UAB

Universitat Autònoma de Barcelona

AIX-MARSEILLE UNIVERSITÉ
ÉCOLE DOCTORALE 352
FACULTÉ DES SCIENCES
CENTRE DE PHYSIQUE DES PARTICULES DE
MARSEILLE

UNIVERSIDAD AUTÓNOMA DE BARCELONA
PROGRAMA DE DOCTORADO EN FÍSICA
DEPARTAMENTO DE FÍSICA
FACULTAD DE CIENCIAS
INSTITUTO DE FÍSICA DE ALTAS ENERGÍAS

Ph.D. thesis

Discipline:
PHYSIQUE ET SCIENCES DE LA MATIÈRE

Specialization:
Physique des Particules et Astroparticules

Yulia RODINA

Search for the Higgs boson decaying to a pair of bottom quarks and produced in association with a pair of top quarks in the LHC Run 2 data with the ATLAS detector using a likelihood technique

Thesis director in AMU:
Laurent VACAVANT

Thesis director in UAB:
Aurelio JUSTE ROZAS

Thesis tutor in UAB:
Maria Pilar CASADO LECHUGA

Defended 23 November 2017 in front of the thesis committee:

Henri	BACHACOU	Rapporteur
Florencia	CANELLI	Rapporteur
Daniel	BLOCH	Examinateur
Stefan	GUINDON	Examinateur
Aurelio	JUSTE ROZAS	Directeur de thèse
Laurent	VACAVANT	Directeur de thèse

Contents

Abstract

Following the discovery of the Higgs boson by the ATLAS and CMS Collaborations at the Large Hadron Collider (LHC) in 2012, the attention has been focussed on studying the properties of the newly discovered particle to test the predictions of the Standard Model (SM). An object of particular interest is the top quark Yukawa coupling - the coupling of the Higgs boson to the top quark, which is predicted to be close to unity in the SM and at the same time very sensitive to the possible effects of new physics beyond the SM. The production of the Higgs boson in association with a pair of top quarks, $t\bar{t}H$, is the mode which gives direct access to the top quark Yukawa coupling. The decay of the Higgs boson into a pair of b-quarks, $H \to b\bar{b}$, is dominant in the SM (its branching ratio is approximately 58%). This decay channel also allows measuring the b-quark Yukawa coupling - the second largest coupling of the Higgs boson to a fermion in the SM.

In this dissertation the search for $t\bar{t}H$ ($H \to b\bar{b}$) in the single-lepton channel, resulting from the semileptonic decay of the $t\bar{t}$ system, is presented. The analysis is based on 36.1 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV recorded with the ATLAS detector in 2015 and 2016. The study is performed using a likelihood-based method that exploits kinematic properties of the selected events to separate the signal from the background, which is dominated by $t\bar{t}$ produced in association with additional jets. This search relies on the high multiplicity of jets originating from b-quarks (b-jets), so identification of these jets (b-tagging) is crucial. A study on the optimisation of b-jet identification algorithms in ATLAS is also presented in this dissertation. The ratio of the measured $t\bar{t}H$ cross-section to the SM expectation is found to be $\mu = 0.88^{+0.64}_{-0.61}$, assuming a Higgs boson mass of 125 GeV. This result is consistent with both the background-only hypothesis and the $t\bar{t}H$ SM prediction.

Résumé

Suite à la découverte du boson de Higgs au Large Hadron Collider (LHC) par ATLAS et CMS collaborations en 2012, l'attention a porté sur l'étude des propriétés de la nouvelle particule découverte pour tester les prédictions du modèle standard (MS). Un objet d'intérêt particulier est le couplage de quark Yukawa top - le couplage du boson de Higgs au quark top, qui devrait être proche de l'unitè dans le SM et en même temps très sensible aux effets possibles de nouvelle physique au-delà du MS. La production du boson de Higgs en association avec une paire de quarks top, $t\bar{t}H$, est le canal qui donne un accès direct à au couplage de quark Yukawa top.

La désintégration du boson de Higgs à une paire de quarks b, $H \rightarrow b\bar{b}$, domine dans le MS: sa rapport de branchement est d'environ 58%. Ce canal de la désintégration permet également de mesurer le couplage b-quark Yukawa - le deuxième plus grand couplage de Higgs boson à un fermion dans le MS.

Dans cette dissertation, la recherche de $t\bar{t}H$ ($H \rightarrow b\bar{b}$) dans le canal single-lepton, résultant de la désintégration semi-lepton du système $t\bar{t}$, est présentée. L'analyse est basée sur 36.1 fb^{-1} de pp collision data à $\sqrt{s} = 13$ TeV enregistré avec le détecteur ATLAS en 2015 et 2016. L'étude est réalisée avec une méthode de vraisemblance, qui exploite les propriétés cinématiques des événements sélectionnés pour séparer le signal de le fond, qui est dominé par la paire de quarks supérieurs $t\bar{t}$ produit en association avec des jets supplémentaires. Cette recherche repose sur une grande multiplicité de jets originaires de quarks b (b-jets), pour cette raison l'identification de ces jets (b-tagging) est cruciale. Une étude sur l'optimisation des algorithmes d'identification b-jet dans ATLAS est également présentée dans cette dissertation. Le ratio de la section efficace mesurée de $t\bar{t}H$ à la prévision de MS est $\mu = 0.88^{+0.64}_{-0.61}$, en supposant une masse du boson de Higgs de 125 GeV. Ce résultat est cohérent avec l'hypothèse de fond seulement ainsi que la prédiction du MS $t\bar{t}H$.

Resumen

Tras el descubrimiento del bosón de Higgs por las Colaboraciones ATLAS y CMS en el Gran Colisionador de Hadrones (LHC, por sus siglas en inglés) en 2012, la atención se ha centrado en estudiar las propiedades de la partícula recientemente descubierta para probar las predicciones del Modelo Estándar (SM, por sus siglas en inglés). Un objeto de particular interés es el acoplamiento de Yukawa del quark top - el acoplamiento del bosón de Higgs al quark top, que se prevé que tenga un valor cercano a la unidad en el SM y al mismo tiempo es muy sensible a los posibles efectos de nueva física más allá del SM. La producción del bosón de Higgs en asociación con una pareja de quarks top, $t\bar{t}H$, es el modo que permite medir directamente al acoplamiento de Yukawa del quark top. La desintegración del bosón de Higgs en una pareja de quarks b, $H \to b\bar{b}$, es dominante en el SM (ocurre aproximadamente el 58% de las veces). Este canal de desintegración también permite medir el acoplamiento de Yukawa del quark b - el segundo mayor acoplamiento del bosón de Higgs a un fermión en el SM.

En esta tesis se presenta la búsqueda del proceso $t\bar{t}H$ ($H \to b\bar{b}$) en sucesos con un sólo leptón en el estado final, resultante de la desintegración semileptónica del sistema $t\bar{t}$. El análisis se basa en 36.1 fb^{-1} de datos de colisiones protón-protón a una energía del centro de masas de $\sqrt{s} = 13$ TeV registrados con el detector ATLAS en 2015 y 2016. El estudio se realiza utilizando un método basado en verosimilitud que explora las propiedades cinemáticas de los eventos seleccionados para separar la señal del fondo, que está dominado por $t\bar{t}$ producido en asociación con chorros hadrónicos (jets) adicionales. Esta búsqueda explota la alta multiplicidad de jets originados a partir de quarks b (b-jets), por lo que la identificación de los mismos es crucial. En esta tesis también se presenta un estudio sobre la optimización de los algoritmos de identificación de b-jets en ATLAS. La razón entre la sección eficaz de $t\bar{t}H$ medida y la correspondiente predicción del SM es $\mu = 0.88^{+0.64}_{-0.61}$, asumiendo un bosón de Higgs con una masa de 125 GeV. Este resultado es consistente ambos con la hipótesis de sólo background así como con la predicción del SM incluyendo el proceso $t\bar{t}H$.

Introduction

In ancient times people were searching for answers to fundamental questions such as: "What is the world surrounding us made of?", "What is matter?".

Modern particle physics has found answers to some of these questions, but also has added new ones to the list: "How do elementary particles interact with each other?", "What is common between different physical interactions?", "Why there is more matter than antimatter in the universe?" or "What is the origin of mass?".

A theory that provides a coherent, but not fully complete picture of elementary particles and the interactions among them is the Standard Model (SM). It gives a unified description of three of the four known fundamental forces. Many theoretical predictions of the SM have been verified experimentally beyond reproach since the 1960s, when the model was established.

One of the fundamental problems raised and solved in the SM is the origin of the mass of the elementary particles. A priori the elementary particles described by the theory are expected to be massless, in contradiction with the observation. Therefore a mechanism that allows particles to acquire their mass was introduced to provide an agreement with experimental evidence. This mechanism assumes the existence of a scalar field, whose excitations manifest themselves as new physical particle that was called the Higgs boson. The SM predicts some properties of the Higgs boson, but its mass is a free parameter of the theory and can only be obtained from experiment. The search for this particle has been one of the main goals of the Large Hadron Collider (LHC) - the world's biggest particle accelerator built at CERN. The discovery of the Higgs boson in 2012 by the ATLAS and CMS collaborations was a triumph of the SM: the last particle predicted by this theory had finally been found.

One of possible modes for Higgs boson production at the LHC is production in association with top-quark pairs ($t\bar{t}H$). This production channel has one of smallest production cross sections at the LHC. At the same time it is of particular physical interest: the coupling of Higgs boson to top quarks, that can be directly measured in this channel, is an important property of the SM. If the measured value of this parameter is significantly different from unity predicted by SM, this would be an indication for a new physics beyond the SM (BSM). Therefore observing the Higgs boson production in association with top quarks is one of the most important physics goals of the LHC.

In this dissertation the search for the Higgs boson in the $t\bar{t}H$ ($H \to b\bar{b}$) channel, using proton-proton collisions at $\sqrt{s} = 13$ TeV registered with the ATLAS detector at the LHC in 2015 and 2016, is presented. This analysis is focused on the semileptonic decay of the $t\bar{t}$ system, resulting in a final state with a single lepton and many jets. The work is focussed on my main contribution - - the development and optimisation of a likelihood-based method to distinguish the signal ($t\bar{t}H$) from the background (dominated by $t\bar{t}$ produced in association with additional jets). Particular kinematic features of both signal and background events are used in the method.

Information on the multiplicity of the jets originating from b-quarks (b-jets) is important in the $t\bar{t}H$ ($H \to b\bar{b}$) search, as there are four b-jets in the final state. Therefore identification of these jets, known as b-tagging, plays a key role. My contribution to ATLAS

b-tagging algorithms optimisation for LHC Run 2 is also presented in this dissertation.

This dissertation is organized as follows. Chapter 1 contains theoretical overview of the Standard Model and the physics of the Higgs boson at hadron colliders. Chapter 2 presents an outline of the LHC and the ATLAS detector, as well as description of the reconstruction of various physical objects out of the signals recorded by the detector. Chapter 3 describes the b-tagging algorithms developed in ATLAS and their optimisation for LHC Run 2. My contribution to the optimisation of the algorithms, exploiting information on the track impact parameter (IP2D, IP3D), is presented. Chapter 4 is an overview of the $t\bar{t}H$ ($H \rightarrow b\bar{b}$) analysis and the results obtained, including a detailed description of my main contribution - the likelihood discriminant method.

1 Theoretical background

1.1 The Standard Model

The Standard Model (SM) of particle physics describes elementary particles and their interactions via three of four fundamental physical forces (gravity is not included). The SM was developed in the 1960s and since then it has been successfully tested in many physics experiments. The observation of the Higgs boson at the LHC in 2012 became one of the confirmations of SM predictions.

1.1.1 Elementary particles

According to the SM, there are two types of elementary particles: fermions and bosons. Matter is composed of fermions, that interact through the exchange of bosons, which mediate forces: electromagnetic, strong and weak.

Fermions are classified into quarks and leptons. Quarks carry an attribute referred to as colour (red, green, blue) and are participating in strong interactions. Quarks can be observed only in bound states, making composite particles - hadrons. This phenomenon is referred to as quark confinement. Three generations of quarks are distinguished: up (u) and down (d) quark in the first generation, charm (c) and strange (s) in the second generation and top (t) and bottom (b) in the third. Leptons are particles not participating in strong interaction. They are also divided into three generations. Charged leptons are electron (e), muon (μ) and tau-lepton (τ), neutral are neutrinos: ν_e, ν_μ and ν_τ. In addition for each quark and lepton an antiparticle with the same mass, but opposite charge exists. The SM fermions with values of their mass and charge are presented in table 1.

	Quarks			Leptons		
Generation	Flavour	Mass	Charge (e)	Flavour	Mass	Charge (e)
1	u	2.2 MeV	2/3	e	0.511 MeV	-1
	d	4.7 MeV	-1/3	ν_e	< 2 eV	0
2	c	1.28 GeV	2/3	μ	105.7 MeV	-1
	s	96 MeV	-1/3	ν_μ	< 0.19 MeV	0
3	t	173.1 GeV	2/3	τ	1776.9 MeV	-1
	b	4.18 GeV	-1/3	$\nu\tau$	< 18.2 MeV	0

Table 1: Quarks and leptons with values of their mass and charge according to Particle Data Group [1].

Gauge bosons are responsible for interactions between particles. Photons (γ) are mediating electromagnetic interactions, gluons (g) - strong interaction. Both of these two bosons are massless. The weak interaction carriers are heavy bosons: charged $W^\pm$ and neutral Z. The properties of gauge bosons are listed in table 2.

Boson	Interaction	Mass	Charge (e)
g	Strong	0	0
γ	Electromagnetic	0	0
$W^{\pm}$	Weak	80.39 GeV	± 1
Z		91.19 GeV	0

Table 2: Gauge bosons with types of interaction they mediate and values of their mass and charge according to Particle Data Group [1].

1.1.2 The Standard Model formalism

The SM is based on renormalisable relativistic quantum field theory. The gauge symmetry group of the SM is

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y, \tag{1}$$

- $SU(2)_L \otimes U(1)_Y$ is the symmetry group of the electroweak interaction, according to unified electroweak (EW) theory developed by Glashow, Salam and Weinberg [2, 3, 4]. $U(1)_Y$ is an abelian group that introduces a new conserved quantum number, the hypercharge Y. $SU(2)_L$ is a non-abelian group, describing the weak interaction, with weak isospin $\vec{I}$ as conserved quantity. The electric charge is related to the third component of the weak isospin I_3 and the hypercharge Y by the Gell-Mann Nishijima formula:

$$Q = I_3 + \frac{Y}{2}. \tag{2}$$

- $SU(3)_C$ is the non-abelian group that describes the strong interaction. The colour (C) is the conserved charge for this group. The theory of strong force is described by Quantum Chromodynamics (QCD) [5, 6, 7, 8, 9].

The SM lagrangian can be divided in two terms, one describing electroweak interactions and another one describing the strong interactions:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{QCD}}. \tag{3}$$

1.1.3 The electroweak theory

The starting point for constructing the part of SM lagrangian, that describes electromagnetic interactions, is considering two kinetic terms: one corresponding to the fermions and another one related to the gauge bosons.

The fermions are represented as Dirac fields, composed of left and right components:

$$\psi_L = \frac{1}{2}(1 - \gamma^5)\psi,$$
$$\psi_R = \frac{1}{2}(1 + \gamma^5)\psi,$$

(4)

The left-handed fermions form weak isospin doublets:

$$Q_L^i = \begin{pmatrix} u^i \\ d^i \end{pmatrix}_L, \quad L_L^i = \begin{pmatrix} \nu^i \\ l^i \end{pmatrix}_L,$$

(5)

the right-handed fermions are represented as weak isospin singlets

$$u_R^i, d_R^i, l_R^i,$$

(6)

where $i = 1, 2, 3$ - the generation number. The right-handed fermions are not involved in weak interactions. Therefore, right neutrinos in Standard Model would not participate in any interaction, so they are not considered.

The fermion component of the lagrangian is described as

$$\mathcal{L}_{\text{fermion}} = \sum_{f=l,q} \bar{f} i \gamma^\mu D_\mu f,$$

(7)

where the covariant derivative

$$D_\mu = \partial_\mu - ig\vec{I} \cdot \vec{W}_\mu - ig'\frac{Y}{2} B_\mu$$

(8)

The gauge part of the lagrangian is

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} F_{\mu\nu}^i F^{\mu\nu i} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu},$$

(9)

where

$$F_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g\epsilon_{ijk} W_\mu^j W_\nu^k,$$
$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu,$$

(10)

g and g' are gauge couplings of $SU(2)_L$ and $U(1)_Y$ groups, W_μ^i (where $i = 1, 2, 3$) and B_μ are gauge fields of these groups, ϵ_{ijk} is the totally antisymmetric tensor. B and W_3 fields mix and provide photon and Z-boson.

1.1.4 Brout-Englert-Higgs mechanism

For the lagrangian composed of two terms described above, $\mathcal{L}_{\text{fermion}}$ and $\mathcal{L}_{\text{gauge}}$, the local symmetries are conserved only assuming that particles are massless. Adding explicit mass terms for gauge bosons or fermions will break the local invariance: gauge symmetry for

bosons and chiral symmetry in the case of fermions. However, it is known from experiment that fermions and $W^{\pm}$ and Z bosons have mass.

This problem is solved due to Brout-Englert-Higgs mechanism [10, 11, 12]. An additional field, called Higgs field, is introduced, that allows SM particles to acquire masses via interacting with it.

The Higgs field is a weak isospin doublet of one charged and one neutral complex scalar field:

$$\phi(x) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{11}$$

The lagrangian of this field can be presented as

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \tag{12}$$

where the first term is kinetic, with covariant derivative D_μ given by equation 8, that is conserving the gauge symmetry of the group $SU(2)_L \otimes U(1)_Y$. The second term is Higgs potential:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \tag{13}$$

The first term of potential can be interpreted as mass of field and the second term represents self-interaction of the field. The minimum of this potential is known as vacuum expectation value of the Higgs field.

The Higgs potential depends on two parameters: μ and λ. To provide stable potential minima λ is required to be positive. For μ parameter there are two possibilities: $\mu^2 > 0$ and $\mu^2 < 0$, presented in figure 1.

For $\mu^2 > 0$ the minimum of the potential $V(\phi)$ is at $\langle 0|\phi|0 \rangle \equiv \phi_0 = 0$, in this case the $SU(2)_L \otimes U(1)_Y$ symmetry is conserved. For the case of $\mu^2 < 0$ the minimum of potential is achieved at a non-zero value of ϕ:

$$\langle 0|\phi^2|0 \rangle \equiv \phi_0^2 = -\frac{\mu^2}{2\lambda} = \frac{\nu^2}{2}. \tag{14}$$

The vacuum state of the field is not invariant with respect to $SU(2)_L \otimes U(1)_Y$. This effect is known as spontaneous symmetry breaking.

To satisfy the requirement that photon should be massless, the minimum of potential is chosen to be

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix}. \tag{15}$$

The expression for field ϕ can be transformed and presented as

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + h(x) \end{pmatrix}, \tag{16}$$
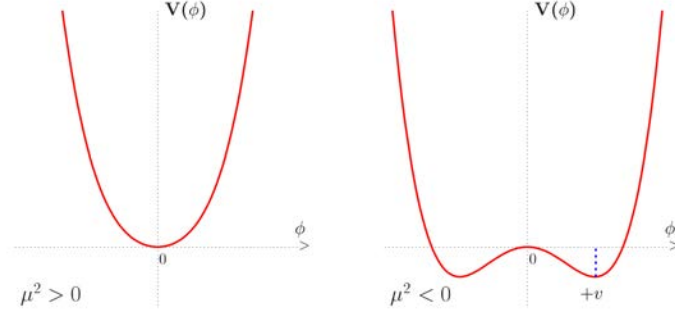
Figure 1: Higgs potential in the case $\mu^2 > 0$ and $\mu^2 < 0$ with $\lambda > 0$ in both cases [14].

where $h(x)$ is describing small deviations from the vacuum state. This is a physical field, associated with the Higgs boson.

Then the part for lagrangian, corresponding to the Higgs field (equation 12) develops into an expression with explicit W and Z mass terms:

$$\mathcal{L}_{\text{Higgs}} = (\partial_\mu h)^2 + \frac{1}{4}g^2 W_\mu W^\mu (v+h)^2 + \frac{1}{8}\left(\sqrt{g^2 + g'^2}\right)^2 Z_\mu Z^\mu (v+h)^2 - V(\frac{1}{2}(v+h)^2), \quad (17)$$

The masses of gauge bosons can then be expressed as

$$m_W = \frac{1}{2}gv, \; m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}. \quad (18)$$

The mass of the Higgs boson can be presented as

$$m_H = \sqrt{2\lambda}v = \sqrt{-2\mu^2}, \quad (19)$$

its value is not predicted, as μ is a free parameter of the SM.

Masses of fermions can be generated using the same scalar field ϕ. For each fermion generation a $SU(2)_L \otimes U(1)_Y$ invariant term, known as Yukawa lagrangian, is introduced:

$$\mathcal{L}_{\text{Yukawa}} = -\lambda_e \bar{L}\phi e_R - \lambda_d \bar{Q}\phi d_R - \lambda_u \bar{Q}\tilde{\phi}u_R + h.c., \quad (20)$$

where ϕ is the scalar field described by equation , $\tilde{\phi} = i\tau_2 \phi*$, τ_2 - the second Pauli matrix.

The fermion masses are then defined as [14]

$$m_e = \frac{\lambda_e v}{\sqrt{2}}, \; m_u = \frac{\lambda_u v}{\sqrt{2}}, \; m_d = \frac{\lambda_d v}{\sqrt{2}}. \quad (21)$$

The full expression for the EW component of the SM lagrangian is

$$\mathcal{L}_{SU(2)_L \otimes U(1)_Y} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}, \quad (22)$$

where the terms are defined by equations 9, 7, 12 and 20.

To summarise, the Brout-Englert-Higgs mechanism allows to obtain masses of gauge bosons and fermions, keeping the $SU(2)_L \otimes U(1)_Y$ invariance. The electromagnetic $U(1)_Q$ and the color symmetry $SU(3)_C$ remain unbroken. Therefore the full lagrangian of the Standard Model is invariant with respect to $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ symmetry.

1.1.5 Quantum chromodynamics

Quantum chromodynamics (QCD) is a non-abelian theory with gauge group $SU(3)_C$ that describes strong interactions. The conserved quantity under the group symmetry trandformations is colour.

The QCD lagrangian in the SM is given by

$$\bar{q}i\gamma^\mu D_\mu q - \frac{1}{4}G^a_{\mu\nu}G^{a\,\mu\nu}, \tag{23}$$

where q are quark fields and D_μ is the covariant derivative, describing interaction between quarks and gluons, given by

$$D_\mu = \partial_\mu - ig_s T_a G^a_\mu, \tag{24}$$

where g_s is the strong coupling constant, T_a ($a = 1,..8$) are $SU(3)_C$ generators, G^a_μ - gluon fields. $G^a_{\mu\nu}$ is field tensor, defined as

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc} G^b_\mu G^c_\nu, \tag{25}$$

with f_{abc} - structure constants of the $SU(3)_C$ group.

Gluons carry colour charge, therefore they can interact with each other. This interaction is described by the last term in equation 25. Due to the gluon self-interaction the strong coupling constant is dependant on the energy scale. This dependency can be approximated as

$$\alpha_S(Q^2) = \frac{12\pi}{(33 - 2n_f)\log\left(\frac{Q^2}{\Lambda^2_{QCD}}\right)} \tag{26}$$

where α_S is related to the strong coupling constant g_S as $\alpha_S = g_S^2/4\pi$, Q - energy scale, n_f - number of active flavour quarks ($m_q < Q$), Λ_{QCD} is the scale, where the perturbative approximation is no longer valid.

With increasing energy (or decreasing distance) α_S decreases. For energy reaching the limit $Q^2 \to \infty$ quarks are becoming free, and this phenomenon is known as asymptotic freedom. With lower energies (or larger distances) α_S increases and diverging at $Q^2 \to 0$. Due to this, quarks and gluons do not exist as free particles. This feature is referred to as confinement. Quarks produced in high-energy interactions tend to create new bound states with quarks with opposite colour charge from vacuum and produce collimated streams of hadrons known as jets.
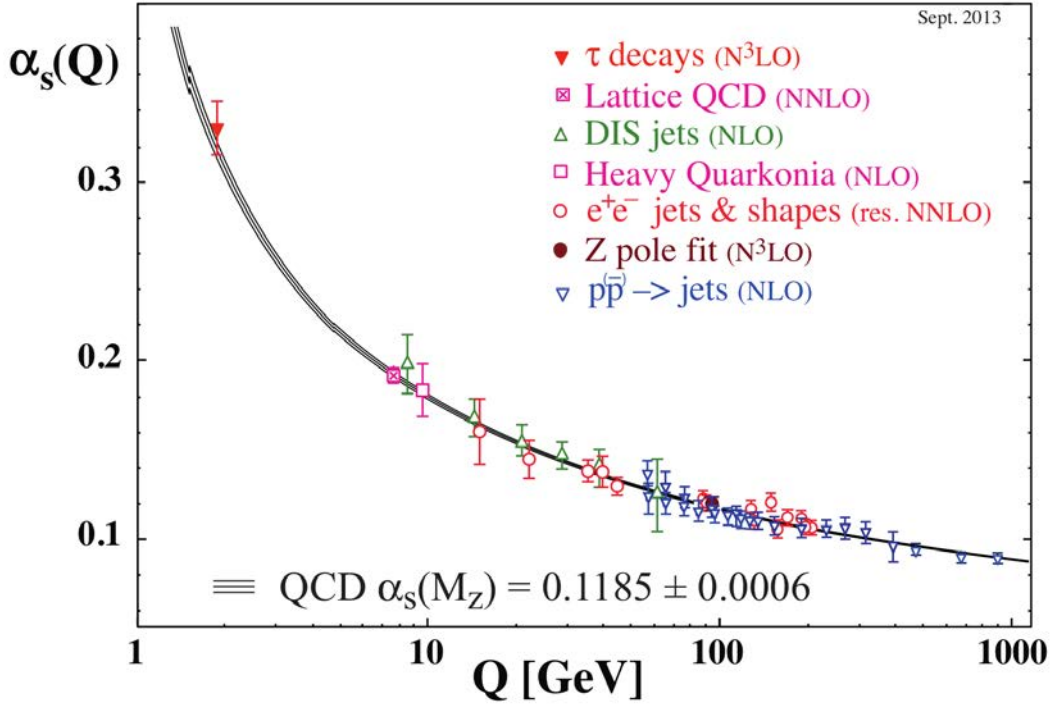
Figure 2: Measurements of α_S as a function of the energy scale Q [1].

1.2 Search for the Higgs boson at the LHC

The search for the Higgs boson as well as studying its properties are very important as it was the last undiscovered particle, predicted by the SM. Since the 1960s, when the SM was developed, and until the discovery in 2012, the Higgs boson searches were performed in several experiments: LEP, Tevatron and LHC. One of the main goals of building of the LHC machine was discovering the Higgs boson (or proving its absence).

1.2.1 Production at hadron colliders

There are four main modes of the Higgs boson production at hadron colliders, which are illustrated in figure 3:

- Gluon-gluon fusion (ggH) - the Higgs boson is produced via gluon-gluon fusion, mediated by a virtual quark loop, where the main contribution is from the top quark, owing to its large Yukawa coupling. This is the main mechanism of Higgs production on Tevatron and LHC.

- Vector boson fusion (VBF): two W or Z bosons originating from the initial quarks are interacting and produce a Higgs boson. This production mode is featuring a special signature characteristics, that allows to distinguish it from the background: the presence of two light jets in the forward and backward region of detector (in
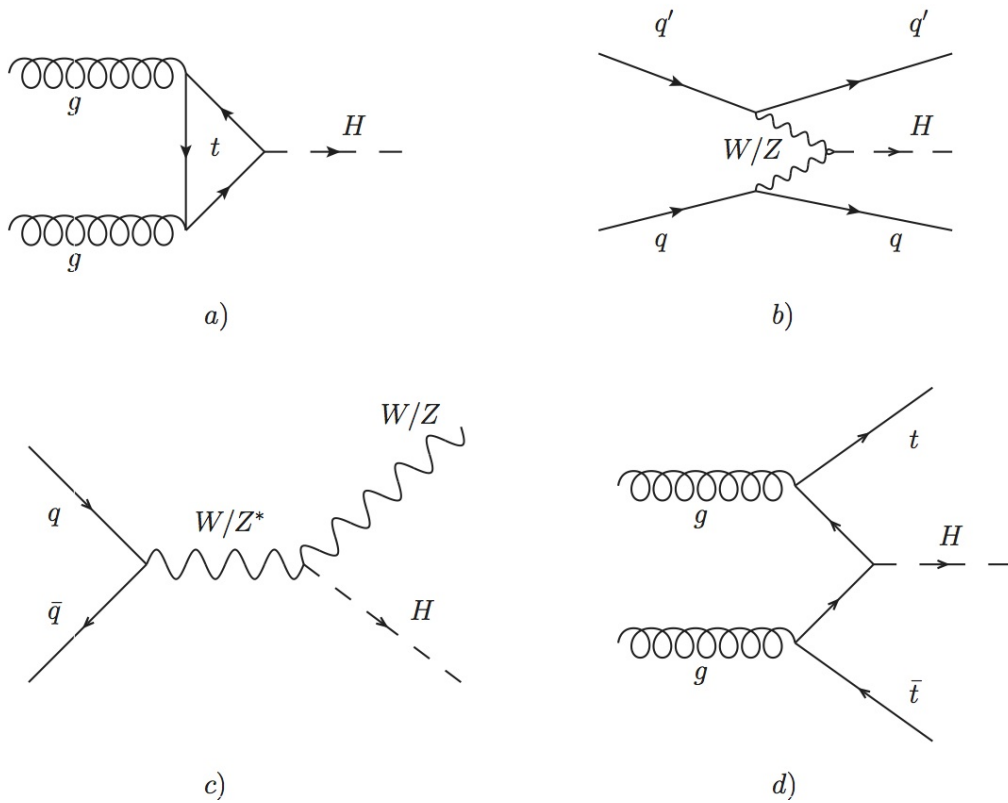
Figure 3: Representative Feynman diagrams for the main production modes for SM Higgs boson production at hadron colliders: (a) gluon-gluon fusion, (b) vector boson fusion, (c) Higgs-strahlung, (d) $\bar{t}tH$ production.

region of pseudorapidity close to ± 3) and maximum transverse momentum about half of the mass of the vector boson.

- Associated production with a vector boson (VH): the Higgs boson is produced in association with a W or Z boson, which is typically required to decay leptonically. This channel was of particular relevance in the search for a light SM Higgs boson at the Tevatron [16].

- $t\bar{t}H$ and $b\bar{b}H$: the Higgs boson is produced in association with a top-antitop quark pair ($t\bar{t}$) or a bottom-antibottom quark pair ($b\bar{b}$). The $t\bar{t}H$ process is of particular importance as it gives direct access in a tree level diagram to Yukawa coupling to top quark. The $t\bar{t}H$ process has smaller cross-section, than three processes described above, but its contribution grows with higher energies of pp collisions. The $b\bar{b}H$ process has comparable cross section to the $t\bar{t}H$ process, but its signature is similar to ggH, since the associated b-quarks are often produced along the beam direction.

The SM cross-sections for the different Higgs-boson production modes in pp collisions as function of the center-of-mass-energy are presented in figure 4. The values of cross-sections, predicted for $\sqrt{s} = 13$ TeV are listed in table 3.
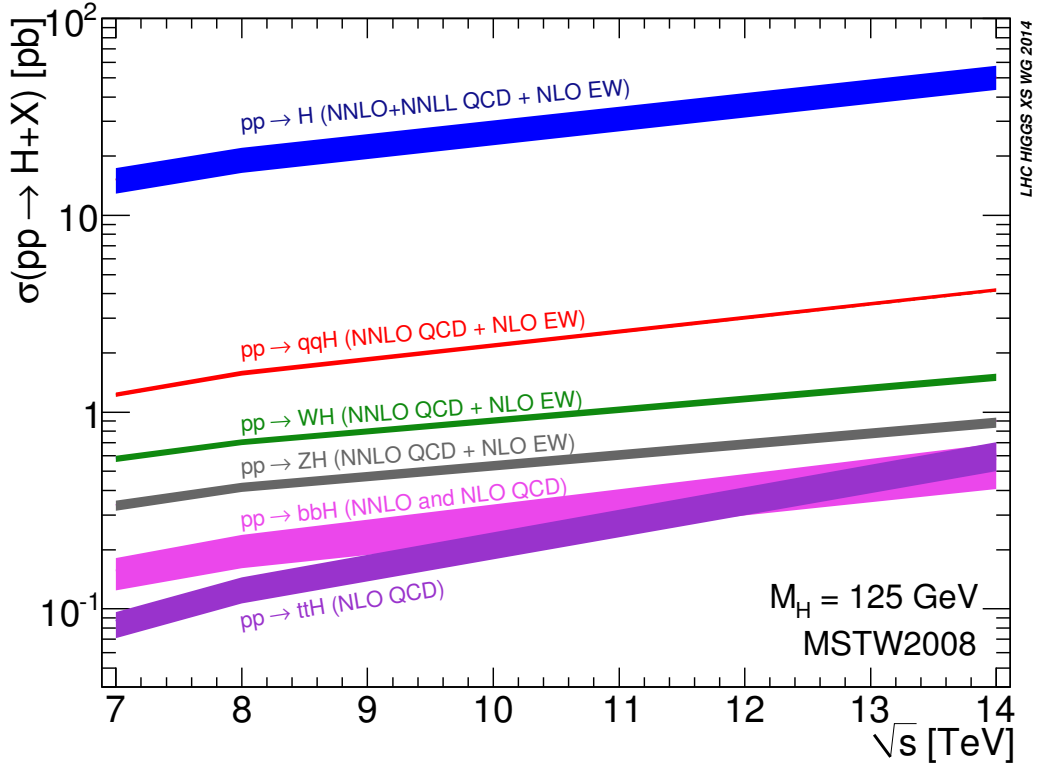

Figure 4: Cross-sections of different Higgs boson production modes in pp collisions as a function of centre-of-mass energy $\sqrt{s}$ assuming a Higgs-boson mass $M_H = 125$ GeV. From Ref. [18].

Production mode	Cross section [pb]
ggH	43.9
VBF	3.75
WH	1.38
ZH	0.87
$t\bar{t}H$	0.509
$b\bar{b}H$	0.512

Table 3: Cross sections of different Higgs-boson production modes predicted by the SM for pp collisions at $\sqrt{s} = 13$ TeV, assuming a value of the Higgs-boson mass of $M_H = 125$ GeV. From Ref. [18])

1.2.2 Decay modes

The Higgs boson decays preferentially to pairs of the heaviest particles kinematically accessible. Loop-induced decays to photons and gluons are also possible. Figure 5 shows branching ratios of the SM Higgs boson as a function of its mass. For $M_H = 125$ GeV the largest branching fraction corresponds to the $H \to b\bar{b}$ channel. Different branching ratios for the SM Higgs boson decays are presented in table 4.

Decay mode	Branching ratio [%]
$H \to b\bar{b}$	57.7
$H \to WW$	21.6
$H \to gg$	8.55
$H \to \tau\bar{\tau}$	6.37
$H \to c\bar{c}$	2.67
$H \to ZZ$	2.6
$H \to \gamma\gamma$	0.229
$H \to Z\gamma$	0.155
$H \to s\bar{s}$	0.044
$H \to \mu\bar{\mu}$	0.022

Table 4: Branching ratios for the different Higgs-boson decay modes assuming $M_H = 125$ GeV. From Ref. [18].
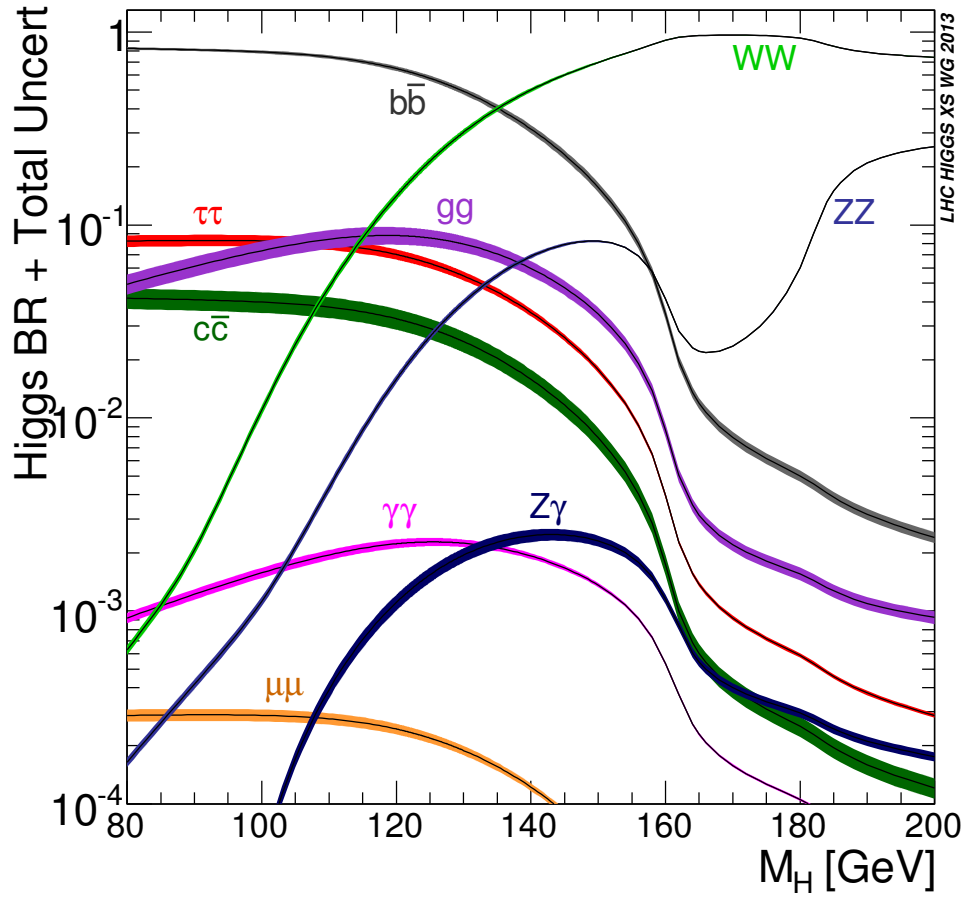
Figure 5: Branching ratios for the different Higgs-boson decay modes as a function of Higgs-boson mass M_H. From Ref. [17].

1.2.3 Discovery

In July 2012 the ATLAS and CMS collaborations announced the discovery of a new particle, compatible with the Higgs boson [19, 20]. The search was performed exploiting the $\gamma\gamma$, $ZZ^* \to 4\ell$ ($\ell = e, \mu$), $W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$, $b\bar{b}$ and $\tau^+\tau^-$ decay modes with data collected at a centre-of-mass energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. The most significant excesses were observed in $\gamma\gamma$ and $ZZ^* \to 4\ell$ channels. The mass of Higgs candidate of 126.0 ± 0.4(stat) ± 0.4(syst) GeV was measured by ATLAS and 125.3 ± 0.4(stat)± 0.5(syst) GeV by CMS. The significance of observation was 5.9σ for ATLAS and 5.8σ for CMS.
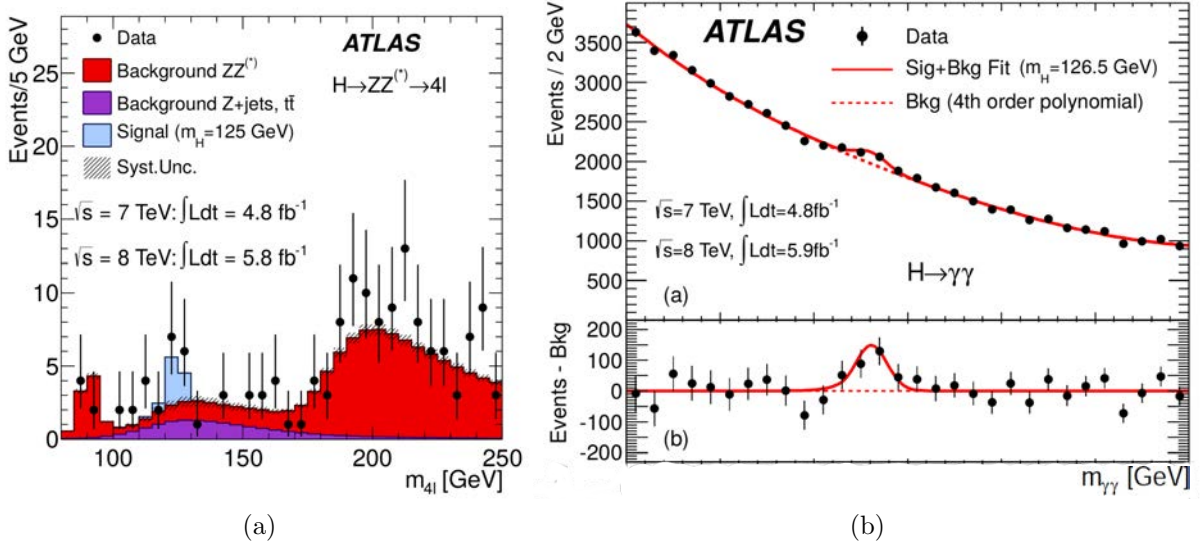
Figure 6: Reconstructed mass distribution in $H \to ZZ* \to 4l$ (a) and $H \to \gamma\gamma$ search by ATLAS. A peak at mass value around 125 GeV is observed [19].

After the discovery the properties of the new particle were further studied to test its compatibility with the SM Higgs boson. Measurements of its couplings to SM particles were found to be consistent to the expectation. Spin and parity of the Higgs candidate were examined and were shown to be consistent with the 0^+ hypothesis, predicted by SM [21].

2 The ATLAS experiment at the Large Hadron Collider

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a circular particle accelerator machine constructed at the European Organization for Nuclear Research (CERN) situated at the French-Swiss border near Geneva. It is accelerating and colliding head-on beams of charged particles: protons and heavy ions. It is placed in the underground tunnel 27 km length and 45-100 m depth that was used previously for the Large Electron-Positron collider (LEP). The LHC was designed to collide protons at a centre-of-mass energy of $\sqrt{s} = 14$ TeV.

The aim of the LHC is to test the predictions of the SM and to search for new phenomena, and this defined its technical parameters. The large centre-of-mass energy allows to study SM processes that occur at this energy scale, as well as to probe various scenarios beyond the Standard Model (BSM). The high luminosity corresponds to large data sets, which are crucial to study rare processes, such as Higgs production and to perform high-precision measurements of the SM parameters.

The choice of proton-proton collisions was made to minimize the loss of energy due to synchrotron radiation when accelerating charged particles in a curved trajectory: for an electron collider this loss would be much larger at such high energy regime. A proton-antiproton machine would not allow to reach the physics goals because of very low rate of antiproton production.

The luminosity describes the rate of collisions produced by a collider. The instantaneous luminosity $\mathcal{L}$ is defined as:

$$\mathcal{L} = \frac{n_1 n_2 n_b f_{\text{rev}} F}{4\pi\sigma_1\sigma_2}, \tag{27}$$

where n_1 and n_2 are numbers of protons per bunch in two beams, n_b is the number of bunches per beam, f_{rev} - the beam revolution frequency, F is a geometric factor representing the crossing angle of the two beams, and σ_1 and σ_2 denote the transverse beam dispersions. The integrated over time luminosity is a proportionality constant between the number of events produced with a certain reaction and the cross-section of this reaction.

$$N = L \times \sigma. \tag{28}$$

The design parameters of the LHC are listed in table 5.

Before protons reach the LHC they go through several acceleration steps to reach the final energy. The CERN accelerator chain is presented at figure 7. First of all protons are produced by ionizing of hydrogen gas and accelerated linearly to the energy of 50 MeV by LINAC II. After that three circular accelerators Booster, Proton Synchrotron (PS) and Super Proton Synchrotron (SPS) increase the energy of the proton beams to 450 GeV. Then the beams go inside the LHC main ring where they reach the final energy of 7 TeV.

The magnet system of the LHC ring are providing a magnetic field of 8.3 T which is needed to keep a beam energy of 7 TeV on the trajectory. The magnets made of
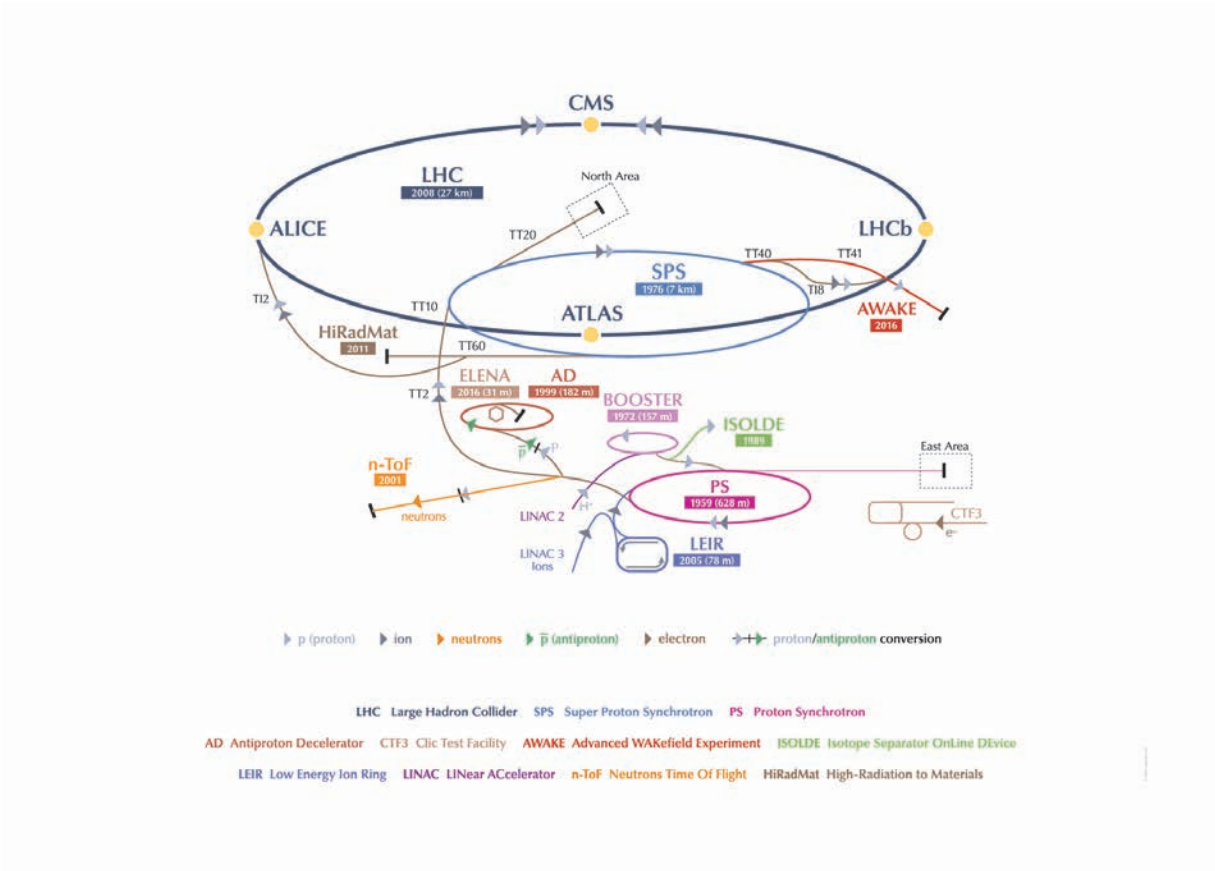
Figure 7: Schematic view of the CERN accelerator complex. The four major LHC experiments are shown at the interaction points.

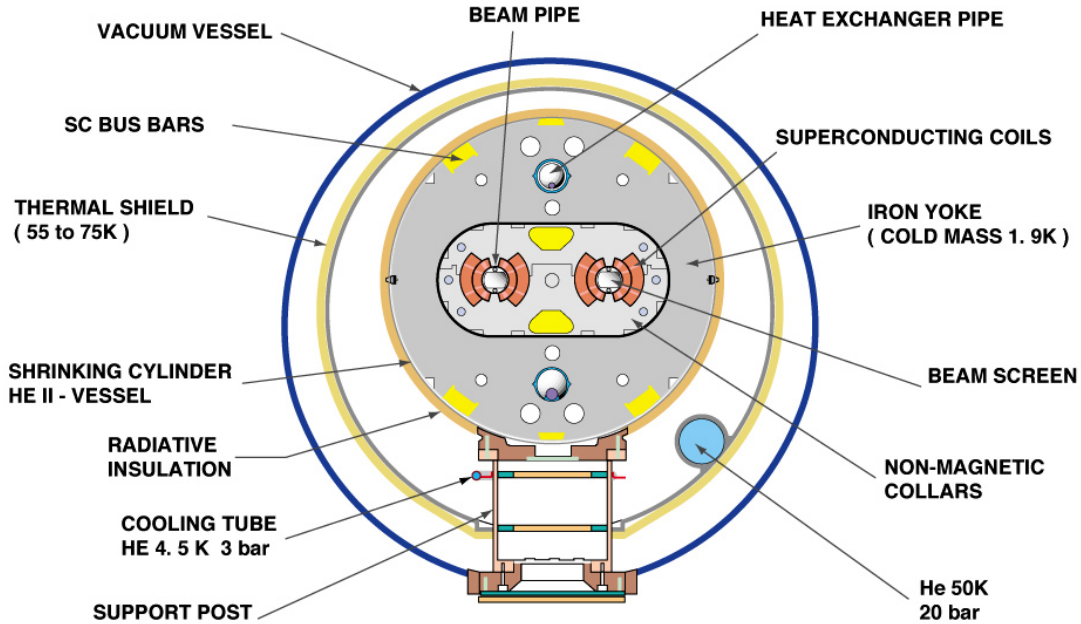|	
Beam energy	7 TeV
Injected beam energy	0.45 TeV
Synchrotron radiation loss per turn	7 keV
Peak luminosity	10^{34} cm^{-2}s^{-1}
Particles per bunch	1.1×10^{11}
Number of bunches	2808
Bunch spacing	24.95 ns
Vertical beam size	18 μm
Horizontal beam size	71 μm
Beam crossing angle	285 μrad
Beam lifetime	13.9 h
Beam energy loss per turn	7 keV
Number of dipole magnets	1232
Max dipole field	8.3 T
Main dipole operation temperature	1.9 K

Table 5: Design parameters of the LHC.

superconductor material are kept at a temperature of 1.9 K. There are two types of magnets: bending and focusing, together they form the desired beam trajectory. There are in total 1232 dipole magnets forming the field, which are supplemented with various multipole components to correct imperfections of the magnetic field at the extremities. As the two proton beams should be accelerated in opposite directions, there are two beam pipes in a dipole magnet, and two opposite sign magnetic fields are created around them. Figure 8 shows the cross-section of the LHC dipole magnet and the beam pipes.

There are four interactions points where LHC beams cross, with detectors deployed at these regions to register the collisions:

- ATLAS (A Toroidal LHC Apparatus) - multipurpose detector, designed to explore various physics processes. This is the largest detector at the LHC (25 m in height and 44 in length) with a weight of 7000 tonnes.

- CMS (Compact Muon Solenoid) - second multipurpose detector, aiming the same broad range of physical processes. With much smaller size than ATLAS (14.6 m in height and 21.6 in length), it weights 12500 tonnes. ATLAS and CMS are placed at opposite intersection points of the LHC ring. Both detector design and physical analyses performed at the two experiments are totally independent, allowing to cross-check the final results.

- ALICE (A Large Ion Collider Experiment) - detector to register heavy ions collisions,

VACUUM VESSEL

BEAM PIPE

HEAT EXCHANGER PIPE

SC BUS BARS

SUPERCONDUCTING COILS

THERMAL SHIELD
(55 to 75K)

IRON YOKE
(COLD MASS 1. 9K)

SHRINKING CYLINDER
HE II - VESSEL

BEAM SCREEN

RADIATIVE
INSULATION

NON-MAGNETIC
COLLARS

COOLING TUBE
HE 4. 5 K 3 bar

SUPPORT POST

He 50K
20 bar

CERN AC _HE107A_ V02/02/98

Figure 8: Cross-section of the LHC superconducting dipole.

that is designed to study the physics of quarks and gluons behaviour at extreme energy densities.

- LHCb (Large Hadron Collider beauty) - detector designed for studying properties of B-hadrons, in particular CP violation in B-meson decay.

- LHCf (Large Hadron Collider forward) - detector to study forward physics (observation of particles at the region of angles close to the proton beam direction). The goal of studying this angle region is exploring the cosmic ray phenomena, as cascades of particles observed there are similar to cosmic rays. It consists of two detectors which are deployed along the LHC beamline, at 140 metres on either side of the ATLAS collision point.

- TOTEM (TOTal Elastic and diffractive cross section Measurement) - experiment whose main purpose is the measurement of the total proton-proton cross-section, as well as deeper understanding of the proton structure through exploring elastic and diffractive scattering. It consists of small detectors that are spread within a distance of 500 m around the CMS interaction point.

- MOEDAL (Monopole and Exotics Detector at the LHC) - experiment whose purpose is to search directly for massive long-lived particles and magnetic monopoles (hypothetical elementary particles with magnetic pole). It is located around the same interaction point as the LHCb.

2.2 The ATLAS detector

ATLAS is a cylindrical forward-backward symmetric detector composed of

- inner detector surrounded by a thin superconducting solenoid

- electromagnetic and hadronic calorimeters surrounded by three large superconducting toroids

- muon spectrometer

A schema of the ATLAS detector is shown in figure 9.

2.2.1 Coordinate system

A right-handed coordinate system is used in ATLAS. The z-axis is defined by the beam direction, x-axis points towards the centre of the LHC ring and the y-axis points upwards.

The asimuthal angle ϕ is defined with respect to x-axis in the $x-y$ plane, $(-\pi < \phi < \pi)$, and θ is the polar angle with respect to the z-axis.

Rapidity variable, which is used for massive objects such as jets, is defined as

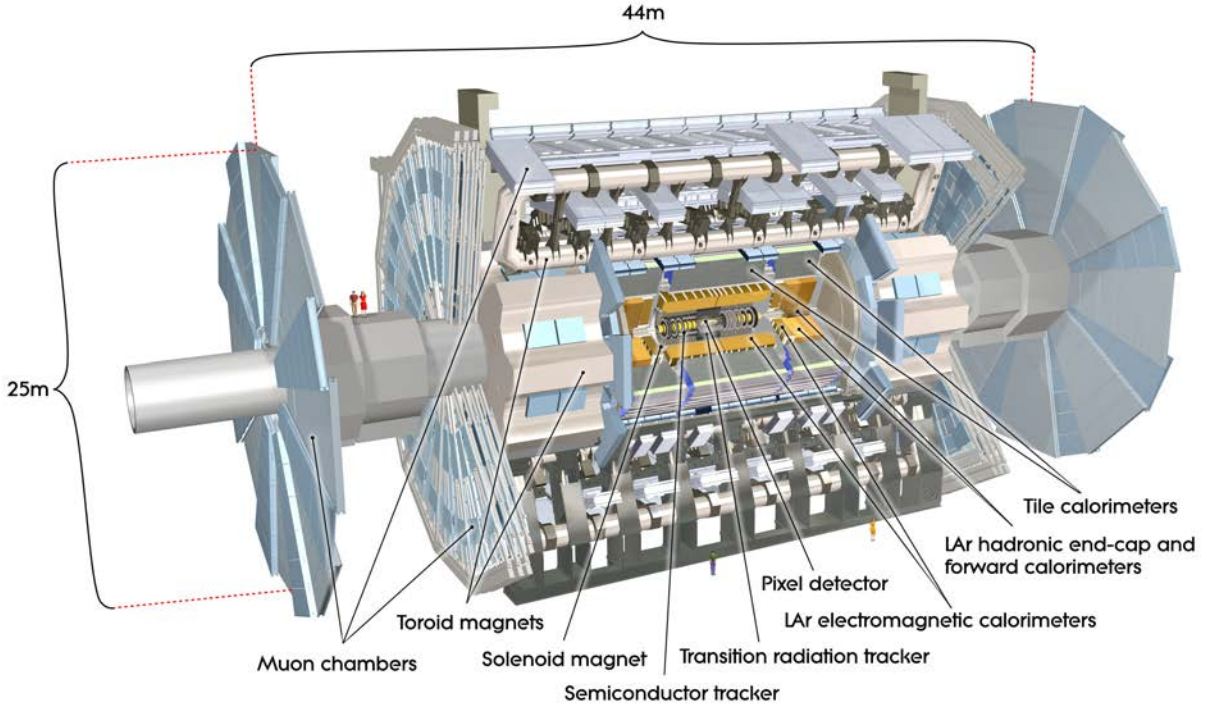$$y = \frac{1}{2}\ln\frac{(E + p_z)}{(E - p_z)}. \tag{29}$$

Figure 9: Cut-away view of the ATLAS detector, showing the different subdetectors and the magnet systems. From Ref. [22].

For those particles which mass can be neglected, the pseudorapidity η is calculated instead:

$$\eta = -\ln(\tan(\frac{\theta}{2})). \tag{30}$$

The distance ΔR in the pseudorapidity-azimuthal angle space is defined as

$$\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}. \tag{31}$$

The tranverse momentum p_T and the transverse energy E_T are defined in the $x - y$ plane.

2.2.2 Magnet system

A magnet field is used in ATLAS for bending particle trajectories. This allows to calculate the particle momentum out of track curvature.

ATLAS has a system of four large superconducting magnets: solenoid and three toroids (one barrel and two end-caps).

The central solenoid is surrounding the inner detector and produces a 2T magnetic field, parallel to the beam axis. It is placed in the same cryostat as calorimeter.

The magnetic field for the muon spectrometer is formed by three toroids: one large in the barrel and two smaller in the end-caps of the detector. These magnets are producing a 4T strength field.

Figure 10 illustrates a scheme of ATLAS magnet system. In figure 11 photograph of the barrel toroid after installation is shown.
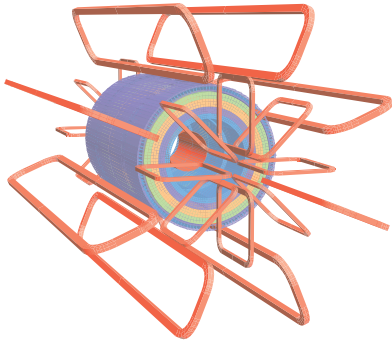

Figure 10: Geometry of magnets (in orange) and tile calorimeter steel. Windings of barrel and end-caps toroid coils are visible, as well as the solenoid, that lies inside the calorimeter volume [22]
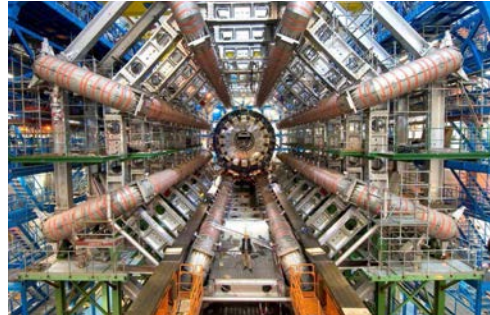

Figure 11: Barrel toroid after installation in the underground cavern [22]

2.2.3 Inner detector

The purpose of the ATLAS inner detector is performing of tracking measurements. With LHC high luminosities the track density in the detector is very large. Thus to provide track resolution as high as required for physical analyses a very high detector granularity is needed, and this demand determined the design of the inner-detector system.

There are three components of the inner detector:

- pixel detector

- silicon microstrip, or semiconductor tracker (SCT)

- transition radiation tracker (TRT)

The layout of the ATLAS inner detector is presented on the figure 12

2.2.3.1 Pixel detector

The pixel detector is the innermost part of the ATLAS tracking system. Being located close to the interaction point, pixel detector plays important role for identification of particles with lifetime, such as b- and c- hadrons and τ-leptons.
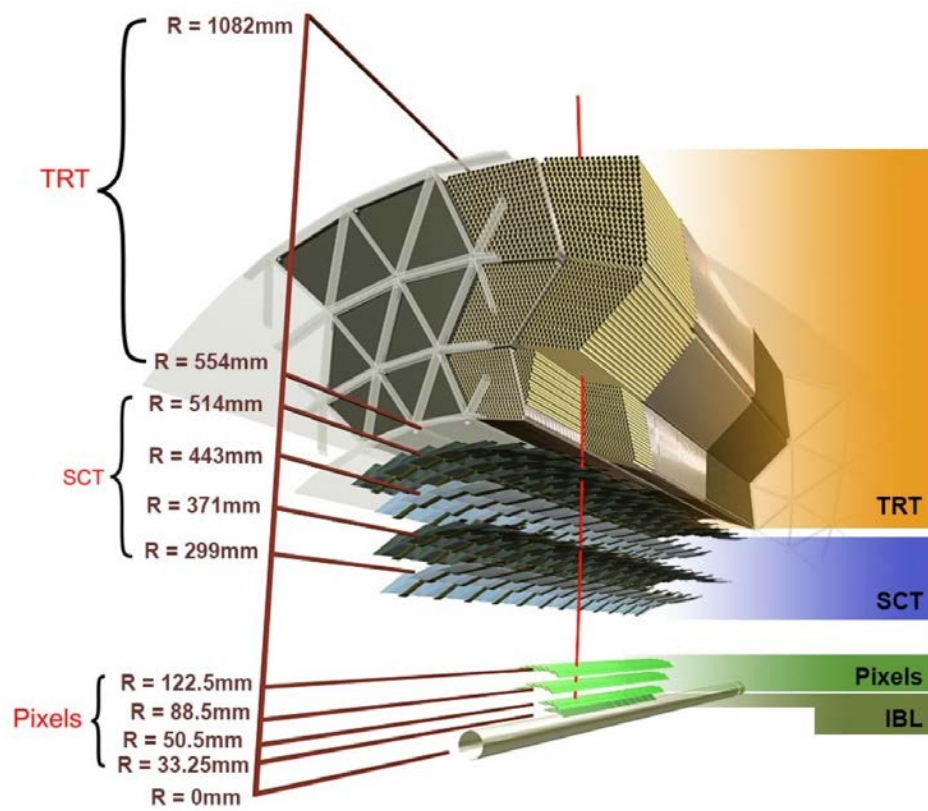
Figure 12: Scheme of the ATLAS inner detector [32]

For Run 1 of the LHC the pixel detector consisted of three cylindrical barrel layers, in 2014 the new forth innermost layer, called the Insertable B-Layer (IBL), was installed. The new layer allowed to improve track parameters measurement resolution for the LHC Run 2 (presented in section 3.2.1). In particular, it had a large positive impact on the b-jet identification performance, as shown in the section 4.3. The barrel layers of the detectors are surrounding the beam axis at r = 33.25, 50.5, 88.5, and 122.5 mm. In the end-cap regions three perpendicular to the beam axis disc layers are installed on each side of the detector.

There are 1744 detector modules, each composed of 46080 pixels. Each sensor is a silicon wafer with the 16 front-end chips for readout. Size of a pixel is $50 \times 400\,\mu$m. Thanks to its high granularity pixel detector provides the highest precision of track parameters measurement among the inner detector components.

The Insertable B-layer (IBL)

The major ATLAS inner detector upgrade for Run-2 is the addition of the Insertable B-Layer (IBL), a fourth pixel layer in order to have better track and vertex reconstruction performance at the higher luminosities [52].

With higher number of interactions per bunch-crossing, or pile-up, the detector readout efficiency is decreased. This affects the innermost detector layer more than other layers and has thereby huge impact on the b-tagging performance.

The pixel detector used during Run-1 was designed for a peak luminosity of $\mathcal{L} \approx 10^{34}\text{cm}^{-2}\text{s}^{-1}$, while during Run-2 peak luminosity is expected to be $\mathcal{L} \approx 1.7\times10^{34}\text{cm}^{-2}\text{s}^{-1}$. Because of that the addition of the new pixel layer was necessary to maintain tracking and b-tagging performance despite the increased pile-up.

The IBL was inserted inside the existing pixel detector at a radius of ≈ 3.3 cm from the beamline, while the next to innermost layer (which was innermost pixel layer in Run-1) is located at a radius of ≈ 5 cm. Another advantage of the IBL is its higher granularity, with pixels of size 50 μm $\times$ 250 μm instead of 50 μm $\times$ 400 μm for the former innermost pixel layer. Finally, the average number of pixel measurements on a single track became 4 instead of 3. This improves the tracking robustness with respect to pile-up and possible pixel module failures.

The Run-2 to Run-1 data comparison showed that IBL significantly improves track impact parameter resolution: by up to a factor of 2 for both transverse and longitudinal components for low-p_T tracks. For a typical 2 GeV track, the transverse impact parameter resolution is now ≈ 30 μm and the longitudinal one ≈ 80 μm [53].

Figure 13 shows Run-2 and Run-1 impact parameter resolution as function of track p_T.

2.2.3.2 Semiconductor Tracker (SCT)

The Semiconductor Tracker (SCT) is based on sensors, made of silicon microstrips. It is using the p-in-n type of semiconductor.
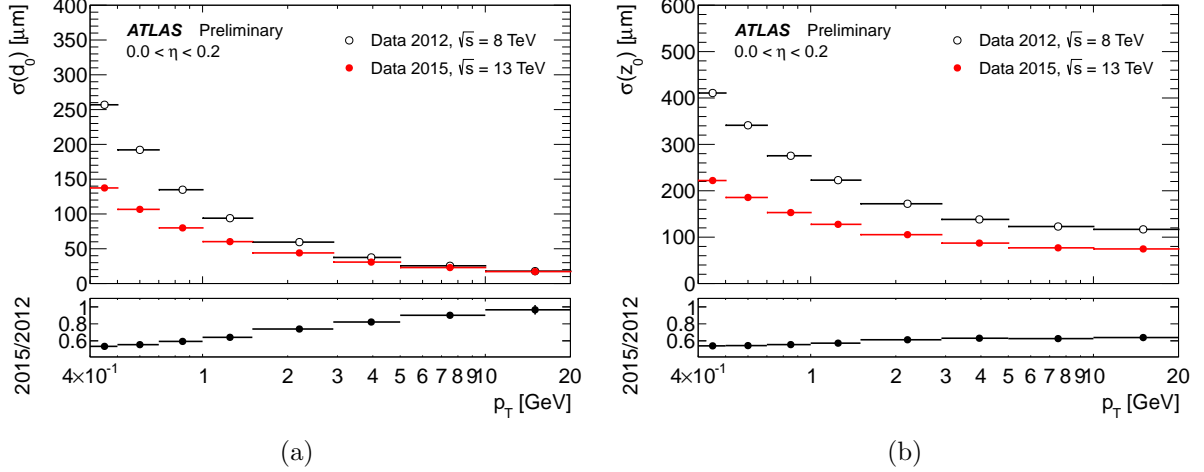
Figure 13: Unfolded transverse (a) and longitudinal (b) impact parameter resolution measured from data in 2015, $\sqrt{s} = 13$ TeV, with the Inner Detector including the IBL, as a function of pT, for values of $0.0 < \eta < 0.3$ compared to that measured from data in 2012, $\sqrt{s} = 8$ TeV

SCT consists of 4 cylindrical layers parallel to the beam axis and 9 end-cap disks with radially oriented strips. There are 2112 detector modules in the barrel layers and 988 in the endcap disks on each side. The barrel layers are located at a distance of 299 - 514 mm from the beam line.

A barrel module components are presented on figure 14. The four sensors, two on the top and two on the bottom side of the module, are rotated by stereo angle of ± 20 mrad, which allows to improve the spatial resolution of the detector. An SCT sensor has pitch of 80 m and consists of a wafer and 768 active microstrips of 12.8 cm length. The wafer is a n-type semiconductor and the strips - p-type.

2.2.3.3 Transition radiation tracker (TRT)

The Transition radiation tracker (TRT) is designed to improve the track resolution in the range of $\eta < 2$ and, at the same time, to measure transition radiation, which is important for electron identification.

The TRT consists of straw tubes filled with gas, embedded in a matrix of polypropylene fibres, which is used as transition radiation material. The transition radiation (photon emission) rate is inversely proportional to the mass of the particle, therefore, among the particles that are leaving a track in the detector, electrons produce the largest amount of photons, so measuring the transition radiation allows to identify them. For electrons with energies above 2 GeV from seven to ten high-threshold hits from transition radiation are expected.

The TRT contains up to 73 layers of straws in the barrel and 160 straw planes in the end-cap. The straws are filled with gas mixture of 70% Xe, 27% CO_2 and 3% O_2. A track with $p_T > 0.5$ GeV and $\eta < 2$ traverses minimum 36 straws (except in the barrel-end-cap

<table>
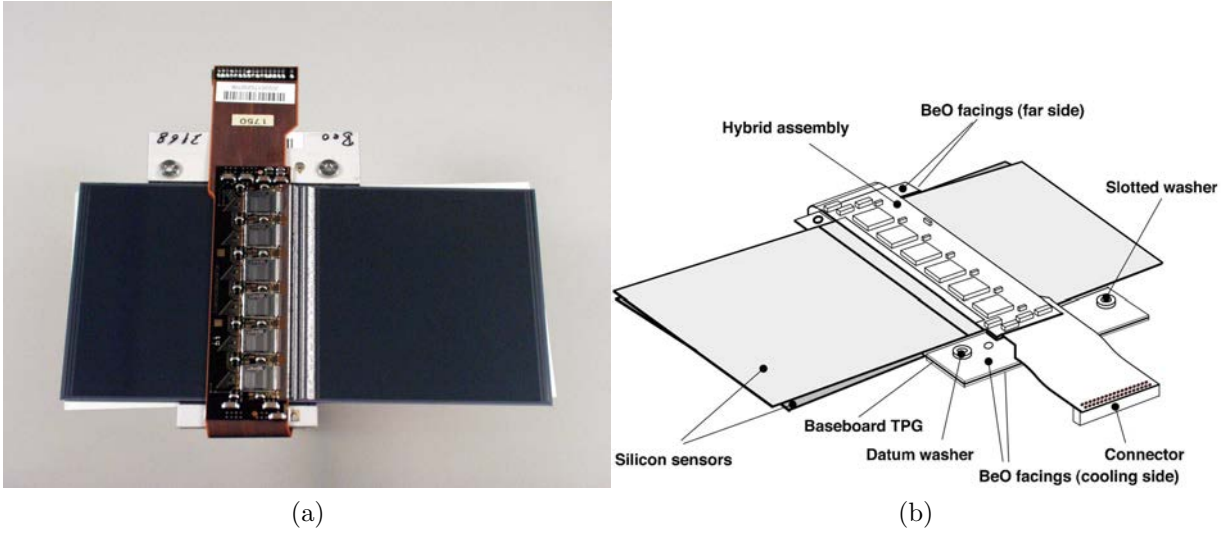<tr><td>(a)</td><td>(b)</td></tr>
</table>

Figure 14: Components of a barrel SCT module [22]

transition region of $0.8 < \eta < 1.0$, where a track crosses minimum 22 straws).

Figure 15 shows a quadrant of the TRT barrel and 16 presents a four-plane TRT end-cap wheel during assembly.

2.2.4 Calorimeters

ATLAS calorimetry system measures the energy of electromagnetically and hadronically interacting particles.

The scheme of the ATLAS calorimetry system is presented on figure 17. It consists of several detectors providing full ϕ-symmetry and coverage over $\eta < 4.9$. The innermost calorimeters are the electromagnetic barrel calorimeter (EMB), the electromagnetic end-cap calorimeter (EMEC), the hadronic end-cap calorimeter (HEC) and the forward calorimeter (FCal). For these calorimeters liquid argon (LAr) is used as the active detector material. The outer calorimeter part is Tile barrel, which consists of scintillator tiles and steel as absorber medium. LAr barrel and EMEC are electromagnetic calorimeters, designed for measurement of electromagnetic showers that electrons and photons produce through the Bremsstrahlung process and pair production. HEC, FCal and Tile are hadronic calorimeters, used to measure energy of hadronic showers (jets).

2.2.4.1 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is based on high granular liquid argon (LAr) technology, that allows to make high-precision measurements of electrons and photons energy. It consists of the barrel section, covering and two end-caps, each housed in own cryostat. The liquid argon was chosen as active detector material due to its stability of response over time and resistance to radiation.

Both barrel and end-cap of ECAL consist of accordion-shaped kapton electrodes and lead absorber plates immersed in liquid argon. An accordion geometry has been chosen
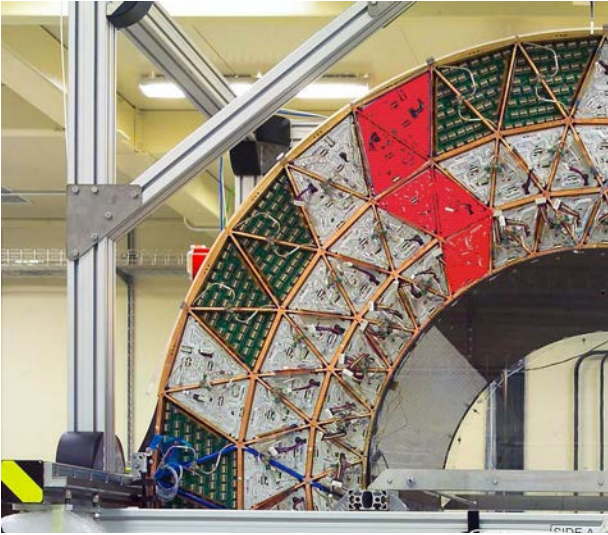
Figure 15: Photograph of a quadrant of the TRT barrel during the integration of the modules at CERN. The shapes of one outer, one middle and one inner TRT module are highlighted. The barrel support structure space-frame can be seen with its triangular sub-structure. [22]

Figure 16: Photograph of four-plane TRT end-cap wheel during assssembly. The inner and outer C-fibre rings can be seen, as well as the first layer of straws and the first stack of polypropylene radiator foils beneath it. Also visible are the plastic end-plugs which are used to position and fix the straws to the C-fibre outer ring. [22]

Figure 17: Cut-away view of the ATLAS calorimeter system [22]

to provide a full coverage in ϕ without gaps, and a fast extraction of the signal at the back and at the front of the electrodes. Photograph of a barrel ECAL module is shown in figure 18.

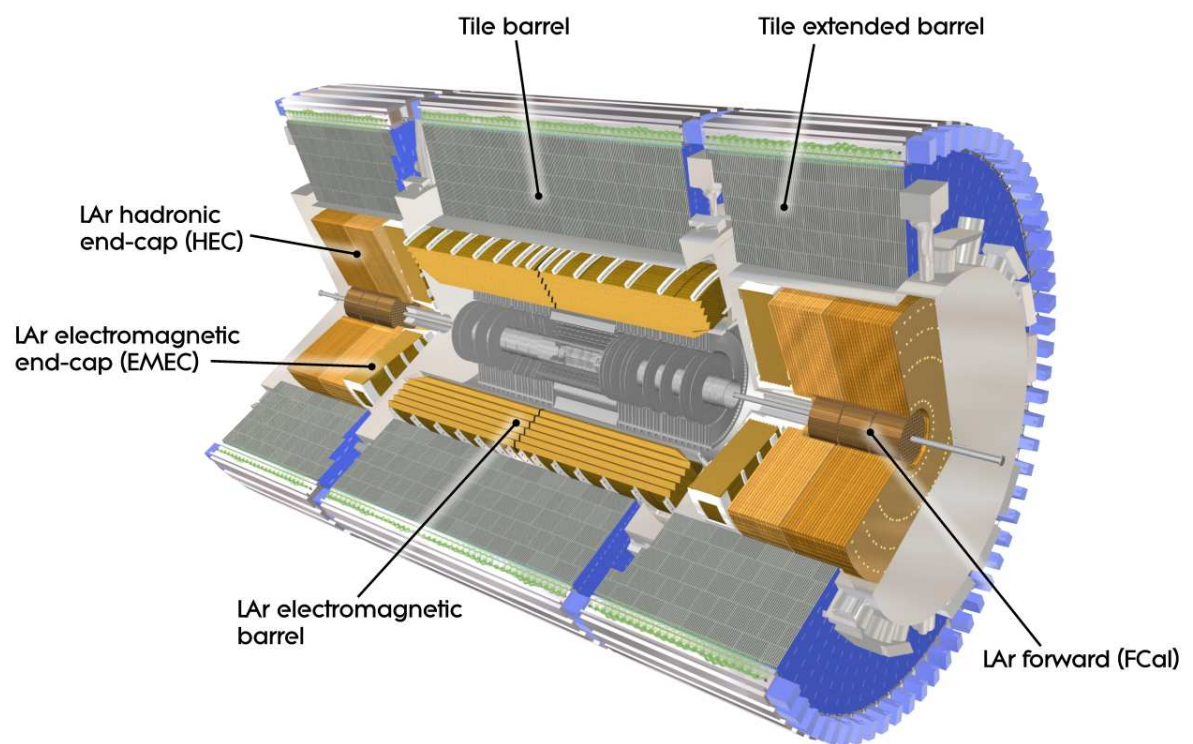The detector consist of three sampling layers: the strip layer, the middle and the back. The electronic readouts are organised in a way, that the volume of detector is divided into virtual cells - elementary energy deposits. The electromagnetic shower is registered in the detector as a cluster of such cells and its energy is measured through summing over energies deposited in each cell. A segment of ECAL scheme is presented in figure 19.
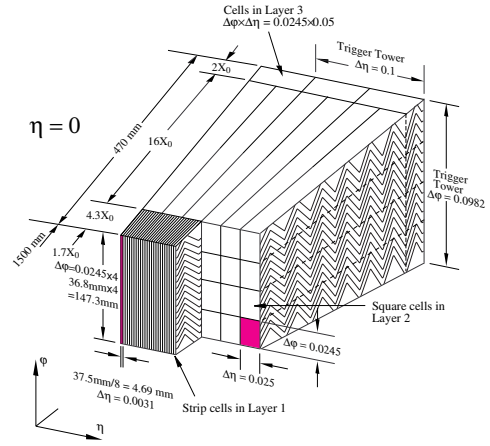

Figure 18: Photograph of a partly stacked barrel electromagnetic LAr module

Figure 19: Sketch of an electromagnetic LAr barrel module

The ECAL energy resolution as function of energy is

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus b, \tag{32}$$

where $a = 10 \pm 0.1\%$ is the stochastic term, $b = 0.17 \pm 0.04\%$ is the constant term and E is expressed in GeV.

2.2.4.2 Hadronic calorimeters

Hadronic calorimeters are designed to measure energy of hadronic showers (jets). AT-LAS hadronic calorimeter system consists of the tile hadronic calorimeter (Tile-Cal), LAr hadronic end-cap calorimeter (HEC) and LAr forward calorimeter (FCal).

Tile calorimeter

Tile calorimeter is a sampling calorimeter that consists of iron plates as absorber material and plastic scintillator tiles as active material. It is composed of a central barrel, covering the region $|\eta| < 1.0$, and two extended barrels, covering the range $0.8 < |\eta| < 1.7$.

Scintillation light produced in the tiles is going through wavelength shifting fibres to photomultiplier tubes (PMTs), where the resulting electronic signal is measured. Figure

20 is showing integration of the mechanical assembly and the optical readout of the tile calorimeter.
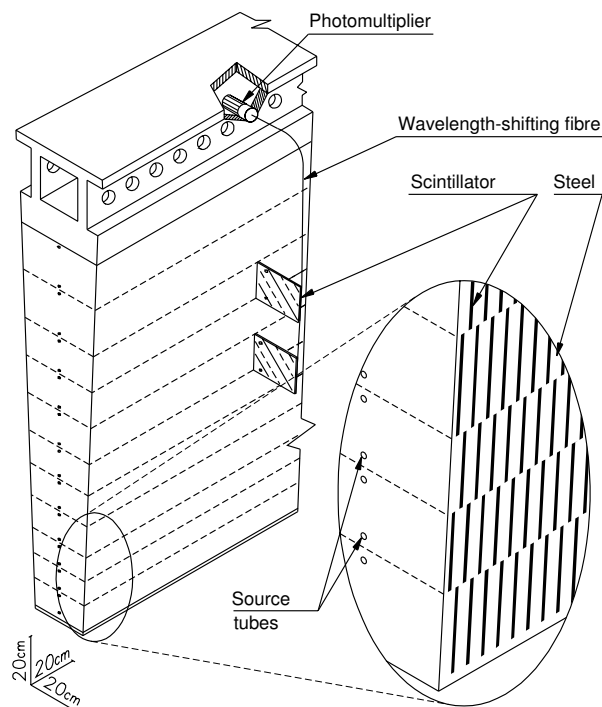

Figure 20: Scheme of the mechanical assembly and the optical readout of the tile calorimeter [22]

LAr hadronic end-cap calorimeter

Hadronic end-cap calorimeter (HEC) is using liquid argon as active material and copper as absorber. It is used to measure the energy of particles in the range $1.5 < |\eta| < 3.2$. HEC consists of two end-cap sections, located behind the electromagnetic calorimeter end-caps in the same cryostat. Each of two hadronic end-cap sections is composed of four layers: two independent wheels, additionally divided into two segment in depth. The wheels that are located closer to interaction point are built of 25 mm copper plates, the outer wheels are using 50 mm copper plates. The 8.5 mm gaps between the plates are filled with liquid argon.

LAr forward calorimeter

The forward calorimeter (FCal) is detecting particles in the very forward region in the range $3.1 < |\eta| < 4.9$. It consists of two sections (one per end-cap), located inside EMEC and HEC. Each FCal section consists of three modules: first is made of copper and designed for electromagnetic measurements, second and third are made of tungsten and measure energy of hadronical showers.

2.2.5 Muon spectrometer

Muons are the only particles detectable by ATLAS, that are getting out of calorimeter. They are measured by the muon spectrometer (MS) - the outermost section of the detector. This detector is designed for identification of muons, reconstruction of their tracks and precision measurement of their momenta. Presence of a muon with certain characteristics can be a sign of an interesting for physics process, therefore information from muon spectrometer is used by ATLAS trigger.

The muon spectrometer is located in the magnetic field formed by toroid magnets: the barrel toroid in the range of $|\eta| < 1.4$ and two end-cap toroids in the range of $1.6 < |\eta| < 2.7$. In the range in between, referred to as transition region, $1.4 < |\eta| < 1.6$, tracks are bent by both barrel and end-caps toroids.

MS consists of four sub-detectors: Monitored Drift Tubes (MDT), Cathode Strip Chambers (CSC), Resistive Plate Chambers (RPC), and Thin Gap Chambers (TGC). Those are muon chambers of two types: first two are designed for track precision measurement (MDT in the barrel and CSC in the end-caps), the other two - for providing the trigger information (RPC in the barrel and TGC).

Scheme of the muon system is presented in figure 21.
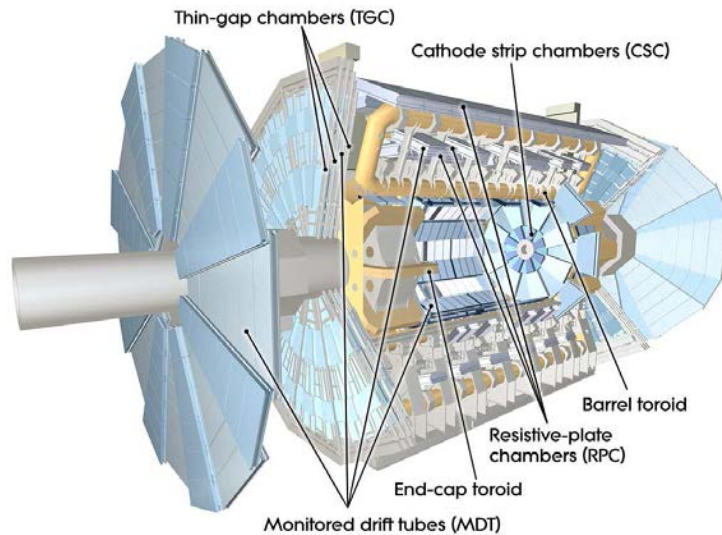
Figure 21: Cut-away of the ATLAS muon system [22]

Tracking chambers

The MDT sub-detector is composed of three to eight layers of drift tubes with gas mixture of Ar (97%) and CO_2 (3%). It covers the range $|\eta| < 2.7$. MDT chambers feature very good spatial resolution 80 μm per tube and 35 μm per chamber, but their counting rate is limited by 150 Hz/cm_2. Therefore in the forward region ($2 < |\eta| < 2.7$) in the innermost tracking layer CSCs are used that provide higher rate capability - up to 1000

Hz/cm$_2$, and better time resolution: their maximum drift time for signal collection is 40 ns compared to 700 ns for the MDTs.

The CSCs are proportional chambers using the same gas mixture as MDT with multiple anode wires oriented in the radial direction and cathodes segmented into strips in orthogonal direction. This allows to reconstruct muons tracks in (η, ϕ) space. The whole CSC system is composed of two disks with eight chambers per disk. Each chamber consists of four planes, which allows to perform four independent measurements in (η) and (ϕ) along each track. Spacial resolution of CSCs if 40 μm in the bending (η) plane and 5 mm in the non-bending (ϕ) plane.

Triggering chambers

The triggering chambers of MS are designed to provide to the trigger system fast and coarse information on muon tracking. Apart from that, the goal of these chambers is to perform track measurement in the non-bending (ϕ) plane additionally to the one done by tracking chambers.

The RPCs are placed in the barrel and cover range of $|\eta| < 1.05$. They consist of two parallel electrode plates (no wires is used), with a 2 mm gap between filled with a gas mixture of $C_2H_2F_4$/Iso-C_4H_{10}/SF_6. The electric field of 4.9 kV/mm is formed between the plates that causes avalanches forming along the ionising tracks. The RPCs allow a good timing resolution of 1.5 ns.

The TGCs are placed in the end-cap wheels, covering the region $1.05 < |\eta| < 2.7$ (2.4 for triggering). They are based on multiwire proportional chamber technique with the copper wires oriented in the radial direction and carbon strips - ϕ direction. Each chamber is filled with a gas mixture of 55% CO_2 and 45% n-C_5H_{12}. TGCs are featuring the wire-to-cathode distance to be smaller than the wire-to-wire distance (1.4 mm compared to 1.8 mm), which allows very fast signal collection, that allows to achive time resolution of 4 ns.

2.2.6 Trigger system

The high luminosity of the LHC requires fast and effective selection of events that are interesting for physics. This is performed by three-level ATLAS trigger system. The scheme of ATLAS trigger and data-acquisition (DAQ) is presented in figure 22

The level-1 trigger (L1) is hardware based: the selection of events is performed by logical electronics. It is decreasing event rate from 40 MHz of the whole collision data to 100 kHz.

L1 is using information from coarse-granularity calorimeter and muon spectrometer. The decision is taken based on the E_T or p_T threshold and the multiplicity of physical objects registered in detectors: electrons, muons, photons, jets, hadronic tau and E_T^{miss}.

The level-2 (L2) and level-3 (L3, event filter) are software based. In Run 1 the L2 trigger reduced the rate of events down to 2-3 kHz and then the L3 was making the final decision, decreasing the rate of events down to 300-400 Hz. In Run 2 these two triggers are merged into a High Level Trigger System (HLT) farm that is using multivariate analysis
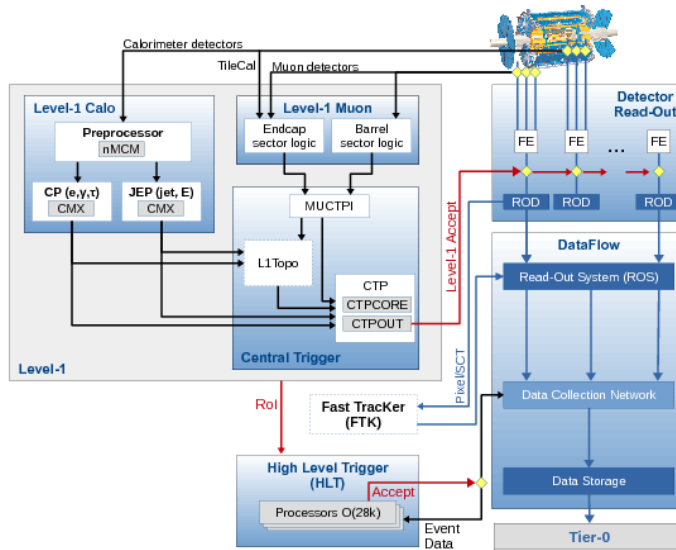
Figure 22: ATLAS trigger and data-acquisition (DAQ) system [35]

techniques. The new approach and various optimisations of the trigger system allowed to simplify and fasten the selection process, which is crucial for high luminosities at Run 2.

Using full event information from different detectors the HLT reduces event rate to 600 Hz to 1.5 kHz.

3   The ATLAS experiment

3.1   Event simulation

Simulation of various physical processes that is used for ATLAS analyses is performed based on theoretical prediction via random number generation - Monte Carlo (MC) method. A physics analysis is based on comparison of theory with experimental data, therefore it is very important to make a good choice of MC modelling for particular processes considered in this analysis.

The full scheme of ATLAS simulation software is presented in figure 23.
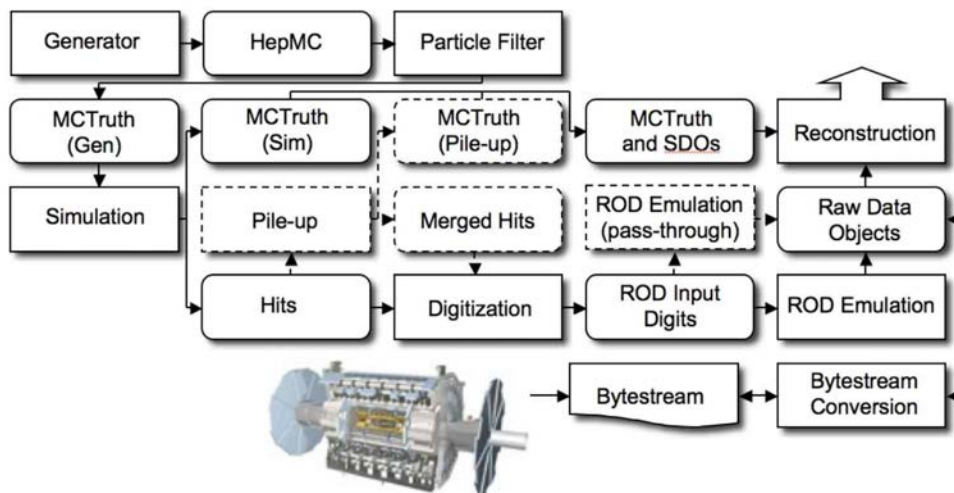

Figure 23: The flow of the ATLAS simulation software from event generators (top left) to reconstruction (top right). Algorithms are placed in square-cornered boxes and persistent data objects are placed in rounded boxes. The optional pile-up portion of the chain, used only when events are overlaid, is dashed. Generators are used to produce data in HepMC format. Monte Carlo truth is saved in addition to energy depositions in the detector (hits). This truth is merged into Simulated Data Objects (SDOs) during the digitization. During the digitization stage, Read Out Driver (ROD) electronics are simulated. [51]

3.1.1   Event generation

According to the factorization theorem, several sub-processes of a physical event that happen at different energy scales can be separated. Thanks to that, simulation of a physical process can be performed in several steps: hard scattering, parton shower and hadronization. The scheme of different steps of event simulation is shown in figure 24.

The momenta of the hard scattering partons inside the colliding protons are obtained via so parton distribution functions (PDFs), that determine probabilities to find a parton of a certain type carrying a certain fraction of proton momentum. The perturbative QCD does not describe the parton content of the proton, therefore the PDFs are obtained from a fit to experimental data in various physical processes.
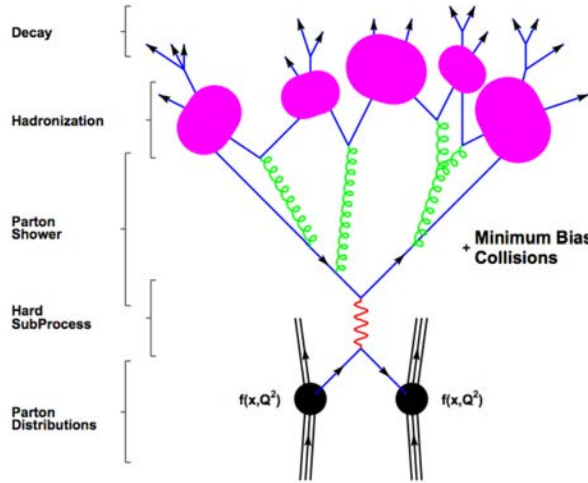
Figure 24: Different steps of a *pp* collision event simulation

The information on the kinematics and flavour partons is then used for the matrix element (ME) calculation - evaluation of the hard process cross section in fixed order perturbation theory. This part of the simulation in ATLAS is performed by so called parton-level generators (or matrix element generators): Alpgen, MadGraph or Powheg.

The second step of simulation is modelling of the parton shower process: partons radiate gluons, which can then further split into other gluons or quark-antiquark pairs etc. A parton shower generator provides a higher order correction to the matrix element calculation due to this radiation: it simulates the emission of quarks and gluons from the partons in the final or initial state.

The third step of event modelling is hadronisation or fragmentation - the partons in the shower, which have reached non-perturbative energy scales, are forming hadrons that further decay. Hadronisation process is simulated using phenomenological models.

For simulation of the parton showering and hadronisation Pythia or Herwig are usually used in ATLAS.

3.1.2 Detector simulation

The final step after generating events is simulation of interaction of the particles with detector. There are two techniques used in ATLAS that are known as fast simulation and full simulation.

The full simulation approach is performed via Geant4 algorithm, that reproduces the interaction between the particles and detector material, resulting in hits in the detector. After that digitisation of hits is performed and the response of detector is simulated. The final step is obtaining tracks and calorimeter clusters using information on the response of the detector. Detailed description of the GEANT4 platform can be found in [50].

Full simulation consumes much computational power, which makes it impossible to obtain the required MC statistics for many physics studies. The alternative fast simulation approach can be used instead in order to save computing resources. The fast simulation

that in used in various ATLAS physics analyses is performed by ATLFAST-II. The reduction of simulation time is achieved by simplifying the detector description used for reconstruction. ATLFAST-II provides the same output as full simulation, which is used in physics analyses in those cases with not enough MC statistics with full simulation [51].

3.2 Event reconstruction

The events that pass trigger selection are further processed to identify which particles were detected. Various algorithms are used to reconstruct physical objects, such as photons, electrons, muons, taus and jets, using tracks and energy clusters registered by sub-detector systems. Neutrinos are not detected, but imbalance of the total transverse momentum in the event tells about presence of neutrino.

3.2.1 Tracks

Charged particles tracks are bended in solenoid magnetic field of the inner detector, obtaining a curvature inversely proportional to their momenta. Reconstruction of a track is performed by extrapolating the hits made by a particle in detector layers with a helix.

There are several track reconstruction algorithms used in ATLAS. The inside-out pattern recognition algorithm is firstly building track seeds in pixel and SCT detectors and then extending the track candidate to the TRT. Most of the tracks used in physics analyses are found with this algorithm.

The outside-in, or back-tracking algorithm is starting with segments finding in the TRT, which are then extrapolated back to the pixel and SCT detectors.

A track in the magnetic field is described by five parameters $(d_0, z_0, \phi, \theta, q/p_T)$, as shown in figure 25.

- azimuthal angle ϕ - the angle between p_T and the x-axis in the transverse plane, $0 < \phi < \pi$

- Signed transverse impact parameters d0 - the closest distance between the track and the beam axis in the transverse plane

- Polar angle θ - the angle between $\vec{p}$ and the z-axis in the r-z plane, $0 < \theta < \pi$

- Signed longitudinal impact parameter z0 - z coordinate of the track at the point of the closest approach

- The electric charge over the transverse momentum q/p_T. It is obtained using the track curvature radius R_{curv}, which is measured by interpolating hits in the detector. The dependence of q/p_T on R_{curv} is defined as

$$q/p_T = (0.3BR_{\mathrm{curv}})^1,$$

(33)
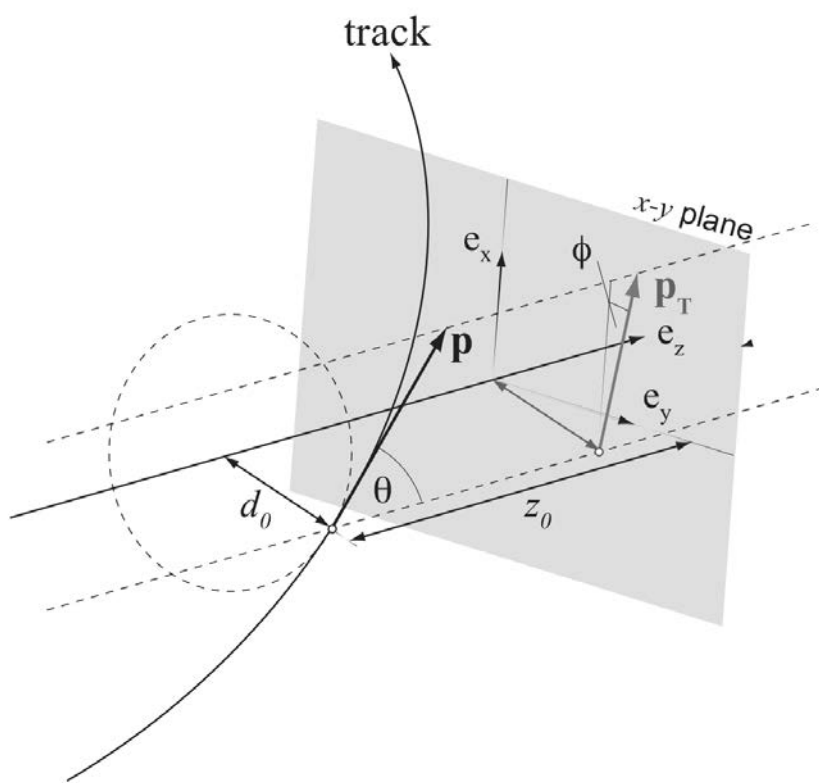
where B is the magnetic field.

Figure 25: Track parameters [37]

3.2.2 Vertices

Extrapolating the tracks allows to determine the position of the vertex - point of initial interaction of protons. But due to the high number of protons per bunch crossing, several vertices can be found for the same event.

First part of vertex reconstruction is the vertex finding algorithm, which is selecting vertex seeds by looking at the local maximum in the distribution of z_0 of the tracks. The second step is the vertex fitting algorithm, that specifies the position and uncertainty of the primary vertex. It takes as input the seed position and the tracks around it and performs χ^2 based fit. Those tracks that are not likely to be coming from the vertex are not rejected, but downgraded. The fit procedure is repeated several times, and outlying tracks are progressively downgraded at each next iteration. Tracks which are incompatible with the vertex by more than 7σ are used to seed a new vertex. [38]

The vertex with highest sum of the squared p_T of its tracks is considered as corresponding to hardest process in the event, others are assumed to be pile-up interactions.

Vertices displaced from the beam collision region are considered as secondary vertices. The secondary vertex reconstruction algorithm that is used for identification of b-jets is described in section 4.1.2.

3.2.3 Electrons

Electron identification is challenging, as hadronic jets and non-prompt electrons (converted photons) behave in detector in a similar way.

Electron reconstruction

Electrons are reconstructed combining information from the inner detector (ID) and the electromagnetic (EM) calorimeter. An electron candidate is build using information on clusters in EM calorimeter that are matched to tracks from ID.

The reconstruction procedure consists of several steps.

- Cluster reconstruction is performed via so called sliding window algorithm. The first step of this algorithm is known as tower building. The calorimeter space is split in squares 0.025×0.025 in the $\eta - \phi$ plane, and for each of them the energy deposits in all EM calorimeter layers are summed. After that a rectangular window is scanning across the elements of towers in order to find seeds with total cluster transverse energy above 2.5 GeV. Then the clusters are reconstructed around the seeds using the clustering algorithm.

- Track reconstruction consists of two steps: pattern recognition and track fit. Pattern recognition algorithm is searching for a track seed, consisting of three hits in different layers of the silicon detectors, with $p_T > 1$ GeV that can be successfully extended to a full track of > 7 hits, so that is matched to a given EM cluster region of interest. (A region of interest is usually defined as a cone-size of $\Delta R = 0.3$ around the seed cluster barycentre). This procedure is applied one or two times per track. For the

first time, the search is done with standard ATLAS pattern recognition that uses the pion hypothesis for energy loss due to interactions with the detector material. If no tracks found, that are consistent with the pion hypothesis, the same procedure, but considering the electron hypothesis is performed. The next step is χ^2 based fit of track candidates (with either pion or electron hypothesis according to the one used in pattern recognition). If fit with pion hypothesis fails for an electron candidate, the second attempt is done with electron hypothesis. This approach of using two hypotheses allowed to make electron track reconstruction complementary to the main track reconstruction procedure, without rerunning the algorithm for all electron track candidates. At the same time the electron performance is improved due to usage of electron-based algorithm for those tracks that cannot be reconstructed via procedure based on pion hypothesis.

- Electron track fit. The electron track candidates that are having ≥ 4 hits are matched to EM clusters in the calorimeter in η and ϕ plane under loose requirement, additionally taking into account energy loss due to bremsstrahlung. The fit is done for those that pass matching criteria. After that matching procedure is repeated with refit tracks under stricter criteria.

Electron candidates that are not matched to a track are removed and considered to be photons.

For Run 2 electron tracks candidate are required to be compatible with the primary vertex of the hard collision, in order to reduce the background from conversions and products of long-lived particle decay, as well as pile-up interactions. Therefore additional restrictions on track impact parameters are applied:

$$d_0/\sigma_{d_0} < 5, \tag{34}$$

and

$$\Delta z_0 \sin\theta < 0.5 \text{ mm}, \tag{35}$$

where d_0 and z_0 are transverse and longitudinal impact parameters, θ is the polar angle of the track (see definitions in 3.2.1) and σ_{d_0} represents uncertainty on d0. d_0/σ_{d_0} is referred to as d_0 significance.

Electron identification

Electron identification algorithms were developed to distinguish electrons of background - hadronic jets and background (non-prompt) electrons mostly originating from photon conversions and heavy flavour hadron decays.

The baseline algorithm for electron identification in Run 2 is based on the likelihood approach. It evaluates several properties of electron candidates using signal and background probability density functions (PDFs). The input variables for this algorithm are quantities related to the electron cluster and track measurements including calorimeter shower shapes, information from the TRT, track-cluster matching related quantities, track properties and variables measuring bremsstrahlung effects. In Run 2 the number of IBL hits

is also used as input, as it helps to discriminate between electrons and converted photons. The full list of variables can be found in [39].

The signal and background probabilities for an electron candidate are evaluated and combined in the final likelihood discriminant:

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B},$$ (36)

where signal and backround probabilities are

$$\mathcal{L}_{S(B)}(\vec{x}) = \prod_{i=1}^{n} P_{s(b),i}(x_i).$$ (37)

In the last equation $\vec{x}$ is the vector of discriminating variable values and $P_{s(b),i}(x_i)$ is the value of the signal (background) probability density function of the ith variable.

Based on the final discriminant $d_{\mathcal{L}}$ distribution three working points are defined, in order of increasing background rejection: Loose, Medium, and Tight. The electron identification efficiency and efficiency of background identification (hadrons identified as electrons) as functions of E_T for these three working points are presented in figure 26.
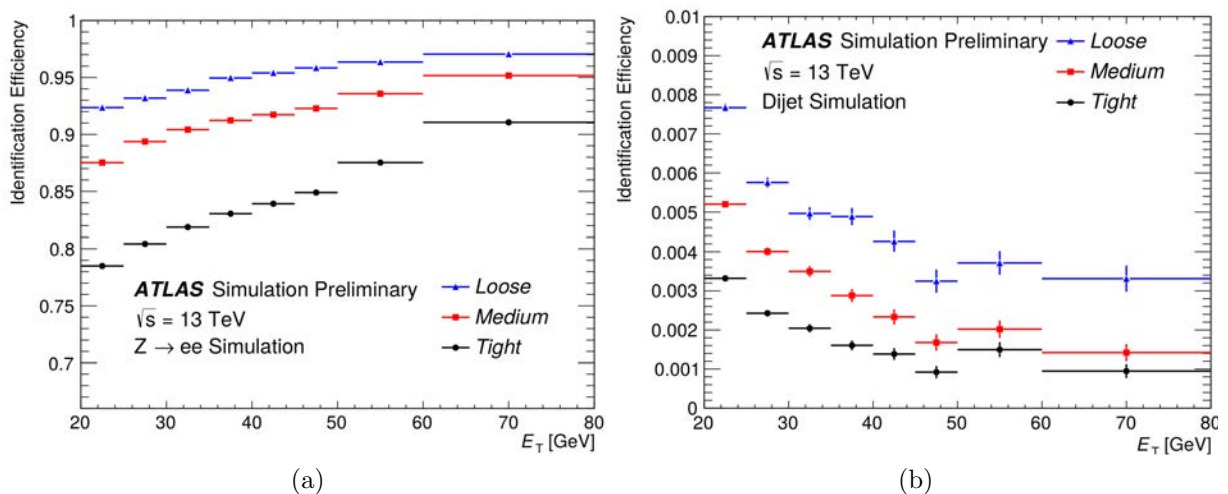

Figure 26: Efficiency of prompt electron identification obtained from $Z \to ee$ decays (a) and efficiency to identify hadrons as electrons from simulated dijet samples (b). The efficiencies are obtained using MC simulations, and are measured with respect to reconstructed electrons [39].

3.2.4   Muons

For reconstruction of muons the measurement from muon spectrometer (MS) is combined to the one from inner detector (ID). Information from EM calorimeter is additionally used.

Muon reconstruction in MS

The first step of the muon reconstruction in the MS is search for hit patterns in each of the muon chambers and building segments out of them. Then the muon track candidates are obtained by fitting together hits from segments in different detector layers. The minimum requirement to build a track is two matching segments, except in the barrelendcap transition region, where one segment is allowed to be used. Then hits associated with a track candidate are fitted with χ^2 based fit.

Combined reconstruction

Track measurement in the MS can be then combined with information from ID and EM calorimeter. Four types of reconstructed muons are used in ATLAS:

- Standalone muons - those reconstructed only in the muon spectrometer (MS)

- Combined muons - using information from both MS and inner detector (ID).

- Calorimeter-tagged muons - muons detected by calorimeter and ID, without using information from MS. They are used for the region $\eta < 0.1$, that is uncovered by MS.

- Segment-tagged muons - when track from ID is matching to a segment of a track in the MS

- Extrapolated (ME) muons - tracks reconstructed using hits in MS detectors only, but with additional requirement on compatibility with originating from the interaction point

Combined muons is the best quality muon type: they have the highest fake muons rejection and the best momentum resolution.

There are two main combined muon reconstruction algorithms: Staco and MuID. Staco is using independently reconstructed tracks from MS and ID and provides their statistical combination based on χ^2 criterion [41]. Muid is performing a combined refit to the hits of both tracks - from MS and ID. Energy loss due to material interaction is additionally taken into account in the fit, using information from calorimeter [42].

Muon identification

After muons are reconstructed, the next challenge is to distinguish prompt muons of those originating from decays of charged hadrons, mostly from pion and kaon decays. For combined muons several discriminating properties are used:

- q/p significance - the absolute value of the difference between the ratio of the charge q and momentum p of the muons measured in the ID and MS, divided by the sum in quadrature of the corresponding uncertainties;

- ρ' - the absolute value of the difference between the p_T in the ID and MS divided by the p_T of the combined track;

- normalised χ^2 of the combined track fit.

Based on distributions of discriminating variables and additional requirements on number of hits in the ID and MS four muons selections are provided: Loose, Medium, Tight and High-p_T. The muon reconstruction efficiencies for these working points were evaluated using $t\bar{t}$ MC sample. In the table 6 efficiencies of signal (muons candidates from W decays) and background (muon candidates from light-hadron decays) are presented [40].

	$4 < p_T < 20$ GeV		$20 < p_T < 100$ GeV	
Selection	$\epsilon_\mu^{MC}[\%]$	$\epsilon_{\text{Hadrons}}^{MC}[\%]$	$\epsilon_\mu^{MC}[\%]$	$\epsilon_{\text{Hadrons}}^{MC}[\%]$
Loose	96.7	0.53	98.1	0.76
Medium	95.5	0.38	96.1	0.17
Tight	89.9	0.19	91.8	0.11
High-p_T	78.1	0.26	80.4	0.13

Table 6: Efficiencies of identifying prompt muons from W decay (signal) and decaying in-flight hadrons misidentified as prompt muons (background) computed using a $t\bar{t}$ MC sample. The results are shown for the four muons selections (Loose, Medium, Tight and High-p_T) for low and high p_T muon candidates with $\eta < 2.5$ [40].

3.2.5 Jets

Quarks and gluons produced in pp collision are hadronising and produce in the detector collimated bunches of particles - hadronic jets. Reconstruction of a jet allows to measure the momentum of the initial parton as each of final state particles carry some fraction of it.

The anti-k_T algorithm is used in ATLAS to reconstruct jets from topological clusters in calorimeter (known as topo-clusters). This algorithm is described in details in [43].

First of all, the algorithm is finding a seed cell with a signal-to-noise ratio is above the threshold $S/N \geq 4$. The noise can have electronic or pile-up origin. Then the cells surrounding the seed are iteratively attached to the seed if they satisfy the requirement $S/N \geq 2$. Finally, the cells lying on the perimeter of resulting cluster with $S/N \geq 0$ are also included to make the edges of the formed cluster smoothed. If there is more than one energy local maximum in a cluster, it can be split into several sub-clusters. This helps to separate jets, that are close, but originating by different particles.

An example of jets clusters from a MC simulated parton-level event reconstructed with the anti-k_T algorithm is shown in figure 27.

Calibration of the cells is performed by local clustering (LCW) method, that allows to discriminate electromagnetic and hadronic topo-clusters. The detailed description of this calibration can be found in [44] and [45].
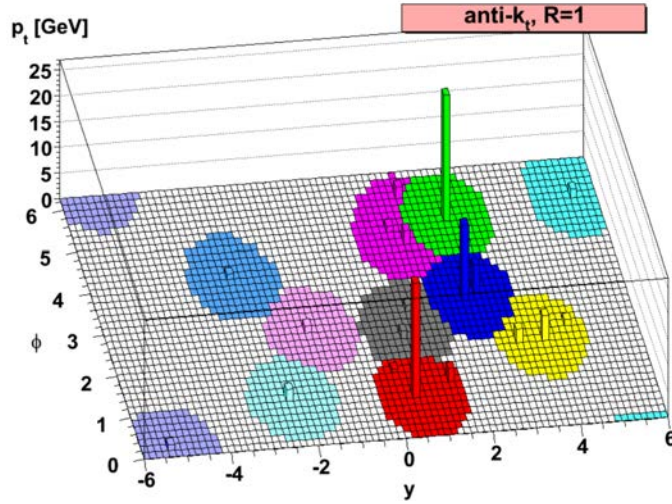
Figure 27: clusters from a MC simulated parton-level event reconstructed with the anti-k_T algorithm [43]

After jet four-momentum is reconstructed from the cluster, corrections are applied to take into accounts various effects. First of all, scale factors that correct the shift in jet energy due to pile-up interactions are applied. Secondly, the jet direction is revised with the requirement to be originated from the primary vertex. Jet energy scale (JES) corrections, that are applied to jet energy and pseudorapidity, allow to take into account energy loss in the detector [46].

Jets that are originating from b-quark (b-jets) can be discriminated of other types of jets, using special properties of B-hadrons. Identification of b-jets, or b-tagging, is important for many physical analyses and in particular for $t\bar{t}H \rightarrow b\bar{b}$. It is described in details in the section 4.1.4.

3.2.6 Missing transverse energy

Neutrinos are not measured by ATLAS detector. Their presence can be determined by imbalance of transverse momentum component of the detected particles, known as missing transverse energy (MET or E_T^{miss}).

The initial partons in hard process event are supposed to have negligible transverse momenta (with respect to the longitudinal component). Due to that for an event with all particles been detected, the transverse component of a vector sum of all particles' momentum is close to 0. If there is imbalance, it is indicating that some of the final state particles were not detected. Among the SM-predicted particles only neutrino is escaping ATLAS detector without interaction. Theories beyond the SM suggest existence of other particles, that participate in weak interaction only, so they can also contribute to the missing energy. Therefore E_T^{miss} measurement is used in exotic search analyses.

Reconstruction of E_T^{miss} takes into account energy deposits in the calorimeters and muons reconstructed in the MS. It is obtained from the negative vector sum of the momenta

of all reconstructed and calibrated physics objects. Their contributions are considered in a specific order: electrons, photons, hadronically decaying τ-leptons, jets and then muons. Soft scale contributions that are not associated with any of these objects (mostly from underlying events and soft radiation) are also considered [47]. If the combined muon momentum is used, the muon energy loss in the calorimeters is subtracted in the calculation in order to to avoid double counting [48].

The x and y components of calculated E_T^{miss} are:

$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{miss,e}} + E_{x(y)}^{\text{miss},\gamma} + E_{x(y)}^{\text{miss},\tau} + E_{x(y)}^{\text{miss,jets}} + E_{x(y)}^{\text{miss,SoftTerm}} + E_{x(y)}^{\text{miss},\mu}. \tag{38}$$

The magnitude of E_T^{miss} is then obtained as

$$E_T^{\text{miss}} = \sqrt{E_x^{\text{miss}} + E_y^{\text{miss}}}. \tag{39}$$

In Run-1 the soft term $E_{x(y)}^{\text{miss,SoftTerm}}$ was calculated using calorimeter information - the soft scale energy deposits not associated with any of reconstructed hard objects. In Run-2 the track-based approach is used: the soft term is calculated from momenta of tracks from the ID. This method is more robust against pile-up interaction contamination, than calorimeter-based approach, as for tracks the pile-up component can be effectively discriminated [49].

The resolution of E_T^{miss} measurement achieved for Run 2 is presented in figure 28.
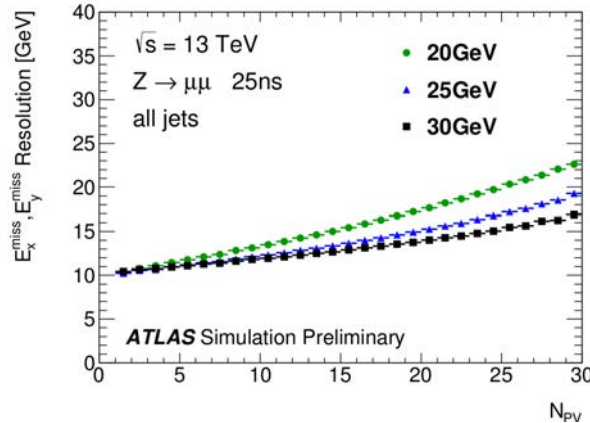

Figure 28: E_T^{miss} resolution in $Z \rightarrow \mu\mu$ MC events for different values of the jet pT threshold, as quantified by (a) the resolution as a function of the number of primary vertices. The E_x^{miss} and E_y^{miss} were found to have similar performance, so two distributions were combined [49]

4 Identification of b-jets

4.1 b-tagging algorithms in ATLAS

Identifying jets containing b-hadrons, a capability known as b-tagging, is useful in various analysis domains with b-jets in final state performed by the ATLAS experiment, like Standard Model measurements (top quark physics and Higgs physics) and beyond the Standard Model searches. In the $t\bar{t}H$ ($H \rightarrow b\bar{b}$) analysis channel there are 6 jets in the final signal state and 4 of them are originating from b-quarks, so it's important to identify them.

For b-tagging various algorithms are used, which rely on special properties of b-hadrons, such as their high mass (≈ 5 GeV) and relatively long lifetime (≈ 1.5 ps).

The goal of b-tagging algorithms is to tag as many b-jets as possible with as high as possible rejection of light jets - jets originating from light (u, d, s) quarks and gluons.

For the Run 2 of LHC with higher center of mass energy (13 TeV) of the proton-proton collisions, the b-tagging algorithms were revisited and optimised. The b-tagging performance during Run-2 was improved thanks to the insertion of a new innermost layer of pixel detector and algorithmic enhancements in both tracking and b-tagging algorithms.

In the ATLAS Run-2 b-tagging scheme there are three basic algorithms:

- Impact parameter-based (IP2D, IP3D), making use of the fact that tracks from the b-hadron decay are not pointing to the primary vertex.

- Secondary vertex finding (SV), reconstructing an inclusive displaced secondary vertex within the jet.

- Decay chain multi-vertex fit (JetFitter), attempting to reconstruct the full b-hadron decay chain.

Several observables from these algorithms are combined with a multivariate algorithm (MV2), that provides the final discriminant between the different jet flavours.

4.1.1 Impact parameter-based (IP2D, IP3D)

IP2D and IP3D algorithms use the signed impact parameter significance of tracks associated to a jet. The impact parameter is the distance of closest approach of a track to the primary vertex (PV). Its sign is defined positive (negative) if the point of closest approach of the track to the primary vertex is in front (behind) the primary vertex with respect to the jet direction as shown at figure 29.

IP2D algorithm is using as input only the transverse impact parameter, while IP3D uses both transverse and longitudinal components and their correlation. Probability density functions (PDFs) of the track's impact parameter are built from simulation for the b- and light-flavour jet hypotheses and are then combined using a log-likelihood ratio method to define a tagging weight for a jet:

$$w_{track} = \frac{p_b}{p_{light}}, \quad w_{jet} = \sum_{tracks} \log w_{track}. \tag{40}$$
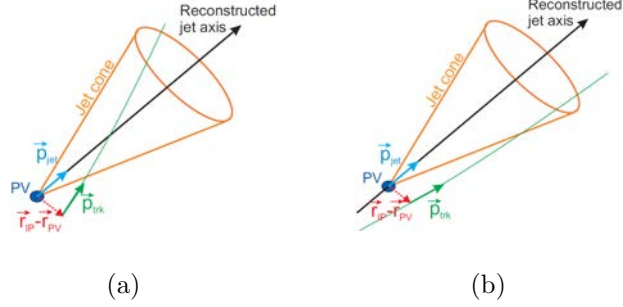
Figure 29: Definition of lifetime of a track: positive (a) and negative (b)

In the IP2D and IP3D algorithms different PDF sets are used for different track categories, depending on the quality of the tracks, which is defined using information on hits in the different silicon layers of the inner detector.

Figure 30 shows transverse (a) and longitudinal (b) impact parameter significance distribution for Run-2 "Good" track category defined in 7. Figure 31 shows the final discriminant of IP2D (a) and IP3D (b).
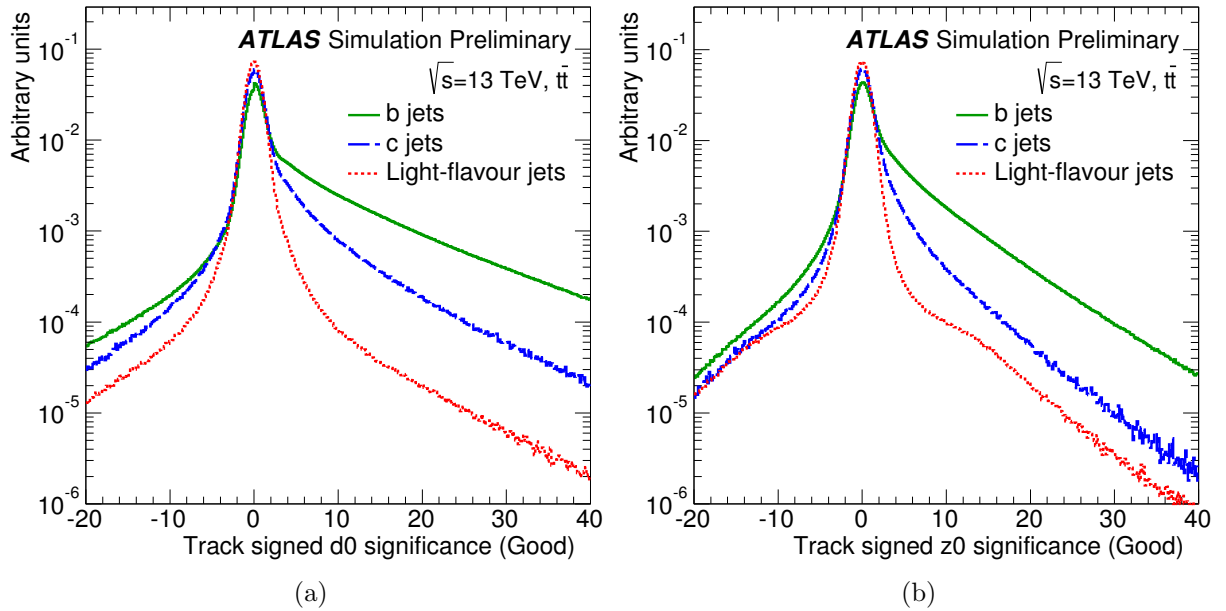

Figure 30: The transverse (a) and longitudinal (b) signed impact parameter significance of tracks in $t\bar{t}$ events for the "Good" category defined in Table 1.

4.1.2 Secondary Vertex Finding Algorithm (SV)

The secondary vertex algorithm attempts to reconstruct the inclusive vertex formed by the decay products of the b-hadron, including those from the subsequent c-hadron decay.
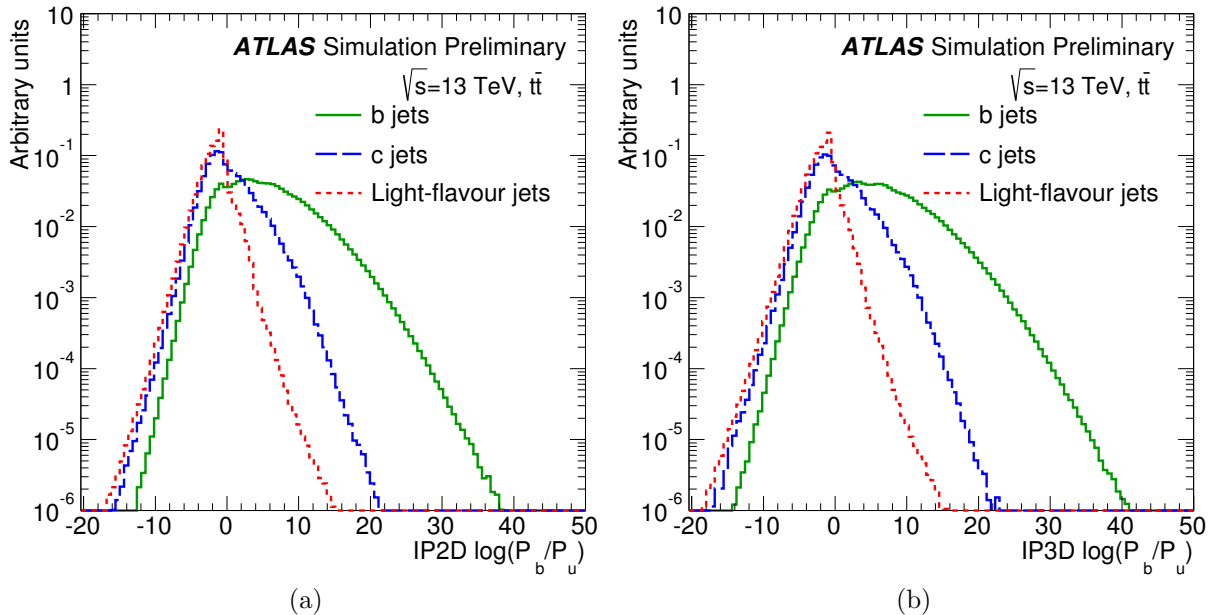
Figure 31: The log likelihood ratio for the IP2D (a) and IP3D (b) b-tagging algorithm in $t\bar{t}$ events.

Firstly it searches for all two-track pairs that form a good vertex, using tracks displaced from the primary vertex. Then the algorithm removes those tracks that are compatible with decays of long lived particles (K_s, Λ etc) or interaction with the detector material. After this selection the algorithm fits an inclusive secondary vertex. Several properties of this vertex are useful to tag b-jets, such as its mass, number of tracks, distance to primary vertex, energy fraction of tracks at vertex with respect to all tracks in the jet.

Figure 32 shows SV vertex reconstruction rate as function of jet p_T and η for b-, c- and light-flavour jets.

4.1.3 Multi-vertex fit (JetFitter)

Another algorithm, called JetFitter, attempts to reconstruct the full PV to b- to c-hadron decay chain. A Kalman filter is used to find a common line on which the primary vertex and all secondary vertices are placed, approximating the b-hadron path. This approach allows to separate b- and c-hadron vertices even if only one track is attached to each of them. Figure 33 shows JetFitter reconstruction rate as function of jet p_T and η for different flavour jets.

4.1.4 Multivariate algorithm (MV2)

Discriminant observables from the above algorithms are combined together into a boosted decision tree (BDT) based algorithm. The default algorithm for Run-2, MV2c20, is a BDT which is trained using b-jets as signal and a mixture of light-flavour jets and c-jets as background (the amount of c-jets in the background is equal to 20% of the amount
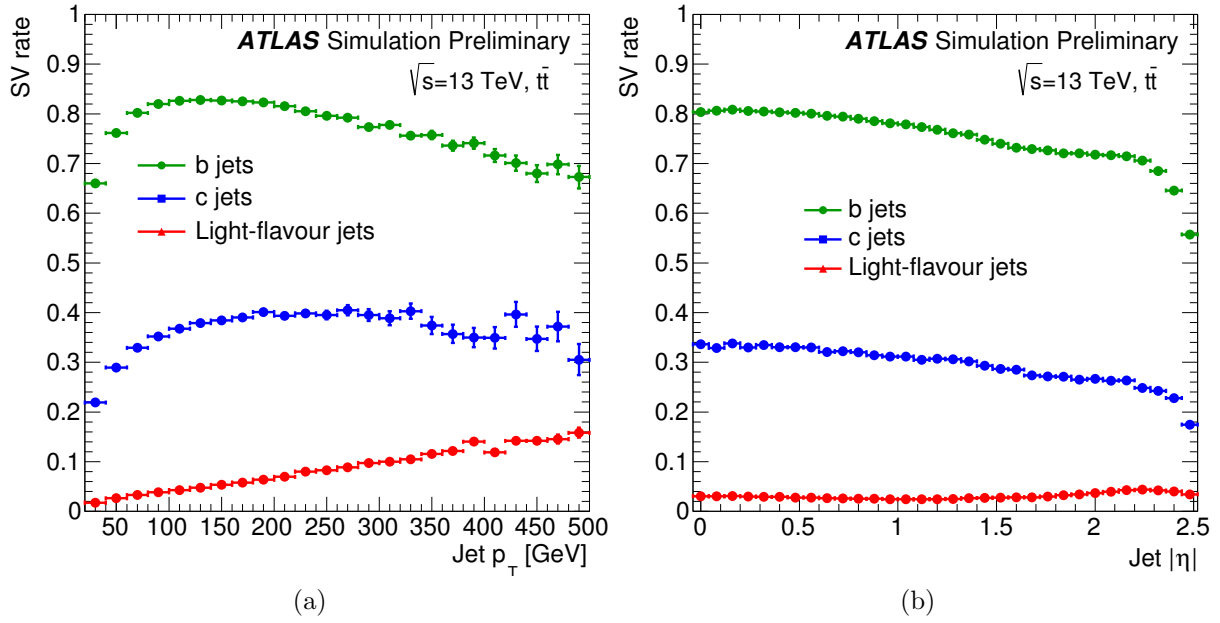
Figure 32: Secondary vertex reconstruction rate as function of jet p_T (a) and η (b) in $t\bar{t}$ events.
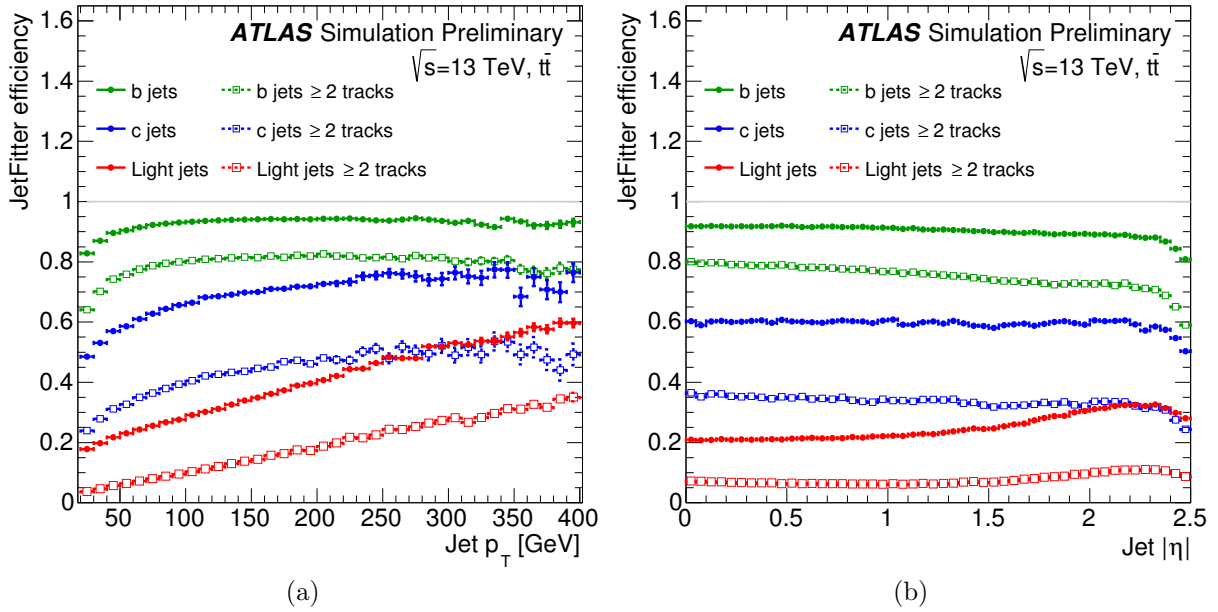

Figure 33: JetFitter vertex reconstruction rate as function of jet p_T (a) and η (b) in $t\bar{t}$ events.

of light-jets). The kinematic properties (p_T and η) of the jets are included in the training to take advantage of correlations with the other input variables.

MV2c20 is an upgrade of Run-1 main b-tagging algorithm MV1, which was combining the outputs of the various b-tagging algorithms using neural network approach.

The MV2c20 algorithm provides better performance and easier retraining and software maintenance.

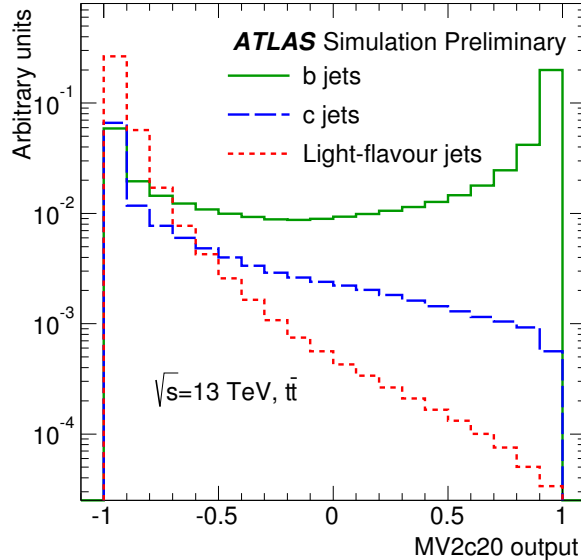Figure 34 shows the MV2c20 output distribution.

Figure 34: The MV2c20 output distribution

4.2 Impact parameter based algorithms optimisation

4.2.1 Introduction

The ATLAS b-tagging algorithms changed significantly at LHC Run-2 with respect to Run-1. One of important developments is IP2D/IP3D algorithms optimisation and in particular new tracking categorisation which gets advantage of the IBL addition. In this chapter the studies on developing the Run-2 b-tagging categorisation and other impact parameter based taggers optimisation is presented.

The studies impact parameter based taggers optimisation were performed using the $t\bar{t}$ events corresponding to 13 TeV protonproton collisions simulated with Powheg+Pythia6 and CT10 parton distribution functions. EvtGen is used to model the decays of b and c-hadrons. Only $t\bar{t}$ decays with at least one lepton from a subsequent W decay are included.

Jets are reconstructed by clustering energy deposits in the calorimeter with the anti-kt algorithm and a radius parameter of 0.4, where clusters are calibrated at the EM-scale and the hadronic scale is obtained through a p_T and η dependent correction factor. Only jets with p_T above 20 GeV and $\eta < 2.5$ are considered.

To reject jets from pileup jet vertex tagger (JVT) algorithm was used. Jets with $p_T < 50$ GeV and $\eta < 2.4$ are rejected if they have a JVT output of less than 0.641. This corresponds to an expected efficiency of about 92% for jets from the hard-scatter and a 2% efficiency for pile-up jets. The JVT selection is close to 100% efficiency for b-tagged b and c-jets.

For simulated jets, a flavour label is assigned by matching jets to the truth-level weakly decaying b and c hadrons with $p_T > 5$ GeV, in a ΔR cone of less than 0.3. The flavour labelling is exclusive, with the hadron matched to the closest jet in the ΔR phase-space. If a b-hadron is found within the cone the jet is labelled as a b-jet. If no b-hadron is found, the search is repeated for c-hadrons, then for τ leptons. If no match is found for b, c, or τ, the jet is labelled as a light flavour jet.

4.2.2 Track categorisation

Some tracks are well reconstructed and therefore having higher resolution of impact parameter significance. But there are also tracks of worse quality: those with missing hit in one of the pixel detector layers, with ambiguities in pattern recognition. Both high and low quality tracks are important for b-tagging, but it is necessary to divide them into categories and treat them separately.

Run-2 track categorisation was developed to be different from the one used in Run-1, making use of several new tracking variables (including those related to the presence of the IBL). Everywhere in this chapter the following notation is used: L0 - the IBL, the innermost pixel detector layer for Run-2; L1 - the next to the innermost layer for Run-2 (which was the innermost layer for Run-1). The variables used for Run-2 categorisation are:

- The number of hits in the innermost (L0) and the next to innermost (L1) layers of pixel detector. Those with missing hit(s) are expected to be worse reconstructed than the others, so they should be treated separately from the "better" tracks.

- Information on whether the hit in L0 and L1 is expected or not (based on the detector coverage and dead module maps). The hit in a layer is expected when the curve obtained by interpolation of hits in other pixel detector layers is crossing a working module of this layer (Figure 36). Therefore, the case when there is no hit in a layer, while it is expected, is indicating that the track is poorly reconstructed. These tracks are supposed to be the "worst" in classification and should be separated from the "better" tracks with no hit in a layer and not expected one (those are mostly tracks crossing a region outside detector coverage or a dead module). This is a new tracking variable, not used in Run-1.

- The number of shared hits (clusters which are shared among more than one track) in L0, L1 as well as in other pixel detector layers and the SCT

- The number of split hits (clusters which have been identified with help of neural network (NN) as coming from different particles and have therefore been split into sub-clusters) in L0, L1 and other pixel detector layers.

To study first impact parameter resolution effects, only tracks from light jets were considered. Track were initially divided into 18 exclusive categories and distributions of transverse and longitudinal impact parameter significance were studied in order to find optimum categorization. These categories are listed below, figure 35 illustrates the impact parameter significance distributions for each category.

1. No hits in first two layers (L0 and L1), while hits are expected in both. The transverse impact parameter distribution is smeared with a double peak structure (Figure 35). The tracks in this category are mostly those coming from long-lived particles (such as K_s, λ), photon conversions and interactions with detector material. These tracks are not having hits in the first layers as the particles they are originating from are decaying beyond these layers (while the hit might still be "expected" if the interpolated curve is crossing working modules of these layers). The impact of each of these different track types is described in 4.2.3.

2. No hits in first two layers, while hit is expected in L0 and not expected in L1.

3. No hits in first two layers, while hit is not expected in L0 and expected in L1.

4. No hits in first two layers and not expected in both. These tracks are mostly lying outside detector coverage. These tracks are better reconstructed than those in categories 1, 2 and 3 where some of the hits are expected, but missing, and the impact parameter resolution is significantly higher (Figure 35).

5. No hit in L0, while expected, hit is present in L1.

6. No hit in L0 and not expected, hit is present in L1. The resolution of the impact parameter resolution is higher, then for category 5.

7. No hit in L1, while expected, hit is present in L0.

8. No hit in L1 and not expected, hit is present in L0. This category has larger fraction of tracks with respect to category 6 because of the dead modules of L1.

9. Shared hit in both L0 and L1

10. Shared hits in L0 (and not in L1)

11. Shared hits in L1 (and not in L0)

12. Shared hits in other pixel layers

13. Two or more shared hits in the SCT

14. Split hits in both L0 and L1

15. Split hits in L0 (and not in L1)

16. Split hits in L1 (and not in L0)

17. Split hits in other pixel layers

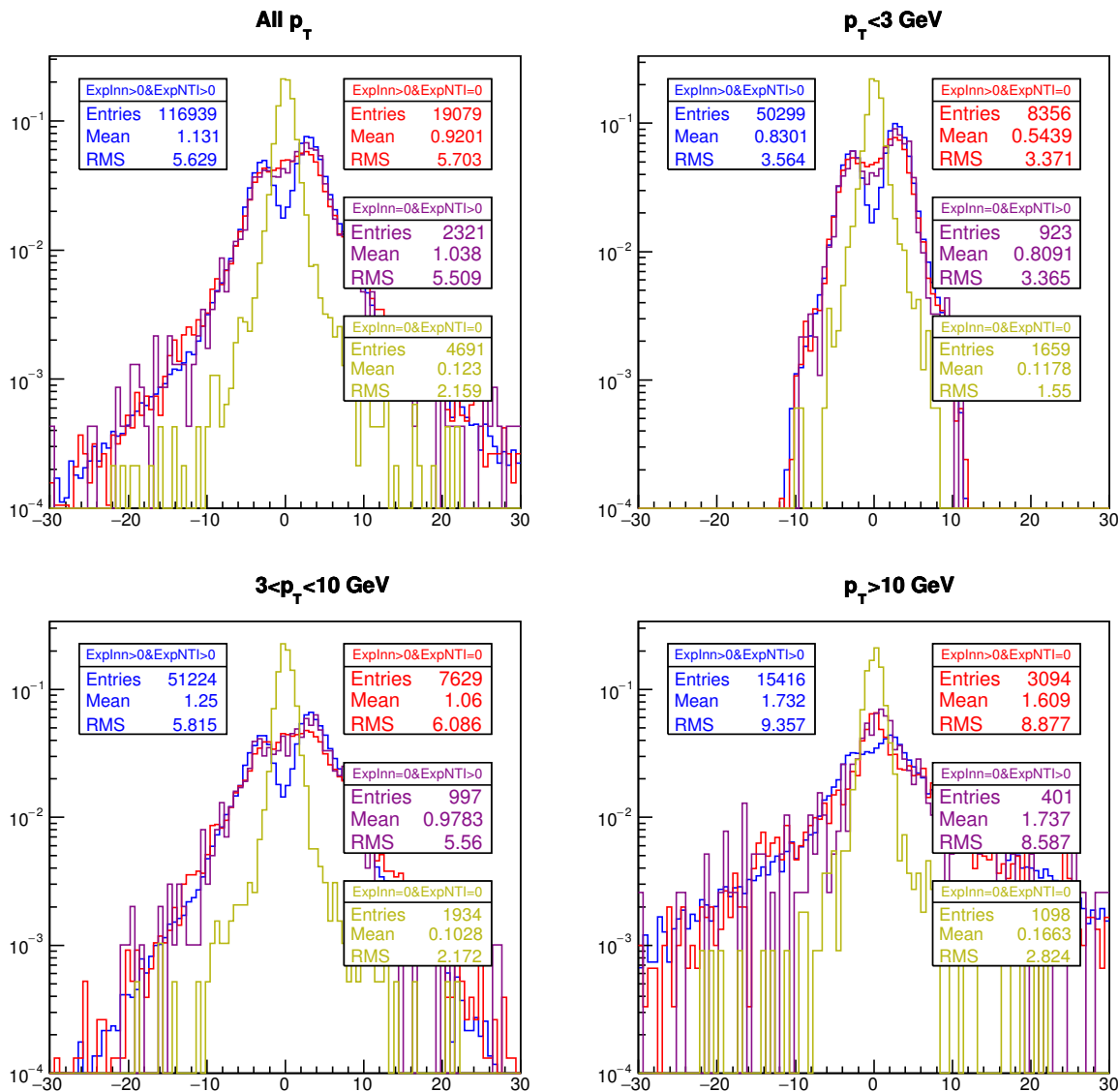18. "Good" tracks: tracks not in any of the above categories

Figure 35: Transverse impact parameter significance of tracks associating with light jets in categories with no hit in neither innermost (L0) nor next to the innermost (L1) layer of pixel detector

Some of the categories with close impact parameter distribution were decided to be merged in order to avoid low statistics categories. So, categories 9, 10 and 11 from the list above were merged into one category ("Shared hit in either L0, or L1, or other pixel layer"), categories 14, 15, 16 were merged into category "Split hit in either L0, or L1, or other pixel layer". Reducing the number of categories from 18 to 14 did not affect the
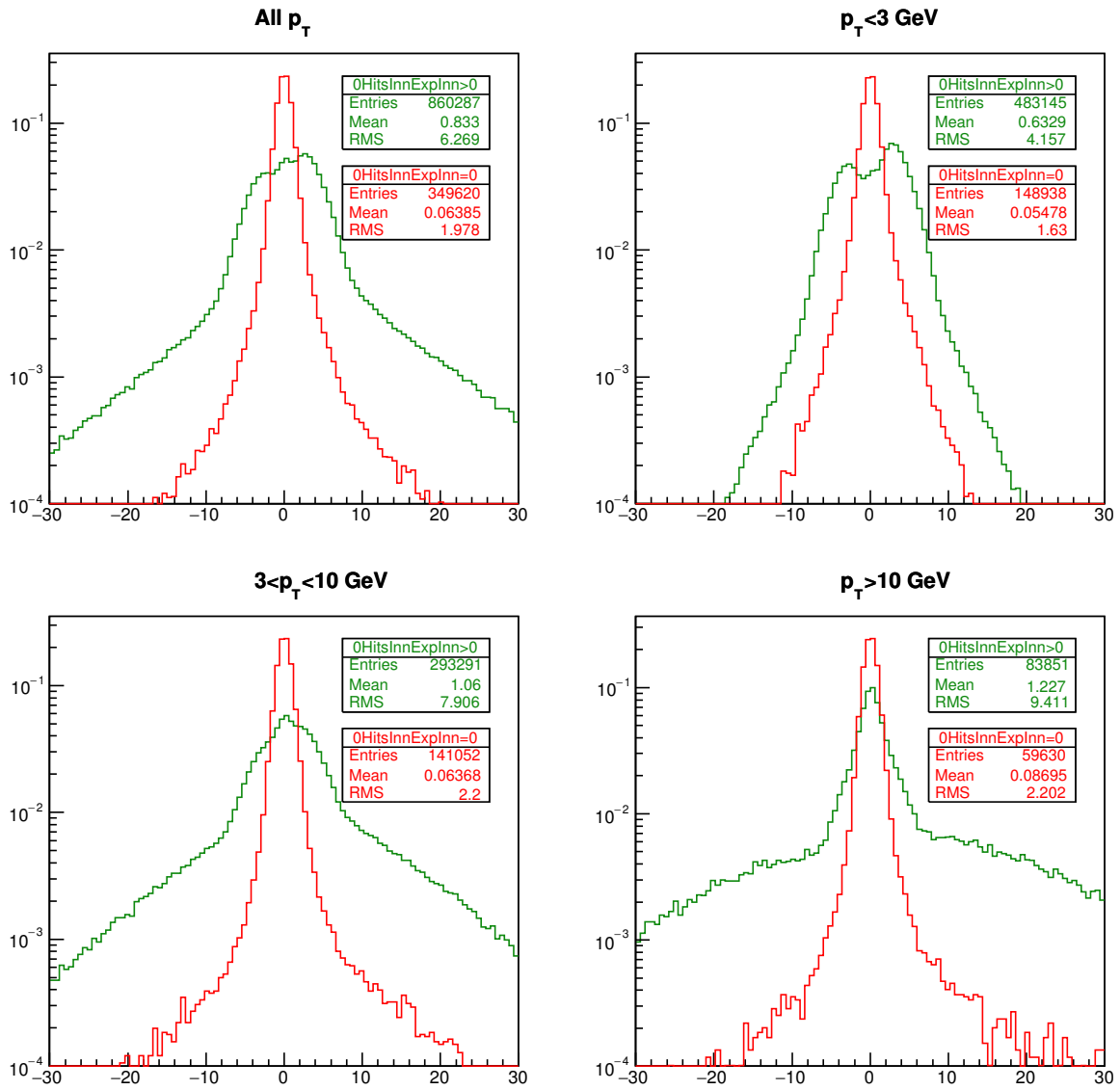
Figure 36: Transverse impact parameter significance of tracks associating with light jets in categories with no hit in the innermost (L0) layer of pixel detector
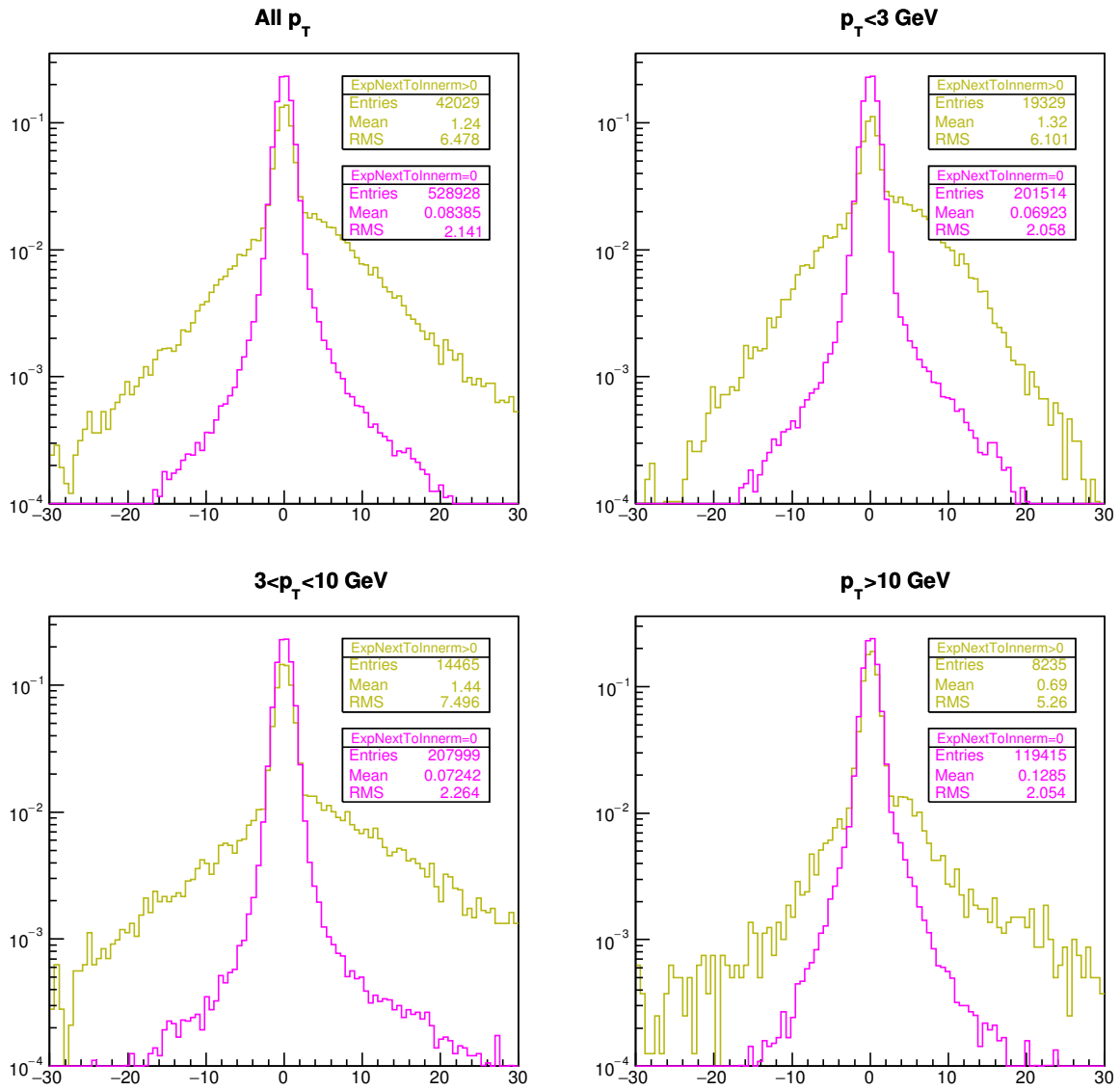
Figure 37: Transverse impact parameter significance of tracks associating with light jets in categories with no hit in the next to innermost (L1) layer of pixel detector
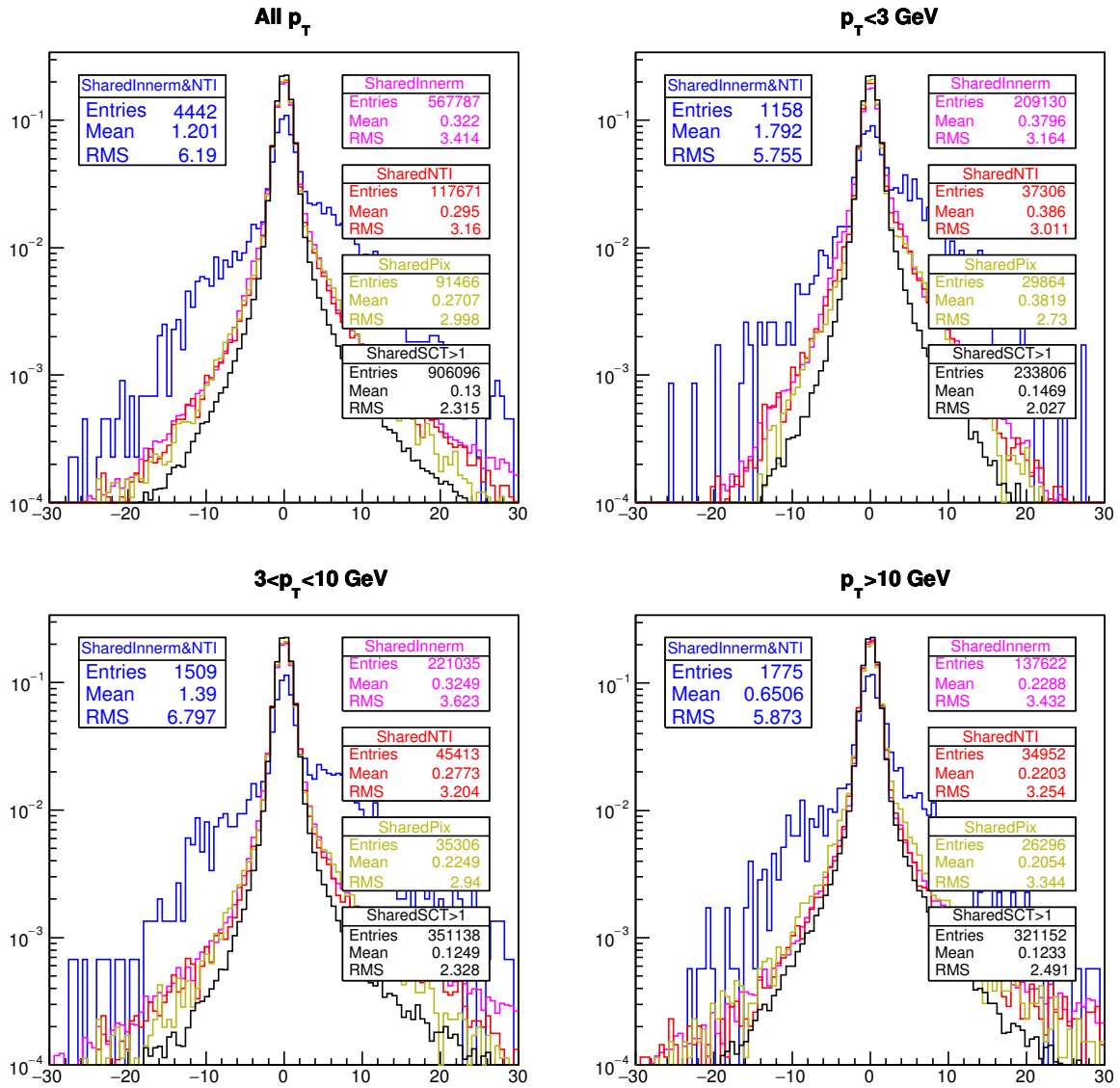
Figure 38: Transverse impact parameter significance of tracks associating with light jets in categories with shared hits
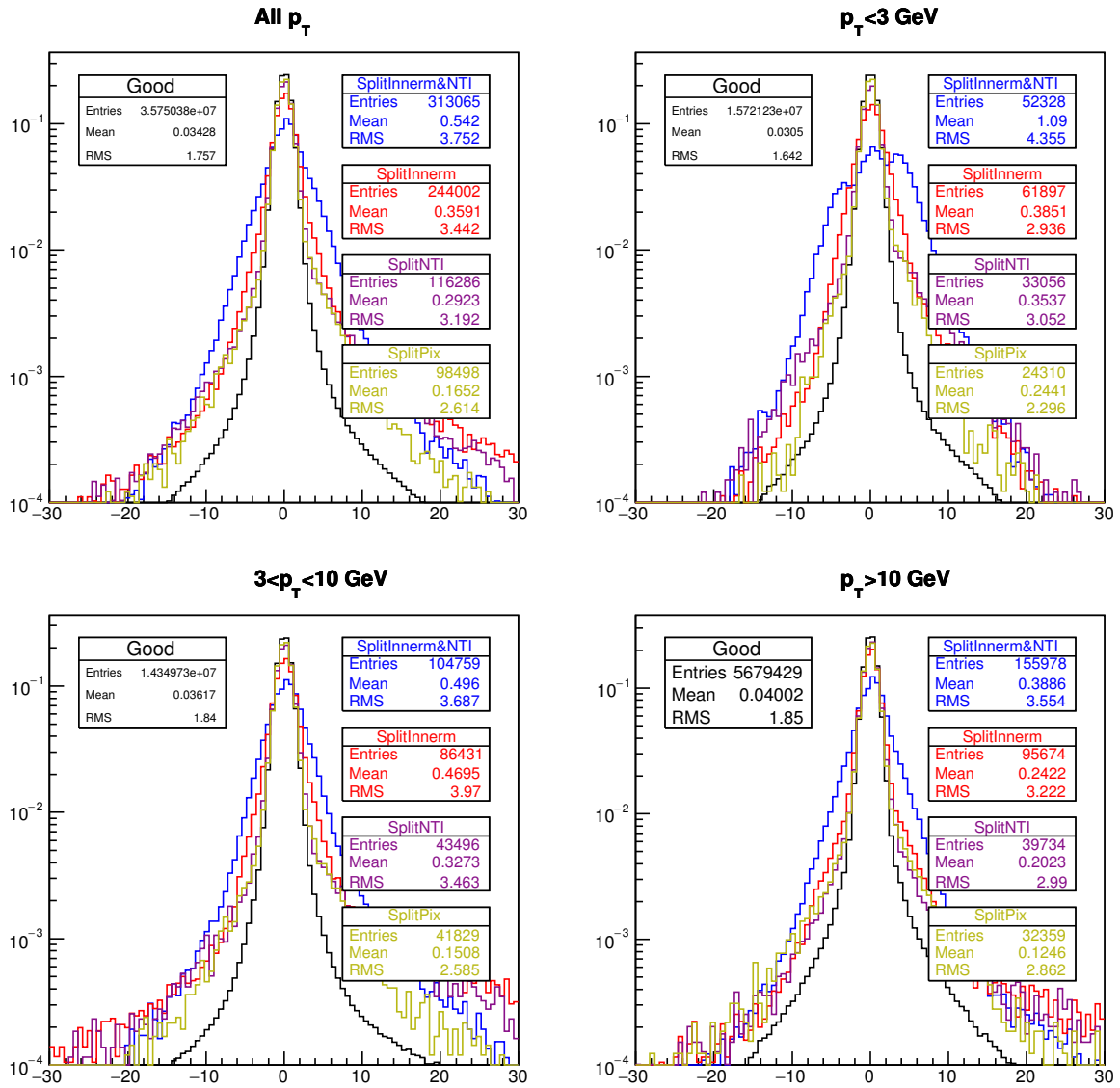
Figure 39: Transverse impact parameter significance of tracks associating with light jets in categories with shared hits

performance of IP3D algorithm. This is shown at figure 40, where light jet rejection vs b-jet efficiency of IP3D algorithm tested and trained with 14 track categories is compared with the one with 18 categories. Rejection is the number of light jets over the number of light jets tagged as b-jets.

The final track categorisation used in Run-2 and percentage of tracks from b-, c- and light flavour jets in each category are presented in the table 7.

The impact on each of the 14 categories was evaluated by retraining the IP3D algorithm with one or another category removed (so if track belong to this category, it is rejected and it's impact is not taken into account in the IP3D calculation). Performance of versions of the IP3D algorithm with each of the 14 categories removed with respect to the default one are presented in figure 41.

Table 7: Run-2 IP2D and IP3D track categories and fraction of tracks from b-, c- and light flavour jets in each category for the $t\bar{t}$ sample [54].

#	Category	light jets	b-jets	c-jets
0	No hits in first two layers; expected hit in both L0 and L1	1.6%	1.5%	1.6%
1	No hits in first two layers; exp. hit in L0 and no exp. hit in L1	0.1%	0.1%	0.1%
2	No hits in first two layers; no exp. hit in L0 and exp. hit in L1	0.03%	0.03%	0.03%
3	No hits in first two layers; no exp. hit in L0 and L1	0.02%	0.03%	0.03%
4	No hit in L0; exp. hit in L0	2.1%	2.4%	2.3%
5	No hit in L0; no exp. hit in L0	0.9%	0.9%	0.9%
6	No hit in L1; exp. hit in L1	0.5%	0.5%	0.5%
7	No hit in L1; no exp. hit in L1	2.3%	2.4%	2.4%
8	Shared hit in both L0 and L1	0.04%	0.01%	0.01%
9	Shared hits in other pixel layers	1.8%	2.1%	1.6%
10	Two or more shared SCT hits	2.2%	2.4%	2.2%
11	Split hits in both L0 and L1	0.8%	1.2%	1.1%
12	Split hits in other pixel layers	1.1%	2.1%	1.6%
13	Good: a track not in any of the above categories	86.6%	84.3%	85.5%

The comparison of the IP3D algorithm performance at Run-2 with new tracking categories with respect to the old one used in Run-1 is presented in Figure 42: light jet rejection vs b-jet efficiency (a) and light jet rejection as a function of jet p_T given a fixed b-jet efficiency of 70% in each bin (b). Light jet rejection is the number of light jets over the number of light jets tagged as b-jets.

The rejection of light jets corresponding to 70% b-jet efficiency is expected to be $\approx 15\%$ higher for the new IP3D configuration with respect to the old one.
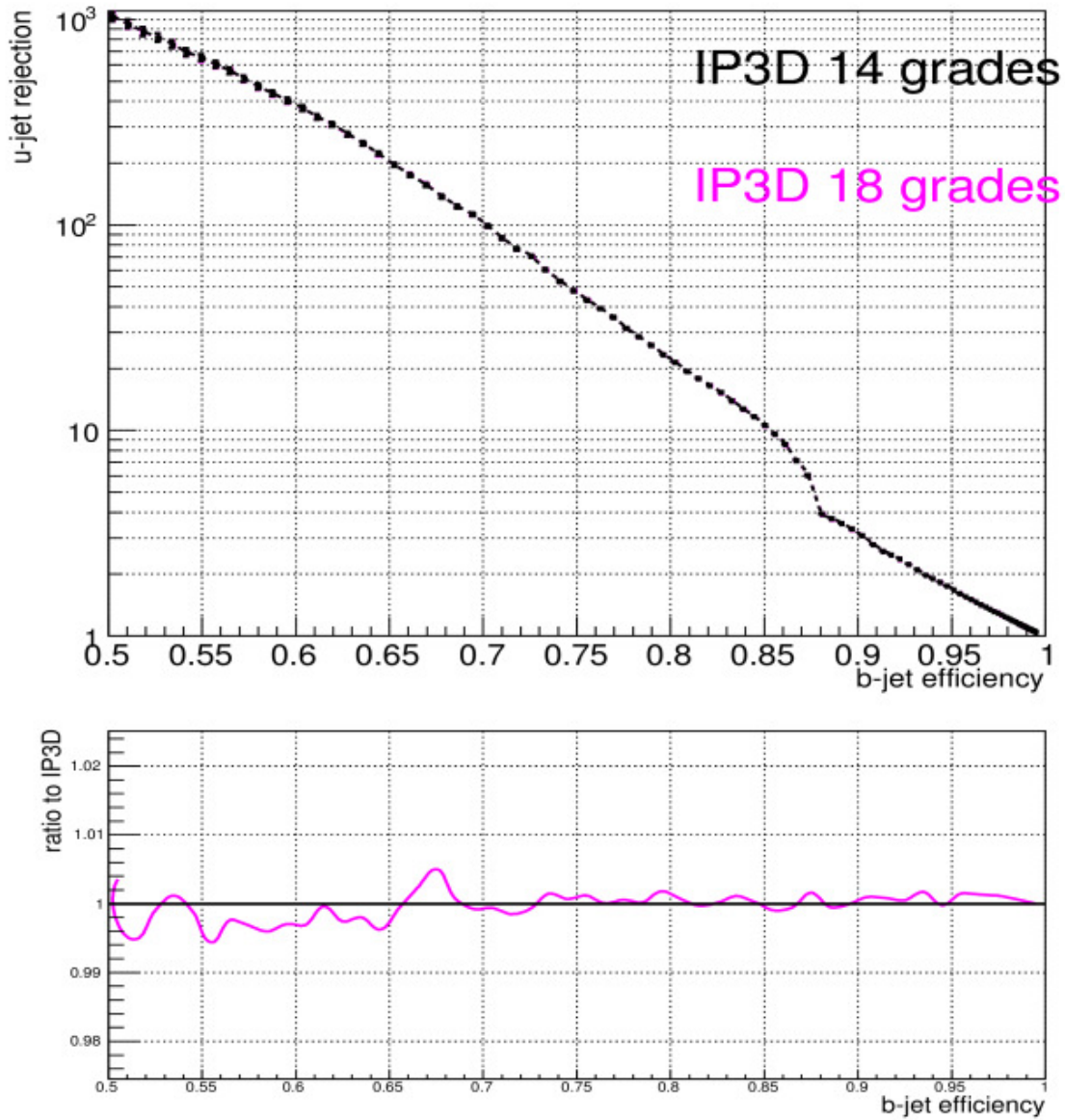
Figure 40: IP3D performance with 14 and 18 categories: light jet rejection vs b-jet efficiency
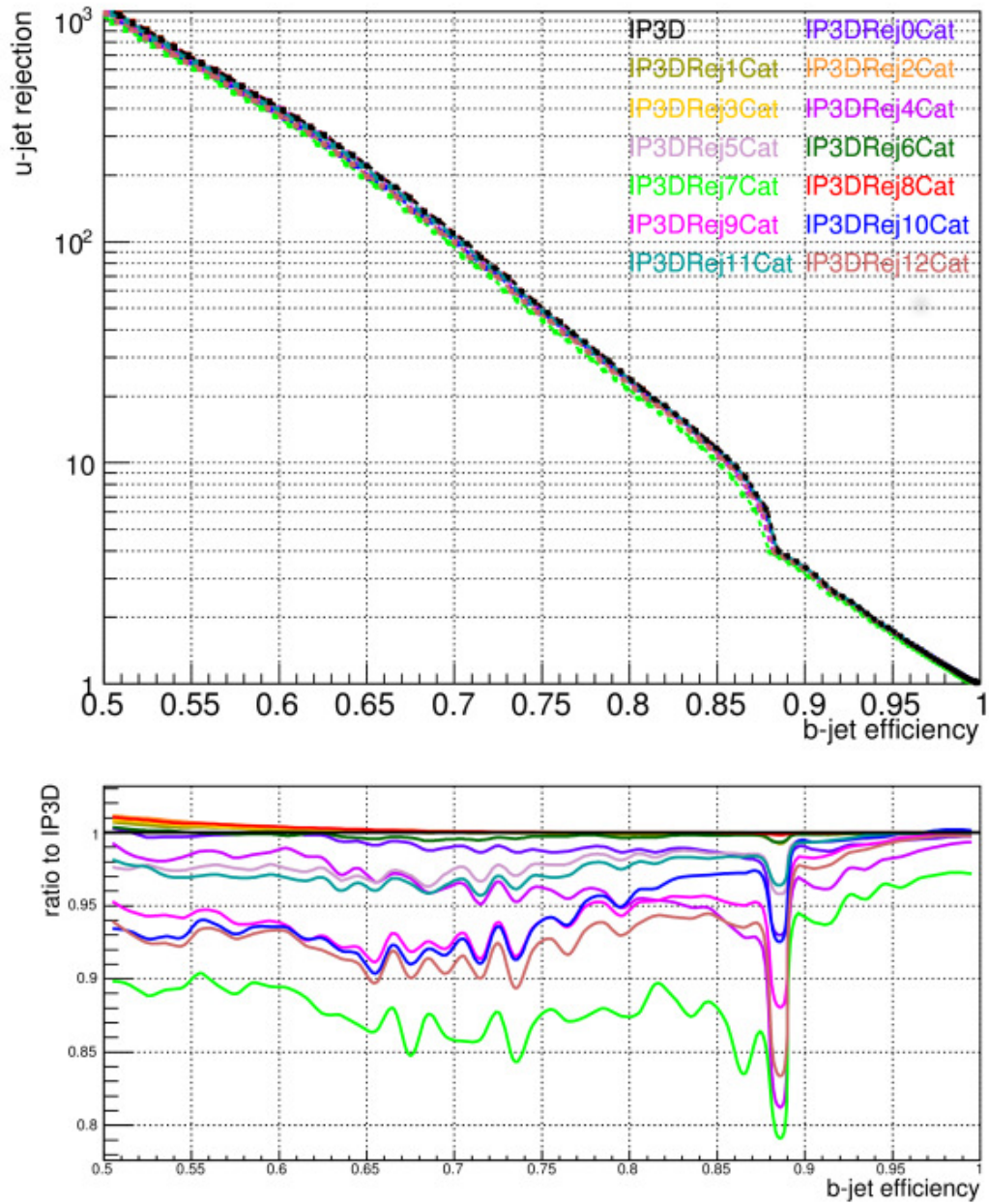
Figure 41: Performance of versions of the IP3D algorithm with each of the 14 categories rejected in the calculation
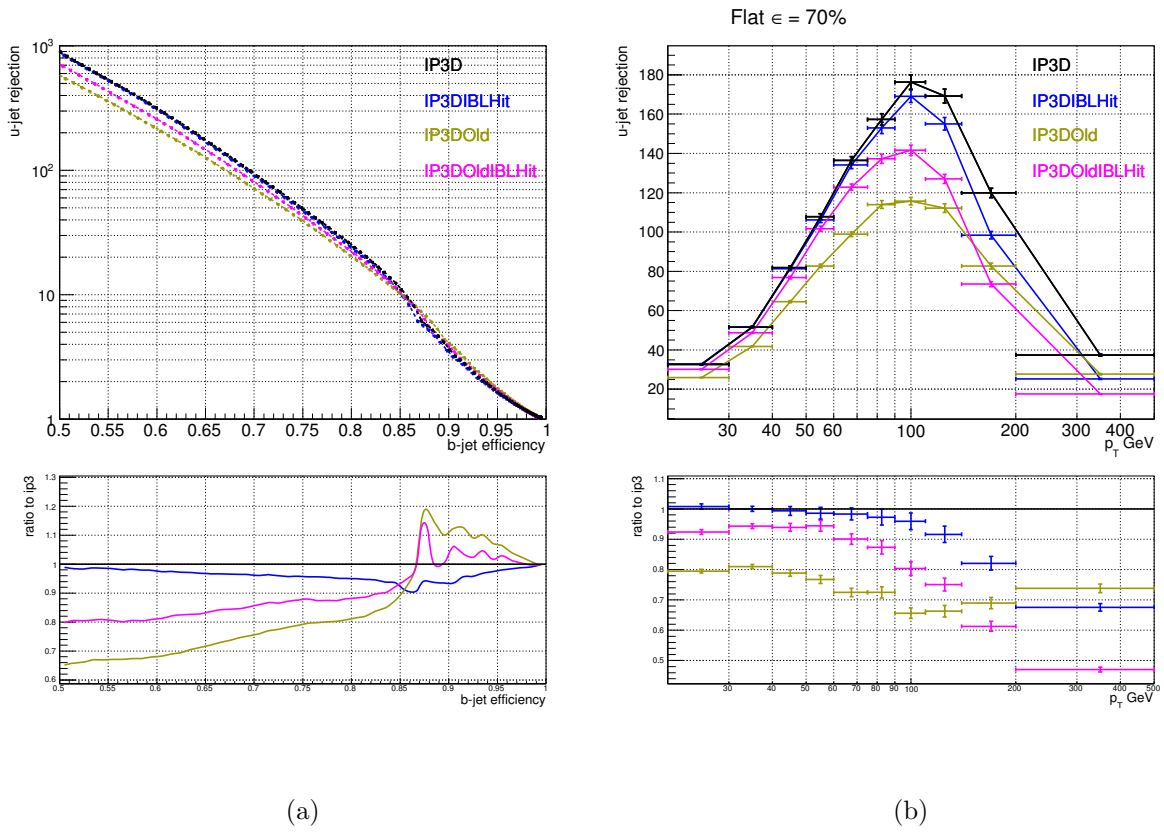
Figure 42: Performance of IP3D algorithm with old grading (Run-1 configuration) and the new Run-2, making use of the IBL and other new tracking variables, in simulated $t\bar{t}$ events: light jet rejection vs b-jet efficiency (a) and light jet rejection as a function of jet p_T for a fixed b-jet efficiency of 70% in each bin (b)

4.2.3   Track selection

Another part of impact parameter based algorithms optimization is considering the selection of tracks that are taking into account in the calculation. Most tracks originating from light jets are having impact parameter close to zero and this allow to distinguish them from tracks from b-jets, which are having larger impact parameter due to long lifetime of B-hadrons. However, in both light and b-jets there is contamination of so called "bad" tracks - tracks coming from long-lived particles, such as K_s, Λ, interaction with the detector material and photon conversions ($\gamma \to e^+e^-$ decay). These tracks are having large impact parameter, so they can be damaging for b-jets identification.

To reduce the negative effect of "bad" tracks one needs to identify and reject them. In order to do that, the SV algorithm identify whether a track is likely to be coming from a long-lived particle.

The Monte Carlo study presented in this section was performed to evaluate the contamination of "bad" tracks of different origin for each of track categories described in 4.2.2, the expected gain for b-tagging of removing the "bad" tracks and the efficiency of SV "bad" track identification algorithm.

In this section the following notation is used:

- "fakes" - tracks with a low matching probability with a MC particle: those with less than 75% of hits coming from the same MC particle (based on Geant4 information).

- "pile-up tracks" - tracks with no association link with a MC particle: most of these tracks are originating from a pile-up interaction rather than hard scattering and information about the MC particles they are coming from is not stored.

- "primary" tracks - tracks originated from particles that are coming directly from primary vertex. Those tracks have impact parameter close to zero.

Table 43 shows composition of tracks from light jets in ttbar events with different origin for all Run-2 b-tagging categories. In the largest "good" tracks category $\approx 97\%$ of the tracks are "primary", while there are categories with significantly different compound. Thus, the categories 0, 1, 2 (with missing hit in both L0 and L1, while at least one of them is expected) and 4 (missing hit in L0 while expected) majority of tracks are coming from photon conversions. In some categories the amount of "fake" tracks is significant: this contamination is especially large for category 8 (shared hits in both L0 and L1) - $\approx 33\%$.

In order to evaluate the expected degrading of b-tagging due to different types of "bad" tracks a Monte Carlo study was performed. The contaminations of "bad" tracks of different types were rejected from the MC samples in training and calculating the IP3D likelihood. Performance of different IP3D algorithm versions (the default one and those with each type of "bad" tracks removed) was compared (Figure 44).

The most damaging for b-tagging are tracks coming from K_s decay: is those particles are removed, the light jet rejection at 70% b-jet efficiency is $\approx 23\%$ higher than for default algorithm (while this difference is even larger for lower efficiencies). The tracks coming from Λ decay are having less significant effect: relative gain of removing those particles

from IP3D calculation is $\approx 5\%$. The tracks coming from photon conversions are not giving any negative impact, but to the contrary, giving a slight improvement ($\approx 1 - 2\%$ at 70% b-jet efficiency region). This can be explained due to the fact that some of the electrons, originating from a photon conversion, are having transverse energy close to the initial photon's, so the transverse impact parameter is not big enough to be harmful for b-tagging. The tracks from the rest of long-lived particles, including interaction with detector material, are also having significant impact: $\approx 23\%$ at 70% b-jet efficiency as well as for lower efficiencies.

		Fakes	Pileup	Ks	Photon Convers.	Λ	Other Secondary	Primary
0	0I&NTIBothExp>0	3.7%	0.4%	3.1%	80.9%	3.3%	3.0%	5.6%
1	0I&NTIExpI>0ExpNTI=0	2.8%	0.6%	2.7%	55.7%	2.8%	3.4%	32.0%
2	0I&NTIExpI=0ExpNTI>0	3.8%	0.4%	2.0%	66.2%	2.7%	2.9%	21.9%
3	0I&NTIExpI=0ExpNTI=0	0.5%	0.4%	0.7%	4.5%	0.4%	0.6%	92.8%
4	0HitsInnExpInn>0	4.4%	0.8%	2.5%	57.3%	2.0%	4.0%	29.0%
5	0HitsInnExpInn=0	0.9%	0.8%	0.5%	2.7%	0.2%	0.5%	94.4%
6	0HitsNTIExpNTI>0	7.6%	1.0%	0.7%	22.4%	0.6%	2.3%	65.4%
7	0HitsNTIExpNTI=0	0.8%	0.9%	0.4%	0.8%	0.2%	0.4%	96.5%
8	SharedInnerm&NTI	32.6%	0.0%	0.3%	8.0%	0.1%	1.3%	57.6%
9	SharedInnerm	3.5%	0.8%	0.7%	10.6%	0.4%	1.7%	82.3%
10	SharedNTI	5.3%	0.8%	0.4%	4.5%	0.2%	1.1%	87.8%
11	SharedPix	8.0%	1.0%	0.4%	2.5%	0.2%	0.6%	87.3%
12	SharedSCT>1	4.9%	0.8%	0.3%	1.1%	0.1%	0.5%	92.4%
13	SplitInnerm&NTI	4.5%	0.1%	0.2%	17.7%	0.2%	1.3%	75.9%
14	SplitInnerm	2.0%	0.4%	1.3%	8.3%	0.6%	2.2%	85.2%
15	SplitNTI	3.3%	0.4%	0.5%	3.0%	0.2%	1.8%	90.8%
16	SplitPix	4.9%	0.4%	0.3%	1.2%	0.1%	0.8%	92.3%
17	Good	1.0%	1.1%	0.4%	0.3%	0.2%	0.3%	96.7%
	All	1.3%	1.1%	0.5%	3.2%	0.3%	0.5%	93.1%

Figure 43: Origin of tracks for different IP3D categories

To reduce the damaging effect of "bad" tracks the information of whether a track was identified by SV as likely coming from a long-lived particle or a material interaction was used. Table 45 presents the efficiency of SV "bad" track identification with different origin for all Run-2 b-tagging categories.

Figure 46 presents the performance of IP3D rejecting the tracks identified by SV tagger as "bad": the expected light jet rejection at 70% b-jet efficiency is $\approx 10\%$ higher than for default algorithm.
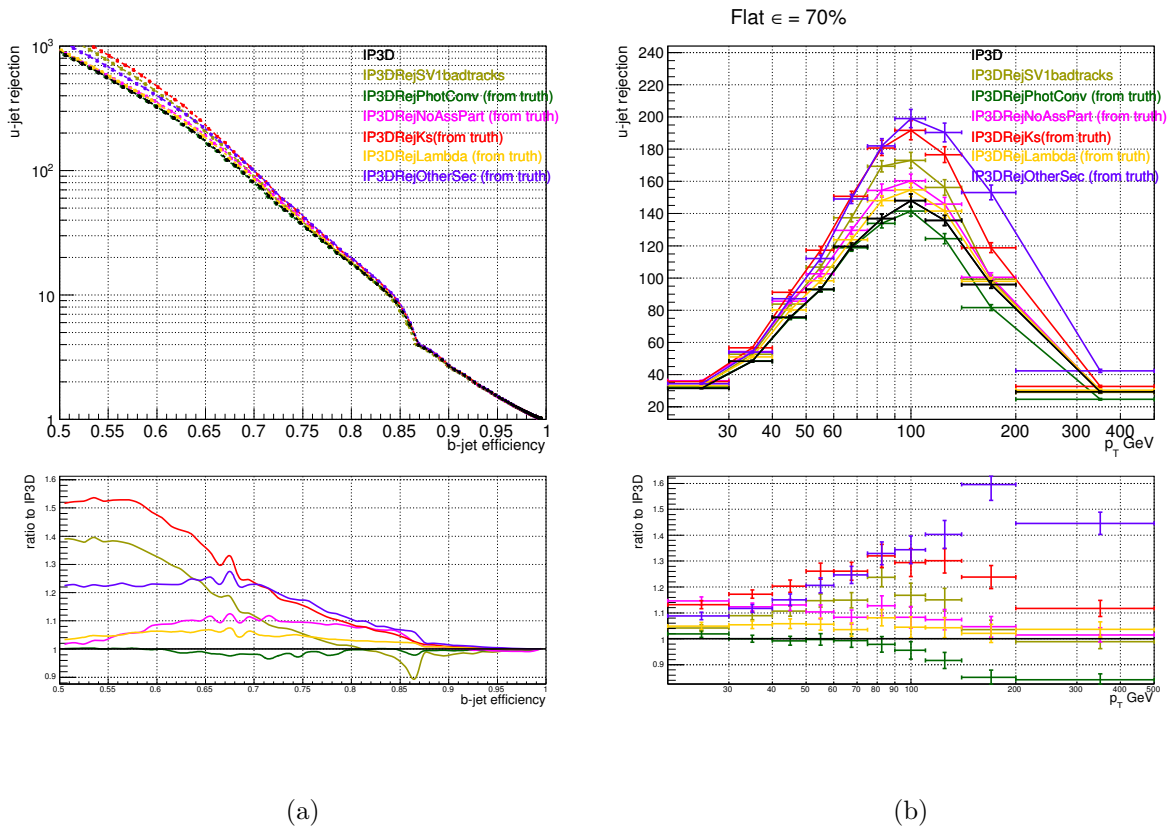
(a)   (b)

Figure 44: Performance of different IP3D algorithm versions (the default one and those with each type of "bad" tracks removed): light jet rejection vs b-jet efficiency (a) and light jet rejection as a function of jet p_T for a fixed b-jet efficiency of 70% in each bin (b) [54].

		Eff Ks	EffPhot Conv	Eff Λ	Eff OtherSec	Eff AllSec	Eff Primary	Purity	% Bad
0	0I&NTIBothExp>0	37.3%	0.7%	31.3%	9.7%	3.3%	0.8%	95.1%	3.1%
1	0I&NTIExpI>0ExpNTI=0	26.5%	0.5%	23.2%	7.1%	3.0%	0.1%	95.6%	2.0%
2	0I&NTIExpI=0ExpNTI>0	38.0%	0.5%	28.0%	13.6%	3.1%	0.3%	94.1%	2.4%
3	0I&NTIExpI=0ExpNTI=0	11.9%	0.2%	17.0%	1.6%	2.7%	0.01%	94.4%	0.2%
4	0HitsInnExpInn>0	41.5%	1.0%	30.6%	12.4%	4.1%	0.4%	90.9%	3.0%
5	0HitsInnExpInn=0	27.7%	0.8%	19.0%	11.4%	6.8%	0.1%	80.3%	0.3%
6	0HitsNTIExpNTI>0	12.5%	5.5%	8.2%	8.0%	6.0%	0.6%	66.2%	2.4%
7	0HitsNTIExpNTI=0	9.4%	0.2%	2.3%	1.4%	2.9%	0.01%	81.4%	0.1%
8	SharedInnerm&NTI	33.3%	8.7%	0%	17.2%	10.6%	5.9%	19.6%	5.9%
9	SharedInnerm	40.6%	11.1%	32.7%	33.2%	16.1%	3.3%	41.6%	5.2%
10	SharedNTI	27.4%	8.5%	18.2%	8.9%	10.0%	1.1%	31.7%	1.8%
11	SharedPix	29.0%	10.3%	14.1%	13.0%	12.7%	0.9%	29.5%	1.5%
12	SharedSCT>1	32.7%	7.1%	17.4%	11.0%	12.9%	0.5%	32.0%	0.8%
13	SplitInnerm&NTI	26.2%	14.7%	26.0%	21.0%	15.4%	11.3%	24.8%	12.1%
14	SplitInnerm	54.0%	12.1%	48.6%	35.5%	22.5%	3.7%	45.3%	6.2%
15	SplitNTI	30.0%	1.4%	11.5%	3.8%	5.3%	0.2%	55.1%	0.5%
16	SplitPix	31.3%	1.4%	12.5%	5.6%	7.3%	0.1%	64.0%	0.3%
17	Good	28.3%	0.3%	15.3%	3.7%	12.9%	0.03%	83.6%	0.2%

Figure 45: Origin of tracks for different IP3D categories

(a)   (b)
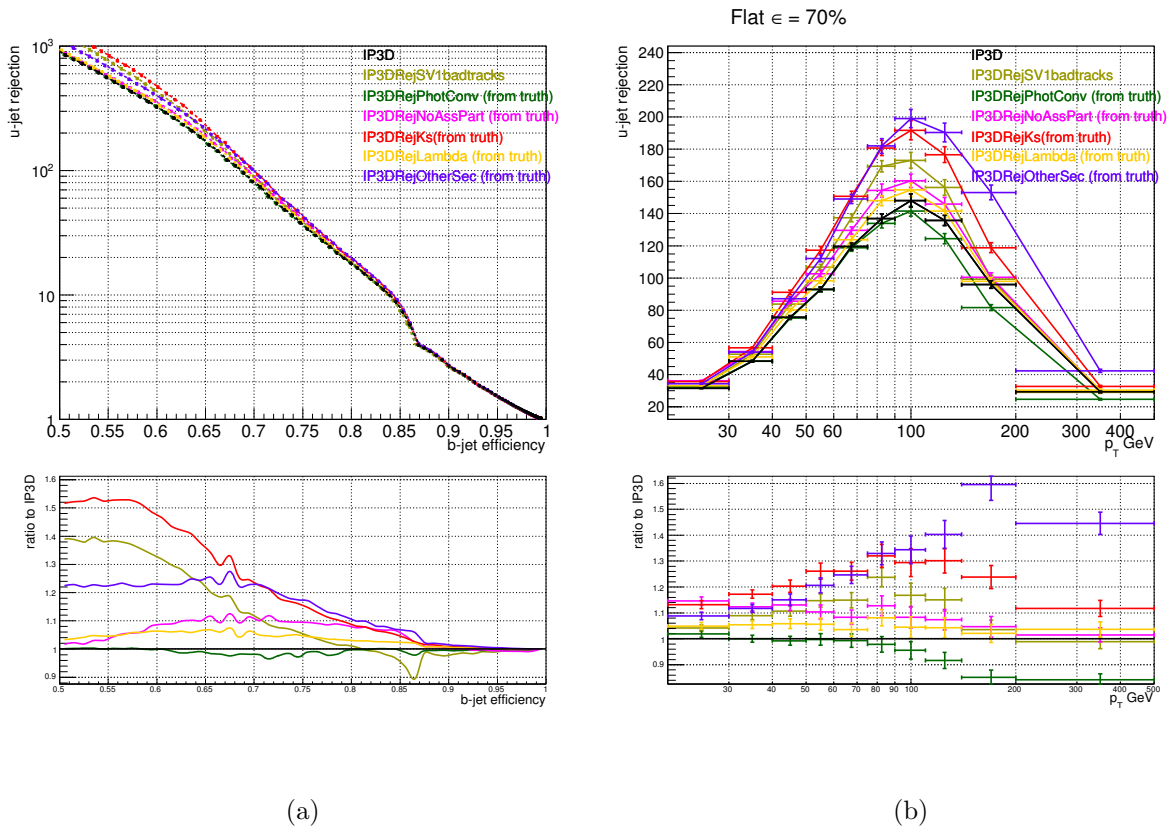
Figure 46: Performance of different IP3D algorithm versions (the default one and those with each type of "bad" tracks removed): light jet rejection vs b-jet efficiency (a) and light jet rejection as a function of jet p_T for a fixed b-jet efficiency of 70% in each bin (b) [54].

4.3 The b-tagging performance for LHC Run-2

The overall b-tagging performance improvement in Run-2 was achieved due to addition of the IBL, updates in track reconstruction [55] and b-tagging algorithms optimisation (the basic taggers and final multivariate algorithm).

Figure 47 show a comparison of the Run-2 b-tagging algorithm MV2c20 and the equivalent Run-1 b-tagging algorithm MV1c: light jet rejection vs b-jet efficiency (a) and light jet rejection as a function of jet p_T given a fixed b-jet efficiency of 70% in each bin (b). Light jet rejection is the number of light jets over the number of light jets tagged as b-jets. Improvement at low and medium p_T is mostly due to the addition of the IBL, while the improvement at high p_T is due to algorithm improvements. At 70% efficiency the light-flavour jet rejection in Run-2 is improved inclusively by a factor of 4 compared to Run-1. This corresponds to a $\approx 10\%$ gain in efficiency at a constant light-jet rejection.
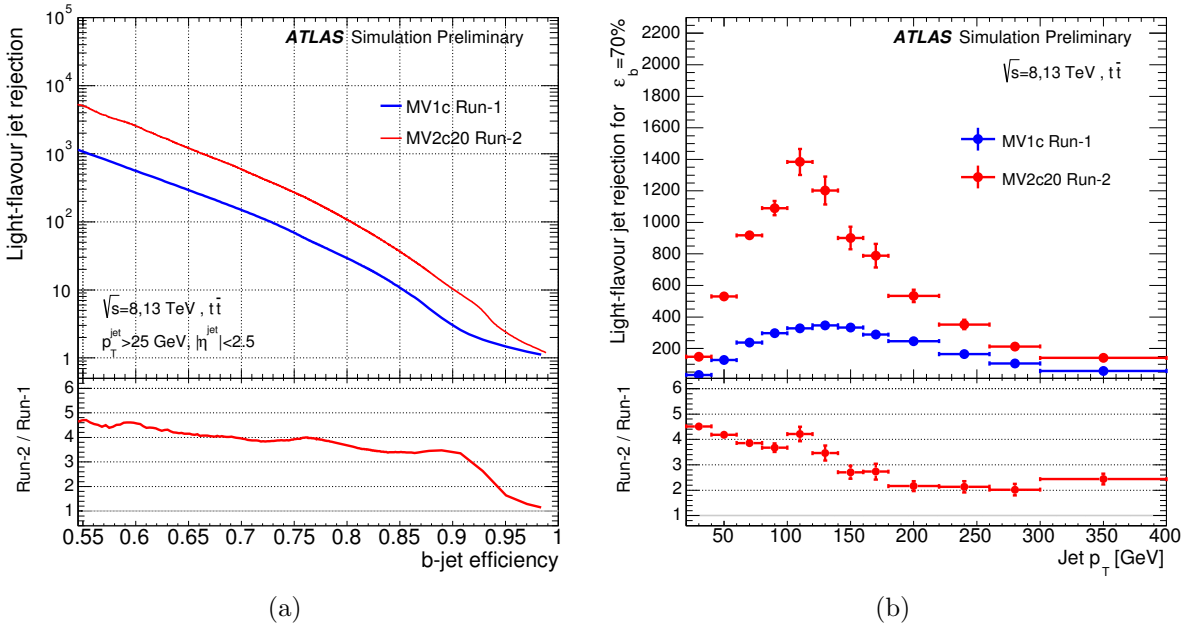

Figure 47: Performance of default Run-2 b-tagging algorithm MV2c20 and the equivalent Run-1 b-tagging algorithm MV1c in simulated $t\bar{t}$ events: light jet rejection vs b-jet efficiency (a) and light jet rejection as a function of jet p_T for a fixed b-jet efficiency of 70% in each bin (b) [54].

To test the Monte Carlo performance, the results in MC simulation were compared to data. The study below was was performed using pp collision data collected by ATLAS at the centre-of-mass energy of 13 TeV with 50 ns bunch-spacing on a high purity b-jet sample of $e + \mu$ di-leptonic $t\bar{t}$ candidate events. Only jets with $p_T > 20$ GeV and $|\eta| < 2.5$ are considered [56].

Input observables from the basic b-tagging algorithms and the output of the multivariate algorithm MV2c20 have been studied. Figure 48 shows the log-likelihood ratio of the IP3D algorithm and the output distribution of the MV2c20 algorithm. On all plots the

data are shown by the points and the simulation by the filled area, divided into b (red), c (light green) and light-flavour (blue) components. The dark green shaded area represents the total systematic and statistical uncertainty on the simulation and the error on the points corresponds to the statistical uncertainty on the data.
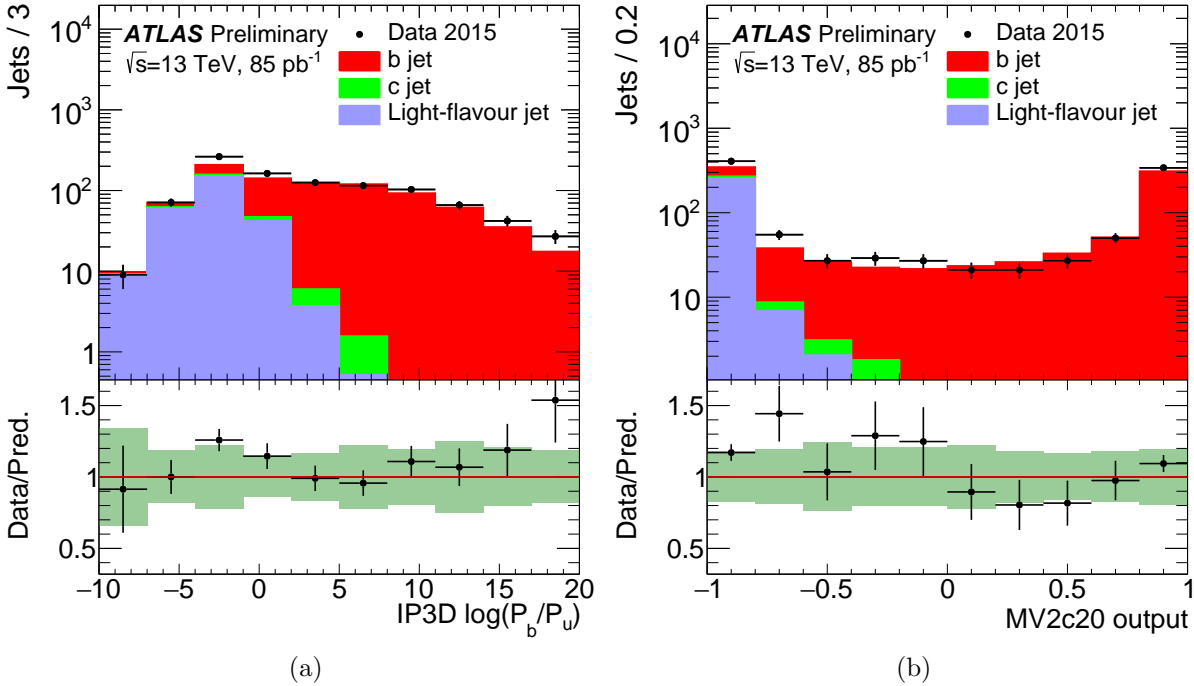

Figure 48: Output distribution of IP3D impact parameter-based algorithm (a) and final MV2c20 algorithm (b) for jets selected from the $t\bar{t}$ dominated $e + \mu$ sample [56].

5 Search for ttH (H → bb)

5.1 Object selection

Physical objects considered in this analysis are electrons, muons, MET and jets. They are reconstructed with ATLAS algorithms as described in section 3.2 and then selected for this analysis channel under additional requirements.

Electrons with "Tight" likelihood identification criteria (see section 3.2.3) are selected, with extra requirements on the transverse momentum $p_T > 27$ GeV, as well as transverse and longitudinal impact parameters:

$$\left| \frac{d_0}{\sigma(d_0)} \right| < 5, \quad |z_0 sin\theta| < 0.5 \text{ mm.} \tag{41}$$

Apart from that, for better discrimination from non-prompts electrons (those coming from conversions and hadron decay), electrons are required to be well isolated from other objects in the event. A tool IsolationSelection was applied to electron candidates. This algorithm using information on the sum of transverse energies of cells around the direction of the candidate in calorimeter as well as scalar sum of the track p_T in a cone of a given size. Then several working points are defined, with respect to efficiency of track isolation. The efficiency numbers were obtained from $Z \to 2l$ MC samples. In this analysis a "Gradient" isolation selection was applied, that corresponds to efficiency $\epsilon = 0.1143 p_T [\text{GeV}] + 92.14\%$ (so the efficiency for a 25 GeV electron is 95%, for a 60 GeV one - 99%).

Muons are reconstructed as explained in (see section 3.2.4) and selected to satisfy Medium quality requirements. Additionally, they are required to have $p_T > 27$ GeV, $\left| \frac{d_0}{\sigma(d_0)} \right| < 3$ and $|z_0 sin\theta| < 0.5$ mm.

As electrons, muons are required to satisfy the "Gradient" isolation selection. Apart from that, muons are demanded to be separated by $\Delta R > 0.4$ from the selected jets. If there is a jet, that doesn't satisfy this criteria, but contains less than three associated tracks, the muon is kept and the jet is removed from the selection.

Jets are reconstructed with the anti-k_T algorithm, presented in section 3.2.5, with a radius parameter $R = 0.4$. They are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. Additional quality criteria are applied to reject jets, coming from non-collision source or detector noise: events containing at least one non-quality jets are removed.

An extra requirement is applied to get rid of jets, coming from secondary pp interactions (pile-up jets), using the Jet Vertex Tagger (JVT) algorithm. The jets with low transverse momentum ($p_T < 60$ GeV) in the central detector region ($|\eta < 2.4|$) are required to have JVT> 0.59.

A procedure known as overlap-removal is applied to avoid double counting jets as electrons. If there is a single jet that contained in a cone $R < 0.2$ around a selected electron, the jet is removed. Then electrons which are lying within $R < 0.4$ of the remaining jets are removed.

Distinguishing b-jets has a key role for this analysis. For this the main b-tagging algorithm MV2c10 (described in details in section 4.1.4) is used.

There are two approaches of using the b-tagging information in a physics analysis.

- cumulative b-tagging - to fix a requirement on the b-tagging algorithm output weight at a certain value of b-tagging efficiency (60%, 70%, 77% and 85% working points (WP) are used in ATLAS). Then the selection of b-jets number is done everywhere in the analysis under this requirement. In this way the Run 1 and previous Run 2 studies were performed.

- pseudo-continuous b-tagging information - using several bins of b-tagging algorithm weight (defined with the same WPs of 60%, 70%, 77% and 85% b-tagging efficiency, and edge points that are interpreted as 0% and 100% efficiencies). For example, if a jet has an MV2c10 weight, that is lying in the bin between 77% and 70% WP, it is considered to be "tagged at 77%". This jet is less likely to be originating from b-quark, than, for instance, the one "tagged at 70%". Thus, jets are divided into 5 grades and this information is used in event categorisation.

E_T^{miss} is reconstructed as described in section 3.2.6 and is used without additional requirements on it.

5.2 Event selection

This search is based on data recorded by the ATLAS detector in pp collisions at 13 TeV in 2015 and 2016. The corresponding integrated luminosity is 36.07 fb^{-1}. Only events for which all detector subsystems were fully operational are considered.

Events are selected with single-electron and single-muon triggers with different p_T thresholds, that are combined in a chain in logical "OR". Those with lower p_T thresholds have additional lepton isolation requirements. For 2015 and 2016 datasets different triggers are used due to the change in data-taking conditions, in particular, the lowest p_T was increased to 26 GeV. All triggers that were used for 2015 and 2016 data in this analysis are listed in table 8.

Events are required to have one lepton and at least five jets, passing the selection described in section 5.1. There are additional requirements based on b-tagging information. For events with 5 jets, at least 3 jets should be b-tagged at 77% WP, for events with ≥ 6 jets two or more jets are required to be b-tagged at 60% WP and at least three at 77% WP.

Events used in other analysis channels search are removed from the selection. Those are, first of all, events with two leptons are removed to avoid overlap with $t\bar{t}H \rightarrow b\bar{b}$ with both W decaying leptonically.

Secondly, events selected for the boosted $t\bar{t}H \rightarrow b\bar{b}$ analysis channel are removed. This search is targeting the final state with the Higgs boson and top quarks produced with a high boost such that their decay products are forming one jet with high ΔR (known as large jet). The event selection for this channel requires at least two large-R jets with some additional criteria.

Finally, events with more than one reconstructed τ-leptons are removed to avoid overlap with $t\bar{t}H$ with 2 τ in the final state.

Type	Name	p_T threshold, [GeV]	Isolation requirement
	2015 data		
	HLT_e24_lhmedium_L1EM20VH	24	yes
	OR		
electron	HLT_e60_lhmedium	60	no
	OR		
	HLT_e120_lhloose	120	no
	HLT_mu20_iloose_L1MU15	20	yes
muon	OR		
	HLT_mu50	50	no
	2016 data		
	HLT_e26_lhtight_nod0_ivarloose	26	yes
	OR		
electron	HLT_e60_lhmedium_nod0	60	no
	OR		
	HLT_e140_lhloose_nod0	140	no
	HLT_mu26_ivarmedium	26	yes
muon	OR		
	HLT_mu50	50	no

Table 8: Single-lepton triggers used for 2015 and 2016 data

74

5.3 Event categorisation

After events are selected, they are split into different categories, called analysis regions, based on jet multiplicity and b-tagging information.

An important characteristic of a region is the expected intensity of signal $t\bar{t}H \to b\bar{b}$ events with respect to the background in the same region. It is usually quantified with two variables: S/B and $S/\sqrt{B}$ ratios, where S - number of expected signal events, B - number of expected background events. Regions that provide high sensitivity to the signal, or signal-enriched regions (SR) are defined as those with $S/B > 1\%$ and $S/\sqrt{B} > 0.3$. The remaining regions are referred to as control regions (CR) - they are used in the fit procedure to constrain systematic uncertainties.

The regions were defined with respect to relative amounts of different background types. It was done taking into account information on jet multiplicity and b-tagging.

Jets were split into 5 grades based on pseudo-continuous b-tagging information as explained in 5.1. Then for both 5 jet and ≥ 6 jet cases, events were split into categories based on grade on four jets with highest b-tagging weight in the events. These fine categories were afterwards merged with certain criteria on relative amount of backgrounds of different types: $t\bar{t}+ \geq 2$ b-jets, $t\bar{t} + 1$ b-jet, $t\bar{t}+ \geq 1$ c-jets and $t\bar{t}+$ light jets. Figure 49 is showing full set of considered b-tagging based categories, with merged regions shown in color. The final classification of signal-enriched and control regions with criteria of the background composition is presented in table 9. Figure 50 is representing sample composition and signal intensity S/B and $S/\sqrt{B}$ for all final analysis regions.

| ≥ 6 jets		5 jets	
Region	Definition	Region	Definition
Signal-enriched regions			
SR1	$> 60\%$ $t\bar{t}+ \geq 2b$	SR1	$> 60\%$ $t\bar{t}+ \geq 2b$
SR2	$> 45\%$ $t\bar{t}+ \geq 2b$	SR2	$> 20\%$ $t\bar{t}+ \geq 2b$
SR3	$> 30\%$ $t\bar{t}+ \geq 2b$		
Control regions			
CR $(t\bar{t} + 1b)$	$> 30\%$ $t\bar{t} + 1b$	CR $(t\bar{t} + 1b)$	$> 20\%$ $t\bar{t} + 1b$
CR $(t\bar{t} + 1c)$	$> 30\%$ $t\bar{t}+ \geq 1c$	CR $(t\bar{t} + 1c)$	$> 20\%$ $t\bar{t}+ \geq 1c$
CR $(t\bar{t}+$ light$)$	Rest	CR $(t\bar{t}+$ light$)$	Rest

Table 9: Analysis regions and their definition with respect to the background composition

5.4 Monte Carlo samples

To estimate signal and background, MC simulation was performed, as described in 3.1. The generators used for in this analysis are summarised in table 10.
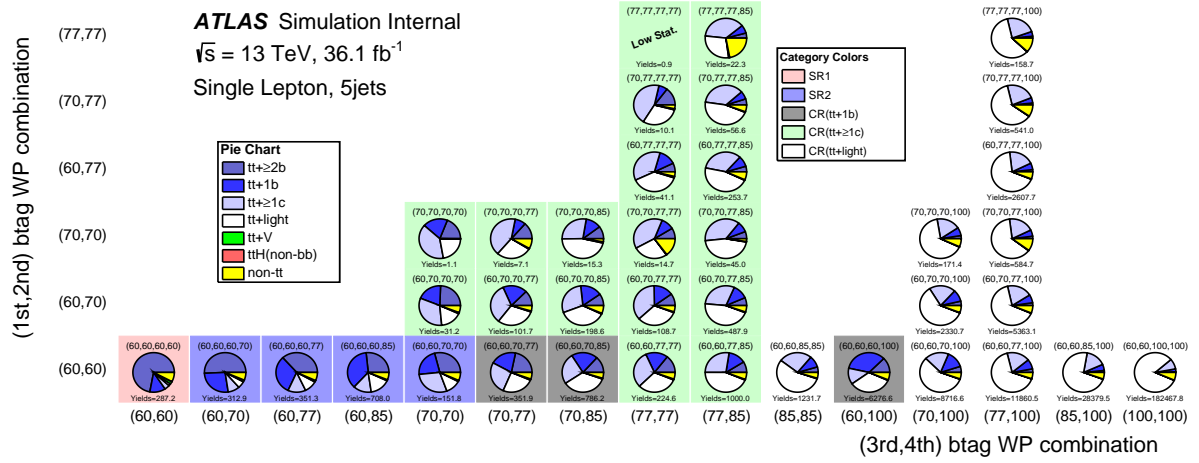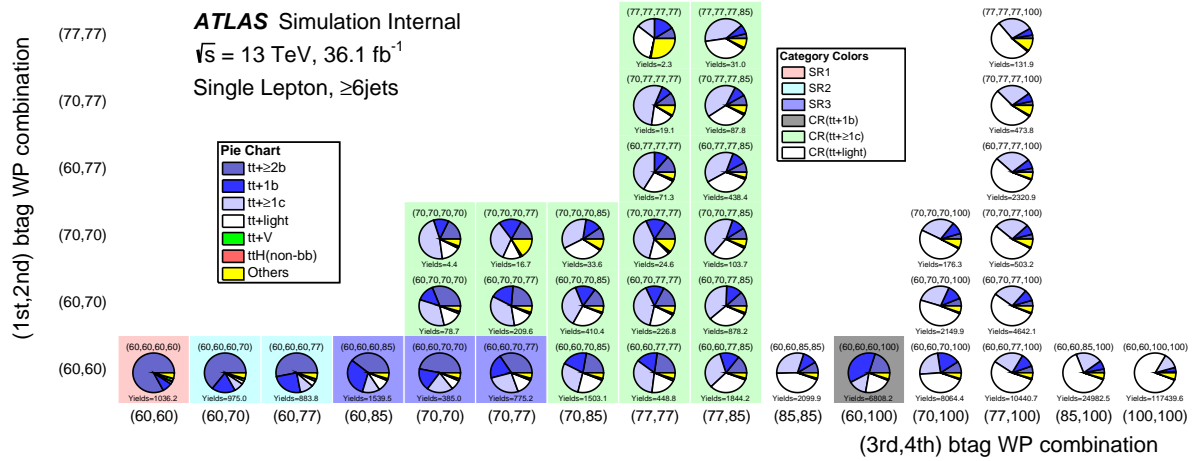
Figure 49: Sample composition for events categories based on b-tagging information in the case of 5 jets (a) and ≥ 6 jets (b). Final categorisation into signal-enriched and control regions is shown in colors.

5.4.1 Signal modelling

The matrix element (ME) calculation for $t\bar{t}H$ modelling was done with MadGraph5_aMC@NLO [57]. The parton distribution function (PDF) set NNPDF3.0NLO [58] is used with factorisation μ_F and renormalisation μ_R scales set to $\mu_F = \mu_R = H_T/2$, where $H_T = \sqrt{p_T^2 + m^2}$ - scalar sum of transverse masses of all final state particles. Various Higgs boson decay modes are produced inclusively. The mass of the Higgs boson is fixed to 125 GeV. The parton shower simulation was performed by Pythia 8.210 [59] using the A14 tune [60] for underlying events modelling. The $t\bar{t}H$ cross section and the Higgs boson decay branching fractions are taken from NLO QCD and NLO QCD + EW theoretical calculations from [17]. For the uncertainty on the signal model a sample interfaced to Herwig++ is used.
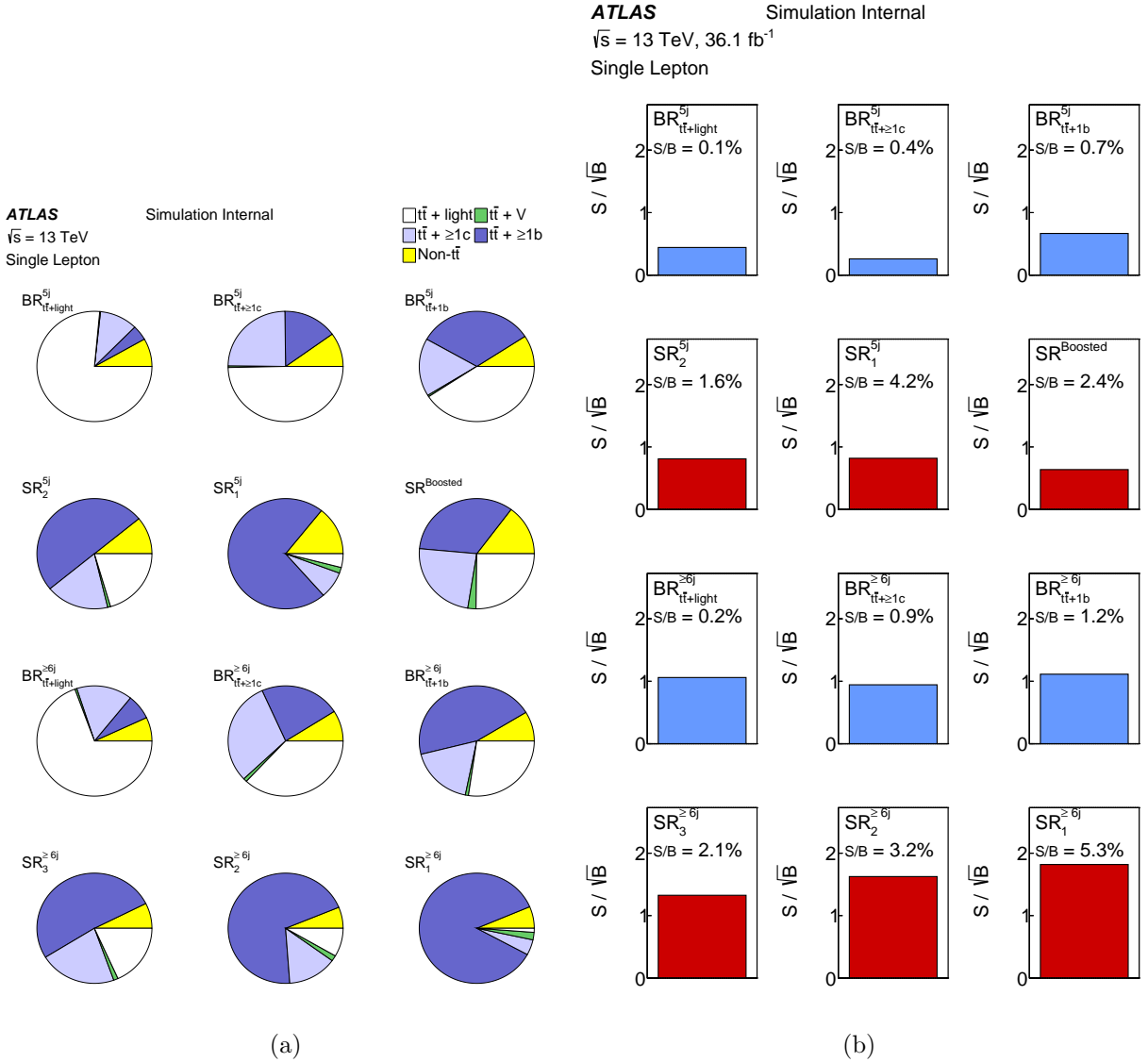
Figure 50: Sample composition (a) and signal intensity S/B and $S/\sqrt{B}$ (b) in different analysis regions.

5.4.2 $t\bar{t}$ + jets background modelling

The $t\bar{t}$ + jets background is dominating in this analysis. It is modelled with Powheg-Box v2 generator [61] with the NNPDF3.0NLO parton distribution function (PDF) set [58]. The hdamp parameter, that regulates the p_T of the first additional emission beyond the Born configuration, is set to 1.5 times the top quark mass $m_t = 172.5$ GeV. Parton shower and hadronisation are modelled by Pythia 8.2 [59] with the appropriate A14 tune [60]. The sample is generated separately for $t\bar{t}$ hadronic and leptonic (including semileptonic and dilepton) W boson decay channels. To reach sufficient statistics in the high b-jet multiplicity regions, that are crucial for this analysis, each of these samples is additionally

Sample	Generator	PDF	Shower	Normalisation
$t\bar{t}H$	MG5_aMC	NNPDF3.0NLO	Pythia 8.2	(N)NLO
$t\bar{t}$ + jets	PowHeg	CTEQ6L1	Pythia 8.2	NNLO+NNLL
W + jets	Sherpa	CT10	Sherpa 2.2.1	NNLO
Z + jets	Sherpa	CT10	Sherpa 2.2.1	NNLO
Single top (s-channel, Wt)	PowHeg	CT10	Pythia 6.428	aNNLO
Single top (t-channel)	PowHeg	CT10f4	Pythia 6.428	aNNLO
$t\bar{t}V$	MG5_aMC	NNPDF3.0NLO	Pythia 8.2	NLO
Diboson	Sherpa	CT10	Sherpa 2.1.1	NLO

Table 10: Generators used for simulating signal and background processes

generated with filters that require additional b-jets (those not originating from top quarks decay).

To simulate bottom and charm hadron decays, the EvtGen v1.2.0 package [62] is used. The sample is normalised to a cross section of 832^{+46}_{-52}, theoretically calculated in NNLO in QCD with top++2.0 [63].

The $t\bar{t}$ + jets events are divided into three categories with respect to the flavour of additional jets: $t\bar{t}+ \geq 1b$, $t\bar{t}+ \geq 1c$ and $t\bar{t}+$ light. To do this classification, matching of bottom and charm hadrons with particle jets is performed. Particle jets are objects created by reconstruction of all stable truth particles, except muons and neutrinos, with the anti-k_T algorithm. Particle jets used in $t\bar{t}$ + jets events classification have radius parameter $\Delta R = 0.4$ and are additionally required to have $p_T > 15$ GeV and $|\eta| < 2.5$. The matching allows to classify events in three categories exclusively.

- First of all, if a particle jet is situated within a cone of $\Delta R < 0.3$ to a B-hadron with $p_T > 5$ GeV, that is not originating from a top quark decay, the event is considered as $t\bar{t}+ \geq 1b$.

- If not, then another matching procedure is performed and if there is a particle jet that is matched to a charm hadron not originating from a W boson decay, the event is defined as $t\bar{t}+ \geq 1c$.

- The events that did not satisfy any of these two matching criteria are labelled as $t\bar{t}+$ light.

A more refined classification can be considered of $t\bar{t}$ + b-jets events, that are of particular importance for this analysis:

- $t\bar{t} + b\bar{b}$ - two particle jets matched to a b-hadron each

- $t\bar{t} + b$ - a single particle jet matched to a single b-hadron

78

- $t\bar{t} + B$ - a single particle jet matched to a b-hadron

- $t\bar{t}+ \geq 3b$ - particle jets are matched to three additional b-hadrons

This approach can be also applied for the $t\bar{t}+$ c-jets event classification in the same way.

It is known that parton shower simulation provided by Sherpa+OpenLoops [64], [65] is modelling the $t\bar{t}+ \geq 1b$ events more precise than Powheg+Pythia8 used in this analysis. In particular, this sample contains tree-level diagrams with up to three additional partons, including b- and c-quarks. To take this into account, so called reweighting procedure is performed. Normalisation of each of the $t\bar{t} + b\bar{b}$, $t\bar{t} + b$, $t\bar{t} + B$ and $t\bar{t}+ \geq 3b$ categories is corrected by applying a weight to match the prediction of Sherpa+OpenLoops.

Alternative MC samples are used to consider systematic uncertainties, connected to the $t\bar{t}$ background. To evaluate the MC generator uncertainty for the hard process a sample produced with Sherpa is used. The showering and hadronization uncertainties are calculated using the Powheg sample interfaced to Herwig 7 [66] is used.

Fractions of various $t\bar{t}$ + jets components (a) and $t\bar{t}+ \geq 1b$ (b) sub-components for different generators are presented in figure 51.

(a) (b)
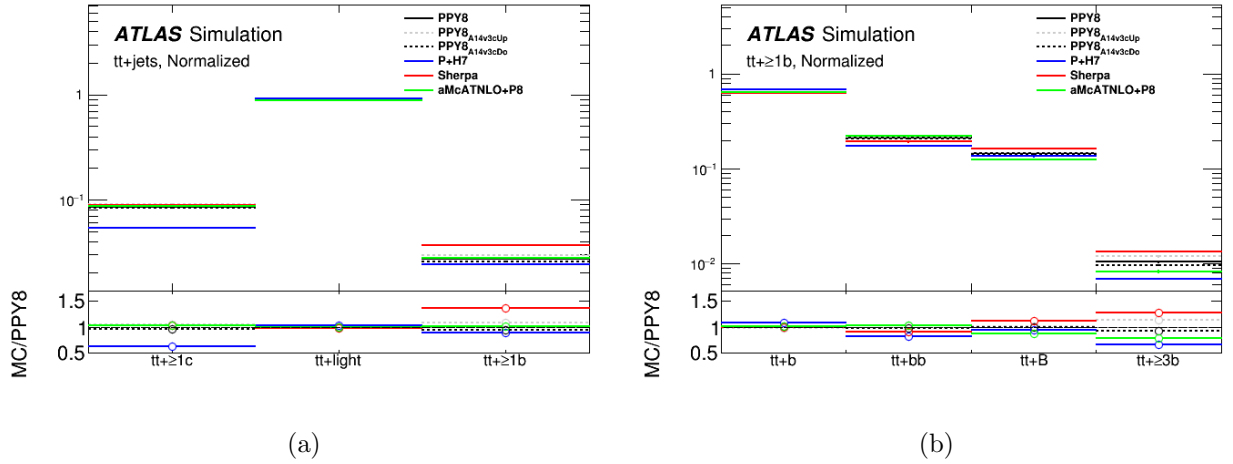
Figure 51: Relative fractions of different $t\bar{t}$ + jets components (a) and $t\bar{t}+ \geq 1b$ (b) sub-components for different generators: nominal Powheg+Pythia8 (black line) and systematic samples: Powheg+Pythia8 sample with more (gray dotted line) and less (black dotted line) radiation, Powheg+Herwig7 (blue), Sherpa 5 FS (red) and aMCAtNLO+Pythia8 (green). The distributions are using particle jets with $p_T > 15$ GeV.

5.4.3 Other background modelling

Other backgrounds considered in this analysis are W/Z and diboson production in association with jets, $t\bar{t}V$ (V - W or Z vector boson), single top quark production (s-channel, t-channel and Wt production).

The samples for W/Z+jets and diboson production in association with jets generated using the Sherpa 2.2.1 generator.

In the case of W/Z+jets samples, matrix element is calculated for up to two partons at NLO and four partons at leading order (LO) using Comix [67] and OpenLoops [65] and merged afterwards with the Sherpa parton shower [68] using the ME+PS@NLO prescription according to [69]. The CT10 PDF set is used with parton shower tuning applied. The resulting W/Z+jets events are then normalised to NNLO cross section prediction.

For diboson+jets a similar approach is used, but in configuration of zero (for WW and WZ) or one (for ZZ) additional partons at NLO and up to three additional partons at LO. The events are normalised afterwards to the NLO cross sections.

The Wt and s-channel single top quark processed are generated with Powheg-Box 2.0 using the CT10 PDF set. To prevent overlap between $t\bar{t}$ and Wt final states, the diagram removal procedure was applied [70]. The t-channel single top-quark samples were produced with the Powheg-Box v1 generator which calculates the NLO matrix element together with CT10f4 PDF set. The parton shower for all single top quark samples are simulated with Pythia 6.428 with the Perugia 2012 underlying-event tune. Bottom and charm decays were modelled with EvtGen v1.2.0. The t- and s-channel samples are normalised to NNLO theoretical cross sections.

In the case of $t\bar{t}V$ samples, the matrix element calculation is performed with Mad-Graph5_aMC@NLO, which is then interfaced to Pythia 8 [59] with the A14 [60] with UE tune [71]. For uncertainties on MC generator for $t\bar{t}V$ alternative samples were used. For therse samples the matrix element was calculated in LO with up to two additional partons using MadGraph5 and interfaced to Pythia 8.

Single top quark produced in association with W boson and Higgs boson (tWH) samples are produced with MadGraph5_aMC@NLO interfaced to Herwig++ with the CTEQ6L1 PDF set. Samples of single top quark produced with Higgs boson and additional jets (tHjb) were generated with Madgraph 5 interfaced to Pythia 8, using the CT10 PDF set. Alternative samples for the case of tHjb are interfaced to Herwig++ with the CTEQ6L1 PDF set.

5.4.4 Misidentified lepton background

Misidentified leptons is a small, but not a negligible background for this analysis. The misidentified electron background has two components: jets or photons misidentified as electrons (known as fake leptons) and non-prompt leptons. For misidentified muons, the source of background is non-prompt muons.

Simulation of these processes is challenging, therefore the misidentified lepton background contribution is estimated from data. A technique known as matrix element is used to obtain misidentified electron yield [72].

Events are categorised into those with loose and tight lepton. The composition of real and fake in two categories is

$$N^{loose/tight} = N^{loose/tight}_{real} + N^{loose/tight}_{fake}. \tag{42}$$

The amount of electrons for the tight category can be rewritten as

$$N^{tight} = \epsilon_{real}N_{real}^{loose} + \epsilon_{fake}N_{fake}^{loose}, \tag{43}$$

where $\epsilon_{real}/\epsilon_{fake}$ - fractions of real/fake leptons for the loose selection that also pass the tight selection.

The fake and real efficiencies depend on properties of the event: kinematics of the lepton as well of number of jets and number of b-jets.

This is taken into account by applying to each event a weight

$$w_i = \frac{\epsilon_f}{\epsilon_r - \epsilon_f}(\epsilon_r - \delta_i), \tag{44}$$

where $\delta_i = 1$ if the loose event i passes the tight event selection and $\delta_i = 0$ otherwise. The background estimate is then is given by the sum of w_i over all events [73].

The real efficiencies ϵ_r are obtained from the $Z \to ee$ and $Z \to \mu\mu$ events. The fake efficiencies ϵ_f are obtained from data samples dominated by non-prompt and fake leptons.

5.5 Analysis strategy

Small signal yields with respect to large amount of irreducible background makes this analysis challenging, therefore effective discriminating techniques are crucial.

To provide better discrimination of signal from background, several methods were developed. Information they provide is combined into the final discriminating variable.

The reconstruction boosted decision trees (BDT) method attempts to reconstruct to $t\bar{t}H$ system by finding the best match between the observed jets and the final-state partons, testing events with a multivariate approach.

The matrix element method is calculating likelihood probabilities for a full parton-level phase space configuration under signal $t\bar{t}H$ and main background $t\bar{t}$ hypotheses. This method is applied only in one signal-enriched region with the highest signal-to-background ratio (6 jets, 4 are b-tagged at 60% WP).

The likelihood discriminant method is exploiting kinematic information of all observed final state objects and testing them under both signal $t\bar{t}H$ and main background $t\bar{t}$ hypotheses, considering all possible matches between the reconstructed jets and the final-state partons. To test these kinematic properties, probability density functions (pdfs) derived from signal and background MC events. Development and optimisation of this method, is the main contribution of this thesis. The detailed description of the technique is presented in section 5.6.

Finally, the classification BDT taking as input information provided by three listed above methods and combines it together with other kinematic variables using a multivariate approach. The output of the classification BDT is the final discriminating variable, used for the fit procedure in all signal-enriched regions.

5.5.1 Reconstruction BDT

The reconstruction method is based on boosted decision trees technique of the TMVA package [74]. A detailed description of the method can be found in [75].

The first step is the training process of the BDT, that is performed using the signal MC samples. The correct jet assignment to the partons they originate from is defined with so called jet-matching procedure. A jet is matched to a parton if it is situated within a cone of $\Delta R < 0.3$, where $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$ - spatial distance. The combination of jets is considered to be correct if all six partons (b-quarks from top quark decay, b-quarks from Higgs boson decay and light and charm quarks from hadronic W decay) are matched with jets, or if all except one quark from W decay are matched. This combination is treated as signal. Other "wrong" jet assignment combinations are considered to be background.

After that the objects such as Higgs boson, top quarks and W boson are reconstructed using kinematic properties of given combination. Distributions of invariant masses as well as spatial angles ΔR between this objects are obtained for both signal (correct jet assignment combinations for $t\bar{t}H$ system) and background (wrong combinations).

After training is done, the BDT is evaluated for each jet assignment combination for $t\bar{t}H$ and $t\bar{t}$ + jets MC samples. The combination with the highest BDT output is considered as the final state reconstruction.

The most powerful discriminating variables between $t\bar{t}H$ and $t\bar{t}$ + jets are properties of the Higgs boson, so a BDT configuration making use of these properties provide high efficiency. But when attempting to reconstruct $t\bar{t}$ + jets events, this configuration biases the mass of Higgs candidate towards $t\bar{t}H$. Therefore two versions of the reconstruction BDT are used. First one is using in training the reconstructed jets that correspond to the six quarks of the $t\bar{t}H$ system. It considers Higgs boson related variables. Second is taking into account only variables corresponding to top quarks and W boson properties. In this case only jets matched to the four quarks from the $t\bar{t}$ are considered in the training.

5.5.2 Matrix element method

The principle of the method is to evaluate a likelihood of an event to be originating from either signal ($t\bar{t}H$) or background ($t\bar{t}$ + b-jets), based on matrix element calculated for Feynman diagrams of these two processes for the parton-level phase space of this event.

For each event two likelihoods are calculated under signal $t\bar{t}H$ and background $t\bar{t}$ + b-jets hypothesis:

$$L_{S/B} = \Sigma \int \frac{f_1(x_1, Q^2) f_2(x_2, Q^2)}{|\vec{q_1}||\vec{q_2}|} \left| M_{S/B}(\mathbf{Y}) \right|^2 T(\mathbf{X}, \mathbf{Y}) \mathrm{d}\mathbf{\Phi}_n(\mathbf{Y}), \tag{45}$$

where sum is over different possible initial states of the system (modes of $t\bar{t}H$ production) and over all final states, considering all possible jet-parton assignments, f_1 and f_2 are parton distribution functions (pdfs) - probability density functions for two initial state partons with momenta q_1 and q_2 to carry energy fraction x_1 and x_2 of the proton in a collision at energy Q. The matrix element $M_{S/B}$ is calculated for a phase space configuration $\mathbf{Y}$ at parton level for either signal $t\bar{t}H$ or background $t\bar{t}$ + b-jets LO Feynman diagrams. The connection between parton-level phase space ($\mathbf{Y}$) and the reconstructed in the detector objects ($\mathbf{X}$) is provided by transfer functions $T(\mathbf{X}, \mathbf{Y})$, that are describing probabilities of reconstructed objects to be originating from this parton phase space configuration. The phase space factor $\mathrm{d}\mathbf{\Phi}_n(\mathbf{Y})$ allows to take into account unknown neutrino parameters.

The final discriminating variable is then given by

$$MEM_{D_1} = \log_{10} L_S - \log_{10} L_B. \tag{46}$$

5.5.3 Classification BDT

The outputs of the reconstruction BDT, likelihood discriminant method and matrix element method are combined together with general kinematic variables and pseudo-continuous b-tagging information into the classification BDT. The full list of variables used as input for the classification BDT is presented in table 11. As reconstruction BDT, it is based on the TMVA package [74].

5.6 Likelihood discriminant

5.6.1 Introduction

The method for discriminating signal from background presented in this chapter is based on the combinatorial likelihood approach. Probabilities of a given event to be signal $P^{sig}(\mathbf{x})$ and background $P^{bkg}(\mathbf{x})$ are computed making use of MC-based probability density functions (pdfs) for signal ($t\bar{t}H$, $H \to b\bar{b}$) and background ($t\bar{t} + jets$) hypotheses (Figure 54). The pdfs are functions of four-momentum vectors $\mathbf{x}$ of reconstructed objects in this event: jets, lepton and neutrino.

Various invariant mass resonances provide important information to separate signal from background. Those are the Higgs mass for the signal hypothesis, masses of the leptonic top, the hadronic top and the hadronic W for both signal and background hypotheses. The pdfs of these invariant masses are the most significant ones used in this method, more details them are presented in section 5.6.2, 5.6.2 and 5.6.2. Other pdfs exploited are additional mass variables, described in 5.6.4, and various angular variables, presented in 5.6.5.

The probability to be signal P^{sig} can be defined as product of the probabilities of the invariant masses in this event (see figure 54 (a)): leptonic top mass $M_{t_l}(l, \nu, b_l)$, hadronic top mass $M_{t_h}(q_1, q_2, b_h)$, hadronic W mass $M_{W_h}(q_1, q_2)$ and Higgs mass $M_H(b_1, b_2)$. Probability to be background P^{bkg} is defined in a very similar way as probability to be signal P^{sig}, but exploiting instead of Higgs mass the pdf of additional jets b_1 and b_2.

The final discriminating variable is defined as

$$D = \frac{P^{sig}}{P^{sig} + P^{bkg}}. \tag{47}$$

The distributions of invariant masses were obtained from simulated signal events and background using the reconstructed lepton and MET four-momentum vectors and jets, which origin is identified by applying so called truth-matching procedure. Jet is defined to be matched to a quark if this quark is within a cone $\Delta R < 0.3$ to the jet. The histograms filled with these mass distributions are normalised to unit area and used as references for the probabilities calculations. A smoothing procedure was applied in the high mass range of the pdfs, where the MC statistics is limited.
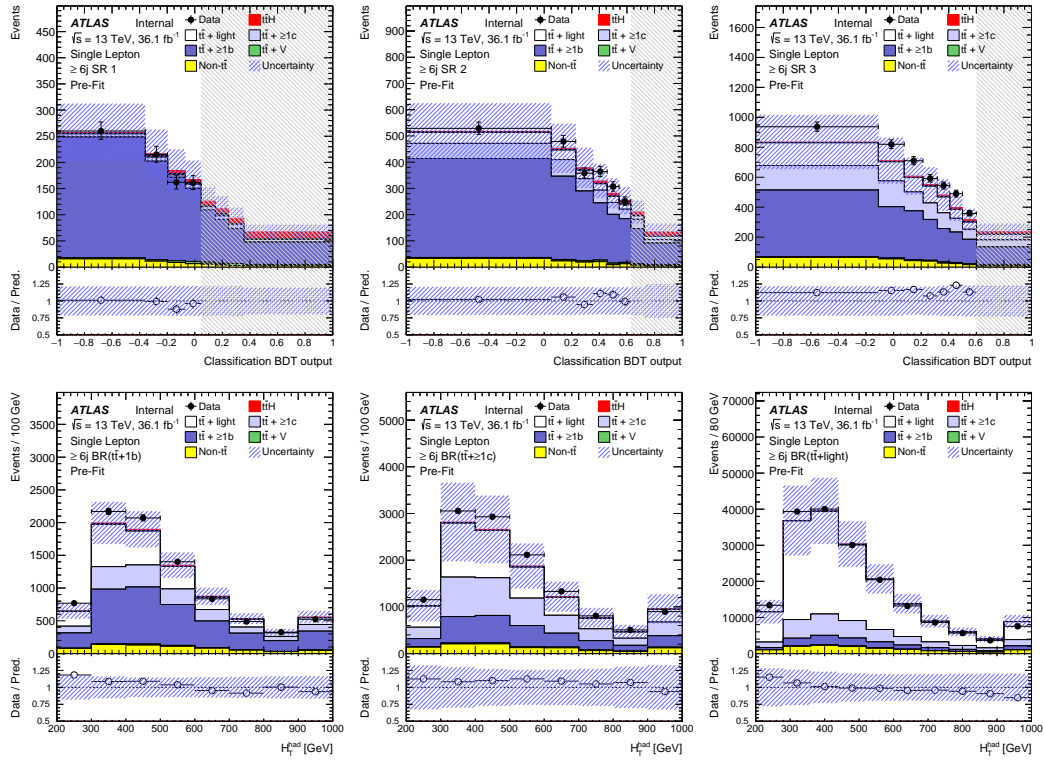
Figure 52: Prefit plots of the discriminant in the signal-enriched (top row) and signal depleted (bottom row) analysis regions with 6 jets.
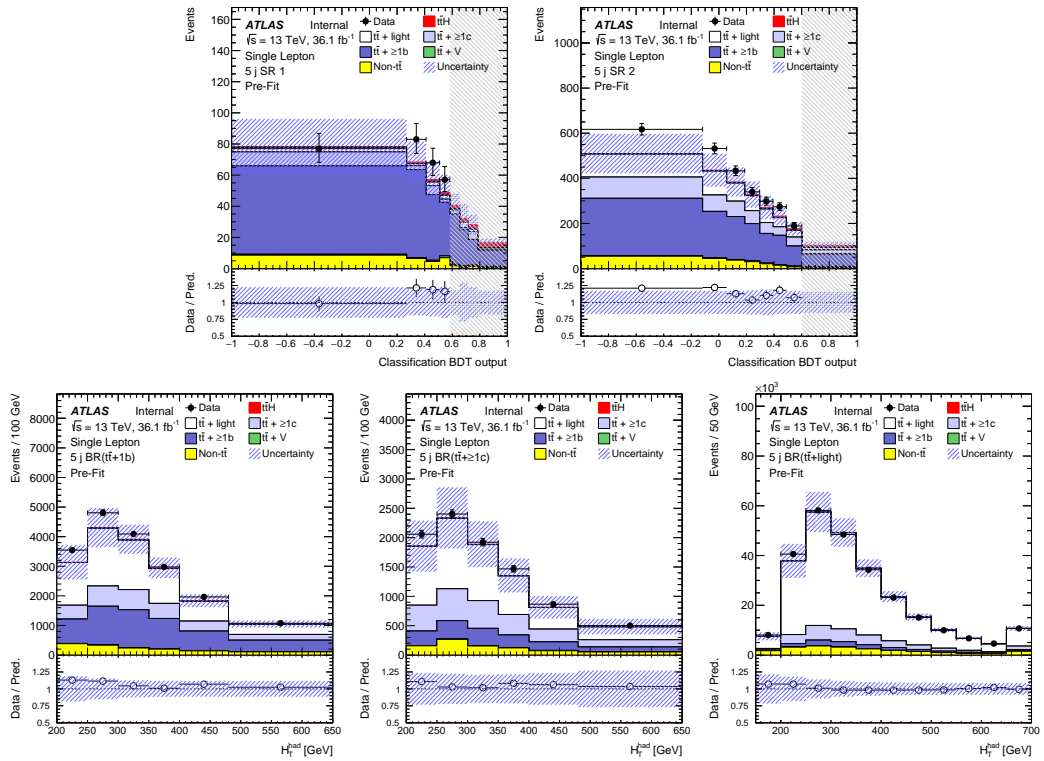

Figure 53: Prefit plots of the discriminant in the signal-enriched (top row) and signal depleted (bottom row) analysis regions with 5 jets.

Variable	Definition	Region	
		$\geq$ 6j	5j
General kinematic variables			
$\Delta R_{bb}^{\mathrm{avg}}$	Average ΔR for all b-tagged jet pairs	✓	✓
$\Delta R_{bb}^{\max\ p_T}$	ΔR between the two b-tagged jets with the largest vector sum p_T	✓	–
$\Delta\eta_{jj}^{\max\ \Delta\eta}$	Maximum $\Delta\eta$ between any two jets	✓	✓
$m_{bb}^{\min\ \Delta R}$	Mass of the combination of the two b-tagged jets with the smallest ΔR	✓	–
$m_{jj}^{\min\ \Delta R}$	Mass of the combination of any two jets with the smallest ΔR	–	✓
$\mathrm{N}_{30}^{\mathrm{Higgs}}$	Number of b-jet pairs with invariant mass within 30 GeV of the Higgs boson mass	✓	✓
H_T^{had}	Scalar sum of jet p_T	–	✓
$\Delta R_{\mathrm{lep}-bb}^{\min\ \Delta R}$	ΔR between the lepton and the combination of the two b-tagged jets with the smallest ΔR	–	✓
Aplanarity	$1.5\lambda_2$, where λ_2 is the second eigenvalue of the momentum tensor built with all jets	✓	✓
$H1$	Second Fox–Wolfram moment computed using all jets and the lepton	✓	✓
Variables from reconstruction BDT output			
BDT	BDT output	✓*	✓*
m_H	Higgs boson mass	✓	✓
$m_{H,b_{\mathrm{lep\ top}}}$	Mass of Higgs boson and b-jet from leptonic top	✓	–
$\Delta R_{\mathrm{Higgs\ bb}}$	ΔR between b-jets from the Higgs boson	✓	✓
$\Delta R_{H,t\bar{t}}$	ΔR between Higgs boson and $t\bar{t}$ system	✓*	✓*
$\Delta R_{H,\mathrm{lep\ top}}$	ΔR between Higgs boson and leptonic top	✓	–
$\Delta R_{H,b_{\mathrm{had\ top}}}$	ΔR between Higgs boson and b-jet from hadronic top	–	✓*
Variable from Likelihood calculation			
D	Likelihood discriminant	✓	✓
Variable from Matrix Method calculation			
MEM_{D1}	Matrix Method	✓	–
Variables from b-tagging			
w_b^H	Sum of binned b-tagging weights of jets from best Higgs candidate	✓	✓
B_{j^3}	3rd jet binned b-tagging weight (sorted by weight)	✓	✓
B_{j^4}	4th jet binned b-tagging weight (sorted by weight)	✓	✓
B_{j^5}	5th jet binned b-tagging weight (sorted by weight)	✓	✓

Table 11: Classification BDT input variables in 6 jets and 5 jets signal-enriched regions. Variables from the reconstruction BDT labeled with * are from the BDT using Higgs boson information, others are from the reconstruction BDT without Higgs boson information. The MEM_{D1} variable is only used in the signal-enriched region with the highest signal-to-background ratio (6 jets, 4 are b-tagged at 60% WP), while b-tagging weights are not used in this region.
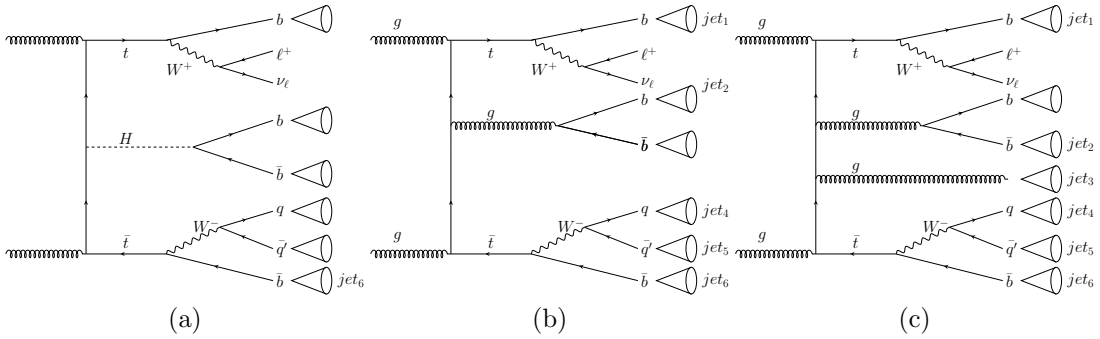
Figure 54: Signal $t\bar{t}H$ (a) and background $t\bar{t} + jets$ hypotheses (b) and (c)

5.6.2 Signal probability

Higgs invariant mass

A very important variable that can be exploited for discriminating signal from background is the Higgs boson mass resonance. The Higgs invariant mass $M_H(b_1, b_2)$ pdf is built using information on the jets truth-matched with two b-quarks from Higgs b_1 and b_2 in the signal MC events, it is shown in figure 55.
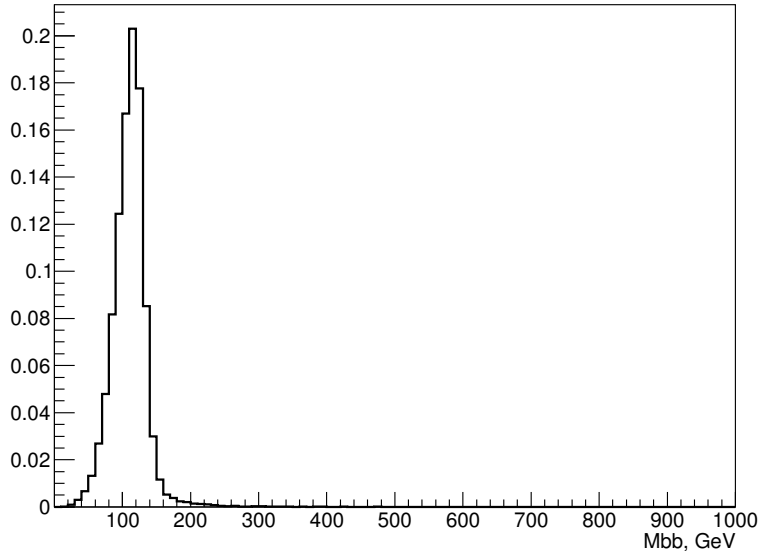

Figure 55: Pdf of Higgs invariant mass in ttH MC events

Leptonic top invariant mass

The invariant mass of leptonic top $M_{t_l}(l, \nu, b_l)$ is reconstructed using the jet matched to b_l and lepton four-momentum vectors and missing transverse energy E_T^{miss}.

86

Transverse energy of neutrino can be measured due to the fact that initial transverse energy of two hard scattering partons is zero. Thus, the imbalanced transverse energy of all reconstructed objects in final state (missing transverse energy, MET, or E_T^{miss}) is considered as transverse energy of neutrino. The longitudinal component of neutrino energy is not known, but can be calculated, using constraint of the empirical value of W-boson mass $M_W = 80.4$ GeV, which provides a quadratic equation with one unknown - p_{z_ν}:

$$M_W^2 = (p_l + p_\nu)^2. \tag{48}$$

In the case of discriminant of this quadratic equation $\Delta > 0$ there are two solutions:

$$p_{z_\nu}^{\pm} = \frac{p_{z_l}\beta \pm \sqrt{\Delta}}{2(E_l^2 - p_{z_l}^2)}, \tag{49}$$

where

$$\beta = M_W^2 - M_l^2 + 2p_{x_l}p_{x_\nu} + 2p_{y_l}p_{y_\nu}, \tag{50}$$

$$\Delta = E_l^2(\beta^2 + (2p_{z_l}p_{T_\nu})^2 - (2E_l p_{T_\nu})^2) \tag{51}$$

In the case of two neutrino solutions they are ordered with respect to the $|p_{z_\nu}|$, so that $|p_{z_\nu 1}| < |p_{z_\nu 2}|$. It was figured out that in $\sim 65\%$ of signal events $p_{z_\nu 1}$ is closer to the truth neutrino distribution than $p_{z_\nu 2}$. Two leptonic top invariant mass pdfs are built and the $P^{sig}(M_{t_l})$ is then constructed using both of them with different weights: 0.65 for $p_{z_\nu 1}$ and 0.35 for $p_{z_\nu 2}$.

Due to the error on MET measurement the quadratic equation 48 does not have a real solution in $\sim 35\%$ signal events. In this case solution p_{z_ν} is approximated: the missing energy is varied until the discriminant of quadratic equation $\Delta = 0$ and one neutrino solution p_{z_ν} is obtained.

Pdfs of leptonic top mass are build separately for two neutrino solutions in the case of real solutions and for approximation in the case of no real neutrino solution. Those three distributions are presented in Figure 56.
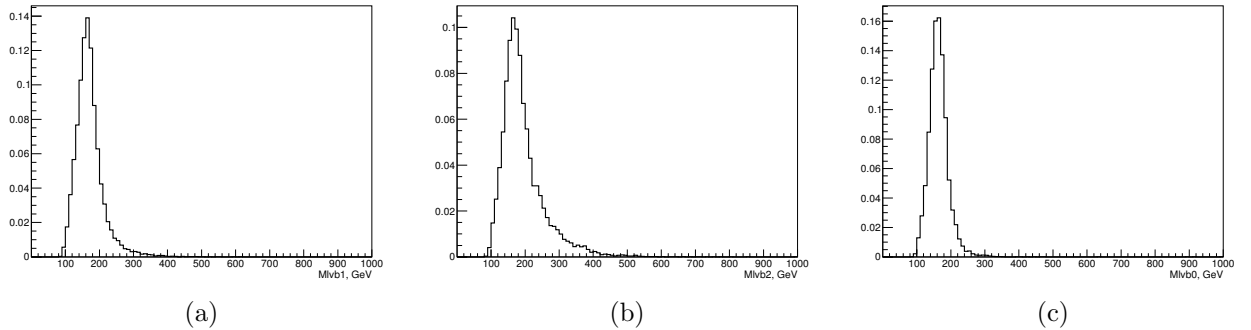
Figure 56: Pdfs of leptonic top mass M_{t_l} in the case of two neutrino solutions: solution 1 (a) and solution 2 (b), and no real solution case (c)

87

Hadronic W and hadronic top invariant masses

Hadronic W $M_{W_h}(q_1, q_2)$ and hadronic top $M_{t_h}(q_1, q_2, b_h)$ invariant masses pdfs are built in a similar way based on information on the jet truth-matching. However, these two masses are correlated, therefore in the final probability calculation instead of the hadronic top mass M_{t_h} the difference between hadronic top and hadronic W masses $M_{t_h} - M_{W_h}$ is used. These three distributions are presented at figure 57. Figure 58 is showing that difference between hadronic top and hadronic W masses $M_{t_h} - M_{W_h}$ is less correlated with hadronic W invariant mass M_{W_h} than hadronic top M_{t_h} mass.

(a) (b) (c)
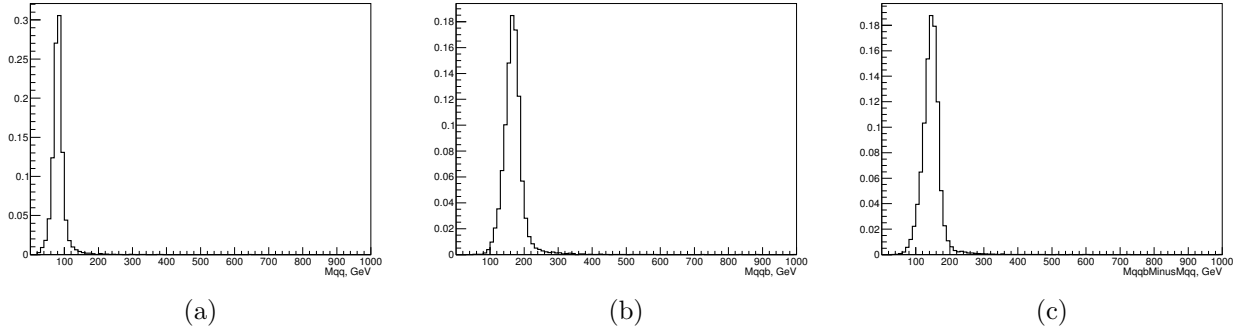
Figure 57: Pdfs of hadronic W invariant mass M_{W_h}(a), hadronic top invariant mass M_{t_h} (b) and their difference $M_{t_h} - M_{W_h}$ (c)

(a) (b)
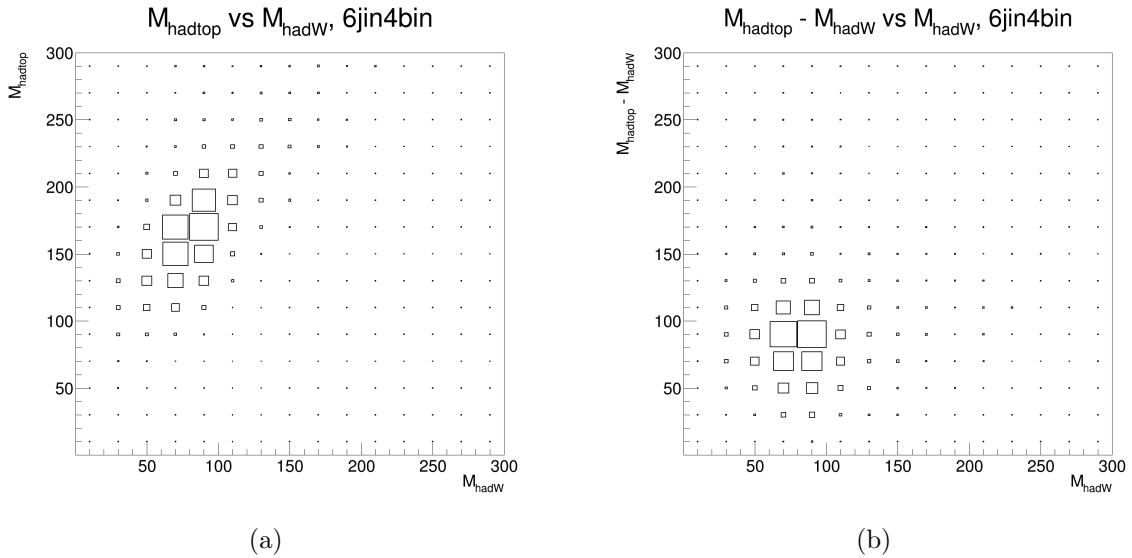
Figure 58: Mass of hadronic top mass vs hadronic W mass (a) and difference between hadronic top and W masses vs W mass (b) in ttH MC events at $>= 6$ jets, $>= 4$ b-tagged jets region

Sum over all jet permutations and b-tagging weights

Assuming that signal probability can be defined as product of invariant masses of the Higgs, tops and W and after replacing hadronic top invariant mass M_{t_h} by $M_{t_h} - M_{W_h}$ the expression for signal probability is becoming

$$P_{kin}^{sig} = P^{sig}(M_H)P^{sig}(M_{t_l})P^{sig}(M_{t_h} - M_{W_h})P^{sig}(M_{W_h}). \tag{52}$$

However, as partonic origin of jets is not known, the signal probability can be calculated summing over all possible jet permutations N_p in the event. The b-tagging information is then used to give different weights to permutations. The expression for signal probability is becoming

$$P^{sig} = \frac{\sum_{k=1}^{N_p} P_{kin}^{sig} P_{btag}^{sig}}{\sum_{k=1}^{N_p} P_{btag}^{sig}}, \tag{53}$$

where P_{kin}^{sig} as shown in equation 52 and the b-tagging term P_{btag}^{sig} is defined as

$$P_{btag}^{sig} = P_b(jet_1)P_b(jet_2)P_b(jet_3)P_l(jet_4)(f_l P_l(jet_5) + f_c P_c(jet_5))P_b(jet_6). \tag{54}$$

In this expression $jet_i(j = 1, ..6)$ is considered reconstructed jet, $P_f(jet_i)$ represent the probability that jet_i is originating from parton of flavour f. These probabilities are computed, using the jet MV2c10 tagger $w(jet_j)$ weight. There are 6 b-tagging WPs used for evaluation: 100%, 85%, 77%, 70%, 60% and 0%. First method of calculation the probabilities is the following: if jet_i has weight $w(jet_i)$ between threshold values for two operating points WP_1 and WP_2: $w_{WP_1} < w(jet_i) < w_{WP_2}$, then $P_f(jet_i) = \epsilon_f^{WP_2} - \epsilon_f^{WP_2}$.

5.6.3 Background probability

The dominating background in the main signal region ($>= 6$ jets, $>= 4$ b-jets) is $t\bar{t}+ >= 1$ additional b-jets. Therefore, two background hypotheses are considered:

A $t\bar{t}+ >= 2$ additional b-jets, which happens in $\sim 80\%$ events (figure 54 (b))

B $t\bar{t} + 1$ additional b-jet, which happens in the $\sim 20\%$ events (figure 54 (c))

The probability of an event to be background P_{bkg} is calculated in similar way to probability to be signal P_{sig}, exploiting the invariant masses of resonances: leptonic top, hadronic top, hadronic W. In order to keep the P_{bkg} in the same dimensionality as P_{sig}, invariant mass of two additional jets $M_{b_1 b_2}$ pdf is used in the same way as pdf of Higgs invariant mass for the signal.

For hypothesis A the pdf is constructed using invariant masses of two leading in p_T additional b-jets (figure 59 (blue)), for hypothesis B the unique b-jet and the leading in p_T additional not b-tagged jet (figure 59 (green)). Two pdfs are then used in the P_{bkg} calculation with different weights, corresponding to fractions of A and B cases in the

background events: for the region ($>= 6$ jets, $>= 4$ b-jets) the weights are $f_A = 0.8$ and $f_B = 0.2$.
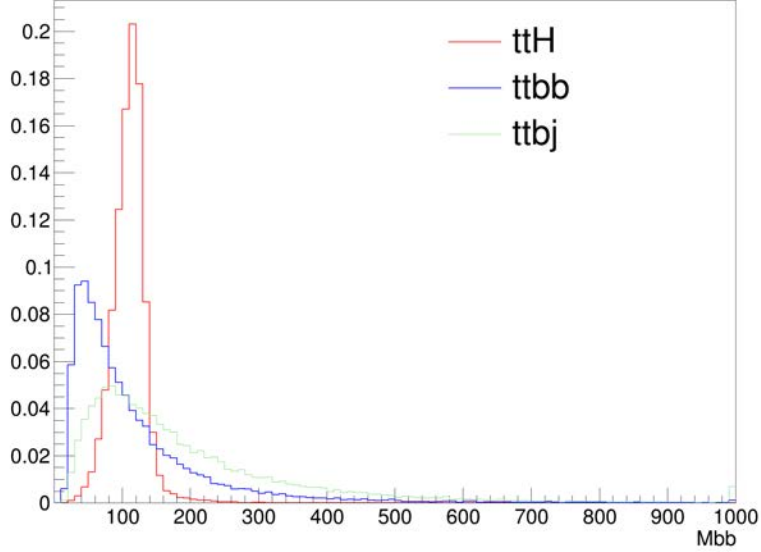
Figure 59: Pdf of Higgs invariant mass (red) in ttH MC events and two extra jets invariant mass in $t\bar{t}+ >= 2$ additional b-jets (blue) and $t\bar{t} + 1$ additional b-jet (green) cases

The background probability is defined in similar way as signal probability, but with additional sum over two hypotheses (A, B):

$$
P^{bkg} = \frac{\sum\limits_{k=1}^{N_p} \sum\limits_{j=A,B} f_k P_{kin}^{bkg,j} P_{btag}^{bkg,j}}{f_k \sum\limits_{k=1}^{N_p} \sum\limits_{j=A,B} P_{btag}^{bkg,j}},
\tag{55}
$$

where kinematic terms for two hypotheses

$$
P_{kin}^{bkg,A} = P^{bkg}(M_{b_1 b_2}) P^{bkg}(M_{t_l}) P^{bkg}(M_{t_h} - M_{W_h}) P^{bkg}(M_{W_h}),
\tag{56}
$$

$$
P_{kin}^{bkg,B} = P^{bkg}(M_{b,j}) P^{bkg}(M_{t_l}) P^{bkg}(M_{t_h} - M_{W_h}) P^{bkg}(M_{W_h}),
\tag{57}
$$

the b-tagging terms P_{btag}^{bkg} are defined as

$$
P_{btag}^{bkg,A} = P_b(jet_1) P_b(jet_2) P_b(jet_3) P_l(jet_4)(f_l P_l(jet_5) + f_c P_c(jet_5)) P_b(jet_6),
\tag{58}
$$

$$
P_{btag}^{bkg,B} = P_b(jet_1) P_b(jet_2)(f_l P_l(jet_3) + f_c P_c(jet_3)) P_l(jet_4)(f_l P_l(jet_5) + f_c P_c(jet_5)) P_b(jet_6),
\tag{59}
$$

probabilities $P_f(jet_i)$ are computed as shown in 5.6.2.

90

5.6.4 Additional invariant mass variables

The kinematic probabilities terms P_{kin}^{sig} and P_{kin}^{bkg} as defined in equations 52, 56, 57 are exploiting invariant masses of different resonance peaks in the event. However, there are additional invariant masses terms that can be used to improve separation between the signal and the background: invariant mass of the $t\bar{t}$ system $M_{t\bar{t}}$ and invariant mass of $t\bar{t} + b\bar{b}$ system $M_{t_h t_l b_1 b_2}$.

These two invariant masses are depending on the neutrino solution p_{z_ν} in the same way as leptonic top invariant mass (see 5.6.2). Thus, the pdfs were derived separately for two solutions in real neutrino solution case and approximate solution for the case of no real neutrino solution. All pdf distributions shown in this paragraph were produced in the "solution 1" case.

The invariant mass of $t\bar{t}$ system $M_{t\bar{t}}$ has correlations with invariant masses of tops M_{t_l} and M_{t_h}. Therefore mass difference $M_{t\bar{t}} - M_{t_l} - M_{t_h}$ is used instead. The two variable distributions $M_{t\bar{t}}$ and $M_{t\bar{t}} - M_{t_l} - M_{t_h}$ are presented in figure 60, their correlations are illustrated on figure 61, it is demonstrating that $M_{t\bar{t}} - M_{t_l} - M_{t_h}$ has significantly less correlations with both mass variables M_{t_l} and M_{t_h} than $M_{t\bar{t}}$. For these variables pdfs are derived in the same way in signal and background events.
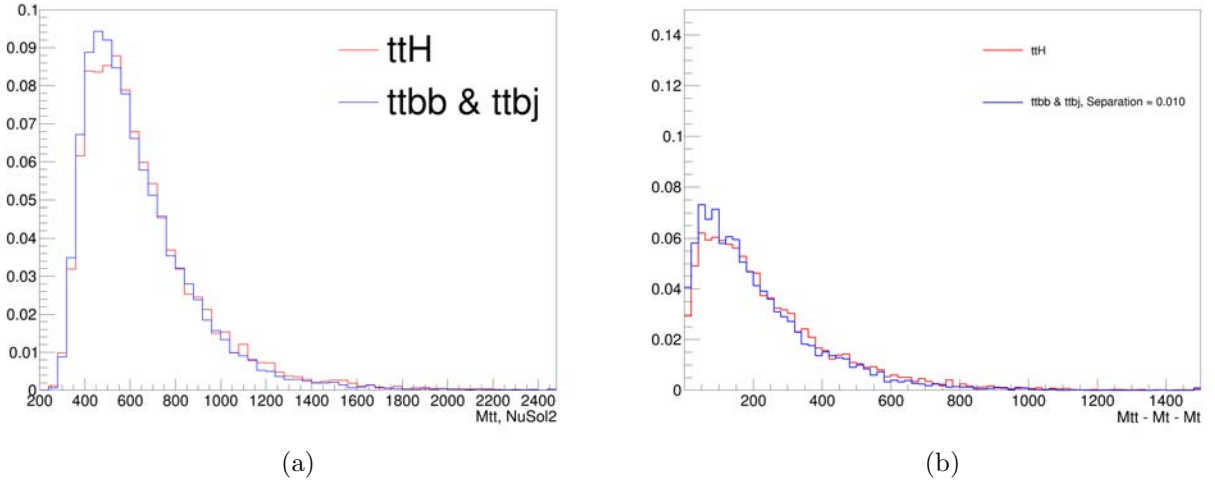
Figure 60: Pdfs of $M_{t_h t_l}$ and $M_{t_h t_l} - M_{t_l} - M_{t_h}$ in ttH MC events at $>= 6$ jets, $>= 4$ b-tagged jets region

The invariant mass of $t\bar{t}+b\bar{b}$ system $M_{t_h t_l b_1 b_2}$ is built differently in signal and background events. In the case of signal it is invariant mass of jets matched to tops and two b-jets b_1 and b_2 matched to Higgs. In the case of background, two pdfs are derived: using two extra jets b_1 and b_2 invariant mass in $t\bar{t}+ >= 2$ extra b-jets case, and unique extra b-jet b and the leading in p_T extra not b-tagged jet j in $t\bar{t} + 1$ extra b-jet case. As all three sets of pdfs (signal hypothesis and two background hypothesis) are built in the same way, for convenience notation $M_{t_h t_l b_1 b_2}$ will refer to all of them.

The correlations of invariant mass of $t\bar{t} + b\bar{b}$ system $M_{t_h t_l b_1 b_2}$ with $M_{t_l t_h}$ and $M_{b_1 b_2}$ are
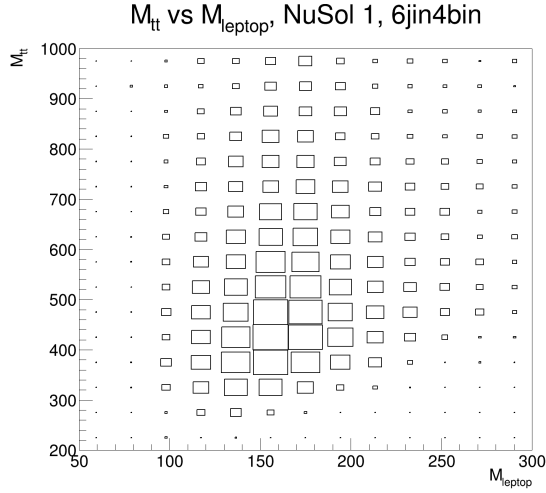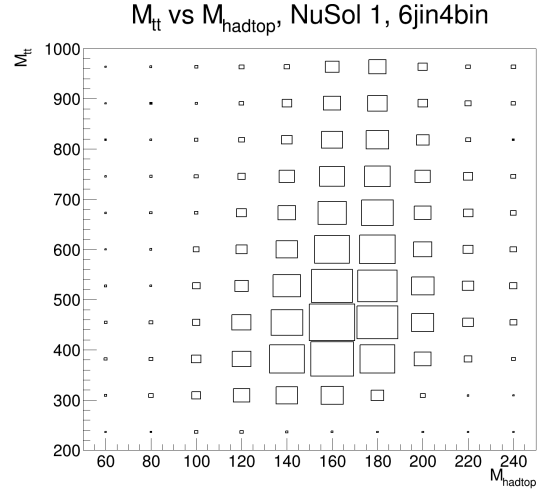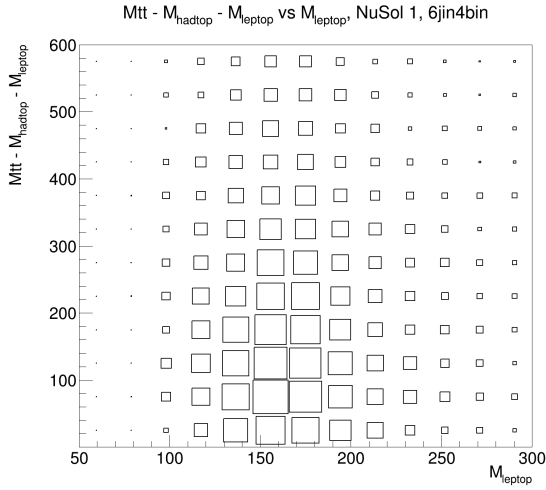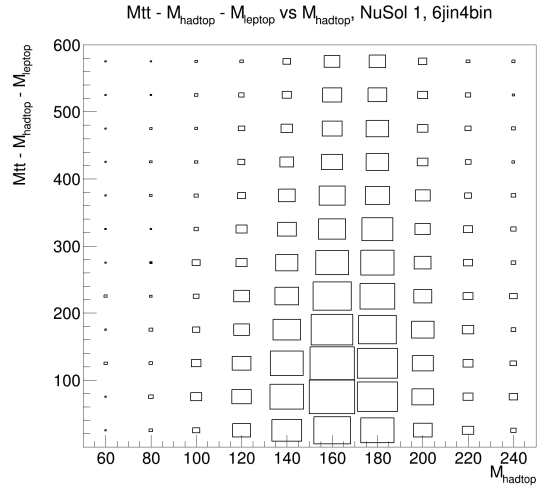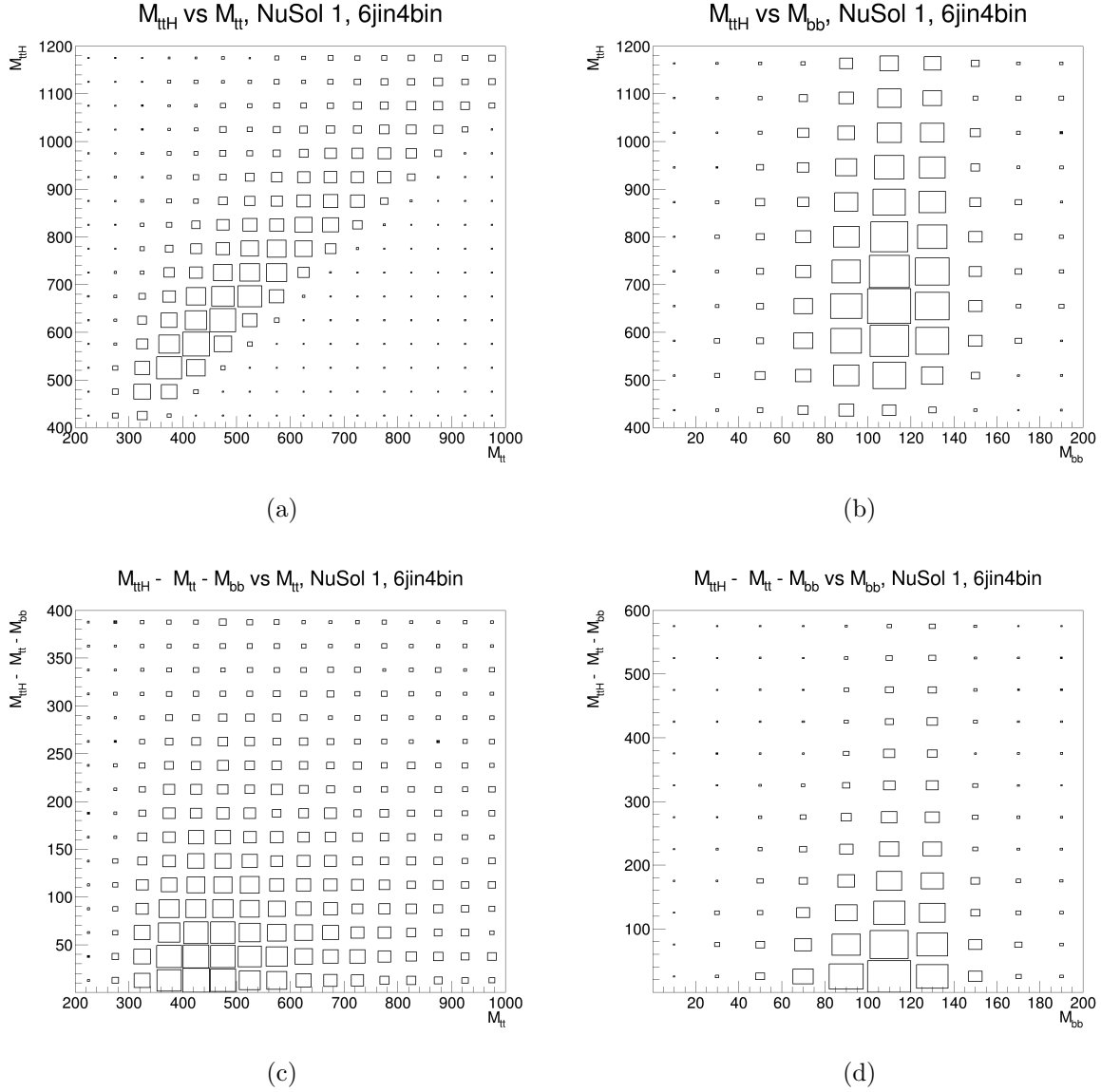
Figure 61: 2D distributions of $M_{t_h t_l}$ vs M_{t_l} (a) and M_{t_h} (b), and $M_{t\bar{t}} - M_{t_l} - M_{t_h}$ vs M_{t_l} (c) and M_{t_h} (d) in ttH MC events at $>= 6$ jets, $>= 4$ b-tagged jets region

reduced by introducing a new variable - $M_{t_h t_l b_1 b_2} - M_{t_h t_l} - M_{b_1 b_2}$. The two pdfs of $M_{t_h t_l b_1 b_2}$ and $M_{t_h t_l b_1 b_2} - M_{t_h t_l} - M_{b_1 b_2}$ are presented in figure 62, their correlations with $M_{t_l t_h}$ and $M_{b_1 b_2}$ are illustrated on figure 63.

<table>
<tr><td>(a)</td><td>(b)</td></tr>
</table>
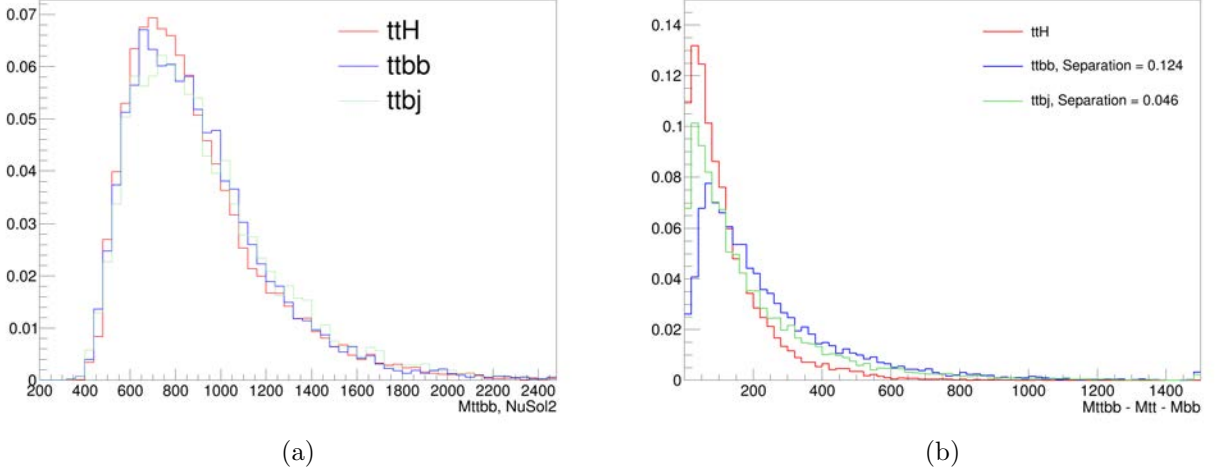
Figure 62: Pdfs of $M_{t_h t_l b_1 b_2}$ and $M_{t_h t_l b_1 b_2} - M_{t_h t_l} - M_{b_1 b_2}$ in ttH (red) and tt+jets (blue and green) MC events at $>= 6$ jets, $>= 4$ b-tagged jets region

The expression for kinematic terms for probability to be signal is becoming

$$
P_{kin}^{sig} = P^{sig}(M_H) P^{sig}(M_{t_l}) P^{sig}(M_{t_h} - M_{W_h}) P^{sig}(M_{W_h}) \times
$$
$$
\times P^{sig}(M_{t_h t_l} - M_{t_h} - M_{t_l}) P^{sig}(M_{t_h t_l b_1 b_2} - M_{t_l t_h} - M_{b_1 b_2}). \quad (60)
$$

and for the background probability

$$
P_{kin}^{bkg} = P^{sig}(M_{b_1 b_2}) P^{sig}(M_{t_l}) P^{sig}(M_{t_h} - M_{W_h}) P^{sig}(M_{W_h}) \times
$$
$$
\times P^{sig}(M_{t_h t_l} - M_{t_h} - M_{t_l}) P^{sig}(M_{t_h t_l b_1 b_2} - M_{t_l t_h} - M_{b_1 b_2}). \quad (61)
$$

5.6.5 Angular variables

Apart from the invariant masses, additional discrimination power can be provided by exploiting information on the spin of daughter particles of various resonances.

The most important variables, that allow to improve separation are

- $\cos \theta^*$ between b_1 in Higgs rest frame and the initial Higgs direction for the signal and $\cos \theta^*$ between j_1 in $j_1 j_2$ system rest frame and the initial $j_1 j_2$ direction for the background

Figure 63: 2D distributions of $M_{t_h t_l b_1 b_2}$ vs $M_{t_h t_l}$ (a) and $M_{b_1 b_2}$ (b), and $M_{t_h t_l b_1 b_2} - M_{t_l} - M_{t_h}$ vs $M_{t_h t_l}$ (c) and $M_{b_1 b_2}$ (d) in ttH MC events at $>= 6$ jets, $>= 4$ b-tagged jets region

- $\cos \theta^*$ between the Higgs in $t\bar{t}+$ Higgs system rest frame and $t\bar{t}+$ Higgs system for the signal and $\cos \theta^*$ between the dijet system $j_1 j_2$ in $t\bar{t}+j_1 j_2$ rest frame and $t\bar{t}+j_1 j_2$ for the background

The distributions of these variables are presented in figure 64.

(a)   (b)
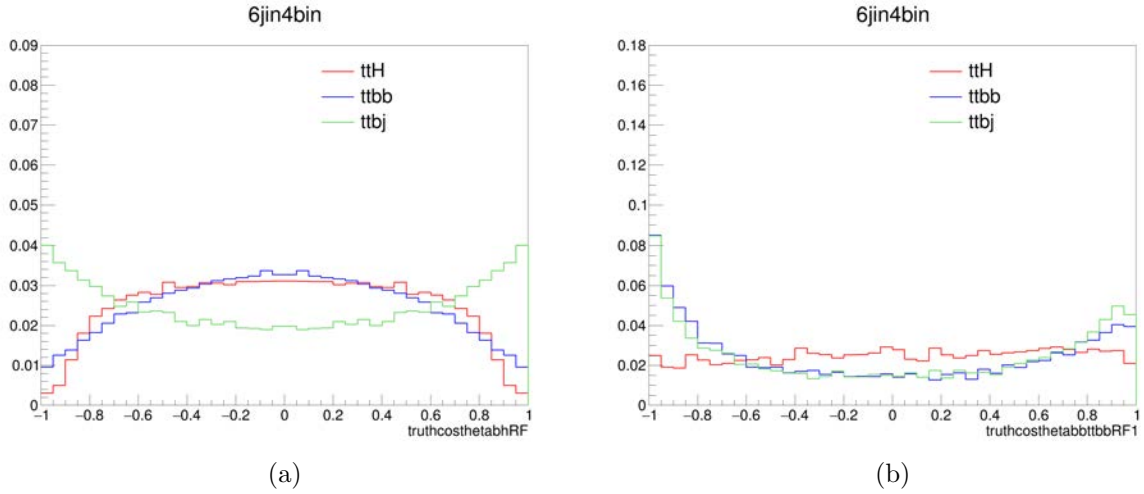
Figure 64: Pdfs of angular variables in ttH and tt + jets MC events at $>= 6$ jets, $>= 4$ b-tagged jets region

Apart from that the impact of following angular variables were tested:

- $\cos \theta^*$ between b_l in the leptonic top rest frame and the initial leptonic top direction.

- $\cos \theta^*$ between b_h in the hadronic top rest frame and the initial hadronic top direction.

- $\cos \theta^*$ between j_1 in the hadronic W rest frame and the initial hadronic W direction.

The pdfs of these additional angular variables are presented in figure 65.

(a)   (b)   (c)
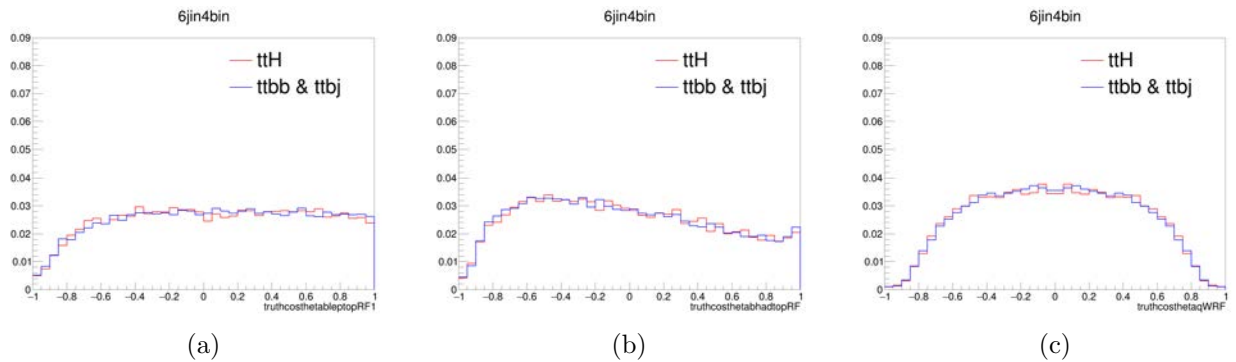
Figure 65: Pdfs of angular variables in ttH and tt + jets MC events at $>= 6$ jets, $>= 4$ b-tagged jets region

The impact of these additional on the final discriminating power was found to be subleading, thus only two angular variables. presented in figure 5.6.5 were used in the calculation.

To make use of the angular variables, the $P_{sig/bkg}$ is now defined as

$$P_{kin}^{sig/bkg} = P_{mass}^{sig/bkg} \times P_{ang}^{sig/bkg}, \tag{62}$$

where $P_{mass}^{sig/bkg}$ is defined as shown in equations 60 and 61, and $P_{ang}^{sig/bkg}$ is

$$P_{ang}^{sig/bkg} = P^{sig/bkg}(cos\theta_{b,bb}^*)P^{sig/bkg}(cos\theta_{bb,ttbb}^*) \tag{63}$$

The figure 66 is showing the correlation coefficients for the final variables set.
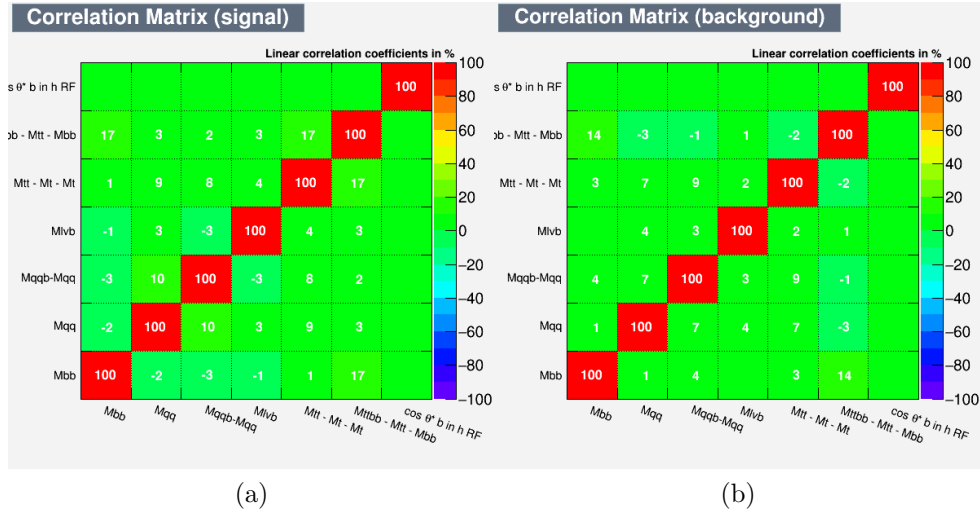
Figure 66: Correlation matrix for the variables used in the final discriminant calculation: signal (a) and background (b)

5.6.6   Missing jet hypothesis

The signal and background probabilities defined in equations 60 - 63 are built based on hypothesis that jets originating from Higgs, tops and W are reconstructed and passing the selection criteria (p_T and η cuts). However, it does not happen always: often there are jets outside acceptance region. It was shown that only in $\sim 40\%$ signal events in $>= 6$ jets, $>= 4$ b-jets channel there are all 6 jets that can be truth-matched to partons. In other $\sim 36\%$ cases only 5 jets are truth-matched, the rest $\sim 24\%$ are < 5 matches (see numbers for other regions in the table 12). This means, that to describe most of the events correctly one needs introducing additional "missing jet" hypothesis and combining it with the existing "all jets matched" hypothesis using as weights fractions of the cases when each of hypotheses has place (according to the table 12).

	All 6 jets	One missing jet	Two or more missing jets
>= 6 jets, >= 4 b-jets	40%	36%	24%
5 jets, >= 4 b-jets	-	58%	42%
>= 6 jets, 3 b-jets	24%	36%	40%

Table 12: Number of truth-matched jets in ttH events in different signal regions

>= 6 jets, >= 4 b-jets

For the main signal region in more than 70% of the cases it is subleading in p_T jet from W, that is missing, as shown in figure 67 (a). Therefore, this assumption is used for the "missing jet hypothesis".
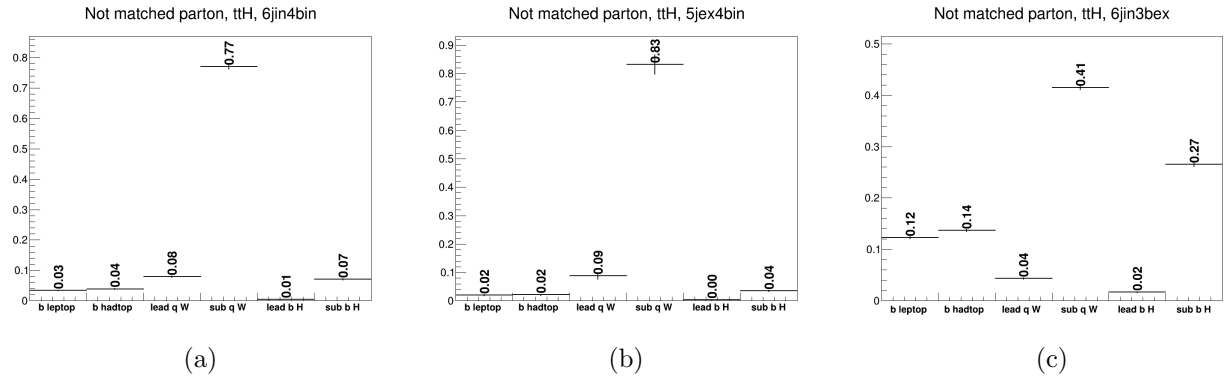
Figure 67: One missing jet: parton origin distribution in >= 6 jets, >= 4 b-jets (a), 5 jets, >= 4 b-jets (b) and >= 6 jets, 3 b-jets (c) regions

The probabilities to be signal and background are defined for this hypothesis in the same way as described by equations 60 - 63, but those probability terms that were calculated using pdfs that contain information from jets from W q_1 and q_2 are now replaced by pdfs build with one jet from W q_W and one extra jet q_X. For example, hadronic W invariant mass $M_{W_h}(q_1, q_2)$ is replaced by invariant mass $M_{W_h'}(q_W, q_X)$. The full list of variables used in signal probability calculation for two hypotheses is presented in table 13 (background probability is calculated accordingly).

5 jets, >= 4 b-jets

Introducing the "missing jet" hypothesis allows to calculate likelihood discriminant in this region. The calculation is performed in the similar way as for the >= 6 jets,>= 4 b-jets region, but with only one hypothesis - "missing jet", which is build on the same assumption as in the previous case that in most of the events the missing jet is the one that is originating from hadronic W (see 67(b)). The full set of variables used for signal probability calculation for 5 jets, >= 4 b-jets region is presented in table 14

All jets	Missing jet
$M_H(b_1, b_2)$	$M_H(b_1, b_2)$
$M_{t_l}(l, \nu, b_l)$	$M_{t_l}(l, \nu, b_l)$
$M_{W_h}(q_1, q_2)$	$M_{W'_h}(q_W, q_X)$
$[M_{t_h} - M_{W_h}](q_1, q_2, b_h)$	$[M_{t'_h} - M_{W'_h}](q_W, q_X, b_h)$
$[M_{t_h t_l} - M_{t_h} - M_{t_l}](l, \nu, b_l, q_1, q_2, b_h)$	$[M_{t'_h t_l} - M_{t'_h} - M_{t_l}](l, \nu, b_l, q_W, q_X, b_h)$
$[M_{t_h t_l b_1 b_2} - M_{t_l t_h} - M_H](b_1, b_2, l, \nu, b_l, q_1, q_2, b_h)$	$[M_{t'_h t_l b_1 b_2} - M_{t_l t'_h} - M_H](b_1, b_2, l, \nu, b_l, q_W, q_X, b_h)$
$cos\theta^*_{bb_{bbRF}}(b_1, b_2)$	$cos\theta^*_{bb_{bbRF}}(b_1, b_2)$
$cos\theta^*_{bb_{ttbbRF}}(b_1, b_2, l, \nu, b_l, q_1, q_2, b_h)$	$cos\theta^*_{bb_{tt'bbRF}}(b_1, b_2, l, \nu, b_l, q_W, q_X, b_h)$

Table 13: Pdfs used in calculation signal probability for "all jets" and "missing jet" hypotheses in $>= 6$ jets, $>= 4$ b-jets and $>= 6$ jets, 3 b-jets regions

Missing jet
$M_H(b_1, b_2)$
$M_{t_l}(l, \nu, b_l)$
$M_{t''_h}(q_W, b_h)$
$[M_{t''_h t_l} - M_{t''_h} - M_{t_l}](l, \nu, b_l, q_W, b_h)$
$[M_{t''_h t_l b_1 b_2} - M_{t_l t''_h} - M_H](b_1, b_2, l, \nu, b_l, q_W, b_h)$
$cos\theta^*_{bb_{bbRF}}(b_1, b_2)$
$cos\theta^*_{bb_{tt''bbRF}}(b_1, b_2, l, \nu, b_l, q_W, b_h)$

Table 14: Pdfs used in calculation signal probability for 5 jets, $>= 4$ b-jets region (One hypothesis: "missing jet")

$>= 6$ jets, 3 b-jets

In this region there are several hypotheses of which jet is missing (see 67(c)). As the hypothesis of jet from W is still the leading one, it was used for this region in this version of method exactly in the same way as done for the main signal region, $>= 6$ jets, $>= 4$ b-jets (with set of variables used as shown in figure 13). Introducing new hypotheses in this region can be considered as optimisation of the calculation.

5.6.7 Final discriminant

The final discriminant defined as shown in equation 47 in $>= 6$ jets, $>= 4$ b-jets, 5 jets, $>= 4$ b-jets and $>= 6$ jets, 3 b-jets signal regions, the distributions for the signal and background ($tt+ >= 1$ b-jet and $t\bar{t}+$ jets inclusive) are presented in figure 68

Distributions of three likelihood variables: logarithms to the base 10 of signal and background probabilities $\log_{10}$ProbSig and $\log_{10}$ProbBkg and the final discriminant in signal-enriched regions are presented in figures 70-74.
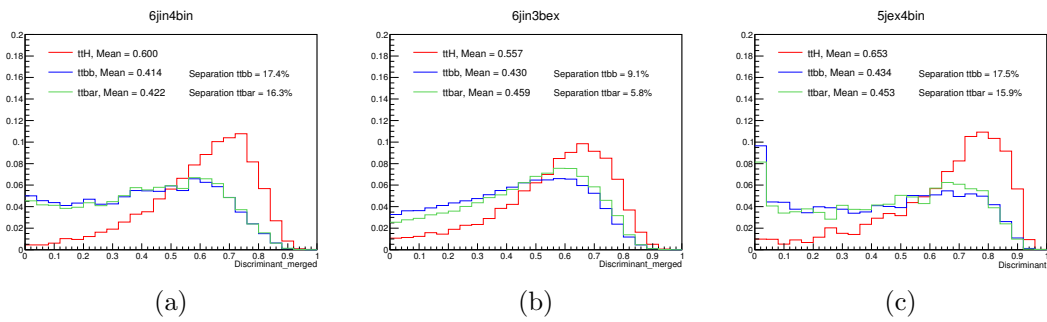
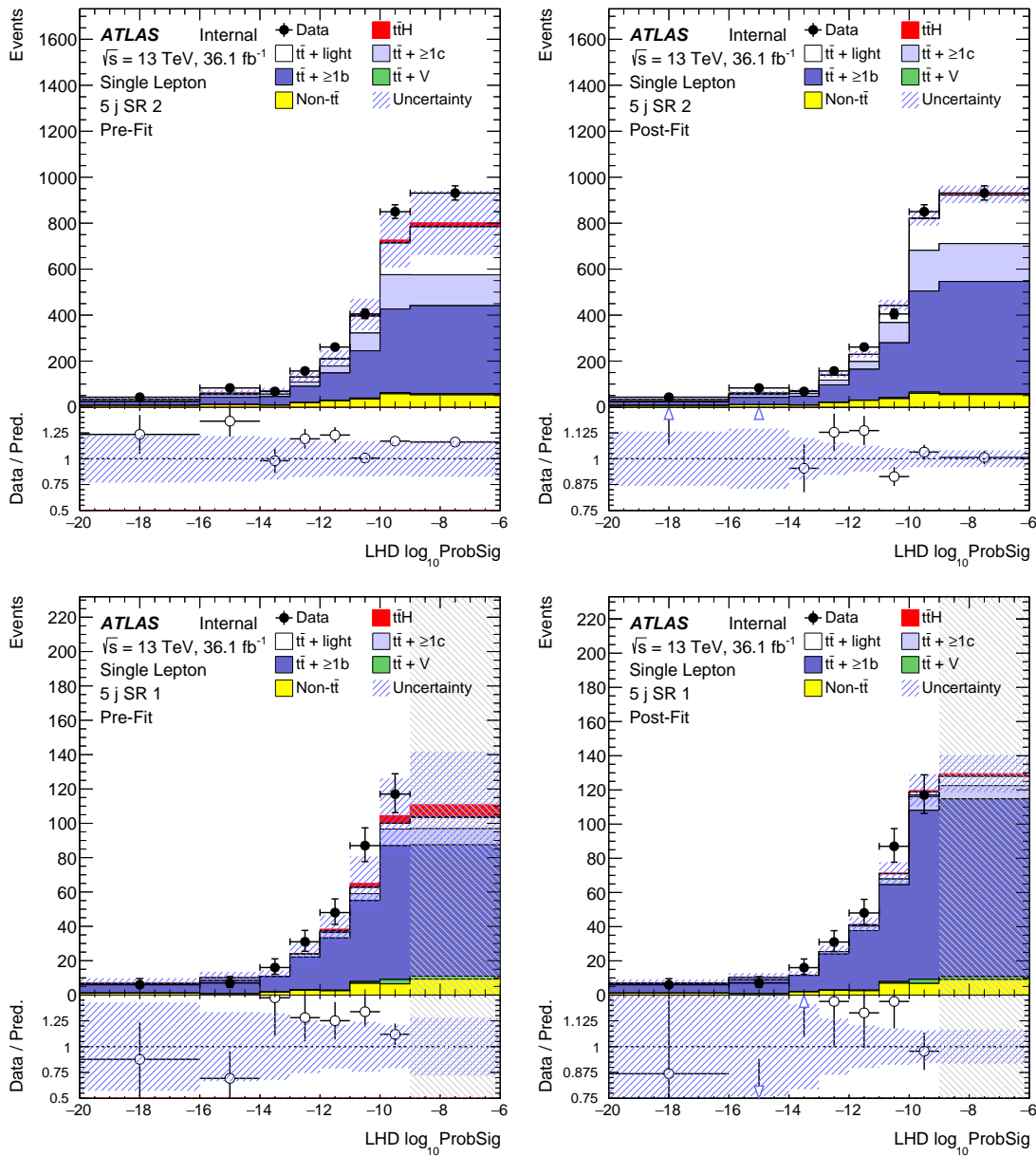Figure 68: Final discriminant in $>= 6$ jets, $>= 4$ b-jets (a), 5 jets, $>= 4$ b-jets (b) and $>= 6$ jets, 3 b-jets (c) regions

Figure 69: Pre-fit and post-fit distributions of $\log_{10}$ProbSig using fit based on background-only fit on distributions of H_T for 5 jets signal-enriched regions
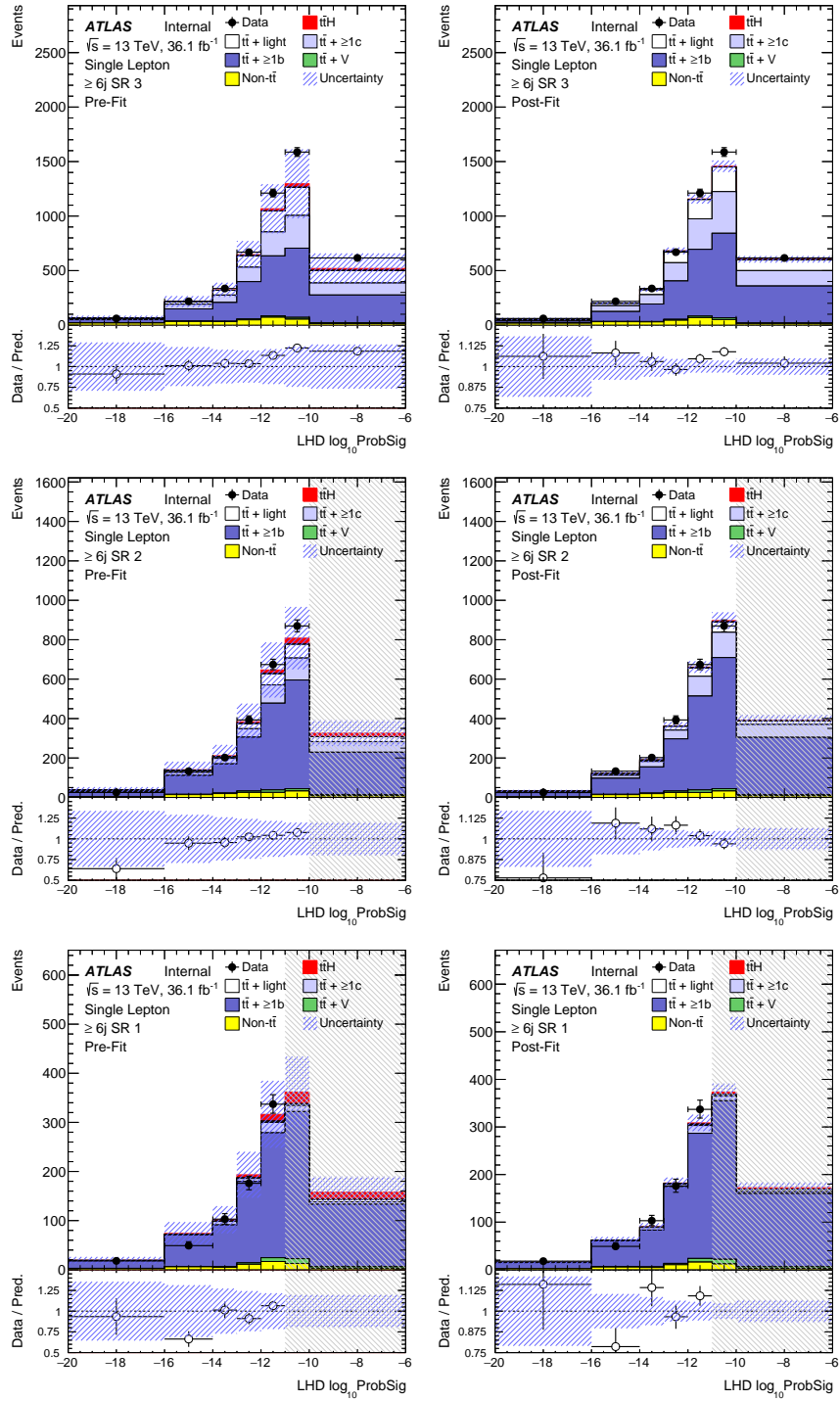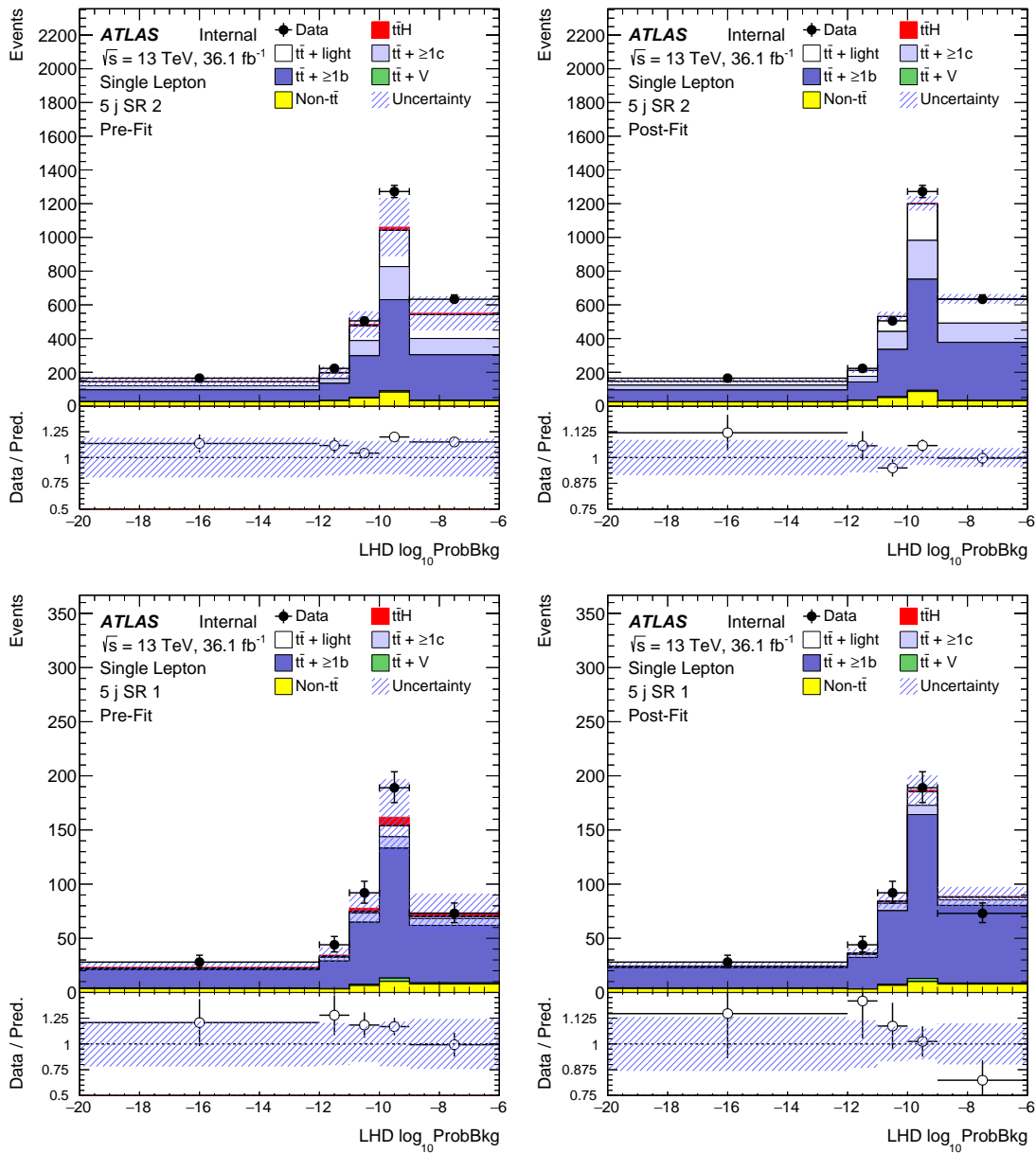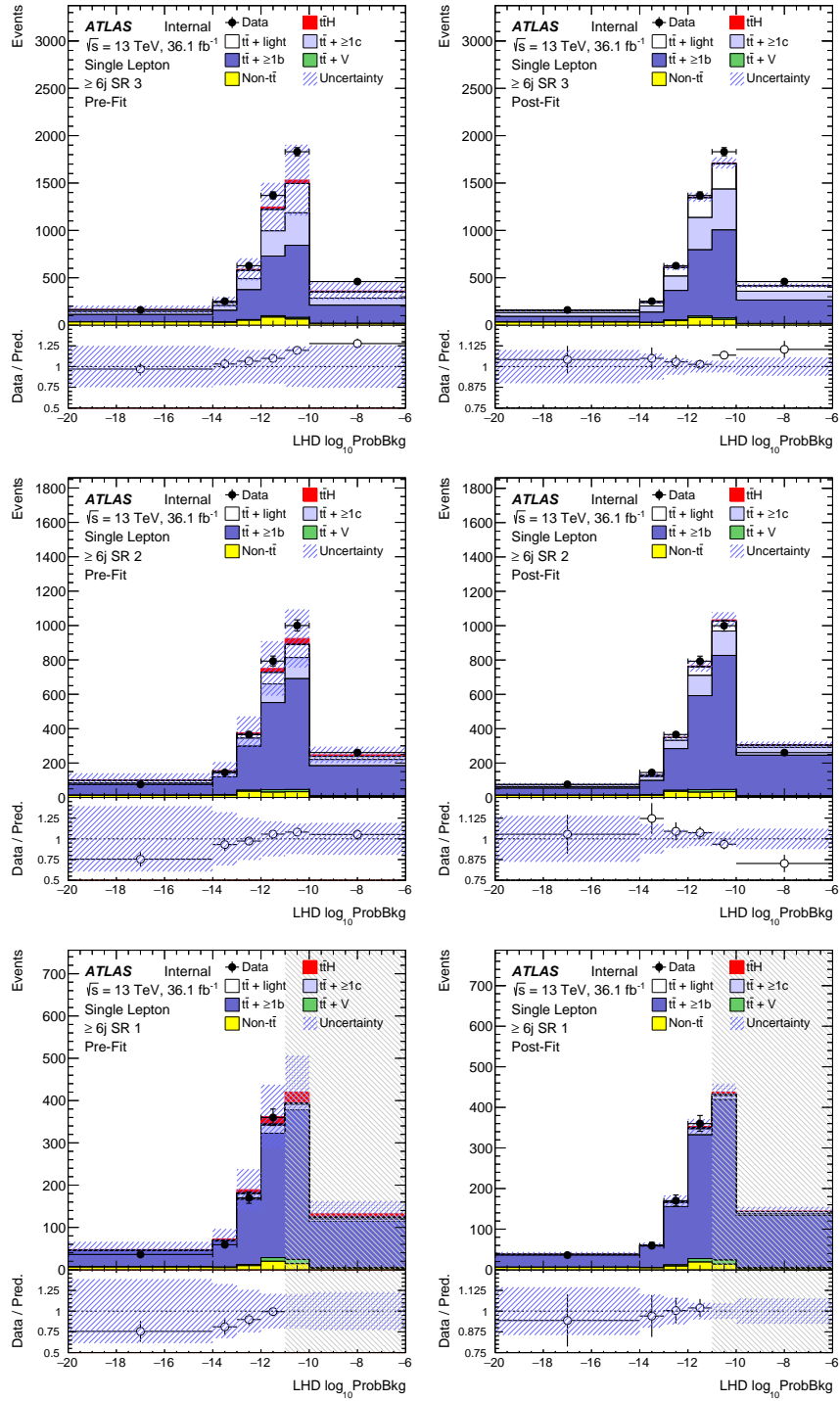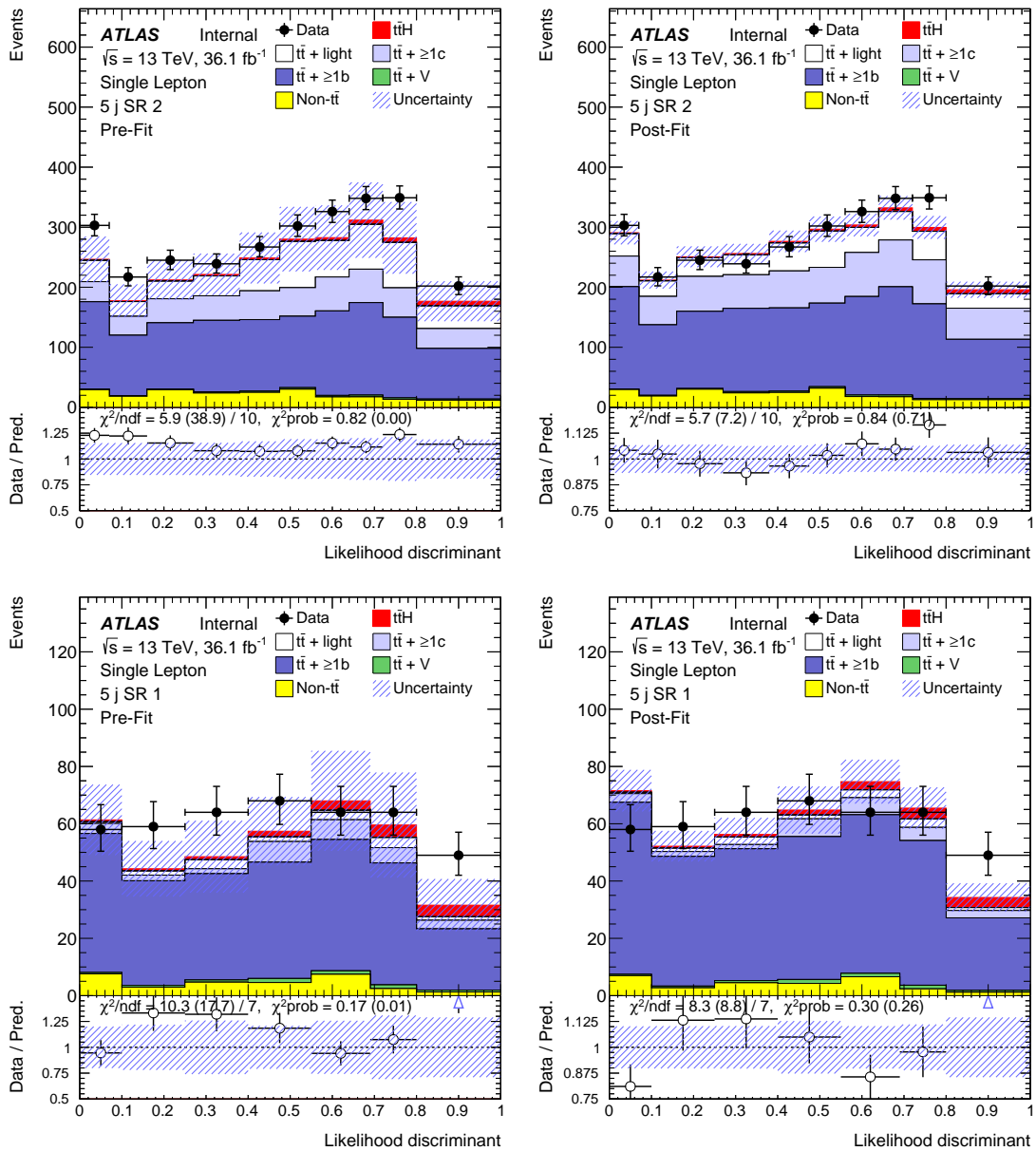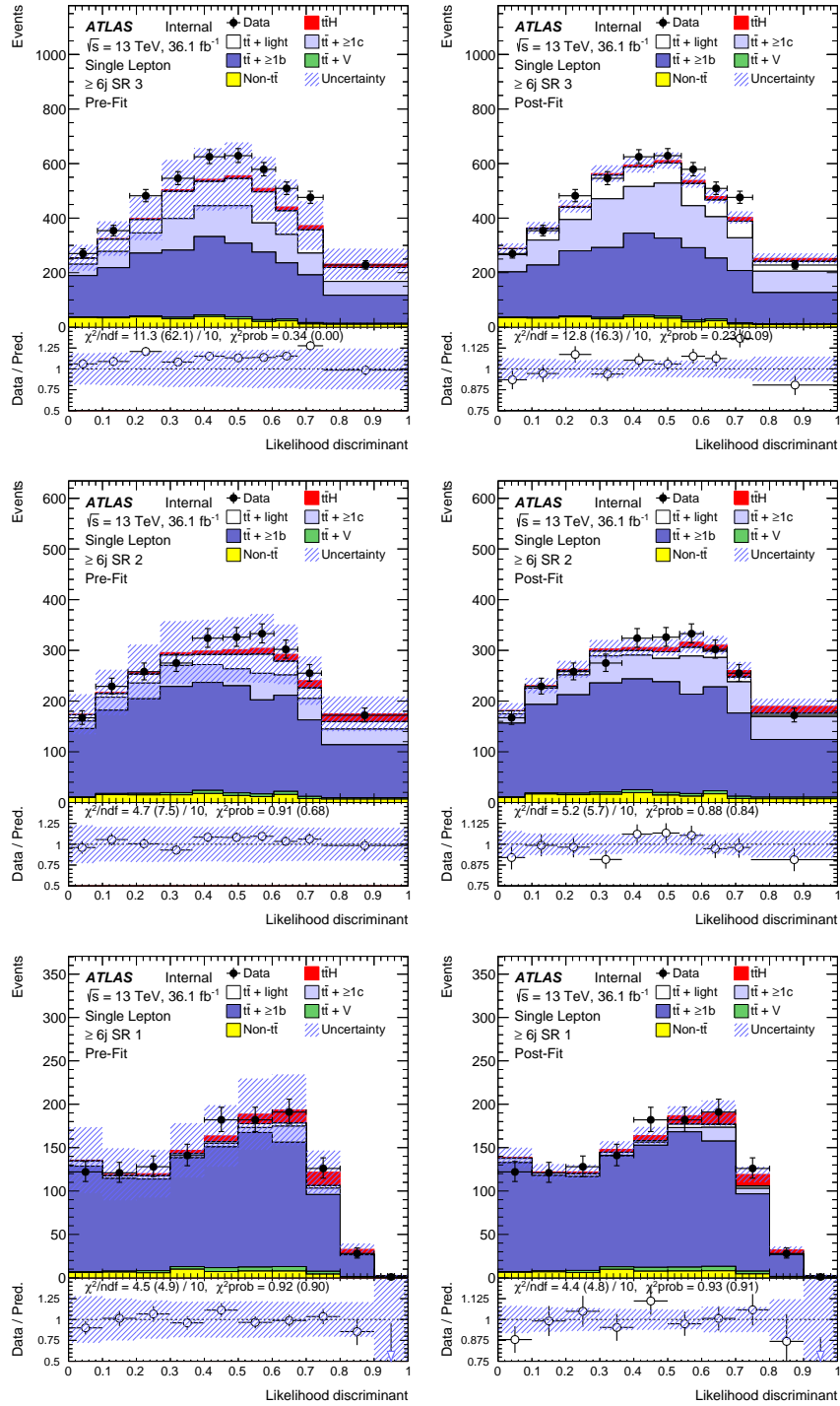
Figure 70: Pre-fit and post-fit distributions of $\log_{10}$ProbSig using fit based on background-only fit on distributions of H_T for ≥ 6 jets signal-enriched regions

Figure 71: Pre-fit and post-fit distributions of $\log_{10}$ProbBkg using fit based on background-only fit on distributions of H_T for 5 jets signal-enriched regions

Figure 72: Pre-fit and post-fit distributions of $\log_{10}$ProbBkg using fit based on background-only fit on distributions of H_T for ≥ 6 jets signal-enriched regions

Figure 73: Pre-fit and post-fit distributions of the final likelihood discriminant using fit based on background-only fit on distributions of H_T for 5 jets signal-enriched regions

Figure 74: Pre-fit and post-fit distributions of the final likelihood discriminant using fit based on background-only fit on distributions of H_T for ≥ 6 jets signal-enriched regions

Variable	Separation [%]
LHD	15.7
Reconstruction BDT output	9.9
$\Delta R_{bb}^{\mathrm{avg}}$	7.9
m_H	6.0
N_{30}^{Higgs}	5.8
$\Delta R_{bb}^{\mathrm{max}\ p_T}$	5.3
$\Delta \eta_{jj}^{\mathrm{max}\ \Delta\eta}$	5.1
$m_{bb}^{\mathrm{min}\ \Delta R}$	5.0
$\Delta R_{H,t\bar{t}}$	5.0
$\Delta R_{H,\mathrm{lep\ top}}$	3.4
Aplanarity	3.0
$\Delta R_{\mathrm{Higgs\ bb}}$	2.9
$m_{H,b_{\mathrm{lep\ top}}}$	2.9
$H1$	1.3
B_{j^5}	0.3

Table 15: Values of separation power for classification BDT input variables in ≥ 6 jets signal-enriched region SR1.

5.6.8   Performance

The LHD final output is becoming an input of classification BDT together with the reconstruction BDT, MEM and other kinematic variables.

Input of a variable is evaluated using the signal-to-background separation power that is defined as

$$S = \frac{1}{2} \sum_i^{N_{bins}} \frac{(N_i^S - N_i^B)^2}{N_i^S + N_i^B} \tag{64}$$

The LHD was found to be the most discriminating single variable in the analysis. Table 15 is presenting all variables that are used as input of the classification BDT and corresponding values of the separation power in ≥ 6 jets signal-enriched region SR1.

The classification BDT distributions without and with LHD as input variable for the $t\bar{t}H$ signal and the $t\bar{t} + $ jets background in ≥ 6 jets signal-enriched regions SR1 and SR2 are shown in figures 75 and 76. Corresponding values of separation power and relative gain due to LHD is summarised in table 16. ROC curves of the classification BDT without and with LHD for the same regions are shown in figure 77.

5.7   Systematic uncertainties

Systematic uncertainties affect normalisation and shape of the MC samples, so they are taken into account as nuisance parameters in the fit procedure. There are two kinds of systematic uncertainties considered: experimental and modelling uncertainties. The main
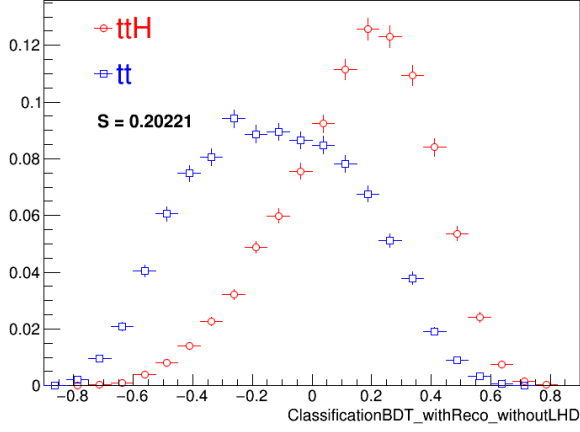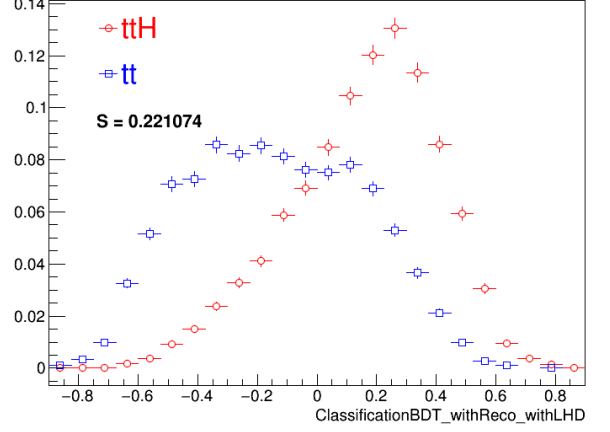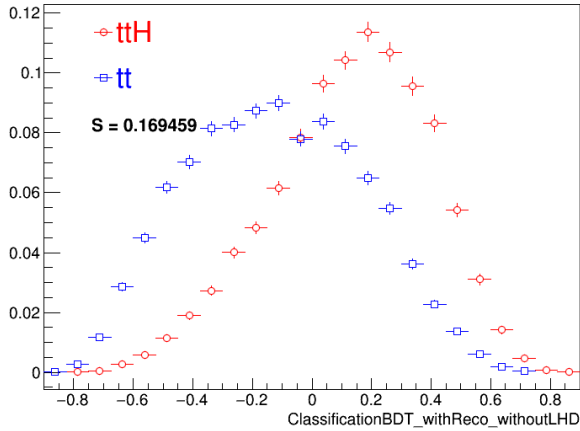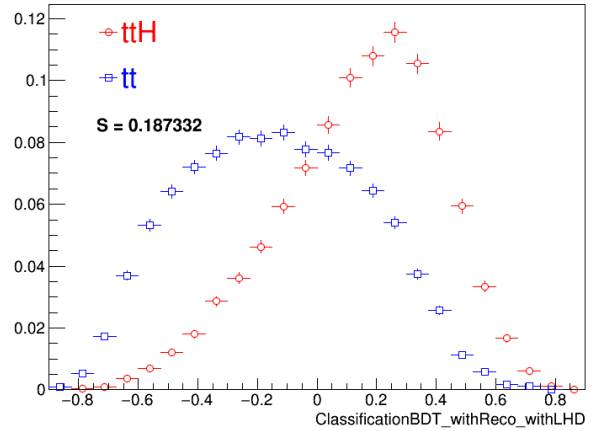
Figure 75: Distributions of the classification BDT without (a) and with (b) LHD output for the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background in ≥ 6 jets signal-enriched region SR1
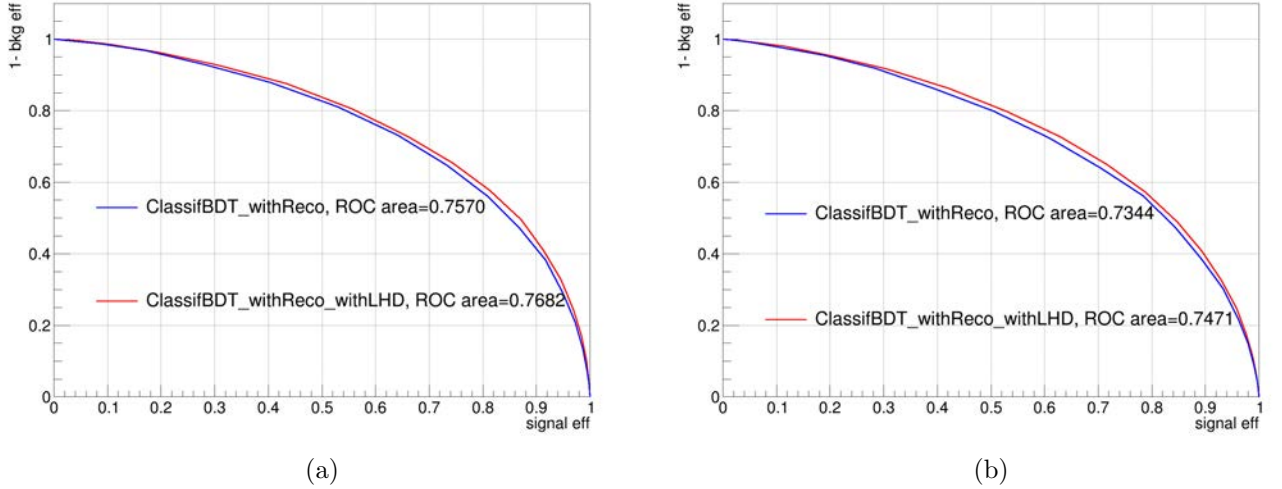

Figure 76: Distributions of the classification BDT without (a) and with (b) LHD output for the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background in ≥ 6 jets signal-enriched region SR2

Region	Separation without LHD [%]	Separation with LHD [%]	Relative gain
≥ 6 jets SR1	20.2	22.1	9.4
≥ 6 jets SR2	16.9	18.7	10.7

Table 16: Values of the separation power of the classification BDT distributions without and with LHD and relative gain due to the LHD for the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background in ≥ 6 jets signal-enriched regions SR1 and SR2.

(a)   (b)

Figure 77: ROC curves for the classification BDTs without and with LHD output for the $t\bar{t}H$ signal and the $t\bar{t}$ + jets background in the signal regions ≥ 6 jets signal-enriched region SR1 (a) and SR2 (b)

source of uncertainties in this analysis are those related to modelling of $t\bar{t}$ + b-jets events. The full list of considered uncertainties is presented in table 17.

5.7.1 Experimental uncertainties

Luminosity

The systematic uncertainty on the 2015+2016 integrated luminosity is 2.1%. The estimation was done using preliminary calibration of the luminosity scale using x-y beam-separation scans performed in August 2015 and May 2016. The detailed description of the estimation method can be found in [76]. The uncertainties for the 2015 and the 2016 datasets are considered to be correlated.

The systematic uncertainty on the luminosity affects normalisation of all MC samples.

Leptons

Uncertainties related to leptons are originating from triggering, reconstruction, identification, isolation, as well as the lepton momentum scale and resolution.

Systematic uncertainty	Type	Components
Luminosity	N	1
Reconstructed Objects		
Electron trigger+reco+ID+isolation	SN	4
Electron energy scale+resolution	SN	2
Muon trigger+reco+ID+isolation	SN	10
Muon momentum scale+resolution+saggita	SN	5
Taus detector, insitu and model	SN	3
Pileup modelling	SN	1
Jet vertex tagger	SN	1
Jet energy scale	SN	20
Jet energy resolution	SN	2
Missing transverse energy scale+resolution	SN	3
b-tagging efficiency	SN	30
c-mistag rate	SN	20
Light-mistag rate	SN	60
Mistag extrapolation $c \to \tau$	SN	1
Background and Signal Model		
$t\bar{t}$ cross section	N	1
$t\bar{t}+ \geq 1c$: normalisation	N (free floating)	1
$t\bar{t}+ \leq 2b$: normalisation	N (free floating)	1
$t\bar{t}+ \geq 3b$: normalisation	N	1
$t\bar{t}+\geq 1b$: NLO Shape	SN	9
$t\bar{t}+\geq 1c$: NLO Shape	SN	1
$t\bar{t}+\geq 1b$: 4F vs 5F Shape	S	1
$t\bar{t}$ modelling: residual Radiation	SN	3
$t\bar{t}$ modelling: residual NLO generator	SN	3
$t\bar{t}$ modelling: residual parton shower+hadronisation	SN	3
W+jets normalisation	N	3
Z+jets normalisation	N	3
Single top cross section	N	1
Single top model	SN	2
Diboson normalisation	N	1
Fakes normalization	SN	6
$t\bar{t}V$ cross section	N	4
$t\bar{t}V$ modelling	SN	2
tZ cross section	N	2
tWZ cross section	N	1
$t\bar{t}WW$ cross section	N	2
4-tops cross section	N	1
$tHjb$ cross section	N	3
WtH cross section	N	2
$t\bar{t}H$ cross section	N	2
$t\bar{t}H$ branching ratios	N	3
$t\bar{t}H$ modelling	SN	1

Table 17: The list of systematic uncertainties. N - the uncertainty considered to be affecting normalisation only, SN - both normalisation and shape of distributions are affected. Some of the uncertainties are split into several components for a more accurate treatment.

The reconstruction, identification and isolation of electrons and muons, as well as the efficiency of the trigger used to record the events, are slightly different between data and simulation. This is taken into account by so called scale factors (SF), which are used as weights applied to the MC events. The SF are obtained from $Z \to l^+l^-$ data and MC events.

Other uncertainties are related difference of lepton momentum scale and resolution in data in MC. To adjust these difference, the corrections are applied, using $Z \to l^+l^-$, $J/\psi \to l^+l^-$ and $W \to e\nu$ events.

The lepton related uncertainties have very small effect for this analysis.

Jets

The uncertainties associated with jets are related to efficiency of jet reconstruction, identification, uncertainties on the jet energy scale (JES), jet energy resolution (JER) and the JVT.

The uncertainty on the jet energy scale is estimated using information from test-beam data, collision data and simulation as described in [77]. It consists of 20 components, corresponding to different uncertainty sources: difference in in-situ techniques of JES calibration (statistical, modelling, detector and mixed), corrections on pile-up mismodelling, flavour of jets (due to the fact that response of the calorimeter is different to jets originated from quarks or gluons), high-p_T jets measurement. The JES uncertainty is about 5.5% for jets with $p_T = 25$ GeV and decreasing with increasing of jet p_T . For central jets with p_T in the range of 100 GeV - 1.5 TeV it is below 1.5%. This is one of the main systematic uncertainties related to reconstructed objects.

The JER uncertainties was measured in Run 1 data and simulated dijet events. They were found to agree within 10% [78]. Additional uncertainties were obtained from extrapolation from Run 1 to Run 2 conditions [77].

Missing transverse energy

The E_T^{miss} uncertainties are propagated from those related to leptons and jet energy scales and resolutions. Additional uncertainties related to the resolution and scale of the soft term of E_T^{miss} are considered. These uncertainties are having small effect on the analysis.

Flavour tagging

Efficiencies of b-, c- and light tagging in simulation are corrected by applying SF to match those in data. The scale factors depend on jet p_T in the case of b- and c-jets and on η for the light-jets. Efficiencies are calibrated for several bins of b-tagging algorithm weight (defined with the WPs of 60%, 70%, 77% and 85% and edge points that are interpreted as 0% and 100% efficiencies).

5.7.2 Modelling uncertainties

tt+jets modelling

Modelling of $t\bar{t}$ +jets events is the main source of systematic uncertainties in this analysis; the full list of corresponding uncertainties is presented in table 18.

Systematic source	How evaluated	$t\bar{t}$ categories
$t\bar{t}$ cross-section	$\pm 6\%$	All, correlated
NLO generator (*residual*)	Powheg-Box +Pythia 8 vs. SHERPA 5FS with b-filter	All, uncorrelated
Radiation (*residual*)	Variations of μ_{R}, μ_{F}, *hdamp* and A14 parameters	All, uncorrelated
PS & hadronisation (*residual*)	Powheg-Box + Pythia 8 vs. Powheg-Box +Herwig7	All, uncorrelated
$t\bar{t} + b\bar{b}$ renorm. scale *reweighting*	Up or down a by factor of two	$t\bar{t}+{\geq}1b$
$t\bar{t} + b\bar{b}$ resumm. scale *reweighting*	Vary μ_{Q} from $H_{\mathrm{T}}/2$ to μ_{CMMPS}	$t\bar{t}+{\geq}1b$
$t\bar{t} + b\bar{b}$ global scales *reweighting*	Set μ_{Q}, μ_{R}, and μ_{F} to μ_{CMMPS}	$t\bar{t}+{\geq}1b$
$t\bar{t} + b\bar{b}$ shower recoil *reweighting*	Alternative model scheme	$t\bar{t}+{\geq}1b$
$t\bar{t} + b\bar{b}$ PDF *reweighting*	CT10 vs. MSTW or NNPDF	$t\bar{t}+{\geq}1b$
$t\bar{t}+ \geq 3b$ reweighting	Up or down by 50%	$t\bar{t}+ \geq 3b$
$t\bar{t}+ \geq 1b$ 4F vs 5F shape	Compare Powheg+Pythia8 with Sherpa 4F shape	$t\bar{t}+ \geq 1b$
$t\bar{t} + b\bar{b}$ MPI	Up or down by 50%	$t\bar{t}+{\geq}1b$
$t\bar{t} + c\bar{c}$ ME calculation	MG5_aMC + Herwig++ inclusive vs. ME prediction	$t\bar{t}+{\geq}1c$

Table 18: The systematic uncertainties on the $t\bar{t}$+jets modelling. For the $t\bar{t}+ \geq 1b$ background, the inclusive $t\bar{t}$ sample is reweighted to a NLO $t\bar{t}+b\bar{b}$ prediction; uncertainties on the inclusive sample are labelled *residual*, while those on the NLO prediction are labelled *reweighting*.

For the inclusive cross-section of $t\bar{t}$ NNLO+NNLL production an uncertainty of $\pm 6\%$ is applied according to [79]. An uncertainty taking into account difference between generators is retrieved by comparing the nominal Powheg-Box + Pythia 8 sample with the one generated with Sherpa. An uncertainty corresponding to the difference in parton shower and hadronisation models is evaluated by comparing the prediction from Powheg-Box generator interfaced to either to Pythia 8 or Herwig7. An uncertainty on modelling of initial and final state radiation is obtained with two Powheg-Box + Pythia 8 samples with different values of hdamp and A14 eigentune parameters. All these uncertainties except the one on the inclusive $t\bar{t}$ cross-section are considered to be uncorrelated for $t\bar{t} + \geq 1b$, $t\bar{t} + \geq 1c$ and $t\bar{t} + $ light.

For alternative samples the fractions of $t\bar{t} + \geq 1b$ subcategories are reweighted to match the prediction of Sherpa+OpenLoops in the same way as the nominal sample, as described in 5.4.2. The remaining differences in the samples are considered as residual uncertainties. Additionally, uncertainties on the reweighting procedure were applied by

varying renormalisation, factorisation and resummation scales of reweighting.

To take into account uncertainty on the choice of PDF, two different sets are considered: nominal NNPDF and alternative MSTW [80]. Another uncertainty is obtained with an alternative shower recoil scheme. Additionally, a 50% uncertainty is associated to the events not included in the original NLO calculation but coming from Multi Parton Interaction (MPI).

Another uncertainty is applied to take into account difference between $t\bar{t} + \geq 1c$ calculated in the matrix element with the default approach of using charm jets produced in the parton shower. This uncertainty is derived by comparing the nominal $t\bar{t} +$ jets sample with $t\bar{t} + c\bar{c}$ NLO matrix element calculation with MadGraph5_aMC@NLO interfaced to Herwig++.

There is no prior uncertainty on normalisation of $t\bar{t} + \geq 1b$ and $t\bar{t} + \geq 1c$, those parameters are let to float freely in the fit.

To take into account significance difference in the $t\bar{t} + \geq 3b$ component with and without reweighting to Sherpa+OpenLoops, an additional 50% uncertainty on the normalisation of $t\bar{t} + \geq 3b$ events is considered.

An additional uncertainty considers difference in two alternative schemes of $\geq 1b$ and $t\bar{t} + \geq 1c$ production, known as four-flavour (4F) and five-flavour (5F) schemes. It is obtained by comparing the shapes of $t\bar{t} + b\bar{b}$, $t\bar{t} + b$, $t\bar{t} + B$ and $t\bar{t} + \geq 3b$ distributions of the nominal PowhegBox + Pythia8 sample with the Sherpa 4F.

Misidentified leptons

For the data-driven non-prompt lepton background estimation an uncertainty of 50% is used. It is considered uncorrelated across 5 jet and 6 jet regions as well as between electron and muon channels.

Other backgrounds modelling

Uncertainties of 40% for W + jets cross-section and 35% for Z + jets cross-section are considered. They are estimated by varying Sherpa generator parameters. In the case of Z + jets, a data-driven scale factor of ~ 1.3 is additionally applied. A 30% uncertainty is considered for W + heavy flavour jets events by comparing the MadGraph5_aMC@NLO and Sherpa samples.

For cross-section of single top production a theoretical uncertainty of $^{+5\%}_{-4\%}$ is applied, according to [82]-[84]. As for the $t\bar{t}$ background, an uncertainty associated with initial and final-state radiation is used. An additional uncertainty is taking into account interference between the $t\bar{t}$ and Wt processes at NLO [70].

For the diboson background, a 50% normalisation uncertainty on cross-section and additional jet production is applied. The uncertainty on the $t\bar{t}V$ NLO cross-section of 15% is used. For the $t\bar{t}t\bar{t}$ background a normalisation uncertainty 50% is considered.

Signal modelling

The uncertainty of $^{+10\%}_{-13\%}$ on the $t\bar{t}H$ signal cross-section is applied. This includes contributions from scale and PDF uncertainties, considered to be uncorrelated [85]-[89]. Uncertainties on the Higgs boson branching ratios are considered, for $H \to b\bar{b}$ it is 2.2%. An uncertainty on the choice of parton shower and hadronisation model is obtained from the difference between MG5_aMC@NLO interfaced to either Pythia 8 (nominal model) or Herwig++.

5.8 Fit procedure

The ratio of measured signal to the Standard Model prediction, or signal strength, $\mu = \sigma/\sigma_{SM}$ is obtained with a fitting procedure based on the ROOSTATS framework [90]. The statistical method used in this analysis is based on a binned maximum likelihood function $\mathcal{L}(\mu, \theta)$, where θ - set of nuisance parameters (NP), corresponding to the considered systematic uncertainties. This function is a product of Poisson probability terms over the bins of the input distributions including the number of data events and expected signal and background yields, taking into account the effects of the systematic uncertainties:

$$\mathcal{L}(\mu, \theta) = \prod_j \prod_{i=bin} \frac{(\mu s_i(j) + b_i(j))^{N_i^{(j)}}}{N_i^{(j)}!} e^{-\mu s_i(j) - b_i(j)} \prod_\theta func(\theta|0, 1), \qquad (65)$$

where func is given by Gaussian or log-normal PDFs, the value $\theta = 0$ corresponds to the nominal value of the prediction, $\theta = \pm 1$ correspond to ± 1 deviation of given systematic uncertainty. $N_i^{(j)}$ is the number of observed events in the i-th bin of the j-th signal region, $s_i(j)$ and $b_i(j)$ are expected numbers of signal and background events, that are given by function of θ.

The test statistic is defined as a profile likelihood ratio

$$q_\mu = -2\ln(\mathcal{L}(\mu, \hat{\hat{\theta}}_\mu)/\mathcal{L}(\hat{\mu}, \hat{\theta})), \qquad (66)$$

where $\hat{\mu}$ and $\hat{\theta}$ are the values of the parameters that maximise the likelihood function (with the constraints $0 \leq \hat{\mu} \leq \mu$), and $\hat{\hat{\theta}}_\mu$ are the values of NP that maximise the likelihood for a given value of μ.

The test statistics is used to determine the compatibility of the data measurement with the background-only hypothesis (μ=0) and predict upper limit on μ using the method called confidence level (CL_S) approach.

A test of a hypothesized value of μ with statistic $q_{/mu}$ is a measure of discrepancy between the data and the hypothesis, with higher values of $q_{/mu}$ corresponding to increasing disagreement. The disagreement is quantified with p-value

$$p_\mu = \int_{q_\mu,obs}^\infty f(q_\mu|\mu)dq_\mu, \qquad (67)$$

where q_μ, obs is the statistic value observed in data, $f(q_\mu|\mu)$ - pdf of q_μ under the signal strength μ assumption.

The compatibility of the result with signal plus background hypothesis is then given by

$$p_{s+b} = f(q \geq q_{obs}|1) = \int_{q_\mu,obs}^{\infty} f(q_\mu|1)dq_\mu, \tag{68}$$

and the one for the background-only hypothesis is

$$p_b = f(q \geq q_{obs}|0) = \int_{-\infty}^{q_\mu,obs} f(q_\mu|0)dq_\mu, \tag{69}$$

Distributions of test statistics under signal plus background and background-only hypotheses and corresponding p-values are presented in figure 78.
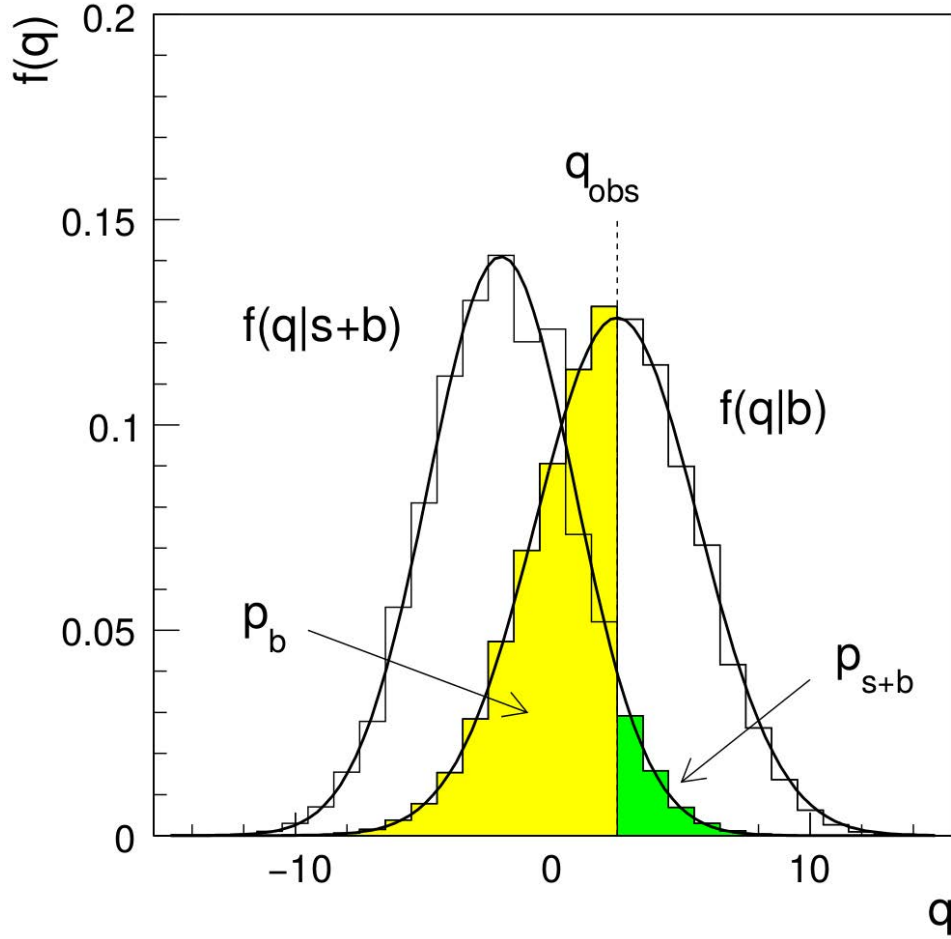

Figure 78: The distribution of the statistic $q_\mu = -2\ln(\mathcal{L}_{s+b}/\mathcal{L}_b)$ under signal plus background ($\mu = 1$) and background-only ($\mu = 0$) hypotheses. The p-values for both hypotheses are also shown with respect to an observed value of the statistic q_{obs} [91].

The confidence level for the signal hypothesis is then defined as

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{p_{s+b}}{1 - p_b}. \tag{70}$$

The values of μ are considered to be are excluded at 95% confidence level, when $CL_s < 0.05$.

The method to obtain test statistics is based on replacing the ensemble of simulated data sets by a single representative, known as Asimov data set. This is a simulated data set with observed parameters set to their expected values. If this data set is used to evaluate the estimators for the parameters, the true parameter values are obtained [91].

5.9   Results

5.9.1   Expected results

The expected results obtained by performing fits to the Asimov data set with the signal plus background (S+B) hypothesis are presented in this section.

Constraints on nuisance parameters (NPs) corresponding to theoretical and instrumental systematic uncertainties resulting from the fits are shown in figures 79 and 80.

The expected uncertainties on the normalisation factors for $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ components and the signal strength are shown in figure 81.

Systematical uncertainties covariance matrix is presented in figure 82.

5.9.2   Fit to data

Fits to data using BDT disctributions are presented in this section. Figures 83 and 84 are showing NP corresponding to theoretical and instrumental systematic uncertainties. Normalisation factors for $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ components are presented in figure 85. Systematical uncertainties covariance matrix is presented in figure 86. Yields comparison between data and prediction after the fit to the measured data is shown in figure 87.

	$\geq 6j$ BR($t\bar{t}$+light)	$\geq 6j$ BR($t\bar{t}+ \geq 1c$)	$\geq 6j$ BR($t\bar{t} + 1b$)	$\geq 6j$ SR 3	$\geq 6j$ SR 2	$\geq 6j$ SR 1
$t\bar{t}$ + light	114 000 ± 4500	4090 ± 451	2040 ± 168	690 ± 106	157 ± 43.7	16.0 ± 6.87
$t\bar{t}$ + $\geq 1c$	39 300 ± 5120	4360 ± 650	1770 ± 286	1130 ± 183	388 ± 89.4	41.3 ± 18.0
$t\bar{t}$ + $\geq 1b$	15 700 ± 1300	3250 ± 250	4200 ± 312	2420 ± 158	1930 ± 122	1070 ± 57.6
$t\bar{t}$ + W	436 ± 59.4	38.4 ± 6.84	15.4 ± 2.22	9.97 ± 1.99	4.54 ± 0.869	1.26 ± 0.298
$t\bar{t}$ + Z	583 ± 68.0	81.3 ± 11.3	53.6 ± 6.69	41.6 ± 5.48	36.6 ± 4.79	24.8 ± 3.32
Wt channel	4480 ± 1300	269 ± 79.4	169 ± 57.9	86.0 ± 41.7	61.2 ± 37.9	21.3 ± 13.8
t channel	1240 ± 190	84.1 ± 53.1	47.5 ± 19.4	38.2 ± 19.4	16.9 ± 9.18	8.55 ± 3.75
Other top	162 ± 13.8	18.0 ± 4.45	11.6 ± 2.71	10.8 ± 3.43	9.53 ± 3.22	6.24 ± 2.11
VV & V + jets	4590 ± 1360	449 ± 156	177 ± 53.0	85.9 ± 27.5	36.5 ± 11.9	15.6 ± 4.73
Fakes & NP (μ)	363 ± 193	29.4 ± 20.2	28.7 ± 19.5	2.10 ± 2.59	0 ± 0	0 ± 0
Fakes & NP (e)	1060 ± 422	129 ± 61.8	130 ± 59.0	30.5 ± 22.1	0 ± 0	0 ± 0
$tHjb$	5.18 ± 0.661	1.10 ± 0.183	1.12 ± 0.167	0.797 ± 0.128	0.853 ± 0.132	0.719 ± 0.109
WtH	18.3 ± 1.79	3.79 ± 0.448	4.22 ± 0.439	2.98 ± 0.369	2.80 ± 0.349	1.58 ± 0.213
$t\bar{t}H$ ($H \to b\bar{b}$)	49.5 ± 140	16.4 ± 44.3	17.8 ± 50.7	15.8 ± 43.2	15.8 ± 43.4	12.1 ± 33.8
$t\bar{t}H$ ($H \to WW$)	22.0 ± 61.7	2.34 ± 6.52	0.787 ± 2.31	0.656 ± 1.85	0.295 ± 0.874	0.0820 ± 0.227
$t\bar{t}H$ ($H \to$ other)	17.4 ± 48.3	2.13 ± 6.04	0.873 ± 2.51	0.703 ± 2.00	0.300 ± 0.880	0.202 ± 0.568
Total	182 000 ± 996	12 800 ± 233	8670 ± 143	4560 ± 96.3	2660 ± 73.6	1220 ± 41.0
Data	181706	12778	8576	4698	2641	1222

Table 19: Post-fit event yields for the regions with ≥ 6 jets for a combined BDT-based fit to data with the S+B hypothesis.

|	5j BR($t\bar{t}$+light)	5j, BR($t\bar{t}$+ $\geq$1c)	5j BR($t\bar{t}$ + 1b)	5j SR 2	5j SR 1	1 Higgs + 1 Top SR
$t\bar{t}$ + light	184 000 $\pm$ 5010	2500 $\pm$ 212	3800 $\pm$ 227	505 $\pm$ 71.4	15.9 $\pm$ 9.49	170 $\pm$ 37.4
$t\bar{t}$ + $\geq$1c	38 500 $\pm$ 5400	1580 $\pm$ 239	2260 $\pm$ 366	513 $\pm$ 94.2	23.1 $\pm$ 12.2	184 $\pm$ 36.8
$t\bar{t}$ + $\geq$1b	13 700 $\pm$ 1210	981 $\pm$ 88.8	4110 $\pm$ 309	1470 $\pm$ 103	345 $\pm$ 24.1	259 $\pm$ 31.5
$t\bar{t}$ + W	270 $\pm$ 36.9	7.29 $\pm$ 1.53	9.89 $\pm$ 1.47	1.69 $\pm$ 0.446	0.173 $\pm$ 0.0557	5.72 $\pm$ 1.14
$t\bar{t}$ + Z	311 $\pm$ 38.0	13.4 $\pm$ 2.10	37.0 $\pm$ 5.10	17.6 $\pm$ 2.34	6.04 $\pm$ 1.15	11.0 $\pm$ 2.02
Wt channel	7340 $\pm$ 1890	135 $\pm$ 27.4	261 $\pm$ 73.4	74.1 $\pm$ 36.9	7.74 $\pm$ 3.66	25.0 $\pm$ 14.6
t channel	2840 $\pm$ 282	56.2 $\pm$ 30.1	87.0 $\pm$ 35.4	23.5 $\pm$ 5.99	7.17 $\pm$ 3.78	1.58 $\pm$ 1.84
Other top	268 $\pm$ 16.8	4.58 $\pm$ 0.725	13.9 $\pm$ 1.27	4.55 $\pm$ 0.688	2.17 $\pm$ 0.381	5.03 $\pm$ 2.00
VV & V + jets	7300 $\pm$ 2270	124 $\pm$ 79.2	252 $\pm$ 84.2	59.7 $\pm$ 21.2	12.1 $\pm$ 6.88	32.2 $\pm$ 9.59
Fakes & NP (μ)	594 $\pm$ 302	5.24 $\pm$ 4.51	18.0 $\pm$ 14.8	8×10^{-5} $\pm$ 2.83×10^{-6}	0 $\pm$ 0	19.4 $\pm$ 12.3
Fakes & NP (e)	3770 $\pm$ 1800	130 $\pm$ 67.1	257 $\pm$ 127	57.7 $\pm$ 31.7	0 $\pm$ 0	11.1 $\pm$ 6.42
$tHjb$	9.25 $\pm$ 1.11	0.829 $\pm$ 0.126	3.22 $\pm$ 0.401	2.01 $\pm$ 0.273	1.04 $\pm$ 0.150	0.0771 $\pm$ 0.0217
WtH	10.1 $\pm$ 1.01	0.724 $\pm$ 0.148	3.10 $\pm$ 0.345	1.32 $\pm$ 0.178	0.374 $\pm$ 0.0739	1.98 $\pm$ 0.248
$t\bar{t}H$ ($H \to b\bar{b}$)	29.2 $\pm$ 83.8	3.04 $\pm$ 8.80	12.4 $\pm$ 35.6	7.66 $\pm$ 21.6	3.07 $\pm$ 8.73	2.92 $\pm$ 8.01
$t\bar{t}H$ ($H \to WW$)	7.06 $\pm$ 20.2	0.200 $\pm$ 0.628	0.267 $\pm$ 0.816	0.0615 $\pm$ 0.195	0.006 60 $\pm$ 0.0220	0.223 $\pm$ 0.641
$t\bar{t}H$ ($H \to$ other)	7.03 $\pm$ 20.1	0.302 $\pm$ 0.924	0.402 $\pm$ 1.19	0.127 $\pm$ 0.380	0.0231 $\pm$ 0.0678	0.254 $\pm$ 0.730
Total	259 000 $\pm$ 1020	5550 $\pm$ 150	11 100 $\pm$ 165	2740 $\pm$ 66.6	424 $\pm$ 22.9	729 $\pm$ 38.8
Data	259320	5465	11095	2798	426	740

Table 20: Post-fit event yields for the regions with 5 jets for a combined BDT-based fit to data with the S+B hypothesis.
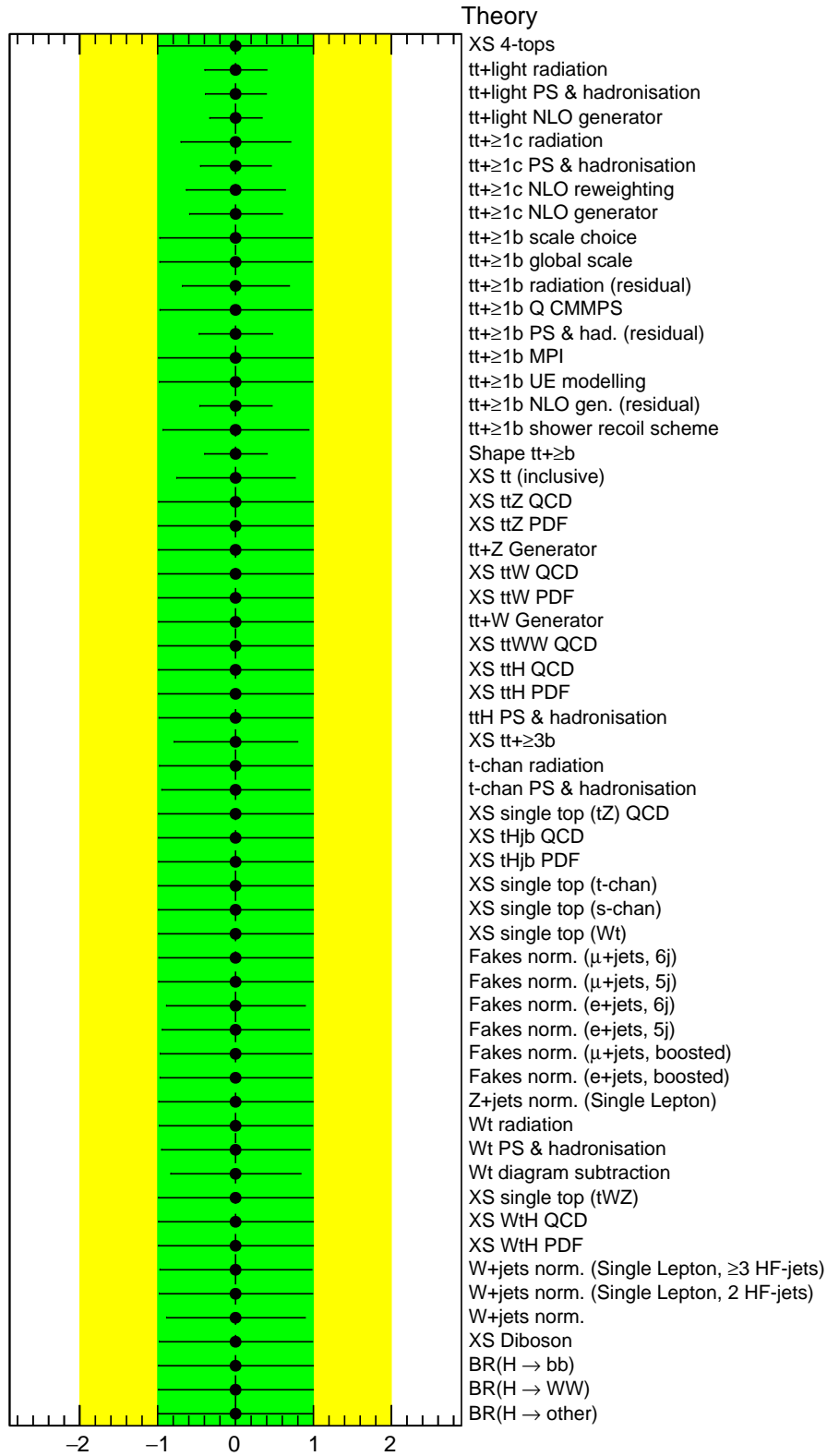
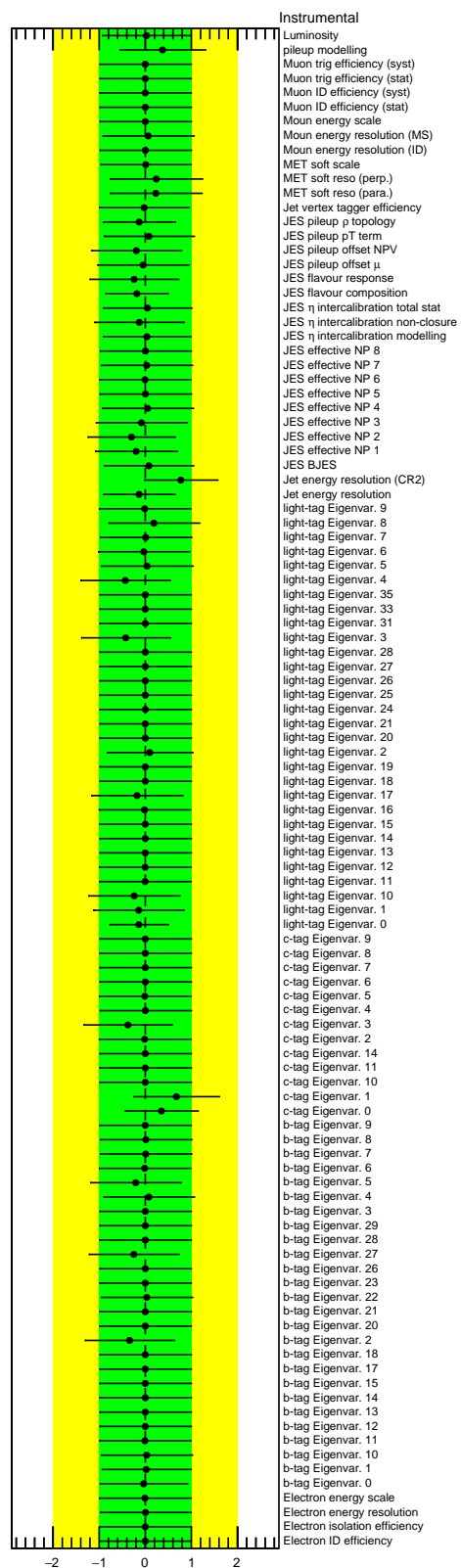Figure 79: Nuisance parameter corresponding to theoretical systematic uncertainties for BDT-based fits to Asimov data set with the S+B hypothesis.
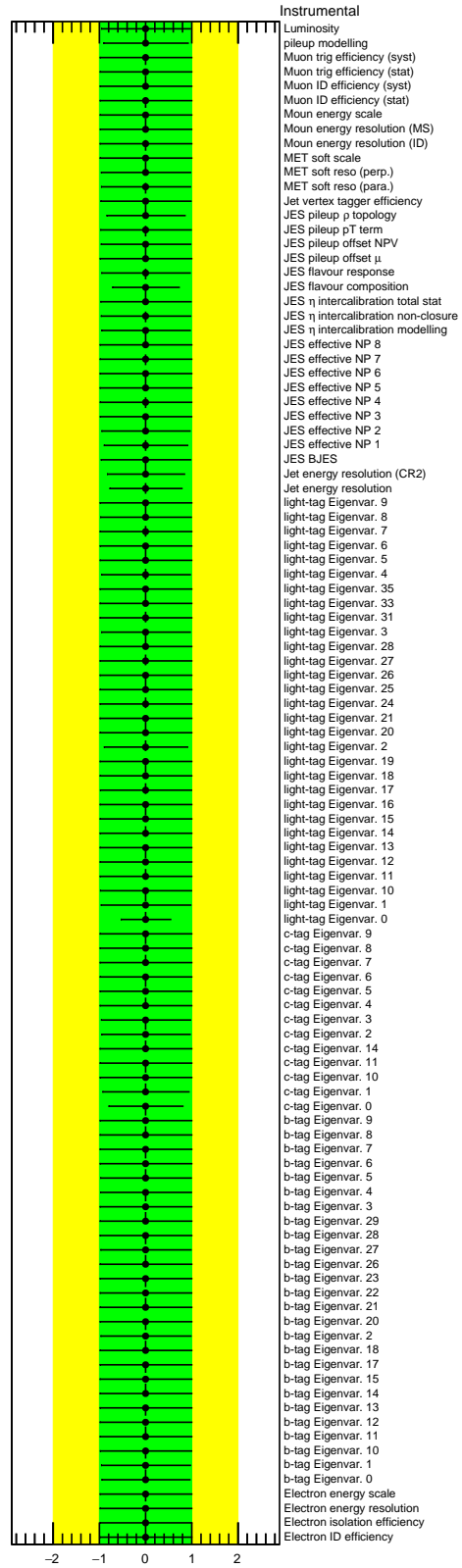
Figure 80: Nuisance parameter corresponding to instrumental systematic uncertainties for BDT-based fits to Asimov data set with the S+B hypothesis.
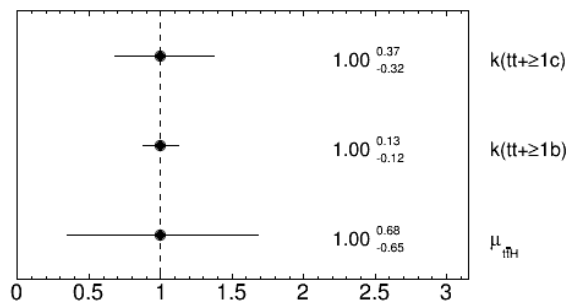
Figure 81: Expected uncertainties on the normalisation factors for $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ components and the signal strength for BDT-based fits to Asimov data set with the S+B hypothesis
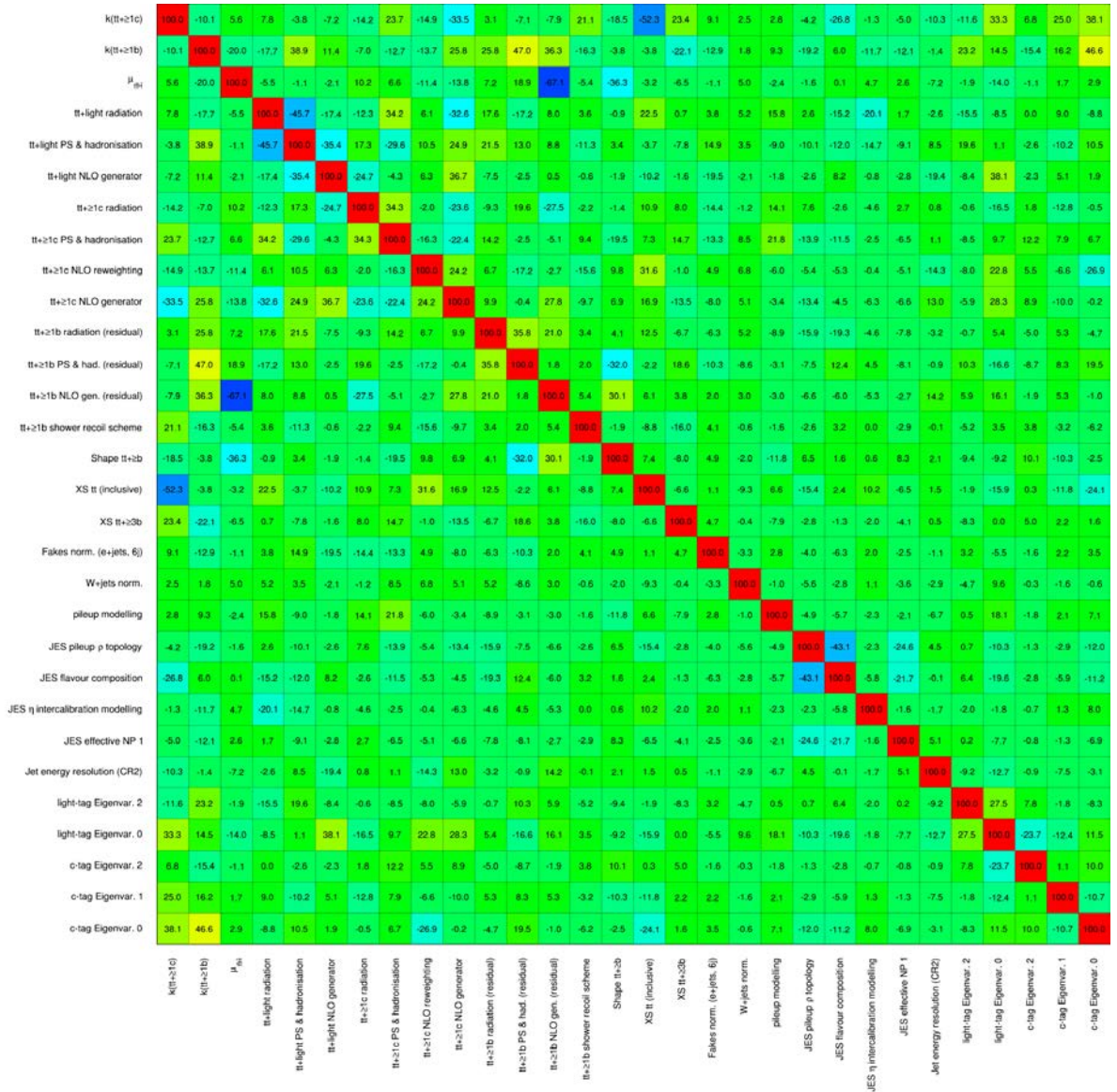
Figure 82: Systematic covariance matrix with BDT-based fit to Asimov data set with the S+B hypothesis.
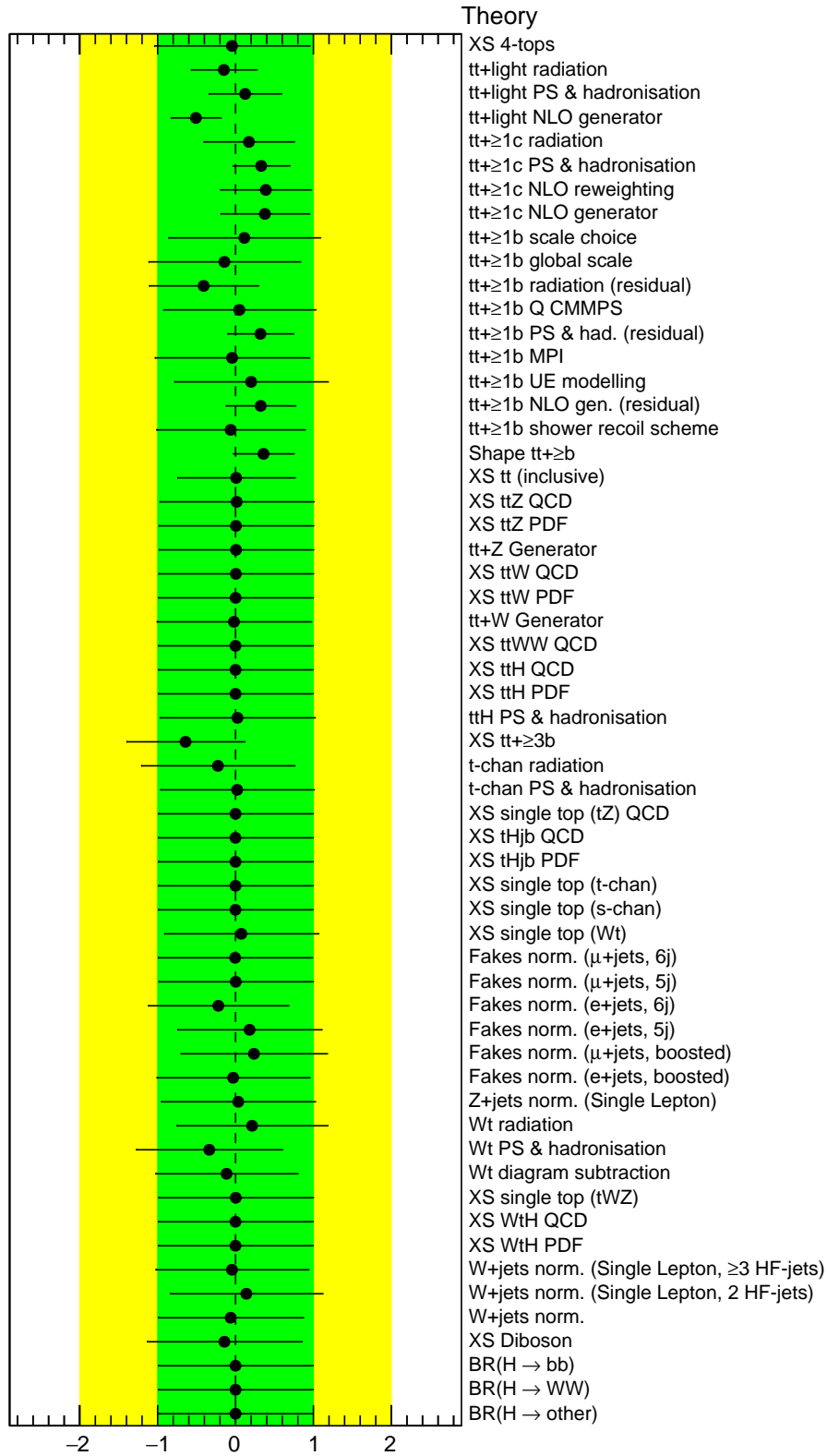
Figure 83: Nuisance parameter corresponding to theoretical systematic uncertainties for BDT-based fits to data.

Figure 84: Nuisance parameter corresponding to instrumental systematic uncertainties for BDT-based fits to data.
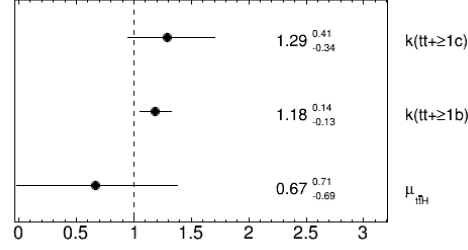
Figure 85: Normalisation factors for $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ components for BDT-based fits to data.
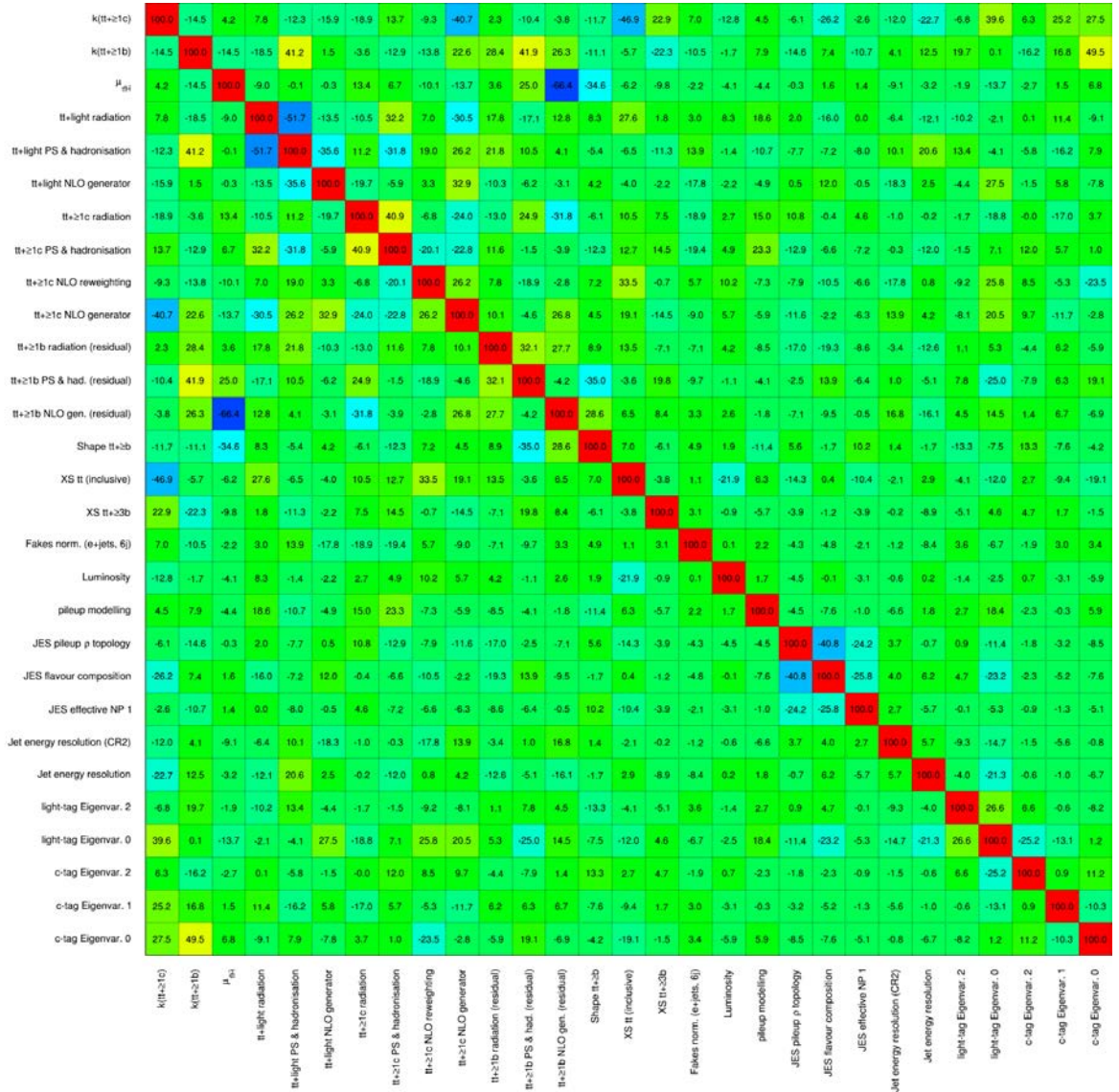
Figure 86: Systematic covariance matrix in BDT-based fit to data.
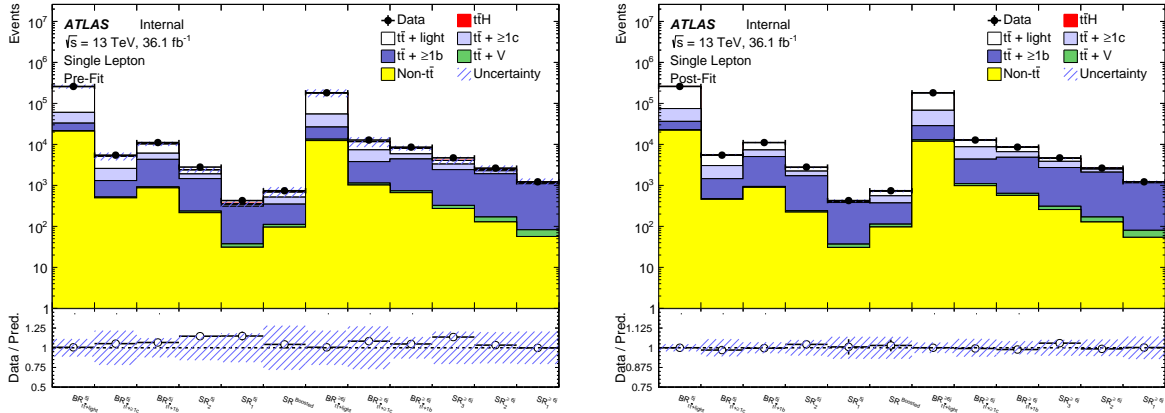
Figure 87: Yield summary prefit and postfit after the fit to the measured data
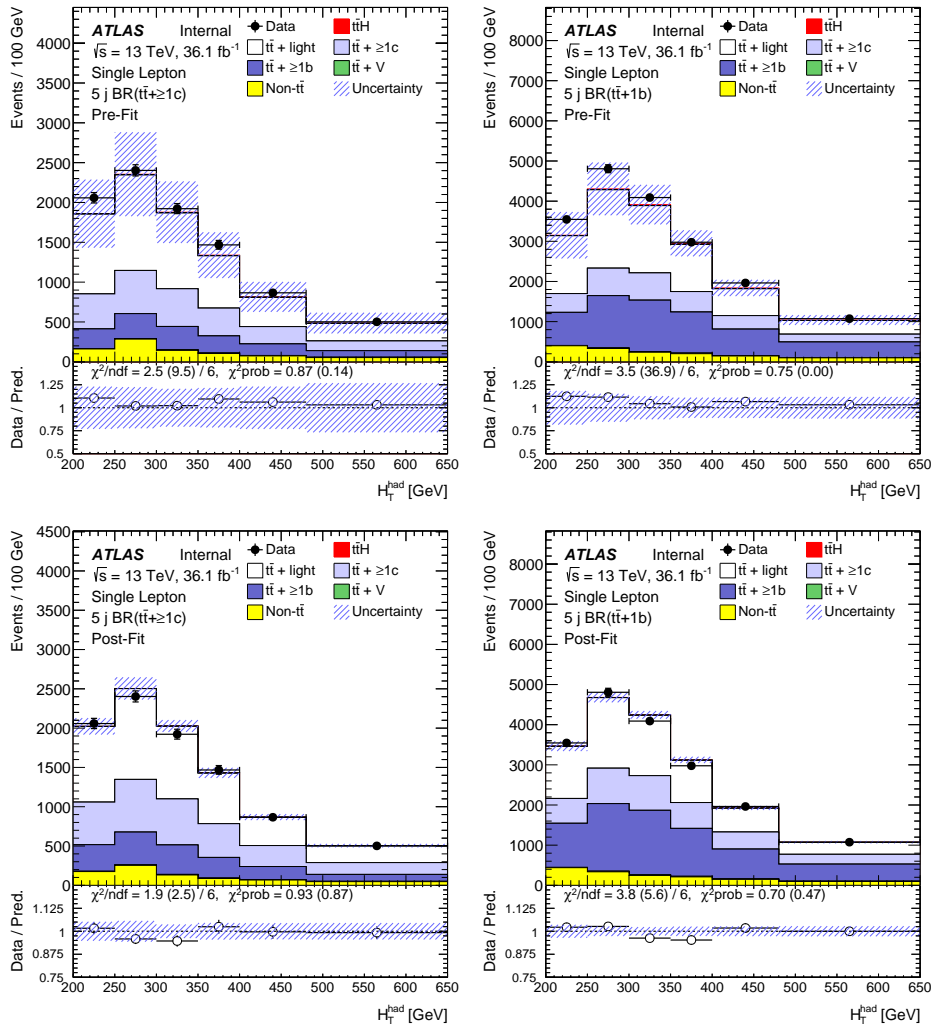
Figure 88: Prefit (top) and postfit (bottom) from fits to data corresponding to a combined BDT-based fit to data with the S+B hypothesis.
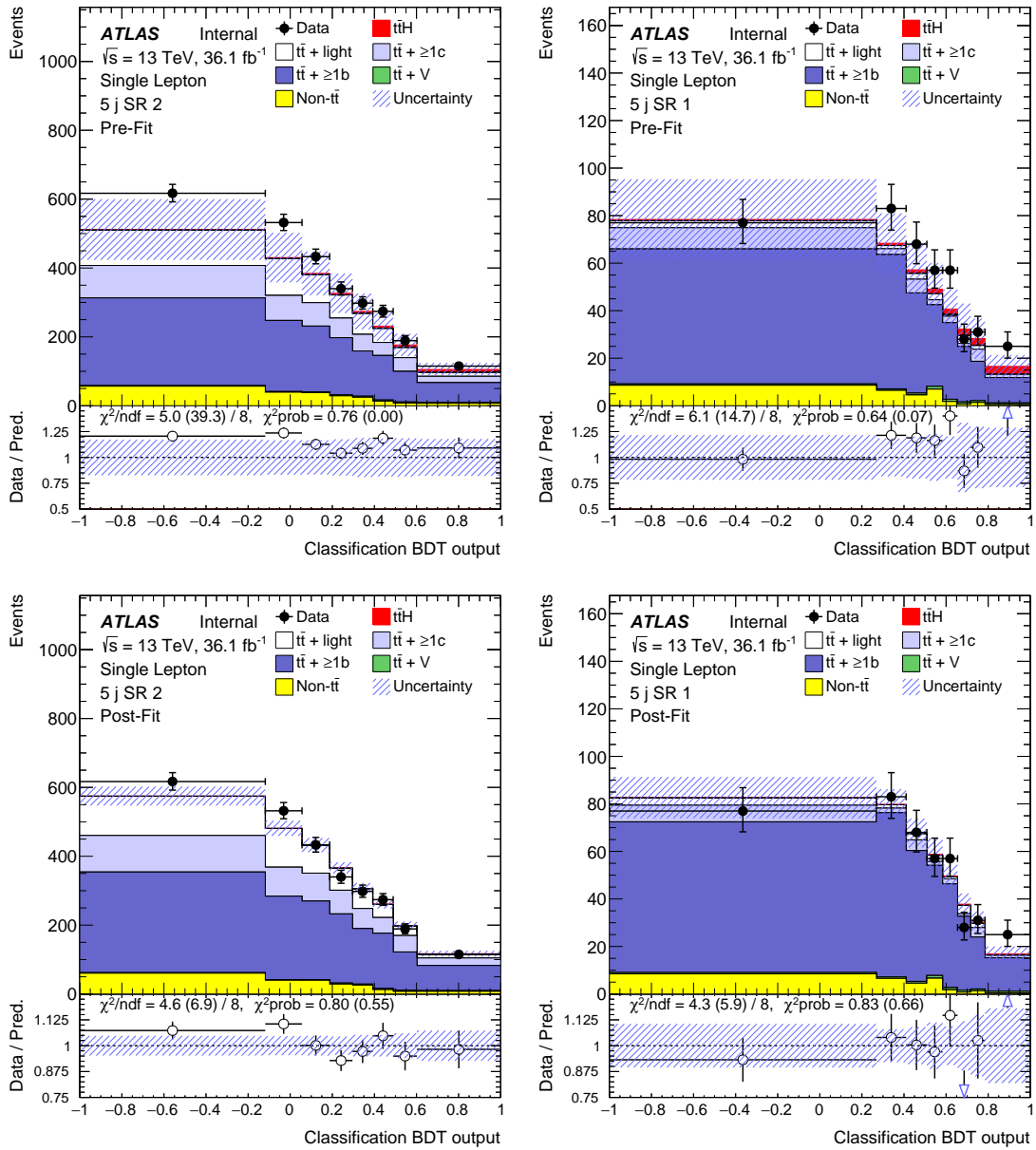
Figure 89: Prefit (top) and postfit (bottom) from fits to data corresponding to a combined BDT-based fit to data with the S+B hypothesis.
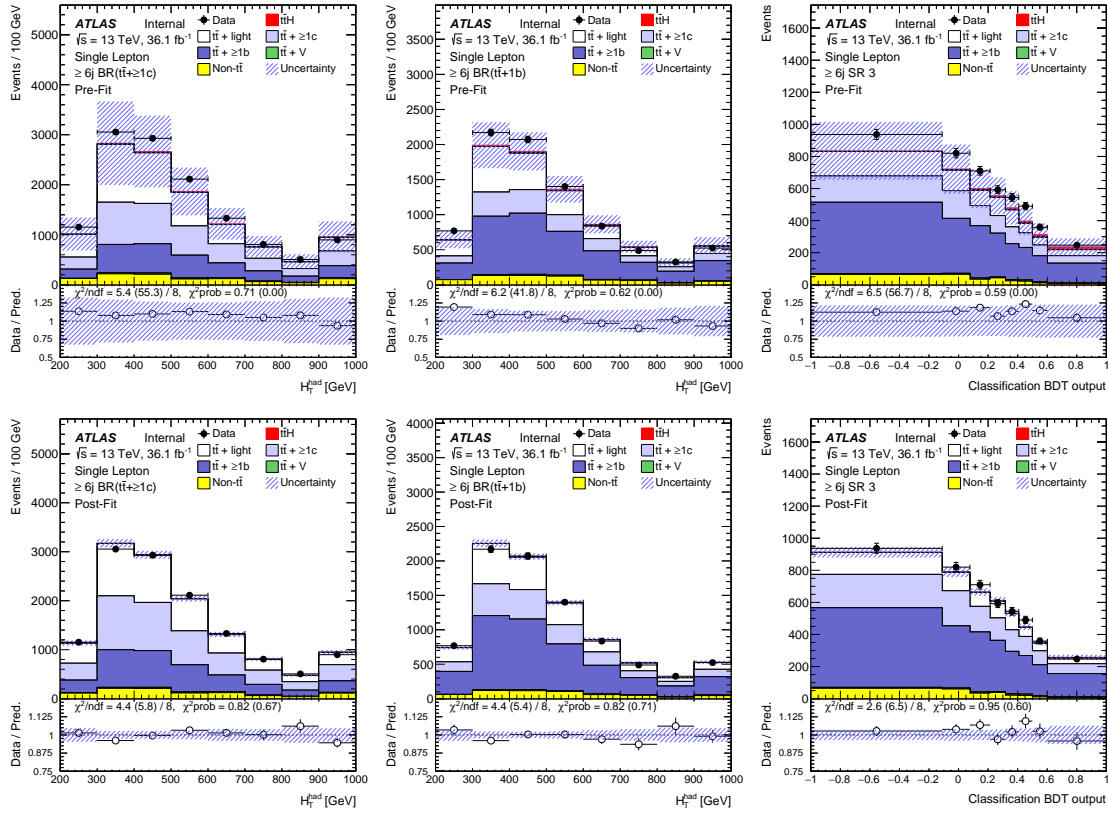
Figure 90: Prefit (top) and postfit (bottom) from fits to data using a combined BDT-based fit to data with the S+B hypothesis.

5.9.3 Combination with the dilepton channel

The results obtained in the single-lepton channel were combined with the dilepton channel. The result for their combination is presented is figure 92.
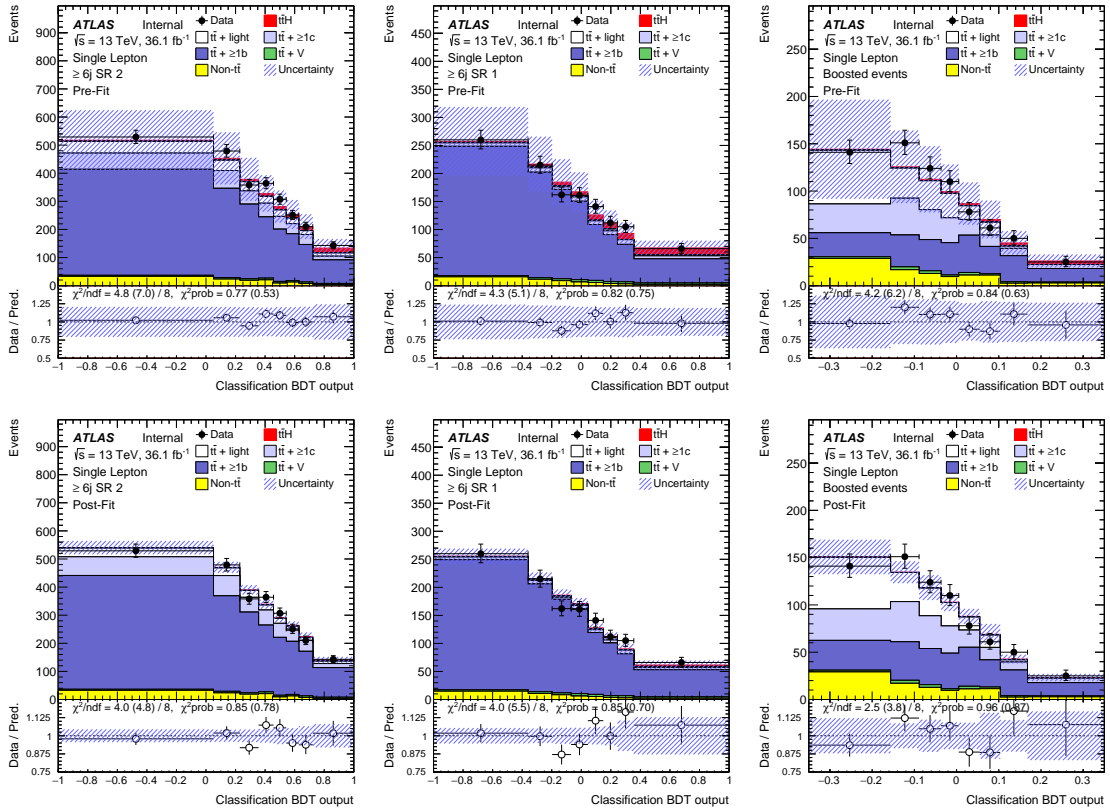
Figure 91: Prefit (top) and postfit (bottom) from fits to data in the most signal-enriched prefit regions using a combined BDT-based fit to data with the S+B hypothesis.
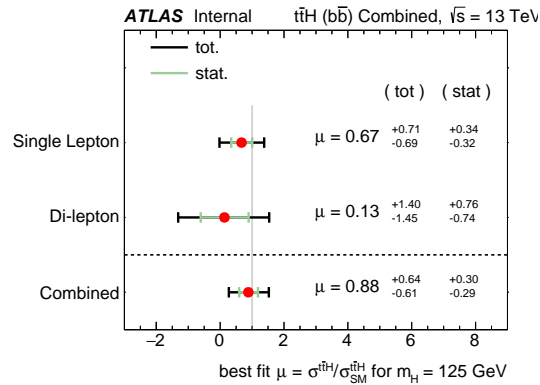
Figure 92: The fitted value of signal strength for single lepton and dilepton channels and their combination

Conclusions

A search for the associated production of the Higgs boson with a pair of top quarks, $t\bar{t}H$ ($H \rightarrow b\bar{b}$), with a single lepton in the final state is presented in this thesis. The search was performed using 36.1 fb^{-1} of pp collision data at a centre-of-mass energy of 13 TeV recorded with the ATLAS detector at the LHC in 2015 and 2016. Measuring the $t\bar{t}H$ cross-section is very important, since it gives direct to the top quark Yukawa coupling, and a significant deviation of this parameter from the SM prediction would indicate a new physics beyond the SM.

The major difficulty of this measurement is that the main background, $t\bar{t}$ with additional b-jets, has the same signature as the signal. To separate the signal from the background several analysis techniques are employed. The main contribution by the author to this analysis, presented in this dissertation, is the development of the likelihood discriminant (LHD) method, that exploits specific kinematic properties of $t\bar{t}H$ ($H \rightarrow b\bar{b}$) and $t\bar{t}$+jets events to distinguish them. The method is used in combination with other discriminating variables via multivariate techniques, in order to achieve improved discrimination between the signal and the background. The LHD was for the first time applied in this analysis and was found to be the most discriminating single variable, providing an improvement in the separation power $\sim 10\%$.

The ratio of the measured $t\bar{t}H$ cross-section to the SM expectation is found to be $\mu = 0.67^{+0.71}_{-0.69}$, assuming a Higgs boson mass of 125 GeV. The result obtained after combining with the dilepton channel is $\mu = 0.88^{+0.64}_{-0.61}$. This result is consistent with both the background-only hypothesis and the $t\bar{t}H$ SM prediction.

The identification of the jets originating from b-quark fragmentation, or b-tagging, plays a key role in this search. Work on the optimisation of the b-tagging impact-parameter-based algorithms (IP2D, IP3D) for LHC Run 2 is also presented in this dissertation. The major contribution made by the author is the development of a new classification of tracks that considers several new tracking variables, in particular, taking advantage of the installation of the IBL - a new pixel detector layer. The expected improvement in light jet rejection at 70% b-jet efficiency of the IP3D algorithm due to the new track categorisation is $\sim 15\%$, while the overall improvement of the optimisations described is $\sim 27\%$.

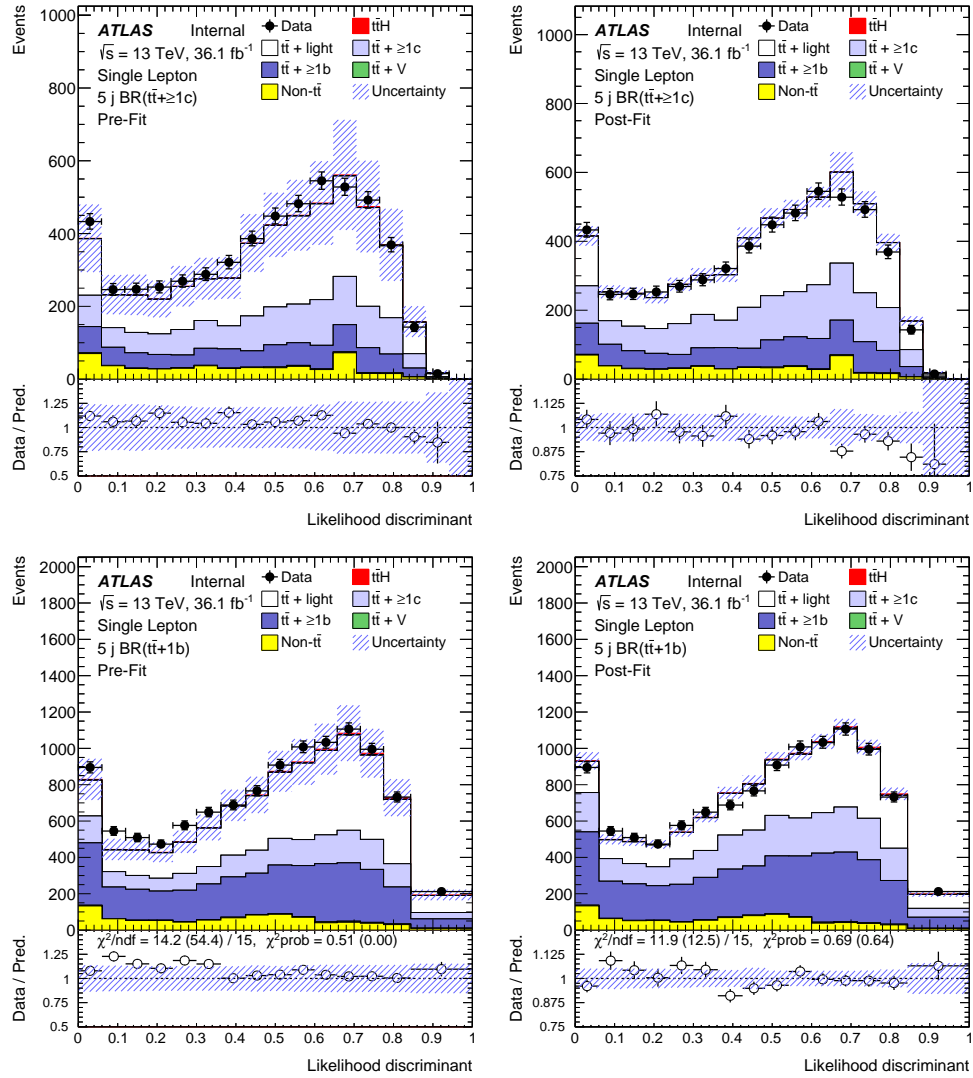# A   Auxiliary materials: BDT input postfit distributions

Figure 93: Pre-fit and post-fit distributions of the final likelihood discriminant for 5 jets backround-enriched regions
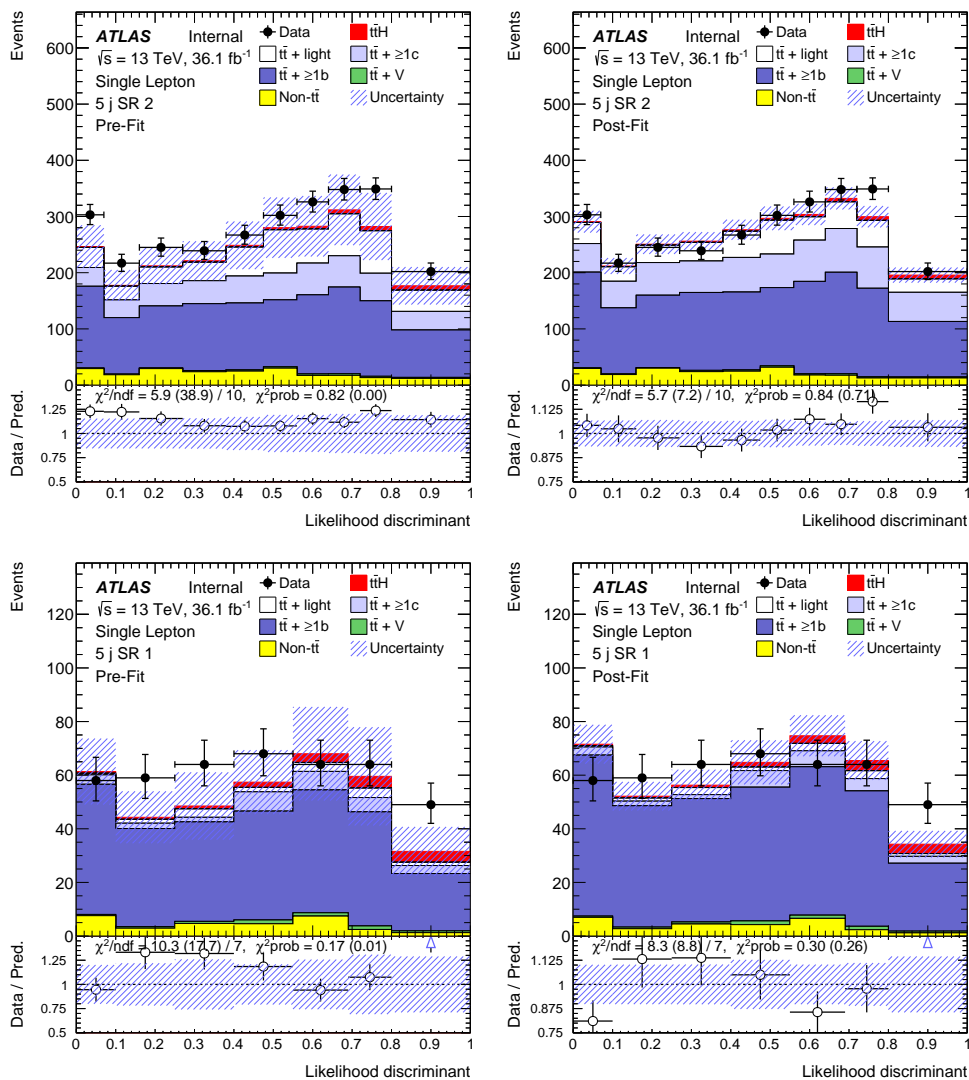
Figure 94: Pre-fit and post-fit distributions of the final likelihood discriminant for 5 jets signal-enriched regions
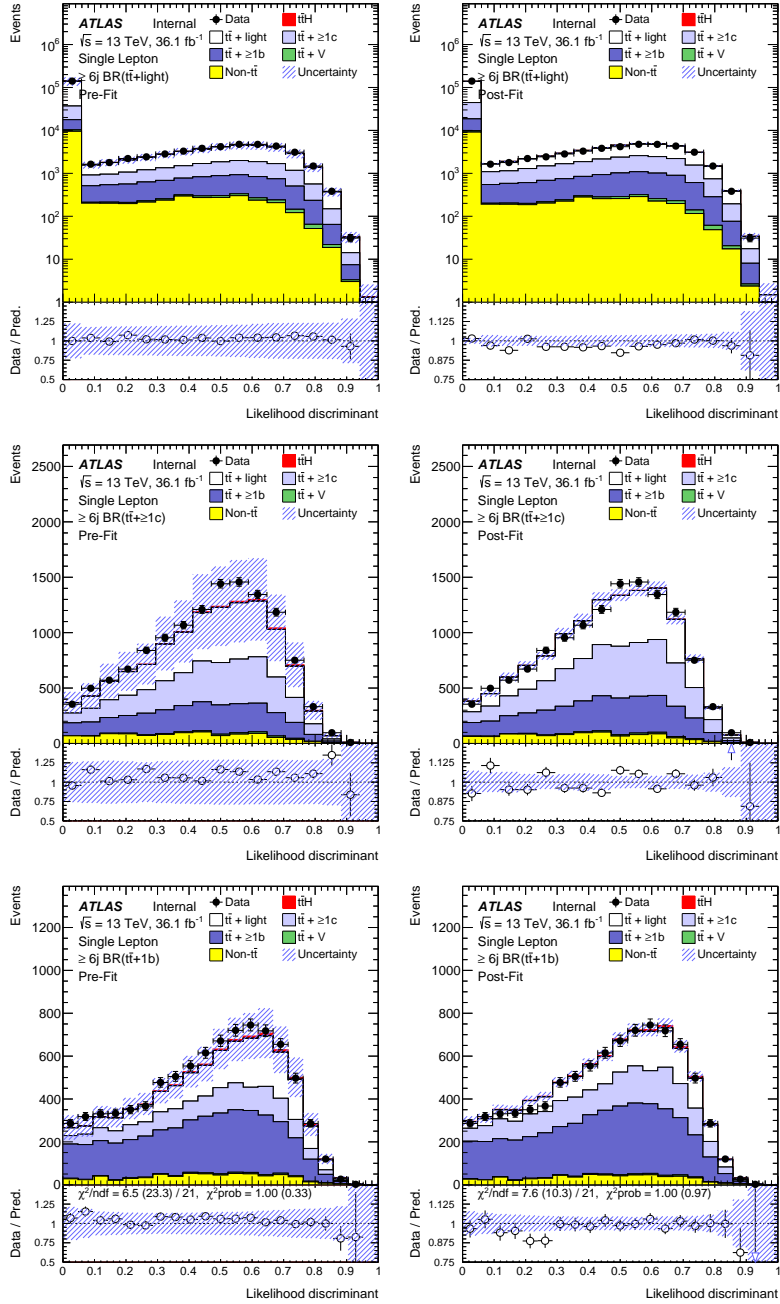
Figure 95: Pre-fit and post-fit distributions of the final likelihood discriminant for ± 6 jets background-enriched regions
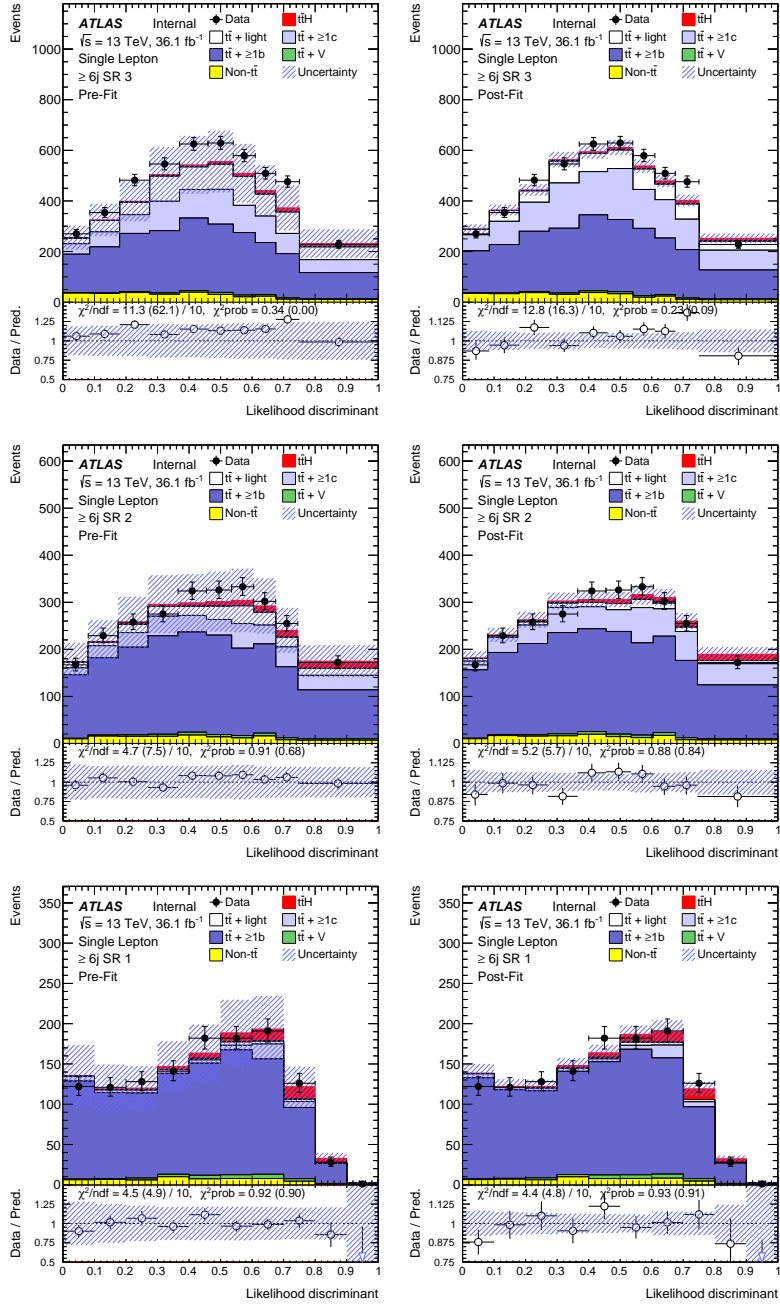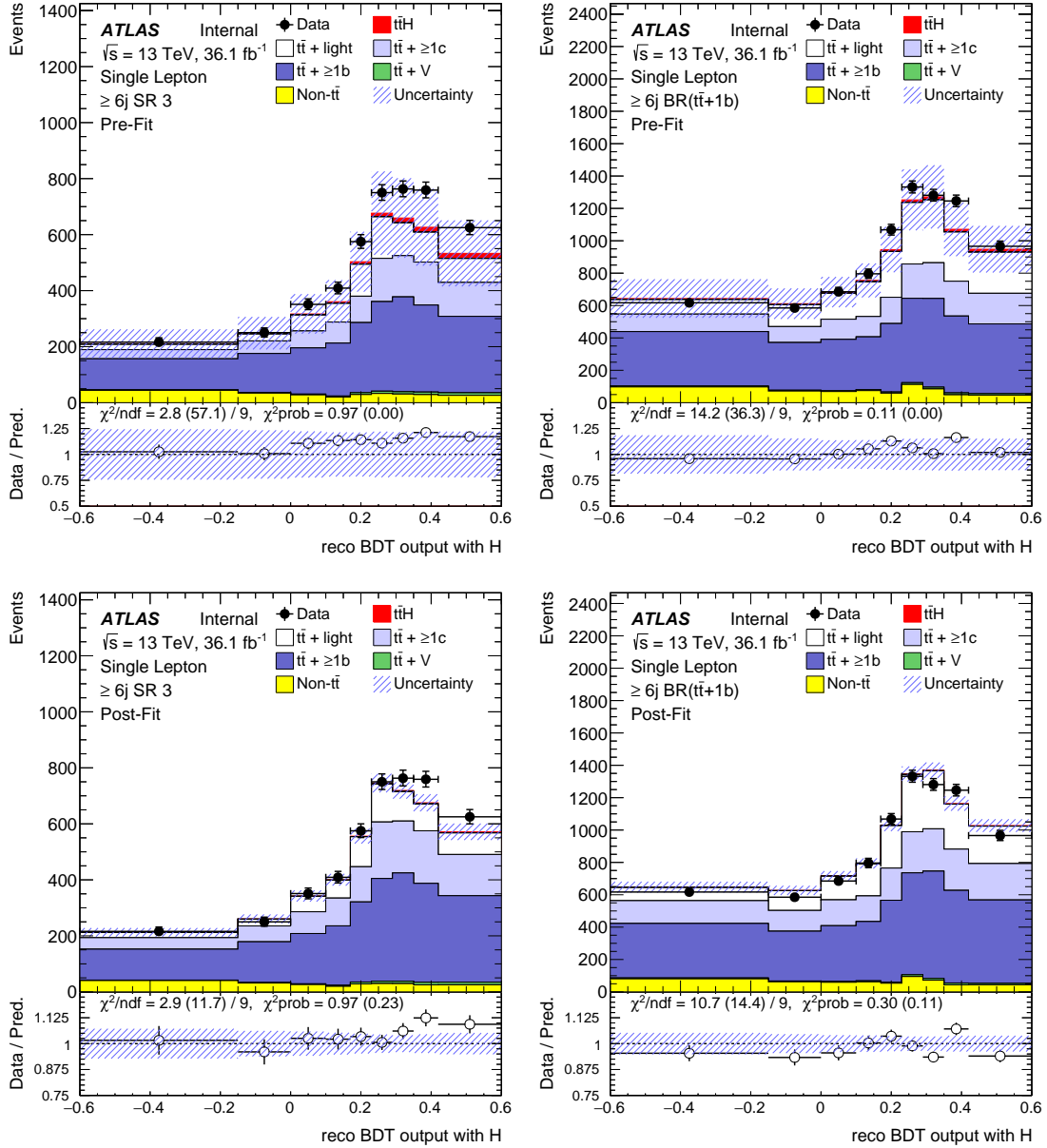
Figure 96: Pre-fit and post-fit distributions of the final likelihood discriminant for $\pm$ 6 jets signal-enriched regions

Figure 97: Pre-fit (top row) and post-fit (bottom row) distributions of the highest reco BDT output for ≥ 6 jets signal-enriched regions. Distributions correspond to the best permutation of the BDT trained using the Higgs related variables
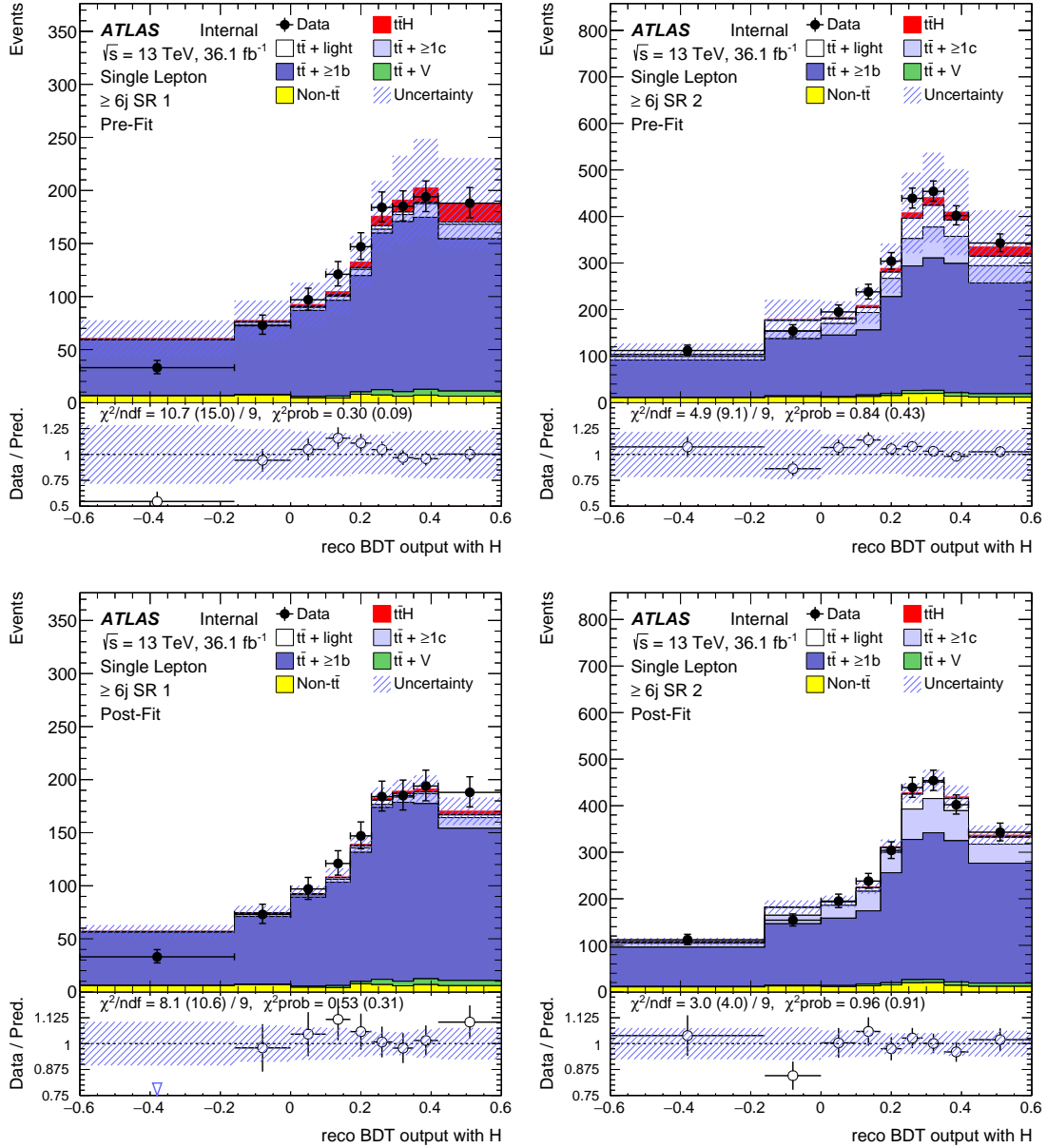
Figure 98: Pre-fit (top row) and post-fit (bottom row) distributions of the highest reco BDT output for ≥ 6 jets nsignal-enriched regions. Distributions correspond to the best permutation of the BDT trained using the Higgs related variables.

Figure 99: Pre-fit (top row) and post-fit (bottom row) distributions of the highest reco BDT output for 5 jets signal-enriched regions. Distributions correspond to the best permutation of the BDT trained using the Higgs related variables. From left to right: $t\bar{t}H$, $t\bar{t}+ \geq 2b$ Hi, and $t\bar{t} + 1b$ Lo. Bins with $\geq 5\%$ signal are blinded. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots.
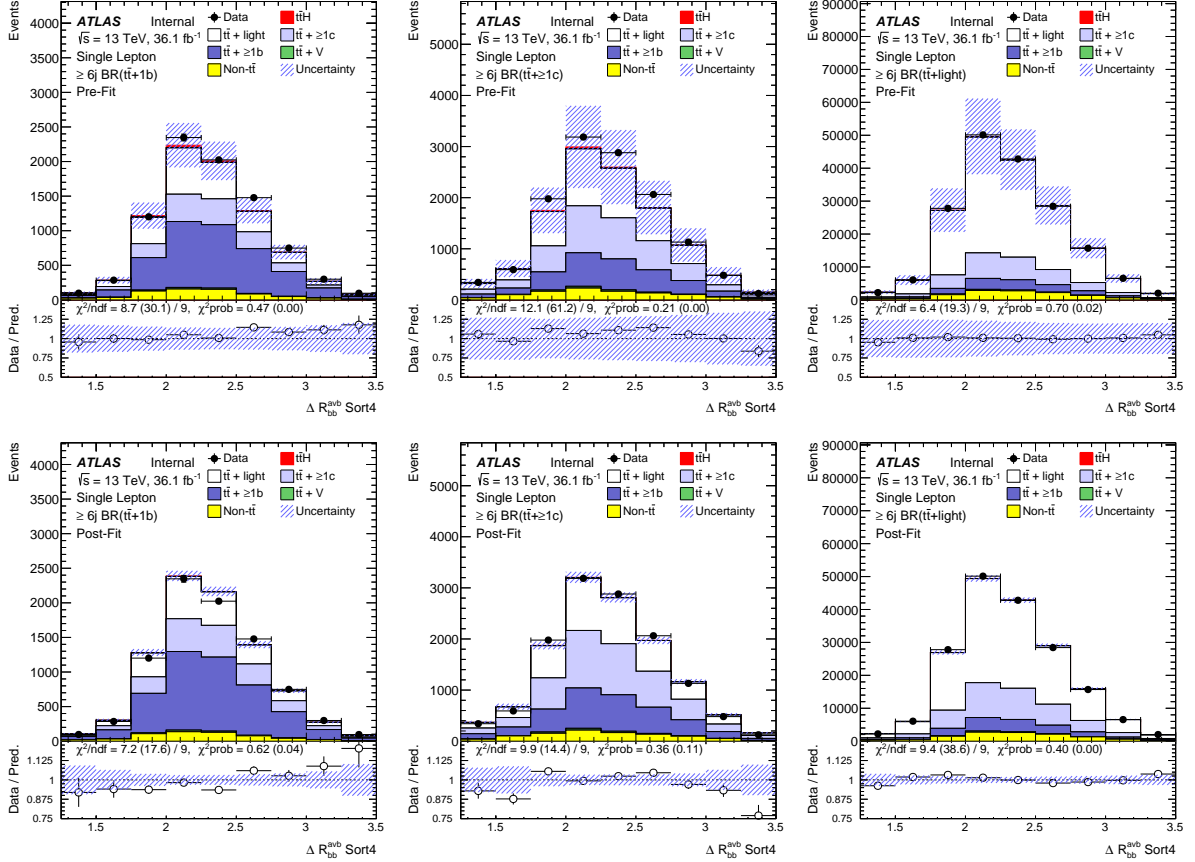
Figure 100: Pre-fit (top row) and post-fit (bottom row) distributions of $\Delta R_{bb,sort4}^{\mathrm{avg}}$ for ≥ 6 jets signal-enriched regions. From left to right: $t\bar{t}H$, $t\bar{t}+ \geq 2b$ Hi, $t\bar{t}+ \geq 2b$ Lo. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots. Post-fit plots are made after unblinded S+B MVA fits.
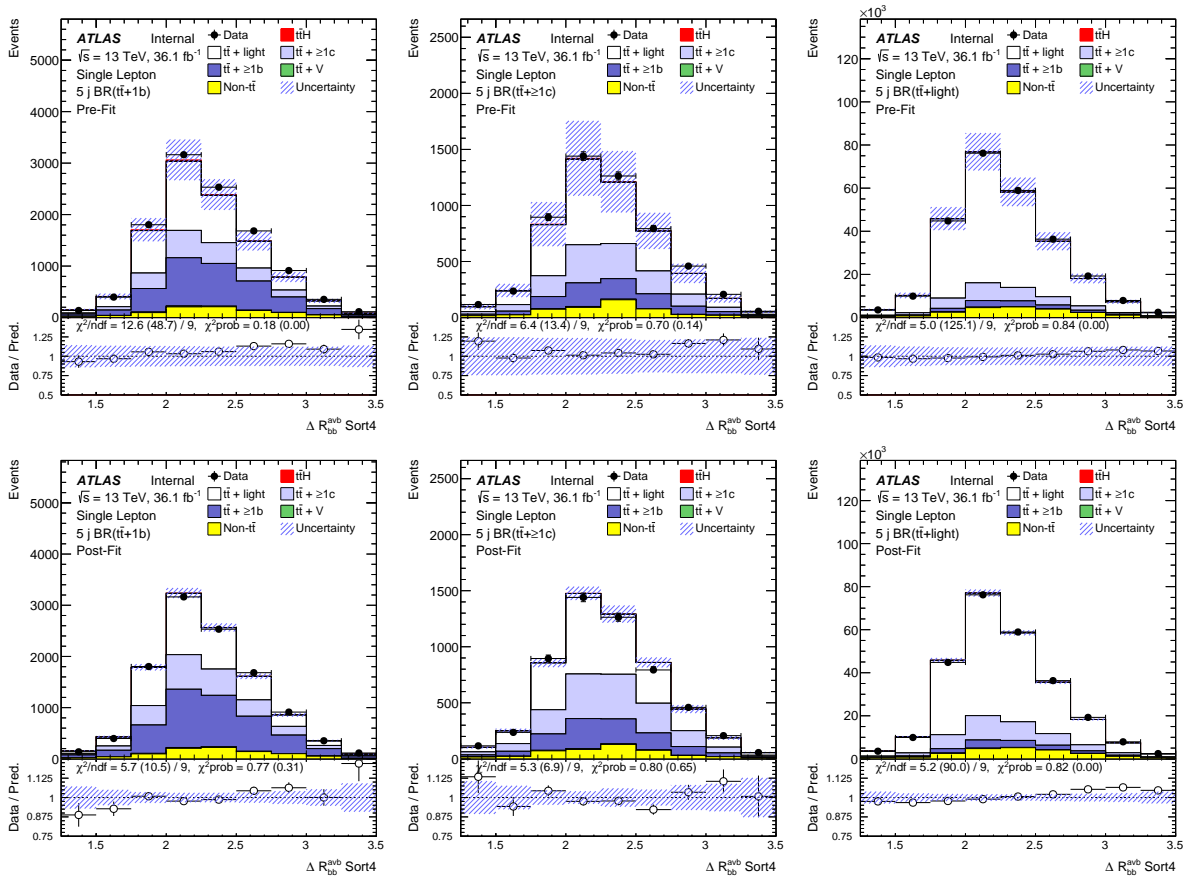
136
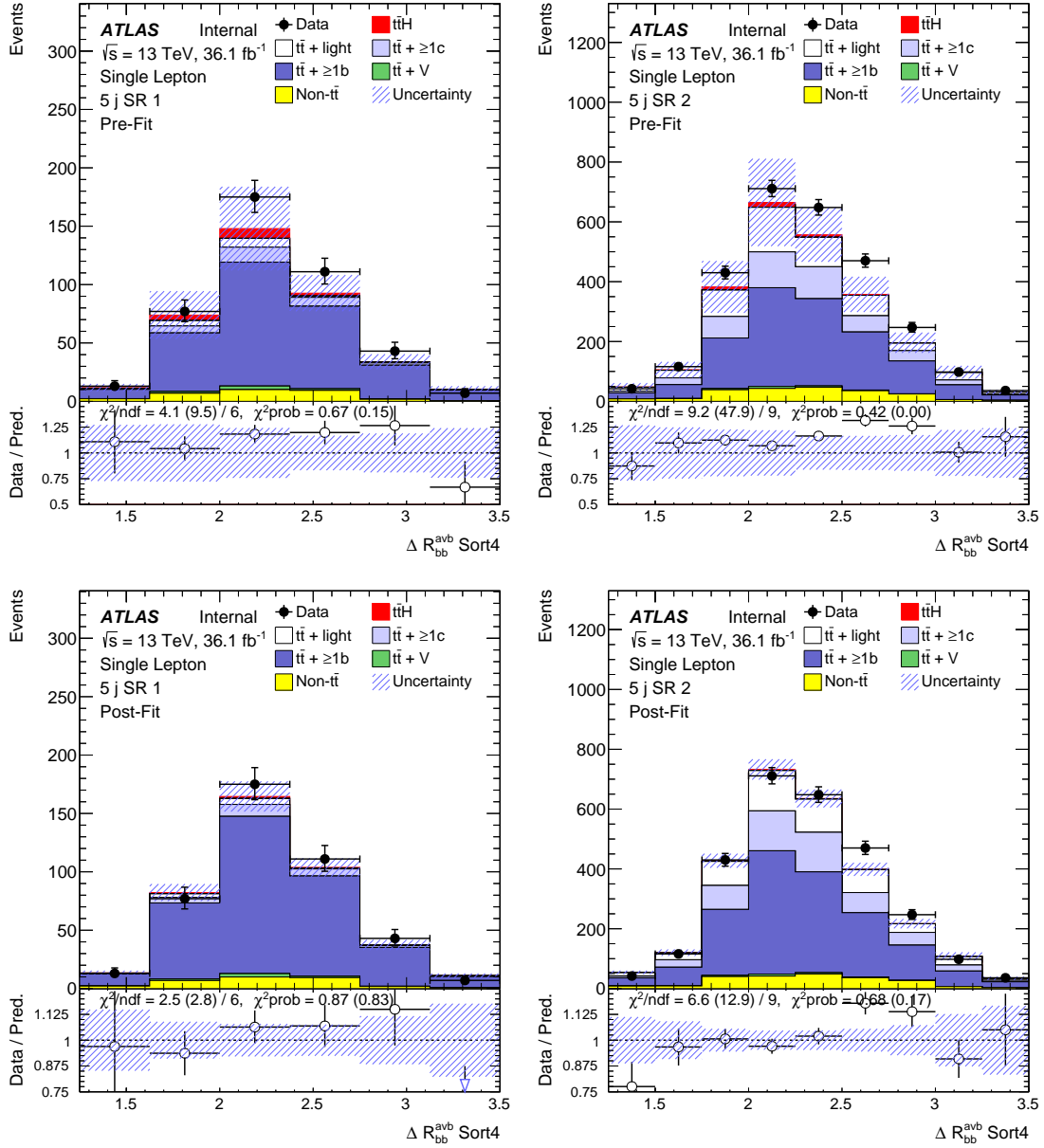
Figure 101: Pre-fit (top row) and post-fit (bottom row) distributions of $\Delta R_{bb,sort4}^{\mathrm{avg}}$ for ≥ 6 jets signal-enriched regions. $t\bar{t}H$ (left), and $t\bar{t}+ \geq 2b$ (right). Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots.

Figure 102: Pre-fit (top row) and post-fit (bottom row) distributions of $\Delta R_{bb,sort4}^{\mathrm{avg}}$ for ≥ 6 jets signal-enriched regions. From left to right: $t\bar{t}1b$, $t\bar{t}+ \geq 1c$, $t\bar{t}+$ light. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots. Post-fit plots are made after unblinded S+B MVA fits.
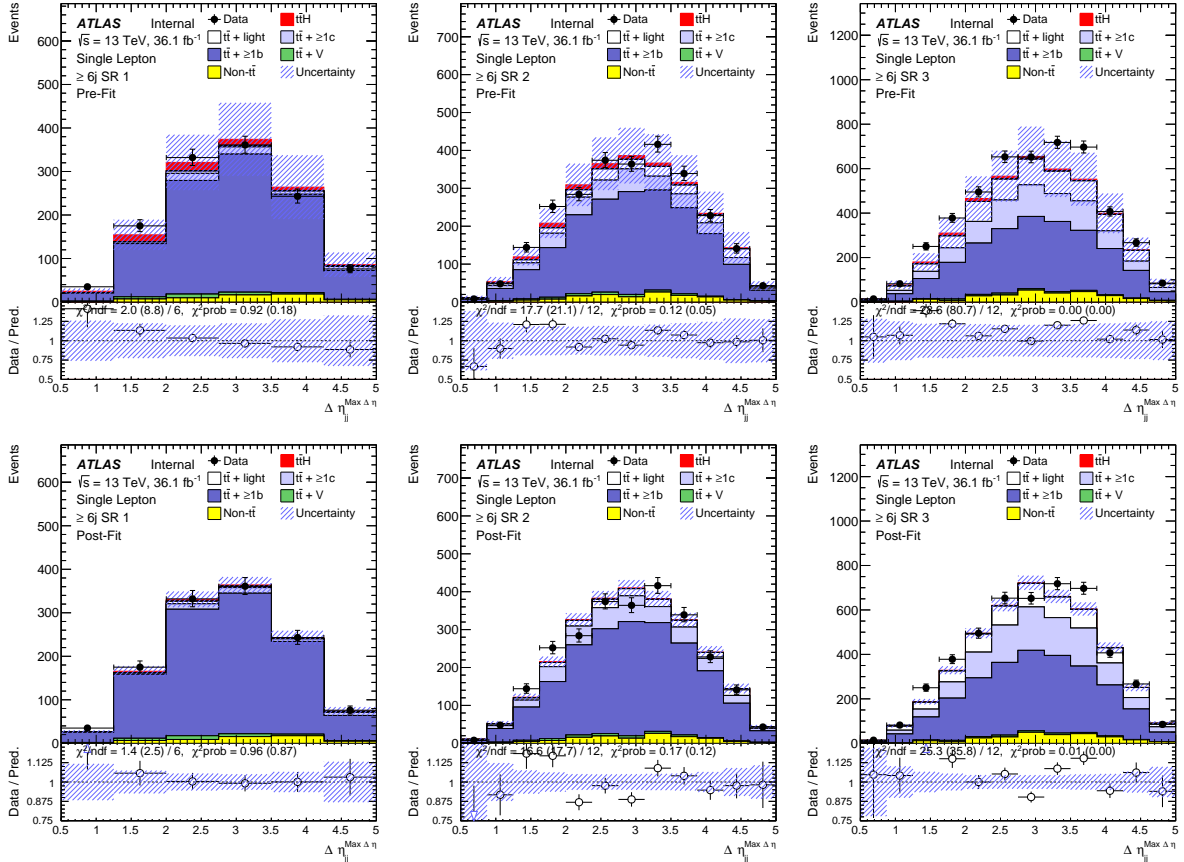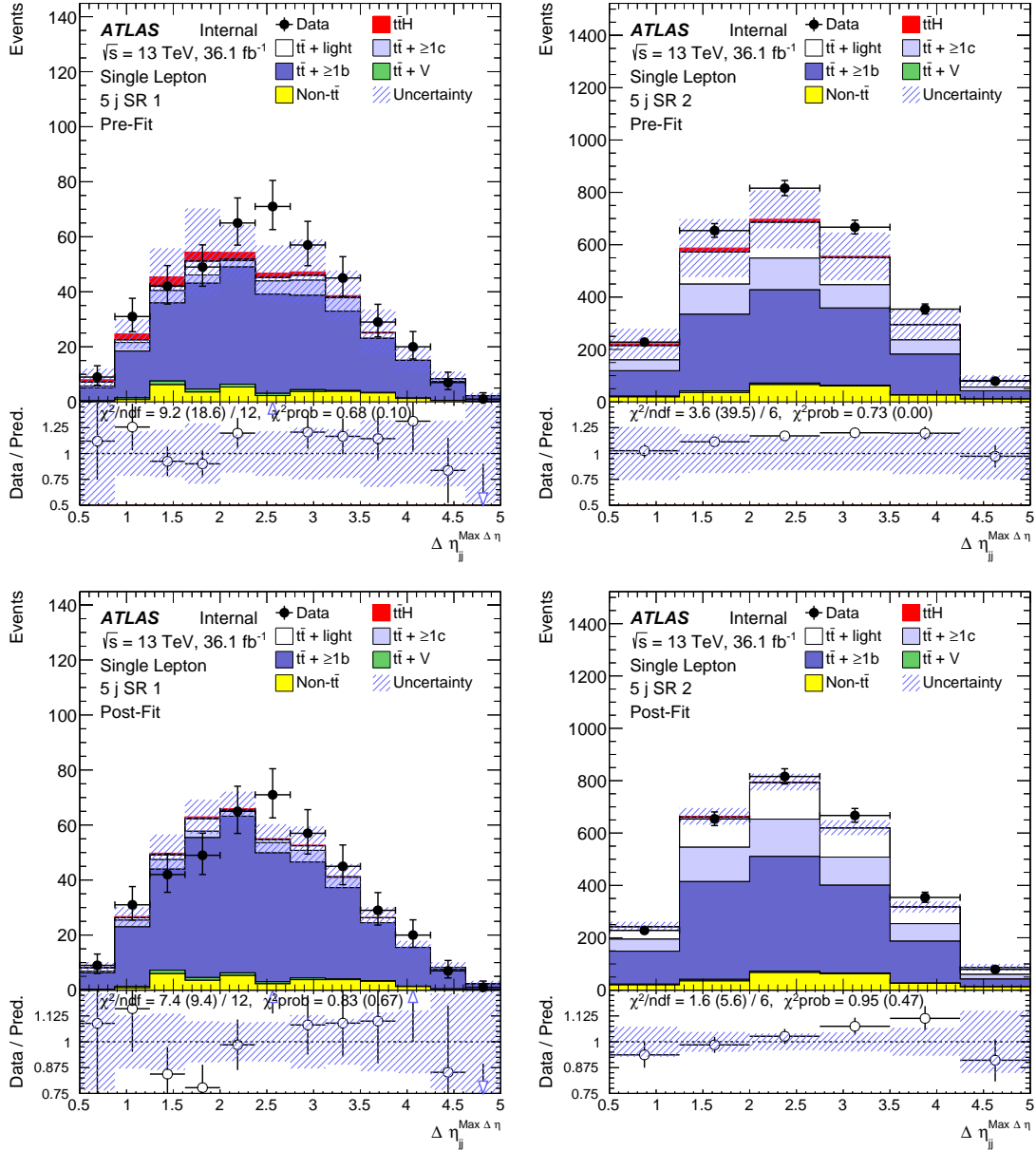
Figure 103: Pre-fit (top row) and post-fit (bottom row) distributions of $\Delta R_{bb,sort4}^{\mathrm{avg}}$ for 5 jets signal-enriched regions. From left to right: $t\bar{t}1b$, $t\bar{t}+ \geq 1c$, $t\bar{t}+$ light. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots.

Figure 104: Pre-fit (top row) and post-fit (bottom row) distributions of $\Delta\eta_{jj}^{\max \Delta\eta}$ for ≥ 6 jets signal-enriched regions. From left to right: $t\bar{t}H$, $t\bar{t}+ \geq 2b$ Hi, $t\bar{t}+ \geq 2b$ Lo. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots. Post-fit plots are made after unblinded S+B MVA fits.

Figure 105: Pre-fit (top row) and post-fit (bottom row) distributions of $\Delta\eta_{jj}^{\max\ \Delta\eta}$ for ≥ 6 jets signal-enriched regions. $t\bar{t}H$ (left), and $t\bar{t}+\geq 2b$ (right). Bins with $\geq 5\%$ signal are blinded. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots.
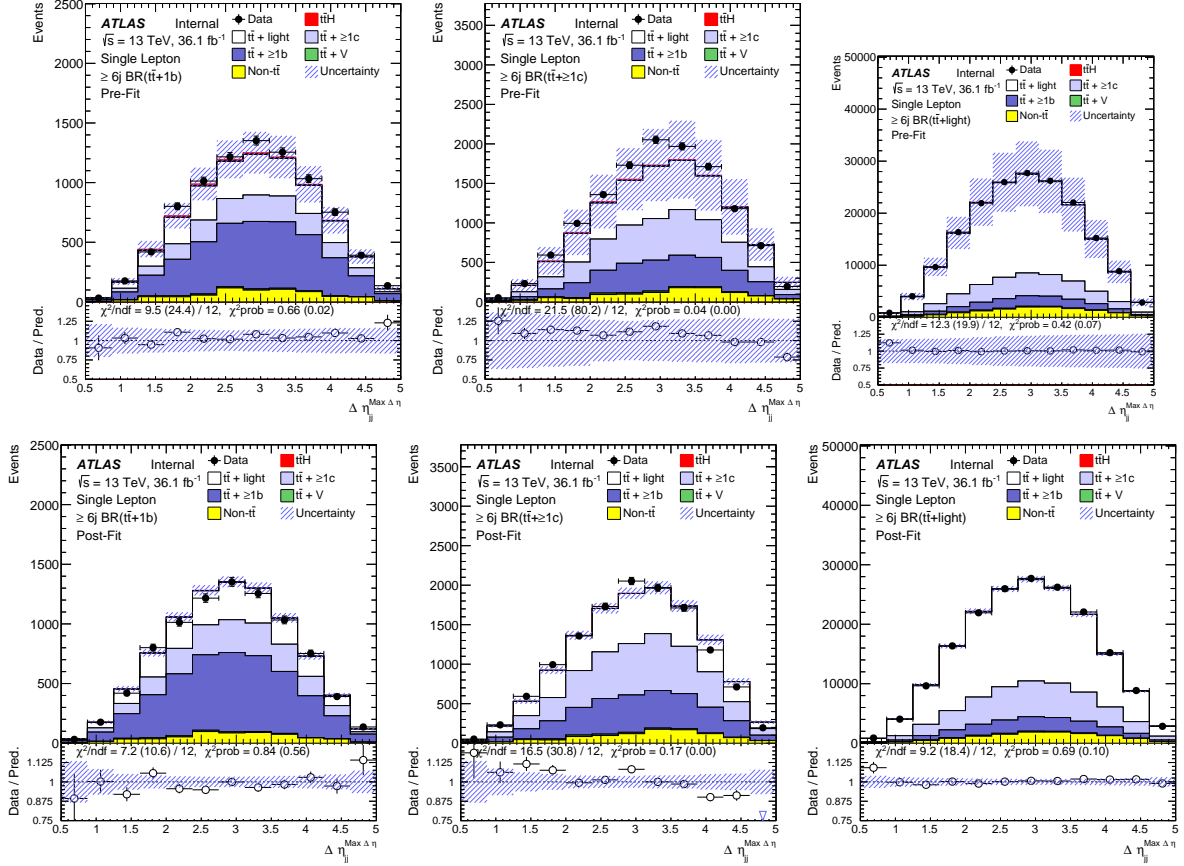
Figure 106: Pre-fit (top row) and post-fit (bottom row) distributions of $\Delta\eta_{jj}^{\max \, \Delta\eta}$ for ≥ 6 jets signal-enriched regions. From left to right: $t\bar{t}1b$, $t\bar{t}+ \geq 1c$, $t\bar{t}+$ light. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots. Post-fit plots are made after unblinded S+B MVA fits.
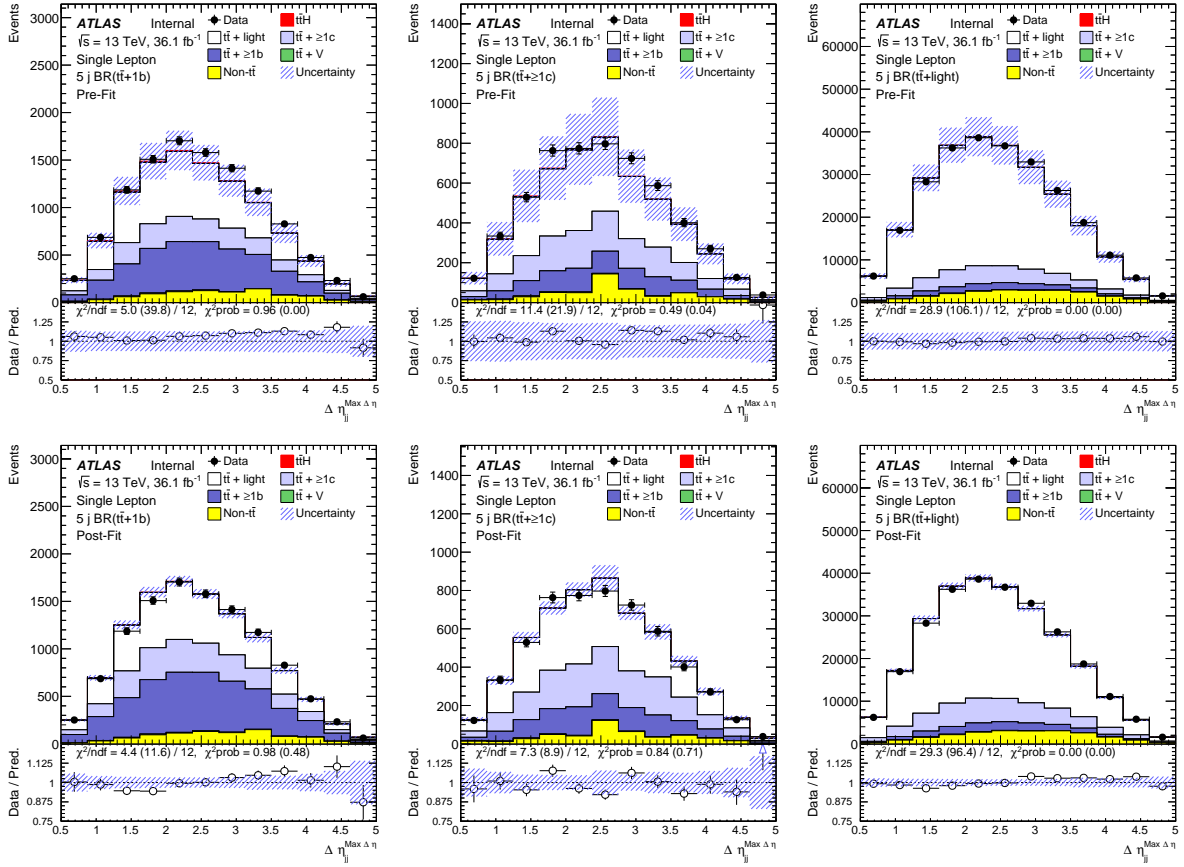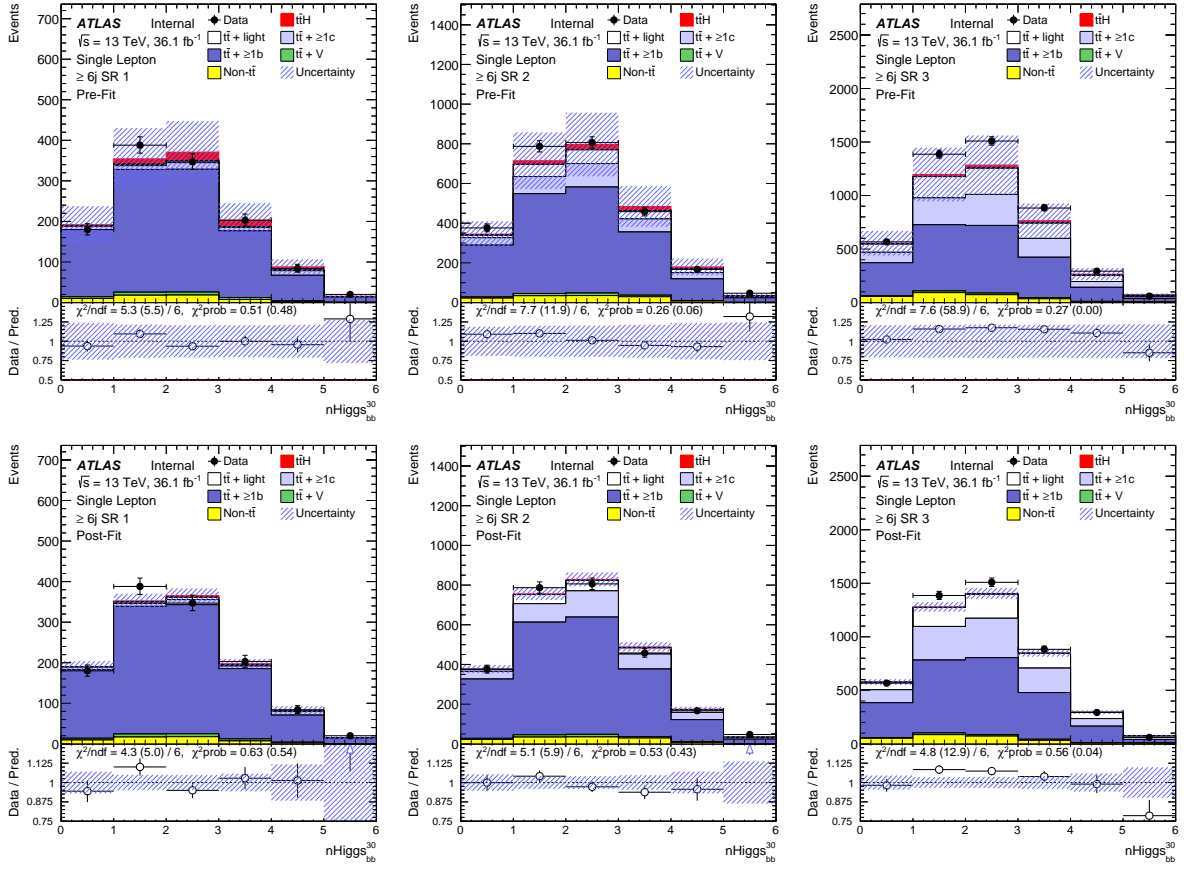
Figure 107: Pre-fit (top row) and post-fit (bottom row) distributions of $\Delta\eta_{jj}^{\max \Delta\eta}$ for 5 jets signal-enriched regions. From left to right: $t\bar{t}1b$, $t\bar{t}+ \geq 1c$, $t\bar{t}+$ light. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots.

Figure 108: Pre-fit (top row) and post-fit (bottom row) distributions of $N_{30}^{\mathrm{Higgs,bbsort4}}$ for ≥ 6 jets signal-enriched regions. From left to right: $t\bar{t}H$, $t\bar{t}+ \geq 2b$ Hi, $t\bar{t}+ \geq 2b$ Lo. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots. Post-fit plots are made after unblinded S+B MVA fits.
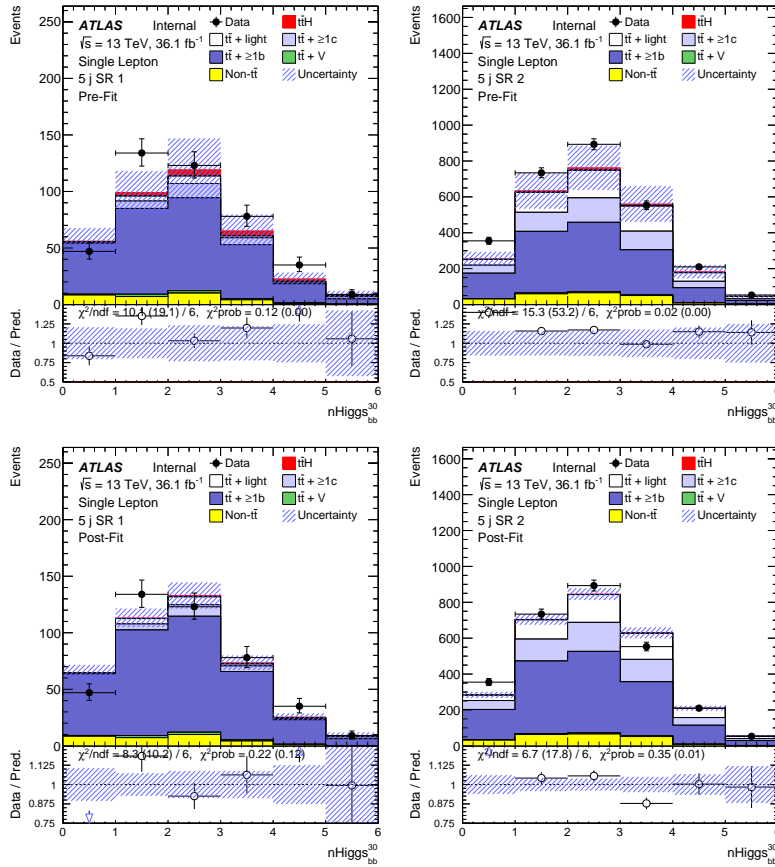
Figure 109: Pre-fit (top row) and post-fit (bottom row) distributions of $N_{30}^{\text{Higgs,bbsort4}}$ for ≥ 6 jets signal-enriched regions. $t\bar{t}H$ (left), and $t\bar{t}+ \geq 2b$ (right). Bins with $\geq 5\%$ signal are blinded. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots.
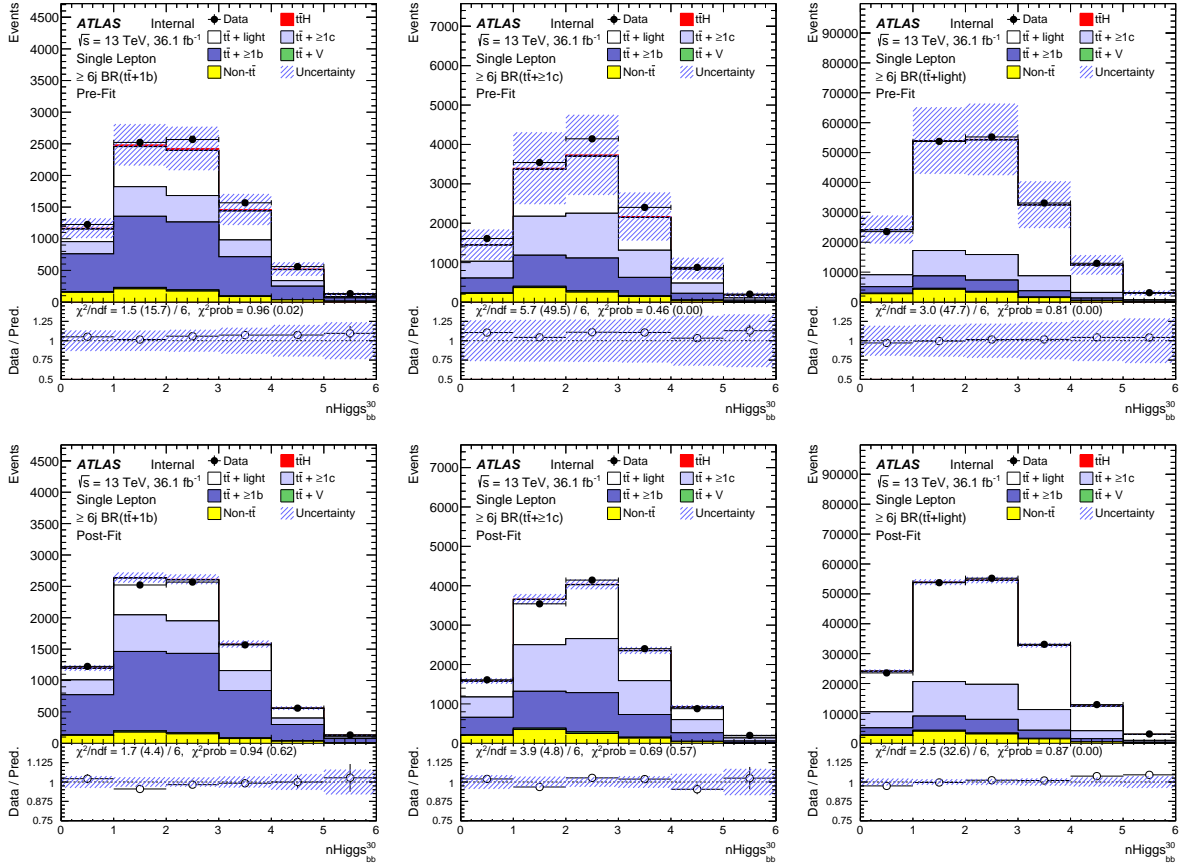
Figure 110: Pre-fit (top row) and post-fit (bottom row) distributions of $N_{30}^{\mathrm{Higgs,bbsort4}}$ for ≥ 6 jets signal-enriched regions. From left to right: $t\bar{t}1b$, $t\bar{t}+ \geq 1c$, $t\bar{t}+$ light. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots. Post-fit plots are made after unblinded S+B MVA fits.
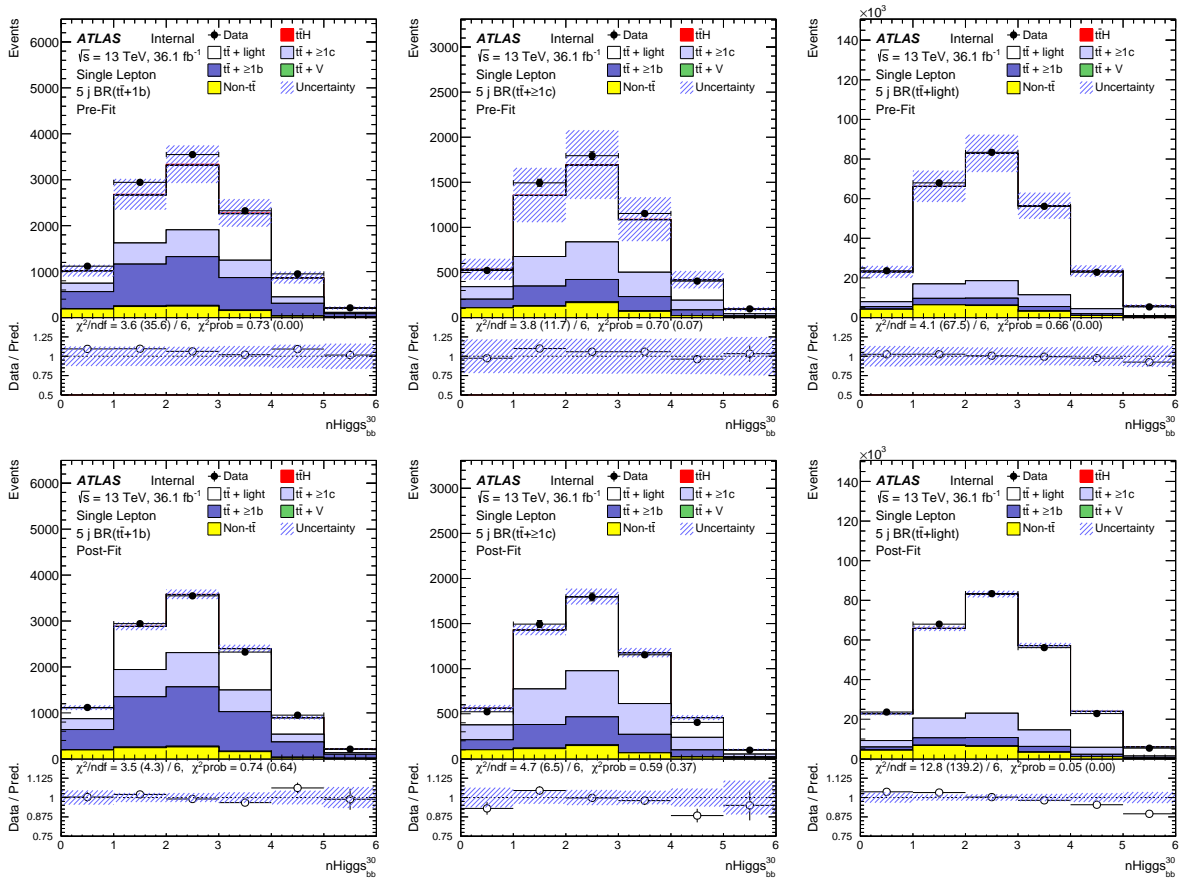
146

Figure 111: Pre-fit (top row) and post-fit (bottom row) distributions of $N_{30}^{\text{Higgs,bbsort4}}$ for 5 jets signal-enriched regions. From left to right: $t\bar{t}1b$, $t\bar{t}+ \geq 1c$, $t\bar{t}+$ light. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots.
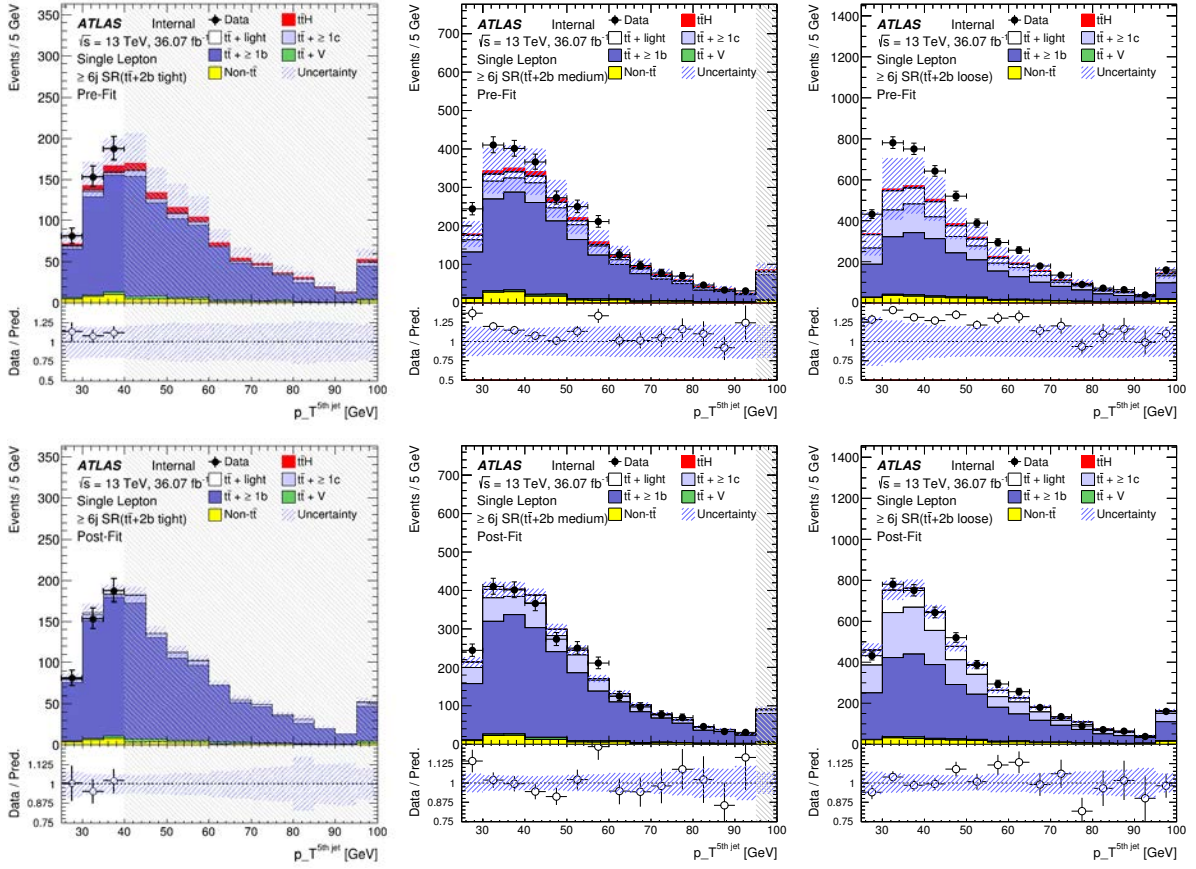
Figure 112: Pre-fit (top row) and post-fit (bottom row) distributions of p_T^{jet5} for ≥ 6 jets signal-enriched regions. From left to right: $t\bar{t}H$, $t\bar{t}+ \geq 2b$ Hi, $t\bar{t}+ \geq 2b$ Lo. Bins with $\geq 5\%$ signal are blinded. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots. Post-fit plots are made after B-only blinded MVA fits.

Figure 113: Pre-fit (top row) and post-fit (bottom row) distributions of p_T^{jet5} for ≥ 6 jets signal-enriched regions. $t\bar{t}H$ (left), and $t\bar{t}+ \geq 2b$ (right). Bins with $\geq 5\%$ signal are blinded. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots.

Figure 114: Pre-fit (top row) and post-fit (bottom row) distributions of p_T^{jet5} for ≥ 6 jets signal-enriched regions. From left to right: $t\bar{t}1b$, $t\bar{t}+ \geq 1c$, $t\bar{t}+$ light. Bins with $\geq 5\%$ signal are blinded. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots. Post-fit plots are made after B-only blinded MVA fits.
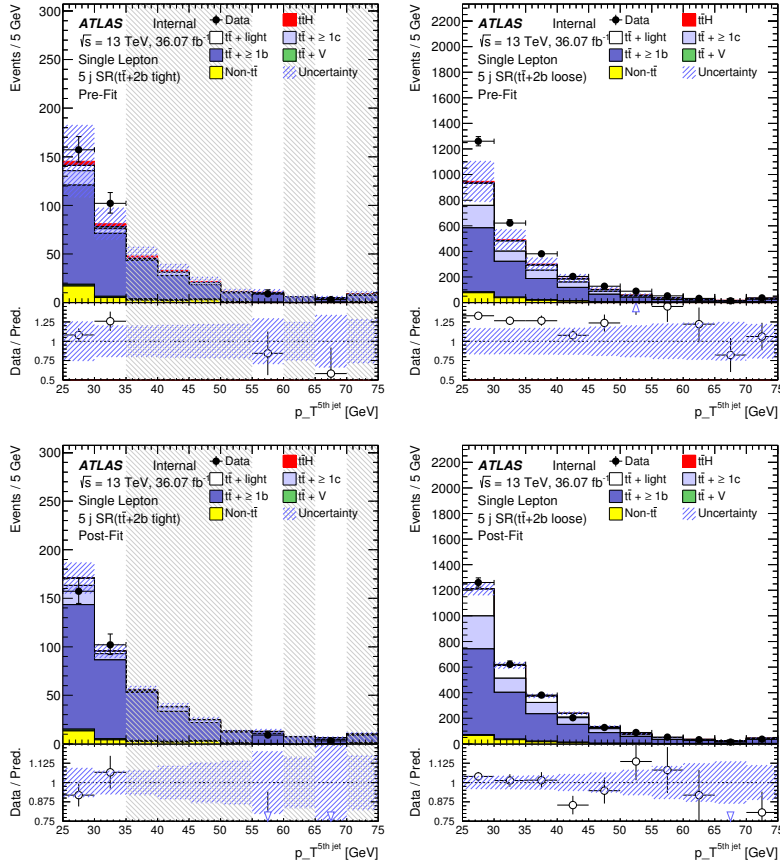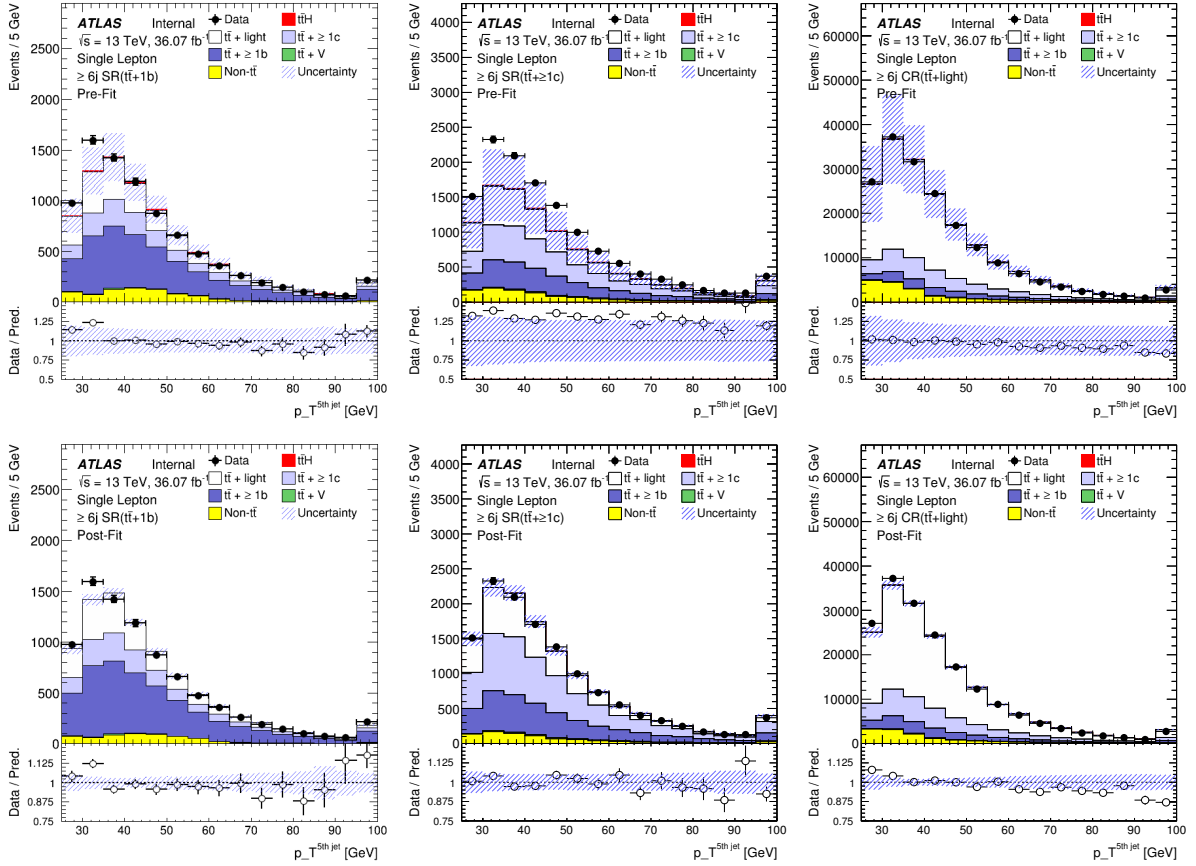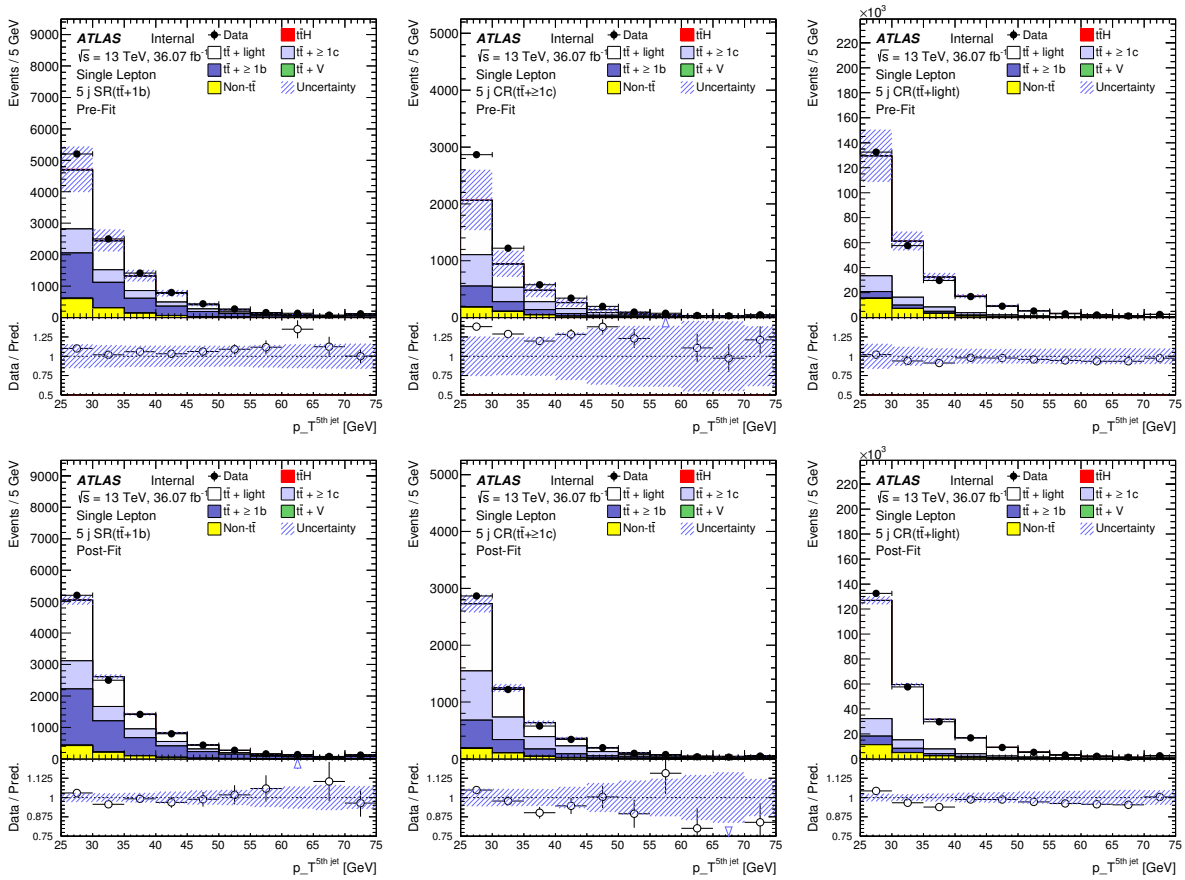
Figure 115: Pre-fit (top row) and post-fit (bottom row) distributions of p_T^{jet5} for 5 jets signal-enriched regions. From left to right: $t\bar{t}1b$, $t\bar{t}+ \geq 1c$, $t\bar{t}+$ light. Bins with $\geq 5\%$ signal are blinded. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots.
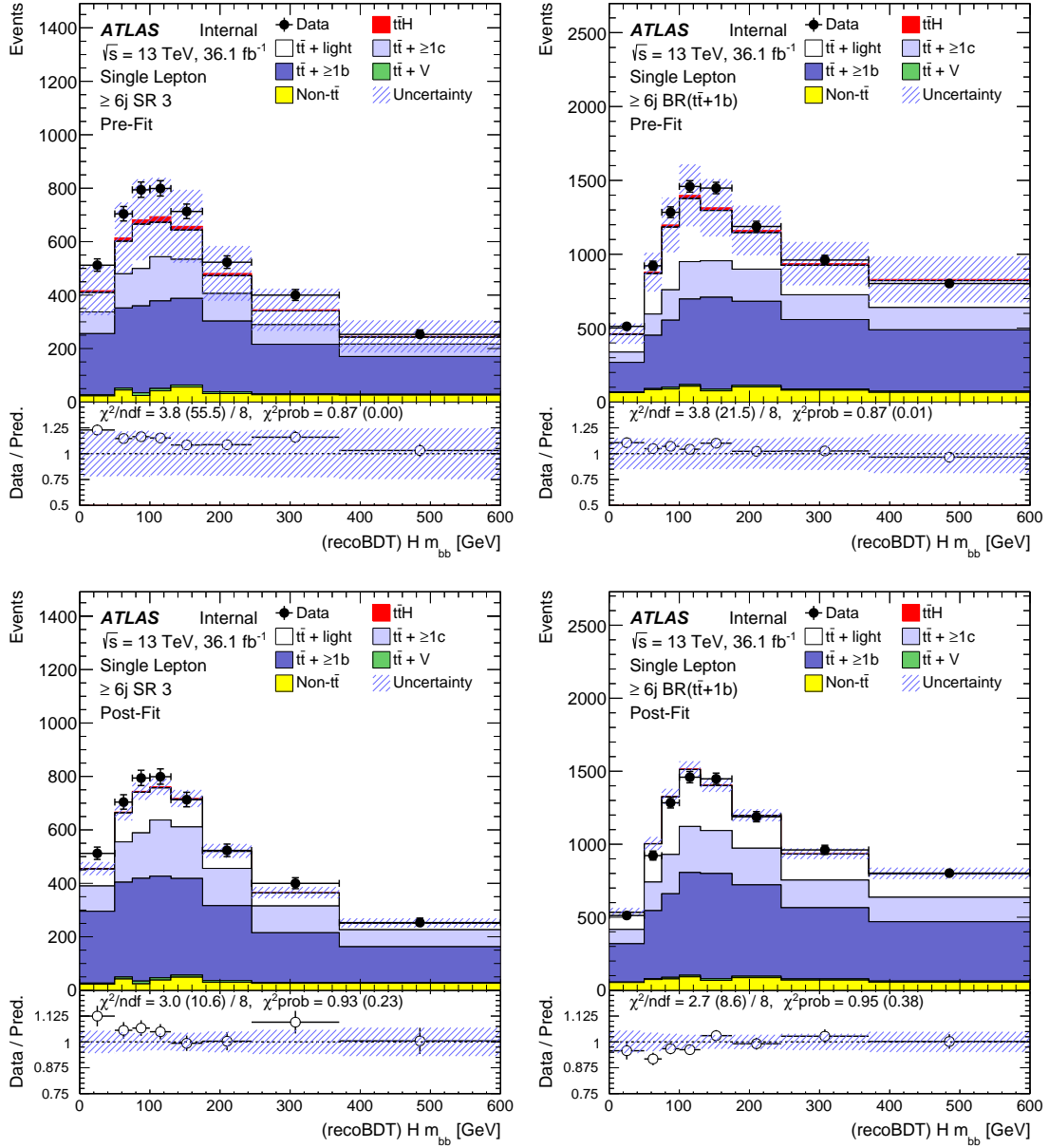
Figure 116: Pre-fit (top row) and post-fit (bottom row) distributions of the $m_{H,b_{\mathrm{lep\ top}}}$ distribution for ≥ 6 jets signal-enriched regions. Distributions correspond to the best reco BDT permutation.
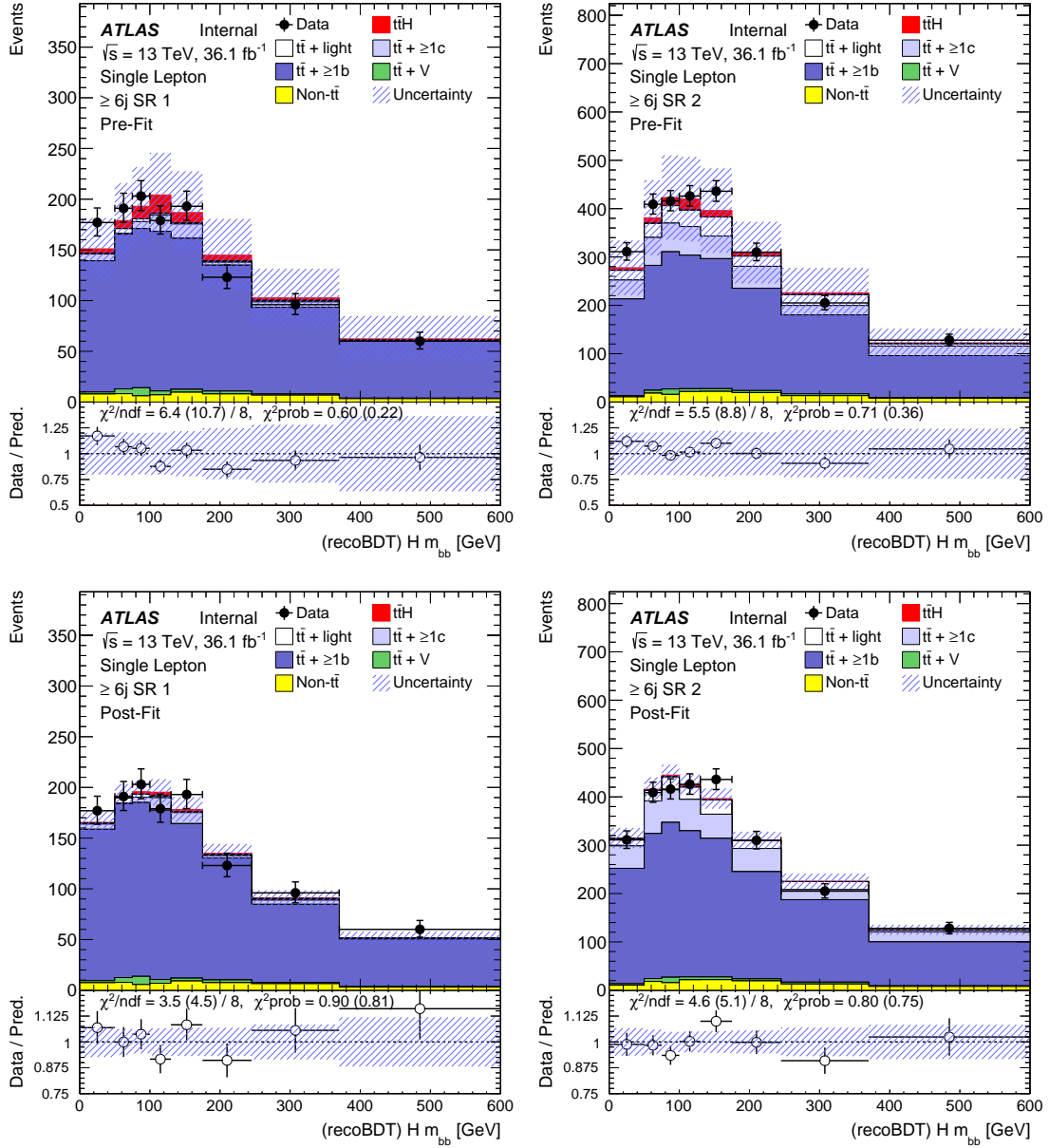
Figure 117: Pre-fit (top row) and post-fit (bottom row) distributions of the $m_{H,b_{\text{lep top}}}$ distribution for ≥ 6 jets signal-enriched regions. Distributions correspond to the best reco BDT permutation.

Figure 118: Pre-fit (top row) and post-fit (bottom row) distributions of the $m_{H,b_{\text{lep top}}}$ distribution for 5 jets signal-enriched regions. Distributions correspond to the best reco BDT permutation. From left to right: $t\bar{t}H$, $t\bar{t}+ \geq 2b$ Hi, and $t\bar{t} + 1b$ Lo. Bins with $\geq 5\%$ signal are blinded. Note that the scale of the y axis of the ratio plot is different for pre-fit and post-fit plots. Post-fit plots are made after unblinded S+B MVA fits.
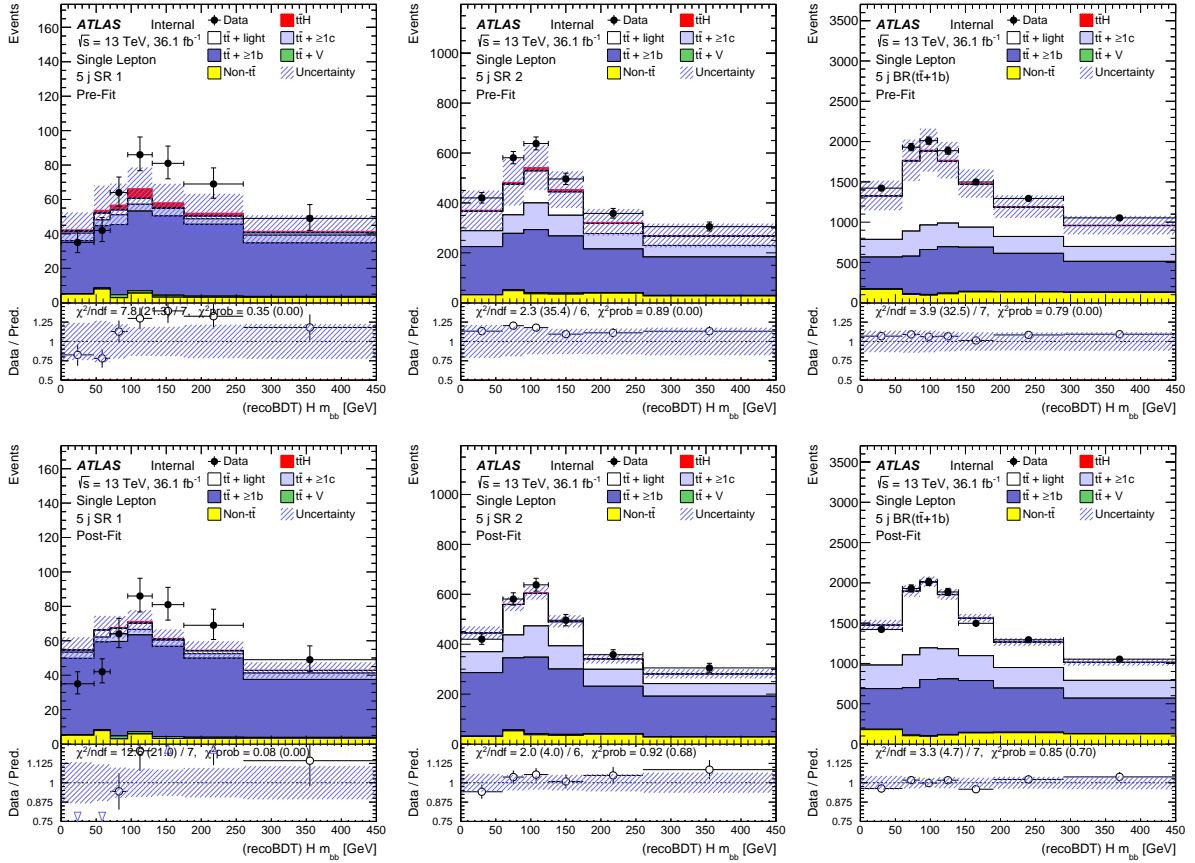
References

[1] C. Patrignani et al. (Particle Data Group), *The Review of Particle Physics (2017)*, Chin. Phys. C, 40, 100001 (2016) and 2017 update. `//pdg.lbl.gov`

[2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. 19 (1967) 12641266.

[3] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. 13 (1964).

[4] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. 22 (1961) 579588.

[5] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. 8 (1964) 214215.

[6] H. Fritzsch, M. Gell-Mann, H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys.Lett. 47B (1973) 365-368.

[7] S. Weinberg, *Nonabelian Gauge Theories of the Strong Interactions*, Phys.Rev.Lett. 31 (1973) 494-497.

[8] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. 30 (1973) 1343.

[9] G. Altarelli, *Partons in Quantum Chromodynamics*, Phys. Rept. 81 (1982) 1.

[10] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. 13 (1964) 321323.

[11] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, hys. Lett. 12 (1964) 132133.

[12] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. 13 (1964) 508509.

[13] P. Langacker, *Structure of the Standard Model*, arXiv:hep-ph/0304186.

[14] A. Djouadi, *The Anatomy of Electro-Weak Symmetry Breaking. I: The Higgs boson in the Standard Model*, Phys. Rept. 457 (2008) 1-216, arXiv:hep-ph/0503172.

[15] J. Baglio, A. Djouadi, *Higgs production at the lHC*, JHEP 1103:055 (2011), arXiv:1012.0530 [hep-ph]

[16] CDF Collaboration, D0 Collaboration, *Evidence for a particle produced in association with weak bosons and decaying to a bottom-antibottom quark pair in Higgs boson searches at the Tevatron*, Phys. Rev. Lett. 109, 071804 (2012), arXiv:1207.6436 [hep-ex].

[17] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties : Report of the LHC Higgs Cross Section Working Group*, CERN-2013-004, arXiv:1307.1347 [hep-ph], FERMILAB-CONF-13-667-T.

[18] LHC Higgs Cross Section Working Group,
`https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG`

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

[20] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B 716 (2012) 30, arXiv:1207.7235 [hep-ex].

[21] ATLAS collaboration, *Measurements of the properties of the Higgs-like boson in the four lepton decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data*, ATLAS-CONF-2013-013.

[22] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3, S08003 (2008).
`https://cds.cern.ch/record/1129811/`

[23] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST 3 (2008) S08002.

[24] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST 3 (2008) S08004.

[25] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST 3 (2008) S08005.

[26] LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, JINST 3 (2008) S08006.

[27] TOTEM Collaboration, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST 3 (2008) S08007.

[28] MoEDAL Collaboration *Technical Design Report of the MoEDAL Experiment*, CERN-LHCC-2009-006, MoEDAL-TDR-001

[29] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013, ATLAS-TDR-19. `https://cds.cern.ch/record/129163`

[30] `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2015-007/`

[31] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022. `https://cds.cern.ch/record/2037697`

[32] ATLAS Collaboration, *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s}$ = 13 TeV*, ATL-PHYS-PUB-2015-018.
`https://cds.cern.ch/record/2037683`

[33] ATLAS Collaboration, *The Optimization of ATLAS Track Reconstruction in Dense Environments,* ATL-PHYS-PUB-2015-006. `https://cds.cern.ch/record/2002609`

[34] ATLAS Collaboration, *Commissioning of the ATLAS b-tagging algorithms using $t\bar{t}$ events in early Run-2 data,* ATL-PHYS-PUB-2015-039. `http://cds.cern.ch/record/2047871`

[35] Y. Nakahama, *The ATLAS Trigger System: Ready for Run-2,* J. Phys. Conf. Ser. 664 no. 8, (2015) 082037.

[36] Gleisberg, T. et al. *Event generation with SHERPA 1.1,* JHEP 0902 (2009) 007 arXiv:0811.4622 [hep-ph] FERMILAB-PUB-08-477-T, SLAC-PUB-13420, ZU-TH-17-08, DCPT-08-138, IPPP-08-69, EDINBURGH-2008-30, MCNET-08-14

[37] T. G. Cornelissen, N. Van Eldik, M. Elsing, W. Liebig, E. Moyse, N. Piacquadio, K. Prokofiev, A. Salzburger, and A. Wildauer, *Updates of the ATLAS Tracking Event Data Model (Release 13),* ATL-SOFT-PUB-2007-003. `https://cds.cern.ch/record/1038095`

[38] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS experiment,* ATLAS-CONF-2010-069. `https://cds.cern.ch/record/1281344`

[39] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data,* ATLAS-CONF-2016-024. `https://cds.cern.ch/record/2157687`

[40] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in protonproton collision data at $\sqrt{s}$=13 TeV,* arXiv:1603.05598 [hep-ex].

[41] S. Hassani, L. Chevalier, E. Lançon, J.-F. Laporte, R. Nicolaidou, A. Ouraou, *A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages,* Nucl.Instrum.Meth. A572 (2007) 77-79.

[42] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics,* arXiv:0901.0512 [hep-ex], SLAC-R-980, CERN-OPEN-2008-020, CERN-OPEN-2008-020.

[43] ATLAS Collaboration, *The anti-k_t jet clustering algorithm,* JHEP 0804:063,2008, arXiv:0802.1189 [hep-ph].

[44] C. Cojocaru et al., *Hadronic Calibration of the ATLAS Liquid Argon End-Cap Calorimeter in the Pseudorapidity Region $1.6 < |\eta| < 1.8$ in Beam Tests,* arXiv:physics/0407009.

[45] T. Barillari et al., *Local Hadronic Calibration,* ATL-LARG-PUB-2009-001.

[46] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV,* arXiv:1112.6426 [hep-ex], CERN-PH-EP-2011-191.

[47] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisio359 ns recorded in 2012 at $\sqrt{s} = 8$ TeV,* ATLAS-CONF-2013-082.

[48] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV with ATLAS,* arXiv:1108.5602 [hep-ex].

[49] ATLAS Collaboration, *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV,* ATL-PHYS-PUB-2015-023.

[50] GEANT4 Collaboration (S. Agostinelli et al.), *GEANT4: A Simulation toolkit,* Nucl.Instrum.Meth. A506 (2003) 250303.

[51] ATLAS Collaboration, *The ATLAS Simulation Infrastructure,* Eur.Phys.J.C70:823-874,2010, arXiv:1005.4568.

[52] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report,* CERN-LHCC-2010-013, ATLAS-TDR-19. `https://cds.cern.ch/record/129163`

[53] `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2015-007/`

[54] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2,* ATL-PHYS-PUB-2015-022. `https://cds.cern.ch/record/2037697`

[55] ATLAS Collaboration, *The Optimization of ATLAS Track Reconstruction in Dense Environments,* ATL-PHYS-PUB-2015-006. `https://cds.cern.ch/record/2002609`

[56] ATLAS Collaboration, *Commissioning of the ATLAS b-tagging algorithms using $t\bar{t}$ events in early Run-2 data,* ATL-PHYS-PUB-2015-039. `http://cds.cern.ch/record/2047871`

[57] J. Alwall at al., *MadGraph 5: Going Beyond,* JHEP 06 (2011) 128, arXiv:1106.0522 [hep-ph].

[58] NNPDF Collaboration, *Parton distributions for the LHC Run II,* JHEP 04 (2015) 040, arXiv:1410.8849 [hep-ph].

[59] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *Introduction to PYTHIA 8.1,* Comput. Phys. Commun. 178 (2008) 852, arXiv:0710.3820 [hep-ph].

[60] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes,* ATL-PHYS-PUB-2012-003, 2012. `http://cds.cern.ch/record/1474107`.

[61] S. Alioli et al. *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP 1006 (2010) 043, arXiv: 1002.2581 [hep-ph].

[62] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A462 (2001) 152155.

[63] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Comput. Phys. Commun. 185 (2014) 2930, arXiv: 1112.5675 [hep-ph].

[64] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP 0902 (2009) 007, arXiv: 0811.4622 [hep-ph].

[65] F. Cascioli, P. Maierhofer and S. Pozzorinis, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. 108 (2012) 111601, arXiv: 1111.5206 [hep-ph]

[66] J. Bellm et al. *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C 76 (2016) 196, arXiv: 1512.01178 [hep-ph].

[67] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP 0812 (2008) 039, arXiv: 0808.3674 [hep-ph].

[68] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP 0803 (2008) 038, arXiv: 0709.1027 [hep-ph].

[69] S. Höche, et al., *QCD matrix elements + parton showers: The NLO case*, JHEP 04 (2013) 027, arXiv: 1207.5030 [hep-ph].

[70] S. Frixione et al., *Single-top hadroproduction in association with a W boson*, JHEP 0807 (2008) 029, arXiv: 0805.3067 [hep-ph].

[71] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B 867 (2013) 244, arXiv: 1207.1303 [hep-ph].

[72] ATLAS collaboration, *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, ATLAS-CONF-2014-058, https://cds.cern.ch/record/1951336

[73] F. Derue, *Estimation of fake lepton background for top analyses using the Matrix Method with the 2015 dataset at $\sqrt{s} = 13$ TeV with AnalysisTop-2.3.41.*, ATL-COM-PHYS-2016-198. https://cds.cern.ch/record/2135116

[74] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS ACAT (2007) 040, arXiv:physics/0703039.

[75] R. E. Ticse Torres, *Search for the Higgs boson in the ttH(H → bb) channel and the identification of jets containing two B hadrons with the ATLAS experiment (PhD thesis)*, CERN-THESIS-2016-123, http://cds.cern.ch/record/2226037/

[76] ATLAS collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C76 (2016) 653, arXiv: 1608.03953 [hep-ex].

[77] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-015 (2015), https://cds.cern.ch/record/2037613

[78] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. C 75 (2015) 17, arXiv: 1406.0076 [hep-ex].

[79] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Collider*, Comput. Phys. Commun. 185 (2014) 2930, arXiv: 1112.5675 [hep-ph].

[80] A. D. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. C63 (2009) 189285, arXiv: 0901.0002 [hep-ph].

[81] ATLAS Collaboration,, *Studies of tt + cc production with MadGraph5_aMC@NLO and Herwig++ for the ATLAS experiment*, ATL-PHYS-PUB-2016-011. https://cds.cern.ch/record/2153876

[82] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, Phys. Rev. D 82 (2010) 054018, arXiv: 1005.4451.

[83] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. D 81 (2010) 054028, arXiv: 1001.5034.

[84] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. D 83 (2011) 091503, arXiv: 1103.2792.

[85] R. Raitio and W. W. Wada, *Higgs Boson Production at Large Transverse Momentum in QCD*, Phys. Rev. D19 (1979) 941.

[86] W. Beenakker et al., *NLO QCD corrections to t anti-t H production in hadron collisions*, Nucl. Phys. B653 (2003) 151, arXiv: hep-ph/0211352.

[87] S. Dawson, C. Jackson, L. H. Orr, L. Reina and D. Wackeroth, *Associated Higgs production with top quarks at the Large Hadron Collider: NLO QCD corrections*, Phys. Rev. D68 (2003) 034022, arXiv: hep-ph/0305087.

[88] Y. Zhang, W.-G. Ma, R.-Y. Zhang, C. Chen and L. Guo, *QCD NLO and EW NLO corrections to $t\bar{t}H$ production with top quark decays at hadron collider*, Phys. Lett. B738 (2014) 1, arXiv: 1407.1110 [hep-ph].

[89] S. Frixione, V. Hirschi, D. Pagani, H.-S. Shao and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, JHEP 06 (2015) 184, arXiv: 1504.03446 [hep-ph].

[90] L. Moneta et al., *The RooStats Project*, PoS ACAT2010 (2010) 057, arXiv:1009.1003 [physics.data-an].

[91] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C 71 (2011) 1554, arXiv:1007.1727v2 [physics.data-an].