Search for dark matter produced in association with a Higgs boson decaying to two bottom quarks in $p-p$ collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector

Teoh Jia Jian

Department of Physics
Osaka University

This dissertation is submitted in partial
fulfillment of the requirements for the degree of
Doctor of Philosophy

August 2017

Abstract

A search for dark matter produced in association with a SM Higgs boson which decays to a pair of bottom quarks using $p-p$ collisions at a center-of-mass energy of 8 TeV is presented. The dataset collected by the ATLAS detector at the LHC corresponds to an integrated luminosity of 20.3 fb^{-1}. The observed data are found to be consistent with the expected Standard Model backgrounds. Exclusion limits are presented for the mass scales of various effective field theory operators that describe the interaction between dark matter particles and the Higgs boson. Model-independent upper limits are also placed on the visible cross-sections for $H(\rightarrow b\bar{b}) + E_{\mathrm{T}}^{\mathrm{miss}}$ events with $E_{\mathrm{T}}^{\mathrm{miss}}$ above 300 GeV and 400 GeV.

Acknowledgements

The completion of this doctoral dissertation was not possible without the support of many people, to whom I must express my sincere gratitude. First and foremost I want to thank my advisors: Kazunori Hanagaki and Taku Yamanaka, without whom I would not have reached this far! They have both consciously and unconsciously taught me how a good experimental physicist should behave. I am indebted to their sacrifices in terms of time as well as their contributions of ideas and funding to make my Ph.D. experience productive and stimulating. I am grateful for the trust and faith they have given me such that I am able to freely conduct and enjoy my research.

My sincere thanks also goes to Shih-Cheih Hsu of the University of Washington, Seattle, who co-supervised me at various phases of this research. How could I forget those gruelling yet rewarding days and nights we spent at CERN to formulate analysis strategies and to perform checks after checks when our results were not making sense. Thanks for transmitting your enthusiasm and knowledge on various subjects to me.

Special thanks to Jike Wang, Nikola Whallon, Yangyang Cheng, and Lynn Marx for their collaboration to make the analysis a reality. To Jike, thanks for teaching me and taking care of the limits calculation. To Nikola, thanks for sharing various tasks in MC sample production, results interpretation and so on. To Yangyang and Lynn, thanks for being so patient and accommodating!

A big thanks also to all the members of the Editorial Board of our analysis, especially Bill Murray and Michael Kagan. Thanks for their incredible stewardship and supports from the very beginning until our paper was published.

My colleagues, Minoru Hirose, Yohei Yamaguchi, Naoki Ishijima, Satoshi Higashino, Kazuki Yajima, and Ryoji Tsuji have all extended their support in a very special way. I gained a lot from them via personal and scholarly interactions. Thank you all.

Last but not least, I would also like to thank the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of the Japanese Government in providing the indispensable financial support. To my parents and friends, for their love, understanding as well as supports and for always being there, especially during the tough times in my Ph.D. pursuit. To everyone who I did not mention but have certainly played a role in my pursuit, thank you. I dedicate this work to all of you.

Cheers!

Table of contents

List of figures

List of tables

Chapter 1

Introduction

The discovery of Higgs at the Large Hadron Collider (LHC) completed the particles predicted in the Standard Model (SM) of elementary particle physics. The SM describes almost everything we see in the laboratory. It is both remarkably simple and very powerful, but it is not an entirely satisfactory theory. The SM is certainly not the end of the story, no one thinks it is.

So what is wrong with the SM? There are fundamental physical phenomena in nature that the SM does not adequately explain. For example, among the open questions one could ask are the following: What causes the difference between matter and antimatter, and is it related to the origin of the matter in the Universe? Why is there a huge difference in the strength of fundamental forces? How does one quantise gravity and how does one unify the fundamental interactions? What is the nature of dark energy and dark matter (DM)?

While all of the questions mentioned above are equally important, this thesis will be focusing on the last, i.e the search for DM - a hypothetical type of matter that is non-luminous and cannot be seen directly in our present observations, but influencing the evolution of the universe only through its gravitational effect. In fact, the latest result from an analysis of Planck's full data [1] indicates that the total mass-energy of the universe consists of 68.3% dark energy and 4.9% ordinary matter. The remaining 26.8% which is about five times as much as the ordinary matter is consisted of DM. Plenty of evidences for the existence of DM have been accumulated from the last century to now. Some of these evidences will be described in more detail in Section 1.1.

The existence of ubiquitous DM has been accepted for years but what constitutes DM is still an open question. A myriad of CDM candidates have been suggested. In this thesis, we focus on one particular candidate, that is the WIMP. More details will be considered in Section 1.2. Various experiments searching for DM have established increasingly strong constraints on the DM mass and interaction cross sections. Section 1.3 attempts to give an overview to some of the recent DM searches and to summarise their status.

We formally state the purpose and the motivation of our works in Section 1.4. In order to perform a mono-Higgs search, first we need know what a DM signal at the LHC will look

like. In doing so, one needs to construct a DM model. For we are interested in conducting a model independent search, we adopted the effective field theory (EFT) formalism, which is discussed in Section 1.5. This is followed by Section 1.6 where we enumerate the possible EFT operators for generating mono-Higgs signatures at the LHC. Next, Section 1.7 will summarised the regions of valid parameter spaces for each of the EFT operator under consideration. Next, Section 1.8 presents a quick overview of the analysis strategy. Finally a brief overview of the thesis will be given in Section 1.9.

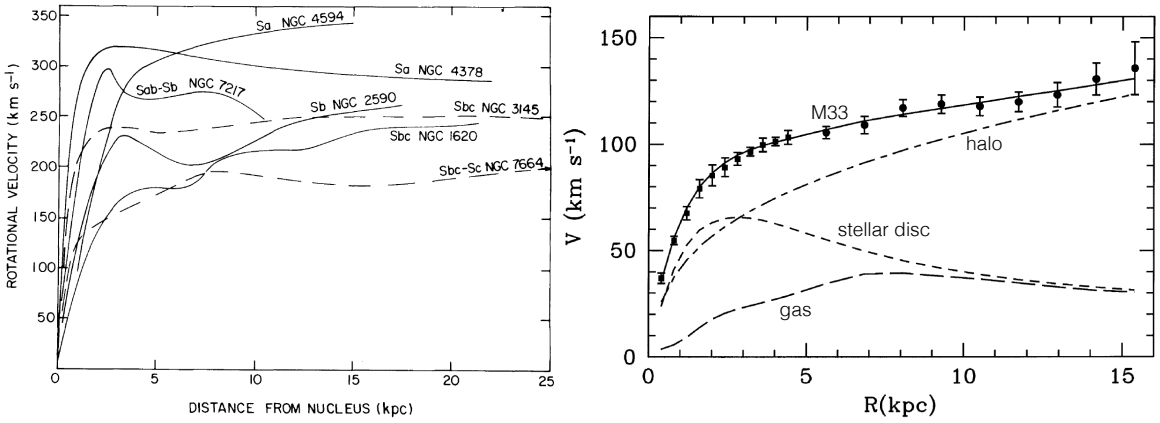
1.1 Evidences of the existence of dark matter

The existence of DM is indirectly inferred from its influence on the evolution of the universe through its gravitation effect. Fritz Zwicky estimated the mass-to-light ratio[1] in the Coma cluster to be $\approx 400 h_o \Upsilon_\odot$ [2], where h_o is the dimensionless Hubble parameter. Compared to the typical value of Υ in the bright central part of galaxies ($\simeq 10-20 h_o \Upsilon_\odot$ [3]), the measured value shows that the total mass of the Coma cluster is at least ten times larger than can be explained by the starlight seen in all the member galaxies. Although initially met with skepticism, the similar observations from the Local Group [4] and the Virgo cluster [5] provide further confirmation to Zwicky's results. That said, a more widely recognised observational evidence in favour of DM in galaxies is the observations of the absence of the "Keplerian fall[2]" in the rotational curves of spiral galaxies. Contrary to the expectation, Rubin et. al. [6–8] and Roberts et. al. [9] in the 80s observed a rather flat rotation curve (see Figure 1.1, suggesting that either Newtonian gravity does not apply universally or that galaxy masses grow approximately linearly with radius well beyond the galactic bulge. The other observations (e.g. [10]) showed that the flat behaviour continued beyond several optical radii. The pioneering theoretical studies such as [11, 12] suggested the presence of a massive halo of DM around galaxies. This DM halo model was further strengthen by the recent observations [13–15].

Another evidence of the existence of DM comes from the study of X-ray emission from hot gas in large elliptical galaxies or intergalactic space in clusters. By comparing the X-ray mass profile (obtained by measuring of the X-ray temperature of the hot gas) with the visible mass profile estimated from the other methods (e.g measurement of the luminosity or optical spectrum) we could obtain the constraints on the contribution of the non-visible mass. Also, hot gas is held in the cluster by gravity. If the mass of the galaxy is not enough to explain the

[1] mass-to-light ratio, Υ, typically expressed in units of solar mass to solar luminosity ($\Upsilon_\odot = M_\odot/L_\odot$), is a good indicator of the presence of DM in a given region. Value of mass-to-light ratio that is greater than unity indicates that most of the matter in these objects does not reside within stars but is present in some form of non-luminous matter.

[2] Under the so-called Keplerian approximation, a test particle is assumed to be so far away as to 'feel' the galaxy as a point-mass. If a system is in virial equilibrium its circular velocity is given by $V^2(R) = AGM(R)/R$, where A is a function of the eccentricity and M is the mass inside the homogeneous spheroid with radius R. Since $M(R) \propto \rho R^3$, we have $V(R) \propto \sqrt{\rho} R$ for $R \leq R_0$ and $V^2(R) \propto M(R_0)/R$ for $R > R_0$. The circular velocity has a Keplerian fall exterior to the mass distribution.



(a) (b)

Fig. 1.1 (a) Rotation velocities of seven galaxies as a function of distance from the galaxy nucleus [7]. The general flatness of the curves are notable. The Keplerian fall is not observed.(b) M33 rotation curve compared with the best-fitting model (continuous line). Also shown are the halo, the stellar disc and the gas contribution (modified from [13]).

presence of this gas, that implies huge amounts of additional, invisible matter are needed for gravity to balance the pressure of the gas. In fact the data from Einstein Observatory [16] have shown that the mass of M87 elliptical galaxy is much larger than the visible mass. Recent data from the NASA's Chandra X-ray observatory [17, 18] has allowed astronomers to tightly constrain the distribution of DM content, which increases smoothly all the way into the central galaxy of the cluster.

Gravitational lensing - the bending of light-rays passing through a gravitational field, can be used to determine the mass of a massive object (such as a cluster of galaxies) or characterise its mean distribution. This technique has the advantage that it is a purely geometrical effect that is free from the astrophysical assumption and can be applied to all matter regardless of their composition and their dynamical states. In particular, the marriage of the X-ray observation and the weak gravitational lensing[1] allows one to use cluster collisions to get important information on the nature of DM. In most regions of the universe ordinary matter and DM are bound together by gravitational attraction, as shown by the large-scale DM distribution survey [19]. Under certain circumstances, such as colliding cluster of galaxies, ordinary matter and DM may be separated. This is exactly what has been observed by Clowe and his colleagues [20] in the Bullet Cluster. The hot plasma cloud which represent most of the ordinary matter in the cluster pair was shocked and decelerated due to electromagnetic interaction between gas

[1] Gravitational lensing can be divided in three regimes: strong lensing, weak lensing, and micro lensing. In strong lensing, the observer sees multiple images or ring-liked structures of the source. Instead, in weak lensing only minute distortion of the background sources are observed. Lastly if the mass is very small, one only observes micro-lensing which involves a magnification of the brightness of the lensed object.

particles. Its distribution from X-rays data are shown in Figure 1.2 as colour map with whiter regions corresponding to higher matter concentration. Contours superimposed over both image are the mass density contours reconstructed from the weak-lensing data. The inner 3 contours (correspond to 68.3%, 95.5%, and 99.7% confidence levels) show the errors on the positions of the mass density peaks. Two major peaks are clearly visible in the reconstruction. Both peaks are offset from the centre of mass of their respective plasma clouds. If models without DM (such as Modified Newtonian Dynamics) were right, the mass density peaks should be near the centre of mass of their respective plasma cloud. Instead, Clowe's team observed an 8 σ spatial offset of the centre of the total mass from the centre of the plasma mass peaks. This result shows that most of the mass in the cluster pair is in the form of DM, which bypassed the gas regions during the collision (since DM is only weakly interacting, other than via the gravitational force). Their observation provides the best evidence to date for the existence of DM and effectively rules out modifications of Newtonian gravity as the explanation of DM.

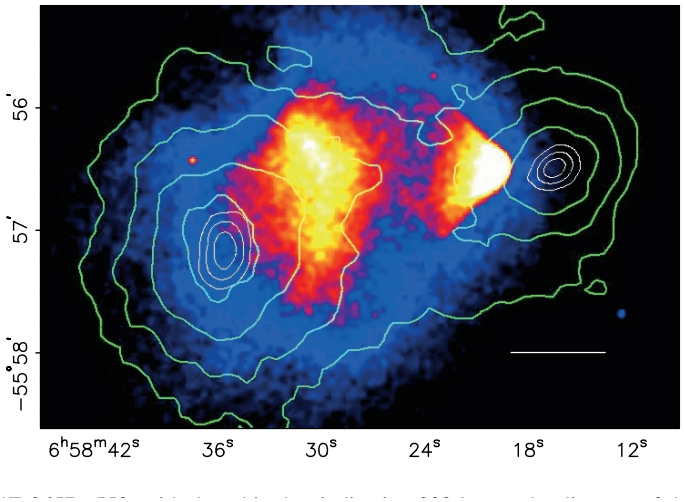


Fig. 1.2 X-ray image from Chandra X-ray Observatory that traces hot plasma of two colliding clusters of galaxies, called Bullet Cluster, are in the process of moving through each other. The spatial offset between the observed mass density peaks and the plasma cloud mass density peaks suggest that most of the mass in the cluster pair is in the form of dark matter. Figures taken from [20].

Another cosmological evidence in favour of DM is from the fluctuations in the Cosmic Microwave Background (CMB) - afterglow radiation left over from the hot Big Bang. The small scale fluctuations can be attributed to the so-called acoustic oscillations. Before the recombination, the matter was distributed almost uniformly in space, with only small variations due to quantum fluctuation. At denser regions, the force of gravity caused matter to fall inward while at the same time outward pressure was exerted by photons due to Thomson scattering with free charge particles. These competing forces caused matter to oscillate in-and-out of the dense regions. In turn, this oscillation caused the matter to heat up when it fell in and to cool off when it flowed out. This process repeats until the recombination and the photon

decoupled from the matter. Depending on the location in the cycle, the photons in thermal equilibrium with the matter emerged vary in temperature. So the temperature variations in the CMB can give us information on the initial density perturbations and information on the amount of different types of matter at the time of recombination.

The temperature fluctuation can be presented in the form of a power spectrum. Typical information that can be obtained from this power spectrum of the CMB includes baryon density, matter density, dark energy density and so on. These parameters are obtained by fitting the observed data with a cosmological model (e.g. ΛCDM) model). The Wilkinson Microwave Anisotropy Probe (WMAP) [21, 22] and the Planck spacecraft [1] provided the most detailed measurements of CMB power spectrum and showed that the existence of DM is favoured. By combining different techniques, such as the Baryonic Acoustic Oscillations [23], CMB, Big Bang Nucleosynthesis(BBN) [24], supernovae [25], and the structure formation [26, 27], better constraints to the dark energy density as well as the matter density, and thus to the DM content of the universe can be obtained.

In conclusion, the evidence for dark matter is very compelling. Having established the need for dark matter, in the next two sections we will discuss possible particle candidates for DM and its properties.

1.2 WIMPs as dark matter

According to the Lambda-Cold Dark Matter (ΛCDM) model, a DM candidates must have the following characteristics:

1. It must be non-relativistic (cold) at freeze-out or else the structure formation would fail. As explained previously, CMB anisotropy is extremely smooth. Particles moving at relativistic speed cannot effectively clump into gravitational wells needed for the initial structure formation.

2. It should interact very weakly with electromagnetic radiation. If DM is not electrically neutral, it will scatter light and thus would not qualify as dark matter.

3. It must be stable (at least with the lifetime exceeding the age of the universe) due to the stability of the halo.

4. Its abundance must be compatible with the observations over a wide range of scales (galactic, galaxy cluster, and cosmological) and experimental techniques.

There is in fact no shortage of candidates for explaining the nature of the DM described previously. While some of these DM candidates are created specifically to account for the missing DM, others emerge quite naturally from the solutions to the problems in the SM. A few of the more popular candidates include axions, sterile neutrinos, and WIMPs. Out of all

the possible DM candidates, we have chosen to focus on WIMPs. WIMPs are not a specific elementary particle, but rather an entire class of possible new fundamental particles. There exists no clear definition of a WIMP, but broadly a WIMP DM candidate must satisfy the criteria mentioned in the previous paragraph. Their mass can lie in the range of a few GeV to 100 TeV.

WIMP looks particularly attractive from the point of view of particle physics and for this analysis. This is because in some scenarios WIMP arises naturally as a by-product with no extra restrictions on the theory to account for DM. Two examples of such scenario are the Supersymmetry (SUSY) and Universal Extra Dimension (UED) models. The conservation of discrete symmetry imposed in such models ensure that the corresponding lightest particle predicted in each model is stable against decay and constitutes a candidate for DM. For example, one of the most studied supersymmetric DM candidate is the lightest neutralino ($\tilde{\chi}_1^0$), which arises naturally in the R-parity conserving models in the Minimal Supersymmetric Standard Model via direct ($\tilde{q} \to q\tilde{\chi}_1^0$) or one-step ($\tilde{q} \to qh\tilde{\chi}_1^0$) decay of squark as shown in Figure 1.3.

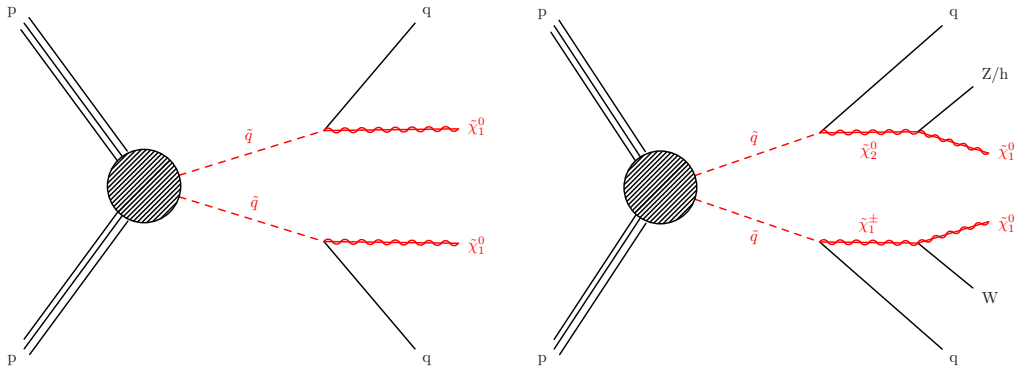


Fig. 1.3 The decay topologies of squark-pair production with (a) direct or (b) one-step decays of squarks to the lightest neutralino.

What is more, as we shall see later, the interest in WIMPs as the most well-motivated DM candidates stems from the fact that WIMPs in thermal equilibrium in the early universe naturally have the right abundance to be DM. The self-annihilation cross section that gives the right WIMP relic density is of the order of weak interaction. This makes the detection of WIMPs possible. The latter aspect is important as it provides a means to test the WIMP hypothesis. Hence, we will restrict our discussion of DM candidates to WIMP.

WIMP relic abundance

The standard well accepted scenario is that WIMP is thermally produced, or in other words, WIMP was in thermal equilibrium with the primordial plasma before freezing out and leaving a relic density. The relics produced this way will be called thermal relics and its final abundance

is set by the standard freeze-out mechanism for large annihilation rates, irrespective of the initial conditions.

During the radiation-dominated epoch and before the big bang nucleosynthesis, the universe was densely filled with WIMP, photons as well as free protons, neutrons and electrons. The temperatures remain too high for the binding of electrons to nuclei. This cosmic soup was in thermal equilibrium, i.e. the annihilation and production rate of each particle species is equal to each other. The non-relativistic WIMP number density n_χ as a function of time, t, can quite adequately be approximated by the Boltzmann equation[1]. Assuming a generic WIMP scenario, where two WIMP particles, χ can annihilate to form two light particles (also in thermal equilibrium) according to the process $\chi\bar{\chi} \leftrightarrow Y\bar{Y}$. The Boltzmann equation can be written in the form of:

$$\frac{dn_\chi}{dt} = -3Hn_\chi - \langle\sigma v\rangle[(n_\chi^2) - (n_\chi^{eq})^2] \tag{1.1}$$

where $\langle\sigma v\rangle$ is the thermal averaged cross section times the relative velocity of the interacting particles and H is the Hubble constant. n_χ^{eq} is the WIMP equilibrium number density and for the non-relativistic WIMP we are considering, n_χ^{eq} can be written as [28, 29]:

$$n_\chi^{eq} = g_* \int \frac{d^3\mathbf{p}}{(2\pi)^3} e^{-\sqrt{|\mathbf{p}|^2+m_\chi^2}/T} = \begin{cases} g_*(\frac{m_\chi T}{2\pi})^{3/2} e^{-m_\chi/T} & \text{if } m_\chi \gg T \\ g_* \frac{T^3}{\pi^2} & \text{if } m_\chi \ll T \end{cases} \tag{1.2}$$

where g_* is the number of internal degrees of freedom of χ (e.g., spin or color). There is then only one unknown, n_χ, the abundance of the WIMP.

Equation 1.1 simply states that the rate of change in the abundance of a WIMP is the difference between the rates for producing and eliminating WIMP particle, taking into account the dilution that comes from the Hubble expansion. In thermal equilibrium the number density tracks the equilibrium number density as in Equation 1.2. The universe cools as it expands. When the temperature dropped below WIMP mass its number density would be suppressed by $e^{-m_\chi/T}$. In fact if the equilibrium were kept until today, there would be no WIMP left. Fortunately, the density of WIMP decreases as the universe expands. However, at some point known as the freeze-out point at the temperature, T_{fr}, the expansion rate outstrips the annihilation rate and the Hubble term becomes dominant. At that point, the WIMP dropped out of equilibrium and its abundance freeze out to a constant value that became the present-day relic density.

Starting at an early time when all particles were in equilibrium, one can integrate Equation 1.1 and obtains the number density at $t = 0$ (present time). Once we obtained the current number density, the relic abundance is simply $\Omega_\chi = \rho_\chi/\rho_{cr}$ where current mass density $\rho_\chi = m_\chi n_\chi$ and the critical density of the Universe, $\rho_{cr} = 3H_0^2/8\pi$. An approximate solution

[1]The distinction between Bose, Fermi and Boltzmann statistics is negligible by the time the WIMP species freezes out, i.e. the temperature of the plasma approaches the particle's mass.

of the Boltzmann equation provides the current WIMP relic density as [28]:

$$\Omega_\chi h^2 \simeq \frac{0.1 pb \cdot c}{\langle \sigma v \rangle} \tag{1.3}$$

Now to get the current WIMP relic abundance of ~ 0.1, we need a cross section of order 1 pb, which is about what one would expect from a weakly interacting particle. This strong link between the sizeable interactions between WIMP particles and SM states and the idea of thermal freeze-out is sometimes referred to as the "WIMP miracle". It justifies the excitement for WIMP searches at the LHC. Figure 1.4 illustrates how the freeze-out relic density changes with increasing $\langle \sigma v \rangle$. Particles with larger cross sections freeze out later, and this later freeze-out carries along with it a lower relic abundance.

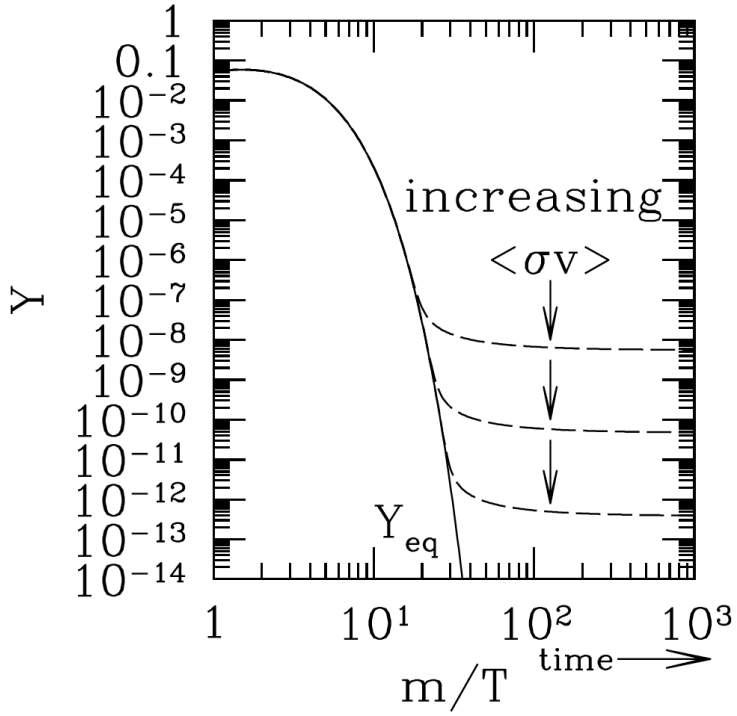
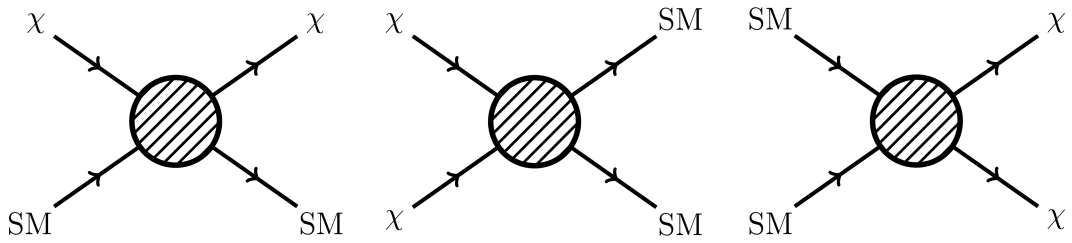


Fig. 1.4 Evolution of a WIMP number density in the early universe. Y is a dimensionless quantities defined as $Y \equiv n_\chi / T^3$. Y tracks its equilibrium value Y_{eq} until $x \sim 10$, and then levels off to a frozen-out constant. WIMP with larger annihilation cross section end up with smaller densities. If Y continue to track its equilibrium value without frozen out, there will be no such particles in our universe today. Figure is taken from Reference [30].

At this point, it is important to point out that starting from this point onward and throughout this thesis, the term "DM" will always refers to WIMP as the candidates unless explicitly specified otherwise.

1.3 Experimental searches for dark matter

In general the search for DM can be broadly divided into three classes depending on the interaction of DM with SM particles, namely direct detection, indirect detection and collider production. All three of them are related and complementary to each other as can be seen from the diagrams in Figure 1.5. The direct detection experiments rely on the detection of signal when DM scatters off an atomic nucleus, while the indirect detection experiments aim to detect annihilation products of DM. On the other hand, the collider experiments look for DM production in high energy particles collision through the telltale sign of missing energy accompanied by a recoiling object. The latter approach has been adopted in this thesis. Below, we will briefly discuss the detection principle and the current status of each approach.



(a) Direct detection (b) Indirect searches (c) Collider production

Fig. 1.5 Three types of generic DM search, (a) direct detection, (b) indirect detection and (c) collider production. They are complementary to each other. The blob at the centre is the mediator that connect DM to SM particles.

1.3.1 Direct detection

In DM Standard Halo Model, it is speculated that our Milky Way Galaxy is embedded in a much larger, roughly spherical halo of DM. Even though the DM are weakly interacting, a small but potentially measurable fraction will elastically scatter off nuclei in a terrestrial detector as they pass through the Earth. The expected interaction rate mainly depends on the DM mass and the DM-nucleus interaction cross section. In general the interaction cross section can be decomposed into spin-independent (SI) and spin-dependent (SD) parts, depending fundamentally on the nature of the couplings in the Lagrangian which describes the DM interaction with quarks. The contributions to the spin-independent part of the cross section originates from scalar and vector couplings to quarks, whereas the contribution to the spin-dependent cross section arises from axial-vector couplings. Typically SD cross sections are smaller than SI cross sections by a factor approximately equal to the square of the mass of the nucleus. As a result, the bounds on the latter are considerably better than those on the former.

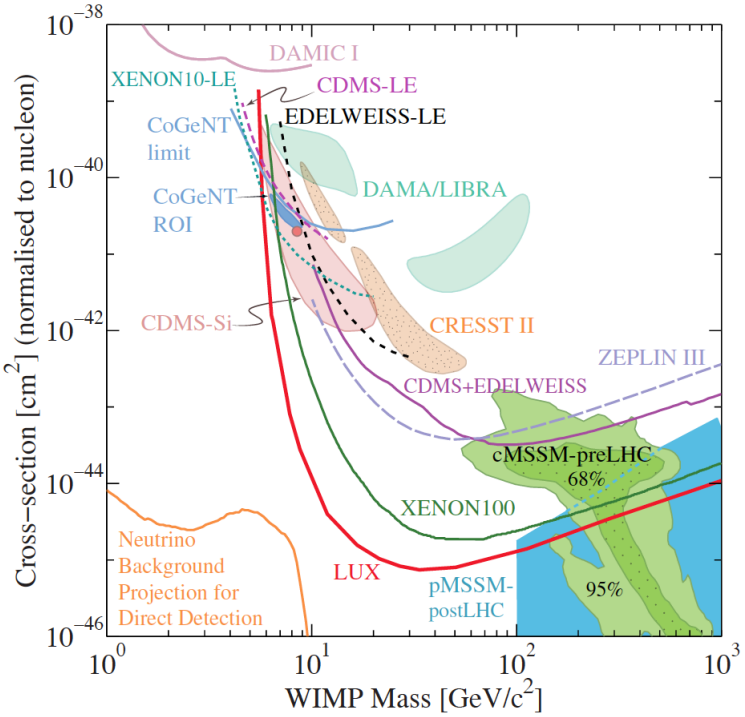


Fig. 1.6 DM cross sections (normalised to a single nucleon) for spin-independent coupling versus mass. Plot taken from [31]. In the orange region on the lower left corner, coherent neutrino scattering becomes an important background [32].

Experimental signatures for direct DM detection are strong daily asymmetry of the nuclear recoil direction and small annual modulation of the recoil rate. The expected recoil rates is typically much smaller than background radiation, ranging from about one event [33] to less than 10^{-3} events per kilogram detector material per year. With the expected DM masses in the range of 10 GeV to 10 TeV, the typical nuclear recoil energy is of order between 1 to 100 keV. Thus, to observe a DM-induced nuclear recoil spectrum, a low energy threshold, an ultra-low background noise and a relatively large target mass are essential. To achieve necessary sensitivity, usually direct detection experiments are located deep underground to shield against cosmic ray backgrounds.

A variety of experiments have been employed to search for the tiny nuclear recoil signal but none of them gave a conclusive confirmation of DM detection. Figure 1.6 illustrate the exclusion limits on DM scattering cross sections, normalised to scattering on a single nucleon, for spin independent couplings as functions of DM mass. The current best limit is set by the Large Underground Xenon dark matter experiment, or LUX [34, 35]. The next generation experiments such as SuperCDMS-SNOLAB, DARk matter WImp search with Nobel liquids (DARWIN) [36], Large Underground Xenon-ZonEd Proportional scintillation in LIquid Noble gases (LUX-ZEPLIN) [37] are expected to push the sensitivity further down and towards lower DM mass region.

1.3.2 Indirect searches

Indirect searches refers to the observation of annihilation or decay products of WIMPs which includes neutrinos, gamma rays, positrons and antiprotons. Since the flux of annihilation products scales with ρ_χ^2, the regions with large DM concentration, such as the core of stars and the galactic centre are of highest interest. One of the potential signatures for indirect detection is mono-energeticc photons from DM annihilation in the halo. Furthermore, excess of antiparticles coming from DM annihilation over the expected antipaticle spectrum also can provide a possible evidence for the presence of DM. Lastly, higher annihilation rate due to trapped DM in the Sun core will give rise to excess of high energy neutrinos over the solar neutrino background.

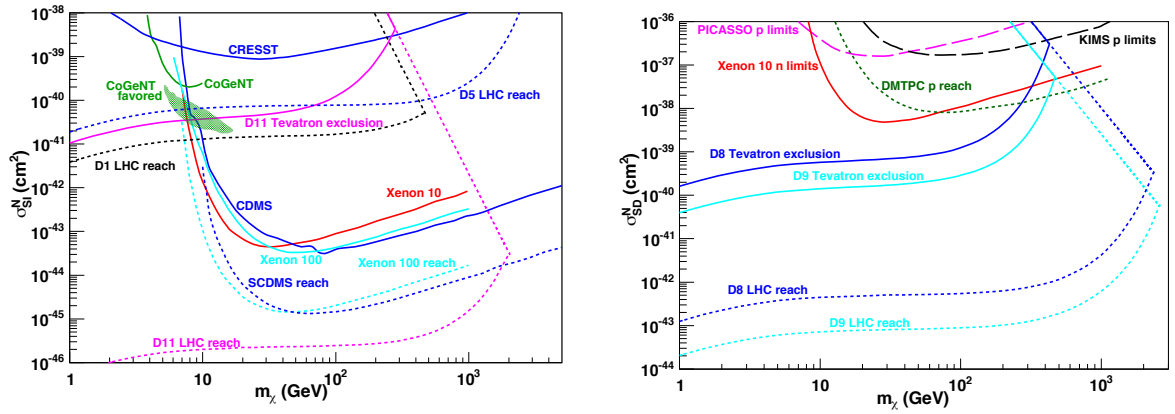
Several existing observational anomalies have caused some excitement. The previous Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics (PAMELA) measurement of the positron fraction between 1.5 and 100 GeV showed the deviation from the background models [38]. The recent measurements [39, 40] also show a rise of the positron fraction between 10 and 200 GeV. The first result of Alpha Magnetic Spectrometer (AMS) [41] showed an unexplained excess of high-energy positrons in Earth-bound cosmic rays. The further measurement of AMS [42] confirms the positron excess. These observations can in principle be explained through the annihilation of DM particles in space, but not yet sufficiently conclusive to rule out other explanation. More data at higher energy will be needed to confirm a DM origin. Another observation from Fermi-Large Area Telescope (FERMI/LAT) showed an excess of mono-energetic photons with energy of 133 GeV. However, with only a local significance of 3.3σ or a global significance of 1.5σ, this result is below the significance required of an unambiguous signal.

1.3.3 Collider production

Another way to search for DM is to produce and detect the signature of DM at particle collider such as the Tevatron and the LHC. One of the main advantages of collider searches is that they do not suffer from astrophysical uncertainties. For example, the direct detection rate depends on the local DM density whose current best estimate suffers from a systematic uncertainty up to 40% [43]. On the other hand, the indirect detection rates rely critically on the distribution of DM and on other astrophysical properties such as the galactic magnetic and radiation fields.

Other important strengths of the collider searches compared to the direct detection experiments can best be illustrated in Figure 1.7, which show the effective collider constraints on the parameter space of DM-nucleon interaction cross section ($\sigma^N_{SI/SD}$) and DM mass (m_χ) for the different operators that couple to SM particles. The most striking feature of the collider constraints is the fact that they provide a strong bound for DM lighter than a few GeV (colliders are able to copiously produce the light DM provided that the production cross section is not very small). This is highly contrast to direct detection constraints which have limited sensitivity

to light DM due to their low energy threshold requirement ($E_{recoill} \sim \mathcal{O}(10\,\text{keV})$). Furthermore if the DM primarily couples to gluons (labeled as D11 in Figure 1.7a), the constraints from colliders become especially strong comparing to any SI search. In addition, unlike the case for direct detection, collider results are insensitive to the the spin content of nucleus (or Lorentz structure of the interaction). This allows the collider searches to have complementary coverage at intermediate masses, as evidenced in Figure 1.7.



(a) Spin-independent limit

(b) Spin-dependent limits

Fig. 1.7 Experimental limits comparison (as of 2011) for (a) SD DM direct detection from CRESST, CDMS, Xenon 10, CoGeNT, Xenon 100 and SCDMS and (b) SI DM direct detection from Picasso, DMTPC, KIMS and Xenon 10. Also shown are the collider constraints from Tevatron and the LHC. Solid lines represent the current result for each experiment while dash lines show the future reach estimates. Plots are taken from Reference [44]. For a complete list of operators coupling DM to SM particles considered in these plots, see Reference [44].

Due to their extremely weak interaction with SM particles, DM will essentially pass through the detector undetected, just like neutrinos. The resulting signature, the missing transverse energy $E_{\text{T}}^{\text{miss}}$ can be used to infer the presence of DM. One approach is to search for the $E_{\text{T}}^{\text{miss}}$ together with multi-parton or multi-lepton final states from a cascade decays of heavier states. Such method is often used in the search for the lightest supersymmetry particles. An alternative approach is to look for pair-produced DM with one recoiling object as trigger for the event. Such event is dubbed mono-object or mono-X event. These searches are characterised by a common strategy that require high $E_{\text{T}}^{\text{miss}}$ and a high quality recoiling object in final state to constitute a clean and distinctive signature. The recoiling object can be a hadronic jet [45] or heavy-flavor jet [46], a single photon [47], or a W or Z boson [48–50], mostly stemming from initial state radiation (ISR). So far, there is no evidence for DM from all these LHC searches. Most of them are limited by statistics and therefore promise gains in sensitivity with increased luminosity.

1.4 Analysis statement and motivation

The discovery of Higgs boson in 2012 has opened up new possibilities in searches for new physics. If DM is indeed associated with the scale of electroweak symmetry breaking, Higgs-related signatures in the LHC are a natural place to search for it. This thesis described a search for DM pair production in association with a Higgs boson, $h\chi\chi$, which is dubbed "mono-Higgs" search. What makes mono-Higgs stand out as being of crucial importance, in contrast to other mono-X processes is that Higgs boson is preferentially emitted as a part of the effective vertex coupling of DM to the SM since Higgs boson coupling to quarks is small. In comparison, a jet, a photon, a W boson or a Z boson may either be emitted as part of the effective vertex coupling or directly from quarks as initial state radiation. This means that by studying mono-Higgs process, we can probe the structure of the effective DM-SM coupling.

At ATLAS, the mono-Higgs search using Higgs to two photons decay channel has been performed [51]. The result is consistent with the SM expectation. We believe that the mono-Higgs channel involves a b-quark pair in the final state is more promising. This is because the Higgs$\rightarrow b\bar{b}$ decay channel gives the best statistics for the signal since it has the largest branching ratio for a Higgs boson with mass m_h = 125 GeV, Br(Higgs$\rightarrow b\bar{b}$) $\approx$ 0.577. Nonetheless, since it has more backgrounds to deal with compared to the two photon channel, a good background rejection is of crucial importance. To achieve this, we capitalise on the event topology where the Higgs boson is produced with sufficiently high transverse momentum that the two b-jets from its decay are very collimated. We can then draw from a number of newly developed novel techniques to effectively identify such Higgs bosons with high transverse momentum.

On the issue of what theoretical framework should be used to model the the production of mono-Higgs event at the LHC, our ignorance of the particle physics nature of DM compels us to take the model independent approach. We want to make as few assumptions as possible. This can be achieved via the effective field theory framework. This approach is desirable from the practical perspective as only a minimal number of new parameters are introduced.

1.5 Effective field theory framework

In the effective field theory (EFT) framework, DM is assumed to be the only accessible new degree of freedom. The coupling of DM to SM particles is parameterised with a set of non-renormalizable (effective) higher-dimensional operators that are generated after integrating out the heavy mediator. The process of integrating out the heavy mediator can be viewed in terms of the expansion of the heavy particle propagator which result in a set of effective Lagrangian with the form of:

$$\mathscr{L}_{eff} = \mathscr{L}_{SM} + \sum_{d>4} \frac{f^{(d)}}{\Lambda^{d-4}} \mathscr{O}_d \tag{1.4}$$

where $\mathscr{L}_{SM}$ is the SM Lagrangian, $f^{(d)}$ is a dimensionless coefficient, $\Lambda = M_\phi$ is the suppression scale and $\mathscr{O}_d$ are operators or interaction terms of mass dimension d that describes contact interactions between SM particles and particles of the extension. The effects of the heavy field in the processes at low momentum transfer, $Q_{tr}^2 \ll \Lambda^2 \sim \Lambda^2$ are encoded by a series of interaction terms which scale as $(Q_{tr}/\Lambda)^{d-4}$. We see that the interaction terms will give smaller contribution to the result when d is larger. The suppression scale Λ gives us an estimate of the importance of the interaction terms. It sets the maximum energy at which the operator expansion in the EFT can be trusted. In other words, as the suppression scale is approached the observables computed within the EFT framework become less accurate as an approximation of the true observables in the unknown UV theory. Generally speaking, the condition of the validity of an EFT approach is that the momentum transfer in the relevant process one wants to describe must be less than the suppression scale:

$$Q_{tr} \lesssim \Lambda \tag{1.5}$$

1.6 Effective operators and LHC observables

Our current goal is to consider the production of Higgs boson at the LHC in association with a pair of DM particles. We need to write down a set of effective operators (i.e expanding the second term in Equation 1.4) that can possibly generate this experimental signature. We focus on operators that give rise to a coupling of DM to h, Z, γ or new mediator particles such as a Z' or scalar singlet S, allowing the production of DM through the process as illustrated in Figure 1.8. Throughout this thesis, we label the DM field as χ. Before moving further, it is best to lay out the grounding assumptions we used in constructing such model that may lead to mono-Higgs signals at the LHC.

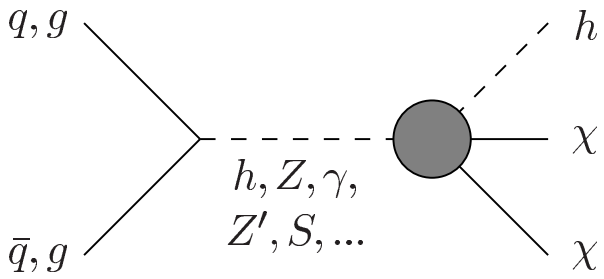


Fig. 1.8 Diagram illustrating an effective interaction (gray circle) that couples χ to the Higgs boson and gives rise to a mono-Higgs signature in a collider experiment. This process can be mediated by electroweak bosons (h, Z, γ) or new mediator particles such as a Z' or scalar singlet, S.

We assume that the interactions between SM hadrons and the particles that constitute cosmological DM exist. In order to describe the interaction between χ and the SM particles in terms of EFT framework, we also assume that χ is the only new degree of freedom beyond the

SM. One may argue that the DM sector may be more complex than a single particle given that it constitutes five times as much as the normal matter. But one must not forget that quite often in the discoveries of SM particles there is only one mediator or search channel that might play the dominant role. Any new particle species in addition to χ shall have a mass much larger than χ.

Next, we also assume that χ is a gauge singlet under $SU(3)_C \times SU(2)_L \times U(1)_Y$. If χ is a gauge singlet, then the gauge invariance implies the operator must also include other electroweak doublets. As obvious as it is, we assume that χ does not interact with the detector and is stable on collider timescales. Stability of this gauge singlet is guaranteed if a continuous or discrete symmetry exists under which these gauge singlets are the lightest particles. In all cases we consider below, we limit our discussion to scalar and fermion DM. Keeping these assumptions in mind, we can now examine dimension by dimension the possible mono-Higgs operators. A more thorough discussion of all the following operators can be found in Reference [52].

Dimension-4 and dimension-5 operators (Higgs portal)

The absolute minimal modification of the SM (under our assumptions) consists of the addition of a singlet scalar DM to those of the SM, using only renormalisable interactions. This singlet scalar can interact with the SM sector through the triple and quartic scalar interactions such as $|H|^2\chi$ and $|H|^2\chi^2$. We focus on the latter. Or more formally, for scalar DM we have a renormalisable interaction at dimension-4:

$$\mathscr{L}_4 = \lambda |H|^2 \chi^2 \tag{1.6}$$

where χ is a real scalar DM and λ is the coupling of the scalar DM to the Higgs boson. H represents the Higgs doublet field. Recall from the considerations that in the unitary gauge there is only one physical Higgs after spontaneous symmetry breaking, H can be conveniently written in the form:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{1.7}$$

where h is the physical Higgs field and $v \approx 246$ GeV is the Higgs vacuum expectation value.

By replacing the gauge singlet scalar with a gauge singlet Dirac fermion, the lowest dimensional operator we may obtain is five. The singlet fermion interacts with the SM sector only through non-renormalisable interactions among which the leading interaction terms are given by the dimension five terms suppressed by a new physics scale, Λ. These dimension-5 operators can be written as

$$\mathscr{L}_5 = \frac{1}{\Lambda}|H|^2 \bar{\chi}\chi + \frac{1}{\Lambda}|H|^2 \bar{\chi} i\gamma_5 \chi \tag{1.8}$$

The interaction described by the first term in Equation 1.8 is essentially ruled out by the XENON100 experiment [53]. Hence, for our study, the dimension-5 operator we consider is the second term of Equation 1.8.

The operators in both Equation 1.6 and Equation 1.8 are usually grouped as the "Higgs-portal" type operators. These operators describe direct couplings between DM particles and the Higgs boson. More detailed studies can be found in References [54–60]. Mono-Higgs signal can arise via $pp \to h^* \to h\chi\chi$ through these operators. Depending on the DM mass, the interactions are different. If DM is light enough, $m_\chi \leq m_h/2$, then these interactions lead to invisible Higgs boson decays. The partial widths calculated at tree level for scalar and fermionic DM can be obtained independently in analytic form [59]:

$$\Gamma(h \to \chi\chi) = \frac{\lambda^2 v^2}{64\pi m_h}\sqrt{1-\frac{4m_\chi^2}{m_h^2}} \qquad \text{scalar } \chi \tag{1.9a}$$

$$\Gamma(h \to \chi\bar{\chi}) = \frac{v^2 m_h}{32\pi\Lambda^2}\left(1-\frac{4m_\chi^2}{m_h^2}\right)^{3/2} \qquad \text{fermion } \chi \tag{1.9b}$$

The requirement that the invisible decay width of the Higgs to satisfy $\mathscr{B}_{h\text{inv}} < 38\%$ obtained in Reference [61] sets a upper (lower) bound on the couplings $\lambda \lesssim 0.016$ ($\Lambda \gtrsim 10$ TeV) for scalar (fermion) DM. Since the couplings are strongly suppressed, the leading mono-Higgs signal is from di-Higgs production where one of the Higgs bosons decay invisibly. On the other hand, for DM with the mass, $m_\chi > m_h/2$ the DM-Higgs couplings can be much larger as the invisible Higgs boson decay is kinematically blocked.

Dimension-6 operators

Next, at dimension-6 there are several operators that give rise to mono-Higgs signals through an effective h-Z-DM coupling. For scalar DM, we have

$$\mathscr{L}_6 = \frac{1}{\Lambda^2}\chi^\dagger i\overleftrightarrow{\partial^\mu}\chi H^\dagger i D_\mu H \tag{1.10}$$

while for fermionic DM we have

$$\mathscr{L}_6 = \frac{1}{\Lambda^2}\bar{\chi}\gamma^\mu\chi H^\dagger i D_\mu H + \frac{1}{\Lambda^2}\bar{\chi}\gamma^\mu\gamma_5\chi H^\dagger i D_\mu H \tag{1.11}$$

where we define a covariant derivative $D_\mu = -i\partial^\mu - (g/2)\sigma^a W_\mu^a - (g'/2)B_\mu$. Note that the Z coupling term arises in the unitary gauge from the covariant derivative. Using

Equation 1.7 and by taking $2m_Z = v\sqrt{g^2+g'^2}$, we can rewrite Equation 1.11 as [62]:

$$\begin{aligned}\mathscr{L}_6 = &\frac{m_Z}{\Lambda^2} hZ_\mu \bar{\chi}\gamma^\mu\chi + \frac{m_Z}{\Lambda^2} hZ_\mu \bar{\chi}\gamma^\mu\gamma_5\chi \\ &+ \frac{m_Z}{2v\Lambda^2} h^2 Z_\mu \bar{\chi}\gamma^\mu\chi + \frac{m_Z v}{2v\Lambda^2} h^2 Z_\mu \bar{\chi}\gamma^\mu\gamma_5\chi \\ &+ \frac{m_Z v}{2\Lambda^2} Z_\mu \bar{\chi}\gamma^\mu\chi + \frac{m_Z v}{2\Lambda^2} Z_\mu \bar{\chi}\gamma^\mu\gamma_5\chi\end{aligned} \tag{1.12}$$

where $Z_\mu = (g^2+g'^2)^{-1/2}(gW^3_\mu - g'B_\mu)$. Each line of Equation 1.12 involves various interactions of DM with different particle species. The mono-Higgs signature comes from the terms in the first line of Equation 1.12. The second line leads to a final state with two Higgs bosons plus DM pairs or a Z boson, a Higgs boson plus DM pairs. Lastly the third line gives an effective coupling of Z boson to the DM. These dimension-6 operators generate mono-Higgs signals via the process $q\bar{q} \to Z^* \to h\chi\bar{\chi}$.

For light enough DM, $m_\chi \leq m_Z/2$, invisible Z boson decay are kinematically open. The bound from invisible Z boson decay can become very constraining. The partial invisible Z width for scalar DM can be obtained analytically as

$$\Gamma(Z \to \chi\chi) = \frac{g_2^2 v^4 m_Z}{768\pi c_W^2 \Lambda^4}\sqrt{1-\frac{4m_\chi^2}{m_z^2}} \qquad \text{scalar } \chi \tag{1.13}$$

while the partial invisible Z width for fermion DM is four times larger [52]. If these invisible decays are kinematically open, it is required that $\Lambda \gtrsim 400$ GeV (550 GeV) for scalar (fermion) DM due to constraints on the $\mathscr{B}_{\text{Zinv}}$ [31]. For our study, the dimension-6 operator we consider is that of Equation 1.10.

Dimension-7 operators

At dimension 7, there are several operators that involve Higgs doublets and their derivatives. Since they are just higher-order $1/\Lambda$ corrections to the Higgs portal type operators discussed above, we therefore do not consider them in our studies.

Dimension-8 operators

Finally at dimension-8, many operators can be generated by combining a DM factor and additional SM fields or their duals. Examples of SM field strength tensor include $B^{\mu\nu}$ and $W^a_{\mu\nu}$, which are the SM $U(1)_Y$ and $SU(2)_L$ field strength tensor respectively. For scalar DM the factors can be $i\chi^\dagger \overleftrightarrow{i\partial^\mu}\chi$ or $\chi^\dagger \overleftrightarrow{i\partial^\mu}\chi$ whereas for fermion DM, the factor can be either $\bar{\chi}\gamma^\mu\chi$ or the axial current $\bar{\chi}\gamma^\mu\gamma_5\chi$. We therefore restrict our attention to just one such operator:

$$\mathscr{L}_8 = \frac{1}{\Lambda^4}\bar{\chi}\gamma^\mu\chi B_{\mu\nu}H^\dagger D^\nu H \tag{1.14}$$

This operator generates the mono-Higgs signal via the process $q\bar{q} \to Z^*/\gamma^* \to h\chi\chi$. It has several advantages. First, this operator is not constrained by invisible decays. Secondly, since it involves derivative coupling[1], i.e $\partial_\mu Z_\nu \partial^\nu h$, the $E_{\mathrm{T}}^{\mathrm{miss}}$spectrum is harder and thus gives the best kinematic acceptance efficiency in comparison to the other operators. Third, this operator is weakly constrained by the direct detection since the direct detection signals, arising at one-loop order, are expected to be strongly suppressed.

Essential information of the operators that we are considering in this studies are summarised in Table 1.1. For each operator we have assigned a short name, of which these operators will be referred to throughout this thesis.

Table 1.1 The Lagrangian and free parameters for all operators under consideration. The MadGraph parameters and other information are tabulated as well for reference (v is the vacuum expectation value of the Higgs). The short naming convention is used to refer to the operators throughout the text.

Short Name	Lagrangian	Dim.	Parameters	Madgraph Parameters	S_χ	DM type
xxhh	$\lambda \lvert H\rvert^2 \chi\chi$	4	λ	$g_{DM} = v \cdot \lambda$	0	scalar
xxhhg5	$\frac{1}{\Lambda}\lvert H\rvert^2 \bar{\chi} i\gamma_5 \chi$	5	Λ	$g_{DM} = \frac{v}{\Lambda}$	1/2	fermion
xdxhDh	$\frac{1}{\Lambda^2}\chi^\dagger i \overleftrightarrow{\partial^\mu} \chi H^\dagger i D_\mu H$	6	Λ	$y_\chi = \frac{1}{\Lambda^2}$	0	scalar
xgxFhDh	$\frac{1}{\Lambda^4}\bar{\chi}\gamma^\mu \chi B_{\mu\nu} H^\dagger D^\nu H$	8	Λ	$y_\chi = \frac{1}{\Lambda^4}$	1/2	fermion

1.7 Valid Parameter Space

The valid parameter spaces of the EFT operators considered here are summarised in Table 1.2. Perturbativity constraints require that the couplings of the operators be less than 4π [63] (albeit relies on a somewhat arbitrary definition of "non-perturbativity", there could be more stringent constraints). For the `xxhh` operator, the unitless coupling is simply λ, while for the `xxhhg5` operator, it is v/Λ. In the `xdxhDh` operator, after spontaneous symmetry breaking, an effective coupling between DM and the Z boson, $g_{Z_{\mathrm{eff}}}$, can be constrained as discussed in Reference [52]. For the `xgxFhDh` operator, it is more difficult to define a coupling with concrete perturbativity constraints. Instead, truncation methods [63] can be used to constrain the `xgxFhDh` operator. This method is discussed in more details in Section 8.3.1.

In addition to perturbativity constraints, operators which predict the Higgs/Z boson decaying to DM are constrained by the $\mathscr{B}_{\mathrm{inv}}$ limits on the Higgs/Z boson when $m_\chi < m_{h,Z}/2$.

[1] from Feynman rules, any ∂_μ in the interaction Lagrangian introduces factor of momenta for the corresponding vertex in the momentum space.

Table 1.2 The valid parameter spaces for the EFT operators.

Name	Perturbativity Requirement	$\mathscr{B}_{inv}$ Requirement
xxhh	$\lambda < 4\pi$	$m_\chi < \frac{m_h}{2} \rightarrow \lambda \lesssim 0.016$
xxhhg5	$\Lambda > \frac{v}{4\pi}$	$m_\chi < \frac{m_h}{2} \rightarrow \Lambda \gtrsim 10$ TeV
xdxhDh	$g_{Z_{\text{eff}}} < 4\pi$ ($\Lambda \gtrsim 30$ GeV)	$m_\chi < \frac{m_Z}{2} \rightarrow \Lambda \gtrsim 400$ GeV
xgxFhDh	Use Truncation	N/A

1.8 Analysis overview

In this analysis, we are looking for the process $pp \rightarrow h\chi\bar{\chi} \rightarrow b\bar{b}\chi\bar{\chi}$. The energetic DM particles produced by the collider will escape detection because they cannot deposit any measurable amount of energy in a calorimeter, but could be registered by means of large transverse momentum or energy imbalance (see Section 4.10). This large missing transverse energy can serve as a unique signature to trigger the event. One can observe the missing transverse energy generated by the escaped DM pairs only if the DM pair is accompanied by a detectable Higgs boson, produced almost back-to-back with the $E_{\text{T}}^{\text{miss}}$. Such event topology is illustrated in Figure 1.9 To reconstruct the Higgs boson, we search for events that contain a jet with reconstructed mass consistent with the Higgs boson mass. Furthermore, the Higgs candidate should have two associated b-tagged jets (see Section 4.9). Since no lepton is expected in the final state, any event with a lepton is vetoed.

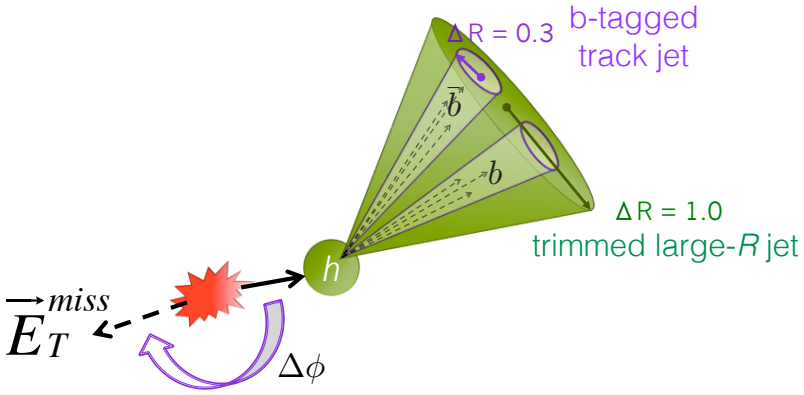


Fig. 1.9 Schematic diagrams depicting the signal event topology.

As we shall see in Section 5.2, the EFT mono-H signal in general have a broad $E_{\text{T}}^{\text{miss}}$ distribution, extending up to a very high $E_{\text{T}}^{\text{miss}}$ value. The p_{T} of the recoiling Higgs boson exhibits similar trend. Thus, in order to ensure high signal acceptance, unlike similar mono-Higgs searches [51, 64], this analysis focus on high momentum regime where the Higgs boson is highly boosted. At very high boost, the hadronic showers from Higgs decays can become so close such that the ability to resolve the individual hadronic decay products by using standard

jet reconstruction algorithms with small jet radii begins to degrade. Furthermore, under such dense environment the performance of the b-tagging algorithm will also suffers. Under such adverse circumstance, in order to recover or even enhance the signal selection efficiency and background rejection power, we employed novel boosted object reconstruction techniques (Section 4.7) and high-p_{T} b-tagging algorithm (Section 4.9).

Figure 1.10-Figure 1.14 shows the typical Feynman diagrams of the backgrounds which can mimic the $E_{\mathrm{T}}^{\mathrm{miss}} + b\bar{b}$ final state. These backgrounds include:

- QCD multi-jet production may have missing transverse energy if one or more of the jets are mis-measured by the calorimeter although there is no $E_{\mathrm{T}}^{\mathrm{miss}}$ in principle. QCD multi-jet events with large $E_{\mathrm{T}}^{\mathrm{miss}}$ (i.e. $>$ 300 GeV) do not occur very often due to the fact that substantial fluctuations in calorimeter jet energy measurements are rare.

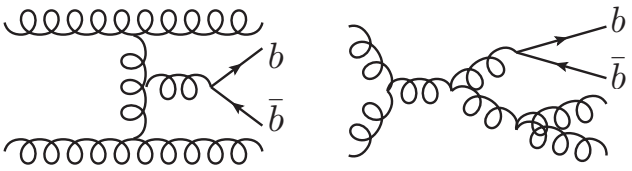


Fig. 1.10 Typical Feynman diagrams for the QCD multi-jet induced $E_{\mathrm{T}}^{\mathrm{miss}} + b\bar{b}$ background. Diagrams with external quark lines instead of gluons are not shown.

- $Z(\rightarrow \ell\ell)$+jets and single top backgrounds may contaminate the SR when one or more leptons from the vector boson decay are not identified. However, due to the high identification efficiency of electrons and muons, these two backgrounds are expected to be negligible.

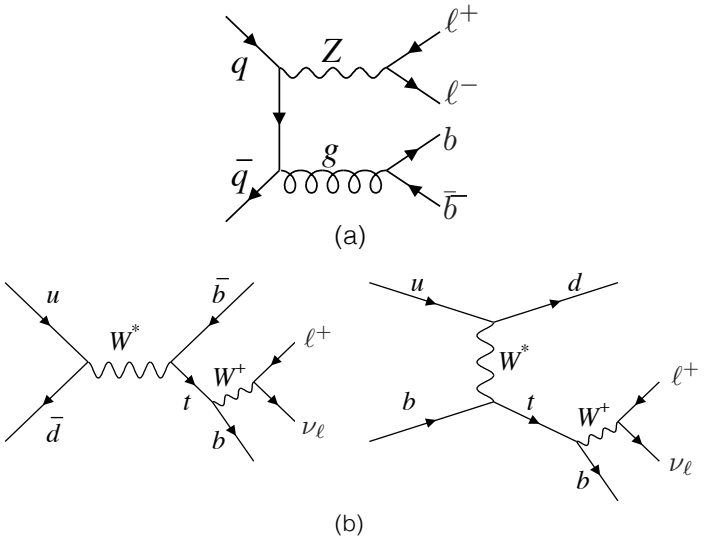


Fig. 1.11 Representative Feynman diagrams for (a) $Z(\rightarrow \ell\ell)$+jetsand (b) single top s-channel and t-channel processes, with leptonic W boson decays.

- Diboson and SM VH backgrounds may contaminate the SR when one or more leptons from vector boson decay are not identified. In addition, the Z boson in these backgrounds may also decays invisibly. This makes it harder to remove completely the diboson and SM VH backgrounds.

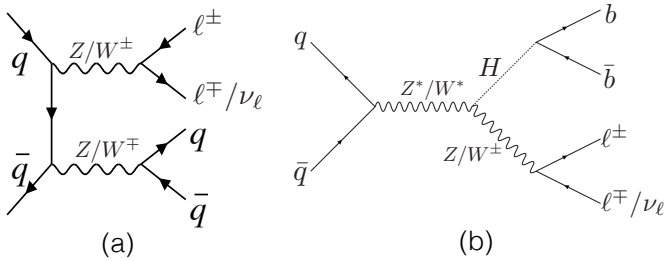


Fig. 1.12 Representative Feynman diagrams for (a) diboson and (b) SM VH productions.

- The $t\bar{t}$ and $W(\to \ell\nu)$+jets production can mimic the $E_{\mathrm{T}}^{\mathrm{miss}} + b\bar{b}$ signal when the lepton from W boson decay is not detected.

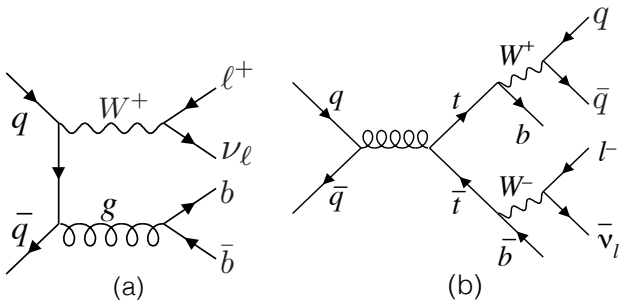


Fig. 1.13 Representative Feynman diagrams for (a) $W(\to \ell\nu)$+jets and (b) $t\bar{t}$ background processes, which has a small contribution to the $E_{\mathrm{T}}^{\mathrm{miss}} + b\bar{b}$ final state.

- The $Z(\to \nu\nu)$+jets events are the dominant background in this analysis. It has exactly the same final state as the targeted signature of $E_{\mathrm{T}}^{\mathrm{miss}} + b\bar{b}$, making it impossible to entirely eliminate its contribution from the total event count. Thus, accurate estimation of this background is crucial.

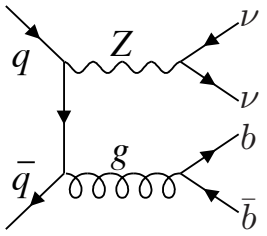


Fig. 1.14 Feynman diagram for the dominant $Z(\to \nu\nu)$+jets backgrounds.

Event selection proceeds via a robust cut-based method which has the advantage of being readily grasped and easily described. To avoid any subtle biases stemming from the analyser preconceptions, we perform a blind analysis by masking the data in the signal region. Selection criteria are determined and optimised by using only the simulated events without looking at the

data in the signal region. After applying the complete event selection, the DM signal (if exist) is expected to manifest as an anomalous excess in the $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum.

1.9 Thesis overview

The remainder of our work is outlined as follows. Section 2 provides the description of the experimental setup at the Large Hadron Collider and the ATLAS experiment. It is followed by Section 3 in which signal and background Monte Carlo processes as well as the dataset used in this analysis are discussed. Section 4 defines the different physics objects under investigation. In this section, the details of each object reconstruction are discussed together with the imposed selection criteria. The exact event selection and optimisation is the topic of Section 5. Section 6 then bring the focus to the background estimation methods, followed by Section 7 in which the systematic uncertainties associated with both simulated background and signal processes are evaluated. The results of the search and the statistical interpretation are presented in Section 8. Section 9 then demonstrate possible improvement that can be done and the outlook of for future mono-Higgs search. Lastly, the thesis is concluded in Section 10.

Chapter 2

The Large Hadron Collider and The ATLAS Experiment

About one hundred meters beneath the border between France and Switzerland, there is a circular particle accelerator that might reveal to us the secrets of the universe by colliding protons or heavy ions at a velocity approaching the speed of light. Along the circumference of the LHC there are four major detectors, each conducting different or sometimes overlapping experiments that aim to uncover important new information about how our universe works. One of them is known as A Toroidal LHC ApparatuS (ATLAS) experiment on which the work of this thesis is based.

This chapter will attempt to give a brief introduction to the LHC as well as the ATLAS detector. The first section of this chapter will give an overview of the LHC. Section 2.2 and its subsections are devoted for the discussion of the ATLAS detector as a whole and its sub-detector systems. In Section 2.6 we will see how the trigger and the data acquisition system work. The principal references on the design of the LHC main ring, the infrastructure and the injector chain are [65–69]. On the other hand, more comprehensive information on the design and the performance of ATLAS detector can be found in References [70–75].

2.1 The Large Hadron Collider (LHC)

LHC lies in a tunnel 26.7 kilometres in circumference that was originally built for Large Electron Positron (LEP) collider. One of the main goals of the LHC is to allow physicists to reveal the physics beyond the SM by producing yet to be discovered rare events through proton-proton ($p-p$) collisions. In order to produce these rare events (whose probability varies with collision energy, among others) in the first place, high beam energy is required. To produce enough of them requires high luminosity. Therefore the LHC is designed to collide proton beams with a centre-of-mass energy of 14 TeV and an unprecedented luminosity of $10^{34} cm^{-2} s^{-1}$.

The existing accelerator complexes at CERN, including the LHC is illustrated in Figure 2.1. Before being injected into the LHC, proton beams will pass through a succession of machines which bring the beam to increasingly higher energies. These steps are shown on the left of Figure 2.1. The protons begin their journey in the LINear ACcelerator (LINAC) 2 with 50 MeV of energy. They are then injected into Proton Synchrotron Booster (PSB) which takes over to accelerate the protons to an energy of 1.4 GeV before being fed to the PS where they are accelerated to 25 GeV. Finally, having accumulated an energy of 450 GeV in the Super Proton Synchrotron (SPS), the protons are then injected into the LHC, where they acquire nominal energy of 7 TeV.

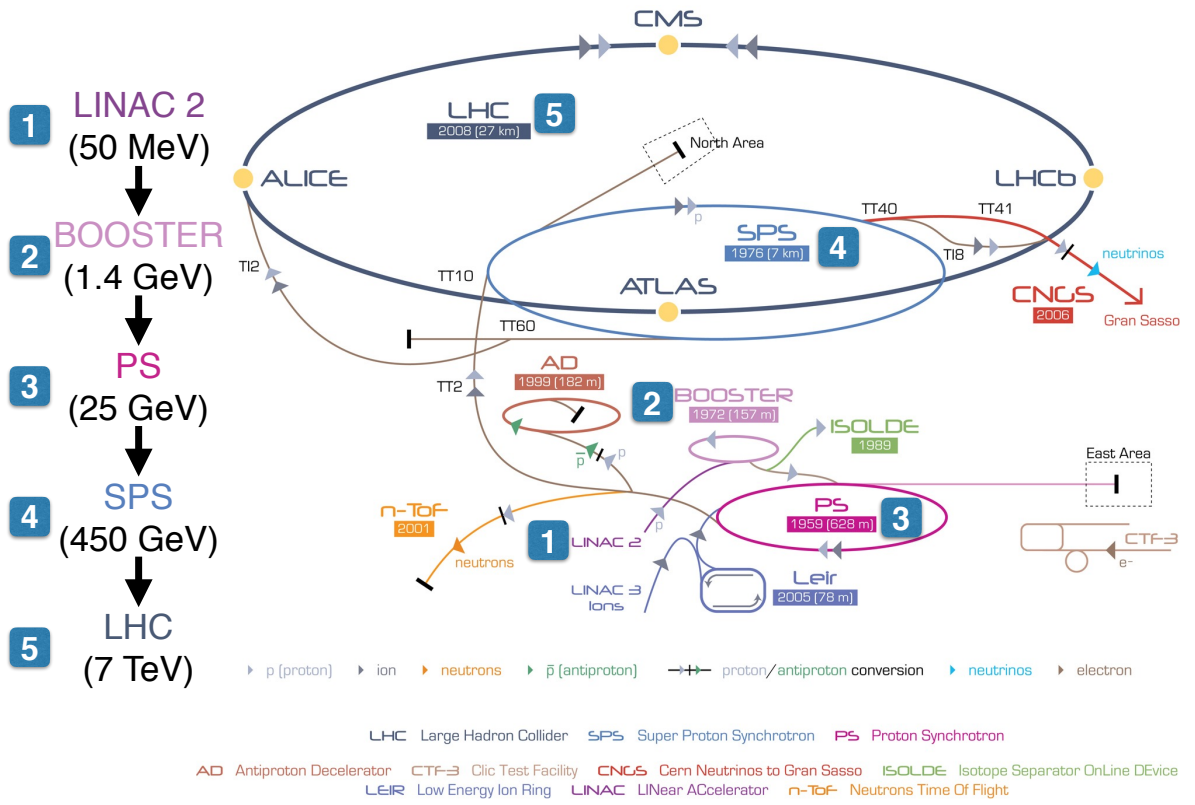


Fig. 2.1 CERN accelerator complexes and experiments. On the left is a schematic representation of the different steps necessary to inject protons into the LHC ring. For each succession of machines, the beam's energy is increased by many fold. Figure adapted from CERN-DI-0606052 [76] ©CERN Geneva.

Unlike particle-antiparticle colliders (e.g Tevatron) that can have both circulating beams sharing the common magnet structure and vacuum chamber, being a particle-particle collider LHC's counter-rotating beams need opposite magnetic fields. However, due to limited space in the tunnel and partly as a cost saving measure, the LHC adopted the twin-bore magnet design

Table 2.1 An overview of performance-related parameters during LHC operations in 2012. In comparison, the design values are also shown. [77]

Parameters	Value in 2012	Design value
Beam energy [TeV]	4	7
β^* at interaction point 1 (ATLAS)	0.6	0.55
Bunch spacing [ns]	50	25
Number of bunches	1374	2808
Average bunch intensity [protons per bunch]	$1.6-1.7\times10^{11}$	1.15×10^{11}
Normalised emittance at start of fill [mm.mrad]	2.5	3.75
Peak luminosity $[cm^{-2}s^{-1}]$	7.7×10^{33}	1×10^{34}

which consist of two sets of coils and beam channels within the same mechanical structure and cryostat.

The maximum beam energy (momentum) attainable depends on the bending radius of the magnet and the strength of the dipole magnetic field that keeps particles in their orbits. The bending radius of the LHC main dipoles at 1.9K is 2803.98 m. The maximum beam energy that can be reached in the LHC is therefore limited by the peak dipole magnet field in the storage ring that bends the trajectory of the protons. A nominal operating dipole magnetic field of 8.33 T thus gives a maximum 7 TeV proton beam. To keep the protons on their circular path and in focus, the LHC is fitted out with 1232 main dipole bending magnets and 392 main quadrupole focusing magnets that are immersed in a bath of superfluid helium at 1.9 K. Driven by the requirements for a long beam life time and to minimise the backgrounds to the experiments, beam vacuum system must maintained a stringent vacuum as low as $10^{15}\mathrm{H_2m^{-3}}$ (equivalent hydrogen gas densities at cryogenic temperature).

For physicists, besides the beam energy, the other most important parameter is the number of interesting events per second that can be generated in LHC collisions. It depends on the machine luminosity [68]:

$$\mathscr{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\varepsilon_n \beta^*} F \tag{2.1}$$

where N_b is the number of particles in each bunch, n_b is the number of bunches per beam, γ_r is the relativistic gamma factor, f_{rev} is the number of revolution per second, ε_n is the normalised transverse beam emittance (a measure of the spread of the beam in transverse phase space), β^* is the beta or amplitude function at the interaction point, and F is the geometric luminosity reduction factor due to the crossing angle at the interaction point. This relation assumes round beams and equal values of the beta function for both beams in both planes. The corresponding operating value (in 2012) and design value for these parameters are listed in Table 2.1.

At the LHC, each proton beam at full intensity will consist of 2808 bunches per beam with a nominal bunch spacing of $\sim$ 7.5m or 25 ns. Each bunch will contain 1.15×10^{11} protons per bunch at the start of nominal fill. At near light-speed, a proton in the LHC will make 11

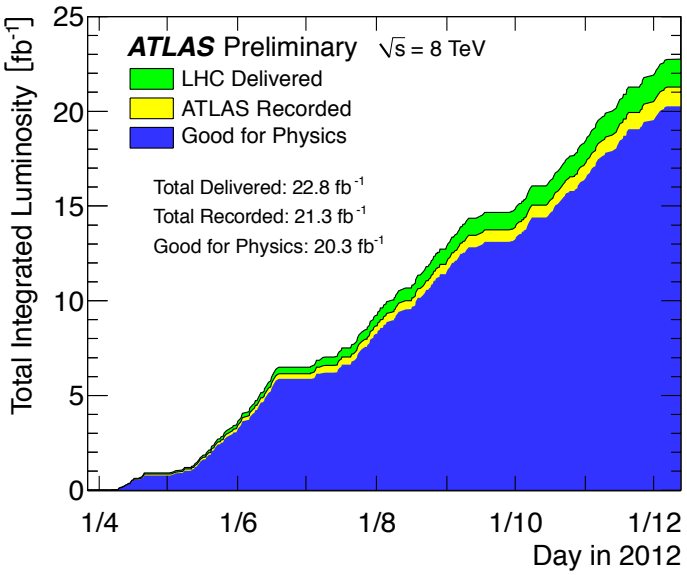


Fig. 2.2 Cumulative luminosity versus time delivered to (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams and for $p-p$ collisions at 8 TeV centre-of-mass energy in 2012. Plot taken from [78].

245 revolution per second. With this huge number of particles, what is left for LHC to do in order to achieve a peak design luminosity of $10^{34} cm^{-2} s^{-1}$ is to squeeze those particles into the smallest amount of space around the interaction region. This can be done by making the emittance and beta function as small as possible.

The LHC is first started up with first beams circulated successfully on 10 September 2008. Unfortunately a magnet quench incident seriously delayed the commissioning of the machine. The LHC beam did not see beam again before November 2009. On 30th March 2010 first collisions at a centre of mass energy of 7 TeV took place. Then the rest of year 2010 mainly devoted to commissioning. 2011 marked the first production year with integrated luminosity of more than 5 fb^{-1}delivered by the LHC. In 2012, the beam energy was increased to 4 TeV per beam at a 50 ns bunch spacing with around 1380 bunches. At the end of 2012 $p-p$ collision run, the LHC managed to delivered a total integrated luminosity of 22.8 fb^{-1}. In total, ATLAS recorded 21.3 fb^{-1}where 20.3 fb^{-1}of data is certified to be good quality data. This is the dataset which this analysis is based on. Cumulative integrated luminosity versus day in 2012 are plotted in Figure 2.2.

The high instantaneous luminosity of the LHC leads to the overlap of several $p-p$ interactions in a single collision event. Depending on the length of the read-out window of a sub-detector, signals from neighbouring bunch crossings can also be present simultaneously when the detector is reading out. This kind of additional proton-proton collisions is called "pile-up" events. The impact of interactions from the previous or next bunches is referred to as out-of-time pile-up, whereas in-time pile-up results from the presence of multiple $p-p$

interactions in the same bunch crossing. A common measure for the amount of pile-up in a collision event is the number of interactions per bunch crossing, μ. It is calculated as

$$\mu = \frac{L_{bunch} \times \sigma_{inel}}{f_{rev}} \tag{2.2}$$

where L_{bunch} is the per bunch instantaneous luminosity, σ_{inel} is the inelastic cross section which is taken to be 73 mb and f_{rev} is the LHC revolution frequency. The number of $p-p$ inelastic interactions per bunch crossing follows a Poisson distribution with mean value $<\mu>$. This value changes during data-taking as a function of time. It decreases with decreasing beam intensity and increasing emittance. The highest value is at the start of the stable beam period of the fill. The $<\mu>$ was found to be around 20 in 2012, as shown in Figure 2.3.

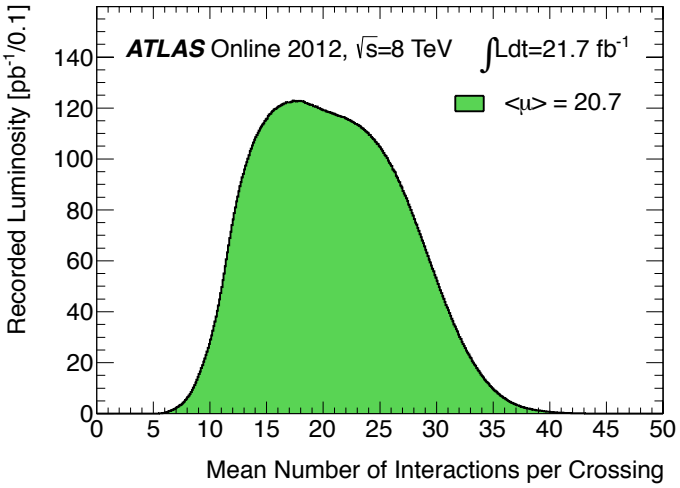


Fig. 2.3 The luminosity-weighted distribution of the mean number of interactions per crossing for 2012 full $p-p$ collisions dataset. The average number of interactions per crossing corresponds to the mean of the poisson distribution of the number of interactions per crossing for each bunch. Plot taken from [78].

2.2 The ATLAS detector

The ATLAS experiment is one of the two general purpose experiments located at the LHC. The ATLAS detector is forward-backward symmetric with respect to the interaction point (IP). It is a hermetic detector designed to cover as large as possible the area around the interaction point (IP). ATLAS incorporates multiple types of sub-detectors. It has a cylindrical shape with concentric layers stacked onto each other. Each of the layers detects different properties of particles. When particles from the $p-p$ collisions are produced in the centre of ATLAS, they move through the detector and are detected by its successive layers.

Closest to the IP is the Inner Detector (ID). It measures the trajectories of charged particles. Surrounding the ID is the calorimeter systems which are designed to measure the energy of electrons, photons, and hadrons. Most particles except muons and neutrinos are stopped by the calorimeter system. The Muon Spectrometer (MS) surrounds the calorimeters. The MS is designed to measure the trajectories of highly penetrating muons leaving the calorimeters. These sub-detectors are arranged around a unique magnet system that bends charged particles in the ID and the MS, allowing their momenta to be measured. Each of this sub detector system will be discussed in more detail in the following subsections.

Figure 2.4 shows the overall layout of the ATLAS detector. Its main performance goals are listed in Table 2.2.

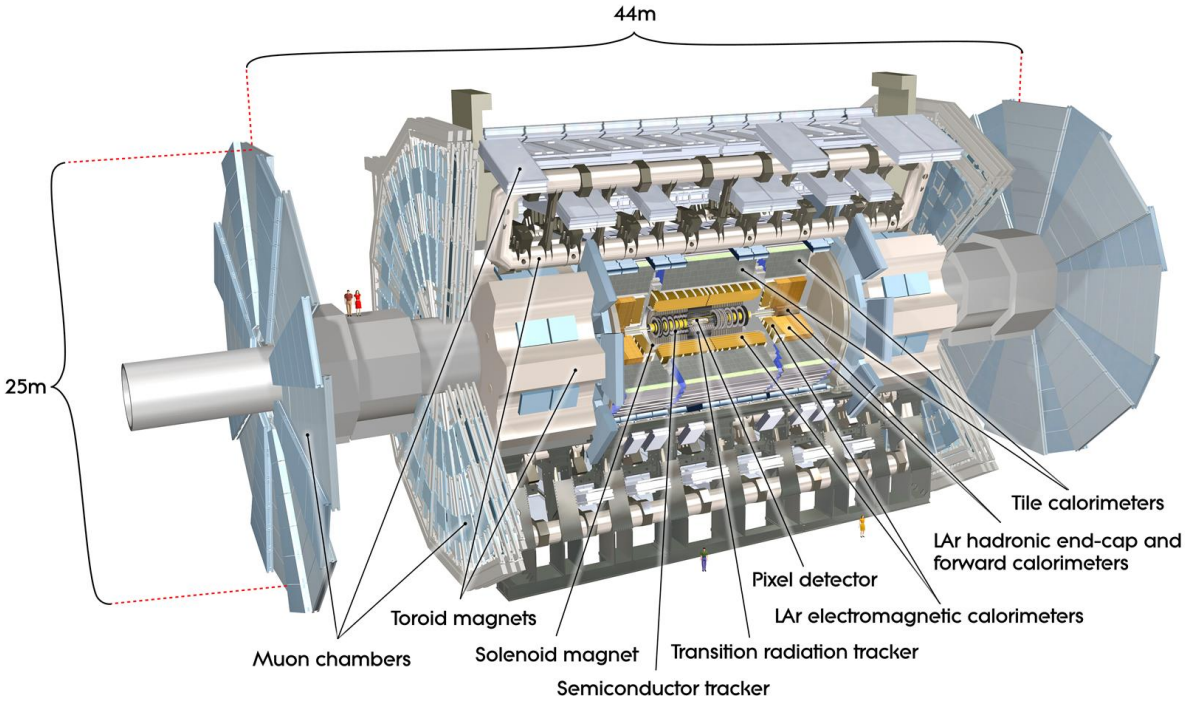


Fig. 2.4 Overview of the ATLAS detector [CERN-GE-0803012 ©CERN Geneva].

2.3 Inner detector

The main function of the ID [79, 80] is to reconstruct tracks and vertices as well as to measure the momentum of charged particles. It also provides particle identification. A solenoidal magnetic field of 2 T, parallel with the beam line, surround the ID. The ID comprise of three independent but complementary sub-detectors, namely the pixel detector [81], the semiconductor tracker (SCT) [82] and the transition radiation tracker (TRT) [83]. The ID layout is

Table 2.2 Designed momentum and energy resolution and pseudorapidity coverage of the ATLAS detector.

Detector component	Required resolution [GeV]	η coverage	
		for measurement	for trigger
Inner detector	$\frac{\sigma_{p_T}}{p_T} = 0.05\% p_T \oplus 1\%$	± 2.5	-
EM calorimetry	$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry (jets) :			
•barrel and end-cap	$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$	± 3.2	± 3.2
•forward	$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\frac{\sigma_{p_T}}{p_T} = 10\%$ at p_T= 1000	± 2.7	± 2.4

shown in Figure 2.5 and Figure 2.6. The expected performance of each component, which is required even at the highest luminosities expected from LHC collisions, is listed in Table 9.1b.

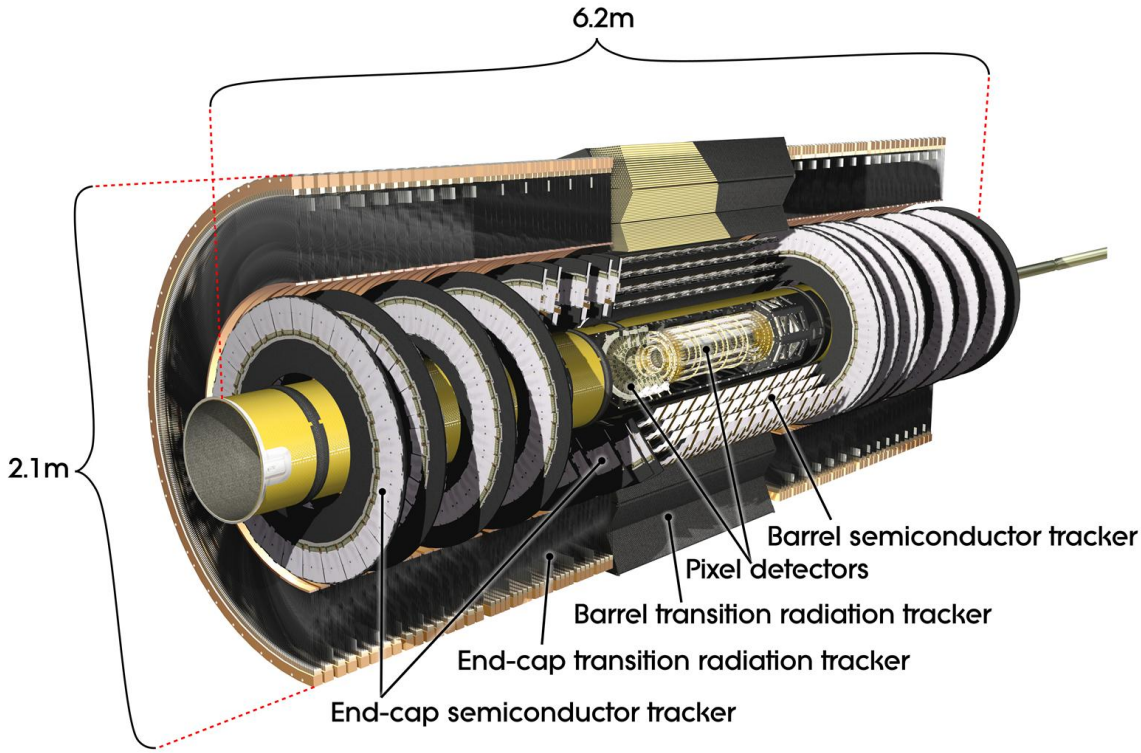


Fig. 2.5 Cut-away view of the ATLAS inner detector. [CERN-GE-0803013 ©CERN Geneva].

The innermost part of the ID, immediately outside the LHC beam pipe, is the pixel detector. It provides the highest granularity with the minimum pixel size on a sensor is $50 \times 400\ \mu m^2$. All of the 1744 pixel sensors have approximately 80 million readout channels and are distributed

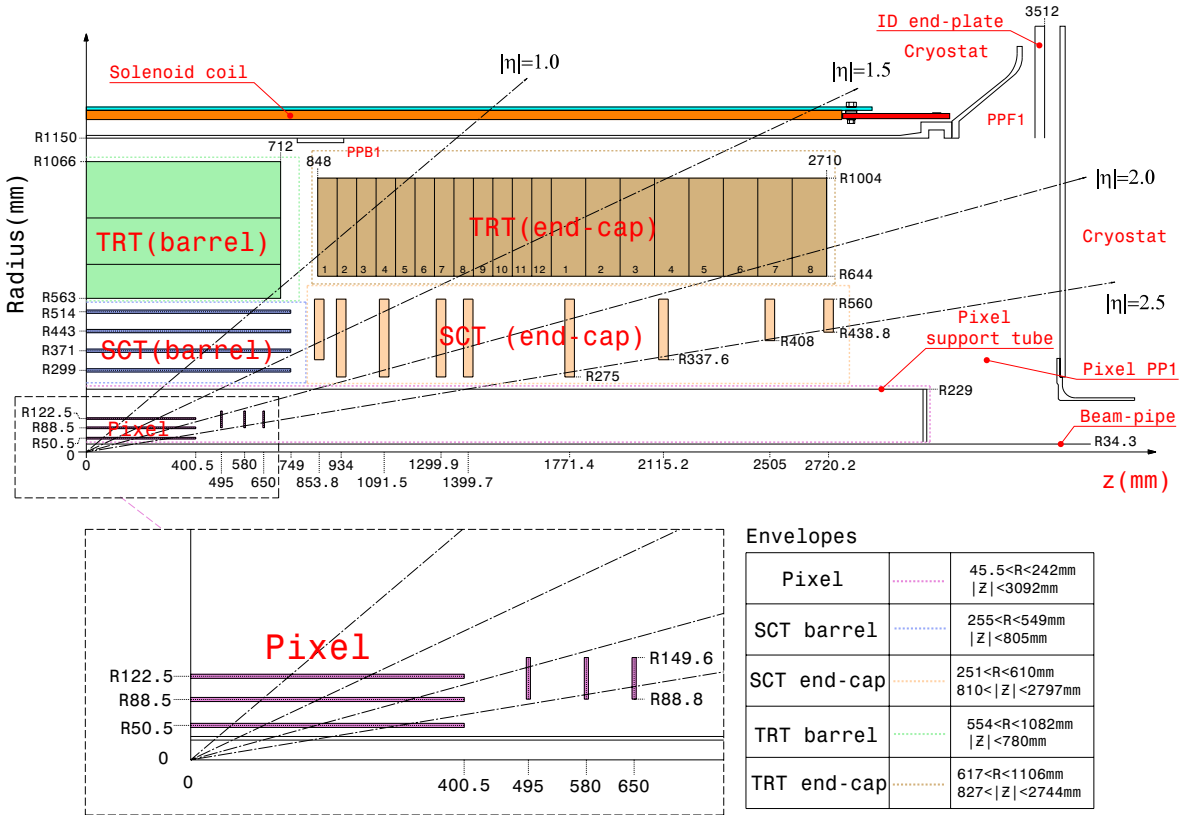


Fig. 2.6 Plan view of a quarter-section of the ATLAS inner detector showing each of the pixel, SCT and TRT elements and their coverage. Figure taken from Reference [70].

in three cylindrical layers and three disk layers on each side. It is designed to provide uniform coverage within the pseudorapidity[1] range $|\eta| < 2.5$. The intrinsic position resolution of individual pixel detector module in r-ϕ (z) direction is expected to be around 10 μm (115 μm). The intense radiation environment of the LHC, especially at small radii, mandates all pixel detectors elements to have a high radiation hardness requirement. Hence they are designed to withstand a lifetime dose of 500 kGy or a total particle fluence (normalized to the equivalent of the damage of 1 MeV neutrons per non-ionising energy loss) of about $10^{15} n_{eq} cm^{-2}$.

Similar to the pixel detector, the SCT detectors use semiconductor sensor to provide precision space-point coordinates. The SCT system consists of four barrel layers and two end-caps each of nine disks. In total there are 8448 barrel and 7104 end-cap microstrip sensors,

[1]The ATLAS coordinate system is a right-handed system with the x-axis pointing to the center of the LHC ring and the y-axis pointing upwards. The polar angle is measured with respect to the LHC beam-line. The azimuthal angle is measured with respect to the x-axis. The rapidity is defined as $y = 0.5 \times ln[(E + p_z) = (E - p_z)]$, where E denotes the energy and p_z is the component of the momentum along the beam direction. The pseudorapidity is an approximation for rapidity y in the high energy limit, and it is related to the polar angle, θ as $\eta = -ln(tan\frac{\theta}{2})$. Transverse momentum and energy are defined as $p_T = p \times sin\theta$ and $E_T = E \times sin\theta$, respectively.

giving approximately 6.3 million readout channels. The sensor is 285 μm thick. For the barrel layers, each microstrip sensor has 768 readout strips at a constant pitch of 80 μm. They are assemble pairwise, glued back-to-back with a relative rotation of 40 mrad with respect to each other, with one set of strips in each layer parallel to the beam direction. For the end-cap layers, the pair of microstrip sensors are arranged such that a set of strips running radially and the other at an angle of 40 mrad. Because of the wedge-shaped geometry of these end-cap sensor, the pitch is not constant but the mean is still approximately 80 μm. Similar to the pixel detector, the SCT provides coverage within the pseudorapidity range $|\eta| < 2.5$. The intrinsic position resolution of the microstrip sensor is around 17 μm in $R\phi$ and 580 μm in z. All the SCT sensors components are designed to be able to operate up to $\sim 2 \times 10^{14}\ n_{eq}cm^{-2}$

Finally, the TRT with 73 barrel layers and 160 end-cap layers constitutes the outermost part of the ID. The detecting elements are made out of 4 mm inner diameter thin-walled proportional drift tubes with 30 μm diameter gold-plated tungsten wire. The barrel parts contain 52 544 straws 144 cm in length oriented parallel to the beam while two end-caps has 122 880 straws 37 cm in length oriented perpendicular to the beam axis. Each straw is filled with gas mixture $Xe(70\%) + CO_2(27\%) + O_2(3\%)$ that was chosen to provide advantages such as efficient transition radiation absorption, high electron drift velocities and a wide operating plateau, among others. The TRT provide in average 35-40 space points on a track for pseudorapidity rage $|\eta| < 2.0$. The intrinsic drift-time accuracy of TRT is around 130 μm per straw.

2.4 Calorimeters

The ATLAS calorimeters consist of an inner electromagnetic (EM) calorimeter and an outer hadronic (HAD) calorimeter. Both of the calorimeters are classified as sampling calorimeters that consist of alternating dense absorber material (e.g. lead, copper or iron) to induce either EM or hadronic showers and active material (e.g. liquid argon or scintillating tiles) to sample the output signal by the particle shower, which is proportional to the incident particle's energy. When a particle travel through the calorimeter only some fraction of the shower energy is sampled by the active material. The full energy of the original particle is reconstructed offline through layer weighting technique. To achieve excellent missing energy determination, the ATLAS calorimeters have an almost 4π acceptance. An overview of the ATLAS calorimeters is illustrated in Figure 2.7.

The ATLAS EM calorimeter is a liquid argon (LAr) sampling detector with lead absorbers. Between the absorbers are the readout electrodes which consist of three conductive copper layers separated by insulating polyimide sheets. The calorimeter is divided into a barrel part and two end-caps at each side. The barrel calorimeter cover the pseudorapidity range $|\eta| < 1.475$, with a total thickness of varying from 22 to 33 radiation length, X_0. On the other hand, the end-cap parts cover the region $1.375 < |\eta| < 3.2$ with active depth ranging from 24 to 38 X_0. Both the barrel and the end-cap EM calorimeter have accordion shaped absorbers and

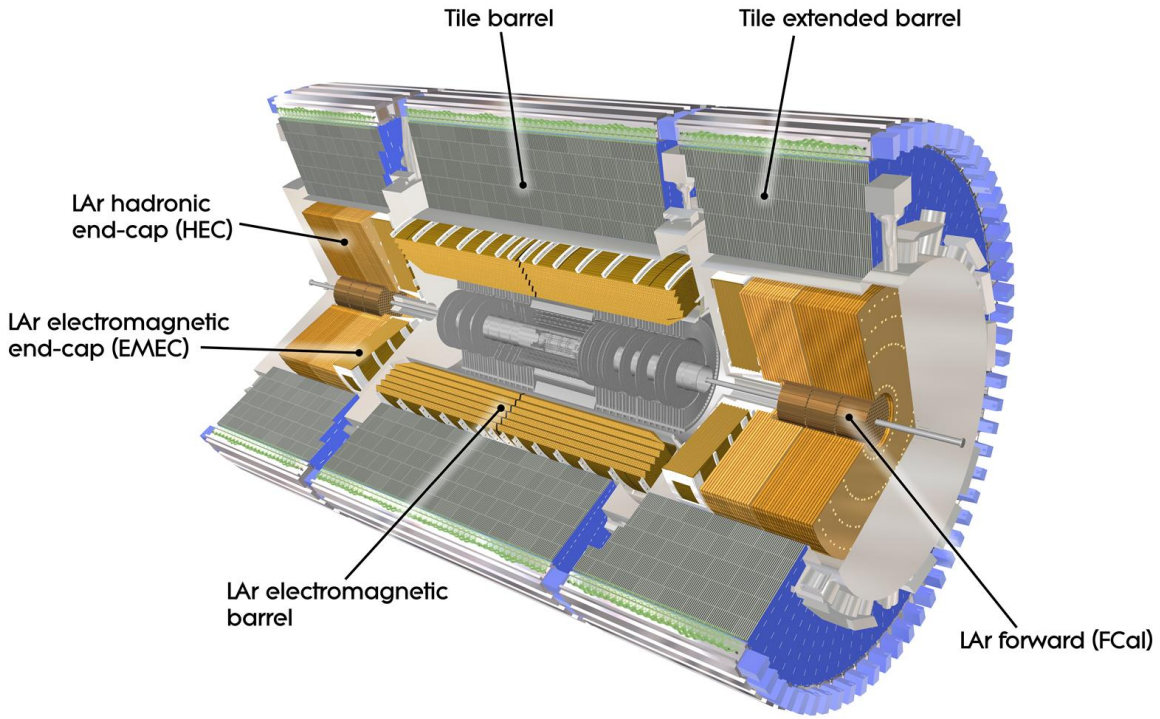


Fig. 2.7 Overview of the ATLAS calorimetry. Near the beam-pipe the tracker is visible, surrounding it is the EM calorimeter and beyond the hadronic calorimeter. Both barrel and end-caps elements are displayed. [CERN-GE-0803015 © CERN Geneva].

electrodes. Such a geometry was chosen because it provides a full coverage in ϕ without any cracks. These calorimeter parts are further divided into several active layers in depth. For region $0 < |\eta| < 2.5$ there are three longitudinal segments while for region $2.5 < |\eta| < 3.2$ there are two. The sketch of the barrel module with three longitudinal segments and with different $\eta - \phi$ granularities is shown in Figure 2.8. The first layer is finely segmented with strips along η to allow accurate position measurement. The second layer collects the largest fraction of energy of the electromagnetic shower. Together, the first and the second layer enable the measurement of the electromagnetic shower in η direction. The third layer collects only the tail of the electromagnetic shower. It is therefore less segmented in η. The active depth of the three samplings are 4.3 X_0, 16 X_0, and 2 X_0 respectively, at $\eta = 0$. In the region $|\eta| < 1.8$, the EM calorimeter is preceded by a pre-sampler (a separate thin instrumented LAr layer) to correct for the energy lost in the upstream material (inner detector, supporting structure, cryostat, superconducting coil etc.). Finally in the forward region ($3.1 < |\eta| < 4.9$), the LAr EM calorimeter has longitudinal samplings with copper rods and matrix.

Surrounding the EM calorimeter and in the region $0 < |\eta| < 1.6$, the hadronic calorimeter consists of scintillator tiles as the active medium and steel as the passive absorber medium.

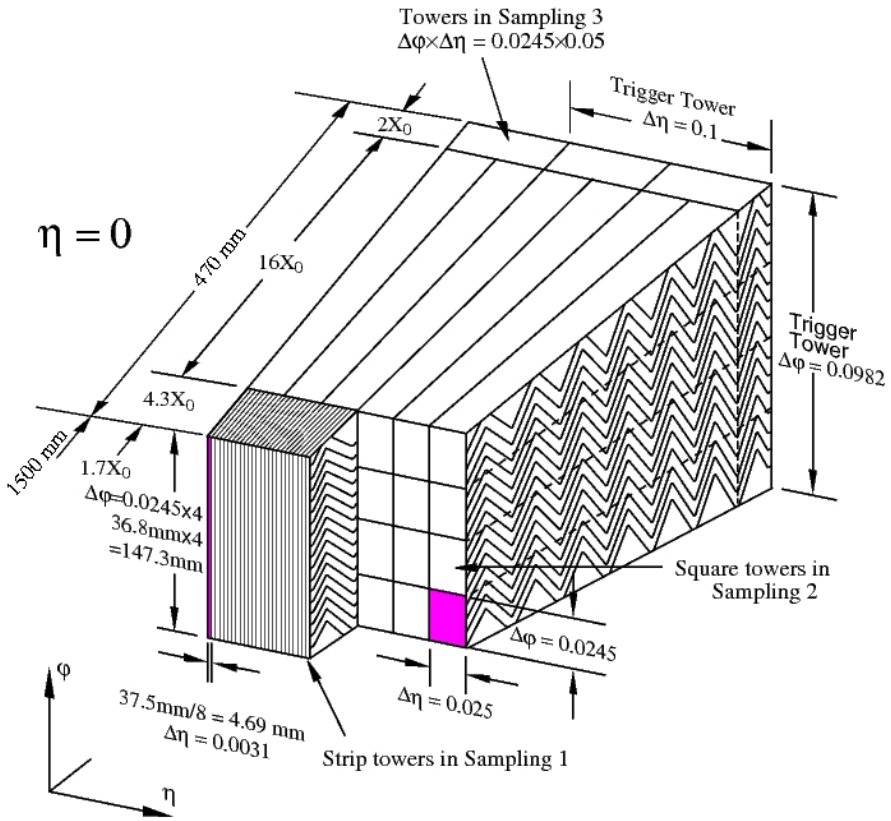


Fig. 2.8 Schematic view of a section of the EM barrel detector, displaying the granularity of the three sampling and the accordion structure. Figure taken from Reference [84]

The tile calorimeter is approximately 7.4 interaction length (λ) thick and the scintillator tiles are oriented radially and normal to the beam line. It is composed of one central barrel and two extended barrels. Both sides of the scintillator tile are read by wavelength-shifting fibres into two separate photomultiplier tubes. The readout cells are projective (pointing) in ϕ and only pseudo-projective (semi-pointing) in η. For region $1.5 < |\eta| < 3.2$, the hadronic end-cap calorimeter (HEC) is an LAr sampling calorimeter with copper as the absorber medium. The parallel copper plate is orthogonal to the beam axis. The HEC consists of two consecutive wheels with absorber thickness of 25 and 50 mm respectively. Finally the hadronic calorimeter is extended to $3.1 < |\eta| < 4.9$ by hadronic forward calorimeter. This hadronic forward calorimeter has longitudinal samplings with tungsten rods and matrix.

2.5 Muon spectrometers

Muon is one of the only few detectable particles that can penetrate through the calorimeters and reach the outermost part of ATLAS, known as the muon spectrometer (MS) [85]. Apart from

serving the functions of simply identifying and independently triggering on muon, the MS is also design for stand-alone (independent of the inner detector) high performance transverse momentum measurement as required by the physics programs of ATLAS. More specifically, the MS is expected to be able to measure the transverse momentum of 10-200 GeV muons with 2-4% accuracy and 1 TeV muons with 10% accuracy. The MS is a tremendously large tracking system. It is composed of the large superconducting air-core toroids to bend muon trajectory in the $r-z$ plane, the precision tracking chambers for accurate momentum resolution and a set of triggering chambers with fast response. An illustration of the muon chambers and magnets is shown in Figure 2.9.

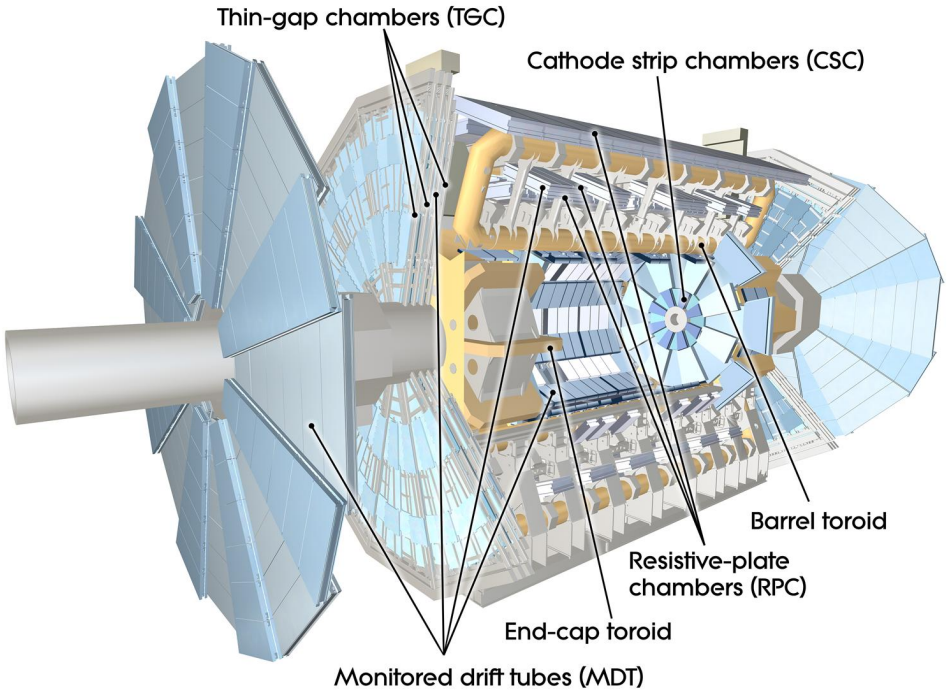


Fig. 2.9 Overview of the ATLAS muon spectrometer [CERN-GE-0803017 © CERN Geneva].

The precision tracking chambers in the barrel region are arranged in three concentric cylindrical shells (or stations) around the beam axis. Two types of precision tracking chamber are used. They are the Monitored Drift Tube (MDT) and Cathode-Strip Chamber (CSC). The MDTs cover the pseudorapidity range $|\eta| < 2.7$. These chambers consist of three to eight layers of pressurised aluminium drift tubes with tungsten-rhenium wires. When a muon passes through the volume it knocks electrons off the atoms of the gas which then drift to the wire under the influence of the applied electric field. By registering the drift times of these electron, one can determine the coordinates of a track in the plane of the layer and in the direction across the tubes. The average resolution is 80 μm per tube, or about 35 μm per chamber. The CSCs are multi wire proportional chambers with the wires oriented in the radial direction. Each chamber has four cathode strip planes at right angles to the wire direction, resulting in four

independent measurements in η and ϕ along each track. The position of the track is obtained by interpolation between the charges induced on neighbouring cathode strips. The CSCs provide a resolution of 60 μm per CSC plane along the bending direction.

The trigger chambers of the MS are designed to provide functions such as discrimination on muon transverse momentum, bunch-crossing identification, and measurement of the muon coordinate in the direction orthogonal to that determined by the precision-tracking chambers. The trigger system covers the pseudorapidity range $|\eta| < 2.4$. For the barrel regions ($|\eta| < 1.05$), the Resistive Plate Chamber (RPC) is used while for the end-cap regions ($1.05 < |\eta| < 2.4$) the Thin Gap Chamber (TGC) is used. Each RPC consist of two parallel plates with opposite charges, both made of a very high resistivity plastic material and separated by a gas volume. Each plate measures coordinate in η and ϕ. A track going through all three RPC layers thus delivers typically six measurements in η and ϕ. These measurement can be then used in a coincidence scheme of the trigger logic to reject fake tracks from noise hits. The TGC operates on the same principle as multi-wire proportional chambers, providing good time resolution to tag the beam-crossing with high efficiency and high rate capability. The TGCs are arranged in four layers orthogonal to the beam axis. The trigger information is generated by a system of fast coincidences between the three last layers along the trajectory of the muon particle.

2.6 Trigger and data acquisition system

When operating at the design luminosity of $10^{34} cm^{-2} s^{-1}$ the LHC will have a 40 MHz bunch crossing rate, with an average of 25 $p-p$ collision per bunch crossing. It is both technologically impossible and cost prohibitive to transfer and record all the collision information given an event rate of about 1 GHz and an event size of approximately 1.5 MB. Moreover, the collision rate is dominated by non-interesting physics (as a comparison, Higgs boson production rate is only $\sim$0.6 Hz at the design luminosity) and most of them can be rejected. Still, all of the interesting data needs to be stored. To achieve this, ATLAS has developed a highly efficient, flexible and robust trigger system to make the online selection of particle collisions potentially containing interesting physics. The purpose of the trigger system is to constrain the an output rate of recording to be less than 200 Hz (limited by the computing resources for offline storage and processing of the data).

A schematic diagram of the ATLAS trigger is shown in Figure 2.10. The ATLAS trigger system [86] consists of three triggers, namely the Level-1 (L1) trigger [87], the Level-2 (L2) trigger and the Event Filter (EF). Together, L2 and EF are called the High Level Trigger (HLT). At each step, more refined and complex decisions (longer latencies) are made to successively lower data rates. As long as the buffers do not fill up (overflow), no additional dead-time is introduced.

The L1 trigger is a hardware-based, pipelined system designed to reduce the rate to a maximum of 75 kHz. It uses signals coming primarily from the dedicated trigger hardware in

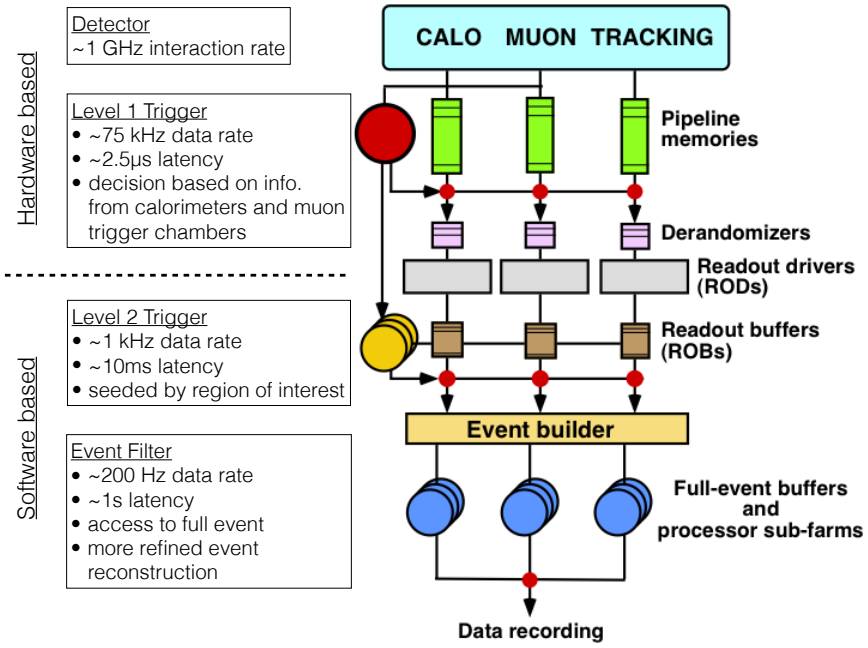


Fig. 2.10 Schematic diagram of the ATLAS trigger system. Figure adapted from Reference [86]

the calorimeter and the muon sub-detectors. It is not practical to make a trigger decision in the time between bunch crossings because of the very short (25 ns) bunch-crossing interval. Hence the information for all detector channels has to be retained in the pipeline memories on the detector electronics while the trigger decision is being formed. The L1 latency, which is the time from the collision until an event is accepted by the L1 trigger is set at 2.5 μs (with a target latency of 2.0 μs and 0.5 μs as contingency). In order to achieve a latency of less than 2.5 μs, the L1 trigger system is implemented in the fast custom electronics.

Most of the rate reductions must be accomplished by the trigger system identifying events of interest. For this, the L1 trigger is supplied with a trigger menu, which is a list of up to 256 criteria upon which to determine if an event is accepted or not. These trigger criteria or items include combinations of candidate physics objects (signatures) such as muons, electrons, photons, jets, and τ-leptons decaying into hadrons (with specific transverse momentum threshold), as well as large missing and total transverse energy. Events from the pipelines that passed the L1 trigger selection are then transferred off the detector to the next stages of the detector-specific functional elements of the front-end systems called the Readout Drivers (ROD).

The L2 trigger is software-based system designed to reduce the L1 output rate to below 3.5 kHz with an average processing time of 40 ms/event. After the L1 trigger selection, the data rates are reduced but can still be massive ($\sim$100 GB/s). To cope with this high data rates, the L2 trigger selection is seeded by the Region of Interest (RoI) information provided by the L1 trigger. RoI is the region in the detector where the L1 trigger has identified interesting

features within the event. Thus L2 trigger can use the full granularity and precision of all the available detector data within the RoI. The resulting total amount of RoI data is minimal (a few percent of the L1 trigger throughput). L2 uses specialised fast reconstruction algorithms (for example the p_{T} of the L2 stand-alone-muon is reconstructed base on simple parameterised functions) to compute event feature quantities associated with the RoI. It is optimised for timing performance. Events passing the L2 trigger are then transferred to the event-builder system and subsequently to the EF for final selection.

The final stage of the event selection is carried out by the EF. The EF is designed to reduce the rate to 200 Hz with an an event processing time of 4 s/event, averaged overall events. It gets full event information from the event builder which assembles all event fragments from the readout buffer. The EF is mostly based on the same offline reconstruction algorithms that analyse data on large computing farms. It reconstructs the event with greater precision than at the L2 trigger stage. Finally events selected by the EF are formatted as raw data prior to being moved to permanent storage at the CERN computer centre.

Chapter 3

Data and Monte Carlo samples

In Section 3.1, we will provide details about the datasets that are used in this analysis. To make predictions for signal and backgrounds processes in a $p-p$ collision, Monte Carlo (MC) event generators are used. The process in which these signal and background events are simulated as well as the generators that are used for the event generation will be clarified in Section 3.3 and 3.4 respectively.

3.1 Collision data

The data used in this analysis were recorded by the ATLAS experiment during the 2012 LHC Run I with a centre-of-mass energy of 8 TeV. It corresponds to an integrated luminosity of 20.3 fb^{-1} [88]. The data were recorded by the ATLAS detector between April and December of 2012. Datasets are labeled as A, B, C, D, E, G, H, I, J, and L for different data taking periods[1]. These datasets are summarised in Table 3.1 with details of the data taking periods, their corresponding run numbers[2], integrated luminosities and total number of events for each data taking period. For the datasets to be deemed good for physics analysis, they must pass ATLAS Data Quality (DQ) requirements. There are more than 100 such DQ criteria, which can be roughly categorised in several groups:

- Global monitoring — This criterion confirm that DQ information has been evaluated and reviewed by experts. It also acts as an indication that no anomalous behaviour of the magnets and luminosity detectors are observed.

[1] Data taking periods are designed such that they represent data with a coherent configuration of the detector and the trigger. Any significant changes to either the detector configuration/calibration or to the trigger should usually cause the definition of a new period.

[2] The data recorded by the ATLAS detector in each data taking period is further organised into runs. Each run usually corresponding to a fill of the LHC. The runs are subdivided into luminosity blocks which typically has a time interval of one minute length during which the beam and detector conditions is approximately constant. This allows quantities such as the number of interactions per bunch crossing, $\langle\mu\rangle$, averaged over a luminosity block, to be calculated. Luminosity blocks are also used in the evaluation of data quality criteria.

Table 3.1 List of data samples used in the analysis.

Period	Dates	Run Numbers	L [$pb^{-}1$]	Number of Events
A	Apr-04 : Apr-20	200804 : 201556	794.80	187,250,442
B	May-01 : Jun-18	202660 : 205113	5104.55	652,495,436
C	Jul-01 : Jul-24	206248 : 207397	1408.33	278,991,901
D	Jul-24 : Aug-23	207447 : 209025	3296.18	455,776,019
E	Aug-23 : Sep-17	209074 : 210308	2531.50	372,901,528
G	Sep-26 : Oct-08	211522 : 212272	1276.43	190,494,845
H	Oct-13 : Oct-26	212619 : 213359	1447.39	272,481,044
I	Oct-27 : Nov-02	213431 : 213819	1019.28	155,753,940
J	Nov-02 : Nov-26	213900 : 215091	2603.29	415,159,353
L	Nov-30 : Dec-06	215414 : 215643	824.12	133,748,710
A-L	Apr-04 : Dec-06	200804 : 215643	20323.9	3,115,053,218

- Detectors status — Data are required to satisfy conditions ensuring all essential elements of the ATLAS detector were operational while the data were collected. The monitoring are based on detector control conditions such as nominal voltages, temperature, humidity, etc. The dataset will be flagged if there is any possible hardware and data-taking problem in any of the detector sub-systems.

- Trigger status — Mainly for assessing if the L1-triggers were working at a reasonable efficiency. It also monitors any timing, consistency, synchronisation or data problems.

- Objects reconstruction performance — Quality criteria are imposed to ensure that electrons, photons, muons, taus, jets, and missing transverse energy reconstruction algorithms as well as b-tagging algorithms behaved as expected. Checks are performed on the relevant variables to ensure that they show sensible distributions.

The physics events collected by the detector are subdivided into streams according to the classes of trigger signatures, e.g. minimum bias (`physics_MinBias` stream), jets/taus/missing energy (`physics_JetTauEtmiss` stream), electron/photon (`physics_Egamma` stream), muons (`physics_Muons` stream). An event can have more than one trigger signature. Hence it can flow in several streams. In this analysis, three different data streams selected by using the lowest unprescaled[1] triggers are used. To select the signal `physics_JetTauEtmiss` stream is used. For backgrounds estimation, `physics_Egamma` stream, and `physics_Muons` stream are used. The details of the trigger used in each data stream will be further explained in Section 4.11.

[1] An artificial random data drop at the trigger decision level to reduce the total trigger rate so that it stays below the acceptable level.

3.2 Event simulation

To simulate a complete event that may happen in a $p-p$ collision, MC generators are used. The simulation procedures can be roughly grouped into three steps (as illustrated in Figure 3.1), that is event generation, followed by detector and physics interaction simulation and finally digitisation of the energy deposited in the detector into voltages and currents for comparison to the readout of the ATLAS detector. In this section we will briefly discuss what happens in each of these steps. For a more complete discussion of the ATLAS simulation infrastructure, see Reference [89].

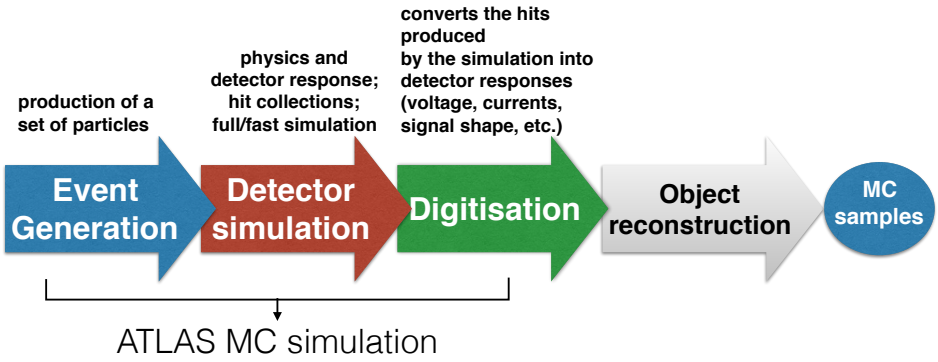


Fig. 3.1 A flow chart depicting the ALTAS MC samples production steps.

Event generation

Event generation consists of simulating $p-p$ collisions to produce hypothetical particles with the distribution predicted by the theory. This includes the simulation of the hard scatter process, initial state radiation (ISR), parton showers, multiple interactions, the hadronization processes, underlying events, and unstable particle decays. The MC generators that are used to simulate signals and backgrounds processes considered in this work include PYTHIA8.175 [90], SHERPA [91], ACERMC [92], POWHEG [93], MC@NLO [94–96], and MADGRAPH [97].

Detector simulation

The generated particle-level events are fed into the ATLAS simulation software packages through interfaces in the ATLAS software framework [98, 99]. Different simulation strategies with varying degrees of accuracy and simulation speed are used to perform the simulation of the physics processes and detector response. Two such strategies relevant for generating the MC samples used in this analysis are the standard ATLAS full simulation (FULLSIM) [100, 101] and the ATLAS fast simulation (ALTFAST-II) [102].

FULLSIM uses full description (geometry, disabled parts, temperatures, high-voltage settings, and etc.) of ATLAS detector. It gives a complete description of the decays and the interactions of all the generated particles with all the active and passive detector material. The energy, position, and time associated with each hit are recorded. FULLSIM is accurate but very resource intensive. The approach adopted by the ATLFAST-II simulation is to use the full simulation only for the inner detector, muon system, and the other detector subsystems, while using a simplified modelling of the calorimeter geometry and replacing the simulation of the development of particle showers (energy deposition and shower properties, including fluctuations and correlation) in the calorimeter by parametrisations. This approach is intrinsically less accurate, but the simulation time is considerably reduced. The result of ALTFAST-II simulation has been validated against the Geant4 based full simulation for electrons, jets, and missing transverse energy [89, 103, 104]. The difference is found to be tolerable especially when the details of the ATLAS detector performance (shower energy profiles, cracks in the calorimeters, etc...) are not of crucial importance.

Digitisation

The ATLAS digitisation converts the hits produced by the simulation into signal that represent the detector responses. The signal may include records that the voltage or current on a particular readout channel has exceeded a preconfigured threshold within a particular time-window, or in some cases, the detailed signal shape over the time-window. Within a digitisation job, hits from the underlying and pile-up interaction are incorporated on top of that from the hard scatter. The distinct properties associated with each sub-detector (charge collection efficiencies, cross-talk, noise, and channel dependent variations in detector response) are modelled in sub-detector specific digitisation software. They are tuned to reproduce the detector response seen in test beam and cosmic ray data. The signal from the detector is converted to the same raw data format as that is written out from the detector's readout drivers. Finally these raw data is passed through the same trigger chain and is subjected to the same reconstruction algorithms that are applied to the real ATLAS data.

3.3 Simulated signal samples

MC signal samples are generated with MADGRAPH, interfaced to PYTHIA v8.175 using the AU2 parameter settings (tune) [105] for parton showering, hadronization, and underlying event simulation. The simulated samples are processed with a full ATLAS detector simulation. For all signal samples, the Higgs boson mass is fixed to 125 GeV. The CTEQ6L1 PDF set [106] is used for all the EFT samples. Furthermore, samples are produced for scalar and fermionic DM particle masses ranging from 1 to 1000 GeV. The couplings (Λ or λ) for each sample are chosen such that the kinematics of the EFT samples become only a function of the DM mass.

Table 3.2 List of signal samples used in the analysis. All signals were generated with MADGRAPH.

Operators	Masses [GeV]	Generator	σ_0 [pb]	Parameters
xxhh	$m_\chi = 1$	MADGRAPHv1.5.1+PYTHIA v8.175	1.52E-007	$\lambda = 0.0001$
	$m_\chi = 65$	MADGRAPHv1.5.1+PYTHIA v8.175	8.31E-012	$\lambda = 0.0001$
	$m_\chi = 100$	MADGRAPHv1.5.1+PYTHIA v8.175	3.05E-013	$\lambda = 0.0001$
	$m_\chi = 500$	MADGRAPHv1.5.1+PYTHIA v8.175	3.59E-017	$\lambda = 0.0001$
	$m_\chi = 1000$	MADGRAPHv1.5.1+PYTHIA v8.175	5.88E-020	$\lambda = 0.0001$
xxhhg5	$m_\chi = 1$	MADGRAPHv1.5.1+PYTHIA v8.175	1.09E-003	$\Lambda = 10$ TeV
	$m_\chi = 65$	MADGRAPHv1.5.1+PYTHIA v8.175	2.00E-007	$\Lambda = 10$ TeV
	$m_\chi = 100$	MADGRAPHv1.5.1+PYTHIA v8.175	2.44E-008	$\Lambda = 10$ TeV
	$m_\chi = 500$	MADGRAPHv1.5.1+PYTHIA v8.175	5.32E-011	$\Lambda = 10$ TeV
	$m_\chi = 1000$	MADGRAPHv1.5.1+PYTHIA v8.175	3.02E-013	$\Lambda = 10$ TeV
xdxhDh	$m_\chi = 1$	MADGRAPHv1.5.1+PYTHIA v8.175	1.08E-001	$\Lambda = 100$ GeV
	$m_\chi = 65$	MADGRAPHv1.5.1+PYTHIA v8.175	1.55E-002	$\Lambda = 100$ GeV
	$m_\chi = 100$	MADGRAPHv1.5.1+PYTHIA v8.175	9.33E-003	$\Lambda = 100$ GeV
	$m_\chi = 500$	MADGRAPHv1.5.1+PYTHIA v8.175	1.71E-004	$\Lambda = 100$ GeV
	$m_\chi = 1000$	MADGRAPHv1.5.1+PYTHIA v8.175	2.86E-006	$\Lambda = 100$ GeV
xgxFhDh	$m_\chi = 1$	MADGRAPHv1.5.1+PYTHIA v8.175	1.44E+001	$\Lambda = 100$ GeV
	$m_\chi = 65$	MADGRAPHv1.5.1+PYTHIA v8.175	1.34E+001	$\Lambda = 100$ GeV
	$m_\chi = 100$	MADGRAPHv1.5.1+PYTHIA v8.175	1.23E+001	$\Lambda = 100$ GeV
	$m_\chi = 500$	MADGRAPHv1.5.1+PYTHIA v8.175	2.34E+000	$\Lambda = 100$ GeV
	$m_\chi = 1000$	MADGRAPHv1.5.1+PYTHIA v8.175	1.46E-001	$\Lambda = 100$ GeV

Table 3.2 lists the MC signal samples used in this study, along with their mass parameter values and cross sections.

3.4 Simulated background samples

Table 3.3 summarises the various event generators and parton distribution function (PDF) sets, as well as parton shower and hadronization software used to simulate various background processes. Sometimes a filter is applied during event generation such that the generator will only generate events that pass some selection to increase statistics in a particular region of phase space. In this case, a single multiplicative factor called filter efficiency (that takes into account the fact we are not generating the full cross-section of that process) need to be applied during the analysis to get the correct normalisation. Some MC generators described above generate cross sections for the requested hard process that are correct at LO. However, for processes (e.g. W+jets and SM Higgs) where the cross-section at a higher order is non-negligible, this approximation is insufficient. Hence, cross sections are normalised or scaled by using a higher-order correction factor, called k-factor. It is defined as the higher-order cross section calculation divided by the lower-order cross section calculation for a related inclusive cross section. Further

details (i.e. filter, cross section, k-factor, filter efficiency, and number of weighted events) for each background MC sample are presented in Table A.1-Table A.7 in Appendix A.

The $W/Z/\gamma$+jet events are divided into non-overlapping "B filter", "C-jet filter & B veto", and "B & C-jet veto" samples, where the B filter requires the existence of B-hadrons within $|\eta| < 4$ and the C-jet filter demands the presence of C-hadron flavoured jets with $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 3$. This allowed the statistics of the critical $W/Z/\gamma$+heavy-flavor samples to be increased using less computational resources. In addition, these sample are produced in slices of vector boson transverse momentum: $p_{\mathrm{T}}^{V} > 0$ GeV, $70 < p_{\mathrm{T}}^{V} < 140$ GeV, $140 < p_{\mathrm{T}}^{V} < 280$ GeV, $280 < p_{\mathrm{T}}^{V} < 500$ GeV, and $p_{\mathrm{T}}^{V} > 500$ GeV. This is to ensure that the statistics in the most sensitive high p_{T}^{V} bins could be increased. To avoid overlap between the $p_{\mathrm{T}}^{V} > 0$ GeV with higher p_{T} extension samples, high p_{T} events in $p_{\mathrm{T}}^{V} > 0$ GeV sample are veto. For $t\bar{t}$ events, a filter which requires at least one W boson from top decays into charged leptons (e, μ, τ) is used. Similarly, the W boson produced in the single top events via the s- and the t-channel are required to decay leptonically, while the events of the Wt production are simulated inclusively. Higgs boson mass of 125 GeV is used consistently to generate all the SM Higgs backgrounds in $ZH \to \nu\bar{\nu}b\bar{b}$, $WH \to \ell\nu b\bar{b}$, and $ZH \to \ell^{+}\ell^{-}b\bar{b}$ channels, where $\ell = e, \mu, \tau$. The $H \to b\bar{b}$ decay branching ratios come from Reference [107, 108]. Furthermore, the samples are normalised using the central values of the cross sections calculated at next-to-next-to-leading-order (NNLO) in QCD corrections [109] and next-to-leading-order (NLO) in electroweak (EW) corrections [110]. No filter is used for the WW sample. For the ZZ sample, lepton, and missing transverse energy filters are used to increase the statistic of the ZZ samples.

Table 3.3 Summary of MC event generators, PDF sets, and parton shower and hadronization models utilized in the analyses for both the signal and background processes.

Model / Process	Generator	PDF	Parton Shower / Hadronization
$W/Z/\gamma$+jets	SHERPA v1.4.3	CT10	SHERPA v1.4.3
$t\bar{t}$	POWHEG r2129	CT10	PYTHIA v6.427 with P2011C tune
Single top (s-ch., Wt)	MC@NLOv3.31	CT10	JIMMY v4.31 with AUET2 tune
Single top (t-ch.)	ACERMCv3.8	CTEQ6L1	PYTHIA v6.426 with AUET2B tune
$WW/WZ/ZZ$	POWHEG r2330.3	CTEQ6L1	PYTHIA v8.175 with AU2 tune
$q\bar{q} \to Vh$	PYTHIA v8.175	CTEQ6L1	PYTHIA v8.175 with AU2 tune
$gg \to Zh$	POWHEG r2330.3	CT10	PYTHIA v8.175 with AU2 tune
QCD multi-jet	PYTHIA v8.160	CT10	PYTHIA v8.160 with AU2 tune

Chapter 4

Physics objects reconstruction and identification

Before selecting the $E_{\mathrm{T}}^{\mathrm{miss}} + h(\rightarrow b\bar{b})$ event it is essential to reconstruct, identify, and select the basic objects such as track, leptons, jets and so on. In each of the following section, we first examine how each object is reconstructed and identified. This is followed by a brief discussion of the role each object plays in this analysis as well as the requirements used to select them.

4.1 Track and vertex

The default track reconstruction employ an "inside-out"[1] strategy [111, 112]. The first step of the track reconstruction is the creation of space-point objects—three-dimensional representations of the pixel and SCT sub-detectors measurements. Three or more space-points are combined to form a track candidate. As the track candidate is propagated outwards from the innermost layer of the pixel detector, successive silicon hits are added to the candidate. Outlier hits that contribute to the large χ^2 of the track fit are removed. The candidate track is scored in a reward-penalty scheme with respect to one another. Each additional hit leads to a better score value. Hit from less precise detector parts is scored less. The χ^2 of the track fit is also used to penalise poor-quality candidate. After the reconstruction of track in the pixel and the SCT detectors, the successful candidate is extrapolated into the TRT volume and associated with its drift circles. The full track is refitted by combining the full information of all three ID sub detectors. Finally the newly fitted track is kept if it has a quality score which is higher than that of the original track candidate.

After the successful reconstruction, the track candidates passing the criteria listed in Table 4.1 are used as the input to the vertex reconstruction algorithms. The track reconstruction

[1] There is also a complementary "outside-in" track finding strategy. It searches for unused track segments in the TRT. Such segments are extended into the SCT and pixel detectors to improve the tracking efficiency for secondary tracks from conversions or decays of long-lived particles

efficiency with this selection is between 75% and 85% for tracks with p_{T} above 500 MeV and $|\eta| < 1.5$.

Table 4.1 Tracks selection criteria [113]. Tracks that pass these requirements are used as inputs to reconstruct primary vertices.

Observable	Cut value
p_{T} (MeV)	> 400
The number of SCT hits	≥ 4
The number of silicon detectors (pixel or SCT) hits	≥ 9
The number of missing hits in the pixel detector	$== 0$
Transverse impact parameter with respect to the primary vertex, $\lvert d_0 \rvert$ [mm]	< 1.5
Transverse impact parameter resolution, $\sigma(d_0)$ [mm]	< 5
Longitudinal impact parameters resolution, $\sigma(z_0)$ [mm]	< 10

To reconstruct the primary vertices, a seed is first formed around the beam spot using the reconstructed tracks at their respective points of closest approach to the centre of the beam spot. The optimal position of the primary vertices (PV) is found by performing an iterative χ^2 minimisation algorithm with the seed position as the starting point and parameters of reconstructed tracks as input. A weight (as a function of the fitted χ^2) is associated to each track to assess its compatibility with the vertex candidate. The vertex position is recalculated using the weighted tracks, and then the procedure is repeated, recalculating track weights with respect to the new vertex position. Tracks which are incompatible with the vertex by more than approximately seven standard deviations are removed from the vertex candidate. Finally all vertices with at least two associated tracks are retained as valid PV candidates.

In this analysis, the PV is required to have at least 5 associated tracks. Of all the reconstructed vertices, the one with the highest $\sum p_T^2$ of the associated tracks is selected as the hard-scatter vertex. The rest are categorised as the pile-up vertices.

4.2 Isolation

For prompt electrons, photons, and muons produced in the hard process, typically there is no energy deposited around them apart from low-energy activities coming from the underlying event, multiple interactions, and pile-up collisions. On the other hand, fake or non-prompt electrons, photons, and muons are typically accompanied by some additional energy coming from the jet. In order to select prompt objects, we define the "isolation" variables below, which are required to be smaller than a certain threshold.

i. *Etcone_XX* (*TopoEtcone_XX*) — This variable is defined as the sum of calorimeter cell energies (topological cluster transverse energies at EM scale[1]) inside a cone with radius

[1] See Appendix B.1 for more description.

$\Delta R = 0.XX$ around the cluster barycenter (from which the electron, photon or muon candidate is reconstructed). To remove the energy of the prompt electron, photon or muon, a $\Delta\eta \times \Delta\phi = 5 \times 7$ grid of cells (in the middle layer of EM calorimeter) cantered on the electron, photon or muon axis are excluded. All layers from the electromagnetic and hadronic calorimeters are used, except the crack scintillators.

ii. *Ptcone_XX* — The track track isolation variable, computed by summing the p_{T} of all ID tracks within a cone of $\Delta R = 0.XX$, centred around the lepton track direction. Electron or muon track are removed. All the tracks are required to pass the selection in Table 4.2. This variable is robust against pile-up for electrons and muons because of the impact parameter cuts which constrain the tracks to come from the same vertex associated to them.

Table 4.2 Tracks selection criteria. Tracks that pass these requirements are used as inputs to compute track isolation.

Observable	Cut value
p_{T} (MeV)	> 400
$\lvert\eta\rvert$	< 2.5
The number of pixel detector hits	≥ 1
The number of SCT hits	≥ 9
The number of missing hits in the pixel detector	$== 0$
Transverse impact parameter with respect to the primary vertex, $\lvert d_0\rvert$ [mm]	< 1.5
Longitudinal impact parameters with respect to the primary vertex, $\lvert z_0 * sin\theta\rvert$ [mm]	< 1.0

4.3 Electron

The reconstruction of electrons in the central detector region, $|\eta| < 2.47$, consists of three steps:

1. seed-cluster reconstruction — Seed cluster with a local maximum transverse energy, $E_T > 2.5$ GeV is searched for by a sliding-window with a size of 3×5 (in unit of 0.025×0.025, corresponding to the EM calorimeter middle layer granularity) in the $\eta - \phi$ space.

2. track-cluster matching — Selected tracks are extrapolated to the EM cluster barycentre. Tracks associated to the clusters are refitted using the Gaussian Sum Filter (GSF) [114] by taking into account the non-linear bremsstrahlung effect.

3. electron candidate reconstruction — An electron is reconstructed if at least one track is matched to the seed closest to the seed cluster is chosen. The four-momentum of the electrons is computed by taking the final corrected EM cluster energy [115, 70] as its energy and the $\eta - \phi$ direction of the best track matched to the original seed cluster as its direction.

The electron identification used in this analysis is based on sequential cuts on a total of 13 discriminating variables as detailed in Table B.1 of Appendix B. Those variables include the energy leakage into hadronic calorimeter, longitudinal and transverse shapes of the EM showers in the EM calorimeters, the properties of the tracks in the ID, the matching between tracks and energy clusters, and various other. Electrons (or photons) deposit their energy primarily in the EM calorimeter. Therefore the energy leakage to the hadron calorimeter is required to be small. Besides, hadrons tend to produce a broader transverse energy deposit in the calorimeter than electrons. Hence, the shower shape in the calorimeter can be used to discriminate electrons from jets. Further rejection of hadronic backgrounds can be achieved by imposing tighter requirements on the quality of the electron track, track-cluster matching, transverse impact parameter, and transition radiation in the TRT. To reject electrons from Dalitz decay and photon conversions, we require electron candidates to have tracks with a hit in the innermost layer of the pixel detector as well as the other layers of the silicon detectors.

The tag-and-probe method is used to measure the combine electron identification and reconstruction efficiency. $Z \to e\bar{e}$ and $J/\psi \to e\bar{e}$ events are selected using the strict selection criteria on one of the two electrons candidates (called "tag electron"). The second electron candidate (called "probe electron") can then be used for the efficiency measurements as the electron ID is not applied. The efficiency is defined as the fraction of probe electrons passing the tested criteria, which may include the reconstruction, the identification, and the additional components (e.g. trigger and isolation). The combined efficiency, which is the product of all components, is around 76% for electron with $7 \leq E_T \leq 20$ GeV and range from 76-90% for electron with $E_T > 20$ GeV, respectively [116–118].

The electron energy scale is calibrated using $Z \to e\bar{e}$ invariant mass peak while the resolution is calibrated by using the Z invariant mass width [119]. The energy scale determination is accurate to within 0.3×10^{-3} for $|\eta| < 1.37$, 2×10^{-3} for $1.37 < |\eta| < 1.82$, and 0.5×10^{-3} for $|\eta| > 1.82$. The resolution are about 0.8% on average in the barrel EM calorimeter, and about 1% in the endcap. The resolution determination is accurate to within 0.3% and 0.5% on average in the barrel and endcap EM calorimeter, respectively.

In this analysis, two categories of electrons, "baseline" and "isolated" are defined. Their definitions are summarised in Table 4.3. The baseline electrons are used in the object removal procedure and as an event veto during the signal event selection as well as during the construction of $t\bar{t}$, $W(\to \mu\nu)$+jets, γ+jets, and $Z(\to \mu\mu)$+jets control regions. The isolated electrons are used in the overlap removal procedure to remove fake jets from electrons candidates.

Table 4.3 Selection used to define the baseline and the isolated electrons.

Cut	Baseline electron (e)	Isolated electron (e_{iso})
$\|\eta\|$ range	< 2.47	< 2.47
p_{T} range [GeV]	> 7	> 20
Transverse impact parameter significance, $\|d_0/\sigma(d_0)\|$	-	< 5
Longitudinal impact parameter with respect to the primary vertex, $\|z_0 \times sin\theta\|$ [mm]	-	< 0.4
Ptcone_30/p_{T} electron	-	< 0.16
Etcone_30/p_{T} electron	-	< 0.18

4.4 Photon

Photons are classified into two main categories: converted and unconverted photons. The converted photons are those which convert in or upstream of the TRT tracker, while the unconverted photons are those which arrive at the EM calorimeter without creating an electron-positron pair. Both types of photons are reconstructed using essentially the same algorithm for electron reconstruction as described in Section 4.3. EM clusters without matching tracks are classified as the unconverted photon. On the other hand, the converted photons are characterised by the presence of at least one tracks which match to the EM clusters originating from a conversion vertex candidates inside the tracker volume.

To distinguish real prompt photons from background photons, a series of cuts on several discriminating variables as listed in Table B.2 in Appendix B are applied. One of the variables exploit the lateral and longitudinal shower shape of prompt photons in the EM calorimeter, which is typically narrower compared to the background. In addition, photons also have smaller leakage to the hadronic calorimeter compared to fake photons from jets, due to the presence of additional hadrons near the photon candidate in the latter case. Photons from isolated π^0 decays can be distinguished from the isolated prompt photons by identifying two separate local energy maxima in the finely segmented strips of the first layer of the EM calorimeter.

The photon identification efficiency is measured using a sample of photon candidates passing the isolation requirement, $Etcone\_04 < 4$ GeV. The identification efficiency is thus defined as the ratio of the number of photons passing the identification selection to the total number of isolated photons. The efficiency is shown to increases steeply from 50-65% (45-55%) for unconverted (converted) photons before reaching a plateau of about 94-100% for $E_T \gtrsim 40$ GeV [120–123]. The results from each method are consistent with each other within the uncertainties for the E_T regions in which the different measurements overlap.

The photon energy scale and energy resolution are calibrated by exploiting the Z resonance mass peak and width from large angle radiative Z decays ($Z \to \ell\ell\gamma$) [119]. The energy scale on average is of the order of 1%. Its determination is accurate to within 0.3% across the whole E_T

range. The energy resolution decreases from about 3% at $E_T \approx 10$ GeV to less than 1.5% for $E_T > 40$ GeV. The uncertainty for the energy resolution determination is about 10% at $E_T \approx 40$ GeV and increases steadily for higher E_T.

In this analysis, reconstructed photon are divided into "baseline" and "isolated" categories using the definition in Table 4.4. The baseline photon are only used in the overlap removal procedure whereas isolated photons are also only used to estimate the $Z(\rightarrow \nu\nu)$+jets background.

Table 4.4 Selection used to define sets of baseline and isolated photons.

Cut	Baseline photon (γ)	Isolated photon (γ_{iso})
$\|\eta\|$ range	< 2.37	< 2.37
p_{T} range [GeV]	> 10	> 125
TopoEtcone_40 [GeV]	-	< 5

4.5 Muon

The muon reconstruction starts by forming standalone tracks using the MS information. Once the standalone muon track is formed, it is extrapolated back to the vertex and match to the nearby ID tracks. The matched ID and MS tracks are combined to form the so called "combined muons" and are removed from the list. This process is repeated until no acceptable combination is obtained anymore. In order to recover the acceptance in the un-instrumented regions of the MS and for low p_{T} muons, the ID tracks that are not used as combined muon are extrapolated and matched to the track segment in the first station of the MS. The matching is derived by the χ^2 test. If an ID track is associated with at least one local track segment in the MS, the ID track is classified as a "segment-tagged muon" candidate. The segment-tagged muon kinematics is determined from the ID track.

The ID track associated with the CB and ST muon is required to pass a series of quality requirements in order to suppress fake tracks and discriminate against muons from π/K decays. These requirements include

- the number of pixel hits + the number of known dead pixel sensors crossed by the track[1] > 0
- the number of SCT hits + the number of known dead SCT sensors crossed by the track[1] > 4
- the number of missing hits in the pixel detector + the number of missing hits in the SCT < 3

[1] To reduce inefficiencies due to known inoperative sensors.

- the number of TRT hits on the muon track > 5 (for muons with $0.1 < |\eta| < 1.9$)
- the number of TRT hits nearby the muon track but not included in the track fit[1] $<$ 0.9*Number of TRT hits on the muon track (for muons with $0.1 < |\eta| < 1.9$)

The muon momentum scale and resolution has been studied in detail using $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ decays [124], similar to the electron energy scale and energy resolution measurement. The momentum scale is known within an uncertainty of $\pm 0.05\%$ for muon with $|\eta| < 1$. It increases to $\lesssim 0.2\%$ for muon with $|\eta| > 2.3$. The momentum resolution ranges from $\sim 1.7\%$ in the central region and at low p_{T} to $\sim 4\%$ at large η and $p_{\mathrm{T}} = 100$ GeV.

Similarly, "baseline" and "isolated" muons are defined. Their definition are listed in Table 4.5. Baseline muons are used in the overlap removal procedures and for rejecting events during signal event selection. Isolated muons are used in the construction of $t\bar{t}$, $W(\to \mu\nu)$+jets, and $Z(\to \mu\mu)$+jets control regions.

Table 4.5 Selection used to define sets of baseline and isolated muons.

Cut	baseline muon (μ)	Isolated muon (μ_{iso})
$\|\eta\|$ range	< 2.4	< 2.4
p_{T} range [GeV]	> 6	> 20
$\|z_0 \times sin\theta\|$ [mm]	-	< 0.4
$\|d_0/\sigma(d_0)\|$	-	< 3
Ptcone_30/p_{T} muon	-	< 0.12
Etcone_30/p_{T} muon	-	< 0.12

4.6 Jet

Jets are the experimental signatures of quarks and gluons. They are "spray" of roughly collinear colourless hadrons that are formed when colored partons from the $p-p$ collision evolve via fragmentation and hadronization process. Hence, jets manifest themselves as localised clusters of energy. Jets are the primary objects used for reconstructing the Higgs boson in this analysis. Hence high quality and highly efficient jet reconstruction is desired.

The input to the jet reconstruction algorithm can be any object with valid four-momentum representation, i.e. MC simulated objects such as partons and particles or reconstructed detector objects such as tracks and calorimeter clusters. All the jets that are used in this analysis,

[1] Also referred as TRT outliers. They can appear in two forms in the track reconstruction. First, as straw tubes with a signal from tracks other than the one in consideration. Second, as a set of TRT measurements which fail to form a smooth trajectory with the extrapolated track (formed from the pixel and SCT measurements). The latter case is typically attributed to a hadron decay-in-flight. It can be rejected by requiring that the outlier fraction (the ratio of outliers to total TRT hits) is less than 90%.

unless specified otherwise, are reconstructed from topological cell clusters (or topo-cluster). The topo-clusters are initially reconstructed at the EM scale, which correctly measures the energy deposited in the calorimeter by particles produced in EM showers. To correct this EM energy scale of the topo-clusters to the hadronic scale, we apply the so called "local hadronic cell weighting" (LCW) correction scheme. Henceforth, these corrected topo-clusters will be referred as "LCW topo-cluster". A detailed description on the topological cell cluster and the LCW correction scheme is given in Appendix B.1. The hadronic jet energy scale correction is a cluster-by-cluster correction, i.e. it does not make any assumption about the origin of the jet. To make sure that the final reconstructed jet closely represents the energy-momentum content of its original particles, some corrections are made by using both the data and MC. More detailed descriptions of these corrections are given in Appendix B.2. Following these corrections, the reconstructed jets are said to have been corrected to "LCW+JES" scale.

We use the anti-k_t algorithm [125] to reconstruct jets. It successively recombines pairs of input objects in an iterative procedure in order of decreasing or increasing relative transverse momentum. The clustering or combination procedure is terminated based upon a condition defined by a "distance" variable between two objects, d_{ij}, which is defined as

$$d_{ij} = min\left(p_{ti}^{2p}, p_{tj}^{2p}\right) \times \frac{\Delta R_{ij}^2}{R} \tag{4.1}$$

where $R_{ij}^2 = \Delta\eta_{ij}^2 + \Delta\phi_{ij}^2$ is the distance between the two objects in $\eta - \phi$ space. The radius parameter, R which determines the final size of the jet, is the only free parameter of the algorithm. The parameter p governs the relative power of the energy versus geometrical ΔR_{ij}^2 scales. Three variation of jet algorithms are possible depending on the value of variable p in the exponent of p_{ti} in Equation 4.1 and in d_{iB}. For the anti-k_t algorithm we use p = -1.

First the algorithm finds the minimum of the entire set $\{d_{ij}, d_{iB}\}$. Here $d_{iB} = p_{ti}^{2p}$ is the distance between the i-th object and the beam (thus the subscript B). If d_{ij} is the minimum then object i and j are combined into one new object k by using summation of four-momentum. Then object i and j are removed from the list of objects while object k is added. On the other hand, if d_{iB} is the minimum, object i is considered as a jet by itself. It is then removed from the list. This process is repeated for a new set of $\{d_{ij}, d_{iB}\}$ until either all objects are either jets by themselves or part of a jet.

Because of the $p = -1$ factor, the anti-k_t algorithm prefers to start the recombination from hardest objects. Furthermore, d_{ij} between similarly separated soft objects will generally be much larger than that between a similarly separated hard and soft objects pair. Therefore softer objects tend to cluster with harder ones long before they cluster among themselves, without modifying much either the shape or the momentum of the final jet. The regularity of the jets produced with such a method are illustrated in Figure 4.1.

The jet energy scale (JES) of all the reconstructed jets are validated using several in-situ (data-driven) methods. A combination of γ+jet and $Z(\rightarrow \ell\ell)$+jet events [126, 127] as well as

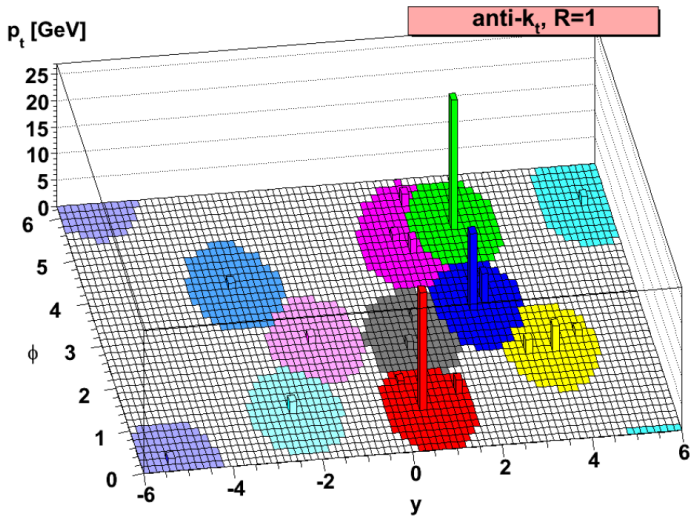


Fig. 4.1 Figure illustrates the regularity of the active catchment areas of the resulting hard jets obtained with the anti-k_t algorithm ($R = 1.0$). The jets are resulting from the clustering of a sample parton-level event made with a few hard particles and a large number of very soft particles that are randomly distributed. The hard jets are all circular, only the softer jets have more complex shapes. Figure taken from Reference [125].

multi-jet events [128] is used to cover the full momentum range, $20 \leq p_{\mathrm{T}} \leq 2000$ GeV. While the exact ways involved in each method may differ, they all rely on the same basic principle, i.e. exploiting transverse momentum balance between a jet and a well-measured reference object (γ, Z bosons or other jets) that recoils against the jet under study. More precisely, the average p_{T} response of the jet being studied and the reference object is compared in both the data and the MC to obtain the ratio of the jet p_{T} responses

$$\frac{\mathscr{R}_{data}}{\mathscr{R}_{MC}} = \frac{\langle p_T^{jet} / p_T^{ref} \rangle_{data}}{\langle p_T^{jet} / p_T^{ref} \rangle_{MC}} \tag{4.2}$$

where $\mathscr{R}$ is the jet p_{T} response, p_T^{jet} is the p_{T} of the jet and p_T^{ref} is the p_{T} of the well-known reference object. This quantity defines the final JES of the jet as well as the residual correction which is applied simply as a multiplicative factor to the p_{T}^{jet} in MC.

Figure 4.2 summarises the result of the Z+jet, γ+jet, and multi-jet balance analyses for jet in the central region. The plot shows the ratio of jet response in data and MC. It can be seen that the agreement between the data and MC is at 1% level across the whole p_{T} range. There is in general a good agreement between the three different in-situ methods in the regions of phase space where they overlap. A systematic uncertainty is assigned for these residual corrections.

Similar to the determination of the JES, the jet energy resolution (JER) is determined by studying the momentum balance between the jet and a well-measured photon or Z boson [127]

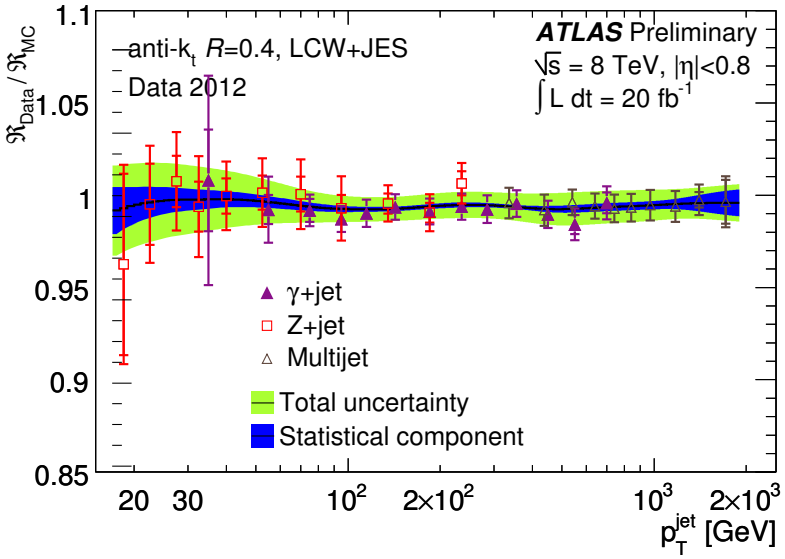


Fig. 4.2 Ratio of the response measured in data to the response measured in MC for each of the in-situ methods. Also shown are the uncertainties. Figure taken from Reference [129].

or another jet [128]. The JER can be estimated by using the width of the p_{T} response distribution between the jet and the reference object, $\$\sigma(p_{\mathrm{T}}) = \sigma(p_{\mathrm{T}}^{jet}/p_{\mathrm{T}}^{\gamma/Z/jet})$. The JER is determined to range from 25% at 20 GeV to 5% near 1 TeV. The uncertainty on the JER is less than 3% at 20 GeV and below 1% above 100 GeV.

To reject fake jets, a series of cuts on several discriminating variables as listed in Table B.3 in Appendix B.4 are applied. Calorimeter instrumental noise can lead to fake energy deposits in calorimeter cells, which can sometimes be reconstructed as fake jets. As energy deposits arising from real particles showering in the calorimeters produce a characteristic pulse shape, it can be used to separate ionisation signals from noise. To reject fake jets originating from beam-induced, and cosmic muons background, the jet energy deposits in the direction of the shower development can be employed. Furthermore, since real jets containing charged hadrons are usually also reconstructed by the tracking system while fake jets typically have no associated tracks, the jet charged particle fraction, defined as the ratio of the scalar sum of the p_{T} of the tracks associated with the jet divided by jet p_{T}, is another powerful variable to discriminate collision jets from fake jets. The jet quality selection efficiency is better than 99.8% for real jet with $p_{\mathrm{T}} > 20$ GeV while at the same time rejects most of the fake jets [130, 131].

Throughout the remaining of this thesis, jets, labelled as lower-case j, are defined as those jet that are reconstructed using the anti-k_t algorithm with a radius parameter $R = 0.4$ and LCW topo-clusters as the input. Jet with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 4.5$ are used in the analysis. In order to select jet that comes from the primary hard-scatter vertex but not a pile-up vertex, the

Jet Vertex Fraction (JVF) [132], defined as

$$\mathrm{JVF}(jet_j, vtx_k) = \frac{\sum_i p_T^{trk_{ijk}}}{\sum_k \sum_i p_T^{trk_{ijk}}}, \tag{4.3}$$

where trk_{ijk} is the i^{th} track matched to the j^{th} jet and associated to the k^{th} vertex, is required to to be greater than 0.5 for jets with $p_{\mathrm{T}} \leq 50$ GeV and $|\eta| < 2.4$.

4.7 Large-R jet

For a particle with a significant Lorentz boost, its decay products can be very collimated, resulting in a single jet if they are reconstructed by using the standard jet reconstruction algorithms with the radius parameter of 0.4. In order to accurately interpret the hadronic final state, it is essential to be able to resolve smaller angular distance. In principle, this can be done by pursuing smaller jet radius parameter but this approach is ultimately limited by the granularity of the calorimeter cell. Hence we opt to to reconstruct the parent particles with a larger jet radius parameter ($R = 1.0$). A larger jet radius is important as it allows the jet to capture enough of the hadronised particles for the accurate calculation of the jets mass and energy. Furthermore, compared to narrower jets, large-R jet has the advantage that it can be decomposed into subjets of varying sizes. Subsequently the information from the large-R jet properties and the substructure observables can be utilised to distinguish which parent particle the jet originates from. LCW topo-clusters are used as the input to reconstruct the large-R jets. Jet trimming is applied to remove the contribution from soft radioation. The trimming procedure first forms subjets of size $R_{sub} = 0.3$ from the constituents of a large-R jet by using the kt-algorithm[1]. Any subjet with p_{T} fraction relative to the parent jet less than 0.05 are removed. This allows us to remove the contributions from pile-up and the underlying events are preferentially removed. The remaining constituents form the trimmed jet. This procedure is illustrated in Figure 4.3. Similarly to the small radius jets, the reconstructed large-R jets need to go through a series of corrections. The details are provided as a supplement in Appendix B.3.

The JES and jet mass scale (JMS) of the large-R jets are measured using the γ+jets and dijet events similar to the in-situ techniques used for the small radius jet [133, 129]. Figure 4.4 shows the large-R jet average p_{T} response distribution as a function of p_{T}^{ref} (which is the p_{T} of γ) for the data and the MC simulation. A small differences of less than 1% for $p_{\mathrm{T}}^{ref} > 65$ GeV and in the central region ($\eta < 0.8$) is observed between the data and MC. The difference increases to about 2% for $0.8 < \eta < 1.2$. This difference is used as a correction factor to restore the agreement between the data and MC.

[1] $p = 1$ in Equation 4.1. Jet constituents with the smallest p_{T} tend to be clustered first, so that the highest p_{T} constituents are combined last.

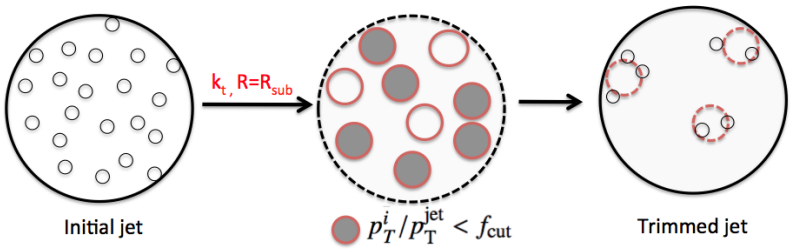
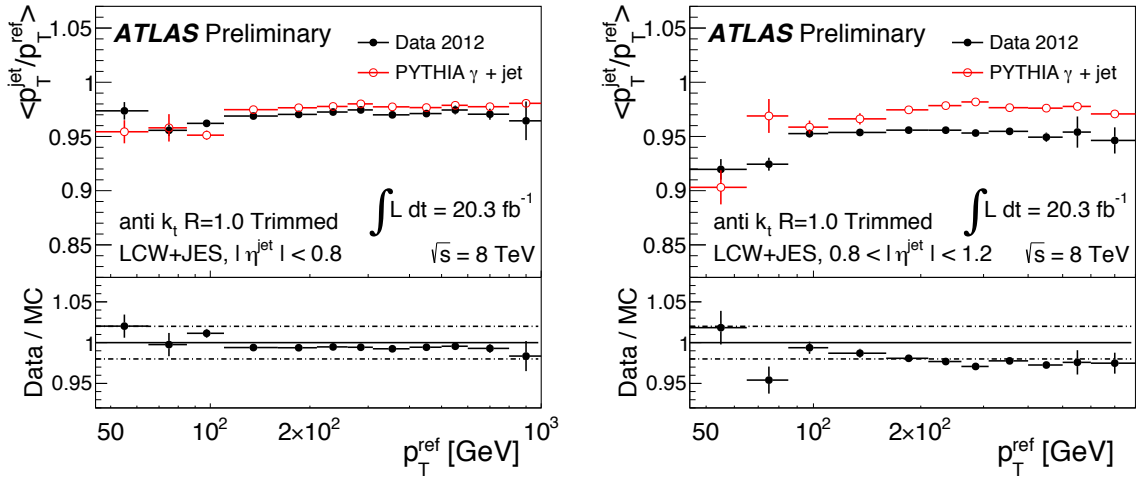


Fig. 4.3 A schematic diagram depicting the jet trimming procedure. The parameters $f_{cut} = 0.05$, and $R_{sub} = 0.3$. Figure taken from Reference [133].



(a) (b)

Fig. 4.4 The average jet energy response of large-R jets as a function of the reference object's (γ in this case) transverse momentum, p_{T}^{ref} for (a) $|\eta| < 0.8$ and (b) $0.8 < |\eta| \leq 1.2$. Only statistical uncertainties are shown. Figure taken from Reference [127].

The preliminary measurement using MC shows that the relative JER ranges between 5% for large-R jets with p_{T} around 300 GeV and around 3.5% for $p_{\mathrm{T}} > 900$ GeV for central jets within $|\eta| < 0.8$ [134]. The jet mass resolution (JMR) is also determined to be around 7-8% over the whole p_{T} and $|\eta|$ range [135] using MC study.

Throughout this thesis, the large-R jet is labelled as upper-case J. Large-R jets must have $p_{\mathrm{T}} > 300$ GeV and $|\eta| < 2.0$. The latter requirement is to ensure that the entire jet cone falls within the ID volume. A large-R jet must also pass the overlap removal rules specified in Section 4.12. On top of that, the leading large-R jet must have at least two track jets associated to its ungroomed parent jet. Track jet and its association technique will be explained in the next section.

4.8 Track Jets

Track jets, labelled as j^{trk}, are also reconstructed by using the anti$-k_t$ algorithm with the radius parameter, R=0.3. The input to the jets clustering algorithm is the inner detector tracks which satisfy the criteria listed in Table 4.6.

Table 4.6 Tracks selection criteria. Tracks that pass these requirements are used as input to track jet reconstruction.

Observable	Cut
p_{T} (MeV)	> 500
$\lvert\eta\rvert$	< 2.5
Pixel detector hits	≥ 1
SCT hits	≥ 6
Transverse impact parameter with respect to the primary vertex, $\lvert d_0\rvert$ [mm]	< 1.5
Longitudinal impact parameters with respect to the primary vertex, $\lvert z_0 \sin\theta\rvert$ [mm]	< 1.5

The impact parameter requirements ensure that only tracks matched to the primary vertex in the event are used for jet clustering. In this way the track jets are robust against the pile-up. Finally, only track jets with at least two tracks, $p_{\mathrm{T}} > 7$ GeV and $\lvert\eta\rvert < 2.5$ are considered for this analysis.

We need to know which track jet belongs to which large-R jet. For this purpose, we adopt a technique called "ghost-association" [136, 137] to associate track jets to a large-R jet. A track jet is treated as infinitesimally soft by artificially setting its 4-vector to a very small value (hence the name "ghost"). The calorimeter clusters and "ghost" are then reclustered using the anti$-k_t$ algorithm with the radius parameter, R=1.0. The low energy "ghosts" does not change the energy and the direction of the large-R jet even after the reclustering. If the track jet, treated as a "ghost" in the reclustering, is clustered into a given large-R jet, it is considered to be matched with the large-R jet.

4.9 b-tagged track jets

b-tagging—the identification of jets containing b hadrons is of paramount importance in this analysis. In this analysis, a neural network-based b-tagging algorithm called "MV1" is used. The inputs to this neural network is discussed below.

A b-hadron has a relatively long lifetime, of the order of 1.5 ps. With a moderate boost in the transverse direction, the average traveling distance before it decays can reach a few

millimetres. This leads to topologies with displaced secondary vertex[1] from the primary vertex. Several discriminant variables can be derived based on the following characteristics.

1. The signed impact parameter significance, $S \equiv (d_0/\sigma_{d_0};\ z_0/\sigma_{z_0})$ where the d_0 (z_0) is the transverse (longitudinal) impact parameter, while σ_{d_0} and σ_{z_0} is the uncertainty on the d_0 and z_0, respectively. The charged-particle tracks from b-hadron decay products tend to have large impact parameters with respect to the primary vertex. The impact parameters are given a sign, defined as

$$sign(d_0) = (\overrightarrow{P}_{jet} \times \overrightarrow{P}_{trk}) \cdot \left(\overrightarrow{P}_{trk} \times (\overrightarrow{X}_{pv} - \overrightarrow{X}_{trk})\right) \tag{4.4}$$

$$sign(z_0) = (\eta_{jet} - \eta_{trk}) \times z_0^{trk} \tag{4.5}$$

   where $\overrightarrow{P}_{jet}$ is the jet direction, $\overrightarrow{P}_{trk}$ and $\overrightarrow{X}_{trk}$ are the direction and the position of the track at the point of closet approach to the primary vertex, and $\overrightarrow{X}_{pv}$ is the position of the primary vertex. The track get a positive sign if its direction intersects the jet axis in front of the primary vertex and vice versa.

2. The number of secondary vertices with at least two tracks.

3. The secondary vertex mass, defined as the invariant mass of all charged particle tracks used to reconstruct the vertex.

4. The energy fraction, defined as the sum of energy of all tracks associated to the secondary vertex divided by the sum of the energies of all charged particles associated to the jet.

5. The number of tracks associated with the secondary vertices.

6. The number of additional single track vertices on the b-hadron flight axis.

7. The flight length significance, $L_{3D}/\sigma_{L_{3D}} > 2$, where L_{3D} is the three dimensional distance between the primary vertex and the point of closest approach of the track associated to the secondary vertex and $\sigma_{L_{3D}}$ is its resolution.

Base on the log-likelihood ratio (LLR) method, the measured value of a discriminating variable described above is compared to pre-defined probability density functions, $b(x_i)$ and $\ell(x_i)$ obtained from simulation for both the b- and light-flavour jet hypotheses. Each track or vertex can be assigned a weight which is defined as the ratio of the probabilities, for example,

[1] The secondary vertex is reconstructed from vertex with at least two tracks which are associated to the jet and far enough from the primary vertex ($L_{3D}/\sigma_{L_{3D}} > 2$). All tracks from the remaining two-track vertices are combined into a single inclusive vertex, using an iterative procedure to remove the worst track until the χ^2 of the vertex fit is less than a predefined threshold.

$b(x_i)/\ell(x_i)$. By taking the sum of weight of the tracks or vertices associated to a jet, the jet weight, a_{jet} can be calculated

$$a_{jet} = \sum_{i=1}^{N} ln\frac{b(x_i)}{\ell(x_i)} \tag{4.6}$$

where N is the number of tracks or vertices. This jet weight gives a better discriminating power between b- and light-flavour jets. The MV1 neural network takes a_{jet} constructed from the first four discriminant as inputs. In addition a neural network is trained by using the jet weight, a_{jet} constructed from the first discriminant, plus the last three discriminants as inputs. The outputs of this neural net constitute the other inputs to the MV1 neural network. The MV1 neural network is trained with b-jets as signal and light-flavour jets as background. A weight, w_{MV1} for each jet is computed by the neural network.

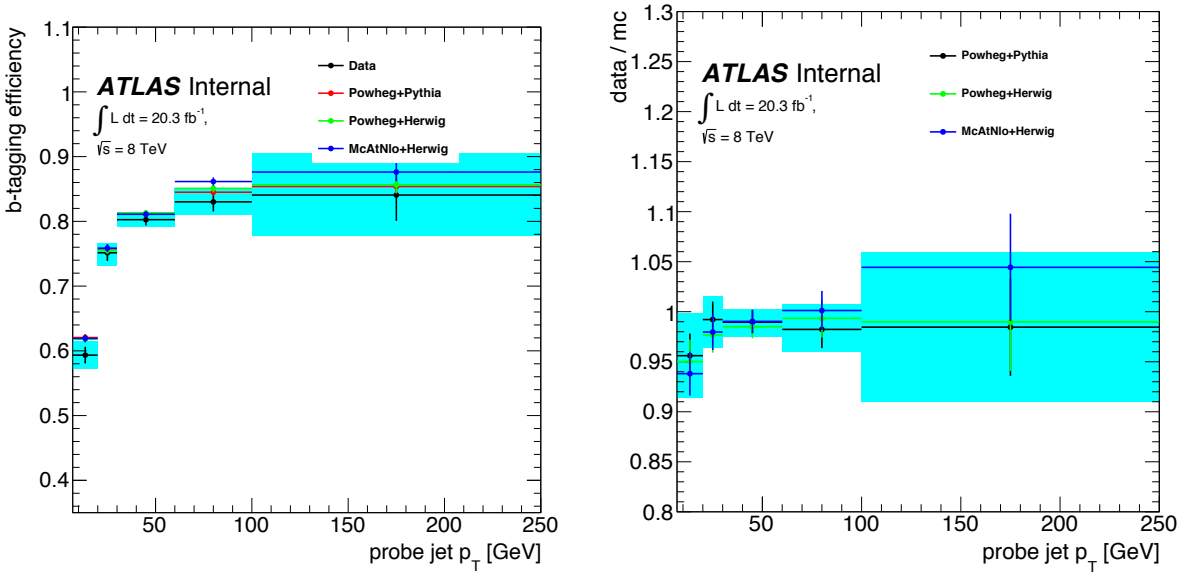
The MV1 algorithm can be used to tag either the calorimeter jet or track jet. The latter case is used in this analysis. One of the main advantages of using track jets for b-tagging is due its inherently better angular resolution than the calorimeter jets. This better angular resolution can greatly improve the b-tagging performance in dense regime. Since track jets are chosen to be originating from the primary vertex, this may reduce the dependence of b-tagging performance on pile-up. Other than that, we can also avoid introducing additional jet energy scale and jet energy resolution systematic uncertainties in b-tagging correction as track jets can be corrected independently from the calorimeter jets. In addition, track jets can recover low-p_{T} b-hadrons which otherwise will be removed as a result of a higher p_{T}-threshold imposed on the large-R calorimeter jets in the grooming procedures.

We require the MV1 b-tagging weight, $w_{MV1} > 0.7$, which gives on average an inclusive efficiency of 73.8% to tag a track jet that comes from b-hadron decays. The measurement of the b-tagging efficiency for track jets with radius R=0.3 has been performed using dilepton $t\bar{t}$ events selected from the complete 2012 ATLAS $p-p$ collision dataset [138]. The b-tagging efficiency is extracted based on the tag-and-probe method [139]. Events with with an opposite-sign eμ pair in the final state and exactly two jets are selected. One jet in the event is required to be b-tagged, allowing the second jet to be used as the "probe" jet, without biasing the b-tagging weight. These probe jets are used to measure the b-tagging efficiency as a function of jet p_{T} and η. The b-tagging efficiencies in data, MC, and the data/MC ratio as a function of track jet p_{T} and η are shown in Figure 4.5 and Figure 4.6.

By taking the ratio of the measured efficiency, ε, in data to that in MC, a per-jet scale factor is calculated as

$$SF_{flavour}(p_{\mathrm{T}},\ \eta) = \frac{\varepsilon^{data}_{flavor}(p_{\mathrm{T}},\ \eta)}{\varepsilon^{MC}_{flavor}(p_{\mathrm{T}},\ \eta)}, \quad flavour = b,\ c, \text{or } light \tag{4.7}$$

To correct for the b-tagging rate in MC to that in data, for each selected track jet in this analysis a jet weight, w_{jet}, is applied, depending on its flavour, p_{T}, and η. If the track jet passes the $w_{MV1} > 0.7$ requiremnt, the jet weight is simply the $SF_{flavor}(p_{\mathrm{T}},\ \eta)$, whereas if the track jet is



(a) (b)

Fig. 4.5 (a) MV1 b-tagging efficiency as a function of the jet p_{T} (for cut $w_{MV1} > 0.7$) measured in data and simulation (3 MC $t\bar{t}$ samples are compared) extracted with the tag-and-probe method. (b) The ratio, or scale factors, for R = 0.3 track jets. Error bars are statistical, shaded regions are statistical and systematics uncertainties added in quadrature. Figure taken from Reference [138].

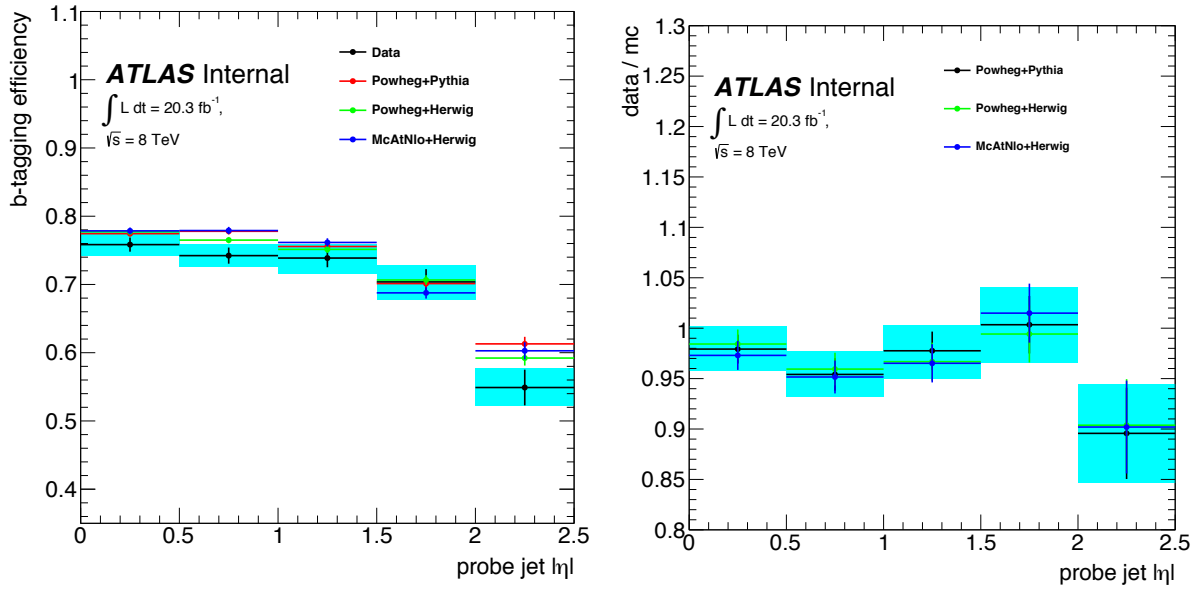
not tagged [140]

$$w_{jet} = \frac{1-\varepsilon^{data}_{flavor}(p_{\mathrm{T}},\ \eta)}{1-\varepsilon^{MC,\ i}_{flavor}(p_{\mathrm{T}},\ \eta)} = \frac{1-SF_{flavor}(p_{\mathrm{T}},\ \eta)\varepsilon^{MC,\ i}_{flavor}(p_{\mathrm{T}},\ \eta)}{1-\varepsilon^{MC}_{flavor}(p_{\mathrm{T}},\ \eta)} \tag{4.8}$$

where $\varepsilon^{MC,\ i}_{flavor}(p_{\mathrm{T}},\eta)$ corresponds to the tagging rate for the specific MC sample under consideration. While the ratio of efficiencies defining $SF_{flavor}(p_{\mathrm{T}},\eta)$ is less dependent on MC sample, the $\varepsilon^{MC,\ i}_{flavor}(p_{\mathrm{T}},\ \eta)$ is more dependent on possible variation in event topology (e.g jet multiplicity, effect of nearby jets) or differences between implementations of parton showering, hadronisation, and decays. The latter form in Equation 4.8 ensures any sample-dependence of the tagging efficiency is taken into account. Finally an event weight is computed as the product of all jet weights for all selected track jets

$$w_{event} = \prod_{track\,jets} w_{jet} \tag{4.9}$$

and applied to each event.



(a) (b)

Fig. 4.6 (a) MV1 b-tagging efficiency as a function of the jet $|\eta|$ (for cut $w_{MV1} > 0.7$) measured in data and simulation (3 MC $t\bar{t}$ samples are compared) extracted with the tag-and-probe method. (b) The ratio, or scale factors, for R = 0.3 track jets. Error bars are statistical, shaded regions are statistical and systematics uncertainties added in quadrature. Figure taken from Reference [138].

4.10 $E_{\mathrm{T}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$

The $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed based on energy deposits in the calorimeter. Its components along the x and y axes is calculated as:

$$\vec{E}_{x,y}^{miss} = \sum_{e} \vec{E}_{x,y}^{miss,e} + \sum_{\gamma} \vec{E}_{x,y}^{miss,\gamma} + \sum_{\tau} \vec{E}_{x,y}^{miss,\tau} + \sum_{jets} \vec{E}_{x,y}^{miss,jets} + \sum_{\mu} \vec{E}_{x,y}^{miss,\mu} + \sum_{SoftTerm} \vec{E}_{x,y}^{miss,SoftTerm} \tag{4.10}$$

The symbol $E_{\mathrm{T}}^{\mathrm{miss}}$ is used to denote its magnitude. In the object reconstruction, calorimeter cells may be shared between different objects. To avoid energy double-counting, calorimeter cells are associated preferentially with a parent physics object in a specific order: electrons (e), photons (γ), the visible parts of hadronically decaying τ-leptons, jets, and finally muons (μ). If a calorimeter cell is already assigned to one object, it is not reassigned or counted a second time. All energy deposits in the calorimeter that are not associated to the objects that are already used in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, are assigned as the calorimeter soft term (CST) in Equation 4.10.

The resolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ has been evaluated by using $Z(\rightarrow ll)$+jets events [141, 142] where no genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected. The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution is estimated from the width of the combined distribution of the E_x^{miss} and E_y^{miss} components. At the average pile-up condition of

2012 run period, the resolution for $E_{x,y}^{miss}$ is about 20 GeV for total transverse energy in the detector, $\sum E_T > 50$ GeV. The agreement between the data and MC is within 5% for the whole $\sum E_T$ range. In addition, the balance of $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ against the transverse momentum vector of the Z boson, $\vec{p}_T^Z$, is used to evaluate the $E_{\mathrm{T}}^{\mathrm{miss}}$ response, R, defined as

$$R = \frac{\vec{Z} \cdot \hat{u}_Z}{p_{\mathrm{T}}^Z} \tag{4.11}$$

where $\vec{Z} = \vec{E}_{\mathrm{T}}^{\mathrm{miss}} + \vec{p}_T^Z$[1] and $\hat{u}_Z$ is the unit vector of $\vec{p}_T^Z$. A bias in the $E_{\mathrm{T}}^{\mathrm{miss}}$ response thus implies a systematic under- or over-estimation of the terms used in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. The $E_{\mathrm{T}}^{\mathrm{miss}}$response measured in the $Z \to ll$ inclusive jet sample shows an overall value of around 0.9 for $p_T^Z > 50$ GeV. The agreement between data and MC is within 5%.

While $E_{\mathrm{T}}^{\mathrm{miss}}$ is measured based on the calorimeters information, $p_{\mathrm{T}}^{\mathrm{miss}}$ is calculated by using the isolated track momenta measured with the Inner Detector. Since $p_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed from isolated tracks that are associated to the primary vertex where the hard interaction has occurred, this quantity provides a less pile-up dependent description of the physics in the event and more correlated to the true $E_{\mathrm{T}}^{\mathrm{miss}}$ of the event. This is in contrast to the calorimeter-based $E_{\mathrm{T}}^{\mathrm{miss}}$ which includes calorimeter deposits originating from all interaction vertices (therefore dependent on the number of vertices in the event). However, there are some disadvantages. $p_{\mathrm{T}}^{\mathrm{miss}}$ has a smaller geometrical coverage and does not contain the information of neutral particles. To ensure a good association to a primary vertex, tracks that are considered in $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation are required to satisfy the criteria listed in Table 4.7. With this collection of tracks, the transverse momentum imbalance of one collision event is calculated as the negative sum of all selected tracks:

$$\vec{p}_{x,y}^{miss} = -\sum_{tracks} \vec{p}_{x,y}^{track} \tag{4.12}$$

where the $\vec{p}_x^{miss}$ and $\vec{p}_y^{miss}$ is the x and y components of the nominal $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ repectively.

Table 4.7 Track selection for $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation.

Observable	Requirement
p_{T} (MeV)	> 500
$\lvert\eta\rvert$	< 2.5
Pixel detector hits	≥ 1
SCT hits	≥ 6
Transverse impact parameter with respect to the primary vertex, $\lvert d_0\rvert$ [mm]	< 1.5
Longitudinal impact parameters with respect to the primary vertex, $\lvert z_0 \sin\theta\rvert$ [mm]	< 1.5

[1] Since the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$includes a negative vector sum over the lepton momenta from Z, the addition of $\vec{p}_T^Z$ removes its contribution.

4.11 Triggers

The sequence of triggers (L1→L2→EF) that is used to select events with a specific detector signature, for example events with a single high-p_{T} muon, is referred to as a trigger chain. The trigger chains that are used to select muons, photons, and $E_{\mathrm{T}}^{\mathrm{miss}}$ in this analysis are described in the following paragraphs.

In this analysis, events containing single muon are used for the estimation and validation of $Z(\rightarrow \ell\ell)$+jets, $t\bar{t}$, single top, and $W(\rightarrow \ell\nu)$+jets backgrounds (as explained in Section 6). Events with muons are primarily selected using logical OR between two lowest un-prescaled single lepton triggers chain, namely the mu24i and mu36 trigger chain. The requirements for each trigger chain are listed in Table 4.8. The mu24i trigger chain is designed to select isolated muons with $p_{\mathrm{T}} > 25$ GeV while the mu36 trigger chain is designed to collect muons with large p_{T} without making an isolation requirement.

Table 4.8 Single muon trigger chain for selecting events containing muon(s).

Trigger Chain Name	Trigger	Selection criteria
mu24i	L1	Muon $p_{\mathrm{T}} > 15$ GeV and coincidence of hits across 3 layers of the RPCs or TGCs
	L2	One or more combined-muon with $p_{\mathrm{T}} > 22$ GeV
	EF	One or more combined-muon with $p_{\mathrm{T}} > 24$ GeV and Ptcone_20/muon $p_{\mathrm{T}} < 0.12$
mu36	L1	Muon $p_{\mathrm{T}} > 15$ GeV and coincidence of hits across 3 layers of the RPCs or TGCs
	L2	One or more combined-muon with $p_{\mathrm{T}} > 22$ GeV
	EF	One or more combined-muon with $p_{\mathrm{T}} > 36$ GeV

In this analysis, γ+jets events are used to estimate the $Z \rightarrow \nu\nu$ + jets background as discussed in Section 6.3. The high p_{T} photon are selected using the un-prescale g120 trigger chain. This trigger has almost full efficiency for photons with p_{T} greater than 125 GeV. The requirements for g120 trigger chain [143] are listed in Table 4.9. L1 photon trigger utilises the third sampling layer (trigger towers) of the EM calorimeter (see Figure 2.8) to identify the position of the region of interests. EM clusters are formed by identifying local maxima using a sliding a window with size $\eta \times \phi = 4 \times 4$. If any pair of towers of the central 2×2 trigger towers has a combined energy that passes the threshold of 30 GeV, the trigger is satisfied. At the L2 and EF level, the reconstructed photons have to pass the photon identification requirements.

Table 4.9 Single photon trigger chain for selecting events containing photon.

Trigger Chain Name	Trigger	Selection criteria
g120	L1	Sum E_T of any pairs of trigger tower at the window's core-region > 30 GeV
	L2	One or more photon with $p_{\mathrm{T}} > 120$ GeV and pass photon ID
	EF	One or more photon with $p_{\mathrm{T}} > 120$ GeV and pass photon ID (tighter than L2)

To select our signal events, a un-prescale $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger is used. The trigger chain is labelled as xe80_tclcw. The requirements for this trigger chain [144–146] are listed in Table 4.10. At

L1 $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated from calorimeter cell information. However, at L2 and EF trigger level, topo-clusters corrected to LCW scale are used instead. In all cases, muon information are not included in the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$during 2012 data taking. That means muons will thus show up as a source of $E_{\mathrm{T}}^{\mathrm{miss}}$in the trigger.

Table 4.10 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger chain for selecting events containing large $E_{\mathrm{T}}^{\mathrm{miss}}$.

Trigger Chain Name	Trigger	Selection criteria
xe80_tclcw	L1	$E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV
	L2	$E_{\mathrm{T}}^{\mathrm{miss}} > 55$ GeV
	EF	$E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV

As $E_{\mathrm{T}}^{\mathrm{miss}}$ is the main signature of our signal events, it is crucial to make sure that the $E_{\mathrm{T}}^{\mathrm{miss}}$EF trigger efficiency in data is reproduced in the MC. Hence, a study to estimate the offline EF trigger efficiency is performed. The goals are:

- to ensure the trigger is fully efficient at $E_{\mathrm{T}}^{\mathrm{miss}}$ above 200 GeV.
- to derive correction factor if necessary.
- to treat correctly any systematic uncertainties related to the trigger inefficiencies

We select an uncorrelated sample with respect to the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger from $W(\rightarrow \mu\nu)$+jets events which are triggered by the muon triggers. To select $W(\rightarrow \mu\nu)$+ jets events, the event must pass the selection as listed in Table 4.11. The trigger efficiency, $\varepsilon(E_{\mathrm{T}}^{\mathrm{miss}})$ is defined as the ratio

Table 4.11 Selection used to select a set of $W(\rightarrow \mu\nu)$+ jets events for the study of $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency.

Object	Selection
Muon triggers	pass mu24i || mu36 trigger (see Section 4.11)
Number of isolated muon, $\mathrm{N}_{\mu,iso}$	== 1
Muon's p_{T} [GeV]	> 25
Baseline electron, n_ℓ	== 0
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j_i)\rvert$	> 1.0
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert$	$< \pi/2$
Number of large-R jet, n_J	≥ 1
Leading large-R jet p_T, $p_T^{J_1}$ [GeV]	> 100
Number of associated track jet, $n_{j^{trk}}$	≥ 2
Number of associated b-tagged track jet, $n_{j_b^{trk}}$	== 0

of the events that are selected by the $E_{\mathrm{T}}^{\mathrm{miss}}$ xe80_tclcw EF trigger to the total event passing selection in Table 4.11:

$$\varepsilon(E_{\mathrm{T}}^{\mathrm{miss}}) = \frac{\text{events that passed xe80\_tclcw EF trigger}}{\text{total events}} \tag{4.13}$$

Figure 4.7 shows the comparison of 2012 data and MC trigger efficiency as a function of the $E_{\mathrm{T}}^{\mathrm{miss}}$ for W+jets events and its backgrounds such as diboson, top quarks, and their combination (all of which are selected using the same selection listed in Table 4.11). Note that while muons are included in the offline $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction, they are not used in the determination of $E_{\mathrm{T}}^{\mathrm{miss}}$ in both L2 and EF active triggers in 2012. Hence, to get a more accurate representation of the $E_{\mathrm{T}}^{\mathrm{miss}}$used in EF trigger, the muon contribution is removed from the offline $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction (cf. Equation 4.10). For $E_{\mathrm{T}}^{\mathrm{miss}}$ greater than 200 GeV, the trigger is fully efficient. Moreover, the data and MC agree almost perfectly. Therefore we chose $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV in the offline analysis without the need to apply any correction factors. The efficiency curve has been fitted using the error function:

$$f(x) = \frac{1}{2}\left(1 + Erf\left(\frac{E_{\mathrm{T}}^{\mathrm{miss}} - threshold}{\sqrt{2} \times width}\right)\right) \tag{4.14}$$

The fit range for each level is 100-500 GeV. The threshold (50% efficiency point) and the width of the turn on curve are estimated by the fit. The sample size before and after the trigger selection, the $E_{\mathrm{T}}^{\mathrm{miss}}$ value where the turn on curve reaches its plateau, the $E_{\mathrm{T}}^{\mathrm{miss}}$ value corresponds to an efficiency of 98%, as well as the fitted parameters for data and each MC background sample are tabulated in Table 4.12.

Table 4.12 Event yield, EF_xe80_tclcw trigger efficiency values as well as the fitting results for data and MC background samples.

	Data	W+jets	Diboson	top-quarks	Combination
Sample Size	6.05x10^6	5.04x10^6	4.22x10^4	4.61x10^5	5.54x10^6
Before Trig.	5.55x10^5	5.67x10^5	4.70x10^3	1.75x10^4	5.90x10^5
After Trig.	2.94x10^5	3.48x10^5	3.16x10^3	1.10x10^4	3.62x10^5
Max Eff. [%]	100.00	100.00	100.00	100.00	100.00
Etmiss @ 98%	153.62	146.30	146.64	164.90	147.25
chi^2/dof	7.59	14.49	1.52	2.39	15.51
threshold	96.95 $\pm$ 0.17	86.78 $\pm$ 0.19	87.74 $\pm$ 0.92	92.97 $\pm$ 0.84	86.50 $\pm$ 0.20
width	27.60 $\pm$ 0.19	28.98 $\pm$ 0.16	28.68 $\pm$ 0.78	35.02 $\pm$ 0.80	29.58 $\pm$ 0.17

Further studies on the properties of the xe80_tclcw EF trigger are performed in order to understand the correlations between the trigger efficiency and kinematic or pile-up conditions. The dependency of trigger efficiency on large-R jet multiplicity, leading large-R jet p_{T}, and number of primary vertices are shown in Figure 4.8. The efficiency shows a weak dependency on all three variables at low $E_{\mathrm{T}}^{\mathrm{miss}}$ region. For $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV, the agreement between the data and MC is very good.

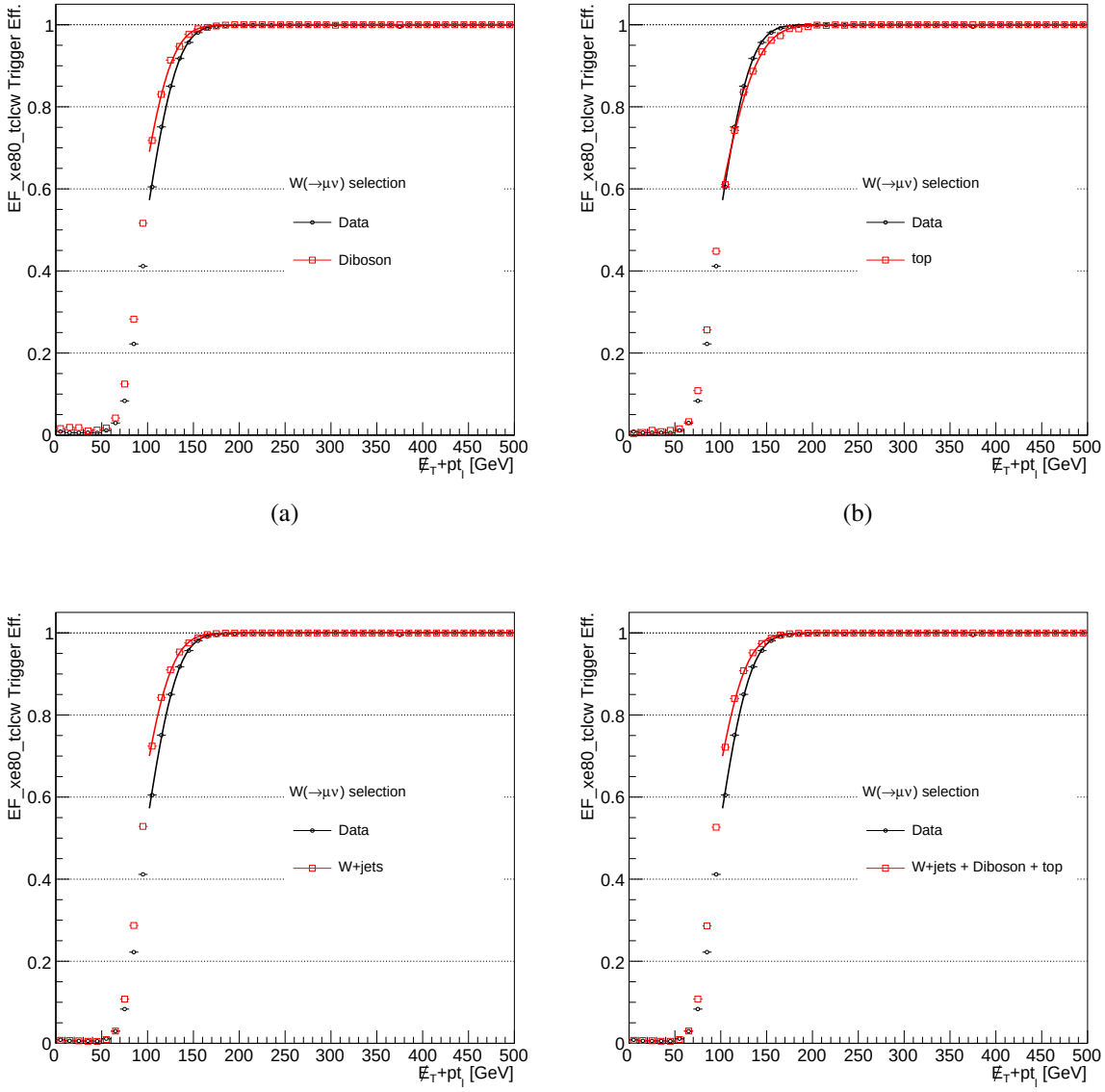


Fig. 4.7 xe80_tclcw EF trigger efficiency comparison between data and (a) diboson, (b) top ($t\bar{t}$ selection in Table 4.11), (c) inclusive $W(\rightarrow l\nu)$+jets and (d) combined MC.

4.12 Overlap Removal

The same calorimeter cell or track could be used in the reconstruction of multiple physics objects since the reconstruction process for each object proceeds independently. In addition, two separate but close-by objects also potentially introduce a bias in the reconstruction. To address these problems, a sequential procedures called “overlap removal" is implemented. The

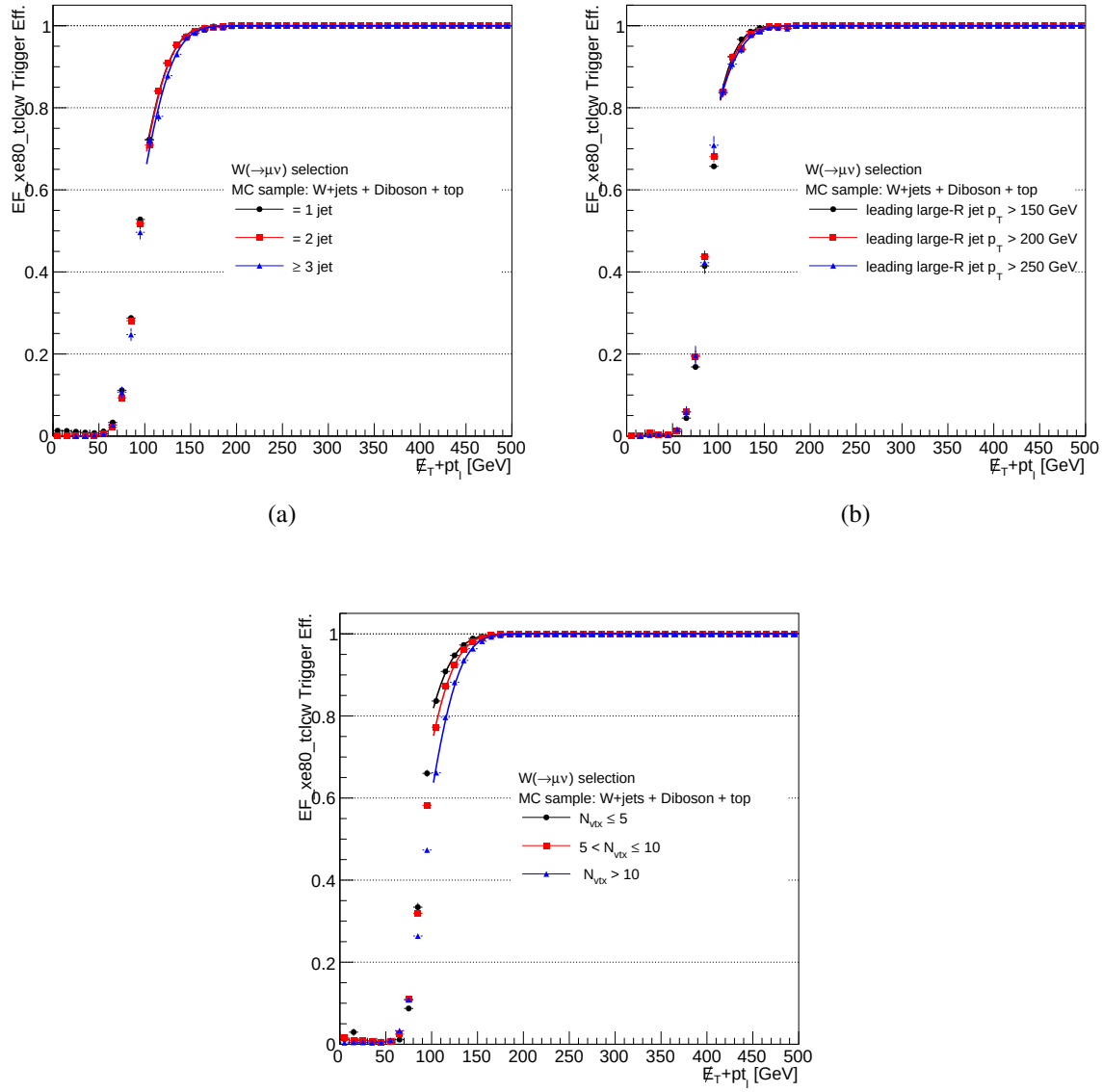


Fig. 4.8 EF_xe80_tclcw trigger efficiency dependency on (a) the large-*R* jet multiplicity, (b) leading large-*R* jet p_{T} and (c) the number of primary vertices.

geometrical proximity, ΔR, is used as the figure of merit to quantify the compatibility of two objects. Object pairs considered for the overlap removal include lepton-jet pairs, electron-muon pairs, and electron-electron pairs.

The electron-jet overlap removal aims at to remove reconstructed jets that are actually (isolated) electrons. To preserve heavy-flavour jets with semi-leptonic decays, the baseline electrons or muons in the proximity of a jet is classified as non-isolated and removed. While

electron-electron duplication is not frequent, the softer or equally energetic electron from the electron pairs is removed. Furthermore, large-R jets are eliminated if an isolated photon is found within $\Delta R < 1.0$ of the large-R jet. Track-jets are discarded if an isolated electron or an isolated muon is found within $\Delta R < 0.1$ of the track-jet. A muon-electron duplication may occur when the collinear final state radiation produces a photon very close to the muon track. In this analysis, the event is discarded to protect against a bias in the reconstruction of the muon momentum in presence of hard photon radiation. The hierarchical sequence in which the ambiguities are resolved and the ΔR requirement is summarised in Table 4.13.

Table 4.13 Overlap hierarchy and the order in which the ambiguity is checked.

Order	Case	Keep γ	Keep μ	Keep e	Keep jet	Keep event
1	$\Delta R(\text{jet}, e_{iso}) < 0.2$	–	–	Yes	No	–
2	$\Delta R(\text{jet}, \gamma_{iso}) < 0.2$	Yes	–	–	No	–
3	$\Delta R(\text{jet}, e) < 0.4$	–	–	No	Yes	–
4	$\Delta R(\text{jet}, \mu) < 0.4$	–	No	–	Yes	–
5	$\Delta R(e, e) < 0.05$	–	–	Yes (more energetic e)	–	–
6	$\Delta R(\text{large-}R\text{ jet}, \gamma_{iso})$	Yes			No	–
7	$\Delta R(\text{track jet}, e_{iso})$	–	–	Yes	No	–
8	$\Delta R(\text{track jet}, \mu_{iso})$	–	Yes	–	No	–
9	$\Delta R(\mu, e) < 0.01$	–	–	–	–	No

Chapter 5

Event selection

In this section, we will lay out the exact treatment and criteria to select the signal event. First, a general quality requirements, referred to as preselection criteria, are applied (Section 5.1). We take a short detour in Section 5.2 to investigate the unique signal event topology in the high-p_{T} (boosted) regime which motivates the use of large-R jets and track jets. The full selection criteria specific to the $E_{\mathrm{T}}^{\mathrm{miss}}$+2 b-jets final state are applied (Section 5.3). The optimisation procedures are outlined in Section 5.4. We also consider two event level corrections that need to be applied to all MC events so that it better describes the data (Section 5.5). Finally we show the signal selection efficiency in Section 5.6.

5.1 Event Preselection

Recall that the dataset used in this analysis must have passed the data quality requirement (Section 3.1). By design the data quality requirement does not identify issues which are sporadic in nature or issues which are noticeable only after integrating over a longer data taking period. This kind of problems are dealt with by imposing specific cleaning cuts, which include:

- In the luminosity block after a detector reset (to recover certain detector busy conditions), some events might be incomplete (due to some missing detector information in the event). Any events flagged with such error are rejected.
- Noise bursts events which show a substantial fraction of cells in the LAr calorimeter with unexpected signal shapes and high signals are discarded.
- Problematic events recorded when there are functioning errors in the tile calorimeter and events containing data corruption from one particular tile channel are vetoed

In addition, each event needs to be checked for badly reconstructed jets which may affect the accuracy of the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement. Event is rejected if:

- it contains any jet with p_{T} above 20 GeV that does not pass the jet quality requirements.

- it contains a jet (before the lepton-jet overlap removal) with $p_{\mathrm{T}} > 40$ GeV that has the relative energy of all non-operational calorimeter cells in a jet, $BCH_{jet} = \sum_{i}^{N_{badcells}} E^{i}_{cell}/E_{jet} > 0.05$ and points roughly in a similar transverse direction with the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ (i.e $\Delta\phi(jet, E_{\mathrm{T}}^{\mathrm{miss}}) < 0.3$) are removed [126].

To avoid using cosmic muons, events containing a muon that survives overlap removal and has a transverse (longitudinal) impact parameter with respect to the primary vertex larger than 0.2 (1) mm are rejected. Finally, to ensure the event contains a hard collision, at least one primary vertex with more than 4 associated tracks is required.

5.2 Event topology and boosted higgs selection

Figure 5.3 shows the comparison of the true missing transverse energy, $E_T^{miss,truth}$, distribution for each sample and for five different DM masses, where $E_T^{miss,truth}$ is defined as the magnitude of the vector sum of p_{T} of all non-interacting true particles with $|\eta| < 4.5$, which is the coverage of the ATLAS calorimeter. As expected, the $E_{\mathrm{T}}^{\mathrm{miss}}$ increases as DM mass increases. The other observation we can make is that the $E_T^{miss,truth}$ spectrums are very broad and generally larger than 200 GeV. Figure 5.2 shows the $\Delta\phi(E_T^{miss,truth}, Higgs)$ distribution. It is the separation between the $E_T^{miss,truth}$ and Higgs boson in the transverse plane. These $\Delta\phi(E_T^{miss,truth}, Higgs)$ distributions illustrate that for every sample the DM ($E_T^{miss,truth}$) is produced back-to-back with the Higgs boson. This also indicates that the p_{T} distribution of the Higgs boson is similar to that of the $E_{\mathrm{T}}^{\mathrm{miss}}$, as shown in Figure 5.3.

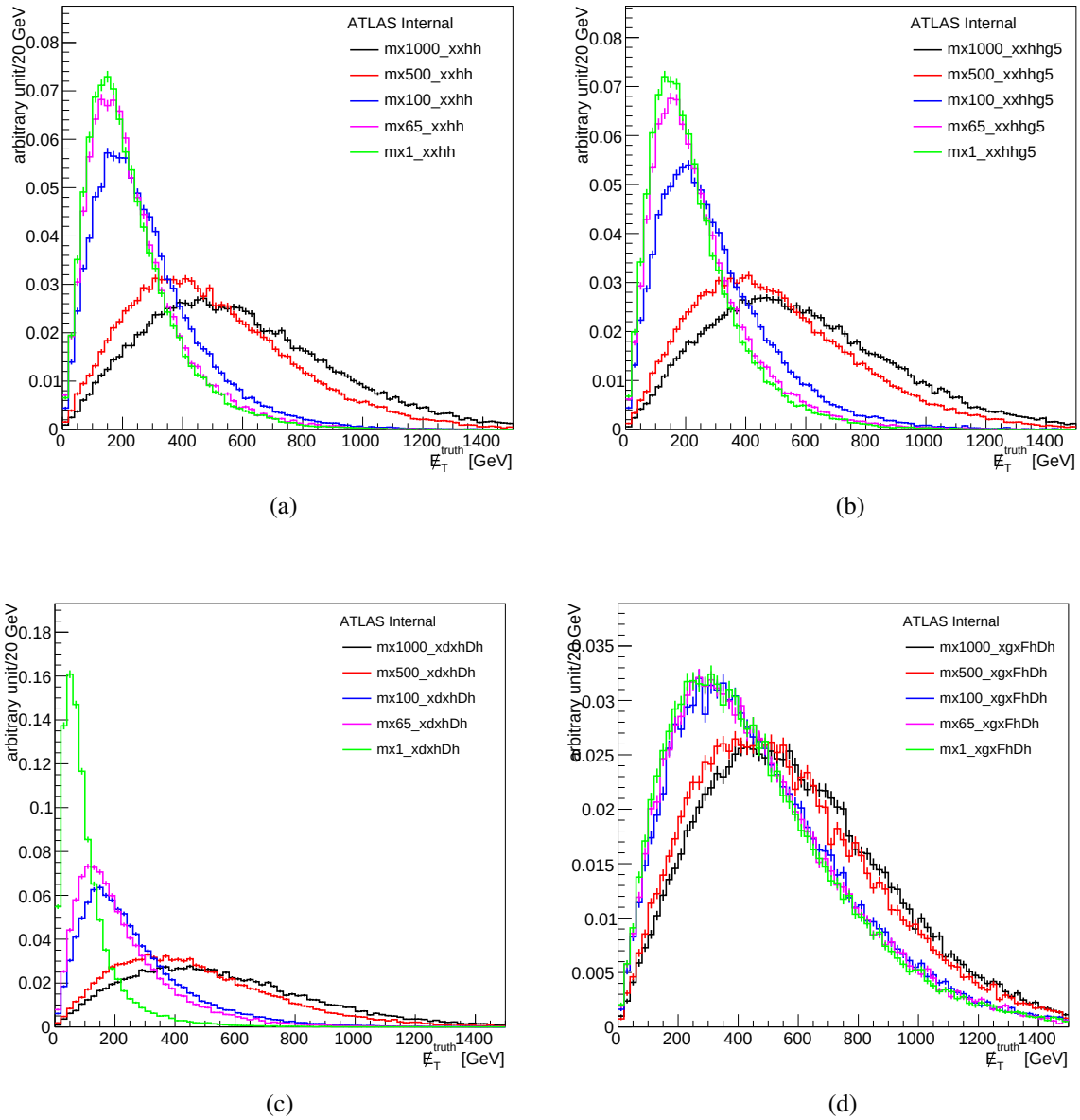
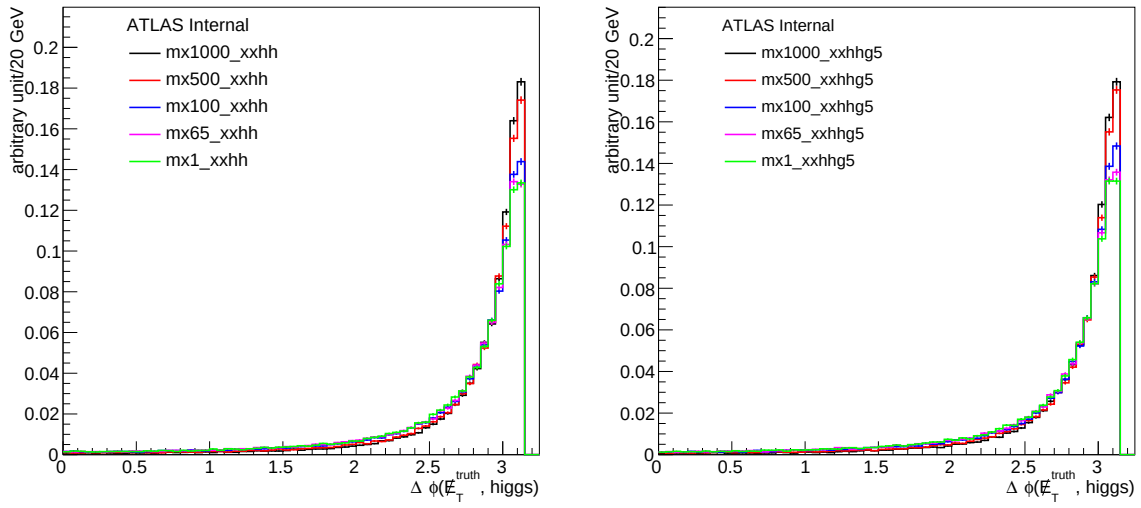
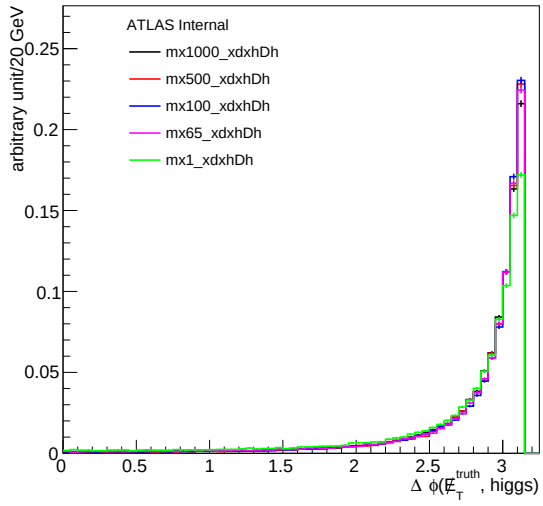


Fig. 5.1 $E_T^{miss,truth}$ distributions for five different DM masses (1, 65, 100, 500 and 1000 GeV) for sample (a) xxhh, (b) xxhhg5, (c) xdxhDh and (d) xgxFhDh.

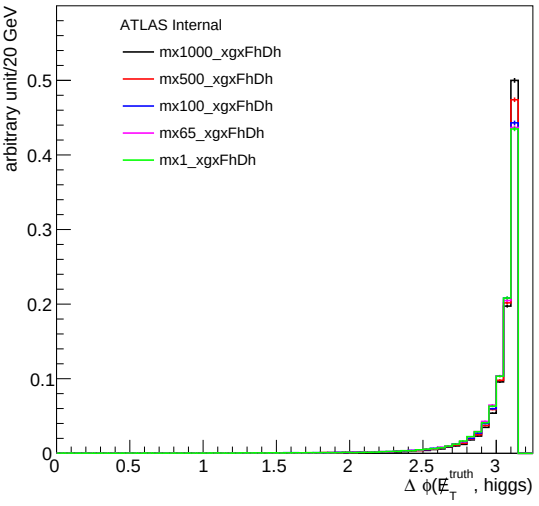


(a)

(b)

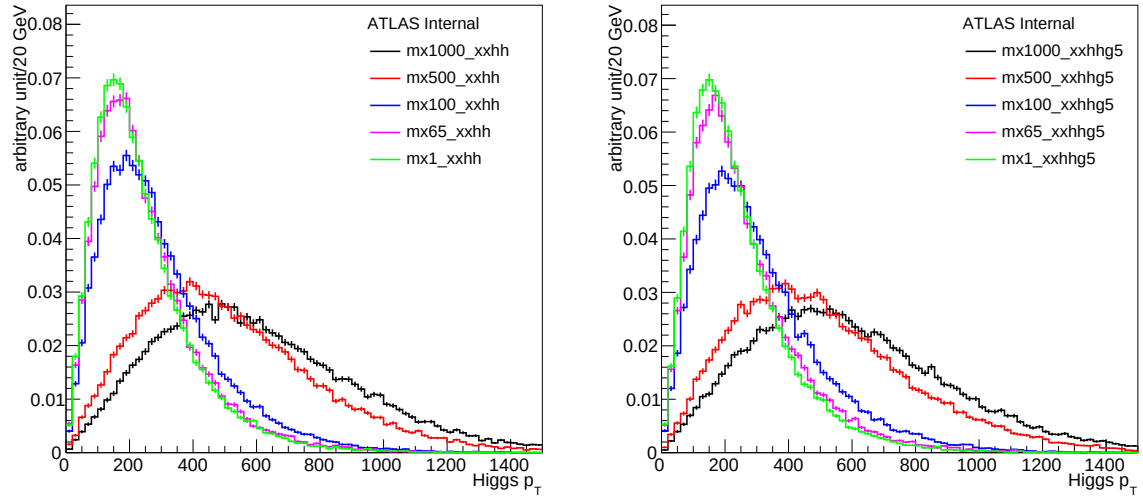


(c)



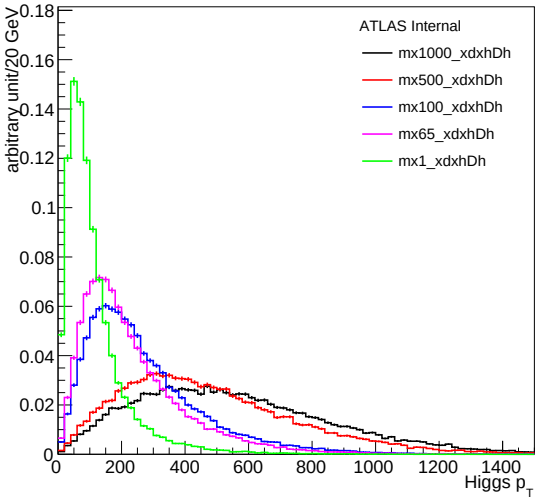
(d)

Fig. 5.2 $\Delta\phi(E_T^{miss,truth}, Higgs)$ distributions for five different DM masses (1, 65, 100, 500 and 1000 GeV) for sample (a) xxhh, (b) xxhhg5, (c) xdxhDh and (d) xgxFhDh.

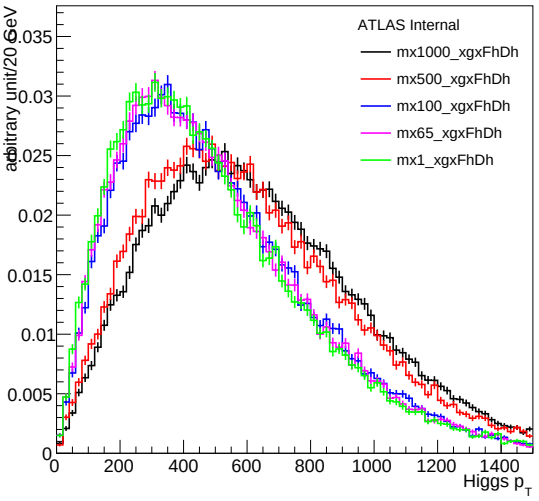


(a)

(b)



(c)



(d)

Fig. 5.3 Higgs p_{T} distributions for five different DM masses (1, 65, 100, 500 and 1000 GeV) for sample (a) xxhh, (b) xxhhg5, (c) xdxhDh and (d) xgxFhDh. The recoiling Higgs boson p_{T} distribution is very similar to $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in Figure 5.3 as Higgs and $E_{\mathrm{T}}^{\mathrm{miss}}$ are mostly back-to-back.

Based on these MC information, we can anticipate that the event topology with $E_{\mathrm{T}}^{\mathrm{miss}} + h(\rightarrow b\bar{b})$ final state to be similar to that illustrated in Figure 1.9. Because of the highly boosted (large p_{T}) Higgs, its decay products is expected to be very collimated. Hence, in order to fully capture these decay products we make use of the trimmed large-R jets with R = 1.0. Two b-quarks from the decay of Higgs boson are identified by associating the leading large-R jet with two small radius (R=0.3) b-tagged track jets.

As described in Section 4.7, large-R jet offers many advantages in boosted object reconstruction. To further demonstrate and justify its usage in this analysis, let us consider the decay of a highly boosted Higgs boson into a b-quark pair. The angular separation between two decay products of a heavy particles, ΔR[1] defined as the distance in azimuthal angle ϕ and in pseudorapidity η, can be approximated by [147]:

$$\Delta R \approx \frac{2m}{p_t} \tag{5.1}$$

where p_{T} and m are the transverse momentum and the mass of the parent particle, respectively. Figure 5.4 shows the separation of the two b-hadrons as a function of Higgs p_{T} at the parton level in MC simulation. Clearly the majority of b-hadrons resulting from the decay of a boosted Higgs boson with $p_{\mathrm{T}} > 250$ GeV would be contained within an R = 1.0 calorimeter jet. At higher Higgs p_{T} the separation between the two b-hadrons would be less than 0.8. This means the jets reconstructed with the standard distance parameter (R=0.4) would begin to merge, ultimately reduces the efficiency to identify two b-jets. This is where the track jet b-tagging (see Section 4.9) plays a central role to recover the efficiency.

Figure 5.5 shows the number of b-tagged jets (reconstructed at MC particle level) that are associated to the leading large-R calorimeter jet for `xgxFhDh` signal sample with m_{χ}=1000 GeV. Events are selected after event preselection plus at least one large-R jet requirement. Three different type of jets are shown. The black histogram belongs to the standard calorimeter jet with R=0.4. Only about 50% of the events have two b-tagged calorimeter jets with R=0.4 associated to the large-R jet, indicating that the two b-hadrons ended up inside a single R = 0.4 jet for a majority of the events. The performance is roughly the same when R = 0.4 track jets are used. However, by decreasing the track jet distance parameter from R = 0.4 to R = 0.3 drastically increases the number of b-tagged jets to around 70%.

5.3 Signal Selection

The full signal selection is given in Table 5.1. After the preselection mentioned in the previous section, events are required to pass the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger (refer to Section 4.11). The leptons veto is applied since no leptons are expected in the signal events. Two anti QCD multi-jet cuts based on the angular variables of $E_{\mathrm{T}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$ are applied (see Section 4.10 and Section 6.4 for

[1] $\Delta R(a,b) = \sqrt{(\phi_a - \phi_b)^2 + (\eta_a - \eta_b)^2}$

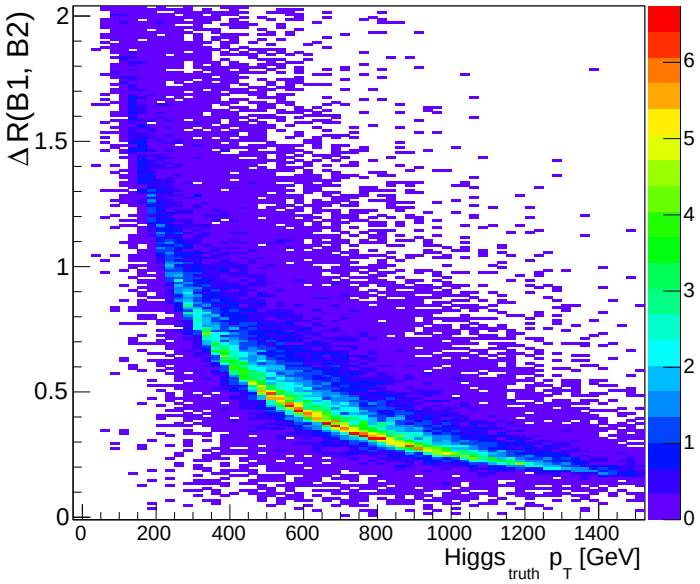


Fig. 5.4 Separation between b-hadrons as a function of Higgs p_{T} obtained using MC particle level information.

Table 5.1 The DM+Higgs($\rightarrow b\bar{b}$) event selection. Superscript index i of each jet collection means the i-th jet in descending order of the transverse momentum. Whereas subscript b in each jet collection means the jet is b-tagged.

Selection	Requirement	Note
Event quality preselection	pass	
$E_{\mathrm{T}}^{\mathrm{miss}}$xe80_tclcw EF trigger	pass	
Number of baseline leptons, n_ℓ	$== 0$	Anti QCD multi-jet cut
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)\rvert$	> 1.0	
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert$	$< \pi/2$	
Number of large-R jet, n_J	≥ 1	
Leading large-R jet p_T, $p_T^{J_1}$ [GeV]	> 350	Anti top cut
Number of associated track jet, $n_{j^{trk}}$	≥ 2	
Number of associated b-tagged track jet, $n_{j_b^{trk}}$	$== 2$	$h \rightarrow b\bar{b}$ reconstruction
Large-R jet mass, m_{J^1} [GeV]	$90 < m_{J^1} < 150$	
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	> 300	Signal region 1
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	> 400	Signal region 2

more discussion). To reconstruct the Higgs boson later on, we require at least one trimmed large-R jet. To ensure that the top quarks are more often fully-contained in the large-R jet, the leading jet is required to have $p_{\mathrm{T}} > 350$ GeV. These boosted top quarks have a reconstructed jet mass distribution that peaks at $\sim$170 GeV (see Reference [148] for more information). Thus they can be rejected more efficiently when combined with the leading jet mass cut. To identify

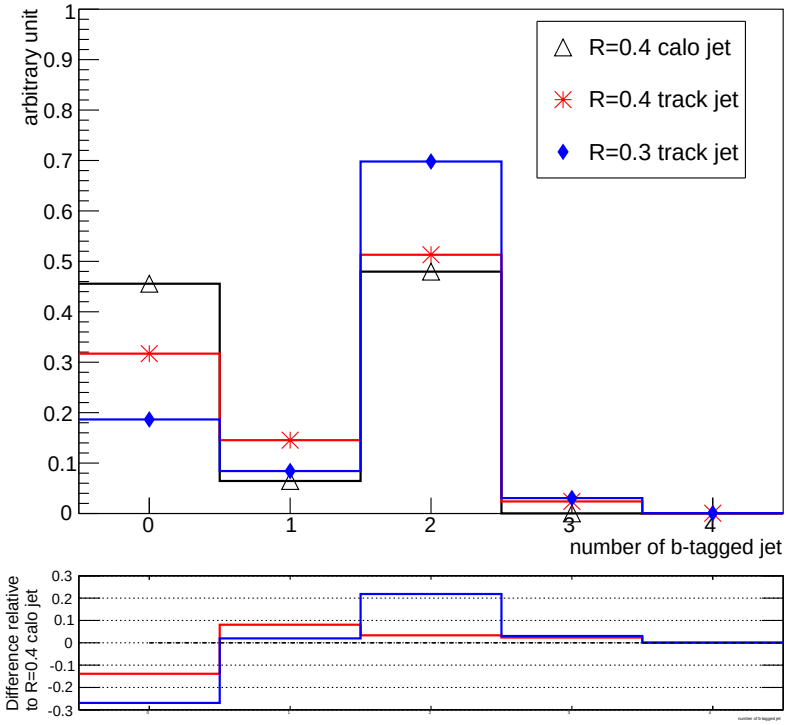


Fig. 5.5 Number of b-tagged jets (reconstructed at MC particle level) associated to the leading large-R jet calorimeter jets for `xgxFhDh` signal sample with m_χ=1000 GeV. The lower plot shows the difference in the fraction of event selected with both types of track jet relative to the fraction of event selected by the calorimeter jet with R=0.4.

the Higgs candidate, the leading large-R jet must have exactly two b-tagged track jets associated to it. Furthermore, the invariant mass of the leading jet should fall within a Higgs mass window between 90 and 150 GeV. Finally, two signal regions are defined, one with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV and the other with $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV. Each signal region is optimised for different mono-Higgs operators and different DM masses to achieve the best signal sensitivity.

5.4 Optimization

Figure 5.6 shows the distributions of several discriminant variables. All signal selections are applied except the cut on the discriminant variable itself. Two representative signal samples (the highest and the lowest m_χ) for each mono-H operator are plotted against the total MC background. The cut value on the kinematics variables: $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$, leading large-$R$ jet mass and $E_{\mathrm{T}}^{\mathrm{miss}}$ should be chosen to yield the optimal compromise between the signal efficiency and the background rejection. The optimal selection criteria for these three variables

are determined by iteratively maximising the signal significance, S, defined as

$$S = \frac{N_s}{\sqrt{N_s + N_b + (0.2N_b)^2}} \tag{5.2}$$

where N_s and N_b are the expected number of signal and background events, respectively. The systematic uncertainties on the total background prediction of about 20% is assumed. While the value that is under optimisation is being varied, all the other cut values are held constant. The value which gives the highest signal significance is chosen. All the backgrounds events used in the optimisation process are MC samples (see Section 3.4). For each mono-H operator, two representative MC signal samples (the highest and the lowest m_χ) are used to optimise the selection values.

Figure 5.7 shows the significance distribution for $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$, leading large-$R$ jet mass and $E_{\mathrm{T}}^{\mathrm{miss}}$. From Figure 5.7a, we require $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ to be greater than 1. For the leading large-R jet invariant mass the lower bound is fixed at $\sim 2\sigma$ (2 times the measured large-R jet mass resolution [149, 150]) away (90 GeV) from the W boson mass peak. The upper edge of the jet mass window is about 150 GeV as demonstrated in Figure 5.7b.

5.5 Event level corrections

5.5.1 Pileup rescaling

The absolute luminosity of a $p-p$ collider can be expressed as:

$$\mathscr{L} = \frac{\mu n_b f_r}{\sigma_{inel}} = \frac{\mu^{vis} n_b f_r}{\varepsilon \sigma_{inel}} = \frac{\mu^{vis} n_b f_r}{\sigma_{inel}^{vis}} \tag{5.3}$$

where f_r is the collider revolution frequency, n_b is the number of bunches, ε is the efficiency for one inelastic $p-p$ collision to satisfy the event-selection criteria (inlcuding detector acceptance) and μ^{vis} is the average number of visible inelastic interactions per bunch crossing (i.e. the mean number of $p-p$collisions per bunch crossing that pass that event selection). The visible inelastic cross section is related to the inelastic cross section by this relation $\sigma_{inel}^{vis} = \varepsilon\sigma_{inel}$.

The absolute luminosity can be inferred from the measured accelerator parameters via the van der Meer scans (vdM) [151, 152] without a priori knowledge of the inelastic $p-p$ cross section or of detector efficiencies. Similarly, μ^{vis} is also a directly measurable quantity. It can be measured using a set of dedicated detectors (such as the Beam Conditions Monitor (BCM) and LUminosity measurement using a Cherenkov Integrating Detector (LUCID) at ATLAS) and algorithms which allow bunch-by-bunch measurements. By comparing the known luminosity delivered in the vdM scan to the measured μ^{vis}, the visible inelastic cross section, σ_{inel}^{vis} can be determined from Equation 5.3. In order to test the reliability of MC simulation, the visible cross sections obtained from the data has been compared to that predicted by the

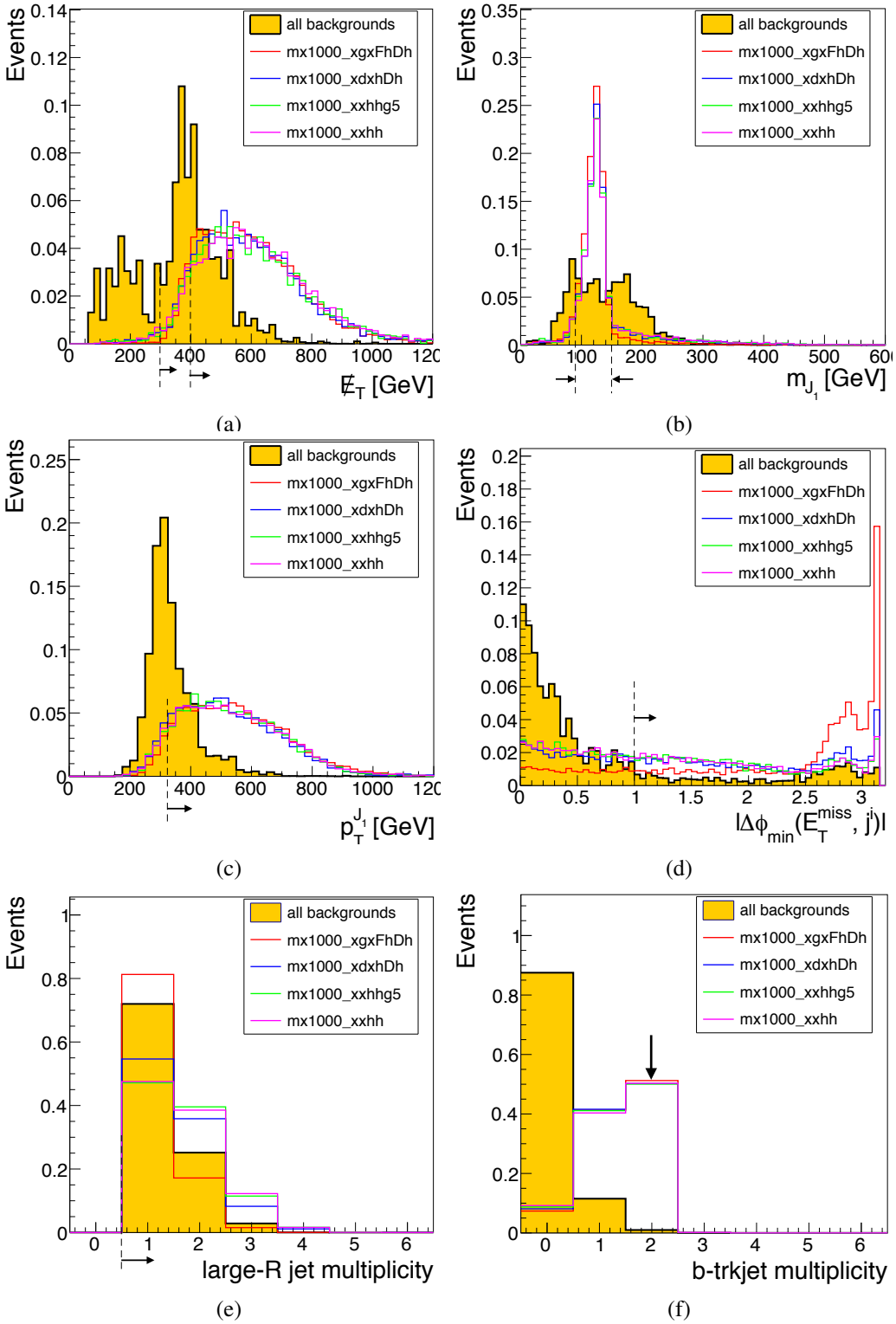
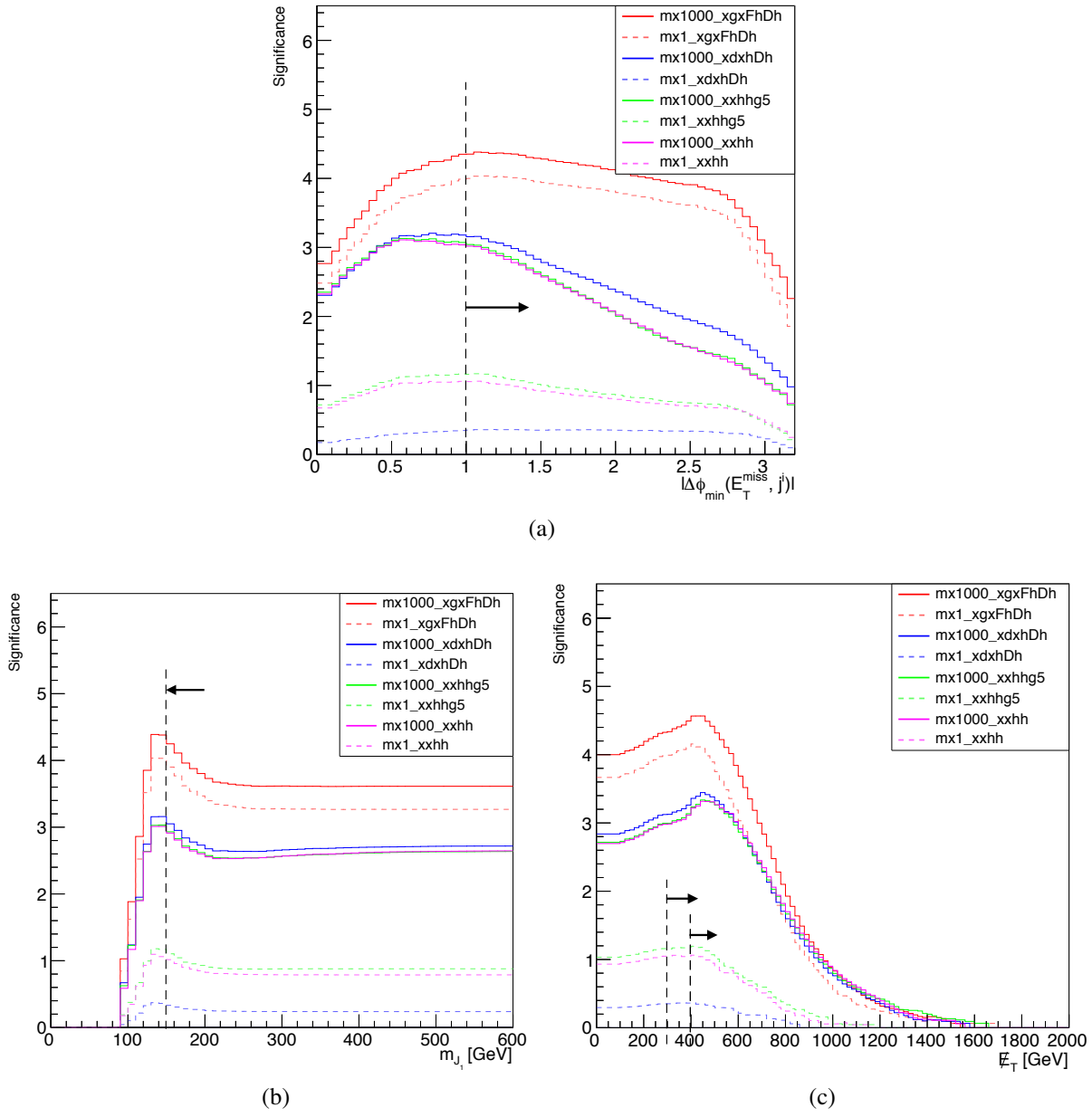


Fig. 5.6 The distribution of cut variables: (a) $E_{\mathrm{T}}^{\mathrm{miss}}$, (b) leading large-$R$ jet mass, (c) leading large-R jet p_{T} , (d) $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$, (e) large-$R$ jet multiplicity and (f) b-tagged track jet multiplicity for several signal samples and the total backgrounds. All distributions are normalised to unity. The arrow(s) indicate the cut value(s).



(a)

(b)

(c)

Fig. 5.7 Significance, S distribution for (a) $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$, (b) large-$R$ jet mass and (c) $E_{\mathrm{T}}^{\mathrm{miss}}$. The arrow(s) indicate the cut value(s). Since all signal cross section are scaled to 10fb^{-1} the values of S are just relative values and do not represent the true absolute significance.

MC [88, 153]. Small discrepancy between the PYTHIA8 event generator and the data has been observed. To account for this difference, the μ in MC has to be rescale. The rescaling factor is calculated by comparing the ratio of the visible cross section to the inelastic cross section, $\varepsilon = \sigma^{vis}_{inel}/\sigma_{inel}$ for data with that for simulation. The cross sections in data, σ_{inel} and σ^{vis}_{inel} are computed from independent measurements performed at the TOTEM experiment [154] and the ATLAS experiment [155]. The resulting scale factor, $\varepsilon^{data}/\varepsilon^{MC} = 1/1.09$ is applied as a multiplicative factor to all MC events. A systematics uncertainty is assigned to this correction as described in Section 7.3.1

5.5.2 Pileup reweighting

Each event in a MC sample is simulated with an assumption about what the average pile-up value will be under a particular simulated detector condition. However, it is impossible to predict exactly how much luminosity takes which pile-up value in reality.

Pile-up can affect things like reconstruction efficiency or the kinematics of the reconstructed objects. Hence it is important to make sure that the pile-up value in MC for a given fractional amount of luminosity matches the data. To derive this correction, the integrated luminosity from the data are binned according to the discrete pile-up values used in the reference pile-up distributions of the MC. Then correction factors for each pile-up bin, i, for an MC event simulated at a given detector condition, A is given by [156]

$$w_{pile-up} = \frac{L^A_i}{L} \cdot \frac{N}{N^A_i} \tag{5.4}$$

where L is the total integrated luminosity of the data, L^A_i is the integrated luminosity of all data assigned to the detector condition, A, in bin i as that found in the given MC event, N is the sum of the generator weight[1] of the whole MC sample and N^A_i is the sum of the generator weight of the events in the sample with the same detector condition, A, in the same bin, i. The correction factor is applied as a multiplication factor to all MC events in the same way as the pile-up rescaling factor mentioned above.

5.6 Signal selection efficiency

Figure 5.8 shows the selection efficiency (detector acceptance, A times reconstruction efficiency, ε) as a function of DM mass for each of the signal sample. More details on the absolute and relative selection efficiency for all signal samples are given in Table C.6-Table C.9 in the Appendix C.3. The selection efficiency after the full set of selection requirements varies from approximately 1% to 14% depending on the mono-H operator and DM mass.

[1] To avoid duplicating events, an event generator weight (whose value depends on generator) may be associated to each event.

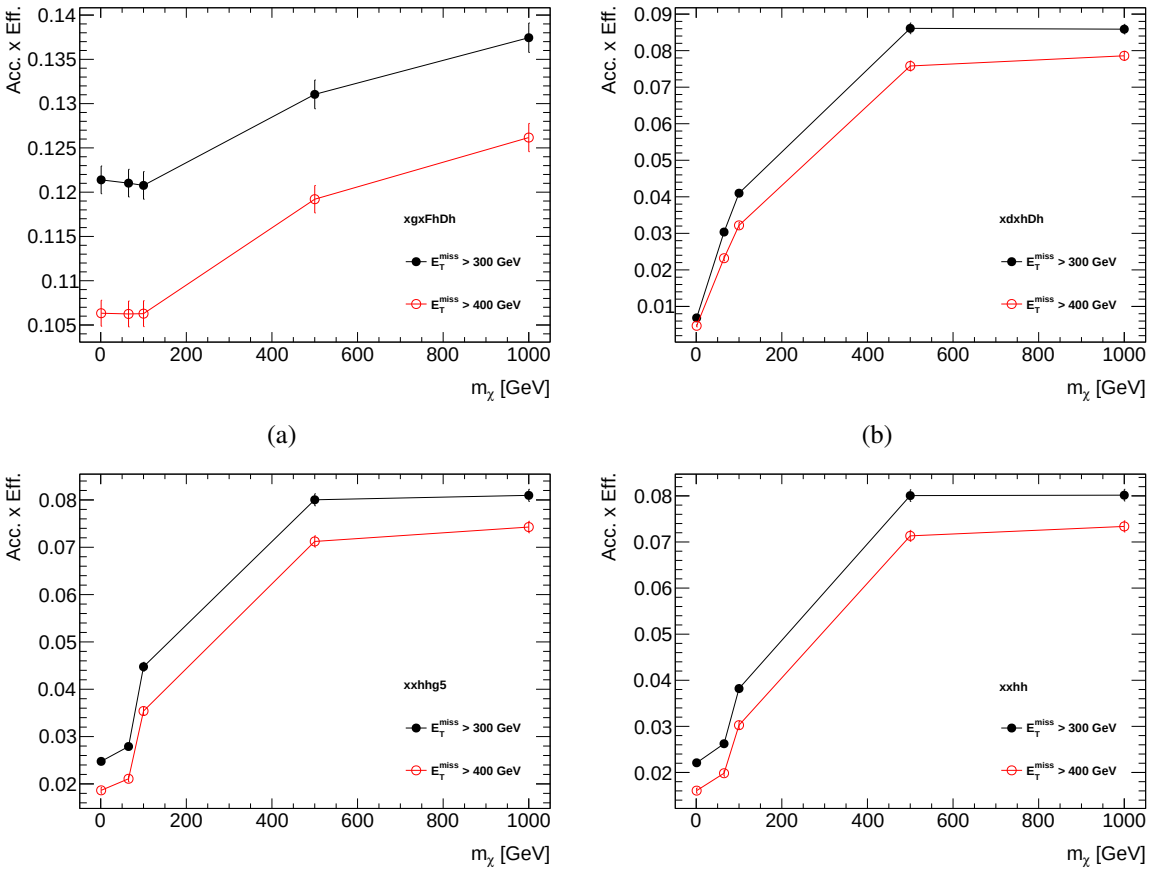


Fig. 5.8 Acceptance times efficiency as a function of DM mass for EFT signal samples in the SRs.

Chapter 6

Background estimation

The methods based on purely MC, the mixture of data and MC, and only data are used to estimate these backgrounds. To facilitate the estimation of each background, statistically independent control regions (CRs) are defined based on the number of leptons and the number of b-tagged track jets. In principle, the number of background events in the SRs can be extrapolated from the CRs based on the following relation:

$$N^{bkg^i}_{SR,data} = N^{bkg^i}_{SR,MC} \times \frac{N^{bkg^i}_{CR,data}}{N^{bkg^i}_{CR,MC}} \tag{6.1}$$

For this to works, each CR must be designed carefully such that each CR should contain a high purity of the specific background under consideration. At the same time the CR should also be free of signal contamination. This makes it possible that the dominant process in each CR can be controlled and compare to the data. In addition to the CRs, a validation region (VR) is constructed to verify the model used to predict the number of background events in the signal region or SR. The definitions of CRs, VR, and SR are as shown in Figure 6.1.

In the following sections, the procedures to estimate the $Z(\to \ell\ell)$+jets (Section 6.1), the $W(\to \ell\nu)$+jets, $t\bar{t}$, and single top (Section 6.2), QCD multi-jet (Section 6.4) and $Z(\to \nu\nu)$+jets (Section 6.3) are presented. Since no suitable CR can be defined for the diboson and the SM VH backgrounds, they are estimated directly from the MC simulations. The validation of the overall background modelling is shown in Section 6.5.

6.1 Estimation of $Z(\to \ell\ell)$+jets events

The estimation of Diboson, $Z(\to \ell\ell)$+jets, and SM VH backgrounds relies purely on the MC simulations. These backgrounds are normalised according to their production cross sections and the integrated luminosity. To check the modelling of $Z(\to \ell\ell)$+jets events, a 2-muon control region is defined. The definition of this control region is shown in Table 6.1. The events

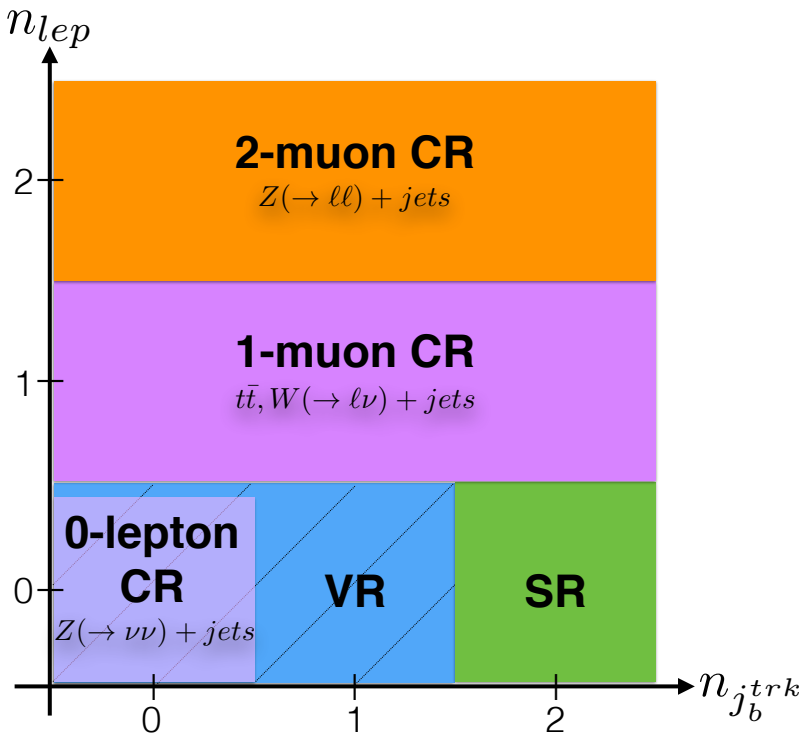


Fig. 6.1 Definitions of CRs, VR and SR in term of the number of lepton and the number of b-tagged track jet.

need to pass the single muon trigger requirement. There must be exactly 2 isolated muons with opposite charge. Each muon should have p_{T} greater than 25 GeV. Either of them need to be trigger matched to the single muon trigger. Furthermore, the reconstructed invariant mass of the dimuon system is required to be within 30 GeV of the Z boson mass (90 GeV). No events should contain any baseline electron. Since the Z boson is not expected to decay to neutrinos in these events, the $E_{\mathrm{T}}^{\mathrm{miss}}$ are expected to be small (mainly coming from fake $E_{\mathrm{T}}^{\mathrm{miss}}$ from the mis-reconstruction of jets). To increase the statistics, however, no $E_{\mathrm{T}}^{\mathrm{miss}}$ cut is applied. Similarly, to ensure that there are enough events, no $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ cut is applied. Finally, the selected events also need to pass the cuts on the large-R jet multiplicity, large-R jet p_{T}, and associated track jet multiplicity.

Table 6.1 The definition of the 2-muon conrtol region to check the modeling of $Z(\to \ell\ell)$+jets background.

2-muon control region	
Cuts	Values
Event quality preselection	pass
Single muon trigger	pass
Number of isolated muon, $n_{\mu_{iso}}$	==2
Trigger matching (either μ)	==true
Muons are oppositely charge	== true
Both muon's p_{T}[GeV]	>25
|Invariant mass of dimuons (m_{ll}) - 90| [GeV]	<30
Number of baseline electron, n_e	==0
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert$	$< \pi/2$
Number of large-R jet, n_J	≥ 1
Leading large-R jet p_T, $p_T^{J_1}$ [GeV]	> 300
Number of associated track jet, $n_{j^{trk}}$	≥ 2

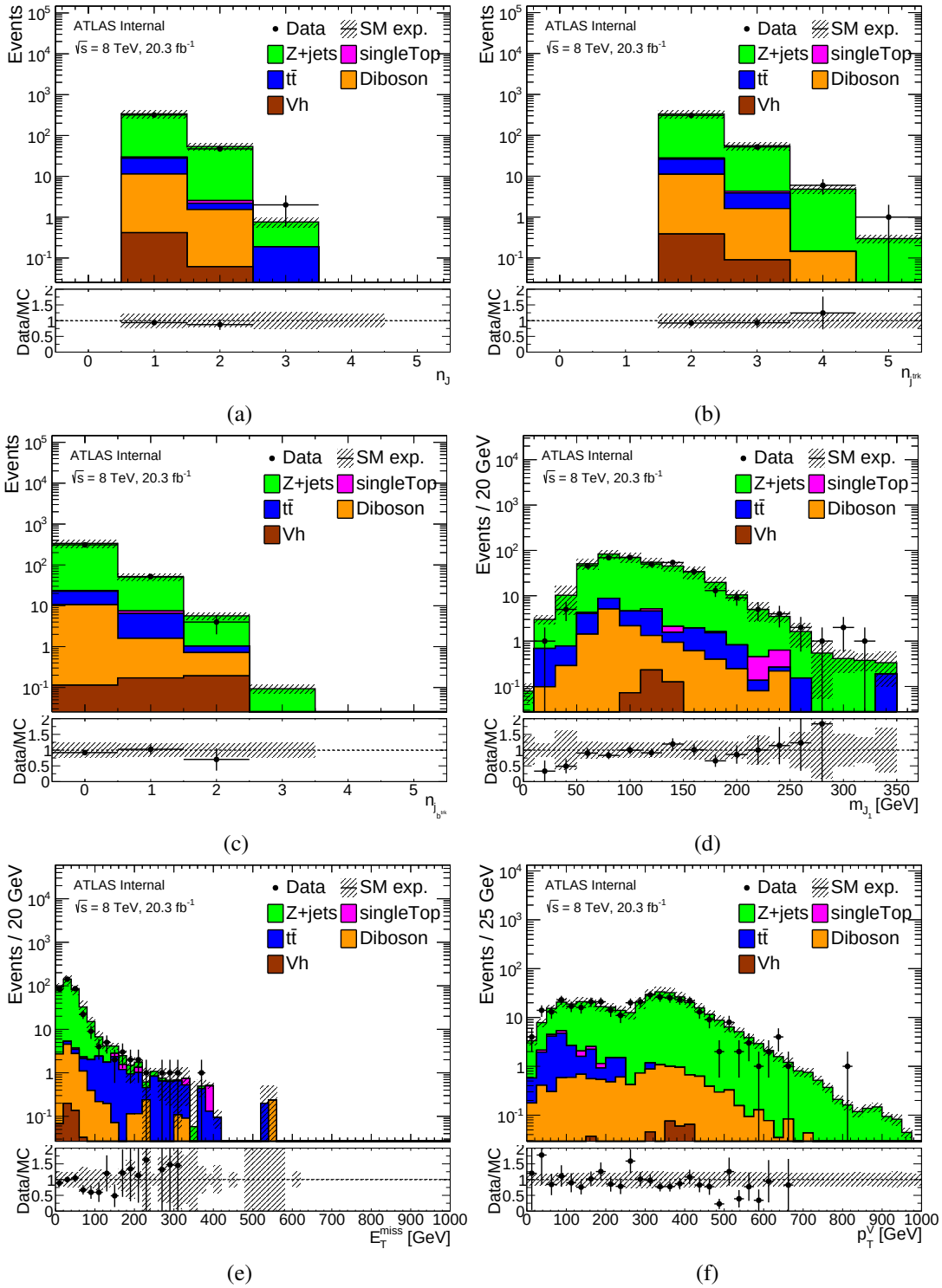


Fig. 6.2 Kinematic distributions for 2-muon CR. (a) trimmed AntiKt10 jet multiplicity, (b) the number of associated track jets, (c) the number of associated *b*-tagged track jets, (d) leading large-R jet mass, (e) $E_{\mathrm{T}}^{\mathrm{miss}}$ and (f) p_{T}^{V}. The systematic uncertainties are shown as hatched band while the statistical uncertainties are given as error bars.

Several kinematic distributions of the 2-muon CR are shown in Figure 6.2. After all the cuts, more than 94% of the selected events are $Z(\rightarrow \ell\ell)$+jets events. The contribution from $t\bar{t}$, single top and VH backgrounds are basically negligible. Good agreement is achieved between data and background. Most importantly, the modelling of the invariant mass of the reconstructed large-R jet in the range between $50 < m_J < 170$ GeV is very satisfactory.

6.2 Estimation of $t\bar{t}$, $W(\rightarrow \ell\nu)$+jets and single top events

The estimation of $W(\rightarrow \ell\nu)$+jets, $t\bar{t}$ and single top background in the SR proceed via a semi data-driven approach. We trust the MC simulation to correctly reproduce the shape of the kinematic distributions of these backgrounds. What we need to estimate is the total number of events for each background in the SRs. To do this, the matrix method is used. A set of linear equations with two unknowns, i.e. the scale factors of $W(\rightarrow \ell\nu)$+jets (SF_W) and $t\bar{t}$ +single Top (SF_{Top}), can be written as

$$SF_W * N^{MC}_{CR_W}(W) + SF_{Top} * N^{MC}_{CR_W}(\text{Top}) = N^{Data}_{CR_W} - N^{MC}_{CR_W}(\text{Non-}W\text{/Top}) \quad (6.2)$$

$$SF_W * N^{MC}_{CR_{Top}}(W) + SF_{Top} * N^{MC}_{CR_{Top}}(\text{Top}) = N^{Data}_{CR_{Top}} - N^{MC}_{CR_{Top}}(\text{Non-}W\text{/Top}) \quad (6.3)$$

where N^{Data} (N^{MC}) are the yield of the observation (prediction) from the different sources that are present in the $W(Top)$ CRs as indicated by the subscript CR_W (CR_{Top}).

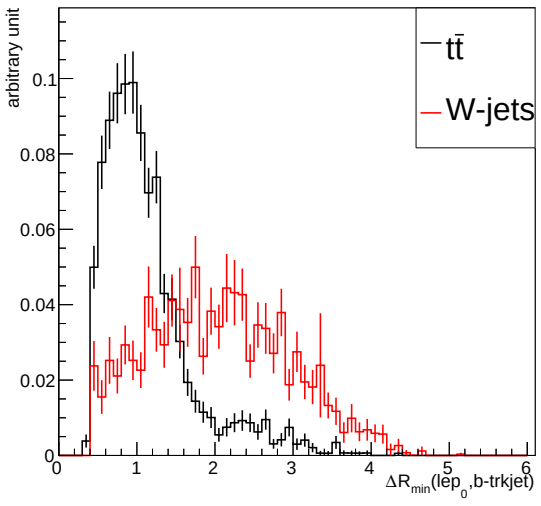
In order to solve this Equation 6.3, we first need to construct a Top CR (where $t\bar{t}$ and single top backgrounds are combined) and a $W(\rightarrow \ell\nu)$+jets CR. These two CRs are nearly identical to the SRs, but the lepton veto is reversed by requiring exactly one isolated muon and no electrons in the final state. The full selection of each CR is listed in Table 6.2. As shown in the table, all the selections except the last are common between the two CRs. At the last selection stage, the remaining events are separated into two categories by introducing a cut on the number of b-tagged track jets, which are not associated to the leading large-R jet[1], near the selected isolated muon. The separation between the muon and the non-associated b-tagged track jets, $\Delta R(l,$ non-associated b-tagged track jet) should be less than 1.5. The distributions of this variable for $t\bar{t}$ and $W(\rightarrow \ell\nu)$+jets are shown in Figure 6.3. One can imagine that for highly boosted leptonically decay top quarks, there should be a b-quark nearby the muon from the W boson, whereas this is less common for $W(\rightarrow \ell\nu)$+jets events. The purity of $t\bar{t}$ plus single top MC events in the Top CR after these cuts is ~94%. The purity of $W(\rightarrow \ell\nu)$+jets events in the $W(\rightarrow \ell\nu)$+jets CR is ~72%. The contamination from Top process ($t\bar{t}$ plus single top) in this region is ~23%.

From the two CRs we constructed, we get the value for all the $N^{Data/MC}_{CR}$ parameters. The scale factor for W+jets and top background can then be derived by solving Equation 6.3

[1]Here, non-association simply means those b-tagged track jets and the leading large-R jet must have a separation of $\Delta R(J_1, j^{trk}) > 1.0$.

Table 6.2 The event selection of the CRs for the scale factor calculation. Note that "non-associated b-tagged track jet" means b-tagged track jets that have $\Delta R(J_1, j^{trk}) > 1.0$.

Selection	Requirement	
	Top CR	$W(\to \ell\nu)$+jets CR
Event quality preselection	pass	
Single muon trigger	==true	
Number of isolated muon, $N_{\mu,iso}$	==1	
Muon's p_T [GeV]	> 25	
Nmber of baseline electron, N_e	==0	
E_T^{miss} [GeV]	> 200	
Number of large-R jet, N_J	≥ 1	
Leading large-R jet p_T, $J^1_{p_T}$ [GeV]	> 350	
Number of associated track jet, $N_{j^{trk}}$	≥ 2	
any b-tagged track jet near lepton with $\Delta R(l$, non-associated b-tagged track jet$) <$ 1.5	True	False



(a)

Fig. 6.3 Comparison of the distributions of ΔR separation between the leading muon and b-tagged track jet for $t\bar{t}$ (black) and $W(\to \ell\nu)$+jets (red) events. All the track jets considered here must have $\Delta R(J_1, j^{trk}) > 1.0$. Both distributions are normalised to unity.

simultaneously. The derived scale factors and the contribution of each background in this CR are listed in Table 6.3. As a cross check, two additional CRs are created by varying track jet and associated track-jet b-tagging requirements. The scale factors and the contribution of each background in these two additional CRs are shown in Table 6.4. All scale factors in all three CRs agree with each other within 1σ of statistical uncertainty.

Table 6.3 Data yield and MC prediction (before correction) in the W+jets and top CRs at ≥ 2 track jet selection stage. Top CR includes contribution from single top and $t\bar{t}$. Only the statistical error is shown.

Selection	CR	Diboson	W+jets	Z+jets	Single Top	$t\bar{t}$	ZH	$Z\nu\nu_{MC}$	Data	Scale Factor
same as Tab. 6.2	Top	1.7 ± 0.3	17.6 ± 1.1	1.5 ± 0.2	19.1 ± 2.8	293.6 ± 8.3	0.1 ± 0.0	0.0 ± 0.0	297 ± 17	0.894 ± 0.062
	$W(\rightarrow \ell\nu)$+jets	51.7 ± 2.1	1313.4 ± 10.2	27.8 ± 0.8	61.1 ± 6.0	317.9 ± 8.7	1.1 ± 0.1	0.0 ± 0.0	1490.0 ± 38.6	0.815 ± 0.045

Table 6.4 Data yield and MC prediction (before correction) in the $W(\rightarrow \ell\nu)$+jets and top CRs at 2 other selection stages. Top CR includes contribution from single top and $t\bar{t}$. Only the statistical error is shown.

Selection	CR	Diboson	W+jets	Z+jets	Single Top	$t\bar{t}$	ZH	$Z\nu\nu_{MC}$	Data	Scale Factor
same as Tab. 6.2 except $N_{jets} \geq 1$	Top	4.3 ± 0.5	53.1 ± 1.9	5.1 ± 0.3	63.4 ± 7.1	466.5 ± 10.6	0.2 ± 0.0	0.0 ± 0.0	500.0 ± 22.4	0.841 ± 0.048
	$W(\rightarrow \ell\nu)$+jets	108.5 ± 3.1	4461.5 ± 21.1	97.0 ± 1.6	145.5 ± 10.4	489.0 ± 10.9	1.6 ± 0.1	0.0 ± 0.0	4497.0 ± 67.1	0.842 ± 0.020
same as Tab. 6.2 plus $N_{jets}^{b} == 0$	Top	1.5 ± 0.3	14.2 ± 1.0	1.3 ± 0.2	13.5 ± 2.4	123.9 ± 5.5	0.0 ± 0.0	0.0 ± 0.0	132.0 ± 11.5	0.857 ± 0.094
	$W(\rightarrow \ell\nu)$+jets	42.7 ± 2.0	1161.3 ± 9.8	24.6 ± 0.8	21.5 ± 3.8	133.4 ± 5.7	0.1 ± 0.0	0.0 ± 0.0	1135.0 ± 33.7	0.805 ± 0.032

The derived scale factor, 0.894 (0.815) is used to scale $t\bar{t}$ and single top ($W(\rightarrow \ell\nu)$+jets) events. To check the kinematic distributions after applying these scale factors, a region combining both the Top and $W(\rightarrow \ell\nu)$+jets CRs where all events that pass the selection up to the number of associated track jet as listed in Table 6.2 is defined. Figure 6.4 shows the kinematic distributions for these events after applying the scale factors derived in Table 6.3. The shapes of the distributions between the data and MC are in an good agreement.

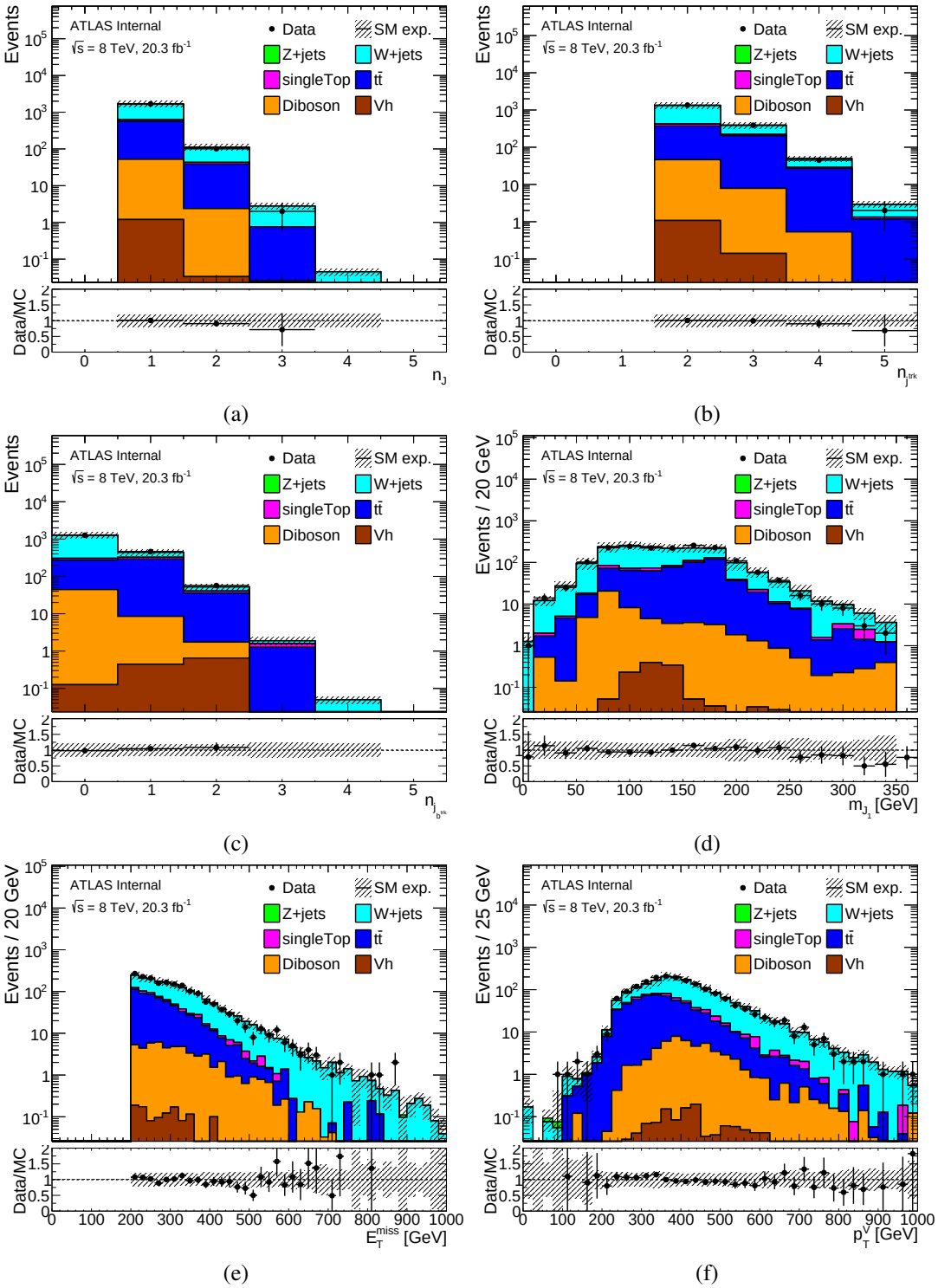


Fig. 6.4 Kinematic distributions for $W(\to \ell\nu)$+jets and top CR after correction. (a) trimmed AntiKt10 jet multiplicity, (b) the number of associated track jets, (c) the number of associated b-tagged track jets, (d) leading large-R jet mass, (e) $E_{\mathrm{T}}^{\mathrm{miss}}$ and (f) W's p_{T}. The systematic uncertainties are shown as hatched band while the statistical uncertainties are given as error bars.

6.3 Data-driven background estimate of $Z(\to \nu\nu)$+jets

An accurate description of $Z(\to \nu\nu)$+jets background is crucial to achieve better sensitivity for DM pair production in this analysis. Estimation based on MC alone suffers from large uncertainties; hence we opted for a data-driven method whose details are described below.

6.3.1 Method overview

The $Z(\to \nu\nu)$+jets background can be estimated via a translation from a well-understood and cleanly measured reference process. This method is viable if the two processes are kinematically similar in some regime. The translation is performed via a transfer function (TF) derived by taking the ratio of a kinematics variable between the $Z(\to \nu\nu)$+jets and the reference process. Symbolically, this translation from process B to A can be expressed as

$$A(x) \equiv TF_{B\to A}(x) \cdot B(x) \tag{6.4}$$

where x is a kinematic variable, for e.g $E_{\mathrm{T}}^{\mathrm{miss}}$ or p_{T} of the object of interest. This method has an advantage such that it is dependent only on one parameter, x. Furthermore, it also provides the cancellation of uncertainties among the two processes that are under studies.

With this idea in mind, the next step is to determine which reference process we want to use to estimate the $Z(\to \nu\nu)$+jets process. As we shall explain later, γ+jets is chosen as the reference process. Next, to derive the transfer function, we first need to decide which kinematic variable we want to use. In our case, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is the natural choice. Equipped with the transfer function, we can then convert γ+jets events to the $Z(\to \nu\nu)$+jets events. All these steps are explained in more detailed in the following sections.

6.3.2 γ+jets as the reference process

We chose to use γ+jets as the reference process[1]. The main reason is that the γ+jets process offers higher statistics especially at higher $E_{\mathrm{T}}^{\mathrm{miss}}$ or p_{T} region since there is no branching ratio suppression for prompt photon production. Using γ+jets as the reference process is viable provided that the theoretical uncertainties can be kept at a competitive level. In fact, they have been shown to be within 5% [157]. The key is that the Z boson and the γ production processes

[1] In principle, it is wise to choose a well understood reference process as similar as possible to the process to be estimated (so that the theoretical uncertainty is minimised). For the $Z(\to \nu\nu)$+jets, the corresponding reference process would naturally be $Z(\to \ell\ell)$+jets ($\ell = e, \mu$), where the theoretical cross section ratio only differs by the branching ratio. Nonetheless, the problem is that the statistics of $Z(\to \ell\ell)$+jets events in high $E_{\mathrm{T}}^{\mathrm{miss}}$ or p_{T} region are too small. The precision of this method will suffer when we extrapolate from looser selection in the CRs to harder selections in the SR. It is also possible to use $W(\to \ell\nu)$+jets as the reference process. But the modest increase in statistics is offset by larger theoretical uncertainties with respect to the $Z(\to \nu\nu)$+jetsprocess. Also, the lower purity of $W(\to \ell\nu)$+jets may also reduce the precision of the estimation.

have very similar production mechanisms. These processes become increasingly similar at high p_{T} ($> m_Z$) as the effect of the non-zero Z mass slowly disappears. At high p_{T}^V, only their electroweak couplings to quarks contributes to the difference in their rate. This similarity (where the ratio is relatively constant with respect to event kinematics) is shown in Figure 6.5. These reasons, coupled with the fact that the signal selection of this analysis are characterised by high-p_{T} requirements, γ+jets process is chosen to predict the $Z(\to \nu\nu)$+jets background.

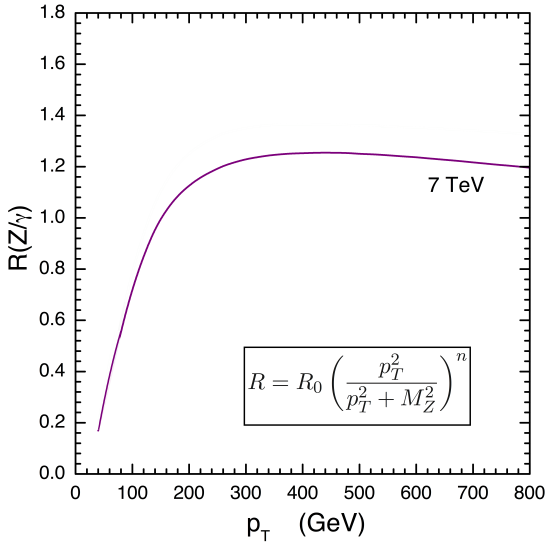


Fig. 6.5 Ratio of the differential cross section between the inclusive Z+1 jet and γ+1 jet processes as a function of the vector boson p_{T}Ḟigure adapted from Reference [157].

6.3.3 Deriving transfer function, $TF_{\gamma\to Z_{\nu\nu}}$, from γ+jets to $Z(\to \nu\nu)$+jets

To derive the transfer function, $TF_{\gamma\to Z_{\nu\nu}}$, two templates are needed. One is the γ+jets template and the other is the $Z(\to \nu\nu)$+jets template. These two templates are obtained from MC samples as explained below.

γ+jets and $Z(\to \nu\nu)$+jets templates

To obtain a sample of pseudo $Z(\to \nu\nu)$+jets events from γ+jets events, we need to treat the p_{T} of γ as an estimator of the $E_{\mathrm{T}}^{\mathrm{miss}}$ caused by two neutrinos decayed from Z. We define $\not{E}_{\mathrm{T}}^{\gamma}$ as the vector sum of the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector and the p_{T} of photon. Throughout this thesis, any reference to $E_{\mathrm{T}}^{\mathrm{miss}}$ of an event from the γ+jets selection is taken to mean $\not{E}_{\mathrm{T}}^{\gamma}$.

To obtain an enriched sample of high p_{T} photons, events must pass the requirements shown in Table 6.5. In addition, events are required to pass the lowest unprescaled photon trigger g120 EF trigger, which is fully efficient for $p_{\mathrm{T}}^{\gamma} > 125$ GeV. Similarly a $Z(\to \nu\nu)$+jets template is selected by imposing a set of cuts listed in Table 6.6.

Table 6.5 Selection for γ+jets template

Selection	Requirement
Event quality preselection	pass
Single photon trigger	pass
Number of photon, N_γ	== 1
Photon's p_{T}, p_{T}^{γ} [GeV]	> 125
Number of baseline leptons, n_ℓ	== 0
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)\rvert$	> 1.0
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert$	$< \pi/2$
Number of large-R jet, n_J	≥ 1
Leading large-R jet p_T, $p_T^{J_1}$ [GeV]	> 350
Number of associated track jet, $n_{j^{trk}}$	≥ 2

Table 6.6 Selection for $Z(\to \nu\nu)$+jets template

Selection	Requirement
Event quality preselection	pass
$E_{\mathrm{T}}^{\mathrm{miss}}$xe80_tclcw EF trigger	pass
Number of baseline leptons, n_ℓ	== 0
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)\rvert$	> 1.0
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert$	$< \pi/2$
Number of large-R jet, n_J	≥ 1
Leading large-R jet p_T, $p_T^{J_1}$ [GeV]	> 350
Number of associated track jet, $n_{j^{trk}}$	≥ 2

6.3.4 Fitting the transfer function

The $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions of $Z(\to \nu\nu)$+jets and γ+jets templates are shown in the upper plot in Figure 6.6. The transfer function is defined as the ratio of the of the $Z(\to \nu\nu)$+jets $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution to the γ+jets $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. We model this ratio with the function form of

$$TF_{\gamma\to Z_{\nu\nu}} = R_0\left(\frac{E_T^{miss\ 2}}{E_T^{miss\ 2} + M_Z^2}\right)^n, \tag{6.5}$$

where M_Z is Z boson mass. This function is motivated by the approximation of cross-section ratio between $Z(\to \nu\nu)$+jets and γ+jets, which is found to be proportional to the n^{th} power of the ratio of two propagators: $1/(p_T^2 + M_Z^2)$ and $1/p_T^2$ for $Z(\to \nu\nu)$+jets and γ+jets, respectively. The R_0 and the power n are taken as free parameters during fitting. The fitting range is between 200 GeV and 1000 GeV. Further details of this parameterisation can be found in Reference [157]. The fit yields a $\chi^2/d.o.f.$ of 1.32. The fitted curve is also shown in the lower plot in Figure 6.6.

To check if the transfer function affects other kinematic properties of the event, a closure test is performed. An event-by-event weight, $w_{\gamma\to Z_{\nu\nu}}$, extracted from the transfer function is

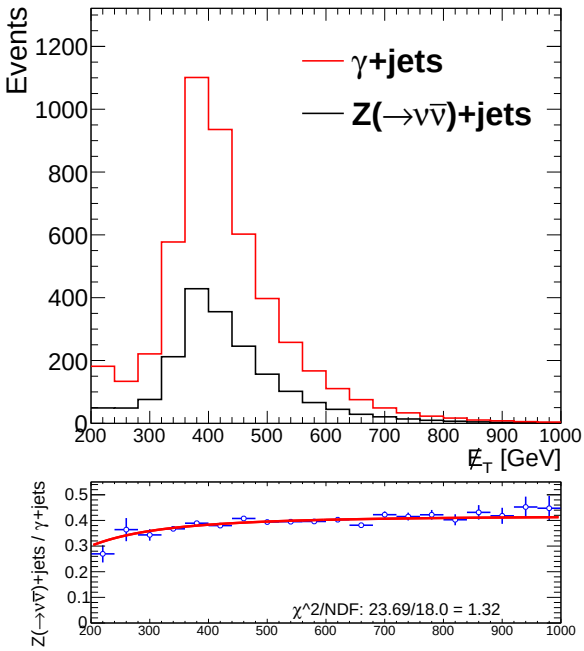


Fig. 6.6 $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution for γ+jets and $Z(\rightarrow \nu\nu)$+jets. The ratio of $Z(\rightarrow \nu\nu)$+jets to γ+jets is the transfer function.

applied to each selected γ+jets event. The $w_{\gamma\rightarrow Z_{\nu\nu}}$ for the i^{th} event is defined as

$$w^{i}_{\gamma\rightarrow Z_{\nu\nu}} = TF_{\gamma\rightarrow Z_{\nu\nu}}(E_T^{miss,\ i}) \tag{6.6}$$

Figure 6.7 shows comparisons of various kinematic variables between reweighted γ+jets events to the $Z(\rightarrow \nu\nu)$+jets MC events. As can be seen in the ratio plots of Figure 6.7, an overall good agreement is observed within 1σ uncertainty by taking into account both statistical and systematic variations. The treatment of the systematic uncertainties for this method will be discussed in Section 7.4.2.

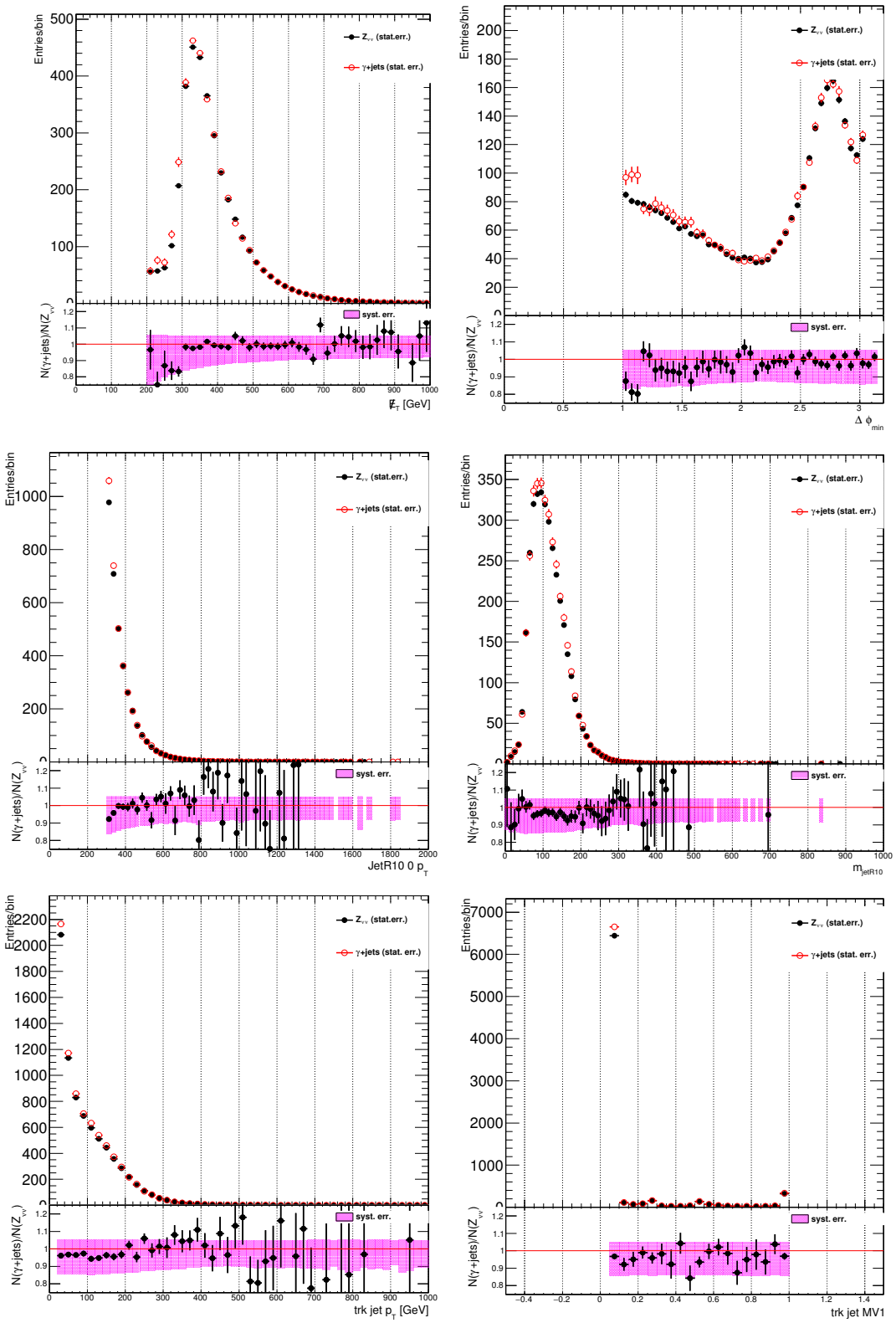


Fig. 6.7 As a closure test, each selected γ+jets MC event is reweighted using the fitted transfer function. Starting from top left is the $E_{\mathrm{T}}^{\mathrm{miss}}$, $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$, leading large-$R$ jet p_T, leading large-R jet mass, track jet p_T and track jet MV1 value. The shaded band in the systematic uncertainty from the transfer function variation described in the Section 7.4.2.

6.3.5 Reweight γ+jets to $Z(\rightarrow \nu\nu)$+jets using template from data

Having derived the transfer function, the next step is to obtain the estimate of $Z(\rightarrow \nu\nu)$+jets in data. A γ+jet template is constructed from data by imposing the selections listed in Table 6.5. The event yields are presented in Table 6.7. The first 6 columns in Table 6.7 are the yields of the other non-γ+jets backgrounds estimated from MC simulation. These non-γ+jets backgrounds have negligible contribution. That means the uncertainties of MC prediction for these backgrounds have little effect on $Z(\rightarrow \nu\nu)$+jets backgrounds estimation. The last column is the yield of γ+jets events after subtracting all other non-γ+jets contribution from the data. These γ+jets events selected from the data is reweighted by applying an event weight $w_{\gamma\rightarrow Z_{\nu\nu}}$ as defined in Equation 6.6.

Table 6.7 Events that passed γ+jets selection. Only the statistical error is shown.

Diboson	W+jets	Z+jets	Single Top	$t\bar{t}$	ZH	γ+jets (Data)
0.5 ± 0.1	10.5 ± 0.5	0.4 ± 0.1	0.1 ± 0.1	0.5 ± 0.2	0.0 ± 0.0	1557.7 ± 24.4

Finally kinematic distributions of the estimate of the $Z(\rightarrow \nu\nu)$+jets events, together with the other background contributions in the $Z(\rightarrow \nu\nu)$+jets CR (selection same as in Table 6.6 plus additonal requirement that $n_{j_b^{trk}} == 0$) are plotted in Figure 6.8. This region is dominated by W/Z+light jets events. Even though their contribution to the signal region is negligible, it still can serve as an excellent region to identify subtle discrepancies between data and simulation due to its high statistics. In general, the shape of the MC distributions agree well with the data. Nonetheless there is still a residual difference of around $\sim$10% in the data/MC ratio. This discrepancy originates from the mismodeling of $Z(\rightarrow \nu\nu)$+jets MC template used to derive the transfer function. To correct for this difference, a scale factor for the $Z(\rightarrow \nu\nu)$+jets estimate need to be derived.

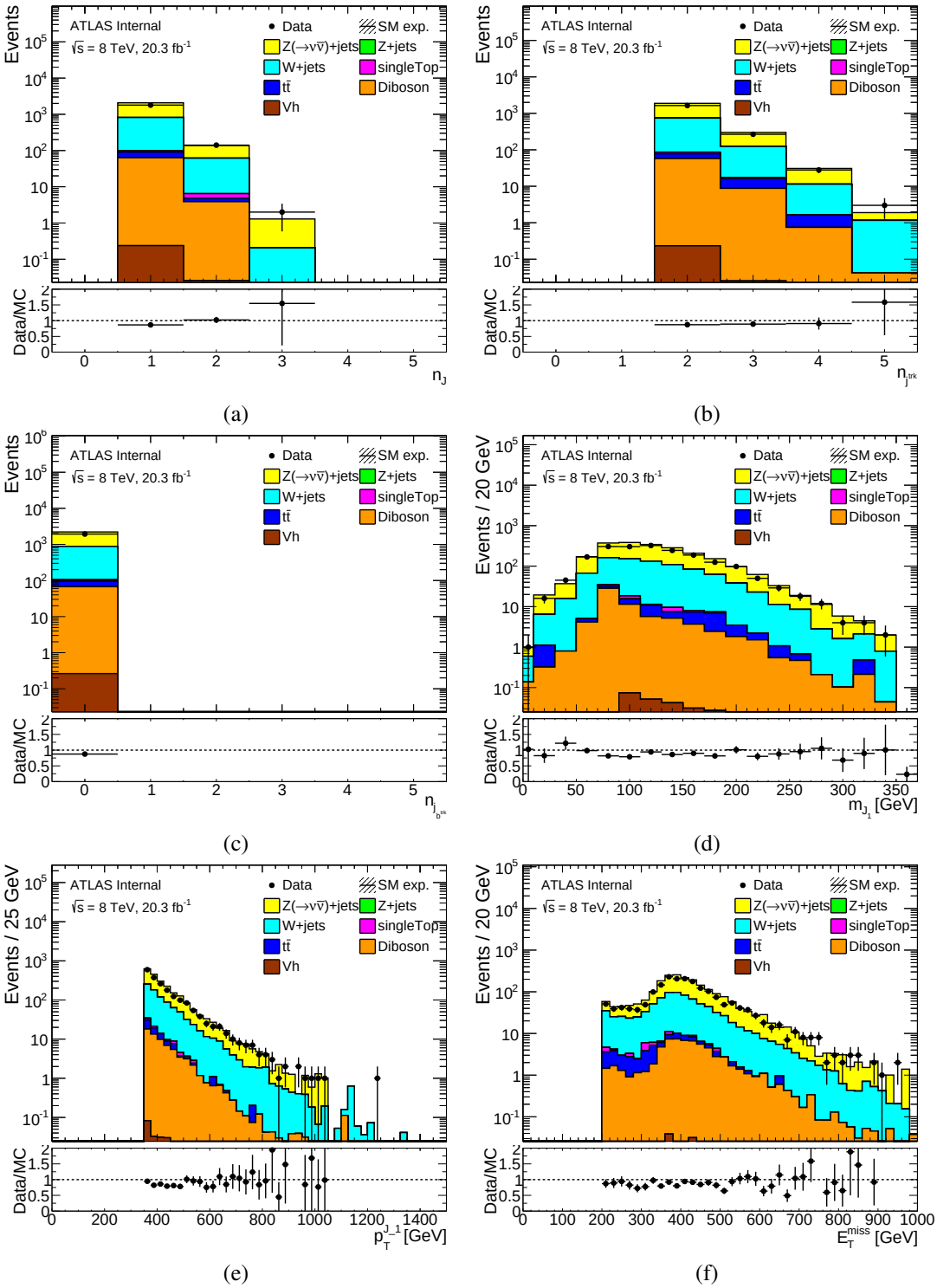


Fig. 6.8 Kinematic distributions in the $Z(\to \nu\nu)$+jets CR. The $Z(\to \nu\nu)$+jets template is obtained using Equation 6.6 and before applying scale factor. (a) trimmed AntiKt10 jet multiplicity, (b) the number of associated track jets, (c) the number of associated b-tagged track jets, (d) leading large-R jet mass, (e) leading large-R jet p_{T} and (f) $E_{\mathrm{T}}^{\mathrm{miss}}$. Only the statistical error is shown.

6.3.6 Scale factor for $Z(\rightarrow \nu\nu)$+jets

The scale factor, SF, for $Z(\rightarrow \nu\nu)$+jets can be derived by solving the equation

$$SF = \frac{N^{data} - N^{MC}_{\text{non } Z(\rightarrow \nu\bar{\nu})\text{+jets}}}{N_{Z(\rightarrow \nu\bar{\nu})\text{+jets}}} \tag{6.7}$$

The scale factor derived from this $Z(\rightarrow \nu\nu)$+jets CR (selection same as in Table 6.6 plus additional requirement that $n_{j_b^{trk}} == 0$) is shown in the last column of Table 6.8. The deviation of the scale factor from the unity is assigned as the symmetric uncertainty on the $Z(\rightarrow \nu\nu)$+jets normalisation.

To test the robustness of this scale factor, the scale factors derived using the same procedure with slightly different b-tagging requirement are compared in Table 6.9. Both scale factors agree with the nominal scale factor within 1σ uncertainty.

Table 6.8 Normalization scale factor derived for $Z(\rightarrow \nu\nu)$+jets background. The W+jets and top scale factors derived earlier in Table 6.3 are applied. Only the statistical error is shown.

Selection	Diboson	W+jets	Z+jets	Single Top	$t\bar{t}$	ZH	$Z(\rightarrow \nu\nu)$+jets	data	Scale Factor
same as Tab. 6.6 plus $N_{j_b^{trk}} == 0$	67.0 ± 1.6	629.9 ± 8.4	3.2 ± 0.3	8.3 ± 2.1	25.6 ± 2.6	0.3 ± 0.0	1333.0 ± 22.7	1938.0 ± 44.0	0.896 ± 0.037

Table 6.9 As a cross check, additional two normalisation scale factors are derived for $Z(\rightarrow \nu\nu)$+jets background with slightly different b-tagging requirement. The W+jets and top scale factors derived earlier in Table 6.3 are applied. Only the statistical error is shown.

Selection	Diboson	W+jets	Z+jets	Single Top	$t\bar{t}$	ZH	$Z(\rightarrow \nu\nu)$+jets	data	Scale Factor
same as Tab. 6.6	83.6 ± 1.8	700.1 ± 8.8	3.7 ± 0.3	16.7 ± 3.2	69.7 ± 4.2	2.4 ± 0.1	1545.6 ± 24.4	2270.0 ± 47.6	0.895 ± 0.034
same as Tab. 6.6 plus $N_{j_b^{trk}} == 1$	13.2 ± 0.7	63.3 ± 2.4	0.3 ± 0.1	6.9 ± 2.1	38.1 ± 3.1	0.9 ± 0.1	194.4 ± 8.6	294.0 ± 17.1	0.875 ± 0.098

6.3.7 Kinematic distribution of $Z(\rightarrow \nu\nu)$+jets CR after correction

Figure 6.9 shows the comparison of modelling of the $Z(\rightarrow \nu\nu)$+jets to data in the $Z(\rightarrow \nu\nu)$+jets control region after applying event by event reweighting to the γ+jets template obtained from data minus all other non-γ+jets backgrounds. The overall scale factor derived in Table 6.8 is also applied to the estimate of the $Z(\rightarrow \nu\nu)$+jets events. A good agreement is observed in various kinematic variables.

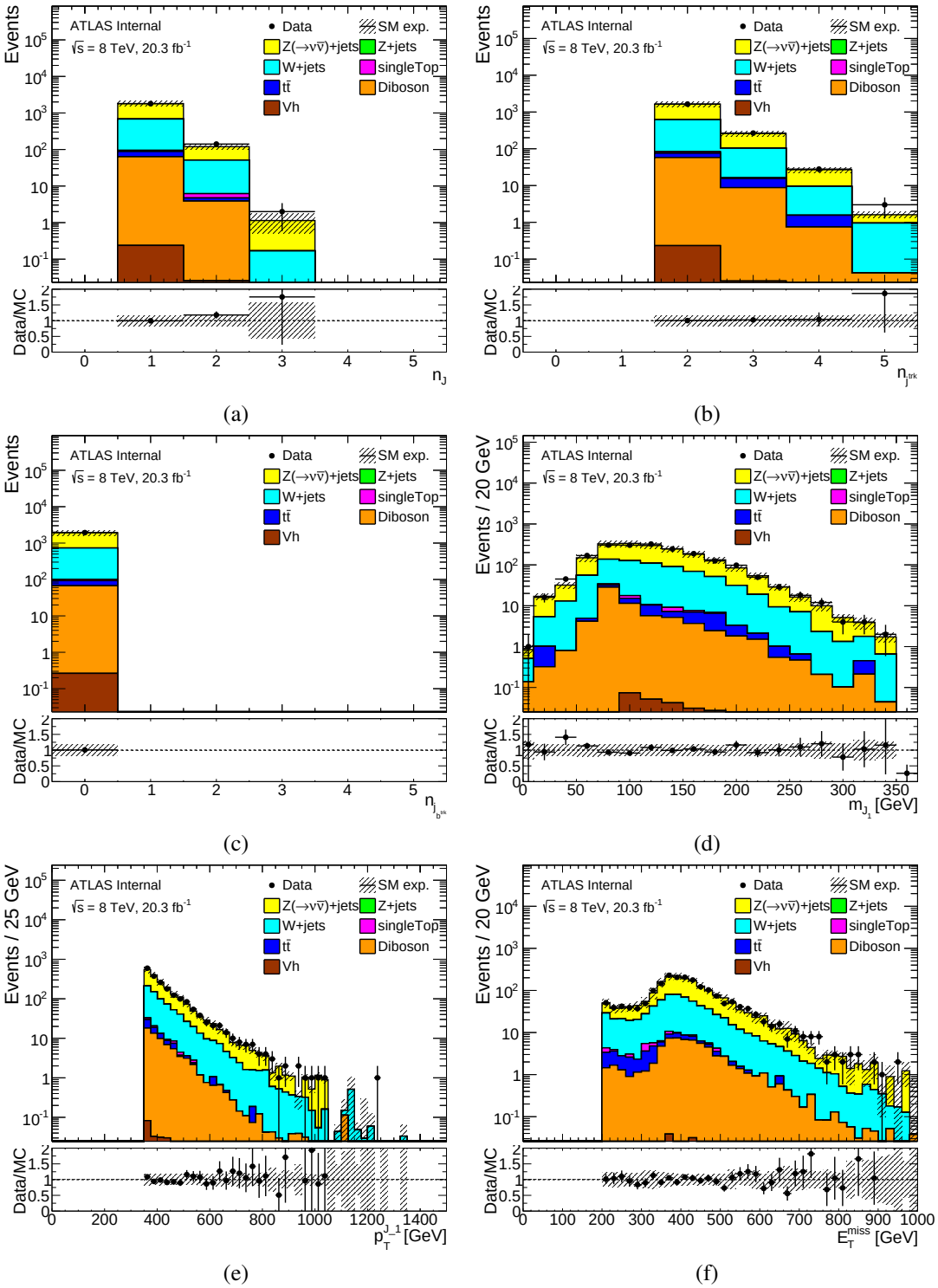


Fig. 6.9 Kinematic distributions in the $Z(\to \nu\nu)$+jets CR after reweighting and after applying scale factor. (a) trimmed AntiKt10 jet multiplicity, (b) number of associated track jet, (c) number of associated b-tagged track jet, (d) leading large-R jet mass, (e) leading large-R jet p_{T} and (f) $E_{\mathrm{T}}^{\mathrm{miss}}$. Both the statistical error and systematics uncertainties are shown.

6.4 Data-driven background estimate of QCD multi-jet

To investigate and estimate the contribution by QCD event due to the resolution or mis-measurement of jet energy, we examine the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in a QCD multi-jet enriched CR. The requirements for this CR are listed in Table 6.10. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the QCD CR is shown in Figure 6.10. Contribution from other events such as $t\bar{t}$ and W+jets are minuscule. The rest of the events are expected to come from QCD multi-jet events. A large fraction of the QCD multi-jet sample have $E_{\mathrm{T}}^{\mathrm{miss}}$ less than 300 GeV.

Table 6.10 Selection for QCD multi-jets enriched region.

Selection	Requirement
Event quality preselection	pass
$E_{\mathrm{T}}^{\mathrm{miss}}$ xe80_tclcw EF trigger	pass
$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	> 100
Number of baseline leptons, n_ℓ	$== 0$
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)\rvert$	< 1.0
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert$	$> pi/2$
Number of large-R jet, n_J	≥ 1
Leading large-R jet p_T, $p_T^{J_1}$ [GeV]	> 350
Number of associated track jet, $n_{j^{trk}}$	≥ 2

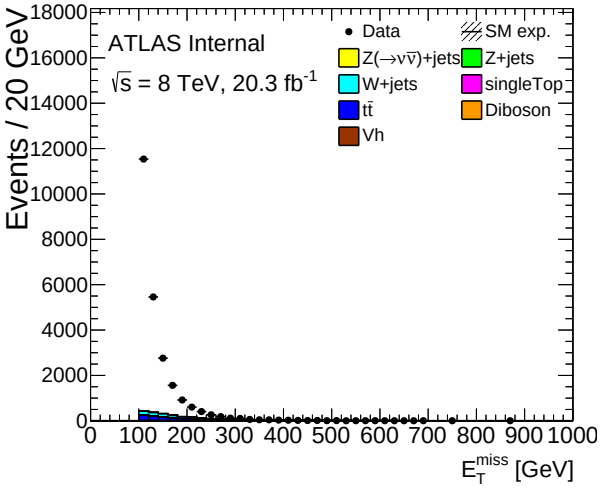


Fig. 6.10 . $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of the QCD control region. No scale factors are applied to MC samples.

Although it is rare for QCD multi-jet events to have high $E_{\mathrm{T}}^{\mathrm{miss}}$, due to the extremely large cross section of QCD multi-jet production compared to that of the $E_{\mathrm{T}}^{\mathrm{miss}} + b\bar{b}$, QCD multi-jet events may still form a considerable background in this analysis. In order to evaluate the QCD multi-jet contribution, we adopt the data-driven approach called "ABCD" method, because in general the QCD multi-jet production is not well modelled in simulation In addition, it is

difficult for MC simulation to reproduce mis-measurement of jet energy that causes fake $E_{\mathrm{T}}^{\mathrm{miss}}$. Furthermore, due to its high production cross section, a considerable amount of MC events would need to be generated in order to match the statistics as collected in the data.

6.4.1 ABCD method

The principle of ABCD method is similar to the CR approach introduced early in the beginning of this chapter. Here the CRs are simply labeled as "A", "B", "C", and "D". To construct these regions, we need two nearly uncorrelated kinematic variables in which the QCD multi-jet production has a markedly different distribution from the other backgrounds (i.e W/Z+jets and $t\bar{t}$ and e.t.c):

1. $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})|$: azimuthal separation between $E_{\mathrm{T}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$
2. $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$: minimum azimuthal separation between $E_{\mathrm{T}}^{\mathrm{miss}}$ and any signal jet

Figure 6.11 shows the $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})|$ versus $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ plane, where the four regions of interest A, B, C, and D can be defined. The sketch of jet, $E_{\mathrm{T}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$ in each region serve as a visual aid to understand the relation of these variables with the composition of the type of events in each region. The back to back structure of the $E_{\mathrm{T}}^{\mathrm{miss}}$and $p_{\mathrm{T}}^{\mathrm{miss}}$ in azimuthal direction (regions C and D) is more indicative of a calorimeter (or an inner-detector) mis-measurement. When both the $E_{\mathrm{T}}^{\mathrm{miss}}$and $p_{\mathrm{T}}^{\mathrm{miss}}$ point in a similar azimuthal direction (regions A and B), this usually signal the present of true $E_{\mathrm{T}}^{\mathrm{miss}}$ originating from non-interacting particles like neutrino and DM. When $E_{\mathrm{T}}^{\mathrm{miss}}$, $p_{\mathrm{T}}^{\mathrm{miss}}$ and the direction of jet are roughly align (region B), it can signify the presence of jet coming from semi-leptonic b-quark decay. In ABCD method, the prediction for QCD multi-jet events in the signal region (region A) can be calculated as:

$$N_{A(SR)}^{QCD} = \frac{N_B}{N_D} \times N_C, \tag{6.8}$$

where N_B, N_C and N_D represent the number of events in data minus the contributions from the other non-multi-jet MC backgrounds in regions B, C and D, respectively.

Figure 6.12 shows the distributions of $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})|$ and $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ for QCD multi-jet and the other events selected based on similar requirements in Table 6.10. The contribution from other events such as $t\bar{t}$ and W+jets are small. The rest of the events are expected to come from QCD multi-jet events. Figure 6.12a shows that the QCD multi-jet background completely dominates in region where $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})| > \pi/2$, while Figure 6.12b shows that the QCD multi-jet background dominates the low $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ region.

The two variables must have non-structured distribution in this 2-D plane for the ABCD method to work. One way to check this assumption is to show the similarity of the $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ distributions for both $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})| < \pi/2$ and $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})| > \pi/2$. This similarity of $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ distributions for QCD dijet MC sample is demonstrated in Figure 6.13. The

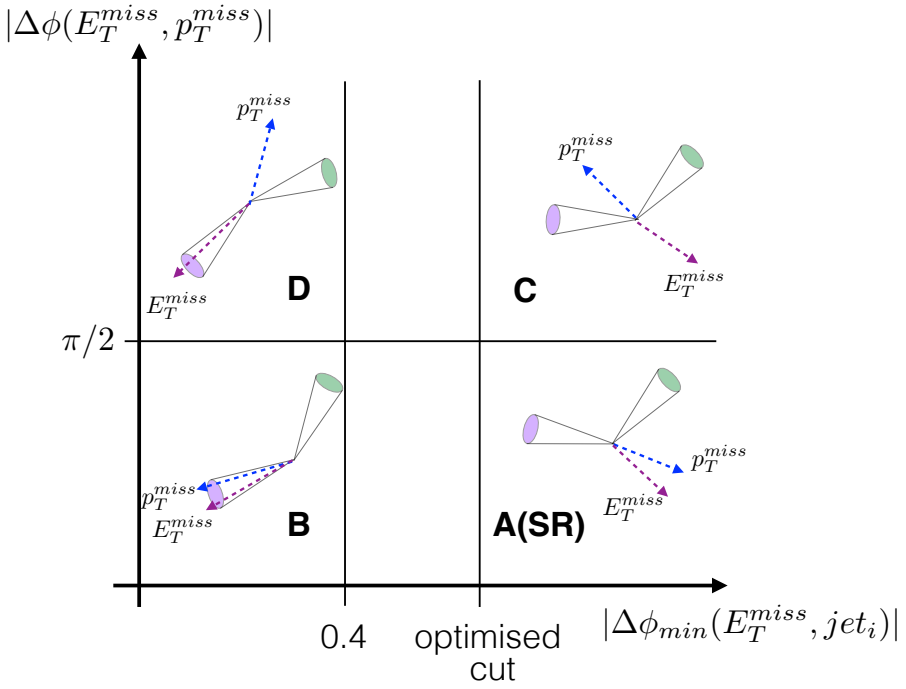
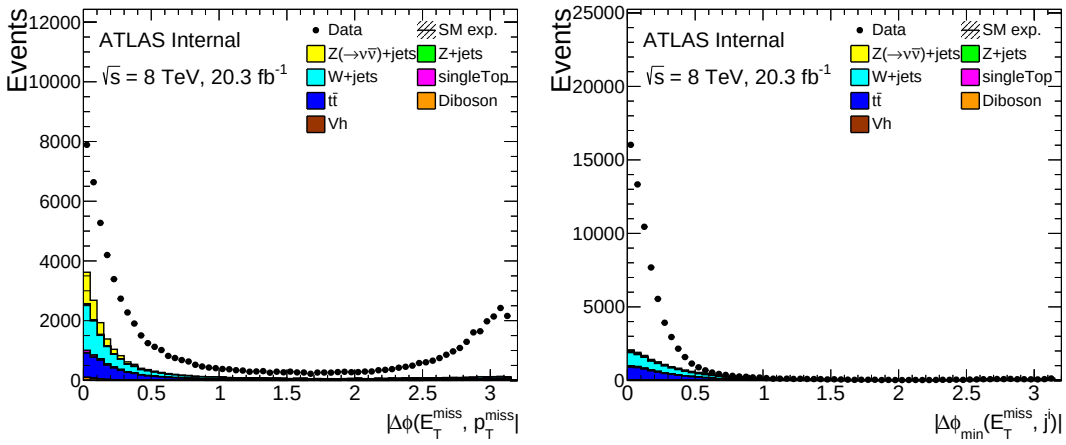


Fig. 6.11 Two-dimensional kinematic plane for estimating QCD multi-jet background in the signal region (region A). The sketch of jets (cones), $E_{\mathrm{T}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\mathrm{miss}}$ in each region help to visualize the relation of the variables between QCD multi-jet, non-multi-jet backgrounds and the signal events.



(a) $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})$ (b) $min|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, jet)|$

Fig. 6.12 $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})$ and $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ distribution in QCD multi-jet CR. Selection listed in Table 6.10 are applied. No scale factors are applied to any MC samples.

$|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ variable is shown to have the same shape above and below $\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}}) = \pi/2$ which supports that the two variables are not correlated. To further build confidence that the $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ variable and the $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})$ are sufficiently de-correlated, a cross check is performed using templates obtained from data after subtracting away contribution from non-multi-jet processes estimated from the simulation. This is illustrated in Figure 6.14. Similarly, both distribution are normalised to the same unit area. The shapes of the two $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ distributions are quite similar.

The number of events in regions B, C and D after all the signal selection (Table 5.1) are too small to produce a robust estimation. Hence a loose selection without the requirement of b-tagging and Higgs mass window is used. In return, a selection rate, R, is introduced to define the probability that an event passes both the two b-tagged jet and Higgs mass window requirements. Equation 6.8 is then rewritten as:

$$N_{A(SR)}^{QCD} = \frac{N_B}{N_D} \times N_C \times R, \tag{6.9}$$

In order to estimate the QCD multi-jet background, we first have to measure the value for R from either region B or region D. We chose to measure R from region B as it has higher statistics. We also loosen the $E_{\mathrm{T}}^{\mathrm{miss}}$requirement to be greater than 100 GeV. By taking the ratio of the number of events after applying both the two b-tagged jet and Higgs mass window requirements to the number of events before applying the cuts (cross reference to Table 5.1), we get $R = 0.007 \pm 0.0005$. With this, we can proceed to calculate $N_{A(SR)}^{QCD}$.

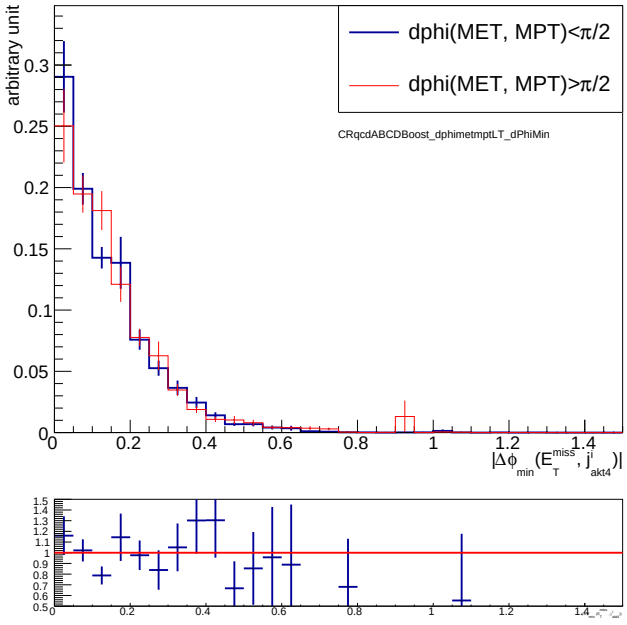


Fig. 6.13 Correlation test for two variables in QCD multi-jet estimation by using ABCD method. This plot are made using PYTHIA8 QCD dijet MC samples. The events are required to pass the event selection in Table 6.10 except the cuts on $\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}}) = \pi/2$ and $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$.

6.4.2 Cross check

Before we calculate $N_{A(SR)}^{QCD}$, it is sound to check that this method works. We checked the method by comparing the predicted number of QCD multi-jet events in different $E_{\mathrm{T}}^{\mathrm{miss}}$ and $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ regions.

First, the ABCD method is repeated in the following three regions with different $E_{\mathrm{T}}^{\mathrm{miss}}$ values:

- $100 < E_{\mathrm{T}}^{\mathrm{miss}} < 150$ GeV,
- $150 < E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV, and
- $100 < E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV.

In order to test the method with larger statistics, $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)| > 0.4$ is used to define regions A and C. Table 6.11 shows that the predicted yields in region A with $100 < E_{\mathrm{T}}^{\mathrm{miss}} < 150$ GeV and $100 < E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV is $23-31\%$ higher than the directly observed yields. For $150 < E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV, the prediction is consistent with the observed yield in region A. However, since the QCD fraction is only 11%, we consider this result less reliable.

We performed two more tests in two different $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ regions:

- $0.4 < |\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)| < 0.9$, and

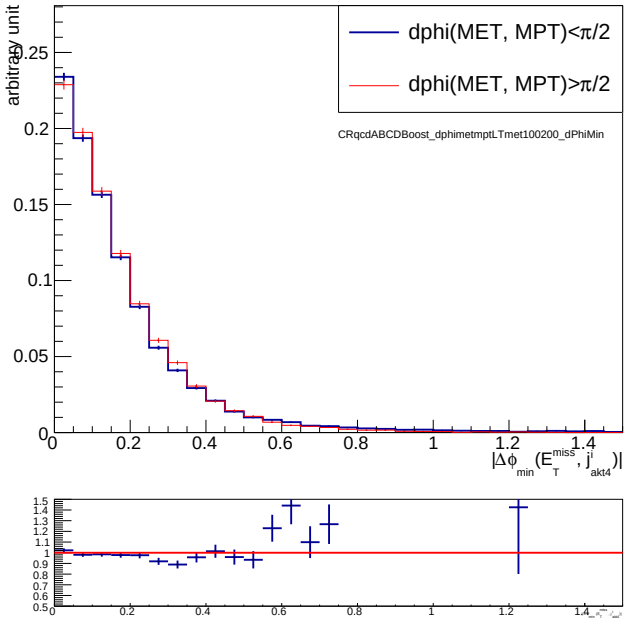


Fig. 6.14 Correlation test for two variables in QCD multi-jet estimation by using ABCD method. This plot is made by subtracting data from non-QCD MC backgrounds. Only events with $100 < E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV are shown here.

- $0.9 < |\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)| < 1.5$

and with $100 < E_{\mathrm{T}}^{\mathrm{miss}} < 200$ GeV. The results are shown in Table 6.12. The predicted yield in the $0.4 < |\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)| < 0.9$ region is 31% higher than the observed number of multi-jet events. In higher $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ region, the QCD fraction decreases drastically, such that for $0.9 < |\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)| < 1.5$ the QCD fraction is only around 5%. This means the estimate will be very sensitive to the systematic variation in other non-QCD backgrounds. Here, the estimated multi-jet yield is consistent with the observed yields.

In summary, these tests suggest that the method is sound. The predicted QCD events in region A with smaller $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ is systematically larger than the statistical uncertainty. On the other hand, the accuracy of QCD estimate with large $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ is limited by the available QCD events. A conservative systematics uncertainty of 31% is assigned to this method itself by taking the largest deviation from the closure test result.

6.4.3 Result for QCD multi-jet background estimation in the SR

To estimate the QCD multi-jet background in the SRs, we apply signal event selections up to the number of associated track jets (see Table 5.1). The boundary of regions A, B, C and D are defined by $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)| = 1.0$ and $|\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})| = \pi/2$. The event yields for each background in each region and with two different $E_{\mathrm{T}}^{\mathrm{miss}}$ values are listed in Table 6.13.

Table 6.11 QCD background estimation closure test. Non-QCD backgrounds are substracted from data. The uncertainties include statistical error only.

Region		N_B	N_D	N_C	N_A^{obs}	N_A^{pred}	N_A^{pred}/N_A^{obs}
$100\ GeV < E_{\mathrm{T}}^{\mathrm{miss}} < 150\ GeV$	Diboson	26 ± 1	13 ± 1	5 ± 1	18 ± 1		
	$W(\to \ell\nu)$+jets	489 ± 11	217 ± 8	115 ± 5	418 ± 10		
	$t\bar{t}$	949 ± 14	367 ± 9	98 ± 4	323 ± 8		
	$Z(\to \nu\nu)$+jets	190 ± 8	58 ± 5	44 ± 6	252 ± 10		
	Single Top	64 ± 5	27 ± 3	9 ± 2	21 ± 3		
	$Z(\to \ell\ell)$+jets	32 ± 2	20 ± 3	4 ± 1	8 ± 1	1622 ± 56	1.23 ± 0.06
	γ+jets	68 ± 14	56 ± 17	1 ± 0	9 ± 2		
	Total non-QCD bkg	1818 ± 25	757 ± 21	275 ± 9	1048 ± 17		
	Data	23518 ± 153	17146 ± 131	1500 ± 39	2365 ± 49		
	QCD	21700 ± 155	16389 ± 133	1225 ± 40	1317 ± 51		
$150\ GeV < E_{\mathrm{T}}^{\mathrm{miss}} < 200\ GeV$	Diboson	28 ± 1	9 ± 1	3 ± 0	17 ± 1		
	$W(\to \ell\nu)$+jets	536 ± 10	145 ± 5	59 ± 3	347 ± 8		
	$t\bar{t}$	782 ± 13	195 ± 6	51 ± 3	285 ± 8		
	$Z(\to \nu\nu)$+jets	227 ± 6	40 ± 4	24 ± 2	252 ± 6		
	Single Top	53 ± 5	14 ± 3	2 ± 1	20 ± 3		
	$Z(\to \ell\ell)$+jets	36 ± 1	16 ± 1	2 ± 0	5 ± 1	159 ± 33	1.4 ± 0.5
	γ+jets	23 ± 9	9 ± 4	0 ± 0	1 ± 0		
	Total non-QCD bkg	1684 ± 20	428 ± 10	140 ± 5	927 ± 13		
	Data	8041 ± 90	3457 ± 59	216 ± 15	1039 ± 32		
	QCD	6357 ± 92	3029 ± 60	76 ± 16	112 ± 35		
$100\ GeV < E_{\mathrm{T}}^{\mathrm{miss}} < 200\ GeV$	Diboson	53 ± 2	22 ± 1	8 ± 1	35 ± 1		
	$W(\to \ell\nu)$+jets	1025 ± 15	362 ± 9	174 ± 6	766 ± 13		
	$t\bar{t}$	1731 ± 19	561 ± 11	149 ± 6	608 ± 11		
	$Z(\to \nu\nu)$+jets	417 ± 10	98 ± 6	68 ± 6	504 ± 11		
	Single Top	117 ± 7	41 ± 4	11 ± 2	40 ± 4		
	$Z(\to \ell\ell)$+jets	68 ± 2	36 ± 3	5 ± 1	13 ± 1	1879 ± 65	1.31 ± 0.07
	γ+jets	91 ± 16	65 ± 17	1 ± 0	10 ± 2		
	Total non-QCD bkg	3502 ± 32	1186 ± 24	416 ± 11	1975 ± 21		
	Data	31559 ± 178	20603 ± 144	1716 ± 41	3404 ± 58		
	QCD	28057 ± 181	19417 ± 145	1300 ± 43	1429 ± 62		

Upon closer inspection on Table 6.13 we know that the QCD fraction decreases as the $E_{\mathrm{T}}^{\mathrm{miss}}$ increases. Region A and C consistently have negative QCD fraction as $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)| > 1.0$ cut combined with high $E_{\mathrm{T}}^{\mathrm{miss}}$ cut had essentially removed all the QCD events. This also implies that QCD background is expected to be negligible in the SRs. Due to this fact, we have to assign an upper limit for QCD background in our SRs.

$E_{\mathrm{T}}^{\mathrm{miss}} > 200\ GeV$ region which has the largest QCD fraction is used for the upper limit derivation. The 68% C.L. upper limit for QCD event estimated in region A is:

$$N^{QCD}_{A(SR),\ 68\%C.L} = \frac{N_B}{N_D} \times N_{C,\ 68\%C.L} \times R \tag{6.10}$$

where $N_{C,\ 68\%C.L} = 2.4$ is the one standard deviation of the statistical uncertainty of N_C predicted in Table 6.13. Plugging in the value for $N_B = 2650$, $N_D = 1262$ and $R = 0.007$, the QCD background upper limit in the SRs is estimated to be 0.04. The systematic error for this estimation is described in more detailed in Section 7.4.3.

Table 6.12 QCD background estimation cross check using different $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ range for region A and region C in $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV bin. For region B and D, $|\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)|$ is fixed at < 0.4. Non-QCD backgrounds are substracted from data. The uncertainties include statistical error only.

	$100\ GeV < E_{\mathrm{T}}^{\mathrm{miss}} < 200\ GeV$						
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, j^i_{akt4})\rvert$ definition for region A(C)		N_B	N_D	N_C	N_A^{obs}	N_A^{pred}	N_A^{pred}/N_A^{obs}
$0.4 < \lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i_{akt4})\rvert < 0.9$	Diboson	53 ± 2	22 ± 1	6 ± 1	22 ± 1	1801 ± 62	1.31 ± 0.07
	$W(\to \ell\nu)$+jets	1025 ± 15	362 ± 9	134 ± 5	523 ± 11		
	$t\bar{t}$	1731 ± 19	561 ± 11	135 ± 5	531 ± 11		
	$Z(\to \nu\nu)$+jets	417 ± 10	98 ± 6	58 ± 6	305 ± 9		
	Single Top	117 ± 7	41 ± 4	9 ± 2	35 ± 4		
	$Z(\to \ell\ell)$+jets	68 ± 2	36 ± 3	4 ± 1	10 ± 1		
	γ+jets	91 ± 16	65 ± 17	1 ± 0	5 ± 2		
	Total non-QCD bkg	3502 ± 32	1186 ± 24	348 ± 10	1432 ± 18		
	Total Data	31559 ± 178	20603 ± 144	1595 ± 40	2811 ± 53		
	QCD	28057 ± 181	19417 ± 145	1247 ± 41	1379 ± 56		
$0.9 < \lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i_{akt4})\rvert < 1.5$	Diboson	53 ± 2	22 ± 1	1 ± 0	11 ± 1	63 ± 15	1.09 ± 0.55
	$W(\to \ell\nu)$+jets	1025 ± 15	362 ± 9	28 ± 3	213 ± 7		
	$t\bar{t}$	1731 ± 19	561 ± 11	11 ± 2	74 ± 4		
	$Z(\to \nu\nu)$+jets	417 ± 10	98 ± 6	9 ± 1	171 ± 7		
	Single Top	117 ± 7	41 ± 4	1 ± 1	6 ± 2		
	$Z(\to \ell\ell)$+jets	68 ± 2	36 ± 3	1 ± 0	3 ± 1		
	γ+jets	91 ± 16	65 ± 17	0 ± 0	2 ± 0		
	Total non-QCD bkg	3502 ± 32	1186 ± 24	51 ± 4	479 ± 11		
	Total Data	31559 ± 178	20603 ± 144	95 ± 10	537 ± 23		
	QCD	28057 ± 181	19417 ± 145	44 ± 10	58 ± 25		

6.5 Kinematic distribution in 0-lepton validation region

Finally all the background estimations are combined in a 0-lepton validation region in order to check the overall modelling of various backgrounds. Events in this VR are required to pass the selection listed in Table 6.6. $E_{\mathrm{T}}^{\mathrm{miss}}$ cut loosened to 200 GeV to allow for more statistics. In addition, no jet mass cut is applied. Only events that contain 0 and 1 b-tagged track jet are considered. The distributions are shown in Figure 6.15. All distributions show that the background predictions agree well with the data.

Table 6.13 Data and non-QCD background yield in each region B, C and D for two $E_{\mathrm{T}}^{\mathrm{miss}}$ bins. The uncertainties include statistical error only.

Region		N_B	N_D	N_C	N_A^{obs}
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$ *GeV*	Diboson	140 ± 3	15 ± 1	2 ± 0	84 ± 2
	$W(\to \ell\nu)$+jets	2314 ± 14	260 ± 6	37 ± 2	700 ± 7
	$t\bar{t}$	1724 ± 19	211 ± 7	3 ± 1	70 ± 4
	$Z(\to \nu\nu)$+jets	864 ± 8	37 ± 3	1 ± 0	1880 ± 8
	Single Top	184 ± 11	24 ± 3	2 ± 1	17 ± 3
	$Z(\to \ell\ell)$+jets	145 ± 2	40 ± 1	2 ± 0	4 ± 0
	γ+jets	8 ± 3	13 ± 6	0 ± 0	0 ± 0
	Total non-QCD bkg	5379 ± 27	600 ± 12	47 ± 2	2754 ± 12
	Total Data	8029 ± 90	1862 ± 43	45 ± 7	2270 ± 48
	QCD	2650 ± 94	1262 ± 45	-2 ± 7	-484 ± 49
$E_{\mathrm{T}}^{\mathrm{miss}} > 250$ *GeV*	Diboson	113 ± 3	10 ± 1	1 ± 0	79 ± 2
	$W(\to \ell\nu)$+jets	1846 ± 11	167 ± 4	30 ± 1	644 ± 6
	$t\bar{t}$	1116 ± 15	109 ± 5	3 ± 1	60 ± 4
	$Z(\to \nu\nu)$+jets	661 ± 6	19 ± 2	1 ± 0	1819 ± 8
	Single Top	146 ± 10	14 ± 3	2 ± 1	13 ± 3
	$Z(\to \ell\ell)$+jets	111 ± 2	29 ± 1	1 ± 0	3 ± 0
	γ+jets	3 ± 2	8 ± 5	0 ± 0	0 ± 0
	Total non-QCD bkg	3996 ± 22	355 ± 9	38 ± 2	2618 ± 11
	Total Data	4632 ± 68	765 ± 28	34 ± 6	2145 ± 46
	QCD	636 ± 72	410 ± 29	-4 ± 6	-473 ± 48

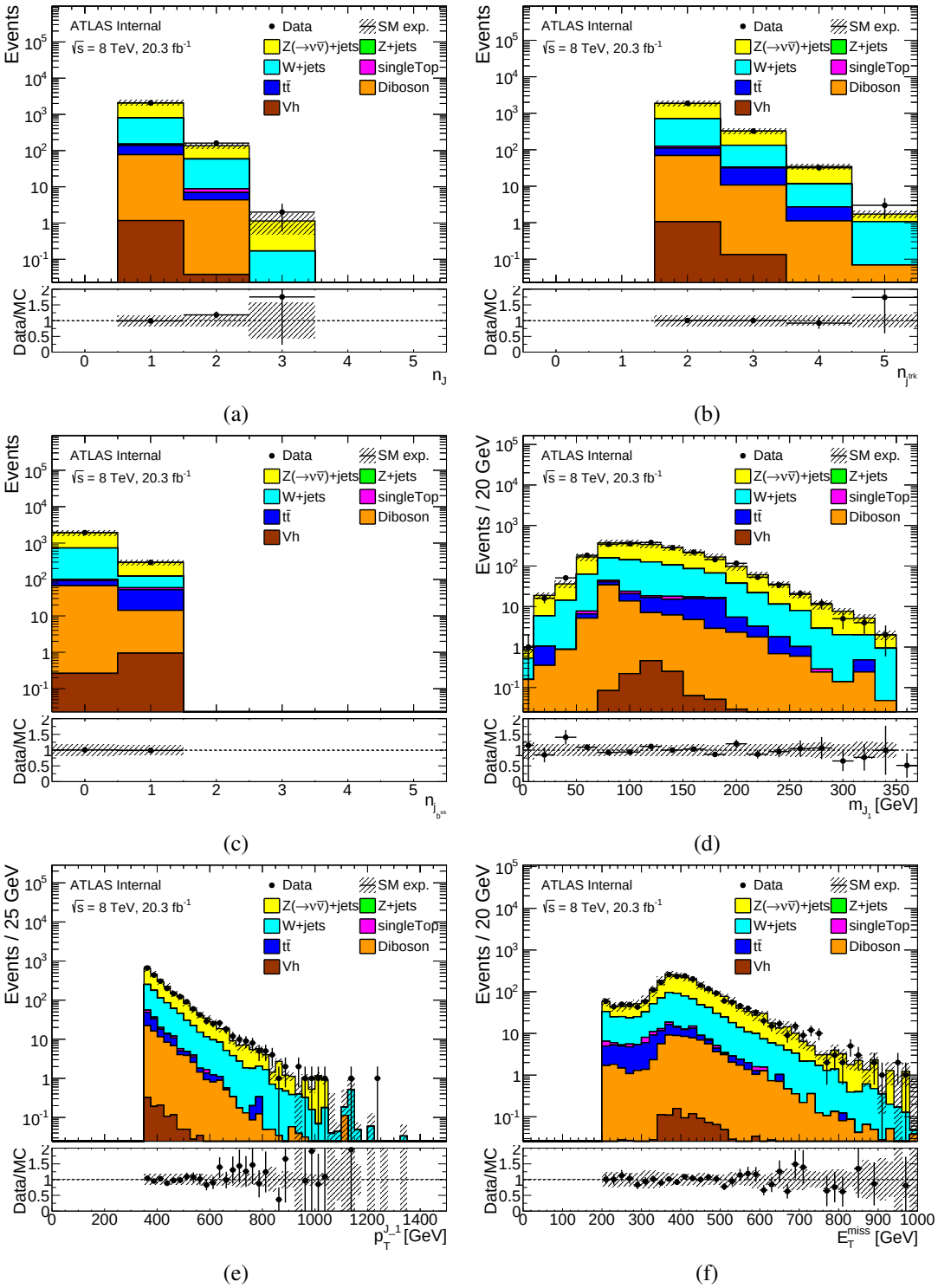


Fig. 6.15 Kinematic distributions for 0-lepton VR after reweighting and applying all scale factors. (a) trimmed AntiKt10 jet multiplicity, (b) number of associated track jet, (c) number of associated *b*-tagged track jet, (d) leading large-*R* jet mass, (e) leading large-*R* jet p_{T} and (f) $E_{\mathrm{T}}^{\mathrm{miss}}$. Both the statistical error and systematics uncertainties are shown.

6.6 Estimated background yield in the signal regions

After applying the full event selection, the total event yields of each background are shown in Table 6.14. The full cutflow for each of the background can be found in Table C.1 in Appendix C. The event yield for all the backgrounds has been scaled based on the integrated luminosity. In addition, the event yield for W+jets, top, and $Z(\to \nu\nu)$+jets backgrounds have also been scaled based on the scale factors derived in previous sections. The largest background contribution comes from $Z(\to \nu\nu)$+jets backgrounds. The QCD multi-jet background number quoted in the table is the upper limit as derived in Section 6.4.

Table 6.14 Event yield of each background in the final SRs. The uncertainties include only the MC statistical errors.

	$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	
Background	> 300	> 400
$Z(\to \nu\nu)$+jets	7.00 ± 1.6	5.2 ± 1.4
W$(\ell\nu)$+jets	1.4 ± 0.2	0.8 ± 0.2
(W/Z)H	1.0 ± 0.1	0.6 ± 0.1
Diboson	0.9 ± 0.1	0.6 ± 0.1
Top	0.8 ± 0.4	0.6 ± 0.3
QCD multi-jet	$< 0.0 \pm 0.04$	$< 0.0 \pm 0.04$
Total bkg.	11.2 ± 1.7	7.7 ± 1.4

Chapter 7

Systematics Uncertainties

This section describes the sources of systematic uncertainty considered in the analysis. These uncertainties are divided into three categories: the theoretical uncertainties (Section 7.1), the uncertainties due to the detector modelling and the reconstruction, collectively called the experimental uncertainties (Section 7.3), and the uncertainties on the background estimation method (Section 7.4).

7.1 Theoretical Uncertainties

The theoretical uncertainties represent our ignorance in the description of a physics process or its implementation in a MC generator. In this analysis, we consider two sources which are expected to have the largest impact, i.e the PDF uncertainty and the cross section uncertainty.

7.1.1 PDF uncertainties

The cross sections in proton-proton collisions are expressed as

$$\sigma^{pp\to X,Y,\dots} = \sum_{i_A,j_B=q(g),\bar{q}(g)} \int dx_1 dx_2 \mathrm{PDF}_i^A(f_1,x_1,\mathscr{Q}^2,\dots)\mathrm{PDF}_j^B(f_2,fx_2,\mathscr{Q}^2,\dots)\sigma^{i_A j_B\to X,Y,\dots}(x_1,x2,\alpha_s,\dots) \tag{7.1}$$

where $\sigma^{i_A j_B\to X,Y,\dots}$ is the parton-parton cross section induced by i_A-th and j_B-th partons in the incoming protons A and B, respectively. PDF_i^A and PDF_j^B are the i-th and j-th PDF of proton A and B, which depend on the flavor of the interacting partons (fl_1 and fl_2), the momentum of each interacting parton (x_1 and x_2) and the momentum transfer $\mathscr{Q}^2$. Since the PDFs are determined empirically using data with some uncertainties, the inclusion of any PDF into the cross section calculation introduces a systematic error. Thus, it is important to estimate the impact of PDF uncertainty on the event selection efficiency.

There are two components of PDF uncertainty that must be considered:

- Intra-PDF uncertainty: this is the uncertainty within a given PDF set. Usually a central PDF that is used for the generation of a sample is accompanied by a set of error PDFs representing the uncertainties (e.g. fit uncertainty) within a given PDF family. The prescription to evaluate this type of uncertainty is PDF-dependent.

- Inter-PDF uncertainty: this is the variation when switching from one PDF set to another PDF set. The comparison is made using the result obtained by each PDF and comparing the variation of the observables.

The total PDF uncertainty is the combination of the inter- and intra-PDF uncertainties. It is obtained by taking the envelope of the variations and uncertainties. Our signal samples are produced with the CTEQ6L1 [158] LO PDF set using MADGRAPH. The CTEQ6L1 PDF does not have associated error sets. We use another two PDF sets (i.e. the MSTW2008LO PDF set and NNPDF2.1 PDF) to estimate the PDF systematic uncertainty. The exact procedures to evaluate each component and their combination will be discussed in the following paragraphs.

Calculating the intra-PDF uncertainty

The MSTW2008LO PDF set consists of 1 central PDF and 40 other PDF sets representing the uncertainty[1]. To evaluate the intra-PDF uncertainty for MSTW2008LO PDF set, we use the asymmetric Hessian method [159]. For an observable X let us denote its value using the central PDF as X_0. The X^+_{2i} (X^-_{2i-1}) is the value of the observable obtained by the PDF corresponding to the upward (downward) fluctuation for the i-th PDF error set. The asymmetric errors for MSTW2008LO PDF set are calculated using the following formulae

$$\begin{aligned} \Delta X^+_{max} &= \sqrt{\sum_{i=1,\ldots N/2}^{N=40} \Big[max(X^+_{2i} - X_0, X^-_{2i-1} - X_0, 0) \Big]^2} \\ \Delta X^-_{max} &= \sqrt{\sum_{i=1,\ldots N/2}^{N=40} \Big[max(X_0 - X^+_{2i}, X_0 - X^-_{2i-1}, 0) \Big]^2} \end{aligned} \tag{7.2}$$

where N is the number of PDF error sets (40 for MSTW2008LO PDF set).

Next we evaluate the intra-PDF uncertainty for the NNPDF2.1 PDF set. Unlike MSTW2008LO PDF set which provide a central value and some error PDFs with parameters varied, the NNPDF2.1 PDF set consist of one central PDF and a MC ensemble of 100 PDF replicas of the original dataset. The data points of the replicas are distributed according to the central values and uncertainties specified by the measurements. The best-fit PDF is constructed for each replica. This procedure is repeated N times, resulting in an ensemble of N PDFs. Hence

[1] A PDF can have n uncorrelated parameters. Each parameter can be varied independently by +/- 1 sigma and a new (error) PDF is calculated. This new PDF is basically what should be used to evaluated the systematic effects on an observable.

the intra-PDF uncertainty for NNPDF2.1 PDF set is calculated as the standard deviation of the ensemble,

$$\Delta X = \sqrt{\frac{1}{N-1}\sum_{i=1,2,...N}(X_i - X_0)^2} \tag{7.3}$$

where N is the total number of PDF in the ensemble.

Calculating the total uncertainty

As per the PDF4LHC prescription [160], the total PDF uncertainty of an observable is the envelope of the combination of the inter- and intra-PDF uncertainties. That is half of the extremum (min and max) of all variations. It is calculated as

$$\begin{aligned}\Delta X_{total} = \frac{1}{2}\Bigg[& max\left(X_0^{NNPDF} + \Delta X,\ X_0^{MSTW2008LO} + \Delta X^{+}_{max}\right) - \\ & min\left(X_0^{NNPDF} - \Delta X,\ X_0^{MSTW2008LO} + \Delta X^{-}_{min}\right)\Bigg]\end{aligned} \tag{7.4}$$

The resulting envelope of maximum deviation is then symmetrised. To obtain the relative uncertainty, ΔX_{total} is divided by the nominal value X_0 obtained using the CTEQ6L1 PDF.

Obtain the variation of an observable given a new PDF

With the uncertainty calculation for the MWST and NNPDF families defined, now the questions are what observable X we should use, and how to obtain the value for the observable given a new PDF or error PDF. In our case, the observable is taken to be the selection efficiency, ε. Ideally, whenever we change to a new PDF or error PDF, we have to generate a new MC sample using the new PDFs and its associated generator tunes (see Section 3.2). We then re-run the whole analysis to obtain the new selection efficiency. This method is more accurate (since it goes through the whole chain of the sample production, i.e. the event generation, the detector simulation, the digitisation and the reconstruction) but it is very resource consuming. Hence, we adopt a second method in which we reweight the MC samples with a relative event weight, given a new PDF[1]. The general idea and procedures are described as follows.

Suppose that our MC events are generated using a nominal of central PDF, call it PDF_0, in a generator. The event selection efficiency for this PDF, ε_0 after applying all the event selections

[1]Compared to the first approach, each event is simulated only once, so the kinematics do not change and there is no residual statistical variation in uncertainty. The concern involved with this method is that re-weighting events only affects the hard process. It does not correctly take into account the Sudakov form factors which is used for parton showers calculations. However, the impact of this was shown to be negligible as demonstrated in Reference [161].

is defined as

$$\varepsilon_0 = \frac{N_0^{cuts}}{N_0^{gen}} \tag{7.5}$$

where N_0^{gen} is the total number of events generated by the generator and N_0^{cuts} is the number of selected events after applying all the selection criteria.

Now suppose that we use the same generator but with new PDFs to generate our MC events. This new PDF can be an error PDF or a new central value PDF. Let us denote the new PDFs as PDF_i, where $i = 0, 1, 2, ...N$ denotes the i-th PDF. The event weight (relative probability of producing a particular event), w_n^i for each event n can be defined as

$$w_n^i = \frac{PDF_i(f_1, x_1, Q^2)}{PDF_0(f_1, x_1, Q^2)} \times \frac{PDF_i(f_2, x_2, Q^2)}{PDF_0(f_2, x_2, Q^2)}. \tag{7.6}$$

The new event selection efficiency corresponding to the alternative PDF can be obtained by using w_n^i as in

$$\varepsilon_i = \frac{N_i^{cuts}}{N_i^{gen}} = \frac{\sum_{n=1}^{N_0^{cuts}} w_n^i}{\sum_{n=1}^{N_0^{gen}} w_n^i} \tag{7.7}$$

Result

An example distribution of PDF systematics variation in the $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV SR for `mx1000_xdxhDh` signal sample is depicted in Figure 7.1. The first point (or bin) is the acceptance (marked by the solid red line) for the nominal PDF (CTEQ6L1). The next 41 points covered by the magenta band are the acceptance for the MSTW2008LO68cl PDF set while those inside the green band are for the NNPDF2.2 PDF set. The total PDF uncertainty calculated using Equation 7.4 is listed in Table 7.1 and Table 7.2 for SRs with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV, respectively. The acceptance obtained using the central value for the MSTW2008LO68cl and mean of the NNPDF2.2 PDF sets differ only slightly. However the intra-PDF uncertainties for MSTW2008LO68cl PDF set are larger than that for NNPDF2.2 PDF set. Hence, the total PDF uncertainty obtained by the envelope method is mostly dominant by the uncertainties from the MSTW2008LO68cl PDF set.

The systematic uncertainties in the SRs due to the choice of the PDFs are also taken into account for the diboson (5.9%), W/Z+jets (5%), $t\bar{t}$ (6%) and VH (2.5%) processes. Their values are cited from Reference [46] with the similar phase space. Because the $Z(\rightarrow \nu\nu)$+jets estimation is data-driven, we assume no additional PDF uncertainties, and the theoretical uncertainties are absorbed in the uncertainty on the transfer function, whose derivation is as detailed in Section 7.4.2.

Table 7.1 PDF uncertainty calculated in terms of acceptance for signal samples after $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV cut.

	CTEQ6L1 (%)	MSTW2008LO (%)	NNPDF2.1 (%)	total uncertainty (%)
mx1_xgxFhDh	12.14 ± 0.17	$12.04^{+0.57}_{-0.65}$	12.04 ± 0.17	5.0
mx65_xgxFhDh	12.10 ± 0.17	$12.03^{+0.41}_{-0.58}$	12.03 ± 0.15	4.1
mx100_xgxFhDh	12.08 ± 0.17	$11.98^{+0.45}_{-0.60}$	11.98 ± 0.15	4.4
mx500_xgxFhDh	13.12 ± 0.17	$13.33^{+2.51}_{-1.63}$	13.41 ± 0.59	15.8
mx1000_xgxFhDh	13.76 ± 0.18	$14.41^{+4.41}_{-3.91}$	14.50 ± 1.57	30.2
mx1_xdxhDh	0.68 ± 0.04	$0.72^{+0.15}_{-0.14}$	0.72 ± 0.03	20.9
mx65_xdxhDh	3.04 ± 0.08	$3.12^{+0.49}_{-0.44}$	3.12 ± 0.11	15.3
mx100_xdxhDh	4.10 ± 0.09	$4.19^{+0.61}_{-0.52}$	4.19 ± 0.13	13.7
mx500_xdxhDh	8.63 ± 0.15	$8.38^{+0.25}_{-0.27}$	8.38 ± 0.07	3.0
mx1000_xdxhDh	8.58 ± 0.14	$8.33^{+1.30}_{-0.57}$	8.40 ± 0.44	10.9
mx1_xxhhg5	2.48 ± 0.07	$2.68^{+0.52}_{-0.52}$	2.68 ± 0.11	20.9
mx65_xxhhg5	2.79 ± 0.08	$2.99^{+0.56}_{-0.54}$	2.99 ± 0.12	19.7
mx100_xxhhg5	4.48 ± 0.10	$4.67^{+0.64}_{-0.65}$	4.67 ± 0.14	14.4
mx500_xxhhg5	8.00 ± 0.13	$8.07^{+0.21}_{-0.20}$	8.06 ± 0.07	2.6
mx1000_xxhhg5	8.10 ± 0.13	$8.18^{+0.63}_{-0.28}$	8.24 ± 0.44	6.2
mx1_xxhh	2.21 ± 0.07	$2.39^{+0.45}_{-0.45}$	2.38 ± 0.10	20.4
mx65_xxhh	2.62 ± 0.07	$2.82^{+0.50}_{-0.52}$	2.82 ± 0.11	19.5
mx100_xxhh	3.82 ± 0.09	$4.02^{+0.58}_{-0.61}$	4.02 ± 0.12	15.5
mx500_xxhh	8.01 ± 0.13	$8.06^{+0.14}_{-0.18}$	8.05 ± 0.04	2.0
mx1000_xxhh	8.02 ± 0.13	$8.10^{+0.57}_{-0.27}$	8.14 ± 0.34	5.4

Table 7.2 PDF uncertainty calculated in terms of acceptance for signal samples after $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV cut.

	CTEQ6L1 (%)	MSTW2008LO (%)	NNPDF2.1 (%)	total uncertainty (%)
mx1_xgxFhDh	10.63 ± 0.15	$10.48^{+0.51}_{-0.57}$	10.48 ± 0.14	5.1
mx65_xgxFhDh	10.62 ± 0.15	$10.50^{+0.37}_{-0.54}$	10.50 ± 0.14	4.3
mx100_xgxFhDh	10.63 ± 0.15	$10.50^{+0.52}_{-0.70}$	10.50 ± 0.18	5.7
mx500_xgxFhDh	11.93 ± 0.16	$12.06^{+2.00}_{-1.32}$	12.12 ± 0.47	13.9
mx1000_xgxFhDh	12.63 ± 0.17	$13.13^{+3.32}_{-3.00}$	13.27 ± 1.35	25.0
mx1_xdxhDh	0.47 ± 0.03	$0.49^{+0.12}_{-0.11}$	0.49 ± 0.03	23.6
mx65_xdxhDh	2.32 ± 0.07	$2.39^{+0.46}_{-0.40}$	2.39 ± 0.10	18.5
mx100_xdxhDh	3.22 ± 0.08	$3.31^{+0.60}_{-0.49}$	3.31 ± 0.12	17.0
mx500_xdxhDh	7.59 ± 0.14	$7.40^{+0.16}_{-0.16}$	7.40 ± 0.05	2.4
mx1000_xdxhDh	7.85 ± 0.13	$7.63^{+0.91}_{-0.43}$	7.67 ± 0.29	8.5
mx1_xxhhg5	1.86 ± 0.06	$2.03^{+0.45}_{-0.45}$	2.03 ± 0.10	24.1
mx65_xxhhg5	2.11 ± 0.07	$2.28^{+0.49}_{-0.48}$	2.27 ± 0.10	23.0
mx100_xxhhg5	3.54 ± 0.09	$3.72^{+0.62}_{-0.61}$	3.71 ± 0.13	17.4
mx500_xxhhg5	7.12 ± 0.12	$7.18^{+0.17}_{-0.20}$	7.17 ± 0.05	2.6
mx1000_xxhhg5	7.43 ± 0.13	$7.48^{+0.18}_{-0.17}$	7.49 ± 0.17	2.4
mx1_xxhh	1.61 ± 0.06	$1.75^{+0.39}_{-0.38}$	1.75 ± 0.08	24.1
mx65_xxhh	1.98 ± 0.06	$2.15^{+0.45}_{-0.45}$	2.15 ± 0.10	22.8
mx100_xxhh	3.03 ± 0.08	$3.20^{+0.55}_{-0.56}$	3.20 ± 0.12	18.3
mx500_xxhh	7.13 ± 0.12	$7.18^{+0.20}_{-0.21}$	7.17 ± 0.04	2.8
mx1000_xxhh	7.34 ± 0.13	$7.40^{+0.25}_{-0.19}$	7.41 ± 0.18	3.0

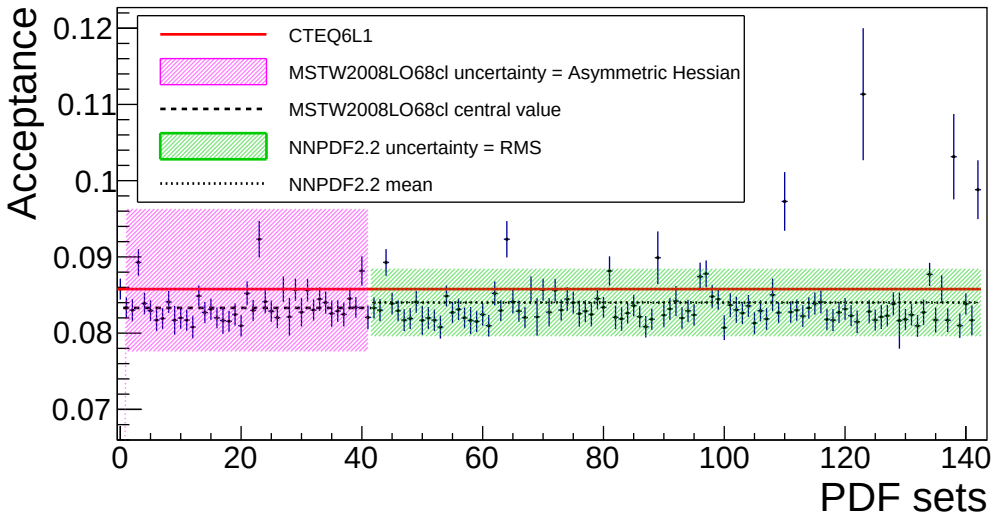


Fig. 7.1 Example of PDF systematic uncertainty calculation. The PDF variation shown here belongs to signal sample `mx1000_xdxhDh`.

7.1.2 Cross section uncertainties

The MC signals samples are produced at leading order (LO). To take into account the uncertainty due to the NLO corrections, an estimated value of 10% is assigned as the uncertainty. The value is cited from the phenomenology study on the $E_{\mathrm{T}}^{\mathrm{miss}}$+jets search for DM [162]. A large uncertainty of 20% is assumed to be associated with the W+jets background. The value is taken from the recent ATLAS measurement of W+jets production with b-jets [163]. Given that the mismodelling (18%) of the W+jets background observed in the CR is fully covered by this cross section uncertainty, no additional systematic uncertainty is assigned to account for the impact of correction on the W+jets normalisation. For $t\bar{t}$ production, an uncertainty of 7% is cited from the theoretical calculations [164], which is consistent with the ATLAS measurement of top quark pair production [165]. However, a 10% non-closure between the data and MC events in the top CR is observed. Consequently, the largest of them is taken as the systematic uncertainty on the $t\bar{t}$ production. The uncertainty on the simulated diboson background cross-section increases from 20% for $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV to 30% for $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV [45]. For vector boson plus Higgs boson production, an uncertainty of 3.1% on the cross-section is estimated from the theoretical calculations [166] and is applied here.

7.2 Luminosity

The uncertainty on the integrated luminosity is $\pm 2.8\%$. It is derived from a preliminary calibration of the luminosity scale derived from the beam-separation scans performed in November 2012, following the same methodology detailed in Reference [167]. If the assumed luminosity varies, the amount of predicted signal or background events varies simultaneously

while the shape of distributions are not affected. It is not only applied as a constant shift on the overall normalisation of each simulated signal and backgrounds ($W(\to \ell\nu)$+jets, $Z(\to \ell\ell)$+jets, $t\bar{t}$ and single top) processes, but also propagated through the data-driven backgrounds (QCD multijet and $Z(\to \nu\nu)$+jets) estimation during the background (non QCD multijet, non γ+jets) subtraction process.

7.3 Detector related uncertainties

In this section the systematic uncertainties related to the detector modelling will be described. These uncertainties affect the reconstruction of the physics objects (leptons, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$). The detector related systematic uncertainties are evaluated by using a common procedure. Each source of uncertainty is varied upward and downward by one standard deviation compared to the nominal settings. The typical amount of variation for each systematic uncertainty is described below. The effect of each systematic variation is then propagated through the complete chain of analysis, including the event reconstruction, the correction and the selection. The difference between the nominal result and the result with systematic variation is assigned as the uncertainty. For one-sided uncertainties, as in the case of the jet energy resolution uncertainty, the relative uncertainty on the final variable is symmetrized under the assumption that the resulting variation is of the same size in both directions.

Notice that in order to enlarge the statistics of $t\bar{t}$ background to allow meaningful systematic studies, we relaxed the b-tagging requirement such that at least one b-tagged track jets are required. To ensure that the kinematics of $t\bar{t}$ events are not biased by this change, from a sample of events passing the two track jet selection (see Table 5.1), we checked the ratio of several kinematic distributions after 2 b-tag and ≥ 1 b-tag selection. As shown in Figure 7.2, no noticeable bias in the distributions are observed.

7.3.1 Pile-up rescaling

As explained in Section 5.5.1 the pileup rescaling is applied to all the events with the nominal value equals to 1.09. The uncertainty for this scale factor is 4%. It is calculated as the sum in quadrature of the uncertainties in the cross-section (σ_{inel}^{vis} and σ_{inel}) measurements, the uncertainty in the extrapolation from 7 TeV to 8 TeV measurement, and the uncertainty in the extrapolation of the scale factor to the inner detector acceptance.

7.3.2 Jet vertex fraction (JVF) uncertainty

The nominal JVF cut value is 0.5. It is varied by ± 0.1 for the systematic evaluation. Because the JVF cut is applied only to jets with p_T <50 GeV and $|\eta|$ <2.4, the size of the uncertainty on the total background yield in the SR is expected to be small ($< 1\%$).

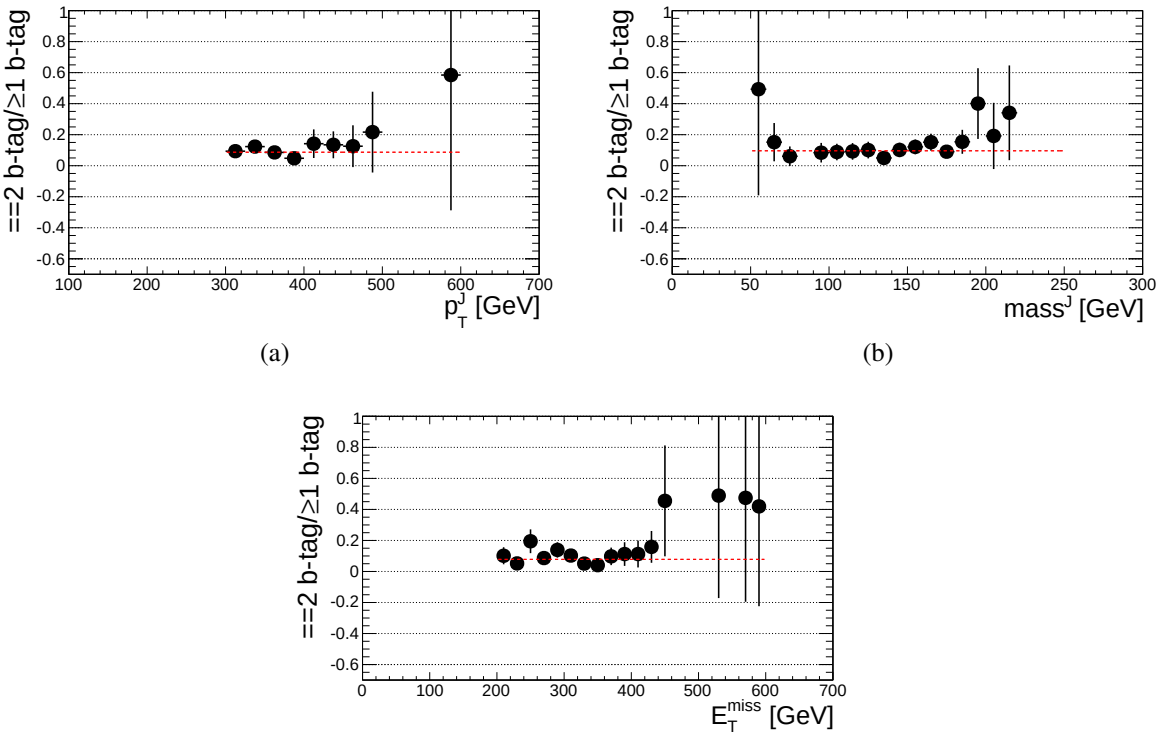
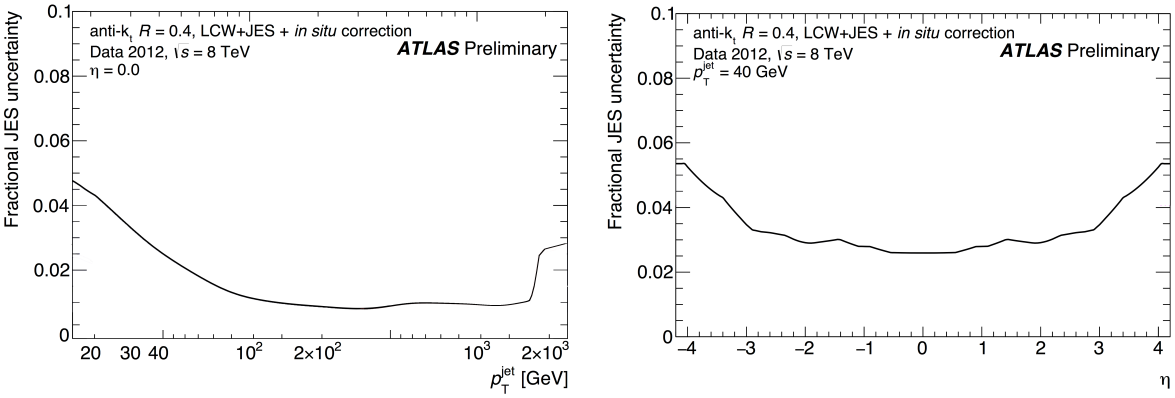


Fig. 7.2 The ratio of (a) the leading large-R jet p_{T}, (b) leading large-R jet mass and (c) $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions between 2 b-tag and ≥ 1 b-tag.

7.3.3 Jets energy scale and resolution uncertainties

The evaluation and the correction of JES have been described in Appendix B.2. There is an uncertainty associates to each correction step, as well as uncertainties due to different jet response for different jet flavour (gluon, light-quark and b-jets). Each uncertainty is treated independently of the others and fully correlated in p_{T} and η bins. The quadratic sum of all the uncertainty sources constitute the total JES uncertainty. Their values as a function of jet p_{T} and η are shown in Figure 7.3. The JES variations are performed simultaneously for all the jets in an event to preserve the information about the correlations between the p_{T} and η bins. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is also recomputed by using the four momentum of the jets with p_{T} varied by $\pm 1\sigma$. Because the number of jets present in the selected events are small, a small sensitivity of the measurement to variations of the JES is expected.



(a) (b)

Fig. 7.3 The total JES uncertainty as a function of (a) p_{T} for central jets and (b) η for jets with $p_{\mathrm{T}} = 40$ GeV. The plots are adapted from Reference [129].

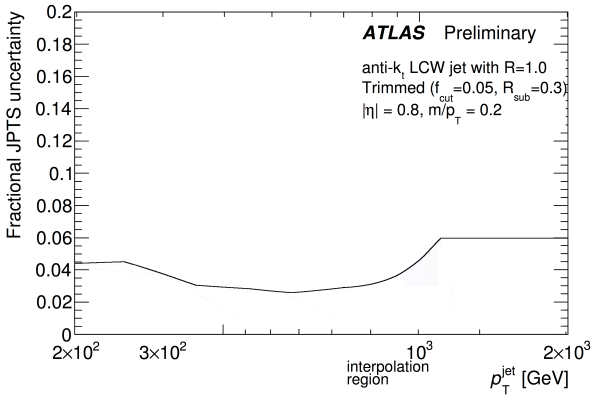
To assess the impact of the JER on the final result, a smearing factor and its associated uncertainty is applied as a scale factor to the energy of all jets in the event, thus worsening the jet energy resolution. The smearing factor is randomly sampled from a Gaussian distribution centered at 1, with width

$$\sigma(p_{\mathrm{T}}^{j},\eta) = \sqrt{(\sigma_{MC}(p_{\mathrm{T}}^{j},\eta) + \Delta\sigma_{data}(p_{\mathrm{T}}^{j},\eta))^2 - \sigma_{MC}^2(p_{\mathrm{T}}^{j},\eta)} \tag{7.8}$$

where $\sigma(p_{\mathrm{T}}^{j},\eta)$ is the measured JER, and $\Delta\sigma(p_{\mathrm{T}}^{j},\eta)$ is the corresponding uncertainty. The difference between the nominal and smeared results is taken as the JER systematic uncertainty. The effect on the final variable is then symmetrised to obtain the a symmetric error.

7.3.4 Jets energy scale and mass scale uncertainties for large-R jet

As discussed in Section 4.7, JES for large-R jet is measured in-situ by comparing the jet energy to that of a well calibrated reference object (photon and track jet) as shown in Equation 4.2. This JES uncertainty is shown as a function of p_{T} in Figure 7.4. Figure 7.5 shows the JMS uncertainties for large-R jets in three different detector regions. The JMS uncertainties are derived in different bins of p_{T}, $|\eta|$ and m/p_{T}. It serves merely as an illustration of the magnitude of the overall JMS uncertainty. The JES and JMS uncertainty are applied to every selected large-R jet in an event in a fully correlated way. The $E_{\mathrm{T}}^{\mathrm{miss}}$is not recalculated for each JES variation.



(a)

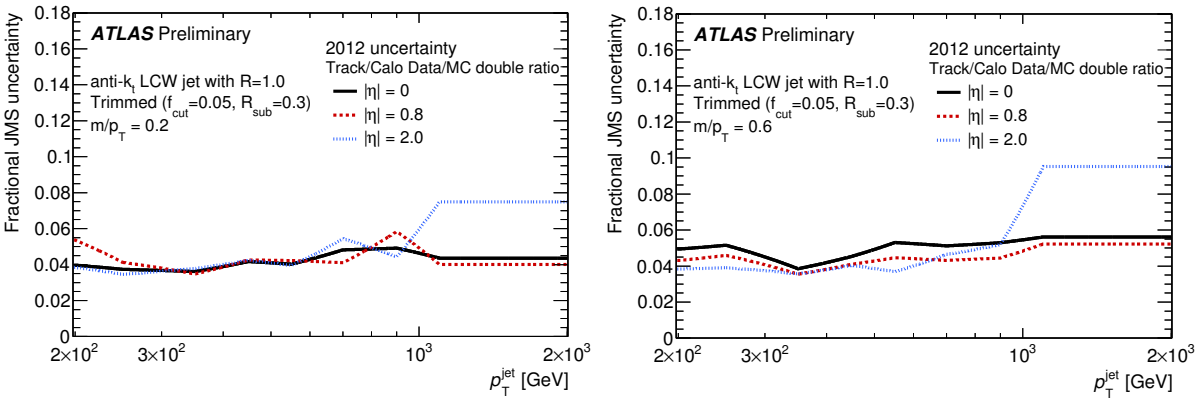
Fig. 7.4 The total JES uncertainty as a function of p_{T} for large-R jets in the region with $\eta = 0.8$. These numbers serve merely as an illustration of the magnitude of the overall JES uncertainty. The plot is adapted from Reference [129].

7.3.5 Jet energy resolution and mass resolution uncertainties for large-R jets

The impact of JER (JMR) uncertainty is obtained in this analysis by smearing the large-R jets energy (mass) by a Gaussian such that the intrinsic resolution is increased by 20% [135, 150, 149]. The effect on the event yield in the SR is then symmetrised to obtain the a two-sided error.

7.3.6 Uncertainties of the $E_{\mathrm{T}}^{\mathrm{miss}}$

The calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ has been discussed in Section 4.10. Due to the object based approach, the uncertainties on the physics objects (jets, muons, electrons and photons) are directly propagated into the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$. Apart from that, the additional uncertainties



(a) (b)

Fig. 7.5 JMS uncertainties for large-R jet in different detector regions for two values of m/p_T: (a) m/p_T= 0.2 and (b) m/p_T= 0.6. The plots are taken from Reference [133].

on the soft term (calculated from calorimeter clusters which have not been associated with the reconstructed object) have to be accounted for. They include the variations of the scale and resolution of the soft term. The size of these uncertainties are negligible.

7.3.7 Flavor tagging uncertainties for track jet

The uncertainties on the b-tagging scale factor for c- (light-) flavour track jets range from 7.7%-15% (15%-31%) [168]. On the other hand, the uncertainties of the b-tagging efficiency on b-flavor track jets with $p_T < 250$ GeV is approximately 2.6%-7.5% depending on p_T. Not enough statistics is available for track jets with $p_T > 250$ GeV in data to accurately determine the scale factor uncertainties. Therefore, the extrapolation is needed. The extrapolation is performed in MC sample by varying the inputs to b-tagging and the detector related parameters. The variations observed with respect to the last track jet p_T bin (100-250 GeV) are taken as the extrapolation uncertainty of the variation. The final extrapolation uncertainties for high p_T track jet varies from 8.1%-22.5% [169]. The largest contributions to the systematic uncertainty on the scale factors for b-flavor track jets come predominantly from the choice of the MC generator, the difference in hadronization models, the modeling of initial and final state radiation and as well as from the uncertainties on the track reconstruction efficiency. The final systematic uncertainty due to the b-tagging is obtained by summing the contribution of each jet flavor in quadrature.

7.3.8 Photon energy scale (PES) and resolution (PER) uncertainties

Photons are used in the estimation of $Z(\rightarrow \nu\nu)$+jets background during the construction of the transfer function. To restore agreement of photon energy scales between the data and MC, a

correction scale factor is applied. The correction has a set of systematic uncertainties which include the uncertainties arising from the extrapolation of the energy scale of electrons to that of photons, uncertainty on the presampler scale, uncertainty on the amount of material upstream of the calorimeter. A full detail of the systematic uncertainty sources can be found in Reference [119]. To assess the impact of the PES (PER) on the final event yield in the SRs, the photon energy scale (resolution) is varied (smeared) by 1σ to get the systematics uncertainty.

7.3.9 Photon identification uncertainty

Similarly the uncertainty of photon isolation enter into our final result via only the $Z(\rightarrow \nu\nu)$+jets background. An overall conservative estimate of 4%, following the studies of the 8 TeV “Mono-photon" search (Ref. [170]) are used to estimate its impact on this analysis.

7.4 Background modelling uncertainties

7.4.1 Uncertainty for top p_{T} reweighting

The differential cross sections of $t\bar{t}$ production measured by the ATLAS experiment shows that the transverse momentum of the top quark, p_T^{top} is mismodeled in POWHEG+PYTHIAMC [171]. As can be seen in Figure 7.6, there is a general trend of MC being harder in p_T^{top} above 200 GeV compared to all MC generators.

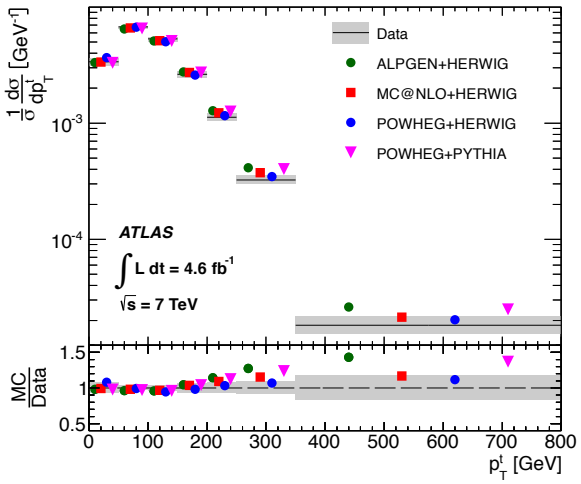


Fig. 7.6 Normalised differential cross-sections for the transverse momentum of the top quark at parton level. The markers are offset within each bin to allow for better visibility. The lower part of figure shows the ratio of the generator predictions to data. The last bins includes events in the overflow bin. Figure taken from Reference [172].

To remedy this, the correction factors derived from the measurement are used to reweight the $t\bar{t}$ events as a function of the average p_{T} of the top and anti-top quarks[1] before parton showering[2]. In order to take into account the kinematic difference between the measurement and this analysis, half the correction is assigned as a systematic error.

Table 7.3 The correction factor for the p_T^{top} from the unfolding measurement [172]. The total uncertainties from the measurement are also shown. These factor are used to reweight the $t\bar{t}$ events.

p_T^{top} [GeV]	0-50	50-100	100-150	150-200	200-250	250-350	350-
Correction factor	$1.051^{+0.088}_{-0.083}$	$1.029^{+0.028}_{-0.026}$	$0.999^{+0.038}_{-0.038}$	$0.976^{+0.049}_{-0.054}$	$0.915^{+0.057}_{-0.068}$	$0.830^{+0.067}_{-0.073}$	$0.753^{+0.116}_{-0.116}$

7.4.2 $Z(\nu\nu)$+jet background

As the transfer function is derived from $Z(\rightarrow \nu\nu)$+jets SHERPA MC sample divided by γ+jets SHERPA MC sample, each of the experimental related systematics described previously in Section 7.3 is propagated to the transfer functions. As the effects of the systematics tend to cancel out in the dividing process, the systematics uncertainty due to the detector is small. Reweighting is performed on γ+jet data, with other backgrounds (top, W+jets, diboson) subtracted from MC. In this case, the experimental systematic variations is applied to the MC samples used in the subtraction step as well. This means for each of the experimental systematics, the uncertainty in $Z(\rightarrow \nu\nu)$+jets is anti-correlated from the rest of the background processes.

To investigate the consistency of the transfer function for different selections, the ratio of $Z(\rightarrow \nu\nu)$+jets over γ+jet $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is constructed for each selection stage (c.f Table 6.5 and Table 6.6) starting from the leading large-R jet p_T cut. Three additional selection stages involving the number of b-tagged track jets are also added. Various ratio distributions are compared in Figure 7.7. The nominal transfer function is a fit from the ratio template constructed at 2-track jet selection. The systematic uncertainty is assigned as the variation when changing the event selection. In each $E_{\mathrm{T}}^{\mathrm{miss}}$ bin, the largest (smallest) ratio from various selection stages and its corresponding statistical uncertainty are chosen to construct the upper bound (lower bound). The upper (lower) range of transfer function is obtained by fitting the upper bound (lower bound) distribution. The resulting upper and lower bands are indicated by the green band in Figure 7.7. The uncertainty of the transfer function determined from this procedure is much larger than the statistical uncertainty of the fit parameters from the nominal fitting. Hence, it is taken as an estimate of the uncertainty for $Z(\rightarrow \nu\nu)$+jets estimation.

[1] Reference [172] shows that there is a small difference depending on top or anti-top is used to make the correction. Therefore the middle way where the average p_{T} of the top and anti-top quarks is used.

[2] Ideally the best variable to correct should be the top momentum after the parton shower. However, PYTHIA does not store the generator information after parton shower. Therefore we have to use the top momentum before

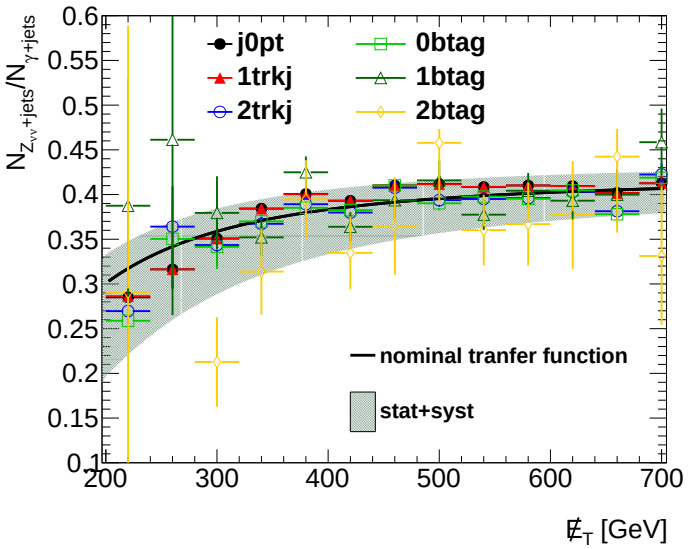


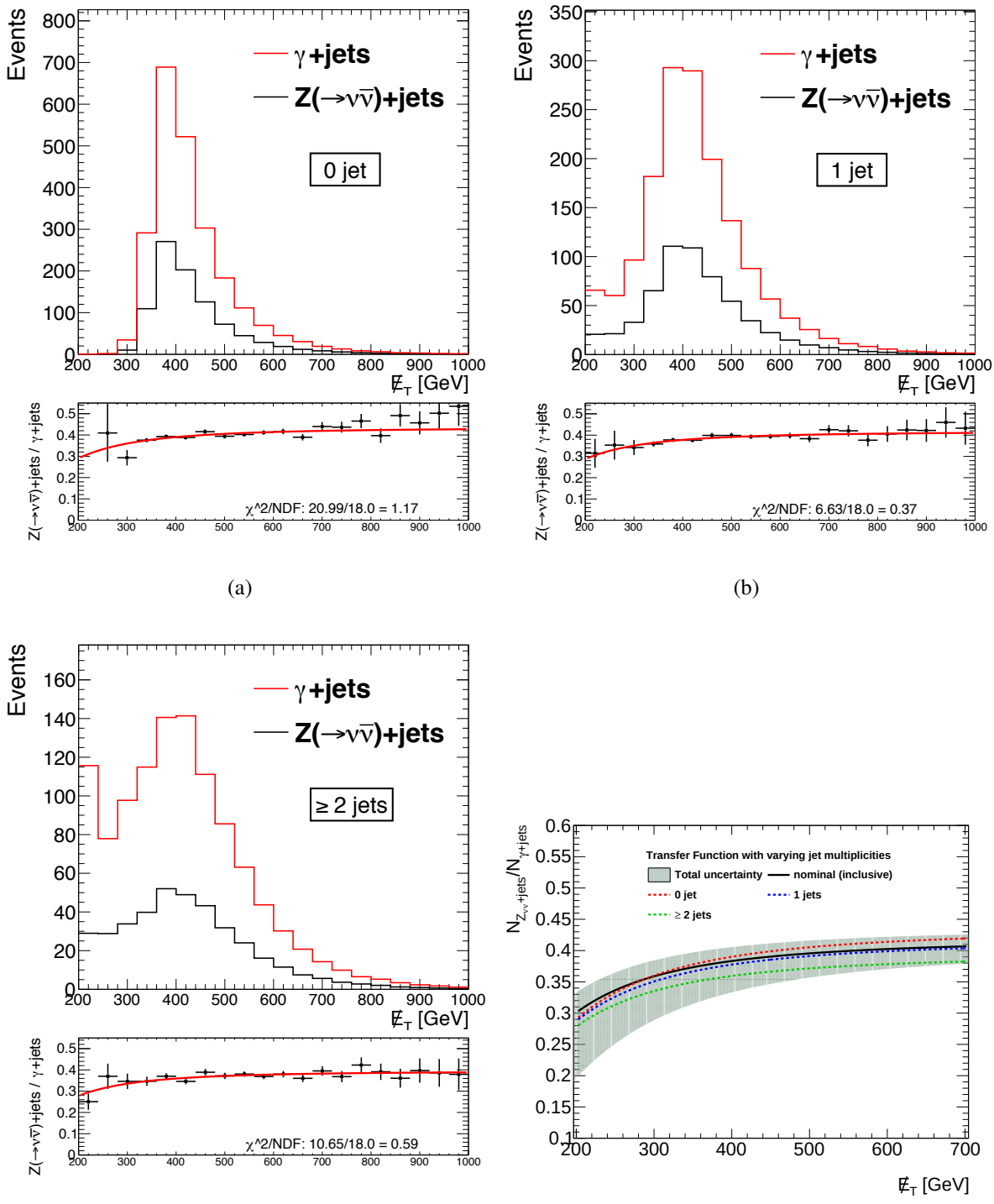
Fig. 7.7 Transfer function $N_{Z_{\nu\bar{\nu}}+jets}/N_{\gamma+jets}$ fitted with Equation 6.5. The green band is the one sigma variation of the transfer function determined from the maximum variation of the ratio at different selection stage. Different markers correspond to different event selection stage (see Table 5.1).

One thing worth noticing is that the transfer function and its systematic uncertainty are derived for inclusive jet multiplicity. Reference [157] pointed out that the net effect of higher jet multiplicity (> 1) is an overall downward shift in the Z/γ ratio. Bear in mind that although we do not explicitly reject events with higher jet multiplicity, the fact that we are selecting a back-to-back event topology means that most of the events in the SR will consist of 1 large-R jet events (such that the $E_{\mathrm{T}}^{\mathrm{miss}}$ is balanced by the single large-R jet). In other words, the systematic uncertainty associated with the transfer function as derived above should in principle include the uncertainty due to higher jet multiplicity. A check is performed to validate this claim as follows.

First the $Z(\rightarrow \nu\nu)$+jets and γ+jets templates are selected using the same criteria as listed in Table 6.6 and Table 6.5, respectively. However, each template is further divided in to different categories with different jet multiplicities, i.e 0 jet, 1 jet and ≥ 2 jets. Here the extra jets are defined as those jets that do not overlap with the leading large-R jet, or, in other words, jets with $\Delta R(large-R\ jet, jet) > 1.0$. These templates and the transfer functions are shown in Figure 7.8a-7.8c. An obvious difference exists in the lower tail between the template of each jet multiplicity. As anticipated, for template with 0 extra jet the $E_{\mathrm{T}}^{\mathrm{miss}}$ template has a peak around 350 GeV to balance the large-R jet p_{T}. Going to higher extra jet multiplicities (1 or ≥ 2 jets), the $E_{\mathrm{T}}^{\mathrm{miss}}$ is diluted as the extra jet share the balancing of the large-R jet p_{T}. Having obtained the transfer functions for each category, we then compared them with the estimated systematics uncertainty. This is illustrated in Figure 7.8d. The nominal transfer function which is derived

the parton shower. Studies performed in Reference [172] has shown that the effect to reweight the events based on top p_{T} before or after the parton shower by using different shower generator is within 2%.

at inclusive jet multiplicity is also shown. As expected, the transfer function shifted down as jet multiplicity increases. However, the variation is still fully covered by the uncertainty band. Hence, we concluded that no additional uncertainty is necessary to be assigned for the variation induce by different jet multiplicity in the event.



(a) (b)

(c) (d)

Fig. 7.8 $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of $Z(\rightarrow \nu\nu)$+jets and γ+jets for different jet multiplicities: (a) 0 jet, (b) 1 jet and (c) at least 2 jets. These extra jet is defined as jets that gives $\Delta R(large-R\ jet, jet) > 1.0$. In (d) the transfer function for different jet multiplicities are compared. The nominal transfer function derived from inclusive jets events are also shown.

7.4.3 Multi-jet background

As shown in Section 6.4, the predicted yield is larger than the observed yield by 23%-31%. Hence systematic uncertainty of 31% is assigned to account for the pathological bias of the ABCD method. This is the largest contribution to the QCD multijet background uncertainty. The other detector related uncertainties are negligible in comparison. Notice that because the estimated number of QCD multijet event in the SR is essentially zero, the uncertainty on this background is negligible. Thus it will not be included in the summary table.

7.5 Summary of the systematic uncertainty in the SR

We summarise the influence of each systematic uncertainty source on the yield of one representative signal sample and each background in the SR in Table 7.4. By all means the discussion in the following paragraphs is equally applicable to the SR with $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV. Only systematic uncertainty sources that have more than 1% effect on the total background yield in the SRs are shown. For a more comprehensive breakdown of the contribution of each systematics uncertainty source for all signals and backgrounds in the SRs, refer to Appendix F.

At first glance, the dominant sources of experimental systematic uncertainties on the background (approximately the same for signal) are due to the b-tagging, JES and JER for large-R jets, pile-up rescaling as well as the uncertainty due to $Z(\to \nu\nu)$ transfer function. The rest of the experimental systematic uncertainties yield minimal impact. All the experimental uncertainties are treated as fully correlated for the diboson, $t\bar{t}$, W+jets and VH backgrounds and anti-correlated for $Z(\to \nu\nu)$+jets (as the other MC backgrounds are subtracted from the γ+jets data). While the relative uncertainty of each individual background (i.e diboson, $t\bar{t}$, W+jets and VH) is large[1], the effect on the total background yield in the SRs are very small as each background only constitutes less than around 10% of the total background. Conversely, the cancelation of most systematic uncertainties in the γ+jets to $Z(\to \nu\nu)$+jets transfer function plus the high purity of γ+jets data template lead to minimal susceptibility of $Z(\to \nu\nu)$+jets estimate to the systematic variations. However, since $Z(\to \nu\nu)$+jets background constitutes a large proportion (>60%) of the total background, its contribution to the uncertainty on the total background is the largest. The cross-section and the PDF uncertainties are treated as uncorrelated among the background processes. To sum up the various contributions, the uncorrelated uncertainties from each background process are added in quadrature. Large-R JES (Large-R JER) and JES (JER) are considered uncorrelated. Overall, the total systematic uncertainties on the total background (signal) yield in the SR is around 14% (20%).

[1] Validation for the relatively huge JMS (JES) uncertainty for the large-R jets for diboson (W+jets) is given in Appendix F.3.

Table 7.4 Percent impact of the systematics uncertainties on the signal yield (mx1000_xdxhDh) and background yield in the signal region with $E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$.

Systematic source	Variation	$\Delta_i = \Delta N_i / N_i$						$\Delta^{total} = \frac{\Delta N^{total}}{N^{total}} = \sum_i \frac{\Delta N_i}{N^{total}}$
		mx1000_xdxhDh	Diboson	W+jets	$t\bar{t}$	SM Higgs	Z$\nu\nu$	Total Bkg.
b-tagging	up	12.5	22.4	27.4	9.0	11.2	1.9	6.5
	down	-11.8	10.3	10.4	4.9	10.0	0.5	3.5
JES of the large-R jets	up	2.5	13.2	36.1	6.8	11.7	-0.4	8.2
	down	-2.1	-9.9	-6.1	-10.7	-11.4	1.1	-1.9
JER of the large-R jets		0.1	2.8	14.4	-2.8	-2.4	1.0	4.3
JMR of the large-R jets		-0.7	5.1	8.6	-3.9	-0.8	-0.2	-1.0
JER		-0.5	1.3	1.2	-0.5	-1.1	0.1	-1.7
Pile-up rescaling	up	-0.4	1.7	11.8	-1.6	2.4	-4.4	-1.6
	down	-0.2	1.8	4.9	-0.6	1.6	-4.7	-2.4
Photon ID		-	-	-	-	-	4.0	2.5
top p_{T} correction	up	-	-	-	15.5	-	-	1.1
	down	-	-	-	-15.5	-	-	-1.1
$Z \to \nu\nu$ transfer function	up	-	-	-	-	-	5.1	3.2
	down	-	-	-	-	-	-12.0	-7.5
Cross section		10.0	26.0	20.0	7.0	3.1	11.0	7.4
PDF		10.9	5.9	5.0	6.0	2.5	-	1.8
Total Sys., $\sqrt{\Delta_i^2}$	up	20.0	53.8	54.4	27.1	18.9	13.9	14.4
	down	18.9	48.7	29.9	24.5	17.1	17.7	12.8

Chapter 8

Results

To avoid any experimentalist bias we perform a blind analysis by masking the data in the SRs. Up to this stage all the selection criteria and the whole analysis chain have been fixed. As we have seen in Section 6, the good agreement between the data and the expectation is obtained in various CRs and VRs. Furthermore the possible systematic uncertainties have also been estimated. In this section the events in the SR are unblinded to get the final result. The numbers of observed data and estimated background events in the SR are presented in Section 8.1. The kinematic distributions in the SR will also be shown. The statistical interpretation of the results is discussed from Section 8.2 to Section 8.2.2.

8.1 Event yield in the SRs

Table 8.1 shows the summary of the number of observed events in data together with the predicted yield for individual background and their associated total uncertainties in the SRs. For SR with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV ($E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV), 20 (9) events are observed in the data compared to the total background expectation of 11 ± 2.3 (7.7 ± 1.7) events. Some kinematic distributions in the SRs are illustrated in Figures 8.1-8.2. Two representative signal models are superimposed in the plots. Their production cross sections are scaled to 10 fb to aid in comparison.

From the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution we see that the excess exist in the lower $E_{\mathrm{T}}^{\mathrm{miss}}$ region closer to the cut value where most of the backgrounds are. Furthermore, if the excess does come from the signal, the excess should be observed as a peak structure in leading large-R jet mass distribution. The fact that it is not means the excess events are not signal-like. Further statistical quantification of the results are discussed in the following sections.

Upon a closer examination of the $E_{\mathrm{T}}^{\mathrm{miss}}$ or large-R jet p_{T} distribution in Figure 8.1a or Figure 8.1b, we find that there is a sharp cutoff of $Z(\rightarrow \nu\nu)$+jets estimate at $E_{\mathrm{T}}^{\mathrm{miss}}(p_{\mathrm{T}}) > 550$ GeV. This cutoff is due to the lack of γ+jets events in data with $E_{\mathrm{T}}^{\mathrm{miss}}(p_{\mathrm{T}}) > 550$ GeV, as shown in Figure 8.3. One may argue that this lack of γ+jets events in data contradicts the motivation

of using γ+jets to estimate $Z(\to \nu\nu)$+jets background in the first place. Under such condition, the statistical uncertainty of this data-driven estimation is large, causing this estimation method to be less reliable than we previously claim. In fact, the estimation $Z(\to \nu\nu)$+jets background completely based on MC does not suffer from this problem, as can be seen in Figure 8.4. In comparison to the data-driven estimate (Figure 8.1a), the agreement between the data and MC is slightly better. Moreover, there is no cutoff at $E_{\mathrm{T}}^{\mathrm{miss}} > 550$ GeV. In addition, by comparing the total $Z(\to \nu\nu)$+jets background yield based on data-driven and MC in the SR with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV, which is 7.0±1.6 (stat.) and 7.25±0.42 (stat.) respectively, we see that the statistical uncertainty of the latter method is smaller.

However, one must not forget that the trade off of background estimation fully based on MC is the large systematic uncertainties. This fact is evident as shown in Table 8.2. All the uncertainties for MC based estimate are larger than the one based on data-driven method, in particular the uncertainty on b-tagging efficiency. The total systematic uncertainty on the background yield in the SR based on MC is 3.5 times larger than the data-driven based estimate. Base on this fact, we have concluded that the choice of using data-driven $Z(\to \nu\nu)$+jets estimation is appropriate.

Table 8.1 Event yield of each background and the observed event in data for the final SRs. The uncertainties include both the statistical and the systematic errors.

	$E_{\mathrm{T}}^{\mathrm{miss}}$ [GeV]	
Background	> 300	> 400
$Z(\to \nu\nu)$+jets	7.00 ± 2.0	5.2 ± 1.6
QCD multi-jet	$< 0.0 \pm 0.04$	$< 0.0 \pm 0.04$
Top	0.8 ± 0.5	0.6 ± 0.4
W($\ell\nu$)+jets	1.4 ± 0.7	0.8 ± 0.4
Diboson	0.9 ± 0.5	0.6 ± 0.3
(W/Z)H	1.0 ± 0.2	0.6 ± 0.1
Total bkg.	11.2 ± 2.3	7.7 ± 1.7
Data	20	9

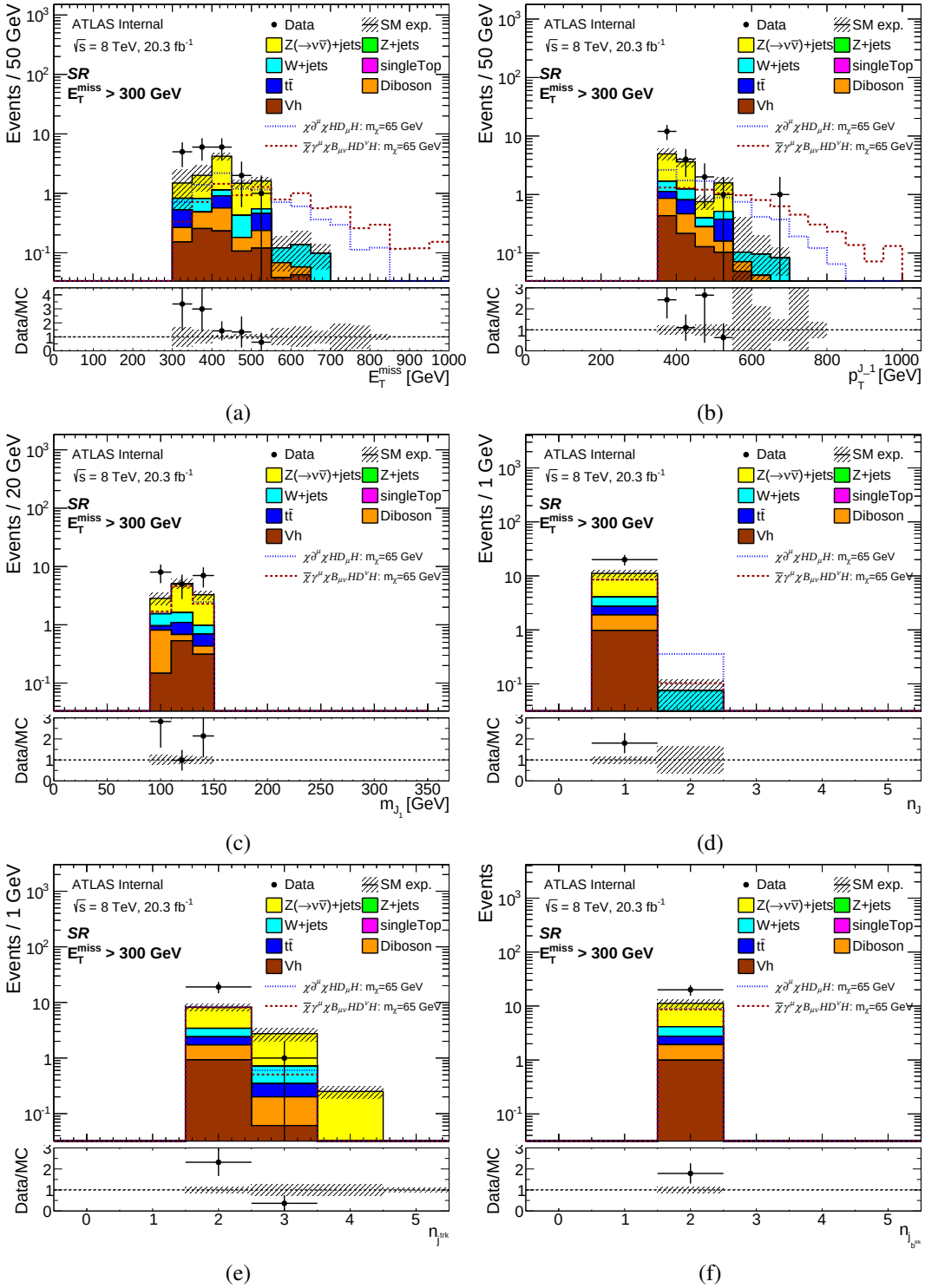


Fig. 8.1 Kinematic distributions in the signal region with $E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$. Two superimposed signal samples' cross sections are scaled to 10 fb for comparison purpose. (a) $E_{\mathrm{T}}^{\mathrm{miss}}$ (b) leading large-R jet p_T, (c) leading large-R jet mass, (d) large-R jet multiplicity, (e) the number of associated track jet and (f) the number of associated b-tagged track jet.

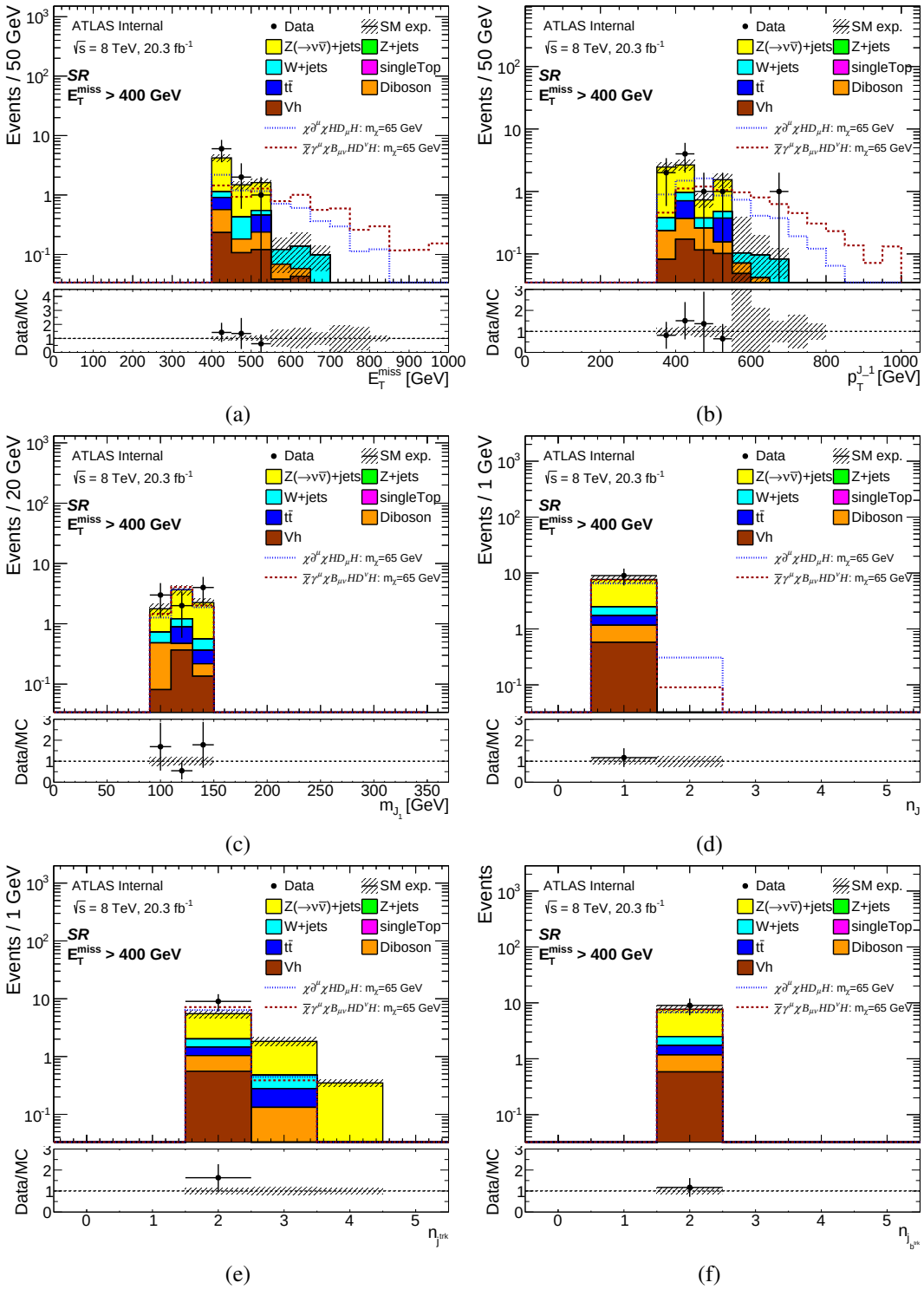


Fig. 8.2 Kinematic distributions in the signal region with $E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$. Two superimposed signal samples' cross sections are scaled to 10 fb for comparison purpose. (a) $E_{\mathrm{T}}^{\mathrm{miss}}$ (b) leading large-R jet p_T, (c) leading large-R jet mass, (d) large-R jet multiplicity, (e) the number of associated track jet and (f) the number of associated b-tagged track jet.

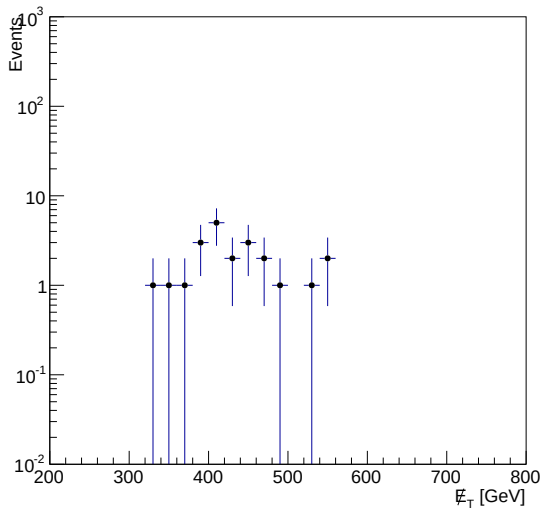


Fig. 8.3 γ+jets events in the SR selected from data.

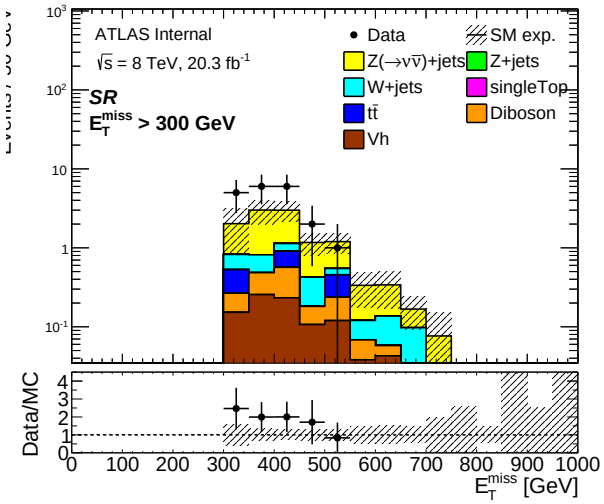


Fig. 8.4 $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the SR. The $Z(\to \nu\nu)$+jets background events are estimated fully based on MC simulation. The hash band include both the statistical and systematics uncertainties.

8.2 Statistical interpretation

With the acquired set of predictions, observations, and systematic variations in the SRs, the next logical step is to perform hypothesis testing. Frequentist approach is used in all of the hypothesis tests discussed below. For our purpose of searching for a new signal process (DM), we define two hypotheses. First is the null hypothesis (also referred to as the background-only

Table 8.2 The comparison of major systematic uncertainties (%) on the yield of $Z(\rightarrow \nu\nu)$+jets background in the SR with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV for data-driven based (left) and MC based (right) method.

	$Z(\rightarrow \nu\nu) + jets^{data}$	$Z(\rightarrow \nu\nu) + jets^{MC}$
b-taging	1.9	44.9
large-R jet JES	-0.4	14.7
large-R jet JER	1.0	2.8
large-R jet JMS	0.6	-3.3
large-R jet JMR	-0.2	1.0
Photon energy scale	-0.4	-
Photon energy resolution	-0.4	-
Photon ID	4.0	-
$Z(\rightarrow \nu\bar{\nu})$+jets transfer function	5.1	-
Total systematic uncert.	13.9	48.8

hypothesis), H_0, which assumes that only SM processes contribute to the measurements, here designated as backgrounds. Next we have the alternative hypothesis (also referred to as the signal-plus-background hypothesis), H_1, which includes both backgrounds as well as the signal. In the following sub-sections, depending on the statement we wish to test, the role of null and alternative hypotheses may get switched. The parameter of interest used in these hypothesis tests is the signal strength parameter, μ. It is defined as the ratio of the expected cross-section according to the assumed theory and the actual observed cross-section. $\mu = 0$ corresponds to the background only hypothesis, while $\mu = 1$ reproduces the assumed signal hypothesis.

Having defined the hypotheses, the next step is to define a test statistic designed to discriminate signal-like events from backgrounds. The most common test statistic that is used in the LHC is the profile likelihood ratio. More specifically, we employ a binned likelihood function, $L(\mu, \theta_i)$ parameterised by the signal strength parameter, μ and a set of nuisance parameters (such as the shape of the distributions, the total background rate, reconstruction efficiencies, systematic uncertainties and so on), collectively denoted as $\theta_j (j = 1, ...K)$. Given a data outcome, the likelihood function can be constructed as the product of Poisson probability terms associated to the i^{th} bin of the histogram[1]. Since we are performing a single bin counting analysis, L is expressed as:

$$L(data, \theta^0 \mid \mu, s, b, \theta) = \frac{(\mu s + b)^n e^{-(\mu s + b)}}{n!} \times \prod_{j \in syst} G(\theta_j^0, \theta_j) \tag{8.1}$$

where n is the number of events in the SR, s and b are and the number of expected signal and background events in the same SR, respectively. Systematic uncertainties are described in

[1] We consider each bin content, n_i as a random variable that are independent of each other and have poissonian distributions. The total number of events is not fixed and fluctuates between one experiment and another.

terms of the nuisance parameters, as encoded in the second Gaussian term, $G(\theta_j^0, \theta_j)$ where θ_j^0 are the central values of the auxiliary measurements around which θ_j can be varied.

With the definition of the likelihood, we can write down the test statistic as [173]

$$\tilde{q}_\mu = \begin{cases} -2\ln\left(\frac{L(\mu,\,\hat{\hat{\theta}}(\mu))}{L(0,\,\hat{\hat{\theta}}(0))}\right) & \hat{\mu} < 0 \\ -2\ln\left(\frac{L(\mu,\,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\,\hat{\theta})}\right) & \hat{\mu} \geq 0 \end{cases} \tag{8.2}$$

where $\hat{\mu}$ and $\hat{\theta}$ are the maximum likelihood estimator of both the μ and θ while $\hat{\hat{\theta}}(\mu)$ is the maximum likelihood estimator of θ when μ is fixed at a particular value. $L(\hat{\mu}, \hat{\theta})$ is the maximised (unconditional) likelihood function whereas $L(\mu, \hat{\hat{\theta}}(\mu))$ is the maximised (conditional) likelihood function for a particular fixed value of μ. $L(0, \hat{\hat{\theta}}(0))$ is the conditional likelihood for the value of $\mu = 0$. The constraint $\hat{\mu} \geq 0$ is dictated by the assumption that the presence of a new signal can only increase the event rate beyond what is expected from background alone (i.e $\mu \geq 0$). Under such assumption, if any dataset gives a result with $\hat{\mu} < 0$, then the best level of agreement between the data and any physical value of μ occurs for $\mu = 0$ (hence the $L(0, \hat{\hat{\theta}}(0))$ in the denominator for $\hat{\mu} < 0$).

In order to perform hypothesis tests, we need the sampling distribution for the test statistic. Two approaches are considered: the toy MC method and the asymptotic approximation[1]. In the case of the toy MC generation, the distribution can be obtained by generating pseudo experiments that randomise the number of observed events and the central values of the auxiliary measurements. Ideally one should scan μ and nuisance parameters to generate a sufficient number of pseudo experiments for each set of auxiliary measurement (as the true values of the auxiliary measurement are unknown). However, this procedure becomes impractical when there are too many auxiliary measurements (in our case, 15 in total) to consider. Instead a technique called profile maximisation in which the nuisance parameters are “profiled" on the observed data. We first obtained the maximum estimator of the nuisance parameters by maximum likelihood fitting based on the observed data and the hypothesised value of μ. Their values are then used to set the true value of the auxiliary measurements. By doing this, pseudo experiments that are expected to maximise the p-value over the auxiliary measurements can be generated.

The asymptotic approximation works instead by approximating Equation 8.2 numerically [2]. By using the Wilks [174] and Wald [175] theorems, it has been shown in Reference [173] that in the large sample limit the likelihood ratio, $\tilde{q}_\mu$ is independent of the values of the nuisance parameters such that it is asymptotically distributed following a χ_s^2 probability distribution

[1] Whenever practical, the toy MC is used but the test result is confirmed using the asymptotic calculation and vice versa.

[2] Besides the possibility to eliminate the nuisance parameters by profiling the pdf, the key reason for choosing the test statistics based on the profile likelihood ratio is that it can be asymptotically approximated.

functions (pdf). The degrees of freedom s are equal to the difference between the number of maximisation parameters at denominator and numerator. The asymptotic approximation holds reasonably well for cases with O(10) events [176].

8.2.1 Discovery hypothesis test

Given that we observe a small excess in the SR with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV, we need to evaluate statistically how significant the observed events in the SRs deviates from the SM background only hypothesis. The level of agreement is quantified by computing the background only p-value, which is defined as

$$p_0 = \int_{q_0^{obs}}^{\infty} f(q_0) dq_0 \tag{8.3}$$

where q_0^{obs} is the value of q_0 obtained using the observed data and $f(q_0)$ is the pdf of the test statistics assuming $\mu = 0$. The test statistic used is a one-sided profile likelihood ratio [177, 173]:

$$q_0 = \begin{cases} -2\ln\left(\frac{L(0,\hat{\hat{\theta}}_0)}{L(\hat{\mu},\hat{\theta})}\right) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{8.4}$$

In order to obtain p_0, 3000 pseudo-experiments are generated with the background only model. The distribution of the test statistic, q_0, is built up from the values of the generated pseudo-experiments. For our SRs with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV, the corresponding p_0 values is 0.03 and 0.37, consistent with the null hypothesis.

8.2.2 Exclusion limit

Since no statistically significant deviation from the background model hypothesis is observed, we proceed to set the exclusion limit. Two exclusion limits are calculated: the cross section limits for DM production and the model independent limit (or the BSM visible cross section limit). An upper limit is determined by performing inversion of the hypothesis test. The general idea is that we repeatedly run hypothesis tests with varying values of the signal strength until the corresponding p-value (of the background only or signal+background hypothesis) drops below a certain threshold (in our case 0.05, for the typical 95% confidence level).

To calculate the p-value for limit setting we follow the "CL_s" methods, in which a signal model is regarded as excluded if one finds [177, 173]:

$$CL_s \equiv \frac{p_{s+b}}{1 - p_b} < 0.05 \tag{8.5}$$

where p_{s+b} is defined as the probability of the signal+background model to produce data with a value of q_μ equal or lesser compatibility than that observed, i.e,

$$p_{s+b} = P_{s+b}(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs} \mid s+b) = \int_{\tilde{q}_\mu^{obs}}^{\inf} f(\tilde{q}_\mu \mid s+b) d\tilde{q}_\mu \tag{8.6}$$

whereas p_b is defined as the probability of the background only model to produce data with the same or more compatibility with the signal+background model as that observed, i.e,

$$p_b = P_b(\tilde{q}_\mu \leq \tilde{q}_\mu^{obs} \mid b) = \int_{-\inf}^{\tilde{q}_\mu^{obs}} f(\tilde{q}_\mu \mid b) d\tilde{q}_\mu \tag{8.7}$$

The test statistic used to determine p_{s+b} and p_b is also a one-sided profile likelihood ratio defined as [177, 173]:

$$\tilde{q}_\mu = \begin{cases} -2\ln\left(\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(0,\hat{\hat{\theta}}(0))}\right) & \hat{\mu} < 0 \\ -2\ln\left(\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\hat{\theta})}\right) & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{8.8}$$

The $\hat{\mu}$ is constrained to be greater than or equal to zero as explained before. The constraint $\hat{\mu} \leq \mu$ implies that we do not consider the upward fluctuations of the data ($\hat{\mu} > \mu$) as the evidence against the signal+background hypothesis. It is imposed to ensure a one-sided confidence interval.

Limits for DM production

The first type of limit we want to compute is the model-specific cross sections upper limit. It must be calculated for each signal sample considered in this study. For a 95% CL upper limit the value of the signal strength needs to be evaluated for which the CL_s value falls below 5%. This is done for the data point in the case of the observed limit, for the median value of the background-only distribution in the case of the expected limit and for the points which contain 68% or 95% of the background-only distribution in the case of the ± 1 and $\pm 2\sigma$ bands on the expected limits.

The obtained upper limit on the signal strength, μ_{lim} can then easily be transformed into an upper limit on the excluded cross section of the signal model via the relation

$$\sigma_{lim}(95\%\ C.L.) = \mu_{lim}\sigma_0 \tag{8.9}$$

where σ_0 is the production cross section for the generated signal samples based on some specific model. Table 8.3 and Table 8.4 show the expected and observed 95% CL cross-section upper

Table 8.3 Expected cross section limits for EFT models in the SRs with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 400$GeV. The expected upper limit is the upper limit one would theoretically observe assuming the hypothetical cross section (hence the number of signal events) is true. The smallest cross section limits for each model are highlighted in red.

Model	m_χ [GeV]	$E_{\mathrm{T}}^{\mathrm{miss}} > 300$GeV σ_{lim} [fb]	$E_{\mathrm{T}}^{\mathrm{miss}} > 400$GeV σ_{lim} [fb]
xxhh	1	$19.7^{+9.8}_{-5.3}$	$25.3^{+9.1}_{-5.8}$
xxhh	65	$17.2^{+7.5}_{-4.3}$	$18.7^{+9.7}_{-4.5}$
xxhh	100	$12.2^{+5.4}_{-2.9}$	$13.1^{+5.6}_{-3.5}$
xxhh	500	$7.1^{+2.2}_{-2.2}$	$5.9^{+3.2}_{-1.1}$
xxhh	1000	$7.2^{+2.3}_{-2.3}$	$5.5^{+3.5}_{-0.7}$
xxhhg5	1	$18.1^{+9.4}_{-4.6}$	$19.6^{+9.4}_{-5.2}$
xxhhg5	65	$14.9^{+8.8}_{-3.1}$	$17.1^{+9.9}_{-3.7}$
xxhhg5	100	$10.0^{+5.0}_{-1.3}$	$10.8^{+4.7}_{-1.8}$
xxhhg5	500	$6.7^{+2.7}_{-1.8}$	$5.9^{+3.2}_{-1.1}$
xxhhg5	1000	$6.9^{+2.4}_{-2.0}$	$6.1^{+2.8}_{-1.3}$
xdxhDh	1	$63.3^{+34.1}_{-15.6}$	$74.6^{+43.4}_{-21.4}$
xdxhDh	65	$14.5^{+7.9}_{-4.5}$	$17.0^{+7.0}_{-3.0}$
xdxhDh	100	$10.6^{+5.4}_{-1.4}$	$11.9^{+6.2}_{-2.6}$
xdxhDh	500	$8.3^{+1.5}_{-3.3}$	$7.8^{+2.0}_{-2.8}$
xdxhDh	1000	$6.0^{+3.0}_{-1.2}$	$5.0^{+3.8}_{-0.2}$
xgxFhDh	1	$3.7^{+1.5}_{-1.1}$	$3.5^{+1.8}_{-1.0}$
xgxFhDh	65	$3.7^{+1.3}_{-1.1}$	$3.5^{+1.7}_{-0.7}$
xgxFhDh	100	$3.9^{+1.8}_{-1.4}$	$3.5^{+1.8}_{-0.9}$
xgxFhDh	500	$3.4^{+1.5}_{-1.1}$	$3.2^{+1.6}_{-0.7}$
xgxFhDh	1000	$3.2^{+1.6}_{-1.0}$	$2.9^{+1.6}_{-0.8}$

limits, respectively, for each of the EFT model in both the SRs. The $E_{\mathrm{T}}^{\mathrm{miss}}$ cut with the best expected medium limit in each parameter space of the models is highlighted in red.

Table 8.5 summarises the best 95% CL model dependent cross-section limits for each EFT operator and the corresponding $E_{\mathrm{T}}^{\mathrm{miss}}$ value. These limits, and their theoretical interpretation, are discussed in Section 8.3.

Model independent upper limit

The second limit we compute is the limit on the BSM visible cross section, σ_{vis}, defined as

$$\sigma_{vis} = \sigma_{BSM} \times \varepsilon \tag{8.10}$$

Table 8.4 Observed cross-section limits for EFT operator and for each DM mass point in the SRs with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 400$GeV. The results highlighted in red are the values associated to the best expected limit in Table 8.3.

Model	m_χ [GeV]	$E_{\mathrm{T}}^{\mathrm{miss}} > 300$GeV σ_{lim} [fb]	$E_{\mathrm{T}}^{\mathrm{miss}} > 400$GeV σ_{lim} [fb]
xxhh	1	44.8	30.3
xxhh	65	37.1	26.4
xxhh	100	25.1	17.2
xxhh	500	13.6	8.6
xxhh	1000	13.7	8.5
xxhhg5	1	39.2	25.5
xxhhg5	65	34.6	24.0
xxhhg5	100	22.1	14.5
xxhhg5	500	13.5	8.6
xxhhg5	1000	13.6	8.4
xdxhDh	1	144.3	111.2
xdxhDh	65	32.3	21.1
xdxhDh	100	25.0	15.0
xdxhDh	500	14.8	9.4
xdxhDh	1000	12.5	8.2
xgxFhDh	1	8.1	4.8
xgxFhDh	65	8.2	4.7
xgxFhDh	100	8.2	4.6
xgxFhDh	500	7.6	4.3
xgxFhDh	1000	7.4	4.1

Table 8.5 The best expected cross-section upper limits and their corresponding observed cross-section limit as well as sliding $E_{\mathrm{T}}^{\mathrm{miss}}$ signal region for each EFT operator and for each DM mass point.

			$\sigma_{BSM}^{95\%}$ [fb]	
Model	m_χ[GeV]	$E_{\mathrm{T}}^{\mathrm{miss}}$ cut [GeV]	Expected	Observed
xxhh	1	> 300	19.7	44.8
xxhh	65	> 300	17.2	37.1
xxhh	100	> 300	12.2	25.1
xxhh	500	> 400	5.9	8.6
xxhh	1000	> 400	5.5	8.5
xxhhg5	1	> 300	18.1	39.2
xxhhg5	65	> 300	14.9	34.6
xxhhg5	100	> 300	10.0	22.1
xxhhg5	500	> 400	5.9	8.6
xxhhg5	1000	> 400	6.1	8.4
xdxhDh	1	> 300	63.3	144.3
xdxhDh	65	> 300	14.5	32.3
xdxhDh	100	> 300	10.6	25.0
xdxhDh	500	> 400	7.8	9.4
xdxhDh	1000	> 400	5.0	8.2
xgxFhDh	1	> 400	3.5	4.8
xgxFhDh	65	> 400	3.5	4.7
xgxFhDh	100	> 400	3.5	4.6
xgxFhDh	500	> 400	3.2	4.3
xgxFhDh	1000	> 400	2.9	4.1

Table 8.6 Visible (model-independent) cross-section upper limits. Left to right: SR sliding $E_{\mathrm{T}}^{\mathrm{miss}}$ cut, number of observed events, number of expected background events, 95% CL upper limits on the visible cross section ($\langle\varepsilon\sigma\rangle_{\mathrm{obs}}^{95}$) and the number of non-SM events ($N_{BSM}{}_{\mathrm{obs}}^{95}$). The sixth column($N_{BSM}{}_{\mathrm{exp}}^{95}$) shows the 95% CL upper limit on the number of non-SM events, given the expected number (and $\pm1\sigma$ excursions on the expectation) of background events. The last column shows the p-value for background-only hypothesis ($p(s=0)$).

$E_{\mathrm{T}}^{\mathrm{miss}}$ **cut**	N_{obs}	N_{bkgd}	$\langle\varepsilon\sigma\rangle_{\mathrm{obs}}^{95}$[fb]	$N_{BSM}{}_{\mathrm{obs}}^{95}$	$N_{BSM}{}_{\mathrm{exp}}^{95}$	$p(s=0)$
>300 GeV	20	11.2	0.90	18.4	$10.1^{+4.1}_{-3.2}$	0.03
>400 GeV	9	7.7	0.45	9.1	$7.8^{+3.4}_{-2.3}$	0.37

where σ_{BSM} is the production cross section of some new physics, and ε is the selection efficiency (including reconstruction efficiency and the detector acceptance). There is no signal model considered here, so this result only depends on the SM background modelling.

For each $E_{\mathrm{T}}^{\mathrm{miss}}$ cut in the final signal selection, the expected background, including its statistical and systematic uncertainties, is fit to the number of observed events.The number of signal events in the signal region is added as a parameter to the fit. The 95% CL upper limits on the visible cross-section, the number of observed and expected non-SM events, as well as the p-value for background-only hypothesis are given in Table 8.6.

8.3 Signal interpretations for EFT's models

For each of our MC signal sample with the coupling constant, λ_0 (or the mass scale, $\Lambda 0$) and the cross section, σ_0, the limit on the coupling constant (or mass scale) can be calculated as

$$\lambda_{lim} = \lambda_0 \left(\frac{\sigma_{lim}}{\sigma_0} \right)^{\frac{1}{n}} \tag{8.11}$$

where the power n is determined by the structure of the Lagrangian of each EFT operator. Their values are listed in Table 8.7. The 95% C.L. limits on the coupling parameter or the mass scale for the four classes of EFT operators are shown in Fig. 8.5. The limits set by the ATLAS search in the DM+Higgs($\rightarrow \gamma\gamma$) [51] are also shown. In comparison to the results obtained by DM+Higgs($\rightarrow \gamma\gamma$) analysis, our limits are a few times stronger for all four operators. Other important constraints such as invisible h or Z decays, as well as the perturbativity requirements are also included. The bound of invisible Higgs decay are highly constraining. For $m_\chi < m_h/2$, mono-Higgs signal can only be observed if LHC sensitivities can be improved by several order of magnitudes. As a comparison, the recent bounds on the spin-independent direct detection cross section from LUX [178] are mapped onto the same parameter space for both the operators `xxhh` (Figure 8.5a) and `xdxhDh` (Figure 8.5c). While the LUX bounds are highly constraining for these two operators, we must bear in mind that these direct detection constraints may be avoided if DM is inelastic[1] [179–181]. In this scenario, there exists an excited state $\chi*$ in addition to the DM particle χ, with a mass splitting $\delta = m_{\chi*} - m_\chi$. If the mass splitting is of the O(MeV) or more, elastic scattering off the nucleus, i.e $\chi N \rightarrow \chi N$, is highly suppressed, compared to the inelastic scattering, i.e $\chi N \rightarrow \chi * N$. This is because only those χ with sufficient kinetic energy to be excited into the heavier state will scatter off nuclei. For the `xxhhg5` operator (Figure 8.5b), the SI cross section are suppressed by powers of the WIMP velocity [52, 60], generically expected to be of order $\sim 10^{-3}$. This leads to a very weak direct detection constraint. For dimension-8 operator `xgxFhDh` (Figure 8.5d) we see that it is constrained neither by invisible decays nor direct detection. In the latter case, the weak constraint is due to the fact that direct detection signal arises at one-loop order, which are expected to be highly suppressed.

8.3.1 Validity of EFT model

As has been briefly discussed in Section 1.5 the EFT description is only justified when there is a clear separation between the energy scale of the process to describe and the scale of the underlying microscopic interactions. At the energy scales and coupling strengths accessible to the LHC, we must check if the EFT approach is still valid to interpret our model dependent

[1] Inelastic dark matter (iDM) was originally proposed to reconcile the DAMA annual modulation observation and null results from CDMS. It is just a simple extension of the standard WIMP model.

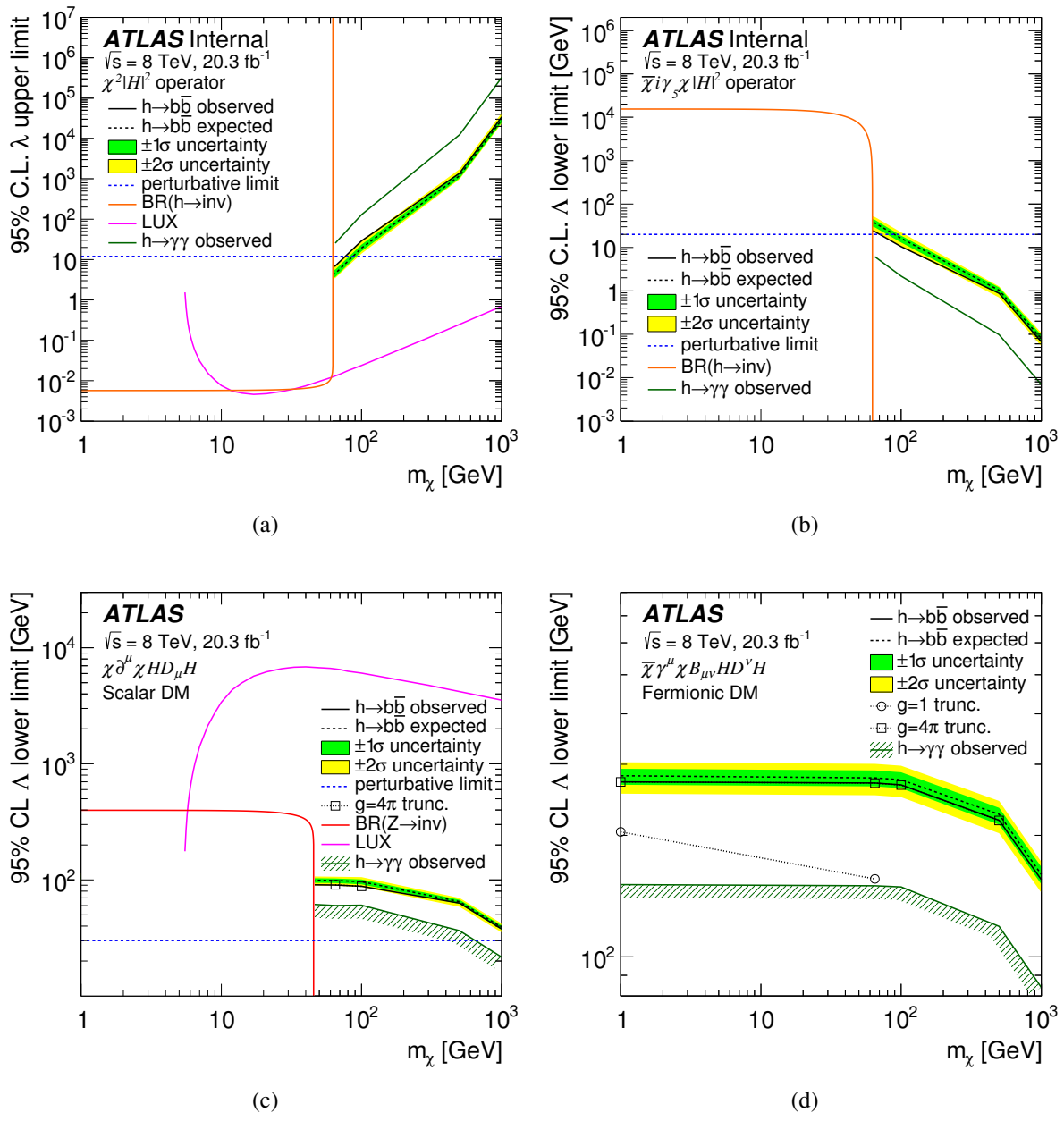


Fig. 8.5 95% CL limits on the coupling parameter/mass scale for the four EFT models considered in this study, shown as the black solid lines. Limits set by the ATLAS search in the mono-H $\gamma\gamma$ channel [51] are shown as dark green contours. Where applicable, the truncated limits (refer to Section 8.3.1) for events passing $Q_{\chi\chi} < m_V$, assuming unit and maximal couplings are shown as dotted line with markers. Magenta contours denote limits set by LUX [178] while orange (red) contours are exclusion limits from invisible h (Z) decays.

Table 8.7 The mono-Higgs cross section as a function of parameters of each operator and their domains of validity for all operators considered. Outside the valid domain, the cross section must be calculated explicitly using MadGraph. σ_0 and Λ_0 are defined in the text.

Short Name	$\sigma_{h\chi\bar{\chi}}$(parameters)	Valid Domain
xxhh	$\begin{cases} \sigma_0 \cdot \left(\frac{\lambda}{\lambda_0}\right)^2 & \lambda \lesssim 1 \\ \sigma_0 \cdot \left(\frac{\lambda}{\lambda_0}\right)^4 & \lambda \gtrsim 1 \end{cases}$	$\lambda < 4\pi \cap m_\chi < \frac{m_h}{2} \rightarrow \lambda \lesssim 0.016$
xxhhg5	$\begin{cases} \sigma_0 \cdot \left(\frac{\Lambda}{\Lambda_0}\right)^{-2} & \lambda \lesssim 30\text{GeV} \\ \sigma_0 \cdot \left(\frac{\Lambda}{\Lambda_0}\right)^{-4} & \Lambda \gtrsim 30\text{GeV} \end{cases}$	$\Lambda > \frac{v}{4\pi} \cap m_\chi < \frac{m_h}{2} \rightarrow \Lambda \gtrsim 10\text{TeV}$
xdxhDh	$\begin{cases} \sigma_0 \cdot \left(\frac{\Lambda}{\Lambda_0}\right)^{-8} & \Lambda \lesssim 100\text{GeV} \\ \sigma_0 \cdot \left(\frac{\Lambda}{\Lambda_0}\right)^{-4} & \Lambda \gtrsim 100\text{GeV} \end{cases}$	$g_{Z_{\text{eff}}} < 4\pi(\Lambda \gtrsim 30\text{GeV}) \cap m_\chi < \frac{m_Z}{2} \rightarrow \Lambda \gtrsim 400\text{GeV}$
xgxFhDh	$\sigma_0 \cdot \left(\frac{\Lambda}{\Lambda_0}\right)^{-8}$	

limits, in particular those of the higher dimensional (>6) operators, i.e the `xdxhDh` and `xgxFhDh` operators. To do this, first we need to clarify under what circumstances the EFT approach is valid. Next we introduce some quantities to assess which of our MC signal events are in the regime where EFT is valid. Those events that are not are simply removed. Then the model dependent limit is recalculated using only the events satisfying the EFT validity condition. More details are discussed in the following paragraphs.

To illustrate the range of validity of our EFT operators, let us consider a simple scenario where the DM interact with the SM via a $s-$channel exchange of a heavy mediator with mass M_{med}. The procedure of integrating out the heavy mediator and retaining the operator of lowest dimension can be viewed in terms of the expansion of the heavy particle propagator [63, 182]

$$\frac{g_1 g_2}{(Q_{tr}^2 - M_{med}^2)} = -\frac{g_1 g_2}{M_{med}^2}\left(1 + \frac{Q_{tr}^2}{M_{med}^2} + \mathcal{O}\left(\frac{Q_{tr}^4}{M_{med}^4}\right)\right) \tag{8.12}$$

where Q_{tr} is the momentum transfer of the process, g_1 and g_2 are the couplings of the propagator. In the limit $Q_{tr}^2 \ll M^2$ the higher order terms can be reasonably dropped. The suppression scale are connected with the heavy mediator via the relation $\Lambda\sqrt{g_1 g_2} = M_{med}$. The mediator must carry at least enough energy to produce a DM pair, therefore $Q_{tr} > m_{\chi\chi}$. The validity condition for the EFT then implies

$$\Lambda\sqrt{g_1 g_2} > Q_{tr} > m_{\chi\chi} \tag{8.13}$$

The details of the couplings are not known a priori. Here we investigate two possible cases [183, 184]:

1. $g_1 = g_2 = 1$ — the unit couplings, typical theory assumption of natural scale.

2. $g_1 = g_2 = 4\pi$ — best case for colliders while remaining in the perturbative regime.

To ensure that our signal events are in the valid EFT regime, the events with $m_{\chi\chi} > \Lambda\sqrt{g_1 g_2}$ are removed. We refer this step as "truncation". The acceptance obtained using the remaining events are thus called "truncation acceptance". One example distribution of the truncation acceptance for the case where $g_1 = g_2 = 1$ and for `xgxFhDh` operator are shown in Figure 8.6. We also compare two other truncation acceptances, where the events are subjected to the same signal selections (but the MC generated objects instead of the reconstructed objects are used to define the selection variables). As expected, more events failed the validity test for larger DM masses at any $\Lambda\sqrt{g_1 g_2}$ values. We also see that for large DM masses, the truncation acceptance is the same for inclusive events (only the $m_{\chi\chi} < \Lambda\sqrt{g_1 g_2}$ cuts is applied) and events which pass the signal selection. For smaller DM masses, the difference in truncation acceptances becomes slightly larger.

We recalculate the limits on Λ for the `xdxhDh` and `xgxFhDh` operators after applying the truncation with g_1 and g_2 equal to 1 or 4π. The results are summarised in Table 8.8. For all DM masses of `xdxhDh` operator, no events pass the validity requirement when $g_1 = g_2 = 1$. However when $g_1 = g_2 = 4\pi$, `xdxhDh` operator satisfies the validity requirement for two of the DM masses. While for `xgxFhDh` operator, the validity requirement for the unit couplings and the maximal perturbative couplings are satisfied for $m_\chi = 1, 65$ GeV and $m_\chi = 1, 65, 100, 500$ GeV respectively. We also note that the observed limit without truncation coincide with the truncated limit with the maximal perturbative couplings. These "truncated" limits on the suppresion scale, $\Lambda_{trunc.}$ are drawn as additional lines in the limit plots shown in Figure 8.5.

Table 8.8 The observed limits on Λ without truncation, Λ_0, and truncated limits, $\Lambda_{trunc.}$, after applying the truncation method to the `xdxhDh` and `xgxFhDh` operators.

xdxhDh			
m_χ [GeV]	Λ_0 [GeV]	$\Lambda_{trunc.}$ [GeV], $g_1 = g_2 = 1$	$\Lambda_{trunc.}$ [GeV], $g_1 = g_2 = 4\pi$
65	91	-	91
100	88	-	88
500	64	-	-
1000	38	-	-
xgxFhDh			
m_χ [GeV]	Λ_0 [GeV]	$\Lambda_{trunc.}$ [GeV], $g_1 = g_2 = 1$	$\Lambda_{trunc.}$ [GeV], $g_1 = g_2 = 4\pi$
1	274	204	274
65	271	156	271
100	268	-	268
500	220	-	219
1000	157	-	-

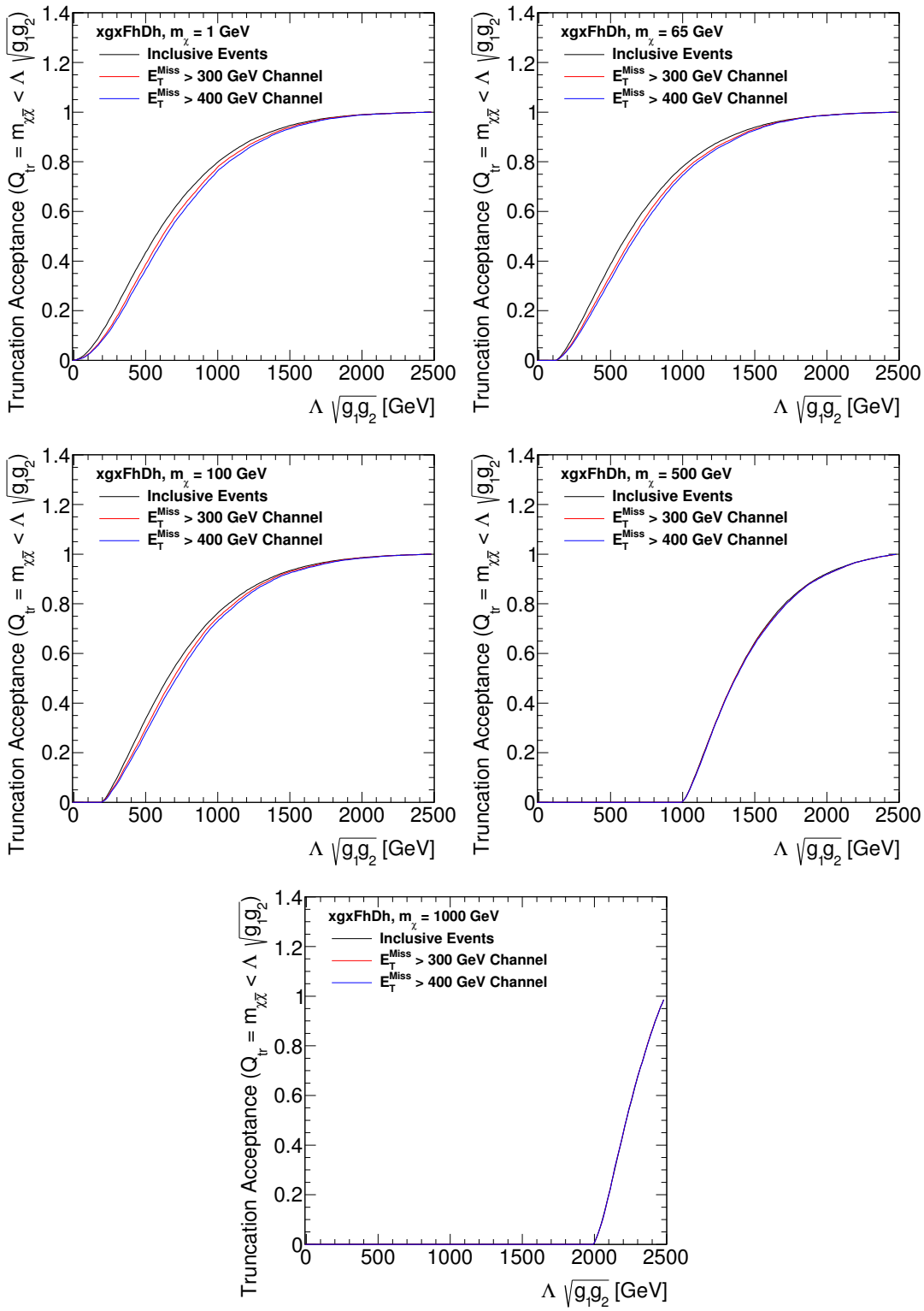


Fig. 8.6 Truncation efficiency of MC particle-level events satisfying $m_{\chi\chi} < \Lambda_0\sqrt{g_1 g_2}$ for xgxFhDh operator at different DM mass, assuming unit couplings ($g_1 = g_2 = 1$). Truncation acceptance are obtained for event

Chapter 9

Disccusion

While the analysis presented so far has yielded tangible and legitimate results, there are still rooms for improvement to be made. This section will explore the issues on how to increase the sensitivity (Section 9.1) and the constraint on the limit (Section 9.2) via advanced techniques. While the case of validity of the EFT model in LHC energy scale has been justified for 8 TeV data, the same may not be true for higher collision energy. Then in Section 9.3 we will discuss the inevitable departure from the EFT model to give way to more acceptable representation, i.e. the simplified model, at higher collision energy at the LHC.

9.1 Increase sensitivity via machine learning (ML) techniques

From Table 6.14 and Table 7.4 we know that the sensitivity of this analysis is limited by the statistics. To increase the sensitivity of this search, the obvious solution is to find ways to decrease the dominant background, $Z(\rightarrow \nu\nu)$+jets which constitute more than 60% of the total background. We note that although our cut-based selection provides a simple yet robust means to select our candidates events, further improvement can still be made. This is because in most cases the individual cuts in each observable are not able to exploit possible correlations among the different observables. Furthermore, a signal event will inevitably be misclassified as background in a cut-based analysis even if it look background-like in only a single observable. A multivariate classification approach (or ML), however, may correctly classify this type of events by compensating this one background-like feature by exploiting all the other observables that might look very signal-like.

Machine learning techniques have been widely used in high energy physics. Two popular methods based on traditional "shallow" architectures are Artificial Neural Network and the Boosted Decision Tree. Their performance have been scrutinised and proven in many cases [185–187] but here I propose to investigate the Artificial Neural Network with deep

architecture[1]" [188]. Compared to their "shallow" counterpart, deep neural networks have the potential to compute complex functions more efficiently than the shallow neural networks. Historically, these networks were considered impractical, in part due to the vanishing gradient problem [189–191] — the more layers you add, the harder it becomes to train. However, the advances primarily in computing power and available training data have made these methods significantly more feasible. Several phenomenological studies have benchmarked deep learning performance for jet tagging [192, 193], as well as for the search of Higgs boson and super-symmetric particles [194, 195] but none have yet to investigate boosted Higgs($\rightarrow b\bar{b}$) signal topology. Hence, here I will explore the feasibility of using Deep Neural Network (DNN)[2] to discriminate our signal against the $Z(\rightarrow \nu\nu)$+jets background. The DNN is built by using TensorFlow [196] software library.

First, let us define the objectives we want to achieve through this study:

- to discriminate our "signal" jet, the large-R jet originating from Higgs($\rightarrow b\bar{b}$) against "background' jet that comes from the overwhelming QCD multi jet (e.g gluon$\rightarrow b\bar{b}$).
- to compare the performance between the current cut-based selection with that of using DNN.
- to quantify the improvement in signal selection efficiency and background rejection.

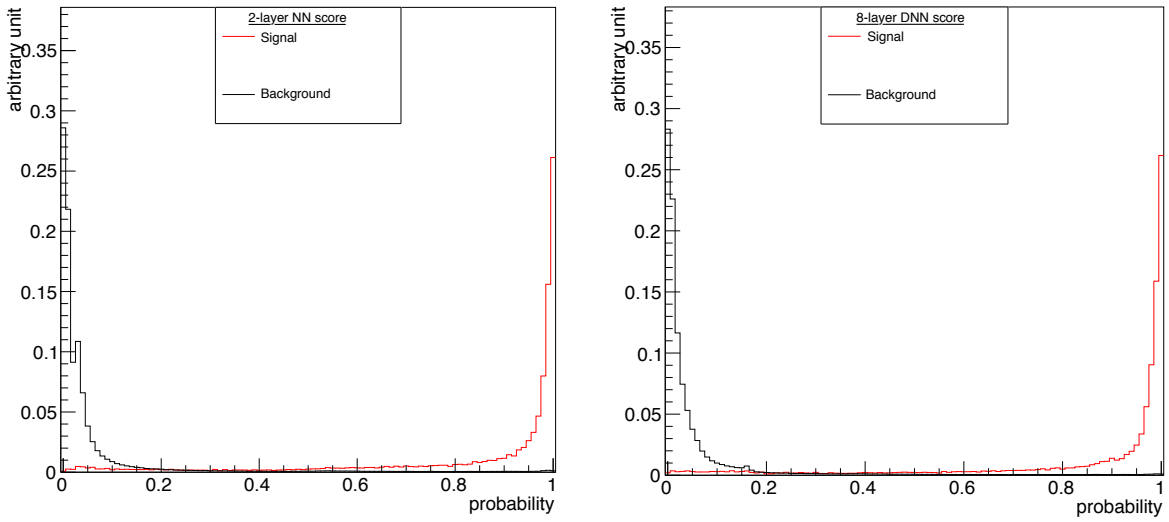
The mono-Higgs($\rightarrow b\bar{b}$) signal and $Z(\rightarrow \nu\nu)$+jets background MC samples described in Section 3 are used for this study. Leading large-R jet is matched to truth b-hadron by requiring $\Delta R(J^1,\text{b-hadron}) < 1.0$ where the b-hadron direction is taken from the MC information. All signal jets are required to have two matching b-hadrons whereas background jets are demanded to be associated to either 0/1/2 b-hadrons. For training, all leading large-R jets are required to have $p_{\mathrm{T}} > 200$ GeV. From a total of 505332 (3171982) signal (background) events, 65%, 30% and 5% of the events are assigned as the training, testing and validation set, respectively.

The discriminant variables used as input to the DNN are large-R jet mass, large-R jet p_{T}, the number of track jets and the number of b-tagged track jets that are associated to the leading large-R jet. The DNN is constructed from 8 hidden layers, containing 10, 20, 20, 10, 10, 10, 10, and 5 neurons, respectively. For this study, no tuning of the hyper-parameters (learning rate, the number of hidden layers, the number of neurons in each layer e.t.c) is performed. The gradient computations were made on mini-batches of size 128 by using Root Mean Square Propagation algorithm (a generalisation of resilient back-propagation algorithm [197]). The validation is performed every 1000 training steps. Early stopping (a form of regularisation used to avoid over-fitting) is implemented. The training is stopped when the accuracy of the trained

[1] There is no universally agreed upon threshold of depth dividing shallow neural network from deep neural network, but most researchers in the field agree that the latter has multiple nonlinear layers (> 2)

[2] The discussion of this section assumes prior knowledge on machine learning techniques. There is a wealth of literature addressing the topic of DNN. Hence the detailed explanation of their architectures and working principle will be omitted.

model does not improve over a period of 10000 steps. As a comparison, another two shallow 2-layer NNs, one with 4 and 2 neurons in each layer and the second contains 10 and 5 neurons in each layer are also constructed. The training procedure is similar to the training of DNN. The output of each NN is a probability score (between 0 and 1) for each large-R jet. Value closer to one means the large-R jet has higher probability to be of the signal jet that originating from Higgs($\rightarrow b\bar{b}$) and vice versa. The distribution for signal and background jets is shown in Figure 9.1. We see that there is a clear separation between signal and background jets. By applying a cut on this probability score, a desire signal selection efficiency and background rejection can be obtained.



(a) (b)

Fig. 9.1 Distribution of (a) 2-layer NN and (b) 8-layer DNN score for each large-R jet.

To compare the performance with our current cut-based analysis, test samples are constructed from those events subjected to the same signal event selection up to large-R jet multiplicity ($n_J \geq 1$) in Table 5.1. The total events after $n_J \geq 1$ selection is assigned as the denominator when assessing the selection efficiency. On the other hand, the numerator will be the leftover events after the jet mass cut (cut-based method) and the DNN score cut (DNN method). To visualise the performance of each NN, the background rejection (inverse of the background efficiency) is plotted against the signal efficiency at each cut on probability score, as shown in Figure 9.2. The selection efficiency for the cut-based method is plotted as a red star in the same plot. The DNN performance surpasses that of the cut based method for all the DM masses. However, the performance of the 2-layer NNs are not as good. Further optimisation of the network architecture or parameters is needed.

The background rejection of each NN and the cut based analysis are compared by fixing the signal efficiencies and background rejection to that of the cut-based method. The relative

changes are tabulated in Table 9.1. From the second and forth column of Table 9.1, we learn that the improvement as large as 24-36% (16-28%) in the background rejection (signal efficiency) is achieved by using the 8-layer DNN.

Table 9.1 Signal selection efficiency and background rejection comparison for mono-H operator `xgxFhDh` with different DM mass (m_χ =1-1000 GeV) and $Z(\to \nu\nu)$+jets.

	relative change (%)			
operator	NN background rej. at cut-based signal eff.		NN signal eff. at cut-based background rej.	
	2-layer NN 10-5	8-layer DNN	2-layer NN 10-5	8-layer DNN
mx1_xgxFhDh	-30.78	24.10	-18.90	16.19
mx65_xgxFhDh	-26.44	25.28	-16.26	17.12
mx100_xgxFhDh	-18.77	27.91	-11.96	18.81
mx500_xgxFhDh	13.99	35.10	9.82	27.83
mx1000_xgxFhDh	19.97	35.64	14.96	28.45

Further improvement can be achieved by using lowest-level highest-dimensionality variables as input to the NN. One such example is the so called "jet-image" [198, 193]. A jet-image is formed by taking the constituents of a jet (tracks, calorimeter clusters, truth particles, etc.) and discretisation its energy into pixels/grids in $\eta - \phi$ plane, with the intensity of each pixel given by the sum of the energy of all constituents of the jet inside that pixel. As a low-level jet description, the jet-image has the advantage that it uses all available information for later discriminatory techniques rather than compressing the information into a set of derived variables. High-level features can be extracted by borrowing the technique in the field of computer vision, or more specifically the deep Convolutional Neural Network (CNN). For readers who are interested, the detailed information of the construction of this discriminant will be presented in Appendix E. There, we will show that the use of low level discriminants such as jet-image, combined with track jet, b-tagged track jet multiplicity or jet mass, yields even better performance in classifying $H \to b\bar{b}$ jets.

The basic algorithm presented here only serve as a general proof of principle. Several questions need to be answers before it can be deployed. For example, what is the quantitative criteria to optimised? How to formally include systematic uncertainties? How can ML help us in understanding or protecting us from systematics? Can we teach a ML algorithm to be insensitive to a given mis-modelled observable? How to visualised or understand the underlying physics that has been learned by the algorithm? What is the theoretical relationship between classification and test sensitivity? All in all this technique can be more deeply developed in future studies.

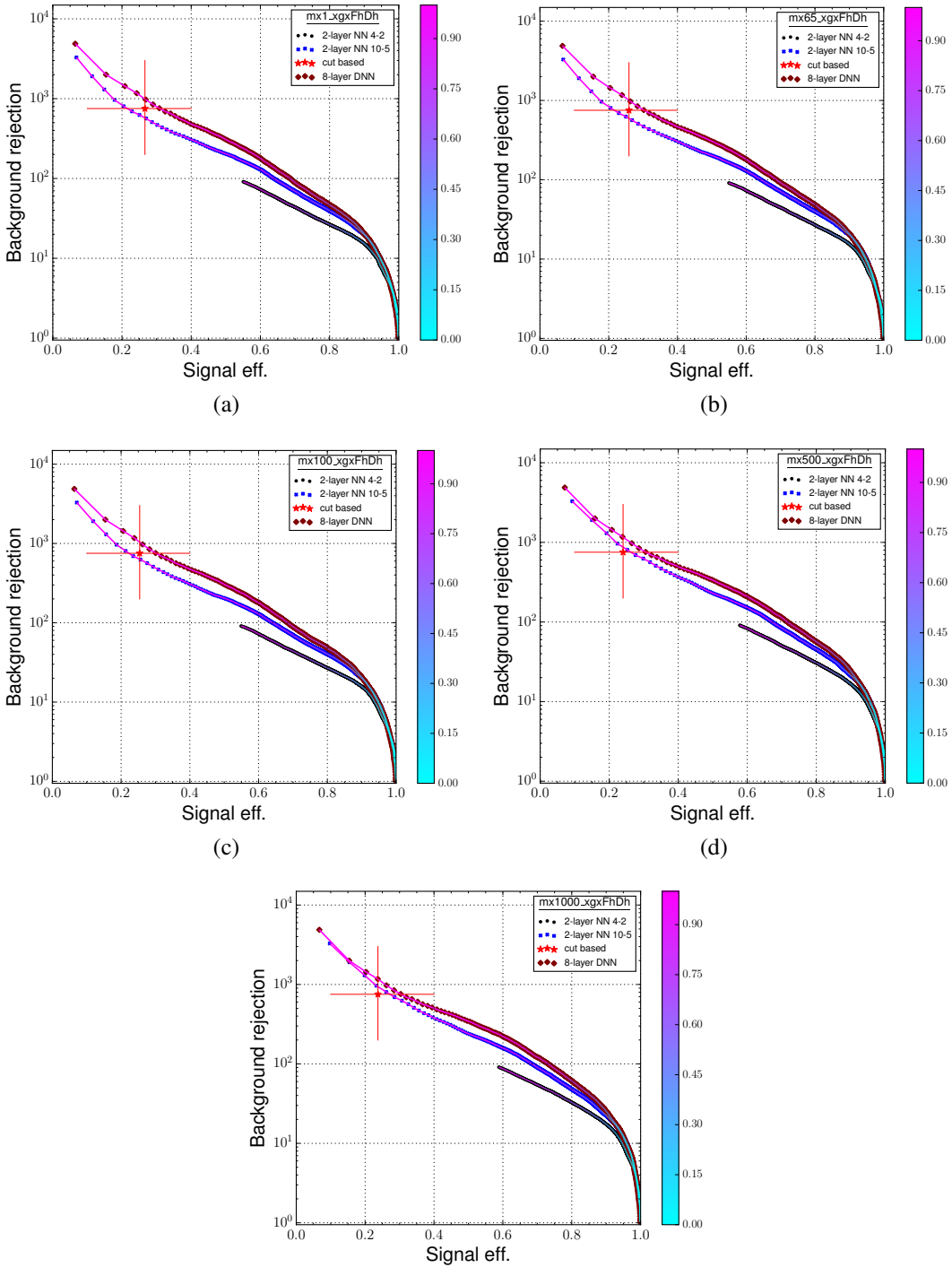


Fig. 9.2 Background rejection (inverse of efficiency) versus signal efficiency for 2-layer neural networks and 8-layer deep networks trained with leading large-R jet p_{T} leading large-R jet mass, track jet and b-tagged track jet multiplicity. The red star represents the cut-based signal efficiency and background rejection values. The colour map (z-axis) represents the threshold on the neural network output that is used to calculate the signal and background efficiencies.

9.2 Beyond single bin counting experiment

The construction of the likelihood function can be easily extended to include various CRs as additional constraint so that the likelihood can be rewritten as

$$\begin{aligned}\mathscr{L}(\mu,\theta) &= P_{SR}\times P_{CR}\times G_{syst} \\ &= \prod_{i=1}^{M}\frac{(\mu s_i+b_i)_i^n e^{-(\mu s_i+b_i)}}{n_i!}\times\prod_{j\in CR}^{K}\frac{u(\theta)_j^{m_j}e^{u(\theta)_j}}{m_j!}\times\prod_{k\in syst}G(\theta_k,\theta)\end{aligned} \tag{9.1}$$

where $n_i(i=1,...,M)$ is the number of events in the histogram bin i, s_i and b_i are and the number of expected signal and background events in the same histogram bin. The CRs can be considered as K additional bins with contents m_j (j=1,...,K) and expected values $u_j(\theta)$ depending not on μ but on the nuisance parameters θ. The statistically independent CRs and SRs, can then be modelled by separate PDFs and combined into a simultaneous fit to data. By doing so we can constrain the nuisance parameters in such a way to reduce their uncertainty and hence to reduce their impact on the final result μ. Furthermore, recall that in our fit strategy described in Section 8.2 the SR is constructed as a single bin region. In principle, by using multiple bins of a signal-sensitive observable in the definition of the SRs and CRs we can gain additional handle from the shape information of its distribution. Combining these two strategies, i.e simultaneous CRs fit and the inclusion of shape information, it is expected that we can get a better sensitivity.

9.3 Complementing EFT with simplified models and SUSY

While the EFT framework is a convenient tool for interpreting and comparing different DM searches, it is only justified whenever there is a clear separation between the energy scale of the process to describe and the scale of the underlying microscopic interactions. Our EFT truncation results indicate that the use of EFT for DM searches in a highly energetic environment should be handled with care. Figure 8.6 tells us that for the full set of generated events (for `xgxFhDh` operator) to fulfil the EFT validity condition the cutoff scale Λ need to be at least larger than a few TeV, unless the couplings constants involved in the processes are close to the non-perturbative regime (however, we must not forget that for large enough couplings constants, computations based on perturbation theory become unreliable). So, even at center of mass energy of 8 TeV the validity of EFT approach may already cause some concern, it is not hard to imagine the the challenge of higher center of mass energy might post.

For this reason, for future searches at the LHC perhaps a better strategy is to compliment the EFT approach with the simplified models whenever the scenarios under study involve mediating scale that is not too large. Such models are specifically designed to involve only the most important states (while other are integrated out) that mediates the DM interactions with

the SM particles, as well as the DM particle itself. By allowing the inclusion of these dynamical mediator state(s), simplified models allow one to more accurately describe the kinematics of DM production at the LHC. However, this comes with the price that they typically involve a handful of parameters that characterise the dark sector and its coupling to the visible sector. As an example, let us consider a scenario where the DM and the SM sectors are coupled through a new massive vector mediator, Z'. The production proceeds via the exchange of a Z' with mass $M_{Z'}$ in the s-channel. The Z' radiates a Higgs boson and decays into two DM particles. A thorough discussion of this model and its parameters can be found in Reference [52]. A diagram illustrating this process is shown in Figure 9.3. In exchange of the modification the EFT contact interactions into single particle s-channel, besides the DM mass, the model now becomes dependent on the couplings (g_q and g'_Z) and the mediator mass. Since this process has the same final state as our EFT signals, our results can easily be re-interpreted in the context of this simplified model.

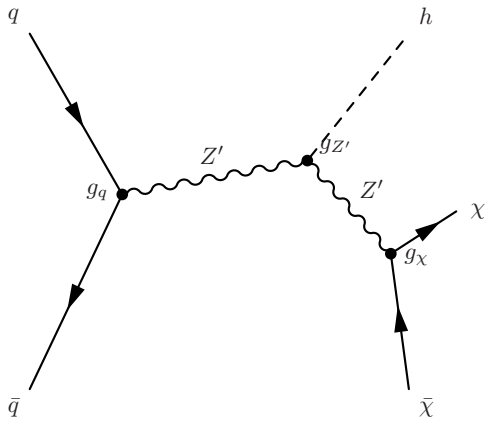


Fig. 9.3 Diagram showing collider production mode in a simplified model including a Z' boson which decays to DM pair.

Apart from the simplified model, another well motivated path is to aim at identifying DM explicitly assuming that the underlying theory is the so called $B-L$ Supersymmetric Standard Model[1] (BLSSM) [199–202]. BLSSM predicts similar mono-Higgs final state via the processes such as $q\bar{q} \to \tilde{\chi}_1^0\tilde{\chi}_i^0 \to \tilde{\chi}_1^0\tilde{\chi}_1^0 h$ and $q\bar{q} \to Z^{(\prime)} \to \tilde{\chi}_1^0\tilde{\chi}_i^0 \to \tilde{\chi}_1^0\tilde{\chi}_1^0 h$ as shown in Figure 9.4. We can study the consequence of our mono-Higgs search on the BLSSM parameter space by reinterpreting the mono-Higgs search data in the context of the BLSSM, thereby establishing a potential signal of SUSY DM or imposing strong bounds on their existence.

The point is, none of these approaches should stand in isolation. Given that the particle nature of DM and its interactions are still unknown, it is important that we include constraints that cover as broad a range of DM models as possible and at the same time in a way that is as

[1] Under this model, right-handed neutrino superfields are introduced (through an additional $U(1)B-L$ symmetry) in order to implement a Type I seesaw mechanism, which provides an elegant solution for the existence and smallness of the left-handed neutrino masses.

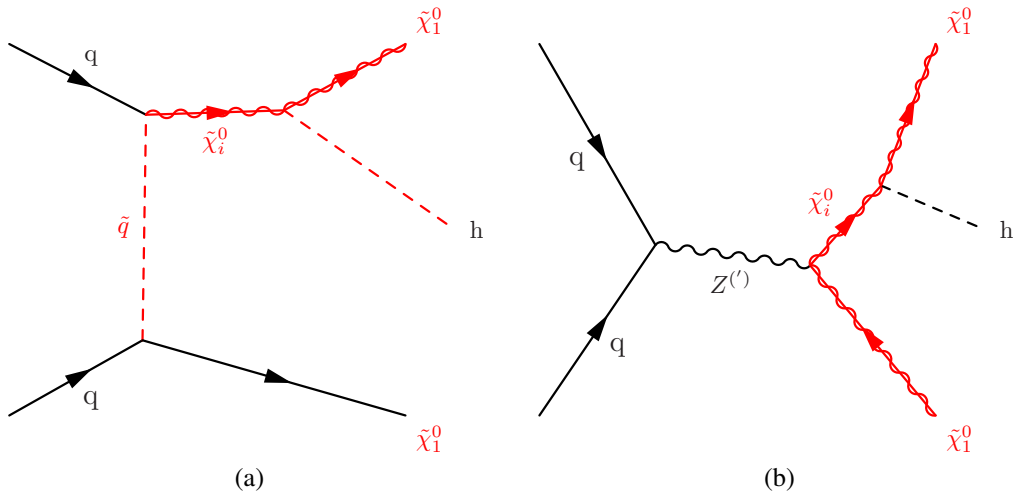


Fig. 9.4 Mono-Higgs final state arises via (a) $q\bar{q} \to \tilde{\chi}_1^0\tilde{\chi}_i^0 \to \tilde{\chi}_1^0\tilde{\chi}_1^0 h$ via a $\tilde{q}$ exchange and (b) $q\bar{q} \to Z^{(\prime)} \to \tilde{\chi}_1^0\tilde{\chi}_i^0 \to \tilde{\chi}_1^0\tilde{\chi}_1^0 h$ process within the BLSSM.

model-independent as possible. Furthermore, the various LHC searches have yet to provide a signal of a direct production of the heavy mediators such as Z' ($m_{Z'} \lesssim 3$ TeV excluded [203]), we are left with the conclusion that the EFT is valid as long as the energy scale of the process involving the DM and the SM particles is small compared to the energy scale associated to the heavy mediator. In addition, there may also be cases in which neither a ultraviolet completion nor other simplified models yielding mono-Higgs signal with the similar kinematic distributions are available. Under such circumstances, the EFT approach will allow us to target new signal regions and help to develop new simplified models (by validating the contact interaction limit) to complete these specific EFT operators.

Chapter 10

Conclusion

A search for dark matter production in association with the Higgs boson decaying to two bottom quarks has been presented. The Higgs boson candidates are reconstructed using trimmed anti$-k_t$ $R = 1.0$ jets matched to two b-tagged anti$-k_t$ $R = 0.3$ track-jets. This novel analysis technique probes new kinematic regions of mono-Higgs production where $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum is hard and the Higgs p_{T} is high.

The analysis is performed using data from pp collisions collected at $\sqrt{s} = 8$ TeV with the ATLAS experiment at the LHC at CERN, for an integrated luminosity of 20.3 fb^{-1}. Both the simulated and the data-driven methods are used to describe the background processes, reaching good agreement of data and expected background across all the control regions.

The observed (predicted) yields of the signal regions with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV are 20 (11.2$\pm$3.2) and 9 (7.7$\pm$2.5), respectively where the uncertainties of the predicted yields include both statistical and systematic uncertainty. The p-value of the former (latter) signal region is 0.03 (0.37). Hence we conclude that no excess is observed. The results are interpreted in the framework of the effective field theory in four operators and for DM mass up to 1 TeV. In general the stronger limits have been achieved compared to the previous mono-$H(\rightarrow \gamma\gamma)$ analysis at 8 TeV. In particular, the value of Λ up to 270 GeV has been excluded for operator `xgxFhDh`.

References

[1] Planck Collaboration, P. A. R. Ade et al., *Planck 2015 results. XIII. Cosmological parameters*, arXiv:1502.01589 [astro-ph.CO].

[2] F. Zwicky, *Republication of: The redshift of extragalactic nebulae*, General Relativity and Gravitation **41** no. 1, (2009) 207–224. http://dx.doi.org/10.1007/s10714-008-0707-4.

[3] D. Mihalas and P. Routly, *Galactic Astronomy*. A Series of books in astronomy and astrophysics. W. H. Freeman, 1968. https://books.google.co.jp/books?id=sfAInQEACAAJ.

[4] F. D. Kahn and L. Woltjer, *Intergalactic Matter and the Galaxy.*, Astrophys. J. **130** (1959) 705. http://adsabs.harvard.edu/abs/1959ApJ...130..705K. Provided by the SAO/NASA Astrophysics Data System.

[5] S. Smith, *The Mass of the Virgo Cluster*, Astrophys. J. **83** (1936) 23. http://adsabs.harvard.edu/abs/1936ApJ....83...23S. Provided by the SAO/NASA Astrophysics Data System.

[6] V. C. Rubin and W. K. Ford, Jr., *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, Astrophys. J. **159** (1970) 379.

[7] V. C. Rubin, N. Thonnard, and W. K. Ford, Jr., *Extended rotation curves of high-luminosity spiral galaxies. IV - Systematic dynamical properties, SA through SC*, Astrophys. J. **225** (1978) L107–L111.

[8] V. C. Rubin, W. K. J. Ford, and N. . Thonnard, *Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/*, Astrophys. J. **238** (1980) 471–487.

[9] M. S. Roberts and R. N. Whitehurst, *The rotation curve and geometry of M31 at large galactocentric distances.*, Astrophys. J. **201** (1975) 327–346.

[10] A. Bosma, *21-cm line studies of spiral galaxies. II. The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types.*, Astrophys. J. **86** (1981) 1825–1846.

[11] E. I. Gates, G. Gyuk, and M. S. Turner, *The Local Halo Density*, Astrophys. J. Lett. **449** (1995) L123, astro-ph/9505039.

[12] M. Weber and W. de Boer, *Determination of the local dark matter density in our Galaxy*, Astronomy and Astrophysics **509** (2010) A25, arXiv:0910.4272 [astro-ph.CO].

References

[13] E. Corbelli and P. Salucci, *The Extended Rotation Curve and the Dark Matter Halo of M33*, Mon. Not. Roy. Astron. Soc. **311** (2000) 441–447, arXiv:astro-ph/9909252 [astro-ph].

[14] T. G. Brainerd, *Anisotropic distribution of SDSS satellite galaxies: Planar (not polar) alignment*, Astrophys. J. **628** (2005) L101–L104, arXiv:astro-ph/0408559 [astro-ph].

[15] F. Iocco, M. Pato, and G. Bertone, *Evidence for dark matter in the inner Milky Way*, Nature Phys. **11** (2015) 245–248, arXiv:1502.03821 [astro-ph.GA].

[16] D. Fabricant, M. Lecar, and P. Gorenstein, *X-ray measurements of the mass of M87*, Astrophys. J. **241** (1980) 552–560.

[17] J. S. Arabadjis, M. W. Bautz, and G. P. Garmire, *Chandra observations of the lensing cluster emss 1358+6245: implications for self-interacting dark matter*, Astrophys. J. **572** (2002) 66, arXiv:astro-ph/0109141 [astro-ph].

[18] A. D. Lewis, D. A. Buote, and J. T. Stocke, *Chandra observations of Abell 2029: The Dark matter profile at < 0.01 R(VIR) in an unusually relaxed cluster*, Astrophys. J. **586** (2003) 135–142, arXiv:astro-ph/0209205 [astro-ph].

[19] R. Massey et al., *Dark matter maps reveal cosmic scaffolding*, Nature **445** (2007) 286, arXiv:astro-ph/0701594 [astro-ph].

[20] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, *A direct empirical proof of the existence of dark matter*, Astrophys. J. **648** (2006) L109–L113, arXiv:astro-ph/0608407 [astro-ph].

[21] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright, *Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results*, Astrophys. J. **208** (2013) 19, arXiv:1212.5226.

[22] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, J. Dunkley, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright, *Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results*, Astrophys. J. **208** (2013) 20, arXiv:1212.5225.

[23] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, R. C. Nichol, R. Scranton, H.-J. Seo, M. Tegmark, Z. Zheng, S. F. Anderson, J. Annis, N. Bahcall, J. Brinkmann, S. Burles, F. J. Castander, A. Connolly, I. Csabai, M. Doi, M. Fukugita, J. A. Frieman, K. Glazebrook, J. E. Gunn, J. S. Hendry, G. Hennessy, Z. Ivezić, S. Kent, G. R. Knapp, H. Lin, Y.-S. Loh, R. H. Lupton, B. Margon, T. A. McKay, A. Meiksin, J. A. Munn, A. Pope, M. W. Richmond, D. Schlegel, D. P. Schneider, K. Shimasaku, C. Stoughton, M. A. Strauss, M. SubbaRao, A. S. Szalay, I. Szapudi, D. L. Tucker, B. Yanny, and D. G. York, *Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies*, Astrophys. J. **633** (2005) 560–574, astro-ph/0501171.

[24] K. Jedamzik and M. Pospelov, *Big Bang nucleosynthesis and particle dark matter*, New Journal of Physics **11** no. 10, (2009) 105028.
http://stacks.iop.org/1367-2630/11/i=10/a=105028.

[25] R. Amanullah, C. Lidman, D. Rubin, G. Aldering, P. Astier, K. Barbary, M. S. Burns, A. Conley, K. S. Dawson, S. E. Deustua, M. Doi, S. Fabbro, L. Faccioli, H. K. Fakhouri, G. Folatelli, A. S. Fruchter, H. Furusawa, G. Garavini, G. Goldhaber, A. Goobar, D. E. Groom, I. Hook, D. A. Howell, N. Kashikawa, A. G. Kim, R. A. Knop, M. Kowalski, E. Linder, J. Meyers, T. Morokuma, S. Nobili, J. Nordin, P. E. Nugent, L. Östman, R. Pain, N. Panagia, S. Perlmutter, J. Raux, P. Ruiz-Lapuente, A. L. Spadafora, M. Strovink, N. Suzuki, L. Wang, W. M. Wood-Vasey, N. Yasuda, and T. Supernova Cosmology Project, *Spectra and Hubble Space Telescope Light Curves of Six Type Ia Supernovae at 0.511 < z < 1.12 and the Union2 Compilation*, Astrophys. J. **716** (2010) 712–738, arXiv:1004.1711 [astro-ph.CO].

[26] G. R. Blumenthal, S. M. Faber, J. R. Primack, and M. J. Rees, *Formation of galaxies and large-scale structure with cold dark matter*, Nature **311** (1984) 517–525.

[27] J. R. Primack, *Dark matter and structure formation*, in *Midrasha Mathematicae in Jerusalem: Winter School in Dynamical Systems Jerusalem, Israel, January 12-17, 1997*. 1997. arXiv:astro-ph/9707285 [astro-ph].
http://alice.cern.ch/format/showfull?sysnb=0254285.

[28] E. W. Kolb and M. S. Turner, *The Early Universe*, vol. 69 of *Frontiers in Physics*. Addison-Wesley, Redwood City, CA, 1990.

[29] S. Dodelson, *Modern Cosmology*. Academic Press, Amsterdam, 2003.
http://www.slac.stanford.edu/spires/find/books/www?cl=QB981:D62:2003.

[30] G. Gelmini and P. Gondolo, *DM Production Mechanisms*, arXiv:1009.3690 [astro-ph.CO].

[31] Particle Data Group Collaboration, K. A. Olive et al., *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.

[32] J. Billard, L. Strigari, and E. Figueroa-Feliciano, *Implication of neutrino backgrounds on the reach of next generation dark matter direct detection experiments*, Phys. Rev. **D89** no. 2, (2014) 023524, arXiv:1307.5458 [hep-ph].

[33] J. R. Ellis, K. A. Olive, and C. Savage, *Hadronic Uncertainties in the Elastic Scattering of Supersymmetric Dark Matter*, Phys. Rev. **D77** (2008) 065026, arXiv:0801.3656 [hep-ph].

[34] (LUX Collaboration) Collaboration, D. S. Akerib, H. M. Araújo, X. Bai, A. J. Bailey, J. Balajthy, S. Bedikian, E. Bernard, A. Bernstein, A. Bolozdynya, A. Bradley, D. Byram, S. B. Cahn, M. C. Carmona-Benitez, C. Chan, J. J. Chapman, A. A. Chiller, C. Chiller, K. Clark, T. Coffey, A. Currie, A. Curioni, S. Dazeley, L. de Viveiros, A. Dobi, J. Dobson, E. M. Dragowsky, E. Druszkiewicz, B. Edwards, C. H. Faham, S. Fiorucci, C. Flores, R. J. Gaitskell, V. M. Gehman, C. Ghag, K. R. Gibson, M. G. D. Gilchriese, C. Hall, M. Hanhardt, S. A. Hertel, M. Horn, D. Q. Huang, M. Ihm, R. G. Jacobsen, L. Kastens, K. Kazkaz, R. Knoche, S. Kyre, R. Lander, N. A. Larsen, C. Lee,

D. S. Leonard, K. T. Lesko, A. Lindote, M. I. Lopes, A. Lyashenko, D. C. Malling, R. Mannino, D. N. McKinsey, D.-M. Mei, J. Mock, M. Moongweluwan, J. Morad, M. Morii, A. S. J. Murphy, C. Nehrkorn, H. Nelson, F. Neves, J. A. Nikkel, R. A. Ott, M. Pangilinan, P. D. Parker, E. K. Pease, K. Pech, P. Phelps, L. Reichhart, T. Shutt, C. Silva, W. Skulski, C. J. Sofka, V. N. Solovov, P. Sorensen, T. Stiegler, K. O'Sullivan, T. J. Sumner, R. Svoboda, M. Sweany, M. Szydagis, D. Taylor, B. Tennyson, D. R. Tiedt, M. Tripathi, S. Uvarov, J. R. Verbus, N. Walsh, R. Webb, J. T. White, D. White, M. S. Witherell, M. Wlasenko, F. L. H. Wolfs, M. Woods, and C. Zhang, *First Results from the LUX Dark Matter Experiment at the Sanford Underground Research Facility*, Phys. Rev. Lett. **112** (2014) 091303. http://link.aps.org/doi/10.1103/PhysRevLett.112.091303.

[35] LUX Collaboration, D. S. Akerib et al., *Results from a search for dark matter in the complete LUX exposure*, Phys. Rev. Lett. **118** no. 2, (2017) 021303, arXiv:1608.07648 [astro-ph.CO].

[36] L. Baudis, *DARWIN: dark matter WIMP search with noble liquids*, PoS **IDM2010** (2011) 122, arXiv:1012.4764 [astro-ph.IM].

[37] LZ Collaboration, D. S. Akerib et al., *LUX-ZEPLIN (LZ) Conceptual Design Report*, arXiv:1509.02910 [physics.ins-det].

[38] PAMELA Collaboration, O. Adriani et al., *An anomalous positron abundance in cosmic rays with energies 1.5-100 GeV*, Nature **458** (2009) 607–609, arXiv:0810.4995 [astro-ph].

[39] PAMELA Collaboration, O. Adriani et al., *The cosmic-ray electron flux measured by the PAMELA experiment between 1 and 625 GeV*, Phys. Rev. Lett. **106** (2011) 201101, arXiv:1103.2880 [astro-ph.HE].

[40] PAMELA Collaboration, O. Adriani et al., *Cosmic-Ray Positron Energy Spectrum Measured by PAMELA*, Phys. Rev. Lett. **111** (2013) 081102, arXiv:1308.0133 [astro-ph.HE].

[41] AMS Collaboration, M. Aguilar et al., *First Result from the Alpha Magnetic Spectrometer on the International Space Station: Precision Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5?350 GeV*, Phys. Rev. Lett. **110** (2013) 141102.

[42] AMS Collaboration, L. Accardo et al., *High Statistics Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5?500 GeV with the Alpha Magnetic Spectrometer on the International Space Station*, Phys. Rev. Lett. **113** (2014) 121101.

[43] R. Catena and P. Ullio, *A novel determination of the local dark matter density*, Journal of Cosmology and Astroparticle Physics **2010** no. 08, (2010) 004. http://stacks.iop.org/1475-7516/2010/i=08/a=004.

[44] J. Goodman, M. Ibe, A. Rajaraman, W. Shepherd, T. M. P. Tait, and H.-B. Yu, *Constraints on Dark Matter from Colliders*, Phys. Rev. **D82** (2010) 116010, arXiv:1008.1783 [hep-ph].

[45] ATLAS Collaboration, ATLAS Collaboration, *Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 299, arXiv:1502.01518 [hep-ex].

[46] ATLAS Collaboration, ATLAS Collaboration, *Search for dark matter in events with heavy quarks and missing transverse momentum in pp collisions with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 92, arXiv:1410.4031 [hep-ex].

[47] ATLAS Collaboration, ATLAS Collaboration, *Search for new phenomena in events with a photon and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Phys. Rev. **D91** (2015) 012008, arXiv:1411.1559 [hep-ex].

[48] ATLAS Collaboration, ATLAS Collaboration, *Search for dark matter in events with a Z boson and missing transverse momentum in pp collisions at $\sqrt{s}$=8 TeV with the ATLAS detector*, Phys. Rev. **D90** (2014) 012004, arXiv:1404.0051 [hep-ex].

[49] ATLAS Collaboration, ATLAS Collaboration, *Search for dark matter in events with a hadronically decaying W or Z boson and missing transverse momentum in pp collisions at $\sqrt{s}$=8 TeV with the ATLAS detector*, Phys. Rev. Lett. **112** (2014) 041802, arXiv:1309.4017 [hep-ex].

[50] ATLAS Collaboration, G. Aad et al., *Search for new particles in events with one lepton and missing transverse momentum in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector*, JHEP **09** (2014) 037, arXiv:1407.7494 [hep-ex].

[51] ATLAS Collaboration, G. Aad et al., *Search for Dark Matter in Events with Missing Transverse Momentum and a Higgs Boson Decaying to Two Photons in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, Phys. Rev. Lett. **115** no. 13, (2015) 131801, arXiv:1506.01081 [hep-ex].

[52] L. Carpenter, A. DiFranzo, M. Mulhearn, C. Shimmin, S. Tulin, et al., *Mono-Higgs: a new collider probe of dark matter*, arXiv:1312.2592 [hep-ph].

[53] A. Djouadi, O. Lebedev, Y. Mambrini, and J. Quevillon, *Implications of LHC searches for Higgs–portal dark matter*, Phys. Lett. **B709** (2012) 65–69, arXiv:1112.3299 [hep-ph].

[54] J. McDonald, *Gauge singlet scalars as cold dark matter*, Phys.Rev. **D50** (1994) 3637–3649, arXiv:hep-ph/0702143 [HEP-PH].

[55] C. Burgess, M. Pospelov, and T. ter Veldhuis, *The Minimal model of nonbaryonic dark matter: A Singlet scalar*, Nucl.Phys. **B619** (2001) 709–728, arXiv:hep-ph/0011335 [hep-ph].

[56] B. Patt and F. Wilczek, *Higgs-field portal into hidden sectors*, arXiv:hep-ph/0605188 [hep-ph].

[57] Y. G. Kim and K. Y. Lee, *The Minimal model of fermionic dark matter*, Phys.Rev. **D75** (2007) 115012, arXiv:hep-ph/0611069 [hep-ph].

References

[58] J. March-Russell, S. M. West, D. Cumberbatch, and D. Hooper, *Heavy Dark Matter Through the Higgs Portal*, JHEP **0807** (2008) 058, arXiv:0801.3440 [hep-ph].

[59] I. Low, P. Schwaller, G. Shaughnessy, and C. E. Wagner, *The dark side of the Higgs boson*, Phys.Rev. **D85** (2012) 015009, arXiv:1110.4405 [hep-ph].

[60] L. Lopez-Honorez, T. Schwetz, and J. Zupan, *Higgs portal, fermionic dark matter, and a Standard Model like Higgs at 125 GeV*, Phys. Lett. **B716** (2012) 179–185, arXiv:1203.2064 [hep-ph].

[61] G. Belanger, B. Dumont, U. Ellwanger, J. Gunion, and S. Kraml, *Status of invisible Higgs decays*, Phys.Lett. **B723** (2013) 340–347, arXiv:1302.5694 [hep-ph].

[62] A. A. Petrov and W. Shepherd, *Searching for dark matter at LHC with Mono-Higgs production*, arXiv:1311.1511 [hep-ph].

[63] G. Busoni, A. De Simone, E. Morgante, and A. Riotto, *On the Validity of the Effective Field Theory for Dark Matter Searches at the LHC*, Phys. Lett. **B728** (2014) 412–421, arXiv:1307.2253 [hep-ph].

[64] Y. Cheng and M. Shochet, *Search for Dark Matter Produced in Association with a Higgs Boson Decaying to Two Bottom Quarks at ATLAS*. PhD thesis, Chicago U., Nov, 2015. http://cds.cern.ch/record/2102935. Presented 11 Nov 2015.

[65] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN, Geneva, 2004. https://cds.cern.ch/record/782076.

[66] O. Buning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report. 2. The LHC infrastructure and general services*,.

[67] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl, *LHC Design Report. 3. The LHC injector chain*,.

[68] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[69] LHC Study Group Collaboration, T. S. Pettersson and P. Lefèvre, *The Large Hadron Collider: conceptual design*, Tech. Rep. CERN-AC-95-05-LHC, CERN, Geneva, Oct, 1995. https://cds.cern.ch/record/291782.

[70] ATLAS Collaboration Collaboration, G. Aad et al., *Expected performance of the ATLAS experiment: detector, trigger and physics*. CERN, Geneva, 2009. https://cds.cern.ch/record/1125884.

[71] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[72] ATLAS Collaboration, A. Airapetian et al., *ATLAS: Detector and physics performance technical design report. Volume 1*,. http://cds.cern.ch/record/391176.

[73] ATLAS Collaboration, A. Airapetian et al., *ATLAS: Detector and physics performance technical design report. Volume 2*,. http://cds.cern.ch/record/391177.

[74] ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS Detector using First Collision Data*, JHEP **09** (2010) 056, arXiv:1005.5254 [hep-ex].

[75] ATLAS Collaboration, G. Aad et al., *Studies of the performance of the ATLAS detector using cosmic-ray muons*, Eur. Phys. J. **C71** (2011) 1593, arXiv:1011.6665 [physics.ins-det].

[76] C. Lefèvre, "The CERN accelerator complex. Complexe des accélérateurs du CERN." Dec, 2008.

[77] M. Lamont, *The First Years of LHC Operation for Luminosity Production*,. https://cds.cern.ch/record/2010134.

[78] "Luminosity public results." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[79] ATLAS Collaboration, *ATLAS inner detector: Technical design report. Vol. 1*,.

[80] ATLAS Collaboration, *ATLAS inner detector: Technical design report. Vol. 2*,.

[81] G. Aad et al., *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007.

[82] A. Ahmad et al., *The Silicon microstrip sensors of the ATLAS semiconductor tracker*, Nucl. Instrum. Meth. **A578** (2007) 98–118.

[83] ATLAS TRT Collaboration, E. Abat et al., *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance*, JINST **3** (2008) P02013.

[84] ATLAS Collaboration Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996. https://cds.cern.ch/record/331061.

[85] ATLAS Collaboration, *ATLAS muon spectrometer: Technical design report*,.

[86] ATLAS Collaboration, G. Aad et al., *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. **C72** (2012) 1849, arXiv:1110.1530 [hep-ex].

[87] ATLAS Collaboration, W. Buttinger, *The ATLAS Level-1 Trigger system*, J. Phys. Conf. Ser. **396** (2012) 012010.

[88] ATLAS Collaboration, *Luminosity Determination in pp Collisions at* $\sqrt{s} = 7$ TeV *Using the ATLAS Detector at the LHC*, Eur. Phys. J. C **71** (2011) 1630, arXiv:1101.2185 [hep-ph].

[89] ATLAS Collaboration, ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[90] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput.Phys.Commun. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[91] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[92] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, Comput.Phys.Commun. **184** (2013) 919–985, arXiv:hep-ph/0405247 [hep-ph].

[93] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **0711** (2007) 070, arXiv:0709.2092 [hep-ph].

[94] S. Frixione and B. R. Webber, *The MC@NLO 3.3 Event Generator*, arXiv:hep-ph/0612272 [hep-ph].

[95] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, arXiv:hep-ph/0204244 [hep-ph].

[96] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, *Single-top production in MC@NLO*, JHEP **03** (2006) 092, arXiv:hep-ph/0512250 [hep-ph].

[97] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5 : going beyond*, JHEP **06** (2011) 128, arXiv:1106.0522 [hep-ph].

[98] ATLAS Collaboration Collaboration, *ATLAS Computing: technical design report*. Technical Design Report ATLAS. CERN, Geneva, 2005. http://cds.cern.ch/record/837738.

[99] P. Calafiura, W. Lavrijsen, C. Leggett, M. Marino, and D. Quarrie, *The athena control framework in production, new developments and lessons learned*, in *Computing in high energy physics and nuclear physics. Proceedings, Conference, CHEP'04, Interlaken, Switzerland, September 27-October 1, 2004*. 2005. http://doc.cern.ch/yellowrep/2005/2005-002/p456.pdf.

[100] A. Rimoldi and A. Dell'Acqua, *The Full detector simulation for the ATLAS experiment: Status and outlook*, eConf **C0303241** (2003) TUMT001, arXiv:physics/0306086 [physics].

[101] D. Costanzo, A. Dell'Acqua, M. Gallas, A. Nairz, N. Benekos, A. Rimoldi, J. Boudreau, and V. Tsulaia, *Validation of the GEANT4-Based Full Simulation Program for the ATLAS Detector: An Overview of Performance and Robustness*, Tech. Rep. ATL-SOFT-PUB-2005-002. ATL-COM-SOFT-2005-004. CERN-ATL-SOFT-PUB-2005-002, CERN, Geneva, Mar, 2005. https://cds.cern.ch/record/830149.

[102] E. Richter-Was, D. Froidevaux, and L. Poggioli, *ATLFAST 2.0 a fast simulation package for ATLAS*, Tech. Rep. ATL-PHYS-98-131, CERN, Geneva, Nov, 1998.

[103] ATLAS Collaboration, A. Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*,.

[104] ATLAS Collaboration, W. Lukas, *Fast Simulation for ATLAS: Atlfast-II and ISF*, J. Phys. Conf. Ser. **396** (2012) 022031.

[105] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, Tech. Rep. ATL-PHYS-PUB-2012-003, CERN, Geneva, Aug, 2012. http://cds.cern.ch/record/1474107.

[106] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, et al., *Implications of CTEQ global analysis for collider observables*, Phys.Rev. **D78** (2008) 013004, arXiv:0802.0007 [hep-ph].

[107] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, CERN-2011-002 (CERN, Geneva, 2011), arXiv:1101.0593 [hep-ph].

[108] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: A program for Higgs boson decays in the Standard Model and its supersymmetric extension*, Comput. Phys. Commun. **108** (1998) 56–74.

[109] O. Brein, A. Djouadi, and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, Phys.Lett. **B579** (2004) 149–156, arXiv:hep-ph/0307206 [hep-ph].

[110] M. Ciccolini, S. Dittmaier, and M. Kramer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, Phys.Rev. **D68** (2003) 073003, arXiv:hep-ph/0306234 [hep-ph].

[111] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, and A. Salzburger, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, Tech. Rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar, 2007. https://cds.cern.ch/record/1020106.

[112] T. Cornelissen, M. Elsing, I. Gavrilenko, W. Liebig, E. Moyse, and A. Salzburger, *The new ATLAS track reconstruction (NEWT)*, J. Phys. Conf. Ser. **119** (2008) 032014.

[113] ATLAS Collaboration, M. Aaboud et al., *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC*, arXiv:1611.10235 [physics.ins-det].

[114] ATLAS Collaboration Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, Tech. Rep. ATLAS-CONF-2012-047, CERN, Geneva, May, 2012. https://cds.cern.ch/record/1449796.

[115] 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*, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008. https://cds.cern.ch/record/1099735.

[116] ATLAS Collaboration, G. Aad et al., *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. **C72** (2012) 1909, arXiv:1110.3174 [hep-ex].

[117] J. Alison, C. Anastopoulos, G. Artoni, E. Benhar Noccioli, A. Bocci, B. Brellier, K. Brendlinger, F. Buhrer, T. Cornelissen, M. Delmastro, O. Ducu, F. Dudziak, M. Elsing, O. Fedin, R. Fletcher, D. Froidevaux, I. Gavrilenko, T. Guillemin, S. Heim, F. Hubaut, L. Iconomidou-Fayard, M. Karnevskiy, T. Koffas, J. Kretzschmar, J. Kroll, C. Lester, K. Lohwasser, J. Maurer, A. Morley, G. Pasztor, E. Richter-Was, A. Schaffer,

T. Serre, P. Sommer, K. Tackmann, E. Tiouchichine, G. Unal, K. Whalen, and H. Williams, *Supporting document on electron efficiency measurements using the 2012 LHC proton-proton collision data*, Tech. Rep. ATL-COM-PHYS-2013-1295, CERN, Geneva, Sep, 2013. https://cds.cern.ch/record/1598905. Supporting document of electron efficiency measurements for 2012.

[118] ATLAS Collaboration, T. A. collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*,.

[119] ATLAS Collaboration, G. Aad et al., *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. **C74** no. 10, (2014) 3071, arXiv:1407.5063 [hep-ex].

[120] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data*, Eur. Phys. J. C **76** no. CERN-EP-2016-110. 12, (2016) 666. 40 p, arXiv:1606.01813 [hep-ex]. https://cds.cern.ch/record/2158117.

[121] K. Liu, Y. Liu, G. Marchiori, and E. Soldatov, *Measurement of the identification efficiency of isolated prompt photons using radiative $Z \to \ell\ell\gamma$ decays in 4.9 fb^{-1} of ATLAS data*, Tech. Rep. ATL-COM-PHYS-2012-382, CERN, Geneva, Apr, 2012. https://cds.cern.ch/record/1437004.

[122] K. Liu, Y. Liu, and G. Marchiori, *Measurement of the identification efficiency of isolated prompt photons using the matrix method and 4.9 fb^{-1} of ATLAS data*, Tech. Rep. ATL-COM-PHYS-2012-242, CERN, Geneva, Mar, 2012. https://cds.cern.ch/record/1428909.

[123] M. Jimenez and K. Tackmann, *Photon identification efficiency extrapolated from electrons in $Z \to e^+e^-$ decays*, Tech. Rep. ATL-COM-PHYS-2012-241, CERN, Geneva, Mar, 2012. https://cds.cern.ch/record/1428906.

[124] ATLAS Collaboration, G. Aad et al., *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton?proton collision data*, Eur. Phys. J. **C74** no. 11, (2014) 3130, arXiv:1407.3935 [hep-ex].

[125] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[126] ATLAS Collaboration, G. Aad et al., *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 17, arXiv:1406.0076 [hep-ex].

[127] *Determination of the jet energy scale and resolution at ATLAS using Z/γ–jet events in data at $\sqrt{s} = 8$ TeV*, Tech. Rep. ATLAS-CONF-2015-057, CERN, Geneva, Oct, 2015. https://cds.cern.ch/record/2059846.

[128] *Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8$ TeV*, Tech. Rep. ATLAS-CONF-2015-017, CERN, Geneva, Apr, 2015. https://cds.cern.ch/record/2008678.

[129] *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*, Tech. Rep. ATLAS-CONF-2015-037, CERN, Geneva, Aug, 2015. https://cds.cern.ch/record/2044941.

[130] ATLAS Collaboration Collaboration, *Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data*, Tech. Rep. ATLAS-CONF-2012-020, CERN, Geneva, Mar, 2012. https://cds.cern.ch/record/1430034.

[131] ATLAS Collaboration, G. Aad et al., *Characterisation and mitigation of beam-induced backgrounds observed in the ATLAS detector during the 2011 proton-proton run*, JINST **8** (2013) P07004, arXiv:1303.0223 [hep-ex].

[132] "Jet vertex fraction." https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JVFUncertaintyTool.

[133] ATLAS Collaboration, G. Aad et al., *Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, JHEP **09** (2013) 076, arXiv:1306.4945 [hep-ex].

[134] "$t\bar{t}$ resonance (single lepton) analysis package." https://svnweb.cern.ch/trac/atlasphys/browser/Physics/Top/PhysAnalysis/ttResoSingleLepton.

[135] ATLAS Collaboration, G. Aad et al., *Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment*, JHEP **05** (2012) 128, arXiv:1203.4606 [hep-ex].

[136] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, Phys. Lett. **B659** (2008) 119–126, arXiv:0707.1378 [hep-ph].

[137] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) 005, arXiv:0802.1188 [hep-ph].

[138] W. Van Den Wollenberg, M. Kagan, E. Thompson, and L. Zhou, *Measurement of the Track-Jet b-Tagging Efficiency of the MV1 algorithm in pp Collisions at $\sqrt{s} = 8$ TeV using eμ Dilepton $t\bar{t}$ events*, Tech. Rep. ATL-COM-PHYS-2015-009, CERN, Geneva, Jan, 2015. https://cds.cern.ch/record/1981533.

[139] R. Hawkings, *Measurement of the b-tagging efficiency of the MV1 algorithm in pp collisions at $\sqrt{s} = 8$ TeV using eμ dilepton $t\bar{t}$ events*, Tech. Rep. ATL-COM-PHYS-2013-381, CERN, Geneva, Apr, 2013. https://cds.cern.ch/record/1537694.

[140] M. Bosman, L. Fiorini, C. Helsens, A. Juste, L. Mir, J. Nadal, and V. Vorwerk, *Weighting method to propagate heavy-flavor tagging calibrations and related uncertainties*, Tech. Rep. ATL-COM-PHYS-2010-331, CERN, Geneva, Jun, 2010. https://cds.cern.ch/record/1269912.

[141] *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV*, Tech. Rep. ATLAS-CONF-2013-082, CERN, Geneva, Aug, 2013. https://cds.cern.ch/record/1570993.

[142] ATLAS Collaboration, G. Aad et al., *Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions in the ATLAS detector*, Eur. Phys. J. **C77** no. 4, (2017) 241, arXiv:1609.09324 [hep-ex].

[143] ATLAS Collaboration Collaboration, *Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at sqrts = 7 TeV in 2011*, Tech. Rep. ATLAS-CONF-2012-048, CERN, Geneva, May, 2012. https://cds.cern.ch/record/1450089.

[144] D. Casadei, I. Aracena, S. Banerjee, P.-H. Beauchemin, S. Calvet, K. Cranmer, D. Damazio, R. Djilkibaev, T. Eifert, S. Hillier, J. Idarraga, K. Johns, V. Kaushik, R. Konoplich, R. Kowalewski, X. Lei, A. Mann, P. Mermod, A. Mincer, J. Morel, P. Nemethy, A. Pinder, R. Taylor, A. Watson, and L. Zhao, *The implementation of the ATLAS missing Et triggers for the initial LHC operation*, Tech. Rep. ATL-DAQ-PUB-2011-001, CERN, Geneva, Feb, 2011. https://cds.cern.ch/record/1331180.

[145] ATLAS Collaboration Collaboration, *Performance of the ATLAS transverse energy triggers with initial LHC runs at sqrt(s) = 7 TeV*, Tech. Rep. ATLAS-CONF-2011-072, CERN, Geneva, May, 2011. https://cds.cern.ch/record/1351836.

[146] *The ATLAS transverse-momentum trigger performance at the LHC in 2011*, Tech. Rep. ATLAS-CONF-2014-002, CERN, Geneva, Feb, 2014. https://cds.cern.ch/record/1647616.

[147] G. P. Salam, *Towards Jetography*, Eur. Phys. J. **C67** (2010) 637–686, arXiv:0906.1833 [hep-ph].

[148] *Performance of boosted top quark identification in 2012 ATLAS data*, Tech. Rep. ATLAS-CONF-2013-084, CERN, Geneva, Aug, 2013. https://cds.cern.ch/record/1571040.

[149] K. Johns, R. Nayyar, and J. Veatch, *Jet mass scale and resolution uncertainty measurement for large-R jets at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Tech. Rep. ATL-COM-PHYS-2014-1468, CERN, Geneva, Nov, 2014. https://cds.cern.ch/record/1969519.

[150] A. R. Davison and I. Ochoa, *Validating the ATLAS Large R-parameter Jet Mass Scale and Resolution with Hadronically Decaying Boosted W Bosons Using the 2012 Dataset*, Tech. Rep. ATL-COM-PHYS-2013-865, CERN, Geneva, Jun, 2013. https://cds.cern.ch/record/1557901.

[151] S. van der Meer, *Calibration of the effective beam height in the ISR*, Tech. Rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968. http://cds.cern.ch/record/296752.

[152] H. Burkhardt and P. Grafström, *Absolute Luminosity from Machine Parameters*, Tech. Rep. LHC-PROJECT-Report-1019. CERN-LHC-PROJECT-Report-1019, 2007. https://cds.cern.ch/record/1056691.

[153] ATLAS Collaboration, M. Aaboud et al., *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C76** no. 12, (2016) 653, arXiv:1608.03953 [hep-ex].

[154] TOTEM Collaboration, G. Antchev et al., *First measurement of the total proton-proton cross section at the LHC energy of $\sqrt{s}$ =7 TeV*, Europhys. Lett. **96** (2011) 21002, arXiv:1110.1395 [hep-ex].

[155] ATLAS Collaboration, G. Aad et al., *Measurement of the Inelastic Proton-Proton Cross-Section at $\sqrt{s} = 7$ TeV with the ATLAS Detector*, Nature Commun. **2** (2011) 463, arXiv:1104.0326 [hep-ex].

[156] W. Buttinger, *Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data*, Tech. Rep. ATL-COM-SOFT-2015-119, CERN, Geneva, May, 2015. https://cds.cern.ch/record/2014726.

[157] S. Ask, M. A. Parker, T. Sandoval, M. E. Shea, and W. J. Stirling, *Using gamma+jets Production to Calibrate the Standard Model Z(nunu)+jets Background to New Physics Processes at the LHC*, JHEP **10** (2011) 058, arXiv:1107.2803 [hep-ph].

[158] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, D. Stump, W.-K. Tung, and C. P. Yuan, *Implications of CTEQ global analysis for collider observables*, Phys. Rev. **D78** (2008) 013004, arXiv:0802.0007 [hep-ph].

[159] J. M. Campbell, J. W. Huston, and W. J. Stirling, *Hard interactions of quarks and gluons: a primer for LHC physics*, Reports on Progress in Physics **70** no. 1, (2007) 89. http://stacks.iop.org/0034-4885/70/i=1/a=R02.

[160] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*, arXiv:1101.0538 [hep-ph].

[161] S. Gieseke, *Uncertainties of Sudakov form-factors*, JHEP **01** (2005) 058, arXiv:hep-ph/0412342 [hep-ph].

[162] U. Haisch, F. Kahlhoefer, and E. Re, *QCD effects in mono-jet searches for dark matter*, JHEP **12** (2013) 007, arXiv:1310.4491 [hep-ph].

[163] ATLAS Collaboration, ATLAS Collaboration, *Measurement of the cross-section for W boson production in association with b-jets in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, JHEP **06** (2013) 084, arXiv:1302.2929 [hep-ex].

[164] S. Moch and P. Uwer, *Heavy-quark pair production at two loops in QCD*, Nucl. Phys. Proc. Suppl. **183** (2008) 75–80, arXiv:0807.2794 [hep-ph].

[165] ATLAS Collaboration, ATLAS Collaboration, *Measurement of the top pair production cross section in 8 TeV proton-proton collisions using kinematic information in the lepton+jets final state with ATLAS*, Phys. Rev. **D91** (2015) 112013, arXiv:1504.04251 [hep-ex].

[166] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:1307.1347 [hep-ph].

[167] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C73** no. 8, (2013) 2518, arXiv:1302.4393 [hep-ex].

References

[168] *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, Tech. Rep. ATLAS-CONF-2014-046, CERN, Geneva, Jul, 2014. https://cds.cern.ch/record/1741020.

[169] J. K. Behr, M. Bellomo, M. Kagan, E. N. Thompson, L. Zhou, S. Willocq, and W. Van Den Wollenberg, *Search for a resonance in the boosted di-Higgs to 4b final state*,. https://cds.cern.ch/record/1741546.

[170] L. Carminati, D. Cavalli, M.-H. Genest, V. Ippolito, A. Nelson, L. Kashif, M. Perego, C. Pizio, M. G. Ratti, S. Resconi, C. Shimmin, F. Wang, D. Whiteson, M. Wu, S. L. Wu, and N. Zhou, *Search for new phenomena with the ATLAS detector in monophoton events from proton-proton collisions at sqrt(s)=8TeV*, Tech. Rep. ATL-PHYS-INT-2014-036, CERN, Geneva, Nov, 2014. https://cds.cern.ch/record/1971252.

[171] ATLAS Collaboration, G. Aad et al., *Measurements of normalized differential cross sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$??TeV using the ATLAS detector*, Phys. Rev. **D90** no. 7, (2014) 072004, arXiv:1407.0371 [hep-ex].

[172] G. Aad, B. Allbrooke, D. Buescher, A. Buckley, D. Cinca, Y. Coadou, P. Conde Muino, I. Connelly, B. Cooper, A. Davison, C. Debenedetti, Y. Enari, G. Facini, S. Fracchia, P. Francavilla, G. Gaycken, V. Giangobbe, R. Goncalo, G. Gonzalez, H. Gray, J. Grivaz, C. Gwilliam, S. Hageböck, G. Halladjian, M. Jackson, D. Jamin, K. Kiuchi, V. Kostyukhin, J. Lee, W. Lockman, K. Lohwasser, D. Lopez Mateos, L. Ma, A. Maio, J. Maneira, M. Martinez, U. Mallik, A. Mehta, K. Mercurio, K. Mochizuki, N. Morange, Y. Ming, Y. Nagai, J. Nielsen, I. Ochoa, H. Otono, G. Piacquadio, E. Pinto, M. Proissl, M. Sanders, T. Scanlon, B. Smart, P. Sommer, V. Sorin, J. Therhaag, J. Thomas-Wilsker, P. Thompson, L. Vacavant, J. Wang, C. Wang, S. Wang, C. Weiser, R. Zaidan, and L. Zhang, *Search for the Standard Model Higgs boson in associated production with a vector boson and decaying to bottom quarks with the ATLAS detector*, Tech. Rep. ATL-COM-PHYS-2013-465, CERN, Geneva, Apr, 2013. https://cds.cern.ch/record/1543067.

[173] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, arXiv:1007.1727 [physics.data-an].

[174] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Annals Math. Statist. **9** no. 1, (1938) 60–62.

[175] A. Wald, *Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large*, Transactions of the American Mathematical Society **54** no. 3, (1943) 426–482. http://www.jstor.org/stable/1990256.

[176] M. Baak, G. J. Besjes, D. Côte, A. Koutsman, J. Lorenz, and D. Short, *HistFitter software framework for statistical data analysis*, Eur. Phys. J. **C75** (2015) 153, arXiv:1410.1280 [hep-ex].

[177] A. L. Read, *Presentation of search results: The CL(s) technique*, J.Phys. **G28** (2002) 2693–2704.

[178] LUX Collaboration, D. Akerib et al., *First results from the LUX dark matter experiment at the Sanford Underground Research Facility*, Phys.Rev.Lett. **112** (2014) 091303, arXiv:1310.8214 [astro-ph.CO].

[179] D. Tucker-Smith and N. Weiner, *The Status of inelastic dark matter*, Phys. Rev. **D72** (2005) 063509, arXiv:hep-ph/0402065 [hep-ph].

[180] D. Tucker-Smith and N. Weiner, *Inelastic dark matter*, Phys. Rev. **D64** (2001) 043502, arXiv:hep-ph/0101138 [hep-ph].

[181] L. J. Hall, T. Moroi, and H. Murayama, *Sneutrino cold dark matter with lepton number violation*, Phys. Lett. **B424** (1998) 305–312, arXiv:hep-ph/9712515 [hep-ph].

[182] A. De Simone and T. Jacques, *Simplified models vs. effective field theory approaches in dark matter searches*, Eur. Phys. J. **C76** no. 7, (2016) 367, arXiv:1603.08002 [hep-ph].

[183] M. Endo and Y. Yamamoto, *Unitarity Bounds on Dark Matter Effective Interactions at LHC*, JHEP **06** (2014) 126, arXiv:1403.6610 [hep-ph].

[184] I. M. Shoemaker and L. Vecchi, *Unitarity and Monojet Bounds on Models for DAMA, CoGeNT, and CRESST-II*, Phys. Rev. **D86** (2012) 015023, arXiv:1112.5457 [hep-ph].

[185] ATLAS Collaboration Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys.Lett. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[186] C. Adam-Bourdarios, G. Cowan, C. Germain-Renaud, I. Guyon, B. Kégl, and D. Rousseau, *The Higgs Machine Learning Challenge*, Journal of Physics: Conference Series **664** no. 7, (2015) 072015. http://stacks.iop.org/1742-6596/664/i=7/a=072015.

[187] ATLAS Collaboration, G. Aad et al., *Performance of b-Jet Identification in the ATLAS Experiment*, arXiv:1512.01094 [hep-ex].

[188] L. Deng and D. Yu, *Deep Learning: Methods and Applications*, Foundations and Trends® in Signal Processing **7** no. 3–4, (2014) 197–387. http://dx.doi.org/10.1561/2000000039.

[189] Y. Bengio, P. Simard, and P. Frasconi, *Learning long-term dependencies with gradient descent is difficult*, IEEE Transactions on Neural Networks **5** no. 2, (1994) 157–166.

[190] S. Hochreiter, Y. Bengio, and P. Frasconi, *Gradient Flow in Recurrent Nets: the Difficulty of Learning Long-Term Dependencies*,.

[191] H. Larochelle, Y. Bengio, J. Louradour, and P. Lamblin, *Exploring Strategies for Training Deep Neural Networks*, J. Mach. Learn. Res. **10** (2009) 1–40. http://dl.acm.org/citation.cfm?id=1577069.1577070.

[192] D. Guest, J. Collado, P. Baldi, S.-C. Hsu, G. Urban, and D. Whiteson, *Jet Flavor Classification in High-Energy Physics with Deep Neural Networks*, Phys. Rev. **D94** no. 11, (2016) 112002, arXiv:1607.08633 [hep-ex].

[193] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, *Jet-images ? deep learning edition*, JHEP **07** (2016) 069, arXiv:1511.05190 [hep-ph].

[194] P. Baldi, P. Sadowski, and D. Whiteson, *Searching for Exotic Particles in High-Energy Physics with Deep Learning*, Nature Commun. **5** (2014) 4308, arXiv:1402.4735 [hep-ph].

[195] P. Sadowski, J. Collado, D. Whiteson, and P. Baldi, *Deep Learning, Dark Knowledge, and Dark Matter*, in *Proceedings of the 2014 International Conference on High-Energy Physics and Machine Learning - Volume 42*. 2014. http://dl.acm.org/citation.cfm?id=2996850.2996855.

[196] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*, ArXiv e-prints (2016), arXiv:1603.04467 [cs.DC].

[197] M. Riedmiller and H. Braun, *A direct adaptive method for faster backpropagation learning: the RPROP algorithm*,.

[198] J. Cogan, M. Kagan, E. Strauss, and A. Schwartzman, *Jet-Images: Computer Vision Inspired Techniques for Jet Tagging*, JHEP **02** (2015) 118, arXiv:1407.5675 [hep-ph].

[199] P. Fileviez Perez and S. Spinner, *The Fate of R-Parity*, Phys. Rev. **D83** (2011) 035004, arXiv:1005.4930 [hep-ph].

[200] P. Fileviez Perez, S. Spinner, and M. K. Trenkel, *The LSP Stability and New Higgs Signals at the LHC*, Phys. Rev. **D84** (2011) 095028, arXiv:1103.5504 [hep-ph].

[201] Z. M. Burell and N. Okada, *Supersymmetric minimal B-L model at the TeV scale with right-handed Majorana neutrino dark matter*, Phys. Rev. **D85** (2012) 055011, arXiv:1111.1789 [hep-ph].

[202] W. Abdallah, A. Hammad, S. Khalil, and S. Moretti, *Search for Mono-Higgs Signals at the LHC in the B-L Supersymmetric Standard Model*, Phys. Rev. **D95** no. 5, (2017) 055019, arXiv:1608.07500 [hep-ph].

[203] ATLAS Collaboration Collaboration, *Search for heavy resonances decaying to a W or Z boson and a Higgs boson in final states with leptons and b-jets in 36.1 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2017-055, CERN, Geneva, Jul, 2017. http://cds.cern.ch/record/2273871.

[204] *Expected electron performance in the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2011-006, CERN, Geneva, Apr, 2011. https://cds.cern.ch/record/1345327.

[205] ATLAS Collaboration, G. Aad et al., *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, Eur. Phys. J. **C74** no. 7, (2014) 2941, arXiv:1404.2240 [hep-ex].

[206] *Expected photon performance in the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2011-007, CERN, Geneva, Apr, 2011. https://cds.cern.ch/record/1345329.

[207] ATLAS Collaboration, G. Aad et al., *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, arXiv:1603.02934 [hep-ex].

[208] T. Barillari, E. Bergeaas Kuutmann, T. Carli, J. Erdmann, P. Giovannini, K. J. Grahn, C. Issever, A. Jantsch, A. Kiryunin, K. Lohwasser, A. Maslennikov, S. Menke, H. Oberlack, G. Pospelov, E. Rauter, P. Schacht, F. Spanó, P. Speckmayer, P. Stavina, and P. Strízenec, *Local Hadronic Calibration*, Tech. Rep. ATL-LARG-PUB-2009-001-2. ATL-COM-LARG-2008-006. ATL-LARG-PUB-2009-001, CERN, Geneva, Jun, 2008. http://cds.cern.ch/record/1112035. Due to a report-number conflict with another document, the report-number ATL-LARG-PUB-2009-001-2 has been assigned.

[209] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, Eur.Phys.J. **C73** (2013) 2304, arXiv:1112.6426 [hep-ex].

[210] ATLAS Collaboration, T. A. collaboration, *Pile-up subtraction and suppression for jets in ATLAS*,.

[211] ATLAS Collaboration, T. A. collaboration, *Jet global sequential corrections with the ATLAS detector in proton-proton collisions at sqrt(s) = 8 TeV*,.

[212] ATLAS Collaboration, G. Aad et al., *Readiness of the ATLAS Liquid Argon Calorimeter for LHC Collisions*, Eur. Phys. J. **C70** (2010) 723–753, arXiv:0912.2642 [physics.ins-det].

[213] C.-Y. Lee, S. Xie, P. Gallagher, Z. Zhang, and Z. Tu, *Deeply-Supervised Nets*, ArXiv e-prints (2014), arXiv:1409.5185 [stat.ML].

[214] I. J. Goodfellow, D. Warde-Farley, M. Mirza, A. Courville, and Y. Bengio, *Maxout Networks*, ArXiv e-prints (2013), arXiv:1302.4389 [stat.ML].

[215] M. D. Zeiler and R. Fergus, *Stochastic Pooling for Regularization of Deep Convolutional Neural Networks*, ArXiv e-prints (2013), arXiv:1301.3557 [cs.LG].

[216] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. A. Riedmiller, *Striving for Simplicity: The All Convolutional Net*, CoRR **abs/1412.6806** (2014). http://arxiv.org/abs/1412.6806.

Appendix A

Detailed lists of background MC samples

Table A.1 W+jets samples used in the analysis. The dataset ID, the dataset description, MC generator, production cross section, k-factor and filter efficiency are shown for each sample. For the DS IDs 167740-167748, a truth $p_{\mathrm{T}}^{W} < 70$ GeV cut is applied in order to combine the samples.

DS ID	Process	Generator	$\sigma \times$ BR [pb]	k-factor	$\varepsilon_{\text{filter}}$
167740	$W \to e\nu$, $p_{\mathrm{T}}^{W} > 0$ GeV, B filter	SHERPA	10973.00	1.11	0.012778
167741	$W \to e\nu$, $p_{\mathrm{T}}^{W} > 0$ GeV, C-jet filter & B veto	SHERPA	10971.00	1.11	0.049039
167742	$W \to e\nu$, $p_{\mathrm{T}}^{W} > 0$ GeV, B & C-jet veto	SHERPA	10987.00	1.11	0.938040
167743	$W \to \mu\nu$, $p_{\mathrm{T}}^{W} > 0$ GeV, B filter	SHERPA	10973.00	1.11	0.0128230
167744	$W \to \mu\nu$, $p_{\mathrm{T}}^{W} > 0$ GeV, C-jet filter & B veto	SHERPA	10970.00	1.11	0.042540
167745	$W \to \mu\nu$, $p_{\mathrm{T}}^{W} > 0$ GeV, B & C-jet veto	SHERPA	10981.00	1.11	0.944610
167746	$W \to \tau\nu$, $p_{\mathrm{T}}^{W} > 0$ GeV, B filter	SHERPA	10974.00	1.11	0.012791
167747	$W \to \tau\nu$, $p_{\mathrm{T}}^{W} > 0$ GeV, C-jet filter & B veto	SHERPA	10971.00	1.11	0.046082
167748	$W \to \tau\nu$, $p_{\mathrm{T}}^{W} > 0$ GeV, B & C-jet veto	SHERPA	10969.00	1.11	0.940650
167761	$W \to e\nu$, $70 < p_{\mathrm{T}}^{W} < 140$ GeV, B filter	SHERPA	250.5500	1.11	0.045931
167762	$W \to e\nu$, $70 < p_{\mathrm{T}}^{W} < 140$ GeV, C-jet filter & B veto	SHERPA	250.7100	1.11	0.200990
167763	$W \to e\nu$, $70 < p_{\mathrm{T}}^{W} < 140$ GeV, B & C-jet veto	SHERPA	250.4300	1.11	0.752980
167764	$W \to \mu\nu$, $70 < p_{\mathrm{T}}^{W} < 140$ GeV, B filter	SHERPA	250.5500	1.11	0.045919
167765	$W \to \mu\nu$, $70 < p_{\mathrm{T}}^{W} < 140$ GeV, C-jet filter & B veto	SHERPA	250.5700	1.11	0.198890
167766	$W \to \mu\nu$, $70 < p_{\mathrm{T}}^{W} < 140$ GeV, B & C-jet veto	SHERPA	250.7700	1.11	0.758550
167767	$W \to \tau\nu$, $70 < p_{\mathrm{T}}^{W} < 140$ GeV, B filter	SHERPA	250.5700	1.11	0.045942
167768	$W \to \tau\nu$, $70 < p_{\mathrm{T}}^{W} < 140$ GeV, C-jet filter & B veto	SHERPA	250.6100	1.11	0.198890
167769	$W \to \tau\nu$, $70 < p_{\mathrm{T}}^{W} < 140$ GeV, B & C-jet veto	SHERPA	250.6000	1.11	0.754850
167770	$W \to e\nu$, $140 < p_{\mathrm{T}}^{W} < 280$ GeV, B filter	SHERPA	31.15500	1.11	0.063159
167771	$W \to e\nu$, $140 < p_{\mathrm{T}}^{W} < 280$ GeV, C-jet filter & B veto	SHERPA	31.18900	1.11	0.221960
167772	$W \to e\nu$, $140 < p_{\mathrm{T}}^{W} < 280$ GeV, B & C-jet veto	SHERPA	31.11200	1.11	0.714960
167773	$W \to \mu\nu$, $140 < p_{\mathrm{T}}^{W} < 280$ GeV, B filter	SHERPA	31.16400	1.11	0.063069
167774	$W \to \mu\nu$, $140 < p_{\mathrm{T}}^{W} < 280$ GeV, C-jet filter & B veto	SHERPA	31.16500	1.11	0.216470
167775	$W \to \mu\nu$, $140 < p_{\mathrm{T}}^{W} < 280$ GeV, B & C-jet veto	SHERPA	31.17300	1.11	0.720300
167776	$W \to \tau\nu$, $140 < p_{\mathrm{T}}^{W} < 280$ GeV, B filter	SHERPA	31.16200	1.11	0.063078
167777	$W \to \tau\nu$, $140 < p_{\mathrm{T}}^{W} < 280$ GeV, C-jet filter & B veto	SHERPA	31.15100	1.11	0.220150
167778	$W \to \tau\nu$, $140 < p_{\mathrm{T}}^{W} < 280$ GeV, B & C-jet veto	SHERPA	31.17600	1.11	0.716090
167779	$W \to e\nu$, $280 < p_{\mathrm{T}}^{W} < 500$ GeV, B filter	SHERPA	1.841300	1.11	0.082886
167780	$W \to e\nu$, $280 < p_{\mathrm{T}}^{W} < 500$ GeV, C-jet filter & B veto	SHERPA	1.837000	1.11	0.234540
167781	$W \to e\nu$, $280 < p_{\mathrm{T}}^{W} < 500$ GeV, B & C-jet veto	SHERPA	1.842600	1.11	0.682000
167782	$W \to \mu\nu$, $280 < p_{\mathrm{T}}^{W} < 500$ GeV, B filter	SHERPA	1.838000	1.11	0.082902
167783	$W \to \mu\nu$, $280 < p_{\mathrm{T}}^{W} < 500$ GeV, C-jet filter & B veto	SHERPA	1.839500	1.11	0.228450
167784	$W \to \mu\nu$, $280 < p_{\mathrm{T}}^{W} < 500$ GeV, B & C-jet veto	SHERPA	1.843300	1.11	0.687760
167785	$W \to \tau\nu$, $280 < p_{\mathrm{T}}^{W} < 500$ GeV, B filter	SHERPA	1.836200	1.11	0.083026
167786	$W \to \tau\nu$, $280 < p_{\mathrm{T}}^{W} < 500$ GeV, C-jet filter & B veto	SHERPA	1.839500	1.11	0.232710
167787	$W \to \tau\nu$, $280 < p_{\mathrm{T}}^{W} < 500$ GeV, B & C-jet veto	SHERPA	1.836800	1.11	0.683970
167788	$W \to e\nu$, $p_{\mathrm{T}}^{W} > 500$ GeV, B filter	SHERPA	0.101880	1.11	0.099655
167789	$W \to e\nu$, $p_{\mathrm{T}}^{W} > 500$ GeV, C-jet filter & B veto	SHERPA	0.101010	1.11	0.244400
167790	$W \to e\nu$, $p_{\mathrm{T}}^{W} > 500$ GeV, B & C-jet veto	SHERPA	0.100930	1.11	0.657410
167791	$W \to \mu\nu$, $p_{\mathrm{T}}^{W} > 500$ GeV, B filter	SHERPA	0.101630	1.11	0.100040
167792	$W \to \mu\nu$, $p_{\mathrm{T}}^{W} > 500$ GeV, C-jet filter & B veto	SHERPA	0.102100	1.11	0.238520
167793	$W \to \mu\nu$, $p_{\mathrm{T}}^{W} > 500$ GeV, B & C-jet veto	SHERPA	0.101860	1.11	0.658370
167794	$W \to \tau\nu$, $p_{\mathrm{T}}^{W} > 500$ GeV, B filter	SHERPA	0.102080	1.11	0.099663
167795	$W \to \tau\nu$, $p_{\mathrm{T}}^{W} > 500$ GeV, C-jet filter & B veto	SHERPA	0.101390	1.11	0.242210
167796	$W \to \tau\nu$, $p_{\mathrm{T}}^{W} > 500$ GeV, B & C-jet veto	SHERPA	0.102010	1.11	0.660040

Table A.2 Z+jets samples used in the analysis. The dataset ID, the dataset description, MC generator, production cross section, k-factor and filter efficiency are shown for each sample. For the DS IDs 167749-167760, a truth $p_{\mathrm{T}}^{Z} < 70$ GeV cut is applied in order to combine the samples.

DS ID	Process	Generator	$\sigma \times$ BR [pb]	k-factor	$\varepsilon_{\mathrm{filter}}$
167749	$Z \to ee$, $p_{\mathrm{T}}^{Z} > 0$ GeV, B filter	SHERPA	1110.700	1.12	0.028034
167750	$Z \to ee$, $p_{\mathrm{T}}^{Z} > 0$ GeV, C filter & B veto	SHERPA	1109.600	1.12	0.283410
167751	$Z \to ee$, $p_{\mathrm{T}}^{Z} > 0$ GeV, B & C veto	SHERPA	1107.100	1.12	0.686210
167752	$Z \to \mu\mu$, $p_{\mathrm{T}}^{Z} > 0$ GeV, B filter	SHERPA	1109.800	1.12	0.027996
167753	$Z \to \mu\mu$, $p_{\mathrm{T}}^{Z} > 0$ GeV, C filter & B veto	SHERPA	1112.000	1.12	0.283070
167754	$Z \to \mu\mu$, $p_{\mathrm{T}}^{Z} > 0$ GeV, B & C veto	SHERPA	1108.700	1.12	0.689700
167755	$Z \to \tau\tau$, $p_{\mathrm{T}}^{Z} > 0$ GeV, B filter	SHERPA	1109.100	1.12	0.027820
167756	$Z \to \tau\tau$, $p_{\mathrm{T}}^{Z} > 0$ GeV, C filter & B veto	SHERPA	1110.200	1.12	0.283730
167757	$Z \to \tau\tau$, $p_{\mathrm{T}}^{Z} > 0$ GeV, B & C veto	SHERPA	1112.100	1.12	0.688840
167758	$Z \to \nu\nu$, $p_{\mathrm{T}}^{Z} > 0$ GeV, B filter	SHERPA	5991.100	1.12	0.029
167759	$Z \to \nu\nu$, $p_{\mathrm{T}}^{Z} > 0$ GeV, C filter & B veto	SHERPA	5991.200	1.12	0.283730
167760	$Z \to \nu\nu$, $p_{\mathrm{T}}^{Z} > 0$ GeV, B & C veto	SHERPA	5991.100	1.12	0.688840
167797	$Z \to ee$, $70 < p_{\mathrm{T}}^{Z} < 140$ GeV, B filter	SHERPA	29.49400	1.12	0.082517
167798	$Z \to ee$, $70 < p_{\mathrm{T}}^{Z} < 140$ GeV, C filter & B veto	SHERPA	29.48700	1.12	0.354970
167799	$Z \to ee$, $70 < p_{\mathrm{T}}^{Z} < 140$ GeV, B & C veto	SHERPA	29.49100	1.12	0.562620
167800	$Z \to \mu\mu$, $70 < p_{\mathrm{T}}^{Z} < 140$ GeV, B filter	SHERPA	29.49100	1.12	0.082585
167801	$Z \to \mu\mu$, $70 < p_{\mathrm{T}}^{Z} < 140$ GeV, C filter & B veto	SHERPA	29.44700	1.12	0.354880
167802	$Z \to \mu\mu$, $70 < p_{\mathrm{T}}^{Z} < 140$ GeV, B & C veto	SHERPA	29.52100	1.12	0.561960
167803	$Z \to \tau\tau$, $70 < p_{\mathrm{T}}^{Z} < 140$ GeV, B filter	SHERPA	29.48900	1.12	0.082563
167804	$Z \to \tau\tau$, $70 < p_{\mathrm{T}}^{Z} < 140$ GeV, C filter & B veto	SHERPA	29.49900	1.12	0.355090
167805	$Z \to \tau\tau$, $70 < p_{\mathrm{T}}^{Z} < 140$ GeV, B & C veto	SHERPA	29.49400	1.12	0.562470
167806	$Z \to \nu\nu$, $70 < p_{\mathrm{T}}^{Z} < 140$ GeV, B filter	SHERPA	166.63	1.12	0.082563
167807	$Z \to \nu\nu$, $70 < p_{\mathrm{T}}^{Z} < 140$ GeV, C filter & B veto	SHERPA	166.63	1.12	0.355090
167808	$Z \to \nu\nu$, $70 < p_{\mathrm{T}}^{Z} < 140$ GeV, B & C veto	SHERPA	166.63	1.12	0.562470
167809	$Z \to ee$, $140 < p_{\mathrm{T}}^{Z} < 280$ GeV, B filter	SHERPA	3.990100	1.12	0.095235
167810	$Z \to ee$, $140 < p_{\mathrm{T}}^{Z} < 280$ GeV, C filter & B veto	SHERPA	3.981100	1.12	0.369190
167811	$Z \to ee$, $140 < p_{\mathrm{T}}^{Z} < 280$ GeV, B & C veto	SHERPA	3.989000	1.12	0.534310
167812	$Z \to \mu\mu$, $140 < p_{\mathrm{T}}^{Z} < 280$ GeV, B filter	SHERPA	3.984200	1.12	0.095389
167813	$Z \to \mu\mu$, $140 < p_{\mathrm{T}}^{Z} < 280$ GeV, C filter & B veto	SHERPA	3.991100	1.12	0.369990
167814	$Z \to \mu\mu$, $140 < p_{\mathrm{T}}^{Z} < 280$ GeV, B & C veto	SHERPA	3.984100	1.12	0.534410
167815	$Z \to \tau\tau$, $140 < p_{\mathrm{T}}^{Z} < 280$ GeV, B filter	SHERPA	3.987800	1.12	0.095807
167816	$Z \to \tau\tau$, $140 < p_{\mathrm{T}}^{Z} < 280$ GeV, C filter & B veto	SHERPA	3.988000	1.12	0.369530
167817	$Z \to \tau\tau$, $140 < p_{\mathrm{T}}^{Z} < 280$ GeV, B & C veto	SHERPA	3.987100	1.12	0.533280
167818	$Z \to \nu\nu$, $140 < p_{\mathrm{T}}^{Z} < 280$ GeV, B filter	SHERPA	22.14	1.12	0.095807
167819	$Z \to \nu\nu$, $140 < p_{\mathrm{T}}^{Z} < 280$ GeV, C filter & B veto	SHERPA	22.14	1.12	0.369530
167820	$Z \to \nu\nu$, $140 < p_{\mathrm{T}}^{Z} < 280$ GeV, B & C veto	SHERPA	22.14	1.12	0.533280
167821	$Z \to ee$, $280 < p_{\mathrm{T}}^{Z} < 500$ GeV, B filter	SHERPA	0.241820	1.12	0.108510
167822	$Z \to ee$, $280 < p_{\mathrm{T}}^{Z} < 500$ GeV, C filter & B veto	SHERPA	0.241280	1.12	0.387440
167823	$Z \to ee$, $280 < p_{\mathrm{T}}^{Z} < 500$ GeV, B & C veto	SHERPA	0.241580	1.12	0.506170
167824	$Z \to \mu\mu$, $280 < p_{\mathrm{T}}^{Z} < 500$ GeV, B filter	SHERPA	0.242190	1.12	0.108020
167825	$Z \to \mu\mu$, $280 < p_{\mathrm{T}}^{Z} < 500$ GeV, C filter & B veto	SHERPA	0.241690	1.12	0.386430
167826	$Z \to \mu\mu$, $280 < p_{\mathrm{T}}^{Z} < 500$ GeV, B & C veto	SHERPA	0.242720	1.12	0.505490
167827	$Z \to \tau\tau$, $280 < p_{\mathrm{T}}^{Z} < 500$ GeV, B filter	SHERPA	0.241200	1.12	0.106530
167828	$Z \to \tau\tau$, $280 < p_{\mathrm{T}}^{Z} < 500$ GeV, C filter & B veto	SHERPA	0.241020	1.12	0.384810
167829	$Z \to \tau\tau$, $280 < p_{\mathrm{T}}^{Z} < 500$ GeV, B & C veto	SHERPA	0.241470	1.12	0.507200
167830	$Z \to \nu\nu$, $280 < p_{\mathrm{T}}^{Z} < 500$ GeV, B filter	SHERPA	1.3528	1.12	0.106530
167831	$Z \to \nu\nu$, $280 < p_{\mathrm{T}}^{Z} < 500$ GeV, C filter & B veto	SHERPA	1.3528	1.12	0.384810
167832	$Z \to \nu\nu$, $280 < p_{\mathrm{T}}^{Z} < 500$ GeV, B & C veto	SHERPA	1.3528	1.12	0.507200
167833	$Z \to ee$, $p_{\mathrm{T}}^{Z} > 500$ GeV, B filter	SHERPA	0.013235	1.12	0.115730
167834	$Z \to ee$, $p_{\mathrm{T}}^{Z} > 500$ GeV, C filter & B veto	SHERPA	0.013454	1.12	0.398460
167835	$Z \to ee$, $p_{\mathrm{T}}^{Z} > 500$ GeV, B & C veto	SHERPA	0.013307	1.12	0.484800
167836	$Z \to \mu\mu$, $p_{\mathrm{T}}^{Z} > 500$ GeV, B filter	SHERPA	0.013161	1.12	0.114080
167837	$Z \to \mu\mu$, $p_{\mathrm{T}}^{Z} > 500$ GeV, C filter & B veto	SHERPA	0.013480	1.12	0.398570
167838	$Z \to \mu\mu$, $p_{\mathrm{T}}^{Z} > 500$ GeV, B & C veto	SHERPA	0.013264	1.12	0.486890
167839	$Z \to \tau\tau$, $p_{\mathrm{T}}^{Z} > 500$ GeV, B filter	SHERPA	0.013231	1.12	0.115240
167840	$Z \to \tau\tau$, $p_{\mathrm{T}}^{Z} > 500$ GeV, C filter & B veto	SHERPA	0.013308	1.12	0.393160
167841	$Z \to \tau\tau$, $p_{\mathrm{T}}^{Z} > 500$ GeV, B & C veto	SHERPA	0.013284	1.12	0.485620
167842	$Z \to \nu\nu$, $p_{\mathrm{T}}^{Z} > 500$ GeV, B filter	SHERPA	0.0732	1.12	0.115240
167843	$Z \to \nu\nu$, $p_{\mathrm{T}}^{Z} > 500$ GeV, C filter & B veto	SHERPA	0.073	1.12	0.393160
167844	$Z \to \nu\nu$, $p_{\mathrm{T}}^{Z} > 500$ GeV, B & C veto	SHERPA	0.073	1.12	0.485620

Table A.3 γ+jets samples used for the analysis. The columns from left to right describe MC generation ID, the dataset description, the MC generator used, the sample cross section in pb, the NLO/LO K-factor (multiplicative to the cross section) and the filter efficiency of the requested sample within the ATLAS simulation.

ID	Name	Generator	σ [pb]	K-factor	ε
177574	single γ, $100 < p_{\mathrm{T}}^{\gamma} < 140$ GeV, B & C-jet veto	SHERPA	430.80	1.0	0.49984
177575	single γ, $140 < p_{\mathrm{T}}^{\gamma} < 280$ GeV, B & C-jet veto	SHERPA	137.98	1.0	0.49377
177576	single γ, $280 < p_{\mathrm{T}}^{\gamma} < 500$ GeV, B & C-jet veto	SHERPA	5.9250	1.0	0.48087
177577	single γ, $p_{\mathrm{T}}^{\gamma} > 500$ GeV, B & C-jet veto	SHERPA	0.28715	1.0	0.46776
177578	single γ, $100 < p_{\mathrm{T}}^{\gamma} < 140$ GeV, C-jet filter & B veto	SHERPA	430.40	1.0	0.42401
177579	single γ, $140 < p_{\mathrm{T}}^{\gamma} < 280$ GeV, C-jet filter & B veto	SHERPA	137.80	1.0	0.42182
177580	single γ, $280 < p_{\mathrm{T}}^{\gamma} < 500$ GeV, C-jet filter & B veto	SHERPA	5.9223	1.0	0.41936
177581	single γ, $p_{\mathrm{T}}^{\gamma} >$500GeV, C-jet filter & B veto	SHERPA	0.28811	1.0	0.41931
177582	single γ, $100 < p_{\mathrm{T}}^{\gamma} < 140$ GeV, B filter	SHERPA	428.83	1.0	0.075821
177583	single γ, $140 < p_{\mathrm{T}}^{\gamma} < 280$ GeV, B filter	SHERPA	137.31	1.0	0.08437
177584	single γ, $280 < p_{\mathrm{T}}^{\gamma} < 500$ GeV, B filter	SHERPA	5.9024	1.0	0.09983
177585	single γ, $p_{\mathrm{T}}^{\gamma} >$500GeV, B filter	SHERPA	0.28729	1.0	0.11408

Table A.4 $t\bar{t}$ and single top samples used for the analysis. The columns from left to right describe MC generation ID, the dataset description, the MC generator used, the sample cross section in pb, the NLO/LO K-factor (multiplicative to the cross section) and the filter efficiency of the requested sample within the ATLAS simulation.

ID	Name	Generator	σ [pb]	K-factor	ε
108343	SingleTopSChanWenu	McAtNloJimmy	0.56444	1.074	1.0
108344	SingleTopSChanWmunu	McAtNloJimmy	0.56426	1.074	1.0
108345	SingleTopSChanWtaunu	McAtNloJimmy	0.56404	1.074	1.0
108346	SingleTopWtChanIncl	McAtNloJimmy	20.658	1.083	1.0
117360	singletop_tchan_e	AcerMCPythia	8.604	1.1	1.0
117361	singletop_tchan_mu	AcerMCPythia	8.604	1.1	1.0
117362	singletop_tchan_tau	AcerMCPythia	8.604	1.1	1.0
105200	ttbar_LeptonFilter	McAtNloJimmy	238.06	1.0	0.543
105204	ttbar_allhad	McAtNloJimmy	238.06	1.0	0.457

Table A.5 Diboson samples used for the analysis. The columns from left to right describe MC generation ID, the dataset description, the MC generator used, the sample cross section in pb, the NLO/LO K-factor (multiplicative to the cross section) and the filter efficiency of the requested sample within the ATLAS simulation.

ID	Name	Generator	σ [pb]	K-factor	ε
181966	ZZ_2l2had_mll020	PowhegPythia8	1.207	1.0	1.0
181967	ZZ_2nu2had_mll020	PowhegPythia8	2.081	1.0	1.0
181968	WZ_hadhad2l_mll020	PowhegPythia8	1.5944	1.0	1.0
181969	WZ_hadhad2nu_mll020	PowhegPythia8	2.7766	1.0	1.0
181970	WZ_lnu2had_mll020	PowhegPythia8	4.8696	1.0	1.0
181971	WW	PowhegPythia8	52.44	1.0	1.0

Table A.6 VH samples used for the analysis, considered in SR only. The columns from left to right describe MC generation ID, the data set description, the MC generator used, the sample cross section in pb, the NLO/LO K-factor (multiplicative to the cross section) and the filter efficiency of the requested sample within the ATLAS simulation.

ID	Name	Generator	σ [pb]	K-factor	ε
161805	$W(\to \ell\nu)H(\to b\bar{b})$	Pythia8	0.7046	0.324	0.577
161827	$Z(\to \ell\ell)H(\to b\bar{b})$	Pythia8	0.38284	0.100971	0.577
161849	$Z(\to \nu\nu)H(\to b\bar{b})$	Pythia8	0.38284	0.2	0.577
161871	$tt(\to \ell\ell + X)H(\to b\bar{b})$	Pythia8	0.00557	1.0	1.0

Table A.7 Multi-jet samples used for the analysis. The columns from left to right describe MC generation ID, the data set description, the MC generator used, the sample cross section in pb, the NLO/LO K-factor (multiplicative to the cross section), the filter efficiency of the requested sample within the ATLAS simulation and the effective number of events for normalization.

ID	Name	Generator	σ [pb]	K-factor	ε
147910	JZ0W	Pythia8	72.85×10^9	1.0	0.98557
147911	JZ1W	Pythia8	72.85×10^9	1.0	0.000129
147912	JZ2W	Pythia8	26.359×10^6	1.0	0.003994
147913	JZ3W	Pythia8	544.19×10^3	1.0	0.001219
147914	JZ4W	Pythia8	6445.3	1.0	0.000708
147915	JZ5W	Pythia8	39.739	1.0	0.002152
147916	JZ6W	Pythia8	0.41609	1.0	0.004684
147917	JZ7W	Pythia8	0.040636	1.0	0.0146

Appendix B

Supplements to objects reconstruction

B.1 Topological cell cluster as input to the jet reconstruction

The principle signal definition for use in the jet reconstruction is the clusters of topologically connected calorimeter cell signals or "topo-clusters". Topo-clusters are three-dimensional energy blobs that represent the shower development of each particle entering the calorimeter. The topological clustering works by grouping neighbouring cells that have signals higher than the expected noise into a cluster. More specifically the cell signal significance, ζ_{cell}^{EM} is defined as [207]:

$$\zeta_{cell}^{EM} = \frac{E_{cell}^{EM}}{\sigma_{cell,noise}^{EM}} \tag{B.1}$$

where E_{cell}^{EM} is the cell signal and $\sigma_{cell,noise}^{EM}$ is the expected noise in the cell. Cluster growth starts at seed cells that have $|\zeta_{cell}^{EM}|$ above a large threshold, t_{seed}. Neighbouring cells that have $|\zeta_{cell}^{EM}|$ greater than a medium threshold, $t_{neighbour}$, are added to the cluster. Finally all direct neighbouring cells on the outer perimeter that have $|\zeta_{cell}^{EM}|$ above a low threshold, t_{cell}, are added to the cluster. The default values for t_{seed}, $t_{neighbour}$ and t_{cell} are 4, 2 and 0, respectively [207]. The higher thresholds for seeds and neighbours are meant for suppressing both electronics and pile-up noise. Note that $t_{cell} = 0$ means that all cells at the outer perimeter will end up in a cluster, regardless of their energies. This low threshold at the perimeter ensures that tails of showers (cells with signals that are close to the noise levels) are not discarded. This implicit noise suppression improves the energy and spatial resolutions in the presence of pile-up. In the end, a topo-cluster is defined as massless pseudo-particle with a four-momentum $(E, \overrightarrow{p})$. Its energy is equal to the energy sum of all the included cells. Whereas its reconstructed direction is defined as that of a unit vector originating from the geometric centre of the ATLAS coordinate system pointing to the energy-weighted topo-cluster barycenter in $\eta - \phi$ space.

Both E_{cell}^{EM} and $\sigma_{cell,noise}^{EM}$ in Equation B.1 are reconstructed in EM energy scale which measures the energy deposited by electrons and photons. This energy scale does not include any corrections for hadrons energy loss (e.g. including escaping energy carried by invisible

Table B.1 Variables used in the electron identification [204, 70, 116, 205, 70, 118].

Type	Description	Cut value
Acceptance	• Central region of the detector.	$\|\eta\| < 2.47$
Hadronic leakage	• The ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster. (for EM cluster with $\|\eta\| < 0.8$ and $\|\eta\| > 1.37$) .	η and E_T dependent cuts.
Front (strip) layer of the EM calo.	• The shower width, defined as: $\sqrt{(\sum E_i(i-i_{max})^2/(\sum E_i))}$, where i runs over all strips in a window of $\Delta\phi \times \Delta\eta = 0.2 \times 0.0625$, corresponding typically to 20 strips in η, and i_{max} is the index of the strip with the highest-energy. • The ratio of the energy difference between the largest and second largest energy deposits in the cluster over their energies sum.	η and E_T dependent cuts. η and E_T dependent cuts.
Middle layer of the EM calo.	• The ratio of energies in 3×7 cells over 7×7 cells centred at the electron cluster position. •The lateral shower width, defined as: $\sqrt{(\sum E_i\eta_i^2)/(\sum E_i) - ((\sum E_i\eta_i)/(\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells.	η and E_T dependent cuts. η and E_T dependent cuts.
Back layer of the EM calo.	• The ratio of the energy in the back layer to the total energy in the EM calorimeter η and E_T dependent cuts.	
Track quality	• The number of hits in the first layer of the pixel detector (for $\|\eta\| < 2.01$) • The number of hits in the pixel detector (for $\|\eta\| > 2.01$) • The number of hits in the silicon detectors • The track's transverse impact parameter	$N^{hits}_{1st\,pixel} \geq 1$ $N^{hits}_{pixel} \geq 2$ $N^{hits}_{pixel+SCT} \geq 7$ $\|d_0\| < 5$ mm
Track-cluster matching	•$\|\Delta\eta\|$ between the cluster position in the first layer and the extrapolated track	$\|\Delta\eta\| < 0.005$
TRT track quality	• The ratio of the hits on track that exceed the TRT high threshold (for signal collected at the TRT wire) to the total number of hits on track	η dependent cuts.

particles and energy required for breaking up nucleus). Because of this, EM particle (electrons or photons) will leave different signal in the calorimeter than hadronic particles (such as charge pions) even when they have identical energies, i.e $E^{EM}_{clus} > E^{HAD}_{clus}$. Thus, in order to correctly estimate the energy deposit by hadronic particles, appropriate local correction (cluster-by-cluster) to the hadronic-like cluster has to be applied. Furthermore, due to the intrinsic noise suppression scheme applied during the clustering sequence, a large amount of true signal may be lost. This effect is particularly relevant when the amount of pile-up event is large (causing increase in the cell-signal baseline fluctuations which in turn cause more of the energy tail of the shower profile to drop below the noise level). Different amount of correction need to be

Table B.2 Variables used in the photon identification [120, 206].

Type	Description	Cut value
Acceptance	• Central region of the detector.	$\|\eta\| < 2.37$ (excluding $1.37 < \|\eta\| < 1.52$)
Hadronic leakage	• The ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster. (for EM cluster with $\|\eta\| < 0.8$ and $\|\eta\| > 1.37$) .	η and E_T dependent cuts.
	• The ratio of E_T in whole hadronic calorimeter to E_T of the EM cluster (for EM cluster within $0.8 < \|\eta\| < 1.37$).	η and E_T dependent cuts.
Front (strip) layer of the EM calo.	• The shower width, defined as: $\sqrt{(\sum E_i(i-i_{max})^2/(\sum E_i))}$, where i runs over all strips in a window of $\Delta\phi \times \Delta\eta = 0.2 \times 0.0625$, corresponding typically to 20 strips in η, and i_{max} is the index of the strip with the highest-energy.	η and E_T dependent cuts.
	• The shower width calculated from three strips around the strip with maximum energy deposit.	η and E_T dependent cuts.
	• The energy outside the core of the three central strips but within seven strips divided by energy within the three central strips. It quantifies the lateral containment of the shower, along the η direction.	η and E_T dependent cuts.
	• The difference between the energy of the strip cell with the second largest energy and the energy in the strip cell with the lowest energy found between the largest and the second largest energy.	η and E_T dependent cuts.
	• The ratio of the energy difference between the largest and second largest energy deposits in the cluster over their energies sum.	η and E_T dependent cuts.
Middle layer of the EM calo.	• The ratio of energies in 3×7 cells over 7×7 cells centred at the electron cluster position.	η and E_T dependent cuts.
	• The ratio of energies in 3×3 cells over 3×7 cells centred at the electron cluster position.	η and E_T dependent cuts.
	•The lateral shower width, defined as: $\sqrt{(\sum E_i\eta_i^2)/(\sum E_i) - ((\sum E_i\eta_i)/(\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells.	η and E_T dependent cuts.

applied to EM and hadronic cluster to account for different shower profiles. On top of that, signal may also be lost due to energy lost in inactive (dead) or uninstrumented material in front of or between the calorimeter modules. This energy loss need to be recovered.

To address these three main sources of calorimeter signal inefficiencies, ATLAS uses the so-called "local hadronic cell weighting" (LCW) correction scheme [207]. The LCW scheme starts by classifying each topo-cluster according to the nature of the energy deposit—

HAD or EM, so that the most appropriate mix of EM and HAD correction can be applied. This classification takes advantages of the differences between the HAD and EM showers characteristics. The later are in general have early starting point in the calorimeter, relatively small shower-by-shower fluctuations in their shower development and compacter shower shape. More specifically, the probability for a topo-cluster to be of EM origin, $\mathscr{P}^{EM}$ can be determined by measuring the probability of detecting an EM-like cluster in bins of 4 observables, i.e the cluster direction, $|\eta_{clus}|$, the cluster energy, E_{clus} [MeV], the logarithmic depth of the shower center, $log_{10}(\lambda_{clus}$ [mm]) and the logarithmic energy-weighted average cell signal density minus the logarithmic cluster energy, $log_{10}(\rho_{clus}$ [MeV/mm^{-3}])-$log_{10}(E^{EM}_{clus}$ [MeV]). This probability is derived by using MC simulations of single charged and neutral pions at various incident energies and entering the calorimeters at various pseudorapidities. $\mathscr{P}^{EM}$ in each variable bin, $ijkl$, is defined as [207, 208]

$$\mathscr{P}^{EM}_{ijkl} = \frac{f^{\pi^0}_{ijkl}}{f^{\pi^0}_{ijkl} + 2f^{\pi^\pm}_{ijkl}} \tag{B.2}$$

where $f^{\pi^0,\pm}_{ijkl}$ is the fraction of neutral or charged pions in a given ij $(E^{EM}_{clus}, \eta_{clus})$ bin from the simulation. Any cluster falls in a phase-space bin with $\mathscr{P}^{EM}_{ijkl} < 0.5$ is classified as hadronic cluster.

For each type of correction κ, a weight w^{κ}_{cell} for each cell is separately determined for EM and hadronic showers, binned in terms of the 4 observables. At any of the correction steps the signal weights, w^{cal}_{cell}, applied to cell signals in the topo-cluster are calculated as:

$$w^{cal}_{cell} = \mathscr{P}^{EM}_{clus} \cdot w^{EM-\kappa}_{cell} + (1 - \mathscr{P}^{EM}_{clus}) \cdot w^{HAD-\kappa}_{cell} \tag{B.3}$$

where the weights $w^{EM-\kappa}_{cell}$ and $w^{HAD-\kappa}_{cell}$ represent the EM or HAD correction factors applied to the cell signal. Thus the cumulative weight w^{LCW}_{cell} for each cell is then given by:

$$w^{LCW}_{cell} = \prod_{\kappa \in \{cal, ooc, dm\}} \left[\mathscr{P}^{EM}_{clus} \cdot w^{EM-\kappa}_{cell} + (1 - \mathscr{P}^{EM}_{clus}) \cdot w^{HAD-\kappa}_{cell} \right] \tag{B.4}$$

where the subscript $\{cal, occ, dm\}$ corresponds to the hadronic correction, correction for out-of-cluster signal losses and correction for dead material respectively. The final LCW cluster energy, E^{LCW}_{clus} can then be calculated as:

$$E^{LCW}_{clus} = \sum_{i \in clus} w^{LCW}_{cell,i} E^{EM}_{cell,i} \tag{B.5}$$

After these corrections, the clusters' four-momentum are recalculated.

B.2 Jet energy scale (JES) corrections

The LCW correction scheme described previously is performed on a cluster-by-cluster basis without any particular assumption about the kind of object. This means that after the jets are formed, they are also corrected to this scale. In order to restore the JES to that of jets reconstructed from stable simulated particles (MC truth particle level), additional jet-by-jet corrections need to be applied. The JES correction [126, 209, 129] procedure used in this analysis is shown schematically in Figure B.1. This procedure consists of the following steps:

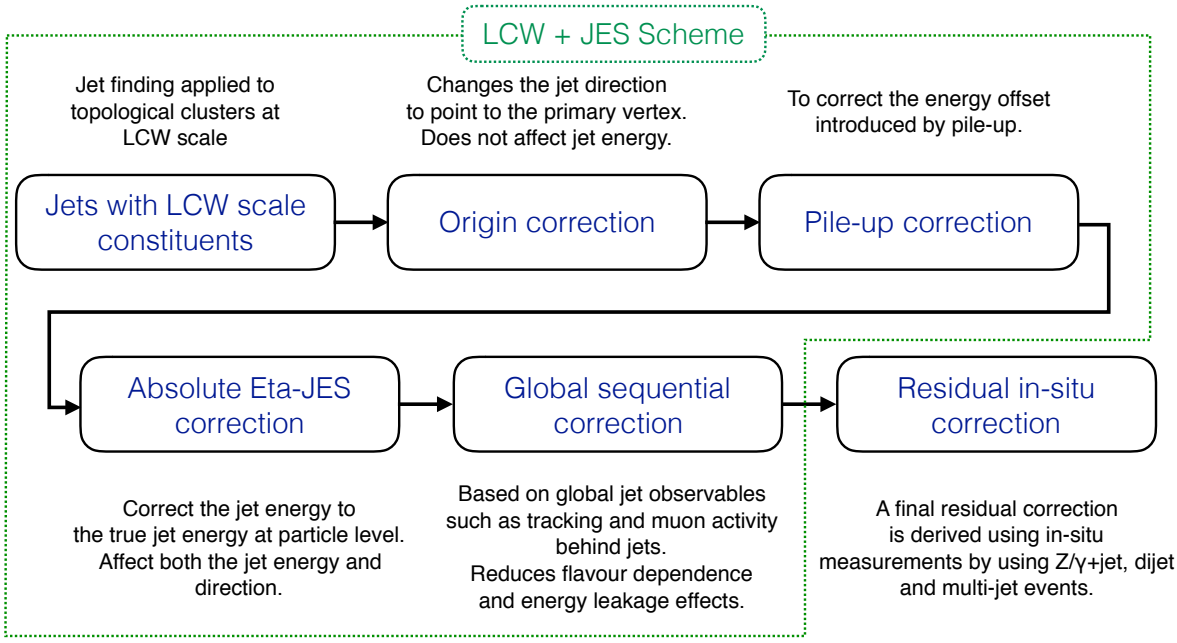


Fig. B.1 The correction steps for LCW jets.

1. First the direction of the jet is corrected such that it points to the primary vertex of the interaction instead of the geometrical centre of ATLAS detector. The kinematic observables of each topo-cluster are recalculated relative to the primary vertex. This correction improves the angular resolution. It leaves the jet energy unchanged.

2. There are two ways in which pile-up can contribute energy to an event: either by forming new clusters, or by overlapping with signal from the triggering event. While the former case has already been partially accounted for during the formation of the topo-clusters, the latter still can contribute to the signal clusters if they overlap with other deposits which survive noise suppression. To mitigate this effect on the reconstructed jet, a jet area [137] based subtraction method [210] is employed. The general idea of this approach is to calculate the amount of p_{T} contribution from pile-up in a jet in MC simulation,

and subtract this offset from the reconstructed jet p_{T}. This method works by treating the pile-up as a uniform and diffuse background that add contribution to jets. Thus a average pile-up density, ρ, times the jet area in the $\eta - \phi$ plane, A, gives the amount of pile-up inside the jet. The effect of pile-up can then be removed by simply subtracting it away. Remaining small dependence of the jet p_{T} on pile-up is parameterised in terms of the number of primary vertices, N_{PV}, and the average number of interactions per bunch crossing, $\langle\mu\rangle$, such that both the residual in-time pile-up dependence, characterised by N_{PV} (at fixed $\langle\mu\rangle$) and out-of-time pile-up, characterised by $\langle\mu\rangle$ (at fixed N_{PV}) can be captured. The corrected p_{T} after pile-up subtraction is therefore given by

$$p_{\mathrm{T}}^{corrected} = p_{\mathrm{T}}^{LCW} - \rho \times A - \alpha(\eta)\,(N_{PV} - 1) - \beta(\eta)\langle\mu\rangle \tag{B.6}$$

where p_{T}^{LCW} is the jet p_{T} at the LCW topo-cluster scale, $\alpha(\eta)$ and $\beta(\eta)$ are the correction constants derived from MC.

3. The next correction is aiming at restoring the reconstructed jet energy, E_{jet}^{reco} to the energy of the MC truth jet[1], E_{jet}^{truth}. For each reconstructed calorimeter jet-truth jet pair (the pair is formed if $\Delta R(jet^{reco}, jet^{truth}) < 0.3$), the average jet energy response, defined as

$$R_{jet} = \langle E_{jet}^{reco} / E_{jet}^{truth} \rangle \tag{B.7}$$

in bins of the E_{jet}^{truth} and the original reconstructed jet pseudorapidity before the origin correction, η_{det}, is measured. Similarly for each $(E_{jet}^{truth}, \eta_{det})$-bin, the average reconstructed calorimeter jet energy, $E_{jet}^{reco,\ avg}$, defined as the mean of the E_{jet}^{reco} distribution, are determined. Then the correction functions, $C(E_{jet}^{reco})_k$ is derived using a fit to the $(E_{jet,\ k}^{reco,\ avg}, R_{jet,\ k})$ points for a given η_{det}-bin, k. The corrected jet energy, $E_{jet}^{reco,\ corr}$, is simply obtained by multiplying the uncorrected jet energy with $1/C$, as

$$E_{jet}^{reco,\ corr} = \frac{E_{jet}^{reco}}{C(E_{jet}^{reco})_k} \tag{B.8}$$

It is found that the reconstructed jet direction has a small tendency to be biased to the better instrumented calorimeter regions as topo-clusters are reconstructed with a higher energy with respect to poorly instrumented calorimeter regions (e.g. barrel-endcap transition region). This effect has to be accounted for. The η-correction is derived as the average difference $\Delta\eta = \eta_{truth} - \eta_{origin}$ in each (E_{truth}, η_{det})-bins. Here η_{origin} is jet η after the jet origin correction. $\Delta\eta$ is simply added back to the η_{origin} to get the corrected jet direction. The correction is generally very small $(\Delta\eta < 0.01)$ for most regions of the calorimeter but larger in the transition regions.

[1] Truth jets are built from stable particles produced by the fragmentation model in the physics generator. Stable particles are defined as those with a laboratory frame lifetime of about 10 picoseconds or more.

4. Finally a correction called the global sequential correction (GSC) [211] are applied to the jets. It is simply a multi-variate extension of the previous jet energy correction. The general idea is that any variable x that is correlated with the detector response to the jet can be used for GSC. Corrections are applied sequentially. For jet reconstructed with LCW topo-clusters, three such variables are used (in order):

   i. The number of tracks associated to a given jet, n_{trk}.

   ii. The p_{T}-weighted transverse width of the jet, W_{trk}, defined as the average $\Delta R(jet^{reco}, track^{i})$ distance weighted by the track p_{T}. It quantifies the transverse structure of the jet, which is sensitive to different jet flavours, i.e. gluon or quark initiated jet. A gluon initiated jet tends to have a wider transverse profile, more particles and a lower energy response than a jet originating from a light quark with the same p_{T} and η.

   iii. The number of muon segments behind non fully contained calorimeter jets, $N_{segments}$. It quantifies the activity in the muon chambers behind high energy jets that penetrates the full depth of the calorimeter. The reconstructed energies of these high-p_{T} unconfined calorimeter jets are smaller, causing more spread in the jet energy response and thus worse jet energy resolution.

   Similar to Step 3 above, a multiplicative jet energy correction function, $C(p^{reco}_{T,jet}$ or $E^{reco}_{jet}, x)_j$ is derived for each variable and for a given η_{det}-bin, k. However, in this case C is constructed using a two dimensional Gaussian kernel whose kernel-width parameters are chosen to capture the shape of the average jet response (R_{jet}) across η and $p_{T,jet}$ as well as to ensure that the average energy is not affected by the correction. The GSC corrected jet p_{T} (or jet energy) is given by the initial jet p_{T} multiplied by the product of the N corrections:

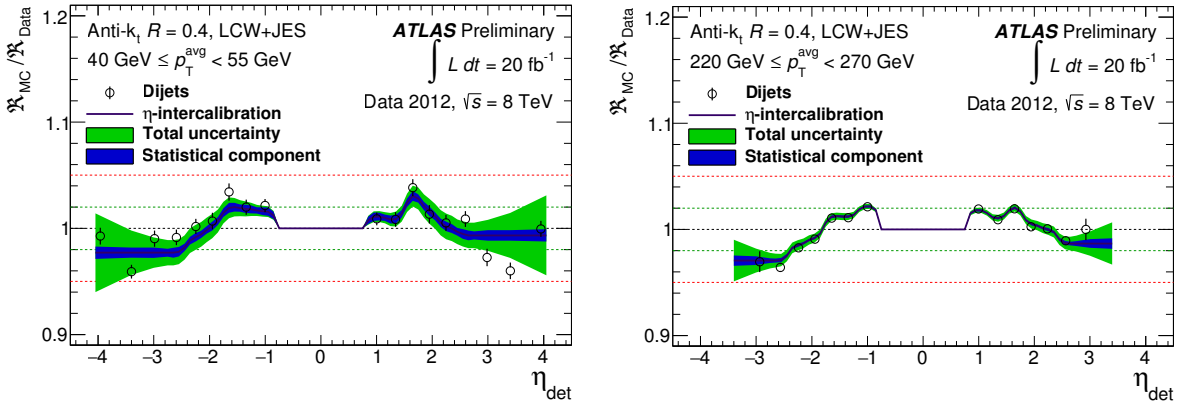
$$p^{reco,GSC}_{T,jet} = p^{reco}_{T,jet} \prod_{i=1}^{N} \frac{1}{C(p^{reco}_{T,jet},\, x)_k} \tag{B.9}$$

   After a successful application, the dependence of the response on the variable x is largely removed without changing the average energy. As a result the spread of the reconstructed jet energy is reduced, which in turn leads to improvement in the energy resolution.

The jet energy correction outlined above is dependent on the MC simulation to correct the response of the calorimeters. However, the calorimeters response to jet may not be fully simulated due to reasons such as the incorrect description of the amounts of dead material in front of the calorimeters in MC. Hence additional correction factor derived from real data are needed to ensure that the impact of mismodelling in the simulation is minimised.

This correction is termed η-intercorrection. It is is performed by using di-jet events [128]. The η_{det}-dependent correction factors are derived by comparing in dijet events the transverse momentum of a well-calibrated jet in the central reference region ($|\eta_{det}| < 0.8$) to a jet in the other regions under investigation. This method relies on the fact that at leading order

in QCD, the two jets are expected to have equal transverse momentum. Any imbalance in the transverse momentum would therefore be expected to arise from the mismodeling of the different jet responses in different calorimeter regions. This correction aims to remove any residual pseudorapidity difference in the jet response between the data and MC prediction following the MC based LCW+JES corrections. It quantifies the balance between a pair of jets in different detector regions without evaluating the absolute scale of either jet. Figure B.2 shows the relative jet response as a function of the η_{det} for the data and the MC simulations. The black solid line shows the derived η-intercorrection factors with the bands showing the uncertainty on this correction. The η-intercorrection factors are generally small (below 3%). After all these corrections, the resulting calorimeter jets built from LCW-scale topo-clusters are referred to as corrected with the LCW+JES scheme.



(a) (b)

Fig. B.2 Relative jet response as a function of η_{det} for jets reconstructed with anti-k_t algorithm with radius parameter, R = 0.4. The jets are corrected with the LCW+JES scheme. Results are shown separately for jets with average transverse momentum (of the two jets in the dijet system) range of $40 < p_T^{avg} < 55$ GeV and $220 < p_T^{avg} < 270$ GeV. Figure taken from Reference [129].

B.3 Jet energy scale and jet mass scale corrections for large-R jet

The flow chart of the JES plus jet mass scale (JMS) correction for large-R jet is illustrated in Figure B.3. Similar to the small-R jets, the large-R jets used in this analysis is corrected with the LCW+JES scheme described in the previous section. Nonetheless, there is no explicit origin offset correction as opposed to the small-R jets. Furthermore, there is also no explicit pile-up correction applied to the large-R jets as their contribution has been removed during the trimming procedure. In addition, no GSC is applied. The JES and η correction are derived

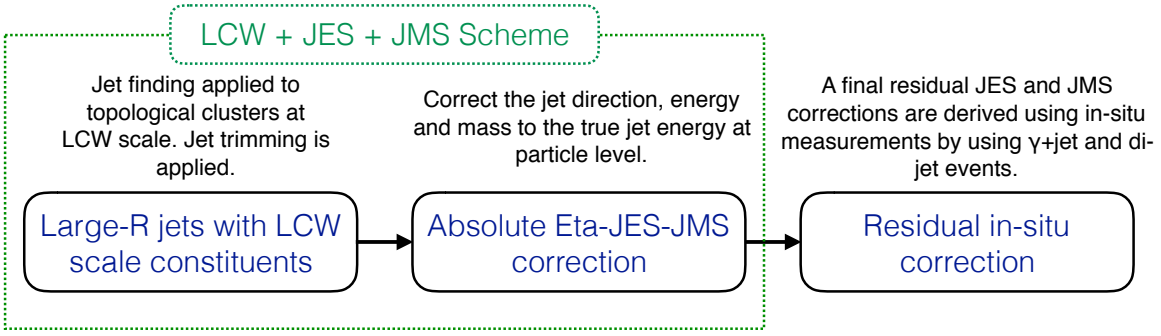


Fig. B.3 A schematic diagram depicting the JES+JMS correction procedure for large-R jet.

using similar strategies as described in Section B.2. Since the primary goal of the use of the large-R jet is to reconstruct accurately the jet mass, an explicit JMS correction [126] is performed on top of the standard energy and η correction. This JMS correction is based on the di-jet MC events. The procedure is the same as that employed for the JES correction. Figure B.4 shows the average jet energy response (see Equation B.7 for definition) and average jet mass response (same definition as in Equation B.7 but jet energy is replaced with jet mass) as a function of the jet's η_{det} in different jet energy bins before and after JES and JMS correction for trimmed large-R jets. Prior to the correction the energy and mass distribution show considerable structure due to the position of the different calorimeter boundaries. After the correction, the position dependence is clearly reduced. A uniform response across the full energy and η range is restored.

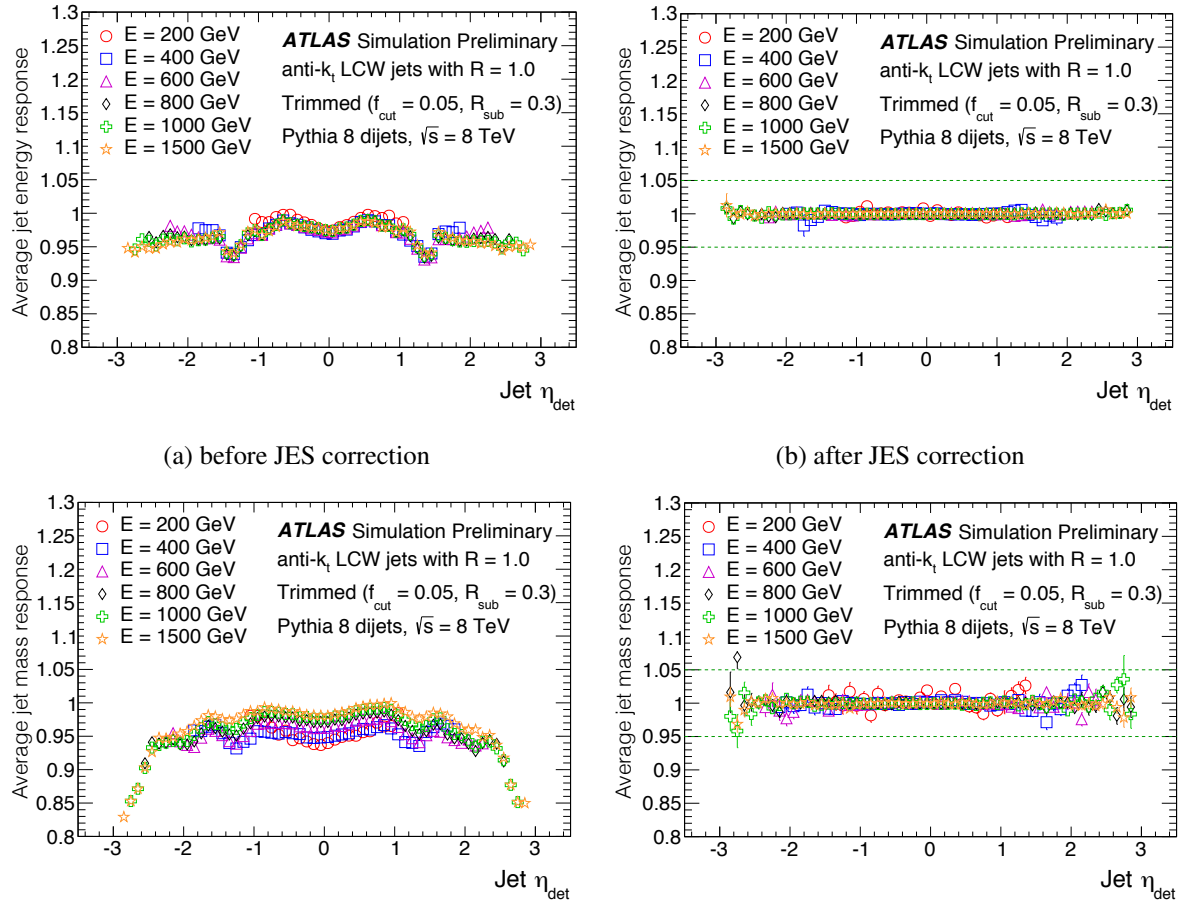
B.4 Jet quality requirements

Reconstructed jet candidates arising from high energy objects produced in a collision event are labelled as "good jet" while jet candidates coming from background processes are called "fake jet". The main sources of fake jets include:

i. Calorimeter noise[1]:

- Coherent noise due to electromagnetic interference (either pick-up from the on board digital activity, power supply or external sources) which affects coherently large groups of EM calorimeter cells.
- Sporadic noise burst or spike in the hadronic calorimeter where almost all of the jet energy come from just a few noisy calorimeter cells.

[1] While most of the noise is already identified and rejected by the data quality requirement, a small fraction of calorimeter noise remains undetected. The noise needs to be rejected by additional criteria.



(a) before JES correction

(b) after JES correction

(c) before JMS correction

(d) after JMS correction

Fig. B.4 Jet energy response of large-R jets as a function of the jet's η (relative to the geometric centre of the detector) before (a) and after (b) the JES correction and the jet mass response before (c) and after (d) the JMS correction. Figure adapted from Reference [129].

ii. Beam-induced backgrounds such as beam-gas and beam halo events.

iii. Photons produced by cosmic ray muons overlapping in-time with collision events.

Calorimeter instrumental noise can lead to fake energy deposits not due to particles in calorimeter cells, which can sometimes be reconstructed as fake jets. As energy deposits arising from real particles showering in the calorimeters produce a characteristic pulse shape, it can be used to separate ionisation signals from noise. By comparing the measured pulse with the expectation from simulation of the electronics response, a pulse quality of the LAr or HEC

calorimeter cells can be defined [130, 212]

$$\mathscr{Q}_{cell} = \sum_{j=1}^{N} \left(s_j - A g_j^{phys} \right)^2 \tag{B.10}$$

where A is the measured amplitude of the signal (computed with a number of sample cells, N), s_j is the amplitude (in ADC counts) of each sample pulse, j, and g_j^{phys} is the normalised predicted ionisation pulse shape. $\mathscr{Q}_{cell}$and several other jet-level quantities are used to discriminate noise from real energy deposits, they include

- $\langle Q \rangle$ — The normalised average jet quality, defined as the energy-squared weighted average of the pulse quality of the calorimeter cells in the jet.
- f_Q^{LAr} — The fraction of the energy in LAr calorimeter cells with poor signal shape quality ($Q_{cell} > 4000$).
- f_Q^{HEC} — The fraction of the energy in the HEC calorimeter cells with poor signal shape quality ($Q_{cell} > 4000$).
- $f_{LAr \text{ or } HEC}$ — The energy deposited in the calorimeter, divided by the total jet energy.
- E_{neg} — Energy of the jet originating from cells with negative energy due to the cell-signal baseline fluctuations introduced predominantly by pile-up and, to a lesser extent, by electronic noise.

Sporadic noise bursts in the HEC can produce large signal, typically contained in just a few calorimeter cells Jets reconstructed from these problematic cells usually have large f_{HEC}, f_Q^{HEC}, and $\langle Q \rangle$ values. Due to the capacitive coupling between channels, the neighbouring calorimeter cells will have an apparent negative energy E_{neg}. Likewise, fake jets stemming from the coherent noise in the LAr EM calorimeter are characterised by large f_{LAr}, f_Q^{LAr}, as well as $\langle Q \rangle$ values.

Table B.3 Variables used in the jet quality selection [130, 131].

Background type	Cut value
LAr EM coherent noise	$f_{LAr} > 0.95$ AND $\lvert f_Q^{LAr} \rvert > 0.8$ AND $\langle Q \rangle > 0.8$ AND $\lvert \eta \rvert < 2.8$
HEC noise burst	$f_{HEC} > 0.5$ AND $\lvert f_Q^{HEC} \rvert > 0.5$ AND $\langle Q \rangle > 0.8$ OR $E_{neg} > 60$ GeV
Beam-induced and cosmic muons	$f_{LAr} < 0.05$ AND $f_{chrg} < 0.05$ AND $\lvert \eta \rvert < 2$ OR $f_{LAr} < 0.05$ AND $\lvert \eta \rvert \geq 2$ OR $f_{max} > 0.99$ AND $\lvert \eta \rvert < 2$

To reject fake jets originating from beam-induced and cosmic muons background, the jet energy deposits in the direction of the shower development can be employed. Examples of discriminating variables are the maximum energy fraction in any single calorimeter layer (relative to the total jet energy), f_{max} as well as $f_{LAr \text{ or } HEC}$. In addition, the jet charged particle fraction, f_{chrg}, defined as the ratio of the scalar sum of the p_{T} of the tracks associated with the jet divided by jet p_{T}, is another powerful variable to discriminate collision jets from fake jets. This is because real jets contain charged hadrons are usually also reconstructed by the tracking system while fake jets typically have no associated tracks. The threshold values for these discriminating variables are summarised in Table B.3.

The jet quality selection efficiency is better than 99.8% for real jet with $p_{\mathrm{T}} > 20$ GeV while at the same time rejects most of the fake jets [130, 131]. Its performance in data is well reproduced by the MC simulation.

Appendix C

Cutflow and selection efficiency

C.1 Cutflow for background

Table C.1 The expected event yields for each background at each selection stage. The uncertainties include statistical error only.

	Diboson	Z+jets	SingleTop	Vh
0-lepton preselection	25179.2 ± 38.1	34419.7 ± 157.1	24820.4 ± 135.5	316.0 ± 1.5
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$	4122.4 ± 15.1	4315.1 ± 35.9	3252.8 ± 46.7	67.7 ± 0.7
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)\rvert > 1.0$	2061.7 ± 9.5	575.5 ± 25.8	792.2 ± 22.6	41.4 ± 0.5
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert < pi/2$	2028.6 ± 9.4	454.2 ± 25.5	743.2 ± 22.0	40.9 ± 0.5
$n_J \geq 1$	282.5 ± 3.3	29.4 ± 1.2	71.9 ± 6.9	6.4 ± 0.2
$J^1 p_T > 350\ GeV$	145.6 ± 2.4	12.6 ± 0.8	36.4 ± 4.7	3.4 ± 0.1
$n_{j^{trk}} \geq 1$	145.3 ± 2.3	12.6 ± 0.8	36.4 ± 4.7	3.3 ± 0.1
$n_{j^{trk}} \geq 2$	83.6 ± 1.8	3.7 ± 0.3	16.7 ± 2.8	2.4 ± 0.1
$n_{j_b^{trk}} \geq 1$	16.3 ± 0.7	0.5 ± 0.1	8.1 ± 2.0	2.1 ± 0.1
$n_{j_b^{trk}} == 2$	3.0 ± 0.2	0.1 ± 0.1	1.2 ± 0.6	1.2 ± 0.1
$90\ GeV < m_J < 150\ GeV$	1.0 ± 0.1	0.1 ± 0.1	0.3 ± 0.3	1.1 ± 0.1
$E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$	0.9 ± 0.1	0.0 ± 0.0	0.0 ± 0.0	1.0 ± 0.1
$E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$	0.6 ± 0.1	0.0 ± 0.0	0.0 ± 0.0	0.6 ± 0.1

	W+jets	$t\bar{t}$	$Z\nu\nu$	QCD multi-jet
0-lepton preselection	1298683.2 ± 762.9	139522.5 ± 185.3	116886.2 ± 187.4	-
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$	118012.8 ± 171.4	17222.0 ± 63.1	116886.1 ± 187.4	-
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)\rvert > 1.0$	55277.7 ± 130.2	3524.8 ± 29.2	93917.6 ± 167.8	-
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert < pi/2$	53534.7 ± 128.3	3397.2 ± 28.6	90505.2 ± 164.8	-
$n_J \geq 1$	5654.3 ± 28.4	250.1 ± 7.3	10935.3 ± 60.5	-
$J^1 p_T > 350\ GeV$	2532.1 ± 16.4	87.8 ± 4.3	5113.5 ± 41.9	-
$n_{j^{trk}} \geq 1$	2513.0 ± 15.0	87.8 ± 4.3	5107.7 ± 41.9	-
$n_{j^{trk}} \geq 2$	700.1 ± 7.2	69.7 ± 3.8	1383.3 ± 21.8	-
$n_{j_b^{trk}} \geq 1$	68.7 ± 2.0	42.6 ± 2.9	190.3 ± 8.1	-
$n_{j_b^{trk}} == 2$	5.3 ± 0.5	4.5 ± 0.9	16.0 ± 2.4	-
$90\ GeV < m_J < 150\ GeV$	1.7 ± 0.2	1.1 ± 0.5	8.5 ± 1.7	-
$E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$	1.4 ± 0.2	0.8 ± 0.4	7.0 ± 1.6	$< 0.0 \pm 0.04$
$E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$	0.8 ± 0.2	0.6 ± 0.3	5.2 ± 1.4	$< 0.0 \pm 0.04$

	Total Bkgds
0-lepton preselection	1639827.4 ± 834.2
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$	263878.8 ± 268.6
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)\rvert > 1.0$	156190.8 ± 217.4
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert < pi/2$	150704.1 ± 213.7
$n_J \geq 1$	17229.9 ± 67.7
$J^1 p_T > 350\ GeV$	7931.3 ± 45.5
$n_{j^{trk}} \geq 1$	7906.1 ± 45.0
$n_{trkj} \geq 2$	2259.5 ± 23.5
$n_{j_b^{trk}} \geq 1$	328.6 ± 9.1
$n_{j_b^{trk}} == 2$	31.3 ± 2.7
$90\ GeV < m_J < 150\ GeV$	13.7 ± 1.8
$E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$	11.2 ± 1.7
$E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$	7.7 ± 1.4

C.2 Cutflow for signals

	mx1_xgxFhDh	mx65_xgxFhDh	mx100_xgxFhDh	mx500_xgxFhDh	mx1000_xgxFhDh
Intial	202.8 ± 0.9	202.8 ± 0.9	202.8 ± 0.9	202.8 ± 0.9	202.8 ± 0.9
0-lepton preselection	174.3 ± 0.9	175.3 ± 0.9	175.8 ± 0.9	180.2 ± 0.9	181.7 ± 0.9
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$	154.4 ± 0.8	156.6 ± 0.8	157.9 ± 0.8	168.9 ± 0.8	172.2 ± 0.9
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i_{akt4})\rvert > 1.0$	125.6 ± 0.7	127.2 ± 0.7	127.9 ± 0.7	135.9 ± 0.8	140.2 ± 0.8
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert < pi/2$	125.6 ± 0.7	127.2 ± 0.7	127.9 ± 0.7	135.8 ± 0.8	140.2 ± 0.8
$n_{j_{akt10Trimmed}} \geq 1$	96.2 ± 0.6	98.1 ± 0.6	99.9 ± 0.6	114.8 ± 0.7	122.0 ± 0.7
$j^0_{akt10Trimmed} p_T > 350\ GeV$	84.8 ± 0.6	86.5 ± 0.6	88.6 ± 0.6	105.4 ± 0.7	113.1 ± 0.7
$n_{trkj} \geq 1$	84.8 ± 0.6	86.4 ± 0.6	88.6 ± 0.6	105.4 ± 0.7	113.1 ± 0.7
$n_{trkj} \geq 2$	57.4 ± 0.5	57.9 ± 0.5	58.8 ± 0.5	65.8 ± 0.5	68.7 ± 0.5
$n_{btrkj} \geq 1$	51.2 ± 0.5	51.8 ± 0.5	52.6 ± 0.5	58.4 ± 0.5	60.9 ± 0.5
$n_{btrkj} == 2$	28.3 ± 0.3	28.2 ± 0.3	28.5 ± 0.3	31.1 ± 0.4	32.1 ± 0.4
$90\ GeV < m_J < 150\ GeV$	24.8 ± 0.3	24.6 ± 0.3	24.6 ± 0.3	26.6 ± 0.3	27.9 ± 0.3
$E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$	24.6 ± 0.3	24.5 ± 0.3	24.5 ± 0.3	26.6 ± 0.3	27.9 ± 0.3
$E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$	21.6 ± 0.3	21.5 ± 0.3	21.6 ± 0.3	24.2 ± 0.3	25.6 ± 0.3

Table C.2 xgxFhDh signal yield (signal cross section scaled to 10 fb). The uncertainties include statistical errors only.

	mx1_xdxhDh	mx65_xdxhDh	mx100_xdxhDh	mx500_xdxhDh	mx1000_xdxhDh
Intial	203.5 ± 0.9	202.7 ± 0.9	203.0 ± 0.9	202.4 ± 1.0	203.0 ± 0.9
0-lepton preselection	59.4 ± 0.5	135.4 ± 0.8	148.5 ± 0.8	173.7 ± 1.0	176.8 ± 0.9
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$	22.6 ± 0.3	83.1 ± 0.6	101.7 ± 0.7	155.6 ± 0.9	163.7 ± 0.8
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i_{akt4})\rvert > 1.0$	17.6 ± 0.3	62.0 ± 0.5	75.9 ± 0.6	101.1 ± 0.7	100.7 ± 0.7
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert < pi/2$	17.6 ± 0.3	61.9 ± 0.5	75.9 ± 0.6	101.1 ± 0.7	100.6 ± 0.7
$n_{j_{akt10Trimmed}} \geq 1$	5.9 ± 0.2	27.8 ± 0.3	37.5 ± 0.4	77.1 ± 0.6	83.5 ± 0.6
$j^0_{akt10Trimmed} p_T > 350\ GeV$	4.0 ± 0.1	20.2 ± 0.3	27.9 ± 0.3	67.5 ± 0.6	75.6 ± 0.6
$n_{trkj} \geq 1$	4.0 ± 0.1	20.2 ± 0.3	27.9 ± 0.3	67.4 ± 0.6	75.6 ± 0.6
$n_{trkj} \geq 2$	3.1 ± 0.1	15.3 ± 0.3	20.9 ± 0.3	47.5 ± 0.5	51.4 ± 0.5
$n_{btrkj} \geq 1$	2.8 ± 0.1	13.6 ± 0.2	18.3 ± 0.3	41.1 ± 0.5	44.1 ± 0.4
$n_{btrkj} == 2$	1.6 ± 0.1	7.4 ± 0.2	10.0 ± 0.2	22.1 ± 0.3	22.6 ± 0.3
$90\ GeV < m_J < 150\ GeV$	1.4 ± 0.1	6.2 ± 0.2	8.5 ± 0.2	17.7 ± 0.3	17.6 ± 0.3
$E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$	1.4 ± 0.1	6.2 ± 0.2	8.3 ± 0.2	17.5 ± 0.3	17.4 ± 0.3
$E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$	1.0 ± 0.1	4.7 ± 0.1	6.5 ± 0.2	15.4 ± 0.3	15.9 ± 0.3

Table C.3 xdxhDh signal yield (signal cross section scaled to 10 fb). The uncertainties include statistical error only.

	mx1_xxhhg5	mx65_xxhhg5	mx100_xxhhg5	mx500_xxhhg5	mx1000_xxhhg5
Intial	202.8 ± 0.9	202.8 ± 0.9	202.8 ± 0.9	202.8 ± 0.9	202.8 ± 0.9
0-lepton preselection	141.6 ± 0.8	146.6 ± 0.8	157.1 ± 0.8	175.7 ± 0.9	179.2 ± 0.9
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$	87.9 ± 0.6	95.9 ± 0.6	117.9 ± 0.7	161.1 ± 0.8	168.6 ± 0.8
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^{i}_{akt4})\rvert > 1.0$	58.5 ± 0.5	63.1 ± 0.5	76.8 ± 0.6	94.8 ± 0.6	95.9 ± 0.6
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert < pi/2$	58.5 ± 0.5	63.1 ± 0.5	76.8 ± 0.6	94.7 ± 0.6	95.8 ± 0.6
$n_{j_{akt10Trimmed}} \geq 1$	24.9 ± 0.3	28.6 ± 0.3	41.6 ± 0.4	75.7 ± 0.6	82.4 ± 0.6
$j^{0}_{akt10Trimmed} p_T > 350 GeV$	17.3 ± 0.3	20.4 ± 0.3	31.4 ± 0.4	67.5 ± 0.5	76.1 ± 0.6
$n_{trkj} \geq 1$	17.3 ± 0.3	20.4 ± 0.3	31.4 ± 0.4	67.4 ± 0.5	76.1 ± 0.6
$n_{trkj} \geq 2$	12.9 ± 0.2	15.2 ± 0.3	23.3 ± 0.3	47.3 ± 0.4	50.8 ± 0.5
$n_{btrkj} \geq 1$	11.2 ± 0.2	13.1 ± 0.2	20.2 ± 0.3	40.6 ± 0.4	43.3 ± 0.4
$n_{btrkj} == 2$	6.5 ± 0.2	7.2 ± 0.2	11.3 ± 0.2	21.4 ± 0.3	21.8 ± 0.3
$90\ GeV < m_J < 150\ GeV$	5.2 ± 0.1	5.9 ± 0.2	9.3 ± 0.2	16.5 ± 0.3	16.6 ± 0.3
$E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$	5.0 ± 0.1	5.7 ± 0.2	9.1 ± 0.2	16.2 ± 0.3	16.4 ± 0.3
$E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$	3.8 ± 0.1	4.3 ± 0.1	7.2 ± 0.2	14.4 ± 0.2	15.1 ± 0.2

Table C.4 xxhhg5 signal yield (signal cross section scaled to 10 fb). The uncertainties include statistical error only.

	mx1_xxhh	mx65_xxhh	mx100_xxhh	mx500_xxhh	mx1000_xxhh
Intial	202.8 ± 0.9	202.8 ± 0.9	202.8 ± 0.9	202.8 ± 0.9	202.8 ± 0.9
0-lepton preselection	141.5 ± 0.8	144.4 ± 0.8	154.4 ± 0.8	174.6 ± 0.9	178.5 ± 0.9
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$	87.9 ± 0.6	92.9 ± 0.6	111.4 ± 0.7	159.7 ± 0.8	168.1 ± 0.8
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^{i}_{akt4})\rvert > 1.0$	58.5 ± 0.5	62.3 ± 0.5	73.0 ± 0.6	94.2 ± 0.6	96.6 ± 0.6
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert < pi/2$	58.5 ± 0.5	62.3 ± 0.5	72.9 ± 0.6	94.1 ± 0.6	96.5 ± 0.6
$n_{j_{akt10Trimmed}} \geq 1$	24.1 ± 0.3	26.5 ± 0.3	37.2 ± 0.4	74.5 ± 0.6	82.6 ± 0.6
$j^{0}_{akt10Trimmed} p_T > 350 GeV$	16.8 ± 0.3	18.8 ± 0.3	27.5 ± 0.3	66.2 ± 0.5	76.1 ± 0.6
$n_{trkj} \geq 1$	16.8 ± 0.3	18.8 ± 0.3	27.5 ± 0.3	66.2 ± 0.5	76.1 ± 0.6
$n_{trkj} \geq 2$	12.3 ± 0.2	14.1 ± 0.2	20.6 ± 0.3	46.8 ± 0.4	50.6 ± 0.5
$n_{btrkj} \geq 1$	10.6 ± 0.2	12.2 ± 0.2	17.7 ± 0.3	40.0 ± 0.4	42.9 ± 0.4
$n_{btrkj} == 2$	5.7 ± 0.2	6.6 ± 0.2	9.8 ± 0.2	21.0 ± 0.3	22.0 ± 0.3
$90\ GeV < m_J < 150\ GeV$	4.7 ± 0.1	5.5 ± 0.1	8.0 ± 0.2	16.5 ± 0.3	16.5 ± 0.3
$E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$	4.5 ± 0.1	5.3 ± 0.1	7.7 ± 0.2	16.2 ± 0.3	16.3 ± 0.3
$E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$	3.3 ± 0.1	4.0 ± 0.1	6.1 ± 0.2	14.5 ± 0.2	14.9 ± 0.2

Table C.5 xxhh signal yield (signal cross section scaled to 10 fb). The uncertainties include statistical error only.

C.3 Signal selection efficiency

Table C.6 Absolute and relative signal selection efficiency for mono-H operator xgxFhDh with different DM mass (m_χ =1-1000 GeV).

	mx1_xgxFhDh		mx65_xgxFhDh		mx100_xgxFhDh		mx500_xgxFhDh		mx1000_xgxFhDh	
	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.
0-lepton preselection	0.86	0.86	0.86	0.86	0.87	0.87	0.89	0.89	0.90	0.90
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$	0.76	0.89	0.77	0.89	0.78	0.90	0.83	0.94	0.85	0.95
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)\rvert > 1.0$	0.62	0.81	0.63	0.81	0.63	0.81	0.67	0.81	0.69	0.82
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert < pi/2$	0.62	1.00	0.63	1.00	0.63	1.00	0.67	1.00	0.69	1.00
$n_J \geq 1$	0.48	0.77	0.48	0.77	0.49	0.78	0.57	0.85	0.60	0.87
$p_T^{J_1} > 350\ GeV$	0.42	0.88	0.43	0.88	0.44	0.89	0.52	0.92	0.56	0.93
$n_{j^{trk}} \geq 1$	0.42	1.00	0.43	1.00	0.44	1.00	0.52	1.00	0.56	1.00
$n_{j^{trk}} \geq 2$	0.28	0.68	0.29	0.67	0.29	0.66	0.32	0.62	0.34	0.61
$n_{j_b^{trk}} \geq 1$	0.25	0.89	0.26	0.89	0.26	0.89	0.29	0.89	0.30	0.89
$n_{j_b^{trk}} == 2$	0.14	0.55	0.14	0.55	0.14	0.54	0.15	0.53	0.16	0.53
$90\ GeV < m_{J^1} < 150\ GeV$	0.12	0.88	0.12	0.87	0.12	0.86	0.13	0.86	0.14	0.87
$E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$	0.12	0.99	0.12	1.00	0.12	1.00	0.13	1.00	0.14	1.00
$E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$	0.11	0.88	0.11	0.88	0.11	0.88	0.12	0.91	0.13	0.92

Table C.7 Absolute and relative signal selection efficiency for mono-H operator xdxhDh with different DM mass (m_χ =1-1000 GeV).

	mx1_xdxhDh		mx65_xdxhDh		mx100_ xdxhDh		mx500_xdxhDh		mx1000_xdxhDh	
	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.
0-lepton preselection	0.29	0.29	0.67	0.67	0.73	0.73	0.86	0.86	0.87	0.87
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$	0.11	0.38	0.41	0.61	0.50	0.68	0.77	0.90	0.81	0.93
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)\rvert > 1.0$	0.09	0.78	0.31	0.75	0.37	0.75	0.50	0.65	0.50	0.62
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert < pi/2$	0.09	1.00	0.31	1.00	0.37	1.00	0.50	1.00	0.50	1.00
$n_J \geq 1$	0.03	0.34	0.14	0.45	0.19	0.50	0.38	0.76	0.41	0.83
$p_T^{J_1} > 350\ GeV$	0.02	0.68	0.10	0.73	0.14	0.74	0.33	0.87	0.37	0.91
$n_{j^{trk}} \geq 1$	0.02	1.00	0.10	1.00	0.14	1.00	0.33	1.00	0.37	1.00
$n_{j^{trk}} \geq 2$	0.02	0.77	0.08	0.76	0.10	0.75	0.24	0.71	0.25	0.68
$n_{j_b^{trk}} \geq 1$	0.01	0.89	0.07	0.88	0.09	0.88	0.20	0.86	0.22	0.86
$n_{j_b^{trk}} == 2$	0.01	0.58	0.04	0.55	0.05	0.55	0.11	0.54	0.11	0.51
$90\ GeV < m_{J^1} < 150\ GeV$	0.01	0.89	0.03	0.84	0.04	0.84	0.09	0.80	0.09	0.78
$E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$	0.01	0.98	0.03	0.99	0.04	0.98	0.09	0.99	0.09	0.99
$E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$	0.00	0.69	0.02	0.76	0.03	0.79	0.08	0.88	0.08	0.91

Table C.8 Absolute and relative signal selection efficiency for mono-H operator xxhhg5 with different DM mass (m_χ =1-1000 GeV).

	mx1_xxhhg5		mx65_xxhhg5		mx100_xxhhg5		mx500_xxhhg5		mx1000_xxhhg5	
	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.
0-lepton preselection	0.70	0.70	0.72	0.72	0.77	0.77	0.87	0.87	0.88	0.88
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$	0.43	0.62	0.47	0.65	0.58	0.75	0.79	0.92	0.83	0.94
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)\rvert > 1.0$	0.29	0.67	0.31	0.66	0.38	0.65	0.47	0.59	0.47	0.57
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert < pi/2$	0.29	1.00	0.31	1.00	0.38	1.00	0.47	1.00	0.47	1.00
$n_J \geq 1$	0.12	0.43	0.14	0.45	0.21	0.54	0.37	0.80	0.41	0.86
$p_T^{J_1} > 350\ GeV$	0.09	0.69	0.10	0.71	0.16	0.76	0.33	0.89	0.38	0.92
$n_{j^{trk}} \geq 1$	0.09	1.00	0.10	1.00	0.16	1.00	0.33	1.00	0.38	1.00
$n_{j^{trk}} \geq 2$	0.06	0.75	0.08	0.74	0.12	0.74	0.23	0.70	0.25	0.67
$n_{j_b^{trk}} \geq 1$	0.06	0.87	0.07	0.87	0.10	0.87	0.20	0.86	0.21	0.85
$n_{j_b^{trk}} == 2$	0.03	0.58	0.04	0.55	0.06	0.56	0.11	0.53	0.11	0.50
$90\ GeV < m_{J^1} < 150\ GeV$	0.03	0.81	0.03	0.82	0.05	0.83	0.08	0.77	0.08	0.76
$E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$	0.02	0.96	0.03	0.96	0.04	0.97	0.08	0.98	0.08	0.99
$E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$	0.02	0.76	0.02	0.75	0.04	0.79	0.07	0.89	0.07	0.92

Table C.9 Absolute and relative signal selection efficiency for mono-H operator xxhh with different DM mass (m_χ =1-1000 GeV).

	mx1_xxhh		mx65_xxhh		mx100_xxhh		mx500_xxhh		mx1000_xxhh	
	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.	Abs.	Rel.
0-lepton preselection	0.70	0.70	0.71	0.71	0.76	0.76	0.86	0.86	0.88	0.88
$E_{\mathrm{T}}^{\mathrm{miss}} > 200$	0.43	0.62	0.46	0.64	0.55	0.72	0.79	0.91	0.83	0.94
$\lvert\Delta\phi_{min}(E_{\mathrm{T}}^{\mathrm{miss}}, j^i)\rvert > 1.0$	0.29	0.67	0.31	0.67	0.36	0.66	0.47	0.59	0.48	0.58
$\lvert\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})\rvert < pi/2$	0.29	1.00	0.31	1.00	0.36	1.00	0.47	1.00	0.48	1.00
$n_J \geq 1$	0.12	0.41	0.13	0.43	0.18	0.51	0.37	0.79	0.41	0.86
$p_T^{J_1} > 350\ GeV$	0.08	0.70	0.09	0.71	0.14	0.74	0.33	0.89	0.38	0.92
$n_{j^{trk}} \geq 1$	0.08	1.00	0.09	1.00	0.14	1.00	0.33	1.00	0.38	1.00
$n_{j^{trk}} \geq 2$	0.06	0.73	0.07	0.75	0.10	0.75	0.23	0.71	0.25	0.67
$n_{j_b^{trk}} \geq 1$	0.05	0.86	0.06	0.87	0.09	0.86	0.20	0.85	0.21	0.85
$n_{j_b^{trk}} == 2$	0.03	0.54	0.03	0.55	0.05	0.55	0.10	0.53	0.11	0.51
$90\ GeV < m_{J^1} < 150\ GeV$	0.02	0.82	0.03	0.83	0.04	0.82	0.08	0.79	0.08	0.75
$E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$	0.02	0.96	0.03	0.97	0.04	0.97	0.08	0.98	0.08	0.98
$E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$	0.02	0.72	0.02	0.75	0.03	0.79	0.07	0.89	0.07	0.92

Appendix D

Acceptance and Selection Efficiency

As described in Section 8.3 we may compute a specific BSM production cross section by evaluating the total efficiency $(A \times \varepsilon)^i$ of the analysis selection. We cab then set a limit on:

$$\sigma^i_{BSM} = \frac{\sigma_{vis}}{(A \times \varepsilon)^i} \tag{D.1}$$

The product $(A \times \varepsilon)$ potentially contains both theoretical and experimental uncertainties, and care must be taken to handle correlations between these systematics and those of the SM Higgs yield prediction.

If one is not able to compute the fully-reconstructed total selection efficiency, the fiducial cross section $\sigma_{fid} = (\sigma_{BSM} \times A)$ may be used instead. Because the fiducial cross section already has the experimental uncertainty from reconstruction efficiency folded in, one only needs to calculate the parton-level fiducial acceptance A:

$$\sigma^i_{BSM} = \frac{\sigma_{fid}}{A^i} \tag{D.2}$$

However, care must be taken to ensure that the reconstruction efficiency used in the definition of the fiducial cross section is valid for the model under consideration.

The fiducial acceptance and total efficiencies for each operator considered in this study are listed in Table D.1 and Table D.2.

Signal Model	Acceptance*Eff (%)	Acceptance (%)	Efficiency (%)
mx1_xdxhDh	0.69± 0.04	2.00± 0.06	34.28± 2.16
mx65_xdxhDh	3.04± 0.08	9.57± 0.14	31.72± 0.94
mx100_xdxhDh	4.10± 0.09	12.83± 0.16	31.95± 0.81
mx500_xdxhDh	8.61± 0.15	21.71± 0.21	39.66± 0.78
mx1000_xdxhDh	8.59± 0.13	28.24± 0.24	30.41± 0.53
mx1_xgxFhDh	12.14± 0.16	35.07± 0.26	34.62± 0.52
mx65_xgxFhDh	12.10± 0.16	35.42± 0.27	34.17± 0.51
mx100_xgxFhDh	12.08± 0.16	35.98± 0.27	33.57± 0.50
mx500_xgxFhDh	13.10± 0.16	39.82± 0.28	32.91± 0.47
mx1000_xgxFhDh	13.74± 0.17	41.35± 0.29	33.24± 0.46
mx1_xxhhg5	2.48± 0.07	7.82± 0.13	31.67± 1.04
mx65_xxhhg5	2.79± 0.07	9.32± 0.14	29.93± 0.92
mx100_xxhhg5	4.48± 0.09	13.86± 0.17	32.30± 0.79
mx500_xxhhg5	8.00± 0.13	26.36± 0.23	30.36± 0.55
mx1000_xxhhg5	8.10± 0.13	27.70± 0.24	29.23± 0.52
mx1_xxhh	2.21± 0.07	7.65± 0.12	28.87± 0.99
mx65_xxhh	2.62± 0.07	8.45± 0.13	31.04± 0.98
mx100_xxhh	3.82± 0.09	12.18± 0.16	31.36± 0.83
mx500_xxhh	8.01± 0.13	26.10± 0.23	30.67± 0.56
mx1000_xxhh	8.02± 0.13	27.54± 0.23	29.10± 0.52

Table D.1 Fiducial acceptances and selection efficiencies for EFT for $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV.

Signal Model	Acceptance*Eff (%)	Acceptance (%)	Efficiency (%)
mx1_xdxhDh	0.47± 0.03	1.19± 0.05	39.57± 3.06
mx65_xdxhDh	2.32± 0.07	6.22± 0.11	37.26± 1.28
mx100_xdxhDh	3.22± 0.08	8.81± 0.13	36.52± 1.06
mx500_xdxhDh	7.58± 0.14	18.44± 0.19	41.11± 0.86
mx1000_xdxhDh	7.86± 0.13	25.22± 0.22	31.16± 0.57
mx1_xgxFhDh	10.63± 0.15	29.01± 0.24	36.65± 0.59
mx65_xgxFhDh	10.62± 0.15	29.61± 0.24	35.88± 0.58
mx100_xgxFhDh	10.63± 0.15	30.16± 0.25	35.23± 0.56
mx500_xgxFhDh	11.92± 0.15	35.02± 0.26	34.04± 0.51
mx1000_xgxFhDh	12.62± 0.16	36.93± 0.27	34.16± 0.50
mx1_xxhhg5	1.86± 0.06	5.03± 0.10	37.07± 1.43
mx65_xxhhg5	2.11± 0.07	6.13± 0.11	34.39± 1.23
mx100_xxhhg5	3.54± 0.08	9.90± 0.14	35.76± 0.99
mx500_xxhhg5	7.12± 0.12	22.85± 0.21	31.18± 0.60
mx1000_xxhhg5	7.43± 0.12	25.17± 0.22	29.51± 0.55
mx1_xxhh	1.61± 0.06	4.74± 0.10	33.87± 1.38
mx65_xxhh	1.98± 0.06	5.52± 0.11	35.90± 1.33
mx100_xxhh	3.03± 0.08	8.47± 0.13	35.72± 1.07
mx500_xxhh	7.13± 0.12	22.66± 0.21	31.47± 0.61
mx1000_xxhh	7.34± 0.12	25.00± 0.22	29.35± 0.55

Table D.2 Fiducial acceptances and selection efficiencies for EFT for $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV.

Appendix E

Jet image with deep Convolutional Neural Network

In Section 9.1 we have seen that the DNN provides a rather powerful discrimination between large-R jets originating from $H \to b\bar{b}$ by using several high-level physics-inspired discriminating variables (jet mass, jet p_{T} jet multiplicity, b-tagging and e.t.c) as inputs. These variables are carefully engineered with guidance from physics theory. As powerful as they can be, there could be a considerable wealth of information inside a jet that is not yet fully exploited. In the attempt to extract this "hypothesised" low-level information contained in a jet, we borrow an advanced Computer Vision technique called Convolutional Neural Network (CNN).

CNN is a type of feed-forward artificial neural network that have proven to be very effective in areas such as image recognition and classification. It is the state of the art tool for most machine learning practitioners today for identifying images. The primary purpose of CNN is to automatically extract features from the input image. This is in contrast to the more conventional neural network where extensive feature engineering is often necessary. In order to capitalise this aspect of the CNN for jet tagging, we first need to represent a jet as an image. This procedure is detailed in Section E.1 below. We also use the output of CNN as an new input to the DNN. The performance of the CNN discriminant and the combination of both CNN and DNN are compared. The result will be shown in Section E.4.

E.1 Jet images as input to CNN

The jet images are built from LCW topo-clusters. As explained in Appendix B.1, topo-clusters are three-dimensional "energy blobs" representing the showers developing for each particle entering the calorimeter. However, they can be approximated (projected in $\eta - \phi$ plane) as a single layered grid of cluster with size of $\Delta\eta \times \Delta\phi = 0.1x0.1$ (to match the hadronic calorimeter granularity), without loss of generality. The grid spans [-1.2, 1.2] in $\eta - \phi$ plane with its centre coincide with the centre of the large-R jet axis. This yields a total of 25x25 elements or pixels

(henceforth, we will refer to each element within a jet-image as pixels). The intensity of each pixel is given by the sum of the energy of all clusters of the jet inside that pixel, normalised to the total jet energy. Some example of jet images constructed in this way are illustrated in Figure E.1-Figure E.4.

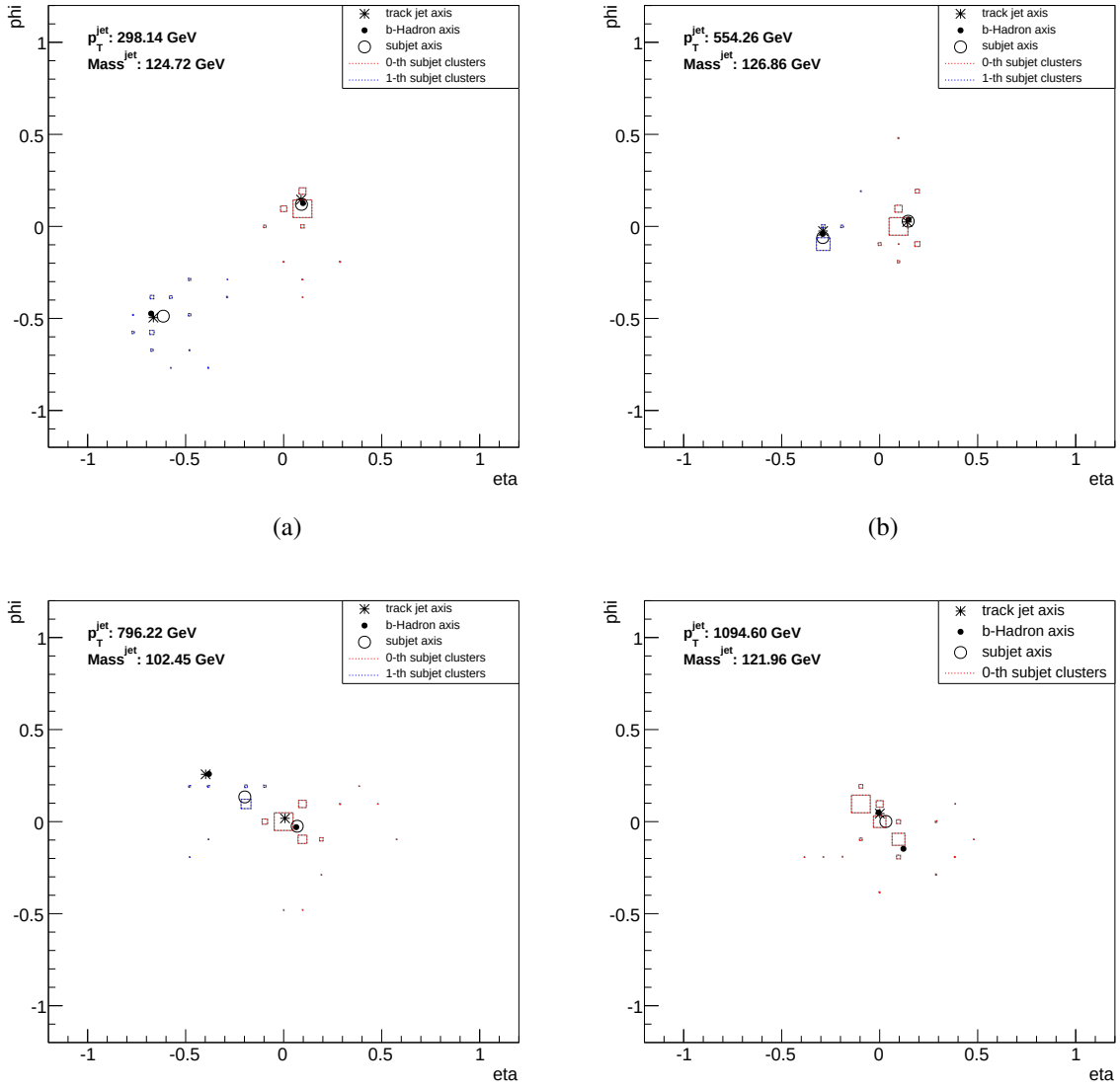
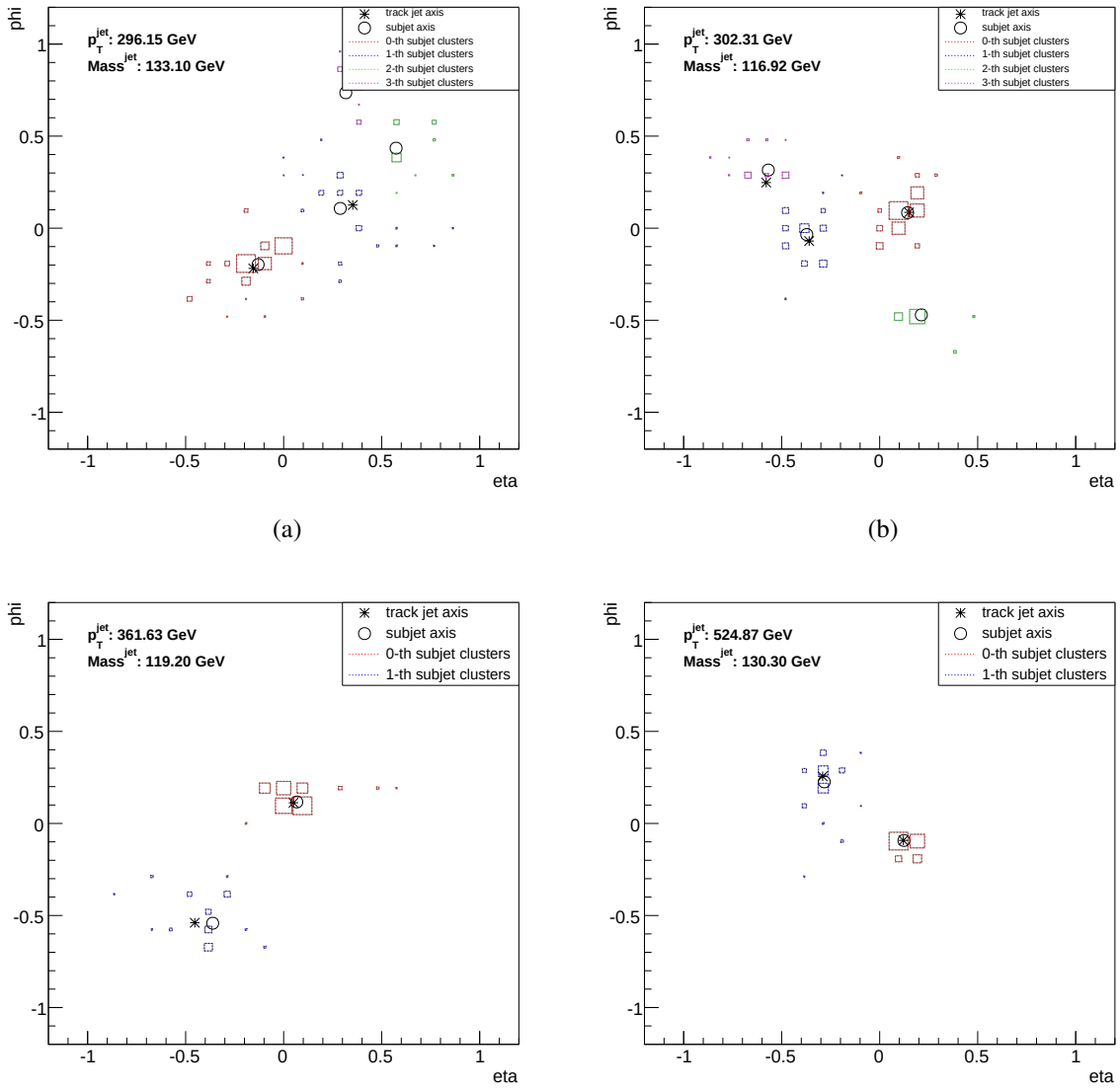


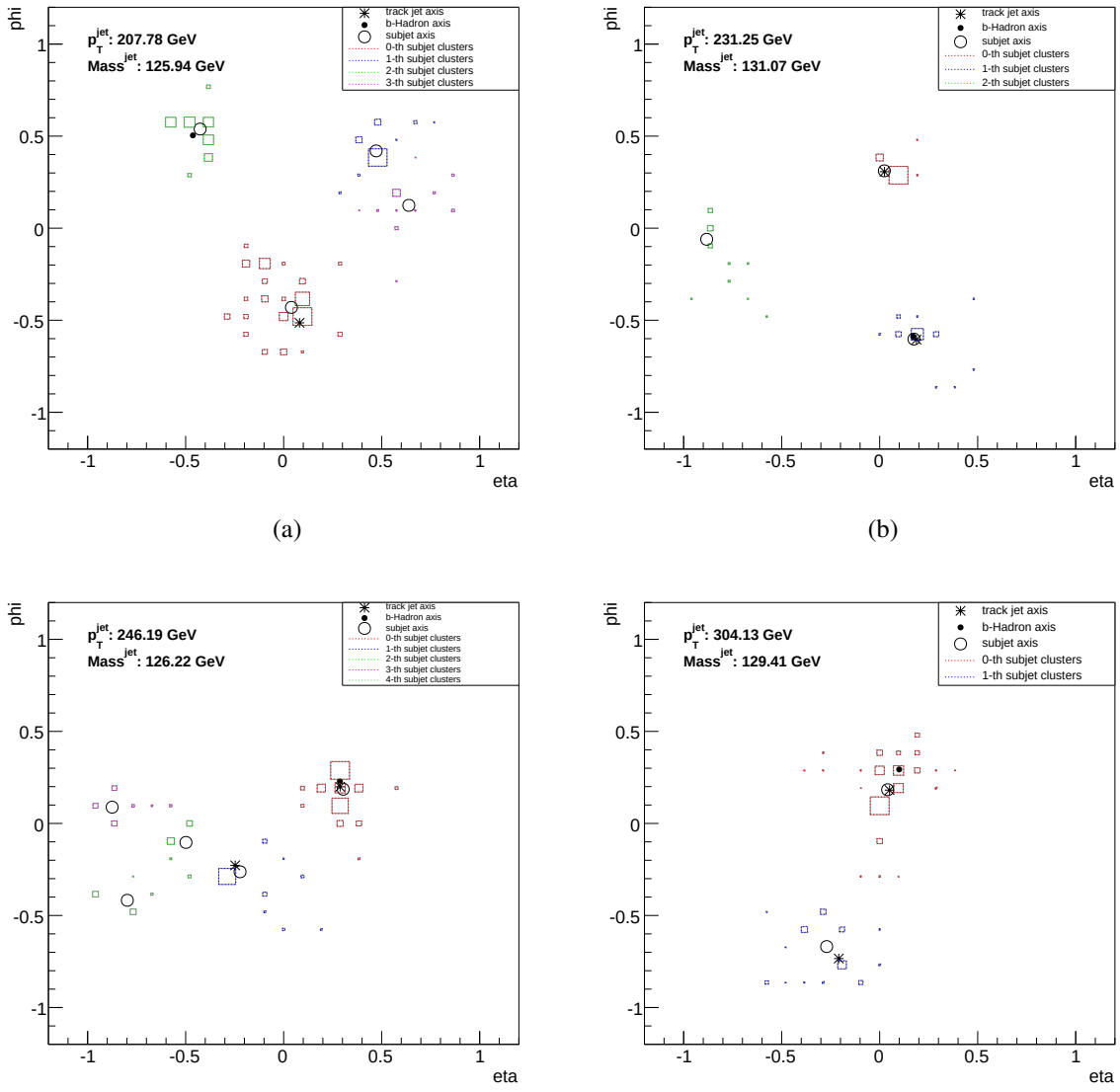
Fig. E.1 The jet image for signal $H \to b\bar{b}$ large-R jet. The size of the boxes represent the fractional jet energy contained in the corresponding pixel. The large-R jet axis is at the center of the image. Track jet, B-hadron and subjet axis are superimpose for reference. These axis are not used as input to CNN training.



(a) (b)

(c) (d)

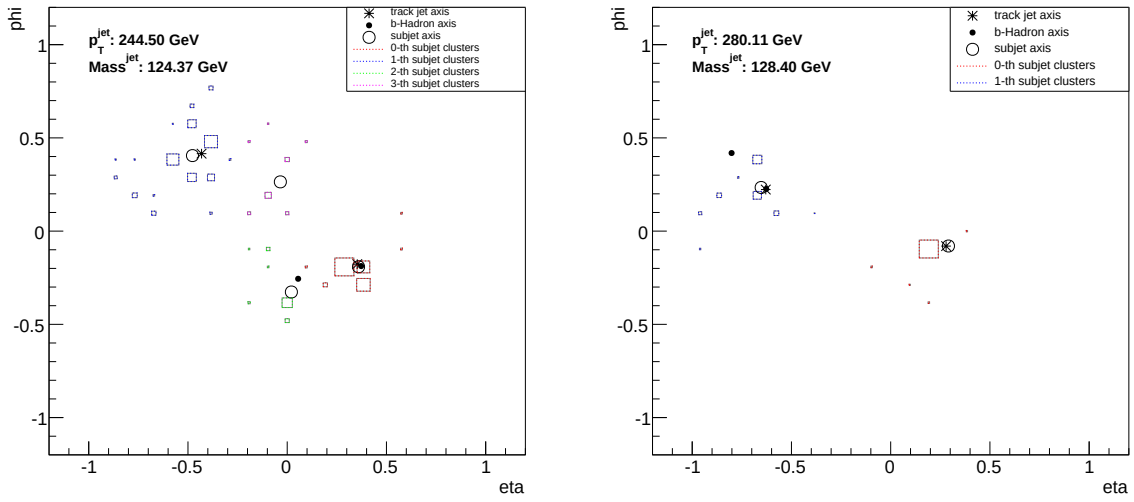
Fig. E.2 The jet image for background jet with 0 B-hadron from $Z(\rightarrow \nu\nu)$+jets sample. The size of the boxes represent the fractional jet energy contained in the corresponding pixel. Track jet, B-hadron and subjet axis are superimpose for reference. These axis are not used as input to CNN training.



(a) (b)

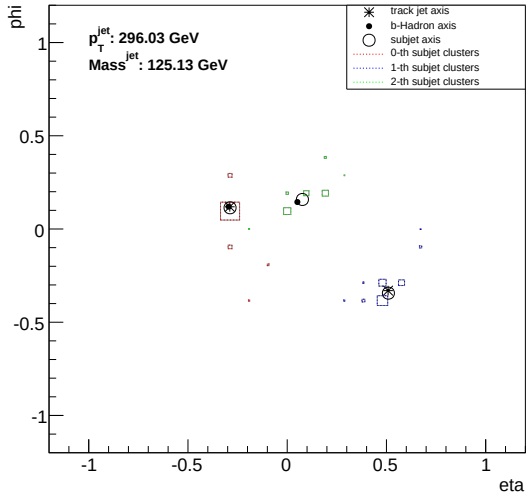
(c) (d)

Fig. E.3 The jet image for background jet with 1 B-hadron from $Z(\rightarrow \nu\nu)$+jets sample. The size of the boxes represent the fractional jet energy contained in the corresponding pixel. Track jet, B-hadron and subjet axis are superimpose for reference. These axis are not used as input to CNN training.

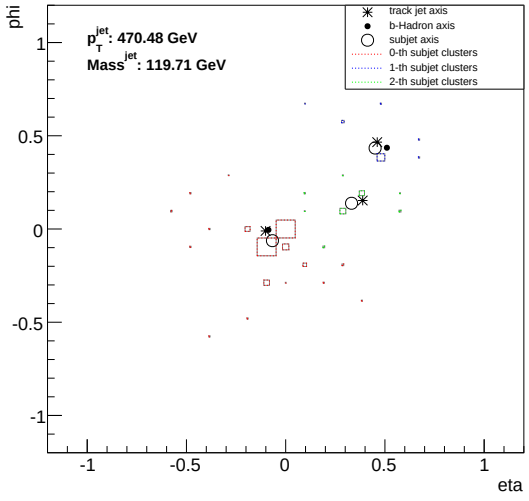


(a)

(b)



(c)



(d)

Fig. E.4 The jet image for background jet with 2 B-hadron from $Z(\rightarrow \nu\nu)$+jets sample. The size of the boxes represent the fractional jet energy contained in the corresponding pixel. Track jet, B-hadron and subjet axis are superimpose for reference. These axis are not used as input to CNN training.

Essentially, every image can be represented as a matrix of pixel values. For example, a colour image from a standard digital camera will have three channels, i.e. red, green and blue. We can imagine those as three 2D-matrices stacked over each other (one for each colour), each having pixel values in the range 0 to 255. In some sense, a jet image represents a grayscale image which has just one channel. The value of each pixel in the matrix will range from 0 to 1 (zero indicating black and 1 indicating white). This 2-D matrix is exactly what serves as an

input to the CNN. Having constructed the input to CNN, now we turn to the architecture of the CNN itself.

E.2 CNN architecture

The vast majority of modern CNN used for object recognition are built using the same basic architecture or scheme, which is composed of a stack of alternating convolution (CONV), rectified linear unit (ReLU) and max-pooling (POOL) layers followed by a small number of fully connected (FC) layers. This basic architecture can be represented as [INPUT-CONV-RELU-POOL-FC]. In more detail:

- INPUT is a [25x25x1] matrix which will hold the raw pixel values of the image.
- CONV layer has a primary purpose to extract features from the input image. It will compute the output of neurons that are connected to local regions in the input, each computing a dot product (convolved feature) between their weights and a small region they are connected to in the input volume. The CONV layers parameters consist of a set of learnable filters. Every filter is small spatially (along width and height), but extends through the full depth of the input volume. We specify a CONV layer with nomenclature [width, height, number of filters].
- RELU layer will apply an element wise (applied per pixel) activation function, such as the $max(0,x)$[1] thresholding at zero. It replaces all negative pixel values in the convolved feature by zero. The purpose is to introduce non-linearity in our network, otherwise the network would only ever be able to compute a linear function.
- POOL layer will perform a downsampling operation along the spatial dimensions (width, height). It reduces the dimensionality of each convolved feature but retains the most important information. It also makes the network invariant to small transformations, distortions and translations in the input image.
- FC layer consist of ordinary Neural Networks that use a Softmax activation function in the output layer. As the name implies, each neuron in this layer will be connected to all the numbers in the previous volume. This layer will compute the class scores.

Three hyper-parameters control the size of the output volume: the depth, stride and zero-padding. First, the depth parameter corresponds to the number of filters we would like to use for each CONV layer, each learning to look for something different in the input. Next, the stride parameter specify how far we slide the filter. When the stride is N then the filters jump N pixels

[1] $max(0,x)$ is used because it can be computed much more efficiently compared to the other more conventional activation functions like the sigmoid and hyperbolic tangent, without making a significant difference to generalisation accuracy.

Table E.1 The CNN used to classify a jet as originating from $H \to b\bar{b}$ or not. Each row correspond to a module. A module consist of a CONV layer followed by RELU layer. Zero-padding is used in each CONV layer.

Input 25] $\times$ 25 grayscale image
[3, 3, 96] CONV-RELU with stride = 1
[3, 3, 96] CONV-RELU with stride = 1
[3, 3, 96] CONV-RELU with stride = 2
[3, 3, 96] CONV-RELU with stride = 1
[3, 3, 96] CONV-RELU with stride = 1
[3, 3, 96] CONV-RELU with stride = 2
[3, 3, 192] CONV-RELU with stride = 1
[1, 1, 192] CONV-RELU with stride = 1
[1, 1, 10] CONV-RELU with stride = 2
FC-RELU
FC-RELU
FC-Softmax linear

at a time as we slide them around. Larger stride value will produce smaller output volumes spatially. Finally, zero-padding allow us to control the spatial size of the output volumes by padding the input matrix with zeros around the border.

Different variant of CNN can be built by staking multiple [CONV-RELU] and/or [POOL] layers or by using more complicated activation functions. In general it is a good idea to construct a larger and deeper networks. It is because the more number of layers we have, the more image features get extracted and the better our network becomes at recognising patterns in the input images. Also, a plethora of extensions [213–215] which follow this basic scheme but with more complex activation functions were recently proposed to enhance networks performance. In our analysis, we have chosen the former. More precisely, we follow a similar all-convolutional-net architecture proposed by Jost et. al. [216]. The proposed network replace all max-pooling layer by a convolutional layer. It offers a much simpler network architecture yet its performance matches or even slightly outperforms the other state of the art networks which include explicit (max-)pooling operations in a network. The exact all-convolutional-net architecture is listed in Table E.1.

E.3 Implementation and training

The CNN is built by using TensorFlow [196] software library. The training and testing sample are prepared based on the same procedure as described for DNN analysis in Section 9.1. To recap, the same mono-Higgs($\to b\bar{b}$) signal and $Z(\to \nu\nu)$+jets background MC samples described in Section 3 are used for this study. Leading large-R jets are matched to truth b-hadron by requiring $\Delta R(J^1,\text{b}-Hadron) < 1.0$ where b-hadrons direction is taken from the MC

information. Large-R jet originating from Higgs($\rightarrow b\bar{b}$) are labeled as "signal" jets whereas those coming from the QCD multi jet (e.g gluon$\rightarrow b\bar{b}$) are labeled as "background' jets. All signal jets are required to have two matching b-hadrons whereas background jets are demanded to be associated to either 0/1/2 b-hadrons. For training, all leading large-R jets are required to have $p_{\mathrm{T}} > 200$ GeV. From a total of 505332 (3171982) signal (background) events, 65% are randomly selected for training. Another 5% of the total events are randomly selected for validation purpose. The rest are assigned as testing set.

The CNN was trained for a total of 40000 steps. The output of the final FC layer is a probability score (between 0 and 1) for each large-R jet. Value closer to one means the large-R jet has higher probability to be of the signal jet that originating from Higgs($\rightarrow b\bar{b}$) and vice versa. The distributions of the output score the for signal and background jets are shown in Figure E.5. A separation between signal and background jets can be clearly seen. By applying a cut on this probability score, a desire signal selection efficiency and background rejection can be obtained.

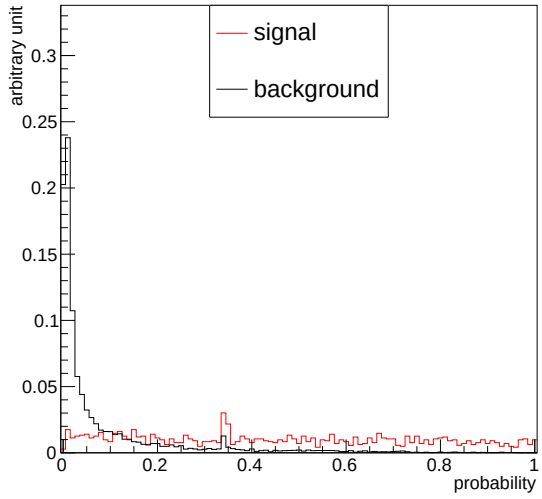


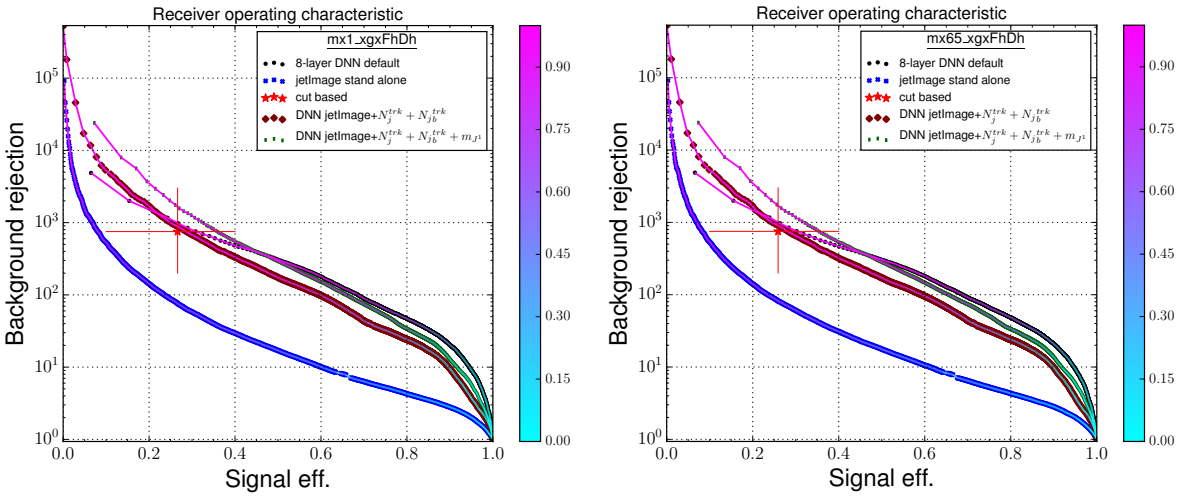
Fig. E.5 Distribution of the output score of CNN for signal and background large-R jet.

E.4 Performance

To visualise the performance of each CNN, a Receiver Operating Characteristic (ROC) curve in which background rejection, the inverse of background efficiency is plotted against the signal efficiency at each cut on probability score, can be used. Similar to the procedures explained in Section 9.1, to provide a direct performance comparison with our current cut-based analysis, test samples are constructed from those events subjected to the same signal event selection up to large-R jet multiplicity ($n_J \geq 1$) in Table 5.1. The total events after $n_J \geq 1$ selection is

assigned as the denominator when assessing the selection efficiency. On the other hand, the numerator will be the leftover events after jet mass cut (cut-based method) and CNN score cut (CNN method). We also compare the performance with the result obtained using 8-layer DNN, which has been discussed in Section 9.1. Furthermore, we also treat the CNN output as an input to the DNN. Two set of DNNs with different inputs are compared. One is the CNN output plus number of track-jets and number of b-tagged track-jets. By adding the large-R jet mass, another set is formed. The ROC curves for CNN as well as each DNN variant are shown in Figure E.6. The selection efficiency for cut-based method is plotted as a red star in the same plot. The CNN output alone does not perform well but once it is combined with the other variables, the DNN yields competitive results or even outperform the 8-layer DNN.

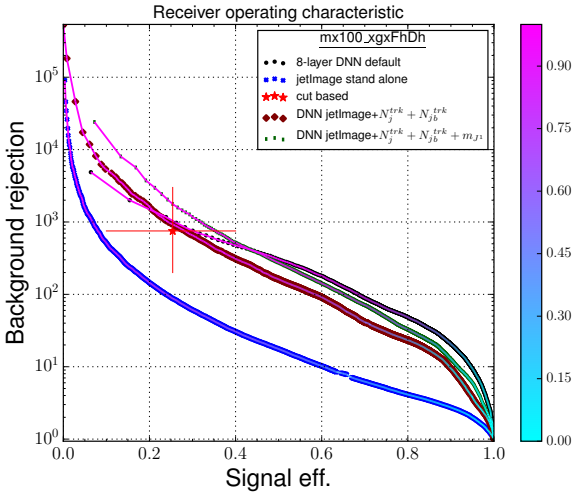
The fact that the CNN output alone does not perform well might be due to the way the CNN is trained. No separated training is performed for each category of background jets (associated with 0/1/2 b-hadrons). If separate training is performed (i.e. signal jets against background jets with 0/1/2 b-hadron), more discriminating power can be recovered as evidenced by the CNN output distributions shown in Figure E.7. Compare to the distribution in Figure E.5, the signal and background distributions exhibit more pronounced separation. All in all, further investigate and optimisation is needed as I believe this technique is very promising.


Receiver operating characteristic
mx1_xgxFhDh
8-layer DNN default
jetImage stand alone
cut based
DNN jetImage+$N_j^{trk}+N_{jb}^{trk}$
DNN jetImage+$N_j^{trk}+N_{jb}^{trk}+m_{j\iota}$
Background rejection
Signal eff.

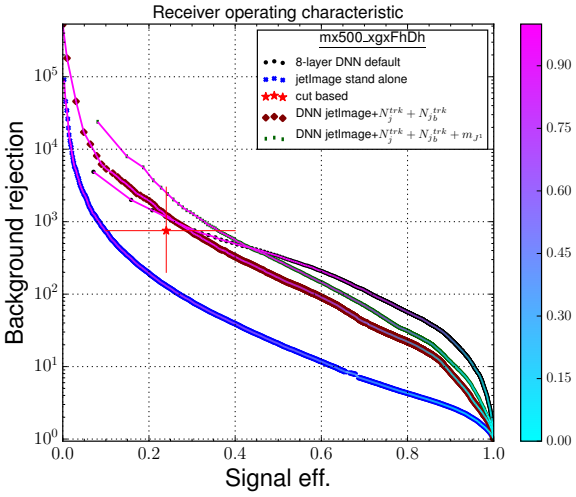
(a)

Receiver operating characteristic
mx65_xgxFhDh
8-layer DNN default
jetImage stand alone
cut based
DNN jetImage+$N_j^{trk}+N_{jb}^{trk}$
DNN jetImage+$N_j^{trk}+N_{jb}^{trk}+m_{j\iota}$
Background rejection
Signal eff.

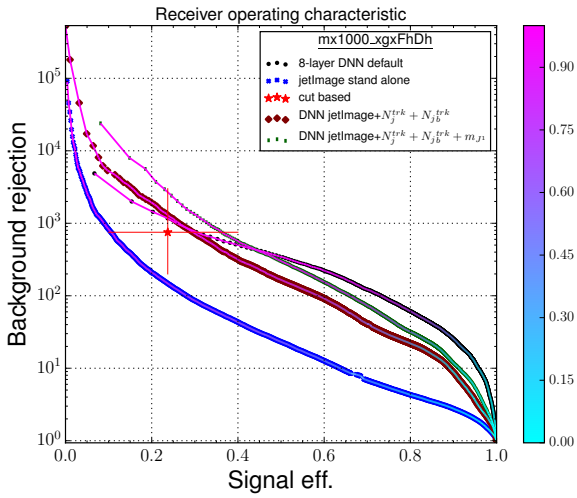
(b)


Receiver operating characteristic
mx100_xgxFhDh
8-layer DNN default
jetImage stand alone
cut based
DNN jetImage+$N_j^{trk}+N_{jb}^{trk}$
DNN jetImage+$N_j^{trk}+N_{jb}^{trk}+m_{j\iota}$
Background rejection
Signal eff.

(c)

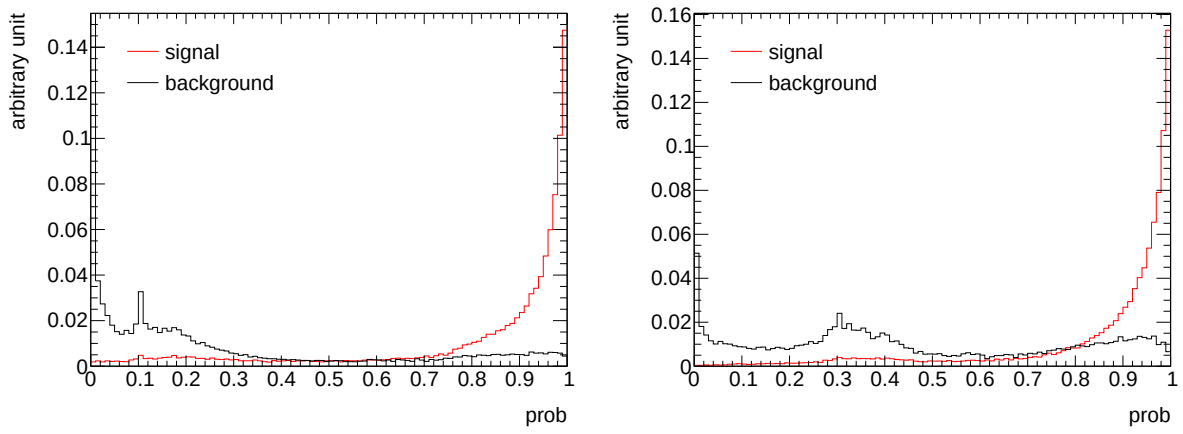

Receiver operating characteristic
mx500_xgxFhDh
8-layer DNN default
jetImage stand alone
cut based
DNN jetImage+$N_j^{trk}+N_{jb}^{trk}$
DNN jetImage+$N_j^{trk}+N_{jb}^{trk}+m_{j\iota}$
Background rejection
Signal eff.

(d)


Receiver operating characteristic
mx1000_xgxFhDh
8-layer DNN default
jetImage stand alone
cut based
DNN jetImage+$N_j^{trk}+N_{jb}^{trk}$
DNN jetImage+$N_j^{trk}+N_{jb}^{trk}+m_{j\iota}$
Background rejection
Signal eff.

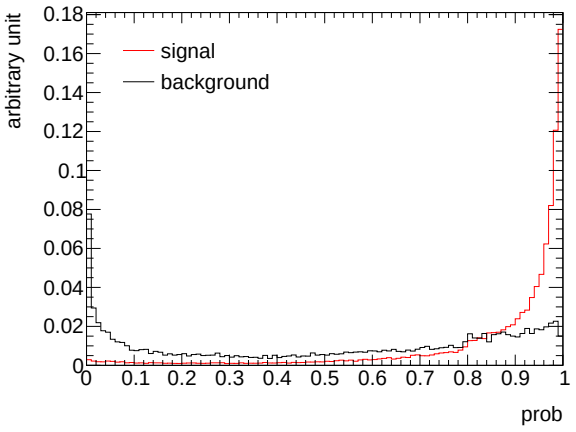
(e)

Fig. E.6



(a) 0 *b*-hadron

(b) 1 *b*-hadron



(c) 2 *b*-hadrons

Fig. E.7 CNN output distribution when training is done separately for signal jets against background jets with a) 0 *b*-hadron, b) 1 *b*-hadrons and c) 2 *b*-hadrons.

Appendix F

Detailed systematics uncertainties

F.1 Systematic uncertainties for SR with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeV

Table F.1 Percent impact of the systematics uncertainties on the total background in the signal region in $E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$ bin

	Diboson	W+jets	$t\bar{t}$	SM Higgs	Z$\nu\nu$	Total Bkg.
JetFlavB up	22.4	27.4	9.0	11.2	1.9	6.5
JetFlavB down	10.3	10.4	4.9	10.0	0.5	3.5
JES large-R up	13.2	36.1	6.8	11.7	-0.4	8.2
JES large-R down	-9.9	-6.1	-10.7	-11.4	1.1	-1.9
JER large-R	2.8	14.4	-2.8	-2.4	1.0	4.3
JMS large-R up	38.8	-0.2	-14.6	-5.4	0.6	0.8
JMS large-R down	-38.0	-2.9	10.6	0.8	-0.0	0.5
JMR large-R	5.1	8.6	-3.9	-0.8	-0.2	-1.0
JES up	-1.6	6.3	-6.2	-2.5	0.1	0.5
JES down	2.4	-1.1	3.2	2.2	-0.3	-0.6
JER	1.3	1.2	-0.5	-1.1	0.1	-1.7
JVF up	1.0	3.0	4.1	1.3	0.0	0.6
JVF down	-0.0	-0.6	0.0	-0.3	-0.1	-0.2
METScaleSoftTerms up	0.0	0.0	0.0	-0.1	-0.0	-0.0
METScaleSoftTerms down	0.0	0.0	0.0	0.0	0.0	0.0
METResoSoftTerms	0.0	0.0	0.0	-0.1	-0.1	-0.1
PileUpRs up	1.7	11.8	-1.6	2.4	-4.4	-1.6
PileUpRs down	1.8	4.9	-0.6	1.6	-4.7	-2.4
PES up	-	-	-	-	-0.3	-0.2
PES down	-	-	-	-	0.0	0.0
PER	-	-	-	-	-0.3	-0.2
PhotonID	-	-	-	-	4.0	2.5
ttbarPtRW up	-	-	15.5	-	-	1.1
ttbarPtRW down	-	-	-15.5	-	-	-1.1
ZnnTF up	-	-	-	-	5.1	3.2
ZnnTF down	-	-	-	-	-12.0	-7.5
Cross section	26.0	20.0	7.0	3.1	11.0	7.4
PDF & α_s	5.9	5.0	6.0	2.5	-	1.8
Total Sys. Up	53.8	54.4	27.1	18.9	13.9	14.4
Total Sys. Down	48.7	29.9	24.5	17.1	17.7	12.8

Table F.2 Percent impact of the systematics uncertainties on the signal xgxFhDh in $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ *GeV* bin

	mx1_xgxFhDh	mx65_xgxFhDh	mx100_xgxFhDh	mx500_xgxFhDh	mx1000_xgxFhDh
JetFlavB up	12.1	12.2	12.2	12.6	12.7
JetFlavB down	-11.5	-11.6	-11.6	-11.9	-12.0
JES large-*R* up	2.6	3.0	3.1	2.3	2.1
JES large-*R* down	-3.4	-3.1	-3.1	-2.0	-2.2
JER large-*R*	-0.5	-0.1	-0.0	0.2	-0.0
JMS large-*R* up	-1.2	-1.4	-1.4	-0.9	-0.9
JMS large-*R* down	-1.0	-1.0	-0.8	-1.0	-1.9
JMR large-*R*	-1.0	-1.3	-0.8	-0.5	-0.9
JES up	-1.0	-1.0	-1.1	-1.3	-0.8
JES down	1.3	1.2	1.3	1.2	1.1
JER	-0.3	0.0	-0.1	0.2	-0.4
JVF up	0.4	0.5	0.4	0.4	0.3
JVF down	-0.2	-0.2	-0.2	-0.4	-0.3
METScaleSoftTerms up	0.0	0.0	0.0	0.0	-0.0
METScaleSoftTerms down	0.0	0.0	0.0	0.0	0.0
METResoSoftTerms	0.0	0.0	-0.0	0.0	0.0
PileUp	-0.5	-0.8	-0.3	-0.7	0.0
PileDo	-0.3	-0.3	-0.1	-0.2	0.0
Total Sys. Up	12.6	12.8	12.7	12.9	13.0
Total Sys. Down	12.2	12.1	12.1	12.2	12.5

	mx1_xdxhDh	mx65_xdxhDh	mx100_xdxhDh	mx500_xdxhDh	mx1000_xdxhDh
JetFlavBUp	11.2	11.3	11.2	12.2	12.5
JetFlavBDo	-10.7	-10.7	-10.7	-11.5	-11.8
JESUp	5.4	6.8	4.7	2.9	2.5
JESDo	-10.5	-6.2	-6.2	-2.9	-2.1
JER	-0.9	0.4	-0.4	-0.2	0.1
JMSUp	-3.7	-1.5	-2.3	-1.2	-1.7
JMSDo	-0.2	0.1	0.1	-0.3	-0.1
JMR	-1.7	-1.2	-1.3	-1.0	-0.7
Akt4JESUp	-1.7	-1.7	-1.8	-1.8	-2.1
Akt4JESDo	3.0	0.8	1.6	2.0	1.4
Akt4JER	0.6	0.2	-0.0	-0.4	-0.5
JVFUp	0.3	0.3	0.6	0.6	0.7
JVFDo	0.0	-0.4	-0.3	-0.3	-0.5
BCHUp	0.6	0.4	0.2	0.2	0.3
BCHDo	0.0	-0.7	-0.3	-0.6	-0.4
METScaleSoftTermsUp	0.0	0.0	0.0	-0.0	0.0
METScaleSoftTermsDo	0.3	-0.1	-0.0	-0.1	0.0
METResoSoftTerms	0.0	-0.3	-0.1	-0.0	0.1
PileUp	0.9	1.2	1.0	-0.5	-0.4
PileDo	0.2	0.5	0.4	-0.1	-0.2
Total Sys. Up	13.3	13.5	12.7	12.8	13.1
Total Sys. Down	15.4	12.5	12.6	12.1	12.2
Stat_err_nominal	5.4	2.6	2.2	1.7	1.5

Table F.3 Percent impact of the systematics uncertainties on the signal xdxhDh in $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ *GeV* bin

	mx1_xxhhg5	mx65_xxhhg5	mx100_xxhhg5	mx500_xxhhg5	mx1000_xxhhg5
JetFlavBUp	11.2	11.3	11.3	12.4	12.5
JetFlavBDo	-10.7	-10.7	-10.7	-11.7	-11.9
JESUp	6.5	6.6	5.5	2.1	2.1
JESDo	-6.0	-5.9	-6.3	-2.6	-2.2
JER	0.9	0.8	-0.2	-0.4	0.0
JMSUp	-2.6	-1.7	-2.4	-1.3	-1.0
JMSDo	0.5	1.0	-0.0	-0.2	-0.1
JMR	-1.6	-0.4	-1.1	-0.7	-0.6
Akt4JESUp	-2.0	-2.0	-1.5	-2.3	-1.8
Akt4JESDo	2.2	3.3	1.6	2.3	2.5
Akt4JER	-1.2	-0.6	-0.5	-0.2	0.1
JVFUp	0.7	0.8	0.7	1.1	1.0
JVFDo	-0.8	-0.6	-0.5	-0.4	-0.7
BCHUp	0.2	0.1	0.1	0.4	0.2
BCHDo	-0.4	-0.4	-0.4	-0.5	-0.6
METScaleSoftTermsUp	-0.0	-0.2	-0.0	-0.1	-0.1
METScaleSoftTermsDo	0.0	0.0	0.2	0.1	0.1
METResoSoftTerms	-0.1	0.3	0.0	-0.1	-0.1
PileUp	-0.3	-0.1	-0.5	-0.4	-1.0
PileDo	-0.2	-0.1	-0.2	-0.2	-0.4
Total Sys. Up	13.6	13.4	12.9	12.9	13.0
Total Sys. Down	12.7	12.8	12.6	12.3	12.4
Stat_err_nominal	2.9	2.7	2.1	1.6	1.6

Table F.4 Percent impact of the systematics uncertainties on the signal xxhhg5 in $E_{\mathrm{T}}^{\mathrm{miss}} > 300\ GeV$ bin

	mx1_xxhh	mx65_xxhh	mx100_xxhh	mx500_xxhh	mx1000_xxhh
JetFlavBUp	11.1	11.2	11.1	12.3	12.8
JetFlavBDo	-10.5	-10.6	-10.6	-11.6	-12.1
JESUp	5.8	6.7	6.2	2.7	2.0
JESDo	-6.3	-6.2	-6.1	-3.0	-2.1
JER	-0.3	-0.2	0.3	-0.6	-0.1
JMSUp	-2.0	-2.5	-2.6	-1.6	-1.5
JMSDo	2.0	-0.8	0.9	-0.4	0.1
JMR	-0.8	-2.1	-0.9	-1.0	-0.7
Akt4JESUp	-2.0	-1.9	-2.1	-1.9	-1.8
Akt4JESDo	2.4	2.0	1.5	2.4	1.8
Akt4JER	-1.5	0.2	-1.3	-0.3	0.5
JVFUp	0.6	0.5	1.0	0.7	0.8
JVFDo	-0.9	-0.9	-0.5	-0.4	-0.4
BCHUp	0.2	0.2	0.4	0.3	0.4
BCHDo	-0.2	-0.8	-0.5	-0.5	-0.5
METScaleSoftTermsUp	-0.2	0.0	-0.0	-0.2	-0.0
METScaleSoftTermsDo	0.0	-0.1	0.0	0.0	0.0
METResoSoftTerms	0.0	-0.1	0.0	-0.0	0.0
PileUp	1.0	0.6	0.4	-0.1	-0.1
PileDo	0.5	0.3	0.2	0.0	-0.1
Total Sys. Up	13.0	13.6	13.3	12.9	13.2
Total Sys. Down	12.9	12.8	12.5	12.3	12.4
Stat_err_nominal	3.0	2.8	2.3	1.6	1.6

Table F.5 Percent impact of the systematics uncertainties on the signal xxhh in $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ *GeV* bin

F.2 Systematic uncertainties for SR with $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV

Table F.6 Percent impact of the systematics uncertainties on the total background in the signal region in $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ *GeV* bin

	Diboson	W+jets	$t\bar{t}$	SM Higgs	Z$\nu\nu$	Total Bkg.
JetFlav up	19.0	37.9	10.9	12.3	0.3	7.0
JetFlav down	11.6	11.2	5.9	10.9	0.3	3.2
JES large-R up	0.7	0.0	0.0	1.0	-0.5	-0.2
JES large-R down	-2.5	-0.0	-6.5	-0.0	0.1	-0.1
JER large-R	-1.9	-0.0	-8.2	-0.3	0.1	-0.1
JMS large-R up	38.5	-0.3	-32.8	-4.6	1.0	1.3
JMS large-R down	-37.4	4.6	21.8	-0.1	-0.0	1.1
JMR large-R	-3.0	7.5	-8.0	-1.5	-0.0	-1.5
JES up	4.3	0.5	9.1	3.2	-1.1	-0.1
JES down	-7.2	1.4	-1.9	-4.1	0.6	-2.2
JER	0.0	18.9	-6.4	3.0	-1.1	-1.5
JVF up	1.0	-0.0	0.0	2.1	0.0	0.2
JVF down	-0.0	-0.0	0.0	0.0	-0.3	-0.2
METScaleSoftTerms up	0.5	0.5	0.0	-0.0	0.0	0.1
METScaleSoftTerms down	-0.7	0.0	0.0	0.6	0.0	0.0
METResoSoftTerms	-0.6	-6.9	0.0	-0.2	-0.1	-0.8
PileUpRs up	1.7	-1.5	-3.4	0.0	-0.8	-1.9
PileUpRs down	2.1	-0.6	-2.0	0.6	-1.0	-1.4
PES up	-	-	-	-	-0.4	-0.3
PES down	-	-	-	-	0.0	0.0
PER	-	-	-	-	-0.4	-0.3
PhotonID	-	-	-	-	4.0	2.7
ttbarPtRW up	-	-	16.4	-	-	1.3
ttbarPtRW down	-	-	-16.4	-	-	-1.3
ZnnTF up	-	-	-	-	5.1	3.4
ZnnTF down	-	-	-	-	-11.4	-7.7
Cross Section	30.0	20.0	7.0	3.1	11.0	7.7
PDF & α_s	5.9	5.0	6.0	2.5	-	1.6
Total Sys. Up	52.7	48.3	42.5	15.7	13.2	11.8
Total Sys. Down	50.0	32.3	32.8	13.9	16.6	12.2

	mx1_xgxFhDh	mx65_xgxFhDh	mx100_xgxFhDh	mx500_xgxFhDh	mx1000_xgxFhDh
JetFlavBUp	12.5	12.6	12.5	12.9	12.9
JetFlavBDo	-11.8	-11.9	-11.9	-12.2	-12.3
JESUp	0.5	0.6	0.5	0.4	0.4
JESDo	-0.7	-0.7	-0.6	-0.5	-0.6
JER	-0.1	-0.1	-0.1	0.0	-0.2
JMSUp	-1.2	-1.4	-1.3	-0.9	-1.0
JMSDo	-1.1	-1.2	-1.0	-1.1	-1.9
JMR	-1.0	-1.4	-0.8	-0.6	-1.0
Akt4JESUp	0.9	0.9	0.9	0.3	0.5
Akt4JESDo	-1.2	-0.6	-0.6	-0.2	-0.6
Akt4JER	-0.4	0.0	-0.3	0.2	-0.7
JVFUp	0.4	0.5	0.4	0.4	0.3
JVFDo	-0.2	-0.3	-0.2	-0.3	-0.3
BCHUp	0.3	0.3	0.2	0.4	0.3
BCHDo	-0.4	-0.4	-0.4	-0.4	-0.4
METScaleSoftTermsUp	-0.1	-0.0	0.0	-0.0	-0.1
METScaleSoftTermsDo	0.0	0.1	0.1	0.0	0.0
METResoSoftTerms	0.1	0.0	-0.1	0.0	-0.1
PileUp	-0.9	-1.0	-0.7	-0.9	-0.2
PileDo	-0.4	-0.4	-0.2	-0.3	-0.1
Total Sys. Up	12.7	12.8	12.7	13.0	13.1
Total Sys. Down	12.0	12.1	12.0	12.3	12.5
Stat_err_nominal	1.4	1.4	1.4	1.3	1.3

Table F.7 Percent impact of the systematics uncertainties on the signal xgxFhDh in $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ *GeV* bin

	mx1_xdxhDh	mx65_xdxhDh	mx100_xdxhDh	mx500_xdxhDh	mx1000_xdxhDh
JetFlavBUp	12.0	11.8	11.7	12.5	12.7
JetFlavBDo	-11.4	-11.2	-11.1	-11.8	-12.0
JESUp	0.4	1.8	1.1	1.1	0.7
JESDo	-1.4	-1.2	-1.2	-0.9	-1.2
JER	-0.5	0.4	-0.2	-0.1	-0.2
JMSUp	-3.2	-1.4	-2.5	-1.4	-1.8
JMSDo	-0.4	-0.2	0.5	-0.2	0.1
JMR	-0.9	-1.0	-1.2	-0.9	-0.7
Akt4JESUp	3.8	1.4	1.4	0.5	-0.7
Akt4JESDo	0.1	-2.9	-2.9	-0.3	-0.1
Akt4JER	2.0	0.4	-0.4	-0.8	-0.2
JVFUp	0.0	0.3	0.6	0.7	0.7
JVFDo	0.0	-0.3	-0.2	-0.4	-0.5
BCHUp	0.4	0.5	0.2	0.2	0.3
BCHDo	0.0	-0.7	-0.2	-0.6	-0.4
METScaleSoftTermsUp	-0.1	0.1	-0.1	-0.1	0.0
METScaleSoftTermsDo	0.0	-0.1	-0.2	-0.0	0.0
METResoSoftTerms	0.4	-0.2	-0.3	-0.1	0.1
PileUp	1.7	1.7	0.1	-0.6	-0.8
PileDo	0.6	0.7	-0.1	-0.1	-0.3
Total Sys. Up	13.3	12.3	12.2	12.7	12.9
Total Sys. Down	11.7	11.8	11.6	12.0	12.1
Stat_err_nominal	6.6	2.9	2.5	1.8	1.6

Table F.8 Percent impact of the systematics uncertainties on the signal xdxhDh in $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ *GeV* bin

	mx1_xxhhg5	mx65_xxhhg5	mx100_xxhhg5	mx500_xxhhg5	mx1000_xxhhg5
JetFlavBUp	11.8	11.7	11.6	12.7	12.8
JetFlavBDo	-11.2	-11.1	-11.0	-12.0	-12.1
JESUp	2.4	1.6	1.7	1.0	0.9
JESDo	-1.5	-1.4	-1.9	-1.1	-1.2
JER	0.3	-0.2	-0.1	-0.2	0.0
JMSUp	-3.1	-2.2	-2.5	-1.2	-0.9
JMSDo	0.2	0.8	-0.4	-0.1	-0.2
JMR	-1.9	-0.7	-1.1	-0.7	-0.6
Akt4JESUp	1.1	1.7	0.7	-1.0	-0.5
Akt4JESDo	-2.0	-0.9	-1.3	0.6	0.9
Akt4JER	-1.6	-0.3	0.1	-0.1	0.1
JVFUp	0.4	0.3	0.7	1.1	0.9
JVFDo	-0.6	-0.5	-0.6	-0.4	-0.6
BCHUp	0.2	0.2	0.1	0.4	0.2
BCHDo	-0.5	-0.5	-0.4	-0.5	-0.6
METScaleSoftTermsUp	-0.6	0.1	-0.2	-0.1	-0.1
METScaleSoftTermsDo	0.1	0.8	0.4	0.2	0.1
METResoSoftTerms	-0.5	-0.1	0.3	-0.1	-0.1
PileUp	-1.9	-0.5	-0.6	-0.7	-1.0
PileDo	-1.0	-0.1	-0.3	-0.3	-0.4
Total Sys. Up	12.9	12.2	12.1	12.9	12.9
Total Sys. Down	11.8	11.3	11.4	12.1	12.2
Stat_err_nominal	3.3	3.1	2.4	1.7	1.6

Table F.9 Percent impact of the systematics uncertainties on the signal xxhhg5 in $E_{\mathrm{T}}^{\mathrm{miss}} > 400\ GeV$ bin

	mx1_xxhh	mx65_xxhh	mx100_xxhh	mx500_xxhh	mx1000_xxhh
JetFlavBUp	11.6	11.7	11.5	12.6	13.0
JetFlavBDo	-11.0	-11.1	-11.0	-11.9	-12.3
JESUp	2.3	1.7	1.4	1.3	0.9
JESDo	-2.0	-2.4	-2.1	-1.5	-1.1
JER	-0.6	-0.8	-0.1	-0.4	-0.1
JMSUp	-2.1	-1.8	-2.9	-1.5	-1.5
JMSDo	2.4	-1.2	0.6	-0.5	0.3
JMR	-0.4	-2.6	-0.7	-1.0	-0.8
Akt4JESUp	0.8	2.4	1.0	-0.3	-0.7
Akt4JESDo	-1.1	-1.9	-1.2	1.0	0.6
Akt4JER	-1.6	0.2	-1.1	-0.3	0.4
JVFUp	0.4	0.3	1.1	0.7	0.8
JVFDo	-1.0	-0.8	-0.5	-0.4	-0.3
BCHUp	0.0	0.2	0.5	0.3	0.3
BCHDo	-0.1	-1.0	-0.5	-0.6	-0.5
METScaleSoftTermsUp	-0.1	0.1	-0.1	-0.1	-0.2
METScaleSoftTermsDo	0.0	0.3	0.3	0.0	0.1
METResoSoftTerms	-0.2	0.3	0.5	-0.1	-0.0
PileUp	1.3	-0.9	0.3	-0.3	-0.3
PileDo	0.5	-0.2	0.2	-0.1	-0.2
Total Sys. Up	12.3	12.5	12.2	12.8	13.2
Total Sys. Down	11.7	11.9	11.4	12.1	12.4
Stat_err_nominal	3.5	3.2	2.6	1.7	1.7

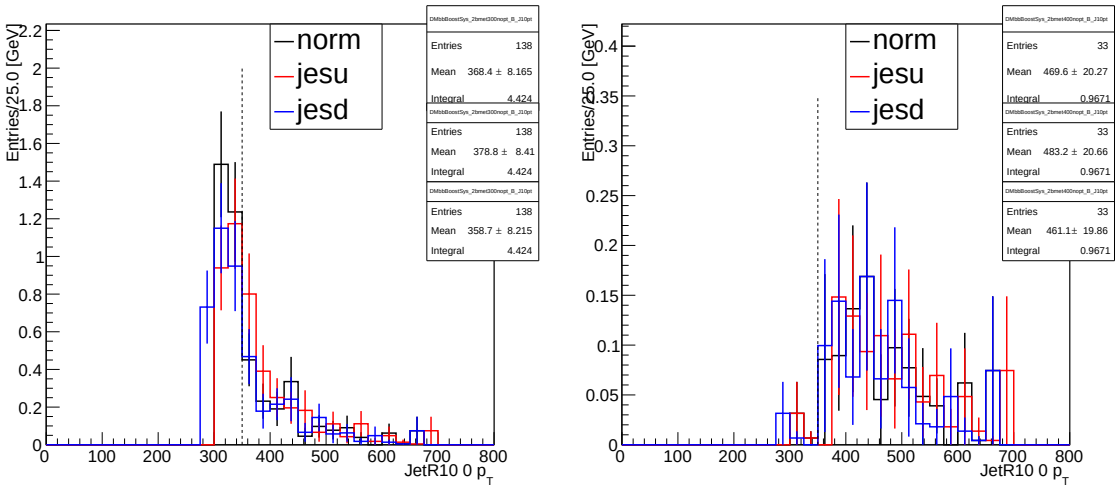
Table F.10 Percent impact of the systematics uncertainties on the signal xxhh in $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ *GeV* bin

F.3 Validation of the JES (JMS) uncertainty for large-R jets for W+jets (diboson) background

W+jets

From Table F.1, a large JES uncertainty for large-R jet is observed for W+jets background. This can be explained by looking at the large-R jet p_T distribution in which all event selection criteria are applied except jet p_T cut. As can be seen in Figure F.1a the signal jet p_T cut of 350 GeV is at the falling edge of the distribution. An increase in JES shifts a considerable amount of lower pt jets to higher pt bin. On the other hand, effect of lower JES on large-R jet pt is less profound.

While examining JES systematics uncertainties for W+jets background in Table F.6, a stark contrast can be observed. Its value in $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV bin is clearly much less than the corresponding number in $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ GeVbin. The reason can be due to the fact that a $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ GeV cut practically makes the leading large-R jet p_T distribution to plateau around 400 GeV. After all the event selections, the statistic is less and the distribution is considerably flatter as illustrates in Figure F.1b. This renders the JES variation to have less effect.



(a) $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ *GeV* (b) $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ *GeV*

Fig. F.1 Leading large-R jet p_T distribution in SR with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ *GeV* and $E_{\mathrm{T}}^{\mathrm{miss}} > 400$ *GeV* for MC W+jets sample. All event selection criteria are applied except jet p_T cut. The vertical dotted lines show the cut value of large-R jet's p_T. Three distributions with systematic variation up (red), variation down (blue) and nominal (black) are shown.

Diboson

As shown in Table F.1, JMS systematic variation on diboson can be as large as $\sim$40%. This is due to the fact that the lower jet mass boundary of 90 GeV falls right at the steep falling edge of the large-R jet mass distribution. Changing the JMS up results in a large migration of lower mass jet to higher jet mass bin and vice versa as can be seen in Figure F.2.

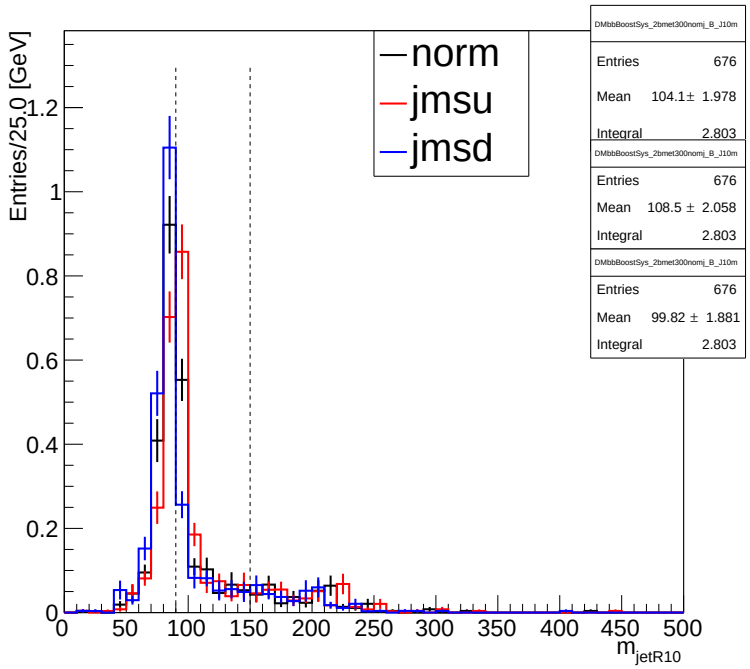


Fig. F.2 Leading large-R jet mass distribution in SR with $E_{\mathrm{T}}^{\mathrm{miss}} > 300$ *GeV* for MC Diboson sample. All event selection criteria are applied except jet mass cut. Two vertical dotted lines corresponds to the boundary of the jet mass window. Three distibution with systematic variation up (red), variation down (blue) and nominal (black) are shown.