LUND UNIVERSITY
Faculty of Science

Derivation of a transfer function to estimate the $Z(\nu\bar{\nu}) + \text{jets}$ background from $\gamma + \text{jets}$ in a search for dark matter at the ATLAS detector

Emelie Olsson

Thesis submitted for the degree of Master of Science

Project duration: 12 months

Supervised by Ruth Pöttgen and Torsten Åkesson

Department of Physics
Division of Particle Physics
May, 2019

Abstract

It has for long been known that dark matter must exist in our universe. Despite massive efforts to detect dark matter, its constituents remain a mystery. One strategy to unveil this mystery is to produce dark matter by colliding Standard Model particles. The most common signature of dark matter production is an excess of missing transverse energy ($\not{E}_{\mathrm{T}}$). Such an event would be very rare, which is why it is important to have a precise background estimation. One irreducible background in the search for dark matter is a Z boson decaying to two neutrinos. In this thesis, events with a photon are used to create a transfer function that can be used in the estimation of this background. First, events are selected with the same characteristics as the $Z(\nu\bar{\nu})$ background events, but with a photon instead of the Z. The two samples are compared by adding the photon $\mathrm{p_T}$ to the missing energy, to replace the missing transverse energy from the Z boson decaying to two neutrinos. The two samples show similar distributions for all studied variables. Next, the transfer function is created, by fitting the ratio of MC samples for the two bosons. The transfer function reflects the relation between the production cross sections and selection differences for the Z and γ and changes slightly with boson $\mathrm{p_T}$. It is for the range 500-2000 GeV determined to $0.28\text{-}5 \cdot 10^{-5} \not{E}_{\mathrm{T}}$ [GeV], with an uncertainty of 18%. To make a $Z(\nu\bar{\nu})$ background estimate the transfer function can be multiplied by γ + jets data. After further validation, this transfer function will be used in the background estimation of the $Z(\nu\bar{\nu})$ background for the mono-Higgs($b\bar{b}$) ATLAS analysis, aiming to improve its sensitivity in the full run II.

Acknowledgements

I would like to express my very great appreciation to Ruth, who has guided me every step of the way. You are probably the best role model a student can wish for.

My grateful thanks are also extended to Eleni, who has helped me countless times. The commitment from your side has been remarkable.

I am thankful for the support of the other master students. Having your company and sharing this experience the past year has been invaluable. A special thanks to Alexander, for always being helpful.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

For almost a century the existence of dark matter has been known. It was discovered by Knut Lundmark in 1930, who observed the velocities of stars in galaxies, realizing more matter was needed to explain gravitational forces on large scale. For a long time, this "Dunkle Materie" as he called it, was thought to be dark stars, planets, and clouds which could not be observed. Today we know it is a new type of matter, different from everything we can see in the universe.

The nature of dark matter is one of the biggest questions in physics, and massive efforts are put into solving the dark matter puzzle. It is a vital part of the quest to understanding the Universe. Over the years, many different dark matter candidates have been ruled out, and its constituents remain a mystery. We only know that dark matter has to be a new type of matter. The bad news is that we do not know if it is possible to determine the nature of dark matter since the only evidence we have now is gravitational. The possibility to detect dark matter and investigate its constituents relies on one important assumption: that dark matter and ordinary matter interact in more ways than through gravity. This is an assumption we are forced to make in order to search for dark matter constituents.

One strategy in the search for dark matter particles is to try to create them in highly energetic particle collisions. In such a collision, huge amounts of ordinary particles are created. To be able to distinguish any dark matter particles in this chaos of particles, comparisons to theoretical predictions are made. The predictions are based on the Standard Model, the theory that describes the elementary particles and their interactions. Any excess could be the detection of a new particle. Very precise predictions are needed since a dark matter particle will be extremely rare. Therefore, these background estimations are a crucial part in the search for dark matter.

For some processes, it is hard to use simulations from purely theoretical predictions in the comparison. Then data or predictions from other processes can be helpful. In this thesis, the aim is to derive a transfer function, used in the background estimation of one of the major irreducible backgrounds in the search for dark matter, the $Z(\nu\bar{\nu})$ + jets background. Improving this background estimate is important for the sensitivity of the dark matter search. Simulations from another process, γ + jets, will be used to derive the transfer function. This transfer function can be applied to γ + jets data to get the $Z(\nu\bar{\nu})$ + jets background estimation.

The derivation of the transfer function will be based on simulations with a pile-up profile from 2017. The transfer function will be determined specifically for the selection in the *Search for Dark Matter Produced in Association with a Higgs Boson decaying to* $b\bar{b}$ analysis [1], for missing transverse energy above 500 GeV. A region to select γ + jets events will be implemented, to use in the calculation of the transfer function. The method from this thesis for deriving a transfer function will be used in the mono-Higgs($b\bar{b}$) analysis of the full run II data (from 2015-2018), if the background estimation from the transfer function improves the uncertainties compared to previous estimations based on other processes. Other searches with the $Z(\nu\bar{\nu})$ + jets background and the same requirements on missing energy could use the transfer function derived in this thesis.

Chapter 2

Theory

This chapter revises basic theory and concepts that are needed for the remainder of this thesis. An overview is given of the Standard Model, shortly describing its particles and forces. Also, the existence of dark matter is motivated and some of its known properties are discussed.

2.1 The Standard Model of Particle Physics

The Standard Model is the theory that describes the elementary particles and how they interact. It is remarkably accurate in describing experimental results. One of the most famous examples is the Higgs boson, discovered in 2012 [2, 3], but theorized in the 1960's [4, 5, 6]. The Standard Model has fermions, which make up matter, gauge bosons, that mediate the forces with which fermions interact, and one scalar boson, the Higgs boson.

Fermions make up all the matter we can detect, such as neutrons or protons. The two types of fermions are called quarks and leptons, with six particles of each. Quarks are pointlike particles, and can only exist in combinations of two or three, while leptons can propagate on their own. In Fig. 2.1, quarks are given in red and leptons in green. The 12 fermions are paired in three generations, with two quarks and two leptons in each generation. The first generation consists of the two lightest quarks (up and down-quarks) and the electron and electron neutrino (ν_e). Only particles of the first generation are stable and are the building blocks of the matter we see in the Universe. All charged particles of the two heavier generations decay into particles of the first generation.

The gauge bosons, given in blue in Fig. 2.1, are the force carriers of three of the four fundamental forces. Gravity is not part of the Standard Model yet, and it is a great challenge to include it. The mediators of the strong force are the bosons called gluons. The strong force affects particles with color charge, namely quarks and the gluons themselves. It becomes asymptotically weaker as energy increases. This property is called asymptotic freedom. For low energies, the strong force is so strong that it leads to confinement. The energy required to separate two quarks is so large that a new quark pair is instead created [8]. Therefore quarks and gluons cannot exist on their own, but only in colorless

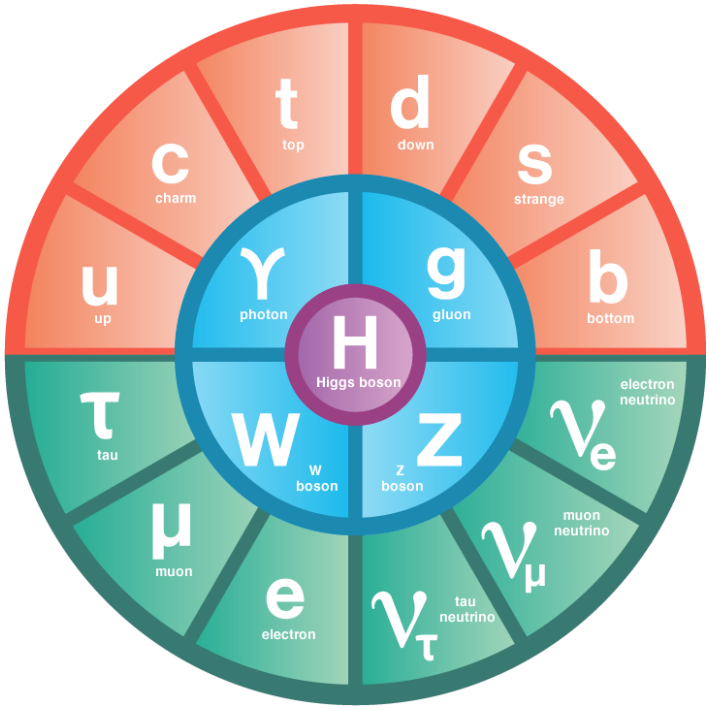


Figure 2.1: The Standard Model of particles. Image from [7].

combinations, called hadrons. There are two types of hadrons, mesons with color and anti-color (two quarks) and baryons with three different colors (three quarks). The quarks and gluons that make up a hadron are called partons.

The weak force is carried by the Z- and $W^{\pm}$-bosons. One of their important properties is that they are massive (91 GeV and 80 GeV respectively), making the range for the weak force very short. Both the Z and the $W^{\pm}$ decay into lighter particles. The weak force violates the charge conjugation and parity conservation laws, and also their combination, the so-called CP-violation [8]. A process and its CP conjugate both occur, but their probability is not equal. This has been observed in the weak decay of B- and K-mesons, particles consisting of a quark-antiquark pair where one of them is a bottom quark (B-meson) or a strange quark (K-meson). The neutral photon is the force carrier of the electromagnetic interaction, which affects charged particles. Since the photon is massless the force has infinite range.

Predicted by the Standard Model, and found in 2012, is the Higgs boson. It is an excitation of the underlying Higgs field, which is responsible for giving most of the elementary particles mass by the Brout-Englert-Higgs mechanism. The Higgs field is non-zero everywhere, and the strength of a particle's coupling to it determines its mass. Therefore, the Higgs boson couples stronger to heavier particles. The Higgs boson is itself at 125 GeV heavier than the other bosons.

For every particle in the Standard Model there exists an antiparticle, which has exactly the same properties but with opposite charges. It is often denoted with a bar above its corresponding letter, e.g. $\bar{u}$ for anti-u. A particle-antiparticle pair can annihilate, or be created from energy in e.g. collisions where the quantum numbers add up to zero. The neutral photon, gluon, and Z-boson are their own antiparticles.

2.2 Dark matter

The particles described by the Standard Model make up the ordinary (luminous) matter in the Universe. This is about 5% of the total mass-energy content of the Universe. The rest is dark matter (27%) and dark energy (68%) [9]. Dark matter interacts gravitationally with ordinary matter, and hopefully also with some additional force, allowing the detection of dark matter in a particle physics experiment. Dark energy is responsible for the accelerating expansion of the Universe and is not a form of matter. The only similarity between dark energy and dark matter is that we know both must exist to explain the Universe.

There are several convincing pieces of evidence for dark matter. If one considers only the ordinary matter in galaxies it is more dense towards the center, and it is therefore expected that the orbital velocities of the stars fall off further from the galactic center. This is not what is observed - instead, the velocities are constant for stars further from the center. To explain this, there has to be some additional matter smoothly distributed over the whole galaxy, and with a total mass larger than that of the ordinary matter. In Fig. 2.2 it is shown how a halo of dark matter would explain the velocity distribution observed in the NGC6503 galaxy.

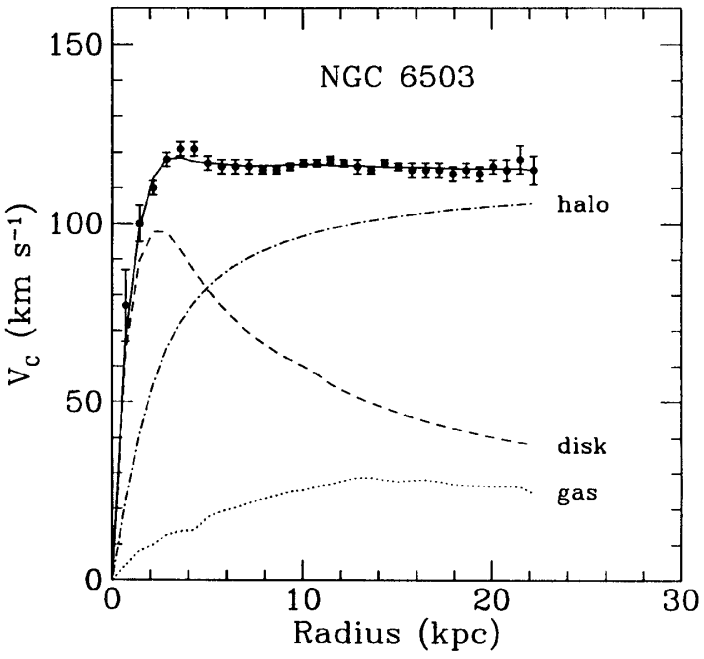


Figure 2.2: Galactic rotation curve for the galaxy NGC 6503. The dashed lines show the contributions from the disk and gas of the galaxy, and the dotted-dashed line is the dark matter halo needed to match the data. Image from [10].

Another example of evidence for dark matter is from astronomers observing a collision of two galaxy clusters in the Bullet Cluster. Since dark matter interacts with ordinary matter gravitationally, and collisions are at most extremely rare, its movement in the galactic collision differs significantly from that of ordinary matter. This can be seen when

comparing observations of the ordinary matter with data measuring the gravitational lensing. Also, the amount of gravitational lensing cannot be explained by only ordinary matter [11].

Signs of dark matter can also be found in the early Universe. After recombination, when the first atoms formed, it became possible for photons to propagate. Gravity pulled atoms together, but the pressure from the photons kept them apart, causing oscillations. These oscillations made it hard for any lumps to form. Dark matter would form lumps more efficiently since it is unaffected by the photon pressure. These lumps would grow and cause the large scale structure we observe in the Universe today. We can actually observe the photons from the recombination as the Cosmic Microwave Background (CMB), where we observe patterns caused by the lumps, confirming the existence of dark matter [12].

More recently, galaxies without dark matter have been found [13]. This is strong evidence for the existence of dark matter as a form of matter, instead of the gravitational theories being wrong. If gravity worked differently on large scales, there would be no exceptions to this. But if dark matter is some form of matter, some galaxies can have more than others.

There are some properties of the dark matter particle that are common in different models. It has to be electrically neutral since we know it does not interact electromagnetically. The particle also has to be non-relativistic, so-called cold, to be able to form the structures we see in the Universe today. If the dark matter is too hot, only large objects with enough gravitational impact would form first, and this is inconsistent with observations. Also, the particles should be stable or long-lived, the latter is simply having a lifetime longer than the age of the Universe. The dark matter particles need to be collisionless, meaning that they cannot have large self-interaction. Bounds on the self-interaction can be set by observing cluster collisions [11, 14].

A basic assumption about dark matter particles is that they have some kind of weak interaction with luminous matter. It would then have been possible for dark matter to form from and annihilate into luminous matter in the early Universe because the temperature was high enough. As the Universe expanded and cooled, only the annihilation of dark matter particles continued until the dark matter number density was too low. The amount of dark matter particles today is called the relic abundance, which is constant. From the relic abundance, constraints on the annihilation cross section and interaction with luminous matter for dark matter can be derived.

One prominent dark matter candidate is the Weakly Interacting Massive Particle (WIMP). As mentioned above, dark matter particles can be at most weakly interacting with luminous matter. This could be by the Standard Model weak force, or by some new, weaker interaction. The WIMP is assumed to interact with ordinary matter in some way, to make detection possible. The dark matter particles have to be massive, to give a gravitational impact. The WIMP is also assumed to move at non-relativistic velocities. This agrees better with the formation of the Universe because at some point the temperature in the Universe was so low that WIMPs could no longer form from lighter particles. But the WIMPs could still annihilate and did so until their density was low enough, giving a roughly constant amount of WIMPs today. Cold WIMPs would then clump together, instead of being too fast and escape gravity's pull. Combining the WIMPs' annihilation cross section and velocity matches the cross section for the weak force. New physics at

the electroweak scale could then give particles, which also are dark matter candidates. This coincidence is called the WIMP Miracle [15].

Other examples of dark matter candidates are sterile neutrinos and axions. The Standard Model neutrino is weakly interacting and stable, without any electromagnetic or strong interaction. These are important criteria fulfilled, but the Standard Model neutrino is too light to explain the large scale structure of the Universe [15]. A heavier, sterile neutrino could be a dark matter candidate, with sterile meaning not interacting weakly either. The sterile neutrino would have a coupling to the Standard Model neutrinos, enabling detection. The axion is a very light particle that appears when introducing a new symmetry to explain the very small effect of CP-violation in QCD [15]. It would have a mass around the μeV scale and is detectable by the decay to two photons.

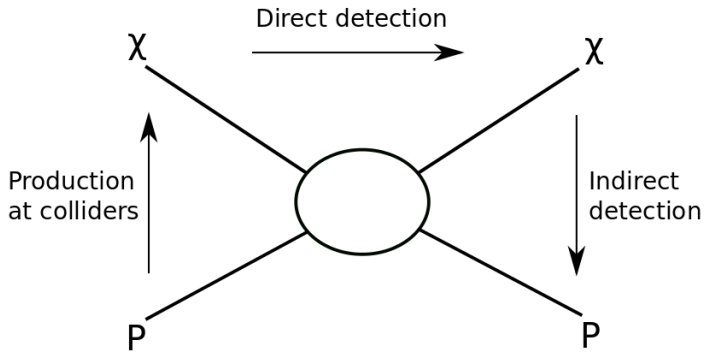


Figure 2.3: Detection methods for dark matter particles. Image from [16].

Today, researchers try to detect WIMP dark matter particles with three major strategies, see Fig. 2.3. Direct detection experiments search for the scattering of dark matter particles on ordinary nuclei, in large underground detectors. Indirect detection instead focuses on detecting the Standard Model particles created when dark matter particles annihilate. The collider search does the opposite, aiming to produce dark matter out of Standard Model particles collided at very high energies.

2.3 Dark Matter Simplified Model

A theoretical connection between Standard Model particles and dark matter is needed to enable production in proton-proton collisions. In the so-called "simplified model approach" one introduces new heavy mediators as a degree of freedom [17]. In the two-Higgs-doublet model, an extra U(1) gauge symmetry gives a heavy boson Z'. From the two Higgs doublets, five different Higgs bosons are derived: h (the Standard Model Higgs), a scalar H, a pseudoscalar A and two charged scalars $H^{\pm}$. The h is the lightest one. The production mechanism is shown in Fig. 2.4, where Z' couples to both the initial state and decays into a Standard Model Higgs boson h and a pseudoscalar Higgs boson A. In short, the signal we are looking for is $pp \to Z' \to Ah$. A then decays to two Dark Matter particles ($\chi\bar{\chi}$), leaving the detector without a trace, and the h decays into $b\bar{b}$.

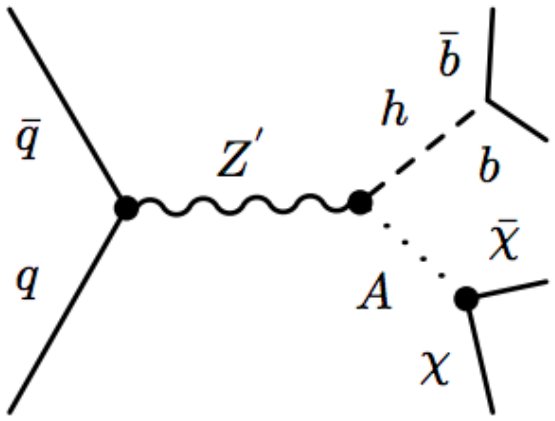


Figure 2.4: Feynman diagram of the production process, where the Z' couples to the initial state and both Higgs bosons, A and h. The Higgs bosons then decays to $\chi\bar{\chi}$ and $b\bar{b}$, respectively. Image from [18].

There are five parameters of the model. Three of them are the masses $m_{Z'}$, m_A, and m_χ. Also, tanβ, which is the ratio of the vacuum expectation values for the two Higgs fields (their coupling to up-type and down-type quarks), and $g_{Z'}$, which is a parameter in the mixing between Z' and the Standard Model Z boson. A is assumed to be produced on shell ($m_A/2 \geq m_\chi$), and to have a large branching ratio to $\chi\bar{\chi}$. tanβ has constraints from electroweak and dijet searches, and $g_{Z'}$ is fixed to its upper limit as a function of Z' mass (95 % confidence level) [19].

2.4 Z/γ ratio

The aim with the transfer function derived in this thesis is to estimate one major irreducible background in the dark matter search, the $Z(\nu\bar{\nu})$ + jets background. Using a transfer function based on the ratio between $Z(\nu\bar{\nu})$ + jets and γ + jets MC samples, γ + jets data can be reweighted to estimate the $Z(\nu\bar{\nu})$ + jets background.

The ratio between MC samples will be given by the number of $Z(\nu\bar{\nu})$ + jets events divided by the number of γ + jets events. Then, the transfer function will basically be the cross-section ratio of the production of the two bosons and their selection differences, times the branching ratio to $\nu\bar{\nu}$ (20%) for the Z boson. Fig. 2.5 (a) and (b) shows two Feynman diagrams for the production of a Z or γ and one jet. At high $\not{E}_{\mathrm{T}}$, meaning high boson p_{T}, the mass difference between Z and γ is negligible and the difference in production cross section will come from the quark coupling differences. Based on the quark coupling differences, an approximate value of the transfer function and its dependence on $\not{E}_{\mathrm{T}}$ will be determined below.

Starting from the relevant part of the Lagrangian, the interaction between Z bosons and fermions (the neutral current) is given by

$$\frac{g_W}{\cos\theta_W}[\bar{f}_L\gamma^\mu f_L(T_3 - Q_f \sin^2\theta_W)) + \bar{f}_R\gamma^\mu f_R(-Q_f \sin^2\theta_W)],$$

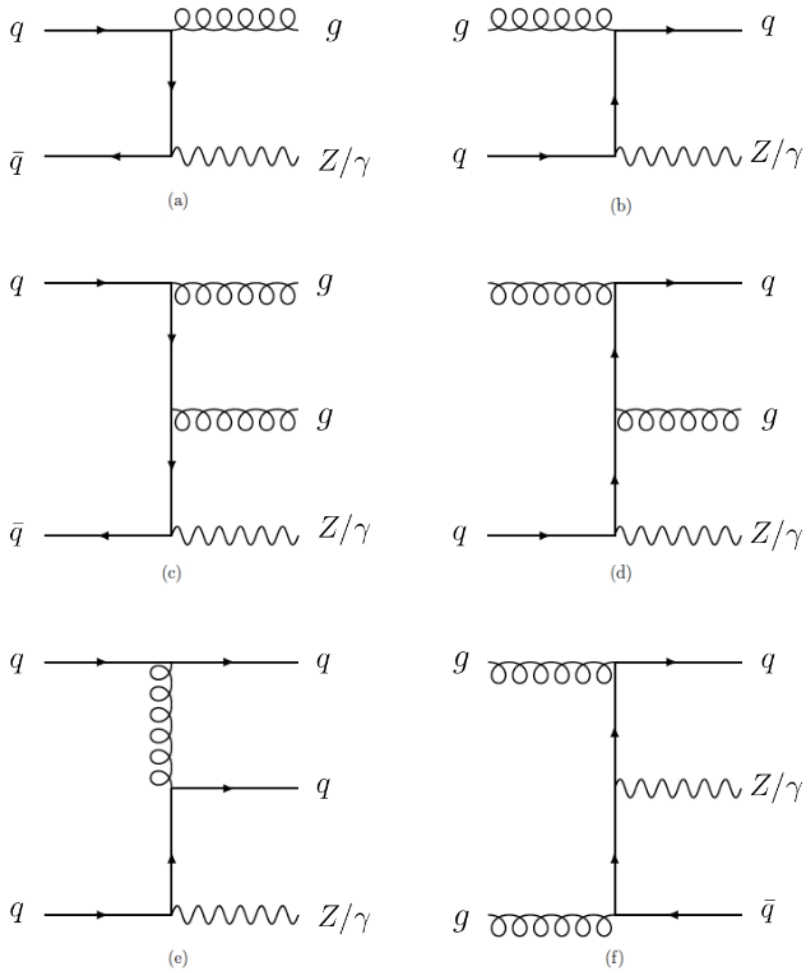


Figure 2.5: Feynman diagrams of the production of one boson (Z or γ) and one (a and b) or two jets (c-f).

where g_W is the weak coupling, θ_W the weak mixing angle, f_L and f_R the right and left handed parts of the fermion spinors, T_3 the weak isospin and Q_f the fermion charge [8].

Rewriting this expression using $v_f = T_3 - 2Q_f \sin^2\theta_W$ and $a_f = T_3$ (vector and axial neutral weak couplings), and $g_W = e/\sin\theta_W$, gives

$$\frac{e}{2\sin\theta_W\cos\theta_W}[\bar{f}\gamma^\mu(v_f - a_f\gamma^5)f].$$

The factor then associated with the fermion vertex is

$$\frac{-ie}{2\sin\theta_W\cos\theta_W}[\gamma^\mu(v_f - a_f\gamma^5)].$$

For photons, the part of the Lagrangian corresponding to interactions between photons and fermions is

$$eQ_f(\bar{f}\gamma^\mu f),$$

and the factor associated with the vertex is $-ieQ_f\gamma^\mu$.

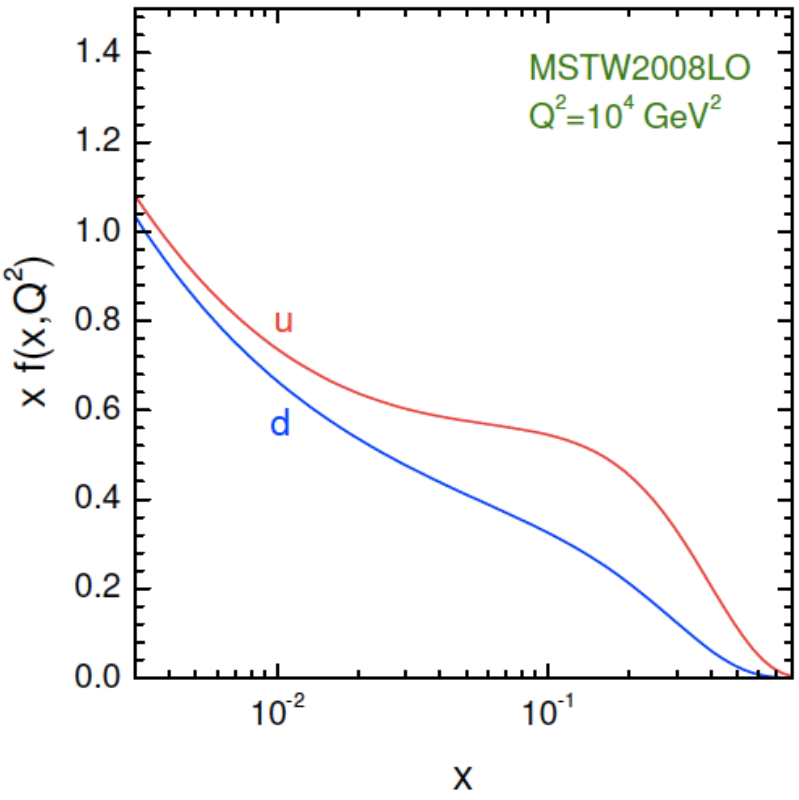


Figure 2.6: Parton distribution functions at Q = 100 GeV for u and d quarks. Figure from [20]

To get the cross section, the matrix element that includes the vertex factor is squared. Only including the vertex factors in the matrix element is as mentioned an approximation, since the mass difference is neglected. At high p_T, the parts depending on the mass difference should anyway cancel. The Z and γ production cross section ratio is then given by the ratio of the fermion couplings, seen in Eq. 2.1.

$$R_f = \frac{v_f^2 + a_f^2}{4\,Q_f^2 \sin^2\theta_W \cos^2\theta_W} \tag{2.1}$$

For this thesis, only the quark couplings are interesting, since it is by these couplings the bosons are produced (see Fig. 2.5). The ratio R_f will be different depending on quark flavour. If only u-type quarks are present (u, c, t), $v_u = 1/2 + 4/3\sin^2\theta_W$, $a_u = 1/2$ and $Q_u = +2/3$, giving $R_u = 0.906$ with $\sin^2\theta_W = 0.2315$. For d-type quarks (d, s, b), the corresponding values are $v_d = -1/2 + 2/3\sin^2\theta_W$, $a_d = -1/2$ and $Q_d = -1/3$, giving $R_d = 4.673$.

In reality, the quark-coupling ratio in Eq. 2.1 will get contributions from both quark types, and will therefore be a weighted average of R_u and R_d. How often a specific quark type will take part in the boson production will be given by the parton distribution functions of the proton. A parton distribution function, $f(x, Q^2)$, is the probability density for finding a parton with momentum fraction x at resolution scale Q^2. Fig. 2.6 shows the parton

distribution functions of the proton, for u and d quarks, at Q = 100 GeV. Assuming $x = 0.1$ gives $d/u = 0.6$ and an average quark coupling ratio of ≈ 1.4 [20]. Defining R_{avg} as the average quark coupling ratio (depending on x), including the 20% branching ratio of Z to two neutrinos, gives $R_{avg} \approx 0.28$ for $x = 0.1$.

As x increases, the d/u ratio will decrease, meaning that we will get more u-quarks compared to d-quarks. Since R_u is lower than R_d, the weighted average R_{avg} will also decrease. Higher x-values are selected by requiring higher p_T of the vector boson, since then only partons with higher momentum fractions take part in the boson production. This means that R_{avg} decreases for higher boson p_T. Higher p_T corresponds to higher $\not{E}_T$, giving that R_{avg} decreases with increasing $\not{E}_T$.

To summarize the discussion, R_{avg} is based on the quark coupling differences for Z and γ, averaging depending on the PDF's of the proton, and includes the 20% branching ratio of the Z boson decaying to two neutrinos. R_{avg} is expected to have a value around 0.28 and to decrease with increasing $\not{E}_T$. It will later in this thesis be used as an theoretical approximation of the transfer function.

The theoretical results above are valid for the production of a Z or γ together with one jet. When including more jets, additional diagrams (see eg. Fig. 2.5 (c)-(f)) will influence the ratio. It is then also possible to have a collinear singularity for γ when it is emitted close to one other jet. Allowing the γ and the jet to get close will increase the matrix element, and decrease the Z/γ production ratio. There is no collinear singularity for the Z boson since it is massive.

Chapter 3

Experimental facilities

The ATLAS detector is one of four major experiments at the Large Hadron Collider (LHC). It is a multi-purpose detector, aiming to detect all kinds of Standard Model particles in all directions. It is 46 meters long and weighs 7000 tonnes. When the Higgs boson was discovered in 2012, ATLAS and CMS (another of the main experiments at the LHC) were the experiments that reported signs of the new elementary particle. About 3000 physicists around the world work on the ATLAS collaboration [21].

3.1 Large Hadron Collider

The LHC is the world's largest particle accelerator, in operation since 2008. It is circular, almost 27 km in circumference. Protons are pre-accelerated and then transferred to the large ring, where superconducting magnets guide protons. They are further accelerated with radiofrequency cavities, almost to the speed of light. They travel around the ring in two separate beamlines, one beam going in each direction. The beams are not continuous, but consist of bunches of particles. With additional magnets that squeeze the beams together, these bunches are forced to collide.

Depending on the number of bunches in each beamline and how many particles are in each bunch, the number of collisions will vary. Only the most energetic parton collisions are of interest, called hard scattering. Additional parton collisions in the same proton are called underlying events, and less energetic proton-proton interactions in the same bunch crossing are called pileup.

The amount of possible interactions is called instantaneous luminosity and it is a property of the collider. It is combined with the cross section for a process to get the collision rate. Higher instantaneous luminosity is advantageous since it means more statistics, but it also means more pile-up. Run II at the LHC, with data recorded from 2015 until 2018, had a collision energy of 13 TeV and a total integrated luminosity of 146 fb^{-1} for proton-proton collisions.

3.2 ATLAS detector structure

In order to detect as many particles as possible, the ATLAS detector has several layers. Each layer is designed to detect a certain type of particles and to help distinguish them from others. The main parts are the inner detector, calorimeters, muon spectrometer and magnet systems, see Fig. 3.1. To get a full picture of the collision, the detector covers all angles except for where the beamlines enter. A cross section of the ATLAS detector is seen in Fig. 3.2.

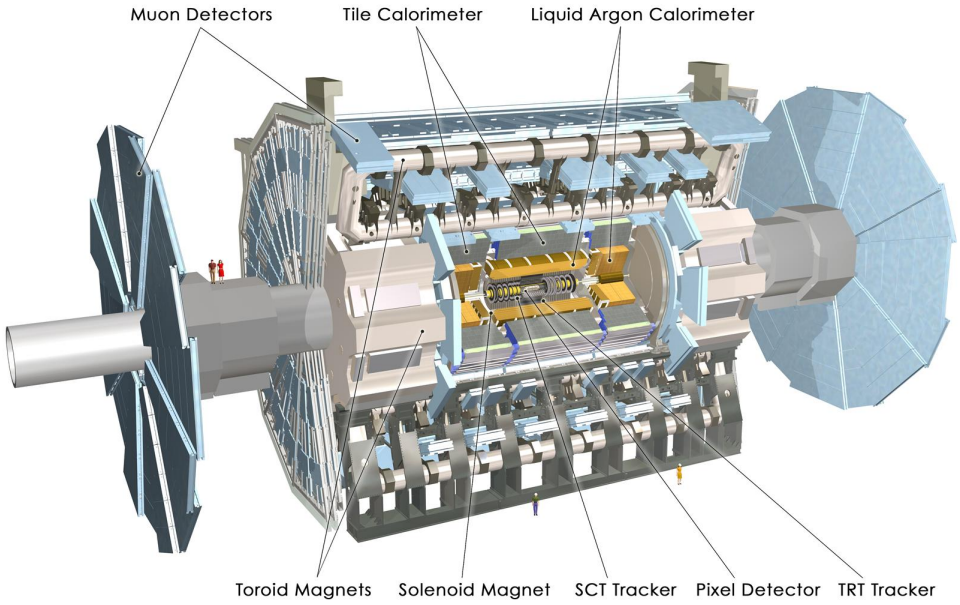


Figure 3.1: Illustration of the ATLAS detector, where the main parts are shown. Image from [22].

The inner detector starts a few centimeters from the beam line, and its main purpose is to track charged particles. It is surrounded by a solenoid magnet, producing a 2 T magnetic field in the inner detector. The magnetic field bends the trajectories of the charged particles, making it possible to calculate their momentum. Particles are tracked by their interactions with the detector. Innermost in the inner detector are pixel detectors, with 80 million readout channels. This is half of the readout channels in all of ATLAS, needed to get as precise tracking as possible close to the collision center. Outside the pixel detector is a semiconductor tracker, which has strips instead of pixels, and each readout covers a larger area. These strips are important for the tracking perpendicular to the beam. Outermost in the inner detector is the transition radiation tracker. It has long drift tubes filled with gas, and between them other materials with varying refractive indices. Particles with high velocities will emit radiation when they pass a boundary between two materials with different refractive indices. The amount of radiation depends on the γ-factor of the particle, and thereby electrons and positrons can be identified since they are the lightest of the charged particles.

The two calorimeters are situated outside the solenoid magnet. Their purpose is to measure the energy of particles by absorbing them. This happens in a series of interactions

with the calorimeter material, giving a particle shower in the detector. The calorimeters have good spatial coverage to make a complete measurement of the energy. Innermost is the electromagnetic calorimeter, also called the Liquid Argon Calorimeter in Fig. 3.1. It has absorbing plates made of lead and stainless steel, with liquid argon in between. Here, the energy of photons and electrons are measured. They can be separated by the fact that the electron left a track in the inner detector, while the photon did not since it is not charged. Outside the Liquid Argon calorimeter is the hadronic tile calorimeter, consisting of less expensive steel and plastic scintillator plates. It measures the energies of hadrons, which are energetic enough to pass through the Liquid Argon calorimeter. If a gluon or quark is created in the proton collision it will form hadrons along its trajectory, a process called hadronization. This hadron shower is seen as a cone in the calorimeter, called a jet.

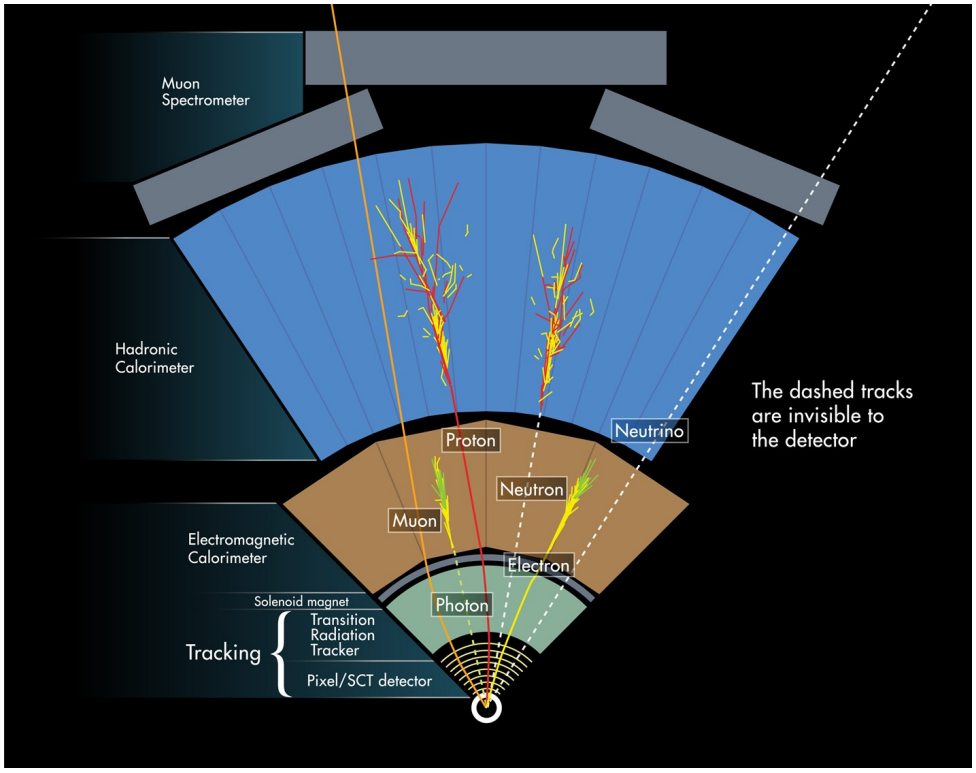
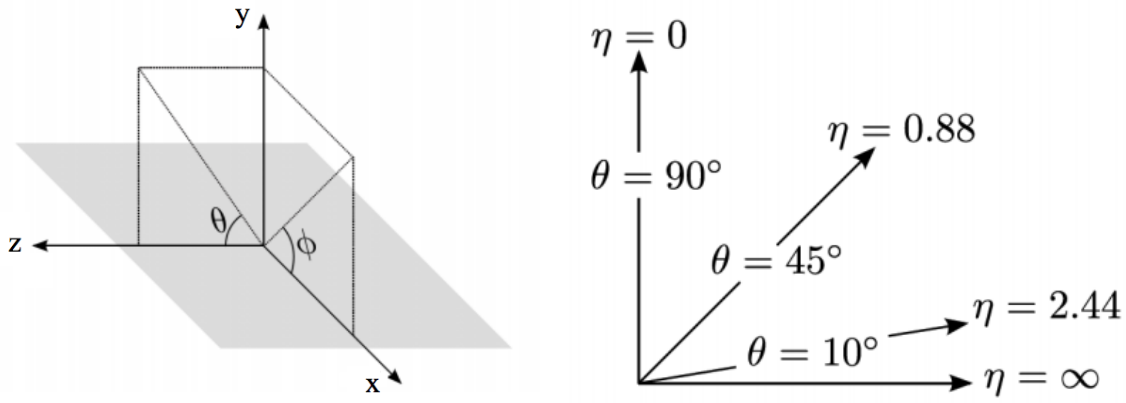


Figure 3.2: Cross-sectional view of the ATLAS detector, showing where in the detector different particles are visible. Image from [23].

Surrounding all of the above mentioned parts is the muon spectrometer. It has its own magnetic field from 3 toroidal magnets, 1200 tracking chambers for the spatial resolution and triggering chambers for the time resolution. Few particles reach the muon spectrometer, of which most are muons. However, as the energy in the collisions increases, it becomes more probable for other particles than muons to reach the muon spectrometer. The muons are not stopped by the spectrometer, but their momentum is measured since the magnetic field bends their trajectories. A muon should leave a trace in all of the detector parts within, and can therefore be identified.

The opposite is true for the neutrino, it only interacts weakly with other particles and leaves no trace in any part of ATLAS. Instead, conservation of momentum is used to



(a) (b)

Figure 3.3: The coordinate system used for the ATLAS detector in (a) and an illustration of the relation between θ and η in (b). Figures from [24].

calculate the energy that is missing. Before the collision, the incoming protons only have momentum along the beam, and the transverse momentum component, p_T, should remain zero when the transverse momentum is added up after the collision. The amount of missing transverse energy in the calorimeters (combined with the muon energies) then gives the energy carried away by neutrinos. Another possibility is the existence of other particles outside the standard model, also invisible to the detector, responsible for the missing energy. If it is possible to create dark matter in collisions at the LHC, the particles would not be detected by ATLAS, but give missing transverse momentum.

The ATLAS detector has a defined coordinate system, used to describe the directions in which the particles move. The z-axis is in the direction of the beam, the x-axis towards the center of the LHC and the y-axis vertically upwards, as seen in Fig. 3.3(a). The azimuthal angle ϕ starts from the x-axis and goes in the xy-plane perpendicular to the beam direction. The polar angle θ is in the yz-plane and starts from the z-axis. θ is rarely used when describing the directions of particles. Instead, η (pseudorapidity) is used, defined in Eq. 3.1. The relation between θ and η is visualized in Fig. 3.3(b).

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{3.1}$$

As mentioned above, directions in the ATLAS detector are given in ϕ and η. The distance from one particle to another, ΔR, is given by Eq. 3.2.

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{3.2}$$

3.3 Data flow

Every second about 40 million bunch crossings take place in the ATLAS detector. This creates an enormous amount of data, and it would be impossible to store every event. In

the end, about 1 in 400 000 events are saved.

The first selection, where most events are discarded, is done by the level 1 trigger. Since it is not possible to first store all the events before deciding which ones to keep, the first step has to be fast. It is based on a quick readout from the calorimeters, that has reduced granularity. The trigger works with simple high p_T-triggers, for example high p_T electrons, photons or jets, or high missing p_T. The trigger options can be used in different combinations, either in coincidence or veto.

The level 2 trigger (or High Level Trigger, HLT) works with all the data for each event selected by the level 1 trigger. It has position information for particles and data from the muon spectrometer. For example, particles can be identified with more certainty. The track of an isolated electron in the calorimeter can be matched with a track in the inner detector. Also, the $\not{E}_T$ measurement will be refined with data from the muon spectrometer.

Next, the data is reconstructed into objects. Data from the different parts of the detector is combined to give particles, e.g. charge, momentum and path. A muon reveals itself by being visible in the tracker, through the calorimeters and in the muon spectrometer, where one can determine its momentum by the curvature of the track. Jets can be reconstructed using different algorithms, that determine which energy depositions belong in the same jet. One widely used method is the anti-k_T jet clustering algorithm. It will let the hard (high momentum) particles shape the jet, avoiding clustering of soft particles. The anti-k_T algorithm is therefore good at resolving jets, but not so good at de-clustering, making it a poor choice for jet substructure studies [25].

The data is stored in a format called AOD (Analysis Data Object) and consists of all reconstructed particles and all available variables for every recorded event. It is possible to do the analysis on these AOD files directly, but it is more efficient to have a smaller file. By selecting only certain clean events (skimming), retaining only specific objects within these events (thinning) and only keeping the useful variables for these objects (slimming), the DAOD (Derived Analysis Data Object) is created. It consists of only the most valuable information, very specific for the analysis. At last, the information in the DAOD is restructured into a list, called ntuple. It has one event and the corresponding information on each row. This is very convenient for the analysis where it is easy to loop through these ntuples [26].

3.4 Monte Carlo simulations

The data from the ATLAS detector is compared to computer-generated collision events, Monte Carlo (MC) simulations, based on Standard Model predictions. It is crucial for the analysis that these simulations are accurate, to discover any excess events and set limits for unknown processes. The MC samples are created by generating collision events, where every possible outcome has a certain probability. It simulates both the hard process, parton showers, hadronization and possible decay of short-lived particles. In addition to this, it is possible that more than one parton collision occurs per proton-proton collision, called underlying event. The pileup is also simulated, but since the MC sample is produced before the data taking period, it is not known how much actual pileup there will be in

the data. Hence, the pileup simulation might not be correct. In that case, it is reweighed in the ntuple creation.

Next, the MC sample goes through a simulated detector, to account for how the particles interact with the matter in the detector. The ATLAS detector simulator is based on GEANT4 [27]. After going through GEANT4, the simulated data sample is processed in the same way as the data. The information in the primary generated sample, before going through the detector simulator, is called truth level information. This truth level differs from what is reconstructed in the AOD from the simulated detector output and gives a measure of how efficiently the reconstruction process works.

The MC samples in Table 3.1 will be used in this thesis. These MC samples are created with the Sherpa event generator [28]. They are all MC16d samples, which means they have been simulated with 2017 data taking conditions, and the pile-up profile for 2017 will be used when processing these samples. Both the γ + jets and $Z(\nu\bar{\nu})$ + jets are next-to-next-to-leading order (NNLO), meaning that second order corrections to the Feynman diagrams are included. The γ + jets samples are divided into five different samples depending on the p_T of the γ. The $Z(\nu\bar{\nu})$ + jets samples are max(H_T,p_T^V) filtered, and are also filtered for different quark flavors for the samples up to 500 GeV. The MC samples will be processed using the XAMPPmonoH framework (v.0800) [29].

Table 3.1: Sample list showing the DSID (dataset ID), name, generator and cross section for the samples used in this thesis.

DSID	Sample name	Generator	σ [pb]
364543	SinglePhoton_pty_70_140	Sherpa 2.2.2	4526.5
364544	SinglePhoton_pty_140_280		376.04
364545	SinglePhoton_pty_280_500		21.851
364546	SinglePhoton_pty_500_1000		1.4651
364547	SinglePhoton_pty_1000_E_CMS		0.029794
364142	Znunu_MAXHTPTV0_70_CVetoBVeto	Sherpa 2.2.1	10706.0
364143	Znunu_MAXHTPTV0_70_CFilterBVeto		10705.0
364144	Znunu_MAXHTPTV0_70_BFilter		10705.0
364145	Znunu_MAXHTPTV70_140_CVetoBVeto		592.36
364146	Znunu_MAXHTPTV70_140_CFilterBVeto		606.9
364147	Znunu_MAXHTPTV70_140_BFilter		607.97
364148	Znunu_MAXHTPTV140_280_CVetoBVeto		211.55
364149	Znunu_MAXHTPTV140_280_CFilterBVeto		222.31
364150	Znunu_MAXHTPTV140_280_BFilter		222.38
364151	Znunu_MAXHTPTV280_500_CVetoBVeto		47.423
364152	Znunu_MAXHTPTV280_500_CFilterBVeto		47.401
364153	Znunu_MAXHTPTV280_500_BFilter		47.421
364154	Znunu_MAXHTPTV500_1000		9.9101
364155	Znunu_MAXHTPTV1000_E_CMS		0.818

Chapter 4

Search for dark matter produced in association with a Higgs boson decaying to $b\bar{b}$

In most cases, the search for dark matter at colliders means searching for missing energy. To get a clear signature of missing energy, events are selected depending on some additional process that is well known, which the dark matter particles should recoil against. The search described in this chapter is focused on detecting dark matter simultaneously with a Standard Model Higgs boson (h) decaying to $b\bar{b}$. This gives two signatures: a large amount of missing transverse energy ($\not{E}_{\mathrm{T}}$) and one or two jets from the decaying h. The chapter is based on the dark matter search performed by the mono-Higgs ATLAS analysis group. Recent publications show that data agree with Standard Model predictions, setting tighter limits on the production of dark matter in association with a Higgs boson [1].

4.1 Analysis strategy

The reason for wanting a h recoiling off the DM, instead of W, Z or γ for example, is that the Higgs boson is unlikely to be initial state radiation. This is a result of its weak coupling to the quark in the initial state. A detected Higgs is therefore much more likely to be emitted by the second vertex in Fig. 2.4. The h decay to $b\bar{b}$ has the highest branching ratio among the h decay modes and is therefore chosen. Also, it is possible to identify b-jets in the ATLAS detector. Hadrons containing a b-quark have a lifetime long enough for them to travel a small distance from the collision center, remaining in the tracker when they decay. If a jet originates from a place other than the collision, the so-called secondary vertex, this can be where a b-hadron decayed.

The event selection starts with a trigger on $\not{E}_{\mathrm{T}}$. The amount of $\not{E}_{\mathrm{T}}$ determines how the Higgs candidate is reconstructed. The two b-jets from the h decay are considered resolved for $\not{E}_{\mathrm{T}} < 500$ GeV, each with a small radius (small-R). In this resolved region, events that have at least two central small-R b-tagged jets are selected. The mass of the Higgs is then

reconstructed using the two jets with the highest p_T. For high $\not{E}_T$ (> 500 GeV) the jets will not be separated but seen as one jet with a larger radius (large-R). This is called the merged region, where events with at least one large-R jet with two b-tagged associated tracks are selected. Only the highest p_T jet is used in the reconstruction of the Higgs mass [18].

Finally, the spectra of the reconstructed Higgs masses are examined in search of any excess events. In the case of no excess, upper limits on the cross section times branching ratio for the process are set.

4.2 Background processes

The $\not{E}_T$ + jets signature is not limited to the Z' decay seen in Fig. 2.4. Standard Model processes that give the same signature are called background processes. The estimated contribution from them is vital to the analysis, and better background estimates mean more precise searches for new physics. Background estimations are often based on the MC simulations described in Section 3.4. But for some background processes, the MC simulations have large uncertainty, and other methods are used to constrain them.

For the $\not{E}_T$ + jets signature, the main backgrounds are $t\bar{t}$ production and V + jets (where $V = Z$ or W), but there are also Standard Model Higgs, multijet and diboson backgrounds. The t-quark decay $t \to b + W^+ \to b + l^+ + \nu$ (and the corresponding process for $\bar{t}$) gives the signature if the leptons are missed in the detector. If it is not missed, a lepton veto ($e/\mu/\tau$) will remove the events to suppress the background. To reduce the $t\bar{t}$ background further, events with b-tagged jets outside of the Higgs candidate jets are vetoed.

The Standard Model Higgs background is made from events where a h and a vector boson are produced, where $Zh \to \nu\bar{\nu}b\bar{b}$, $Wh \to l\nu b\bar{b}$, or $Zh \to l^+l^-b\bar{b}$. The contributions from these processes are estimated using MC simulations. Also to model the diboson background MC events are used. This background consists of the production of either ZZ, ZW or WW, where one of the bosons decays leptonically and the other hadronically. Multijet events can also contribute to the background, if the energy of one or more jets is incorrectly measured, giving $\not{E}_T$. This makes high $\not{E}_T$ multijet events unusual since large measurement errors are required. The multijet production at lower energy is very high. MC simulation models multijet events well, but the amount of simulations needed to get an estimate for higher $\not{E}_T$ is too big. There are simply not enough computing resources. A data-driven estimate from a multijet-enriched region is instead used.

For the V+jets background, the decay of the boson can give one or two neutrinos, either by $W^\pm \to l^\pm + \nu$ or $Z \to \nu\bar{\nu}$. With a jet or two on the recoiling side (coming from a quark or gluon) it gives the $\not{E}_T$ + jets signature. The V+jets background is estimated using a so-called control region (CR). In this region, there are no signal events, but it should otherwise be as similar as possible to the signal region (SR). For this analysis, the SR has a lepton veto, and the CR's have a requirement of 1 or 2 leptons (e/μ). It should here be mentioned that it is assumed that the Z' does not couple to leptons, but to the other bosons. Then the W + jets CR is given by a 1 lepton selection to pick out the events with

$W^{\pm} \rightarrow l^{\pm} + \nu$, giving an estimate of the background. For the $Z(\nu\bar{\nu})$ + jets background, it is not possible to measure the neutrinos directly. The background estimate in the Run II analysis 2015-2017 was derived from a 2 lepton control region, selecting $Z(\mu^+\mu^-)$ events. Then the difference in branching ratio for the $Z(\nu\bar{\nu})$ and $Z(\mu^+\mu^-)$ processes was used to get the background estimate ([1], [18]). In this thesis, γ + jets events will be used to create a transfer function that can be used in the $Z(\nu\bar{\nu})$ + jets background estimate.

Chapter 5

Method for deriving the transfer function

In this chapter, it is described how the transfer function for the estimation of the $Z(\nu\bar{\nu})$ + jets background is derived. The idea is to use MC samples to create the transfer function, that can be used to reweight γ + jets data to estimate the $Z(\nu\bar{\nu})$ + jets background. The transfer function is given by a fit to the ratio of the missing energy from MC samples of the two processes. At the beginning of this chapter, motivation is given to why γ + jets events are used. Next, the object selections used in the analysis are defined, and the event selections on the different MC samples are listed. Finally, the transfer function is defined and explained, and the systematic uncertainties discussed.

5.1 Motivation

$Z(\nu\bar{\nu})$ + jets is one major irreducible background in the search for dark matter. It is an experimentally challenging background because of the neutrinos, for which it is not possible to measure the individual values of the energy or momentum, only the combined $\not{E}_{\mathrm{T}}$. In the Standard Model, only neutrinos contribute to the $\not{E}_{\mathrm{T}}$, but the measurement is dependent on all other measurements in the calorimeters. It is therefore likely to have contributions from mismeasurements of jets.

The signal region in this search, having large $\not{E}_{\mathrm{T}}$ and multiple jets, explores an extreme part of phase space. This requires including many orders of corrections in the MC samples to model the region well, but the MC samples have limits to the number of corrections. Also, this region of phase space is dominated by statistical uncertainties. It is then desirable to also use data in the background estimation. The transfer function itself is based only on simulations, but since it will be multiplied by γ + jets data to make the background estimate, it is called a semi data-driven estimate.

As stated in the above Section 4.2, the $Z(\nu\bar{\nu})$ + jets background has previously been estimated by extrapolating $Z(\mu^+\mu^-)$ + jets events using the branching ratio difference. One problem with using $Z(\mu^+\mu^-)$ is the lower branching ratio than the decay to two neutrinos, meaning even fewer statistics. $W(l\nu)$ can also be used, it has better statistics,

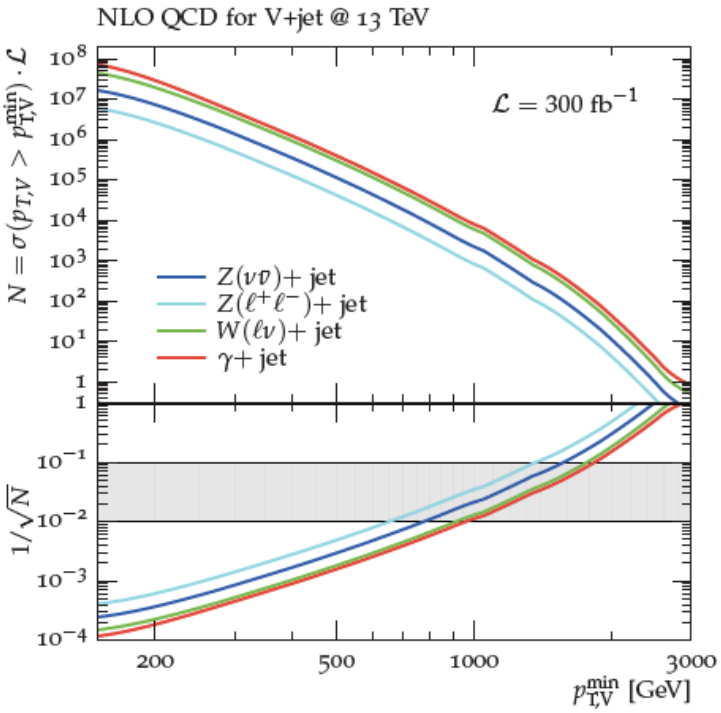


Figure 5.1: Production rates for V + jets processes, for minimum vector boson p_T, normalized to 300 fb^{-1} of LHC data at $\sqrt{s}$ = 13 TeV. Lower panel showing the statistical uncertainties, where the grey band is 1-10%. Image from [30].

but instead a larger theoretical uncertainty when going from the production process of one boson to another.

Another option is to use γ + jets, which has about a factor of 10 more events than $Z(\mu^+\mu^-)$ and marginally more than $W(l\nu)$(see Fig. 5.1). Therefore, it would be the best option if the theoretical uncertainties can be constrained [30]. For high p_T, the mass of the Z boson can be neglected and the production processes for Z and γ become similar. The difference in production rate for Z and γ then comes from the electroweak couplings. It is natural to analyze the cross-section ratio in the p_T parameter, as higher p_T means similar production processes and that the ratio of the two becomes almost constant. Only the region with boson p_T ($\not{E}_T$) > 500 GeV will be regarded in the background estimation in this thesis.

5.2 Object selection

Here follow the definitions of the objects used in this analysis, namely jets, leptons, photons, and $\not{E}_T$. Also, a short description of the overlap removal between these objects is given. All definitions but the photon are the same as in the most recent publication from the *Search for Dark Matter Produced in Association with a Higgs Boson decaying to* $b\bar{b}$ analysis group [31].

There are three types of jets used in this analysis: small-R jets, large-R jets, and variable-

radius (VR) track jets. The small-R jets are reconstructed with a radius parameter of 0.4 from energy deposits in the calorimeters, using the anti-k_T algorithm. Central small-R jets are required to have $|\eta| < 2.5$ and $p_T > 20$ GeV. To reduce jets from pile-up, small-R jets with $|\eta| < 2.4$ and $p_T < 60$ GeV have to be originating from the primary vertex. Small-R jets with $2.5 < |\eta| < 5.5$ and $p_T > 30$ GeV are defined as forward jets. Small-R b-jets are identified using an algorithm with 77% average efficiency [31].

For VR track jets the anti-k_T algorithm is used again, now using the tracks from the inner detector, required to have $|\eta| < 2.5$ and $p_T^{track} > 0.5$ GeV. In combination with the calorimeter information, the radius of the jet is determined to a value between 0.02 and 0.4, depending on its p_T. It is by definition also a small-R jet, and is used to improve the b-tagging. The large-R jets are reconstructed using the same anti-k_T algorithm as for the small-R and VR jets, with a radius parameter of 1.0. For these large-R jets, $|\eta| < 2.0$ and $p_T > 200$ GeV. To calculate the mass of the large-R jet information from the calorimeter is used, see Eq. 5.1. The sum is performed over all energy measured in the calorimeter within the jet.

$$m^{calo} = \sqrt{\Big(\sum_{i\in j} E_i\Big)^2 - \Big(\sum_{i\in j} p_i\Big)^2} \tag{5.1}$$

Only the baseline selection for leptons is described here since they are vetoed in all event selections relevant for this thesis. Electrons and muons are identified using loose identification and isolation requirements. For electrons, $|\eta| < 2.47$ and $p_T > 7$ GeV, reconstructed from deposits in the electromagnetic calorimeter only if a matching track can be found in the inner detector. Muons are reconstructed using tracks in the inner detector ($|\eta| < 2.5$) and the muon spectrometer ($|\eta| < 2.7$), also with $p_T > 7$ GeV. Both electrons and muons should originate from the primary vertex. The selection of taus uses both tracking and calorimeter information in a boosted decision tree discriminant, using a loose working point here as well. Taus should have $|\eta| < 2.5$ (excluding $1.37 < |\eta| < 1.52$) and $p_T > 20$ GeV.

Baseline photons are selected using the tight identification definition, $p_T > 25$ GeV and are required to have $|\eta| < 2.37$. The baseline selection is used for regions with a photon veto. Signal photons, used when selecting events containing photons, have additional requirements on tight isolation and $p_T > 150$ GeV.

The $\not{E}_T$ in an event is calculated as the negative sum of the p_T vectors of the reconstructed objects described above, including a soft term. All objects fulfilling the baseline selections are used in the calculation. The soft term consists of the p_T from tracks associated with the primary vertex, but not associated with any reconstructed objects. The $\not{E}_T$ is dependent on how well all other objects are measured in the detector and will depend on mismeasurements and resolution effects, so-called fake $\not{E}_T$. It is not possible to know how much of the $\not{E}_T$ is real and how much is fake. For γ + jets samples, the p_T of the photon will be removed from the negative sum in the $\not{E}_T$ calculation, and thereby be a part of the $\not{E}_T$ in the event.

If an object has been reconstructed as more than one object, it has to be decided which of these objects to keep in the event. This is called overlap removal and is done in the

same order for all events. Electrons are removed if they share tracks with an electron with higher p_T. If a tau is within $\Delta R = 0.2$ from an electron or muon, it is removed. Then, electrons sharing tracks with muons are rejected, and photons sharing tracks with electrons or muons are rejected. If an electron is close to a small-R jet (or VR track jet) the jet is removed. Small-R jets are removed if they are less than $\Delta R = 0.2$ from a muon, if the jet has less than three tracks or $p_T^{muon} > p_T^{jet}/2$ and $p_T^{muon} > 0.7 \cdot p_T^{associated\,tracks}$. Furthermore, if the small-R jet is not removed by the last selection, the muon is removed if it is close to the small-R jet (or VR track jet). The opposite is true for photon and small-R jet case, where priority is given to the photon if it is closer than $\Delta R = 0.4$ to a small-R jet. In addition, large-R jets are rejected in favor of electrons if they are closer than $\Delta R = 0.1$. Photons are also kept if they have less than $\Delta R = 0.1$ to a large-R jet. No overlap removal is done between small-R jets and large-R jets.

5.3 Signal region

As described in Section 4.1, the signature in this dark matter search is two b-jets and missing energy. By requiring this signature, with additional cuts to reduce backgrounds, the signal region is defined. It is divided into two subregions, the resolved and the merged, depending on the amount of missing energy. The reason for this is that in the merged region where $\not{E}_T$ is above 500 GeV, the Higgs boson is boosted and the two b-jets are seen as one large jet, leading to different selection requirements.

In Table 5.1 the skimming cuts for the signal region are listed. A skimming cut means that only if an event passes the selection it is stored in the ntuple. It is therefore not possible to compare samples without this selection applied. Table 5.1 has two columns, one for the resolved and one for the merged region. The resolved and merged regions have for the skimming cuts not completely disjoint sets of events, but in the final selection below they are separated. An event is stored if it passes either of the selections in the two columns, meaning that the first four cuts are always applied, but for the ones below, it is a logically inclusive or.

Table 5.1: Skimming cuts for the signal region. Events that pass either the resolved or merged cuts are stored in the ntuples.

Skimming Cuts	
Resolved	**Merged**
$\not{E}_T$ Trigger above 110 GeV (`HLT_xe110_pufit_L1XE55` and `HLT_xe110_pufit_L1XE50`)	
Photon Veto	
Lepton Veto ($e/\mu/\tau$-veto)	
$\not{E}_T > 150$ GeV	
N(central small-R jets) $\geqslant 2$ N(central small-R b-jets) > 0 Invariant mass of two leading small-R jets > 40 GeV	N(central large-R jets) ≥ 1 Invariant mass of leading large-R jet > 40 GeV

Table 5.2 shows the full event selection for the merged signal region. It starts with triggering on $\not{E}_{\mathrm{T}}$ above 110 GeV, followed by vetoing all events with charged leptons. The photon veto is not implemented in the signal region yet, but the $Z(\nu\bar{\nu})$ + jets MC does not contain any high $\mathrm{p_T}$ photons. At least one large-R jet is requested, and the leading large-R jet should have a mass above 40 GeV. The leading large-R jet is most likely the Higgs candidate in the events, so it has to have the right mass to be a Higgs candidate (50-270 GeV for the merged region). The cut at 40 GeV is to have a lower limit on the leading large-R jet mass in all regions, to make signal region and control regions more similar.

Next, the $\not{E}_{\mathrm{T}}$ is cut at 500 GeV, which is the cut that defines the merged region. The following cut on $\Delta\phi$ is to reduce the multijet background. If one or more jets are mismeasured, it gives $\not{E}_{\mathrm{T}}$ in the direction of the jets. Setting a limit to the angle between the $\not{E}_{\mathrm{T}}$ and the closest of the leading three jets ($\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$) then reduces the multijet background.

Table 5.2: Event selection in the merged signal region.

Merged Signal Region
$\not{E}_{\mathrm{T}}$ Trigger above 110 GeV
Photon Veto
Lepton Veto ($e/\mu/\tau$-veto)
N(central large-R jets) $\geqslant 1$
Mass of leading large-R jet > 40 GeV
$\not{E}_{\mathrm{T}} > 500$ GeV
$\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3}) > 0.35$
Additional b-jet veto
$\mathrm{H_T}$ Ratio $\leqslant 0.57$
N(associated tack-jets) $\geqslant 2$ & ΔR ratio > 1
N(associated b-tagged track-jets) $= 2$

The additional b-jet veto is applied to cut all events with a b-tagged track not associated with the leading large-R jet. The cut on the $\mathrm{H_T}$ Ratio is to make sure that the leading large-R jet carries the largest fraction of the total jet $\mathrm{p_T}$. The $\mathrm{H_T}$ Ratio is defined as the total $\mathrm{p_T}$ of the small-R jets not associated with the leading large-R jets divided by the total $\mathrm{p_T}$ of the leading large-R jet and the non-associated small-R jets. This cut will reduce the $t\bar{t}$ background, where the $\mathrm{p_T}$ will be more evenly distributed between the leading large-R jets and the small-R jets. Furthermore, at least two tracks associated with the leading large-R jet are required, with a separation of the two jets if they are VR-jets (ΔR ratio). The radius of VR-jets depends on their $\mathrm{p_T}$, which is why there is a risk that they are concentric and a separation cut is needed to optimize the b-tagging efficiency. Finally, events are selected where the two associated tracks are b-tagged.

5.4 Photon control region

The photon CR is created in order to select the γ + jets events from MC and data that is used in the background estimation. The region is similar to the SR and the orthogonality to the SR is ensured with a photon veto in the SR and a photon selection in the CR.

For a MC sample, all real missing energy comes from neutrinos. But it is also possible to get fake missing energy from mismeasurements as mentioned previously.

$$\not{E}_{\mathrm{T}} = \not{E}_{\mathrm{T}}^{real} + \not{E}_{\mathrm{T}}^{fake} \tag{5.2}$$

In the $Z(\nu\bar{\nu})$ case, part of the $\not{E}_{\mathrm{T}}^{real}$ in Eq. 5.2 is the Z-boson momentum, $\mathrm{p}_{\mathrm{T}}^{Z}$, giving

$$\not{E}_{\mathrm{T}} = \mathrm{p}_{\mathrm{T}}^{Z} + \not{E}_{\mathrm{T}}^{other\,\nu} + \not{E}_{\mathrm{T}}^{fake}.$$

The $\not{E}_{\mathrm{T}}^{other\,\nu}$ is missing energy from other neutrinos than the Z decay, for example from the b-quark decay. To make the γ + jets sample comparable, the photon p_{T} is added to the $\not{E}_{\mathrm{T}}$, replacing the $\mathrm{p}_{\mathrm{T}}^{Z}$. The $\not{E}_{\mathrm{T}}$ for the photon sample is then renamed according to Eq. 5.3, to make it easier to separate from the $\not{E}_{\mathrm{T}}$ in the $Z(\nu\bar{\nu})$ case.

$$\not{E}_{\mathrm{T}}^{\gamma} = \mathrm{p}_{\mathrm{T}}^{\gamma} + \not{E}_{\mathrm{T}}^{other\,\nu} + \not{E}_{\mathrm{T}}^{fake} \tag{5.3}$$

Since both samples have the same amount of $\not{E}_{\mathrm{T}}^{fake}$ and $\not{E}_{\mathrm{T}}^{other\,\nu}$, $\not{E}_{\mathrm{T}}$ and $\not{E}_{\mathrm{T}}^{\gamma}$ are directly comparable [20].

The skimming cuts are (Table 5.3) the same for the photon CR as for the SR. Instead of the $\not{E}_{\mathrm{T}}$ trigger, the lowest unprescaled photon trigger is used, and the $\not{E}_{\mathrm{T}}$ is replaced by $\not{E}_{\mathrm{T}}^{\gamma}$. The $Z(\nu\bar{\nu})$ + jets MC and the γ + jets MC are compared after the respective skimming cuts in the SR and photon CR to make sure the two samples are comparable.

Table 5.3: Skimming cuts for the photon control region. A photon trigger is used instead of the $\not{E}_{\mathrm{T}}$ trigger in the SR, and all $\not{E}_{\mathrm{T}}$ cuts are instead applied to $\not{E}_{\mathrm{T}}^{\gamma}$. Events that pass either the resolved or merged cuts are stored in the ntuples.

Skimming Cuts	
Resolved	**Merged**
$\mathrm{p}_{\mathrm{T}}^{\gamma}$ Trigger above 140 GeV (`HLT_g140_loose`)	
N(photons) = 1 with $\mathrm{p}_{\mathrm{T}}^{\gamma} \geq 150$ GeV	
Lepton Veto ($e/\mu/\tau$-veto)	
$\not{E}_{\mathrm{T}}^{\gamma} > 150$ GeV	
N(central small-R jets) $\geqslant 2$ N(central small-R b-jets) > 0 Invariant mass of two leading small-R jets > 40 GeV	N(central large-R jets) ≥ 1 Invariant mass of leading large-R jet > 40 GeV

In Table 5.4, the full event selection for the photon CR is listed. The cuts are described in Section 5.3. The full event selection is used in the calculation of the transfer function, as described in Section 5.5.

Table 5.4: Event selection in the photon control region. A photon trigger is used instead of the $\not{E}_{\mathrm{T}}$ trigger in the SR, and all $\not{E}_{\mathrm{T}}$ cuts are instead applied to $\not{E}_{\mathrm{T}}^{\gamma}$.

Photon Control Region
p_{T}^{γ} Trigger above 140 GeV
N(photons) = 1 with $\mathrm{p}_{\mathrm{T}}^{\gamma} \geq 150$ GeV
Lepton Veto (e/μ-veto)
N(central large-R jets) $\geqslant 1$
Mass of leading large-R jet > 40 GeV
$\not{E}_{\mathrm{T}}^{\gamma} > 500$ GeV
$\Delta\phi_{min}(\not{E}_{\mathrm{T}}^{\gamma}, \mathrm{jet}_{1-3}) > 0.35$
Additional b-jet veto
$\mathrm{H_T}$ Ratio $\leqslant 0.57$
N(associated track-jets) $\geqslant 2$ & ΔR ratio > 1
N(associated b-tagged track-jets) $= 2$

5.5 Transfer function definition

The transfer function is based on the assumption that the ratio of number of events for $Z(\nu\bar{\nu})$ + jets and γ + jets is the same in MC and data. Then, the number of $Z(\nu\bar{\nu})$ + jets events in data are given by multiplying the number of γ + jets events in data by the ratio of the MC samples, see Eq. 5.4. The transfer function is defined as a fit for this ratio of MC samples.

$$\frac{N_{Z,SR}^{Data}}{N_{\gamma,CR}^{Data}} = \frac{N_{Z,SR}^{MC}}{N_{\gamma,CR}^{MC}} \rightarrow N_{Z,SR}^{Data} = \frac{N_{Z,SR}^{MC}}{N_{\gamma,CR}^{MC}} \cdot N_{\gamma,CR}^{Data} \tag{5.4}$$

All estimations of $Z(\nu\bar{\nu})$ + jets and γ + jets are done in the SR and photon CR respectively. As mentioned in Section 5.1, the ratio will become fairly constant for high boson p_T, where the affect on the ratio of the mass of the Z boson is negligible. Therefore, the transfer function is derived using the $\not{E}_T$ variable. This is also the variable that defines the signal region.

The discussion following Eq. 2.1 in Section 2.4, gives a theoretical approximation of the transfer function, R_{avg}. It concludes that R_{avg} should have a value around 0.28, and also that it should decrease slightly with increasing $\not{E}_T$. The decrease comes from the difference in the couplings between quarks and the two bosons. This behaviour is then also expected for the transfer function, and motivates that the fit function that defines the transfer function should be a first degree polynomial.

Different triggers are used for $Z(\nu\bar{\nu})$ + jets and γ + jets, $\not{E}_T$ and p_T^γ triggers respectively, and their efficiency differences should be regarded in the transfer function. Also, the difference in trigger efficiency between data and MC should be taken into consideration. Since the transfer function is derived for $\not{E}_T$ above 500 GeV, where the trigger efficiency becomes constant, all of these differences are assumed to be very small and are therefore neglected.

The transfer function is derived for a one jet selection, to compare with values from the theoretical approximation in Section 2.4. Next, it is derived for the selection stages in Table 5.2 and 5.4 to ensure consistency. The three first cuts are skimming cuts, and are therefore applied already. It is also derived for different b-jet multiplicities, where the last cuts in the SR and photon CR are changed to select 0, 1 and 2 b-jets. The transfer functions for different b-jet multiplicities are compared to see if any effect of the difference between the coupling of Z and γ to quarks can influence the b-jet multiplicity selection. Finally, a closure test is performed by applying the transfer function to γ + jets MC samples. The reweighted samples are compared to $Z(\nu\bar{\nu})$ + jets MC samples, which by construction should be the same as the reweighted γ + jets MC sample.

5.6 Systematic uncertainties

The systematic uncertainties for the background estimation come from both the fit of the transfer function itself and from experimental and theoretical uncertainties.

The uncertainty of the fit function is determined from the error of the fit itself and by varying the fit range to see if the fit parameters change. The most deviating values are given as uncertainties.

No systematic uncertainties are included in the ntuple production of the MC samples. Many of the experimental uncertainties should cancel in the ratio, so the impact on the transfer function is small. What remains as a source of uncertainty is the photon selection in the photon CR, namely the identification and reconstruction efficiency, energy scale, resolution and isolation. Benchmark values of these uncertainties are taken from [32] and added in quadrature, giving an overall 5% uncertainty for the photon selection.

The theoretical uncertainty on the γ + jets MC sample is 15-25% [32]. Because of the similarities between γ and Z production processes, many of the theoretical uncertainties are expected to cancel in the ratio. The remaining theoretical uncertainty is roughly estimated to be 10%.

Chapter 6

Results and Discussion

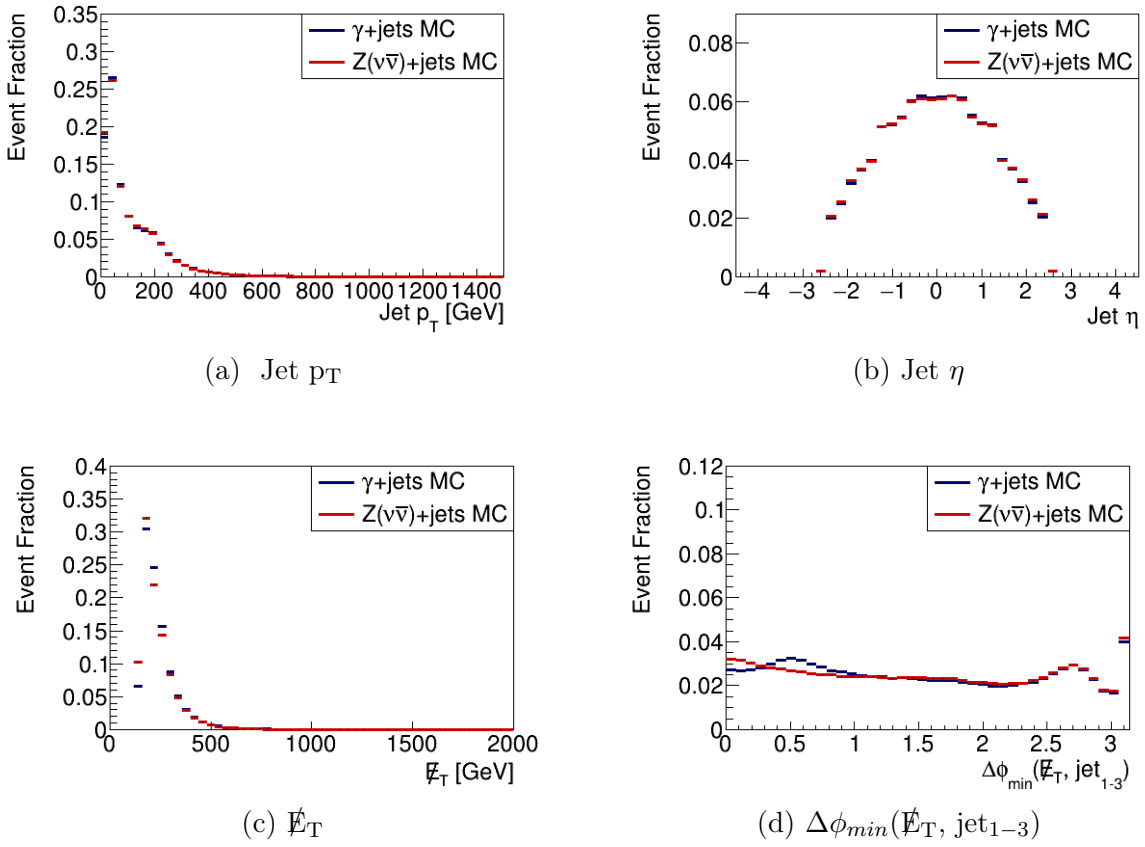
This chapter begins with the result of a comparison between the two MC samples. Then, the transfer function is calculated for a one jet selection, to make the transfer function comparable to the theoretical approximation described in Section 2.4. To ensure consistency, the transfer function is derived for every selection stage in the signal and control region selection. It is also compared when changing the final 2 b-jet selection to 0 or 1 b-jets. A closure test is performed, where $\gamma+$ jets MC samples are reweighted with the transfer function, and compared to $Z(\nu\bar{\nu})$ + jets. Last, the systematic uncertainty for the fit of the transfer function is determined.

6.1 Cut variable comparison

The $Z(\nu\bar{\nu})$ + jets and $\gamma+$ jets MC samples are first compared after the skimming cuts, meaning that the cuts in Table 5.1 and 5.3 have already been applied. Four distributions are seen in Fig. 6.1; the jet $\mathrm{p_T}$, jet η, $\not{E}_{\mathrm{T}}$($\not{E}_{\mathrm{T}}^{\gamma}$) and $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$. All distributions in Fig. 6.1 have been normalized to 1. The $Z(\nu\bar{\nu})$ + jets and $\gamma+$ jets samples show similar shapes for the variables in Fig. 6.1, with exception for low $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ (d). This distribution shows the $\Delta\phi$ between the $\not{E}_{\mathrm{T}}$ and the closest of the three leading jets. Here, the overlap removal between γ and jets ($\Delta R < 0.4$) plays in, removing γ and thereby also $\not{E}_{\mathrm{T}}^{\gamma}$ close to jets. Additional variable distributions can be seen in Appendix A.

Table 6.1 shows each cut in the full signal and control region selection and how much is removed by it. The cut efficiency (2nd and 3rd column) is given by taking the ratio of the number of events before and after a cut. All cuts have a similar impact on the two samples. The last three cuts are for different b-jet multiplicities and show that most events have 0 b-jets (88%), and only 1% have 2 b-jets.

The largest difference in cut effect is for the $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ variable, which selects more events for the γ + jets sample. The cut efficiency difference can be explained by Fig. 6.1d, where the distribution decreases for γ + jets at low $\Delta\phi_{min}(\not{E}_{\mathrm{T}}^{\gamma}, \mathrm{jet}_{1-3})$ because of the overlap removal between γ and jets.



(a) Jet p_T (b) Jet η

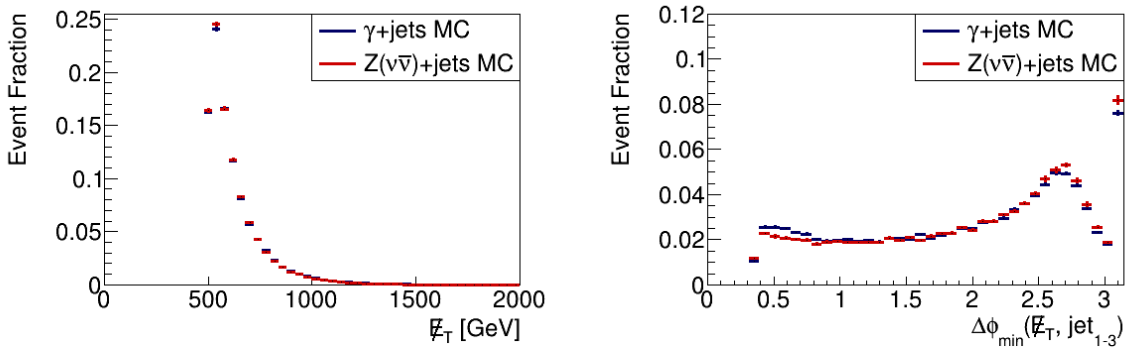
(c) $\not{E}_T$ (d) $\Delta\phi_{min}(\not{E}_T, \text{jet}_{1-3})$

Figure 6.1: Comparison of variables after skimming cuts only. All distributions are normalized to 1 and have statistical error bars.

Table 6.1: Effect of each cut for the $Z(\nu\bar{\nu})$ + jets (2nd column) and γ+ jets (3rd column) MC samples. Each row corresponds to a specific cut and gives the ratio of number of events before and after the cut for the two samples. The last three cuts on b-jet multiplicity are not successive cuts and are one by one compared to the N(associated jets) & ΔR ratio cut.

	$Z(\nu\bar{\nu})$ + jets	γ + jets
Skimming cuts	1	1
N(central large-R jets) $\geqslant 1$	0.707	0.750
Mass of leading large-R jet > 40 GeV	0.859	0.874
$\not{E}_T(\not{E}_T^\gamma) > 500$ GeV	0.030	0.029
$\Delta\phi_{min}(\not{E}_T, \text{jet}_{1-3}) > 0.35$	0.869	0.904
Additional b-jet veto	0.923	0.922
H_T Ratio $\leqslant 0.57$	0.975	0.976
N(associated jets) $\geqslant 2$ & ΔR ratio > 1	0.907	0.908
N(associated b-jets) $= 2$	0.008	0.007
N(associated b-jets) $= 1$	0.109	0.108
N(associated b-jets) $= 0$	0.883	0.885

Fig. 6.2 shows $\not{E}_{\mathrm{T}}$ and $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ distributions with all cuts in the signal and control region applied except for the b-jet multiplicity requirement (the last cut in Table 6.1). All distributions are normalised to 1 and have statistical error bars. Comparing $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ in Fig. 6.1d and 6.2b the distribution has been pushed to higher $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$. This is primarily caused by the $\not{E}_{\mathrm{T}} > 500$ GeV cut, which selects events where $\not{E}_{\mathrm{T}}$ and the jets are more back to back.



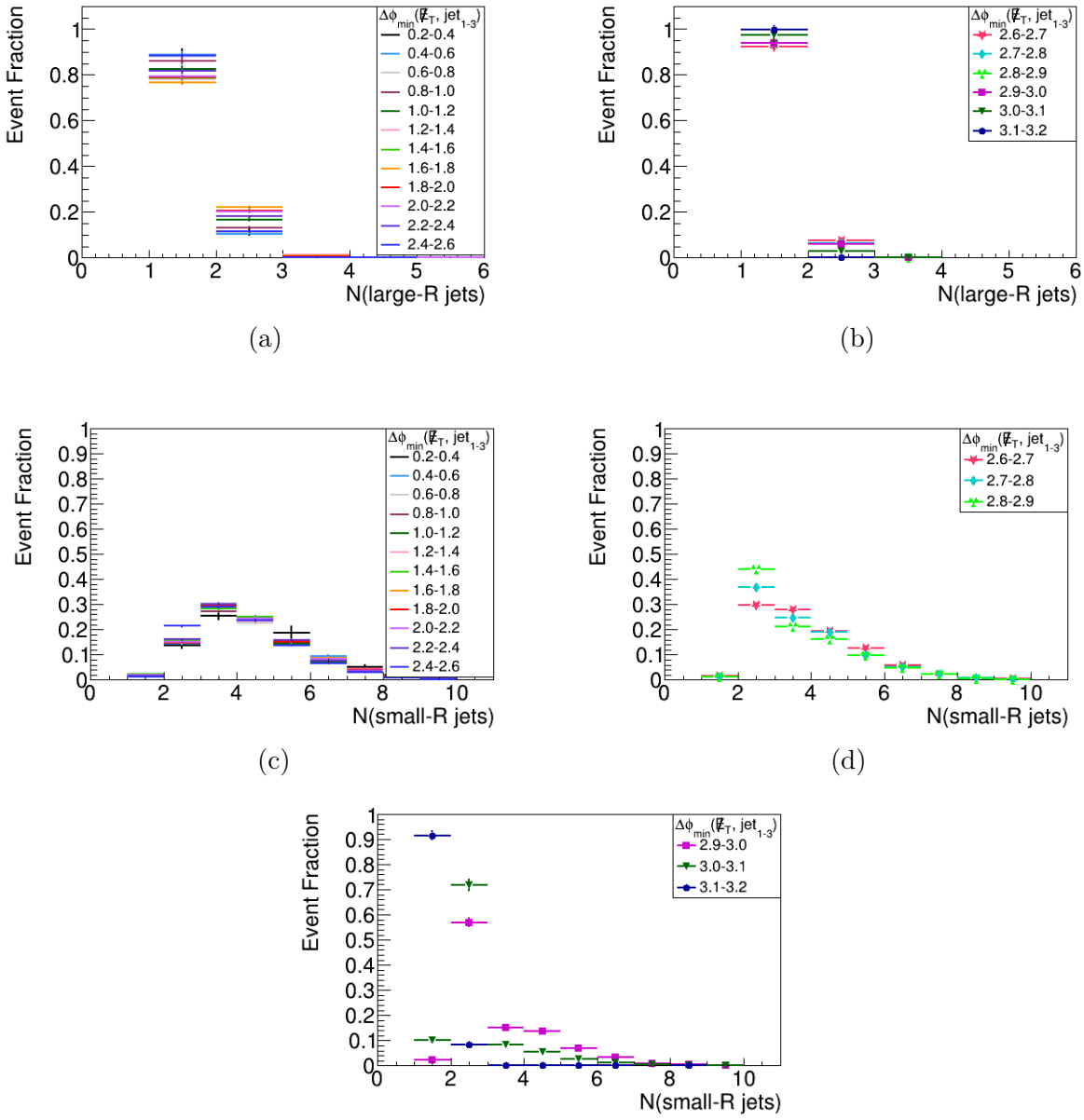
(a) (b)

Figure 6.2: $\not{E}_{\mathrm{T}}$ and $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ where all signal or control region cuts have been applied, except for the b-jet multiplicity cut. The distributions are normalised to 1 and have statistical error bars.

In Fig. 6.2b the $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ distribution first rises towards higher values, and then a decrease of events starts around 2.7, with a minimum for the bin ending at 3.1. The rightmost bin has the largest event fraction, where the $\not{E}_{\mathrm{T}}$ and closest jet are back to back. The jet multiplicity for each $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ bin is shown in Fig. 6.3, using only events from the γ + jets sample shown in Fig. 6.2b. The combinations of large-R jets and small-R jets are similar for $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ ranges below 2.6 (6.3a and c), where most events have one large-R jet and 2-5 small-R jets.

Fig. 6.3b shows that all events with $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3}) > 3.1$ have one large-R jet, and in 6.3e we see that such events also have one, or sometimes two, small-R jets. It is expected to have one small-R jet in the same direction as the large-R jet since there is no overlap removal between the two. Having two small-R jets in a back to back event indicates that they have similar $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$, but are separated in η.

The jet combinations at $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ 2.6-3.1 are in most cases one large-R jet and 2 small-R jets (Fig. 6.3b, d, e). Disregarding the large-R jet, since there is no overlap removal between large-R and small-R jets anyway, going from $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ 2.7 to 3.1 means having the two small-R jets closer and closer together in ϕ. The two small-R jets can still be separated in η, which is why it is possible to have events with two small-R jets with radius 0.4 and $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ above 2.7 at all. The closer the two jets are in ϕ, the less likely it is for them to still be separated in η, because they have less space available. This gives lower and lower event fraction for higher $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ up to 3.1.



(a) (b)

(c) (d)

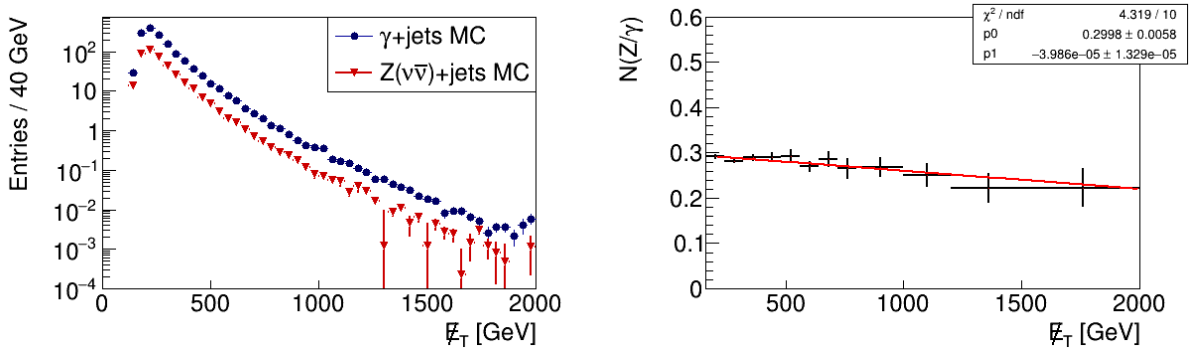
(e)

Figure 6.3: Large-R (a, b) and small-R (c, d, e) jet multiplicities for different $\Delta\phi_{min}(\not{E}_T, \text{jet}_{1-3})$ ranges, for the γ + jets MC with all cuts in the control region applied, except for the b-jet multiplicity cut.

6.2 Transfer function for events with one jet

This section compares events with only one jet to the theoretical approximation of the ratio from Section 2.4. In order to select one jet events, the skimming cuts in Table 5.1 and 5.3 are followed by requiring exactly one small-R jet and one large-R jet. These jets should be the same, and only have different radii, since there is no overlap removal between jets. No events with only one small-R jet can be selected because they are removed by the skimming cuts where at least two small-r jets are required in the resolved selection. Events with one small-R and one large-R jet can be part of the signal and control region if the large-R jet has two b-tags associated to it.

Fig. 6.4a shows the $\not{E}_{\mathrm{T}}$ distributions for the γ + jets and $Z(\nu\bar{\nu})$ + jets samples, for events with one jet. The ratio of the two $\not{E}_{\mathrm{T}}$ distributions is shown in Fig. 6.4b, together with the transfer function, which is the fit to the ratio. Here, the fit is done from 200 to 2000 GeV, with a first order polynomial. The ratio decreases with $\not{E}_{\mathrm{T}}$, as expected from section 2.4, since for higher $\not{E}_{\mathrm{T}}$, higher momenta per parton in the interaction is needed. This means going towards higher x in the parton distribution function in Fig. 2.6, reducing the d/u relation. Having more u-type quarks will reduce the cross section ratio because γ couples stronger to them than Z relative d-type quarks, as discussed in section 2.4. In the same section, the ratio is calculated to be 0.28 assuming $x = 0.1$. From Fig. 6.4b, given the parameters from the fit, the ratio is 0.29 for $\not{E}_{\mathrm{T}}$ 200 GeV, and 0.28 for 500 GeV.



(a) (b)

Figure 6.4: $\not{E}_{\mathrm{T}}$ comparison (a) and their ratio (b), also showing the transfer function, for events with one jet.

6.3 Transfer function at different selection stages

To ensure consistency the transfer function is derived at every selection stage in the signal and control region. If one cut alters the relation between the $\not{E}_{\mathrm{T}}$ distributions in the γ + jets and $Z(\nu\bar{\nu})$ + jets samples, the transfer function should change when that cut is applied. Thereby, the transfer function is derived after each consecutive selection in the signal and photon control region, see Table 5.2 and 5.4. Every selection includes all preceding cuts in the table. The b-jet multiplicity cut is not included here but is instead

addressed in the next section. The transfer functions from each selection stage are seen in Fig. 6.5, where the transfer function is the fit of a first order polynomial between 500 and 2000 GeV. The slope for each transfer function is given in the legend, before the name of each cut.

When the $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ cut is applied, the transfer function decreases overall, and increases in slope. It means that the $\not{E}_{\mathrm{T}}$ ratio decreases in total and even more for high $\not{E}_{\mathrm{T}}$. All other cuts show no drastic changes to the transfer function.

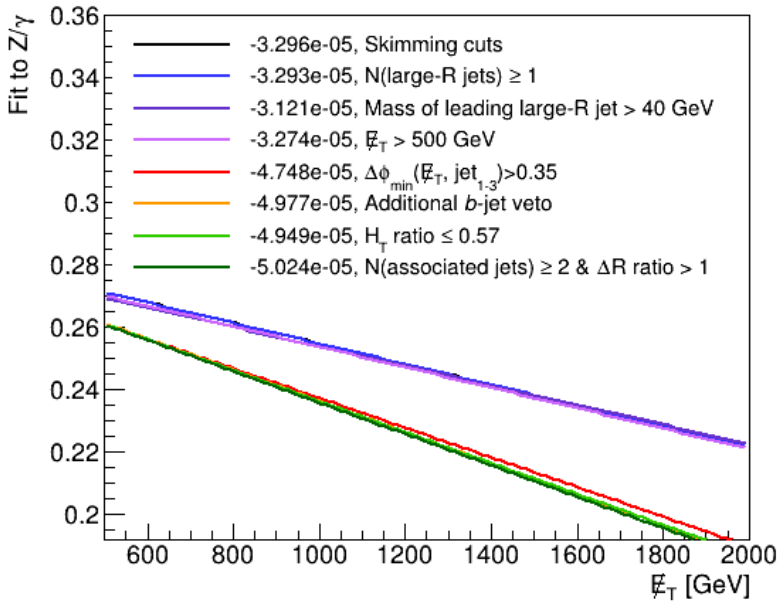


Figure 6.5: Transfer function derived for each selection stage in the signal region and photon control region. The slope of each transfer function is given in the legend.

The effect of $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ to the transfer function is studied by comparing $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ distributions for different $\not{E}_{\mathrm{T}}$ ranges. Fig. 6.6 shows the lowest range of $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ for different $\not{E}_{\mathrm{T}}$ ranges between 300 and 1100 GeV. For γ + jets samples all $\not{E}_{\mathrm{T}}$ ranges show similar behavior since the overlap removal between γ and jets will reduce the number of events at low $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ significantly. For $Z(\nu\bar{\nu})$ + jets, the distribution rises at low $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ for higher $\not{E}_{\mathrm{T}}$. Why the topology changes have not been further investigated. It could be due to a depletion of the ϕ region around $\pi/2$ from the Z for higher $\not{E}_{\mathrm{T}}$, when the Z and recoiling jets become more back to back.

When cutting $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ at 0.35, it will remove more events for $Z(\nu\bar{\nu})$ + jets relative γ + jets, and this difference will increase for higher $\not{E}_{\mathrm{T}}$. This introduces the increasing slope and decreasing value of the transfer function for this cut seen in Fig. 6.5.

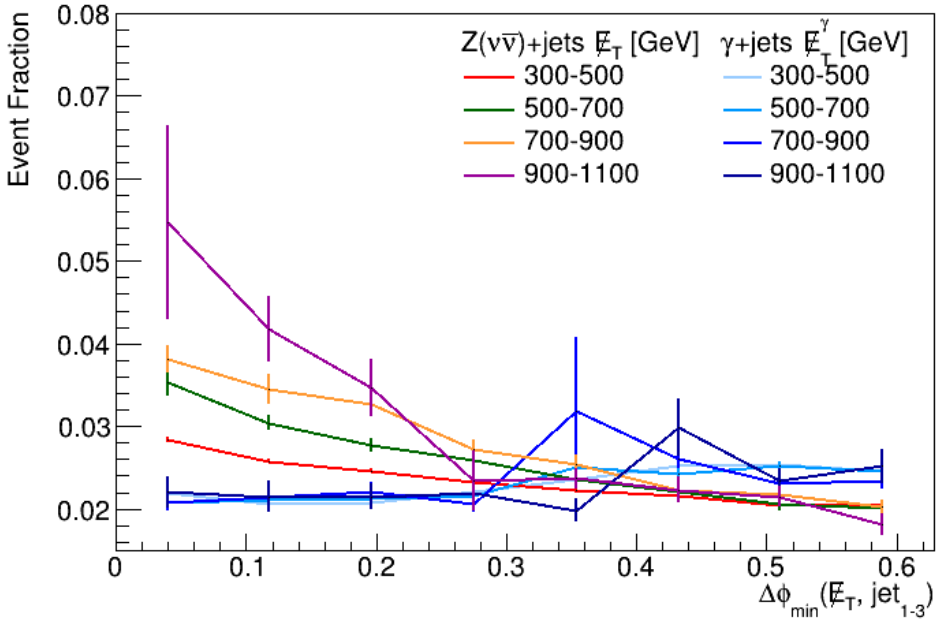


Figure 6.6: Lowest range of the $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ distribution, where the jets are closest to the $\not{E}_{\mathrm{T}}$, for different ranges of $\not{E}_{\mathrm{T}}$. The error bars are the statistical uncertainty.

6.4 Transfer functions for different *b*-jet multiplicities

The final cut in the signal and control regions is selecting events with exactly 2 b-jets. Here, the transfer function for the 2 b-jet final selection is compared with transfer functions corresponding to requiring 0 or 1 b-jets instead.

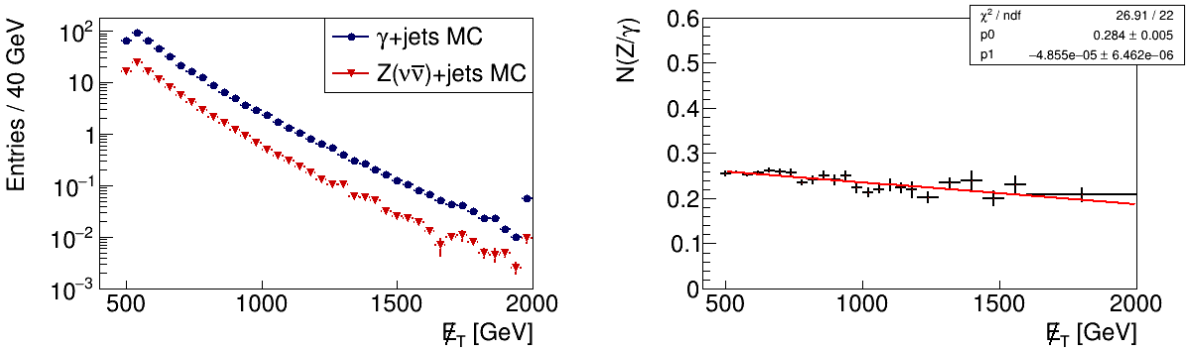
Fig. 6.7a, 6.8a and 6.9a shows the $\not{E}_{\mathrm{T}}$ distributions for the γ + jets and the $Z(\nu\bar{\nu})$ + jets samples, for 0, 1 and 2 b-jets respectively. The $\not{E}_{\mathrm{T}}$ ratio for each selection and the corresponding transfer function is seen in Fig. 6.7b, 6.8b and 6.9b. The fit range of the transfer functions is from 500 to 2000 GeV. For the 0 and 1 b-jet case, the binning is determined by requiring a maximum relative error of 10% per bin for the $\not{E}_{\mathrm{T}}$ ratio. The relative error is the ratio of the bin error and the bin content. For the 2 b-jets $\not{E}_{\mathrm{T}}$ ratio a 20% maximum relative error is used to determine the binning.

Compared to the transfer function for the one jet selection in Fig. 6.4 (which is compared to theory), the transfer functions for events with multiple jets are slightly lower and have a little less slope. These differences come from additional Feynman diagrams included to produce multiple jet events, that influence the transfer function. Example Feynman diagrams with 2 jets can be seen in Fig. 2.5(c-f). Another aspect that could play in when having events with multiple jets is the possibility of a collinear singularity for γ when a jet is produced close to it, mentioned in Section 2.4. Having both overlap removal for γ and jets, and the cut on $\Delta\phi_{min}(\not{E}_{\mathrm{T}}, \mathrm{jet}_{1-3})$ at 0.35, should reduce the influence of the collinear singularity.

The three transfer functions for 0, 1 and 2 b-jets are similar, but the one for 2 b-jets has less slope than the other two. The transfer function difference is caused by low statistics at high $\not{E}_{\mathrm{T}}$ for the 2 b-jet selection, causing the rightmost bin to be very large. As mentioned above, the binning is determined by the relative error, which should stay below 20%. The transfer function is sensitive to the binning that is chosen at lower $\not{E}_{\mathrm{T}}$ as well, because of low statistics in the whole $\not{E}_{\mathrm{T}}$ range, also giving the parameters of the transfer function large errors. Changing the binning alters the values of the fit parameters, but they stay within the error of the transfer function in Fig. 6.9b.

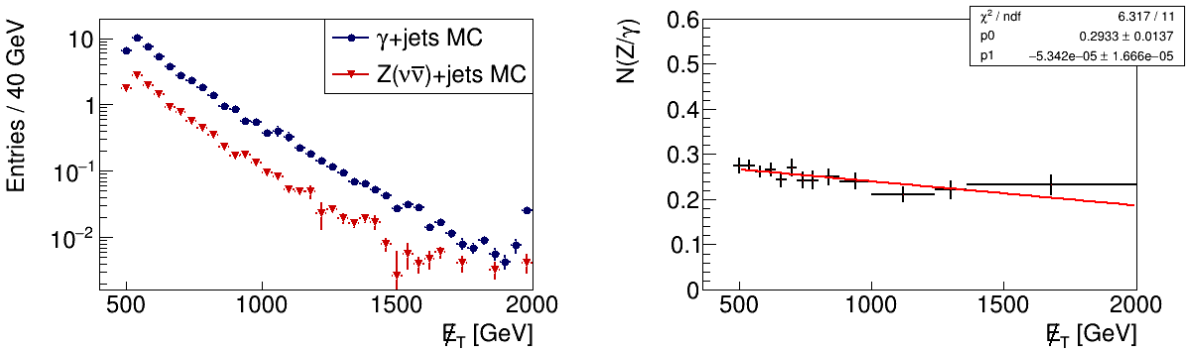
To determine if it is possible to use the 0 b-jet region to derive the transfer function instead of the 2 b-jet region, it is investigated how similar the three b-jet multiplicity selections are. Four comparisons of the b-jet multiplicity selections follow. It would be favorable to use the higher statistics 0 b-jet region when deriving the transfer function, which is the reason to make additional comparisons between the b-jet selections.

Changing the fit range to between 500 and 1200 GeV, ignoring the bin at highest $\not{E}_{\mathrm{T}}$, gives a transfer function for 2 b-jets with slope better comparable to the others. The highest $\not{E}_{\mathrm{T}}$ bin is ignored to avoid the bins with 0 entries for $Z(\nu\bar{\nu})$ + jets (Fig. 6.9a). The transfer functions with the shorter fit range are shown in Fig. 6.10, with a smaller slope difference between b-jet multiplicities because of the shorter fit range. The shorter range transfer function comparison indicates that the three b-jet multiplicity cases are indeed similar.



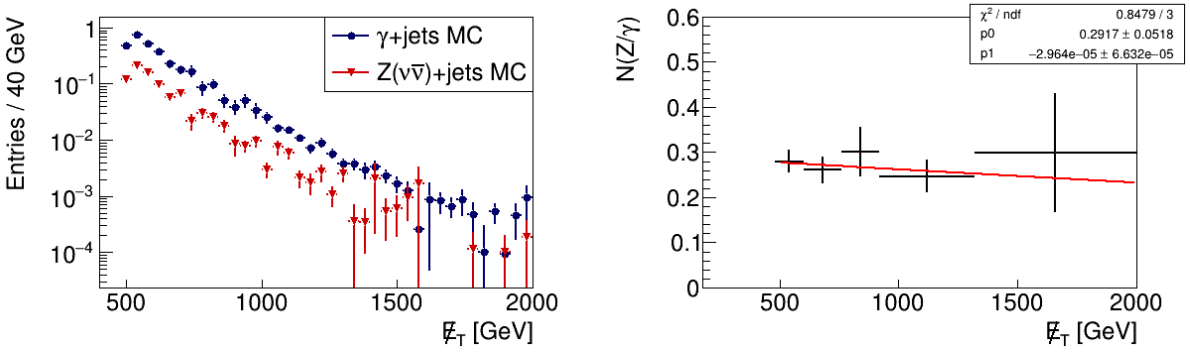
(a) (b)

Figure 6.7: $\not{E}_T$ comparison in (a) for selecting 0 b-jets in the final cut. The ratio of the $\not{E}_T$ distributions in (b), also showing the transfer function.



(a) (b)

Figure 6.8: $\not{E}_T$ comparison in (a) for selecting 1 b-jet in the final cut. The ratio of the $\not{E}_T$ distributions in (b), also showing the transfer function.



(a) (b)

Figure 6.9: $\not{E}_T$ comparison in (a) for selecting 2 b-jets in the final cut. The ratio of the $\not{E}_T$ distributions in (b), also showing the transfer function.

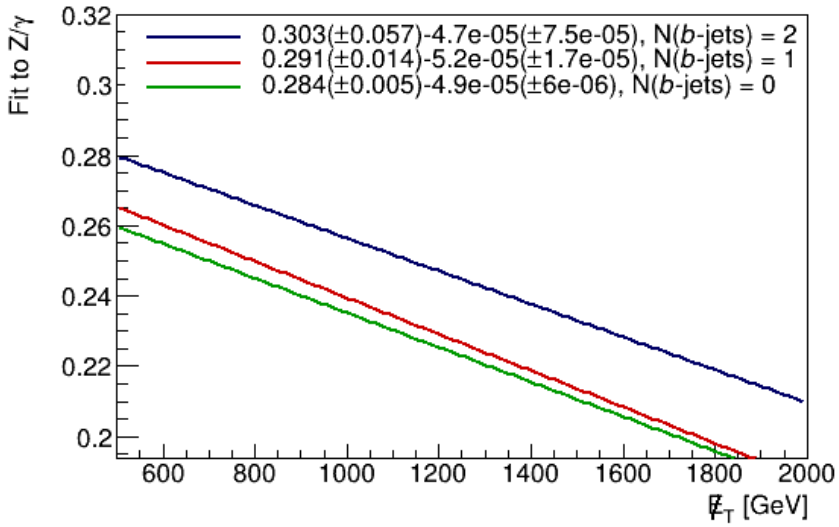


Figure 6.10: Transfer function derived for different b-jet multiplicities. The legend shows the fit parameters of the transfer function with errors.

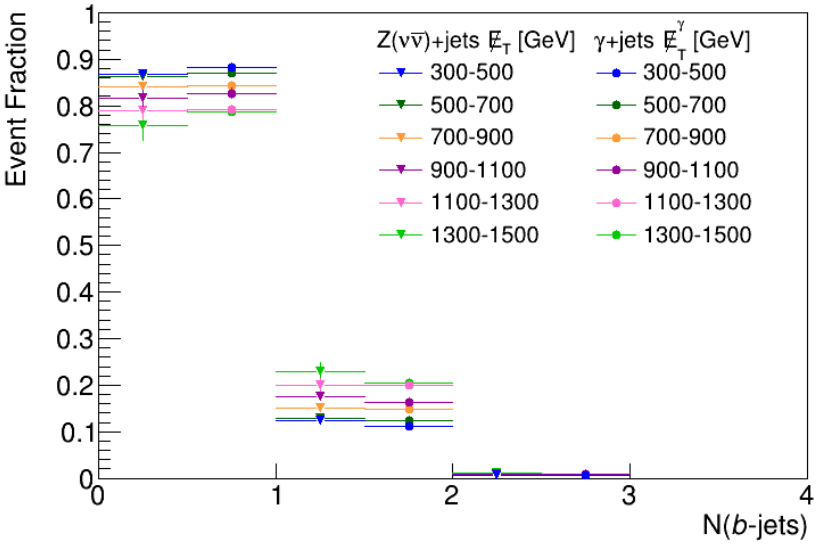


Figure 6.11: b-jet multiplicity distributions for different $\not\!\!E_T$. The $Z(\nu\bar{\nu})$ + jets distributions are shown in the left part of each multiplicity bin and the γ + jets in the right.

In order to determine if the b-jet selection introduces a difference in the slope of the transfer function, the number of b-jets for different $\not\!\!E_T$ ranges are compared. To motivate a slope difference, the γ + jets and $Z(\nu\bar{\nu})$ + jets b-jet multiplicity distributions should differ between $\not\!\!E_T$ ranges. Then one of the b-jet selections would cut away more events from one sample, leading to a change of the slope of the transfer function. As seen in Fig. 6.11, this is not the case. All $\not\!\!E_T$ selections have the similar event fraction for both samples. An overall change of the transfer function for different b-jet multiplicities can also be ruled out from Fig. 6.11. Then one of the samples should have an overall different distribution, regardless of the $\not\!\!E_T$ selection. This is more apparent in Fig. 6.12.

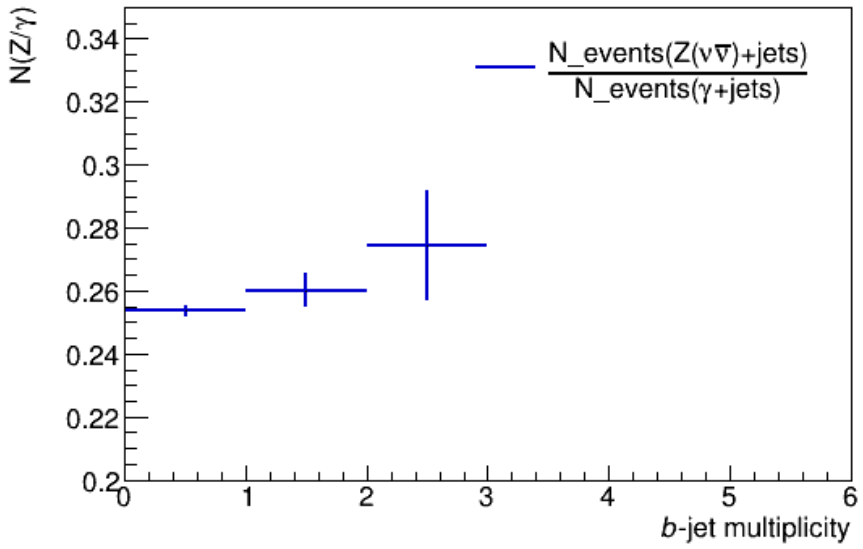


Figure 6.12: Event ratio per b-jet multiplicity.

Fig. 6.12 gives the ratio of the total number of events for the three b-jet multiplicity selections, where the error bar is the statistical uncertainty. The ratio is about 0.25, which corresponds well to the pure cross sectional ratio, calculated in Section 2.4 to reach 0.28 for single jet events, and then decreasing with $\not{E}_{\mathrm{T}}$. This figure shows that the cross section ratio is within uncertainties for the three b-jet multiplicity selections, and no difference that should change the value of the transfer function can be inferred.

To further study the three b-jet selections, the small-R, and large-R jet multiplicities are seen in Fig. 6.13, for both γ + jets and $Z(\nu\bar{\nu})$ + jets samples. No differences between the three cases are recognized, which shows that the three different b-jet selections have the same event structure, not revealing any reason for the transfer function to differ between b-jet multiplicities.

The comparison of the three b-jet multiplicity selections shows no differences. Based on this, the transfer function for the 0 b-jet selection will be used in the background estimation instead of the 2 b-jets transfer function. It is beneficial since it has more statistics, and smaller fit errors. From Fig. 6.7b, the transfer function is determined to $0.28\text{-}5 \cdot 10^{-5} \not{E}_{\mathrm{T}}$ [GeV].

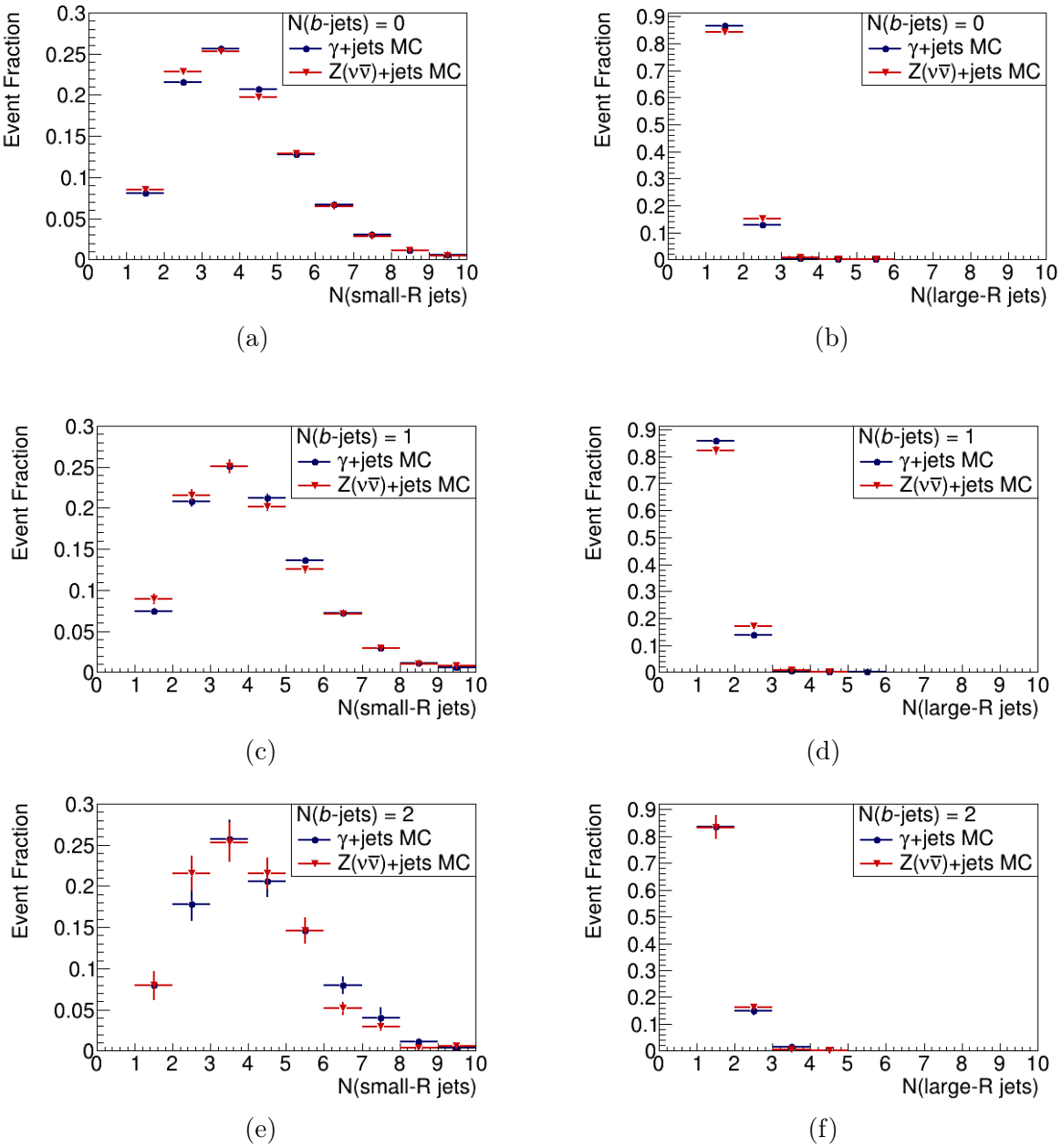


Figure 6.13: Small-R and large-R jet multiplicities when requiring 0, 1 and 2 b-jets.

6.5 Closure test

In this section, the transfer function is used to reweight γ + jets MC samples, which is compared to $Z(\nu\bar{\nu})$ + jets MC. The transfer function is applied for every bin, by evaluating it for the bin center and then multiply with the bin content.

First, the transfer function is applied to γ + jets MC for the 0 b-jet selection, seen in the upper part of Fig. 6.14. The ratio of the two distributions (lower part of Fig. 6.14) shows that the reweighted $\not{E}_{\mathrm{T}}^{\gamma}$ distribution matches the $\not{E}_{\mathrm{T}}$ from $Z(\nu\bar{\nu})$ well. The error bars are the statistical uncertainty only.

Fig. 6.15 shows the transfer function applied to γ + jets MC for the 2 b-jet selection. The transfer function is, as defined in the last section, derived for the 0 b-jet selection. Again, the overall agreement to the $Z(\nu\bar{\nu})$ + jets $\not{E}_{\mathrm{T}}$ distribution is good.

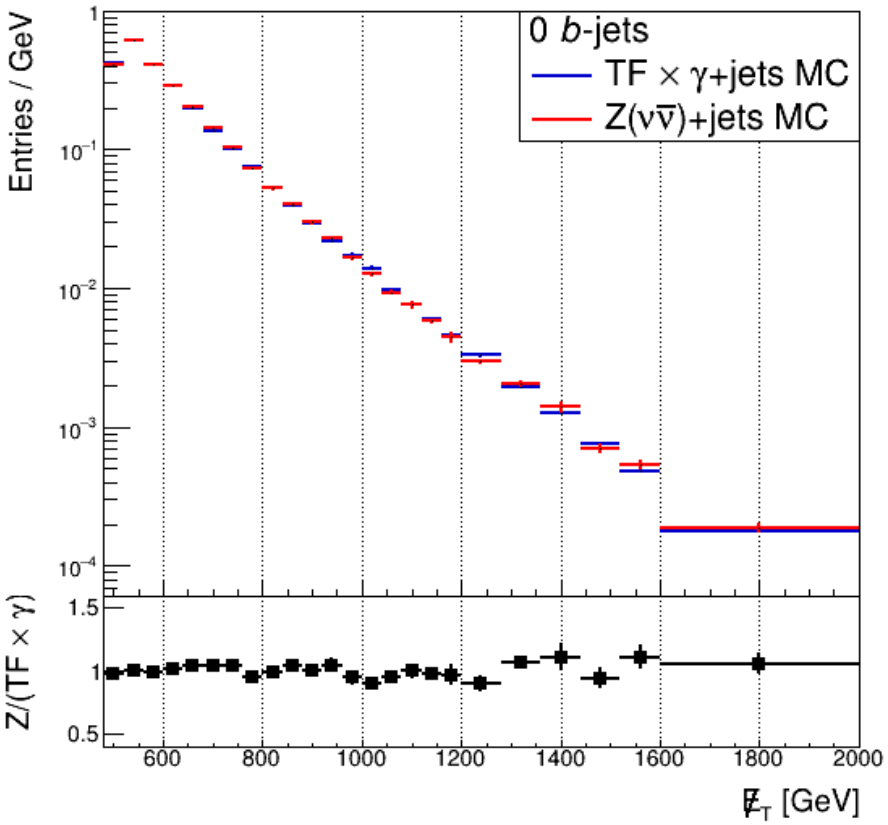


Figure 6.14: Transfer function applied to γ + jets MC for the 0 b-jet selection, compared to $Z(\nu\bar{\nu})$ + jets MC (upper) and the ratio of the two $\not{E}_{\mathrm{T}}$ distributions (lower).

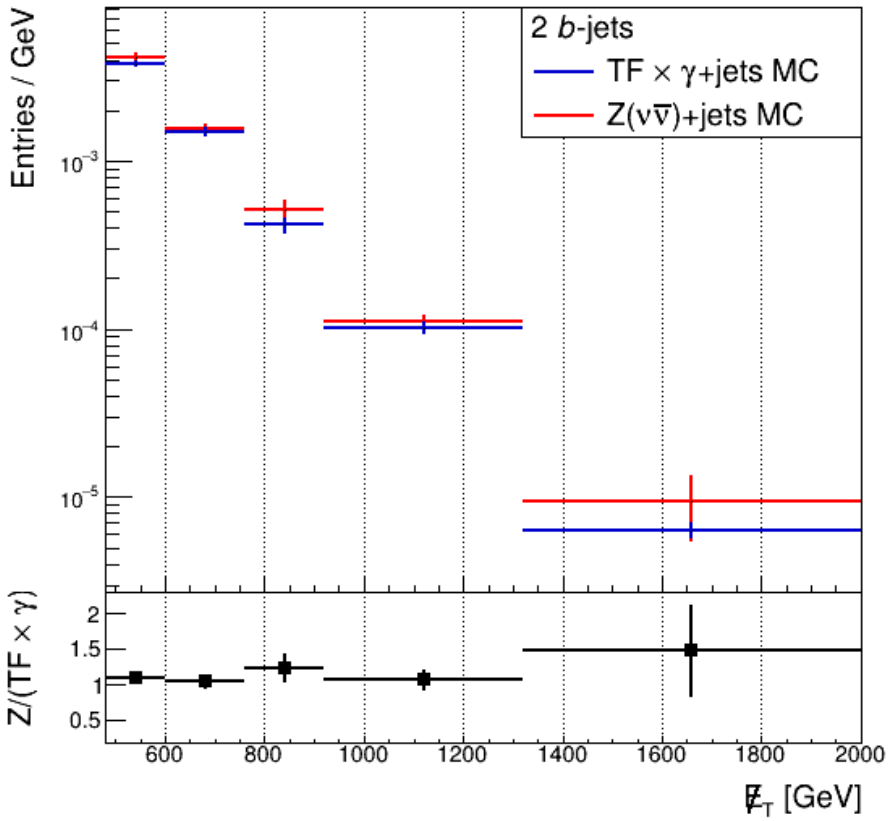


Figure 6.15: Transfer function applied to γ + jets MC for the 2 b-jet selection, compared to $Z(\nu\bar{\nu})$ + jets MC (upper) and the ratio of the two $\not{E}_{\mathrm{T}}$ distributions (lower).

6.6 Determination of systematic uncertainties

To determine if there is an additional uncertainty on the transfer function from the fit range, the range of the fit is varied by changing either the lower or upper limit. The lower limit, originally at $\not{E}_T = 500$, is set to 600, 700 and 800 GeV. For the upper limit, usually set to $\not{E}_T = 2000$, the corresponding values are 1200, 1500, 1700 GeV. When changing the range of the fit, the values of the fitting parameters stay within error for the parameters of the transfer function with fit range 500-2000 GeV.

Fig. 6.16 shows the transfer function for the 0 b-jet selection in red (the same as Fig. 6.7b), with its fit error in blue. The constant parameter in the fit has an error of about 2% and the parameter for the slope an error of 14%.

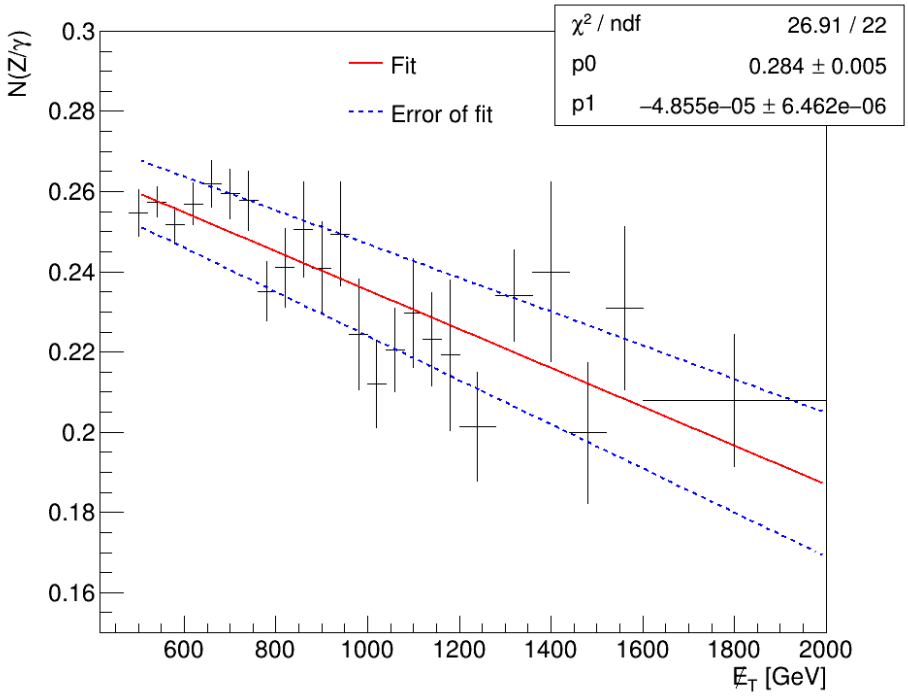


Figure 6.16: $\not{E}_T$ ratio for 0 b-jets, with the transfer function in red and its corresponding error in blue.

As mentioned in Section 5.6, the photon selection has a 5% uncertainty, and the theoretical uncertainty on the ratio of the MC samples is roughly estimated to 10%. Adding these to the fit uncertainty gives an overall systematic uncertainty of 18%.

Chapter 7

Outlook

In this thesis a transfer function from γ + jets to $Z(\nu\bar{\nu})$ + jets has been derived, which can be used in the background estimation of the irreducible $Z(\nu\bar{\nu})$ + jets background. First, a control region for the γ + jets samples was created, where the photon p_T is added to the $\not{E}_T$ to replace the Z boson decaying to two neutrinos. The transfer function is defined as a first order polynomial fit to the ratio of number of $Z(\nu\bar{\nu})$ + jets events in the signal region and γ + jets events in the control region, using MC samples.

The transfer function was first derived for a one jet selection to enable comparison to a theoretical approximation, that confirms the values of the fit parameters. It was also derived for the selection stages of the signal and control regions, and any changes of the transfer function were explained. When comparing transfer functions for different b-jet multiplicity selections, they appeared to be very similar. For this reason, the transfer function for 0 b-jets was used in the background estimation, instead of the one for 2 b-jets, because of higher statistics. The transfer function for 0 b-jets was determined to $0.28\text{-}5 \cdot 10^{-5}\not{E}_T$ [GeV] with a total uncertainty of 18%.

The results from this thesis are useful for anyone who wants to make a $Z(\nu\bar{\nu})$ + jets background estimation. The transfer function itself is derived with a selection specific for the dark matter search performed by the mono-H($b\bar{b}$) ATLAS analysis group. With further work, this method of estimating the $Z(\nu\bar{\nu})$ + jets background can be used in the full Run II analysis.

Examples of future work can be to perform a comparison between data and MC in the control region defined in this thesis. An estimate of the contamination of the other backgrounds in the control region is also needed. Validation of the transfer function has to be done in data, where the transfer function is applied to γ + jets control region data and the outcome compared to $Z(\nu\bar{\nu})$ + jets signal region data. To take part in the full Run II analysis, the transfer function has to be derived using MC samples for all of Run II.

References

[1] ATLAS Collaboration. "Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ using 36 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector". In: *Phys. Rev. Lett.* 119.18 (2017). DOI: 10.1103/PhysRevLett.119.181804. arXiv: 1707.01302.

[2] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Rev. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214.

[3] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[4] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[5] Peter W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Phys. Lett.* 12 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9.

[6] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[7] Symmetry Magazine Kurt Riesselman. *Symmetry Dimensions of Particle Physics.* URL: https://www.symmetrymagazine.org/standard-model/ (visited on 11/27/2018).

[8] Gordon Kane. *Modern elementary particle physics: explaining and extending the standard model; 2nd ed.* Cambridge: Cambridge University Press, 2017.

[9] N. Aghanim et al. "Planck 2018 results. VI. Cosmological parameters". In: (2018). arXiv: 1807.06209 [astro-ph.CO].

[10] Katherine Freese. "Review of Observational Evidence for Dark Matter in the Universe and in upcoming searches for Dark Stars". In: *EAS Publ. Ser.* 36 (2009), pp. 113–126. DOI: 10.1051/eas/0936016. arXiv: 0812.4005 [astro-ph].

[11] Douglas Clowe, Anthony Gonzalez, and Maxim Markevitch. "Weak-Lensing Mass Reconstruction of the Interacting Cluster 1E 0657-558: Direct Evidence for the Existence of Dark Matter". In: *The Astrophysical Journal* 604.2 (2004), p. 596. URL: http://stacks.iop.org/0004-637X/604/i=2/a=596.

[12] Mark H. Jones, Robert J. A. Lambourne, and Stephen Serjeant. *An Introduction to Galaxies and Cosmology (2nd ed)*. Cambridge: Cambridge University Press/Open University, 2015, pp. 287–289. URL: http://oro.open.ac.uk/44361/.

[13] Pieter van Dokkum et al. "A galaxy lacking dark matter". In: *Nature* 555 (Mar. 2018). DOI: 10.1038/nature25767. URL: https://doi.org/10.1038/nature25767.

[14] Giorgio Busoni et al. "Making the Most of the Relic Density for Dark Matter Searches at the LHC 14 TeV Run". In: *JCAP* 1503.03 (2015), p. 022. DOI: 10.1088/1475-7516/2015/03/022. arXiv: 1410.7409.

[15] Gianfranco Bertone and Dan Hooper. "History of dark matter". In: *Rev. Mod. Phys.* 90 (4 2018), p. 045002. DOI: 10.1103/RevModPhys.90.045002. URL: https://link.aps.org/doi/10.1103/RevModPhys.90.045002.

[16] Teresa Marrodán Undagoitia and Ludwig Rauch. "Dark matter direct-detection experiments". In: *J. Phys.* G43.1 (2016), p. 013001. DOI: 10.1088/0954-3899/43/1/013001. arXiv: 1509.08767.

[17] Asher Berlin, Tongyan Lin, and Lian-Tao Wang. "Mono-Higgs Detection of Dark Matter at the LHC". In: *JHEP* 06 (2014), p. 078. DOI: 10.1007/JHEP06(2014)078. arXiv: 1402.7074 [hep-ph].

[18] ATLAS Collaboration. "Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ at $\sqrt{s} = 13$ TeV with the ATLAS Detector". In: ATL-COM-PHYS-2016-1484 (2016). URL: https://cds.cern.ch/record/2225941.

[19] ATLAS Collaboration. "Search for dark matter produced in association with a Higgs boson decaying to two bottom quarks in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Phys. Rev.* D93.7 (2016). DOI: 10.1103/PhysRevD.93.072007. arXiv: 1510.06218.

[20] Tanya Sandoval. "Estimation of the $Z \to \nu\nu$ background to New Physics searches in ATLAS". PhD thesis. Cambridge U., 2012. DOI: 10.17863/CAM.16591. URL: https://www.repository.cam.ac.uk/handle/1810/244264.

[21] A. Airapetian et al. "ATLAS: Detector and physics performance technical design report. Volume 1". In: *CERN-LHCC-99-14, ATLAS-TDR-14* (1999).

[22] Joao Pequenao. *Computer generated image of the whole ATLAS detector*. 2008. URL: https://cds.cern.ch/record/1095924 (visited on 09/24/2018).

[23] Joao Pequenao and Paul Schaffner. *An computer generated image representing how ATLAS detects particles*. 2013. URL: https://cds.cern.ch/record/1505342 (visited on 09/24/2018).

[24] Anders Floderus. "Luminosity determination and searches for supersymmetric sleptons and gauginos at the ATLAS experiment". PhD thesis. Lund University, 2014.

[25] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-ktjet clustering algorithm". In: *Journal of High Energy Physics* (2008), p. 63. DOI: 10.1088/1126-6708/2008/04/063.

[26] James Catmore. *The ATLAS data processing chain: from collisions to papers*. URL: https://indico.cern.ch/event/472469/contributions/1982677/attachments/1220934/1785823/intro_slides.pdf (visited on 12/18/2018).

[27] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[28] T. Gleisberg et al. "Event generation with SHERPA 1.1". In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622.

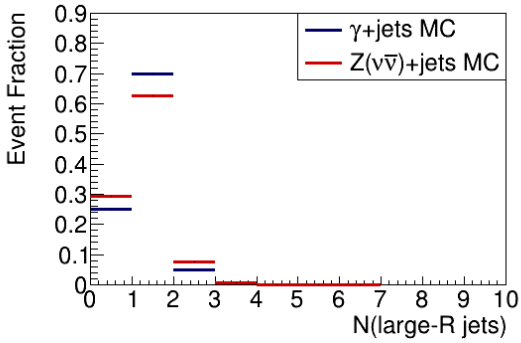
[29] Mono-Higgs ATLAS analysis group. *XAMPPmonoH: Implementation of the DM mono-Higgs (bb) Search*. URL: https://gitlab.cern.ch/atlas-mpp-xampp/XAMPPmonoH (visited on 05/02/2019).

[30] J.M. Lindert et al. "Precise predictions for V+jets dark matter backgrounds". In: *Eur. Phys. J. C* 77.CERN-TH-2017-102. 12 (2017), 829. 36 p. URL: https://cds.cern.ch/record/2264512.

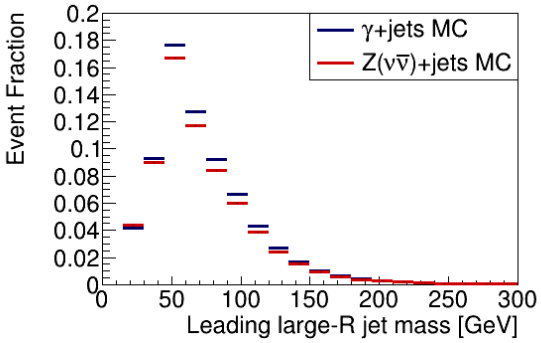
[31] *Search for Dark Matter Produced in Association with a Higgs Boson decaying to $b\bar{b}$ at $\sqrt{s} = 13$ TeV with the ATLAS Detector using 79.8 fb^{-1} of proton-proton collision data*. Tech. rep. ATLAS-CONF-2018-039. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2632344.

[32] ATLAS Collaboration. "Measurement of the cross section for isolated-photon plus jet production in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector". In: *Phys. Lett.* B780 (2018), pp. 578–602. DOI: 10.1016/j.physletb.2018.03.035. arXiv: 1801.00112.

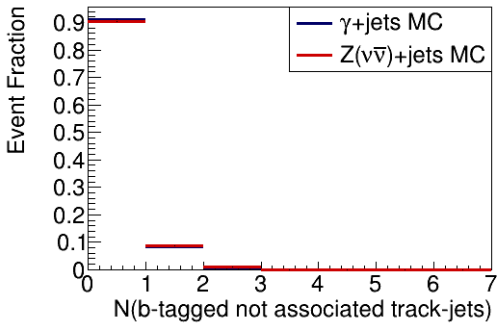
Appendix A: Additional plots



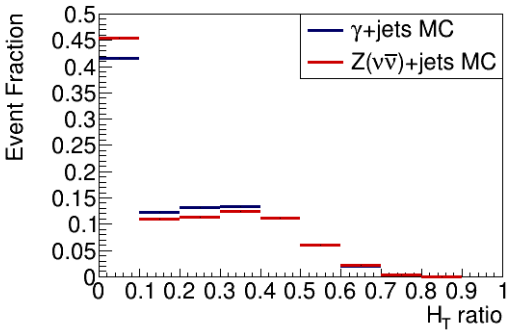
(a)



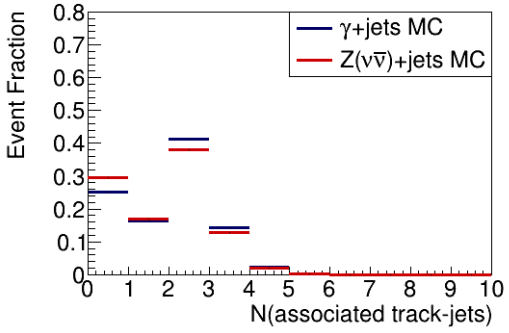
(b)



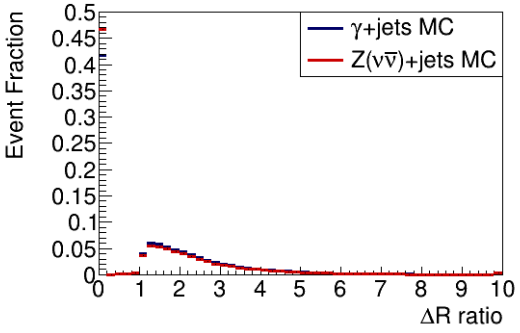
(c)



(d)



(e)



(f)

Figure A.1: Additional cut variable distributions at skimming cuts.

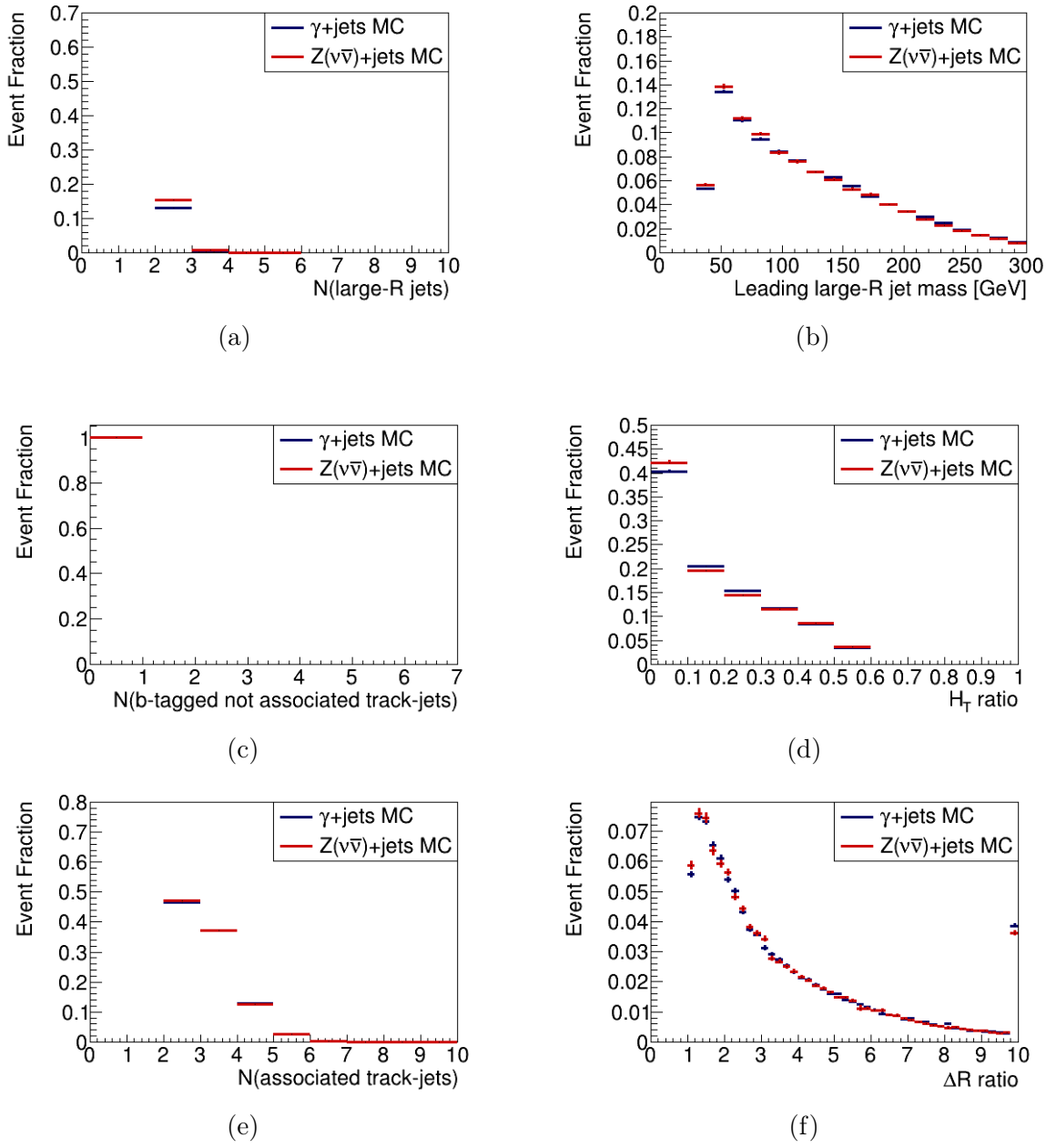


Figure A.2: Additional cut variable distributions with all cuts in signal and control region applied, except for the b-jet multiplicity cut.