$
Tevatron: $p\bar{p}$ (1.96 TeV)	$q\bar{q} \to t\bar{t}$	$\sim 7 \cdot 10^4 \times L$
LHC: pp (14 TeV)	$gg \to t\bar{t}$	$\sim 9 \cdot 10^5 \times L$
ILC: e^+e^- (400 GeV)	$ee \to t\bar{t}$	$\sim 800 \times L$
single top	dominant reaction	$(N_t + N\bar{t})$
Tevatron:	$u + b \to d + t$	$\sim 3 \cdot 10^3 \times L$
LHC:	$u + b \to d + t$	$\sim 3.3 \cdot 10^5 \times L$

Table 1.5 Top quark production models.

1.7.1 Discovery of the Top Quark

The top is the last of the fundamental constituents of subnuclear matter that theories of the strong, weak, and electromagnetic interactions. Theoretically, the top's existence was required to make the electroweak theory internally consistent. Experimentally, the top's existence was signaled by the pattern of disintegration of the b quark and by the way b quarks interacted with other particles via the exchange of Z bosons, leading in February 1995 to the announcement that the top quark [61] had been observed in two experiments at the Tevatron [62] proton-antiproton collider at Fermilab.

The top quark is a remarkable particle. Although a single top quark weighs about as much as an atom of gold, we expect that it is structureless down to a scale of at least 10^{-18} m. Its expected lifetime of about 0.4×10^{-24} s, introduces the argument that they were produced in an insignificant number in the moment of the Big Bang after disintegrated in the merest fraction of a second, and then vanished from the scene.

1.7.2 Lifetime

Because its mass is so large compared to other quarks, it can decay into on-shell W bosons, the dominant mode in the SM prediction. The average proper lifetime of the t quark is extremely short, $\tau_t = 1/\Gamma_t \cong 5 \times 10^{-25}$ s.

For comparison, the mean lifetime of b hadrons, which involve the next most massive quark, is almost 13 orders of magnitude larger, $\tau_{bhadron} \cong 1.5 \times 10^{-12}$ s.

1.7.3 Spin

The top quark, as observed at the Tevatron, is a spin 1/2 fermion. The observed decay $t \to bW$, the known spins of W and b, and the conservation of total angular momentum imply that the top quark is a fermion. If the spin of the top quark were 3/2, the $t\bar{t}$ cross section at the Tevatron would be much larger than the measured one.

A direct experimental evidence for the top quark having spin 1/2 would be the observation of the resulting polarization and spin-correlation effects. Another possibility is the measurement of the differential cross section $\frac{d\sigma}{M_{t\bar{t}}}$ near the $t\bar{t}$ production threshold.

1.7.4 Colour and electric charge

Top quarks, like the other quarks, carry color charge – they transform as a color triplet under the $SU(3)_C$ gauge group of the strong interactions. Colour-confinement precludes the direct measurement of this quantum number.

The top quark is the $I_3 = \frac{1}{2}$ weak-isospin partner of the b quark.

Measurements of the $t\bar{t}$ production cross section are consistent with the SM predictions for a colour-triplet and antitriplet quark-antiquark pair.

The electric charge of the top quark, which is $Q_t = \frac{2}{3}$ in units of the positron charge $e > 0$ according to the SM, has so far not been measured.

The top-quark charge can be directly determined by measuring the production rate of $t\bar{t}$ plus a hard photon and taking the ratio $\frac{\sigma_{t\bar{t}\gamma}}{\sigma_{t\bar{t}}}$. At the LHC this ratio is approximately proportional to Q_t^2 because $t\bar{t}$ and $t\bar{t}\gamma$ production is dominated by gluon fusion.

1.7.5 Spin correlations

Although the top and antitop quarks produced via the strong interaction at the hadron colliders are unpolarized, their spins are correlated. New physics models beyond the SM can change the spin correlation of the top and the antitop quark by either changing the spin of their decay products or by changing the production mechanism of the $t\bar{t}$ pair [65].

1.7.6 Charge Asymmetry

The $t\bar{t}$ production is predicted to be symmetric under charge conjugation at LO in the SM. At NLO, however, for the $q\bar{q}$ and qg production modes, there is a small preference to produce top (antitop) quarks in the direction of the incoming quark (antiquark) in the partonic rest frame.

At the Tevatron, this asymmetry is equivalent to a forward-backward asymmetry, with top (antitop) quarks more abundant in the direction of the incoming proton (antiproton).

At the LHC, this forward-backward asymmetry vanishes trivially because the pp initial state (dominated by gluon-gluon fusion) is symmetric. Nevertheless, QCD predicts a slight preference for centrally produced antitop quarks at the LHC, while top quarks are more abundant at large positive and negative rapidities. This is due to the proton composition in terms of quarks and antiquarks: quarks carry a large momentum fraction while antiquarks carry a small momentum fraction.

1.7.7 Couplings

The top quark couples to other SM fields through its gauge and Yukawa interactions. The coupling of the top quark to the W boson can be examined either by looking at the decay of the top quark or from single top quark production. With the advent of large data samples with top quarks at the LHC, the processes where bosons (g, Z or H) are produced in association with top quarks become accessible. The first evidence on the couplings of the top quark to these particles (will) come from the production rate of these processes.

1.7.8 Top quark pair production

In the SM, the dominant production mechanism for top quark pairs is mediated by the strong interaction, and since the top quark mass is much larger than Λ_{QCD}, $t\bar{t}$ production at LHC can be successfully described in terms of quantum chromodynamics (QCD), the theory of the strong interaction.

At leading order (LO) QCD, the processes $gg \to t\bar{t}$ and $q\bar{q} \to t\bar{t}$ contribute as seen in 1.8.

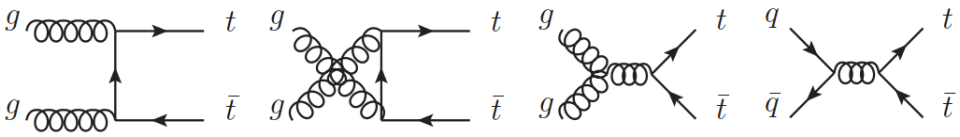


Fig. 1.8 Feynman diagrams for $t\bar{t}$ production at leading order QCD

While Next to Leading Order (NLO) corrections are entirely known, the Next to Next Leading Order (NNLO) are partly known, approximate NNLO results and the production at LHC and Tevatron being depicted in Figure 1.9.

	Tevatron 1.96 TeV	LHC 14 TeV
qq gg	85% 15%	5% 95%
σ (pb)	7.0 pb	887 pb

Fig. 1.9 Top Quark Production

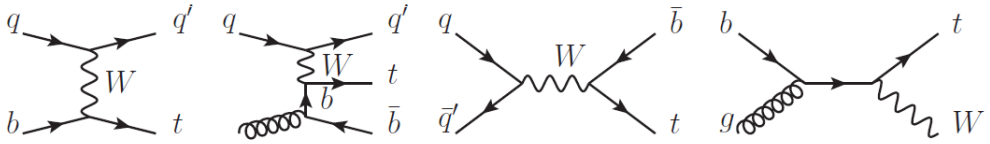
1.7.9 Production of single top quarks

Single top quarks can be produced through the electroweak interaction by W_{tb} vertex (almost exclusively). Three different production modes 1.10 are considered:

1. the t-channel mode, space-like W-boson scatters off a b-quark
2. the s-channel mode, a time-like W-boson is produced from two quarks belonging to an isospin doublet
3. the tW-channel mode, which is also called **associated production**, the top quark is produced in association with a close-to-real W-boson

1.7.10 Top quark decays

The top quark mass is above the threshold for (top) decays to a W boson and a down-like quark, which is, therefore, the dominant decay.



Example Feynman diagrams for single top quark production at LO QCD. From left to right: t-channel production as flavor excitation and as W-gluon fusion; s-channel production; tW-channel production.

Fig. 1.10 Single Top Quark Production.

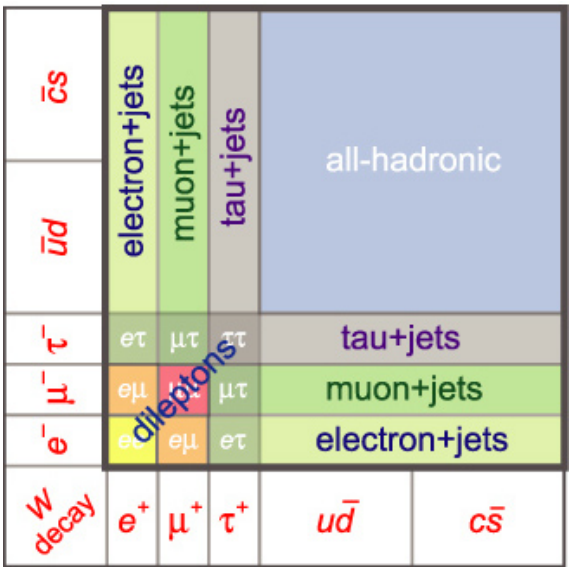


Fig. 1.11 Top Pair decay channels

The decays of top quark pairs (Figure 1.11, Figure 1.12) can be classified according to the decay of the W-bosons:

- Di-lepton channel: both W-bosons decay into lepton (electron, muon, tau) and neutrino, $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \bar{l}\nu_l bl'\bar{\nu}_{l'}\bar{b}$. The branching fraction is $BR(di-lepton;e;\mu;\tau) = 10.3\%$ while $BR(di-lepton;e;\mu) = 4.7\%$.

- Lepton+jets channel: one W-boson decays into lepton and neutrino, the other one into a quark - anti-quark pair, $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'bl\bar{\nu}_l\bar{b}$. The branching fraction is $BR(lepton+jets;e;\mu;\tau) = 43.5\%$, approximately evenly split for W-boson decays into electron, muon and tau.

- Hadronic channel: both W-bosons decay into a quark - anti-quark pair, $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'bq''\bar{q}''\bar{b}$. The branching fraction is $BR(hadronic) = 46.2\%$

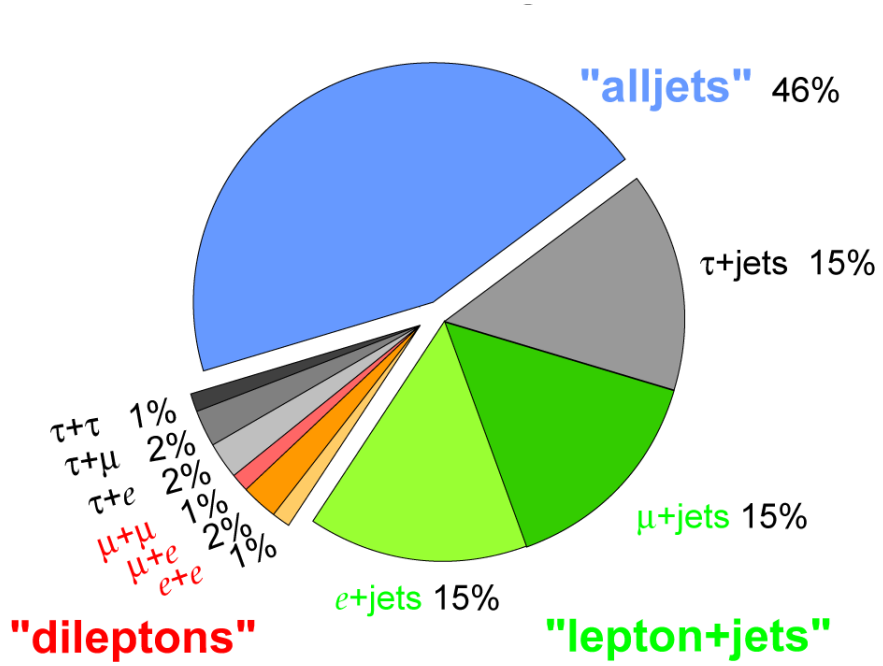


Fig. 1.12 Top quark pair branching fractions into decay channels and their rate.

1.8 Conclusion

In 2012, a missing piece of the Standard Model puzzle was added by the discovery of the Higgs boson by the ATLAS and CMS experiments at the Large Hadron Collider (LHC) [46]. Higgs properties measurements are essential to provide further validation of the Standard Model, one key parameter being the top Yukawa coupling.

The measurement of the $t\bar{t}$H production rate is directly linked to the top Yukawa coupling. It provides a direct test of the Higgs couplings but also serves a test for the new physics beyond SM. The $t\bar{t}$H production will be presented in more details in Chapter 4.

Run 2 will bring a much higher available statistics, opening a possibility to reach a discovery of $t\bar{t}$H production and allow a good precision on the coupling measurements. In addition, complementary approaches to disentangle possible BSM effects will be continued in Run 2 and will provide higher accuracy.

For many reasons we do not believe that the SM is the final "theory of everything". Firstly, SM is not joined with General Relativity. In general, few open questions within the SM needs

better understanding: the origin of symmetries; the origin of EW and QCD energy scales; the origin of 3 fermion generations; the origin of neutrino masses; the hierarchy of leptons and fermions masses; the absence of strong CP violation in the QCD sector; confinement in QCD, and so on. There are also some phenomenological issues: the baryon asymmetry in the Universe the dark matter; the dark energy; the proton charge radius.

The main decay and production properties of the observed Higgs boson are consistent with predictions of the SM. It may, however, have other decay channels beyond those anticipated in the SM. I will, therefore, present a beyond Standard Model search namely a model of doubly charged scalar bosons decaying to $W^{\pm}\,W^{\pm}$ in Chapter 5, model based on an extension of the SM.

With this elegant mathematical framework in our hands, the Standard Model, we can move further to the experimental physics, the only way a theory could be validated or excluded. Therefore, I will continue with a description of the CERN LHC experiment, with particular attention given to the ATLAS detector, setting, therefore, the experimental part of my research.

A complete description of the ATLAS detector and the object selection is given, with the emphasis on the calorimeter system which is of significant importance within my performance work (studying the electron identification efficiency using the W tag and probe method).

Chapter 2

The LHC experiment and the ATLAS detector

2.1 Introduction

There is a crucial link between theory and experimental physics. In experimental physics, physicists rely on a test theory to construct and design the experimental platform. Then, experimental physicists can perform particular comparisons between specific theories or classes of theory with the outcome of the experiment, validating or excluding, therefore, the theory tested.

This idea-driven the necessity to test and validate the Standard Model of particles, using accelerator physics, in particular, using the LHC experiment.

2.2 What is CERN?

Founded in 1954, the international research institution CERN [1] [2] [32] ("Conseil Européen pour la Recherche Nucléaire"), aims to probe the fundamental structure of the universe.

Making use of dedicated accelerators, which enable particles to collide together with a velocity close to the speed of light, physicists can study the fundamental constituents of matter and the forces acting between them.

The CERN laboratory sits astride the Franco-Swiss border near Geneva with a total of 22 member states involved in the collaboration, with Romania joining in 2016.

2.3 The Large Hadron Collider (LHC)

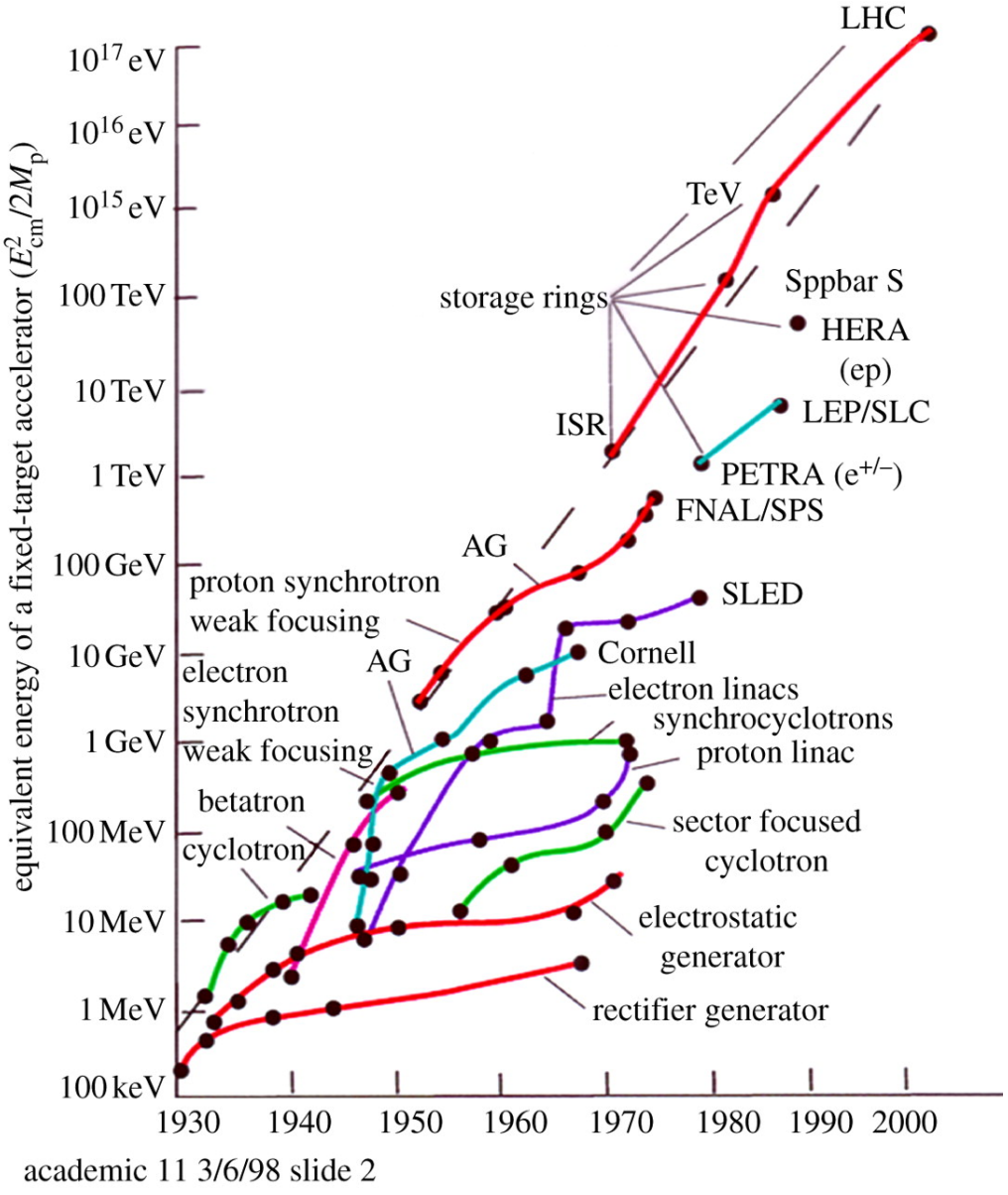


Fig. 2.1 Schematic representation of 80 years of particle accelerator physics [84].

Particle accelerators and detectors have existed and been developed for more than 80 years, ranging the equivalent energy of a fixed-target accelerator from 100 keV to 10^{17} eV

(as can be seen in Figure 2.1). The accelerators and detectors have increased enormously in complexity, size, cost, reaching to the present first TeV (7 TeV/beam) and largest proton collider, the LHC. More about the history of particle detectors can be found in [85].

The Large Hadron Collider (LHC) is a wonderful experimental facility which provides us an opportunity to test the existence of new types of sub-atomic particles other than proton, neutrons, and electrons.

Started on 10 September 2008, the Large Hadron Collider (LHC) [1] [2] [32] is the world's largest and most powerful particle accelerator, running at TeV scale energy, consisting of a 27-kilometer ring of superconducting magnets with additional accelerating structures to boost the energy of the particles along the way (Figure 2.2).

Protons accelerated through the accelerator complex at CERN follow a simple schema [32]:

- Hydrogen atoms are taken from a bottle containing hydrogen. From this, protons are obtained by stripping electrons from hydrogen atoms.
- Protons are injected into the PS Booster (PSB) at an energy of 50 MeV from Linac2.
- The booster accelerates them to 1.4 GeV and sends the beam to the Proton Synchrotron (PS) where it is accelerated to 25 GeV.
- Protons are then sent to the Super Proton Synchrotron (SPS) where they further accelerate to 450 GeV.
- Proton beam is finally transferred to the LHC (both in a clockwise and an anticlockwise direction) where it is accelerated for 20 minutes to 6.5 TeV. Under normal operating conditions, this process can last for many hours inside the LHC beam pipes.

When designing LHC, the selected parameters for run operation at 7 TeV are shown in Table 2.1.

Inside the accelerator, strong magnetic fields are driven by the superconducting state of the magnets guide the two high-energy particle beams around the accelerator ring with a velocity close to the speed of light, in opposite directions.

Parameter	Value
Circumference	26 659 m
Dipole operating temperature	1.9 K (-271.3°C)
Number of magnets	9593
Number of main dipoles	1232
Number of main quadrupoles	392
Number of RF cavities	8 per direction
Energy, protons	7 TeV
Energy, ions	2.56 TeV/u
Peak magnetic dipole field	7.74 T
Distance between bunches	7.5 m
Luminosity (protons)	Peak Luminosity $\approx 1.2 \times 10^{34} cm^{-2} s^{-1}$
No. of bunches per proton beam	$\approx$ 2808 in 2018, $\approx$ 2000 in 2017
No. of protons per bunch (at start)	1.2×10^{11}
Number of turns per second	11 245
Number of collisions per second	1 billion

Table 2.1 LHC parameters.

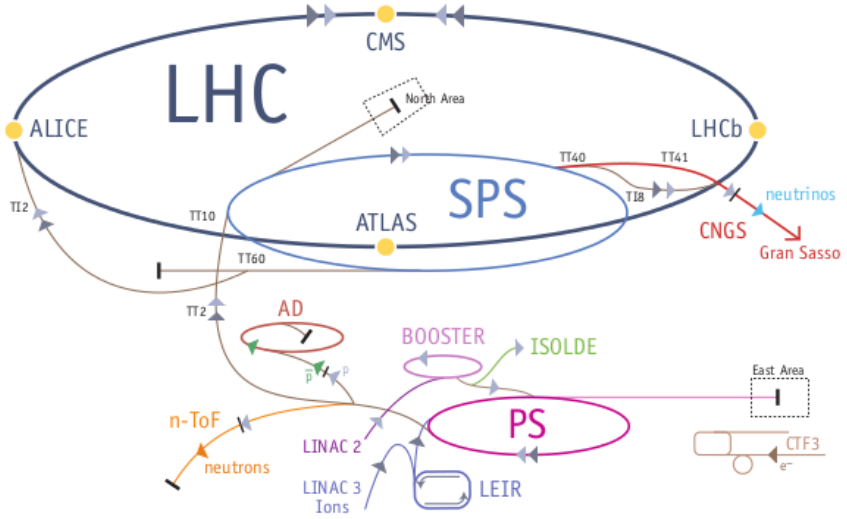


Fig. 2.2 Schematic representation of LHC accelerator complex [32]

2.4 LHC experiments

In the LHC site, seven distinct experiments are using detectors to analyze the particles produced by collisions in the accelerator.

Four detectors sit underground in huge caverns on the LHC ring. From these, the biggest are two independent general purpose detectors, ATLAS [3] and CMS [4], aimed, in an independent and complementary manner, to study the Standard Model, to search for the Higgs boson and new phenomena beyond the Standard Model. ALICE [5] and LHCb [6] are detectors specialized in underlying specific phenomena. ALICE aims at studying quark-gluon plasma, a state of matter thought to have formed just after the big bang while the LHCb experiment plans to discover why we live in a universe that appears to be composed almost entirely of matter, but no antimatter.

The smallest experiments are TOTEM [7] and LHCf, which focus on "forward particles" using detectors positioned on either side of the CMS interaction point, while LHCf is made up of two detectors which sit along the LHC beamline. The LHCf experiment uses particles thrown forward by LHC collisions to simulate cosmic rays. It is complementary to The 'Total, elastic and diffractive cross-section measurement' (TOTEM) experiment which studies particles in the forward region by collisions in the LHC. MoEDAL [8] detectors are deployed near LHCb to search for a hypothetical primordial particle called the magnetic monopole.

2.5 Purpose of the LHC facility

Although the Standard Model (SM) of particles and forces (see Chapter 1) summarizes our present knowledge of particle physics successfully, many unsolved questions survive, which the LHC will hopefully help to answer.

Within the SM framework, the origin of mass is considered to be a consequence of the so-called Higgs mechanism. This theory lays on the theoretical ground that space holds the 'Higgs field', and by interacting with this field intensely (heavy particles) or feebly (light particles), particles acquire their masses accordingly. However, a missing puzzle of the Standard Model [9] is the explanation of the mass hierarchy observed between particles.

Another weakness of the SM theory is the non-inclusion of gravity among the fundamental forces. One theory constructed to facilitate the unification of fundamental forces is the supersymmetry (SUSY) [15], a theory that hypothesizes the existence of more massive partners of the standard particles we know. Within SUSY theory, one writes an effective Lagrangian (or superpotential) in terms of a certain set of hidden sector fields that lead to supersymmetry breaking in the ground state. The messenger sector includes the following possibilities [79]: (i) gravity mediated; (ii) gauge mediated and (iii) anomalous U(1) mediated. The proof of the SUSY theory through the existence of lightest supersymmetric particles should also be a groundbreaking discovery at the LHC future runs.

From cosmological and astrophysical observations [16] we know that all of the visible matter accounts for only 4 % of the Universe (see Figure 2.3). The hunting for the remaining particles or phenomena responsible for dark matter (23%) and dark energy (73%) is still ongoing. For example, the Optical Search for QED Vacuum Bifringence, Axions, and Photon Regeneration (OSQAR) experiment [28], along with LHCb, looks for particles that could be a component of dark matter and explain why our universe is made of matter instead of antimatter. The Alpha Magnetic Spectrometer [27] looks for dark matter, antimatter and missing matter from a module on the International Space Station.

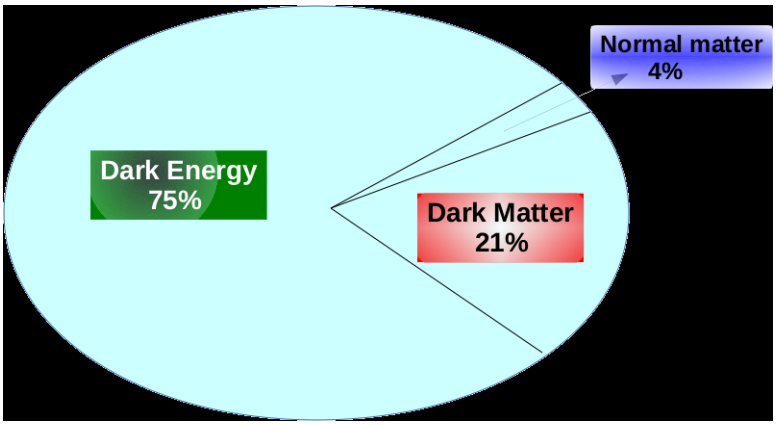


Fig. 2.3 Universe structure

Another purpose of LHC is to investigate the mystery of antimatter. Contrary to the time of Big Bang, when matter and antimatter were supposed to be produced in the same amounts, our current Universe is made only of matter [40]. So one fundamental question still surviving nowadays is why is there far more matter than antimatter in the Universe and what produced this discrepancy? LHCb, OSQAR, AMS experiments plan to tackle this problem

by analyzing the matter-antimatter properties.

2.6 LHC road map

Run-1 (see Figure 2.4) has been a fantastic success, few achievements that are worthy to be mentioned being:

1. Discovery of the Higgs boson [42], [43];
2. Severe constraints on physics BSM [44], [81];
3. $>$ 500 papers published on a vast variety of topics.

Upon Higgs discovery in Run 1, a new era of precision measurements and search for hints of New Physics (NP) or Physics Beyond Standard Model (BSM) has begun, with high luminosity upgrades that would significantly enhance the achievable precision.

Considering the proposed LHC future (see Figure 2.4), knowledge of TeV scale physics will be improved dramatically by future LHC running:

1. Going to full energy and increase L by factor 100;
2. Higgs couplings will be measured with a precision of 2-10 %;
3. Searches for new particles will extend.

Initially conceived in the 1980s, the LHC took another 25 years to come to life in the form we know today.

The necessity to upgrade the LHC has given rise to the HL-LHC project (see Figure 2.4 for the High Luminosity LHC project roadmap). HL-LHC relies on some important innovative technologies, representing exceptional technological challenges, such as cutting-edge 13 Tesla superconducting magnets, very compact and ultra-precise superconducting cavities for beam rotation, and 300-meter-long high-power superconducting links with zero energy dissipation.

The HL (High Luminosity) LHC [29] is CERN's number-one priority and will increase the number of collisions accumulated in the experiments by a factor of ten from 2024 onwards.

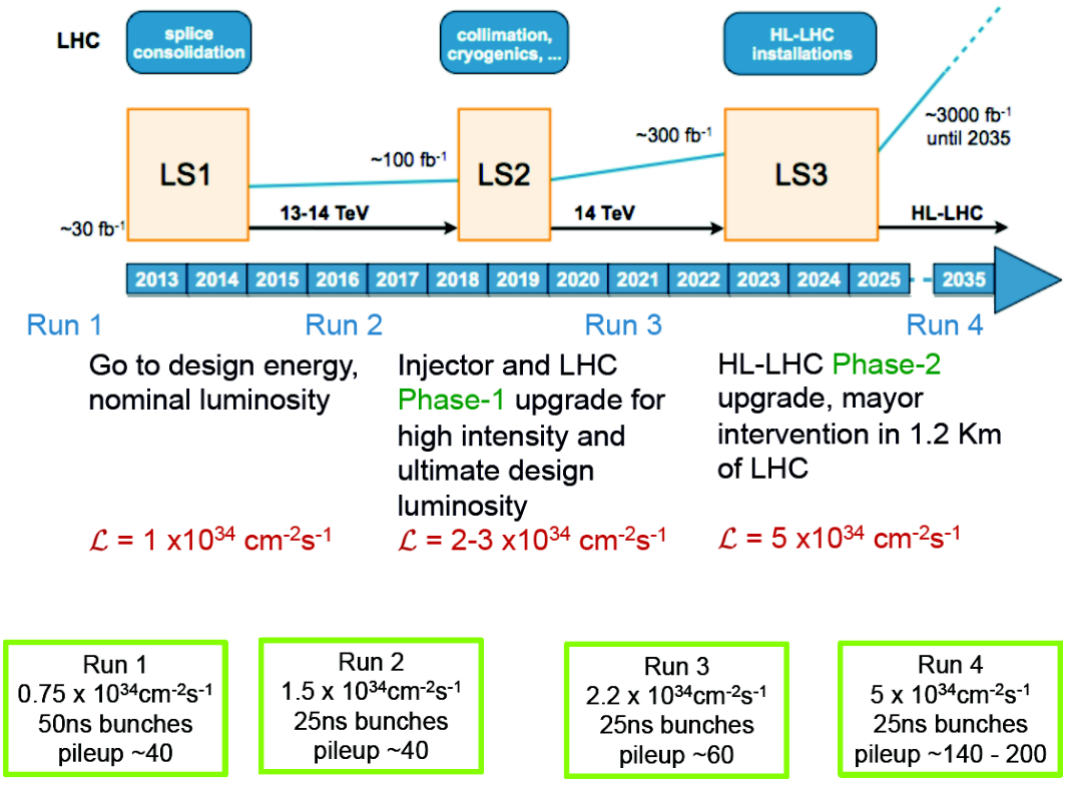


Fig. 2.4 Run periods within LHC Roadmap

To extend its discovery potential, the LHC will need a significant, extended over 10 years upgrade starting from 2020 to increase its luminosity (rate of collisions) by a factor of 10 beyond the original design value (from 300 to 3000 fb^{-1}).

Besides the HL schedule, CERN is now initiating an exploratory study for a future long-term project centered on a new-generation circular collider with a circumference of 80 to 100 kilometers. A worthy successor to the LHC, whose collision energies will reach 14 TeV, such an accelerator would allow particle physicists to push back the boundaries of knowledge even further. The Future Circular Collider (FCC) programme [30] will focus primarily on studies for a hadron collider, similar to the LHC, capable of reaching unprecedented energies in the region of 100 TeV. The FCC will thus run in parallel with another study that has already been underway for a number of years, the Compact Linear Collider, or "CLIC" [31] , another option for a future accelerator at CERN. The CLIC study aims to investigate the potential of a linear collider based on a novel accelerating technology.

ATLAS has planned a 3 stage program (Figure 2.5) to optimize the physics reach at each LHC phase. The goal is keeping similar performance in a harsher environment than that of the current detector (see Figure 2.5 for reference).

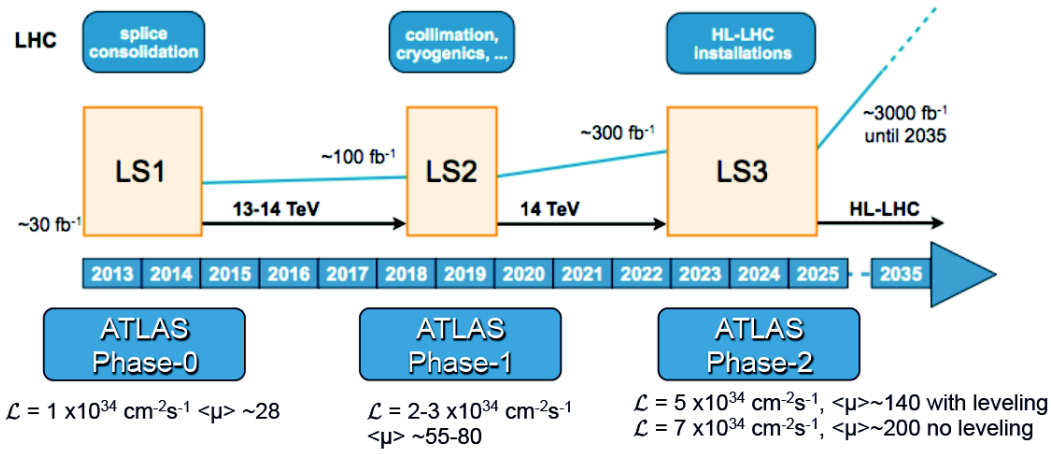


Fig. 2.5 ATLAS Roadmap

2.7 ATLAS experiment

ATLAS [3] is about 45 meters long, more than 25 meters high, and weighs about 7,000 tons. It consists of four major components (see Figure 2.8):

- inner detector (yellow) - measures the momentum of each charged particle
- calorimeter (orange and green) - measures the energies carried by the particles
- muon spectrometer (blue) - identifies and measures the momenta of muons
- magnet system (grey) - provide the bending of the charge particles, allowing the measurement the momentum out of the charge particle

ATLAS is a particle physics experiment at the Large Hadron Collider at CERN that is searching for discoveries in the head-on collisions of protons at energies in the center-of-mass of up to 14 TeV.

Beside the fundamental studies about the elementary forces of our Universe, ATLAS also attacks unknowns about extra dimensions of space, unification of fundamental forces, evidence for dark matter candidates in the Universe, etc. Following the discovery of the Higgs boson, further data will allow in-depth investigation of the boson's properties and thereby of the origin of mass [3] .

ATLAS schedule was the following:

1. Late 2009 $\rightarrow$ Startup of LHC and first event collisions at a total energy of 0.9 TeV and later at 2.36 TeV (above the previous world record from Tevatron).
2. March 2010 $\rightarrow$ Event collisions at a total energy of 7 TeV. This led to about eight months of data taking before a few weeks of heavy ion collisions and the usual winter shutdown. Many papers with early results have come as a result of the 2010 run.
3. March 2011 $\rightarrow$A year of intensive data taking was followed by a few weeks of heavy ion collisions and a winter shutdown. Event collisions at a total energy of 7 TeV.
4. April 2012 $\rightarrow$ Event collisions at a total energy of 8 TeV. A year of intensive data taking was followed by a few weeks of heavy ion collisions.
5. 2013 $\rightarrow$ A long shutdown to prepare for an increase of the total energy towards 13 TeV.

6. **Next 15-20 years** → Continued data taking with the publication of results on an ongoing basis.

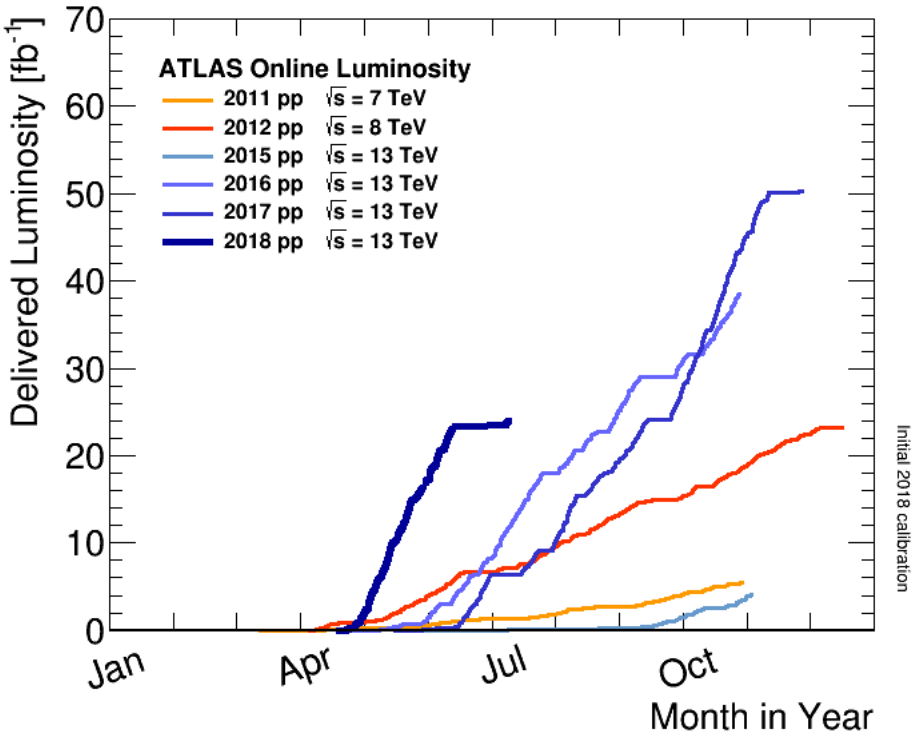


Fig. 2.6 Delivered Luminosity versus time for 2011-2018 (p-p data only)

The increased LHC collision energy to $\sqrt{s}$ = 13 TeV in 2015 opens up new research opportunities in Run 2, an era of precision physics and hope of discoveries. In Run 2, LHC bunches filled with protons collide at 40MHz, providing an instantaneous luminosity which peaked at $20.6 \times 10^{33} cm^{-2} s^{-1}$ (see Figure 2.6). This leads to a large number of p-p interactions (see Figure 2.7) leading to an increase of interesting signatures.

To fully exploit the new LHC capabilities, an upgrade program was launched during the long shutdown.

The ATLAS detector is now more efficient in detecting the tracks thanks to the fourth layer of pixels in its pixel tracker, the subdetector closest to the collisions and whose function is to reconstruct the particle trajectories. Improvements were also made to the muon detectors and calorimeters. Sections of the beam pipe were replaced to reduce the high radiation background noise. Moreover, more efficient trigger and data-acquisition systems were employed,

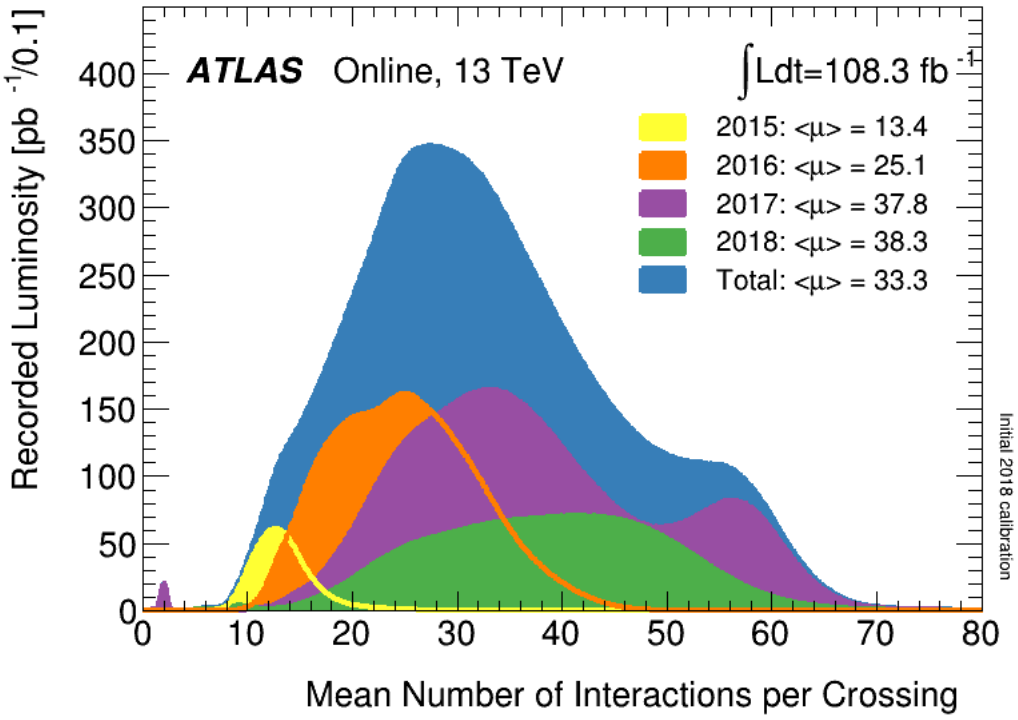


Fig. 2.7 Number of Interactions per Crossing, showing the 13 TeV data from 2015 - 2018

capable of recording a thousand events every second – more than double its capacity during Run 1. In parallel, upgrades on the simulation, reconstruction, and data-analysis software were carried out.

The basic design criteria of the detector include the following [59]:

- Very good electromagnetic calorimetry for electron and photon identification and energy measurement, complemented by full-coverage hadronic calorimetry for accurate jet and missing transverse energy measurements;
- High-precision muon momentum measurements, with the capability to guarantee accurate measurements at the highest luminosity using the external muon spectrometer alone;
- Efficient tracking at high luminosity for high-p_T lepton-momentum measurements, electron, and photon identification, (tau) τ-lepton and heavy-flavor identification, and full event reconstruction capability at lower luminosity;
- Large acceptance in pseudorapidity (η) with almost full azimuthal angle (Φ) coverage everywhere.
- Triggering (on missing energy as well) and measurements of particles at low-p_T thresholds, providing high efficiencies for most physics processes of interest at LHC.

The interactions in the ATLAS detectors creates an enormous amount of data-flow. To analyze this data we need:

- the trigger system which records 1000 interesting events per second on disk
- the data acquisition system that channels the data from the detectors to the storage
- the computing system in charge of analyzing 1000 Million events recorded per year

After Run 1 of the LHC between 2010 and 2013, the accelerator has been upgraded and prepared for running with higher collision energy and intensity in Run 2 (started in June 2015). The Run 2 conditions significantly increase the demands on detector trigger systems, requiring upgrades and new approaches to be employed [58]. Comparison of Run 1 and Run 2 conditions is shown in Table 2.2.

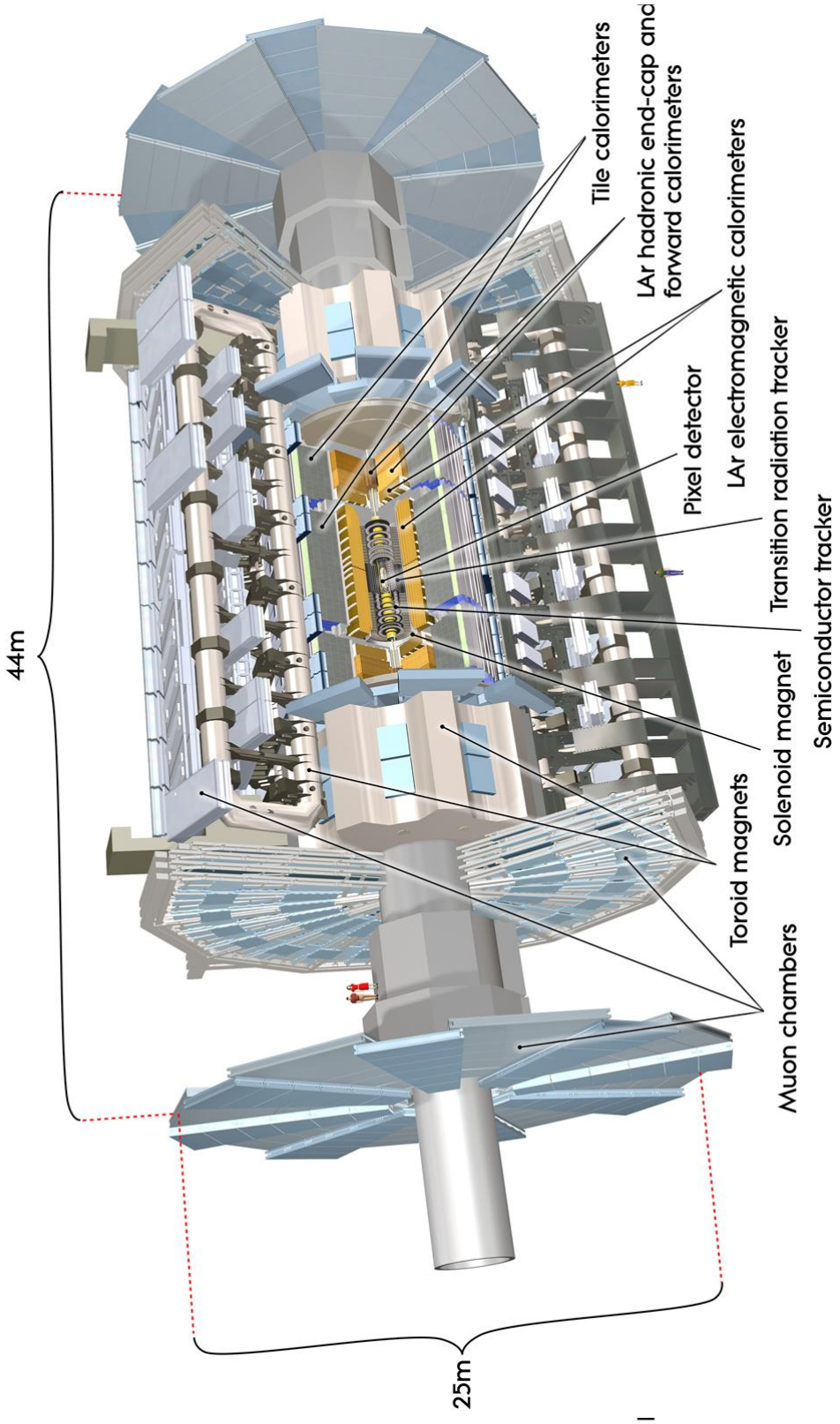


Fig. 2.8 ATLAS Detector

LHC parameter	Unit	Run 1	Run 2
E_{CM}	[TeV]	7	13
Bunch separation	[ns]	50	25
Peak luminosity	[cm^{-2} s^{-1}]	7×10^{33}	2×10^{34}
Interactions per crossing (at peak luminosity)	[interactions]	≈ 21	≈ 55
ATLAS detector input rate	[MHz]	20	40

Table 2.2 Comparisons of LHC run parameters for Run 1 and Run 2.

2.8 Particle identification

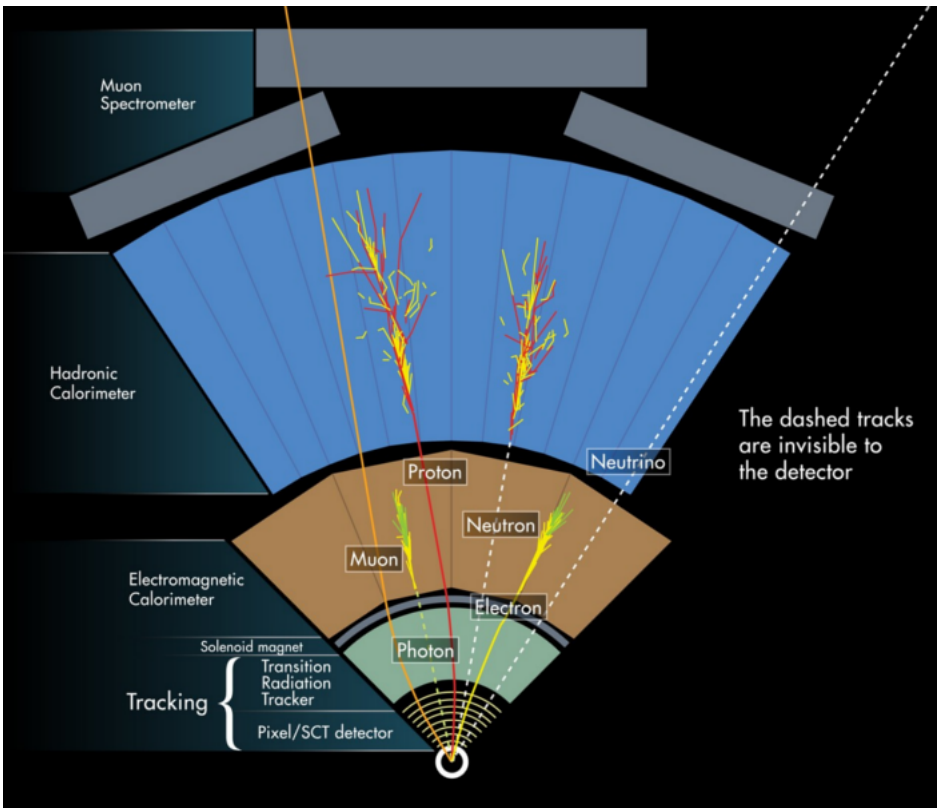


Fig. 2.9 Particle identification in LHC experiment.

Particle Identification (PID) is a crucial aspect of most High Energy Physics (HEP) experiments. Therefore, the reconstruction of events with high accuracy relies on:

- Tracking detectors which determine whether the particles are charged, and in conjunction with a magnetic field, measure the sign of the charge and the momentum of the particle

- Calorimeters which detect neutral and charged particles, measure the energy of particles and determine whether they have electromagnetic or hadronic interactions

Each particle gives a different signature in the ATLAS detector, making, therefore, their identification possible. The signatures for the most important particles are summarised in Figure 2.9.

An electron gives a signal in the inner detector, losing a small part of its energy while depositing its remaining energy in the calorimeter (mainly in the electromagnetic calorimeter).

Photons also give a signal in the calorimeter, but not in the inner detector (unless a photon converts into an electron-positron pair).

Muons traverse the calorimeter, giving a signal in the inner detector and muon spectrometer.

Charged hadrons give a signal in the inner detector and the calorimeters. Hadrons shower gives a signal into the calorimeter in both the electromagnetic and hadronic calorimeter.

Jets are a combination of hadronic and leptonic particles, causing a signal in the inner detector and calorimeters and eventually also in the muon spectrometer.

Neutrinos cannot be detected and leave only a signal of missing transverse energy.

2.9 ATLAS sub-detectors

The dimensions of the ATLAS sub-detectors are summarized in Table 2.3.

The ATLAS detector consists of two types of detector components:

- tracking detectors, which measure the trajectory of a charged particle,
- calorimeters, which measure the energy of a particle.

.

The ATLAS detector is split into a barrel part, where detector layers are positioned on cylindrical surfaces around the beam axis, and two end-cap parts, where detector layers are

component	**radius [m]**	**length [m]**	**η-coverage**
barrel muon spectrometer	11	26	$\|\eta\| < 1.4$
end-cap muon spectrometer	11	2.8	$1.1 < \|\eta\| < 2.8$
barrel hadronic calorimeter	4.25	12.2	$\|\eta\| < 1.0$
end-cap hadronic calorimeter	2.25	2.25	$1.5 < \|\eta\| < 3.2$
barrel em-calorimeter	2.25	6.42	$\|\eta\| < 1.4$
end-cap em-calorimeter	2.25	0.63	$1.4 < \|\eta\| < 3.2$
forward/backward calorimeter	integrated in end-cap	integrated in end-cap	$3.1 < \|\eta\| < 4.9$
barrel + end-cap inner detector	1.15	6.8	$\|\eta\| < 2.4$

Table 2.3 Comparisons of LHC run parameters for Run 1 and Run 2.

positioned in planes of constant z perpendicular to the beam pipe.

The calorimeter also consists of a forward and a backward part, extending up to a pseudorapidity of $|\eta| = 4.9$.

In the following subsections, a detailed description of the ATLAS sub-components is provided.

2.9.1 ATLAS Inner Detector

The inner detector is the first part of the ATLAS detector and measures the direction, momentum, and charge of electrically-charged particles produced in each proton-proton collision. Very compact and highly sensitive, it consists of three different systems of sensors all immersed in a magnetic field of ≈ 2 Tesla (T) parallel to the beam axis.

The main components of the Inner Detector (Figure 2.10) are: Pixel Detector, Semiconductor Tracker (SCT), and Transition Radiation Tracker (TRT).

The ATLAS Insertable B-Layer (IBL) (Figure 2.11) [53] is the innermost layer of pixel detectors. Installed in May 2014 at a radius of 3.3 cm from the beam axis, between the existing pixel detector and a new smaller radius beam-pipe, it is the first large scale application of 3D sensors and CMOS 130 nm technology also built to cope with high radiation and occupancy.

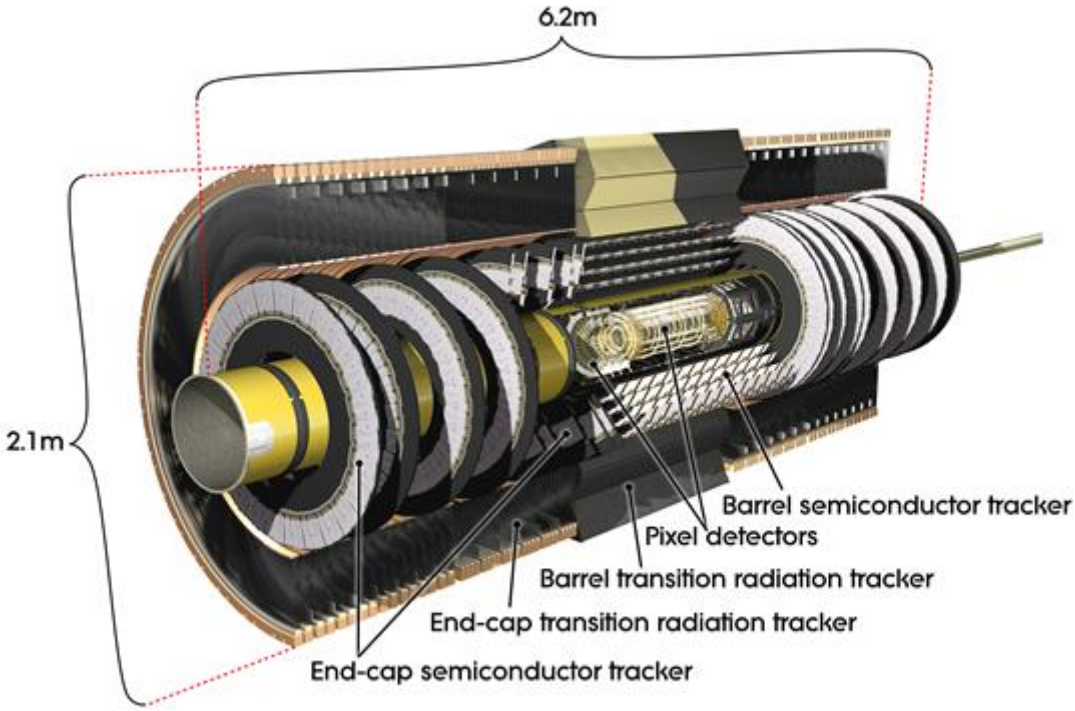


Fig. 2.10 ATLAS inner detector

Fig. 2.11 The ATLAS IBL detector prior the insertion

The Pixel Detector consists of 80 million pixels (80 million channels) of a total area of 1.7 m^2, each pixel size being 50 $\times$ 400 μm^2 with a resolution of 14 $\times$ 115 μm^2. The barrel has 1,744 modules (10 cm^2) with 46,080 readout channels per module while the three-pixel disks (in each endcap) have 6.6 million channels.

The Semiconductor Tracker is a silicon microstrip tracker consisting of 4,088 two-sided modules and over 6 million implanted readout strips (6 million channels). Readout strips at every 80 μm^2 on the silicon, allow the positions of charged particles to be recorded to an accuracy of 17 μm^2 per layer (in the direction transverse to the strips).

The Transition Radiation Tracker consists of 350,000 read-out channels of total volume 12 m^3 that allows precision measurement of 0.17 mm (particle track to the wire) and provides additional information on the particle type that flew through the detector (example if it is an electron or pion).

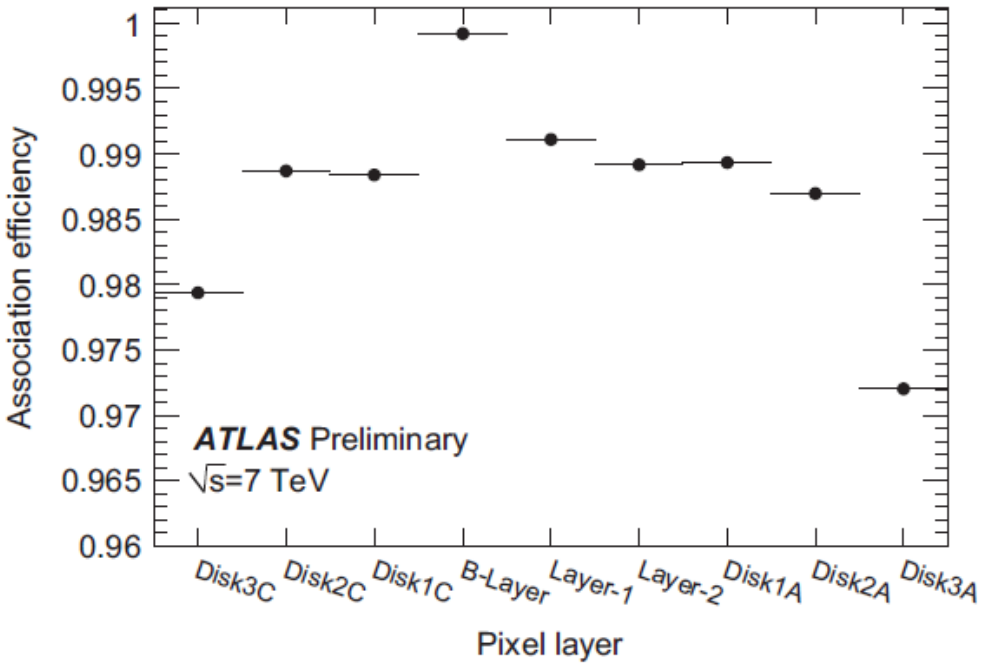


Fig. 2.12 Efficiency of tracks having hits associated with the different pixel layers.The B-Layer efficiency is 100% due to the track requirement. Lower outer disk efficiencies arise from a higher percentage of dead pixels.

The Pixel Detector into the ATLAS Detector delivered in the 2012 run 99.9% good quality data [18]. The efficiency of tracks having associated hits in the different Pixel Detector layers is shown in Figure 2.12 and is around 99%. The slightly lower outer disk efficiency is due to the higher percentage of dead pixels in these modules, which have been placed there on purpose. The track resolution is significantly improved by the analog readout of the charge compared to a binary readout.

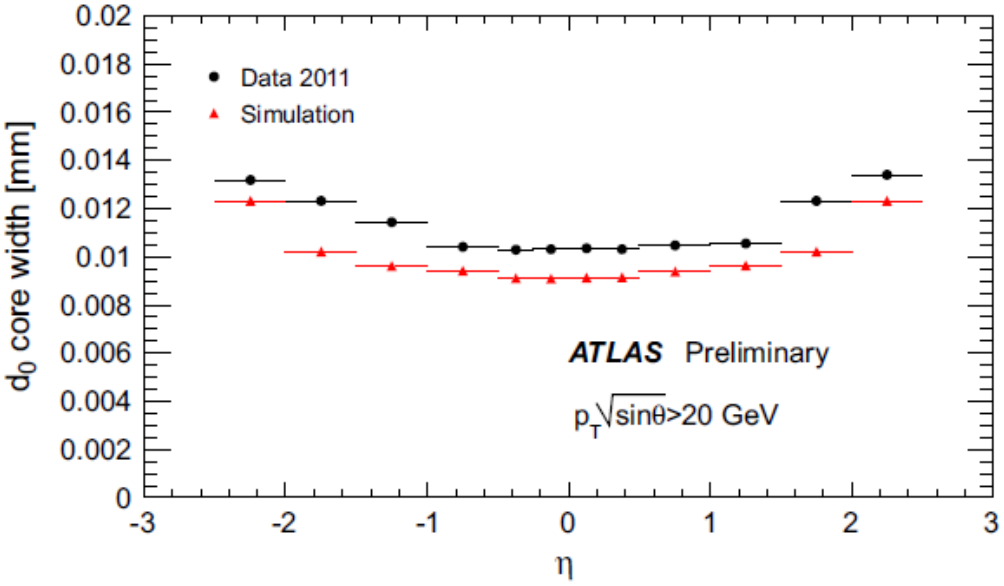


Fig. 2.13 Transverse impact parameter d_0 resolution for 2011 data (back points) and simulation (red triangles).

The overall tracking performance can be described by a transverse impact parameter d_0 resolution as shown in Figure 2.13, where a precision of $\sigma(d_0)$ of the order of 10 μm over the complete pseudo-rapidity range is achieved. This high precision vertexing is crucial for successful and highly efficient, b-tagging, a key experimental feature in most of the physics analysis.

One of the operational problems of the pixel detector is the desynchronization of modules with respect to the Read Out Driver, making the data of these modules unusable for reconstruction. To synchronize the module, a configuration ROD-level monitoring was implemented. The improvement is shown in Figure 2.14. While the ROD is reconfiguring the desynchronized module, data taking is continued normally on all other modules.

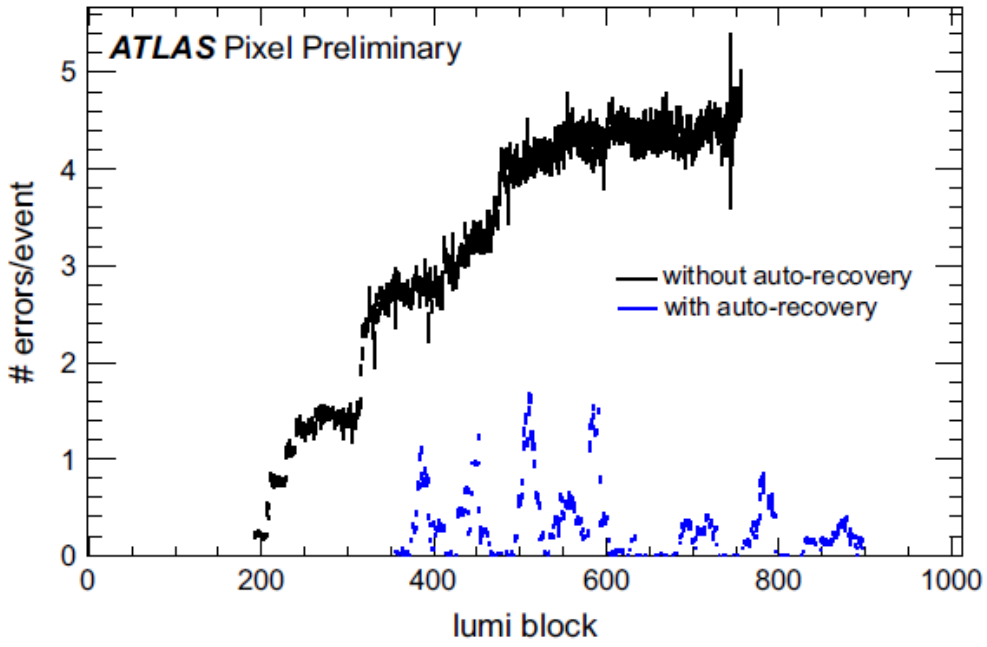


Fig. 2.14 Number of desynchronization errors over the cause of one luminosity block before and after implementation of a ROD-level monitoring.

2.9.2 The Calorimeter

Calorimeters measure the energy the particle losses as it passes through the detector. It is usually designed to stop entire or "absorb" most of the particles coming from a collision, forcing them to deposit all of their energy within the detector. Calorimeters typically consist of layers of "passive" or "absorbing" high-density material.

Electromagnetic calorimeters measure the energy of electrons and photons as they interact with matter. Hadronic calorimeters sample the energy of hadrons (particles that contain quarks, such as protons and neutrons) as they interact with atomic nuclei. Calorimeters can stop most known particles except muons and neutrinos.

The components of the ATLAS calorimetry system are: the Liquid Argon (LAr) Calorimeter, shown in Figure 2.15 (left) and, the Tile Hadronic Calorimeter shown in Figure 2.15 (right).

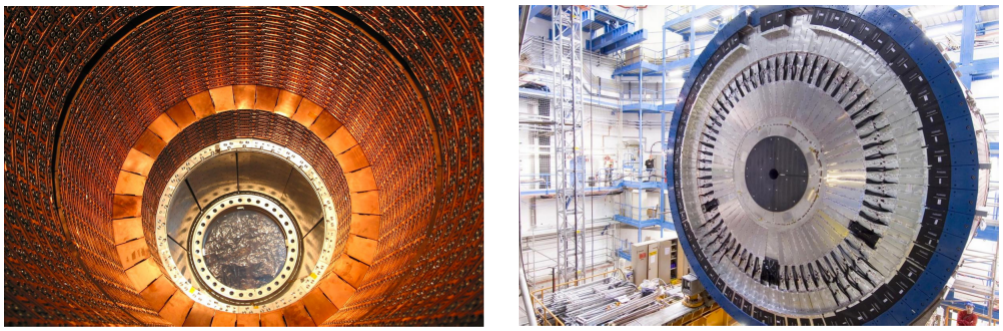


Fig. 2.15 ATLAS LAr Calorimeter (left) and Tile Hadronic Calorimeter (right).

The Liquid Argon (LAr) Calorimeter

ATLAS EM calorimetry [11] is based on liquid argon calorimeters and consists of an accordion-type calorimeter for $|\eta|$ <3.2 and a compact design in the forward region.

Accordion calorimeter The main part of ATLAS EM calorimeter is a lead–liquid argon (LAr) sampling detector with accordion-shaped electrodes and lead absorber plates over its full coverage. The calorimeter is divided into a Barrel part and two End-Caps (see Table 2.4).

Accordion calorimeter	**Barrel**	**End-Cap**
Coverage	$\|\eta\|$<1.475	1.375<$\|\eta\|$<3.2
Longitudinal segmentation	3 samplings	2–3 samplings
Granularity ($\Delta\eta \times \Delta\phi$) Sampling 1	0.003 × 0.1	0.003 × 0.1
Sampling 2	0.025 × 0.025	0.025 × 0.025
Sampling 3	0.05 × 0.025	0.05× 0.025
Lead thickness (mm)	1.1-1.5	1.7-2.2
LAr gap thickness (mm)	2.1	0.9-3.1
nb of channels	102,400	62,208 per End-Cap
Presampler	Barrel	End-Cap
Coverage	$\|\eta\|$<1.475	1.5<$\|\eta\|$<1.8
Longitudinal segmentation	1 sampling	1 sampling
Granularity ($\Delta\eta \times \Delta\phi$)	0.025 × 0.1	0.025 × 0.1
nb of channels	7808	768 per End-Cap

Table 2.4 EM accordion calorimeter and presampler main parameters.

There are three longitudinal samplings for $|\eta|$ <2.5 and two for 2.5<$|\eta|$ <3.2. The first sampling consists of η strips, to allow rejection of π_0 and jet background. This geometry offers a complete ϕ symmetry without azimuthal crack. The active depth of the three samplings are about $6X_0$; $16X_0$ and $3X_0$; respectively, at $\eta = 0$.

In the range $|\eta|$ <1.8, a presampler continues the calorimeter structure to recover the energy lost in the upstream material (cryostat, superconducting coil, inner detector, etc.). The active depth of the presampler is 11 mm of liquid argon in the Barrel and 4 mm in the End-Cap (which amounts to $0.08X_0$ and $0.03X_0$ respectively).

In the forward region (3.1<$|\eta|$<4.9), the EM calorimetry is composed by the Forward Calorimeter (FCAL) which consists of copper rods parallel to the beam axis inside an outer tube. The EM Forward Calorimeter main parameters are listed in Table 2.5. FCAL has to cope with extremely high radiation levels (doses up to 5×10^6 Gy/yr), having an active depth of the EM module of about $28X_0$.

Coverage	3.1<$\|\eta\|$<4.9
Longitudinal segmentation	1 samplings
Granularity ($\Delta\eta \times \Delta\phi$)	$\approx 0.1 \times 0.1$
nb of channels	1024 per End-Cap

Table 2.5 EM Forward Calorimeter main parameters.

The performance of the Liquid Argon (LAr) Calorimeter [36] has been excellent throughout the commissioning and first two years of data-taking. In 2011, LAr data were of sufficiently good quality for data analysis for 98% of the data collected during stable beam periods.

Periodic studies on the collision data also ensure that the whole LAr calorimeter system is uniformly aligned in time. A timing alignment for all readout channels better than 1ns was then measured (Figure 2.16 left). Moreover, studies using reconstructed electrons from $W \to e\nu$ candidates have demonstrated that a 300 ps medium layer channel timing resolution can be achieved for large energy deposit in the EMB as shown in Figure 2.16 right.

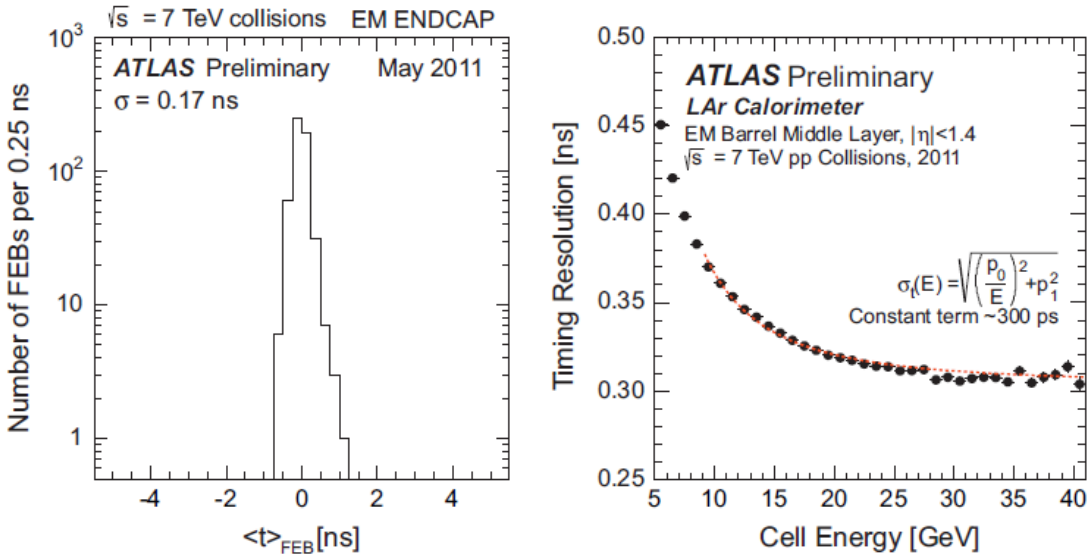


Fig. 2.16 Left: Average time per FEB in the EMEC computed with 7 TeV collision data on May 2011. Right: Single channel time resolution as function of cell energy.

The Hadronic Calorimeter

In the range $|\eta|$<1.6, the ATLAS hadronic calorimeter is an iron-scintillating tiles calorimeter, followed in $|\eta|$>1.6 regions by a the LAr calorimeter like structure.

Hadronic Tile	**Barrel**	**Extended Barrel**
Coverage	$\|\eta\|$<1.0	0.8<$\|\eta\|$<1.7
Longitudinal segmentation	3 samplings	3 samplings
Granularity ($\Delta\eta \times \Delta\phi$) Sampling 1 and 2	0.1 × 0.1	0.1 × 0.1
Sampling 3	0.2 × 0.1	0.2× 0.1
nb of channels	5760	1792 per Extended Barrel

Table 2.6 Hadronic Tile calorimeter main parameters.

The Hadronic Tile calorimeter is a sampling calorimeter, located behind the solenoid coil and the EM calorimeter. It uses iron as absorber material and scintillating tiles as active material. The calorimeter is split in a Barrel and two Extended Barrel parts. The Hadronic Tile calorimeter main parameters are listed in Table 2.6.

Coverage	1.5<$\|\eta\|$<3.2
Longitudinal segmentation	4 samplings
Granularity ($\Delta\eta \times \Delta\phi$) 1.5<$\|\eta\|$<2.5	0.1×0.1
Granularity ($\Delta\eta \times \Delta\phi$) 2.5<$\|\eta\|$<3.2	0.2×0.2
nb of channels	3072 per End-Cap

Table 2.7 The Hadronic End-Cap calorimeter main parameters.

The Hadronic End-Cap calorimeter (HEC) is a LAr sampling calorimeter which provides hadronic coverage for 1.5<$|\eta|$<3.2 (see its parameter in Table 2.7). It has parallel Cu plate absorbers orthogonal to the beam axis and consists of two consecutive wheels with absorber thickness of 25 and 50 mm. The thickness of the active part of the calorimeter is about 12 interaction length.

Coverage	3.1<$\|\eta\|$<4.9
Longitudinal segmentation	2 samplings
Granularity ($\Delta\eta \times \Delta\phi$)	0.2×0.2
nb of channels	768 per End-Cap

Table 2.8 Hadronic Forward calorimeter main parameters.

In the forward region (3.1<$|\eta|$<4.9), the hadronic calorimetry (see Table 2.8 for its parameters) is continued by the hadronic calorimeter, equivalent of the EM FCAL. It is located just behind the EM FCAL module, which amounts for 2.6 interaction length. Total active length of the FCAL calorimeter is 9.5 interaction length. A passive shield is located behind the FCAL calorimeter at the transition with the most forward muon chambers of ATLAS.

Figure 2.17 shows the three layers and the dimensions of the optimized cells to obtain a structure of projective towers.

The Tile Hadronic Calorimeter performance [35] with cosmic muons is shown in Figure 2.18 where the difference between the time offsets as measured with cosmic muons and with 2008 single beam splash events. It shows an agreement at the level of 1 ns, confirming the goodness of the analysis (Figure 2.18 left). The comparison between cosmic data and Monte-Carlo prediction is shown in Figure 2.18 right. A global uniformity of the response of cells of BC layer as a function of pseudorapidity and azimuth angle within 3% is observed

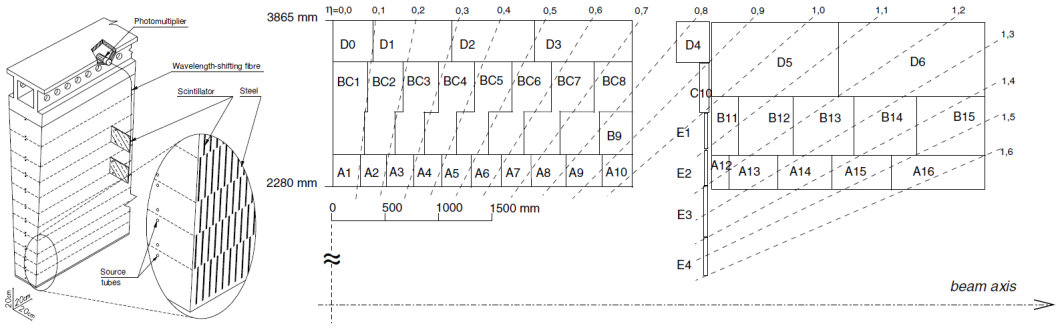


Fig. 2.17 Schematic of one of the 64 azimuthal modules of TileCal (left) showing the system of signal collection and drawing of half of the calorimeter divided into a barrel and an extended barrel part with the cell division scheme depicted (right).

for all the layers.

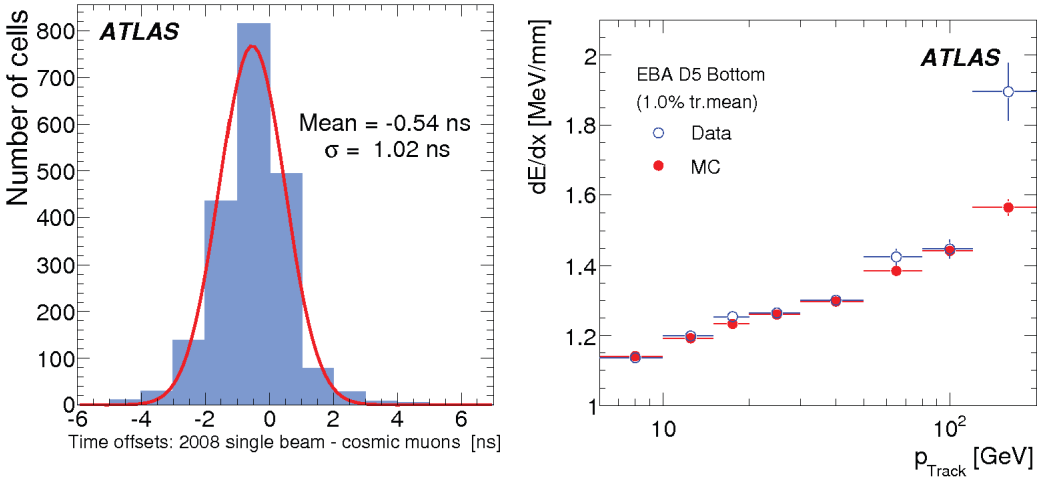


Fig. 2.18 Left: difference between time offsets from 2008 test beam and cosmic muons analysis. Right: muon energy loss per path length dE/dx as a function of momentum of the track (measured in the Inner Detector).

The performance with collision data is shown in Figure 2.19 as the uniformity of cell response to first collisions at a center-of-mass energy of 7 TeV. The energy scale uncertainty, which was successfully extrapolated from test beam to ATLAS, is conservatively considered to be 4%. The time synchronization between cells is well below 1 ns and has been verified with a single beam and cosmic muons.

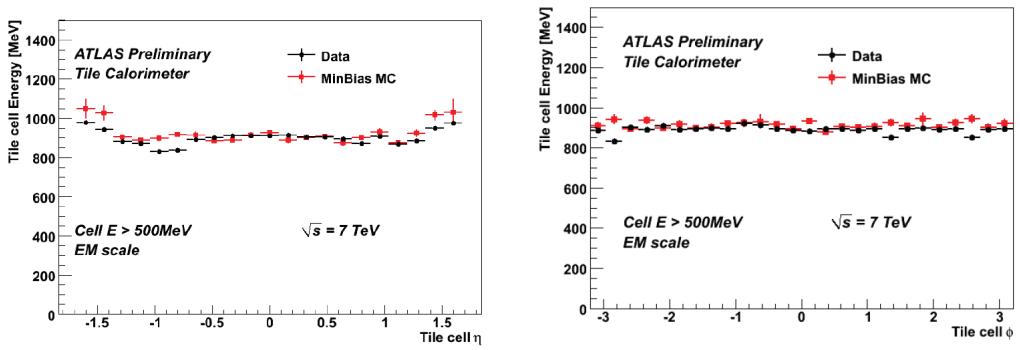


Fig. 2.19 Average energy deposition in TileCal cells as a function of η (left) and ϕ (right), for data and MonteCarlo

2.9.3 Muon Spectrometer

Muons are particles that usually pass through the Inner Detector and Calorimeter undetected. The muon spectrometer is made up of 4,000 individual muon chambers using four different technologies to identify and measure the momenta of muons.

Subsections of the Muon System are: Thin Gap Chambers Figure 2.20 (top left), Resistive Plate Chambers Figure 2.20 (top right), Monitored Drift Tubes Figure 2.20 (bottom left), and Cathode Strip Chambers Figure 2.20 (bottom right).

Thin Gap Chambers consists of 440,000 channels used for triggering and second coordinate measurement (non-bending direction) at ends of the detector.

Resistive Plate Chambers of electric field 5,000 V/mm are used for triggering and second coordinate measurement in the central region.

Monitored Drift Tubes measure curves of tracks at a resolution of 80 μm and it consists of 1,171 chambers with total 354,240 tubes (3 cm diameter, 0.85-6.5 m long).

Cathode Strip Chambers measure precision coordinates at ends of the detector with a resolution 60 μm, making use of 70,000 channels.

The muon reconstruction efficiency [37] is close to 99 % over most of the pseudorapidity range of $|\eta| < 2.5$ for $p_T > 5$ GeV. The $Z \to \mu\mu$ sample enables a measurement of the efficiency with a precision at the 0.2 % level for $p_T > 20$ GeV. The $J/\Psi \to \mu\mu$ sample

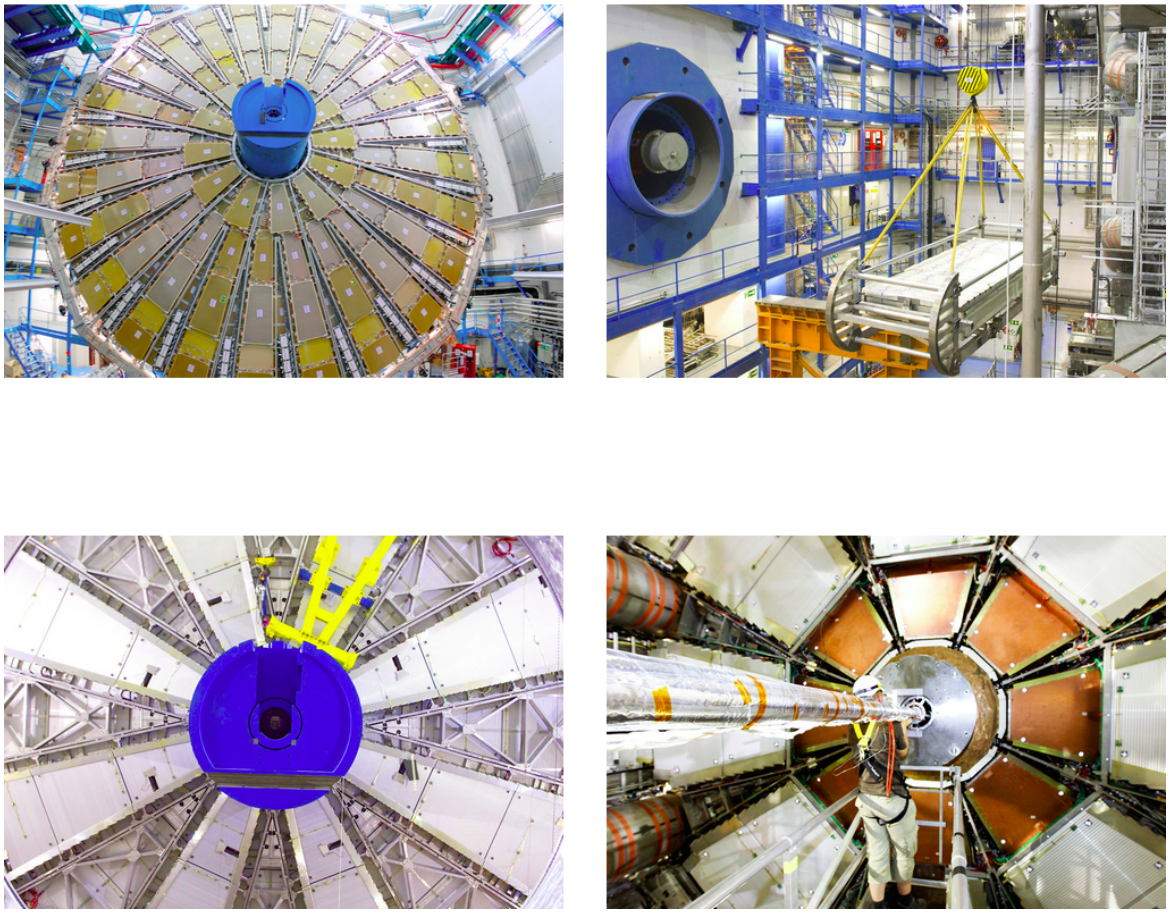


Fig. 2.20 ATLAS Muon Spectrometer subsections.

provides a measurement of the reconstruction efficiency between 5 and 20 GeV with a precision better than 1 %.

The isolation efficiency varies between 93 and 100 % depending on the selection and on the momentum of the particle and is well reproduced in the simulation.

The muon momentum scale and resolution have been studied in $J/\Psi \to \mu\mu$ and $Z \to \mu\mu$ decays. For $Z \to \mu\mu$ decays, the uncertainty in the momentum scale varies from a minimum of 0.05% for $|\eta| < 1$ to a maximum of 0.3 % for $|\eta| \approx 2.5$. This corresponds, after applying momentum corrections, to the p_T resolution in data and simulation agreement to better than 5 % for most of the η range.

2.9.4 Magnet System

ATLAS Magnetic System features a unique hybrid system of large superconducting magnets of 22 m in diameter and 26 m in length, with a stored energy of 1.6 GJ.

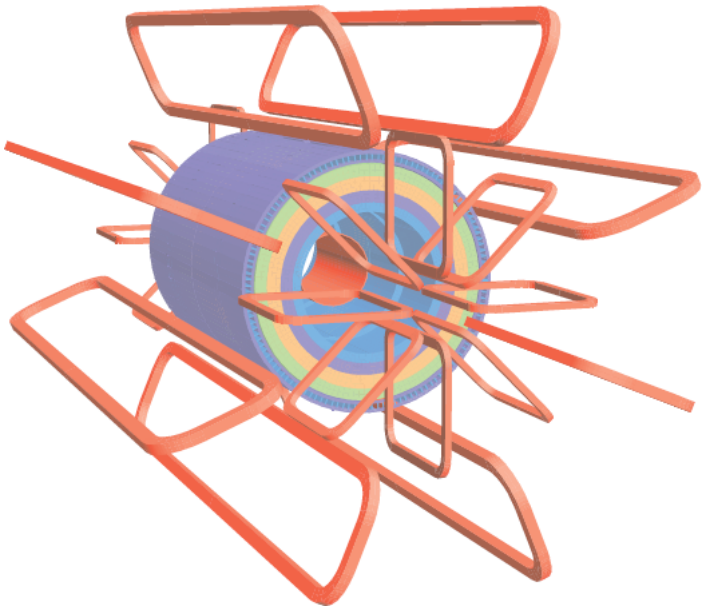


Fig. 2.21 Geometry of magnet windings and tile calorimeter steel.

Figure 2.21 shows the general layout, the four main layers of detectors and the four superconducting magnets which provide the magnetic field over a volume of approximately

12,000 m^3 (defined as the region in which the field exceeds 50 mT).

Property	Feature	Unit	Solenoid	Barrel toroid	End-cap toroids
Size	Inner diameter	m	2.46	9.4	1.65
	Outer diameter	m	2.56	20.1	10.7
	Axial length	m	5.8	25.3	5.0
	Number of coils		1	8	2×8
Mass	Conductor	t	3.8	118	2×20.5
	Cold mass	t	5.4	370	2×140
	Total assembly	t	5.7	830	2 x 239
Coils	Turns per coil		1154	120	116
	Nominal current	kA	7.73	20.5	20.5
	Magnet stored energy	GJ	0.04	1.08	2 x 0.25
	Peak field in the windings	T	2.6	3.9	4.1
	Field range in the bore	T	0.9–2.0	0.2–2.5	0.2–3.5
Conductor	Overall size	mm^2	30 x 4.25	57 x 12	41 x 12
	Ratio Al:Cu:NbTi		15.6:0.9:1	28:1.3:1	19:1.3:1
	Number of strands (NbTi)		12	38–40	40
	Strand diameter (NbTi)	mm	1.22	1.3	1.3
	Critical current (at 5 T and 4.2 K)	kA	20.4	58	60
	Operating/critical-current ratio at 4.5 K	%	20	30	30
	Residual resistivity ratio (RRR) for Al		> 500	> 800	> 800
	Temperature margin	K	2.7	1.9	1.9
	Number of units $\times$ length	m	4×2290	$8 \times 4 \times 1730$	$2 \times 8 \times 2 \times 800$
	Total length (produced)	km	10	56	2 x 13
Heat load	At 4.5 K	W	130	990	330
	At 60–80 K	kW	0.5	7.4	1.7
	Liquid helium mass flow	g/s	7	410	280

Fig. 2.22 Main parameters of the ATLAS magnet system.

The ATLAS magnet system, whose main parameters are listed in Table 2.22, consists of:

- a solenoid which is aligned on the beam axis and provides a 2 T axial magnetic field for the inner detector, while minimizing the radiative thickness in front of the barrel electromagnetic calorimeter;

- a barrel toroid and two end-cap toroids, which produce a toroidal magnetic field of approximately 0.5 T and 1 T for the muon detectors in the central and end-cap regions, respectively.

2.9.5 Trigger and Data Acquisition System

ATLAS is designed to observe up to one billion proton-proton collisions per second ($\sim$ 60 million megabytes per second). However, only a few percentages of these events will contain interesting characteristics that might lead to discoveries and worth analyzing. To reduce the flow of data to manageable levels, ATLAS uses a specialized multi-level computing system - the Trigger System - which selects events with distinguishing characteristics that make them interesting for physics analyses.

In Run 2 [54], the trigger system consists of a hardware-based first level trigger (Level-1) and a software-based high-level trigger (HLT).

The Level-1 trigger uses custom electronics to determine Regions-of-Interest (RoIs) in the detector, taking as input coarse granularity calorimeter and muon detector information. The Level-1 trigger reduces the event rate from the LHC bunch crossing rate of approximately 30MHz to 100 kHz. The decision time for a Level-1 accept is 2.5 μs.

The HLT reduces the rate from the Level-1 output rate of 100 kHz to approximately 1 kHz on average within a processing time of about 200 ms.

A schematic overview of the upgraded ATLAS trigger and data acquisition system is shown in Figure 2.23.

The ATLAS trigger system was successfully deployed and commissioned during the early data-taking Run 1 phase. It has successfully operated over six orders of magnitude rise in luminosity recording data with an efficiency exceeding 95%.

2.9.6 The Computing System

The ATLAS Computing System analysis the data produced by the ATLAS detector. To improve the performance of the detector, continuous work on developing and improving computing software is done to process and analyze the vast amounts of collision data.

All members of the ATLAS Collaboration have equal access possibilities to all ATLAS data, independently of their geographical location (over 130 computing centers worldwide), thanks to the Worldwide LHC Computing Grid [60].

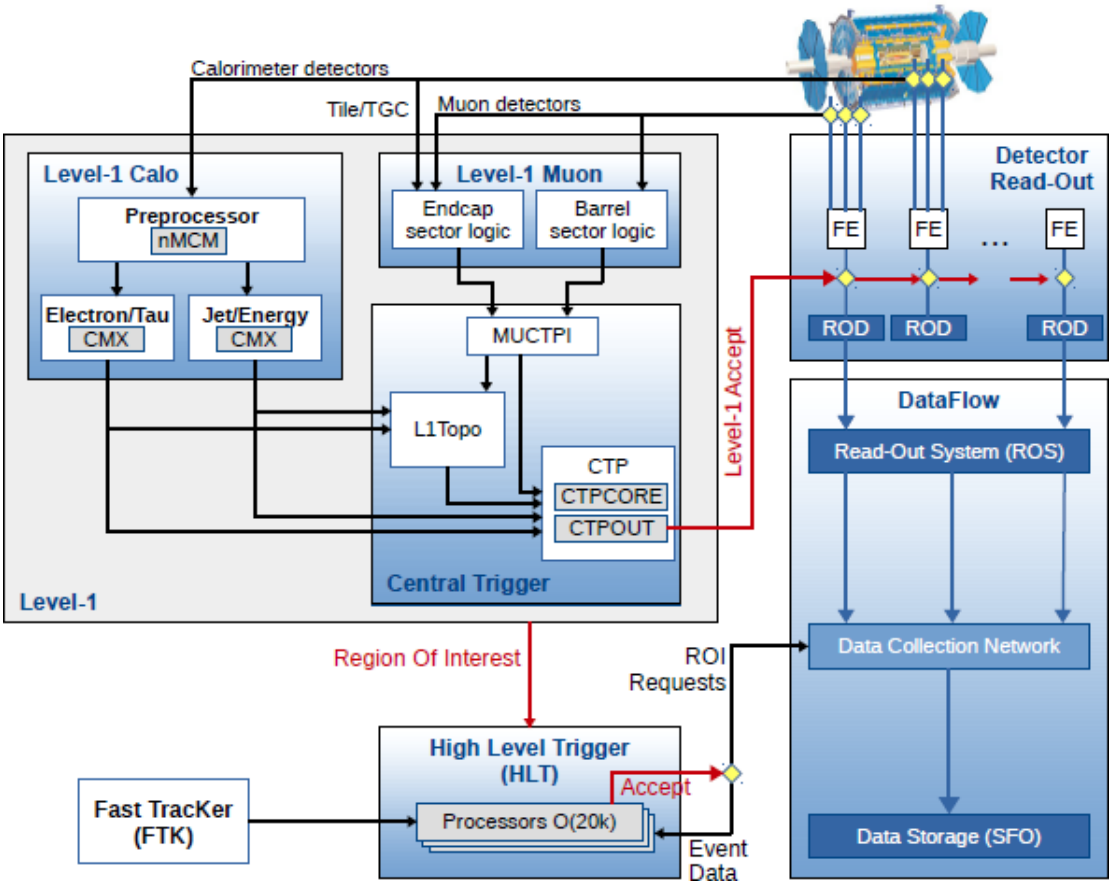


Fig. 2.23 ATLAS trigger and data acquisition system in Run-2.

Established in 2007, the ATLAS Distributed Computing project aims to deliver infrastructure for the needs of the ATLAS experiment in terms of data handling and data processing. It consists of two main working areas: the software development and computing operations.

The ATLAS distributed computing system relies on resources spread over approximately 100 computer centers around the world and embraced the Grid paradigm to enable data movement and data processing across resources.

The ATLAS data management system was capable of delivering an aggregated traffic exceeding 10 GB/s, without the loss of even a single RAW event, with availabilities larger than 99.9% over the last several years and a very high level of automation.

The LHC gives 1 PB/s of data generated by the detectors, up to 100 PB/year of stored data. To handle those data, a distributed computing infrastructure of half a million cores working 24/7 was set in place with an average of 40M jobs/month. A continuous data transfer rate of 6 GB/s (600TB/day) across the Worldwide LHC Grid (WLCG) is then analyzed by the scientist.

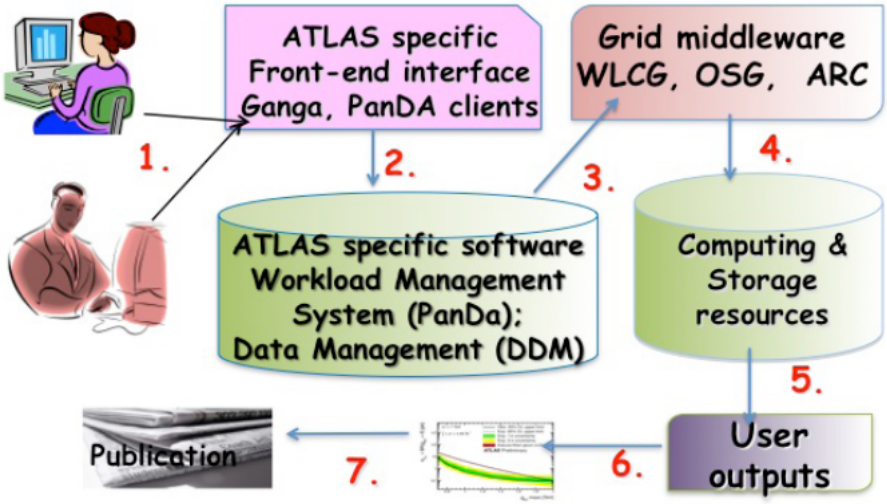


Fig. 2.24 ATLAS Distributed data Analysis workflows.

The ATLAS Workload Management System builds on top of the core service for Production and Distributed Analysis (PanDA). The system is based on pilot jobs with a central queue. PanDA consists of a workflow engine for job brokering and dispatching to the distributed computing resources, based on the concept of late binding of jobs to CPUs. ATLAS

Distributed Data Analysis (DA) is based on a complex and evolving system with many components interacting with each other. Figure 2.24 summarizes the various workflow paths, from user submission to DA until the results publication [17].

The Run 2 computing system is robust, mature, efficient and much lighter in operational effort than the system used during Run-1, easily coping with the high LHC performance. Each subsystem was tested to sustain 5-10 times larger load than currently used in Run 2. However, in 2016, the processing power, storage capacities and network were used to their limits. Therefore, careful planning of CPU and storage usage for further running campaigns will be necessary, cleaning the unused data more aggressively and better optimizing the network usage to avoid the saturation.

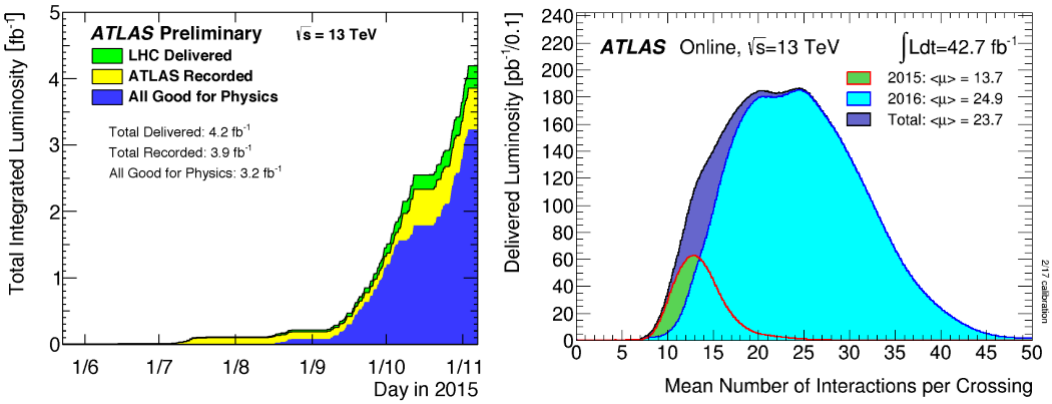


Fig. 2.25 Total integrated luminosity recorded by ATLAS in 2015 (left) corresponding to the Number of Interactions per Crossing (pileup) of 17-20 (right, Green color).

2.10 ATLAS Monte-Carlo Simulations

In this section, I will present the Monte Carlo simulations [72] of the ATLAS detector. The emphasis will be on the physics processes and the algorithms involved while the actual software implementation will not be treated.

Every detector simulation has to start with a description of the detector geometry. For the ATLAS detector, almost every detail is included in the detector description. The GEANT4 [71] environment is used for the geometry description and the full detector simulation in-

cluding the tracing of the particles and the electronics response of the active detector elements.

For the Inner Detector, it is essential to accurately simulate the most important physics processes: multiple scattering, continuous energy loss, bremsstrahlung for electrons, conversions for photons and nuclear interactions for the hadrons. It is thus necessary to have both the correct distribution and composition of the materials.

In the calorimeters, the accurate simulation of the shower processes is the most important. The showers are simulated by following all the particles created in the showering process and then particles are traced down to an energy of 100 keV at which point they stop and deposit their remaining energy.

Next step is the digitization phase: for each sub-detector, there is a model of how the ionization in the active detector element is changed into the digital output of the readout electronics. The simulated data should at this point be equivalent in format to the data that will eventually be recorded with the ATLAS detector. The simulation of noisy and dead channels in the electronics is also a part of the digitization phase.

The reconstruction step in a detector simulation involves the reconstruction of the kinematic information and particle identification. For the Inner Detector and the muon system, tracks are reconstructed from the hits in the individual detector elements. In the calorimeter cells with deposited energy are grouped in clusters. At a later stage in the reconstruction, all information can be combined to obtain the kinematic information of the event.

2.11 ATLAS Physics Objects: Performance at 13 TeV

In the interval of spring and summer 2015, the ATLAS experiment recorded data of a total integrated luminosity of 133 pb^{-1} and 127 pb^{-1} at a center-of-mass energy of 13 TeV. The bunch-crossing time intervals varied from 50 ns and 25 ns afterward with a pileup, from 20 to 17, respectively. The overall data taking efficiency was close to 90% (see Figure 2.25). This data set is used for the performance studies presented below.

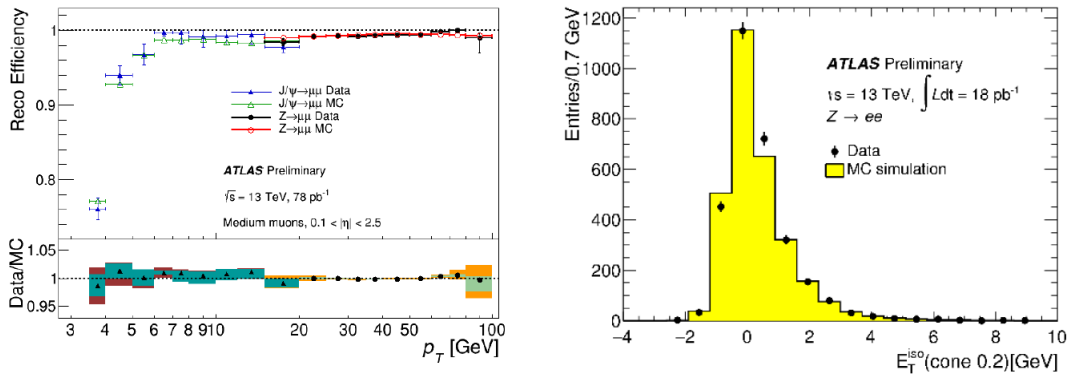


Fig. 2.26 Left: The muon reconstruction efficiency for both data and MC simulation; Right: The electron isolation energy in a cone of radius of 0.2 is shown for both data and MC simulation.

2.11.1 Electron and muon performance

Electron identification efficiencies are measured using a tag-and-probe method. This method will be described in details in Chapter 3. The efficiencies vary between 75% and 95% depending on the electron candidate energy, with an uncertainty of 3 to 5% [19]. The left plot in Figure 2.26 shows the reconstruction efficiency for muons identified in the central region of the detector using the medium criteria as a function of the transverse muon momentum for both data and Monte Carlo, for all tag and probe methods. The muon momentum scale is understood with a precision of $\sim 0.2\%$.

An essential feature of any analysis that uses leptons is the isolation of the lepton. The calo-based electron isolation is defined as the sum of clusters' transverse energies in a cone of a radius of 0.2 around the electron after excluding cells corresponding to the electron (see the right plot of Figure 2.26).

2.11.2 Jet and missing transverse energy performance

An efficient reconstruction of jets plays a vital role in many physics analysis. ATLAS reconstructs jets using the anti-kt algorithm with a radius parameter of 0.4 starting from topological clusters. This is the most common algorithm used by ATLAS collaboration. Jets are calibrated after applying corrections as a function of their transverse momentum and pseudo-rapidity. Overall, a good agreement with Monte Carlo simulations is found.

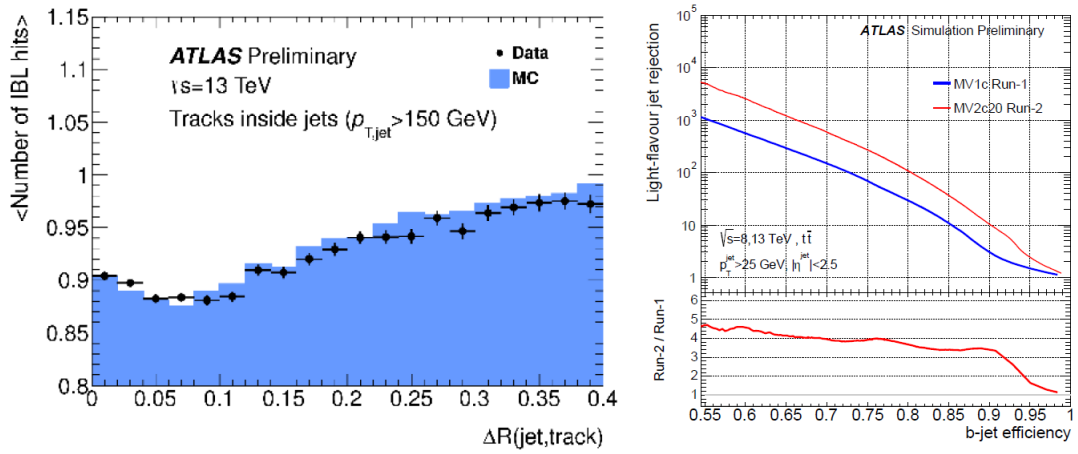


Fig. 2.27 Left: The number of IBL hits on track as a function of $\Delta R(jet, track)$ in data and simulation, for jets with $p_T > 150$ GeV. Right: The comparison of the expected light-flavour jets rejection as a function of b-jet tagging efficiency for both MV1c and MV2c20, the b-jet tagging algorithms used in Run 1 and Run 2 respectively. The ratio of MV2c20 and MV1c rejections is shown at the bottom.

Various properties of the pixel and IBL hits on tracks inside the cones of reconstructed jets are compared between data and simulation. The left plot in Figure 2.27 shows the number of IBL hits of tracks in jets with $p_T > 150$ GeV as a function of $\Delta R(jet, track)$, the separation between the track and the jet axis. Good agreement between data and Monte Carlo is obtained.

The missing transverse energy uses selected calibrated hard objects to measure the missing transverse momentum in an event. Its soft term is reconstructed from detector signal objects not associated with any selected hard object. Using a selected sample of Z decays into two muons, the distribution of missing transverse energy is compared with Monte Carlo simulation finding an agreement within 20% [19].

2.11.3 b-tagging and tau performance

The identification of jets originated from b-quarks (b-jets), commonly known as b-tagging, is very important for the ATLAS physics program. Different algorithms making use of multivariate discriminants were developed to identify b-jets in Run 1. The performance of the new Run 2 multi-variate discriminant MV2c20, namely measuring the rejection of light-flavor jets as a function of the b-jet tagging efficiency, has been studied and compared to the one used in Run 1, MV1. Monte Carlo simulation studies show that the rejection of light-flavor

jets is expected to improve by a factor of four for a wide range of b-jet tagging efficiencies (shown on the right of Figure 2.27), as a result of the installation of the highly granular IBL and new techniques adopted by the collaboration.

The identification of hadronically decaying tau leptons has also been studied using the initial Run 2 data. Reasonably good agreement with simulation has been achieved.

2.12 Conclusion

The CERN Large Hadron Collider (LHC) was built to collide 7 TeV protons or heavy ions of equivalent rigidity. The first proton run (2010-2013) was carried out at 3.5 TeV–4 TeV/beam at a bunch spacing of 50 ns, instead of the nominal 25 ns. This implied fewer bunches with considerable intensity and hence a high peak luminosity but larger than nominal pileup. Run 1 resulted in about 25 fb^{-1} of proton data and important physics results, most notably the discovery of the Higgs boson.

Run 1 was followed by a long shutdown. Run 2 started in 2015 and is planned to continue until the end of 2018, with a goal to produce more than 150 fb^{-1} of data at higher energy and using the nominal 25 ns bunch spacing.

The LHC was and is performing exceptionally well offering a broad range of experimental results from precise measurements to extending the limits and constraints of various models. Run 2 and beyond opened a new era, expecting an exciting future in both theory and experimental physics.

Having set the "experimental playground", I will enter further into the performance part of my thesis namely **Electron identification using W Tag and Probe Method**, followed by **the ttH analysis research** within SM and a study based on a BSM extension of the SM.

Chapter 3

Electron identification using W Tag and Probe Method

3.1 Introduction to electron identification

Electrons and positrons, collectively referred to as electrons, are identified in ATLAS detector as tracks in the inner detector and energy deposits in the electromagnetic calorimeter. First level trigger system (L1) calorimeter signals are combined with tracks using reconstruction algorithms in the high-level trigger (HLT) decision algorithms.

The electron candidates are discriminated against background using several sets of identification criteria with different levels of background rejection and signal efficiency. These identification criteria rely on the shapes of electromagnetic showers in the calorimeter as well as on tracking and track-to-cluster matching quantities. Additionally, isolation criteria may be required on electrons to distinguish them from background objects further.

The accuracy of the Monte Carlo (MC) detector simulation to model the electron measurement efficiency plays an important role. However, the MC simulations are not 100% accurate. For reliable results, the MC samples are corrected to reproduce the efficiencies measured with data.

The efficiency measurements are based on the tag-and-probe method requiring the presence of an isolated identified electron as the probe which is unbiased and pure.

The efficiencies are calculated for data and Monte-Carlo as a function of the pseudorapidity η and the transverse energy E_T. In order to use the results in analyses, Scale Factors (SF) corrects Monte-Carlo to match the results from data. They are defined as:

$$SF(\eta, E_T) = \frac{\varepsilon_{data}(\eta, E_T)}{\varepsilon_{MC}(\eta, E_T)} \tag{3.1}$$

Measuring as precisely as possible the identification efficiency (along with all efficiencies like reconstruction, isolation, etc.) is a mandatory ingredient in almost all physics analysis (in particular it is an essential tool for ttH, the analysis presented in Chapter 4).

The samples of electron candidates used to calculate the electron identification efficiency in a range of transverse momentum $p_T > 7$ GeV, in the central region of the detector $|\eta| < 2.47$, are selected using the tag-and-probe method, both on data and on Monte Carlo simulations in the same phase space.

To cover all the energy range, three complementary decays are used:

- low p_T range: $J/\Psi \to ee$
- $Z \to ee$
- $W \to e\nu$

Numerous advantages come from using $W \to e\nu$ events to measure the electron identification efficiency. I can mention here that the measurements with W and Z bosons cover both a large kinematic range (as can be seen in Figure 3.1 that shows the transverse energy distributions of probe electron candidates satisfying the Tight identification criterion in the $W \to e\nu$, $Z \to ee$ and the $J/\Psi \to ee$ samples). The $W \to e\nu$ events could help in understanding the uncertainty by offering different systematics from Z. Also, W provides additional statistics for the low E_T range (15-25 GeV) that can complement the J/Ψ channel and offers interesting experimental signatures due to its large statistics.

Each analysis has its own advantage and limits. The official $Z \to ee$ and the $J/\Psi \to ee$ events are easier to reconstruct since in the final state there are two leptons, compared to the $W \to e\nu$ events where the modeling of the missing transverse energy could be problematic. However, the combination of the three channels will be valuable. First, this will allow increase of statistics over the entire kinematic range. Secondly, it will improve the systematic uncertainty that arises in each method (signal optimization, background templates,

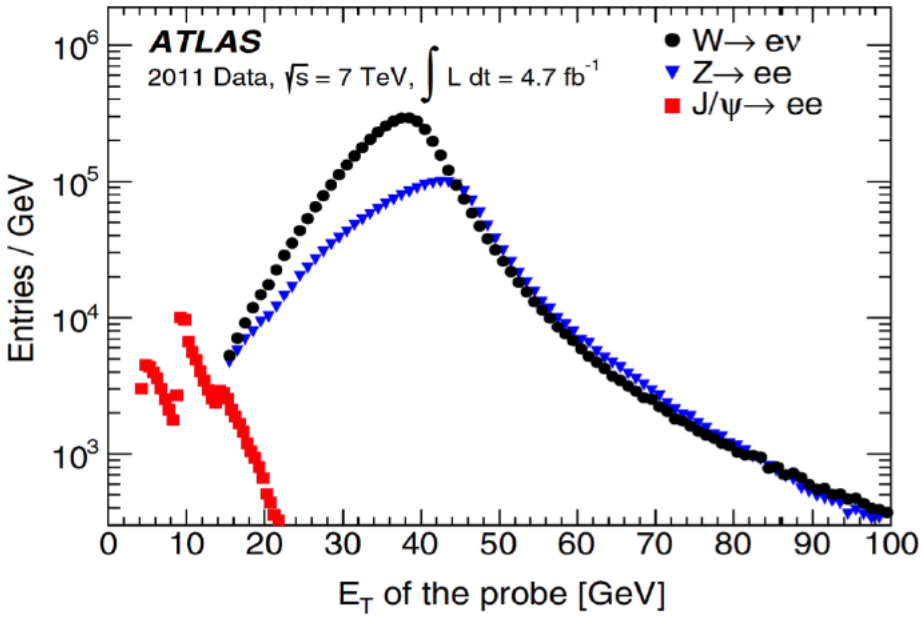


Fig. 3.1 The transverse energy distributions of probe electron candidates satisfying the Tight identification criterion in the $W \to e\nu$ (full circles), $Z \to ee$ (blue triangles) and the $J/\Psi \to ee$ (red squares) samples using 2011 data set.

fit procedure). The combination of the three channels will lead to a more accurate description of the electron identification namely on the efficiency and Scale Factors that are applied on Monte-Carlo samples to simulate the data from the ATLAS detector better.

In the following sections, I will present the discriminating variables used by ATLAS algorithms to disentangle between Signal and Background. I will present after the Tag and Probe method to extract the electron identification efficiency, and how it is applied on $W \to e\nu$ events.

One of the limits of the 2015-2016 analysis on $W \to e\nu$ events was the large background observed in the "signal pure" region. Therefore I will focus on a few sections on how the background estimation and rejection was handled using two complementary methods: cut-based and multivariate technique.

In the end, I will present the results obtained in the $W \to e\nu$ events on efficiency and SF and also give a comparison with the official ATLAS $Z \to ee$ values.

3.1.1 Discriminating Variables for Electron Identification

In order to have a sample of pure reconstructed electrons, additional selection criteria using both the ID and the calorimeter information are necessary. The aim is to provide a highly efficient electron selection, with large background rejection using a set of variables are generically referred to as "shower-shapes" [12] [13].

Figure 3.2 shows a sketch of electromagnetic and hadronic calorimeter along with the electron discriminating variables that will be described in this section.

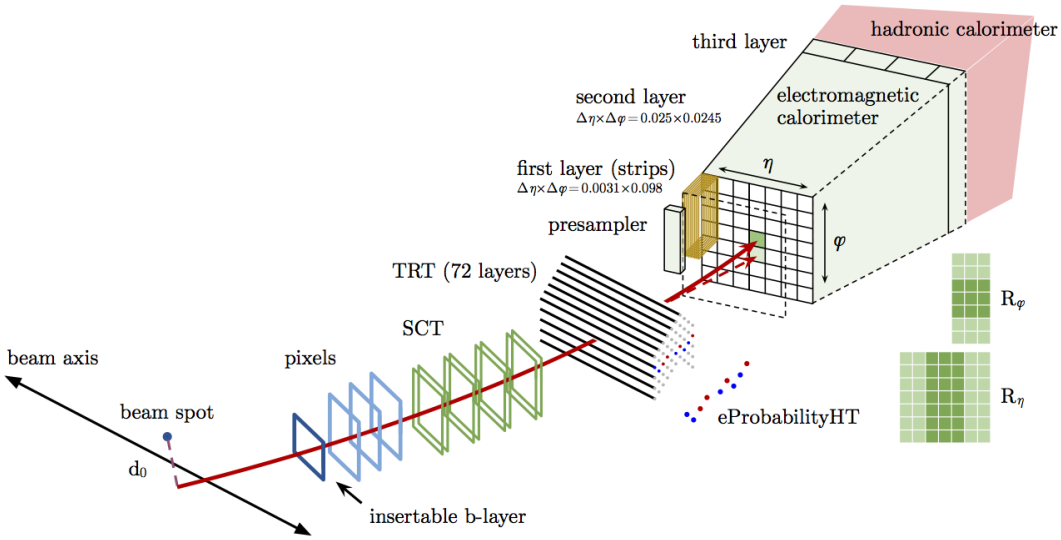


Fig. 3.2 A sketch of electromagnetic and hadronic calorimeter and the ID.

Figure 3.3 shows the shower-shapes variables for true electrons labeled as "Isolated electrons", hadrons, conversions labeled "Background electrons", and semi-leptonic heavy-flavor decays labeled "Non-isolated" electrons.

In the Figure 3.3 a) we see the distribution of the hadronic leakage R_{had1}. It is defined as the ratio of the energy in the first sampling of the hadronic calorimeter, behind the electron cluster, to the energy of the electron cluster. Real electrons deposit most of their energy in the EM calorimeter. Thus they tend to have small values of R_{had1}, larger values indicating hadronic activity associated with the electron cluster.

In the Figure 3.3 b) we see the distribution of the width od the shower in the second sampling, $w_{\eta 2}$. It measures the width of the shower in η as the energy-weighted RMS of the

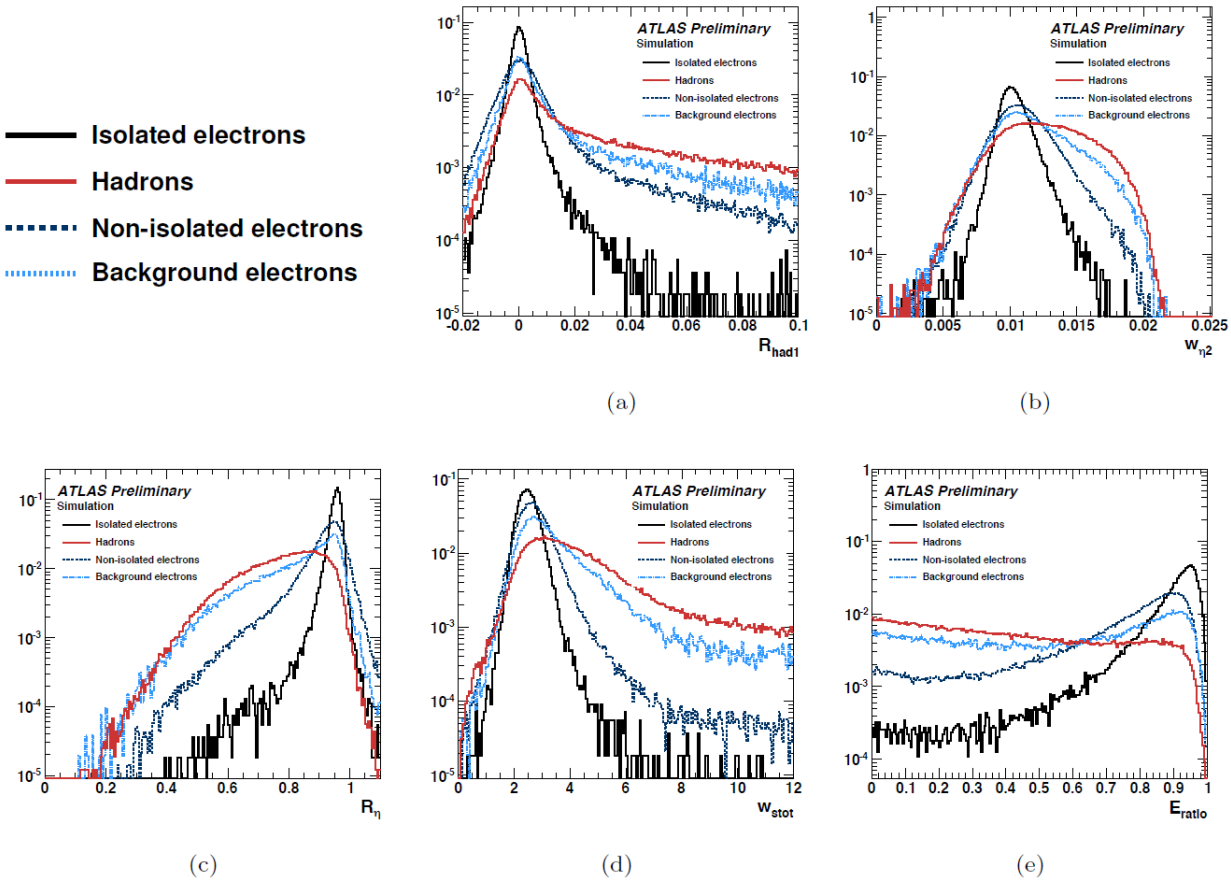


Fig. 3.3 "Shower-shapes" variables, shown separately for signal and the various background types: true electrons labeled as "Isolated electrons", hadrons, conversions labeled "Background electrons", and semi-leptonic heavy-flavor decays labeled "Non-isolated" electron. The variables shown are a) hadronic leakage R_{had1}, b) width in eta in the second sampling $w_{\eta 2}$, c) R_η, d) width in eta in the strips $w_{s,tot}$, and e) E_{ratio}.

η distribution of cells in the second sampling:

$$w_{\eta 2} = \sqrt{\frac{\Sigma_i(E_i\eta_i^2)}{\Sigma_i(E_i)} - (\frac{\Sigma_i(E_i\eta_i)}{\Sigma_i(E_i)})^2} \tag{3.2}$$

where $E_i(\eta_i)$ is the energy/pseudorapidity of the i^{th} cell, and the sum runs over the cells in a 3×5 window of the second sampling, centered on the electron. To suppress background from jets and photon conversions, which tend to have wider showers than true electrons, narrow shower widths are required.

In the Figure 3.3 c) we see the distribution of R_η, defined as the ratio of cell energies in a 3×7 ($\eta \times \phi$) window to that of a 7×7 window, in the second sampling. The backgrounds candidates have lower values of R_η, making it the most powerful variables for background separation.

In the Figure 3.3 d) we see the distribution of $w_{s,tot}$, the width of the shower in the first layer of the calorimeter, or strips, defined as:

$$w_{s,tot} = \sqrt{\frac{\Sigma_i E_i(i - i^{max})^2}{\Sigma_i(E_i)}} \tag{3.3}$$

where, E_i is the energy in the i^{th} strip, i is the strip index, and i^{max} is the index of the strip with the most energy. The discrimination power comes from the fact that the shower width in the strips is larger for background than for signal.

Finally, in the Figure 3.3 e) we see the distribution of E_{ratio} defined using the cells corresponding to the two highest energy maxima in the strips:

$$E_{ratio} = \frac{E^s_{1st-max} - E^s_{2nd-max}}{E^s_{1st-max} + E^s_{2nd-max}} \tag{3.4}$$

Jet background tends to have multiple incident particles associated to the reconstructed cluster, leading to lower values of E_{ratio} than for true electrons.

Similar to R_{had1}, the energy fraction of the EM sampling to the cluster energy in the third sampling, f_3, is another calorimeter variable used to discriminate between electrons and background: it tends to be smaller for electrons than for background, which penetrates

deeper into the calorimeter.

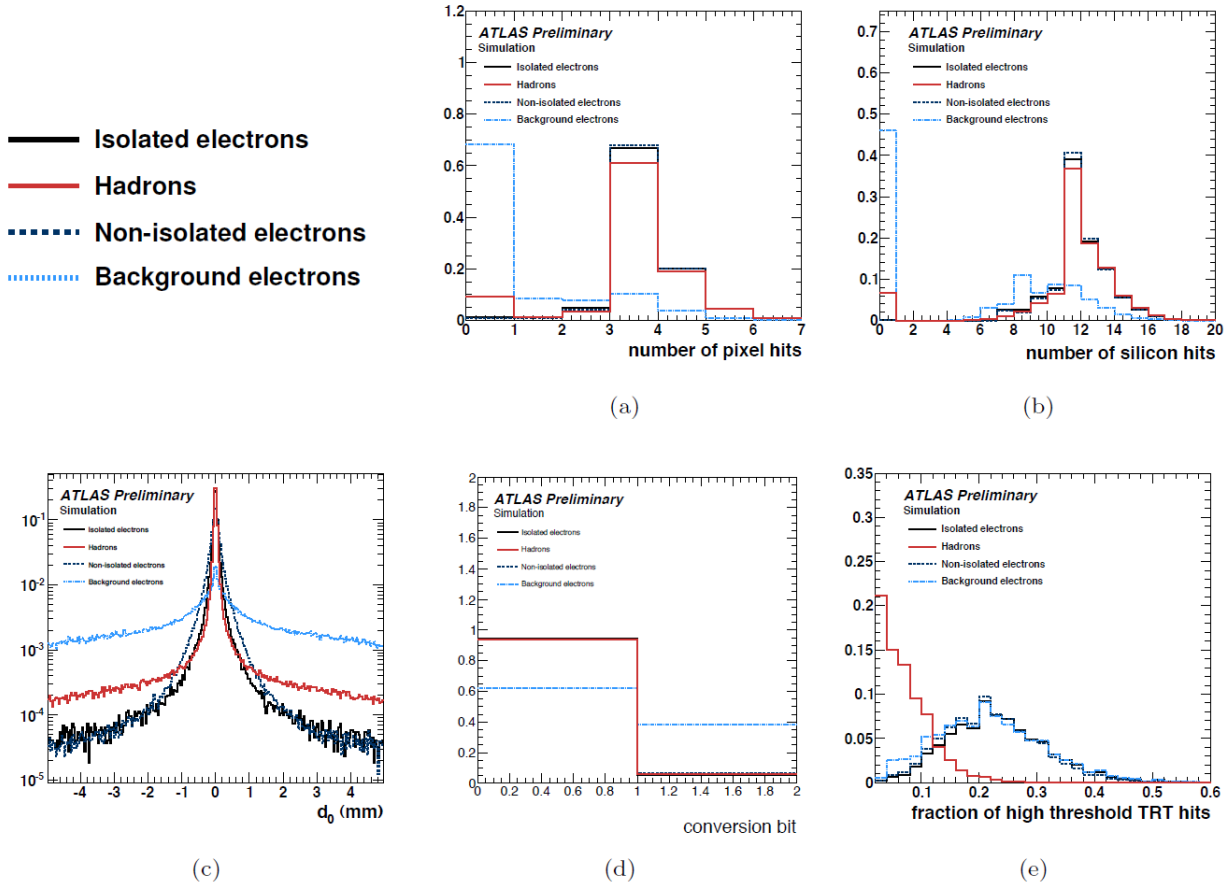


Fig. 3.4 Electron identification variables for true electrons labeled as "Isolated electrons", hadrons, conversions labeled "Background electrons", and semi-leptonic heavy-flavor decays labeled "Non-isolated" electron: a) number of hits in the Pixel detector, b) combined number of hits in the Pixel and SCT detectors, c) transverse impact parameter d_0, d) conversion flag, or "conversion bit", and e) fraction of high-threshold hits in the TRT.

The ID variables also provide discriminating power used in electron identification. They are shown in Figure 3.4 for a data sample in Run II . The ID variables are often uncorrelated from the calorimeter measurements, allowing the signal purity of the calorimeter (tracking) variables to be enhanced, in an unbiased way.

Figure 3.4 a) and Figure 3.4 b) shows the number of hits in the Pixel and SCT detectors associated to the electron track. By requiring electron tracks to have pixel hits and a significant number of SCT hits (track-quality requirement), the background from conversions can

be suppressed with little loss in signal efficiency.

The transverse impact parameter distribution, d_0 (Figure 3.4 c), measures the distance of closest approach of the electron track to the primary vertex in the transverse plane. The discrimination power of this variable is primarily against conversions, which have tracks significantly displaced from the interaction point, and against heavy-flavor decays because of the large b-hadron lifetime.

The conversion bit (Figure 3.4 d) is set if the electron track is matched to a conversion vertex: single-leg or double-leg. Electrons are flagged as double-leg conversions if there is another ID track that forms a secondary vertex with the electron track consistent with coming from a photon conversion. An electron is flagged as a single-leg conversion if it is missing a hit in the b-layer. If conversion bit is required to be 0, then it removes a significant fraction of reconstructed electrons from conversions.

Figure 3.4 e shows the fraction of high threshold hits in the TRT. In the case of transition radiation (TR) photons, we will expect a high threshold TRT hits. Also, the high threshold fraction is one of the most powerful discriminating variables against the background from hadrons.

Track-cluster matching variables provides additional background discrimination (Figure 3.5).

Figure 3.5 a) shows the difference in η of the track and the cluster, narrowest for real electrons. A similar variable, the track-cluster matching in $\Delta\phi$, is shown in Figure 3.5 b). It is assigned to be positive or negative based on the electron charge. Matching in $\Delta\phi$, particularly on the positive side of the distribution, can be used to suppress background, analogously to $\Delta\eta$.

Another variable related to track-cluster matching is E/p, shown in Figure 3.5 c), defined as the ratio of the electron energy measured in the calorimeter to the track momentum determined from the ID. The signal electron distribution peaks at one and has a long tail in E/p while for hadrons it peaks at lower values of E/p and for conversions it tends to have larger values.

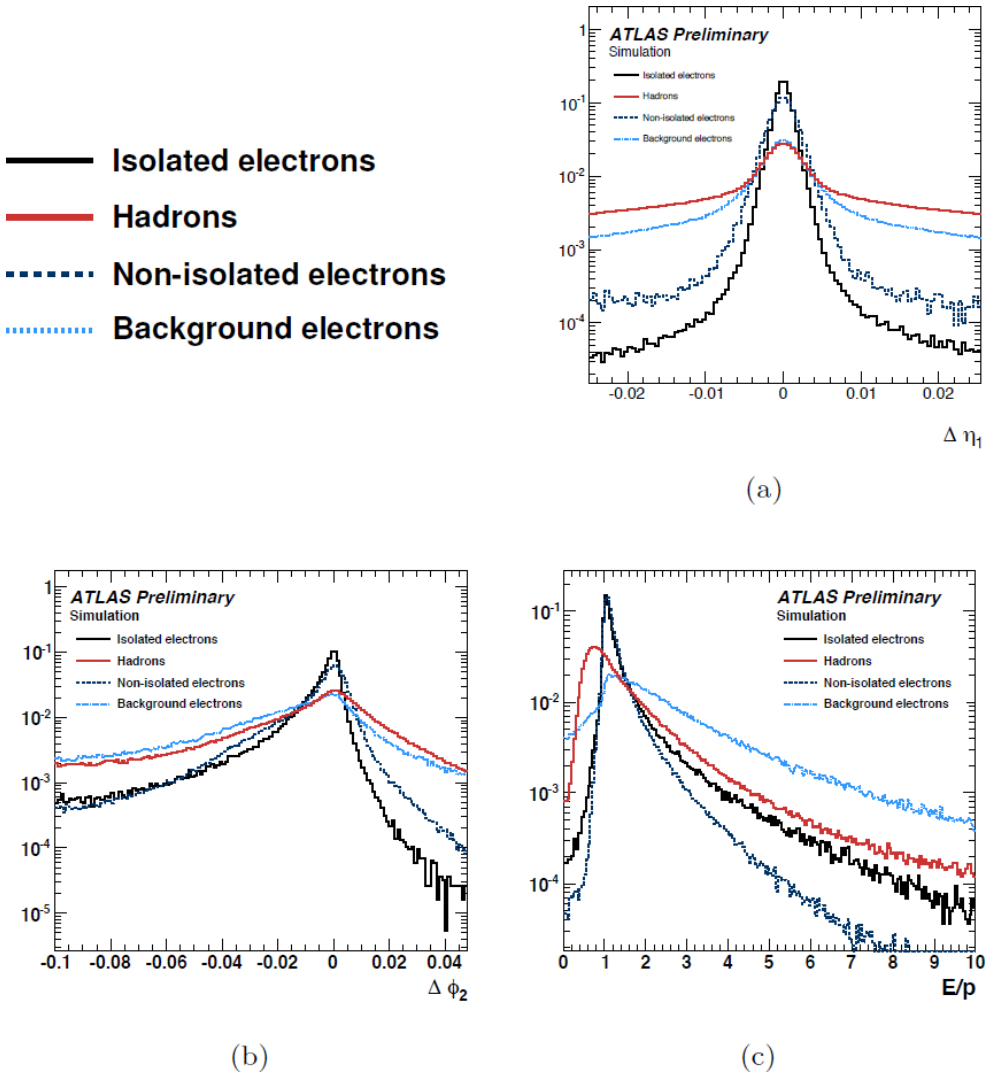


Fig. 3.5 Track-Cluster matching variables for true electrons labeled as "Isolated electrons", hadrons, conversions labeled "Background electrons", and semi-leptonic heavy-flavor decays labeled "Non-isolated" electron: a) the difference in track and cluster η, b) the difference in track and cluster ϕ, and c) ratio of the energy measured in calorimeter to the momentum measured in the tracker.

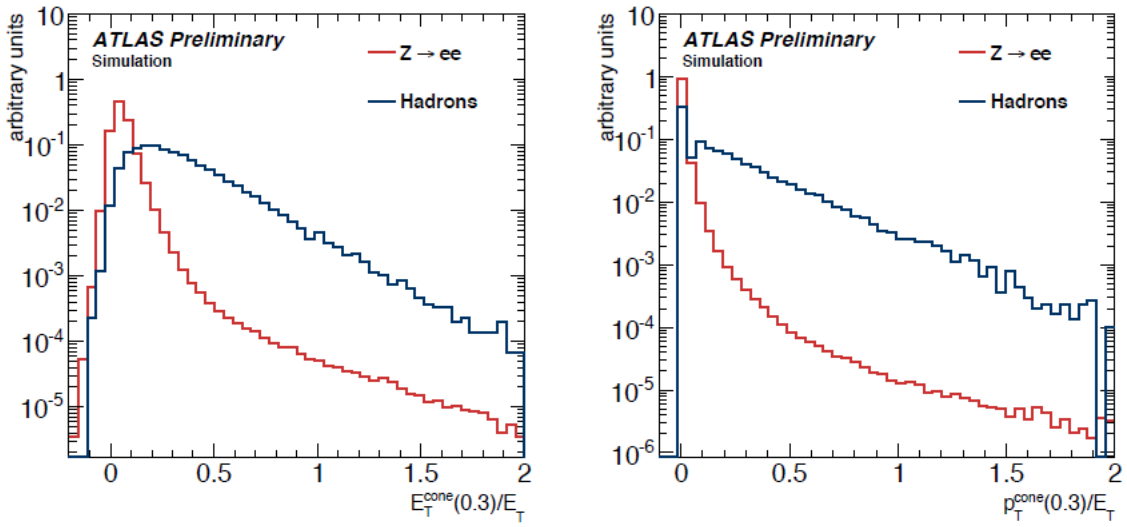


Fig. 3.6 Electron isolation variables for true electrons labeled as "Isolated electrons", hadrons, conversions labeled "Background electrons", and semi-leptonic heavy-flavor decays labeled "Non-isolated" electron: a) relative calorimeter isolation in a cone of $\Delta R < 0.3$; b) relative track isolation in a cone of $\Delta R < 0.3$.

The final class of discriminating variables is the isolation, as shown in Figure 3.6 at $\sqrt{s}$=13 TeV. Isolation measures the amount of energy near the reconstructed electron. Since background electrons tend to be produced in association with other particles, they will have larger values of isolation compared to signal electrons. The isolation is calculated by summing the energy in a cone centered around the electron, ΔR, of typical size of 0.2, 0.3 or 0.4. Figure 3.6 a) shows the calorimeter isolation formed from the energy deposits in the cone of ΔR = 0.3 around the electron candidate divided by the E_T of the electron candidate. Figure 3.6 b) shows the relative track isolation, using a cone size of 0.3. The distributions are shown for signal electrons and the hadron background.

3.1.2 The electron efficiency measurement

The production cross-section times branching-ratio of a certain physics process, (for example W decay), is measured using the formula:

$$\sigma_W \times BR(W \to e\nu) = \frac{N^{data} - N^{bkgrd}}{A_W \cdot C_W \cdot L_{int}} \tag{3.5}$$

with:

- N^{data} the number of events observed in data;
- N^{bkgrd} the number of expected background events;
- A_W is the correction factor to account for the geometrical and kinematic acceptance of the detector;
- C_W is the correction factor to account for the efficiency of the event being reconstructed in the detector;
- L_{int} the integrated luminosity.

The acceptance A_W is the fraction of Monte-Carlo events that are generated in the geometrical and kinematic range of the detector, the so-called fiducial phase-space. This number is obtained using generator-level information only. The efficiency C_W is the fraction of these events that get reconstructed correctly and it can be further decomposed into:

$$C_W = \varepsilon_{event} \times \varepsilon_{reco} \times \varepsilon_{ID} \times \varepsilon_{trigg} \times \varepsilon_{isol} \tag{3.6}$$

where ε_{event} is the efficiency of the signal events passing the event preselection cuts, ε_{reco} accounts for the differences between applying the geometrical and kinematic selection at generator or reconstruction level (the reconstruction efficiency), ε_{ID} is the efficiency of an electron passing the identification criteria, ε_{trigg} is the efficiency of an event being triggered and ε_{isol} is the efficiency of possible isolation cuts.

3.1.3 The identification algorithms

ATLAS has developed a set of identification requirements (referred to as the "cut value") used to select electrons. The identification algorithms are based on a set of cuts on shower shape variables, track variables and the combination of both, as described in 3.1.1. They intend to reject a large number of background candidates with little loss of real signal electrons. There are three sets of criteria to accommodate a broad range of physics topics. They are referred to, in order of increasing background rejection as Loose, Medium, and Tight, and are applied on top of the electron reconstruction and successively built upon each other. The variables used for the three identification criteria are listed in Table 3.1 for Loose selection, 3.2 for Medium selection and 3.3 for Tight selection.

The cuts on the shower shape variables are optimized in pseudo-rapidity η and transverse energy E_T by a multivariate analysis program, using 10 bins in E_T and 20 bins in η being considered:

- η binning: "-2.47" , "-2.37", "-2.01" , "-1.81" , "-1.52" , "-1.37" , "-1.15" , "-0.8" , "-0.6" , "-0.1" , "0" , "0.1" , "0.6" , "0.8" , "1.15" , "1.37" , "1.52" , "1.81" , "2.01" , "2.37" , "2.47";
- E_T binning: "15", "20", "25", "30", "35", "40", "45", "50", "60", "80", "150" ;

All these 200 bins will be used to calculate the electron identification efficiencies (section 3.8) and Scale Factors (SF) (section 3.9).

Type in Loose selection	**Description**	**Name**
Acceptance	$\|\eta\| < 2.47$	
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| < 0.8$ and $\|\eta\| > 1.37$) Ratio of E_T of the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| > 0.8$ and $\|\eta\| < 1.37$)	R_{had1} R_{had}
Middle layer of EM calorimeter	Ratio of the energy in 3×7 cells over the energy in 7×7 cells centred at the electron cluster position Lateral width of the shower	R_η $W_{\eta 2}$
Track quality Track-cluster matching	Number of hits in the pixel detector (≥ 1) Number of total hits in the pixel and SCT detectors (≥ 7) $\Delta\eta$ between the cluster position in the strip layer and the extrapolated track ($\|\Delta\eta\| < 0.015$)	n_{pixel} n_{Si} $\Delta\eta$
Strip layer of EM calorimeter	Total shower width Ratio of the energy difference between the largest and second largest energy deposits in the cluster over the sum of these energies	W_{stot} E_{ratio}

Table 3.1 List of variables used for the identification criterion Loose

3.1.4 W triggers in 2015-2016

Dedicated $W \to e\nu$ tag-and-probe triggers are used to record the data for the efficiency measurement. In Z T&P analysis, one can use single-electron triggers, as long as it is the

Type in Medium selection	**Description**	**Name**
Shower-shapes	All Loose shower shape cuts but at tighter values	
Track-cluster matching	Tighter $\Delta\eta$ track-cluster matching $\lvert\Delta\eta\rvert < 0.005$	$\Delta\eta$
Track quality	Transverse impact parameter ($\lvert d_0\rvert < 5$ mm)	d_0
	Number of b-layer hits (≥ 1) for $\lvert\eta\rvert < 2.01$	n_{BL}
	Stricter requirements for pixel hits (> 1) for $\lvert\eta\rvert > 2.01$	n_{pixel}
TRT	Ratio of the number of high-threshold hits to the total number of hits in the TRT	f_{HT}

Table 3.2 List of variables used for the identification criterion Medium

Type in Tight selection	**Description**	**Name**
Track-cluster matching	Tighter $\Delta\Phi$ between the cluster position in the middle layer and extrapolated trackg $\lvert\Delta\Phi\rvert < 0.02$	$\Delta\Phi$
	Ratio of the cluster energy to the track momentum	E/p
	Tighter $\Delta\eta$ requirement ($\lvert\Delta\eta\rvert < 0.005$)	
Track quality	Tighter transverse impact parameter requirement ($\lvert d_0\rvert < 1$ mm)	d_0
	Number of b-layer hits (≥ 1) for full $\lvert\eta\rvert$ range	n_{BL}
TRT	Total number of hits in the TRT	n_{TRT}
Conversions	Number of hits in the b-layer (≥ 1)	n_{BL}
	Veto electron candidates matched to reconstructed photon conversions	

Table 3.3 List of variables used for the identification criterion in addition to Medium criteria

tag electron that set the trigger. In the case of W T&P, the missing transverse energy (or its significance) is used to select events of interest.

The ATLAS online data processing reconstructs and identifies electron candidates both at the L1 trigger and at the HLT. At L1, the electron triggers use the signals recorded in the electromagnetic (EM) and hadronic calorimeters. The ET thresholds used in the trigger decision can be set differently for different η regions, allowing a coarse tuning of thresholds taking into account different energy responses in different regions.

At the HLT, electron candidates are reconstructed and selected in several steps, requiring passing different background rejection cuts based on the topology of the events (for example

in W T&P we apply at HLT level cuts on xe, xs, m_T, see later of definition) to enhance the signal, thereby reducing the event rate to a level where more precise offline-like algorithms can be applied in the allowed latency range.

For each Monte-Carlo run a weight of the delivered luminosity, divided by the total luminosity in the corresponding data period, is assigned to the different triggers. The weight assigned to a given event is the sum of trigger weights of the triggers the event passed.

In order to reduce the event rate to prevent automatic pre-scaling, for W T&P, the missing transverse energy (or its significance) is combined with a low E_T EM cluster trigger element with $E_T > 10$ GeV and $E_T > 13$ GeV.

We can also further reduce the event rate by applying additional cuts such as on the azimuthal separation of the missing transverse energy from the two/six leading jets (2dphi or 6dphi) at >0.5 or >1.5.

In the early Run 2 data taking, the triggers used for $W \to e\nu$ are divided, based on EM cluster selection and W topology, into:

- e13/e18 triggers: the triggers with E_T of electron >13/18 GeV at HLT level:
  - "HLT_e13_etcut_trkcut_L1EM12",
  - "HLT_e13_etcut_trkcut_xs15_mt25",
  - "HLT_e13_etcut_trkcut_xs15_j20_perf_xe15_2dphi05_mt25",
  - "HLT_e13_etcut_trkcut_j20_perf_xe15_2dphi05_mt25",
  - "HLT_e13_etcut_trkcut_xs15_j20_perf_xe15_6dphi05_mt25",
  - "HLT_e13_etcut_trkcut_j20_perf_xe15_6dphi05_mt25",
  - "HLT_e13_etcut_trkcut_xs30_xe30_mt35",
  - "HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_6dphi05_mt35",
  - "HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_6dphi15_mt35",
  - "HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_2dphi05_mt35",
  - "HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_2dphi15_mt35",
  - "HLT_e18_etcut_trkcut_L1EM15",
  - "HLT_e18_etcut_trkcut_xs20_mt35",

  - "HLT_e18_etcut_trkcut_xs20_j20_perf_xe20_6dphi15_mt35",
  - "HLT_e18_etcut_trkcut_j20_perf_xe20_6dphi15_mt35",
  - "HLT_e18_etcut_trkcut_xs30_xe30_mt35",
  - "HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_6dphi05_mt35",
  - "HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_6dphi15_mt35",
  - "HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_2dphi05_mt35",
  - "HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_2dphi15_mt35"

- High-MET triggers: the triggers that require a cut on missing tranverse energy (xe at HLT level or XE at L1 level):

  - "HLT_j80_xe80",
  - "HLT_xe80_tc_lcw_L1XE50",
  - "HLT_xe90_mht_L1XE50",
  - "HLT_xe90_tc_lcw_wEFMu_L1XE50",
  - "HLT_xe90_mht_wEFMu_L1XE50"

To explain the configuration of a trigger let's take as an example a HLT_e18 trigger, HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_2dphi05_mt35. For this specific trigger at HLT is applied a threshold for the electron p_T to be > 18 GeV. The etcut_trkcut refers to passing track requirements, xs30 represents a cut on missing transverse energy significance (xs) to be > 3, j15 requires to have at least one jet with p_T >15 GeV, xe30 requires a missing transverse energy (MET or xe) to be >30 GeV, 2dphi05 takes the 2 highest p_T jets and the angle between the jets and probe electron to be >0.5, mt35 requires the transverse mass to be >35 GeV.

Another example from High-MET triggers, if we take HLT_j80_xe80, the j80 requires at least one jet with p_T>80 GeV and missing transverse energy >80 GeV.

The W T&P dedicated triggers assure by their construction a selection of probe electron and tag MET with the highest signal to background ratio as possible, depending on the requirements of the trigger. For the low energy threshold triggers, useful for J/Ψ T&P combination, for example, additional requirements on the tag are applied to assure a reduction of event rate by keeping the signal purity as high as possible. The High-MET triggers provide a low statistical sample of pure high energy $W \rightarrow e\nu$ candidates, useful for E_T>60 GeV electron efficiency measurement, a range where W T&P analysis could be combined with Z

T&P efficiency measurements.

3.2 Tag and probe method

The method uses electrons produced in the decays of known resonances that are selected based on a discriminant variable (i.e., the electron-positron invariant mass). In the $W \to e\nu$, the missing transverse momentum is used for tagging the event. The probe electron must have a good track quality all along the particle path (at least one-pixel hit and 7 Si hits).

The basic idea is to select an event by setting strict requirements on one of the two final state particles, the tag object. The other particle serves as a probe for the measurement. The cut of interest as identification criterion can be tested on this unbiased probe.

In the simple case of a counting experiment the efficiency ε of a cut is given by:

$$\varepsilon = \frac{N_P}{N_0} \tag{3.7}$$

where N_0 is the number of total events and N_P is the number of events passing a certain cut.

The electron efficiency measurement is based on the Tag and Probe method which uses the isolation distribution as main ingredient to subtract the background. It exploits the isolation properties of the W boson decay products specifically $W \to e\nu$. The probe isolation E^T_{cone} is used as a discriminating variable, which is a measure of the energy deposit of topological clusters within a specific cone around the electron. This variable disentangles electron from heavy resonance decays (such as $W \to e\nu, Z \to e^+e^-$) from soft electrons originating from photon conversions or heavy flavor hadron decays from light hadrons miss-identified as electrons.

E_{cone} is defined as the sum of transverse momenta of positive energy topological clusters, calibrated at the electromagnetic scale, within a cone of $\Delta R = 0.3$ of the candidate electron cluster. The transverse energy contained in a rectangular cluster of size $\Delta\eta \times \Delta\Phi = 0.125 \times 0.175$ centered around the electron cluster barycenter is subtracted.

Isolated electrons from the W boson decay should be well separable from fake electrons in the isolation distribution, making possible to estimate the fake electron background.

The Cut Based (named after the way they are defined) background templates are constructed by inverting cuts on identification and shower shape variables and scaled to the background dominated region. There are two background templates constructed in this way:

- Template1 (T1): define as probe electron failing the cuts on the variables $w_{s,tot}$ and $ePROBHT$ (the probability of an electron to have a high-threshold hit in the TRT.)
- Template6 (T6): constructed by requiring at least two selection cuts not to be fulfilled (excepting R_{had})

Another background template is constructed by requiring the probe electron to fail the Very Loose Likelihood requirement (labeled as !VL LH Background Template), the Run II version of the Run I Loose definition.

To clean the background, all events are required to pass basic preselection and quality cuts:

- Pass W triggers defined in Section 3.1.4
- Missing Transverse Energy (MET) for the tag (the neutrino) $> 25GeV$

while for the probe electron, stronger cuts are applied to further reject the background and have a cleaner probe sample:

- $p_T > 15GeV$
- $|\eta| < 2.47$
- Transverse mass $(m_T) > 40GeV$
- $R_{had} < 0.1$

Also, a jet veto is applied on the probe electron, rejecting all probes inside a radius of $\Delta R < 0.2$ from an AntiKt4 jet with energy E_T larger 15 GeV.

The missing transverse energy (labeled as MET or E_T^{miss}) is the key signature of neutrinos that escape the detector without interacting. It is defined as the negative vectorial sum of all energy deposits in the detector:

$$E_x^{miss} = -\Sigma_i E_i sin\theta_i cos\phi_i \tag{3.8}$$

$$E_y^{miss} = -\Sigma_i E_i sin\theta_i sin\phi_i \tag{3.9}$$

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \tag{3.10}$$

E_T^{miss} is calculated from three-dimensional topological clusters with energies E_i (with $E_i >$ 10 GeV).

The E_T^{miss} significance (labeled as XS) is defined as the ratio of the missing transverse energy to its resolution:

$$XS = \frac{E_T^{miss}}{a \cdot (\sqrt{\Sigma E_T} - b)} \tag{3.11}$$

Depending on the period and trigger level, values of 0.46-1.12 are used for the parameter a and values of 0.5-1.28 for the parameter b [86]. The trigger ΣE_T has some positive bias which is taken care by the b parameter, whereas a takes care of the energy scale

The transverse mass (labeled as m_T) is calculated from E_T^{miss} and the selected electron E_T:

$$m_T = \sqrt{2E_T E_T^{miss}(1 - cos\Delta\phi)} \tag{3.12}$$

with the azimuthal separation between the electron and the missing transverse energy $\Delta\phi$.

3.3 Event selection in 2015-2016 data

The $W \to e\nu$ events contain an isolated electron as well as a neutrino (referred to as the missing transverse energy). In 2015-2016 data of total integrated luminosity of $L = 25.21 fb^{-1}$, each event is required to have at least one reconstructed vertex with more than 2 associated tracks with $p_T > 500$ MeV .

In 2015, the events were selected by prescaled triggers which require to pass several cuts (see Table 3.4) like MET (xe) , MET significance (xs), m_T etc.

In 2015, the triggers were heavily prescaled and noisy (background dominated), so the background rejection at HLT level with the cuts on the variables from the table 3.4 was not enough to provide a statistically significant sample with high signal purity feasible to be used for efficiency studies.

Variables used in Egamma Triggers	Definition	Cut(s) applied
Transverse mass (m_T)	The transverse mass is calculated from the selected electron candidate and the missing transverse energy as: $m_T = \sqrt{2E_T E_T^{miss}(1-\cos\theta)}$	$m_T >$25, 35 *GeV*
E_T threshold	at HLT level, the electron is required to have an energy above certain threshold in order to reject the background candidates	$E_T >$ 13, 18 *GeV*
MET significance (xs)	xs is defined as the E_T^{miss} calculated as the vectorial sum of energy deposits in the calorimeters, divided by the scalar sum of all energy deposits in the calorimeter $xs = \frac{E_T^{miss}}{0.5 \times \sqrt{\sum E_T}}$	$xs >$ 1.5, 2
MET energy (xe)	Threshold on the missing transverse energy	$xe >$ 15, 20 *GeV*
$\Delta\Phi$ (e-Jet)	A jet reconstructed close to the electron and thus possibly be caused by energy deposits of the electron is targeted, the 6 leading (6dphi) or 2 (2dphi) leading p_T jets being considered	$\Delta\Phi >$ 0.5, 1.5
j20	Only jet with $p_T > 20$ *GeV* are considered	p_T of the jet $>$ 20 *GeV*

Table 3.4 Table with variables used by Egamma Triggers in 2015 data

Following this reason and the analysis performed on background rejection (see section 3.4), we proposed a new set of triggers to be used in 2016, with harsher cuts on electron identification variables (see section 3.6).

3.4 Background rejection in 2015-2016 using Cut Based Method

In this section, I will present how we can perform background rejection using Cut Based Method. As the name suggests, the aim is to identify discrimination variables and apply a cut on them to disentangle between Signal and Background.

First, if we look at the E_T distribution for each W trigger, we can see a significant reduction from event level to probe level (after applying the Tag and Probe Method) and

further from Loose to Tight menu (Figure 3.7). The statistics are reduced drastically from Probe (containing Background+Signal events) to Tight (considered Pure Signal events) menu, therefore, in 2015-2016, the Background (BG) Rejection performed at HLT level was not enough to provide a clean signal data sample. So, to enhance the Signal to Background ratio, we proposed a new set of cuts to be applied on top of the HLT level, offline in the analysis stage.

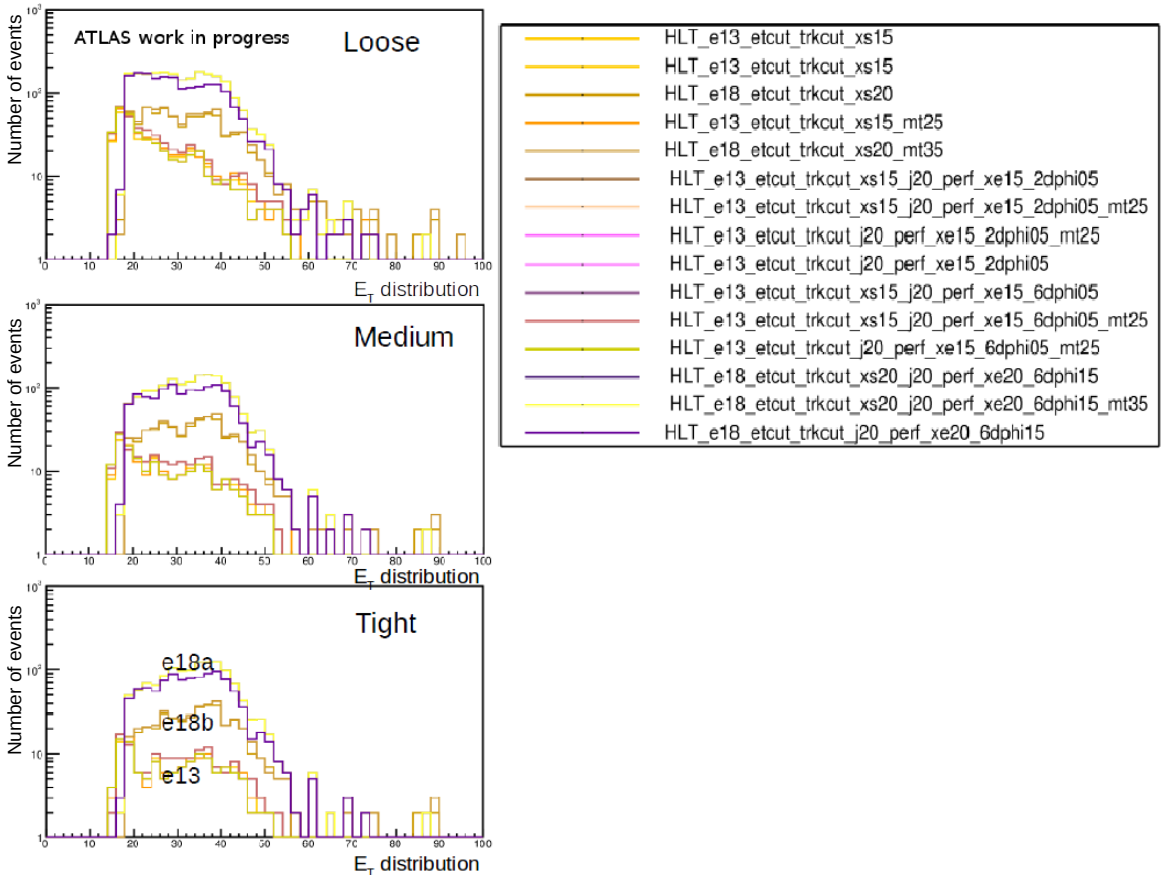


Fig. 3.7 PreScaled E_T distributions in logarithmic scale, for all W triggers considered (see Section 3.1.4) at Loose(L) (top) /Medium(M) (middle) /Tight(T) (bottom) Level.

In order to further reduce the background, we test the angular distribution in the transverse plane (see Figure 3.8) for which only Jets with $p_T > 5$, 10 or 15 GeV were tested and with a map distribution in Φ and η plane between the electron and the first leading p_T Jet of Δ R (e-jet) > 0.2.

$$\Delta R = \sqrt{\Delta\Phi^2(e-Jet) + \Delta\eta^2(e-Jet)} \tag{3.13}$$

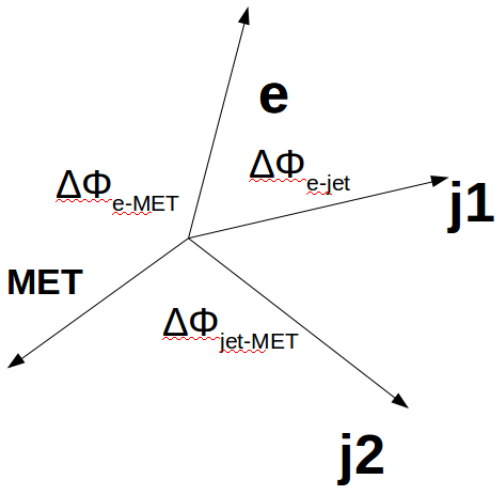


Fig. 3.8 W topology in electron ID angular variables in W boson rest frame.

Using a Monte-Carlo (MC) sample of high purity $W \rightarrow e\nu$ signal, we can compare the MC distribution of several variables to the data distributions, the difference between them being related to the background contamination in data. Looking at Figure 3.9, by comparing data (left plots) to MC (right plots) from Probe (black line) to Tight(Cyan line) we can see that the background is visible around a $\Delta\Phi$ (e-Jet) > 2.5 and $\Delta\Phi$ (MET-Jet) < 0.7. In addition, the correlation for the $HLT\_e18\_etcut\_trkcut\_j20\_perf\_xe20\_6dphi15$ trigger between the angles mentioned (see Figure 3.10) indicates the fact that most of the statistics at trigger level are in the background region, so the proposed cuts on the angular variable to reduce the background will not affect the signal.

In conclusion, angular distributions illustrate background features that can be rejected by applying the $\Delta\Phi$ (e-Jet) > 2.5 and $\Delta\Phi$ (MET-Jet) < 0.7. With these cuts, the Signal to Background ratio being improved by 10-20 %.

Lastly, we have check the distributions in terms of the shower shapes variables (described in section 3.1.1) and apply the following cuts:

- $R_{had} < 0.1$
- $R_{\eta} > 0.65$ and $R_{had} < 0.1$
- $w_{stot} < 4$
- $R_{\eta} > 0.7$

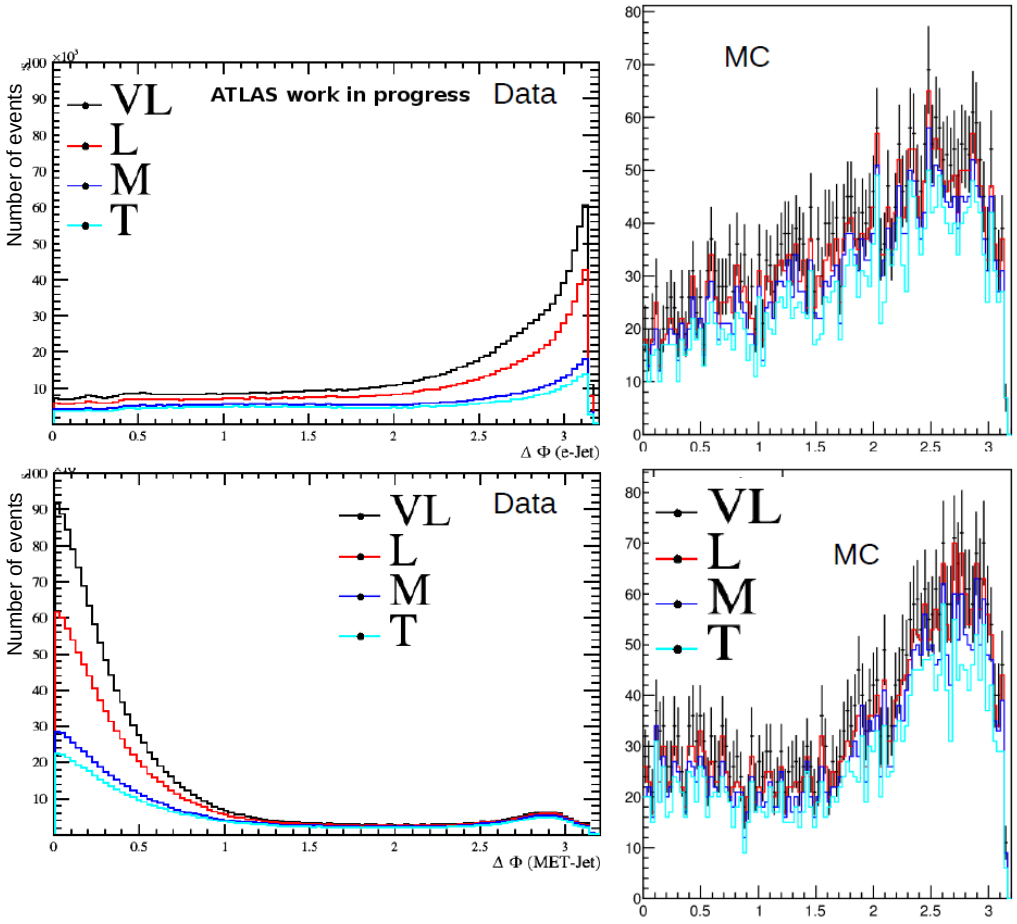


Fig. 3.9 The ΔΦ distributions ($\Delta\Phi_{e-jet}$ on top and $\Delta\Phi_{jet-MET}$ on bottom) for jets with $p_T >$ 15 in data (left) and MC (right), requiring to pass the W T&P triggers. The color labels represent different selection mode: black for Very Loose, Red for Loose, Blue for Medium and Cyan for Tight selection.

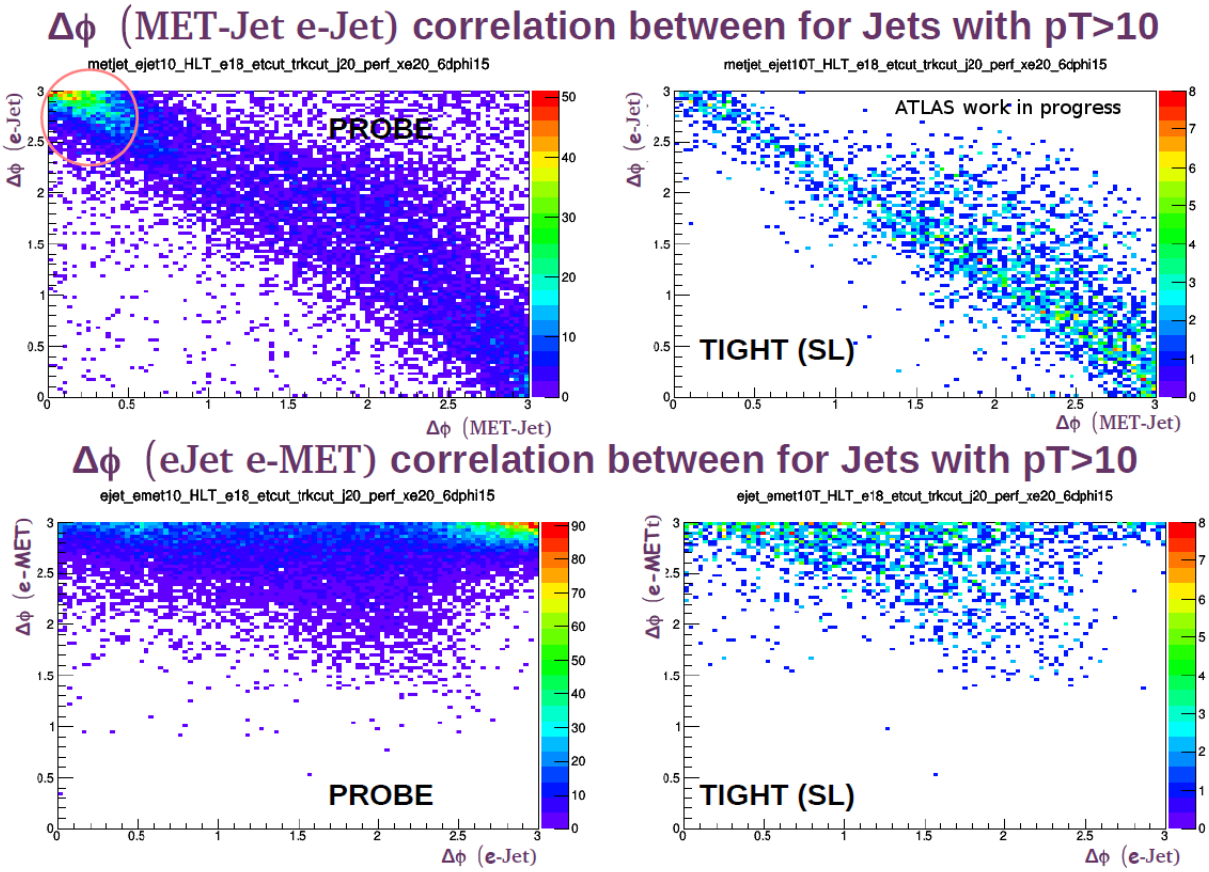


Fig. 3.10 The correlation between $\Delta\Phi_{jet-MET} - \Delta\Phi_{e-jet}$ on top and $\Delta\Phi_{e-jet} - \Delta\Phi_{e-MET}$ on bottom between the 6 leading in p_T jets with $p_T > 10$ and the transverse missing energy/probe electron for Probe (left) and Tight (right) selection.

- $R_\Phi > 0.7$.

The R_Φ is the ratio of the energy in 3×3 cells over the energy in 3×7 cells centered at the electron cluster position.

Although the Signal to Background ratio improves by applying cuts on different variables, this is not enough to have a pure signal sample in Tight requirement, some background events still surviving which need to be subtracted. Therefore, another method using the TMVA technique was tested, as will be presented in the following section.

3.5 TMVA techniques to further reduce the background

TMVA (Toolkit for Multivariate Analysis) provides a ROOT-integrated machine learning environment for the processing and parallel evaluation of sophisticated multivariate classification techniques [14].

Currently, supported methods in TMVA includes cuts (optimizers: MC sampling, Genetic alg, simulated annealing), 1-dimensional Likelihood Estimator, Multi-dimensional Probability Density Estimator (range search), Fisher Discriminant, Artificial Neural Network, Boosted/bagged Decision Trees.

For my analysis, I am using a (Boosted) Decision Tree (Gradient) method (BDT and BDTG, which is a BDT that uses on top of Gradient Boost). The structure classifier includes S (signal) and B (background) leafs (see Figure 3.11). It classifies events by following a sequence of decisions depending on the events variable content until a pure S or B leaf is reached.

In order to reduce the background especially at low p_T, two techniques are applied:

- Cut Based Selection: with the kinematics selection presented in the previous section
- BDTG based on kinematics and tunned in bins of p_T (in view of the much better performances for TMVA cuts)

  The variables used for the TMAVA training are selected in such a way to give the best signal to background discrimination, keeping in the analysis only the variables with the highest separation power. They are only related to kinematics (not related to isolation distribution used in the efficiency calculation to not introduce a bias in the electron

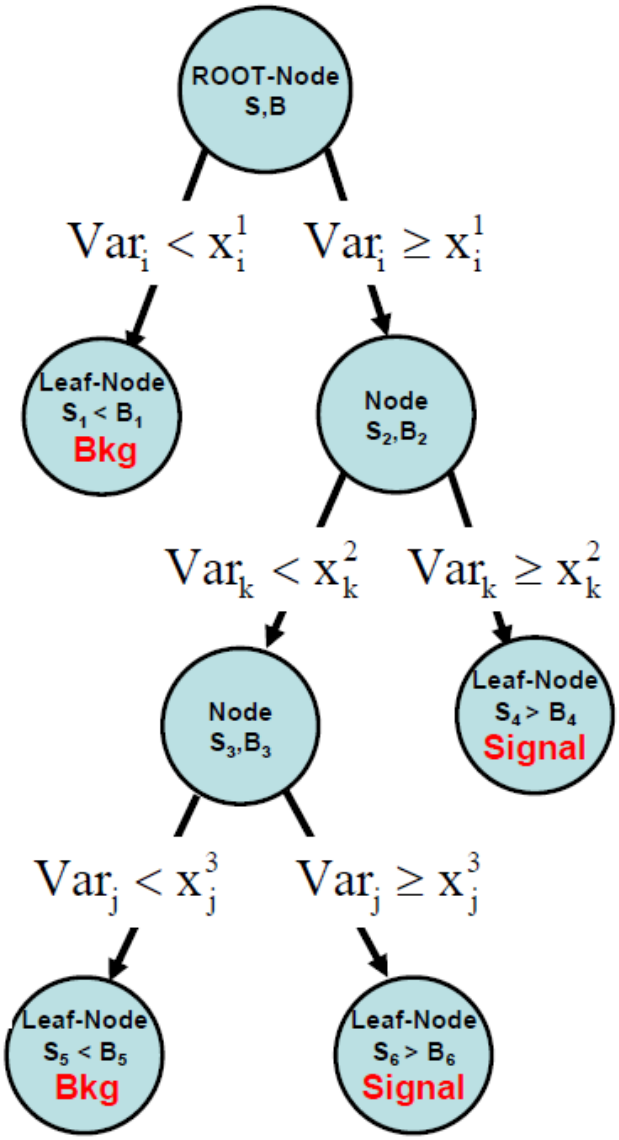


Fig. 3.11 The (Boosted) Decision Trees classifier.

efficiency measurement). Another requirement is for the variables to be uncorrelated.

After optimization studies on the discrimination power of each variable, the list of variables chosen for the TMVA training are:

1. Transverse mass
2. MET
3. MET_Sumet (ΣE_T)
4. MET significance (xs)
5. $\Delta\Phi_{JETMET}$ (the angle between the missing transverse energy direction and the highest p_T jet)
6. $\Delta\Phi_{JETEle}$ (the angle between the probe electron direction and the highest p_T jet)
7. $\Delta\Phi_{METEle}$ (the angle between the missing transverse energy direction and the probe electron direction)
8. p_T of the probe electron (variable used also for the binning of the identification efficiency measurement)

The distributions of the variables used in TMVA are shown in Figure 3.12 for e13/e18 energy threshold triggers (as explained previously in section 3.1.4) in the $20< p_T <25$ range for !VL LH (probe electron fail Very Loose selection) background. The background correlation is presented in Figure 3.13 and the signal correlation matrix in Figure 3.14). The correlations between TMVA variables will introduce a bias in the efficiency measurement. Therefore, the list of variables presented was chosen in such a way to provide the most powerful discrimination without introducing a bias that will be difficult to control.

For the training of the Signal, the sample that passes the "Tight requirement" is used while for the background, the sample that passes a dedicated background template condition is used.

The BDTG output of the two background templates (fail VeryLoose LH, !VL LH and Template1, labeled as T1) is shown for one p_T bin in the Figure 3.15 for the e13/e18 energy threshold triggers and in Figure 3.16 for high MET triggers. From this, we observe similar TMVA output behavior, the difference between them being included in the systematic error. Therefore, for consistency and simplicity, in the following results only the !VL LH background template TMVA output is used to discriminate the signal events from the background events.

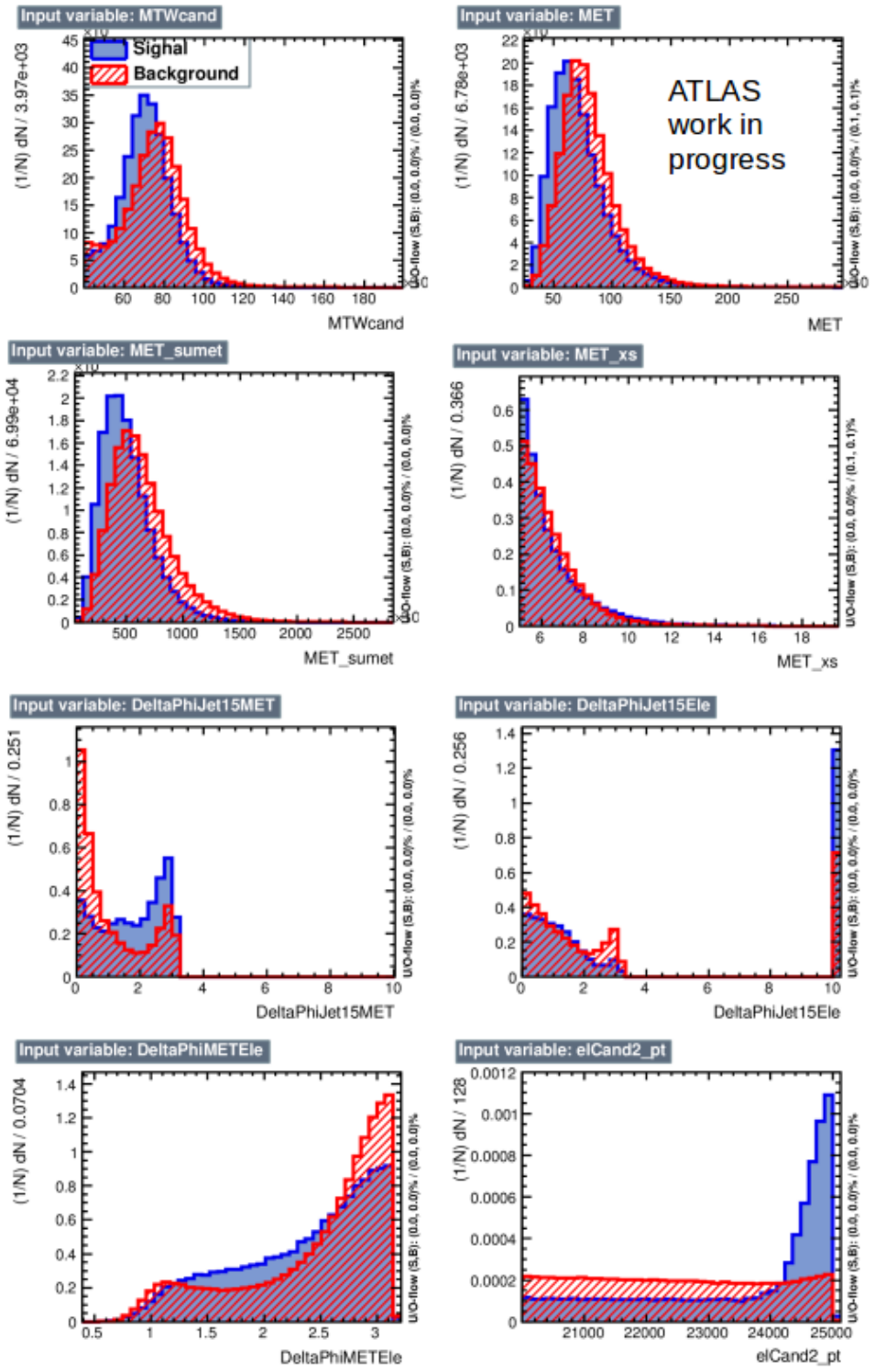


Fig. 3.12 TMVA variables distribution for background (red) and signal (blue) for e13/e18 energy threshold triggers in the 20< p_T <25 GeV range for !VL LH background

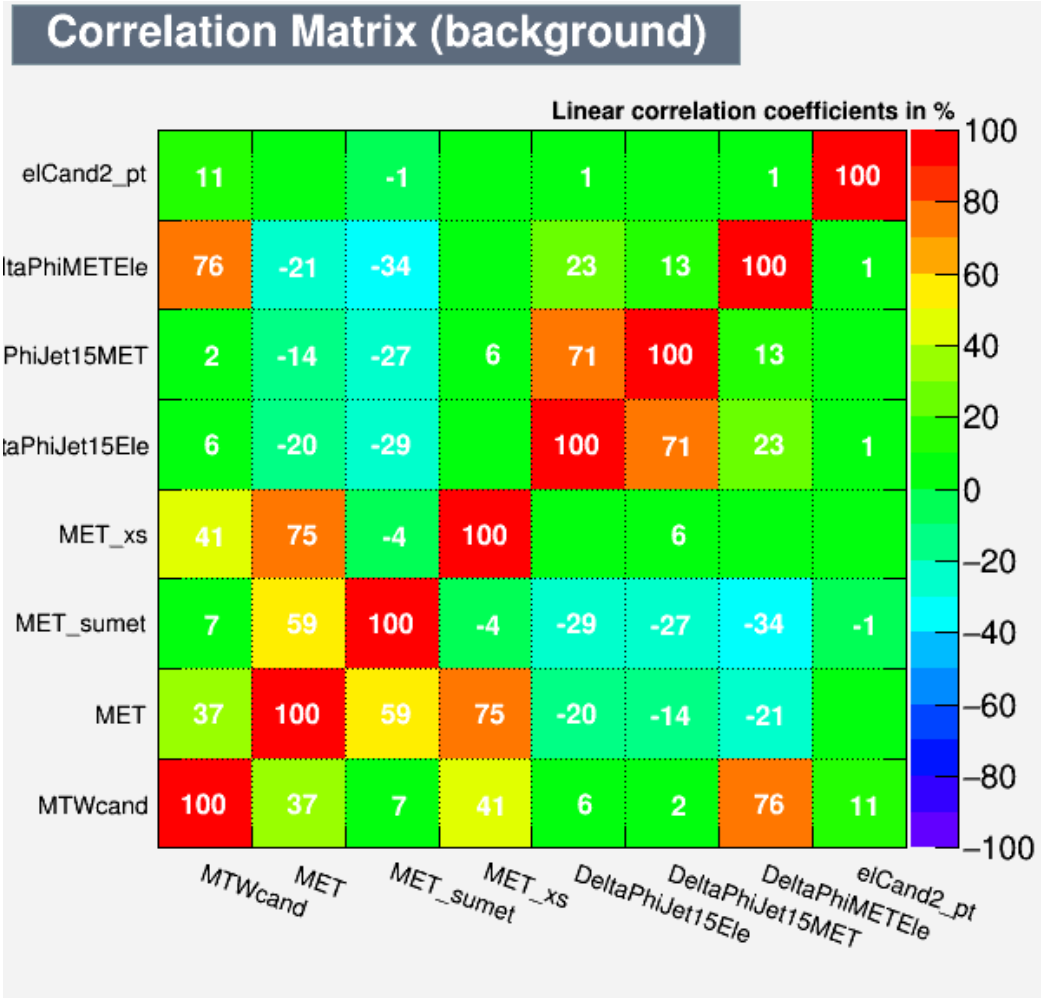


Fig. 3.13 TMVA variables background correlation for e13/e18 energy threshold triggers in the 20< p_T <25 GeV range

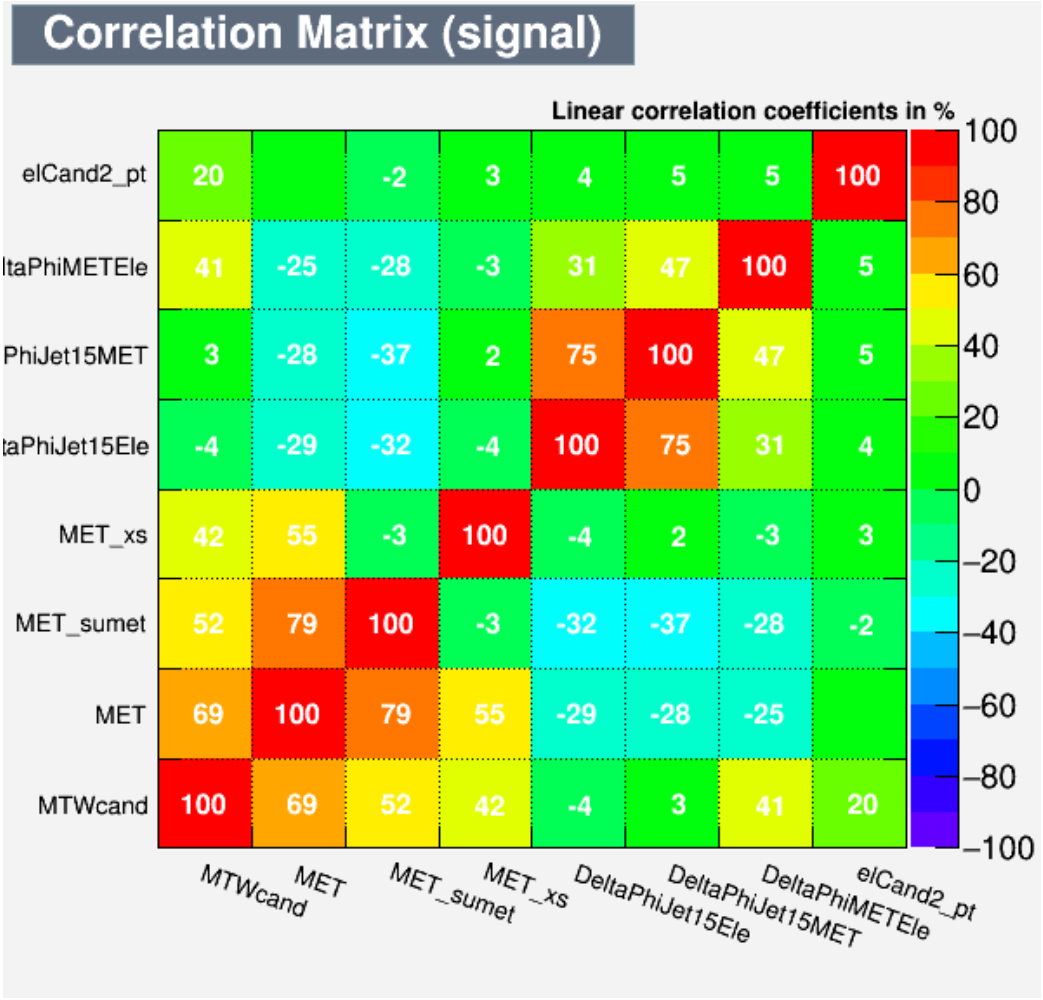


Fig. 3.14 TMVA variables signal correlation for e13/e18 energy threshold triggers in the 20< p_T <25 GeV range

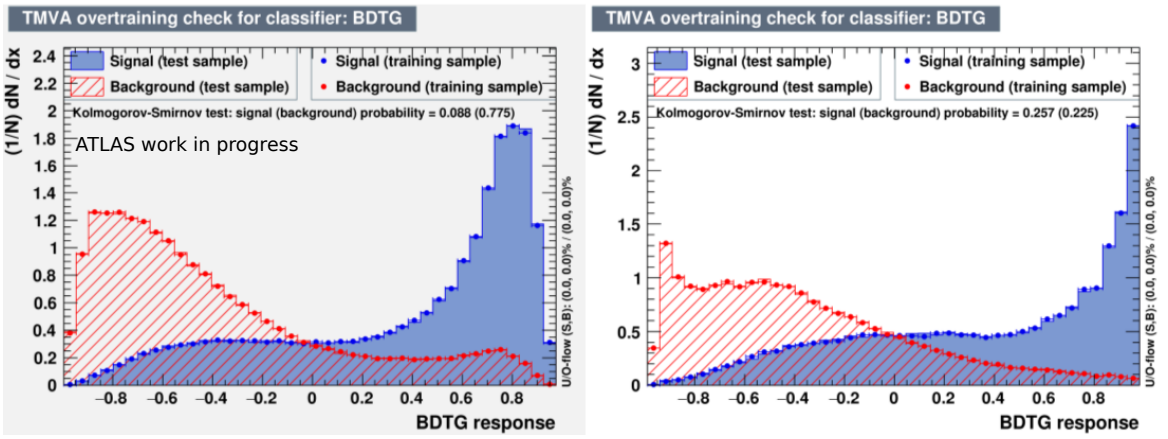


Fig. 3.15 TMVA training on 20< p_T <25 GeV Tight probes against !VL LH background template (left) and T1 (right) for the e13/e18 energy threshold triggers

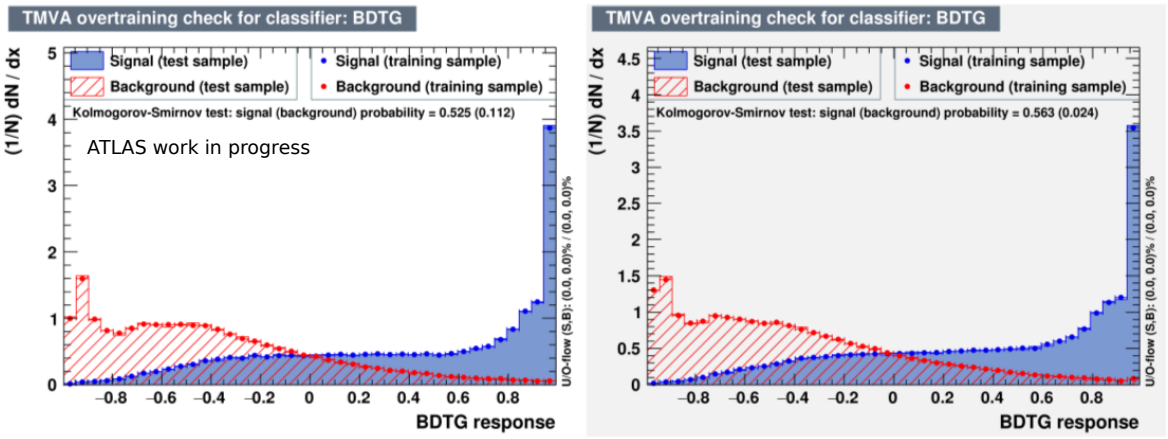


Fig. 3.16 TMVA training on 20< p_T <25 GeV Tight probes against !VL LH background template (left) and T1 (right) for the high MET triggers

One technical way to assure the consistency of the training, the cross-validation training and testing were checked (train on odd events and test on even and train on even events and test on odd events). The default BDTG setting is to choose randomly half of the events for training and the other half left for the testing. In the double-inverse training and testing, we choose to train on odd events and apply the TMVA output on even events and train on odd events and apply the TMVA output on even events. This allows us to use the entire statistics available in the sample and check if a bias is introduced by a random default setting. The BDTG output is shown on the left in Figure 3.17 for training on odd events and in Figure 3.18 for training on even events, while on the right I show the ROC curves (the signal efficiency as a function of background rejection) for the two configurations. The difference between the outputs is negligible, so the default setting (randomly half of the events training and the other half left is used for the testing) was chosen for simplicity.

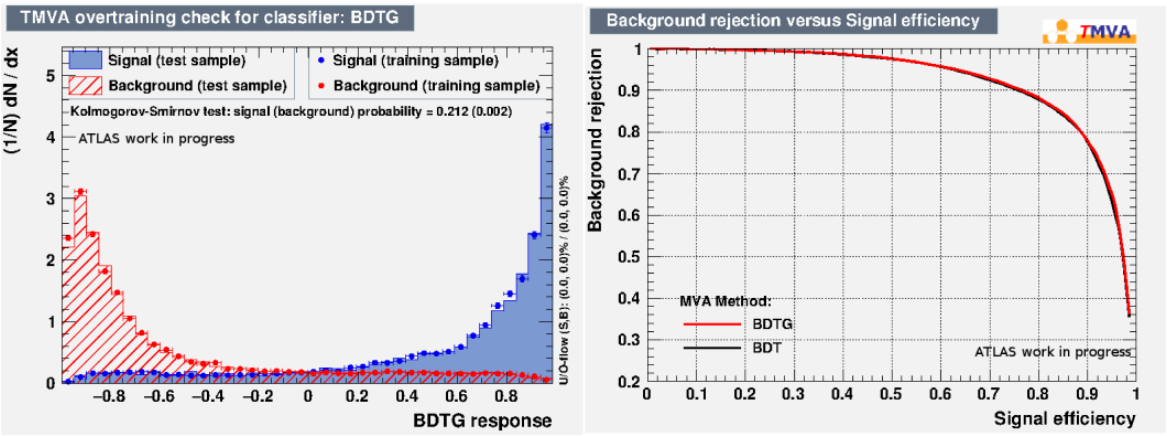


Fig. 3.17 TMVA BDTG output on the left and ROC curve on the right for training on odd events, with 20<p_T<25 GeV, with Signal Tight (Red) against !VL LH background template (Blue).

The difference between the cut based and TMVA based techniques regarding the background rejection is overwhelming especially at low p_T (background dominated region) but also in the rest of the transverse momentum range (see Figure 3.19 for e13/e18 energy threshold triggers, Figure 3.20 for high MET triggers in the 15< p_T <20 GeV range and 0.1< η <0.6 for !VL LH background and Figure 3.21 for e13/e18 energy threshold triggers + high MET triggers in 0.1< η <0.6 and in the 60< p_T <80 GeV range for !VL LH background template). The colors for the left top plot are associated to Signal selection menu (Tight represented in Pink) and Background menus (Blue for T1, Red for T6 and Green for !VL LH), the Probe menu being represented in Black. On the left, we see the variables we

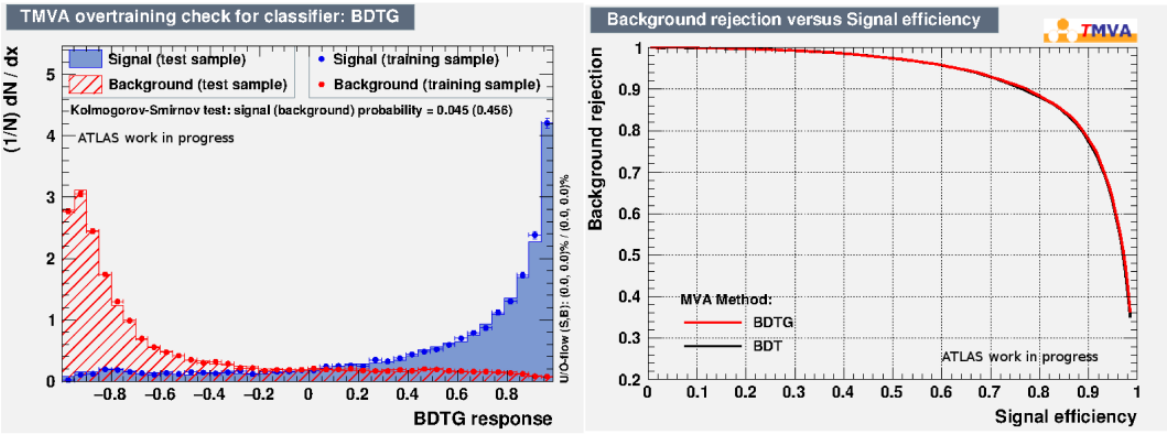


Fig. 3.18 TMVA BDTG output on the left and ROC curve on the right for training on even events, with 20<p_T<25 GeV, with Signal Tight (Red) against !VL LH background template (Blue).

are cutting on: top left for the Cut-Based method I represent the $\Delta\Phi_{JETMET}$ with the cut on 0.7. On the left bottom, we see the TMVA output for !VL LH background (Red) and 10 x Signal Tight (Blue), as well as the cut chosen of TMVA>0. On right part we see the isolation distribution on Probe level, the background is represented in Blue. The Signal is therefore situated between the blue template and data points. The biggest the white part, the highest signal to background ratio.

This result is a strong argument to use a TMVA technique to increase the signal purity in W T&P analysis. The TMVA was never used in performance analysis, my studies providing a novel method of background rejection, more efficient than a cut based method, that could be used by all users in the collaboration.

The optimal cut on the TMVA is decided accordingly to the maximum "significance" defined for simplicity as $\frac{S}{\sqrt{(S+B)}}$ by varying the TMVA cut point from -1 to 1 as can be seen as an example in Figure 3.22 for the 30< p_T <35 range where a cut on 0.2 is set (for which $S/\sqrt{(S+B)}$=129.46). The 100 bins represents the TMVA range from -1 to 1 for a step of 0.02 (-1+i*0.02, i=0,100). This procedure of finding the optimal cut on TMVA is applied in each p_T bin.

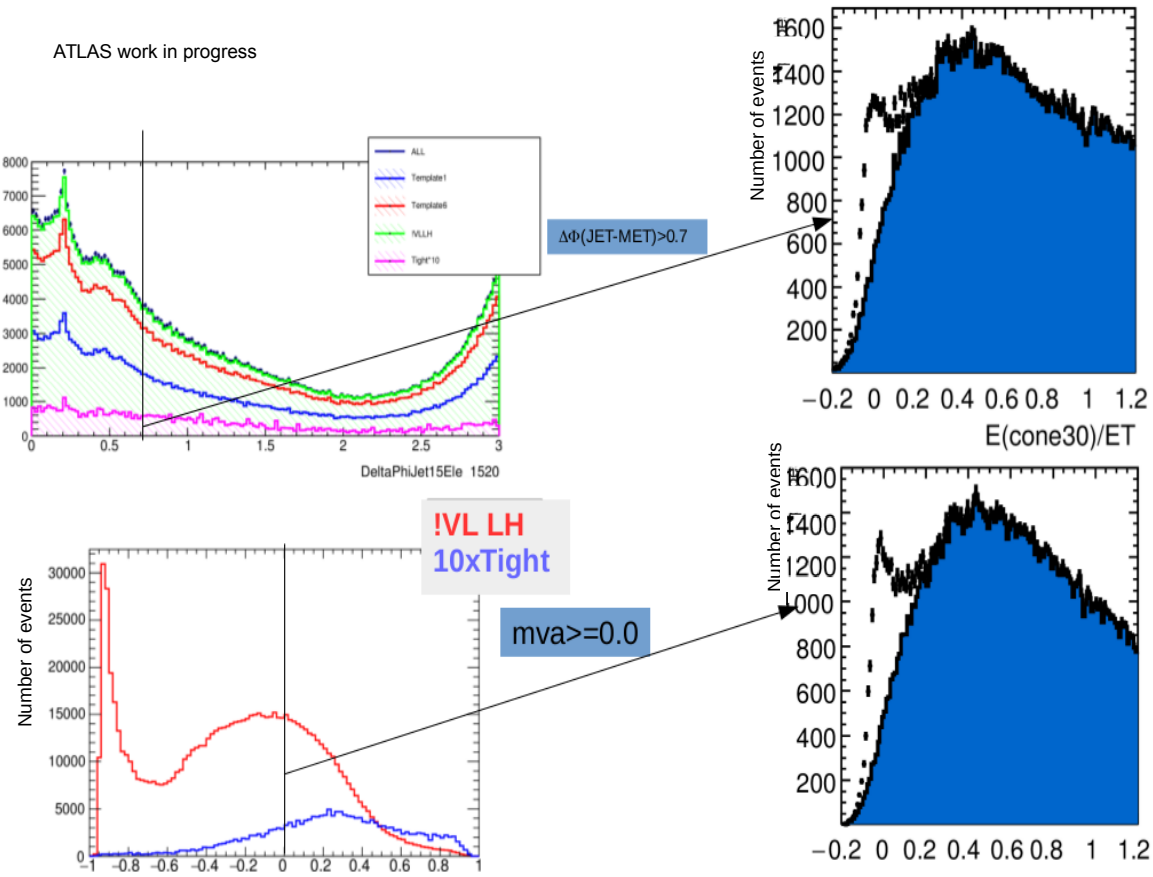


Fig. 3.19 The comparison between Cut-Based and TMVA background reject: On the top left for the Cut-Based method is represented the distribution of $\Delta\Phi_{JETMET}$ with the cut on 0.7. The colors for the left top plot are associated to Signal selection menu (Tight represented in Pink) and Background menus (Blue for T1, Red for T6 and Green for !VL LH), the Probe menu being represented in Black. On the left bottom, it is presented the TMVA output for !VL LH background (Red) and 10 x Signal Tight (Blue), as well as the cut chosen of TMVA>0. On right part we see the isolation distribution on Probe level, the background is represented in Blue. The Signal is therefore lying between the blue template and data points. The training is performed by matching the probe electron to e13/e18 energy threshold triggers in 0.1< η <0.6 and in the 15< p_T <20 GeV.

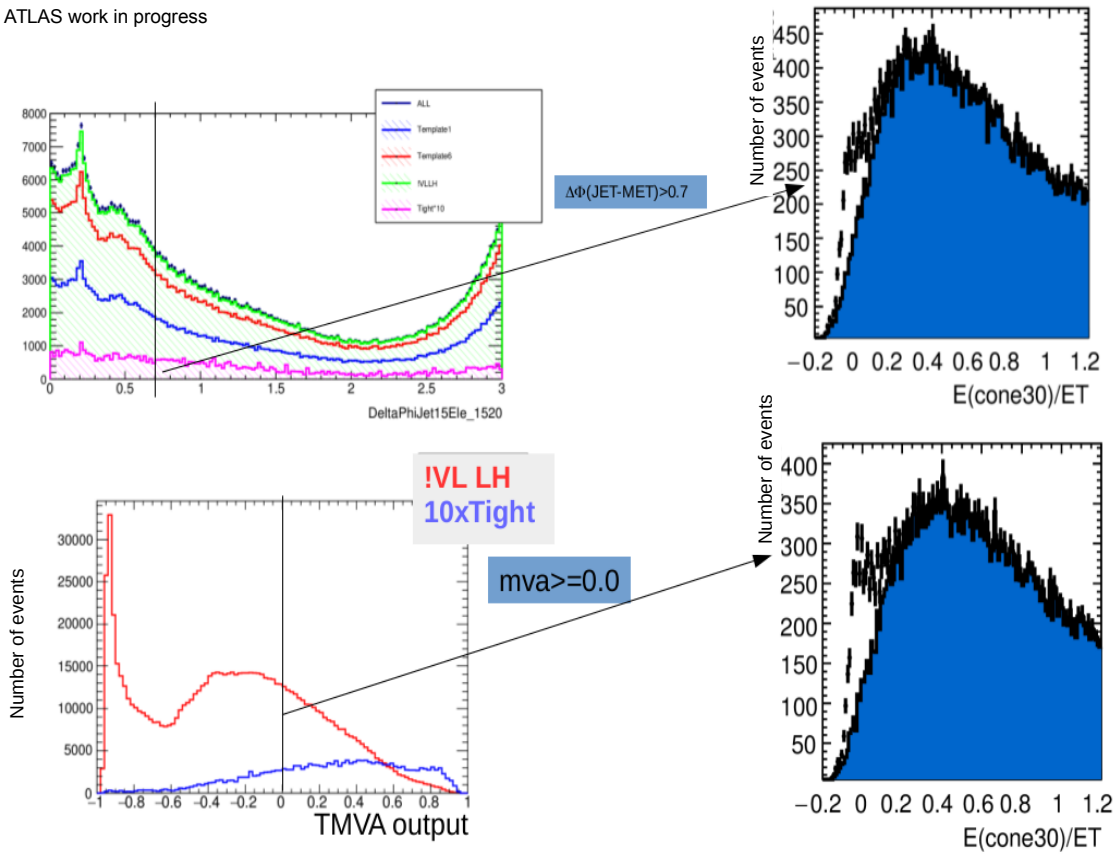


Fig. 3.20 The comparison between Cut-Based and TMVA background reject: On the top left for the Cut-Based method is represented the distribution of $\Delta\Phi_{JETMET}$ with the cut on 0.7. The colors for the left top plot are associated to Signal selection menu (Tight represented in Pink) and Background menus (Blue for T1, Red for T6 and Green for !VL LH), the Probe menu being represented in Black. On the left bottom, it is presented the TMVA output for !VL LH background (Red) and 10 x Signal Tight (Blue), as well as the cut chosen of TMVA>0. On right part we see the isolation distribution on Probe level, the background is represented in Blue. The Signal is therefore lying between the blue template and data points. The training is performed by matching the probe electron to high-MET triggers in $0.1< \eta <0.6$ and in the $15< p_T <20$ GeV.

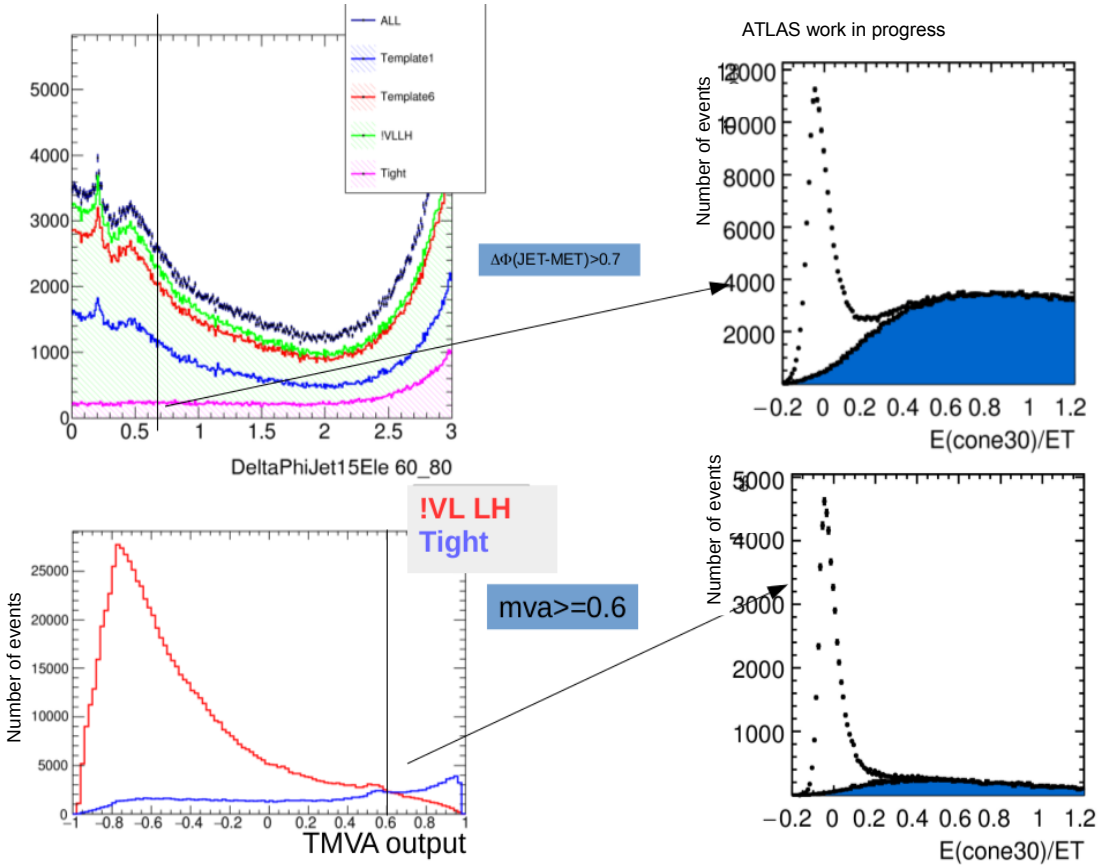


Fig. 3.21 The comparison between Cut-Based and TMVA background reject: On the top left for the Cut-Based method is represented the distribution of $\Delta\Phi_{JETMET}$ with the cut on 0.7. The colors for the left top plot are associated to Signal selection menu (Tight represented in Pink) and Background menus (Blue for T1, Red for T6 and Green for !VL LH), the Probe menu being represented in Black. On the left bottom, it is presented the TMVA output for !VL LH background (Red) and 10 x Signal Tight (Blue), as well as the cut chosen of TMVA>0. On right part we see the isolation distribution on Probe level, the background is represented in Blue. The Signal is therefore lying between the blue template and data points. The training is performed by matching the probe electron to e13/e18 energy threshold + high MET triggers in 0.1< η <0.6 and 60< p_T <80 GeV range.

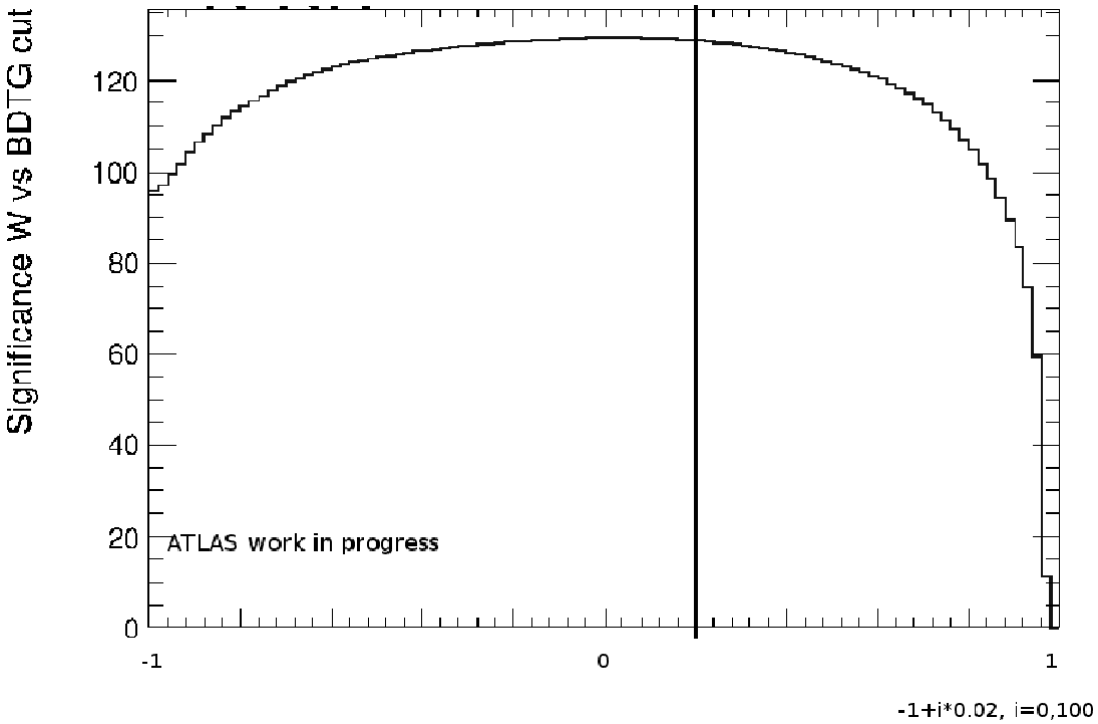


Fig. 3.22 The signal significance ($\frac{S}{\sqrt{(S+B)}}$) as a function of the TMVA cut for the e13/e18 energy threshold triggers + high MET triggers in the 30< p_T <35 GeV range for !VL LH background template. The 100 bins represents the TMVA range from -1 to 1 for a step of 0.02 (-1+i*0.02, i=0,100).

3.6 Analysis procedure to measure the electron ID efficiency using Isolation in 2015-2016 data

The Isolation Method is used for $W \to e\nu$, using as a discriminating variable between signal and background, the calorimetric isolation E_T^{cone}/E_T for low and middle energy electrons ($E_T < 45$ GeV) and $E_T^{cone}/min(E_T^{cone}, 25GeV)$ for high energy electrons ($E_T > 45$ GeV) in order to avoid a complete lack of events in the control region used to normalized the background.

To assess the background, a model is constructed representing the shape of the isolation distribution for fake electrons. For the construction of a background template, cut based (for the T1 and T6 background templates) and likelihood menus (for the !VL LH background template to fail the VeryLoose identification) are applied. The use of more than one background templates allows a straightforward cross-check of the analysis. The difference between the templates are included in the systematic uncertainties. The residual signal contamination in the background templates is expected to be very small (in particular for cut-based BG templates). However, it is subtracted as will be discussed in this section.

To provide a cross-check for our analysis, we have performed the Isolation method to extract the efficiency using simple background template subtraction using $Z \to e^+e^-$ tag and probe as described in the following subsection.

3.6.1 $Z \to ee$ T&P cross-check

For the Z analysis, the full Derived dataset (DAOD) EGAM1 at full luminosity was tested.

The Z Selection require two electrons to pass specific Z Triggers. For the tag and probe electron it is required to pass $p_T > 15$ GeV and $|\eta| < 2.47$. The dilepton mass of the electron pair should be in the mass window of 100 GeV $> m_{ll} >$80 GeV.

For the Z Isolation, the simple background template subtraction is performed. The probe electrons in each E_T and η binning are subdivided into a signal dominated and a background dominated region, used as signal and control region respectively, at a certain isolation threshold which is chosen to be 0.4.

All electrons with isolation energies higher than this threshold are taken to be background electrons. Therefore the number of electrons in the selected background samples above the isolation threshold is normalized to the number of events in data above the threshold, using a range in the isolation variable of up to 1.5. The number of background electrons with isolation energy smaller than the threshold is subtracted from the data, taking into account the normalization in the control region. An example of the Isolation Distribution (E_{cone30}/E_T) is depicted in Figure 3.23 where it is visible the difference between the Z and W analysis both on statistics but also in shape.

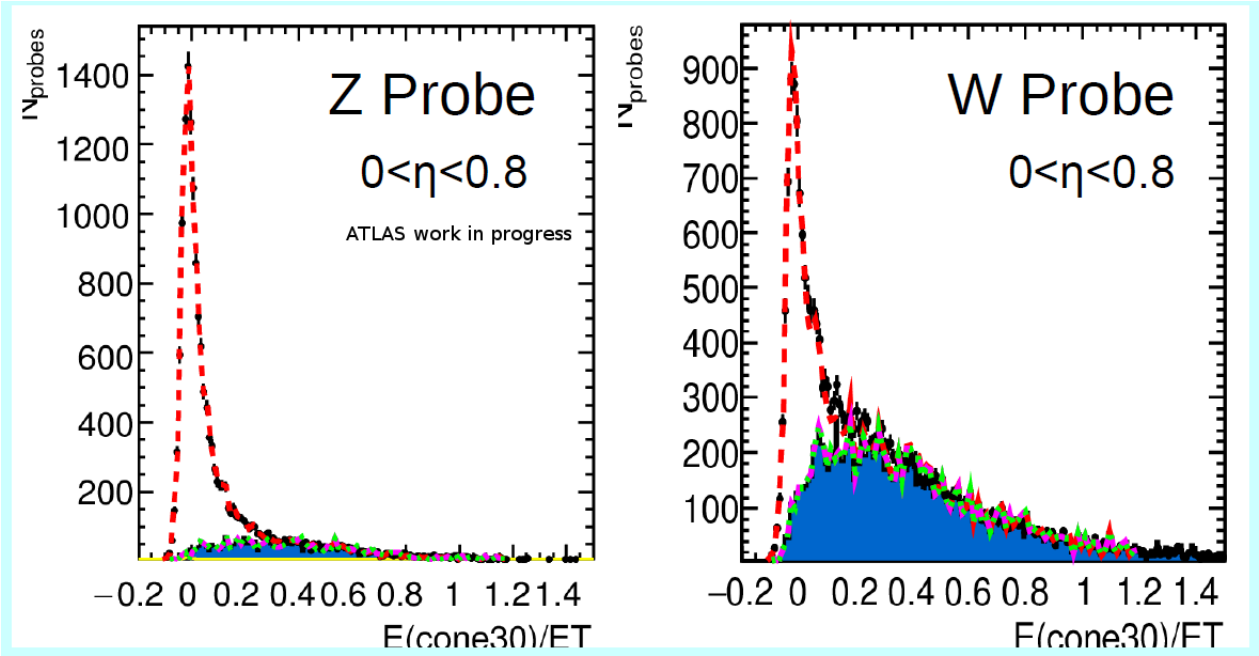


Fig. 3.23 Example of Isolation Distribution at Probe Level of $Z \to ee$ events (left) and $W \to e\nu$ at $0.0 < \eta < 0.80$ for all p_T bins. The background template used is !VL LH (blue color).

3.6.2 Background Template Construction

The three Background Templates from Cut-Based and Likelihood menus constructed as discussed above are used in the function to fit the isolation distributions, their comparison revealing puzzling effects that are not yet understood completely. The difference between the background templates is taken as systematic uncertainty.

First, we look in the W T&P analysis at the ratio of the number of background events to the number of probes (N_{BG}/N_{Probe}) in the Signal Region, defined in the range of (-0.2,0.4) of the isolation distribution, after background normalization and signal like events extraction, using the 3 different background templates: !VL LH background is represented in Black, T6 in Red and fail Medium background (fail the Medium Identification criterion) in Green (see Figure 3.24). We can conclude that the stability of the background estimation is worse on W than on Z (Figure 3.25) namely the !VH LH template has significantly different behavior than the cut-based ones, the Template6 being the most stable both in W and Z analyzes.

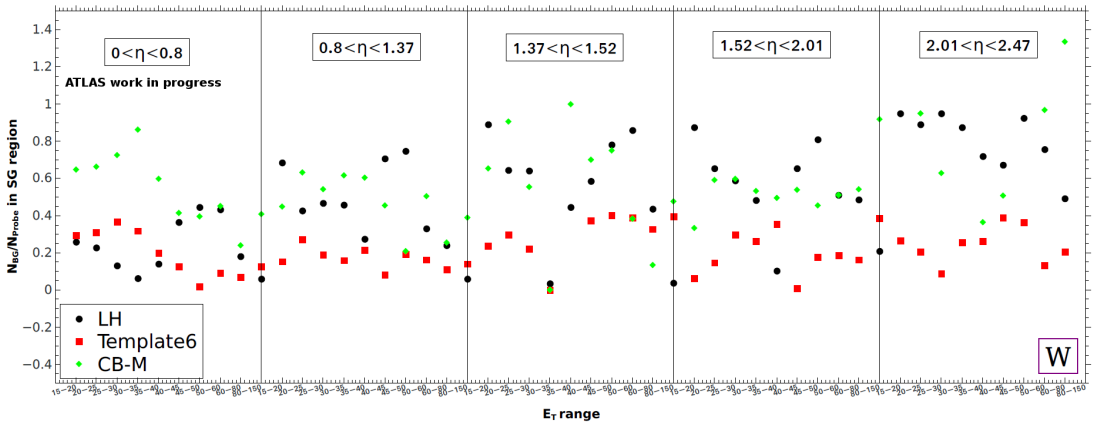


Fig. 3.24 N_{BG}/N_{Probe} for W analysis in the Signal Region (-0.2,0.4) calculated using the 3 different background templates: !VL LH background is represented in Black, T6 in Red and fail Medium background in Green.

The background templates are built by inverting some electron-background discrimination variables. The background does contain a small contribution from the true electrons collected by the background selection criteria. These true electrons presented in the background distributions distort the background template by enhancing the background shape in the signal region and can, therefore, lead to an overestimation of the efficiency. This signal contribution should, therefore, be removed from the background template.

For each bin and each background template conditions, the rate with which a true electron is selected as background can be estimated using Monte Carlo simulations using the following formula:

$$\varepsilon_B = \frac{N_B^{matched}}{N_{Probe}^{matched}} \tag{3.14}$$

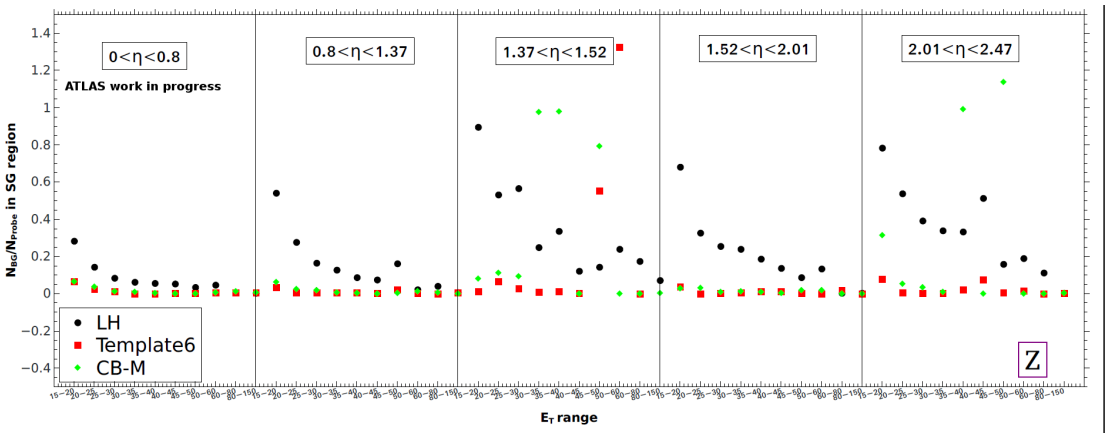


Fig. 3.25 N_{BG}/N_{Probe} for Z analysis in the Signal Region (-0.2,0.4) calculated using the 3 different background templates: !VL LH background is represented in Black, T6 in Red and fail Medium background in Green.

where $N_B^{matched}$ is the number of electrons, matched to a truth electron and selected by the respective background template conditions, while $N_{Probe}^{matched}$ is the total number of electron candidates in the given bin without other condition (”matched probe level”).

The ε_B using the three background templates for W and Z can be seen in Figure 3.27 for W analysis and Figure 3.28 for Z analysis. An example of the signal like events in the background template for W (the yellow distribution) is shown in Figure 3.26.

For a given background enriched sample, in given phase space and analysis bin, the absolute contribution of the signal to the background enriched sample (background template) should be given by: $\varepsilon_B \times N^{all}$ where N^{all} is the total number of electrons in the given phase space. Those events have to be subtracted from the background enriched sample (”template”) before any fits on data are performed. The total number of electrons N^{all} is estimated from the ”Tight” distribution, corrected by the respective efficiency:

$$BG_{corrected} = BG - \frac{\varepsilon_B}{\varepsilon_T} \times T \tag{3.15}$$

Here $BG_{corrected}$, B, and T are isolation distributions. It should be underlined that this subtraction should be done before any other renormalization of the template. The formula for background correction is not ”closed”, since the tight selection efficiency is needed as input. Therefore an iterative approach will be needed, for which the most straightforward approach is to calculate first an efficiency without the signal subtraction from the background template.

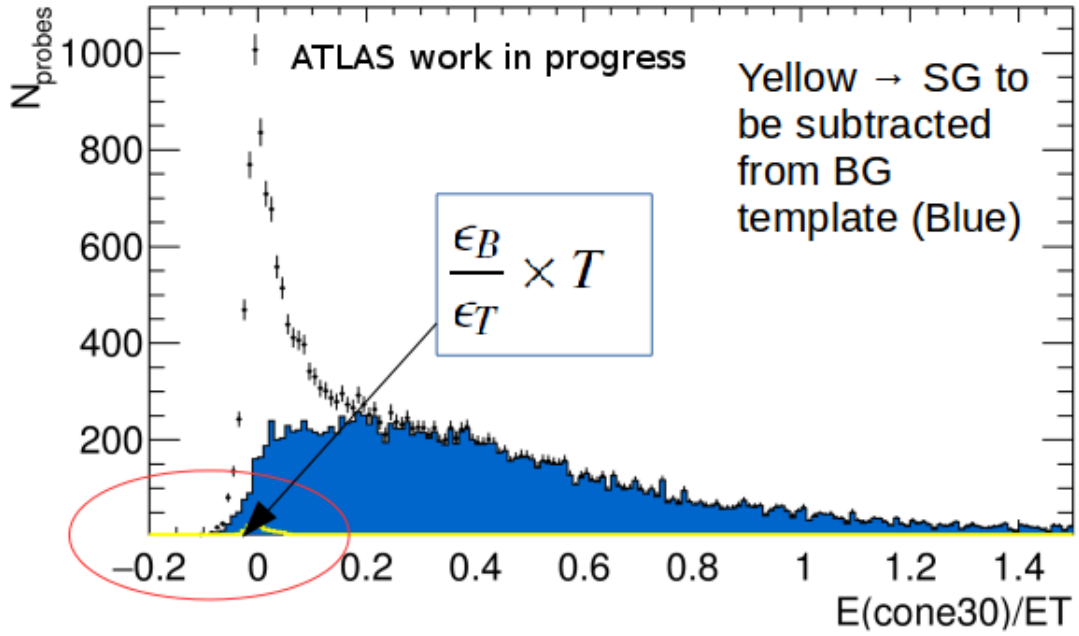


Fig. 3.26 The amount of signal like events in the background templates (yellow) for W analysis.

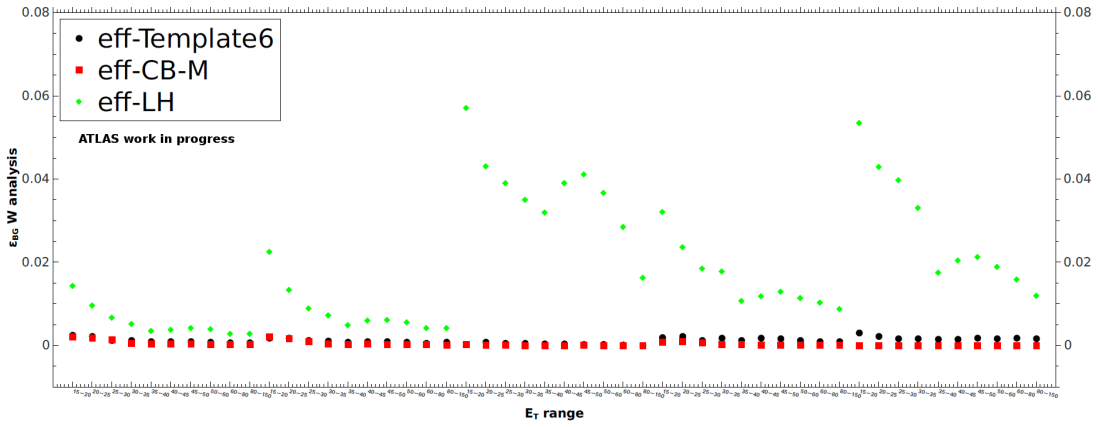


Fig. 3.27 The ε_{BG} using the 3 background templates for W T&P: !VL LH background is represented in Black, T6 in Red and fail Medium background in Green.

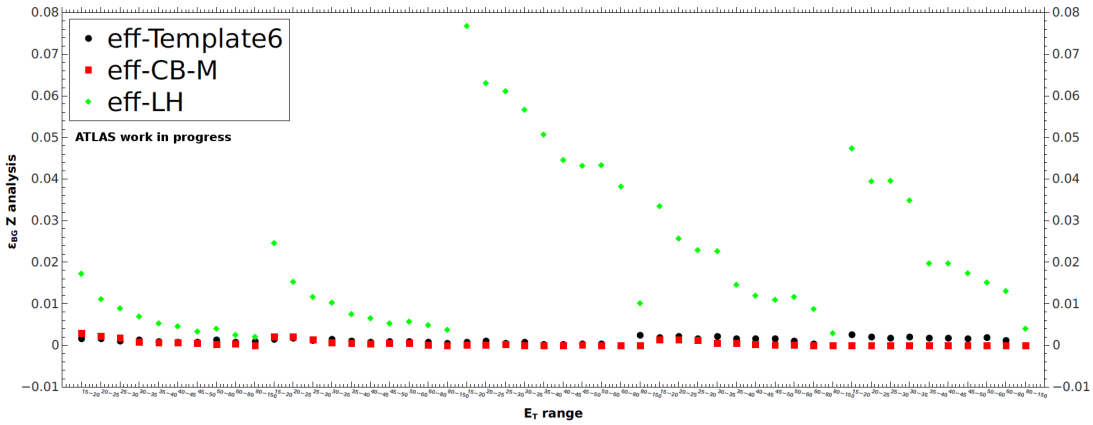


Fig. 3.28 The ε_{BG} using the 3 background templates for Z T&P: !VL LH background is represented in Black, T6 in Red and fail Medium background in Green.

This approximated efficiency is then used to obtain the corrected background template.

For each selection level A (VL, L, M, T), the efficiency is then calculated.

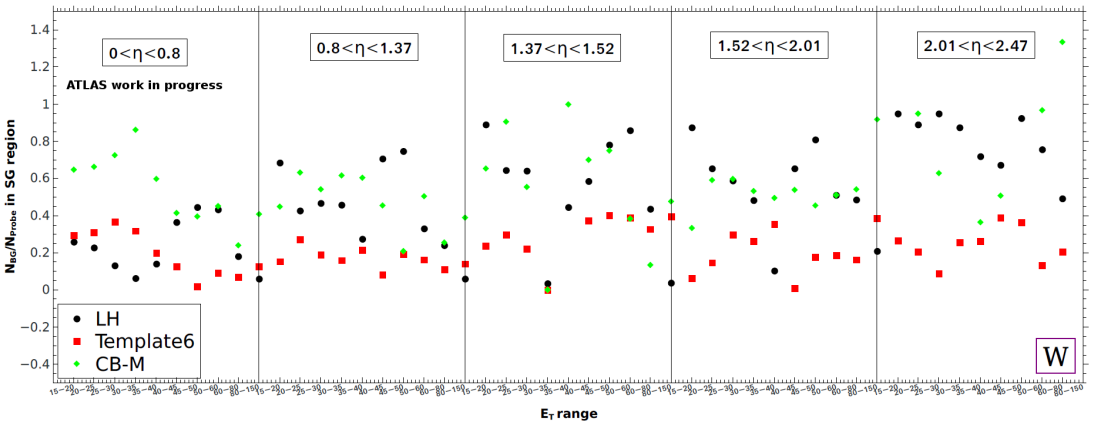


Fig. 3.29 The N_{BG}/N_{Probe} ratio using the 3 background templates for W T&P: !VL LH background is represented in Black, T6 in Red and fail Medium background in Green.

To have a hint on the number of signal like events in the background templates and the impact on the subtraction of this amount, one can see in Figure 3.29 and Figure 3.30 for W and Z analysis that, the ratio of Number of Background events to Number of Probe events (the blue template divided to the white part from the 3.26 distribution) is more stable for Cut-Based menus (Red and Green points) than for the !LH template (Black points), both

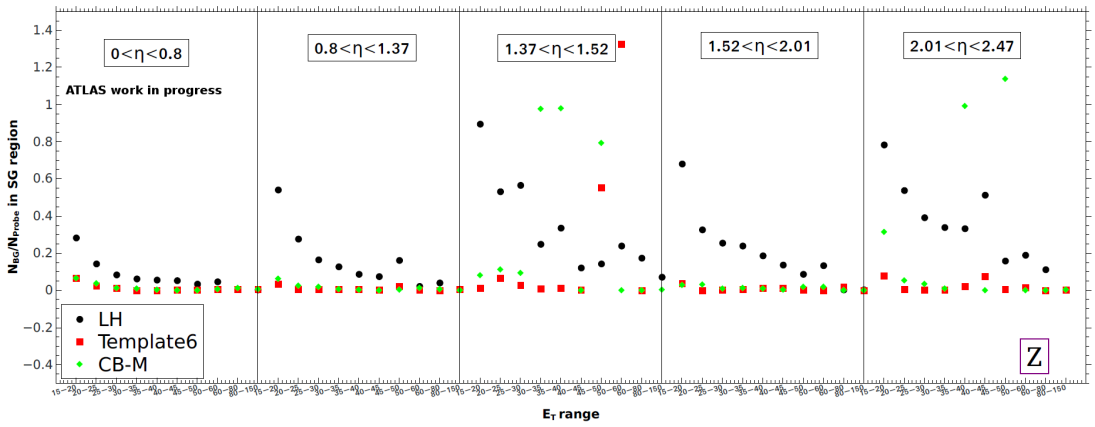


Fig. 3.30 The N_{BG}/N_{Probe} ratio using the 3 background templates for Z T&P: !VL LH background is represented in Black, T6 in Red and fail Medium background in Green.

for W and Z analysis. This is also in agreement with the Efficiencies of the Background Template distributions which tends to be stable and in agreement for Template6 and CB-M and puzzling for !LH template.

For MC samples, a fail truth ID Background Templates is constructed, for both W and Z analysis. In this template, we are sure that only non-signal electrons will survive, as can be seen in Figure 3.31 where the majority of candidates are background electrons and hadrons.

In conclusion, the Background Template for !VL LH is still problematic, the number of signal like events surviving in the background templates is much higher than for the cut-based templates. In addition, the background template shape is distorted and different, especially in the Signal Region, making impossible after renormalization to fit both the LH and CB menus in the same manner.

3.6.3 Signal Templates Construction in the Cut Based Method

For the Signal Templates, two (considered background-free) distributions are used:

- The isolation distribution from $Z \rightarrow ee$ events
- The isolation distribution from the Tight identification menu

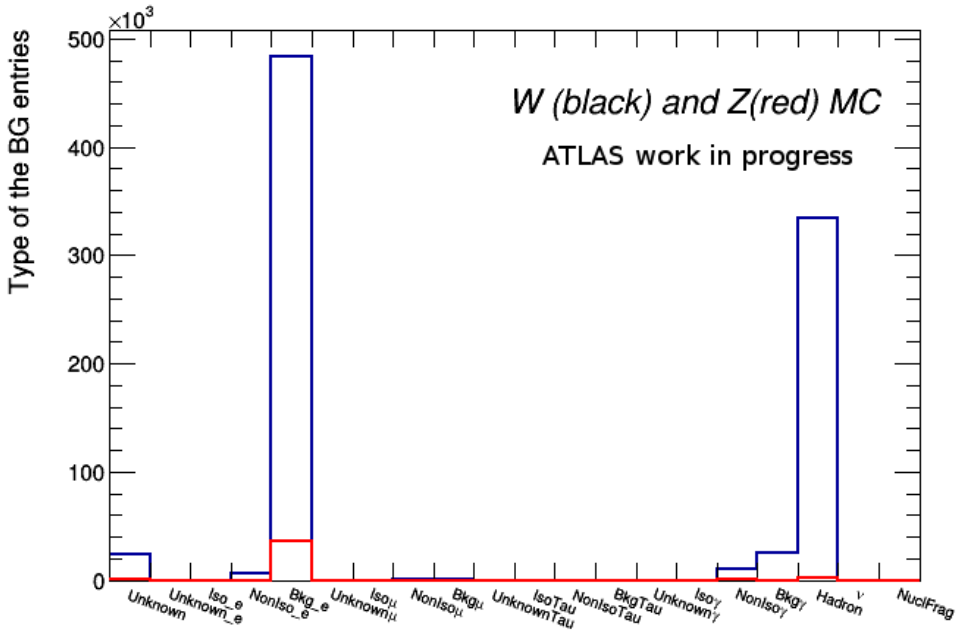


Fig. 3.31 The type of the particles that enters into fail electron Truth Type Background. The blue distribution reflects the W TP analysis while the red color belongs to Zee TP analysis.

with the option of adding an extra Gaussian to the function fit to compensate the gap between data and Background (BG) template.

However, it is worth mentioning that both of the Signal Templates used are not entirely background free and also that from Tight to L, M, VL and Probe menus, the shape of the isolation distribution is also changing (per thousand effects) so to be sharply correct in the analysis, other signal templates may be chosen to not bias further the efficiency results.

Several analysis have been performed concerning the choice of the signal templates and their contribution to the efficiency extraction.

Firstly, the choice of the signal templates should not influence the background templates. To verify that, in the Figure 3.32 we see the W efficiency obtain with (yellow line)/without (black line) signal like events extraction from the background templates using both signal templates from Tight Isolation Distributions (continues line) and from ZIso Distributions (dotted line) All identification menus are presented: VL in Black, Loose in Red, Medium in Blue and Tight in Pink. The variations had a per thousand effect so no BG normalization

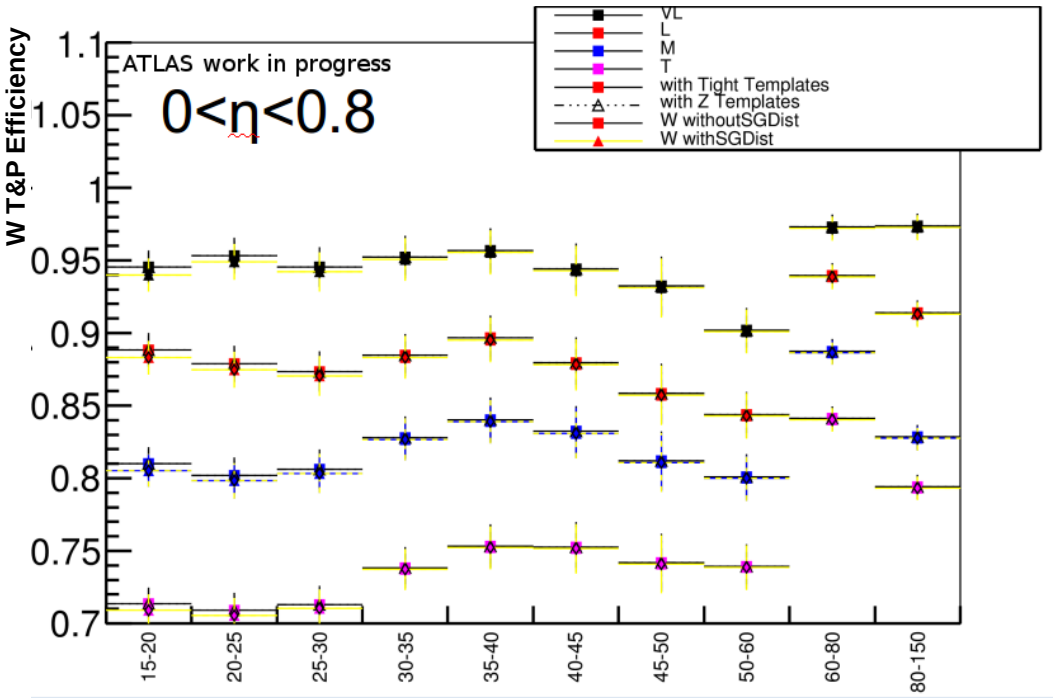


Fig. 3.32 The W T&P efficiency obtain with (yellow line)/without (black line) signal like events extraction from the background templates using both signal templates from Tight Isolation Distributions (continues line) and from ZIso Distributions (dotted line). All identification menus are presented: VL in Black, Loose in Red, Medium in Blue and Tight in Pink.

dependency on SG template can be observed.

Secondly, the W T&P Efficiencies in Template6 and !VL LH background template is shown in Figure 3.33 and Figure 3.34. All identification menus are presented: VL in Black, Loose in Red, Medium in Blue and Tight in Pink. We observe that, for the central η region, the 2 signal templates (Tight Isolation on the left and from ZIso Distributions right) are in agreement, and by applying the SG subtraction from BG templates (yellow line), efficiency varies up to 1%, behavior kept along with all η ranges. For all BG templates, the discrepancy is significantly higher at low p_T indicating the background domination in the low energy region coming from the lack of purity in the data samples.

In conclusion, considering the compatibility between the 2 signal templates, the Tight distribution was used as a Signal template for final efficiency calculation, the difference in the fit functions being included as systematics.

3.6.4 The signal like events in the background templates comparison between cut-based method and TMVA

By applying the TMVA cut, the reduction of the background will also influence the number of signal like events in the background templates that need to be subtracted by using the MC truth ID. Figure 3.35 shows the amount of the signal like events in the !VL LH background in both the TMVA and Cut-Based method for all 200 E_T - η bins (on the x-axis is the η binning while each pad represents an E_T bin) while in Figure 3.36 the ratio of the number of signal like events in the !VL LH background between the TMVA and Cut-Based method. We can conclude that by applying the TMVA method, we can benefit from the reduction of the signal like events in the background templates.

3.6.5 Fitting the Isolation Distributions

The isolation distributions are fitted using a linear combination between a Signal Function and a Background Template.

The Signal Region (SR) is chosen from (-0.2 , 0.4) while the Background Region (BR) is chosen from (0.4 , 1.4). The variation on the S or/and B regions are included in systematic errors, as discussed in section 3.7.

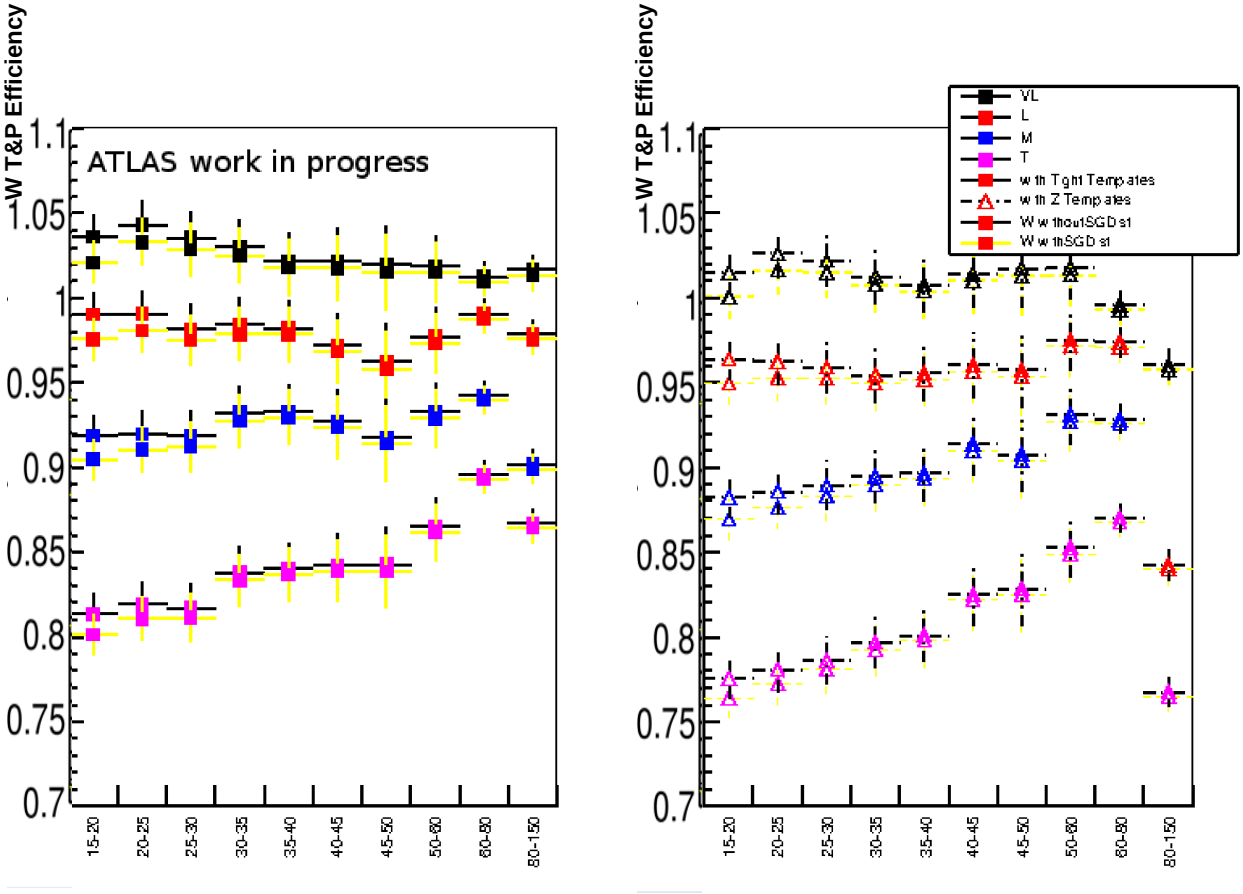


Fig. 3.33 The W T&P efficiency using the 2 signal templates (Tight Isolation on left and from ZIso Distributions right) for Template6 background template, in the $0 < |\eta| < 0.8$ region. All identification menus are presented: VL in Black, Loose in Red, Medium in Blue and Tight in Pink.

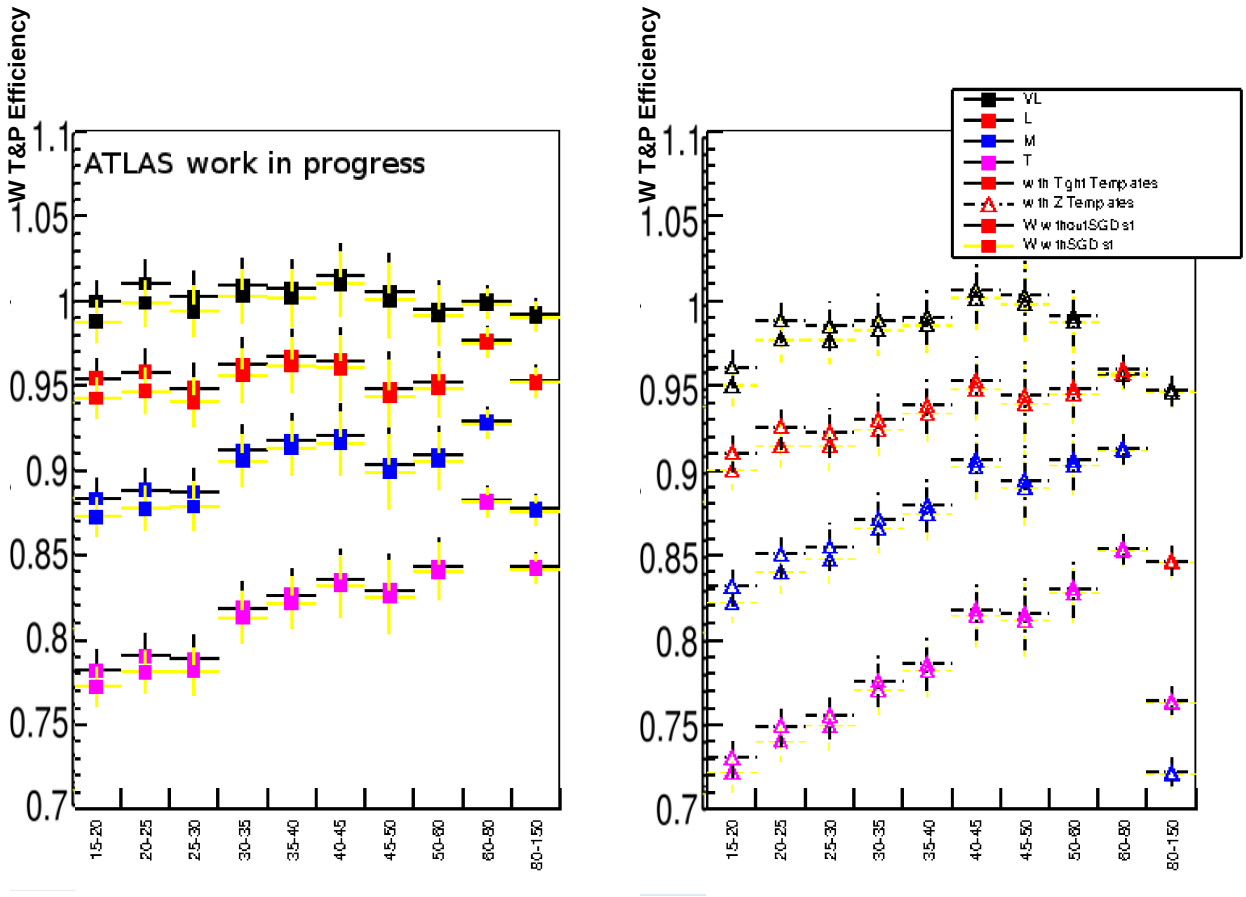


Fig. 3.34 The W T&P efficiency using the 2 signal templates (Tight Isolation on left and from ZIso Distributions right) for !VL LH background template, in the $0 < |\eta| < 0.8$ region. All identification menus are presented: VL in Black, Loose in Red, Medium in Blue and Tight in Pink.

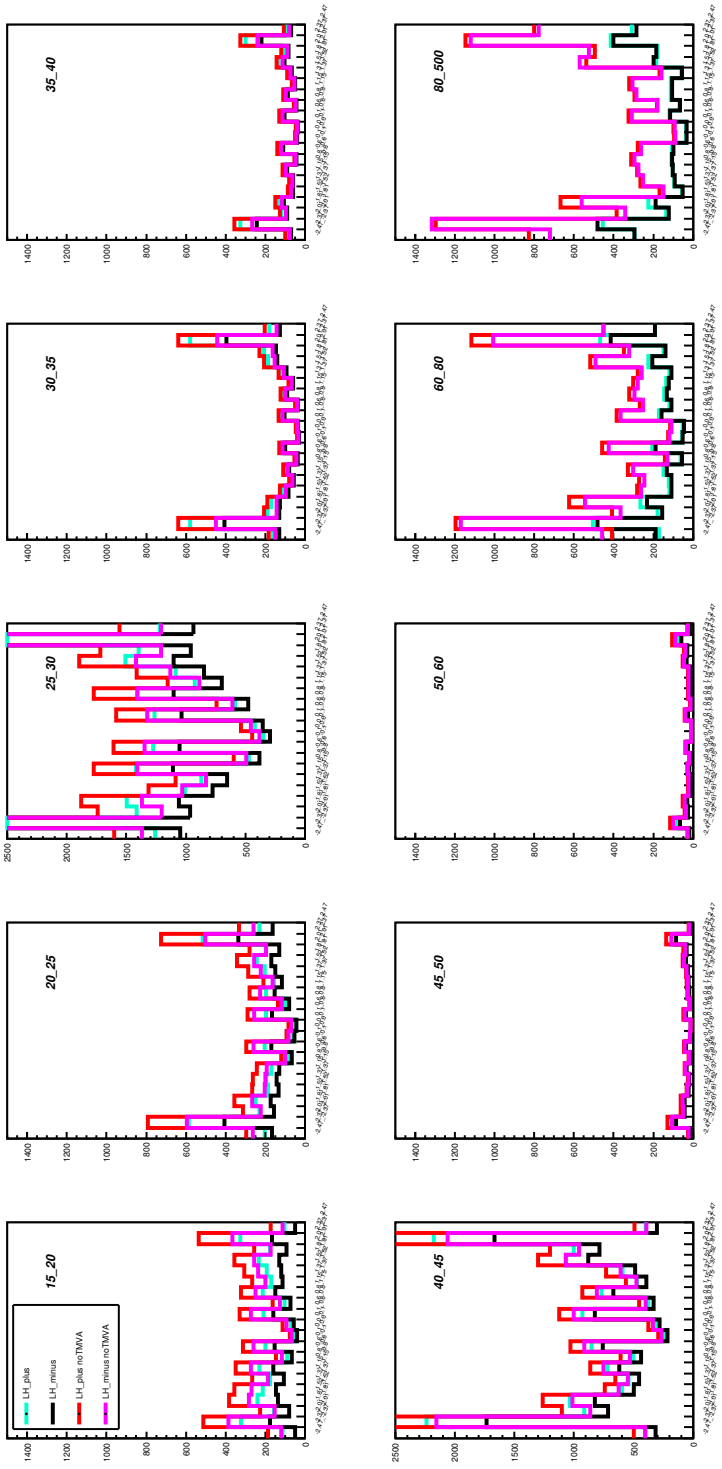


Fig. 3.35 Amount of the signal like events in the !VL LH background using the TMVA and Cut-Based method. Cyan color depicts !VL LH background for $W^+ \to e^+\nu$ analysis, Black depicts !VL LH background for $W^- \to e^-\nu$ analysis by using TMVA. Red (W+) and Pink (W-) are the colors used to depict the cut-based background rejection method for the same background template.

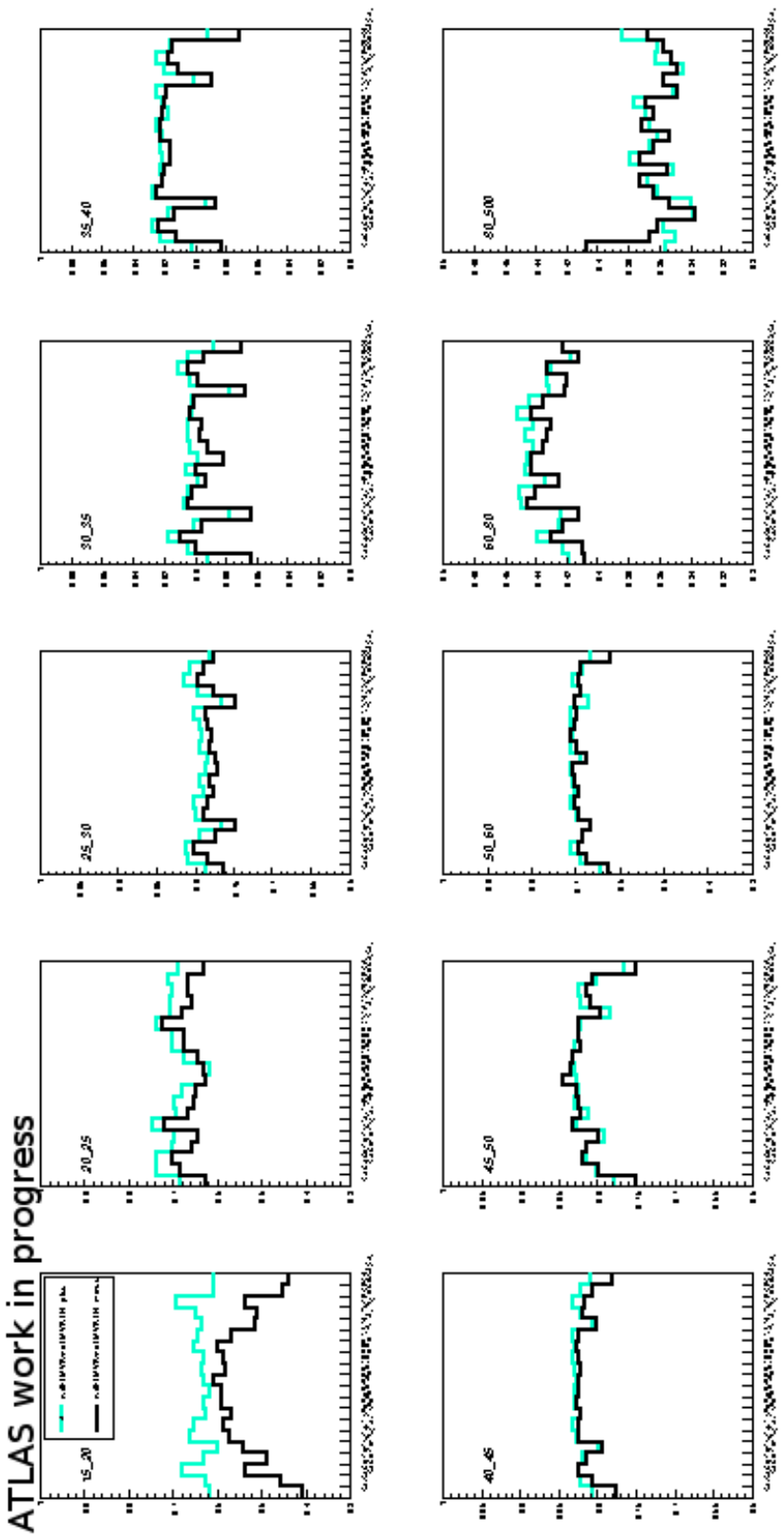


Fig. 3.36 Ratio of the signal like events in the !VL LH background between the TMVA and Cut-Based method. Cyan color is used for $W^+ \to e^+\nu$ analysis, Black for $W^- \to e^-\nu$ analysis.

In the SR, from (0.1, 0.4) one option is added namely usage of an extra Gaussian in the Signal Function for the fit function to account for the unphysical gap between the signal and background distributions.

Further, we can extract the signal from the signal fit integral in the signal region and estimate the systematic uncertainty of the fit as 1/2 of the Gaussian contribution in that range.

As an example, the efficiencies in the $0 < \eta < 0.8$ central region using Tight Templates, no Gaussian without/with the signal like events extraction from background templates in the signal region in left/right is shown in Figure 3.37. The three different background templates are considered.

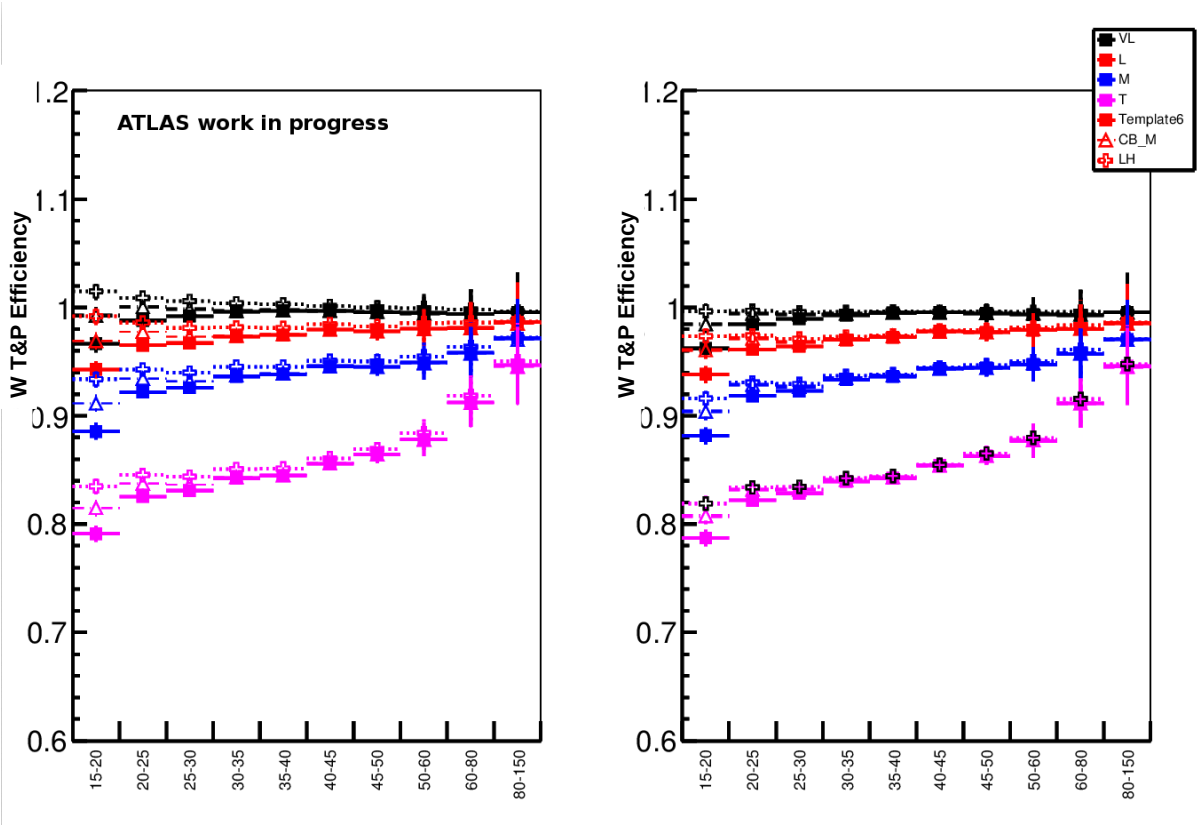


Fig. 3.37 Example of efficiency measurement using Z Isolation Distribution as Signal Template in η central region $0.0 < |\eta| < 0.80$

3.7 Systematics

Efficiency measurements using the isolation method were performed having as systematics:

- The three Background Templates from Cut-Based and Likelihood menus
- Variation on cone size around the selected electron for the isolation distribution
- The 2 Signal Templates constructed using Z Isolation distributions or Tight W distributions (both of them considered background free samples)
- Variations on Signal and Background regions
- extra Gaussian in the fit function to the signal with the average in the range 0.1 - 0.4 and estimate systematics as one half of the Gaussian contribution in that range
- extra Crystal Ball function in the fit function to the signal with the average in the range 0.1 - 0.4 and estimate systematics as one half of the Gaussian contribution in that range

To further evaluate the systematics, the analysis is divided into W+ and W- according to the probe electron charge.

Additionally, the following systematics are considered for the Scale Factors (SF):

- varying the kinematic variables: MET ($\pm$10 GeV) ; mT ($\pm$10 GeV); xs ($\pm$1); DeltaPhiJet15MET ($\pm$0.1) and DeltaPhiMETEle ($\pm$0.1) in MC before the TMVA reader
- varying the TMVA ($\pm$ 5% around the maximum significance value) in MC before the TMVA training

The systematic uncertainties are shown in Figure 3.38 for Loose probe electron selection, Figure 3.39 for Medium probe electron selection and Figure 3.40 for Tight probe electron selection for W^- (left), W^+ (middle), Ziso (right). In each Figure we can see the SF (top), Stat Err (2nd raw), Syst Err due to background variation (!VL and Template1) 3rd raw, Kinematic variation 4th raw for W, 5th Syst Err due to upper signal limit in the fit (from 0.4 to 0.3 and 0.5), 6th raw: systematic error due to fit model default + Crystal Ball, 7th raw: systematic error. We can observe the low impact on the SF from the statistical and systematics variations considered.

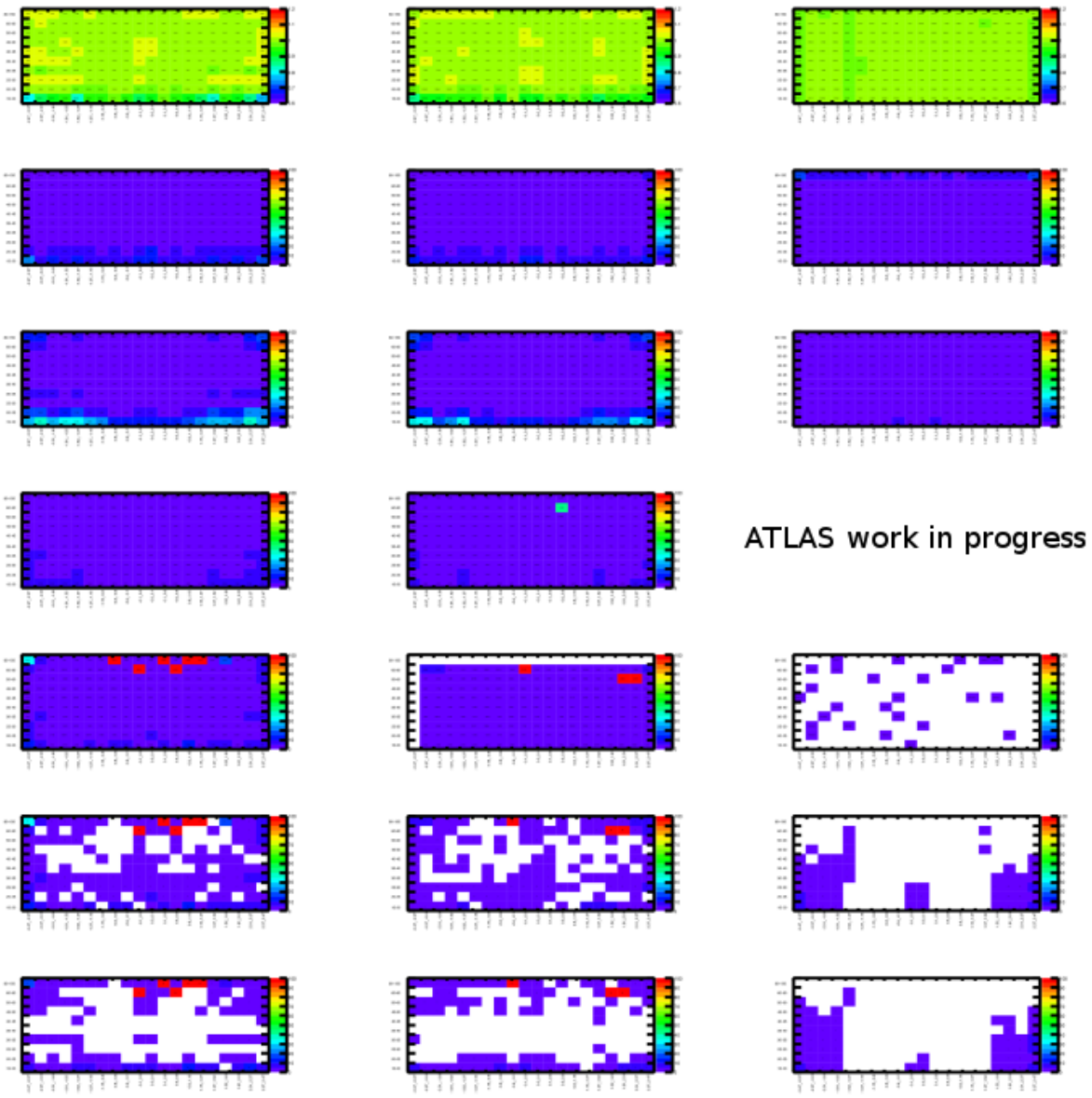


Fig. 3.38 The SF (top) and the systematic uncertainties for !VeryLoose LH background Templates and Loose probe electron selection for W^- (left), W^+ (middle), Ziso (right) and SF (top), Stat Err (2nd raw), Syst Err due to background variation (!VL and Template1) 3rd raw, Kinematic variation 4th raw for W, 5th Syst Err due to upper signal limit in the fit (from 0.4 to 0.3 and 0.5), 6th raw: systematic error due to fit model default + Crystal Ball, 7th raw: systematic error.

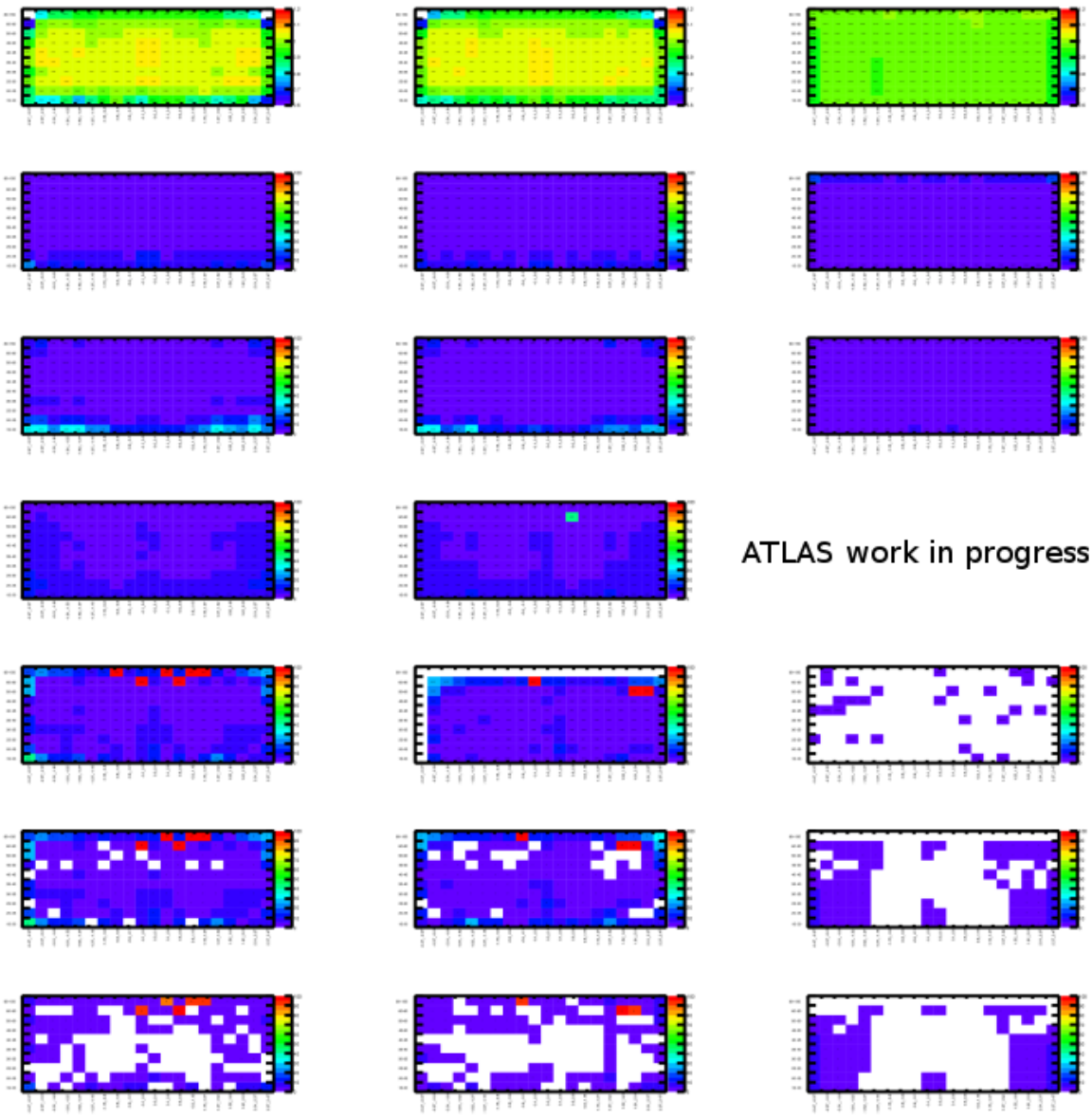


Fig. 3.39 The SF (top) and the systematic uncertainties for !VeryLoose LH background Templates and Medium probe electron selection for W^- (left), W^+ (middle), Ziso (right) and SF (top), Stat Err (2nd raw), Syst Err due to background variation (!VL and Template1) 3rd raw, Kinematic variation 4th raw for W, 5th Syst Err due to upper signal limit in the fit (from 0.4 to 0.3 and 0.5), 6th raw: systematic error due to fit model default + Crystal Ball, 7th raw: systematic error.

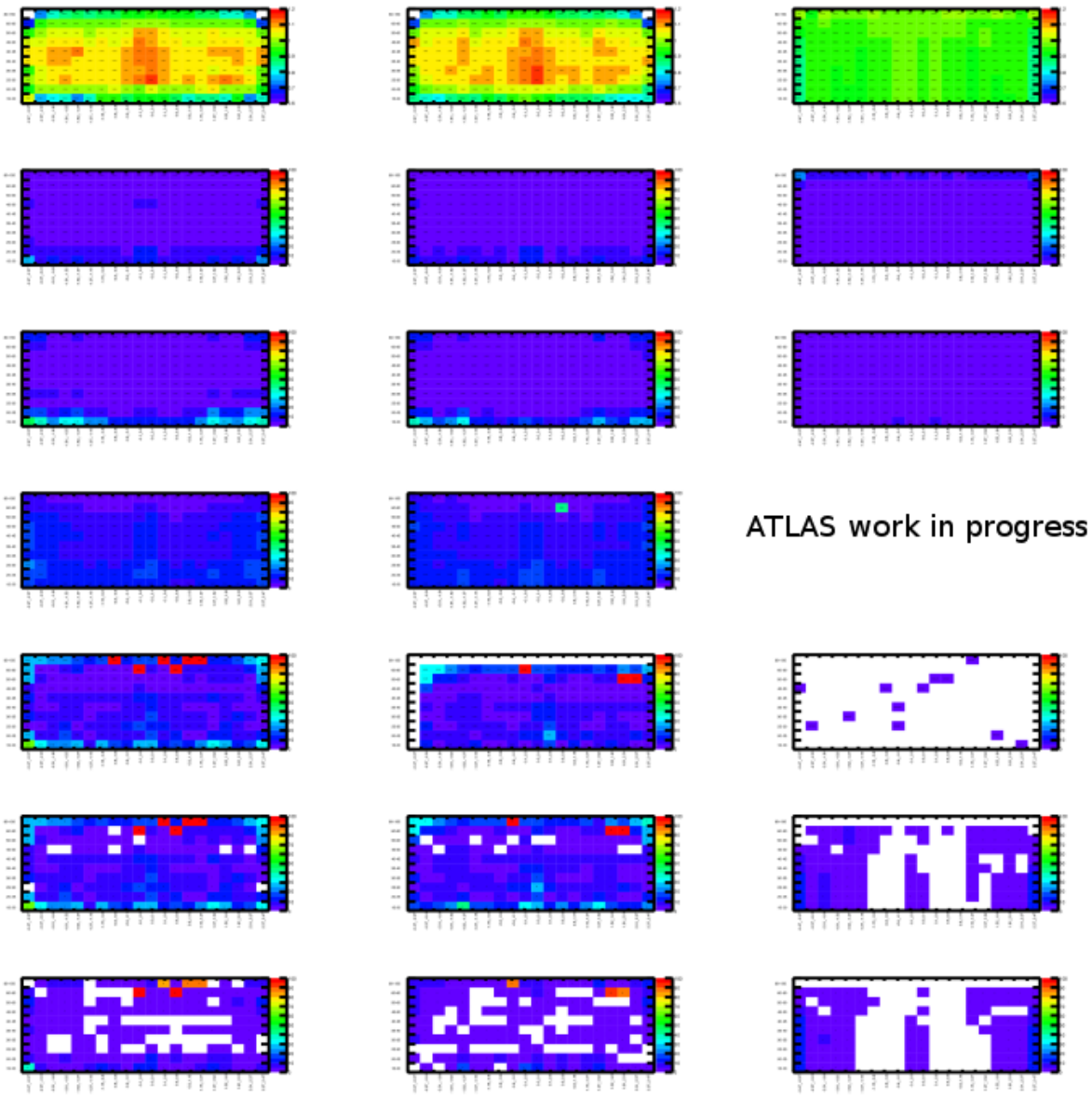


Fig. 3.40 The SF (top) and the systematic uncertainties for !VeryLoose LH background Templates and Tight probe electron selection for W^- (left), W^+ (middle), Ziso (right) and SF (top), Stat Err (2nd raw), Syst Err due to background variation (!VL and Template1) 3rd raw, Kinematic variation 4th raw for W, 5th Syst Err due to upper signal limit in the fit (from 0.4 to 0.3 and 0.5), 6th raw: systematic error due to fit model default + Crystal Ball, 7th raw: systematic error.

3.8 Results on the efficiency

In the following subsection, I will present the results on the efficiency of W T&P analysis for MC samples in 3.8.1 and data in 3.8.2. I will only present the results for the Medium menu for simplicity and to be situated in a region with middle selection on the signal and background. Same conclusions are valid for Loose and Tight selection menus.

The efficiencies are calculated using Tight distribution for the Signal Template and !VL LH background Template.

The SFs are presented for the Signal Template and T1, T6 and !VL LH background Templates. A cross-check using Z T&P is also shown.

3.8.1 Results on the MC efficiency

The Monte-Carlo W Tag and Probe uses a dedicated W T&P MC $W \rightarrow e\nu$ + 0-4 jets: mc15_13TeV.361523.MadGraphPythia8EvtGen_A14NNPDF23LO_Wenu_Npx, x = 0/1/2/3/4. This MC sample is the default one used in the analysis. It is generated using one generator named MadGraph, interfaced with Pythia8 and A14 tune for showering.

Another MC samples tested is mc16_13TeV.361103.PowhegPythia8EvtGen_Wminus (plus)_enu, generated using Powheg v2.0 and interfaced with Pythia8 for the parton showering and fragmentation. This sample is used as cross-check for validation of robustness of the efficiency measurement.

The Monte-Carlo W Tag and Probe efficiencies for !VL LH background template are given in Figure 3.41 for a background rejection based on a cut based procedure (in red), based on a TMVA (in green) and compared the Ziso method (in black). The 200 bins represents the binning in 10 p_T bins, for each p_T bin we have 20 η bins, as discussed in section 3.1.3. A good agreement is seen between the two procedure, and the use of the TMVA does not introduce any bias.

We can observe a good compatibility between the W tag and probe analysis and Z analysis efficiencies, within errors, making, therefore, a validation of the W Tag and Probe method.

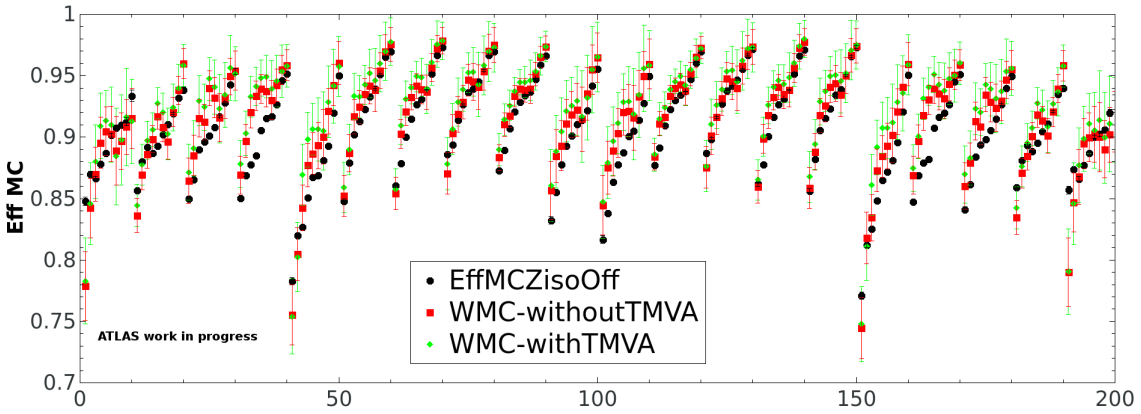


Fig. 3.41 The MC W T&P efficiencies using the Medium ID for Cut-Based menu (red) and with TMVA (green) as well as Ziso efficiency (black), using Tight distribution for the Signal Template and !VL LH background Template. The 200 bins represents the binning in 10 p_T bins, for each p_T bin we have 20 η bins.

3.8.2 Results on the Data efficiency

For the data, the W Tag and Probe efficiencies using Tight distribution for the Signal Template and !VL LH background Template, using Cut-Based procedure to remove the background is shown in Figure 3.42 for two dedicated W T&P MC configurations ($W \rightarrow e\nu$ + 0-4 jets, $W^{\pm} \rightarrow e^{\pm}\nu$). Each of them were used for the signal like events subtraction from the background templates, the negligible impact of the efficiency being a validation of the efficiency procedure calculation.

The Data W Tag and Probe efficiencies using Tight distribution for the Signal Template and !VL LH background Template and using TMVA procedure to remove the background is shown in Figure 3.43 for two MC samples mentioned above.

In both Figures, we can observe an agreement between the efficiencies using the two MC configurations. This indicates a robust efficiency measurement in W T&P data.

The Data W T&P Efficiencies using TMVA procedure to remove the background is shown in Figure 3.44 for two MC samples mentioned above for 2015+2016 data samples (red and green) and 2017 data samples (black). We can observe an agreement between the efficiencies using the two MC samples between 2015+2016 data and high-pileup 2017 data, which uses also the new configuration of triggers as will be discussed in 3.10. Therefore, no pile-up dependence can be attributed to data W T&P efficiencies. Also, the agreement

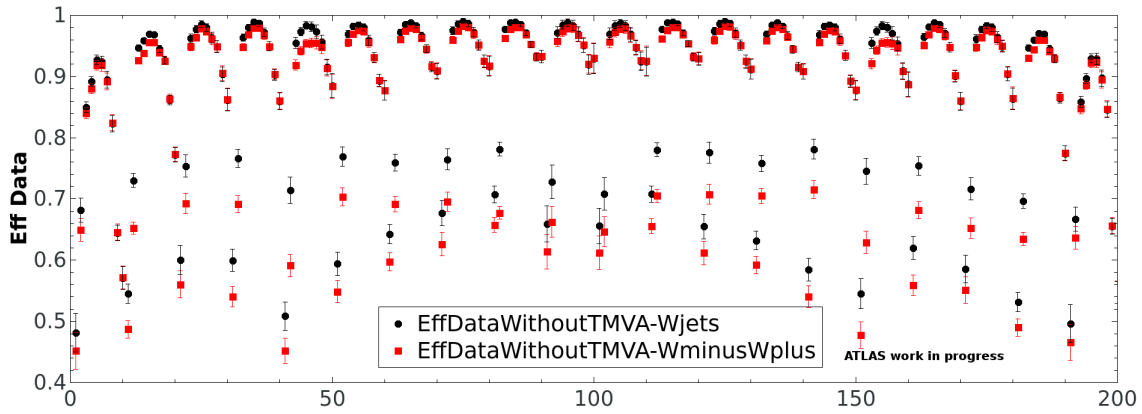


Fig. 3.42 The data efficiencies using the Medium ID and Cut-Based procedure to remove the background for two MC samples, using Tight distribution for the Signal Template and !VL LH background Template. The 200 bins represents the binning in 10 p_T bins, for each p_T bin we have 20 η bins. The Black points uses the MadGraphPythia8EvtGen Monte-Carlo sample to subtract the signal like events from the background samples while the Red points uses the PowhegPythia8EvtGen Monte-Carlo sample for extracting true signal like events from the background template.

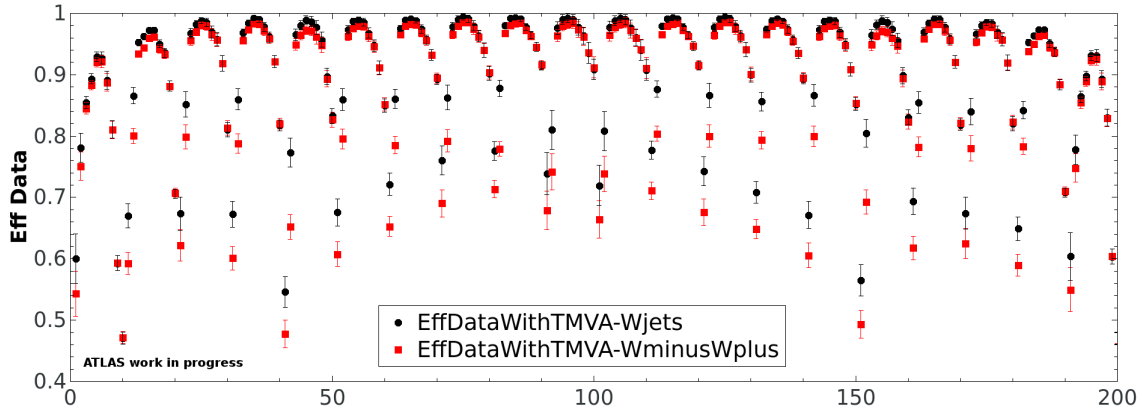


Fig. 3.43 The data efficiencies using the Medium ID and TMVA procedure to remove the background for two MC samples, using Tight distribution for the Signal Template and !VL LH background Template. The 200 bins represents the binning in 10 p_T bins, for each p_T bin we have 20 η bins. The Black points uses the MadGraphPythia8EvtGen Monte-Carlo sample to subtract the signal like events from the background samples while the Red points uses the PowhegPythia8EvtGen Monte-Carlo sample for extracting true signal like events from the background template.

between 2015+2016 and 2017 data sets is a strong argument for the continuity of the W T&P analysis using higher statistics in Run II.

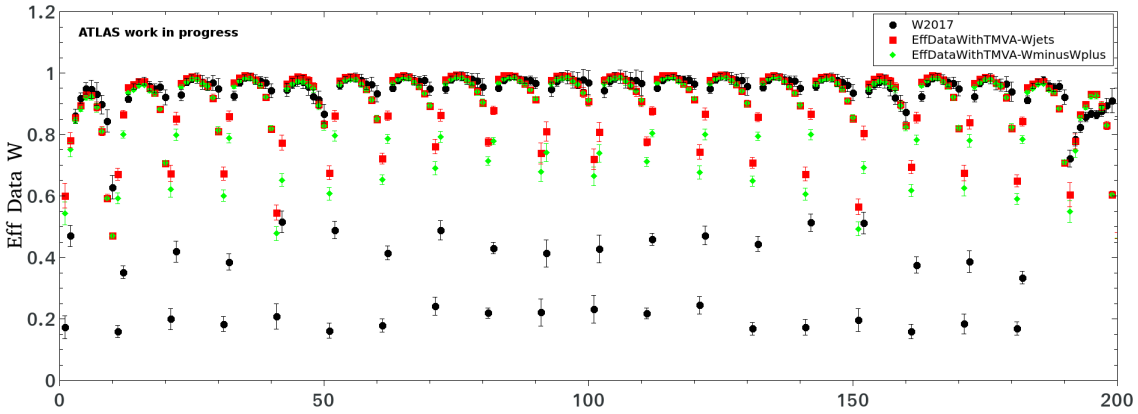


Fig. 3.44 The Data W T&P Efficiencies using the Medium ID and TMVA procedure to remove the background for two MC samples for 2015+2016 data and 2017 data samples, using Tight distribution for the Signal Template and !VL LH background Template. The 200 bins represents the binning in 10 p_T bins, for each p_T bin we have 20 η bins. The Black points represents the 2017 data set, the Red points represents the 2015+2016 data set that uses the MadGraphPythia8EvtGen Monte-Carlo sample to subtract the signal like events from the background samples while the Green points uses the PowhegPythia8EvtGen Monte-Carlo sample for extracting true signal like events from the background template.

3.9 Results on the Scale Factors (SF)

The final output of the performance studies is the SF that is used in all physics analysis on MC sample to provide a better description of data.

The SF using the W tag and probe method (labeled as W:) are compared with $Z \to e^+e^-$ tag and probe (labeled as Z:), for all background templates, in Figure 3.45, as a function of all 200 bins that represents the dependence on probe electron p_T and η. We can observe 10 % agreement between the W and Ziso. This is a strong argument towards the possibility for the W T&P analysis to be included in future performance combination along with $Z \to ee$ and the $J/\Psi \to ee$ analysis, once the systematic uncertainties are fully understood.

However, the disagreement between the templates is visible, especially between Template6/Template1 and !VL LH. Moreover, the discrepancy is visible between the templates

in the low p_T and low/high η range (regions dominated by background). Therefore further studies will be needed on the background treatment as well as on fitting procedure from which the electron identification efficiencies are calculated.

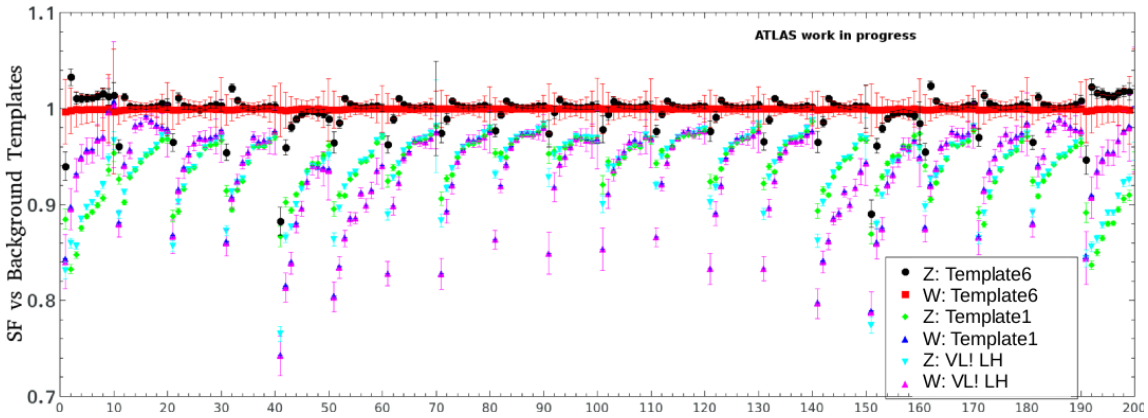


Fig. 3.45 SF using the Medium ID for W T&P and Ziso T&P, using Tight distribution for the Signal Template and T6 (Black for W T&P and Red for Z T&P), T1 (Green for W T&P and Blue for Z T&P) and !VL LH (Cyan for W T&P and Pink for Z T&P) background Templates. The 200 bins represents the binning in 10 p_T bins, for each p_T bin we have 20 η bins.

3.10 New triggers proposed for 2016-2017 data taking

On top of the background rejection procedure that was developed in order to purify the signal sample, the problem of low statistics is also solved by changing the trigger configuration, aiming for the following objectives:

- Complement Z analysis at low p_T
- Measure the 25/30-80 GeV range; with statistics competitive with Z
- Catch the highest E_T tail to improve precision from Z; Unprescaled trigger L1EM50/60 and MET>35 GeV can be used.

To test a new trigger by an emulation procedure [39], it is essential to consider three important factors: Efficiency, rate and CPU time. The trigger emulation aims to re-compute the decision of the TriggerDecisionTool (a tool of ATLAS collaboration) based on selections in plain C++. After this, the rates of this triggers are calculated, and a given prescale is given to each trigger in order to have the complete set of triggers within a given bandwidth

available for this analysis.

Prescales are variables set in order to reduce the trigger output rate for a given stage (trigger chain or item). The physics trigger prescale is defined as prescale of Level 1 item x prescale of Level 2 chain x prescale of Event Filter chain. In 2015-2016 data taking, the e13/e18 W Tag and Probe triggers mentioned in section 3.1.4 were heavily prescaled.

Moreover, an overwhelming background (especially at low p_T) was observed at HLT level for this triggers, so more selective triggers were needed aiming at reducing the background contamination while keeping a satisfactory signal rate.

Therefore, I was appointed to design a new set of triggers at low and high energy E_T threshold. This was performed in emulation studies, checking the impact of the W T&P analysis variables on the background rejection, trigger rate and prescale. The variables tested were :

- xs + xe
- $\Delta\Phi$ (eJet/eMET) $>$ 0.5/1.5
- RATIO2 requirements: ex: 90RO2: XE2 $\geq$ 9HT (see B for more details), jets with p_T $>$10 GeV at HLT
- m_T

The new set of triggers and their prescales are presented in Figure 3.47 and described more in Appendix B:

Looking at the trigger rate at globally level (defined as pass trigger/all events in the whole phase space) and in the W T&P analysis (defined as pass trigger/all events in the W analysis phase space) in Figure 3.47, we can observe trigger rates up to 70-90%. However, some triggers were proposed to be disabled due to the fact that they are background dominated and have high prescale (see Table 3.5), enabling at the same time the high E_T triggers so catch the highest E_T tail to improve precision from Z:

- HLT_e60_etcut_trkcut_L1EM24VHIM_xe60_mt35 at rate of 338 Hz
- HLT_e60_etcut_trkcut_j15_perf_xe60_6dphi05_mt35 at rate of 37 Hz
- HLT_e60_etcut_trkcut_L1EM24VHIM_j15_perf_xe60_6dphi15_mt35 at rate of 0.5 Hz

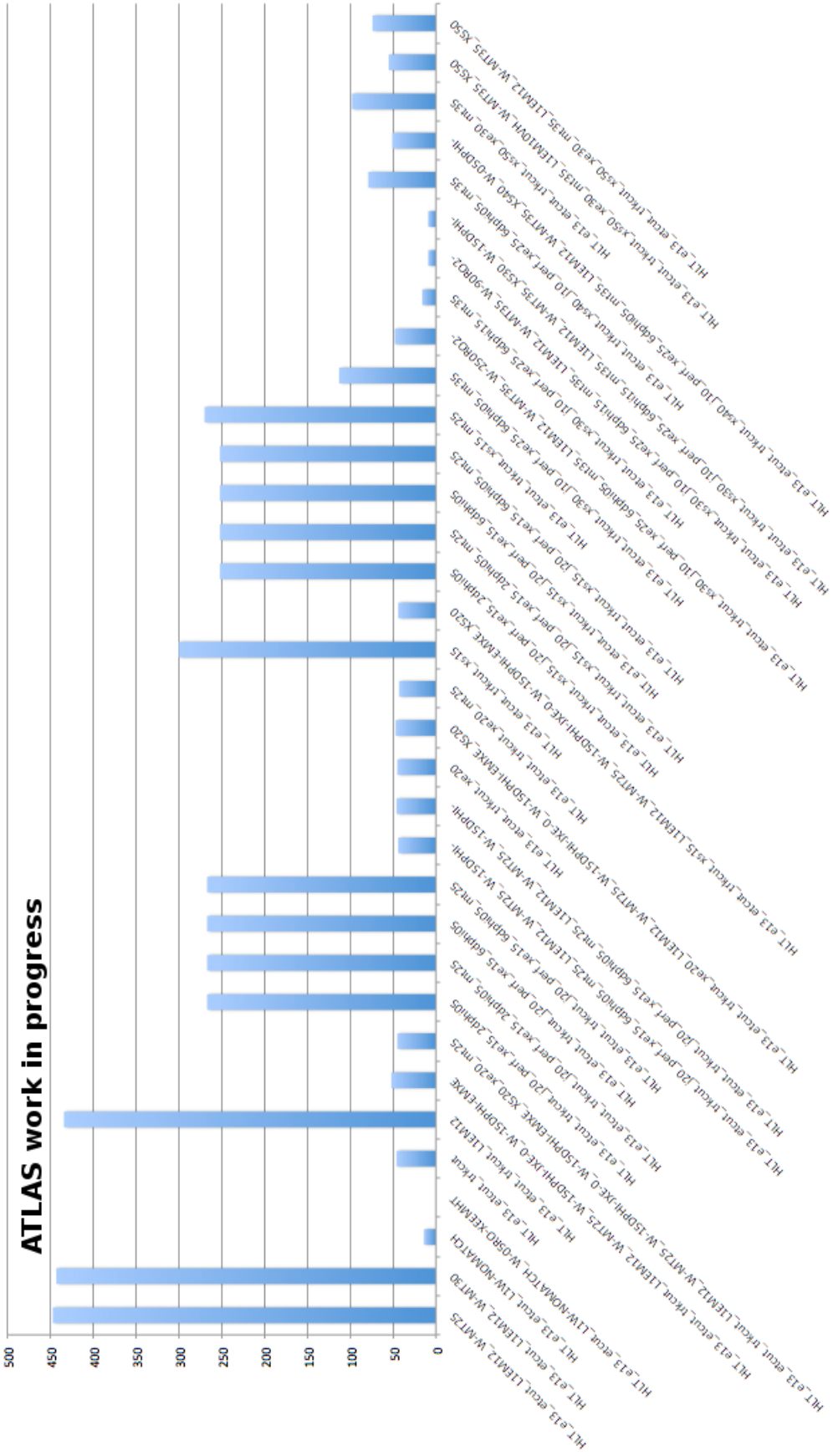


Fig. 3.46 New set of triggers prescales for 2016-2017 data taking.

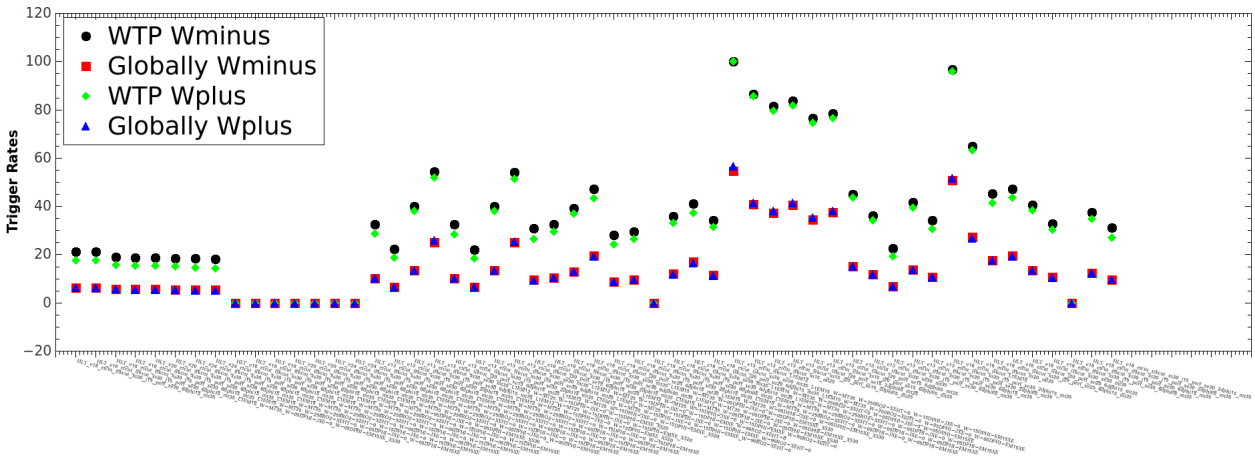


Fig. 3.47 Trigger rate (%) at Global level and W T&P analysis phase space.

- HLT_e60_etcut_trkcut_L1EM24VHIM_xs30_xe30_mt35 at rate of 722 Hz
- HLT_e60_etcut_trkcut_xs30_j15_perf_xe30_6dphi05_mt35 at rate of 91 Hz
- HLT_e60_etcut_trkcut_L1EM24VHIM_xs30_j15_perf_xe30_6dphi15_mt35 at rate of 1.3 Hz

Trigger	Prescale
HLT_e18_etcut_trkcut_L1EM15	384.6
HLT_e18_etcut_trkcut_xs20_mt35	182.4
HLT_e13_etcut_trkcut_j20_perf_xe15_2dphi05_mt25	219.6
HLT_e13_etcut_trkcut_j20_perf_xe15_6dphi05_mt25	128.3
HLT_e13_etcut_trkcut_L1EM12	384.6
HLT_e13_etcut_trkcut_xs15_j20_perf_xe15_2dphi05_mt25	154.1
HLT_e13_etcut_trkcut_xs15_j20_perf_xe15_6dphi05_mt25	91.8
HLT_e13_etcut_trkcut_xs15_mt25	206.6

Table 3.5 Table with variables used by Egamma Triggers in 2015 data

These new triggers provide a better background rejection compared to the 2015-2016 configuration, allowing us to have a cleaner signal sample at the HLT level. The new set of triggers were implemented for 2016-2017 data taking, and future studies will confirm a more efficient electron identification using W T&P analysis.

3.11 Conclusion

I presented in this chapter the electron identification efficiency measurement using $W \to e\nu$ decays. Numerous advantages come for using $W \to e\nu$ events to measure the electron identification efficiency: W and Z bosons cover both a large kinematic range, as can be seen in Figure 3.1, but $W \to e\nu$ events could help in understanding the uncertainty by offering different systematics from Z. Also, W selected samples provides additional statistics for the low E_T range (15-25 GeV) that can complement the J/Ψ channel and offers interesting experimental signatures due to its large statistics.

The main difficulty of this decay channel is represented by the background. In order to have a clean signal, optimization studies were performed using cut-based and TMVA methods as discussed in the sections. Further, for the 2017-2018 a new set of trigger configuration was proposed and validated, making the $W \to e\nu$ an interesting channel to provide cross-check on the scale factors obtained using $Z \to ee$ events.

The SF using the W tag and probe method and a comparison with Ziso tag and probe for all background templates were presented, with a 10 % agreement between the W and Ziso especially in bulk region (central η) where the background is smaller compared to low p_T region. This represents the first measurement on the Run 2 data for the electron efficiency using $W \to e\nu$ events.

My work could be valuable for the ATLAS collaboration due to the novelty of the measurement: the optimization and background studies, the first look on the efficiency and scale factors using $W \to e\nu$ events in Run 2 data, the trigger emulation studies. Future studies with a higher statistics could overcome the drawbacks seen on the background templates, the signal purity, and fitting procedure. The 2017-2018 new set of trigger configuration will allow future studies for the $W \to e\nu$ to be included in official recommendations along using Z and J/Ψ channels.

In the following chapters, I will present the work performed on the $t\bar{t}H$ channel within the Standard Model and a search for doubly charged Higgs boson in the beyond Standard Model framework, focusing on the $4l$ channel for which I was the main analyzer. Both analyses rely on a good lepton identification. Therefore the performance study presented in this chapter is a valuable ingredient to achieve accurate results.

Chapter 4

$t\bar{t}$H to 4l channel

4.1 Introduction to $t\bar{t}$H

All current measurements of the Higgs boson properties are consistent with the SM. The Higgs couplings to the fermions through Yukawa interactions are essential parameters in the SM that could be directly measured in the $t\bar{t}$H production. The top quark mass is a consequence of the electroweak symmetry breaking, the Yukawa coupling being proportional to fermion mass (see the Feynman diagram in Figure 4.1).

The Yukawa coupling can also be a test for the new physics or physics beyond SM (BSM) as new, colored particles could be present in the loop of ggH production process, dominated by top quarks. Thus, comparison of the gluon-fusion and the $t\bar{t}$H modes would allow for disentangling the effects of these possible new particless.

The top quark is heaviest fermion in the SM therefore it has the largest Yukawa coupling: $\lambda_t = \sqrt{2}m_t/v \approx 1$.

Due to the addition of new data by increase in the cross section of the $t\bar{t}$H production, Run II is expected to provide better measurements and $t\bar{t}$H production and the top Yukawa coupling, compared to the Run I results mentioned in section 1.6.4.

Also, the $t\bar{t}$H spin could also disentangle from a SM signal versus possible non-SM effects such as the SM scalar component from a pseudoscalar contribution in the top-Higgs coupling. Although in the SM top quark and antiquark pairs are mostly unpolarized in $t\bar{t}$H production at hadron colliders, their spins are strongly correlated as confirmed by experimental studies

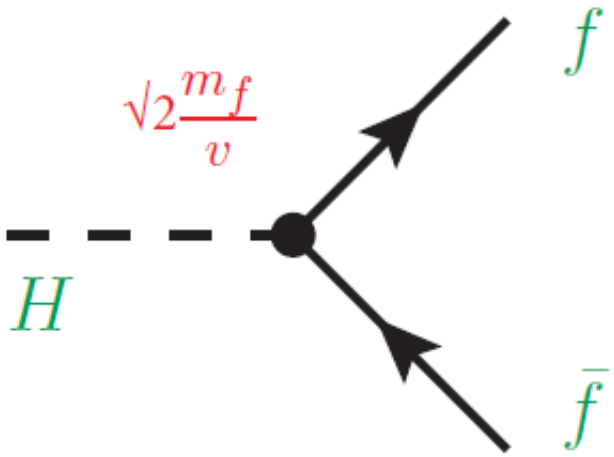


Fig. 4.1 The Feynman diagram for Higgs production in association with a pair of fermions.

at the Tevatron and the LHC Run I [65].

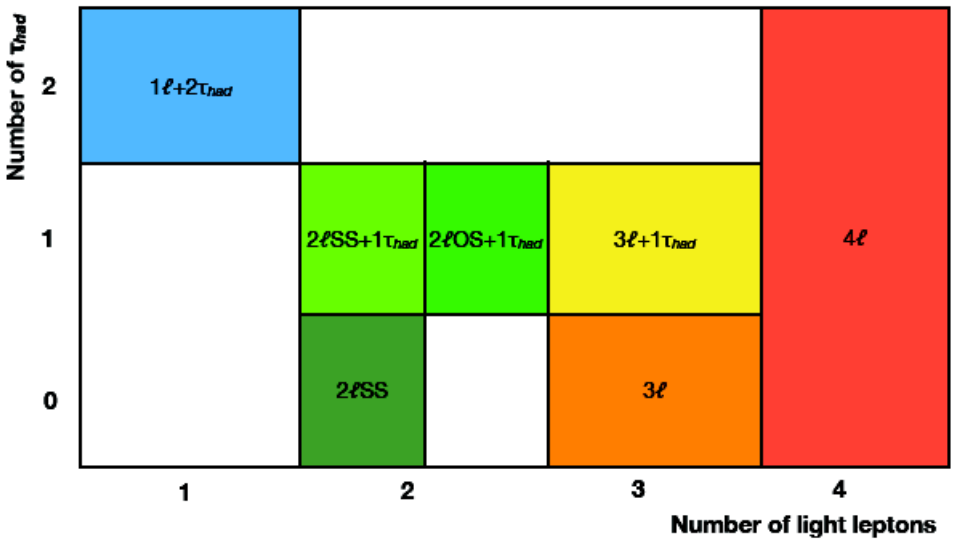


Fig. 4.2 The $t\bar{t}$H to multileptons final state channels.

In the next sections I will give a detailed description of the 4L channel for which I was the main analyzer. This channel is included in the $t\bar{t}H \rightarrow$ multileptons analysis that uses light leptons (l) and hadronic taus (τ_{had}) to define several orthogonal Signal Regions (SR) (see Figure 4.2):

- 2lSS and 2lSS+1τ_{had}

- $3l$ and $2l$OS+$1\tau_{had}$
- $4l$ and $3l$+$1\tau_{had}$
- $1l$ and $2\tau_{had}$

Here SS denotes same sign lepton pair and OS denotes opposite sign lepton pairs.

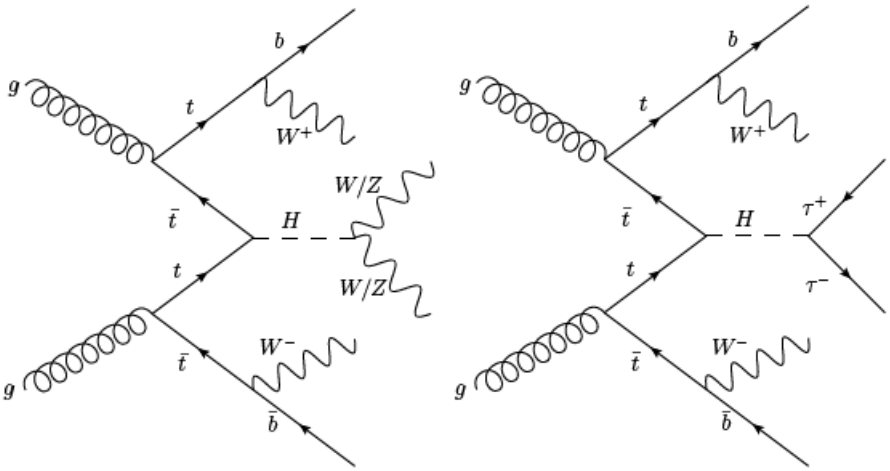


Fig. 4.3 The Feynman diagram for the final state of $t\bar{t}H \rightarrow$ multileptons analysis that targets $H \rightarrow WW; \tau\tau; ZZ$ decay modes.

The $t\bar{t}H \rightarrow$ multileptons analysis targets $H \rightarrow WW; \tau\tau; ZZ$ decay modes, combined with leptonic $t\bar{t}$ decays (see the Feynman diagram for the final state on Figure 4.3).

The $t\bar{t}H$ background composition and S/B are shown in Figure 4.4. Different background contributions are taken into account:

- Fake/non-prompt light and τ_{had} leptons
- Irreducible backgrounds: $t\bar{t}$ W, $t\bar{t}$ Z, other rare SM processes

The signal to background ratio ranges from few % to >40% in some channels ($4l$, $3l$+$1\tau_{had}$).

The $t\bar{t}H \rightarrow$ multileptons [68] analysis is then combined with other $t\bar{t}H$ final states to provide a final signal strength result:

- Multilepton

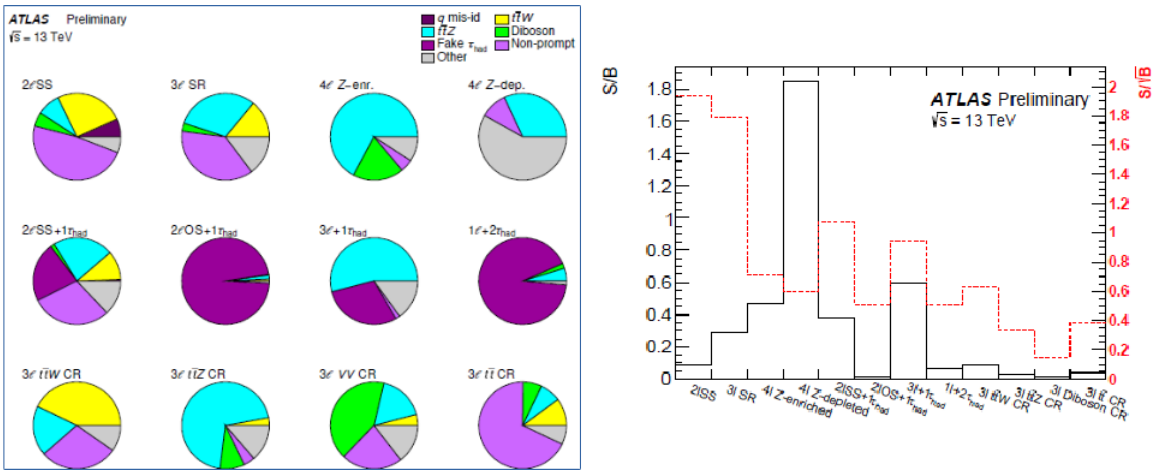


Fig. 4.4 The Background composition (left) and S/B ratio on right plot for all multilepton channels.

- $H \to b\bar{b}$
- $H \to \gamma\gamma$
- $H \to 4l$

The combined result gave the best-fit value of the signal strength of $\mu_{t\bar{t}H} = 1.17 \pm 0.19(stat)^{+0.27}_{-0.23}(syst)$, with a corresponding value of cross-section of $\sigma_{t\bar{t}H} = 590^{+160}_{150} fb$ ($\sigma_{t\bar{t}H \to multileptons} = 507^{+35}_{50}$ fb) and significance of 4.2σ at an expected value of 3.8σ.

4.1.1 Object reconstruction and selection

All analysis channels share a common jet, lepton, and overall event preselection. Events with significant noise in the calorimeters or data corruption are removed.

Events are required to pass a single lepton triggers (single electron, single muon) or dilepton triggers (ee, $e\mu$, $\mu\mu$) defined in Figure 4.5.

Light leptons

The performance study presented in chapter 3 enters in the object definition and selection used for the $t\bar{t}H$ analysis.

	Single lepton triggers (2015)
μ	HLT_mu20_iloose_L1MU15, HLT_mu50
e	HLT_e24_lhmedium_L1EM20VH, HLT_e60_lhmedium, HLT_e120_lhloose
	Dilepton triggers (2015)
$\mu\mu$ (asymm.)	HLT_mu18_mu8noL1
ee (symm.)	HLT_2e12_lhloose_L12EM10VH
$e\mu, \mu e$ (∼symm.)	HLT_e17_lhloose_mu14

	Single lepton triggers (2016)
μ	HLT_mu26_ivarmedium, HLT_mu50
e	HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0, HLT_e140_lhloose_nod0
	Dilepton triggers (2016)
$\mu\mu$ (asymm.)	HLT_mu22_mu8noL1
ee (symm.)	HLT_2e17_lhvloose_nod0
$e\mu, \mu e$ (∼symm.)	HLT_e17_lhloose_nod0_mu14

Fig. 4.5 List of triggers used for the whole 2015-2016 data taking.

Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeter that are associated with charged particle tracks reconstructed in the inner detector. Only candidates with $p_T > 10$ GeV and $|\eta| < 2.47$ are kept, while removing the crack region ($1.37 < |\eta| < 1.52$). The crack region is the region where the barrel accordion meets the endcap.

A multivariate likelihood discriminant combining shower shape and track information is used to distinguish real electrons from hadronic showers (fake electrons), as described in chapter 3.

For the object preselection, a loose electron discriminant working point is used, and no isolation is required.

Muon candidates are reconstructed by combining inner detector tracks with track segments or full tracks in the muon spectrometer. Candidates are required to satisfy $p_T > 10$ GeV and $|\eta| < 2.5$.

	e					μ			
	L	L†	L*	T	T*	L	L†	L*	T
Isolation	No	Yes				No	Yes		
Non-prompt lepton BDT	No	Yes				No	Yes		
Identification	Loose	Tight				Loose			
Charge mis-assignment veto	No	Yes				N/A			
Tranverse impact parameter significance $\lvert d_0\rvert/\sigma_{d_0}$	< 5					< 3			
Longitudinal impact parameter $\lvert z_0 \sin\theta\rvert$	< 0.5 mm								

Fig. 4.6 Loose (L), loose and isolated (L†), less loose (L*), tight (T) and very tight (T*) light lepton definitions. Selections for the tighter leptons are applied in addition to the looser ones.

The loose and tight WP on electrons and muons used in the $t\bar{t}H$ as well as in $H^{\pm\pm}$ analysis are shown in Figure 4.6.

Jets

Jets are reconstructed from topological clusters built from energy deposits in the calorimeters, using the anti-kt algorithm with a radius parameter R = 0.4. Only jets satisfying p_T >25 GeV and $|\eta|$<2.5 are used in this analysis. For jets with p_T <60 GeV and $|\eta|$<2.4, a jet-track association algorithm is used to confirm that the jet originates from the selected primary vertex, giving an average efficiency of this association of 92% per jet.

B-tagged jets

Jets containing b-hadrons are identified (b-tagged) via a multivariate discriminant combining information from the impact parameters of displaced tracks with topological properties of secondary and tertiary decay vertices reconstructed within the jet. The working point used for this search corresponds to an average efficiency of 70% for b-jets with $p_T > 20$ GeV and $|\eta| < 2.5$.

Missing transverse energy

Missing transverse momentum (E_T^{miss}) is defined as minus the vector sum of the transverse momenta of all reconstructed and calibrated physics objects and remaining unclustered tracks not associated to any object. The value of E_T^{miss} for an event is the magnitude of this vector.

4.1.2 Overlap removal

The reconstruction algorithms for electrons, muons, and jets in ATLAS are all independent.

In order to remove this duplication when performing analysis, a procedure called "overlap removal" is applied to the objects defined previously. For example for a genuine electron signal, one ends up typically with two reconstructed objects: an electron and a jet. Double-counting is not wanted, so the typical overlap removal procedure is used to chose which object should be kept while the other is discarded.

The overlap removal follows the steps:

- Remove any jet with Δ R < 0.2 with any electron.
- Remove any electron with ΔR < 0.4 with any jet.
- Remove any muon with ΔR < 0.4 with any jet.
- Remove any electron with ΔR < 0.1 with any muon.

4.1.3 Data and Monte Carlo samples

This analysis uses 36.1 fb^{-1} of data collected from proton-proton collision recorded by the ATLAS detector at $\sqrt{s}$ = 13 TeV during 2015 and 2016. The data set has been collected with a bunch crossing of 25 ns.

The analysis uses data being prepared with xAOD format and further produced to DxAOD format using HIGG8D1 derivation framework that applies a skimming procedure to reduce the data. The skimming implies requires at least two light leptons passing loose identification criteria with leading lepton p_T > 15 GeV and subleading lepton p_T > 5GeV, within $|\eta| < 2.6$.

Monte Carlo samples for each event generator programs and configurations used for simulating the signal and background processes are summarized in Table 4.7. In all cases, Higgs mass at 125 GeV is assumed. All Monte Carlo samples are processed through a complete simulation of the ATLAS detector response based on Geant4. All simulated events were processed using the same reconstruction algorithms and analysis chain as the data. The MC samples are corrected so that the object reconstruction and identification efficiencies,

Process	Generator (alternative)	Parton Shower (alternative)	PDF	Tune
$t\bar{t}H$	MG5_aMC	Pythia 8	NNPDF 3.0 NLO [10]/ NNPDF 2.3 LO [11]	A14
	(-)	(Herwig++)		
$tHqb$	MG5_aMC	Pythia 8	CT10 [12]	A14
tHW	MG5_aMC	Herwig++	CT10 /CTEQ6L1 [14, 15]	UE-EE-5 [13]
$t\bar{t}W$	MG5_aMC	Pythia 8	NNPDF 3.0 NLO /2.3 LO	A14
	(Sherpa)	(Sherpa)		
$t\bar{t}(Z/\gamma^*)$	MG5_aMC	Pythia 8	NNPDF 3.0 NLO /2.3 LO	A14
	(Sherpa)	(Sherpa)		
$t(Z/\gamma^*)$	MG5_aMC	Pythia 6	CTEQ6L1	Perugia2012 [16]
$tW(Z/\gamma^*)$	MG5_aMC	Pythia 8	NNPDF 2.3 LO	A14
$t\bar{t}t, t\bar{t}t\bar{t}$	MG5_aMC	Pythia 8	NNPDF 2.3 LO	A14
$t\bar{t}W^+W^-$	MG5_aMC	Pythia 8	NNPDF 2.3 LO	A14
$t\bar{t}$	Powheg-Box [17]	Pythia 8	CT10/CTEQ6L1	Perugia2012
$t\bar{t}\gamma$	MG5_aMC	Pythia 8	NNPDF 2.3 LO	A14
s-, t-channel, Wt single top	Powheg-Box [18, 19]	Pythia 6	CT10 /CTEQ6L1	Perugia2012
$VV, qqVV, VVV$	Sherpa 2.1.1 [20]	Sherpa	CT10	Sherpa default
$Z \to \ell^+\ell^-$	Sherpa 2.2	Sherpa	NNPDF 3.0 NLO	Sherpa default

Fig. 4.7 The configurations used for event generation of signal and background processes.

energy scales and energy resolutions match those determined from data.

Several MC versions were tested, corresponding to different versions of the ntuple production which are linked to different selection procedure and luminosity:

- v20/04: boosted Leptons true and not using ghost-matched info in μ-jet OLR + loosening the impact parameter cuts, L = 13.2 fb^{-1}
- v20/02: boostedLeptons true, criteria to reject leptons in el-jet and mu-jet OLR: dR $<$ $\min(0.4, 0.04 + 10GeV/pT)$ instead of dR$<$0.4, L = 13.2 fb^{-1}
- v21, same as v20/02 but with L = 22.1 fb^{-1};
- v24 at L = 36.47 fb^{-1}
- v26 at L = 36.47 fb^{-1} with a new standard Run1-OLR, no isolation required on muons. This version is taken as default (when a specific MC version is not mentioned).

In the following sections, I will present the full analysis performed on the 4l channel namely: the optimization studies, the definition of Signal Regions and Control Regions, the fake estimation and fit procedure. I will conclude with the overall multilepton result obtained by combining all the channels mentioned in 4.1.

4.2 The 4ℓ channel baseline selection

The $t\bar{t}$H production is captured only with a very small rate (0.2 %) into the four-lepton channel (see the Feynman diagram in Figure 4.9), which benefits from mild background conditions. The branching ratio of different final states used for the direct measurements of the top Yukawa coupling in the $t\bar{t}$H production is shown in Figure 4.8.

As can be seen in Figure 4.10 depicting the 4l background composition and the signal percentage. The prompt lepton backgrounds are dominated by the ZZ production followed by the double leptonic decays, while the non-prompt background is expected to originate from $t\bar{t}$ production, where at least two fake leptons are associated to the leptons originating from the two W bosons (considering 100 % BR of top to W and b).

The signal region construction starts from the simple conditions related to the leptons and jets (as described in section 4.1.1), applying on top in order to reject the background, a

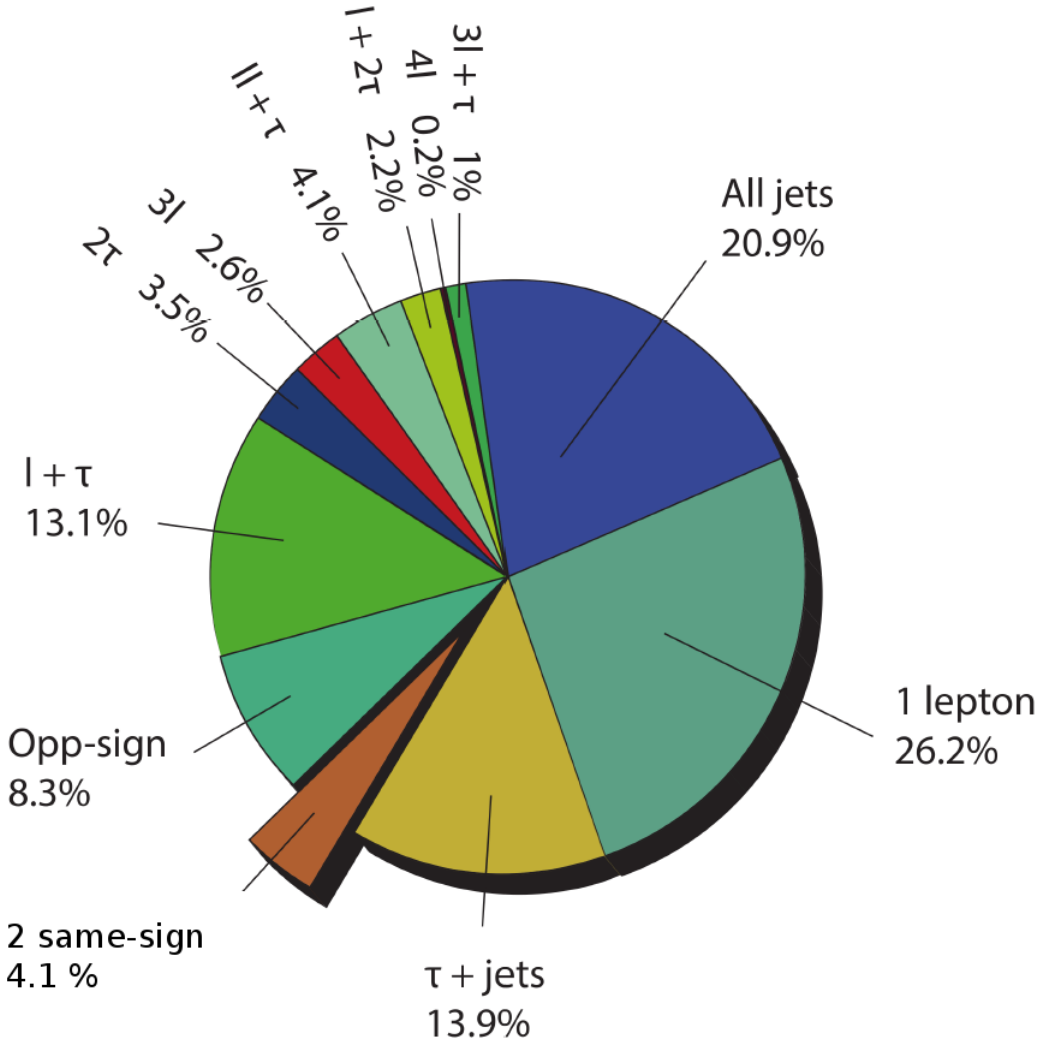


Fig. 4.8 The branching ratio of different final states used for the direct measurements of the top Yukawa coupling in the $t\bar{t}$H production.

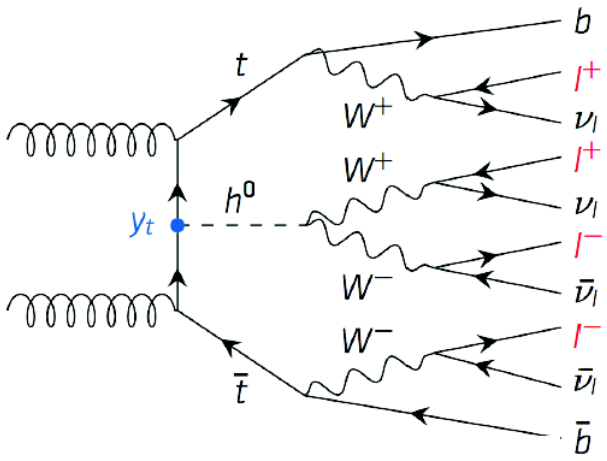


Fig. 4.9 The topology of the 4*l* final states.

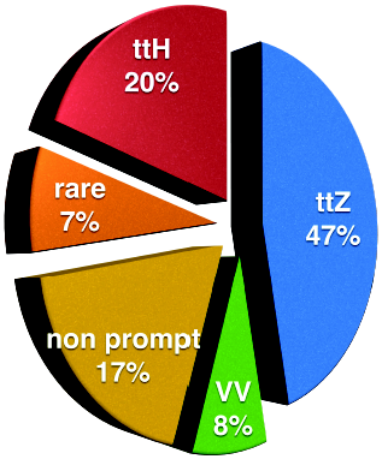


Fig. 4.10 The background composition in the 4*l* final states.

Z-boson veto (against ZZ) and the lepton isolation (against $t\bar{t}$).

The baseline event selection corresponding to the signal region (SR) is the following:

- exactly 4 light leptons (sorted by decreasing p_T value), with $p_T > 10$ GeV
- total charge to be 0
- the 3rd and 4th lepton should pass the Prompt Lepton Isolation (PLI) (see Appendix C) cut (lep promptLeptonIso TagWeight <- 0.5, a value obtained as a result of various optimization studies described in [68]), no isolation requirement on the first two leptons. The isolation criteria was applied only to the 3rd and 4th lepton due to the high probability of this leptons to be categorized as fake compared to the two highest p_T leptons (lepton 1 and lepton 2).
- at least one of 4 leptons to be trigger matched (the list of triggers are shown in 4.5) with p_T great than 25 (21) GeV for electron (muon)
- Z veto: the invariant mass of any opposite sign (OS) same flavor (SF) lepton pairs is required to be outside the Z mass window [81, 101] GeV
- Higgs veto: the invariant mass of any opposite sign (OS) same flavor (SF) lepton pairs are required to be outside the H mass window [120,130] GeV. This cut is applied to remove contribution from dilep and Z events.
- di-lepton resonances veto: inv. mass of any two OS-SF lepton pairs great than 12 GeV to further remove contribution from di-leptonic and Z events
- at least 2 jets
- at least 1 jet to be b-tagged. The cuts on jet and b-jet multiplicity allows further suppression on ttbar, ttV and VV background.

The goal of applying this selection cuts is to achieve the maximum purity of the $t\bar{t}$H signal while rejecting the MC background (described in section 4.1.3). The 4l background includes VV, ttbar, ttZ/W, Z/Wgam, DY (Drell-Yann) process, four top, single top. In addition, particular attention is given to the reduction of the $t\bar{t}$ contribution.

The following investigations have been pursued and detailed in further subsections:

- Different isolation algorithms have been tested (different isolation working points are described in Table 4.1)

- The isolation requirements have been studied for different event topologies, to preserve/recover the signal efficiency:
  - As a function of b-jet multiplicity: the fake leptons are mostly originating from b-jets. The $t\bar{t}$ background is therefore concentrated at low b-jet multiplicity $n_{b-jet} = 1$, where the isolation conditions are most effective.
  - As a function of lepton charge/flavor (SFOS: same flavor opposite sign pairs).

4.2.1 Isolation Selection

Isolation consists in assessing the activity surrounding the trajectory of the particle in the tracker and the calorimeters, being a powerful tool to discriminate signal against the background.

Two kinds of variables are defined in ATLAS. The track-based isolation estimates the surrounding activity by summing up the transverse momenta of the close-by tracks. The calorimeter-based isolation takes the energy of clusters of cells close to the particle in the calorimeters instead.

The calorimeter and tracking isolation variables E_T^{cone20} and $p_T^{varcone20}/p_T^{varcone30}$ are combined to separate isolated lepton candidates (prompt leptons) from non-isolated electron candidates such as leptons from heavy flavour hadron decays, electrons from photon conversion and light hadrons mis-identified as electrons.

WPs are then defined consisting of cuts applied to the two isolation variables. These cut values are fixed or vary to target specific signal efficiencies as a function of E_T.

The current isolation working points (WP) considered in the case of leptons final state are listed in Table 4.1. A new isolation WP was considered in the analysis namely the lepton MVA called Prompt Lepton Isolation (PLI) (see Appendix C) that uses lepton variables related to isolation, vertexing, nearest jet, and detector-specification as inputs to discriminate prompt from non-prompt leptons.

Two phase space were tested, providing a cross-check of the isolation application as well as to determine the region with highest signal significance:

- with a fixed Signal Region definition as presented previously

Working point	**Objects**	**Calo isolation**	**Track isolation**	**Combined calo+track isolation**
Tight	all leptons	96%	99%	95%
LooseTrackOnly	all leptons	-	99%	99%
Loose	all leptons	99%	99%	99%
Gradient	all leptons	$\varepsilon = (0.1143 * pT + 92.14)\%$	$\varepsilon = (0.1143 * pT + 92.14)\%$	$\varepsilon(25GeV) = 90\%, \varepsilon(60GeV) = 99\%$
GradientLoose	all leptons	$\varepsilon = (0.057 * pT + 95.57)\%$	$\varepsilon = (0.057 * pT + 95.57)\%$	$\varepsilon(25GeV) = 95\%, \varepsilon(60GeV) = 99\%$
FixedCutTight	electrons	E_T^{cone20}/E_T<0.06	$p_T^{varcone20}/p_T$<0.06	-
FixedCutTightTrackOnly	muons	-	$p_T^{varcone30}/p_T$<0.06	-
FixedCutTightTrackOnly	electrons	-	$p_T^{varcone20}/p_T$<0.06	-
FixedCutLoose	electrons	E_T^{cone20}/E_T<0.2	$p_T^{varcone20}/p_T$<0.15	-
FixedCutLoose	muons	E_T^{cone20}/E_T<0.3	$p_T^{varcone20}/p_T$<0.15	-

Table 4.1 The electron and muon isolation operating points definitions.

- within an optimized Signal Region by varying the cuts on the kinematic variables like:
  - Missing Transverse Energy (MET) $> 10, 20, 25, 30, 35$ GeV
  - $|M_{\ell\ell} - 91| > -1, 10, 13, 15$ (where -1 corresponds to no Z veto)
  - Inferior cut on Mass_4l_min $100, 120, 150$ GeV;
  - Superior cut on Mass_4l_max $500, 600, 10000$ GeV;
  - Cut on $P_T^{lep_0} = 25, 30, 35$ GeV;
  - Cut on $P_T^{lep_1} = 13, 15, 18, 20$ GeV ;

Different lepton isolation WPs have been explored, focusing on gradient and PLI ("Prompt Lepton MVA Isolation"), the WPs that provide the best signal significance. The basic principle of applying the isolation requirement is to suppress non-prompt as much as possible keeping high signal sensitivity.

In the case of the 4l channel, the choice to apply on the 3rd and 4th leptons the Prompt Lepton Isolation (PLI) (lep promptLeptonIso TagWeight <- 0.5) with no isolation require-

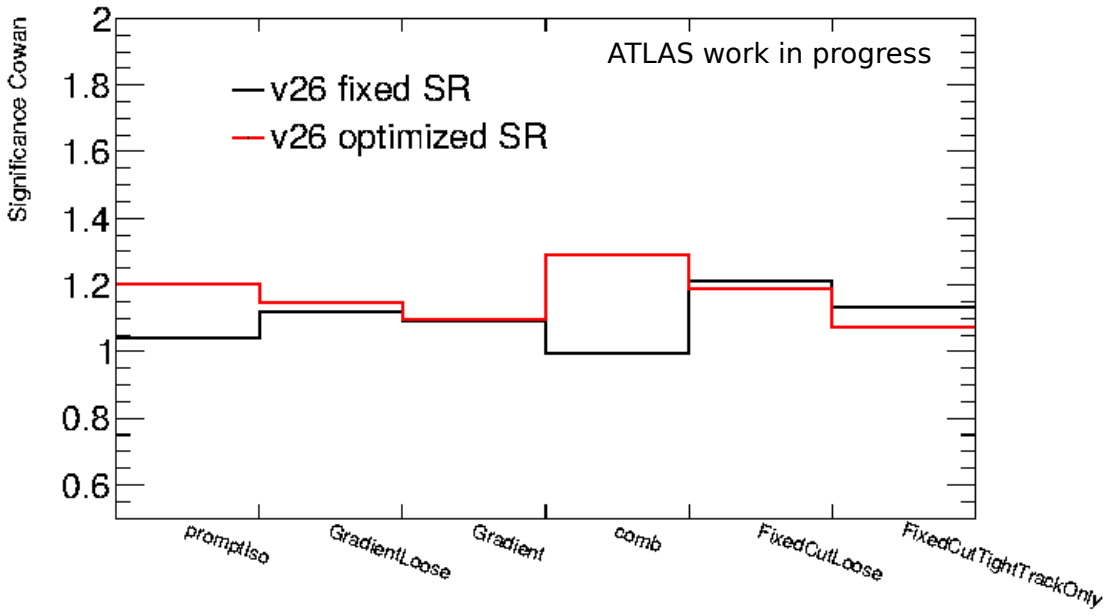


Fig. 4.11 Maximum significance of ttH signal calculated using Cowan definition, with respect to background samples obtained using different isolation WP. The comb in the Figure means the application of FixedCutLoose for electrons + FixedCutTightTrackOnly for muons.

Samples	**Baseline + Isolation Gradient on lepton 3 and 4**	**Baseline+PLI on lepton 3 and 4**
$t\bar{t}$H	1.59±0.22	1.37±0.19
$t\bar{t}$Z	2.97±0.49	2.68±0.44
$t\bar{t}$W	0.03±0.01	0.04±0.02
VV	0.23±0.12	0.20±0.11
other	0.45±0.23	0.38±0.20
$t\bar{t}$	1.22±0.61	1.05±0.32
Total	6.52±0.88	4.69±0.55
Signal Strength	$1^{+1.97}_{-1.47}$	$1^{+1.90}_{-1.39}$

Table 4.2 The yields in signal and background samples and the Signal Strength associated.

ment on the first two leptons was decided.

This choice was set by looking at the impact on signal significance (as shown in Figure 4.11) as well on the signal strength uncertainty with respect to the isolation WP while keeping the signal to background ratio as high as possible.

In Table 4.2 we can see the reduction of 28% in the total background by applying the PLI on the 3rd and 4th leptons with respect to the Isolation Gradient (the 4th WP in Table 4.1) on the same leptons (the second WP with highest signal significance). The reduction of the signal is 14%. The final impact on the signal strength uncertainty is of $\approx 5\%$.

Finally, the 4L SR is categorized as follows:

- baseline selection
- Split into Z-depleted (no same flavor opposite sign lepton pairs) and Z-enriched regions (at least one same flavor opposite sign lepton pair), in order to increase signal sensitivity (see Table 4.3). In both region, the Z veto requirement is applied.

For the Z enriched region, we apply on top a BDT cut to improve the result on signal strength as will be discussed in the following.

4.2.2 Optimization studies using TMVA

In High Energy Physics (HEP) we often need to discriminate from Signal like events and Background like events. This can be achieved by "Multivariate Analysis" (MVA) and in this case particularly the Gradient Boosted Decision Tree, a machine learning algorithm that makes uses of discriminating variables as input conditions, finding the one for which with a single cut gives the best improvement in signal purity.

For the 4L Signal Region, in order to make an optimal use of the existing MC statistics of the samples mentioned in section 4.1.3, the training is performed in the following regions:

- ttH: 4L + isoWP + Zveto + njet$\geq$2 + bjet$\geq$1
- ttV: 4L + isoWP + Zveto + njet$\geq$2 + bjet$\geq$1 + the invariant mass of the 4 leptons $>$100 GeV

The default BDTG setting is to choose randomly half of the events for training and the other half left for the testing. In the 4l channel, we use double-inverse training and testing to benefit from the entire statistics available in the sample (same procedure as described in 3).

The variables used for the TMVA training are selected in such a way to give the best signal to background discrimination, keeping in the analysis only the variables with the highest separation power.

The variables used are (from top left to right bottom) as seen in Figure 4.12:

- Missing Transverse Energy;
- The best other M_{ll} : constructed using the lepton pair with mass second closest to the Z mass;
- The best M_{ll} mass: constructed using the lepton pair with mass closest to the Z mass;
- The invariant mass of the 4 leptons;
- The $\sqrt{p_T}$ of the sub-leading in p_T jet;
- The leading (in p_T) jet;
- The ratio between ΔR and p_T of the pair lepton 1 and lepton 4;
- The ratio between ΔR and p_T of the pair lepton 1 and lepton 3;
- The transverse momentum scalar sum HT.

The variable distribution is shown in the Figure 4.12 for training against ttV background on odd events and Figure 4.13 for training against ttV background on even events respectively.

The correlation matrix for the variables used is shown in the Figure 4.14 for training against ttV background on even events (on the left for Signal and respectively on the right for Background events). We can observe no correlation between the variable chosen. Therefore, we can be sure that no bias is included by applying the TMVA procedure.

The output of the TMVA is shown in the Figure 4.15 for training against ttV background on odd events and Figure 4.16 for training against ttV background on even events respectively. On the left is shown the TMVA over-training checks while on the right the background rejection versus signal efficiency (ROC curve).

Region	**Signal Strength**
PLI34 (baseline)	$1^{+1.90}_{-1.39}$
PLI34 + Z-depleted and Z-enriched	$1^{+1.84}_{-1.16}$
Option BDTG on Z enriched (cut on TMVA_ttV>-0.2)	$1^{+1.652}_{-1.124}$

Table 4.3 The Fit on Signal Strength in optimized SR.

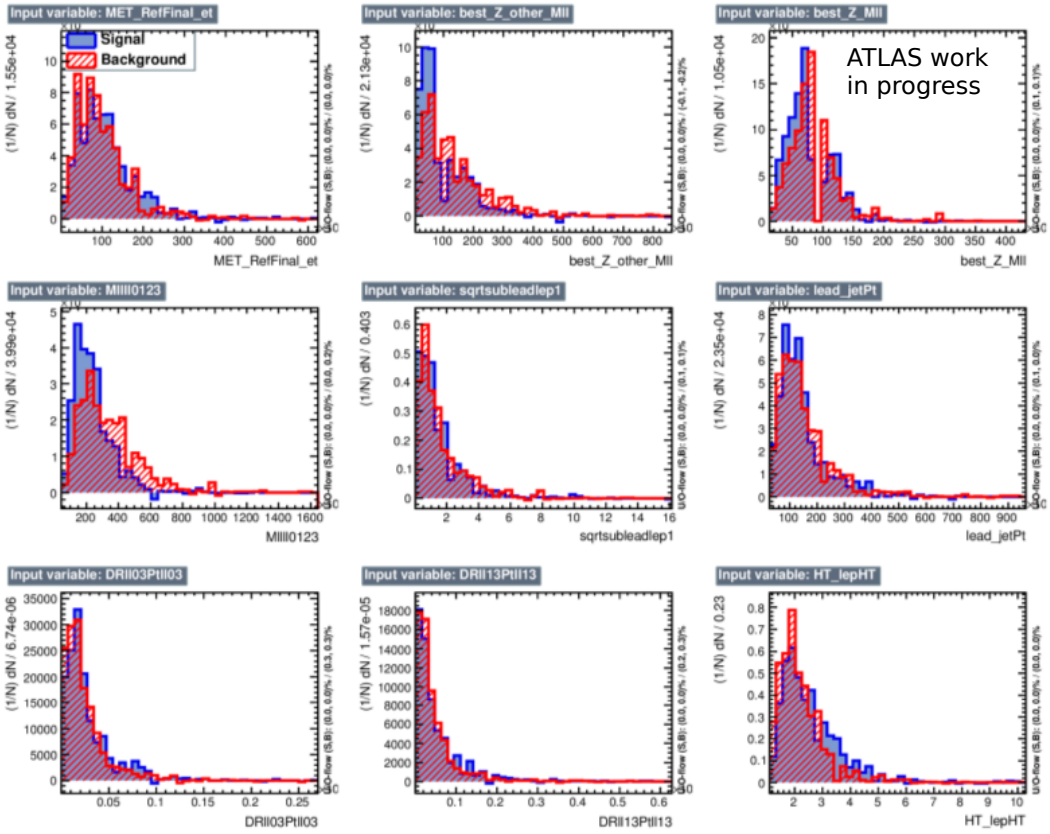


Fig. 4.12 Variables used for TMVA training ttH against ttV, training on odd events.

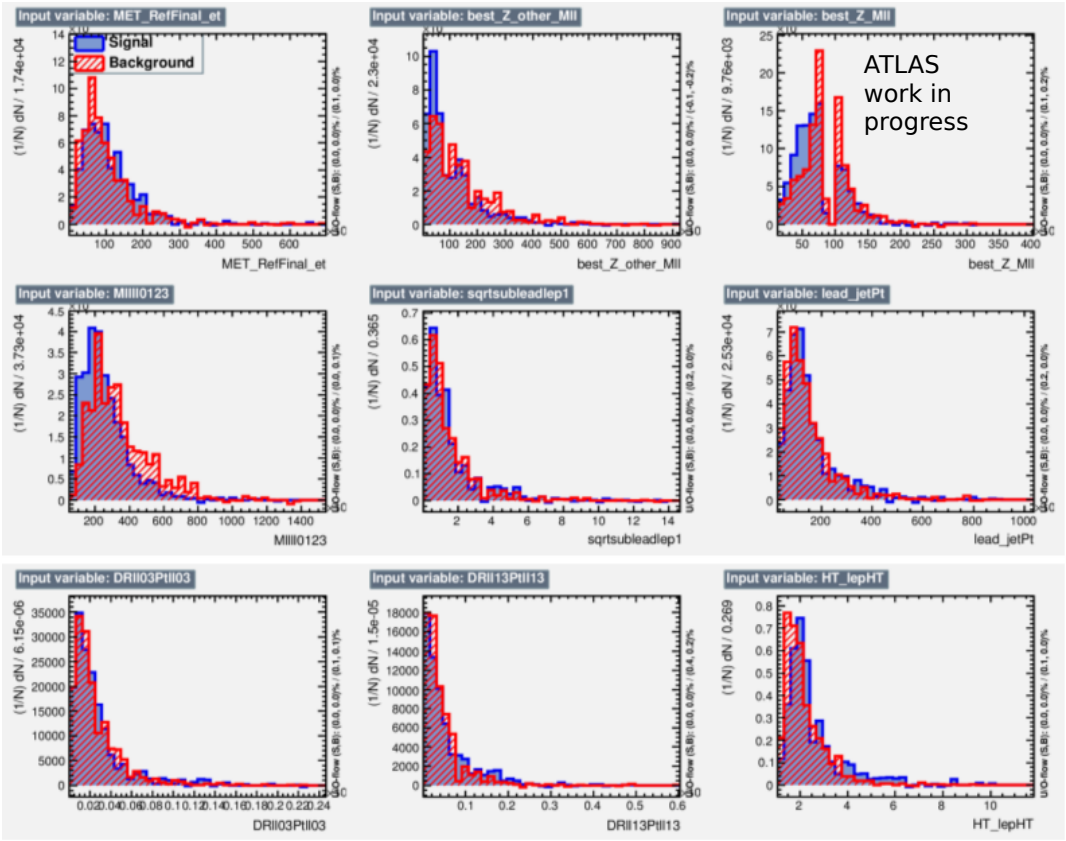


Fig. 4.13 Variables used for TMVA training ttH against ttV, training on even events.

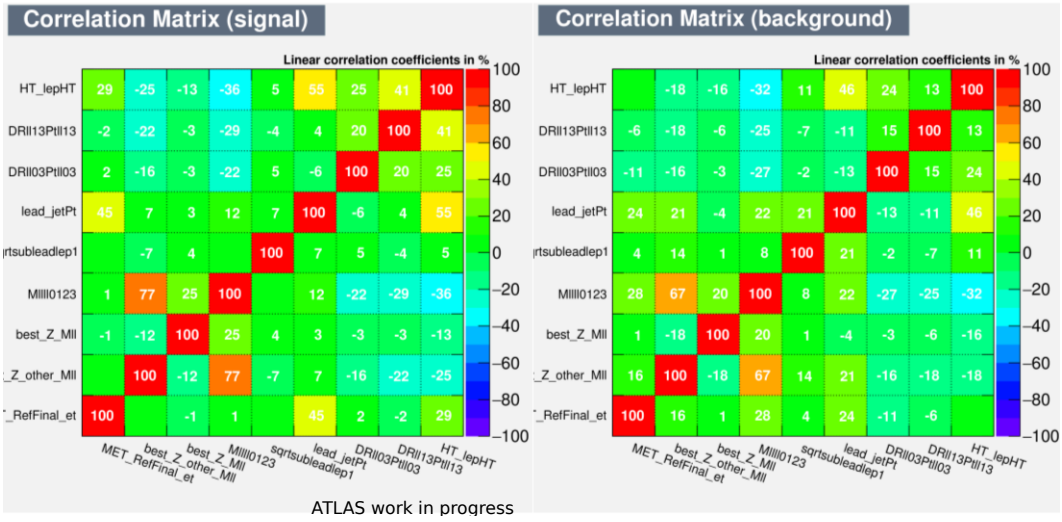


Fig. 4.14 Correlation matrix between the variables used for TMVA training ttH against ttV, training on even events.

Due to lack of statistics, the signal (blue) to background (red) separation is limited. However, looking at Table 4.3, applying the TMVA procedure gives a lower uncertainty on the signal strength compared to the Cut-Based analysis. The similar output between the even-odd training/testing assures a consistent TMVA procedure.

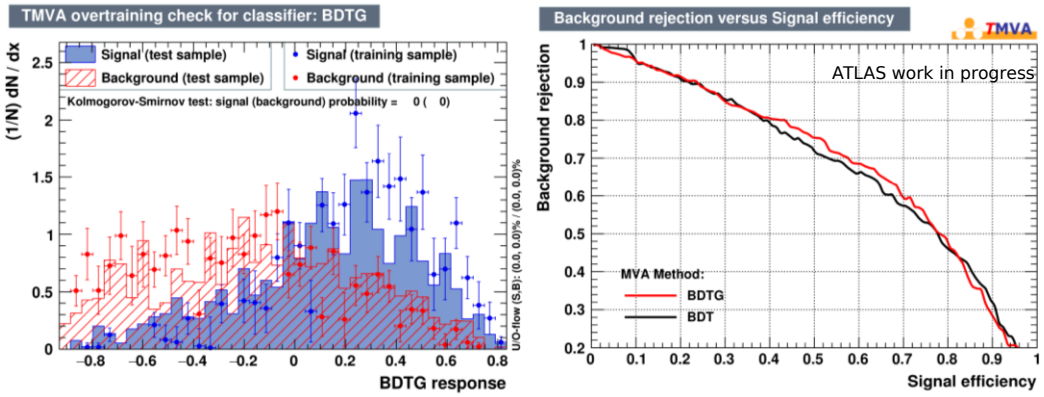


Fig. 4.15 TMVA BDTG output ttH against ttV, training on odd events.

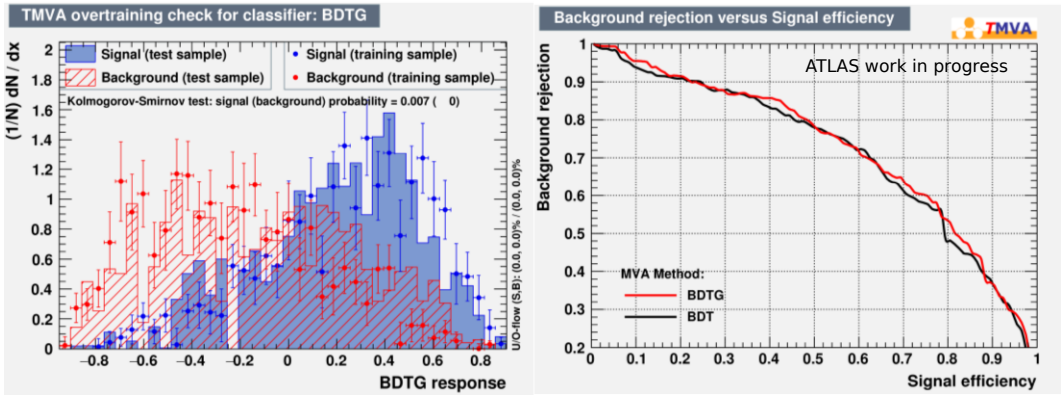


Fig. 4.16 TMVA BDTG output ttH against ttV, training on even events.

The optimal cut on the TMVA is set when achieving the maximum significance defined for simplicity as $\frac{S}{\sqrt{(S+B)}}$ when varying the TMVA cut point from -1 to 1 using 100 bins for a step of 0.02 (-1+i*0.02, i=0,100). The best cut was chosen for TMVA>-0.20 where the signal strength is $1^{1.652}_{-1.124}$.

The result on the fit on the signal strength by using the baseline selection, the split on the Z-depleted and Z-enriched region and finally the result after applying the BDT on the Z

enriched region is shown in Table 4.3.

As can be seen, the split on the Z-depleted and Z-enriched regions gives an improvement on the uncertainty of signal strength while adding the BDTG cut on Z enriched region gives the optimal option.

Further optimization studies are presented in Appendix B.4. They were done on the BDTG configuration in order to check the consistency of the method by varying the settings of the BDTG or the variables used in the training/testing. It is also presented a study of TMVA on other background samples as well as an extension of the 4L in the $H \rightarrow ZZ \rightarrow ll$ qq / bb channel. However, due to lack of statistics, the improvements of the signal strength by this additional studies is insignificant.

In conclusion, two optimization procedure were tested: cut-based by applying the baseline selection + Z-depleted and Z-enriched regions splitting and TMVA technique. Looking at the uncertainty of signal strength in Table 4.3, cut-based Z-depleted + BDTG cut on Z enriched region gives the optimal result. This option is considered as the 4*l* final SR region.

4.3 ttH Fake Estimates in 4l channel

There are several different sources of lepton misidentification depending on lepton type defined as “fake leptons”, or “fakes”.

For electrons, fakes can arise from charged hadrons, photon conversions, or semi-leptonic heavy-flavor decays. In the case of photon conversions and semi-leptonic heavy-flavor decays, an actual electron is present in the final state. These electrons are still considered fake in the sense that they are not produced in isolation as part of the prompt decay of a particle of interest. In the following, the term fake applies to both hadrons misidentified as leptons and to leptons from non-prompt sources. Prompt leptons produced in isolation are referred to as “real” or “true” leptons.

In the case of muons, almost all fakes come from either semi-leptonic heavy flavor decays or meson decays in flight. As above, these muons are referred to as fake even though an actual muon is present in the final state.

The fake factor method is a data-driven procedure for modeling background arising from misidentification. The method provides a measurement of the yield and the kinematic distributions of fake background, applicable to any physics analysis in which particle-level selection criteria are used to suppress background.

4.3.1 Fake Factor Method and CR definitions

The fundamental idea of the fake factor method is: select a control sample (also called control region CR) of events enriched in the background being estimated and then use an extrapolation factor to relate these events to the background in the signal region. The method is data-driven provided the control sample is selected in data, and the extrapolation factor is measured with data. The extrapolation factor is referred to as the "fake factor".

The control region is defined to select the background being estimated. The type of background considered with the fake factor method arises from particle misidentification. To collect this type background more efficiently, the particle selection in the signal region is replaced with a particle selection for which the misidentification rate is higher. This alternative particle selection criterion is referred to as the "denominator selection" or the "denominator definition"; particles passing this criterion are referred to as "denominator objects" or simply "denominators". The control region is then defined to be the same as the signal region, except a denominator object is required in place of the full particle selection in the signal region (SR).

Due to lack of data statistics in the 4l control regions CR, a semi data-driven technique is developed aiming to correct MC predictions (fake factors) in terms of contributions from different sources of non-prompt leptons. It uses the 3l CR to calculate the fake factors and then apply to 4l events as described in section 4.3.2. The procedure relies on the validated fact that the fake factors are compatible between the 3l and 4l regions.

The CRs are regions enriched in the Z+jets and $t\bar{t}$ background, which have a non-negligible contribution in 4l channel.

The Z+jets CR is defined by requiring the following selection cuts:

- exactly 3l loose leptons (with $p_T > 10$ GeV) abs(total charge) = 1;
- Pass the isolationGradient WP;

- one opposite sign same flavor (OSSF) pair within the Z-window ($|Mll - M_Z| \leq$10 GeV);
- exactly 1 or 2 jets with $p_T > 10$ GeV
- $E^T_{miss} < 50$ GeV and $m_T < 50$GeV (for the fake lepton)

In this case the fake lepton is the lepton which does not belong to the best Z candidate.

The $t\bar{t}$ CR is defined by requiring the following selection cuts:

- exactly 3l loose leptons ($p_T > 10$ GeV) + abs(total charge) = 1 +isolationGradient
- no OSSF lepton pair
- Exactly one or two jet (with one $p_T >$30 GeV)

In this case, the fake lepton is the Same Sign (SS) lepton which has the lowest p_T.

4.3.2 Fake Scale Factors

Starting from the CR defined above, one can calculate the fake scale factors by using the formula:

$$N^{f,CR}_{data} - N^{f,CR}_{others} = \lambda_{tt} \cdot N^{f,CR}_{tt} + \lambda_{Zjets} \cdot N^{f,CR}_{Zjets} \tag{4.1}$$

where CR is the considered Control Region, f is the flavor of the considered leptons, N_{data} is the number of selected data events, N_{others} are estimating the MC sample excepting tt and Zjets and $N_{tt/Zjets}$ are the yields in the CR enrich background samples tt/Zjets.

We can, therefore, calculate simultaneously from 4 independent equations the λ factors for each flavor (fake muon $\lambda^{\mu}_{b}, \lambda^{\mu}_{l}$ or fake electron $\lambda^{e}_{b}, \lambda^{e}_{l}$).

The resulting scale factors, obtained in the 3*l* CR defined previously, are listed in Table 4.4 where the error is the total error calculated as $\sqrt{stat^2 + sys^2}$. The systematic error was taken to be 50% of nominal value, following the conclusions from the section 4.3.3. The table shows the agreement between the scaling factors calculated using different MC versions (see section 4.1.3 for their definition).

The consistency between the fake scaling factors in different ntuple versions assures a validation of the method, which is robust across different selection procedures performed

at ntuple versions. It is also robust with the increase of luminosity available at each ntuple production time. These arguments support the fact that the systematic uncertainty taking into consideration for the fake estimation covers the biases that could arise from different ntuple production steps.

Type	v21	v20/04	v20/02	v24	v26
λ_b^e	0.92 ± 0.47	1.00 ± 0.51	0.76 ± 0.39	1.15 ± 0.58	1.07 ± 0.54
λ_l^e	1.18 ± 0.60	1.24 ± 0.63	1.11 ± 0.57	1.74 ± 0.89	1.47 ± 0.75
λ_b^μ	0.86 ± 0.45	1.10 ± 0.56	0.50 ± 0.27	1.07 ± 0.54	0.99 ± 0.51
λ_l^μ	0.71 ± 0.37	0.85 ± 0.44	0.56 ± 0.29	1.69 ± 0.88	1.13 ± 0.58

Table 4.4 Fake Scale Factors using the 2 CR defined in 4.3.1

4.3.3 Scale Factors Systematics

For the systematics of the scaling factors, two dependencies were tested: on the b-tagging multiplicity and the fake p_T threshold.

Scale factors dependency on b-jets multiplicity For the dependency on bjet algorithm, we have checked 4 disjoint regions:

1. Nominal = No bjet req
2. A = number of bjet=0 for Zjets and number of bjet=0 for $t\bar{t}$
3. B = number of bjet>0 for Zjets and number of bjet=0 for $t\bar{t}$
4. C = number of bjet=0 for Zjets and number of bjet>0 for $t\bar{t}$
5. D = number of bjet>0 for Zjets and number of bjet>0 for $t\bar{t}$

As can be seen in Figure 4.17 for v26 alone or in Figure 4.18 for all ntuple versions tested, a systematics of 50% (the blue interval) will cover the dependency of the scaling factor on the bjet multiplicity.

Scale factors dependency on fake p_T For the dependency on the fake p_T, we vary the fake momentum:

1. PT>10 GeV

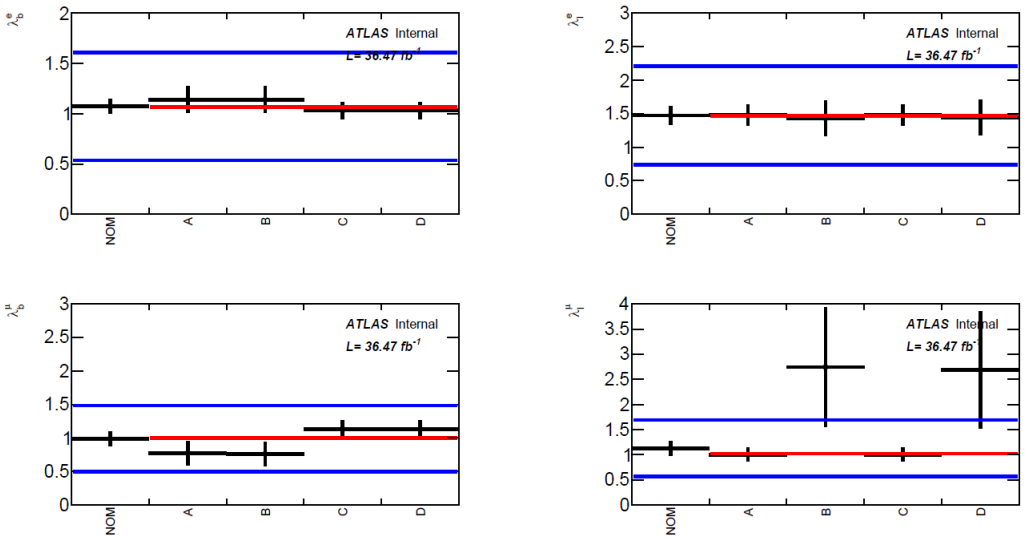


Fig. 4.17 The Scaling Factor dependency on bjet multiplicity

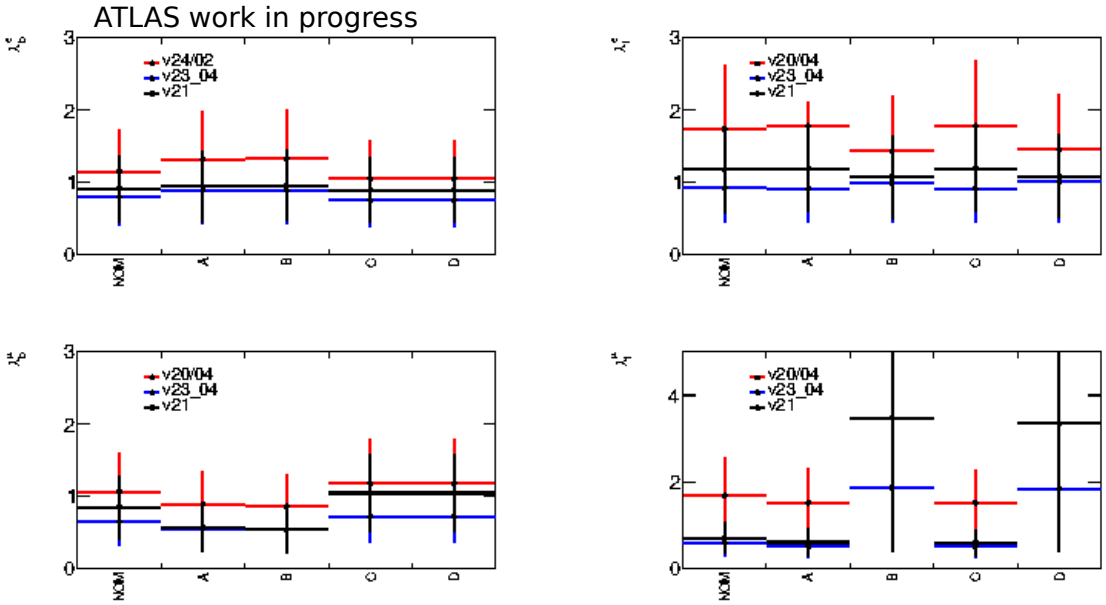


Fig. 4.18 The Scaling Factor dependency on bjet multiplicity for various ntuples versions

2. PT>12 GeV
3. PT>15 GeV
4. PT>18 GeV

The dependency on the fake PT for the two considered options can be seen in Figure 4.19 for v26 alone or in Figure 4.20 for all ntuple versions tested.

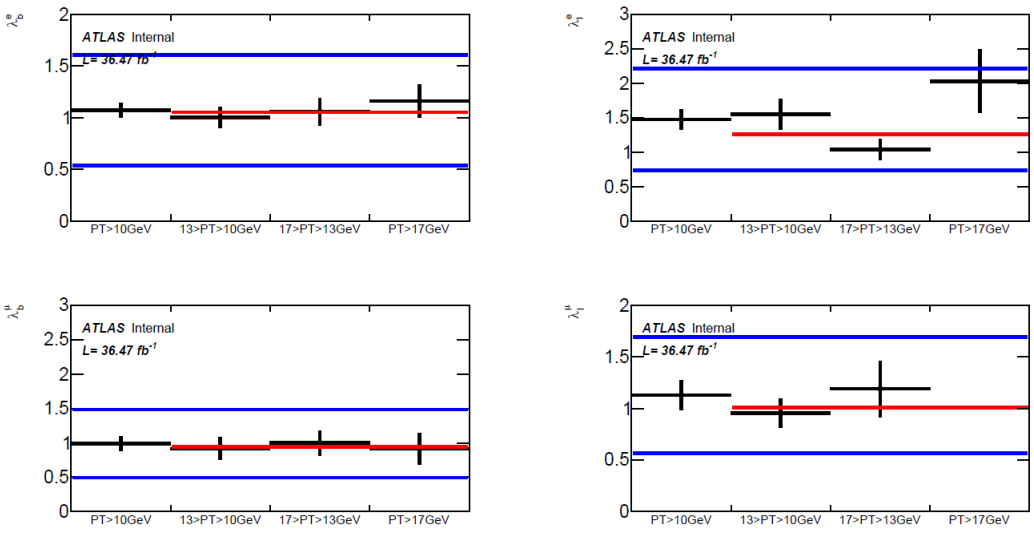


Fig. 4.19 The Scaling Factor dependency on fake lepton transverse momentum

In conclusion, the fake scale factors (FSF) are fairly stable within different ntuple versions tested. This assures a procedure of fake estimation without any bias from the ntuple production. An impact of up to 20% in the background can be observed when a conservative systematic error of 50% is applied to fakes, independent of the methods. This systematic error of 50% of nominal value covers all the dependencies seen of the FSFs with respect to b-jet multiplicity and fake lepton transverse momentum.

4.3.4 The scaling factor application

The scaling is applied as a function of the nature and origin of the leptons:

- A fake lepton is a reconstructed non-prompt lepton (thus not coming from Higgs boson, t, Z, W or τ)

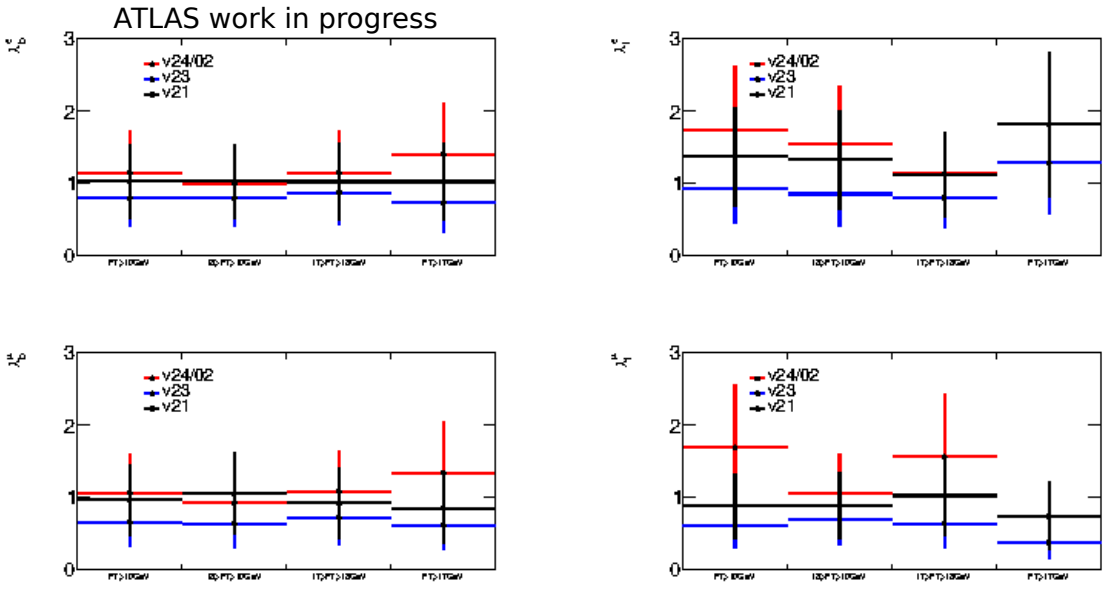


Fig. 4.20 The Scaling Factor dependency on fake lepton transverse momentum for various ntuples versions

- An event with at least a fake lepton is a fake event.
- FSFs are applied to fake events according to their nature (Monte Carlo events with a true top quark referred to “heavy" environment, all others to “light" one) and the flavor of the leptons.
- Events with N expected fakes would be reweighted by $\Pi_{i=1}^{N} FSF_i$

Another method of fake estimation used in Run I is a data-driven approach. In this method, one defines the same control regions enriched in $t\bar{t}$ and Z+jets, constructed with 2 tight leptons plus 2 anti-tight leptons (which explicitly fail one or a set of selection criteria). Identical event-level selections to the signal region are applied, up to and including the Z-veto (i.e., jet-inclusive).

This method is detailed in Annex B.5. However, using the tight/anti-tight CR in $4l$, the low statistics is problematic, requiring more phase space for fakes (looser selection for leptons). Therefore the data-driven method approach was not used for the final $4l$ result but could provide a valid framework for future studies at higher statistics in Run II.

4.3.5 Fakes estimation in 4l

Three methods of calculating the correction due to fakes on the yields have been tested.

In Method 1 (labeled as M1), the identification of fakes is made by requiring that the lepton origin be different from TauLep, top, W or Z boson, Higgs, and diboson. The scaling factors are applied differently for ZJets sample (light fakes) and the top samples (heavy fakes).

In the Method 2 (labeled as M2), besides the origin of the leptons, it is also taken into consideration the type, requiring to be a non-isolated lepton. First, the same scaling procedure as above is applied for the different samples (Zjets and top) and after factorize if more than 1 fake lepton.

In the Method 3 (labeled as M3), it is used the variable lep_isFake (which gives a positive value if the lepton is considered as fake) to separate fakes and apply the same procedure as in Method 2 to scale the MC events.

In conclusion, the results of the fake contribution over the yields in the 4l channel are summarized in Table 4.5 for ttH and 4.6 for the MC background. The fake rates obtained applying various methods vary within 13-14%, within the overall systematic attributed to fakes estimate of 50%.

Selection	**M1**	**M2**	**M3**
Raw	1.672	1.672	1.672
Prompt	1.525	1.525	1.563
Fakes	0.15	0.15	0.11
Fakes_Scaled	0.16	0.16	0.13
Total_Corrected	1.688	1.688	1.694
Fakes_Scaled + bjet>0	0.17	0.18	0.13
Total_Corrected + bjet>0	1.695	1.705	1.688

Table 4.5 Fake estimation on the 4l signal for different fake calculation methods.

Selection	M1	M2	M3
Raw	3.282	3.282	3.282
Prompt	2.824	2.824	2.848
Fakes	0.46	0.46	0.43
Fakes_Scaled	0.71	0.76	0.72
Total_Corrected	3.532	3.588	3.573
Fakes_Scaled + bjet>0	0.88	0.7	0.58
Total_Corrected + bjet>0	3.706	3.526	3.43

Table 4.6 Fake estimation on the 4l background for different fake calculation methods.

4.4 Final Results and Conclusion

The Yukawa coupling of the Higgs boson to the top quark is a key parameter of the Standard Model (SM) opening a potential window to signs of new physics that can modify the $t\bar{t}$H production cross section relative to the SM expectation. This parameter can be determined from the ratio of the top quark mass and Higgs field vacuum expectation value, from the cross-section of gg $\rightarrow$ H production through a top quark loop, or from the cross-section of the three-level process gg/qq $\rightarrow t\bar{t}H$.

A search for $t\bar{t}H$ production in multilepton final states has been performed using 36.1 fb^{-1} of proton-proton collision data at $\sqrt{s}$=13 TeV recorded by the ATLAS experiment at the Large Hadron Collider. Seven final states, targeting Higgs decays to WW*, $\tau\tau$, and ZZ*, are examined for the presence of the Standard Model (SM) Higgs boson: one light lepton and two hadronically-decaying τ lepton candidates ($1l$+$2\tau_{had}$); two same-charge light leptons without any hadronically decaying tau lepton ($2l$SS); two same-charge light leptons with one hadronically decaying tau lepton ($2l$SS+$1\tau_{had}$); two opposite-charge light leptons with one hadronically decaying tau lepton ($2l$OS+$1\tau_{had}$); three light leptons without any hadronically decaying tau lepton ($3l$); three light leptons with one hadronically decaying tau lepton ($3l$+$1\tau_{had}$); and four leptons ($4l$).

A maximum-likelihood fit is performed on all twelve categories simultaneously (7 channels split into 12 categories depending on different region definition) to extract the signal strength, μ, with the assumption of a Higgs boson mass of 125 GeV. The signal strength is defined as the ratio of the measured $t\bar{t}H$ signal yield to the expectation from the Standard Model. In the case of a SM Higgs boson μ =1, but if it is absent then μ = 0. The statistical analysis of the data uses a binned likelihood function $\mathscr{L}(\mu;\vec{\theta})$, which is constructed from a prod-

uct of Poisson probability terms, to estimate μ, which scales the $t\bar{t}H$ production cross section.

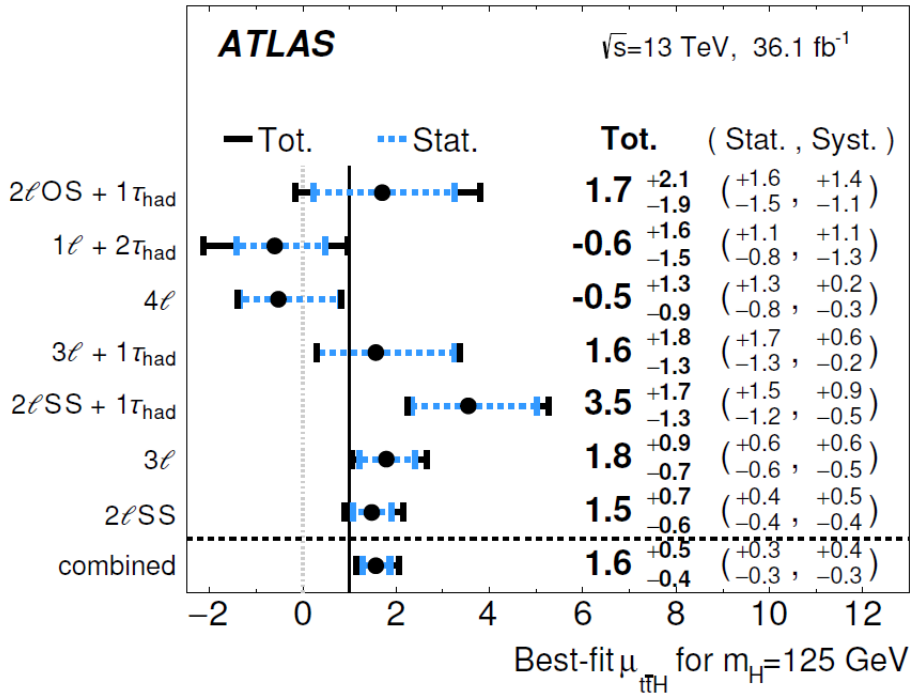


Fig. 4.21 The observed best fit values of the $t\bar{t}H$ signal strength $\mu_{t\bar{t}H}$ and their uncertainties by final state channels and combined. The SM prediction is μ =1

The observed (expected) best-fit value of $\mu_{t\bar{t}H}$, combining all channels, is $1.6^{+0.3}_{-0.3}$ (stat) $^{+0.4}_{-0.3}$ (syst), as shown in Figure 4.21 for each individual final state category and the combination of all channels. The observed (expected) significance with respect to the no-$t\bar{t}H$ hypothesis is 4.1 σ (2.8 σ) while a model-dependent extrapolation is made to the inclusive phase space, and the measured $t\bar{t}H$ production cross section is σ ($t\bar{t}H$) = 790^{+150}_{-150}(stat)$^{+170}_{-150}$(syst) fb. The predicted theoretical cross section is σ ($t\bar{t}H$) = 507^{+35}_{-50} fb.

The impact of the most important groups of systematic uncertainties on the measured value of μ is shown in Table 4.7. The uncertainties with the most substantial impact are those on the signal modeling, the jet energy scale and resolution and the non-prompt light lepton estimate. The uncertainties on the non-prompt light lepton estimates, the fake τ_{had} estimates and the charge miss-assignment have large statistical components due to the data statistics. The small difference between the sum in quadrature of the individual groups and the total uncertainty is due to rounding effects and small correlations between the individual groups.

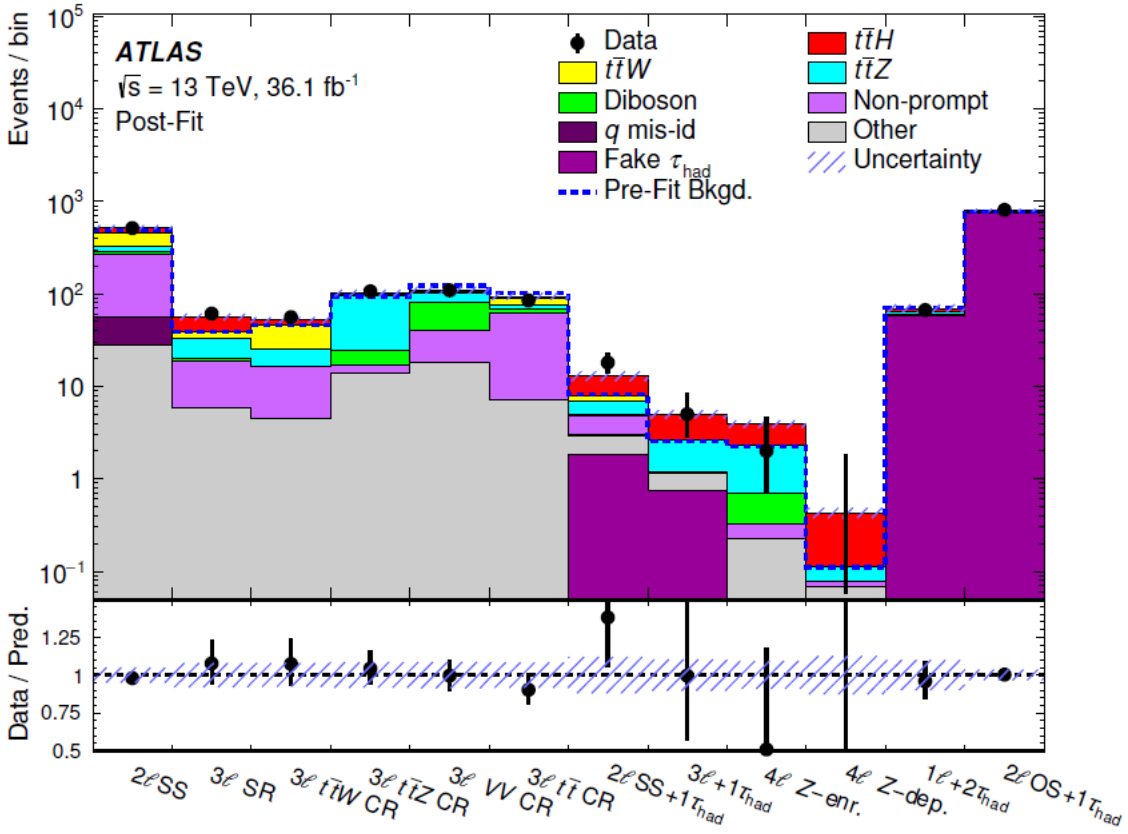


Fig. 4.22 Comparison of prediction to data after the fit in the eight signal and four control regions entering the profile likelihood fit. The systematic uncertainties on the predicted yields are indicated by the hashed blue bands.

Uncertainty Source	$\Delta\mu$
$t\bar{t}H$ modelling (cross section)	+0.20 -0.09
Jet energy scale and resolution	+0.18 -0.15
Non-prompt light-lepton estimates	+0.15 -0.13
Jet flavour tagging and τ_{had} identification	+0.11 -0.09
$t\bar{t}W$ modelling	+0.10 -0.09
$t\bar{t}Z$ modelling	+0.08 -0.07
Other background modelling	+0.08 -0.07
Luminosity	+0.08 -0.06
$t\bar{t}H$ modelling (acceptance)	+0.08 -0.04
Fake τ_{had} estimates	+0.07 -0.07
Other experimental uncertainties	+0.05 -0.04
Simulation statistics	+0.04 -0.04
Charge misassignment	+0.01 -0.01
Total systematic uncertainty	+0.39 -0.30

Table 4.7 Systematic uncertainties on the measured value of μ.

Channel	Best fit $\mu_{t\bar{t}H}$		Significance	
	Observed	Expected	Observed	Expected
2ℓOS+1τ_{had}	$1.7\ ^{+1.6}_{-1.5}$ (stat.) $^{+1.4}_{-1.1}$ (syst.)	$1.0\ ^{+1.5}_{-1.4}$ (stat.) $^{+1.2}_{-1.1}$ (syst.)	0.9σ	0.5σ
1ℓ+2τ_{had}	$-0.6\ ^{+1.1}_{-0.8}$ (stat.) $^{+1.1}_{-1.3}$ (syst.)	$1.0\ ^{+1.1}_{-0.9}$ (stat.) $^{+1.2}_{-1.1}$ (syst.)	-	0.6σ
4ℓ	$-0.5\ ^{+1.3}_{-0.8}$ (stat.) $^{+0.2}_{-0.3}$ (syst.)	$1.0\ ^{+1.7}_{-1.2}$ (stat.) $^{+0.4}_{-0.2}$ (syst.)	-	0.8σ
3ℓ+1τ_{had}	$1.6\ ^{+1.7}_{-1.3}$ (stat.) $^{+0.6}_{-0.2}$ (syst.)	$1.0\ ^{+1.5}_{-1.1}$ (stat.) $^{+0.4}_{-0.2}$ (syst.)	1.3σ	0.9σ
2ℓSS+1τ_{had}	$3.5\ ^{+1.5}_{-1.2}$ (stat.) $^{+0.9}_{-0.5}$ (syst.)	$1.0\ ^{+1.1}_{-0.8}$ (stat.) $^{+0.5}_{-0.3}$ (syst.)	3.4σ	1.1σ
3ℓ	$1.8\ ^{+0.6}_{-0.6}$ (stat.) $^{+0.6}_{-0.5}$ (syst.)	$1.0\ ^{+0.6}_{-0.5}$ (stat.) $^{+0.5}_{-0.4}$ (syst.)	2.4σ	1.5σ
2ℓSS	$1.5\ ^{+0.4}_{-0.4}$ (stat.) $^{+0.5}_{-0.4}$ (syst.)	$1.0\ ^{+0.4}_{-0.4}$ (stat.) $^{+0.4}_{-0.4}$ (syst.)	2.7σ	1.9σ
Combined	$1.6\ ^{+0.3}_{-0.3}$ (stat.) $^{+0.4}_{-0.3}$ (syst.)	$1.0\ ^{+0.3}_{-0.3}$ (stat.) $^{+0.3}_{-0.3}$ (syst.)	4.1σ	2.8σ

Fig. 4.23 The observed and expected best-fit values of the signal strength and associated significance with respect to the SM background-only hypothesis for all channels.

Figure 4.22 shows a comparison in the eight signal and four control regions of the post-fit background and signal yields to data. The observed and expected best-fit values of the signal strength and associated significance with respect to the SM background-only hypothesis for all channels is shown in Figure 4.23.

4.5 Conclusion

In this chapter I focused on the $4l$ channel, part of the $t\bar{t}H \rightarrow multilepton$ analysis, presenting all procedure steps in the analysis: optimization studies, the definition of Signal Regions and Control Regions, fake estimation, fit procedure.

$4l$ channel is included in the seven final states considered in the final combination, targeting Higgs decays to WW*, $\tau\tau$, and ZZ*, are examined for the presence of the Standard Model (SM) Higgs boson: one light lepton and two hadronically-decaying τ lepton candidates ($1l$+$2\tau_{had}$); two same-charge light leptons without any hadronically decaying tau lepton ($2l$SS); two same-charge light leptons with one hadronically decaying tau lepton ($2l$SS+$1\tau_{had}$); two opposite-charge light leptons with one hadronically decaying tau lepton ($2l$OS+$1\tau_{had}$); three light leptons without any hadronically decaying tau lepton ($3l$); three light leptons with one hadronically decaying tau lepton ($3l$+$1\tau_{had}$); and four leptons ($4l$).

The default data and Monte-Carlo version is v26 at integrated luminosity of 36.47 fb^{-1} with a new standard Run1-OLR (overlap removal) and no isolation required on muons. The analysis (optimization, fake estimation, fitting procedure) is robust between the different n-tuple tested, validating therefore the various tools developed by the group and the changes performed by ATLAS collaboration.

The observed (expected) best-fit value of $\mu_{t\bar{t}H}$, combining all channels, is $1.6^{+0.3}_{-0.3}$ (stat) $^{+0.4}_{-0.3}$ (syst). The observed (expected) significance with respect to the no-$t\bar{t}H$ hypothesis is 4.1 σ (2.8 σ) while a model-dependent extrapolation is made to the inclusive phase space, and the measured $t\bar{t}H$ production cross section is σ ($t\bar{t}H$) = 790^{+150}_{-150}(stat)$^{+170}_{-150}$(syst) fb. The predicted theoretical cross section is σ ($t\bar{t}H$) = 507^{+35}_{-50} fb.

This result represents a first step towards the ambitious goal of constraining the measurement of the top Yukawa coupling at the LHC in a direct and most model-independent way in the $t\bar{t}H$ production mode.

The CMS collaboration [80] published observation in Run 1+2, at 25+36 fb^{-1}. They observed an excess of events, with a significance of 5.2 standard deviations, over the expectation from the background-only hypothesis. The corresponding expected significance from the standard model for a Higgs boson mass of 125.09 GeV is 4.2 standard deviations. The combined best-fit signal strength normalized to the standard model prediction is $1.26^{+0.31}_{-0.26}$.

In the summer of 2018, the ATLAS collaboration announced the observation of ttH production in ATLAS with a expected/observed significance of $4.9\sigma/5.8\sigma$ using only Run II results and $5.1\sigma/6.3\ \sigma$ using the combined Run I and Run II results (see Figure 4.24).

Analysis	Integrated luminosity [fb^{-1}]	Expected significance	Observed significance
$H \to \gamma\gamma$	79.8	3.7σ	4.1σ
$H \to$ multilepton	36.1	2.8σ	4.1σ
$H \to b\bar{b}$	36.1	1.6σ	1.4σ
$H \to ZZ^* \to 4\ell$	79.8	1.2σ	0σ
Combined (13 TeV)	36.1–79.8	4.9σ	5.8σ
Combined (7, 8, 13 TeV)	4.5, 20.3, 36.1–79.8	5.1σ	6.3σ

Fig. 4.24 The observed and expected significances and their associated luminosity with respect to the SM background-only hypothesis for all ttH production modes.

In the following chapter I will study the multilepton final state in the pair production $H^{\pm\pm}$ mechanism $p \to \gamma^*, Z^* \to H^{\pm\pm}H^{\mp\mp}$, within a beyond SM framework namely a doublet triplet model (DTM).

Chapter 5

Search for doubly charged scalar bosons decaying to $W^{\pm}\,W^{\pm}$ at LHC

5.1 Introduction

This analysis focuses on BSM searches in $H^{\pm\pm} \rightarrow W^{\pm *}W^{\pm *}$, with a BR = 100%, an analysis never done in ATLAS/CMS. It is complementary to the exotics $H^{\pm\pm} \rightarrow l^{\pm}l^{\pm}$ search [88], studied at the low v_t (the vacuum expectation value of the triplet) region (see section 5.4).

For this analysis, we rely on the SM $t\bar{t}H$ analysis presented in the previous chapter: the objects definition and selection 4.1.1, trigger configuration, some selection for the preselection region.

I will first give the motivation and present the theoretical framework after I will focus on the experimental part of the analysis in the $4l$ channel for which I was the main analyzer. In the end, I will present the outcome of the analysis by combining the $2l$, $3l$ and $4l$ channels.

5.2 Motivation

Although many theoretical aspects of the Standard Model were successfully tested, there are still many theories to be confirmed within the SM or beyond SM framework. Here, we can mention the scalar sector, one of the main candidates to host new particles and forces beyond the Standard Model, accessible by the increased amount of data collected at high energies at LHC Run 2. The addition of a scalar triplet is mentioned in the context of the

so-called "see-saw" models [75], a generic model directed to explain the relative sizes of observed neutrino masses (of the order of eV), much lower compared to the masses of quarks and charged leptons. Despite its theoretical beauty, the see-saw mechanism is difficult to test experimentally and typically contains more parameters than physical observables.

The model used in my analysis [87] is based on the type 1 mechanism which is an extension of the SM, the SM scalar doublet being complemented by a hypercharge Y = 2 scalar triplet, Δ. Five couplings describe the interaction of the fields while the electroweak symmetry breaking is defined by two vacuum expectation values: v_d = 246 GeV and v_t, constrained to the sub-GeV range. The symmetry breaking leads to a rich phenomenology, including seven scalars, $H^{\pm\pm}$, $H^{\pm}$, A^0 (CP-odd), H^0 (CP even), h^0 (CP even). One of the neutral scalars h^0 (scalar) is identified with the SM Higgs boson.

5.3 Phenomenology

The doublet triplet model (DTM) extends the scalar sector to include a hypercharge Y = 2 scalar triplet, Δ , together with the SM scalar doublet, H, leading to the following Lagrangian [74]:

$$\mathscr{L} = (D_\mu H)^\dagger (D^\mu H) + Tr(D_\mu \Delta)^\dagger (D^\mu \Delta) - V(H,\Delta) + \mathscr{L}_{Yukawa} \tag{5.1}$$

with the renormalizable scalar potential $V(H,\Delta)$ is defined in its most general form by:

$$\begin{aligned} V(H,\Delta) = -m_H^2 H^\dagger H + \frac{\lambda}{4}(H^\dagger H)^2 + m_\Delta^2 Tr(\Delta^\dagger \Delta) + [\mu (H^\dagger i\sigma^2 \Delta^\dagger H) + h.c.] + \\ \lambda_1 (H^\dagger H) Tr(\Delta^\dagger \Delta) + \lambda_2 (Tr(\Delta^\dagger \Delta))^2 + \lambda_3 Tr(\Delta^\dagger \Delta)^2 + \lambda_4 H^\dagger \Delta \Delta^\dagger H \end{aligned} \tag{5.2}$$

where the constants μ and λ_{1-4} define the scalar sector couplings and therefore determine the Higgs bosons phenomenology after the electroweak symmetry breaking, $\mathscr{L}_{Yukawa}$ contains the neutrino mass term in addition to the usual fermionic terms and σ represent the Pauli matrices.

The additional Yukawa term to generate neutrino masses is given by

$$\mathscr{L}_{Yukawa} \supset -Y_\nu L^T C \bigotimes i\sigma^2 \Delta L \tag{5.3}$$

Where $\mathscr{L}$ denotes the $SU(2)_L$ lepton doublets and Y_ν denotes the neutrino Yukawa couplings, C denotes the charge conjugation operator and σ^2 denotes the second Pauli matrix. If the triplet acquires a vacuum expectation value, this term will provide non-zero mass of the neutrinos.

Electroweak Symmetry Breaking (EWSB) can be achieved if the neutral components of H and Δ acquire vacuum expectation values of v_d and v_t, respectively, at the minimum of the potential. This causes a mixing between these fields and results in seven scalar bosons: $H^{\pm\pm}$, $H^{\pm}$, A^0 (CP-odd), H^0 (CP even), h^0 (CP even). One of the neutral scalars h^0 (scalar) is identified with the SM Higgs boson.

The fact that the SM Higgs potential is a function of $H^\dagger H$, known as custodial symmetry, implies that $\rho \equiv \frac{M_W^2}{M_Z^2 cos^2\theta_W} = 1$ at tree level, a relation verified experimentally with a very good precision. Any extension of the SM should therefore accommodate only small distortions from this symmetry. In the present doublet-triplet model, a modified ρ at tree level is given by:

$$\rho = \frac{v_d^2 + 2v_t^2}{v_d^2 + 4v_t^2} \tag{5.4}$$

In the limit $v_d \ggg v_t$, $\rho \approx 1 - 2\frac{v_t^2}{v_d^2}$. At 2σ level, the measured $\rho_0 = 1.0004 \pm 0.00048$ [9] leads to an upper bound on v_t of about 1.6 GeV.

5.4 The production and decay modes of $H^{\pm\pm}$

The production mechanisms of $H^{\pm\pm}$ are:

1. pair production $pp \to \gamma^*, Z^* \to H^{\pm\pm}H^{\mp\mp}$
2. associated production $pp \to W^{*+} \to H^{\pm\pm}H^{\mp}$
3. single production $pp \to W^{\pm *}W^{\pm *} \to H^{\pm\pm}$

The single production mode proceeds via vector-boson fusion and is proportional to v_t, hence negligible. The pair- and associated-productions are illustrated in Figure 5.1.

The doubly charged Higgs bosons can decay in two channels: $H^{\pm\pm} \to l^\pm l^\pm$, $H^{\pm\pm} \to W^\pm W^\pm$ with their branching ratios depending on v_t: the WW mode is favoured at higher values of v_t, while the $l^\pm l^\pm$ is dominant at lower values of v_t (see Figure 5.2).

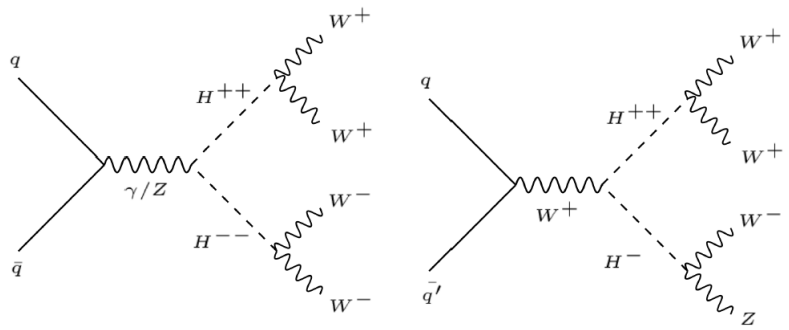


Fig. 5.1 The dominant production modes of doubly charged Higgs bosons in the doublet-triplet model: pair production (left) and associated production (right).

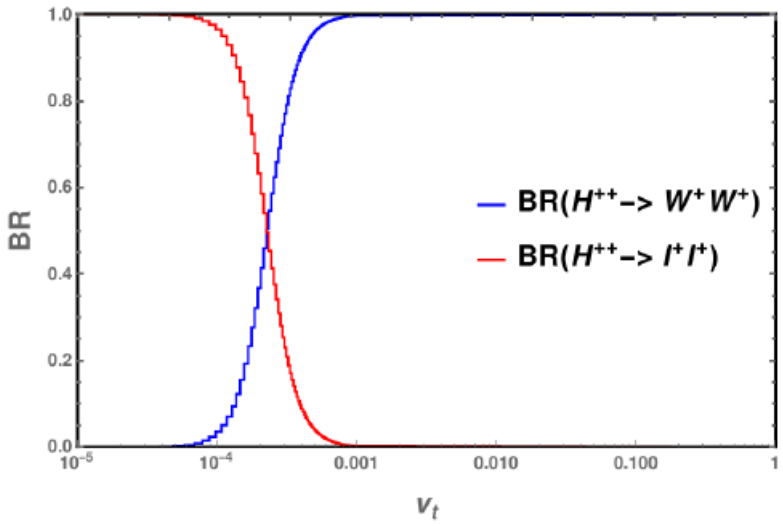


Fig. 5.2 Dependence of Branching Ratio (BR) on the vacuum expectation value of the triplet (v_t).

For the present analysis we focus on the bosonic decay mode $H^{\pm\pm} \to W^{\pm}W^{\pm}$, assumes v_t = 0.1 GeV (to have a BR=1 for $W^{\pm *}W^{\pm *}$ final state, see Figure 5.2) and 10^{-4} value for the mixing between the CP-even Higgses leading to a SM Higgs behavior for the h^0 boson.

$H^{\pm\pm}H^{\mp\mp} \to 4W$	**BR (in %)**
$\to$ 0l + 8 jets	20.1
$\to$ 1l + E_T^{miss} + 6 jets	39.7
$\to$ 2l + E_T^{miss} + 4 jets	29.3
$\to$ 3l + E_T^{miss} + 2 jets	9.6
$\to$ 4l + E_T^{miss} + no jets	1.2

Table 5.1 Pair-production of $H^{\pm\pm} \to W^{\pm}W^{\pm}$

The branching ratios of the various final states containing light leptons (electrons and muons) are shown in Table 5.1. The main difficulty due to large multi-jet processes and experimental issues (like fake lepton pre-dominance, etc.) is found in the channels with 0 or 1 lepton, not included for now for this analysis.

Monte Carlo (MC) simulation samples were produced for signal and background processes using the full ATLAS detector simulation based on Geant4 or a fast simulation using a parameterization of the calorimeter response and Geant4 for the tracking system.

The signal events containing $H^{\pm\pm}$ pairs were simulated with the CalcHEP version 3.4 [77] generator which is at leading order in QCD. Event samples for the process were simulated for $H^{\pm\pm}$ masses in the range from 200 to 700 GeV with steps of 100 GeV. The production cross-section decreases rapidly with $m_{H^{\pm\pm}}$ as shown in Table 5.2.

A filter to select at least two light leptons (electron or muons) with $p_T > 10$ GeV and $|\eta| < 10$ is applied at the truth level, aiming to reduce the rate while enhancing the signal for this analysis in particular since we use $\geq 2l$ final states. The production cross sections and the filter efficiencies are shown in Table 5.2.

$H^{\pm\pm}$**mass (Gev)**	**cross section (fb)**	**filter efficiency**
200	64.58	0.2858
300	13.34	0.3031
400	3.998	0.3198
500	1.466	0.3264
600	0.610	0.3362
700	0.276	0.3451

Table 5.2 Cross sections and filter efficiencies of the signal samples

5.5 Analysis procedure

For this analysis, we rely on $t\bar{t}H$ analysis presented in the previous chapter 4.1.1: the objects definition and selection, trigger configuration, some selection for the preselection region. We use 36.1 fb^{-1} of data collected from proton-proton collision recorded by the ATLAS detector at $\sqrt{s}$ = 13 TeV during 2015 and 2016. The data set has been collected with a bunch crossing of 25 ns.

The multi-boson final states are identified by the presence of leptons, missing transverse energy E_T^{miss} and hadronic jets. For the preselection region, all samples include a requirement that at least one of the leptons be trigger matched and have a transverse momentum above 30 GeV, to ensure a high trigger efficiency. To further reduce the contamination from processes containing top quarks, a veto on b-jets tagged is applied. The jets originating from W's are in general produced by light quarks udsc, leading therefore to events with no b-jets.

On top of kinematic quantities and lepton/jet identification, to further optimize the signal region selection, several event level variables are defined:

- M_{jj}^{W}: the invariant mass of two jets closest to the W-boson mass.
- M_{jets}: the invariant mass of all jets in the event. When there are more than 4 jets in the event, only the leading 4 jets are considered for the calculation of the invariant mass.
- $M_{l_i l_j}$: di-lepton invariant mass. While a complete reconstruction of the $H^{\pm\pm}$ bosons is not possible using leptons due to escaping neutrinos, the variable still reflects the mass scale of the boson.
- $\Delta R_{l_i l_j}$: the distance in $\eta - \phi$ between two same-sign leptons. The same sign leptons tend to close for the signal, while some part the background favors a back-to-back topology. In the 4L channel, two such variables can be calculated per event, and the

minimal and maximal value is used to discriminate the signal from the background. This and the previous variables can be used for any lepton pair in all analysis channels.

- $\Delta\Phi(ll;MET)$: difference in azimuth between the di-lepton system and the E_T^{miss}. This variable is used in the 2LSS analysis.
- RMS: this is a variable used by the 2LSS channel to describe the "spreads" of the azimuthal angles of leptons, E_T^{miss} , and jets.
- ΔR_{l_i-jet}: the minimal distance in eta-phi between a lepton and leading or sub-leading jet.
- $\Delta\Phi_{MET-jet}$: the distance in azimuth between the MET and the leading jet.
- M_{3l}: the invariant mass of three leptons (for the 3L channel).
- M_{4l}: the invariant mass of four leptons (for the 4L channel).

In the next sections, I will give an overall description of the $4L$ channel for which I was the main analyzer. It is a channel combined with $2L$ and $3L$ channels to provide a final result.

I will present the full analysis performed on the $4l$ channel namely: the optimization studies, the definition of Signal Regions and Control Regions, the fake estimation and fit procedure.

5.6 Four-lepton channel

The event topology of the $4l$ signal process is shown in Figure 5.3.

Similar to SM $t\bar{t}$H analysis, the $4l$ channel is low in statistics but with a cleaner background (see the background composition in Figure 5.4). The SM contributions are expected to be dominated by the prompt ZZ production. Smaller contributions are expected from rare processes like $t\bar{t}H$ or four tops.

The signal four lepton signature (Figure 5.3) is expected to have no extra hard jets, large transverse missing momentum, and a typical topology due to the same-sign W's originating from the same $H^{\pm\pm}$ boson.

The preselection sample is defined using the following criteria:

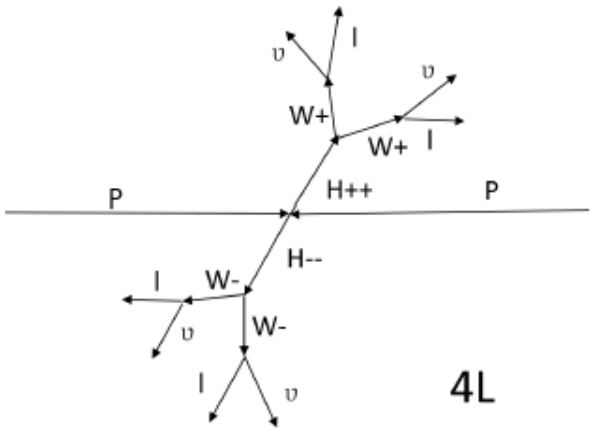


Fig. 5.3 The event topologies of the 4*l* signal process.

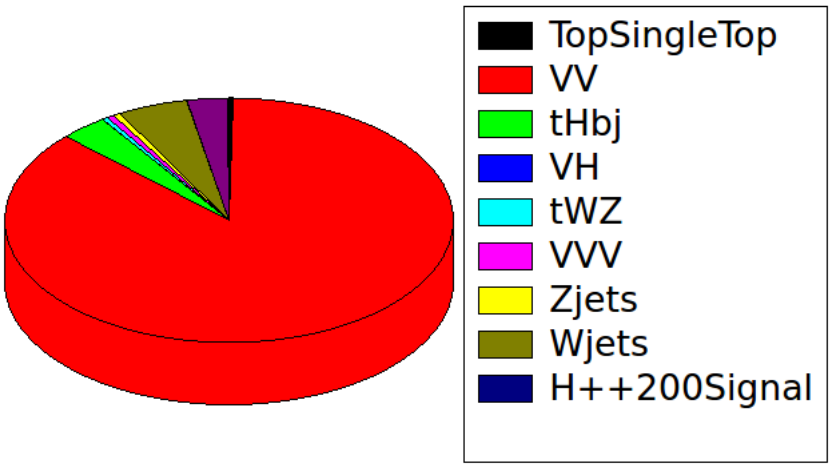


Fig. 5.4 The event composition for the 4*l* channel without any selection applied.

- A) Event are selected if they contain exactly four leptons which have a transverse momentum of at least 10 GeV and pass the "loose" selection. In order to match the trigger requirement, the so-called trigger match is required for at least one lepton that should also have a transverse momentum above 30 GeV. The total charge of the four leptons is requested to be 0. The event cleaning includes the vertex and trigger requirements.

- B) Processes containing Z-bosons: a veto is applied on events which contain at least one same-flavor opposite-sign lepton pair with an invariant mass within a $\pm$ 10 GeV window around the Z-boson mass.

- C) Low mass di-leptons are suppressed: any opposite-sign same flavor lepton pair is required to have an invariant mass exceeding 12 GeV.

- D) $E_T^{miss} > 30$ GeV is required since it further suppresses the inclusive Z+jets production.

- E) ttbar process reduction: events with one or more jets tagged as b-jet using a multivariate identification algorithm are rejected.

5.7 Fake leptons contributions

5.7.1 Definition of the method

The fake contribution uses a semi-data-driven technique as the one used in $t\bar{t}$H analysis described in section 4.3, namely the “Fake Scale Factor" (FSF) method.

In order to increase the statistics, the FSF method is applied in 3 lepton channel and used to correct Monte Carlo prediction of fake rates according to the (candidate fake) lepton flavour and the nature of the environment in which the fake lepton is produced (whether a true top quark is present in the event or not). Therefore, the Fake Scale Factor are λ_L^e, λ_H^e, λ_L^μ, λ_H^μ where L and H stand for light and heavy environment (defined by the Control Regions).

The regions used to calculate the FSF are called Control Regions (CR). They are defined by requiring three leptons, one enriched in Z+jets events (light-jet environment) and the other one with $t\bar{t}$ events (heavy-jet environment), defined as:

Z+jets enriched:

- Exactly 3l loose leptons with P_T > 10 GeV, total charge = $\pm$ 1.
- one Same Flavor Opposite Sign (SFOS) lepton pair with an invariant mass within the Z-mass window $|M_{l^+l^-} - M_Z|$ < 10 GeV.
- Either 1 or 2 jets with transverse momenta above 25 GeV.
- $E_T^{miss} <$ 50 GeV and the transverse mass M_T < 50 GeV.

$t\bar{t}$ enriched

- Exactly 3l loose leptons with P_T > 10 GeV and total charge = $\pm$ 1.
- No SFOS lepton pair, to avoid Z boson contributions.
- Either one or two jets, with $P_T^{(1,2)}$ > 30(25) GeV.

For the Fake Factor calculation the scaling factors are applied as follows:

$$electrons(e) : N_{Data-Prompt} = \lambda_L^e N_{Z+jets} + \lambda_H^e N_{t\bar{t}} \tag{5.5}$$

$$muons(\mu) : N_{Data-Prompt} = \lambda_L^\mu N_{Z+jets} + \lambda_H^\mu N_{t\bar{t}} \tag{5.6}$$

For each flavour of the leptons (electrons and muons) and for each CR, we obtain the following scaling factors, compatible with unity within 10% :

- λ_L^e = 1.12 $\pm$ 0.05(stat) $\pm$ 0.56(syst)
- λ_H^e = 1.02 $\pm$ 0.07(stat) $\pm$ 0.51(syst)
- λ_L^μ = 1.11 $\pm$ 0.05(stat) $\pm$ 0.55(syst)
- λ_H^μ = 0.94 $\pm$ 0.07(stat) $\pm$ 0.47(syst)

where the first errors is statistical and the second are systematics (taken as a conservative value of 50% to cover the fluctuations described below) .

In order to check the stability to sample contributions and evaluate the systematics, the $t\bar{t}$ and $Z + jets$ samples are further split as a function of the b-jets multiplicity:

- A : N_{b-jets} = 0 for $Z + jets$ and N_{b-jets} = 0 for $t\bar{t}$
- B :N_{b-jets} > 0 for $Z + jets$ and N_{b-jets} = 0 for $t\bar{t}$

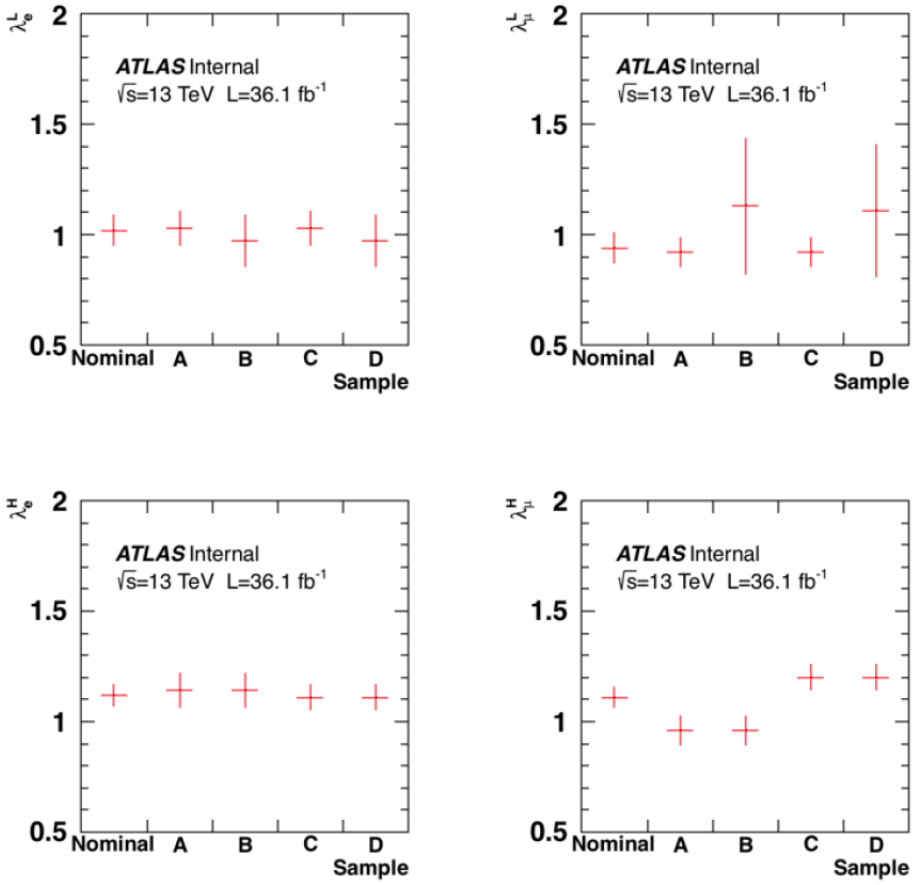


Fig. 5.5 The scaling factors in the CR split as a function of the b-jets multiplicity.

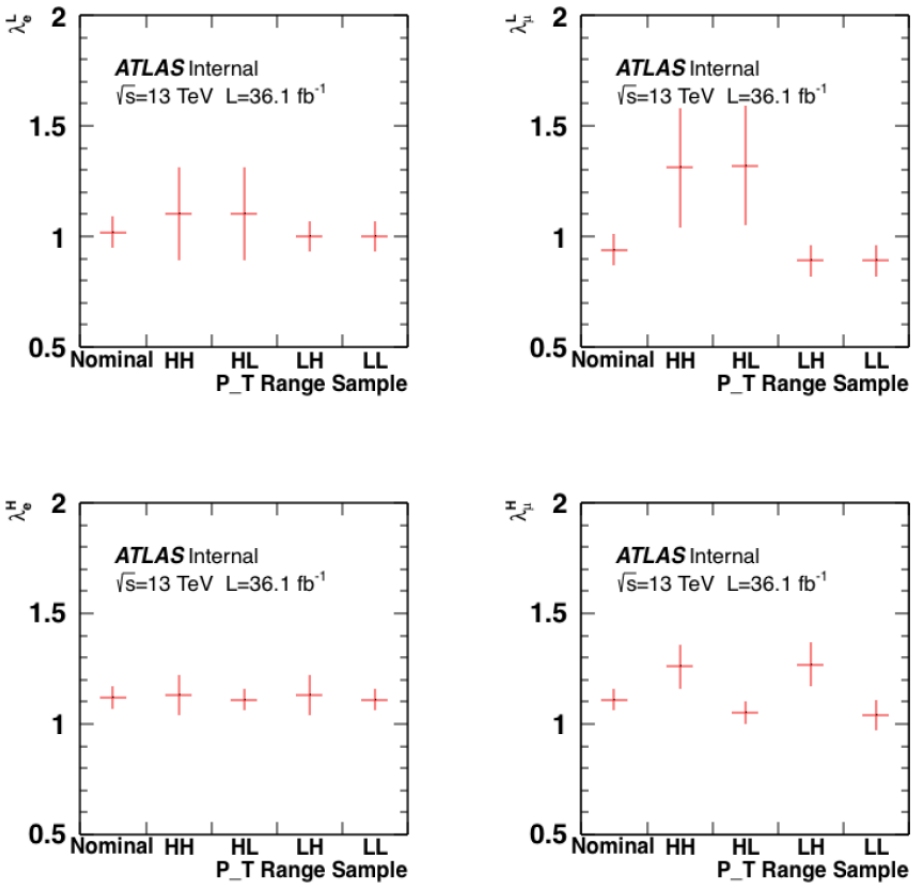


Fig. 5.6 The scaling factors in the CR split as a function of the "fake" lepton candidates transverse momentum.

- C : $N_{b-jets} = 0$ for $Z + jets$ and $N_{b-jets} > 0$ for $t\bar{t}$
- D : $N_{b-jets} > 0$ for $Z + jets$ and $N_{b-jets} > 0$ for $t\bar{t}$

The scaling values observed from this study remain compatible with the unity within an uncertainty of 50%, value taken to be consistent with the $t\bar{t}H$ analysis procedure, as shown in Figure 5.5 that presents the variation of all scaling factors as a function of b-jets multiplicity.

In order to further check the P_T dependence of the scaling factors, the $t\bar{t}$ and $Z + jets$ CR are splited as a function of the "fake" lepton candidates transverse momentum as follows:

- P_T of fake lepton > 18 GeV for $Z + jets$ and P_T of fake lepton > 18 GeV for $t\bar{t}$ (HH)
- P_T of fake lepton > 18 GeV for $Z + jets$ and P_T of fake lepton < 18 GeV for $t\bar{t}$ (HL)
- P_T of fake lepton < 18 GeV for $Z + jets$ and P_T of fake lepton > 18 GeV for $t\bar{t}$ (LH)
- P_T of fake lepton < 18 GeV for $Z + jets$ and P_T of fake lepton < 18 GeV for $t\bar{t}$ (LL)

The scaling values observed from this study remain compatible with the unity within an uncertainty of 50%, as shown in Figure 5.6 that presents the variation of all scaling factors as a function of the "fake" lepton transverse momentum.

5.7.2 The scaling factor application

The scaling is applied as a function of the nature and origin of the leptons, following the same procedure as for $t\bar{t}$H analysis:

- A fake lepton is a reconstructed non-prompt lepton (thus not coming from Higgs boson, t, Z, W or τ)
- An event with at least a fake lepton is a fake event.
- FSFs are applied to fake events according to their nature (Monte Carlo events with a true top quark referred to “heavy" environment, all others to “light" one) and the flavor of the leptons.
- Events with N expected fakes would be reweighted by $\Pi_{i=1}^{N} FSF_i$

In conclusion, by applying the scaling to the 4L analysis samples, using the nominal values and the rescaling procedure, the fakes contributions vary from 21.0 ± 1.7 to 24.8 ± 2.1.

5.7.3 Signal region optimisation

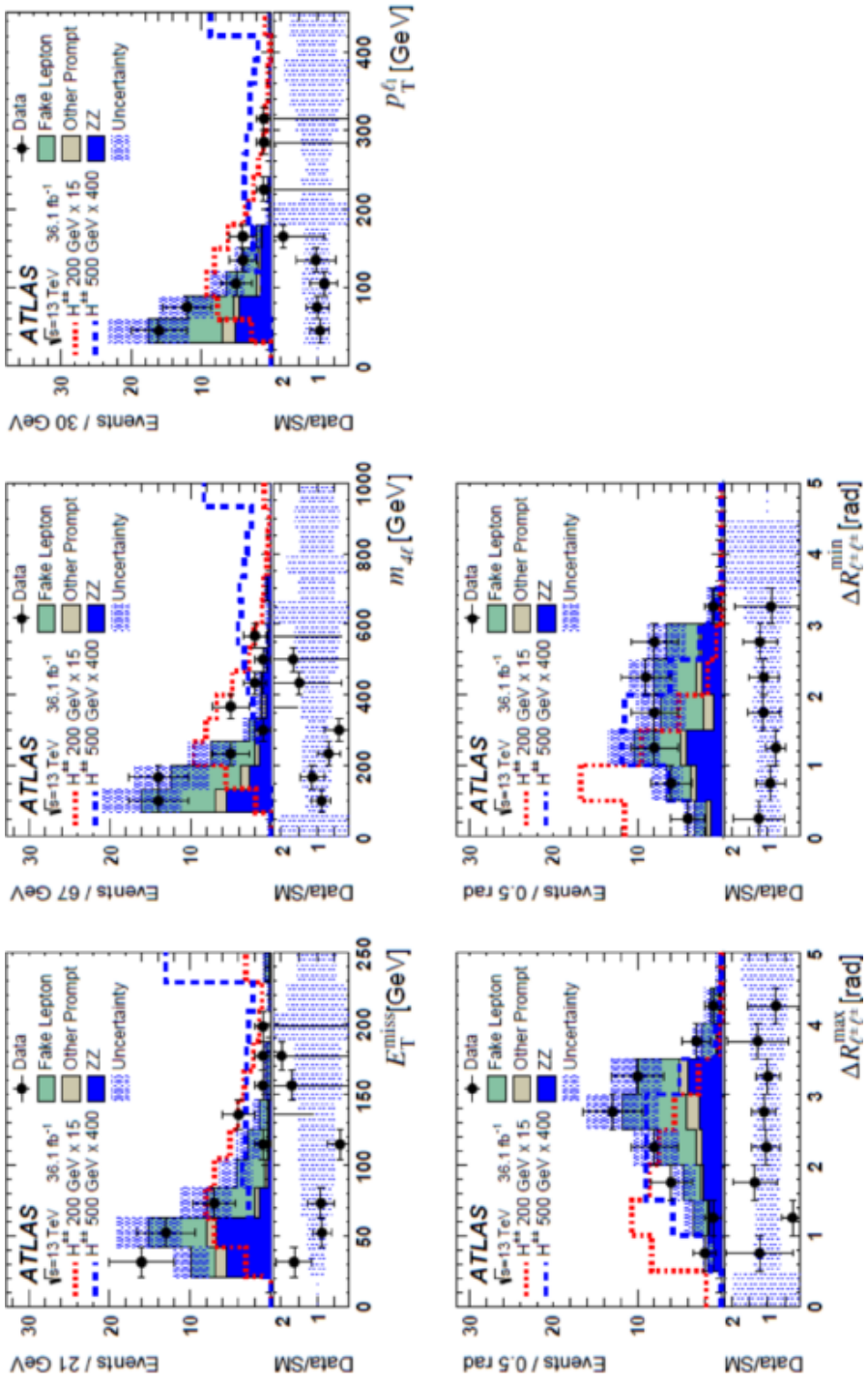


Fig. 5.7 Discriminating variable distributions in the $4l$ final state.

In order to discriminate the Signal from the Background region, a set of variables were tested, and the ones with the best separation power were chosen: E_T^{miss} , the invariant mass of the four leptons, the transverse momentum of the highest p_T lepton, and the minimal $\Delta Rmin_{ll}$ and maximal $\Delta Rmax_{ll}$ distance in $\eta - \Phi$ plane between the same-sign lepton pairs in each event. Their distribution is shown in Figure 5.7. No flavor information is used in the optimization since the split on the electron/muon flavor will not bring an increase in the signal significance due to low statistics available. Therefore the signal region is flavor independent.

The multi-variate cuts optimization algorithm is based on TMVA framework (option CutsSA) and uses the set of variables defined previously. The optimization procedure takes into account the systematic uncertainties.

The algorithm procedure is the same as the one used in Chapter 3 or Chapter 4:

- Discriminating variables selection;
- The TMVA package (CutsSA option) is used to scan the cuts, starting from the preselection level;
- The out-of-range values are regularised (for instance very large upper limits for E_T^{miss} are set to a constant value equivalent to infinity, and similarly for lower limits not activated by the optimization).
- The training performed for each mass is after assigned a target efficiency, obtained at the maximum significance, calculated using two methods "Cowan" and "Roofit", taking into account a systematic uncertainty on the Standard Model prediction.

The optimised cuts are shown as a function of the $H^{\pm\pm}$ mass in Figure 5.8. The efficiencies to select the signal in the preselection region for various doubly charged Higgs masses are:

- ($M_{H^{\pm\pm}}$ = 200 GeV)= 0.40,
- ($M_{H^{\pm\pm}}$ = 300 GeV)=0.58,
- ($M_{H^{\pm\pm}}$ = 400 GeV)=0.65
- ($M_{H^{\pm\pm}}$ = 500 GeV)= 0.65,
- ($M_{H^{\pm\pm}}$ = 600 GeV)=0.71,

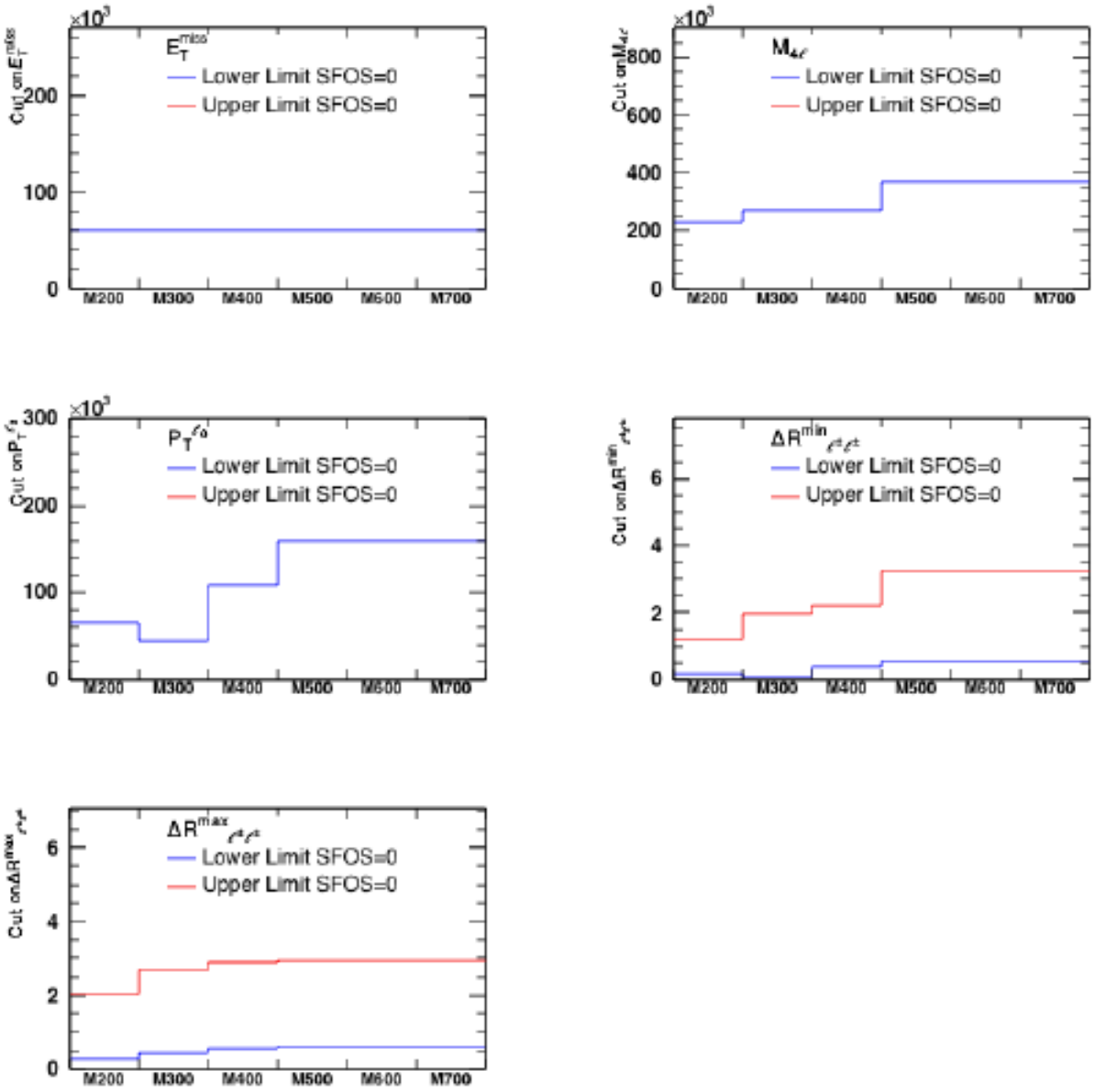


Fig. 5.8 The Lower (Blue) and Upper (Red) cuts obtained for 4L channel after the optimisation TMVA (CutsSA option) procedure.

- ($M_{H^{\pm\pm}}$ = 700 GeV)=0.71.

In the end, the 4L Signal Regions definition and yields for each mass points are listed in Table 5.3 for the 200 GeV doubly charged Higgs mass, Table 5.4 for the 300 GeV doubly charged Higgs mass, Table 5.5 for the 400 GeV doubly charged Higgs mass and Table 5.6 for the 500-700 GeV doubly charged Higgs mass (the 500-700 GeV mass points are combined to minimize the problem of the lack of available statistics).

Cuts	**Data**	**Total MC**	**Prompt**	**Fakes**	$H^{\pm\pm}$ **signal**
Preselection	44	43.2±2.9	18.4±2.0	24.8±2.1	2.27±0.04
E_T^{miss}> 60 GeV	17	20.0±1.9	4.2±0.7	15.7±1.7	1.81±0.04
M_{4L} > 230 GeV	8	4.5±0.8	2.2±0.6	2.3±0.6	1.37±0.04
$p_T^{l_0}$> 65 GeV	8	4.2±0.8	2.1±0.6	2.1±0.6	1.36±0.03
0.16 < $\Delta R_{l^\pm l_\pm}^{min}$ < 1.21 GeV	0	0.5±0.1	0.34±0.04	0.2±0.1	1.08±0.03
0.27 < $\Delta R_{l^\pm l_\pm}^{max}$ < 2.03 GeV	0	0.1±0.0	0.07±0.02	0.03±0.01	0.90±0.03

Table 5.3 The cutflow for the 200 GeV signal region. Only statistical errors are shown.

Cuts	**Data**	**Total MC**	**Prompt**	**Fakes**	$H^{\pm\pm}$ **signal**
Preselection	44	43.2±2.9	18.4±2.0	24.8±2.1	2.27±0.04
E_T^{miss}> 60 GeV	17	20.0±1.9	4.2±0.7	15.7±1.7	0.51±0.01
M_{4L} > 270 GeV	7	2.9±0.7	1.9±0.6	0.9±0.3	0.44±0.01
$p_T^{l_0}$> 45 GeV	7	2.9±0.7	1.9±0.6	0.9±0.3	0.44±0.01
0.09< $\Delta R_{l^\pm l_\pm}^{min}$ < 1.97 GeV	1	1.4±0.3	1.1±0.3	0.4±0.2	0.39±0.01
0.44 < $\Delta R_{l^\pm l_\pm}^{max}$ < 2.68 GeV	0	0.4±0.2	0.25±0.04	0.2±0.2	0.34±0.01

Table 5.4 The cutflow for the 300 GeV signal region. Only statistical errors are shown.

Cuts	**Data**	**Total MC**	**Prompt**	**Fakes**	$H^{\pm\pm}$ **signal**
Preselection	44	43.2±2.9	18.4±2.0	24.8±2.1	2.27±0.04
E_T^{miss}> 60 GeV	17	20.0±1.9	4.2±0.7	15.7±1.7	0.51±0.01
M_{4L} > 270 GeV	7	2.9±0.7	1.9±0.6	0.9±0.3	0.17±0.00
$p_T^{l_0}$> 110 GeV	7	2.6±0.6	1.9±0.5	0.7±0.3	0.17±0.00
0.39 < $\Delta R_{l^\pm l_\pm}^{min}$ < 2.22 GeV	1	1.3±0.3	1.1±0.3	0.2±0.1	0.15±0.00
0.55 < $\Delta R_{l^\pm l_\pm}^{max}$ < 2.90 GeV	1	0.4±0.1	0.33±0.05	0.05±0.02	0.13±0.00

Table 5.5 The cutflow for the 400 GeV signal region. Only statistical errors are shown.

Cuts	**Data**	**Total MC**	**Prompt**	**Fakes**	$H^{\pm\pm}$ **signal**
Preselection	44	43.2±2.9	18.4±2.0	24.8±2.1	2.27±0.04
E_T^{miss}>60 GeV	17	20.0±1.9	4.2±0.7	15.7±1.7	0.51±0.01
M_{4L} > 370 GeV	6	1.7±0.5	1.3±0.5	0.4±0.2	0.07±0.00
$p_T^{l_0}$> 160 GeV	6	1.2±0.5	0.9±0.5	0.4±0.2	0.06±0.00
0.53 < $\Delta R_{l^{\pm}l^{\pm}}^{min}$ < 3.24 GeV	6	1.2±0.5	0.9±0.5	0.4±0.2	0.06±0.00
0.59< $\Delta R_{l^{\pm}l^{\pm}}^{max}$ < 2.94 GeV	3	0.4±0.2	0.24±0.05	0.2±0.1	0.05±0.00

Table 5.6 The cutflow for the 500-700 GeV signal region. Only statistical errors are shown.

5.8 Systematics

Two types of uncertainties are involved in this search: theoretical and experimental.

Experimental uncertainties include effects due to detector simulation and data-driven methods that are used to estimate the instrumental backgrounds.

Contributions to experimental uncertainties are related to the trigger efficiency, lepton reconstruction and identification, jet calibration, continuous b-tagging and the global event activity. The systematic experimental treatments are evaluated by ATLAS performance groups and are used in the multilepton analysis either as an overall event re-weighting or as a rescaling of the object energy and momentum.

The lepton/jet energy scale and resolution, E_T^{miss} uncertainties are 3-8% for signal and 10-30% for background. The efficiency of reco/ID/trigger uncertainties are 4-6% for the signal, 2-5% for the background.

The luminosity and pileup uncertainty corresponding to the dataset collected by the ATLAS experiment in 2015 and 2016 is 5%.

Theoretical uncertainties include mainly:

- MC simulations: affect the prompt leptons identification in the signal process and the background
- PDF uncertainties of the signal yields, found to be in the range of 2.5% to 4.5%
- The signal cross section variation on the factorization scale ranges from 1.5 to 8.7%.

- Normalizations and the acceptances (Signal normalization give a 17% contribution while the theoretical uncertainties on background account for 17-25%)

The PDF uncertainties of the signal yields are evaluated using the LHAPDF6 library [78]. The uncertainties are taken as symmetric (average of up-down variations) for simplicity. The study is performed for the inclusive, 2LSS, 3L and 4L phase spaces, by applying the corresponding pre-selections to the truth level quantities of the signal samples.

The variations due to the factorization scale (taken to be equal to the renormalization scale in this study) follows the study of SUSY SS2L and 3L analysis [79], giving an overall uncertainty of 15% for the signal normalization.

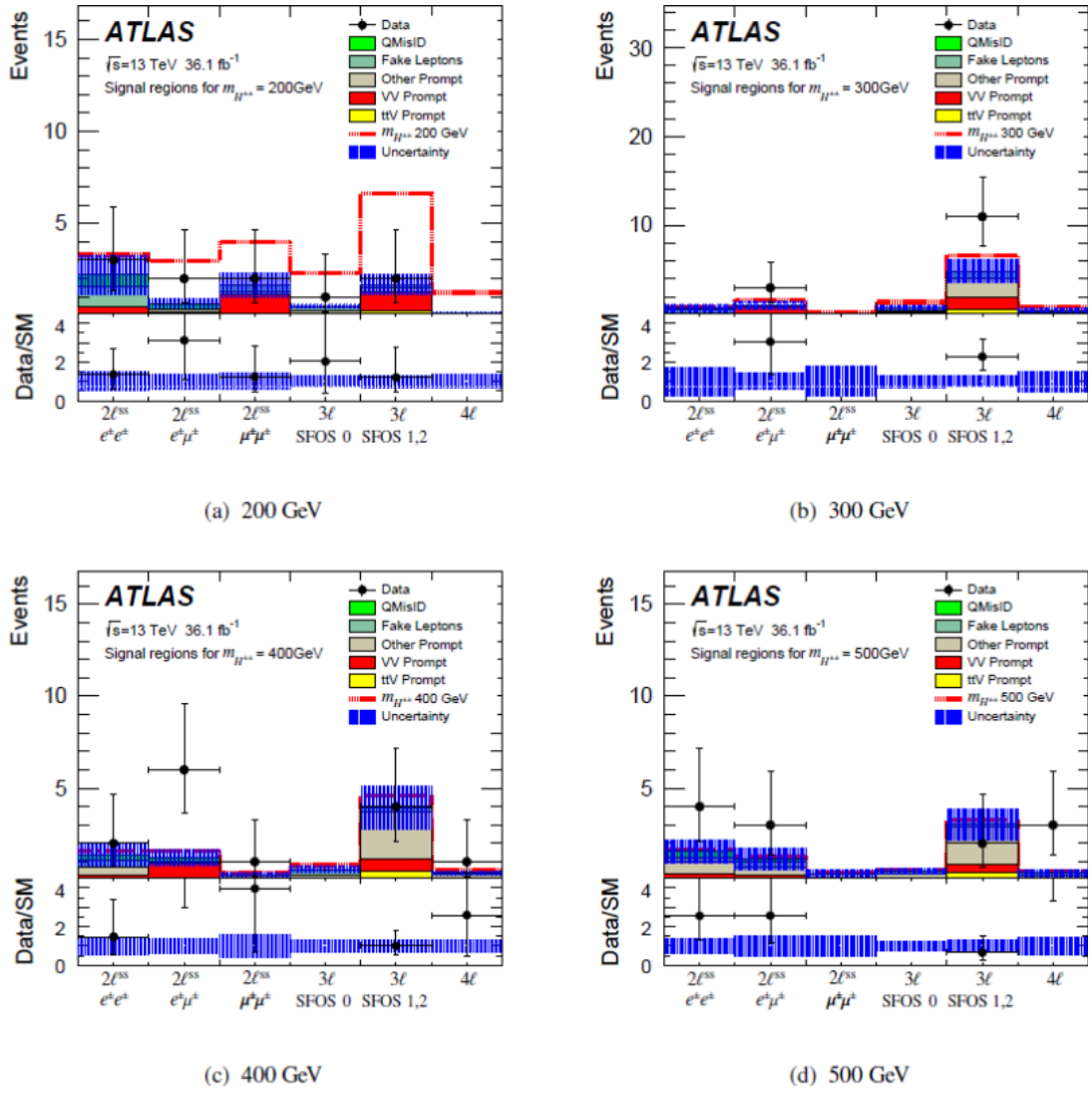


Fig. 5.9 The yields obtained in all channels at $m_H++ = 200$ (up left), 300 (up right), 400 (down left) and 500 (down right) GeV. The error bars represent the full error (statistic and systematic)

5.9 Results

The 4l channel is combined with the 2lSS and 3l channels to provide the final result. To increase the signal significance, the other two channels are divided as a function of SFOS (same flavor opposite sign pairs). Figure 5.9 shows the expected and observed event yields in the signal regions for the 200 - 500 GeV mass points. The data and SM prediction yields obtained for $m_{H^{++}}$ = 500 GeV are also valid for the other two mass points $m_{H^{++}}$ = 600 and 700 GeV. We can observe a good performance in all channels. The error bars represent the full error (statistic and systematic). No significant excess was observed in data. Therefore, we can set a limit on the model considered.

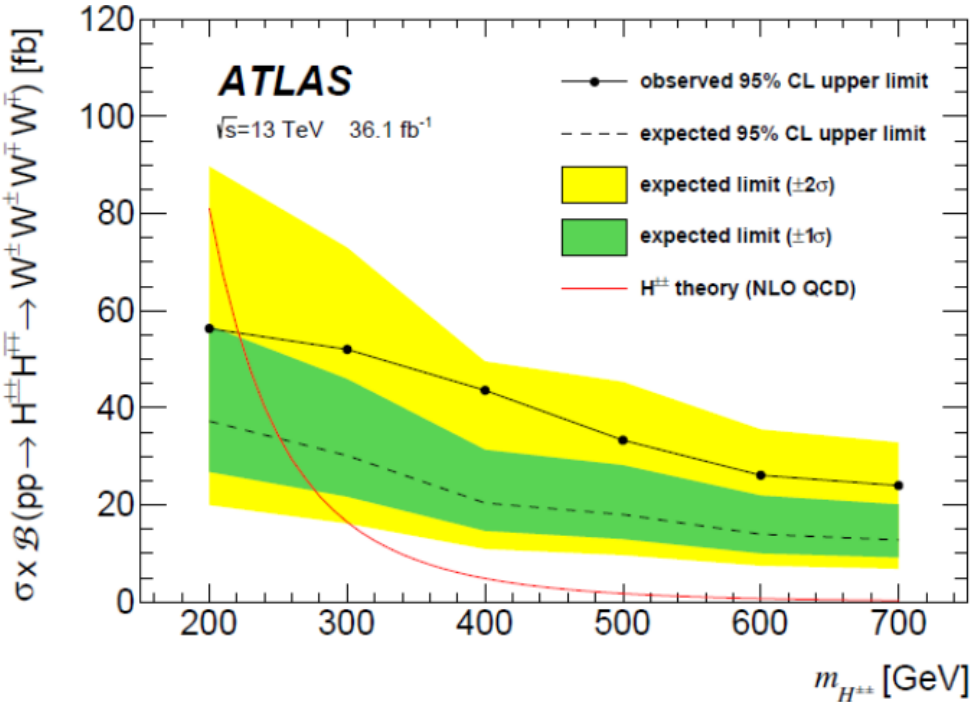


Fig. 5.10 Expected and observed limits for the combination of 2LSS, 3L and 4L channels.

The statistical analysis of the results is done with TtHFitter [73], the parameter of interest being the production cross-section $\times$ the branching ratio. The systematic uncertainties are treated as nuisance parameters implemented in the likelihood function. The likelihood function is constructed by Poisson statistics of counting experiments in the 6 signal regions (2,3,4 leptons + SFOS splitting ($e^{\pm}e^{\pm}, e^{\pm}\mu^{\mp}, \mu^{\pm}\mu^{\pm}$).

The expected and observed significances as a function of the mass of H^{++} are shown in Figure 5.11. Combining all of the 3 channels, the expected significance exceeds 3 standard

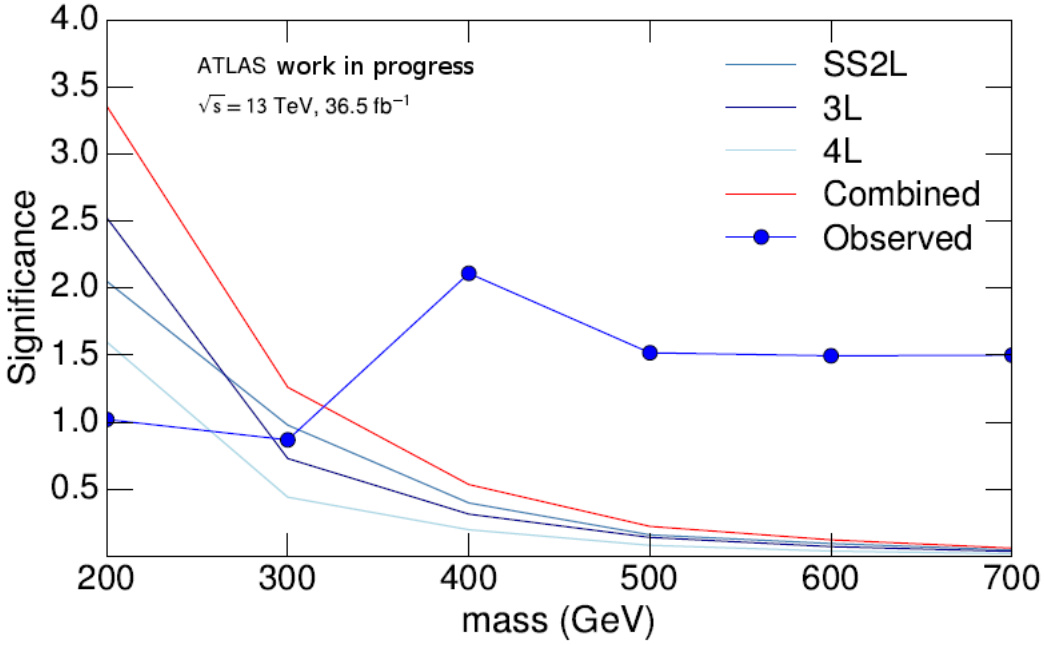


Fig. 5.11 Expected and observed significances as a function of the mass of H^{++} for all 3 channels.

deviations for the mass point of 200 GeV. However, no significant excess is observed.

The expected and observed limits of cross section times branching fraction at 95% C.L. combining 2LSS, 3L and 4L channels is shown in Figure 5.10. The theory is represented by Red line, while the exclusion at 2σ and 1σ are represented in Yellow and Green. No significant excess is observed in data. Combining those channels, the model is excluded at 95% C.L. for masses between 200 and 220 GeV.

5.10 Conclusion

In this chapter, I presented the study of the multilepton final state in the pair production $H^{\pm\pm}$ mechanism $p \to \gamma^*, Z^* \to H^{\pm\pm}H^{\mp\mp}$, within a beyond SM framework namely a doublet triplet model (DTM). This analysis consider a v_t = 0.1 GeV (to have a BR=1 for $W^{\pm *}W^{\pm *}$ final state, see Figure 5.2) and 10^{-4} value for the mixing between the CP-even Higgses.

This analysis focuses on multilepton final state with ≥ 2 leptons, sharing in common with the $t\bar{t}H$ analysis presented in the previous chapter 4.1.1 the object's definition and selection,

trigger configuration, some selection for the preselection region.

The signal samples have been simulated with $m_{H^{++}}$ = 200, 300, 400, 500, 600 and 700 GeV respectively.

I focused on the 4l channel, presenting all procedure steps in the analysis: optimization studies, the definition of Signal Regions and Control Regions, fake estimation, fit procedure.

Putting together all the 6 signal regions (2,3,4 leptons + SFOS splitting), this study offers the first limits on the production and decay channels considered, excluding at 95% C.L. for $H^{\pm\pm}$ masses between 200 and 220 GeV.

Chapter 6

Conclusions

The data sample collected in 2015-2017 by ATLAS, has been used to measure the efficiencies for the electron identification using the tag and probe methods. The identification algorithms are based on a multivariate likelihood discriminator that has been optimized to cope with the dependence of the pileup conditions and is well reproduced by the Monte Carlo simulation. The differences between data and MC simulation have been incorporated into data-to-MC ratios (scale-factors), obtained as a function of the transverse energy and the pseudorapidity of the electrons.

I presented in this thesis the electron identification efficiency using $W \to e\nu$ decays. Numerous advantages come for using $W \to e\nu$ events to measure the electron identification efficiency: W and Z bosons cover both a large kinematic range (as can be seen in Figure 3.1) but $W \to e\nu$ events could help in understanding the uncertainty by offering different systematics from Z. Also, W provides additional statistics for the low E_T range (15-25 GeV) that can complement the J/Ψ channel and offers interesting experimental signatures due to its large statistics.

The main difficulty of this decay channel comes from the background. To have a clean signal, optimization studies were performed using cut-based and TMVA methods. Further, for the 2017-2018 a new set of trigger configuration was proposed and validated, making the $W \to e\nu$ an interesting channel to provide cross-check on the scale factors obtained using $Z \to ee$ events.

The SF using the W tag and probe method as a comparison with Ziso tag and probe for all background templates were presented, with a 10 % agreement between the W and Ziso in the bulk region (central η). This represents the first measurement on the Run 2 data for the

electron efficiency using $W \to e\nu$ events.

My work was valuable for the ATLAS collaboration due to the novelty of the measurement: the optimization and background studies, the first look on the efficiency and scale factors using $W \to e\nu$ events in Run 2 data, the trigger emulation studies.

Among the main goals of the Large Hadron Collider was the discovery of the Higgs boson, a significant achievement of Run I. Upon discovery of a 125 GeV Higgs boson, a large effort is dedicated in Run II to the measurement of its properties, especially the Higgs boson coupling to quarks. Because the top-quark is the heaviest known elementary particle, it has the largest coupling to the Higgs boson. To directly asses the measurement of this coupling at the LHC experiment, one can search the associated production of a Higgs boson with a pair of top-quarks ($t\bar{t}H$). This thesis presents a search of the $t\bar{t}H$ production mode in the multileptonic final state based on 36.1 fb^{-1} of data collected in 2015 and 2016 by the ATLAS detector at the LHC at $\sqrt{p}$ = 13 TeV.

Seven final states, targeting Higgs boson decays to WW*, $\tau\tau$, and ZZ*, are examined for the presence of the Standard Model (SM) Higgs boson: one light lepton and two hadronically-decaying τ lepton candidates ($1l$+$2\tau_{had}$); two same-charge light leptons without any hadronically decaying tau lepton ($2l$SS); two same-charge light leptons with one hadronically decaying tau lepton ($2l$SS+$1\tau_{had}$); two opposite-charge light leptons with one hadronically decaying tau lepton ($2l$OS+$1\tau_{had}$); three light leptons without any hadronically decaying tau lepton ($3l$); three light leptons with one hadronically decaying tau lepton ($3l$+$1\tau_{had}$); and four leptons ($4l$).

In this thesis I focused on the $4l$ channel, presenting all procedure steps in the analysis: optimization studies, the definition of Signal Regions and Control Regions, fake estimation and fit procedure.

The observed (expected) best-fit value of $\mu_{t\bar{t}H}$, combining all channels, is $1.6^{+0.3}_{-0.3}$ (stat) $^{+0.4}_{-0.3}$ (syst). The observed (expected) significance with respect to the no-$t\bar{t}H$ hypothesis is 4.1 σ (2.8 σ) while a model-dependent extrapolation is made to the inclusive phase space, and the measured $t\bar{t}H$ production cross section is σ ($t\bar{t}H$) = 790^{+150}_{-150}(stat)$^{+170}_{-150}$(syst) fb. The predicted theoretical cross section is σ ($t\bar{t}H$) = 507^{+35}_{-50} fb.

This result represents a first step towards the ambitious goal of constraining the measurement of the top Yukawa coupling at the LHC in a direct and most model-independent way in the $t\bar{t}H$ production mode.

Adding a scalar triplet to the Standard Model framework directly explain the observed non-zero neutrino masses (of the order of eV). The model is based on the type 1 mechanism which is an extension of the SM, the SM scalar doublet being complemented by a hypercharge Y = 2 scalar triplet, Δ. Five couplings describe the interaction of the fields while the electroweak symmetry breaking is defined by two vacuum expectation values: v_d = 246 GeV and v_t, constrained to the sub-GeV range. The symmetry breaking leads to a rich phenomenology, including seven scalars, $H^{\pm\pm}$, $H^{\pm}$, A^0 (CP-odd), H^0 (CP even), h^0 (CP even). One of the neutral scalars h^0 (scalar) is identified with the SM Higgs boson.

This model is used in the search for the doubly charged Higgs analysis, in the multilepton final state (2LSS, 3L and 4L channels) within the pair production mechanism followed by the decay of $H^{\pm\pm} \rightarrow W^{\pm}W^{\pm}$. In this thesis I focused on the $4l$ channel, presenting all procedure steps in the analysis: optimization studies, the definition of Signal Regions and Control Regions, fake estimation and fit procedure. Combining all final state channels, this study offers the first limits on the production and decay channels considered, excluding at 95% C.L. the $H^{\pm\pm}$ masses between 200 and 220 GeV.

Appendix A

Theoretical Aspects

A.1 Gauge symmetries

The gauge groups are important bricks to construct the quantum field theory and, in particular, the field theory of elementary particles. The invariance under gauge transformations requires the introduction of gauge vector fields, which are interpreted as the quanta mediating the interactions among the fermions. The mechanism of spontaneous symmetry breaking (SSB) uses the invariance of the Lagrangian of the system under a specific gauge symmetry, keeping a lower symmetry for the solutions of the equations of motion (in particular, the lowest energy state, i.e., the vacuum). Through this mechanism, some of the fields, including the massless gauge vector bosons, will acquire a mass.

In the Standard Model, at Fermi scale, weak and electromagnetic interactions have been unified at a specific energy scale (of the order of 10^3 GeV), in the $SU(2) \bigotimes U(1)$ gauge field theory.

A.1.1 Invariance under group transformations and conservation laws

An important concept of quantum field theory is the correspondence between conservation laws and symmetry properties expressed mathematically by Noether's theorem. For example, the symmetries and invariance properties of the geometrical kind are:

- the invariance with respect to time translation $\rightarrow$ the conservation of energy
- the invariance of the laws of nature under space translations of the coordinate frame of reference $\rightarrow$ conservation of momentum

- invariance under rotation of the coordinate system about an arbitrary axis → the conservation of angular momentum
- Equivalence of all frames of reference in uniform relative motion → the Lorentz transformations and to the laws of special relativity
- Equivalence of left-handed and right-handed coordinate systems → conservation of parity

However, the internal symmetries that describe dynamical aspects of particle interactions cannot be described in geometrical terms.

We start by considering a physical system described by a Lagrangian L, expressed as a spatial integral of a Lagrangian density $\mathscr{L}(\phi, \partial_\mu \phi)$. We can write then the action as:

$$S = \int L dt = \int \mathscr{L}(\phi, \partial_\mu \phi) d^4x \tag{A.1}$$

The principle of least action states that, the evolution of the system between two times t_1 and t_2 corresponds to an extremum:

$$\delta S = 0 \tag{A.2}$$

From this we derive the **Euler-Lagrange equation**, which describes the motion of the system:

$$\partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\right) - \frac{\partial \mathscr{L}}{\partial \phi} = 0 \tag{A.3}$$

Starting from Noether's theorem an infinitesimal variation of form $\phi(x) \to \phi'(x) = \phi(x) + \delta\phi(x)$ is a symmetry of the system if it leaves the action (therefore the Euler-Lagrange equation), invariant.

Since $\delta\mathscr{L} = \partial_\mu \mathscr{J}^\mu(x)$ we can obtain the current from:

$$\begin{aligned} \delta\mathscr{L} &= \frac{\partial \mathscr{L}}{\partial \phi}\delta\phi + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\delta(\partial_\mu \phi) \\ &= \left(\frac{\partial \mathscr{L}}{\partial \phi} - \partial_\mu\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\right)\right)\delta\phi + \partial_\mu\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\delta\phi\right) \\ &\to j^\mu = (x) = \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\delta\phi - \mathscr{J}^\mu(x) \end{aligned} \tag{A.4}$$

which satisfies the condition $\partial_\mu j^\mu(x) = 0$ leading to a conservation law in integral form as:

$$Q = \int j^0(x) d^3x \tag{A.5}$$

expression constant in time:

$$\frac{dQ}{dt} = \int \partial_0 j^0(x) d^3x = -\int \nabla \mathbf{j}(x) d^3x = 0 \tag{A.6}$$

A.1.2 The mechanism of spontaneous symmetry breaking

The mechanism of spontaneous symmetry breaking (SSB), first introduced by Higgs, could explain the appearance of heavy bosons as exchange particles although, in cases of QED and QCD, the gauge vector bosons are massless, a fact in contrast with the short-range nature of weak interactions.

A.1.3 Spontaneous symmetry breaking of a discrete symmetry

We start from the Lagrangian density of a real scalar field of form:

$$\mathscr{L} = T - V = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - V(\phi) \tag{A.7}$$

where $V(\phi)$, the potential, contains the mass and the self-interaction terms.
In its simplest form, it is given by:

$$V(\phi) = \frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4 \tag{A.8}$$

with $\lambda > 0$ to ensure the existence of a spectrum of stable bound states.

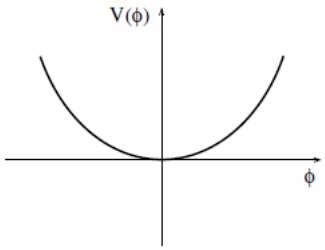


Fig. A.1 Case of a discrete symmetry: the potential $V(\phi)$ for $\mu^2 > 0$.

We want to analyze the properties of the ground state which corresponds to the vacuum state. We consider here two cases:

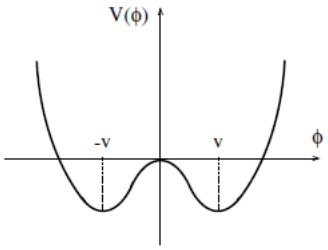


Fig. A.2 Case of a discrete symmetry: the potential $V(\phi)$ for $\mu^2 < 0$.

- In the case $\mu^2 > 0$ (Figure A.1) , the behavior of the potential describes a scalar field of mass μ, with the ground state identified by $\phi = 0$ and excited levels obtained through a perturbative expansion around $\phi = 0$.

- In the case $\mu^2 < 0$ (Figure A.2), the potential $V(\phi)$ exhibits two limits $\phi_{min} = \pm v$ with $v = \sqrt{\frac{-\mu^2}{\lambda}}$

If now we want to describe the spectrum of the system and choose for example $\phi_{min} = \pm v$ as ground state and shift the field $\phi(x)$ in such a way that the ground state occurs at the zero of the new field $\eta(x)$: $\phi(x) \rightarrow \eta(x) + v$, we have the new form of the Lagrangian density as:

$$\mathscr{L}' = \frac{1}{2}\partial_\mu \eta \partial^\mu \eta - v^2 \lambda \eta^2 - v\lambda \eta^3 - \frac{1}{4}\lambda \eta^4 + const \tag{A.9}$$

From here, we can see that the new field $\eta(x)$ has the mass:

$$m_\eta = \sqrt{2v^2 \eta} = \sqrt{-2\mu^2} \tag{A.10}$$

Now the reflexion symmetry is hidden, the Lagrangian densities $\mathscr{L}$ and $\mathscr{L}'$ are completely equivalent; it is only the specific choice of the ground state that breaks the symmetry. This is called the **mechanism of spontaneous symmetry breaking**.

A.1.4 Spontaneous symmetry breaking of a continuous global symmetry

We start with the Lagrangian density of a complex scalar field $\phi(x)$, to which we add a self-coupling term:

$$\mathscr{L} = (\partial_\mu \phi) * (\partial^\mu \phi) - \mu^2 \phi * \phi - \lambda (\phi * \phi)^2 \tag{A.11}$$

which is invariant under the global transformations of the group $U(1)$:

$$\phi(x) \to e^{i\alpha}\phi(x), \phi^*(x) \to e^{-i\alpha}\phi^*(x) \tag{A.12}$$

which gives rise to an additive conserved quantity, usually interpreted as the electric charge of the system.

The potential of the system is $V(\phi) = \mu^2|\phi|^2 + \lambda|\phi|^4$ and in the case $\mu^2 > 0$, the usual minimum of the potential is given by $|\phi_{min}| = 0$. If $\mu^2 < 0$, the minimum is obtained for $|\phi|_{min} = -\frac{\mu^2}{2\lambda}$ which represents a continuum set of solutions, since the phase of ϕ at the minimun is completely arbitrary: $\phi_{min} = |\phi|_{min}e^{i\gamma}$. This solution represents an infinite set of degenerate vacuum states: the choice of a particular vacuum state gives rise to spontaneous symmetry breaking (in here we choose the value $\gamma = 0$). Then, by transforming $\phi = \frac{1}{\sqrt{2}}[\phi_1(x) + i\phi_2(x)]$, the minimum of the potential is given by $(\phi_1)_{min} = \sqrt{\frac{-\mu^2}{\lambda}} \equiv v$, v is called vacuum expectation value (v.e.v.) of the scalar field.

As in the case of the discrete symmetry, we shift the two fields ϕ_1, ϕ_2 according to our choice of the minimum: $\phi_1(x) = \eta(x) + v$ and $\phi_2(x) = \zeta(x)$. Then, the Lagrangian density becomes:

$$\mathscr{L}' = \frac{1}{2}\partial_\mu\eta\,\partial^\mu\eta + \frac{1}{2}\partial_\mu\zeta\,\partial^\mu\zeta - \frac{1}{2}(-2\mu^2)\eta^2 - \frac{1}{4}\lambda(\eta^2+\zeta^2)^2 - \lambda v(\eta^2+\zeta^2)\eta \tag{A.13}$$

The Lagrangian densities $\mathscr{L}$ and $\mathscr{L}'$ are utterly equivalent although written differently. Thus, in principle, the use of either of the Lagrangian densities should give the same physical results; however, this would be true only if one could get an exact solution of the theory.

The first configuration, for $\mu^2 > 0$, contains a complex scalar field ϕ with squared mass μ^2. In the second expression, $\mu^2 < 0$ implies the presence of real scalar field $\eta(x)$ which represents the quantum excitations and has squared mass $m_\eta^2 = -2\mu^2 = 2\lambda v^2$, and a massless real scalar field ζ.

As a consequence, changing μ^2 continuously with the energy scale from positive to negative values leads to a change in the behavior of the system as a phase transition, from a state which is invariant under the transformations of the group U(1) to a new state in which this invariance is lost. This is a simple example of the appearance of massless scalar bosons called **Goldstone bosons** in the case of spontaneous symmetry breaking (SSB) of a

continuous symmetry.

A.2 Dirac equation

In nonrelativistic quantum mechanics particles are described by the **Schrodinger** equation; in relativistic quantum mechanics particles of spin 0 are described by the **Klein-Gordon** equation, particles of spin 1/2 by the **Dirac** equation, and particles of spin 1 by the **Proca** equation.

The Schrödinger equation is given by:

$$\frac{-\hbar}{2m}\nabla^2\Psi + V\Psi = i\hbar\frac{\partial\Psi}{\partial t} \tag{A.14}$$

The Klein-Gordon equation can be obtained in the same way, by requiring the relativistic energy-momentum relation, $E^2 - p^2c^2 = m^2c^4$, or $p^\mu p_\mu - m^2c^2 = 0$.

Making the substitution $p_\mu \rightarrow i\hbar\partial_\mu$ we can write the Klein-Gordon equation:

$$-\hbar^2\partial^\mu\partial_\mu\Psi - m^2c^2\Psi = 0 \tag{A.15}$$

Dirac looked for a wave equation that was first order in both space and time derivatives:

$$\widehat{E}\Psi = (\alpha\cdot\widehat{p} + \beta m)\Psi \tag{A.16}$$

Where the Dirac Hamiltonian operator is:

$$\widehat{H} = (\alpha\cdot\widehat{p} + \beta m) \tag{A.17}$$

It must be Hermitian in order to have real eigenvalues, the α and β matrices also must be Hermitian.The Dirac Hamiltonian is a 4×4 matrix of operators that must act on a four-component wavefunction, namely the Dirac spinor:

$$\Psi = \begin{pmatrix}\Psi_1\\ \Psi_2\\ \Psi_3\\ \Psi_4\end{pmatrix} \tag{A.18}$$

The algebra of the Dirac equation is fully defined by the relations:

$$\alpha_x^2 = \alpha_x^2 = \alpha_x^2 = \beta^2 = I \tag{A.19}$$

$$\alpha_j\beta + \beta\alpha_j = 0 \tag{A.20}$$

$$\alpha_j\alpha_k + \alpha_k\alpha_j = 0 (j \neq k) \tag{A.21}$$

The conventional choice of describing matrices α and β is the Dirac–Pauli representation, based on the Pauli spin-matrices:

$$\beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{A.22}$$

and

$$\alpha = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \tag{A.23}$$

with

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{A.24}$$

A.2.1 Solutions to the Dirac equation

By looking for free-particle plane wave solutions of the form

$$\Psi(x,t) = u(E,p)e^{i(p\cdot x - Et)} \tag{A.25}$$

where u(E, p) is a four-component Dirac spinor and the overall wavefunction $\Psi(x,t)$ satisfies the Dirac equation written in the following form:

$$(i\gamma^\mu \partial_\mu - m)\Psi(x) = 0 \tag{A.26}$$

By defining the four Dirac γ-matrices as

$$\gamma_0 = \beta \tag{A.27}$$

$$\gamma_i = \beta\alpha_i, i = (x,y,z) \tag{A.28}$$

The Dirac equation provides an elegant mathematical framework for the relativistic quantum mechanics of spin-half fermions in which the properties of spin and magnetic moments emerge naturally. However, the presence of negative energy solutions is unavoidable. It is,

therefore, necessary to provide a physical interpretation of the negative energy solutions namely the **Dirac sea interpretation**. Dirac proposed that the vacuum corresponds to the state where all negative energy states are occupied. In this "Dirac sea" picture, the Pauli exclusion principle prevents positive energy electrons from falling into the fully occupied negative energy states.

A.3 The Klein–Gordon equation

The Lorentz invariant Lagrangian density:

$$\mathscr{L} = \frac{1}{2}[g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - m^2\phi^2] = \frac{1}{2}[\partial_\mu\phi\partial^\mu\phi - m^2\phi^2] \tag{A.29}$$

where $\phi(x)$ is a real scalar field. The field equation becomes:

$$-\partial_\mu\partial^\mu\phi - m^2\phi = 0 \tag{A.30}$$

or:

$$(-\frac{\partial^2}{\partial t^2} + \nabla^2 - m^2)\phi = 0 \tag{A.31}$$

This equation is known as **the Klein–Gordon equation**. The equation has wave-like solutions:

$$\phi(r,t) = acos(k\cdot r - \omega_k t + \theta_k) \tag{A.32}$$

where the frequency ω_k is related to the wave vector k by the dispersion relation $\omega_k^2 = k^2 + m^2$ and θ_k is an arbitrary phase angle.

Appendix B

Experimental aspects

B.1 EGAM Trigger Configuration

L1Topo menu naming convention The Topological Processor for the Level-1 Topological Trigger (L1Topo) performs real-time event selection based on the geometric and kinematic relationships between Trigger OBjects (TOBs) (i.e., electrons/photons, muons, jets, and taus), as well as event-level quantities such as missing transverse energy (E_T^{miss}).

The L1Topo latency is $\sim$ 300 ns = 12 Bunch Crossings (BCs) and between 4-5 BCs are reserved for L1Topo algorithms. Out of these 4 - 5 BCs, 1 - 2 BCs are used for preparing the sorted and abbreviated lists of TOBs that the algorithms operate on. This uses the largest amount of resources on the FPGAs (see Figure B.1). That leaves 2-4 BCs for executing the topological algorithms which are foreseen to consume a relatively small fraction of the resources available.

The currently proposed list of topological calculations in the L1Topo module includes $\Delta\eta$ (DETA), $\Delta\phi$ (DPHI), ΔR (DR) (actually combination of $\Delta\eta$ and $\Delta\phi$ (DETA-DPHI) cuts is favoured for low cut values), invariant mass, H_T and HC_T.

B.1.1 Definition of the variables used in L1 topo

Angular requirements

- $\Delta\eta$ (DETA): ensures that a given TOB is inside a given η range, e.g. MU.ETA10 means that this muon is in the barrel region $|\eta|$ (MU) < 1.05.

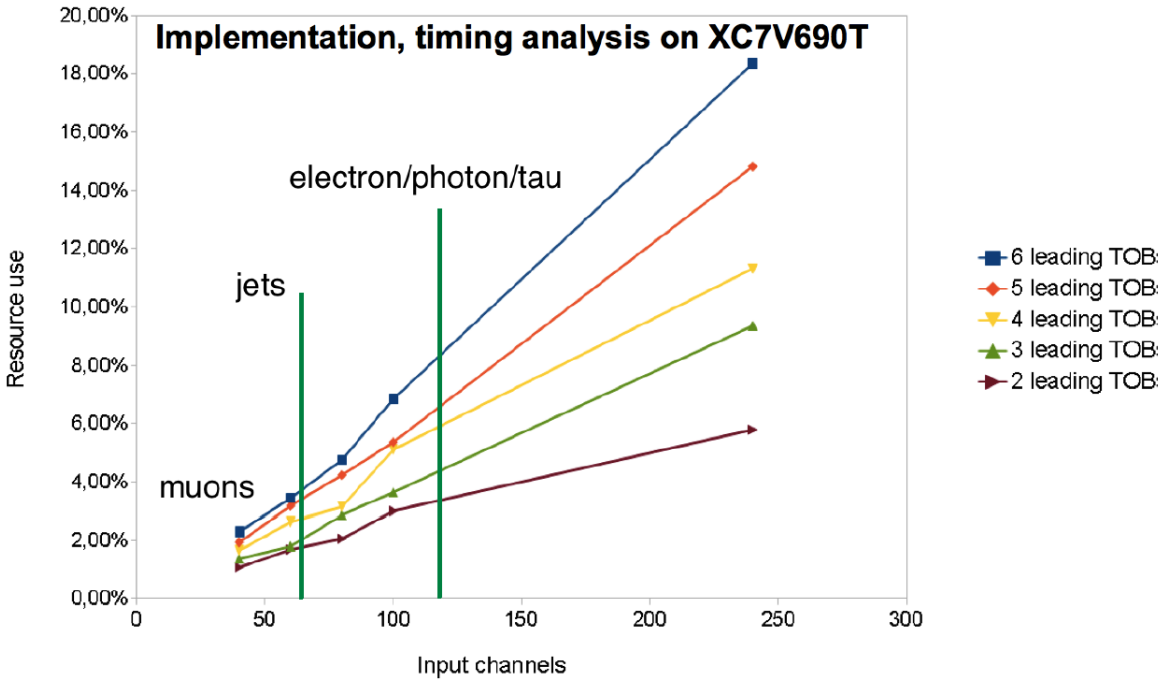


Fig. B.1 Resource usage to sort different Trigger OBjects (TOBs): muon, electron/photon/tau and jet candidates

- $\Delta\eta$ and $\Delta\phi$ (DPHI): takes two input objects with given granularities and ranges, calculates the difference and makes decision.
- ΔR (DR): actual calculation implemented in L1Topo is $\Delta R^2 = \Delta\eta^2 + \Delta\phi^2$
- Box cut (DETA-DPHI): can be used instead of ΔR. This combination uses less resources than the traditional ΔR, and for small angles, is equivalent given the granularity at Level 1.
- LAR: this is a dedicated algorithm to select an EM or Jet objects in the region in the Liquid Argon (LAr) EM calorimeter ($0.0 < |\eta| < 1.4$ and $9/16\pi < \phi < 11/16\pi$) where a new LAr Upgrade Demostrator is located.

Another set of variables refers to the total energy deposited and indicate the "hardness" of the event:

- H_T: the sum of the p_T of all (or 5-6) jets above a certain threshold : Σp_T (jets).
- HC_T: the sum of the p_T of all (or 5-6) central jets above a certain threshold:Σp_T(central jets).
- M_{eff}: the sum of the p_T of all jets in a preferred list plus the missing energy:E_T^{miss} +Σp_T (jets).
- MHT: alternative to a MET trigger, where $MHT_x = \Sigma p_x$ (jet), $MHT_y = \Sigma p_y$(jet) and $MHT = \sqrt{MHT_x^2 + MHT_y^2}$.

Mass requirements

- M: actual usage is M^2 and is defined as: $2E_T^1 \cdot E_T^2(cosh\Delta\eta - cos\Delta\phi)$.
- M_T : actual usage is M_T^2 and is defined as: $2E_T^1 \cdot E_T^{miss}(1 - cos\Delta\phi)$
- M_{CT}: actual usage is M_{CT}^2 and is defined as: $2E_T^1 \cdot E_T^{miss}(1 + cos\Delta\phi)$

Dedicated algorithms

- Matching of the TOBs (MATCH): There is a possibility to compare ET of various TOBs which are at the same geometrical location.
- NOT triggers: NOT triggers are expected to veto on the presence of for example an EMRoI associated with a TAU RoI.

- RATIO triggers: RATIO triggers are expected to select on the ratio of the two object energies or two event quantities.
  example: 10RATIO20-HT-XE would mean $1 \cdot XE < HT < 2 \cdot XE$ while 40RATIO90-HT-2XE would mean $2 \cdot XE^2 < HT < 3 \cdot XE^2$.

- Late particles (LATE): There is a possibility to trigger events using information from the following bunch-crossing, this will be done by dedicated LATE algorithms.

B.2 L1Topo menu request from egamma in 2016/2017 data taking

W T&P triggers in 2016 Current triggers running online (L1 triggers prescaled to 500 Hz, HLT triggers to 0.5 Hz):

- L1_EM12_XS20_HLT_e13_etcut_trkcut_xs15_mt25
- L1_EM12_XS20_HLT_e13_etcut_trkcut_j20_perf_xe15_2dphi05_mt25
- L1_EM12_XS20_HLT_e13_etcut_trkcut_j20_perf_xe15_6dphi05_mt25
- L1_EM12_XS20_HLT_e13_etcut_trkcut_xs15_j20_perf_xe15_2dphi05_mt25
- L1_EM12_XS20_HLT_e13_etcut_trkcut_xs15_j20_perf_xe15_6dphi05_mt25
- L1_EM15_XS30_HLT_e18_etcut_trkcut_xs20_mt35
- L1_EM15_XS30_HLT_e18_etcut_trkcut_j20_perf_xe20_6dphi15_mt35
- L1_EM15_XS30_HLT_e18_etcut_trkcut_xs20_j20_perf_xe20_6dphi15_mt35

Triggers with tighter cuts requested based on studies from the T&P group:

- L1_EM12_XS20_HLT_e13_etcut_trkcut_xs30_xe30_mt35
- L1_EM12_XS20_HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_2dphi05_mt35
- L1_EM12_XS20_HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_2dphi15_mt35
- L1_EM12_XS20_HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_6dphi05_mt35
- L1_EM12_XS20_HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_6dphi15_mt35
- L1_EM15_XS30_HLT_e18_etcut_trkcut_xs30_xe30_mt35

- L1_EM15_XS30_HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_2dphi05_mt35
- L1_EM15_XS30_HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_2dphi15_mt35
- L1_EM15_XS30_HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_6dphi05_mt35
- L1_EM15_XS30_HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_6dphi15_mt35

L1Topo W T&P triggers for 2017

L1_EM12_W-MT25_W-15DPHI-JXE-0_W-15DPHI-EMXE_XS20

1. HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_6dphi05_mt35
2. HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_6dphi15_mt35
3. HLT_e13_etcut_trkcut_j20_perf_xe15_6dphi05_mt25
4. HLT_e13_etcut_trkcut_xs30_xe30_mt35

L1_EM12_W-MT25_W-15DPHI-JXE-0_W-15DPHI-EMXE_W-90RO2-XEHT-0

1. HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_2dphi05_mt35
2. HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_2dphi15_mt35
3. HLT_e13_etcut_trkcut_j20_perf_xe15_6dphi05_mt25
4. HLT_e13_etcut_trkcut_xs30_xe30_mt35

L1_EM15_W-MT35_W-05DPHI-JXE-0_W-05DPHI-EM15XE_XS30

1. HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_6dphi05_mt35
2. HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_6dphi15_mt35
3. HLT_e18_etcut_trkcut_j20_perf_xe20_6dphi15_mt35
4. HLT_e18_etcut_trkcut_xs30_xe30_mt35

L1_EM15_W-MT35_W-250RO2-XEHT-0_W-05DPHI-JXE-0_W-05DPHI-EM15XE

1. HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_2dphi05_mt35
2. HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_2dphi15_mt35
3. HLT_e18_etcut_trkcut_j20_perf_xe20_6dphi15_mt35
4. HLT_e18_etcut_trkcut_xs30_xe30_mt35

B.3 New triggers proposed for 2016-2017 data taking

Hight MET triggers:

- HLT_xe80_tc_lcw_L1XE50
- HLT_xe90_mht_L1XE50
- HLT_xe90_tc_lcw_wEFMu_L1XE50
- HLT_xe90_mht_wEFMu_L1XE50
- HLT_j80_xe80

e13 triggers:

- HLT_e13_etcut_trkcut_L1EM12
- HLT_e13_etcut_trkcut_xs15_mt25
- HLT_e13_etcut_trkcut_xs15_j20_perf_xe15_2dphi05_mt25
- HLT_e13_etcut_trkcut_j20_perf_xe15_2dphi05_mt25
- HLT_e13_etcut_trkcut_xs15_j20_perf_xe15_6dphi05_mt25
- HLT_e13_etcut_trkcut_j20_perf_xe15_6dphi05_mt25
- HLT_e13_etcut_trkcut_xs30_xe30_mt35
- HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_6dphi05_mt35
- HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_6dphi15_mt35
- HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_2dphi05_mt35

- HLT_e13_etcut_trkcut_xs30_j15_perf_xe30_2dphi15_mt35

e18 triggers:

- HLT_e18_etcut_trkcut_L1EM15
- HLT_e18_etcut_trkcut_xs20_mt35
- HLT_e18_etcut_trkcut_xs20_j20_perf_xe20_6dphi15_mt35
- HLT_e18_etcut_trkcut_j20_perf_xe20_6dphi15_mt35
- HLT_e18_etcut_trkcut_xs30_xe30_mt35
- HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_6dphi05_mt35
- HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_6dphi15_mt35
- HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_2dphi05_mt35
- HLT_e18_etcut_trkcut_xs30_j15_perf_xe30_2dphi15_mt35

Hight ET triggers:

- HLT_e60_etcut_trkcut_L1EM24VHIM_j15_perf_xe60_6dphi15_mt35
- HLT_e60_etcut_trkcut_L1EM24VHIM_xe60_mt35
- HLT_e60_etcut_trkcut_L1EM24VHIM_xs30_j15_perf_xe30_6dphi15_mt35
- HLT_e60_etcut_trkcut_L1EM24VHIM_xs30_xe30_mt35
- HLT_e60_etcut_trkcut_j15_perf_xe60_6dphi05_mt35
- HLT_e60_etcut_trkcut_xs30_j15_perf_xe30_6dphi05_mt35
- HLT_e70_lhloose_nod0_L1EM24VHIM_xe70noL1
- HLT_e70_lhloose_nod0_xe70noL1

B.4 Further options analyzed in the $t\bar{t}$H $4l$ channel

B.4.1 Dependence on the BDTG ttH against ttV configuration

The option 1 for configuration of the BDTG was chosen as (see the TMVA tutorial for more information on the definition of the variables [89]) :

- NTrees=500
- MinNodeSize=5%
- Shrinkage=0.10
- nCuts=25
- MaxDepth=2

while for option 2 configuration was chosen:

- NTrees=500
- MinNodeSize=5%
- Shrinkage=0.05
- nCuts=12
- MaxDepth=3

If the changes in the BDTG configuration provide negligible differences in the outputs, then we can validate the stability of the training/testing.

The output of the TMVA is shown in the Figure B.2 for training against ttV background on odd events and Figure B.3 for training against ttV background on even events respectively. On the left is shown the TMVA over-training checks while on the right the background rejection versus signal efficiency (ROC curve). Option 1 config BDTG is on top while Option 2 config BDTG is on the bottom.

The result on the fit on the signal strength by using the baseline selection, the split on the Z-depleted and Z-enriched region and finally the result after applying the BDT on the Z enriched region is shown in Table B.1.

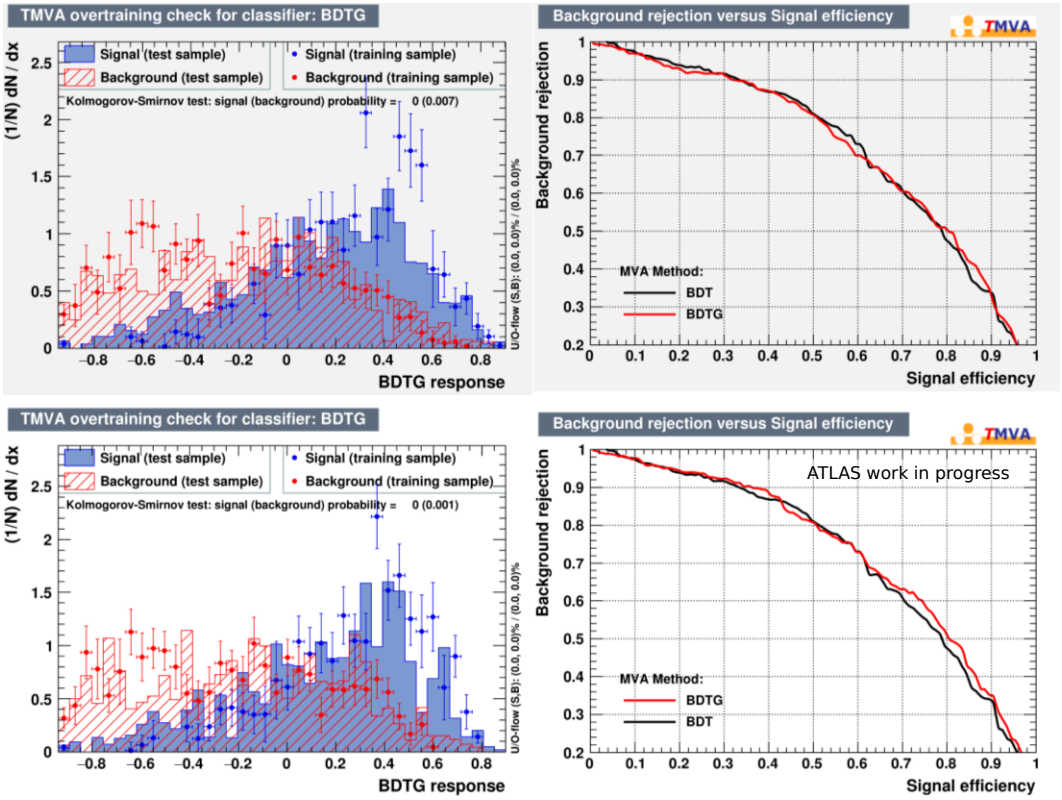


Fig. B.2 TMVA BDTG output ttH against ttV, training on odd events.

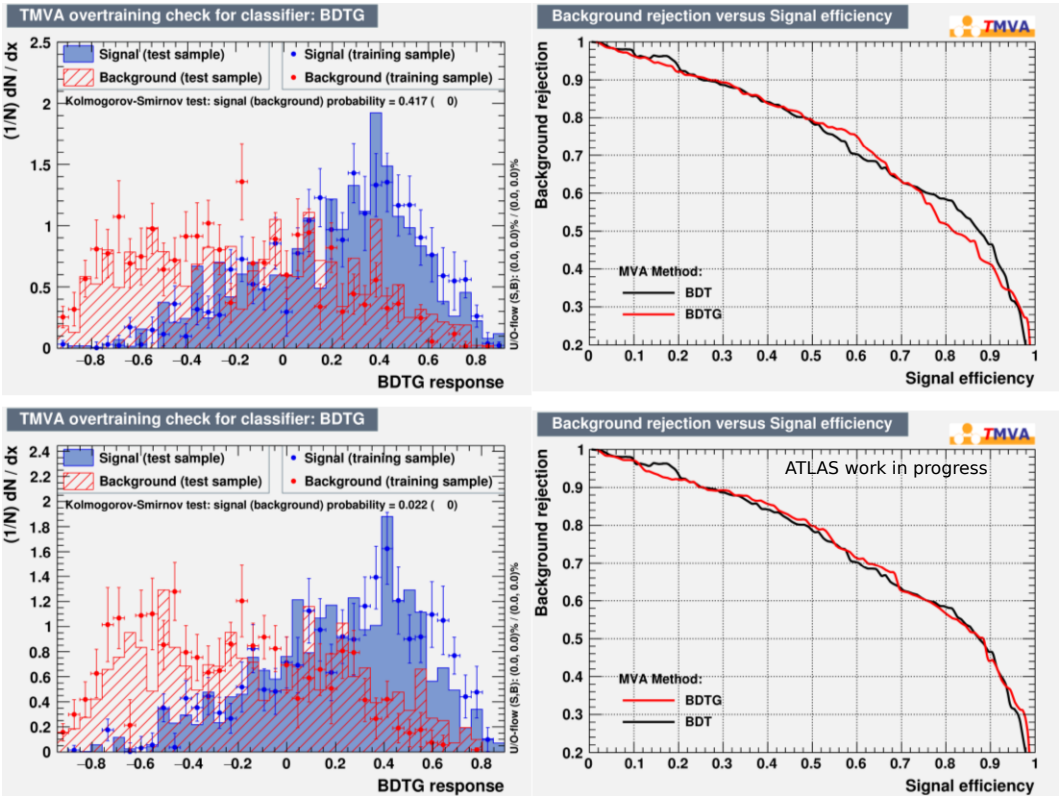


Fig. B.3 TMVA BDTG output ttH against ttV, training on even events.

Region	**Signal Strength**
PLI34 (baseline)	$1^{+1.90}_{-1.39}$
PLI34 + Z-depleted and Z-enriched	$1^{+1.84}_{-1.16}$
Option 1 BDTG on Z enriched (cut on *mva_ttV* >-0.3)	$1^{+1.64}_{-1.12}$
Option 2 BDTG on Z enriched (cut on *mva_ttV* >-0.14)	$1^{+1.64}_{-1.08}$

Table B.1 The Fit on Signal Strength in different SR

The output of the two configurations give a compatible signal strength uncertainty therefore we can assume that the separation power given by one BDTG configuration reached the limit of optimization in the training samples tested.

B.4.2 Optimization studies using TMVA ttH against VV

For the 4L Signal Region, in order to make an optimal use of the existing MC statistics, the training is performed in the following regions:

- ttH: 4L + isoWP + Zveto + njet≥2 + bjet≥1
- VV: 4L + isoWP + Zveto + njet≥2

The output of the TMVA is shown in the Figure B.4, on left is shown the TMVA overtraining checks while on the right the background rejection versus signal efficiency (ROC curve).

The variable distribution is shown in the Figure B.5.

One option is to further add along with the BDTG cut on ttV, a cut on VV (mva>0 for example). However, the improvement by adding this new cut on the BDTG ttH against VV is insignificant for the moment.

B.4.3 Further extension of the 4L in the H → ZZ → ll qq / bb channel

Considering the Higgs decaying modes (see Figure B.6), we plan to capture the 4L when 2L originate from one Z by relaxing the Z veto. To select Z->ll Z->qq, we increase the b-jet/jet multiplicity requirement as shown in Figure B.7.

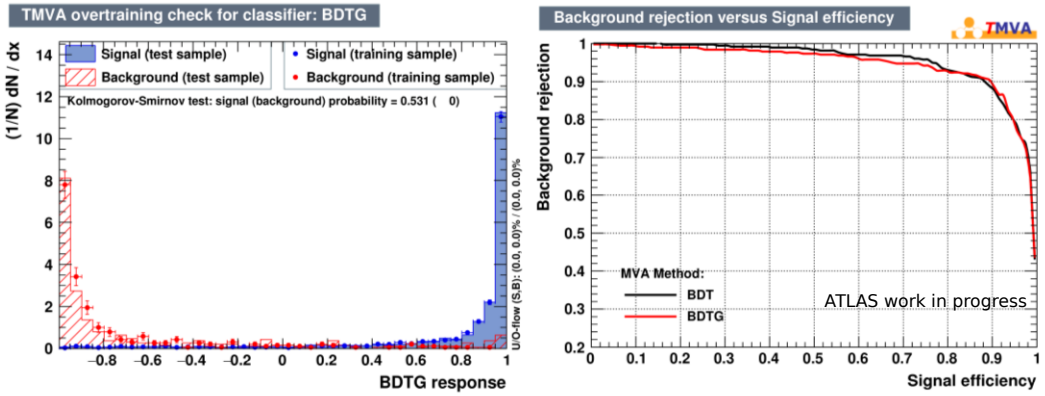


Fig. B.4 TMVA BDTG output ttH against ttV, training on odd events.

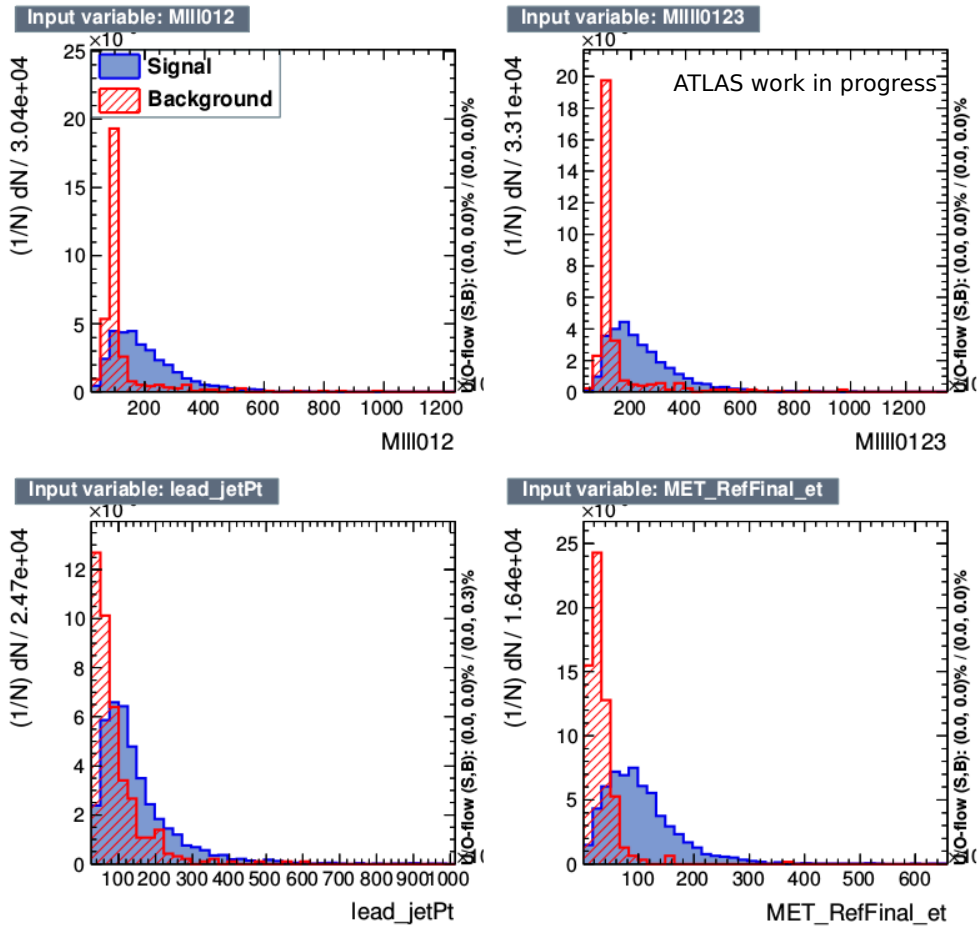


Fig. B.5 Variables used for TMVA training ttH against VV.

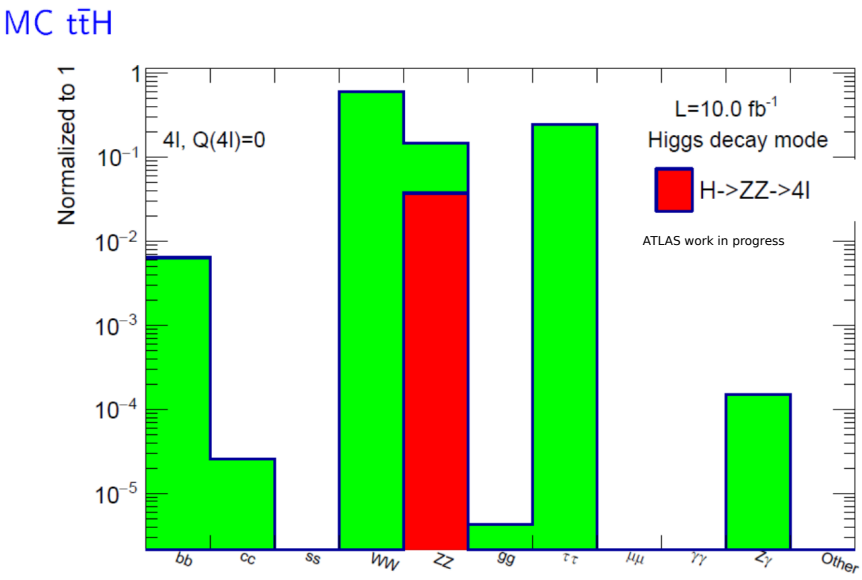


Fig. B.6 The contribution of each Higgs decay mode to the 4l channel. The $H \rightarrow ZZ \rightarrow 4l$ is highlighted in Red.

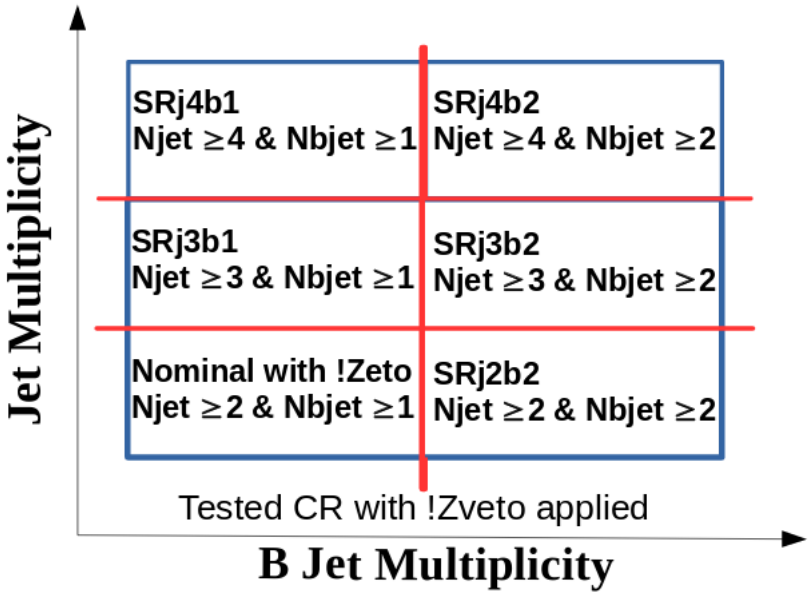


Fig. B.7 Higgs decay modes in ttH MC sample.

In the TMVA training we apply the following cuts:

- ttH_Pythia against ttV: 4L + isoGrad + any of the OSSF pairs within Z mass + njet≥2
- ttH_Pythia against VV: 4L + isoGrad + any of the OSSF pairs within Z mass + njet≥2

The output of the TMVA is shown in the Figure B.8 for training against ttV background (right) and training against VV background (left).

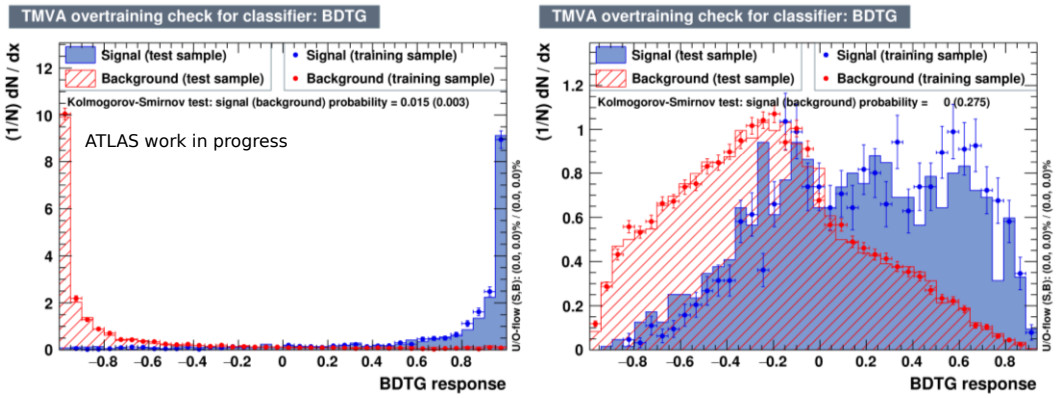


Fig. B.8 TMVA BDTG output ttH against ttV background (right) and training against VV background (left).

Combining the output of the two BDTGs, we search for the point with maximum significance (see Figure B.9) and compare the results with the 4L SR (see Table B.2). In the j3b2l‖j4b1 region (defined as number of jets ≥3 and number of b-jets ≥2 or number of jets ≥4 and number of b-jets ≥1) we observe a 25% loss in ttH with a 70% reduction in the background. The rest of the regions present ∼ 50% reduction in ttH with a ∼ 90% reduction in the background whilst the lower jets multiplicity region may add 0.1 to the signal but are higher in background events.

However, the improvement of the signal strength by adding this ZZ SR is insignificant.

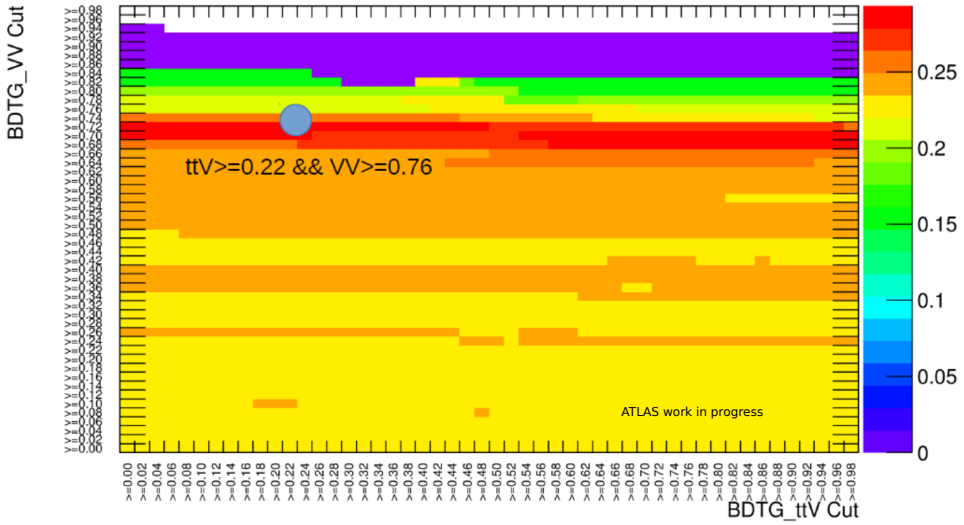


Fig. B.9 2D Significance distribution ttH against ttV and VV BDTG combined. The point of maximum significance obtained by varying simultaneously both BDTG from 0 to 1 was obtained when applying a cut on the TMVA for ttH against ttV >0.22 and on the TMVA for ttH against VV >0.76.

Region	**NLO+ttV**	**Dibosons**	**ttH_Pythia**
Nominal 4L SR (Zveto+njet>=2+nbjet>=1)	2.40±0.12	0.022±0.009	1.21±0.05
j3b2||j4b1	13.3±0.3	0.38±0.04	0.48±0.04
TMVAj3b2||j4b1	4.50±0.16	0.023±0.010	0.29±0.03
j3b1	9.23±0.25	0.17±0.03	0.38±0.04
TMVAj3b1	0.22±0.04	0.004±0.002	0.058±0.015

Table B.2 The TMVA application the H $\rightarrow$ ZZ phase space in the nominal 4*l* SR and the two extensions j3b2||j4b1 region with/without the TMVA cut and j3b1 with/without the TMVA cut .

B.5 Data Driven Fake Estimates

Due to low statistics in the 4l channel, the data-driven estimates need to rely on three lepton analysis.

Inspired by the Run 1 ttH to multilepton final state, one defines the same control regions enriched in $t\bar{t}$ and Z+jets, constructed with 2 tight leptons plus 2 anti-tight leptons (which explicitly fail one or a set of selection criteria). Identical event-level selections to the signal

region are applied, up to and including the Z-veto (i.e., jet-inclusive).

For the tt contribution we consider $t\bar{t}$ + $t\bar{t}$-dilep + ttV and $t\bar{t}$ and Z+jets events while the MC estimate from prompt backgrounds (primarily ZZ and t$\bar{t}$Z) is subtracted from the data yield and the result is taken as the fake contribution. This fake contribution in each CR is extrapolated into the jet-inclusive signal region (comprised of 4 tight leptons plus same cuts to the CRs) by the application of extrapolation factors (per flavor) related to the anti-tight leptons passing the tight lepton selection criteria.

Finally, a second extrapolation is done from the jet-inclusive SR to the full selection by a separate extrapolation factor which is derived from the jet spectra of two lepton events with at least two b-tagged jets and represents the efficiency of $t\bar{t}$ and Z+jets events passing the SR Njet requirement.

The tested SR' is the 4L selection relaxed (with no jet and no 4L mass requirements):

- exactly 4 leptons ($p_T > 10$ GeV)
- total charge to be 0.
- at least one of 4 leptons to be trigger matched
- Z veto: inv. mass of any two OS-SF lepton pairs to be out of Z mass window [81, 101] GeV
- di-lepton resonances veto: inv. mass of any two OS-SF lepton pairs great than 12 GeV

B.5.1 The Tight and Anti-Tight definition

Commonly between the three lepton and four lepton channels, the anti-tight leptons are obtained for both electrons and muons by reversing or relaxing isolation requirements are used to define the objects.

For Tight electrons and muon the selection can be seen in Table B.3 while for anti-leptons the selection is in Table B.4.

With this definitions, we define 3 CR:

1. CR1: 2 tight leptons + 2 anti-tight electrons
2. CR2: 2 tight leptons + 2 anti-tight muons

Selection	**Electron**	**Muon**
pt	> 10 GeV	> 10 GeV
η	$\lvert\eta\rvert < 2.47$, and not $1.37 < \lvert\eta\rvert < 1.52$	$\lvert\eta\rvert < 2.5$
Selection and Isolation WP	pass Tight ID, pass IsolationGradient	pass IsolationGradient and MuQuality $\rightarrow 1$
$\lvert z0sin\theta\rvert$	< 2 mm	< 2 m
$\lvert d0significance\rvert$	< 10	< 10

Table B.3 Tight lepton definition

Selection	**Electron**	**Muon**
ETcone20/pT	$\geq$0.05	$\geq$0.1
pTcone20/pT	< 1.0	< 1.0
Further selection	passing all other tight e selections	passing all other tight μ selections

Table B.4 Anti-Tight lepton definition

3. CR3: 2 tight leptons + 1 anti-tight electron + 1 anti-tight muon

This method aims to deduce the fakes in the signal region using CR 1,2,3, and Fake Factors deduced from the 3L study, as presented in the main body of chapter 4.

B.5.2 Fake Factors in 3L channel

Methodology of this method is the following:

- Extrapolation into the SR: A transfer factor θ for the events in the CR to the SR is evaluated as the ratio of the number of $t\bar{t}$ background events with 3 leptons all tight, over the number of 2 tight + 1 anti-tight of the CR:

$$\theta = \frac{N_{lll}}{N_{ll\not{l}}} \tag{B.1}$$

- Finally, the estimated number of top events in SR is :

$$N_{estimatedtop,SR} = [N_{data,CR} - N_{nontop,MC}] \times \theta \tag{B.2}$$

Common base between 3L and 4L channels are the 3 Tight leptons (with $p_T > 10$ GeV), the sum of the charges of which should be ± 1, trigger matching, di-lepton resonances veto + Z veto. Further we define control samples at high jets (((nJets≥4 + bjets≥1) || (nJets=3 + bjets≥2))) and low (!high) jets multiplicity where the θ factor is calculated and after applied to the 4L SR'.

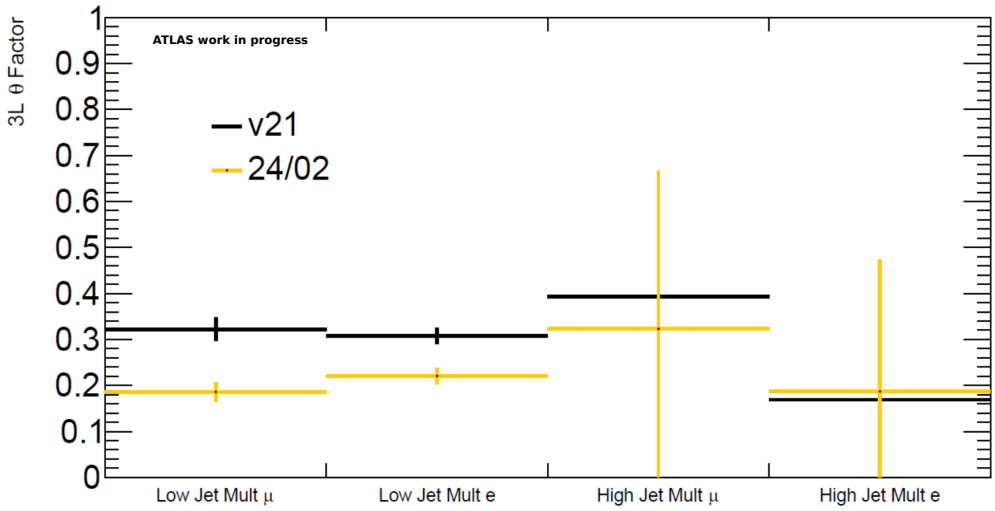


Fig. B.10 Fake Factor in the 3L channel.

The θ factor for the v24 and v21 is illustrated in Figure B.10 for the high and low jets multiplicity per flavor of the leptons.

Starting from the 3L channel, the extrapolation to the 4L channel and the fake estimation is done using the following formula:

$$N_{Fakes} = (N_{data} - N_{prompt})CR_1 \cdot \theta_e^2 + (N_{data} - N_{prompt})CR_2 \cdot \theta_\mu^2 + (N_{data} - N_{prompt})CR_3 \cdot \theta_e \cdot \theta_\mu \quad \text{(B.3)}$$

The result of the fake estimation using data-driven method is listed in Table B.5. The number of fakes using the B.3 formula is found to be around $N_{Fakes} = 0.404$.

Using the tight/anti-tight CR in 4L, the low statistics is problematic, requiring more phase space for fakes (looser selection for leptons).

Sample	**4Tight**	**CR1**	**CR2**	**CR3**
Data	9	3	5	3
prompt	7.98±2.04	0.81±0.27	0.19±0.15	0.00±0.23
ttbar-dilep	0.00±0.00	3.69±0.78	0.60±0.27	2.24±0.58
ttH_Herwig	1.72±0.10	0.18±0.04	0.09±0.03	0.12±0.04

Table B.5 Fake estimation in the 4L channel using data driven method.

Appendix C

Prompt Lepton Tagger

C.1 Introduction

Non-prompt leptons usually have as origin the decays of hadrons containing b or c quarks, therefore, the majority fail standard isolation selection criteria. In the case where the non-prompt lepton passes these criteria, the nearest track jet to the lepton can be investigated and determine its origin. Two possible scenarios are found: first one where the lepton may carry a large fraction of the energy of the bottom or charm jet, leaving the remaining jet not energetic enough for the reconstructed lepton to fail the imposed isolation requirements and secondly the lepton direction of decay is far enough away from the remaining jet components, therefore, the isolation cone is not large enough to capture the energy carried by the jet.

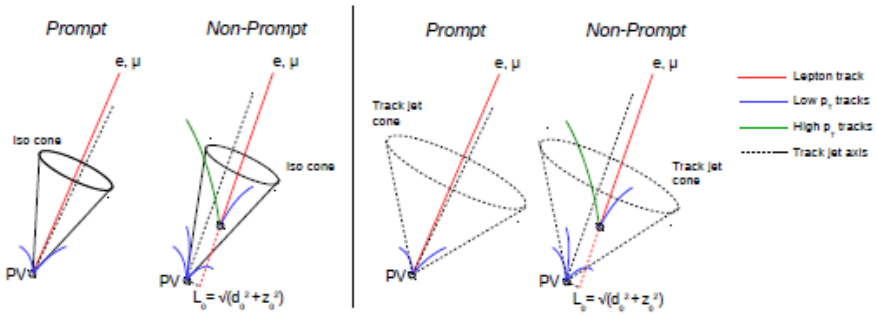


Fig. C.1 A schematic diagram of prompt and non-prompt leptons that pass basic impact parameter and isolation cuts. d0 and z0 correspond to the transverse and longitudinal impact parameters respectively, with L0 corresponding to the secondary vertex decay length from the primary vertex (PV)

Figure C.1 illustrates a schematic representation of the identification and veto the non-prompt leptons [69]. In the case of non-prompt identification, the high p_T tracks (green line) from the secondary decay vertex can be missed by the isolation cone. To get information about these high p_T tracks, one can analyze the reconstructed track jet which nearly always includes the reconstructed lepton track due to a larger radius that collects the tracks than the isolation variable. The most powerful discrimination between prompt and non-prompt cases comes from the secondary vertex within the track jet by using standard b-tagging techniques on track jets containing the "isolated" lepton track.

The object selection is the one used by the Run 2 ATLAS multilepton ttH analysis [68]: tight restrictions on the p_T, isolation and impact parameters of the leptons to reduce the non-prompt lepton background. In order to keep a statistically significant sample of tt MC events for a BDT training algorithm two sets of lepton selection criteria are employed: loose for training the BDT (training selection) and the tight selection criteria (testing selection).

The reconstruction of the track jet candidates with $p_T > 15$ GeV and $|\eta| < 2.5$ is based on the anti-kt [10] clustering algorithm with a ΔR cone of 0.4 to cluster tracks within the jet. The distribution of $\Delta R(lepton, jet)$ can be used as a discriminating variable as shown in Figure C.2.

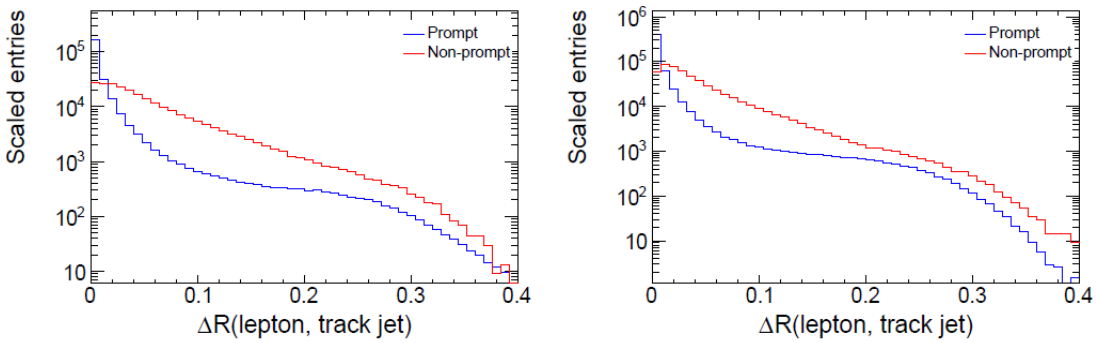


Fig. C.2 ΔR between the nearest track jet to a training electron (left) and training muon (right). The non-prompt lepton distribution is scaled to the number of prompt leptons

C.2 Training Variables

The separating variables used to train the BDT algorithm are listed in Table C.1. They use information of the track jets initiated by b or c quarks, lepton isolation and also impact parameter, secondary vertex reconstruction, and multi-vertex reconstruction algorithms information following the b-tagging procedure.

b-tagging in ATLAS combined mentioned algorithms into a BDT to create a standard b-tagging variable, called MV2.

IP2D is an impact parameter based algorithm that uses the signed transverse impact parameter to distinguish between b, c, and light jets. IP3D is a similar algorithm that includes the use of the longitudinal impact parameter in its training. The logarithm of the likelihood ratio between b and light and c and light is used as the discriminating variables. SV1 is a secondary vertex algorithm that aims to resolve a secondary vertex within the jet. JetFitter is a multi-vertex algorithm that assumes the primary vertex and the resulting b and c vertices lie along a common line, allowing for vertices to be resolved with only a single track coincident. The JetFitterComb variable is a Neural Network combination of the JetFitter and IP3D algorithms. This algorithm is a relic of Run 1, with the variable being superseded by MV2.

There are three main MV2 algorithms; MV2c00, MV2c10, and MV2c20. The algorithms are trained with different admixtures of c-jets as background; MV2c00 is trained with 100% light jet background, MV2c10 with 90% light jet, 10% c-jet background and MV2c20 with 80% light jet, 20% c-jet background.

Variable	Description
N_{track} in the track jet	Number of tracks collected by the track jet
IP2 $\log(P_b/P_{light})$	Log-likelihood ratio between the b and light jet hypotheses with the IP2D algorithm
IP3 $\log(P_b/P_{light})$	Log-likelihood ratio between the b and light jet hypotheses with the IP3D algorithm
$N_{TrkAtVtx}$ SV + JF	Number of tracks used in the secondary vertex found by the SV1 algorithm in addition with the number of tracks from secondary vertices found by the JetFitter algorithm with at least two tracks
$p_T^{lepton}/P_b/P_{track\,jet}$	The ratio of the lepton pT and the track jet pT
ΔR(lepton, track jet)	ΔR between the lepton and the track jet axis
$p_T^{VarCone30}/p_T$	Lepton track isolation, with track collecting radius of $\Delta R < 0.3$
$E_T^{TopoCone30}/p_T$	Lepton calorimeter isolation, with topological cluster collecting radius of $\Delta R < 0.3$

Table C.1 Variables used in the training of PromptLeptonIso.

C.3 Electron Prompt Lepton Tagger

The electron PromptLeptonIso distribution is shown in Figure C.3 and for muon in Figure C.4. Figures C.5 and C.6 compares the efficiency of selecting prompt electrons/muons with the efficiency of rejecting non-prompt electrons (ROC curves) for PromptLeptonIso and the best performing general b-tagging algorithms: JetFitterComb and MV2c10 .

For all prompt efficiency working points, PromptLeptonIso has greater non-prompt rejection power than MV2c10 and JetFitterComb.

Table C.2 and C.3 details the non-prompt rejection efficiencies for the 90% and 95% prompt electron/muon selection efficiency working points.

Algorithm	**90% prompt eff cut**	**Bkg. rejection (%)**	**95% prompt eff cut**	**Bkg. rejection (%)**
JetFitterComb	<3.25	54.1	<2.45	44.9
MV2c10	<0.905	53.2	<0.845	46.2
PromptLeptonIso	<0.6375	78.4	<0.3375	67.9

Table C.2 Non-prompt electron rejections for the 90% and 95% prompt efficiency working points.

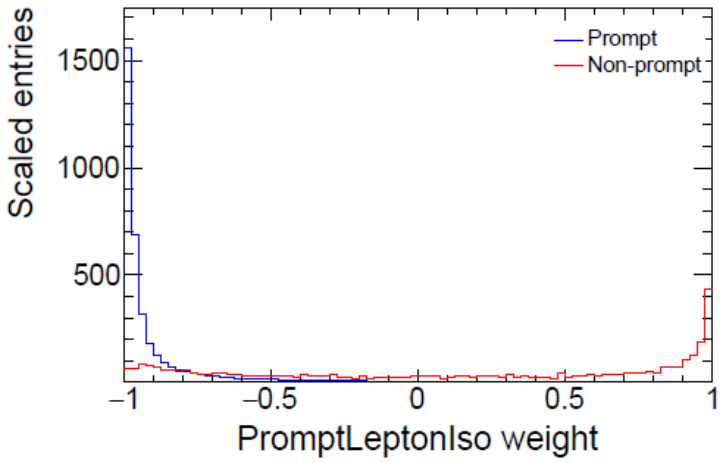


Fig. C.3 Distribution of PromptLeptonIso for prompt and non-prompt testing electrons. The non-prompt distribution is scaled to the number of prompt leptons

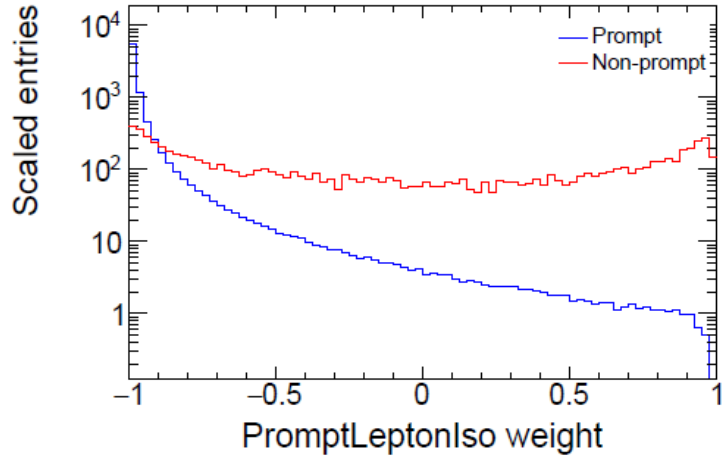


Fig. C.4 Distribution of PromptLeptonIso for prompt and non-prompt testing muons with a log scale. The non-prompt distribution is scaled to the number of prompt leptons

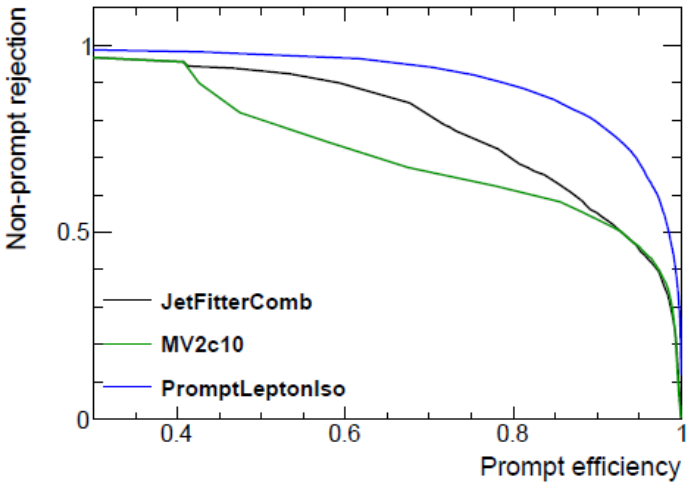


Fig. C.5 ROC curves describing the prompt electron efficiency versus the non-prompt electron rejection for PromptLeptonIso, JetFitterComb and MV2c10.

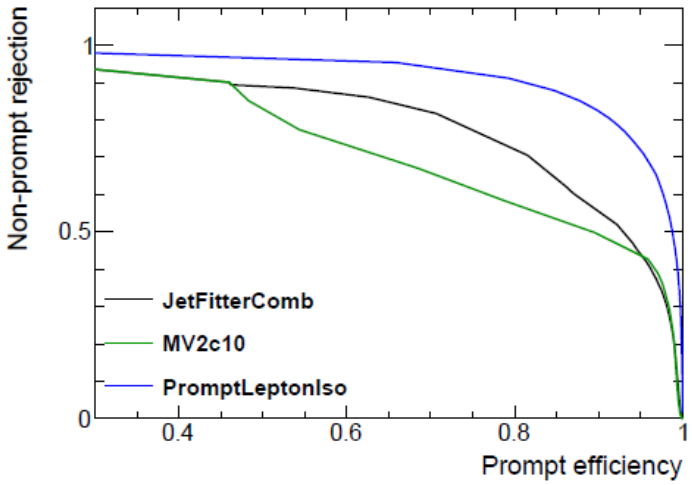


Fig. C.6 ROC curves describing the prompt electron efficiency versus the non-prompt muon rejection for PromptLeptonIso, JetFitterComb and MV2c10.

Algorithm	**90% prompt eff cut**	**Bkg. rejection (%)**	**95% prompt eff cut**	**Bkg. rejection (%)**
JetFitterComb	<4.25	52.0	< 3.95	43.5
MV2c10	<0.925	42.7	<0.925	42.7
PromptLeptonIso	<0.8375	80.6	<0.6875	71.0

Table C.3 Non-prompt muon rejections for the 90% and 95% prompt efficiency working points.

C.4 Conclusion

PromptLeptonIso aims to separate isolated prompt leptons from the non-prompt lepton. It consists of two BDTs, trained separately for electrons and muons, which utilize the b-quark like features of the track jet. In a top - anti-top MC sample, a non-prompt rejection of $\sim 70\%$ is seen with a tightly isolated lepton selection at a 95% prompt efficiency working point.

Appendix D

Nomenclature

The beam direction defines the z-axis, and the x-y plane is the plane transverse to the beam direction. The positive x-axis is defined as pointing from the interaction point to the centre of the LHC ring, and the positive y-axis is pointing upwards. The azimuthal angle Φ is measured around the beam axis, and the polar angle θ is the angle from the beam axis. The pseudorapidity is defined as $\eta = -ln \tan(\theta/2)$. The transverse momentum p_T and the transverse energy E_T ,as well as the missing transverse energy E_T^{miss} and other transverse variables, are defined in the x-y plane unless stated otherwise.

The distance ΔR in the pseudorapidity-azimuthal angle space is defined as $\Delta R = \sqrt{\Delta^2\eta + \Delta^2\Phi}$. Trajectories of charged particles can be described by five helix parameters in an ideal uniform magnetic field. The following helix parametrisation is used in ATLAS, with all quantities measured at the point of closest approach to the nominal beam axis x=0, y=0. Parameters in x-y plane are:

- $1/p_T$: Reciprocal of the transverse momentum with respect to the beam-axis.
- Φ: Azimuthal angle, where $\Phi = \tan p_y/p_x$
- d_0: Transverse impact parameter, defined as the transverse distance to the beam axis at the point of closest approach; signed according to the reconstructed angular momentum of the track about the axis.

Parameters in the R-z plane are:

- $\cot\theta$: Cotangent of the polar angle, where $\cot\theta = p_z/p_T$

- z_0: Longitudinal impact parameter, defined as the z position of the track at the point of closest approach.

L1 trigger Level 1 trigger looks for signatures of interesting processes using measurements of muons, electrons/gamma particles or jets. It can also select events using information about E_T^{miss}. L1 uses only reduced detector information from the calorimeter and from muon detectors designed for the triggering.

The L1 Calorimeter Trigger (L1Calo) identifies high-E_T objects (photons, electrons, jets, taus) as well as missing transverse momentum. Isolation may be applied - then it is checked if the energy deposit around the considered cone is not larger than a given threshold.

The information from the muon detectors; i.e. RPC (Resistive Plate Chambers) in the barrel and TGC (Thin-Gap Chambers) in the end-caps, are used for muon-based trig- gers. There are six independent muon p_T thresholds. Muons should point back to the interaction point and at least one of them should be over given threshold. Muon double counting in the region of overlaping detectors is resolved. The final trigger decision (accept/reject) for every event is made by the Central Trigger Processor (CPT) which combines information from the different subsystems. While the L1 decision is based only on the multiplicity of objects that passed a given requirement, the information about the spacial position of the objects is saved as Regions of Interests (RoI) and passed to the HLT. The output rate from L1 is about 75 kHz and the decision from L1 reaches the front-end electronics after 2.5 μs.

HLT trigger During the L1 2.5μs latency the detector readout is stored in the buffer of each specific detector subsystem. If the event is accepted by the L1, the data is transferred through Readout Links to the HLT. But before, information from L1 subsystems is merged by the RoI builder into a single data structure. Each subsystem provides also information about the RoI. Further selections are applied at the L2, that uses information from all subdetectors (including the tracker) around RoI. The main part of the L2 trigger is a processing farm that consists of 500 processors. The output rate is about 3.5 kHz.

Events accepted by L2 are processed by the Event Filter that consists of 1600 processors. EF incorporates alignment and the full detector readout is available. Events that pass all three levels of the trigger system are saved to the mass storage. The final output rate is about 200 Hz.

References

[1] CERN website: $http://home.web.cern.ch$

[2] LHC website: $http://home.web.cern.ch/topics/large-hadron-collider$

[3] ATLAS Experiment website: $http://atlas.ch/$

[4] CMS Experiment website: $http://cms.web.cern.ch/$

[5] ALICE Experiment website: $http://aliweb.cern.ch/$

[6] LHCb Experiement website: $http://lhcb.web.cern.ch/lhcb/$

[7] TOTEM Experiment website: $http://totem.web.cern.ch/Totem/$

[8] MoEDAL Experiment website: $http://moedal.web.cern.ch/$

[9] *T. Morii, C. S. Lim, S. N. Mukherjee*, **The Physics of the Standard Model and Beyond**, ISBN 981-02-4571-8

[10] *W. N. COTTINGHAM and D . A . GREENWOOD*, University of Bristol, UK, **AN INTRODUCTION TO THE STANDARD MODEL OF PARTICLE PHYSICS**, Published in the United States of America by Cambridge University Press, New York

[11] *P. Puzo*, **"ATLAS calorimetry"**, Nuclear Instruments and Methods in Physics Research A 494 (2002) 340–345

[12] *W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, et al.*, **"Calorimeter clustering algorithms: Description and performance"**, tech. rep., 2008.

[13] *ATLAS Collaboration*, **"Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung"**, Tech. Rep. ATLAS-CONF-2012-047, CERN, Geneva, May, 2012.

[14] http://tmva.sourceforge.net/

[15] $http://home.web.cern.ch/about/physics/supersymmetry$

[16] NASA Science.Astrophysics: $http://science.nasa.gov/astrophysics/focus-areas/what-is-dark-energy/$

[17] *Farida Fassi on behalf of the ATLAS Collaboration*, **"ATLAS Distributed Data Analysis: challenges and performance"**, ATL-SOFT-PROC-2015-001

[18] *Timon Heim, On behalf of the ATLAS Collaboration*, **"Status and performance of the ATLAS Pixel Detector after 3 years of operation"**, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 765, 21 November 2014, Pages 227-231

[19] *I. Riu, On behalf of the ATLAS Collaboration*, **"ATLAS Physics Objects: Status and Performance at 13 TeV**", 8th International Workshop on Top Quark Physics, TOP2015 14-18 September, 2015, Ischia, Italy

[20] *R. N. MOHAPATRA*, **SUPERSYMMETRIC GRAND UNIFICATION: Lectures at TASI97**

[21] *F.J. Hasert et al*, **Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment**, Phys. Lett. 46B (1973) 138

[22] *P. S¨oding, B. Wiik, G. Wolf, Sau Lan Wu*, **The First Evidence for Three-Jet Events in e^+e^- Collisions at PETRA – First Direct Observation of the Gluon** (presented by S.L.Wu), Proc. Int. Europhysics Conf. on High Energy Physics (EPS HEP 95) (Brussels, 1995), p. 3

[23] *G. Arnison et al. (UA1 Collaboration)*, Phys. Lett. B 122, 103

[24] *Harari H. (1977).* **"Three generations of quarks and leptons"** (PDF). In E. van Goeler, Weinstein R. (eds.). Proceedings of the XII Rencontre de Moriond. p. 170. SLAC-PUB-1974.

[25] *The ATLAS Collaboration*, **Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC**, arXiv:1207.7214 [hep-ex]

[26] *The CMS Collaboration*, **Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC**, arXiv:1207.7235 [hep-ex]

[27] https://home.cern/about/experiments/ams

[28] https://home.cern/about/experiments/osqar

[29] http://hilumilhc.web.cern.ch/

[30] https://fcc.web.cern.ch/Pages/default.aspx

[31] http://clic-study.web.cern.ch/

[32] CERN-Brochure-2009-003

[33] *S.M.Bilenky, J.Hosek*, **"Glashow-Weinberg-Salam theory of electroweak interactions and the neutral currents"**, Physics Reports, Volume 90, Issue 2, October 1982, Pages 73-157

[34] http://pdg.lbl.gov/

[35] *Antonella Succurro, on behalf of the ATLAS Collaboration*, **"TIPP 2011 - Technology and Instrumentation for Particle Physics 2011, The ATLAS Tile Hadronic Calorimeter performance in the LHC collision era"**, Physics Procedia 37 (2012) 229 – 237

[36] *L. AperioBella*, **"Status of the ATLAS Liquid Argon calorimeter and its performance after two years of LHC operation"**, Nuclear Instruments and Methods in Physics Research A 718(2013)60–62

[37] *ATLAS Collaboration*, **"Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$=13 TeV"**, Eur. Phys. J. C (2016) 76:292

[38] *S. Weinberg*, **"The making of the Standard Model"**, Eur. Phys. J. C 34, 5–13 (2004)

[39] *Jon Burr*, **How to emulate the trigger**, University of Oxford, October 25, 2016

[40] *The LHCb collaboration*, Measurement of matter–antimatter differences in beauty baryon decays, Nature Physics volume 13, pages 391–396 (2017) doi:10.1038/nphys4021

[41] *MARK THOMSON*, **Modern Particle Physics**, Published in the United States of America by Cambridge University Press, New York, ISBN 978-1-107-03426-6 Hardback

[42] *ATLAS Collaboration*, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Physics Letters B, Volume 716, Issue 1, 17 September 2012, Pages 1-29

[43] *CMS Collaboration*, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, Physics Letters B, Volume 716, Issue 1, 17 September 2012, Pages 30-61

[44] *John Ellis*, Physics Beyond the Standard Model, 10.1016/j.nuclphysa.2009.05.034

[45] *BarbaraClerbaux, ATLAS and CMS Collaborations*,Searches for physics beyond the Standard Model at ATLAS and CMS, Nuclear Physics B - Proceedings Supplements, Volume 245, December 2013, Pages 86-93

[46] Higgs discovery at CERN: $http://home.web.cern.ch/topics/higgs-boson$

[47] *Michele Maggiore*, **A Modern Introduction to Quantum Field Theory**, Published in the United States by Oxford University Press Inc., New York, ISBN 0 19 852073 5

[48] *Paul Langacker*, **Introduction to the Standard Model and Electroweak Physics**, arXiv:0901.0241

[49] *Giovanni Costa, Gianluigi Fogli*, **Symmetries and Group Theory in Particle Physics**, DOI $10.1007/978-3-642-15482-9$

[50] *Michael Riordan*, **The Discovery of Quarks**, SLAC-PUB-5724, April 1992

[51] *Luis Anchordoqui and Francis Halzen*, **Lessons in Particle Physics**, http://arxiv.org/abs/0906.1271v2

[52] *ALESSANDRO BETTINI*, **INTRODUCTION TO ELEMENTARY PARTICLE PHYSICS**, Published in the United States of America by Cambridge University Press, New York, ISBN-13 978-0-521-88021-3

[53] *A. La Rosa,on behalf of ATLAS collaboration*, **The ATLAS Insertable B-Layer: from construction to operation**, arXiv:1610.01994v3 [physics.ins-det] 22 Nov 2016

[54] *Arantxa Ruiz Martinez, on behalf of the ATLAS Collaboration*, **The Run-2 ATLAS Trigger System**, ATL-DAQ-PROC-2016-003

[55] $http://lhc-commissioning.web.cern.ch/lhc-commissioning/2015/2015-commissioning-outline.htm$

[56] *Dr Antonio Limosani*, **Run 2 of the Large Hadron Collider: Upgrades and prospects for new discoveries**, AIP 2014 Congress, December 7-12, ANU, Canberra

[57] *Beate Heinemann*, **Future of LHC**, PiTP, July 2013

[58] *C Kilby on behalf of the ATLAS Collaboration*, The design and performance of the ATLAS Inner Detector trigger for Run 2 LHC Collisions at $\sqrt{s}$ = 13 TeV, ATL-DAQ-PROC-2016-004, 29/02/2016

[59] *ATLAS Collaboration*, ATLAS Detector and Physics Performance, ATLAS TDR 14, CERN/LHCC 99-14

[60] Worldwide LHC Computing Grid: http://wlcg.web.cern.ch/

[61] http://lutece.fnal.gov/Papers/PhysNews95.html

[62] http://www-d0.fnal.gov/Run2Physics/top/

[63] "Top quark physics at the LHC", *Werner Bernreuther*, arXiv:0805.1333v1 [hep-ph] 9 May 2008

[64] *FRANK-PETER SCHILLING*, TOP QUARK PHYSICS AT THE LHC: A REVIEW OF THE FIRST TWO YEARS, International Journal of Modern Physics A, arXiv:1206.4484v2 [hep-ex] 17 Apr 2013

[65] *S. Biswas, R. Frederix, E. Gabrielli, B. Mele*, Enhancing the $t\bar{t}$H signal through top-quark spin polarization effects at the LHC, JHEP07(2014)020

[66] *R. Alkofer, J. Greensite*, Quark Confinement: The Hard Problem of Hadron Physics, arXiv:hep-ph/0610365

[67] *ATLAS Collaboration*, Search for the associated production of the higgs boson with a top quark pair in multilepton final states with the atlas detector, Physics Letters B, 749:519–541, 2015.

[68] *ATLAS Collaboration*, Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector, CERN-EP-2017-281; arXiv:1712.08891.-Geneva : CERN, 2018-04-09 - 44 p. - Published in : Phys. Rev. D 97 (2018) 072003

[69] *ATLAS Collaboration*, Tagging non-prompt electrons and muons, Rustem Ospanov, Rhys Roberts, Terry Wyatt.

[70] *ATLAS, CMS Collaborations*, **"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"**, arXiv:1606.02266

[71] CN division Application Software group., GEANT Detector Description and Simulation Tool., CERN Program Library Long Writeup W5013, 1993.

[72] *Ulrik Egede*, **"The search for a standard model Higgs at the LHC and electron identification using transition radiation in the ATLAS tracker"**, Lund University,

[73] *Shreya Saha*, Event reconstruction of tt and ttH using the Kinematic Likelihood Fitter in final states with two same-sign electric charge leptons, CERN-THESIS-2017-149

[74] *A. Arhrib et al.*, The Higgs Potential in the Type II Seesaw Model, Phys. Rev. D84 (2011) 095005, arXiv: 1105.1925 [hep-ph].

[75] *P. Fileviez Perez et al.*, Neutrino Masses and the CERN LHC: Testing Type II Seesaw, Phys. Rev. D78 (2008) 015018, arXiv: 0805.3536 [hep-ph].

[76] *Stephen F. King*, Minimal predictive see-saw model with normal neutrino mass hierarchy, arXiv:1304.6264

[77] *A. Belyaev, N. D. Christensen and A. Pukhov*, **"CalcHEP 3.4 for collider physics within and beyond the Standard Model"**, Comput. Phys. Commun. 184 (2013) 1729, arXiv: 1207.6082

[78] *A. Buckley et al.*, **"LHAPDF6: parton density access in the LHC precision era"**, The European Physical Journal C 75 (2015) 132, issn: 1434-6052

[79] *ATLAS Collaboration*, **"Search for supersymmetry in final states with two same-sign or three leptons and jets using 36 fb of s=13 TeV pp collision data with the ATLAS detector"**, tech. rep. ATLAS-CONF-2017-030, CERN, 2017, url: http://cds.cern.ch/record/2262920.

[80] *The CMS Collaboration*, **"Observation of ttH production"**, Physical Review Letters as doi:10.1103/PhysRevLett.120.231801.

[81] *C.T. Potter*, **"ATLAS Searches for Beyond the Standard Model Higgs Bosons"**, Department of Physics, University of Oregon, arXiv:1310.0515v1 [hep-ex] 1 Oct 2013

[82] *Stephen P. Martin*. **A Supersymmetry primer**. 1997.

[83] *B. W. Lee and S. Weinberg*, **Cosmological Lower Bound on Heavy Neutrino Masses**, Phys. Rev. Lett. 39 (1977) 165.

[84] *Steve Myers*, **The engineering needed for particle physics**,Published 16 July 2012.DOI: 10.1098/rsta.2011.0053

[85] *P.J. Bryant* ,**A BRIEF HISTORY AND REVIEW OF ACCELERATORS**, CERN, Geneva, Switzerland

[86] https://twiki.cern.ch/twiki/bin/viewauth/Atlas/AtlasHltMetSlice.

[87] *ATLAS Collaboration*, **"Search for doubly charged scalar bosons decaying into same-sign W boson pairs with the ATLAS detector"**, Eur. Phys. J. C, arXiv:1808.01899

[88] *ATLAS Collaboration*, **"Search for doubly charged Higgs boson production in multi-lepton final states with the ATLAS detector using proton–proton collisions at $\sqrt{s}$ = 13 TeV"**, Eur. Phys. J. C78 (2018) 199 DOI: 10.1140/EPJC/S10052-018-5661-Z

[89] https://root.cern.ch/doc/v608/group_tutorial_tmva.html

List of articles

Papers

HIGG-2017-02: HTop ttH multileptons 2015+2016 Analysis/Editor:Optimisation and fake estimates in 4L channel.

HIGG-2016-09: Search for doubly-charged scalar bosons decaying to same sign W-bosons at LHC Analysis/Editor:Optimisation and fake estimates in 4L channel.

PERF-2017-01: To be submitted to: EPJC: Electron reconstruction and identification in the ATLAS Experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s}$ =13 TeV Analysis/Editor: Electron identification

Conference Notes

ATLAS-CONF-2016-058: HTop ttH multileptons ICHEP 2016 ; Analysis/Editor:Optimisation of 4l channel and fake estimates.

ATLAS collaboration

- **Electron efficiency measurements with the ATLAS detector using 2012 LHC proton-proton collision data**, *ATLAS Collaboration*, CERN-EP-2016-262; arXiv:1612.01456.- Geneva : CERN, 2017-03-27 - 64 p. - Published in : Eur. Phys. J. C 77 (2017) 195, Article from SCOAP3: PDF; Preprint: PDF; Springer Open Access article: PDF;
- **Performance of the ATLAS Trigger System in 2015**, *ATLAS Collaboration*, CERN-EP-2016-241; arXiv:1611.09661.- Geneva : CERN, 2017-05-18 - 76 p. - Published in : Eur. Phys. J. C 77 (2017) 317, Article from SCOAP3: PDF; Preprint: PDF; Springer Open Access article: PDF;

- **Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data**, *ATLAS Collaboration*, Eur.Phys.J. C77 (2017) no.3, 195,2017

- **Performance of the ATLAS Trigger System in 2015**, *ATLAS Collaboration*, Eur.Phys.J. C77 (2017) no.5, 317,2017

- **Search for doubly charged Higgs boson production in multi-lepton final states with the ATLAS detector using proton–proton collisions at $\sqrt{s}$=13TeV**, *ATLAS Collaboration*, CERN-EP-2017-198; arXiv:1710.09748.- Geneva : CERN, 2018-03-10 - 52 p. - Published in : Eur. Phys. J. C 78 (2018) 199, Article from SCOAP3

- **Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector**, *ATLAS Collaboration*, CERN-EP-2018-138 ; arXiv:1806.00425 ; CERN-EP-2018-138. - 2018. - 33 p.

- **Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector**, *ATLAS Collaboration*, CERN-EP-2017-281; arXiv:1712.08891.- Geneva : CERN, 2018-04-09 - 44 p. - Published in : Phys. Rev. D 97 (2018) 072003 Article from SCOAP3

- **Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector**, *ATLAS Collaboration*, CERN-EP-2017-291; arXiv:1712.08895.- Geneva : CERN, 2018-05-01 - 44 p. - Published in : 10.1103/PhysRevD.97.072016 Article from SCOAP3

- **Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s}$=8 TeV with the ATLAS detector**, *ATLAS Collaboration*, JHEP 1605 (2016) 160,2016

- **Search for the Standard Model Higgs boson produced by vector-boson fusion in 8 TeV pp collisions and decaying to bottom quarks with the ATLAS detector**, *ATLAS Collaboration*, JHEP 1611 (2016) 112,2016

- **Search for the Standard Model Higgs boson produced in association with a vector boson and decaying into a tau pair in pp collisions at $\sqrt{s}$=8TeV with the ATLAS detector**, *ATLAS Collaboration*, Phys.Rev. D93 (2016) no.9, 092005,2016

Presentations and posters at international conferences:

- The European Physical Society Conference on High Energy Physics (EPS-HEP), Poster: "Search for the associated production of a Higgs boson and a top quark pair in multilepton (2 leptons, no hadronically-decaying τ lepton candidates and 4 leptons) final states with the ATLAS detector."
- Thirteenth Conference on the Intersections of Particle and Nuclear Physics (CIPANP 2018), the Hyatt Regency Indian Wells from May 28 - June 3, 2018, Presentation: “Searches for nonstandard-model Higgs bosons”

Local conferences and workshops:

- Horia Hulubei National Institute of Physics and Nuclear Engineering - IFIN HH celebration days, 17 December 2015
- Horia Hulubei National Institute of Physics and Nuclear Engineering - IFIN HH celebration days, 20-21 December 2016
- Horia Hulubei National Institute of Physics and Nuclear Engineering - IFIN HH celebration days, 19.12-20.12.2017, Study of the Higgs production in association with tt quarks. Search for doubly-charged Higgs boson production in multi-lepton final states,
- ATLAS Romania-annual meeting, 4-5 July 2016, “Transylvania University of Brasov”, Romania
- ATLAS Romania-annual meeting, 17-18 March 2017, “Alexandru Ioan Cuza” University, Iasi, Romania
- Journées de Rencontre des Jeunes Chercheurs 2016 , 4-10 December 2016, Anjou , France
- ATLAS Physics France workshop (PAF), 29-31 March, Vogue, France, presentation

- Summer School and Workshop on the Standard Model and Beyond September 1 - 11, 2015, Corfu Summer Institute
- ATLAS Software Tutorial, 14-18 September 2015, CERN, Switzerland
- ATLAS e/gamma workshop, 9-13 November 2015, Annecy, France
- Poster at ATLAS week, 22-26 February 2016, CERN, Switzerland
- Presentation at ATLAS e/gamma workshop, 7-10 November 2016, Thessaloniki, at the University's Research Dissemination Center,Greece
- The 2016 European School of High Energy Physics, 15-28 June 2016, Skeikampen, Norway
- Htop workshop, 3-5 October 2016, CERN
- Journées de Rencontre des Jeunes Chercheurs 2016, 4-10 December 2016, Anjou, France
- Htop workshop, 22-24 March 2017, Marseille, France
- Journées Scientifiques de l'Ecole Doctorale Physique et Sciences de la Matière 2-3 May 2016, ST-CHARLES, Marseille, France
- Poster at Journées Scientifiques de l'Ecole Doctorale Physique et Sciences de la Matière 2-3 May 2017, ST-CHARLES, Marseille, France
- Presentation entitled "ttHML: 4l" at workshop PAF2017 (Physique ATLAS France), 29-31 March, Domaine Lou Capitelle, Vogüé, France
- Htop workshop, 22-24 March 2017, Marseille, France
- Presentation entitled "ttHML: 4l" at workshop PAF2017 (Physique ATLAS France), 29-31 March, Domaine Lou Capitelle, Vogüé, France
- The European Physical Society Conference on High Energy Physics (EPSHEP), Venice, Italy on 5-12 July, https://indico.cern.ch/event/466934/overview
- Egamma workshop, 6-9 November 2017, DESY Hamburg, Electron reco and ID efficiencies in Rel. 21,
- Htop workshop, 4-6 April 2018, CERN

- Outreach communications ; 6-10 March ATLAS Masterclass, CPPM, France,
- Lecturer for Physics for both infinities : L'École d'été France Excellence 2017 (https://indico.in2p3.fr/event/14866/page/683-lecturers-list)
- Lecturer for Physics for both infinities : L'École d'été France Excellence 2018 (https://indico.in2p3.fr/event/17629/timetable/?

REZUMAT

Lucrarea stiintifică prezentată contine datele provenind de la detectorul Large Hadron Collider (LHC) prin coliziunea proton-proton produsa la energia de 13 TeV în centrul masei, date înregistrate între 2015 si 2017 de către detectorul ATLAS.

După o scurtă introducere la programul de fizică de la LHC, primul capitol oferă o prezentare globală a contextului teoretic. Ulterior, o descriere completă a detectorului ATLAS si a contextului experimental este prezentata, cu un accent deosebit pe diferitele obiecte fizice utilizate în această cercetare.

Al treilea capitol prezinta performanta calorimetrului electromagnetic, si anume studiul eficientei de identificare a electronilor folosind dezintegrarea W în electroni si neutrini.

În următoarele capitole voi prezenta nucleul acestei lucrări si anume studiul productiei de bosoni Higgs asociată cu perechea top quark, asa-numita productie ttH, în cadrul modelului standard si dincolo de acesta. Observarea productiei ttH cu detectorul ATLAS anuntata recent în 2018 reprezintă o mare realizare, semnificativă pentru domeniul fizicii la energii înalte.

Pentru a lărgi spectrul dincolo de SM, acest canal 4L a fost studiat, de asemenea, în contextul căutării unui boson dublu Higgs ($H^{\pm\pm}$) ce de dezintegreaza în $W^{\pm}$ $W^{\pm}$.

În concluzii, voi face un rezumat al analizelor prezentate si voi prezenta o scurta perspectiva în cadrul Run 2 pentru analizele mentionate în cadrul Modelului Standard precum si în contextului unei fizici noi.

THESIS SUMMARY

The scientific work presented in this thesis is based on proton proton collision produced, at 13 TeV in the center of mass, created by the Large Hadron Collider (LHC) and recorded between 2015 and 2017 by the ATLAS detector.

After a short introduction on the physics program at LHC, the first chapter gives a comprehensive overview of the theoretical context and motivation. Then a complete description of the ATLAS detector and more generally the experimental setup is given with a highlight on the different physics objects used in this research.

The third chapter is dealing with a measure of the performances of the ATLAS electromagnetic calorimeter, namely the study of the electron identification efficiencies using W decaying into electrons and neutrinos events.

In the following chapters, the core of this thesis work, the search and study of the associated Higgs boson production with top quark pair, so-called ttH production, within the Standard Model schema and beyond is treated. The observation for ttH production with the ATLAS detector recently announced in 2018 represents a significant milestone for the High-Energy Physics field.

In order to broaden the spectrum of this ttH analysis beyond the SM, this 4L channel final states was also studied in the context of the search for doubly charged Higgs boson decaying into $W^{\pm}W^{\pm}$, $H^{\pm\pm} \rightarrow W^{\pm}W^{\pm}$.

In conclusions, future prospects in Standard Model (SM) as well as Beyond SM (BSM) scientific results that could be reached using the full Run1 and Run2 combined available statistics are given.

RÉSUMÉ

Le travail scientifique présenté dans cette thèse s'appuie sur des collisions proton proton, à 13 TeV dans le centre de masse, produite par le grand collisionneur de Hadron (LHC) et enregistrées entre 2015 et 2017 par le détecteur ATLAS.

Après une brève introduction sur le programme de physique du LHC, le premier chapitre donne un aperçu complet du contexte et des motivations théorique. Ensuite une description complète du détecteur ATLAS et plus généralement de l'ensemble expérimental est donnée en détaillant les différents objets physiques utilisés dans cette recherche.

Le troisième chapitre traite d'une des mesures des performances du calorimètre électromagnétique ATLAS, à savoir l'étude des efficacités identification électronique à l'aide de W se désintégrant en électrons et neutrinos.

Dans les chapitres suivants, le coeur de cette thèse, la recherche et l'étude de la production associée de boson de Higgs avec une paire de quark top, appelée production ttH, est traitée dans le schéma du modèle Standard et au-delà. L'observation de la production ttH avec le détecteur ATLAS a été récemment annoncée en 2018 et représente une étape importante pour le domaine de la physique des hautes énergies.

Afin d'élargir le spectre de cette analyse tth au-delà du SM, cet état final du canal 4L a aussi été étudié dans le cadre de la recherche de bosons de Higgs doublement chargés se désintégrant en $W^{\pm}W^{\pm}$, $H^{\pm\pm} \rightarrow W^{\pm}W^{\pm}$.

En conclusions, les perspectives sur les futurs résultats scientifiques dans le modèle Standard (SM) ainsi que au-delà de SM (BSM) qui pourraient être atteint en utilisant les statistiques complètes des Run1 et Run2 combinés sont donnés.