Palacký University Olomouc
Faculty of Science
Department of Experimental physics

DOCTORAL THESIS

Study of Heavy Particles Production and Properties with the ATLAS Experiment

Author:	**Mgr. Petr Baroň**
Study programme:	Physics
Specialization:	Applied Physics
Study form:	Daily
Supervisor:	Mgr. Jiří Kvita, Ph.D.
Date of submission:	October 2024

Prohlášení

Prohlašuji, že jsem předloženou disertační práci vypracoval samostatně pod vedením Mgr. Jiřího Kvity, Ph.D. a že jsem použil zdrojů, které cituji a uvádím v seznamu použitých pramenů.

V Olomouci dne .....

.....

Mgr. Petr Baroň

Poděkování

Rád bych poděkoval svému školiteli, Mgr. Jiřímu Kvitovi, Ph.D., za jeho vedení, ochotu a vstřícný přístup během mého studia. Získané znalosti uplatním ve své další vědecké kariéře. Dále bych rád poděkoval vedení a zaměstnancům Společné Laboratoře Optiky SLO za podporu a příjemné pracovní prostředí. Velké díky patří také mé rodině za jejich podporu a trpělivost, kterou mi projevovali v průběhu celého studia.

Za podporu děkuji projektům IGA_PrF_2024_004 University Palackého, Získávání nových poznatků o mikrosvětě v infrastruktuře CERN (Projekt MŠMT I-E-T LTT17018), Výzkumná infrastruktura pro experimenty v CERN (Projekty MŠMT LM2018104 a LM2023040), Nové techniky rekonstrukce boostovaných top kvarků pro hledání nové fyziky na LHC (GAČR č. GA19-21484S), and Aplikace strojového učení ve fyzice vysokých energií a astročásticové fyzice (GAČR č. 23-07110S).

Acknowledgements

I would like to express my sincere thanks to my supervisor, Mgr. Jiří Kvit, Ph.D., for his guidance, willingness, and supportive approach throughout my studies. The knowledge I have gained will be invaluable in my future scientific career. I would also like to thank the management and staff of the Joint Laboratory of Optics (JLO) for their support and the pleasant working environment. My heartfelt thanks also go to my family for their support and patience during my studies.

The author gratefully acknowledge the support from the projects IGA_PrF_2024_004 of Palacky University, and Getting new knowledge of the microworld using the CERN infrastructure (Project MSMT I-E-T LTT17018), Research infrastructure for experiments at CERN (Projects MSMT LM2018104 and LM2023040), Novel techniques for boosted top quarks reconstruction for new physics searches at LHC (GACR 19-21484S), and Machine learning applications in high-energy and astroparticle physics (GACR 23-07110S).

Bibliografická identifikace

Jméno a příjmení	Mgr. Petr Baroň
Název práce	Zkoumání produkce a vlastností těžkých částic v experimentu ATLAS
Typ práce	Disertační
Pracoviště	Katedra experimentální fyziky
Vedoucí práce	Mgr. Jiří Kvita, Ph.D.
Rok obhajoby práce	2024
Abstrakt	Cílem této disertační práce je popsat proces měření účinného průřezu produkce párů top kvarů ve srážkách proton-olovo při energii srážek na jeden nukleon $\sqrt{s_{\mathrm{NN}}} =$ 8,16 TeV v experimentu ATLAS na urychlovači LHC. Vzhledem k tomu, že top kvark má největší hmotnost ze všech elementárních částic, slouží jako sonda pro jaderné partonové distribuční funkce (nPDF) při vysokých hodnotách Bjorkenova x. Jaderné partonové distribuční funkce jsou běžně používány ve fyzice těžkých iontů k popisu jevů jako například stínění, anti-stínění a Fermiho pohyb ve srovnání s distribučními funkcemi ve srážkách proton-olovo. Měření produkce párů top kvarků ve srážkách proton-olovo představuje další datový bod pro fitování jaderných partonových fukncí, a je proto velmi důležité provést měření s co nejmenšími možnými nejistotami. Bylo tak potřeba provést analýzu včetně dodatečných studií pro prostředí s nízkou mírou pile-up efektu ve srážkách proton-olovo, jako je například odvození škálovacích faktorů pro elektrony a miony, energetické korekce jetů, systematiky tvaru fake lepton pozadí nebo provedení párování mezi dvěmi kolekcemi jetů. Strategie měření a oblasti fázového prostoru pro fitování musely být pečlivě zvoleny, aby bylo dosaženo maximální statistiky a poměru signálu k pozadí. Jako referenční měření pro analýzu slouží měření produkce párů top kvarů v kanále lepton+jets ve srážkách proton-olovo, zveřejněné experimentem CMS. Studie v této disertační práci představuje první analýzu, která měří páry top kvarků nejen v lepton+jets, ale také v dileptonovém kanále. Tato práce poskytuje naměřený účinný průřez produkce párů top kvarků ve srážkách proton-olovo v dileptonovém a lepton+jets kanále v experimentu ATLAS na urychlovači LHC v laboratoři CERN.
Klíčová slova	top kvark, srážky proton-olovo, nPDF, ATLAS experiment, velký hadronový urychlovač, měření účinného průřezu, fyzika těžkých iontů
Počet stran	144
Jazyk	anglický

Bibliographical identification

First name and surname	Mgr. Petr Baroň
Title	Study of Heavy Particles Production and Properties with the ATLAS Experiment
Type of thesis	Dissertation
Department	Department of Experimental Physics
Supervisor	Mgr. Jiří Kvita, Ph.D.
The year of presentation	2024
Abstract	The aim of this thesis is to describe the process of measuring the cross-section of the top quark pairs production in proton-lead collisions at the energy per nucleon $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV with the ATLAS experiment at the Large Hadron Collider (LHC) at CERN. Since the top quark has the largest mass of all elementary particles it serves as a probe of nuclear parton distribution functions (nPDFs) at high Bjorken-x. The nPDF are commonly used in Heavy Ion physics and describe different phenomena (e.g. shadowing, anti-shadowing, Fermi movement) compared to proton-proton collision processes. As the measurement of the $t\bar{t}$ production in p+Pb collisions represents a new data point in the fit of nPDF sets, it is very important to perform the analysis with the lowest uncertainties possible. The cross-section measurement had to be performed along additional studies for low pile-up in the p+Pb environment, namely derivation of electron and muon scale factors, jet energy scale and jet energy resolution corrections, shape systematics of the fake lepton background or performing matching between two jet collections. The strategy and phase-space regions of the global fit to the signal strength had to be chosen carefully to gain the maximum statistics and signal to background ratio. The reference analysis is the measurement of $t\bar{t}$ production in the lepton+jets channel in p+Pb collisions released by the CMS experiment, however, the study presented in this thesis is the first analysis to search for the $t\bar{t}$ production not only in lepton+jets channel, but also in the dilepton channel. This thesis provides the total measured cross-section of $t\bar{t}$ pairs in p+Pb collisions in lepton+jets and dilepton channels with the ATLAS experiment by LHC at CERN.
Keywords	top quark, proton-lead collisions, nPDF, ATLAS experiment, Large Hadron Collider, cross-section measurement, heavy-ion physics
Number of pages	144
Language	English

Table of Contents

Introduction

This thesis focuses on measuring the cross-section of the $t\bar{t}$ production in proton-lead collisions at the center-of-mass energy of $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV using the ATLAS detector at the Large Hadron Collider (LHC). Top-quark pair production in proton-lead collisions has not been extensively studied, and this measurement represents a step toward constraining nPDFs, particularly at high Bjorken-x, where theoretical uncertainties remain large.

The first chapter of this thesis, titled "Theory", provides the needed theoretical background. It begins with a discussion of the Standard Model, focusing on the interactions and particles within the model, including the role of the top quark in both electroweak and strong interactions. Following this, the properties of the top quark are explored, highlighting its mass, short lifetime, and significance as a probe for new physics, especially in relation to $t\bar{t}$ production. The Lagrangian of the Quantum Chromodynamics (QCD) is then introduced, explaining the interactions between quarks and gluons. The chapter also covers the parton shower and hadronization processes, describing how partons evolve into hadrons and how Monte Carlo simulations model these processes along with detector responses.

In the second chapter "Experimental Tools" the design and operation of the Large Hadron Collider are described, particularly its ability to collide protons and lead ions at high energies. Lastly, the chapter provides an overview of the ATLAS detector, emphasizing key subsystems such as tracking, calorimetry, and muon detectors.

The third chapter, "Measurement of $t\bar{t}$ production in lepton+jets and dilepton channels in p+Pb collisions," describes the experimental setup and analysis approach of the key measurement of this thesis. It starts by outlining the scientific motivation for studying $t\bar{t}$ production in p+Pb collisions and provides details on the datasets used for the analysis, including proton-lead collision data from the ATLAS experiment. The criteria for object selection, such as leptons, jets, and missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$), are defined, followed by the event selection criteria to optimize the signal and background separation. Background estimation methods, particularly for fake leptons, are explained. The performance of electron and muon reconstruction, including efficiency and scale factor corrections, is discussed, along with the jet performance, focusing on the jet energy scale (JES) and resolution corrections. The chapter also outlines various sources of systematic uncertainties affecting the measurement and discusses the statistical methods used to extract the signal strength through a fit procedure. Finally, the results of the analysis, including the measured $t\bar{t}$ production cross-section in the lepton+jets and dilepton channels, are presented and compared with theoretical predictions and previous measurements.

Theory

This chapter covers main points necesary to understand top quark pair production, decay modes, detector signature, differences between pp and p+Pb collisions environment with respect of the measurement described in the Chapter 3. By the end of this chapter the motivation and importance of the measurement as a new data point in unexplored kinematic region for the fit of nPDFs should emerge as well as the effort to keep measurement systematic uncertainties as low as possible, but at the same time keep uncertainties conservative enough and under control of analysers.

1.1 Standard Model

The Standard Model (SM) of particle physics is a theory that describes three of the four fundamental forces of nature: electromagnetism, the weak nuclear force, and the strong nuclear force. It provides a framework for understanding the behavior of elementary particles (see Figure 1.1), which are the building blocks of matter. These particles are divided into two categories: fermions, which make up matter, and bosons, which mediate the fundamental forces.

Fermions are further divided into quarks and leptons. There are six types of quarks (up, down, charm, strange, top, and bottom) and six types of leptons (electron, muon, tau, and their corresponding neutrinos). Quarks combine to form hadrons, e.g. protons and neutrons, through the strong interaction, which is mediated by gluons. The gauge bosons associated with the electromagnetic, weak, and strong interactions are the photon, W and Z bosons, and gluons, respectively. Additionally, the Higgs boson theoretized in 1964 by Peter Higgs [1] and François Englert and Robert Brout [2], discovered in 2012 [3], is responsible for imparting mass to the particles through the Higgs mechanism.

1.2 Top Quark

The top quark discovered in 1995 at Fermilab's Tevatron in $p\bar{p}$ collisions [5] is the heaviest elementary particle known to date and therefore an important topic to study in the context of the Standard Model at the Large Hadron Collider. Top quarks at the LHC are predominantly produced as top quark-antiquark pairs $t\bar{t}$ (or four top production) via the strong interaction, but can also be produced as single top quarks via the electroweak interaction. Since the top quark almost always decays into a W boson and a bottom quark before hadronization, the final state is characterized by the subsequent decay of the W boson, which can either happen into a quark-antiquark pair or into a charged lepton and the corresponding neutrino. Due to their abundant production at the LHC and their clear experimental signature top quarks provide a

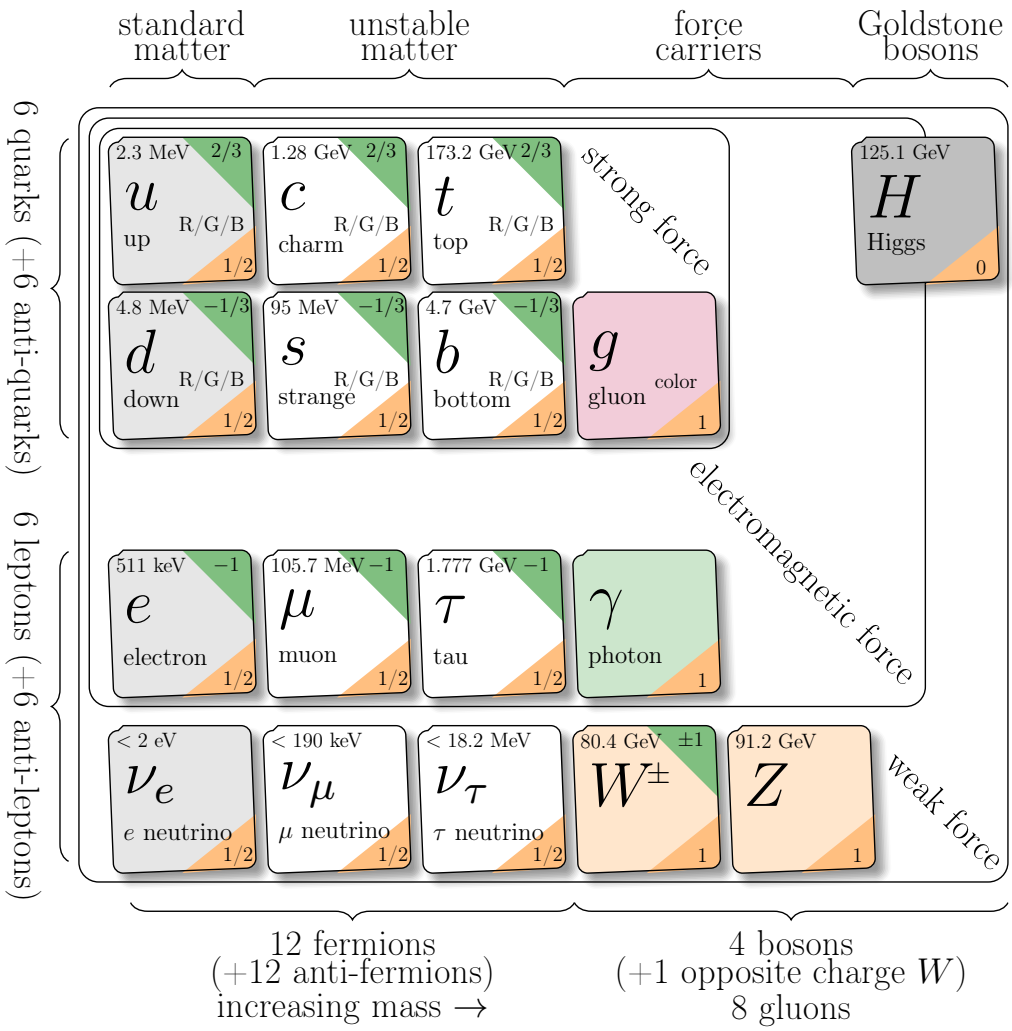


Figure 1.1: Elementary paricles sorted by their properties. Generated using TikZ-Feynman package [4].

unique tool for probing the Standard Model parameters and search for deviations to theoretical predictions, which would be an indication for physics beyond the SM [6].

1.3 Lagrangian of Quantum Chromo-dynamics

The Lagrangian of the Quantum Chromo-dynamics (QCD) can be built up by requiring invariance of the quark fields under local gauge symmetry of SU(3)

$$\psi' = e^{i\Theta_a(x)T_a}\psi \equiv S\psi \tag{1.1}$$

on the Dirac Lagrangian

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi \tag{1.2}$$

where ψ is a quark field covering new degree of freedom, the color (labelled as red, green, blue) as well as spinor components. Substituting term 1.1 into Lagrangian 1.2

yields

$$\begin{aligned}
\mathcal{L}'_{\text{Dirac}} &= \bar{\psi}'(i\gamma^\mu\partial_\mu - m)\psi' \\
&= \bar{\psi}e^{-i\Theta_a(x)T_a}(i\gamma^\mu\partial_\mu - m)e^{i\Theta_a(x)T_a}\psi \\
&= \bar{\psi}e^{-i\Theta_a(x)T_a}i\gamma^\mu\partial_\mu(e^{i\Theta_a(x)T_a}\psi) - m\bar{\psi}\psi \\
&= \bar{\psi}e^{-i\Theta_a(x)T_a}i\gamma^\mu\left[e^{i\Theta_a(x)T_a}(\partial_\mu\psi + iT_a\partial_\mu\Theta_a(x)\psi)\right] - m\bar{\psi}\psi \\
&= \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - \bar{\psi}\gamma^\mu T_a\partial_\mu\Theta_a(x)\psi. \qquad (1.3)
\end{aligned}$$

To cancel out the last term of Equation 1.3 the $\mathcal{L}_{\text{Dirac}}$ can be modified:

$$\mathcal{L}_{\text{modif}} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi \qquad (1.4)$$

defining the covariant derivative

$$D_\mu = \partial_\mu - ig_sT^aA^a_\mu \qquad (1.5)$$

by introducing new eight vector fileds A^a_μ transforming as

$$D'_\mu = \partial_\mu - ig_sT^aA'^a_\mu. \qquad (1.6)$$

To satisfy invariance of $\mathcal{L}_{\text{modif}}$ and find transformation of gauge fileds A^a_μ the following terms have to be equal (denoting $\Theta_a(x)$ as Θ_a for better readability)

$$\begin{aligned}
D'_\mu\psi' &= \partial_\mu\psi' - ig_sT^aA'^a_\mu\psi' \\
&= \partial_\mu(e^{i\Theta_aT_a}\psi) - ig_sT^aA'^a_\mu(e^{i\Theta_aT_a}\psi) \\
&= (\partial_\mu e^{i\Theta_aT_a})\psi + e^{i\Theta_aT_a}\partial_\mu\psi - ig_sT^aA'^a_\mu(e^{i\Theta_aT_a}\psi) \\
&= iT_ae^{i\Theta_aT_a}\partial_\mu(\Theta_a)\psi + e^{i\Theta_aT_a}\partial_\mu\psi - ig_sT^aA'^a_\mu(e^{i\Theta_aT_a}\psi) \\
&\stackrel{?}{=} e^{i\Theta_aT_a}(\partial_\mu\psi - ig_sT^aA^a_\mu\psi) \qquad (1.7)
\end{aligned}$$

The condition in Equation 1.7 is fulfilled once the term $T_aA'^a_\mu$ is transformed as

$$T^aA'^a_\mu = e^{i\Theta_aT_a}T^aA^a_\mu e^{-i\Theta_aT_a} + \frac{i}{g_s}e^{i\Theta_aT_a}\partial_\mu(\Theta_a)(-iT_a)e^{-i\Theta_aT_a} \qquad (1.8)$$

or in a more compact way

$$\begin{aligned}
A_\mu &\equiv T^aA^a_\mu \qquad (1.9)\\
S &\equiv e^{i\Theta_aT_a} \qquad (1.10)\\
A'_\mu &= SA_\mu S^{-1} + \frac{i}{g_s}S\partial_\mu S^{-1}. \qquad (1.11)
\end{aligned}$$

Taking into account only infinitesimal transformation of Equation 1.10,

$$\begin{aligned}
S &= 1 + i\Theta^aT^a \qquad (1.12)\\
S^{-1} &= 1 - i\Theta^aT^a \qquad (1.13)
\end{aligned}$$

and pluging-in into Equation 1.11

$$\begin{aligned}
A'_\mu &\simeq (1+i\Theta^a T^a)A_\mu(1-i\Theta^a T^a) + \frac{i}{g_s}(1+i\Theta^a T^a)\partial_\mu(1-i\Theta^a T^a) \\
&= (A_\mu + i\Theta^a T^a A_\mu)(1-i\Theta^a T^a) + \frac{i}{g_s}(-i)T^a\partial_\mu(\Theta^a) + \\
&+ \frac{i}{g_s}\Theta^a T^a\partial_\mu(\Theta^a) \\
&= A_\mu - iA_\mu\Theta^a T^a + i\Theta^a T^a A_\mu + (\Theta_a)(T^a)A_\mu(\Theta_a)(T^a) + \frac{1}{g}\partial_\mu(\Theta^a) \\
&+ \frac{i}{g_s}\Theta^a T^a\partial_\mu(\Theta^a) \\
&= A_\mu + i\left[\Theta^a T^a, A_\mu\right] + (\Theta_a)(T^a)A_\mu(\Theta_a)(T^a) + \frac{1}{g}\partial_\mu(\Theta^a) \\
&+ \frac{i}{g_s}\Theta^a T^a\partial_\mu(\Theta^a) \\
&= A_\mu + i\Theta^a A^b_\mu(if^{abc}T^c) + (\Theta_a)(T^a)A_\mu(\Theta_a)(T^a) + \frac{1}{g}\partial_\mu(\Theta^a) \\
&+ \frac{i}{g_s}\Theta^a T^a\partial_\mu(\Theta^a) \\
&= |\text{neglecting } O(\Theta^2) \text{ \& cycl. } f^{abc}| = A_\mu - f^{abc}\Theta^b A^c_\mu + \frac{1}{g}\partial_\mu(\Theta^a) \qquad (1.14)
\end{aligned}$$

One additional remark, having just one generator the Equation 1.11 would lead to photon field transformation as used in Quantum Electro-dynamics (QED)

$$A'_\mu = A_\mu + \frac{1}{g}\partial_\mu\Theta. \qquad (1.15)$$

The Lagrangian $\mathcal{L}_{\text{modif}}$ (1.4) is gauge-invariant under transformations (1.1), (1.6), and (1.11). However, the term for gluon fields has not been yet introduced. Being inspired by QED the part of kinetic term $F_{\mu\nu}F^{\mu\nu}$ is given by

$$\begin{aligned}
-ig_s F_{\mu\nu} f &= \left[D_\mu, D_\nu\right] f \\
&= (\partial_\mu - ig_s A_\mu)(\partial_\nu - ig_s A_\nu)f - (\partial_\nu - ig_s A_\nu)(\partial_\mu - ig_s A_\mu)f \\
&= \partial_\mu\partial_\nu f - ig_s\partial_\mu(A_\nu f) - ig_s A_\mu\partial_\nu f - g_s^2 A_\mu A_\nu f \\
&- \partial_\nu\partial_\mu f + ig_s\partial_\nu(A_\mu f) + ig_s A_\nu\partial_\mu f + g_s^2 A_\nu A_\mu f \\
&= \partial_\mu\partial_\nu f - ig_s\partial_\mu(A_\nu)f - ig_s A_\nu\partial_\mu f - ig_s A_\mu\partial_\nu f - g_s^2 A_\mu A_\nu f \\
&- \partial_\nu\partial_\mu f + ig_s\partial_\nu(A_\mu)f + ig_s A_\mu\partial_\nu f + ig_s A_\nu\partial_\mu f + g_s^2 A_\nu A_\mu f \\
&= -ig_s(\partial_\mu A_\nu - \partial_\nu A_\mu)f - g_s^2\left[A_\mu, A_\nu\right] f \\
&= -ig_s(\partial_\mu A_\nu - \partial_\nu A_\mu - ig_s\left[A_\mu, A_\nu\right])f. \qquad (1.16)
\end{aligned}$$

The components of strength field tensor thus reads

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_s f^{abc} A^b_\mu A^c_\nu. \qquad (1.17)$$

Having all terms prepared, the QCD Lagrangian is given by:

$$\begin{aligned}
\mathcal{L}_{\text{QCD}} &= \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}F^a_{\mu\nu}F^{\mu\nu}_a && (1.18)\\
&= \bar{\psi}(i\not{\partial} - m)\psi + g_s\bar{\psi}\gamma^\mu A^a_\mu T^a\psi \\
&- \frac{1}{4}(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_s f^{abc}A^b_\mu A^c_\nu)(\partial^\mu A^\nu_a - \partial^\nu A^\mu_a + g_s f_{abc}A^\mu_b A^\nu_c) \\
&= \bar{\psi}(i\not{\partial} - m)\psi + g_s\bar{\psi}\gamma^\mu A^a_\mu T^a\psi \\
&- \frac{1}{4}(\partial_\mu A^a_\nu\partial^\mu A^\nu_a - \partial_\mu A^a_\nu\partial^\nu A^\mu_a + g_s\partial_\mu A^a_\nu f_{abc}A^\mu_b A^\nu_c - \partial_\nu A^a_\mu\partial^\mu A^\nu_a + \partial_\nu A^a_\mu\partial^\nu A^\mu_a \\
&- g_s\partial_\nu A^a_\mu f_{abc}A^\mu_b A^\nu_c + g_s f^{abc}A^b_\mu A^c_\nu\partial^\mu A^\nu_a - g_s f^{abc}A^b_\mu A^c_\nu\partial^\nu A^\mu_a \\
&+ g_s^2 f^{abc}A^b_\mu A^c_\nu f_{abc}A^\mu_b A^\nu_c) \\
&= \underbrace{\bar{\psi}(i\not{\partial} - m)\psi - \frac{1}{2}\partial_\mu A^a_\nu\partial^\mu A^\nu_a + \frac{1}{2}\partial_\mu A^a_\nu\partial^\nu A^\mu_a}_{\mathcal{L}_0} && (1.19)\\
&+ \underbrace{g_s\bar{\psi}\gamma^\mu A^a_\mu T^a\psi}_{\mathcal{L}_{qqg}} && (1.20)\\
&- \underbrace{\frac{g_s f_{abc}}{4}(\partial_\mu A^a_\nu A^\mu_b A^\nu_c - \partial_\nu A^a_\mu A^\mu_b A^\nu_c + A^b_\mu A^c_\nu\partial^\mu A^\nu_a - A^b_\mu A^c_\nu\partial^\nu A^\mu_a)}_{\mathcal{L}_{3g}} && (1.21)\\
&+ \underbrace{\frac{g_s^2}{4}f^{abc}A^b_\mu A^c_\nu f_{ade}A^\mu_d A^\nu_e}_{\mathcal{L}_{4g}} && (1.22)
\end{aligned}$$

where as a consequence of requiring local gauge symmetry the quark-gluon and gluon-gluon interaction terms emerged $\mathcal{L}_{qqg}$ (1.20), $\mathcal{L}_{3g}$ (1.21), and $\mathcal{L}_{4g}$ (1.22).

Having the $\mathcal{L}_{\text{QCD}}$ the quark propagator is defined as an inverse operator of a solution of the equation of motion of free quark field in momentum space.

$$\begin{aligned}
\frac{\partial\mathcal{L}_{0,\text{Dirac.}}}{\partial\bar{\psi}} - \frac{\partial\mathcal{L}_{0,\text{Dirac.}}}{\partial_\mu(\partial_\mu\bar{\psi})} &= 0 && (1.23)\\
(i\not{\partial} - m)\psi &= 0. && (1.24)\\
& && (1.25)
\end{aligned}$$

To solve for the field ψ a Green's function $S_F(x-y)$ has to fulfill the equation

$$(i\not{\partial} - m)S_F(x-y) = \delta^4(x-y) \tag{1.26}$$

in momentum space given by

$$\begin{aligned}
(i(-i)\not{p} - m)\left(\int\frac{d^4p}{(2\pi)^4}e^{-ip(x-y)}\tilde{S}_F(p)\right) &= \int\frac{d^4p}{(2\pi)^4}e^{-ip(x-y)} && (1.27)\\
\left(\int\frac{d^4p}{(2\pi)^4}e^{-ip(x-y)}(\not{p} - m)\tilde{S}_F(p)\right) &= \int\frac{d^4p}{(2\pi)^4}e^{-ip(x-y)} && (1.28)\\
(\not{p} - m)\tilde{S}_F(p) &= 1 && (1.29)
\end{aligned}$$

which gives, with the Feynman $+i\epsilon$ prescription, the quark propagator $\tilde{S}_F(p)$

$$\begin{aligned}
\tilde{S}_F(p) &= \frac{i}{\not{p} - m + i\epsilon} && (1.30)\\
\text{or} \qquad \tilde{S}_F(p) &= \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon}, && (1.31)
\end{aligned}$$

where the $+i\epsilon$ term ensures the correct formulation of the propagator as a Green's function that respects time-ordering in quantum field theory. For the gluon propagator the gauge fixing term $-\frac{1}{2}(\partial_\mu A^{a\mu})^2$ is added to free gluon field

$$\mathcal{L}_{0,\text{gluon}} = \frac{1}{2}(\partial_\mu A^a_\nu \partial^\nu A^\mu_a - \partial_\mu A^a_\nu \partial^\mu A^\nu_a) - \frac{1}{2}(\partial_\mu A^{a\mu})^2. \tag{1.32}$$

Solving the Euler-Lagrangian equation of motion gives

$$\frac{\partial \mathcal{L}_{0,\text{gluon}}}{\partial A^a_\nu} - \partial_\mu \left(\frac{\partial \mathcal{L}_{0,\text{gluon}}}{\partial(\partial_\mu A^a_\nu)} \right) = 0 \tag{1.33}$$

$$0 - \partial_\mu \frac{\partial}{\partial(\partial_\mu A^a_\nu)} \left(\frac{1}{2}\partial_\alpha A^a_\beta \partial^\beta A^{a\alpha} \right) + \partial_\mu \frac{\partial}{\partial(\partial_\mu A^a_\nu)} \left(\frac{1}{2}\partial_\alpha A^a_\beta \partial^\alpha A^{a\beta} \right) + \partial_\mu \frac{\partial}{\partial(\partial_\mu A^a_\nu)} \left(\frac{1}{2}(\partial_\alpha A^{a\alpha})^2 \right) = 0 \tag{1.34}$$

$$\partial_\mu \left(\frac{1}{2}\delta^\mu_\alpha \delta^\nu_\beta \partial^\beta A^{a\alpha} - \frac{1}{2}\delta^\mu_\alpha \delta^\nu_\beta \partial^\alpha A^{a\beta} - \frac{1}{2}2(\partial^\nu A^{a\nu})\delta^\mu_\alpha \delta^\nu_\alpha \right) = 0 \tag{1.35}$$

$$\partial_\mu \left(\frac{1}{2}\partial^\nu A^{a\mu} - \frac{1}{2}\partial^\mu A^{a\nu} - \partial^\nu A^{a\nu} \right) = 0 \tag{1.36}$$

$$\frac{1}{2}(\partial^\nu \partial_\mu A^{a\mu} - \partial_\mu \partial^\mu A^{a\nu}) - \partial_\mu \partial^\nu A^{a\nu} = 0. \tag{1.37}$$

By applying the Lorentz gauge condition $\partial_\mu A^{a\mu} = 0$, equation 1.37 becomes

$$\Box A^{a\nu} = 0 \tag{1.38}$$

which in momentum space reads

$$-k^2 \tilde{A}^{a\nu}(k) = 0. \tag{1.39}$$

Based on equation 1.39 the gluon propagator after introducing Feynman $+i\epsilon$ prescription is given as

$$(-k^2 g^{\mu\nu}) D^{ab}_{\mu\nu}(k) = i\delta^{ab} \tag{1.40}$$

$$\Rightarrow D^{ab}_{\mu\nu}(k) = \frac{-i\delta^{ab} g_{\mu\nu}}{k^2 + i\epsilon}, \tag{1.41}$$

where a and b are the color indecies. The Feynman-rules factor for qqg vertex is given by Largangian term $\mathcal{L}_{qqg}$ (1.20)

$$-ig_s \gamma^\mu T^a. \tag{1.42}$$

The three gluon ggg vertex is determined by the Lagrangian term $\mathcal{L}_{3g}$ (1.20) in a form

$$\begin{aligned}
\mathcal{L}_{3g} &= -\frac{g_s f_{abc}}{4}(\underbrace{\partial_\mu A^a_\nu A^\mu_b A^\nu_c}_{1.\text{ term}} \underbrace{-\partial_\nu A^a_\mu A^\mu_b A^\nu_c}_{2.\text{ term}} \underbrace{+A^b_\mu A^c_\nu \partial^\mu A^\nu_a}_{3.\text{ term}} \underbrace{-A^b_\mu A^c_\nu \partial^\nu A^\mu_a}_{4.\text{ term}}) \\
&\left| \int (\partial_\mu A^a_\nu) A^\mu_b A^\nu_a = \underbrace{\int \partial_\mu (A^a_\nu A^\mu_b A^\nu_a)}_{=0} - \int A^a_\nu (\partial_\mu A^\mu_b A^\nu_c + A^\mu_b \partial_\mu A^\nu_c) \right| \\
&\Rightarrow \underbrace{\partial_\mu A^a_\nu A^\mu_b A^\nu_c}_{1.\text{ term}} = -A^a_\nu \partial_\mu A^\mu_b A^\nu_c - A^a_\nu A^\mu_b \partial_\mu A^\nu_c \\
&\Rightarrow \underbrace{-\partial_\nu A^a_\mu A^\mu_b A^\nu_c}_{2.\text{ term}} = A^a_\mu \partial_\nu A^\mu_b A^\nu_c + A^a_\mu A^\mu_b \partial_\nu A^\nu_c \\
&\Rightarrow \underbrace{A^b_\mu A^c_\nu \partial_\mu A^\nu_a}_{3.\text{ term}} = -\partial^\mu A^b_\mu A^c_\nu A^\nu_a - A^b_\mu (\partial^\mu A^c_\nu) A^\nu_a \\
&\Rightarrow \underbrace{-A^b_\mu A^c_\nu \partial_\nu A^\mu_a}_{4.\text{ term}} = \partial^\nu A^b_\mu A^c_\nu A^\mu_a + A^b_\mu (\partial^\nu A^c_\nu) A^\mu_a \\
&= -\frac{g_s f_{abc}}{4}(-A^a_\nu \partial_\mu A^\mu_b A^\nu_c - \partial^\mu A^b_\mu A^c_\nu A^\nu_a + A^a_\mu \partial_\nu A^\mu_b A^\nu_c + \partial^\nu A^b_\mu A^c_\nu A^\mu_a \\
&\qquad -A^a_\nu A^\mu_b \partial_\mu A^\nu_c - A^b_\mu (\partial^\mu A^c_\nu) A^\nu_a + A^a_\mu A^\mu_b \partial_\nu A^\nu_c + A^b_\mu (\partial^\nu A^c_\nu) A^\mu_a) \qquad (1.43)
\end{aligned}$$

Performing Fourier transformation into momentum space (p, q, r stands for gluon momenta) the Lagrangian $\mathcal{L}_{3g}$ simplifies to

$$\begin{aligned}
\mathcal{L}_{3g} &= -\frac{g_s f_{abc}}{4}(iq_\mu A^a_\nu A^\mu_b A^\nu_c + iq_\mu A^b_\mu A^c_\nu A^\nu_a - iq_\nu A^a_\mu A^\mu_b A^\nu_c - iq_\nu A^b_\mu A^c_\nu A^\mu_a \\
&\qquad +ir_\mu A^a_\nu A^\mu_b A^\nu_c + ir_\mu A^b_\mu A^c_\nu A^\nu_a - iq_\nu A^a_\mu A^\mu_b A^\nu_c - ir_\nu A^b_\mu A^c_\nu A^\mu_a) \\
&\quad \mid p+q+r=0 \mid \\
&= -\frac{g_s f_{abc}}{4}(\cancel{iq_\mu A^a_\nu A^\mu_b A^\nu_c} + \cancel{iq_\mu A^b_\mu A^c_\nu A^\nu_a} - iq_\nu A^a_\mu A^\mu_b A^\nu_c - iq_\nu A^b_\mu A^c_\nu A^\mu_a \\
&\qquad -ip_\mu A^a_\nu A^\mu_b A^\nu_c \cancel{-iq_\mu A^a_\nu A^\mu_b A^\nu_c} - ip_\mu A^b_\mu (A^c_\nu) A^\nu_a \cancel{+iq_\mu A^b_\mu A^c_\nu A^\nu_a} \\
&\qquad -iq_\nu A^a_\mu A^\mu_b A^\nu_c - ir_\nu A^b_\mu A^c_\nu A^\mu_a) \\
&= -\frac{ig_s f_{abc}}{4}([p-q]_\lambda g_{\mu\nu} A^\mu_a A^\nu_b A^\lambda_c + [q-r]_\mu g_{\nu\lambda} A^\mu_a A^\nu_b A^\lambda_c \\
&\qquad + [r-p]_\nu g_{\lambda\mu} A^\mu_a A^\nu_b A^\lambda_c \,). \qquad (1.44)
\end{aligned}$$

Based on the Equation 1.44 the factor for three gluon vertex is

$$g_s f_{abc} = ([p-q]_\lambda g_{\mu\nu} + [q-r]_\mu g_{\nu\lambda} + [r-p]_\nu g_{\lambda\mu}) \qquad (1.45)$$

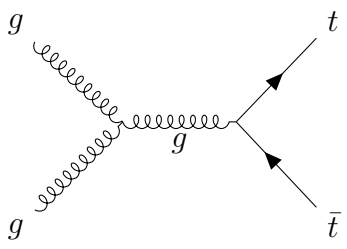
The four gluon vertex $4g$ is given by last Lagrangian term $\mathcal{L}_{4g}$ (1.22) as

$$-ig_s^2 \left[f^{abe} f^{cde} (g_{\mu\rho} g_{\nu\sigma} - g_{\mu\sigma} g_{\nu\rho}) + f^{ace} f^{bde} (g_{\mu\nu} g_{\rho\sigma} - g_{\mu\sigma} g_{\nu\rho}) + f^{ade} f^{bce} (g_{\mu\nu} g_{\rho\sigma} - g_{\mu\rho} g_{\nu\sigma}) \right]. \qquad (1.46)$$

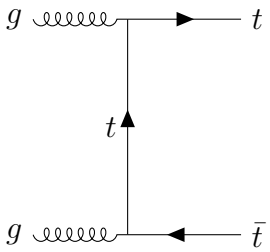
Due to gluon self-interactions, the strong force does not diminish with distance, but instead increases, preventing quarks and gluons from existing in isolation. This is known as color confinement, where only color-neutral hadrons are observable. At very short distances or high energies, the effective strong coupling α_s decreases, causing quarks and gluons to behave almost like free particles. This phenomenon is opposite to QED, where the coupling increases at shorter distances.

1.3.1 Top Quark Pair Production

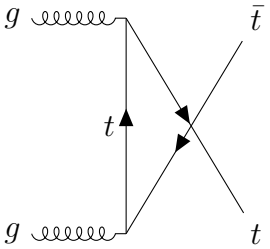
Top quark pairs are produced at leading order via gluon-gluon fusion and quark anti-quark annihilation, see Feynman diagrams in Figure 1.2.



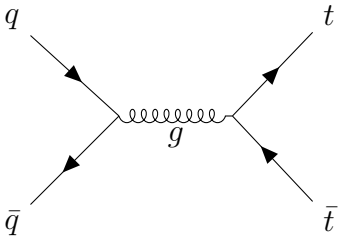
(a) $gg \to t\bar{t}$, s-channel.



(b) $gg \to t\bar{t}$, t-channel.



(c) $gg \to t\bar{t}$, u-channel.



(d) $q\bar{q} \to t\bar{t}$, s-channel.

Figure 1.2: Feynman diagrams of the production of $t\bar{t}$ pairs at the leading order. Generated using the TikZ-Feynman package [4].

At Tevatron, the proton-antiproton collider, top quark pairs were produced in 90% [7] via the quark anti-quark annihilation, due to high probability of colliding quark originating from proton with anti-quark of the same flavor originating from anti-quark and relatively high Bjorken-x (momentum fraction of the partons compare to hadrons), while at the LHC the gluon-gluon fusion is the dominant production mechanism (80% [7], $\sqrt{s} = 7$ TeV in pp collisions) due to low Bjorken-x and its large gluon density in this region.

In the rest of this section the calculation of Bjorken-x is shown to demonstrate its relatively low values in case of production of $t\bar{t}$ pairs in p+Pb collisions. By approximating partons to be massless compare to centre-of-mass energy (CME), assuming asymetric proton-lead collisions, the Bjorken-x's of two partons can be written as

$$
\begin{aligned}
s_{\text{hadr.}} &= (P_1 + P_2)^2 = (E_1 + E_2,\, 0,\, 0,\, E_1 - E_2)^2 = 4E_1E_2 && (1.47)\\
p_{\text{part.}} &= x \cdot E_{\text{hadr.}}(1,\, 0,\, 0,\, \pm 1) && (1.48)\\
s_{\text{part.}} &= (p_1 + p_2)^2 = (x_1E_1 + x_2E_2,\, 0,\, 0,\, x_1E_1 - x_2E_2)^2 = 4x_1x_2E_1E_2 && (1.49)\\
&= x_1x_2s_{\text{hadr.}} && (1.50)\\
\Rightarrow x_1 &= \frac{s_{\text{part.}}}{s_{\text{hadr.}}}\frac{1}{x_2}, && (1.51)
\end{aligned}
$$

where $E_{1,2}$ are the energies and $P_{1,2}$ four-momenta of colliding proton and lead ion, while $p_{1,2}$ is the four-momentum of the partons. Using pseudorapidity of $t\bar{t}$ system $\eta_{t\bar{t}}$

the x_2 can be expressed as a function of x_1:

$$\eta_{t\bar{t}} = \frac{1}{2}\log\left(\frac{E+p_z}{E-p_z}\right) = \frac{1}{2}\log\left(\frac{(E_1x_1+E_2x_2+E_1x_1-E_2x_2)}{(E_1x_1+E_2x_2-E_1x_1+E_2x_2)}\right) \quad (1.52)$$

$$= \frac{1}{2}\log\left(\frac{E_1x_1}{E_2x_2}\right) \quad (1.53)$$

$$\Rightarrow \frac{1}{x_2} = \frac{e^{2\eta_{t\bar{t}}}E_2}{x_1E_1}. \quad (1.54)$$

By inserting Eq. 1.54 to the Eq. 1.51 the formulae of Bjorken-x is given as

$$x_1 = \sqrt{\frac{s_{\text{part}}E_2}{s_{\text{hadr.}}E_1}}e^{\eta_{t\bar{t}}}. \quad (1.55)$$

In case of studied $t\bar{t}$ production at $\sqrt{s_{\text{NN}}} = 8.16$ TeV in proton-lead collisions the Bjorken-x ranges within the interval [0.003 - 0.517] as the ATLAS detector covers pseudorapidity $\eta \in [-2.5,2.5]$

$$x_1 = \frac{2m_t}{\sqrt{s_{\text{hadr.}}}}e^{\eta_{t\bar{t}}} = \frac{2\cdot 172.69}{8.16\cdot 10^3}e^{\{-2.5,-0.465,0,0.465,2.5\}} \simeq \{0.003,0.026,0.042,0.068,0.517\} \quad (1.56)$$

Given the Bjorken-x region interval [0.003 - 0.517] roughly corresponds to the range of the x-axis in Figure 1.3 shown in the publication [8].

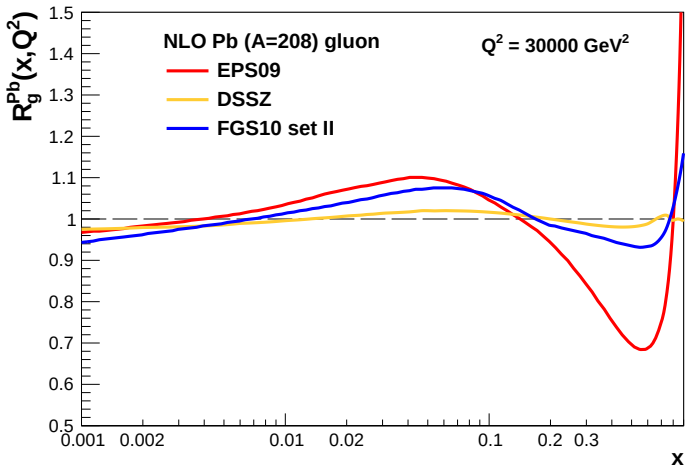


Figure 1.3: Ratio of the lead-to-proton gluon densities in the antishadowing ($x \approx 0.05-0.1$) and the EMC effect (named by the European Muon Collaboration) ($x \approx 0.1-0.6$) regions probed by $t\bar{t}$ production at virtualities $Q^2 = m_t^2 \approx 3\cdot 10^4$ GeV^2 in p+Pb collisions at the LHC, for three different NLO nuclear PDF sets: EPS09 [9], DSSZ [10], and FGS10 [11].

1.3.2 Leading Order Matrix Elements

This section follows the derivation of matrix elements and partonic cross-sections given in [12], however the equations were recalculated either by the author or using the Mathematica software [13]. Finally the leading order hadronic cross-section was integrated using own Python code given in Appendix B.

Quark Anti-quark annihilation

The matrix element for quark antiquark annihilation is given as

$$-iM_{q\bar{q}} = \bar{v}(k',s')c_j^\dagger(-ig_s\gamma^\mu T_a)u(k,s)c_i\frac{-ig_{\mu\nu}\delta_{ab}}{(k+k')^2}\bar{u}(p,s'')c_k^\dagger(-ig_s\gamma^\nu T_b)v(p',s''')c_l, \quad (1.57)$$

where

- c_i,$c_j^\dagger$,c_k,$c_l^\dagger$ are the color indices (for quarks and anti-quarks),
- $u(k,s)$ and $v(k',s')$ are the quark and anti-quark spinors, respectively, with momenta k and k', and spin states s and s',
- T^a,T^b are Gell-Mann matrices (generators of SU(3)),
- g_s is the strong coupling constant,
- $g_{\mu\nu}$ is the Minkowski metric tensor,
- δ^{ab} ensures color conservation between the gluons.

Its adjoint is

$$\begin{aligned} iM_{q\bar{q}}^\dagger &= -ig_s^2 c_l^\dagger v^\dagger(p',s''')(\gamma^\sigma)^\dagger T_b^\dagger (c_k^\dagger)^\dagger \bar{u}^\dagger(p,s'')\frac{g_{\sigma\lambda}\delta_{ab}}{s}c_i^\dagger u^\dagger(k,s)(\gamma^\lambda)^\dagger T_a^\dagger (c_j^\dagger)^\dagger \bar{v}^\dagger(k',s') \\ &= -ig_s^2 c_l v^\dagger(p',s''')\gamma^0\gamma^\sigma\gamma^0 T_b c_k \gamma^0 u(p,s'')\frac{g_{\sigma\lambda}\delta_{ab}}{s}c_i^\dagger u^\dagger(k,s)\gamma^0\gamma^\lambda\gamma^0 T_a c_j \gamma^0 v(k',s') \\ &= -i\frac{g_s^2\delta_{ab}}{s}(T_b)_{lk}(T_a)_{ij}\bar{v}(p',s''')\gamma^\sigma u(p,s'')\bar{u}(k,s)\gamma^\lambda v(k',s'). \quad (1.58) \end{aligned}$$

The color factor, arising from the trace over the color matrices, can be computed as

$$\begin{aligned} \frac{1}{N^2}\sum_{i,j,k,l,a,b}|M_{q\bar{q}}|^2 &\propto \frac{1}{N^2}\left(c_j^\dagger T_a c_i\right)\left(c_k^\dagger T_b c_l\right)\left(c_l^\dagger T_b^\dagger c_k\right)\left(c_i^\dagger T_a^\dagger c_j\right)\delta_{ab} \quad (1.59) \\ &= \frac{1}{N^2}(T_a)_{ji}(T_b)_{kl}(T_b)_{lk}(T_a)_{ij}\delta_{ab} \quad (1.60) \\ &= \frac{Tr(T_aT_a)^2}{N^2}\delta_{aa} = \frac{T_R^2}{N^2}\delta_{aa} \quad (1.61) \\ &= \frac{T_R^2}{N^2}(N^2-1) = \frac{2T_R^2 C_F}{N}. \quad (1.62) \end{aligned}$$

Using the completeness relations

$$\begin{aligned} \sum_s u(p,s)\bar{u}(p,s) &= \not{p}+m \quad (1.63) \\ \sum_s v(p,s)\bar{v}(p,s) &= \not{p}-m \quad (1.64) \end{aligned}$$

the matrix element becomes, after averaging over spins,

$$\begin{aligned} \frac{1}{4}\sum_{s,s',s'',s'''}|M_{qq}|^2 &\propto \frac{1}{4}\mathrm{Tr}(\not{k}'\gamma^\mu\not{k}\gamma^\lambda)\mathrm{Tr}((\not{p}+m_t)\gamma_\mu(\not{p}'-m_t)\gamma_\lambda) \\ &= \frac{1}{4}k'_\alpha k_\beta \mathrm{Tr}(\gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\lambda)(-m_t^2\mathrm{Tr}(\gamma_\mu\gamma_\lambda)+p_\alpha p'_\beta\mathrm{Tr}(\gamma_\alpha\gamma_\mu\gamma_\beta\gamma_\lambda)) \\ &= k'_\alpha k_\beta(g^{\alpha\mu}g^{\beta\lambda}-g^{\alpha\beta}g^{\mu\lambda}+g^{\alpha\lambda}g^{\mu\beta}) \\ &\quad (-4m_t g_{\mu\lambda}+4p^\alpha p'^\beta(g_{\alpha\mu}g_{\beta\lambda}-g_{\alpha\beta}g_{\mu\lambda}+g_{\alpha\lambda}g_{\mu\beta})) \\ &= (k'^\mu k^\lambda-(k'\cdot k)g^{\mu\lambda}+k'^\lambda k_\mu)(4m_t^2 g_{\mu\lambda}+4(p_\mu p'_\lambda-(p\cdot p')g_{\mu\lambda})+p_\lambda p'_\mu) \\ &= 8m_t^2(k\cdot k')+8(p\cdot k')(p'\cdot k)+8(k'\cdot p')(k\cdot p). \quad (1.65) \end{aligned}$$

Which can be expressed using Mandelstam variables as

$$\overline{|M_{qq}|^2} = \frac{g_s^4}{s^2} T_R^2 C_F \left(2m_t^4 + u^2 + t^2 + 2m_t^2(s-u-t)\right). \tag{1.66}$$

Gluon-Gluon Fusion

The gluon-gluon fusion matrix elements for each channel are given as

$$-iM_1 = \bar{u}(p',s')c_i^\dagger \left(-ig_s\gamma^\mu T_a\right) \frac{i(\not{p}' - \not{k}' + m_t)}{(p'-k')^2 - m_t^2} \left(-ig_s\gamma^\nu T_b\right) \bar{v}(p,s)c_j \epsilon_\nu^{(\lambda)}(k)\epsilon_\mu^{(\lambda')}(k') \tag{1.67}$$

$$-iM_2 = \bar{u}(p',s')c_i^\dagger \left(-ig_s\gamma^\nu T_b\right) \frac{i(\not{k}' - \not{p} + m_t)}{(k'-p)^2 - m_t^2} \left(-ig_s\gamma^\mu T_a\right) \bar{v}(p,s)c_j \epsilon_\nu^{(\lambda)}(k)\epsilon_\mu^{(\lambda')}(k') \tag{1.68}$$

$$-iM_3 = \bar{u}(p',s')c_i^\dagger \left(-ig_s\gamma^\rho T_d\right) \bar{v}(p,s)c_j \frac{-ig_{\rho\sigma}\delta_{cd}}{(k+k')^2} \left(-gf_{abc}\right) M_\sigma^{\nu\mu} \epsilon_\nu^{(\lambda)}(k)\epsilon_\mu^{(\lambda')}(k') \tag{1.69}$$

$$\begin{aligned} -iM_{\text{gh},1} &= \bar{u}(p',s')c_i^\dagger \left(-ig_s\gamma^\rho T_d\right) v(p,s)c_j \frac{-ig_{\lambda\sigma}\delta_{cd}}{(k+k')^2} g_s f_{abc}(-k^\sigma) \\ -iM_{\text{gh},2} &= \bar{u}(p',s')c_i^\dagger \left(-ig_s\gamma^\rho T_d\right) v(p,s)c_j \frac{-ig_{\lambda\sigma}\delta_{cd}}{(k+k')^2} g_s f_{abc}(-k'^\sigma) \end{aligned}$$

with color factors

$$\begin{aligned} \overline{|M_1|^2} &\propto \frac{1}{(N^2-1)^2}(T_a)_{ij}(T_a)_{jk}(T_b)_{kl}(T_b)_{li} \\ &= \frac{T_R^2}{(N^2-1)^2}\left(\delta_{ik}\delta_{jj} - \frac{1}{N}\delta_{ij}\delta_{jk}\right)\left(\delta_{ki}\delta_{ll} - \frac{1}{N}\delta_{kl}\delta_{li}\right) \\ &= \frac{T_R^2}{N(N^2-1)^2}\left(N^4 - 2N^2 + 1\right) = \frac{T_R^2}{N} \end{aligned} \tag{1.70}$$

$$\overline{|M_2|^2} \propto \frac{1}{(N^2-1)^2}\,\text{Tr}(T_bT_aT_aT_b) = \frac{T_R^2}{N} \tag{1.71}$$

$$\begin{aligned} \overline{|M_3|^2} &\propto \frac{1}{(N^2-1)^2}(T_c)_{ij}(T_e)_{ji}f_{abc}f_{abc} \\ &= \frac{T_RC_A}{N^2-1}\,\text{Tr}(T_cT_c)\delta_{cc} = \frac{T_RC_A}{N^2-1} \end{aligned} \tag{1.72}$$

$$\begin{aligned} \overline{M_1M_2^\dagger} &\propto \frac{1}{(N^2-1)^2}(T_aT_b)_{ij}(T_aT_b)_{ji} \\ &= \frac{1}{(N^2-1)^2}(T_a)_{ik}(T_a)_{jl}(T_b)_{kj}(T_b)_{li} \\ &= \frac{T_R^2}{(N^2-1)^2}\left(\delta_{il}\delta_{kj} - \frac{1}{N}\delta_{ik}\delta_{jl}\right)\left(\delta_{ki}\delta_{jl} - \frac{1}{N}\delta_{kj}\delta_{li}\right) \\ &= \frac{T_R^2}{N(N^2-1)}\left(1-N^2\right) = -\frac{T_R^2}{N(N^2-1)} \end{aligned} \tag{1.73}$$

$$\tag{1.74}$$

$$\begin{aligned}
\overline{M_1 M_3^\dagger} &\propto \frac{-i}{(N^2-1)^2}\,(T_aT_b)_{ij}(T_c)_{ij}f_{abc} = \frac{-i}{(N^2-1)^2}\,\mathrm{Tr}(T_aT_bT_c)f_{abc} \\
&= \frac{C_AT_R}{2(N^2-1)}\,\mathrm{Tr}(T_aT_a) = \frac{-iC_AT_R}{2(N^2-1)} \qquad (1.75)\\
\overline{M_2 M_3^\dagger} &\propto \frac{-i}{(N^2-1)^2}\,(T_bT_a)_{ij}(T_c)_{ji}f_{abc} = \frac{-i}{(N^2-1)^2}\,\mathrm{Tr}(T_bT_aT_c)f_{abc} \\
&= \frac{-i}{(N^2-1)^2}\,\mathrm{Tr}(T_bT_aT_c)(-f_{cba}) = \frac{-C_AT_R}{2(N^2-1)} \qquad (1.76)\\
\overline{|M_{\mathrm{gh},12}|^2} &\propto \frac{1}{(N^2-1)^2}\,(T_d)_{ij}f_{abc}(T_d)_{ji}f_{abc} = \frac{-i}{(N^2-1)^2}\,\mathrm{Tr}(T_dT_d)f_{abc}f_{abc} \\
&= \frac{-i}{(N^2-1)^2}\,T_RC_A\delta_{dd} = \frac{C_AT_R}{2(N^2-1)} \qquad (1.77)
\end{aligned}$$

and traces evaluated using Mathematica [13] and FeynCalc [14] packages (code in Appendix A)

$$\begin{aligned}
\overline{|M_1|^2} &\propto \frac{1}{4}\mathrm{Tr}\left((\not{p}'+m_t)\gamma^\mu(\not{p}'-\not{k}'+m_t)\gamma^\nu(\not{p}-m_t)\gamma_\nu(\not{p}'-\not{k}'+m_t)\gamma_\mu\right) \qquad (1.78)\\
\overline{|M_2|^2} &\propto \frac{1}{4}\mathrm{Tr}\left((\not{p}'+m_t)\gamma^\nu(\not{k}'-\not{p}+m_t)\gamma^\mu(\not{p}-m_t)\gamma_\mu(\not{k}'-\not{p}+m_t)\gamma_\nu\right) \qquad (1.79)\\
\overline{|M_3|^2} &\propto \frac{1}{4}\mathrm{Tr}\left((\not{p}'+m_t)\gamma^\sigma(\not{p}-m_t)\gamma^\delta\right)M_\sigma^{\nu\mu}M_{\nu\mu}^\delta \qquad (1.80)\\
\overline{M_1M_2^\dagger} &\propto \frac{1}{4}\mathrm{Tr}\left((\not{p}'+m_t)\gamma^\mu(\not{p}'-\not{k}'+m_t)\gamma^\nu(\not{p}-m_t)\gamma_\mu(\not{k}'-\not{p}+m_t)\gamma_\nu\right) \qquad (1.81)\\
\overline{M_1M_3^\dagger} &\propto \frac{1}{4}\mathrm{Tr}\left((\not{p}'+m_t)\gamma^\mu(\not{p}'-\not{k}'+m_t)\gamma^\nu(\not{p}-m_t)\gamma^\delta\right)M_{\nu\mu}^\delta \qquad (1.82)\\
\overline{M_2M_3^\dagger} &\propto \frac{1}{4}\mathrm{Tr}\left((\not{p}'+m_t)\gamma^\nu(\not{k}'-\not{p}+m_t)\gamma^\mu(\not{p}-m_t)\gamma^\delta\right)M_{\nu\mu}^\delta \qquad (1.83)\\
\overline{|M_{\mathrm{gh},1}|^2} &\propto \frac{1}{4}\mathrm{Tr}\left((\not{p}'+m_t)\not{k}(\not{p}-m_t)\not{k}\right) \qquad (1.84)\\
\overline{|M_{\mathrm{gh},2}|^2} &\propto \frac{1}{4}\mathrm{Tr}\left((\not{p}'+m_t)\not{k}'(\not{p}-m_t)\not{k}'\right) \qquad (1.85)
\end{aligned}$$

read

$$\begin{aligned}
\overline{|M_1|^2} &= -\frac{2g_s^4T_R^2}{N}\frac{m_t^4+m_t^2(3t+u)-tu}{(t-m_t^2)^2} \qquad (1.86)\\
\overline{|M_2|^2} &= -\frac{2g_s^4T_R^2}{N}\frac{m_t^4+m_t^2(t+3u)-tu}{(u-m_t^2)^2} \qquad (1.87)\\
\overline{|M_3|^2} &= -\frac{C_Ag_s^4T_R}{(N^2-1)}\frac{27m_t^4-19m_t^2(t+u)+4t^2+3tu+4u^2}{s^2} \qquad (1.88)\\
\overline{M_1M_2^\dagger} &= \frac{2g_s^4m_t^2T_R^2}{N(N^2-1)}\frac{2m_t^2+t+u}{(t-m_t^2)(u-m_t^2)} \qquad (1.89)\\
\overline{M_1M_3^\dagger} &= -\frac{C_Ag_s^4T_R}{(N^2-1)}\frac{3m_t^4-m_t^2(3t+u)+t^2}{s(t-m_t^2)} \qquad (1.90)\\
\overline{M_2M_3^\dagger} &= -\frac{C_Ag_s^4T_R}{(N^2-1)}\frac{3m_t^4-m_t^2(t+3u)+u^2}{s(u-m_t^2)} \qquad (1.91)\\
\overline{|M_{\mathrm{gh},1}|^2} = \overline{|M_{\mathrm{gh},2}|^2} &= \frac{C_Ag_s^4T_R}{2(N^2-1)}\frac{(m_t^2-t)(m_t^2-u)}{s^2}. \qquad (1.92)
\end{aligned}$$

Hadronic Cross-section at the Leading Order of QCD

Assuming accelerator with colliding particles in opposite direction the four vectors of incoming partons are given as

$$k = E\begin{pmatrix}1\\ \overrightarrow{e}_z\end{pmatrix}, \quad k' = E\begin{pmatrix}1\\ -\overrightarrow{e}_z\end{pmatrix}, \tag{1.93}$$

and outgoing quarks as

$$p = \begin{pmatrix}E\\ \sqrt{E^2-m_t^2}\sin(\theta)\\ 0\\ \sqrt{E^2-m_t^2}\cos(\theta)\end{pmatrix}, \quad p' = \begin{pmatrix}E\\ -\sqrt{E^2-m_t^2}\sin(\theta)\\ 0\\ -\sqrt{E^2-m_t^2}\cos(\theta)\end{pmatrix}. \tag{1.94}$$

The Mandelstam variables t and u can be expressed as a function of $\cos\theta$ and s

$$\begin{aligned} t &= (k-p)^2 = -2k\cdot p + m_t^2 = -2(E^2 - E\sqrt{E^2-m_t^2}\cos\theta) + m_t^2 \\ &= -\frac{s}{2} + m_t^2 + \cos\theta\sqrt{\frac{E^2}{4}(E^2-m_t^2)} = -\frac{s}{2} + m_t^2 + \cos\theta\sqrt{s\left(\frac{s}{4}-m_t^2\right)} \end{aligned} \tag{1.95}$$

and

$$\begin{aligned} u &= (k-p')^2 = -2k\cdot p' + m_t^2 = -2(E^2 + E\sqrt{E^2-m_t^2}\cos\theta) + m_t^2 \\ &= -\frac{s}{2} + m_t^2 - \cos\theta\sqrt{\frac{E^2}{4}(E^2-m_t^2)} = -\frac{s}{2} + m_t^2 - \cos\theta\sqrt{s\left(\frac{s}{4}-m_t^2\right)}. \end{aligned} \tag{1.96}$$

The partonic cross-section can be expressed as

$$\frac{d\sigma}{d\Omega} = \frac{\overline{|M|}^2\sqrt{1-\frac{4m_t^2}{s}}}{64\pi^2 s} \tag{1.97}$$

which for quark anti-quark annihilation and gluon fusion assuming $N=3$, $T_R=\frac{1}{2}$, $C_F=\frac{4}{3}$, $C_A=3$, $g_s=\sqrt{4\pi\alpha_s}$, $\beta=\sqrt{1-\frac{4m_t^2}{s}}$, and $\rho=\frac{4m_t^2}{s}$ gives

$$\frac{d\sigma_{qq}}{d\Omega} = \frac{\alpha_s^2\beta}{18s}\left(\cos^2(\theta)\beta^2+\rho+1\right) \tag{1.98}$$

$$\begin{aligned} \frac{d\sigma_{gg}}{d\Omega} &= \frac{\alpha_s^2\beta}{192s(1-\cos^2(\theta)\beta^2)^2}[-9\cos^6(\theta)\beta^6 - \cos^4(\theta)\beta^4(18\rho+7) \\ &+ 7\left(-2\rho^2+2\rho+1\right) + \cos^2(\theta)\left(18\rho^3-22\rho^2-5\rho+9\right)]. \end{aligned} \tag{1.99}$$

Integrating over $\cos\theta$ the partonic cross-sections become

$$\sigma_{qq} = \frac{4\pi\alpha_s^2\beta(\rho+2)}{27s} \tag{1.100}$$

$$\sigma_{gg} = \frac{\pi\alpha_s^2}{48s}\left((2\rho^2+32\rho+32)\,\mathrm{artanh}(\beta) - \beta(31\rho+28)\right) \tag{1.101}$$

which is identical to results in [15].

To obtain hadronic cross-section a convolution with PDFs, i.e. the integrals have to be evaluated

$$\sigma_{\mathrm{hadr}} = \int_{s_{\min}}^{S} ds \int_{s/S}^{1} dx_1 \int_{s_{\min}/Sx_1}^{1} dx_2 \;\; \delta(s-x_1x_2S) f_a(x_1,\mu_F) f_b(x_2,\mu_F)\sigma(s) \tag{1.102}$$

where parameters were set as $m_t = 172.5$ GeV, $\alpha_s \approx 0.122$ (taken from pdf at m_t), $s_{\min} = 4m_t^2$, $S = (8160 \text{ GeV})^2$, and parton distribution function CT18LO [16] was chosen. The Python [17] code for numerical integration (using packages Scipy [18], Numpy [19], and Pandas [20]) is given in Appendix B, providing hadronic cross-sections

$$\sigma_{qq,\,\text{hadr}} = 36.81 \pm 2.89 \text{ pb} \tag{1.103}$$

$$\sigma_{gg,\,\text{hadr}} = 138.74 \pm 5.77 \text{ pb} \tag{1.104}$$

$$\Rightarrow \sigma_{\text{LO, tot, hadr}} = 175.55 \pm 6.46 \text{ pb}, \tag{1.105}$$

where the uncertainties represent the precision of the numerical integration performed by the SciPy package. The result is in agreement with the Powheg [21] Monte Carlo generator (configuration file in Appendix C) which with the same parameters setup and pdfs CT18LO [16] and CT18NLO [16] provides

$$\sigma_{\text{Powheg, LO}} = 181.05 \ \pm \ 5.26 \text{ pb} \tag{1.106}$$

$$\sigma_{\text{Powheg, NLO}} = 228.99 \ \pm \ 3.34 \text{ pb}. \tag{1.107}$$

To obtain the next-to-next-to-leading (NNLO) correction with next-to-next-to-leading logarithmic (NNLL) accuracy, the k-factor

$$k \equiv \sigma_{\text{NNLO}}/\sigma_{\text{NLO}} = 271.99/230.41 = 1.180 \tag{1.108}$$

calculated using TOP++package [22] with pdf sets CT18NLO [16] and CT18NNLO [16] (configuration file in Appendix E) is applied also to other NLO predictions.

Focusing on proton-lead collisions, the MCFM [23] generator was used to calculate the NLO cross-section (configuration file in Appendix D) setting one colliding particle as lead and the other as proton with pdf sets CT18NLO [16] and EPPS21nlo_CT18Anlo_Pb208 nuclear pdf [24] resulting in

$$\sigma_{\text{pPb,MCFM, NLO}} = 254.81 \pm 0.28 \text{ pb}. \tag{1.109}$$

Multiplying by the k-factor and the number of nucleons within the lead $A = 208$, the next-to-next-to-leading cross-section of $t\bar{t}$ production in proton-lead collisions at $\sqrt{s_{\text{NN}}} = 8.16$ TeV is estimated

$$\sigma_{\text{pPb,MCFM, NNLO}} = 254.81 \cdot 1.180 \cdot 208 = 62.54 \text{ nb}. \tag{1.110}$$

1.4 Parton Shower, Hadronization, and Detector-Level Simulation

Event simulation proceeds through multiple stages, from the initial hard scattering of partons to the final interaction of particles with the detector. This process involves simulating the hard scattering, parton showers, hadronization, and detector response. Several computational tools are used in combination to achieve realistic event simulations. Key tools include the Powheg [21] generator for simulating the hard scattering at NLO, and Pythia [25] and Herwig [26] generators for modeling the parton shower and hadronization, and Geant4 [27] for simulating the interaction of particles with the detector.

1.4.1 Hard Scattering and NLO Corrections with Powheg

The simulation of particle collisions begins with the hard scattering process between partons (quarks or gluons) from the incoming protons. This hard process is typically calculated using perturbative QCD at NLO accuracy. Powheg (Positive Weight Hardest Emission Generator) is a tool designed to simulate the hard scattering and the hardest parton emission with NLO precision [21].

Powheg has several key advantages:

- NLO Accuracy: Powheg includes NLO order corrections, providing a more accurate cross-section calculation than leading-order (LO) generators.
- Hardest Emission: It generates the hardest parton emission in a manner consistent with NLO QCD, ensuring that the highest-energy radiation is modeled accurately.
- Positive Weights: Unlike some other NLO methods, Powheg guarantees positive weights for all events, simplifying event generation and analysis.

Once Powheg generates the hard scattering event and the hardest parton emission, the simulation proceeds to the next stage, where softer parton emissions are modeled.

1.4.2 Parton Shower and Hadronization

Following the hard scattering, the outgoing partons undergo a parton shower, where they radiate additional quarks and gluons, creating a cascade of lower-energy partons. This parton shower process is modeled by tools like Pythia or Herwig.

Pythia

Pythia simulates the parton shower by generating successive emissions of softer partons, followed by the hadronization process, where the partons combine into color-neutral hadrons such as pions, kaons, and protons. Pythia uses a leading-logarithmic approximation to handle soft and collinear QCD emissions, ensuring a realistic parton shower and hadronization [25].

Herwig

Alternatively, Herwig (Hadron Emission Reactions With Interfering Gluons) is another popular tool for simulating parton showers and hadronization. Herwig models the parton shower through angular-ordered emissions, which emphasizes the coherence of soft gluon radiation. Like Pythia, Herwig also handles the hadronization process using a cluster model, where color-singlet clusters of partons fragment into hadrons [26].

The combination of Powheg with either Pythia or Herwig ensures that the hard scattering is treated with NLO precision, while the parton shower and hadronization are handled by leading-logarithmic approximations. This provides a complete simulation of the event at the parton and hadron level, see illustration in Figure 1.4.

1.4.3 DGLAP Evolution

The Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations describe the behavior of partons as they radiate or absorb quarks and gluons, impacting the parton shower, they are given as

$$\frac{df_i(x,Q^2)}{d\ln Q^2} = \frac{\alpha_s(Q^2)}{2\pi}\sum_j \int_x^1 \frac{dz}{z} P_{ji}(z) f_j\left(\frac{x}{z},Q^2\right), \tag{1.111}$$

where Q^2 is the parton's virtuality, x stands for Bjorken-x, $f_i(x,Q^2)$ represents the parton distribution functions, $\alpha_s(Q^2)$ is the strong coupling constant, and $P_{ji}(z)$ are the splitting functions given by

$$P_{qq}(z) = C_F\left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\,\delta(1-z)\right] \tag{1.112}$$

$$P_{qg}(z) = C_F\left[\frac{1+(1-z)^2}{z}\right] \tag{1.113}$$

$$P_{gq}(z) = T_R\left[z^2+(1-z)^2\right] \tag{1.114}$$

$$P_{gg}(z) = 2\,C_A\left[\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z)\right] + \delta(1-z)\left(\frac{11}{6}C_A - \frac{2}{3}n_f T_R\right). \tag{1.115}$$

It can be understood as any parton i, identified in the parent hadron at the scale Q^2, could have originated from the branching of another parton j, resolved at scale $Q^2 + \mathrm{d}Q^2$. The transition from parton j to parton i is naturally associated with the production of an extra parton, ensuring momentum and flavor conservation. The extra particle is omitted for the PDF evolution calculation. In a Monte-Carlo event generator, it is considered as an extra final-state particle, and the production mechanism is referred as initial-state radiation (ISR) [28].

These equations play a crucial role in understanding how the parton shower evolves in both Pythia and Herwig.

1.4.4 Detector-Level Simulation with Geant4

Once the parton shower and hadronization processes are complete, the resulting final-state particles (hadrons, leptons, photons, etc.) are passed to a software toolkit for simulating the interaction of particles with the detector, see Figure 1.5.

Geant4 [27] simulates various physical processes as the particles pass through different layers of the detector:

- Tracing: Geant4 traces the passage and interaction of charged particles with the detector material, simulating energy loss, scattering, and ionization, etc.
- Shower Simulation: Geant4 handles both electromagnetic showers (caused by electrons and photons) and hadronic showers (caused by protons, neutrons, etc.) as particles interact with calorimeters and other detector elements.
- Signal Generation: Geant4 models the energy deposition and signal generation in detector components, which are then processed to simulate real experimental conditions.

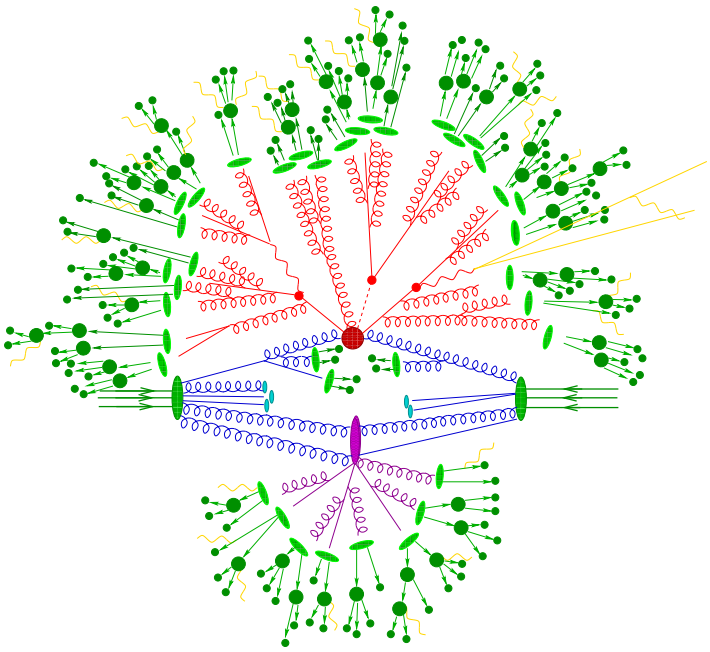


Figure 1.4: Sketch of a hadron-hadron collision as simulated by a Monte-Carlo event generator. The red blob in the center represents the hard collision, surrounded by a tree-like structure representing FSR as simulated by parton showers. The purple blob indicates a secondary hard scattering event. Parton-to-hadron transitions are represented by light green blobs, dark green blobs indicate hadron decays, while yellow lines signal soft photon radiation [28].

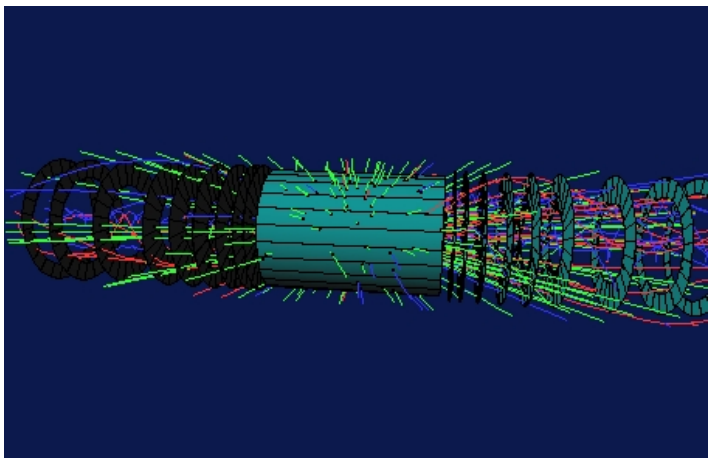


Figure 1.5: Example of a Geant4 Simulation of an event in the ATLAS Inner Detector [29].

By combining Powheg, Pythia (or Herwig), and Geant, the simulation pipeline provides a full event description from the hard scattering at NLO to the interactions with the detector, allowing for realistic comparisons between simulation and experimental data.

1.4.5 Top Quark Decay

The top quark is unique among quarks because it is the only quark that decays exclusively via the weak interaction to a W boson, and it decays before hadronizing

due to its extremely large mass and short lifetime. The top quark has a width of approximately $\Gamma_t \approx 1.35\,\text{GeV}$, corresponding to an exceptionally short lifetime of about $\tau_t \approx 0.5 \times 10^{-24}\,\text{seconds}$ [30]. The decay of a top quark (t) occurs predominantly via the channels

$$t \to W^+ b, \quad \bar{t} \to W^- \bar{b}$$

Each W-boson can decay either [30]

- Hadronically (into quarks): $\text{BR}(W \to q\bar{q}') = 67.6\%$;
- Leptonically (into leptons e,μ,τ): $\text{BR}(W \to \ell\nu_\ell) = 32.4\%$;
  which is split equally among the three lepton flavors:

$$\text{BR}(W \to e\nu_e) \approx 10.8\%, \quad \text{BR}(W \to \mu\nu_\mu) \approx 10.8\%, \quad \text{BR}(W \to \tau\nu_\tau) \approx 10.8\%.$$

All-Hadronic Channel

Both W-bosons decay hadronically (see Figure 1.6a). The branching ratio is

$$\text{BR}(\text{all-hadronic}) = \text{BR}(W \to q\bar{q}') \times \text{BR}(W \to q\bar{q}')$$

$$\text{BR}(\text{all-hadronic}) = 0.676 \times 0.676 = 0.457 \quad (45.7\%).$$

Lepton+Jets Channel

In this channel, one W-boson decays hadronically, and the other decays leptonically (see Figure 1.6b). The total branching ratio for the lepton+jets channel is

$$\text{BR}(\text{lepton+jets}) = 2 \times (\text{BR}(W \to q\bar{q}') \times \text{BR}(W \to \ell\nu_\ell))$$

$$\text{BR}(\text{lepton+jets}) = 2 \times (0.676 \times 0.324) = 0.438 \quad (43.8\%)$$

This can be further divided into each lepton flavor

$$\text{BR}(e + \text{jets}) = 2 \times (0.676 \times 0.108) = 0.146 \quad (14.6\%)$$

$$\text{BR}(\mu + \text{jets}) = 0.146 \quad (14.6\%)$$
$$\text{BR}(\tau + \text{jets}) = 0.146 \quad (14.6\%)$$

Dilepton Channel

In cases when both W-bosons decay leptonically (see Figure 1.6c). The branching ratio is

$$\text{BR}(\text{dilepton}) = \text{BR}(W \to \ell\nu_\ell) \times \text{BR}(W \to \ell\nu_\ell)$$

$$\text{BR}(\text{dilepton}) = 0.324 \times 0.324 = 0.105 \quad (10.5\%)$$

Splitted further into the different lepton combinations (e.g. $e+e$, $e+\mu$, $\mu+\tau$, etc.), the probabilities for each specific pair of leptonic decays are calculated.

$\Rightarrow$ Same-Flavored Lepton Pair (e.g. $e + e$, $\mu + \mu$, $\tau + \tau$)

The probability of both W-bosons decaying into the same lepton flavor is the product of the individual branching ratios for that flavor

$$\text{BR}(e+e) = \text{BR}(W \to e\nu_e) \times \text{BR}(W \to e\nu_e) = 0.108 \times 0.108 = 0.0117 \quad (1.2\%)$$

Similarly, for $\mu + \mu$ and $\tau + \tau$, each is also $\equiv 1.2\%$. They added to 3.6%.

$\Rightarrow$ Different-Flavored Lepton Pair (e.g. $e + \mu$, $e + \tau$, $\mu + \tau$)

The probability of two different lepton flavors decaying from the two W-bosons is the product of the branching ratio for each flavor. Since there are two possible ways to get a different pair (e.g. $W^+ \to e\nu_e$ and $W^- \to \mu\nu_\mu$, or $W^+ \to \mu\nu_\mu$ and $W^- \to e\nu_e$), we multiply the product by 2

$$\text{BR}(e+\mu) = 2 \times (0.108 \times 0.108) = 2 \times 0.0117 = 0.0234 \quad (2.3\%).$$

The same applies for other different-flavored pairs like $e + \tau$ and $\mu + \tau$, which are also 2.3% each,

$$\text{BR}(e+e) \approx 1.2\%, \quad \text{BR}(\mu+\mu) \approx 1.2\%, \quad \text{BR}(\tau+\tau) \approx 1.2\%$$

$$\text{BR}(e+\mu) \approx 2.3\%, \quad \text{BR}(e+\tau) \approx 2.3\%, \quad \text{BR}(\mu+\tau) \approx 2.3\%.$$

Figure 1.7a (left) shows a pie chart summarizing the branching ratios of $t\bar{t}$ decay, while Figure 1.7b on the right represents the branching ratios after the τ-leptons have decayed and their contributions have been redistributed into other channels.

1.4.6 Jets

Since quarks cannot exist freely in nature due to the color confinement, they undergo a process called hadronization shortly after their production in high-energy collisions. During hadronization, quarks transform into a collimated stream of hadrons, known as jets. This process is particularly relevant in the decay of the top quark. When a top quark decays, it produces a bottom quark associated with lighter quarks or leptons. These quarks subsequently hadronize, forming jets. Studying these jets provides crucial insights into the properties of the top quark and the dynamics of the strong interaction.

Several algorithms [32] are used to reconstruct jets from the particles observed in a detector. These algorithms can be broadly classified into three categories: cone algorithms, sequential recombination algorithms, and modern approaches that combine elements of both.

Cone Algorithms

- **Iterative Cone Algorithm**: Iteratively searches for stable cones of particles. A cone is defined by a radius parameter R in the rapidity-azimuth plane.
- **Seedless Infrared Safe Cone Algorithm (SISCone)**: An advanced cone algorithm designed to be infrared and collinear safe. An observable is infrared safe if it remains well-defined when the energy of emitted particles becomes small, such as soft gluon emission and collinear safe if it does not change when particles become collinear, meaning they move in nearly the same direction. These criteria ensure that the observable does not diverge, making the process calculable and physically meaningful while including higher-order QCD corrections.

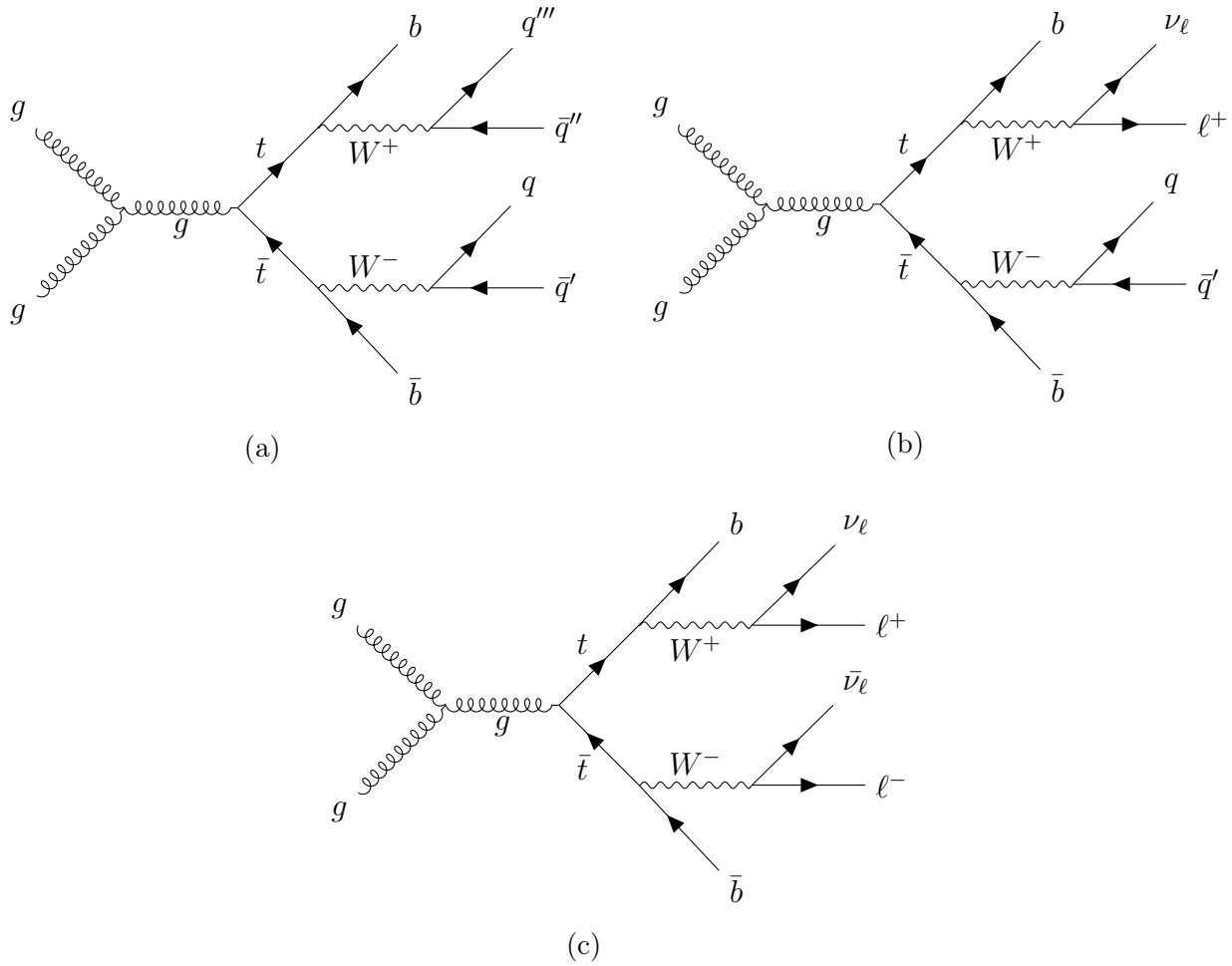


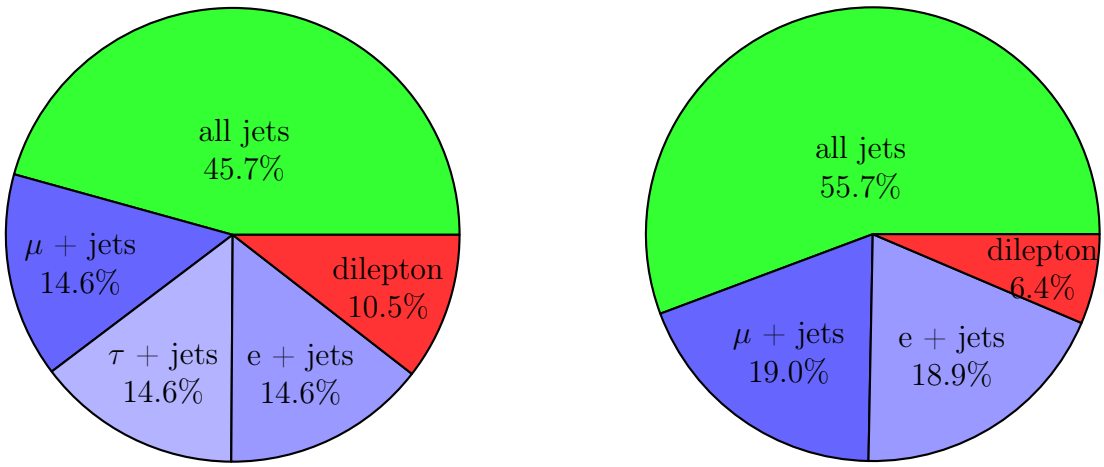
Figure 1.6: Feynman diagrams for $t\bar{t}$ production via gluon-gluon fusion in the s-channel, followed by different decay modes of the top and anti-top quarks. (a) Lepton+jets decay: one W boson decays leptonically ($W^+ \to \ell^+\nu_\ell$) and the other hadronically ($W^- \to q\bar{q}'$). (b) Dilepton decay: both W^+ and W^- bosons decay leptonically ($W^+ \to \ell^+\nu_\ell$ and $W^- \to \ell^-\bar{\nu}_\ell$). (c) All-hadronic decay: both W bosons decay hadronically ($W^+ \to q\bar{q}'$ and $W^- \to \bar{q}q'$). In all cases, the top and anti-top quarks decay to a W boson and a bottom quark ($t \to W^+b$, $\bar{t} \to W^-\bar{b}$). Generated using the TikZ-Feynman package [4].

Sequential Recombination Algorithms

- **The k_t Algorithm**: Merges pairs of particles based on their relative transverse momentum k_t. Pairs with the smallest k_t are merged first. This algorithm tends to produce less circular jets in the $\eta \times \phi$ plane.
- **Cambridge-Aachen (CA) Algorithm**: Similar to the k_t algorithm but uses angular separation instead of transverse momentum for merging.
- **Anti-k_t Algorithm**: Uses an inverse k_t measure, producing more regular, cone-like jets. It is the standard algorithm used in many high-energy physics experiments.

The measure used to judge what jet constituents to merge is given by

$$d_{ij} = \min(p_{T,i}^{2p}, p_{T,j}^{2p}) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = p_{T,i}^{2p}$$



(a) (b)

Figure 1.7: Decay channels perentage in $t\bar{t}$ production (a) with τ-leptons (b) with τ-leptons decayed (calculation using Python [17] code and Numpy [19] package in Appendix F). Generated using pdf-pie [31] package.

Where d_{ij} is the distance measure between particles i and j, $p_{T,i}$ is the transverse momentum of particle i, ΔR_{ij} is the distance in the rapidity-azimuth plane, and p is a parameter that determines the type of algorithm:

- $p = 1$ for the k_t algorithm.
- $p = 0$ for the CA algorithm.
- $p = -1$ for the anti-k_t algorithm.

Figure 1.8 presents jet shapes of four jet reconstruction algorithms.

Grooming Techniques

Jet grooming techniques, such as the "Soft Drop" and "Recursive Soft Drop" algorithms, are used to groom jets by removing soft radiation or pile-up. These techniques clean up the jet's internal structure, allowing for better analysis of its origin and properties.

Particle Flow Algorithms

Particle Flow algorithms combine information from different parts of the detector (e.g. calorimeters, trackers) to improve jet reconstruction accuracy.

1.4.7 Nuclear Parton Distribution Functions

Four distinct regions (shown in Figure 1.9) observed in nuclear parton distribution functions (nPDFs) correspond to well-known phenomena that modify the momentum distribution of quarks and gluons in nuclei, compared to free protons [35, 36].

- In the small $x < 0.01$ region, parton densities in nuclei are reduced relative to those of free nucleons. This phenomenon, known as shadowing, occurs because gluons from different nucleons overlap and interfere inside the nucleus. Shadowing effects are prominent in high-energy scattering processes, particularly in heavy nuclei.

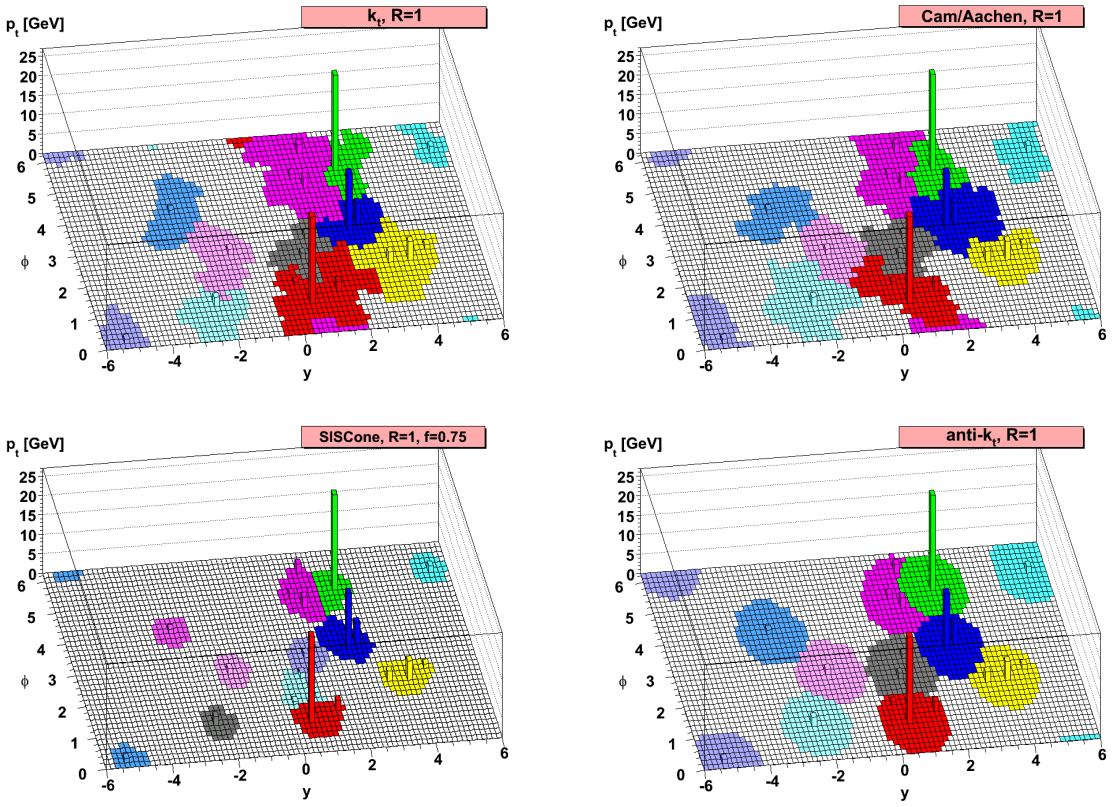


Figure 1.8: An example of parton-level event (generated with Herwig [33]), together with many random soft particles called "ghosts", clustered with four different jets algorithms, illustrating the "active" catchment areas of the resulting hard jets. For k_t and Cam/Aachen the detailed shapes are in part determined by the specific set of ghosts used, and change when the ghosts are modified. [34]

- At intermediate $0.1 < x < 0.3$, nuclear parton densities are enhanced compared to free nucleons. This compensatory effect, known as anti-shadowing, ensures the momentum sum rule for partons inside nucleons is satisfied. Anti-shadowing primarily affects gluons and sea quarks.

- The European Muon Collaboration (EMC) effect refers to the suppression of parton densities in the moderate $0.3 < x < 0.7$ region. The exact cause of the EMC effect remains a topic of debate, with several models attributing it to modifications in quark confinement and binding energy within bound nucleons.

- At high $x > 0.7$, parton distributions in nuclei are affected by the motion of nucleons within the nucleus, known as the Fermi motion. This shifts the parton momentum distribution to higher x, leading to an enhancement in parton densities, particularly for valence quarks.

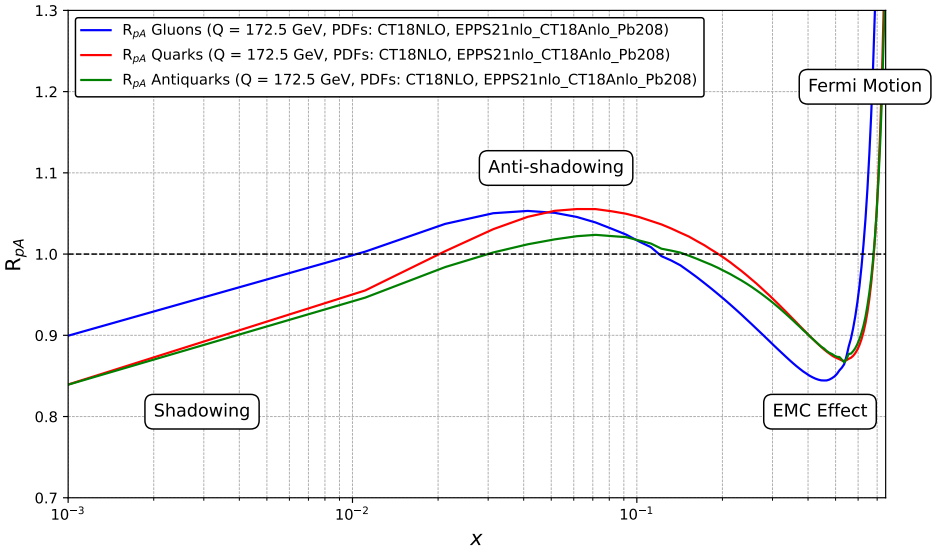


Figure 1.9: Nuclear modification factor R_{pA} of the parton distribution functions for gluons (blue), quarks (red), and antiquarks (green) as a function of the Bjorken x in Pb^{208} at a fixed scale $Q = 172.5$ GeV. The plot shows four distinct regions that reflect different nuclear effects: shadowing at low x, anti-shadowing at intermediate x, the EMC effect at moderate x, and Fermi motion at high x. The nuclear PDFs used are from the EPPS21nlo_CT18Anlo set, and the free proton PDFs are from CT18NLO. The solid horizontal line at $R_{pA} = 1$ indicates no nuclear modification.

Experimental Tools

2.1 Large Hadron Collider

The Large Hadron Collider at CERN (European Organization for Nuclear Research) is the world's largest and most energetic particle accelerator. Located beneath the ground at the Franco-Swiss border near Geneva, it represents a monumental achievement in particle physics. The LHC is designed to explore the fundamental forces and particles that constitute the universe. Operational since 2008, it has contributed significantly to our understanding of the universe's fundamental components, most notably through the discovery of the Higgs boson in 2012, confirming the Brout-Englert-Higgs (BEH) mechanism [37, 1], which explains how particles acquire mass.

The LHC is located in a circular tunnel with a circumference of 27 kilometers, located approximately 100 meters underground. It accelerates two beams of protons (or, alternatively, heavy ions like lead) to nearly the speed of light before colliding them at four primary interaction points of detectors: ATLAS, CMS, ALICE, and LHCb. The superconducting magnets used in the LHC bend and focus the beams. These magnets operate at temperatures close to absolute zero, around 1.85 K, utilizing liquid helium to maintain superconductivity. Currently the LHC achieves collision energy of 13.6 TeV [38, 39] with currently integrated luminostity 181 fb^{-1} [40].

2.1.1 Detectors

The LHC operates with four main detectors, each optimized for specific experiments:

- **ATLAS** (A Toroidal LHC ApparatuS): A general-purpose detector that investigates a wide range of phenomena, including searches for new particles beyond the Standard Model. It was integral to the discovery of the Higgs boson [3].
- **CMS** (Compact Muon Solenoid): Another general-purpose detector, CMS uses a different design approach than ATLAS. It contributed with ATLAS detector to confirm the discovery of the Higgs boson [41].
- **ALICE** (A Large Ion Collider Experiment): ALICE focuses on the study of heavy-ion collisions, such as lead-lead collisions. These collisions mimic the conditions of the early universe and allow scientists to study quark-gluon plasma, a state of matter that existed shortly after the Big Bang [38].
- **LHCb** (Large Hadron Collider beauty): LHCb is designed to study particles containing beauty (or b) quarks. Its research into matter-antimatter asymmetries aims to explain why the universe is predominantly made of matter despite theoretical predictions of equal amounts of matter and antimatter [38].

2.2 The ATLAS Detector

The ATLAS (A Toroidal LHC ApparatuS) detector(see Figure 2.1) is one of the two general-purpose detectors at the Large Hadron Collider, the other being the CMS detector. This section provides an overview of the structure, functionality, and purpose of each layer of the ATLAS detector.

The ATLAS detector is the largest volume particle detector ever constructed, measuring 46 meters in length, 25 meters in diameter, and weighing around 7,000 tonnes [42], it covers nearly the entire solid angle around the interaction point.

The ATLAS detector uses a right-handed coordinate system with its origin at the interaction point in the center of the cylindrical detector, and the z-axis aligned along the beamline. The x-axis points from interaction point to the center of the LHC ring, and the y-axis points upwards from interaction point. In the transverse plane, cylindrical coordinates (r,ϕ) are used, where ϕ represents the azimuthal angle around the z-axis. The pseudorapidity, η, is defined in relation to the polar angle θ by the expression $\eta = -\ln\tan(\theta/2)$. Angular distance is defined as $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

The detector comprises multiple layers of different subsystems, each designed to measure specific properties of particles produced in proton-proton collisions at the LHC. The overall structure of ATLAS can be divided into several key subsystems, which are described some more in detail below.

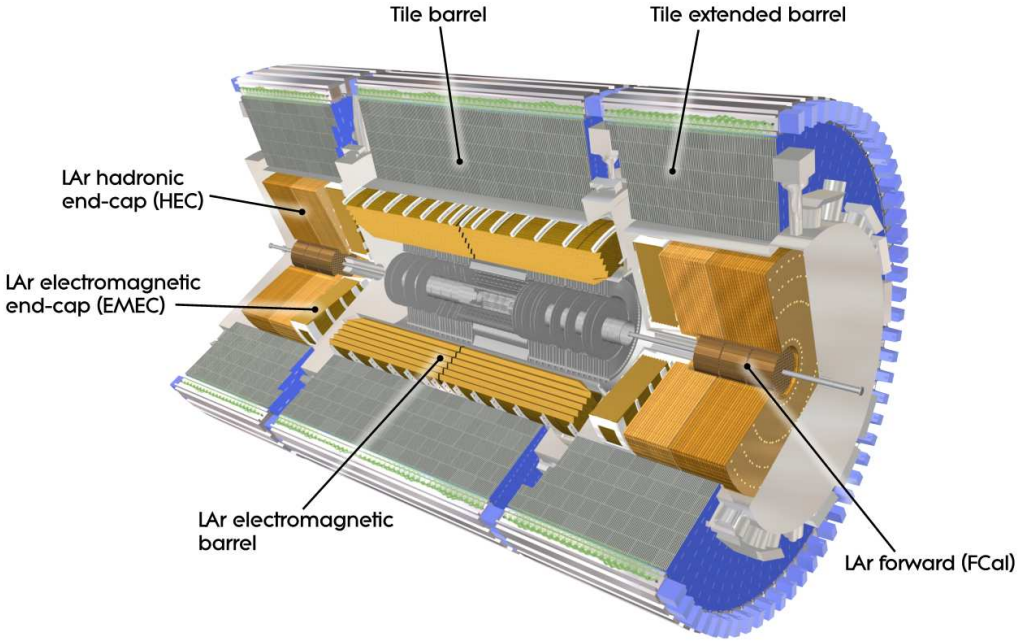


Figure 2.1: A cut-away drawing of the ATLAS inner detector and calorimeters. Tile Calorimeter consists of one barrel and two extended barrel sections and surrounds the Liquid Argon (LAr) barrel electromagnetic and endcap hadronic calorimeters. [43]

2.2.1 Inner Detector (ID)

The innermost region of the ATLAS detector is instrumented by the Inner Detector (ID), which is responsible for tracking the paths of charged particles as they propagate from the collision point. The ID is immersed in a magnetic field of 2 T produced by

a solenoid magnet, which curves the trajectories of charged particles, allowing their momentum to be measured. The ID consists of three primary subsystems:

- **Pixel Detector**: This is the layer closest to the beam pipe, located only a few centimeters away from the collision point. It is composed of millions of silicon pixels arranged in several layers. The pixel detector provides high-precision tracking information, especially important for identifying particle tracks near the collision point [44].
- **Semi-Conductor Tracker (SCT)**: Surrounding the pixel detector, the SCT consists of silicon strip detectors that provide additional tracking precision. It measures the position of charged particles with greater coverage and complements the pixel detector [42].
- **Transition Radiation Tracker (TRT)**: The TRT is the outermost part of the inner detector. It consists of thousands of thin straw tubes filled with gas, capable of detecting charged particles and distinguishing between electrons and heavier charged particles based on the radiation they emit when transitioning through different materials [45].

The ID is optimized for reconstructing particle trajectories with high spatial resolution, allowing physicists to precisely measure particle momentum and identify short-lived particles, such as the B hadrons via the reconstruction of secondary vertices.

2.2.2 Calorimeters

The next layer of the ATLAS detector is the system of calorimeters, which are designed to measure the energy of particles. There are two types of calorimeters in ATLAS:

- **Electromagnetic Calorimeter**: This system measures the energy of particles that interact electromagnetically, such as electrons and photons. It uses layers of lead and liquid argon to induce electromagnetic showers, where particles repeatedly emit energy in the form of secondary particles in an EM cascade. The energy of these showers is then measured, providing an accurate determination of the initial particle's energy [45].
- **Hadronic Calorimeter**: The hadronic calorimeter is located outside the electromagnetic calorimeter and is responsible for measuring the energy of hadrons (such as protons, neutrons, and pions). It consists of alternating layers of steel and scintillating tiles or liquid argon to absorb the hadrons' energy and measure the resulting particle shower energy [42].

The calorimeters play a crucial role in particle identification by differentiating between electrons, photons, and hadrons, and in measuring the energy carried by these particles.

2.2.3 Muon Spectrometer (MS)

The outermost and one of the most critical systems in the ATLAS detector is the Muon Spectrometer (MS). Muons are heavy, stable charged particles that can traverse large distances without losing much of energy through interactions with the inner layers of

the detector. Therefore, the MS is located at the outer edge of the detector, where it can accurately measure the properties of muons.

The MS is comprised of three types of detectors [46]:

- **Monitored Drift Tubes (MDT)**: These tubes are the primary tracking devices in the MS. They consist of wire chambers filled with gas that detect muons by measuring the ionization caused by the passing muon and recording the drift time of the resulting electrons.
- **Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC)**: These fast-response detectors are positioned in the endcap regions and are used to trigger muon detection in real-time. Their primary function is to signal that a muon has passed through and to provide coarse position measurements for triggering purposes.
- **Cathode Strip Chambers (CSC)**: These detectors are used in regions of high muon flux (near the beam pipe) due to their precision and ability to handle high particle rates.

The MS is designed to operate in a large toroidal magnetic field of B = 0.5–1 T generated by eight superconducting magnets that surround the detector. This magnetic field causes the muons to bend as they pass through, allowing their momentum to be determined accurately. The MS's ability to identify and measure muons with high precision is vital for many key physics analyses at ATLAS, such as searches for new particles and rare decay processes [42] which often result in final states which contain muons.

2.2.4 Magnetic Field Systems

Two main magnetic field systems are used in the ATLAS detector:

- **Central Solenoid Magnet**: A solenoid magnet is located around the inner detector and provides a magnetic field of B = 2 T that curves the paths of charged particles, enabling precise momentum measurements [45].
- **Toroidal Magnets**: eight large superconducting toroidal magnets generate a strong magnetic field of B = 0.5–1 T in the muon spectrometer. The toroidal field allows the MS to measure muon momentum with high accuracy over a wide range of particle energies [46].

The combination of these magnetic systems ensures that ATLAS can track particles with high precision over a large momentum range and measure their momenta effectively.

2.2.5 Trigger

Due to the high rate of proton collisions at the LHC (around 40 million per second), ATLAS cannot record every event. Instead, it uses a sophisticated Trigger System [42] to filter out events of interest, such as those potentially involving new physics or Higgs boson production. The trigger system consists of two levels:

- **Level-1 Trigger (L1)**: A hardware-based system that processes basic information from the calorimeters and muon spectrometer to make rapid decisions. The L1 rate is typically around 100 kHz.
- **High-Level Trigger (HLT)**: A software-based system that performs more detailed analysis of event data. The HLT rate is reduced to about 1 kHz, which is finally written to disk.

The combination of the L1 and HLT ensures that only a small fraction of events—those most likely to involve interesting or rare phenomena—are recorded for further analysis. This is crucial given the enormous volume of data produced by the LHC.

Measurement of $t\bar{t}$ production in lepton+jets and dilepton channels in p+Pb collisions

3.1 Motivation

The $t\bar{t}$ production process is a valuable probe of the heavy ion environment that has not yet been fully exploited. A measurement of $t\bar{t}$ production in lepton+jets and dilepton channels in proton-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV (see event display in Figure 3.1) could reduce the uncertainties in the gluon distribution at intermediate and small x, where the shadowing and antishadowing effects are most prominent are expected to play a major role in nuclear collisions.

Figure 3.2 demonstrates a significant reduction in the uncertainty on the gluon distribution after incorporating proton-lead pseudodata, especially in the momentum fraction range $10^{-3} < x < 10^{-1}$.

Even if the data consists of a single cross section, it can still provide meaningful constraints on the normalization of the gluon distribution in nuclei. This is especially important because gluon-gluon interactions dominate top quark pair production. The procedure of reweighting shows that such data can shift the central value of the gluon distribution and significantly tighten the uncertainty bands, improving the precision of nPDFs [47].

3.2 Data

The data utilized for this measurement were obtained using the ATLAS detector during the collision period involving proton-lead (p+Pb) interactions in 2016, accounting for an integrated luminosity of 165 nb^{-1}. The energy of the nominal proton beam and ion was 6.5 TeV, which leads in energy per nucleon of lead 2.56 TeV given by equation

$$E_{\mathrm{Pb}/A} \equiv \frac{E}{A} = \mathrm{E}_p \cdot \frac{Z}{A} = 6.5 \text{ TeV} \cdot \frac{82}{82+126} \doteq 2.56 \text{ TeV}, \tag{3.1}$$

where Z stands for number of protons and A for number of nucleons in lead ion, considering only lead isotop with 126 neutrons used by LHC. In total once proton collides with proton or neutron within the lead ion the center-of-mass energy per nuclean reads

$$\sqrt{s_{\mathrm{NN}}} = \sqrt{(P_p + P_{\mathrm{Pb}/A})^2} = \sqrt{(6.5+2.56)^2 - (6.5-2.56)^2} \doteq 8.16 \text{ TeV}. \tag{3.2}$$

Another interesting fact is that the collided system is boosted in the proton direction, since the energy of the proton 6.5 TeV is roughly two a half larger than energy of the

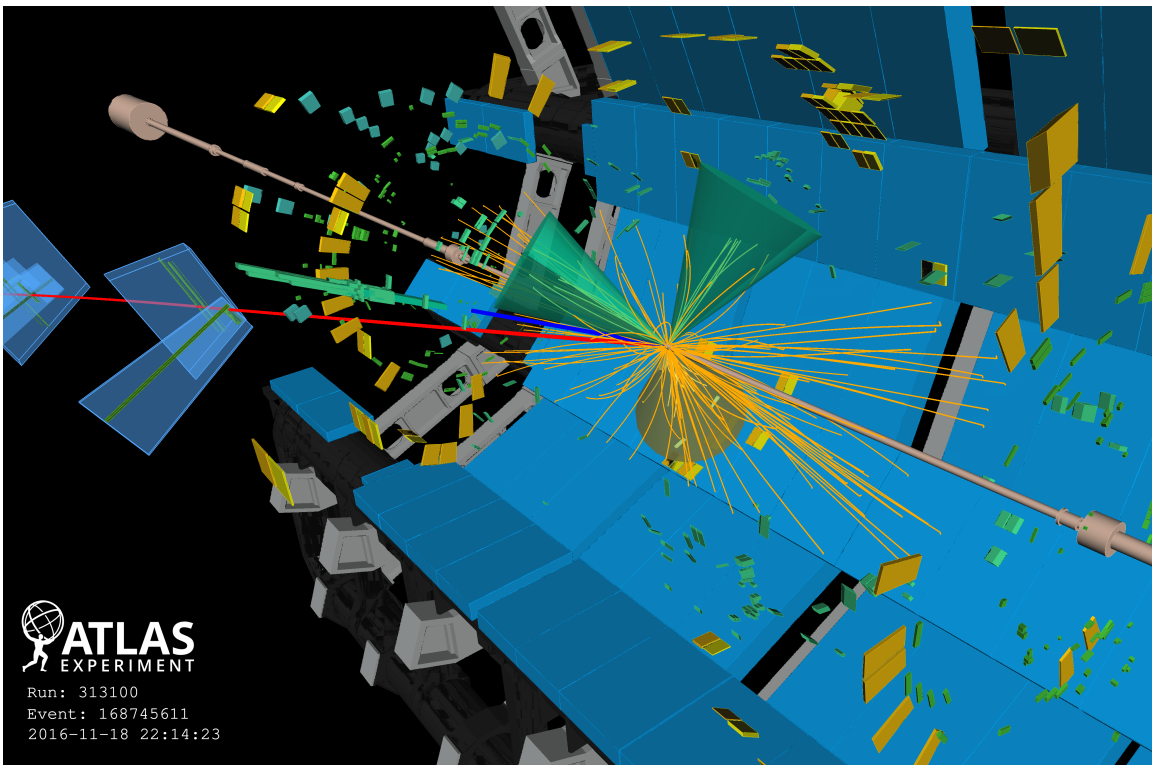


Figure 3.1: Event display of a candidate event for $t\bar{t}$ production in proton-lead ion collision decaying in the dilepton channel with the reconstructed electron track (blue line) and its associated energy deposit in the electromagnetic calorimeter (green boxes), the reconstructed muon track (red line) and its associated muon chambers (blue boxes; the bars are the related measurements), one b-tagged jet (yellow cone) and two non-b-tagged jets (green cones), and tracks in the inner detector (orange lines). Yellow rectangles correspond to energy deposits in cells of the hadron calorimeter [48].

proton/neutron in lead ion 2.56 TeV. Boost factor $\vec{\beta}$ given by the equation

$$|\vec{\beta}| = \left| \frac{\vec{p}_p + \vec{p}_{\mathrm{Pb}}}{E_p + E_{\mathrm{Pb}}} \right| = \frac{6.5 - 2.56}{6.5 + 2.56} \doteq 0.43 \tag{3.3}$$

shifts the rapidity distribution of outcomming objects by

$$y' = y \pm \operatorname{arctanh}(\beta) = y \pm \operatorname{arctanh}(0.43) \doteq y \pm 0.465. \tag{3.4}$$

The $\pm$ sign is given, because two beam configurations in the Monte Carlo (MC) simulations were considered – p+Pb and Pb+p. The effect is shown in Figure 3.3, where means of the lepton (electrons and muons) pseudorapidity distributions are shifted by the value $\beta = \pm 0.465$ in case of the $t\bar{t}$ signal MC (red and blue lines) with respect to the ATLAS laboratory frame. However, in data (black line) the shift is more flat comapre to signal MC and tends more to the Pb+p beam configuration, because about twice more luminosity in Pb+p (107.79 nb^{-1}) configuration than in p+Pb (56.76 nb^{-1}) were collected and also due to effects of other background contributions. Data were collected in a low-pileup environment, characterized by an average of $\mu_{\mathrm{PU}} = 0.18$ hadronic interactions per bunch crossing.

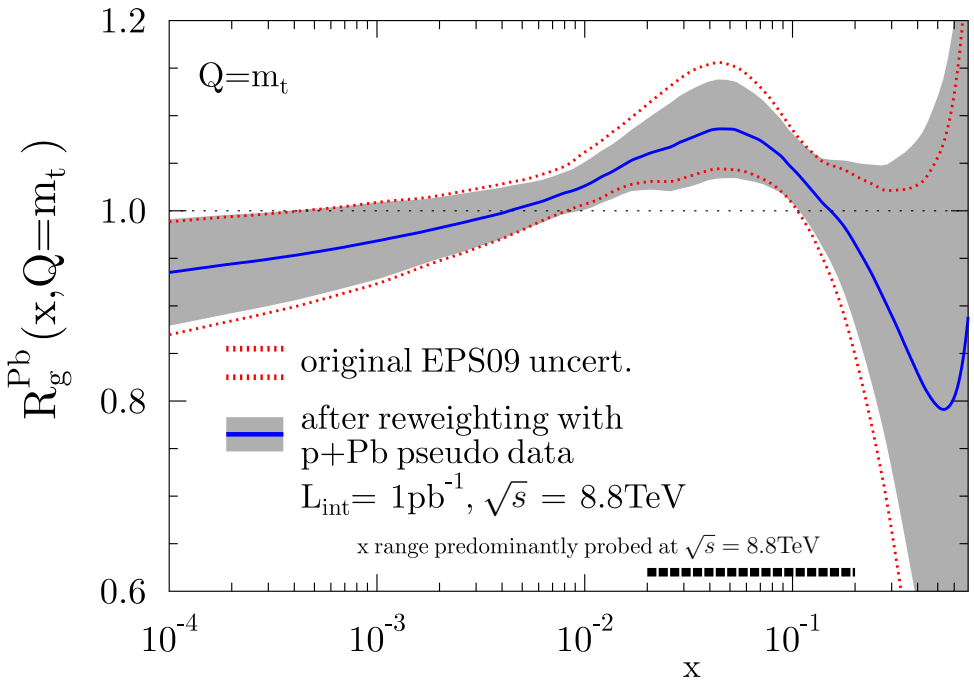


Figure 3.2: Ratio of nuclear-over-proton gluon densities, R_g^{Pb} evaluated at $Q = m_t$, for the p+Pb at $\sqrt{s_{NN}} = 8.8$ TeV at LHC for original EPS09 uncertainty (band enclosed by red dotted lines) and for the reweighted EPS09 using pseudodata (blue curve with grey band) [47].

3.2.1 Monte Carlo Simulation

Samples were generated at a center-of-mass energy of $\sqrt{s} = 8.16$ TeV with setups used in analyses at $\sqrt{s} = 13$ TeV. To account for isospin effects, two configurations—proton-proton (pp) and proton-neutron (pn) collisions were simulated.

The events for signal and background processes in individual isospin combinations were overlaid with real-data events collected during the 2016 p+Pb run, producing "data overlay" samples. This embedding technique is widely adopted in ATLAS heavy-ion measurements involving hard scattering processes, enabling accurate descriptions of the underlying event in p+Pb collisions. However, samples overlaid by data became challenge to process with Athena software and several modifications and development of the sotfware had to be introduced.

All simulated samples were normalized using the most accurate theoretical cross-sections and k-factor corrections. The MC events were processed through the full ATLAS detector simulation framework, based on Geant4. Given the negligible pileup in the 2016 p+Pb data, no additional pileup reweighting was applied. Furthermore, all samples underwent the same reconstruction algorithms and analysis chain as real data. Minor corrections were applied to the simulated lepton trigger and reconstruction efficiencies, derived from comparisons between data and simulation at $\sqrt{s} = 8.16$ TeV, to ensure better alignment with the actual detector response.

The normalization of MC samples involved scaling the pp samples by $Z_{\mathrm{Pb}} = 82$ and the pn samples by $N_{\mathrm{Pb}} = 126$. Additionally, the p+Pb and Pb+p beam configurations were scaled according to the integrated luminosity of the respective ATLAS data-taking campaigns.

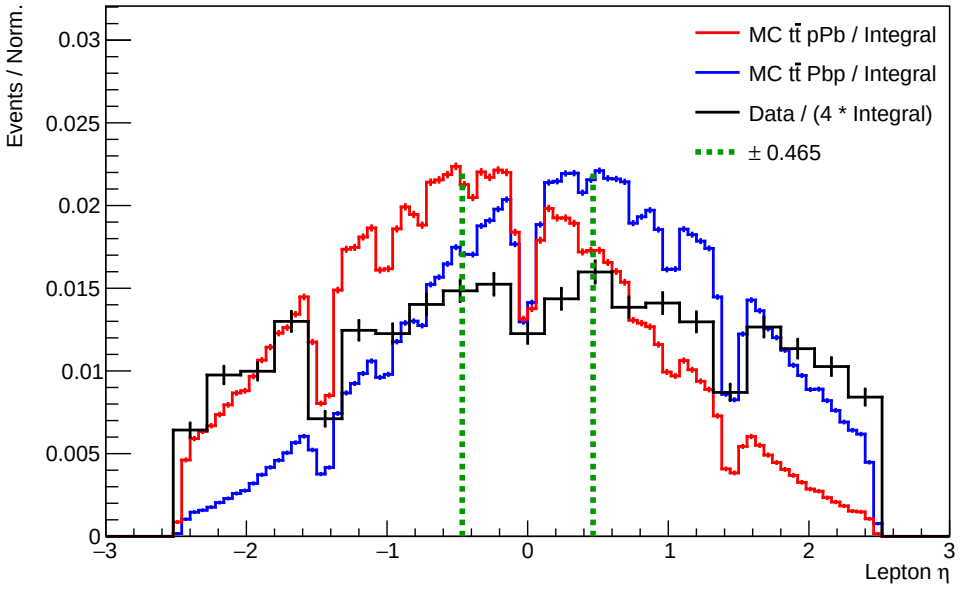


Figure 3.3: The lepton (electrons and muons) pseudorapidity distributions of data (black line) and signal $t\bar{t}$ in p+Pb (red line) and Pb+p (blue line) beam configurations. The rapidity shift of center-mass-system by value $y = \pm 0.465$ (green dashed lines) compare to the ATLAS laboratory system. Distributions are normalized, and on top of that data are scaled by value 0.25 due to the four times coarser binning compare to MC.

MC samples for $t\bar{t}$ signal

The nominal signal $t\bar{t}$ MC samples at NLO were generated using the PowHeg Box-v2 [49, 21]. The matrix element (ME) generator uses NNPDF3.0 [50] pdf with Pythia8 generator that simulates parton shower, fragmentation and the underlying event. The hard process factorisation scale μ_f and renormalisation scale μ_r were set to the default value: $\mu = (m_t^2 + p_{\mathrm{T,top}}^2)^{1/2}$, where $m_t = 172.5$ GeV and $p_{\mathrm{T,top}}$ are the top quark mass and transverse momentum, respectively, evaluated for the underlying Born configuration. The PowHeg model resummation damping parameter $h_{\mathrm{damp}} = 1.5 m_t^2$ controls matrix element to parton shower matching and effectively regulates the high-p_{T} radiation. The A14 tune [51] with NNPDF2.3LO pdf set was applied for Pythia8 showering. Samples in are generated in two isospin configurations for pp/pn interactions and two beam configurations p+Pb and Pb+p , see Table 3.1.

The nominal PowHeg $t\bar{t}$ MC samples are normalised to the total inclusive $t\bar{t}$ cross-section at a centre-of-mass energy of $\sqrt{s} = 8.16$ TeV calculated at NNLO in QCD including resummation of the next-to-next-to-leading logarithmic (NNLL) soft-gluon terms with the Top++ (v2.0) program [22]. It predicts the total cross-section $\sigma_{t\bar{t}}^{\mathrm{NNLO}} = 272.6^{+18.3}_{-18.9}$ pb with total uncertainties indicated and corresponding to a relative precision of $^{+6.9}_{-6.7}\%$ [52]. The pdf and α_S uncertainties were calculated following the PDF4LHC2015 [53] prescription with the MMHT2014 [54] and CT14 NNLO [55] pdf sets. The first uncertainty is calculated from the envelope of predictions with the QCD renormalisation and factorisation scales varied independently up or down by a factor of two from their default values of $\mu_r = \mu_f = m_t$, whilst never letting them differ by more than a factor of two. The second uncertainty is from variations in the PDF fol-

lowing the PDF4LHC2015 prescription using NNPDF3.0, MMHT, CT14, PDF4LHC pdf variations with nominal scale variation.

Cross-section values for other PDF sets recommended by the PDF4LHC are given in Appendix G.

nominal	Conf.	N. Evt	ME Gen.	PS Model	ME PDF	Tune
$t\bar{t}$	Pbp/pn	3888871	Powheg	Pythia8	NNPDF3.0 NLO	MMHT 2014 LO, CT14, PDF4LHC
$t\bar{t}$	pPb/np	2069897				
$t\bar{t}$	Pbp/pp	2496722				
$t\bar{t}$	pPb/pp	1376346				

Table 3.1: PowHeg +Pythia8 $t\bar{t}$ production with PowHeg h_{damp} parameter equal to $1.5m_t$, A14 tune, at least one lepton filter used for the nominal samples.

Alternative $t\bar{t}$ samples were generated to evaluate the modeling uncertainties associated with this process, as detailed in Table 3.2. These samples explore the effects of initial- and final-state radiation (ISR/FSR) and the impact of missing higher-order QCD corrections. This is achieved using an alternative PowHeg-v2 + Pythia8 sample, where the h_{damp} parameter is set to $3m_t$. Additionally, in the nominal samples, independent variations of the renormalisation and factorisation scales, as well as the effective strong coupling constant that controls ISR and FSR, are performed.

To compare with a different parton shower (PS) model, an alternative $t\bar{t}$ sample is produced using the same PowHeg Box configuration as the nominal sample, but with parton showering, hadronisation, underlying event (UE), and multiple parton interactions (MPI) simulated by Herwig7 [26]. Systematic uncertainties arising from the use of different hard-scattering generators, while keeping the PS model consistent, are assessed by generating a sample with Madgraph5_aMC@NLO [56], interfaced with Pythia8 using the A14 tune.

Alt.	Conf.	N. Evt	ME Gen.	PS Model	ME PDF	Tune
$3m_t$	pPb/np	1782708	Powheg	Pythia8	NNPDF 3.0 NLO	A14 NNPDF 2.3 LO
	Pbp/pn	3879231				
	pPb/pp	1386583				
	Pbp/pp	2387300				
$t\bar{t}$	pPb(Pbp)/pp	3139759	aMC @NLO	Pythia8		
	pPb(Pbp)/pp	6207050				
$1.5m_t$	pPb(Pbp)/pp	3169553	Powheg	Herwig7		H7.2-Default MMHT2014lo68cl
	pPb(Pbp)/pp	6266799				

Table 3.2: Alternative MC samples for the $t\bar{t}$ process used in systematic studies. At the time of production, the Herwig7 and aMC@NLO generators did not support the generation of pn isospin combinations.

V+jets background samples

Events containing W or Z bosons ($V = W/Z$) with associated jets, $V+ \geq 1b$ (BFilter), $V+ \geq 1c$ (CFilterBVeto), and $V+l$ (light-flavour jets, CvetoBVeto), were simulated using the Sherpa 2.2.10 [57] generator (see Feynman diagrams in Figure 3.4). The matrix elements are calculated for up to two jets at NLO and four jets at LO, utilizing the Comix [58] and OpenLoop [59] generators. These are then merged with the Sherpa parton shower [60] following the ME+PS@NLO prescription [61], using the NNPDF 3.0 NNLO PDF set and a dedicated parton shower developed by the Sherpa authors. The W/Z+jets events are normalized to the NNLO cross sections [62]. A summary of the W/Z+jet background MC samples can be found in Appendix H, Tables 5.2–5.4.

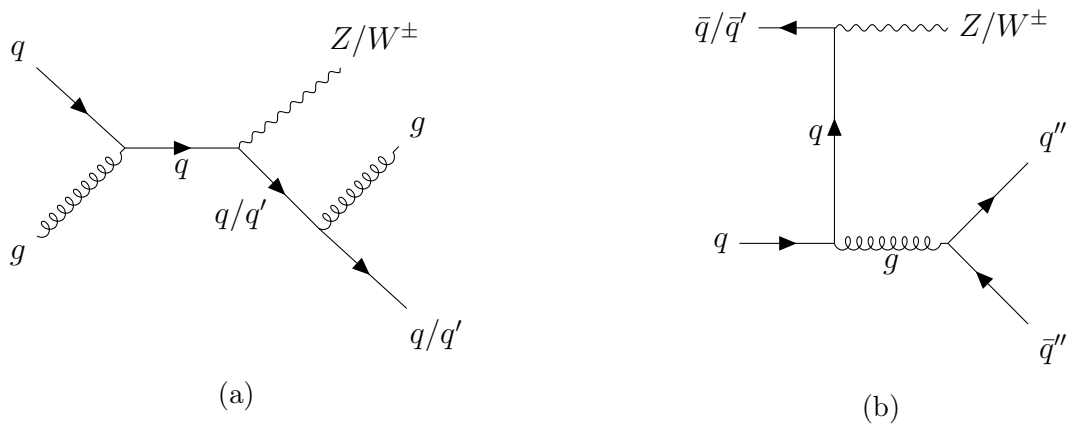


Figure 3.4: Examples of Feynman diagrams of V+jets background. Generated using the TikZ-Feynman package [4].

Signle top background samples

The single top quark production (see leading-order Feynman diagrams in Figure 3.5) is included as the $t\bar{t}$ background, because single top events can have the same detector signature of top quark pair production. Specifically, the Wt-channel at next-to-leading order (see in Figure 3.6) can produce a final state with a W boson and jets that is hard to distinguish from $t\bar{t}$ production. Although interference between Wt and $t\bar{t}$ production exists, current simulations do not typically account for it due to the complexity of modeling. Instead, the Wt-channel is generated separately, called diagram removal method, where matrix element is treated as

$$|M^{\text{removal}}_{Wt-\text{channel}}|^2 = |M_{Wt-\text{single top}}|^2 + \cancel{|M_{t\bar{t}}|^2} + \cancel{2\text{Re}(M_{Wt-\text{single top}}\, M^*_{t\bar{t}})}, \tag{3.5}$$

while as a systematic uncertainty the digram subtraction method is used by accounting also for an interference term

$$|M^{\text{subtraction}}_{Wt-\text{channel}}|^2 = |M_{Wt-\text{single top}}|^2 + \cancel{|M_{t\bar{t}}|^2} + 2\text{Re}(M_{Wt-\text{single top}}\, M^*_{t\bar{t}}). \tag{3.6}$$

The t-channel is another important background for $t\bar{t}$ production as it produces a top quark alongside light jets. On the other hand, the s-channel single top process has a relatively small cross-section and its contribution is negligible compared to the tW and t-channel processes. Therefore, the s-channel is excluded from simulations.

Single top and anti-top contributions in the tW-channel and t-channel were generated using the PowHeg-v2 matrix element (ME) generator [49, 21], interfaced with Pythia8 and the A14 tune. For these processes, the top quarks were decayed using MadSpin [63], ensuring that all spin correlations were preserved. The ME generator employed the NNPDF 3.0 NLO PDF set and was connected to Pythia8 [25], which handled the parton shower, fragmentation, and underlying event. The decays of bottom and charm hadrons were simulated using the EvtGen v1.2.0 program [64].

For the electroweak t-channel single top-quark events, the 4-flavour scheme was applied with the fixed four-flavour PDF4LHC PDF set for NLO matrix element calculations.

A total of eight samples were produced across two beam configurations (p+Pb and Pb+p) and two isospin combinations (pp and pn); see Table 3.3.

Process	Conf.	NEvt	ME Gen.	PS Model	ME PDF	Tune
Single top tW	pPb/pp	49539	Powheg	Pythia8	NNPDF 3.0 NLO	MMHT, CT14, PDF4LHC
Single top tW	Pbp/pp	99491				
Single (anti)top tW	pPb/pp	49539				
Single (anti)top tW	Pbp/pp	98496				
Single (anti)top tW	pPb/np	49539				
Single top tW	pPb/np	49539				
Single (anti)top tW	Pbp/np	99491				
Single top tW	Pbp/np	99491				
Single (anti)top t-chan.	pPb/np	49539	Powheg	Pythia8	NNPDF 3.0 NLO	MMHT, CT14, PDF4LHC
Single top t-chan.	pPb/np	49539				
Single (anti)top t-chan.	Pbp/np	99491				
Single top t-chan.	Pbp/np	99491				
Single (anti)top t-chan.	pPb/pp	49539				
Single (anti)top t-chan.	Pbp/pp	99491				
Single top t-chan.	pPb/pp	49539				
Single top t-chan.	Pbp/pp	99491				

Table 3.3: PowHeg + Pythia8 + EvtGen MC simulation of single top and anti-top quark processes in the tW and t-channels, used for generating the single top-quark samples.

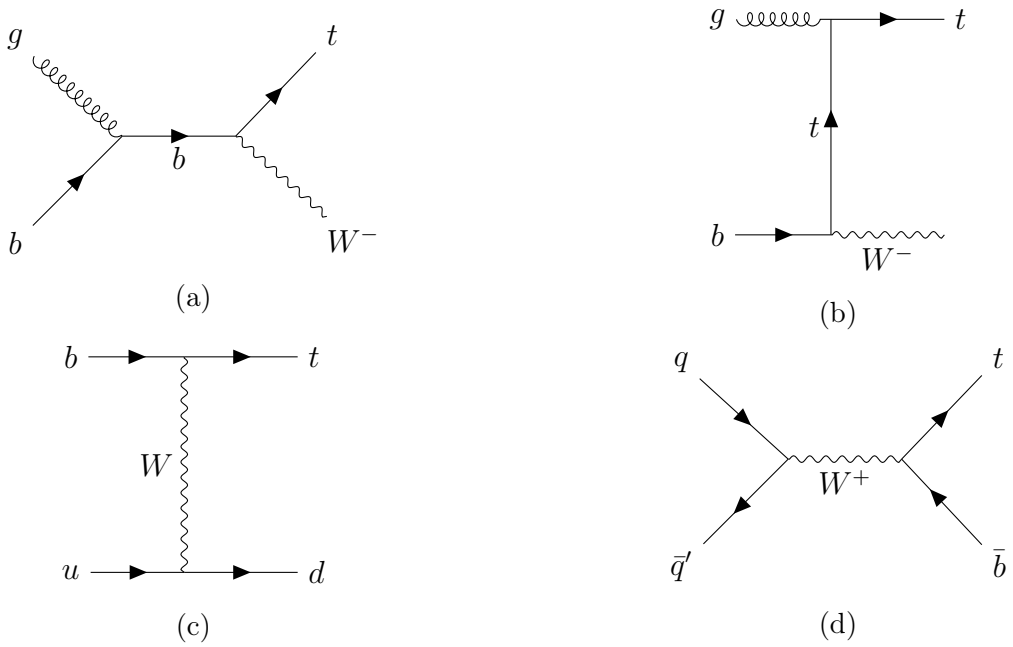


Figure 3.5: Leading order Feynman diagrams of single top background, (a)-(b) Wt-channel, (c) t-channel, and (d) s-channel. Generated using the TikZ-Feynman package [4].

VV+jets background samples

Diboson production can also lead to final states with the same objects as in the case of the $t\bar{t}$ production, see Figure 3.7. VV+jets background samples were simulated using Sherpa 2.2.11, with matrix elements calculated at NLO accuracy in QCD for up to one additional jet and at LO accuracy for up to three additional jets summarized in Table 3.4. The CT14 PDF set was chosen, along with dedicated parton shower tuning developed by the Sherpa authors.

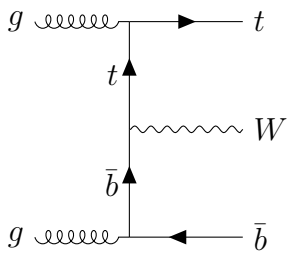


Figure 3.6: Example of next-to-leading order Feynman diagram of single top background Wt-channel. Generated using the TikZ-Feynman package [4].

Process	Conf.	NEvt	ME Gen.	PS Model	Tune
$VV \to \ell\ell\nu\nu$	Pbp/pp	49539	Sherpa-2.2.11	Sherpa-2.2.11	CT14
$VV \to \ell\ell\nu\nu$	pPb/pp	49539			
$VV \to \ell\ell\nu\nu$	Pbp/pn	49539			
$VV \to \ell\ell\nu\nu$	pPb/np	49539			

Table 3.4: V+jet samples for a dilepton channel with two charged leptons and two neutrinos.

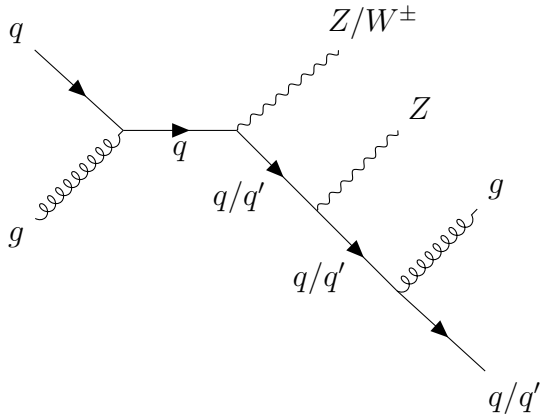


Figure 3.7: Example of Feynman diagram of VV+jets background. Generated using the TikZ-Feynman package [4].

Samples for performance studies

Table 3.5 provides a summary of $Z \to e^+e^-$, $Z \to \mu^+\mu^-$ MC samples simulated using PowHegwith showering Pythia8 using tune AZNLO CTEQ6L1 and Photos PDF sets. These samples are futher used for estimating lepton scale factors using tag-and-probe method.

Process	Conf.	NEvt	ME Gen.	PS Model	Tune
$Z \to e^+e^-$	pPb/pp	434533	Powheg	Pythia8	AZNLO CTEQ6L1
$Z \to e^+e^-$	Pbp/pp	804721			
$Z \to e^+e^-$	Pbp/pn	1233406			
$Z \to e^+e^-$	pPb/np	663363			
$Z \to \mu^+\mu^-$	pPb/pp	336597	Powheg	Pythia8	AZNLO CTEQ6L1
$Z \to \mu^+\mu^-$	Pbp/pp	626680			
$Z \to \mu^+\mu^-$	Pbp/pn	953827			
$Z \to \mu^+\mu^-$	pPb/np	514786			

Table 3.5: POWHEG +PYTHIA8 generated events of Z decays to leptons with overlay of data events.

3.3 Object Selection

In this section, detector object definitions with event pre-selection based on these objects are described in detail.

3.3.1 Electrons

Electron candidates are reconstructed from energy deposits in the central region of the electromagnetic calorimeter (EMC) and are associated with tracks reconstructed from the Inner Detector (ID). The selection of electron candidates is based on requirements:

- Candidates have to pass Medium likelihood-based working point [65] and have a transverse momentum $p_{\mathrm{T}} > 18$ GeV, absolute pseudorapidity $|\eta| < 2.47$, and a transverse energy $E_{\mathrm{T}} > 18$ GeV. The transition region of the EMC (crack region), defined as $1.37 < |\eta| < 1.52$, is excluded from this selection.
- Electron candidates must originate from the primary vertex, which imposes requirements on the transverse impact parameter significance: $|d_0|/\sigma_{d_0} < 5$, and the longitudinal impact parameter: $|\Delta z_0 \sin\theta| < 0.5$ mm.
- The Gradient isolation working point [65] is applied to suppress background contributions from QCD jets misidentified as prompt leptons (fake leptons) and from semileptonic decays of heavy-flavour hadrons (non-prompt real leptons).

In the EM calorimeter, the isolation variable is defined as the sum of the transverse energy of topo-clusters [65] within a cone of radius $\Delta R = 0.2$ around the electron cluster direction, divided by the electron p_{T}. In the ID, the isolation variable is defined as the sum of the p_{T} of tracks within a cone of variable size around the electron track, divided by the electron p_{T} ($\Delta R = \min[0.2, 10/p_{\mathrm{T}}]$).

3.3.2 Muons

Muon candidates were required to have a reconstructed track in the muon spectrometer (MS) combined with a track in the ID called "combined muon".

Muon candidates are selected with the following requirements:

- Candidates have to pass Medium ID criteria [66], with a transverse momentum of $p_{\mathrm{T}} > 18$ GeV, absolute pseudorapidity and $|\eta| < 2.5$.
- Muons are required to be associated to the primary vertex with the transverse impact parameter significance $|d_0|/\sigma_{d_0} < 3$, and on the longitudinal impact parameter of $|\Delta z_0 \sin\theta| < 0.5$ mm.
- To reduce the background of muons originating from heavy-flavour decays inside jets, muon candidates are required to be isolated using the track-quality and isolation criteria similar to those applied to electrons except the maximum cone size used in the ID isolation, which is 0.3 [66].

Lepton tracks have to pass criteria described in [67].

3.3.3 Jets

Jets are built using the anti-k_t algorithm [34] with a radius of $R = 0.4$, implemented through the FastJet package [68]. The algorithm utilizes a four-momentum recombination scheme and is applied using two different approaches.

PFlow jets

In the first approach, topo-clusters [65] serve as the inputs for jet reconstruction. These jets are calibrated to the hadronic energy scale, with p_{T} and η dependent correction factors derived from simulation. These jets are labelled as PF jets.

Heavy Ion jets

The second approach for jet reconstruction uses massless calorimeter towers [69] with a size of $\Delta\eta \times \Delta\phi = 0.1 \times \pi/32$ to construct jets, referred to as HI jets. To suppress the background from the underlying event (UE), a background subtraction is applied to each tower to reduce the rate of fake jets.

The average transverse energy density is estimated iteratively, with regions containing jets excluded from the estimation process. A detailed explanation of this procedure can be found in [69]. These jets are labelled as HI jets.

Two types of calibrations are applied to HI jets:

- EtaJES MC-based calibration: Corrects to the truth energy scale in Monte Carlo simulations as a function of jet η.
- HI/EMTopo cross-calibration and in-situ/η-intercalibration: Accounts for differences between Monte Carlo simulations and data.

The EtaJES correction has larger impact of the two corrections, and is derived using dijet events simulated by Pythia 8. The second calibration is applied exclusively to data to address discrepancies between simulation and real data.

The in-situ and η-intercalibrations provide residual corrections to the jet energy scale. The in-situ correction relies on the p_{T} balance between vector boson Z and jet events, while the η-intercalibration uses dijet events, where jets at mid-rapidity are used to calibrate jets in the forward rapidity region.

b-tagging

An important selection criteria for the analysis is the identification of jets containing b-quarks. PF jets containing B-hadrons are tagged using the Deep Learning DL1r algorithm [70].

The algorithm is trained on $t\bar{t}$ MC events, assigning b-jets as signal and a mixture of light-flavour and c-jets as background. Including c-jets in the training enhances c-jet rejection with minimal impact on light-jet rejection.

Jets are classified as b-tagged if the DL1r weight exceeds a fixed working point cut value of 1.27 corresponding to 85.43% b-tagging efficiency for b-jets in $t\bar{t}$ events. The corresponding rejection factors for charm quarks and light jets are 0.346 and 0.0248, respectively [71].

3.3.4 Missing Transverse Energy

Missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ (or MET) is a measure of the momentum carried by escaping neutrinos. It is calculated as

$$E_T^{\text{miss}} = \left| -\sum_i \vec{p}_{T,i} \right| \tag{3.7}$$

where $\vec{p}_{T,i}$ represents the transverse momentum of each reconstructed and calibrated physics object (e.g. electrons, muons, photons, hadronically decaying τ-leptons, small-R jets) within $|\eta| < 4.9$. A soft term is also included, comprising tracks associated with the primary vertex that are not matched to these objects.

The $E_{\text{T}}^{\text{miss}}$ is rebuilt in all events. The algorithm uses information from objects present in the event e.g. medium electrons and medium muons. Also PF jets are used in the $E_{\text{T}}^{\text{miss}}$ calculation. They are also used to derive systematic uncertainties associated with MET. Unfortunatelly, for technical reasons there is no MET calculation available for HI jets. Due to this limitation, the nominal analysis does not use selection on MET. However, the MET requirement is used to define a dedicated control region for fake-lepton background evaluation as explained in Section 3.5.

3.3.5 Jet Matching

Since the HI jet collection is not calibrated for b-tagging, both PF and HI jet collections must be used together. The matching process between HI jets and PF jets is based on two criteria. First, events are matched using a run and an event number criterion. Then, ΔR matching is applied to pair jets within the same event.

The b-tagging information is extracted from the PF jets and assigned to matched HI jets. Kinematic variables in the analysis are derived from HI jets, with the candidates required to have JES-corrected $p_{\text{T}} > 20$ GeV and $|\eta| < 2.5$. The same kinematic cuts are applied to PF jets.

The ΔR matching is performed by considering all possible combinations of HI jets and PF jets in an event. The distance between each pair of jets in pseudorapidity and azimuthal angle is calculated using the ΔR metric.

Jet pairs are then sorted based on the smallest ΔR, and the algorithm loops through the sorted pairs, matching jets and removing their indices. This results in $N_{\text{jets}} = \min(N_{\text{HIjets}}, N_{\text{PFjets}})$ matched jets. If a matched pair has $\Delta R < 0.3$ (chosen based on Figure 3.8) and the PFlow jet is b-tagged, the corresponding HI jet is also labelled as b-tagged.

Feature	**PF jets**	**HI jets**
b-tagging	b-tag info available	b-tag info not available
MET calculation	MET available	No MET available
Calibration	High-pileup (pp, 13 TeV)	Low-pileup (p+Pb, 8.16 TeV)
Jet energy scale	Affected by underlying events (UE)	Corrected by UE subtraction

Table 3.6: Comparison between PF and HI jets.

The jet matching algorithm combines the strengths of both PF and HI jet collections, but it introduces additional systematic uncertainties.

Varying ΔR

Two matching systematic uncertainties are evaluated by shifting $\Delta R = 0.3$ criteria for matched jets by ± 0.1. The effect on $H_{\text{T}}^{\ell,j}$ (the scalar sum of the transverse momenta of all jets and selected isolated leptons in the event) distributions is shown for one signal region $\geq$2b (defined in the Section 3.4 in Table 3.9) in Figures 5.3c–5.3d. The relative differences in yields of MC (background + signal) are within 2.5%. The effect on $H_{\text{T}}^{\ell,j}$ in other signal and cotrol regions are shown in Appendix I.

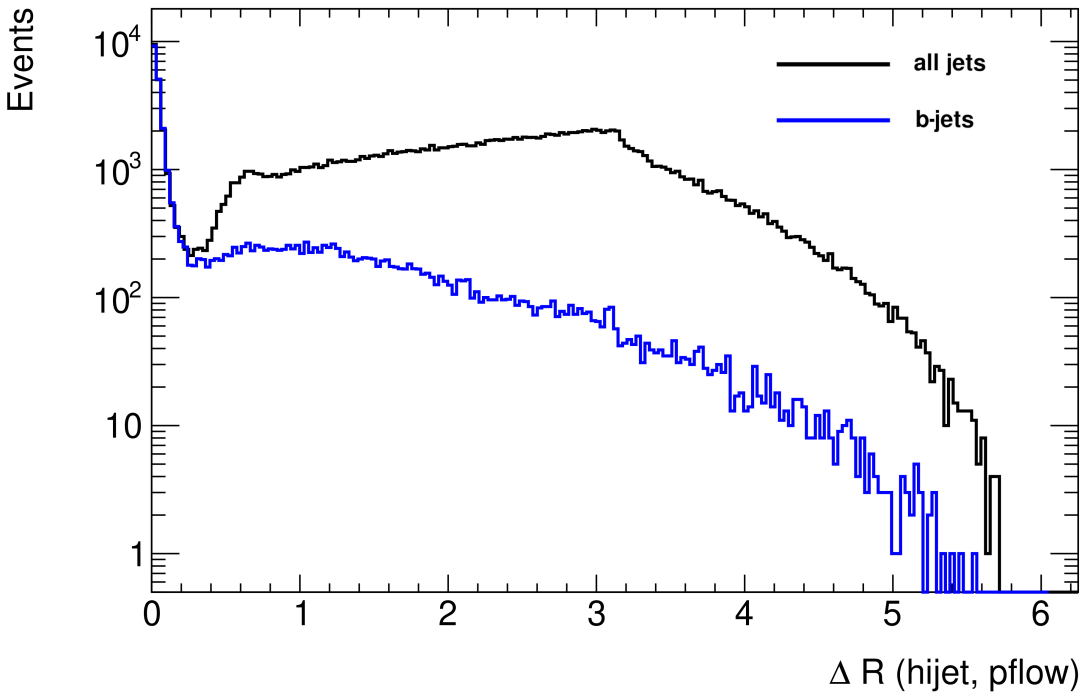


Figure 3.8: ΔR between HI and PF jets in signal MC samples. The black line represents all jets, while the blue line corresponds to the jets that have a b-tagged PF jet counterpart.

HI jets missing PF counterpart

In cases when number of HI jets is larger than PF jets, shown by bins above a diagonal in Figure 3.9, the HI jet is tagged as non-b-tagged in nominal samples. However, in principle such jet could have been b-tagged. The other extreme would be to assing all HI jets which are missing PF counterpart as b-tagged. More realistic approach is to assign HI jets which are missing PF counterpart randomly as b-jets based on light-flavour jet rejection function in Figure 3.10 as a function of jet p_{T}.

Tables 3.7–3.8 show an important cross-check of how often HI jets are missing its PF counterpart with respect to studied regions and consistency between data and MC. The effect on $H_{\mathrm{T}}^{\ell,j}$ distribution in one signal region $\geq$2b (defined in the following Section 3.4 in Table 3.9) is shown in Figure 5.3b and does not exceed relative difference of 10% to the nominal sample. The effect of $H_{\mathrm{T}}^{\ell,j}$ in other signal and cotrol regions are shown in Appendix I.

b-jet Region	Data	MC (background + signal)
≥ 0 b-jets	17.97% ($\pm$0.30)	14.89% ($\pm$0.29)
$= 1$ b-jet	14.80% ($\pm$0.58)	12.10% ($\pm$0.55)
≥ 2 b-jets	8.70% ($\pm$0.57)	6.35% ($\pm$0.53)

Table 3.7: Percentage of HI jets missing the PF counterpart caused by jet matching algorithm in data and MC (background + signal) samples for different b-jet regions.

b-jet Region	Data	MC (background + signal)
≥ 0 b-jets	-5.51% (± 0.33)	-5.77% (± 0.34)
$= 1$ b-jet	-7.30% (± 0.58)	-7.32% (± 0.57)
≥ 2 b-jets	-4.16% (± 0.39)	-4.54% (± 0.40)

Table 3.8: The decrease of the number of b-tagged jets caused by jet matching algorithm in data and MC (background + signal) samples for different b-jet regions.

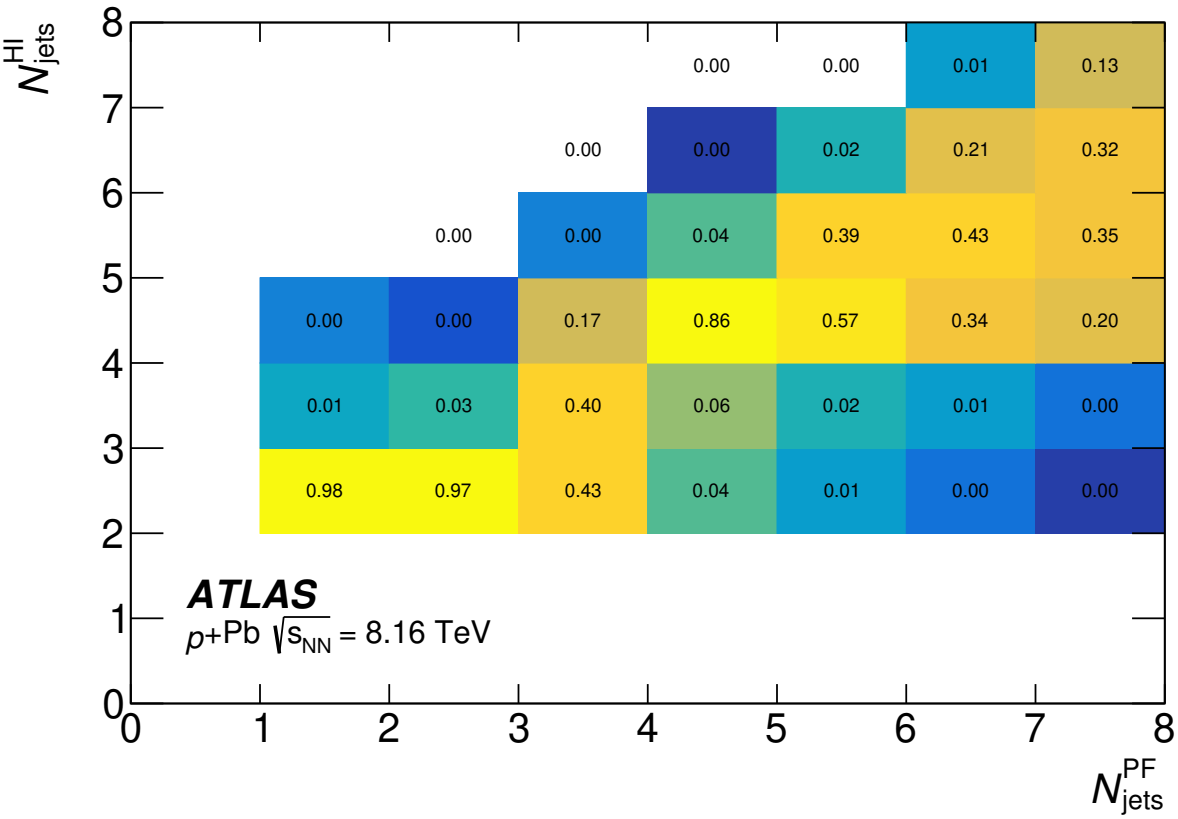
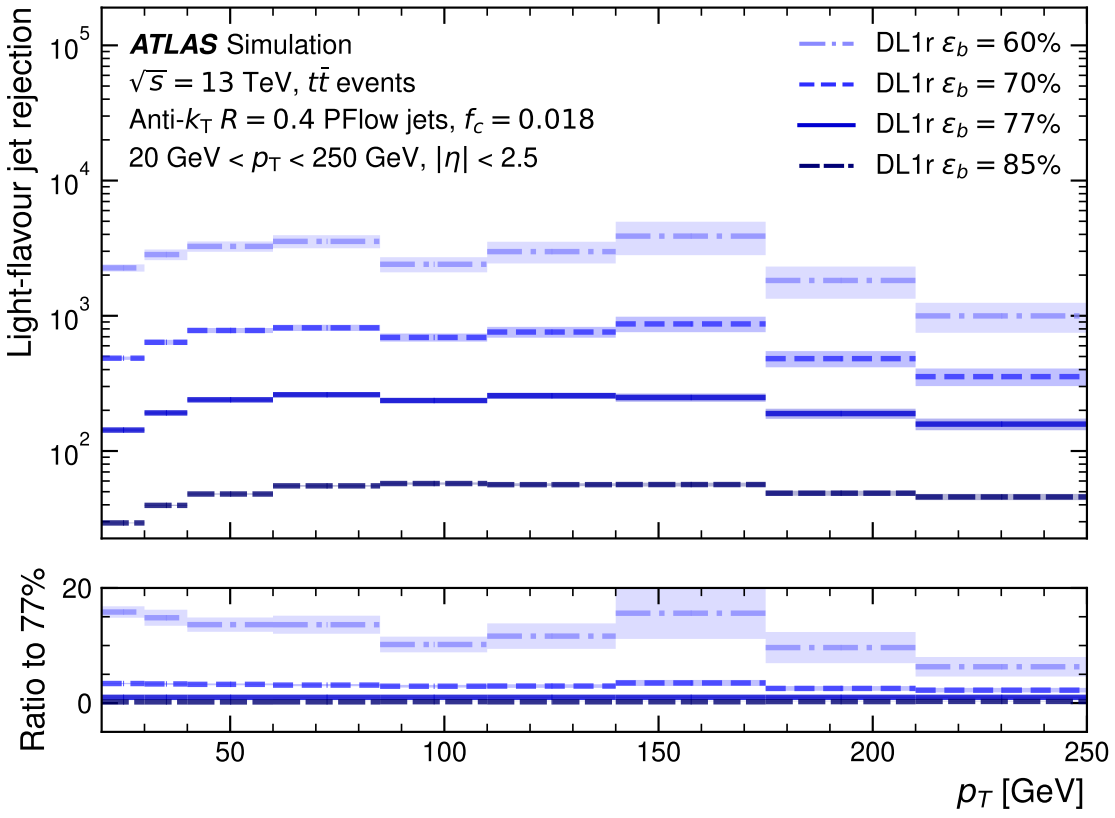
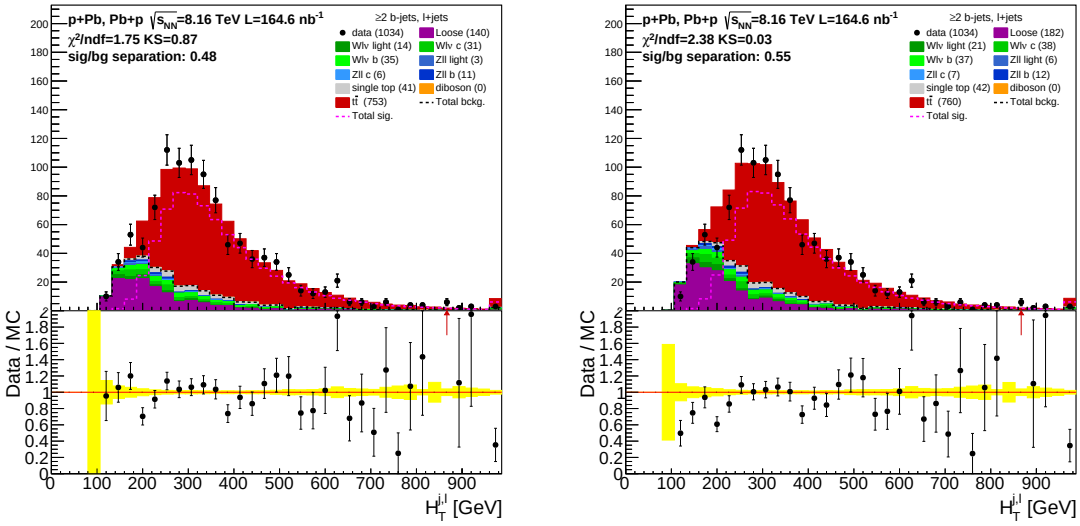


Figure 3.9: Number of particle flow (PF) jets versus number of heavy ion (HI) jets using jet matching algorithm normalised by columns in MC signal + background samples. Edge bins include overflow events [48].



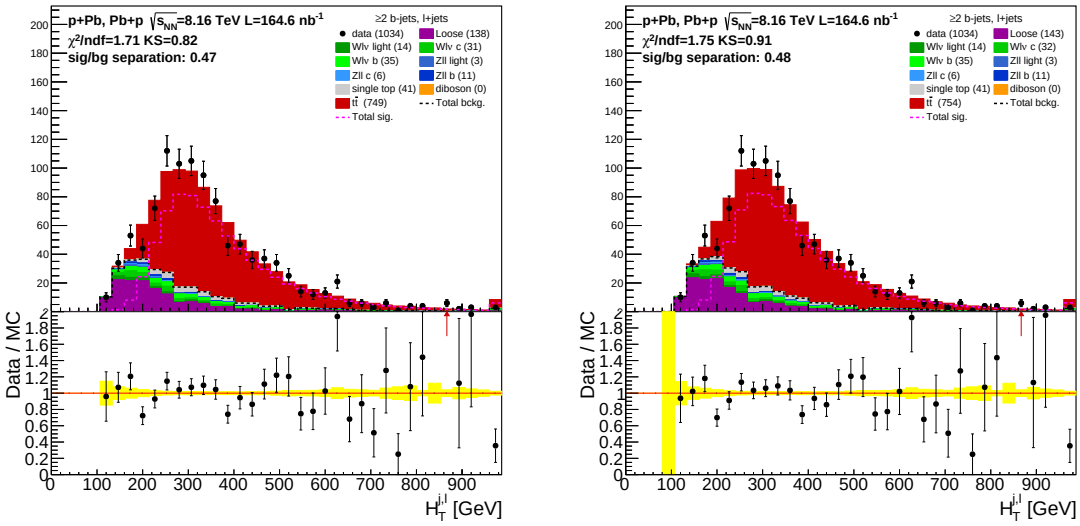
(a)

Figure 3.10: The background-jet rejection factors for different b-tagging efficiency ϵ_b operating points as a function of jet p_{T}. Plot shows the light-flavour b-tag jet rejection rate for a wide range of jet p_{T} bins and for efficiency operating points commonly used in ATLAS analyses of LHC Run 2 pp data. The lower panels show the ratio of each operating point's performance to that of the 77% operating point. The statistical uncertainties of the efficiency (rejection) are calculated using binomial uncertainties and are indicated as coloured bands [71].



(a)

(b)



(c)

(d)

Figure 3.11: Control plots of the $H_{\mathrm{T}}^{\ell,j}$ variable in the $>= 2b$ region (a) - default $\Delta R < 0.3$ between PF and HI jets for b-tagged matching, (b) systematic variation of assigning extra HI jets using rejection function, (c) systematic variation down $\Delta R < 0.2$, (d) systematic variation up $\Delta R < 0.4$. The yellow hatched band represents MC statistical uncertainties. The last bin includes overflow events.

3.3.6 Overlap Removal

After selecting objects described in previous sections (electrons, muons, and jets), the following points are applied to ensure the objects do not overlap:

- If a selected electron shares a track with a selected muon, the electron is removed.
- If an HI jet is within $\Delta R < 0.2$ of a reconstructed electron, the HI jet is removed.
- If an electron is within $\Delta R < 0.4$ of an HI jet, the electron is removed (it helps to reduce the impact of non-prompt leptons).
- If an HI jet has fewer than three tracks and is within $\Delta R < 0.4$ of a muon, the HI jet is removed. Conversely, the muon is removed if it is within $\Delta R < 0.4$ of an HI jet with three or more tracks.

3.4 Event Selection

The dilepton channel is defined by events containing exactly two opposite-sign leptons. Events with same-flavour lepton pairs (e^+e^- or $\mu^+\mu^-$) with an invariant mass ($m_{\ell\ell}$) within the Z boson mass window (80 GeV $< m_{\ell\ell} <$ 100 GeV) are excluded to suppress the Z boson background.

The invariant mass $m_{\ell\ell}$ has to exceed 15 GeV in the $e\mu$ channel and 45 GeV in the ee and $\mu\mu$ channels. This requirement ensures compatibility with the phase space of the Z+jets simulation samples while having minimal impact on the final results.

To define the signal region (SR) of the dilepton channel, events must also contain at least two HI jets, including at least one b-tagged jet. The signal region is further categorized based on the number of b-tagged jets: events with exactly one b-tagged jet are labelled as $2\ell 1b$, while those with two or more b-tagged jets are labelled as $2\ell 2b$incl.

In the ℓ+jets channel, events are selected if they contain exactly one lepton and at least four HI jets, including at least one b-tagged jet, to define the signal region. This signal region is subdivided based on the lepton flavour and the number of b-tagged jets: one electron or muon with exactly one b-tagged jet ($1\ell 1b$), and one electron or muon with two or more b-tagged jets ($1\ell 2b$incl).

Table 3.9 summarizes the six signal regions and three control regions (CR) which have the same selection as SR except the number of b-tagged jets is equal to zero.

3.5 Fake Background Estimation

The fake background is composed of non-prompt leptons and actual leptons faked in the detector by jets. Non-prompt leptons originating in decays of light-flavor hadrons or heavy-flavor hadrons, often found within jets.

The number of charge-flip electrons is estimated by reconstructing the charge of electrons/positrons and comparing it with the true generated charge. This type of background is particularly relevant in regions defined by two leptons with the same electric charge.

In the analysis, prompt lepton backgrounds originating from processes such as W and Z boson production, and $t\bar{t}$ production are also evaluated and subtracted. The fake lepton efficiencies for electrons and muons are measured for transverse momenta $p_T > 18$ GeV, within the pseudorapidity ranges $|\eta| \in (0,1.37) \cup (1.52,2.47)$ outside the detector's crack region[1].

Apart from charge misidentification, all the mentioned background contributions are measured using the Matrix Method (MM), a data-driven approach [72] that relies on experimental

[1]The crack region lies between the barrel ($|\eta| < 1.475$) and end-cap detectors ($1.375 < |\eta| < 1.52$).

Channel	SR Region Label	Selection Criteria
Dilepton	$2\ell 1b$	2 OS leptons, 1 b-tagged jet
Dilepton	$2\ell 2b$incl	2 OS leptons, $\geq$ 2 b-tagged jets
ℓ+jets (electron)	$1\ell 1b$ e+jets	1 electron, 1 b-tagged jet, $\geq$ 4 jets
ℓ+jets (electron)	$1\ell 2b$incl e+jets	1 electron, $\geq$ 2 b-tagged jets, $\geq$ 4 jets
ℓ+jets (muon)	$1\ell 1b$ μ+jets	1 muon, 1 b-tagged jet, $\geq$ 4 jets
ℓ+jets (muon)	$1\ell 2b$incl μ+jets	1 muon, $\geq$ 2 b-tagged jets, $\geq$ 4 jets
	CR Region Label	
Dilepton	$2\ell 0b$	2 OS leptons, 0 b-tagged jet
ℓ+jets (electron)	$1\ell 0b$ e+jets	1 electron, 0 b-tagged jet, $\geq$ 4 jets
ℓ+jets (muon)	$1\ell 0b$ μ+jets	1 muon, 0 b-tagged jet, $\geq$ 4 jets

Table 3.9: Summary of signal (SR) and control (CR) regions for the dilepton and ℓ+jets channels based on the number of leptons and b-tagged jets. The OS stands for opposite charged leptons.

data and MC simulations. Two key inputs for this method are the efficiencies of real and fake leptons.

3.5.1 Real and Fake Lepton Efficiencies

The approach to estimate the normalization and shapes of the fake background contributions, based on lepton isolation information, is the Matrix Method (MM). This technique relies on defining two distinct working points for lepton isolation and identification.

The data sample containing N events is divided into two categories. The first consists of leptons that pass the tight selection , denoted as T, and the second category includes leptons passing the loose selection but not fulfilling the tight requirements, denoted as L. A summary of these definitions is provided in Table 3.10.

Electrons		
Property	Loose	Tight
Electron identification level	LooseAndBLayerLH	MediumLH
Electron isolation working point	None	Gradient
Minimal electron p_T in [GeV]	18	18
Maximal electron d_0 in [mm]	5	5
Maximal electron z_0 in [mm]	0.5	0.5
Muons		
Property	Loose	Tight
Muon identification level	Medium	Medium
Muon isolation working point	None	FCTight_FixedRad
Minimal Muon p_T in [GeV]	18	18
Maximal muon d_0 in [mm]	7	3
Maximal muon z_0 in [mm]	0.5	0.5

Table 3.10: Definitions of loose and tight leptons.

- LooseAndBLayerLH refers to a loose selection criterion, requiring hits in the innermost pixel detector layer (B-layer), along with a likelihood-based identification (LH) method, which uses multiple variables to distinguish real electrons from background [65].

- MediumLH working point applies stricter selection criteria than the loose one, improving the purity of selected electrons by applying more stringent requirements in the likelihood-based identification [73].

- None means no isolation criteria are applied, allowing for electrons that may be closer to other objects such as jets. These electrons are more likely to include fakes or non-prompt electrons.

- Gradient is a dynamic isolation criterion that adjusts based on the electron's transverse momentum (p_T). Lower p_T electrons are required to be more isolated in terms of surrounding track activity, while higher p_T electrons can be less isolated. This technique finds a balance between signal efficiency and background rejection [73].

- FCTight_FixedRad stands for FixedCone Tight with a fixed radius. This means that a fixed cone around the muon is used to calculate the isolation, and the Tight requirement enforces a strict cut to ensure that the muon is isolated from other particles within the cone. This criterion is used to reduce background contamination from non-prompt or fake muons [74, 75].

An alternative categorization divides events into two groups: one containing real leptons (R) and the other containing fake leptons (F). The total number of events is then constrained by the relation:

$$N_T + N_L = N_R + N_F \,, \tag{3.8}$$

where N_T represents tight leptons and N_L represents loose leptons. The expected number of events in the control regions (CR), using the Matrix Method (MM), can be described by the matrix equation:

$$\begin{pmatrix} \langle N_T \rangle \\ \langle N_L \rangle \end{pmatrix} = \begin{pmatrix} \epsilon_r & \epsilon_f \\ 1-\epsilon_r & 1-\epsilon_f \end{pmatrix} \begin{pmatrix} N_R \\ N_F \end{pmatrix}, \tag{3.9}$$

where:

- $\langle N_T \rangle$ and $\langle N_L \rangle$ are the expected number of leptons in the tight and loose groups, respectively;

- ϵ_r and ϵ_f represent the real and fake lepton efficiencies, i.e., the probability for a real (fake) lepton to pass the tight selection, $\epsilon_r = P(T|R)$ and $\epsilon_f = P(T|F)$;

- N_R and N_F are the number of real and fake leptons, respectively.

To estimate the number of real N_R and fake N_F leptons, Equation 3.9 must be inverted. Assuming $\epsilon_r \neq \epsilon_f$, the inversion becomes

$$\begin{pmatrix} N_R \\ N_F \end{pmatrix} = \frac{1}{\epsilon_r - \epsilon_f} \begin{pmatrix} 1-\epsilon_f & -\epsilon_f \\ \epsilon_r - 1 & \epsilon_r \end{pmatrix} \begin{pmatrix} \langle N_T \rangle \\ \langle N_L \rangle \end{pmatrix}. \tag{3.10}$$

From Equation 3.10, the estimators for the number of real and fake leptons passing the tight selection, N_R^T and N_F^T, can be constructed as

$$N_R^T = \epsilon_r N_R = \frac{\epsilon_r}{\epsilon_r - \epsilon_f} \left(N_T - \epsilon_f (N_T + N_L)\right), \tag{3.11}$$

$$N_F^T = \epsilon_f N_F = \frac{\epsilon_f}{\epsilon_r - \epsilon_f} \left(\epsilon_r (N_T + N_L) - N_T\right), \tag{3.12}$$

where N_T and N_L are the observed numbers of tight and loose leptons, respectively.

Both real ϵ_r and fake ϵ_f lepton efficiencies are derived as two-dimensional distributions in lepton pseudorapidity η and transverse momentum p_T. The background estimation is

performed by summing weights for each event within a certain bin of the 2D map. The MM weights $w_{T,i}$ and $w_{L,i}$ for events passing the tight and loose selections are defined based on Equation 3.12 as

$$w_{T,i} \equiv \frac{\epsilon_f}{\epsilon_r - \epsilon_f}\left(\epsilon_r - 1\right), \tag{3.13}$$

$$w_{L,i} \equiv \frac{\epsilon_r \epsilon_f}{\epsilon_r - \epsilon_f}. \tag{3.14}$$

Fake efficiencies are computed using W/Z+jets MC samples, while real efficiencies are measured using $Z \to \ell\ell$ ($ee/\mu\mu$) MC samples, based on truth efficiencies and the electron/muon scale factor (SF) ratio.

For the dilepton channel, due to limited statistics (10–15 events), the 4×4 matrix approach could not have been used.

Measurement the fake and real of efficiencies

Preselection cuts and prompt subtraction were applied to suppress W/Z+jets background. Criteria imposed on the control region (CR) enriched with fake leptons are

- Exactly one lepton
- $\mathrm{N_{jet,18}} \geq 1$ for PF/HI jet collection
- $E_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV
- $\mathrm{N_{bjet}} = 0{,}1, \geq 2$

where $\mathrm{N_{jet,18}}$ is a number of jets with transverse momentum greater than 18 GeV, and $\mathrm{N_{bjet}}$ stands for a number of b-tagged jets (jet matching algorithm imposed).

Fake lepton efficiency [76] is defined as ratio of tight and loose non-promt leptons calculated by taking number of leptons in data and subtracting prompt and charge-flipped leptons scaled to data luminosity

$$\epsilon_f = \frac{N_T^{\mathrm{data}} - \sum_{i=1}^{4} w_i \cdot N_{T,i}^{\mathrm{MC,prompt}} - \sum_{i=1}^{4} w_i \cdot N_{T,i}^{\mathrm{CF}}}{N_L^{\mathrm{data}} - \sum_{i=1}^{4} w_i \cdot N_{L,i}^{\mathrm{MC,prompt}} - \sum_{i=1}^{4} w_i \cdot N_{L,i}^{\mathrm{CF}}} \tag{3.15}$$

where:

- N_T^{data} (N_L^{data}) is a number of events with tight (loose) leptons in the data sample. Index i corresponds to two isospin configurations for pp and pn interactions in two beam configurations p+Pb system and Pb+p;
- w_i is a normalization factor given as

$$w_i = \frac{L_{\mathrm{data}}}{L_{\mathrm{MC}}} \tag{3.16}$$

  where

  - L_{data} – is the integrated luminosity of the data sample in beam configuration reflecting MC conditions (p+Pb or Pb+p beam configurations);
  - L_{MC} – is the integrated luminosity of the simulated MC sample;

- $N_{T,i}^{\mathrm{CF}}$ ($N_{L,i}^{\mathrm{CF}}$) is a number of events with charged flipped leptons passing tight (loose) selection;

- $N_{T,i}^{\mathrm{MC,promt}}$ ($N_{L,i}^{\mathrm{MC,promt}}$) is a number of events with tight (only loose) prompt leptons. The subtracted terms in Equation 3.15 are estimated from MC simulation using truth information.

The real lepton efficiencies are derived from MC (with a truth-mathing) and multiplied by the scale factor SF-ratio SF(right)/SF(loose) [76]:

$$\epsilon_r = \epsilon_{\mathrm{truth}} \times \frac{\mathrm{SF}(\texttt{Tight})}{\mathrm{SF}(\texttt{Loose})}, \; \mathrm{SF} = \frac{\epsilon_{\mathrm{data}}}{\epsilon_{\mathrm{MC}}} \tag{3.17}$$

where SF-ratio denotes the scale factors derived for tight and loose electrons/muons, respectively. An each term in definition of SF-ratio depends on the complete chain of SFs: SF(`Reconstruction`) × SF(`Identification`) × SF(`Isolation`) to tight and loose leptons and according to their definitions given in Table 3.10.

The resulting electron and muon efficiencies are shown in Figures 3.12, and 3.13 respectively.

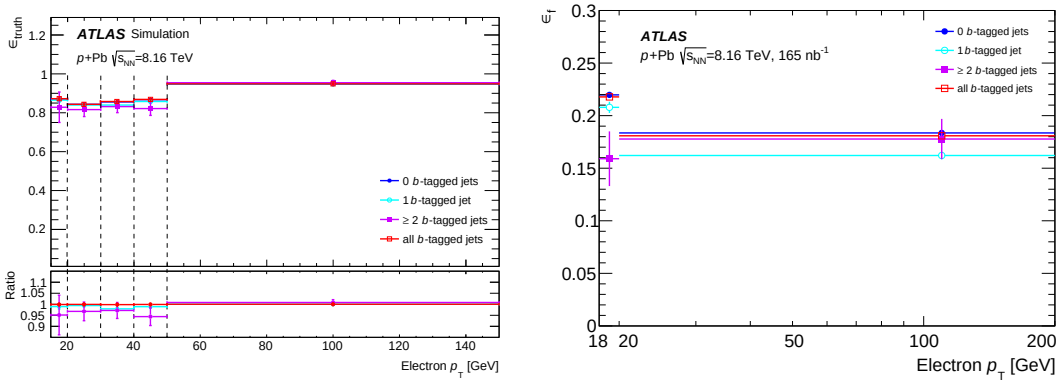


Figure 3.12: Efficiencies ϵ_r (top) and ϵ_f (bottom) for electrons as a function of p_T for events with a varing number of b-tagged jets in the event [48].

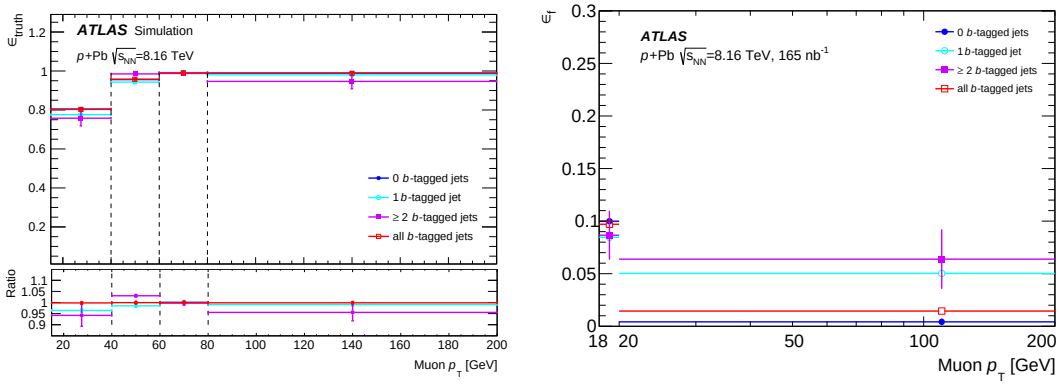


Figure 3.13: Efficiencies ϵ_r (top) and ϵ_f (bottom) for muons as a function of p_T. The efficiencies are shown separately for three selection requirements [48].

The two-dimensional efficiencies in p_T and η are used in the estimation of the background contribution by summing weights of each event according to Equations 3.13, 3.14.

3.6 Muon Performance

The goal of muon performance studies is to derive scale factors based on recommendations and tools given by Muon Combined Performance group and summarized in the publication [66].

The settings remained default as for the *pp* analysis except that no pile-up correction was introduced. The challage was to modify recommended software to the MC samples overlaid with data characteristic for proton-lead collisions. Three types of scale factors were derived

- reconstruction/identification,
- isolation,
- trigger,

joining reconstruction and identification efficiencies into one efficiency. The total scale factor is

$$\begin{aligned} \mathrm{SF}_{\mathrm{total}} &= \frac{\varepsilon_{\mathrm{total,\ data}}}{\varepsilon_{\mathrm{total,\ MC}}} = \frac{\varepsilon_{\mathrm{reco/id,\ data}} \times \varepsilon_{\mathrm{iso,\ data}} \times \varepsilon_{\mathrm{trig,\ data}}}{\varepsilon_{\mathrm{reco/id,\ MC}} \times \varepsilon_{\mathrm{iso,\ MC}} \times \varepsilon_{\mathrm{trig,\ MC}}} \\ &= \mathrm{SF}_{\mathrm{reco/id}} \times \mathrm{SF}_{\mathrm{iso}} \times \mathrm{SF}_{\mathrm{trig}}, \end{aligned} \tag{3.18}$$

while $Z \to \mu\mu$ data and MC samples were used. Due to the ratio, possible biases introduced by the measurement method which appear both in data and MC simulation cancel out. The total scale factor quantifies the deviation of the simulation from the real detector behaviour, and is therefore used in physics analysis to correct the simulation used as a weight [66].

3.6.1 Reconstruction/Identification Efficiency

The tag-and-probe method was used to measure the reconstruction and identification efficiency. For the muon reconstruction the information from inner detector (ID), Muon spectrometer (MS), and calorimeters were taken. Based on reconstruction algorithm the five muon types are defined [66]:

- combined muon (CB) - matches MS to ID tracks and performs a combined fit based, the energy loss from calorimeters is taken into account;
- inside-out combined muon (IO) - extrapolates ID to the MS tracks and tries to find at least 3 loosely-aligned MS hits;
- MS extrapolated muon (ME) - defined when MS track cannot be matched to an ID track, MS track parameters are extrapolated to the beamline;
- segment-tagged muon (ST) - ID track extrapolated to the MS with tight angular matching requirements applied to at least one reconstructed MS segment;
- calorimeter-tagged muon (CT) - extrapolates ID tracks to the calorimeters to search for energy deposits consistent with a minimum-ionising particle.

Working points

The ratio of charge and momentum of muons expressed at interaction point called q/p compatibility and is defined as

$$q/p = \frac{|q/p_{\mathrm{ID}} - q/p_{\mathrm{MS}}|}{\sqrt{\sigma^2(q/p_{\mathrm{ID}}) + \sigma^2(q/p_{\mathrm{MS}})}}, \tag{3.19}$$

where $\sigma(q/p)$ are uncertainties of the measurement. Three major muon woring points are defined as

- The *Medium* working point refer to muons CP and IO within $|\eta| < 2.5$ with $q/p < 7$, must have at least two precision stations (parts of MS with precise measurements of the position and angle of a muon's trajectory), except muons within $|\eta| < 0.1$ are required to have 1 precision station. In $t\bar{t}$ events prompt *Medium* muons are by 98% CB muons.

- The *Loose* working point are the muons passing the *Medium* working point with additional CT and ST muons in pseudorapidity $|\eta| < 0.1$, where the gap in the MS causes loss of efficiency for CB muon reconstruction. To increase the efficiency of the *Loose* criteria for low-p_{T} muons, IO muons having $p_{\mathrm{T}} < 7$ GeV and only one precision station are accepted in $|\eta| < 1.3$. Requiring that the purity of IO muons increases.

- The *Tight* working point are the *Medium* muons, but only those CB and IO muons having at least two precision stations. The normalised χ^2 of the combined track fit is required to be less than 8 to reject tracks due to hadron decays. Futher optimisated cuts on the q/p and ρ' ($|p_{\mathrm{T,ID}} - p_{\mathrm{T,MS}}|/p_{\mathrm{T,CB}}$) depending on the p_{T} and $|\eta|$ of the muon are applied to supress backgournd from non-prompt muons [66].

Event selection

The invariant mass of dimuon (tag and probe) candidates has to be within the range of $m_{\mu\mu} \in \{61,121\}$ GeV having opposite charge. The tag has to satisfy *Medium* identification with $p_{\mathrm{T}} > 18$ GeV, $|\eta| < 2.5$, and the single-muon trigger requirements. Additionally, tag has to fulfill *Tight* isolation criteria

$$p_{\mathrm{T}}^{\mathrm{varcone30}} < 0.04 \cdot p_{\mathrm{T}}^{\mu}, \quad E_{\mathrm{T}}^{\mathrm{topoetcone20}} < 0.15 \cdot p_{\mathrm{T}}^{\mu}, \quad p_{T}^{\mathrm{track}} > 1 \text{ GeV} \tag{3.20}$$

where p_{T}^{μ} stands for transverse momentum of the muon, $p_{\mathrm{T}}^{\mathrm{varcone30}}$ is the sum of p_{T} tracks aound the muon within the cone of

- $\Delta R < \min(10 \text{ GeV}/p_{\mathrm{T}}^{\mu}, 0.3)$, for $p_{\mathrm{T}}^{\mu} <= 50$ GeV,

- $\Delta R < 0.2$, for $p_{\mathrm{T}}^{\mu} > 50$ GeV.

Similarly $E_{\mathrm{T}}^{\mathrm{topoetcone20}}$ is the sum of E_{T} of tracks around the muon within $\Delta R < 0.2$.

Furthermore, the vertex association criteria ensure a maximal purity of tags originating from the hard-scattering proton–proton collision. The criteria includes:

- Transverse impact parameter significance: $|d_0/\sigma(d_0)| < 3$,

- Longitudinal impact parameter: $|z_0| \sin\theta < 0.5\,\mathrm{mm}$.

where d_0 is the transverse impact parameter (shortest distance between the muon track and the primary vertex in the xy-plane), $\sigma(d_0)$ is the uncertainty of d_0, the z_0 longitudinal impact parameter is the shortest distance between the muon track and the primary vertex along the z-axis.

The probe muons have looser selection than tag muons

- $p_{\mathrm{T}} > 18$ GeV

- $|\eta| < 2.5$

- Transverse impact parameter significance: $|d_0/\sigma(d_0)| < 3$,

- Longitudinal impact parameter: $|z_0| \sin\theta < 10\,\mathrm{mm}$.

and probe isolation is more stringent calorimeter-based and looser track-based than the tag *Tight* isolation.

The identification and reconstruction efficiency for *Medium* woring point is defined as

$$\epsilon\,(\mathit{Medium}) = \epsilon\,(\mathrm{ID}) \times \epsilon\,(\mathit{Medium}|\mathrm{ID}) \simeq \epsilon\,(\mathit{Medium}|\mathrm{CT}) \times \epsilon\,(\mathrm{ID}|\mathrm{MS})\,. \tag{3.21}$$

since efficiency $\epsilon\,(\mathrm{ID})$ cannot be measured directly, it is repaced by the conditional efficiency $\epsilon\,(\mathrm{ID}|\mathrm{MS})$ for a muon reconstructed by the MS to be also reconstructed in the ID. To reduce

the background contamination, the $\epsilon\,(\mathit{Medium}|\mathrm{ID}) \simeq \epsilon\,(\mathit{Medium}|\mathrm{CT})$ approximation was used, replacing ID probes with the more pure CT probes, and a systematic uncertainty was assigned to cover for the small bias introduced. The validity of Equation 3.21 is guaranteed by the independecy of the track reconstruction in the ID and in the MS and can be defined for other woring points (*Loose* *Tight*, etc.).

The are few options of types of muon probes which test different kind of efficiencies listed below [66]. The probes reconstructed by one detector subsystem test efficiency of the other independent subsystem.

- ID probes are ID tracks used to measure the reconstruction efficiency in the MS, or of specific identification algorithms.
- MS probes are ME tracks used to test the efficiency of the ID reconstruction.
- CT probes are ID tracks also satisfying the calo-tagging reconstruction algorithm described in Section 4 of [66]. In the same way as the ID probes, they are used to measure the reconstruction efficiency in the MS, or of specific muon identification algorithms.
- ST probes are ID tracks also satisfying the segment-tagging reconstruction algorithm described in Section 4 of [66]. In the same way as the ID probes, they are used to measure the reconstruction efficiency in the MS, or of specific muon identification algorithms.
- *Two-track* probes are MS tracks required to be within $\Delta R = 0.05$ of an ID track. They are used to measure the combined reconstruction efficiency of a muon candidate with ID and MS tracks, or the efficiency of specific identification criteria.
- *Loose* probes are muon candidates satisfying the *Loose* identification requirements. They are used to measure the isolation and vertex association efficiencies.

The efficiency of a certain algorithm is measured using a matching requirement of $\Delta R < 0.05$ between the probe and any muon candidate reconstructed and identified with the algorithm of interest. The efficiency is then computed as the number of probes P that are successfully matched to a muon reconstructed and identified according to the X criterion, N_P^X, divided by the total number of selected probes N_P^{All}:

$$\epsilon\,(X|P) = \frac{N_P^X}{N_P^{\mathrm{All}}}. \tag{3.22}$$

Probes are counted in data events after the subtraction of the backgrounds. In simulation, to eliminate the background contamination, both the tag and the probe muons are required to be a prompt muon at generator level [66].

Background subtraction

The background in data are is estimated by fit in the range $m_{\mu\mu} \in \{61{,}121\}$ GeV separately for the all selected and matched probes. The background contribution (non-prompt muons) is modelled using the form

$$f(m_{\mathrm{tag-probe}}) = \left(1 - \frac{m_{\mathrm{tag-probe}}}{\Lambda}\right)^{p_1} \cdot \left(\frac{m_{\mathrm{tag-probe}}}{\Lambda}\right)^{p_2}, \tag{3.23}$$

where the Λ parameter is approximately the energy necessary to produce the dimuon pair, set to 2.5 times of upper mass range of $m_{\mu\mu}$ ($\Lambda = 2.5 \times 121 = 302.5$ GeV) and the p_1 and p_2 are obtained using a separate fit with a sample of same-charge tag-and-probe pairs, satisfying all the selection criteria except the isolation requirements.

Systematic uncertainties

The list of systematic uncertainties contributing the most to the total uncertainty is listed below [66]

- *T&P method*: Possible biases in the tag-and-probe approach, such as different kinematic distributions between reconstructed probes and generated muons, or ID-MS efficiency correlations, are evaluated using simulation. This is done by comparing the reconstructed efficiency to the fraction of generator-level muons that are successfully reconstructed. Since this bias affects both data and simulation similarly, it cancels out in the scale factor (SF) calculation. Half of the observed difference is assigned as the SF uncertainty to conservatively cover potential imperfections in the simulation. Utilizing *two-track* probes lowers this uncertainty to below approximately 0.1%, which further decreases as p_{T} increases.

- *Probe matching*: the default ΔR-based matching procedure is varied in order to assess an uncertainty in how much a given probe type contributed to a certain type of reconstructed muon candidate. This is done by comparing the nominal fraction of matched probes with the fraction of probe tracks for which muon candidate reconstruction is successful.

- *Template shape*: the uncertainty in the shape of the template modelling the non-prompt muon background is evaluated by simultaneously varying the p_1 and p_2 parameters in Eq. (3.23) by their fit uncertainties. The consequent deviation of the SFs from their nominal value is taken as the systematic uncertainty.

- Λ-*SC*: the numerical value of the Λ parameter in Eq. (3.23) guarantees a well-behaved, smooth function across $m_{\mathrm{tag-probe}}$. Possible effects on the SFs are estimated by varying its value by $\pm 20\%$.

- *Background fit*: to cover effects associated with the fitting procedure used to extract the contribution of the non-prompt muon background, the change in the SFs obtained when varying the fitted non-prompt muon background by its corresponding fit uncertainty is assigned as a systematic uncertainty.

Muon reco/ID scale factors

Reconstruction/identification scale factors are derived in the *Medium* working point. The data efficiency is larger compare to Monte Carlo which leads into scale factor exceeding unity by up to 5% in Fig. 3.14. This result was discussed with muon CP experts and result was cross-checked using independed code, see in Appendix J Figures 5.5e–5.7e.

3.6.2 Isolation Efficiency

To measure the muon isolation efficiencies the probe muons with $p_{\mathrm{T}} > 18$ GeV are used. In order to improve the background rejection at low p_{T}, the *Loose* identification criteria with the standard vertex association requirements are applied to the probe. Additionally, the tag muon within the isolation code of the probe ($\Delta R_{\mu\mu} < 0.3$) is rejected.

The background fit remains the same as in case of reconstruction efficiencies with adjustment of the $m_{\mu\mu}$ interval to $m_{\mu\mu} \in \{81{,}101\}$ GeV. If a jet overlaps with a selected probe within $\Delta R(\mathrm{jet},\mu) < 0.4$, the probe is discarded if either of the following conditions is met:

- the probe-to-jet p_T ratio is less than 0.5, or

- the probe-to-jet-tracks p_T ratio (calculated as the probe's p_T divided by the summed vector p_T of all tracks associated with the jet) is less than 0.7.

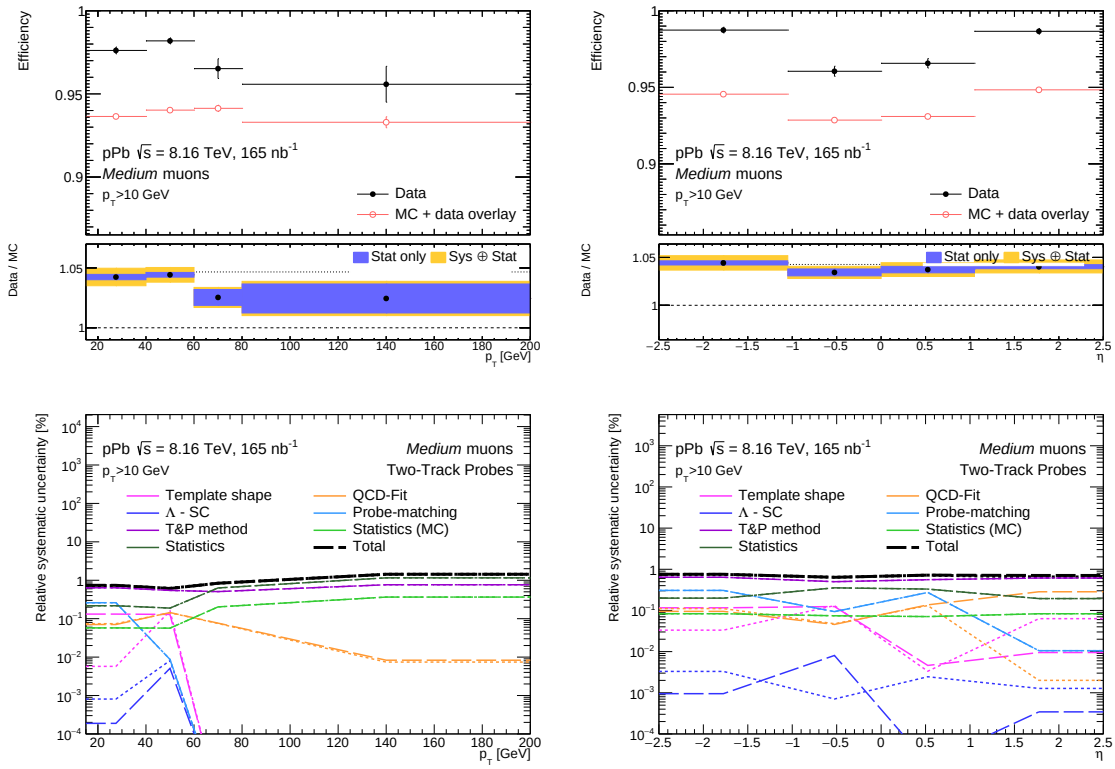


Figure 3.14: The muon reconstruction/identification ϵ(Medium) efficiency as a function of p_T and η (below are shown individual scale factor uncertainties).

This rejection reduces contamination from bottom and charm hadron decays.

Systematic uncertainties

All uncertainties discussed in Sec. 3.6.1 also apply to the measurement of the isolation efficiency scale factors, except for the *T&P method* and *Probe matching* uncertainties.

In contrast, additional sources of systematic uncertainty are introduced for measuring the isolation efficiencies and SFs:

- *Probe PID*: The selected probe identification working point affects both background contamination and signal yield, creating a systematic uncertainty. This is evaluated by comparing the SFs for *Loose* vs. *LowPt* probes for $p_{\mathrm{T}} < 15$ GeV, or *Loose* vs. *Tight* for $p_{\mathrm{T}} > 15$ GeV.
- *Mass window*: Varying the $m_{\mathrm{tag-probe}}$ range in the template fit can change the SF values. The largest observed shift is assigned as the uncertainty, comparing windows of 81–101 GeV, 86–96 GeV, and 71–111 GeV.
- $\Delta R\,(\mathrm{jet},\mu)$: as the isolation efficiencies and scale factors are found to depend on the angular distance between the probe and the closest jet, the procedure used to resolve muon–jet reconstruction ambiguities is a source of systematic uncertainty. To account for it, the criteria for the *probe-to-jet* and *probe-to-jet-tracks* p_{T} ratios are independently dropped, and the largest change in the SFs is taken as the uncertainty.

Muon isolation scale factors

The isolation scale factors are shown in *FCTight_FixedRad* and *Loose* working point in Figure 3.15 exceed unity by about 2%.

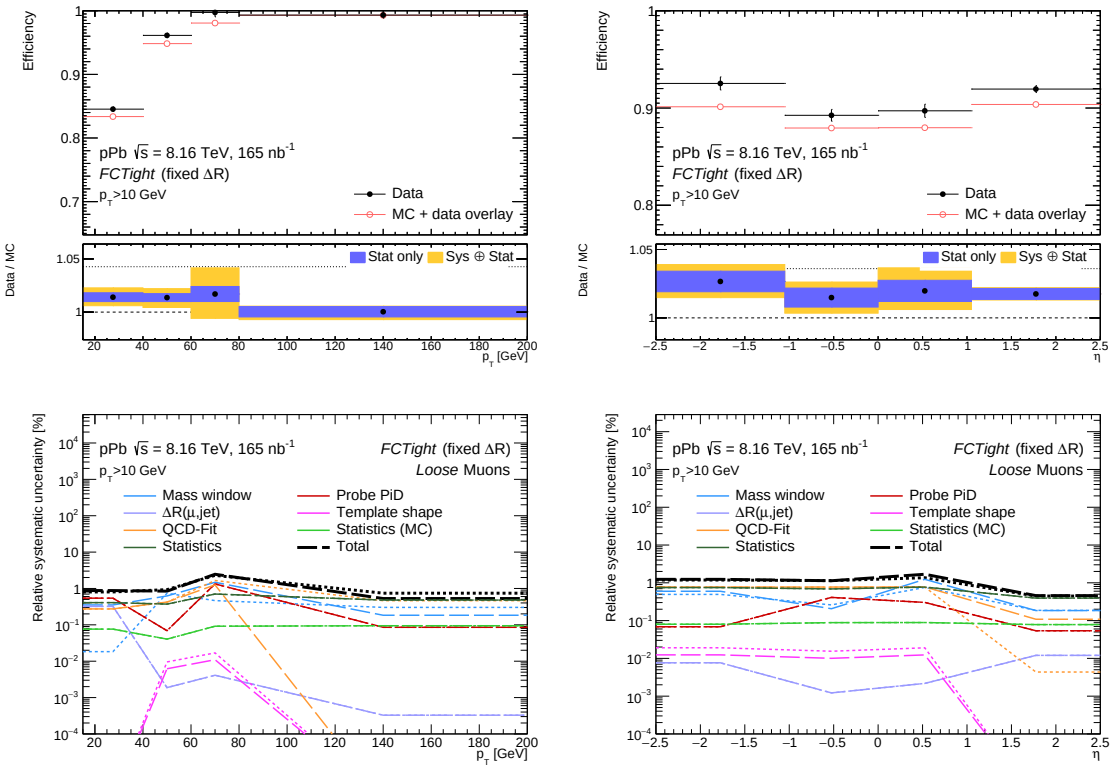


Figure 3.15: The isolation efficiency *FCTight_FixedRad*, Loose working point as a function of p_T and η (below are shown individual scale factor uncertainties - dotted lines stand for down variation).

3.6.3 Trigger Efficiency

The trigger efficiency is evaluated using three triggers combined into one trigger HLT_mu15_comb = HLT_mu15 or HLT_mu15L1MU6 or HLT_mu15L1MU10 using Equation 3.22 which can be written as

$$\epsilon^{\text{Medium}}_{\text{HLT\_mu15\_comb}} = \frac{N^{\text{Medium}}_{\text{probe passing HLT\_mu15\_comb}}}{N^{\text{Medium}}_{\text{tag passing HLT\_mu15\_comb}}}. \tag{3.24}$$

Systematic uncertainties

The trigger systematic uncertainties are named (as appearing in Figures 3.16) asd defined as

- dphill - Detector symmetry: Selecting back-to-back probe and tag muons $\Delta\Phi < \pi - 0.1$;
- isoTight, isoTightTrackOnly - Isolation;
- mll - Z-mass window: Loosen invariant mass requirement to Z mass $\pm$ 5 GeV;
- muneg, mupos - Muon Charge: Force probes to be all negative or positive;
- noIP - Interaction parameter: Remove requirement that probe, tag come from same vertex;
- ptup, ptdw - p_T dependence: Selecting only "low p_T <= 30 GeV" or "high p_T > 30 GeV" events.

Muon trigger scale factors

The trigger scale factors of HLT_mu15_comb trigger in *Medium* working point are shown in Figure 3.16. The agreement with unity is satisfactory, except in the ATLAS detector "foot" region (bin $-2 < \phi < -1$ and $-1 < \eta < 1$, see Appendix J, Figure 5.17) where the SF values exceeds unity by 20%. This effect is being expected.

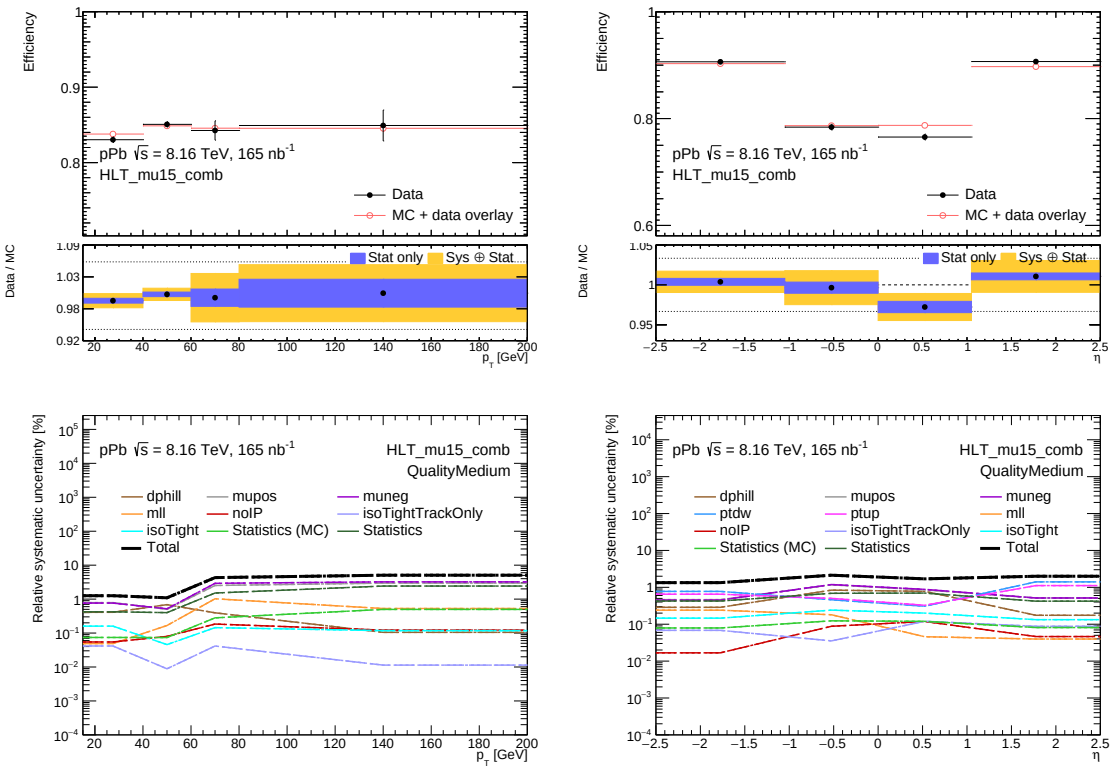


Figure 3.16: The trigger efficiency HLT_mu15_comb in medium working point as a function of p_T and η (below are shown individual scale factor uncertainties).

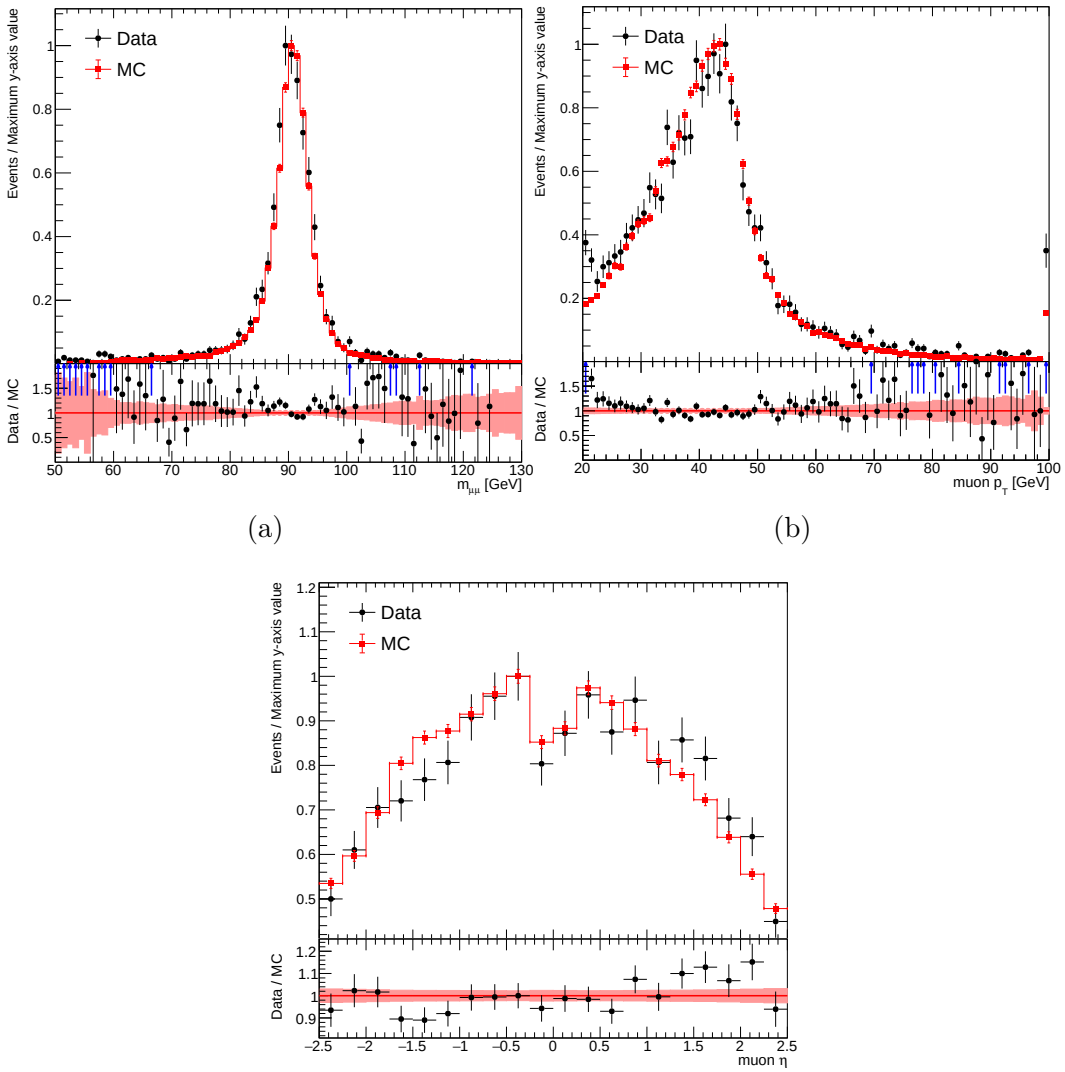
3.6.4 Control Plots with $Z \to \mu\mu$ events

Validation of muon scale factors has been done with $Z \to \mu^+\mu^-$ events using PowHeg MC simulation. Events with two Medium muons and FCTight_FixedRad isolation have been selected. Two muons had to be of an opposite sign. A requirement on the invariant mass of the dimuon system was imposed to be between 50-130 GeV. No requirement on jets was imposed.

Figure 3.17 shows distributions for the dimuon system invariant mass, muon p_{T}, and muon pseudo-rapidity. A good description of the data by MC simulation with muon scale factors applied is found with some features that can be attributed to background in data or to muon scale and resolution differences. The 2D maps of scale factors which were applied the the MC of proton-lead analysis can be found in Appendix J.

3.7 Electron Performance

Similar as for muons the aim of electron performance studies is to correct for any potential detector mismodelling using four sets of scale factors (SF) in Equation 3.25 related to electron



(a) (b)

(c)

Figure 3.17: Illustration of control distributions for $Z \to \mu^+\mu^-$ events compared with MC simulation (100k events) for the signal process. Only the statistical uncertainties are shown.

reconstruction $\varepsilon_{\text{reco}}$, identification ε_{id}, isolation ε_{iso} and trigger $\varepsilon_{\text{trig}}$ with both statistical and systematic uncertainties.

$$\begin{aligned} \text{SF}_{\text{total}} &= \frac{\varepsilon_{\text{total, data}}}{\varepsilon_{\text{total, MC}}} = \frac{\varepsilon_{\text{reco, data}} \times \varepsilon_{\text{id, data}} \times \varepsilon_{\text{iso, data}} \times \varepsilon_{\text{trig, data}}}{\varepsilon_{\text{reco, MC}} \times \varepsilon_{\text{id, MC}} \times \varepsilon_{\text{iso, MC}} \times \varepsilon_{\text{trig, MC}}} \\ &= \text{SF}_{\text{reco}} \times \text{SF}_{\text{id}} \times \text{SF}_{\text{iso}} \times \text{SF}_{\text{trig}}. \end{aligned} \tag{3.25}$$

These efficiencies may be estimated directly from data or MC simulation using the tag-and-probe method. The events are selected based on the electron pair invariant mass m_{ee} from $Z \to ee$ process (covering a range of electron p_{T} 15–200 GeV.), unbiased samples of electrons (probes) by using strict selection requirements on the second object (tags) produced from the particle's decay. Events are selected on the basis of the electron–positron invariant mass. The efficiency of a given requirement can then be determined by applying it to the probe sample after accounting for residual background contamination [65].

To properly estimate the signal, a background subtraction has to be performed. The background contribution is estimated by reversing the selection criteria for the probe electron and normalizing to the data using the sidebands of the m_{ee} distribution. To evaluate

the systematic uncertainties, two background templates are introduced. For electrons with transverse energy

$$E_{\mathrm{T}} = \sqrt{p_{\mathrm{T}}^2 + m^2} < 30 \text{ GeV} \tag{3.26}$$

the template normalization is modified by considering the invariant mass region of 60 GeV$<$ $m_{ee} < 70$ GeV, along with the nominal range of 120 GeV$<$ $m_{ee} < 250$ GeV. For electrons with $E_{\mathrm{T}} > 30$ GeV, the nominal normalization is applied, while alternative selection criteria are used to introduce variation.

In the numerator the background template is normalised to the data containing tag-and-probe pairs with the same charge in the tail region of the invariant mass distribution. Same-charge data are used to reduce the contribution of signal electrons. For the denominator, a similar approach is used, but the opposite-charge data are used. Some signal events remain in the background template, which are subtracted using Monte Carlo (MC) simulation.

To derive efficiencies and their uncertainties, multiple variations of the selection criteria are considered. These variations include adjustments to parameters such as the tag electron selection, the mass window around the Z-boson peak, and the background template. The final efficiency value is obtained by averaging the results across all variations. Systematic uncertainty is then computed as the root mean square of these individual results, while statistical uncertainty is averaged from the statistical errors of the variations.

The reconstruction efficiencies are presented in a detail in following section 3.7.2, while identification, isolation, and trigger efficiencies are discussed in Appendix K.

3.7.1 Tag-and-probe Selection

Events are selected using a single electron trigger with an E_{T} threshold of 15 GeV and Loose identification requirements [73]. In the next step, the events have to have at least two reconstructed electrons with opposite charges, in the $|\eta| < 2.47$ region of the detector.

The tag electron candidate is required to have transverse energy $E_{\mathrm{T}} > 24.5$ GeV and be located outside the electromagnetic calorimeter transition region ($1.37 < |\eta| < 1.52$). Additionally, the tag must meet identification and isolation criteria. In cases where both electron candidates satisfy the tagging criteria, the event contributes two potential electron pairs.

For the probe electron, the minimum requirement is $E_{\mathrm{T}} > 15$ GeV, along with successful reconstruction of the electron.

3.7.2 Electron Reconstruction Efficiency

An algorithm of the electron reconstruction is divided into the following steps: EM cluster reconstruction, track reconstruction, loose electron track fit and final electron reconstruction [77].

The reconstruction efficiency is defined as the ratio of the number of reconstructed electrons to the number of EM clusters. The efficiency of creating an EM cluster for an electron with E_{T} greater than 15 GeV is measured to be above 99% [78].

The probes are required to be EM clusters, while the tag has to pass tighter electron requirements as specified in Section 3.7.1. Tag-and-probe pairs have no requirements on the charge.

The reconstruction efficiency ϵ_{reco} is defined for both MC and data and is represented in terms of background-subtracted number of events passing

$$\epsilon_{\mathrm{reco}} = \frac{N_{\mathrm{track}}^{\mathrm{good}} - B_{\mathrm{track}}^{\mathrm{good}}}{(N_{\mathrm{track}}^{\mathrm{good}} - B_{\mathrm{track}}^{\mathrm{good}}) + (N_{\mathrm{track}}^{\mathrm{bad}} - B_{\mathrm{track}}^{\mathrm{bad}}) + (N_{\mathrm{track}}^{\mathrm{no}} - B_{\mathrm{track}}^{\mathrm{no}})}, \tag{3.27}$$

where

- $N^{\text{good}}_{\text{track}}(B^{\text{good}}_{\text{track}})$ is a number of reconstructed electrons (background events) with a good track quality. The track is matched to the EM cluster, and minimum seven precision hits and one pixel hit are required (see Figure 3.18),
- $N^{\text{bad}}_{\text{track}}(B^{\text{bad}}_{\text{track}})$ is a number of reconstructed EM clusters (background events) with a bad track quality. Track-cluster matching is imposed. The track has to include between four and six silicon hits and no pixel hit or minimum four silicon hits and one pixel hit (see Figure 3.18),
- $N^{\text{no}}_{\text{track}}(B^{\text{no}}_{\text{track}})$ is a number of reconstructed EM clusters (background events) without the track. The track is only associated to TRT hits or no track is associated at all (see Figure 3.18).

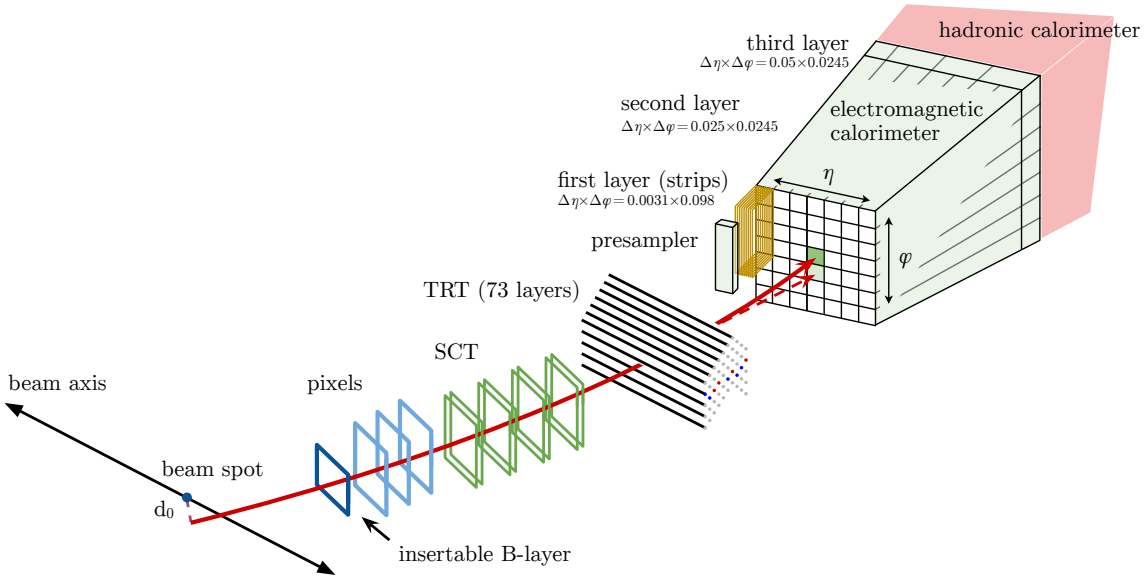


Figure 3.18: A schematic illustration of the path of an electron through the detector. The red trajectory shows the hypothetical path of an electron, which first traverses the tracking system (pixel detectors, then silicon-strip detectors and lastly the TRT) and then enters the electromagnetic calorimeter. The dashed red trajectory indicates the path of a photon produced by the interaction of the electron with the material in the tracking system [73].

The contribution of electron-positron pairs from background processes is estimated directly from data. Background templates are generated for electron candidates with good track quality ($B^{\text{good}}_{\text{track}}$) and poor track quality ($B^{\text{bad}}_{\text{track}}$).

For electron candidates without an associated track ($B^{\text{no}}_{\text{track}}$), the background template is modeled using a third-order polynomial, which is fitted to the invariant mass distribution of trackless electron candidates. Systematic variations are defined by applying different regions of the fit.

Figure 3.19 shows the electron reconstruction efficiency as a function of E_{T} and η, for both data and Monte Carlo (MC) simulation. The efficiency increases with E_{T}, starting from 93% at $E_{\text{T}} = 15$ GeV, and reaches a plateau of around 98% for E_{T} values near 50 GeV. The derived scale factors are found to be consistent with unity.

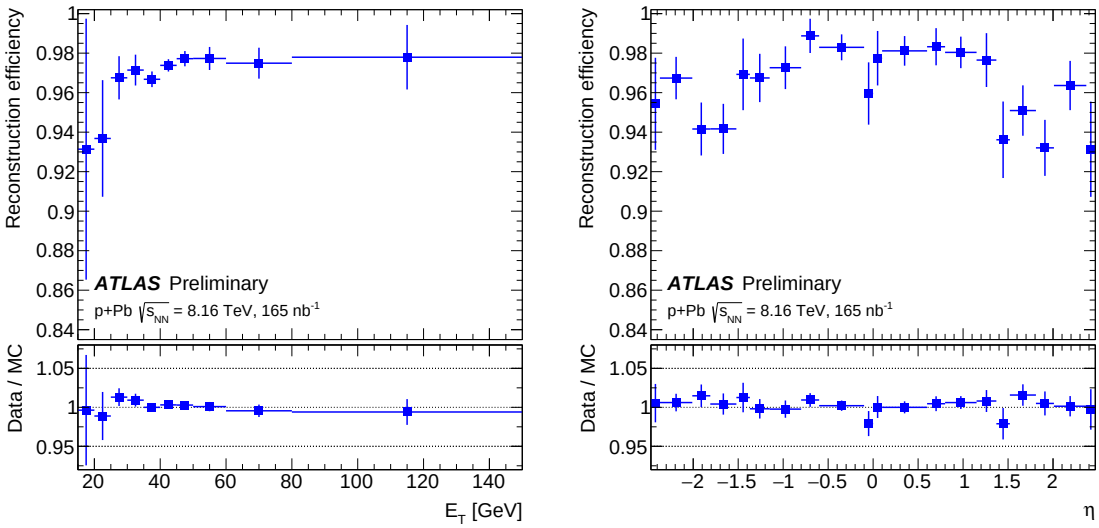


Figure 3.19: Electron reconstruction efficiency as a function of electron E_{T} (left) and η (right) evaluated in the data [79].

3.8 Jet Performance

The goal of jet performance studies is to validate jet energy scale and resolution for both PF and HI jet collection by a proxy variable that can be defined for both data and MC, along for in-situ data / MC comparison. The recommended systematic uncertainties are used for each PF and HI jet collection. A reasonable agreement between data and MC is observed.

3.8.1 Event selection

For the jet performance analysis, events containing a Z boson candidate decaying into two leptons are selected. The selection is divided into two distinct channels: one for the $Z \to ee$ process and the other for $Z \to \mu\mu$.

All electron candidates must satisfy Medium identification criteria and meet the Gradient isolation requirements. Similarly, muon candidates must fulfill the Medium identification standard along with the PflowTight_FixedRad isolation criteria.

In both the ee and $\mu\mu$ channels, at least one jet with a transverse momentum (p_{T}) exceeding 15 GeV is required. The selected events must contain exactly two leptons—either two electrons with no muons, or two muons with no electrons. Both leptons are required to have p_{T} greater than 18 GeV, and they must have opposite electric charges. The invariant mass of the lepton pair is constrained to lie within the Z boson mass window, specifically $m_{\ell\ell} \in$ (66,116) GeV.

The performance of two jet collections, PF and HI, is studied independently. PF jets are calibrated using the nominal high-pileup proton-proton ($\sqrt{s} = 13$ TeV) calibration, whereas HI jets are calibrated using the specialized low-pileup proton-lead (p+Pb) calibration at $\sqrt{s} =$ 8.16 TeV.

3.8.2 Z-jet Balance

To evaluate the systematic uncertainties related to jet energy scale (JES) and jet energy resolution (JER), a Z-jet momentum balance method is used. The method involves a jet recoiling against the Z boson, which decays into either electron or muon pairs. The p_{T} of the Z boson is reconstructed using lepton kinematics and serves as the reference for the leading jet p_{T}.

A requirement of $|\Delta\phi(\text{Z,jet})| > 2.8$ is imposed to ensure back-to-back emission of the jet agains the Z boson. In the Z-jet balance method, the per-event jet response is determined as $p_{\mathrm{T}}^{\mathrm{reco}}/p_{\mathrm{T}}^{\mathrm{ref}}$, where the reference transverse momentum $p_{\mathrm{T}}^{\mathrm{ref}}$ is the projection of the Z boson $p_{\mathrm{T}}^{\mathrm{Z}}$ along the jet axis, given by the formula:

$$p_{\mathrm{T}}^{\mathrm{ref}} = p_{\mathrm{T}}^{\mathrm{Z}} |\cos \Delta\phi(\text{Z,jet})|. \tag{3.28}$$

The JES is quantified as the average jet energy response $R^{\mathrm{ref}} = \langle p_{\mathrm{T}}^{\mathrm{reco}}/p_{\mathrm{T}}^{\mathrm{ref}} \rangle$, determined as the mean of a Gaussian function fitted to the jet response distribution. The jet response is measured in several reference transverse momentum $p_{\mathrm{T}}^{\mathrm{ref}}$ and reconstructed jet pseudorapidity η^{reco} intervals.

The JER is evaluated as the standard deviation over the mean, $\sigma^{\mathrm{ref}}/R^{\mathrm{ref}}$, from the same Gaussian fit. Data-to-MC ratios are derived with both statistical and systematic uncertainties. A default flavour fraction of 0.5 ± 0.5 is used to estimate flavour uncertainties. The systematic uncertainties related to JER include MC-only smearing and are symmetrized, due to missing pseudo-data for HI jets.

Figure 3.20 shows the jet response and jet resolution as functions of $p_{\mathrm{T}}^{\mathrm{ref}}$ for the PF jet collection in the combined lepton channel from the Z boson decays. The JES reaches up to 1.0 and drops below unity to 0.88 at higher $p_{\mathrm{T}}^{\mathrm{ref}}$, while the JER varies from 0.4 to 0.24. Data-to-MC ratios for the JES are consistent with unity within uncertainties, while small MC non-closure is observed for the JER.

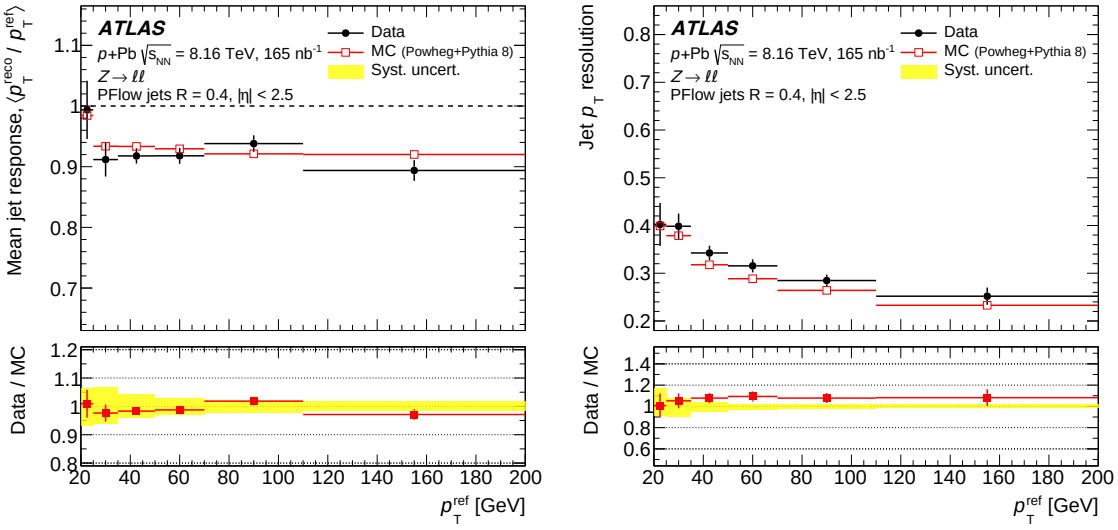


Figure 3.20: JES (left) and JER (right) for the PF jet collection, evaluated in the combined $\ell\ell$ channel from the Z boson decays for the data (full markers) and MC simulation (open markers). Bottom panels show data-to-MC ratios with error bars and yellow boxes representing statistical and systematic uncertainties, respectively [48].

In Figure 3.21 the JES and JER as functions of $p_{\mathrm{T}}^{\mathrm{ref}}$ is presented for the HI jet collection in the combined $\ell\ell$ channel from the Z boson decays. The JES is rising with $p_{\mathrm{T}}^{\mathrm{ref}}$ from 0.77 to 0.9. The JER is larger compared to the PF jet collection, ranging from 0.69 to 0.24. Data-to-MC ratios for the JES are consistent with unity within uncertainties. Differences between data and MC for the JER are covered by systematic uncertainties at $p_{\mathrm{T}}^{\mathrm{ref}} < 50$ GeV, while some MC non-closure is found at higher $p_{\mathrm{T}}^{\mathrm{ref}}$.

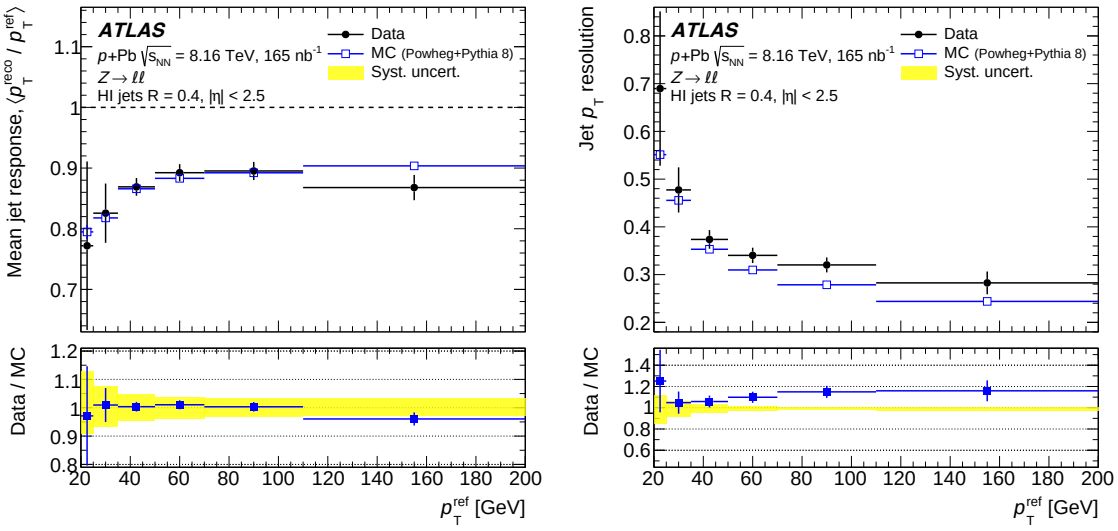


Figure 3.21: JES (left) and JER (right) for the HI jet collection, evaluated in the combined $\ell\ell$ channel from the Z boson decays for the data (full markers) and MC simulation (open markers). Bottom panels show data-to-MC ratios with error bars and yellow boxes representing statistical and systematic uncertainties, respectively [48].

3.9 Systematic Uncertainties

Major systematic uncertainties to the analysis of the $t\bar{t}$ production in p+Pb production at $\sqrt{s_{\mathrm{NN}}}$=8.16 TeV arise from the following sources:

- reconstruction of electrons, muons and jets,
- b-tagging,
- fake-lepton background,
- MC signal + background modelling,
- integrated luminostity.

Reconstruction of electrons, muons and jets

Uncertainties in electron reconstruction, identification, isolation, and trigger efficiencies are derived from $Z \to e^+e^-$ events as described in Section 3.7. The uncertainty in the low pile-up energy calibration is evaluated in line with the methodology presented in the publication [80].

For muons, uncertainties in the momentum scale and resolution are based on the results from the publication [66]. Additionally, uncertainties in the muon scale factors covering reconstruction, isolation, and trigger efficiencies are derived from $Z \to \mu^+\mu^-$ events as described in Section 3.6.

Jet-related uncertainties are inferred by in situ studies of the calorimeter response [81], their application to jets in heavy-ion (HI) data [82], and comparisons of the simulated jet response across samples from different generators.

b-tagging

The b-tagging systematic uncertainties are computed by varying the data-to-MC correction factors within their uncertainties [83, 84, 85]. To assess uncertainties in the HI-PF jet matching, two additional systematic variations are introduced in Section 3.3.5 and are negligible with respect of other uncertainties.

The fake-lepton background

Systematic uncertainties in the fake-lepton background arise from statistical and systematic variations of the real and fake lepton efficiencies, and are evaluated using the Matrix Method. Based on distributions in $0b$ control region the conservative normalization uncertainties of 100% in the μ+jets and 50% in the e+jets and the dilepton signal regions are set as uncorrelated uncertainties.

Shape variations of this background in the ℓ+jets channel were evaluated also in the $0b$ CR. To derive them, all background contributions except the fake-lepton events were subtracted from the data. The difference was normalised to the number of fake-lepton events. A ratio is constructed of such subtracted and scaled data to the fake-lepton contribution as a function of the azimuthal angle $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}},\ell)$ between the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ a variable sensitive to fake lepton background. This ratio is fitted by a second-order polynomial. Shape variations of the fake-lepton background in $1b$ and $\geq 2b$ ℓ+jets SRs are defined as up and down fit shape variations using the fit parameter uncertainties. However, the shape of the $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}},\ell)$ variable is not correlated to the shape of the final fit variable $H_{\mathrm{T}}^{\ell,j}$.

MC signal + background modelling

Signal modeling uncertainties are established using alternative $t\bar{t}$ MC samples based on different choice of parton-shower, hadronization models, and matrix-element algorithms of matching to the parton shower. The uncertainty due to initial-state radiation of partons (ISR) is estimated by variations of α_s for ISR in the A14 tune [51]. Further effects on the ISR are evaluated by varying the renormalisation (μ_{r}) and factorisation scales (μ_{f}) in the matrix-element calculation as well as the Powheg h_{damp} parameter. The μ_{r} and μ_{f} are varied independently by factors of 0.5 and 2.0, avoiding the same side variations of the scales. The effect of final-state radiation of partons (FSR) uncertainties is evaluated by modifying the μ_{r} for emissions from the parton shower by factors of 0.5 and 2.0. The PDF uncertainties affecting the $t\bar{t}$ signal are evaluated using the PDF4LHC15 Hessian uncertainties [53].

Additional background systematic uncertainties related to the normalisation of the V+jets samples are determined using the Berends scaling technique [86, 87, 88]. Single-top-quark diagram removal and diagram subtraction variation samples are used to assess the uncertainties from the interference between the $t\bar{t}$ and tW processes [89]. A conservative uncertainty of 9.5% is considered for the normalisation of both the tW and t-channel single-top-quark processes [90]. The diboson background normalisation is allowed to vary by 50% [91].

Integrated luminostity

The uncertainty in the integrated luminosity of the combined data sample is 2.4%. It is derived from the calibration of the luminosity scale using x-y beam-separation scans, following a methodology similar to that detailed in Ref. [92], and using the LUCID-2 detector for the baseline luminosity measurements [93].

3.10 Fit Procedure

Due to the limited statistics, a differential cross-section measurement or unfolding procedure was not feasible. Instead, the fit strategy focused on determining the signal strength, denoted as $\mu_{\text{fit}}^{\text{data}}$, which quantifies the scaling factor required to match the $t\bar{t}$ Monte Carlo (MC) prediction to the observed data.

The fit to the signal strength variable μ was performed simultaneously across the six signal regions (two dilepton and four ℓ+jets regions) described in Section 3.4 (Table 3.9). The chosen observable for the final fit was $H_{\text{T}}^{\ell,j}$ (see pre-fit plots in each region in Figure 3.22), defined as the scalar sum of the transverse momenta of all jets and selected isolated leptons in each event. This variable offers sensitivity to the kinematic properties of the $t\bar{t}$ events. The combined data yield from all channels amounted to 3024 events. Figure 3.23 provides the number of events in each channel, where the background composition and signal purity vary.

Pre-fit plots display the predicted distributions of observables ($H_{\text{T}}^{\ell,j}$, p_{T}, η, etc.) based on the simulations before any fit to the data. The predictions use nominal values of the model parameters without adjustments to match data. Post-fit plots display the same distributions after the model parameters have been fitted to match the data. This is done through likelihood minimization. The fit refines the parameters to better describe the observed data. Post-fit uncertainties are usually smaller compare to pre-fit because the data constrains the model parameters, reducing the overall uncertainty. The fitting process narrows the range of possible values for parameters.

The total $t\bar{t}$ production cross-section was then computed using the relation

$$\sigma_{t\bar{t}}^{\text{measured}} = \frac{\mu_{\text{fit}}^{\text{data}}}{\text{BR}_{\text{non-allhad}}} \sigma_{\text{MC}}^{\text{t}\bar{\text{t}},\text{non-allhad}}, \tag{3.29}$$

where $\mu_{\text{fit}}^{\text{data}}$ represents the signal strength obtained from the fit, $\sigma_{\text{MC}}^{\text{t}\bar{\text{t}},\text{non-allhad}}$ is the theoretical cross-section (including k-factors) used to normalize the $t\bar{t}$ MC samples in the non-allhadronic decay modes, and $\text{BR}_{\text{non-allhad}}$ is the branching ratio of $t\bar{t}$ to non-allhadronic final states.

The likelihood function used in the fit is

$$\begin{aligned} L(\vec{n},\, \vec{\theta_0} \mid \mu,\, \vec{\theta}) \;\; = \;\; & \prod_{i}^{N_{\text{bins}}} \frac{(\mu \cdot s_i(\vec{\theta}) + b_i(\vec{\theta}))^{n_i} e^{-(\mu \cdot s_i(\vec{\theta}) + b_i(\vec{\theta}))}}{n_i!} \cdot \\ & \prod_{\theta_k \in \vec{\theta}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\theta_k - \theta_k^0)^2}{2\sigma^2} \right). \end{aligned} \tag{3.30}$$

In this expression, n_i is the observed number of events in bin i, $s_i(\vec{\theta})$ and $b_i(\vec{\theta})$ are the signal and background expectations, respectively (in bin i), and $\vec{\theta}$ represents the set of nuisance parameters (systematic uncertainties in case of the measurement with their central values $\vec{\theta_0}$). The second term in the product accounting for the Gaussian constraints on the nuisance parameters, ensuring the simultaneous treatment of systematic uncertainties. However, for positively defined parameters, such as luminosity and cross-section, a log-normal prior is used to maintain positivity. The signal strength μ represents the physics free parameter of interest. The best estimate for μ and $\vec{\theta}$ is obtained by maximizing the likelihood 3.30 using the TRExFitter package [94].

In total, 140 systematic uncertainties were considered in the analysis. For one particular uncertainty, symmetrization was applied as follows:

$$\text{UP}_{\text{symmetric}} = \text{NOMINAL} + \frac{\text{UP} - \text{DOWN}}{2} \tag{3.31}$$

$$\text{DOWN}_{\text{symmetric}} = \text{NOMINAL} - \frac{\text{UP} - \text{DOWN}}{2}. \tag{3.32}$$

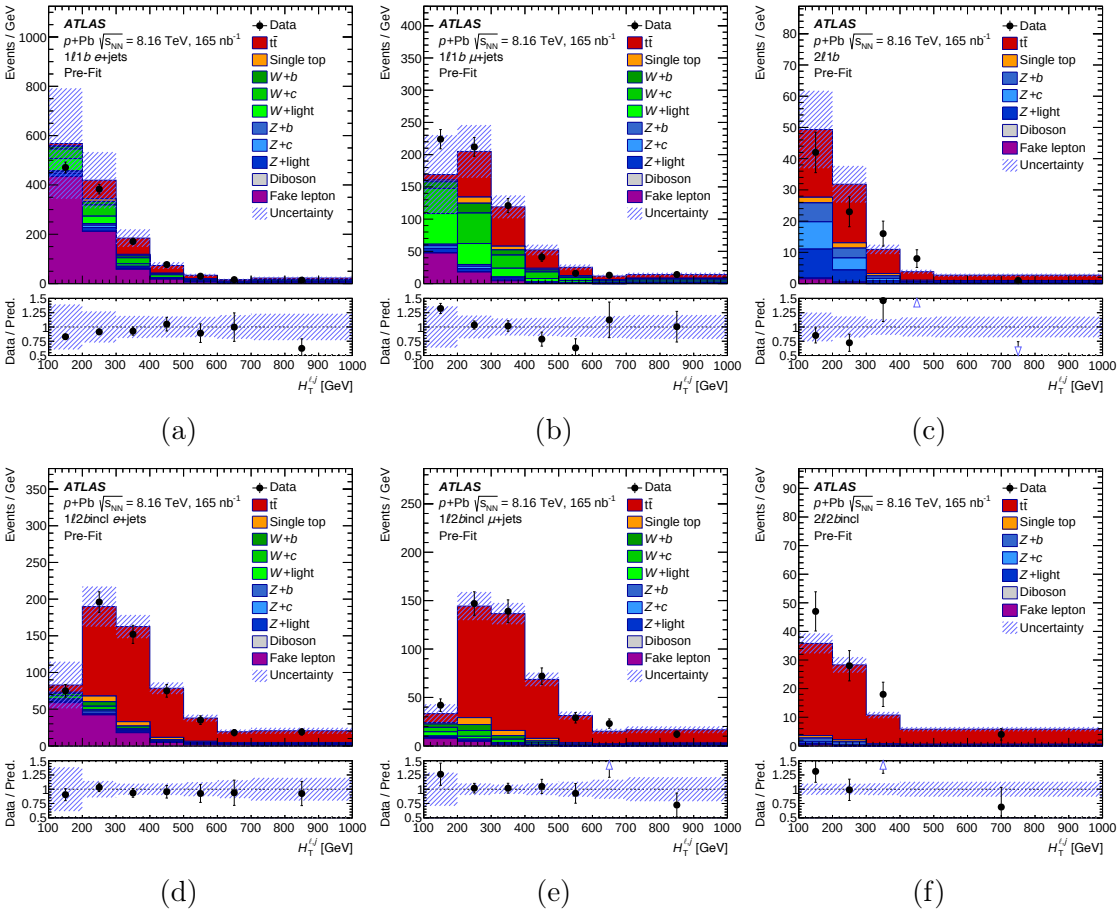


Figure 3.22: Comparison of data and total pre-fit prediction for the $H_{\mathrm{T}}^{\ell,j}$ distribution in each of the six SRs (e+jets: (a) $1\ell 1b$ and (d) $1\ell 2b$incl, μ+jets: (b) $1\ell 1b$ and (e) $1\ell 2b$incl, dilepton: (c) $2\ell 1b$ and (f) $2\ell 2b$incl), with total uncertainties in the prediction represented by the hatched area. The full markers in the bottom panels show a ratio of data and a sum of predictions. Open triangles indicate bins with entries which are outside the ratio range. The first and last bins include underflow and overflow events, respectively. The vertical order of the individual contributions forming the total prediction is the same as in the legend. The Z+jets contribution is negligible in the ℓ+jets$\geq 2b$ regions [48].

Additionally, one-sided UP and DOWN variations for uncertainties (e.g. single-top diagram removal/subtraction, $t\bar{t}$ modelling, and JES) were symmetrized using the formula:

$$\mathrm{UP}_{\mathrm{symmetric}} = \mathrm{NOMINAL} + \mathrm{UNC} \tag{3.33}$$

$$\mathrm{DOWN}_{\mathrm{symmetric}} = \mathrm{NOMINAL} - \mathrm{UNC} \tag{3.34}$$

To reduce computational time during the fit, a pruning technique was applied, which removes systematic variations that have an impact below 0.5% from the nominal value. Using this technique, approximately 25% of the systematic uncertainties were pruned.

During the analysis, the value of signal strength μ was blinded using the TRExFitter setup to ensure that the analyzers remained unbiased until all performance studies and the complete analysis chain were approved by a dedicated group of experts within the ATLAS Collaboration.

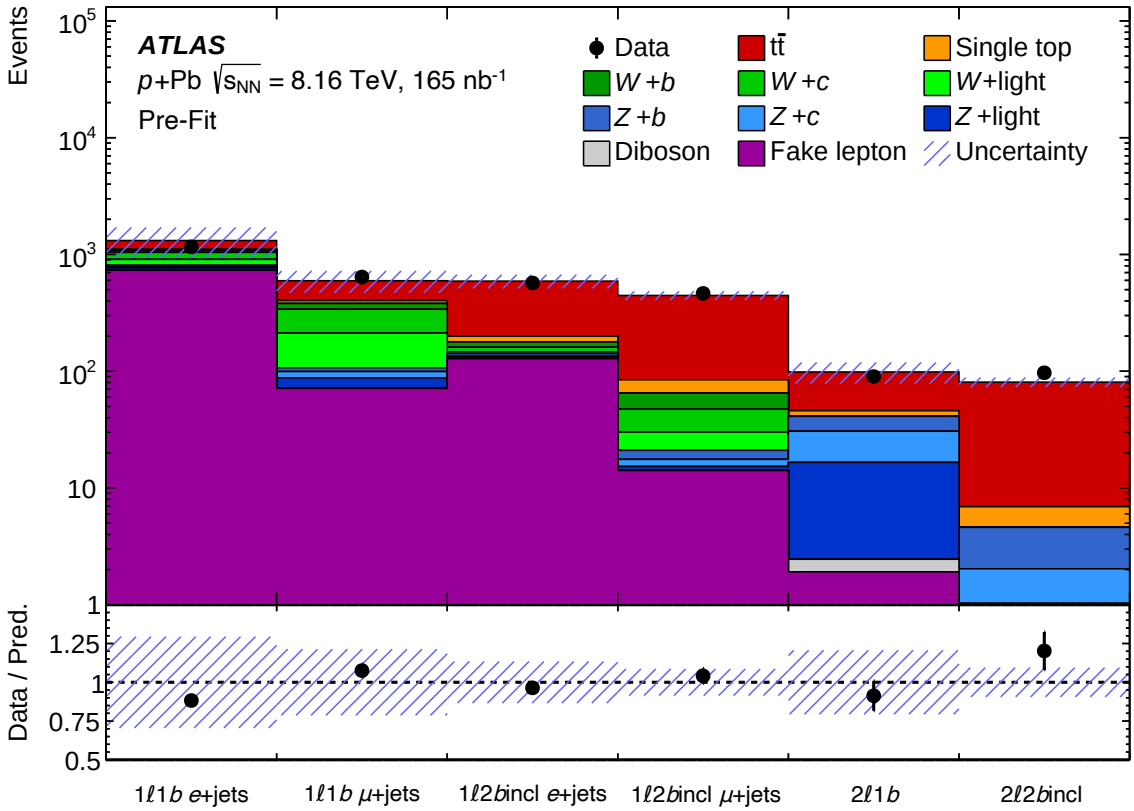
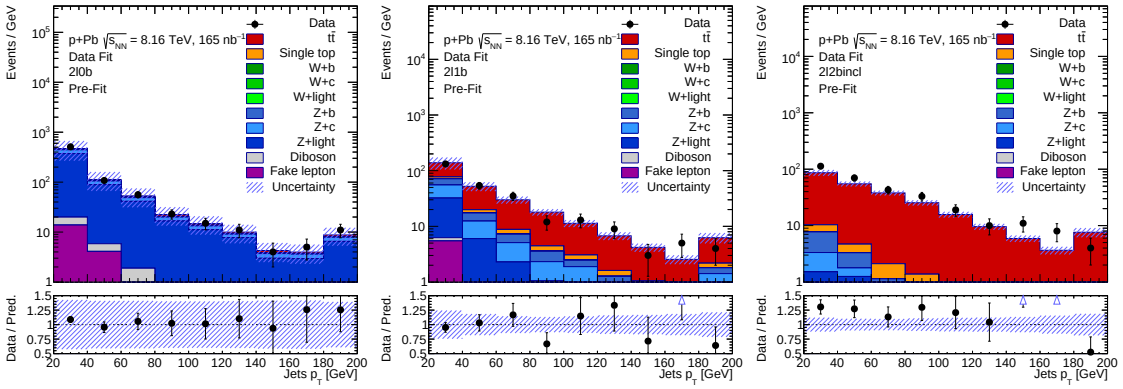


Figure 3.23: Data and predicted pre-fit event yields in each of the six signal regions. The background contributions, before fit, are depicted as filled histograms. The size of the combined statistical and systematic uncertainty in the sum of the signal and background is indicated by the blue hatched band. The ratio of the data to the sum of the signal and background is presented in the lower panel [48].

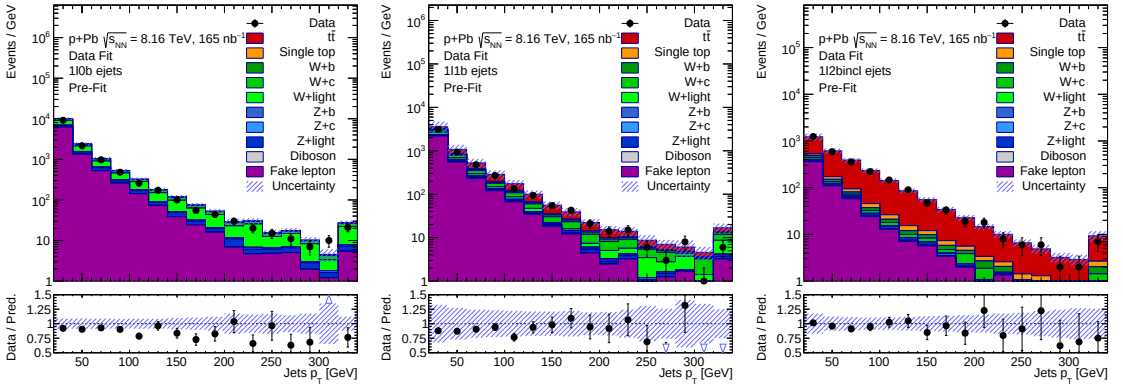
3.10.1 Control Plots

Before performing the fit, the agreement between data and MC simulations was studied across various kinematic observables and regions to ensure consistency within the uncertainties. Figures 5.28– 5.30 shows the distribution of the jet p_{T} for all channels. The control regions with zero b-tagged jets (e.g. in Figures 3.24a–3.26a) generally lack a significant portion of the systematic uncertainties due to their unavailability. However, they still played a crucial role in estimating the shape and normalization uncertainties of the fake background, as discussed in Section 3.9. The mass of the hadronically decaying top quark and W boson are shown in Figures 5.35–5.36, and Figures 5.37–5.38, respectively, showing the expected resonant behaviour with peaks around the top and W masses. Additional observables, such as lepton p_{T}, and missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$, are provided in Appendix L. A good agreement between data and MC simulations is observed within the uncertainties, with the exception of a few fluctuating data points.



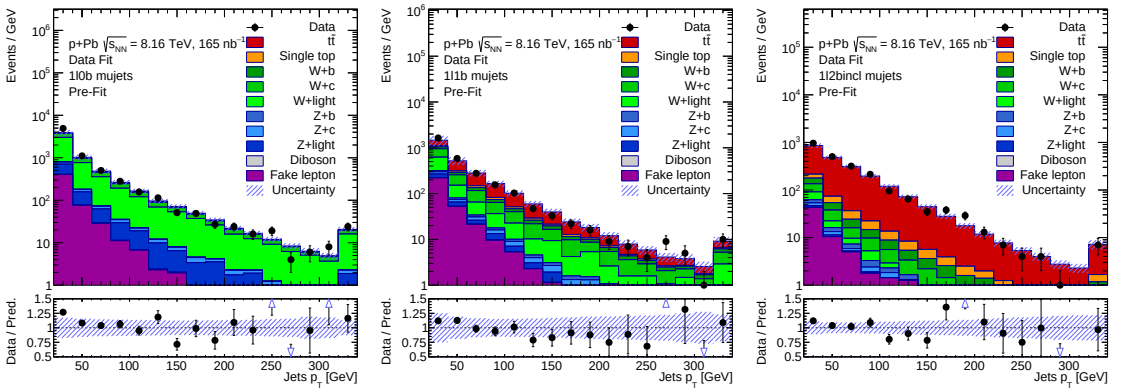
(a) (b) (c)

Figure 3.24: Pre-fit jet p_T distributions in dilepton channel: (a) $2\ell 0b$, (b) $2\ell 1b$, (c) $2\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



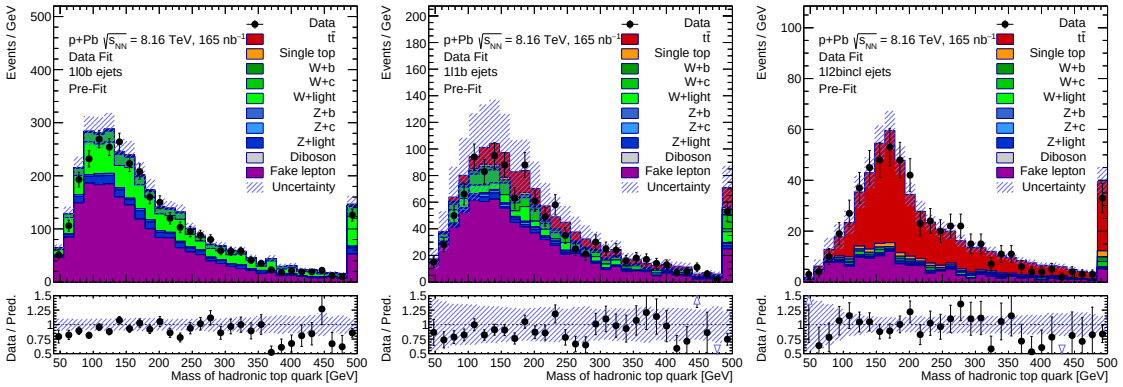
(a) (b) (c)

Figure 3.25: Pre-fit jet p_T distributions in ℓ+jets events (electron channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



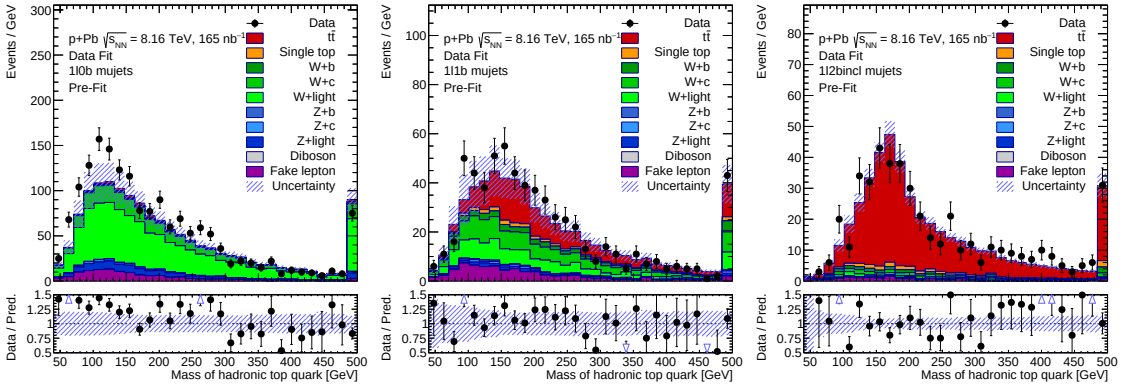
(a) (b) (c)

Figure 3.26: Pre-fit jet p_T distributions in ℓ+jets events (muon channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



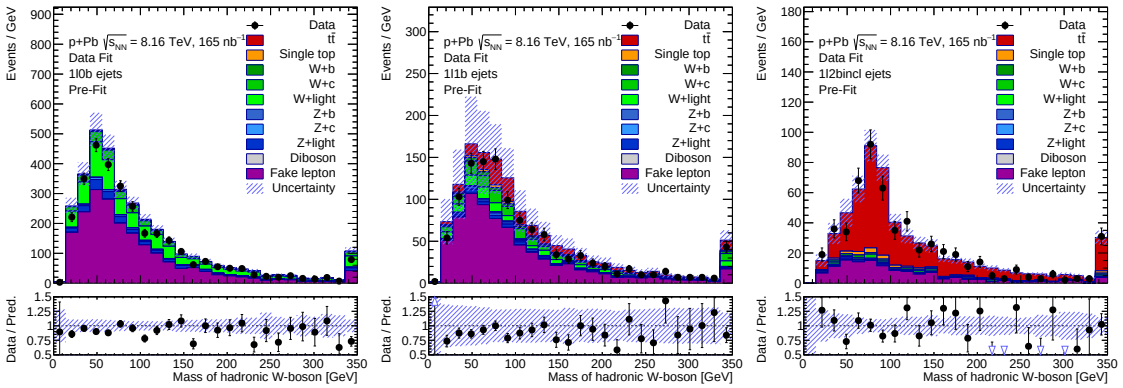
(a) (b) (c)

Figure 3.27: Pre-fit distribution of the mass of the hadronically decaying top quark in ℓ+jets events (electron channel): (a) 1ℓ0b, (b) 1ℓ1b, (c) 1ℓ2$b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



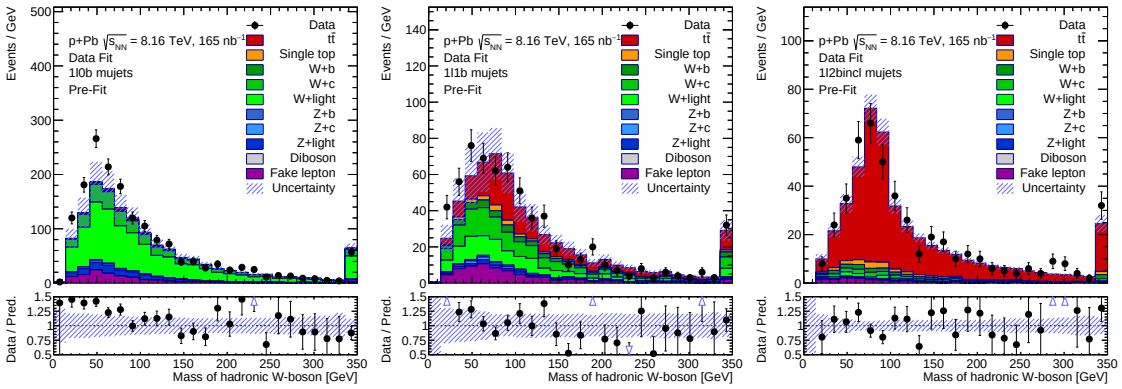
(a) (b) (c)

Figure 3.28: Pre-fit distribution of the mass of the hadronically decaying top quark in ℓ+jets events (muon channel): (a) 1ℓ0b, (b) 1ℓ1b, (c) 1ℓ2$b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



(a) (b) (c)

Figure 3.29: Pre-fit distribution of the mass of the hadronically decaying W in ℓ+ +jets events (electron channel): (a) 1ℓ0b, (b) 1ℓ1b, (c) 1ℓ2$b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



(a) (b) (c)

Figure 3.30: Pre-fit distribution of the mass of the hadronically decaying W in ℓ+ +jets events (muon channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.

3.11 Results

The TRExFitter package not only provides the fitted value of the signal strength μ, but also offers convenient options configurable via a configuration file. For this measurement it generates key pre-fit and post-fit plots for each systematic uncertainty, a correlation matrix of systematic uncertainties, and a ranking plot. The best fit value of the signal strength $\mu_{\text{fit}}^{\text{data}}$ is find to be

$$\mu_{\text{fit}}^{\text{data}} = 1.04^{+0.034}_{-0.035}\,(\text{stat.})^{+0.088}_{-0.087}\,(\text{syst.}) = 1.04^{+0.094}_{-0.094}. \tag{3.35}$$

A ranking plot in Figure 3.31 is used to assess the impact of different systematic uncertainties on the signal strength μ in the fit. It helps to identify which systematic uncertainties have the largest effect on the result. To construct a ranking plot a fit is performed to determine the best estimate of the signal strength μ with all systematic uncertainties included. Each systematic uncertainty is individually shifted up and down by its fitted uncertainty, while all other uncertainties are kept free. This assesses how much the μ changes due to this individual shift. Pre-fit impact shows how much each systematic uncertainty would affect μ without any constraint from the data while the post-fit impact displays the effect after the fit, where correlations between systematic uncertainties and the data's ability to constrain those uncertainties are accounted for. The post-fit impact is typically smaller as the fit constrains the nuisance parameters and therefore results in smaller uncertainties also for μ, a technique known as profiling. Systematics are ranked based on their contribution to the uncertainty in μ. The most impactful uncertainties appear at the top of the plot, and bars or markers indicate the magnitude of their influence. By comparing pre-fit and post-fit impacts, one can determine which systematic uncertainties are well-constrained by the fit and which remain dominant.

According to the ranking plot in Figure 3.31, the dominant systematic uncertainty is the fake-lepton background in the μ + jets 1b and e + jets 1b channels, which was conservatively set to 100% and 50%, respectively, based on the $E_{\text{T}}^{\text{miss}}$ control plot in Figure 3.32a. Therefore, corresponding nuisance parameters are constrained by the fit. Although the $E_{\text{T}}^{\text{miss}}$ variable is derived from PF jets, which were not used for kinematic measurements (HI jets were used for kinematics), the aim was to account for the observed disagreement. The fake-lepton background appears to be underestimated in the μ + jets 1b region and overestimated in the e + jets 1b region, therefore the corresponding nuisance parameters are shifted in opposite direction from each other in the ranking plot.

Since $E_{\text{T}}^{\text{miss}}$ is used as an inverted cut in the derivation of fake efficiencies, there is a possibility of non-orthogonality in the selection. As a result, an additional shape and normalization uncertainty was introduced to cover this potential issue. It is not surprising that these systematic uncertainties are the leading ones, given their strong dependence on background modeling and control region discrepancies. The fitted $H_{\text{T}}^{\ell,j}$ distributions are not strongly dependent of the $E_{\text{T}}^{\text{miss}}$ variable, except through the fake background estimation, which was one of the key reasons for using it as the primary fit distribution.

The effect of systematic uncertainties, sorted into groups based on their origin, is provided in Table 3.11, where jet energy scale, signal modeling, and fake background uncertainties are dominant.

Figure 3.34 shows the post-fit distributions of $H_{\text{T}}^{\ell,j}$ (yields given in Figure 3.35), which have smaller uncertainties compared to the pre-fit plots in Figure 3.22. Figure 3.36 shows correlations between systematic uncertainties, where e.g. the fake background shows the strongest correlation between e+jets and μ+jets channels in 1b region, see pre-fit plots in Figures 3.22a–3.22b and at the same time anti-correlates with W+light jets background systematic uncertainties as these compete against each other in the fit.

The total cross-section is calculated using Equation 3.29 as follows

Source	$\Delta\sigma_{t\bar{t}}/\sigma_{t\bar{t}}$	
	unc. up [%]	unc. down [%]
Jet energy scale	+4.6	-4.1
$t\bar{t}$ generator	+4.5	-4.0
Fake-lepton background	+3.1	-2.8
Background	+3.1	-2.6
Luminosity	+2.8	-2.5
Muon uncertainties	+2.3	-2.0
W+jets	+2.2	-2.0
b-tagging	+2.1	-1.9
Electron uncertainties	+1.8	-1.5
MC statistical uncertainties	+1.1	-1.0
Jet energy resolution	+0.4	-0.4
$t\bar{t}$ PDF	+0.1	-0.1
Systematic uncertainty	+8.3	-7.6

Table 3.11: Summary of the impact of the systematic uncertainties on the $t\bar{t}$ cross-section measurement grouped into different categories. The quoted uncertainties are obtained by repeating the fit with a group of nuisance parameters fixed to their fitted values and subtracting in quadrature the resulting total uncertainty from the uncertainty of the complete fit. However, the total uncertainty is not the quadratic sum of the grouped impacts, as this approach neglects the correlation among the different groups. [90]

$$\begin{aligned}
\sigma_{t\bar{t}}^{\text{measured}} &= \frac{\mu_{\text{fit}}^{\text{data}}}{\text{BR}_{\text{non-allhad}}}\,\sigma_{\text{MC}}^{t\bar{t},\text{non-allhad}} \\
&= \frac{1.04168}{0.543} \times 127.8212 \times 1.139 \times 208 \\
&= 58093.09 \text{ pb} \\
&= 58.1 \pm 2.0 \text{ (stat.)} \,^{+4.8}_{-4.4} \text{ (syst.) nb}
\end{aligned} \tag{3.36}$$

Here, the NLO non-all-hadronic $pp \to t\bar{t}$ cross-section $\sigma_{\text{NLO}} = 127.8212\,\text{pb}$, evaluated at $\sqrt{s_{\text{NN}}} = 8.16\,\text{TeV}$, with a k-factor of 1.139, is scaled to the p+Pb system by the number of nucleons $A = 208$. The non-all-hadronic branching $\text{BR}_{\text{non-allhad}}$ ratio is taken as $1 - \text{BR}_{\text{allhad}} = 1 - 0.457$ [95], accounting to 0.543. The relative total uncertainty of the measured cross-section is 9%.

Figure 3.37a presents a comparison between the observed $\sigma_{t\bar{t}}$ and the CMS measurement in p+Pb collisions at $\sqrt{s_{\text{NN}}} = 8.16$ TeV [96]. The two results are consistent within 1.4 standard deviations. Additionally, the figure includes the most precise measurement of the $t\bar{t}$ production cross-section in pp collisions at $\sqrt{s} = 8$ TeV from the ATLAS and CMS combination [97]. This value is extrapolated to the centre-of-mass energy used in this study using the Top++ v2 prediction, and scaled by A_{Pb} to account for the p+Pb system. The extrapolated cross-section carries a 2.5% relative uncertainty and is independent of any nPDF assumptions. The extrapolation factor is calculated to be 1.0528 ± 0.0005 (PDF) $^{+0.0001}_{-0.0013}$ (scale). Furthermore, the measured cross-section is compared with NLO calculations generated by MCFM [98], scaled to NNLO precision in QCD using the same k-factor, obtained from the Top++ v2 generator. Four different nPDF sets were used as input to the MCFM calculations: EPPS21 [99], nCTEQ15HQ [36, 100], nNNPDF30 [101, 102], and TUJU21 [103] [90].

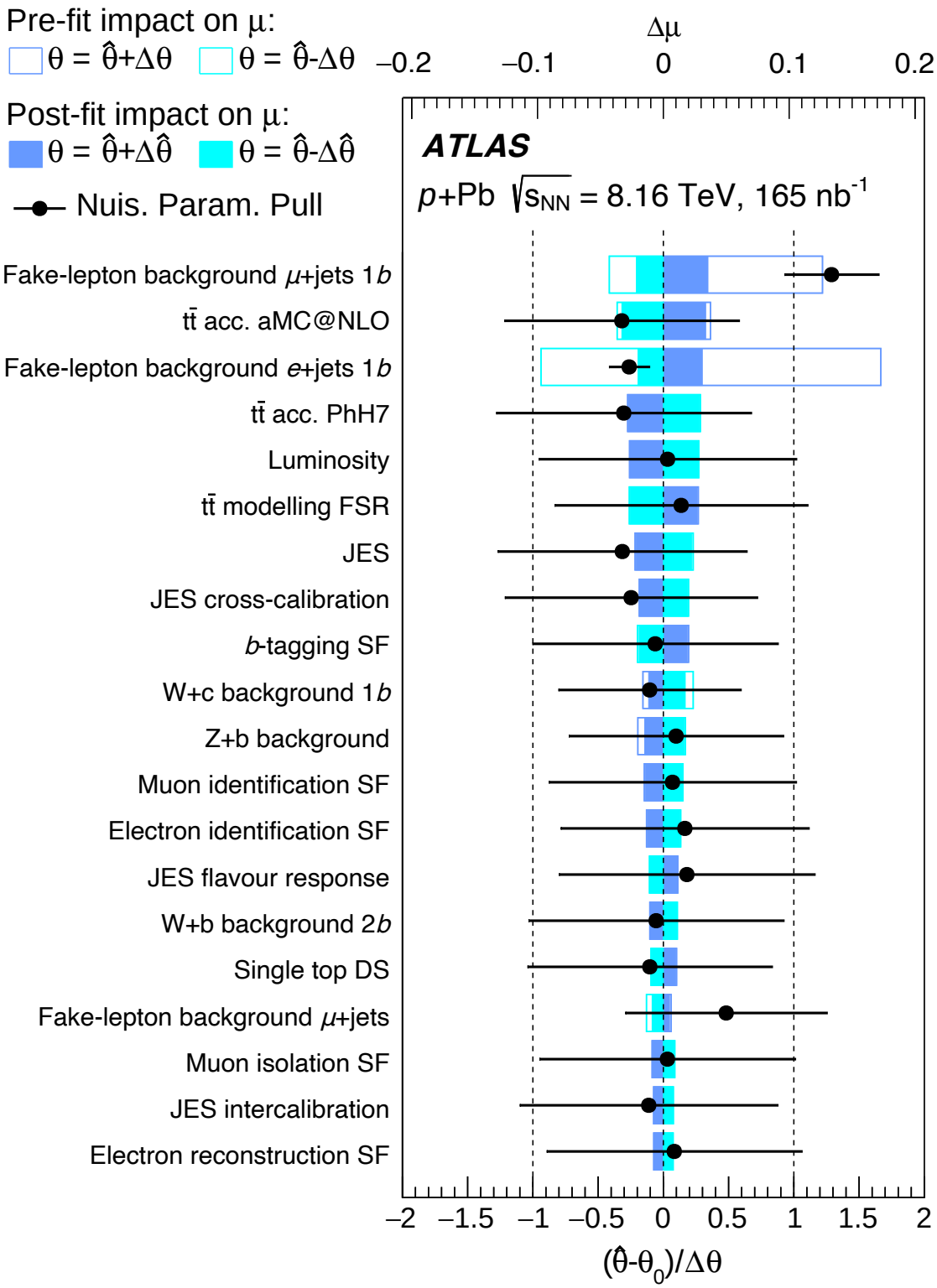
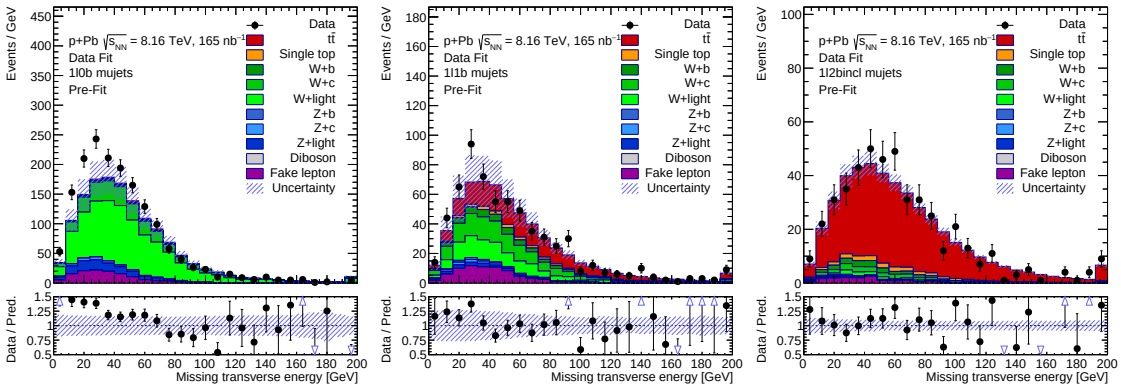
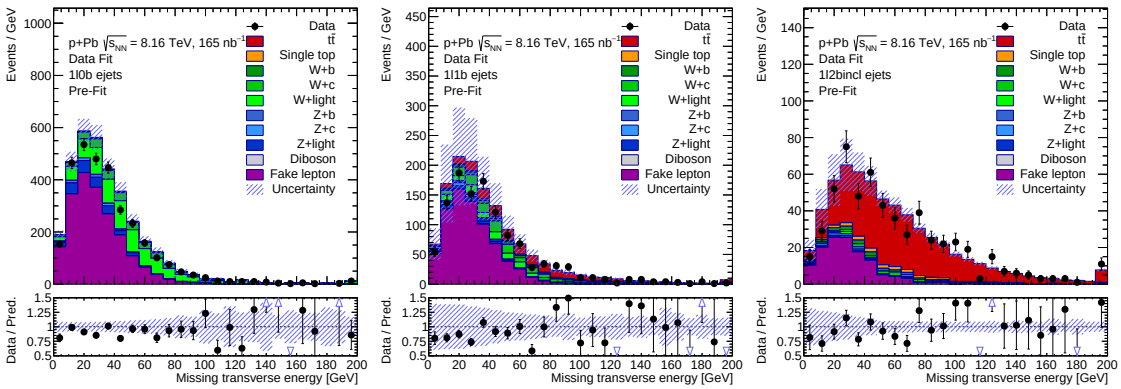


Figure 3.31: The impact of systematic uncertainties on the fitted signal-strength parameter $\hat{\mu}$ for the combined fit of all channels is shown. Filled blue boxes represent the variations of $\hat{\mu}$ from the central value, $\Delta\hat{\mu}$ (top x-axis), when fixing the corresponding individual nuisance parameter, θ, to its post-fit value $\hat{\theta}$, modified upwards or downwards by its post-fit uncertainty, and repeating the fit. Empty blue boxes show the corresponding pre-fit impact. Black points represent the fitted values and uncertainties of the nuisance parameters relative to their pre-fit values, θ_0, and uncertainties, $\Delta\theta$ (bottom x-axis). Black lines represent the post-fit uncertainties of the nuisance parameters relative to their nominal uncertainties, as indicated by the dashed line [48].



(a) (b) (c)

Figure 3.32: Pre-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in ℓ+jets events (muon channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



(a) (b) (c)

Figure 3.33: Pre-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in ℓ+jets events (electron channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.

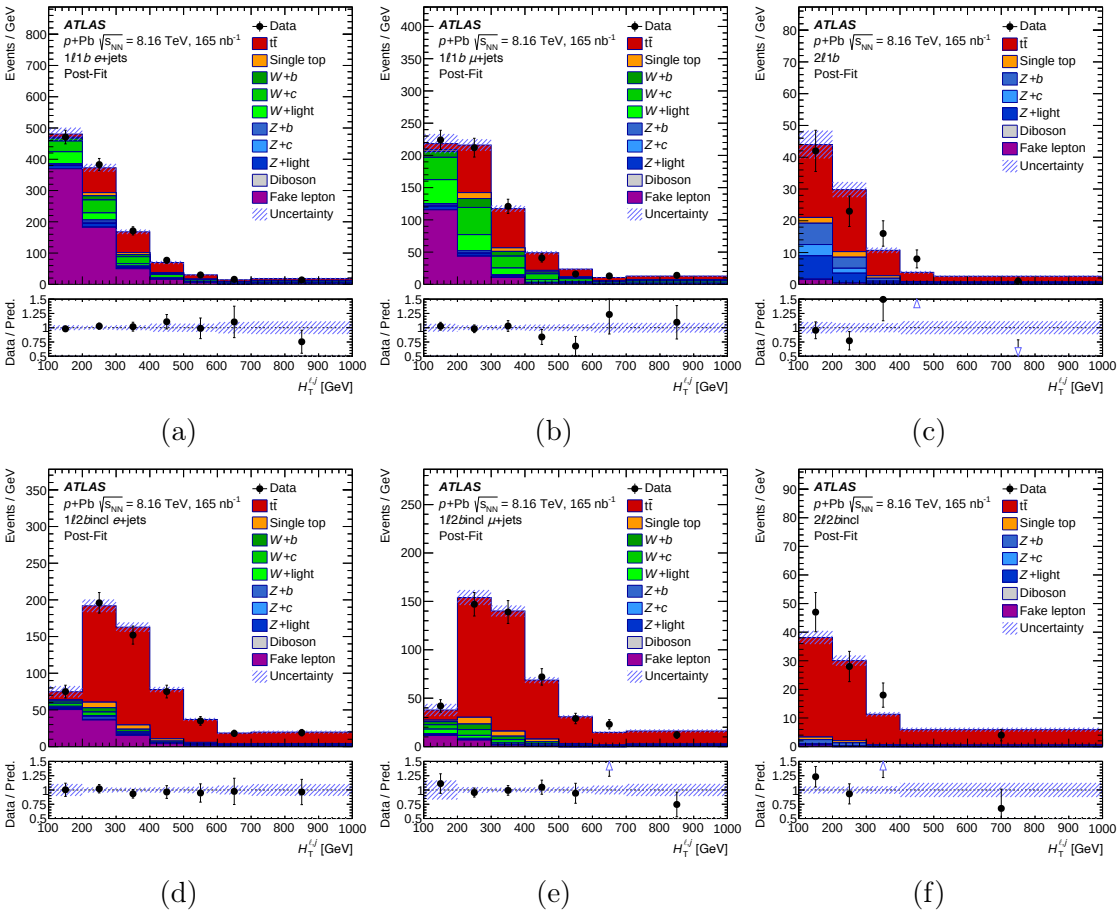


Figure 3.34: Comparison of data and total post-fit prediction for the $H_{\mathrm{T}}^{\ell,j}$ distribution in each of the six SRs (e+jets: (a) $1\ell 1b$ and (d) $1\ell 2b$incl, μ+jets: (b) $1\ell 1b$ and (e) $1\ell 2b$incl, dilepton: (c) $2\ell 1b$ and (f) $2\ell 2b$incl), with total uncertainties in the prediction represented by the hatched area. The full markers in the bottom panels show a ratio of data and a sum of predictions. Open triangles indicate bins with entries which are outside the ratio range. The first and last bins include underflow and overflow events, respectively. The vertical order of the individual contributions forming the total prediction is the same as in the legend. The Z+jets contribution is negligible in the ℓ+jets$\geq 2b$ regions [48].

	$1\ell 1b$ e+jets	$1\ell 1b$ μ+jets	$1\ell 2b$incl e+jets	$1\ell 2b$incl μ+jets	$2\ell 1b$	$2\ell 2b$incl
$t\bar{t}$	214 ± 24	194 ± 21	405 ± 21	373 ± 19	55 ± 6	79 ± 5
t-channel	6.9 ± 1.0	6.4 ± 1.0	7.7 ± 0.9	7.1 ± 0.9	0 ± 0	0 ± 0
W+b	37 ± 19	37 ± 19	16 ± 8	17 ± 9	–	–
W+c	120 ± 40	110 ± 40	14 ± 7	17 ± 8	–	–
W+light	80 ± 40	80 ± 40	4.8 ± 3.1	9 ± 5	–	–
Z+b	16 ± 13	8 ± 7	8 ± 7	3.7 ± 3.0	12 ± 9	2.9 ± 2.4
Z+c	9 ± 14	5 ± 7	1.7 ± 2.6	0.9 ± 1.4	6 ± 9	0.4 ± 0.6
Z+light	28 ± 16	12 ± 7	1.2 ± 1.1	0.9 ± 0.5	11 ± 6	0.34 ± 0.25
Diboson	0.32 ± 0.16	0.29 ± 0.15	0.055 ± 0.029	0.039 ± 0.02	0.53 ± 0.27	0.049 ± 0.025
tW	17.1 ± 3.0	15.5 ± 2.7	13.6 ± 3.2	12.1 ± 2.9	5.1 ± 2	2.4 ± 1.2
Fake lepton	630 ± 50	170 ± 40	110 ± 19	21 ± 12	1.9 ± 1	0.51 ± 0.27
Total	1154 ± 34	648 ± 24	582 ± 21	462 ± 18	91 ± 7	85 ± 5
Data	1162	641	570	464	90	97

Figure 3.35: Data and predicted post-fit event yields in each of the six signal regions. The total uncertainty is a quadrature sum of statistical and systematic uncertainties. Owing to rounding effects and small correlations between the different sources of uncertainties, the total systematic uncertainty is different from the sum in quadrature of the individual sources [48].

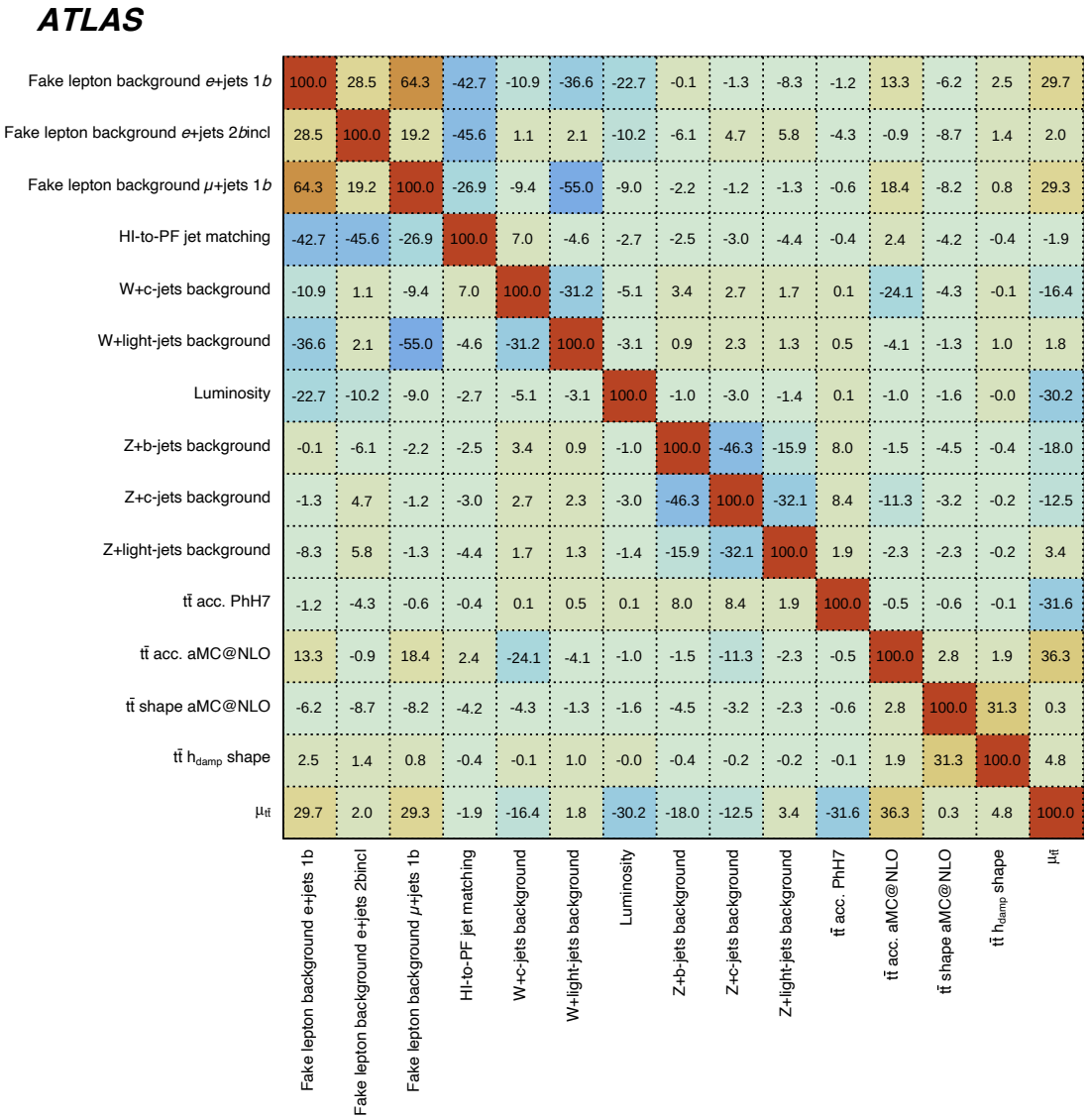


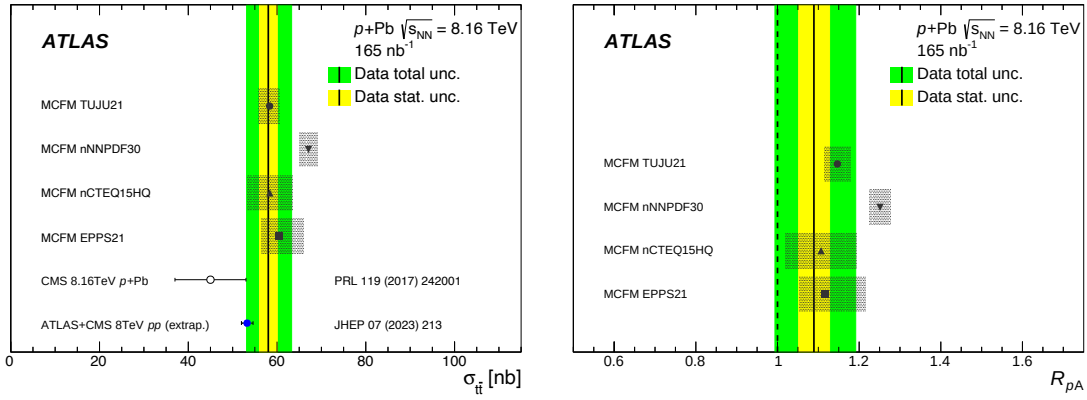
Figure 3.36: Correlation matrix of the fit parameters for the combined fit to data: only parameters with at least one correlation factor greater than 30% are shown [48].

3.11.1 Nuclear Modification Factor

A nuclear modification factor as a single number is defined as the ratio of the cross-section for the $t\bar{t}$ production in proton-lead p+Pb collisions to the cross-section for the $t\bar{t}$ production in proton-proton pp collisions, scaled by the atomic mass number A_{Pb} of the lead nucleus.

$$\begin{aligned} R_{p\mathrm{A}} &= \frac{\sigma_{t\bar{t}}^{p+\mathrm{Pb}}}{A_{\mathrm{Pb}} \cdot \sigma_{t\bar{t}}^{pp}} \\ &= \frac{58093.09}{208 \cdot 243.3 \cdot 1.0528} \\ &= 1.090 \ \pm \ 0.039 \ (\mathrm{stat.}) \ ^{+0.094}_{-0.087} \ (\mathrm{syst.}), \end{aligned} \tag{3.37}$$

where the $\sigma_{t\bar{t}}^{p+\mathrm{Pb}}$ is the measured value of the $t\bar{t}$ cross-section in p+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV and $\sigma_{t\bar{t}}^{pp}$ is the measured $t\bar{t}$ cross-section in pp collisions at $\sqrt{s} = 8$ TeV [97] extrapolated to the same centre-of-mass energy $\sqrt{s} = 8.16$ TeV by factor 1.0528 and scaled by number of nucleons in lead (208). A comparison between the measured $R_{p\mathrm{A}}$ from data and simulation is shown in Figure 3.37b, where R_{pA} was calculated at NNLO precision using the MCFM code [98] scaled to the p+Pb system for four different nPDF sets. The uncertainty related to the baseline PDF for pp interactions is treated as fully correlated across the predictions and, as a result, cancels out in the ratio. The remaining uncertainty reflects the contribution from the nPDF. All nPDF calculations yield R_{pA} values greater than unity.



(a) (b)

Figure 3.37: Comparison between the measured and predicted values of (a) $\sigma_{t\bar{t}}$ and (b) R_{pA}. $\sigma_{t\bar{t}}$ is also compared with the existing measurement in p+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV [96], and the combined measurement of the $t\bar{t}$ production cross-section in pp collisions at $\sqrt{s} = 8$ TeV from ATLAS and CMS collaborations [97]. The latter is extrapolated to the centre-of-mass energy of this measurement and scaled by the factor $A_{\mathrm{Pb}} = 208$. Predictions are calculated at NLO precision using the MCFM code [98] scaled to the p+Pb system and by the k-factor 1.139, given for different nPDF sets. The uncertainty in the predictions represents the internal PDF uncertainty. The solid black line indicates the measured value. The combined statistical and systematic uncertainty of the measurement is represented by the outer band around the central value, while the statistical component is depicted as the inner band [48].

3.11.2 Observation in the Dilepton Channel

In Figure 3.38 the signal strength μ measured in each of the six fitted regions individually is shown while the last row represents the combined fitted value of μ. The fitted values of

μ in the individual channels are consistent with each other within uncertainties and align with the Standard Model prediction. The precision of the μ measurement is primarily limited by systematic uncertainties in the ℓ+jets signal regions (SRs), while statistical uncertainties dominate in the dilepton SRs.

The significance is determined by performing separate fits of μ in the combined four ℓ+jets SRs and the combined two dilepton SRs, with both exceeding five standard deviations. This confirms the observation of the $t\bar{t}$ production in both the ℓ+jets and dilepton channels separately, with the latter being observed for the first time in p+Pb collisions at the LHC.

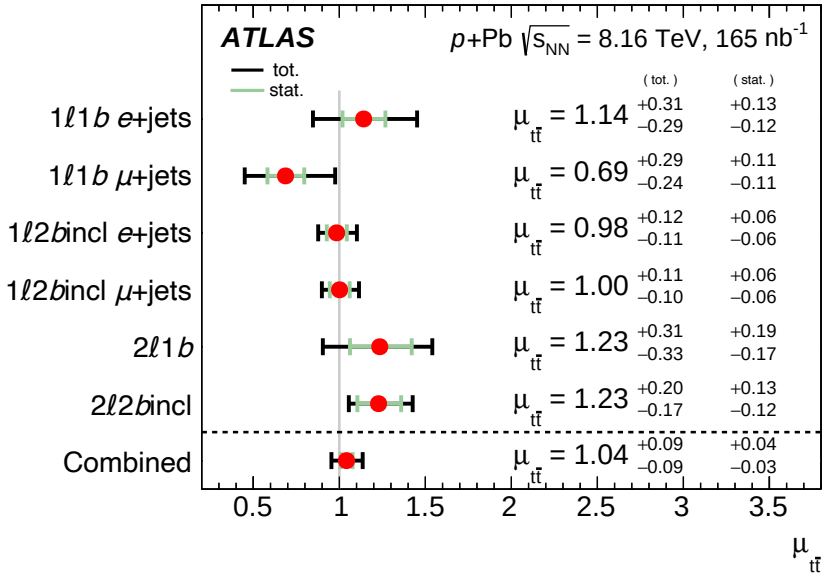


Figure 3.38: The observed best-fit values of the signal strength $\mu_{t\bar{t}}$ and their uncertainties by final-state category and combined. The individual $\mu_{t\bar{t}}$ values for the channels are obtained from a simultaneous fit with the signal-strength parameter for each channel allow to vary independently. The SM prediction is $\mu_{t\bar{t}} = 1$ [48].

3.11.3 Differences to the CMS Measurement

The CMS Collaboration released publication of the measurement of the $t\bar{t}$ production at 8.16 TeV in the ℓ+jets channel [96] in 2017 using a fit in $m_{jj'}$ distribution as shown in Appendix 5.41. Table 3.12 summarizes main differences in analysis conditions, where the ATLAS experiment gains in statistics using lower cuts on lepton and jet transverse momenta, larger phase-space in pseudorapidity η, and by using the dilepton channel. The CMS recorded slightly higher luminostity, however, due to wider selection cuts and other effects, the signal yield in ℓ+jets channel is about 67% higher at ATLAS analysis compared to CMS. Table 3.13 presents a comparison of uncertainties between the ATLAS and CMS measurements, showing that the ATLAS experiment achieved overall twice the precision of CMS.

Analysis conditions	ATLAS	CMS
Channel	lepton+jets, dilepton	lepton+jets
Fitted Distribution	$H_{\mathrm{T}}^{\ell,j}$	$m_{jj'}$
Min lepton p_T [GeV]	18	30
Lepton $\|\eta\|$	< 2.4 (2.47)	< 2.1
Min jet p_T [GeV]	20	25
Integrated luminosity [nb^{-1}]	164.6	174

Table 3.12: Comparison of analysis conditions between ATLAS and CMS of the $t\bar{t}$ production in proton-lead collisions at 8.16 TeV.

Uncertainties [%]	ATLAS	CMS
Integrated luminosity	2.4	5
Statistical	3	8
B-tagging + JES	5	13
Extra JES	-	4
Background	4	7
Lepton trigger and reconstruction	3	4
Total uncertainty	9	18

Table 3.13: Comparison of uncertainties between the ATLAS and CMS experiments in the $t\bar{t}$ production cross-section in proton-lead collisions at 8.16 TeV.

3.12 Conclusion

This thesis presents the measurement of top quark pair production in proton-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV using data recorded by the ATLAS detector. This measurement explores the behavior of the top quark, the heaviest particle in the Standard Model, in a unique nuclear environment. By studying $t\bar{t}$ production in p+Pb collisions, we probe nuclear parton distribution functions (nPDFs), especially at high Bjorken-x, where phenomena like shadowing and anti-shadowing may play a role.

The analysis focuses on both the lepton+jets and dilepton channels, marking the first observation of the $t\bar{t}$ production in the dilepton channel in p+Pb collisions. The measured cross-section of the top quark pair production is $\sigma_{t\bar{t}} = 58.1 \pm 2.0$ (stat.) $^{+4.8}_{-4.4}$ (syst.) nb, with a total relative uncertainty of 9%. This result represents the most precise measurement of the $t\bar{t}$ cross-section in heavy-ion collisions to date. Systematic uncertainties dominate the precision in the lepton+jets channel, while statistical uncertainties are the primary limitation in the dilepton channel.

This measurement aligns with earlier observations made by the CMS collaboration, but extends the study into a new decay channel. By combining both the lepton+jets and dilepton channels, this study provides a more comprehensive understanding of the $t\bar{t}$ production in heavy-ion collisions. The result can be used for future fits of nPDFs and parton distribution functions in a previously unexplored kinematic region.

In addition to providing valuable insights into the nuclear environment, this analysis highlights the importance of systematic control, such as b-jet tagging techniques and optimized lepton selection criteria, which significantly improved the measurement precision. These findings will contribute to ongoing efforts to better understand the behavior of quarks and gluons in high-energy collisions and provide a stepping stone for future investigations in both proton-lead and lead-lead collisions.

Author's Contributions

The author contributed the most to the tasks described in the Muon Performance Section 3.6 which was his qualification task to become an author within the ATLAS collaboration, and topics in Jet Matching Section 3.3.5 which he developed and implemented within the analysis code. He futher provided cross-checks related to the content of the Fake Background Estimation Section 3.5 where matching between HI and PF jets had to be implemented as well, in the Fit Procedure Section 3.10 provided independent runs and implemented and provided validation plots. The author took over the analysis code and was responsible for the final ntuple production. The author has fixed various issues aring from ATLAS analysis softwere Athena code, to adapt smooth processing of MC samples with the data-based pile-up overlay.

The author would like to thank to his colleagues Patrycja Potępa, Yuriy Volkotrub, and Santu Mondal for providing results in Sections 3.5, 3.7, 3.8, 3.10 under supervision and leadership of prof. Iwona Grabowska-Bołd, PhD. and Mgr. Jiří Kvita, PhD. and help of other experts from the ATLAS collaboration. The measurement lead to the publication [90].

In addition to the work on measuring top-quark pair production in proton-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV using data from the ATLAS detector, the author continued his research after completing his master's studies. He contributed to the analysis of variations in the single top and W+jets backgrounds (see Appendix N) within the measurements of top-quark pair differential and double-differential cross-sections in the ℓ+jets channel using pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector [104]. His contributions earned him exceptional authorship in the ATLAS Collaboration prior to the formal assignment of his qualification task. Another study focused on extending Fully Bayesian Unfolding by regularization (see Appendix O) is summarized in the publication [105], while the publication investigating machine learning applications of top quark and W boson tagging in four top quark final states is currently under submission process [106] (see Appendix P).

Appendix

A Calculation of Traces using Mathematica

```
Import["https://raw.githubusercontent.com/FeynCalc/feyncalc/master/install.m"];
InstallFeynCalc[]

CalculateTrace[expr_] := Module[
  {simplifiedExpr, traceResultSubstituted, simplifiedFinal},
  simplifiedExpr = DiracSimplify[expr];
  traceResultSubstituted = simplifiedExpr /. {
        Pair[Momentum[p], Momentum[pc]] -> (s - 2 m^2)/2,
        Pair[Momentum[kc], Momentum[k]] -> s/2,
        Pair[Momentum[k], Momentum[p]] -> (-t + m^2)/2,
        Pair[Momentum[kc], Momentum[pc]] -> (-t + m^2)/2,
        Pair[Momentum[k], Momentum[pc]] -> (-u + m^2)/2,
        Pair[Momentum[kc], Momentum[p]] -> (-u + m^2)/2,
        Pair[Momentum[p], Momentum[p]] -> m^2,
        Pair[Momentum[pc], Momentum[pc]] -> m^2,
        Pair[Momentum[k], Momentum[k]] -> 0,
        Pair[Momentum[kc], Momentum[kc]] -> 0,
        s + t + u -> 2*m^2
  };
  simplifiedFinal = Simplify[DiracSimplify[
     traceResultSubstituted], s = -t - u + 2*m^2];
  simplifiedFinal
];

MFI1 = DiracTrace[(DiracSlash[pc] + m) . DiracMatrix[mu] . (
   DiracSlash[pc] - DiracSlash[kc] + m) . DiracMatrix[nu] . (
   DiracSlash[p] - m) . DiracMatrix[nu] . (DiracSlash[pc] -
   DiracSlash[kc] + m) . DiracMatrix[mu]];
MFI2 = DiracTrace[(DiracSlash[pc] + m) . DiracMatrix[nu] . (
   DiracSlash[kc] - DiracSlash[p] + m) . DiracMatrix[mu] . (
   DiracSlash[p] - m) . DiracMatrix[mu] . (DiracSlash[kc] -
   DiracSlash[p] + m) . DiracMatrix[nu]];
M0[nu0_, mu0_, sigma_, k0_, k1_, k2_] := MT[nu0, sigma] FV[k0
   - k1, mu0] + MT[nu0, mu0] FV[k1 - k2, sigma] + MT[mu0,
   sigma] FV[k2 - k0, nu0];
M1[nu0_, mu0_, sigma_, k0_, k1_, k2_] := MT[nu0, mu] FV[k0 -
   k1, sigma] + MT[nu0, sigma] FV[k1 - k2, mu] + MT[sigma, mu
   ] FV[k2 - k0, nu0];
MFI3 = DiracTrace[((DiracSlash[pc] + m) . DiracMatrix[sig] . (
   DiracSlash[p] - m) . DiracMatrix[del])] . M0[mu, nu, sig,
   kc, k, -(kc + k)] . M0[mu, nu, del, k, kc, -(kc + k)];
MFI4 = DiracTrace[(DiracSlash[pc] + m) . DiracMatrix[mu] . (
   DiracSlash[pc] - DiracSlash[kc] + m) . DiracMatrix[nu] . (
```

```
            DiracSlash[p] - m) . DiracMatrix[mu] . (DiracSlash[kc] -
            DiracSlash[p] + m) . DiracMatrix[nu]];
30          MFI5 = DiracTrace[((DiracSlash[pc] + m) . DiracMatrix[mu] . (
            DiracSlash[pc] - DiracSlash[kc] + m) . DiracMatrix[nu] . (
            DiracSlash[p] - m) . DiracMatrix[del])] . M0[mu, nu, del,
            kc, k, -(kc + k)];
31          MFI6 = DiracTrace[((DiracSlash[pc] + m) . DiracMatrix[nu] . (
            DiracSlash[kc] - DiracSlash[p] + m) . DiracMatrix[mu] . (
            DiracSlash[p] - m) . DiracMatrix[del])] . M0[mu, nu, del,
            kc, k, -(kc + k)];
32          MFI7 = DiracTrace[(DiracSlash[pc] + m) . DiracSlash[k] . (
            DiracSlash[p] - m) . DiracSlash[k]];
33          MFI8 = DiracTrace[(DiracSlash[pc] + m) . DiracSlash[kc] . (
            DiracSlash[p] - m) . DiracSlash[kc]];
34
35          resultMFI1 = (g^4) * TTR * Simplify[CalculateTrace[MFI1] / 4]
            / (Nn * (t - m^2)^2);
36          resultMFI2 = (g^4) * TTR * Simplify[CalculateTrace[MFI2] / 4]
            / (Nn * (u - m^2)^2);
37          resultMFI3 = CCA * (g^4) * TTR * Simplify[CalculateTrace[MFI3]
             / 4] / ((Nn^2 - 1) * s^2);
38          resultMFI4 = -(g^4) * TTR^2 * Simplify[CalculateTrace[MFI4] /
            4] / (((Nn - 1) * Nn) * (t - m^2) * (u - m^2));
39          resultMFI5 = -(g^4) * TTR * CCA * Simplify[CalculateTrace[MFI5
            ] / 4] / ((Nn^2 - 1) * s * (t - m^2));
40          resultMFI6 = -(g^4) * TTR * CCA * Simplify[CalculateTrace[MFI6
            ] / 4] / ((Nn^2 - 1) * s * (u - m^2));
41          resultMFI7 = (g^4) * TTR * CCA * Simplify[CalculateTrace[MFI7]
             / 4] / ((Nn^2 - 1) * s);
42          resultMFI8 = (g^4) * TTR * CCA * Simplify[CalculateTrace[MFI8]
             / 4] / ((Nn^2 - 1) * s);
```

B Calculation of Hadronic Cross-Sections

B.1 Python Code: LO Cross-Section Calculation

```
import lhapdf
import scipy.integrate as spi
import numpy as np

# Constants and variables
mt = 172.5  # Top quark mass in GeV
S = 8160**2  # Total energy squared
s_min = 4 * mt**2
s_max = S
Q = mt  # Scale in GeV (top quark mass)
conversion_factor = 0.38937936561 * 10**9  # GeV^2 to picobarns

# Load the LHAPDF set (e.g., CT18LO)
pdf = lhapdf.mkPDF("CT18LO", 0)  # Central member

# Redefine sigma_gg using dynamically running or fixed alpha_s
def sigma_gg(s):
    rho = 4 * mt**2 / s
    beta = np.sqrt(1 - rho)
    alpha_s = pdf.alphasQ(Q)  # Fixed alpha_s at the top quark mass scale (can also use dynamic)
    return (np.pi * alpha_s**2 / (48 * s)) * ((2 * rho**2 + 32 * rho + 32) * np.arctanh(beta) - beta * (31 * rho + 28))

# Define the gluon-gluon integrand using sigma_gg and LHAPDF for gluon PDFs
def integrand_gg(x1, x2):
    s = x1 * x2 * S  # Partonic center-of-mass energy squared
    if s < s_min or s > s_max:
        return 0
    fa = pdf.xfxQ(21, x1, Q) / x1  # Gluon PDF for x1 at scale Q
    fb = pdf.xfxQ(21, x2, Q) / x2  # Gluon PDF for x2 at scale Q
    return fa * fb * sigma_gg(s)

# Perform the double integration over x1 and x2 for gg contribution
result_gg, error_gg = spi.dblquad(
    integrand_gg,
    0, 1,  # x1 limits
    lambda x1: 4 * mt**2 / (S * x1), lambda x1: 1  # x2 limits depend on x1
)

# Convert to picobarns
result_gg_scaled = result_gg * conversion_factor
error_gg_scaled = error_gg * conversion_factor

# Redefine sigma_qq using dynamically running or fixed alpha_s
def sigma_qq(s):
    rho = 4 * mt**2 / s
    beta = np.sqrt(1 - rho)
    alpha_s = pdf.alphasQ(Q)  # Fixed alpha_s at the top quark mass scale (can also use dynamic)
```

```
    return (4 * np.pi * alpha_s**2 * beta * (rho + 2)) / (27 * s)

# Define the quark-antiquark integrand using sigma_qq and LHAPDF for quark PDFs
def integrand_qq(x1, x2):
    s = x1 * x2 * S
    if s < s_min or s > s_max:
        return 0
    # Sum of quark PDFs (1 to 5) for x1 and antiquark PDFs (-1 to -5) for x2
    quark_pdf_sum_x1 = sum([pdf.xfxQ(flavor, x1, Q) for flavor in [1, 2, 3, 4, 5]]) / x1
    antiquark_pdf_sum_x2 = sum([pdf.xfxQ(-flavor, x2, Q) for flavor in [1, 2, 3, 4, 5]]) / x2
    return quark_pdf_sum_x1 * antiquark_pdf_sum_x2 * sigma_qq(s)

# Perform the double integration over x1 and x2 for qq contribution
result_qq, error_qq = spi.dblquad(
    integrand_qq,
    0, 1,  # x1 limits
    lambda x1: 4 * mt**2 / (S * x1), lambda x1: 1  # x2 limits depend on x1
)

# Convert to picobarns (dividing by 2 due to the counting of quark/antiquark cases twice)
result_qq_scaled = result_qq * conversion_factor / 2
error_qq_scaled = error_qq * conversion_factor / 2

# Print results with combined errors
print(f"Cross-section (qq): {result_qq_scaled:.2f} pm {error_qq_scaled:.2f} pb")
print(f"Cross-section (gg): {result_gg_scaled:.2f} pm {error_gg_scaled:.2f} pb")
print(f"Total cross-section: {result_qq_scaled + result_gg_scaled:.2f} pm "
      f"{(error_qq_scaled**2 + error_gg_scaled**2) ** 0.5:.2f} pb")
```

C Powheg Configuration File

```
numevts 10000    ! number of events to be generated 10000
ih1   1          ! hadron 1
ih2   1          ! hadron 2
ebeam1 4080d0     ! energy of beam 1
ebeam2 4080d0     ! energy of beam 2
qmass  172.5     ! mass of heavy quark in GeV
facscfact 1    ! factorization scale factor: mufact=muref* facscfact
renscfact 1    ! renormalization scale factor: muren=muref* renscfact
fixedscale 1    ! use ref. scale=qmass (default 0, use running scale)
topdecaymode 22222
!semileptonic 1
hmass 125
hwidth 0.4171529E-02
bwcutoff 100
tdec/wmass 80.4  ! W mass for top decay
tdec/wwidth 2.141
tdec/bmass 5
tdec/twidth  1.31
tdec/elbranching 0.108
tdec/emass 0.00051
tdec/mumass 0.1057
tdec/taumass 1.777
tdec/dmass   0.100
tdec/umass   0.100
tdec/smass   0.200
tdec/cmass   1.5
tdec/sin2cabibbo 0.051
! To be set only if using LHA pdfs
lhans1 13200 ! 13100 for nlo     ! 91900pdf set for hadron 1 (LHA numbering)
lhans2 13200 ! 13100 for nlo     ! pdf set for hadron 2 (LHA numbering)
! To be set only if using internal mlm pdf
ndns1 131         ! pdf set for hadron 1 (mlm numbering)
ndns2 131         ! pdf set for hadron 2 (mlm numbering)
! To be set only if using different pdf sets for the two incoming hadrons
! QCDLambda5  0.25 ! for not equal pdf sets
! Parameters to allow or not the use of stored data
use-old-grid    1 ! if 1 use old grid if file pwggrids.dat is present (<> 1 regenerate)
use-old-ubound  1 ! if 1 use norm of upper bounding function stored in pwgubound.dat, if present; <> 1 regenerate
higgsfixedwidth 1 ! (default 0), If 1 uses standard, fixed width Breit-Wigner
                                    ! formula, if 0 it uses the running width Breit-Wigner
bornsuppfact 1 ! (default 1), If 1 the Born suppression factor is included.
                                  ! Weighted events are generated. If 0 no suppression
```

```
                                        ! factor is included, and
                                           events are unweighted. A
                                        ! generation cut bornktmin>0
                                           must be supplied in this
                                           case.
        ckkwscalup 1 ! (default 1), If 1 compute the scalup scale for
           subsequent
                                ! shower using the smallest kt in
                                   the final state;
                                ! If 0, use the standard POWHEG BOX
                                   scalup
        runningscales 0  ! (default 0), if 0 use hmass as central
                                        ! factorization and
                                           renormalization scale;
                                        ! if 1 use the Ht/2
        ncall1   3000      ! number of calls for initializing the
           integration grid 3000
        itmx1    5           ! number of iterations for initializing
           the integration grid
        ncall2   500000     ! number of calls for computing the
           integral and finding upper bound 500000
        itmx2    5         ! number of iterations for computing the
           integral and finding upper bound
        fastbtlbound 1    ! (default 0) if 1 use fast btilde bound
        foldcsi   1       ! number of folds on csi integration
        foldy     1       ! number of folds on  y  integration
        foldphi   1       ! number of folds on phi integration
        nubound 200000    ! number of bbarra calls to setup norm of
           upper bounding function

        ! OPTIONAL PARAMETERS
        testplots  1       ! (default 0, do not) do NLO and PWHG
           distributions
        bornonly   1     ! (default 0) if 1 do Born only
        LOevents   0     ! (default 0) if 1 do Born only
        bornktmin 20    ! Minimum transverse momentum of the Higgs at
           the underlying Born level
        storeinfo_rwgt 1   ! store info to allow for reweighting
        flg_debug      1   ! store extra event info for debugging
        minlo 0            ! default 0, set to 1 to use minlo
        sudscalevar 1      ! (default 1) scale variation also in
           Sudakov form factors in minlo
        turnNegPDFsPos 1
        manyseeds 1
        parallelstage 1
        xgriditeration 1
        novirtual 0
        btlscalereal 1
        btlscalect 1
        alphas_from_pdf 1
        maxseeds 5000
        rivetWeight 0
```

D MCFM Configuration File

```
mcfm_version = 10.1
[general]
    nproc = 157
    part = nlo
    runstring = 8TeV
    rundir = Ttbar8p16TeV
    sqrts = 8160
    ih1 = +82208
    ih2 = +1
    zerowidth = .false.
    removebr = .false.
    ewcorr = none
[nnlo]
[resummation]
    usegrid = .true.
    makegrid = .false.
    gridoutpath = PDFs/
    gridinpath = PDFs/
    res_range = 0.0 80.0
    resexp_range = 1.0 80.0
    fo_cutoff = 1.0
    transitionswitch = 0.4
[pdf]
    pdlabel = 'lhapdf'
[lhapdf]
        lhapdfset = CT18NLO     #for proton
        lhapdfset = EPPS21nlo_CT18Anlo_Pb208 #for lead
    lhapdfmember = 0
    dopdferrors = .false.
[scales]
    renscale = 172.5
    facscale = 172.5
    ;renscale = 1.0
    ;facscale = 1.0
    ;dynamicscale = sqrt(M^2+pt34^2)
    dynamicscale = none
    doscalevar = .false.
    maxscalevar = 6
[masses]
    hmass = 125
    mt = 172.5
    mb = 4.66
    mc = 1.275
[basicjets]
    inclusive = .true.
    algorithm = ankt
    Rcutjet = 0.5
[masscuts]
    m34min = 0
    m56min = 0
    m3456min = 0
[cuts]
    makecuts = .false.
    ptleptmin = 20
    etaleptmax = 2.4
    etaleptveto = 0.0 0.0
    ptminmiss = 30.0
```

```
    ptlept2min = 20.0
    etalept2max = 2.4
    etalept2veto = 0.0 0.0
    m34transmin = 0.0
    Rjlmin = 0.0
    Rllmin = 0.0
    delyjjmin = 0.0
    jetsopphem = .false.
    lbjscheme = 0
    ptbjetmin = 0.0
    etabjetmax = 99.0
[photon]
    fragmentation = .false.
    fragmentation_set = GdRG__LO
    fragmentation_scale = 1.0
    gammptmin = 40
    gammrapmax = 2.5
    gammpt2 = 25
    gammpt3 = 25
    Rgalmin = 0
    Rgagamin = 0.4
    Rgajetmin = 0
    cone_ang = 0.4
    epsilon_h = 0.5
    n_pow = 1
[histogram]
    writetop = .true.
    writetxt = .true.
[integration]
    initcallslord = 500
    initcallsnloreal=10000
    initcallsnlovirt=2000
    initcallsnnlobelow=2000
    initcallsnnlovirtabove=4000
    initcallsnnlorealabove=20000
    initcallsnloresummed=10
    initcallsnloresabove=2000
    usesobol = .true.
    seed = 0
    precisiongoal = 0.1
    readin = .false.
    writeintermediate = .true.
    warmupprecisiongoal = 0.25
    warmupchisqgoal = 2.5
[singletop]
    c_phiq = 0
    c_phiphi = 0.0 0.0
    c_tw = 0.0 0.0
    c_bw = 0.0 0.0
    c_tg = 0.0 0.0
    c_bg = 0.0 0.0
    lambda = 1000
    enable_lambda4 = .false.
    disable_sm = .false.
    mode_anomcoup = .false.
[anom_wz]
    enable = .false.
    delg1_z = 0
    delk_z = 0
```

```
    delk_g = 0
    lambda_z = 0
    lambda_g = 0
    h1Z = 0
    h1gam = 0
    h2Z = 0
    h2gam = 0
    h3Z = 0
    h3gam = 0
    h4Z = 0
    h4gam = 0
    tevscale = 2.0
[hjetmass]
    mtex = 0
[anom_higgs]
    hwidth_ratio = 1.0
    cttH = 1.0
    cWWH = 1.0
```

E Top++ Configuration File

```
/================================== General setup (Collider, pdf, F.O.
   vs RES.)
Collider LHC
ECMLHC 8160
WithResummation YES #NO for NLO
PDFuncertainty NO
PDFset CT18NNLO #CT18NLO for nlo
/================================== mt(GeV);
Mtop 172.5
MtopLimit -1.0
MtopStep 1
/================================== Resummation
OrderFO NNLO #NLO for nlo
OrderRES NNLL #NLL for nlo
/================================== Fixed Order
LO YES
NLO YES
NNLO YES #NO for nlo
/================================== Setup parameters
Precision 2
NPdfGrid 200
```

F Decay of τ Lepton

```
import numpy as np

tau_to_e = 0.1782
tau_to_mu = 0.1739
tau_to_hadr = 0.648

all_had  = 0.457
e_jets   = 0.146
mu_jets  = 0.146
tau_jets = 0.146
ee       = 0.012
mumu     = 0.012
tautau   = 0.012
emu      = 0.023
etau     = 0.023
mutau    = 0.023

new_all_had = round(100*(all_had + tau_jets*tau_to_hadr +
    tautau*np.power(tau_to_hadr,2)
                                        ) , 1)
new_e_jets  = round(100*(e_jets  + tau_jets*tau_to_e    +
    mutau*tau_to_hadr               + 2*tautau*tau_to_e *
    tau_to_hadr ) , 1)
new_mu_jets = round(100*(mu_jets + tau_jets*tau_to_mu   + etau
    *tau_to_hadr                + 2*tautau*tau_to_mu*
    tau_to_hadr ) , 1)
new_ee      = round(100*(ee      + etau*tau_to_e        +
    tautau*np.power(tau_to_e,2)
                                          ) , 1)
new_mumu    = round(100*(mumu    + emu *tau_to_mu       +
    tautau*np.power(tau_to_mu,2)
                                        ) , 1)
new_emu     = round(100*(emu     + mutau*tau_to_e       + etau
    *tau_to_mu                  + 2*tautau*tau_to_e*tau_to_mu
       ) , 1)

print("new_all_had = ", new_all_had, " %")
print("new_e_jets  = ", new_e_jets , " %")
print("new_mu_jets = ", new_mu_jets, " %")
print("new_ee      = ", new_ee     , " %")
print("new_mumu    = ", new_mumu   , " %")
print("new_emu     = ", new_emu    , " %")
print("sum         = ", new_all_had + new_e_jets + new_mu_jets
    + new_ee + new_mumu + new_emu)
```

G Cross sections for $t\bar{t}$ production in pp collisions at 8.16 TeV

PDF set	$\sigma_{t\bar{t}}^{\mathrm{pp}}$, $\sqrt{s}$=8.16 TeV [pb]
Run-2 style PDF4LHC (from param.)	266.23 ± 12.13 (PDF+α_S) $^{+6.44}_{-9.36}$ (Scale)
NNPDF3.1 notop (explicitly eval.)	$268.79\ ^{+7.89}_{-7.17}$ (PDF+α_S)
PDF4LHC21 (explicitly eval.)	$269.47^{+8.37}_{-8.37}$ (PDF+α_S)$^{+7.00}_{-9.43}$ (Scale)
(Run-3 style recommendation)	
ATLASpdf21 T-3 (explicitly eval.)	$286.95^{+15.21}_{-14.24}$ (Total PDF) $^{+13.30}_{-12.60}$ (Exp PDF) $^{+7.37}_{-6.57}$ (Model PDF)$^{+0.00}_{-0.92}$ (Para PDF)
MSHT20 (explicitly eval.)	$269.30^{+8.64}_{-6.74}$ (PDF+α_S)
NNPDF 4.0 (explicitly eval.)	$258.83^{+4.96}_{-5.85}$ (PDF+α_S)

Table 5.1: Total NNLO+NNLL cross section at $\sqrt{s}$=8.16 TeV, for the PDF4LHC recommendation, including PDF+α_S uncertainties. The top-quark mass used is 172.5 GeV.

H V+jets Background Samples

Process	Conf.	N. Evt	ME Gen.	PS Model	ME PDF
$Z \to e^+e^-$	pPb/pp	198038	Sherpa-2.2.10	Sherpa-2.2.10	NNPDF3.0 NLO
$Z \to e^+e^-$	pPb/np	297025			
$Z \to e^+e^-$	Pbp/pp	298468			
$Z \to e^+e^-$	Pbp/pn	397934			
$W \to e\nu_e$	pPb/pp	297025			
$W \to e\nu_e$	pPb/np	415745			
$W \to e\nu_e$	Pbp/pp	517271			
$W \to e\nu_e$	Pbp/pn	795779			
$Z \to \mu^+\mu^-$	pPb/pp	5150440	Sherpa-2.2.10	Sherpa-2.2.10	NNPDF3.0 NLO
$Z \to \mu^+\mu^-$	pPb/np	7904035			
$Z \to \mu^+\mu^-$	Pbp/pp	994741			
$Z \to \mu^+\mu^-$	Pbp/pn	1591518			
$W \to \mu\nu_\mu$	pPb/pp	1980847			
$W \to \mu\nu_\mu$	pPb/np	3070519			
$W \to \mu\nu_\mu$	Pbp/pp	4078279			
$W \to \mu\nu_\mu$	Pbp/pn	6366214			
$Z \to \tau^+\tau^-$	pPb/pp	1564764	Sherpa-2.2.10	Sherpa-2.2.10	NNPDF3.0 NLO
$Z \to \tau^+\tau^-$	pPb/np	2475144			
$Z \to \tau^+\tau^-$	Pbp/pp	3183008			
$Z \to \tau^+\tau^-$	Pbp/pn	4754814			
$W \to \tau\nu_\tau$	pPb/pp	6517120			
$W \to \tau\nu_\tau$	pPb/np	9865100			
$W \to \tau\nu_\tau$	Pbp/pp	5968388			
$W \to \tau\nu_\tau$	Pbp/pn	8951559			

Table 5.2: Z/W+jet samples in the electron, muon, and tau channels associated to the light-flavour jet candidate.

Process	Conf.	N. Evt	ME Gen.	PS Model	ME PDF
$Z \to e^+e^-$	pPb/pp	198038	Sherpa-2.2.10	Sherpa-2.2.10	NNPDF3.0 NLO
$Z \to e^+e^-$	pPb/np	297025			
$Z \to e^+e^-$	Pbp/pp	298468			
$Z \to e^+e^-$	Pbp/pn	397934			
$W \to e\nu_e$	Pbp/pn	297025			
$W \to e\nu_e$	pPb/np	415745			
$W \to e\nu_e$	Pbp/pp	517271			
$W \to e\nu_e$	Pbp/pn	795779			
$Z \to \mu^+\mu^-$	pPb/pp	653499	Sherpa-2.2.10	Sherpa-2.2.10	NNPDF3.0 NLO
$Z \to \mu^+\mu^-$	pPb/np	970460			
$Z \to \mu^+\mu^-$	Pbp/pp	1293155			
$Z \to \mu^+\mu^-$	Pbp/pn	1969505			
$W \to \mu\nu$	pPb/pp	3149682			
$W \to \mu\nu_\mu$	pPb/np	4836464			
$W \to \mu\nu_\mu$	Pbp/pp	6046948			
$W \to \mu\nu_\mu$	Pbp/pn	8922724			
$Z \to \tau^+\tau^-$	pPb/pp	188101	Sherpa-2.2.10	Sherpa-2.2.10	NNPDF3.0 NLO
$Z \to \tau^+\tau^-$	pPb/np	306866			
$Z \to \tau^+\tau^-$	Pbp/pp	396936			
$Z \to \tau^+\tau^-$	Pbp/pn	591881			
$W \to \tau\nu_\tau$	pPb/pp	970460			
$W \to \tau\nu_\tau$	pPb/np	1366714			
$W \to \tau\nu_\tau$	Pbp/pp	1770501			
$W \to \tau\nu_\tau$	Pbp/pn	2775268			

Table 5.3: Z/W+jet samples in the electron, muon, and tau channels associated to the c-jet candidate.

Process	Conf.	N. Evt	ME Gen.	PS Model	ME PDF
$Z \to e^+e^-$	pPb/pp	128683	Sherpa-2.2.10	Sherpa-2.2.10	NNPDF3.0 NLO
$Z \to e^+e^-$	pPb/np	207888			
$Z \to e^+e^-$	Pbp/pp	258669			
$Z \to e^+e^-$	Pbp/pn	396938			
$W \to e\nu_e$	pPb/pp	267230			
$W \to e\nu_e$	pPb/np	415745			
$W \to e\nu_e$	Pbp/pp	507327			
$W \to e\nu_e$	Pbp/pn	795281			
$Z \to \mu^+\mu^-$	pPb/pp	128683	Sherpa-2.2.10	Sherpa-2.2.10	NNPDF3.0 NLO
$Z \to \mu^+\mu^-$	pPb/np	207888			
$Z \to \mu^+\mu^-$	Pbp/pp	258669			
$Z \to \mu^+\mu^-$	Pbp/pn	387986			
$W \to \mu\nu_\mu$	pPb/pp	247554			
$W \to \mu\nu_\mu$	pPb/np	396043			
$W \to \mu\nu_\mu$	Pbp/pp	487323			
$W \to \mu\nu_\mu$	Pbp/pn	785832			
$Z \to \tau^+\tau^-$	pPb/pp	39612	Sherpa-2.2.10	Sherpa-2.2.10	NNPDF3.0 NLO
$Z \to \tau^+\tau^-$	pPb/np	59382			
$Z \to \tau^+\tau^-$	Pbp/pp	69631			
$Z \to \tau^+\tau^-$	Pbp/pn	109433			
$W \to \tau\nu_\tau$	pPb/pp	69232			
$W \to \tau\nu_\tau$	pPb/np	118748			
$W \to \tau\nu_\tau$	Pbp/pp	149241			
$W \to \tau\nu_\tau$	Pbp/pn	228820			

Table 5.4: Z/W+jet samples in the electron, muon, and tau channels associated to the b-jet candidate.

I Matching Systematics

This section provides detailed plots of the effect of matching systematics on $H_{\mathrm{T}}^{\ell,j}$ distribution shown in Figures 5.1-5.4.

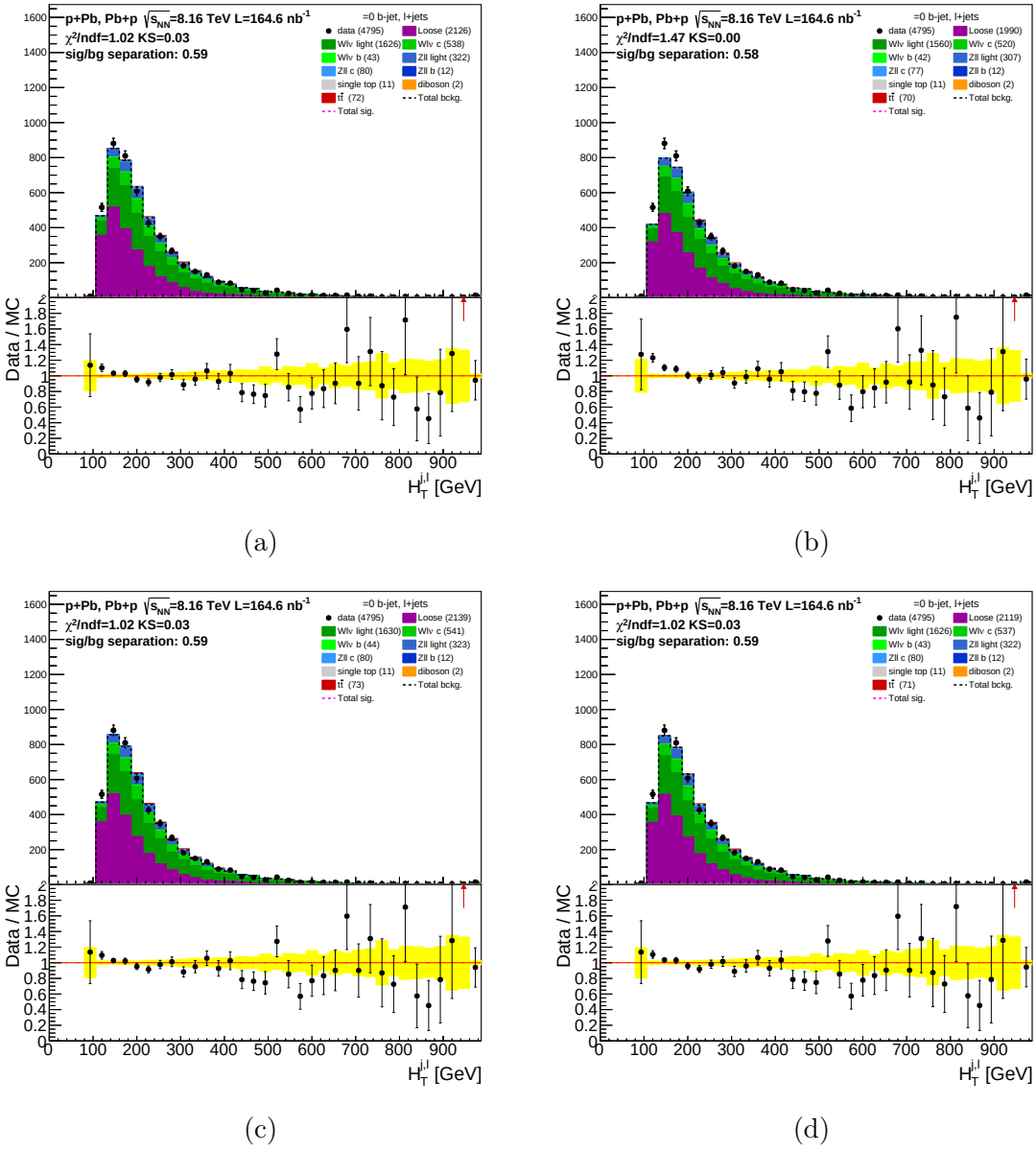
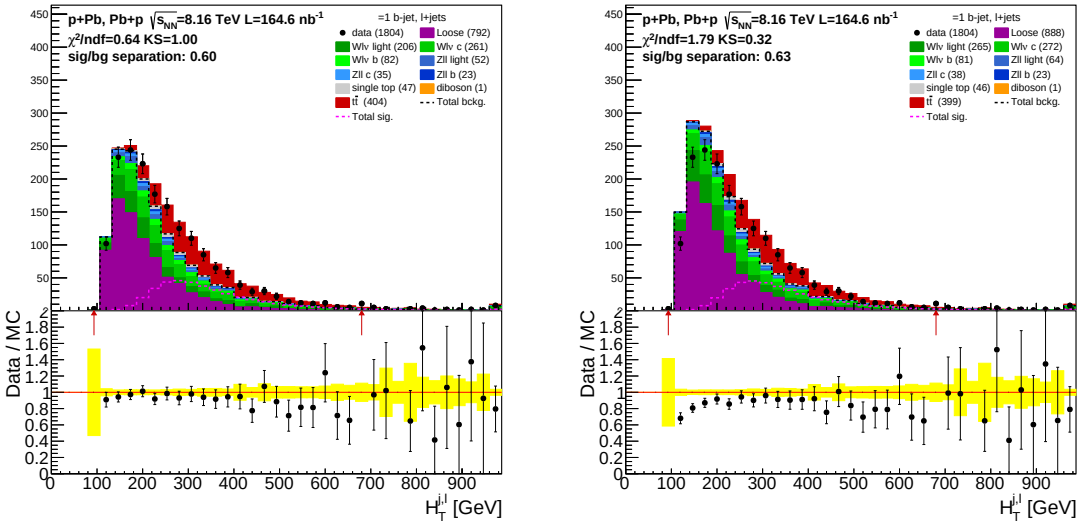


Figure 5.1: Control plot of the $H_{\mathrm{T}}^{\ell,j}$ variable in the $=0$b region (a) – default $\Delta R < 0.3$ for the b-tagged matching, (b) systematics of assigning extra HI jets using rejection function, (c) systematics down $\Delta R < 0.2$, (d) systematics up $\Delta R < 0.4$.



(a) (b)

(c) (d)

Figure 5.2: Control plot of the $H_{\mathrm{T}}^{\ell,j}$variable in the = 1b region (a) – default $\Delta R < 0.3$ for the b-tagged matching, (b) systematics of assigning extra HI jets using rejection function, (c) systematics down $\Delta R < 0.2$, (d) systematics up $\Delta R < 0.4$.

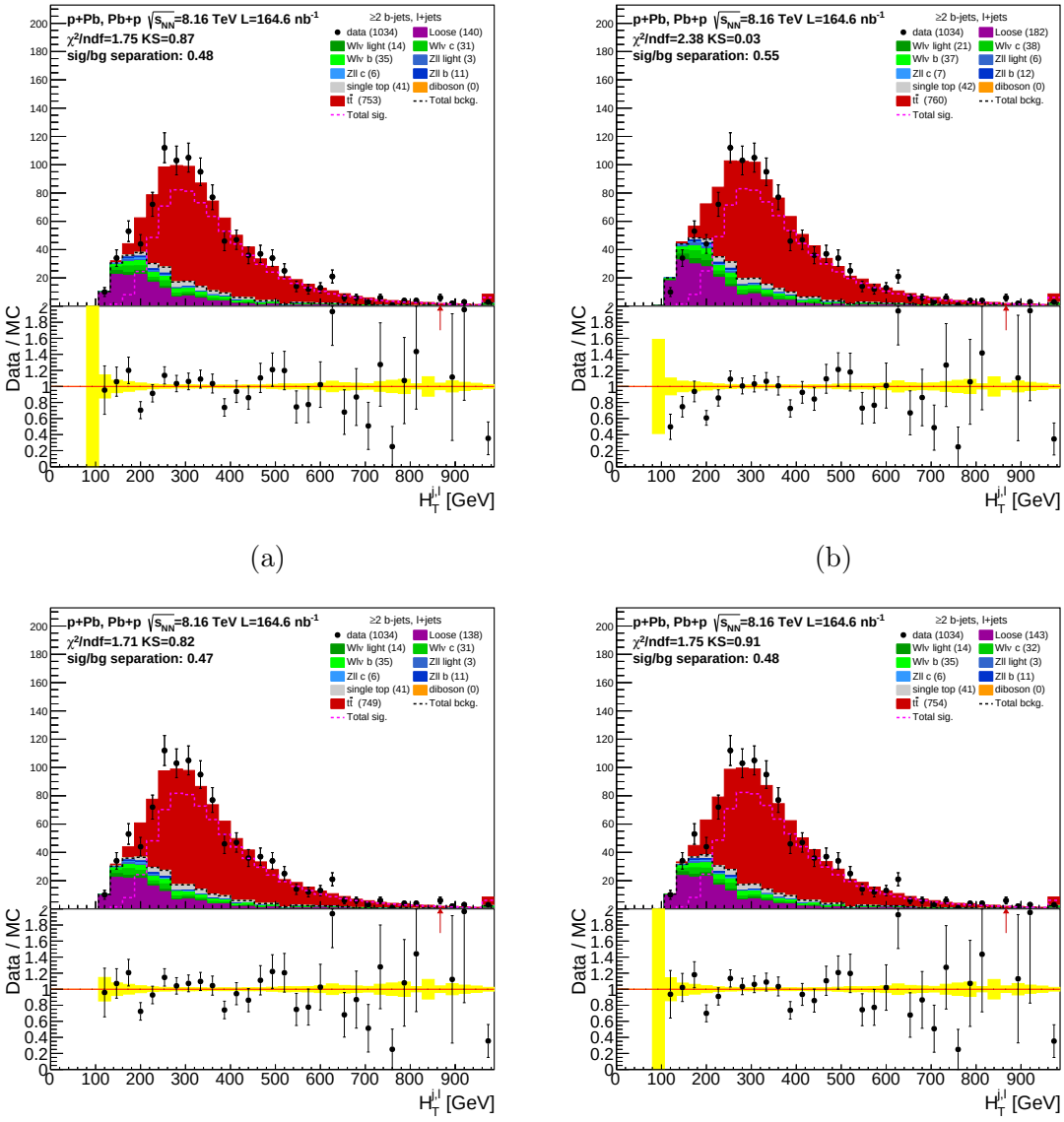


Figure 5.3: Control plot of the $H_{\mathrm{T}}^{\ell,j}$variable in the >= 2b region (a) – default $\Delta R < 0.3$ for the b-tagged matching, (b) systematics of assigning extra HI jets using rejection function, (c) systematics down $\Delta R < 0.2$, (d) systematics up $\Delta R < 0.4$.

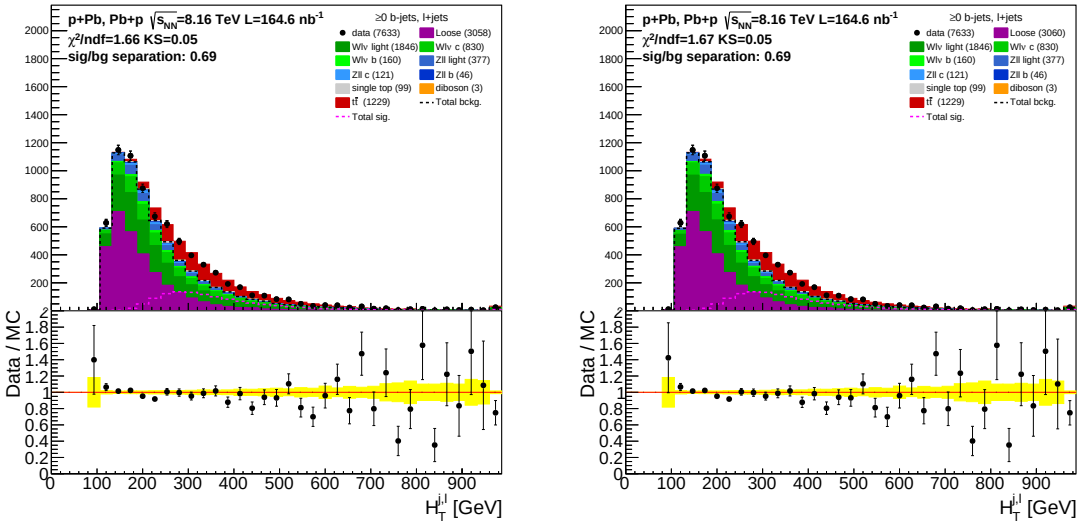
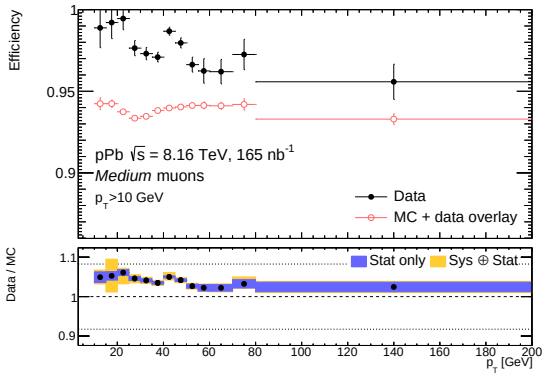


Figure 5.4: Control plot of the $H_{\mathrm{T}}^{\ell,j}$variable in the >= 0b region (a) – default $\Delta R < 0.3$ for the b-tagged matching, (b) systematics of assigning extra HI jets using rejection function, (c) systematics down $\Delta R < 0.2$, (d) systematics down $\Delta R < 0.4$.

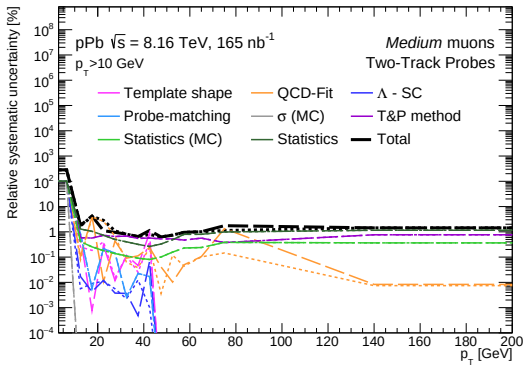
J Muon performance

In this section finer binning and 2D maps of individual scale factors are shown.

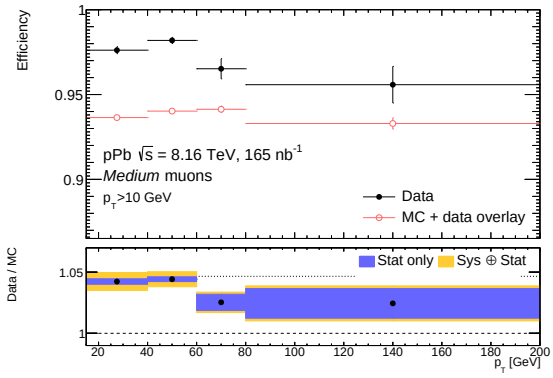
J.1 Muon Reconstruction/Identification Efficiency



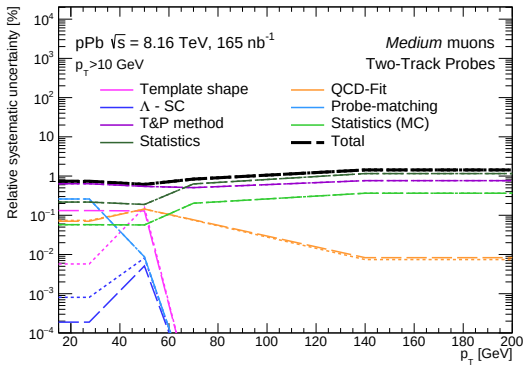
(a)



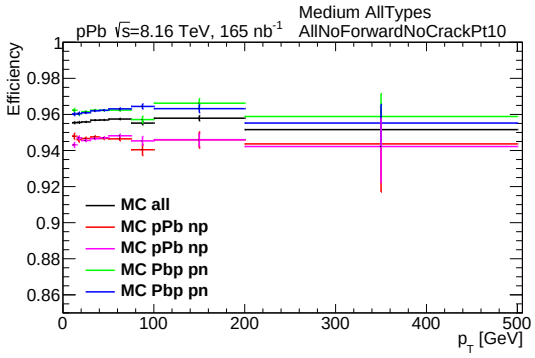
(b)



(c)

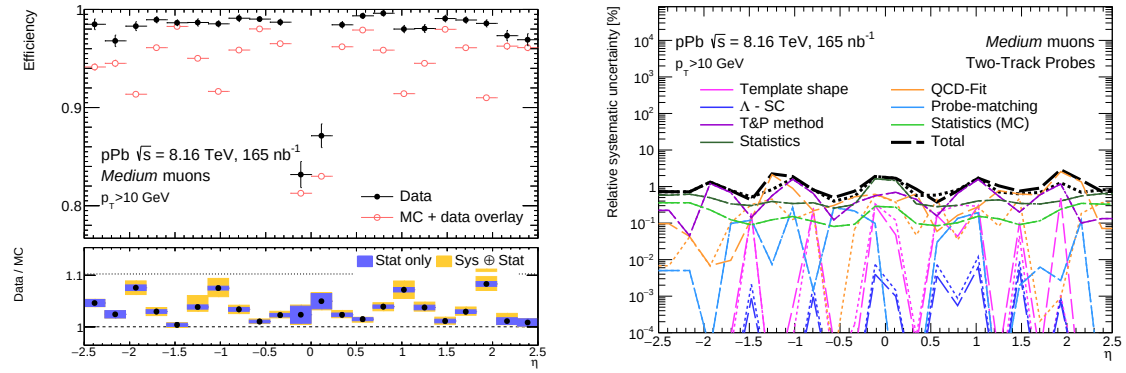


(d)



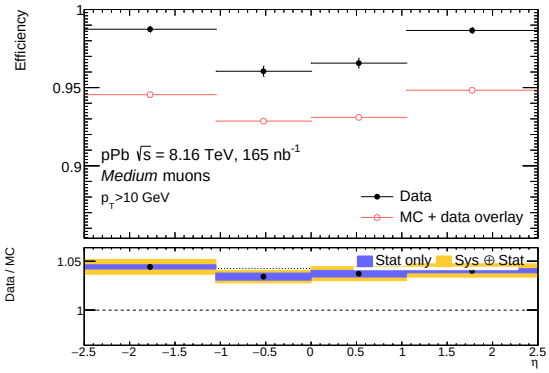
(e) Cross-check of MC efficiencies using independed code.

Figure 5.5: The muon reconstruction/identification ϵ(Medium) efficiency as a function of p_T (in the right pad are shown individual systematic uncertainties). In Figure 5.5e is the cross-check of MC efficiencies using independed code.

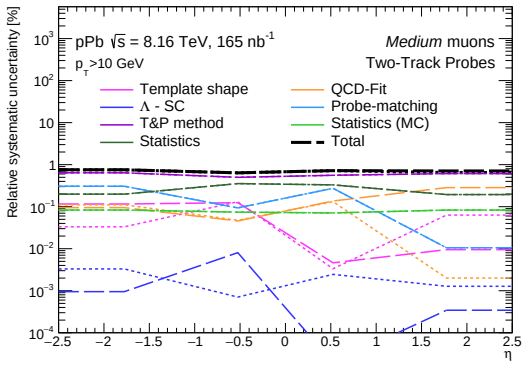


(a)

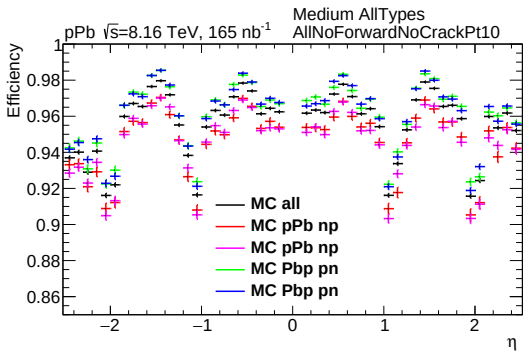
(b)



(c)

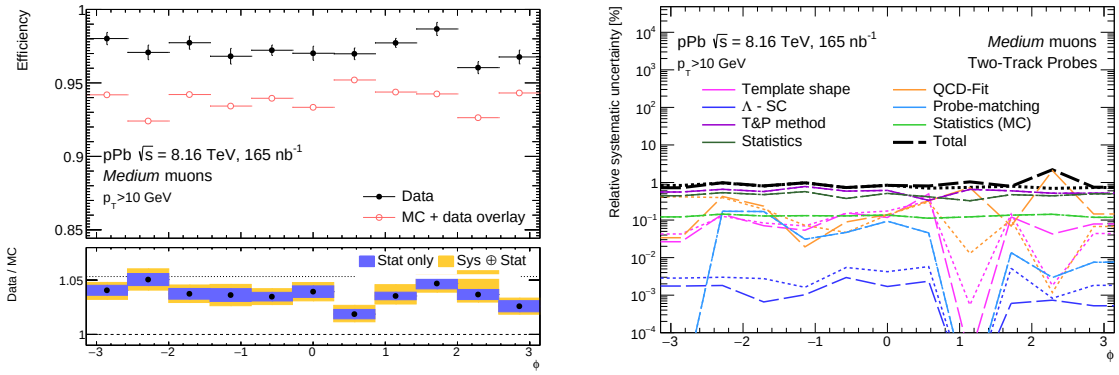


(d)



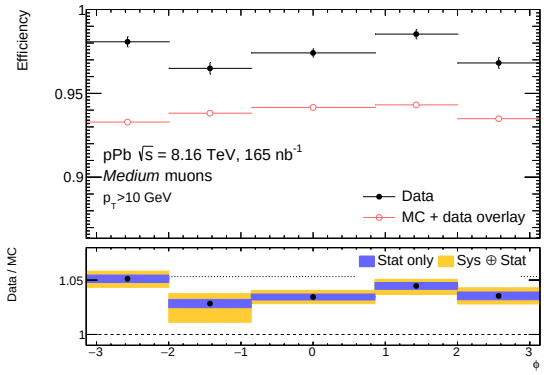
(e) Cross-check of MC efficiencies using independed code.

Figure 5.6: The muon reconstruction/identification ϵ(Medium) efficiency as a function of η (in the right pad are shown individual systematic uncertainties). In Figure 5.6e is the cross-check of MC efficiencies using independed code.

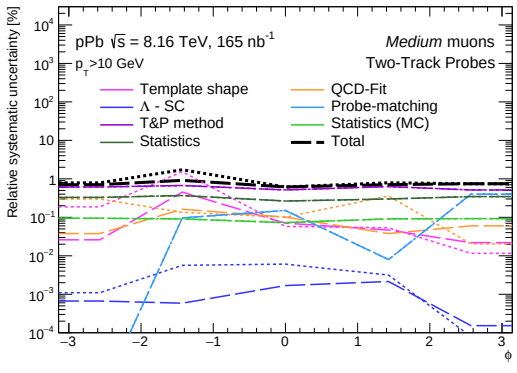


(a)

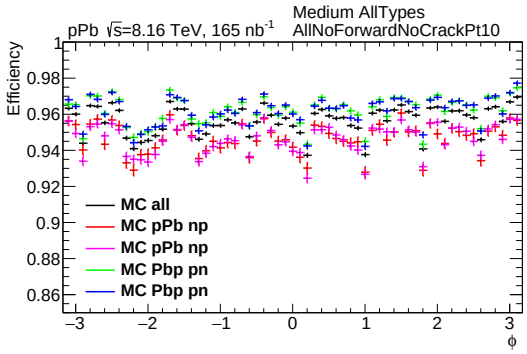
(b)



(c)



(d)



(e) Cross-check of MC efficiencies using independed code.

Figure 5.7: The muon reconstruction/identification ϵ(Medium) efficiency as a function of ϕ (in the right pad are shown individual systematic uncertainties). In Figure 5.7e is the cross-check of MC efficiencies using independed code.

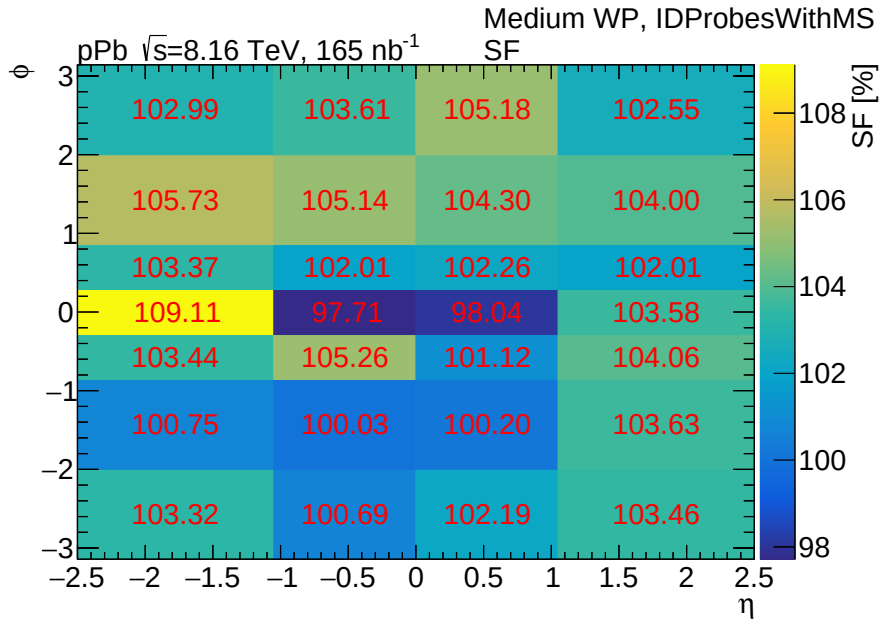


Figure 5.8: The 2D $\eta - \phi$ reconstruction scale factor SF(Medium) in the AllPt10 detector region.

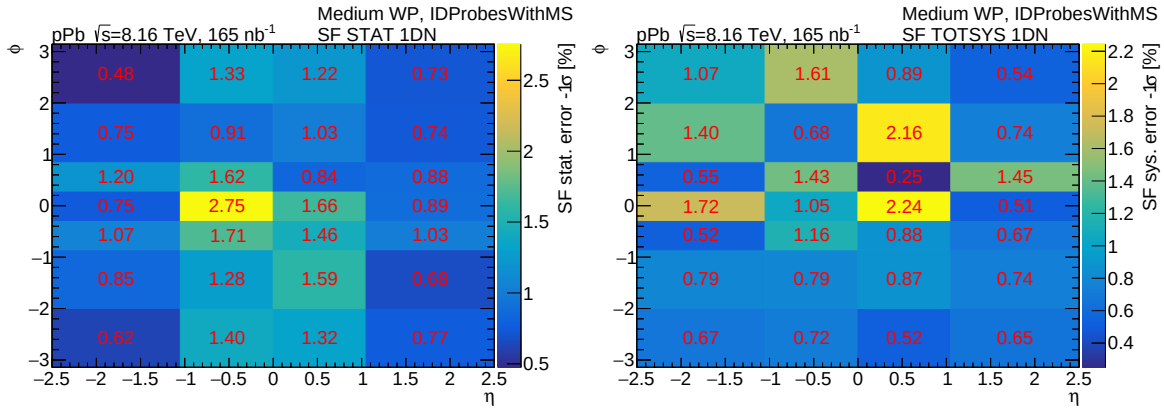


Figure 5.9: The 2D $\eta - \phi$ reconstruction scale factor SF(Medium) statistical (left) and systematic (right) uncertainties (down) in the AllPt10 detector region.

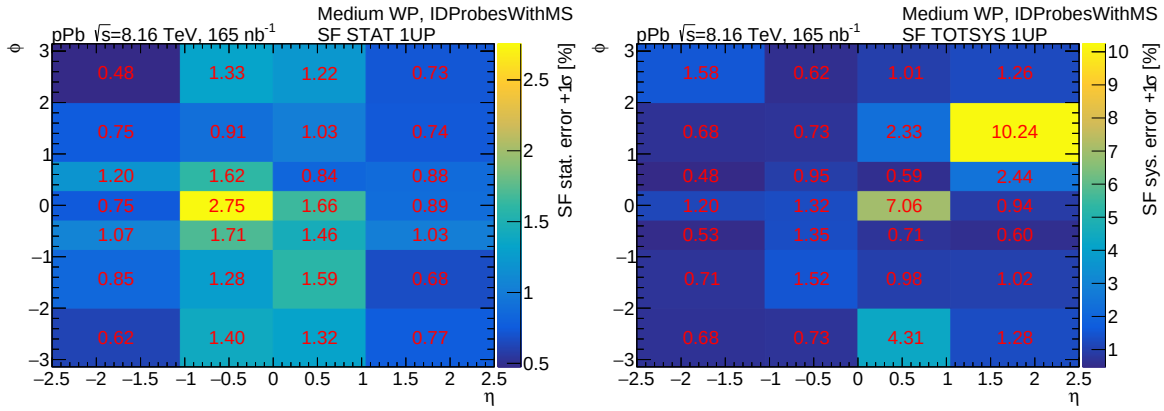


Figure 5.10: The 2D $\eta - \phi$ reconstruction scale factor SF(Medium) statistical (left) and systematic (right) uncertainties (up) in the AllPt10 detector region.

J.2 Muon Isolation Efficiency

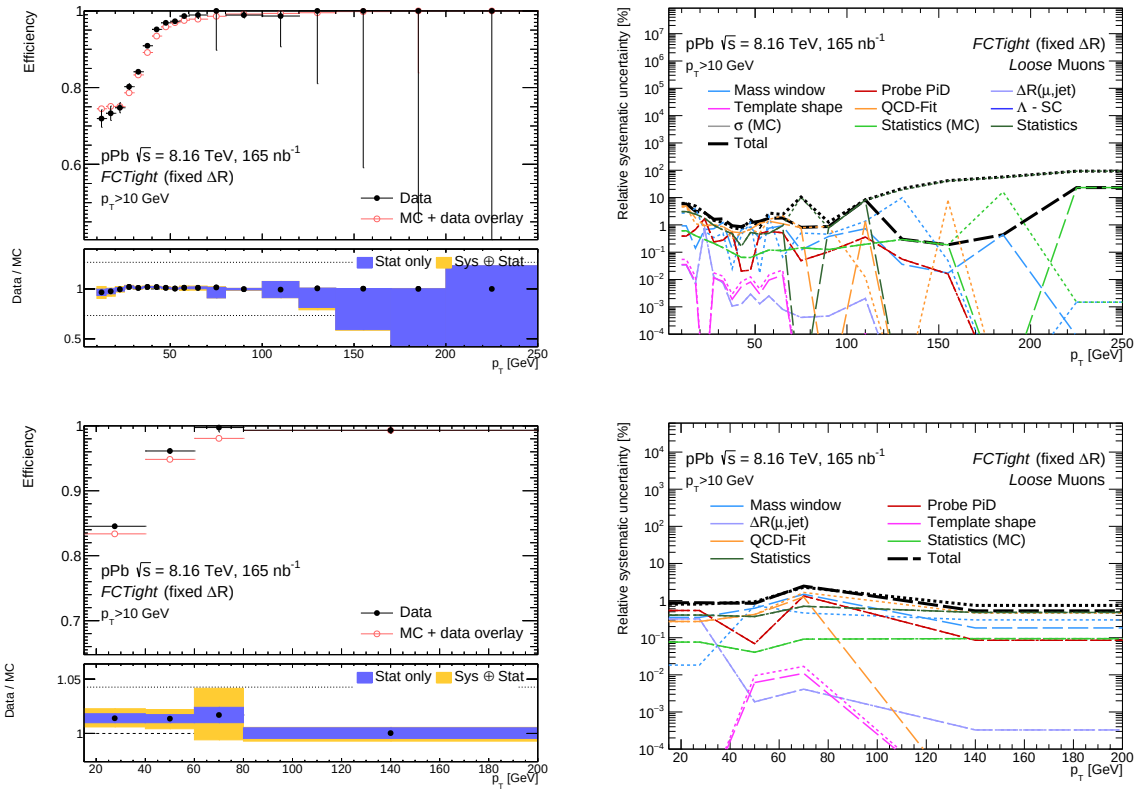


Figure 5.11: The muon isolation efficiency *FCTight_FixedRad*, Loose muon ID working point as a function of p_T (in the right pad are shown individual systematic uncertainties).

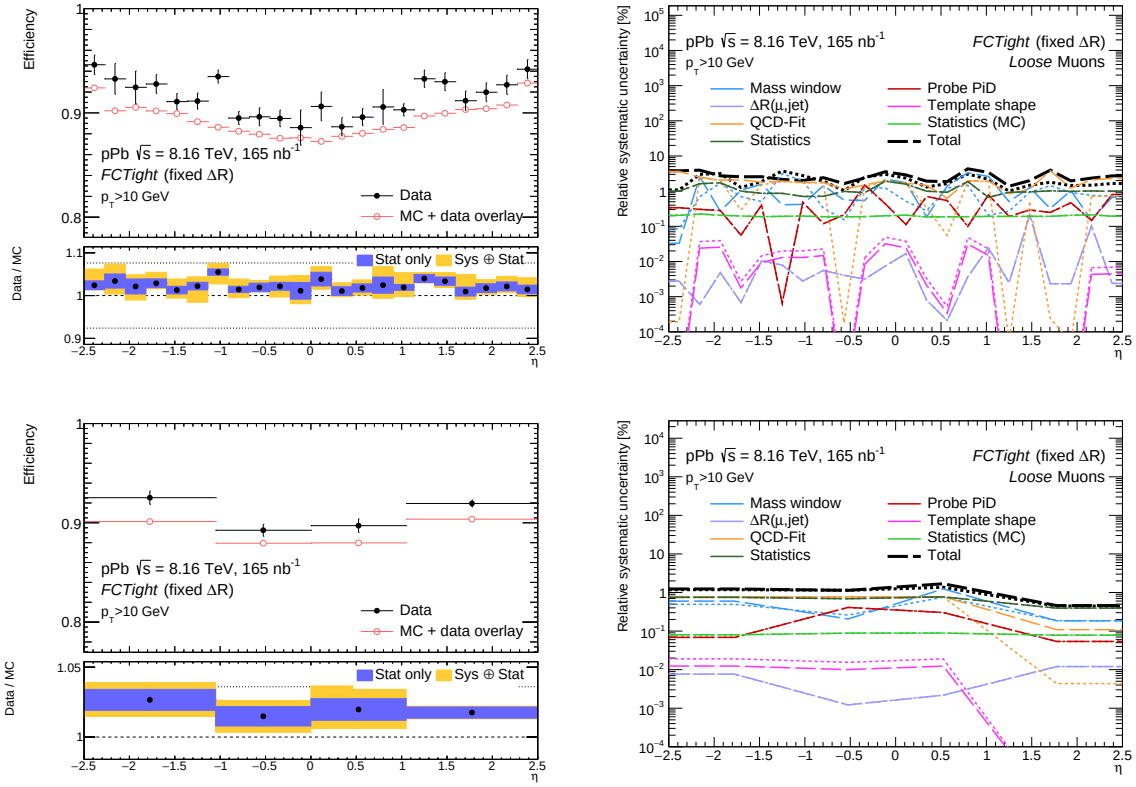


Figure 5.12: The muon isolation efficiency *FCTight_FixedRad*, Loose muon ID working point as a function of η (in the right pad are shown individual systematic uncertainties).

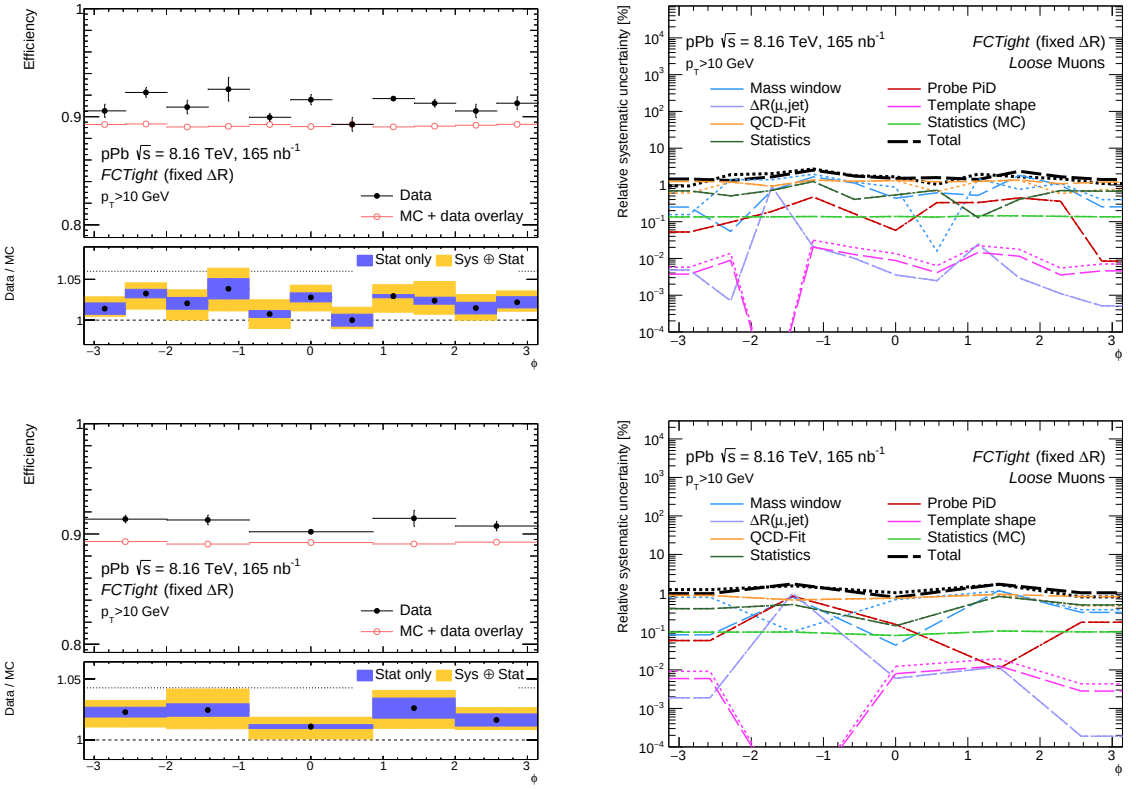


Figure 5.13: The muon isolation efficiency *FCTight_FixedRad*, Loose muon ID working point as a function of ϕ (in the right pad are shown individual systematic uncertainties).

J.3 Muon Trigger Efficiency

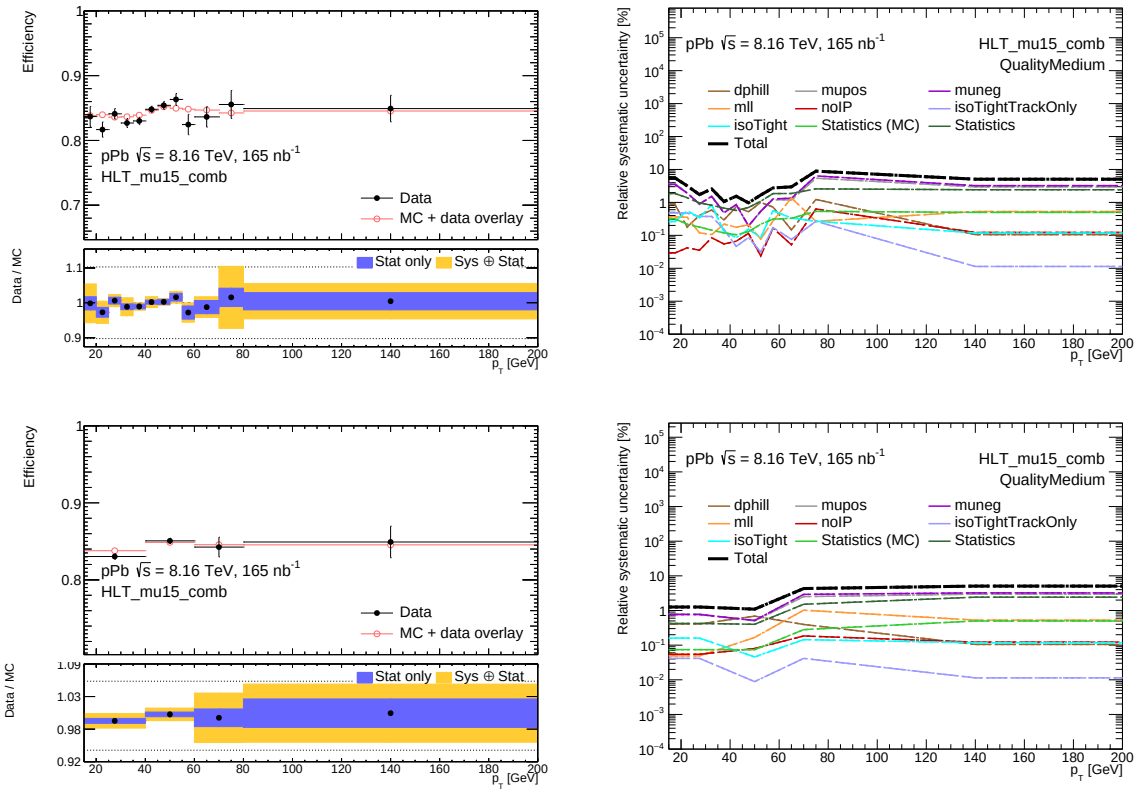


Figure 5.14: The muon trigger efficiency for the HLT_mu15_comb trigger for the medium muon ID working point as a function of p_T (in the right pad are shown individual systematic uncertainties).

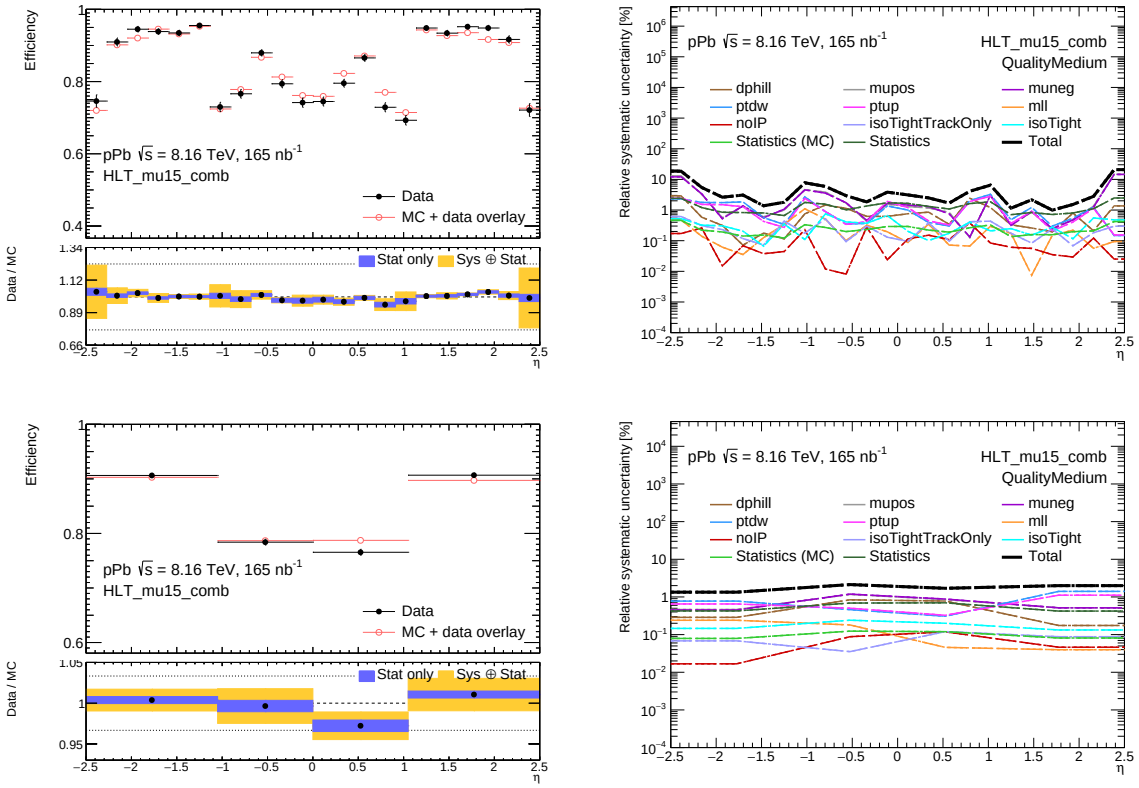


Figure 5.15: The muon trigger efficiency for the HLT_mu15_comb trigger for the medium muon ID working point as a function of η (in the right pad are shown individual systematic uncertainties).

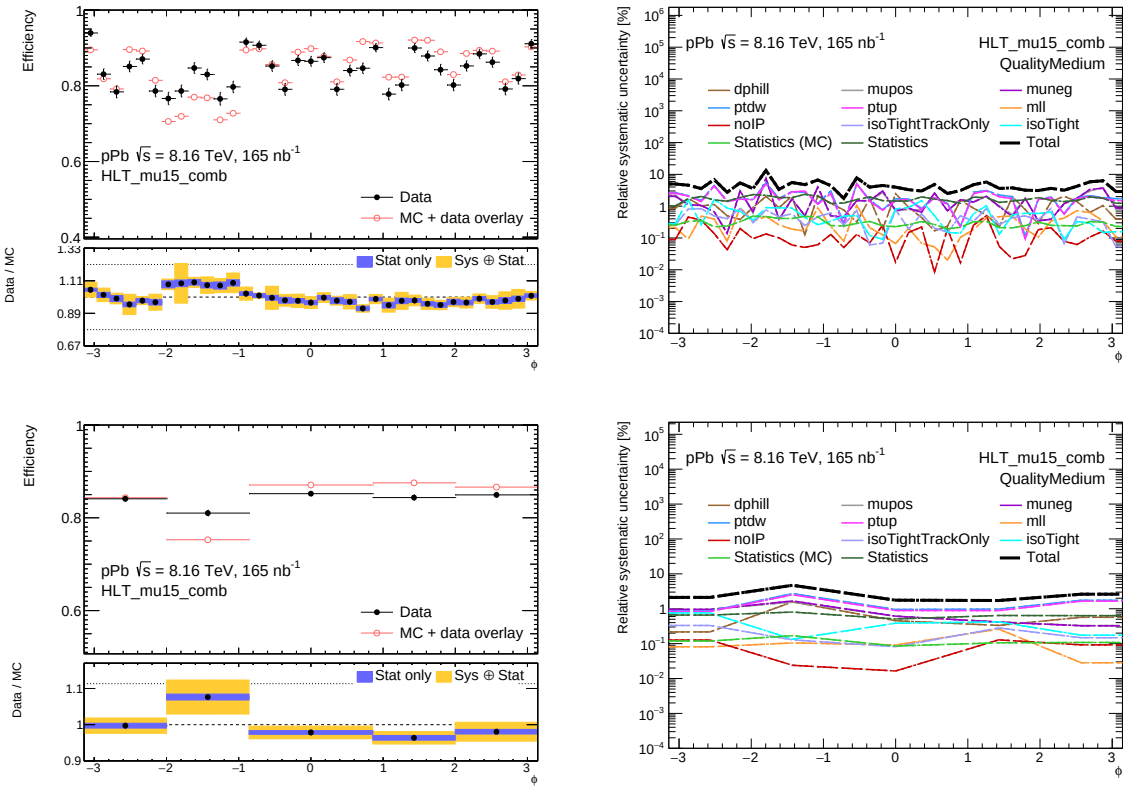


Figure 5.16: The muon trigger efficiency for the HLT_mu15_comb trigger for the medium muon ID working point as a function of ϕ (in the right pad are shown individual systematic uncertainties).

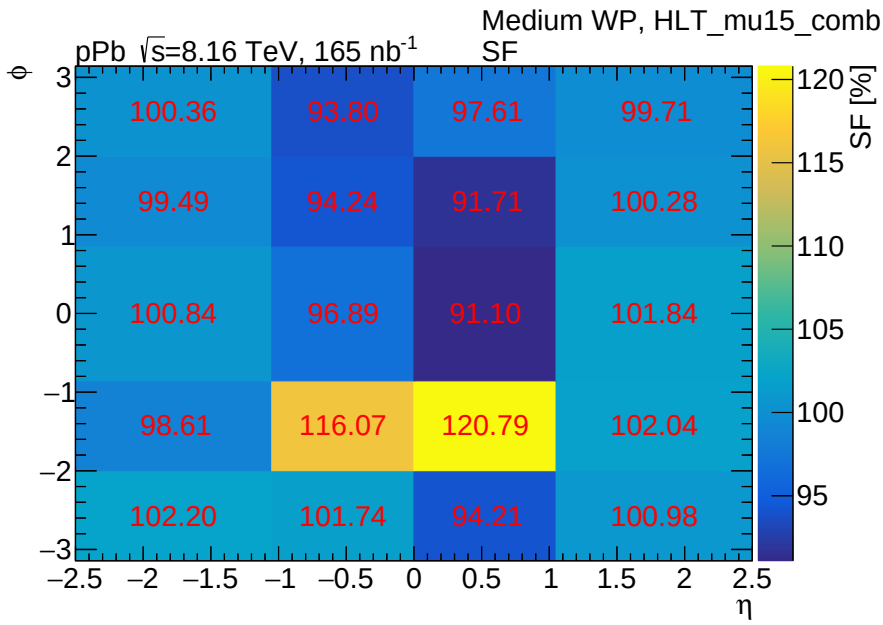


Figure 5.17: The 2D $\eta - \phi$ muon trigger scale factor HLT_mu15_comb in medium working point.

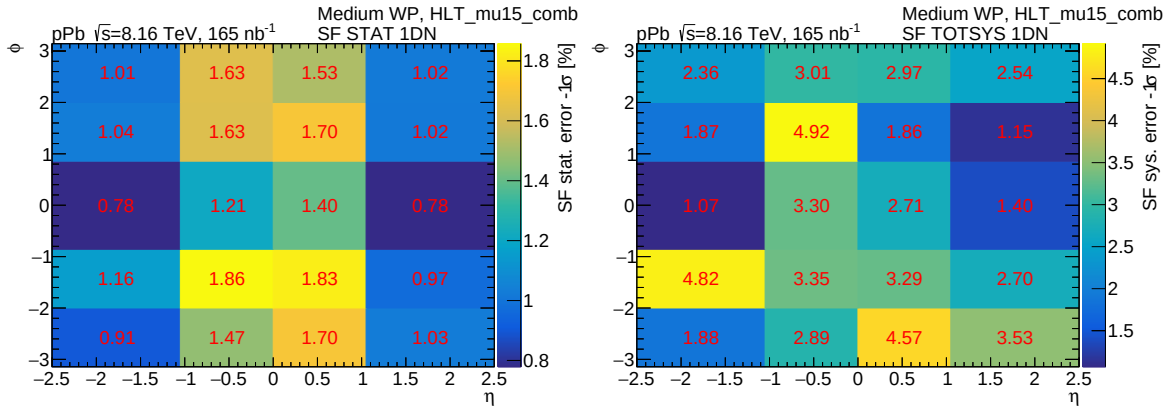


Figure 5.18: The 2D $\eta - \phi$ muon trigger scale factor staticical (left) and systematic (right) uncertainties (up and down are symetrical) HLT_mu15_comb in medium working point.

K Electron performance

K.1 Electron Identification Efficiency

The application of electron identification algorithms allows to determine whether the electron candidates originate from signal or background. Two of the main identification algorithms are cut-based and likelihood-based. The latter has many advantages over the former, and is currently used as a baseline identification algorithm by the ATLAS Collaboration [73].

The signal and background likelihoods L_S and L_B are the products of n probability density functions (PDFs)

$$L_{S(B)}(\vec{x}) = \prod_{i=1}^{n} P_{S(B),i}(x_i), \tag{5.1}$$

where $\vec{x}$ is the vector of electron discriminating quantities and $P_{S(B),i}(x_i)$ is the signal (background) PDF value for the i^{th} quantity evaluated at x_i. A discriminant d_L is determined as

$$d_L = \frac{L_S}{L_S + L_B}. \tag{5.2}$$

The discriminant d_L for signal and background has a sharp peak at unity or zero for signal and background, respectively. Therefore, it is convenient to define working points using a log-transformed discriminant d'_L

$$d'_L = -\tau \ln(d_L^{-1} - 1), \tag{5.3}$$

where τ parameter is fixed to 15 [107]. Compared to d_L, which varies between zero and unity, the range of d'_L values is extended to all real numbers.

Based on d'_L values for signal and background, four identification working points are defined, referred to as LooseLLH, LooseAndBLayerLLH, MediumLLH and TightLLH. For consecutive selections, increasing thresholds are defined for the likelihood discriminant. Looser working points have higher efficiencies and lower background rejection, while tighter ones provide better background reduction at lower electron efficiency. In order to determine numerical values of the d'_L discriminant, MC simulation is used.

Electron identification efficiencies as a function of E_{T} and η for four working points are shown in Figure 5.19 for data and MC simulation. The efficiencies increase with E_{T} from 68% (82)% for TightLLH (MediumLLH) at $E_{\text{T}} = 15$ GeV and reach the plateau for E_{T} at around 60 GeV with 87% (92%) for TightLLH (MediumLLH).

The identification efficiencies in Figure 5.19 in MC simulation is 1–5% higher than in the data which results in scale factors which are below unity. For central pseudorapidities $|\eta| < 0.5$ scale factors are consistent with unity. They are significantly below unity for $|\eta| > 1$.

K.2 Electron Isolation Efficiency

To further distinguish signal electrons from background the isolation criteria are introduced. Isolation variables are designed to quantify the level of surrounding activity near the electron candidate. Two key variables are utilized for this purpose:

- Calorimeter-based isolation, $E^{\text{iso}}_{\text{T,cone}}$: This is defined as the sum of the transverse energy of EM clusters whose centers lie within an isolation cone of radius $\Delta R = 0.2$ around the electron candidate. The transverse energy of the electron candidate's core cluster, with a size of 0.125×0.175 in $\eta \times \phi$ space, known as core energy ($E_{\text{T,core}}$), is subtracted. Additional corrections are applied to account for core energy leakage and pileup effects.

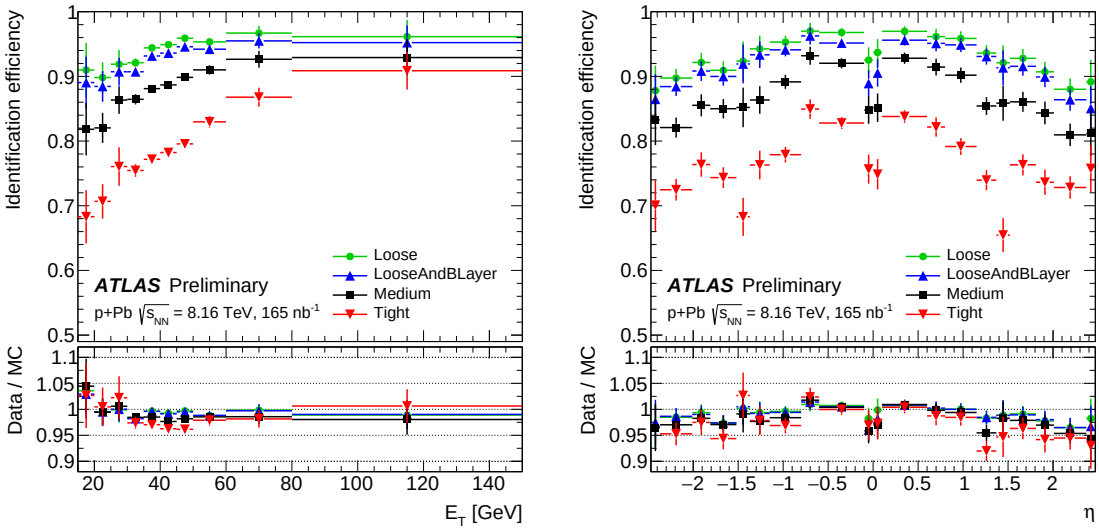


Figure 5.19: Electron identification efficiency as a function of electron E_{T} (left) and η (right) for four working points evaluated in the data [79].

- Track-based isolation, $p_{\mathrm{T,var}}^{\mathrm{iso}}$: This is the sum of the transverse momenta of tracks within an isolation cone of radius $\Delta R = \min(0.2, 10\ \mathrm{GeV}/p_{\mathrm{T}})$, centered on the electron candidate's track. The cone originates from the primary vertex of the proton-lead (p+Pb) interaction. Tracks associated with the electron, such as those from e^+e^- conversions or bremsstrahlung photons, are excluded from the sum.

Using these variables, various isolation working points are defined. Isolation requirements fall into two categories: 1) Efficiency-targeted and 2) Fixed-requirements. Efficiency-targeted working points adjust isolation criteria to achieve a desired efficiency, which can be constant or vary with the electron's p_{T}, here denoted as Gradient working point.

Fixed-requirement working points, on the other hand, use fixed thresholds for the isolation variables. Efficiency-targeted isolation is typically applied in high-E_{T} analyses where high signal efficiency is prioritized, whereas fixed-requirement isolation is more suited to low-E_{T} analyses, where stronger background rejection is necessary. Three working points are denoted as FixedCutLoose, FixedCutTight and FixedCutHighPtCaloOnly.

Electron isolation efficiencies have been measured as a function of E_{T} and η for four working points. They are shown in Figure 5.20 for data and MC simulation. The efficiencies range between 65–96% at $E_{\mathrm{T}} = 15$ GeV for various working points. They also reach a plateau region at various E_{T} values. Scale factors do not show significant deviations from unity.

K.3 Electron Trigger Efficiency

The analysis utilizes the e15_lhloose and e15_lhloose_nod0 trigger efficiencies of which were measured using tag-and-probe method. With these triggers, an electron candidate passing likelihood loose identification with $E_{\mathrm{T}} > 15$ GeV is selected. Term _nod0 means that the impact parameter d_0 is not used in the identification process [108]. Trigger efficiency is defined as a ratio of the number of probes passing trigger requirements to the number of all probe electrons [109]. Probes for trigger efficiency are required to pass identification and isolation selections. Therefore, the trigger efficiency ϵ_{trig} is defined as

$$\epsilon_{\mathrm{trig}} = \frac{N^{\mathrm{id+iso+trig}} - B^{\mathrm{id+iso+trig}}}{N^{\mathrm{id+iso}} - B^{\mathrm{id+iso}}}, \tag{5.4}$$

where $N^{\mathrm{id+iso+trig}}(B^{\mathrm{id+iso+trig}})$ is the number of Z boson electron candidates (background events) passing requirements of identification, isolation and trigger, and $N^{\mathrm{id+iso}}(B^{\mathrm{id+iso}})$ is a

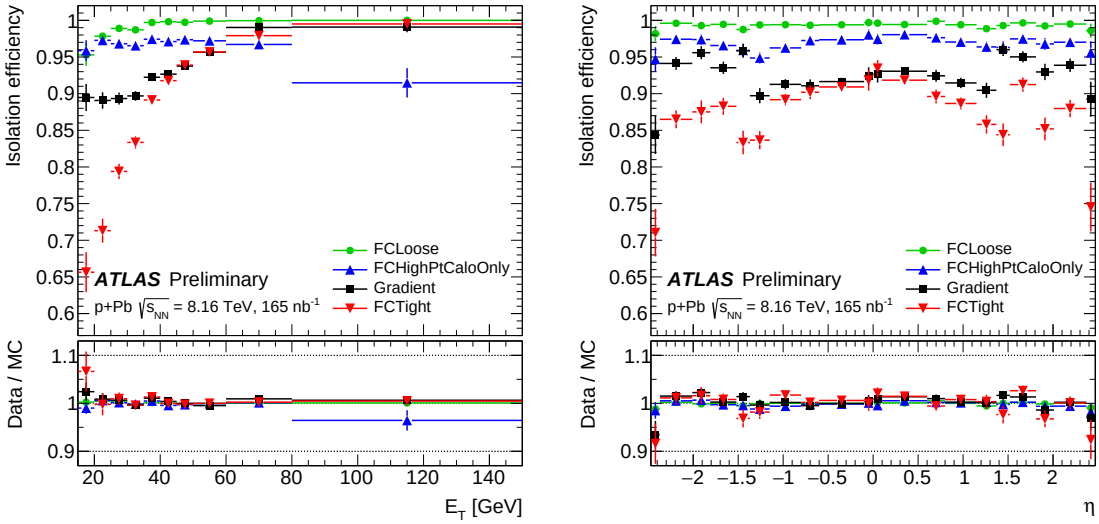


Figure 5.20: Electron isolation efficiency as a function of electron E_{T} (left) and η (right) for five working points evaluated in the data (full markers) [79].

number of Z boson electron candidates (background events) passing requirements of identification and isolation only.

Electron trigger efficiencies as a function of E_{T} and η for two electron triggers are shown in Figure 5.21 for data and MC simulation. The efficiencies raise with E_{T} from 82% at $E_{\mathrm{T}} = 15$ GeV and reach the plateau for E_{T} at around 45 GeV with 98%. Scale factors are consistent with unity except for $p_{\mathrm{T}} < 20$ GeV, the EM calorimeter transition region, $1.37 < |\eta| < 1.52$, and for $|\eta| < 0.1$.

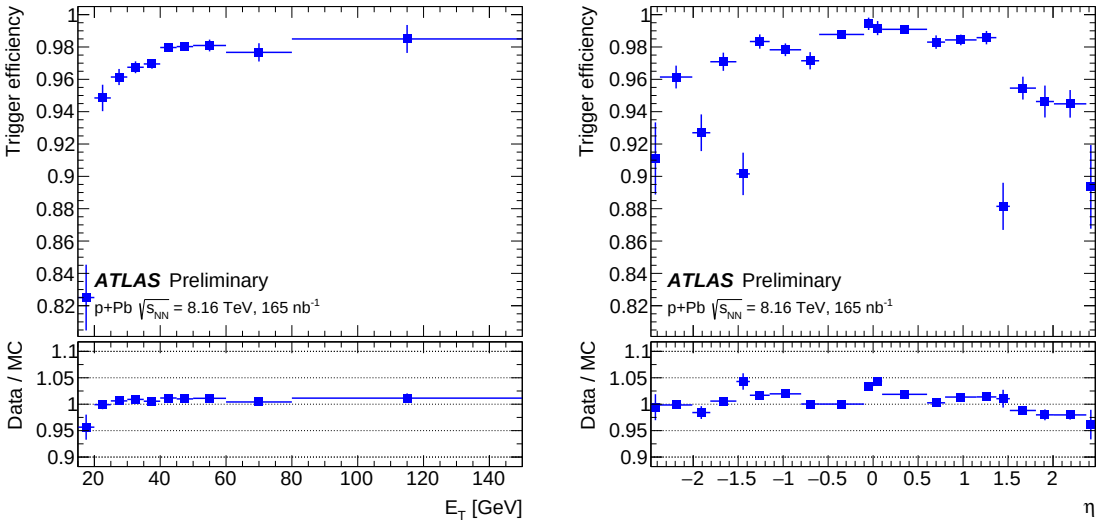


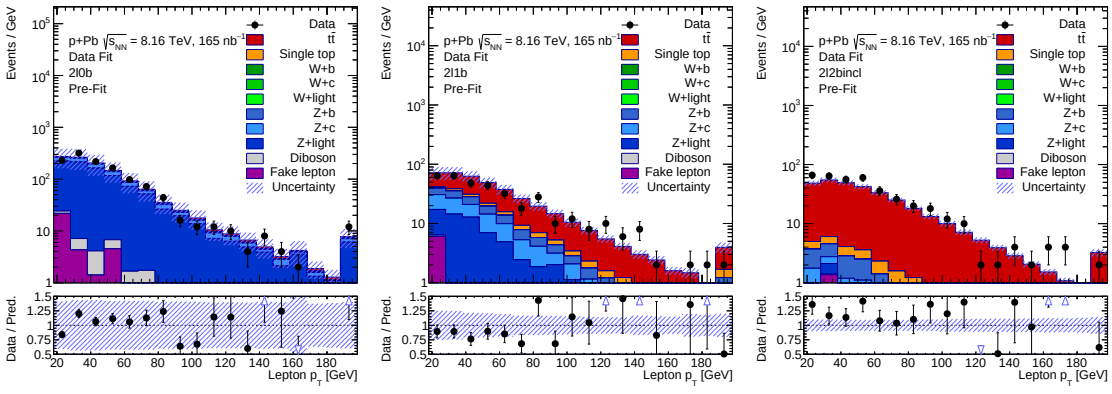
Figure 5.21: Electron trigger efficiency as a function of electron E_{T} (left) and η (right) for two triggers evaluated in the data (full markers) [79].

L Control Plots

The control plots presented in this section are used to validate the modeling of key kinematic distributions. They help assess how well the data agrees with the MC before and after fitting. Two categories of plots are provided: pre-fit and post-fit distributions, each offering insights into the data and simulation behavior.

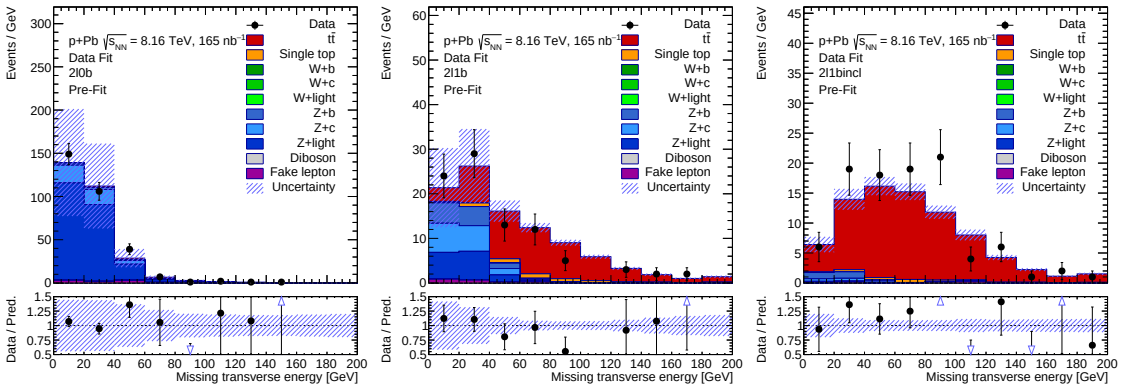
L.1 Pre-fit Plots

Pre-fit plots in Figures 5.31–5.40, show distributions before applying any constraints from the fit to data. These comparisons highlight the performance of the MC in describing data using initial predictions, with uncertainties represented by the blue hatched areas.



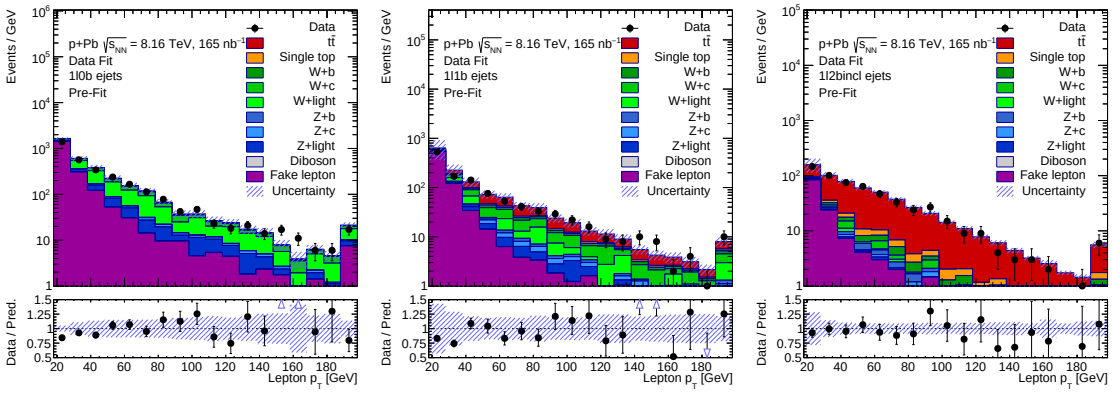
(a) (b) (c)

Figure 5.22: Pre-fit lepton p_T distributions in dilepton channel: (a) $2\ell 0b$, (b) $2\ell 1b$, (c) $2\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



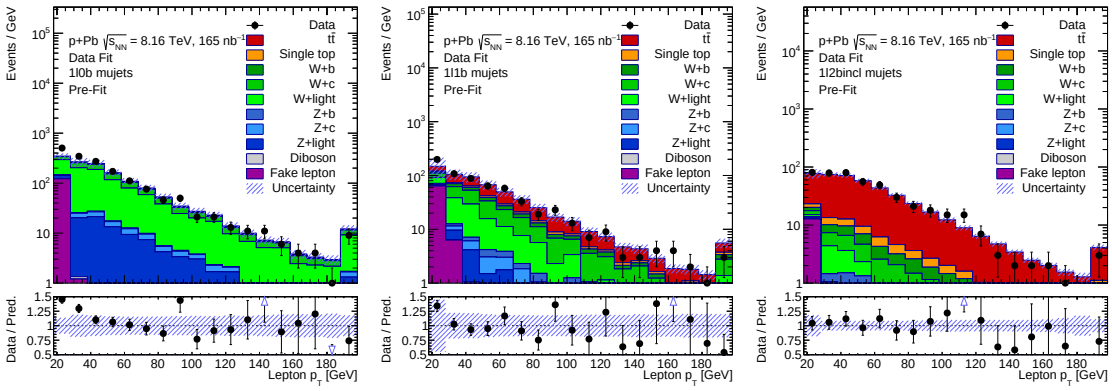
(a) (b) (c)

Figure 5.23: Pre-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in dilepton channel: (a) $2\ell 0b$, (b) $2\ell 1b$, (c) $2\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



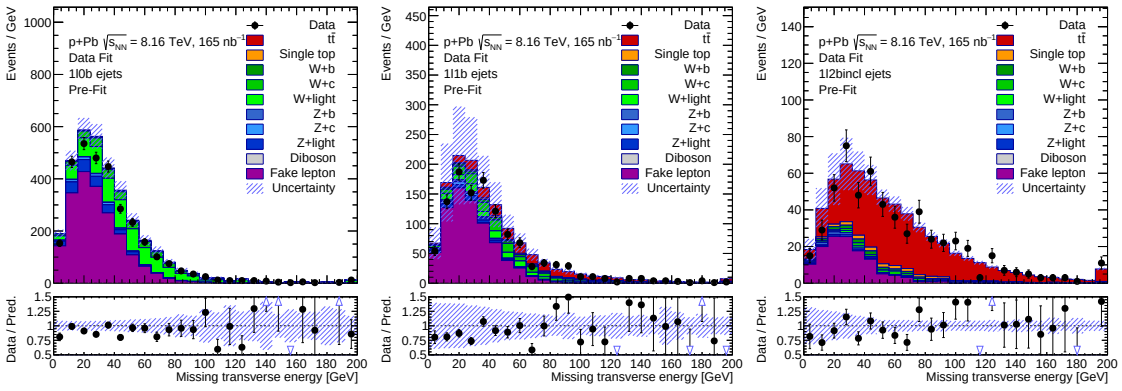
(a) (b) (c)

Figure 5.24: Pre-fit lepton p_T distributions in ℓ+jets events (electron channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



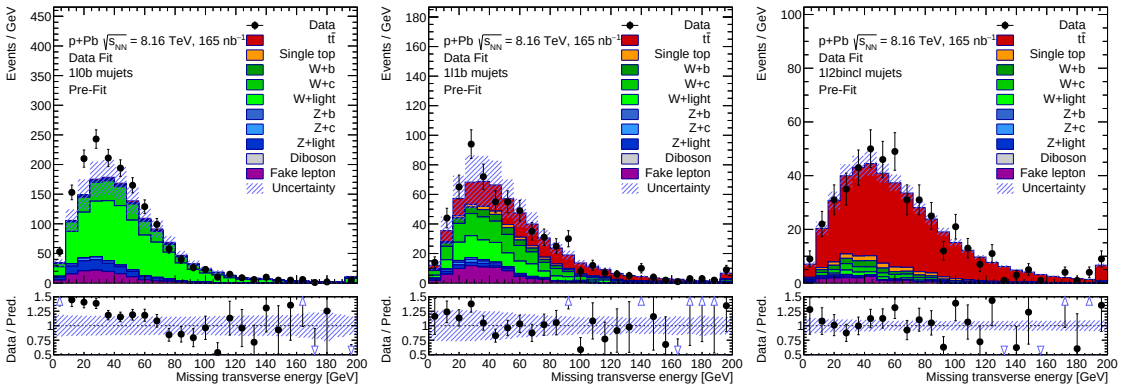
(a) (b) (c)

Figure 5.25: Pre-fit lepton p_T distributions in ℓ+jets events (muon channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



(a) (b) (c)

Figure 5.26: Pre-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in ℓ+jets events (electron channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.

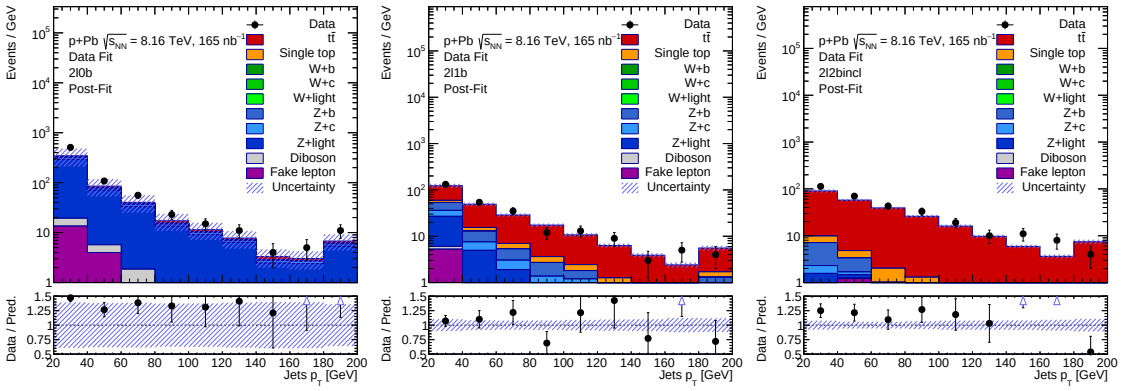


(a) (b) (c)

Figure 5.27: Pre-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in ℓ+jets events (muon channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.

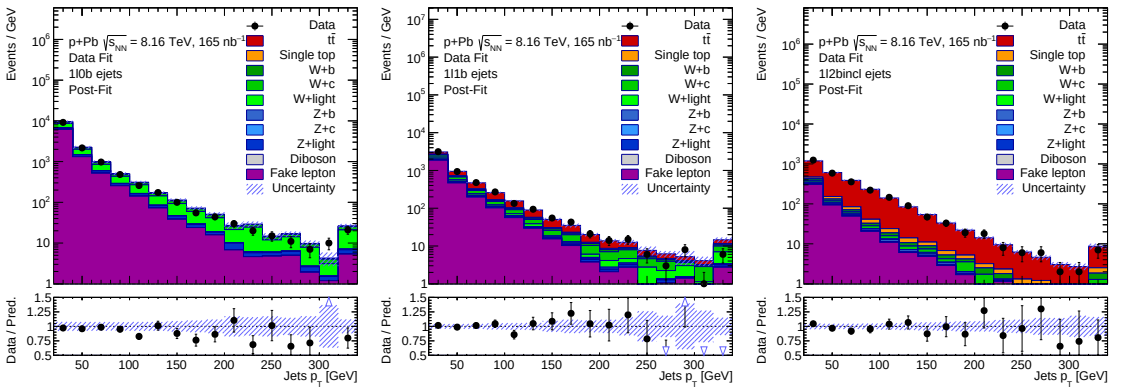
L.2 Post-fit Plots

Post-fit plots in Figures 5.28–5.40, display distributions after the fit procedure has been applied. These plots reflect the improved agreement between data and the adjusted MC, incorporating information from both the fit and data uncertainties.



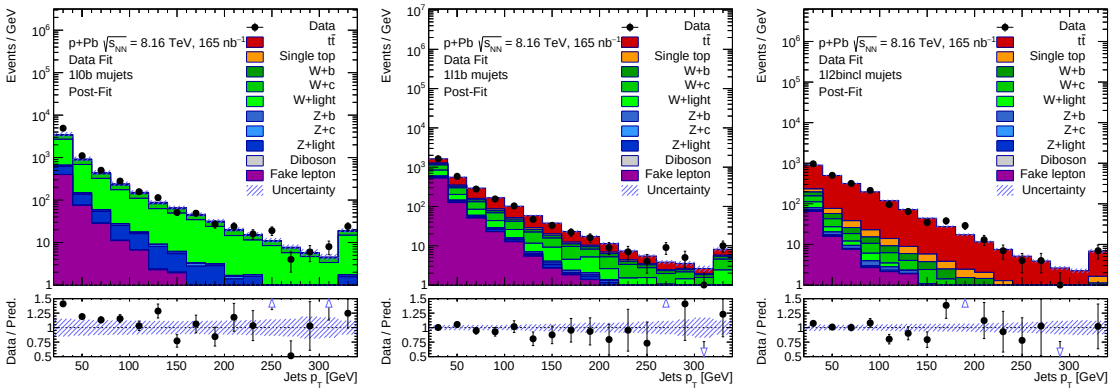
(a) (b) (c)

Figure 5.28: Post-fit jet p_T distributions in dilepton channel: (a) $2\ell 0b$, (b) $2\ell 1b$, (c) $2\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



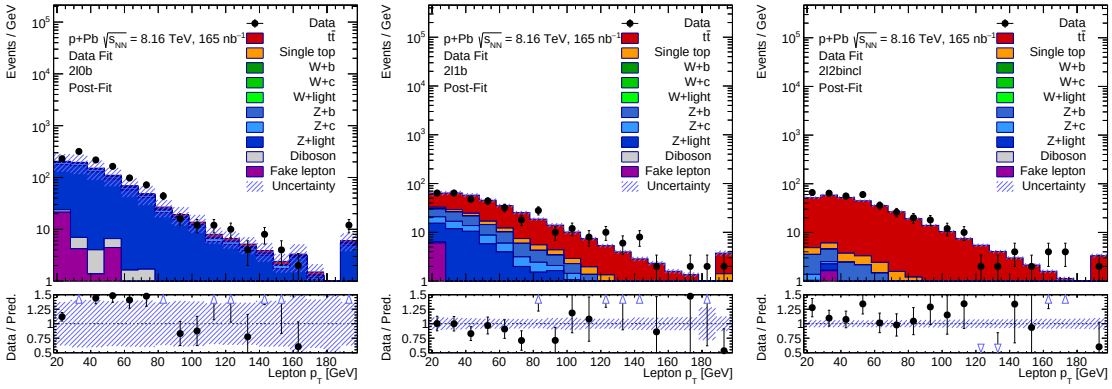
(a) (b) (c)

Figure 5.29: Post-fit jet p_T distributions in ℓ+jets events (electron channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



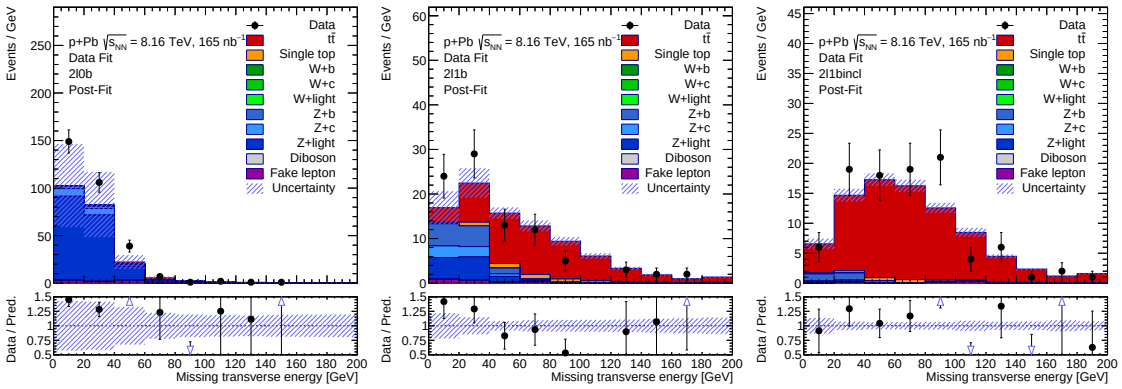
(a) (b) (c)

Figure 5.30: Post-fit jet p_T distributions in ℓ+jets events (muon channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



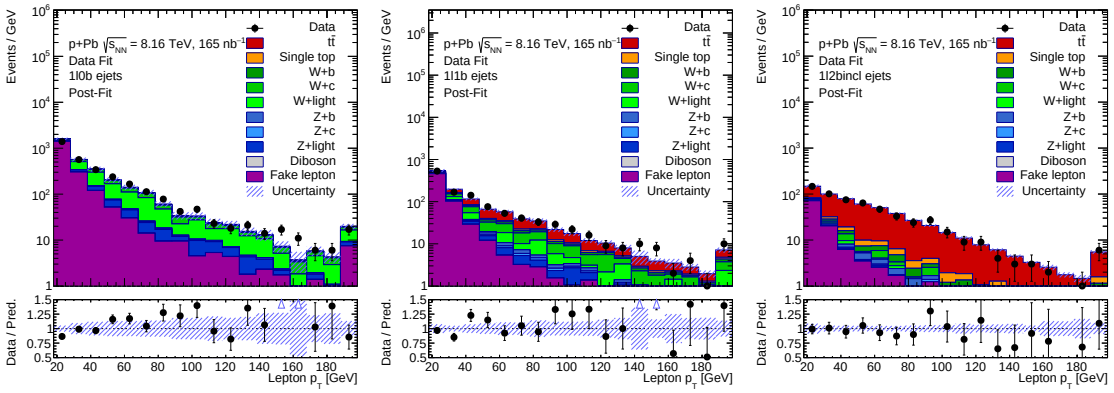
(a) (b) (c)

Figure 5.31: Post-fit lepton p_T distributions in dilepton channel: (a) $2\ell 0b$, (b) $2\ell 1b$, (c) $2\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



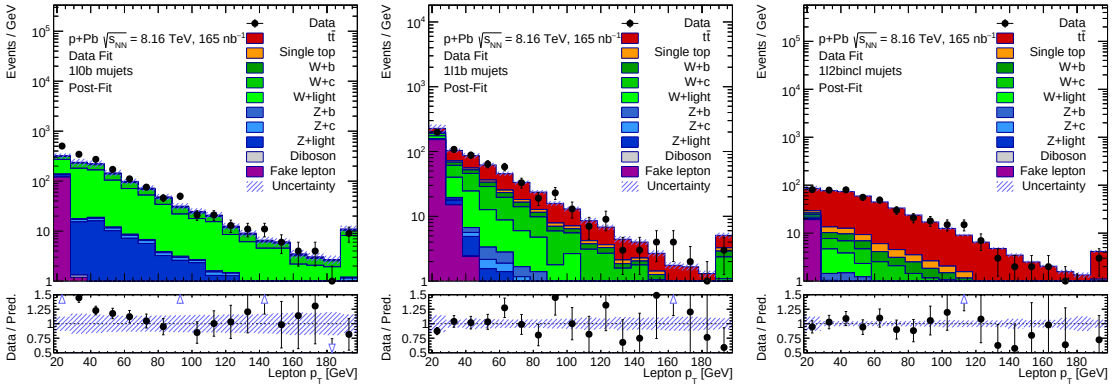
(a) (b) (c)

Figure 5.32: Post-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in dilepton channel: (a) $2\ell 0b$, (b) $2\ell 1b$, (c) $2\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



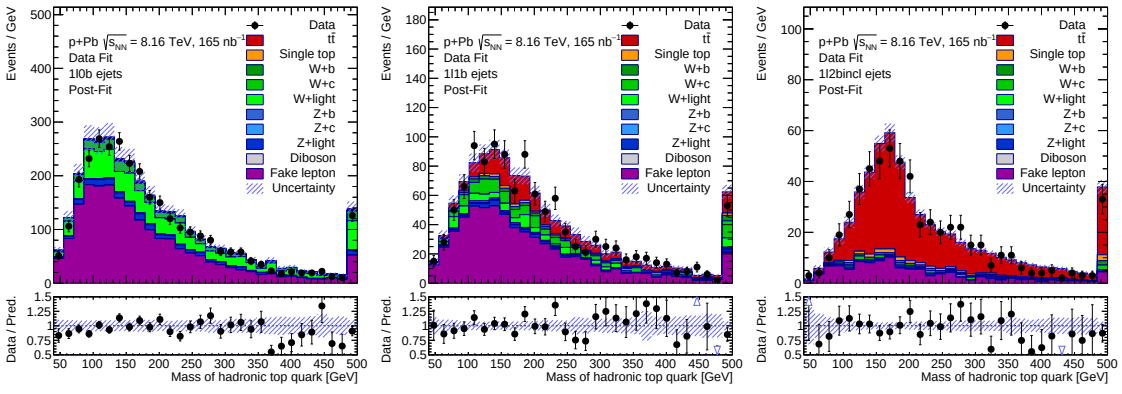
(a) (b) (c)

Figure 5.33: Post-fit lepton p_T distributions in ℓ+jets events (electron channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



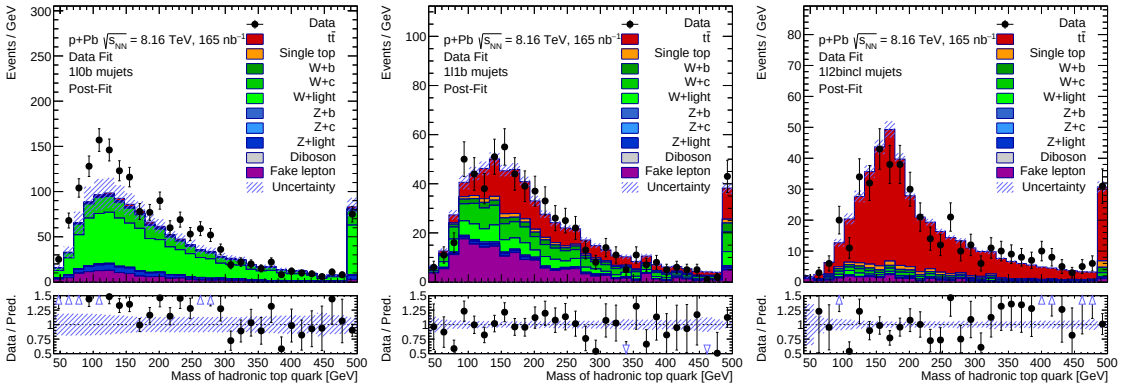
(a) (b) (c)

Figure 5.34: Post-fit lepton p_T distributions in ℓ+jets events (muon channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



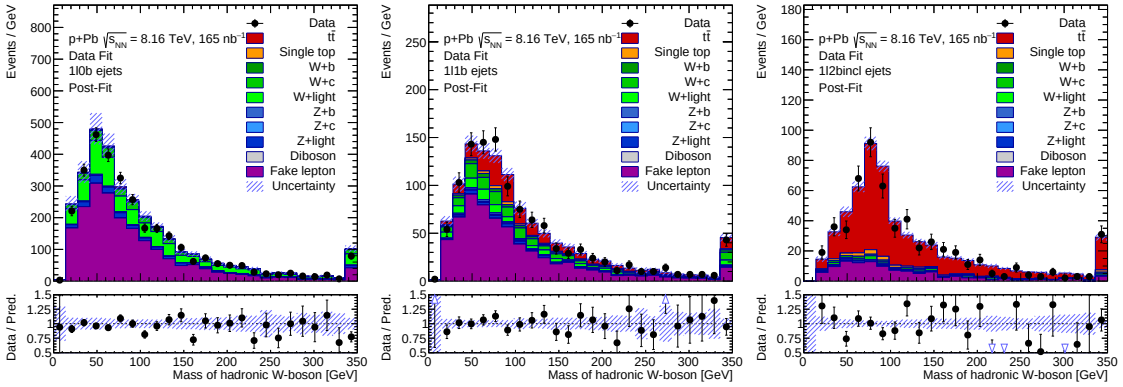
(a) (b) (c)

Figure 5.35: Post-fit of the mass of the hadronically decaying top quark in ℓ+jets events (electron channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



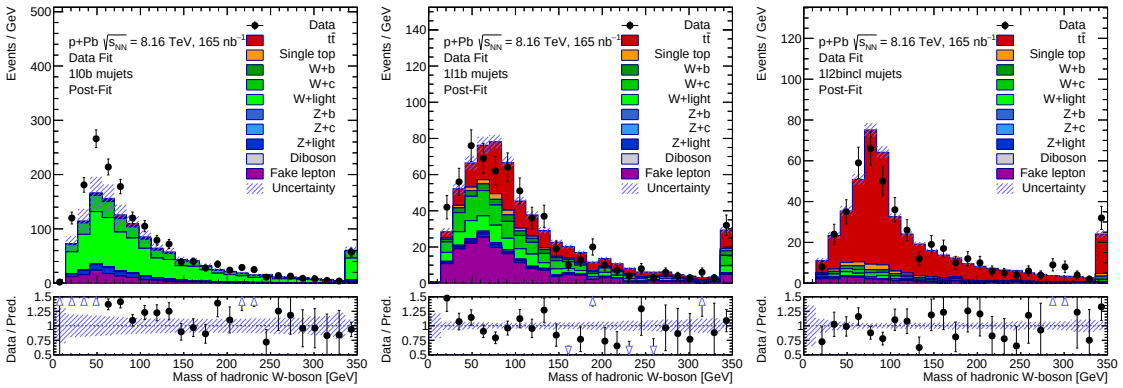
(a) (b) (c)

Figure 5.36: Post-fit of the mass of the hadronically decaying top quark in ℓ+jets events (muon channel): (a) 1ℓ0b, (b) 1ℓ1b, (c) 1ℓ2$b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



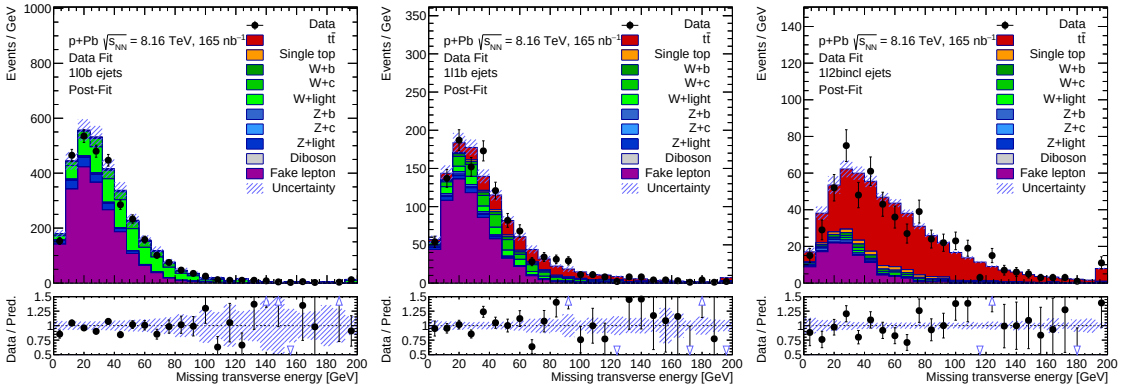
(a) (b) (c)

Figure 5.37: Post-fit of the mass of the hadronically decaying W in ℓ+jets events (electron channel): (a) 1ℓ0b, (b) 1ℓ1b, (c) 1ℓ2$b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



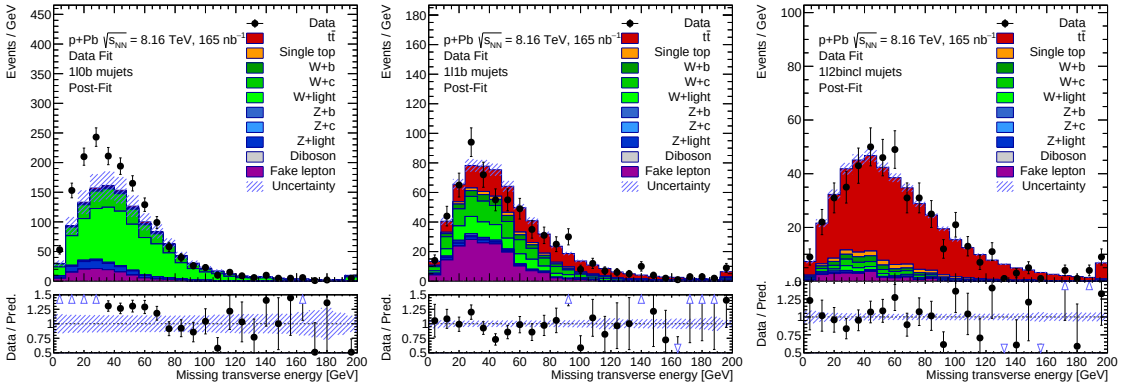
(a) (b) (c)

Figure 5.38: Post-fit of the mass of the hadronically decaying W in ℓ+jets events (muon channel): (a) 1ℓ0b, (b) 1ℓ1b, (c) 1ℓ2$b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



(a) (b) (c)

Figure 5.39: Post-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in ℓ+jets events (electron channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.



(a) (b) (c)

Figure 5.40: Post-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in ℓ+jets events (muon channel): (a) $1\ell 0b$, (b) $1\ell 1b$, (c) $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.

M Invariant Mass Distributions of the W Candidate used in CMS Measurement

Figure 5.41 presents the invariant mass distributions of the W boson candidate, denoted as $m_{jj'}$, in events with 0, 1, and 2 b-tagged jets. These distributions were used in the CMS cross-section measurement titled "Observation of Top Quark Production in Proton-Nucleus Collisions" [96]. The red and orange shaded areas represent the signal contributions, distinguishing between correctly and incorrectly assigned jets, while the blue region accounts for non-top background events. The error bars reflect the statistical uncertainties in the measurements.

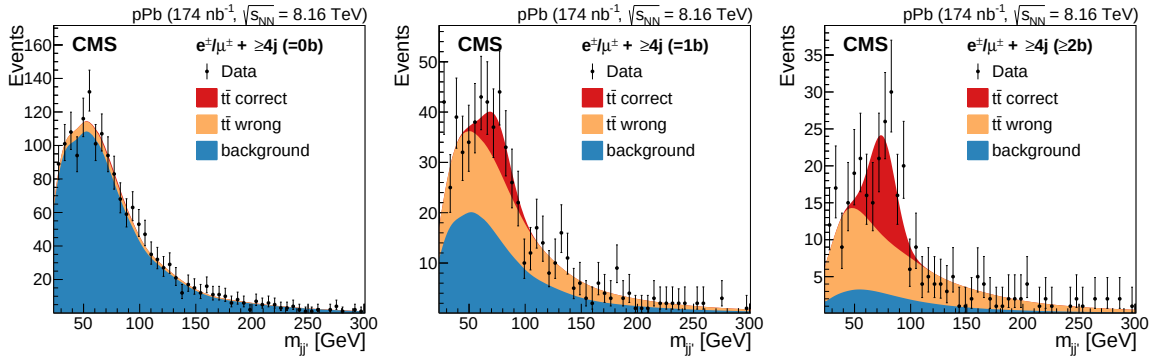


Figure 5.41: Invariant mass distributions of the W candidate, $m_{jj'}$, in the 0 (left), 1 (center), and 2 (right) b-tagged jet categories after all selections. The red and orange areas correspond to the signal simulation (correct and wrong assignments, respectively) while the blue one corresponds to the estimated non-top background contributions. The error bars indicate the statistical uncertainties [96].

N $W+$ Jets Uncertainties

This section discusses the uncertainties in the determination of the reconstructed $t\bar{t}$ kinematic variable shapes for events from Sherpa samples, focusing on the $W+$jets background. This section relates to the measurements of top-quark pair differential and double-differential cross-sections in the $\ell+$jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector [104].

Figure 5.42 illustrate the variations in the spectra when reweighting events based on different renormalization (μ_R) and factorization (μ_F) scales, as well as different values of α_S and PDF choices. The largest impact arises from scale variations, while PDF and α_S variations have a smaller effect.

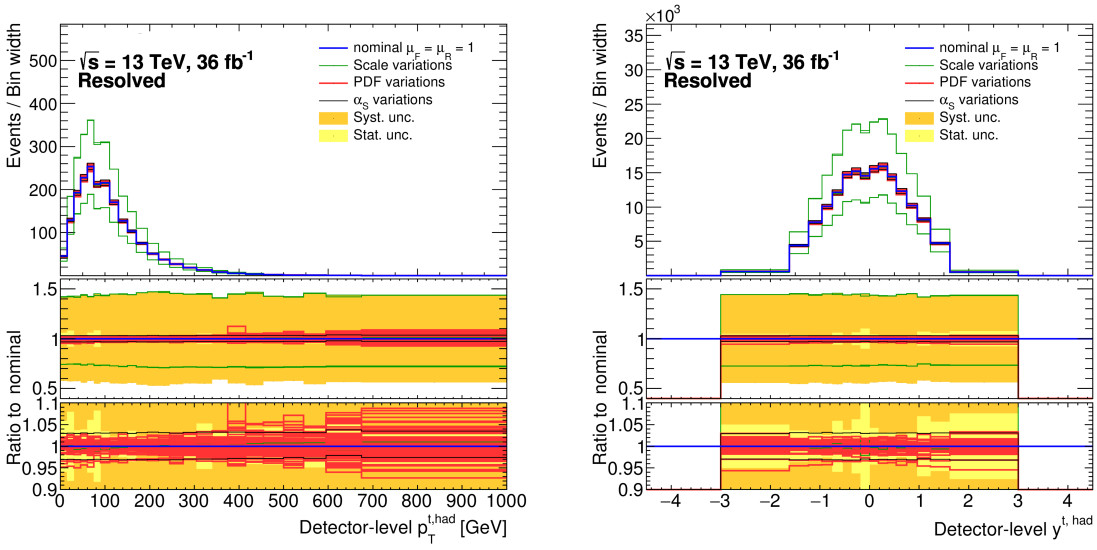


Figure 5.42: The p_{T} (left) and rapidity (right) distributions of the hadronic top quark candidate in the resolved regime for the Sherpa $W+$jets sample with variations in the scale, PDF and α_S choices.

O Extending the Fully Bayesian Unfolding with Regularization Using a Combined Sampling Method

Fully Bayesian Unfolding (FBU) is a method to correct for detector effects and recover the true particle-level spectrum from measured data. It uses Bayes' theorem to combine prior information with the likelihood of the data, producing a posterior distribution for the true spectrum.

The basic unfolding formula in FBU is:

$$P(\vec{T}|\vec{D}) = \frac{P(\vec{D}|\vec{T}) \cdot P(\vec{T})}{P(\vec{D})}, \tag{5.5}$$

where $P(\vec{T}|\vec{D})$ is the posterior probability of the true spectrum $\vec{T}$ given the data $\vec{D}$, $P(\vec{D}|\vec{T})$ is the likelihood of observing the data given $\vec{T}$, and $P(\vec{T})$ is the prior distribution for $\vec{T}$.

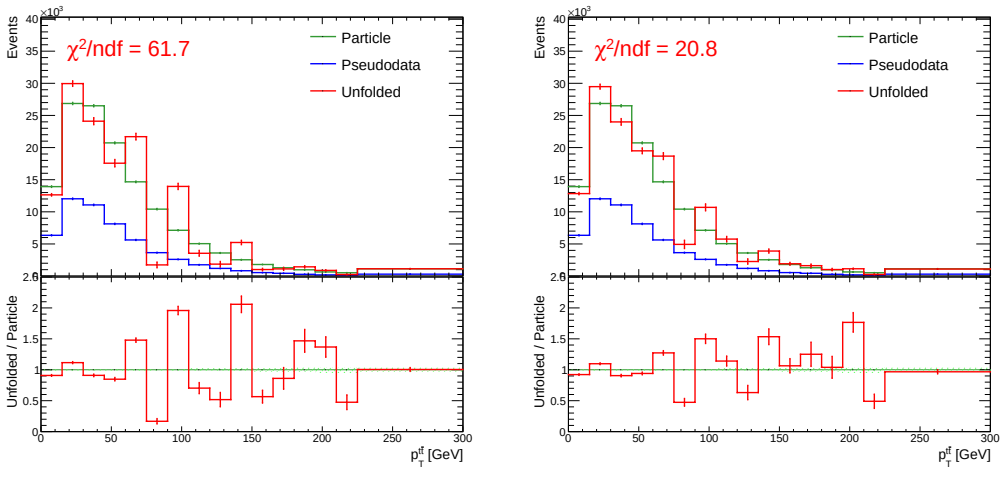
When regularization is introduced the convenient way is to use an exponential function with regularization strength parameter τ and an inner function $S(\vec{T})$

$$P(\vec{T}|\vec{D}) = \frac{P(\vec{D}|\vec{T}) \cdot e^{-\tau S(\vec{T})}}{\text{Norm.}}. \tag{5.6}$$

If the parameter $\tau = 0$, the prior $P(\vec{T}) = 1$, and no regularization is applied; on the other hand, the higher value of τ, the more dominant the regularization term. The Entropy-based regularization given as

$$S(\vec{T}) = -\left[-\sum_{t=1}^{N} \frac{T_t}{\sum T_{t'}} \log\left(\frac{T_t}{\sum T_{t'}}\right)\right]. \tag{5.7}$$

is used in this example, smoothing the unfolded spectrum and reducing statistical noise. The strength of the regularization, τ, is varied to minimize the χ^2/ndf, balancing bias and variance.



(a) (b)

Figure 5.43: Unfolding result (a) without regularization and (b) with entropy regularization of the $p_T^{t\bar{t}}$ spectrum [105].

Figure 5.43 shows the unfolding of the $p_T^{t\bar{t}}$ spectrum (a) without regularization and (b) with entropy regularization. Regularization significantly reduces fluctuations, leading to a

more stable spectrum. The optimal regularization parameter τ_{opt} is found by minimizing χ^2/ndf, ensuring a balance between smoothness and accuracy.

This method provides a robust unfolding framework, particularly in cases with limited data, and can be extended to other measurements. Additional details can be found in publication [105].

P Application of Machine Learning Based Top Quark and W Jet Tagging to Hadronic Four-Top Final States Induced by SM and BSM Processes

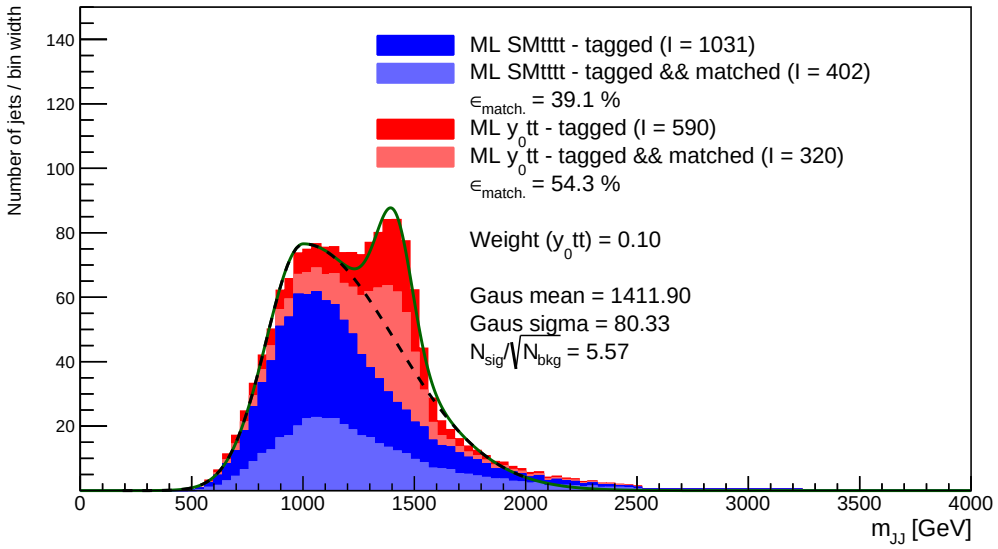
In this analysis, both machine learning (ML)-based and traditional cut-based techniques are used for jet tagging. ML method uses Gradient Boostingon three input features jet mass and jet substructure variables τ_{21} and τ_{32}, which are ratios of N-subjettiness variables that describe how well a jet matches a hypothesis of having 2 or 3 subjets, respectively. These variables provide a clear distinction between top quarks and W bosons from light jets, helping to reduce mistag rates.

The cut-based method applies simple thresholds on the same input variables:

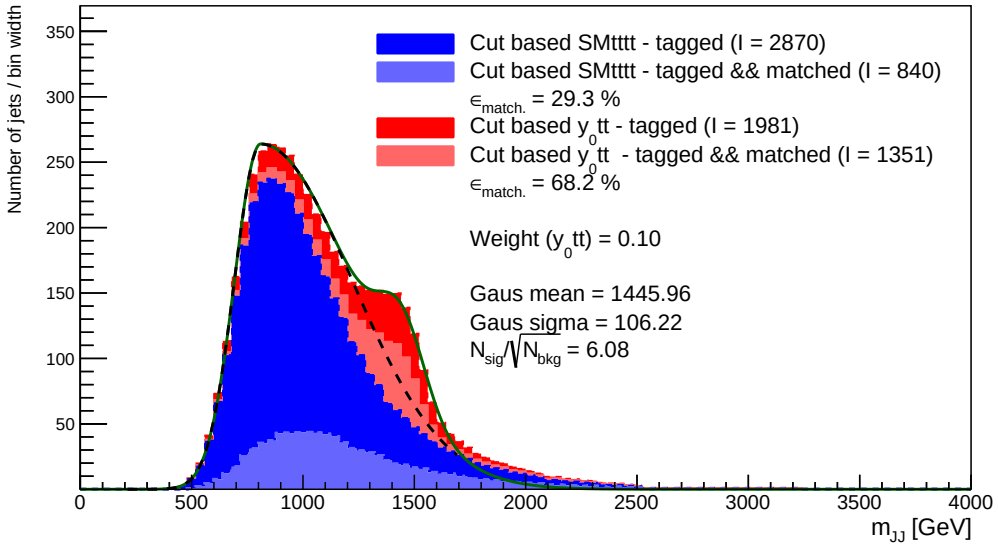
- W-jets if $0.10 < \tau_{21} < 0.60 \wedge 0.50 < \tau_{32} < 0.85 \wedge m_J \in [60{,}110]\,\text{GeV}$;
- top-jets if $0.30 < \tau_{21} < 0.70 \wedge 0.30 < \tau_{32} < 0.80 \wedge m_J \in [138{,}208]\,\text{GeV}$.

Figure 5.44 compares the performance of these two approaches. The invariant mass distributions of two t-tagged jets are shown for the SM $t\bar{t}t\bar{t}$ process (blue area) and the BSM signal process $t\bar{t}y_0 \to t\bar{t}t\bar{t}$ (red area), scaled to 10%. The light red and blue areas represent the tagged and matched jets, indicating tagging efficiency for both methods. The background fit is modeled with a Bifurcated Gaussian (black line), while the signal is fitted with a Gaussian (green line).

The ML-based method (5.44a) shows a tighter signal peak and reduced background, highlighting its superior performance over the cut-based method (5.44b), which shows a wider mass peak and higher background contamination. Additional details of the study can be found in [106].



(a) ML-based method.



(b) Cut-based method.

Figure 5.44: Invariant mass of two t-tagged jets (all possible combinations) for the process of SM $t\bar{t}t\bar{t}$ (blue area) representing background process with the stacked signal process $t\bar{t}y_0 \rightarrow t\bar{t}t\bar{t}$ (red area) scaled to its 10%. The light red and blue areas show tagged and matched jets to highlight the tagging efficiencies. The background fit is given by black line using Bifurcated Gaussian and green line is the Gaussian signal fit.

References

[1] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Physical Review Letters*, 13:508–509, October 1964. https://doi.org/10.1103/PhysRevLett.13.508.

[2] François Englert and Robert Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Physical Review Letters*, 13:321–323, 1964. https://doi.org/10.1103/PhysRevLett.13.321.

[3] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, A. A. Abdelalim, et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716:1–29, September 2012. http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[4] Joshua P. Ellis. TikZ-Feynman: Feynman diagrams with TikZ. *Computer Physics Communications*, 210:103–123, 2017. https://www.sciencedirect.com/science/article/pii/S0010465516302521.

[5] S. Abachi, B. Abbott, M. Abolins, et al. Observation of the Top Quark. *Physical Review Letters*, 74:2632–2643, 1995. https://doi.org/10.1103/PhysRevLett.74.2632.

[6] Nils Faltermann. Top quark mass and cross section at ATLAS and CMS, 2024. https://arxiv.org/abs/2405.09182.

[7] Michele Gallinaro. Top quark physics: A tool for discoveries. *Journal of Physics: Conference Series*, 447(1):012012, jul 2013. https://dx.doi.org/10.1088/1742-6596/447/1/012012.

[8] David d'Enterria. Top-quark pair production cross sections at NNLO+NNLL in pPb collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV, 2017. https://arxiv.org/abs/1706.09521.

[9] K.J. Eskola, H. Paukkunen, and C.A. Salgado. EPS09 — A new generation of NLO and LO nuclear parton distribution functions. *Journal of High Energy Physics*, 2009(04):065, apr 2009. https://dx.doi.org/10.1088/1126-6708/2009/04/065.

[10] Daniel de Florian, Rodolfo Sassot, Pia Zurita, and Marco Stratmann. Global analysis of nuclear parton distributions. *Phys. Rev. D*, 85:074028, Apr 2012. https://link.aps.org/doi/10.1103/PhysRevD.85.074028.

[11] L. Frankfurt, V. Guzey, and M. Strikman. Leading twist nuclear shadowing phenomena in hard processes with nuclei. *Physics Reports*, 512(4–5):255–393, March 2012. http://dx.doi.org/10.1016/j.physrep.2011.12.002.

[12] Luca Paolo Wiggering. *Monte Carlo Generators For $t\bar{t}$ Production.* Bachelor's thesis, University of Münster, 2019. https://www.uni-muenster.de/imperia/md/content/physik_tp/theses/klasen/wiggering_bsc.pdf.

[13] Wolfram Research, Inc. *Mathematica, Version 12.1.* Wolfram Research, Inc., Champaign, IL, 2020. https://www.wolfram.com/mathematica/.

[14] FeynCalc Developers. FeynCalc Installation. `InstallFeynCalc[]`, 2024. https://raw.githubusercontent.com/FeynCalc/feyncalc/master/install.m.

[15] P. Nason, S. Dawson, and R.K. Ellis. The total cross section for the production of heavy quarks in hadronic collisions. *Nuclear Physics B*, 303(4):607–633, 1988. https://doi.org/10.1016/0550-3213(88)90422-1.

[16] Pavel M. Nadolsky Jon Pumplin Tie-Jiun Hou, Joey Huston et al. New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC, 2019. https://arxiv.org/abs/1912.10053.

[17] Python Software Foundation. Python Language Reference, version 3.x, 2023. https://www.python.org/.

[18] Pauli Virtanen and et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. https://doi.org/10.1038/s41592-019-0686-2.

[19] Charles R. Harris and et al. Array programming with NumPy. *Nature*, 585:357–362, 2020. https://doi.org/10.1038/s41586-020-2649-2.

[20] Wes McKinney. Data Structures for Statistical Computing in Python. pages 56 – 61, 2010. https://doi.org/10.25080/Majora-92bf1922-00a.

[21] S. Alioli, P. Nason, C. Oleari, and E. Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *Journal of High Energy Physics*, 2010:043, 2010. https://doi.org/10.1007/jhep06(2010)043.

[22] Michal Czakon and Alexander Mitov. Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders. *Comput. Phys. Commun.*, 185:2930, 2014. https://doi.org/10.1016/j.cpc.2014.06.021.

[23] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Precision phenomenology with MCFM. *JHEP*, 12:034, 2019. https://doi.org/10.1007/JHEP12(2019)034.

[24] K. J. Eskola, Petja Paakkinen, Hannu Paukkunen, and Carlos A. Salgado. EPPS21: A global QCD analysis of nuclear PDFs. *Eur. Phys. J. C*, 82:413, 2022. http://dx.doi.org/10.1140/epjc/s10052-022-10359-0.

[25] Torbjörn Sjöstrand et al. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015. https://arxiv.org/abs/1410.3012.

[26] J. Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *European Physical Journal C*, 76:196, 2016. https://arxiv.org/abs/1512.01178.

[27] S. Agostinelli et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Methods A*, 506:250–303, 2003. https://doi.org/10.1016/S0168-9002(03)01368-8.

[28] M. Markovych and A. Tandogan. Introduction to Parton-Shower Event Generators. *arXiv preprint arXiv:1411.4085*, 2014. https://arxiv.org/abs/1411.4085.

[29] ATLAS Collaboration. ATLAS Inner Detector Geant4 Simulation. `https://atlas-computing.web.cern.ch/atlas-computing/packages/simulation/geant4/images/AtlasInnerDetector_9.jpg`. Accessed: September 17, 2024.

[30] R. L. Workman and Others. Review of Particle Physics. *PTEP*, 2022:083C01, 2022. http://dx.doi.org/10.1093/ptep/ptac097.

[31] Jakub Janda. pgf-pie: Pie charts with PGF/TikZ. CTAN, 2015. https://ctan.org/pkg/pgf-pie.

[32] R. Atkin. Review of jet reconstruction algorithms. *J. Phys.: Conf. Ser.*, 645:012008, 2015. High Energy Particle Physics Workshop (HEPPW2015), 11–13 February 2015, Johannesburg, South Africa, https://doi.org/10.1088/1742-6596/645/1/012008.

[33] G. Corcella et al. `Herwig` 6.5 release note. *arXiv preprint hep-ph/0210213*, 2002. https://arxiv.org/abs/hep-ph/0210213.

[34] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *Journal of High Energy Physics*, 2008:063, 2008. arXiv:0802.1189, https://arxiv.org/abs/0802.1189.

[35] K.J. Eskola, P. Paakkinen, H. Paukkunen, and C.A. Salgado. EPPS16: nuclear parton distributions with LHC data. *Eur. Phys. J. C*, 77:163, 2017. https://doi.org/10.1140/epjc/s10052-017-4725-9.

[36] K. Kovarik et al. nCTEQ15 - Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework. *Phys. Rev. D*, 93:085037, 2016. https://arxiv.org/abs/1509.00792.

[37] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Physical Review Letters*, 13:321–323, August 1964. https://doi.org/10.1103/PhysRevLett.13.321.

[38] The Large Hadron Collider, 2019. https://home.cern/science/accelerators/large-hadron-collider.

[39] L. Evans and P. Bryant. LHC Machine. *Journal of Instrumentation*, 3:S08001, August 2008. http://dx.doi.org/10.1088/1748-0221/3/08/S08001.

[40] ATLAS Collaboration. Luminosity delivered to ATLAS during Run 3 (2024), 2024. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun3.

[41] S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716:30–61, September 2012. http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[42] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3:S08003, August 2008. http://dx.doi.org/10.1088/1748-0221/3/08/S08003.

[43] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/PublishedTilecalFigures.

[44] ATLAS Pixel Detector: Technical Design Report, 2010. https://cds.cern.ch/record/381263.

[45] The ATLAS Detector, 2008. https://home.cern/science/experiments/atlas.

[46] The ATLAS Muon Spectrometer: Technical Design Report, 1997. https://cds.cern.ch/record/331068.

[47] David d'Enterria, Krisztián Krajczár, and Hannu Paukkunen. Top-quark production in proton-nucleus and nucleus-nucleus collisions at LHC energies and beyond. *Nucl. Phys. A*, 933:256–287, 2015. https://arxiv.org/pdf/1501.05879.

[48] ATLAS Collaboration. Supplementary Figures from: Observation of $t\bar{t}$ production in lepton+jets and dilepton channels in p+Pb collisions at $\sqrt{s_{\mathrm{NN}}}$ = 8.16 TeV with the ATLAS detector. 2023. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TOPQ-2023-32/.

[49] P. Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *Journal of High Energy Physics*, 2004:040, 2004. https://dx.doi.org/10.1088/1126-6708/2004/11/040.

[50] R. D. Ball and others (NNPDF Collaboration). Parton distributions for the LHC Run II. *Journal of High Energy Physics*, 2015:40, 2015. https://doi.org/10.1007/JHEP04(2015)040.

[51] ATLAS Collaboration. ATLAS Pythia 8 tunes to 7 TeV data. *ATL-PHYS-PUB-2014-021*, 2014. https://cds.cern.ch/record/1966419.

[52] David d'Enterria. High-precision α_s measurements from LHC to FCC-ee. *Proceedings of the International Workshop on Future Colliders*, 2017:69–78, 2017. https://arxiv.org/abs/1512.05194.

[53] Jonathan M. Butterworth, Amanda Cooper-Sarkar, Anna Hall, et al. PDF4LHC recommendations for LHC Run II. *Journal of Physics G: Nuclear and Particle Physics*, 43(2):023001, 2016. https://arxiv.org/abs/1510.03865.

[54] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne. Parton distributions in the LHC era: MMHT 2014 PDFs. *European Physical Journal C*, 75:204, 2015. https://doi.org/10.1140/epjc/s10052-015-3397-6.

[55] S. Dulat, T. J. Hou, J. Gao, et al. New parton distribution functions from a global analysis of quantum chromodynamics. *Physical Review D*, 93:033006, 2016. https://doi.org/10.1103/PhysRevD.93.033006.

[56] Johan Alwall, Rikkert Frederix, Stefano Frixione, Valentin Hirschi, Fabio Maltoni, Olivier Mattelaer, Hua-Sheng Shao, Tim Stelzer, Paolo Torrielli, and Marco Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics (JHEP)*, 07:079, 2014. https://arxiv.org/abs/1405.0301.

[57] Enrico Bothmann and Others. Event generation with Sherpa 2.2. *Science Reports*, 7:081, 2019. https://arxiv.org/abs/1905.09127.

[58] Tanju Gleisberg, Stefan Hoeche, Frank Krauss, Marek Schonherr, Steffen Schumann, Frank Siegert, and Jan Winter. Event generation with SHERPA and COMIX. *Journal of High Energy Physics (JHEP)*, 02:007, 2009. https://arxiv.org/abs/0811.4622.

[59] F. Cascioli, P. Maierhöfer, and S. Pozzorini. Scattering Amplitudes with Open Loops. *Physical Review Letters*, 108:111601, 2012. https://arxiv.org/abs/1111.5206.

[60] Steffen Schumann and Frank Krauss. A parton shower algorithm based on Catani-Seymour dipole factorisation. *Journal of High Energy Physics (JHEP)*, 03:038, 2008. https://dx.doi.org/10.1088/1126-6708/2008/03/038.

[61] Stefano Catani, Frank Krauss, Bryan R. Webber, and R. Kuhn. QCD Matrix Elements + Parton Showers. *Journal of High Energy Physics (JHEP)*, 11:063, 2001. https://dx.doi.org/10.1088/1126-6708/2001/11/063.

[62] J. Butterworth et al. Single Boson and Diboson Production Cross Sections in *pp* Collisions at $\sqrt{s} = 7$ TeV. Technical Report CERN-2010-010, CERN, 2010. https://cds.cern.ch/record/1287902.

[63] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Richard Rietkerk. Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations. *Journal of High Energy Physics (JHEP)*, 03:015, 2013. https://dx.doi.org/10.1007/JHEP03(2013)015.

[64] David J. Lange. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 462:152–155, 2001. https://dx.doi.org/10.1016/S0168-9002(01)00089-4.

[65] ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. *Journal of Instrumentation*, 14:P12006, 2019. https://dx.doi.org/10.1088/1748-0221/14/12/P12006.

[66] ATLAS Collaboration. Muon reconstruction and identification efficiency in ATLAS using the full Run 2 *pp* collision data set at $\sqrt{s} = 13$ TeV. *European Physical Journal C*, 81:578, 2021. https://dx.doi.org/10.1140/epjc/s10052-021-09233-2.

[67] ATLAS Collaboration. Measurement of the $t\bar{t}$ production cross-section in *pp* collisions at $\sqrt{s} = 5.02$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2023:138, 2023. https://dx.doi.org/10.1007/JHEP06(2023)138.

[68] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *European Physical Journal C*, 72:1896, 2012. https://dx.doi.org/10.1140/epjc/s10052-012-1896-2.

[69] T. T. Rinn, A. M. Sickles, and P. A. Steinberg. Measurement of Dijet Asymmetry in Pb+Pb and *pp* Collisions at 5.02 TeV, 2020. https://cds.cern.ch/record/2711618.

[70] ATLAS Collaboration. Deep Sets Based Neural Networks for Impact Parameter Flavour Tagging in ATLAS, 2020. https://cds.cern.ch/record/2718948.

[71] ATLAS Collaboration. ATLAS flavour-tagging algorithms for the LHC Run 2 *pp* collision dataset, 2023. https://cds.cern.ch/record/2842028.

[72] ATLAS Collaboration. Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC. *Journal of Instrumentation*, 18:T11004, 2023. https://dx.doi.org/10.1088/1748-0221/18/11/T11004.

[73] ATLAS Collaboration. Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 79(8):639, 2019. https://doi.org/10.1140/epjc/s10052-019-7140-6.

[74] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 76(5):292, 2016. https://doi.org/10.1140/epjc/s10052-016-4120-y.

[75] ATLAS Collaboration. Performance of the ATLAS muon triggers in Run 2. *Journal of Instrumentation*, 15(09):P09015, sep 2020. https://dx.doi.org/10.1088/1748-0221/15/09/P09015.

[76] F. Cardillo et al. Tools for estimating fake/non-prompt lepton backgrounds in ATLAS. Technical report, CERN, 1999. https://cds.cern.ch/record/2686919.

[77] ATLAS Collaboration. Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data. 2016. https://cds.cern.ch/record/2157687.

[78] ATLAS Collaboration. Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data. Tech. Rep. ATLAS-CONF-2014-032, CERN, 2014. https://cds.cern.ch/record/1706245.

[79] ATLAS Collaboration. Figures: Electron performance in 2016 p+Pb collisions. 2022. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2022-01.

[80] ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. *JINST*, 14:P12006, 2019. https://doi.org/10.1088/1748-0221/14/12/P12006.

[81] ATLAS Collaboration. Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 81(8):689, 2021. https://doi.org/10.1140/epjc/s10052-021-09402-3.

[82] ATLAS Collaboration. Jet energy scale and its uncertainty for jets reconstructed using the ATLAS heavy ion jet algorithm. 2015. https://cds.cern.ch/record/2008677.

[83] ATLAS Collaboration. ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 79(12):970, 2019. https://doi.org/10.1140/epjc/s10052-019-7450-8.

[84] ATLAS Collaboration. Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector. *Eur. Phys. J. C*, 82(2):95, 2022. https://doi.org/10.1140/epjc/s10052-022-10047-2.

[85] ATLAS Collaboration. Calibration of the light-flavour jet mistagging efficiency of the b-tagging algorithms with Z+jets events using 139 fb^{-1} of ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 83(8):728, 2023. https://arxiv.org/abs/2301.06319.

[86] C. F. Berger et al. Precise Predictions for W+4-Jet Production at the Large Hadron Collider. *Phys. Rev. Lett.*, 106:092001, 2011. https://doi.org/10.1103/PhysRevLett.106.092001.

[87] F. A. Berends, H. Kuijf, B. Tausk, and W. T. Giele. On the production of a W and jets at hadron colliders. *Nucl. Phys. B*, 357:32–64, 1991. https://doi.org/10.1016/0550-3213(91)90458-A.

[88] S. D. Ellis, R. Kleiss, and W. J. Stirling. W's, Z's and Jets. *Phys. Lett. B*, 154:435–440, 1985. https://doi.org/10.1016/0370-2693(85)90425-3.

[89] E. Re. Single-top Wt-channel production matched with parton showers using the POWHEG method. *Eur. Phys. J. C*, 71:1547, 2011. https://doi.org/10.1140/epjc/s10052-011-1547-z.

[90] ATLAS Collaboration. Observation of $t\bar{t}$ production in the lepton+jets and dilepton channels in $p + Pb$ collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV with the ATLAS detector. 2023. https://arxiv.org/abs/2405.05078.

[91] ATLAS Collaboration. Measurement of the $t\bar{t}$ cross section and its ratio to the Z production cross section using pp collisions at $\sqrt{s} = 13.6$ TeV with the ATLAS detector. *Phys. Lett. B*, 848:138376, 2024. https://doi.org/10.1016/j.physletb.2023.138376.

[92] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J. C*, 83:982, 2023. https://doi.org/10.1140/epjc/s10052-023-11747-w.

[93] G. Avoni et al. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *JINST*, 13:P07017, 2018. https://doi.org/10.1088/1748-0221/13/07/P07017.

[94] TRExFitter Collaboration. TRExFitter Documentation, 2024. https://trexfitter-docs.web.cern.ch/trexfitter-docs/.

[95] P.A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020. https://academic.oup.com/ptep/article/2020/8/083C01/5891211.

[96] CMS Collaboration. Observation of top quark production in proton-nucleus collisions. *Phys. Rev. Lett.*, 119:242001, 2017. https://arxiv.org/abs/1709.07411.

[97] ATLAS and CMS Collaborations. Combination of inclusive top-quark pair production cross-section measurements using ATLAS and CMS data at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 07:213, 2023. https://arxiv.org/abs/2205.13830.

[98] J. M. Campbell, D. Wackeroth, and J. Zhou. Study of weak corrections to Drell-Yan, top-quark pair, and dijet production at high energies with MCFM. *Phys. Rev. D*, 94:093009, 2016. https://arxiv.org/abs/1608.03356.

[99] K. J. Eskola, P. Paakkinen, H. Paukkunen, and C. A. Salgado. Towards EPPS21 nuclear PDFs. *SciPost Phys. Proc.*, 8:033, 2022. https://arxiv.org/abs/2106.13661.

[100] P. Duwentäster et al. Impact of heavy quark and quarkonium data on nuclear gluon PDFs. *Phys. Rev. D*, 105:114043, 2022. https://arxiv.org/abs/2204.09982.

[101] M. Ubiali. Towards the NNPDF3.0 parton set for the second LHC run. In *PoS DIS2014*, volume 041, 2014. https://arxiv.org/abs/1407.3122.

[102] R. Abdul Khalek et al. nNNPDF3.0: evidence for a modified partonic structure in heavy nuclei. *Eur. Phys. J. C*, 82:507, 2022. https://arxiv.org/abs/2201.12363.

[103] I. Helenius, M. Walt, and W. Vogelsang. TUJU21: nuclear PDFs with electroweak-boson data at NNLO. In *29th International Workshop on Deep-Inelastic Scattering and Related Subjects*, 2022. https://arxiv.org/abs/2207.04654.

[104] G. Aad, B. Abbott, D. C. Abbott, et al. Measurements of top-quark pair differential and double-differential cross-sections in the ℓ+jets channel with pp collisions at $\sqrt{s} =$ 13 TeV using the ATLAS detector. *European Physical Journal C*, 79(12):1028, 2019. https://doi.org/10.1140/epjc/s10052-019-7525-6.

[105] J. Baron P., Kvita. Extending the Fully Bayesian Unfolding with Regularization Using a Combined Sampling Method. *Symmetry*, 12:2100, 2020. https://doi.org/10.3390/sym12122100.

[106] P. Baroň, J. Kvita, R. Přívara, M. Machalová, R. Vodák, and J. Tomeček. Application of Machine Learning Based Top Quark and W Jet Tagging to Hadronic Four-Top Final States Induced by SM and BSM Processes. *arXiv preprint arXiv:2410.13904*, 2024. https://arxiv.org/abs/2410.13904.

[107] A. Hoecker et al. TMVA: Toolkit for Multivariate Data Analysis. *PoS ACAT 040*, 2007. https://arxiv.org/abs/physics/0703039.

[108] ATLAS Collaboration. Performance of electron and photon triggers in ATLAS during LHC Run 2. 2019. https://arxiv.org/abs/1909.00761.

[109] P. A. Janus. *Measurement of W boson production in Pb+Pb collisions at 5.02 TeV with the ATLAS detector*. PhD thesis, CERN, 2019. http://cds.cern.ch/record/2690570.

Palacký University Olomouc
Faculty of Science
Department of Experimental physics

SUMMARY OF DOCTORAL THESIS

Study of Heavy Particles Production and Properties with the ATLAS Experiment

Author:	**Mgr. Petr Baroň**
Study programme:	Physics
Specialization:	Applied Physics
Study form:	Daily
Supervisor:	Mgr. Jiří Kvita, Ph.D.
Opponent 1:	prof. RNDr. Stanislav Tokár, DrSc.
Opponent 2:	Mgr. Peter Berta, Ph.D.
Opponents's reviews available:	Study Department of the Faculty of Science, Palacký University Olomouc, 17. listopadu 1192/12, 779 00 Olomouc, Czech Republic

Bibliografická identifikace

Jméno a příjmení	Mgr. Petr Baroň
Název práce	Zkoumání produkce a vlastností těžkých částic v experimentu ATLAS
Typ práce	Disertační
Pracoviště	Katedra experimentální fyziky
Vedoucí práce	Mgr. Jiří Kvita, Ph.D.
Rok obhajoby práce	2024
Abstrakt	Cílem této disertační práce je popsat proces měření účinného průřezu produkce párů top kvarů ve srážkách proton-olovo při energii srážek na jeden nukleon $\sqrt{s_{\mathrm{NN}}} =$ 8,16 TeV v experimentu ATLAS na urychlovači LHC. Vzhledem k tomu, že top kvark má největší hmotnost ze všech elementárních částic, slouží jako sonda pro jaderné partonové distribuční funkce (nPDF) při vysokých hodnotách Bjorkenova x. Jaderné partonové distribuční funkce jsou běžně používány ve fyzice těžkých iontů k popisu jevů jako například stínění, anti-stínění a Fermiho pohyb ve srovnání s distribučními funkcemi ve srážkách proton-olovo. Měření produkce párů top kvarků ve srážkách proton-olovo představuje další datový bod pro fitování jaderných partonových fukncí, a je proto velmi důležité provést měření s co nejmenšími možnými nejistotami. Bylo tak potřeba provést analýzu včetně dodatečných studií pro prostředí s nízkou mírou pile-up efektu ve srážkách proton-olovo, jako je například odvození škálovacích faktorů pro elektrony a miony, energetické korekce jetů, systematiky tvaru fake lepton pozadí nebo provedení párování mezi dvěmi kolekcemi jetů. Strategie měření a oblasti fázového prostoru pro fitování musely být pečlivě zvoleny, aby bylo dosaženo maximální statistiky a poměru signálu k pozadí. Jako referenční měření pro analýzu slouží měření produkce párů top kvarů v kanále lepton+jets ve srážkách proton-olovo, zveřejněné experimentem CMS. Studie v této disertační práci představuje první analýzu, která měří páry top kvarků nejen v lepton+jets, ale také v dileptonovém kanále. Tato práce poskytuje naměřený účinný průřez produkce párů top kvarků ve srážkách proton-olovo v dileptonovém a lepton+jets kanále v experimentu ATLAS na urychlovači LHC v laboratoři CERN.
Klíčová slova	top kvark, srážky proton-olovo, nPDF, ATLAS experiment, velký hadronový urychlovač, měření účinného průřezu, fyzika těžkých iontů
Počet stran	54
Jazyk	anglický

Bibliographical identification

First name and surname	Mgr. Petr Baroň
Title	Study of Heavy Particles Production and Properties with the ATLAS Experiment
Type of thesis	Dissertation
Department	Department of Experimental Physics
Supervisor	Mgr. Jiří Kvita, Ph.D.
The year of presentation	2024
Abstract	The aim of this thesis is to describe the process of measuring the cross-section of the top quark pairs production in proton-lead collisions at the energy per nucleon $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV with the ATLAS experiment at the Large Hadron Collider (LHC) at CERN. Since the top quark has the largest mass of all elementary particles it serves as a probe of nuclear parton distribution functions (nPDFs) at high Bjorken-x. The nPDF are commonly used in Heavy Ion physics and describe different phenomena (e.g. shadowing, anti-shadowing, Fermi movement) compared to proton-proton collision processes. As the measurement of the $t\bar{t}$ production in p+Pb collisions represents a new data point in the fit of nPDF sets, it is very important to perform the analysis with the lowest uncertainties possible. The cross-section measurement had to be performed along additional studies for low pile-up in the p+Pb environment, namely derivation of electron and muon scale factors, jet energy scale and jet energy resolution corrections, shape systematics of the fake lepton background or performing matching between two jet collections. The strategy and phase-space regions of the global fit to the signal strength had to be chosen carefully to gain the maximum statistics and signal to background ratio. The reference analysis is the measurement of $t\bar{t}$ production in the lepton+jets channel in p+Pb collisions released by the CMS experiment, however, the study presented in this thesis is the first analysis to search for the $t\bar{t}$ production not only in lepton+jets channel, but also in the dilepton channel. This thesis provides the total measured cross-section of $t\bar{t}$ pairs in p+Pb collisions in lepton+jets and dilepton channels with the ATLAS experiment by LHC at CERN.
Keywords	top quark, proton-lead collisions, nPDF, ATLAS experiment, Large Hadron Collider, cross-section measurement, heavy-ion physics
Number of pages	54
Language	English

Statement

The doctoral thesis was carried out under the full-time doctoral program Physics in the specialization Applied Physics in the period 2018–2024. The research institute at which the thesis was prepared was the Joint Laboratory of Optics of UP. The program was maintained by the Department of Experimental Physics of the Faculty of Science at Palacký University in Olomouc.

The summary of the thesis has been sent to distribution on .....

The oral defense in front of the committee for the Ph.D. study program Physics and the specialization Applied Physics will take place on ..... at the Joint Laboratory of Optics, room ....., 17. listopadu 50A, Olomouc.

The doctoral thesis is available in the library at the Faculty of Science of Palacký University in Olomouc.

Acknowledgements

I would like to express my sincere thanks to my supervisor, Mgr. Jiří Kvita, Ph.D., for his guidance, willingness, and supportive approach throughout my studies. The knowledge I have gained will be invaluable in my future scientific career. I would also like to thank the management and staff of the Joint Laboratory of Optics (JLO) for their support and the pleasant working environment. My heartfelt thanks also go to my family for their support and patience during my studies.

The author gratefully acknowledge the support from the projects IGA_PrF_2024_004 of Palacky University, and Getting new knowledge of the microworld using the CERN infrastructure (Project MSMT I-E-T LTT17018), Research infrastructure for experiments at CERN (Projects MSMT LM2018104 and LM2023040), Novel techniques for boosted top quarks reconstruction for new physics searches at LHC (GACR 19-21484S), and Machine learning applications in high-energy and astroparticle physics (GACR 23-07110S).

Table of Contents

Introduction

This thesis focuses on measuring the cross-section of the $t\bar{t}$ production in proton-lead collisions at the center-of-mass energy of $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV using the ATLAS detector at the Large Hadron Collider (LHC). Top-quark pair production in proton-lead collisions has not been extensively studied, and this measurement represents a step toward constraining nPDFs, particularly at high Bjorken-x, where theoretical uncertainties remain large.

The first chapter of this thesis, titled "Theory", provides the needed theoretical background. It begins with exploring the properties of the top quark, its decay, mass, short lifetime, and significance as a probe for new physics, especially in relation to $t\bar{t}$ production.

The second chapter, "Measurement of $t\bar{t}$ production in lepton+jets and dilepton channels in p+Pb collisions," describes the experimental setup and analysis approach of the key measurement of this thesis. It starts by outlining the scientific motivation for studying $t\bar{t}$ production in p+Pb collisions and provides details on the datasets used for the analysis, including proton-lead collision data from the ATLAS experiment. The criteria for object selection, such as leptons, jets, and missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$), are defined, followed by the event selection criteria to optimize the signal and background separation. The performance of muon reconstruction, including efficiency and scale factor corrections, is discussed. The chapter also outlines various sources of systematic uncertainties affecting the measurement and discusses the statistical methods used to extract the signal strength through a fit procedure. Finally, the results of the analysis, including the measured $t\bar{t}$ production cross-section in the lepton+jets and dilepton channels, are presented and compared with theoretical predictions and previous measurements.

Theory

This chapter covers main points necesary to understand top quark pair production, decay modes, detector signature, differences between pp and p+Pb collisions environment with respect of the measurement described in the Chapter 2. By the end of this chapter the motivation and importance of the measurement as a new data point in unexplored kinematic region for the fit of nPDFs should emerge as well as the effort to keep measurement systematic uncertainties as low as possible, but at the same time keep uncertainties conservative enough and under control of analysers.

1.1 Top Quark

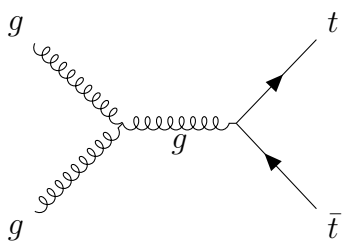
The top quark discovered in 1995 at Fermilab's Tevatron in $p\bar{p}$ collisions [1] is the heaviest elementary particle known to date and therefore an important topic to study in the context of the Standard Model at the Large Hadron Collider. Top quarks at the LHC are predominantly produced as top quark-antiquark pairs $t\bar{t}$ (or four top production) via the strong interaction, but can also be produced as single top quarks via the electroweak interaction. Since the top quark almost always decays into a W boson and a bottom quark before hadronization, the final state is characterized by the subsequent decay of the W boson, which can either happen into a quark-antiquark pair or into a charged lepton and the corresponding neutrino. Due to their abundant production at the LHC and their clear experimental signature top quarks provide a unique tool for probing the Standard Model parameters and search for deviations to theoretical predictions, which would be an indication for physics beyond the SM [2].

1.1.1 Top Quark Pair Production

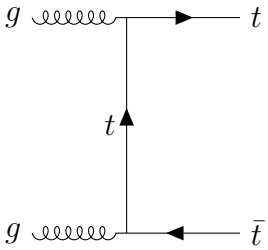
Top quark pairs are produced at leading order via gluon-gluon fusion and quark anti-quark annihilation, see Feynman diagrams in Figure 1.1.

At Tevatron, the proton-antiproton collider, top quark pairs were produced in 90% [4] via the quark anti-quark annihilation, due to high probability of colliding quark originating from proton with anti-quark of the same flavor originating from anti-quark and relatively high Bjorken-x (momentum fraction of the partons compare to hadrons), while at the LHC the gluon-gluon fusion is the dominant production mechanism (80% [4], $\sqrt{s} = 7$ TeV in pp collisions) due to low Bjorken-x and its large gluon density in this region.

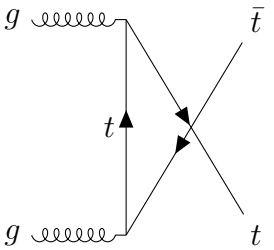
In the rest of this section the calculation of Bjorken-x is shown to demonstrate its relatively low values in case of production of $t\bar{t}$ pairs in p+Pb collisions. By approximating partons to be massless compare to centre-of-mass energy (CME), assuming



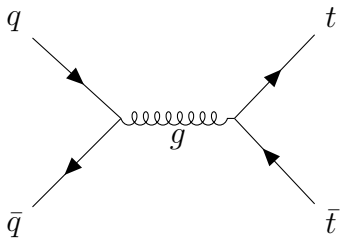
(a) $gg \to t\bar{t}$, s-channel.



(b) $gg \to t\bar{t}$, t-channel.



(c) $gg \to t\bar{t}$, u-channel.



(d) $q\bar{q} \to t\bar{t}$, s-channel.

Figure 1.1: Feynman diagrams of the production of $t\bar{t}$ pairs at the leading order. Generated using the TikZ-Feynman package [3].

asymetric proton-lead collisions, the Bjorken-x's of two partons can be written as

$$
\begin{aligned}
s_{\text{hadr.}} &= (P_1 + P_2)^2 = (E_1 + E_2,\, 0,\, 0,\, E_1 - E_2)^2 = 4E_1E_2 && (1.1)\\
p_{\text{part.}} &= x \cdot E_{\text{hadr.}}(1,\, 0,\, 0,\, \pm 1) && (1.2)\\
s_{\text{part.}} &= (p_1 + p_2)^2 = (x_1E_1 + x_2E_2,\, 0,\, 0,\, x_1E_1 - x_2E_2)^2 = 4x_1x_2E_1E_2 && (1.3)\\
&= x_1x_2s_{\text{hadr.}} && (1.4)\\
\Rightarrow x_1 &= \frac{s_{\text{part.}}}{s_{\text{hadr.}}}\frac{1}{x_2}, && (1.5)
\end{aligned}
$$

where $E_{1,2}$ are the energies and $P_{1,2}$ four-momenta of colliding proton and lead ion, while $p_{1,2}$ is the four-momentum of the partons. Using pseudorapidity of $t\bar{t}$ system $\eta_{t\bar{t}}$ the x_2 can be expressed as a function of x_1:

$$
\begin{aligned}
\eta_{t\bar{t}} &= \frac{1}{2}\log\left(\frac{E + p_z}{E - p_z}\right) = \frac{1}{2}\log\left(\frac{(E_1x_1 + E_2x_2 + E_1x_1 - E_2x_2)}{(E_1x_1 + E_2x_2 - E_1x_1 + E_2x_2)}\right) && (1.6)\\
&= \frac{1}{2}\log\left(\frac{E_1x_1}{E_2x_2}\right) && (1.7)\\
\Rightarrow \frac{1}{x_2} &= \frac{e^{2\eta_{t\bar{t}}}E_2}{x_1E_1}. && (1.8)
\end{aligned}
$$

By inserting Eq. 1.8 to the Eq. 1.5 the formulae of Bjorken-x is given as

$$
x_1 = \sqrt{\frac{s_{\text{part}}E_2}{s_{\text{hadr.}}E_1}}e^{\eta_{t\bar{t}}}. \tag{1.9}
$$

In case of studied $t\bar{t}$ production at $\sqrt{s_{\text{NN}}} = 8.16$ TeV in proton-lead collisions the Bjorken-x ranges within the interval [0.003 - 0.517] as the ATLAS detector covers pseudorapidity $\eta \in [-2.5,2.5]$

$$
x_1 = \frac{2m_t}{\sqrt{s_{\text{hadr.}}}}e^{\eta_{t\bar{t}}} = \frac{2 \cdot 172.69}{8.16 \cdot 10^3}e^{\{-2.5,-0.465,0,0.465,2.5\}} \simeq \{0.003,0.026,0.042,0.068,0.517\} \tag{1.10}
$$

Given the Bjorken-x region interval [0.003 - 0.517] roughly corresponds to the range of the x-axis in the publication [5].

Hadronic Cross-section at the Leading Order of QCD

Assuming accelerator with colliding particles in opposite direction the four vectors of incoming partons are given as

$$k = E\begin{pmatrix}1\\ \overrightarrow{e}_z\end{pmatrix}, \quad k' = E\begin{pmatrix}1\\ -\overrightarrow{e}_z\end{pmatrix}, \tag{1.11}$$

and outgoing quarks as

$$p = \begin{pmatrix}E\\ \sqrt{E^2-m_t^2}\sin(\theta)\\ 0\\ \sqrt{E^2-m_t^2}\cos(\theta)\end{pmatrix}, \quad p' = \begin{pmatrix}E\\ -\sqrt{E^2-m_t^2}\sin(\theta)\\ 0\\ -\sqrt{E^2-m_t^2}\cos(\theta)\end{pmatrix}. \tag{1.12}$$

The Mandelstam variables t and u can be expressed as a function of $\cos\theta$ and s

$$\begin{aligned} t &= (k-p)^2 = -2k\cdot p + m_t^2 = -2(E^2 - E\sqrt{E^2-m_t^2}\cos\theta) + m_t^2 \\ &= -\frac{s}{2} + m_t^2 + \cos\theta\sqrt{\frac{E^2}{4}(E^2-m_t^2)} = -\frac{s}{2} + m_t^2 + \cos\theta\sqrt{s\left(\frac{s}{4}-m_t^2\right)} \end{aligned} \tag{1.13}$$

and

$$\begin{aligned} u &= (k-p')^2 = -2k\cdot p' + m_t^2 = -2(E^2 + E\sqrt{E^2-m_t^2}\cos\theta) + m_t^2 \\ &= -\frac{s}{2} + m_t^2 - \cos\theta\sqrt{\frac{E^2}{4}(E^2-m_t^2)} = -\frac{s}{2} + m_t^2 - \cos\theta\sqrt{s\left(\frac{s}{4}-m_t^2\right)}. \end{aligned} \tag{1.14}$$

The partonic cross-section can be expressed as

$$\frac{d\sigma}{d\Omega} = \frac{\overline{|M|}^2\sqrt{1-\frac{4m_t^2}{s}}}{64\pi^2 s} \tag{1.15}$$

which for quark anti-quark annihilation and gluon fusion assuming $N=3$, $T_R=\frac{1}{2}$, $C_F=\frac{4}{3}$, $C_A=3$, $g_s=\sqrt{4\pi\alpha_s}$, $\beta=\sqrt{1-\frac{4m_t^2}{s}}$, and $\rho=\frac{4m_t^2}{s}$ gives

$$\frac{d\sigma_{qq}}{d\Omega} = \frac{\alpha_s^2\beta}{18s}\left(\cos^2(\theta)\beta^2+\rho+1\right) \tag{1.16}$$

$$\begin{aligned} \frac{d\sigma_{gg}}{d\Omega} &= \frac{\alpha_s^2\beta}{192s(1-\cos^2(\theta)\beta^2)^2}[-9\cos^6(\theta)\beta^6 - \cos^4(\theta)\beta^4(18\rho+7) \\ &+ 7\left(-2\rho^2+2\rho+1\right) + \cos^2(\theta)\left(18\rho^3-22\rho^2-5\rho+9\right)]. \end{aligned} \tag{1.17}$$

Integrating over $\cos\theta$ the partonic cross-sections become

$$\sigma_{qq} = \frac{4\pi\alpha_s^2\beta(\rho+2)}{27s} \tag{1.18}$$

$$\sigma_{gg} = \frac{\pi\alpha_s^2}{48s}\left((2\rho^2+32\rho+32)\,\mathrm{artanh}(\beta) - \beta(31\rho+28)\right) \tag{1.19}$$

which is identical to results in [6].

To obtain hadronic cross-section a convolution with PDFs, i.e. the integrals have to be evaluated

$$\sigma_{\text{hadr}} = \int_{s_{\min}}^{S} ds \int_{s/S}^{1} dx_1 \int_{s_{\min}/Sx_1}^{1} dx_2 \ \delta(s - x_1 x_2 S) f_a(x_1,\mu_F) f_b(x_2,\mu_F) \sigma(s) \tag{1.20}$$

where parameters were set as $m_t = 172.5$ GeV, $\alpha_s \approx 0.122$ (taken from pdf at m_t), $s_{\min} = 4m_t^2$, $S = (8160 \text{ GeV})^2$, and parton distribution function CT18LO [7] was chosen. The result of the numerical integration gives hadronic cross-sections

$$\sigma_{qq,\text{ hadr}} = 36.81 \pm 2.89 \text{ pb} \tag{1.21}$$

$$\sigma_{gg,\text{ hadr}} = 138.74 \pm 5.77 \text{ pb} \tag{1.22}$$

$$\Rightarrow \sigma_{\text{LO, tot, hadr}} = 175.55 \pm 6.46 \text{ pb}, \tag{1.23}$$

where the uncertainties represent the precision of the numerical integration performed by the SciPy package. The result is in agreement with the Powheg [8] Monte Carlo generator which with the same parameters setup and pdfs CT18LO [7] and CT18NLO [7] provides

$$\sigma_{\text{Powheg, LO}} = 181.05 \ \pm \ 5.26 \text{ pb} \tag{1.24}$$

$$\sigma_{\text{Powheg, NLO}} = 228.99 \ \pm \ 3.34 \text{ pb}. \tag{1.25}$$

To obtain the next-to-next-to-leading (NNLO) correction with next-to-next-to-leading logarithmic (NNLL) accuracy, the k-factor

$$k \equiv \sigma_{\text{NNLO}}/\sigma_{\text{NLO}} = 271.99/230.41 = 1.180 \tag{1.26}$$

calculated using TOP++package [9] with pdf sets CT18NLO [7] and CT18NNLO [7] is applied also to other NLO predictions.

Focusing on proton-lead collisions, the MCFM [10] generator was used to calculate the NLO cross-section setting one colliding particle as lead and the other as proton with pdf sets CT18NLO [7] and
EPPS21nlo_CT18Anlo_Pb208 nuclear pdf [11] resulting in

$$\sigma_{\text{pPb,MCFM, NLO}} = 254.81 \pm 0.28 \text{ pb}. \tag{1.27}$$

Multiplying by the k-factor and the number of nucleons within the lead $A = 208$, the next-to-next-to-leading cross-section of $t\bar{t}$ production in proton-lead collisions at $\sqrt{s_{\text{NN}}} = 8.16$ TeV is estimated

$$\sigma_{\text{pPb,MCFM, NNLO}} = 254.81 \cdot 1.180 \cdot 208 = 62.54 \text{ nb}. \tag{1.28}$$

1.1.2 Top Quark Decay

The top quark is unique among quarks because it is the only quark that decays exclusively via the weak interaction to a W boson, and it decays before hadronizing due to its extremely large mass and short lifetime. The top quark has a width of approximately $\Gamma_t \approx 1.35\,\text{GeV}$, corresponding to an exceptionally short lifetime of about $\tau_t \approx 0.5 \times 10^{-24}\,\text{seconds}$ [12]. The decay of a top quark (t) occurs predominantly via the channels

$$t \to W^+ b, \quad \bar{t} \to W^- \bar{b}$$

Each W-boson can decay either [12]

- Hadronically (into quarks): BR($W \to q\bar{q}'$) = 67.6%;
- Leptonically (into leptons e,μ,τ): BR($W \to \ell\nu_\ell$) = 32.4%;

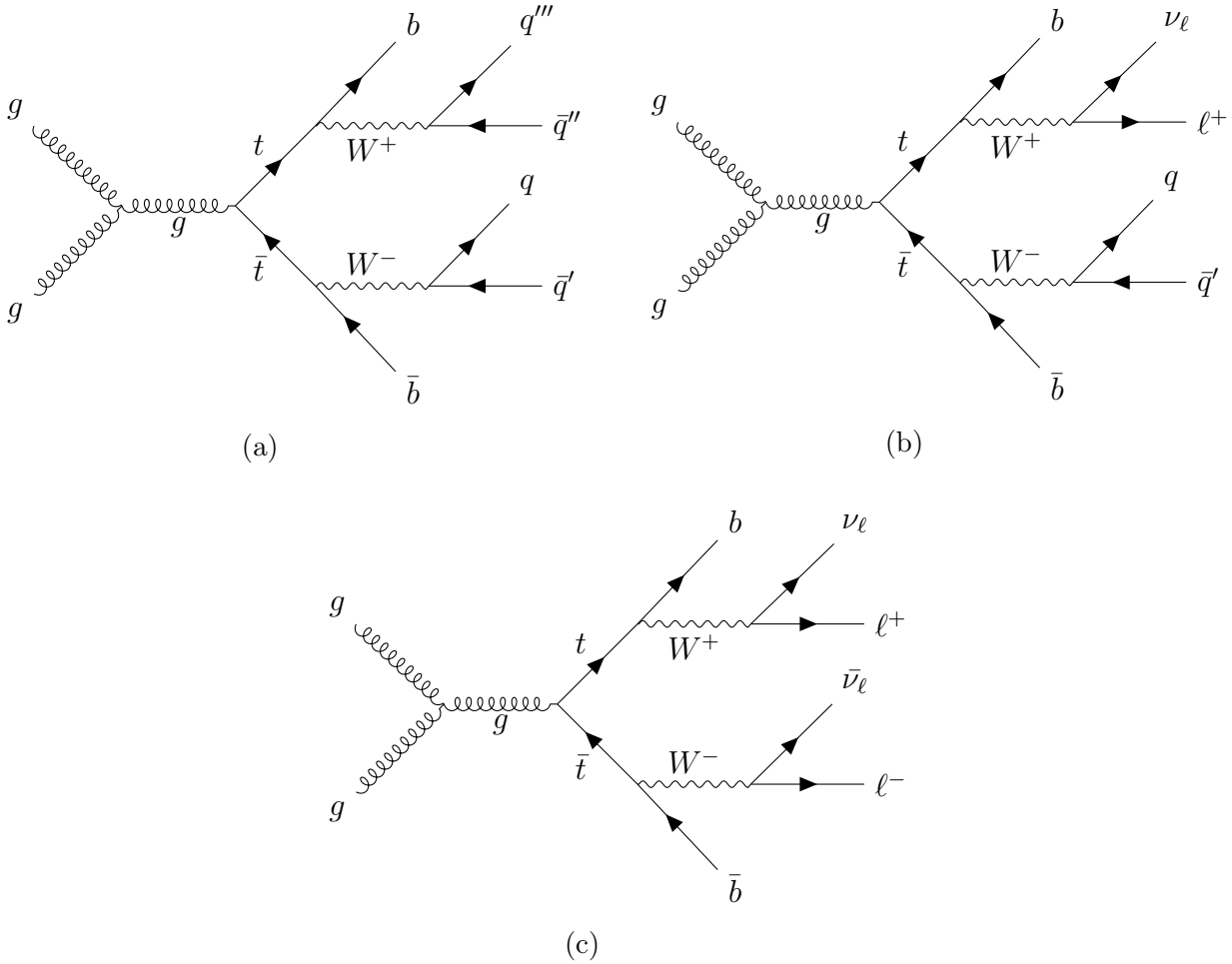


Figure 1.2: Feynman diagrams for $t\bar{t}$ production via gluon-gluon fusion in the s-channel, followed by different decay modes of the top and anti-top quarks. (a) Lepton+jets decay: one W boson decays leptonically ($W^+ \to \ell^+\nu_\ell$) and the other hadronically ($W^- \to q\bar{q}'$). (b) Dilepton decay: both W^+ and W^- bosons decay leptonically ($W^+ \to \ell^+\nu_\ell$ and $W^- \to \ell^-\bar{\nu}_\ell$). (c) All-hadronic decay: both W bosons decay hadronically ($W^+ \to q\bar{q}'$ and $W^- \to \bar{q}q'$). In all cases, the top and anti-top quarks decay to a W boson and a bottom quark ($t \to W^+b$, $\bar{t} \to W^-\bar{b}$). Generated using the TikZ-Feynman package [3].

Figure 1.3 shows a pie chart summarizing the branching ratios of $t\bar{t}$ decay.

1.2 Nuclear Parton Distribution Functions

Four distinct regions (shown in Figure 1.4) observed in nuclear parton distribution functions (nPDFs) correspond to well-known phenomena that modify the momentum distribution of quarks and gluons in nuclei, compared to free protons [14, 15].

- In the small $x < 0.01$ region, parton densities in nuclei are reduced relative to those of free nucleons. This phenomenon, known as shadowing, occurs because

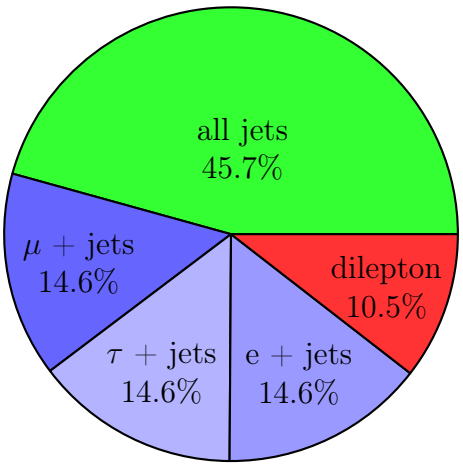


Figure 1.3: Decay channels perentage in $t\bar{t}$ production. Generated using pdf-pie [13] package.

gluons from different nucleons overlap and interfere inside the nucleus. Shadowing effects are prominent in high-energy scattering processes, particularly in heavy nuclei.

- At intermediate $0.1 < x < 0.3$, nuclear parton densities are enhanced compared to free nucleons. This compensatory effect, known as anti-shadowing, ensures the momentum sum rule for partons inside nucleons is satisfied. Anti-shadowing primarily affects gluons and sea quarks.

- The European Muon Collaboration (EMC) effect refers to the suppression of parton densities in the moderate $0.3 < x < 0.7$ region. The exact cause of the EMC effect remains a topic of debate, with several models attributing it to modifications in quark confinement and binding energy within bound nucleons.

- At high $x > 0.7$, parton distributions in nuclei are affected by the motion of nucleons within the nucleus, known as the Fermi motion. This shifts the parton momentum distribution to higher x, leading to an enhancement in parton densities, particularly for valence quarks.

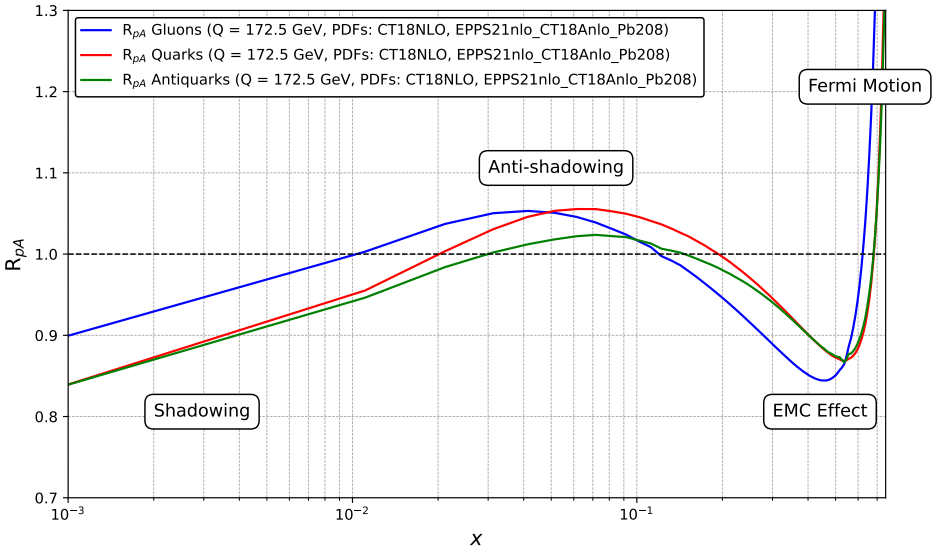


Figure 1.4: Nuclear modification factor R_{pA} of the parton distribution functions for gluons (blue), quarks (red), and antiquarks (green) as a function of the Bjorken x in Pb^{208} at a fixed scale $Q = 172.5$ GeV. The plot shows four distinct regions that reflect different nuclear effects: shadowing at low x, anti-shadowing at intermediate x, the EMC effect at moderate x, and Fermi motion at high x. The nuclear PDFs used are from the EPPS21nlo_CT18Anlo set, and the free proton PDFs are from CT18NLO. The solid horizontal line at $R_{pA} = 1$ indicates no nuclear modification.

Measurement of $t\bar{t}$ production in lepton+jets and dilepton channels in p+Pb collisions

2.1 Motivation

The $t\bar{t}$ production process is a valuable probe of the heavy ion environment that has not yet been fully exploited. A measurement of $t\bar{t}$ production in lepton+jets and dilepton channels in proton-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV (see event display in Figure 2.1) could reduce the uncertainties in the gluon distribution at intermediate and small x, where the shadowing and antishadowing effects are most prominent are expected to play a major role in nuclear collisions.

Figure 2.2 demonstrates a significant reduction in the uncertainty on the gluon distribution after incorporating proton-lead pseudodata, especially in the momentum fraction range $10^{-3} < x < 10^{-1}$.

Even if the data consists of a single cross section, it can still provide meaningful constraints on the normalization of the gluon distribution in nuclei. This is especially important because gluon-gluon interactions dominate top quark pair production. The procedure of reweighting shows that such data can shift the central value of the gluon distribution and significantly tighten the uncertainty bands, improving the precision of nPDFs [16].

2.2 The ATLAS Detector

The ATLAS (A Toroidal LHC ApparatuS) detector is one of the two general-purpose detectors at the Large Hadron Collider, the other being the CMS detector. This section provides an overview of the structure, functionality, and purpose of each layer of the ATLAS detector.

The ATLAS detector is the largest volume particle detector ever constructed, measuring 46 meters in length, 25 meters in diameter, and weighing around 7,000 tonnes [18], it covers nearly the entire solid angle around the interaction point.

The ATLAS detector uses a right-handed coordinate system with its origin at the interaction point in the center of the cylindrical detector, and the z-axis aligned along the beamline. The x-axis points from interaction point to the center of the LHC ring, and the y-axis points upwards from interaction point. In the transverse plane, cylindrical coordinates (r,ϕ) are used, where ϕ represents the azimuthal angle around the z-axis. The pseudorapidity, η, is defined in relation to the polar angle θ by the expression $\eta = -\ln\tan(\theta/2)$. Angular distance is defined as $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

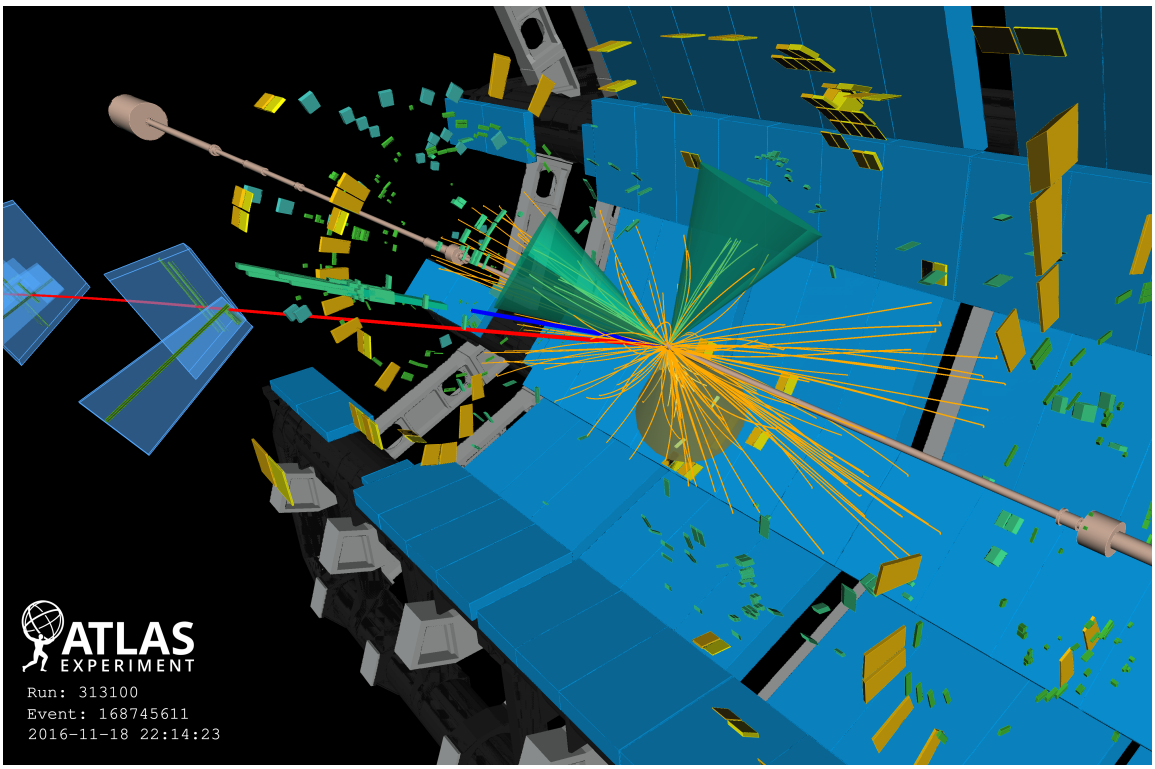


Figure 2.1: Event display of a candidate event for $t\bar{t}$ production in proton-lead ion collision decaying in the dilepton channel with the reconstructed electron track (blue line) and its associated energy deposit in the electromagnetic calorimeter (green boxes), the reconstructed muon track (red line) and its associated muon chambers (blue boxes; the bars are the related measurements), one b-tagged jet (yellow cone) and two non-b-tagged jets (green cones), and tracks in the inner detector (orange lines). Yellow rectangles correspond to energy deposits in cells of the hadron calorimeter [17].

The detector comprises multiple layers of different subsystems, each designed to measure specific properties of particles produced in proton-proton collisions at the LHC. The overall structure of ATLAS can be divided into several key subsystems (inner Detector, electromagnetic and hadronic calorimeter, muon spectrometer).

2.3 Data

The data utilized for this measurement were obtained using the ATLAS detector during the collision period involving proton-lead (p+Pb) interactions in 2016, accounting for an integrated luminosity of 165 nb^{-1}. The energy of the nominal proton beam and ion was 6.5 TeV, which leads in energy per nucleon of lead 2.56 TeV given by equation

$$E_{\mathrm{Pb}/A} \equiv \frac{E}{A} = \mathrm{E}_p \cdot \frac{Z}{A} = 6.5 \text{ TeV} \cdot \frac{82}{82+126} \doteq 2.56 \text{ TeV}, \tag{2.1}$$

where Z stands for number of protons and A for number of nucleons in lead ion, considering only lead isotop with 126 neutrons used by LHC. In total once proton collides with proton or neutron within the lead ion the center-of-mass energy per nuclean reads

$$\sqrt{s_{\mathrm{NN}}} = \sqrt{(P_p + P_{\mathrm{Pb}/A})^2} = \sqrt{(6.5+2.56)^2 - (6.5-2.56)^2} \doteq 8.16 \text{ TeV}. \tag{2.2}$$

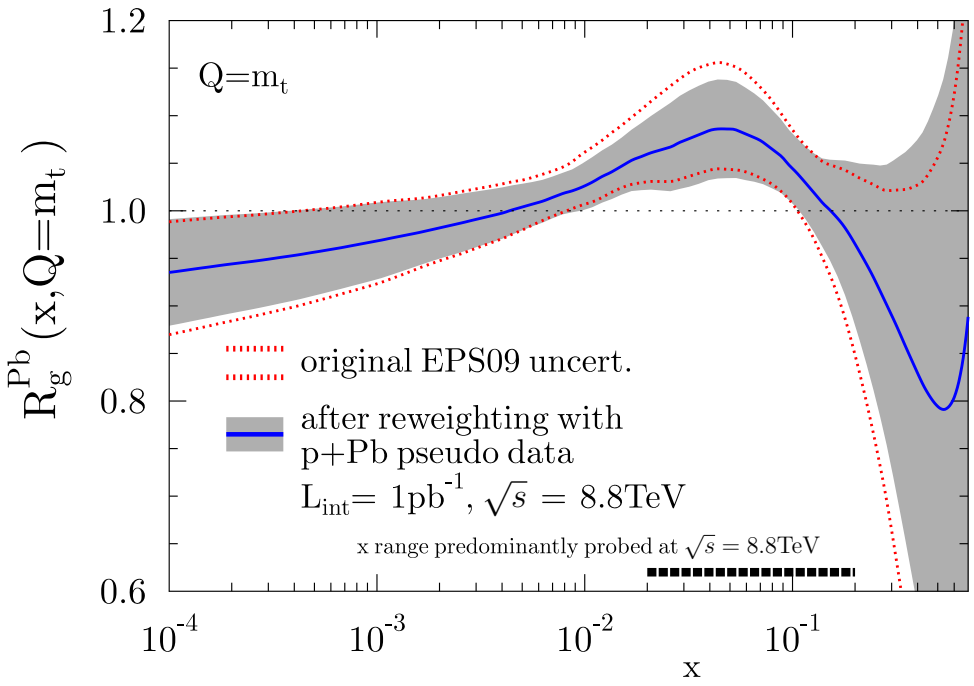


Figure 2.2: Ratio of nuclear-over-proton gluon densities, R_g^{Pb} evaluated at $Q = m_t$, for the p+Pb at $\sqrt{s_{NN}} = 8.8$ TeV at LHC for original EPS09 uncertainty (band enclosed by red dotted lines) and for the reweighted EPS09 using pseudodata (blue curve with grey band) [16].

Another interesting fact is that the collided system is boosted in the proton direction, since the energy of the proton 6.5 TeV is roughly two a half larger than energy of the proton/neutron in lead ion 2.56 TeV. Boost factor $\vec{\beta}$ given by the equation

$$|\vec{\beta}| = \left| \frac{\vec{p}_p + \vec{p}_{\mathrm{Pb}}}{E_p + E_{\mathrm{Pb}}} \right| = \frac{6.5 - 2.56}{6.5 + 2.56} \doteq 0.43 \tag{2.3}$$

shifts the rapidity distribution of outcomming objects by

$$y' = y \pm \operatorname{arctanh}(\beta) = y \pm \operatorname{arctanh}(0.43) \doteq y \pm 0.465. \tag{2.4}$$

The $\pm$ sign is given, because two beam configurations in the Monte Carlo (MC) simulations were considered – p+Pb and Pb+p. The effect is shown in Figure 2.3, where means of the lepton (electrons and muons) pseudorapidity distributions are shifted by the value $\beta = \pm 0.465$ in case of the $t\bar{t}$ signal MC (red and blue lines) with respect to the ATLAS laboratory frame. However, in data (black line) the shift is more flat comapre to signal MC and tends more to the Pb+p beam configuration, because about twice more luminosity in Pb+p (107.79 nb^{-1}) configuration than in p+Pb (56.76 nb^{-1}) were collected and also due to effects of other background contributions. Data were collected in a low-pileup environment, characterized by an average of $\mu_{\mathrm{PU}} = 0.18$ hadronic interactions per bunch crossing.

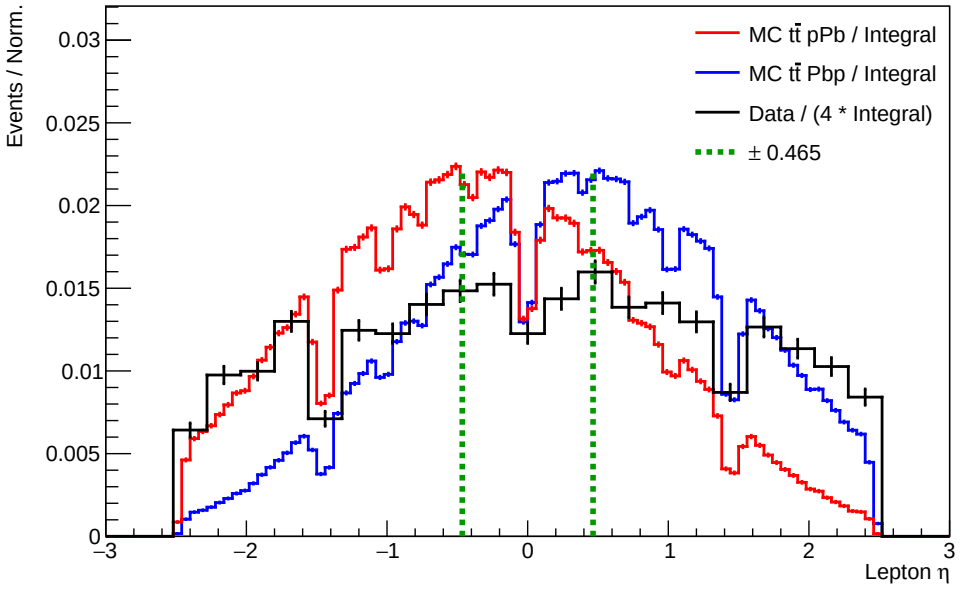


Figure 2.3: The lepton (electrons and muons) pseudorapidity distributions of data (black line) and signal $t\bar{t}$ in p+Pb (red line) and Pb+p (blue line) beam configurations. The rapidity shift of center-mass-system by value $y = \pm 0.465$ (green dashed lines) compare to the ATLAS laboratory system. Distributions are normalized, and on top of that data are scaled by value 0.25 due to the four times coarser binning compare to MC.

2.3.1 Monte Carlo Simulation

Samples were generated at a center-of-mass energy of $\sqrt{s} = 8.16$ TeV with setups used in analyses at $\sqrt{s} = 13$ TeV. To account for isospin effects, two configurations—proton-proton (pp) and proton-neutron (pn) collisions were simulated.

The events for signal and background processes in individual isospin combinations were overlaid with real-data events collected during the 2016 p+Pb run, producing "data overlay" samples. This embedding technique is widely adopted in ATLAS heavy-ion measurements involving hard scattering processes, enabling accurate descriptions of the underlying event in p+Pb collisions. However, samples overlaid by data became challenge to process with Athena software and several modifications and development of the sotfware had to be introduced.

All simulated samples were normalized using the most accurate theoretical cross-sections and k-factor corrections. The MC events were processed through the full ATLAS detector simulation framework, based on Geant4. Given the negligible pileup in the 2016 p+Pb data, no additional pileup reweighting was applied. Furthermore, all samples underwent the same reconstruction algorithms and analysis chain as real data. Minor corrections were applied to the simulated lepton trigger and reconstruction efficiencies, derived from comparisons between data and simulation at $\sqrt{s} = 8.16$ TeV, to ensure better alignment with the actual detector response.

The normalization of MC samples involved scaling the pp samples by $Z_{\mathrm{Pb}} = 82$ and the pn samples by $N_{\mathrm{Pb}} = 126$. Additionally, the p+Pb and Pb+p beam configurations were scaled according to the integrated luminosity of the respective ATLAS data-taking campaigns.

MC samples for $t\bar{t}$ signal

The nominal signal $t\bar{t}$ MC samples at NLO were generated using the PowHeg Box-v2 [19, 8]. The matrix element (ME) generator uses NNPDF3.0 [20] pdf with Pythia8 generator that simulates parton shower, fragmentation and the underlying event. The hard process factorisation scale μ_f and renormalisation scale μ_r were set to the default value: $\mu = (m_t^2 + p_{\mathrm{T,top}}^2)^{1/2}$, where $m_t = 172.5$ GeV and $p_{\mathrm{T,top}}$ are the top quark mass and transverse momentum, respectively, evaluated for the underlying Born configuration. The PowHeg model resummation damping parameter $h_{\mathrm{damp}} = 1.5 m_t^2$ controls matrix element to parton shower matching and effectively regulates the high-p_{T} radiation. The A14 tune [21] with NNPDF2.3LO pdf set was applied for Pythia8 showering. Samples in are generated in two isospin configurations for pp/pn interactions and two beam configurations p+Pb and Pb+p , see Table 2.1.

The nominal PowHeg $t\bar{t}$ MC samples are normalised to the total inclusive $t\bar{t}$ cross-section at a centre-of-mass energy of $\sqrt{s} = 8.16$ TeV calculated at NNLO in QCD including resummation of the next-to-next-to-leading logarithmic (NNLL) soft-gluon terms with the Top++ (v2.0) program [9]. It predicts the total cross-section $\sigma_{t\bar{t}}^{\mathrm{NNLO}} = 272.6^{+18.3}_{-18.9}$ pb with total uncertainties indicated and corresponding to a relative precision of $^{+6.9}_{-6.7}\%$ [22]. The pdf and α_S uncertainties were calculated following the PDF4LHC2015 [23] prescription with the MMHT2014 [24] and CT14 NNLO [25] pdf sets. The first uncertainty is calculated from the envelope of predictions with the QCD renormalisation and factorisation scales varied independently up or down by a factor of two from their default values of $\mu_r = \mu_f = m_t$, whilst never letting them differ by more than a factor of two. The second uncertainty is from variations in the PDF following the PDF4LHC2015 prescription using NNPDF3.0, MMHT, CT14, PDF4LHC pdf variations with nominal scale variation.

nominal	Conf.	N. Evt	ME Gen.	PS Model	ME PDF	Tune
$t\bar{t}$	Pbp/pn	3888871	Powheg	Pythia8	NNPDF3.0 NLO	MMHT 2014 LO, CT14, PDF4LHC
$t\bar{t}$	pPb/np	2069897				
$t\bar{t}$	Pbp/pp	2496722				
$t\bar{t}$	pPb/pp	1376346				

Table 2.1: PowHeg +Pythia8 $t\bar{t}$ production with PowHeg h_{damp} parameter equal to $1.5 m_t$, A14 tune, at least one lepton filter used for the nominal samples.

Alternative $t\bar{t}$ samples were generated to evaluate the modeling uncertainties associated with this process, as detailed in Table 2.2. These samples explore the effects of initial- and final-state radiation (ISR/FSR) and the impact of missing higher-order QCD corrections. This is achieved using an alternative PowHeg-v2 + Pythia8 sample, where the h_{damp} parameter is set to $3 m_t$. Additionally, in the nominal samples, independent variations of the renormalisation and factorisation scales, as well as the effective strong coupling constant that controls ISR and FSR, are performed.

To compare with a different parton shower (PS) model, an alternative $t\bar{t}$ sample is produced using the same PowHeg Box configuration as the nominal sample, but with parton showering, hadronisation, underlying event (UE), and multiple parton interactions (MPI) simulated by Herwig7 [26]. Systematic uncertainties arising from the use of different hard-scattering generators, while keeping the PS model consistent, are assessed by generating a sample with Madgraph5_aMC@NLO [27], interfaced with Pythia8 using the A14 tune.

Alt.	Conf.	N. Evt	ME Gen.	PS Model	ME PDF	Tune
$3m_t$	pPb/np Pbp/pn pPb/pp Pbp/pp	1782708 3879231 1386583 2387300	Powheg	Pythia8	NNPDF 3.0 NLO	A14 NNPDF 2.3 LO
$t\bar{t}$	pPb(Pbp)/pp pPb(Pbp)/pp	3139759 6207050	aMC @NLO	Pythia8		
$1.5m_t$	pPb(Pbp)/pp pPb(Pbp)/pp	3169553 6266799	Powheg	Herwig7		H7.2-Default MMHT2014lo68cl

Table 2.2: Alternative MC samples for the $t\bar{t}$ process used in systematic studies. At the time of production, the Herwig7 and aMC@NLO generators did not support the generation of pn isospin combinations.

V+jets background samples

Events containing W or Z bosons ($V = W/Z$) with associated jets, $V+ \geq 1b$ (BFilter), $V+ \geq 1c$ (CFilterBVeto), and $V+l$ (light-flavour jets, CvetoBVeto), were simulated using the Sherpa 2.2.10 [28] generator (see Feynman diagrams in Figure 2.4). The matrix elements are calculated for up to two jets at NLO and four jets at LO, utilizing the Comix [29] and OpenLoop [30] generators. These are then merged with the Sherpa parton shower [31] following the ME+PS@NLO prescription [32], using the NNPDF 3.0 NNLO PDF set and a dedicated parton shower developed by the Sherpa authors. The W/Z+jets events are normalized to the NNLO cross sections [33].

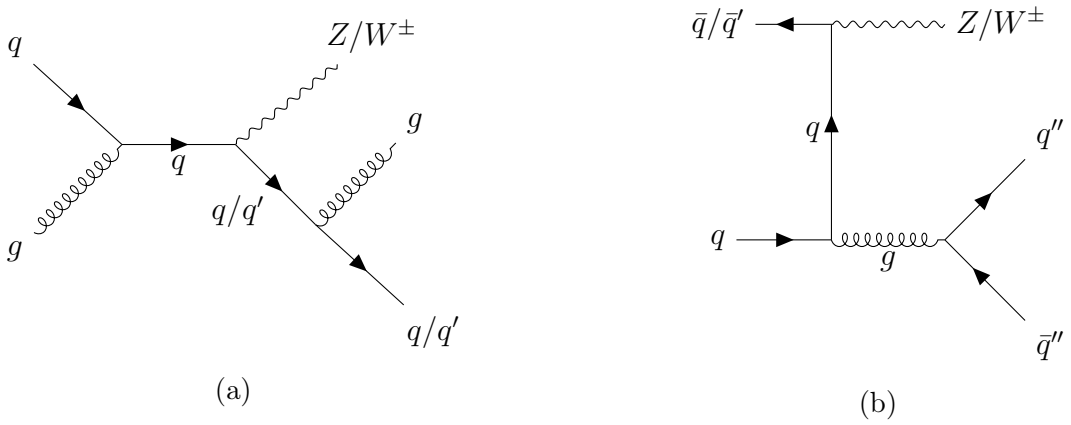


Figure 2.4: Examples of Feynman diagrams of V+jets background. Generated using the TikZ-Feynman package [3].

Signle top background samples

The single top quark production (see leading-order Feynman diagrams in Figure 2.5) is included as the $t\bar{t}$ background, because single top events can have the same detector signature of top quark pair production. Specifically, the Wt-channel at next-to-leading order (see in Figure 2.6) can produce a final state with a W boson and jets that is hard to distinguish from $t\bar{t}$ production. Although interference between Wt and $t\bar{t}$ production exists, current simulations do not typically account for it due to the complexity of modeling. Instead, the Wt-channel is generated separately, called diagram removal method, where matrix element is treated as

$$|M^{\text{removal}}_{Wt-\text{channel}}|^2 = |M_{Wt-\text{single top}}|^2 + \cancel{|M_{t\bar{t}}|^2} + \cancel{2\text{Re}(M_{Wt-\text{single top}}\, M^*_{t\bar{t}})}, \tag{2.5}$$

while as a systematic uncertainty the digram subtraction method is used by accounting also for an interference term

$$|M^{\text{subtraction}}_{Wt-\text{channel}}|^2 = |M_{Wt-\text{single top}}|^2 + \cancel{|M_{t\bar{t}}|^2} + 2\text{Re}(M_{Wt-\text{single top}}\, M^*_{t\bar{t}}). \tag{2.6}$$

The t-channel is another important background for $t\bar{t}$ production as it produces a top quark alongside light jets. On the other hand, the s-channel single top process has a relatively small cross-section and its contribution is negligible compared to the tW and t-channel processes. Therefore, the s-channel is excluded from simulations.

Single top and anti-top contributions in the tW-channel and t-channel were generated using the PowHeg-v2 matrix element (ME) generator [19, 8], interfaced with Pythia8 and the A14 tune. For these processes, the top quarks were decayed using MadSpin [34], ensuring that all spin correlations were preserved. The ME generator employed the NNPDF 3.0 NLO PDF set and was connected to Pythia8 [35], which handled the parton shower, fragmentation, and underlying event. The decays of bottom and charm hadrons were simulated using the EvtGen v1.2.0 program [36].

For the electroweak t-channel single top-quark events, the 4-flavour scheme was applied with the fixed four-flavour PDF4LHC PDF set for NLO matrix element calculations.

A total of eight samples were produced across two beam configurations (p+Pb and Pb+p) and two isospin combinations (pp and pn); see Table 2.3.

Process	Conf.	NEvt	ME Gen.	PS Model	ME PDF	Tune
Single top tW	pPb/pp	49539	Powheg	Pythia8	NNPDF 3.0 NLO	MMHT, CT14, PDF4LHC
Single top tW	Pbp/pp	99491				
Single (anti)top tW	pPb/pp	49539				
Single (anti)top tW	Pbp/pp	98496				
Single (anti)top tW	pPb/np	49539				
Single top tW	pPb/np	49539				
Single (anti)top tW	Pbp/np	99491				
Single top tW	Pbp/np	99491				
Single (anti)top t-chan.	pPb/np	49539	Powheg	Pythia8	NNPDF 3.0 NLO	MMHT, CT14, PDF4LHC
Single top t-chan.	pPb/np	49539				
Single (anti)top t-chan.	Pbp/np	99491				
Single top t-chan.	Pbp/np	99491				
Single (anti)top t-chan.	pPb/pp	49539				
Single (anti)top t-chan.	Pbp/pp	99491				
Single top t-chan.	pPb/pp	49539				
Single top t-chan.	Pbp/pp	99491				

Table 2.3: PowHeg + Pythia8 + EvtGen MC simulation of single top and anti-top quark processes in the tW and t-channels, used for generating the single top-quark samples.

VV+jets background samples

Diboson production can also lead to final states with the same objects as in the case of the $t\bar{t}$ production, see Figure 2.7. VV+jets background samples were simulated using Sherpa 2.2.11, with matrix elements calculated at NLO accuracy in QCD for up to one additional jet and at LO accuracy for up to three additional jets summarized in Table 2.4. The CT14 PDF set was chosen, along with dedicated parton shower tuning developed by the Sherpa authors.

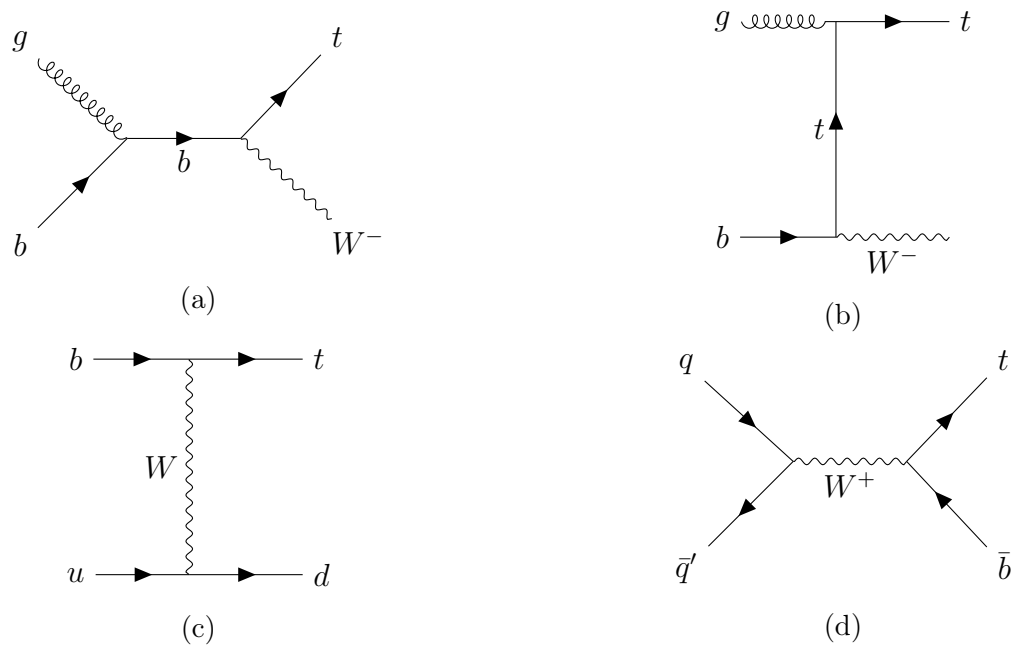


Figure 2.5: Leading order Feynman diagrams of single top background, (a)-(b) Wt-channel, (c) t-channel, and (d) s-channel. Generated using the TikZ-Feynman package [3].

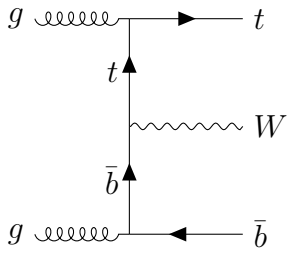


Figure 2.6: Example of next-to-leading order Feynman diagram of single top background Wt-channel. Generated using the TikZ-Feynman package [3].

Process	Conf.	NEvt	ME Gen.	PS Model	Tune
$VV \to \ell\ell\nu\nu$	Pbp/pp	49539	Sherpa-2.2.11	Sherpa-2.2.11	CT14
$VV \to \ell\ell\nu\nu$	pPb/pp	49539			
$VV \to \ell\ell\nu\nu$	Pbp/pn	49539			
$VV \to \ell\ell\nu\nu$	pPb/np	49539			

Table 2.4: V+jet samples for a dilepton channel with two charged leptons and two neutrinos.

Samples for performance studies

Table 2.5 provides a summary of $Z \to e^+e^-$, $Z \to \mu^+\mu^-$ MC samples simulated using PowHegwith showering Pythia8 using tune AZNLO CTEQ6L1 and Photos PDF sets. These samples are futher used for estimating lepton scale factors using tag-and-probe method.

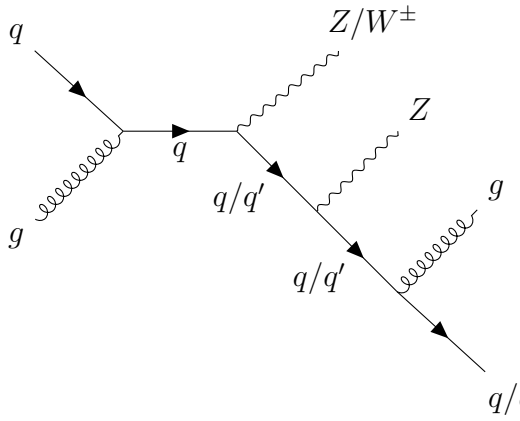


Figure 2.7: Example of Feynman diagram of VV+jets background. Generated using the TikZ-Feynman package [3].

Process	Conf.	NEvt	ME Gen.	PS Model	Tune
$Z \to e^+e^-$	pPb/pp	434533	Powheg	Pythia8	AZNLO CTEQ6L1
$Z \to e^+e^-$	Pbp/pp	804721			
$Z \to e^+e^-$	Pbp/pn	1233406			
$Z \to e^+e^-$	pPb/np	663363			
$Z \to \mu^+\mu^-$	pPb/pp	336597	Powheg	Pythia8	AZNLO CTEQ6L1
$Z \to \mu^+\mu^-$	Pbp/pp	626680			
$Z \to \mu^+\mu^-$	Pbp/pn	953827			
$Z \to \mu^+\mu^-$	pPb/np	514786			

Table 2.5: PowHeg +Pythia8 generated events of Z decays to leptons with overlay of data events.

2.4 Object Selection

In this section, detector object definitions with event pre-selection based on these objects are described in detail.

2.4.1 Electrons

Electron candidates are reconstructed from energy deposits in the central region of the electromagnetic calorimeter (EMC) and are associated with tracks reconstructed from the Inner Detector (ID). The selection of electron candidates is based on requirements:

- Candidates have to pass Medium likelihood-based working point [37] and have a transverse momentum $p_{\mathrm{T}} > 18$ GeV, absolute pseudorapidity $|\eta| < 2.47$, and a transverse energy $E_{\mathrm{T}} > 18$ GeV. The transition region of the EMC (crack region), defined as $1.37 < |\eta| < 1.52$, is excluded from this selection.
- Electron candidates must originate from the primary vertex, which imposes requirements on the transverse impact parameter significance: $|d_0|/\sigma_{d_0} < 5$, and the longitudinal impact parameter: $|\Delta z_0 \sin\theta| < 0.5$ mm.
- The Gradient isolation working point [37] is applied to suppress background contributions from QCD jets misidentified as prompt leptons (fake leptons) and from semileptonic decays of heavy-flavour hadrons (non-prompt real leptons).

In the EM calorimeter, the isolation variable is defined as the sum of the transverse energy of topo-clusters [37] within a cone of radius $\Delta R = 0.2$ around the electron cluster direction, divided by the electron p_{T}. In the ID, the isolation variable is defined as the sum of the p_{T} of tracks within a cone of variable size around the electron track, divided by the electron p_{T} ($\Delta R = \min[0.2, 10/p_{\mathrm{T}}]$).

2.4.2 Muons

Muon candidates were required to have a reconstructed track in the muon spectrometer (MS) combined with a track in the ID called "combined muon".

Muon candidates are selected with the following requirements:

- Candidates have to pass Medium ID criteria [38], with a transverse momentum of $p_{\mathrm{T}} > 18$ GeV, absolute pseudorapidity and $|\eta| < 2.5$.
- Muons are required to be associated to the primary vertex with the transverse impact parameter significance $|d_0|/\sigma_{d_0} < 3$, and on the longitudinal impact parameter of $|\Delta z_0 \sin\theta| < 0.5$ mm.
- To reduce the background of muons originating from heavy-flavour decays inside jets, muon candidates are required to be isolated using the track-quality and isolation criteria similar to those applied to electrons except the maximum cone size used in the ID isolation, which is 0.3 [38].

Lepton tracks have to pass criteria described in [39].

2.4.3 Jets

Jets are built using the anti-k_t algorithm [40] with a radius of $R = 0.4$, implemented through the FastJet package [41]. The algorithm utilizes a four-momentum recombination scheme and is applied using two different approaches.

PFlow jets

In the first approach, topo-clusters [37] serve as the inputs for jet reconstruction. These jets are calibrated to the hadronic energy scale, with p_{T} and η dependent correction factors derived from simulation. These jets are labelled as PF jets.

Heavy Ion jets

The second approach for jet reconstruction uses massless calorimeter towers [42] with a size of $\Delta\eta \times \Delta\phi = 0.1 \times \pi/32$ to construct jets, referred to as HI jets. To suppress the background from the underlying event (UE), a background subtraction is applied to each tower to reduce the rate of fake jets.

The average transverse energy density is estimated iteratively, with regions containing jets excluded from the estimation process. A detailed explanation of this procedure can be found in [42]. These jets are labelled as HI jets.

Two types of calibrations are applied to HI jets:

- EtaJES MC-based calibration: Corrects to the truth energy scale in Monte Carlo simulations as a function of jet η.
- HI/EMTopo cross-calibration and in-situ/η-intercalibration: Accounts for differences between Monte Carlo simulations and data.

The EtaJES correction has larger impact of the two corrections, and is derived using dijet events simulated by PYTHIA 8. The second calibration is applied exclusively to data to address discrepancies between simulation and real data.

The in-situ and η-intercalibrations provide residual corrections to the jet energy scale. The in-situ correction relies on the p_{T} balance between vector boson Z and jet events, while the η-intercalibration uses dijet events, where jets at mid-rapidity are used to calibrate jets in the forward rapidity region.

b-tagging

An important selection criteria for the analysis is the identification of jets containing b-quarks. PF jets containing B-hadrons are tagged using the Deep Learning DL1r algorithm [43].

The algorithm is trained on $t\bar{t}$ MC events, assigning b-jets as signal and a mixture of light-flavour and c-jets as background. Including c-jets in the training enhances c-jet rejection with minimal impact on light-jet rejection.

Jets are classified as b-tagged if the DL1r weight exceeds a fixed working point cut value of 1.27 corresponding to 85.43% b-tagging efficiency for b-jets in $t\bar{t}$ events. The corresponding rejection factors for charm quarks and light jets are 0.346 and 0.0248, respectively [44].

2.4.4 Missing Transverse Energy

Missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ (or MET) is a measure of the momentum carried by escaping neutrinos. It is calculated as

$$E_T^{\mathrm{miss}} = \left| -\sum_i \vec{p}_{T,i} \right| \tag{2.7}$$

where $\vec{p}_{T,i}$ represents the transverse momentum of each reconstructed and calibrated physics object (e.g. electrons, muons, photons, hadronically decaying τ-leptons, small-R jets) within $|\eta| < 4.9$. A soft term is also included, comprising tracks associated with the primary vertex that are not matched to these objects.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ is rebuilt in all events. The algorithm uses information from objects present in the event e.g. medium electrons and medium muons. Also PF jets are used in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. They are also used to derive systematic uncertainties associated with MET. Unfortunatelly, for technical reasons there is no MET calculation available for HI jets. Due to this limitation, the nominal analysis does not use selection on MET. However, the MET requirement is used to define a dedicated control region for fake-lepton background evaluation.

2.4.5 Jet Matching

Since the HI jet collection is not calibrated for b-tagging, both PF and HI jet collections must be used together. The matching process between HI jets and PF jets is based on two criteria. First, events are matched using a run and an event number criterion. Then, ΔR matching is applied to pair jets within the same event.

The b-tagging information is extracted from the PF jets and assigned to matched HI jets. Kinematic variables in the analysis are derived from HI jets, with the candidates required to have JES-corrected $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$. The same kinematic cuts are applied to PF jets.

The ΔR matching is performed by considering all possible combinations of HI jets and PF jets in an event. The distance between each pair of jets in pseudorapidity and azimuthal angle is calculated using the ΔR metric.

Jet pairs are then sorted based on the smallest ΔR, and the algorithm loops through the sorted pairs, matching jets and removing their indices. This results in $N_{\mathrm{jets}} = \min(N_{\mathrm{HIjets}}, N_{\mathrm{PFjets}})$

matched jets. If a matched pair has $\Delta R < 0.3$ (chosen based on Figure 2.8) and the PFlow jet is b-tagged, the corresponding HI jet is also labelled as b-tagged.

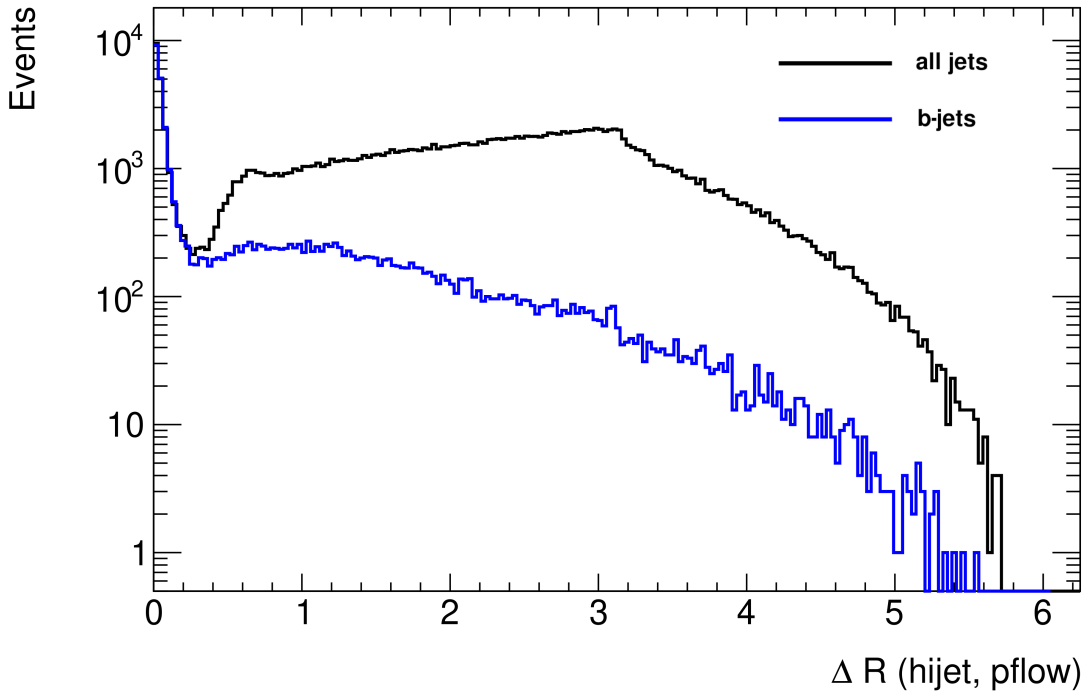


Figure 2.8: ΔR between HI and PF jets in signal MC samples. The black line represents all jets, while the blue line corresponds to the jets that have a b-tagged PF jet counterpart.

Feature	**PF jets**	**HI jets**
b-tagging	b-tag info available	b-tag info not available
MET calculation	MET available	No MET available
Calibration	High-pileup (pp, 13 TeV)	Low-pileup (p+Pb, 8.16 TeV)
Jet energy scale	Affected by underlying events (UE)	Corrected by UE subtraction

Table 2.6: Comparison between PF and HI jets.

The jet matching algorithm combines the strengths of both PF and HI jet collections, but it introduces additional systematic uncertainties.

Varying ΔR

Two matching systematic uncertainties are evaluated by shifting $\Delta R = 0.3$ criteria for matched jets by ± 0.1. The relative differences in yields of MC (background + signal) are within 2.5%.

HI jets missing PF counterpart

In cases when number of HI jets is larger than PF jets, shown by bins above a diagonal in Figure 2.9, the HI jet is tagged as non-b-tagged in nominal samples. However, in principle such jet could have been b-tagged. The other extreme would be to assing all HI jets which are missing PF counterpart as b-tagged. More realistic approach is to assign HI jets which

are missing PF counterpart randomly as b-jets based on light-flavour jet rejection function as a function of jet p_{T}.

Tables 2.7–2.8 show an important cross-check of how often HI jets are missing its PF counterpart with respect to studied regions and consistency between data and MC. The effect on $H_{\mathrm{T}}^{\ell,j}$ distribution in one signal region ≥2b does not exceed relative difference of 10% to the nominal sample.

b-jet Region	Data	MC (background + signal)
≥ 0 b-jets	17.97% (±0.30)	14.89% (±0.29)
$= 1$ b-jet	14.80% (±0.58)	12.10% (±0.55)
≥ 2 b-jets	8.70% (±0.57)	6.35% (±0.53)

Table 2.7: Percentage of HI jets missing the PF counterpart caused by jet matching algorithm in data and MC (background + signal) samples for different b-jet regions.

b-jet Region	Data	MC (background + signal)
≥ 0 b-jets	−5.51% (±0.33)	−5.77% (±0.34)
$= 1$ b-jet	−7.30% (±0.58)	−7.32% (±0.57)
≥ 2 b-jets	−4.16% (±0.39)	−4.54% (±0.40)

Table 2.8: The decrease of the number of b-tagged jets caused by jet matching algorithm in data and MC (background + signal) samples for different b-jet regions.

2.4.6 Overlap Removal

After selecting objects described in previous sections (electrons, muons, and jets), the following points are applied to ensure the objects do not overlap:

- If a selected electron shares a track with a selected muon, the electron is removed.
- If an HI jet is within $\Delta R < 0.2$ of a reconstructed electron, the HI jet is removed.
- If an electron is within $\Delta R < 0.4$ of an HI jet, the electron is removed (it helps to reduce the impact of non-prompt leptons).
- If an HI jet has fewer than three tracks and is within $\Delta R < 0.4$ of a muon, the HI jet is removed. Conversely, the muon is removed if it is within $\Delta R < 0.4$ of an HI jet with three or more tracks.

2.5 Event Selection

The dilepton channel is defined by events containing exactly two opposite-sign leptons. Events with same-flavour lepton pairs (e^+e^- or $\mu^+\mu^-$) with an invariant mass ($m_{\ell\ell}$) within the Z boson mass window (80 GeV $< m_{\ell\ell} <$ 100 GeV) are excluded to suppress the Z boson background.

The invariant mass $m_{\ell\ell}$ has to exceed 15 GeV in the $e\mu$ channel and 45 GeV in the ee and $\mu\mu$ channels. This requirement ensures compatibility with the phase space of the Z+jets simulation samples while having minimal impact on the final results.

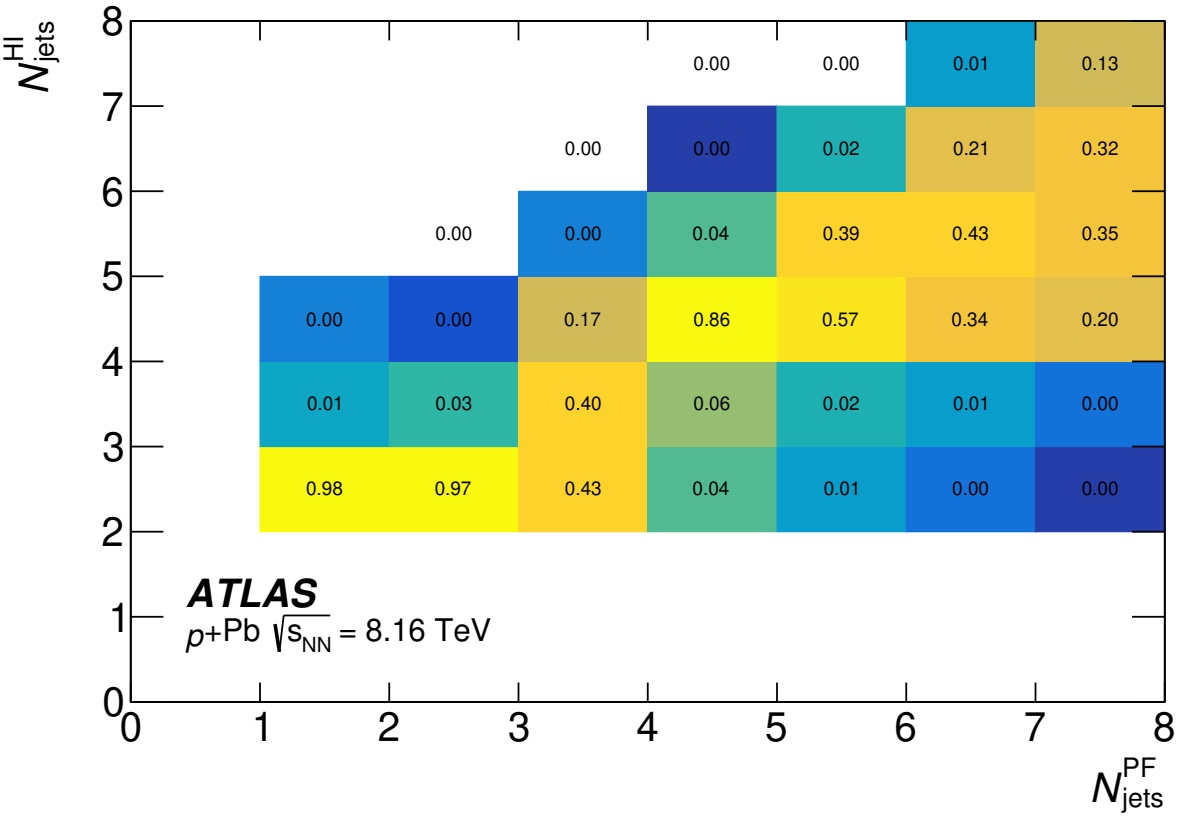


Figure 2.9: Number of particle flow (PF) jets versus number of heavy ion (HI) jets using jet matching algorithm normalised by columns in MC signal + background samples. Edge bins include overflow events [17].

To define the signal region (SR) of the dilepton channel, events must also contain at least two HI jets, including at least one b-tagged jet. The signal region is further categorized based on the number of b-tagged jets: events with exactly one b-tagged jet are labelled as $2\ell 1b$, while those with two or more b-tagged jets are labelled as $2\ell 2b$incl.

In the ℓ+jets channel, events are selected if they contain exactly one lepton and at least four HI jets, including at least one b-tagged jet, to define the signal region. This signal region is subdivided based on the lepton flavour and the number of b-tagged jets: one electron or muon with exactly one b-tagged jet ($1\ell 1b$), and one electron or muon with two or more b-tagged jets ($1\ell 2b$incl).

Table 2.9 summarizes the six signal regions and three control regions (CR) which have the same selection as SR except the number of b-tagged jets is equal to zero.

2.6 Muon Performance

The goal of muon performance studies is to derive scale factors based on recommendations and tools given by Muon Combined Performance group and summarized in the publication [38]. The settings remained default as for the pp analysis except that no pile-up correction was introduced. The challage was to modify recommended software to the MC samples overlaid with data characteristic for proton-lead collisions. Three types of scale factors were derived

- reconstruction/identification,
- isolation,

Channel	SR Region Label	Selection Criteria
Dilepton	$2\ell 1b$	2 OS leptons, 1 b-tagged jet
Dilepton	$2\ell 2b$incl	2 OS leptons, ≥ 2 b-tagged jets
ℓ+jets (electron)	$1\ell 1b$ e+jets	1 electron, 1 b-tagged jet, ≥ 4 jets
ℓ+jets (electron)	$1\ell 2b$incl e+jets	1 electron, ≥ 2 b-tagged jets, ≥ 4 jets
ℓ+jets (muon)	$1\ell 1b$ μ+jets	1 muon, 1 b-tagged jet, ≥ 4 jets
ℓ+jets (muon)	$1\ell 2b$incl μ+jets	1 muon, ≥ 2 b-tagged jets, ≥ 4 jets
	CR Region Label	
Dilepton	$2\ell 0b$	2 OS leptons, 0 b-tagged jet
ℓ+jets (electron)	$1\ell 0b$ e+jets	1 electron, 0 b-tagged jet, ≥ 4 jets
ℓ+jets (muon)	$1\ell 0b$ μ+jets	1 muon, 0 b-tagged jet, ≥ 4 jets

Table 2.9: Summary of signal (SR) and control (CR) regions for the dilepton and ℓ+jets channels based on the number of leptons and b-tagged jets. The OS stands for opposite charged leptons.

- trigger,

joining reconstruction and identification efficiencies into one efficiency. The total scale factor is

$$\begin{aligned} \mathrm{SF}_{\mathrm{total}} &= \frac{\varepsilon_{\mathrm{total,\ data}}}{\varepsilon_{\mathrm{total,\ MC}}} = \frac{\varepsilon_{\mathrm{reco/id,\ data}} \times \varepsilon_{\mathrm{iso,\ data}} \times \varepsilon_{\mathrm{trig,\ data}}}{\varepsilon_{\mathrm{reco/id,\ MC}} \times \varepsilon_{\mathrm{iso,\ MC}} \times \varepsilon_{\mathrm{trig,\ MC}}} \\ &= \mathrm{SF}_{\mathrm{reco/id}} \times \mathrm{SF}_{\mathrm{iso}} \times \mathrm{SF}_{\mathrm{trig}}, \end{aligned} \tag{2.8}$$

while $Z \to \mu\mu$ data and MC samples were used. Due to the ratio, possible biases introduced by the measurement method which appear both in data and MC simulation cancel out. The total scale factor quantifies the deviation of the simulation from the real detector behaviour, and is therefore used in physics analysis to correct the simulation used as a weight [38].

2.6.1 Reconstruction/Identification Efficiency

The tag-and-probe method was used to measure the reconstruction and identification efficiency. For the muon reconstruction the information from inner detector (ID), Muon spectrometer (MS), and calorimeters were taken. Based on reconstruction algorithm the five muon types are defined [38]:

- combined muon (CB) - matches MS to ID tracks and performs a combined fit based, the energy loss from calorimeters is taken into account;
- inside-out combined muon (IO) - extrapolates ID to the MS tracks and tries to find at least 3 loosely-aligned MS hits;
- MS extrapolated muon (ME) - defined when MS track cannot be matched to an ID track, MS track parameters are extrapolated to the beamline;
- segment-tagged muon (ST) - ID track extrapolated to the MS with tight angular matching requirements applied to at least one reconstructed MS segment;
- calorimeter-tagged muon (CT) - extrapolates ID tracks to the calorimeters to search for energy deposits consistent with a minimum-ionising particle.

Working points

The ratio of charge and momentum of muons expressed at interaction point called q/p compatibility and is defined as

$$q/p = \frac{|q/p_{\mathrm{ID}} - q/p_{\mathrm{MS}}|}{\sqrt{\sigma^2(q/p_{\mathrm{ID}}) + \sigma^2(q/p_{\mathrm{MS}})}}, \tag{2.9}$$

where $\sigma(q/p)$ are uncertainties of the measurement. Three major muon woring points are defined as

- The *Medium* working point refer to muons CP and IO within $|\eta| < 2.5$ with $q/p < 7$, must have at least two precision stations (parts of MS with precise measurements of the position and angle of a muon's trajectory), except muons within $|\eta| < 0.1$ are required to have 1 precision station. In $t\bar{t}$ events prompt *Medium* muons are by 98% CB muons.
- The *Loose* working point are the muons passing the *Medium* working point with additional CT and ST muons in pseudorapidity $|\eta| < 0.1$, where the gap in the MS causes loss of efficiency for CB muon reconstruction. To increase the efficiency of the *Loose* criteria for low-p_{T} muons, IO muons having $p_{\mathrm{T}} < 7$ GeV and only one precision station are accepted in $|\eta| < 1.3$. Requiring that the purity of IO muons increases.
- The *Tight* working point are the *Medium* muons, but only those CB and IO muons having at least two precision stations. The normalised χ^2 of the combined track fit is required to be less than 8 to reject tracks due to hadron decays. Futher optimisated cuts on the q/p and ρ' ($|p_{\mathrm{T,ID}} - p_{\mathrm{T,MS}}|/p_{\mathrm{T,CB}}$) depending on the p_{T} and $|\eta|$ of the muon are applied to supress backgournd from non-prompt muons [38].

Event selection

The invariant mass of dimuon (tag and probe) candidates has to be within the range of $m_{\mu\mu} \in \{61,121\}$ GeV having opposite charge. The tag has to satisfy *Medium* identification with $p_{\mathrm{T}} > 18$ GeV, $|\eta| < 2.5$, and the single-muon trigger requirements. Additionally, tag has to fulfill *Tight* isolation criteria

$$p_{\mathrm{T}}^{\mathrm{varcone30}} < 0.04 \cdot p_{\mathrm{T}}^{\mu}, \quad E_{\mathrm{T}}^{\mathrm{topoetcone20}} < 0.15 \cdot p_{\mathrm{T}}^{\mu}, \quad p_T^{\mathrm{track}} > 1 \text{ GeV} \tag{2.10}$$

where p_{T}^{μ} stands for transverse momentum of the muon, $p_{\mathrm{T}}^{\mathrm{varcone30}}$ is the sum of p_{T} tracks aound the muon within the cone of

- $\Delta R < \min(10 \text{ GeV}/p_{\mathrm{T}}^{\mu}, 0.3)$, for $p_{\mathrm{T}}^{\mu} <= 50$ GeV,
- $\Delta R < 0.2$, for $p_{\mathrm{T}}^{\mu} > 50$ GeV.

Similarly $E_{\mathrm{T}}^{\mathrm{topoetcone20}}$ is the sum of E_{T} of tracks around the muon within $\Delta R < 0.2$.

Furthermore, the vertex association criteria ensure a maximal purity of tags originating from the hard-scattering proton–proton collision. The criteria includes:

- Transverse impact parameter significance: $|d_0/\sigma(d_0)| < 3$,
- Longitudinal impact parameter: $|z_0| \sin\theta < 0.5\,\mathrm{mm}$.

where d_0 is the transverse impact parameter (shortest distance between the muon track and the primary vertex in the xy-plane), $\sigma(d_0)$ is the uncertainty of d_0, the z_0 longitudinal impact parameter is the shortest distance between the muon track and the primary vertex along the z-axis.

The probe muons have looser selection than tag muons

- $p_{\mathrm{T}} > 18$ GeV
- $|\eta| < 2.5$
- Transverse impact parameter significance: $|d_0/\sigma(d_0)| < 3$,
- Longitudinal impact parameter: $|z_0| \sin\theta < 10\,\mathrm{mm}$.

and probe isolation is more stringent calorimeter-based and looser track-based than the tag *Tight* isolation.

The identification and reconstruction efficiency for *Medium* woring point is defined as

$$\epsilon\,(\mathit{Medium}) = \epsilon\,(\mathrm{ID}) \times \epsilon\,(\mathit{Medium}|\mathrm{ID}) \simeq \epsilon\,(\mathit{Medium}|\mathrm{CT}) \times \epsilon\,(\mathrm{ID}|\mathrm{MS})\,. \tag{2.11}$$

since efficiency $\epsilon\,(\mathrm{ID})$ cannot be measured directly, it is repaced by the conditional efficiency $\epsilon\,(\mathrm{ID}|\mathrm{MS})$ for a muon reconstructed by the MS to be also reconstructed in the ID. To reduce the background contamination, the $\epsilon\,(\mathit{Medium}|\mathrm{ID}) \simeq \epsilon\,(\mathit{Medium}|\mathrm{CT})$ approximation was used, replacing ID probes with the more pure CT probes, and a systematic uncertainty was assigned to cover for the small bias introduced. The validity of Equation 2.11 is guaranteed by the independecy of the track reconstruction in the ID and in the MS and can be defined for other woring points (*Loose* *Tight*, etc.).

The are few options of types of muon probes which test different kind of efficiencies listed below [38]. The probes reconstructed by one detector subsystem test efficiency of the other independent subsystem.

- ID probes are ID tracks used to measure the reconstruction efficiency in the MS, or of specific identification algorithms.
- MS probes are ME tracks used to test the efficiency of the ID reconstruction.
- CT probes are ID tracks also satisfying the calo-tagging reconstruction algorithm described in Section 4 of [38]. In the same way as the ID probes, they are used to measure the reconstruction efficiency in the MS, or of specific muon identification algorithms.
- ST probes are ID tracks also satisfying the segment-tagging reconstruction algorithm described in Section 4 of [38]. In the same way as the ID probes, they are used to measure the reconstruction efficiency in the MS, or of specific muon identification algorithms.
- *Two-track* probes are MS tracks required to be within $\Delta R = 0.05$ of an ID track. They are used to measure the combined reconstruction efficiency of a muon candidate with ID and MS tracks, or the efficiency of specific identification criteria.
- *Loose* probes are muon candidates satisfying the *Loose* identification requirements. They are used to measure the isolation and vertex association efficiencies.

The efficiency of a certain algorithm is measured using a matching requirement of $\Delta R < 0.05$ between the probe and any muon candidate reconstructed and identified with the algorithm of interest. The efficiency is then computed as the number of probes P that are successfully matched to a muon reconstructed and identified according to the X criterion, N_P^X, divided by the total number of selected probes N_P^{All}:

$$\epsilon\,(X|P) = \frac{N_P^X}{N_P^{\mathrm{All}}}. \tag{2.12}$$

Probes are counted in data events after the subtraction of the backgrounds. In simulation, to eliminate the background contamination, both the tag and the probe muons are required to be a prompt muon at generator level [38].

Background subtraction

The background in data are is estimated by fit in the range $m_{\mu\mu} \in \{61,121\}$ GeV separately for the all selected and matched probes. The background contribution (non-prompt muons) is modelled using the form

$$f(m_{\text{tag-probe}}) = \left(1 - \frac{m_{\text{tag-probe}}}{\Lambda}\right)^{p_1} \cdot \left(\frac{m_{\text{tag-probe}}}{\Lambda}\right)^{p_2}, \tag{2.13}$$

where the Λ parameter is approximately the energy necessary to produce the dimuon pair, set to 2.5 times of upper mass range of $m_{\mu\mu}$ ($\Lambda = 2.5 \times 121 = 302.5$ GeV) and the p_1 and p_2 are obtained using a separate fit with a sample of same-charge tag-and-probe pairs, satisfying all the selection criteria except the isolation requirements.

Systematic uncertainties

The list of systematic uncertainties contributing the most to the total uncertainty is listed below [38]

- *T&P method*: Possible biases in the tag-and-probe approach, such as different kinematic distributions between reconstructed probes and generated muons, or ID-MS efficiency correlations, are evaluated using simulation. This is done by comparing the reconstructed efficiency to the fraction of generator-level muons that are successfully reconstructed. Since this bias affects both data and simulation similarly, it cancels out in the scale factor (SF) calculation. Half of the observed difference is assigned as the SF uncertainty to conservatively cover potential imperfections in the simulation. Utilizing *two-track* probes lowers this uncertainty to below approximately 0.1%, which further decreases as p_{T} increases.

- *Probe matching*: the default ΔR-based matching procedure is varied in order to assess an uncertainty in how much a given probe type contributed to a certain type of reconstructed muon candidate. This is done by comparing the nominal fraction of matched probes with the fraction of probe tracks for which muon candidate reconstruction is successful.

- *Template shape*: the uncertainty in the shape of the template modelling the non-prompt muon background is evaluated by simultaneously varying the p_1 and p_2 parameters in Eq. (2.13) by their fit uncertainties. The consequent deviation of the SFs from their nominal value is taken as the systematic uncertainty.

- Λ-*SC*: the numerical value of the Λ parameter in Eq. (2.13) guarantees a well-behaved, smooth function across $m_{\text{tag-probe}}$. Possible effects on the SFs are estimated by varying its value by $\pm 20\%$.

- *Background fit*: to cover effects associated with the fitting procedure used to extract the contribution of the non-prompt muon background, the change in the SFs obtained when varying the fitted non-prompt muon background by its corresponding fit uncertainty is assigned as a systematic uncertainty.

Muon reco/ID scale factors

Reconstruction/identification scale factors are derived in the *Medium* working point. The data efficiency is larger compare to Monte Carlo which leads into scale factor exceeding unity by up to 5% in Fig. 2.10. This result was discussed with muon CP experts and result was cross-checked using independed code.

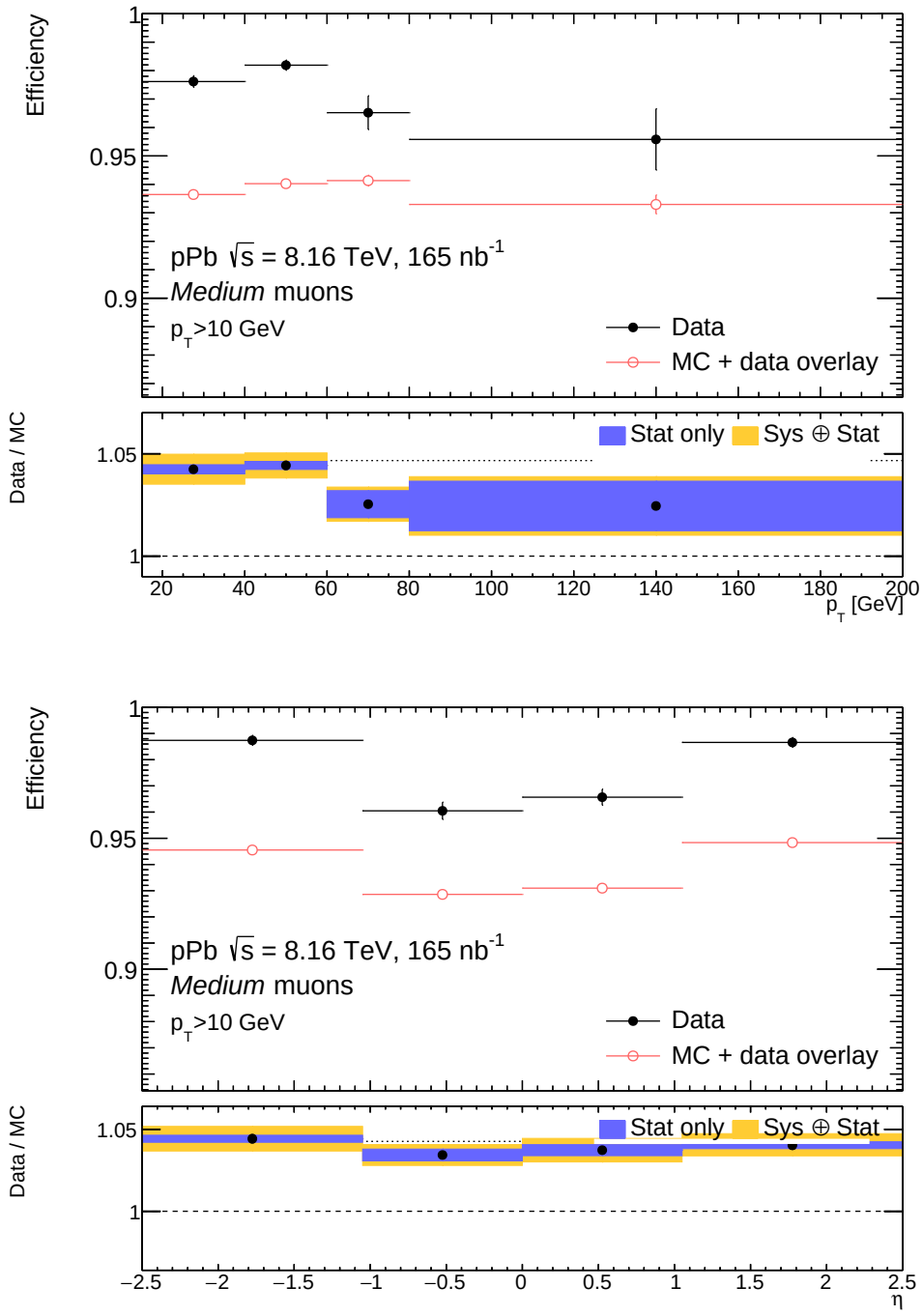


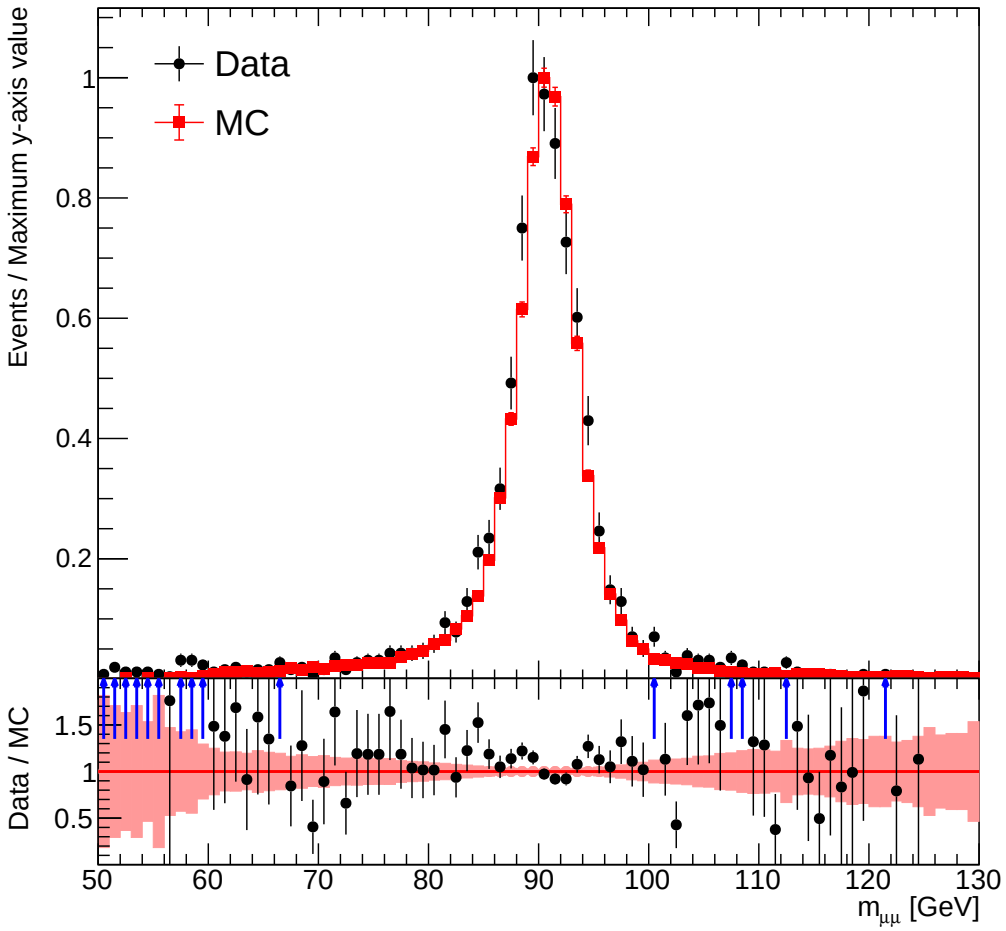
Figure 2.10: The muon reconstruction/identification ϵ(Medium) efficiency as a function of p_T and η.

2.6.2 Control Plots with $Z \to \mu\mu$ events

Validation of muon scale factors has been done with $Z \to \mu^+\mu^-$ events using PowHeg MC simulation. Events with two Medium muons and FCTight_FixedRad isolation have been selected. Two muons had to be of an opposite sign. A requirement on the invariant mass of the dimuon system was imposed to be between 50-130 GeV. No requirement on jets was imposed.

Figure 2.11 shows distributions for the dimuon system invariant mass. A good description

of the data by MC simulation with muon scale factors applied is found with some features that can be attributed to background in data or to muon scale and resolution differences.



(a)

Figure 2.11: Illustration of control distributions for $Z \to \mu^+\mu^-$ events compared with MC simulation (100k events) for the signal process. Only the statistical uncertainties are shown.

2.7 Systematic Uncertainties

Major systematic uncertainties to the analysis of the $t\bar{t}$ production in p+Pb production at $\sqrt{s_{\mathrm{NN}}}$=8.16 TeV arise from the following sources:

- reconstruction of electrons, muons and jets,
- b-tagging,
- fake-lepton background,
- MC signal + background modelling,
- integrated luminostity.

Reconstruction of electrons, muons and jets

Uncertainties in electron reconstruction, identification, isolation, and trigger efficiencies are derived from $Z \to e^+e^-$ events. The uncertainty in the low pile-up energy calibration is evaluated in line with the methodology presented in the publication [45].

For muons, uncertainties in the momentum scale and resolution are based on the results from the publication [38]. Additionally, uncertainties in the muon scale factors covering reconstruction, isolation, and trigger efficiencies are derived from $Z \to \mu^+\mu^-$ events as described in Section 2.6.

Jet-related uncertainties are inferred by in situ studies of the calorimeter response [46], their application to jets in heavy-ion (HI) data [47], and comparisons of the simulated jet response across samples from different generators.

b-tagging

The b-tagging systematic uncertainties are computed by varying the data-to-MC correction factors within their uncertainties [48, 49, 50]. To assess uncertainties in the HI-PF jet matching, two additional systematic variations are introduced in Section 2.4.5 and are negligible with respect of other uncertainties.

The fake-lepton background

Systematic uncertainties in the fake-lepton background arise from statistical and systematic variations of the real and fake lepton efficiencies, and are evaluated using the Matrix Method. Based on distributions in $0b$ control region the conservative normalization uncertainties of 100% in the μ+jets and 50% in the e+jets and the dilepton signal regions are set as uncorrelated uncertainties.

Shape variations of this background in the ℓ+jets channel were evaluated also in the $0b$ CR. To derive them, all background contributions except the fake-lepton events were subtracted from the data. The difference was normalised to the number of fake-lepton events. A ratio is constructed of such subtracted and scaled data to the fake-lepton contribution as a function of the azimuthal angle $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}},\ell)$ between the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ a variable sensitive to fake lepton background. This ratio is fitted by a second-order polynomial. Shape variations of the fake-lepton background in $1b$ and $\geq 2b$ ℓ+jets SRs are defined as up and down fit shape variations using the fit parameter uncertainties. However, the shape of the $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}},\ell)$ variable is not correlated to the shape of the final fit variable $H_{\mathrm{T}}^{\ell,j}$.

MC signal + background modelling

Signal modeling uncertainties are established using alternative $t\bar{t}$ MC samples based on different choice of parton-shower, hadronization models, and matrix-element algorithms of matching to the parton shower. The uncertainty due to initial-state radiation of partons (ISR) is estimated by variations of α_s for ISR in the A14 tune [21]. Further effects on the ISR are evaluated by varying the renormalisation (μ_{r}) and factorisation scales (μ_{f}) in the matrix-element calculation as well as the Powheg h_{damp} parameter. The μ_{r} and μ_{f} are varied independently by factors of 0.5 and 2.0, avoiding the same side variations of the scales. The effect of final-state radiation of partons (FSR) uncertainties is evaluated by modifying the μ_{r} for emissions from the parton shower by factors of 0.5 and 2.0. The PDF uncertainties affecting the $t\bar{t}$ signal are evaluated using the PDF4LHC15 Hessian uncertainties [23].

Additional background systematic uncertainties related to the normalisation of the V+jets samples are determined using the Berends scaling technique [51, 52, 53]. Single-top-quark diagram removal and diagram subtraction variation samples are used to assess the uncertainties from the interference between the $t\bar{t}$ and tW processes [54]. A conservative uncertainty of

9.5% is considered for the normalisation of both the tW and t-channel single-top-quark processes [55]. The diboson background normalisation is allowed to vary by 50% [56].

Integrated luminostity

The uncertainty in the integrated luminosity of the combined data sample is 2.4%. It is derived from the calibration of the luminosity scale using x-y beam-separation scans, following a methodology similar to that detailed in Ref. [57], and using the LUCID-2 detector for the baseline luminosity measurements [58].

2.8 Fit Procedure

Due to the limited statistics, a differential cross-section measurement or unfolding procedure was not feasible. Instead, the fit strategy focused on determining the signal strength, denoted as $\mu_{\text{fit}}^{\text{data}}$, which quantifies the scaling factor required to match the $t\bar{t}$ Monte Carlo (MC) prediction to the observed data.

The fit to the signal strength variable μ was performed simultaneously across the six signal regions (two dilepton and four ℓ+jets regions) described in Section 2.5 (Table 2.9). The chosen observable for the final fit was $H_{\text{T}}^{\ell,j}$ (see pre-fit plots in each region in Figure 2.12), defined as the scalar sum of the transverse momenta of all jets and selected isolated leptons in each event. This variable offers sensitivity to the kinematic properties of the $t\bar{t}$ events. The combined data yield from all channels amounted to 3024 events. Figure 2.13 provides the number of events in each channel, where the background composition and signal purity vary.

Pre-fit plots display the predicted distributions of observables ($H_{\text{T}}^{\ell,j}$, p_{T}, η, etc.) based on the simulations before any fit to the data. The predictions use nominal values of the model parameters without adjustments to match data. Post-fit plots display the same distributions after the model parameters have been fitted to match the data. This is done through likelihood minimization. The fit refines the parameters to better describe the observed data. Post-fit uncertainties are usually smaller compare to pre-fit because the data constrains the model parameters, reducing the overall uncertainty. The fitting process narrows the range of possible values for parameters.

The total $t\bar{t}$ production cross-section was then computed using the relation

$$\sigma_{t\bar{t}}^{\text{measured}} = \frac{\mu_{\text{fit}}^{\text{data}}}{\text{BR}_{\text{non-allhad}}} \sigma_{\text{MC}}^{\text{t}\bar{\text{t}},\text{non-allhad}}, \tag{2.14}$$

where $\mu_{\text{fit}}^{\text{data}}$ represents the signal strength obtained from the fit, $\sigma_{\text{MC}}^{\text{t}\bar{\text{t}},\text{non-allhad}}$ is the theoretical cross-section (including k-factors) used to normalize the $t\bar{t}$ MC samples in the non-allhadronic decay modes, and $\text{BR}_{\text{non-allhad}}$ is the branching ratio of $t\bar{t}$ to non-allhadronic final states.

The likelihood function used in the fit is

$$\begin{aligned} L(\vec{n},\, \vec{\theta_0} \mid \mu,\, \vec{\theta}) \;=\; & \prod_{i}^{N_{\text{bins}}} \frac{(\mu \cdot s_i(\vec{\theta}) + b_i(\vec{\theta}))^{n_i} e^{-(\mu \cdot s_i(\vec{\theta}) + b_i(\vec{\theta}))}}{n_i!} \cdot \\ & \prod_{\theta_k \in \vec{\theta}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\theta_k - \theta_k^0)^2}{2\sigma^2}\right). \end{aligned} \tag{2.15}$$

In this expression, n_i is the observed number of events in bin i, $s_i(\vec{\theta})$ and $b_i(\vec{\theta})$ are the signal and background expectations, respectively (in bin i), and $\vec{\theta}$ represents the set of nuisance parameters (systematic uncertainties in case of the measurement with their central values $\vec{\theta_0}$).

The second term in the product accounting for the Gaussian constraints on the nuisance parameters, ensuring the simultaneous treatment of systematic uncertainties. However, for positively defined parameters, such as luminosity and cross-section, a log-normal prior is used to maintain positivity. The signal strength μ represents the physics free parameter of interest. The best estimate for μ and $\vec{\theta}$ is obtained by maximizing the likelihood 2.15 using the TRExFitter package [59].

In total, 140 systematic uncertainties were considered in the analysis. For one particular uncertainty, symmetrization was applied as follows:

$$\mathrm{UP}_{\mathrm{symmetric}} = \mathrm{NOMINAL} + \frac{\mathrm{UP} - \mathrm{DOWN}}{2} \tag{2.16}$$

$$\mathrm{DOWN}_{\mathrm{symmetric}} = \mathrm{NOMINAL} - \frac{\mathrm{UP} - \mathrm{DOWN}}{2}. \tag{2.17}$$

Additionally, one-sided UP and DOWN variations for uncertainties (e.g. single-top diagram removal/subtraction, $t\bar{t}$ modelling, and JES) were symmetrized using the formula:

$$\mathrm{UP}_{\mathrm{symmetric}} = \mathrm{NOMINAL} + \mathrm{UNC} \tag{2.18}$$

$$\mathrm{DOWN}_{\mathrm{symmetric}} = \mathrm{NOMINAL} - \mathrm{UNC} \tag{2.19}$$

To reduce computational time during the fit, a pruning technique was applied, which removes systematic variations that have an impact below 0.5% from the nominal value. Using this technique, approximately 25% of the systematic uncertainties were pruned.

During the analysis, the value of signal strength μ was blinded using the TRExFitter setup to ensure that the analyzers remained unbiased until all performance studies and the complete analysis chain were approved by a dedicated group of experts within the ATLAS Collaboration.

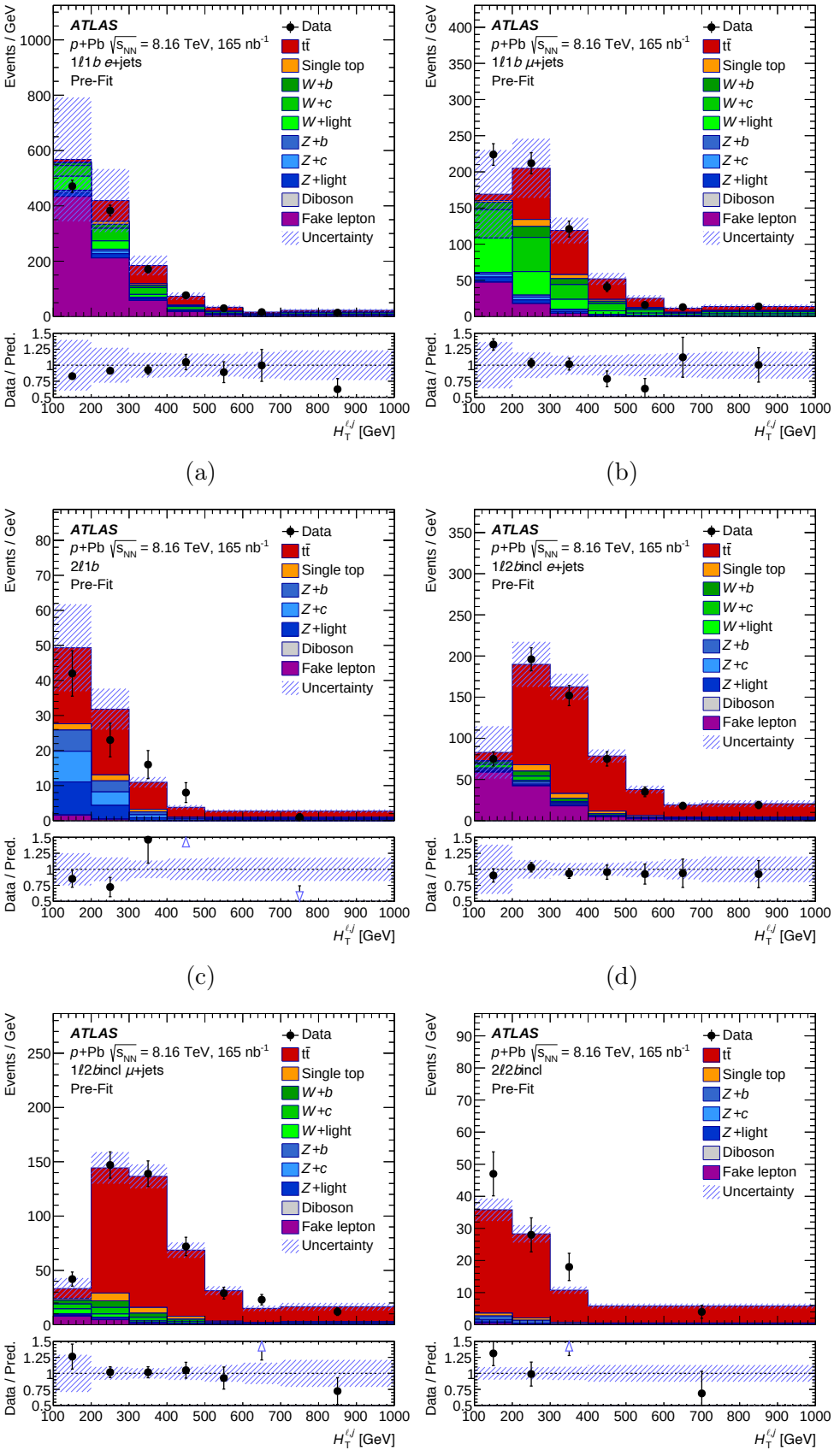


Figure 2.12: Comparison of data and total pre-fit prediction for the $H_{\mathrm{T}}^{\ell,j}$ distribution in each of the six SRs (e+jets: (a) $1\ell 1b$ and (d) $1\ell 2b$incl, μ+jets: (b) $1\ell 1b$ and (e) $1\ell 2b$incl, dilepton: (c) $2\ell 1b$ and (f) $2\ell 2b$incl), with total uncertainties in the prediction represented by the hatched area. The last bin includes overflow events [17].

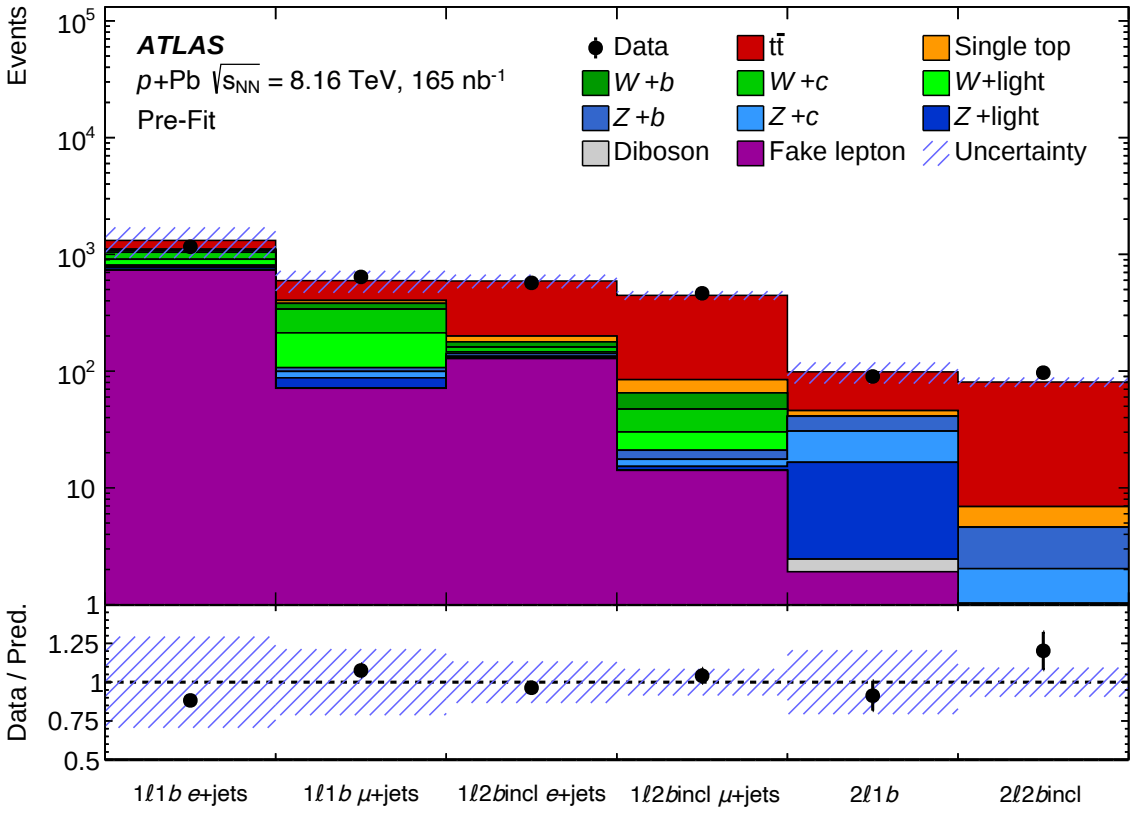


Figure 2.13: Data and predicted pre-fit event yields in each of the six signal regions. The background contributions, before fit, are depicted as filled histograms. The size of the combined statistical and systematic uncertainty in the sum of the signal and background is indicated by the blue hatched band. The ratio of the data to the sum of the signal and background is presented in the lower panel [17].

2.9 Results

The TRExFitter package not only provides the fitted value of the signal strength μ, but also offers convenient options configurable via a configuration file. For this measurement it generates key pre-fit and post-fit plots for each systematic uncertainty, a correlation matrix of systematic uncertainties, and a ranking plot. The best fit value of the signal strength $\mu_{\text{fit}}^{\text{data}}$ is find to be

$$\mu_{\text{fit}}^{\text{data}} = 1.04^{+0.034}_{-0.035}\,(\text{stat.})^{+0.088}_{-0.087}\,(\text{syst.}) = 1.04^{+0.094}_{-0.094}. \tag{2.20}$$

A ranking plot in Figure 2.14 is used to assess the impact of different systematic uncertainties on the signal strength μ in the fit. It helps to identify which systematic uncertainties have the largest effect on the result. To construct a ranking plot a fit is performed to determine the best estimate of the signal strength μ with all systematic uncertainties included. Each systematic uncertainty is individually shifted up and down by its fitted uncertainty, while all other uncertainties are kept free. This assesses how much the μ changes due to this individual shift. Pre-fit impact shows how much each systematic uncertainty would affect μ without any constraint from the data while the post-fit impact displays the effect after the fit, where correlations between systematic uncertainties and the data's ability to constrain those uncertainties are accounted for. The post-fit impact is typically smaller as the fit constrains the nuisance parameters and therefore results in smaller uncertainties also for μ, a technique known as profiling. Systematics are ranked based on their contribution to the uncertainty in μ. The most impactful uncertainties appear at the top of the plot, and bars or markers indicate the magnitude of their influence. By comparing pre-fit and post-fit impacts, one can determine which systematic uncertainties are well-constrained by the fit and which remain dominant.

According to the ranking plot in Figure 2.14, the dominant systematic uncertainty is the fake-lepton background in the μ + jets 1b and e + jets 1b channels, which was conservatively set to 100% and 50%, respectively, based on the $E_{\text{T}}^{\text{miss}}$ control plot in Figure 2.15a. Therefore, corresponding nuisance parameters are constrained by the fit. Although the $E_{\text{T}}^{\text{miss}}$ variable is derived from PF jets, which were not used for kinematic measurements (HI jets were used for kinematics), the aim was to account for the observed disagreement. The fake-lepton background appears to be underestimated in the μ + jets 1b region and overestimated in the e + jets 1b region, therefore the corresponding nuisance parameters are shifted in opposite direction from each other in the ranking plot.

Since $E_{\text{T}}^{\text{miss}}$ is used as an inverted cut in the derivation of fake efficiencies, there is a possibility of non-orthogonality in the selection. As a result, an additional shape and normalization uncertainty was introduced to cover this potential issue. It is not surprising that these systematic uncertainties are the leading ones, given their strong dependence on background modeling and control region discrepancies. The fitted $H_{\text{T}}^{\ell,j}$ distributions are not strongly dependent of the $E_{\text{T}}^{\text{miss}}$ variable, except through the fake background estimation, which was one of the key reasons for using it as the primary fit distribution.

The effect of systematic uncertainties, sorted into groups based on their origin, is provided in Table 2.10, where jet energy scale, signal modeling, and fake background uncertainties are dominant.

Figure 2.16 shows the post-fit distributions of $H_{\text{T}}^{\ell,j}$ (yields given in Figure 2.17), which have smaller uncertainties compared to the pre-fit plots in Figure 2.12. Figure 2.18 shows correlations between systematic uncertainties, where e.g. the fake background shows the strongest correlation between e+jets and μ+jets channels in 1b region, see pre-fit plots in Figures 2.12a–2.12b and at the same time anti-correlates with W+light jets background systematic uncertainties as these compete against each other in the fit.

The total cross-section is calculated using Equation 2.14 as follows

Source	$\Delta\sigma_{t\bar{t}}/\sigma_{t\bar{t}}$	
	unc. up [%]	unc. down [%]
Jet energy scale	+4.6	-4.1
$t\bar{t}$ generator	+4.5	-4.0
Fake-lepton background	+3.1	-2.8
Background	+3.1	-2.6
Luminosity	+2.8	-2.5
Muon uncertainties	+2.3	-2.0
W+jets	+2.2	-2.0
b-tagging	+2.1	-1.9
Electron uncertainties	+1.8	-1.5
MC statistical uncertainties	+1.1	-1.0
Jet energy resolution	+0.4	-0.4
$t\bar{t}$ PDF	+0.1	-0.1
Systematic uncertainty	+8.3	-7.6

Table 2.10: Summary of the impact of the systematic uncertainties on the $t\bar{t}$ cross-section measurement grouped into different categories. The quoted uncertainties are obtained by repeating the fit with a group of nuisance parameters fixed to their fitted values and subtracting in quadrature the resulting total uncertainty from the uncertainty of the complete fit. However, the total uncertainty is not the quadratic sum of the grouped impacts, as this approach neglects the correlation among the different groups. [55]

$$\begin{aligned}
\sigma_{t\bar{t}}^{\text{measured}} &= \frac{\mu_{\text{fit}}^{\text{data}}}{\text{BR}_{\text{non-allhad}}}\,\sigma_{\text{MC}}^{t\bar{t},\text{non-allhad}} \\
&= \frac{1.04168}{0.543} \times 127.8212 \times 1.139 \times 208 \\
&= 58093.09 \text{ pb} \\
&= 58.1 \pm 2.0 \text{ (stat.)} \,^{+4.8}_{-4.4} \text{ (syst.) nb}
\end{aligned} \tag{2.21}$$

Here, the NLO non-all-hadronic $pp \to t\bar{t}$ cross-section $\sigma_{\text{NLO}} = 127.8212\,\text{pb}$, evaluated at $\sqrt{s_{\text{NN}}} = 8.16\,\text{TeV}$, with a k-factor of 1.139, is scaled to the p+Pb system by the number of nucleons $A = 208$. The non-all-hadronic branching $\text{BR}_{\text{non-allhad}}$ ratio is taken as $1 - \text{BR}_{\text{allhad}} = 1 - 0.457$ [60], accounting to 0.543. The relative total uncertainty of the measured cross-section is 9%.

Figure 2.19a presents a comparison between the observed $\sigma_{t\bar{t}}$ and the CMS measurement in p+Pb collisions at $\sqrt{s_{\text{NN}}} = 8.16$ TeV [61]. The two results are consistent within 1.4 standard deviations. Additionally, the figure includes the most precise measurement of the $t\bar{t}$ production cross-section in pp collisions at $\sqrt{s} = 8$ TeV from the ATLAS and CMS combination [62]. This value is extrapolated to the centre-of-mass energy used in this study using the Top++ v2 prediction, and scaled by A_{Pb} to account for the p+Pb system. The extrapolated cross-section carries a 2.5% relative uncertainty and is independent of any nPDF assumptions. The extrapolation factor is calculated to be 1.0528 ± 0.0005 (PDF) $^{+0.0001}_{-0.0013}$ (scale). Furthermore, the measured cross-section is compared with NLO calculations generated by MCFM [63], scaled to NNLO precision in QCD using the same k-factor, obtained from the Top++ v2 generator. Four different nPDF sets were used as input to the MCFM calculations: EPPS21 [64], nCTEQ15HQ [15, 65], nNNPDF30 [66, 67], and TUJU21 [68] [55].

2.9.1 Nuclear Modification Factor

A nuclear modification factor as a single number is defined as the ratio of the cross-section for the $t\bar{t}$ production in proton-lead p+Pb collisions to the cross-section for the $t\bar{t}$ production in proton-proton pp collisions, scaled by the atomic mass number A_{Pb} of the lead nucleus.

$$\begin{aligned} R_{p\mathrm{A}} &= \frac{\sigma_{t\bar{t}}^{p+\mathrm{Pb}}}{A_{\mathrm{Pb}} \cdot \sigma_{t\bar{t}}^{pp}} \\ &= \frac{58093.09}{208 \cdot 243.3 \cdot 1.0528} \\ &= 1.090 \ \pm \ 0.039 \ (\mathrm{stat.}) \ ^{+0.094}_{-0.087} \ (\mathrm{syst.}), \end{aligned} \tag{2.22}$$

where the $\sigma_{t\bar{t}}^{p+\mathrm{Pb}}$ is the measured value of the $t\bar{t}$ cross-section in p+Pb collisions at $\sqrt{s_{\mathrm{NN}}} =$ 8.16 TeV and $\sigma_{t\bar{t}}^{pp}$ is the measured $t\bar{t}$ cross-section in pp collisions at $\sqrt{s} = 8$ TeV [62] extrapolated to the same centre-of-mass energy $\sqrt{s} = 8.16$ TeV by factor 1.0528 and scaled by number of nucleons in lead (208). A comparison between the measured $R_{p\mathrm{A}}$ from data and simulation is shown in Figure 2.19b, where R_{pA} was calculated at NNLO precision using the MCFM code [63] scaled to the p+Pb system for four different nPDF sets. The uncertainty related to the baseline PDF for pp interactions is treated as fully correlated across the predictions and, as a result, cancels out in the ratio. The remaining uncertainty reflects the contribution from the nPDF. All nPDF calculations yield R_{pA} values greater than unity.

2.9.2 Observation in the Dilepton Channel

In Figure 2.20 the signal strength μ measured in each of the six fitted regions individually is shown while the last row represents the combined fitted value of μ. The fitted values of μ in the individual channels are consistent with each other within uncertainties and align with the Standard Model prediction. The precision of the μ measurement is primarily limited by systematic uncertainties in the ℓ+jets signal regions (SRs), while statistical uncertainties dominate in the dilepton SRs.

The significance is determined by performing separate fits of μ in the combined four ℓ+jets SRs and the combined two dilepton SRs, with both exceeding five standard deviations. This confirms the observation of the $t\bar{t}$ production in both the ℓ+jets and dilepton channels separately, with the latter being observed for the first time in p+Pb collisions at the LHC.

2.9.3 Differences to the CMS Measurement

The CMS Collaboration released publication of the measurement of the $t\bar{t}$ production at 8.16 TeV in the ℓ+jets channel [61] in 2017 using a fit in $m_{jj'}$ distribution. Table 2.11 summarizes main differences in analysis conditions, where the ATLAS experiment gains in statistics using lower cuts on lepton and jet transverse momenta, larger phase-space in pseudorapidity η, and by using the dilepton channel. The CMS recorded slightly higher luminostity, however, due to wider selection cuts and other effects, the signal yield in ℓ+jets channel is about 67% higher at ATLAS analysis compared to CMS. Table 2.12 presents a comparison of uncertainties between the ATLAS and CMS measurements, showing that the ATLAS experiment achieved overall twice the precision of CMS.

Analysis conditions	ATLAS	CMS
Channel	lepton+jets, dilepton	lepton+jets
Fitted Distribution	$H_{\mathrm{T}}^{\ell,j}$	$m_{jj'}$
Min lepton p_T [GeV]	18	30
Lepton $\|\eta\|$	< 2.4 (2.47)	< 2.1
Min jet p_T [GeV]	20	25
Integrated luminosity [nb^{-1}]	164.6	174

Table 2.11: Comparison of analysis conditions between ATLAS and CMS of the $t\bar{t}$ production in proton-lead collisions at 8.16 TeV.

Uncertainties [%]	ATLAS	CMS
Integrated luminosity	2.4	5
Statistical	3	8
B-tagging + JES	5	13
Extra JES	-	4
Background	4	7
Lepton trigger and reconstruction	3	4
Total uncertainty	9	18

Table 2.12: Comparison of uncertainties between the ATLAS and CMS experiments in the $t\bar{t}$ production cross-section in proton-lead collisions at 8.16 TeV.

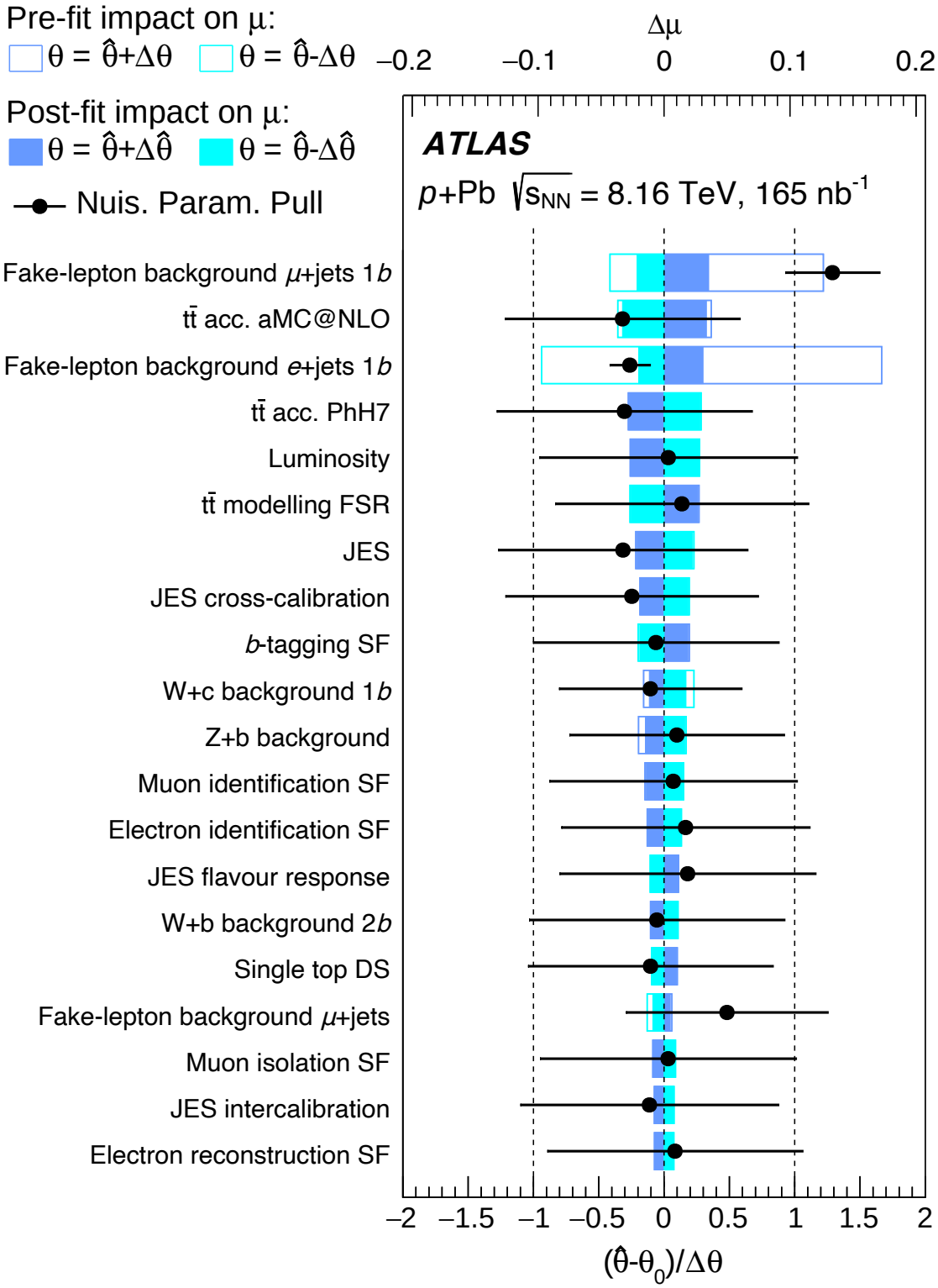


Figure 2.14: The impact of systematic uncertainties on the fitted signal-strength parameter $\hat{\mu}$ for the combined fit of all channels is shown. Filled blue boxes represent the variations of $\hat{\mu}$ from the central value, $\Delta\hat{\mu}$ (top x-axis), when fixing the corresponding individual nuisance parameter, θ, to its post-fit value $\hat{\theta}$, modified upwards or downwards by its post-fit uncertainty, and repeating the fit. Empty blue boxes show the corresponding pre-fit impact. Black points represent the fitted values and uncertainties of the nuisance parameters relative to their pre-fit values, θ_0, and uncertainties, $\Delta\theta$ (bottom x-axis). Black lines represent the post-fit uncertainties of the nuisance parameters relative to their nominal uncertainties, as indicated by the dashed line [17].

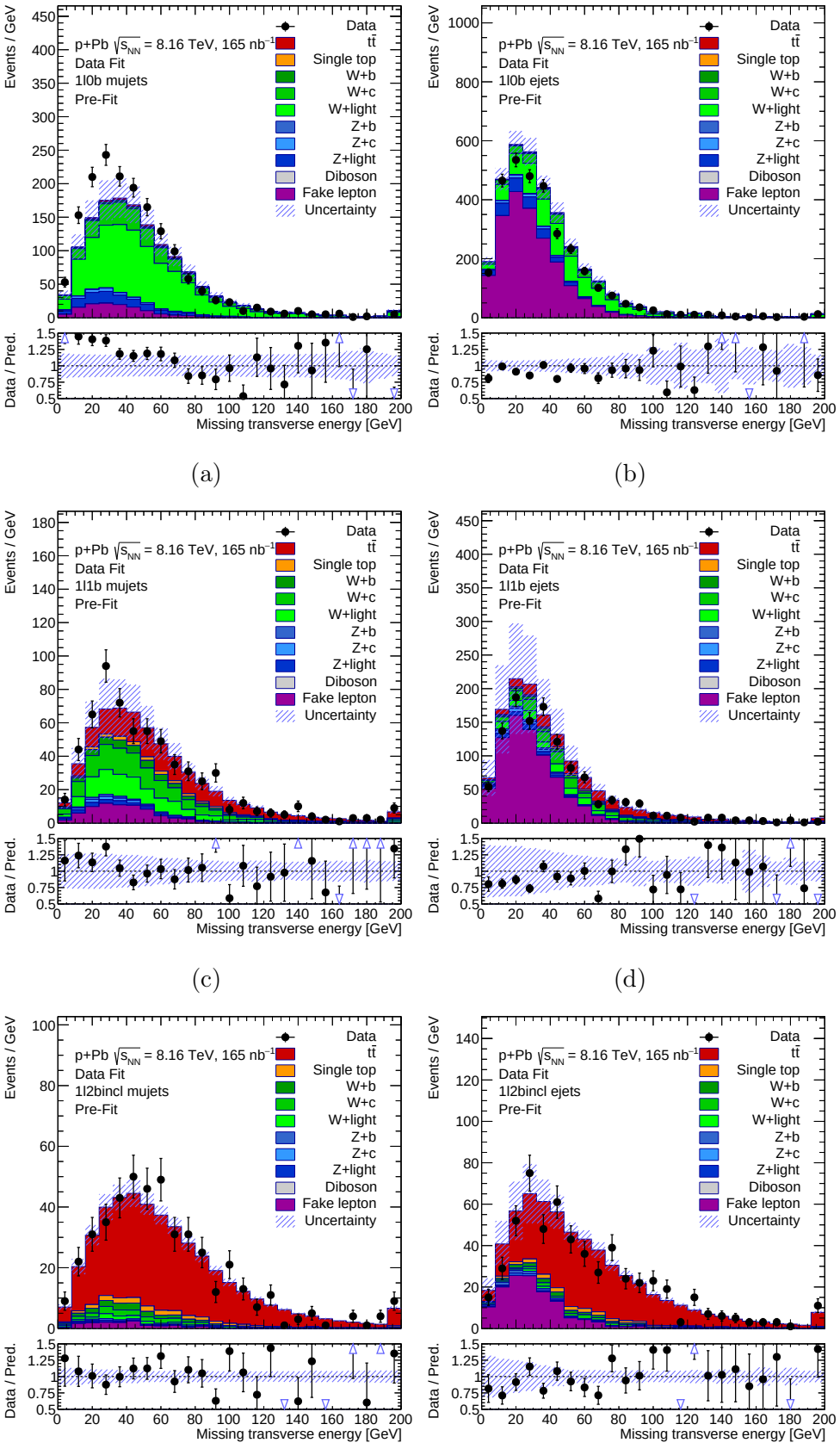


Figure 2.15: Pre-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in ℓ+jets events (left column muon channe, right column electron channel): (a)–(b) $1\ell 0b$, (c)–d $1\ell 1b$, (e)–f $1\ell 2b$incl. The blue hatched area represents the total uncertainties. The last bin includes overflow events.

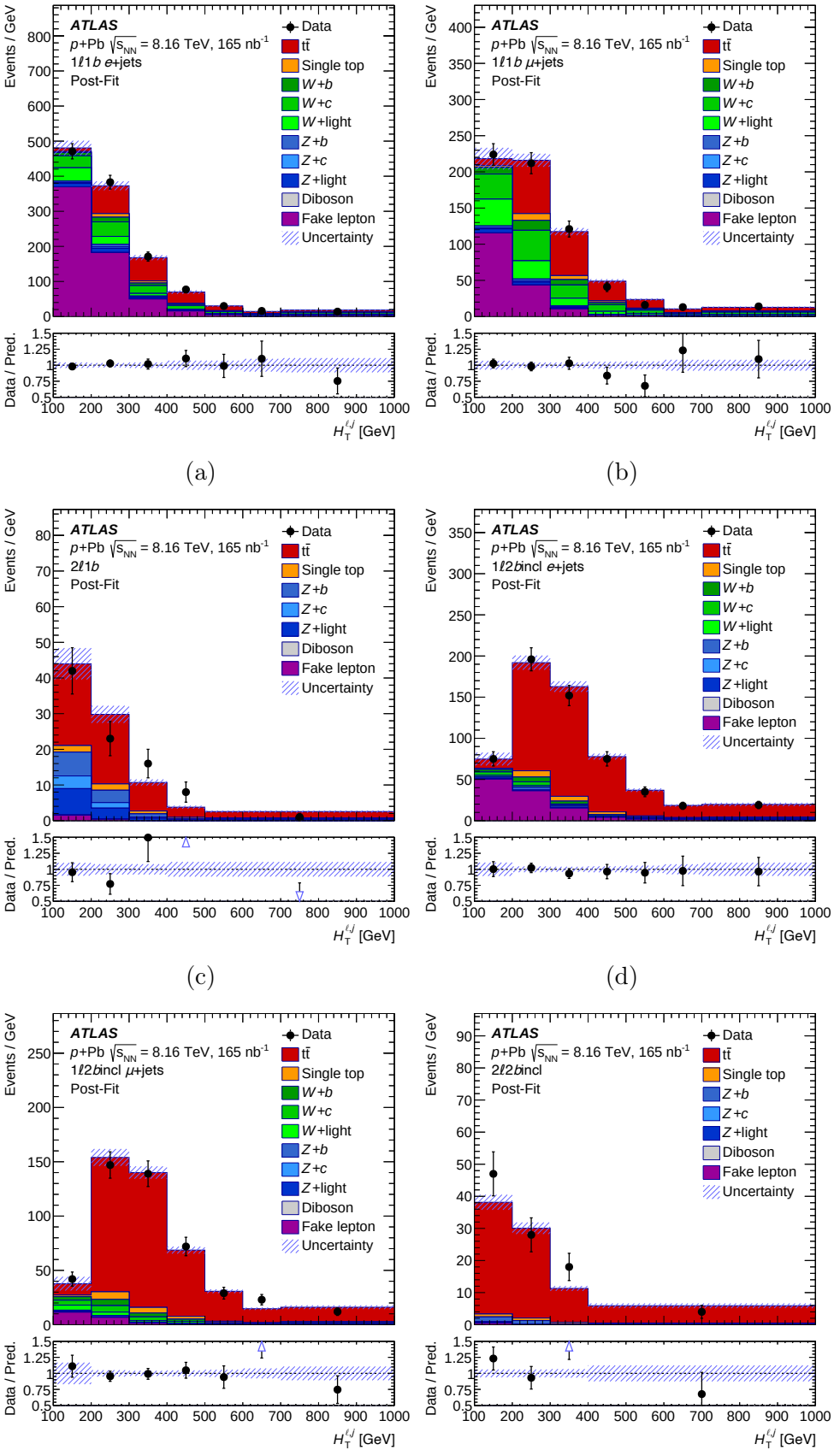


Figure 2.16: Comparison of data and total post-fit prediction for the $H_{\mathrm{T}}^{\ell,j}$ distribution in each of the six SRs (e+jets: (a) $1\ell 1b$ and (d) $1\ell 2b$incl, μ+jets: (b) $1\ell 1b$ and (e) $1\ell 2b$incl, dilepton: (c) $2\ell 1b$ and (f) $2\ell 2b$incl), with total uncertainties in the prediction represented by the hatched area. The last bin includes overflow events [17].

	$1\ell 1b$ e+jets	$1\ell 1b$ μ+jets	$1\ell 2b$incl e+jets	$1\ell 2b$incl μ+jets	$2\ell 1b$	$2\ell 2b$incl
$t\bar{t}$	214 ± 24	194 ± 21	405 ± 21	373 ± 19	55 ± 6	79 ± 5
t-channel	6.9 ± 1.0	6.4 ± 1.0	7.7 ± 0.9	7.1 ± 0.9	0 ± 0	0 ± 0
W+b	37 ± 19	37 ± 19	16 ± 8	17 ± 9	–	–
W+c	120 ± 40	110 ± 40	14 ± 7	17 ± 8	–	–
W+light	80 ± 40	80 ± 40	4.8 ± 3.1	9 ± 5	–	–
Z+b	16 ± 13	8 ± 7	8 ± 7	3.7 ± 3.0	12 ± 9	2.9 ± 2.4
Z+c	9 ± 14	5 ± 7	1.7 ± 2.6	0.9 ± 1.4	6 ± 9	0.4 ± 0.6
Z+light	28 ± 16	12 ± 7	1.2 ± 1.1	0.9 ± 0.5	11 ± 6	0.34 ± 0.25
Diboson	0.32 ± 0.16	0.29 ± 0.15	0.055 ± 0.029	0.039 ± 0.02	0.53 ± 0.27	0.049 ± 0.025
tW	17.1 ± 3.0	15.5 ± 2.7	13.6 ± 3.2	12.1 ± 2.9	5.1 ± 2	2.4 ± 1.2
Fake lepton	630 ± 50	170 ± 40	110 ± 19	21 ± 12	1.9 ± 1	0.51 ± 0.27
Total	1154 ± 34	648 ± 24	582 ± 21	462 ± 18	91 ± 7	85 ± 5
Data	1162	641	570	464	90	97

Figure 2.17: Data and predicted post-fit event yields in each of the six signal regions. The total uncertainty is a quadrature sum of statistical and systematic uncertainties. Owing to rounding effects and small correlations between the different sources of uncertainties, the total systematic uncertainty is different from the sum in quadrature of the individual sources [17].

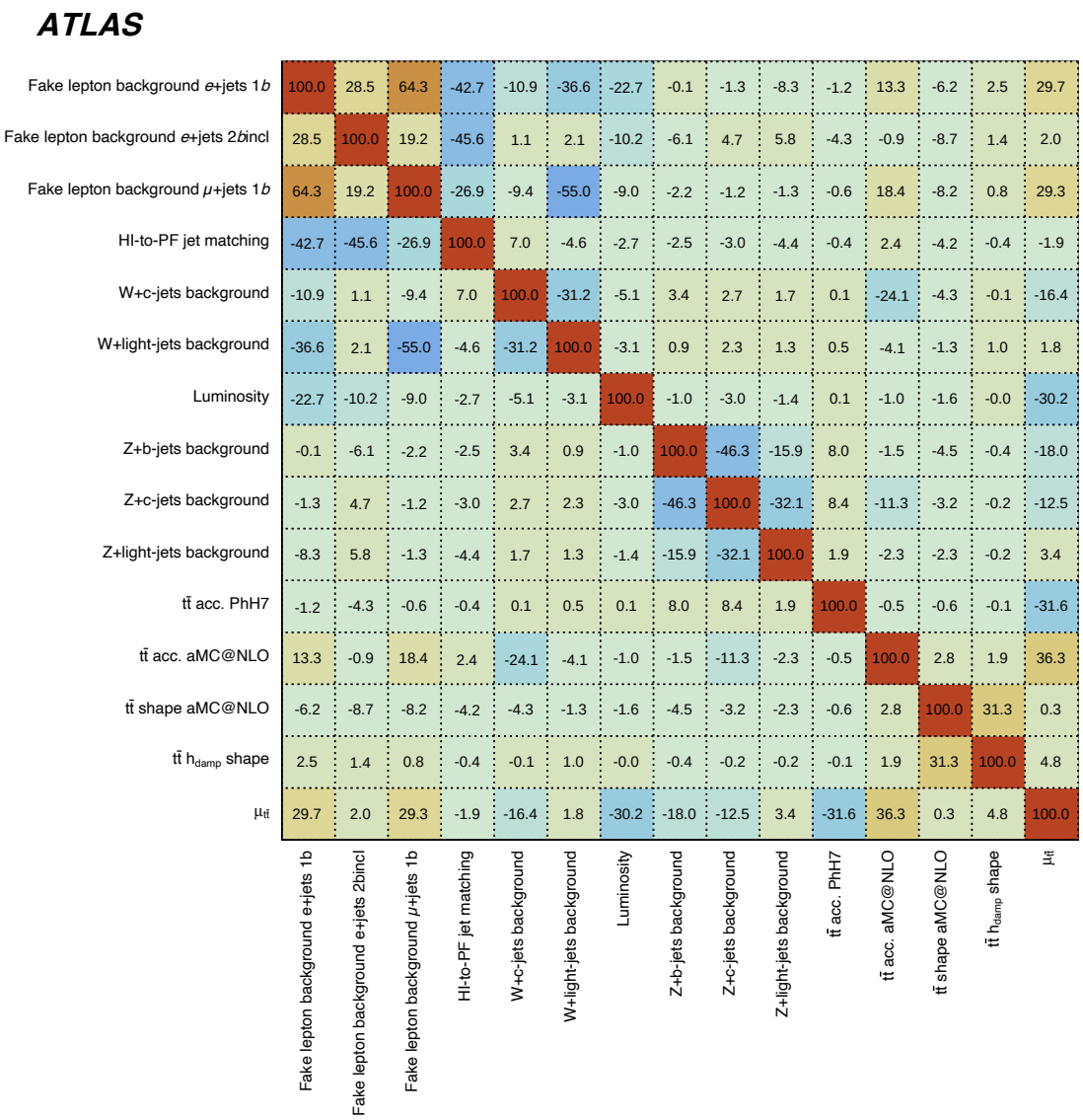
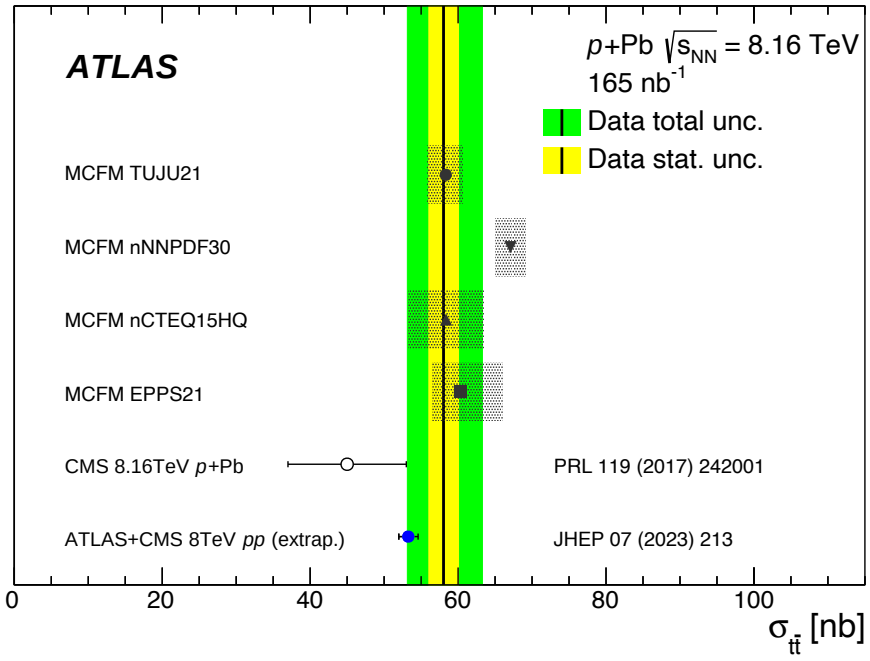
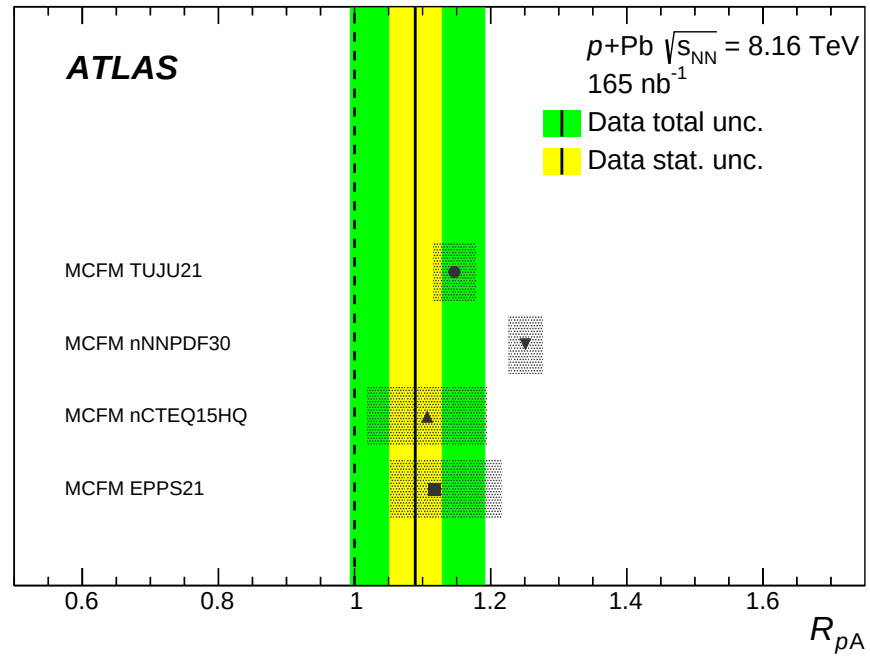


Figure 2.18: Correlation matrix of the fit parameters for the combined fit to data: only parameters with at least one correlation factor greater than 30% are shown [17].



(a)



(b)

Figure 2.19: Comparison between the measured and predicted values of (a) $\sigma_{t\bar{t}}$ and (b) R_{pA}. $\sigma_{t\bar{t}}$ is also compared with the existing measurement in p+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV [61], and the combined measurement of the $t\bar{t}$ production cross-section in pp collisions at $\sqrt{s} = 8$ TeV from ATLAS and CMS collaborations [62]. The latter is extrapolated to the centre-of-mass energy of this measurement and scaled by the factor $A_{\mathrm{Pb}} = 208$. Predictions are calculated at NLO precision using the MCFM code [63] scaled to the p+Pb system and by the k-factor 1.139, given for different nPDF sets. The uncertainty in the predictions represents the internal PDF uncertainty. The solid black line indicates the measured value. The combined statistical and systematic uncertainty of the measurement is represented by the outer band around the central value, while the statistical component is depicted as the inner band [17].

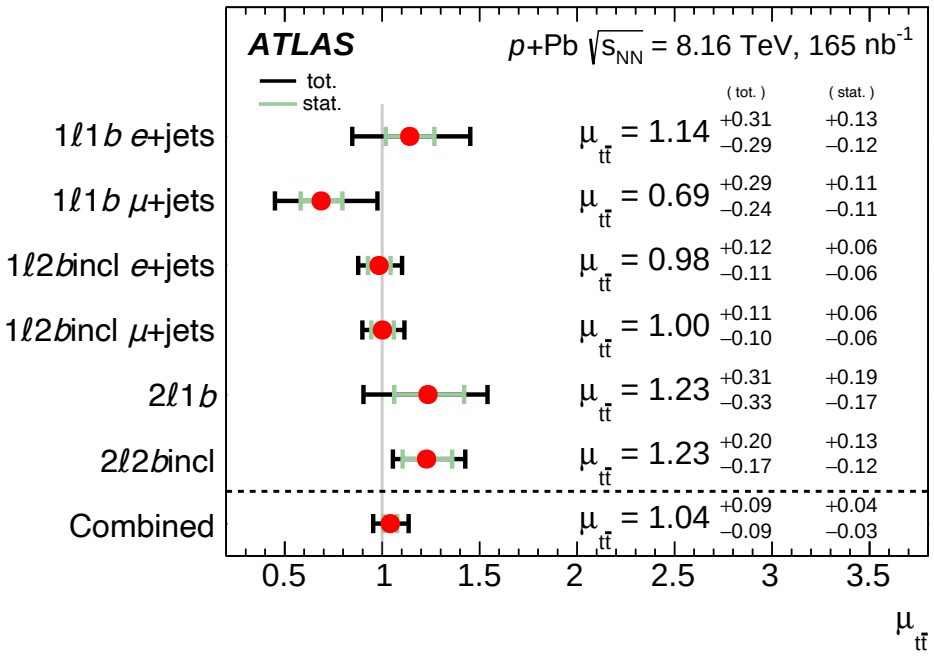


Figure 2.20: The observed best-fit values of the signal strength $\mu_{t\bar{t}}$ and their uncertainties by final-state category and combined. The individual $\mu_{t\bar{t}}$ values for the channels are obtained from a simultaneous fit with the signal-strength parameter for each channel allow to vary independently. The SM prediction is $\mu_{t\bar{t}} = 1$ [17].

2.10 Conclusion

This thesis presents the measurement of top quark pair production in proton-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV using data recorded by the ATLAS detector. This measurement explores the behavior of the top quark, the heaviest particle in the Standard Model, in a unique nuclear environment. By studying $t\bar{t}$ production in p+Pb collisions, we probe nuclear parton distribution functions (nPDFs), especially at high Bjorken-x, where phenomena like shadowing and anti-shadowing may play a role.

The analysis focuses on both the lepton+jets and dilepton channels, marking the first observation of the $t\bar{t}$ production in the dilepton channel in p+Pb collisions. The measured cross-section of the top quark pair production is $\sigma_{t\bar{t}} = 58.1 \pm 2.0$ (stat.) $^{+4.8}_{-4.4}$ (syst.) nb, with a total relative uncertainty of 9%. This result represents the most precise measurement of the $t\bar{t}$ cross-section in heavy-ion collisions to date. Systematic uncertainties dominate the precision in the lepton+jets channel, while statistical uncertainties are the primary limitation in the dilepton channel.

This measurement aligns with earlier observations made by the CMS collaboration, but extends the study into a new decay channel. By combining both the lepton+jets and dilepton channels, this study provides a more comprehensive understanding of the $t\bar{t}$ production in heavy-ion collisions. The result can be used for future fits of nPDFs and parton distribution functions in a previously unexplored kinematic region.

In addition to providing valuable insights into the nuclear environment, this analysis highlights the importance of systematic control, such as b-jet tagging techniques and optimized lepton selection criteria, which significantly improved the measurement precision. These findings will contribute to ongoing efforts to better understand the behavior of quarks and gluons in high-energy collisions and provide a stepping stone for future investigations in both proton-lead and lead-lead collisions.

References

[1] S. Abachi, B. Abbott, M. Abolins, et al. Observation of the Top Quark. *Physical Review Letters*, 74:2632–2643, 1995. https://doi.org/10.1103/PhysRevLett.74.2632.

[2] Nils Faltermann. Top quark mass and cross section at ATLAS and CMS, 2024. https://arxiv.org/abs/2405.09182.

[3] Joshua P. Ellis. TikZ-Feynman: Feynman diagrams with TikZ. *Computer Physics Communications*, 210:103–123, 2017. https://www.sciencedirect.com/science/article/pii/S0010465516302521.

[4] Michele Gallinaro. Top quark physics: A tool for discoveries. *Journal of Physics: Conference Series*, 447(1):012012, jul 2013. https://dx.doi.org/10.1088/1742-6596/447/1/012012.

[5] David d'Enterria. Top-quark pair production cross sections at NNLO+NNLL in pPb collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV, 2017. https://arxiv.org/abs/1706.09521.

[6] P. Nason, S. Dawson, and R.K. Ellis. The total cross section for the production of heavy quarks in hadronic collisions. *Nuclear Physics B*, 303(4):607–633, 1988. https://doi.org/10.1016/0550-3213(88)90422-1.

[7] Pavel M. Nadolsky Jon Pumplin Tie-Jiun Hou, Joey Huston et al. New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC, 2019. https://arxiv.org/abs/1912.10053.

[8] S. Alioli, P. Nason, C. Oleari, and E. Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *Journal of High Energy Physics*, 2010:043, 2010. https://doi.org/10.1007/jhep06(2010)043.

[9] Michal Czakon and Alexander Mitov. Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders. *Comput. Phys. Commun.*, 185:2930, 2014. https://doi.org/10.1016/j.cpc.2014.06.021.

[10] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Precision phenomenology with MCFM. *JHEP*, 12:034, 2019. https://doi.org/10.1007/JHEP12(2019)034.

[11] K. J. Eskola, Petja Paakkinen, Hannu Paukkunen, and Carlos A. Salgado. EPPS21: A global QCD analysis of nuclear PDFs. *Eur. Phys. J. C*, 82:413, 2022. http://dx.doi.org/10.1140/epjc/s10052-022-10359-0.

[12] R. L. Workman and Others. Review of Particle Physics. *PTEP*, 2022:083C01, 2022. http://dx.doi.org/10.1093/ptep/ptac097.

[13] Jakub Janda. pgf-pie: Pie charts with PGF/TikZ. CTAN, 2015. https://ctan.org/pkg/pgf-pie.

[14] K.J. Eskola, P. Paakkinen, H. Paukkunen, and C.A. Salgado. EPPS16: nuclear parton distributions with LHC data. *Eur. Phys. J. C*, 77:163, 2017. https://doi.org/10.1140/epjc/s10052-017-4725-9.

[15] K. Kovarik et al. nCTEQ15 - Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework. *Phys. Rev. D*, 93:085037, 2016. https://arxiv.org/abs/1509.00792.

[16] David d'Enterria, Krisztián Krajczár, and Hannu Paukkunen. Top-quark production in proton-nucleus and nucleus-nucleus collisions at LHC energies and beyond. *Nucl. Phys. A*, 933:256–287, 2015. https://arxiv.org/pdf/1501.05879.

[17] ATLAS Collaboration. Supplementary Figures from: Observation of $t\bar{t}$ production in lepton+jets and dilepton channels in p+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV with the ATLAS detector. 2023. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TOPQ-2023-32/.

[18] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3:S08003, August 2008. http://dx.doi.org/10.1088/1748-0221/3/08/S08003.

[19] P. Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *Journal of High Energy Physics*, 2004:040, 2004. https://dx.doi.org/10.1088/1126-6708/2004/11/040.

[20] R. D. Ball and others (NNPDF Collaboration). Parton distributions for the LHC Run II. *Journal of High Energy Physics*, 2015:40, 2015. https://doi.org/10.1007/JHEP04(2015)040.

[21] ATLAS Collaboration. ATLAS Pythia 8 tunes to 7 TeV data. *ATL-PHYS-PUB-2014-021*, 2014. https://cds.cern.ch/record/1966419.

[22] David d'Enterria. High-precision α_s measurements from LHC to FCC-ee. *Proceedings of the International Workshop on Future Colliders*, 2017:69–78, 2017. https://arxiv.org/abs/1512.05194.

[23] Jonathan M. Butterworth, Amanda Cooper-Sarkar, Anna Hall, et al. PDF4LHC recommendations for LHC Run II. *Journal of Physics G: Nuclear and Particle Physics*, 43(2):023001, 2016. https://arxiv.org/abs/1510.03865.

[24] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne. Parton distributions in the LHC era: MMHT 2014 PDFs. *European Physical Journal C*, 75:204, 2015. https://doi.org/10.1140/epjc/s10052-015-3397-6.

[25] S. Dulat, T. J. Hou, J. Gao, et al. New parton distribution functions from a global analysis of quantum chromodynamics. *Physical Review D*, 93:033006, 2016. https://doi.org/10.1103/PhysRevD.93.033006.

[26] J. Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *European Physical Journal C*, 76:196, 2016. https://arxiv.org/abs/1512.01178.

[27] Johan Alwall, Rikkert Frederix, Stefano Frixione, Valentin Hirschi, Fabio Maltoni, Olivier Mattelaer, Hua-Sheng Shao, Tim Stelzer, Paolo Torrielli, and Marco Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics (JHEP)*, 07:079, 2014. https://arxiv.org/abs/1405.0301.

[28] Enrico Bothmann and Others. Event generation with Sherpa 2.2. *Science Reports*, 7:081, 2019. https://arxiv.org/abs/1905.09127.

[29] Tanju Gleisberg, Stefan Hoeche, Frank Krauss, Marek Schonherr, Steffen Schumann, Frank Siegert, and Jan Winter. Event generation with SHERPA and COMIX. *Journal of High Energy Physics (JHEP)*, 02:007, 2009. https://arxiv.org/abs/0811.4622.

[30] F. Cascioli, P. Maierhöfer, and S. Pozzorini. Scattering Amplitudes with Open Loops. *Physical Review Letters*, 108:111601, 2012. https://arxiv.org/abs/1111.5206.

[31] Steffen Schumann and Frank Krauss. A parton shower algorithm based on Catani-Seymour dipole factorisation. *Journal of High Energy Physics (JHEP)*, 03:038, 2008. https://dx.doi.org/10.1088/1126-6708/2008/03/038.

[32] Stefano Catani, Frank Krauss, Bryan R. Webber, and R. Kuhn. QCD Matrix Elements + Parton Showers. *Journal of High Energy Physics (JHEP)*, 11:063, 2001. https://dx.doi.org/10.1088/1126-6708/2001/11/063.

[33] J. Butterworth et al. Single Boson and Diboson Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV. Technical Report CERN-2010-010, CERN, 2010. https://cds.cern.ch/record/1287902.

[34] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Richard Rietkerk. Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations. *Journal of High Energy Physics (JHEP)*, 03:015, 2013. https://dx.doi.org/10.1007/JHEP03(2013)015.

[35] Torbjörn Sjöstrand et al. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015. https://arxiv.org/abs/1410.3012.

[36] David J. Lange. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 462:152–155, 2001. https://dx.doi.org/10.1016/S0168-9002(01)00089-4.

[37] ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. *Journal of Instrumentation*, 14:P12006, 2019. https://dx.doi.org/10.1088/1748-0221/14/12/P12006.

[38] ATLAS Collaboration. Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV. *European Physical Journal C*, 81:578, 2021. https://dx.doi.org/10.1140/epjc/s10052-021-09233-2.

[39] ATLAS Collaboration. Measurement of the $t\bar{t}$ production cross-section in pp collisions at $\sqrt{s} = 5.02$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2023:138, 2023. https://dx.doi.org/10.1007/JHEP06(2023)138.

[40] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *Journal of High Energy Physics*, 2008:063, 2008. arXiv:0802.1189, https://arxiv.org/abs/0802.1189.

[41] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *European Physical Journal C*, 72:1896, 2012. https://dx.doi.org/10.1140/epjc/s10052-012-1896-2.

[42] T. T. Rinn, A. M. Sickles, and P. A. Steinberg. Measurement of Dijet Asymmetry in Pb+Pb and pp Collisions at 5.02 TeV, 2020. https://cds.cern.ch/record/2711618.

[43] ATLAS Collaboration. Deep Sets Based Neural Networks for Impact Parameter Flavour Tagging in ATLAS, 2020. https://cds.cern.ch/record/2718948.

[44] ATLAS Collaboration. ATLAS flavour-tagging algorithms for the LHC Run 2 *pp* collision dataset, 2023. https://cds.cern.ch/record/2842028.

[45] ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data. *JINST*, 14:P12006, 2019. https://doi.org/10.1088/1748-0221/14/12/P12006.

[46] ATLAS Collaboration. Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 81(8):689, 2021. https://doi.org/10.1140/epjc/s10052-021-09402-3.

[47] ATLAS Collaboration. Jet energy scale and its uncertainty for jets reconstructed using the ATLAS heavy ion jet algorithm. 2015. https://cds.cern.ch/record/2008677.

[48] ATLAS Collaboration. ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 79(12):970, 2019. https://doi.org/10.1140/epjc/s10052-019-7450-8.

[49] ATLAS Collaboration. Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector. *Eur. Phys. J. C*, 82(2):95, 2022. https://doi.org/10.1140/epjc/s10052-022-10047-2.

[50] ATLAS Collaboration. Calibration of the light-flavour jet mistagging efficiency of the b-tagging algorithms with Z+jets events using 139 fb^{-1} of ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 83(8):728, 2023. https://arxiv.org/abs/2301.06319.

[51] C. F. Berger et al. Precise Predictions for W+4-Jet Production at the Large Hadron Collider. *Phys. Rev. Lett.*, 106:092001, 2011. https://doi.org/10.1103/PhysRevLett.106.092001.

[52] F. A. Berends, H. Kuijf, B. Tausk, and W. T. Giele. On the production of a W and jets at hadron colliders. *Nucl. Phys. B*, 357:32–64, 1991. https://doi.org/10.1016/0550-3213(91)90458-A.

[53] S. D. Ellis, R. Kleiss, and W. J. Stirling. W's, Z's and Jets. *Phys. Lett. B*, 154:435–440, 1985. https://doi.org/10.1016/0370-2693(85)90425-3.

[54] E. Re. Single-top Wt-channel production matched with parton showers using the POWHEG method. *Eur. Phys. J. C*, 71:1547, 2011. https://doi.org/10.1140/epjc/s10052-011-1547-z.

[55] ATLAS Collaboration. Observation of $t\bar{t}$ production in the lepton+jets and dilepton channels in $p + Pb$ collisions at $\sqrt{s_{\mathrm{NN}}} = 8.16$ TeV with the ATLAS detector. 2023. https://arxiv.org/abs/2405.05078.

[56] ATLAS Collaboration. Measurement of the $t\bar{t}$ cross section and its ratio to the Z production cross section using pp collisions at $\sqrt{s} = 13.6$ TeV with the ATLAS detector. *Phys. Lett. B*, 848:138376, 2024. https://doi.org/10.1016/j.physletb.2023.138376.

[57] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J. C*, 83:982, 2023. https://doi.org/10.1140/epjc/s10052-023-11747-w.

[58] G. Avoni et al. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *JINST*, 13:P07017, 2018. https://doi.org/10.1088/1748-0221/13/07/P07017.

[59] TRExFitter Collaboration. TRExFitter Documentation, 2024. https://trexfitter-docs.web.cern.ch/trexfitter-docs/.

[60] P.A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020. https://academic.oup.com/ptep/article/2020/8/083C01/5891211.

[61] CMS Collaboration. Observation of top quark production in proton-nucleus collisions. *Phys. Rev. Lett.*, 119:242001, 2017. https://arxiv.org/abs/1709.07411.

[62] ATLAS and CMS Collaborations. Combination of inclusive top-quark pair production cross-section measurements using ATLAS and CMS data at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 07:213, 2023. https://arxiv.org/abs/2205.13830.

[63] J. M. Campbell, D. Wackeroth, and J. Zhou. Study of weak corrections to Drell-Yan, top-quark pair, and dijet production at high energies with MCFM. *Phys. Rev. D*, 94:093009, 2016. https://arxiv.org/abs/1608.03356.

[64] K. J. Eskola, P. Paakkinen, H. Paukkunen, and C. A. Salgado. Towards EPPS21 nuclear PDFs. *SciPost Phys. Proc.*, 8:033, 2022. https://arxiv.org/abs/2106.13661.

[65] P. Duwentäster et al. Impact of heavy quark and quarkonium data on nuclear gluon PDFs. *Phys. Rev. D*, 105:114043, 2022. https://arxiv.org/abs/2204.09982.

[66] M. Ubiali. Towards the NNPDF3.0 parton set for the second LHC run. In *PoS DIS2014*, volume 041, 2014. https://arxiv.org/abs/1407.3122.

[67] R. Abdul Khalek et al. nNNPDF3.0: evidence for a modified partonic structure in heavy nuclei. *Eur. Phys. J. C*, 82:507, 2022. https://arxiv.org/abs/2201.12363.

[68] I. Helenius, M. Walt, and W. Vogelsang. TUJU21: nuclear PDFs with electroweak-boson data at NNLO. In *29th International Workshop on Deep-Inelastic Scattering and Related Subjects*, 2022. https://arxiv.org/abs/2207.04654.