SEARCHES FOR NEW PHYSICS WITH TOP QUARKS IN THE

FINAL STATE AND MEASUREMENT OF THE DIFFERENTIAL

CROSS-SECTION OF $t\bar{t}$ PRODUCTION IN ASSOCIATION WITH

ADDITIONAL HEAVY FLAVOR JETS

By

EGOR ANTIPOV

Bachelor of Engineering in Engineering Physics
National Research Nuclear University "MEPhI"
Moscow, Russia
2016

SEARCHES FOR NEW PHYSICS WITH TOP QUARKS IN THE

FINAL STATE AND MEASUREMENT OF THE DIFFERENTIAL

CROSS-SECTION OF $t\bar{t}$ PRODUCTION IN ASSOCIATION WITH

ADDITIONAL HEAVY FLAVOR JETS

Dissertation Approved:

Dr. Flera Rizatdinova

Dissertation Advisor

Dr. Joseph G. Haley

Dr. Kaladi S. Babu

Dr. James E. Stine

ACKNOWLEDGMENTS

First of all, I want to thank the whole Oklahoma State ATLAS group and my scientific advisor Dr. Flera Rizatdinova in particular. Thank you very much for giving me an opportunity to become a physicist and for the invaluable help and support in both research and everyday life difficulties. I also want to thank our engineer Steven Welch for being a perfect mentor in hardware projects and a good friend.

I did my first physics analysis with the ATLAS multi-b jet analysis group. It has been a great pleasure to work with this amazing group of physicists. I want to express my gratitude to Dr. Maximilian Swiatlowski for his mentorship and coordination of the analysis. I want to thank Dr. Giordon Stark for his help in coding issues. And many thanks to Dr. Chiara Rizzi for her valuable input to the discussions of my work.

I want to appreciate an opportunity to be a part of the $t\bar{t}$+jets ATLAS group. It has been a pleasure to work with you! In particular, I want to thank Dr. Nedaa-Alexandra Asbah, Dr. Mahsana Ahsan, and Lucas Klein for awesome collaborative work in the analysis.

During my PhD, I also had a great experience visiting US national laboratories. I want to express my appreciation to the Lawrence Berkeley National Laboratory ATLAS group, and in particular to Dr. Maurice Garcia-Sciveres for his guidance and to Dr. Aleksandra Dimitrievska for her help with my project during my 6-months visit.

I want to express my deepest gratitude to Dr. Charles Young, Dr. Su Dong, Dr. Caterina Vernieri, and the whole SLAC ATLAS group for the opportunity to work with you for 9 months, it was a great experience!

During my PhD at Oklahoma State University, I had a chance to work with many

outstanding scientists from the all around the world. Many thanks and appreciation to all of you for being with me along the way and making the thesis possible!

I also want to recognize my undergraduate scientific adviser Dr. Evgeny Krastelev, who helped me to start a career in science by mentoring my research diploma project and by broadening my views to engineering and science. Thank you very much!

I want to thank my whole family: my Mom, Dad, Brother, and my Wife. Thanks a lot for your support and patience! Last but not least, I want to thank all my friends who also supported me here: Dr. Anton Avramov, Dr. Bryan and Andrea Hayes, Evan Van de Wall, Jonny Riggs, and many others.

Name: EGOR ANTIPOV

Date of Degree: May 2022

Title of Study: SEARCHES FOR NEW PHYSICS WITH TOP QUARKS IN THE FINAL STATE AND MEASUREMENT OF THE DIFFERENTIAL CROSS-SECTION OF $t\bar{t}$ PRODUCTION IN ASSOCIATION WITH ADDITIONAL HEAVY FLAVOR JETS

Major Field: PHYSICS

Abstract:

This thesis presents the results of the search for the pair-production of gluinos $\tilde{g}$ decaying via supersymmetric top quarks into the top quark and the lightest neutralino $\tilde{\chi}_1^0$. It utilizes an LHC proton-proton collision dataset of the centre-of-mass energy $\sqrt{s} = 13$ TeV with an integrated luminosity of 139 fb^{-1} collected by the ATLAS detector in the $2015-2018$ data-taking period. The search is performed in events containing large missing transverse momentum and several energetic jets, at least three of which are identified as originating from b-quarks. The analysis is done in two final states: the first is required to have at least one lepton, and the second one with a veto on leptons. No excess of data above the predicted background is found. The results are presented in terms of the combined 95% CL upper limit on the gluino and neutralino masses $m_{\tilde{g}}$ and $m_{\tilde{\chi}_1^0}$. The highest excluded $m_{\tilde{g}}$ is approximately 2240 GeV, which corresponds to the massless neutralino, and the highest excluded $m_{\tilde{\chi}_1^0}$ is 1300 GeV, which corresponds to approximately $m_{\tilde{g}} = 2100$ GeV.

The production of top-antitop pairs in association with additional b-jets was found to be the most challenging process to discriminate from the hypothetical gluino production signal. This process is not well-described by the current Monte Carlo generators. This thesis presents differential cross-section measurements of $t\bar{t}+b$ and $t\bar{t}+b\bar{b}$ production using machine learning techniques. The studies will help to improve the $t\bar{t}$ production simulation with future Monte Carlo generators.

The Large Hadron Collider will be upgraded to high luminosity. To address the increased data rate, the entire ATLAS inner detector will be replaced with a new all-silicon tracker, ITk. This thesis presents serial powering and data readout tests of demonstrators of the ITk pixel readout chips.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

xv

CHAPTER I

Introduction

Particle physics studies properties of elementary particles as well as interactions between them. The beginning of the experimental particle physics goes back to 1897, when J.J. Thompson discovered [1] the first elementary particle - the electron. Since then, much of effort was put into the experimental discoveries of new particles and development of theoretical models. The most successful theory of particle physics so far is the Standard Model [2, 3, 4]. It was formulated during the second half of the 20$^\text{th}$ century with the combined work of many scientists from around the world. The Standard model was not only successful in describing known particles and interactions, but also predicted new ones.

As time passed, the experimental setups became more and more complicated in order to search for new particles and to measure properties of the known ones more precisely. Although to discover an electron in 1897 it was enough to build just a table-top setup, to run qualitative analyses modern particle physics experiments require more complexity. One of the examples is the Large Hadron Collider [5] at CERN with two general purpose detectors, ATLAS [6] and CMS [7], where large international collaborations of physicists explore the cutting edge of physics. In 2012, 115 years after the discovery of the electron, the scientific society celebrated the greatest achievement of the ATLAS and CMS experiments − the discovery [8] [9] of the missing piece of the Standard Model, the Higgs boson.

Although the Standard Model was successful in providing experimental predictions, there are several phenomena that it cannot explain. For example, why are there three generations of particles? Or, why is there more matter than antimatter? What are the dark matter and the dark energy? Why do the weak, electromagnetic and string forces have different strengths? Another problem is related to the Higgs boson mass, which must be corrected by an artificial parameter to prevent it from divergency. The latter is known as the hierarchy problem. There has been made quite a few attempts to extend the Standard Model in order to resolve the problems. One of the promising extensions is Supersymmetry. Supersymmetry with R-parity conservation resolves the hierarchy and the dark matter problems. However, Supersymmetry hasn't been discovered yet, but many physics analyses at ATLAS and CMS are dedicated to searches for it.

The thesis presents one of the searches for the Minimal Supersymmetric Standard Model at ATLAS. The theoretical overview of the Standard Model and the Minimal Supersymmetric Standard Model are presented in Chapter II. The Large Hadron Collider and the ATLAS detector are described in Chapter III. Chapter IV explains data-taking process at the ATLAS

detector and Monte Carlo simulations of pp interactions as well as modeling of interactions of resulting particles with the ATLAS detector. The detector records events as a collection of responses from independent systems; reconstruction of physics objects from the signals is presented in Chapter V. The search for the Minimal Supersymmetric Standard Model with R-parity conservation with top quarks and large missing transverse momentum in the final state using data from the ATLAS detector is presented in Chapter VI. Many searches at ATLAS, including the presented analysis, have $t\bar{t}$ production as the dominant background, therefore they rely on good description of process by the Monte Carlo modeling. An attempt to correct the modeling to better describe kinematic of resulting b-jets using machine learning techniques is presented in Chapter VII. Since most of the searches for new physics are statistically limited, evolution of the analyses strategies and Monte Carlo modeling is not enough for moving further in the analyses. To allow for collection of larger statistic, the Large Hadron Collider, as well as all the general purpose detectors, will be upgraded. Tests of readout systems for the upgraded ATLAS silicon tracker are presented in Chapter VIII.

CHAPTER II

Theoretical Background

The Standard Model (SM) of particle physics is the most successful theory so far. It classifies all known elementary particles and interactions between them. This chapter presents an introduction to the SM of particle physics and an extension the SM, named Supersymmetry (SUSY). Section 2.1 provides an overview of the SM with open questions of the theory, and section 2.2 presents SUSY, which answers some of the questions.

2.1 The Standard Model of Particle Physics

There are four known fundamental forces in the Universe: electromagnetic, weak, strong, and gravitational forces. The unified theory of electromagnetic, weak and Higgs interactions is called the electroweak model, and the theory of strong interaction is called quantum chromodynamics. The SM combines the electroweak model and quantum chromodynamics while omitting gravity, the weakest fundamental interaction.

The standard model is a gauge quantum field theory that combines three different Lie groups into one: $SU_C(3) \times SU_L(2) \times U_Y(1)$. The $SU_L(2) \times U_Y(1)$ product represents the Glashow-Weinberg-Salam theory of electroweak interactions which is discussed in section 2.1.3. $SU_C(3)$ term represents the Quantum Chromodynamics theory, and the theory is discussed in section 2.1.5. The SM predicts a number of elementary particles that can be sorted into two families: fermions and bosons. Fermions have half-integer spin values and constitute matter. Fermions of the SM are classified into quarks and leptons. Gauge bosons have integer spin values and act as force carriers for fermions. Figure 1 summarized particles of the SM. The discovery of the Higgs boson in 2012 at the LHC [8, 9] completed the SM of particle physics. In the following year after the discovery, Peter Higgs and François Englert were awarded with a Noble Prize in Physics for their development of the Higgs mechanism in 1964.
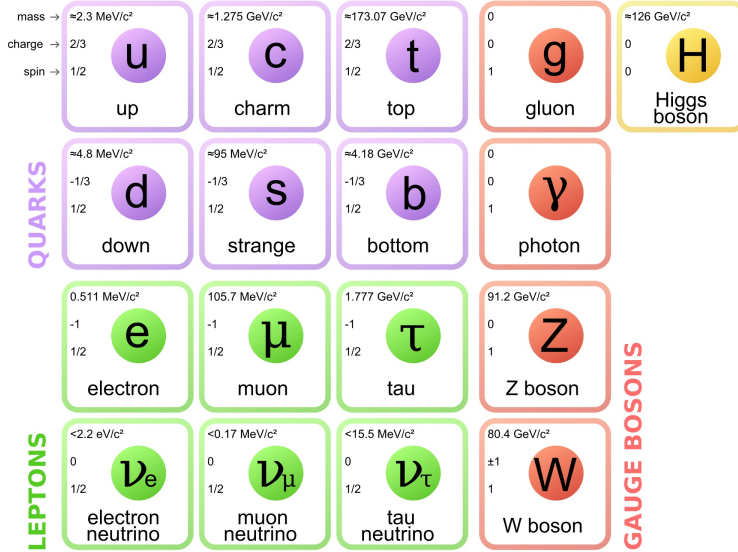
Figure 1: Particles of the SM. Quarks and leptons are fermions with spin 1/2, highlighted in green and purple respectively. Gauge bosons have integer spin, and divided into two categories: gluon, photon and W/Z bosons with spin values 1 (in red), and the Higgs bosons with spin value 0. Masses, charges and spin values are provided by the Particle Data Group [10].

2.1.1 Standard Model Particles and Fields

As discussed above, all the particles of the SM are categorized into two families based on their spin values: fermions and bosons. SM fermions are spin-$\frac{1}{2}$ particles, therefore they obey the Fermi-Dirac statistics. Fermions make up all the visible matter of the Universe. SM fermions are categorized into quarks and leptons.

The family of quarks consists of six particles arranged into three generations, and each generation consist of two quarks. Quarks interact via strong interaction, and also experience electroweak interaction. The lightest quark is the *down* quark with mass of ≈ 4.8 MeV, and the heaviest is the *top* with $m_{top} \approx 173.07$ GeV. *Top* quark was discovered at Fermilab in the CDF and D0 experiments in 1995 [11, 12].

Leptons include electrons, muons, and tau, as well as three associated neutrinos. The first three particles are electrically charged, and the associated neutrinos are neutral. The charged leptons experience electromagnetic and weak interactions, while the neutral leptons interact only weakly. Similar to quarks, the six leptons are arranged into three generations with two particles in each.

There are several similarities and differences between quarks and leptons. All fermions have left-handed doublets ψ_L within which two particles in one generation differ in charge, and only charged fermions have right-handed singlets ψ_R. Due to so-called color confinement, quarks cannot exist in a free state; instead, they constitute bound states named hadrons consisting of at least two quarks. Hadrons made up of two quarks are called mesons, and

made up of three − baryons. Leptons don't interact strongly; therefore, the color confinement doesn't affect them, and leptons can exist and be observed in a free state.

Bosons are spin-integer particles, and thus they obey Bose-Einstein statistics. Bosons in the SM are force carriers, and all interactions between the fermions are carried by an exchange of bosons. For example, all electroweak interactions between two electrons are due to photon exchange.

Particles	Electromagnetic	Weak	Strong	Gravity
Leptons	Yes	Yes	No	Yes
Quarks	Yes	Yes	Yes	Yes
Gauge Bosons	Photon	W/Z	Gluon	Graviton*

Table 1: Summary of interactions of fundamental particles and force carriers names. *graviton hasn't been discovered yet.

2.1.2 Gauge Invariance

Modern particle physics theories are formulated within the Quantum Field Theory (QFT) framework, where a Lagrangian $\mathcal{L}$ controls the dynamics and kinematics of the theory. One can describe a *particle* as a *localized entity*. In classical mechanics, the *position* of a particle is described as a function of time; for example, $x(t)$, $y(t)$ and $z(t)$ in the Cartesian coordinate system. In contrast, particles in QFT are described in terms of *dynamic fields* that exist in space-time. The SM of particle physics is constructed using the QFT in two steps. First, a set of symmetries is postulated. Second, the most general renormalizable Lagrangian is written from its *fields* (particles) that observes postulated symmetries.

In High Energy Physics it is very useful to exploit so-called natural systems of units ($c = \hbar = 1$), because it makes all equations look lighter and easier to read. Usually, the c's and $\hbar$'s are added in the final equation for the units to come out right. In the natural units, the free Lagrangian for fermions and the massless electromagnetic vector A^μ is are given by:

$$\mathcal{L}_{fermion}(x) = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x) \tag{2.1.1}$$

$$\mathcal{L}_{vector}(x) = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} \tag{2.1.2}$$

$$(F^a_{\mu\nu} = \partial_\mu A^a_\nu + \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu) \tag{2.1.3}$$

In the equations above ψ is the fermion *field*, γ are the 4×4 Dirac matrices, ∂_{mu} is the four-vector derivative, f^{ijk} are the structure constant of the Lie algebra of the generators of the gauge groups. Lagrangians 2.1.1 and 2.1.3 must be invariant under the local and global

gauge transformations. *Global* gauge transformation is defined as a global (i.e. independent from x) change in the phase of a fermion field:

$$\psi'(x) = e^{-ig\tau\theta}\psi(x) \tag{2.1.4}$$

where g is the coupling constant and τ is the group's generator. If the phase change parameter depends on x, such gauge transformation is called *local*:

$$\psi'(x) = e^{-ig\tau\theta(x)}\psi(x) \tag{2.1.5}$$

The kinetic term in the $\mathcal{L}_{fermion}$ is not invariant under local gauge transformations: $\partial_{mu}\theta(x) \neq 0$. To make the larganmian invariant under the transformation, a covariant derivative D_μ is added to the theory in place of ∂_{mu}:

$$D_\mu = \partial_{mu} - ig\Sigma_j\tau^a A_\mu^a \tag{2.1.6}$$

where A_μ^a are massless vector gauge fields. The massless vector gauge fields transform under local gauge transformations as

$$A_\mu^{a\prime} = A_\mu^a - \frac{1}{g}\partial_\mu\theta^a + f_{abc}\theta^b A_\mu^c \tag{2.1.7}$$

By substituting both 2.1.6 and 2.1.7 into 2.1.1 and 2.1.2 one can obtain Lagrangians $\mathcal{L}_{fermion}(x)$ and $\mathcal{L}_{vector}(x)$ that are invariant under the local transformations.

To summarize, the SM is locally invariant under the $SU_C(3) \times SU_L(2) \times U_Y(1)$ local gauge symmetry. The $SU_L(2) \times U_Y(1)$ symmetry gives rise to the theory of electroweak interactions, while the theory of strong interactions follows from the $SU_C(3)$ symmetry. Noether's theorems state that every continuous symmetry under which the Lagrangian (or Hamiltonian) is invariant in form is associated with a conservation law [13]. The SM local gauge invariances imply the conservation of isospin, hypercharge, and color charge.

2.1.3 Quantum Electrodynamics

Quantum electrodynamics (QED) is a QFT of electrodynamics. The theory describes how light and matter interact with each other. QED is the first theory that archived full agreement between special relativity and quantum mechanics. Mathematically, QED is a local gauge theory in which all Lagrangians are invariant under the $U(1)$ group symmetry. The gauge field is called the *electromagnetic* field. The photons mediate interactions between spin-$\frac{1}{2}$ fields. The QED Lagrangian is given by:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.1.8}$$

where

- ψ is field of a spin-$\frac{1}{2}$ particles;

- $\bar{\psi} = \psi^\dagger\gamma^0$ is a Dirac adjoint of ψ;

6

- γ^μ are Dirac matrices;

- D_μ is the gauge covariant derivative: $D_\mu = \partial_\mu + ieA_\mu + ieB_\mu$;

- e is the coupling constant which equals to the electric charge of the spin-$\frac{1}{2}$ particle;

- A_μ is the covariant four-potential of the electromagnetic field generated by the spin-$\frac{1}{2}$ particle itself;

- B_μ is the external field due to an external source;

- m is the mass of the spin-$\frac{1}{2}$ particle;

- $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor.

The first part of the Lagrangian represents the fermionic field: the kinetic energy, the mass of a fermion, and the interactions between a fermion and a vector field. The last part, $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, describes the kinetic energy of the photon. The local gauge invariance under $U(1)$ transformations leads to massless photons, because the mass term $\frac{m^2}{2}A_\mu A^\mu$ violates the local gauge invariance unless $m = 0$.

Examples of QED interactions are given in figure 2. Feynman diagrams are used to compute the amplitude $\mathcal{M}$ of a process. The amplitudes are used to derive observables, such as the production cross-section. Figure 2a shows the leading order diagram of the interaction between two pairs of leptons by exchanging a photon. Particles in the initial or/and final state can emit a photon. Such processes are called *initial* and *final* state radiation. An example of initial state radiation is given in figure 2b.

(a) The leading order Feynman diagram of a QED interaction.

(b) Feynman diagram of a QED interaction with initial state radiation.
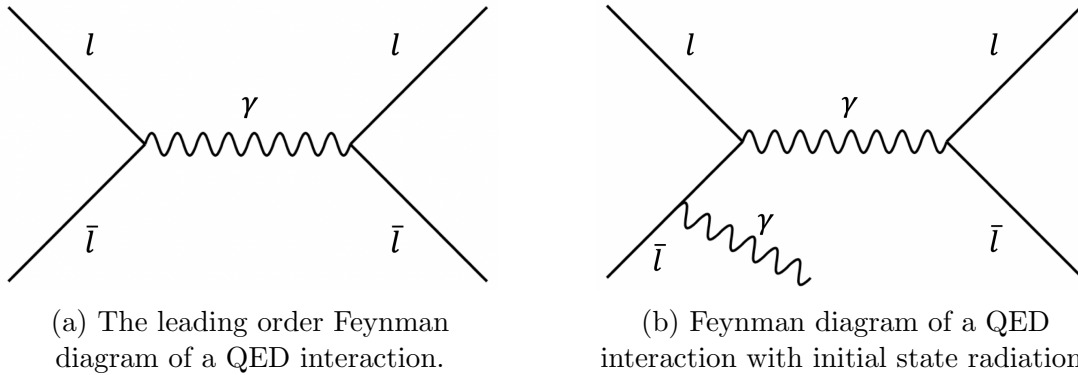
Figure 2: Examples of QED interactions.

2.1.4 Electroweak Theory

Electroweak interaction unifies electromagnetic and weak interactions. The unified theory models electromagnetic and weak forces as two different aspects of one force. Although the two forces appear different at low energies, they merge into one force at energies above the so-called unification energy, which approximately equals to 246 GeV. This energy corresponds to the temperature of the order of 10^{15} K, which haven't been achieved since shortly after

the Big Bang.

The unified electroweak is named Glashow-Weinberg-Salam (GWS) after three theorists, Sheldon Glashow, Steven Weinberg and Abdus Salam. They were awarded with the Noble Prize in Physics in 1979 for their contributions to the unification of weak and electromagnetic interactions.

The theory of electroweak interaction is a non-Abelian gauge theory that is based on $SU_L(2) \times U_Y(1)$ symmetry. The $SU_L(2)$ group governs weak interaction, and the subscript L denotes the left-handed fermions. Electromagnetic interaction is governed by the $U_Y(1)$ gauge group, where Y indicates the weak hypercharge.

In the GWS theory, leptons are either left-haded doublets or right-handed singlets of the $SU(2)$ group:

$$\chi_L = \begin{pmatrix} \nu_l \\ l \end{pmatrix}_L, \qquad \chi_R = l_R, \qquad (l = 1, \ 2, \ 3) \ , \qquad (2.1.9)$$

where χ_L consists of the left-handed neutrinos and leptons. No right-handed neutrinos has been observed in experiments; therefore, the right-handed leptons χ_R are singlet.

Similar to the leptons, quarks are left-haded doublets and right-handed singlets:

$$Q_{L_i} = \begin{pmatrix} u_i \\ d_i \end{pmatrix}, \qquad Q_{R_i} = u_{R_i}, d_{R_i} \ , \qquad (i = 1, \ 2, \ 3), \qquad (2.1.10)$$

where u and d represent the *up-* and *down-* types quarks respectively.

The covariant derivative for the $SU_L(2) \times U_Y(1)$ group is:

$$D_\mu = \partial_\mu - i\frac{g'}{2}YB_\mu - i\frac{g}{2}\vec{\tau}\vec{A_\mu} \ , \qquad (2.1.11)$$

where $\vec{A_\mu}$ and B_μ are the gauge boson fields associated with the $SU_L(2)$ and U_Y groups respectively; g and g' are the gauge coupling constants; Y is the weak hypercharge; $\vec{\tau}$ is the weak isospin operator defined by the Pauly matrices.

The gauge invariant Lagrangian for GWS model is given by:

$$\mathcal{L}_F = \bar{L}i\gamma^\mu D_\mu L + \bar{R}i\gamma^\mu D_\mu R \ , \qquad (2.1.12)$$

where L and R denote the left- and right-handed fermion fields. The kinetic term $\mathcal{L}_G$ of the gauge field is given by:

$$\mathcal{L}_G = -\frac{1}{4}W^a_{\mu\nu}W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \ , \qquad (a = 1, \ 2, \ 3) \qquad (2.1.13)$$

where $W^i_{\mu\nu}$ and $B_{\mu\nu}$ are the gauge field strength tensors for the weak isospin and hypercharge gauge fields. There are no fermion and gauge boson fields mass terms in equations 2.1.12

and 2.1.13 respectively, which contradicts the experimental results. To make gauge bosons massive, the Higgs mechanism [14, 15] was implemented. The mechanism adds a quantum field, known as the Higgs field, to the SM. Below some threshold temperature, the field causes spontaneous symmetry breaking during interactions, causing the gauge bosons interacting with the field to acquire mass. Fermions also acquire masses via interactions with the Higgs field, but not the same way as gauge bosons.

To spontaneously break the symmetry, the Higgs mechanism introduces a complex scalar field in the spinor representation of $\text{SU}_L(2)$:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} , \tag{2.1.14}$$

where ϕ^+ and ϕ^0 are positively charged and neutral complex scalar fields. The Lagrangian of the scalar Higgs is given by:

$$\mathcal{L}_{Higgs} = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi) , \tag{2.1.15}$$

where $\phi^\dagger$ is the Hermitian conjugate of ϕ. The covariant derivative D^μ was defined in equation 2.1.11. The potential term $V(\phi)$ is given by:

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 , \tag{2.1.16}$$

where μ and λ are real constant parameters. The λ term describes quadratic self-interactions among the scalar field, and vacuum stability requires the parameter to be positive. The values of parameters defines the shape of the potential. If $\mu^2 > 0$ and $\lambda > 0$, the potential has the form of a parabola with one minimum corresponding to $\phi = 0$. However, if $\lambda > 0$, but $\mu^2 < 0$, the potential acquires the form of a Mexican hat, and the ground state forms a circle with radius of $v = \sqrt{\mu^2/\lambda}$ in the (ϕ_1, ϕ_2) phase space. Such shape of potential $V(\phi)$ allows for the spontaneous symmetry breaking, since the ground state is not unique: the symmetry will be broken by arbitrary choice of the complex phase, and the field falls on a position that is randomly chosen on the minimal potential circle.

(a) $\lambda > 0$, $\mu^2 > 0$. One minimum. (b) $\lambda > 0$, $\mu^2 < 0$. Mimina are defined as $v = \sqrt{\mu^2/\lambda}$.
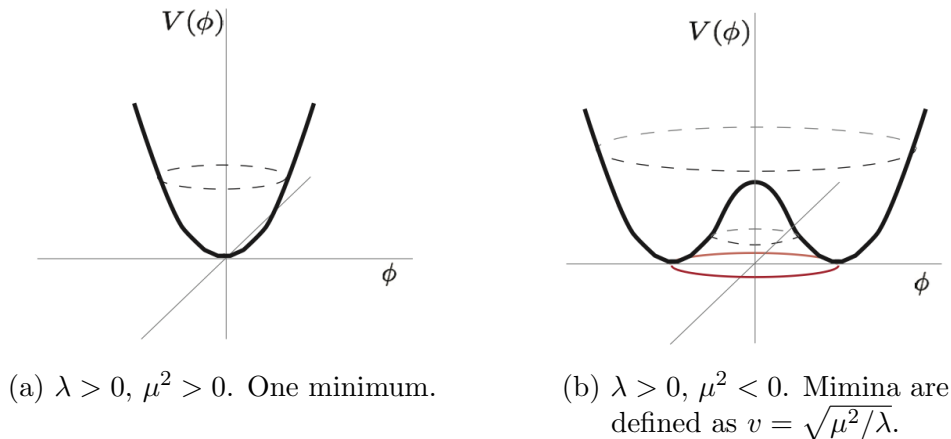
Figure 3: Higgs potentials for different sets of parameters μ^2 and λ.

When the spontaneous symmetry breaking occurs, the scalar field develops a non-zero vacuum expectation value (provided that $\mu > 0$). Since the potential depends only on $\phi\phi^\dagger$, we can arbitrarily choose the vacuum expectation value of the Higgs field as:

$$\langle\phi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} . \tag{2.1.17}$$

Given the vacuum expectation value of ϕ, we can parametrize the scalar doublet in terms of the fields relative to the $\langle\phi\rangle$:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix} , \tag{2.1.18}$$

and the piece of $\mathcal{L}_{Higgs}$ from the equation 2.1.15 that generates the gauge bosons' masses yields:

$$(D^\mu\phi)^\dagger(D_\mu\phi) = \frac{v^2}{8}[g^2((W^1_\mu)^2 + (W^2_\mu)^2) + (gW^3_\mu - g'B_\mu)^2] . \tag{2.1.19}$$

The charged vector boson W^-_μ and its complex conjugate are defined as:

$$W^\pm_\mu \equiv \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) . \tag{2.1.20}$$

Thereby, the g^2 term in equation 2.1.19 becomes:

$$\frac{1}{2} \left(\frac{g\,v}{2}\right)^2 W^\dagger_\mu W^\mu . \tag{2.1.21}$$

yielding the W mass:

$$m_W = \frac{gv}{2} . \tag{2.1.22}$$

The two remaining neutral gauge bosons, Z and A, are defined as follows:

$$Z_\mu \equiv \frac{1}{\sqrt{g^2 + g'^2}}(gW^3_\mu - g'B_\mu) \qquad \text{with} \qquad m_Z = \frac{v}{2}\sqrt{g^2 + g'^2}, \tag{2.1.23}$$

$$A_\mu \equiv \frac{1}{\sqrt{g^2 + g'^2}}(gW^3_\mu + g'B_\mu) \qquad \text{with} \qquad m_A = 0 . \tag{2.1.24}$$

The latter boson, A_μ, is the massless carrier of the electromagnetic force, the photon. Fermions acquire their masses via interacting with the Higgs field as well. Consider the Yukawa Lagrangian for $e^- e^+$:

$$\begin{aligned}
\mathcal{L}_{Yukawa} &= -G_e(\bar{\chi}_L \phi \chi_R + \bar{\chi}_R \phi^\dagger \chi_L) \\
&= -G_e(\bar{e}_L \frac{1}{\sqrt{2}}(v+h)e_R + \bar{e}_R \frac{1}{\sqrt{2}}(v+h)e_L) \\
&= -\frac{G_e v}{\sqrt{2}}\bar{e}e - \frac{G_e}{\sqrt{2}}h\bar{e}e \ .
\end{aligned} \quad (2.1.25)$$

For definitions of χ_L and χ_R refer to the equation 2.1.9. The first term in the equation above represents the fermion mass, $m_e = G_e v/\sqrt{2}$. The second term corresponds to the fermion coupling to the Higgs field, and the coupling is proportional to the fermion mass as $G_e/\sqrt{2} = m_e/v$.

2.1.5 Quantum Chromodynamics

Quantum chromodynamics (QCD) is the theory of strong interactions between the color charged particles, quarks and gluons. The mediator of the strong field is the gluon. There are three types of color charge: red, blue and green. Analogous to the eclectic charge, each of the three color charges have a corresponding anti-charge (i.e. anti-red, anti-blue, anti-green). QCD is a non-Abelian gauge QFT with the symmetry group $SU_C(3)$, where the C refers to the color charge. QCD yields 8 gauge bosons - gluons - which carry color charge; therefore, gluons not only mediate the strong interaction, but also participate in such interactions (unlike the photon in QED). Gluons are massless and carry no electric charge. Figure 4 shows the leading order QCD Feynman diagram in (a) and the sub-leading with the final state radiation in (b).

(a) The leading order Feynman diagram of a QCD interaction.

(b) Feynman diagram of a QCD interaction with final state radiation.
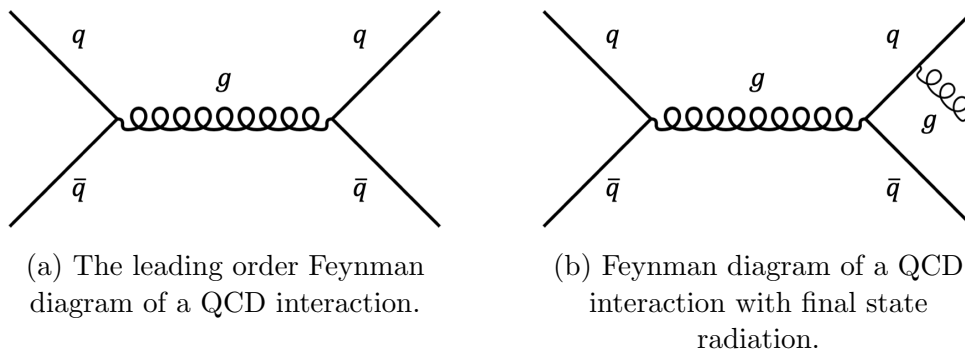
Figure 4: Examples of QED interactions.

The gauge invariant QCD Lagrangian is given by:

$$\mathcal{L}_{QCD} = \Sigma_n \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu a} \ , \quad (2.1.26)$$

where the summation goes over all quark flavors; the covariant derivative D_μ is defined as $D_\mu = \partial_\mu - i\frac{g_s}{2}\lambda^a A^a_\mu$ to have a gauge invariant Lagrangian; $A^a_\mu(a = 1, 2, ..., 8)$ are the gluons;

g_s is the strong coupling constant; λ^i are the 3×3 Gell-Mann matrices. $G_{\mu\nu}^a$ is the gauge invariant filed strengths tensors for gluon fields A_μ^a, and given by:

$$G_{\mu\nu}^a = \partial_\mu A_\mu^a - \partial_\nu A_\mu^a + g f^{abc} A_\mu^b A_\nu^c \ , \tag{2.1.27}$$

where f^{abc} are the structure constants of SU(3) group and defined as $[\lambda^b, \lambda^c] = i f_a^{bc} \lambda^a$. $\mathrm{SU}_C(3)$ group is not Abelian ($f_{abc} \neq 0$), which means couplings for three- or four-vertices gluon interactions are allowed in QCD.

The strength of strong interaction depends on the momentum transfer between the interacting particles, and this dependency can affect gluon-loop and quark-loop corrections at the leading and higher orders. The strong coupling constant is given by:

$$\alpha_S(Q^2) = \frac{4\pi}{\beta_0 \ln(Q/\Lambda^2)} \ , \tag{2.1.28}$$

where $\beta_0 = 11 - \frac{2}{3} n_f$, and n_f is the number of quarks (i.e. $n_f = 6$); Λ is the so-called cut-off scale. Given $n_f = 6$, the running coupling constant $\alpha_S(Q^2)$ decreases with the increase of the energy scale; the behavior is shown in figure 5. At low momentum transfer, the strong coupling is large. Because of the large $\alpha_S(Q^2)$ value, all quarks that are gluons are bounded or *confined* into hadrons. The phenomena is knows as *color confinement*. On the other hand, at high momentum transfer Q^2, the effective coupling between gluons and quarks becomes small; therefore quarks and gluons can behave as free particles. This phenomena is called *asymptotic freedom*.
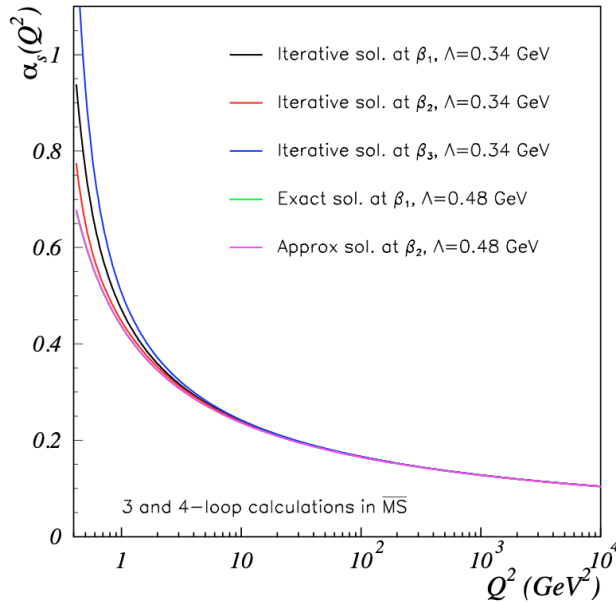
Figure 5: The perturbative coupling $\alpha_S(Q^2)$ computed for different orders of parameters. Figure from [16].

2.1.6 Challenges of the Standard Model

The SM is the most successful theory in particle physics so far. However, it does not explain some phenomena observed in the experiment. The list of the problems includes, but is not limited to, neutrino oscillations, dark matter, hierarchy problem, combination of general relativity and the SM, gravity and the SM, the excess of matter over antimatter, and strong and electroweak forces unification. The dark matter and hierarchy problems in the SM are discussed below.

Dark Matter.

Visible matter accounts for approximately 17% of the total matter in the Universe. The remaining 83% of matter is called dark matter. Dark matter hasn't been observed directly; however, its existence is supported by many side observations.

One of the observations supporting the existence of dark matter is galaxies rotation. The arms of spiral galaxies rotate around their galactic centers. The visible mass density decreases as the distance from the center of the galaxy increases. If visible mass was the only mass in the galaxy, then the rotation speed would decrease with the distance from the center of the galaxy. However, it is observed that the rotation velocity doesn't depend on the distance from the center of the galaxy. That can be explained by the presence of "invisible" matter.

Observation of galaxies rotation, as well as many others (galaxy clusters, gravitational lensing, cosmic microwave background and so on), supports the hypothesis of the existence of dark matter. The SM of particle physics doesn't explain dark matter. One of the solutions to the problem could be new physics as an additional symmetry to the SM with yet-to-be-discovered particles.

Hierarchy Problem.

At the loop level, the Higgs mass receives correction from self interactions, gauge bosons, and fermion loops, as shown in figure 6. The Yukawa coupling $(-\lambda_f \bar{f} H f)$ makes a mass correction to the Higgs boson mass as:

$$\Delta m_H^2 = -\frac{|\lambda_F|^2}{8\pi^2}\Lambda_{UV}^2 + ... \, , \tag{2.1.29}$$

where Λ_{UV} is the ultraviolet momentum cutoff scale. If $\Lambda \sim M_P = 2.4 \times 10^{18}$ GeV, then the quantum correction is much larger than the Higgs mass itself [17]. The issue is also know as quadratic divergence: the Higgs mass is quadratically sensitive to any mass scale of new physics.
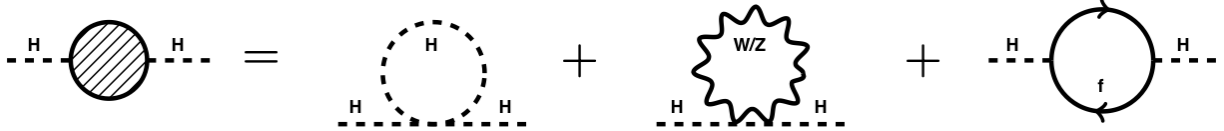
Figure 6: Components of the Higgs mass correction. From left to right: self-interaction, gauge bosons loop, fermion loop.

One way to solve the problem is to introduce new physics with a heavy scalar particle S that interacts with the Higgs in a four-point interaction. Mathematically, such an interaction is represented by $-\lambda_S |H|^2 |S|^2$. A corresponding Feynman diagram is shown in figure 7. Such a scalar field results in a Higgs mass correction of:

$$\Delta m_H^2 = \frac{\lambda_S}{16\pi^2} [\Lambda_{UV}^2 - 2m_S^2 \ln(\Lambda_{UV}/m_S) + (\text{finite})] . \tag{2.1.30}$$
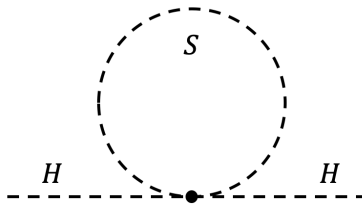
Figure 7: One-loop quantum correction to the Higgs squared mass due to a hypothetical scalar field S with interaction $-\lambda_S |H|^2 |S|^2$.

Taking into account the opposite signs between contributions to Δm_H^2 from the fermion loops and the boson loops, it follows that the proposed solution can be implemented via a new symmetry between fermions and bosons. If the proposed complex scalar field is associated with each of the SM quarks and leptons with $\lambda_S = |\lambda_f|^2$, then the quadratic divergence in the Higgs mass correction will be canceled.

2.2 Supersymmetry in Particle Physics

In QFT, Supersymmetry (SUSY) [18, 19, 20, 21, 22, 23] is a generalization of space-time symmetries. Supersymmetry is not a QFT as is, but an extension to actual theories. SUSY applied to the SM predicts new bosonic partners to the SM fermions and vise versa. To implement the new symmetry in the SM, so-called supercharge operator Q is used. When applied to the SM particles, the operator generates supersymmetric partners as shown below:

$$Q|\text{fermion}\rangle_{SM} = |\text{boson}\rangle_{SUSY}$$
$$Q|\text{boson}\rangle_{SM} = |\text{fermion}\rangle_{SUSY} . \tag{2.2.1}$$

The supersymmetric partners of the SM particles are discussed in section 2.2.1. A realization of SUSY in the SM with the least number of new particles is presented in section 2.2.2. Since no SUSY particles have been observed yet, their masses must be larger than the corresponding SM masses. Therefore, if SUSY exists in nature, it must be a broken symmetry. Some of the breaking mechanisms are presented in section 2.2.3.

2.2.1 Supersymmetric Partners and Supermultiplets

Supersymmetric partners to the SM particles are called superpartners. The single-particle states of supersymmetric theory fall into irreducible representations of the supersymmetry algebra, called *supermultiplets* [24]. Each of the supermultiplets contain both fermion and boson states, which are known as *supermultiplets* of each other. The number of bosonic n_B and fermionic n_F degrees of freedom in a supermultiplet must be equal. There are two types of supermultiplets: *chiral* (or *matter*, or *scalar*) and *gauge* (or *vector*) supermultiplets.

Chiral supermultiplets are the simplest supermultiplets. Each chiral supermultiplet consists of a single Weyl (massless) fermion with two helicity states ($n_F = 2$) and two spin-0 real scalar bosons (n_B=2). The bosonic superpartners of SM fermions are called "sparticles", and denoted by the tilde on top of the designator. For example, the supersummetric partner of the SM top quark t is the stop quark $\tilde{t}$. Since the SM Higgs boson has spin 0, it must be in a chiral supermultiplet. All chiral multiplets for the Minimal Supersummetric Standard Model are summarized in table 2.

Names		spin-0	spin-$\frac{1}{2}$	$SU_C(3)$, $SU_L(2)$, $U_Y(1)$
squarks, quarks, (3 generations)	Q	$(\tilde{u}_L,\ \tilde{d}_L)$	$(u_L,\ d_L)$	$(3,\ 2,\ \frac{1}{6})$
	$\bar{u}$	$\tilde{u}_R^*$	$u_R^\dagger$	$(\bar{3},\ 1,\ \frac{2}{3})$
	$\bar{d}$	$\tilde{d}_R^*$	$d_R^\dagger$	$(\bar{3},\ 1,\ \frac{1}{3})$
sleptons, leptons, (3 generations)	L	$(\tilde{\nu},\ \tilde{e}_L)$	$(\nu,\ e_L)$	$(1,\ 2,\ -\frac{1}{2})$
	$\bar{e}$	$\tilde{e}_R^*$	$e_R^\dagger$	$(1,\ 1,\ 1)$
Higgs, higgsinos	H_u	$(H_u^+,\ H_u^0)$	$(\tilde{H}_u^+,\ \tilde{H}_u^0)$	$(1,\ 2,\ +\frac{1}{2})$
	H_d	$(H_d^0,\ H_d^-)$	$(\tilde{H}_d^0,\ \tilde{H}_d^-)$	$(1,\ 2,\ -\frac{1}{2})$

Table 2: Chiral supermultiplets in the Minimal Supersymmetric Standard Model. The spin-0 fields are complex scalars, and the spin-1/2 fileds are left-handed two-component Weyl fermions. Source: [24].

The other simple type of a supermultiplet contains a spin-1 vector boson, and is called gauge supermultiplet. If the gauge symmetry is not spontaneously broken, the spin-1 boson must be massless. A spin-1 boson has two helicity states ($n_B = 2$). Therefore, its superpartner is a Weyl fermion, which has two helicity states as well ($n_F = 2$). Using spin-3/2 fermions leads to non-renormalizable theory; thus the option cannot be considered. Gauge bosons must transform as the adjoint representation of the gauge group; therefore their fermionic partners must do the same [24]. The fermionic partners of the gauge bosons are called gauginoes, and similar to the sparticles, are denoted by the tilde. For example, superpartner of the SM gluon g is called *gluino* $\tilde{g}$. The gaugino fermions must have the same gauge transformation properties for the left-handed and for the right-handed components,

because the adjoint representation of a gauge group is always its own conjugate [24]. Such combinations with spin-1 gauge bosons and spin-1/2 gauginos is called gauge supermultiplet.

With gravity included with the SM, the spin-2 graviton with two helicity states ($n_B = 2$) has a spin-3/2 superpartner gravitino. If SUSY were not broken, the gravitino would be massless and have 2 helicity states.

2.2.2 Minimal Supersymmetric Standard Model

The Minimal Supersymmetric Standard Model (MSSM) is an extension to the SM that realizes SUSY with the minimum number of new particle states and interactions. The superpotential for the MSSM is given by formula:

$$W_{MSSM} = \bar{\tilde{u}} \vec{y}_u \tilde{Q} H_u = \bar{\tilde{d}} \vec{y}_d \tilde{Q} H_d - \bar{\tilde{e}} \vec{y}_e \tilde{L} H_d + \mu H_u H_d \ , \tag{2.2.2}$$

where H_u, H_d, $\tilde{Q}$, $\bar{\tilde{u}}$, $\bar{\tilde{d}}$, $\bar{\tilde{e}}$ are the chiral superfields corresponding to the chiral supermultiplets from table 2; $\vec{y}_u$, $\vec{y}_d$ and $\vec{y}_e$ are the dimensionless Yukawa coupling parameters that are 3×3 matrices in family space; μ is the SUSY version of the Higgs boson mass in the SM. From equation 2.2.2 it can be seen that both H_u and H_d are needed on order to give Yukava couplings (and thus masses) to all fermions. Examples of supersymmetric interactions proportional to the Yukawa couplings are shown in figure 8.
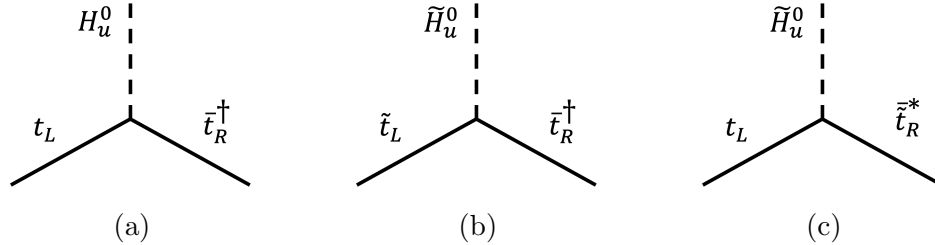
Figure 8: The top quark Yukawa coupling (a) and its "suppersymmetrizations" (b) and (c), all of stenght y_t.

In order to have massive superparticles, SUSY-breaking terms must be added to the MSSM Lagrangian. The general soft-breaking Lagrangian includes mass terms for both the scalar fields and gauginos:

$$\begin{aligned}
\mathcal{L}_{MSSM}^{soft} = & -\frac{1}{2}(M_3 \tilde{g}\tilde{g} + M_2 \tilde{W}\tilde{W} + M_1 \tilde{B}\tilde{B} + c.c) \\
& - (\bar{\tilde{u}}\vec{a}_u H_u - \bar{\tilde{d}}\vec{a}_d \tilde{Q} H_d - \bar{\tilde{e}}\vec{a}_e \tilde{L} H_d + c.c) \\
& - \tilde{Q}^\dagger \vec{m}_{\tilde{Q}}^2 \tilde{Q} - \tilde{L}^\dagger \vec{m}_{\tilde{L}}^2 \tilde{L} - \bar{\tilde{u}}\vec{m}_{\tilde{u}}^2 \bar{\tilde{u}}^\dagger - \bar{\tilde{d}}\vec{m}_{\tilde{d}}^2 \bar{\tilde{d}}^\dagger - \bar{\tilde{e}}\vec{m}_{\tilde{e}}^2 \bar{\tilde{e}}^\dagger \\
& - m_{H_u}^2 H_u^* H_u - m_{H_d}^2 H_d^* H_d - (b H_u H_d + c.c.) \ .
\end{aligned} \tag{2.2.3}$$

where

- M_3, M_2 and M_1 are the gluino, wino and bino mass terms.

- scalar couplings a_u, a_d and a_e are 3×3 matrices in one-to-one correspondence with the Yukawa couplings, making it possible for Higgs-sfermion-sfermion interactions.

- sfermions mass terms $\vec{m}_{\tilde{Q}}^2$, $\vec{m}_{\tilde{L}}^2$, $\vec{m}_{\tilde{u}}$, $\vec{m}_{\tilde{d}}^2$ and $\vec{m}_{\tilde{e}}^2$ are 3×3 matrices in family space.

- $m_{H_u}^2$, $m_{H_d}^2$ and b are squared mass terms; these are SUSY-breaking contributions to the Higgs potential.

Overall, the Lagrangian in equation 2.2.3 introduces many new parameters: 105 masses, phases and mixing angles [25]. That makes SUSY difficult to explore.

Unlike in the SM, the baryon B and lepton L numbers are not conserved in the MSSM. Conservation of the baryon and lepton numbers in the SM follows from the absence of any possible renormalizable Lagrangian terms that would violate B or L. Furthermore, there is no proof that B and L are fundamental symmetries of nature. Therefore, B and L numbers conservations cannot be assumed in the MSSM. To restrict B and L to be conserved, another symmetry is added to the MSSM, called "R-parity". Mathematically, the symmetry eliminates the possibility for B and L violating terms in the renormalizable superpotential. R-parity for a particle is defined as:

$$P_R = (-1)^{3(B-L)+2s} , \qquad (2.2.4)$$

where B is the baryon number, L is the lepton number, and s is the spin number of a particle. As it can be seen from the definition above, B and L numbers are not conserved independently in MSSM, but rather their combination with s is conserved.

R-parity values for the SM particles are $+1$, while for the superpartners it is -1. Some phenomenological consequences follow from the R-parity conservation, in particular:

- There must be an even number of SUSY particles in an interaction vertex. In other words, a SUSY particle cannot be produced from only SM particles (and cannot decay into only SM particles).

- There is the so-called lightest supersymmetric particle (LSP) that must be stable. The LSP is a neutral SUSY particle that interacts only weakly with the matter. The LSP is a candidate for dark matter. All SUSY particles undergo decay chains ending up with an odd number of LSPs.

- MSSM without R-parity allows for the proton to decay. However, R-parity conservation forbids the process.

2.2.3   Supersymmetry Braking Mechanisms

In order to have massive superpartners of SM particles, the MSSM must be broken. However, there are some challenges in creating a phenomenological framework with would yield renormalizable interactions in broken SUSY. One problem is associated with the gluino mass. There is no $scalar-gaugino-gaugino$ coupling which would assign mass to the gluino in SUSY. The other problem is related to the sfermions' masses: at least some of them should

be relatively light and would have already been observed in an experiment. For these reasons, it is assumed that SUSY breaking in MSSM occurs in a "hidden sector" of particles that have very small or no direct coupling to the "visible sector" chiral supermultiplets of the MSSM [24]. However, some interactions that are responsible for mediating SUSY breaking between the two sectors are still possible. The idea of the two sectors is sketched in figure 9.
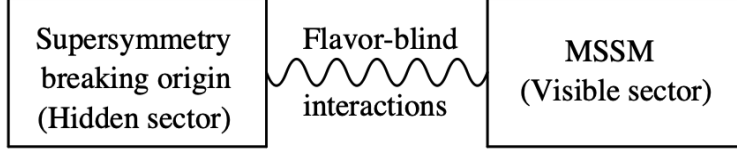
Figure 9: The presumed schematic structure for supersymmtry breaking. Figure from [24].

Two main proposals for what the mediating interactions could be are discussed below.

Plank-scale-mediated SUSY breaking (PSMB).

The first mechanism proposes that the interactions are due to new physics, including gravity, that enter near the Plank scale. In the PSMB scenario, if SUSY is broken in the hidden sector by a vacuum expectation value $\langle F \rangle$, the the so-called soft-terms in the visible sector are given by:

$$m_{soft} = \langle F \rangle / M_P \qquad (M_P = 2.4^{18} \text{ GeV}) , \qquad (2.2.5)$$

from the equation above it follows that m_{soft} vanishes when $\langle F \rangle \to 0$ where SUSY is unbroken, or when $M_P \to \infty$ where gravity is negligible. For $m_{soft} \sim$ a few hundred GeV, the scale of the origin of SUSY breaking in the hidden sector is roughly $\sqrt{\langle F \rangle} \sim 10^{10} \div 10^{11}$ GeV. After writing the SUSY-breaking Lagrangian 2.2.3 and minimizing the form of the normalization of gauge interactions and kinetic energy terms, the soft-terms of $\mathcal{L}_{soft}^{MSSM}$ are determined by the following four parameters:

$$m_{1/2} = f \frac{\langle F \rangle}{M_P} ,$$
$$m_0^2 = (k + n^2) \frac{|\langle F \rangle|^2}{M_P^2} ,$$
$$A_0 = (\alpha + 3n) \frac{\langle R \rangle}{M_P} , \qquad (2.2.6)$$
$$B_0 = (\beta + 2n) \frac{\langle F \rangle}{M_P} ,$$

where the soft-terms with the reference to the $\mathcal{L}_{soft}^{MSSM}$ Lagrangian from equation 2.2.3 are:

- $m_{1/2} = M_1, M_2, M_3$ is the gaugino mass parameters,

- $m_0^2 \vec{1} = \vec{m}_{\tilde{Q}}^2 = \vec{m}_{\tilde{L}}^2 = \vec{m}_{\tilde{u}}^2 = \vec{m}_{\tilde{d}}^2 = \vec{m}_{\tilde{e}}^2$ are the scalar masees,

- A_0 is the supersymmetrization of the Yukawa coupling: $\vec{a}_u = A_0 \vec{y}_u$, $\vec{a}_d = A_0 \vec{y}_d$ and $\vec{a}_e = A_0 \vec{y}_e$,

18

- B_0 defines the Higgs mass parameter: $b = B_0\mu$.

The system described above is also referred as the *minimal supergravity* (MSUGRA) or *Constrained Minimal Supersymmetric Standard Model* (CMSSM) scenario for the soft terms [24]. Based on the soft terms in equation 2.2.6, the entire MSSM spectrum in MSUGRA can be predicted by a set of just five parameters: $m_{1/2}$, m_0^2, A_0, B_0 and μ. In all the SUSY-breaking models, the Higgsino $\tilde{H}$, wino $\tilde{W}^{\pm}$ and zino $\tilde{Z}$ are combined to form neutralinos χ_m^0 and charginos $\chi_m^{\pm1}$, where m is a consecutive integer mass number in the order of increasing of the actuall mass. In MSUGRA, the LSP is electrically neutral.

Gauge-mediated SUSY-breaking (GMSB).

The second possibility for flavor-blind mediating interactions for SUSY breaking is that they are ordinary electroweak and QCD gauge interactions. In GMSB, a new set of chiral supermultiplets, called *messengers*, provides the necessary connection between the visible and hidden sectors. The messengers interact with both MSSM and the source of the SUSY-breaking vacuum expectation value $\langle F \rangle$ via $\mathrm{SU}_C(3) \times \mathrm{SU}_L(2) \times \mathrm{U}_Y(1)$ gauge bosons. Then the soft-term for the visible sector should be as:

$$m_{soft} \sim \frac{\alpha_a}{4\pi} \frac{1}{\langle F \rangle} M_{mess} \tag{2.2.7}$$

where $\alpha/4\pi$ is a loop factor involving gauge interactions, and M_{mess} is a characteristic scale of the masses of the messenger fields. In GMSB, the gravitino is the LSP.

CHAPTER III

The ATLAS Expetiment at LHC

This chapter introduces the Large Hadron Collider (LHC) [5] and the ATLAS experiment [6] at the European Center for Nuclear Researches − CERN (French: *Conseil Européen pour la Recherche Nucléaire*). Section 3.1 briefly describes the accelerator complex and introduces major experiments at LHC, as well as some quantitative parameters of the accelerator. Section 3.2 presents the ATLAS experiment with an eponymous detector at LHC. The section introduces each subsystem of the detector from the innermost to the outermost and discusses physics program at ATLAS.

3.1 The Large Hadron Collider at CERN

CERN is the European research organization that hosts and operates the world's largest and most powerful particle accelerator − the LHC. The CERN scientific community consists of about $2,700$ scientists, technicians, and administrative personnel working for CERN itself, and about $12,400$ users from institutions from more than 70 countries. The CERN laboratory, as well as the LHC, is located on the French-Swiss border near Geneva.

3.1.1 The Accelerator Complex

The LHC is a circular accelerator 27 km in circumference located in a tunnel from 50 to 175 meters underground. The tunnel was previously used by the Large Electron-Positron Collider (LEP) [26] which was running from 1989 until 2000 and was decommissioned in 2001. The LHC accelerates two bunched beams of protons in the opposite directions to up to 99.99999991% the speed of light.

The LHC consists of 1232 15 m long dipole magnets, which bend beams forcing them to follow a circular path, and 392 $5 - 7$ m long quadruple magnets, which squeeze beams in vertical or horizontal directions depending on the magnets system orientation in space, preventing them from free Coulomb repelling. Particularly-strong quadruple magnets are located close to the intersection points to focus the beams to increase the collision rate per bunch crossing. The dipole magnets are made of a niobium-titanium alloy and are cooled by superfluid helium down to 1.9 K [27]. The dipole magnets can provide a magnetic field of more than 8 T. To accelerate particles, 16 400 MHz radio-frequency (RF) cavities are used, each of which produce an accelerating field of 5 MeV/m.

Before entering the LHC, proton beams undergo a few stages of acceleration by a chain of linear and circular particle accelerators. A schematic representation of the accelerator complex is presented in figure 10, and the main components are:

- Protons are created by ionizing hydrogen gas and are accelerated to a kinetic energy of 50 MeV by the linear accelerator Linac2 [28]. The accelerator was decommissioned and replaced by Linac4 [29] in 2020. Linac4 accelerates negative hydrogen ions H^- to kinetic energy of 160 MeV.

- After been accelerated by Linac2, protons are injected into the Proton Synchrotron Booster (PSB). The accelerator is made up of four superimposed synchrotron rings, and accelerats the proton beams from Linac2 to 1.4 GeV. In the updated system, H^- ions are stripped from their two electrons during their injection into the PSB. The resulting protons are accelerated to 2 GeV.

- From the PSB, the protons are injected into the Proton Synchrotron (PS). The PS is the first synchrotron built at CERN. The accelerator is 628 m in circumference and consists of 100 dipoles. It accelerates protons to a kinetic energy of up to 26 GeV.

- The next step in the acceleration chain for the protons is the Super Proton Synchrotron (SPS) [30]. The SPS accelerates protons to kinetic energies of up to 450 GeV. The accelerator consists of 744 dipoles and is 7 km in diameter. Back in 1983, CERN announced the discovery of the W and Z bosons in proton-antiproton collisions at SPS [31]. In 1984, Carlo Rubbia and Simon Van der Meer were awarded with the Nobel Prize for the discovery.

- Lastly, the accelerated proton beams are injected from SPS into two beam pipes of the LHC. The protons in the beam pipes are accelerated in opposite directions and collide at four points along the LHC's circumference. It takes 4 minutes and 20 seconds to produce and deliver proton beams to the LHC, and up to 20 minutes to accelerate the beams in the LHC to their maximum energy of 6.5 TeV.
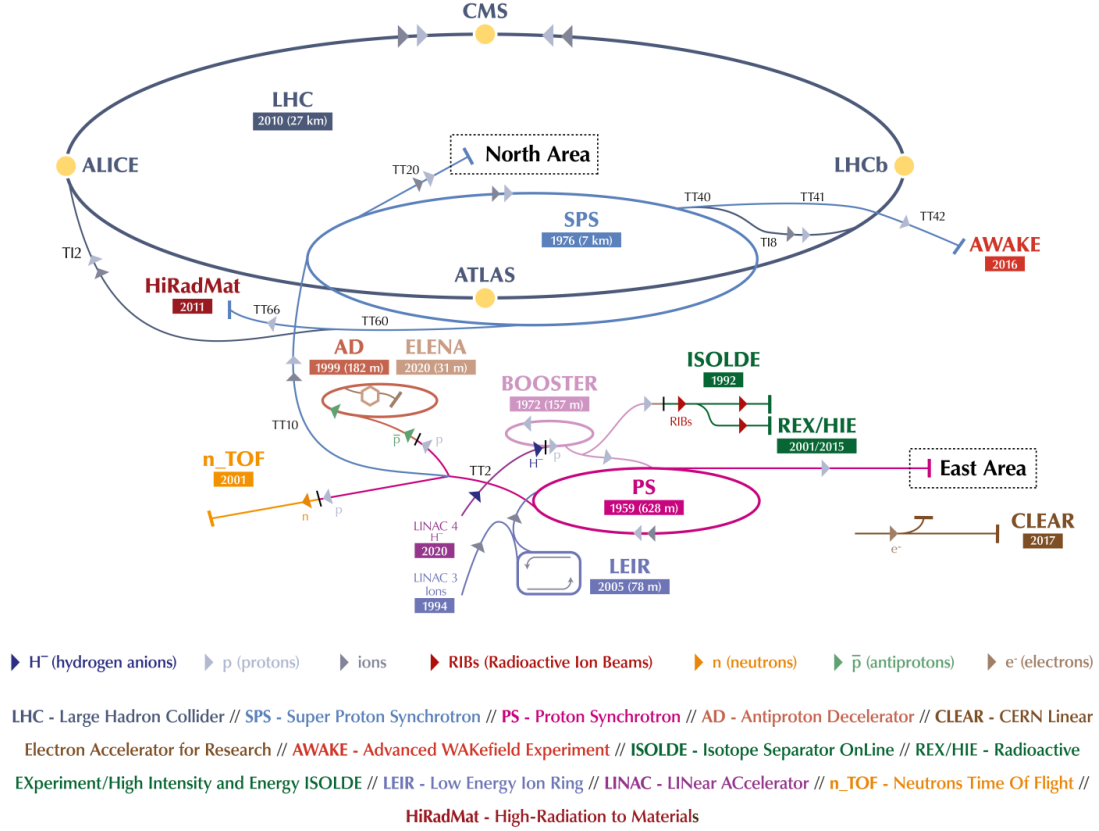
Figure 10: Schematic representation of the CERN acceleration complex. Figure from [32].

3.1.2 LHC Operational Parameters

The most important operational parameters of the LHC are the center-of-mass energy of the colliding protons $\sqrt{s}$ and luminosity $\mathcal{L}$. The LHC is designed to accelerate proton beams to kinetic energies of up to 7 TeV, resulting in a center-of-mass energy of $\sqrt{s} = 14$ TeV. However, the maximal center-of-mass energy hasn't been achieved yet.

There are two measures of luminosity at the LHC: instantaneous luminosity and total integrated luminosity. The instantaneous luminosity is defined as [33]:

$$\mathcal{L} = \frac{f N_b n_1 n_2}{4\pi \sigma_x \sigma_y} \ , \tag{3.1.1}$$

where $f = 11.2455$ kHz is the revolution frequency, N_b is the number of bunches in each beam, n_1 and n_2 are the number of protons in each of the colliding beams, σ_x and σ_y are the transverse beam sizes at the interaction point.

The integrated luminosity is the time integral of the instantaneous luminosity:

$$\mathcal{L}_{int} = \int \mathcal{L} dt \ . \tag{3.1.2}$$

Knowing the integrated luminosity and the production cross-section for a specific process, one can estimate the expected number of events produced at the LHC:

$$N_{events}^{process} = \mathcal{L}^{int} \sigma^{process} \ . \tag{3.1.3}$$

The LHC accelerates and collides bunches of protons rather than continuous beams. The accelerator hosts 2544 bunches of protons with 25 ns bunch spacing [33]. On average, there is more than one pp interaction per bunch crossing happening at the LHC. To describe the interactions multiplicity per bunch crossing, a term *pileup* is introduced. There are two types of pileup events:

- In-time pileup corresponds to signals from pp interactions within the same bunch crossing.

- Out-of-time pileup defines signals from particles resulting from pp interaction from different, usually two consecutive, bunch crossings.

The measure of the pileup, $\langle \mu \rangle$, is defined as the average minimum bias interactions per bunch crossing. The mean number of interactions per bunch crossing equals to:

$$\mu = \mathcal{L} \sigma_{inel} f n_b \tag{3.1.4}$$

where $\mathcal{L}$ is the instantaneous luminosity, σ_{inel} is the inelastic collision cross-section, f is the frequency at which the bunches collide, and n_b is the number of bunches per beam. An example of $\langle \mu \rangle$ distribution measured by the ATLAS detector during the 2015-2018 data taking period is given in figure 11.
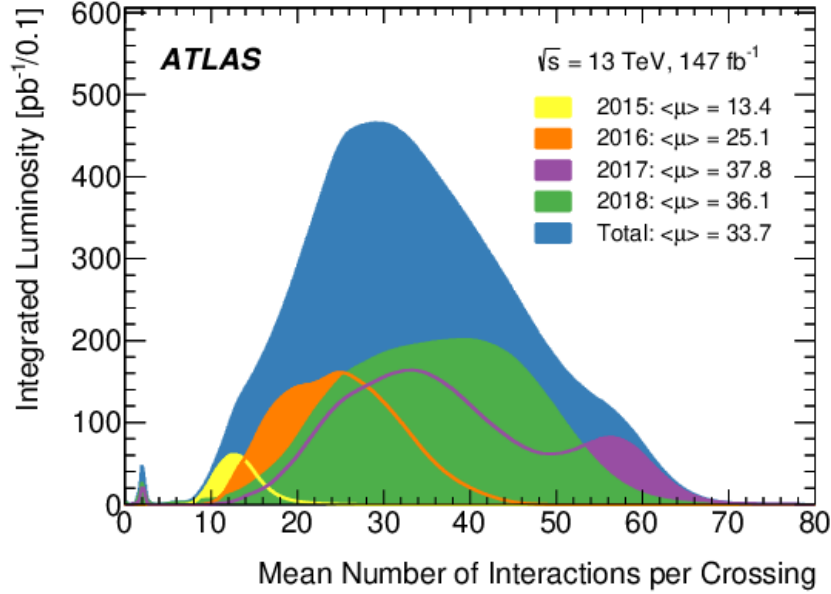
Figure 11: Pileup events distribution measured by the ATLAS detector during 2015-2018 data taking period. Figure from [34].

3.1.3 Experiments at LHC

The following detectors are build at each of the four collision points of the LHC:

- **ATLAS** - The Toroidal LHC Apparatus is one of two general-purpose detectors at the LHC. The detector and scientific program of the collaboration are presented in section 3.2.

- **CMS** [7] − The Compact Muon Solenoid experiment at the LHC. CMS is the second general-purpose detector at the LHC. As for ATLAS, its goal is to measure the properties of the Standard Model with high precision and to search for new particles beyond the Standard Model.

- **ALICE** [35] − A Large Ion Collider Experiment. The ALICE detector is designed to study quark-gluon plasma originating from the collisions of heavy ions. The plasma created in $p-Pb$ or $Pb-Pb$ ions collisions at the LHC is $100,000$ hotter than the core of the Sun.

- **LHCb** [36] − The Large Hadron Collider beauty. The LHCb experiment is specialized in b-physics. Its area of focus is the measurements of the charge parity (CP) violation in interactions of b-hadrons.

In addition, there are four more detector experiments built around the main ones:

- **TOTEM** [37] − TOTal Elastic and diffractive cross section Measurement. The detector is built in the CMS cavity, and the experiment aims for the measurements of the total cross-section, elastic scattering, and diffraction process of pp collisions.

24

- **LHCf** [38] − Large Hadron Collider Forward. The LHCf detector is located in the forward direction of each side of the ATLAS detector. The experiment aims for the studies of high-energy QCD processes to improve the Monte Carlo simulations' modeling of cosmic rays showers.

- **MoEDAL** [39] − Monopole and Exotics Detector at the LHC. The detector shares the cavity with the LHCb. The experiment aims for the searches of magnetic monopoles and highly ionizing stable particles.

- **FASER** [40] − ForwArd Search ExpeRiment. FASER is a new experiment at CERN that was approved in 2019 and will start collecting data during the LHC Run3. The detector is installed in the way of an intense and highly collimated beam of neutrinos, 480 m downstream from the ATLAS detector, in a tunnel that was used to inject beams from the Super Proton Synchrotron to the LEP accelerator. The FASER experiment aims for the studies of the interactions of high-energy neutrinos and searches for new light and weakly interacting particles. There is about 100 of rock and concrete shielding between the detector and the collision point, which should provide a low background environment.

3.1.4 LHC Lifetime and High Luminosity LHC Upgrade

One of the limiting factors in high-energy physics searches for new physics and precision measurements is the amount of collected data. As discussed above, the luminosity is a measure of the expected collisions rate (instantaneous) and the total expected number of collisions (integrated). The higher the luminosity is, the larger the collected number of potentially interesting events is, including rare ones. The other limiting factor is the total energy of the colliding particles, i.e. the center-of-mass energy $\sqrt{s}$. Since no new particles with masses of the order of SM ones have been found, the potential candidates must be heavier. If these particles exist, more energy must be supplied to pp interactions to probe regions with higher energies per particle. Therefore, experimentalists strive for a machine that collides particles at as much as possible rate with the highest energies possible, as well as for a detector that would be able to capture and process the results from the collisions.

To satisfy the need for higher luminosity and center-of-mass energy, the LHC, as well as the detector experiments, have long-term plans for upgrades of the facilities. Figure 12 presents the timeline of the LHC campaigns starting with its first run, going to the current status, and the future plans for runs and upgrades.
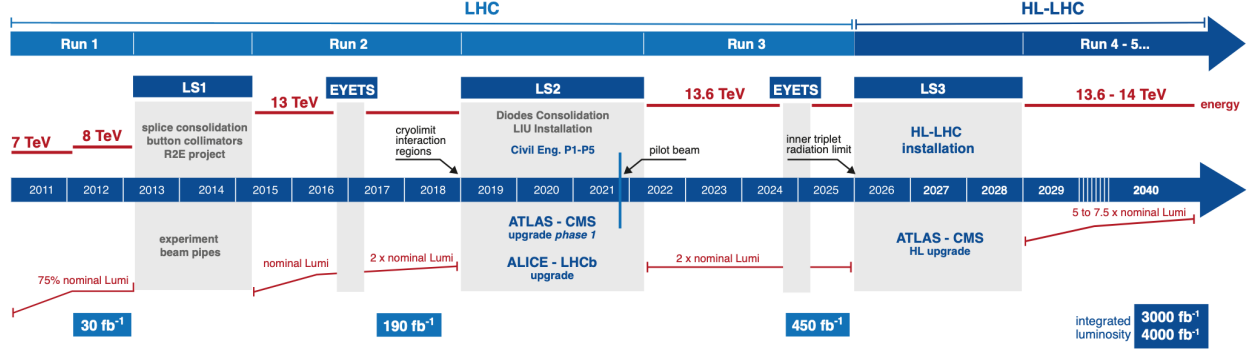
Figure 12: the LHC timeline with major milestones. Figure from [41].

The LHC started working back in 2008. It was tested for some period of time, and the actual scientific program started on March 30th, 2010, 1:06pm CEST, when two beams collided at $\sqrt{s} = 7$ TeV. In 2011, the accelerator was delivering pp collisions at $\sqrt{s} = 7$ TeV, ramping up its instantaneous luminosity to 75% of the nominal level. In 2012, the center-of-mass energy increased to 8 TeV. The 2011-2012 campaign is called Run1 [42]. The accelerator was shut down on February 14th, 2013 for an upgrade; this upgrade period is called the Long Shutdown 1 (LS1). During Run1, the LHC delivered $\mathcal{L} = 30$ fb^{-1} of pp collisions.

On June 3rd, 2015, after almost two years of offline recommisioning, the accelerator started delivering pp collisions with the record-breaking $\sqrt{s} = 13$ TeV. On June 29th, 2016, the accelerator achieved its design value of instantaneous luminosity. On May 24th, 2017, the accelerator achieved twice the value of its designed instantaneous luminosity. On December 10th, 2018, the accelerator was shut down for the Long Shutdown 2 (LS2), which concluded Run2 [43] campaign. During Run2, the LHC delivered $\mathcal{L} = 150$ fb^{-1} of pp collisions. During the LS2, the LHC was upgraded to deliver pp collisions at even higher $\sqrt{s}$, and the four major detectors were upgraded as well.

As of February 2022, the LHC is planned to restart in June 2022. The accelerator is expected to deliver pp collisions at $\sqrt{s} = 13.6$ TeV at twice the nominal instantaneous luminosity. This new data-taking period, named Run3 [44], will last until 2026. It is expected that the LHC will deliver up to 260 fb^{-1} of the integrated luminosity during Run3.

After Run3, the accelerator and the four major detectors will shut down for the third Long Shutdown (LS3). The period is entitled as the "High Luminosity LHC Upgrade". After the upgrade, the High Luminosity LHC (HL-LHC) [45] is expected to deliver pp collisions at $5 - 7.5$ instantaneous luminosity of the original LHC design value and at the center-of-mass energy of up to $\sqrt{s} = 14$ TeV $-$ the maximal achievable value at the LHC. To withhold the increased data rate and radiation doses (especially for the inner layers), all detectors at the LHC will undergo major upgrades.

3.2 The ATLAS Experiment

ATLAS [46] is the general-purpose particle physics experiment at the LHC at CERN [47]. The collaboration consists of over 5500 members and almost 3000 scientific authors from 183 institutions in 38 counties around the world. The ATLAS collaboration is searching for new physics beyond the Standard Model, as well as running precise measurements of the Standard Model itself using the ATLAS detector. In 2012, the ATLAS collaboration along with the CMS collaboration at CERN announced [8, 9] the discovery of the Higgs bosons, the last piece of the standard model.

The ATLAS detector is built around one of the collision points at the LHC. It is the largest-ever volume detector constructed for particle physics experiments. It has the shape of a cylinder 46 m long and 25 m in diameter and weights approximately 7000 tonnes. The cut-away view of the ATLAS detector if shown in figure 13. The detector has a layered structure, and the four main subdetector systems are: the Inner Detector (ID), the Electromagnetic Calorimeter (ECal), the Hadronic Calorimeter (HCal), and the Muon tracking system. There are also two magnetic systems: a solenoid magnet in the inner part of the detector, and toroid magnets in the mid-area.
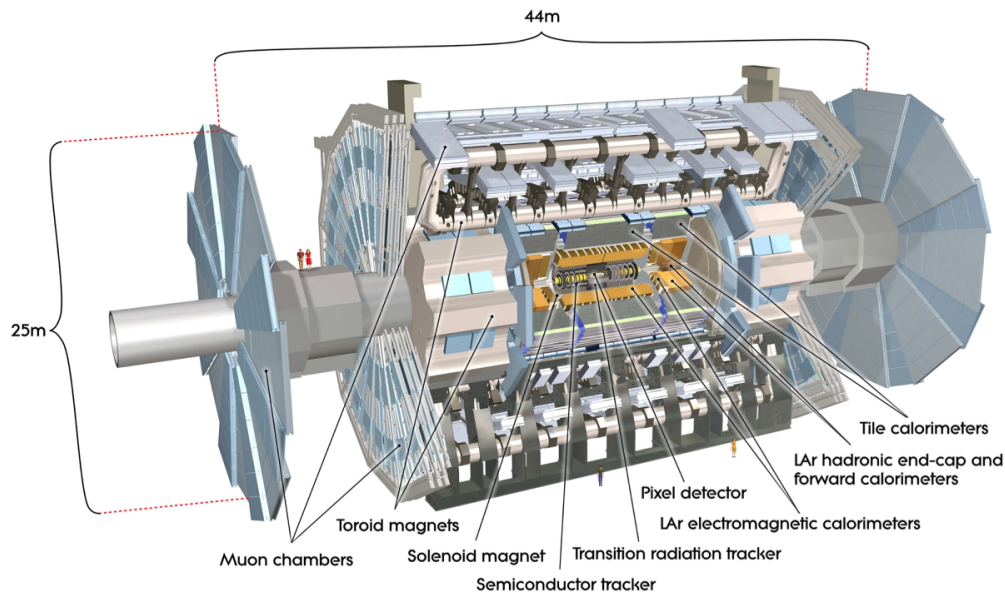
Figure 13: Cut-away view of the ATLAS detector. Figure from [46]. The original length of the detector from 2008 is 44 m; after a series of upgrades, the length increased to 46 m.

A cut-view diagram of the ATLAS detector showing the interactions of different particles with the detector layers is shown in figure 14. Particles produced in pp interactions at the LHC travel outwards from the collision point and interact with certain subdetector systems, depending on the particles' nature.
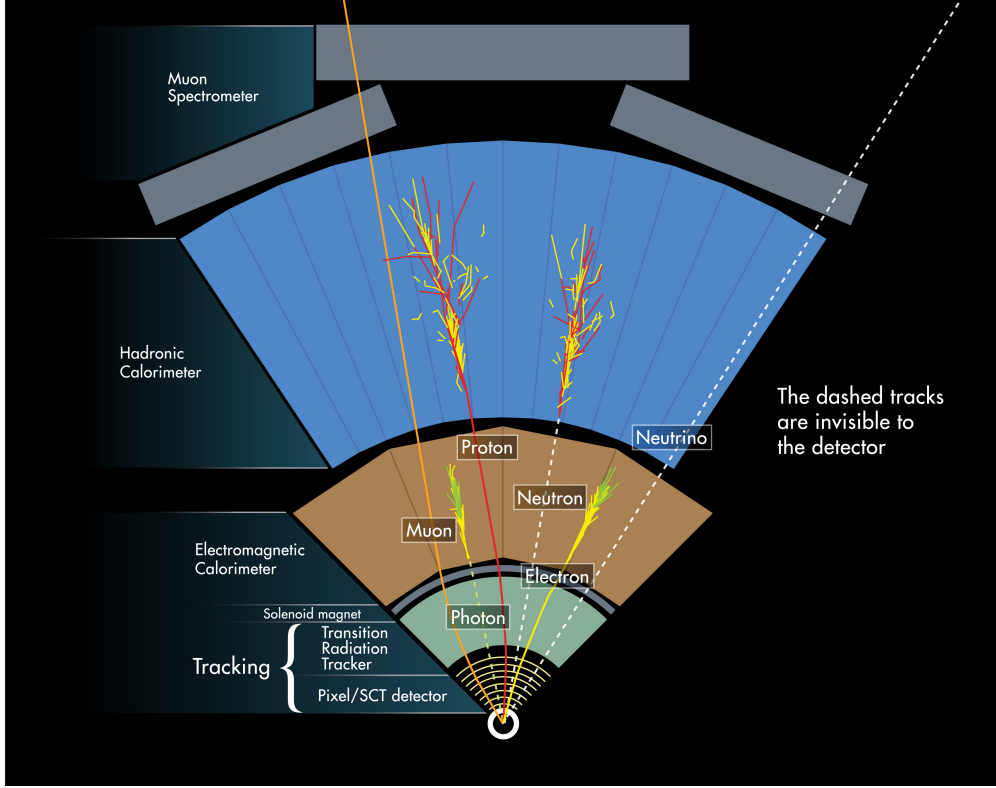
Figure 14: Cut-away diagram of the ATLAS detector showing interactions of different particles with the detector's subsystems . Figure from [48].

3.2.1 ATLAS Coordinate System

The ATLAS detector uses a right-hand coordinate system with its origin set at the nominal interaction point. The beam direction defines the z coordinate with the $x - y$ plane being the transverse plane to the beam axis. The positive z direction points toward side A of the ATLAS detector, the positive x direction points to the center of the LHC accelerator ring, and the positive y points upward. The coordinate system is schematically depicted in figure 15. In the cylindrical coordinate system, $r = \sqrt{x^2 + y^2}$ and ϕ is the azimuth angle measured around the z axis starting from 0 along the positive x axis. In the spherical coordinate system, θ is defined as the polar angle from the beam axis, and it is zero along the positive z axis.

In most of the cases, pseudorapidity η is used instead of θ and defined as:

$$\eta = -\ln(\tan(\frac{\theta}{2})) \ , \tag{3.2.1}$$

$\eta(\theta = 90°) = 0$ corresponding to the region perpendicular to the beam axis (i.e. in the transverse $x - y$ plane), and $\eta(\theta = 90°) = \infty$. Pseudorapidity is an approximation of rapidity y in case of a massless particle:

$$y = \frac{1}{2}\ln[\frac{E + p_z}{E - p_z}] \ , \tag{3.2.2}$$

28

where E and p_z are particle's total energy and momentum along the z axis. Using η is favorable over using θ, because $\Delta\eta$ between two particle is invariant under the Lorentz transformations.
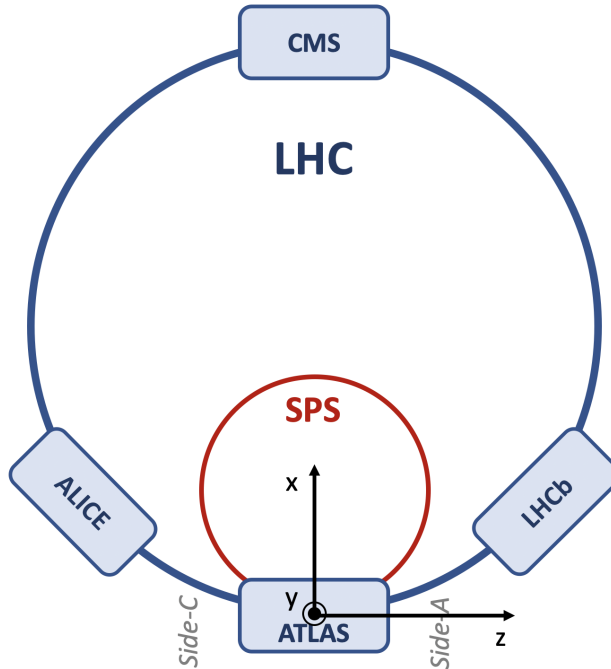
Figure 15: Coordinate system of the ATLAS detector

In the pp collisions, partons (quarks and gluons) constituting the protons interact, and the longitudinal momentum of each parton involved in an interaction is unknown. For this reason, most of the kinematic variables used in ATLAS analyses are computed in the $x - y$ plane; for example, the transverse momentum $p_T = \sqrt{p_x^2 + p_y^2}$, or the missing transverse energy E_T^{miss} (defined is section 5.7).

3.2.2 The Inner Detector and The Magnet System

The Inner Detector (ID) in ATLAS [49, 50] is the closest system to the beam pipe. It is designed to measure the trajectory and momenta of charged particles originating from the pp interactions as well as the positions of the pp collisions. The ID consists of three subsystems: the pixel detector, the semiconductor tracker and the transition radiation tracker. The entire system is surrounded by a superconducting solenoid magnet. The ID is designed to reconstruct particles in the $|\eta| < 2.5$ range. A schematic cut-view of the ID is shown in figure 16, and a 2D scheme of all layers of the ID in the $y - z$ plane in shown in figure 17.
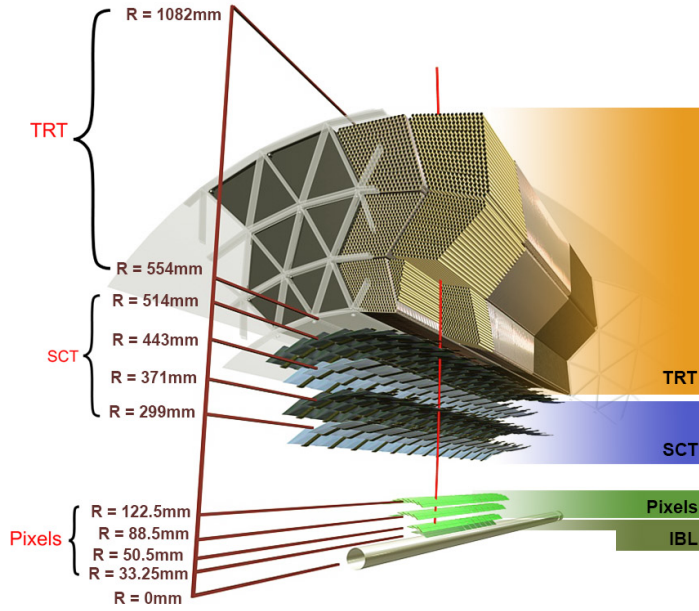
Figure 16: ATLAS Inner Detector layers structure from the innermost to the outermost: the pixel detector, the semiconductor tracker and the transition radiation tracker. The Insertable B-Layer, installed in May 2014, is also shown in the layout. Figure from [51].
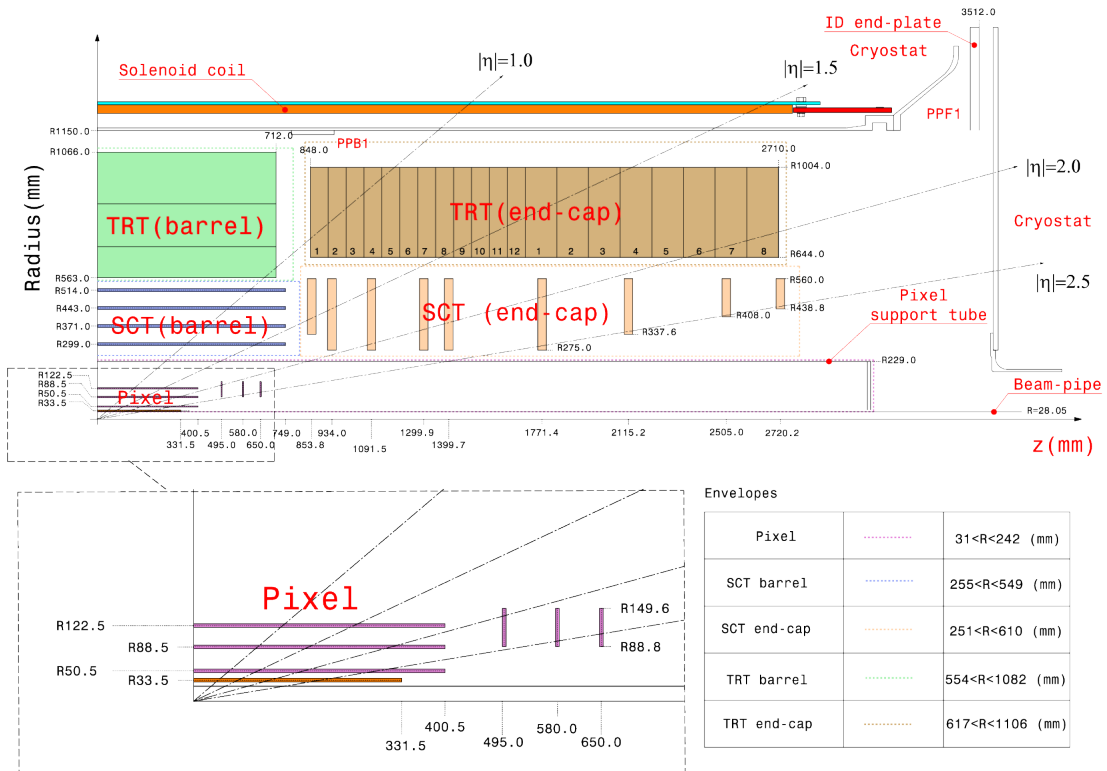
Figure 17: Layers structure of the ATLAS ID for the Run2. Figure from [52].

The Superconducting Solenoid Magnet

The superconducting solenoid magnet generates 2 T magnetic field [53] which bends trajectories of the charged particles. The degree and the direction of the curvature is used to measure particles' charges and momenta. The magnet is 5.8 m long with inner and outer diameters of 2.46 m and 2.56 m respectively, and it weights approximately 5 tonnes. The magnet has a single-wound coil. The superconducting wire is made of Al-stabilized NbTi alloy, and conducts 7.73 kA current.

The Pixel Detector

The Pixel Detector [54] is the closest subdetector to the beam pipe. It consists of a barrel and two end-cap regions. In the barrel region, the detector is made of three concentric cylindrical layers of pixel modules, while the two end-cap regions consist of three disks with modules. The detector covers $|\eta| < 2.5$ region. Each pixel module hosts silicon-based sensor arrays and front-end and readout electronics. One pixel of the silicon sensors represents a reverse-biased diode. A charged particle passing through the sensors generates charge in the pixel; the charge is drained by the electronics and forms measurable current.

Each pixel sensor consists of $41,984$ pixels of $r\phi \times z\ = 50 \times 400\ \mu m^2$ size and $5,284$ pixels of $50 \times 600\ \mu m^2$ size resulting into $47,268$ pixels in total [55]. However, the number of readout channels per sensor is $46,080$. The extra pixels are needed to cover gaps in the structure due to the actual layout of the readout modules. A fraction of the pixels in the edge region are paired such that two pixels are connected to one readout channel, resulting in two-fold ambiguity on the 5% of a sensor surface which is resolved offline. The intrinsic spatial resolution of a pixel is of the order of 10 μm in $r - \phi$ and 115 μm in z coordinates [51].

During the LS1, an additional layer of pixel sensors, named Insertable B-Layer (IBL) [56], was installed; it covers $|\eta| < 3$ region. The IBL added 12 million readout channels to the ID, resulting in a total number of 80 million channels. The IBL significantly improved the tagging of jets originating from b-quarks.

The Semiconductor Tracker (SCT)

The SCT [57] is the second ID subdetector system. Similar to the Pixel Detector, 4088 SCT sensors are arranged into four concentric barrel layers and 9 end-cap disks. The system covers $|\eta| < 2.5$ region. A SCT sensor chip is made of p-type doped silicon on n-type bulk material [58]. The average pitch of the strips is 80 μm. One SCT module consists of 4 silicon chips, and for most of the modules two chips are connected to form a daisy-chain. To achieve high-precision tracking, two sensors are glued together back-to-back at a stereo angle of 40 mrad. The sensors in the barrel region are rectangular in shape with strips aligned to the direction of the magnetic field, while in the end-cap region the sensors located on the disks are trapezoidal in shape and the strips point in the radial direction. The stereo angle and the geometry of the sensors allow the SCT to measure z coordinates in the barrel region and r coordinates in the end-cap. The intrinsic resolution of SCT is 17 μm in the $r - \phi$ plane

and 580 μm in the z and r in the barrel and the end-cap regions respectively.

The Transition Radiation Tracker (TRT)

The TRT [59] is the outermost part of the ID, and it covers $|\eta| < 2$ region. The system is designed to provide tracking information in the $r - \phi$ plane in the barrel and in the $z - \phi$ plane in the end-cap regions. The TRT consists of 298,304 drift tubes [60] 4 mm in diameter each. In the barrel region, the tubes are oriented in the beam direction and form a 520 mm thick layer. In the end-cup region, the tubes are aligned perpendicular to the beam, and mounted on 20 wheels with 8 layers of tubes each.

The walls of the tubes are coated with aluminum on the inside to form a cathode, and the anode is a gold-plated tungsten wire running in the middle of the tube. In Run1, the tubes were filled with a xenon-based gas mixture consisting of 70% Xe, 27% CO_2 and 3% O_2. However, some gas leakage was detected, and the gas was replaced with a 70% Ar, 27% CO_2 and 3% O_2 mixture [61]. The tubes work as a drift chamber. A particle passing through a tube ionizes the gas, the ions are collected by the anode wire, and the total collected charge is measured. With the known properties of the drift tubes, the TRT position measurements accuracy was estimated to be 130 μm for charged particles with $p_T > 0.5$ GeV [60].

The TRT also has particle identification capabilities. The space between the tubes is filled with foil in the end-cap and polymer fiber sheets in the barrel regions. Charged particles passing through the materials emit soft X-rays which are detected by the tubes. The intensity of the X-rays depends on the γ-factor of a particle, and therefore the measurements of the radiation make it possible to differentiate high γ-factor electrons from pions.

3.2.3   Electromagnetic and Hadronic Calorimeters

The ATLAS calorimeters system [6] surrounds the superconducting solenoid magnet and covers $|\eta| < 4.9$ region. It consists of two subdetectors: the inner electromagnetic calorimeter (ECal) [62] and the outer hadronic calorimeter (HCal). The primary goal of the system is to measure energy of particles originating from the pp interactions. In addition, the system shields the muon spectrometer from the majority of particles other than muons. The cut-view of the calorimeters is shown in figure 18.

The calorimeters measure particles' energy by stopping them and absorbing showers caused by electromagnetic or hadronic interactions with the absorbers. Both the ECal and the HCal are sampling calorimeters meaning that the energy absorbers layers are alternated with the active material which measures the deposited energy.
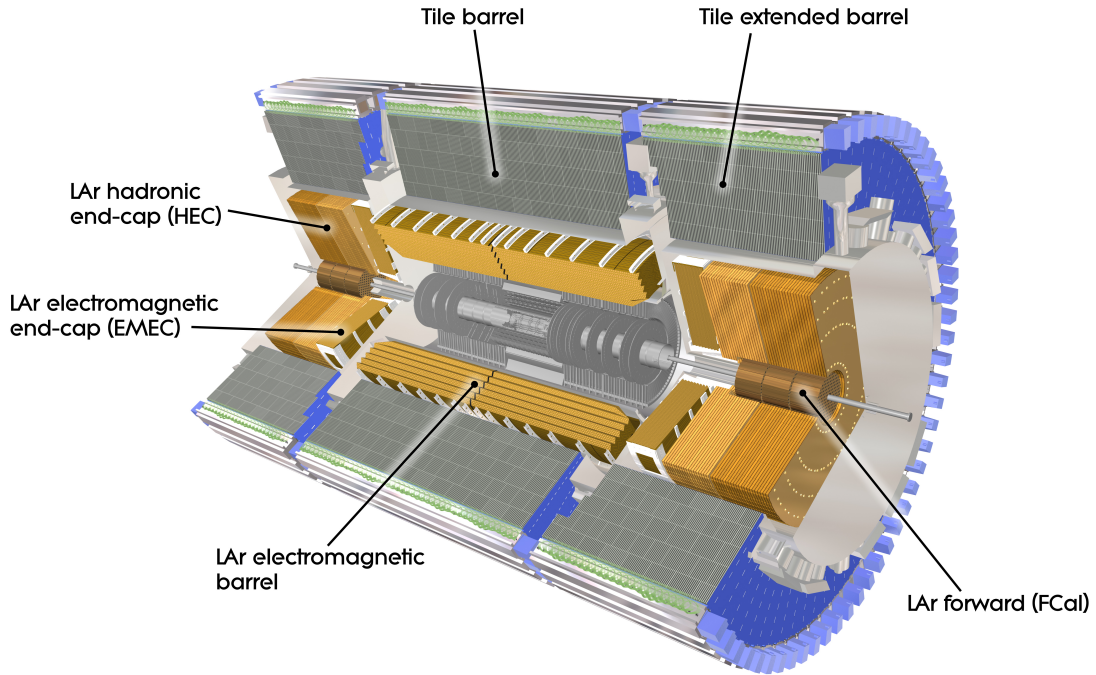
Figure 18: Cut-view of the ATLAS detector calorimeters. The calorimeters surround the superconduction solenoid magned with the ECal being the inner and HCal being the outer structures. Figure from [6].

One of the most important parameters of calorimeters is thickness. The thicker the calorimeter, the more energetic particles it can stop, and thus measure their total energy. Quantitatively it can be described with the radiation length X_0 - the distance which a charged particle travels through the material with atomic number Z and atomic mass A to lose $1/e$ of its energy:

$$X_0(g/cm^2) = 716.4 \ g \ cm^{-2} \frac{A}{Z(Z+1)\ln(287/\sqrt{Z})} \ . \tag{3.2.3}$$

The ECal thickness in the central barrel region increases from $22X_0$ ($|\eta| = 0$) to $30X_0$ ($|\eta| = 0.8$) and in the extensions from $24X_0$ ($|\eta| = 0.8$) to $33X_0$ ($|\eta| = 1.3$). In the end-caps, its thickness goes from $24X_0$ ($|\eta| = 1.475$) to $38X_0$ ($|\eta| = 2.5$) and from $26X_0$ ($|\eta| = 2.5$) to $36X_0$ ($|\eta| = 3.2$).

Hadron calorimeters thickness is described by the nuclear (or hadronic) interaction length: $\lambda \approx 35 \ g \ cm^{-2} A^{1/3}$ - the distance a hadron travels through a medium to lose $1/e$ of its energy.

Electromagnetic Calorimeter

The ECal is made of three systems: the EM barrel (EMB), which covers $|\eta| < 1.475$ region, and two end-caps (EMEC), each of which covers $1.37 < |\eta| < 3.2$ regions on sides A

and C of the ATLAS detector. The gap between the structures, where electronics, cables, coolant and gas tubing are placed, is ofter referred as the transition or crack region. Particles detection is reduced in the region, and because of that ATLAS physics analyses tend not to use objects falling within the transition region.

Both the EMB and the EMEC use liquid argon (LAr) as the active material and lead plates as the absorbers. High energy charged particles passing through the absorber layers of the ECal lose energy via bremsstrahlung [10] and produce cascades of secondary low energy particles. The secondary particles ionize liquid argon in the active layers. High voltage is applied to the LAr regions by copper cathodes, which allows for drift current of ions and free electrons. The drift current is measured and used to evaluate the energy of the initial particle.

The structure of the LAr ECal is shown in figure 19. The system consists of three layers: the innermost strip cells layer with the most precision in η, the middle square cells layer, where the largest fraction of the energy is deposited, and the outermost trigger tower, which collects the tails of the EM showers. Granularities of the layers are $\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$, 0.025×0.0245 and 0.1×0.0982 respectively.
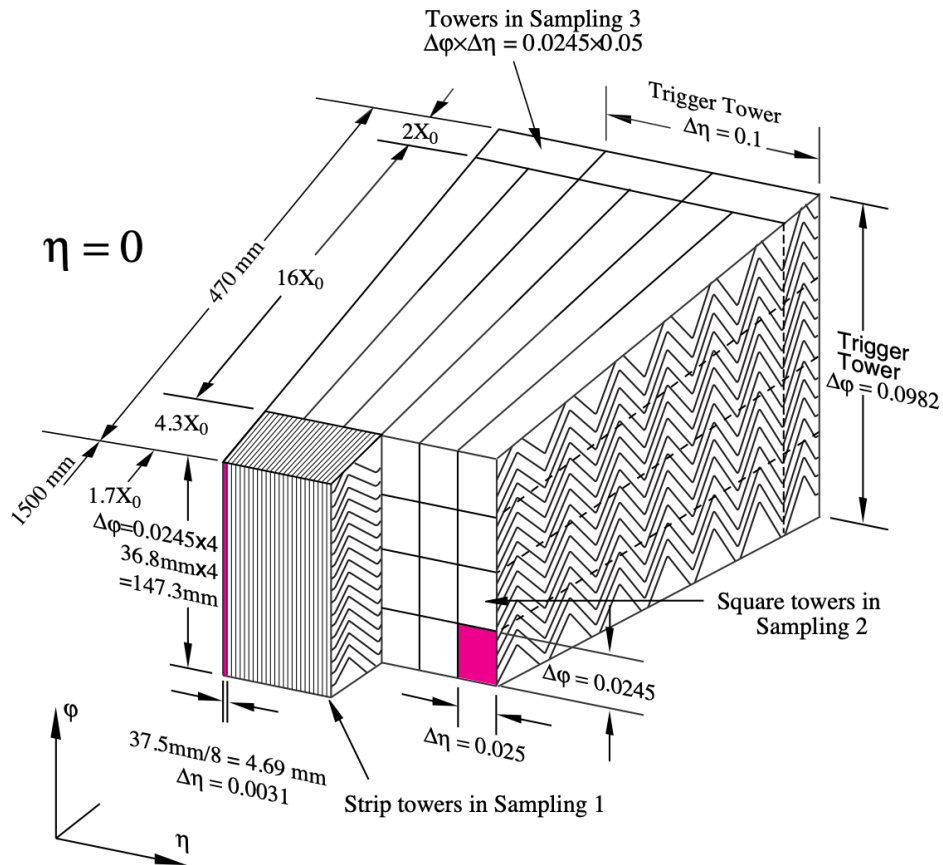
Figure 19: Structure view of the LAr ECal. The detector consist of three regions with resolutions: $\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$ in the inner, $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$ in the middle, and $\Delta\eta \times \Delta\phi = 0.1 \times 0.0982$ in the outer. Figure from [62].

Hadronic Calorimeter (HCal)

The HCal is made of three subsystems: the tile calorimeter [63], the hadronic end-cap (HEC) calorimeters [64], and the forward calorimeter (FCal) [65]. The system is designed to measure energies of hadrons.

The tile calorimeter is located in the barrel region right outside of the ECal. It covers $|\eta| < 1.0$ region with two extended barrels covering $0.8 < |\eta| < 1.7$ regions from both sides. The tile calorimeter uses steel as the absorber with plastic scintillators being the active medium. A hight-energy hadron passing through the tile calorimeter interacts with the absorber material and produces showers of lower-energy secondary particles. The secondary particles interact with the scintillators causing it to radiate light. The light is collected by fibers and transferred to photomultiplier tubes (PMT), and the PMTs convert the collected light into electric current. The tile calorimeter consists of three segments with different granularities. The first two layers have dimensions of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, and the third layer is $\Delta\eta \times \Delta\phi = 0.1 \times 0.2$. The layers are 1.5, 4.1 and 1.8 λ thick.

The HEC calorimeters are two pairs of detector wheels placed on each side of the barrel covering $1.5 < |\eta| < 3.2$ range. HEP calorimeters are sampling calorimeters made of copper absorbers and LAr active material. Granularity of the system is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $|\eta| < 2.5$ and 0.2×0.2 for regions with higher η.

The FCal [65] covers $3.1 < |\eta| < 4.9$ range. One FCal is installed on each side of the detector, and consists of three 45 cm long submodules, FCal1, FCal2 and FCal3, stacked in longitudinal direction. The FCal1 is an EM calorimeter. It is made of copper absorber and filled with LAr. The FCal2 and FCal3 are hadronic calorimeters made of tungsten absorber. Total thickness of the systems is approximately 10λ. The last, fourth, layer is a copper-alloy shield that minimizes punch-through into the muon spectrometer.

3.2.4 Muon Spectrometer

The Muons Spectrometer (MS) [66] is the outermost system of the ATLAS detector, which measures momentum of muons in $|\eta| < 2.7$ range. As the other detector systems, it consists of a barrel ($|\eta| < 1$) and end-caps ($1 \leq |\eta| < 2.7$). Each end-cap system consists of three wheels. Four different detector technologies are used in the MS: the Monitored Drift Tubes (MDT) and the Cathode Strip Chambers (CSC) for precision tracking, and the Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC) for fast muon event trigger calibrations. The MS is built inside of a system of superconducting magnets which bend trajectories of muons passing through the MS. The system of magnets consists of the barrel toroid magnet [67] and the end-cap toroids [68]. Both the barrel and the end-cap magnets produce the magnetic field of 4 T. The magnetic fields are roughly perpendicular to the trajectories of incoming muons, and bend the trajectories in the $r - z$ plane. A cut-view of the MS is shown in figure 20.
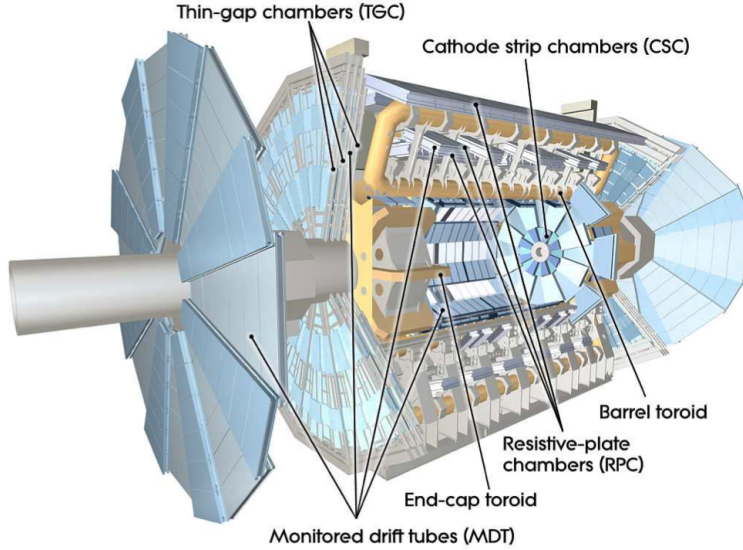
Figure 20: Structure view of the ATLAS muon system. The system consists of the barrel and end-cap detector regions, and barrel toroid and end-cap toroid magnets. The barrel detector region consists of concentric detector layers located inside, within and outside of the barrel magnet. The end-cap system consists of three wheels: the smaller one is inside of the barrel magnet, and the two bigger are outside of the magnet. Figure from [6].

MDT modules are used for precision tracking and precent in both the barrel and the end-cup. An MDT modules consist of $3-4$ layers of drift tubes $30\ mm$ in diameter each. MDT tubes work in a similar way to the TRT tubes. The tubes are filled with an Ar-CO$_2$ gas mixture in 93%/7% ratio, with a $50\ \mu m$ tungsten-ruthenium wire placed in the middle, and high voltage potential is applied to the gas filled space. Passing muons ionize the gas, change is collected on the wire and detected as an electric charge. Shape and duration of the pulse define drift radius which could be measured with $\sim 80\ \mu m$ precision. The MDT modules are placed inside, within and outside of the barrel magnet, and on the two big wheels.

As the MDT, the CSC modules are used for precision tracking measurements. Due to their higher radiation tolerance and higher count rate, they are placed closer to the beam on the small wheels. The CSC modules are multi-wire proportional chambers filled with an Ar-CO$_2$ gas mixture and have radial anode vires and cathode strips. Muons passing through a chamber cause charge avalanches that are collected by cathodes. Resolution of the modules is $\sim 40\ \mu m$ in the bending plane.

The ATLAS muon trigger is designed for identifying events with high momentum muons within bunch spacing time resolution (25 ns). Three layers of TGC modules are located behind the smaller end-cup wheel to cover $1.05 < |\eta| < 2.4$ region for triggering. TGCs are multi-wire proportional chambers filled with a 55%/45% CO$_2$+n-pentane gas mixture. In the barrel region. Three layers of the RPC modules are placed outside of the MDT to measure the ϕ (non-bending) coordinate of muons. As the CSC and TGC, the RPC are multi-wire proportional chambers, and they are filled with a tetrafluoroethane, isobutane, and sulfur hexafluoride gas mixture.

3.2.5 High Luminosity LHC Upgrade of the ATLAS Detector

The HL-LHC will deliver *pp* collisions at the centre-of-mass energy of $\sqrt{s} = 14$ TeV with up to $\mathcal{L} = 7.5 \times 10^{34}$ $cm^{-2}s^{-1}$ luminosity. All detectors at LHC will be upgraded to withhold higher radiation doses, to provide faster readout and improve granularity of their subsystems. Upgrades of the ATLAS detector subsystems [69] are discussed below.

Inner Tracker

The ATLAS ID will be replaced by a new all-silicon system called Inner Tracker (ITk). The ITk will improve momentum resolution for charged particles tracks and extend $|\eta|$ coverage from $|\eta| < 2.5$ to $|\eta| < 4.0$. The new tracker will consist of 5 concentric barrel layers of pixel modules [70] located closer the the beam pipe, and 4 layers of silicon strips sensors [71] located right after the pixels. The new pixel sensor pitch is 50×50 μm^2.

In order to decrease the material budget and allow for as many readout channels as possible, the ATLAS collaboration decided to adopt the serial powering scheme for readout modules. Such a scheme has never been utilized by large physics experiments, and various tests of the powering scheme as well as the readout performance are required. Chapter VIII presents some test of the serial powering scheme and the readout tests for the ITk.

Calorimeters

Both the ECal and tile hadron calorimeter will have new front-end and readout electronics [72, 73]. Also, the tile hadron calorimeter will get new power supplies and interface boards.

Muon Spectrormeter

To achieve higher trigger rates and longer latencies, a fraction of the MS frontend and on- and off-detector readout and trigger electronics will be replaced. To increase the trigger acceptance, maintain muon identification and reconstruction performance additional muon chambers will be installed [74].

Trigger and DAQ

ATLAS will use two-level TDAQ system. The first level is a hardware trigger which will receive signal at up 1 MHz rate, corresponding to $\mathcal{L} = 5 \times 10^{34}$ $cm^{-2}s^{-1}$, and process selected data to the second level, so-called "Event filter", which will output data at 10 kHz rate [75].

High Granularity Timing Detector

The ATLAS High Granularity Timing Detector (HGTD) is designed to precisely measure the timings of charged particles. The system will be installed in front of the LAr calorimeter to reduce background from pileup jets. HGTD will cover $2.4 < |\eta| < 4.0$ region. The expected timing resolution for minimum-ionizing particles is 30 ps [76].

CHAPTER IV

Triggering, Data and Monte Carlo Simulations at the ATLAS Detector

Most of the events produced in pp collisions have no scientific interest, and the total data rate is that large so it is impossible so save all events and clean the statistics off-line. The Trigger and Data Acquisition system (TDAQ) [77] of the ATLAS detector was specifically designed to ensure the optimal data-taking conditions and to select the most interesting collisions. The system combines online and offline levels analyses and records the selected events at approximately 1 kHz rate from 40 kHz of the incoming rate of pp collisions at LHC. The recorded data is further converted into formats that are easy to use in physics analyses. Monte Carlo simulation of pp collisions and response of the ATLAS detector are performed to study the real data in physics analyses. The chapter provides a brief introduction to the ATLAS TDAQ system, ATLAS data and ATLAS events simulation process.

4.1   The TDAQ system

The TDAQ system operated during the LHC Run1 (2009-2013) [78] at instantaneous luminosities of up to 8×10^{33} cm^{-2} s^{-1} and primary centre-of-mass energies of pp collisions of 7 and TeV and 8 TeV. In Run2 (2015-2018), both the instantaneous luminosity and the centre-of-mass energy of pp collisions increased to 10^{34} cm^{-2} s^{-1} and 13 TeV respectively, which led to significant increase of the production rate of interesting physics processes. The ATLAS TDAQ system managed to provide a sufficient level of data recording in both LHC Run1 [78] and Run2 [79]. Block diagram of the ATLAS TDAQ system is shown in figure 21. The system consists of Level-1 online hardware-based trigger and offline software-based High Level Trigger.
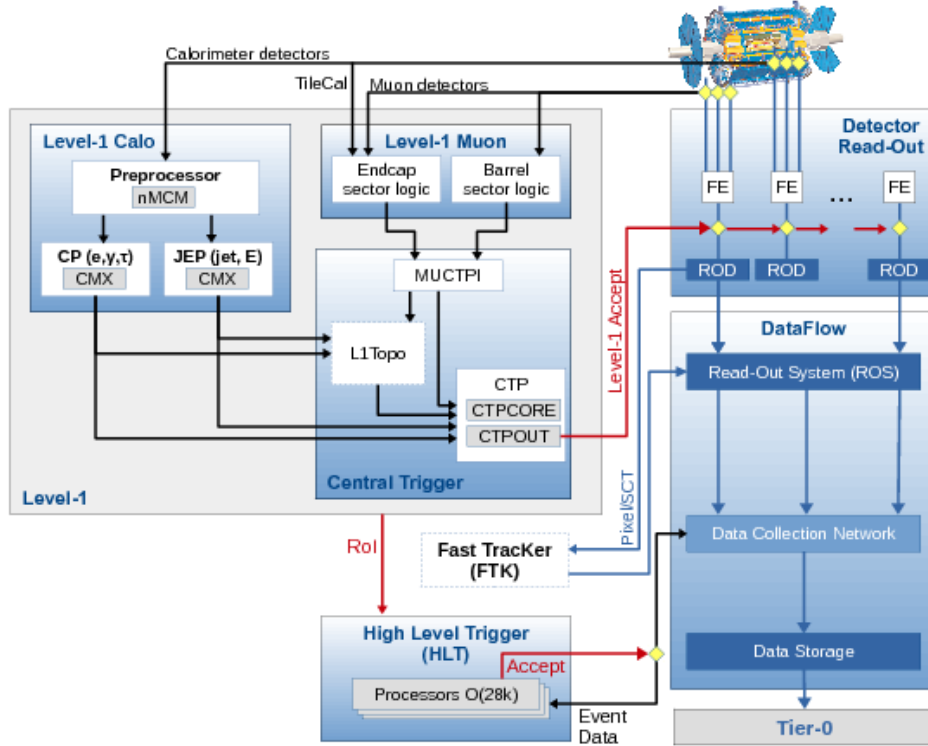
Figure 21: The ATLAS TDAW system in Run2 with emphasis in the components relevant for triggering. Figure from [79]

Level-1 Trigger (L1)

The L1 trigger is a hardware-based system that uses custom electronics to trigger on reduced granularity information from the calorimeter and muon detectors [80]. The trigger consists of three general components: calorimeter (L1Calo) trigger, muon (L1Muon) trigger and the Central Trigger Processor (CTP).

The L1Calo trigger searches for electromagnetic and hadronic energy deposits using the Cluster Processor (CP) and the Jet/Energy-sum Processor (JEP) respectively. In total there are 7000 trigger towers with $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ granularity.

The CP system is designed to identify electrons, photons and τ candidates with transverse momenta above certain programmable threshold. It uses a sliding window algorithm to identify 2×2 clusters with excess in the transverse momenta.

JEP system identifies jets candidates with transverse momenta above some threshold and produces global sums of total detected and missing transverse energy as well. Inputs to JEP are 2×2 electromagnetic and 2×2 hadronic trigger towers. As the CP system, it uses a sliding window algorithm and identifies 4×4 or 8×8 electromagnetic or hadronic clusters with transverse momenta excess above the threshold.

39

The L1Muon trigger is used to determine the deviation of the detected hits pattern in the muon system from that of a muon with infinite momentum [81]. The 1LMuon is also used to reduce the rate of particles not originating from the interaction point by applying consecutive requirements between the outer and inner thin-gap muon chamber stations and between the thin-gap chambers and the tile calorimeter.

The L1 trigger decision is done by the CTP which receives inputs from L1Calo and L1Muon triggers as well as from some other trigger systems from other detector subsystems, such as the LUCID Cherenkov counter [82], the Minimum Bias Trigger Scintillators (MBTS) [83] and the zero-degree calorimeter (ZDC) [84].

The L1 trigger selects events by considering event-level quantities, such as total energy in the calorimeter, the multiplicity of objects or topologies. The trigger accepts events at 40 kHz rate and transfer the selected events to the second level trigger at 100 kHz rate.

High Level Trigger (HLT)

HLT is a software-based system. It analyses events in offline mode using CPU-intensive algorithms. The output rate of the HLT is on average 1.2 kHz which corresponds to 1.2 GB/s data rate. Once an event is accepted by the LHT, the Sub-Farm Output (SFO) sends the data to permanent storage for offline reconstruction [77].

4.2   ATLAS Data

During Run2, the LHC delivered 156 fb^{-1} of pp collisions data at a centre-of-mass energy of 13 TeV. The ATLAS detector recorded 147 fb^{-1} of the data, 139 fb^{-1} of which is "good for physics". Figure 22a shows the delivered, collected and good data from pp collisions over the Run2 data taking period for the ATLAS detector. The ATLAS data is organized into *runs*. Each of the runs corresponds to a certain time interval of data collection, typically a few hours, and has its own identification number. The unit of granularity of the runs is called *lumi block* (LB). One lumi block corresponds to data collected over one minute. Data quality experts analyze the collected data and provide so-called Good Runs List (GRL) which consists of runs that are not affected by data collection problems and approved for using in physics analyses. Figure 22b shows the data quality recorded by the ATLAS detector during Run2 versus the integrated luminosity for each data taking year.
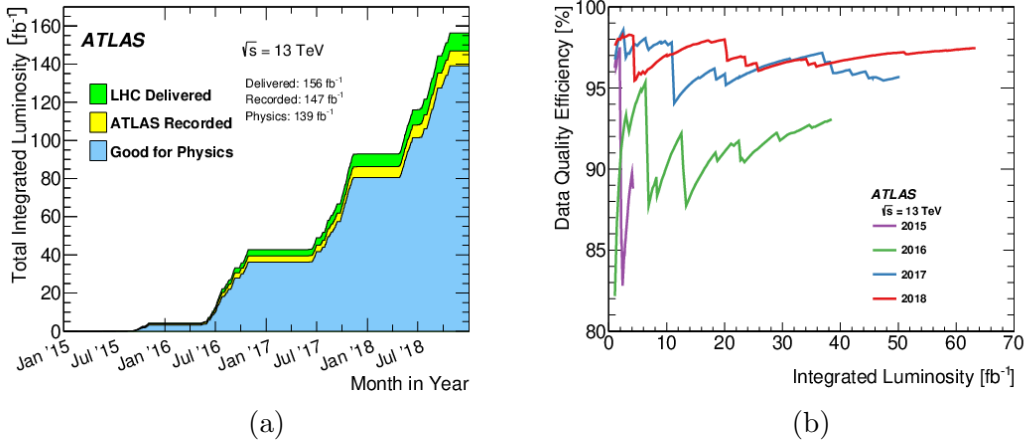
Figure 22: (a) The cumulative integrated luminosity delivered by the LHC and recorder by the ATLAS detector during Run2, and "good for physics" integrated luminosity. (b) Data quantity efficiency versus the integrated luminosity for each year in Run2. Figures from [34].

4.3   ATLAS Monte Carlo Simulations

Generation and simulation of events plays crucial role in physics analyses at ATLAS. It is practically impossible to distinguish between different physics processes with the same final state in data collected by the ATLAS detector. Therefore, for careful estimation of various backgrounds and signal processes in collected data, simulation of inelastic pp collisions at the LHC and the detector response are required. The simulation are done using Monte Carlo (MC) generators [85]. A schematic example of one generated event is shown in figure 23.

MC samples used in the analyses presented in chapters VI and VII consist of events originating from inelastic proton-proton collisions. To prepare MC samples for an analysis, three consecutive simulation process are performed: proton-proton interaction simulation, ATLAS event simulation and detector response simulation. The simulation results are followed by the reconstruction of physics objects which is discussed in chapter V.

Proton-proton interaction simulation

Protons are composite particles which consist of partons: three quarks (two *up* and one *down*), gluons, and so-called "sea of quarks" [86] - many qaurk-antiquark pairs of all flavors appearing from gluons. Therefore, pp collisions are if fact interactions of partons from incoming protons, and the patrons have different momentum distributions. One can estimate the production cross-section for a particular final state using parton distribution functions (PDF). PDFs are functions of the squared energy scale Q^2 of the pp interaction and of the fraction of the proton's energy x carried by the parton. There is no theoretical prediction of PDF as a function of x, therefore PDFs are derived experimentally from multiple measurements. However, one can estimate PDF as a function of Q^2 [87] using DGLAP (Dokshitzer [88], Gribov and Lipatov [89] , Altarelli-Parisi [90]) equations.

ATLAS event simulation

The simulation of ATLAS events goes in a few consecutive stages and managed by the Athena software framework [91]. The three simulation steps are matrix element calculation, parton shower and hadronization simulations. A schematic example of one generated event is shown in figure 23.
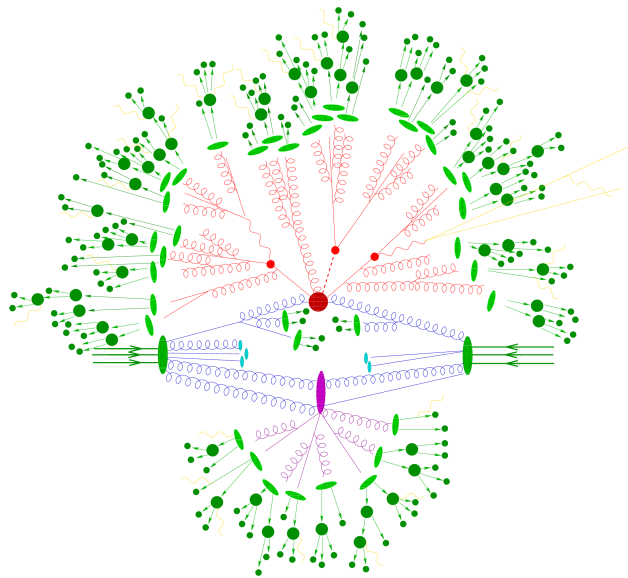
Figure 23: Sketch of an inelastic pp collision as simulated by a MC event generator. Two green blobs on the left and right of the central horizontal axis represent two incoming protons. The red and purple blobs represent inelastic interactions of gluons, the main and the underlying events respectively. Light green blobs represent hadronization of partons, and dark green blobs are hadrons decays. Yellow lines represent soft photon radiations. Figure from [92].

- **Matrix Elements $\mathcal{M}$ (ME)**, also known as amplitudes, are computed using Matrix Element Monte Carlo (MEMC) generators which produce the initial and final state particles of a specific hard scatter collision process. The $\mathcal{M}$'s are produced using the leading-order (LO), the next-to-the-leading-order (NLO) and higher order sets of Feynman diagrams. Each additional level of the xNLO Feynman diagrams introduced to the calculations changes the overall normalization of the cross-section and also may change kinematics. The higher the order of the Feynman diagrams, the more computational power it takes to produce $\mathcal{M}$. The most commonly used ME Monte Carlo generators at ATLAS are Pythia8 [93], Sherpa [94] and MadGraph [95]. While the first two generators are multipurpose, i.e. can be used for both the $\mathcal{M}$ and parton shower generation, MadGraph is only used to compute $\mathcal{M}$.

- **Parton Shower (PS)**. The energetic partons originating from pp collisions emit gluons which decay into quark-antiquark pairs $q\bar{q}$ pair or split into a pair of gluons, each of which can again produce a $q\bar{q}$ or split into two gluons, and so on. The process is named

42

parton showering. The probability for a parton to loose energy and evolve from scale q_1^2 to q_2^2 is estimated using the Sudakov form factors [96].

Soft interactions, such as Initial State Radiation (ISR), Final State Radiation (FSR) and Multi Parton Interaction (MPI) are computed and included at the PS step as well. These processes, often referred as the underlying event, describe additional lower energy radiation within the same pp collision as the main event.

- **Hadronization**. At the hadronization energy scale perturbation theory becomes invalid. Due to the effect of QCD color confinement, color-charged partons form color-singlet bound states. The resulting streams of hadrons are called jets. There are two hadronization models used in MC generators: the *string model* [97] and the *cluster model* [98]. Hadronization processes for the two models are sketched in figure 24 for the same jet.

The string model is based on the $q\bar{q}$ confinement. As the quark and antiquark move away from each other, the gluonic relativistic string between them is stretched, and the potential energy is increased. When the potential energy reaches the scale of hadron masses, the probability for the gluonic string to break increases significantly. Eventually the gluonic string breaks into a new $q\bar{q}$ pair. The original quark creates a bound state with the new antiquark, and so does the original antiquark with the new quark. The process repeats for the new $q\bar{q}$ pairs until the total energy in the system is converted into $q\bar{q}$ pairs. Lund string hadronization model [99] is implemented in Pythia [93] MC generator.

The cluster method is based on the QCD pre-confinement [100]. In the model gluons split into $q\bar{q}$ pairs. Color singlets combinations of $q\bar{q}$ pairs form clusters, the clusters isotropically decay into pairs of observed hadrons. The cluster hadronization model is implemented in Herwig++ [101] MC generator.

(a)   (b)
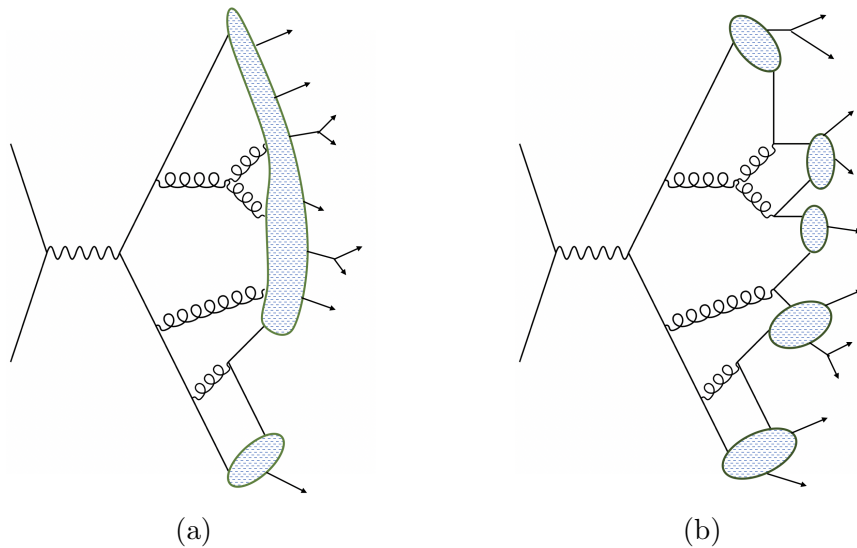
Figure 24: Visualization in the space-time view of (a) string (Pythia) and (b) cluster hadronization (Herwig) models for the same jet.

Detector simulation

The ATLAS detector is modeled in GeoTool [102], and the model is interfaced to the Geant4 [103] framework to simulate the detector response to the passing particles. The framework takes four-vectors of the final state particles simulated in the previous step as the input and models passage of the particles through the detector. Interactions between the particles and the detectors materials and fields are simulated resulting into energy deposits in the detector subsystems. The energy deposits are called *hits*. The hits are further used in the digitalization step where simulated detector output is produced. The output of MC generation chain is of the same format as the data captured by the ATLAS detector. Both the MC simulation and the data digitalized detector outputs undergo identical reconstruction procedure of physical objects, such as jets or leptons. The reconstruction procedure is discussed in the next chapter.

CHAPTER V

Objects Reconstruction and Identification at the ATLAS Detector

The inelastic *pp* collisions at LHC produce many particles, such as quarks, leptons and gauge bosons. Most of the particles are unstable and decay into more stable particles or hadronize before they reach of within the first layers of the ATLAS detector. The final stable particles interact with different layers of the detector depending on the particles' nature. Multiple algorithms are developed to reconstruct particles based on the energy deposits in calorimeters and hits in the inner detector and in the muon spectrometer of the ATLAS detector. The same algorithms applied to both the real data and MC simulated events. The chapter presents reconstruction algorithms for both the low-level (vertices, tracks, calorimeter clusters) and high-level (jets, leptons, missing transverse energy) physics objects.

5.1   Tracks

Charged-particles tracks are reconstructed from hits in the ID [104]. Reconstruction of the tracks is done in multiple steps.

Clusterization

The reconstruction starts with the clusterization procedure. A connected component analysis (CCA) [105] groups pixels and strips that share an edge or a corner in one module with energy deposits above the threshold level into clusters. The clusters are further used to create three-dimensional *space-points* which correspond the points where a charged particle traversed the ID. Clusters can be created by energy deposits from one charged particle of by a few charted particles. Therefore, a space-points due to multiple-particles clusters can be used to reconstruct multiple tracks.

Iterative combinatorial track finding

Next, track seeds are formed using combinations of three space-points. Trajectories are estimated from the seeds by assuming a perfect helical trajectory in a uniform magnetic field [105]. To increase the purity and to remove purely combinatorial tracks, a number of criteria is applied. First, requirements on the momentum and impact parameters are placed based on which sub-detector layer (SCT, pixels, or combined) recorded the space-points. Second, an additional space-point compatible with a particle's trajectory from the seed is required. Next, the combinatorial Kalman filter [106] is used to build track candidates using

space-points which are compatible with the preliminary trajectory from the seed. If there is more than one compatible extension for a trajectory from the seed, multiple track candidates can be created.

Track candidates and ambiguity solving

Clusters are allowed to be shared by two tracks at most. To resolve overlap of space-points between different tracks and to remove incorrectly assigned space-points, track candidates enter the ambiguity solving procedure. *Track score* is applied to each of the track candidates to determine the goodness of a track candidate. The score depends on multiple parameters, such as number of missing but expected clusters, trajectory fit quality and momentum. The score is used to resolve overlaps of space-points between multiple tracks favoring ones with the higher track scores.

Additionally, a neural network is used to minimize the loss due to the restriction in maximum of two tracks per cluster. The neural network is trained to find merged clusters, and uses three parameters: the measured charge, the relative position of pixels and the particles' incident angles provided from the track candidates.

Tracks candidates failing the ambiguity solver or with too many shared clusters are rejected.

Track fit

The last step of the tracks reconstruction procedure is high-resolution fitting. The fitting uses all available information, thus it is the most CPU-intensive process in the entire chain. For this reason fits are performed on the tracks candidates passing ambiguity resolving.

Physics quantities of tracks, such as charge and momentum, are calculated from the curvature of the fitted trajectories in the magnetic field of the detector.

5.2 Primary Vertices

Reconstruction of interaction vertices is important for distinguishing objects originating from different interactions. The primary vertex is the point in space where the inelastic pp collision happens. In ATLAS, reconstruction of primary vertices follow a two-step procedure [107]: vertex finding and vertex fitting [108].

In the first step, a vertex seed is placed in a point in space with transverse coordinates of the centre of the protons beam spot. Next, using the Half-Sample Mode algorithm [109], the z coordinate is calculated from the closest approach of the tracks to the center of the beam spot.

Once the vertex seed is determined, the transverse coordinate of it is computed using the iterative annealing procedure [110]. In order to find the most optimal position of the vertex

seed, the algorithm performs an iterative χ^2 minimization. All tracks initially associated with the vertex seed are assigned with weights which reflect compatibility of a track to the seed. The vertex seed position is computed using the tracks weights, and tracks with higher weights have higher impact on the vertex coordinate calculation. After each iteration, tracks less compatible with the vertex seed are weighed down, and the vertex seed transverse coordinate is recalculated. In the iterative process, the low-weighed tracks are not removed, but they don't have significant impact on the vertex coordinate calculation. After the final fit iteration, the incompatible to the vertex tracks are removed from the vertex.

The two-step procedure described above is then repeated to find a new vertex using the unused and rejected tracks after previously found vertices. The vertices finding stops when there is no unassociated tracks left in the event, or the remaining set of tracks is not sufficient to define a vertex.

5.3   Calorimeter Clusters

When interacting with the calorimeter material, particles from the pp collisions generate showers of secondary particles. The energy deposits due to streams of the secondary particles induce signals in the active components of calorimeters. The collection of neighboring calorimeter cells that measure the energy deposits compose calorimeter clusters. For each cell signal significance of the energy deposit is computed:

$$|\zeta_{cell}^{EM}| = \frac{|E_{cell}^{EM}|}{|\sigma_{noise\ cell}^{EM}|} \tag{5.3.1}$$

where $|\zeta_{cell}^{EM}|$ is the significance, $|E_{cell}^{EM}|$ and $|\sigma_{noise\ cell}^{EM}|$ are the deposited energy and the noice of a cell on the electromagnetic energy scale [111]. Cells with $|\zeta_{cell}^{EM}| > S$ are considered as the primary cells, and cells with $|\zeta_{cell}^{EM}| > N$ constitute the volume for growth of the cluster. Parameters S and N are, in fact, variable with the default values $S = 4$ and $N = 2$ for the ATLAS detector [112]. The cluster formation starts with the highest-$|\zeta_{cell}^{EM}|$ cell, and the adjacent cells with $|\zeta_{cell}^{EM}| > S$ are merged to the cluster seed.

Clusters due to independent showers can merge into one. A splitting algorithm is used to resolve the overlapping clusters. The algorithm analyses the local maxima within one cluster, and splits the cluster into two when needed. The resulting clusters are also referred as topoclusters. Set of clusters is used to reconstruct showers in calorimeters and jet, which are described in the next section.

5.4   Jets

Jets are defined as cone-shaped clusters created by particles depositing energy in the calorimeters. As discussed above, hadrons and leptons interact with the active material of the calorimeters and produce showers of secondary particles. Adjacent calorimeter cells measuring signal above threshold value are combined into clusters. The clusters are further combiner into jets. Once the jets are reconstructed, they must be calibrated.

5.4.1 Jet Reconstruction Algorithms

There are different jets reconstruction algorithms which could be classified as sequential recombination and cone jet algorithms. Examples of the sequential recombination algorithms are k_T [113] and Cambridge/Aachen [114]. One the most known cone jet algorithms is the [115, 116]. Another cone jet algorithm is the anti-k_T algorithm [117], that works in the opposite order to the k_T. Application of any of the algorithms is done it two stages:

- Define two distance measures: the cluster to beam distance d_{iB}, and the cluster to cluster distance d_{ij}:

$$d_{iB} = p_{T,i}^{2p} \, , \tag{5.4.1}$$

$$d_{ij} = min(p_{T,i}^{2p}, \ p_{T,j}^{2p}) \frac{\Delta_{ij}^2}{R^2} \, , \tag{5.4.2}$$

where $p_{T, \ X}$ is the transverse momentum of the cluster X. Δ_{ij}^2 is the angular separation between two clusters defined as $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, where y and ϕ are rapidity and azimuthal coordinates of the clusters. Usually Δ_{ij}^2 is denoted as ΔR_{ij}^2, but the "R" is dropped here to avoid confusion with the regular radius parameter R which sets the radius of a jet. Parameter p is a property of a particular reconstruction algorithm. $p = 1$, $p = 0$ and $p = -1$ correspond to the inclusive k_T, Cambridge/Aachen and the anti-k_T algorithms respectively. $0 < p < 1$ and $-1 < p < 0$ yield similar to the k_T and anti-k_T algorithms results.

- Once d_{iB} and d_{ij} values for all clusters and pairs are defined, the topoclusters are sequentially analyzed. For each cluster, d_{iB} is compared to all possible d_{ij}'s. If $d_{iB} > d_{ij}$ for all j's, the cluster is called a jet and removed; otherwise, clusters i and j are merged and returned for the further analysis. The procedure repeats until no topoclusters are left.

Different numerical values of the p parameter results into different performance of jets clustering by different algorithms. The k_T algorithm starts from merging soft cluster together, and then groups them to the harder clusters. The algorithm builds jets with irregular borders far from circular shapes in circumference. In the anti-k_T algorithm, soft clusters are combined with the harder ones, which results into circular-shaped shapes with hard objects in the center. The Cambridge/Aachen algorithm with $p = 0$ doesn't depend on the transverse momenta, and merges clusters in geometrical order. In ATLAS, jets are reconstructed with the anti-t_T algorithm. Figure 25 shows the difference in jets shapes acquired by using different algorithms for the same event.
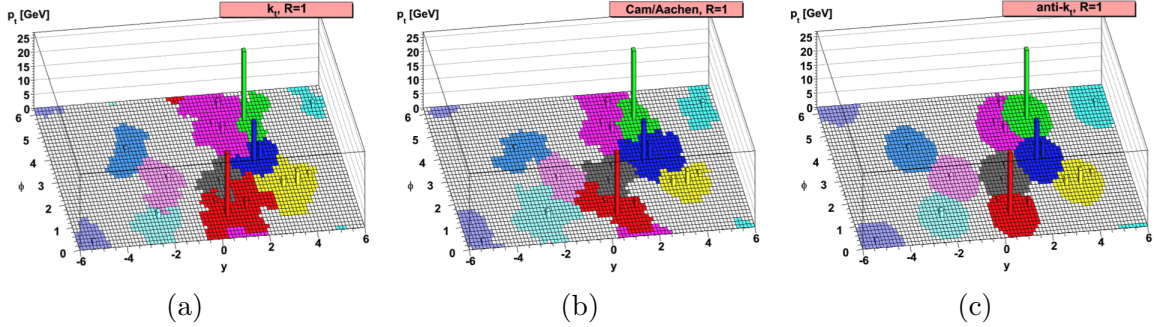
Figure 25: Examples of performance of the (a) k_T, the (b) Cambridge/Aachen and the (c) anti-k_T jets reconstruction algorithms for the same event. The k_T returns jets with irregular shapes, and so does the Cambridge/Aachen. The anti-k_T returns majority of the jets with circular shapes in the circumference clustered around the heavy object in the center, except for heavily overlapping jets. Figures from [117].

5.4.2 Jet Energy Scale and Resolution Calibration

In many physics analyses at ATLAS, the largest sources of systematic uncertainties are due to jets kinematics. The uncertainties can be classified into Jet Energy Scale (JES) and Jet Energy Resolution (JER) systematics. Therefore, precise knowledge the JES and the JER is important for both the SM measurements and new physics searches analyses.

As it was shown in section 5.3, when created, topoclusters are calibrated at the EM scale. The EM-calibrated topoclusters are then used to reconstruct jets. Therefore, to derive more precise measurement of the jets kinematics, a dedicated JES calibration must be performed. Figure 26 shows calibration stages for the EM-scaled jets. The calibration procedure corrects four-momentum and η of jets using simulation-based and data-based analyses.
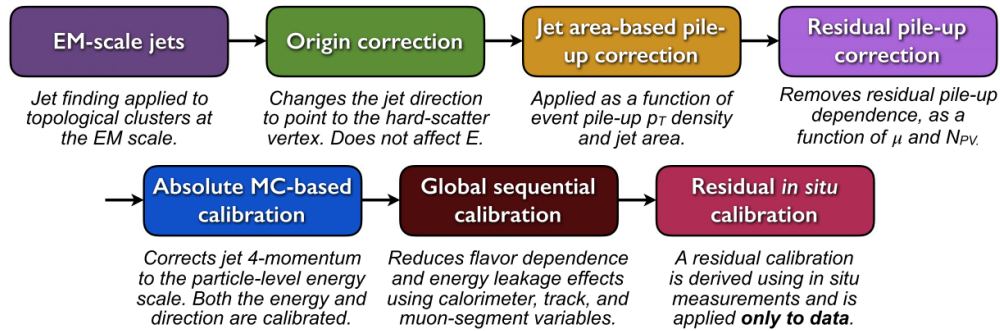
Figure 26: EM-scaled jets JES calibration stages. Figure from [118].

Jet origin correction improves the η resolution of jets. At this stage, four-momentum of a jet is recalculated to point to the primary vertex instead of the center of the detector, and the total jet energy remains unchanged. The pile-up correction is applied to remove the energy excess due to in-time and out-of-time pile-up, and consists of two components: an area-based p_T density subtraction [119] and a residual correction derived from the MC simulation. The absolute MC-based jets calibration corrects jets four-momenta to the

particle level energy scale. The correction is derived from truth jets in multi-jet MC events. The global sequential calibration uses calorimeter, MC and track-based variables to further improve the reconstructed energy and related uncertainties. The *in-situ* calibration addresses the difference in the jet response between data and MC simulation is corrected. It uses well-measured reference objects, such as photons, Z bosons and calibrated jets. In the last step of calibration, a set of 80 JES systematic uncertainties are analyzed for the reduction. 67 of the systematics come from Z/γ+jets and multi-b jets *in-situ* calibrations. All the JES calibration steps are described in detail in [118].

To correct the difference in JER between data and MC simulation, a method based on jets p_T smearing of simulated events is deployed [120]. The η-dependent resolution is derived from data and MC simulations of multi-jet events. The measured resolution considers sources from electronic noise in the detector, pile-up, statistical fluctuations in the deposited energy in calorimeters, constant resolution offset due to energy deposits in the passive components of the detector (readout electronics) and non-linear response from the detector. The detailed procedure of JER calibration can be found in [120].

5.4.3 Re-Clustered Jets

The large centre-of-mass energy of the colliding protons at the LHC sets good conditions for production of Lorentz-boosted heavy particles, whose decay products become highly collimated. The angular separation $\Delta R = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$ (where y and ϕ are the rapidity and azimuthal coordinates of jets i and j) between two jets is estimated [121] as:

$$\Delta R \approx \frac{2m}{p_\mathrm{T}} \tag{5.4.3}$$

where m and p_T are the mass and the transverse momentum of the parent particle. Thus, the more boosted the parent particle, the smaller the angle between the child particles. The jet re-clustering technique allows to effectively reconstruct collimated child particles with one large-R jet ($R = 0.8$), instead of treating them as multiple independent small-R jets ($R = 0.4$). The re-clustering algorithm takes small-R jets as the input, and jet trimming [122] is applied to suppress the effect of the pile-up soft contributions. The trimming procedure cuts jets with p_T f_{cut} times smaller than the analyzed jet. By default, $f_{cut} = 0.1$.

5.4.4 Flavor Tagging

Efficient identification of jets associated with heavy-flavor quarks (*bottom* and *charm*) and with light quarks (*up*, *down* and *strange*) is crucial for many analyses at ATLAS. Examples for such analyses are the SM measurements for the Higgs boson decay in the highest branching ratio channel $H \rightarrow b\bar{b}$, or varieties of BSM physics searches, such as searches for SUSY in channels with multiple b-jets in the final state presented in chapter VI.

All b-quarks produced in the inelastic pp collisions at the LHC hadronize and form a bound states with quarks of another flavor. The bound states are called B mesons. Lifetimes of the neutral (B^0) and charged ($B^\pm$) mesons are approximately 1.5×10^{-12} and 1.6×10^{-12}

seconds [10]. Traveling at speeds close to the speed of light, the B mesons can cover a distance of the order of a millimeter, before they decay.

Low-level b-tagging algorithms have been developed [123] in ATLAS and further combined to construct one discriminator:

- **Impact Parameter 3D Algorithm (IP3D)** is based on a log-likelihood (LLR) discriminant separating tracks assigned to jets according to whether or not they are compatible to the primary vertex hypothesis. The algorithm computes significance of the transverse d_0 and longitudinal z_0 impact parameters of tracks inside of a jet. Typically, tracks originating from the B mesons decays are characterized by larger impact parameters than light tracks.

- **Secondary Vertex Fining Algorithm (SV1)** is another LLR classifier. It tries to reconstruct a single secondary vertex for each jet. The algorithm takes the jet axis direction, the primary vertex position and the list of track associated to the jet as the input and returns the secondary vertex position with a list of tracks associated to the vertex. For a b-jets with both b and c tracks, the SV1 algorithm merges two associated vertices into one, or, if they are far apart, return the one with the largest track multiplicity.

- **Decay Chain Multi-Vertex Algorithm (JetFitter).** The JetFitter is a Kalman Filter based algorithm that reconstructs the decay chain of b- and c-hadrons when it is possible.

The three algorithms with additional kinematic information are used to construct one high-level classifier. One of the high-level classifiers, named MV2c10 [124], is derived using the Boosted Decision Tree (BDT) technique. Another example is DL1r [125], a novel flavor-tagging algorithm based on an artificial deep learning neural network. It uses the same set of the low-level taggers, as the MV2c10, with the addition of the JetFitter c-tagging variables [125].

5.5 Electrons

In the ATLAS detector, electrons leave tracks in the ID, as any other charged particles, and clusters of energy deposits in the electromagnetic calorimeter. Electron identification is based on matching the inner detector tracks to the ECal clusters. The electrons are reconstructed in the central region ($\eta < 2.47$) in a few steps [126]:

- **Seed-cluster reconstruction.** A sliding window of 3×5 cells size corresponding to 0.025×0.025 in the $\eta - \phi$ space runs through the ECal and searches for cluster seeds with $E_T > 2.5$ GeV. Once the seeds are found, clusters are formed around them using a clustering algorithm [127].

- **Track reconstruction.** The electron track reconstruction is done in two stages: pattern recognition and track fit. The first attempt to reconstruct a track is done with the pion hypothesis for the energy loss. The requirements applied to track candidates

are: constructed with three hits in different layers of the inner detector, transverse momentum above 1 GeV, can be extrapolated to at least 7 hits. If no tracks which meet the requirements are reconstructed, the second attempt is made using the electron hypothesis. In the second stage, the tracks are fitted with either the pion or electron hypothesis (whichever were used to reconstruct each track) using the ATLAS Global χ^2 Track Fitter [128].

- **Electron specific track fit.** The electron track candidates are loosely matched to the reconstructed clusters in the ECal using the distance in η and ϕ between the extrapolated track and the calorimeter cluster center.

- **Electron candidate reconstruction.** Track candidates are refit using stronger conditions and matched to the calorimeter clusters. An algorithm [129] is applied to select the "primary" track in cases when more that one track candidate fulfill the matching conditions. The algorithm uses the track-cluster distance R as the main discriminator. Electron candidate tracks with no precision hit tracks are considered as photons and removed. Three working points, *Loose*, *Medium* and *Tight* corresponding to signal efficiencies of 93%, 88% and 78% respectively, are used to apply cuts on the likelihood score.

A likelihood based algorithm is applied to identify signal-like (prompt) electrons and separate them from the background-like objects, such as hadronic jets and converted photons. PDFs of the calorimeter and track related variables are supplied to the algorithm which constructs a likelihood score; the PDFs are generated from $Z \to ee$ and $J/\psi \to ee$ MC simulated events. To address specific needs of different analyses, three working points, *Loose*, *Medium* and *Tight* corresponding to signal efficiencies of 93%, 88% and 78% respectively, are used to apply cuts on the likelihood score.

5.6 Muons

Muons leave tracks in the ID and the MS of the ATLAS detector. Tracks are reconstructed in both subsystems independently, then the information is combined to form muons tracks [130]. The ID tracks are reconstructed in the same way as jets and electrons, as discussed in the previous sections. This section presents the MS muons tracks reconstruction and the combined ID+MS muon reconstruction.

First, segments of hits in different layers of the MS are constructed. Hits in MDT are fitted with a straight line in the bending plane of the detector. Information from RPC or TGC is used to measure the track position in the orthogonal to the bending plane. Segments of hits in the CSC detectors are build using combinatorial search in the η and ϕ planes.

In the second stage of the MS track reconstruction, segments of hits from different layers are fitted together. The fit starts from searches for seeds in the middle layers of MS where more hits are available. Then the search is extended to the inner and outer layers. Matching of segments from different layers is done using criteria based on the hit multiplicity and fit quality. To build a combined track inside of the MS, at least two matching segments are

required, except for the barrel−end-cap transition regions, where one segment with η and ϕ coordinates is enough.

Several MS track candidates may share the same segment. An overlap removal algorithm is applied to select the best matching track to a single segment, or to allow the segment to be shared between two tracks. Moreover, tracks are allowed to share two segments but only if they have no shared hits in the outermost layer.

In the last step, track candidates are fitted with the χ^2 global fit. Hits with large contributions to the χ^2 are removed, and the fit is repeated until the χ^2 will satisfy selection criteria. If additional hits are consistent with the candidacy trajectory, they are added and the χ^2 fit is repeated.

Once the ID and MS tracks are reconstructed, they are matched to create global muons tracks. Four types of muons are defined based on the subdetectors used to reconstruct them:

- **Combined (CB) muons.** Two approaches to construct CB muons are available: an outside-in and an inside-out. In the outside-in muons are reconstructed in the MS, and then the tracks are extrapolated to match ID tracks. The inside-out is a complimentary approach where ID tracks are extrapolated and matched to MS tracks. The combined tracks are then refit, and MS hits can be added ot removed to improve the fit.

- **Segment-tagged (ST) muons.** ID tracks are extrapolated to match at least one *local* MDT or CSC segment. ST muons are used because of the low p_T of the particles or the tracks fall in regions with reduced MS acceptance.

- **Calorimeter-tagged (CT) muons.** ID tracks are matched to the energy deposits in the calorimeters compatible with the minimum−ionizing particle hypothesis. Although the method has the lowest overall purity, it is useful in reconstruction of muons in the specific region with $|\eta| < 0.1$ and $15 < p_\mathrm{T} < 100$ GeV, where other are insufficient.

- **Extrapolated (ME) muons.** Muons trajectories are reconstructed using information from the MS only. The trajectories are required to point to the primary vertex. The method is primarily used to reconstruct muons in $2.5 < \eta < 2.7$ range which is not covered by the ID.

The last step of muons reconstruction is the overlap removal between different types of muons presented above. In cases when two muon candidates share an ID track, preference is given in the CB − ST − CT order. Overlaps with ME muons in the MS are resolved by selecting the track with better χ^2 fit quality.

After muons candidates are reconstructed, a muon identification is performed. The procedure allows for QCD background suppression and increase of the purity of prompt muons. The identification sets requirements on the number of hits, the momentum measured in the MS and ID (except for ME muons) and the fit quality of the trajectory. To address specific needs of different analyses, four working points are defined: *Loose*, *Medium*, *Tight* and *high-p_T* corresponding to 98.1%, 96.1%, 91.8% and 80.4% respectively.

5.7   Missing Transverse Energy

The missing transverse energy E_T^{miss} is a two dimensional vector in the $R - \phi$ plane defined [131] as the negative vector sum of p_T of all reconstructed and calibrated objects in an event: jets, leptons, photons and soft terms. The soft term is computed using ID tracks that are not associated with any reconstructed objet and that are originating form the primary vertex.

$$E_T^{miss} = -(\sum_{jets} p_T + \sum_{leptons} p_T + \sum_{photons} p_T + \sum_{soft} p_T) \tag{5.7.1}$$

E_T^{miss} can arise from neutrinos which leave no footprints in any layer of the ATLAS detector, imperfection of the objects reconstruction or due to the BSM physics particles.

CHAPTER VI

Searches For SUSY in Events with Multiple b-jets and Large Missing Transverse Momentum in the Final State

In this chapter searches for gluino pair production in events with multiple b-jets and large missing transverse momentum in the final state are presented. The analysis targets signal models with the gluinos decaying in top-stop pairs, and the stop quarks decay into the top quark and the LSP. The channel is named Gtt. General aspects of the analysis setup, including the data and Monte Carlo samples, physics objects definitions and definitions of the analysis observables, are given in section 6.1. Results of studies in preselection regions and kinematic reweighing of events with 1 or more leptons are discussed in section 6.2. The cut-and-cout analysis strategy is presented in section 6.3. Systematic uncertainties that affect the analysis are presented in section 6.4. The results of the analysis are discussed in section 6.5. A simplified Feynman diagram of the gluino pair production process is shown in figure 27.
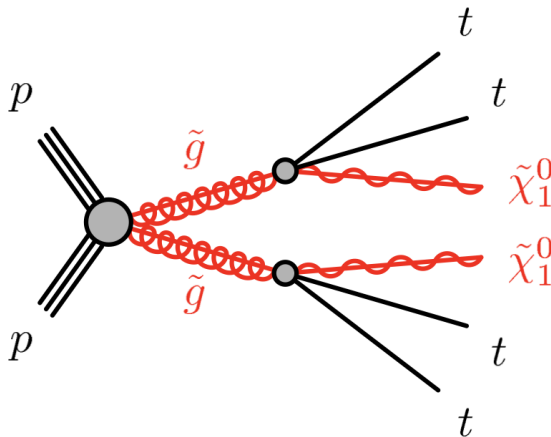
Figure 27: A simplified Feynman diagram of a gluino $\tilde{g}$ production that decays into a pair of gluinos. The process is denoted by the 4-vertex $pp \to \tilde{g}\tilde{g}$. Each of the gluinos decays into a top-stop pair, and each of the stops subsequently decays into a SM top quark and the LSP $\tilde{\chi}_1^0$. The two-step process is assumed to happen immediately, and denoted by the 4-body vertices $\tilde{g} \to tt\tilde{\chi}_1^0$.

6.1 Analysis Setup

6.1.1 Data and Monte Carlo Samples

Three types of samples are used in the analysis. The first type is the real data collected by the ATLAS detector. The second type is called signal samples. Signal samples consist of events with SUSY signatures. The samples are used to optimize the search strategy and provide the result hypotheses. The third type is called background samples. Each SM background process is prepared independently one from another. List of software versions used for signal and background samples is summarized in table 3.

The analysis uses the total amount of data of pp collisions at the centre-of-mass energy of 13 TeV in the 25 ns bunch spacing configuration recorded by the ATLAS detector during $2015 - 2018$ data-taking period which corresponds to an integrated luminosity of $139.0 fb^{-1}$.

Signal samples are produced with MC collision generators [10] in two steps. In the first step matrix elements (ME) for gluino pair production $p + p \rightarrow \tilde{g} + \tilde{g}$ are generated using MadGraph5_aMC@NLO v2.3.3 [95] with up to two additional partons at the leading order (LO) with the NNPDF 2.3 [132] Parton Density Function (PDF) sets. In the second step, modeling of gluino decays to top or bottom quarks pair and the neutralino as well as modeling of parton showering (PS), hadronization and underlying event (UE) is done using Pythia v8.212 [93]. The two stages are further matched using CKKW-L [133] prescription with a matching scale set to one quarter of the gluino mass. As the result of the modeling, 152 Gtt signal samples with different $\tilde{g}, \tilde{\chi}_1^0$ mass pairs were generated. In the samples are produced for gluino masses in the range from $1.1\,TeV$ to $2.8\,TeV$. For each gluino mass the upper bound of the neutralino mass was set with respect to the available kinematic phase space: $m(\tilde{g}) - 2m(top\ quark)$.

The dominant SM background process for the Gtt signal models is the production of the top-antitop quark pairs $(t\bar{t})$ in association with additional high-p_T jets due to the initial and/or the final state radiation. Similar to the signal modeling, the $t\bar{t}$ production is modeled in two steps. In the first step, Powheg-Box v2 [134, 135] generator at the next-to-leading order (NLO) with the NNPDF 3.0 [136] PDF was used to model matrix elements of the $t\bar{t}$ production. In the second step, the matrix element is interfaced to Pythia v8.230 with NNPDF 2.3 PDF set for showering, hadronization and underlying event modeling, using the NNPDF2.3 PDF set [132]. Smaller SM backgrounds come from production of the single top quarks, $t\bar{t}$ pairs production in association with vector or scalar bosons and additional high-p_T jets, vector boson production in association with additional high-p_T jets, and for-top production $(t\bar{t}t\bar{t})$ events, and diboson (WW/WZ/ZZ) events.

Dijet, or pure quantum chromodynamics (QCD) production, is another background to signal events. There is no leptons in the final state of QCD production, and total number of events might be small but potentially non-negligible. MC simulation usually fails to reproduce QCD production in kinematic regions targeted by the analysis. Do estimate the background, the data-driven technique [137] is used, in which a template fit of $\Delta\phi_{min}^{4j}$ is

performed in a region with large statistics ($\Delta\phi_{min}^{4j} < 0.1$), and then extrapolated to the phase-space of the interest of the analysis ($\Delta\phi_{min}^{4j} \geq 0.4$). More details on the data-driven QCD production estimation are given in section 6.2.1.

Process	Generator + fragmentation/hadronization	Tune	PDF set	Cross-section order
Gtt signal	MadGraph5_aMC@NLO v2.3.3 + Pythia V8.212	A14	NNPDF2.3	NNLO_approx +NNLL [138, 139, 140, 141, 142, 143, 144, 145, 146]
t$\bar{\text{t}}$	Powheg-Box v2 +Pythia v8.230	A14	NNPDF3.0 (ME) NNPDF2.3(UE)	NNLO+NNLL [147]
Single top	Powhg-Box v2 +Pythia v8.230	A14	NNPDF3.0(ME) NNPDF2.3(UE)	NLO t/s-channel NLO+NNLL [148] (Wt)
t$\bar{\text{t}}$W/t$\bar{\text{t}}$Z	MadGraph5_aMC@NLO v2.3.3 +Pythia v8.210	A14	NNPDF3.0 (ME) NNPDF2.3 (UE)	NLO [95]
4-tops	MadGraph5_aMC@NLO v2.3.3 +Pythia v8.186	A14	NNPDF2.3	NLO [95]
t$\bar{\text{t}}$h	Powheg-Box v2 +Pythia v8.230	A14	NNPDF3.0 (ME) NNPDF2.3 (UE)	NLO [149]
Dibosons *WW, WZ, ZZ*	Sherpa v2.2.1 [94]	Default	NNPDF3.0	NLO [150, 151]
W/Z+jets	Sherpa v2.2.1	Default	NNPDF3.0	NNLO [152]

Table 3: List of generators used for the different processes. Information is given about the underlying event tunes, the PDF sets and the pQCD highest-order accuracy used for the normalization of the different samples.

6.1.2 Physics Objects Definitions

As discussed in Chapter III, the ATLAS detector captures energy deposits of different particles with its different subsystems. The energy deposits from different layers of the detector are reconstructed into tracks of particles and calorimeters clusters. The general information about objects reconstruction at ATLAS is discussed in Chapter V. This section discusses definitions of particular objects used in the analysis.

All objects can be sorted into two categories: prompt objets, arising from the primary hard-scattering pp collisions, and secondary objets, originating from further processes. Some secondary objects are discarded using an overlap removal procedure, which is discussed in the section as well.

- **Primary vertex**. Interaction vertices are reconstructed from at least two tracts with $p_T > 0.4\ GeV$. The vertex with the largest sum of squares of p_T from all associated tracks is identified as the primary vertex of an event.

- **Small-radius jets**. Partons produced in the pp hard scattering process at the LHC undergo hadronization, and the final state jets consist of many hadrons. Thus, one needs to correctly identify all hadrons associated with a particular parton and construct a jet out of the objects. Reconstruction of hadronic jets is done in two step. In the first step, topological clusters ("topoclusters") are created with three-dimensional clustering of energy deposits measured by calorimeter cells. The procedure resolves direction and energy of single hadrons [111]. In the second step, the topoclusters are used as an input to the anti-k_T sequential recombination algorithm [117] with a radius of $R = 0.4$. Lastly, the energy of the reconstructed jets is calibrated to account for experimental effects using a combination of MC-based and *in situ* techniques [153].

- **Large-radius jets**. The large radius jets are reconstructed in a similar way to the small-radius jets with. Radius of $R = 0.8$ used in the anti-k_T algorithm. These "fat" jets are needed to construct some variables sensitive to the large-scale structure of an events, which is very useful in Gtt analysis with large mass splitting between the gluino $\tilde{g}$ and the neutralino $\tilde{\chi}_1^0$

- **b-jets**. The b-jets are jets associated with a b hadron. The Gtt channel signature is featured by large multiplicity of b-jets in the final state. Thus, a good b-jet identification is required for the analysis. The MV2c10 b-tagging algorithm at the 77% working point is used in the analysis.

- **Leptons**. There are three generations of leptons, and only two are reconstructed with two sets of criteria each: baseline and signal.

  - **Electrons**. Two kinds of electrons are considered in the analysis: the baseline electrons and the signal electrons. The baseline electrons are reconstructed using a high-efficiency and low-rejection algorithm using the information from the tracker and the electromagnetic calorimeters, and required to meet $p_T > 20\ GeV$, $| \eta | < 2.47$ criteria as well as requirements on the uncertainties on the transverse and longitudinal impact parameters $| d_0/\sigma_{d_0} | < 5$ and $| \Delta z \sin\theta | < 0.5$. The baseline electrons are used to define all-hadronic channels vetoing leptonic events. The signal electrons are used to define the leptonic channel. The signal electrons are defined by an additional set of criteria applied on top of the baseline electrons identification, which results into a lower-efficiency and higher-rejection operation point [154].

  - **Muons**. Similar to the electrons, there are two kinds of muons defined: the baseline and the signal muons. The only difference for the muons is the requirement to the transverse impact parameter of the muons tracks: $| d_0/\sigma_{d_0} | < 3$.

  - **τ-leptons**. Because of their large mass, τ-leptons tend to decay to hadrons. For this reason, no attempt was made to reconstruct the leptons.

- **Overlap Removal**. Overlaps between candidate objects are removed sequentially. First, electron candidates sharing a track with a muon are assumed to arise from muons bremsstrahlung and thus discarded. Second, overlaps between electrons and jet candidates are resolved. This stage targets two kinds of objets: jets lying within $\Delta R < 0.2$ cone from electrons are assumed to be originating from prompt electrons showering and thus discarded while corresponding electrons are kept; electrons with $p_T < 50$ GeV that are lying within a $\Delta R < 0.4$ cone from a jet are assumed to be produced in the decay chain of hadrons and thus discarded while corresponding jets are kept. To increase acceptance for boosted events, the latter removal is performed using a threshold of $\Delta R = min(0.4, \ 0.04 + 10 \ \mathrm{GeV}/p_T)$. Third, overlaps between muons and jet candidates are resolved. Identical to the electrons, muons lying close to jests candidates are removed and jets are kept. Jets originating from muons bremsstrahlung are expected to have few matching inner detector tracks and lay close to the muon. Therefore, if the angular distance $\Delta R(jet, muon) < 0.2$, and the jet candidate has three or fewer inner detector tracks, such a jet is discarded and the corresponding muon is kept.

- **Missing transverse energy,** E_T^{miss}, is a two dimensional vector in the $R - \phi$ plane. The vector is defined as the negative vector sum of p_T of all selected and calibrated objects described above in an event plus the soft term.

6.1.3   Definition of Analysis Observables

Event-level kinematic variables are introduced to the analysis to use in a various ways.

- **N$_{\mathbf{jets}}$**: the number of small-radius jets with $p_T > 30$ GeV and $\mid \eta \mid < 2.8$.

- **N$_{\mathbf{b-jets}}$**: the number of b-jets with $p_T > 30$ GeV and $\mid \eta \mid < 2.5$.

- **N$_{\mathbf{lep}}$**: the number of signal leptons.

- **$\mathbf{\Delta\phi_{min}^{4j}}$**: the minimum $\Delta\phi$ between the four jets with the highest p_T and the E_T^{miss}:

$$\Delta\phi_{min}^{4j} = min(|\phi_1 - \phi_{E_T^{miss}}|, \ ... \ , \ |\phi_4| - \phi_{E_T^{miss}}) \tag{6.1.1}$$

- **m$_{\mathbf{eff}}$**: the effective mass is defined as the sum of the p_T of all jets (with $p_T > 30$ GeV and $\mid \eta \mid < 2.8$), p_T of all signal leptons and E_T^{miss}:

$$m_{eff} = \sum_{j \leq n} p_T^{jet \ j} + \sum_{k \leq m} p_T^{lep \ k} + E_T^{miss} \tag{6.1.2}$$

- **m$_{\mathbf{T}}$**: is defined as the transverse mass between the leading lepton and E_T^{miss}:

$$m_T = \sqrt{2p_T^{leading \ lep.} E_T^{miss}\{1 - \cos(\Delta\phi(E_T^{miss}, \ lepton))\}} \tag{6.1.3}$$

In the $W + jets$ and $t\bar{t}$ events, the lepton and E_T^{miss} originate from the same vertex, and m_T will be around $m(W)$. Therefore, the variable is a powerful in discriminating SUSY events from the two backgrounds.

- $\mathbf{m_{T,min}^{b-jets}}$: is defined as the minimum transverse mass between the E_T^{miss} and the three leading b-jets:

$$m_{T,min}^{b-jets} = \sqrt{(E_T^{miss} + p_T^{jet_i}) - (E_x^{miss} + p_x^{jet_i}) - (E_y^{miss} + p_y^{jet_i})} \qquad (6.1.4)$$

The variable is useful for separating signal and control/validation regions. It has a kinematic endpoint near the top mass for $t\bar{t}$ backgrounds. The $\tilde{\chi}_1^0$ produces large E_T^{miss} in SUSY events, and it is independent from the b-kinematics. Therefore, value of $m_{T,min}^{b-jets}$ in signal events can be much larger than in background.

- $\mathbf{M_J^{\Sigma,4}}$: is the total jet mass variable, defined as:

$$M_J^{\Sigma,4} = \sum_{i \leq 4} m_{J,i} \qquad (6.1.5)$$

where m_J refers to the mass of large-radius re-clustered jets in the event. The sum is taken over the first 4 leading re-clustered jest in an events. The variable is sensitive to jets originating from the signal events. Boosted and semi-boosted top quarks in the signal events leads to the formation of high-p_T massive jets. Thus, large $M_J^{\Sigma,4}$ might indicate a potentially SUSY event.

6.2 Events Preselection

After the definition of the analysis objects and variables, the next step is the estimation of the SM background. This section discusses the main aspects of the procedure. The background estimate in the analysis relies on usage of MC samples where possible, and on data-assisted techniques when MC cannot provide with a reliable simulation results.

- **Estimate from MC simulation**. Due to its large cross-section in pp collisions at the LHC, $t\bar{t}$ production is expected to be the dominant background in all signal regions of the analysis. To eliminate possible discrepancies in the combined SM background yields from data, only the $t\bar{t}$ background will be scaled is so-called control regions designed specifically for each signal region. Detailed definitions of the control and signal regions are given in section 6.3. All other backgrounds are estimated directly from MC simulations. To fix the data/MC discrepancy in some analysis variables, a kinematic reweighing procedure is applied to all MC samples. The reweighing affects events with at least one lepton and does not change overall normalization. More details on the procedure are given in subsection 6.2.3.

- **Data-driven estimate**. QCD estimation is the only background that is not taken from MC simulations. Details on its estimate are given in section 6.2.1.

6.2.1 QCD Estimation is the 0L Regions

As it was highlighted in section 6.1.1, MC simulation for QCD production with zero leptons fails to provide with a reliable results in kinematic regions targeted by the analysis.

However, the production of the dijet events has extremely large cross-section at the LHC [155]. In contrast, requirements in $E_\mathrm{T}^{miss} \geq 200\,GeV$ and in at least one signal lepton result in negligible contribution of QCD background in such regions. Therefore, the QCD background is estimated only in the zero-lepton region. Following the strategy from [156], a template method is used to fit $\Delta\phi_{min}^{4j}$ distribution. The distribution is fitted in the $\Delta\phi_{min}^{4j} < 0.1$ region, and the other SM background production from MC is subtracted from the data. Then, the fit is validated in the $0.1 \leq \Delta\phi_{min}^{4j} < 0.2$ region and propagated to the $\Delta\phi_{min}^{4j} > 0.4$ region. The fitted distributions for different data taking campaigns are shown in figure 28, and validations of the N_{jets} and E_T^{miss} are shown in figure 29. Since it wasn't possible to validate the QCD background in the signal region of $\Delta\phi_{min}^{4j} > 0.4$, it was decided to assign a systematic uncertainty of 300% to the background estimation. QCD is the minor background, and the large relative systematic uncertainty will not affect sensitivity of the analysis.

(a) mc16a

(b) mc16d

(c) mc16e
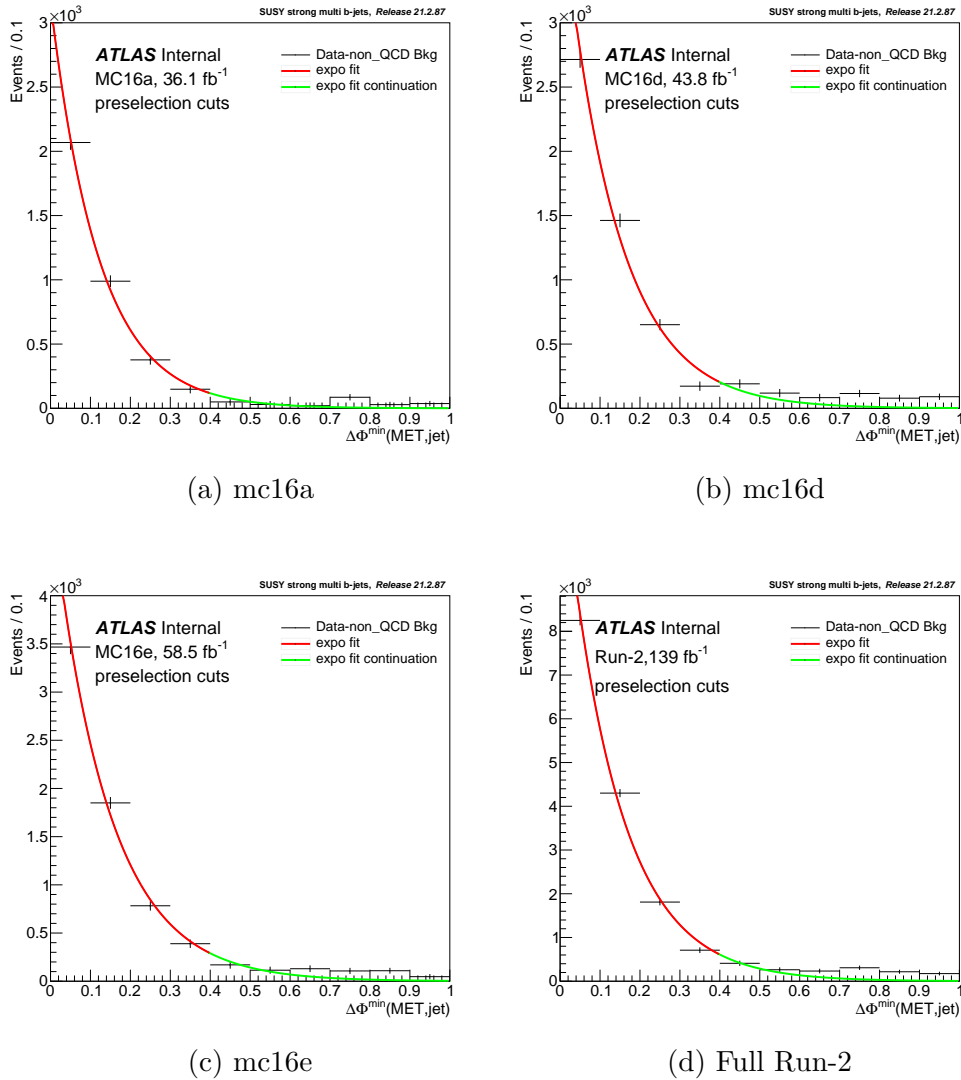
(d) Full Run-2

Figure 28: Results of the exponential fits of $\Delta\phi_{min}^{4j}$ for three consecutive data-taking campaigns in 28a, 28b, 28c and for the full Run-2 in 28d.
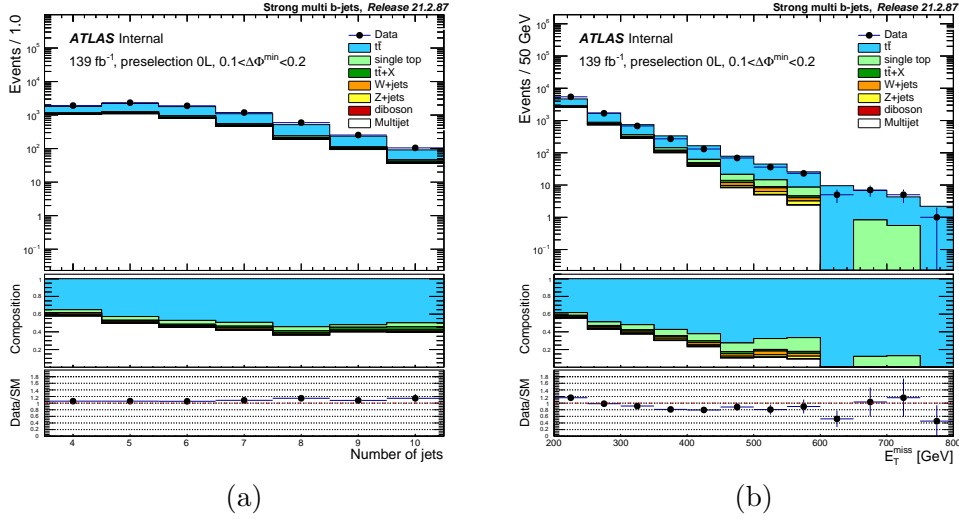
Figure 29: QCD estimate validation for in the 0L background for N_{jets} and E_T^{miss} in 29a and 29b respectively.

6.2.2 Pre-Fit Data/MC Comparisons

To make sure that potential excess in the data above the SM backgrounds in regions with large expected cross-section of the signal is not caused by inaccuracy of the MC modeling, the modeling is verified in preselection regions. The regions are characterized by sets of loose cuts. Due to the loose cuts, the expected cross-section of signal events in the regions is much smaller than the total cross-section of the combined SM background. Therefore, any large disagreements between the data and the SM background in the preselection regions can be spotted and corrected where needed. Selections for two different kinematic regions, with veto on leptons number (Loose 0L) and with at least one lepton (Loose 1L) in the final state, are summarized in table 4.

Region name	N_{lep}	N_{jets}	N_{b-jets}	E_T^{miss}, GeV	$\Delta\phi_{min}^{4j}$, rad
Loose 0L	$= 0$	≥ 4	≥ 3	≥ 200	≥ 0.4
Loose 1L	≥ 1	≥ 4	≥ 2	≥ 200	X

Table 4: Sets of selection applied to two preselection regions: Loose 0L and Loose 1L

Spectra of some analysis variables in the Loose 0L and Loose 1L regions are given in figure 30 and figure 31 respectively, and distributions for all analysis variables are given in Appendix 0.1. While the 0L distributions show no significant disagreement between the data and the total MC simulated SM background, several kinematic variables, such as E_T^{miss} and p_T of jets and leptons, show moderate disagreement in events with one or more leptons. The disagreement appears as downward slope in the data/MC ratio. Similar behavior is observed for all jets and leptons p_T and in the E_T^{miss}. This disagreement has strong impact on the

analysis, thus it must be corrected with a specifically derived reweighing that brings the MC prediction closer to the data. The procedure is described in the next section.

(a) Jets multiplicity

(b) E_T^{miss}

(c) leading small-R jet p_T

(d) m_{eff}
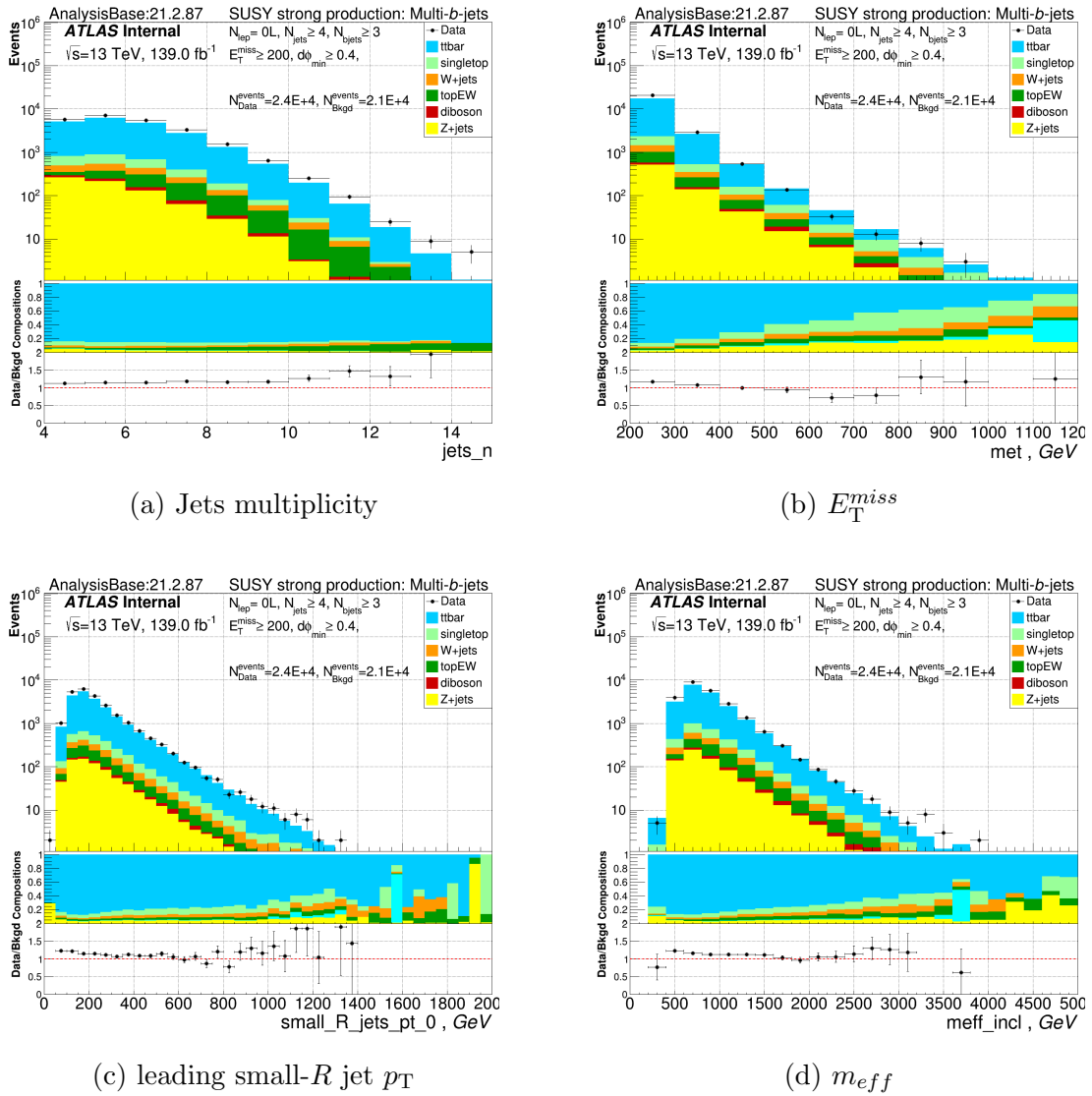
Figure 30: Spectra of some analysis variables in events with no leptons. Top canvas: Data (black dots) and SM background (in color) yields. Middle canvas: background compositions. Bottom canvas: data to background ratios.
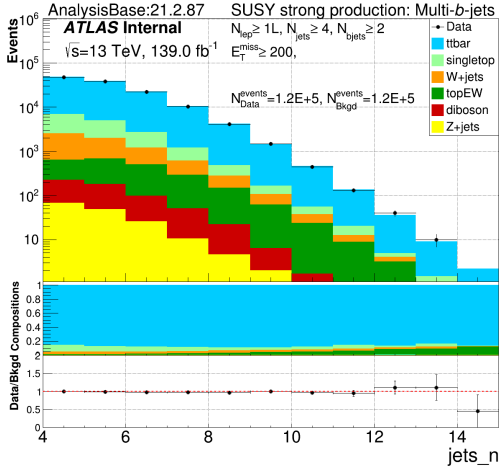
(a) jets multiplicity
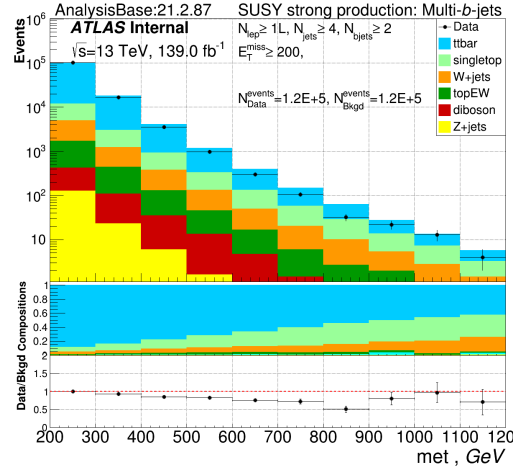
(b) E_T^{miss}
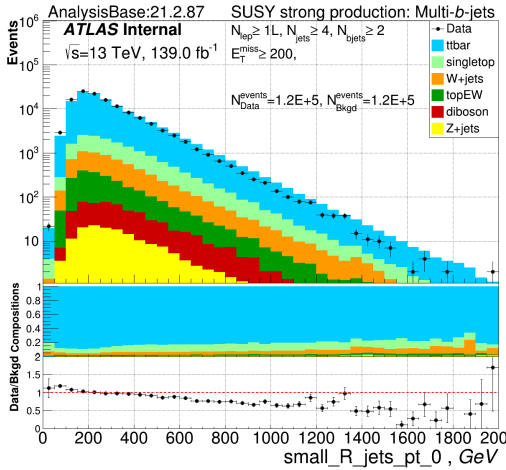
(c) leading small-R jet p_T
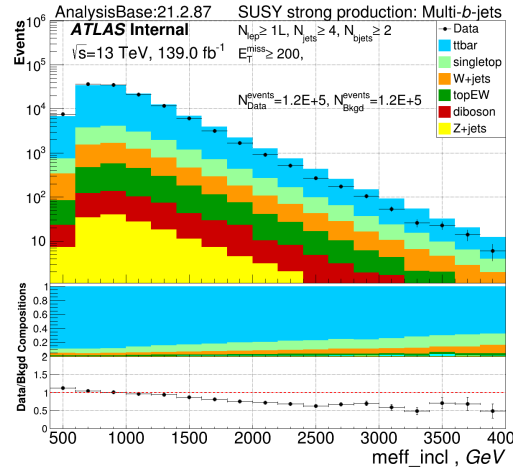
(d) m_{eff}

Figure 31: Spectra of some analysis variables in events with at least one lepton. Top canvas: Data (black dots) and SM background (in color) yields. Middle canvas: background compositions. Bottom canvas: data to background ratios. For all "p_T-related" variables discrepancy in data to background ratio is observed.

6.2.3 Kinematic Reweighing in 1L Region

The data to the SM background MC comparison on the preselection level in events with at least one lepton shows disagreement in all p_T-related variables: small-R jets p_T, leptons p_T, E_T^{miss} and the combination of the variables $- m_{eff}$. The disagreement appears as downward slope in the data/MC ratio. The trend in the data/MC ratio in the m_{eff} distribution and its components is observed in most of the generated SM background samples. Therefore, independent reweighing factors are required for different SM processes. Four mutually orthogonal control regions enriched with $t\bar{t}$, single top, W+jets and Z+jets events are constructed. The control regions are orthogonal to the rest of the analysis through

requirement in exactly two or zero b-tagged jets. Selections for the control regions are summarized in table 5.

Criteria common to all regions: $N_{jets} \geq 4$, $E_{\mathrm{T}}^{miss} \geq 200 \ GeV$				
Control region	N_{lep}	N_{b-jets}	$m_{\mathrm{T,min}}^{b-jets} \ GeV$	$M_Z \ GeV$
$t\bar{t}$-enriched	$= 1$	$= 2$	≤ 350	$-$
single-top-enriched (ST)	$= 1$	$= 2$	> 350	$-$
W-enriched	$= 1$	$= 0$	$-$	$-$
Z-enriched	$= 2$	$= 0$	$-$	$[60, 120]$

Table 5: Definitions of the control regions used to derive the kinematic reweighing factors. The N_{b-jets} requirement makes the control regions orthogonal with rest the analysis which uses $N_{b-jets} \geq 3$ regions.

Even though the control regions in table 5 are constructed in the way to maximize number of specific events, some contamination with other SM events is possible. It is assumed that kinematic reweighing factors depend only on the control region, and are independent from specific event types within a control region. In other words, there is only one reweighing factor per one control region. To derive the reweighing factors, a matrix equation 6.2.1 must be solved for specific segments in m_{eff}. Components of the 4 by 4 matrix on the left side of the equation are fractions of events of each type in each of the control regions. Vector w consists of reweighing factors that needs to be derived. Components of the vector in the right side of the equation are the data/MC ratios. The selection on the number of segments in m_{eff} and their length is driven by statistics. The most optimal selection was found to be 12 equal segments on $4200 \ GeV$ range (i.e. each segment is 300 GeV wide). Reweighing factors for each control region from each segment of m_{eff} are further combined and fitted with the exponential function. The resulting exponential fits are used to rescale the MC events. The results of the reweighing is shown in figure 32. From the figure it follows that the new reweighing factors corrected disagreement in data/MC m_{eff} distribution, but induced disagreement in N_{jets} appearing as a positive slope in data/MC. Thus, more sophisticated method is needed.

$$
\begin{pmatrix}
f_{CR-t\bar{t}}^{t\bar{t}} & f_{CR-t\bar{t}}^{W+jets} & f_{CR-t\bar{t}}^{single-top} & f_{CR-t\bar{t}}^{Z+jets} \\
f_{CR-W+jets}^{t\bar{t}} & f_{CR-W+jets}^{W+jets} & f_{CR-W+jets}^{signle-top} & f_{CR-W+jets}^{Z+jets} \\
f_{CR-single-top}^{t\bar{t}} & f_{CR-ST}^{W+jets} & f_{CR-ST}^{single-top} & f_{CR-ST}^{Z+jets} \\
f_{CR-Z+jets}^{t\bar{t}} & f_{CR-Z+jets}^{W+jets} & f_{CR-Z+jets}^{single-top} & f_{CR-Z+jets}^{Z+jets}
\end{pmatrix}
\begin{pmatrix}
w^{t\bar{t}} \\
w^{W+jets} \\
w^{single-top} \\
w^{Z+jets}
\end{pmatrix}
=
\begin{pmatrix}
(\frac{Data}{Bkgd})_{CR-t\bar{t}} \\
(\frac{Data}{Bkgd})_{CR-W+jets} \\
(\frac{Data}{Bkgd})_{CR-ST} \\
(\frac{Data}{Bkgd})_{CR-Z+jets}
\end{pmatrix}
$$

(6.2.1)
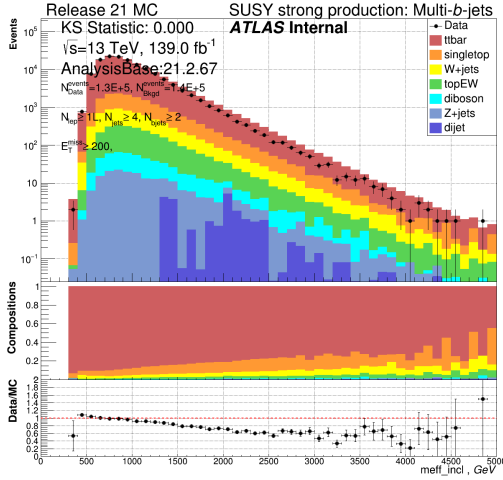
(a) m_{eff} in events with at least 1 signal lepton before reweighing
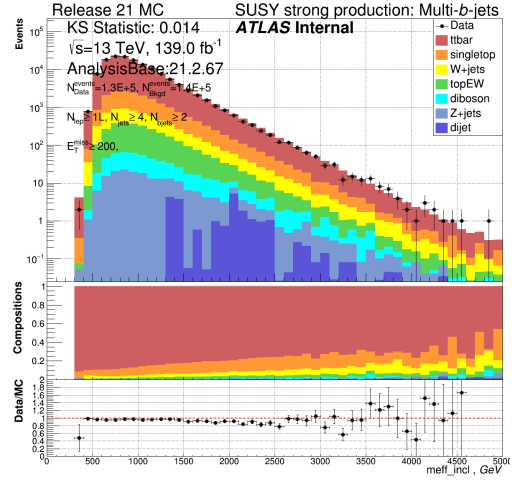
(b) m_{eff} in events with at least 1 signal lepton after reweighing
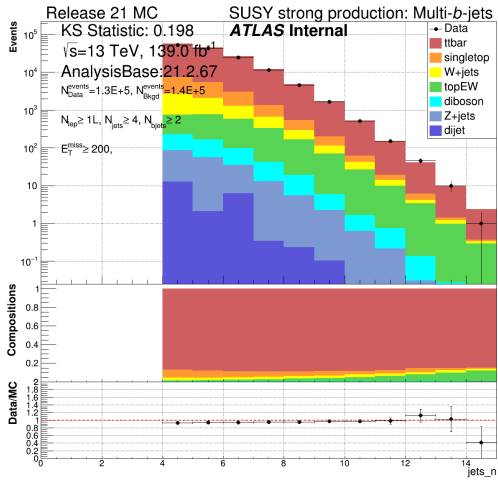
(c) N_{jets} in events with at least 1 signal lepton before reweighing
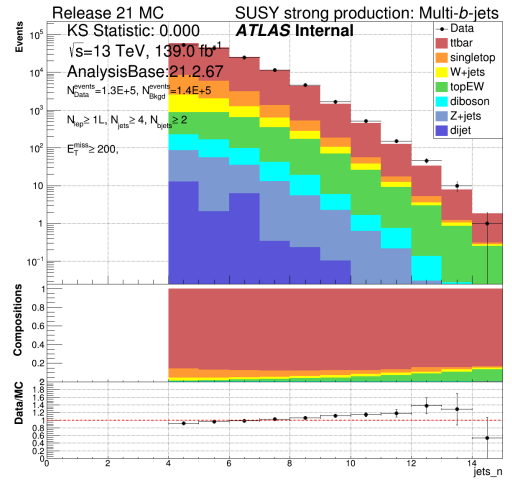
(d) N_{jets} in events with at least 1 signal lepton after reweighing

Figure 32: (32a , 32b)m_{eff} and (32c , 32d) N_{jets} before and after the kinematic reweighing . The new reweighing factors corrected disagreement in the data/MC m_{eff} distribution, but induced disagreement in the N_{jets} appearing as a positive slope in data/MC.

One way to protect the N_{jets} distribution from the induced disagreement is derivation of kinematic reweighing in the $\{m_{eff}, N_{jets}\}$ phase-space. Due to the limited statistic in the control regions, event-specific weights are derived in bins of $N_{jets} = 4$, 5, 6 and ≥ 7 for the $t\bar{t}$ and W+jets enriched regions and in the $N_{jets} \geq 4$ bin for the single top and Z+jets enriched regions. Similar to the original reweighing approach, the m_{eff} distribution divided into 12 segments of 300 GeV each covering the total range of 4200 GeV. The m_{eff} spectra for each control region and each N_{jets} bin are shown in figure 33.
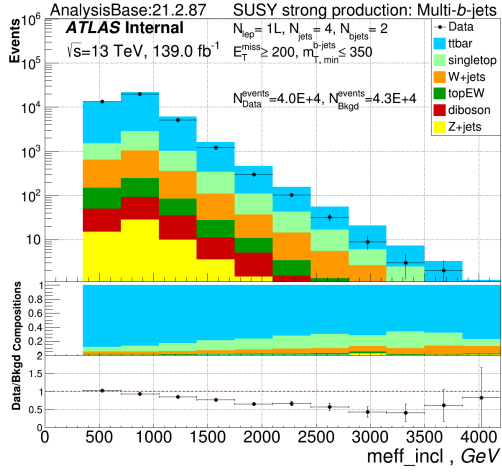
(a) $t\bar{t}$ CR, $N_{jets} = 4$
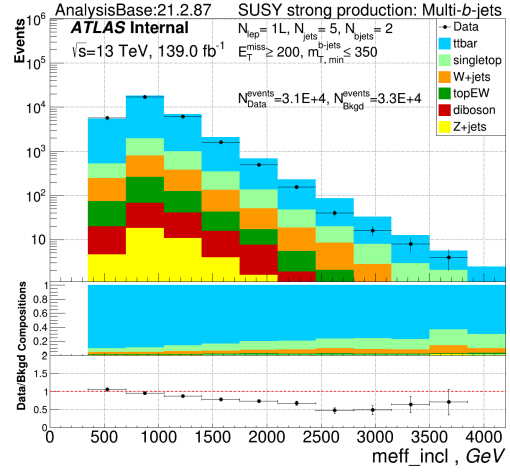
(b) $t\bar{t}$ CR, $N_{jets} = 5$
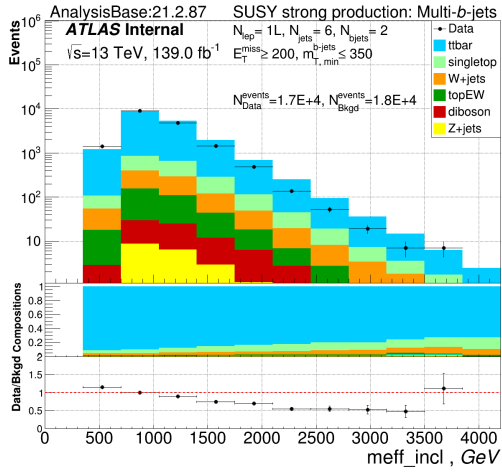
(c) $t\bar{t}$ CR, $N_{jets} = 6$
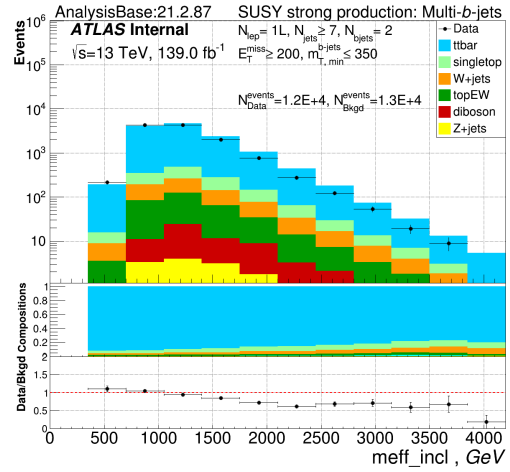
(d) $t\bar{t}$ CR, $N_{jets} \geq 7$
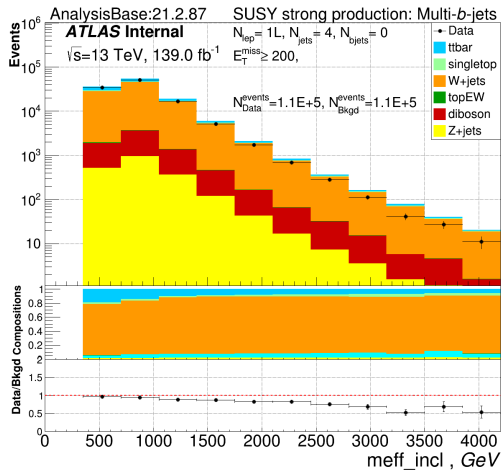
(e) W+jets CR, $N_{jets} = 4$
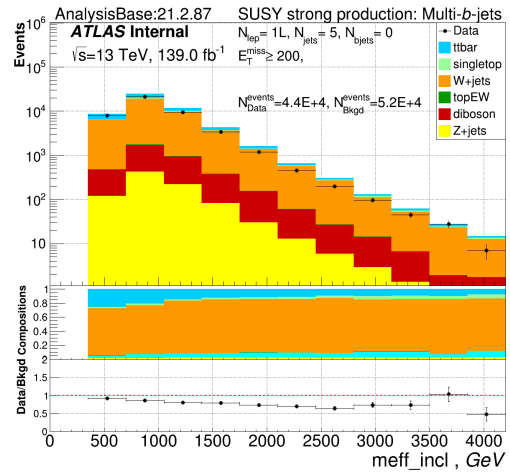
(f) W+jets CR with $N_{jets} = 5$
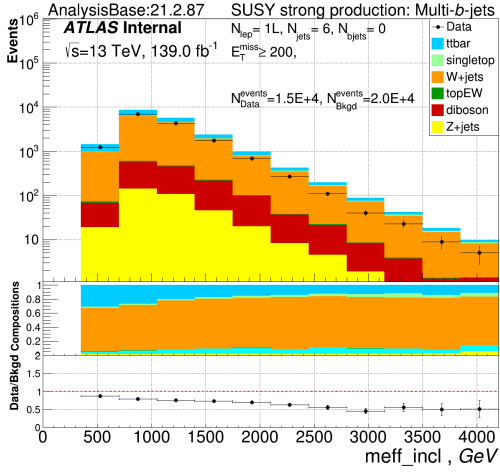
(g) W+jets CR, $N_{jets} = 5$
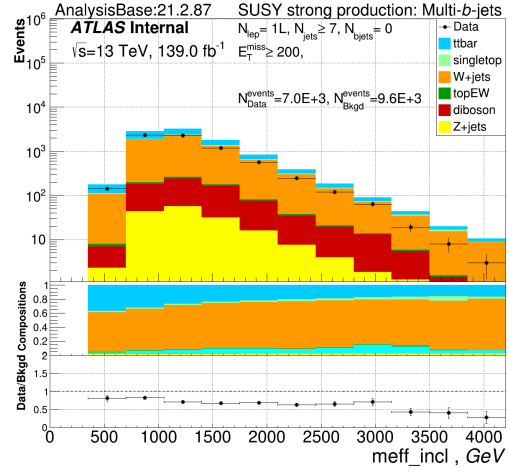
(h) W+jets CR, $N_{jets} \geq 7$
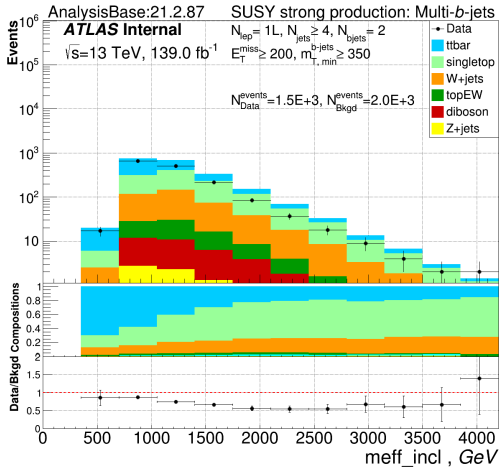
(i) single top CR
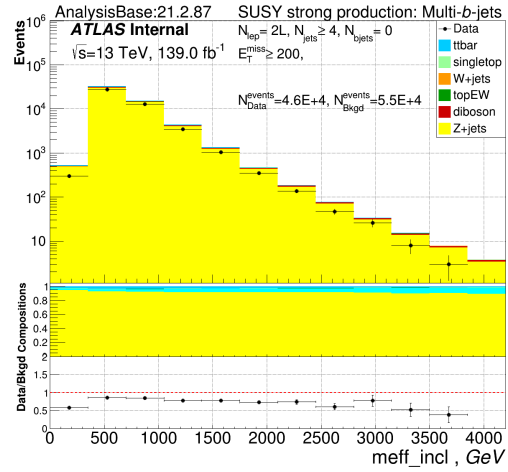
(j)]Z+jets CR

Figure 33: m_{eff} spectra for all four control regions, binned in N_{jets} for the $t\bar{t}$ and W+jets.

In the first step, reweighting factors for the $t\bar{t}$ and W+jets control regions are derived with respect to equation 6.2.2. The equation is solved for all m_{eff} segments for all the N_{jets} bins consequently. The data/MC ratios for the vector in the right side of the equation are estimated from fits. Fits are performed in the m_{eff} regions that satisfy two conditions:

- There are at least 10 events in both the data and the weighted MC in each consequent bin of the fit region.

- Both the data and the MC spectra follow exponential behavior in the fit region.

Given the solutions in all 12 m_{eff} segments for each N_{jets} bin, weight functions are fitted with exponential functions. The fits are performed in regions where the derived weights follow exponential behavior, and then extended to the $[0, \infty]$ range. Weights for events in the MC samples are then evaluated from the extended exponential functions $f(m_{eff})$. An

68

example of the data and the MC m_{eff} spectra in $N_{jets} = 4$ bin for the $t\bar{t}$ enriched control region with the data/MC ratio and the resulting weight function are given in figure 34.

In the second step, weights for the single top and Z+jets enriched regions are derived with respect to equation 6.2.3. The newly derived reweighting factors are applied to the $t\bar{t}$ and W+jets events.The procedure is done with the same approach as for the $t\bar{t}$ and W+jets control regions.

The results of the two-dimensional kinematic reweighing can be found in figure 35. As it follows from the figures, the m_{eff} data/MC ratio was corrected, while no significant disturbance was introduced to the N_{jets} distribution: the data/MC ratio follows flat pattern in the low-N_{jets} region, and ratio $= 1$ is covered by the uncertainty bars in the higher-N_{jets} region. Since the aim of the procedure was to correct shapes of analysis distributions but not the overall normalization, reweighing functions applied to each of the MC samples are scaled in a way to keep the overall yield unchanged. All functions for the kinematic reweighing are shown in Appendix 0.2.

$$
\begin{pmatrix} f_{CR-t\bar{t}}^{t\bar{t}} & f_{CR-t\bar{t}}^{W+jets} \\ f_{CR-W+jets}^{t\bar{t}} & f_{CR-W+jets}^{W+jets} \end{pmatrix} \begin{pmatrix} w^{t\bar{t}} \\ w^{W+jets} \end{pmatrix} = \begin{pmatrix} (\frac{Data}{Bkgd})_{CR-t\bar{t}} - f_{CR-t\bar{t}}^{single-top} - f_{CR-t\bar{t}}^{Z+jets} \\ (\frac{Data}{Bkgd})_{CR-W+jets} - f_{CR-W+jets}^{signle-top} - f_{CR-W+jets}^{Z+jets} \end{pmatrix}
$$
$$(6.2.2)$$

$$
\begin{pmatrix} f_{CR-ST}^{single-top} & f_{CR-ST}^{Z+jets} \\ f_{CR-Z+jets}^{single-top} & f_{CR-Z+jets}^{Z+jets} \end{pmatrix} \begin{pmatrix} w^{ST} \\ w^{Z+jets} \end{pmatrix} = \begin{pmatrix} (\frac{Data}{Bkgd})_{CR-ST} - f_{CR-single-top}^{t\bar{t}} - f_{CR-ST}^{W+jets} \\ (\frac{Data}{Bkgd})_{CR-Z+jets} - f_{CR-Z+jets}^{t\bar{t}} - f_{CR-Z+jets}^{W+jets} \end{pmatrix}
$$
$$(6.2.3)$$
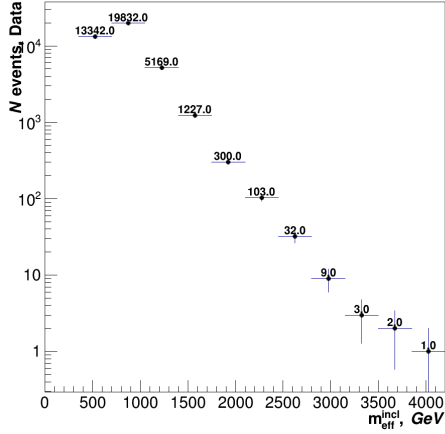
(a) Data m_{eff} spectrum in $t\bar{t}$ enriched control region with $N_{jets} = 4$
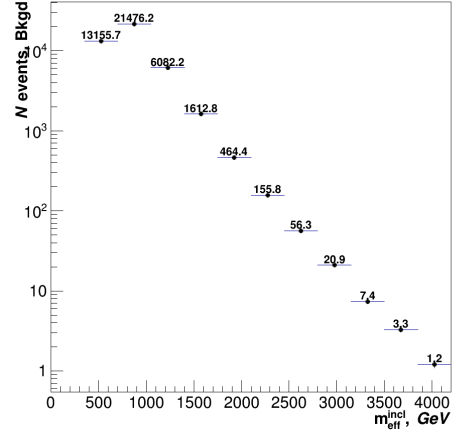
(b) Total MC m_{eff} spectrum in $t\bar{t}$ enriched control region with $N_{jets} = 4$
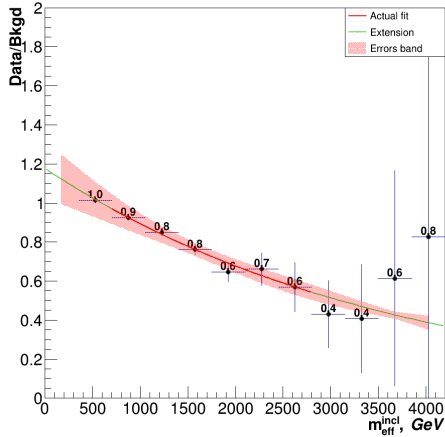
(c) Data/MC m_{eff} in $t\bar{t}$ enriched control region with $N_{jets} = 4$
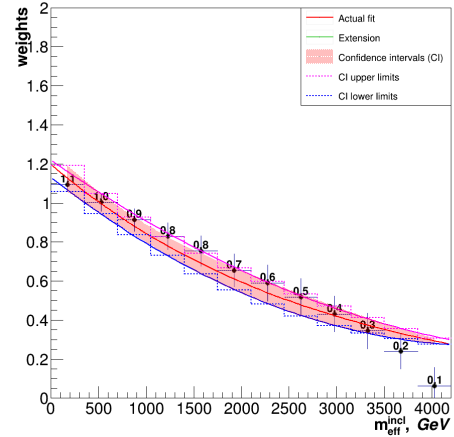
(d) Resulting weights function for $t\bar{t}$ events with $N_{jets} = 4$

Figure 34: (34a) Data and (34b) MC m_{eff} spectra, (34c) their ratio and (34d) resulting kinematic reweighing function derived in $t\bar{t}$ $N_{jets} = 4$ bin.
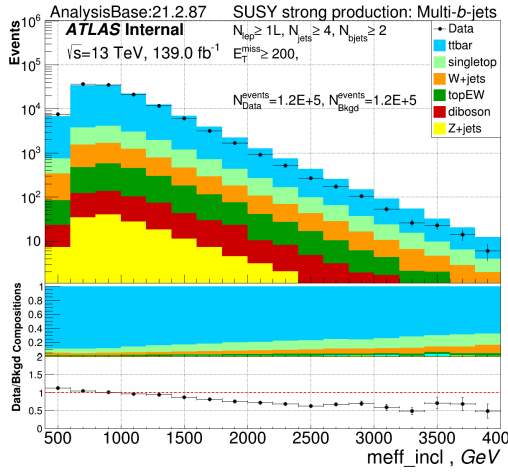
(a) m_{eff} in events with at least one signal lepton before 2D reweighing
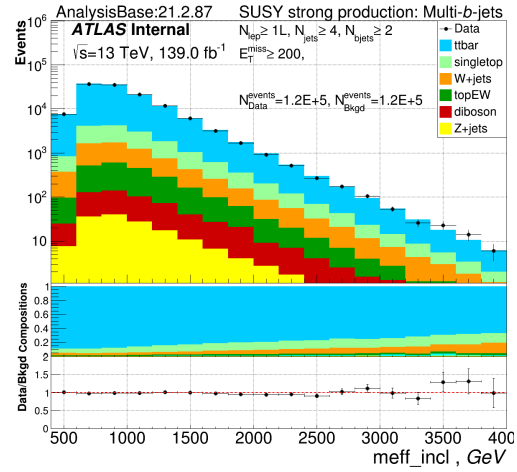
(b) m_{eff} in events with at least one signal lepton after 2D reweighing
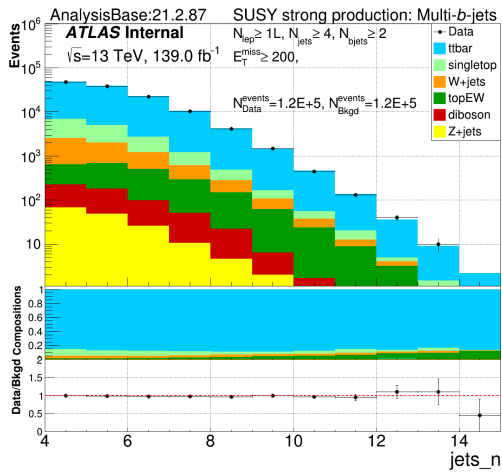
(c) N_{jets} in events with at least one signal lepton before 2D reweighing
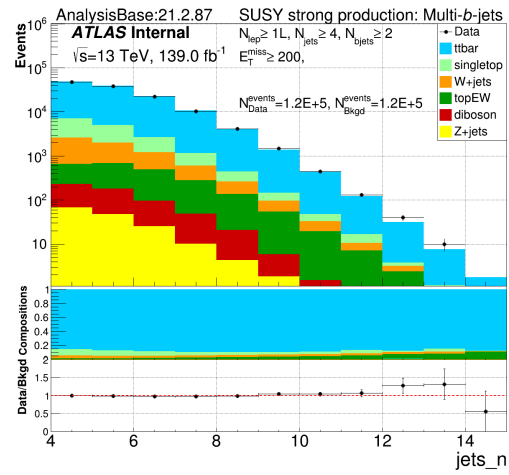
(d) N_{jets} in events with at least one signal lepton after 2D reweighing

Figure 35: (35a, 35b) m_{eff} and (35c, 35d) N_{jets} spectra before and after the 2D kinematic reweighing. While m_{eff} data/MC agreement was corrected, no significant disturbance was interceded to N_{jets}.

6.3 Cut-and-Count Analysis

This section discusses the cut-and-count analysis strategy and presents definitions of the Signal, Control and Validation regions for the analysts. General aspects of the regions definitions are established in section 6.3.1. Procedures and specific details of the regions definitions in the 0L and 1L channels are given in sections 6.3.2 and 6.3.3 respectively. The regions definitions are summarized in tables 6 and 7.

6.3.1 General Aspects

Signal Regions (SR). The cut-and-count analysis makes use of the comparison of the expected and observed numbers of events in the data after applying some selection criteria (cuts) on several analysis variables. The regions of the phase-space selected by these cuts are called Signal Regions. SRs are optimized in a way to achieve the greatest signal significance for each entry of the $\{m_{\tilde{g}};\ m_{\tilde{\chi}_1^0}\}$ model phase-space. The real data is not used in the SR construction process. Instead, only the SM background and signal MC simulations are used to estimate the number of expected background and signal events. The optimization is done using Histfitter [157] tool. The tool takes a user-provided list of variables with their minimal, maximal and step values, applies all possible combinations of selections, computes significance for each of the combinations and returns the set selections that results into the highest significance. In cases with more than one combination of selections resulting into the same highest signal significance, a region with $t\bar{t}$ production being the dominant background are preferred over others. The SRs are used for signal events searches in the real data. For the analysis of the real data it is extremely important to estimate the expected number of SM background events precisely.

Control Regions (CR). For correct estimation of the number of the SM background events, a normalization procedure is applied. The goal of the procedure is to derive scale factors for different SM background processes in so-called Control Regions, each of which corresponds to one SR. The CRs must be kinematically close but orthogonal to corresponding SRs. Each scale factor introduces additional systematic uncertainty to an analysis. Therefore, a balance between the accuracy of the number of renormalized background events and the total systematic uncertainty must be found. In the presented analysis, $t\bar{t}$ production is the dominant SM background, hence only this background will be renormalized. Each CR in the analysis must satisfy a set of conditions. The CRs are required to have $t\bar{t}$ purity of at least 80%, have at least 10 reweighted total background events and have signal contamination of 1% at most.

Validation Regions (VR). Before propagating to the SRs, the scale factors are tested in so-called Validation Regions, which are orthogonal to both corresponding SRs and CRs. The VRs are build from SRs definitions inverting one selection. The same requirements in $t\bar{t}$ purity and signal contamination as for CRs are applied. To meet the two criteria, selections

in E_{T}^{miss} or/and m_{eff} can be relaxed.

Tables 6 and 7 summarize definitions of SRs, CRs and VRs in the 0L and 1L channels respectively. The entire $m_{\tilde{g}}/m_{\tilde{\chi}_1^0}$ mass plane was divided into four kinematic regions based on similar selections in groups of SRs. The kinematic regions group SRs by mass splitting between the gluino and the neutralino $\Delta m = m_{\tilde{g}} - m_{\tilde{\chi}_1^0}$ as following:

- **Boosted** regions correspond to large Δm. In these regions events are characterized by large values of energy-related variables: E_{T}^{miss}, m_{eff}, $M_J^{\Sigma,4}$.

- **Moderate** regions target signal models with intermediate Δm. However, the range of the "intermediate" Δm is large, and optimization of the SRs yields two moderate subregions, named Moderate-1 for the larger Δm and Moderate-2 for the smaller Δm.

- **Compressed** regions target signal models close to the production limit of the particles of the final state: $m_{\tilde{g}} \approx m_{\tilde{\chi}_1^0} + 2 * m_{top\ quark}$.

Further sections describe derivation of selections for all regions in more details.

Gtt 0-lepton

Criteria common to all regions: $p_{\mathrm{T}}^{leading\ jet} \geq 30$ GeV

Targeted kinematics	Type	N_{lep}	N_{jets}	N_{b-jets}	E_{T}^{miss}	$\Delta\phi_{min}^{4j}$	m_{eff}	m_{T}	$m_{\mathrm{T,min}}^{b-jets}$	$M_J^{\Sigma,4}$
Region B (Boosted, Large Δm)	SR	$= 0$	≥ 5	≥ 3	≥ 600	≥ 0.4	≥ 2900	–	≥ 120	≥ 300
	CR	$= 1$	≥ 4	≥ 3	≥ 200	–	≥ 2000	< 200	–	≥ 150
	VR	$= 0$	≥ 5	≥ 3	≥ 250	≥ 0.4	≥ 2000	–	–	< 300
Region M (Moderate-1 Δm)	SR	$= 0$	≥ 9	≥ 3	≥ 600	≥ 0.4	≥ 1700	–	≥ 120	≥ 300
	CR	$= 1$	≥ 8	≥ 3	≥ 200	–	≥ 1100	< 150	–	≥ 150
	VR	$= 0$	≥ 9	≥ 3	≥ 300	≥ 0.4	≥ 1400	–	–	< 300
Region M (Moderate-2 Δm)	SR	$= 0$	≥ 10	≥ 3	≥ 500	≥ 0.4	≥ 1100	–	≥ 120	≥ 200
	CR	$= 1$	≥ 9	≥ 3	≥ 200	–	≥ 800	< 150	–	≥ 100
	VR	$= 0$	≥ 10	≥ 3	≥ 300	≥ 0.4	≥ 800	–	–	< 200
Region C (Compressed, small Δm)	SR	$= 0$	≥ 10	≥ 4	≥ 400	≥ 0.4	≥ 800	–	≥ 180	≥ 100
	CR	$= 1$	≥ 9	≥ 4	≥ 200	–	≥ 800	< 150	–	≥ 100
	VR	$= 0$	≥ 10	≥ 4	≥ 200	≥ 0.4	≥ 800	–	–	< 100

Table 6: Definitions of the Gtt 0-lepton SRs, CRs and VRs of the cut-and-count analysis. All kinematic variables are expressed in GeV except $\Delta\phi_{min}^{4j}$, which is in radians. The $p_{\mathrm{T}}^{leading\ jet}$ requirements is also applied to b-tagged jets.

Gtt 1-lepton

Criteria common to all regions: ≥ 1 signal lepton, $p_T^{leading\ jet} \geq 30$ GeV, $N_{b-jets} \geq 3$

Targeted kinematics	Type	N_{jets}	E_T^{miss}	m_{eff}	m_T	$m_{T,min}^{b-jets}$	$M_J^{\Sigma,4}$
Region B (Boosted, Large Δm)	SR	≥ 4	≥ 600	≥ 2300	≥ 150	≥ 120	≥ 200
	CR	$= 4$	≥ 200	≥ 1500	< 150	$-$	$-$
	VR-m_T	≥ 4	≥ 200	≥ 1500	≥ 150	$-$	< 200
	VR-$m_{T,min}^{b-jets}$	> 4	≥ 200	≥ 1200	< 150	≥ 120	< 200
Region M (Moderate-1 Δm)	SR	≥ 5	≥ 600	≥ 2000	≥ 200	≥ 120	≥ 200
	CR	$= 5$	≥ 200	≥ 1200	< 200	$-$	$-$
	VR-m_T	≥ 5	≥ 200	≥ 1200	≥ 200	$-$	< 200
	VR-$m_{T,min}^{b-jets}$	> 5	≥ 200	≥ 1000	< 200	≥ 120	≥ 100
Region M (Moderate-2 Δm)	SR	≥ 8	≥ 500	≥ 1100	≥ 200	≥ 120	≥ 100
	CR	$= 8$	≥ 200	≥ 800	< 200	$-$	$-$
	VR-m_T	≥ 8	≥ 200	≥ 800	≥ 200	$-$	< 100
	VR-$m_{T,min}^{b-jets}$	> 8	≥ 200	≥ 800	< 200	≥ 120	≥ 100
Region C (Compressed, small Δm)	SR	≥ 9	≥ 300	≥ 800	≥ 150	≥ 120	$-$
	CR	$= 9$	≥ 200	≥ 800	< 150	$-$	$-$
	VR-m_T	≥ 9	≥ 200	≥ 800	≥ 150	< 120	$-$
	VR-$m_{T,min}^{b-jets}$	> 9	≥ 200	≥ 800	< 150	≥ 120	$-$

Table 7: Definitions of the Gtt 1-lepton SRs, CRs and VRs of the cut-and-count analysis. All kinematics variables are expressed in GeV except $\Delta\phi_{min}^{4j}$, which is in radians. The $p_T^{leading\ jet}$ requirements is also applied to b-tagged jets.

6.3.2 0L Channel Regions

Signal regions. Selections with their minimal, maximal and step values that were used in the signal grid scan in the 0L channel are summarized in table 8. Two selections were fixed: $\Delta\phi_{min}^{4j} \geq 0.4$ radians and $p_T^{leading\ jet} \geq 30$ GeV. To make $t\bar{t}$ production the dominant background, for signal models with smaller Δm the maximal selection in $m_{T,min}^{b-jets}$ was set to 180 GeV, and 120 GeV for signal models with larger values of Δm.

Gtt 0-lepton								
Parameters:	N_{jets}	N_{b-jets}	$p_{\mathrm{T}}^{leading\ jet}$	E_{T}^{miss}	$\Delta\phi_{min}^{4j}$	m_{eff}	$m_{\mathrm{T,min}}^{b-jets}$	$M_J^{\Sigma,4}$
Min value:	4	3	30	200	0.4	800	0	0
Max value:	10	4	X	600	X	3200	120/180	400
Step:	1	1	X	100	X	300	40	100

Table 8: Ranges of selections in variables used in Gtt 0L SR optimization. All kinematic variables are expressed in GeV except for $\Delta\phi_{min}^{4j}$ which is in radians.

An example of the resulting m_{eff} selections from the most optimal sets for each signal model are given in figure 36, and the rest of the plots are given in Appendix 0.3. The selections follow diagonal pattern which is dictated by Δm. Definitions of four kinematic regions discussed at the end of section 6.3.1 follow directly from the figure:

- Boosted: $\Delta m \geq 1800$ GeV;

- Moderate-1: $900 \leq \Delta m < 1800$ GeV;

- Moderate-2: $400 < \Delta m < 900$ GeV;

- Compressed: $\Delta m \leq 400$ GeV.

Figure 37 shows the number of expected signal events and estimated significance for each evaluated signal model. Some signal models return the estimated significance above 5σ. These signal models would be considered as candidates for discovery in High Energy Physics. However, previous studies [158] set an exclusion limit that covers the models with that high expected significances.
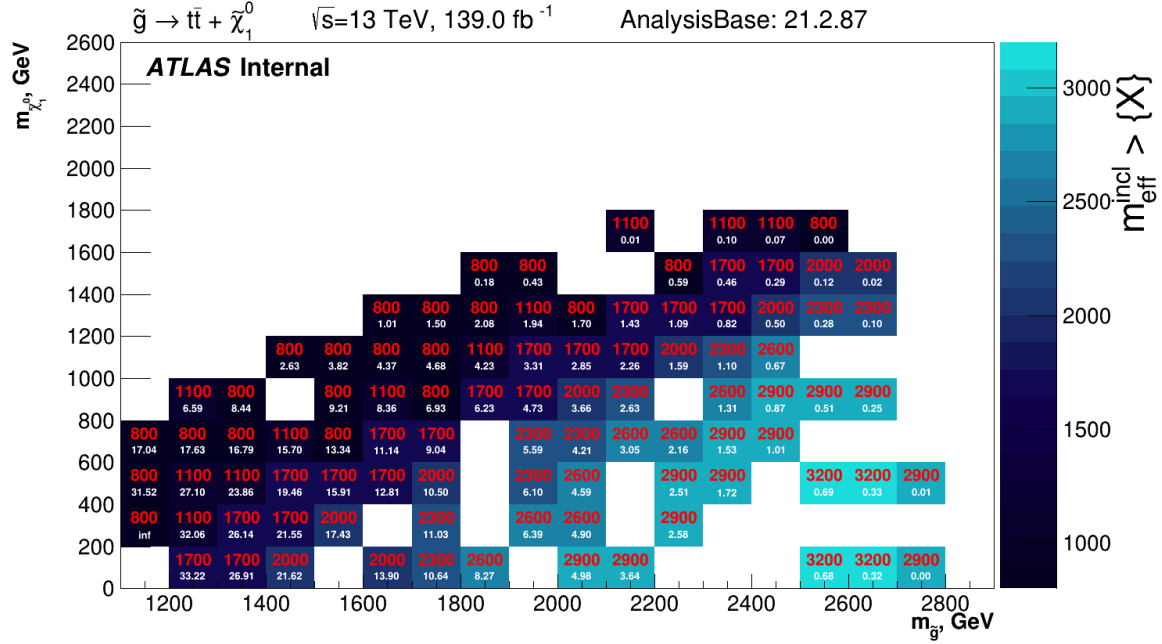
Figure 36: The most optimal cuts for m_{eff} in the 0L channel for SRs. Numbers in red represents selection in m_{eff} and number is white - the estimated signal significance.
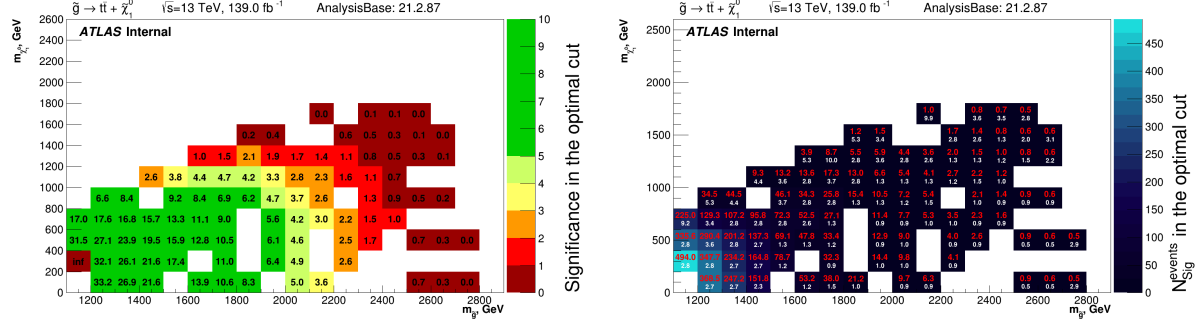
(a) Expected significance

(b) Expected number of events

Figure 37: (a) The expected significances for each model of the Gtt 0L signal grid, (b) The expected number of signal events (in red) and backround events (in white).

Control regions. Semileptonic $t\bar{t}$ events, in which the lepton is outside the acceptance or is hadronically decaying τ-lepton, dominate in the SRs. For this reason, the Gtt 0L control regions make use of the 1L channel, requiring the presence of exactly one signal lepton and decreasing the number of jets by 1. The requirement in exactly 1 signal lepton makes the CRs orthogonal to the Gtt 0L SRs. An inverted selection on m_T is applied to exclude any overlapping with the Gtt 1L SRs. The same requirement in m_T ensures a small signal contamination in the CRs. Selection on $\Delta\phi_{min}^{4j}$ is removed, and selection on kinematic variables are relaxed to ensure at least 100 reweighted events in each CR. So-called "N-1" plots of all the 0L channel CRs binned in E_T^{miss} can be found in Figure 38.
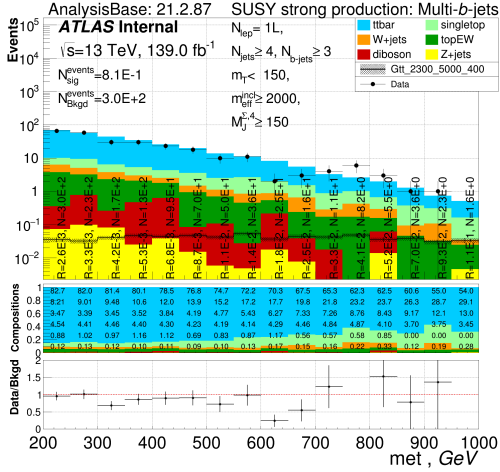
(a) Boosted Δm region
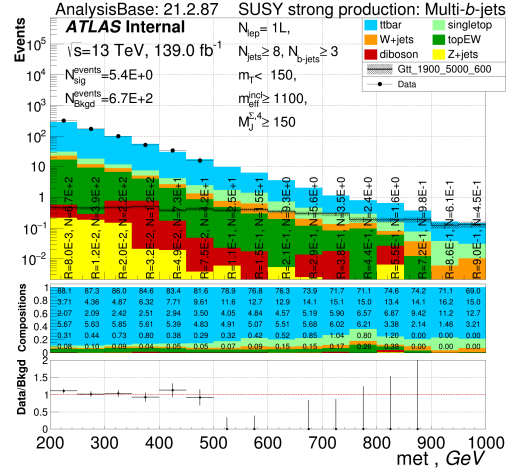
(b) Moderate-1 Δm region
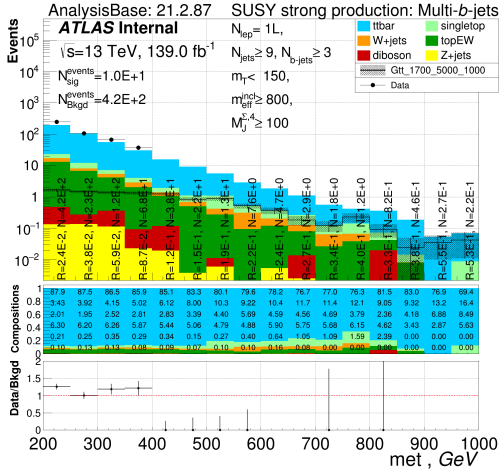
(c) Moderate-2 Δm region
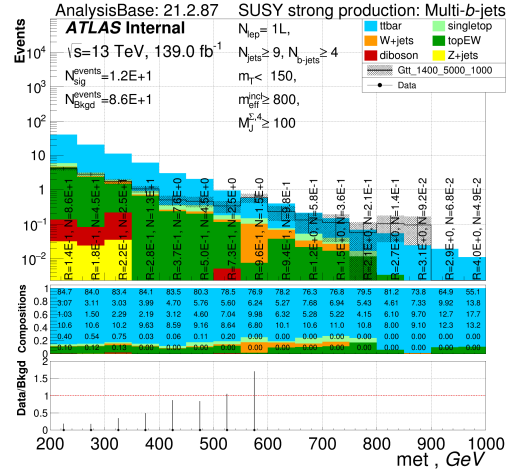
(d) Compressed Δm region

Figure 38: N-1 plots for the Gtt 0L CRs: (a) Boosted, (b) Moderate-1, (c) Moredate-2 and (d) Compressed Δm regions. Top canvas: normalized to unity background (in color), data (black dots) and signal (black line with shaded errors band) distributions; vertical numbers in bins indicate (R) signal contamination and (N) number of background events as if the '$E_T^{miss} \geq$' selection was set at the bin. Middle canvas: background compositions; numbers in the bins indicate compositions as if the '$E_T^{miss} \geq$' selection was set at the bin. Bottom canvas: data to background ratio per bin.

Validation regions are constructed to validate the background prediction in the high-m_T region. An inverted selections on $M_J^{\Sigma,4}$ and signal leptons veto makes the VRs mutually exclusive with the SRs and the CRs. To keep the VRs kinematically close to the corresponding SRs, selections on N_{jets}, N_{b-jets} and $\Delta\phi_{min}^{4j}$ are taken to be the same as in the SRs. To achieve the highest possible number of background events and preserve small signal contamination, selection on $m_{T,min}^{b-jets}$ is removed, and E_T^{miss} and m_{eff} selections are relaxed. N-1 plots of all the VRs for the Gtt 0L channel binned in E_T^{miss} can be found in Figure 39.
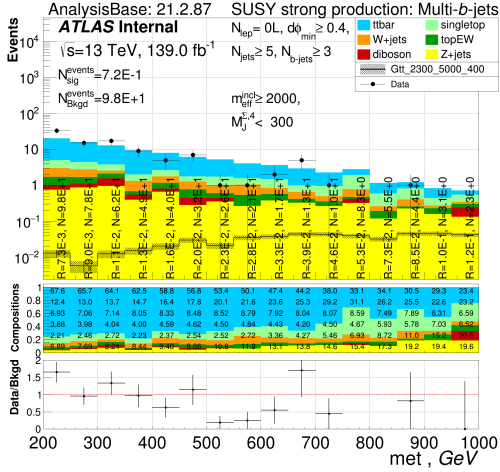
(a) Boosted Δm region
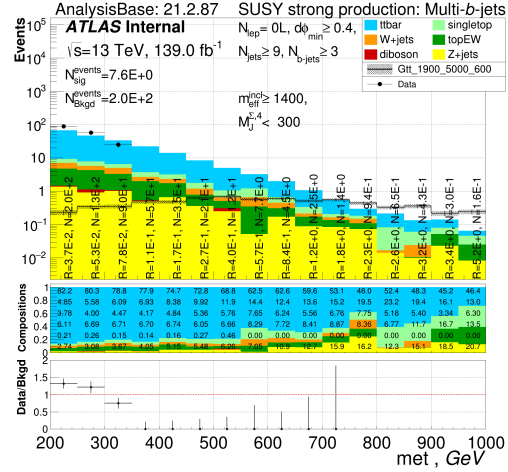
(b) Moderate-1 Δm region
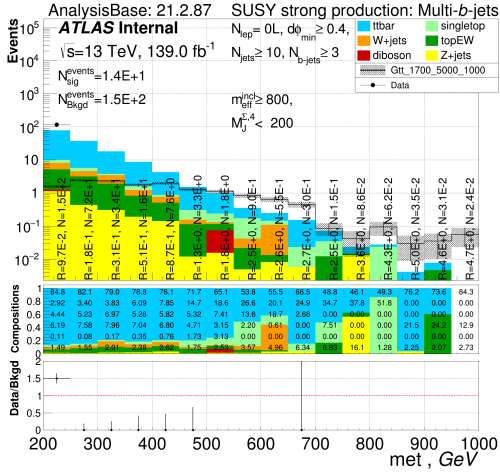
(c) Moderate-2 Δm region
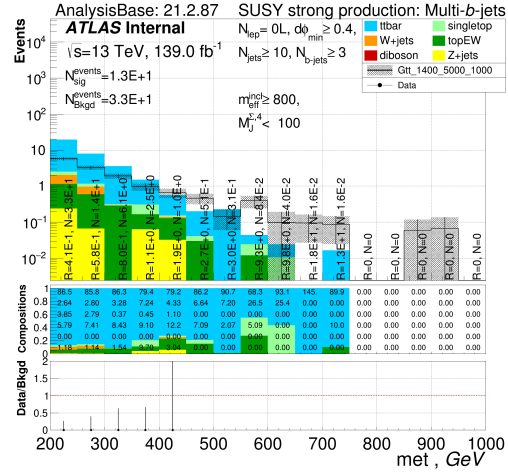
(d) Compressed Δm region

Figure 39: N-1 plots for the Gtt 0L VRs: (a) Boosted, (b) Moderate-1, (c) Moredate-2 and (d) Compressed Δm regions. Top canvas: normalized to unity background (in color), data (black dots) and signal (black line with shaded errors band) distributions; vertical numbers in bins indicate (R) signal contamination and (N) number of background events as if the '$E_T^{miss} \geq$' selection was set at the bin. Middle canvas: background compositions; numbers in the bins indicate compositions as if the '$E_T^{miss} \geq$' selection was set at the bin. Bottom canvas: data to background ratio per bin.

6.3.3 1L Channel Regions

Signal regions. Signal regions were optimized in the 1L channel using the same strategy as in the 0L channel. A set of variables with their minimal, maximal and step values are summarized in table 9. Three selections were fixed: $p_T^{leading\ jet} \geq 30$ GeV, $m_{T,min}^{b-jets} \geq 120$ GeV and $N_{b-jets} \geq 3$.

Gtt 1-lepton								
Parameters:	N_{jets}	N_{b-jets}	$p_\text{T}^{leading\ jet}$	E_T^{miss}	m_{eff}	m_T	$m_\text{T,min}^{b-jets}$	$M_J^{\Sigma,4}$
Min value:	4	3	30	200	800	100	0	0
Max value:	10	X	X	600	3200	200	120	300
Step:	1	X	X	100	300	50	40	100

Table 9: Ranges of selections in variables used in Gtt 1L SR optimization. All kinematic variables are expressed in GeV.

An example of the resulting m_{eff} selections from the most optimal sets for each signal model are given in figure 40, and the rest of the plots are given in Appendix 0.3. As in the 0L channel, the most optimal sets of selections follow diagonal pattern in the signal grid and form 4 kinematic regions with similar selection within each of them:

- Boosted: $\Delta m \geq 1700$ GeV;

- Moderate-1: $1000 \leq \Delta m < 1700$ GeV;

- Moderate-2: $400 < \Delta m < 1000$ GeV;

- Compressed: $\Delta m \leq 400$ GeV.

Figure 41 shows the number of expected signal events and estimated significance for each evaluated signal model. Analogues to the 0L channel, signal models that return the estimated significance above 5σ are excluded by the previous studies [158].
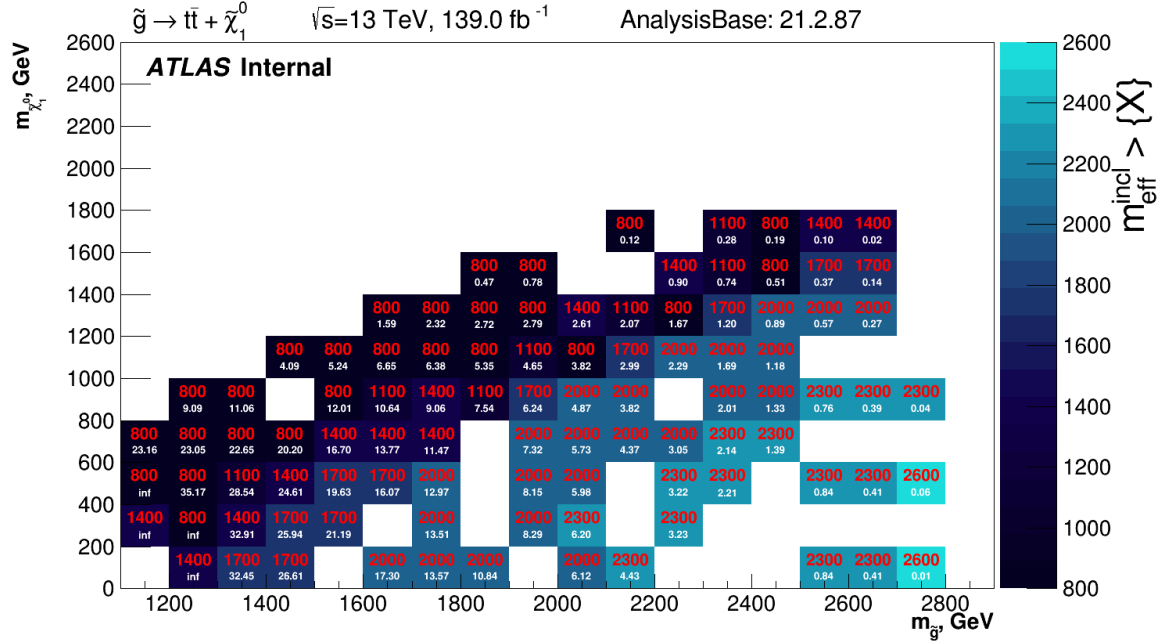
Figure 40: The most optimal cuts for m_{eff} in the 1L channel for SRs. Numbers in red represents selection in m_{eff} and number is white - the estimated signal significance.
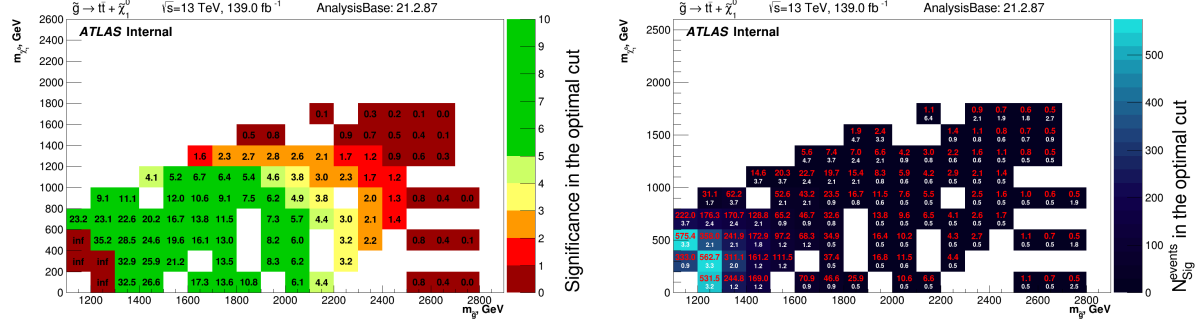
(a) Expected significance

(b) Expected number of events

Figure 41: (a) The expected significances for each model of the Gtt 1L signal grid, (b) The expected number of signal events (in red) and backround events (in white).

Control regions. An inverted selection on m_T is applied to make CRs orthogonal with the corresponding SRs. Selections in N_{jets} are set to be exactly the lowest value of the selection in the corresponding SRs in order to leave a degree of freedom for orthogonality with Gtt 1L VRs. Selections on $m_{T,min}^{b-jets}$ and $M_J^{\Sigma,4}$ are removed and selections on other kinematic variables are relaxed to ensure at least 100 reweighed events in each CR. N-1 plots of all the CRs for Gtt 1L channel binned in E_T^{miss} are shown in figure 42.
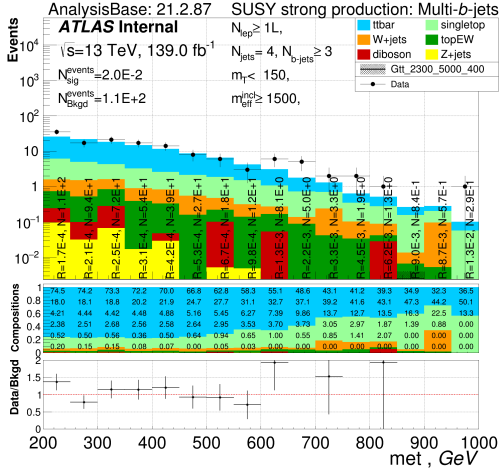
(a) Boosted Δm region
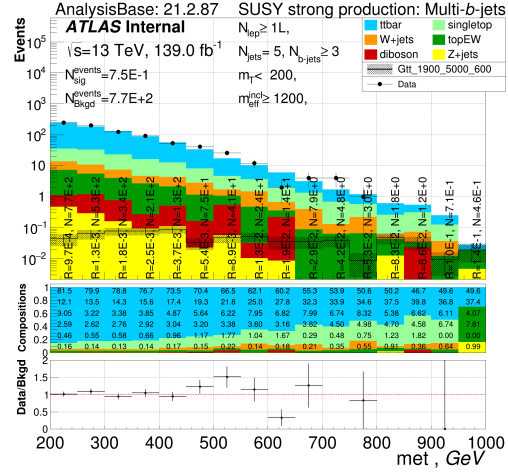
(b) Moderate-1 Δm region
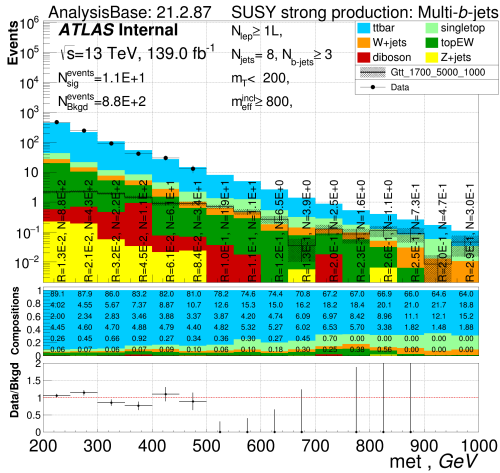
(c) Moderate-2 Δm region
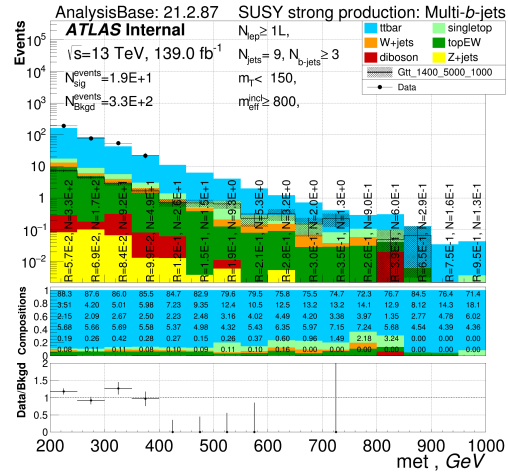
(d) Compressed Δm region

Figure 42: N-1 plots for the Gtt 1L CRs: (a) Boosted, (b) Moderate-1, (c) Moredate-2 and (d) Compressed Δm regions. Top canvas: normalized to unity background (in color), data (black dots) and signal (black line with shaded errors band) distributions; vertical numbers in bins indicate (R) signal contamination and (N) number of background events as if the '$E_T^{miss} \geq$' selection was set at the bin. Middle canvas: background compositions; numbers in the bins indicate compositions as if the '$E_T^{miss} \geq$' selection was set at the bin. Bottom canvas: data to background ratio per bin.

Validation regions. Two types of validation are constructed in the 1L channel. The regions, named VR-m_T and VR-$m_{T,min}^{b-jets}$, validate the background predictions with scale factors applied to $t\bar{t}$ in the high-m_T and high $m_{T,min}^{b-jets}$ regims correspondingly. VR-m_T is kept mutually exclusive with the corresponding SR and CR by an inverted selection on $M_J^{\Sigma,4}$ or $m_{T,min}^{b-jets}$. Orthogonality of VR-$m_{T,min}^{b-jets}$ to other regions is achieved by adjusted selections on m_T and N_{jets}. Other kinematic variables are relaxed to ensure that there are more than 100

81

reweighed events in the VRs. N-1 plots of VR-m_T for all four kinematic regions for the Gtt 1L channel binned in E_T^{miss} can be found in figure 43 and for VR-$m_{T,min}^{b-jets}$ - in figure 44.

(a) Boosted Δm region

(b) Moderate-1 Δm region

(c) Moderate-2 Δm region

(d) Compressed Δm region
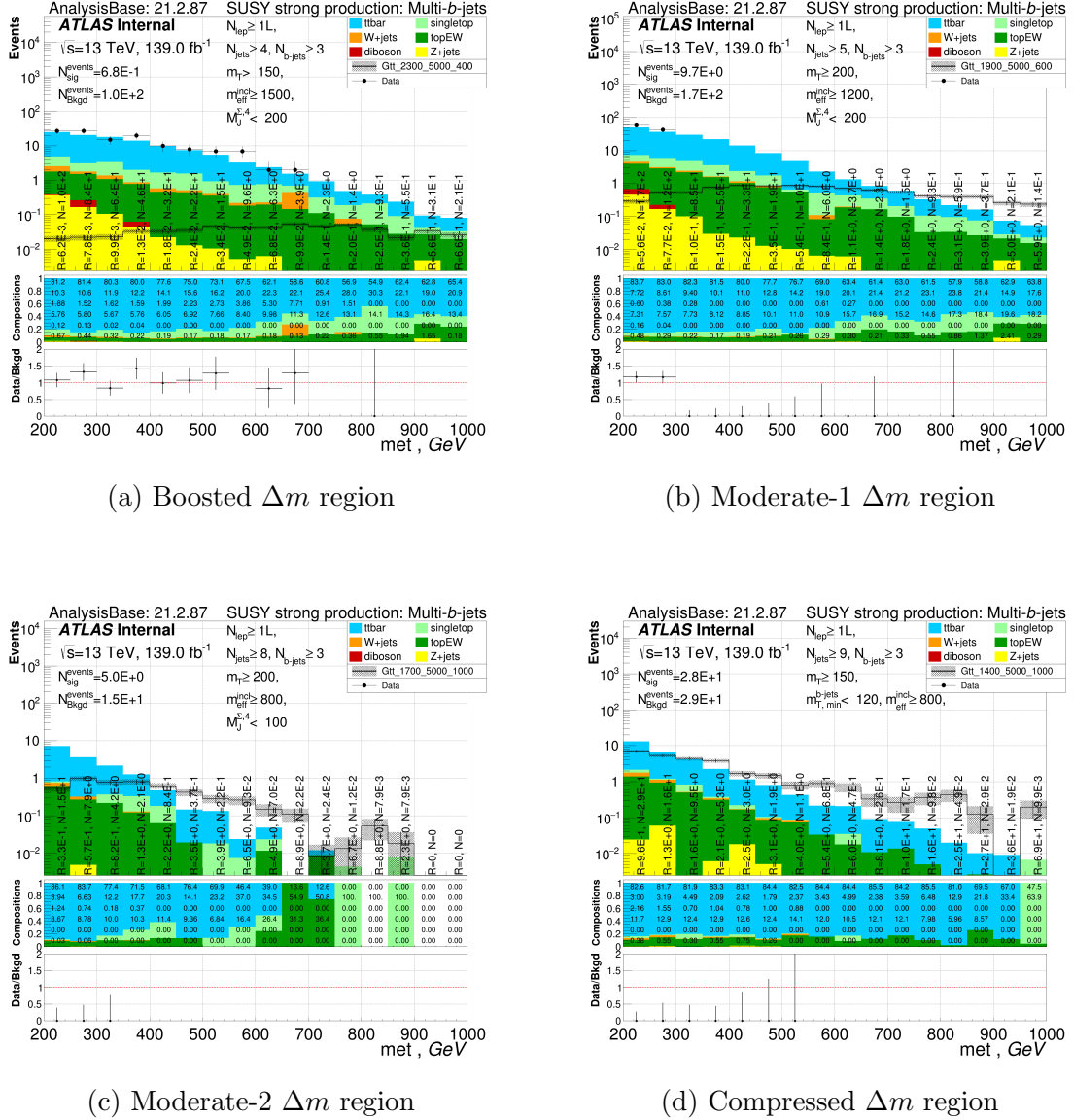
Figure 43: N-1 plots for the Gtt 1L VR-m_T: (a) Boosted, (b) Moderate-1, (c) Moderate-2 and (d) Compressed Δm regions. Top canvas: normalized to unity background (in color), data (black dots) and signal (black line with shaded errors band) distributions; vertical numbers in bins indicate (R) signal contamination and (N) number of background events as if the '$E_T^{miss} \geq$' selection was set at the bin. Middle canvas: background compositions; numbers in the bins indicate compositions as if the '$E_T^{miss} \geq$' selection was set at the bin. Bottom canvas: data to background ratio per bin.
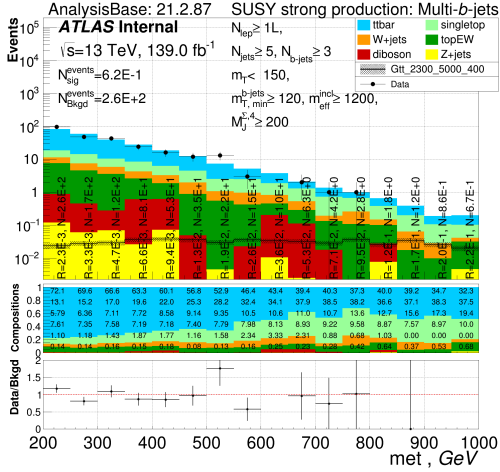
(a) Boosted Δm region
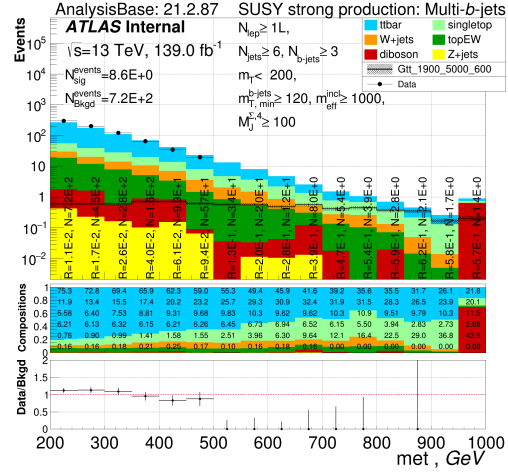
(b) Moderate-1 Δm region
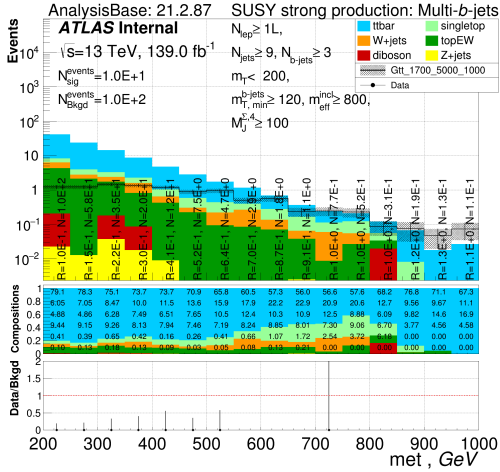
(c) Moderate-2 Δm region
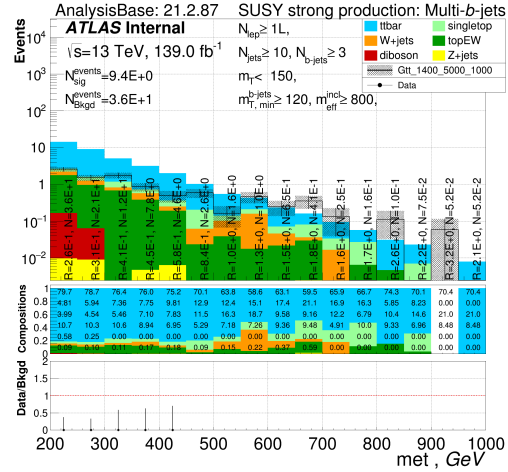
(d) Compressed Δm region

Figure 44: N-1 plots for the Gtt 1L VR-$m_{\mathrm{T,min}}^{b-jets}$: (a) Boosted, (b) Moderate-1, (c) Moredate-2 and (d) Compressed Δm regions. Top canvas: normalized to unity background (in color), data (black dots) and signal (black line with shaded errors band) distributions; vertical numbers in bins indicate (R) signal contamination and (N) number of background events as if the '$E_{\mathrm{T}}^{miss} \geq$' selection was set at the bin. Middle canvas: background compositions; numbers in the bins indicate compositions as if the '$E_{\mathrm{T}}^{miss} \geq$' selection was set at the bin. Bottom canvas: data to background ratio per bin.

6.4 Analysis Uncertainties

This section covers information about uncertainties that affect the analysis. There are two categories of the uncertainties: statistical and systematic. The statistical uncertainties are due to finite statistic in the data and the MC samples. The uncertainties are estimated as square root of the total number of events $\sqrt{N}$. The systematic uncertainties are categorized

into experimental and theoretical. The experimental systematic uncertainties are detector-related, they come from the imperfection of modeling of the detector performance. These uncertainties are discussed in section 6.4.1. The theoretical systematic uncertainties depend on theoretical modeling and discussed in section 6.4.2. Tables with breakdowns of uncertainties on background estimates are summarized in Appendix 0.5.

6.4.1 Experimental Systematic Uncertainties

Each of the objects in the MC simulations are calibrated to accurately describe the data. The recommendations from the dedicated performance groups are used to estimate modeling uncertainties. Following the recommendations, variations of systematic parameters by 1σ up and down are used to estimate the effect of the systematics. All sources of the experimental systematic uncertainties that affect the analysis are summarized below.

Jet Kinematics. One the largest sources of systematic uncertainties in the analysis are due to jets kinematics, such as Jet Energy Scale (JES), Jet Energy Resolution (JER) and Jet Mass Scale (JMS). By applying the JES and JMS energies and masses of the reconstructed jets are corrected to the particle level. JER is estimated from the dijet events using the asymmetry distribution [159] between the reference and the probe jets. Sets if nuisance parameters (NP) are used for each of the uncertainties calculations.

b−tagging efficiency. To correct the difference between the b-tagging efficiency in MC simulation and data, a scale factor is applied to each simulated event. The b-tagging uncertainty is evaluate by varying p_{T}, η and truth flavor dependent scale factors. The variations are applied separately to b, c and *light* jets resulting into three uncorrelated uncertainties.

Lepton-related uncertainties. Following lepton-related uncertainties are taken into account in the analysis: electron energy scale (1 NP), electron energy resolution (1 NP), muon momentum scale (1 NP), muon momentum resolution (2 NPs). The systematic uncertainties affecting the lepton efficiencies and isolation are accounted for and found to be negligible in the analysis.

$E_{\mathrm{T}}^{\mathrm{miss}}$ Track Soft Term (TST) uncertainties. The uncertainties associated to E_{T}^{miss} TST [160] are categorized into scale and resolution. The variations of the objects described above are also propagated to the E_{T}^{miss} systematic. The uncertainties are negligible in the analysis.

Kinematic reweighing uncertainty. The uncertainty is estimated from the data to background ratios. As discussed in section 6.2.3, for each bin of m_{eff} distribution a system of equations is needed to be solved. The data to total MC ratios in the right side were treated as Gaussian functions as shown in equation 6.4.1.

$$\frac{data}{total\ MC} = p_0 * exp(-\frac{1}{2} * (\frac{x - p_1}{p_2})^2) \qquad (6.4.1)$$

where $p_0 = 1/(2\pi\sigma)$, p_1 is the mean value, i.e. the data to the total MC ratio, and p_2 is the σ, i.e. the uncertainty on the data to the total MC ratio. The systems of equations from section 6.2.3 are solved 5000 times for each bin of m_{eff} distribution each time with a new random value taken from the Gaussian distribution of the data to MC ratios. The 5000 resulting solutions for kinematic weights in each bin of the m_{eff} distribution are fitted by the Gaussian distribution with p_1 being the mean value of the resulting kinematic weight in the bin and $p_2 -$ the resulting uncertainty of the value. The final exponential fit is performed on the p_1 values with their 1σ uncertainties (i.e 67% CI) form the Gaussian fits of solutions for kinematic weights. Thus, for each kinematic reweighing function there is a known 1σ variation band, and the systematic uncertainty is estimated as the 1σ up/down variation in the band.

QCD uncertainties. 300% uncertainty on the QCD background is applied in all regions.

Luminosity. The systematic uncertainty on the integrated luminosity for the full Run-2 ($139~fb^{-1}$) is 1.7% [33] obtained using the LUCID-2 detector [161].

6.4.2 Theory Uncertainties

The theory systematic uncertainties are estimated by comparison between the nominal MC samples and the alternative generators. The theory systematics are estimated for all the MC background processes (i.e. for all except for QCD which is produced with the data-driven method) and the signal.

$t\bar{t}$-production related uncertainties. Three components of $t\bar{t}$ MC production were taken into account in the estimation of the uncertainties: the generator, the parton shower and the Initial State Radiation (ISR). Uncertainties associated with the generator, the nominal Powheg+Pythia8 sample was compared to Madgraph5+Pythia8. To estimate the effect of the showering uncertainty, the nominal sample was compared to the Powheg+Herwig7. The ISR uncertainty was estimated from the nominal sample by varying the h_{damp} parameter [162] which is responsible for the matrix element to parton shower matching.

Another source of theory uncertainties is the normalization factor derived specifically for each of the signal regions of the analysis. The uncertainty is represented by a transfer factor which is computed as the ratio of SR over CR yields (or VR to CR when the uncertainty is evaluated for validation regions). For each SR and VR, the relative uncertainty on the transfer factor was computed for the generator, parton shower and ISR variations. Then, for each SR and VR the three components are summed up in quadratures resulting in a single nuisance parameter.

An additional uncertainty was applied to the $t\bar{t}b\bar{b}$ and $t\bar{t}+c\bar{c}$ components by varying their normalization by 30%.

The summary of the numerical values for the $t\bar{t}$ modeling uncertainties in all regions of the analysis is shown in figure 45.

85
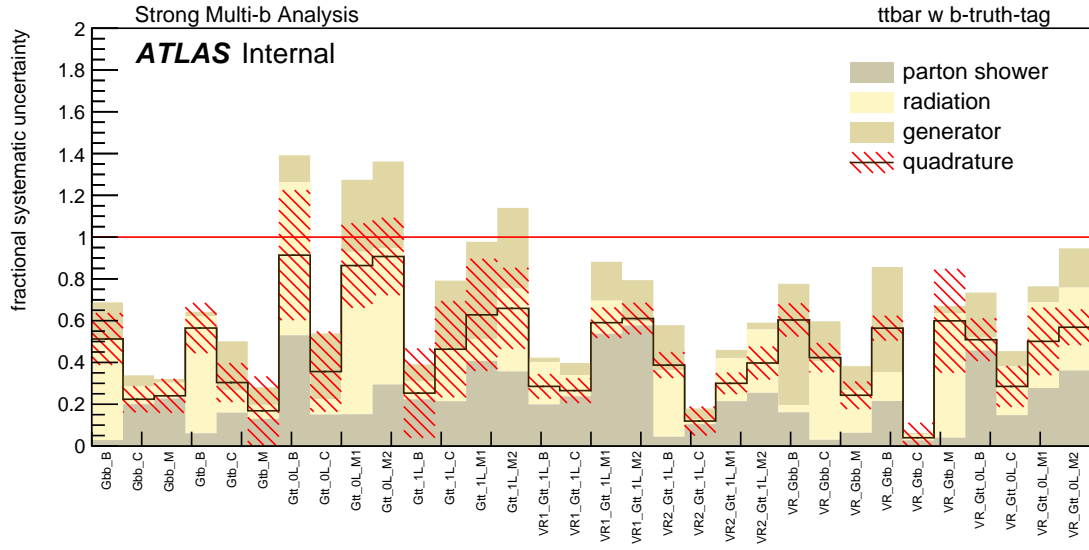
Figure 45: Summary of the $t\bar{t}$ production modeling uncertainties.
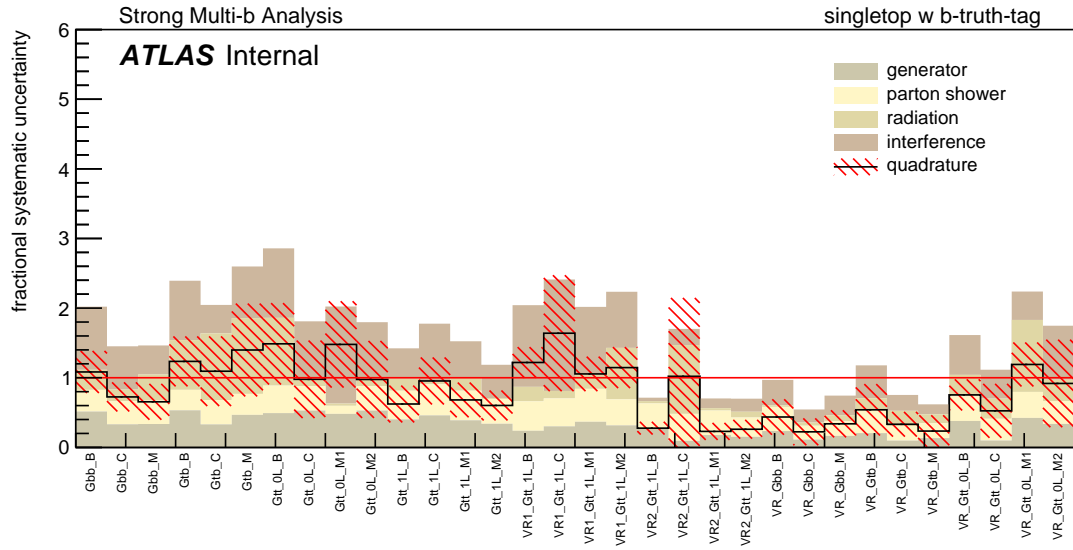
Figure 46: Summary of the single top production modeling uncertainties.

Single top production related uncertainties. As in case with the $t\bar{t}$ production, three components of single top MC production are considered: the generator, the parton showering the the radiation. An additional uncertainty component due to the interference between Wt and $t\bar{t}$ was also considered and estimated by comparing the yields of $WWbb$ to the $Wt + t\bar{t}$ truth samples. The numerical values of the uncertainty was computed with formula 6.4.2. The four components are added up in quadrature for each SR and VR of the analysis. Figure 46 shows he summary of the numerical values for the single top modeling uncertainties.

$$1 - \frac{WWbb - t\bar{t}}{Wtb} \tag{6.4.2}$$

W/Z+jets production related uncertainties. The nominal samples for the V+jets backgrounds were generated using Sherpa 2.2.1. The uncertainties on the renormalization μ_R and factorization μ_F scales [163] were configured in the generator. The two variations are correlated, and 6 sets of the variations were compared to the nominal values of the parameters resulting into 6 independent nuisance parameters.

t$\bar{t}$+X and diboson production related uncertainties. A systematic uncertainty of 50% was applied to the background sources following the previous experience [158].

Signal models production uncertainties. A few sources of uncertainties in the signal modeling were considered: factorization and renormalization (one "up" and "down" pair), merging scale (one "up" and "down" pair), parton shower tuning and radiation uncertainty (five "up" and "down" pairs), resulting into 14 variations. To evaluate the uncertainties, samples for each variation were produced and compared to the nominal. The uncertainties were combined following PDF4LHC recommendations [164].

6.5 Results

This chapter presents the statistical analysis of the data in regions defined in Chapter 6.3. The analysis strategy follows two-step procedure. A background-only fit is performed in CRs to derive normalization factors for $t\bar{t}$ background - $\mu_{t\bar{t}}$. Then the normalization factors are propagated to corresponding SRs, and the data to the total SM background ratios are studied for potential excess. If a significant excess of data events is observed, a discovery can be claimed. Otherwise, an exclusion fit is performed using the CL_s [165] to test which signal models can be excluded by the analysis.

6.5.1 Likelihood Function

The likelihood function $L(\boldsymbol{\theta}) = P(\mathbf{x} \mid \boldsymbol{\theta})$ for a set of parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$ answers the questions what is the probability that $\mathbf{x}$ is coming from a probability distribution $P(\mathbf{X}, \boldsymbol{\theta})$ given that we measured $\mathbf{X} = \mathbf{x}$.

The likelihood function used for the background-only fit is defined in equation 6.5.1:

$$L = \prod_{CR} P_{CR} \times C_{syst} \tag{6.5.1}$$

where the probability densities P_{CR} are Poisson distributions as shown in equation 6.5.2.

$$P_{CR} = Pois(n_{CR}^{observed} \mid n_{CR}^{expected}) \tag{6.5.2}$$

and the expected number of events $n_{CR}^{expected}$ is the sum of the renormalized number of $t\bar{t}$ events with other backgrounds:

$$n_{CR}^{expected} = (\mu_{t\bar{t}} \cdot n_{CR}^{t\bar{t}} + n_{CR}^{other}) \tag{6.5.3}$$

The expected yields can change due to the effects of applied systematic uncertainties. A profiling technique [10] is applied to express the likelihood as a function of the parameters of interest only. The parameters of interest are nuisance parameters mentioned in the Experimental Systematic Uncertainties section (6.4.1). The expected number of events takes form of:

$$n^{expected} = n^{expected} \cdot (1 + \sum_i \kappa_i \cdot \alpha_i) \tag{6.5.4}$$

where κ_i equals to the effect on the $n^{expected}$ of a one-sigma deviation in the α_i nuisance parameter. The values for α_i are set by a fit procedure and constrained by the C_{syst} term that is given by formula 6.5.5.

$$C_{syst} = \prod_i N(0 \mid \alpha_i, \ 1) \tag{6.5.5}$$

The result of the fit are the most probable values of the nuisance parameters given the observed data. The resulting nuisance parameters are applied in equation 6.5.4, and the number of observed events is compared to the number of expected events. If no significant is observed, an exclusion fit is performed taking into account the signal regions:

$$L = \prod_{SR} P_{SR} \cdot \prod_{CR} P_{CR} \cdot C_{syst} \tag{6.5.6}$$

Similar to the background-only fit, the profiling procedure is performed to express the likelihood as a function of the parameters of interest. In the exclusion fit, the parameter of interest is the signal strength μ_{sig}. From the Neyman-Pearson lemma [166], the most statistically powerful test statistic is the profile likelihood ratio with the $q(\mu_{sig})$ [167] variable:

$$q_{\mu_{sig}} = -2 \cdot \ln \frac{\hat{L}(\mu_{sig})}{\hat{L}} \tag{6.5.7}$$

where $\hat{L}(\mu_{sig})$ is the result of the profiling fit with a fixed value of μ_{sig}, while $\hat{L}$ allows μ_{sig} to vary to maximize the likelihood. The $q_{\mu_{sig}}$ values are computed for all considered signal models. Models with small $q_{\mu_{sig}}$ values (i.e. $\hat{L}(\mu_{sig}) \sim \hat{L}$) are called signal-like, and

models with large $q_{\mu_{sig}}$ values (i.e. $\hat{L}(\mu_{sig}) >> \hat{L}$) are called background-like. Results of the hypothesis test are used to produce 95% confidence level exclusion contours in the $(m_{\tilde{g}}, m_{\tilde{\chi}_1^0})$ phase-space using CL_s prescription.

6.5.2 Background Only Fit

The background-only fit results are shown in figures 47-49. Figure 47 shows the observed numbers data events and the expected numbers of the SM background events from the MC simulations in the control regions of the analysis, and the resulting normalization factors $\mu_{t\bar{t}}$ that vary from 0.9 to 1.2 depending on the region. The data to MC comparisons after applying the scale factors in the validation regions are shown in figures 48 and 49 for the 0L and 1L channels respectively. There is no significant deviation of the rescaled background from the observed data. Therefore, the scale factors can be propagated to the signal regions. Figure 50 shows comparisons of the numbers of the observed data and the total expected background events in the signal regions of the analysis. The fit results show no significant excess of data above the expected background. Therefore, a discovery cannot be claimed.
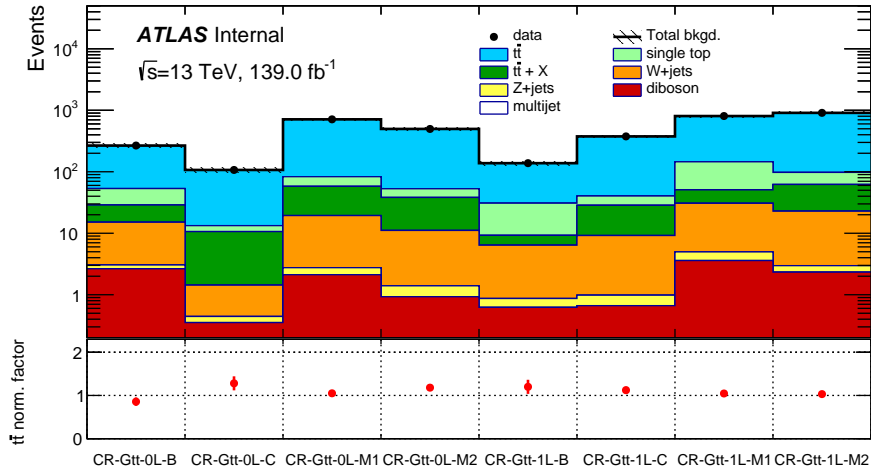
Figure 47: Results of the background-only fit in the control regions. Top panel: the observed number of data events shown in black dots, the expected number of background events shown in colors. Bottom panel: resulting normalization factors $\mu_{t\bar{t}}$ for $t\bar{t}$ background.
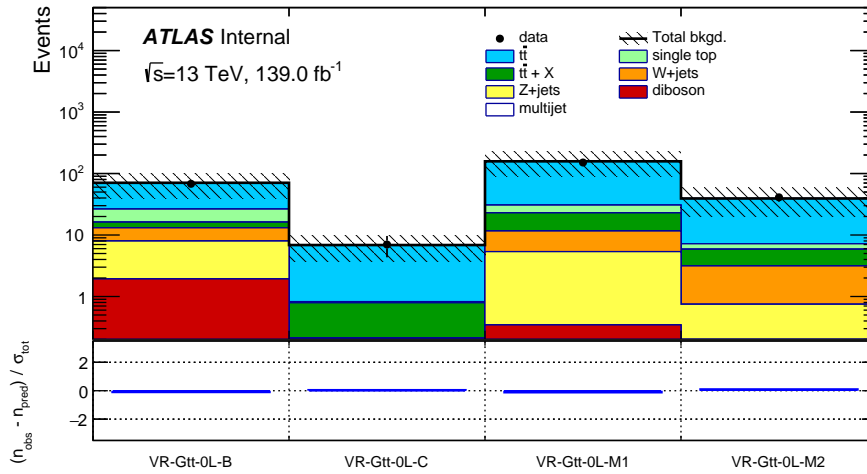
Figure 48: Results of the background-only fit in 0L channel validation regions. Top panel: the observed number of data events shown in black dots, the expected number of background events shown in colors. Bottom panel: difference between the observed number of events and the total expected background divided by the total uncertainty.
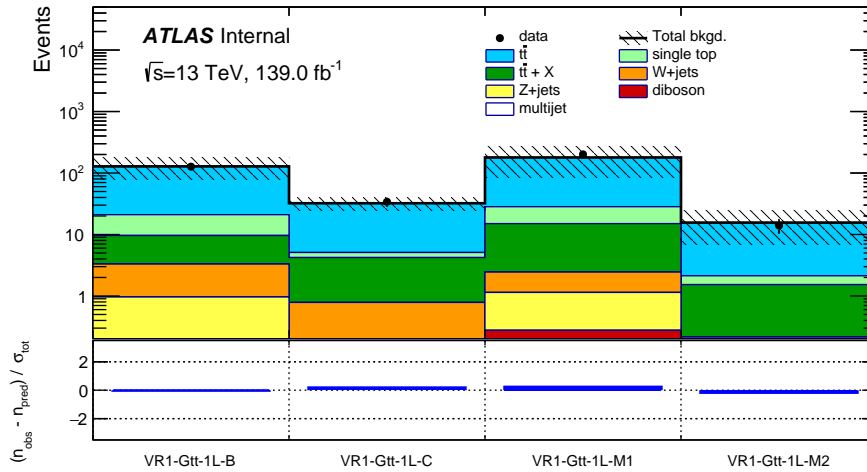
(a)
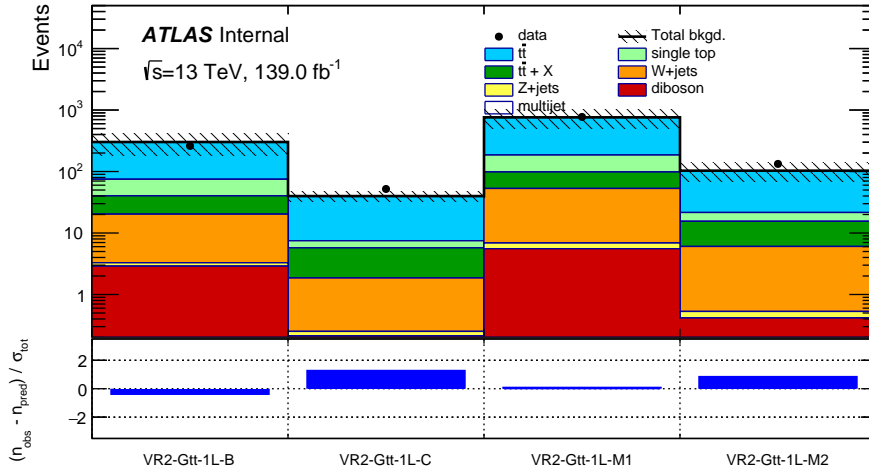
(b)

Figure 49: Results of the background-only fit in 1L channel validation regions: (a) VR-m_T and (b) VR-$m_{T,\min}^{b-jets}$. Top panel: the observed number of data events shown in black dots, the expected number of background events shown in colors. Bottom panel: difference between the observed number of events and the total expected background divided by the total uncertainty.
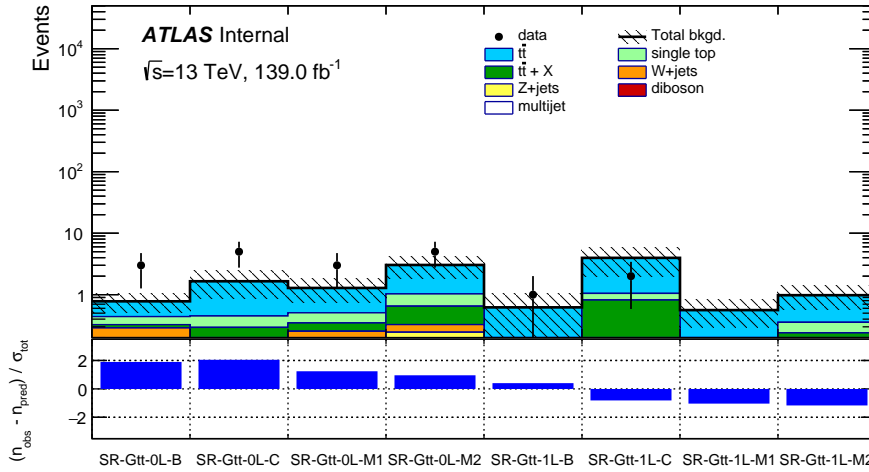
Figure 50: Results of the background-only fit signal regions. Top panel: the observed number of data events shown in black dots, the expected number of background events shown in colors. Bottom panel: difference between the observed number of events and the total expected background divided by the total uncertainty.

91

6.5.3 Model Dependent Upper Limit

Since no significant excess of the data above the expected background is observed in the signal regions, the data is used to derive exclusion limits for each signal region of the analysis, and the limits are further combined into one shown in figure 51. The highest excluded gluino mass $m_{\tilde{g}}$ is approximately 2240 GeV which corresponds to the massless neutralino, and the highest excluded neutralino mass $m_{\tilde{\chi}_1^0}$ is 1300 GeV which corresponds to approximately $m_{\tilde{g}} = 2100$ GeV.

The result derived using a dataset corresponding to the integrated luminosity of 139 fb^{-1} shows improvement compared to the previous iteration of the analysis [158] that used the total dataset of 79.9 fb^{-1}. The observed limit in $m_{\tilde{g}}$ was pushed by approximately 90 GeV, and the limit in $m_{\tilde{\chi}_1^0}$ by approximately 100 GeV. Compared to the results of similar searches obtained by the CMS collaboration [168] that used a dataset corresponding to the integrated luminosity of 137 fb^{-1}, the observed limit in $m_{\tilde{g}}$ is higher by approximately 90 GeV and by 100 GeV in $m_{\tilde{\chi}_1^0}$.
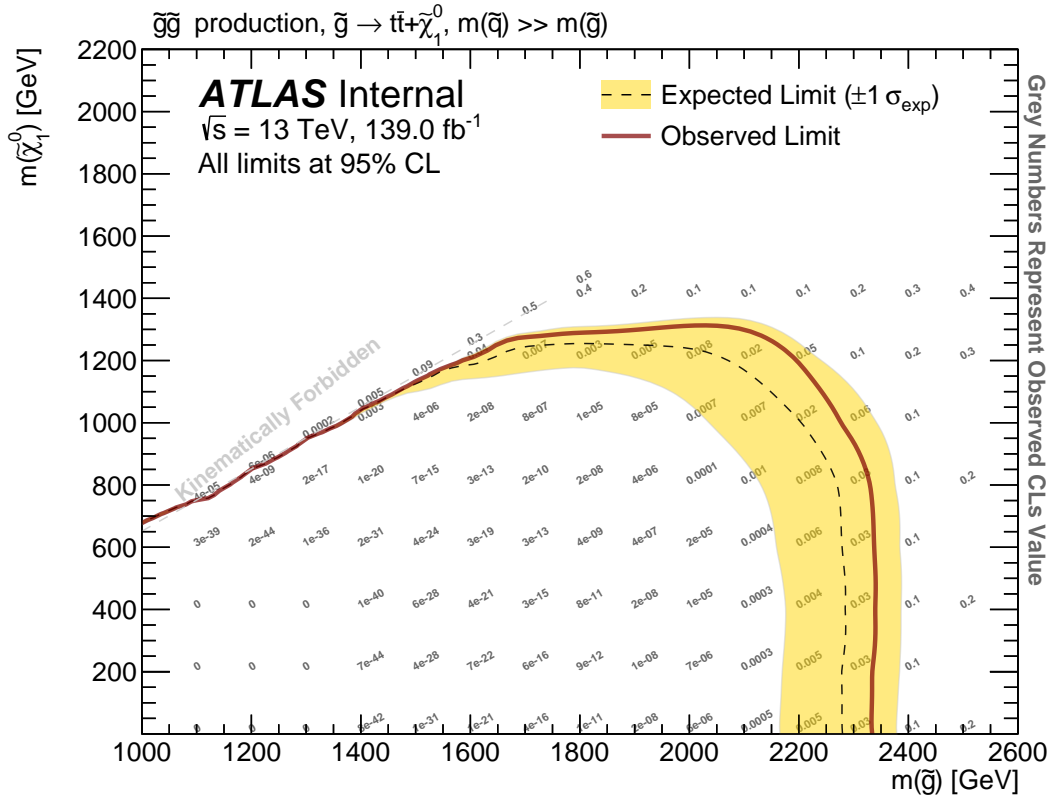
Figure 51: Exclusion limit for Gtt signal models. Models with $m_{\tilde{g}} < 2240$ GeV and $m_{\tilde{\chi}_1^0} < 1300$ GeV are excludes, i.e. no excess of data above combined MC was found. Models with higher masses of the gluino and neurtalino cannot be excluded due to the limited statistic in data.

CHAPTER VII

Differential Cross-Section Measurements of $t\bar{t}$ Production with Additional Heavy Flavor Jets

Production of top-antitop quarks pairs in association with additional heavy flavor jets ($t\bar{t}b$) is the dominant background for many BSM physics analyses [169] as well as for some of the SM measurements. A particular example of a BSM analysis with $t\bar{t}b$ being the dominant background is presented in Chapter VI of the dissertation, and an example of the SM study is the production of the Higgs boson in the $t\bar{t}H(b\bar{b})$ channel [170]. The uncertainties associated with the $t\bar{t}b$ modeling are usually the largest in the analyses. Therefore, a good understating of the $t\bar{t}b$ process is needed.

Theory predictions for the additional heavy flavor jets in the $t\bar{t}b$ production are poor. To model $t\bar{t}b$ production, event generators are usually employed. Measurements of the $t\bar{t}b$ production can be used to test various $g \to b\bar{b}$ splitting models, therefore they can provide a feedback for event generators tuning.

The analysis targets $t\bar{t}b$ events with exactly two leptons of different flavors (electron and muon) of the opposite signs; the channel is denoted as $e\mu - OS$. The selection of dilepton channel eliminates the contamination of events with additional jets from W hadronic decay. Requirement in the different flavors of the leptons significantly minimizes the $Z+jets$ production contamination.

Section 7.1 presents the data and MC samples used in the analysis. Section 7.2 explains the detector-level object and event selection, and the particle level definitions. An overview of the analysis signal regions and observables is presented in section 7.3. Section 7.4 shows the detector-level and MC prediction comparisons. Section 7.5 presents an origin-identification algorithm for b-jets derived using Boosted Decision Trees (BDT) technique. Unfolding results are presented in section 7.6. Section 7.7 discusses systematic uncertainties. And section 7.8 summarizes the results of the analysis.

7.1 Data and Monte-Carlo Samples

The study uses the total amount of the data from pp collisions at the centre-of-mass energy of 13 TeV in the 25 ns bunch spacing configuration recorded by the ATLAS detector during the $2015 - 2018$ data-taking period which corresponds to an integrated luminosity of 139.0 fb^{-1}.

7.1.1 Signal Modeling

The signal process in the study is the $t\bar{t}$ production in association with additional jets. The nominal sample was generated using the Powheg Box next-to-leading order (NLO) generator [171, 135, 134, 172] with the NNPDF3.0NLO [173] set in the matrix element calculation. Pythia 8.230 was subsequently used for parton shower, fragmentation and underlying event simulations. A set of alternative samples are generated with one change in modeling configuration of the nominal samples and are used to estimate systematic uncertainties due to the particular choice of the nominal MC modeling. Table 10 summarizes the nominal and alternative setups for the signal events modeling.

Some modeling parameters remain the same for all the samples. The mass of the top quark is set to 172.5 GeV. All the samples are normalized to the cross-section prediction at the next-to-next-to-leading order (NNLO) in QCD. The normalization also includes the resummation of the next-to-next-to-leading logarithmic (NNLL) soft-gluon terms calculated using TOP++ 2.0 [174, 175, 176, 177, 178, 147, 179]. For pp collisions at a centre-of-mass energy of 13 TeV, this cross-section corresponds to 832 ± 51 pb [179]. The total uncertainty in the theoretical cross-section comes from variations of factorisation, renormalization scales, variations in PDFs and α_s [180, 181, 182, 183, 132]. The independent components of the uncertainty are added in quadrature resulting into the total uncertainty on the theoretical cross-section.

MC samples	Generator	Matrix-element order	PDF set hard process	Parton Shower	Matching	Tune
inclusive $t\bar{t}$ MC						
Powheg+Pythia8	Powheg Box v2	$t\bar{t}$ NLO	NNPDF3.0NLO	Pythia 8.230	Powheg $h_{damp} = 1.5 m_{top}$	A14
Powheg+Pythia8 h_{damp}	Powheg Box v2	$t\bar{t}$ NLO	NNPDF3.0NLO	Pythia 8.230	Powheg $h_{damp} = 3.0 m_{top}$	A14
Powheg+Herwig 7.0.4 [184]	Powheg Box v2	$t\bar{t}$ NLO	NNPDF3.0NLO	Herwig 7.0.4	Powheg $h_{damp} = 1.5 m_{top}$	H7UE
Powheg+Herwig 7.1.3 [185]	Powheg Box v2	$t\bar{t}$ NLO	NNPDF3.0NLO	Herwig 7.1.3	Powheg $h_{damp} = 1.5 m_{top}$	Herwig7.1
aMC@NLO+Phytia8	MC@NLO v2.6.0	$t\bar{t}$ NLO	NNPDF3.0NLO	Pythia 8.230	MC@NLO	A14
aMC@NLO+Herwig 7.1.3	MC@NLO v2.6.0	$t\bar{t}$ NLO	NNPDF3.0NLO	Herwig 7.1.3	MC@NLO	Herwig7.1
Sherpa 2.2.10	Sherpa 2.2.10	$t\bar{t}$+1jet NLO +4 jets LO	NNPDF3.0NNLO	Sherpa	MEPS@NLO	Sherpa

Table 10: Summary of MC setups used for modelling the $t\bar{t}$ signal for the data analysis and for comparisons to the unfolded data.

7.1.2 Background Modeling

All the background samples in the study, except for the fake lepton and the miss-identified b-jet in $t\bar{t}$, are estimated from the MC modeling. Table 11 summarizes setups for various backgrounds.

The non-prompt and fake leptons backgrounds are estimated using a partially data-driven approach in events with the electron and muon of the same sign. Correction factors derived in the same sign region are applied in the $e\mu - OS$ signal region. These factors are determined from non-allhadronic MC simulations.

In some $t\bar{t}$ production events additional light or c-jets can be miss-tagged as additional b-jets. To take the effect into account, scale factors to correct for the composition of light-jets and c-jets in the $t\bar{t} + b$ jets events were derived using data-driven technique.

Process	Generator	Type	Version	PDF	Tune
$t\bar{t} + V$	MC@NLO + Pythia8	NLO + PS	2.3.3 8.230	NNPDF3.0NLO NNPDF2.3LO	- A14
$t\bar{t} + H$	Powheg Box + Pythia 8	NLO + PS	v2 8.230	NNPDF3.0NLO NNPDF2.3LO	- A14
Single top	Powheg Box + Pythia 8	NLO + PS	v2 8.230	NNPDF3.0NLO NNPDF2.3LO	- A14
W/Z+jets	Sherpa	$\leq 2j$ @ NLO, $\leq 4j$ @LO + PS	2.2.1	NNPDF3.0NNLO	Sherpa tune
WW, ZZ, WZ	Sherpa	$\leq 1j$ @ NLO, $\leq 3j$ @LO + PS	2.2.2	NNPDF3.0NNLO	Sherpa tune

Table 11: MC samples used for modeling the different backgrounds.

7.2 Objects Definitions and Event Selection

There are two levels at which the objects and events are defined: the Detector-level, also known as "reco" level, and the Particle or "truth" level (PL).

7.2.1 Detector-Level Object Selection

- **Primary vertex.** Interaction vertices are recostructed from at least two two tracks with $p_{\mathrm{T}} > 0.4$ GeV. The vertex with the largest $\sum p_{\mathrm{T}}^2$ from all associated tracks is called the primary vertex. Event with zero reconstructed vertices are removed from consideration.

- **Jets** are reconstructed with a particle flow algorithm [186] using topological clusters with three-dimensional clustering of energy deposits measured by calorimeter cells and tracking information resulting in so-called "pflow"-jets. The pflow-jets are subsequently used as an input to the anti-t_T sequential recombination algorithm with a radius of $R = 0.4$. In the last step energy of the reconstructed jets is calibrated to account for experimental effects using a combination of MC-based and in situ techniques [153]

  The calibrated jets are required to have $p_{\mathrm{T}} > 25$ GeV and $\mid \eta \mid < 2.5$. Low energetic jets with $p_{\mathrm{T}} < 60$ GeV and $\mid \eta \mid < 2.4$ are also required to originate from the primary vertex.

- **b-tagged jets** are identified using the DL1r b-tagger at 77% working point.

- **Leptons.**

  - **Elections** are reconstructed using the tracker and the electromagnetic calorimeter information. The reconstructed electrons are required to have $p_T > 25$ GeV, $| \eta | < 2.47$, $| d_0/\sigma_{d_0} | < 5$, $| \Delta z \sin(\eta) | < 0.5$ mm and originate from the primary vertex. Electron candidates that fall in the transition region between the barrel and the endcup $(1.37 < | \eta | < 1.52)$ are not considered in the study due to poor reconstruction performance in the region.

  - **Muons** are reconstructed using the tracker and the muon spectrometer information. The reconstructed muons are required to have $p_T > 25$ GeV, $| \eta | < 2.47$, $| d_0/\sigma_{d_0} | < 3$, $| \Delta z \sin(\eta) | < 0.5$ mm and originate from the primary vertex.

  - **τ-leptons**. Electrons and muons from the τ decays are considered in the analysis.

- **Overlap removal.** The procedure is exactly the same as discussed in section 6.1.2 for the objects in the SUSY searches analysis with only one change. Electrons are removed if they are located at $\Delta R = min(0.2, \ 10 \ \mathrm{GeV}/p_T)$ from a reconstructed jet.

7.2.2 Particle-Level Object Selection

The PL objects are selected in a similar manner to the detector level objects. PL objects are defined using stable particles with lifetime greater than 30 ps.

- **Jets** are clustered using the anti-k_T sequential recombination algorithm with the radius parameter of $R = 0.4$. These jets don't include jets from the pile-up events, but include jets from the underlying events. Therefore, decay products of hadronic decays of τ-leptons are included. All reconstructed jets are required to have $p_T > 25$ GeV and $| \eta | < 2.5$.

- **Flavor tagging of jets.** Jets are identified as b-jets if there is at least one b-hadron with $p_T > 5$ GeV is matched to the jet. In a similar way, c-jets are identified as jets with at least one c-hadron with $p_T > 5$ GeV matched to the jet and exactly zero b-hadrons. If there are zero b- and c- hadrons with $p_T > 5$ GeV matched to a jet, the jet is identified as a *light* jet.

- **Leptons.** Electrons and muons are required to have $p_T > 25$ GeV and $| \eta | < 2.5$. The leptons may radiate photons that carry away some energy. To account for the final-state radiation (FSR) effect, photons that are not originating from hadrons and located in a $R = 0.1$ cone around a lepton are identified as FSR photons with respect to the lepton, and four-momenta of the photos are added to the four-momentum of the corresponding lepton.

7.3   Observables

Two phase-spaces are studied in the analysis: one with at least 3 b-jets, shortly named $3j3b$, and the other with at least 4 b-jets, shortly named $3j4b$. Selections on variables for the region are summarized below:

- Leptons multiplicity: $N_e = 1$, $N_\mu = 1$.

- Total charge of the two leptons: $Q_e + Q_\mu = 0$.

- Electron, muon: $p_T > 28$ GeV, $| \eta | < 2.5$.

- Electrons and muons from τ decay are included.

- b-jets multiplicity: $3b$ channel with $N_{b-jets} \geq 3$ and $4b$ channel with $N_{b-jets} \geq 4$.

- Jets and b-jets: $p_T > 25$ GeV, $| \eta | < 2.5$.

- $\Delta R(lepton, \; jet) > 0.4$.

Two sets of variables related to b-jets are distinguished: ones that are sensitive to the modeling top quark decay and the other that are related to the properties of jets from gluon emissions. There is also the third set of variables that describes the overall event modeling. The sets of variables are summarized below:

- **Top decay kinematics:**

   – b-jets kinematics: $p_T(b_1^{top})$, $p_T(b_2^{top})$, $| \eta(b_1^{top}) |$, $| \eta(b_2^{top}) |$.
   – b-jets correlations: $m_{inv}(b_1^{top}, \; b_2^{top})$, $\Delta R(b_1^{top}, \; b_2^{top})$, $p_T(b_1^{top}, \; b_2^{top})$, $\Delta\phi(b_1^{top}, \; b_2^{top})$.
   – global: $m_{inv}(b_1^{top}b_2^{top}e\mu)$, $\Delta R(b_1^{top}b_2^{top}e\mu, \; b_1^{add.})$.

- **Additional b-jet properties:**

   – b-jets kinematics: $p_T(b_1^{add.})$, $| \eta(b_1^{add.}) |$, $p_T(b_2^{add.})$, $| \eta(b_2^{add.}) |$ (the latter two in the $4j4b$ channel only).
   – b-jets correlations (in the $4b$ channel only): $m_{inv}(b_1^{add.}, \; b_2^{add.})$, $\Delta R(b_1^{add.}, \; b_2^{add.})$, $p_T(b_1^{add.}, \; b_2^{add.})$, $\Delta\phi(b_1^{add.}, \; b_2^{add.})$.

Information of the jets origin is not available on the detector level. Therefore, an algorithm that would rely on observables other than listed above is needed for the jet origin identification. A multi variable algorithm based on the Boosted Decision Trees was developed for the analysis and presented in section 7.5.

7.4   Detector-Level Data and MC Predictions Comparisons

The section presents the data to MC comparison in the region with at least two b-tagged jets, one electron and one muon of the opposite signs. No significant discrepancy is observed in η_{jet}, ϕ_{jet}, N_{jets} and m^{b-jet}_{inv} distributions. Excess of data events in N_{b-jets} is a known issue of $t\bar{t}$ production modeling and fixed by normalization of yields in the $3j3b$ and $4j4b$ signal regions.
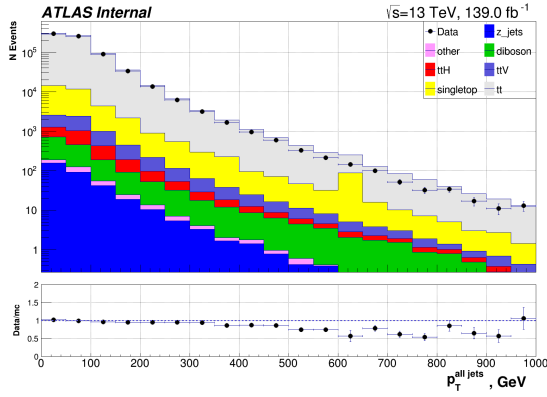
(a) p_{T} of all jets in an event
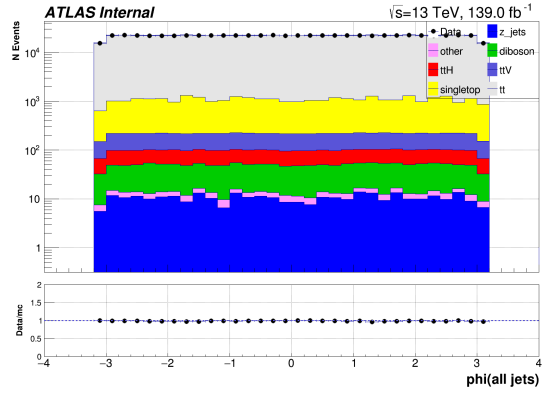
(b) ϕ of all jets in an event
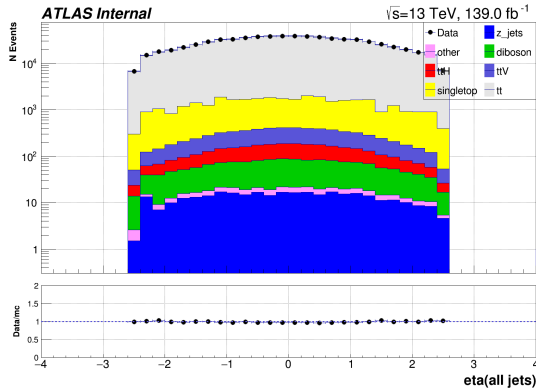
(c) η of all jets in an event
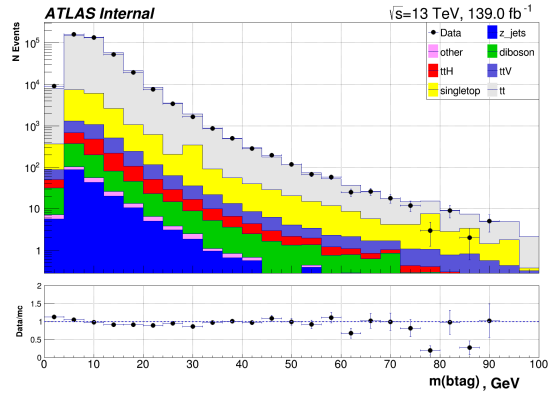
(d) m_{inv} of b-jets
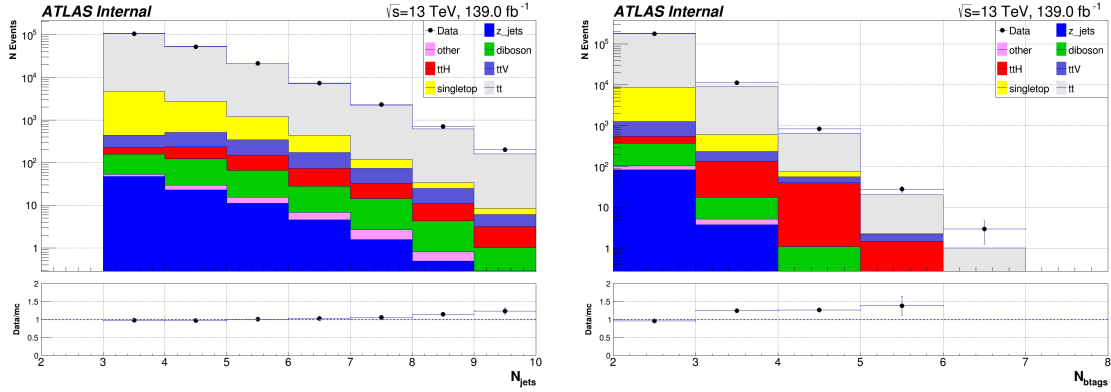
(e) jets multiplicity

(f) b-jets multiplicity

Figure 52: Kinematic distributions in events with exactly 2 b-jets, one electron and one muon of the opposite signs. Top panel: data (black dots) and MC (in color) yields. Bottom panel: data to the total MC ratio.

7.5 b-jets Origin Identification Using BDT

Observables listed in section 7.3 include ones related to the top decay kinematics and the others related to the additional b-jets due to gluon emissions. Since the information about jets origin is not available on the detector level data, a discriminator what would distinguish b-jets originating from top quark decays from the additional b-jets is needed. It is possible to design such a discriminator using the particle level MC modeling information along with observables that present on both the detector and particle level, and propagate the discriminator to the detector level events. The section presents a discriminator developed using Boosted Decision Trees (BDT) algorithm. TMVA package [187] for CERN ROOT was used to setup a framework for training and testing of the BDT algorithm.

7.5.1 BDT Simple Example

BDT is a machine learning technique for optimizing the predictive value of a model through successive steps in the learning process. An example of BDT with four variables $c1$, $c2$, $c3$, and $c4$ used to differentiate between signal and background events is shown in figure 53. The tree is build according to the following steps:

- Sort all events with respect to each variable independently and construct four arrays.

- For each of the arrays find the cut value of the corresponding variable with the best separation power. Separation $\langle S^2 \rangle$ of a classifier y is defined in formula 7.5.1, where $\hat{y}_S$ and $\hat{y}_B$ are the signal and background PDFs of y, respectively. The separation is zero for identical signal and background shapes, and it is one for shapes with no overlap [187].

99

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_S(y) - \hat{y}_B(y))^2}{\hat{y}_S(y) + \hat{y}_B(y)} dy \qquad (7.5.1)$$

- The variable with the largest separation power ($c1$) is used in the first node of the tree, and the events are organized into two branches: signal-like and background-like.

- Construct the second level nodes for the signal- and background-like events separately. As for the first node, the events are ranged with respect to each variable independently, and the variable with the largest separation value is used in a node, and the events are further divided into the signal- and background-like collections.

- Iterate until one of the stopping criteria is reached:

  - Maximal tree depth (i.e. the total number of nodes).
  - Insufficient improvement from the further splitting.
  - Minimum leaf size: the minimal number of jets ending up in the same leaf, needed for statistical significance.
  - Perfect classification: all jets in the leaf belong to the same class (signal or background).

- The result of the training is a function that is used to assign a BDT score to any event with certain values of observables $c1$, $c2$, $c3$, and$c4$ used in the training. The higher the score, the more likely the event is signal-like, and vise versa.

One of the main advantages of BDT classification is that in training a true signal event failing a cut in one observable still can be classified as signal due to other observables, while in the simple cut-and-count method such event would be assigned to background.
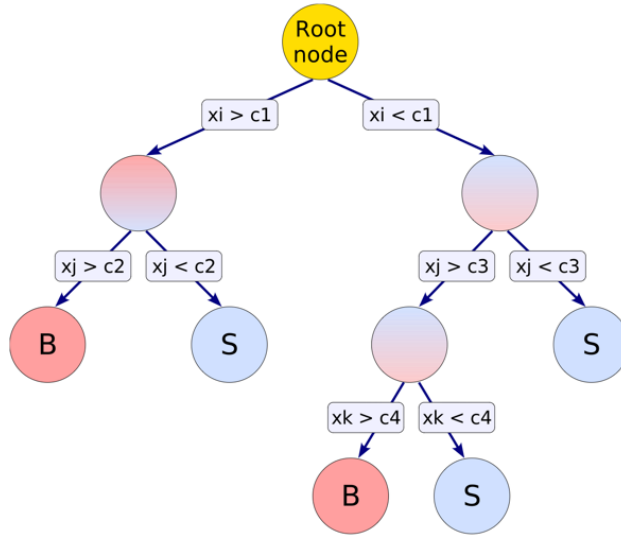
Figure 53: Example of a tree from the BDT algorithm. Figure from [187].

7.5.2 BDT Setup For b-jets Origin Identification

The BDT discriminant was constructed using a set of jets' kinematic variables selected in two steps:

- The MC modeling distributions of a large set of variables were compared to the data to test for potential miss modeling in relevant variables ranges. Those distributions with discrepancies in shapes were removed from the list of candidates. Variables listed in section 7.3 that to be unfolded were not included in the initial list.

- The MC distributions for b-jets from top quarks, henceforth named as signal, and the additional b-jets, henceforth named as background, were compared. Those with complete overlap were removed from the list of candidates.

The final list of discriminative variables used for BDT training consists of 5 observables:

- $m_{inv}(b$-jet, electron): invariant mass of a b-jet and the electron. The mass should peak around top quark mass for the background b-jets, and have no pronounced peak for the signal. The neutrino's contribution to the invariant mass wasn't computed.

- $m_{inv}(b$-jet, muon): invariant mass of a b-jet and the muon. Same behavior as for the invariant mass of a b-jet and the electron is expected.

- $\Delta R(b$-jet, lead. lepton) $= \sqrt{\Delta\phi^2 + \Delta\eta^2}$: angular distance between a b-tag and the leading-p_{T} lepton. The background b-jets are expected to be closer to a lepton coming from the same top quark decay.

- $\Delta R(b$-jet, sub-lead. lepton) $= \sqrt{\Delta\phi^2 + \Delta\eta^2}$: angular distance between a b-jet and the sub-leading-p_{T} lepton. Same behavior as for the angular distance between a b-jet and the leading-p_{T} lepton is expected.

- $\Delta R_{min}(b$-jet, lepton) $= \sqrt{\Delta\phi^2 + \Delta\eta^2}$: angular distance between a b-jet and the closest lepton. Same behavior as for the first two angular distance is expected with even smaller overlap of shapes.

The data/MC and the signal/background ratios for the discriminative variabled can be found in figures 54 and 55 respectively. The normalized data to MC-detector to MC-particle levels distributions are shown in figure 56. Good shapes agreement is observed between the data, the MC-detector and the MC-particle levels. Since the unfolding will be done to the particle level, and to remove potential uncertainty from b-tagger systematics, BDT was trained and tested using truth b-jets and truth leptons from the particle level events.
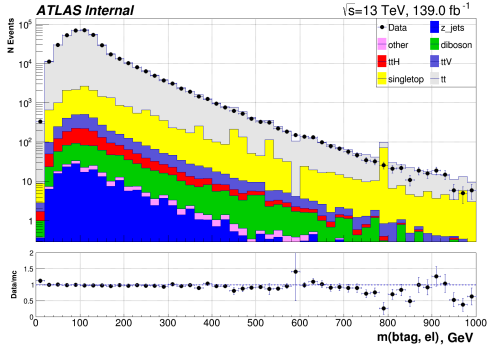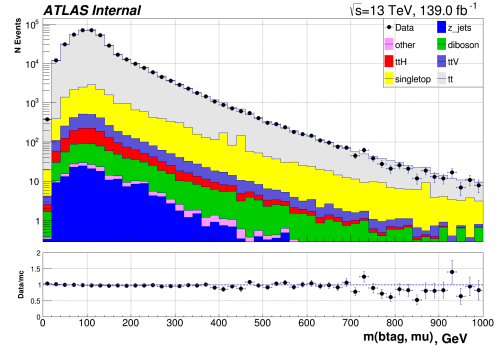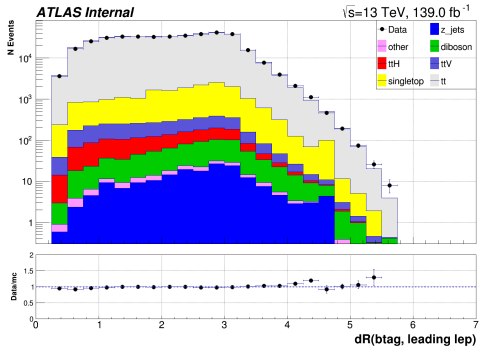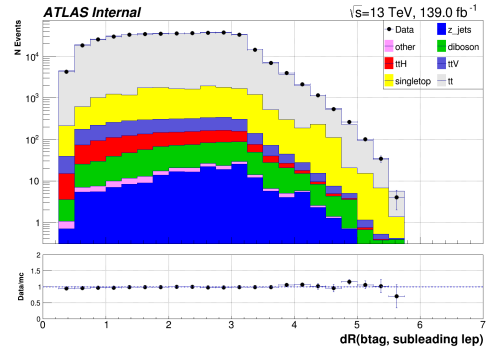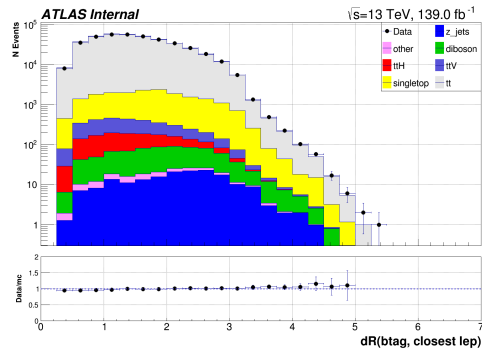
Figure 54: Data to MC comparisons of the discriminative variables used in the BDT: (a) invariant mass of the (*b*-jet, electron) system, (b) invariant mass of the (*b*-jet, muon) system, (c) ΔR between a *b*-jet and the leading lepton, (d) ΔR between a *b*-jet and the leading lepton, (e) ΔR between a *b*-jet and the closest lepton.
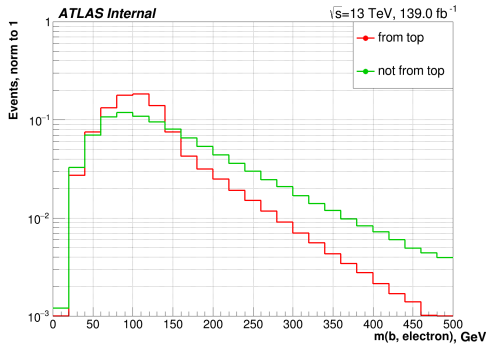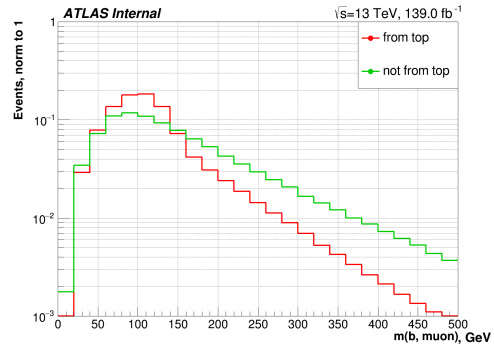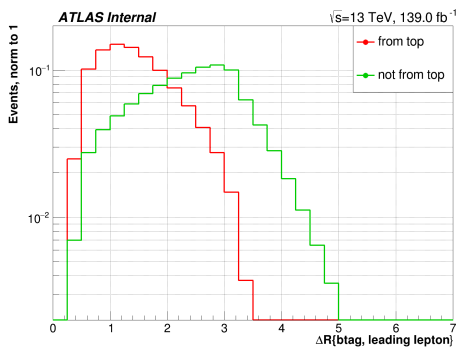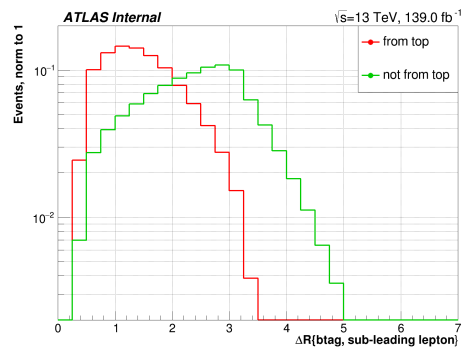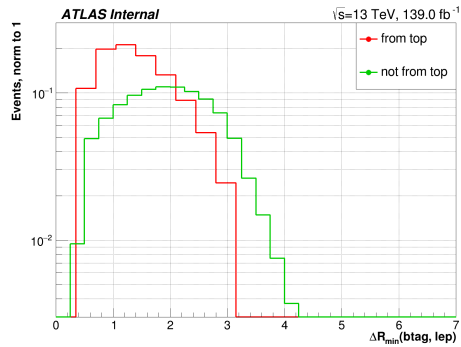
Figure 55: Signal to Background MC comparisons of the discriminative variables used in the BDT: (a) invariant mass of the (*b*-jet, electron) system, (b) invariant mass of the (*b*-jet, muon) system,(c) ΔR between a *b*-jet and the leading lepton, (d) ΔR between a *b*-jet and the leading lepton, (e) ΔR between a *b*-jet and the closest lepton.
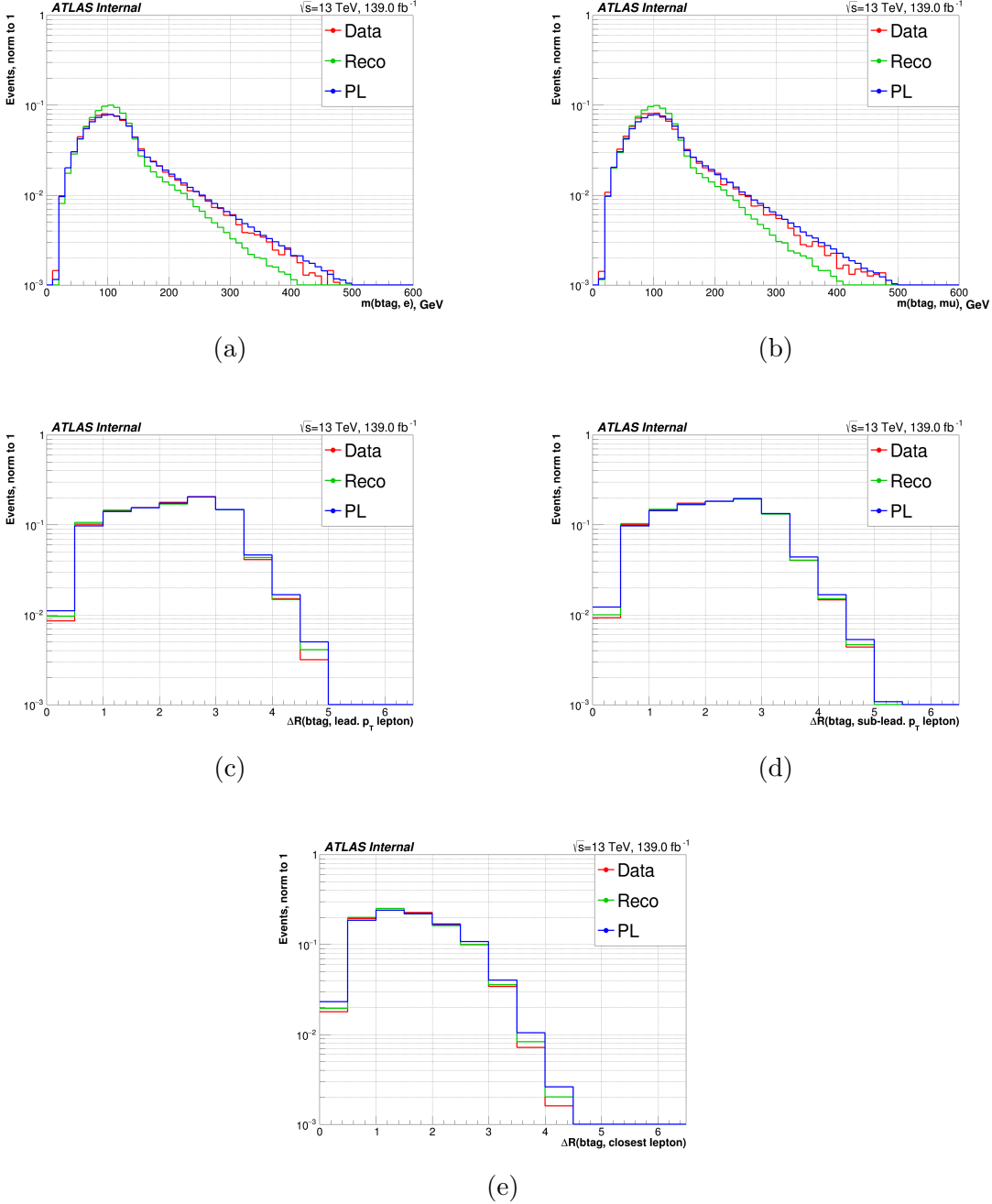
Figure 56: Data to MC-detector to MC-particle level comparisons of the discriminative variables used in the BDT: (a) invariant mass of the (b-jet, electron) system, (b) invariant mass of the (b-jet, muon) system, (c) ΔR between a b-jet and the leading lepton, (d) ΔR between a b-jet and the leading lepton, (e) ΔR between a b-jet and the closest lepton.

7.5.3 Performance of the BDT Discriminator

The list of discriminative variables ranked with respect to their separation values $\langle S^2 \rangle$ is given in table 12. Correlation matrices for the variables are shown in figure 57 for (a) the signal and the (b) background b-jets. Although there is some level of correlation between

$\Delta R_{min}(b\text{-jet, lepton})$ and both $\Delta R(b\text{-jet, lead. lepton})$ and $\Delta R(b\text{-jet, sub-lead. lepton})$, its separation power is of the same order compared to the rest of discriminative observables. It was found that presence of the variable in the list increases efficiency of the discriminator.

Variable	Separation $\langle S^2 \rangle$
$m_{inv}(b\text{-jet, electron})$	$1.134 * 10^{-1}$
$m_{inv}(b\text{-jet, muon})$	$8.571 * 10^{-2}$
$\Delta R_{min}(b\text{-jet, lepton})$	$6.931 * 10^{-2}$
$\Delta R(b\text{-jet, lead. lepton})$	$6.911 * 10^{-2}$
$\Delta R(b\text{-jet, sub-lead. lepton})$	$5.542 * 10^{-2}$

Table 12: Variables used to train BDT classifies ranked with respect to their separation $\langle S^2 \rangle$

(a)   (b)
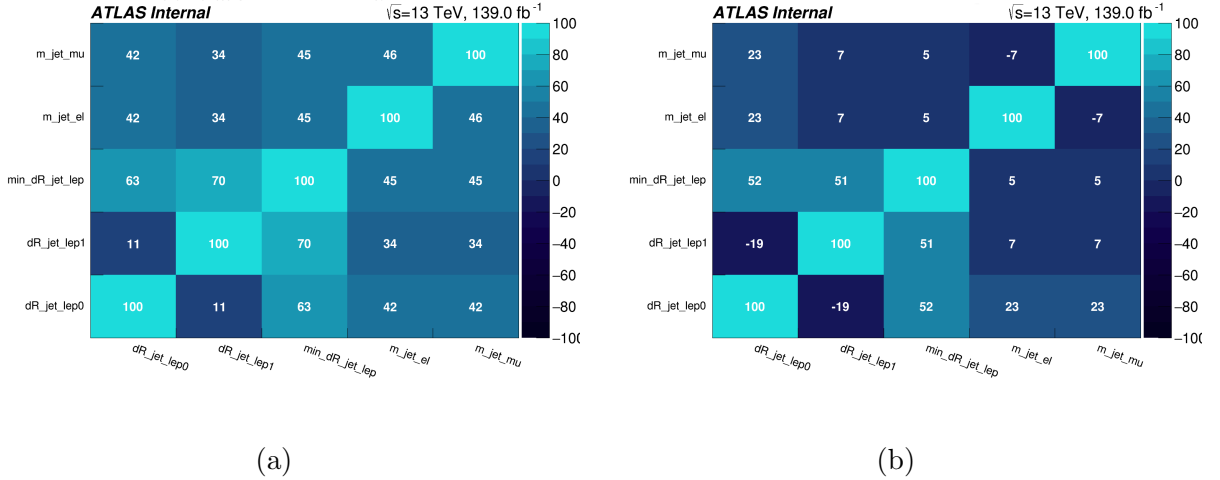
Figure 57: Correlation matrices for the set of discriminative variables used in BDT for (a) signal and (b) background b-jets.

BDT scores spectra for the signal and background b-jets are presented in figure 58. From the figure it follows that the lower the BDT score, the higher probability for a b-jet to be from top quark decay, and vise versa. Assuming that all b-jets with BDT scores below a certain value originate from top quarks decays, efficiency and purity of the BDT discriminator to assign b-jets to top quarks were evaluated and are shown in figure 59. The plots indicated that the efficiency and purity are in odds, and there is no such value of BDT score as the most optimal cut. Another disadvantage of a fixed cut in BDT score is the possibility to assign more than two b-jets to top quarks. However, regardless the BDT scores of b-jets in a $t\bar{t} + b$ event, it is expected to have exactly two b-jets originating from top quark decay and at least one additional b-jet.
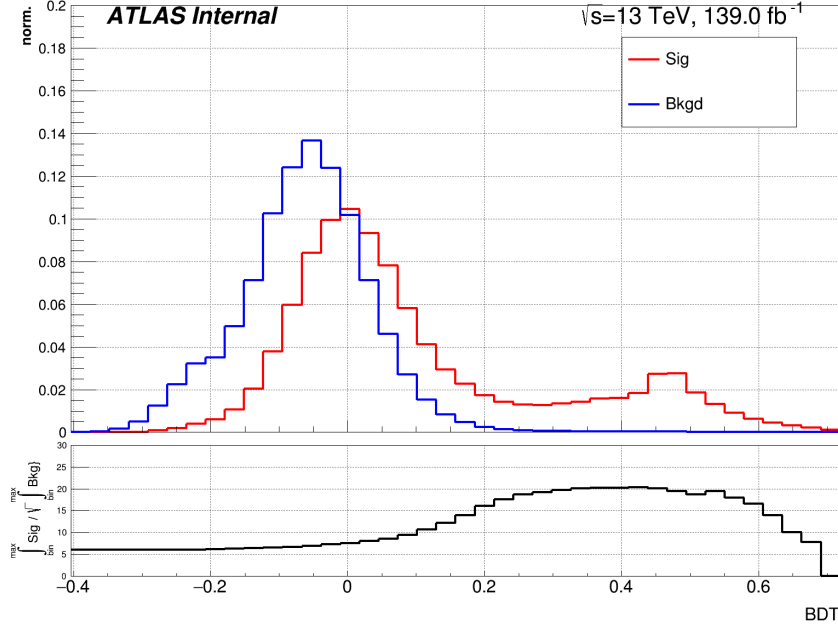
Figure 58: BDT score spectra for b-jets originating from top (signal) and additional b-jets (background).

(a)   (b)
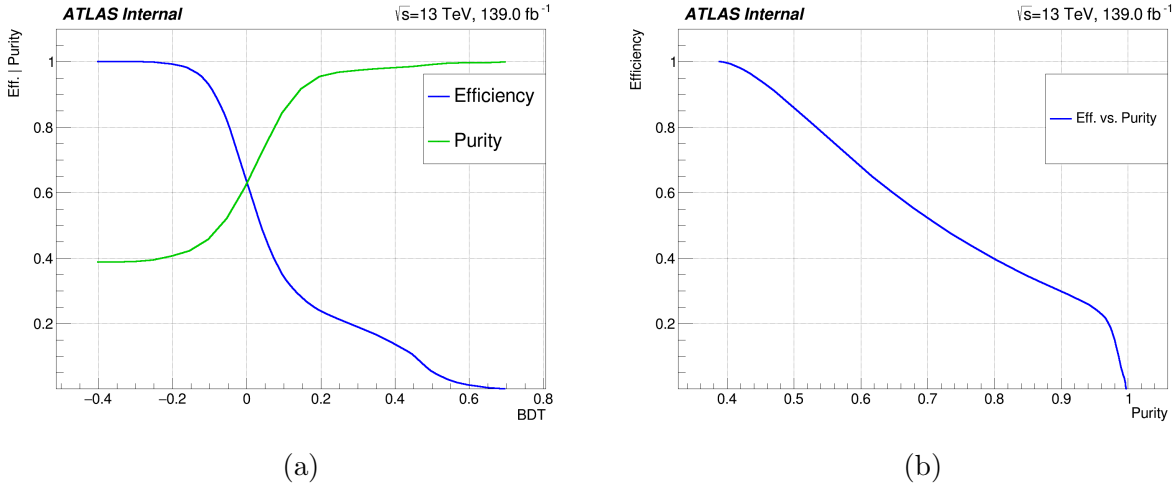
Figure 59: (a) Efficiency to select all b-jet originating from top quark decay and purity of the selected statistic as functions of BDT cut, and (b) efficiency versus purity.

For the reasons discussed above, another approach to label b-jets was adopted: in each event two b-jets with the smallest BDT scores are assigned to top quarks decays, and the rest of the b-jets being additional. The approach excludes possibility to assign non-physical number of b-jets to top quarks decays in the $t\bar{t}$ production events, as it can happen with a fixed cut in the BDT score.

BDT scores derived with the particle level b-jets were applied to the detector level b-tagged jets. Efficiencies to correctly label the leading, sub-leading p_T from top and additional b-jets

are presented in figure 60 for both the $3j3b$ and the $4j4b$ channels. In the $3j3b$ channel, the efficiency for the leading p_T b-jet from top is as hight as 80%, and 68% for the sub-leading p_T b-jet from top. The efficiency to identify the origin of additional b-jet is approximately 64%. In the $4j4b$ channel, the efficiency for signal leading and sub-leading p_T and background b-jets are: 74% and approximately 59%, 65% and approximately 70%. The efficiencies in both the $3j3b$ and the $4j4b$ channels are almost identical for both the particle and detector level jets of each kind of origin and p_T order.

Since some of the observables for the unfolding compose of information from both b-jets originating from top quark decays or the two leading p_T additional, efficiency to correctly assign the pair of b-jets to tops (and thus the pair of leading p_T additional in $4j4b$) was estimated and shown in figure 61. The efficiency is 60% in the $3b$ channel and approximately 51% in the $4j4b$. The efficiencies are almost identical for the particle and detector level b's. There are up to 10% of events with less than two b-jets from top quarks decays in the $3j3b$ channel, and $1 - 4\%$ in the $4j4b$. A loss of a b-jet from top quark decay might happen due to imperfection the b-tagger in the detector level events or due to smearing of two close b-jets by the anti-k_T jets reconstruction algorithm.

(a)   (b)
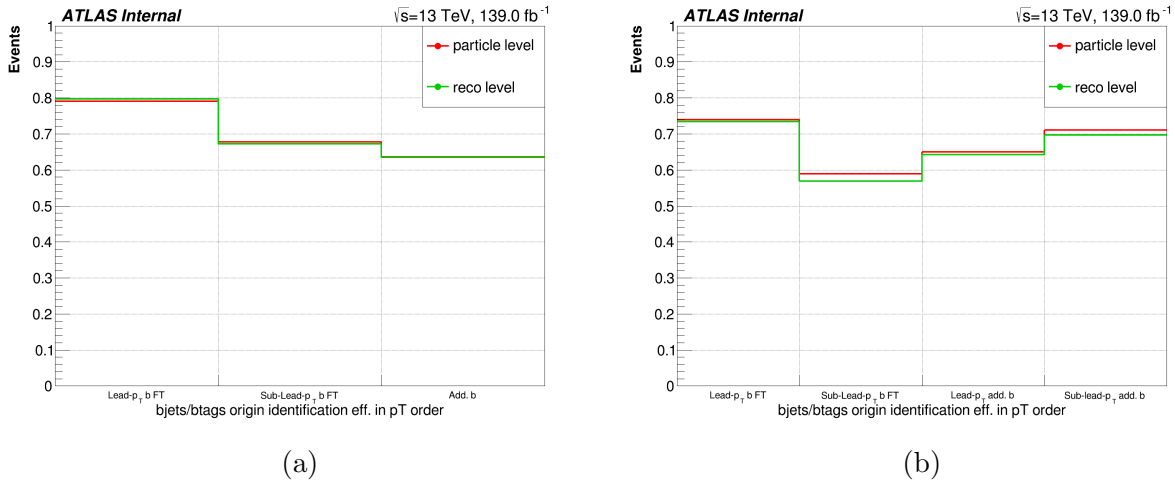
Figure 60: Efficiencies to correctly assign b-jets to top quark decay and additional b-jets in their p_T orders in the (a) $3j3b$ and (b) $4j4b$ channels in $t\bar{t}$ events. "FT" stands for "originating From Top", and "add." - additional or not from top.
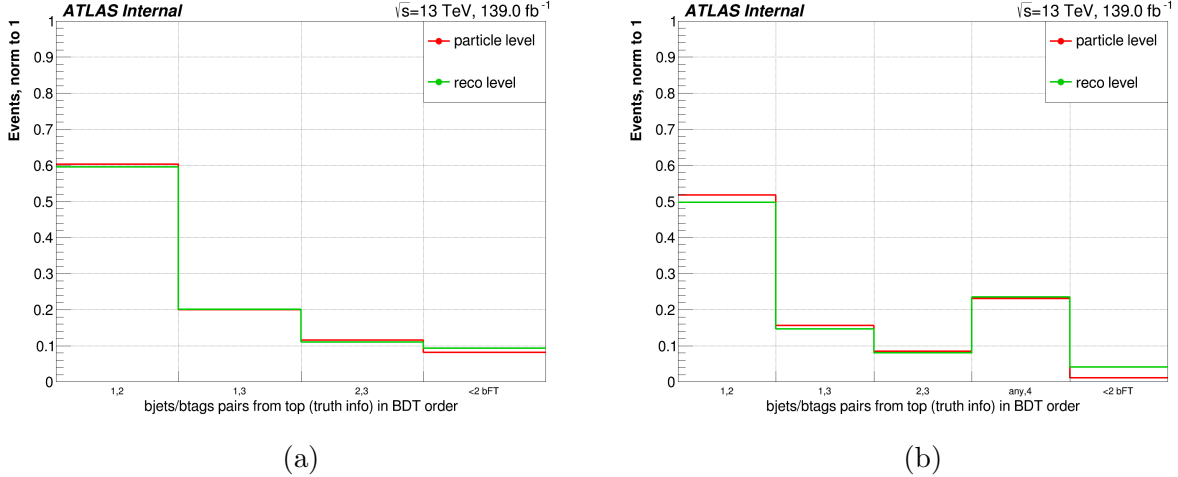
Figure 61: Efficiency to correctly identify pair of b-jets originating from top quarks decays in the (a) $3j3b$ and (b) $4j4b$ channels in $t\bar{t}$ events. The first bin shows the fraction of events with the two smallest BDT b-jets (1 and 2) originating from top on the truth level. The last bin shows the fraction of event with less than two b-jets originating from top on the truth level. It might happen due to jets smearing or miss-tagging as non-b (on the detector level). The other bins show other combinations of b-jets labeled in BDT order where at leas one is incorrectly assigned as from top.

7.6 Unfolding

The measured distributions at detector level are unfolded to the particle level. The unfolding procedure corrects for the detector efficiency and acceptance, and for resolution effects. An iterative Bayesian unfolding technique [188], as implemented in the RooUnfold software package [189], is used.

First, the number of non-$t\bar{t}$ and $t\bar{t}V$, $t\bar{t}H$ backgrounds (N_{bkg}^k) are subtracted from the data (N_{data}^k) distribution at the detector level in each bin k. This leaves a mixture of signal and $t\bar{t}$-related background with mis-tagged b-jets. Following the subtraction of non-$t\bar{t}$ background, the data are first corrected for mis-tagged events by applying a correction factor ($f_{t\bar{t}b}^k$) as defined below separately for $t\bar{t}b$ and $t\bar{t}b\bar{b}$ signal fiducial regions.

For the $t\bar{t}b$ signal region:

$$f_{t\bar{t}b}^k = \frac{\mathcal{S}_{t\bar{t}b,\mathrm{reco}}^k}{\mathcal{S}_{t\bar{t}b,\mathrm{reco}}^k + \mathcal{B}_{t\bar{t}b,\mathrm{reco}}^k} \tag{7.6.1}$$

where the signal ($\mathcal{S}_{t\bar{t}b,\mathrm{reco}}^k$) and background ($\mathcal{B}_{t\bar{t}b,\mathrm{reco}}^k$) predicted events are defined as:

$$\mathcal{S}_{t\bar{t}b,\mathrm{reco}}^k = \sum_r \left(\alpha_{b;r}^{3\mathrm{j}} N_{t\bar{t}b,\mathrm{reco};r}^{3\mathrm{j};k} \right) + \sum_{r'} \left(\alpha_{b;r'}^{4\mathrm{j}} N_{t\bar{t}b,\mathrm{reco};r'}^{4\mathrm{j};k} \right) + \sum_{r'} \left(\alpha_{bb;r'}^{4\mathrm{j}} N_{t\bar{t}b\bar{b},\mathrm{reco};r'}^{k} \right) \tag{7.6.2}$$

$$\mathcal{B}^k_{t\bar{t}b,\text{reco}} = \sum_r \left(\alpha^{3j}_{c;r} N^{3j;k}_{t\bar{t}c,\text{reco};r} \right) + \sum_{r'} \left(\alpha^{4j}_{c;r'} N^{4j;k}_{t\bar{t}c,\text{reco};r'} \right) + \sum_r \left(\alpha^{3j}_{l;r} N^{3j;k}_{t\bar{t}l,\text{reco};r} \right) + \sum_{r'} \left(\alpha^{4j}_{l;r'} N^{4j;k}_{t\bar{t}l,\text{reco};r'} \right)$$

$$(7.6.3)$$

where, the scale factors $\alpha^{3j}_{b;r}$, $\alpha^{4j}_{bb;r'}$, $\alpha^{3j}_{c;r'}$, $\alpha^{4j}_{c;r'}$, $\alpha^{3j}_{l;r}$ and $\alpha^{4j}_{l;r'}$ for a given p_T bin r or r' in the presented analysis are set unity. The $N^{3j;k}_{t\bar{t}b,\text{reco};r}$ and $N^{4j;k}_{t\bar{t}b,\text{reco};r'}$ are the number of reconstructed $t\bar{t}b$ events in k bin with 3 jets, or ≥ 4 jets as predicted by the $t\bar{t}$ simulation, respectively. The $N^k_{t\bar{t}b\bar{b},\text{reco};r'}$ are the number of reconstructed $t\bar{t}b\bar{b}$ events in k^{th} bin. Similarly, the number of $t\bar{t}c$ background events containing 3 jets or ≥ 4 jets are represented by $N^{3j;k}_{t\bar{t}c,\text{reco};r}$ or $N^{4j;k}_{t\bar{t}c,\text{reco};r'}$, respectively. And the number of $t\bar{t}l$ background events containing 3 jets or ≥ 4 jets are denoted by $N^{3j;k}_{t\bar{t}l,\text{reco};r}$ or $N^{4j;k}_{t\bar{t}l,\text{reco};r'}$, respectively.

For the $t\bar{t}b\bar{b}$ signal region:

$$f^k_{t\bar{t}b\bar{b}} = \frac{\mathcal{S}^k_{t\bar{t}b\bar{b},\text{reco}}}{\mathcal{S}^k_{t\bar{t}b\bar{b},\text{reco}} + \mathcal{B}^k_{t\bar{t}b\bar{b},\text{reco}}} \qquad (7.6.4)$$

where

$$\mathcal{S}^k_{t\bar{t}b\bar{b},\text{reco}} = \sum_{r'} \left(\alpha^{4j}_{bb;r'} N^k_{t\bar{t}b\bar{b},\text{reco};r'} \right) \qquad (7.6.5)$$

$$\mathcal{B}^k_{t\bar{t}b\bar{b},\text{reco}} = \sum_r \left(\alpha^{4j}_{b;r} N^{4j;k}_{t\bar{t}b,\text{reco};r} \right) + \sum_{r'} \left(\alpha^{4j}_{c;r'} N^{4j;k}_{t\bar{t}c,\text{reco};r'} \right) + \sum_{r'} \left(\alpha^{4j}_{l;r'} N^{4j;k}_{t\bar{t}l,\text{reco};r'} \right) \qquad (7.6.6)$$

The $t\bar{t}b$ events with at least four jets are regarded as background according to $t\bar{t}b\bar{b}$ phase space definition if there are less than four b-tagged jets, the number of such events predicted by MC simulation are denoted as $N^{4j;k}_{t\bar{t}b,\text{reco};r}$.

Next, an acceptance correction f^k_{accept} is applied, which corrects for the fiducial acceptance and is defined as the probability of a $t\bar{t}b$ event (or a $t\bar{t}b\bar{b}$ event) passing the detector-level selection in a given bin k to also fall within the fiducial particle-level phase space.

It is estimated as for $t\bar{t}b$ phase space:

$$f^k_{\text{accept},t\bar{t}b} = \frac{\mathcal{S}^k_{t\bar{t}b,\text{reco}\wedge\text{part}}}{\mathcal{S}^k_{t\bar{t}b,\text{reco}}} \qquad (7.6.7)$$

and for $t\bar{t}b\bar{b}$ phase space:

$$f^k_{\text{accept},t\bar{t}b\bar{b}} = \frac{\mathcal{S}^k_{t\bar{t}b\bar{b},\text{reco}\wedge\text{part}}}{\mathcal{S}^k_{t\bar{t}b\bar{b},\text{reco}}} \qquad (7.6.8)$$

where $\mathcal{S}^k_{t\bar{t}b,\text{reco}\wedge\text{part}}$ and $\mathcal{S}^k_{t\bar{t}b\bar{b},\text{reco}\wedge\text{part}}$ are the reweighted number of events passing both the detector-level and particle level selections in $t\bar{t}b$ and $t\bar{t}b\bar{b}$ phase space definition, respectively.

The detector-level objects are required to be matched to the corresponding particle-level objects. The objects are considered matched if the ΔR between the particle-level and detector-level objects is less than 0.4. This requirement leads to a better correspondence between the particle and detector levels. The matching factor f^k_{matching} is defined as:

$$f^k_{\text{matching},t\bar{t}b} = \frac{\mathcal{S}^k_{t\bar{t}b,\text{reco}\wedge\text{part}\wedge\text{matched}}}{\mathcal{S}^k_{t\bar{t}b,\text{reco}\wedge\text{part}}} \tag{7.6.9}$$

$$f^k_{\text{matching},t\bar{t}b\bar{b}} = \frac{\mathcal{S}^k_{t\bar{t}b\bar{b},\text{reco}\wedge\text{part}\wedge\text{matched}}}{\mathcal{S}^k_{t\bar{t}b\bar{b},\text{reco}\wedge\text{part}}} \tag{7.6.10}$$

where $\mathcal{S}^k_{t\bar{t}b,\text{reco}\wedge\text{part}\wedge\text{matched}}$ and $\mathcal{S}^k_{t\bar{t}b\bar{b},\text{reco}\wedge\text{part}\wedge\text{matched}}$ are the subsets of reconstructed events falling in the particle level fiducial volume which are matched to the corresponding particle level objects.

The remaining part of the unfolding procedure consists of effectively inverting the migration matrix $\mathcal{M}$ to correct for the resolution effects and subsequently correcting for detector inefficiencies. An iterative Bayesian unfolding technique [189] implemented in the RooUnfold software package is used. The matrix, $\mathcal{M}$, represents the probability for a particle-level event in bin i to be reconstructed in bin k. The chosen binning is optimised for each distribution to have a migration matrix with a large fraction ($\sim 60\%$) of events on the diagonal and a sufficient number of events in each bin. The Bayesian unfolding technique performs the effective matrix inversion, $\mathcal{M}^{-1}_{ik}$, iteratively. The optimal number of iteration ranges from 3 to 4 depending on the variable.

Finally, the factor f^i_{eff} corrects for the reconstruction efficiency and is defined as

$$f^i_{\text{eff},t\bar{t}b} = \frac{\mathcal{S}^i_{t\bar{t}b,\text{reco}\wedge\text{part}\wedge\text{matched}}}{\mathcal{S}^i_{t\bar{t}b,\text{part}}} \tag{7.6.11}$$

$$f^i_{\text{eff},t\bar{t}b\bar{b}} = \frac{\mathcal{S}^i_{t\bar{t}b\bar{b},\text{reco}\wedge\text{part}\wedge\text{matched}}}{\mathcal{S}^i_{t\bar{t}b\bar{b},\text{part}}} \tag{7.6.12}$$

where $\mathcal{S}^i_{t\bar{t}b,\text{part}}$ and $\mathcal{S}^i_{t\bar{t}b\bar{b},\text{part}}$ are the weighted number of $t\bar{t}b$ and $t\bar{t}b\bar{b}$ events, respectively, passing the particle-level selection in bin i. The $\mathcal{S}^i_{t\bar{t}b,\text{reco}\wedge\text{part}\wedge\text{matched}}$ and $\mathcal{S}^i_{t\bar{t}b\bar{b},\text{reco}\wedge\text{part}\wedge\text{matched}}$ are the weighted number of $t\bar{t}b$ and $t\bar{t}b\bar{b}$ events in bin i, respectively, that also pass the detector-level selection, containing matched objects. These are expressed as below for $t\bar{t}b$:

$$\mathcal{S}^i_{t\bar{t}b,\text{part}} = \sum_r \left(\tilde{\alpha}^{3j}_{b;r} N^{3j;i}_{t\bar{t}b,\text{part};r} \right) + \sum_r \left(\tilde{\alpha}^{4j}_{b;r'} N^{4j;i}_{t\bar{t}b,\text{part};r'} \right) + \sum_{r'} \left(\tilde{\alpha}^{4j}_{bb;r'} N^i_{t\bar{t}b\bar{b},\text{part};r'} \right) \tag{7.6.13}$$

$$\mathcal{S}^i_{t\bar{t}b,\text{reco}\wedge\text{part}\wedge\text{matched}} = \sum_r \left(\alpha^{3j}_{b;r} N^{3j;i}_{t\bar{t}b,\text{reco}\wedge\text{part}\wedge\text{matched};k} \right) + \sum_{r'} \left(\alpha^{4j}_{b;r'} N^{4j;i}_{t\bar{t}b,\text{reco}\wedge\text{part}\wedge\text{matched};r'} \right)$$
$$+ \sum_{r'} \left(\alpha^{4j}_{bb;r'} N^i_{t\bar{t}b\bar{b},\text{reco}\wedge\text{part}\wedge\text{matched};r'} \right) \tag{7.6.14}$$

and for $t\bar{t}b\bar{b}$ phase space:

$$S^i_{t\bar{t}b\bar{b},\text{part}} = \sum_{r'} \left(\tilde{\alpha}^{4j}_{bb;r'} N^i_{t\bar{t}b\bar{b},\text{reco};r'} \right) \tag{7.6.15}$$

$$S^i_{t\bar{t}b\bar{b},\text{reco}\wedge\text{part}\wedge\text{matched}} = \sum_{r'} \left(\alpha^{4j}_{bb;r'} N^i_{t\bar{t}b\bar{b},\text{reco}\wedge\text{part}\wedge\text{matched};r'} \right) \tag{7.6.16}$$

The unfolding procedure for an observable X at particle level can be summarized by the following expressions for $t\bar{t}b$ and $t\bar{t}b\bar{b}$ phase spaces, respectively.

$$\frac{\mathrm{d}\sigma^{\text{fid}}_{t\bar{t}b}}{\mathrm{d}X^i} = \frac{N^i_{\text{unfold},t\bar{t}b}}{\mathcal{L}\,\Delta X^i} = \frac{1}{\mathcal{L}\,\Delta X^i\,f^i_{\text{eff},t\bar{t}b}} \sum_k \mathcal{M}^{-1}_{ik}\, f^k_{\text{matching},t\bar{t}b}\, f^k_{\text{accept},t\bar{t}b}\, f^k_{t\bar{t}b}\, (N^{3b,k}_{\text{data}} - N^{3b,k}_{\text{bkg}}) \tag{7.6.17}$$

$$\frac{\mathrm{d}\sigma^{\text{fid}}_{t\bar{t}b\bar{b}}}{\mathrm{d}X^i} = \frac{N^i_{\text{unfold},t\bar{t}b\bar{b}}}{\mathcal{L}\,\Delta X^i} = \frac{1}{\mathcal{L}\,\Delta X^i\,f^i_{\text{eff},t\bar{t}b\bar{b}}} \sum_k \mathcal{M}^{-1}_{ik}\, f^k_{\text{matching},t\bar{t}b\bar{b}}\, f^k_{\text{accept},t\bar{t}b\bar{b}}\, f^k_{t\bar{t}b\bar{b}}\, (N^{4b,k}_{\text{data}} - N^{4b,k}_{\text{bkg}})$$
$$\tag{7.6.18}$$

where ΔX^i is the bin width, $\mathcal{L}$ is the integrated luminosity and $N^i_{\text{unfold},t\bar{t}b}$ and $N^i_{\text{unfold},t\bar{t}b\bar{b}}$ are the number of events in bin i of the unfolded distribution.

$$N^i_{\text{unfold},t\bar{t}b} = \frac{1}{f^i_{\text{eff},t\bar{t}b}} \sum_k \mathcal{M}^{-1}_{ik}\, f^k_{\text{matching},t\bar{t}b}\, f^k_{\text{accept},t\bar{t}b}\, f^k_{t\bar{t}b}\, (N^{3b,k}_{\text{data}} - N^{3b,k}_{\text{bkg}}) \tag{7.6.19}$$

$$N^i_{\text{unfold},t\bar{t}b\bar{b}} = \frac{1}{f^i_{\text{eff},t\bar{t}b\bar{b}}} \sum_k \mathcal{M}^{-1}_{ik}\, f^k_{\text{matching},t\bar{t}b\bar{b}}\, f^k_{\text{accept},t\bar{t}b}\, f^k_{t\bar{t}b\bar{b}}\, (N^{4b,k}_{\text{data}} - N^{4b,k}_{\text{bkg}}) \tag{7.6.20}$$

The integrated fiducial cross sections $\sigma^{\text{fid}}_{t\bar{t}b}$ and $\sigma^{\text{fid}}_{t\bar{t}b\bar{b}}$ are obtained from

$$\sigma^{\text{fid}}_{t\bar{t}b} = \int \frac{d\sigma^{\text{fid}}_{t\bar{t}b}}{dX} dX = \frac{\sum_i N^i_{\text{unfold},t\bar{t}b}}{\mathcal{L}} \tag{7.6.21}$$

and

$$\sigma^{\text{fid}}_{t\bar{t}b\bar{b}} = \int \frac{d\sigma^{\text{fid}}_{t\bar{t}b\bar{b}}}{dX} dX = \frac{\sum_i N^i_{\text{unfold},t\bar{t}b\bar{b}}}{\mathcal{L}} \ , \tag{7.6.22}$$

respectively.

Results from the unfolded distributions are presented in terms of a relative differential cross-sections as $\frac{1}{\sigma^{\text{fid}}_{t\bar{t}b}} \cdot \frac{d\sigma^{\text{fid}}_{t\bar{t}b}}{dX^i}$ and $\frac{1}{\sigma^{\text{fid}}_{t\bar{t}b\bar{b}}} \cdot \frac{d\sigma^{\text{fid}}_{t\bar{t}b\bar{b}}}{dX^i}$, respectively for $t\bar{t}b$ and $t\bar{t}b\bar{b}$ fiducial regions.

7.7 Systematic Uncertainties

The section present the uncertainties that affect the analysis. Section 7.7.1 discusses statistical systematic uncertainties. Experimental systematic uncertainties are summarized in section 7.7.2. Systematic uncertainties due to $t\bar{t}$ modeling are presented in section 7.7.3.

7.7.1 Statistical Systematics

Statistical uncertainties of the data. The impact of the statistical uncertainties of the data is evaluated using generation of toys. The toys are obtained by Poisson fluctuating the data in each bin i of the reconstructed distribution:

$$N_{toy}^i = Poisson\ _{toy}(N_{obs}^i) \tag{7.7.1}$$

where $Poisson\ _{toy}(N_{obs}^i)$ represents the Poisson distribution with the mean of N_{obs}^i - the number of observed data events in the ith bin. The data statistical uncertainty of a given bin of the unfolded distribution is taken as the RMS of the unfolded toy generator results:

$$\delta N_{unflod}^i = \sqrt{\frac{1}{n-1} \sum_{j=1}^{i} (N_{unfold;\ j}^i - \overline{N_{unfold}^i})^2} \tag{7.7.2}$$

Statistical uncertainties of t$\bar{\text{t}}$ MC. Similar to the data, the impact of the statistical uncertainties to the nominal unfolding correlations due to limited statistics in the $t\bar{t}$ MC samples was estimated using toys generation. The toys are generated by fluctuating each bin of the detector and particle levels distributions and of the migration matrix within their bin errors. The resulting MC statistical uncertainty is estimated as the root mean square of the unfolded results obtained from all generated toys.

Statistical uncertainty of the subtracted background is also estimated using toys. The background predictions in each bin are fluctuated within their bin error, and the root mean square of the unfolded results from toys is taken as the statistical uncertainty.

7.7.2 Experimental Systematics

Most of the systematic uncertainties are similar or event identical (luminosity) to those used in the SUSY searches analyses and discussed in section 6.4.1.

b−tagging efficiency. To correct the difference between the b-tagging efficiency in the MC simulation and the data, a scale factor is applied to each simulated event. The b-tagging uncertainty is evaluated by varying p_T, η and truth flavor dependent scale factors. The variations are applied separately to b, c and $light$ jets resulting into three uncorrelated uncertainties.

Luminosity. The systematic uncertainty on the integrated luminosity for the full Run-2 (139 fb^{-1}) is 1.7% [33] obtained using the LUCID-2 detector [161].

Pile-up. The uncertainty in the pile-up reweighing is estimated by comparing the distributions of the number of primary vertices in the MC simulation and the data as the function of instantaneous luminosity. Differences between these distributions are adjusted by scaling the mean number of pp interactions per bunch crossing in the MC simulation, and $\pm\sigma$ uncertainties are assigned to these scale factors.

7.7.3 Modeling Systematics

t$\bar{t}$ MC Generator. Uncertainties due to the choice of a particular $t\bar{t}$ MC generator are evaluated by unfolding alternative $t\bar{t}$ samples. The list of the alternative samples is given in table 10.

Fragmentation and hadronisation model. Uncertainties due to the choice of the factorization and the hadronisation models are evaluated using the Powheg+Herwig 7.1.3 sample. In the sample only the parton shower and hadronisation are varied compared to the nominal Powheg+Pythia8 sample.

Uncertainty due to the **modeling of the first hard emission** at the matrix element is evaluated using Powheg+Pythia8 sample with h_{damp} parameter set to $3.0m_{top}$, while the nominal value is $1.5m_{top}$.

The uncertainty due to the **choice choice of PDF** is evaluated following the PDF4LHC prescription [146] using event weights that are available in the nominal Powheg+Pythia8 sample.

7.8   Results

The unfolded results are presented as fiducial cross-sections and as normalized differential cross-sections as a function of analysis variables listed in section 7.3. The unfolding is done for the 4th iteration of the Iterative Bayesian Method

The unfolded distributions for the in the $3j3b$ region are presented in figures $62 - 66$. The unfolded distributions describe the observed data well, and the dominant uncertainty is the statistical uncertainty.

Unfolded distributions for the $4j4b$ regions are presented in figures $67 - 71$. Similar to the $3j3b$ region, the unfolded distributions describe the observed data well, and the dominant uncertainty is the statistical uncertainty.
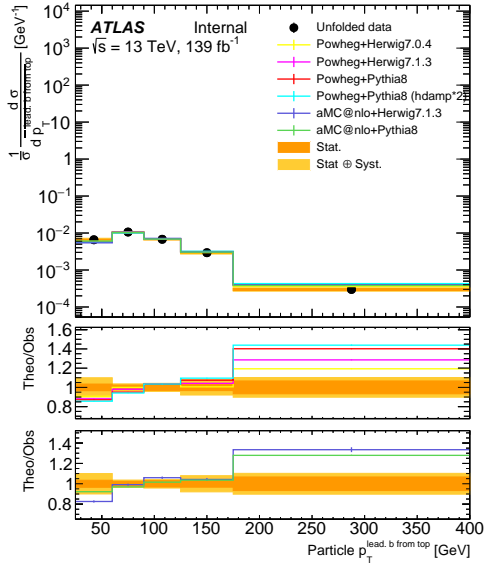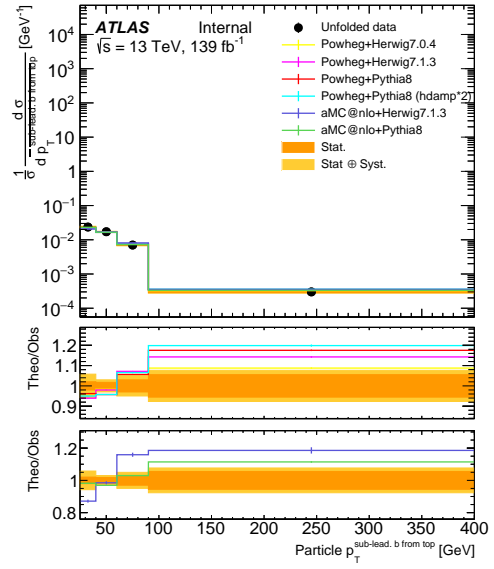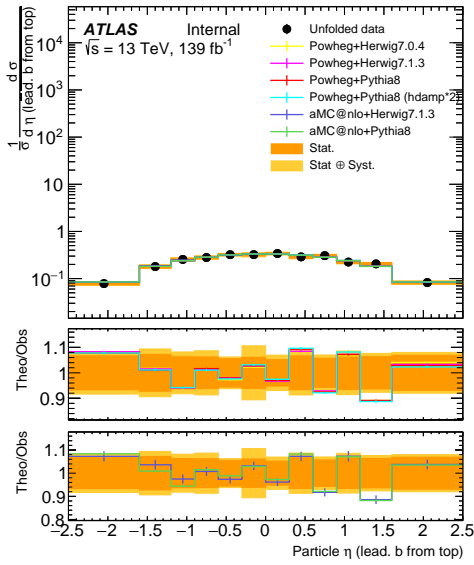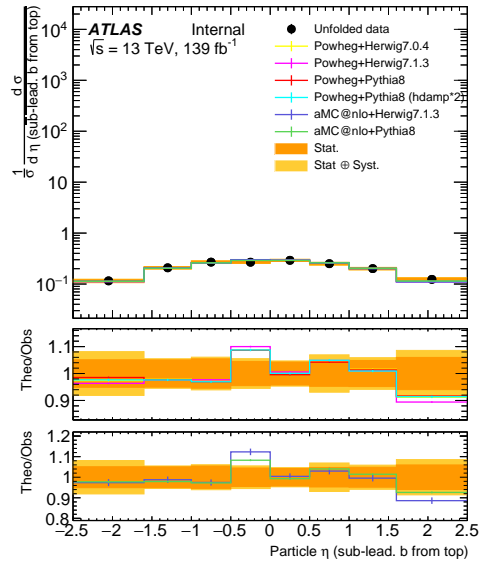
Figure 62: Unfolded distributions in the 3j3b region for: (a) the leading p_T b-jet from top quark decay, (b) the sub-leading p_T b-jet from top, (c) η of the leading b-jet from top, (d) η of the sub-leading b-jet from top.
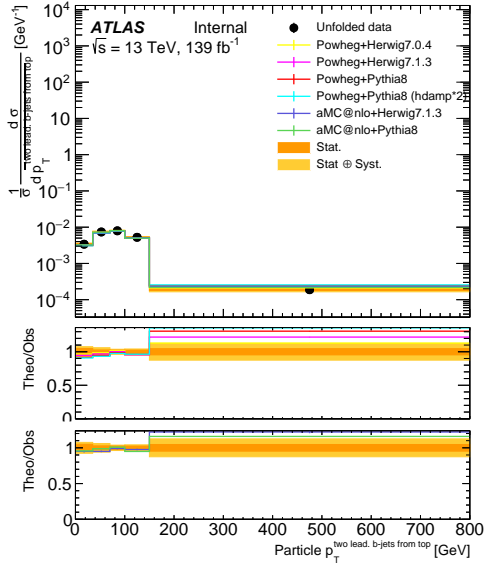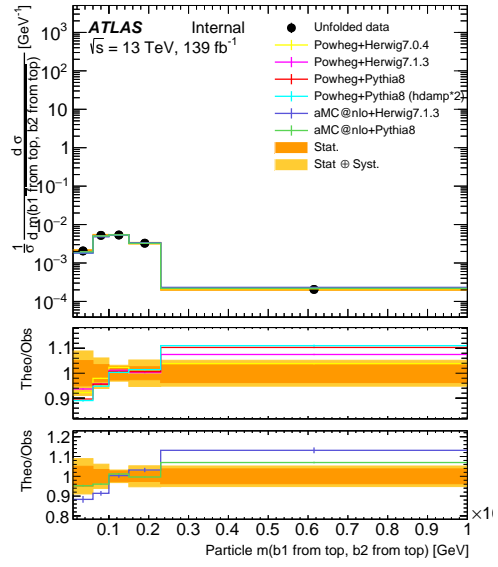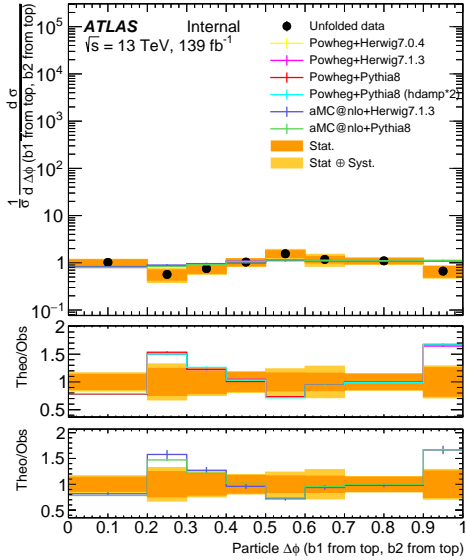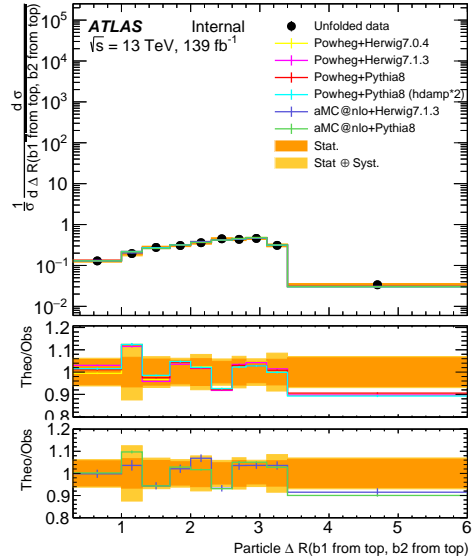
Figure 63: Unfolded distributions in the 3j3b region for: (a) vector sum p_T of the b-jets from top quarks decays, (b) invariant mass of the b-jets from tops, (c) $\Delta\phi$ between the b-jets from tops, (d) ΔR between the b-jets from tops.
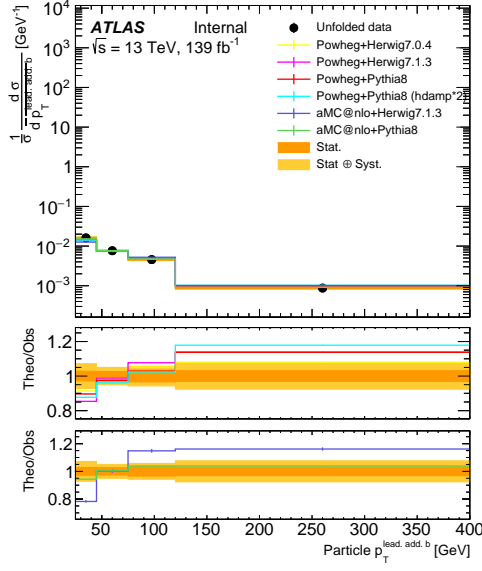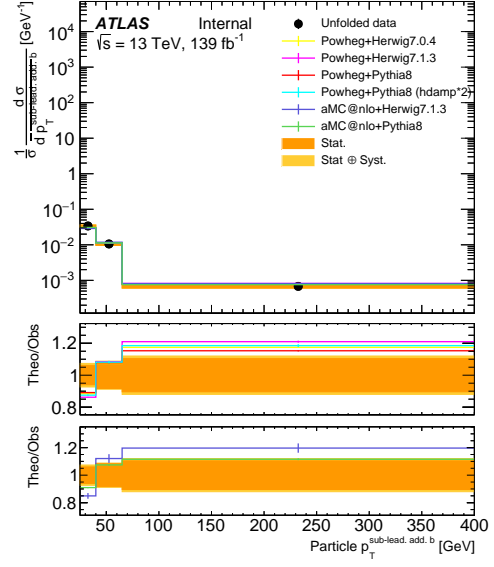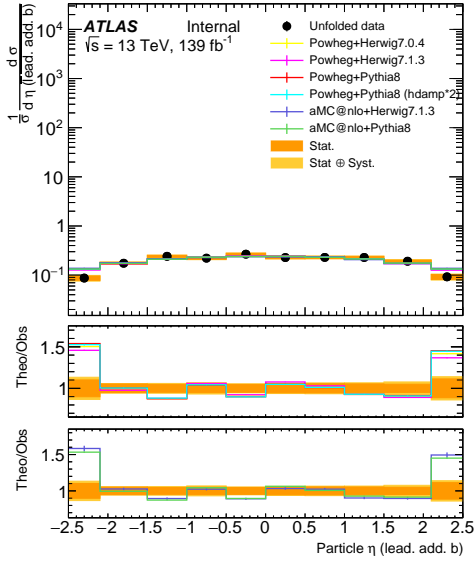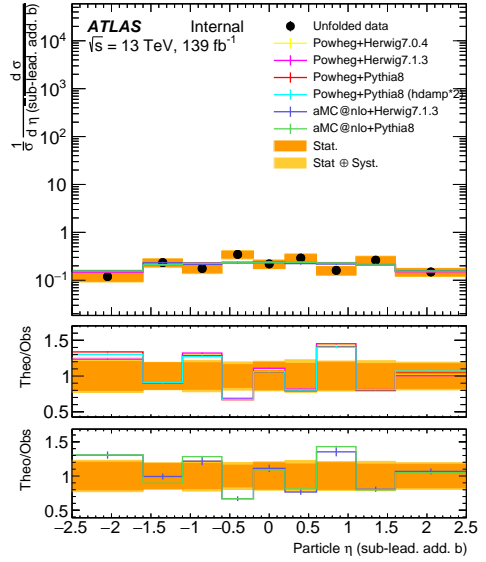
Figure 64: Unfolded distributions in the 3j3b region for: (a) the leading p_T additional b-jet, (b) the sub-leading p_T additional b-jet, (c) η of the leading additional b-jet, (d) η of the sub-leading additional b-jet.
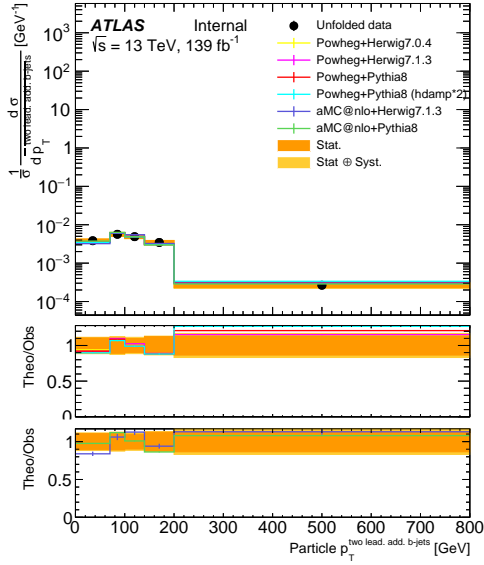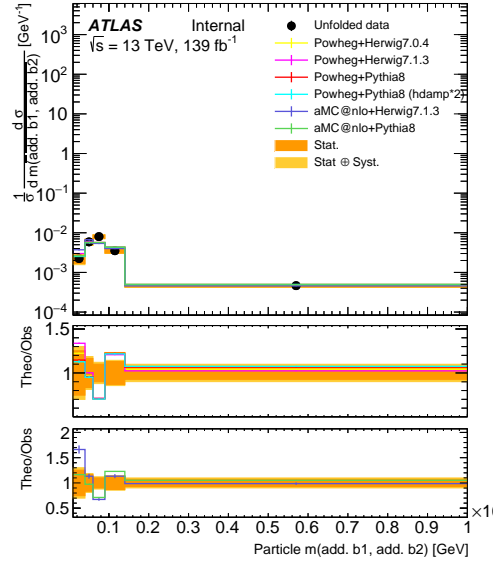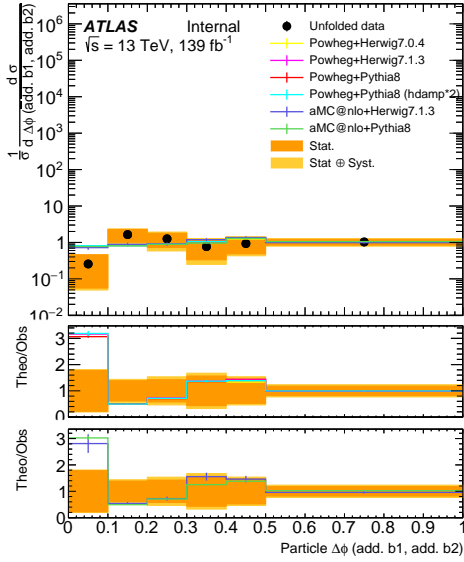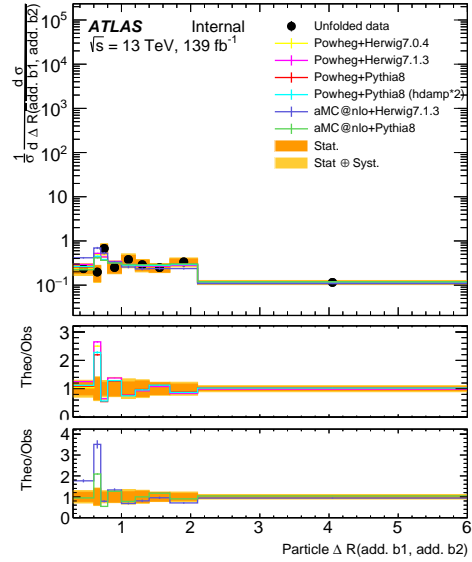
Figure 65: Unfolded distributions in the 3j3b region for: (a) vector sum p_T of the two leading additional b-jets, (b) invariant mass of the two leading additional b-jets, (c) $\Delta\phi$ between the two leading additional b-jets, (d) ΔR between the two additional b-jets.
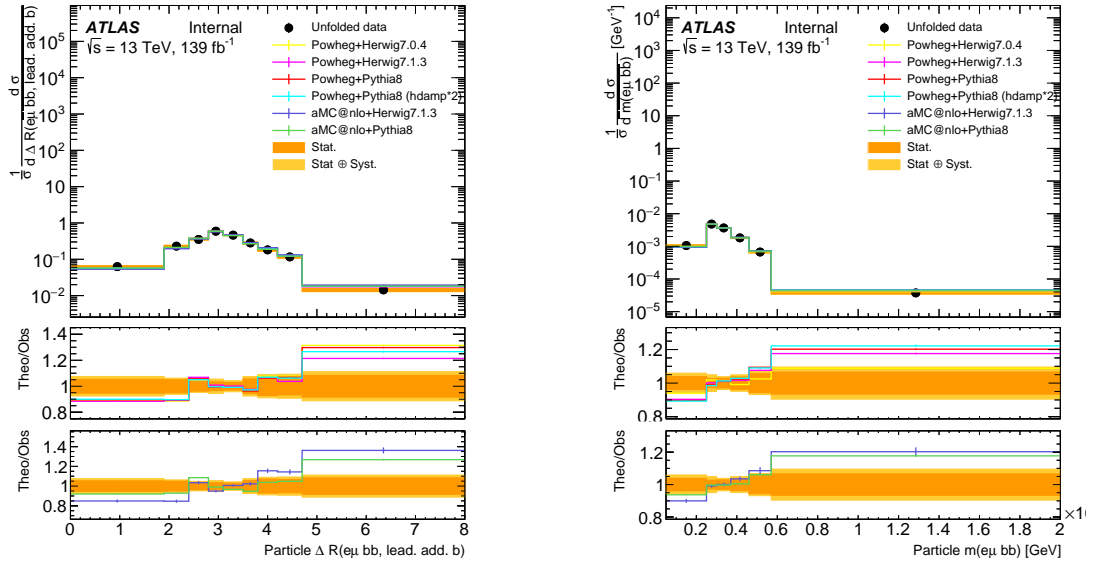
(a)   (b)

Figure 66: Unfolded distributions in the 3j3b region for: (a) ΔR between the $e\mu bb$ system and the leading additional b-jet, (b) invariant mass of the $e\mu bb$ system
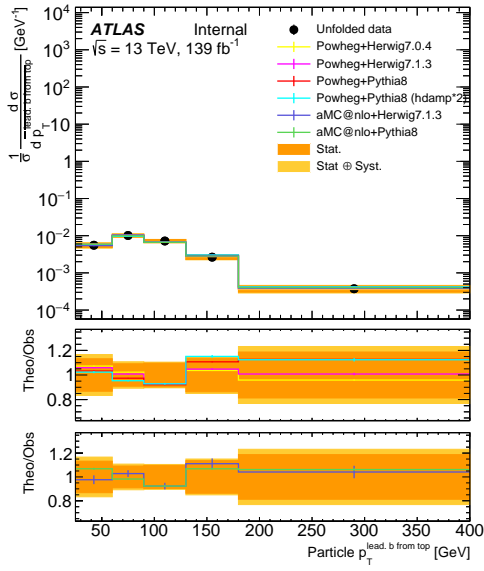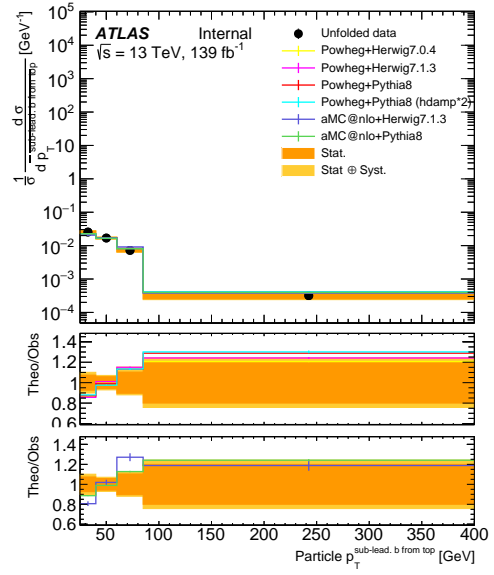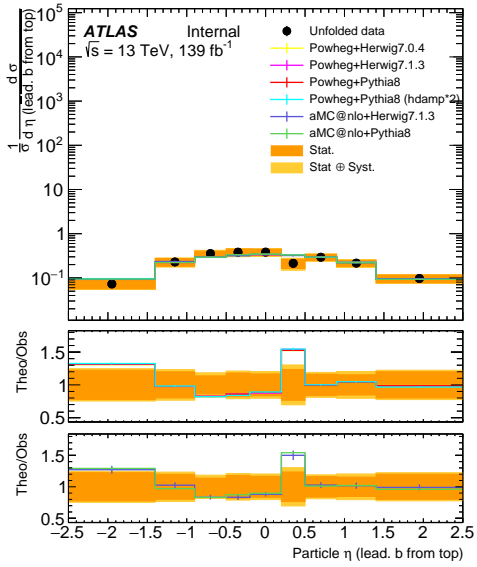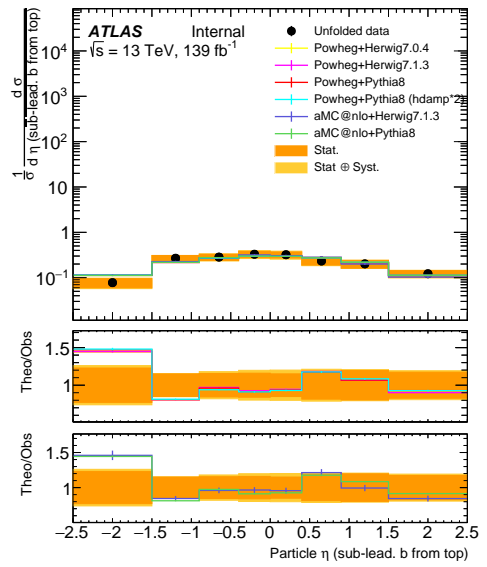
Figure 67: Unfolded distributions in the 4j4b region for: (a) the leading p_{T} b-jet from top quark decay, (b) the sub-leading p_{T} b-jet from top, (c) η of the leading b-jet from top, (d) η of the sub-leading b-jet from top.
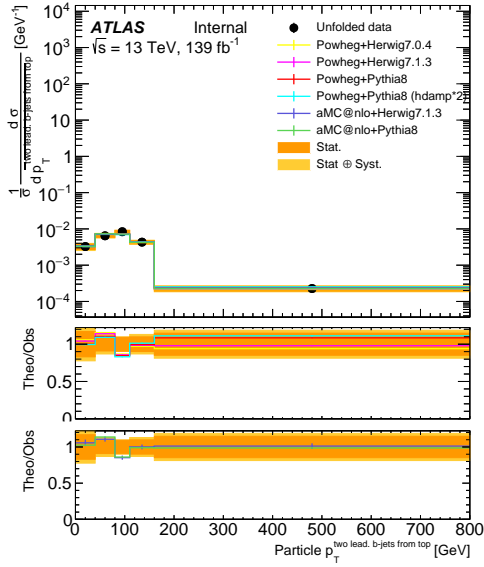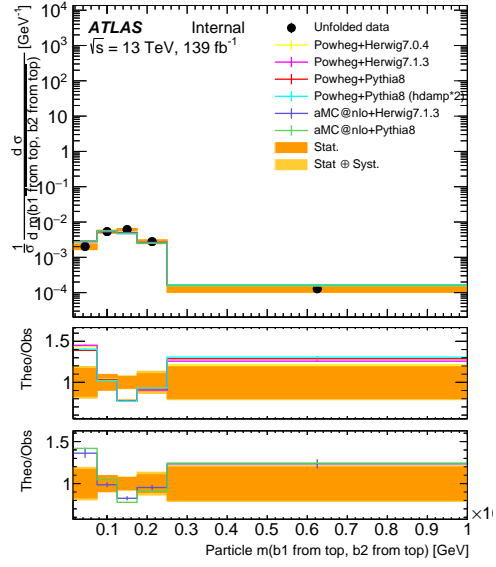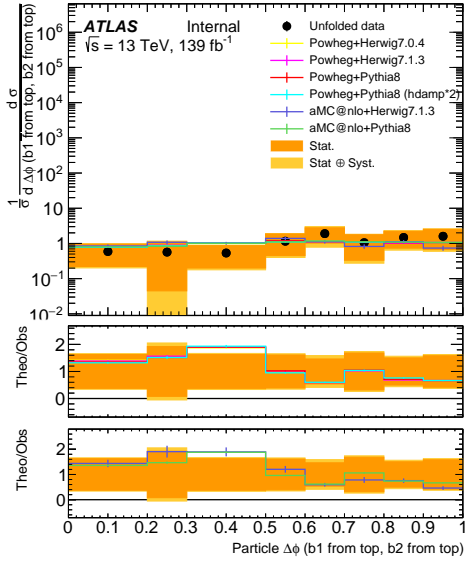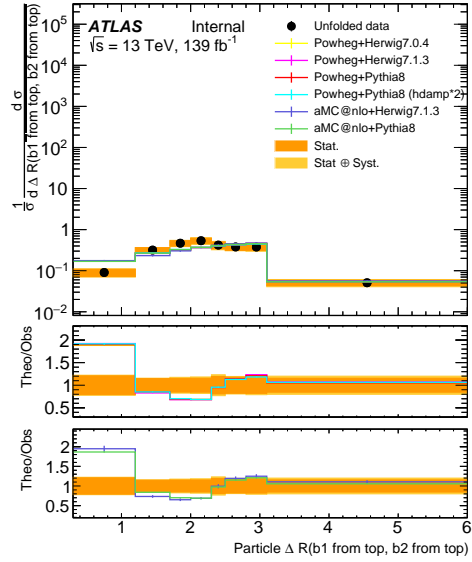
Figure 68: Unfolded distributions in the 4j4b region for: (a) vector sum p_T of the b-jets from top quarks decays, (b) invariant mass of the b-jets from tops, (c) $\Delta\phi$ between the b-jets from tops, (d) ΔR between the b-jets from tops.
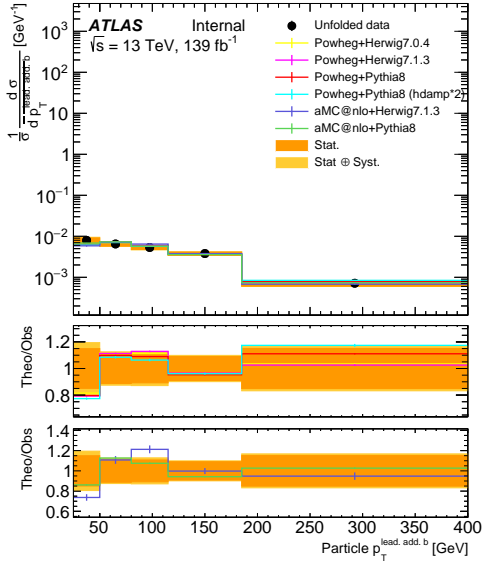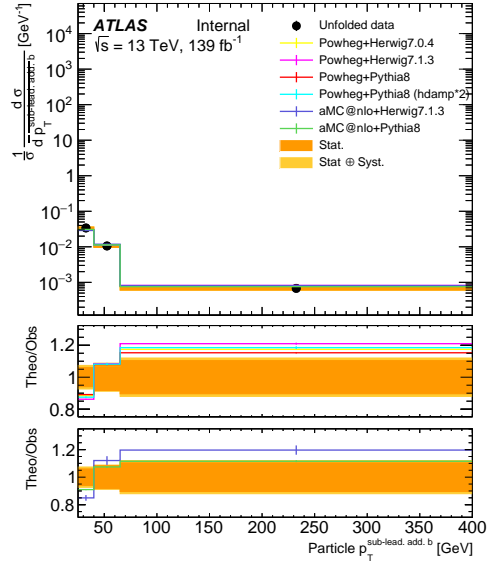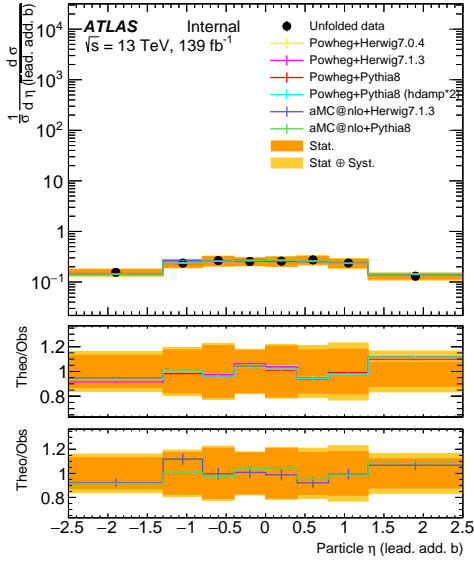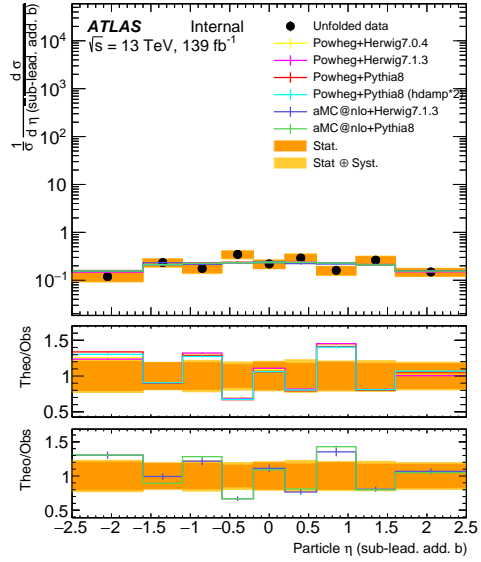
Figure 69: Unfolded distributions in the 4j4b region for: (a) the leading p_T additional b-jet, (b) the sub-leading p_T additional b-jet, (c) η of the leading additional b-jet, (d) η of the sub-leading additional b-jet.
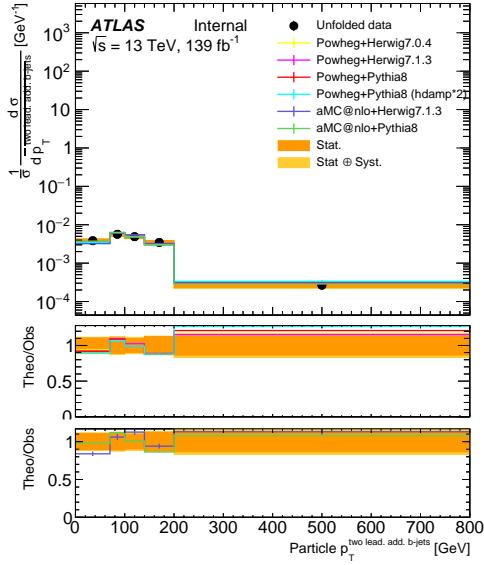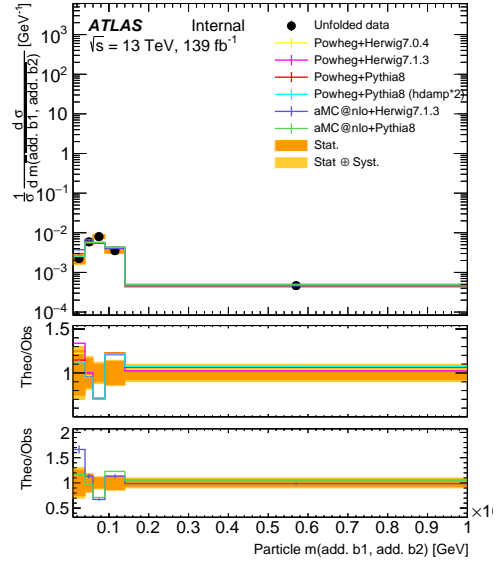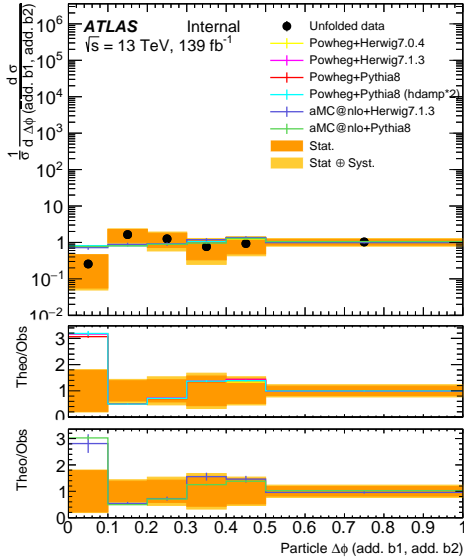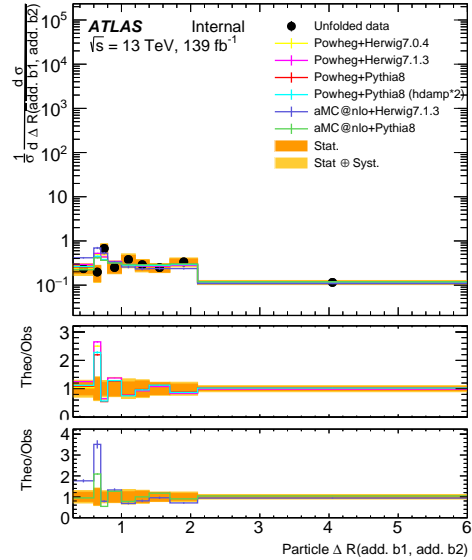
Figure 70: Unfolded distributions in the 4j4b region for: (a) vector sum p_{T} of the two leading additional b-jets, (b) invariant mass of the two leading additional b-jets, (c) $\Delta\phi$ between the two leading additional b-jets, (d) ΔR between the two additional b-jets.
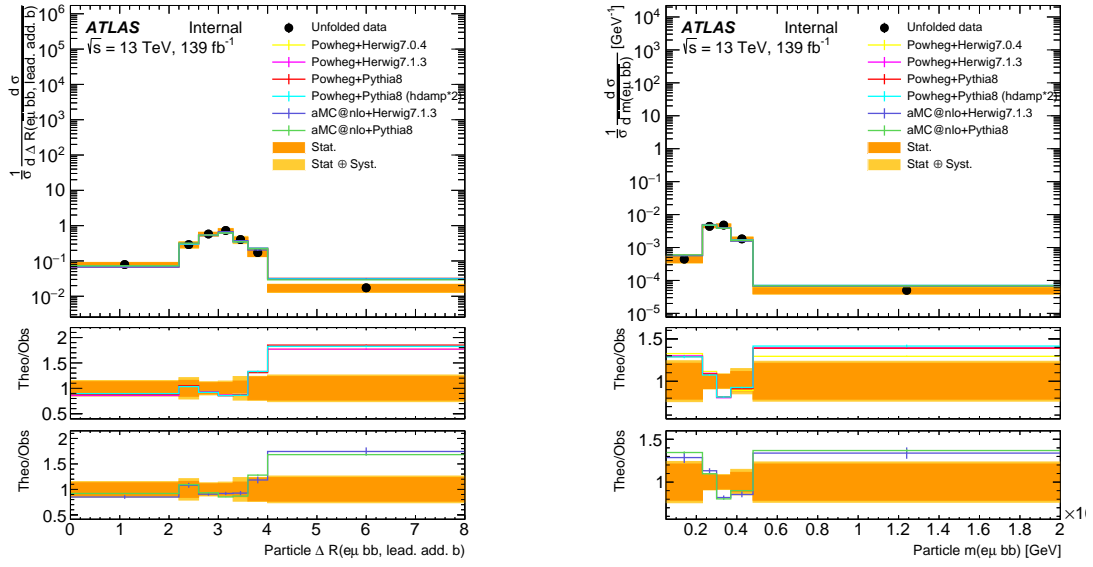
Figure 71: Unfolded distributions in the 4j4b region for: (a) ΔR between the $e\mu bb$ system and the leading additional b-jet, (b) invariant mass of the $e\mu bb$ system

CHAPTER VIII

Test of Pixel Readout Chips for the HL-LHC ATLAS Upgrade

In 2026, after the Run3, the LHC will stop for the third long shut down to be upgraded to the High Luminosity LHC [41, 45]. After the upgrade, the accelerator will deliver $5 - 7.5$ of the nominal luminosity of the original design, compared to the factor of 2 achieved so far. For the detectors at the LHC, higher luminosity results into higher data rate and higher radiation doses. To account for the new conditions, the detectors will also be upgraded. The ATLAS detector upgrade program for the HL-LHC is called Phase II Upgrade. In the upgrade, the entire ID of the ATLAS detector will be replaced with a new all silicon system, called the Inner Tracker. It will consist of 5 concentric layers of pixel detectors and 4 layers of silicon strips sensors. In this chapter, serial powering and DAQ tests of the prototypes of new pixel readout chips are presented.

8.1   RD53A Readout Chip

Pixel readout chips act as analog to digital converters: they receive analog signal from the silicon sensors, digitalize it, and a readout system collects the digital data. Therefore, the two most important components of a pixel readout chip are the analog front-end and the digital matrix. Besides that, there are a powering system and a trigger/clock system, which are also essential for operation of pixel readout chips.

The ATLAS and CMS pixel detector collaborations joint their efforts and started a new collaboration, known as RD53 [190], to design a new generation readout chip for the upgraded pixel detectors. The RD53A [191] is a large-scale demonstrator for the new readout chip. It is produced in 65 nm CMOS technology. The chip has dimensions of 20.0 mm by 11.8 mm, and its matrix is 400 pixels wide and 192 pixels tall with 50×50 μm^2 pixels. The chip is designed to work at up to 1.28 Gbit/s readout speed and to withstand 3 GHz/cm^2 hit-rate after being irradiated up to 500 Mrad. Pictures of the RD53a single chip card (SCC) and the readout chip are shown in figure 72. The SCC is used for laboratory tests of the RD53A readout chip. It utilizes a Display Port as a data interface. The presented information about the RD53A chip is taken from the RD53A Integrated Circuit [191] manual, unless stated otherwise.

(a)   (b)
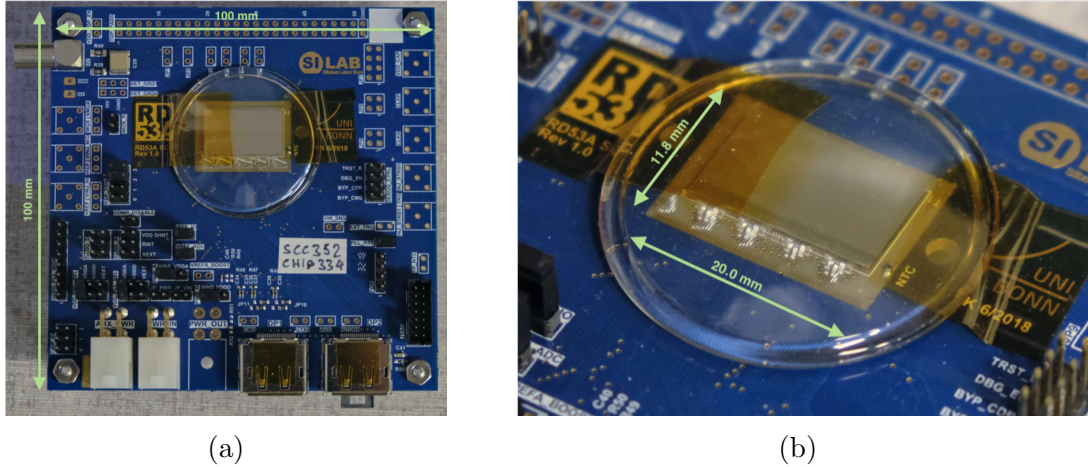
Figure 72: Photo of the (a) RD53A SCC and (b) the readout chip itself. The SCC serves as a host for one RD53A readout chip and allows for different configurations for testing purposes. Data interface is a Display Port cable.

A floorplan organization of the RD53A readout chip is shown in figure 73. The chip uses a 9 metal layer stack with one additional layer for power lines distribution. The top row contains tests pads for debugging, and it won't be included in the final production chip. All global analog and digital circuitry, which is needed to bias, configure, monitor and readout the chip, is placed in the bottom part of the chip. Wire bonding pads are organized in one row and located at the bottom edge of the chip on the padframe.
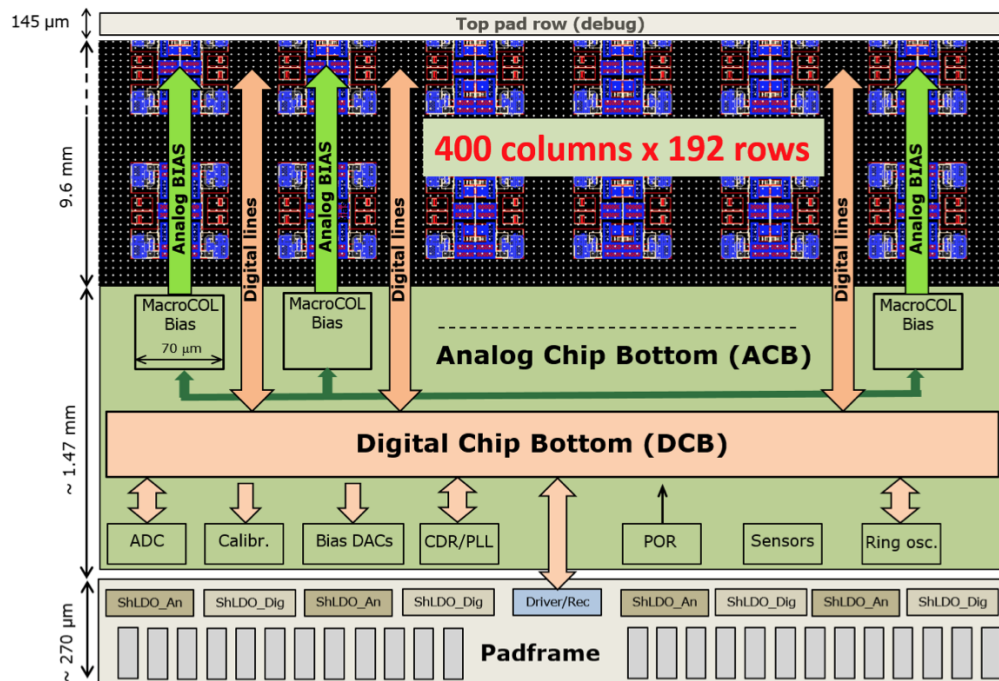
Figure 73: RD53A floorplan, functional view. Figure from [191].

Input Power and ShuLDO Regulator

To reduce passive material budget and allow for higher pixel sensors density, the ATLAS and CMS collaborations decided to adopt the serial powering (SP) scheme for the pixel readout chips. In ATLAS, each SP chain will consist of up to 13 readout modules. One readout module will host 3 or 4 readout chips powered in parallel. To ensure stable powering mode in each chip in one SP chain, special internal powering system must be developed. To stabilize voltage in the chip core and drawn exactly the needed amount of current, pixel readout chips use so-called ShuLDO regulators. Simplified scheme of the regulator is shown in figure 74. ShuLDO regulator is a combination of low-drop linear voltage regulator and a shunt regulator. LDO is used in the conventional voltage based power supply mode, and the the shunt regulator must be enabled in the current based serially powered powering mode.
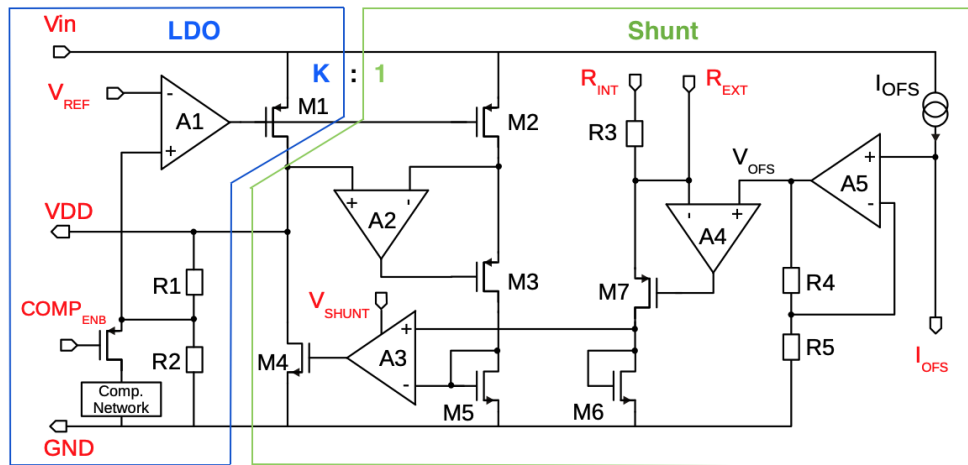
Figure 74: RD53A ShuLDO regulator, simplified scheme. Figure from [191].

The RD53A has two internal rails for powering the chip core, analog (VDDA) and digital (VDDD), each of which uses its own ShuLDO regulator. Voltage in an internal rail is generated from V_{REF}, and it is double of the corresponding reference: $VDD = 2 \times V_{REF}$. The SCCs are able to supply power to RD53A chip in three modes: the conventional voltage based mode using the LDO regulator only, the current based mode with Shunt enabled, and the direct powering mode by bypassing the ShuLDO regulator. Table 14 summarizes typical and maximal powering characteristics of different subsystems of the chip.

Pin	Type	Min	Typical	Max	Description
V_{in}	Power	1.4V		2.0 V	external power input (voltage)
V_{in}	Power	0	0.5 A	2.0 A	external power input (current)
V_{SHUNT}	Power	1.4 V		2.0 V	supply voltage of shunt circuitry
GND	Ground				local ground and shunt current output
VDD	Power	1.0 V	1.2 V	1.32 V	regulator output
V_{REF}	Analog	500 mV	600 mV	660 mV	reference voltage (VDD = $2V_{REF}$
R_{INT}	Analog		V_{IN}		supply/enable internal ref. R
R_{EXT}	Analog	300 Ω			external ref. R to V_{in}
I_{OFS}	Analog		200 kΩ		external offset R to GND
$COMP_{ENB}$	Digital		GND		GND to enable compensation

Table 13: ShuLDO regulator circuit pins legend. From [191]

Patameter	Typical	Maximal
Core direct supply voltage	1.2 V	1.32 V
ShuLDO input voltage	1.5 V	2.0 V
Per pixel analog current	4 μA	8 μA
Per pixel digital current	4 μA	6 μA
RD53a Periphery analog current	30 μA	60 μA
RD53a Periphery digital current	30 μA	60 μA
Output drivers (each)	20 μA	30 μA
Total RD53a current (4 outputs)	0.75 μA	1.3 A

Table 14: Powering characteristics of the RD53A chip. From [191]

Analog Front-End

To study analog front-end performance, the RD53A chip exploits three different designs of the analog front-ends, called Synchronous, Linear and Differential. The final readout chip will use the best design from the three. Figure 75 shows arrangement of the three analog front-ends on the pixels matrix. The Synchronous front end uses a baseline "auto-zeroing" scheme that requires periodic acquisition of a baseline instead of pixel-by-pixel threshold trimming. The Linear front end implements a linear pulse amplification in front of the

discriminator, which compares the pulse to a threshold voltage. The Differential front end uses a differential gain stage in front of the discriminator and implements a threshold by unbalancing the two branches.
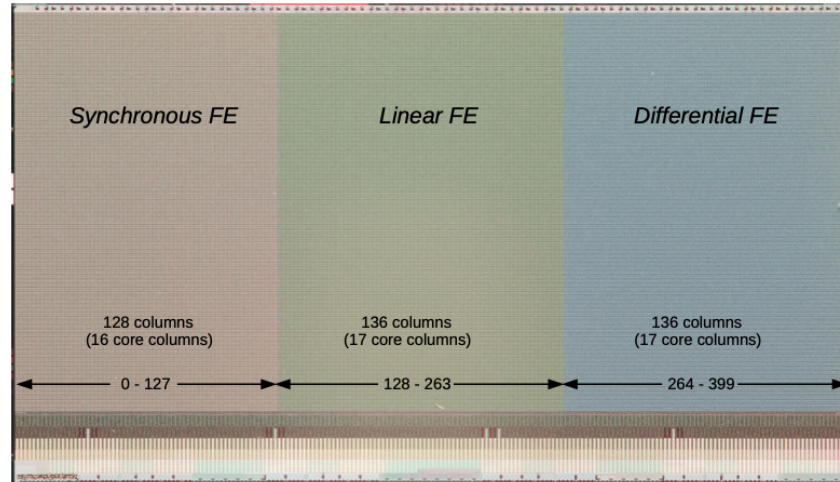
Figure 75: RD53A analog front-ends arrangement on the pixels matrix. Figure from [191].

Digital Matrix

The digital pixel matrix is built of digital cores. Each of the cores contains 8 pixel channels, corresponding to 4×4 analog islands. One digital core is synthesized as one digital circuit. The digital core handles all processing of the binary outputs, including masking, digital injection, Time over Threshold (ToT) counting, storage of ToT values, latency timing, triggering and readout.

Two digital core flavors are implemented in the RD53A chip, called Disturbed Buffer Architecture (DBA) and Central Buffer Architecture (CBA). For each of the core designs, the basic unit is called *region*. The DBA region size is 4 pixels, while CBA regions size is 16 pixels. For both, the timing information of a hit occurrence is stored by a region, not by individual pixels. The designs are different in the way they store ToT. In the DBA region, the ToT is stored by each individual pixel, while the CBA region stores ToT in the common memory for all pixels. Thus, the DBA core takes extra memory to store "meaningless" zero ToT values, but don't need a hip map, while the CBA surpasses zero ToT values resulting into smaller memory allocation, but requires to store the hit map. Therefore, the DBA architecture is efficient in cases of multiple hits per core (fewer zero ToTs), which is the case of small region. The CBA architecture is efficient with few hits and many zero ToTs pixels, which is the case of large regions .

Clock and Trigger

RD53A uses four clocks at different frequencies. Three of them are normally generated internally by the Clock and Data Recovery (CDR): the 160 MHz clock, the 1.28 GHz clock and the 640 MHz fine delay clock. When generated internally, the first and the last are derived from the 1.28 GHz clock. An external clock supply is also possible. In this case, the 160 MHz clock is used to generate the other two; however, external independent supply of 1.28 GHz clock is also possible. The last clock is 40 MHz clock, it is derived from the 160 MHz clock.

The 1.28 GHz clock is locked to the phase of the command input stream transitions, and used to produce the needed output. In operation, the 40 MHz clock must be synchronized with the beam crossings, and it is the only clock supplied to pixels.

Trigger pulses are normally generated by the readout chip, but can also be supplied externally. An internal self-trigger was not implemented in the RD53A chip.

Input, Output and Configuration

The RD53A chip is fully controlled with a single serial input stream with a custom encoding. All data, messages, and configuration read-back are output on a high speed serial port. There are 4 outputs each sending encoded data at 1.28 Gbps nominal bandwidth. Output data are serialized with the Aurora 64/66b protocol and transmitted on the multilane 1.28 Gbps port.

Control of the chip over a single differential serial input has been implemented using a simple custom protocol. The protocol provides encoded clock and commands on a single link. The custom protocol developed runs at 160 Mbps, which makes for better transmission on low mass cables and can be directly driven from GBT e-links.

8.2   Serial Powering Tests

Reliability of the SP chain must be studied before being implemented in the actual detector. SP Task Force raised a concern about effects of oscillating voltage references leading to a noisy chip in one module. Initial studies of the problem are presented in the section. The studies were done with two RD53A SCCs.

Pixel readout modules consists of three (triplets) or four (quads) readout chips. Readout chips inside of one module are powered in parallel, while modules in one SP chain are powered in series. Illustration of the SP chain of RD53A modules is shown in figure 76. To replicate both the *in module* and *two modules* scenarios, the parallel and serial powering configurations of two SCCs were studied.
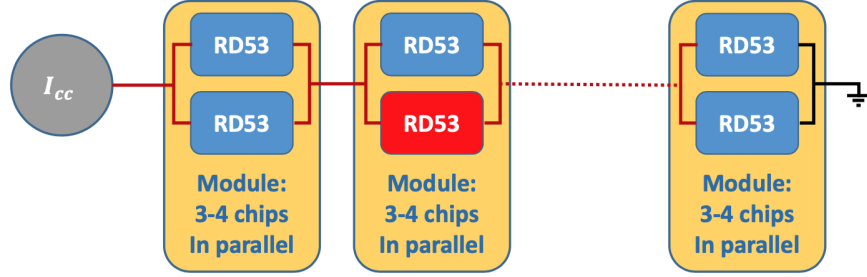
Figure 76: Illustration of a SP chain of RD53A modules powered by a constant current source I_{CC}. Red chip is assumed to have noise in its references.

AC Noise in V_{REF}

First, effects of oscillating V_{REF} in the ShuLDO were studied. AC noise of 50 mV in the amplitude was supplied to V_{REF}, and the frequency varied; an illustration of the experimental setup is shown in figure 77. Amplitude of oscillations in VDDD and ratio of the value to the supplied AC voltage are shown in figure 78. At frequencies below 100 Hz, the oscillations from $V_{REF.D}$ are transmitted into VDDD with the factor of 2 in the amplitude, as it is governed by the LDO functionality. At higher frequencies, the transmitted oscillations faint and approach 0 amplitude at 10 kHz. The noise in VDDD was not transmitted to the power output of the SCC. The study was done in both LDO and ShuLDO powering modes for both VDDD and VDDA, and yielded to similar results. Therefore, AC noise in the voltage reference of one chip in an RD53A module will not affect other chips in the SP chain.
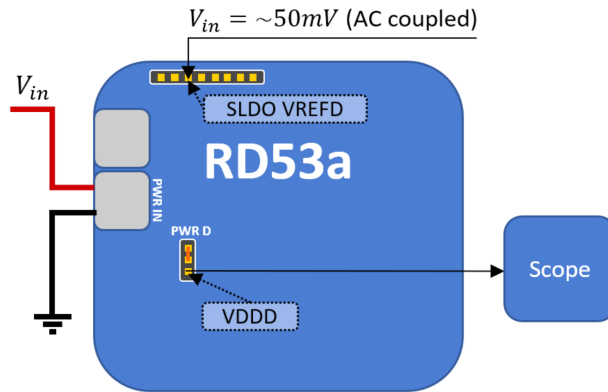
Figure 77: Illustration of the experimental setup for studies of oscillations in V_{REF} to the internal rail and output voltage of RD53A SCC.

(a)   (b)
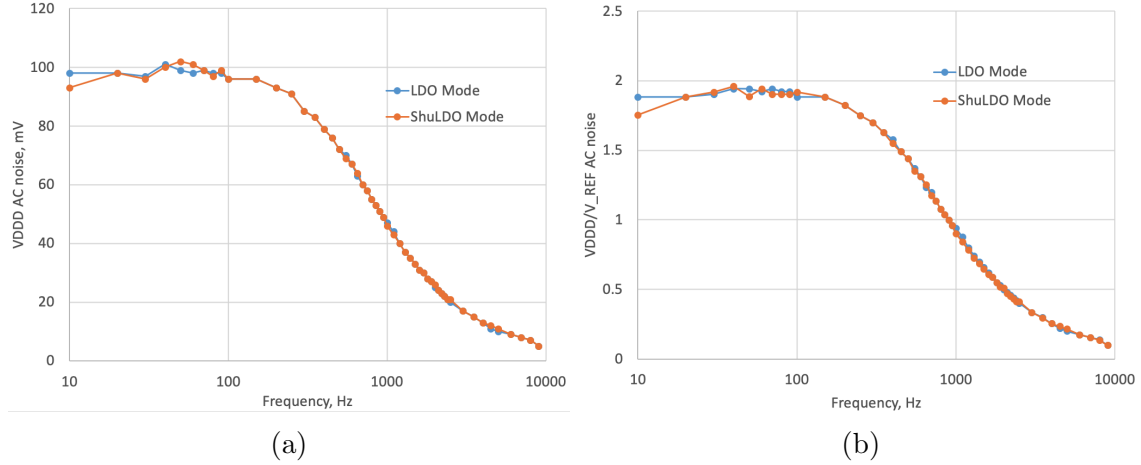
Figure 78: Results of the studies of oscillations in $V_{REF.D}$ to VDDD. (a) Transmitted oscillations from $V_{REF.D}$ to VDDD at different amplitudes; the transmission factor starts at 2, as governed by the LDO functionality, and faints with the increase in the frequency of the oscillations. (b) Ratio of amplitudes of the transmitted to VDDD oscillations to the injected AC noise.

AC Noise in I_{OFS}

The second set of studies was done with AC noise in the offset current of the Shunt regulator. AC noise from an external power supply was injected into I_{OFS} of the analog ShuLDO regulator. An illustration of the experimental setup is shown in figure 79. The injected noise causes oscillations in the output (local GND) of the RD53A SCC with the same frequency, but the amplitude decreases with the increasing frequency. The amplitude vs. frequency graph is shown in figure 80.

(a)   (b)
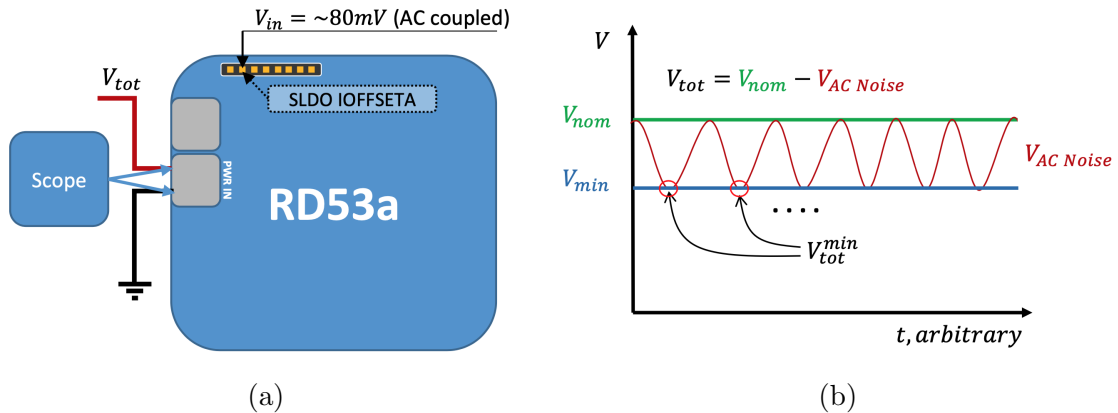
Figure 79: (a) Illustration of the experimental setup for modulation of the AC noise in the output of a RD53A SCC due to noise in the offset current of the ShuLDO regulator. (b) Sketch of the adjustment of modulated AC noise in the output of the RD53A SCC: the drop voltage across the SCC must oscillate such that it doesn't drop below the minimal input value.

131
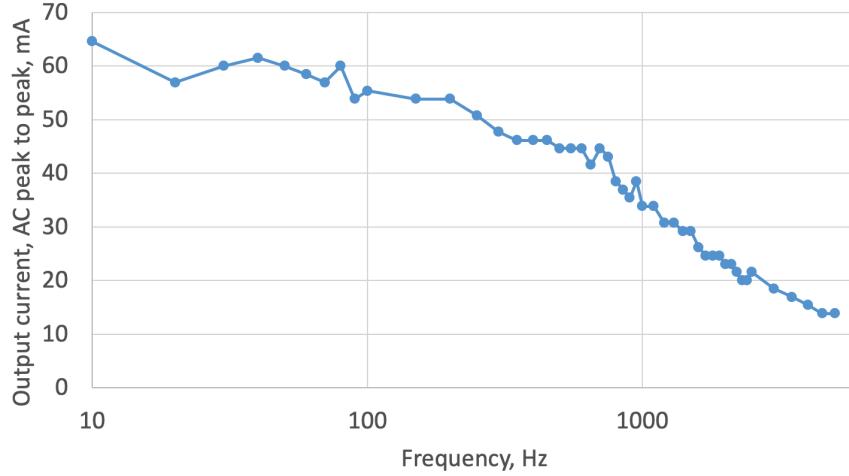
Figure 80: Voltage oscillations in the output of the RD53A SCC due to oscillating current (80 mA max) injected into the analog I_{OFS}.

The effect of oscillating output of RD53A SCC due to AC noise in the I_{OFS} was used to study setups of two SCCs powered in parallel and series. Illustrations of the experimental setups are shown in figure 81. The SCC with AC noise in its output is called the aggressor chip, and the second chip is called the victim chip.
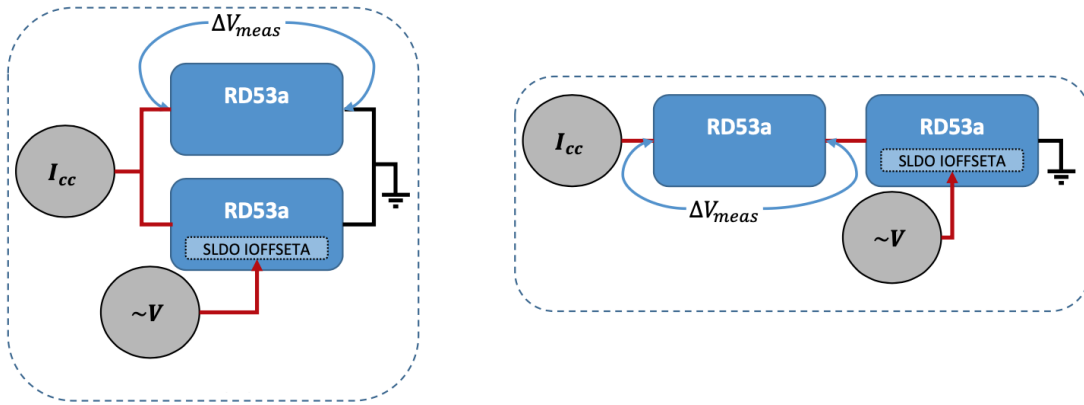
Figure 81: Illustrations of the experimental setups for studies of the effect of AC noise in the output of the RD53A chip if the parallel (left) and serial (right) powering modes.

In the serial powering scheme it was found that AC noise from a SCC propagates to the higher potential side of the chain and does not propagate to the lower. However, the victim chip on the higher potential side from the aggressor remain under the constant voltage drop, because both the input and output voltage levels oscillate at the same frequency with the same amplitude, as shown in figure 82. In the parallel powering mode, both the aggressor and the victim experience the same oscillations in voltage drop across the SCCs.
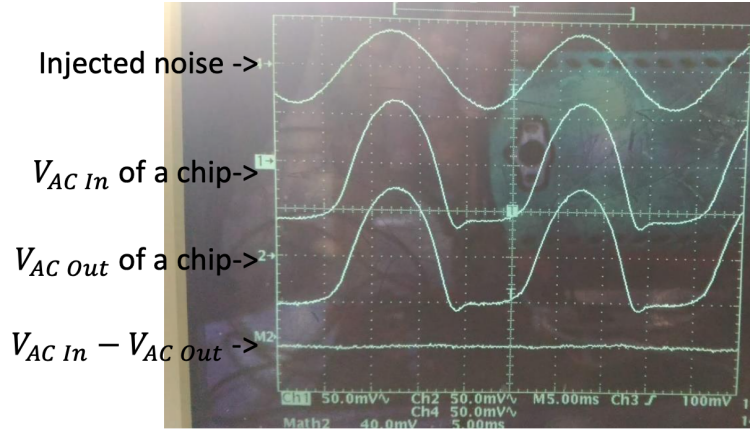
Figure 82: From top to bottom, oscillations in: I_{OFS} of the aggressor chip, input voltage of the victim chip, output voltage of the victim chip, voltage drop across the victim chip.

In each pixel, signal is detected when the collected charge exceeds threshold value. To ensure efficient signal detection, the threshold should be set low enough to sense small signals, but above noise in the electrical circuitry. Effects of AC oscillations in the I_{OFS} and in the powering line to noise in pixels in the differential and linear analog front-end of RD53A matrix were studied.

Figures 83 and 84 summarize result of the tests in both the serial and parallel powering configurations for the aggressor and victim SCCs respectively. The AC oscillations in the studied frequencies range did not affect noise in pixels in the linear and differential analog front-ends. In the serial powering configuration with the AC oscillations at 100 Hz, linear front-end of the aggressor SCCs returned results different from the general trend due to different tuning of the matrix (figure 83d). The linear front-end of the victim SCC showed increase in the average noise in pixels with the decrease in threshold. The effect is due to columns of noisy pixels in the matrix, and not related to the AC noise in the powering line; the analog matrix tuned to 1200 e and 600 e threshold is shown in figure 85.
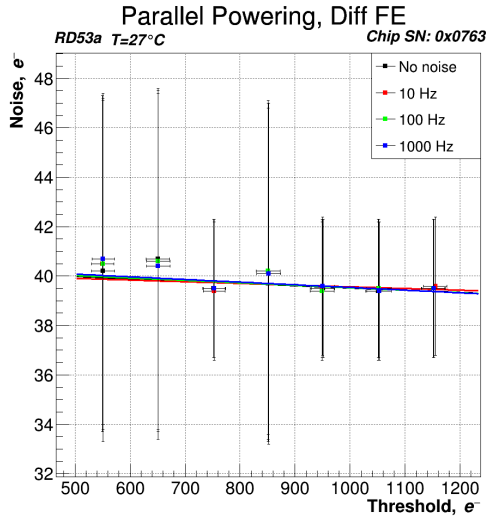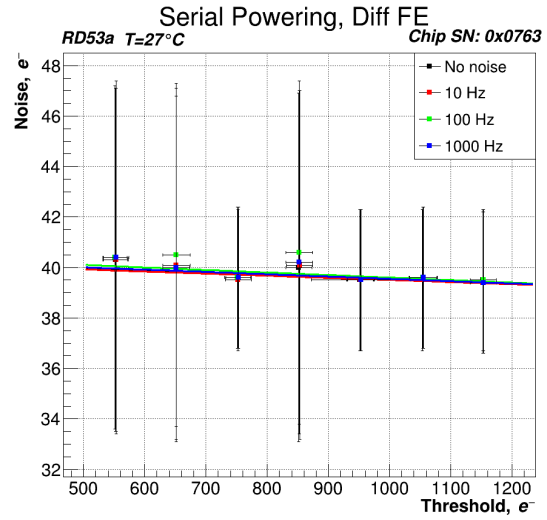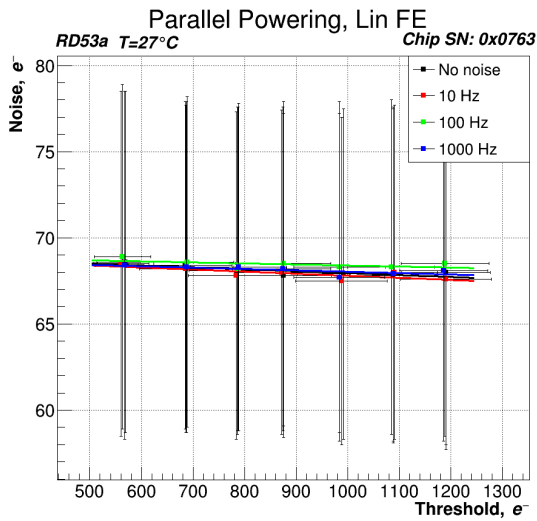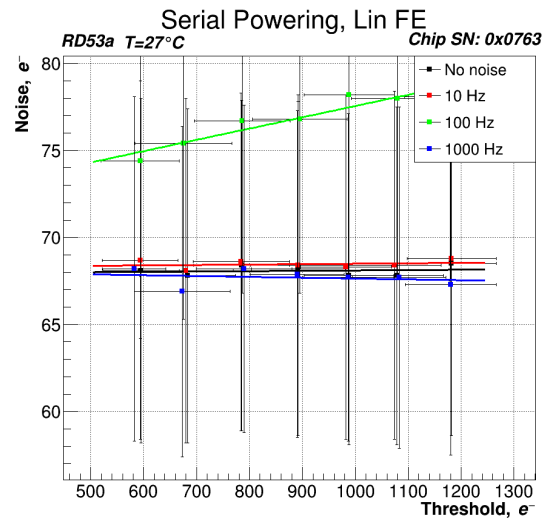
Figure 83: Average noise per pixel in the differential (a, b) and linear (c, d) analog front-ends of the aggressor RD53A SCC in the parallel and serial powering configurations. Difference in the linear front-end at 100 Hz oscillations in the serial powering mode is due to different tuning.

134
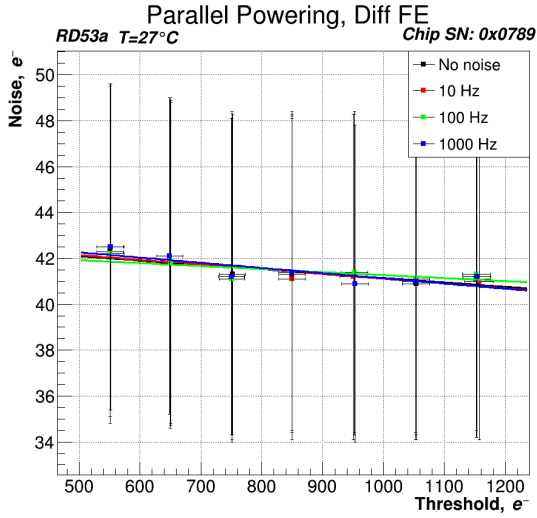
(a)
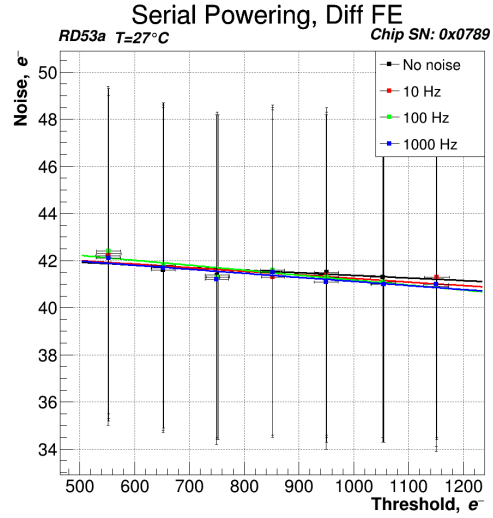
(b)
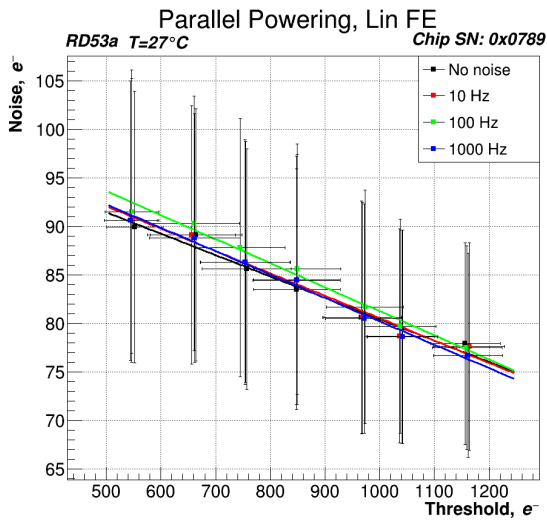
(c)
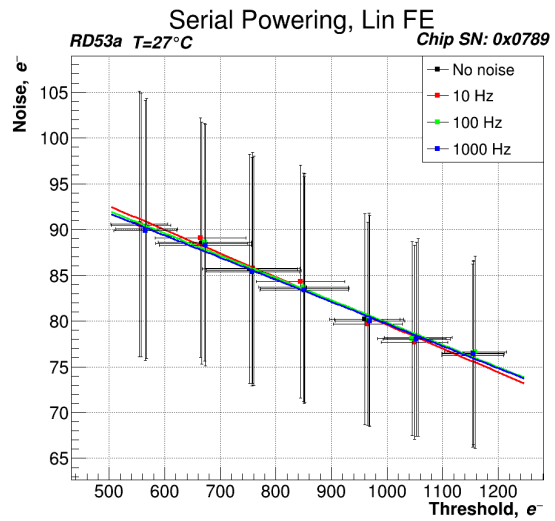
(d)

Figure 84: Average noise per pixel in the differential (a, b) and linear (c, d) analog front-ends of the victim RD53A SCC in the parallel and serial powering configurations. Linear increase of the average noice in the linear front-end is due to columns of noisy pixels, shown in figure 85.
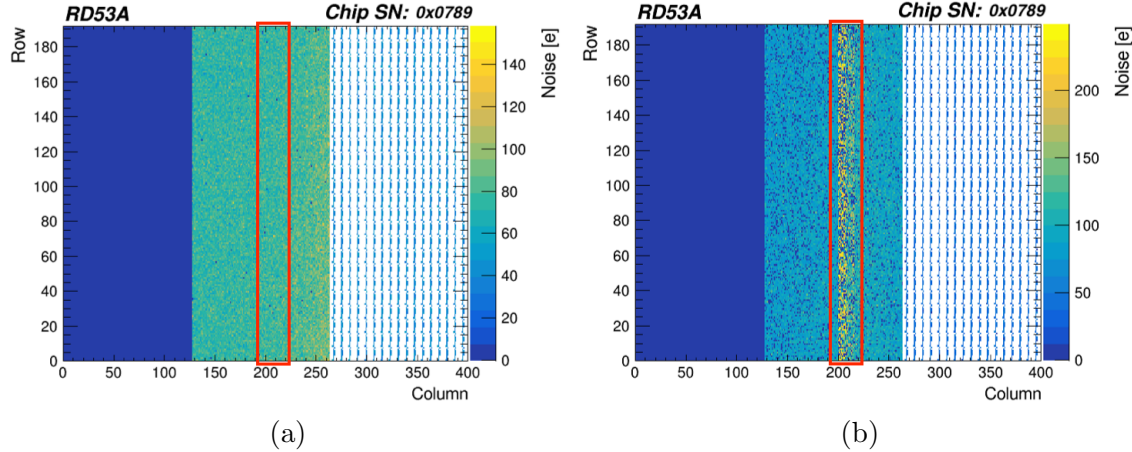
Figure 85: Columns of noisy pixels (red frame) in the liner front-end of one of the studied RD53A. No AC noise was supplied to the chip's references.

8.3 Readout Tests

In the ITk, data cables of a certain length and quality will be used to transfer data from the pixel readout modules. Preliminary specifications for the ITk require a readout system to successfully receive and decode signal from the pixel modules using cables with 18 bB loss before irradiation. Tests of three readout systems, YARR [192], RCE [193] and FELIX [194] are presented in the section.

Each of the readout systems uses display port cables to connect to the RD53A SCC. The cables replicate data cables from the final design of the detector. However, the interface will change. The purpose of the study is to evaluate performance of the three readout systems with losses in the readout cables of up to 18 dB. The study uses a set of 10 feet long display port cables with 6 dB losses in data channels, and one short cable with loss less than 3 dB. Measurements with the latter cable are used as a reference for all readout systems. To study performance of readout systems with cables with losses, one cable of 6 dB loss and chains of two and three cables, resulting into 12 and 18 dB losses combined respectively, were used.

"Yet Another Rapid Readout"

Yet Another Rapid Readout (YARR) is a readout system based on the concept of moving intelligence from FPGA firmware into the host computed software. In YARR, the FPGA serves the role of a reconfigurable input/output, and not a hardware accelerator. Such principle allows for easier access to modifications for developers and testers, and flexibility in the software makes it possible to support a wide range of FPGA platforms.

A test stand for YARR readout system was assembled, and its principle block diagram is shown in figure 86. The test stand is build around a CentOS computer that can host up to 4 FPGA cards. Trenz TEF-1001 card with an Ohio card is used for the setup. The Ohio card is an interface board with 4 mini display ports that is mounted directly on the Trenz

FPGA card. The test stand uses the YARR software [195] and works at 640 Mbps readout speed, which half of the nominal readout speed to the final pixel readout chips in ITk.
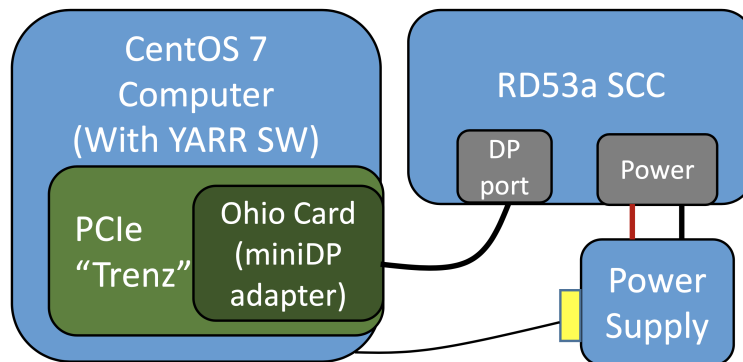
Figure 86: Principal block diagram of the YARR test stand. The test stand consists of a CentOS computer with Trenz TEF-1001 FPGA card. An Ohio Card, which serves an a miniDP adapter, is mounted on the FPGA card. An RD53A SCC is powered by a bench power supply that can be controlled from the computer.

The parameter of evaluation for the YARR readout system was successful performance of a digital scan of an RD53A SCC. Failure to retrieve and decode data from the SCC means that the system cannot work with cables of certain signal loss. The experimental studies showed that YARR readout system is able to work with signal losses in cables of the order of 6 dB, but fails at 12 dB.

Reconfigurable Cluster Element

Reconfigurable Cluster Element (RCE) platform is a general purpose clustered data acquisition system implemented on a custom ATCA compliant blade. The central element is a system-on-chip design based upon the Xilinx Zynq family of FPGAs. The RCE is a combination of software and firmware which is stored locally on an SD-card. To control the readout system, a computer running on Unix operational system is needed.

A principal block diagram of a test stand for RCE is presented in figure 87. The test stand is build around the ZCU102 evaluation board. YARR software optimized for RCE is installed on the board. An adapter card with mini fisplay ports is mounted on ZCU102. A CentOS computer is used to access the board and to control a bench power supply that powers an RD53A SCC under the tests. The system can operate at the full readout speed of pixel chips, 1.28 Gbps.

The parameters of evaluation of the RCE system were successful performance of a digital scan of an RD53A SCC and the eye diagram of the signal received by ZCU102. The experimental measurements showed that RCE is able to work with cables with losses in signal up to 18 dB.
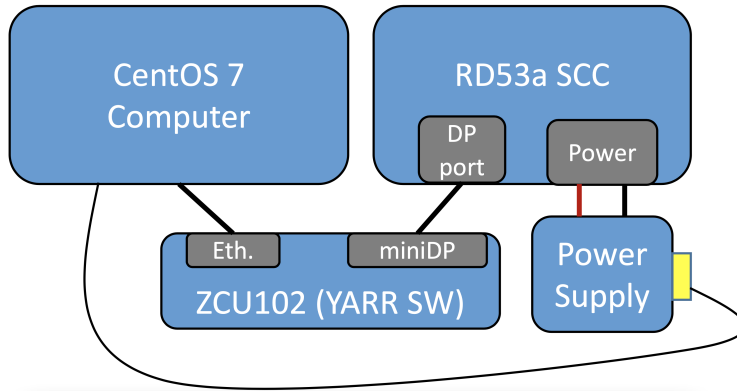
Figure 87: Principal block diagram of the RCE test stand. The test stand consists of a CentOS computer, ZCU102 evaluation board with an adapter card with 4 display ports. An RD53A SCC is powered by a bench power supply that can be controlled from the computer.

Front-End Link Exchange

Front-End Link eXchange (FELIX) is a new detector readout component being developed as part of the ATLAS upgrade effort. FELIX is designed to act as a data router, receiving packets from detector front-end electronics and send it to programmable peers on a commodity high bandwidth network. FELIX is intended to unify all readout across one well supported and flexible platform, compared to previous detector readout implementations relied on diverse custom hardware platforms. Detector data processing is be implemented in software hosted by commodity server systems subscribed to FELIX data, compared to previous detector FPGA-based data processing.

A test stand for direct FELIX readout system was assembled, and its principle block diagram is shown in figure 88. The test stand consists of a few key components. The first is a CentOS computer with a FLX-712 PCIe card. A Timing Mezzanine Card (TMC) is mounted on the PCIe card and serves as an interface for MTP cable. The second component is the Versatile Link Demonstrator Board (VLDB) with a GBT-FPGA chip. The board has an SFP connector for an MTP cable running from the computer and multiple miniHDMI ports. VLDB serves as a generator of digital command signal for RD53A from incoming analog signal from the computer. The USB JTAG dongle is used to install firmware on the VLDB. The third key component of the setup is the Interface Board. The interface board receives digital command from VLDB and transfers it to RD53A SCCs connected to the board. Digital signal generated by the RD53A SCCs is received by the board, and transferred to the computer in analog format (light). A miniPod is used to convert the digital signal into light. FELIX readout system operates at 1.28 Gbps readout speed.
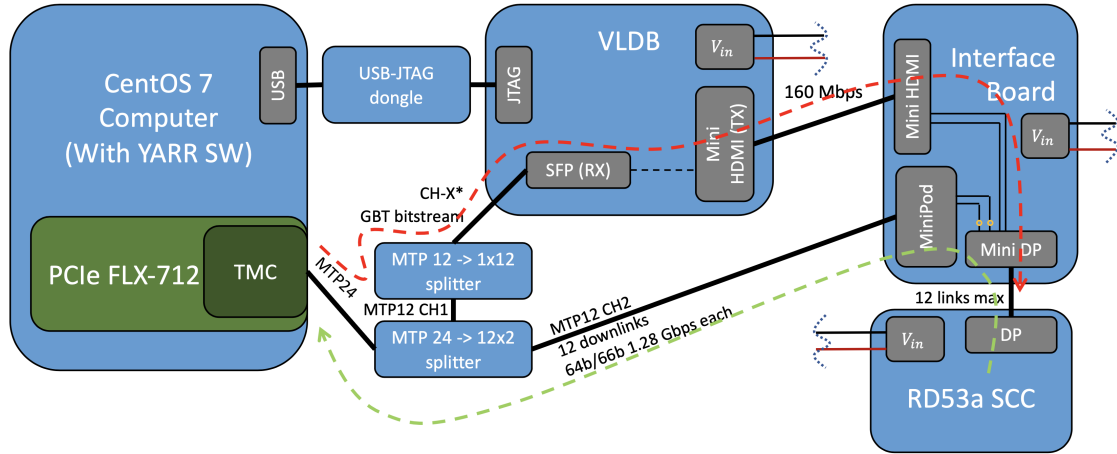
Figure 88: Principal block diagram of the FELIX test stand. The test stand consists of a CentOS computer

The parameters of evaluation of the FELIX system were successful performance of a digital scan of an RD53A SCC and the eye diagram of the signal received by the interface board. Although, the latter cannot serve as an ultimate proof since the signal is being converted into light and amplified by a miniPod on its way to the computer where the software works, it is a good parameter of evaluation in the mid-point of the readout chain. Figures 89 and 90 shows eye diagrams measured on the RD53A SCC and on the interface board in experiments with a short display port cable and a chain of cables with the total loss of 12 dB respectively. Form the figures it follows that the jitter increases significantly with increase in loss of cables, which at some point makes it impossible to decode the signal by the computer. The experimental studies showed that direct FELIX readout system is able to work with signal losses in cables of the order of 6 dB, but fails at 12 dB.
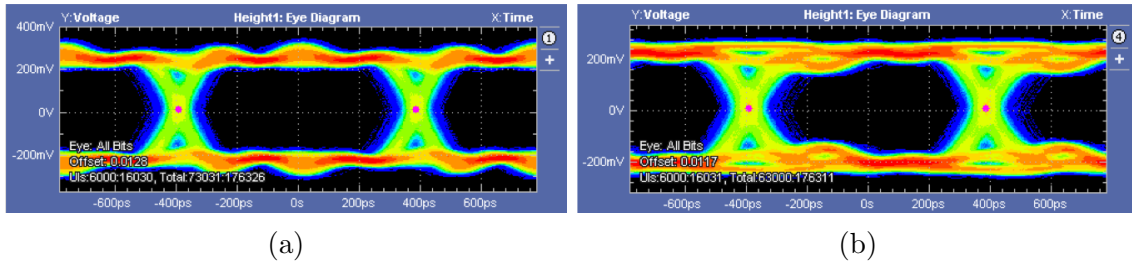
Figure 89: Eye diagrams of signal generated by an RD53A chip measured at (a) the RD53A SCC and (b) the interface board. Some increase of the jitter is observed in the eye diagram measured on the interface board.
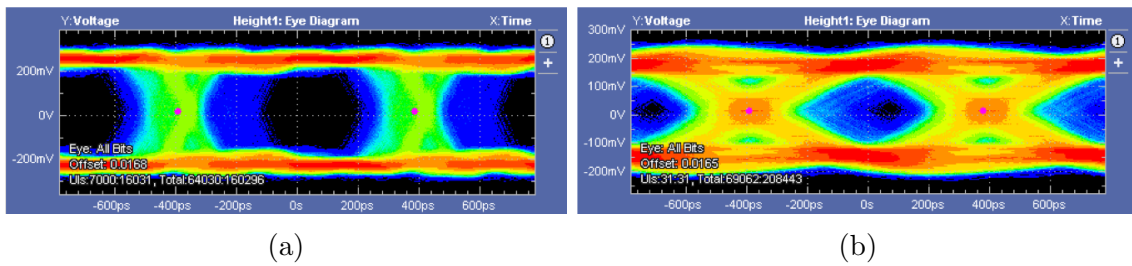
Figure 90: Eye diagrams of signal generated by an RD53A chip measured at (a) the RD53A SCC and (b) the interface board. Significant increase of the jitter is observed in the eye diagram measured on the interface board.

Results of the studies are summarized in table 15. Only the RCE readout system meets requirements of the specifications for losses in data cables. However, the FELIX system is still under the development, and future upgrades will address the problems with the signal recovery with data cables with high signal losses.

System	<3 dB	6 dB	12 dB	18 dB
FELIX	works	works	fails	fails
YARR	works	works	fails	fails
RCE	works	works	works	works

Table 15: Summary of the readout performance tests of three different readout systems with data cables with losses in signal lines.

140

REFERENCES

[1] J. J. Thomson. Cathode rays. *Phil. Mag. Ser. 5*, 44:293–316, 1897.

[2] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[3] Abdus Salam. Gauge Unification of Fundamental Forces. *Rev. Mod. Phys.*, 52:525–538, 1980.

[4] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

[5] LHC Machine. *JINST*, 3:S08001, 2008.

[6] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[7] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[8] The ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, Sep 2012.

[9] The CMS Collaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61, Sep 2012.

[10] Particle Data Group. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.

[11] The CDF Collaboration. Observation of top quark production in pp collisions with the collider detector at fermilab. *Physical Review Letters*, 74(14):2626–2631, Apr 1995.

[12] The D0 Collaboration. Search for high mass top quark production in pp collisions at $\sqrt{s} = 1.8$ tev. *Physical Review Letters*, 74(13):2422–2426, Mar 1995.

[13] John Daintith. *A Dictionary of Physics*. Oxford University Press, 2009.

[14] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken Symmetries. *Phys. Rev.*, 127:965–970, 1962.

[15] Peter W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12:132–133, 1964.

[16] Alexandre Deur, Stanley J. Brodsky, and Guy F. de Téramond. The qcd running coupling. *Progress in Particle and Nuclear Physics*, 90:1–74, Sep 2016.

[17] Csaba Csáki and Philip Tanedo. Beyond the Standard Model. In *2013 European School of High-Energy Physics*, pages 169–268, 2015.

[18]

[19] D. V. Volkov and V. P. Akulov. Is the Neutrino a Goldstone Particle? *Phys. Lett. B*, 46:109, 1973.

[20] J. Wess and B. Zumino. Supergauge Transformations in Four-Dimensions. *Nucl. Phys. B*, 70:39, 1974.

[21] J. Wess and B. Zumino. Supergauge Invariant Extension of Quantum Electrodynamics. *Nucl. Phys. B*, 78:1, 1974.

[22] S. Ferrara and B. Zumino. Supergauge Invariant Yang-Mills Theories. *Nucl. Phys. B*, 79:413, 1974.

[23] Abdus Salam and J. A. Strathdee. Supersymmetry and Nonabelian Gauges. *Phys. Lett. B*, 51:353, 1974.

[24] STEPHEN P. MARTIN. A supersymmetry primer. *Advanced Series on Directions in High Energy Physics*, page 1–98, Jul 1998.

[25] Savas Dimopoulos and David Sutter. The supersymmetric flavor problem. *Nuclear Physics B*, 452(3):496–512, Oct 1995.

[26] Stephen Myers. The LEP collider, from design to approval and commissioning. 10 1991.

[27] O. Bruning, H. Burkhardt, and S. Myers. The Large Hadron Collider. *Prog. Part. Nucl. Phys.*, 67:705–734, 2012.

[28] Linear accelerator 2. Sep 2012.

[29] Maurizio Vretenar, J Vollaire, R Scrivens, C Rossi, F Roncarolo, S Ramberger, U Raich, B Puccio, D Nisbet, R Mompo, S Mathot, C Martin, L A Lopez-Hernandez, A Lombardi, J Lettry, J B Lallement, I Kozsar, J Hansen, F Gerigk, A Funken, J F Fuchs, N Dos Santos, M Calviani, M Buzio, O Brunner, Y Body, P Baudrenghien, J Bauche, and T Zickler. *Linac4 design report*, volume 6 of *CERN Yellow Reports: Monographs*. CERN, Geneva, 2020.

[30] The Super Proton Synchrotron. Jan 2012.

[31] P. M. Watkins. DISCOVERY OF THE W AND Z BOSONS. *Contemp. Phys.*, 27:291–324, 1986.

[32] Esma Mobs. The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019. Jul 2019. General Photo.

[33] Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC. Technical report, CERN, Geneva, Jun 2019.

[34] The ATLAS Collaboration. Atlas data quality operations and performance for 2015–2018 data-taking. *Journal of Instrumentation*, 15(04):P04003–P04003, Apr 2020.

[35] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008.

[36] A. Augusto Alves, Jr. et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[37] G. Anelli et al. The TOTEM experiment at the CERN Large Hadron Collider. *JINST*, 3:S08007, 2008.

[38] The LHCf Collaboration. The LHCf detector at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08006–S08006, aug 2008.

[39] James Pinfold. The MoEDAL experiment at the LHC. *EPJ Web Conf.*, 145:12002, 2017.

[40] Michaela Queitsch-Maitland. FASER: Forward Search Experiment at the LHC. *PoS*, ICHEP2020:273. 6 p, 2021.

[41] High Luminosity LHC Project. The hl-lhc project, 2022.

[42] R Alemany-Fernandez, E Bravin, L Drosdal, A Gorzawski, V Kain, M Lamont, A Macpherson, G Papotti, M Pojer, L Ponce, S Redaelli, G Roy, M Solfaroli Camillocci, W Venturini, and J Wenninger. Operation and Configuration of the LHC in Run 1. Nov 2013.

[43] Jorg Wenninger. Operation and Configuration of the LHC in Run 2. Mar 2019.

[44] Fartoukh et al. LHC Configuration and Operational Scenario for Run 3. Technical report, CERN, Geneva, Nov 2021.

[45] High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1. 4/2017, 2017.

[46] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003. 437 p, 2008. Also published by CERN Geneva in 2010.

[47] ATLAS Experiment: Collaboration at the frontiers of science and technology. Jun 2018.

[48] Joao Pequenao and Paul Schaffner. How ATLAS detects particles: diagram of particle paths in the detector. Jan 2013.

[49] *ATLAS inner detector: Technical Design Report, Vol. 1.* Technical design report. ATLAS. CERN, Geneva, 1997.

[50] ATLAS inner detector: Technical Design Report. Vol. 2. 4 1997.

[51] The ATLAS Collaboration. Alignment of the ATLAS Inner Detector in Run-2. *Eur. Phys. J. C*, 80:1194. 41 p, Jul 2020.

[52] The ATLAS Collaboration. Study of the material of the atlas inner detector for run 2 of the lhc. *Journal of Instrumentation*, 12(12):P12009–P12009, Dec 2017.

[53] A Yamamoto, T Kondo, Y Doi, Y Makida, K Tanaka, T Haruyama, H Yamaoka, H H J ten Kate, L Bjorset, K Wada, S Meguro, J S H Ross, and K D Smith. Design and development of the ATLAS central solenoid magnet. *IEEE Trans. Appl. Supercond.*, 9:852–855. 6 p, Nov 1998.

[54] The ATLAS Pixels Collaboration. ATLAS pixel detector electronics and sensors. *JINST*, 3:P07007, 2008.

[55] Norbert Wermes and G Hallewel. *ATLAS pixel detector: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1998.

[56] M Capeans, G Darbo, K Einsweiller, M Elsing, T Flick, M Garcia-Sciveres, C Gemme, H Pernegger, O Rohne, and R Vuillermet. ATLAS Insertable B-Layer Technical Design Report. Technical report, Sep 2010.

[57] The ATLAS collaboration. Operation and performance of the atlas semiconductor tracker. *Journal of Instrumentation*, 9(08):P08009–P08009, Aug 2014.

[58] A. Ahmad et al. The Silicon microstrip sensors of the ATLAS semiconductor tracker. *Nucl. Instrum. Meth. A*, 578:98–118, 2007.

[59] The ATLAS TRT Collaboration. The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance. *JINST*, 3:P02013, 2008.

[60] Bartosz Mindur. ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider. Technical report, CERN, Geneva, Mar 2016.

[61] The ATLAS Collaboration. Performance of the atlas transition radiation tracker in run 1 of the lhc: tracker properties. *Journal of Instrumentation*, 12(05):P05002–P05002, May 2017.

[62] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1996.

[63] *ATLAS tile calorimeter: Technical Design Report*. Technical design report. ATLAS. CERN, Geneva, 1996.

[64] T. Barillari. The atlas liquid argon hadronic end-cap calorimeter: construction and selected beam test results. *Nuclear Physics B - Proceedings Supplements*, 150:102–105, Jan 2006.

[65] R S Orr. The ATLAS Forward Calorimeter. Sep 2011.

[66] *ATLAS muon spectrometer: Technical Design Report.* Technical design report. ATLAS. CERN, Geneva, 1997.

[67] J P Badiou, J Beltramelli, J M Baze, and J Belorgey. *ATLAS barrel toroid: Technical Design Report.* Technical design report. ATLAS. CERN, Geneva, 1997. Electronic version not available.

[68] ATLAS endcap toroids: Technical design report. 4 1997.

[69] Expected performance of the ATLAS detector at the High-Luminosity LHC. Technical report, CERN, Geneva, Jan 2019.

[70] Technical Design Report for the ATLAS Inner Tracker Pixel Detector. Technical report, CERN, Geneva, Sep 2017.

[71] Technical Design Report for the ATLAS Inner Tracker Strip Detector. Technical report, CERN, Geneva, Apr 2017.

[72] ATLAS Liquid Argon Calorimeter Phase-II Upgrade: Technical Design Report. Technical report, CERN, Geneva, Sep 2017.

[73] Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter. Technical report, CERN, Geneva, Sep 2017.

[74] Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer. Technical report, CERN, Geneva, Sep 2017.

[75] Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System. Technical report, CERN, Geneva, Sep 2017.

[76] Technical Proposal: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade. Technical report, CERN, Geneva, Jun 2018.

[77] Peter Jenni, Marzio Nessi, Markus Nordberg, and Kenway Smith. *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report.* Technical design report. ATLAS. CERN, Geneva, 2003.

[78] The ATLAS Collaboration. Performance of the atlas trigger system in 2010. *The European Physical Journal C*, 72(1), Jan 2012.

[79] The ATLAS Collaboration. Performance of the atlas trigger system in 2015. *The European Physical Journal C*, 77(5), May 2017.

[80] Georges Aad et al. Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System. 9 2013.

[81] The ATLAS Collaboration. Performance of the atlas muon triggers in run 2. *Journal of Instrumentation*, 15(09):P09015–P09015, Sep 2020.

[82] G. Avoni et al. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *JINST*, 13(07):P07017, 2018.

[83] A. Sidoti. Minimum Bias Trigger Scintillators in ATLAS Run II. *JINST*, 9(10):C10020, 2014.

[84] Peter Jenni, Marzio Nessi, and Markus Nordberg. Zero Degree Calorimeters for ATLAS. Technical report, CERN, Geneva, Jan 2007.

[85] C. Ay et al. Monte Carlo generators in ATLAS software. *J. Phys. Conf. Ser.*, 219:032001, 2010.

[86] D F Geesaman and P E Reimer. The sea of quarks and antiquarks in the nucleon. *Reports on Progress in Physics*, 82(4):046301, Mar 2019.

[87] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Luigi Del Debbio, Stefano Forte, Patrick Groth-Merrild, Alberto Guffanti, Nathan P. Hartland, Zahari Kassabov, José I. Latorre, Emanuele R. Nocera, Juan Rojo, Luca Rottoli, Emma Slade, and Maria Ubiali. Parton distributions from high-precision collider data. *The European Physical Journal C*, 77(10), Oct 2017.

[88] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977.

[89] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:438–450, 1972.

[90] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys. B*, 126:298–318, 1977.

[91] ATLAS Collaboration. Athena, April 2019.

[92] Stefan Höche. Introduction to parton-shower event generators. In *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, pages 235–295, 2015.

[93] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159, 2015.

[94] Stefano Catani, Leandro Cieri, Giancarlo Ferrera, Daniel de Florian, and Massimiliano Grazzini. Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO. *Phys. Rev. Lett.*, 103:082001, 2009.

[95] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[96] Geoffrey C. Fox and Stephen Wolfram. A Model for Parton Showers in QCD. *Nucl. Phys. B*, 168:285–295, 1980.

[97] X. Artru and G. Mennessier. String model and multiproduction. *Nucl. Phys. B*, 70:93–115, 1974.

[98] J.-C. Winter, F. Krauss, and G. Soff. A modified cluster-hadronisation model. *The European Physical Journal C*, 36(3):381–395, Aug 2004.

[99] Silvia Ferreres-Solé and Torbjörn Sjöstrand. The space–time structure of hadronization in the lund model. *The European Physical Journal C*, 78(11), Nov 2018.

[100] D. Amati and G. Veneziano. Preconfinement as a Property of Perturbative QCD. *Phys. Lett. B*, 83:87–92, 1979.

[101] M. Bähr et al. Herwig++ physics and manual. *Eur. Phys. J. C*, 58:639, 2008.

[102] Marilena Bandieramonte, Riccardo Maria Bianchi, Joseph Boudreau, Andrea Dell'Acqua, and Vakhtang Tsulaia. The GeoModel tool suite for detector description. Technical report, CERN, Geneva, Jun 2021.

[103] S. Agostinelli et al. GEANT4–a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250–303, 2003.

[104] The ATLAS Collaboration. Performance of the atlas track reconstruction algorithms in dense environments in lhc run 2. *The European Physical Journal C*, 77(10), Oct 2017.

[105] Azriel Rosenfeld and John L. Pfaltz. Sequential operations in digital picture processing. *J. ACM*, 13(4):471–494, oct 1966.

[106] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth. A*, 262:444–450, 1987.

[107] The ATLAS Collaboration. Reconstruction of primary vertices at the atlas experiment in run 1 proton–proton collisions at the lhc. *The European Physical Journal C*, 77(5), May 2017.

[108] G Piacquadio, K Prokofiev, and A Wildauer. Primary vertex reconstruction in the ATLAS experiment at LHC. *Journal of Physics: Conference Series*, 119(3):032033, jul 2008.

[109] Tim Robertson and J. D. Cryer. An iterative procedure for estimating the mode. *Journal of the American Statistical Association*, 69(348):1012–1016, 1974.

[110] Wolfgang Waltenberger, Rudolf Frühwirth, and Pascal Vanlaer. Adaptive vertex fitting. *Journal of Physics G: Nuclear and Particle Physics*, 34(12):N343–N356, nov 2007.

[111] The ATLAS Collaboration. Topological cell clustering in the atlas calorimeters and its performance in lhc run 1. *The European Physical Journal C*, 77(7), Jul 2017.

[112] Peter Speckmayer. Energy Measurement of Hadrons with the CERN ATLAS Calorimeter, 2008. Presented on 18 Jun 2008.

[113] Gavin P. Salam. A practical seedless infrared safe cone algorithm, 2007.

[114] Yu.L Dokshitzer, G.D Leder, S Moretti, and B.R Webber. Better jet clustering algorithms. *Journal of High Energy Physics*, 1997(08):001–001, Aug 1997.

[115] Gavin P Salam and Grégory Soyez. A practical seedless infrared-safe cone jet algorithm. *Journal of High Energy Physics*, 2007(05):086–086, May 2007.

[116] Stefan Weinzierl. The siscone jet algorithm optimised for low particle multiplicities. *Computer Physics Communications*, 183(3):813–820, Mar 2012.

[117] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, Apr 2008.

[118] The ATLAS Collaboration. Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ tev with the atlas detector. *Phys. Rev. D*, 96:072002, Oct 2017.

[119] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Physics Letters B*, 659(1–2):119–126, Jan 2008.

[120] The 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*, 81:689. 73 p, Jul 2020.

[121] The ATLAS Collaboration. Performance of jet substructure techniques for large-r jets in proton-proton collisions at $\sqrt{s} = 7$ tev using the atlas detector. *Journal of High Energy Physics*, 2013(9), Sep 2013.

[122] David Krohn, Jesse Thaler, and Lian-Tao Wang. Jet trimming. *Journal of High Energy Physics*, 2010(2), Feb 2010.

[123] G. Aad, B. Abbott, D. C. Abbott, A. Abed Abud, K. Abeling, D. K. Abhayasinghe, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, and et al. Atlas b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ tev. *The European Physical Journal C*, 79(11), Nov 2019.

[124] Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run. Technical report, CERN, Geneva, Jun 2016.

[125] Optimization and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run. Technical report, CERN, Geneva, Jul 2017.

[126] Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data. Technical report, CERN, Geneva, Jun 2016.

[127] W Lampl, S Laplace, D Lelas, P Loch, H Ma, S Menke, S Rajagopalan, D Rousseau, S Snyder, and G Unal. Calorimeter Clustering Algorithms: Description and Performance. Technical report, CERN, Geneva, Apr 2008.

[128] Thijs G. Cornelissen, M. Elsing, I. Gavrilenko, J. F. Laporte, W. Liebig, M. Limper, K. Nikolopoulos, A. Poppleton, and A. Salzburger. The global chi**2 track fitter in ATLAS. *J. Phys. Conf. Ser.*, 119:032013, 2008.

[129] Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data. Technical report, CERN, Geneva, Jun 2014.

[130] The ATLAS Collaboration. Muon reconstruction performance of the atlas detector in proton–proton collision data at $\sqrt{s}$=13 tev. *The European Physical Journal C*, 76(5), May 2016.

[131] The ATLAS Collaboration. Performance of missing transverse momentum reconstruction with the atlas detector using proton–proton collisions at $\sqrt{s} = 13=$ 13 tev. *The European Physical Journal C*, 78(11), Nov 2018.

[132] Richard D. Ball et al. Parton distributions with LHC data. *Nucl. Phys. B*, 867:244, 2013.

[133] Leif Lönnblad. Correcting the colour-dipole cascade model with fixed order matrix elements. *Journal of High Energy Physics*, 2002(05):046–046, May 2002.

[134] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.

[135] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.

[136] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.

[137] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, H. Abramowicz, and et al. Search for pair production of up-type vector-like quarks and for four-top-quark events in final states with multiple b-jets with the atlas detector. *Journal of High Energy Physics*, 2018(7), Jul 2018.

[138] Wim Beenakker, Christoph Borschensky, Michael Krämer, Anna Kulesza, and Eric Laenen. NNLL-fast: predictions for coloured supersymmetric particle production at the LHC with threshold and Coulomb resummation. *JHEP*, 12:133, 2016.

[139] Wim Beenakker, Christoph Borschensky, Michael Krämer, Anna Kulesza, Eric Laenen, Vincent Theeuwes, and Silja Thewes. NNLL resummation for squark and gluino production at the LHC. *JHEP*, 12:023, 2014.

[140] Wim Beenakker, Tim Janssen, Susanne Lepoeter, Michael Krämer, Anna Kulesza, Eric Laenen, Irene Niessen, Silja Thewes, and Tom Van Daal. Towards NNLL resummation: hard matching coefficients for squark and gluino hadroproduction. *JHEP*, 10:120, 2013.

[141] Wim Beenakker, Silja Brensing, Michael Krämer, Anna Kulesza, Eric Laenen, and Irene Niessen. NNLL resummation for squark-antisquark pair production at the LHC. *JHEP*, 01:076, 2012.

[142] Wim Beenakker, Silja Brensing, Michael Krämer, Anna Kulesza, Eric Laenen, and Irene Niessen. Soft-gluon resummation for squark and gluino hadroproduction. *JHEP*, 12:041, 2009.

[143] A. Kulesza and L. Motyka. Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC. *Phys. Rev. D*, 80:095004, 2009.

[144] A. Kulesza and L. Motyka. Threshold resummation for squark-antisquark and gluino-pair production at the LHC. *Phys. Rev. Lett.*, 102:111802, 2009.

[145] W. Beenakker, R. Hopker, M. Spira, and P.M. Zerwas. Squark and gluino production at hadron colliders. *Nucl. Phys. B*, 492:51, 1997.

[146] Jon Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys. G*, 43:023001, 2016.

[147] 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.

[148] Nikolaos Kidonakis. NNLL resummation for s-channel single top quark production. *Phys. Rev. D*, 81:054028, 2010.

[149] S Heinemeyer et all. *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group.* CERN Yellow Reports: Monographs. Jul 2013. Comments: 404 pages, 139 figures, to be submitted to CERN Report. Working Group web page: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections.

[150] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, et al. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009.

[151] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert. QCD matrix elements + parton showers: The NLO case. *JHEP*, 04:027, 2013.

[152] Stefano Catani, Leandro Cieri, Giancarlo Ferrera, Daniel de Florian, and Massimiliano Grazzini. Vector boson production at hadron colliders: A fully exclusive qcd calculation at next-to-next-to-leading order. *Phys. Rev. Lett.*, 103:082001, Aug 2009.

[153] M. Aaboud et al. Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev. D*, 96(7):072002, 2017.

[154] G. Aad, B. Abbott, D.C. Abbott, A. Abed Abud, K. Abeling, D.K. Abhayasinghe, S.H. Abidi, O.S. AbouZeid, N.L. Abraham, H. Abramowicz, and et al. Electron and photon performance measurements with the atlas detector using the 2015–2017 lhc proton-proton collision data. *Journal of Instrumentation*, 14(12):P12006–P12006, Dec 2019.

[155] Matthew Tamsett. Measurement of multi-jet cross sections at atlas, 2011.

[156] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, H. Abramowicz, and et al. Search for pair production of up-type vector-like quarks and for four-top-quark events in final states with multiple b-jets with the atlas detector. *Journal of High Energy Physics*, 2018(7), Jul 2018.

[157] M. Baak, G. J. Besjes, D. Côté, A. Koutsman, J. Lorenz, and D. Short. Histfitter software framework for statistical data analysis. *The European Physical Journal C*, 75(4), Apr 2015.

[158] Chiara Rizzi, Maximilian J Swiatlowski, Jean-Francois Arguin, Davide Costanzo, Otilia Anamaria Ducu, Trisha Farooque, Louis-Guillaume Gagnon, Aurelio Juste Rozas, Alexander Khanov, Evangelos Kourlitis, Matthew Edgar Leblanc, Calum Michael Macdonald, David Miller, Yuji Minegishi, Tuan Nguyen Manh, Flera Rizatdinova, Giordon Holtsberg Stark, Junichi Tanaka, Loic Valery, Emma Sian Kuwertz, Meisam Ghasemi Bostanabad, Josu Cantero, Carlos Moreno Martinez, Matthew Berg Epland, Ayana Arce, and Justin Albert. Search for gluino-mediated stop and sbottom pair production in events with *b*-jets and large missing transverse momentum. Technical report, CERN, Geneva, Mar 2018.

[159] The ATLAS Collaboration. Jet energy measurement and its systematic uncertainty in proton–proton collisions at $\sqrt{s} = 7$ tev with the atlas detector. *The European Physical Journal C*, 75(1), Jan 2015.

[160] Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, Jul 2015.

[161] G. Avoni, M. Bruschi, G. Cabras, D. Caforio, N. Dehghanian, A. Floderus, B. Giacobbe, F. Giannuzzi, F. Giorgi, P. Grafström, V. Hedberg, F. Lasagni Manghi, S. Meneghini, J. Pinfold, E. Richards, C. Sbarra, N. Semprini Cesari, A. Sbrizzi, R. Soluk, G. Ucchielli, S. Valentinetti, O. Viazlo, M. Villa, C. Vittori, R. Vuillermet, and A. Zoccoli. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *Journal of Instrumentation*, 13(07):P07017–P07017, jul 2018.

[162] Comparison of Monte Carlo generator predictions to ATLAS measurements of top pair production at 7 TeV. Technical report, CERN, Geneva, Jan 2015.

[163] Enrico Bothmann, Gurpreet Singh Chahal, Stefan Höche, Johannes Krause, Frank Krauss, Silvan Kuttimalai, Sebastian Liebschner, Davide Napoletano, Marek Schönherr, Holger Schulz, Steffen Schumann, and Frank Siegert. Event generation with sherpa 2.2. *SciPost Physics*, 7(3), Sep 2019.

[164] Michiel Botje et al. The PDF4LHC Working Group Interim Recommendations. 1 2011.

[165] A L Read. Presentation of search results: theCLstechnique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693–2704, sep 2002.

[166] J. Neyman and E. S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231:289–337, 1933.

[167] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), Feb 2011.

[168] CMS Collaboration. Search for supersymmetry in pp collisions at $\sqrt{s} = 13$ tev with 137 fb-1 in final states with a single lepton using the sum of masses of large-radius jets. *Physical Review D*, 101(5), Mar 2020.

[169] Junpei Maeda. BSM searches with top final states. Jun 2021.

[170] The ATLAS Collaboration. Search for the standard model higgs boson produced in association with top quarks and decaying into a bb pair in pp collisions at $\sqrt{s} = 13$ tev with the atlas detector. *Physical Review D*, 97(7), Apr 2018.

[171] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.

[172] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction. *JHEP*, 09:126, 2007.

[173] Richard D. Ball et al. Parton distributions for the LHC run II. *JHEP*, 04:040, 2015.

[174] M. Beneke, P. Falgari, S. Klein, and C. Schwinn. Hadronic top-quark pair production with NNLL threshold resummation. *Nucl. Phys. B*, 855:695–741, 2012.

[175] Matteo Cacciari, Michal Czakon, Michelangelo Mangano, Alexander Mitov, and Paolo Nason. Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation. *Phys. Lett. B*, 710:612–622, 2012.

[176] Michal Czakon and Alexander Mitov. NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels. *JHEP*, 12:054, 2012.

[177] Michal Czakon, Paul Fiedler, and Alexander Mitov. Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$. *Phys. Rev. Lett.*, 110:252004, 2013.

[178] Peter Bärnreuther, Michal Czakon, and Alexander Mitov. Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t} + X$. *Phys. Rev. Lett.*, 109:132001, 2012.

[179] LHCTopWG. NNLO+NNLL top-quark-pair cross sections, 2015.

[180] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J. C*, 63:189, 2009.

[181] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections. *Eur. Phys. J. C*, 64:653–680, 2009.

[182] H.-L. Lai et al. New parton distributions for collider physics. *Phys. Rev. D*, 82:074024, 2010.

[183] J. Gao et al. CT10 next-to-next-to-leading order global analysis of QCD. *Phys. Rev. D*, 89:033009, 2014.

[184] Johannes Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J. C*, 76(4):196, 2016.

[185] Johannes Bellm et al. Herwig 7.1 Release Note. 2017.

[186] The ATLAS Collaboration. Jet reconstruction and performance using particle flow with the atlas detector. *The European Physical Journal C*, 77(7), Jul 2017.

[187] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, M. Backes, T. Carli, O. Cohen, A. Christov, D. Dannheim, K. Danielowski, S. Henrot-Versille, M. Jachowski, K. Kraszewski, A. Krasznahorkay Jr. au2, M. Kruk, Y. Mahalalel, R. Ospanov, X. Prudent, A. Robert, D. Schouten, F. Tegenfeldt, A. Voigt, K. Voss, M. Wolter, and A. Zemla. Tmva - toolkit for multivariate data analysis, 2009.

[188] Giulio D'Agostini. A multidimensional unfolding method based on bayes' theorem. *Nucl. Instrum. Meth. A*, 362(2):487 – 498, 1995.

[189] Tim Adye. Unfolding algorithms and tests using RooUnfold. In *Proceedings, 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding (PHYSTAT 2011)*, pages 313–318.

[190] RD53 Collaboration. https://rd53.web.cern.ch.

[191] Maurice Garcia-Sciveres. The RD53A Integrated Circuit. Technical report, CERN, Geneva, Oct 2017.

[192] Timon Heim. YARR - a PCIe based readout concept for current and future ATLAS pixel modules. *Journal of Physics: Conference Series*, 898:032053, oct 2017.

[193] R. Herbst, R. Claus, M. Freytag, G. Haller, M. Huffer, S. Maldonado, K. Nishimura, C. O'Grady, J. Panetta, A. Perazzo, B. Reese, L. Ruckman, J. G. Thayer, and M. Weaver. Design of the slac rce platform: A general purpose atca based data acquisition system. In *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pages 1–4, 2014.

[194] FELIX user manual. https://rd53.web.cern.ch.

[195] YARR software. https://gitlab.cern.ch/YARR/YARR.

APPENDICES

0.1 Preselection level plots

Plots of all variables used in the analysis on the 0L and 1L preselection level are given in the appendix. Figure 91 shows the 0L channel. Figure 92 and 93 shows 1L channel plots before (left column) and after (right column) kinematic reweighing.
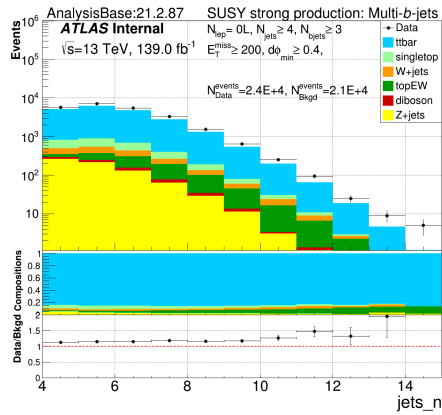
(a) Jets multiplicity
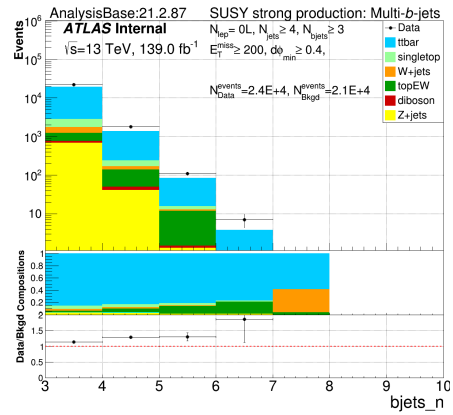
(b) b-jets multiplicity
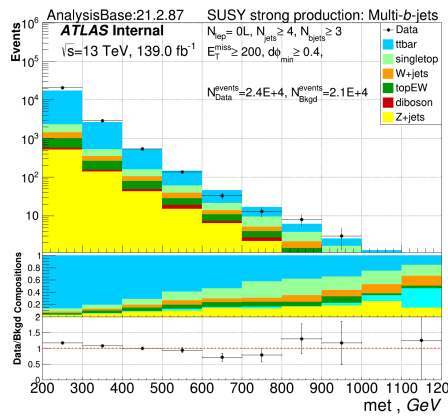
(c) E_T^{miss}
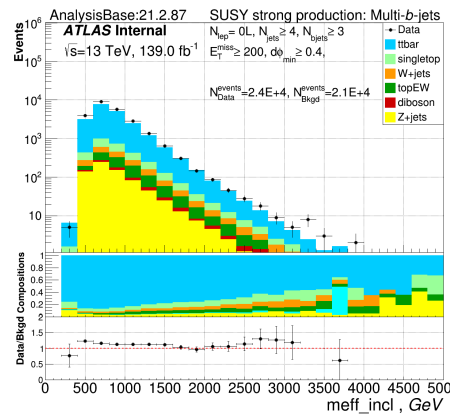
(d) m_{eff}
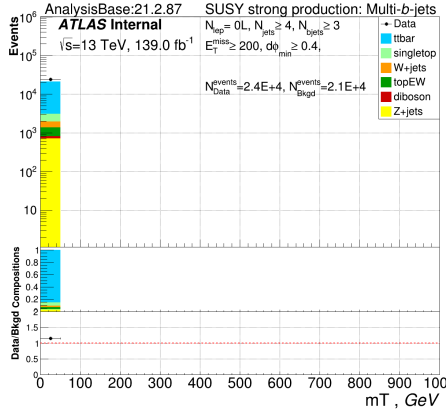
(e) m_{T}
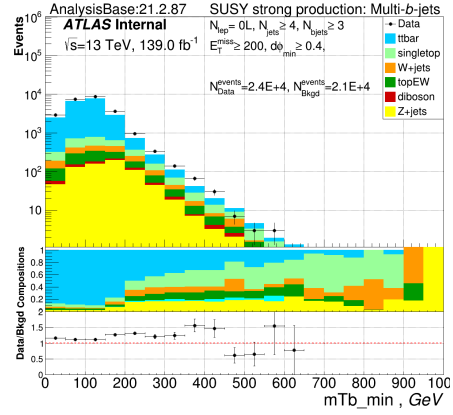
(f) $m_{\text{T,min}}^{b-jets}$
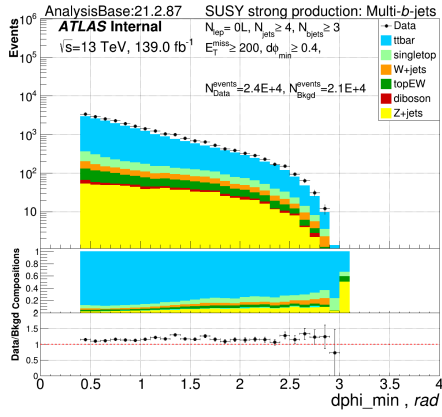
(g) $\Delta\phi_{min}^{4j}$
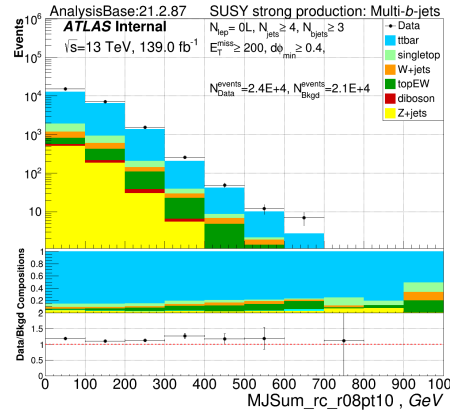
(h) $M_J^{\Sigma,4}$
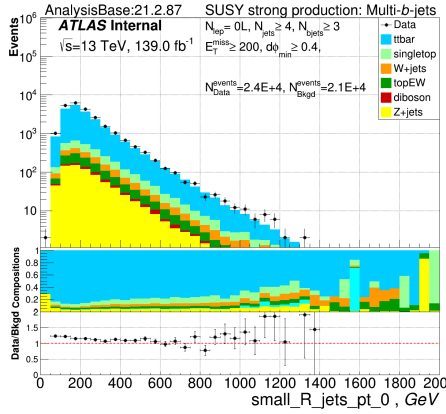
(i) leading small-R jet p_{T}
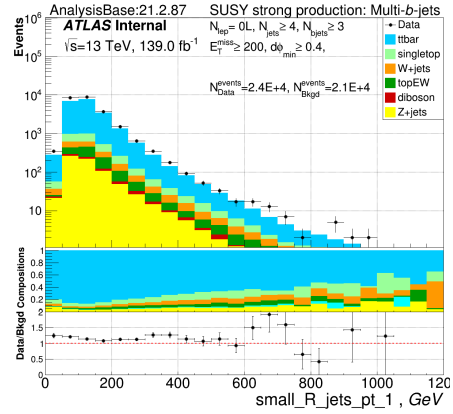
(j) sub-leading small-R jet p_{T}
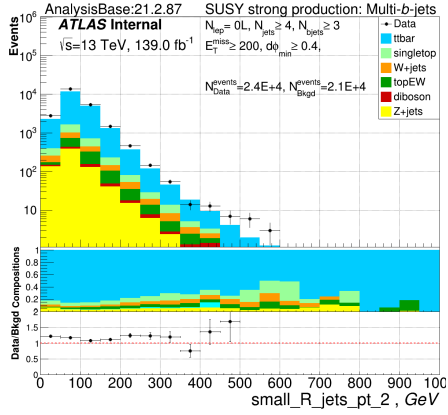
(k) 3rd small-R jet p_T
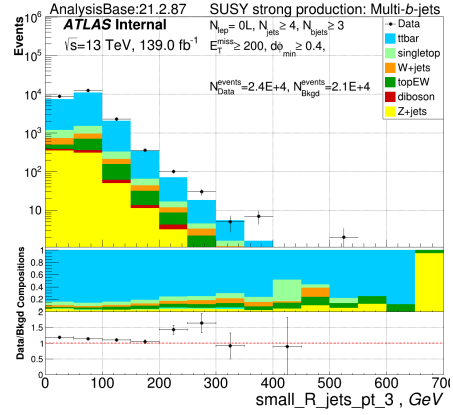
(l) 4th small-R jet p_T
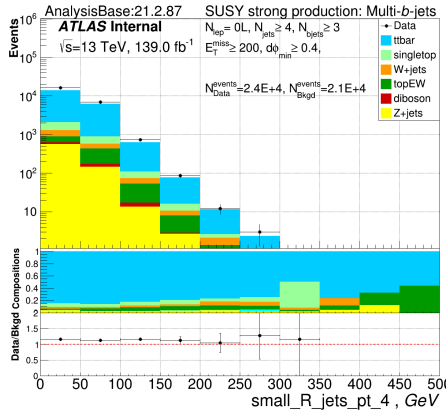
(m) 5th small-R jet p_T
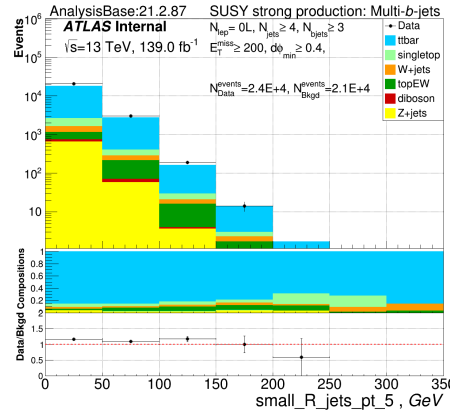
(n) 6th small-R jet p_T
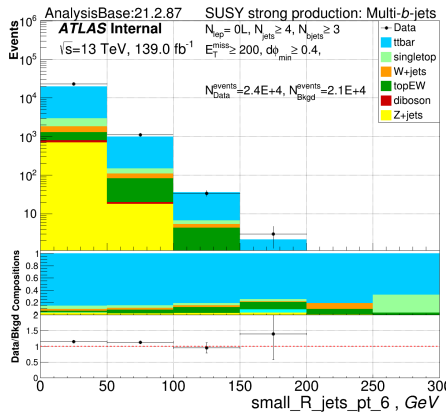
(o) 7th small-R jet p_T

(p) 8th small-R jet p_T

157
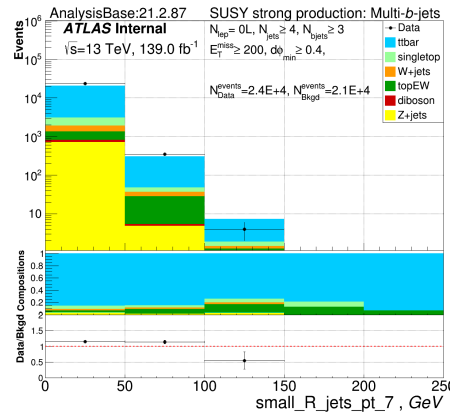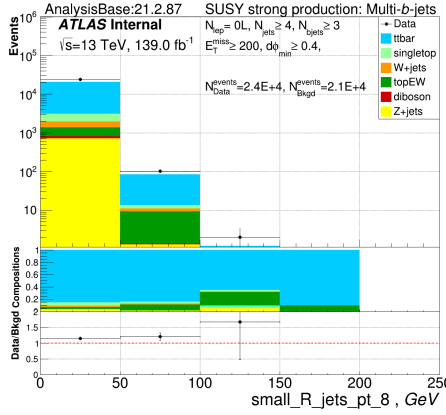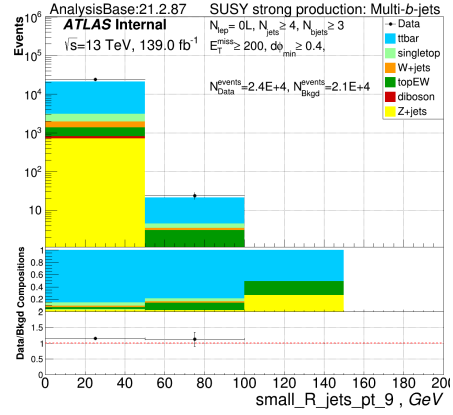
(q) 9th small-R jet p_{T}

(r) 10th small-R jet p_{T}

Figure 91: Kinematic distributions in events with no leptons on the preselection level. Top canvas: Data (black dots) and background (in color) yields. Middle canvas: background compositions. Bottom canvas: data to background ratios.
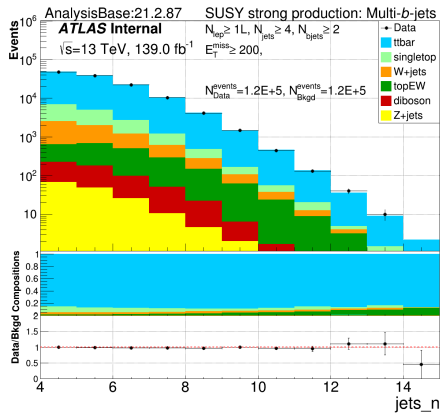
(a) jets multiplicity before
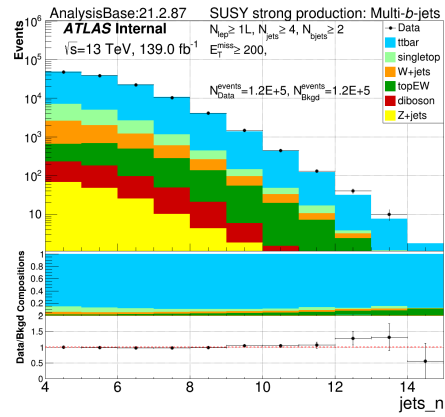
(b) jets multiplicity after
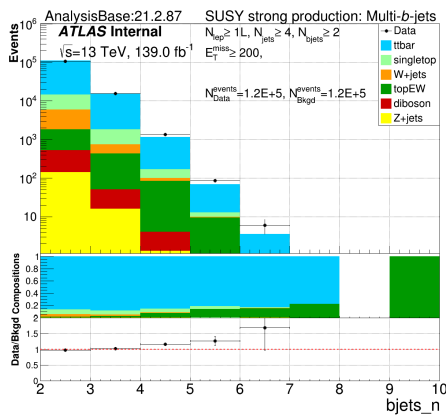
(c) b-jets multiplicity before
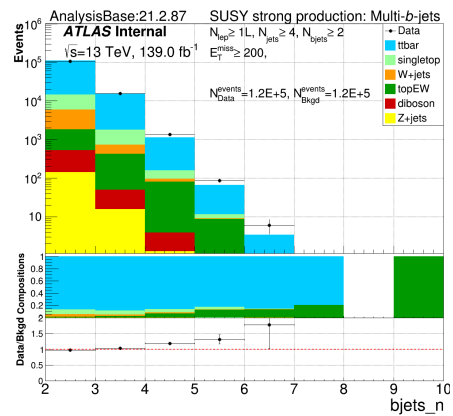
(d) b-jets multiplicity after
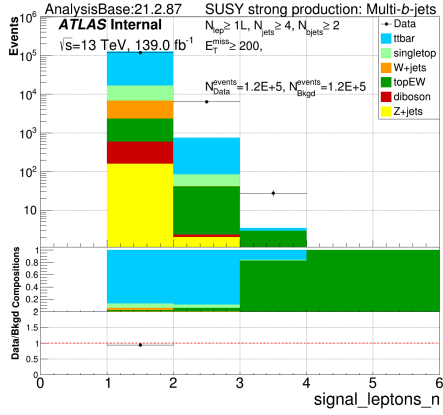
(e) leptons multiplicity before
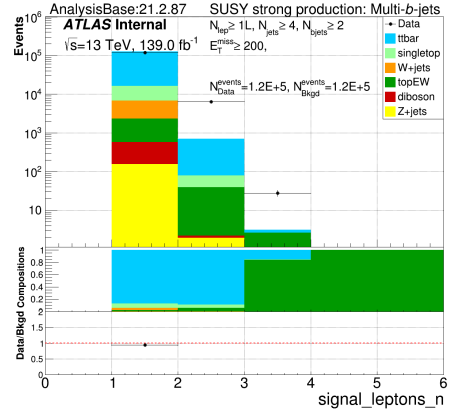
(f) leptons multiplicity after
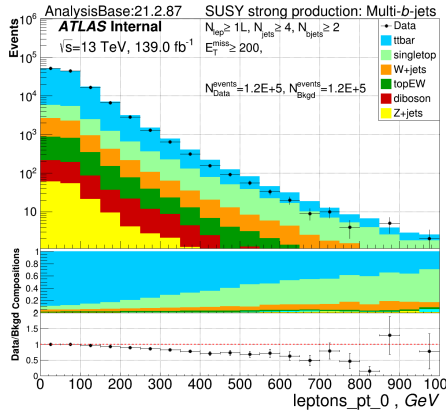
(g) leading lepton p_{T} before
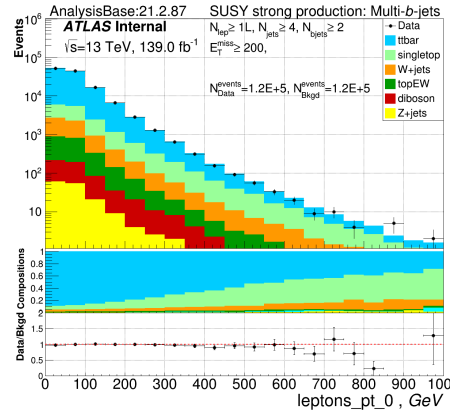
(h) leading lepton p_{T} after
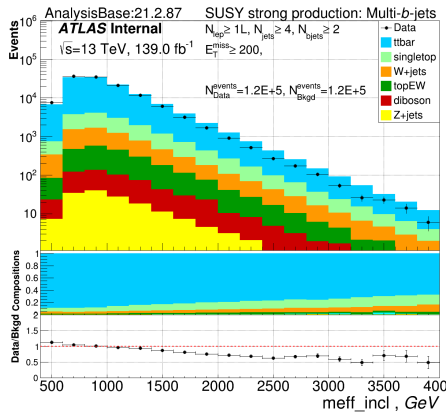
(i) m_{eff} before
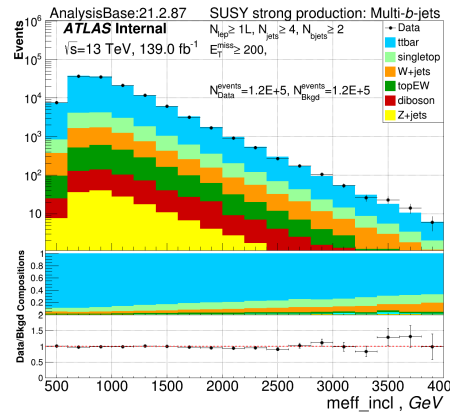
(j) m_{eff} after
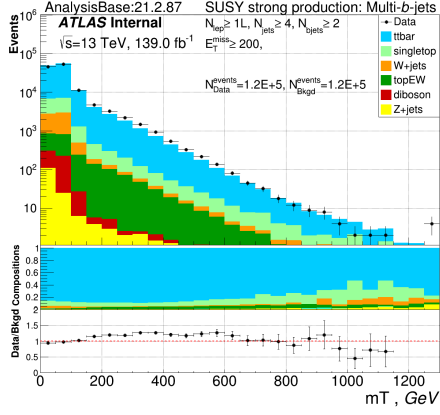
(k) m_{T} before
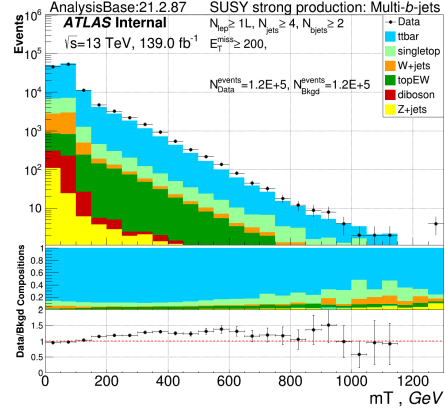
(l) m_{T} after
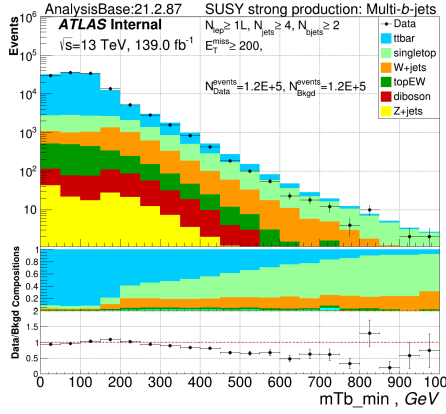
(m) $m_{\mathrm{T,min}}^{b-jets}$ before
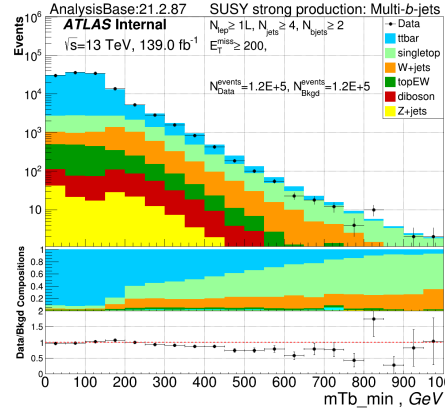
(n) $m_{\mathrm{T,min}}^{b-jets}$ after
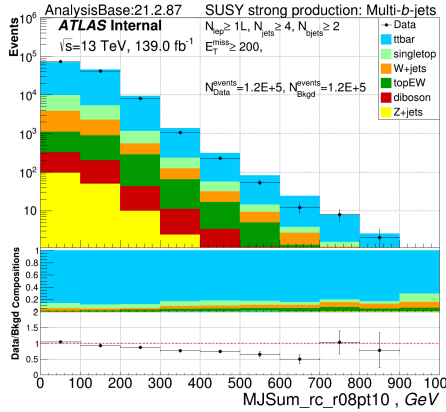
(o) $M_J^{\Sigma,4}$ before
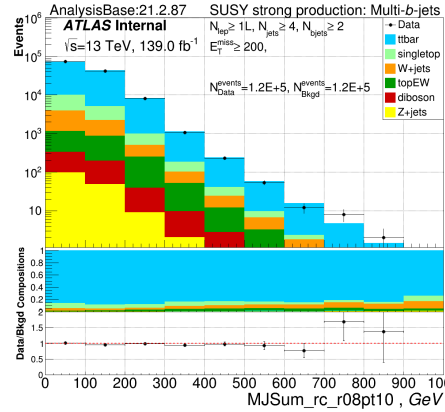
(p) $M_J^{\Sigma,4}$ after

Figure 92: Kinematic distributions in events with at least one leptons on the preselection level. Left figures/susy: before applying kinematic reweighing; right figures/susy: after. Top canvas: Data (black dots) and background (in color) yields. Middle canvas: background compositions. Bottom canvas: data to background ratios.
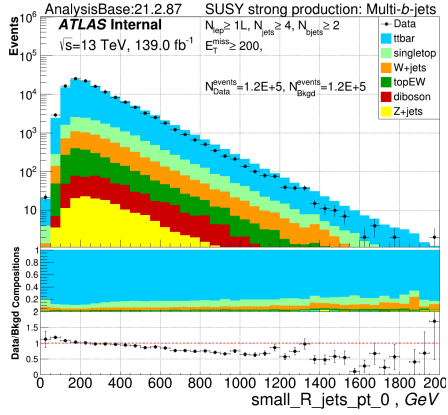
(a) leading small-R jet p_T before
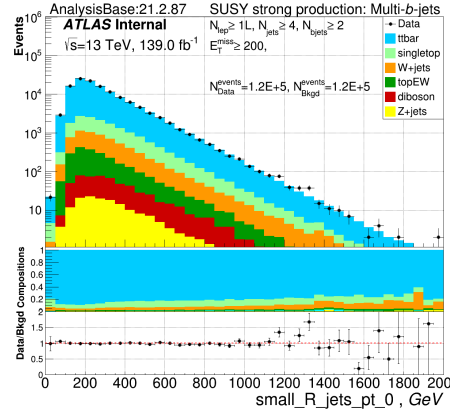
(b) leading small-R jet p_T after
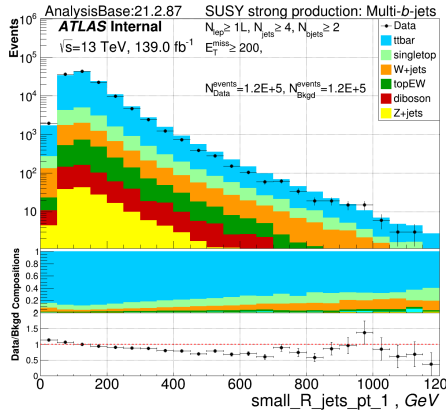
(c) sub-leading small-R jet p_T before
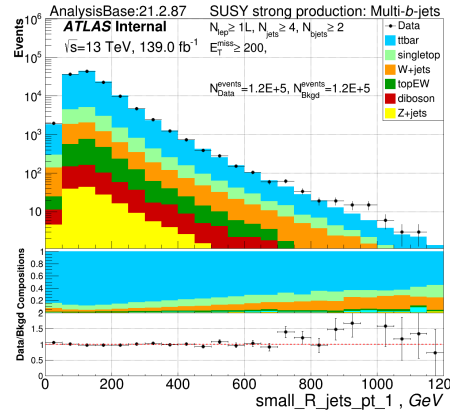
(d) sub-leading small-R jet p_T after
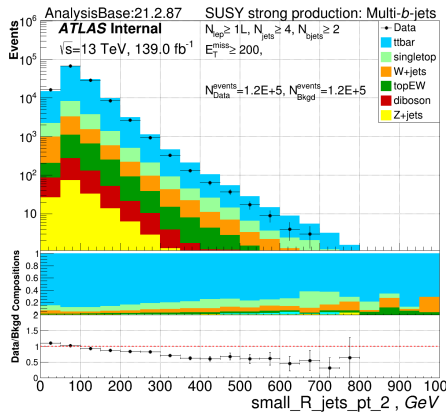
(e) 3rd small-R jet p_T before
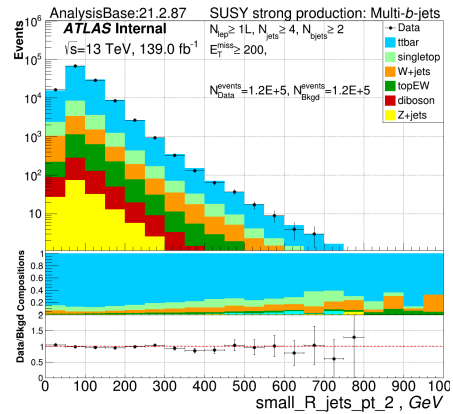
(f) 3rd small-R jet p_T after
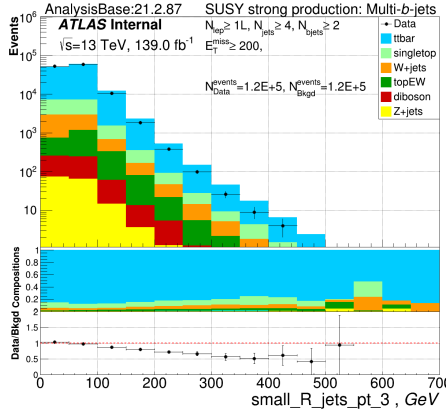
(g) 4th small-R jet p_T before
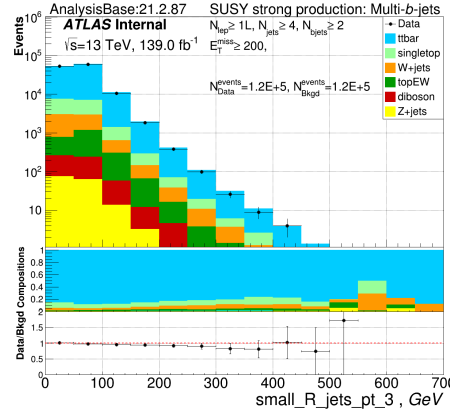
(h) 4th small-R jet p_T after
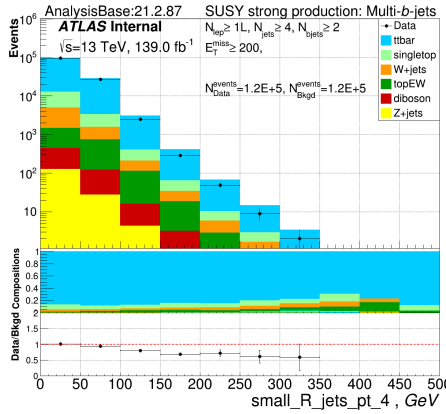
(i) 5th small-R jet p_T before
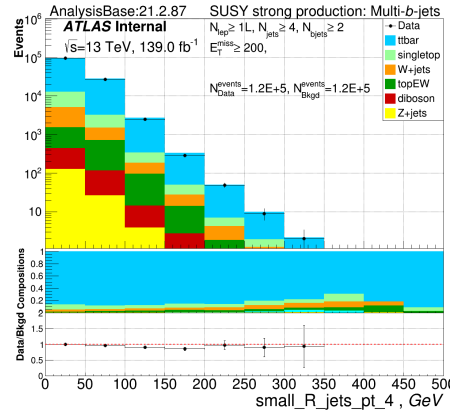
(j) 5th small-R jet p_T after
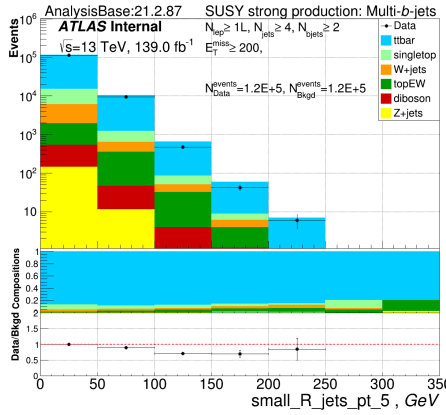
(k) 6th small-R jet p_T before
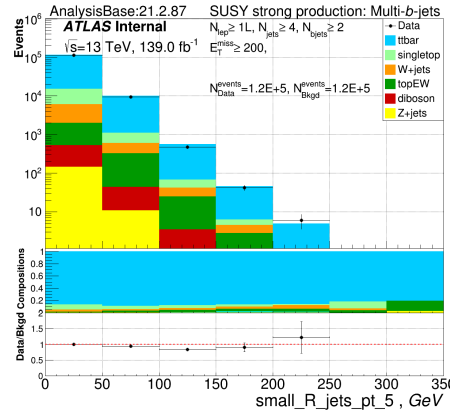
(l) 6th small-R jet p_T after
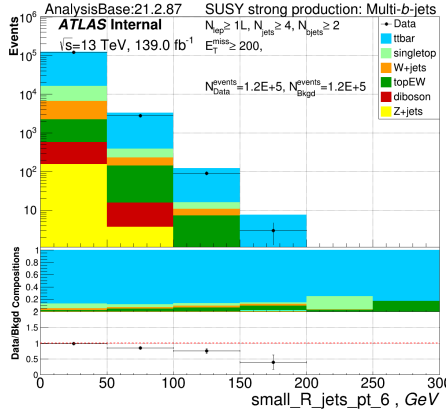
(m) 7th small-R jet p_{T} before
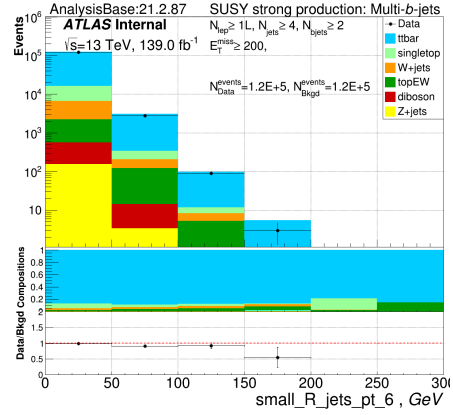
(n) 7th small-R jet p_{T} after
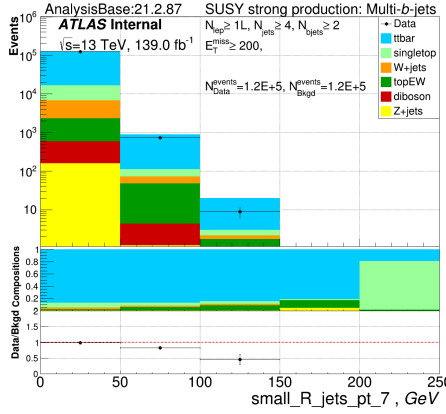
(o) 8th small-R jet p_{T} before
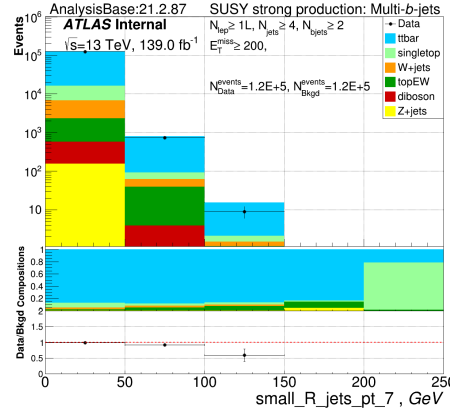
(p) 8th small-R jet p_{T} after
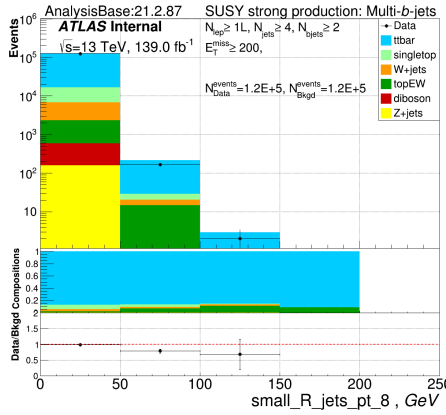
(q) 9th small-R jet p_{T} before
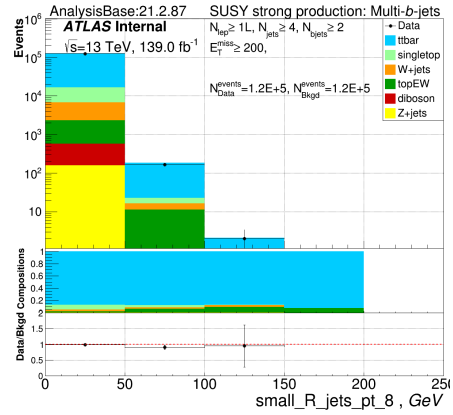
(r) 9th small-R jet p_{T} after

163
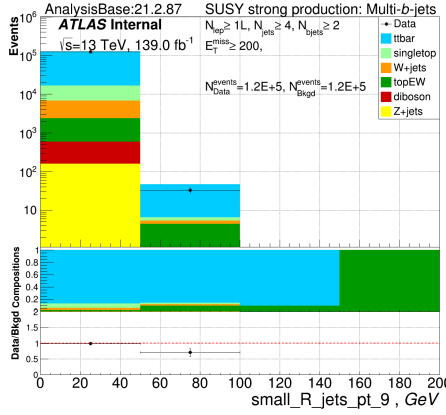
(s) 10th small-R jet p_T before
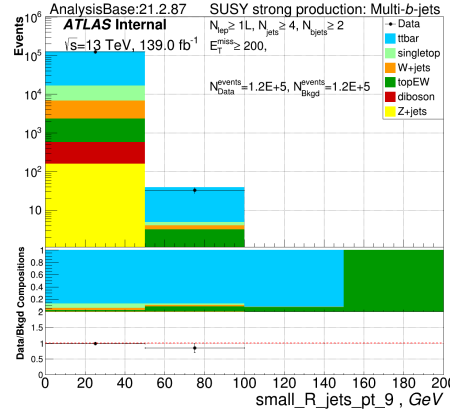
(t) 10th small-R jet p_T after

Figure 93: Jets p_T distributions in events with at least one leptons on the preselection level. Left figures/susy: before applying kinematic reweighing; right figures/susy: after. Top canvas: Data (black dots) and background (in color) yields. Middle canvas: background compositions. Bottom canvas: data to background ratios.

0.2 Kinematic Reweighing Functions

All reweighing functions for events with at least one signal lepton are summarized in this appendix. Motivation and procedure of derivation are given in section 6.2.3. The aim of the procedure was to correct shapes of analysis variables, but not the overall normalization. Therefore, when the functions are applied to a particular MC sample, they are scaled such that the overall yield in the sample doesn't change.
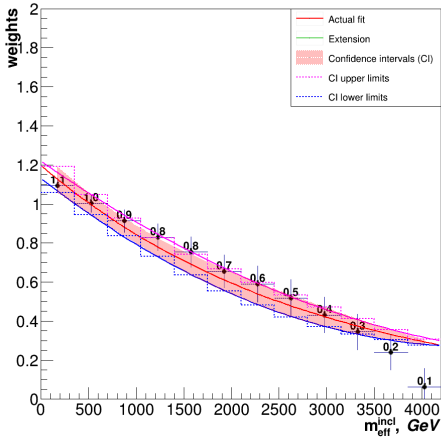
(a) $t\bar{t}$ with $N_{jets} = 4$
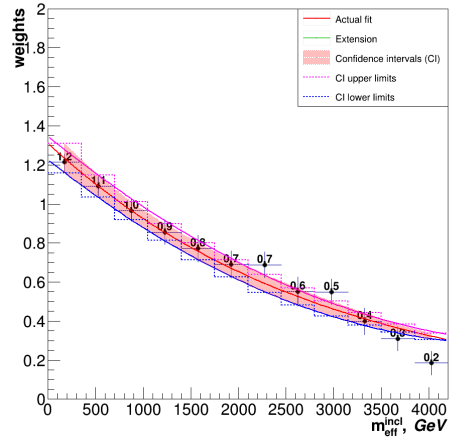
(b) $t\bar{t}$ with $N_{jets} = 5$
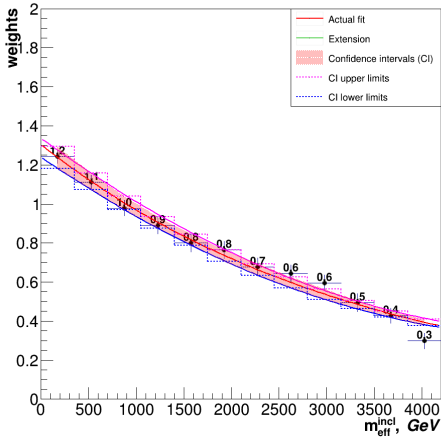
(c) $t\bar{t}$ with $N_{jets} = 6$
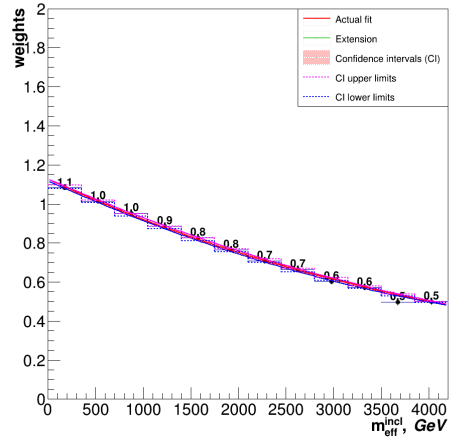
(d) $t\bar{t}$ with $N_{jets} \geq 7$
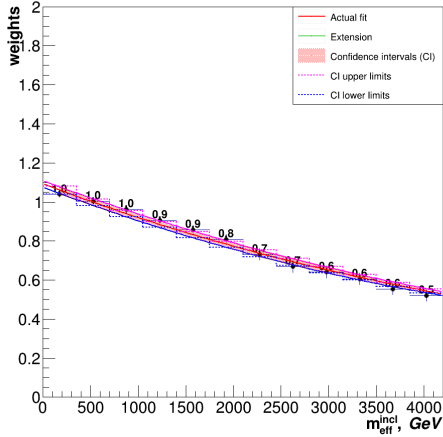
(e) W+jets with $N_{jets} = 4$
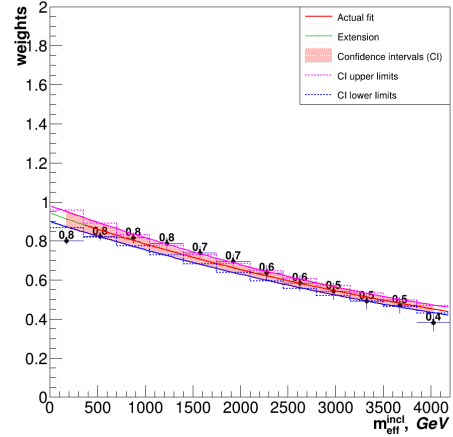
(f) W+jets with $N_{jets} = 5$
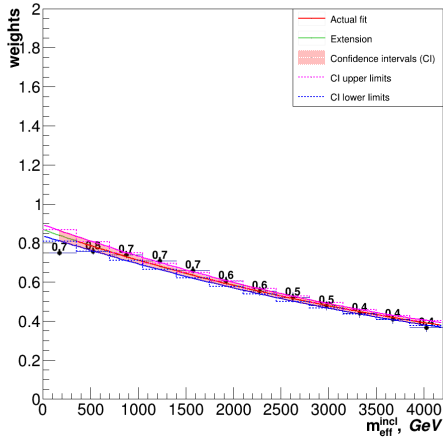
(g) W+jets with $N_{jets} = 6$
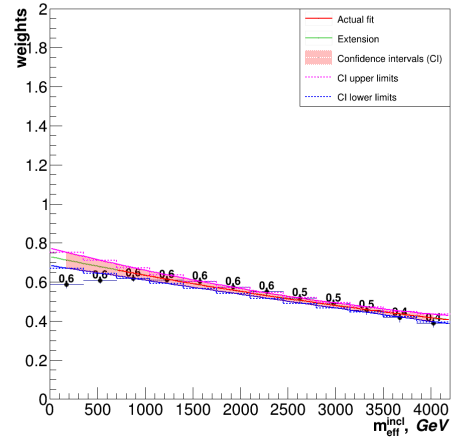
(h) W+jets with $N_{jets} > 7$
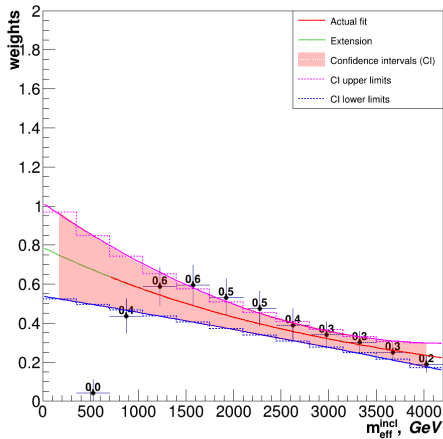
(i) single top with $N_{jets} \geq 4$
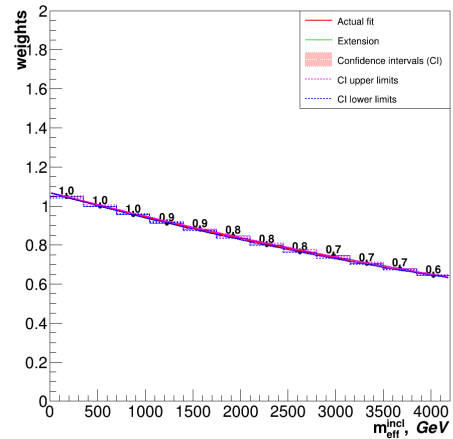
(j) Z+jets with $N_{jets} \geq 4$

Figure 94: Reweighing functions for four main kinds of events with at least one signal lepton.

0.3 Signal Regions Optimizations Grids in 0L and 1L Channels

The appendix summarizes plots with selections in each variable from the sets of the most optimal cuts for each signal model in the $\{m_{\tilde{g}};\ m_{\tilde{\chi}_1^0}\}$ phase-space. Figures 95 and 96 show the selections in the 0L and 1L channels respectively.
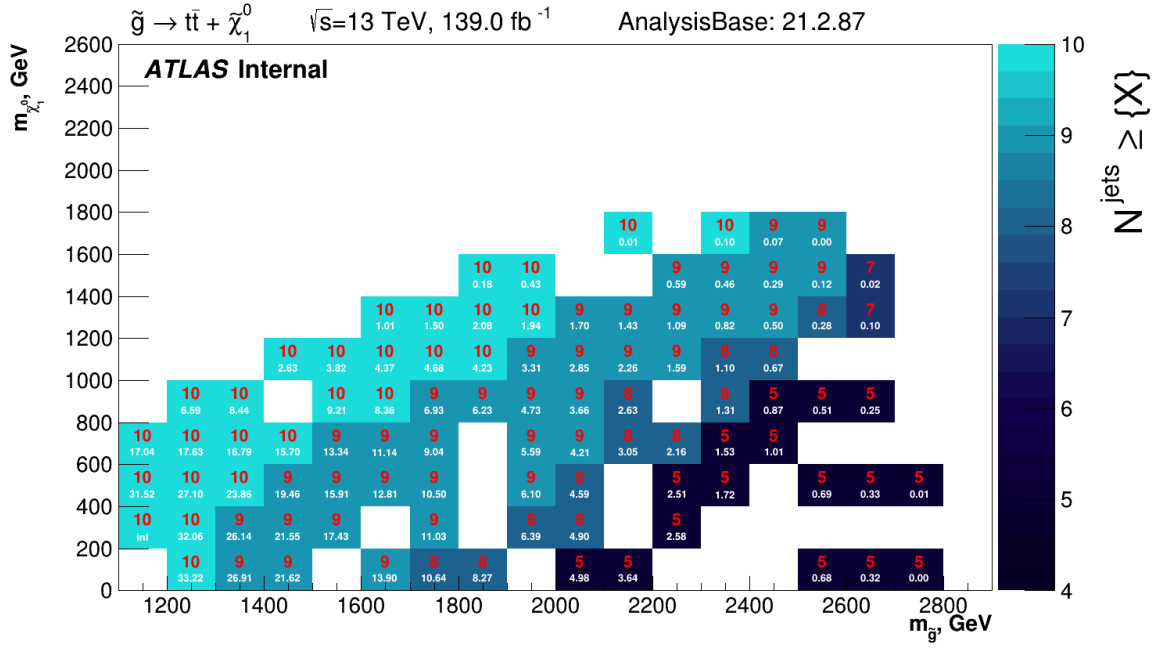
(a) Jets multiplicity
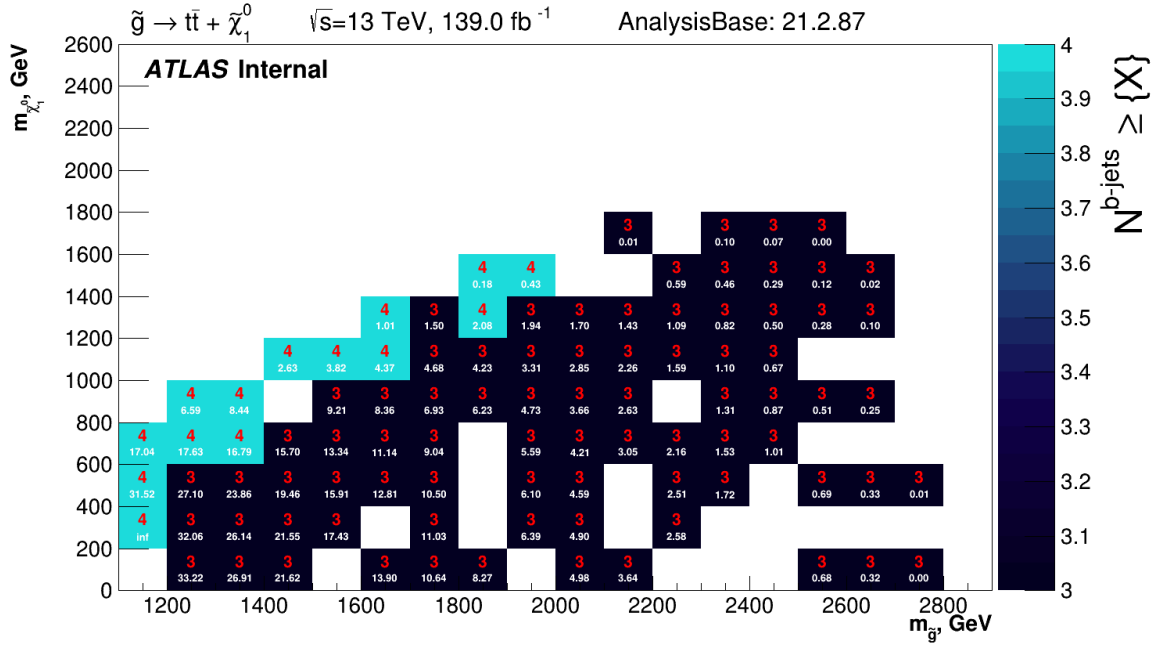
(b) b-jets multiplicity
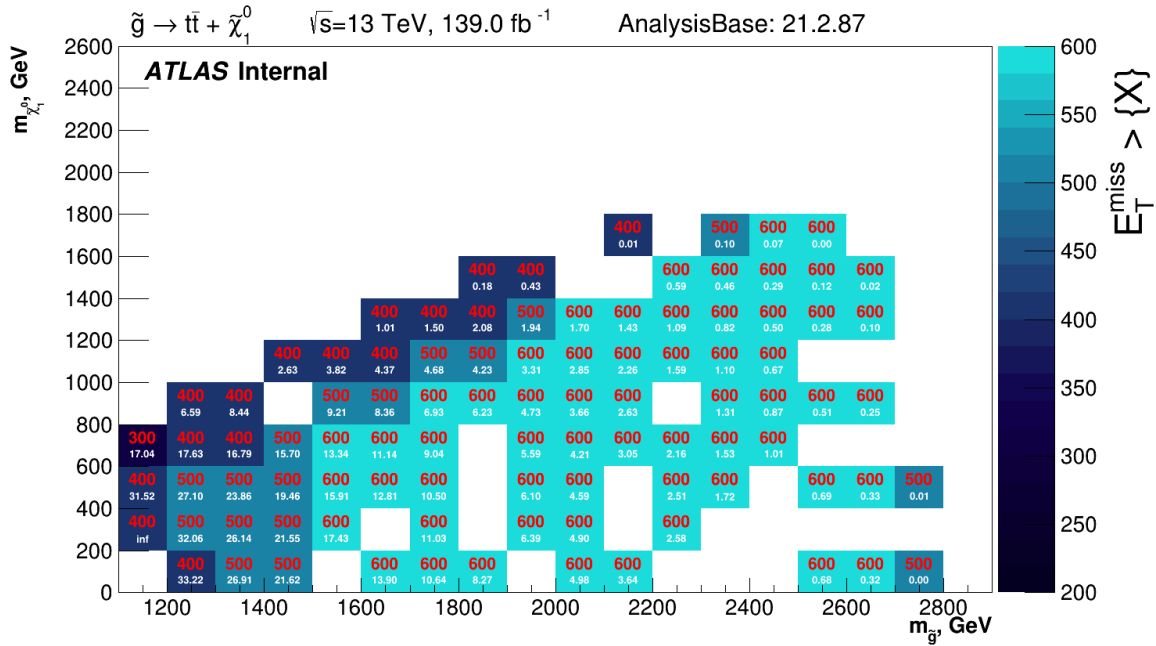
(c) E_T^{miss}
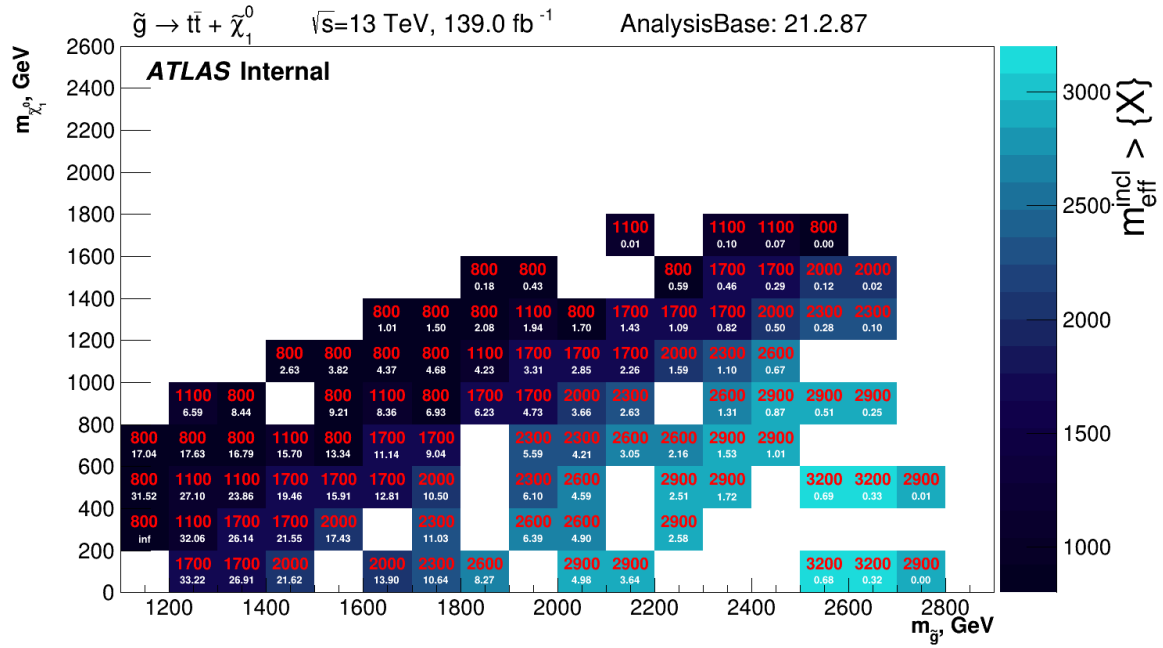
(d) m_{eff}

(e) $m_{\mathrm{T,min}}^{b-jets}$

169
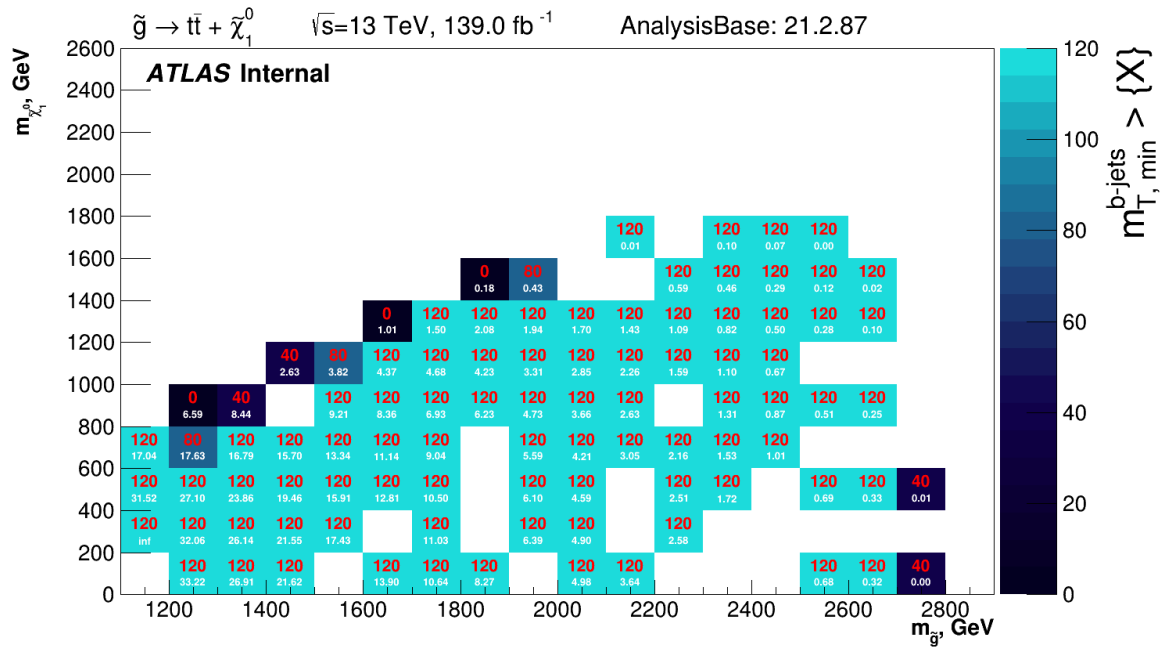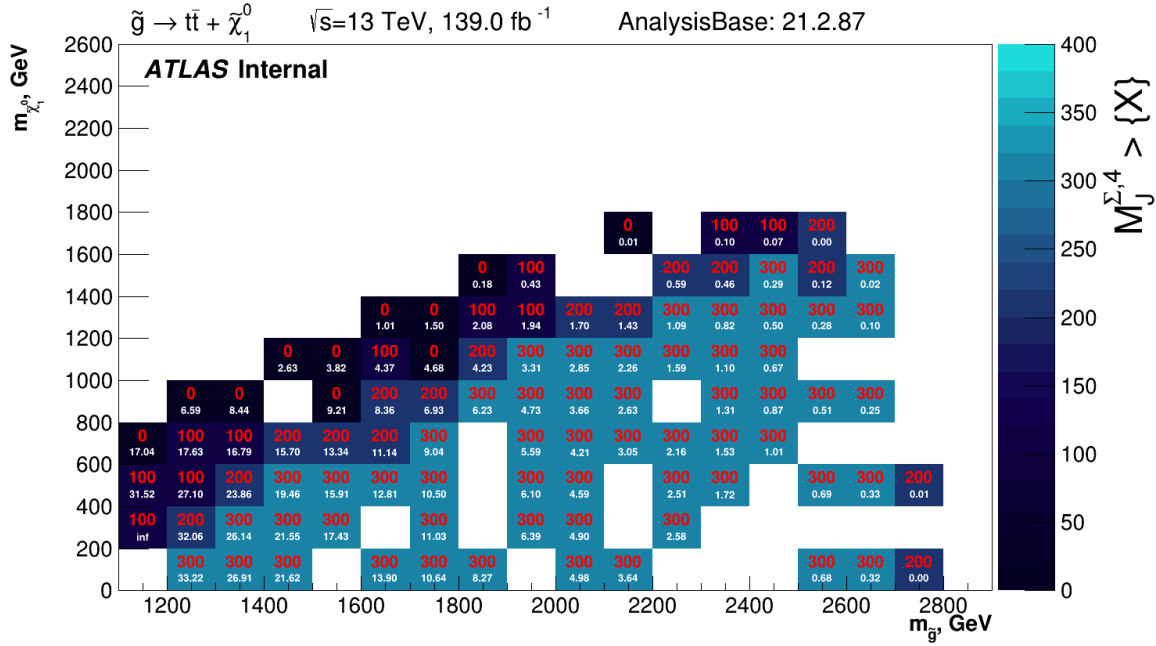
(f) $M_J^{\Sigma,4}$

Figure 95: The most optimal selections for each variable considered in the 0L channel signal models. The variables values are shown in red color, and the evaluated significance - in white.

(a) Jets multiplicity

(b) b-jets multiplicity

(c) E_T^{miss}

(d) m_{eff}

(e) m_{T}

172
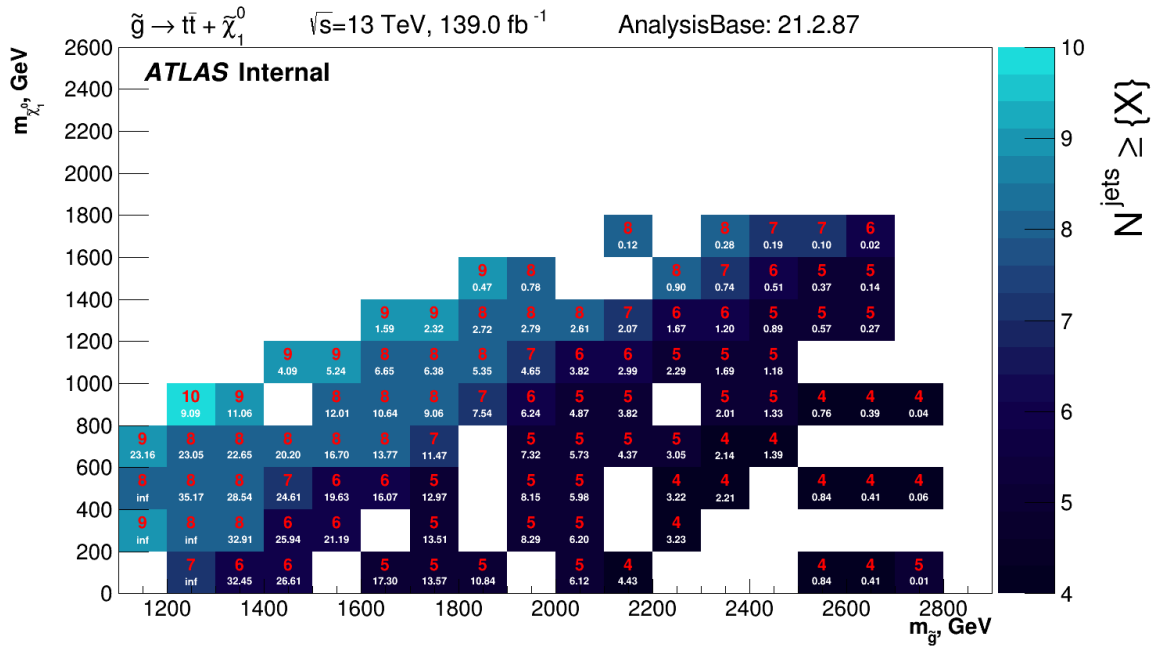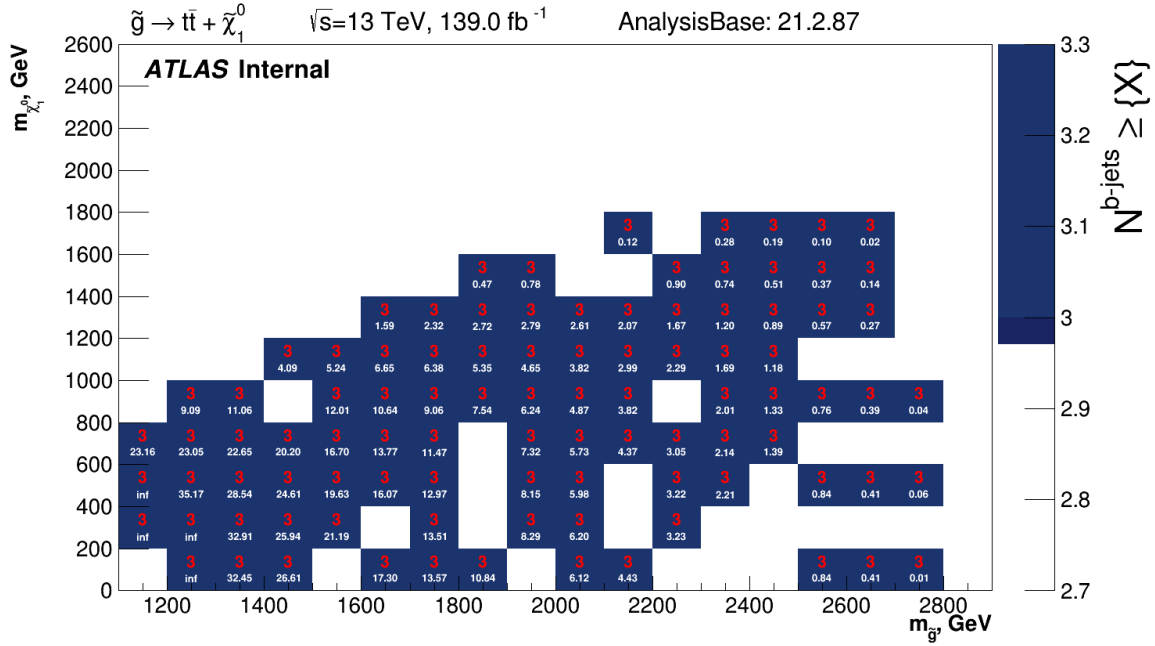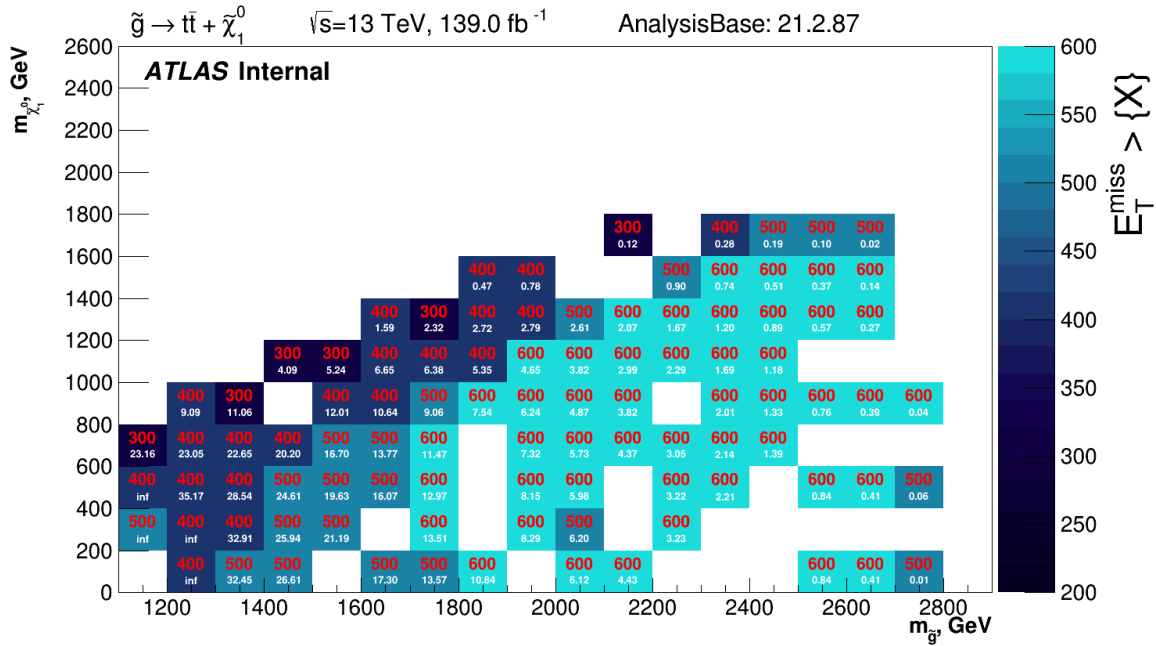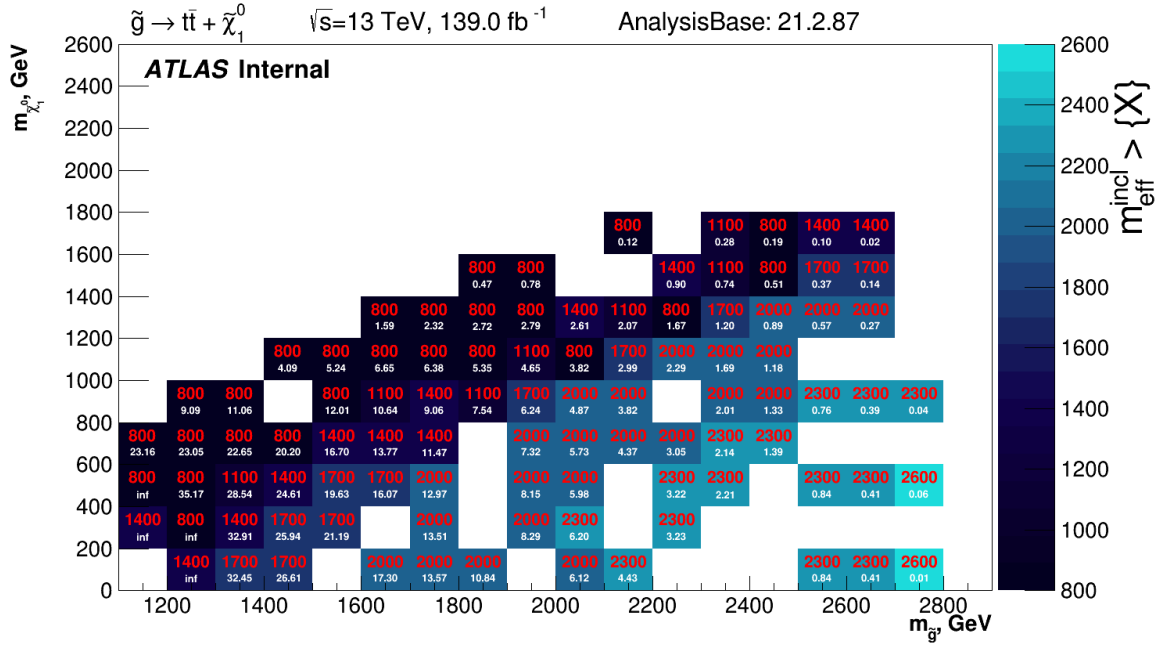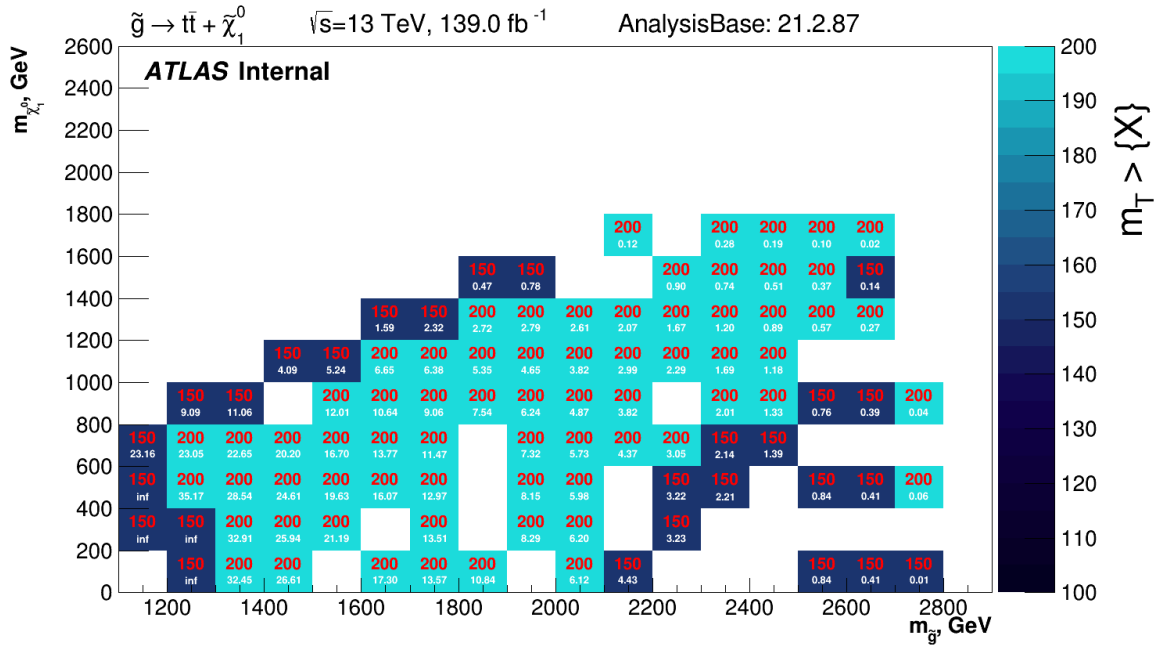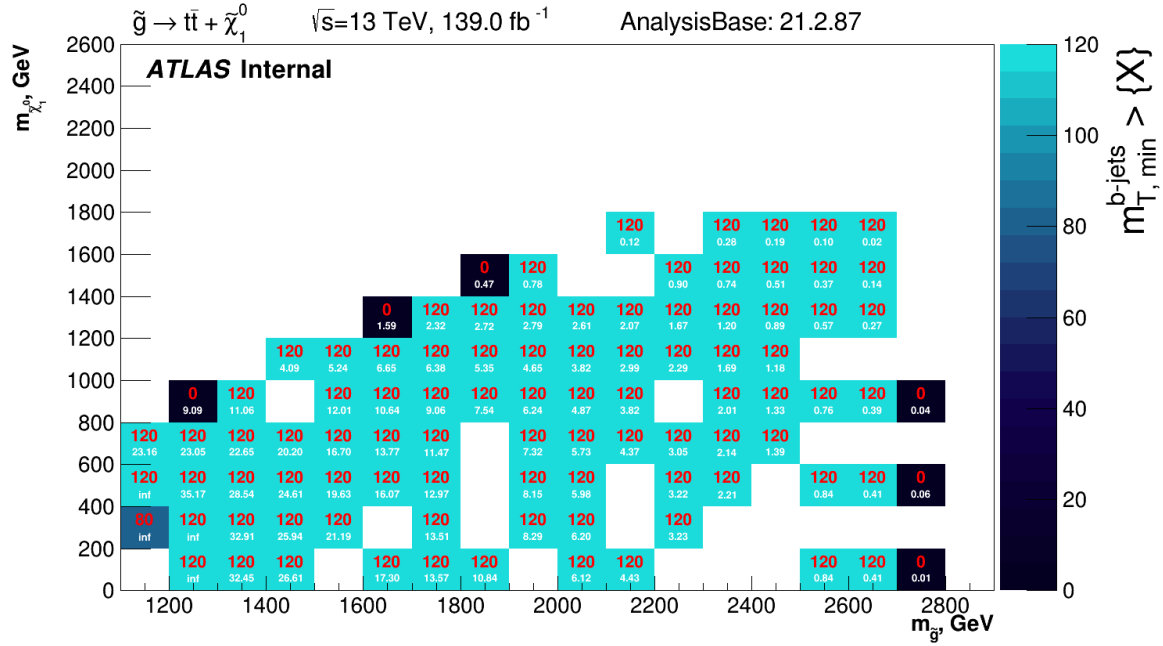
(f) $m_{\mathrm{T,min}}^{b-jets}$
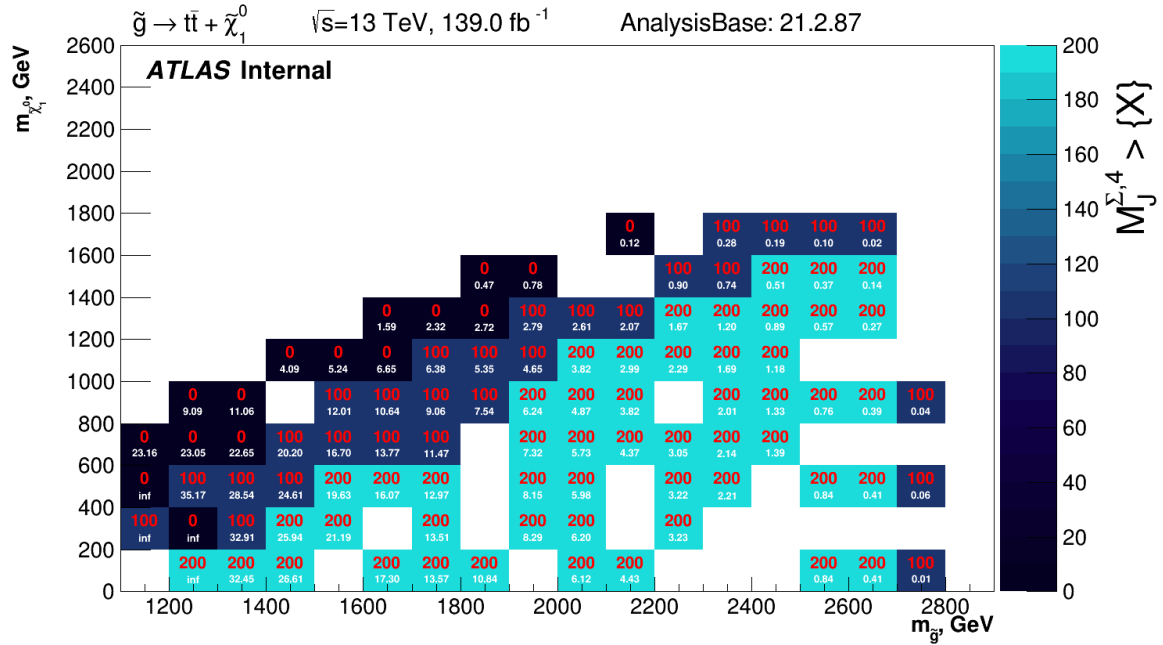
(g) $M_J^{\Sigma,4}$

Figure 96: The most optimal selections for each variable considered in the 1L channel signal models. The variables values are shown in red color, and the evaluated significance - in white.

173

0.4 Event Numbers in Gtt Regions

The appendix summarized tables with the total observed and expected background event numbers in control, validation and signal regions of the SUSY searches analyses in the Gtt channel.

table.results.yields channel	CR_Gtt_0L_B	SR_Gtt_0L_B	VR_Gtt_0L_B
Observed events	267	3	68
Fitted bkg events	266.64 ± 16.34	0.81 ± 0.37	82.90 ± 43.20
Fitted ttbar events	213.33 ± 24.07	0.34 ± 0.21	44.27 ± 30.16
Fitted singletop events	24.51 ± 8.33	0.11 ± 0.09	10.20 ± 7.84
Fitted topEW events	13.67 ± 7.97	$0.04^{+0.05}_{-0.04}$	3.12 ± 1.64
Fitted W_jets events	12.07 ± 5.59	0.15 ± 0.11	5.11 ± 2.42
Fitted Z_jets events	0.40 ± 0.24	0.13 ± 0.08	6.05 ± 2.14
Fitted diboson events	2.64 ± 1.69	0.00 ± 0.00	1.93 ± 1.07
Fitted QCD events	0.00 ± 0.00	$0.03^{+0.20}_{-0.03}$	$12.21^{+29.87}_{-12.21}$
MC exp. SM events	302.66	0.87	90.20
MC exp. ttbar events	249.11	0.40	51.55
MC exp. singletop events	24.72	0.11	10.21
MC exp. topEW events	13.67	0.04	3.12
MC exp. W_jets events	12.11	0.15	5.11
MC exp. Z_jets events	0.41	0.13	6.06
MC exp. diboson events	2.66	0.00	1.93
data-driven exp. QCD events	0.00	0.03	12.21

Table 16: 0L Boosted region

table.results.yields channel	CR_Gtt_0L_M1	SR_Gtt_0L_M1	VR_Gtt_0L_M1
Observed events	712	3	151
Fitted bkg events	711.94 ± 26.70	1.35 ± 0.65	161.65 ± 67.43
Fitted ttbar events	628.79 ± 38.54	0.78 ± 0.46	127.29 ± 65.41
Fitted singletop events	25.26 ± 6.54	0.16 ± 0.10	$7.73^{+8.65}_{-7.73}$
Fitted topEW events	38.50 ± 21.45	$0.09^{+0.12}_{-0.09}$	11.32 ± 6.03
Fitted W_jets events	16.64 ± 4.63	0.10 ± 0.09	6.26 ± 3.09
Fitted Z_jets events	0.64 ± 0.29	0.16 ± 0.08	5.02 ± 2.22
Fitted diboson events	2.12 ± 1.11	0.00 ± 0.00	0.35 ± 0.29
Fitted QCD events	0.00 ± 0.00	$0.06^{+0.25}_{-0.06}$	$3.68^{+8.26}_{-3.68}$
MC exp. SM events	682.41	1.31	155.67
MC exp. ttbar events	599.13	0.74	121.29
MC exp. singletop events	25.29	0.16	7.73
MC exp. topEW events	38.57	0.09	11.32
MC exp. W_jets events	16.66	0.10	6.27
MC exp. Z_jets events	0.64	0.16	5.03
MC exp. diboson events	2.12	0.00	0.35
data-driven exp. QCD events	0.00	0.06	3.68

Table 17: 0L Moderate 1 region

table.results.yields channel	CR_Gtt_0L_M2	SR_Gtt_0L_M2	VR_Gtt_0L_M2
Observed events	497	5	41
Fitted bkg events	497.33 ± 22.31	3.04 ± 1.21	40.02 ± 19.32
Fitted ttbar events	444.41 ± 29.67	2.01 ± 1.13	32.09 ± 18.90
Fitted singletop events	14.71 ± 4.15	0.38 ± 0.22	1.32 ± 1.28
Fitted topEW events	27.09 ± 15.28	0.32 ± 0.20	2.74 ± 1.50
Fitted W_jets events	9.71 ± 3.99	0.08 ± 0.05	2.39 ± 1.34
Fitted Z_jets events	0.47 ± 0.29	0.24 ± 0.14	0.69 ± 0.40
Fitted diboson events	0.93 ± 0.55	0.00 ± 0.00	0.06 ± 0.05
Fitted QCD events	0.00 ± 0.00	0.00 ± 0.00	$0.74^{+1.48}_{-0.74}$
MC exp. SM events	429.38	2.74	35.18
MC exp. ttbar events	376.52	1.71	27.22
MC exp. singletop events	14.73	0.38	1.32
MC exp. topEW events	27.00	0.32	2.74
MC exp. W_jets events	9.74	0.08	2.41
MC exp. Z_jets events	0.47	0.25	0.69
MC exp. diboson events	0.93	0.00	0.06
data-driven exp. QCD events	0.00	0.00	0.74

Table 18: 0L Moderate 2 region

table.results.yields channel	CR_Gtt_0L_C	SR_Gtt_0L_C	VR_Gtt_0L_C
Observed events	107	5	7
Fitted bkg events	107.01 ± 10.34	1.65 ± 0.80	7.06 ± 3.22
Fitted ttbar events	93.71 ± 11.89	1.20 ± 0.75	6.04 ± 3.04
Fitted singletop events	2.65 ± 0.85	0.16 ± 0.11	$0.03^{+0.12}_{-0.03}$
Fitted topEW events	9.21 ± 5.14	0.23 ± 0.15	0.58 ± 0.37
Fitted W_jets events	0.99 ± 0.60	0.03 ± 0.02	0.21 ± 0.08
Fitted Z_jets events	0.09 ± 0.05	$0.03^{+0.08}_{-0.03}$	0.00 ± 0.00
Fitted diboson events	0.35 ± 0.26	0.00 ± 0.00	0.00 ± 0.00
Fitted QCD events	0.00 ± 0.00	0.00 ± 0.00	$0.19^{+0.47}_{-0.19}$
MC exp. SM events	86.53	1.39	5.75
MC exp. ttbar events	73.22	0.94	4.73
MC exp. singletop events	2.66	0.16	0.03
MC exp. topEW events	9.22	0.23	0.58
MC exp. W_jets events	0.99	0.03	0.21
MC exp. Z_jets events	0.09	0.03	0.00
MC exp. diboson events	0.35	0.00	0.00
data-driven exp. QCD events	0.00	0.00	0.19

Table 19: 0L Compressed region

table.results.yields channel	CR_Gtt_1L_B	SR_Gtt_1L_B	VR1_Gtt_1L_B	VR2_Gtt_1L_B
Observed events	138	1	127	261
Fitted bkg events	137.88 ± 11.73	0.62 ± 0.43	128.18 ± 50.40	302.19 ± 117.37
Fitted ttbar events	107.03 ± 14.56	0.45 ± 0.42	107.17 ± 51.20	227.41 ± 120.75
Fitted singletop events	21.61 ± 5.92	0.11 ± 0.08	$11.37^{+12.90}_{-11.37}$	34.59 ± 12.96
Fitted topEW events	2.85 ± 1.58	0.05 ± 0.04	6.33 ± 3.47	19.97 ± 11.05
Fitted W_jets events	5.53 ± 2.49	$0.01^{+0.02}_{-0.01}$	2.34 ± 1.21	16.95 ± 4.57
Fitted Z_jets events	0.25 ± 0.24	0.00 ± 0.00	0.83 ± 0.19	0.39 ± 0.14
Fitted diboson events	0.62 ± 0.42	0.00 ± 0.00	0.14 ± 0.07	2.89 ± 1.53
Fitted QCD events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
MC exp. SM events	120.13	0.55	110.29	264.25
MC exp. ttbar events	89.24	0.38	89.26	189.40
MC exp. singletop events	21.63	0.11	11.39	34.62
MC exp. topEW events	2.85	0.05	6.33	19.98
MC exp. W_jets events	5.54	0.01	2.35	16.97
MC exp. Z_jets events	0.24	0.00	0.83	0.39
MC exp. diboson events	0.62	0.00	0.14	2.89
data-driven exp. QCD events	0.00	0.00	0.00	0.00

Table 20: 1L Boosted region

table.results.yields channel	CR_Gtt_1L_M1	SR_Gtt_1L_M1	VR1_Gtt_1L_M1	VR2_Gtt_1L_M1
Observed events	806	0	201	773
Fitted bkg events	805.54 ± 28.38	0.56 ± 0.27	178.85 ± 95.62	760.26 ± 252.50
Fitted ttbar events	660.33 ± 43.66	0.37 ± 0.22	150.56 ± 93.34	574.97 ± 232.21
Fitted singletop events	94.40 ± 22.36	0.08 ± 0.06	$13.27^{+13.98}_{-13.27}$	87.03 ± 29.64
Fitted topEW events	20.09 ± 10.87	0.10 ± 0.06	12.57 ± 6.88	45.29 ± 25.13
Fitted W_jets events	25.75 ± 9.02	$0.01^{+0.02}_{-0.01}$	1.31 ± 1.28	46.10 ± 13.13
Fitted Z_jets events	1.38 ± 0.51	0.00 ± 0.00	0.87 ± 0.24	1.32 ± 0.31
Fitted diboson events	3.60 ± 1.93	0.00 ± 0.00	0.28 ± 0.15	5.55 ± 3.08
Fitted QCD events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
MC exp. SM events	776.97	0.55	172.07	733.85
MC exp. ttbar events	631.78	0.36	143.79	548.66
MC exp. singletop events	94.33	0.08	13.27	86.99
MC exp. topEW events	20.11	0.10	12.56	45.26
MC exp. W_jets events	25.77	0.01	1.31	46.07
MC exp. Z_jets events	1.38	0.00	0.87	1.32
MC exp. diboson events	3.60	0.00	0.28	5.55
data-driven exp. QCD events	0.00	0.00	0.00	0.00

Table 21: 1L Moderate 2 region

table.results.yields channel	CR_Gtt_1L_M2	SR_Gtt_1L_M2	VR1_Gtt_1L_M2	VR2_Gtt_1L_M2
Observed events	908	0	14	133
Fitted bkg events	907.91 ± 30.15	0.98 ± 0.41	15.61 ± 8.67	103.20 ± 35.35
Fitted ttbar events	810.31 ± 43.12	0.63 ± 0.34	13.48 ± 8.72	81.66 ± 34.14
Fitted singletop events	35.47 ± 9.79	0.12 ± 0.08	$0.60^{+0.80}_{-0.60}$	6.06 ± 2.32
Fitted topEW events	39.28 ± 21.61	0.23 ± 0.15	1.32 ± 0.78	9.45 ± 5.34
Fitted W_jets events	19.89 ± 8.28	0.01 ± 0.00	0.19 ± 0.05	5.51 ± 1.96
Fitted Z_jets events	0.64 ± 0.37	0.00 ± 0.00	$0.02^{+0.02}_{-0.02}$	0.11 ± 0.08
Fitted diboson events	2.33 ± 1.41	0.00 ± 0.00	0.00 ± 0.00	0.42 ± 0.24
Fitted QCD events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
MC exp. SM events	883.60	0.96	15.21	100.78
MC exp. ttbar events	785.99	0.61	13.07	79.22
MC exp. singletop events	35.51	0.12	0.60	6.06
MC exp. topEW events	39.25	0.23	1.32	9.46
MC exp. W_jets events	19.89	0.01	0.19	5.51
MC exp. Z_jets events	0.64	0.00	0.02	0.11
MC exp. diboson events	2.33	0.00	0.00	0.42
data-driven exp. QCD events	0.00	0.00	0.00	0.00

Table 22: 1L Moderate 2 region

table.results.yields channel	CR_Gtt_1L_C	SR_Gtt_1L_C	VR1_Gtt_1L_C	VR2_Gtt_1L_C
Observed events	376	2	34	52
Fitted bkg events	375.81 ± 19.38	3.97 ± 1.96	32.21 ± 8.08	39.85 ± 7.48
Fitted ttbar events	335.47 ± 24.39	2.92 ± 1.77	27.12 ± 7.74	32.38 ± 6.08
Fitted singletop events	11.86 ± 3.38	0.23 ± 0.15	$0.87^{+1.23}_{-0.87}$	$1.74^{+2.18}_{-1.74}$
Fitted topEW events	19.30 ± 10.91	0.76 ± 0.47	3.43 ± 1.96	3.88 ± 2.21
Fitted W_jets events	8.20 ± 3.83	0.05 ± 0.04	0.67 ± 0.33	1.60 ± 0.64
Fitted Z_jets events	0.32 ± 0.21	$0.01^{+0.02}_{-0.01}$	0.11 ± 0.06	$0.04^{+0.04}_{-0.04}$
Fitted diboson events	0.66 ± 0.64	0.00 ± 0.00	0.00 ± 0.00	0.21 ± 0.12
Fitted QCD events	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
MC exp. SM events	339.30	3.66	29.28	36.37
MC exp. ttbar events	298.95	2.61	24.18	28.88
MC exp. singletop events	11.89	0.23	0.88	1.74
MC exp. topEW events	19.25	0.76	3.44	3.89
MC exp. W_jets events	8.22	0.05	0.67	1.61
MC exp. Z_jets events	0.32	0.01	0.11	0.04
MC exp. diboson events	0.67	0.00	0.00	0.21
data-driven exp. QCD events	0.00	0.00	0.00	0.00

Table 23: 1L Compressed region

0.5 Systematics tables

The appendix summarizes tables with breakdowns of uncertainty on background estimates for each of the kinematic regions in both 0L and 1L channels of the analysis. Note that the individual uncertainties can be correlated, and do not necessarily add up quadratically to the total background uncertainty. The percentages show the size of the uncertainty relative to the total expected background.

Uncertainty of channel	CR	VR	SR
Total background expectation	278.07	71.48	0.81
Total statistical ($\sqrt{N_{\mathrm{exp}}}$)	±16.68	±8.45	±0.90
Total background systematic	±16.67 [6.00%]	±15.58 [21.80%]	±0.20 [24.41%]
mu_ttbar_Gtt_0L_B	±18.90	±3.77	±0.03
alpha_kin_RW	±5.93	±4.97	±0.04
alpha_JES0	±3.00	±1.10	±0.04
alpha_JES1	±2.82	±0.41	±0.05
alpha_bTag_L	±2.08	±0.62	±0.01
alpha_JES6	±1.86	±0.05	±0.03
Lumi	±1.56	±0.77	±0.01
alpha_bTag_C	±1.50	±1.32	±0.02
alpha_bTag_B	±1.45	±0.40	±0.00
alpha_JES2	±1.28	±1.00	±0.00
alpha_JER1	±1.27	±1.65	±0.02
alpha_JER3	±1.11	±1.02	±0.04
alpha_JER0	±1.09	±1.89	±0.00
alpha_JER4	±1.06	±1.35	±0.01
alpha_JER6	±1.03	±1.26	±0.01
alpha_JER5	±0.98	±1.06	±0.01
alpha_bTag_extrapol	±0.87	±0.15	±0.00
alpha_JER2	±0.71	±1.44	±0.05
alpha_JER7	±0.68	±0.28	±0.01
alpha_JVT	±0.33	±0.17	±0.00
alpha_bTag_extrapol_charm	±0.17	±0.90	±0.01
alpha_JES4	±0.03	±0.02	±0.00
alpha_JES5	±0.01	±0.02	±0.00
alpha_JES3	±0.00	±0.01	±0.00
alpha_W_jets_syst_SR_Gtt_0L_B	±0.00	±0.00	±0.05
alpha_topEW_syst_VR_Gtt_0L_B	±0.00	±0.94	±0.00
alpha_singletop_syst_VR_Gtt_0L_B	±0.00	±3.06	±0.00
alpha_ttbar_syst_VR_Gtt_0L_B	±0.00	±13.39	±0.00
alpha_Z_jets_syst_VR_Gtt_0L_B	±0.00	±1.81	±0.00
alpha_diboson_syst_VR_Gtt_0L_B	±0.00	±0.58	±0.00
gamma_stat_SR_Gtt_0L_B_cuts_bin_0	±0.00	±0.00	±0.11
alpha_Z_jets_syst_SR_Gtt_0L_B	±0.00	±0.00	±0.04
alpha_ttbar_syst_SR_Gtt_0L_B	±0.00	±0.00	±0.10
alpha_W_jets_syst_VR_Gtt_0L_B	±0.00	±1.66	±0.00
alpha_topEW_syst_SR_Gtt_0L_B	±0.00	±0.00	±0.01
alpha_singletop_syst_SR_Gtt_0L_B	±0.00	±0.00	±0.03

Table 24: Systematics chart in the 0L channel, Boosted region

Uncertainty of channel	CR	VR	SR
Total background expectation	712.00	158.94	1.30
Total statistical ($\sqrt{N_{\mathrm{exp}}}$)	±26.68	±12.61	±1.14
Total background systematic	±26.68 [3.75%]	±41.62 [26.18%]	±0.44 [33.96%]
mu_ttbar_Gtt_0L_M1	±31.46	±6.37	±0.04
alpha_kin_RW	±11.57	±1.26	±0.01
alpha_JES1	±7.31	±6.87	±0.15
alpha_JES0	±6.48	±6.34	±0.14
alpha_JES6	±3.76	±3.99	±0.09
Lumi	±2.33	±0.90	±0.01
alpha_bTag_L	±2.13	±0.02	±0.05
alpha_JER1	±1.84	±5.45	±0.09
alpha_bTag_C	±1.73	±2.64	±0.00
alpha_JER2	±1.58	±3.33	±0.08
alpha_bTag_B	±1.46	±0.50	±0.01
alpha_JER0	±1.35	±7.00	±0.15
alpha_JER7	±1.33	±1.30	±0.03
alpha_JER3	±1.21	±2.05	±0.08
alpha_JER6	±1.14	±2.12	±0.10
alpha_JVT	±0.96	±0.94	±0.00
alpha_JER4	±0.64	±1.69	±0.07
alpha_JES2	±0.59	±1.56	±0.01
alpha_JER5	±0.59	±1.46	±0.09
alpha_bTag_extrapol_charm	±0.25	±2.24	±0.01
alpha_bTag_extrapol	±0.21	±0.30	±0.02
alpha_JES4	±0.00	±0.02	±0.00
alpha_JES5	±0.00	±0.01	±0.00
alpha_diboson_syst_VR_Gtt_0L_M1	±0.00	±0.10	±0.00
alpha_topEW_syst_VR_Gtt_0L_M1	±0.00	±3.40	±0.00
alpha_W_jets_syst_SR_Gtt_0L_M1	±0.00	±0.00	±0.03
alpha_ttbar_syst_VR_Gtt_0L_M1	±0.00	±38.34	±0.00
alpha_singletop_syst_VR_Gtt_0L_M1	±0.00	±2.32	±0.00
gamma_stat_SR_Gtt_0L_M1_cuts_bin_0	±0.00	±0.00	±0.13
alpha_ttbar_syst_SR_Gtt_0L_M1	±0.00	±0.00	±0.23
alpha_Z_jets_syst_VR_Gtt_0L_M1	±0.00	±1.50	±0.00
alpha_JES3	±0.00	±0.00	±0.00
alpha_Z_jets_syst_SR_Gtt_0L_M1	±0.00	±0.00	±0.05
alpha_singletop_syst_SR_Gtt_0L_M1	±0.00	±0.00	±0.05
alpha_W_jets_syst_VR_Gtt_0L_M1	±0.00	±2.02	±0.00
alpha_topEW_syst_SR_Gtt_0L_M1	±0.00	±0.00	±0.03

Table 25: Systematics chart in the 0L channel, Moderate-1 region

Uncertainty of channel	CR	VR	SR
Total background expectation	497.01	39.41	3.06
Total statistical ($\sqrt{N_\mathrm{exp}}$)	±22.29	±6.28	±1.75
Total background systematic	±22.29 [4.48%]	±11.25 [28.55%]	±0.85 [27.89%]
mu_ttbar_Gtt_0L_M2	±25.43	±1.84	±0.12
alpha_kin_RW	±7.75	±0.31	±0.02
alpha_JES1	±5.56	±3.21	±0.12
alpha_JES0	±5.10	±2.99	±0.16
alpha_JES6	±2.70	±1.24	±0.08
alpha_JER0	±1.74	±1.91	±0.16
alpha_JER1	±1.58	±1.71	±0.23
Lumi	±1.49	±0.21	±0.03
alpha_bTag_L	±1.38	±0.00	±0.07
alpha_JER2	±1.36	±0.91	±0.26
alpha_JER6	±1.35	±1.06	±0.21
alpha_JER4	±1.22	±0.76	±0.10
alpha_JER3	±1.19	±1.06	±0.13
alpha_JER7	±1.13	±0.03	±0.06
alpha_bTag_C	±1.08	±0.68	±0.02
alpha_JER5	±1.03	±0.43	±0.02
alpha_bTag_B	±0.98	±0.14	±0.00
alpha_JVT	±0.78	±0.31	±0.01
alpha_JES2	±0.36	±0.20	±0.02
alpha_bTag_extrapol_charm	±0.16	±0.55	±0.00
alpha_bTag_extrapol	±0.12	±0.05	±0.02
alpha_JES4	±0.01	±0.10	±0.00
alpha_JES5	±0.01	±0.05	±0.00
alpha_Z_jets_syst_SR_Gtt_0L_M2	±0.00	±0.00	±0.07
gamma_stat_SR_Gtt_0L_M2_cuts_bin_0	±0.00	±0.00	±0.22
alpha_W_jets_syst_SR_Gtt_0L_M2	±0.00	±0.00	±0.03
alpha_diboson_syst_VR_Gtt_0L_M2	±0.00	±0.02	±0.00
alpha_topEW_syst_VR_Gtt_0L_M2	±0.00	±0.82	±0.00
alpha_ttbar_syst_VR_Gtt_0L_M2	±0.00	±9.66	±0.00
alpha_singletop_syst_VR_Gtt_0L_M2	±0.00	±0.40	±0.00
alpha_Z_jets_syst_VR_Gtt_0L_M2	±0.00	±0.21	±0.00
alpha_JES3	±0.00	±0.00	±0.00
alpha_ttbar_syst_SR_Gtt_0L_M2	±0.00	±0.00	±0.61
alpha_singletop_syst_SR_Gtt_0L_M2	±0.00	±0.00	±0.11
alpha_topEW_syst_SR_Gtt_0L_M2	±0.00	±0.00	±0.10
alpha_W_jets_syst_VR_Gtt_0L_M2	±0.00	±0.72	±0.00

Table 26: Systematics chart in the 0L channel, Moderate-2 region

Uncertainty of channel	CR	VR	SR
Total background expectation	107.00	6.88	1.65
Total statistical ($\sqrt{N_{\mathrm{exp}}}$)	±10.34	±2.62	±1.29
Total background systematic	±10.34 [9.66%]	±3.16 [46.00%]	±0.60 [36.33%]
mu_ttbar_Gtt_0L_C	±10.82	±0.70	±0.14
alpha_kin_RW	±2.16	±0.06	±0.01
alpha_JES0	±1.13	±0.58	±0.14
alpha_JES1	±1.07	±0.45	±0.16
alpha_JER1	±0.72	±0.94	±0.05
alpha_bTag_L	±0.63	±0.17	±0.06
alpha_JER0	±0.59	±0.70	±0.15
alpha_JES6	±0.57	±0.06	±0.15
alpha_JER2	±0.57	±1.27	±0.26
alpha_bTag_B	±0.52	±0.08	±0.02
alpha_JER3	±0.39	±0.76	±0.04
Lumi	±0.38	±0.02	±0.01
alpha_bTag_C	±0.38	±0.13	±0.00
alpha_JER4	±0.33	±0.54	±0.01
alpha_JER5	±0.25	±0.63	±0.01
alpha_JER6	±0.25	±0.67	±0.08
alpha_JVT	±0.20	±0.07	±0.01
alpha_JES2	±0.16	±0.14	±0.05
alpha_JER7	±0.13	±0.44	±0.02
alpha_bTag_extrapol_charm	±0.05	±0.05	±0.00
alpha_bTag_extrapol	±0.04	±0.02	±0.01
alpha_JES4	±0.00	±0.00	±0.00
alpha_JES5	±0.00	±0.00	±0.00
alpha_W_jets_syst_SR_Gtt_0L_C	±0.00	±0.00	±0.01
alpha_topEW_syst_VR_Gtt_0L_C	±0.00	±0.17	±0.00
alpha_singletop_syst_VR_Gtt_0L_C	±0.00	±0.01	±0.00
alpha_ttbar_syst_VR_Gtt_0L_C	±0.00	±1.82	±0.00
gamma_stat_VR_Gtt_0L_C_cuts_bin_0	±0.00	±0.70	±0.00
gamma_stat_SR_Gtt_0L_C_cuts_bin_0	±0.00	±0.00	±0.18
alpha_Z_jets_syst_SR_Gtt_0L_C	±0.00	±0.00	±0.01
alpha_W_jets_syst_VR_Gtt_0L_C	±0.00	±0.06	±0.00
alpha_ttbar_syst_SR_Gtt_0L_C	±0.00	±0.00	±0.36
alpha_topEW_syst_SR_Gtt_0L_C	±0.00	±0.00	±0.07
alpha_JES3	±0.00	±0.00	±0.00
alpha_singletop_syst_SR_Gtt_0L_C	±0.00	±0.00	±0.05

Table 27: Systematics chart in the 0L channel, Compressed region

Uncertainty of channel	CR	VR1	VR2	SR
Total background expectation	138.00	128.24	301.32	0.62
Total statistical ($\sqrt{N_{\mathrm{exp}}}$)	±11.75	±11.32	±17.36	±0.79
Total background systematic	±11.74 [8.51%]	±47.00 [36.65%]	±95.79 [31.79%]	±0.39 [61.94%]
mu_ttbar_Gtt_1L_B	±13.15	±13.15	±27.89	±0.06
alpha_kin_RW	±4.13	±24.40	±28.74	±0.28
alpha_bTag_L	±1.85	±1.27	±4.16	±0.01
alpha_JER0	±1.54	±2.69	±5.33	±0.06
alpha_JER1	±1.49	±2.99	±6.19	±0.07
alpha_JER2	±1.32	±2.53	±4.86	±0.08
alpha_JER6	±1.28	±1.49	±3.14	±0.05
alpha_bTag_C	±1.21	±0.90	±0.94	±0.00
alpha_bTag_B	±1.15	±0.05	±1.41	±0.01
Lumi	±0.88	±0.60	±2.11	±0.00
alpha_JES2	±0.79	±0.54	±0.84	±0.01
alpha_bTag_extrapol	±0.71	±3.11	±7.63	±0.00
alpha_JER4	±0.69	±1.29	±1.46	±0.06
alpha_JER3	±0.63	±0.84	±1.43	±0.04
alpha_JER5	±0.53	±0.58	±0.09	±0.05
alpha_JER7	±0.42	±0.26	±1.35	±0.01
alpha_JES1	±0.36	±6.68	±29.16	±0.03
alpha_bTag_extrapol_charm	±0.16	±2.28	±0.81	±0.00
alpha_JES0	±0.14	±5.96	±24.82	±0.05
alpha_JVT	±0.05	±0.65	±1.87	±0.00
alpha_JES6	±0.04	±3.24	±17.61	±0.02
alpha_JES5	±0.04	±0.04	±0.10	±0.00
alpha_JES4	±0.00	±0.01	±0.14	±0.00
alpha_JES3	±0.00	±0.01	±0.01	±0.00
alpha_topEW_syst_VR2_Gtt_1L_B	±0.00	±0.00	±5.99	±0.00
alpha_ttbar_syst_SR_Gtt_1L_B	±0.00	±0.00	±0.00	±0.14
alpha_ttbar_syst_VR2_Gtt_1L_B	±0.00	±0.00	±68.48	±0.00
alpha_topEW_syst_VR1_Gtt_1L_B	±0.00	±1.90	±0.00	±0.00
alpha_Z_jets_syst_VR2_Gtt_1L_B	±0.00	±0.00	±0.11	±0.00
alpha_Z_jets_syst_VR1_Gtt_1L_B	±0.00	±0.22	±0.00	±0.00
alpha_diboson_syst_VR2_Gtt_1L_B	±0.00	±0.00	±0.87	±0.00
alpha_W_jets_syst_VR1_Gtt_1L_B	±0.00	±0.62	±0.00	±0.00
alpha_ttbar_syst_VR1_Gtt_1L_B	±0.00	±32.27	±0.00	±0.00
gamma_stat_SR_Gtt_1L_B_cuts_bin_0	±0.00	±0.00	±0.00	±0.06
alpha_W_jets_syst_SR_Gtt_1L_B	±0.00	±0.00	±0.00	±0.00
alpha_W_jets_syst_VR2_Gtt_1L_B	±0.00	±0.00	±4.56	±0.00
alpha_singletop_syst_VR1_Gtt_1L_B	±0.00	±3.42	±0.00	±0.00
alpha_diboson_syst_VR1_Gtt_1L_B	±0.00	±0.04	±0.00	±0.00
alpha_singletop_syst_VR2_Gtt_1L_B	±0.00	±0.00	±10.39	±0.00
alpha_topEW_syst_SR_Gtt_1L_B	±0.00	±0.00	±0.00	±0.01
alpha_singletop_syst_SR_Gtt_1L_B	±0.00	±0.00	±0.00	±0.03

Table 28: Systematics chart in the 1L channel, Boosted region

Uncertainty of channel	CR	VR1	VR2	SR
Total background expectation	806.12	178.98	755.60	0.56
Total statistical ($\sqrt{N_{\mathrm{exp}}}$)	±28.39	±13.38	±27.49	±0.75
Total background systematic	±28.39 [3.52%]	±53.42 [29.85%]	±248.43 [32.88%]	±0.19 [33.89%]
mu_ttbar_Gtt_1L_M1	±38.48	±8.76	±33.42	±0.02
alpha_kin_RW	±21.97	±30.03	±161.52	±0.02
alpha_bTag_L	±5.82	±0.46	±1.90	±0.02
alpha_bTag_C	±5.26	±1.41	±0.02	±0.00
alpha_JER1	±4.51	±0.50	±13.98	±0.05
alpha_bTag_B	±4.37	±0.68	±0.37	±0.00
Lumi	±4.13	±0.81	±5.19	±0.01
alpha_JER2	±3.81	±0.85	±11.80	±0.05
alpha_JER0	±3.74	±0.60	±11.94	±0.04
alpha_JES0	±2.54	±5.79	±58.05	±0.03
alpha_JER6	±2.27	±1.09	±5.22	±0.05
alpha_JER3	±2.13	±0.43	±4.07	±0.05
alpha_JES1	±2.09	±7.31	±67.18	±0.02
alpha_JER4	±1.79	±0.36	±5.76	±0.04
alpha_JER7	±1.71	±0.71	±3.72	±0.00
alpha_JES6	±1.67	±3.56	±40.63	±0.02
alpha_JER5	±1.52	±0.51	±3.93	±0.04
alpha_JES2	±1.51	±0.65	±0.61	±0.03
alpha_bTag_extrapol	±1.07	±1.31	±4.75	±0.01
alpha_bTag_extrapol_charm	±0.78	±3.57	±2.77	±0.00
alpha_JVT	±0.49	±0.71	±4.61	±0.00
alpha_JES5	±0.08	±0.04	±0.01	±0.00
alpha_JES4	±0.01	±0.00	±0.13	±0.00
alpha_topEW_syst_VR1_Gtt_1L_M1	±0.00	±3.77	±0.00	±0.00
alpha_Z_jets_syst_VR2_Gtt_1L_M1	±0.00	±0.00	±0.35	±0.00
gamma_stat_SR_Gtt_1L_M1_cuts_bin_0	±0.00	±0.00	±0.00	±0.06
alpha_W_jets_syst_VR2_Gtt_1L_M1	±0.00	±0.00	±12.19	±0.00
alpha_singletop_syst_VR1_Gtt_1L_M1	±0.00	±3.98	±0.00	±0.00
alpha_ttbar_syst_VR2_Gtt_1L_M1	±0.00	±0.00	±172.79	±0.00
alpha_Z_jets_syst_SR_Gtt_1L_M1	±0.00	±0.00	±0.00	±0.00
alpha_diboson_syst_VR1_Gtt_1L_M1	±0.00	±0.08	±0.00	±0.00
alpha_topEW_syst_SR_Gtt_1L_M1	±0.00	±0.00	±0.00	±0.03
alpha_ttbar_syst_SR_Gtt_1L_M1	±0.00	±0.00	±0.00	±0.11
alpha_diboson_syst_VR2_Gtt_1L_M1	±0.00	±0.00	±1.67	±0.00
alpha_JES3	±0.00	±0.00	±0.01	±0.00
alpha_topEW_syst_VR2_Gtt_1L_M1	±0.00	±0.00	±13.58	±0.00
alpha_Z_jets_syst_VR1_Gtt_1L_M1	±0.00	±0.25	±0.00	±0.00
alpha_singletop_syst_VR2_Gtt_1L_M1	±0.00	±0.00	±26.10	±0.00
alpha_singletop_syst_SR_Gtt_1L_M1	±0.00	±0.00	±0.00	±0.02
alpha_W_jets_syst_VR1_Gtt_1L_M1	±0.00	±0.31	±0.00	±0.00
alpha_W_jets_syst_SR_Gtt_1L_M1	±0.00	±0.00	±0.00	±0.00
alpha_ttbar_syst_VR1_Gtt_1L_M1	±0.00	±45.30	±0.00	±0.00

Table 29: Systematics chart in the 1L channel, Moderate-1 region

Uncertainty of channel	CR	VR1	VR2	SR
Total background expectation	908.01	15.62	102.83	0.98
Total statistical ($\sqrt{N_{\mathrm{exp}}}$)	±30.13	±3.95	±10.14	±0.99
Total background systematic	±30.13 [3.32%]	±4.71 [30.12%]	±27.41 [26.66%]	±0.27 [27.23%]
mu_ttbar_Gtt_1L_M2	±36.42	±0.61	±3.67	±0.03
alpha_kin_RW	±15.06	±0.39	±2.95	±0.06
alpha_JES1	±8.43	±0.74	±7.04	±0.06
alpha_JES0	±7.08	±0.63	±6.24	±0.06
alpha_JES6	±3.60	±0.19	±4.62	±0.02
alpha_JER1	±2.76	±0.95	±2.09	±0.06
Lumi	±2.75	±0.06	±0.60	±0.01
alpha_bTag_C	±2.40	±0.02	±0.05	±0.01
alpha_bTag_L	±2.36	±0.14	±1.04	±0.02
alpha_JER6	±2.13	±0.68	±1.86	±0.01
alpha_bTag_B	±1.89	±0.12	±0.17	±0.00
alpha_JER0	±1.77	±0.09	±2.79	±0.04
alpha_JER4	±1.72	±0.41	±1.90	±0.02
alpha_JER3	±1.70	±0.47	±1.86	±0.04
alpha_JER2	±1.64	±1.10	±2.40	±0.03
alpha_JER7	±1.57	±0.28	±0.20	±0.01
alpha_JER5	±1.25	±0.39	±0.63	±0.00
alpha_JVT	±1.15	±0.08	±0.49	±0.00
alpha_JES2	±0.49	±0.05	±0.93	±0.01
alpha_bTag_extrapol_charm	±0.25	±0.22	±0.25	±0.00
alpha_bTag_extrapol	±0.17	±0.03	±0.16	±0.01
alpha_JES4	±0.02	±0.01	±0.03	±0.00
alpha_JES5	±0.01	±0.00	±0.00	±0.00
alpha_singletop_syst_VR2_Gtt_1L_M2	±0.00	±0.00	±1.82	±0.00
alpha_topEW_syst_VR1_Gtt_1L_M2	±0.00	±0.40	±0.00	±0.00
alpha_Z_jets_syst_VR2_Gtt_1L_M2	±0.00	±0.00	±0.03	±0.00
gamma_stat_VR1_Gtt_1L_M2_cuts_bin_0	±0.00	±1.03	±0.00	±0.00
alpha_topEW_syst_VR2_Gtt_1L_M2	±0.00	±0.00	±2.84	±0.00
gamma_stat_SR_Gtt_1L_M2_cuts_bin_0	±0.00	±0.00	±0.00	±0.11
alpha_singletop_syst_VR1_Gtt_1L_M2	±0.00	±0.18	±0.00	±0.00
alpha_W_jets_syst_VR2_Gtt_1L_M2	±0.00	±0.00	±1.47	±0.00
alpha_ttbar_syst_VR2_Gtt_1L_M2	±0.00	±0.00	±24.57	±0.00
alpha_Z_jets_syst_SR_Gtt_1L_M2	±0.00	±0.00	±0.00	±0.00
alpha_ttbar_syst_SR_Gtt_1L_M2	±0.00	±0.00	±0.00	±0.19
alpha_topEW_syst_SR_Gtt_1L_M2	±0.00	±0.00	±0.00	±0.07
alpha_diboson_syst_VR2_Gtt_1L_M2	±0.00	±0.00	±0.12	±0.00
alpha_JES3	±0.00	±0.00	±0.00	±0.00
alpha_Z_jets_syst_VR1_Gtt_1L_M2	±0.00	±0.00	±0.00	±0.00
alpha_singletop_syst_SR_Gtt_1L_M2	±0.00	±0.00	±0.00	±0.03
alpha_W_jets_syst_SR_Gtt_1L_M2	±0.00	±0.00	±0.00	±0.00
alpha_W_jets_syst_VR1_Gtt_1L_M2	±0.00	±0.06	±0.00	±0.00
alpha_ttbar_syst_VR1_Gtt_1L_M2	±0.00	±4.05	±0.00	±0.00

Table 30: Systematics chart in the 1L channel, Moderate-2 region

Uncertainty of channel	CR	VR1	VR2	SR
Total background expectation	376.02	32.28	39.87	3.98
Total statistical ($\sqrt{N_{\exp}}$)	±19.39	±5.68	±6.31	±1.99
Total background systematic	±19.39 [5.16%]	±8.56 [26.53%]	±11.32 [28.39%]	±1.48 [37.31%]
mu_ttbar_Gtt_1L_C	±21.82	±1.76	±2.11	±0.19
alpha_kin_RW	±6.03	±0.78	±1.04	±0.17
alpha_JES1	±4.21	±0.98	±3.75	±0.34
alpha_JES0	±4.03	±0.97	±3.20	±0.34
alpha_JER0	±2.32	±0.01	±0.76	±0.43
alpha_JER1	±2.06	±0.41	±0.86	±0.46
alpha_JES6	±2.04	±0.64	±2.13	±0.16
alpha_JER2	±1.86	±0.69	±0.57	±0.41
alpha_JER4	±1.32	±0.15	±0.45	±0.31
alpha_JER6	±1.29	±0.51	±0.79	±0.36
alpha_bTag_L	±1.18	±0.09	±0.58	±0.02
Lumi	±1.14	±0.15	±0.21	±0.03
alpha_JER7	±0.93	±0.19	±0.35	±0.15
alpha_JER3	±0.90	±0.06	±0.34	±0.27
alpha_bTag_C	±0.86	±0.30	±0.07	±0.03
alpha_JER5	±0.77	±0.10	±0.29	±0.27
alpha_bTag_B	±0.75	±0.09	±0.15	±0.02
alpha_JVT	±0.59	±0.09	±0.18	±0.01
alpha_JES2	±0.15	±0.10	±0.26	±0.04
alpha_bTag_extrapol_charm	±0.15	±0.55	±0.05	±0.01
alpha_bTag_extrapol	±0.10	±0.01	±0.03	±0.02
alpha_JES4	±0.01	±0.00	±0.03	±0.00
alpha_JES5	±0.00	±0.01	±0.09	±0.00
alpha_topEW_syst_VR2_Gtt_1L_C	±0.00	±0.00	±1.17	±0.00
alpha_ttbar_syst_SR_Gtt_1L_C	±0.00	±0.00	±0.00	±0.88
alpha_Z_jets_syst_SR_Gtt_1L_C	±0.00	±0.00	±0.00	±0.00
alpha_ttbar_syst_VR2_Gtt_1L_C	±0.00	±0.00	±9.76	±0.00
gamma_stat_SR_Gtt_1L_C_cuts_bin_0	±0.00	±0.00	±0.00	±0.29
alpha_Z_jets_syst_VR1_Gtt_1L_C	±0.00	±0.03	±0.00	±0.00
alpha_diboson_syst_VR2_Gtt_1L_C	±0.00	±0.00	±0.06	±0.00
alpha_Z_jets_syst_VR2_Gtt_1L_C	±0.00	±0.00	±0.01	±0.00
alpha_W_jets_syst_VR1_Gtt_1L_C	±0.00	±0.19	±0.00	±0.00
alpha_ttbar_syst_VR1_Gtt_1L_C	±0.00	±8.16	±0.00	±0.00
alpha_W_jets_syst_SR_Gtt_1L_C	±0.00	±0.00	±0.00	±0.02
alpha_topEW_syst_VR1_Gtt_1L_C	±0.00	±1.03	±0.00	±0.00
alpha_W_jets_syst_VR2_Gtt_1L_C	±0.00	±0.00	±0.43	±0.00
alpha_singletop_syst_SR_Gtt_1L_C	±0.00	±0.00	±0.00	±0.07
alpha_JES3	±0.00	±0.00	±0.00	±0.00
alpha_singletop_syst_VR1_Gtt_1L_C	±0.00	±0.26	±0.00	±0.00
alpha_singletop_syst_VR2_Gtt_1L_C	±0.00	±0.00	±0.53	±0.00
alpha_topEW_syst_SR_Gtt_1L_C	±0.00	±0.00	±0.00	±0.23

Table 31: Systematics chart in the 1L channel, Compressed region

0.6 Unfolded Distributions Systematics Tables

The appendix summarizes tables with systematics uncertainties of the unfolded spectra presented in section 7.8. The tables show the uncertainties bin-by-bin for each of the distributions.

3j3b channel:

Uncert. (%)	1	2	3	4	5
Data Stat.	3.5	2.4	3.1	3.4	6.7
Matrix Uncert	5.7	1.4	1.5	5.5	5.1
PS Uncert	3.4	0.2	0.8	2.2	2.5
Rad Uncert	1.2	1.8	0.6	1.2	3.0
BkgStat Uncert	0.2	0.1	0.2	0.2	0.4
TrainingStat Uncert	0.6	0.4	0.5	0.8	1.1
BTag	3.5	0.9	1.3	2.2	3.1
JVT	0.8	0.0	0.3	0.4	0.5
Lep	1.2	0.3	0.3	0.7	1.4
PileUp	0.1	0.2	0.3	0.3	0.3
muAlphaSFSR0p5 Uncert	5.3	0.7	1.5	3.1	0.2
muF0p5 Uncert	0.3	0.4	0.2	0.5	1.8
muR0p5 Uncert	0.3	0.5	0.0	0.5	1.4
var3cUp Uncert	0.3	0.2	0.1	0.4	0.2
Total Sys.	9.4	2.7	2.8	7.2	7.8
Total Uncert.	10.0	3.7	4.2	8.0	10.2

Table 32: p_{T} of the leading b-jet from top: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4
Data Stat.	2.0	1.7	3.0	5.5
Matrix Uncert	2.3	0.8	1.7	1.3
PS Uncert	2.3	1.2	1.7	0.1
Rad Uncert	1.0	0.8	0.2	0.1
BkgStat Uncert	0.1	0.1	0.2	0.3
TrainingStat Uncert	0.4	0.4	0.6	1.0
BTag	3.5	1.0	2.7	3.9
JVT	0.3	0.3	0.5	0.7
Lep	1.5	0.2	1.3	1.5
PileUp	0.1	0.2	0.3	0.3
muAlphaSFSR0p5 Uncert	1.8	1.3	0.4	2.5
muF0p5 Uncert	0.1	0.3	0.0	1.0
muR0p5 Uncert	0.3	0.1	0.3	0.8
var3cUp Uncert	0.3	0.1	0.2	0.1
Total Sys.	5.4	2.4	4.0	5.3
Total Uncert.	5.8	3.0	5.0	7.7

Table 33: p_{T} of the sub-leading b-jet from top: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5	6	7	8	9	10		
Data Stat.	6.7	7.2	6.3	6.2	5.6	6.1	5.2	6.7	5.7	6.6	6.3	6.9
Matrix Uncert	3.3	4.3	1.9	1.9	0.3	4.8	1.0	1.1	3.1	1.7	2.6	3.2
PS Uncert	0.6	0.7	2.2	4.3	1.8	3.1	0.0	1.7	0.3	3.8	1.2	1.0
Rad Uncert	2.4	1.5	2.0	2.2	1.4	3.4	0.9	2.0	1.2	0.7	2.1	0.7
BkgStat Uncert	0.5	0.4	0.4	0.4	0.3	0.4	0.4	0.4	0.3	0.4	0.4	0.4
TrainingStat Uncert	1.3	1.3	1.3	1.2	1.1	1.1	1.0	1.1	1.1	1.3	1.2	1.3
BTag	0.5	0.1	0.1	0.3	0.2	0.2	0.2	0.2	0.2	0.1	0.1	0.6
Lep	0.7	0.6	0.5	0.3	0.3	0.2	0.3	0.3	0.3	0.3	0.3	0.3
PileUp	1.0	1.1	0.7	0.4	0.2	1.1	1.6	0.4	0.5	0.3	0.5	1.2
muAlphaSFSR0p5 Uncert	2.0	3.3	3.6	2.7	1.1	5.3	4.0	4.6	1.9	3.1	1.9	0.9
muF0p5 Uncert	0.1	0.2	0.1	0.0	0.1	0.3	0.1	0.3	0.1	0.1	0.1	0.1
muR0p5 Uncert	0.1	0.1	0.2	0.3	0.1	0.4	0.3	0.0	0.3	0.8	0.3	0.3
var3cUp Uncert	0.3	0.7	0.1	0.7	0.4	0.1	0.1	0.0	0.2	0.3	0.0	0.1
Total Sys.	5.0	6.0	5.3	6.0	2.8	8.7	4.7	5.6	4.0	5.5	4.3	4.1
Total Uncert.	8.3	9.4	8.2	8.7	6.2	10.6	7.0	8.7	7.0	8.6	7.6	8.0

Table 34: η of the leading b-jet from top: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5	6	7	8
Data Stat.	5.1	5.0	5.4	4.5	4.3	4.7	4.8	6.0
Matrix Uncert	2.0	1.4	1.7	1.3	1.0	1.8	2.2	1.5
PS Uncert	3.6	0.7	0.6	0.9	0.2	3.8	0.1	1.4
Rad Uncert	3.3	0.3	0.4	1.5	0.4	1.5	0.8	0.1
BkgStat Uncert	0.4	0.3	0.3	0.3	0.2	0.3	0.3	0.4
TrainingStat Uncert	1.1	0.9	0.9	0.8	0.8	0.9	0.9	1.1
BTag	0.9	0.0	0.3	0.4	0.3	0.3	0.0	0.9
Lep	0.5	0.2	0.5	0.2	0.7	0.5	0.2	1.5
PileUp	0.6	0.4	0.7	0.4	0.4	0.3	0.4	0.4
muAlphaSFSR0p5 Uncert	3.0	0.8	1.4	2.1	1.0	2.0	2.1	5.4
muF0p5 Uncert	0.2	0.2	0.2	0.0	0.2	0.2	0.1	0.1
muR0p5 Uncert	0.3	0.0	0.2	0.1	0.2	0.1	0.0	0.2
var3cUp Uncert	0.7	0.2	0.3	0.1	0.3	0.2	0.4	0.5
Total Sys.	6.3	2.1	2.7	3.2	1.9	5.1	3.3	6.2
Total Uncert.	8.1	5.5	6.1	5.5	4.7	6.9	5.9	8.6

Table 35: η of the sub-leading b-jet from top: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5	6
Data Stat.	0.0	4.9	2.8	2.5	2.5	4.7
Matrix Uncert	0.0	4.9	4.0	1.1	2.4	8.7
PS Uncert	0.0	2.1	1.8	0.0	0.1	4.5
Rad Uncert	0.0	0.1	1.6	2.0	1.0	3.6
BkgStat Uncert	0.0	0.3	0.2	0.1	0.1	0.3
TrainingStat Uncert	0.0	0.9	0.5	0.5	0.5	0.8
BTag	0.0	1.3	1.3	0.5	0.7	2.5
JVT	0.0	0.3	0.2	0.1	0.1	0.4
Lep	0.0	0.6	0.6	0.4	0.4	1.4
PileUp	0.0	0.5	0.2	0.1	0.6	0.5
muAlphaSFSR0p5 Uncert	0.0	0.9	2.0	1.3	1.8	2.7
muF0p5 Uncert	0.0	0.6	0.4	0.4	0.1	1.7
muR0p5 Uncert	0.0	0.6	0.8	0.6	0.0	2.4
var3cUp Uncert	0.0	0.2	0.0	0.0	0.1	0.0
Total Sys.	0.0	5.8	5.4	2.9	3.4	11.6
Total Uncert.	0.0	7.6	6.1	3.8	4.2	12.5

Table 36: Vector sum of the leading and sub-leading b-jets from tops p_{T}: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4
Data Stat.	2.8	2.6	3.4	3.1
Matrix Uncert	1.6	0.1	2.1	3.8
PS Uncert	2.3	2.0	1.8	0.9
Rad Uncert	2.5	2.7	1.0	3.8
TrainingStat Uncert	0.6	0.5	0.6	0.6
BTag	3.5	0.9	1.6	2.3
JVT	0.4	0.2	0.2	0.4
Lep	3.9	0.4	1.8	2.1
PileUp	0.2	0.3	0.5	0.4
muAlphaSFSR0p5 Uncert	1.0	2.2	2.3	2.5
muF0p5 Uncert	0.3	0.4	0.2	0.8
muR0p5 Uncert	0.4	0.6	0.5	1.2
var3cUp Uncert	0.6	0.3	0.2	0.7
Total Sys.	6.6	4.3	4.6	6.9
Total Uncert.	7.2	5.0	5.7	7.6

Table 37: p_{T} of the leading additional b-jet: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3
Data Stat.	6.0	8.3	10.4
Matrix Uncert	0.1	0.5	0.4
PS Uncert	0.6	0.4	0.6
Rad Uncert	2.5	0.4	4.6
BkgStat Uncert	0.3	0.5	0.7
TrainingStat Uncert	0.8	1.0	1.3
BTag	3.2	1.9	3.4
JVT	0.2	0.1	0.5
Lep	0.3	0.5	0.2
PileUp	0.3	0.6	0.6
muAlphaSFSR0p5 Uncert	1.0	1.3	0.3
muF0p5 Uncert	0.4	0.0	0.7
muR0p5 Uncert	0.4	0.2	0.8
var3cUp Uncert	0.0	0.2	0.2
Total Sys.	4.3	2.8	6.1
Total Uncert.	7.4	8.8	12.0

Table 38: p_{T} of the sub-leading additional b-jet: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5	6	7	8	9	10
Data Stat.	10.1	5.2	4.8	5.2	4.8	5.4	5.5	5.4	4.9	11.3
Matrix Uncert	4.5	0.3	1.2	3.4	1.2	0.7	2.9	2.7	0.1	6.5
PS Uncert	4.4	0.6	1.6	0.3	1.0	2.2	0.3	1.4	1.7	0.7
Rad Uncert	0.2	1.4	0.1	2.2	0.9	0.1	0.1	1.0	1.1	1.1
BkgStat Uncert	0.8	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.7
TrainingStat Uncert	2.0	1.1	1.0	1.0	0.9	0.9	1.0	1.0	1.0	1.9
BTag	2.7	0.5	0.4	0.7	0.7	0.8	0.6	0.3	0.6	2.7
Lep	2.4	0.8	0.8	0.6	1.1	0.9	0.2	0.9	1.0	2.8
PileUp	0.8	0.2	0.4	0.5	0.4	0.3	0.2	0.5	0.9	0.6
muAlphaSFSR0p5 Uncert	4.1	0.7	2.5	1.1	0.5	1.5	1.4	1.6	5.2	1.8
muF0p5 Uncert	0.2	0.5	0.2	0.3	0.5	0.0	0.1	0.1	0.2	0.7
muR0p5 Uncert	0.1	0.3	0.4	0.2	0.2	0.2	0.4	0.3	0.3	0.2
var3cUp Uncert	1.4	0.2	0.2	0.4	0.4	0.3	0.5	0.1	0.4	0.8
Total Sys.	8.7	2.3	3.5	4.5	2.5	3.2	3.5	3.9	5.9	8.2
Total Uncert.	13.4	5.7	5.9	6.8	5.4	6.3	6.5	6.6	7.7	14.0

Table 39: η of the leading additional b-jet: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5	6	7	8	9
Data Stat.	20.2	18.6	18.5	15.9	19.2	17.6	19.3	17.3	17.2
Matrix Uncert	6.8	0.5	2.2	0.0	4.2	1.2	1.9	8.1	3.0
PS Uncert	2.6	2.2	4.5	6.4	3.7	4.1	2.8	3.8	4.2
Rad Uncert	3.4	1.0	3.6	1.2	1.6	13.3	5.7	1.6	0.0
BkgStat Uncert	1.0	1.0	1.4	0.7	1.2	0.9	1.3	1.0	0.9
TrainingStat Uncert	2.2	2.2	2.3	2.1	2.5	2.2	2.4	2.1	2.0
BTag	0.5	0.1	0.4	0.2	0.4	0.2	0.3	0.2	0.3
Lep	0.7	0.7	0.9	1.1	0.9	0.9	0.5	0.4	1.0
PileUp	1.3	0.7	1.9	1.4	0.6	0.5	0.6	0.9	1.0
muAlphaSFSR0p5 Uncert	6.8	2.1	9.6	9.1	4.0	2.1	3.8	8.4	6.8
muF0p5 Uncert	0.0	0.2	0.2	0.2	0.2	0.1	0.5	0.6	0.1
muR0p5 Uncert	0.1	1.0	0.5	0.9	0.3	0.9	0.2	0.1	0.2
var3cUp Uncert	0.3	0.5	0.4	0.3	0.5	0.4	0.5	0.4	0.8
Total Sys.	10.9	4.3	11.9	11.6	7.7	14.4	8.2	12.6	9.0
Total Uncert.	22.9	19.1	22.0	19.7	20.7	22.7	21.0	21.4	19.4

Table 40: η of the sub-leading additional b-jet: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5	6
Data Stat.	0.0	10.6	11.0	9.7	12.5	13.4
Matrix Uncert	0.0	0.1	1.9	4.9	0.9	6.8
PS Uncert	0.0	2.0	0.4	1.8	0.1	5.1
Rad Uncert	0.0	0.1	4.9	0.3	3.4	1.7
BkgStat Uncert	0.0	0.5	0.7	0.5	0.6	0.9
TrainingStat Uncert	0.0	1.3	1.3	1.3	1.3	1.6
BTag	0.0	3.0	1.0	0.6	1.4	2.9
JVT	0.0	0.3	0.2	0.0	0.2	0.4
Lep	0.0	0.4	0.6	0.4	0.5	0.4
PileUp	0.0	0.6	1.0	0.3	0.3	0.1
muAlphaSFSR0p5 Uncert	0.0	1.6	1.9	2.5	0.7	3.6
muF0p5 Uncert	0.0	0.9	0.2	0.1	0.4	1.1
muR0p5 Uncert	0.0	1.0	0.5	0.2	0.2	1.7
var3cUp Uncert	0.0	0.3	0.5	0.4	0.5	0.8
Total Sys.	0.0	4.5	6.0	6.0	4.2	10.3
Total Uncert.	0.0	11.5	12.5	11.5	13.1	16.9

Table 41: Vector sum of the two additional b-jets p_{T}: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5
Data Stat.	13.2	9.1	9.7	13.7	18.7
Matrix Uncert	4.5	0.2	1.1	5.0	5.5
PS Uncert	0.3	0.2	2.5	0.2	7.8
Rad Uncert	0.2	0.6	0.2	0.8	0.4
BkgStat Uncert	0.8	0.5	0.5	0.7	1.4
TrainingStat Uncert	1.6	1.0	1.1	1.8	2.3
BTag	3.5	1.0	1.3	2.0	3.0
JVT	0.8	0.0	0.2	0.3	0.5
Lep	0.4	0.3	0.4	0.3	0.8
PileUp	0.3	0.2	0.7	0.9	0.2
muAlphaSFSR0p5 Uncert	8.1	5.4	2.1	2.1	8.6
muF0p5 Uncert	0.3	0.4	0.0	0.2	1.6
muR0p5 Uncert	0.2	0.5	0.1	0.4	1.6
var3cUp Uncert	0.7	0.3	0.0	0.1	0.3
Total Sys.	10.2	5.7	3.9	6.3	13.7
Total Uncert.	16.7	10.8	10.5	15.0	23.2

Table 42: p_{T} of the leading b-jet from top: the uncertainties per category and the total uncertainty (in percentage).

4j4b channel:

Uncert. (%)	1	2	3	4
Data Stat.	7.4	5.8	10.4	20.0
Matrix Uncert	2.0	0.7	3.2	2.0
PS Uncert	2.7	0.5	0.4	5.1
Rad Uncert	3.9	1.4	0.4	5.9
BkgStat Uncert	0.4	0.4	0.5	1.0
TrainingStat Uncert	1.0	0.8	1.4	2.0
BTag	3.4	1.1	2.3	3.3
JVT	0.3	0.2	0.4	0.6
Lep	0.5	0.2	0.7	0.6
PileUp	0.6	0.3	0.8	0.8
muAlphaSFSR0p5 Uncert	2.8	3.4	3.9	10.4
muF0p5 Uncert	0.1	0.1	0.1	0.8
muR0p5 Uncert	0.5	0.1	0.5	1.0
var3cUp Uncert	0.5	0.0	0.4	0.4
Total Sys.	7.0	4.1	5.9	13.9
Total Uncert.	10.1	7.1	11.9	24.3

Table 43: p_{T} of the sub-leading b-jet from top: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5	6	7	8	9
Data Stat.	22.5	19.9	13.7	18.1	16.7	23.5	16.3	15.4	20.0
Matrix Uncert	7.7	4.6	5.0	2.9	2.9	0.4	6.2	5.5	6.7
PS Uncert	4.5	1.7	2.8	3.7	2.0	2.0	2.4	1.7	5.4
Rad Uncert	1.4	8.3	0.9	0.8	9.4	2.1	2.4	5.0	2.5
BkgStat Uncert	1.6	0.7	0.9	1.0	0.8	1.3	1.0	1.0	1.1
TrainingStat Uncert	2.6	2.2	1.9	2.3	1.8	2.8	2.1	2.2	2.4
BTag	0.4	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4
JVT	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1
Lep	0.9	0.9	0.4	0.8	0.7	0.8	0.5	0.8	1.5
PileUp	1.0	0.2	1.4	1.1	1.0	1.4	1.5	0.5	1.0
muAlphaSFSR0p5 Uncert	3.2	7.0	10.5	9.3	5.5	18.7	6.1	10.7	3.2
muF0p5 Uncert	0.5	0.6	0.2	0.8	0.3	0.4	0.1	0.1	0.6
muR0p5 Uncert	0.0	0.4	0.2	0.4	0.1	0.1	0.4	0.6	0.6
var3cUp Uncert	0.6	0.2	1.0	0.8	0.2	0.4	0.1	0.9	0.5
Total Sys.	10.2	12.3	12.3	10.9	11.7	19.2	9.7	13.4	10.1
Total Uncert.	24.7	23.4	18.4	21.1	20.4	30.4	19.0	20.4	22.4

Table 44: η of the leading b-jet from top: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5	6	7	8
Data Stat.	22.8	15.0	15.5	16.0	15.1	18.5	19.6	17.3
Matrix Uncert	0.1	1.6	0.6	1.2	4.2	7.0	0.4	4.7
PS Uncert	4.0	0.8	0.7	1.2	2.8	5.5	0.1	4.7
Rad Uncert	4.1	1.0	0.9	10.5	6.8	1.3	2.5	1.4
BkgStat Uncert	1.5	0.8	0.7	0.9	0.9	1.0	0.9	1.1
TrainingStat Uncert	2.7	1.8	2.0	2.0	2.1	2.0	2.1	2.3
BTag	0.7	0.3	0.2	0.4	0.5	0.3	0.1	0.6
Lep	0.9	0.6	0.5	1.2	1.1	0.6	0.5	1.2
PileUp	1.3	0.6	1.6	0.5	1.4	0.6	0.3	2.4
muAlphaSFSR0p5 Uncert	10.5	1.5	8.6	4.0	9.8	4.2	5.5	3.3
muF0p5 Uncert	0.3	0.3	0.3	0.4	0.3	0.3	0.2	0.4
muR0p5 Uncert	1.1	0.1	0.1	0.8	0.7	0.4	0.4	0.6
var3cUp Uncert	0.8	0.8	0.4	0.4	0.2	0.5	0.4	0.4
Total Sys.	12.6	3.4	9.1	11.7	13.3	10.3	6.5	8.5
Total Uncert.	26.1	15.4	18.0	19.8	20.1	21.2	20.7	19.2

Table 45: η of the sub-leading b-jet from top: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5	6
Data Stat.	0.0	17.0	9.1	9.3	9.8	14.8
Matrix Uncert	0.0	2.0	7.3	2.0	4.8	8.9
PS Uncert	0.0	5.8	4.1	0.7	6.3	5.7
Rad Uncert	0.0	7.6	2.2	0.3	2.7	1.0
BkgStat Uncert	0.0	0.9	0.5	0.5	0.6	0.9
TrainingStat Uncert	0.0	2.0	1.1	1.1	1.3	1.7
BTag	0.0	1.6	1.3	0.5	1.0	2.7
JVT	0.0	0.2	0.2	0.1	0.1	0.5
Lep	0.0	0.7	0.3	0.5	0.5	0.5
PileUp	0.0	2.7	0.6	0.5	1.1	0.6
muAlphaSFSR0p5 Uncert	0.0	9.9	3.3	1.6	0.6	0.7
muF0p5 Uncert	0.0	0.7	0.3	0.2	0.4	1.8
muR0p5 Uncert	0.0	1.4	0.6	0.1	0.0	2.3
var3cUp Uncert	0.0	0.5	0.0	0.2	0.2	0.1
Total Sys.	0.0	14.5	9.5	3.1	8.7	11.6
Total Uncert.	0.0	22.3	13.1	9.8	13.1	18.8

Table 46: Vector sum of the b-jets from top p_{T}: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5
Data Stat.	14.9	11.2	10.8	9.0	15.1
Matrix Uncert	4.0	3.8	5.2	2.3	3.9
PS Uncert	8.9	0.8	1.9	1.9	2.2
Rad Uncert	4.1	1.2	3.5	0.9	2.4
BkgStat Uncert	0.7	0.6	0.6	0.6	0.9
TrainingStat Uncert	1.9	1.3	1.5	1.3	1.8
BTag	5.3	1.3	1.0	1.6	2.8
JVT	0.6	0.3	0.1	0.3	0.4
Lep	0.8	0.8	0.3	0.5	0.6
PileUp	0.5	0.9	0.2	0.3	0.6
muAlphaSFSR0p5 Uncert	3.0	3.0	1.5	1.1	3.2
muF0p5 Uncert	1.0	0.5	0.1	0.2	1.1
muR0p5 Uncert	0.8	0.5	0.4	0.2	2.0
var3cUp Uncert	0.6	0.4	0.0	0.3	0.6
Total Sys.	12.5	5.6	7.0	4.0	7.4
Total Uncert.	19.5	12.5	12.9	9.9	16.8

Table 47: p_T of the leading additional b-jet: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3
Data Stat.	6.0	8.3	10.4
Matrix Uncert	0.1	0.5	0.4
PS Uncert	0.6	0.4	0.6
Rad Uncert	2.5	0.4	4.6
BkgStat Uncert	0.3	0.5	0.7
TrainingStat Uncert	0.8	1.0	1.3
BTag	3.2	1.9	3.4
JVT	0.2	0.1	0.5
Lep	0.3	0.5	0.2
PileUp	0.3	0.6	0.6
muAlphaSFSR0p5 Uncert	1.0	1.3	0.3
muF0p5 Uncert	0.4	0.0	0.7
muR0p5 Uncert	0.4	0.2	0.8
var3cUp Uncert	0.0	0.2	0.2
Total Sys.	4.3	2.8	6.1
Total Uncert.	7.4	8.8	12.0

Table 48: p_T of the sub-leading additional b-jet: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5	6	7	8
Data Stat.	13.0	17.3	20.9	17.1	20.5	17.4	18.1	12.1
Matrix Uncert	0.3	4.6	0.9	1.8	3.2	7.2	0.4	4.8
PS Uncert	3.6	1.9	2.1	3.6	1.2	8.6	5.6	7.0
Rad Uncert	1.3	3.6	2.5	1.6	5.0	1.3	10.6	6.4
BkgStat Uncert	0.6	0.9	1.0	1.4	1.0	1.1	0.8	1.0
TrainingStat Uncert	1.7	2.2	2.4	2.3	2.1	2.2	2.1	2.0
BTag	0.5	0.2	0.3	0.6	0.5	0.5	0.4	0.9
Lep	0.4	0.5	0.5	0.8	0.8	0.8	0.5	0.5
PileUp	1.2	1.2	2.8	0.8	1.1	0.4	1.7	0.5
muAlphaSFSR0p5 Uncert	8.4	5.5	7.1	1.4	1.5	0.9	7.1	2.9
muF0p5 Uncert	0.3	0.4	0.4	0.1	0.4	0.1	0.3	0.3
muR0p5 Uncert	0.6	0.2	0.1	0.1	0.0	0.0	0.8	0.0
var3cUp Uncert	0.0	0.2	0.3	0.1	0.2	0.4	0.4	0.1
Total Sys.	9.6	8.7	8.7	5.5	6.8	11.6	14.2	11.2
Total Uncert.	16.1	19.4	22.6	17.9	21.6	21.0	23.0	16.5

Table 49: η of the leading additional b-jet: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5	6	7	8	9
Data Stat.	20.2	18.6	18.5	15.9	19.2	17.6	19.3	17.3	17.2
Matrix Uncert	6.8	0.5	2.2	0.0	4.2	1.2	1.9	8.1	3.0
PS Uncert	2.6	2.2	4.5	6.4	3.7	4.1	2.8	3.8	4.2
Rad Uncert	3.4	1.0	3.6	1.2	1.6	13.3	5.7	1.6	0.0
BkgStat Uncert	1.0	1.0	1.4	0.7	1.2	0.9	1.3	1.0	0.9
TrainingStat Uncert	2.2	2.2	2.3	2.1	2.5	2.2	2.4	2.1	2.0
BTag	0.5	0.1	0.4	0.2	0.4	0.2	0.3	0.2	0.3
Lep	0.7	0.7	0.9	1.1	0.9	0.9	0.5	0.4	1.0
PileUp	1.3	0.7	1.9	1.4	0.6	0.5	0.6	0.9	1.0
muAlphaSFSR0p5 Uncert	6.8	2.1	9.6	9.1	4.0	2.1	3.8	8.4	6.8
muF0p5 Uncert	0.0	0.2	0.2	0.2	0.2	0.1	0.5	0.6	0.1
muR0p5 Uncert	0.1	1.0	0.5	0.9	0.3	0.9	0.2	0.1	0.2
var3cUp Uncert	0.3	0.5	0.4	0.3	0.5	0.4	0.5	0.4	0.8
Total Sys.	10.9	4.3	11.9	11.6	7.7	14.4	8.2	12.6	9.0
Total Uncert.	22.9	19.1	22.0	19.7	20.7	22.7	21.0	21.4	19.4

Table 50: η of the sub-leading additional b-jet: the uncertainties per category and the total uncertainty (in percentage).

Uncert. (%)	1	2	3	4	5	6
Data Stat.	0.0	10.6	11.0	9.7	12.5	13.4
Matrix Uncert	0.0	0.1	1.9	4.9	0.9	6.8
PS Uncert	0.0	2.0	0.4	1.8	0.1	5.1
Rad Uncert	0.0	0.1	4.9	0.3	3.4	1.7
BkgStat Uncert	0.0	0.5	0.7	0.5	0.6	0.9
TrainingStat Uncert	0.0	1.3	1.3	1.3	1.3	1.6
BTag	0.0	3.0	1.0	0.6	1.4	2.9
JVT	0.0	0.3	0.2	0.0	0.2	0.4
Lep	0.0	0.4	0.6	0.4	0.5	0.4
PileUp	0.0	0.6	1.0	0.3	0.3	0.1
muAlphaSFSR0p5 Uncert	0.0	1.6	1.9	2.5	0.7	3.6
muF0p5 Uncert	0.0	0.9	0.2	0.1	0.4	1.1
muR0p5 Uncert	0.0	1.0	0.5	0.2	0.2	1.7
var3cUp Uncert	0.0	0.3	0.5	0.4	0.5	0.8
Total Sys.	0.0	4.5	6.0	6.0	4.2	10.3
Total Uncert.	0.0	11.5	12.5	11.5	13.1	16.9

Table 51: Vector sum of the two leading additional b-jets p_T: the uncertainties per category and the total uncertainty (in percentage). Note that the first bin is of width of 0.1 GeV $(0 - 0.1$ GeV), and serves for debugging purposes. Non events are expected in the bin, therefore systematic uncertainties are 0.

VITA

Egor Antipov

Candidate for the Degree of

Doctor of Philosophy

Dissertation: SEARCHES FOR NEW PHYSICS WITH TOP QUARKS IN THE FINAL STATE AND MEASUREMENT OF THE DIFFERENTIAL CROSS-SECTION OF $t\bar{t}$ PRODUCTION IN ASSOCIATION WITH ADDITIONAL HEAVY FLAVOR JETS

Major Field: Physics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Physics at Oklahoma State University, Stillwater, Oklahoma in May, 2022.

Completed the requirements for the Bachelor of Engineering Physics in Engineering at National Nuclear Research University, Moscow, Russia in 2016.

Professional Membership:

American Physical Society