Search for rare exclusive hadronic decays of the W boson with the ATLAS experiment

Júlia Manuela Cardoso Silva

Supervisors: Prof. Konstantinos Nikolopoulos
Dr. Andrew Stephen Chisholm

Thesis submitted for the degree of
Doctor of Philosophy

UNIVERSITY OF BIRMINGHAM

Particle Physics Group,
School of Physics and Astronomy,
University of Birmingham.

October 2023

Abstract

The search for rare exclusive hadronic decays of the W boson in the $W^{\pm} \to \pi^{\pm}\gamma$, $W^{\pm} \to K^{\pm}\gamma$ and $W^{\pm} \to \rho^{\pm}\gamma$ final states is presented. This search was performed using up to 140 fb^{-1} of proton collision data collected by the ATLAS experiment at a center-of-mass energy of $\sqrt{s} = 13$ TeV. These decays can serve as clean probes of the QCD factorisation formalism and, in the future, could offer a new way to perform the W boson mass measurement, through fully reconstructed final states. Upper limits at 95% CL are set on the branching fractions of the decays: $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma) < 1.9 \times 10^{-6}$, $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma) < 1.7 \times 10^{-6}$ and $\mathcal{B}(W^{\pm} \to \rho^{\pm}\gamma) < 5.4 \times 10^{-6}$. These results correspond to the first experimental constraints on $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma)$ and $\mathcal{B}(W^{\pm} \to \rho^{\pm}\gamma)$, and the most stringent upper limit to date on $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma)$. This analysis was enabled by the use of dedicated triggers and of a novel, data-driven non-parametric background modelling technique. Detailed characterisation and validation studies of this novel background modelling method are also presented.

Declaration of Author's Contribution

While most of the results presented in this thesis are a direct result of my own work, this work was not done in isolation, but as part of different collaborations. Various inputs which required the work of many other people were used, and are summarised here. Moreover, all of the work was performed under the guidance of my supervisor, Kostas Nikolopoulos, and the support of the Birmingham Higgs group.

The first two chapters present theoretical and experimental aspects which are relevant for the work presented in the following chapters, and as such do not include original work. Many scientific publications and textbooks were consulted and reviewed throughout the writing of these chapters, and are referenced accordingly.

Chapters 3 and 5 present work which I performed as a member of the ATLAS collaboration, and therefore benefitted from the work of thousands of people which contributed to the ATLAS experiment over the years. Besides what is presented in this thesis, during my PhD I participated in ATLAS Run 3 data operations, by performing ATLAS control room shifts (Run Control and Trigger). Furthermore, I was responsible for implementing and maintaining the triggers for exclusive Higgs, Z and W searches in the Run 3 trigger menu. Besides this detector operations and performance work, I also contributed to the early Run 3 measurement of the W and Z boson cross sections, at $\sqrt{s} = 13.6$ TeV [1]. My main contributions corresponded to multijet background studies in the $Z \to ee$ and $Z \to \mu\mu$ channels and fit validation studies.

The work presented in Chapter 3 was performed within the ATLAS Clustering and Tracking in Dense Environments (CTIDE) group, and it builds upon a previous ATLAS publication [2]. Section 3.1 details the methodology which was in place when I started working on the analysis. Most of the framework which obtains the cluster $\mathrm{d}E/\mathrm{d}x$ associated with different reconstructed tracks was already in place, and I developed it

further in order to streamline its usage. Section 3.3 lists the improvements introduced in the measurement of the fraction of tracks lost in reconstruction, which represent the bulk of my contribution to this area. Both the framework used to build the templates (following the new definitions detailed in section 3.3) and the statistical analysis, with which one can obtain the results presented in Chapter 3.4, were fully developed by me.

Chapter 5 presents the search for exclusive decays of the W boson, which was performed within the ATLAS Standard Model group, in a small analysis team. The search was performed in two final states: track-photon and tau-photon. I was the main analyser of the track-photon final state and carried out the analysis in this final state. The background modelling technique used in both final states is detailed in Chapter 4, and was developed by the Birmingham group, and shared with the rest of the analysis team. The statistical analysis for both final states was carried out by me. My work benefitted greatly from the tools previously developed in the context of the searches for exclusive Higgs and Z boson decays, undertaken by the Birmingham Higgs group: the data processing framework used for the track-photon final state was the common framework designed for this set of searches, with the necessary additions and updates for this specific analysis: the background modelling framework was adapted from the framework used in the $H \to \rho\gamma$ search [3]; and the statistical analysis was developed by me, but based in great part on the framework used previously in the other exclusive decay searches.

Chapter 4 presents a novel non-parametric data-driven background modelling technique, which was presented in Ref. [4] and includes not only work performed by myself, but also by the co-authors of the publication. Two implementations of the technique are discussed. The focus of this chapter is on the implementation which uses ancestral sampling. This method had been designed and utilised previously in the context of the searches for exclusive decays of the Higgs and Z bosons. It was conceived by Andy Chisholm and Konstantinos Nikolopoulos. However, for the first time this method was fully characterised and validated. This is detailed in section 4.2 and corresponds to my contribution to the publication. A new way of implementing the sampling performed is also detailed in this section and was introduced for Ref. [4], and has not been used before in any other publications. Section 4.3 summarises a

second implementation of the background modelling technique, which uses conditional adversarial networks. While I partook in the discussions during its development as a member of the team, I did not directly work in its setup, optimisation and validation, but it is presented here briefly for completeness.

Acknowledgements

After two months of arduous thesis writing I finally come to the moment in which I look back and reflect on everyone who helped me get through the last four years and come out of it with a Dr. in front of my name (hopefully!). I could never be able to thank absolutely everyone who contributed along the way, but I am going to try my best in the next paragraphs to acknowledge the most important ones.

First of all, the person without whom I would absolutely have no PhD is Kostas. Kostas, once while describing you to someone else I said that being your student was always knowing the smartest/more knowledgeable person in the room is your supervisor. I am very grateful for having had you around these last years. You gave me the freedom of knowing that even if I made a mistake or if I didn't know how to tackle a problem, something would be figured out, because you had my back. I must also thank you (and Dimitra, of course!) for all the kindness you showed me throughout the years, but specially while I was alone in Birmingham, without my family, during Covid.

Secondly, I must thank Andy Chisholm, who also has been putting up with my non-stop non-sense since 2019. Thank you for all the advice (physics and otherwise), all the words of comfort, all the jokes, but most of all thank you for having bought a coffee machine some years ago! But in all seriousness, you are one of the kindest people I have ever met, and I will always be grateful to you for all the help you've given me. There are not enough cakes I can bake, or porto bottles I can offer you to make up for the countless times you have listened to me and answered to my silly questions, even when you didn't really have much time to spare.

I also want to thank everyone else in the Birmingham Higgs group, for all the input and help you have given me, and for being a lovely bunch of people. A special thanks must go to some brothers and sister in arms. António, without I would have never

gone to Birmingham! Thank you for being on my corner all these years. I've kept your "Tu consegues!" in front of me since you wrote it down. Rob, thank you for all the rants, for all the coding help, for all the countless sharing of bugs/walls we were hitting. It always helped knowing we were fighting the same battles and that I could always ask for your help. Mihaela, thank you for being a friend, a savage, an all-around cool person, but also let's be honest, a girl. Gov, thank you for keeping me sane, specially in this last stretch. Whenever you need me, I hope you know you can count on me. Thank you for all the pep-talks, all the wise advise, but most importantly for accepting my weird love of the teletubbies. That is the mark of true friendship. I would also like to thank our Gas Lab extended family. The Higgs+Gas Lab social events are some of my favourite memories of Birmingham. On that note, I would also like to thank Panagiotis for organising the epic event which was culture night.

I would be remiss not to thank the greater ATLAS Birmingham group, particularly my office mates, with whom I shared many hours and many meals: Will, Gareth, Eric, Josh, Chonghao, Harry and Adrien (and others which I already mentioned). I also want to thank three people in particular who really got me through the last months in Birmingham: Will, Lex and Jack. You offered me a shoulder to cry on when I needed it the most, and I will never forget it.

There is one final "work" related thanks I must give, which is to the GOAT, Marc Granado. Marquito, you are the greatest person I met in Birmingham, and I am a much better person for having met you and for having shared these last 4 years of PhD with you. At this point I probably should thank Harry again for telling me you were in a band (thank you, Harry). You are not only warm hearted, but also smart; kind but also extremely cool; silly in all the best moments; serious in all the bittersweet ones. Thank you for all the coffees together through video-call during lockdown, for always being patient while I vent, for always calling me back to earth if needed. Your friendship is without a doubt the best thing I take with me from Birmingham.

I must also thank the people in my life who are not part of the Birmingham Particle Physics group, but certainly played an invaluable part in getting this thesis out. First, I would like to thank you Bruno, for all the times you listened to me practising the same presentations over and over. You are, and will always be my lucky charm. I would

also like to thank Ana Luísa and Ricardo Barrué for always making me feel at home whenever I went to CERN. I hope our paths will forever keep crossing. Gonçalo, thank you for monitoring my sanity until the end. You've been doing it since my masters, and I am yet to earn all the friendship you've shown me through the years. A thank you with a sorry!!! must also go to Filipa and Leonor, with whom I had plans made months ago which I cancelled in the last minute, to finish this thesis. I'm sorry I failed you, but I promise I will make it up to you! Laura, thank you of course for the many playlists, but also for being there when I really needed someone who understood. Your strength helped me be stronger. Enrico, thank you for being the reason why I could never regret my choice of going to Birmingham. If I hadn't gone I wouldn't have met you, and you brought light and the greatest happiness to my life.

Finally, I must thank my family, who are the masterminds behind me. I must thank my aunt and uncle for all their help throughout the years, giving me a home during my university years, but also for always being there to help me with anything I need. I hope I can one day repay your kindness. Thank you Filipa, for being my inspiration for wanting to do research. You will forever be the first Doctor in the family. Thank you Elsa, for always being there with a kind word, a warm hug and more importantly, drugs. I also owe you much money you have lent me for me to be able to pursue my studies around the world (and I'm writing this here so it's forever in writing that I need to pay you back). And last, but not least, thank you to my parents without whom there would be no PhD, no masters, no undergraduate degree, no Júlia. Thank you for teaching me hard work from a young age, but also for teaching me kindness and empathy. Sorry for not being a medical doctor, but I'm a Doctor of Particle Physics. Who wouldn't want one of those in a crisis?

To my sisters.

É mais estranho do que todas as estranhezas
E do que os sonhos de todos os poetas
E os pensamentos de todos os filósofos,
Que as coisas sejam realmente o que parecem ser

It's stranger than every strangeness
And the dreams of all the poets
And the thoughts of all the philosophers,
That things are really what they seem to be

Alberto Caeiro

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This chapter introduces theoretical aspects which are relevant to the main topic of this thesis, the search for exclusive hadronic decays of the W boson. These decays have been proposed as novel clean probes of calculation techniques involving non-perturbative effects of the strong interaction. They could also offer a new way to measure the W boson mass through fully reconstructed decays. These motivations are discussed further in the following, and the experimental status of the searches is summarised.

1.1 Standard Model

The Standard Model (SM) is the theory that describes all known elementary particles and their interactions. Arguably, it is the most successful physics model built so far, having been extensively tested through experimental measurements, which have verified multiple SM predictions with large precision. The particles of the SM, summarised in Fig. 1.1, are divided into two categories: fermions, particles with half-integer spin, which are further grouped into quarks and leptons; and bosons, which are the spin-1 gauge bosons ($W^{\pm}$, Z, γ, g) and the Higgs boson, which has spin zero. Gauge bosons are mediators of three fundamental forces: the $W^{\pm}$ and Z bosons mediate the weak force, the photon (γ) mediates the electromagnetic force, and the gluon (g) mediates the strong force. Quarks interact through all these three forces. Neutral leptons - neutrinos - interact only through the weak force, while charged leptons also participate in electromagnetic interactions. There are three generations of leptons, each comprising

a charged lepton - electron (e), muon (μ) and tau (τ), in order of increasing mass - and a corresponding neutrino (ν_e, ν_μ and ν_τ). Quarks are also divided into three generations, each with a quark with electric charge +2/3 (up-type quark), and another with charge -1/3 (down-type quark). Each fermion has also a counterpart anti-fermion of the same mass and opposite charge. Mathematically, the SM is a quantum field theory which features the local gauge symmetry group $SU(3)_c \times SU(2)_L \times U(1)_Y$. It can be divided into the electroweak sector, which describes the weak and electromagnetic interactions and quantum chromodynamics (QCD) which describes strong interactions. In order to respect local gauge invariance under $SU(3)_c \times SU(2)_L \times U(1)_Y$, the gauge bosons of the SM must be massless. However, while gluons and photons are indeed massless, the same is not true of $W^\pm$ and Z bosons. The Brout-Englert-Higgs mechanism, first hypothesised in 1964 [5, 6], allows the generation of masses for the $W^\pm$ and Z bosons, via spontaneous electroweak symmetry breaking. The fundamental prediction of this theory was the existence of a scalar boson, the Higgs boson, which was discovered decades later, by the ATLAS and CMS experiments, at the Large Hadron Collider (LHC), in 2012 [7, 8]. The Higgs boson was the final missing piece of the SM, and its discovery marked a major milestone for particle physics and a great success for the LHC.

1.1.1 Electroweak theory

The electroweak model (also referred to as the Glashow, Weinberg and Salam model) [10–12] unifies the theory of electromagnetic interactions, quantum electrodynamics (QED), and the theory of weak interactions into one common mathematical framework, based on the $SU(2)_L \times U(1)_Y$ local gauge symmetry group.

The theory of weak interactions is based on local gauge invariance under $SU(2)_L$. There are three gauge fields introduced by this symmetry group, $W^i_\mu = (W^1_\mu, W^2_\mu, W^3_\mu)$. The subscript L denotes the chiral nature of the weak theory: lepton and quark fields are described by their chiral components, and only the left-handed components couple to the W^i_μ gauge bosons[1]. Left-handed fermions are grouped in weak isospin doublets, with an isospin of 1/2. Right-handed fermions are instead singlets, with 0 weak isospin.

[1]Chirality is associated with the operators $P_L = \frac{1-\gamma_5}{2}$ and $P_R = \frac{1+\gamma_5}{2}$. Left and right-handed chiral states are the eigenstates of the P_L and P_R operators, respectively.

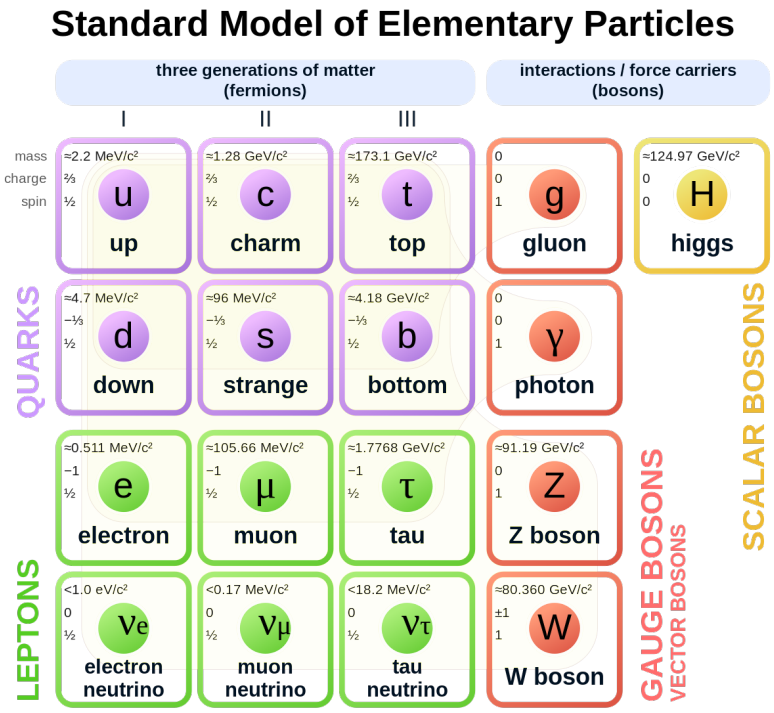


Figure 1.1: Summary of the particles of the SM and their mass, charge and spin [9].

In the case of leptons, there are three left-handed doublets, one for each generation:

$$\begin{pmatrix} I_3 = +1/2 \\ I_3 = -1/2 \end{pmatrix} = \begin{pmatrix} \nu_{eL} \\ e_L^- \end{pmatrix}, \begin{pmatrix} \nu_{\mu L} \\ \mu_L^- \end{pmatrix}, \begin{pmatrix} \nu_{\tau L} \\ \tau_L^- \end{pmatrix}.$$

I_3 stands for the third component of the weak isospin. Left-handed quarks are likewise grouped in isospin doublets, but in this case quark mixing needs to be taken into account. The left-handed quark doublets are, for each generation

$$\begin{pmatrix} I_3 = +1/2 \\ I_3 = -1/2 \end{pmatrix} = \begin{pmatrix} u_L \\ d'_L \end{pmatrix}, \begin{pmatrix} c_L \\ s'_L \end{pmatrix}, \begin{pmatrix} t_L \\ b'_L \end{pmatrix}.$$

d', s' and b' are the weak interaction eigenstates, which relate to mass eigenstates

through the Cabibbo-Kobaiashi-Maskawa (CKM) [13, 14] matrix, in the following way

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.1}$$

This is a unitary 3×3 matrix, which describes quark mixing through charged weak current interactions, mediated by the W^+ and W^- bosons. Using unitarity conditions, this matrix can be re-parametrised using three mixing angles and one phase parameter. These are parameters of the SM which can only be determined through experimental measurements. Each entry of the CKM matrix represents the strength of each possible flavour transition. Their experimentally determined values, taken from Ref. [15], are

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97373 \pm 0.00031 & 0.2243 \pm 0.0008 & (3.82 \pm 0.20) \times 10^{-3} \\ 0.221 \pm 0.004 & 0.975 \pm 0.006 & (40.8 \pm 1.4) \times 10^{-3} \\ (8.6 \pm 0.2) \times 10^{-3} & (41.5 \pm 0.9) \times 10^{-3} & 1.014 \pm 0.029 \end{pmatrix}. \tag{1.2}$$

Diagonal elements are close to unity, showing that the corresponding transitions are favoured. Off-diagonal elements have smaller magnitudes, suppressing the corresponding transitions. Therefore, flavour changing processes between different quark generations are CKM suppressed.

The $\mathrm{SU}(2)_L$ local gauge theory is extended to include electromagnetic interactions by requiring invariance under $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$, where Y stands for weak hypercharge. The hypercharge relates to the electric charge, Q, and I_3 through

$$Y = 2(Q - I_3). \tag{1.3}$$

Local gauge invariance under $\mathrm{U}(1)_Y$ introduces a new field, B_μ. The four aforementioned fields W^i_μ and B_μ are not the physical mediators. The fields corresponding to the physical weak charged current mediators, $W^\pm$, the physical weak neutral current mediator, Z, and the physical electromagnetic interaction mediator, γ, are obtained after spontaneous symmetry breaking, which is discussed in the following section.

1.1.2 The Brout-Englert-Higgs mechanism

Local gauge invariance of the electroweak theory only holds if the gauge fields W^i_μ and B_μ are massless: a term of the form $m^2 A^\mu A_\mu$, with A_μ denoting one of the fields in question, would violate local gauge invariance. The gauge boson mass generation problem in EW theory is solved by spontaneous breaking of the EW symmetry through the Brout-Englert-Higgs (BEH) mechanism [5, 6].

The BEH mechanism breaks the electroweak $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry down to $\mathrm{U}(1)_{EM}$, generating mass terms for the three weak gauge bosons, W^+, W^- and Z, while the photon remains massless. This is achieved through the introduction of a scalar complex field Φ, which is a weak isospin doublet of the form

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{1.4}$$

This is the Higgs field, with four degrees of freedom. The associated potential is

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi \right)^2. \tag{1.5}$$

For $\mu^2 > 0$, $V(\Phi)$ has a minimum at the origin. If instead one considers $\mu^2 < 0$, the potential has an infinite number of minima defined by

$$|\Phi|^2 = \frac{\mu^2}{2\lambda} = \frac{v^2}{2}, \tag{1.6}$$

where the substitution $v = \frac{\mu}{\sqrt{\lambda}}$ was made. The shape of the Higgs potential with multiple minima is represented in Fig. 1.2. v is the vacuum expectation value (VEV). The choice of physical vacuum breaks down the symmetry of the system. After spontaneous electroweak symmetry breaking, the Higgs doublet takes the form

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

The $h(x)$ field is the scalar Higgs boson field, which accounts for one of the four degrees of freedom of Eq. (1.4). There are three broken generators of the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ group so,

according to the Goldstone theorem [16], there are three associated massless Goldstone bosons, accounting for the remaining three degrees of freedom.

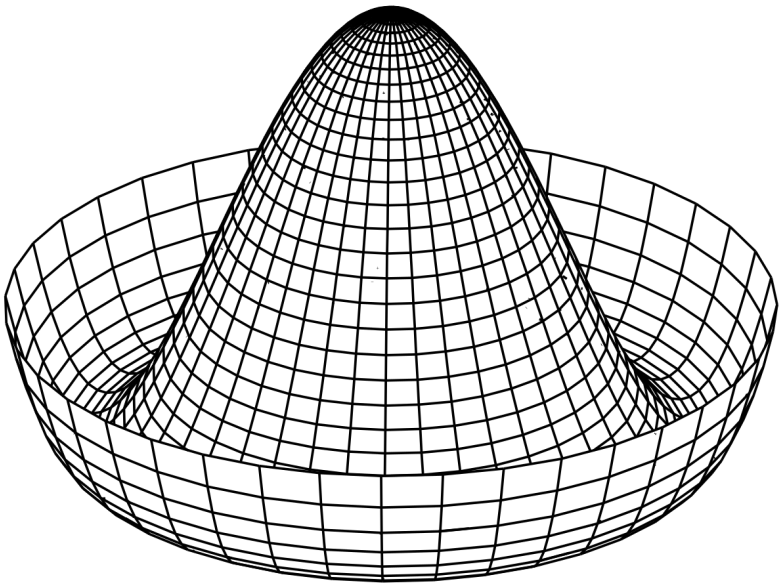


Figure 1.2: Shape of the Higgs potential [17].

The Lagrangian of the Higgs sector is

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi), \tag{1.8}$$

in which the Higgs field couples to W_μ^i and B_μ through the covariant derivative

$$D_\mu \Phi = \left(\partial_\mu + \frac{ig\sigma^a W_\mu^a}{2} + \frac{ig' Y B_\mu}{2} \right) \Phi. \tag{1.9}$$

σ^a are the Pauli matrices, and $i = 1, 2, 3$. The degrees of freedom associated with the Goldstone bosons are absorbed by the gauge fields [18], becoming the longitudinal components of W^+, W^- and Z, which become massive. The physical $W^\pm$, Z, and γ boson fields are given by:

$$W_\mu^\pm \equiv \frac{1}{\sqrt{2}}(W_1 \pm iW_2), \tag{1.10}$$

$$Z_\mu \equiv -B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W, \tag{1.11}$$

$$A_\mu \equiv B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W, \tag{1.12}$$

respectively. The angle θ_W is known as the weak mixing angle, and it is related to the $SU(2)_L$ and $U(1)_Y$ coupling constants, g and g', through the expression

$$\theta_W \equiv \tan^{-1} \frac{g'}{g}, \tag{1.13}$$

and to the elementary electric charge through

$$e = g \sin \theta_W. \tag{1.14}$$

At tree level, the lowest perturbative order, the relation of the weak mixing angle with the $W^{\pm}$ and Z masses is

$$M_W / M_Z = \cos \theta_W. \tag{1.15}$$

The W and Z boson masses relate to v through the expressions:

$$m_W^2 = \frac{g^2 v^2}{4}, \tag{1.16}$$

$$m_Z^2 = \frac{(g'^2 + g^2) v^2}{4}. \tag{1.17}$$

Through their interactions with the Higgs field, fermions acquire a mass given by $mf_i = \frac{y_{f_i} v}{\sqrt{2}}$. In this expression, i runs over the different fermions, and y_{f_i} are the Yukawa couplings, which represent the strength of the interaction between the fermion and the Higgs field.

1.1.3 Quantum Chromodynamics

Quantum chromodynamics is the theory that describes strong interactions. The relevant symmetry group is $SU(3)_c$, where c stands for colour, the charge associated with the strong interactions. Quarks and gluons carry this charge. There are three different possible colours, red, green and blue (R, G, B), and three anti-colours ($\overline{R}$,$\overline{G}$,$\overline{B}$). Quarks have colour (also called positive colour charge) while anti-quarks have anti-colour (negative colour charge). Different colour and anti-colour combinations make up eight

different possible gluons. Since gluons carry colour charge, they can self-couple[2].

One of the parameters of QCD is the strong coupling constant, α_s, which is a function of the energy scale, ${Q_E}^2$. This "running" effect for α_s originates from vacuum polarisation due to quark-antiquark ($q\bar{q}$) and gluon-gluon interactions. In QED, there is an analogous vacuum polarisation effect which arises from the continuous generation of short-lived fermion-antifermion pairs, due to fluctuations of the vacuum. In the presence of a charged particle, these pairs become polarised. For instance, if the particle has negative charge, the positively charged fermions will arrange closer to the fixed particle, while the negatively-charged fermions will be further away. This effectively screens the charged particle, reducing the effect of the electric charge for increasing distance. In the case of QCD, quark-antiquark pairs generate an analogous effect, screening the colour charge. However, due to gluon self-interactions there is a competing anti-screening effect. The resulting coupling constant is given by

$$\alpha_s({Q_E}^2) = \frac{\alpha({\mu_E}^2)}{1 + [\alpha_s({\mu_E}^2)/12\pi](33 - 2n_f)\ln\left(\frac{{Q_E}^2}{{\mu_E}^2}\right)}. \tag{1.18}$$

With this expression, given a known value of α_s for a fixed energy scale ${\mu_E}^2$, one can obtain α_s corresponding to another energy scale ${Q_E}^2$ [17]. n_f is the number of active quark flavours, i.e. flavours that effectively contribute to loops, which are those with mass $m_f < |Q_E|$. One can define a scale ${Q_E}^2 \equiv \Lambda^2_{QCD}$, for which the denominator in Eq. (1.18) is 0 and α_s diverges. This scale is therefore obtained by imposing:

$$\ln\left(\frac{\Lambda^2_{QCD}}{{\mu_E}^2}\right) = -\frac{12\pi}{(33 - 2n_f)\alpha_s({\mu_E}^2)}. \tag{1.19}$$

α_s can be redefined in terms of Λ_{QCD}:

$$\alpha_s({Q_E}^2) = \frac{12\pi}{(33 - 2n_f)\ln\left({Q_E}^2/\Lambda^2_{QCD}\right)}. \tag{1.20}$$

For high ${Q_E}^2$ values (short distances), α_s becomes small. This is the "asymptotic

[2] Z bosons can self-couple as well. However, in QED this is not observed: since the photon is electrically neutral, it cannot couple to itself.

freedom" regime, where quarks become asymptotically free. Perturbative QCD calculations are valid in this regime. On the other hand, as one approaches Λ_{QCD} scale, for small ${Q_E}^2$ (long distances), α_s diverges and perturbative theory breaks down. At these scales, other methods need to be used for calculations, such as lattice QCD. Due to the anti-screening effect arising from gluon self-interactions, the quark-antiquark potential increases with distance. As a $q\bar{q}$ pair are pulled apart, the potential becomes increasingly larger, and it becomes energetically favourable to produce two new $q\bar{q}$ pairs confined in two hadrons. There are different types of hadrons, the most common being mesons, which have a valence quark and a valence antiquark; and baryons, which are composed by three valence quarks. All quark flavours hadronise, except the top quark, which decays before hadronisation can take place.

When calculating cross sections of processes involving hadrons, QCD factorisation theorems are applied [19]. These theorems allow the separation of long distance perturbative effects from short distance non-perturbative effects, which are factorised in the calculations. The cross section for a process with two hadrons h_1 and h_2 in the initial state, and an arbitrary particle X in the final state, of the form $h_1 h_2 \to X$ can be written as

$$\sigma_{h_1 h_2 \to X} = \sum_{a,b} \int_0^1 dx_a dx_b f_{a/h1}(x_a, \mu_F) f_{b/h2}(x_b, \mu_F) \times \sigma_{ab \to X}(\mu_F, \mu_R). \tag{1.21}$$

In this expression, a and b run over the partons in hadrons h_1 and h_2. $\sigma_{ab \to X}$ is the parton-level cross section, associated with the parton scattering process $ab \to X$. This cross section can be written as a perturbative series in terms of σ_s, of the form $[\sigma_0 + \alpha_s(\mu_R^2)\sigma_1 + ...]$. μ_R is the renormalisation scale at which α_s is evaluated. $f_{a/h1}$ and $f_{b/h2}$ are parton distribution functions (PDFs) describing h_1 and h_2. At leading order, they can be interpreted as the probabilities of finding partons a and b with momentum fractions x_a and x_b within h_1 and h_2. These PDFs are determined at a factorisation scale μ_F, which can be interpreted as the scale that separates the long distance from the short distance physical regime. PDFs cannot be calculated in a perturbative manner and are instead determined from fitting experimental data from multiple experiments and processes. On the other hand, a perturbative framework can be used to calculate the evolution of

PDFs with Q^2. This can be achieved using Dokshiter-Gribov-Lipatov-Alterelli-Parisi (DGLAP) equations [20–22]. This means that measurements at specific energy scales can be used to obtain PDFs used in cross section calculations at different energy scales, after DGLAP evolution. The dependence of the cross section on the renormalisation and factorisation scales originates from the need to truncate the perturbative series included in the calculations. In theory, if all orders were included, the cross section would be invariant under different μ_F and μ_R scales. However, in practise it is impossible to include all orders, and different scales will lead to different results for the cross section. Therefore there is an uncertainty associated with the chosen scale, which becomes smaller with the inclusion of higher order terms in the calculations. For the case of the process $h_1 h_2 \to X$, typically the scale used is $\mu_F = \mu_R = M_X$.

1.2 W boson physics

1.2.1 W boson production in pp collisions

At hadron colliders, such as the Large Hadron Collider (LHC), $W^{\pm}$ bosons are created primarily by quark-antiquark annihilation processes. These are called Drell-Yan processes [23]. W boson production cross sections in pp collisions can be calculated using the QCD factorisation formalism, described in the previous section. The total cross section is obtained through the product of non-perturbative proton PDFs and the parton level cross sections $\sigma_{q\bar{q}\to X}$, which can be computed using perturbation theory. Fig. 1.3 shows the Feynman diagrams at leading order for the parton subprocesses $u\bar{d} \to W^+$ and $d\bar{u} \to W^-$, and subsequent $W^{\pm} \to l^{\pm}\nu$ decays. At higher perturbative orders there are contributions with partons being created along with the W boson. These contributions manifest as W + jet events in the data. The transverse component of the W boson momentum arises from these higher order processes. At leading order, and without accounting for non-perturbative effects of the incoming partons, the W bosons are predicted to be produced with all the momentum along the beam line and $p_{\mathrm{T}}(W) = 0$.

At leading order, for $p_{\mathrm{T}}(W) = 0$, the W will be produced mainly in the direction of the higher momentum parton in the hard scattering process. In pp collisions, this

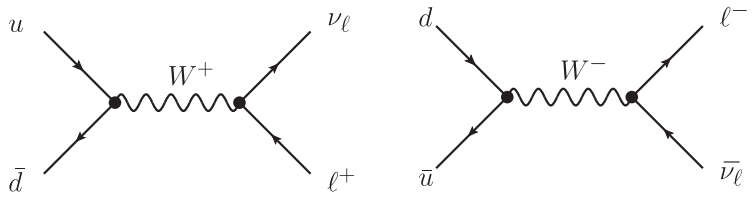


Figure 1.3: Leading order Feynman diagram for the partonic subprocesses $u\bar{d} \to W^+$ and $d\bar{u} \to W^-$, and the subsequent decay into a charged lepton and a neutrino [19].

will likely be the quark and not the anti-quark, since quarks carry on average a higher momentum fraction than anti-quarks (as indicated by proton PDFs). Furthermore, anti-quarks are only present in the sea of the colliding protons. Considering the subprocess in Fig. 1.3, the W^+ is produced in the direction of the u quark and the W^- in the direction of the d quark. Up quarks are more likely to carry higher momentum fractions than down quarks, which creates a difference in the rapidity[3] distributions of W^+ and W^- bosons. Fig. 1.4 shows the predicted rapidity distributions for $W^\pm$ bosons for pp collisions at $\sqrt{s} = 14$ TeV. W^- bosons tend to have lower rapidity, being distributed more centrally than W^+ bosons.

Using similar arguments, one can also draw some conclusions regarding W boson polarisation in pp collisions. Due to the left-handed nature of the electroweak interactions, quarks will be left-handed and anti-quarks will be right-handed in the processes represented in Fig. 1.3. By helicity conservation, and given that the W boson is produced in the direction of the quark, at leading order the W bosons will be approximately purely left-handed. For W + jet production and higher $p_{\mathrm{T}}(W)$, W bosons are produced preferentially with left-handed polarisation at the LHC, but there are also contributions from right-handed and longitudinal polarised W bosons. These are quantified by the polarisation fractions, f_L, f_R, f_0, for left-handed, right-handed and longitudinal polarisation, respectively. Ref. [24] presents predictions for the polarisation fractions of W bosons at the LHC, as a function of $p_{\mathrm{T}}(W)$. These can be found in Fig. 1.5.

W boson production cross section measurements are a benchmark for the understanding of QCD and the fine-tuning of PDFs. A summary of such measurements performed by the ATLAS collaboration, in proton-proton collisions, for different center-of-mass

[3]Rapidity is defined as $\mathrm{y} = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right)$, in which p_z is the z component of the momentum of the particle and E is its energy. This quantity is further discussed in the next chapter.

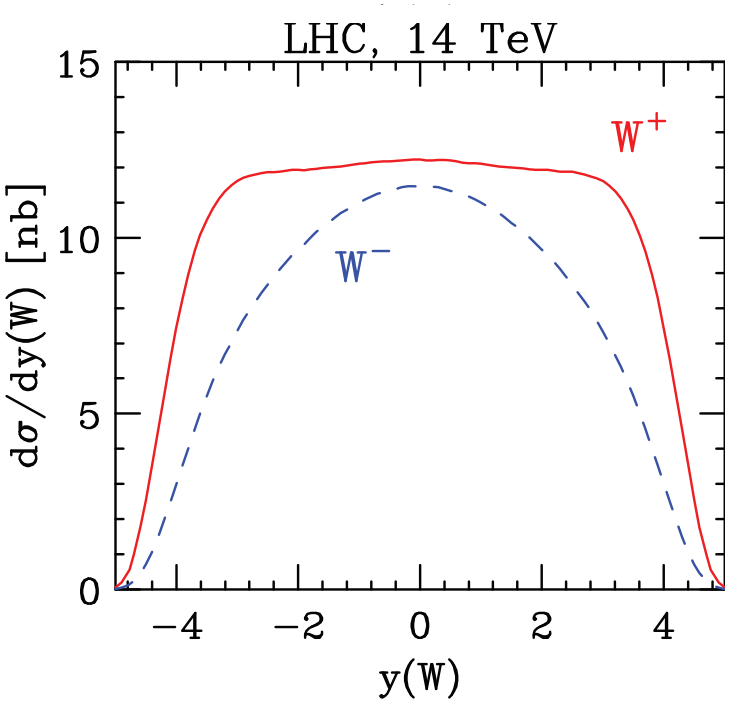


Figure 1.4: Expected rapidity distributions for W^+ and W^- bosons, at the LHC for $\sqrt{s} = 14$ TeV [19]. These predictions are obtained at leading order.

energies ($\sqrt{s}$), can be found in Fig. 1.6. The $W^\pm \to l^\pm \nu$ decay channel was targeted in these measurements. There exists a clear linear dependence on $\sqrt{s}$. The experimental results are compared to next-to-next-to-leading order theory predictions, with very good agreement being observed. At $\sqrt{s} = 13$ TeV, the measured inclusive cross sections are $\sigma_{W^+} \times \mathcal{B}(W \to l\nu) = 11.83 \pm 0.41$ nb and $\sigma_{W^-} \times \mathcal{B}(W \to l\nu) = 8.79 \pm 0.30$ nb [25]. For the same $\sqrt{s}$, the reported predictions at next-to-next-to-leading order (NNLO) are $\sigma_{W^+} \times \mathcal{B}(W \to l\nu) = 11.54 \pm 0.38$ nb and $\sigma_{W^-} \times \mathcal{B}(W \to l\nu) = 8.54 \pm 0.28$ nb. There is a clear charge asymmetry: there are more W^+ bosons produced at the LHC than W^- bosons. This is explained by the fact that the colliding protons are composed by two valence up quarks, which initiate W^+ production; and one valence down quark, which in turn initiates W^- production. The predicted and measured W boson production charge asymmetry at the LHC, defined as the difference between the W^+ and the W^- boson cross sections normalised to the total $W^\pm$ cross section, can be found in Fig. 1.7, as a function of $|y_W|$. This measurement was performed in pp collisions, at $\sqrt{s} = 13$ TeV, by the CMS experiment [26]. As $|y_W|$ increases, so does the charge asymmetry. This is compatible with what is seen in Fig. 1.4.

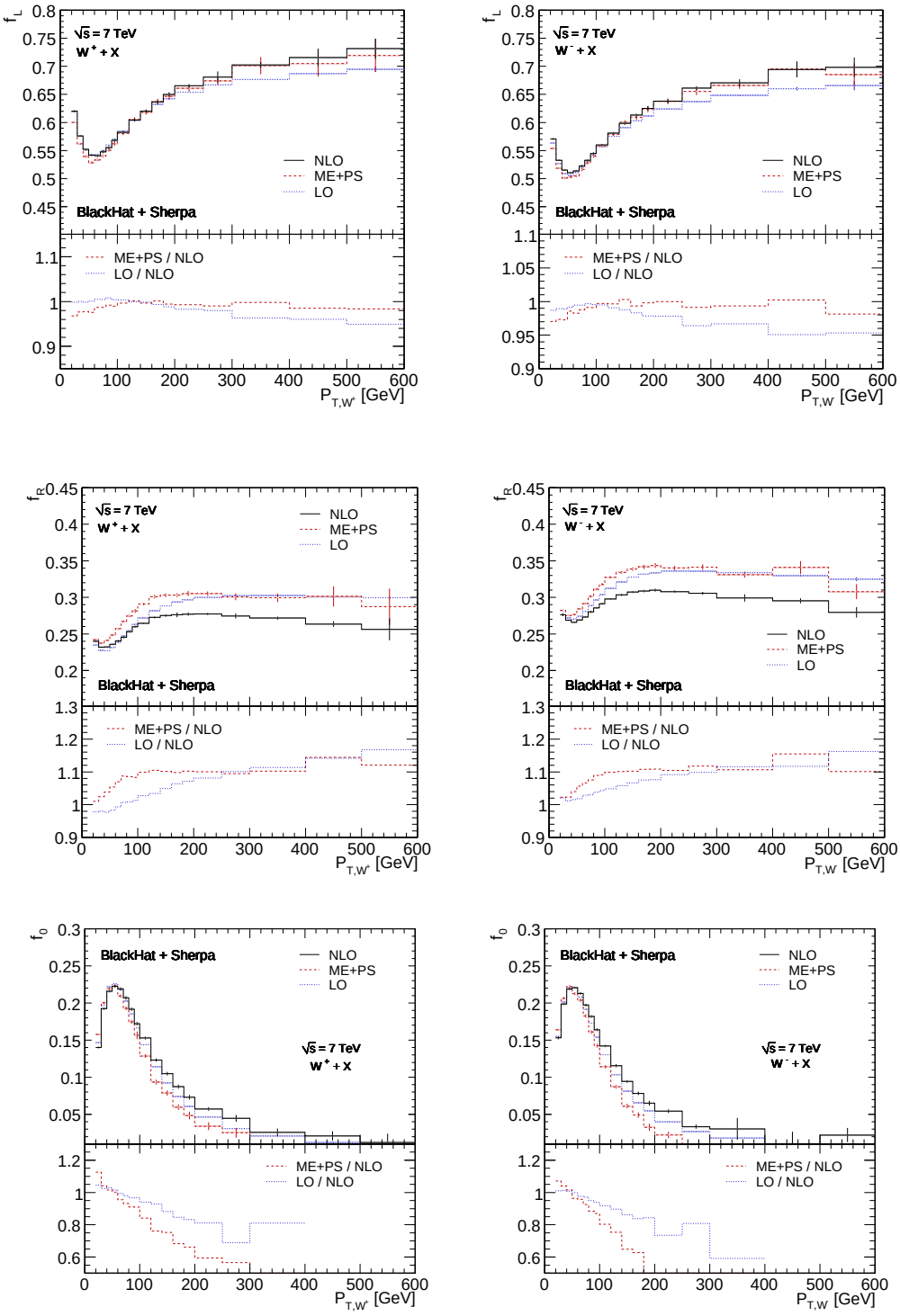


Figure 1.5: Predictions for the f_L (top), f_R (middle) and f_0 (bottom) polarisation fractions for W^+ (left) and W^- (right) bosons, at the LHC for $\sqrt{s} = 7$ TeV, as a function of $p_{\mathrm{T}}(W)$, for $W +$ jet events [24].

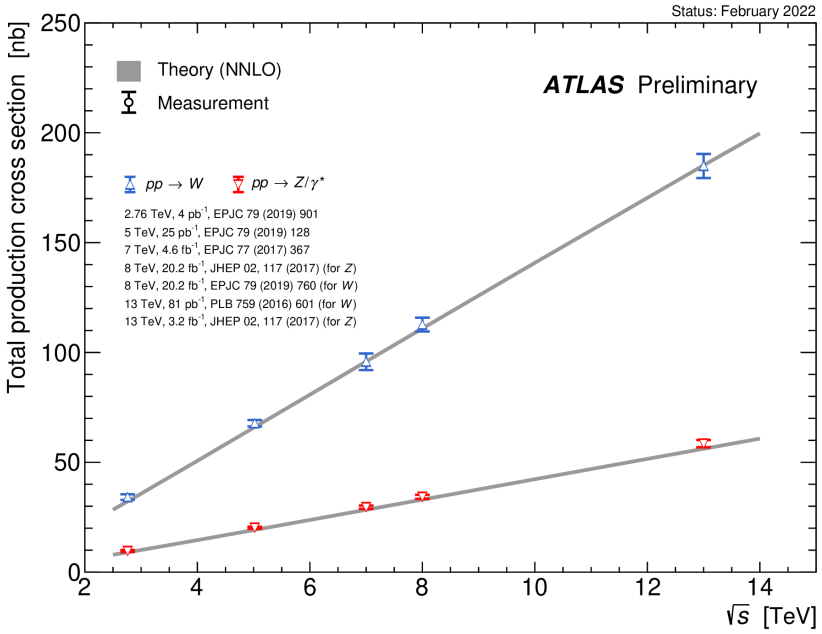


Figure 1.6: Summary of W and Z boson production cross section measurements performed by ATLAS, as a function of center-of-mass energy [27]. The measurements are compared to the next-to-next-to-leading order theory expectation.

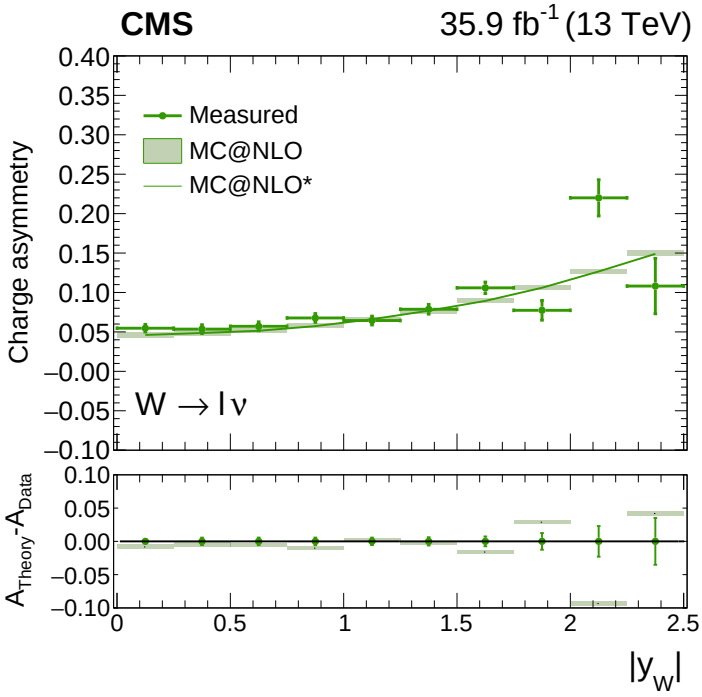


Figure 1.7: W boson production charge asymmetry as a function of $|y_W|$, measured in pp collisions, at $\sqrt{s} = 13$ TeV, by the CMS experiment [26]. The charge asymmetry is defined as the difference between the W^+ and the W^- boson cross sections, normalised to the total $W^\pm$ cross section. Production is measured in $W \to l\nu$ decays. The measured asymmetry is compared with next-to-leading order theory predictions.

1.2.2 Exclusive hadronic decays of the W boson

The W boson has been studied extensively since its discovery, mostly through its leptonic and inclusive hadronic decays. The current picture of the experimental measurements of the decay modes of the W boson is summarised in Table 1.1. None of the exclusive hadronic decays of the W boson, that are predicted by the standard model, have been observed. This is also true for the analogous Z boson decays. A group of these relatively untapped decays, radiative decays of the type $V \to M + \gamma$ (where V is a vector boson), have been proposed as novel probes of the QCD factorisation framework [28].

W decay mode	Fraction (Γ_i/Γ)
$l\nu$	$(10.86 \pm 0.09)\%$
$e\nu$	$(10.71 \pm 0.16)\%$
$\mu\nu$	$(10.63 \pm 0.15)\%$
$\tau\nu$	$(11.38 \pm 0.21)\%$
hadrons	$(67.41 \pm 0.27)\%$
$\pi\gamma$	$< 7 \times 10^{-6}$ at the 95% CL
$D_s\gamma$	$< 6.5 \times 10^{-4}$ at the 95% CL
cX	$(33.3 \pm 2.6)\%$
$c\bar{s}$	$(31^{+13}_{-11})\%$
invisible	$(1.4 \pm 2.9)\%$
$\pi\pi\pi$	$< 1.01 \times 10^{-6}$ at the 95% CL

Table 1.1: Summary of experimental measurements of the decay modes of the W boson [29, 30]. l indicates each type of lepton, not sum over them. X means any particle produced in association with c.

In the context of exclusive processes with individual, highly energetic mesons in the final state, amplitudes are expressed as convolutions of perturbative hard-scattering coefficients with non-perturbative meson light-cone distribution amplitudes (LCDAs), which account for bound-state effects. At leading order, these LCDAs can be interpreted as the amplitudes associated with finding a quark inside the meson in the decay with longitudinal momentum fraction x. They cannot be computed using perturbation theory, so lattice QCD, QCD sum rules or experimental results need to be employed in the calculations. Ultimately, the amplitudes of exclusive decays with a single highly energetic hadron in the final state can be expanded in powers of Λ_{QCD}/E, where E is the energy released to the final state. While leading order terms can be precisely

predicted, the same is not true for higher order corrections, decreasing the precision of predictions. Exclusive hadronic decays of the W and Z bosons are an ideal laboratory to test the QCD factorisation framework, since the high energy release in these processes means higher order terms of Λ_{QCD}/E become negligible. As such, the observation of these decays would offer clean probes of QCD factorisation, making it possible to extract information about the meson LCDAs, and to obtain more rigorous predictions of for processes involving non-perturbative QCD effects. Such predictions become more essential as larger datasets become available and experimental measurements become limited by modelling theoretical uncertainties.

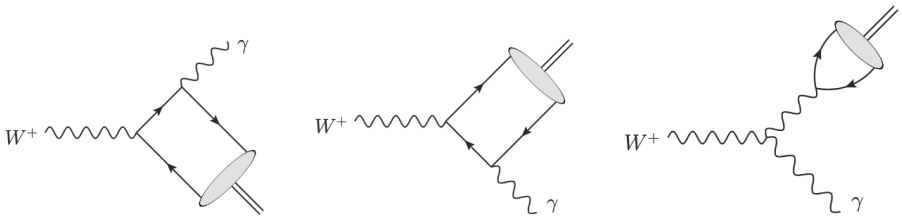


Figure 1.8: Representation of the exclusive hadronic W decays from Ref. [28]. In this case the $W^+ \to M^+\gamma$ decays are shown.

The central focus of this thesis will be on the search for exclusive radiative hadronic decays of the W boson. The formation of the meson in the decays is allowed by the emission of a photon by one of the quarks or the W boson, as is represented in the Feynman diagrams of Fig. 1.8. The latest theoretical predictions for these branching fractions, calculated in the QCD factorisation framework and presented in Ref. [28], are listed in Table 1.2. These predictions have relatively high uncertainties, which are dominated by the uncertainties associated with the meson LCDAs. Furthermore, one can find different predictions in the literature. For instance, the reported predictions for the $W^\pm \to \pi^\pm\gamma$ decay range from $\mathcal{O}(10^{-7})$ to $\mathcal{O}(10^{-9})$ [28, 31–33]. As listed in Table 1.1, experimental constraints have been previously set on the branching fractions of the $W^\pm \to \pi^\pm\gamma$ by the CDF Collaboration [34], and $W^\pm \to D_S^\pm\gamma$ decay channels by a recent LHCb analysis [30]. The remaining decays listed in Table 1.2 have remained largely experimentally unexplored. Limits have also been set on the fully hadronic exclusive decay $W^\pm \to \pi^\pm\pi^\mp\pi^\pm$ [35]. For the Z boson, many such searches have been performed

and many more limits have been set on its exclusive radiative hadronic decays [36]. Measuring these types of decays of the W and Z bosons is particularly challenging not only due to how rare they are but also due to the high multiplicity environment of the LHC, and the associated large backgrounds. These searches represent an interesting triggering, reconstruction and background modelling challenge.

Decay Channel	SM Branching Fraction
$W^{\pm} \to \pi^{\pm}\gamma$	$(4.0 \pm 0.8) \times 10^{-9}$
$W^{\pm} \to \rho^{\pm}\gamma$	$(8.7 \pm 1.9) \times 10^{-9}$
$W^{\pm} \to K^{\pm}\gamma$	$(3.3 \pm 0.7) \times 10^{-10}$
$W^{\pm} \to K^{*\pm}\gamma$	$(4.8 \pm 1.4) \times 10^{-10}$
$W^{\pm} \to D_S^{\pm}\gamma$	$(3.7 \pm 1.6) \times 10^{-8}$
$W^{\pm} \to D^{\pm}\gamma$	$(1.4 \pm 0.5) \times 10^{-9}$
$W^{\pm} \to B^{\pm}\gamma$	$(1.6 \pm 0.8) \times 10^{-12}$

Table 1.2: Predicted branching fractions for various $W \to M\gamma$ decays [28].

The work presented in this thesis targets the $W^{\pm} \to \pi^{\pm}\gamma$, $W^{\pm} \to \rho^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ decays. No experimental constraints have been set on $\mathcal{B}(W^{\pm} \to \rho^{\pm}\gamma)$ and $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma)$, prior to this work. In the case of the $W^{\pm} \to \pi^{\pm}\gamma$ decay, the CDF experiment set a limit of $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma) < 7 \times 10^{-6}$, at the 95% CL [34]. This search was performed using a $p\overline{p}$ collision dataset collected with a center-of-mass energy of $\sqrt{s} = 1.96$ TeV and corresponding to a luminosity of 4.3 fb^{-1}. Inclusive photon triggers were used that target events with at least one electromagnetic calorimeter cluster with a transverse energy larger than 25 GeV. The offline analysis selected events with a back-to-back isolated photon ($E_T > 25$ GeV) and isolated charged pion candidate ($p_{\mathrm{T}} > 25$ GeV). The expected background was estimated in a fit to the W boson invariant mass sidebands, using an exponential function. The same trigger used in the selection was used to target $W^{\pm} \to e^{\pm}\nu$ events, and the limit setting was performed for $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma)/\mathcal{B}(W^{\pm} \to e^{\pm}\nu)$, which allowed to cancel systematic uncertainties common to both analyses. The distribution of data events as a function of the W boson invariant mass, compared to the background expectation, can be found in Fig. 1.9a.

More recently, the CMS experiment also performed a search for $W^{\pm} \to \pi^{\pm}\gamma$, setting an upper limit on its branching fraction of $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma) < 1.5 \times 10^{-5}$ at 95% CL [37], a looser limit than the one previously set by CDF. A dataset collected in pp collisions

at $\sqrt{s} = 13$ TeV was employed, corresponding to an integrated luminosity of 137 fb^{-1}. The search targeted W bosons produced in top quark-antiquark ($t\bar{t}$) events. In these events, two W bosons are created. This choice allowed to use muon and electron triggers, selecting events in which one of the W bosons in the event decays leptonically. The search for the $W^{\pm} \to \pi^{\pm}\gamma$ decay is then performed using the other W boson produced in the event. A multivariate analysis technique, namely a boosted decision tree (BDT) classifier trained on signal and background simulated samples, was employed to discriminate signal and background. The output of the BDT classifier was used to define a control region, which was employed to estimate the shape of the expected background. The functional form determined from a fit to the W boson invariant mass distribution in data in the control region, was subsequently used in the fit to data in the signal region. The results of the fit to the W boson invariant mass in the signal region can be found in Fig. 1.9b. This search was limited by the $t\bar{t}$ production cross section, which is approximately 20 times smaller than the inclusive W boson production cross section [38].

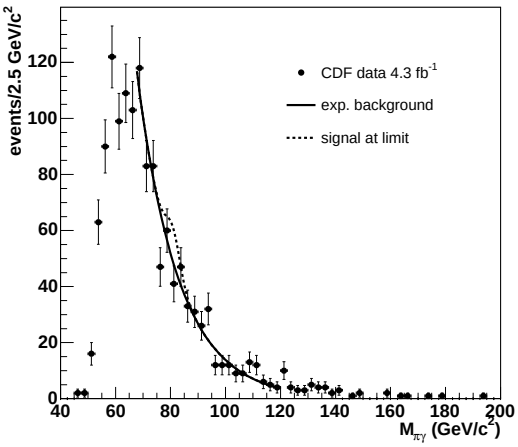
1.2.3 W boson mass measurement

At the lowest perturbative order, the mass of the W boson is given by Eq. (1.15). When taking into account radiative corrections Δr, the W boson mass can be written as

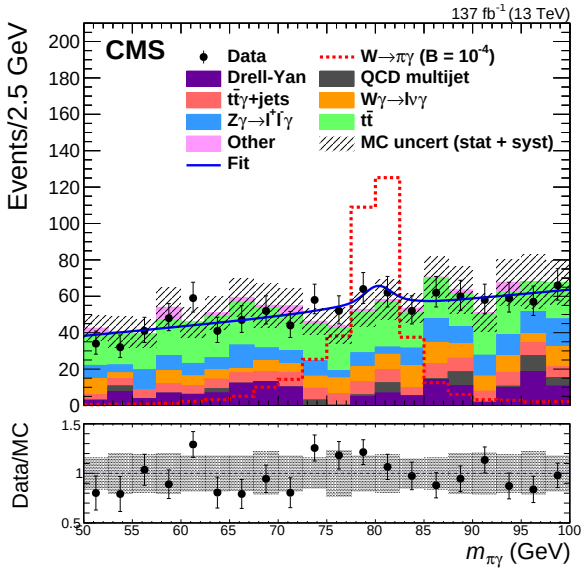
$$M_W^2 = \frac{\sqrt{2}e^2}{8G_F \sin\theta_W (1 - \Delta r)}, \tag{1.22}$$

where G_F is the Fermi constant [39]. These radiative corrections are due to heavy quark and Higgs loops, and introduce additional dependences on the masses of the top quark and the Higgs boson. As a consequence, the measurements of the W and Z boson masses and of the other electroweak parameters made predictions of the top mass possible, before its discovery by CDF in 1995 [40]. Subsequently, with this added information on the top mass, a broad prediction was also available for the Higgs boson mass that aided its eventual discovery in 2012.

The global electroweak fit to measurements of the relevant SM parameters using the state-of-the-art high order theoretical calculations yields a prediction for the W



(a) W boson invariant mass distribution, observed in the $W^{\pm} \to \pi^{\pm}\gamma$ search, performed by the CDF experiment [34]. The expected background and the expected signal at the 95% CL limit are shown.



(b) W boson invariant mass distribution, observed in the $W^{\pm} \to \pi^{\pm}\gamma$ search, performed recently by the CMS experiment [37]. The signal is shown normalised to a 10^{-4} branching fraction. The different background contributions are also shown, stacked on top of each other. The ratio in the lower panel is taken between the data and the expected background from simulation.

Figure 1.9: Distribution of data events as a function of W boson invariant mass, in the search for $W^{\pm} \to \pi^{\pm}\gamma$ performed by CDF (left) and CMS (right).

boson mass of $M_W = 80354 \pm 7$ MeV [41], a result with a precision higher than 0.1 per mille. New particles would introduce further radiative corrections to M_W, changing its value. This provides motivation to measure the W boson mass with the same level of precision as the one of the theoretical prediction. The current world average is $M_W = 80377 \pm 12$ MeV [29], in agreement with the predicted value.

A recent publication from the CDF collaboration [42] reported $M_W = 80433 \pm 9$ MeV [42], a result with a precision higher than the one obtained from the combination of all previous results. This result is significantly above the SM prediction (7σ), and renewed the scientific community's interest in the W boson mass measurement. Since the publication of the CDF result, ATLAS has released an improved W boson mass measurement, after a re-analysis of their $\sqrt{s} = 7$ TeV dataset, reporting $M_W = 80360 \pm 16$ MeV, compatible with the SM value [43]. This measurement achieved a 16% improvement in precision over the previous ATLAS result. An overview of the different M_W measurements and their comparison with the SM prediction can be found in Fig. 1.10.

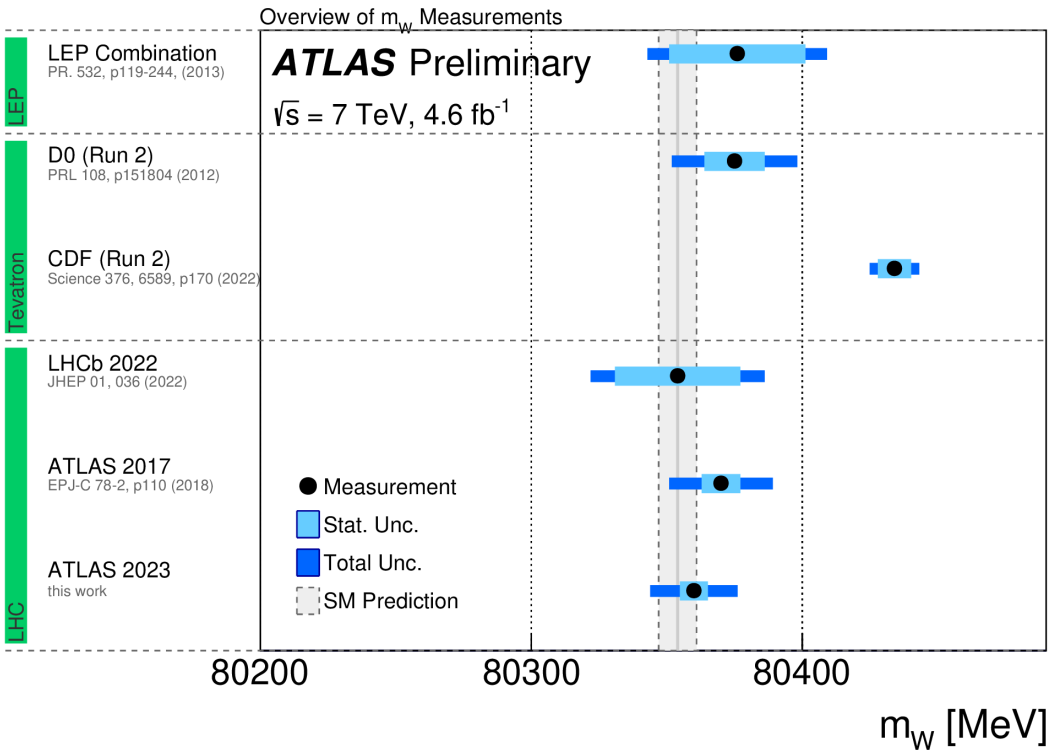


Figure 1.10: Summary of different W boson mass measurements, including the latest results by CDF and ATLAS [43]. The SM prediction is represented in grey.

The aforementioned measurements have exploited W boson leptonic decay channels

($W \rightarrow l\nu$) in which the neutrino is not reconstructed and consequently neither is the W boson invariant mass. Systematics arising from the incomplete kinematics in reconstruction are often leading uncertainties in the analyses [43]. Radiative decays of the W boson could enable a W boson mass measurement with fully reconstructed final states [31, 44]. Specifically, the $W^{\pm} \rightarrow \pi^{\pm}\gamma$ decay includes a track and photon in the final state, and the associated high resolutions at the ATLAS experiment for these objects could be exploited in the mass measurement. The small predicted SM branching fractions for these decays makes it unlikely that such a measurement could be achieved in the next years. Nevertheless, these could become possible as larger datasets become available in the next decades, and at future colliders.

1.3 Thesis Outline

This thesis presents searches for exclusive hadronic decays of the W boson, performed using proton-proton collision data collected by the ATLAS experiment at $\sqrt{s} = 13$ TeV. Furthermore, a novel non-parametric data-driven background modelling technique is presented.

The next chapter contains a description of the ATLAS detector and of the reconstruction of physics objects, with a greater detail given to tracking. The particularities of tracking in dense charged particle environments are also introduced in Chapter 2. Chapter 3 expands on this discussion on tracking in dense environments and presents a data-driven methodology to measure tracking efficiency in dense environments, quantified by the fraction of lost tracks. Improvements to this methodology developed by the author of this thesis and results obtained using 2018 ATLAS data are presented. Chapter 4 discusses the development of a novel non-parametric data-driven background modelling technique. Two implementations of this method are proposed and exemplified using case studies. Furthermore, ensemble tests are employed in order to thoroughly validate the performance of the background modelling method. In Chapter 5, the searches for three exclusive hadronic decays of the W boson, $W^{\pm} \rightarrow \pi^{\pm}\gamma$, $W^{\pm} \rightarrow K^{\pm}\gamma$ and $W^{\pm} \rightarrow \rho^{\pm}\gamma$, are presented. The first experimental constrains on the branching fractions of the latter two decays are set. The most stringent upper limit on $\mathcal{B}(W^{\pm} \rightarrow$

$\pi^{\pm}\gamma$) is also reported. These searches were enabled to a large extent by the use of the background modelling technique presented in Chapter 4.

Chapter 2

The ATLAS experiment at the LHC

In this chapter, an introduction to the Large Hadron Collider and the ATLAS experiment is provided. The reconstruction of physics objects is also briefly discussed. Track reconstruction is presented in more detail, serving as an introduction to the tracking performance work presented in the next chapter.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) with a circumference of approximately 27 km is the world's largest particle collider, which was designed to collide protons up to a record center-of-mass energy (currently $\sqrt{s}$ = 13.6 TeV) [45]. It is operated by the European Organisation for Nuclear Research (CERN) and located at the border between Switzerland and France, near the city of Geneva. The collider resides in an underground tunnel, which previously was used to house the Large Electron-Positron Collider (LEP). Besides proton-proton (pp) collisions, the LHC collides lead nuclei in $PbPb$ collisions, and in pPb. Heavy ion collisions allow the study of QCD in high energy density regimes, where quarks and gluons become de-confined, and form quark-gluon plasma (QGP).

Protons and nuclei pass through a sequence of accelerators and are subsequently injected into the LHC ring. Two counter-rotating beams are focused in four interaction points (IP), where four detectors are located. These detectors are: A Toroidal LHC Apparatus (ATLAS) and the Compact Muon Solenoid (CMS), which are general purpose detectors; the Large Hadron Collider beauty (LHCb), which specialises in flavour

physics involving b hadrons; and A Large Ion Collider Experiment (ALICE), which is optimised for the study of the QGP medium through heavy ion collisions. Fig. 2.1 shows a schematic of the CERN accelerator complex. The series of accelerators through which particles travel are represented, as well as the four interaction points of the LHC.

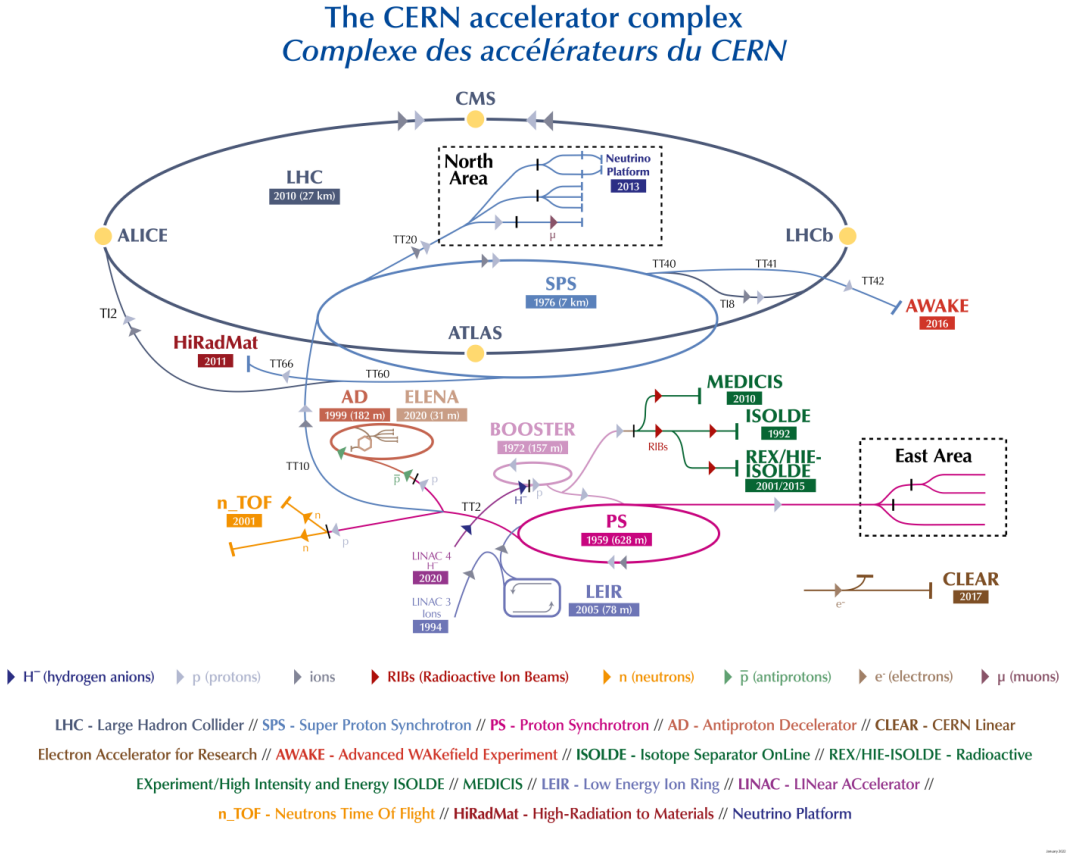


Figure 2.1: Representation of the CERN accelerator complex [46].

In the case of pp collisions, a maximum of 2808 bunches are collided, each with up to 1.15×10^{11} protons, and with a bunch spacing of 25 ns [45] and bunch crossing rate of 40 MHz. The values of the beam parameters define the instantaneous luminosity, $\mathcal{L}_{\text{inst}}$, which is the ratio between the rate of events generated per unit time and the cross section of the process in question. It is given by

$$\mathcal{L}_{\text{inst}} = n_b \frac{\langle \mu \rangle f_R}{\sigma_{\text{inel}}}, \tag{2.1}$$

where σ_{inel} is the inelastic cross section, n_b is the number of colliding bunches, $\langle \mu \rangle$ is the average number of inelastic interactions per bunch crossing, and f_R is the revolution

frequency, which is 11246 Hz for protons [47]. The total integrated luminosity can be obtained by integrating $\mathcal{L}_{\text{inst}}$ in time

$$\mathcal{L} = \int \mathcal{L}_{\text{inst}} dt. \tag{2.2}$$

Out of all pp collisions, only a small fraction is interesting for further analysis, and therefore it is unlikely that two happen in the same bunch crossing. The extra "pile-up" interactions present a challenge to physics objects reconstruction algorithms, as it becomes harder to resolve the particles created in the pp collision of interest from pile-up particles.

The periods of operation of the LHC are called "Runs", which are separated by long shut-down periods. Currently, Run 3 of the LHC is ongoing, and pp collision data is being collected at $\sqrt{s} = 13.6$ TeV. The work presented in this thesis uses pp collision data from Run 2 of the LHC, which started in 2015 and finished in 2018. Collisions with a center-of-mass energy of $\sqrt{s} = 13$ TeV were recorded. Fig. 2.2 shows the cumulative luminosity delivered by the LHC and recorded by ATLAS during Run 2. Events are only considered for physics analysis if they were recorded during data taking periods in which all relevant detector modules were operational. 'Good Run Lists' (GRL), which filter the portion of the recorded dataset that has been identified as "good-for-physics", are employed for this effect [48]. The "good-for-physics" Run 2 dataset corresponds to a total integrated luminosity of 140 fb^{-1} [47]. The pile-up distributions, measured by ATLAS during each year of Run 2 are shown in Fig. 2.3. The average pile-up in Run 2 was $\langle\mu\rangle = 33.7$, which is larger than the Run 1 average ($\langle\mu\rangle = 18.5$), and lower than the Run 3 average ($\langle\mu\rangle = 46.5$).

2.2 A Toroidal LHC Apparatus

The ATLAS detector [50] is one of the general purpose detectors of the LHC. It was designed to be capable to search for the SM Higgs boson, and to study its properties, after discovery. Besides the study of the Higgs sector, ATLAS has a broad physics programme, performing multiple SM precision measurements and searches for physics

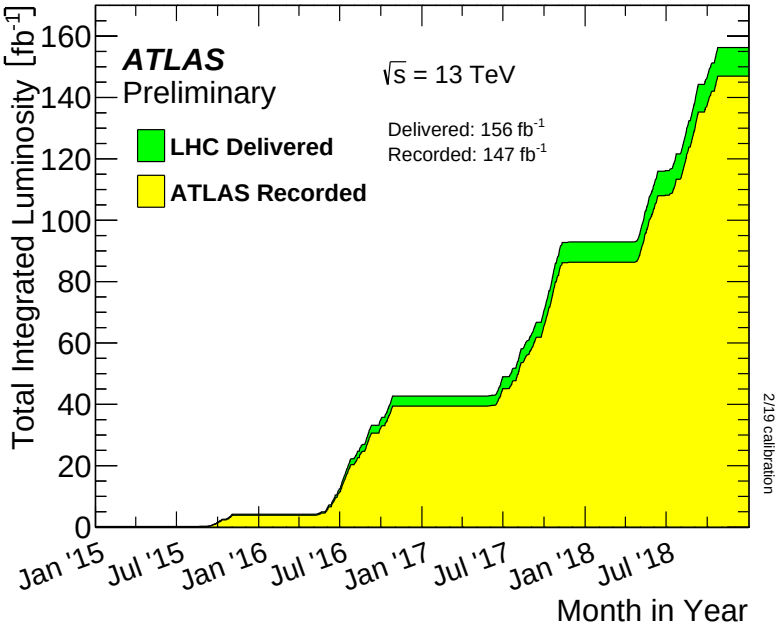


Figure 2.2: Cumulative total luminosity delivered by LHC during Run 2, from 2015 to 2018. The luminosity which was recorded at ATLAS can be seen in yellow [49].

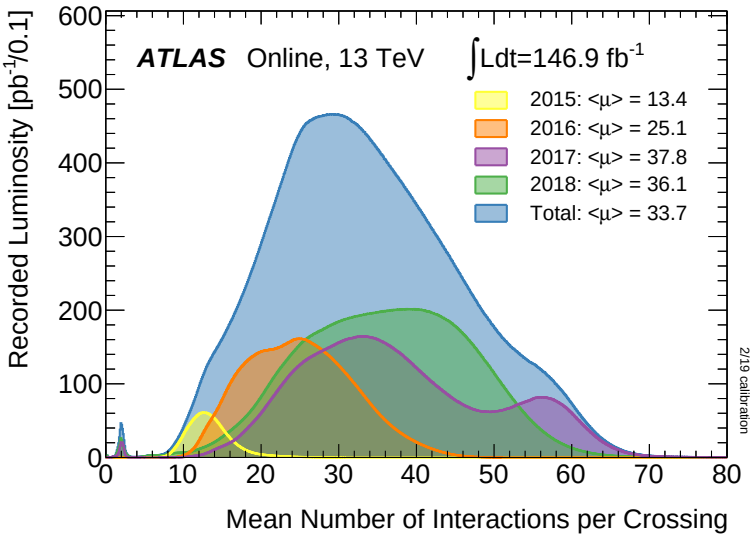


Figure 2.3: Distribution of mean number of interactions per bunch crossing, measured by ATLAS, for each year of Run 2, after luminosity weighting [49].

beyond the SM.

ATLAS comprises multiple sub-detectors, each optimised to measure the properties of different particles which traverse the detector. These sub-detectors are arranged around the beam in a cylindrical manner, forming the barrel region of the ATLAS detector, and in two endcap regions at each end of the barrel (called endcap-A and endcap-C). A system of magnets composed by a central solenoid and three toroid magnets, one in the barrel and two in the endcaps, generates a magnetic field which bends the trajectories of charged particles, allowing their momentum to be measured. The central solenoid surrounds the tracking detector, immersing it in a 2T magnetic field. The three toroid magnets create a magnetic field of up tp 3.5T for the muon detector, improving the measurement of muon momentum. Moving outwards, starting from the beam line, the sub-detector layers of ATLAS are (in the barrel region): the Inner Detector (ID), which measures the position and momentum of charged particles; the Liquid Argon Calorimeter (LAr), an electromagnetic sampling calorimeter which was optimised to provide high resolution energy measurements for electrons and photons; the Tile Hadronic Calorimeter, a sampling calorimeter which provides energy measurements for hadronic jets; and the Muon Spectrometer, which enables the identification of muons and provides an independent momentum measurement for muons beyond the inner detector. A schematic representation of the ATLAS detector and its sub-systems is shown in Fig. 2.4.

In order to fully describe the ATLAS detector, it is important to first define a coordinate system. A right-handed reference frame is used, with origin at the interaction point. The z axis is aligned with the beam direction (positive z is directed towards endcap-A), and the x axis points towards the center of the LHC ring. The y axis completes the right-handed coordinate system, pointing upwards. Fig. 2.5 shows the definitions of these axes. Spherical coordinates (r, θ, ϕ) are defined as follows: the polar angle θ is defined with respect to the z axis; the azimuthal angle ϕ is defined in $x - y$ plane, with respect to the x axis; and r is $\sqrt{x^2 + y^2}$.

The pseudorapidity of a particle is defined as

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right). \tag{2.3}$$

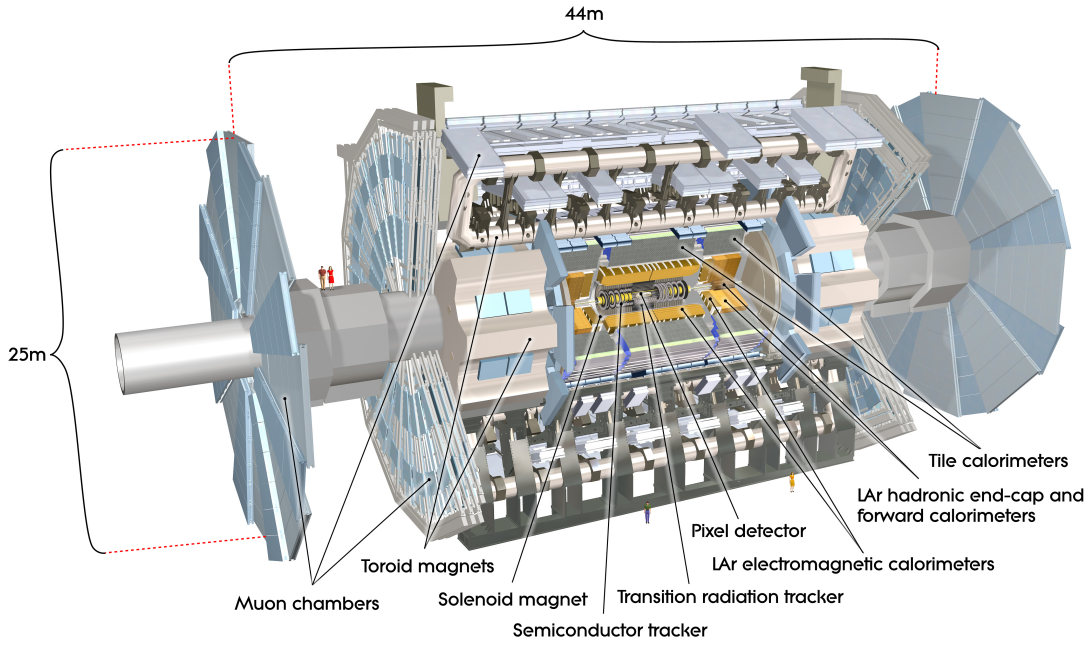


Figure 2.4: Schematic of the ATLAS detector and its sub-detectors [51].

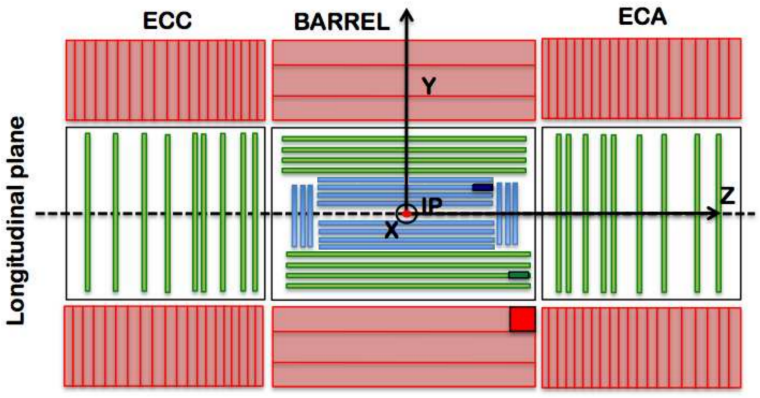


Figure 2.5: Representation of the ATLAS coordinate system, with respect to a longitudinal view of the ATLAS inner detector, described in the following section [52]. ECC and ECA stand for Endcap C and Endcap A, respectively.

Particles with smaller η are more central, while particles with larger η are more forward. A particle traversing the transverse plane has $\eta = 0$. For particles with larger masses, it is more common to use rapidity, defined as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{2.4}$$

in which p_z is the z component of the momentum of the particle and E is its energy. The angular separation between two objects can be defined using the difference between their pseudorapidities and the difference between their azimuthal angles, as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{2.5}$$

2.2.1 Inner Detector

The inner detector (ID) [50] is designed to provide position measurements for charged particles, as well as to allow the reconstruction of primary and secondary vertices. It is surrounded by the superconducting solenoid which generates a 2T magnetic field, which bends the trajectory of charged particles, allowing to measure their momentum from their curvatures. The ID is composed of two silicon detectors, which span the pseudorapidity range of $|\eta| < 2.5$, the pixel detector and the semiconductor tracker (SCT), and one transition radiation tracker (TRT) which spans the range $|\eta| < 2.0$. Each sub-detector is composed by cylindrical layers, which make up the barrel, and cover the range $|\eta| < 1.5$ and end-cap disks, perpendicular to the beam axis, which cover the remaining pseudorapidity range ($1.5 < |\eta| < 2.5$). Fig. 2.6 shows a schematic of a cross-section of the ID detector.

The basic operating principal of the silicon-based detectors is the formation of electron-hole pairs caused by a traversing charged particle [54]. When moving under an electric field, the electron-hole pairs create a signal in the detector. The pixel detector, the silicon sub-detector which is closest to the beam pipe, is segmented into 4 barrel layers and 2 endcaps with three disks each. The insertable B-Layer (IBL) [55] is the innermost barrel layer, being located at a mean radius of about 33.25 mm. The IBL was introduced for Run 2, improving the ATLAS tracking and vertexing performances.

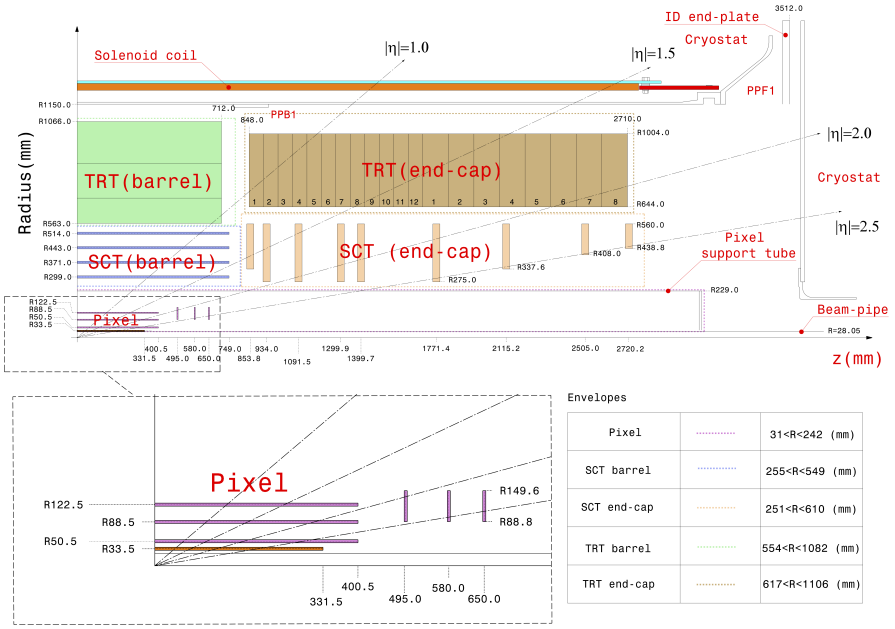


Figure 2.6: r-z cross section of the ATLAS inner detector. The bottom panel shows a zoomed in view of the pixel detector and a summary of the radial positions of the ID sub-detectors [53].

A typical pixel belonging to the IBL has a size of 50 μm $(r-\phi)$ $\times$ 250 μm (z) and a thickness of 200 μm. The intrinsic resolution of the IBL is 10 μm $(r-\phi)$ $\times$ 60 μm (z). The remaining barrel layers are located at mean radii of 50.5, 88.5 and 122.5 mm. A typical pixel belonging to these layers has a size of 50 μm $(r-\phi)$ $\times$ 400 μm (z) and a thickness of 250μm. The expected hit resolution is 10 μm $(r-\phi)$ $\times$ 115 μm (z, for the barrel and r, for the endcaps) [56]. The pixel layer located at 50.5 mm is referred to as the B-Layer. The charge collected by each pixel is measured using the time-over-threshold (TOT) method [57]. The TOT is the length of time a signal generated from a charged particle is above a given threshold and it is proportional to the energy being deposited.

The SCT is a micro-strip silicon detector, which is located in the radial region between 299 mm and 514 mm. It is composed of four barrel layers and 9 disks in each endcap. Each layer consists of two layers of strips with a stereo angle of 40 mrad between them, allowing measurements in both the $r-\phi$ and z directions. Strips in the barrel SCT have an uniform strip pitch of 80 μm. In the end-cap, strips are in the radial direction and have a 161.5 μrad angular pitch. The intrinsic resolution of the micro-strip modules is 17 μm in $r-\phi$ and 580 μm in the z direction for the barrel and the r direction for the

endcap [56].

When a charged particle crosses the boundary between media with different dielectric properties, it emits electromagnetic radiation [54]. The total energy emitted depends on the Lorentz ($\gamma = E/m$) factor of the particle, increasing for larger γ factors. The γ factor of an electron is larger than the γ factor of a pion or a kaon with the same momentum, by about 270 and 970 times, respectively. This fact is exploited by transition radiation detectors to perform electron identification. The ATLAS TRT [58] covers a radial region until 1082 mm and is composed of 350848 gas-filled drift tubes, each with a diameter of 4 mm. The gas mixture filling the tubes is composed mostly of Xe, and a smaller percentage of CO_2 and O_2. Due to gas leaks, some of the TRT modules have been operating with a Argon based gas mixture since 2015, instead of relying on the more expensive Xenon [59]. The Ar based mixture provides a very similar tracking efficiency to the Xe based one, but leads to losses in the electron identification capabilities of the TRT, due to inefficient absorption of transition radiation. A tungsten wire in the center of each tube collects electrons which are created from charged particles passing through and ionising the gas. A polymer material fills the spaces between the tubes generating transition radiation for particles with large γ factor. This additional energy deposition increases the readout signal and provides information for electron/hadron discrimination. There are 3 barrel layers with 32 TRT modules each, and 40 disks in the endcaps. The intrinsic spacial resolution of a TRT module is 130 μm.

2.2.2 LAr Electromagnetic Calorimeter

The ATLAS LAr electromagnetic (EM) calorimeter [50] provides energy measurements for electrons and photons. It has a "destructive" operating principle, designed to stop particles by making them deposit all their energy within the detector. Electromagnetic showers are initiated by electrons and photons when interacting with the detector material, generating multiple secondary particles, with bremsstrahlung and pair production interactions occurring, until the charged particles in the shower are finally stopped through ionisation and photons are absorbed.

The LAr EM calorimeter is a sampling calorimeter, being composed of layers of absorber material (Pb), interspersed with active material (liquid Argon). The detector

modules are disposed in an accordion-like geometry. The Argon is kept in a liquid state by three cryostats which house the LAr calorimeters. Fig. 2.7 displays the ATLAS calorimeter system, including three sections of the electromagnetic calorimeter, the barrel section and two endcap sections. The barrel calorimeter covers $|\eta| < 1.475$, while the two endcap calorimeters cover $1.4 < |\eta| < 3.2$. The LAr EM calorimeter has a large depth, in order to ensure good containment of the EM shower. Each barrel and endcap is divided into three longitudinal segments with different granularities and depth, for $|\eta| < 2.5$: the first one has the finest granularity (0.003×0.1 in the $\Delta\eta \times \phi$ direction, for the barrel), allowing a precise determination of the position of particles, and a thickness ranging between 2 and 5 radiation lengths (X_0)[1]; the second has a granularity of 0.025×0.025 in $\Delta\eta \times \phi$ and a thickness which ranges between 17 and 20 X_0, being responsible for the greatest energy loss for traversing particles; and the third has a granularity of 0.05×0.025 in $\Delta\eta \times \phi$ and a thickness ranging from 2 to 10 X_0. For $2.5 < |\eta| < 3.2$, the LAr calorimeters are divided in only two longitudinal segments. In front of the active calorimeter material, in both barrel and encap regions, there is a thin presampler (PS), composed of liquid argon, used to correct for photon and electron energy losses before traversing the LAr calorimeter. The PS covers $|\eta| < 1.8$, has a thickness of 11 mm in the barrel and 5 mm in the endcaps and has a granularity of 0.025×0.1 in $\Delta\eta \times \phi$.

There is a gap between the barrel and endcaps, for $1.37 < |\eta| < 1.52$, commonly referred to as transition region, which has cables and services for the inner detector and the LAr calorimeters. This amounts to a lot of material, with a thickness which ranges from $5X_0$ to $10X_0$. In order to correct for energy losses in this material, the so-called "E4" scintillators placed in this gap are instrumented [60].

2.2.3 Hadronic Calorimeters

The ATLAS hadronic (HAD) calorimeter [50] is designed to measure the energy of hadrons, which materialise as jets in the calorimeter. In this case, hadronic showers are generated, with multiple interaction processes involved, such as nuclear interactions and ionisations. A fraction of the shower develops electromagnetically, due to electromagnetic

[1] X_0 is the length an electron travels in a material until it only has an energy corresponding to $1/e$ of its initial energy. The mean free path of a photon before undergoing pair conversion is $9/7X_0$.

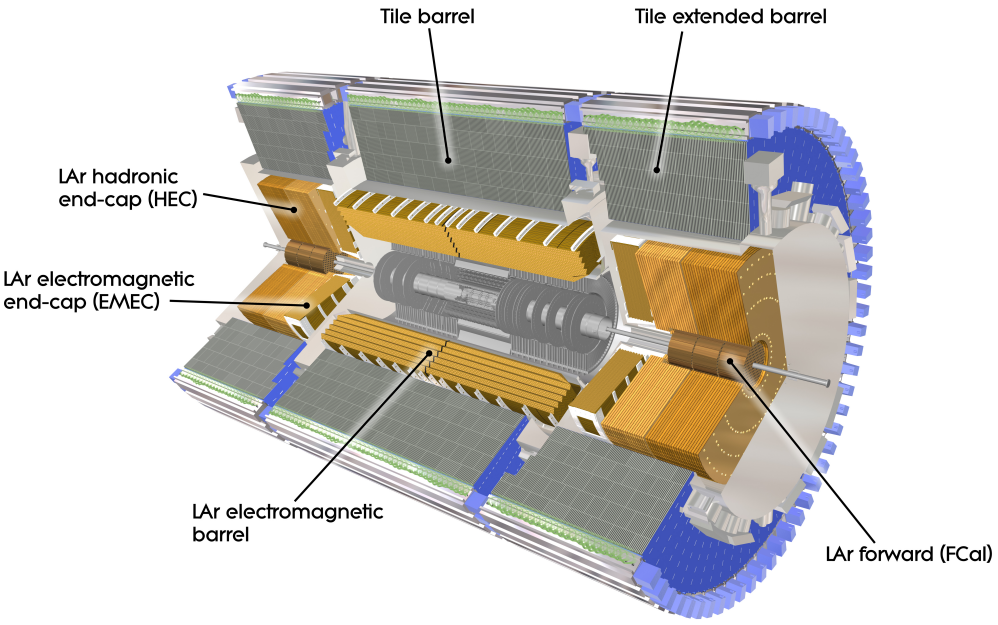


Figure 2.7: Schematic of the ATLAS calorimeters systems, showing the Liquid Argon and Tile Calorimeters [61].

decays of neutral hadrons (for instance $\pi^0 \to \gamma\gamma$ and $\eta^0 \to \gamma\gamma$).

The HAD calorimeter is composed by three Tile calorimeter sections in the barrel region of the ATLAS detector (one Tile barrel and two Tile extended barrel sections); two LAr hadronic calorimeters in each endcap; and two LAr forward calorimeters, also in the endcaps. These sub-sections can be seen in Fig. 2.7.

The Tile calorimeter is a sampling calorimeter with layers of iron alternating with plastic scintillator material. These layers are disposed in planes perpendicular the direction of the beam. The Tile calorimeter is sub-divided into three longitudinal segments, with a total depth of 9.7 interaction lengths (λ)[2]. The three sections of the Tile calorimeter cover $|\eta| < 1.7$.

The LAr hadronic calorimeters at each endcap cover the pseudorapidity region of $1.5 < |\eta| < 3.2$. The absorber material for these segments is copper. In the case of the LAr forward calorimeters, the passive medium is composed of tungsten. These forward sections of the HAD calorimeter further extend the total acceptance to $|\eta| < 4.9$.

[2] λ is the mean free path a hadron travels before an inelastic nuclear interaction occurs.

2.2.4 Muon Spectrometer

Due to the depth of the ATLAS calorimeters, the only particles which reach the outermost layer of the ATLAS detector are muons and neutrinos[3]. The muon spectrometer (MS) [50] comprises a set of gaseous detectors, and was designed to measure the momentum of muons. These measurements are enabled by the toroidal magnets which generate a magnetic field, bending the trajectories of the charged muons traversing the detector volume.

The MS is divided into a barrel section and two end-cap sections, all divided in three layers. It is composed of chambers which provide precision tracking measurements, covering the pseudorapidity range of $|\eta| < 2.7$. Most of these chambers are Monitored Drift Tube Chambers (MDTs), except in the innermost precision tracking layer, for $2 < |\eta| < 2.7$, where Cathode-Strip Chambers (CSCs) are in place. The MS also has trigger capabilities. Two types of trigger chambers are employed: Resistive Plate Chambers (RPCs), in the barrel, and Thin Gap Chambers (TGCs), in the endcaps. Fig. 2.8 displays the different sections of the MS.

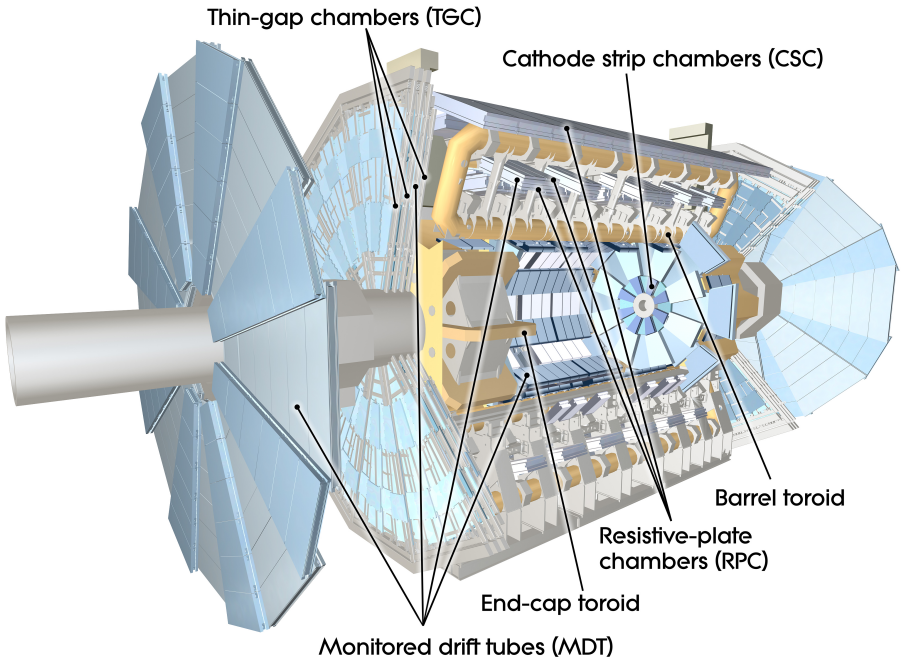


Figure 2.8: Schematic of the ATLAS Muon Spectrometer [62].

[3]However, neutrinos do not interact with the detector material.

2.2.5 Trigger system

An essential component of data acquisition in collider experiments is triggering. The ATLAS trigger system [63] performs real time decisions on which pp collision events to record, selecting events with experimental signatures of interest. The trigger system is crucial during data taking, since the current data transfer rate and data storage limitations make it impossible to record all the hard scatter events produced at the LHC.

Events are selected in two trigger states: Level-1 Trigger and High-Level Trigger. The Level-1 Trigger (L1) is fully hardware based and uses coarse granularity information from the calorimeter and muon systems. Signals from the calorimeters are processed by the L1Calo trigger system, while TGC and RPC hits are processed by the L1 Muon trigger system. These two systems provide inputs to the Central Trigger Processor (CTP), responsible for forming the L1 trigger decision. The maximum L1 accepted event rate is 100 kHz. The L1 trigger seeds the second trigger stage, the High-Level Trigger (HLT), providing it with Regions-of-Interest (ROIs) in ϕ and η.

The HLT is a software based trigger which runs dedicated selection algorithms, taking into account the complete, fine granularity information provided by the detector. These algorithms perform physics object reconstruction, very similar to the offline reconstruction, and apply sets of requirements on the properties of these objects, in order to target different signatures. A series of dedicated HLT algorithms, seeded by a L1 trigger item is known as a trigger chain. Multiple trigger chains are used during ATLAS data taking, selecting events for different physics analysis. On average, the HLT output rate during Run 2 was 1.2 kHz.

2.3 Physics objects reconstruction

2.3.1 Tracks

Charged particles traversing the inner detector interact with its material, in the way described in Section 2.2.1. From the energy deposits (hits) created by these particles, their trajectories can be reconstructed, and track objects are formed, with a set of associated measured parameters. Examples of charged particles which typically do

not decay before arriving to the inner detector and therefore create tracks in the ID are charged hadrons such as pions, kaons and protons. It is important to note that the ATLAS experiment has limited particle identification (PID) capabilities for charged high momentum hadrons. Therefore, pions, kaons and protons cannot be usually distinguished[4].

ATLAS track reconstruction is performed by first building track candidates from silicon hits and extending them into the TRT (primary tracking), and then performing the reconstruction in reverse, starting the pattern recognition in the TRT and extending it into the silicon detectors (back-tracking) [65].

Primary tracking in the silicon detectors [2] starts with clusterization, which is the process of grouping energy deposits associated with adjacent pixels and strips into clusters. These clusters are then converted into three-dimensional space-points, point measurements of the charged particles trajectories. Groups of 3 space-points form track seeds. Track seeds which pass a set of impact parameter and momentum requirements are extended into track candidates using a combinatorial Kalman filter [66], which finds other pixel and SCT clusters which are compatible with the formed trajectory. At this point, the same clusters might have been used in the reconstruction of different track candidates. In order to address this and to reject track candidates built from random combinations of clusters, track candidates are put through an ambiguity solving stage. Each track candidate is assigned a track score, which increases with the track momentum, the amount of clusters belonging to the track, and the χ^2 of the track fit, and decreases with the amount of holes in the tracks (points of the track trajectory which cross a detector sensor but did not leave a hit). Tracks can only contain two shared clusters and only two tracks can share the same cluster. Tracks with a higher score are favoured by the ambiguity solver when resolving track overlaps. When a track shares a cluster with two other higher ranked tracks, the cluster in question is removed from the track candidate, which is re-scored and re-processed. After processing, track candidates are retained by the ambiguity solver if they have: $p_{\mathrm{T}} > 400, \mathrm{MeV}$; $|\eta| < 2.5$; at least 7 clusters (pixel and SCT); no more than one shared pixel cluster or two shared clusters on both sides of a SCT layer; no more than 2 holes in both silicon detectors and no

[4]For low momentum charged hadrons $\mathrm{d}E/\mathrm{d}x$ provides some discrimination [64].

more than one hole in the pixel; $|d_0^{\mathrm{BL}}| < 2$ mm, in which d_0^{BL} is the transverse impact parameter with respect to the beam-line; and $|z_0^{\mathrm{BL}} \sin\theta| < 3$ mm, in which z_0^{BL} is the longitudinal impact parameter with respect to the beam-line.

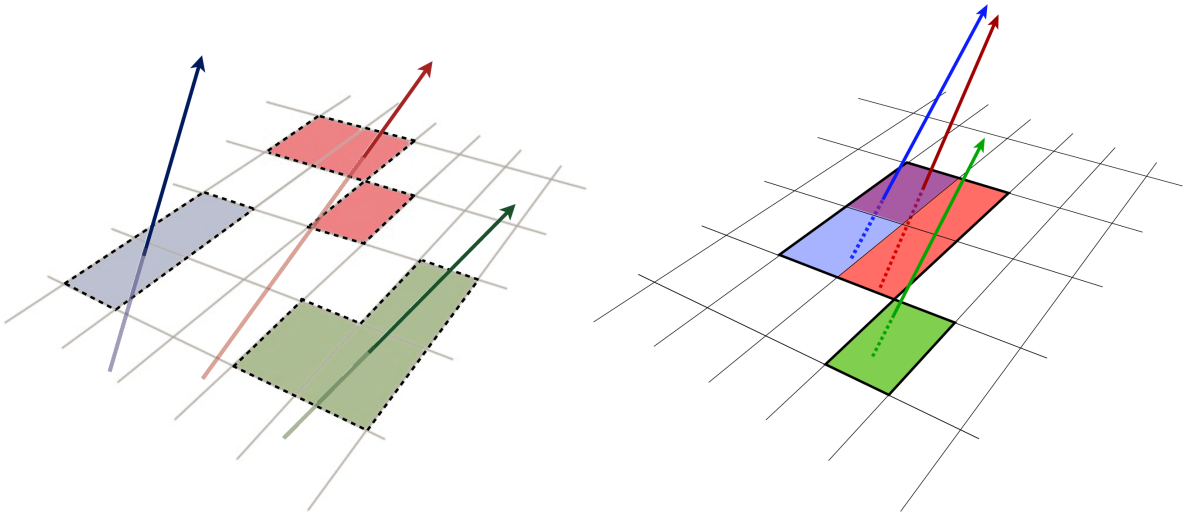
The ambiguity solver algorithm limits the number of clusters shared by multiple tracks, since they can indicate an incorrect cluster-to-track association. However, in dense charged particle environments, it becomes more likely that multiple charged particles contribute to the same cluster, since particles are highly collimated. Clusters created by the energy deposition of multiple particles are referred to as merged clusters. In order to minimise the loss of efficiency due to limitations on the number of shared clusters imposed by the ambiguity solver, a neural network (NN) is deployed to identify merged clusters [67, 68]. The inputs to this NN are: the charges collected in the pixels contained in the cluster; the sizes of the pixels[5]; which layer of the pixel detector the cluster belongs to, and if it is part of the barrel or endcap; and the ϕ and θ of the incident track candidate with respect to the sensor. The NN identifies merged clusters created by 2 (3) charged particles with 90% (85%) efficiency [2]. Merged clusters do not penalise tracks in the ambiguity solving process and there are no requirements on the number of merged clusters during reconstruction [6]. Fig. 2.9 is an illustration of single particle clusters and a merged cluster created by highly collimated charged particles.

The track candidates that are retained following the ambiguity solving stage are re-fitted through a χ^2 method, estimates for the track parameters are calculated and tracks are added to the track collection. Subsequently a TRT extension is attempted using a Kalman filter. Tracks which were successfully extended are once again fitted, improving the parameters estimates. The outside-in reconstruction is then performed extending tracks from TRT measurements. This process increases the tracking efficiency, by for instance, recovering tracks which emerge from the decays of displaced particles or from photon conversions, and as such do not have enough silicon hits to survive the selection applied by the primary tracking algorithm.

Tracks are described by five parameters, with two different parametrisations being

[5] In Section 2.2.1, the typical pixel sizes are detailed. Nevertheless, pixels between the readout chips are longer (600 μm in the longitudinal direction). Inputting pixel size to the NN allows the labelling of the longer pixels during the training.

[6] A cluster associated with multiple charged particles is exclusively classified as merged or shared.



(a) (b)

Figure 2.9: Representation of clusters created by single charged particles (a) and a merged cluster with energy deposition from collimated charged particles (b) [2].

used, global and local [69]. In the global parametrisation the parameters are defined with respect to the perigee, the point of closest approach of the track to the beam-line. Tracks are described by $(d_0, z_0, \phi, \theta, q/p)$, where d_0 and z_0 are the transverse and longitudinal impact parameters, respectively; ϕ and θ are the azimuthal and polar directions; and q/p is the ratio between the particle's charge and its momentum. In the local parametrisation, tracks are described by $(l_x, l_y, \phi, \theta, q/p)$, where ϕ, θ and q/p are the same as in the global frame, while l_x and l_y are defined with respect to a local frame associated to a specific detector module. Fig. 2.10 shows a schematic of these parametrisations, and corresponding parameter definitions. Local ϕ and θ measurements can also be accessed for each cluster belonging to tracks. These angles are once again defined with respect to a specific sensor.

Two selection working points are defined for tracks originating from charged hadrons (different specific working points are derived for electron and muon selection): `Loose` and `TightPrimary` [71]. These working points were optimised for primary tracks. Tracks are categorised as `Loose` if they have:

- $p_{\mathrm{T}} > 500$ MeV;
- $|\eta| < 2.5$;

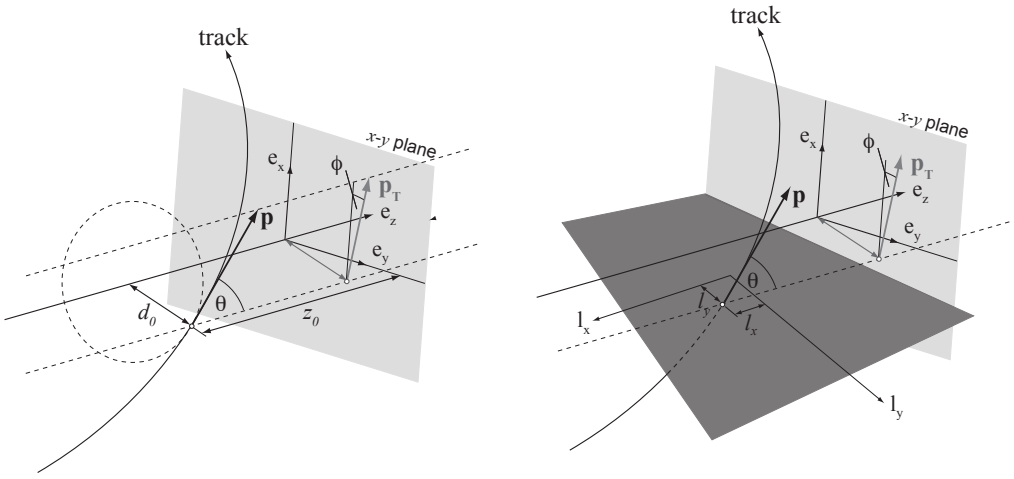


Figure 2.10: Illustration of the global (left) and local (right) track parametrisations [70]. The perigee is the point of closest approach of the track to the beam-line.

- at least 7 silicon hits ($N_{Sil}^{hits} \geq 7$, including pixel and SCT);
- no more than one shared cluster ($N_{mod}^{shared} \leq 1$);
- no more than two holes in both pixel and SCT layers ($N_{Sil}^{hole} \leq 2$);
- no more than one pixel hole ($N_{Pix}^{hole} \leq 1$).

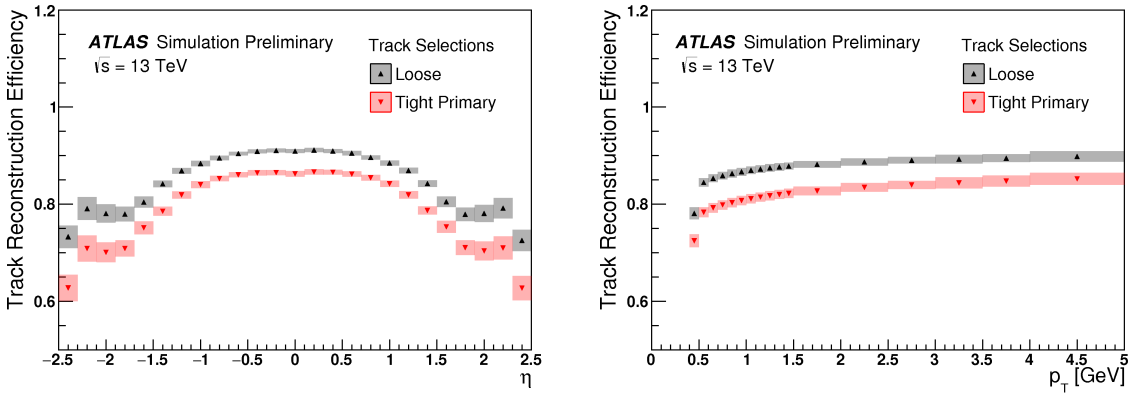
Tracks survive the `TightPrimary` selection if, on top of fulfilling the Loose requirements they have:

- at least 9 silicon hits ($N_{Sil}^{hits} \geq 9$), if $|\eta| \leq 1.65$;
- at least 11 silicon hits ($N_{Sil}^{hits} \geq 11$), if $|\eta| \geq 1.65$;
- no pixel holes ($N_{Pix}^{hole} = 0$);
- at least one hit in the IBL or B-Layer ($N_{IBL}^{hits} + N_{B\text{-}layer}^{hits} > 0$).

The Run 2 track reconstruction efficiencies associated to each working point, estimated using simulated events can be found as a function of track η and p_T in Fig. 2.11.

2.3.2 Photons

As explained in Section 2.2.2, when photons traverse the LAr EM calorimeter, they interact with the Lead absorber material, initiating electromagnetic showers. These



(a) (b)

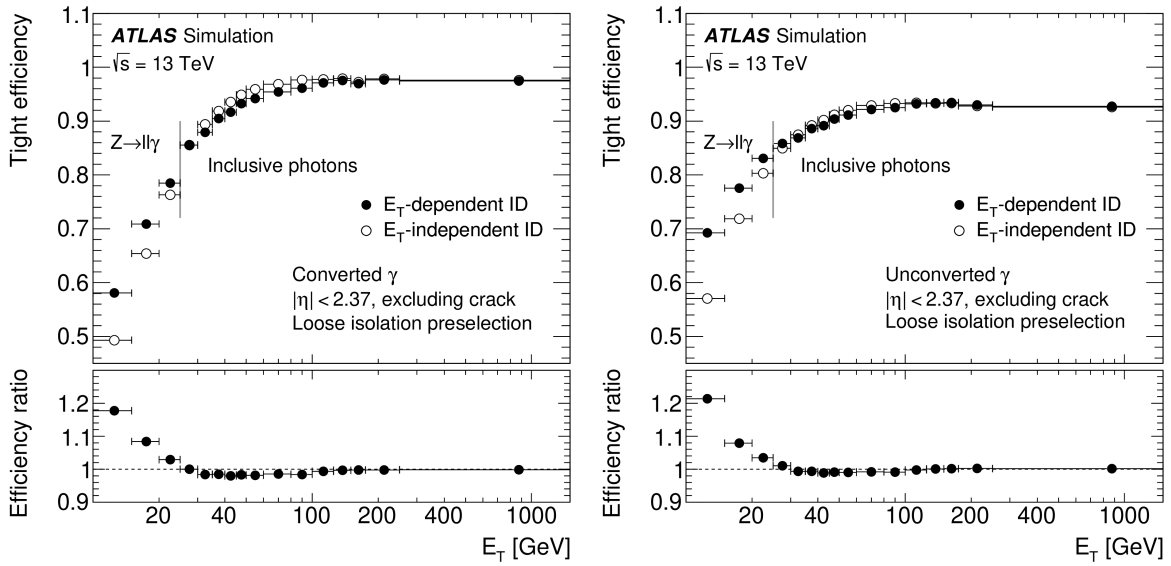
Figure 2.11: Run 2 track reconstruction efficiency as a function of η, for inclusive p_{T} (a) and in bins of p_{T}, for inclusive η (b) [71], for `Loose` and `Tight` tracks.

EM showers generate signals in the cells of the calorimeter, which are grouped into clusters in the reconstruction algorithm. However, there is a probability that photons traversing the ATLAS detector interact with the inner detector material, converting into electron-positron pairs, before reaching the LAr calorimeter. Therefore, photons can be reconstructed either as unconverted or converted. In order to reconstruct photons which undergo conversion in the ID, EM clusters are matched to ID tracks which are consistent with photon conversion vertices. A converted photon can be matched to two opposite charged tracks or just one track. Tracks originating from electrons produced in photon conversions are more likely to have holes in the first layers of the Pixel detector. Furthermore, they are more likely to generate high-threshold signals in the TRT (due to more transition radiation being produced). Selection requirements devised according to these factors are applied to converted photon track candidates, in order to ensure high purity. A full description of photon reconstruction in ATLAS during Run 2 can be found in Ref. [60].

Energy measurements for photons (and electrons) obtained using the LAr EM calorimeter are calibrated using data-MC comparisons for $Z \to e^{-}e^{+}$ events. This procedure is validated through a data-driven test performed with radiative Z boson decays, and an uncertainty on the photon energy scale correction is derived. A detailed description of these procedures is provided in Ref. [60].

Three photon identification working points are commonly used by ATLAS analyses,

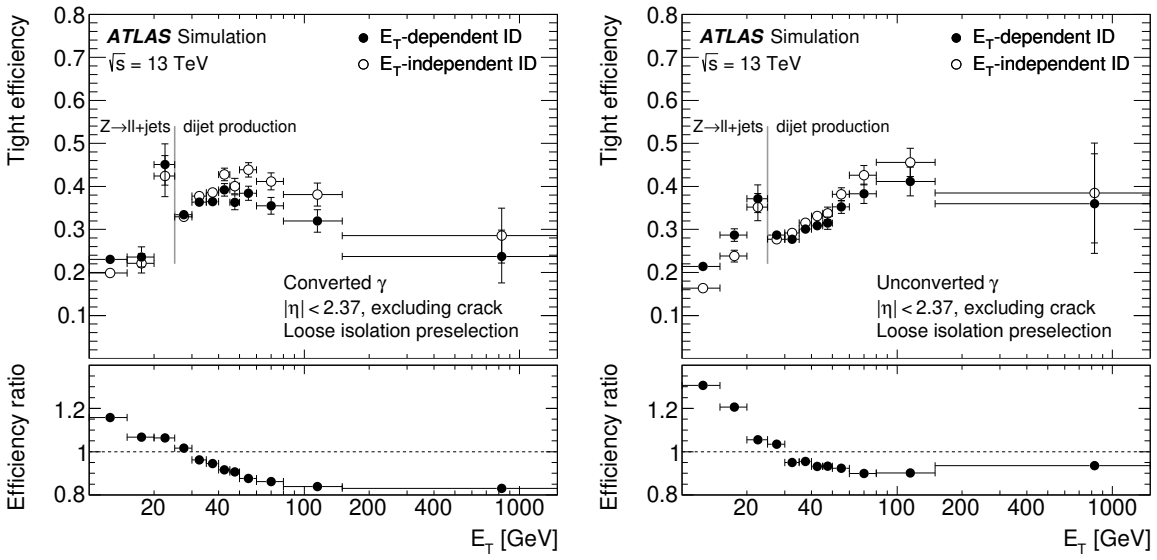
Loose, Medium and Tight, with different sets of criteria required. The application of these three working points rejects hadronic jets faking photons, each with a different level of associated background rejection and signal efficiency. The Tight working point (used in the work presented in this thesis) is optimised using a multivariate analysis technique, relying on shower shape variables. The optimisation is performed in bins of $|\eta|$ and E_{T}. The discriminating variables used are listed in Ref. [60]. The signal and hadronic jet background efficiencies associated with the Tight identification requirement for unconverted and converted photons, as a function of $E_{\mathrm{T}}(\gamma)$, are shown in Fig. 2.12 and Fig. 2.13, respectively. These plots compare the efficiencies of the E_{T} dependent optimisation, with an earlier version which was performed independently of E_{T}. The efficiencies were obtained using simulation: for low E_{T}, $Z \to ll\gamma$ and Z + jets MC samples were used; for high E_{T}, inclusive photon and inclusive dijet production samples were employed.



(a) (b)

Figure 2.12: Signal efficiencies of the Tight photon identification working point as a function of E_{T}, for converted (a) and unconverted photons (b) [60].These efficiencies are calculated with respect to the loose identification and loose isolation pre-selection.

Requirements can also be applied on variables associated with the amount of detected activity surrounding the selected photon, commonly referred to as isolation variables. The isolation is quantified either taking into account EM calorimeter clusters



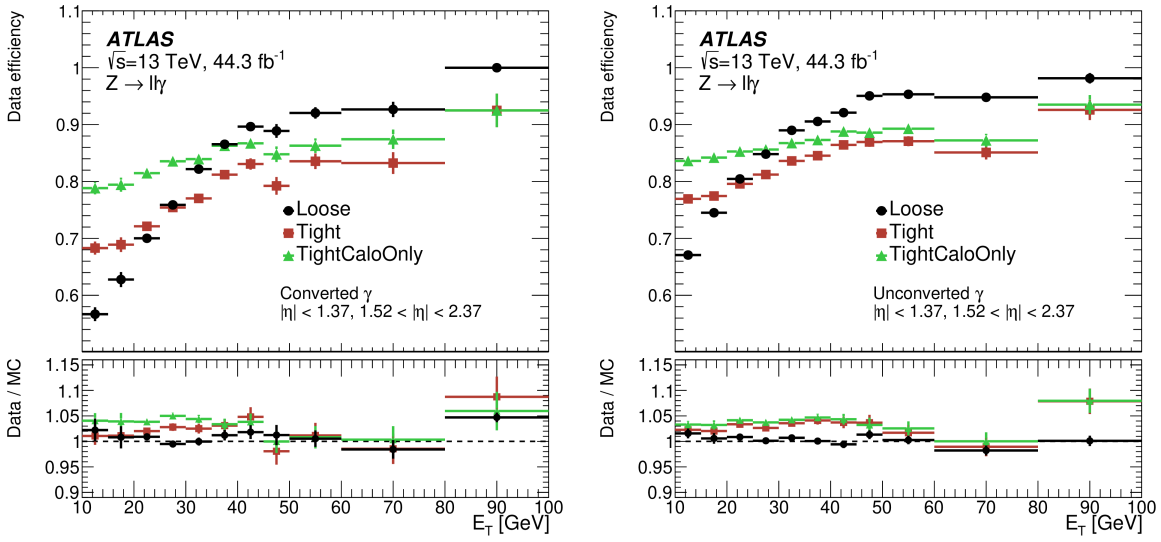
(a) (b)

Figure 2.13: Jet background efficiencies of the `Tight` photon identification working point as a function of E_{T}, for converted (a) and unconverted photons (b) [60]. These efficiencies are calculated with respect to the loose identification and loose isolation pre-selection.

(photon calorimeter isolation) or ID tracks (photon track isolation). For calorimeter isolation, the $E_{\mathrm{T}}^{\mathrm{coneXX}}$ variable is defined, corresponding to the sum of transverse energy associated with clusters contained within a $\Delta R < \mathrm{XX}/100$ cone, centred around the photon cluster[7]. In the case of track isolation, the $p_{\mathrm{T}}^{\mathrm{coneXX}}$ variable is used instead, corresponding to the sum of transverse momentum associated with tracks which are within a cone of $\Delta R < \mathrm{XX}/100$ around the photon direction. The tracks taken into account in the calculation are required to meet a set of loose quality requirements. Three photon isolation working points are pre-defined and applied frequently in ATLAS analyses: `Loose`, `Tight` (also known as `FixedCutTight`) and `TightCaloOnly`. The signal efficiencies of each of these three working points are shown in Fig. 2.14, measured using simulated $Z \to ll\gamma$ events. In the analysis presented in this thesis, the `FixedCutTight` working point is used, defined by the following two requirements:

- $E_{\mathrm{T}}^{\mathrm{cone40}} < 0.022 \times E_{\mathrm{T}} + 2.45$ GeV (photon calorimeter isolation);
- $p_{\mathrm{T}}^{\mathrm{cone20}} / p_{\mathrm{T}} < 0.05$ (photon track isolation).

[7]The E_{T} corresponding to the photon is subtracted and corrections accounting for leakage and pile-up are applied.



(a) (b)

Figure 2.14: Efficiencies of the different photon isolation working points as a function of E_{T}, for converted (a) and unconverted photons (b) [60]. The efficiencies are calculated with respect to a sample of tightly identified photons from simulated $Z \to ll\gamma$ events.

2.3.3 Electrons

Electron reconstruction proceeds in a very similar way to photon reconstruction. In this case, the electron object corresponds to an EM calorimeter cluster matched to an ID track. The electron reconstruction efficiency as a function of E_{T} can be found in Fig. 2.15. A detailed description of electron reconstruction during Run 2 can be found in Ref. [60].

As is the case for photons, shower shape variables have good discriminating power for electron identification. Furthermore, as was already explained in Section 2.2.1, TRT based variables allow to discriminate against tracks created by hadrons. Two variables in particular are employed in the main analysis presented in this thesis, and as such are here defined:

- Rhad, the ratio of E_{T} in the hadronic calorimeter to E_{T} in the electromagnetic cluster, which quantifies the electron's hadronic leakage;

- eProbabilityHT, a likelihood discriminant employed in electron versus charged-hadron discrimination, based on TRT transition radiation.

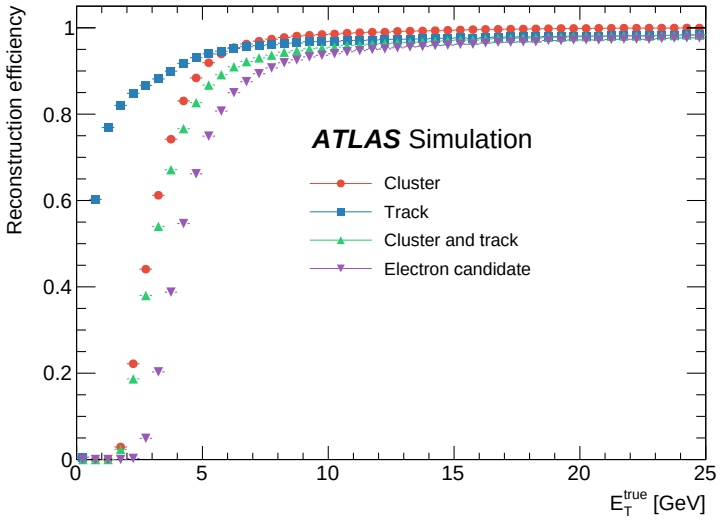


Figure 2.15: Electron candidate reconstruction efficiencies, as a function of electron E_T, estimated using simulation. The individual reconstruction efficiencies of the track and cluster, as well as the combined efficiency of reconstructing both the track and cluster, are also shown [60].

Over the range $|\eta| < 0.8$ and $|\eta| > 1.37$, only the first layer of the hadronic calorimeter is taken into account in the calculation of Rhad.

2.3.4 Jets

Jets are sprays of collimated, energetic hadrons. They can be reconstructed from calorimeter clusters, using the anti-k_t algorithm [72]. The size of the reconstructed jets is characterised by a radius parameter R. A parameter of $R = 0.4$ is typically used in the ATLAS jet reconstruction.

Background jets of non-collision origin can be mis-identified as jets originating from pp collisions. These can be generated by cosmic rays, calorimeter noise, or by muons produced in interactions between protons from beam losses and residual gas or other material away from the IP. A set of selection requirements is applied to jets in order to reject those which are not produced in the hard scatter event. This procedure is referred to as jet cleaning [73]. The jet selection efficiency for two jet cleaning working points, `Loose` (or `BadLoose`), and `Tight` (or `BadTight`), as a function of jet η, can be found in Fig. 2.16. The looser working point is typically used in ATLAS analyses, while the tighter one was devised for use in analyses which have significant contributions

from backgrounds of non-collision origin. Besides non-collision jets, it is important to have discriminating power against jets created in pile-up interactions. For this end, a jet-vertex tagger (JVT) discriminant is employed [74] by ATLAS analyses.

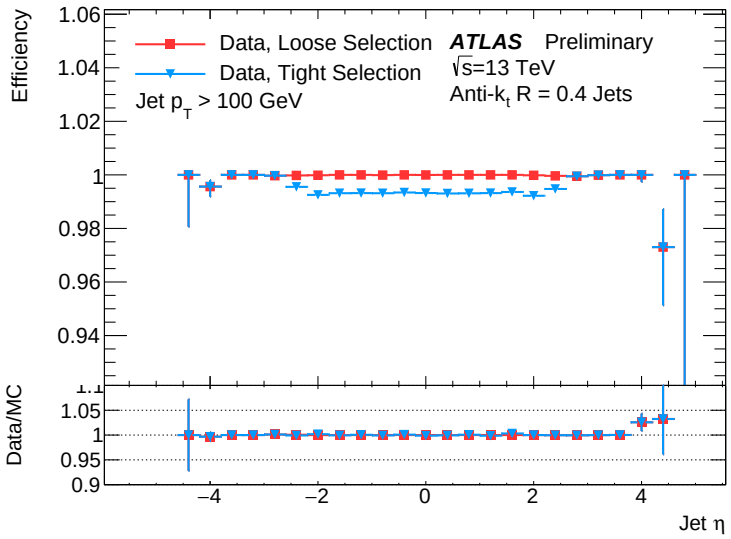


Figure 2.16: Jet quality selection efficiency for the `Loose` and `Tight` criteria, as a function of jet η [73].

2.3.5 Hadronic taus

Typically, tau leptons decay before reaching the Inner Detector. As such, they are reconstructed from their decay products. Approximately 65% of tau lepton decays are hadronic ($\tau \rightarrow \nu_\tau + \text{hadrons}$), and the majority of these decays include one or three charged hadrons and up to two neutral pions. In reconstruction, hadronic taus are split into categories according to the tau decay mode they are the most consistent with, and the number of charged hadrons in the decay (the number of "prongs"). The neutrino in the decays is not reconstructed. Only the visible components are included in the hadronic tau object which, for this reason, is referred to as $\tau_{\text{had-vis}}$.[8]

Hadronic tau reconstruction algorithms [75] are seeded by jet objects with $p_{\text{T}} > 10$ GeV and $|\eta| < 2.5$. ID tracks within $\Delta R < 0.2$ of the jet seed direction (called the core region) are matched to the $\tau_{\text{had-vis}}$ candidate if they meet minimum requirements on kinematic, quality and other track variables. A $\tau_{\text{had-vis}}$ with n associated core

[8] In the case of leptonically decaying tau leptons, the electron/muon decay products are reconstructed.

tracks is known as an n-prong tau. A Particle Flow (PF) algorithm [76] is used in the reconstruction of the charged and neutral hadrons associated with the hadronic tau, allowing the classification of taus into different decay modes. Charged hadrons are reconstructed using tracking ID measurements. The neutral pions are instead reconstructed through their decays into collimated photons ($\pi^0 \to \gamma\gamma$), which create EM showers in the LAr calorimeter. After reconstructing an EM cluster in the core of the $\tau_{\text{had-vis}}$, a correction is applied to the energy of the π^0 candidate to account for EM deposits produced by the charged hadrons in the decay. To ensure high purity for the π^0 candidates a boosted decision tree (BDT) trained on shower shape variables is employed, ultimately improving the efficiency of the hadronic tau mode classification. The performance of this BDT discriminator was evaluated using simulated $Z \to \tau\tau$ events and it quantified in Fig. 2.17, which shows the identification efficiency of signal π^0 versus the corresponding background rejection (obtained for different requirements on the BDT output score).

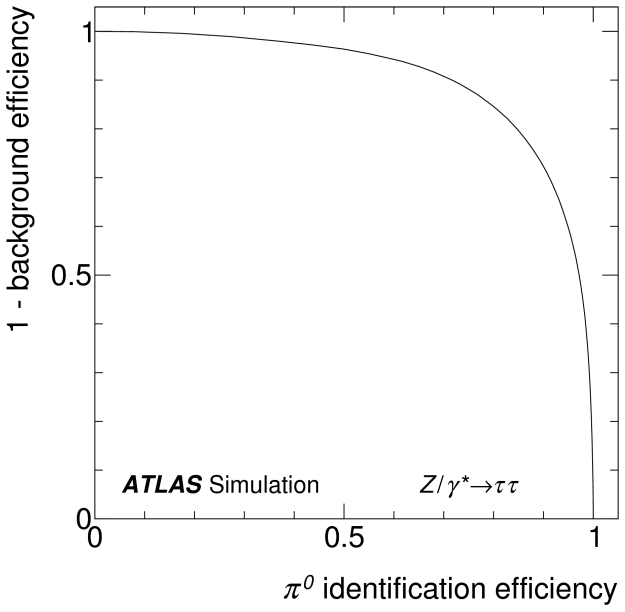


Figure 2.17: Identification efficiency of signal π^0 versus the corresponding background rejection [76]. Simulated $Z \to \tau\tau$ events are used in the study. All π^0 candidates which are not matched to a generated π^0 are considered background.

Different multivariate techniques are used in the identification of hadronic taus. A Recurrent Neural Network (RNN) [77] is used to discriminate against quark and gluon

initiated jets, while a Boosted Decision Tree (BDT) is trained to discriminate against electrons [78]. In both cases, different working points are provided, corresponding to different requirements on the output score of either the RNN or the BDT. The performance of the RNN discriminator is quantified in Fig. 2.18, which shows the jet background rejection as a function of the signal $\tau_{\text{had-vis}}$ efficiency, for different provided working points, `Tight`, `Medium`, `Loose` and `VeryLoose`, for 1 and 3-prong taus.

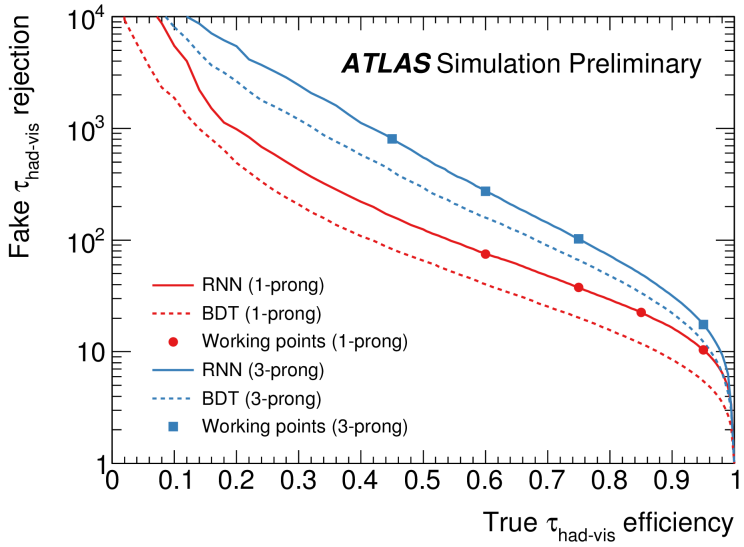


Figure 2.18: Rejection power of fake $\tau_{\text{had-vis}}$ as a function of true $\tau_{\text{had-vis}}$ efficiency, for different requirements on the RNN score [77]. The curves are shown for separately for 1 and 3-prong taus. The four working points, `Tight`, `Medium`, `Loose` and `VeryLoose`, are indicated by the square and circle markers. The RNN performance is compared to the performance of a BDT discriminator, used previously by ATLAS analyses in the identification of hadronic taus.

Chapter 3

Measurement of track reconstruction efficiency in dense enviroments

With the increase of the centre-of-mass energy of pp collisions at the LHC, events including highly energetic jets or decays of highly boosted particles (for instance τ leptons or b-hadrons) into collimated charged particles occur at higher rates. In these dense charged particle environments, the average separation between collimated charged particles is of the order of the granularity of the ATLAS pixel detector (detailed in Section 2.2.1). This is demonstrated in Fig. 3.1, from Ref. [2], which shows the average minimum separation between charged particles produced in different decay processes, at the IBL layer, as a function of the p_{T} of the mother particle. This study was performed using simulations of single particles decaying into final states with different charged particle multiplicities: $\rho^0 \to \pi^+\pi^-$; $\tau \to \pi^+\pi^-\pi^\pm\nu_\tau$; τ-lepton decays into a 5-prong final state; and B^0 meson decays into all possible final states. The collimation between charged particles increases as a function of the p_{T} of the initial particle. These types of signatures are targeted by multiple ATLAS analyses, as is the case for exclusive decays of the Higgs and Z bosons into a highly boosted meson and a photon, for instance $H/Z \to \rho(\pi^+\pi^-)\gamma$ and $H/Z \to \phi(K^+K^-)\gamma$ [3].

As was explained in Section 2.3.1, highly collimated charged particles which cannot be resolved by the pixel detector create merged clusters, and are more likely to fail reconstruction. Therefore, it is important to optimise and fully characterise the track reconstruction performance in dense environments, in both data and MC simulation. During reconstruction, an NN is employed to identify merged clusters, improving the

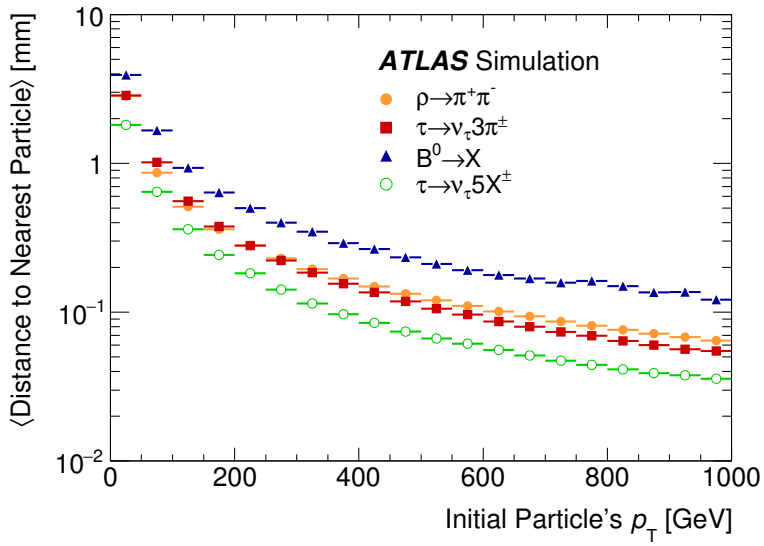


Figure 3.1: Average minimum distance at the IBL, between charged particles created in different decay processes, as a function of the transverse momentum of the initial particle [2]. This study was performed using different single particle simulations.

track reconstruction efficiency. Nonetheless, a residual efficiency loss associated with high charged particle density and collimation still remains. A data-driven method was developed to estimate the fraction of tracks lost in reconstruction (F_{lost}) within high p_T jets. Ref. [2] details this methodology and presents the measured F_{lost} for early run 2 data, collected in 2015. This chapter presents a review of this pre-established methodology and discusses improvements introduced by the author of this thesis. Results obtained using the modified method applied to 2018 data are also presented.

3.1 F_{lost} methodology

Charged particles which cross the ATLAS pixel detector deposit energy in the detector via ionisation, creating electron-hole pairs, which drift under the applied electric field, and are collected by the electrodes of the pixel modules. As explained in Section 2.2.1, this charge is measured through the TOT method for each pixel, if it is above the set threshold. The charge collected in the clusters associated to reconstructed tracks is proportional to the ionisation energy loss of the charged particles (dE/dx). The average energy loss per path length of a charged particle in a medium is described by the Bethe-Bloch formula [15], and depends on the properties of the medium (such as density,

atomic number, mean excitation energy), and the properties of the particle (such as its momentum and charge). The $\mathrm{d}E/\mathrm{d}x$ of different charged particles in silicon, as a function of the particles momentum, can be found in Fig. 3.2.

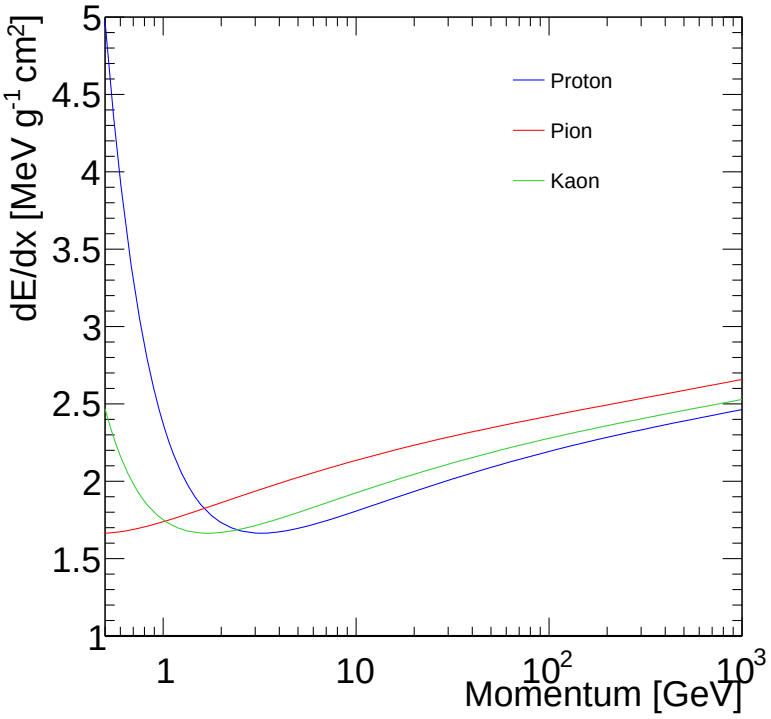


Figure 3.2: Mean ionisation energy loss of per path length for different charged particles in silicon, as a function of their momentum [15]. The $\mathrm{d}E/\mathrm{d}x$ is normalised to the silicon density.

The data-driven method presented in Ref. [2], which allows to estimate the fraction of tracks lost in reconstruction due to the collimation of charged particles, relies on cluster $\mathrm{d}E/\mathrm{d}x$. The $\mathrm{d}E/\mathrm{d}x$ of single charged particles traversing a thin layer is expected to follow a Landau distribution. Two charged particles will, on average, deposit more energy via ionisation than a single charged particle. In fact, the $\mathrm{d}E/\mathrm{d}x$ associated with an n-particle cluster is, on average, n times larger than the $\mathrm{d}E/\mathrm{d}x$ associated to a single-particle cluster (as it will be shown in this chapter). Using this simple principle, it is possible to statistically separate single particle from multiple particle clusters, using cluster $\mathrm{d}E/\mathrm{d}x$ as the discriminating variable. Every time a cluster has a $\mathrm{d}E/\mathrm{d}x$ compatible with n particles but only one track was reconstructed in associated with it, there were $n-1$ lost tracks.

The $\mathrm{d}E/\mathrm{d}x$ associated with a cluster can be calculated from the generated charge,

using the expression

$$\mathrm{d}E/\mathrm{d}x[\mathrm{MeVg}^{-1}\mathrm{cm}^2] = \frac{\text{charge} \times \mathrm{E}_{\text{pair}}}{\rho_{Si} \times l}, \tag{3.1}$$

in which E_{pair} is the average energy to produce an electron-hole pair in silicon; ρ_{Si} is the silicon density; and l is the path traversed by the particle in the sensor. l can be derived using the sensor thickness, t, and the angle between the direction of the track and the normal to the sensor surface, α, using the expression

$$l = \frac{t}{\cos\alpha}. \tag{3.2}$$

The incident angle α can be obtained using local cluster angular variables. Fig. 3.3 is a geometric representation of a charged particle traversing a sensor. The represented ϕ and θ are defined with respect to the local cluster frame (θ_{local} and ϕ_{local}). The relevant trigonometric relations are

$$\begin{aligned} \tan\theta_{\text{local}} &= \frac{b}{t} \\ \tan\phi_{\text{local}} &= \frac{a}{t} \\ \tan\alpha = \frac{d}{t} &= \frac{\sqrt{a^2+b^2}}{t}, \end{aligned}$$

from which one can derive the expression for α

$$\alpha = \arctan\sqrt{\tan^2\theta_{\text{local}} + \tan^2\phi_{\text{local}}}. \tag{3.3}$$

As charged particles traverse the inner detector, they become more separated. Therefore, merged clusters are more frequent for pixel layers which are closer to the beam-pipe. The IBL is the innermost layer of the ID, but it only has 4 bits available for ToT encoding, while 8 bits are available for the remaining layers, providing better ToT resolution. Due to these reasons, only B-layer clusters are considered for the F_{lost} measurement. This is true also for the analysis presented in the following sections.

For the F_{lost} measurement presented in Ref. [2], only two particle clusters were considered. To obtain the F_{lost}, a fit is performed to the $\mathrm{d}E/\mathrm{d}x$ distribution of B-layer

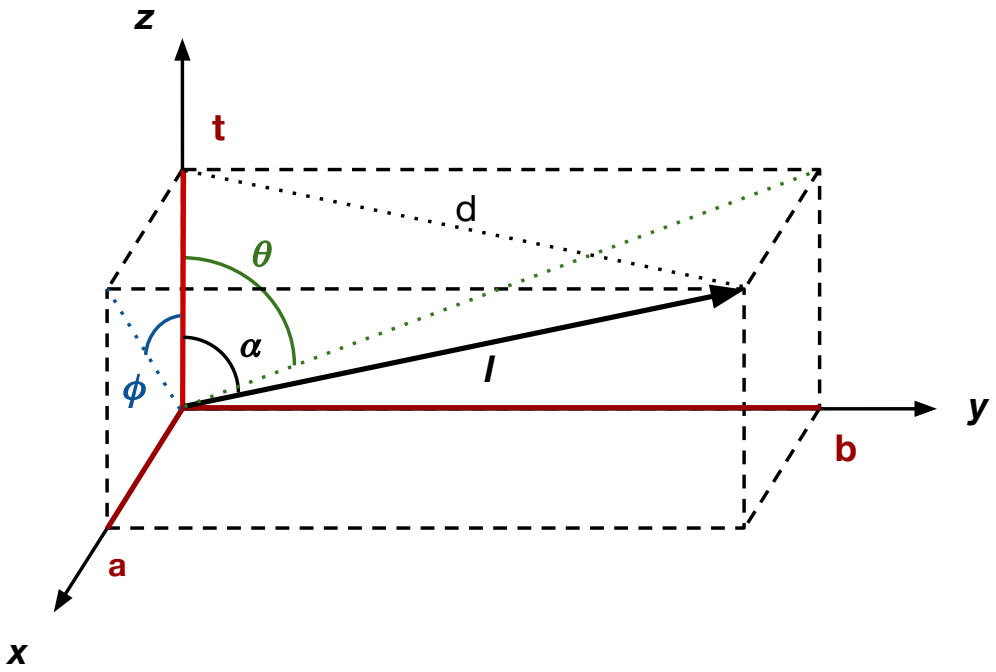


Figure 3.3: Geometric representation of a charged particle crossing a sensor, with an incident angle α. ϕ and θ are the local angles with respect to the module's local frame. The cuboid is defined by the point of entry and exit of the particle in the sensor.

clusters used by a single track matched to a jet within ΔR(jet, trk)<0.05. For these jet regions, the highest density of collimated charged particles is found. This was verified for jets in data and simulation, in Ref. [2], as can be seen in Fig. 3.4, which shows the average number of primary tracks per unit of angular area, as a function of $\Delta R(\text{jet}, \text{trk})$. Different jet p_{T} ranges were considered. The track density within jets increases with the transverse momentum of the jets. The distribution of $\text{d}E/\text{d}x$ for B-layer clusters associated to a single track with ΔR(jet, trk)<0.05 is referred to in this chapter as the measurement distribution.

$\text{d}E/\text{d}x$ templates derived for different charged particle multiplicities are used to model the different contributions to the total distribution. In Ref. [2], the single-track template is defined as the $\text{d}E/\text{d}x$ of singly-used B-layer clusters associated to tracks which are reconstructed further away from the jet core, with ΔR(jet, trk)>0.1. This far from the jet core, there is a relatively low density of collimated charged particles, less merged clusters are created. Clusters reconstructed in this region, in association to one single track are thus expected to truly be so. For the purposes of the F_{lost} measurement, a

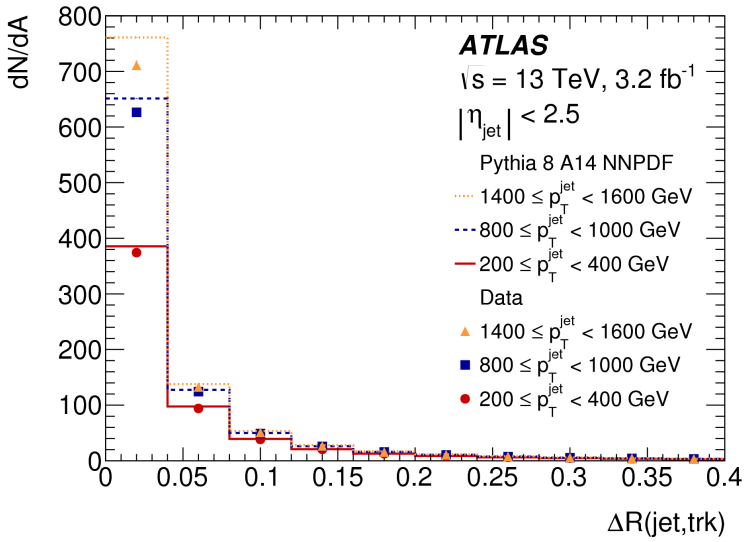


Figure 3.4: Average number of primary tracks per unit of angular area, as a function of $\Delta R(\text{jet}, \text{trk})$, for jets in data and di-jet MC simulation [2]. Different jet p_{T} ranges were considered.

track with ΔR(jet, trk)<0.05 is defined as being inside the jet core and a track with ΔR(jet, trk)>0.1 is said to be outside the jet core. A second $\text{d}E/\text{d}x$ template, a multiple-track template, is defined in Ref. [2] as the $\text{d}E/\text{d}x$ distribution of B-layer clusters used by more than one track from inside the jet core. The definitions of the different templates used in Ref. [2] are summarised in Fig. 3.5. Fig. 3.6 shows an example of the single and multiple-track $\text{d}E/\text{d}x$ templates built following these definitions. The multiple-particle distribution peaks at a $\text{d}E/\text{d}x$ value which is approximately two times the value for which the single particle distribution peaks, as expected. However, there is a smaller peak at around the expected $\text{d}E/\text{d}x$ from a single particle. This contribution arises from clusters which have been wrongly identified as merged. Additionally, this multiple-track template definition includes effects from all multiply-used clusters, not allowing to differentiate between particle multiplicities (two or three particles). This leads to the smaller peak which can be seen for $\text{d}E/\text{d}x > 3.2$ MeVg^{-1}cm^2. The effect from higher charged particle multiplicities was minimised by limiting the fit range to 1.1-3.07 MeVg^{-1}cm^2.

In Ref. [2], F_{lost} was defined as the fraction of tracks which truly contribute to 2 particle clusters from inside the jet core, which are reconstructed as a single track. It can

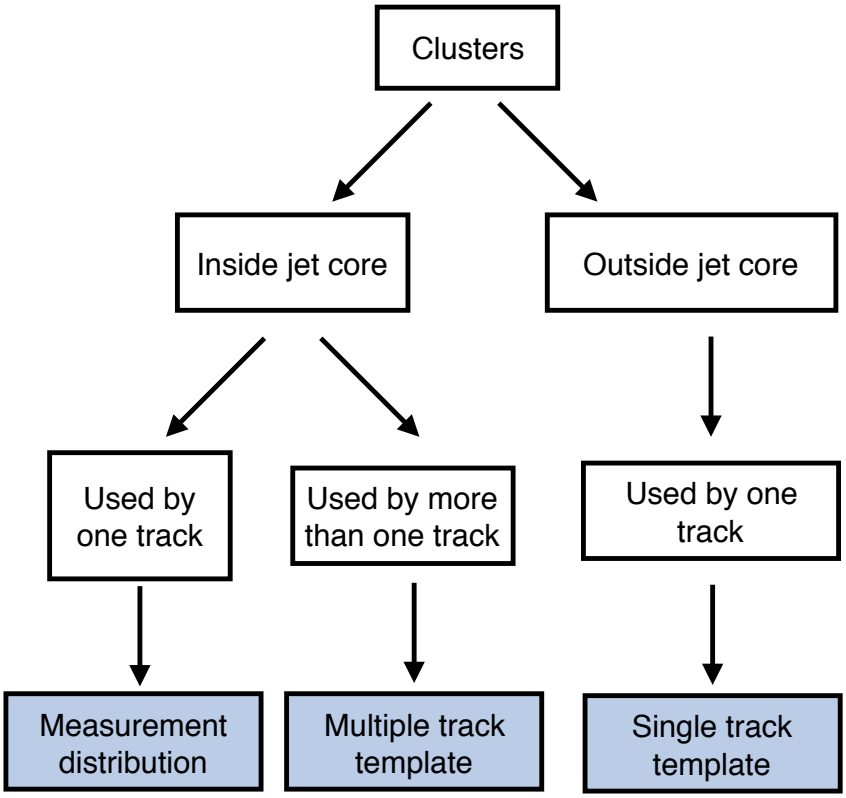


Figure 3.5: Summary of the definition of the templates and measurement distribution used in Ref. [2].

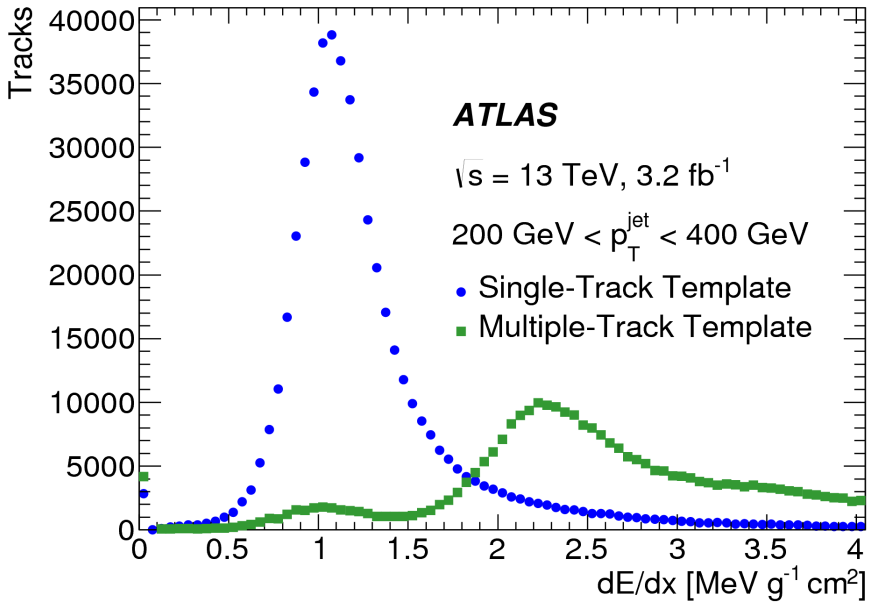


Figure 3.6: An example of the single and multiple particle $\mathrm{d}E/\mathrm{d}x$ templates [2], for $200 < p_{\mathrm{T}}(\mathrm{jet}) < 400$ GeV.

be written as

$$F_{\text{lost2}} = \frac{N_2^{\text{lost}}}{N_2^{\text{true}}}, \tag{3.4}$$

where N_2^{true} is the true number of tracks contributing to 2 particle clusters from inside the jet core, and N_2^{lost} is the subset of those tracks which are lost in reconstruction. N_2^{true} can be approximated by

$$N_2^{\text{true}} \approx N_2^{\text{reco}} + 2N_2^{\text{lost}}, \tag{3.5}$$

in which N_2^{reco} is number of tracks, reconstructed inside the jet core, associated to 2 track clusters. N_2^{lost} is obtained from the $\mathrm{d}E/\mathrm{d}x$ template fit. For each lost track, there are 2 expected tracks, which explains the factor of 2 in Eq. (3.5).

In order to study the dependence of lost tracks on jet p_{T}, the F_{lost} was measured differentially in jet p_{T}, in Ref. [2]. For larger jet p_{T}, the built $\mathrm{d}E/\mathrm{d}x$ templates have lower statistics, due to the existence of fewer jets with larger p_{T}. To mitigate this issue, the $\mathrm{d}E/\mathrm{d}x$ templates corresponding to the lowest jet p_{T} bin considered were used in the fits for all the other jet p_{T} bins. Fig. 3.7 shows the measured F_{lost2}, as a function of jet p_{T}, for 2015 data, compared to simulation. Due to the increase of the density of charged particles for higher jet p_{T}, the fraction of lost tracks increases accordingly, as can be seen in the results. This measurement has significant statistical and systematic uncertainties, with the latter being dominant. Sources of systematic uncertainty include: the choice of fit range, due to varying fractions of contributions from two and three particle clusters being inside the fitting window; using the lowest jet p_{T} templates to fit the measurement distribution from all other jet p_{T} bins; and a bias arising from the residual contribution from mis-identified singly-used clusters to the multiple-track template.

3.2 Data sample and selection requirements

The F_{lost} measurement presented in the following sections uses proton-proton data collected at $\sqrt{s}$ = 13 TeV during 2018. Events are only considered if they were recorded during stable beams and during data taking periods in which all relevant detector modules were operational.

The ATLAS data processing workflow includes a derivation step which reduces the

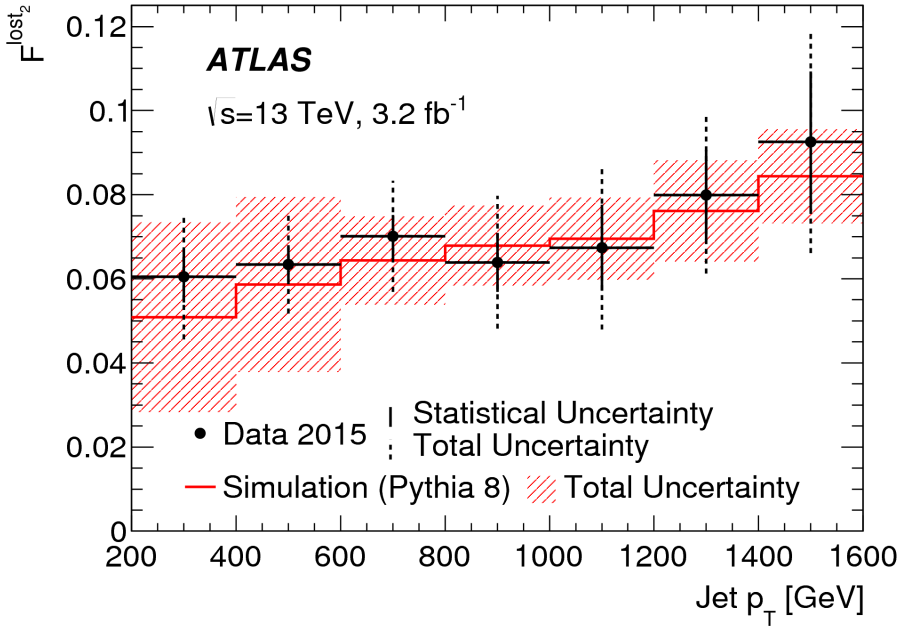


Figure 3.7: Measured F_{lost2} as a function of jet p_T compared to MC [2].

size of the datasets by removing events based on selection requirements (skimming) and by removing objects and variables not used in the specific analysis (thinning and slimming, respectively). Relevant information to the specific analyses might be added (physics objects decorations). The `IDTIDE` derivation was employed for this analysis, which includes additional cluster variables which are essential for the analysis. The derivation also performs skimming based on the triggers used in the analysis.

Events are required to pass a combination of single jet triggers with different p_T thresholds, with the lowest being 100 GeV. Furthermore, events are required to have at least one reconstructed primary vertex with at least three associated tracks. Jets are retained for the analysis if they have p_T >200 GeV and $|\eta| < 2.5$. Jet cleaning is performed, using the `BadLoose` working point. A requirement is also applied on the jet-vertex-tagger discriminant, to reject pile-up jets. A working point of 0.64 is used. Tracks are matched to the jet to which they are closest in ΔR. Jet and track pairs are required to have $\Delta R(\text{trk, jet}) < 0.4$. Only tracks which pass the `TightPrimary` working point defined in Section 2.3.1 are considered for the analysis. Tracks are further required to have one pixel hit per layer for each pixel layer except the IBL, for which they are allowed to have 0 or 1 hit.

3.3 Improvements to F_{lost} method

There are some shortcomings associated to the dE/dx templates described in Section 3.1, which lead to substantial systematic uncertainties on the F_{lost2} measurement. The multiple-track template includes a residual contribution from single particle clusters which are misidentified as merged clusters in reconstruction. Additionally, this multiple-track template includes contributions from clusters used by more than two particles. In order to address these issues, a new procedure was used to obtain multiple-track templates, based on sampling from the single-track template. In this way, the residual one particle cluster contribution is no longer present in the constructed templates, and it is possible to build two and three-track templates separately. The dE/dx distribution of singly-used clusters from outside the jet core is still taken as the single-track template.

As shown in Fig. 3.2, the dE/dx of a charged particle depends on its momentum. In order to take into account the momentum of the tracks when building the two-track templates, the following procedure was devised:

1. Single-track templates are built using single particle clusters from outside the jet core, for different slices of track momentum.

2. Two values are sampled simultaneously from the two-dimensional distribution of leading track momentum and subleading track momentum for tracks associated with two particle clusters from inside the jet core.

3. Depending on the values sampled in the previous step, two values are randomly sampled from the dE/dx distributions from the corresponding momentum slice.

4. The two values of dE/dx are summed. The resulting value is added to the two track dE/dx template.

The results presented in the following section were obtained when considering the momentum slices: $p_{\text{trk}} < 40\,\text{GeV}$, $40 < p_{\text{trk}} < 80\,\text{GeV}$, $80 < p_{\text{trk}} < 120\,\text{GeV}$, $120 < p_{\text{trk}} < 160\,\text{GeV}$ and $p_{\text{trk}} > 160\text{GeV}$.

By repeating this procedure N times, a two-track template with N entries is built. Three-track templates are built in an analogous manner. In order to study the F_{lost}

dependence on $|\eta(\text{trk})|$, as well as $p_{\text{T}}(\text{jet})$, $\text{d}E/\text{d}x$ templates were built for different $|\eta(\text{trk})|$ bins. The updated template definitions are summarised in Fig. 3.9[1]. The single, two and three-track templates derived as described, are shown in Fig. 3.8, for the different $p_{\text{T}}(\text{jet})$ and $|\eta(\text{trk})|$ bins considered in the analysis. Comparing the templates for different $|\eta(\text{trk})|$, for larger $|\eta|$ the width of the $\text{d}E/\text{d}x$ distribution decreases, and the distribution also shifts towards larger values. This effect is expected, since a track with larger $|\eta|$ travels a longer path in the sensor[2].

To reduce the impact of statistical fluctuations, the single-track template is smoothed using kernel density estimation (KDE) [79]. This is implemented using the `Roofit` [80] class `RooKeysPdf`. A fixed high smoothing parameter, $\rho = 3$, is used for the tail of the $\text{d}E/\text{d}x$ distribution, while a reduced smoothing parameter, $\rho = 0.15$, is used for the core, given the amount of statistics in each part of the distribution. Two `RooKeysPdf` objects, providing two smoothing predictions are used - KDE_{core} and KDE_{tail}. The total prediction is given by

$$P(\text{d}E/\ \text{d}x) = (1 - \text{Erf}(\text{d}E/\text{d}x - 3)) \times \text{KDE}_{\text{core}} + ((\text{Erf}(\text{d}E/\text{d}x - 3) + 1) \times \text{KDE}_{\text{tail}}, \tag{3.6}$$

in which the error functions are employed to ensure continuity between the predictions. No mirroring is used at the boundaries in the KDE estimation[3]. The result of the KDE smoothing for the lowest $p_{\text{T}}(\text{jet})$ bin, can be found in Fig. 3.10. The smoothed single-track templates for the remaining bins can be found in Appendix B. Since statistical fluctuations present in the templates built directly from the data are smoothed out, contrary to what was done in Ref. [2], the measurement distribution for each specific $|\eta(\text{trk})|$ and $p_{\text{T}}(\text{jet})$ is fitted using the templates derived for that same bin.

The number of tracks lost in reconstruction is obtained by performing a binned maximum likelihood fit to the $\text{d}E/\text{d}x$ distribution of single particle clusters from inside the jet core. The single-track template used to fit this distribution is built from clusters associated to tracks from outside the jet core. Fig. 3.11 shows the momentum

[1] There are two measurement distributions considered in the updated procedure, and a fit is also performed to the two-track clusters from inside the jet core, as will be explained in the following.

[2] The length of the traversed path depends on local θ, which varies with track $|\eta|$.

[3] The mirroring option pertains to the behaviour of the KDE estimation in the boundaries of the distributions. When mirroring is used, the dataset is mirrored at the boundary.

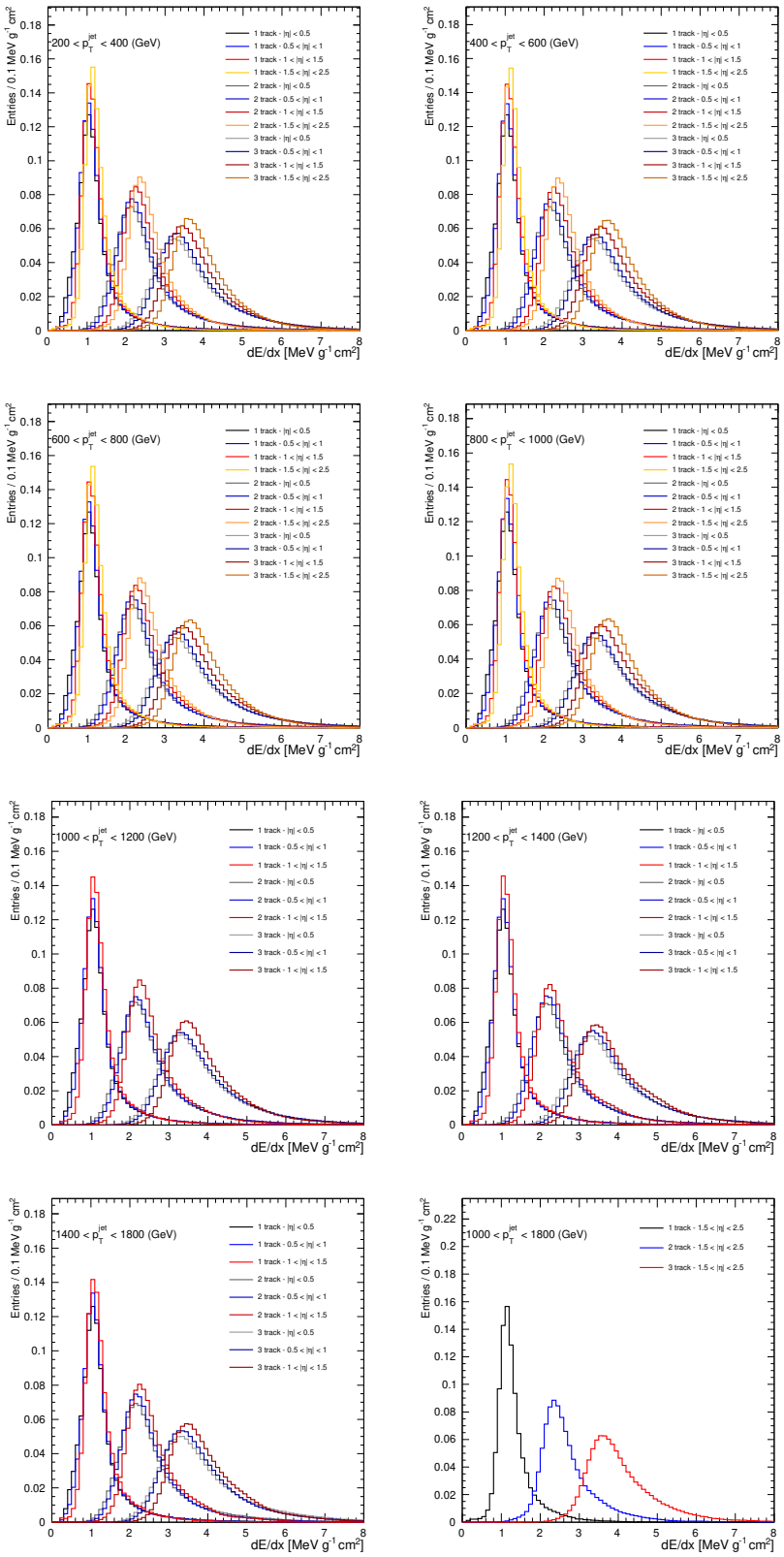


Figure 3.8: Single, two and three particle templates for the $p_{\mathrm{T}}(\mathrm{jet})$ and $|\eta(\mathrm{trk})|$ bins considered.

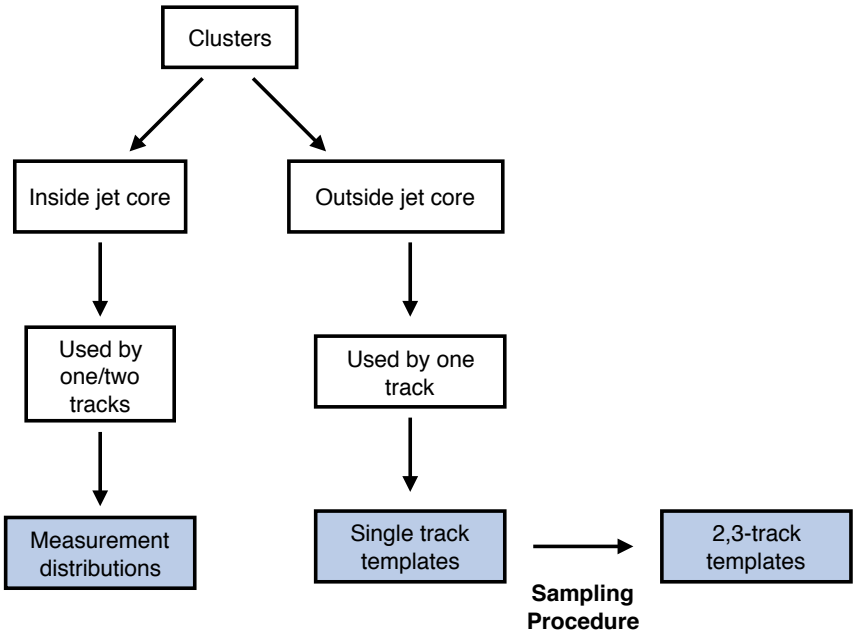


Figure 3.9: Summary of the definition of the templates and measurement distribution used in the updated procedure presented. Two measurement distributions are considered, one using the $\mathrm{d}E/\mathrm{d}x$ of one-track clusters from inside the jet core, as done in Ref. [2], and another using two-track clusters from the same jet region.

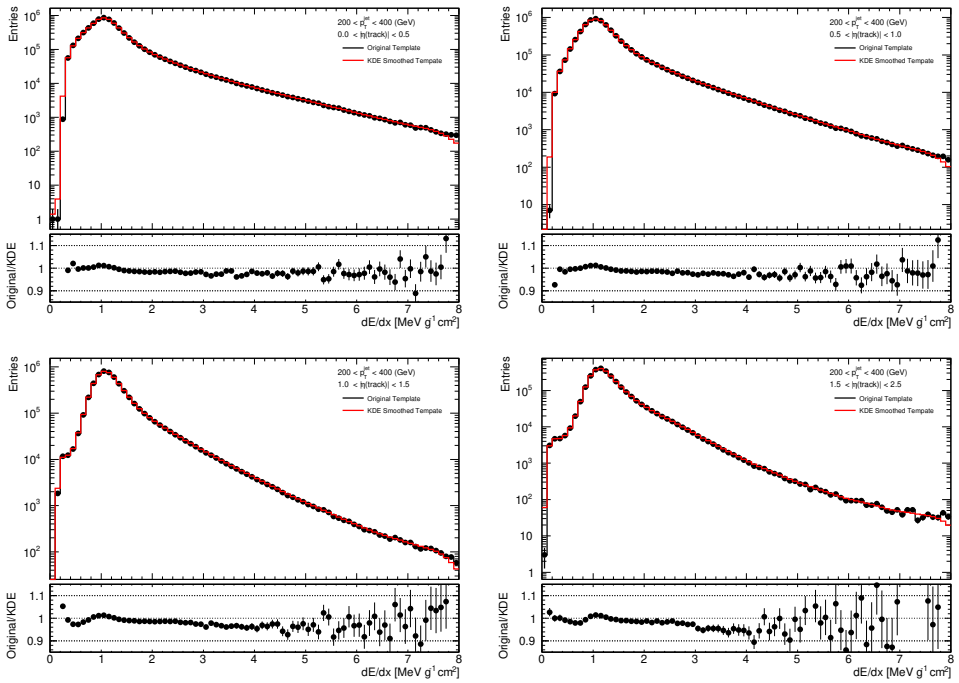


Figure 3.10: Single-track template before and after KDE smoothing, for $200 < p_{\mathrm{T}}(\mathrm{jet}) < 400$ GeV.

distributions for tracks reconstructed inside and outside the jet core. Tracks from outside the jet core tend to have a lower momentum and consequently the $\mathrm{d}E/\mathrm{d}x$ distribution of the clusters associated to those tracks will be shifted to lower values, with respect to the $\mathrm{d}E/\mathrm{d}x$ distribution for tracks inside the jet core. This means that the single-track template is shifted towards lower $\mathrm{d}E/\mathrm{d}x$ values, with respect to the one-particle contribution in data. This effect is also observed for the constructed two and three-track templates. Figure 3.12 shows the comparison between the two-track templates built from the sampling method described and the $\mathrm{d}E/\mathrm{d}x$ distribution of 2 track clusters from inside the jet core. The two particle template is shifted towards lower $\mathrm{d}E/\mathrm{d}x$ values. In order to account for these discrepancies, three free parameters are added to the fit to the measurement distribution, allowing each template to shift in $\mathrm{d}E/\mathrm{d}x$. Variations of the templates were derived, by shifting them to lower and higher $\mathrm{d}E/\mathrm{d}x$ values, and are implemented for each template using a moment morphing technique [81]. In this way, the PDF used to model each k-particle contribution is the result of interpolating between three alternative templates. The nominal and shifted versions of each k-track template are mapped to a parameter μ_k, with $k = 1, 2, 3$, in the following way:

- Default template mapped to $\mu_k = 0$.
- $\mathrm{d}E/\mathrm{d}x$ template artificially shifted to lower values mapped to $\mu_k =$ -1.
- $\mathrm{d}E/\mathrm{d}x$ template artificially shifted to higher values mapped to $\mu_k = 1$.

Figure 3.13 shows the single-track template distributions and their mapping to the shift parameter. The interpolation between them is realized using the `Roofit` class `RooMomentMorph` [81]. The value of μ_k is relative to the pre-fit size of the shift, which is 0.3 $\mathrm{MeVg^{-1}cm^2}$ for the single and two-track templates, and 0.6 $\mathrm{MeVg^{-1}cm^2}$ for the three-track template. The post-fit value of μ_k represents how much the k-track template shifted to adapt to the measurement distribution, with respect to the nominal template (e.g. $\mu_1 = 1$ represents a shift of 0.3 $\mathrm{MeVg^{-1}cm^2}$ of the single track template towards larger $\mathrm{d}E/\mathrm{d}x$ values).

The μ_k parameters are constrained, by adding a Gaussian term to the likelihood. From first principles, since the two and three-track templates are built by repeatedly

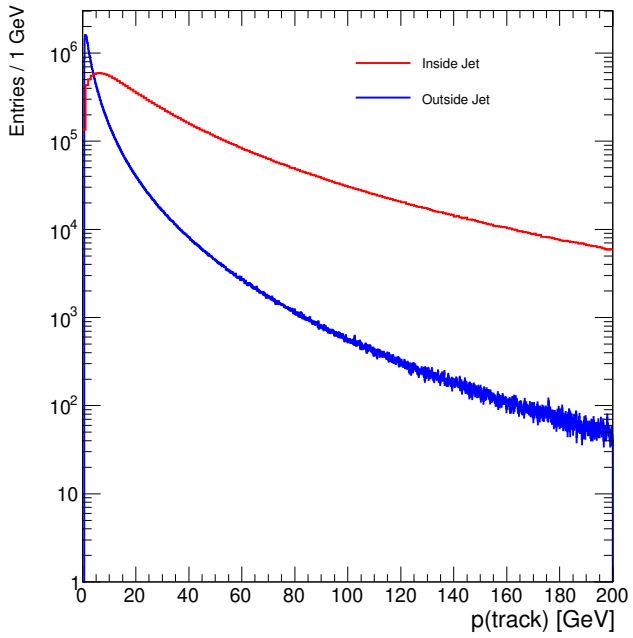


Figure 3.11: Track momentum of tracks from inside the jet core (blue) and outside the jet core (red), for $200 < p_{\mathrm{T}}(\mathrm{jet}) < 400$ GeV and $0 < |\eta(\mathrm{trk})| < 0.5$.

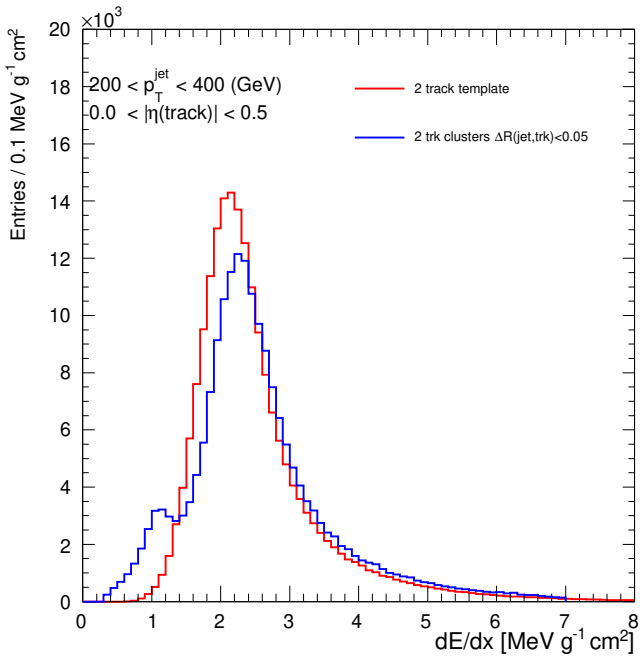


Figure 3.12: Comparison between the 2-track template derived as described (red) and the $\mathrm{d}E/\mathrm{d}x$ distribution of 2-track clusters from inside the jet core (blue), for $200 < p_{\mathrm{T}}(\mathrm{jet}) < 400$ GeV and $0 < |\eta(\mathrm{trk})| < 0.5$.

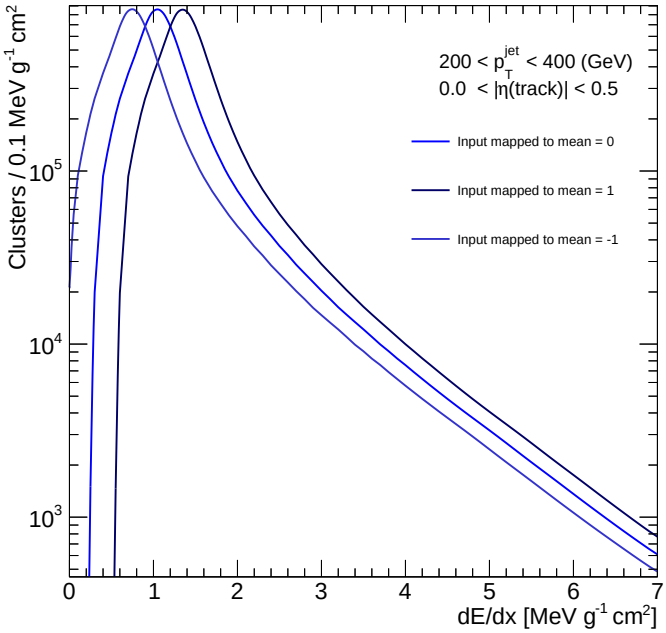


Figure 3.13: Mapping of shifted versions of the single-track template to μ_1 parameter.

sampling the single-track template, the size of their $\mathrm{d}E/\mathrm{d}x$ shift should be two and three times the shift of the single-track template. The μ_k parameters are thus constrained in the fit in the following way:

- μ_1 is constrained by a Gaussian centred at 0, with width 1.
- μ_2 - $2\mu_1$ is constrained by a Gaussian centred at 0, with width 0.1.
- μ_3 - $3/2\mu_1$ is constrained by a Gaussian centred at 0, with width 0.1. The 0.5 factor reflects the difference between the pre-fit size of the shift for μ_3 and μ_1.

The remaining free parameters in the fit are the normalisations of each template - $\mathrm{N}_k^{\mathrm{fit1}}$, in which $k = 1, 2, 3$. $\mathrm{N}_k^{\mathrm{fit1}}$ is the fit estimate for the number of clusters used by k charged particles. Given that the measurement distribution is the $\mathrm{d}E/\mathrm{d}x$ of single-particle clusters, for every cluster compatible with a 2 (3) particle $\mathrm{d}E/\mathrm{d}x$, there was 1 (2) lost tracks. In other words, $\mathrm{N}_1^{\mathrm{fit1}}$ translates to 0 lost tracks, $\mathrm{N}_2^{\mathrm{fit1}}$ to $\mathrm{N}_2^{\mathrm{fit1}}$ lost tracks and $\mathrm{N}_3^{\mathrm{fit1}}$ to 2 $\times \mathrm{N}_3^{\mathrm{fit1}}$ lost tracks.

$\mathrm{F}_{\mathrm{lost2}}$ is calculated using Eq. (3.7). As already explained, $\mathrm{N}_2^{\mathrm{reco}}$, which is used in the denominator in this expression, has a non-negligible contribution from clusters that

were wrongly identified as multiply used. This contribution can be estimated by fitting the $\mathrm{d}E/\mathrm{d}x$ distribution of two track clusters from inside the jet core, and extracting the normalisation of the single-track template ($\mathrm{N}_1^{\mathrm{fit2}}$). This value can be then subtracted to $\mathrm{N}_2^{\mathrm{reco}}$, correcting it. A modified expression of $\mathrm{F}_{\mathrm{lost2}}$ including this correction is used to obtain the results presented in the following section:

$$\mathrm{F}_{\mathrm{lost2}} = \frac{\mathrm{N}_2^{\mathrm{fit1}}}{\mathrm{N}_2^{\mathrm{reco}} - 2\mathrm{N}_1^{\mathrm{fit2}} + 2\mathrm{N}_2^{\mathrm{fit1}}}. \tag{3.7}$$

The $\mathrm{F}_{\mathrm{lost}}$ calculation was further extended to take into account three particle clusters. This means adding two lost tracks to $\mathrm{N}^{\mathrm{lost}}$ and three expected tracks to $\mathrm{N}^{\mathrm{true}}$ by every single track cluster with a $\mathrm{d}E/\mathrm{d}x$ compatible with the energy deposition of three particles (from inside the jet core). In the same way, one lost track must be added to the numerator by every two track cluster from inside the jet core with a three particle $\mathrm{d}E/\mathrm{d}x$. In this case, there are three expected tracks, but two are already accounted for in $\mathrm{N}_{2,3}^{\mathrm{reco}}$, which is the is number of tracks, reconstructed inside the jet core, associated to two or three track clusters. $\mathrm{F}_{\mathrm{lost}}$ is then obtained using the expression

$$\mathrm{F}_{\mathrm{lost}} = \frac{\mathrm{N}_2^{\mathrm{fit1}} + 2\mathrm{N}_3^{\mathrm{fit1}} + \mathrm{N}_3^{\mathrm{fit2}}}{\mathrm{N}_{2,3}^{\mathrm{reco}} - 2\mathrm{N}_1^{\mathrm{fit2}} + 2\mathrm{N}_2^{\mathrm{fit1}} + 3\mathrm{N}_3^{\mathrm{fit1}} + \mathrm{N}_3^{\mathrm{fit2}}}, \tag{3.8}$$

in which the superscript fit1 represents the fit to single track clusters and fit2 the fit to two track clusters. The results are discussed in Section 3.4.

3.4 Results and discussion

Fig. 3.14 shows the result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distributions of single track clusters from inside the jet core, for the lowest jet p_{T} bin considered. The results are shown for different $|\eta(\mathrm{trk})|$ ranges. The fit results for the remaining jet p_{T} bins can be found in Appendix A. The distributions display a clear peak compatible with the $\mathrm{d}E/\mathrm{d}x$ of one charged particle, and then a second, smaller peak at around twice that energy. This second peak arises from track clusters which were identified as singly-used in reconstruction, which are in fact merged clusters. For each one of

these clusters, there was one track lost in reconstruction. The two particle contribution becomes larger for higher $|\eta(\text{trk})|$ and jet p_{T}. The contribution from three particle clusters is negligible for lower jet p_{T} and $|\eta(\text{trk})|$ bins, but becomes significant for high jet p_{T} and $|\eta(\text{trk})|$[4]. Fig. 3.14 and the remaining fit results displayed in Appendix A show that the single, two and three-track templates provide a good description of the measurement distribution. In Fig. 3.15 the result of the fit to the $E/\text{d}x$ distributions of two track clusters from inside the jet core can be found for $200 < p_{\text{T}}(\text{jet}) < 400$ GeV. In this case, the prominent contribution arises from two particle depositions, as expected. The peak compatible with the $\text{d}E/\text{d}x$ of one particle clusters arises from misidentified clusters, as previously discussed. For all bins, there is also a significant contribution from three particle clusters. For high $\text{d}E/\text{d}x$ values the model underestimates the data, indicating that there is a four-particle contribution which is not included in the model is present in the data. In general, for both the fit to single track clusters and two track clusters, the constrained shift parameters are pulled with respect to the value they are being constrained to. This reflects the statistical power of the data to constrain the parameters in the model.

Figure 3.16a shows F_{lost2} (obtained using Eq. 3.7) as a function of jet p_{T} an track $|\eta|$. Figure 3.16b shows F_{lost} (obtained from Eq. 3.8). In general, the fraction of lost tracks increases with jet p_{T}, as was observed in Fig. 3.7. These results also show that tracks with a larger $|\eta|$ are more likely to be lost in reconstruction due to the dense environment. F_{lost2} ranges between approximately 2 and 15%, which is comparable to the values measured for 2015 data. When including three particle clusters, F_{lost} increases, ranging between approximately 4 and 20%. The error bars in Fig. 3.16 only include statistical uncertainties. Besides the measurement done in data, it is important to perform the same measurement for MC simulation, in order to understand how well the inefficiency due to high density of collimated charged particles in high p_{T} jets is modelled. The data-to-MC ratio observed would represent a systematic uncertainty for analyses targeting topologies with highly collimated tracks.

None of the sources of systematic uncertainties listed in Section 3.1 impact the

[4]For $p_{\text{T}}(\text{jet}) < 1000$ GeV and $|\eta(\text{trk})| < 1$, the three particle normalisation was fixed to 0 in the fit, since it was found to be negligible for these bins.

current measurement: there was no need to limit the fit range, since separate two and three-track templates were built; and the KDE smoothing performed to the templates reduced the impact of statistical fluctuations on the templates, such that the templates built for each bin could be used to fit the corresponding measurement distribution. Nevertheless, to fully validate the method and estimate residual biases, a truth-based closure test should be performed, using for instance di-jet simulated events. Using generator level information, the true F_{lost} could be calculated and compared to the estimate obtained through the method being discussed. Each simulated particle is generated with a "barcode". At generator level it is possible to access the list of particle barcodes associated with a cluster. After associating generated particles to clusters using barcodes, one could count the true number of particles contributing to each cluster (N_2^{true}). N^{lost} would be obtained by subtracting the number of reconstructed tracks associated to each cluster to the true number of particles using it. Thus, the true F_{lost} could be calculated and compared to the measurement.

Another possible future study could be performing the fits allowing the widths of the templates to float, giving more freedom for the model to adapt to the data. This change could improve the data-model agreement in the fit to two track clusters from inside the jet core. Further improvements could come from using a finer momentum slicing for the single track-templates which are sampled from in the construction of the two and three-track templates. This would improve the modelling of the dE/dx as a function of momentum and therefore could improve the two and three-track templates. Furthermore, a simultaneous fit to the dE/dx distributions of single and two track clusters from inside the jet core could be performed, with correlated μ_k, allowing to better constrain the shift in dE/dx of the templates.

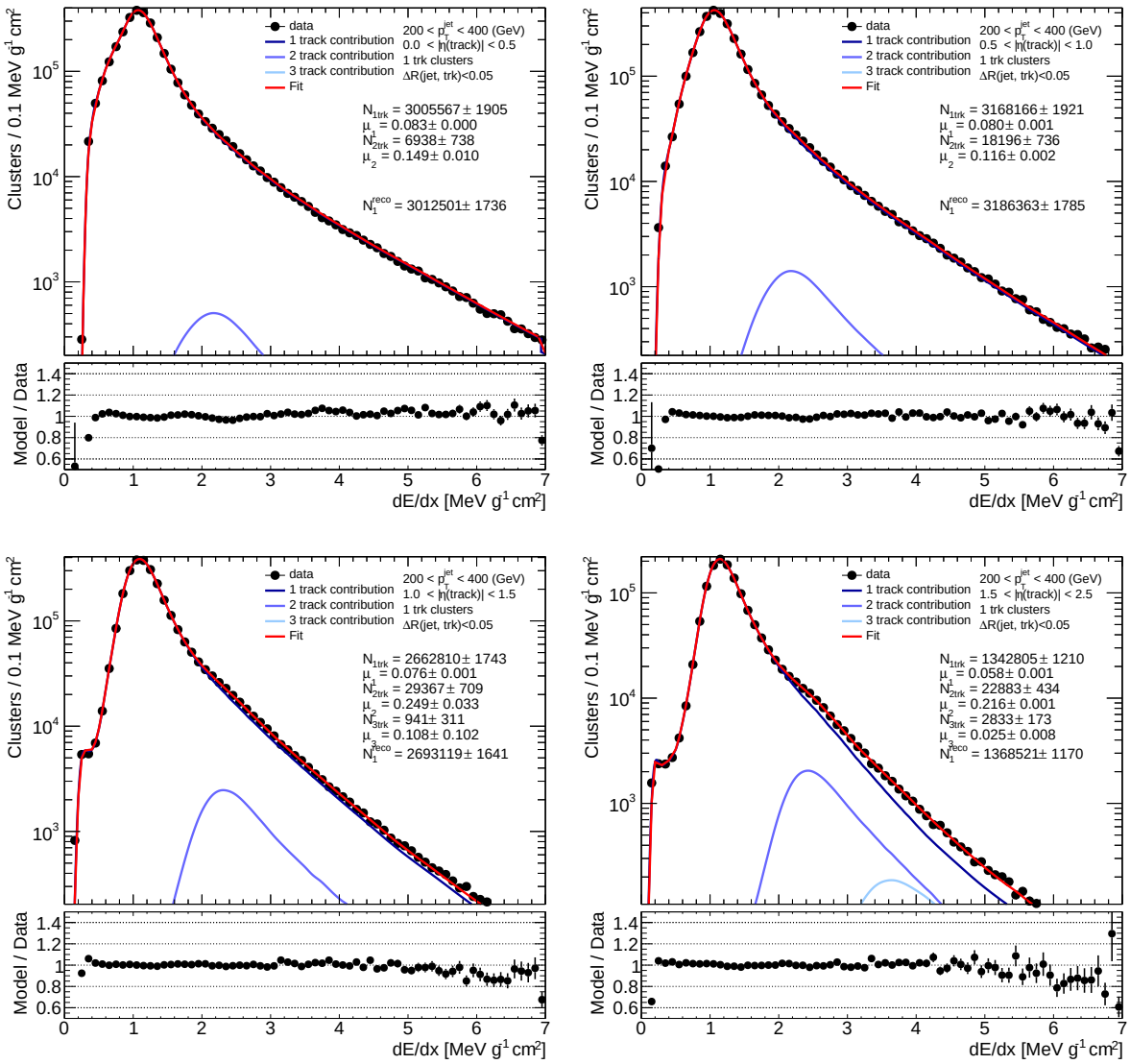


Figure 3.14: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of single track clusters from inside the jet core, for $200 < p_{\mathrm{T}}(\mathrm{jet}) < 400$ GeV and different $|\eta(\mathrm{trk})|$ ranges. The bottom panel shows the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_1^{\mathrm{reco}}$).

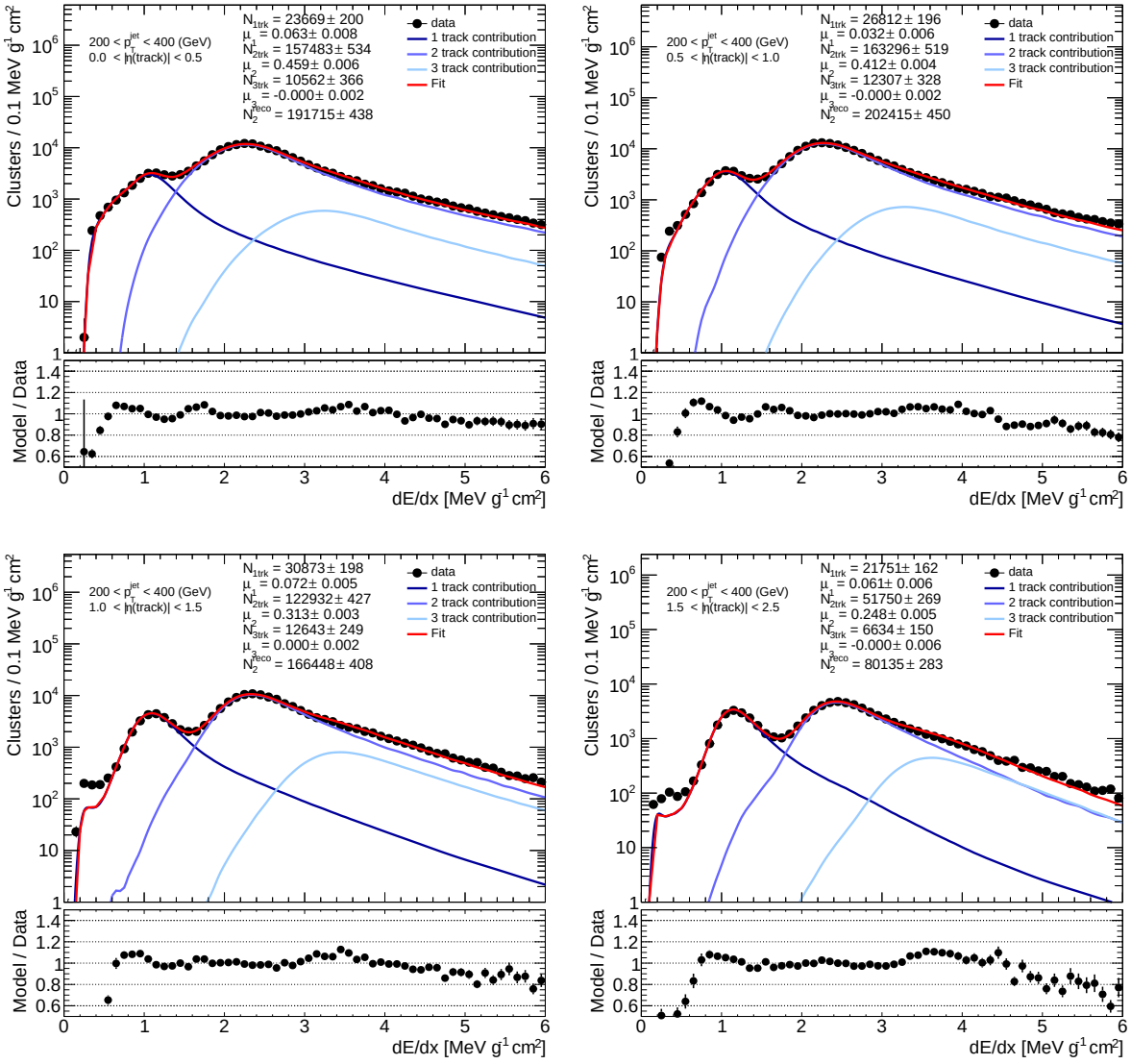
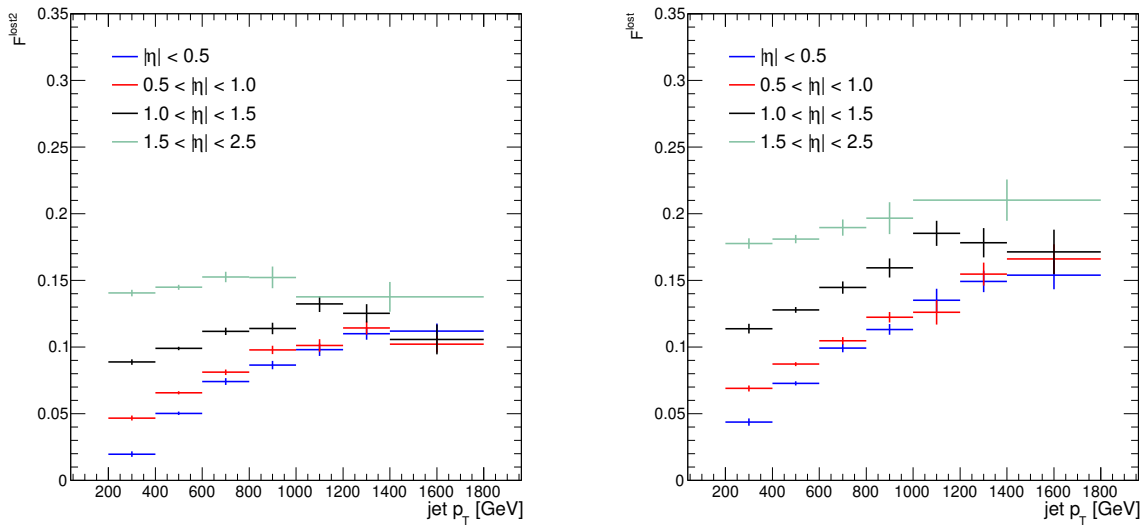


Figure 3.15: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of two track clusters from inside the jet core, for $200 < p_{\mathrm{T}}(\mathrm{jet}) < 400$ GeV and different $|\eta(\mathrm{trk})|$ ranges. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_2^{\mathrm{reco}}$).



(a) F_{lost2} (b) F_{lost}

Figure 3.16: Fraction of lost tracks as a function of $p_T(\text{jet})$ and $|\eta(\text{trk})|$. In (a) F_{lost2} is calculated accounting only for two-particle clusters, while in (b) the calculation included also three-particle clusters. Only statistical uncertainties are shown.

Chapter 4

Non parametric data-driven background modelling

This chapter presents a novel, non-parametric data-driven background modelling technique. In its application to an analysis, this technique employs a background enriched data sample, obtained using a relaxed version of the selection requirements, in order to estimate conditional probability functions. Two implementations of this technique have been developed and presented in Ref. [4]. One relies on ancestral sampling, while the second on conditional generative adversarial networks (cGANs). The first had been already been deployed in several ATLAS searches for exclusive decays of the Higgs and Z bosons [3, 82–84], prior to the publication of Ref. [4]. Nevertheless, the method was further developed, studied and, for the first time, presented in full detail in this latter publication. The author of this thesis performed the work of further developing the method and fully characterising and validating its performance, through its application to a $H \to \phi(K^+K^-)\gamma$ search case study and by performing ensemble tests. For completeness, the implementation using cGANs is also briefly discussed in this chapter.

4.1 Motivation

Background modelling is a critical element of particle physics measurements and searches. Monte Carlo (MC) simulations are often employed for the modelling of background processes. However, in certain cases, it might not be possible to model the background with enough accuracy using MC simulation, as is the case for fully

hadronic backgrounds and "fake" backgrounds, i.e. backgrounds composed of mis-reconstructed physics objects. Furthermore, producing large enough samples may be computationally costly. This becomes a limiting factor for analyses as the datasets delivered by collider experiments become larger, and the production of larger MC samples becomes necessary to match the statistical precision of the data. As a result, theoretical and statistical uncertainties associated with the modelling of background processes are often dominant uncertainties for analyses, limiting the sensitivity of searches or the precision of measurements [85, 86].

Another commonly used class of methods, parametric methods, relies on the choice of a functional form with a number of free parameters to fit the observed data, typically in the signal side-band. Adding free parameters increases its flexibility to adapt to the dataset, but decreases the statistical precision. Ultimately, the choice of parametrisation can potentially bias the signal extraction procedure, since there is no guarantee that it actually describes the underlying background shape. In order to address this issue, "spurious signal" calculations are often performed, which quantify the bias arising from different possible choices of background modelling function [7, 87]. To this end, background-only MC samples are fitted with a signal-plus-background model. Since there is no signal present in this samples, a non-zero post-fit signal yield is considered "spurious" and is taken as the estimate of the bias associated with the specific background model being tested (N_{SP}). This spurious signal effect is illustrated in Fig. 4.1. In this example, data was generated following a second order exponential distribution and fitted using a signal-plus-background model, with a first order exponential being used as the background fitting function and a gaussian being used to model the signal. The choice of background parametrisation does not have enough degrees of freedom to describe the underlying distribution, falsely measuring a significant signal yield. After performing spurious signal studies, different sets of criteria may be used to select the background modelling function, but generally, the tested function with the least associated bias and the least number of degrees of freedom is chosen. The N_{SP} associated with the selected function is then assigned as a systematic uncertainty. Since MC samples are used to perform this test, many times the issue of MC statistics ends up being a limiting factor, even if these samples are not being used directly in the modelling of the

background: the resulting spurious signal just reflecting the statistical fluctuations of the samples, inflating the associated systematic uncertainty, or even leading to a function with more degrees of freedom being chosen, resulting in decreased statistical precision. Moreover, these samples might not reliably describe the background, and often they need to be reweighted to match data control regions, prior to the spurious signal studies being performed. Some examples of recent results with significant spurious signal uncertainties can be found in references [88–90].

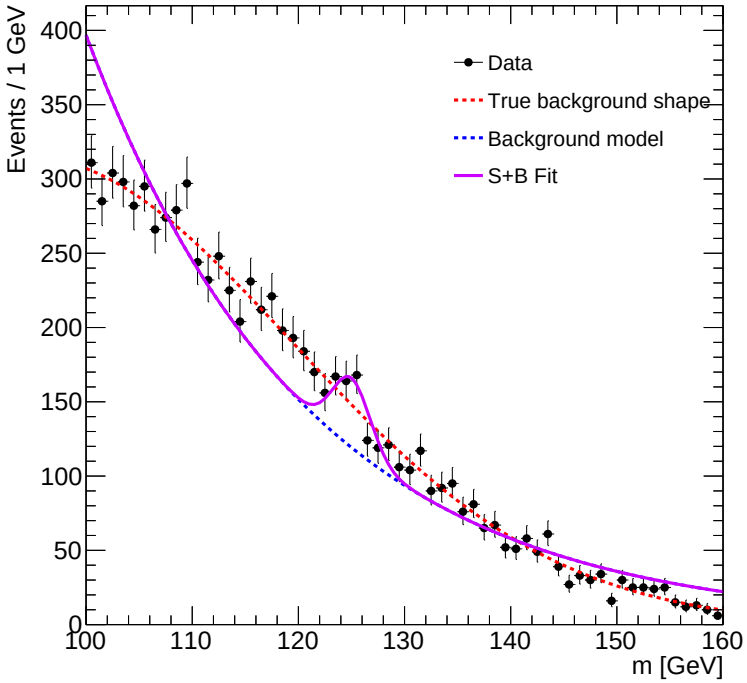


Figure 4.1: Representation of the "spurious signal" effect.

Another way to deal with the inherent uncertainty associated with the choice of background parametrisation is to treat it as a discrete nuisance parameter, in an analogous way to the treatment of other uncertainties as continuous nuisance parameters [91]. In this method, referred to as the discrete profiling method, an ensemble of parametric forms is combined at the likelihood level. Each background modelling function has an associated profile likelihood curve, and instead of considering only one of these when determining the best-fit value and uncertainty of the parameter of interest, the overall minimum of the different profile likelihood curves (envelope) is considered. A correction is further applied to the profile likelihood ratio to penalise functions

with more degrees of freedom. Practical complications arise for analyses including multiple event categories, with correlated systematics between them, since all possible combinations of background functions for different categories must be considered, in theory. In practise, approximations are often taken.

Data-driven alternatives for the modelling of backgrounds are widely used, like ABCD method [92] and sideband methods. Also different non-parametric methods for the modelling of smooth backgrounds have been proposed, for example using Gaussian Processes [93]. A novel approach for non-parametric data-driven background modelling was developed and it is discussed in the following sections. In its application to a specific analysis, this technique relies on a relaxed version of the event selection, allowing to estimate conditional probability density functions (PDFs) in a region enriched in background events. A sample of pseudo-events is generated following this background PDF and the analysis selection requirements are then applied to this background sample. A description of the background shape in a discriminant variable can then be obtained. Two implementations of this general strategy and their application to a case study are presented.

4.2 Ancestral sampling

4.2.1 Overview of the method

In a typical particle physics search, the presence of a signal in a dataset is quantified by performing a fit to the distribution of a discriminant variable in data. Therefore, the accurate description of the shape of the background in that variable is essential. The method presented in this section relies on ancestral sampling [94] from PDFs of the relevant variables to model the background shape in a discriminant variable.

Consider a search for a decay of the form $X \to YZ$. The statistical analysis is often performed using the invariant mass of X, which is a compound variable of the kinematics of the reconstructed YZ system. Through the modelling of the distributions and correlations between the kinematic components of the Y and Z four-vectors one can obtain the distributions of compound variables such as the invariant mass. The most

general distribution describing the Y and Z four-vectors would be an eight dimensional distribution of the kinematic components, which could be built as an eight dimensional histogram. One can also factorise this eight dimensional PDF into the product of lower dimensionality PDFs, by only explicitly describing the largest correlations observed in data. This allows to only use more practical lower dimensionality histograms in the modelling (up to three dimensions). Pseudo-events can then be generated through sequential sampling from the factorised distributions: each event is described by the values sampled for each kinematic component of the Y and Z four-vectors and the compound kinematic properties calculated from them. One must note that other variables, which affect the distribution of the compound variable of interest (which could be for instance the isolation measures of X and Y employed in the selection), should be included in the background PDF being sampled.

Different regions are defined for use in the background modelling process. The Signal Region (SR) is the region which includes all the analysis selection requirements. Ultimately, it is in the SR that the background shape must be described for further analysis. A second region, which will be referred to as Generation Region (GR) is defined by relaxing different SR requirements, enriching the sample in background events. The factorised n-dimensional PDF is constructed to describe the data distributions in this region. Finally, one might also define intermediate Validation Regions (VRs), by applying the SR requirements on top of the GR, one at a time. The variables that are used in selection criteria which is applied on top of the GR defining the SR must be included in the modelling PDF, so that the compound variable of interest is modelled accurately in the SR. The VRs allow to check the individual effect of each requirement on the model.

The construction of the model PDFs can be done using the N events in the GR dataset directly or, alternatively, one can sample N events from the GR with replacement and then construct the PDFs. These two approaches are compared in the following and are found to give similar results.

4.2.2 Application to case study

4.2.2.1 Search for $H \to \phi(K^+K^-)\gamma$

To date, there has been no experimental confirmation of the coupling of the Higgs boson to quarks of the first and second generations. These couplings could be probed through exclusive decays of the Higgs boson into a meson and a photon [95–97]. Decays of this type are rare and are yet to be observed. Various searches for these decays have been performed at the LHC, by the ATLAS [3, 82–84] and CMS experiments [98, 99]. The case study used in this section to demonstrate the application of the method is a search for the Higgs boson decay into a ϕ meson and a photon. The ATLAS collaboration has published such a search, which used a dataset collected at $\sqrt{s} = 13\,\text{TeV}$, corresponding to $36\,\text{fb}^{-1}$ [100]. The ϕ meson is reconstructed in its $\phi \to K^+K^-$ decay. The final state consists of a pair of oppositely charged tracks, with an invariant mass consistent with the ϕ mass, recoiling against an isolated photon. The main contributions of background for this analysis are photon + jet processes and multijet processes, in which a pair of charged particles from a jet form the ϕ meson candidate. It is difficult to reliably model these processes using MC simulation, and as such this search represents an ideal case study for this background modelling technique. A simplified version of the $H \to \phi(K^+K^-)\gamma$ search is employed here, using a sample of simulated background events as the "data". The reconstruction and event selection used are very similar to the ones described in Ref. [100].

Signal and background samples were generated using MC simulation. Only the dominant signal and background processes were considered. Higgs boson production through gluon fusion was simulated with Pythia 8.244 [101]. The CT14nlo PDF set [102] was used. The Higgs boson was decayed using the same generator. No other Higgs production and decay processes were simulated. In the case of the background, only the $\gamma +$ jet contribution was considered. This process was simulated using Sherpa 2.2.10 [103] with the NNPDF3.0 PDF set [104]. Up to two additional jets were simulated at the matrix element level. The response of an LHC type detector is simulated using version 3.4.2 of the Delphes framework [105]. The ATLAS-like configuration card of Delphes was used, with minimal modifications to the charged hadron tracking

resolution. The resolution from version 3.3.3 was used, since it was found to be more representative of the tracking resolution of the ATLAS detector. 300 million events were simulated. Accounting for the effect of the generator weights, this corresponds to an effective integrated luminosity of $20\,\text{fb}^{-1}$.

Events are required to have a photon with $p_{\text{T}}(\gamma) > 35\,\text{GeV}$ and that is reconstructed within $|\eta(\gamma)| < 2.37$, not including the transition region between the barrel and the endcap of the ATLAS calorimeter, $1.37 < |\eta(\gamma)| < 1.52$. ϕ meson candidates are reconstructed from pairs of oppositely charged tracks, which are required to have $p_{\text{T}}(K^{\pm}) > 15\,\text{GeV}$ and $|\eta(K^{\pm})| < 2.5$. Furthermore, the leading track of the pair must fulfil $p_{\text{T}}(K^{\pm}) > 20\,\text{GeV}$. These p_{T} requirements reflect the trigger thresholds used in Ref. [100]. ϕ candidates must satisfy $1.012 < m(\phi) < 1.028\,\text{GeV}$. This considered range accounts for the tracking resolution of the detector. ϕ candidates are also required to be isolated: they must satisfy $I(\phi) < 0.5$, with $I(\phi)$ being defined as the ratio between the sum of the p_{T} of tracks within $\Delta R = 0.2$ of the leading track (excluding the tracks that form the ϕ meson) and $p_{\text{T}}(\phi)$. In each event, the photon candidate with the highest p_{T} and the ϕ candidate with an invariant mass closest to the ϕ meson mass are selected to form the Higgs boson candidate. In order to target decay products that are created back-to-back events are also required to satisfy $\Delta\Phi(\phi,\gamma) > \pi/2$. A requirement on the transverse momentum of the ϕ meson is also applied, depending on the three-body invariant mass $m_{\phi,\gamma}$: for $m(\phi,\gamma) < 91\,\text{GeV}$ a p_{T} threshold of 40 GeV is used; if $m(\phi,\gamma) \geq 140\,\text{GeV}$, this threshold is instead 47.2 GeV; and for $91 \leq m(\phi,\gamma) < 140\,\text{GeV}$ the threshold is linearly varied from 40 GeV to 47.2 GeV as a function of $m(\phi,\gamma)$. This requirement was optimised for the simultaneous search for Higgs and Z bosons decaying to $\phi\,\gamma$. The set of criteria described defines the Signal Region (SR).

The presence of signal in the SR is quantified by performing a maximum likelihood fit to the Higgs boson invariant mass distribution, $m(\phi,\gamma)$. Therefore, the $m(\phi,\gamma)$ shape of the background must be derived. In the present example, the method is applied to the γ + jet simulated sample.

4.2.2.2 Model construction

As described in Section 4.2.1, a background enriched Generation Region is defined by relaxing the SR requirements. In this case, the requirements on $I(\phi)$ and $p_{\mathrm{T}}(\phi)$ are relaxed. Two validation regions, VR1 and VR2 are defined by applying the requirement on $p_{\mathrm{T}}(\phi)$ and $I(\phi)$ on top of the GR, respectively. Table 4.1 summarises the definition of the regions employed.

	Minimum $p_{\mathrm{T}}(\phi)$ requirement	Maximum $I(\phi)$ requirement
GR	35 GeV	Not applied
VR1	Varying from 40 to 47.2 GeV	Not applied
VR2	35 GeV	0.5
SR	Varying from 40 to 47.2 GeV	0.5

Table 4.1: Definition of the regions considered in the background modelling.

The events in the GR (either directly or after sampling with replacement has been performed) are used to construct histograms of the relevant kinematic and isolation variables. The use of 2D and 3D histograms allows to describe the most important correlations between variables which are observed in the data. By sequentially sampling the constructed distributions, an ensemble of background pseudo-events is generated. Each pseudo-event is fully described by the meson and photon four-vectors, which form the Higgs pseudo-candidate, and by $I(\phi)$. Therefore, the components of the meson and photon four-vectors and the $I(\phi)$ of the data events need to be described by the background modelling generation templates. The list of histograms used in the background modelling for this $H \to \phi(K^+K^-)\gamma$ case study is reported in Table 4.2. The sampling procedure from the templates, for each pseudo-event is represented in Fig. 5.23 and goes as follows:

1. Values for $p_{\mathrm{T}}(\phi)$ and $p_{\mathrm{T}}(\gamma)$ are simultaneously sampled from template A.

2. Given the values of $p_{\mathrm{T}}(\phi)$ and $p_{\mathrm{T}}(\gamma)$ sampled in step 1, template B is projected along the $\Delta\Phi(\phi,\gamma)$ dimension and a value for $\Delta\Phi(\phi,\gamma)$ is sampled. One must note that the binning of $p_{\mathrm{T}}(\phi)$ and $p_{\mathrm{T}}(\gamma)$ in template B is chosen to be coarser than in template A.

3. Depending on the value of $p_{\mathrm{T}}(\gamma)$ sampled in step 1, template D is projected along the $I(\phi)$ dimension and a value for $I(\phi)$ is sampled.

4. Given the value of $\Delta\Phi(\phi,\gamma)$ sampled in step 2, template C is projected along the $\Delta\eta(\phi,\gamma)$ dimension and a value for $\Delta\eta(\phi,\gamma)$ is sampled.

5. Templates E and F are independently sampled, and values for $\eta(\gamma)$ and $\Phi(\gamma)$ are defined. Imposing $m(\gamma) = 0$, the photon four-momentum becomes fully defined at this stage.

6. A value for $m(\phi)$ is sampled from generation template G. $\eta(\phi)$ and $\Phi(\phi)$ are calculated from the previously sampled values of $\eta(\gamma)$, $\Phi(\gamma)$, $\Delta\Phi(\phi,\gamma)$ and $\Delta\eta(\phi,\gamma)$. The ϕ four-momentum is fully defined at this stage.

Template Name	Dimensionality	Variable 1	Variable 2	Variable 3
A	2D	$p_{\mathrm{T}}(\phi)$	$p_{\mathrm{T}}(\gamma)$	-
B	3D	$\Delta\Phi(\phi,\gamma)$	$p_{\mathrm{T}}(\gamma)$	$p_{\mathrm{T}}(\phi)$
C	2D	$\Delta\eta(\phi,\gamma)$	$\Delta\Phi(\phi,\gamma)$	-
D	2D	$I(\phi)$	$p_{\mathrm{T}}(\gamma)$	-
E	1D	$\eta(\gamma)$	-	-
F	1D	$\Phi(\gamma)$	-	-
G	1D	$m(\phi)$	-	-

Table 4.2: The definition of the generation templates used in the sampling sequence.

By repeating this steps N times, an ensemble of N background pseudo-events is generated in the GR. Composite variables such as $m(\phi,\gamma)$ and $p_{\mathrm{T}}(\phi,\gamma)$ can be calculated for each event. After applying the SR requirements to the GR pseudo-event sample, one can obtain the $m(\phi,\gamma)$ background shape in SR. Besides obtaining a prediction for the background shape, the normalisation of the background in SR is also predicted. This is done by scaling the distributions by the ratio of the number of data events in GR to the number of generated pseudo-events (N). The validation regions are scaled in the same way. Accounting for the generator weights, the GR contains 30175 events, while the SR contains 11885 events.

The linear correlations between the variables directly included in the modelling and $m(\phi,\gamma)$ can be found in Fig. 5.24, both for the sample of simulated $\gamma+$jet events and for the background pseudo-events. In general, the magnitude and hierarchy of the

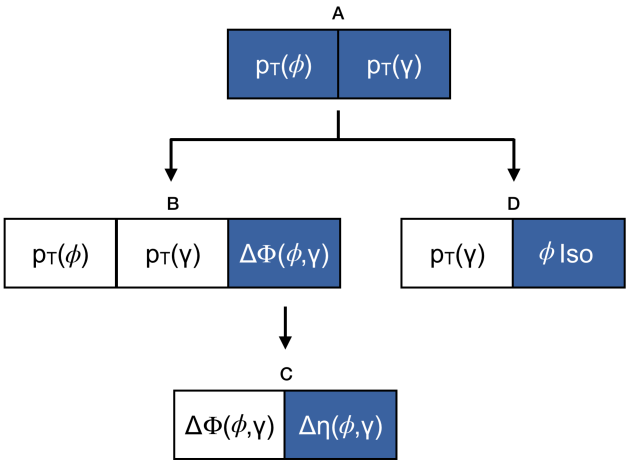
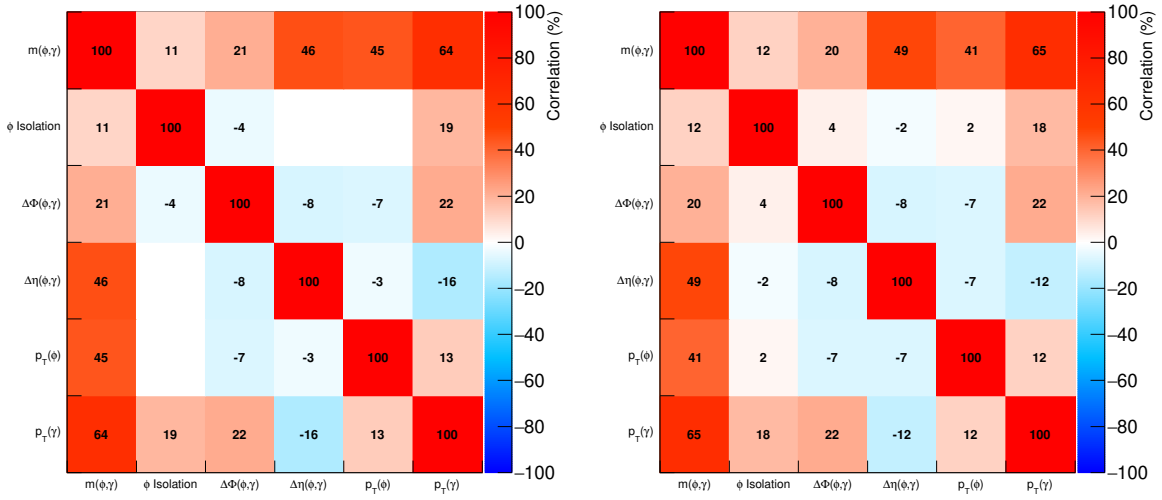


Figure 4.2: Schematic of the sampling sequence followed in the modelling. Variables not included in the schematic are sampled in a factorised, uncorrelated, manner from an 1-dimensional template. Groups of two (three) variables represent two (three)-dimensional templates. Arrows are used to show the sequential order of the steps in the sampling. Variables are highlighted with color at the step in which they are defined for each pseudo-candidate. In this schematic, "Iso" stands for isolation.

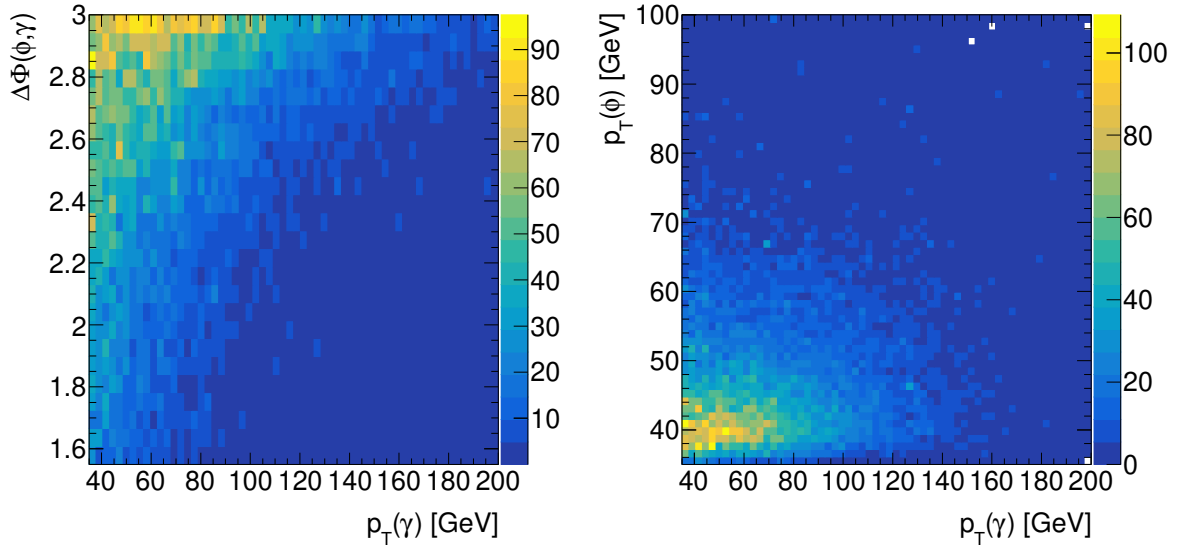
linear correlations measured in data is well described by the model. Two dimensional distributions between two of the most correlated variables in the $\gamma+$jet sample are shown in Fig. 4.4. These distributions show hints of a non-trivial correlation between these variables, which is not strictly linear, and which the modelling tries to capture through the sampling.

Data-model comparisons for the distribution of the compound variable of interest, $m(\phi,\gamma)$, are shown in Fig. 4.5 and Fig. 4.6, for the cases in which the GR dataset is sampled directly and after sampling with replacement, respectively. The distributions are shown in all the regions defined in Table 4.1. The good agreement that can be seen for both the realisations of the modelling and for the different regions, reflects the good description of the relevant correlations between variables in the $\gamma+$jet GR event sample. Some residual mis-modelling can be seen in SR for low values of $m(\phi,\gamma)$. Such small deviations are addressed by the implementation of shape systematic uncertainties, as explained in Section 4.2.2.4. SR distributions of other compound and primary variables of the modelling can be found in Appendix C. Given the observed similar performance between the model when sampling the GR dataset with or without replacement, in the remaining studies of this section results will be shown only for the sampling without



(a) (b)

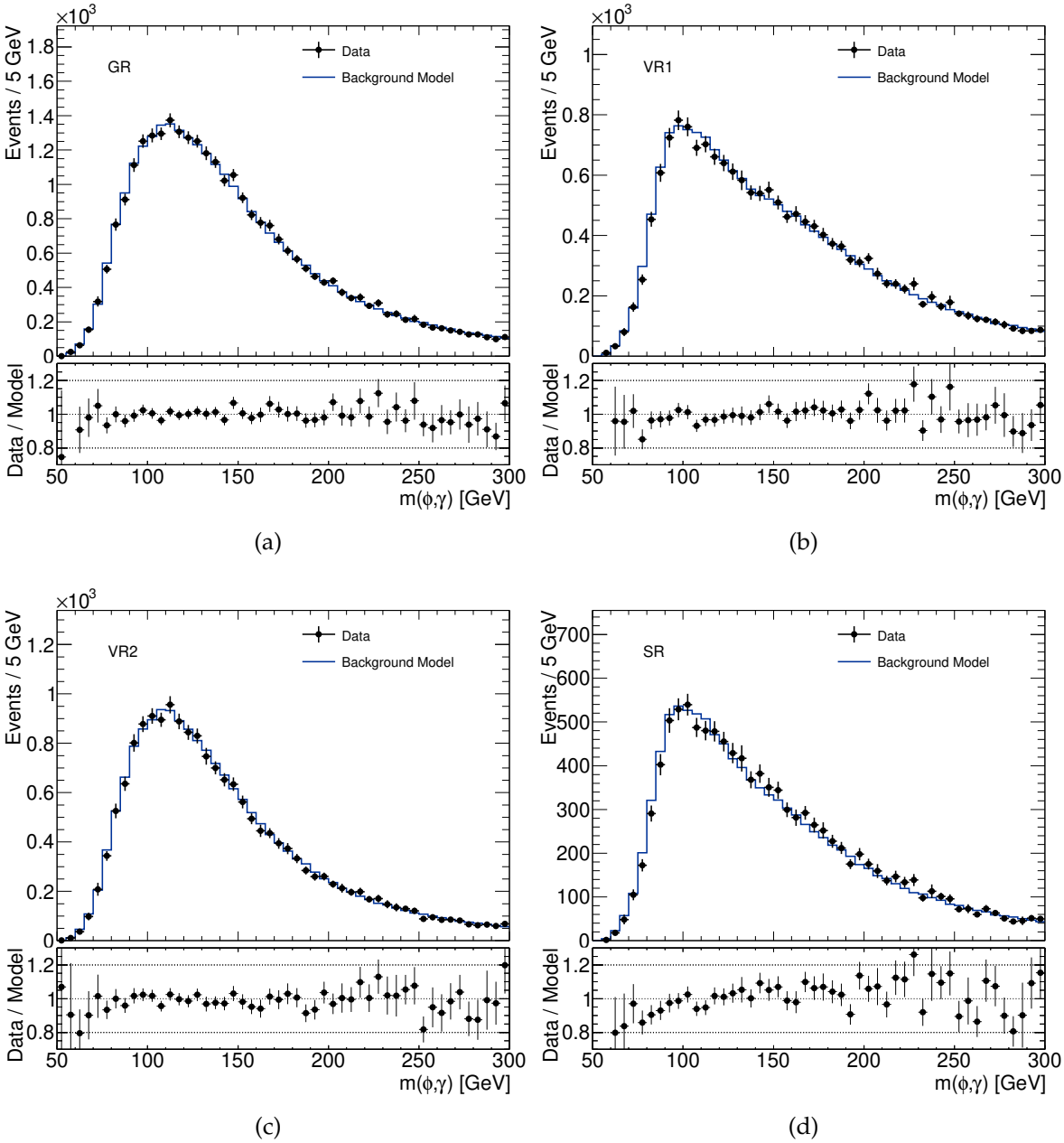
Figure 4.3: Observed linear correlations between the variables included in the background modelling, shown for the γ + jet MC sample (left) and the pseudo-events sample (right), in the GR. The background model was constructed by directly sampling the GR.



(a) (b)

Figure 4.4: Two dimensional distributions of $p_{\mathrm{T}}(\gamma)$ versus $\Delta\Phi(\phi,\gamma)$ (left) and $p_{\mathrm{T}}(\phi)$ (right).

replacement.



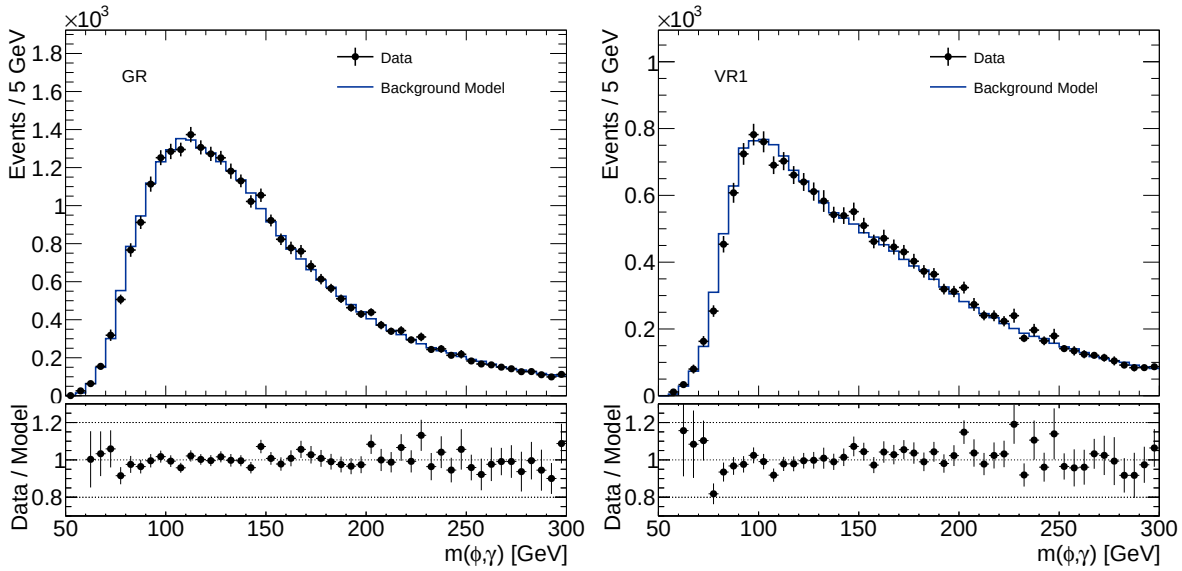
(a) (b)

(c) (d)

Figure 4.5: $m(\phi, \gamma)$ distributions in the regions considered for the γ + jet MC sample and background model built using the events in GR directly.

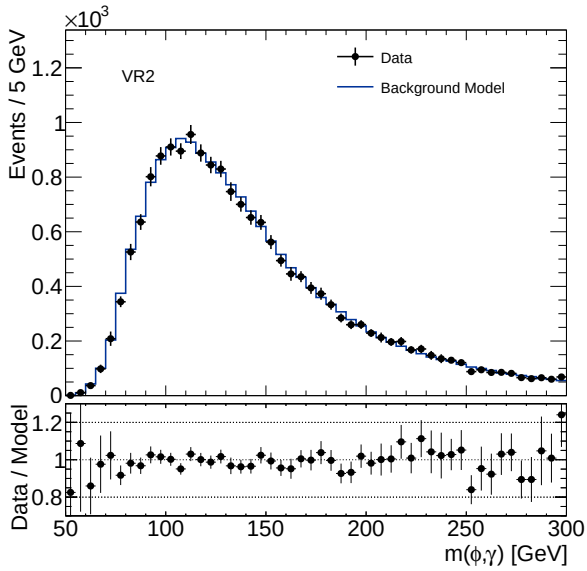
4.2.2.3 Signal injection tests

In the previous section, the modelling was demonstrated using a background-only γ + jet sample. In a real search, a contribution from signal events might be present in the SR dataset. The GR is constructed with the intention of performing the modelling

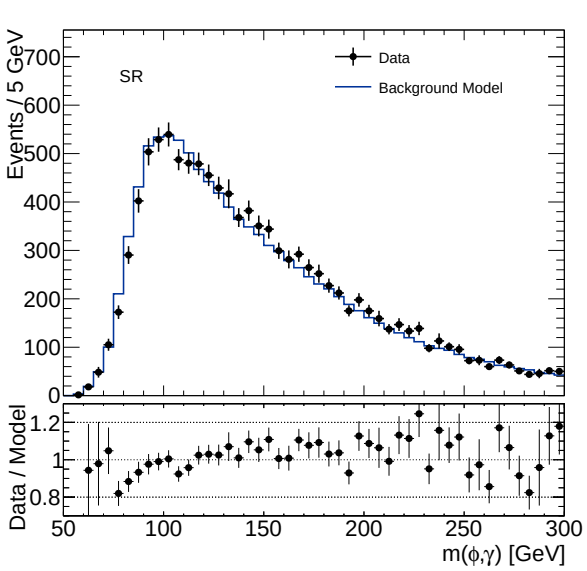


(a)

(b)



(c)



(d)

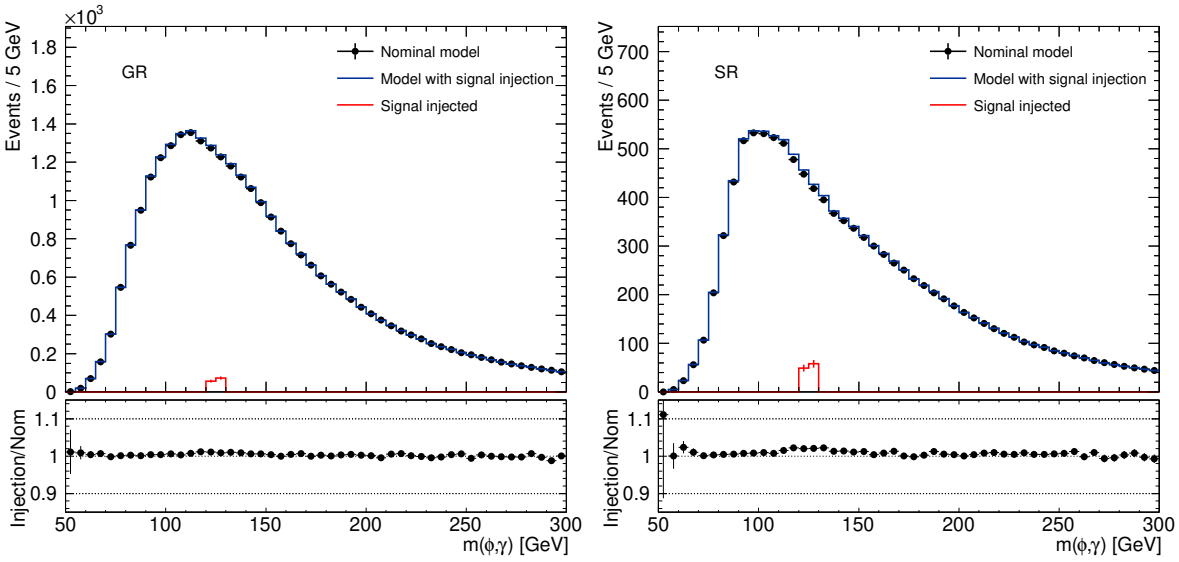
Figure 4.6: $m(\phi, \gamma)$ distributions in the regions considered for the γ + jet MC sample and background model built after sampling with replacement from the GR.

using a dataset dominated by background events. However, in a case in which the sensitivity to the signal is large, and consequently the signal-to-background ratio in GR is large, the performance of the background modelling method could degrade. In order to quantify the effect of a significant signal contribution on the background modelling, signal injection tests were performed. 130 $H \to \phi(K^+K^-)\gamma$ signal events were added into the γ + jet GR dataset, corresponding to 10.4% of the background in the $122.5 < m(\phi,\gamma) < 127.5\,\text{GeV}$ region in which the signal is located. This level of signal contamination in the GR would lead to an observation of the $H \to \phi(K^+K^-)\gamma$ signal, with an estimated 5.5σ significance in SR. After the injection of signal events in the GR, the background modelling PDFs were constructed and a sample of pseudo-events was generated as described in Section 4.2.2.2. The comparison between the resulting background model and the "nominal" model derived without any signal injection can be found in Fig. 4.7. The background estimate between $122.5 < m(\phi,\gamma) < 127.5\,\text{GeV}$ in the SR increased just by 2%. This result shows the robustness of the method under signal contamination. The background model behaves in the same way in the presence of any resonant contribution, since the features of these contributions are diluted by the process of factorising the background n-dimensional PDF to the product of lower dimensionality distributions. This means that any resonant background needs to be modelled separately.

4.2.2.4 Systematic uncertainties

A strategy for the implementation of systematic uncertainties associated with the predicted shape of the background, which allows analyses to account for potential residual mis-modelling, is presented and demonstrated in this section.

Variations of the nominal background model shape are derived, to ultimately be incorporated in the likelihood fit, allowing the background PDF to adapt to the observed dataset. The moment morphing technique described in Ref. [81] is employed, with each variation being controlled by a nuisance parameter. When devising the set of alternative shapes, one should try to capture different modes of deformation of the nominal background shape, trying to minimise the correlations between the derived shapes. The exact size of the variations is not as important as it only needs to be large



(a) (b)

Figure 4.7: Comparison of the $m(\phi, \gamma)$ background model distributions for the GR (a) and the SR (b), derived with and without an injection of signal events in the GR dataset.

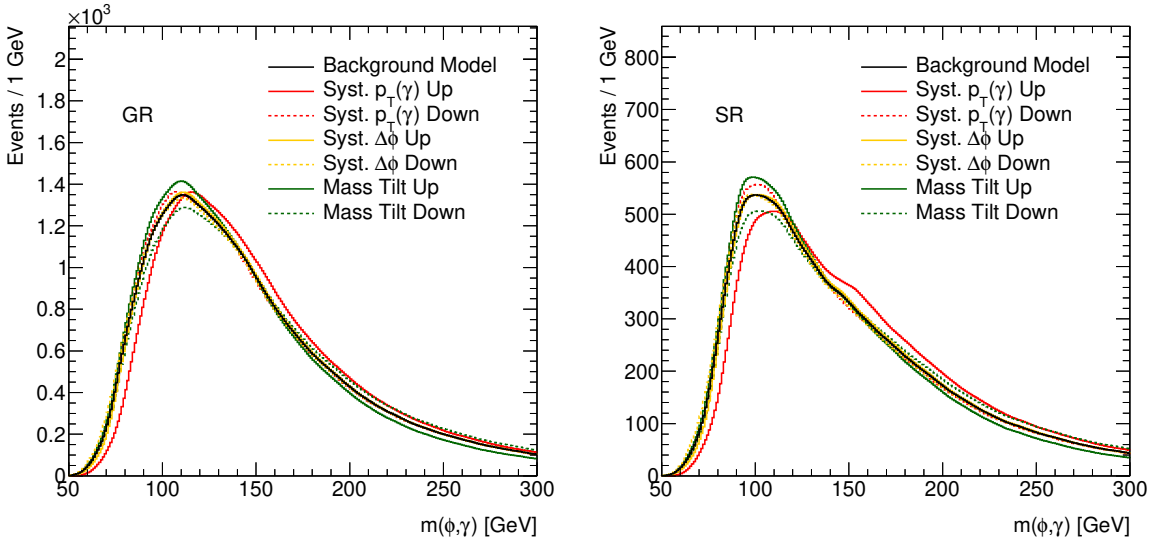
enough to allow the corresponding nuisance parameters to be constrained in the fit to the data. Furthermore, the nuisance parameters might be completely free or constrained by adding a Gaussian term to the likelihood.

"Up/down" pairs of approximately anti-symmetric shapes around the nominal background shape are generated either through modifications of the background modelling procedure or through direct transformations of the nominal shape. For the purposes of this demonstration, the following variations were used:

- A shift by A to the $p_{\mathrm{T}}(\gamma)$ of each generated pseudo-event, $p_{\mathrm{T}}(\gamma) = p_{\mathrm{T}}(\gamma) \pm A$[1].
- A multiplicative transformation of the $\Delta\Phi(\phi, \gamma)/\pi$ distribution by a function of the form $B + C \times \Delta\Phi(\phi, \gamma)$. The up (down) variations are obtained by choosing a positive (negative) value for the coefficient C.
- A direct multiplicative transformation of the $m(\phi, \gamma)$ distribution by a function of the form $D \times m(\phi, \gamma) + E$. The up (down) variations are obtained by choosing a positive (negative) value for the coefficient D.

[1]This variation is applied after the pseudo-event has been fully defined and as such it does not alter the values of the other ϕ and γ properties

Fig. 4.8 shows the effect of these three pairs of shape deformations on the $m(\phi,\gamma)$ distribution. The difference between each variation and the nominal shape is the pre-fit size of the corresponding variation. These variations are implemented in the fit (detailed in the following section) using the `RooFit` class `RooStarMomentMorph`, which allows to interpolate between the alternative shapes as described in Ref. [81]. Each pair of variations is controlled in the fit by a nuisance parameter, α. The up and down templates are mapped in the fit to α = 1 and α = -1, respectively. This means that the post-fit value of each parameter can be interpreted as how much the final shape varied from the nominal shape towards each of the alternative shapes: a negative α means the final shape is more "down-like", while a positive α value shows that it is more "up-like".



(a) (b)

Figure 4.8: Alternative $m(\phi,\gamma)$ distributions for the GR (a) and SR (b) obtained by performing modifications to the modelling procedure or by directly transforming the nominal invariant mass shape.

4.2.2.5 Implementation in statistical analysis

A binned maximum likelihood fit is performed to the $m(\phi,\gamma)$ distribution of the $\gamma+$jet sample in SR. The signal is modelled in the fit by the sum of two Gaussian functions with the same mean and different widths, which are fixed in the fit. The background PDF

results from the interpolation between the shape variations described in the previous section. Each of these variations is modelled by a PDF which is obtained from a linear interpolation of a finely binned histogram, which is the output of the background modelling. Each variation is controlled by an individual nuisance parameter: $\alpha(p_{\mathrm{T}}^{\gamma}\,\mathrm{shift})$, $\alpha(\Delta\Phi\,\mathrm{tilt})$ and $\alpha(\mathrm{mass\ tilt})$. The first two parameters are constrained in the fit, while $\alpha(\mathrm{mass\ tilt})$ is left free. The signal and background strengths are also free in the fit, controlling the normalisation of each contribution. μ_{signal} corresponds to 50 signal events and $\mu_{\mathrm{bkgd}} = 1$ corresponds to the background prediction in SR which is extrapolated from the GR.

Fig. 4.9 shows the result of the fit. The post-fit values of the parameters are listed in Table 4.3. As expected since no signal is present in the dataset being fitted, the post fit value of μ_{signal} is consistent with zero. The fact that the post-fit values of the shape nuisance parameter are non-zero shows the capability of the model to absorb residual background mismodelling. The uncertainties on these parameters show the statistical power of the data to constrain them relative to their pre-fit size.

Parameter	Value	Uncertainty ($\pm1\sigma$)
μ_{signal}	-0.07	±0.54
μ_{bkgd}	1.01	±0.01
$\alpha(p_{\mathrm{T}}^{\gamma}\,\mathrm{shift})$	0.26	±0.15
$\alpha(\Delta\Phi\,\mathrm{tilt})$	0.30	±0.43
$\alpha(\mathrm{mass\ tilt})$	0.10	±0.24

Table 4.3: Post-fit parameter values and corresponding uncertainties.

The impact of each systematic variation on the signal strength was quantified by fixing the corresponding parameters, one at a time, to the respective post-fit value listed in Table 4.3 post or minus the reported uncertainty, and repeating the fit. The resulting change in signal strength, normalised by $\sigma_{\mu_{\mathrm{signal}}}$ is reported for each case in Table 4.4.

4.2.3 Ensemble tests

The performance of the method is validated through an ensemble of independent toy experiments. For each experiment, an independent dataset is used to produce a sample of peudo-events using the background modelling procedure. A binned maximum

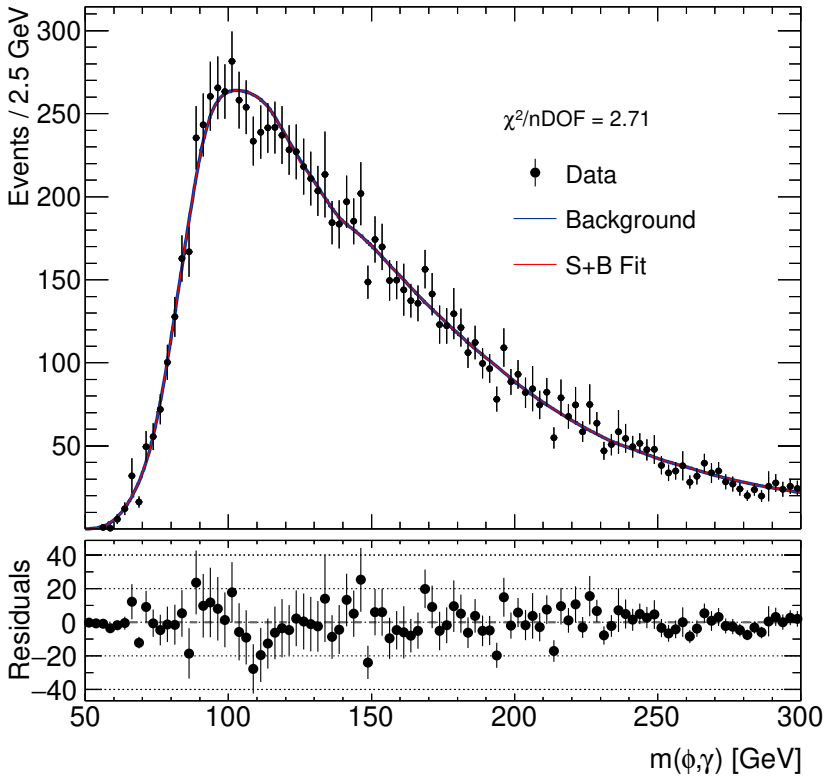


Figure 4.9: Result of a signal-plus-background binned maximum likelihood fit to the γ + jet MC sample using the described background modelling technique. The residuals panel shows the difference between the data and the post-fit signal-plus-background model.

		$\Delta\mu_{\text{signal}}/\sigma_{\mu_{\text{signal}}}$
$\alpha(p_{\text{T}}^{\gamma}\,\text{shift})$	$+1\sigma$	-0.18
	-1σ	+0.17
$\alpha(\Delta\Phi\,\text{tilt})$	$+1\sigma$	+0.07
	-1σ	-0.11
$\alpha(\text{mass tilt})$	$+1\sigma$	-0.22
	-1σ	+0.19

Table 4.4: Effect of the background modelling systematic variations on the signal strength.

likelihood fit is then performed to the corresponding original dataset, using the derived model as described in Section 4.2.2.5.

Since it would be computationally expensive to produce a large enough number of events for the ensemble tests in the manner described in Section 4.2.2.1, instead, an ensemble of synthetic datasets was generated through sampling from analytic functions. The analytic function used for the generation were chosen trying to reproduce the distributions of variables observed for the $\gamma +$ jet MC sample, used in the case study. Synthetic events were generated as follows:

1. A value for $m(\phi, \gamma)$ was sampled from a Landau function.

2. Given the value sampled in the previous step, a value for $p_{\mathrm{T}}(\phi, \gamma)$ was sampled from a Gaussian whose mean and width are functions of $m(\phi, \gamma)$.

3. A value for $\eta(\phi, \gamma)$ was sampled from the sum of two Gaussian functions, with the same width and opposite sign means ($\pm\mu$).

4. A value for $\Phi(\phi\gamma)$ was sampled from an uniform distribution defined in the $\{-\pi, \pi\}$ range. At this stage the synthetic Higgs four momentum vector is fully defined.

5. The synthetic Higgs candidates are decayed using `TGenPhaseSpace`, which performs phase space sampling. With this step, the ϕ and photon four momentum vectors are defined.

6. Depending on the value of $p_{\mathrm{T}}(\phi)$, a value for $I(\phi)$ is sampled from a linear function of $p_{\mathrm{T}}(\phi)$.

7. $\phi \to K^+K^-$ decays are simulated using `TGenPhaseSpace`. The K^+ and K^- four momentum vectors are defined at this stage.

An ensemble of 400 independent event samples was generated. The number of synthetic events in each sample was chosen in order to match the effective statistics of the $\gamma +$ jet MC sample at SR level. Each sample contains N events, with N being sampled from a Poisson distribution centred at 30000. The definitions of the generation,

validation and signal regions used can be found in Table 4.5. After applying the listed requirements, each sample has approximately 23000 events in the GR and 16000 events in the SR.

Region	Requirements
GR	p_{T}(leading track) $> 20\,\mathrm{GeV}$, p_{T}(subleading track) $> 15\,\mathrm{GeV}$
VR1	GR + $p_{\mathrm{T}}(\gamma) > 35\,\mathrm{GeV}$
VR2	GR + $p_{\mathrm{T}}(\phi) > 40\,\mathrm{GeV}$
VR3	GR + $\Delta\Phi(\phi,\gamma)/\pi > 0.5$
VR4	GR + $I(\phi) > 0.5$
SR	All of the above

Table 4.5: Definition of the regions employed in the background modelling for each synthetic background dataset.

The procedure to generate one pseudo-event in each toy experiment, represented in Fig. 4.10, is the following:

1. Values for $p_{\mathrm{T}}(\phi)$ and $p_{\mathrm{T}}(\gamma)$ are simultaneously sampled from a two-dimensional histogram.

2. $\Delta\eta(\phi,\gamma)$ is described in bins of $p_{\mathrm{T}}(\phi)$ and $p_{\mathrm{T}}(\gamma)$, in a three-dimensional histogram. Given the values of $p_{\mathrm{T}}(\phi)$ and $p_{\mathrm{T}}(\gamma)$ sampled in step 1, the template is projected along the $\Delta\eta(\phi,\gamma)$ dimension and a value for $\Delta\eta(\phi,\gamma)$ is sampled.

3. $I(\phi)$ is described in bins of and $p_{\mathrm{T}}(\gamma)$, in a three-dimensional histogram. Given the values of $p_{\mathrm{T}}(\phi)$ and $p_{\mathrm{T}}(\gamma)$ sampled in step 1, the template is projected along the $I(\phi)$ dimension and a value for $I(\phi)$ is sampled.

4. $\Delta\Phi(\phi,\gamma)$ is described in bins of $p_{\mathrm{T}}(\gamma)$ and $\Delta\eta(\phi,\gamma)$, in a three-dimensional histogram. Given the values sampled for $p_{\mathrm{T}}(\gamma)$ and $\Delta\eta(\phi,\gamma)$, the template is projected along the $\Delta\Phi(\phi,\gamma)$ dimension and a value for $\Delta\Phi(\phi,\gamma)$ is sampled.

5. Values for $\eta(\gamma)$ and $\Phi(\gamma)$ are sampled from the corresponding one-dimensional histograms. Imposing $m(\gamma) = 0$, the photon four-momentum of the pseudo-event becomes fully defined at this stage.

6. A value for $m(\phi)$ is sampled from the corresponding one-dimensional histogram. $\eta(\phi)$ and $\Phi(\phi)$ are calculated from the previously sampled values of $\eta(\gamma)$, $\Phi(\gamma)$,

$\Delta\Phi(\phi,\gamma)$ and $\Delta\eta(\phi,\gamma)$. The ϕ four-momentum of the pseudo-event is fully defined.

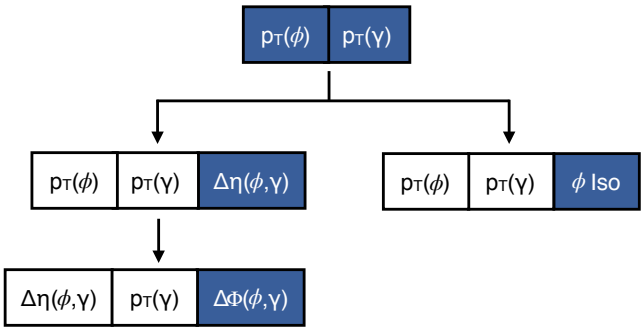
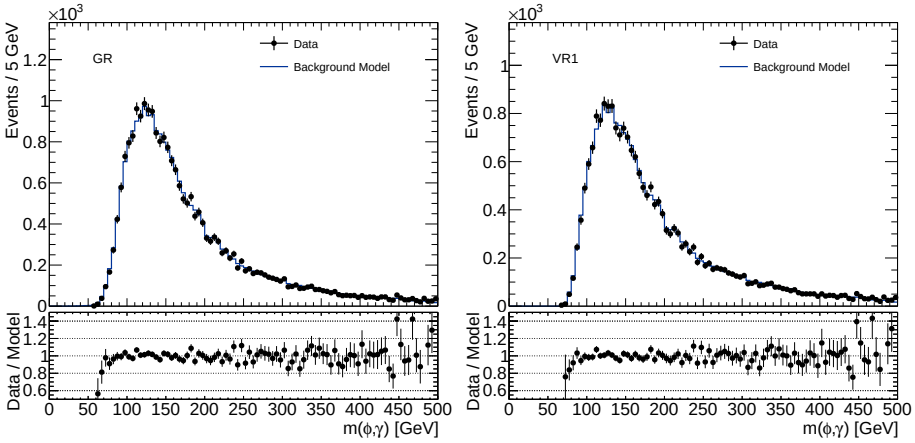


Figure 4.10: Schematic of the sampling sequence followed in the modelling of each independent dataset in the ensemble. Variables not included in the schematic are sampled in a factorised, uncorrelated, manner from an 1-dimensional template, as described in Table 4.2. Groups of two (three) variables represent two (three)-dimensional templates. Arrows are used to show the sequential order of the steps in the sampling. Variables are highlighted with color at the step in which they are defined for each pseudo-candidate.

By repeating the sequential sampling N times, the pseudo-events sample is generated. Fig. 4.11 shows the comparisons between the $m(\phi,\gamma)$ distributions of one of the synthetic datasets and the corresponding generated background model (built using the events in the GR directly), for all the regions defined in Table 4.5. Good agreement can be seen between the dataset and the model.

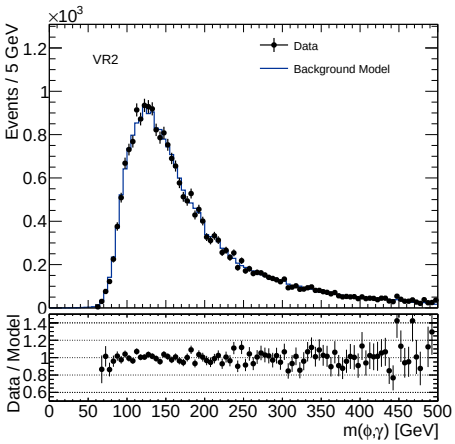
For each toy experiment, a fit was performed to the synthetic dataset as explained in Section 4.2.2.5. Two different means were considered for the double gaussian which is used to model the signal: 125 GeV, locating the signal on the peak of the $m(\phi,\gamma)$ distribution, and 150 GeV, locating the signal in a region in which the background is smoothly falling. These ensemble tests were performed both when the sampling is done directly from the GR, or after sampling with replacement. The distributions of the pull of the signal strength for the ensemble of toy experiments were plotted and can be found in Fig. 4.12. The pull is defined as

$$\text{pull} = \frac{\mu^{\text{fitted}}_{\text{signal}} - \mu^{\text{true}}_{\text{signal}}}{\sigma_{\text{signal}}}. \tag{4.1}$$

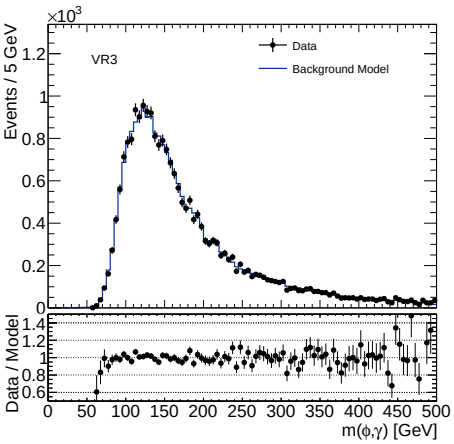


(a)

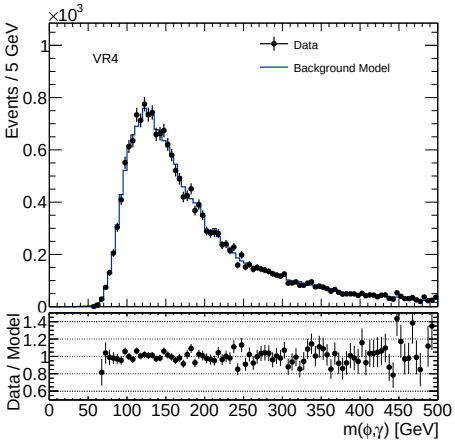
(b)



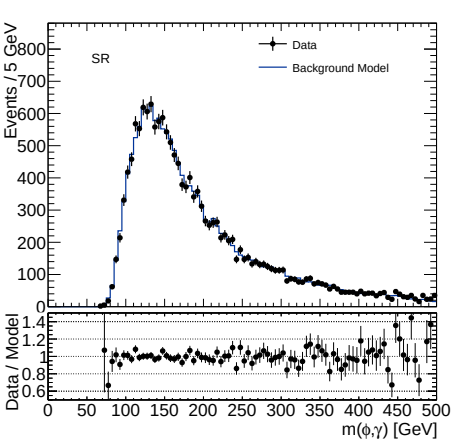
(c)



(d)



(e)



(f)

Figure 4.11: $m(\phi, \gamma)$ distributions in the GR, VR1, VR2, VR3, VR4 and SR for the synthetic dataset and background model built using the events in GR directly, for one of the toy experiments.

Since there is no signal present in each synthetic dataset, $\mu_{\text{signal}}^{\text{true}} = 0$. In all cases, the ensemble tests show that the modelling method provides a sufficiently accurate background description, with no substantial bias to the signal extraction. The maximum pull value observed was 0.18. One must note that assigning an uncertainty on the signal strength corresponding to 18% of its post-fit uncertainty would only increase the total uncertainty by 2%. One can also take as reference the level of bias on the signal extraction which is taken as acceptable in various ATLAS analyses when choosing the analytical form of background modelling shape. Functions with an associated bias up to 20% of the statistical uncertainty on the signal strength are considered [87, 90]. Nevertheless, the acceptable bias associated with the background model should be evaluated on a case-by-case basis, depending on the level of precision of the analysis being performed.

For the ensembles in which the background modelling was performed by sampling directly N events from the GR, the standard deviation of the pull distributions are significantly below one, suggesting that the uncertainty on the signal strength is overestimated. This effect arises because the events in the GR are not independent from the events in the SR, since the SR is subset of the GR. Fig. 4.13a shows the comparison between the synthetic dataset and the generated background model, for two different toy experiments. It can be seen that the background model reproduces the specific fluctuations of each dataset to a certain extent. The impact of this effect depends on the amount of overlap between the events in the SR and the events in the GR, N_{SR}/N_{GR}. The ensemble test were repeated for different values of N_{SR}/N_{GR} and it was verified that when N_{SR}/N_{GR} tends to zero the width of the pull distribution tends to unity, as is shown in Fig. 4.13b. Published analyses that previously used this method [82–84, 100, 106] have a N_{SR}/N_{GR} ratio of approximately 0.1, therefore this effect is negligible. Nevertheless, there might be cases in which it might not be possible to achieve a small GR-to-SR efficiency, due to experimental constraints. In those cases, the model can be generated after sampling with replacement from the GR, which effectively removes the correlations between the original dataset and the background model; alternatively, the signal strength uncertainty can be corrected according to ensemble test studies. Fig. 4.13b shows that for all N_{SR}/N_{GR} considered, the pull distribution of the ensemble

using sampling with replacement has a width significantly closer to unity, comparatively to the experiments in which each event in the GR is only used once.

4.3 Conditional generative adversarial networks

As discussed in the previous sections, the ancestral sampling technique provides an accurate description of a background, in a non-parametric, data-driven way. However, in the case of a search relying on multivariate analysis techniques for the discrimination between signal and background, it would become impractical to use the ancestral sampling approach. Furthermore, even though the method is robust under signal contamination, as was shown in Section 4.2.2.3, no signal region blinding strategy was proposed. Trying to address these issues, a generalisation of the method was developed relying on generative adversarial networks (GANs) [107] trained on data to produce the background model. GANs are machine learning models consisting of a pair of neural networks, a generator and a discriminator. The generator network is trained to generate simulated data by learning the features of an input dataset, while the discriminator network is simultaneously trained to differentiate the input dataset from the generator's output. This is represented in Fig. 4.14. The subject of GANs and their applications to high energy physics has been gaining popularity in the last years, given their capability to generate large samples significantly faster then other techniques [108, 109].

In order to address the possible presence of a signal contribution in the training dataset, this implementation of the background modelling technique was devised to allow the signal region to be blinded during training. For this purpose, a conditional generative adversarial network (cGAN) [110] was used. When trained on a dataset, a cGAN learns a generative model of the conditional probability density distributions of the features of the data, given the value of a variable. In this case, this conditioning variable is chosen to be the blinding variable, for which the window containing the signal is kept blinded during training. If one inputs the full distribution of the conditioning variable after the training, the cGAN is able to extrapolate its prediction of the background shape from the sidebands to the blinded signal region. Fig. 4.15 shows a schematic of a cGAN.

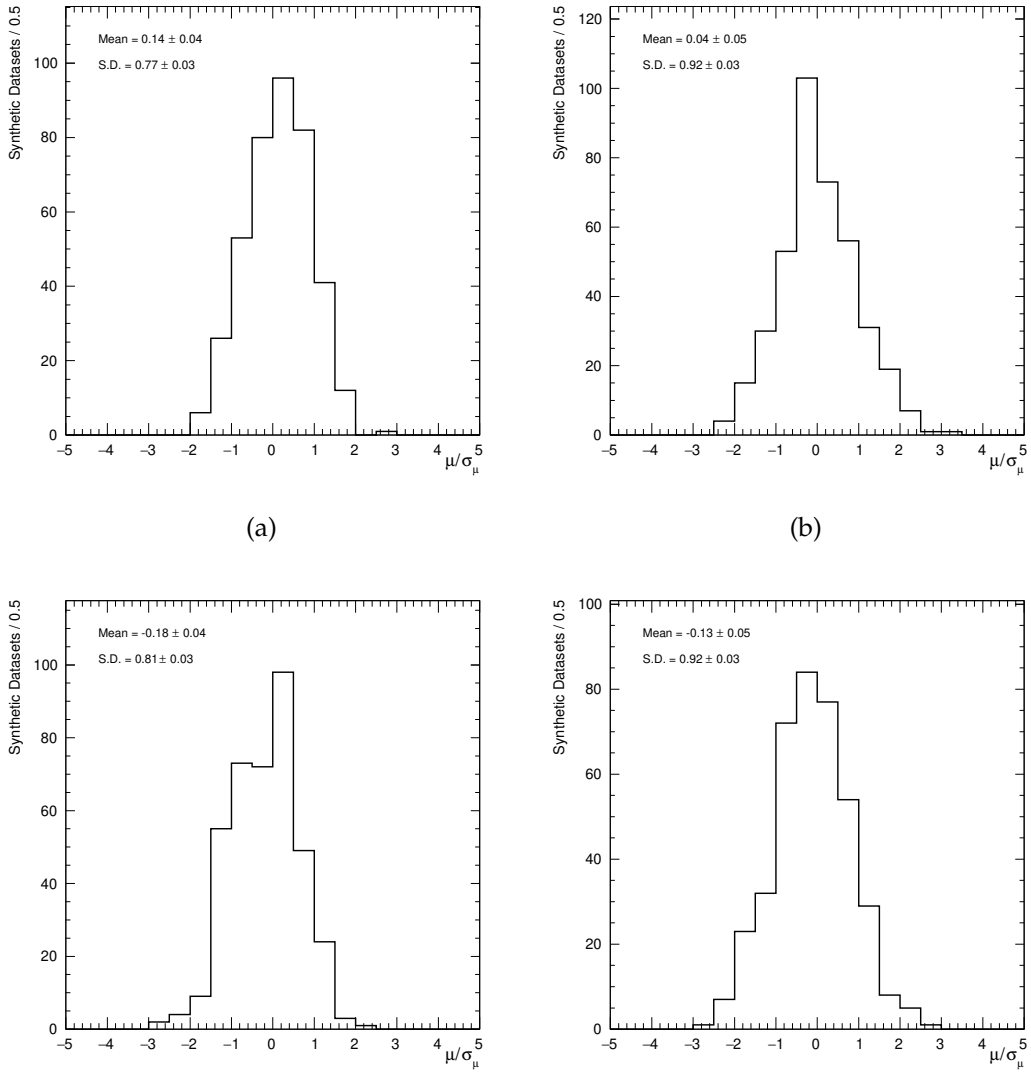
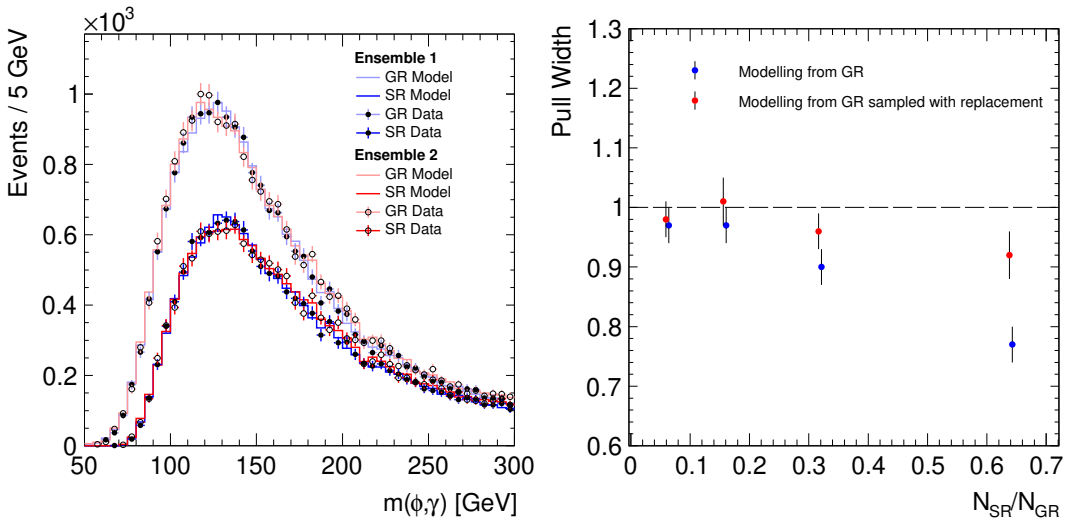


Figure 4.12: Distributions of the signal strength post-fit values for the different experiments in the ensemble, normalised to the corresponding uncertainty, for two mass hypotheses: for Fig. 4.12a and Fig. 4.12b the mean of the signal distribution is 125 GeV; for Fig. 4.12c and Fig. 4.12d the mean of the signal distribution is 150 GeV In Fig. 4.12a and Fig. 4.12c the background model is obtained by sampling the GR directly, while in Fig. 4.12b and Fig. 4.12d sampling with replacement is used. The mean and standard deviation of each distribution is reported in the plots.



(a) (b)

Figure 4.13: (a) Synthetic dataset and background model comparison for two distinct experiments in the ensemble. The sampling from GR utilises each of the N events once. (b) The width of the pull distribution for ensemble tests performed using different N_{SR}/N_{GR}. A shift of -0.5% along the horizontal axis was applied to the red markers, to improve visibility.

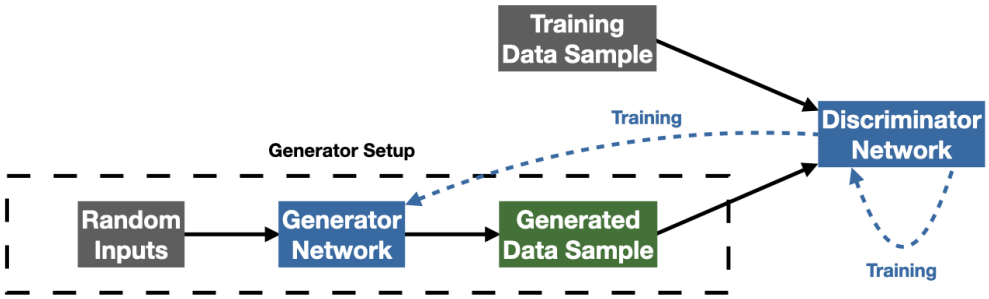


Figure 4.14: Schematic of a generative adversarial network [4].

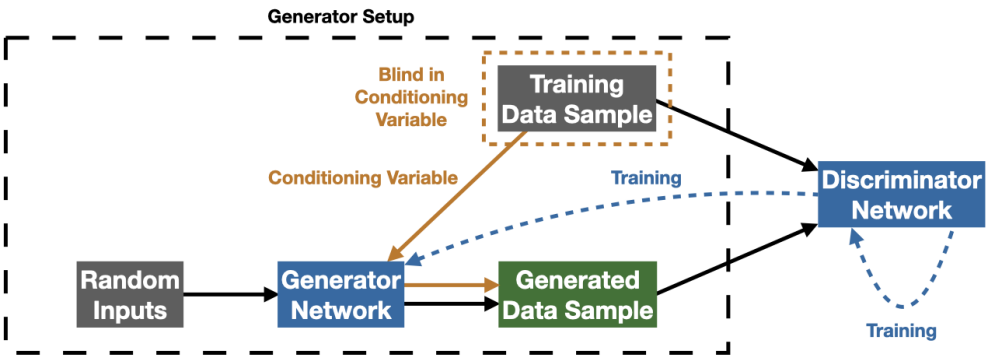


Figure 4.15: Schematic of a conditional generative adversarial network [4].

4.3.1 Case study: search for $H \to Za \to \mu\mu + \text{jet}$

Light pseudo-scalar particles, which can be produced in decays of the Higgs boson, are predicted by several beyond the SM theories [111]. A search for decays of the Higgs boson into a Z boson and an hadronically decaying light scalar, a, has ben recently published by ATLAS [112]. The a particle was reconstructed as a single jet, while the Z boson was reconstructed in its leptonic decays ($Z \to \mu^+\mu^-$ and $Z \to e^+e^-$).

The main background of this search arises from $Z + \text{jets}$ events. The analysis used a modified ABCD method, including a MC-based correction, to estimate this background. Furthermore, a multilayer perceptron (MLP) [113] classifier was employed for signal/background discrimination. The MLP was trained on jet substructure variables, which included: the ΔR between the leading track and the jet axis ($\Delta R_{\text{LeadTrack}}$); the ratio between the p_{T} of the leading track and the sum of the p_{T} of all tracks in the jet ($p_{\text{T}_{\text{LeadTrack}}}/p_{\text{T}_{\text{LeadTrack}}}$); angularity [114]; and U_1, a modified energy correlation function [115]. The use of MVA techniques, allied to the fact that statistical and theoretical MC uncertainties were reported as the dominant uncertainties for the analysis, make this an ideal case study to demonstrate the application of the background modelling method using cGANs.

In the following section, a summary of the application of the method to a simplified version of the $H \to Za \to \mu\mu + \text{jet}$ search is presented. These studies are presented in full detail in [4]. For the purposes of this demonstration, only the $Z \to \mu^+\mu^-$ decays are considered in the analysis. A sample of $Z + \text{jets}$ simulated events is treated as the "data" to be modelled. The $Z + \text{jets}$ sample was generated using Pythia 8.244 with the CT14nlo PDF set. The detector response was simulated using Delphes, as described in Section 4.2.2.1.

4.3.2 Application to case study

The cGAN is trained on the $Z + \text{jets}$ MC sample. The chosen conditioning variable is the invariant mass $m_{\mu\mu\text{j}}$, which is blinded for training in the window $123 < m_{\mu\mu\text{j}} < 135\,\text{GeV}$. The variables included in the training of the cGAN are the four jet substructure variables: $\Delta R_{\text{LeadTrack}}$, $p_{\text{T}_{\text{LeadTrack}}}/p_{\text{T}_{\text{LeadTrack}}}$, angularity and U_1. These are the inputs to the

MLP used in the selection. Before training, the requirement on the MLP classifier is removed, ensuring that the training dataset is dominated by background events. The cGAN learns the generative model of the probability density distributions of the jet substructure variables in data, given the value of $m_{\mu\mu j}$. After training, the full inclusive distribution of the invariant mass in data (without the MLP-based selection) is provided to the cGAN, and the cGAN extrapolates the background prediction for the jet substructure variables from the $m_{\mu\mu j}$ sidebands into the blinded $m_{\mu\mu j}$ region. In this way, a sample of background pseudo-events is generated, each one defined by ($m_{\mu\mu j}$, $\Delta R_{\text{LeadTrack}}$, $p_{\text{T}_{\text{LeadTrack}}}/p_{\text{T}_{\text{LeadTrack}}}$, angularity, U_1). From the values of the jet substructure variables, the output of the MLP classifier can be calculated for each pseudo-event. The requirement on this output is then applied to the sample of pseudo-events. The shape of the background on $m_{\mu\mu j}$ in SR can be obtained in this way, and can be used to model the background in a fit to the data.

In order to reduce the uncertainty due to the training, an ensemble of 100 cGANs is used. The performance of each cGAN is evaluated using a 2D χ^2 metric, based on the difference between the distribution of the data and of the background prediction in $m_{\mu\mu j}$ and the MLP output. 25% of the events of the Z + jets sample are not used in the training, and are kept as the validation dataset, for which the χ^2 is calculated. Ultimately, the background prediction is obtained from the average of the top 5 cGANs. Uncertainties associated the background modelling technique can then be derived from the differences between each individual cGAN and the average.

The performance of the background modelling technique is verified in Fig. 4.16 and Fig. 4.17. Fig. 4.16 shows the distribution of the jet variables for the Z + jets sample and the pseudo-events sample, for events in the blinded invariant mass region. The good agreement that is seen proves the ability of the cGAN to extrapolate the background prediction from the sidebands. Fig. 4.17 shows these same distributions but for event in the sidebands. The low and high mass sidebands are shown separately. The fact that the cGAN can provide a good description for both sidebands shows that the cGAN is able to learn the dependence of the jet variables on $m_{\mu\mu j}$.

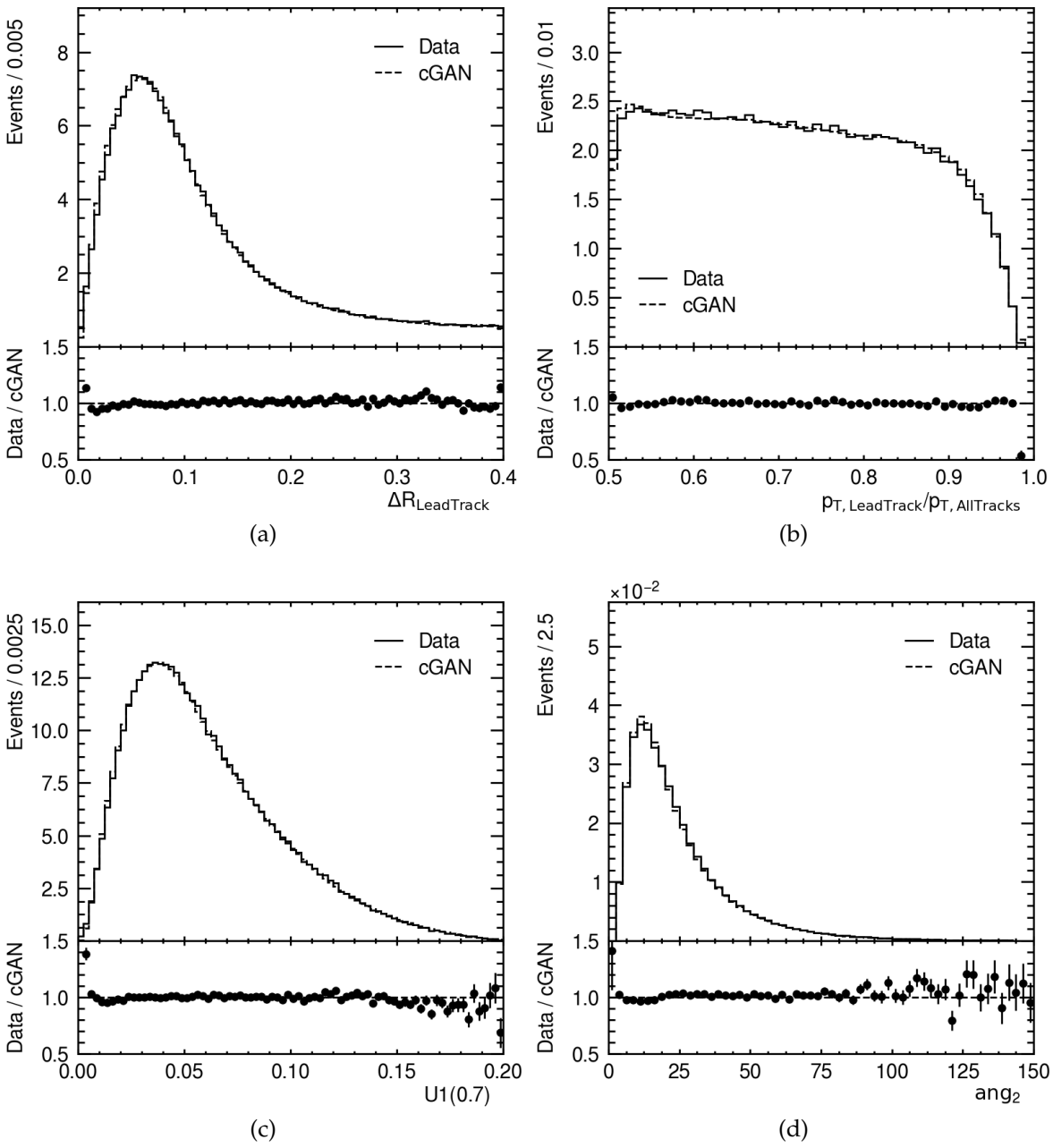


Figure 4.16: Distributions of the jet substructure variables for the Z + jets MC sample and for the background model, for events with $123 < m_{\mu\mu\text{j}} < 135$ GeV.

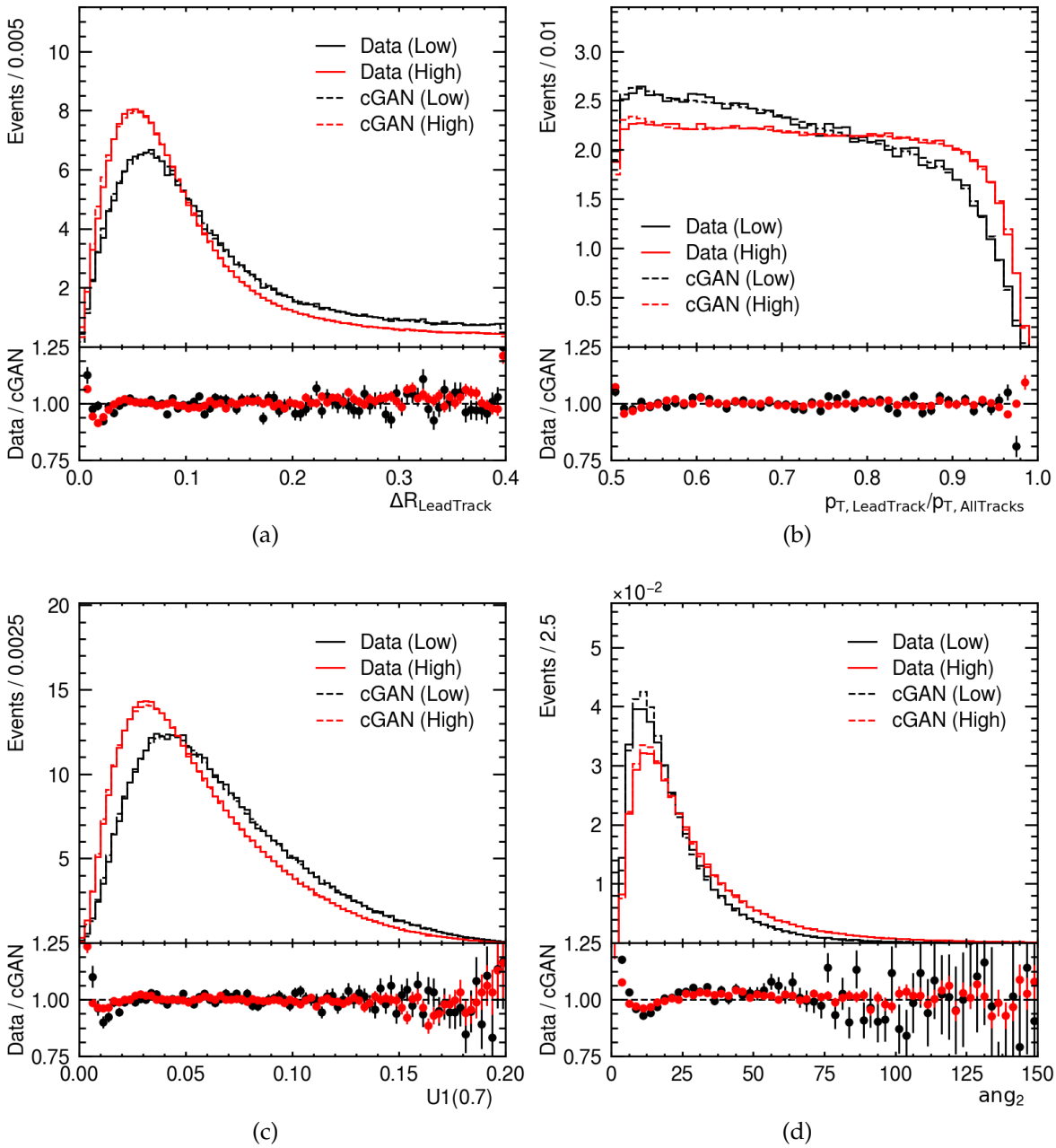


Figure 4.17: Distributions of the jet substructure variables for the Z + jets MC sample and for the background model. The distributions are shown for the low and high $m_{\mu\mu\mathrm{j}}$ sidebands separately.

Chapter 5

Search for exclusive hadronic decays of the W boson

This chapter describes the search for exclusive hadronic decays of the W boson with the ATLAS experiment using Run 2 LHC data, which has been recently submitted to PRL and made public in Ref. [116]. Three decays were targeted: $W^{\pm} \to \pi^{\pm}\gamma$, $W^{\pm} \to \rho^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$. As listed in Table 1.1, the world's best limit on the branching fraction of $W^{\pm} \to \pi^{\pm}\gamma$ is $< 7 \times 10^{-6}$ at 95% CL [34], which corresponds to 1750$\times$ SM value [28]. No experimental constraints had been set on B($W^{\pm} \to \rho^{\pm}\gamma$) and B($W^{\pm} \to K^{\pm}\gamma$) before this search was performed.

The decay channels in question have distinct experimental signatures: in the case of $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ one isolated high p_{T} track recoils against an isolated high p_{T} photon; in the case of the $W^{\pm} \to \rho^{\pm}\gamma$, the $\rho^{\pm}$ meson decays into $\pi^{\pm}\pi^{0}$ within the detector volume, resulting in an additional electromagnetic energy deposition due to the neutral pion. The $\pi^{\pm}$ and $K^{\pm}$ in the final state cannot be distinguished experimentally by the ATLAS detector, due to its limited particle identification capabilities for high momentum hadrons. The analysis employed two different reconstruction strategies for the meson in the decays, which define two different final states. These will be referred to throughout this chapter as the track-photon and tau-photon final states. The track-photon final state was optimised to target the $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ decay channels, and reconstructs the meson as a track. The second final state, tau-photon, was devised to target the $W^{\pm} \to \rho^{\pm}\gamma$ decay channel and reconstructs the $\rho^{\pm}$ as a tau object with exactly one associated charged track and a neutral pion, taking advantage of the similarities between the $\rho^{\pm}$ decay and hadronic τ-lepton decays. A different

selection was applied to events reconstructed in the two different final states, with the largest difference being the trigger selection. Dedicated triggers requiring a high p_{T} track and a high p_{T} photon were employed in the selection of the events reconstructed in the track-photon final state, while triggers requiring two photons were used to select events reconstructed in the tau-photon final state. These two trigger strategies were found to be practically orthogonal, i.e. to select different events. Even though the event selection used in track-photon final state was defined to optimise the sensitivity to the $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ decay channels, this final state is still sensitive to the $W^{\pm} \to \rho^{\pm}\gamma$ decay. The orthogonality between the triggers used in the two final states allowed the combination of the two data event samples in a simultaneous fit to the W boson invariant mass distribution in data, increasing the overall sensitivity of the analysis to the $W^{\pm} \to \rho^{\pm}\gamma$ decay.

Throughout this chapter the analysis strategy employed for each of the two final states will be described, with more detail being given to the track-photon final state analysis which corresponds to the author's main contribution. The statistical analysis in both final states was fully performed by the author and as such will be presented in full. The analysis employs the background modelling technique presented in the previous chapter, in both final states.

5.1 Data and Monte Carlo simulation samples

5.1.1 Data samples

A data sample collected during Run 2 of the LHC at $\sqrt{s} = 13$ TeV is used in the analysis. Only events included in the GRL are considered for the analysis. The dedicated triggers used in the track-photon final state selection only became active from 2016 onwards, and as such, data from 2015 is not analysed in this particular final state. The data collected by the dedicated triggers corresponds to 136.8 fb^{-1}. The full Run 2 dataset is instead used for the tau-photon final state analysis, corresponding to 140 fb^{-1}.

Due to the differences in reconstruction and selection between the two final states,

two derivations were employed: HDBS2 for track-photon and STDM14 for tau-photon[1]. To be retained by the HDBS2 derivation an event must:

- have been recorded by at least one of the dedicated track-photon triggers;
- have at least one photon with $p_T > 30$ GeV and $|\eta| < 2.5$;
- have at least one track with $p_T > 30$ GeV;
- have at least one W candidate with $\Delta\Phi(\text{track}, \gamma) > 0.766$;
- have at least one W candidate with m(track, photon) > 55 GeV.

Instead, the STDM14 derivation relies on di-photon trigger skimming and requires events to have at least one reconstructed photon and one hadronic tau object (τ_{had}).

5.1.2 Signal Monte Carlo samples

$W^{\pm} \to \pi^{\pm}\gamma$, $W^{\pm} \to K^{\pm}\gamma$ and $W^{\pm} \to \rho^{\pm}\gamma$ Monte Carlo (MC) samples were generated with POWHEG-BOXv1 [117] at next-to-leading order (NLO), using CT10 parton distribution functions [118]. $W^{\pm} \to \mu^{\pm}\nu$ events are generated initially; $W^{\pm} \to \pi^{\pm}\gamma$, $W^{\pm} \to K^{\pm}\gamma$ and $W^{\pm} \to \rho^{\pm}\gamma$ events are generated by editing the Les Houches Event files (LHE) [119] on-the-fly and by replacing the muon and neutrino in each event by a meson and a photon. The modelling of the parton shower, hadronisation and underlying event is performed using the PYTHIA 8.243 [120] event generator in the case of the $W^{\pm} \to \pi^{\pm}\gamma$ sample and PYTHIA 8.244 in the case of $W^{\pm} \to K^{\pm}\gamma$ and $W^{\pm} \to \rho^{\pm}\gamma$ samples. The CTEQ6L1 PDF set [121] configured following the AZNLO tune [122] was used. The ATLAS detector response is simulated with GEANT4 [123].

The W^+ and W^- samples are added, weighted by their respective cross-section. The cross sections measured by the ATLAS experiment are taken from [25]: $\sigma_{W^+} \times \mathcal{B}(W \to l\nu) = 11.83 \pm 0.41$ nb and $\sigma_{W^-} \times \mathcal{B}(W \to l\nu) = 8.79 \pm 0.30$ nb, with $\mathcal{B}(W \to \ell\nu) = 0.1086$ [15].

Due to different detector conditions, different samples with different configurations are produced corresponding to each year of Run 2 of the LHC: MC16a for 2016, MC16d

[1] HDBS and STDM stand for Higgs and Diboson Searches and Standard Model, respectively.

for 2017 and MC16e for 2018. The number of events generated in each signal sample can be found in Table 5.1. The samples from the different MC campaigns are weighted according to the total integrated luminosity for the respective year.

	MC16a	MC16d	MC16e
$W^+ \to \pi^+\gamma$	25k	35k	45k
$W^- \to \pi^-\gamma$	25k	35k	45k
$W^+ \to K^+\gamma$	20k	20k	30k
$W^- \to K^-\gamma$	20k	20k	29k
$W^+ \to \rho^+\gamma$	30k	40k	50k
$W^- \to \rho^-\gamma$	30k	40k	50k

Table 5.1: Number of generated events for each simulated decay process.

MC events are re-weighted to match the distribution of the average number of interactions per bunch crossing (pile-up) observed in data. At reconstructed level, scale factors (SF) are applied to account for the mismatch between data and simulation on identification, isolation and trigger efficiencies. Additionally, each event is generated with an associated weight by the MC generator. In summary, each MC event is weighted by a weight w, as follows:

$$w = \frac{w_i^{\mathrm{MC}} \times w_i^{\mathrm{PU}} \times w_i^{\mathrm{Pol}} \times \mathrm{SF} \times \mathcal{L} \times \sigma_W}{\sum_i^{\mathrm{N}} w_i^{\mathrm{MC}}}, \tag{5.1}$$

in which w_i^{MC} is the event generator weight, w_i^{PU} the pileup correction and N the total number of events in the sample. w_i^{Pol} is a weight associated with the polarisation of each of the signal decays, discussed in the following subsection.

5.1.2.1 Polarisation reweighting

The signal samples are generated isotropically and reweighted to account for the predicted angular distribution of the decay products. In order to perform such calculations, one must first choose a reference frame and a polarisation axis, z. Several definitions of the polarisation axis can be found in literature. In the case of a collider experiment, they can be defined in relation to the production plane, which is the plane formed by the momenta of the two colliding beams in the decaying particle's rest frame. Two possible choices of frame, both with the z axis belonging to the production plane are the helicity

frame (HX) and the Collins-Soper frame (CS). In the case of the helicity frame, the z axis is chosen to be in the flight direction of the decaying particle, in the center-of-mass of the colliding beams (the laboratory frame). The Collins-Sopper z axis is instead defined in the direction of the bisector between the momentum of one colliding beam (h1) and the opposite of the other (h2), in the particle's rest frame. Fig. 5.1 illustrates the different possible polarisation axes. For both the helicity and Collins-Sopper frame the y axis is in the direction perpendicular to the production plane. The x axis is the cross product of the defined y and z axes.

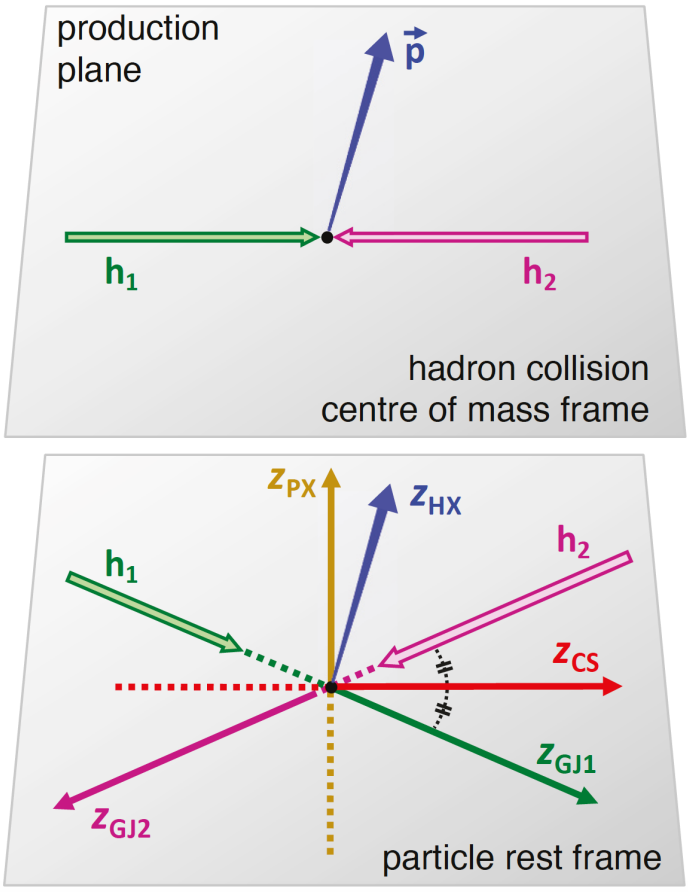


Figure 5.1: Graphical representation of the possible choices of polarisation axis with respect to the momenta of the colliding hadrons h1 and h2 [124].

Two body decays of the form $p_1 \to p_2 p_3$, such as the $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$

decays, have an angular distribution that follows the expression from [125]

$$I(\theta) \sim \sum_{\lambda_1\lambda_2\lambda_3} |d^{s_1}_{\lambda_1,\lambda_2-\lambda_3}|^2 |A_{\lambda_2\lambda_3}|^2, \tag{5.2}$$

where d are the Wigner D-matrix elements, λ the helicities, s the spin of the particle in question, and A the helicity amplitudes. θ is defined with respect to the helicity polarisation frame: it is the angle between the direction of particle 2 in the rest frame of particle 1 and the direction of particle 1 in the lab frame. In the case of the $W^\pm \to \pi^\pm\gamma$ and $W^\pm \to K^\pm\gamma$ decays, the helicity of the photon (particle 3) is $\lambda_3 = \pm 1$, since the photon can only be transversely polarised and the pion/kaon (particle 2) is a spin 0 particle which has $\lambda_2 = 0$. The W boson is a spin 1 particle which can have $\lambda_1 =$ -1,0,1. Nevertheless, to conserve helicity in the decay, in this case only $\lambda_1 = \pm 1$ are allowed. As such, the terms contributing to the angular distribution of $W^\pm \to \pi^\pm\gamma$ and $W^\pm \to K^\pm\gamma$ are $d^1_{1,1} = d^1_{-1,-1} = \frac{1+\cos\theta}{2}$, $d^1_{-1,1} = d^1_{1,-1} = \frac{1-\cos\theta}{2}$, and the resulting angular distribution is

$$I(\theta) \sim (1 + \cos^2\theta). \tag{5.3}$$

This expression is further normalised such that

$$\int_{-1}^{1} I(\theta)\, d\cos\theta = 1. \tag{5.4}$$

The effect of the polarisation reweighting on the kinematics of the decay products can be seen in Figs. 5.2 to 5.4, for the $W^\pm \to \pi^\pm\gamma$ MC events. These are generator level (truth) distributions, before any detector simulation is performed. Fig. 5.4 shows the $\cos\theta$ distribution and serves as a sanity check, one can see the initial isotropic distribution and the final angular distribution after the correction. The largest observed effect from the reweighting is on the p_T distributions of the $\pi^\pm$ and the γ, which become softer. This means that the acceptance to the $W^\pm \to \pi^\pm\gamma$ decays is lower than if the decays were isotropic. $W^\pm \to \pi^\pm\gamma$ and $W^\pm \to K^\pm\gamma$ have the same angular distribution, and as such the same effect is observed after the polarisation reweighting. A comparison of the generator level distributions of the kinematic variables for $W^\pm \to \pi^\pm\gamma$ and $W^\pm \to K^\pm\gamma$ can be found in Figs. 5.5 to 5.6. As expected, the $W^\pm \to \pi^\pm\gamma$ and

$W^{\pm} \to K^{\pm}\gamma$ exhibit the same kinematic behaviour.

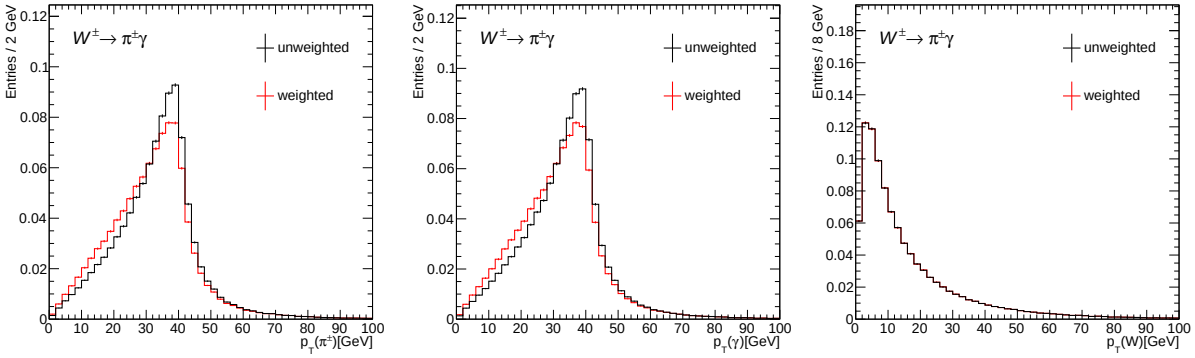


Figure 5.2: Generator level p_T distributions of the $\pi^{\pm}$, γ and W, for $W^{\pm} \to \pi^{\pm}\gamma$, with and without polarisation. The distributions are normalised to unity.

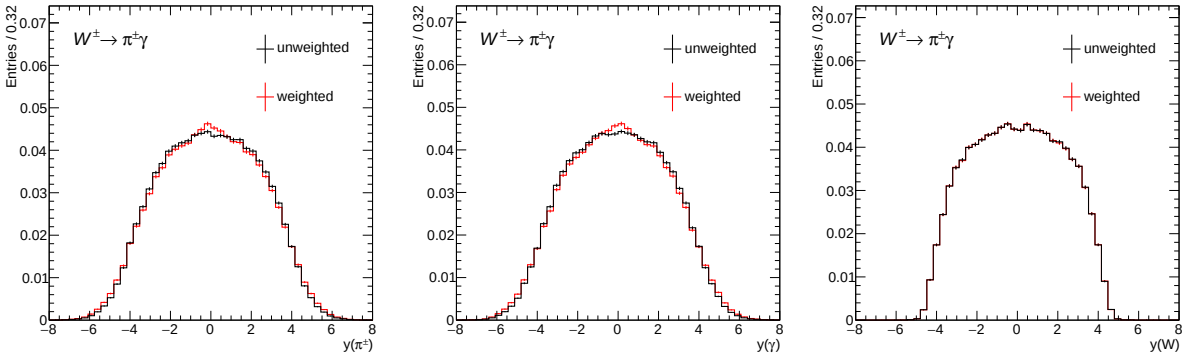


Figure 5.3: Generator level rapidity distributions of the $\pi^{\pm}$, γ and W, for $W^{\pm} \to \pi^{\pm}\gamma$, with and without polarisation. The distributions are normalised to unity.

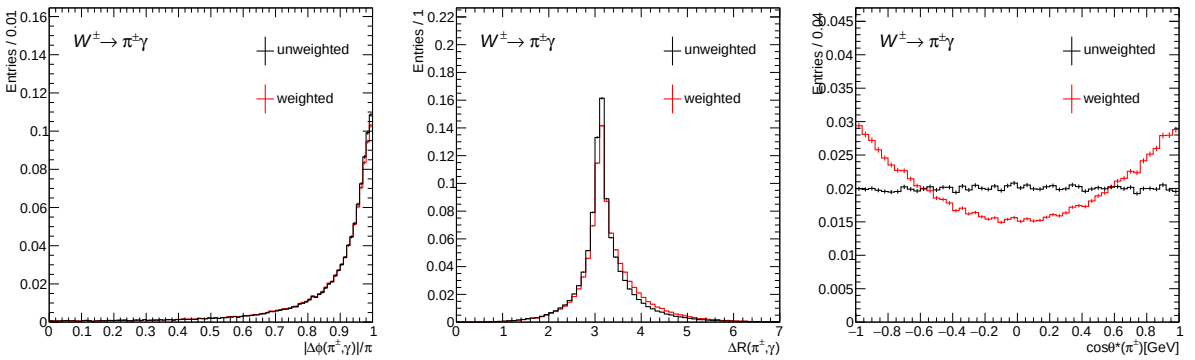


Figure 5.4: Generator level $\Delta\phi(\pi^{\pm},\gamma)$, $\Delta R(\pi^{\pm},\gamma)$ and $\cos\theta$ distributions, for $W^{\pm} \to \pi^{\pm}\gamma$, with and without polarisation. The distributions are normalised to unity.

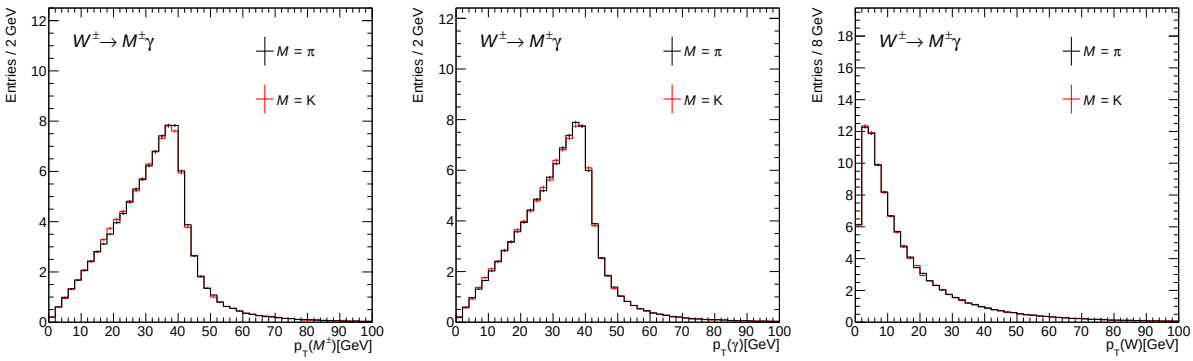


Figure 5.5: Comparison between the generator level p_T distributions of the $M^{\pm}$, γ and W, for $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$. The $W^{\pm} \to K^{\pm}\gamma$ distribution is normalised to the integral of the $W^{\pm} \to \pi^{\pm}\gamma$ distribution.

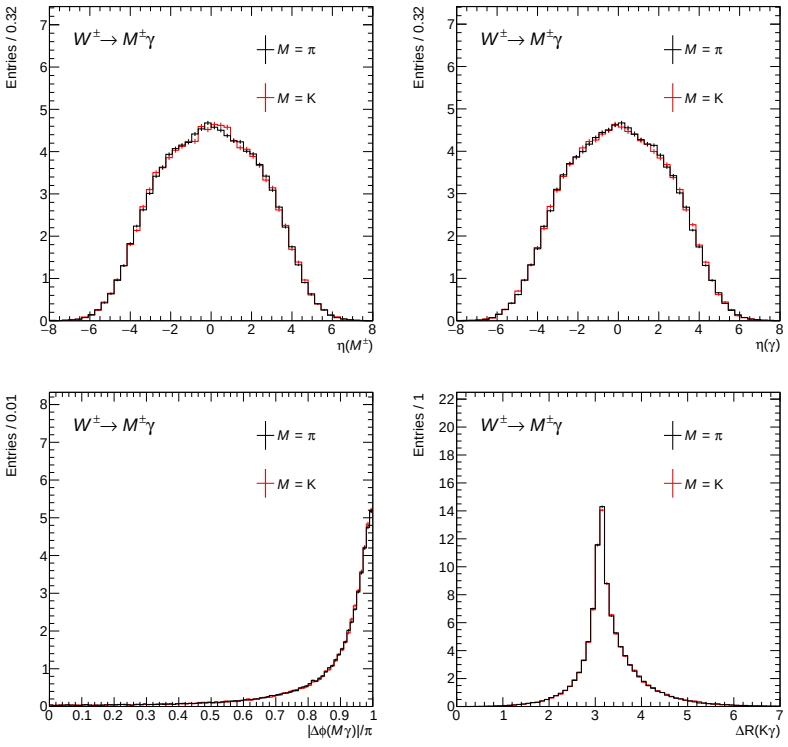


Figure 5.6: Comparison between the generator level distributions of $\eta(\gamma)$, $\eta(M)$, $\Delta\phi(M,\gamma)$ and $\Delta R(M,\gamma)$, for $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$. The $W^{\pm} \to K^{\pm}\gamma$ distribution is normalised to the integral of the $W^{\pm} \to \pi^{\pm}\gamma$ distribution.

In the case of the $W^{\pm} \to \rho^{\pm}\gamma$ there are two sequential decays, and so one needs to consider not only the direction of the $\rho^{\pm}$ meson, but also the direction of the $\pi^{\pm}$ after the second decay. The angular distribution used was defined using the "cloned cascade frame" formalism [124]. In this formalism, the first decay $p_1 \to p_2 p_3$ is described using one of the possible polarisation axes (for example HX or CS), and the second decay ($p_2 \to p_4 p_5$) is described using the geometrical "clone" of the first frame, in the rest frame of 2. Fig. 5.7 illustrates this choice of frame in the case of a cascade process of the form $O \to V + X, V \to l^+ l^-$. The Θ and ϕ angles are defined according to (x, y, z) in the rest frame of O, and the ϑ and φ angles are defined according to (x'', y'', z''), which are exact geometrical clones of (x, y, z) in the rest frame of V.

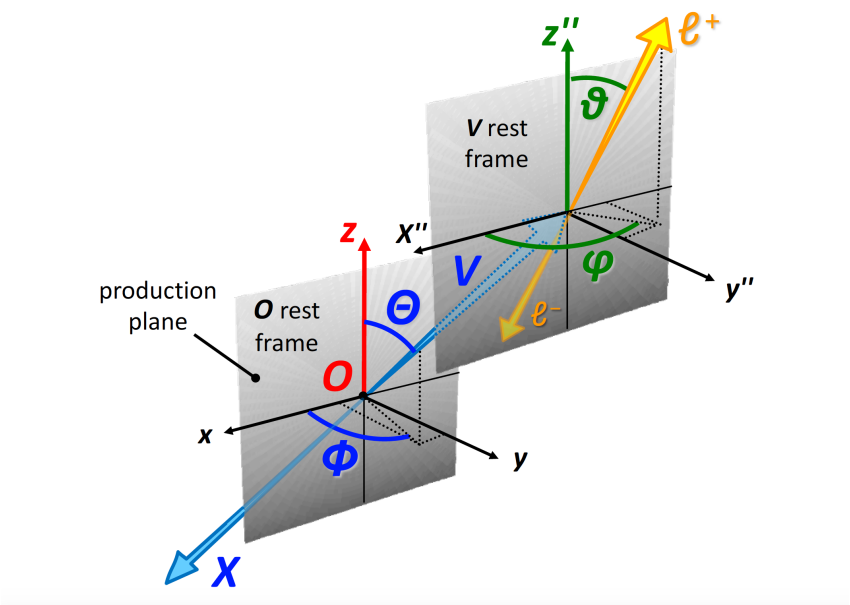


Figure 5.7: Representation of the cloned cascade frame, for a cascade process of the form $O \to V + X, V \to l^+ l^-$ [124].

The angular distribution of $W^{\pm} \to \rho^{\pm}\gamma$ is

$$I(\Theta, \vartheta, \phi, \varphi) \sim 1 - \cos^2\Theta + 2\cos^2\Theta\cos^2\vartheta + \frac{1}{2}\sin 2\Theta \sin 2\vartheta \cos(\phi - \varphi)[124]. \quad (5.5)$$

The Θ and ϕ angles are defined in the W boson Collins-Soper frame. This expression is an approximation of a more detailed expression, also provided in Ref. [124]. The approximation is valid for cases in which the decaying W boson has a p_T which is

significantly smaller than its mass, as is the case for this analysis (as seen in Fig. 5.2). The $\cos\theta$, $\cos\vartheta$ and p_{T} distributions of the decay products before and after reweighting to the correct angular distribution can be found in Fig. 5.8. As in the $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ case, the correct acceptance is lower than it would be if the decays were isotropic.

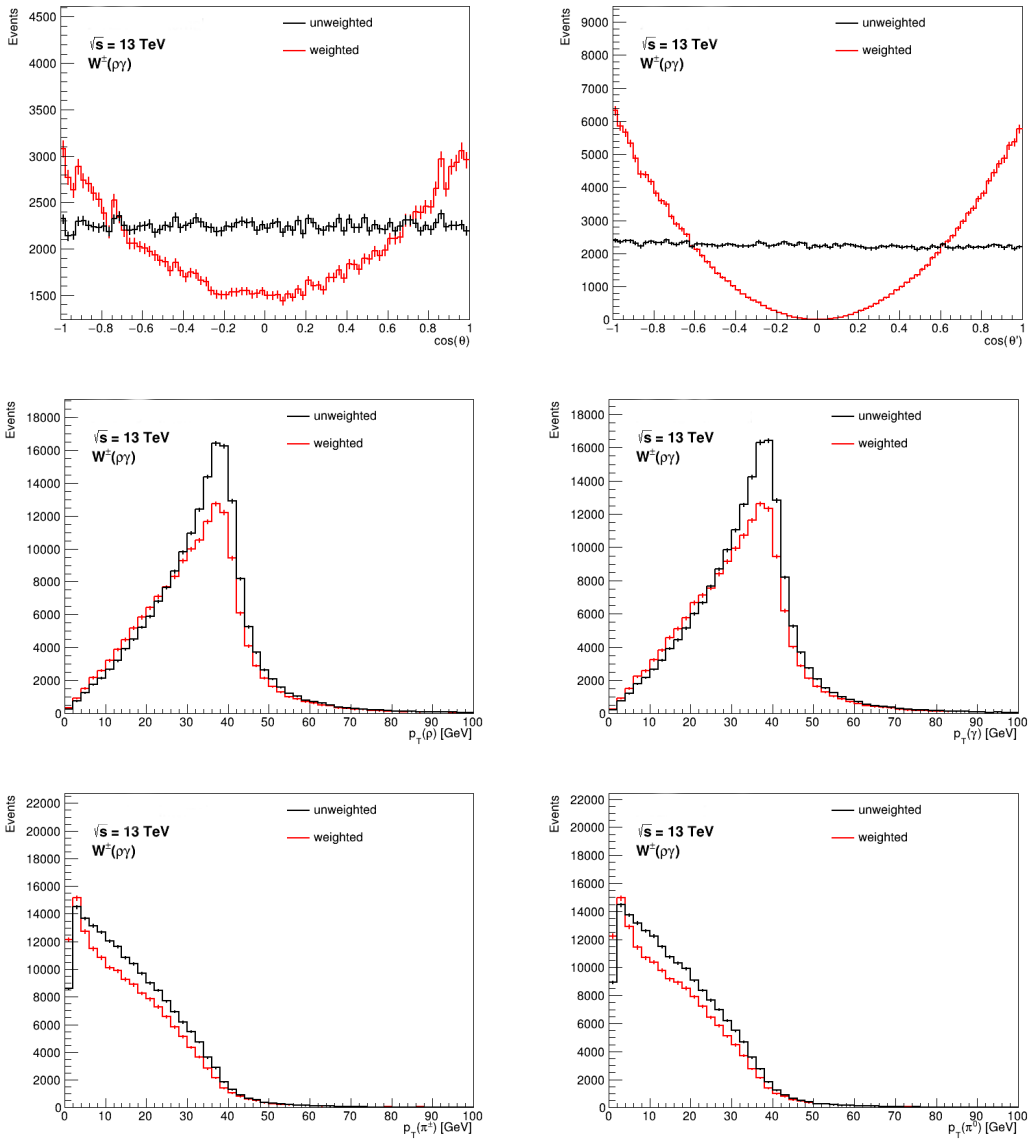


Figure 5.8: $\cos\theta$ and $\cos\theta'$ distribution and the p_T distributions of the ρ, γ, $\pi^{\pm}$, π^0, at generator level, with and without polarisation.

5.1.3 $Z \to e^- e^+$ MC sample

A small fraction of the total background is comprised of mis-reconstructed $Z \to e^- e^+$ events, with one electron being wrongly reconstructed as a photon and the other being reconstructed as a track. This background is modelled using POWHEG MC simulation at NLO precision, using the CTEQ6L1 PDF set with AZNLO tune. The parton shower is modelled using PYTHIA 8.186.

The remaining background processes are di-jet and photon+jet. No MC simulation is employed in the modelling of these processes. The data-driven technique discussed in the previous chapter is used. Contributions to the tau-photon final signal region from $W \to \tau\nu$ events are negligible, being suppressed by requirements on the $\Delta\phi$ separation between the tau and the photon. Cases in which initial or final state radiation is present are also negligible, since the invariant mass of the reconstructed photon and tau would not fall within the W boson mass resonance.

5.2 Trigger strategy

Dedicated triggers selecting events with a high p_{T} track and a high p_{T} photon were designed to record $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ events. They were part of a set of triggers developed for searches for exclusive Higgs, Z and W boson decays. These triggers rely on modified versions of tau trigger algorithms for the meson selection, taking advantage of the similarities between the decay of the mesons in the targeted chains and hadronic tau decays. The trigger algorithm looks for a tau with $p_{\mathrm{T}} > 25$ GeV with exactly one track with $p_{\mathrm{T}} > 30$ GeV and $\Delta R(\mathrm{tau}, \mathrm{trk}) < 0.06$. One further requirement was introduced on the trigger tau object in order to constrain the trigger rate within reasonable values: the $E_T(\mathrm{tau})/p_T(\mathrm{trk})$ is required to be within 0.4 and 0.85. This variable corresponds to the sum of transverse energy from all hadronic and electromagnetic clusters associated to the tau object, divided by the p_{T} of the leading track in the core of the tau object. Events are also required to have a photon with a p_{T} higher than 25 or 35 GeV. Finally, the invariant mass of the photon and "tau" system is required to be larger than 50 GeV.

The dedicated triggers used in the track-photon final state analysis and detailed

above are listed in Table 5.2, as well as their active period and corresponding luminosity. The trigger rates of the listed chains are within 1 and 3 Hz. For most of 2016 the trigger photon p_{T} threshold was 35 GeV, however by the end of the year a new trigger chain with a photon p_{T} threshold of 25 GeV was introduced in the trigger menu. Hence, for 2016, the g35 trigger chain was used until run 310247 and from run 310247 the g25 chain is used instead, given its higher efficiency. This change in trigger is implemented for signal by applying each trigger to the MC16a sample according to the proportion between the corresponding luminosity of the g35/g25 triggers: the g35 trigger is applied for approximately 80% of the events and the g25 is applied for the remaining 20%. Overall, the dedicated triggers have an average efficiency of 57.6% for $W^{\pm} \rightarrow \pi^{\pm}\gamma$ events. This efficiency is calculated with respect to the Signal Region offline selection listed on Table 5.7. The trigger efficiency as a function of track and photon p_{T} can be seen in Fig. 5.9, for the $W^{\pm} \rightarrow \pi^{\pm}\gamma$ signal. As expected, for lower p_{T} values, closer to the trigger p_{T} requirement, the efficiency is lower. As the p_{T} becomes larger there is an efficiency "turn-on" and the efficiency becomes constant within uncertainties. The track-photon trigger efficiency is higher for $W^{\pm} \rightarrow K^{\pm}\gamma$ events, corresponding to 60.1% . The difference in trigger efficiency between the two signals is discussed in section Section 5.6.

Trigger	Start Run	End Run	Luminosity (fb^{-1})
2016			
HLT_g35_medium_L1EM24VHI_tau25_singlepion_tracktwo_L1TAU12	297730	311481	33.367
HLT_g25_medium_tau25_singlepion_tracktwo_50mVis10000	310247	311481	6.228
2017			
HLT_g25_medium_L1EM24VHI_tau25_singlepion_tracktwo_50mVis10000	325713	340453	44.631
2018			
HLT_g25_medium_L1EM24VHI_tau25_singlepion_tracktwo_50mVis10000	348885	364292	58.792

Table 5.2: Summary of the dedicated track-photon triggers used in the track-photon final state analysis. The respective integrated luminosities are provided.

The requirement on $E_T(\mathrm{tau})/p_T(\mathrm{trk})$ is efficient for the $W^{\pm} \rightarrow \pi^{\pm}\gamma$ and $W^{\pm} \rightarrow K^{\pm}\gamma$ decays, since the distribution of this variable is mostly contained within the selection window. In the case of $W^{\pm} \rightarrow \rho^{\pm}\gamma$ decay, the additional electromagnetic energy deposition arising from the π^0 in the decay increases the value of $E_T(\mathrm{tau})/p_T(\mathrm{trk})$, decreasing the acceptance of the trigger to this signal. This behaviour can be seen in Fig. 5.10, which shows the distribution of $E_T(\mathrm{tau})/p_T(\mathrm{trk})$ for the $W^{\pm} \rightarrow \pi^{\pm}\gamma$ and

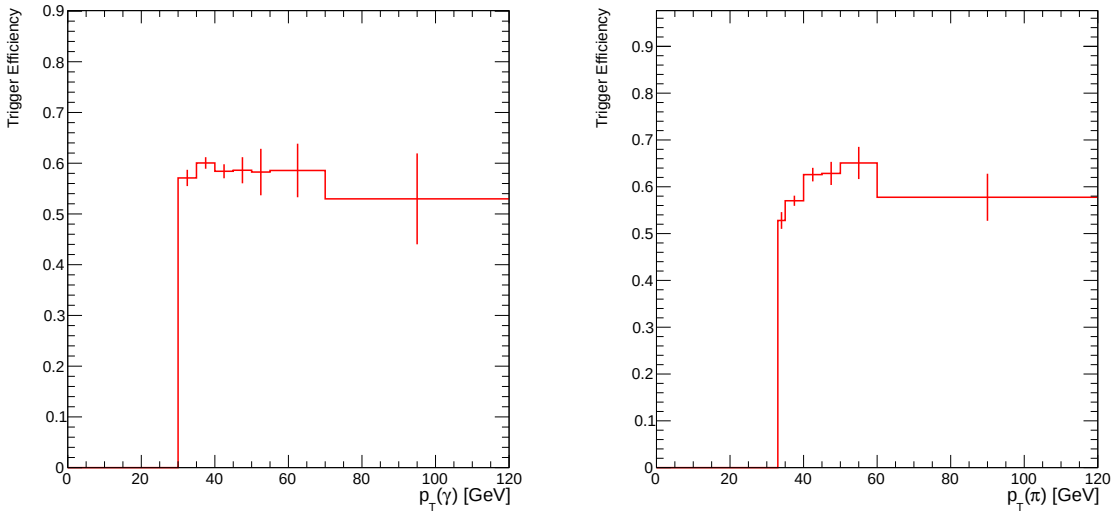


Figure 5.9: Trigger efficiency with respect to the offline selection, as a function of $p_{\mathrm{T}}(\gamma)$ and $p_{\mathrm{T}}(\mathrm{trk})$.

$W^{\pm} \to \rho^{\pm}\gamma$ signal MC, after requiring the presence of a high p_{T} tau and a high p_{T} photon. For this reason different triggers were used in the tau-photon analysis: di-photon triggers were found to be able to select $W^{\pm} \to \rho^{\pm}\gamma$ events, due to the decay of the π^0 into two collimated photons. The triggers used in the tau-photon final state analysis are listed in Table 5.3. Both trigger chains require the presence of two photons with $p_{\mathrm{T}} > 35\,\mathrm{GeV}$ and $p_{\mathrm{T}} > 25\,\mathrm{GeV}$. The chain which was active in 2015 and 2016 required these photons to have a loose ID, while in 2017 and 2018 a medium ID working point was required. The average trigger efficiency with respect to the offline selection summarised in Table 5.8 is 43%, for $W^{\pm} \to \rho^{\pm}\gamma$ events.

Trigger	Luminosity (fb^{-1})
2015	
`HLT_g35_loose_g25_loose`	3.245
2016	
`HLT_g35_loose_g25_loose`	33.402
2017	
`HLT_g35_medium_g25_medium_L12EM20VH`	44.631
2018	
`HLT_g35_medium_g25_medium_L12EM20VH`	58.792

Table 5.3: Triggers employed in the tau-photon final state analysis. The periods of operation and their respective integrated luminosities are given.

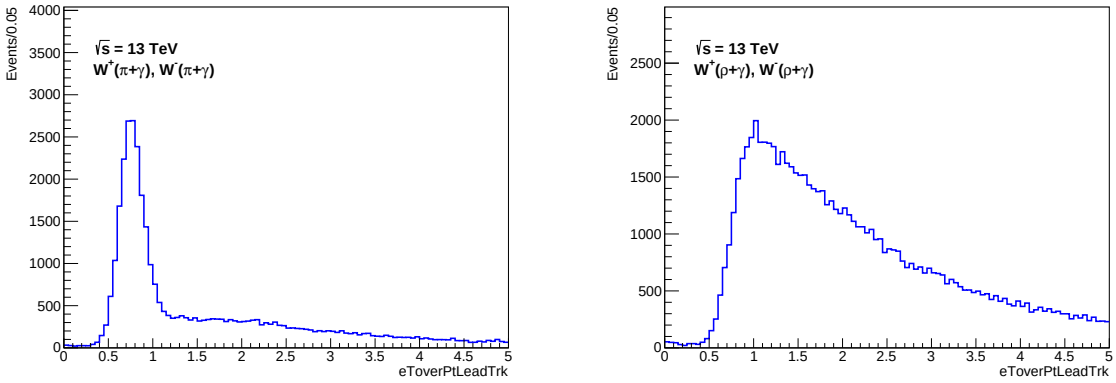


Figure 5.10: $E_T(\text{tau})/p_T(\text{trk})$ distribution associated to the meson for $W^{\pm} \to \pi^{\pm}\gamma$ (left) and $W^{\pm} \to \rho^{\pm}\gamma$ (right) MC events.

5.3 Pre-selection requirements

Track-photon final state

Events reconstructed in the track-photon final state are required to have at least one photon with $p_{\mathrm{T}} > 30$GeV or $p_{\mathrm{T}} > 35$ GeV (depending on the trigger being applied); an $|\eta|$ lower than 2.37 but not within $1.37 < |\eta| < 1.52$ (excludes the transition region between barrel and endcap); and fulfilling the "Tight" identification criteria [126]. They are also required to have at least one track with $p_{\mathrm{T}} > 30$ GeV; $|\eta| < 2.5$ and passing the "Tight Primary" identification working point [71]. For each event, the track and photon with the highest p_{T} are paired to form the $M\gamma$ candidate. Additionally, since the decay products are created back-to-back with a high $\Delta\phi(\text{trk}, \gamma)$ as can be seen in Fig. 5.4, events must satisfy $\Delta\phi(\text{trk}, \gamma) > \pi/2$. If both the selected photon and track have been reconstructed in the endcap they must satisfy $\eta(\text{trk}) \times \eta(\gamma) \geq 0$. The track-photon baseline selection is summarised in Table 5.4. The corresponding cut-flow can be found in Table 5.5. The photon and track selection requirements are the main cause of signal efficiency loss, as expected, since the respective p_{T} distributions peak between 30 and 40 GeV (Fig. 5.2). These p_{T} requirements are limited by the trigger thresholds. The number of data events is only listed on the cut-flow after the application of all the selections present in the derivation, since no data samples were available with no derivation skimming applied.

Trigger
HLT_g35_medium_L1EM24VHI_tau25_singlepion_tracktwo_L1TAU12
HLT_g25_medium_tau25_singlepion_tracktwo_50mVis10000
HLT_g25_medium_tau25_singlepion_tracktwo_50mVis10000
Photon requirements
$p_{\mathrm{T}} > 30\,\mathrm{GeV}$ (or $p_{\mathrm{T}} > 35\,\mathrm{GeV}$ depending on trigger)
$\lvert\eta\rvert < 2.37$ + crack veto
Tight ID
Track requirements
$p_{\mathrm{T}} > 30\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.5$
Tight ID
Global requirements
At least one primary vertex
$\eta(\mathrm{trk}) \times \eta(\gamma) \geq 0$ if track and γ in endcap
$\Delta\Phi(\mathrm{trk}, \gamma) > \pi/2$

Table 5.4: Summary of the track-photon pre-selection requirements.

Requirement	Data	$W^{\pm} \to \pi^{\pm}\gamma$	$W^{\pm} \to K^{\pm}\gamma$
No requirement	—	101.04	8.22
1 good photon	—	34.78	2.83
1 good track	—	19.55	1.61
Trigger	—	8.85	0.78
$p_{\mathrm{T}}(\gamma)$ >30/35 GeV	3439187	8.36	0.73
$\eta(\mathrm{trk}) \times \eta(\gamma) \geq 0$	3331060	8.36	0.73
$\Delta\Phi(\mathrm{trk}, \gamma) > \pi/2$	3024219	8.02	0.69

Table 5.5: Cut-flow of track-photon pre-selection for data, $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ signal. A "good photon" is a photon with a $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, passing the $|\eta|$ selection and "Tight" identification requirement; A "good track" is a "Tight" track with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 2.5$. The signal contributions are scaled according to the SM branching fractions, listed in Table 1.2.

Tau-photon final state

As already mentioned, in the tau-photon final state analysis, the $\rho^{\pm}$ meson is reconstructed as a tau object with exactly one charged pion (1-prong) and one neutral pion. In this way, both products of the $\rho^{\pm}$ decay are reconstructed. This strategy takes advantage of the similarity between the $\rho^{\pm} \to \pi^{\pm}\pi^0$ decay and $\tau \to \pi^{\pm}\pi^0\nu_{\tau}$ decays, which proceed most frequently via an intermediate $\rho^{\pm}$ meson, $\tau \to \rho^{\pm}(\pi^{\pm}\pi^0)\nu_{\tau}$ [15]. The neutrino in the τ-lepton decay is not included in the tau reconstruction, only the visible decay products form a $\tau_{\text{had-vis}}$ object, making the $\tau_{\text{had-vis}}$ well suited to reconstruct the $\rho^{\pm}$ meson.

Events reconstructed in the tau-photon final state are required to have at least one 1-prong $\tau_{\text{had-vis}}$ object, with exactly one neutral PFO. The selected $\tau_{\text{had-vis}}$ is further required to have $p_{\text{T}} > 26\,\text{GeV}$ and to be within the detectors geometrical acceptance, within $|\eta| < 2.5$ and not in the barrel/endcap transition region. A Medium τRNN working point is applied [77], supressing some of the multijet background. Events are further required to have at least one photon with $p_{\text{T}} > 36\,\text{GeV}$, passing the "Tight" identification requirements [126], and within $|\eta| < 2.5$ (excluding the transition region). The $M\gamma$ candidate is the tau and photon pair with the largest $\Delta\Phi(\tau_{\text{had-vis}}, \gamma)$. Finally, events are only retained for further analysis if $\Delta\Phi(\tau_{\text{had-vis}}, \gamma) > 2$. The summary of the pre-selection requirements applied in the tau-photon final state analysis can be found in Table 5.6.

5.4 $Z \to e^-e^+$ background suppression

As already mentioned, a small percentage of the background arises from mis-reconstructed $Z \to e^-e^+$ events. Given the proximity of the $Z \to e^-e^+$ background to the signal and the fact that resonant contributions are not modelled by the inclusive background modelling method as explained in Section 5.8.1, a set of pre-selection criteria was devised to suppress the $Z \to e^-e^+$ background. For this purpose, the $Z \to e^-e^+$ MC sample was used in a dedicated study. The acceptance of the track-photon baseline selection (Table 5.4) to $Z \to e^-e^+$ events was estimated to be 0.06%. Approximately 76000 $Z \to e^-e^+$ events survive this selection, corresponding to 3% of the total background

Triggers
`HLT_g35_loose_g25_loose`
`HLT_g35_medium_g25_medium_L12EM20VH`
Photon requirements
$p_{\mathrm{T}} > 36\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.37$ + crack veto
Tight ID, Tight isolation,
τ **requirements**
$h^{\pm}\pi^{0}$ decay mode
$p_{\mathrm{T}} > 26\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.5$ + crack veto
Medium τRNN score
Global requirements
At least one primary vertex
$\Delta\Phi(\tau_{\text{had-vis}}, \gamma) > 2$

Table 5.6: Summary of the tau-photon pre-selection requirements.

after the pre-selection has been applied. The invariant mass of the $Z \to e^{-}e^{+}$ events reconstructed in the track-photon final state can be found in Fig. 5.11. As expected, the distribution peaks near m_{Z}.

The strategy to suppress the $Z \to e^{-}e^{+}$ background focused on the track component. No specific $Z \to e^{-}e^{+}$ rejection requirements were based on the properties of the photon candidates. For each selected event, a track-electron ΔR matching procedure was performed: an electron was considered matched to the selected track if $\Delta R(\mathrm{trk}, e) < 0.01$[2]. It was verified that 98.5% (69.5%) of the $Z \to e^{-}e^{+}$ ($W^{\pm} \to \pi^{\pm}\gamma$) events contain a reconstructed electron with $\Delta R(\mathrm{trk}, e) < 0.01$. For these events, the following variables showed good discriminating power between $Z \to e^{-}e^{+}$ and $W^{\pm} \to \pi^{\pm}\gamma$ signal: Rhad of the matched electron, defined in Section 2.3.3; $pt(\mathrm{e})/pt(\mathrm{trk})$, which is the ratio between the p_{T} of the electron $\Delta R(\mathrm{trk}, e)$-matched to the selected track and the p_{T} of the selected track; and the eProbabilityHT of the track also defined in Section 2.3.3. Fig. 5.12 shows the distribution of these variables for signal and background. The peak at 0.5 for eProbabilityHT is caused by tracks which fall outside the TRT acceptance. The 2-dimensional distribution of Rhad vs eProbabilityHT for signal and background can be found in Fig. 5.13. Ultimately events in which an electron is matched to the selected track with a $\Delta R(\mathrm{trk}, e) < 0.01$ are retained for further analysis if Rhad $>$

[2]All electrons in the event were considered in the matching procedure, with no selection requirements applied.

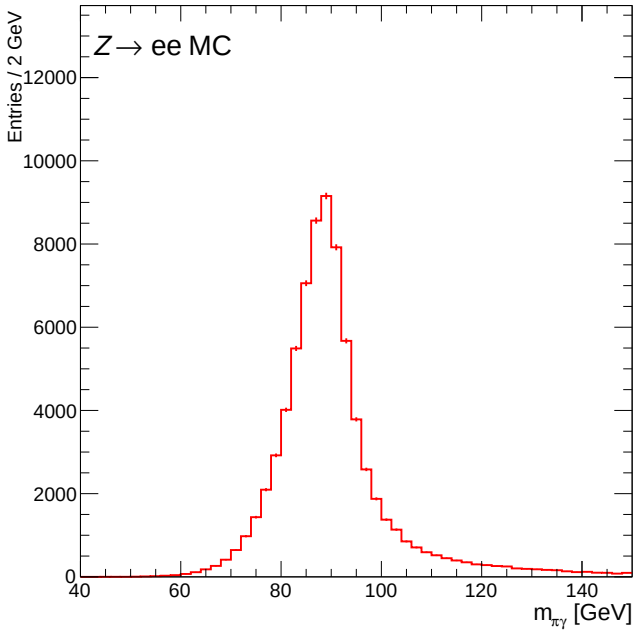


Figure 5.11: $\pi\gamma$ invariant mass distribution of $Z \to e^-e^+$ events.

0.03 and eProbabilityHT > 0.1. If no electron is matched to the selected track with a $\Delta R(\text{trk}, e) < 0.01$, no requirement is applied. The $W^\pm \to \pi^\pm\gamma$ signal efficiency for the requirement is 94%, while only 9% of the 76111 $Z \to e^-e^+$ background events survive. The choice was made to use Rhad in the selection instead of $p_\text{T}(\text{e})/p_\text{T}(\text{trk})$, since the first option yielded a slightly higher signal efficiency for the same $Z \to e^-e^+$ background rejection. 93% of the data events survive the $Z \to e^-e^+$ suppression selection.

Specific $Z \to e^-e^+$ rejection requirements were also introduced in the tau-photon final state analysis. Most of the variables utilised have already been introduced in this thesis: the track associated with the $\tau_\text{had-vis}$ is required to have eProbabilityHT < 0.9 and the $\tau_\text{had-vis}$ object is required to have $E_T(\tau_\text{had-vis})/p_T(\text{trk}) > 2.4$ and $\Delta R_\text{max}(\tau_\text{had-vis}, \text{trk}) > 0.036$. Additionally, the standard tau electron-veto BDT variable trained to discriminate between taus and electrons [78] is used, with a Tight working point being required. No $W^\pm \to \pi^\pm\gamma$ and $W^\pm \to K^\pm\gamma$ survive the tau-photon pre-selection and $Z \to e^-e^+$ suppression requirements, meaning that as expected the tau-photon final state is not sensitive to these two signals.

For both final state analyses, the pre-selection requirements listed in Table 5.4

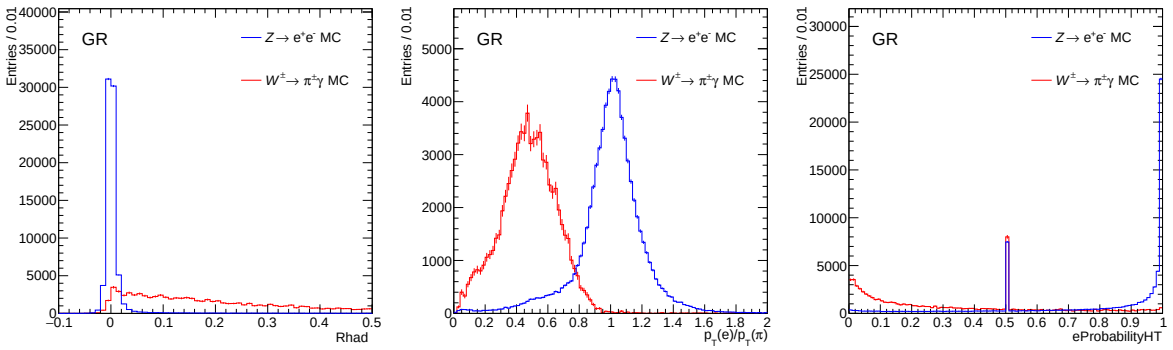


Figure 5.12: Left: Rhad distribution for $W^{\pm} \to \pi^{\pm}\gamma$ signal (blue) and $Z \to e^{-}e^{+}$ events (red). Center: $p_{\mathrm{T}}(\mathrm{e})/p_{\mathrm{T}}(\mathrm{trk})$ distribution for $W^{\pm} \to \pi^{\pm}\gamma$ signal (blue) and $Z \to e^{-}e^{+}$ events (red) ; Right: eProbabilityHT distribution for $W^{\pm} \to \pi^{\pm}\gamma$ signal (blue) and $Z \to e^{-}e^{+}$ events (red).

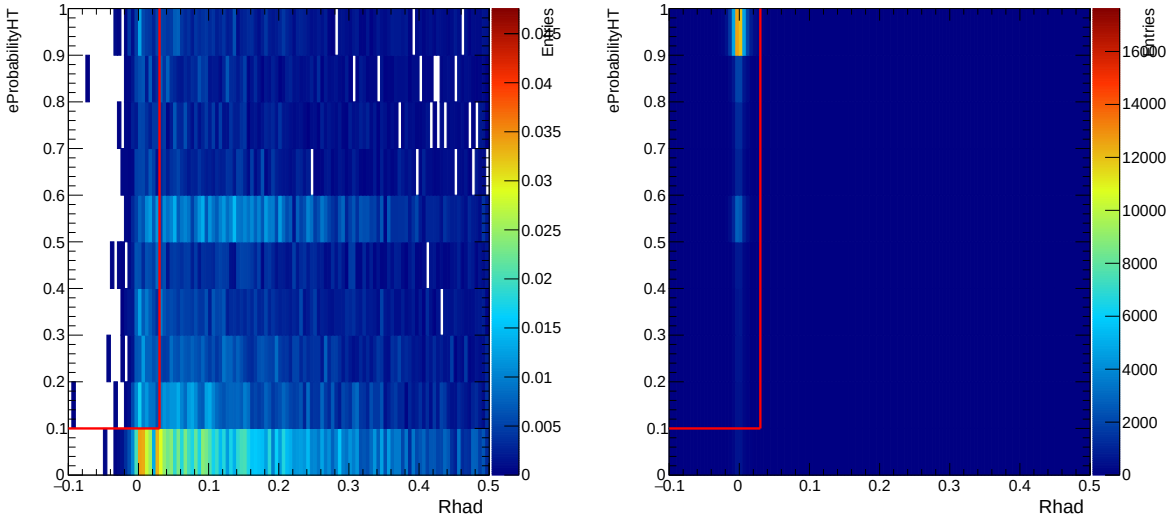


Figure 5.13: 2-dimensional distribution of events as function of Rhad and eProbabilityHT for $W^{\pm} \to \pi^{\pm}\gamma$ signal (left) and for the $Z \to e^{-}e^{+}$ background (right). The red lines delineate the region which is rejected by the selection.

and Table 5.6 plus the set of $Z \to e^-e^+$ suppression requirements define the Generation Region which is defined for use in the background modelling as previously explained in Section 4.2.1.

5.5 Signal and validation regions

Track-photon final state

The Signal Region (SR) in the track-photon final state was defined after successive optimisation of different requirements on kinematic and isolation variables, for the $W^{\pm} \to \pi^{\pm}\gamma$ signal. The figure of merit considered in the optimisation was $S/\sqrt{B}$, with the number of signal events, S, being estimated using the $W^{\pm} \to \pi^{\pm}\gamma$ MC sample, while the number of background events, B, was obtained using the estimate provided by the background model (detailed in Section 5.8.1).

Fig. 5.14 shows the distributions of the different kinematic and isolation variables considered for signal and background, after the application of pre-selection and $Z \to e^-e^+$ suppression requirements. Only events for which the reconstructed W invariant mass is within 75-85 GeV (approximately 2σ either side of m_W) are plotted.

The SR is defined by applying the three requirements listed below on top of the GR:

- $p_{\mathrm{T}}(\mathrm{trk}) > 33$ GeV;
- `FixedCutTight` photon isolation, defined in Section 2.3.2, including a photon calorimeter isolation and a photon track isolation requirement;
- $p_{\mathrm{T}}^{\mathrm{cone20}}(\mathrm{trk})/p_{\mathrm{T}}(\mathrm{trk}) < 0.14$, i.e. the sum of momenta of all tracks within $\Delta R = 0.2$ of the track direction (excluding the reconstructed track momentum) is required to be less than 14 % of $p_{\mathrm{T}}(\mathrm{trk})$;

The specific track p_{T} and isolation requirements correspond to the maximum $S/\sqrt{B}$ in Fig. 5.15. The optimisation of the track isolation requirement was performed after imposing the SR level requirement on $p_{\mathrm{T}}(\mathrm{track})$. The two-dimensional distributions of these two variables can be seen in Fig. 5.16. In the case of the photon isolation

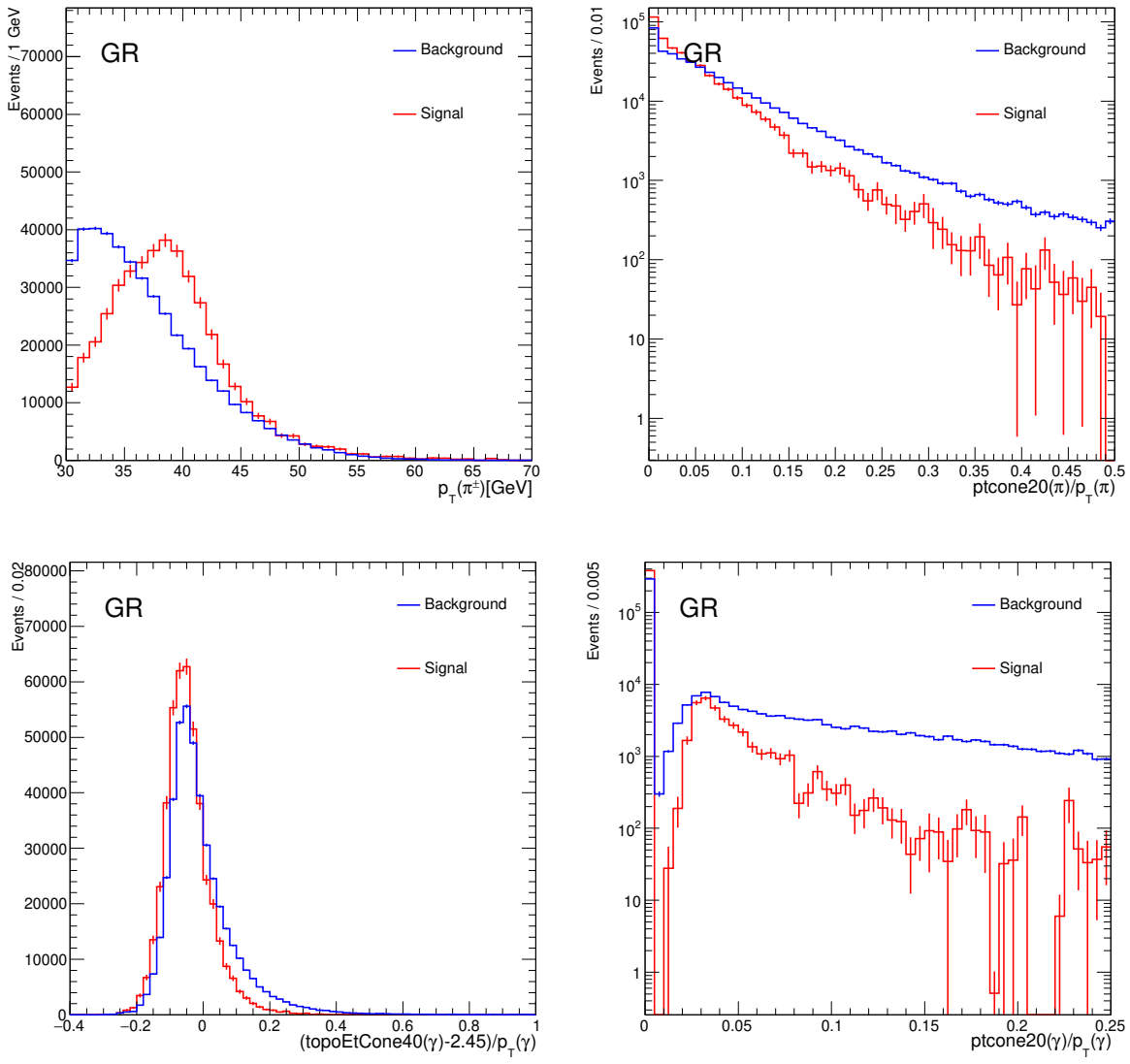


Figure 5.14: Background and signal distributions of kinematic and isolation variables used to define the SR. Top left: p_{T}(trk); Top right: track isolation; Bottom left: photon calorimeter isolation; Bottom right: photon track isolation. The signal distribution has been normalised to the number of background events. Only events with a reconstructed W invariant mass within 75-85 GeV are shown.

requirement the choice was made to use the pre-defined `FixedCutTight` working point [60].

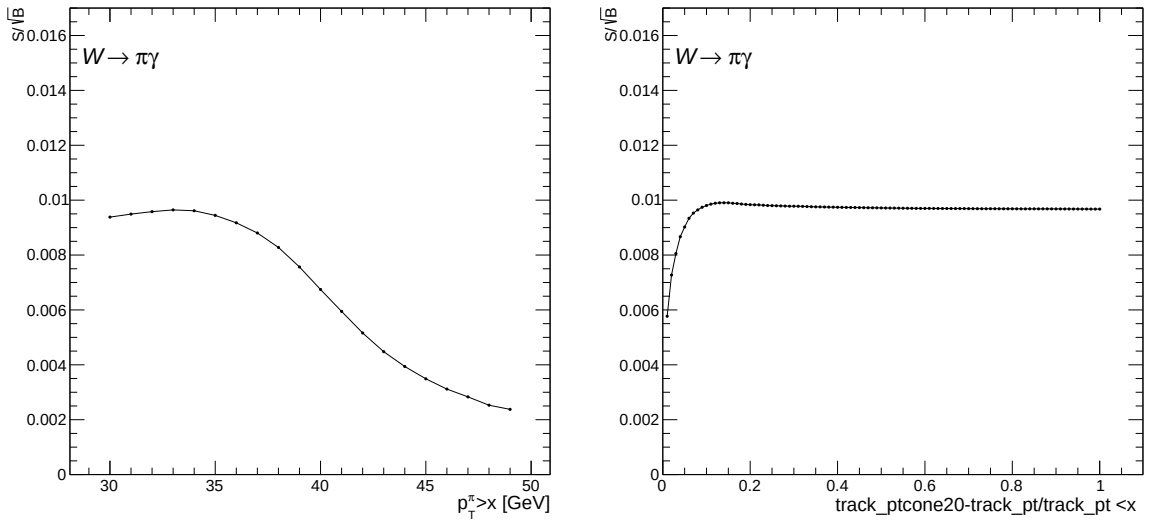


Figure 5.15: $S/\sqrt{B}$ as function of the working point of p_{T}(trk) and track isolation.

Four intermediate validation regions (VRs) are defined in order to verify the performance of the background modelling, as was done in Section 4.2.2.2: VR1 is defined by applying the track p_{T} requirement on top of the GR; VR2a by applying the photon calorimeter isolation requirement; VR2b by applying the photon track isolation requirement; and VR3 by applying the track isolation requirement. The complete SR selection is summarised in Table 5.7. The effect of the SR selection on the photon and meson p_{T} distributions is found in Fig. 5.17, for $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$. The dashed lines correspond to the p_{T} distributions at generator level (for events that pass the geometrical η requirements) and the full lines are the p_{T} distributions at reconstructed level after all selection requirements have been applied. The overall track-photon SR efficiency is 5.2% for $W^{\pm} \to \pi^{\pm}\gamma$ and 5.8% for $W^{\pm} \to K^{\pm}\gamma$. This efficiency difference is discussed in the next section.

Tau-photon final state

The tau-photon final state SR was defined after simultaneous optimisation of the requirements on three variables: $p_{\mathrm{T}}(\tau_{\text{had-vis}})$, ΔR_{max} and $\log(|d_0(\tau_{\text{had-vis}})|)$ where d_0

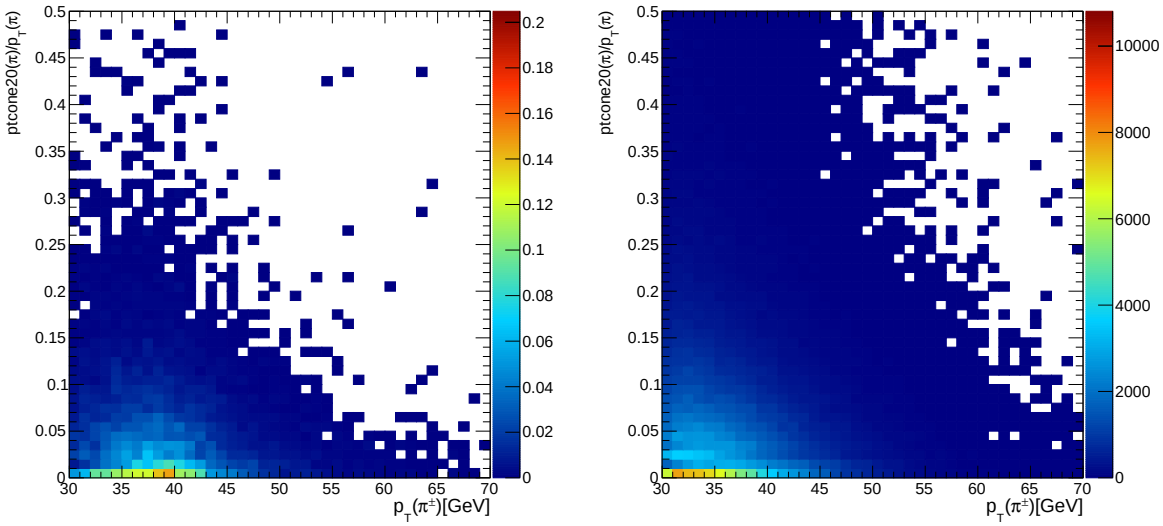
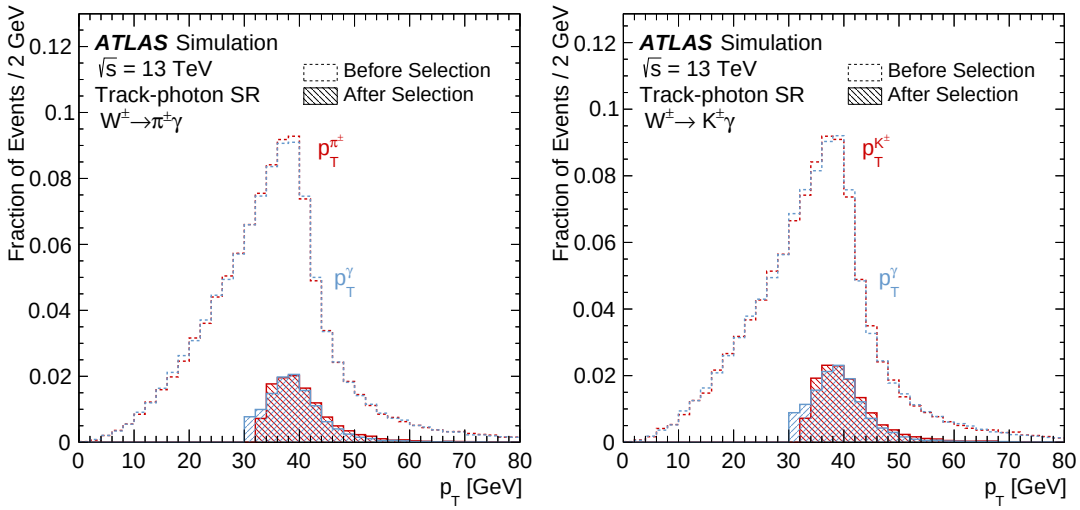


Figure 5.16: Two dimensional distribution of p_{T}(trk) versus track isolation, for $W^{\pm} \to \pi^{\pm}\gamma$ signal (left), and background (right). The signal is modelled by MC simulation and the background is modelled through the method detailed in Section 5.8.1.



(a) (b)

Figure 5.17: p_{T} distributions of the photon and meson for (a) $W^{\pm} \to \pi^{\pm}\gamma$ and (b) $W^{\pm} \to K^{\pm}\gamma$ simulated events. The dashed-line distributions with a clear fill show the events at generator level which fall within the analysis geometric acceptance. The solid-line distributions with a hatched fill show the fraction of these events which pass the track-photon SR event selection. The relative difference between the two sets of distributions corresponds to the effect of the full SR efficiency.

Baseline selection
Trigger:
`HLT_g35_medium_L1EM24VHI_tau25_singlepion_tracktwo_L1TAU12`
`HLT_g25_medium_tau25_singlepion_tracktwo_50mVis10000`
`HLT_g25_medium_tau25_singlepion_tracktwo_50mVis10000`
Photon requirements:
$p_{\mathrm{T}} > 30\,\mathrm{GeV}$ (or $p_{\mathrm{T}} > 35\,\mathrm{GeV}$ depending on trigger)
$\lvert\eta\rvert < 2.37$ + crack veto
Tight ID
Track requirements:
$p_{\mathrm{T}} > 30\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.5$
Tight ID
Global requirements:
$\eta(\mathrm{trk}) \times \eta(\gamma) \geq 0$ if track and γ in endcap
$\Delta\Phi(\mathrm{trk}, \gamma) > \pi/2$
$Z \to e^- e^+$ **suppression requirement**
If at least 1 electron is found with ΔR(trk, e) < 0.01:
Rhad > 0.03 if also eProbabilityHT > 0.1
GR selection: baseline + $Z \to e^- e^+$ suppression requirement
VR1 selection: GR + $p_{\mathrm{T}}(\pi) > 33\,\mathrm{GeV}$
VR2a selection: GR + Photon Calo Isolation
VR2b selection: GR + Photon Track Isolation
VR3 selection: GR + (ptcone20-$p_{\mathrm{T}}(\mathrm{trk})$)/$p_{\mathrm{T}}(\mathrm{trk}) < 0.14$
SR selection: all requirements

Table 5.7: Summary of SR requirements applied in the track-photon final state analysis.

is the transverse impact parameter of the track associated to the $\tau_{\text{had-vis}}$ object. Once again the figure of merit $S/\sqrt{B}$ was used in the optimisation. The validation regions are defined by applying each of the SR requirements on top of the GR. The SR selection requirements are summarised in Table 5.8. Fig. 5.18 shows the effect of the SR selection on the photon and meson p_{T} distributions for $W^{\pm} \to \rho^{\pm}\gamma$. Once again, the dashed-lined distributions correspond to the p_{T} at generator level (for events that pass the geometrical η requirements) and the full lined-distributions are the p_{T} at reconstructed level for the fraction of the events that passes all SR selection requirements. The overall tau-photon SR efficiency is 0.3% for $W^{\pm} \to \rho^{\pm}\gamma$.

Baseline selection
Triggers:
`HLT_g35_loose_g25_loose`
`HLT_g35_medium_g25_medium_L12EM20VH`
Photonrequirements:
$p_{\text{T}} > 36\,\text{GeV}$, $\lvert\eta\rvert < 2.37$ + crack veto
Tight ID, Tight isolation,
Tight ID
Tau requirements:
τ **requirements**
$h^{\pm}\pi^0$ decay mode
$p_{\text{T}} > 26\,\text{GeV}$, $\lvert\eta\rvert < 2.5$ + crack veto
Medium τRNN score
Global requirements:
At least one primary vertex
$\Delta\Phi(\tau_{\text{had-vis}}, \gamma) > 2$
$Z \to e^{-}e^{+}$ **suppression requirement**
Tight electron-veto working point
$E_T(\text{tau})/p_T(\text{trk}) > 2.4$
$\Delta R_{\tau}^{\text{max}} > 0.036$
eProbabilityHT (associated to the tau track) < 0.9
GR selection: baseline + $Z \to e^{-}e^{+}$ suppression requirement
VR1 selection: GR + $p_{\text{T}}(\tau_{\text{had-vis}}) > 30\,\text{GeV}$
VR2 selection: GR + $\Delta R_{\tau}^{\text{max}} < 0.065$
VR3 selection: GR + $\log(\lvert d_0(\tau_{\text{had-vis}})\rvert) < -1.2$
SR selection: all requirements

Table 5.8: Summary of SR requirements applied in the tau-photon final state analysis.

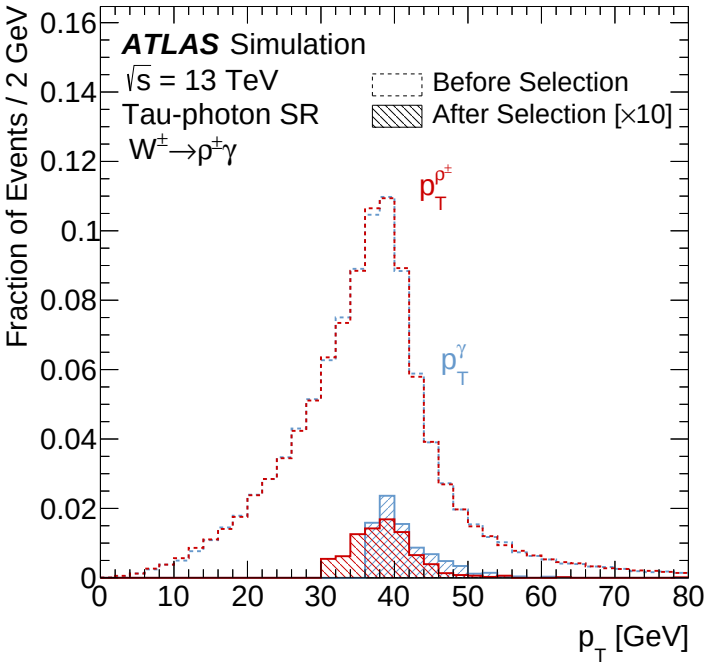


Figure 5.18: p_{T} distributions of the photon and meson for $W^{\pm} \to \rho^{\pm}\gamma$ in the tau-photon final state. The dashed-line distributions with a clear fill show the events at generator level which fall within the analysis geometric acceptance. The solid-line distributions with a hatched fill show the fraction of these events at reconstructed level which pass the SR event selection in the tau-photon final state.

5.6 $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ efficiency difference

Despite the similarities between the kinematics of the two decays (Figs. 5.5 to 5.6), a difference of about 11% has been observed between the $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ efficiencies. Examination of the reconstructed p_{T} distibutions of the decay products did not show differences that could account for the 11% difference (Fig. 5.19). The cut-flow of both channels was compared in order to better understand which requirements are responsible for the difference, and can be found in Table 5.9. The largest difference arises from the trigger selection. In order to compare the response of pions and kaons to the $E_T(\mathrm{tau})/p_T(\mathrm{trk})$ trigger requirement, taus were ΔR matched to the selected tracks in the decays, and the $E_T(\mathrm{tau})/p_T(\mathrm{trk})$ ditribution of the matched taus was plotted. This distribution can be seen in Fig. 5.20. It shows that on average, kaons have a smaller $E_T(\mathrm{tau})/p_T(\mathrm{trk})$ than pions, and that there is an approximately 6.4% higher efficiency for kaons to the $0.4 < E_T(\mathrm{tau})/p_T(\mathrm{trk}) < 0.85$ requirement. These differences are not completely unexpected given that kaons and pions have different nuclear interaction lengths [127]. Differences were also verified in the response to the $Z \to e^{-}e^{+}$ suppression criterion. The track eProbabilityHT and electron Rhad distributions for $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ can be found in Fig. 5.21. Once again the differences in Rhad should arise from subtle differences in the hadronic showers of the kaons and pions. In the case of the eProbabilityHT variable the differences should arise from the different gamma factor for kaons and pions, leading to different amounts of transition radiation being produced in the TRT.

5.7 $W^{\pm} \to \rho^{\pm}\gamma$ events in track-photon SR

Despite having been optimised for the $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ signals, the track-photon SR is found to still have 0.6% efficiency for $W^{\pm} \to \rho^{\pm}\gamma$ events, corresponding to 1.24 events out of the 221 expected. In this case the π^0 of the $\rho^{\pm} \to \pi^{\pm}\pi^0$ decay is not explicitly reconstructed, and as such the corresponding W boson invariant mass distribution is shifted with respect to the $W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma$ distribution, as can be seen in Fig. 5.36 of Section 5.10. This means that a fit to the reconstructed W

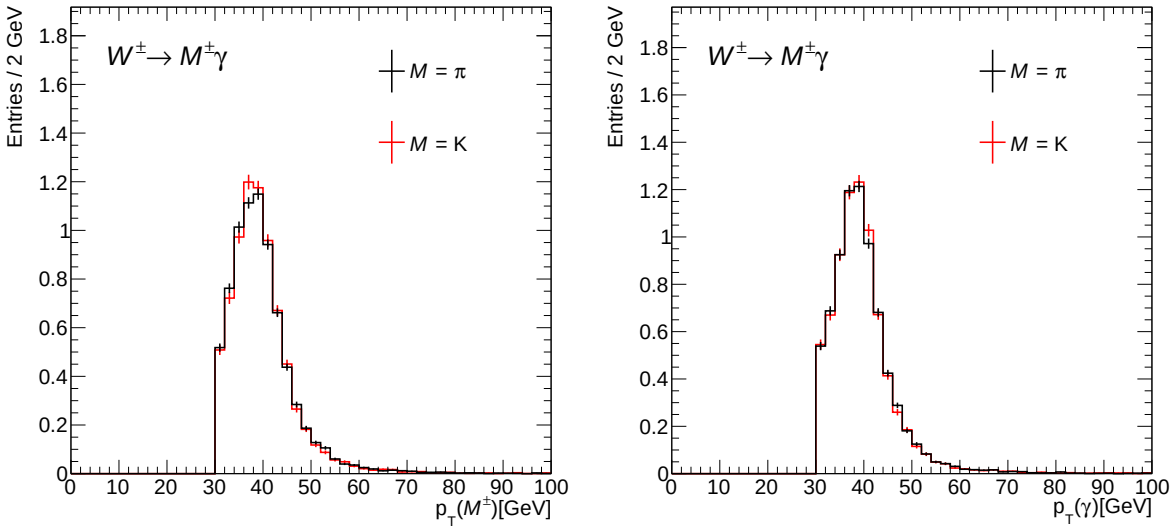


Figure 5.19: $p_{\mathrm{T}}(\mathrm{trk})$ and $p_{\mathrm{T}}(\gamma)$ at reconstructed level, for $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$. The $W^{\pm} \to K^{\pm}\gamma$ distributions are normalised to the integral of the $W^{\pm} \to \pi^{\pm}\gamma$ distributions.

	$W^{\pm} \to \pi^{\pm}\gamma$	$W^{\pm} \to K^{\pm}\gamma$	Efficiency Difference (%)
Total	101.04	8.22	—
1 good photon	34.78	2.83	0
1 good track	19.55	1.61	-1.3
Trigger	8.85	0.78	-8.3
$p_{\mathrm{T}}(\gamma)$ >30/35 GeV	8.36	0.73	-7.3
$\eta(\mathrm{track}) \times \eta(\gamma) \geq 0$	8.36	0.73	-7.3
$\Delta\phi$ requirement	8.02	0.69	-5.8
Electron Veto	7.57	0.67	-8.8
SR	5.25	0.47	-11

Table 5.9: Cut-flow comparison between $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$. The efficiency difference is relative to $W^{\pm} \to \pi^{\pm}\gamma$.

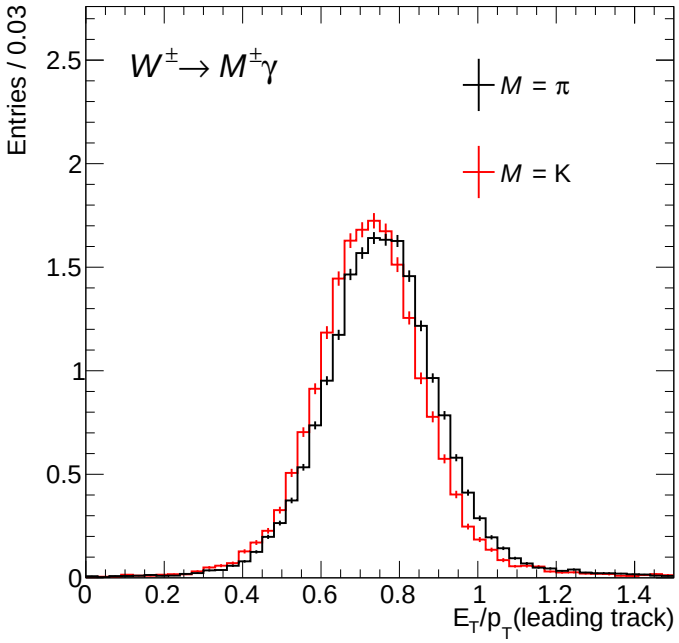


Figure 5.20: Comparison between $E_T(\text{tau})/p_T(\text{trk})$ for taus ΔR matched to the selected tracks, for $W^{\pm} \rightarrow \pi^{\pm}\gamma$ and $W^{\pm} \rightarrow K^{\pm}\gamma$ decays. The $W^{\pm} \rightarrow K^{\pm}\gamma$ distribution is normalised to the integral of the $W^{\pm} \rightarrow \pi^{\pm}\gamma$ distribution.

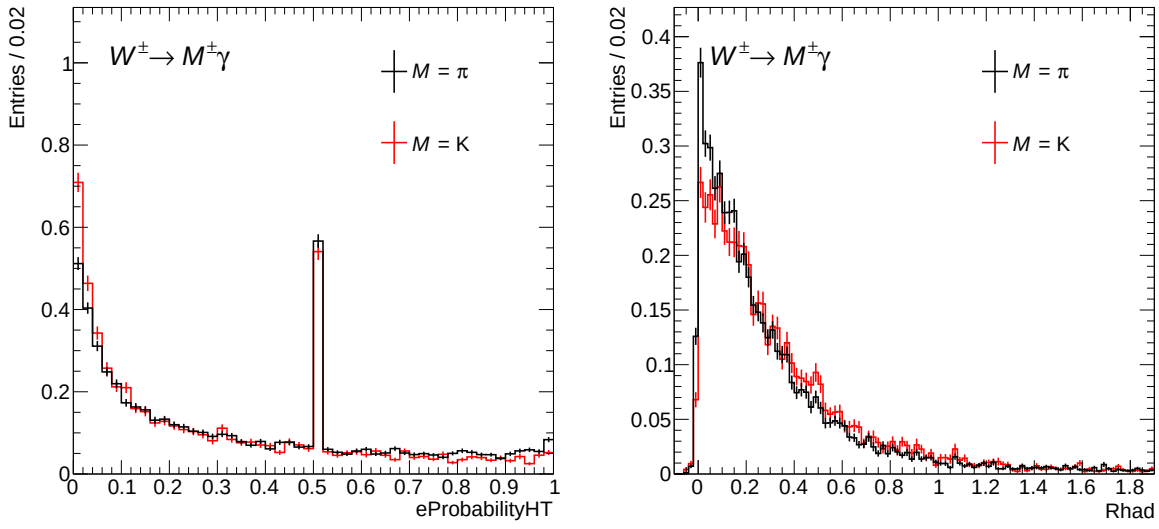


Figure 5.21: Comparison between track eProbabilityHT and electron Rhad (for electrons ΔR matched to the selected tracks), for $W^{\pm} \rightarrow \pi^{\pm}\gamma$ and $W^{\pm} \rightarrow K^{\pm}\gamma$ decays. The $W^{\pm} \rightarrow K^{\pm}\gamma$ distributions are normalised to the integral of the $W^{\pm} \rightarrow \pi^{\pm}\gamma$ distributions.

boson invariant mass distribution in the track-photon SR including can constrain both $W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma$ and $W^{\pm} \to \rho^{\pm}\gamma$ signal strengths at the same time. The effect of the SR selection on the p_{T} distributions of the photon and the charged pion resulting from the $W^{\pm} \to \rho^{\pm}\gamma$ decay can be found in Fig. 5.22. The $p_{\mathrm{T}}(\pi^0)$ distribution is only shown at generator level, since no attempt is made at reconstructing the π^0.

The number of raw simulated $W^{\pm} \to \rho^{\pm}\gamma$ events surviving the track-photon and tau-photon selection requirements is listed in Section 5.7. The overlap between $W^{\pm} \to \rho^{\pm}\gamma$ events in both SRs is $< 0.3\%$. From the table, one can see that the main reason for this lack of overlap is the different trigger strategy employed for the two final state analyses. The two trigger selections are found to be orthogonal due to the different requirements applied to the meson component: the track-photon trigger requires the presence of a high p_{T} track with lower $E_T(\mathrm{tau})/p_T(\mathrm{track})$ ratio, selecting events with a less energetic π^0; the di-photon trigger requires the presence of a high p_{T} photon, taking advantage of the presence of the π^0 in the decay of the $\rho^{\pm} \to \pi^{\pm}\pi^0$ and selecting events with a more energetic π^0. The overlap between data events surviving the track-photon and tau-photon selection requirements was also studied and Section 5.7 lists the number of data events passing both trigger requirements and the track-photon SR selection. These studies show that the overlap of data events in both SRs is $< 0.6\%$. The negligible overlap between the two final states allows a simultaneous fit to be performed using both SRs, allowing to better constrain the $W^{\pm} \to \rho^{\pm}\gamma$ signal strength.

Year	Total MC	mγT	DPT	both triggers	track-photon SR	both SRs
2016	60000	661	3520	15	280	2
2017	80000	1013	2482	5	380	1
2018	100000	1248	3129	8	457	0

Table 5.10: Summary of the $W^{\pm} \to \rho^{\pm}\gamma$ signal overlap studies in the track-photon and tau-photon final state SRs. mγT stands for the dedicated meson-photon trigger and DPT stands for di-photon trigger.

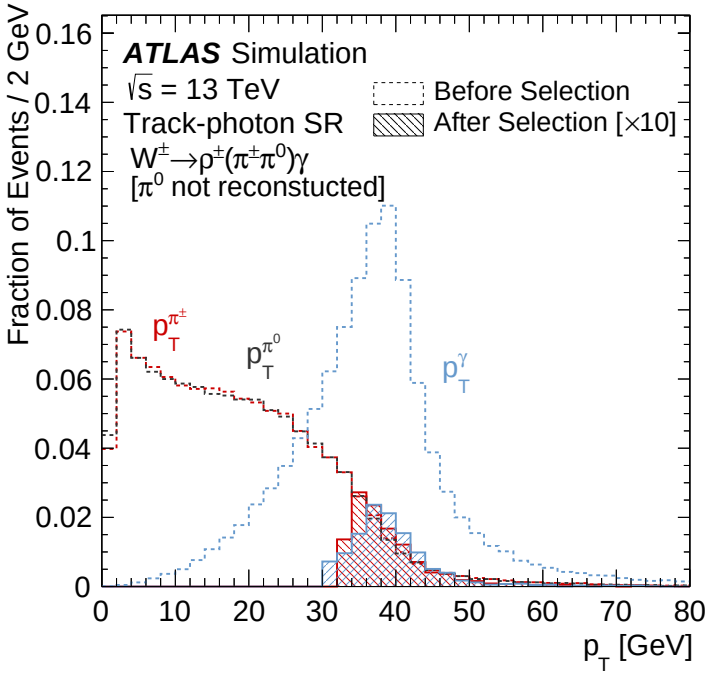


Figure 5.22: p_{T} distributions of the decay products for $W^{\pm} \to \rho^{\pm}\gamma$ in the track-photon final state. The dashed-line distributions with a clear fill show the events at generator level which fall within the analysis geometric acceptance. The solid-line distributions with a hatched fill show the fraction of these events at reconstructed level which pass the SR event selection in the track-photon final state. The reconstructed distribution of the π^0 in the track-photon SR is not shown as the object is not explicitly reconstructed.

Year	mγT	mγT+DPT	track-photon SR	track-photon SR + DPT
2016	680418	32671	232536	2814
2017/2018	2781790	73998	831154	3387

Table 5.11: Summary of the data overlap studies between events passing the track-photon and tau-photon selections. mγT stands for the dedicated meson-photon trigger and DPT stands for di-photon trigger.

5.8 Background modelling

5.8.1 Multijet background

For both final states, the main background contributions arise from di-jet and jet+photon events. This background is modelled using the non-parametric data-driven technique using ancestral sampling described in Section 4.2. A factorised n-Dimensional PDF is constructed in the GRs, which are defined by the requirements listed in Table 5.7 and Table 5.8 and the $Z \to e^- e^+$ suppression requirements. Histograms of up to 3 dimensions are used to model the distributions of relevant properties of the GR data sample. Each multijet pseudo-event is fully described by the meson and photon four-vectors and by the additional variables used in selection criteria applied to GRs that define the SRs. Composite variables like the W boson invariant mass can be constructed from the meson and photon 4-vectors.

In the case of the track-photon final state, the track and photon kinematic and isolation variables are used in the modelling. The sampling procedure used to generate multijet pseudo-events in the track-photon analysis is illustrated in Fig. 5.23 and is the following:

1. $p_{\mathrm{T}}(\mathrm{trk})$ and $p_{\mathrm{T}}(\gamma)$ are sampled simultaneously from their 2D data distribution.

2. Track isolation is described in bins of $p_{\mathrm{T}}(\gamma)$ and $p_{\mathrm{T}}(\mathrm{trk})$, in a 3D distribution. Given the values sampled in step 1, the distribution is projected along the track isolation dimension and a value of track isolation is randomly sampled.

3. Photon calorimeter isolation is described in bins of $p_{\mathrm{T}}(\gamma)$. Given the value of $p_{\mathrm{T}}(\gamma)$ sampled in step 1, the distribution is projected along the photon calorimeter isolation dimension and a value for photon calorimeter isolation is sampled.

4. $\Delta\eta(\mathrm{trk}, \gamma)$ and photon track isolation are described in bins of photon calorimeter isolation using a 3D histogram. The distribution is projected along the $\Delta\eta(\mathrm{trk}, \gamma)$ and photon track isolation dimensions, depending on the value of photon calorimeter isolation sampled in the previous step, and two values are simultaneously sampled from this 2D distribution.

5. $\Delta\phi(\text{trk},\gamma)$ is described in bins of $\Delta\eta(\text{trk},\gamma)$ using a 2D histogram. Based on the value of $\Delta\eta(\text{trk}\gamma)$ obtained in the previous step, the template is projected along the $\Delta\phi(\text{trk},\gamma)$ dimension and a value for $\Delta\phi(\text{trk},\gamma)$ is sampled.

6. Values of $\eta(\text{trk})$ and $\phi(\text{trk})$ are sampled from the corresponding distributions in data.

7. From the values obtained in 6 and the values previously obtained for $\Delta\eta(M,\gamma)$ and $\Delta\phi(M,\gamma)$, $\eta(\gamma)$ and $\phi(\gamma)$ are calculated. The photon and track four momentum vectors are completely defined at this point, imposing $m(\gamma) = 0$ and $m(\text{trk}) = M_{\pi^\pm}$ [29].

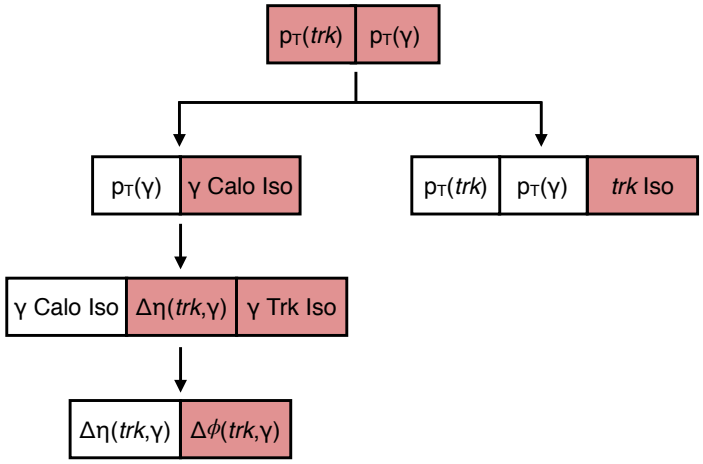


Figure 5.23: Schematic of the sampling sequence used in the generation of multijet pseudo-events in the track-photon final state analysis. Variables not shown explicitly are sampled in a factorized, uncorrelated manner from one-dimensional templates. Groups of two (three) variables represent two (three)-dimensional templates. Arrows are used to show the sequential order of steps in the sampling. Variables are highlighted with color at the step in which they are defined for each pseudo-candidate.

The linear correlations observed between the variables used in the modelling in GR, for the data and background model sample can be found in Fig. 5.24. Two dimensional distributions built for some of the variables which the highest linear correlations can be found in Fig. 5.25. Hints of non-trivial, non-linear correlations can be seen, which the modelling tries to capture. The most important correlations in data are reproduced by the model. After generation of the sample of multijet pseudo-events in GR, the SR

and VR selections listed in Fig. 5.17 are applied to the sample (the exact same selection requirements which are applied to the data). The number of generated pseudo-events is arbitrary, so the sample in each region is scaled according to the ratio between the number of data events and multijet pseudo-events in the GR.

Given the two different photon p_{T} thresholds used in the analysis (depending on the trigger used), two samples of pseudo-events are generated for each of the subsets of data for which the $p_{\mathrm{T}} > 30$ GeV and $p_{\mathrm{T}} > 35$ GeV requirements are applied. These are subsequently summed (after being appropriately scaled to each respective GR) creating the template used in the statistical analysis. Fig. 5.24 refers to the inclusive sample.

As previously explained and shown in Fig. 4.7, this background estimation technique is not expected to model resonant contributions. Therefore, the small fraction of $Z \to e^{-}e^{+}$ background present at GR level (0.2% of the total background) is not modelled by this technique and has a negligible effect in the inclusive background template generated.

The number of multijet background events predicted in all considered regions is listed in Table 5.12. The predicted number of $Z \to e^{-}e^{+}$ events (from the MC simulation), and $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ signal events (using the branching fractions from Ref. [28]) are also listed. Fig. 5.26 shows the pre-fit background prediction for the W boson invariant mass (the discriminant variable used in the fit) compared to the data, for the track-photon GR, SR and VRs. The grey band represents the shape uncertainty on the multijet background estimate which is derived following the procedure described in Section 5.9. Good agreement can be seen between the data and the background, verifying the correct modelling of the most important correlations in the data sample. The distributions of the variables used in the modelling and p_{T}(track, photon) can be found in Appendix D, for both background and signal.

The W boson invariant mass template is smoothed using Kernel Density Estimation [79], in order to obtain a continuous, smooth PDF to be used in the fit to the data. The `Roofit` class `RooKeysPdf` is employed for this smoothing. A fixed smoothing parameter of 0.1 is used, given the high-statistics sample generated during the modelling. No mirroring is used in the KDE. The resulting template can be seen in Fig. 5.27.

The multijet estimate in the tau-photon final state analysis is obtained using the same

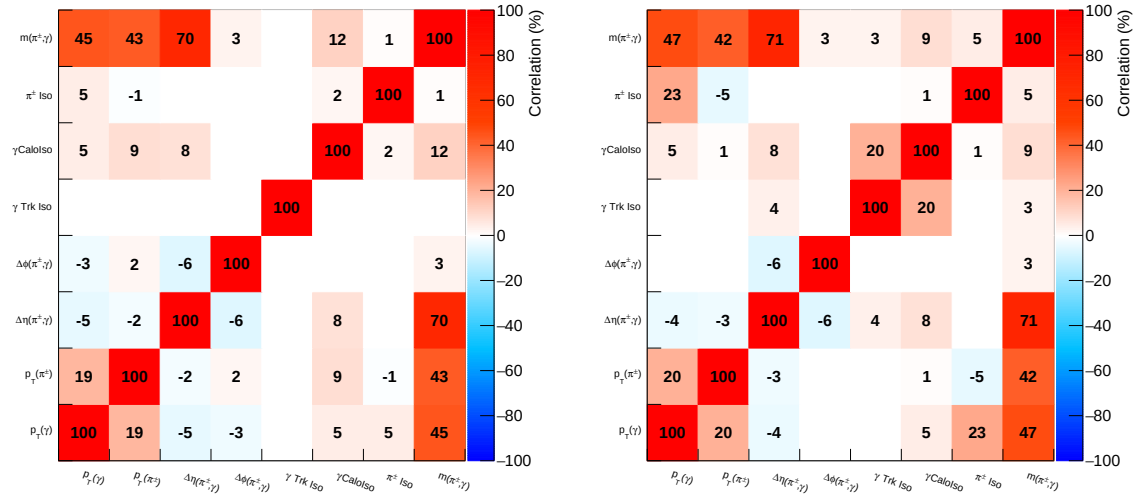


Figure 5.24: Linear correlations between variables used in the background modelling. The matrix is shown for the data (left) and background model (right) in the track-photon GR.

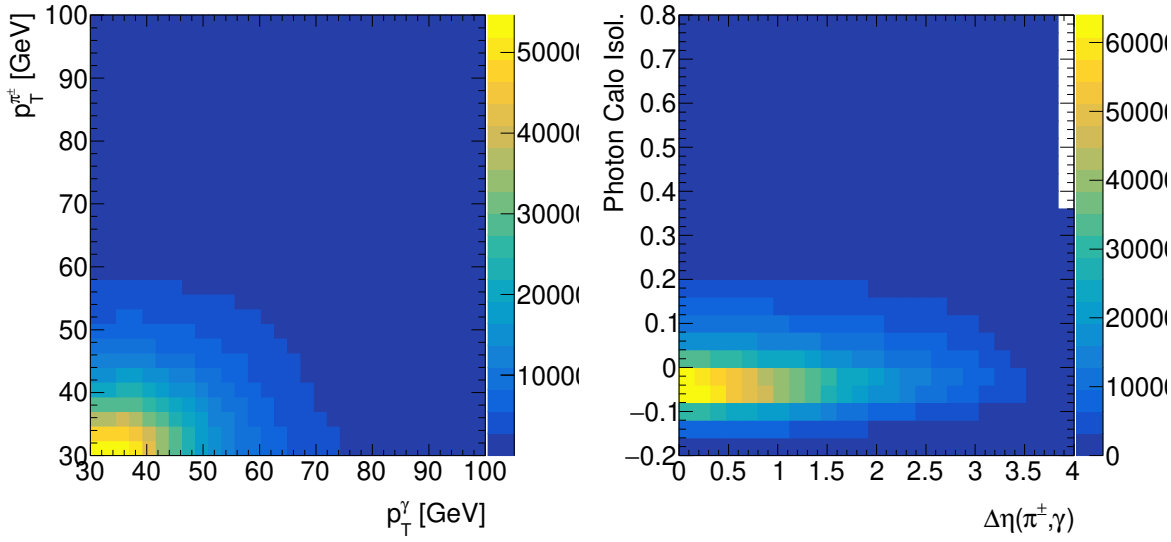
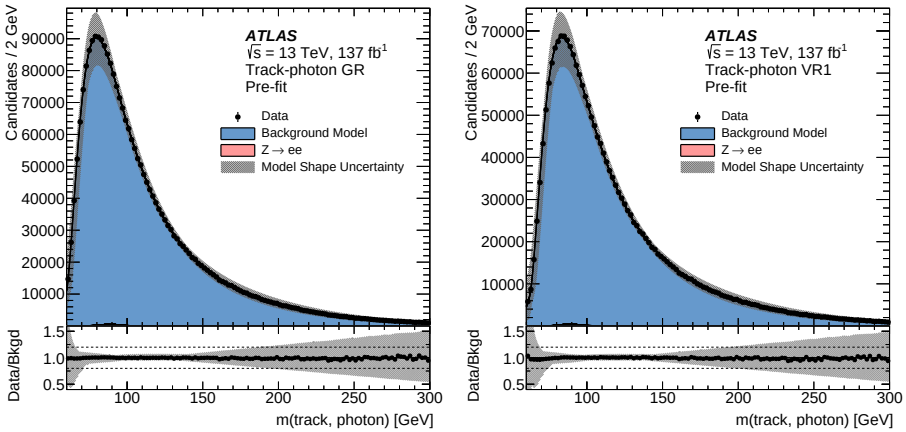
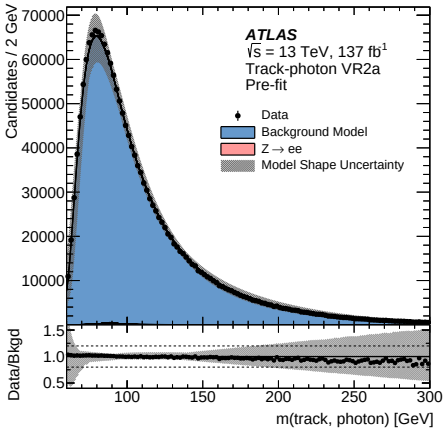


Figure 5.25: Two dimensional distributions of $p_{\mathrm{T}}(\gamma)$ versus $p_{\mathrm{T}}(\mathrm{trk})$(left) and $\Delta\eta(\mathrm{trk},\gamma)$ and photon calorimeter isolation (right).

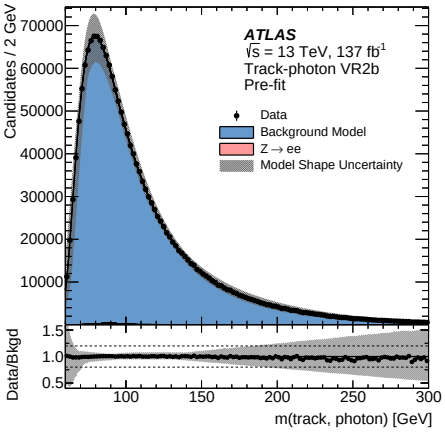


(a)

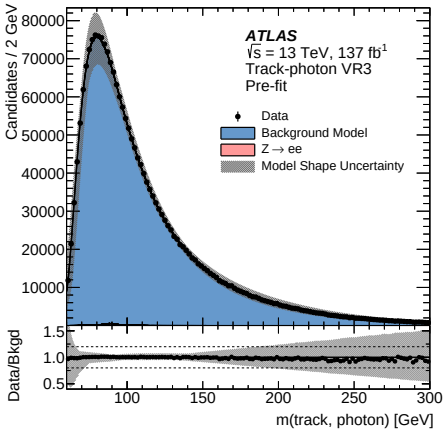
(b)



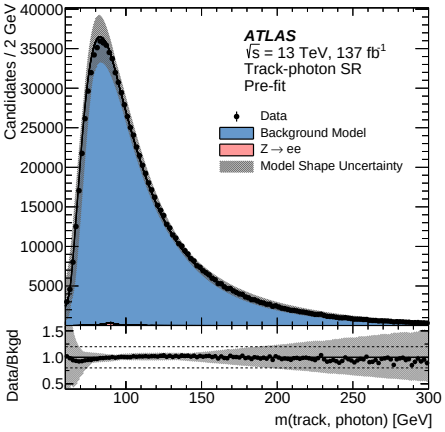
(c)



(d)



(e)



(f)

Figure 5.26: Distributions of the W boson invariant mass for data and background in the track-photon GR, VR1, VR2a, VR2b, VR3 and SR. The Background Model component includes the $Z \to e^-e^+$ contribution, which is also shown overlaid. The bottom panel shows the ratio between the data and the total background prediction. The grey band represents the shape uncertainty on the multijet background estimate.

	Inclusive Mass Range				75 - 85 GeV			
Region	$W^{\pm} \to \pi^{\pm}\gamma$	$W^{\pm} \to K^{\pm}\gamma$	Multijet bkg	$Z \to e^{-}e^{+}$	$W^{\pm} \to \pi^{\pm}\gamma$	$W^{\pm} \to K^{\pm}\gamma$	Multijet bkg	$Z \to e^{-}e^{+}$
GR	7.57 ± 0.06	0.668 ± 0.007	2797824	6611 ± 67	6.10 ± 0.05	0.556 ± 0.006	445877	1206 ± 28
VR1	6.70 ± 0.06	0.594 ± 0.006	2209285 ± 682	5583 ± 61	5.40 ± 0.05	0.493 ± 0.006	329829 ± 293	952 ± 25
VR2a	6.50 ± 0.06	0.581 ± 0.006	1911713 ± 778	5315 ± 59	5.26 ± 0.05	0.485 ± 0.006	321567 ± 299	942 ± 25
VR2b	7.28 ± 0.06	0.643 ± 0.006	1974618 ± 762	5010 ± 57	5.86 ± 0.05	0.535 ± 0.006	332721 ± 291	862 ± 23
VR3	7.10 ± 0.06	0.629 ± 0.006	2325844 ± 626	4415 ± 54	5.78 ± 0.05	0.527 ± 0.006	373578 ± 246	870 ± 23
SR	5.25 ± 0.05	0.474 ± 0.005	1069797 ± 813	2419 ± 54	4.27 ± 0.05	0.397 ± 0.005	176013 ± 326	410 ± 16

Table 5.12: Cut-flow for signal and background in the track-photon final state analysis. Yields are shown for the inclusive mass range and for signal mass window of interest.

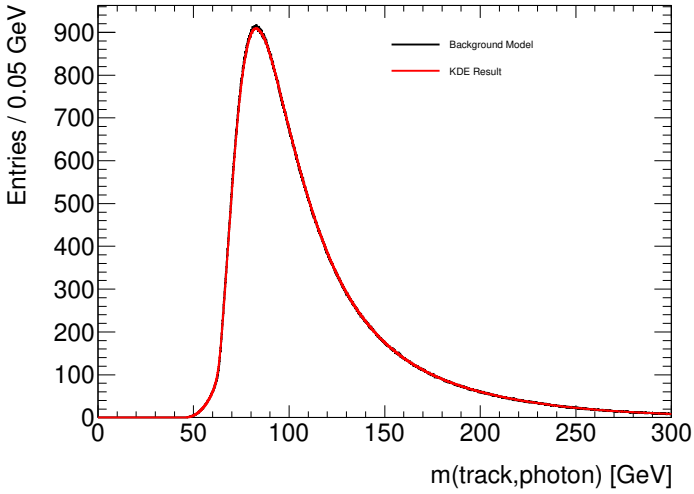


Figure 5.27: Track-photon multijet background template and KDE smoothing result.

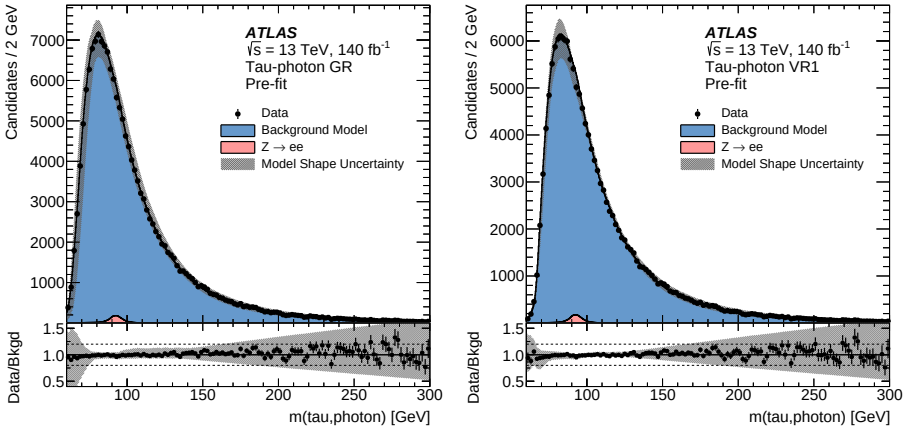
data-driven background modelling technique. In this case the variables included in the modelling at the tau and photon kinematics, $\log(|d_0(\tau_{\text{had-vis}})|)$ and $\Delta R_{\tau}^{\text{max}}$. Validation regions are defined as summarised in Table 5.8. Fig. 5.28 shows the pre-fit background prediction compared to the the data for the W boson invariant mass, for the trau-photon GR, SR and VRs. The total uncertainty on the shape of the multijet background is represented by the grey band and is derived as described in the next section. These plots show good agreement between the data and the background prediction for all regions. The number of multijet background events predicted to be in each tau-photon region is listed in Table 5.13, as well as the expected number of $W^{\pm} \to \rho^{\pm}\gamma$ signal and $Z \to e^{-}e^{+}$ events.

5.8.2 $Z \to e^{-}e^{+}$ background

The $Z \to e^{-}e^{+}$ contribution in SR is modelled in the fit to the data using the MC simulation prediction, in both final state analyses. The m(track, photon) and m(tau, photon) distributions for $Z \to e^{-}e^{+}$ MC events are smoothed through KDE, using the `Roofit` class `RooKeysPdf`. This class provides an adaptive KDE smoothing, which uses a variable smoothing parameter not user-defined. This adaptive method is used for the lower mass range of the distribution. A fixed high smoothing parameter is used for the high mass tail, where there are fewer statistics: a value of 2 is used in the track-photon

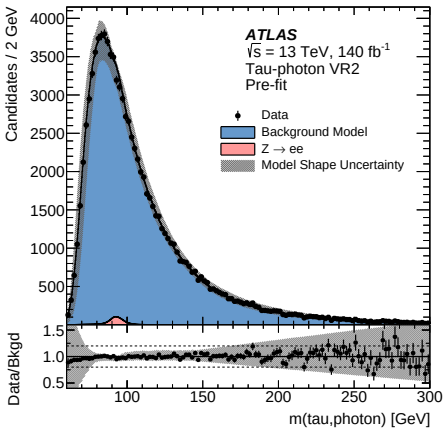
	Inclusive Mass Range			70-90 GeV		
Region	$W^{\pm} \to \rho^{\pm}\gamma$	Multijet bkg	$Z \to e^{-}e^{+}$	$W^{\pm} \to \rho^{\pm}\gamma$	Multijet bkg	$Z \to e^{-}e^{+}$
GR	1.01 ± 0.38	171735	1136 ± 20	1.00	111761	1066
VR1	0.97 ± 0.37	146606 ± 205	1059 ± 19	0.96	91034	999
VR2	0.81 ± 0.34	97728 ± 778	676 ± 15	0.80	58746	633
VR3	0.90 ± 0.36	116416 ± 194	429 ± 12	0.89	75401	386
SR	0.72 ± 0.32	67513 ± 202	193 ± 8	0.71	39263	172

Table 5.13: Cut-flow for signal and background in the tau-photon final state analysis. Yields are shown for the inclusive mass range and for a mass range around the signal mass window.

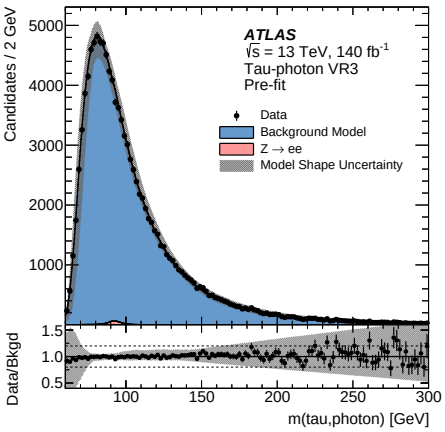


(a)

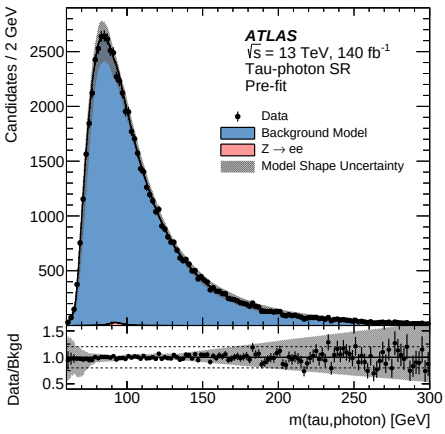
(b)



(c)



(d)



(e)

Figure 5.28: Distributions of the W boson invariant mass for data and background in the tau-photon GR, VR1, VR2, VR3 and SR. The Background Model component includes the $Z \to e^-e^+$ contribution, which is also shown overlaid. The bottom panel shows the ratio between the data and the total background prediction. The grey band represents the shape uncertainty on the multijet background estimate.

final state and 1.5 in the tau-photon case. This is done using two `RooKeysPdf` objects, that provide two smoothing predictions: KDE_{core} and KDE_{tail}. The $Z \to e^- e^+$ m(track, photon) prediction is given by

$$P(m_{trk,\gamma}) = (1 - \text{Erf}(m_{\text{trk},\gamma} - 95)) \times \text{KDE}_{\text{core}} + ((\text{Erf}(m_{\text{trk},\gamma} - 95) + 1) \times \text{KDE}_{\text{tail}}, \quad (5.6)$$

where the error functions are employed to ensure the continuity of the resulting PDF. No mirroring is used at the boundaries during the KDE estimation. The resulting model is shown in Fig. 5.29.

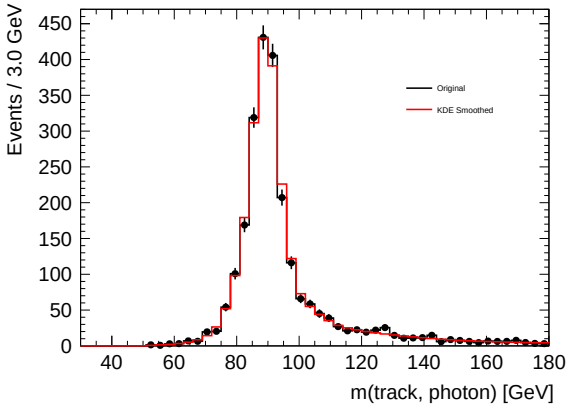


Figure 5.29: $Z \to e^- e^+$ track-photon invariant mass SR distribution, and KDE smoothing result, used to model $Z \to e^- e^+$ component in the fit to data.

In the tau-photon SR the $Z \to e^- e^+$ m(tau, photon) prediction is given by

$$P(m_{\text{tau},\gamma}) = (1 - \text{Erf}(m_{\text{tau},\gamma} - 110)) \times \text{KDE}_{\text{core}} + ((\text{Erf}(m_{\text{tau},\gamma} - 110) + 1) \times \text{KDE}_{\text{tail}}, \quad (5.7)$$

with the resulting distribution being shown in Fig. 5.30.

It is not expected that the $Z \to e^- e^+$ MC correctly models the rate of mis-reconstruction of the electrons as photons and tracks, and as such the normalisation of the $Z \to e^- e^+$ component will be left as a free parameter in the fit to the data.

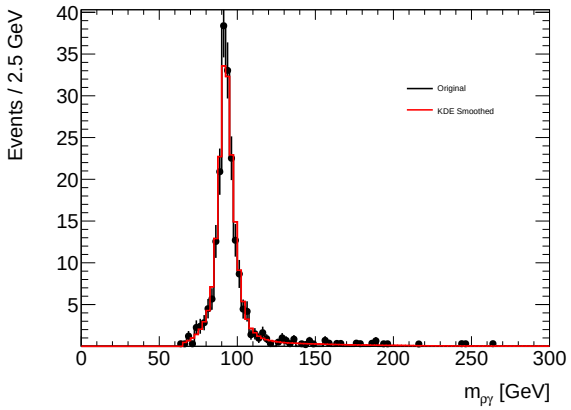


Figure 5.30: $Z \to e^- e^+$ tau+photon invariant mass SR distribution, and KDE smoothing result, used to model $Z \to e^- e^+$ component in the fit to data.

5.9 Background systematic uncertainties

To account for residual mismodelling, alternative W boson invariant mass templates are obtained (following the strategy delineated in Section 4.2.2.4) and implemented in the fit to the data using a moment morphing technique [81]. Two "up/down" pairs of approximately symmetric shapes around the nominal mass shape are generated through modifications of the background modelling procedure: the first one by shifting the $p_{\mathrm{T}}(\gamma)$ of each pseudo-event by ± 3 GeV; the second one by re-scaling the $\Delta\phi(\mathrm{trk},\gamma)/\pi$ distribution by $(1+\Delta\phi/\pi)^{10}$ (up variation) or $1+10\times(1-\Delta\phi/\pi)$ (down variation). The effect of the $p_T(\gamma)$ and $\Delta\phi(\mathrm{trk},\gamma)/\pi$ variations on the respective variables being modified can be seen in Fig. 5.31, for the track-photon background model in SR. The effects of the variations are propagated to the compound variables produced by the model, such as the W boson invariant mass. One more pair of "up/down" alternative shapes is produced by directly applying a multiplicative transformation (an overall tilt) to the $\mathrm{m}(\mathrm{track},\gamma)$ shape, using a linear function: $y = -0.0026 \times \mathrm{m}(\mathrm{track},\gamma) + 1.34$ to obtain the up variation and $y = 0.0026 \times \mathrm{m}(\mathrm{track},\mathrm{photon}) + 0.635$ to obtain the down variation. As was done for the nominal shape, the variations are produced separately for the subsets of data with different photon p_{T} requirements, and are subsequently summed providing one single template for the statistical analysis. The effect of each variation on the $\mathrm{m}(\mathrm{track},\gamma)$ is shown in Fig. 5.32, for the track-photon analysis. As was explained in Section 4.2.2.4, the pre-fit size of the variations is chosen to be large enough

to allow the corresponding nuisance parameters to be constrained in the fit to the data.

The variations are implemented in the fit using `RooStarMomentMorph`. In total, three nuisance parameters are added to the fit, $\alpha(p_{\mathrm{T}}(\gamma))$, $\alpha(\Delta\phi(\mathrm{trk},\gamma))$ and $\alpha(\mathrm{tilt})$. In the case of the $p_{\mathrm{T}}(\gamma)$ and $\Delta\phi(\mathrm{trk},\gamma)$ variations, the up and down templates are mapped to α = 1 and α = -1, respectively. The up and down variations corresponding to the $\mathrm{m}(\mathrm{M},\gamma)$ tilt are mapped to α = 2 and α = -2, for the track-photon final state analysis and α = 1 and α = -1 for the tau-photon final state analysis. The nominal W boson invariant mass template is mapped to α = 0 for all variations.

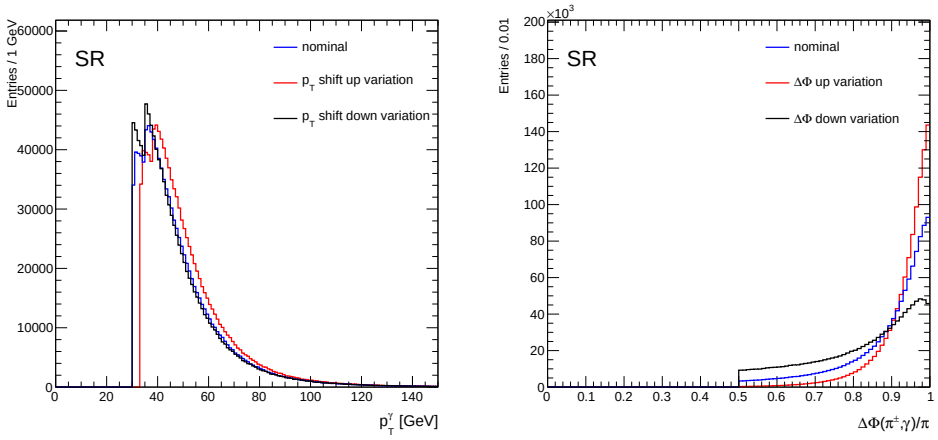


Figure 5.31: Effect of the the $p_T(\gamma)$ shift variation on the $p_T(\gamma)$ distribution (left) and effect of the $\Delta\phi(\mathrm{track},\gamma)$ tilt on the $\Delta\phi(\mathrm{track},\gamma)$ distribution (right).

5.10 Signal modelling

Track-photon final state

The signal is modelled using the MC simulated samples described in Section 5.1.2. The $\mathrm{m}(\mathrm{track},photon)$ shape of $W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma$ is modelled in the fit to the data using the product between the sum of two Voigtian functions and an efficiency curve. A Voigtian function is defined as the convolution between a Breit-Wigner and a Gaussian. The Breit-Wigner distribution models the W boson mass resonance and depends on the W natural width (Γ). The Gaussian function models the experimental resolution (σ). The two Voigtian functions have the same mean (μ) and Γ, but different Gaussian widths. In

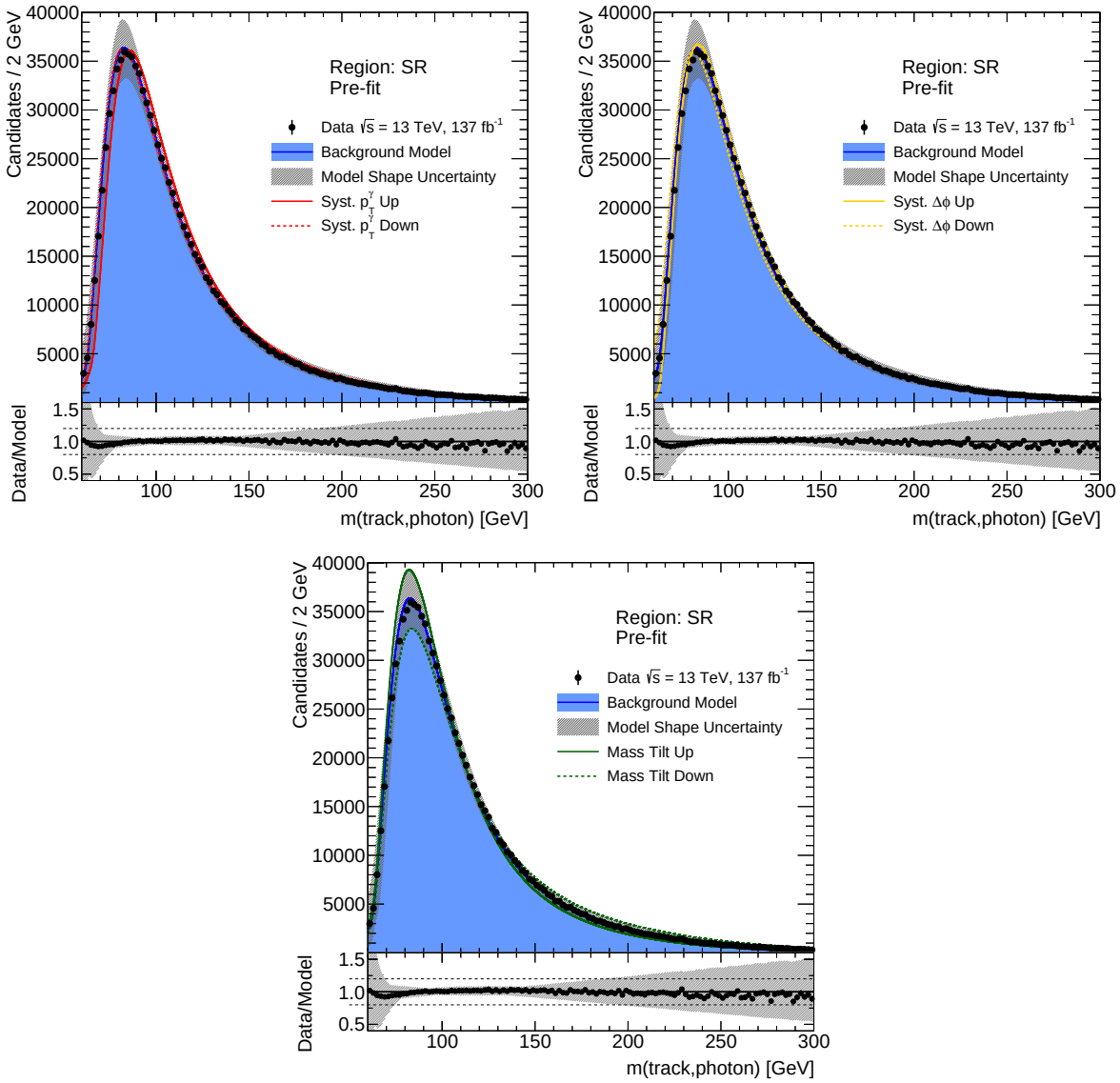


Figure 5.32: Effect of the systematic variations on $m(\text{track}, \gamma)$.The systematic uncertainty band on the background represents the maximum deviation of all the alternative shapes from the nominal model.

summary, the $W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma$ signal shape in the W boson invariant mass is modelled by:

$$\varepsilon(\mathrm{m}) \times [\, f \ \mathrm{Voigt}_1(\mathrm{m}|\mu, \sigma_1, \Gamma = \Gamma_W) + (1-f)\ \mathrm{Voigt}_2(\mathrm{m}|\mu, \sigma_2, \Gamma = \Gamma_W)\ ], \tag{5.8}$$

where the Γ width is fixed to the W natural width (2.085 ± 0.042 GeV) [15], f determines the weight of each Voigtian in the sum, and m represents the W boson invariant mass. The efficiency function $\varepsilon(\mathrm{m})$ accounts for the experimental acceptance to the SR requirements. This efficiency is derived as a function of the W boson invariant mass at generator level, and can be found in Fig. 5.33, for $W^{\pm} \to \pi^{\pm}\gamma$, in the left. $\varepsilon(\mathrm{m})$ is obtained by performing a cubic spline interpolation to the efficiency distribution. The values of the μ, σ_1, σ_2 and f parameters are obtained after performing a fit to the reconstructed W boson invariant mass distribution for $W^{\pm} \to \pi^{\pm}\gamma$ MC events.

Fig. 5.33 shows the efficiency function for $W^{\pm} \to \pi^{\pm}\gamma$, in the track-photon SR, and the comparison with the efficiency for $W^{\pm} \to K^{\pm}\gamma$, after normalising the two distributions to the same integral. The two efficiency curves are compatible within statistical uncertainties. Fig. 5.34 shows the reconstructed W boson invariant mass distribution for both $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$. As can be seen the two distributions are consistent within statistical uncertainties, with just small differences for the high mass tail. Therefore, the same $\mathrm{m}(\mathrm{track}, \gamma)$ shape is used to model the $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ signal processes.

The results of the fit to the W invariant mass distribution for the $W^{\pm} \to \pi^{\pm}\gamma$ signal MC are summarised in Table 5.14. The resulting $\mathrm{m}(\mathrm{track}, \gamma)$ shape is shown Fig. 5.35. The W boson mass resolution for $W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma$ in the track-photon final state is 2.7%.

Table 5.14: Signal model parameter values obtained from fit to $W^{\pm} \to \pi^{\pm}\gamma$ MC W boson invariant mass distribution.

Parameter	Post-fit value
m	80.43 ± 0.03 GeV
σ_1	4.38 ± 0.02 GeV
σ_2	1.68 ± 0.04 GeV
f	0.19 ± 0.01

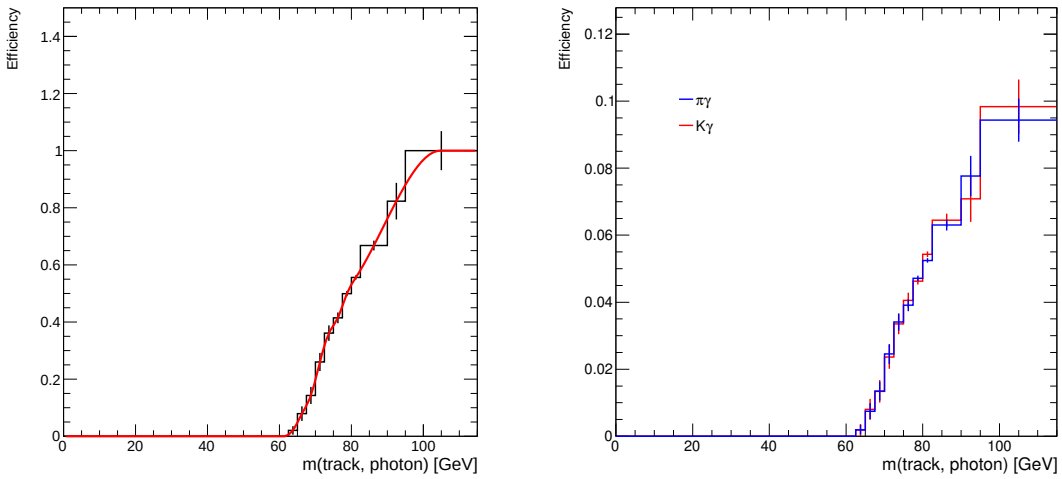


Figure 5.33: Signal efficiency as function of the track-photon W boson invariant mass at generator level for $W^{\pm} \to \pi^{\pm}\gamma$ (left), and $W^{\pm} \to K^{\pm}\gamma$ compared to $W^{\pm} \to \pi^{\pm}\gamma$ (right).

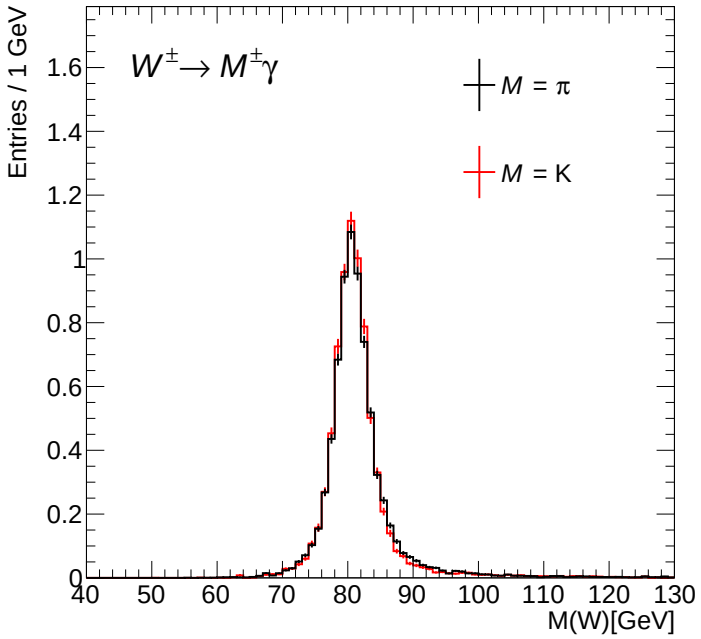


Figure 5.34: W boson invariant mass at reconstructed level, for $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$. The $W^{\pm} \to K^{\pm}\gamma$ distribution is normalised to the integral of the $W^{\pm} \to \pi^{\pm}\gamma$ distribution.

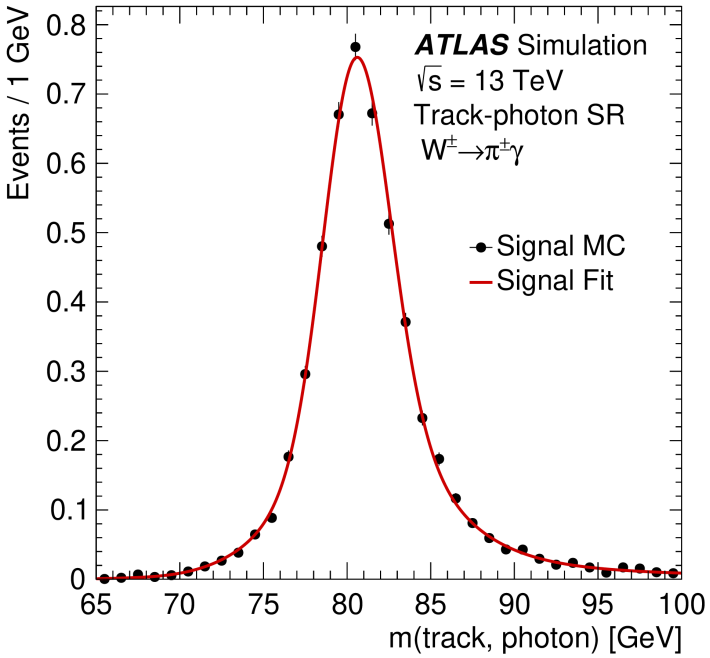


Figure 5.35: Result of fit to W boson invariant mass distribution of $W^{\pm} \to \pi^{\pm}\gamma$ MC events, using the model described by Eq. (5.8).

In the case of the $W^{\pm} \to \rho^{\pm}\gamma$ decay in the track-photon final state, the π^0 is not reconstructed, the track-photon W boson invariant mass distribution is shifted to lower values and the resolution deteriorates. The $W^{\pm} \to \rho^{\pm}\gamma$ signal is thus not modelled in the same way as the $W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma$ signals, and instead its shape is obtained by smoothing the distribution of MC events using KDE. As was done for the modelling of the $Z \to e^{-}e^{+}$ background, two `RooKeysPdf` objects were used: one employs the adaptive method for the smoothing prediction and is used for the core of the invariant mass distribution (KDE_{core}); while the other applies a fixed high smoothing parameter of 2, and is used for the tail of the distribution, where there are fewer events, and the adaptive mode was found to not provide a smooth enough distribution (KDE_{tail}). The $W^{\pm} \to \rho^{\pm}\gamma$ prediction as a function of $m_{\text{track},\gamma}$ is given by:

$$P(m_{\text{track},\gamma}) = (1 - \text{Erf}(m_{\text{track},\gamma} - 115)) \times \text{KDE}_{\text{core}} + ((\text{Erf}(m_{\text{track},\gamma} - 115) + 1) \times \text{KDE}_{\text{tail}}, \tag{5.9}$$

where the error functions are employed to ensure the continuity of the resulting PDF. No mirroring is used at the boundaries during the KDE estimation. The resulting

$W^{\pm} \to \rho^{\pm}\gamma$ signal model is shown in Fig. 5.36. The Gaussian width of this curve is estimated from the full width at half maximum of the distribution (FWHM), by subtracting Γ_W in quadrature. The resulting estimated mass resolution is 3.1%.

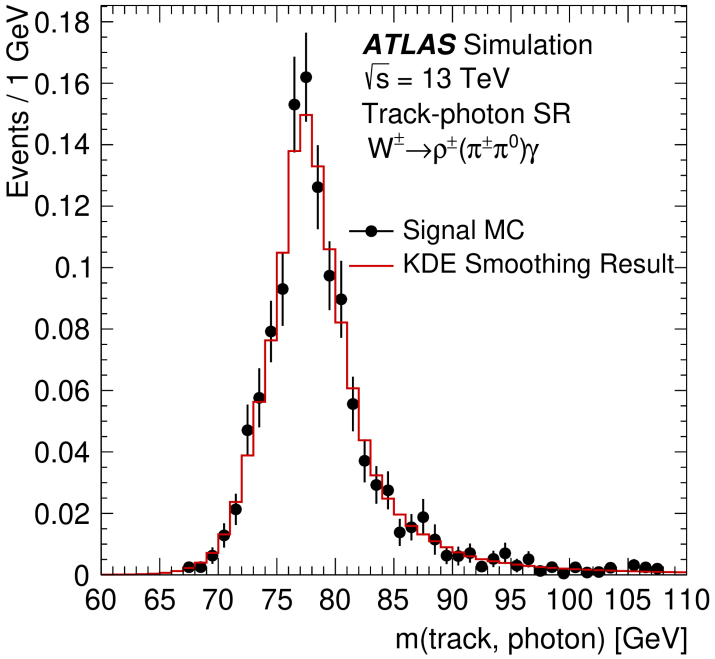


Figure 5.36: $W^{\pm} \to \rho^{\pm}\gamma$ track-photon invariant mass distribution and the result of applying KDE smoothing to this distribution.

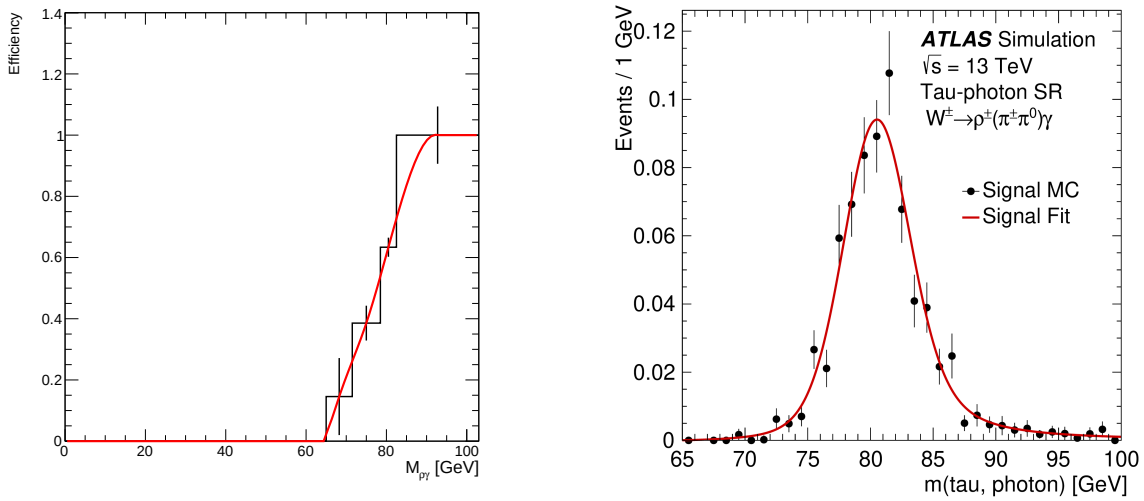
Tau-photon final state

The m(tau, γ) shape of $W^{\pm} \to \rho^{\pm}\gamma$ signal is modelled in the fit to the data in the tau-photon SR using the product between a Voigtian function and an efficiency curve. As in the track-photon final state, the Γ width of the Breit-Wigner is fixed to the W natural width (2.085 $\pm$ 0.042 GeV) [15]. The $W^{\pm} \to \rho^{\pm}\gamma$ tau-photon invariant mass shape is described in the final fit using

$$\varepsilon(\mathrm{m}) \times \mathrm{Voigt}(\mathrm{m}|\mu, \sigma, \Gamma = \Gamma_W), \tag{5.10}$$

where m represents the W boson invariant mass, σ the gaussian width and μ the mean of the Voigtian. $\varepsilon(\mathrm{m})$ is derived as a function of the W boson invariant mass at generator

level, as was done for the $W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma$ modelling in the track-photon final state. The resulting efficiency curve is shown in Fig. 5.37(a). It models the effect of the tau-photon SR selection. The values of the μ and σ parameters are obtained from a fit to the tau-photon invariant mass distribution of $W^{\pm} \to \rho^{\pm}\gamma$ MC events. The result of this fit is shown in Fig. 5.37(b). The post-fit values of the parameters are $\mu = 80.3 \pm 0.2$ GeV and $\sigma = 2.3 \pm 0.2$ GeV. The latter corresponds to a mass resolution of 2.9%.



(a) (b)

Figure 5.37: (a) Signal efficiency as function of the tau-photon W boson invariant mass at generator level for $W^{\pm} \to \rho^{\pm}\gamma$; (b) Result of fit to W boson invariant mass distribution of $W^{\pm} \to \rho^{\pm}\gamma$ MC events, using the model described by Eq. (5.10).

5.11 Signal systematic uncertainties

Different systematic sources can affect the normalisation and shape of the signal distributions. This section presents the signal systematic uncertainties in the track-photon final state in detail. The signal systematic uncertainties in the tau-photon final state are also briefly discussed and summarised inTable 5.18. For each relevant signal systematic uncertainty, an associated nuisance parameter is added to the fit to data. Ultimately, the shape components of the signal uncertainties were found to be negligible and only the normalisation components were taken into account in the fit. Statistical MC uncertainties are also found negligible compared to the uncertainty on the background,

and therefore no bin-by-bin systematics are added to the fit. Given the differences in reconstruction, trigger and selection between the two final states, the effect of most uncertainties is estimated independently for the two final states. In the track-photon final state analysis, signal uncertainties were estimated using the $W^{\pm} \to \pi^{\pm}\gamma$ MC, with the same uncertainty being assigned to the $W^{\pm} \to \rho^{\pm}\gamma$ signal.

The uncertainties on the ATLAS luminosity and W boson cross section measurements result on an uncertainty on the signal yield expectation for both final states. The uncertainty on the integrated luminosity of the full Run 2 proton-proton dataset is 0.83% [47], while the uncertainty on the W boson production cross section is 3.3% [25].

Theoretical modelling uncertainties are also estimated for both analysis final states. They are associated with high order terms of the perturbative expansion of the inclusive $pp \to W$ cross section which are neglected in its calculation. To estimate these, the factorisation (μ_F) and renormalisation scales (μ_R) used in the generation of the signal MC samples are multiplied and divided by a factor of two. Seven variations were considered: $\{\mu_R, \mu_F\} = \{0.5, 0.5\}, \{1.0, 0.5\}, \{0.5, 1.0\}, \{2.0, 1.0\}, \{2.0, 2.0\}, \{2.0, 1.0\}$. For each of these variations, alternative $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to \rho^{\pm}\gamma$ samples were generated, normalised to the W boson production cross section [25]. No reconstruction simulation was performed for these samples. The SR kinematic requirements were applied for each sample at generator level. In the case of the track-photon analysis events were required to meet the following requirements:

- $p_{\mathrm{T}}(\pi^{\pm}) > 33$ GeV;
- $|\eta(\pi^{\pm})| < 2.5$;
- $p_{\mathrm{T}}(\gamma) > 30$ GeV (or 35 GeV for 20% of the events);
- $|\eta(\gamma)| < 2.37$ excluding $1.37 < |\eta(\gamma)| < 1.52$;
- $\Delta\phi(\pi^{\pm}, \gamma) > \pi/2$.

The envelope of all the variations, meaning the maximum difference between the number of events in the different alternative samples and the number of events in the nominal sample, was obtained as a function of the W boson truth mass, and can be seen in Fig. 5.38, for $W^{\pm} \to \pi^{\pm}\gamma$ events. The average relative difference with respect to

nominal is taken as the renormalisation and factorisation scale systematic uncertainty, and corresponds to 6.2%. Since the distribution of these differences is approximately flat, only a normalisation systematic is taken into account in the fit to the data.

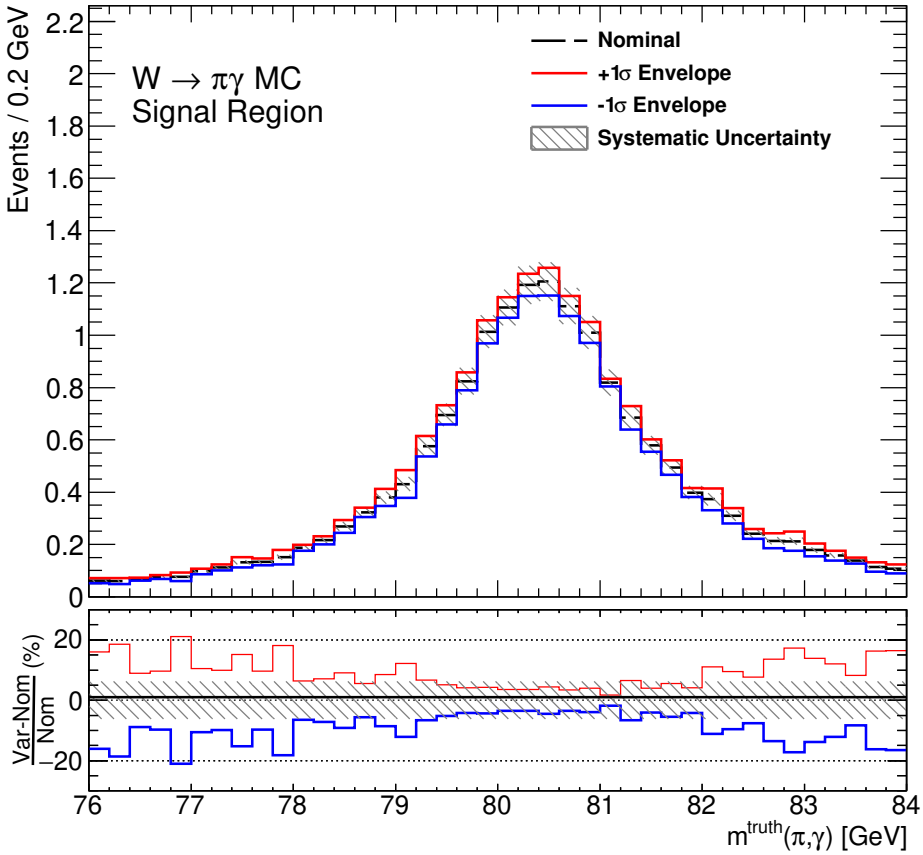


Figure 5.38: Envelope of the factorisation and renormalisation scale variations as a function of the W boson truth mass. The nominal W boson truth mass distribution is represented by the black and dashed line. The blue and red full lines show the envelope of the variations. The bottom panel shows the relative difference from the envelope with respect to the nominal value. The grey band shows the resulting normalisation systematic uncertainty (6.21%).

The uncertainty associated with the pile-up correction applied to the signal MC samples is also taken into account. Events are re-weighted using alternative pileup scale factors corresponding to the $\pm 1\sigma$ variations on the nominal scale factors, and the signal expectation is recalculated. The relative difference from the nominal signal expectation is 2.2% in the track-photon analysis, and is taken as the pileup correction uncertainty.

The impact of uncertainties on the energy calibration and resolution of photons is

evaluated by applying 70 different variations to the photon four momentum vector. Each variation allows to evaluate a different component of the total energy scale and resolution uncertainties. Variations related to the photon energy scale account for [128]: the statistical and systematic uncertainties on the data-to-MC calibration using $Z \to e^-e^+$ decays; uncertainties on the calibration of the layers of the EM calorimeter and of the barrel-endcap gap scintillator; uncertainties associated with the calibration of the readout gains of the EM calorimeter electronics; uncertainties on the knowledge of the detector material between the interaction point and the calorimeter; uncertainties associated with the dependence of the calorimeter's cells energy response on the EM shower width in the η direction; systematics on the extrapolation from the electron-based calibration with $Z \to e^-e^+$ decays to photons, due to differences in lateral shower energy leakage; and uncertainties associated the modelling of the efficiency of the classification of photons as unconverted/converted. Most of the listed sources of uncertainty on energy scale are also relevant for electrons, and as such are classified as electron/γ (EG) scale uncertainties, while the latter two are specific for photons and are classified as photon scale uncertainties. After each of the photon four momentum variations are applied, the resulting $W^\pm \to \pi^\pm\gamma$ distributions are obtained in the track-photon SR and compared to the nominal distribution. Relative differences with respect to nominal are calculated for signal yield, mean and width. The uncertainties resulting from each independent variation are then summed in quadrature, and the total EG and photon scale uncertainties are obtained, listed in Table 5.15. The same procedure is followed to evaluate the effect of the uncertainties associated with the photon energy resolution. The contributions to the total photon resolution uncertainty are the uncertainties on the effects of pile-up and electronics noise; the uncertainty associated with the MC smearing based on $Z \to e^-e^+$ events to match the resolution observed in data; the uncertainty associated with energy losses in the detector material upstream of the EM calorimeter; and uncertainties on the intrinsic energy resolution due to the calorimeter sampling fluctuations [128]. The total effects of the EG resolution uncertainty on the normalisation, mean and width of the $W^\pm \to \pi^\pm\gamma$ signal are listed in Table 5.15. The estimated overall impact of the energy scale and resolution uncertainties on the signal yield and shape is found to be negligible (<1%) and as such

no associated nuisance parameter is added to the fit to data.

Table 5.15: Total estimated systematic uncertainties associated with the EG and photon scale and resolution on the $W^{\pm} \to \pi^{\pm}\gamma$ signal distribution.

Systematic	Signal Yield (%)	Mean (%)	Width (%)
EG Resolution	0.09	0.03	0.80
EG Scale	0.15	0.11	0.91
Photon Scale	0.03	0.05	0.24

Scale factors that correct simulation to match the photon identification and isolation efficiencies observed in data are applied to the signal MC samples, as explained in Section 5.1.2. The effect of the uncertainty on these scale factors on the signal yield in the track-photon SR is estimated to be 2.1%. The photon identification and isolation components are combined by being summed in quadrature.

The uncertainty on the track reconstruction efficiency also affects the expected signal yield. The sources of this uncertainty are the knowledge of the ID material [53] and the modelling of the rate of reconstruction of fake tracks. The impact of this uncertainty on the track-photon analysis is evaluated by applying five different variations that randomly remove tracks from the simulated events, depending on $\eta(\text{trk})$ and $p_{\text{T}}(\text{trk})$, and calculating the resulting relative signal yield difference with respect to nominal. Each of the five variations evaluate uncertainties associated with: the overall knowledge of the ID material; the knowledge of material of the IBL; the knowlegedge of the pixel service region (composed of cables and cooling pipes located between the pixel and SCT detectors); the physics model used in the Geant4 simulation of hadronic interactions in the ID material; and the fake rate modelling. The different contributions are summed in quadrature, yielding an estimated uncertainty due to track reconstruction efficiency of 1.2%.

No scale factors are available to correct the modelling of the dedicated track-photon trigger efficiency in the signal MC samples. Nevertheless, the photon component of the triggers utilises the standard ATLAS photon trigger algorithm and as such the associated scale factors were used, accounting for the photon leg of the trigger. The uncertainty on these photon trigger scale factors is estimated to be 0.6%. A dedicated study was performed to estimate the uncertainty associated with the efficiency of the

track leg of the trigger. It is expected that the largest component of this uncertainty arises from the modelling of the $E_T(\text{tau})/p_T(\text{trk})$ variable used in the trigger, since the largest efficiency loss comes from the requirement applied on it. A recent ATLAS publication [129] studied the energy response of the ATLAS calorimeter to charged pions presenting data-MC comparisons of $E_T(\text{tau})/p_T(\text{trk})$ as a function of $p_{\text{T}}(\text{trk})$ and $\eta(\text{trk})$. Ratios between the mean and width of the distribution of this variable in data and MC are provided in the aforementioned publication, and shown in Fig. 5.39, for tracks with $30 < p_{\text{T}}(\text{trk}) < 50$ GeV. Most of the tracks selected in the track-photon SR are within this p_{T} range. The $E_T(\text{tau})/p_T(\text{trk})$ of $W^{\pm} \to \pi^{\pm}\gamma$ MC events is obtained through ΔR matching of taus in each event to the selected tracks. The distribution of this variable for events that survive offline selection is corrected using the ratios presented in Ref. [129]. Depending on $\eta(\text{trk})$, the value of $E_T(\text{tau})/p_T(\text{trk})$ is corrected according to the expression

$$\frac{E_T(\text{tau})}{p_T(\text{trk})} \to \frac{E_T(\text{tau})}{p_T(\text{trk})} \times \frac{\mu_{\text{Data}}}{\mu_{\text{MC}}} + A, \tag{5.11}$$

in which A is a number sampled from a gaussian of the form

$$\text{Gaus}\left(0, \sigma_{\text{MC}} \times \sqrt{\left(\frac{\sigma_{\text{Data}}}{\sigma_{\text{MC}}}\right) - 1}\right). \tag{5.12}$$

$\frac{\mu_{\text{Data}}}{\mu_{\text{MC}}}$ and $\frac{\sigma_{\text{Data}}}{\sigma_{\text{MC}}}$ are the ratios from the corresponding $\eta(\text{trk})$ bin in Fig. 5.39. Fig. 5.40 shows the comparison of the $E_T(\text{tau})/p_T(\text{track})$ variable before applying the correction (nominal) and after it has been applied (variation). The efficiency difference to the $0.4 < E_T(\text{tau})/p_T(\text{track}) < 0.85$ requirement is -3.6% (relative to nominal). This value is the estimated uncertainty on the trigger efficiency. The signal normalisation is corrected accordingly before the fit, yielding the expected number of signal events listed in Table 5.16.

Table 5.16: Number of signal events after correcting for the $E_T(\text{tau})/p_T(\text{track})$ data-MC discrepancies.

Channel	Before correction	After correction
$W^{\pm} \to \pi^{\pm}\gamma$	5.25	5.06
$W^{\pm} \to K^{\pm}\gamma$	0.47	0.45
$W^{\pm} \to \rho^{\pm}\gamma$	1.24	1.20

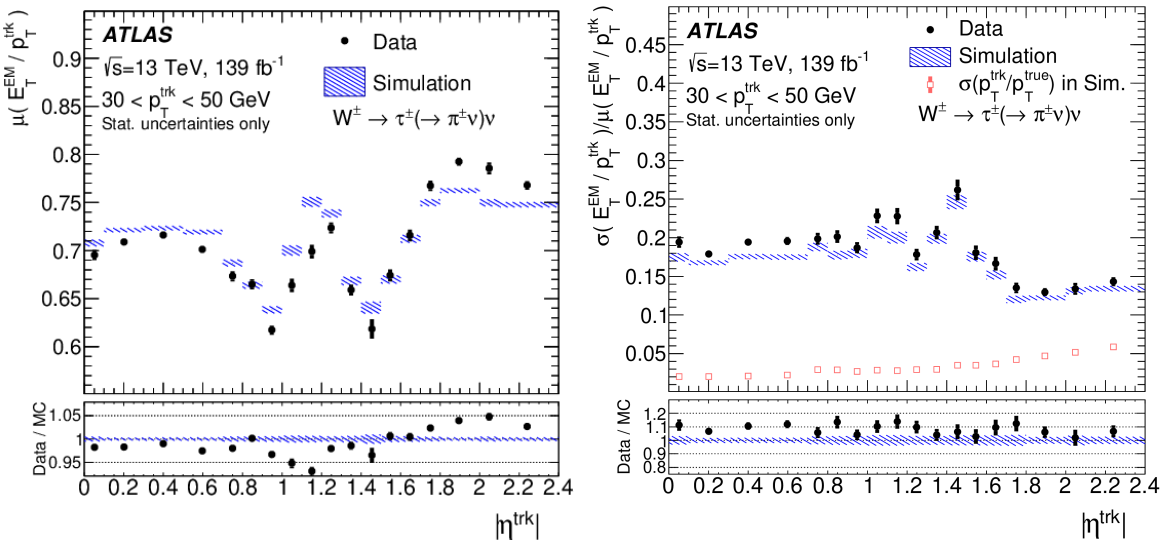


Figure 5.39: Data-MC comparisons of the mean and width of the $E_T(\text{tau})/p_T(\text{trk})$ distribution [129].

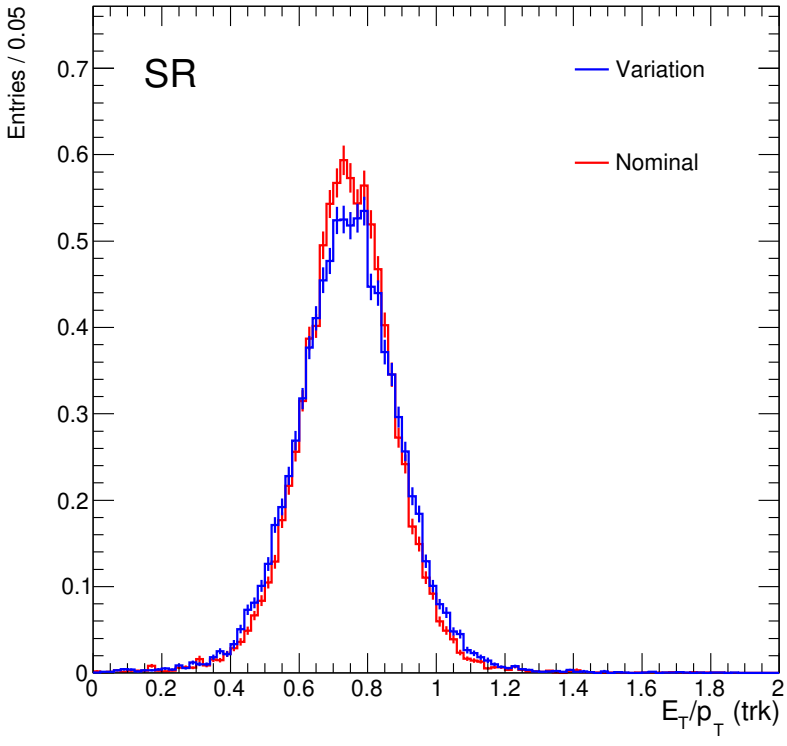


Figure 5.40: $E_T(\text{tau})/p_T(\text{trk})$ before and after correction.

Table 5.17 and Table 5.18 summarise the signal systematic uncertainties in the track-photon and tau-photon final state analyses, respectively. Uncertainties associated with photon ID and isolation, photon scale and resolution, pile-up and factorisation and renormalisation scales are estimated in the tau-photon analysis in the same way as described for the track-photon analysis. In the case of the trigger uncertainty a dedicated study of the photon trigger response to taus is performed. With this goal, a `HLT_g25_medium_mu24` trigger is used to select $Z \to \tau\tau$ events. This trigger uses the same photon selection as the di-photon trigger used in the tau-photon analysis. The trigger response is compared in data and MC, with the MC modelling the photon trigger efficiency within 10%. Uncertainties associated with the efficiency of the tau reconstruction, the tau identification RNN, the electron-tau BDT discriminator used in the selection, and the tau energy scale are also estimated. They have an estimated combined effect on the $W^{\pm} \to \rho^{\pm}\gamma$ signal yield of 13%.

Table 5.17: Signal systematic uncertainties in the track-photon final state, taken into account in the fit to data. These are calculated in terms of the expected signal yield.

Source	Uncertainty (%)
Photon ID/Isolation Efficiency	2.1
Track Reconstruction Efficiency	1.2
Trigger Eficiency	3.6
Renormalisation and Factorisation Scale	6.2
Pileup	2.2
Luminosity	0.8
Cross Section	3.3

5.12 Statistical analysis and results

The presence of a signal is quantified through a binned extended maximum likelihood fit to the reconstructed W boson invariant mass distribution. The parameters of interest of the fit are the $W^{\pm} \to \pi^{\pm}\gamma$, $W^{\pm} \to K^{\pm}\gamma$ and $W^{\pm} \to \rho^{\pm}\gamma$ signal strengths (μ). Fits were performed independently in the track-photon and tau-photon SRs. A simultaneous fit including both SRs was also performed, allowing to better constrain $\mu(W^{\pm} \to \rho^{\pm}\gamma)$. The $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to K^{\pm}\gamma$ signal processes cannot be disentangled, as they have

Table 5.18: Signal systematic uncertainties in the tau-photon final state, taken into account in the fit to data. These are calculated in terms of the expected signal yield.

Source	Uncertainty (%)
Photon ID/Isolation Efficiency	1.9
EG Scale	3.0
EG Resolution	4.9
Photon Scale	1.7
Tau Reconstruction and ID Efficiency	13
Trigger Eficiency	10
Renormalisation and Factorisation Scale	6.5
Pileup	5.5
Luminosity	0.8
Cross Section	3.3

the exact same reconstructed W boson invariant mass shape. Nevertheless, the two signals have different efficiencies and therefore the choice was made to consider only one process, while the other signal strength is assumed to be zero. The normalisations of the background processes are free parameters of the fit, since the $Z \to e^-e^+$ MC is not expected to correctly model the rate of mis-reconstruction of electrons as photons and tracks; and the pre-fit prediction for the number of multijet events in the SR is extrapolated from the number of events in the GR. The PDFs used to model the $W^\pm \to \pi^\pm\gamma / W^\pm \to K^\pm\gamma$ signal in the track-photon SR and the $W^\pm \to \rho^\pm\gamma$ signal in the tau-photon SR are the ones described by Eq. (5.8) and Eq. (5.10), respectively. As already explained, the parameters of these functions are fixed to values obtained from the fit to the W boson invariant mass distribution of MC events. The PDF of the $W^\pm \to \rho^\pm\gamma$ signal in the track-photon final state is obtained after KDE smoothing is applied to the W boson invariant mass distribution of $W^\pm \to \rho^\pm\gamma$ MC events (Fig. 5.36). In practise, a finely binned histogram is the result of the KDE smoothing procedure, and the PDF used in the fit is obtained through linear interpolation of this finely binned histogram. The $Z \to e^-e^+$ background PDF is obtained in the same way. The multijet background is modelled in the fit by the PDF resulting from moment morphing between different PDFs, each corresponding to the background shape variations described in Section 5.9. Each of these shape variations is also KDE smoothed and the corresponding PDF is obtained from linearly interpolating the resulting finely binned histogram. The

morphing between variations is controlled by three background nuisance parameters, $\boldsymbol{\alpha}^{\text{bkg}} = \{\alpha(p_{\text{T}}(\gamma)), \alpha(\Delta\phi(\text{M}, \gamma)), \alpha(\text{tilt})\}$. While the first two parameters are constrained in the fit by a Gaussian, the latter is left unconstrained. The systematic uncertainties on the signal normalisation are also controlled in the fit by nuisance parameters, $\boldsymbol{\alpha}^{\text{sig}}$, all with an associated Gaussian constraint. There are 7 signal nuisance parameters in the track-photon fit, one for each uncertainty listed in Table 5.17. In the tau-photon case, there are 22 signal nuisance parameters, given that the 12 uncertainties which contribute to the combined tau efficiency uncertainty presented in Table 5.18 were treated independently in the fit. Given the differences in reconstruction, trigger and selection between the two final states, most uncertainties were treated in an uncorrelated manner in the combined fit including both final states. Only the uncertainties associated with the luminosity and W boson cross section were correlated in the combined fit.

For a general case, the likelihood in a binned extended maximum likelihood fit is given by

$$L = \prod_{i=1}^{B} \text{Pois}(n_i; \nu_i) = \prod_{i=1}^{B} e^{-\nu_i} \frac{\nu_i^{n_i}}{n_i!}, \tag{5.13}$$

where n_i is the number of observed events in bin i, ν_i is the number of expected events in bin i and B the number of considered bins [130]. Therefore, the likelihood is the product of the bin-by-bin probabilities of observing n_i events. These probabilities fluctuate around the expected number of events ν_i, following a Poisson distribution. For each bin, $\nu_i = \sum_p P_{ip}(x; \boldsymbol{\alpha})\nu_p$, in which p runs over the different signal and background processes. ν_p denotes the expected total number of events for a specific process and P_{ip} denotes the probability that an event of process p is found in bin i, which is a function of the observable x and a set of parameters $\boldsymbol{\alpha}$. For each bin, P_{ip} is calculated by integrating the probability density function $f_p(x; \boldsymbol{\alpha})$ corresponding to process p, in bin i,

$$P_{pi}(x; \boldsymbol{\alpha}) = \int_{x^{\text{low}}}^{x^{\text{up}}} f_p(x; \boldsymbol{\alpha})\, dx. \tag{5.14}$$

In the case at hand, for a background process, ν_b can be written as

$$\nu_b = \mu_b \times N_b, \tag{5.15}$$

while for a signal process, ν_s is

$$\nu_s = \mu_s \times N_s \times \prod_k \left(1 + \alpha_k^{\text{sig}}\right). \tag{5.16}$$

N_s and N_b are the pre-fit numbers of expected signal and background events in SR, respectively; and μ_s and μ_b are the signal and background strengths for processes s and b. For the track-photon final state $s = \{W^\pm \to \pi^\pm\gamma / W^\pm \to K^\pm\gamma, W^\pm \to \rho^\pm\gamma\}$, while for the tau-photon final state s can only be the $W^\pm \to \rho^\pm\gamma$ process. For both final states, $b = \{\text{Multijet}, Z \to e^-e^+\}$. The number of expected signal events can be modified by the signal normalisation uncertainties $\boldsymbol{\alpha}^{\text{sig}}$. The index k runs through the signal uncertainties included in the considered final state. The PDFs of each process are the ones already described, which are function of the W boson invariant mass, all normalised to unity: $f_s(m)$ for the signal processes; $f_{Z\to ee}(m)$ for $Z \to e^-e^+$; and $f_{\text{Multijet}}(m; \boldsymbol{\alpha}^{\text{bkg}})$ for the multijet background, which can be modified by the background shape parameters $\boldsymbol{\alpha}^{\text{bkg}}$. Putting everything together, for each final state considered, one can write the number of expected events in each bin ν_i as

$$\nu_i = \sum_s \left(\mu_s \times N_s \times \prod_k \left(1 + \alpha_k^{\text{sig}}\right) \times P_{s,i}(m)\right) + \mu_{\text{Multijet}} \times N_{\text{Multijet}} \times P_{\text{Multijet},i}(m; \boldsymbol{\alpha}^{\text{bkg}}) + \mu_{Z\to ee} \times N_{Z\to ee} \times P_{Z\to ee,i}(m). \tag{5.17}$$

Adding also the Gaussian constraints on the nuisance parameters (not including $\alpha(\text{tilt})$), with ν_i given by Eq. (5.17), for each final state, the likelihood can be written as

$$L = \prod_{i=1}^{B} e^{-\nu_i} \frac{\nu_i^{n_i}}{n_i!} \times \prod_k \text{Gaus}(\boldsymbol{\alpha}^{\text{sig}}; 0, \boldsymbol{\sigma}_{\boldsymbol{\alpha}}^{\text{sig}}) \times \prod_h \text{Gaus}(\boldsymbol{\alpha}^{\text{bkg}}; 0, \boldsymbol{\sigma}_{\boldsymbol{\alpha}}^{\text{bkg}}). \tag{5.18}$$

In practise, the nuisance parameters controlling the signal systematics are renormalised, such that σ_α corresponds to 1 for each signal uncertainty.

The fit was performed in the 60-110 GeV range, where the data-driven background model was found to provide a more robust prediction of the multijet background shape. Appendix E details results obtained when varying the fitting range in the track-photon final state. A discussion on the final choice of fitting range is also included. The pre-fit

number of expected events in the considered range for the track-photon and tau-photon SRs can be found in Table 5.19.

Table 5.19: Pre-fit number of expected events in both SRs.

Process	Track-photon	Tau-photon
Multijet bkg	648237	43133
$Z \to e^- e^+$	2064	179
$W^\pm \to \pi^\pm \gamma$	4.98	——
$W^\pm \to K^\pm \gamma$	0.45	——
$W^\pm \to \rho^\pm \gamma$	1.18	0.72

Upper limits on the branching fractions of the targeted decays were set at 95% CL, using the modified frequentist CL_s method [131], and the asymptotic approximation of the profile likelihood ratio test statistic, detailed in Ref. [132]. The profile likelihood ratio is defined as

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}, \tag{5.19}$$

where μ represents a given parameter of interest and $\boldsymbol{\theta}$ represents a set of nuisance parameters. The denominator is the global maximised likelihood, with $\hat{\mu}$ and $\hat{\theta}$ being the unconditioned values of the respective parameters that maximize the likelihood. Conversely, the likelihood in the numerator is calculated for a fixed μ value, with $\hat{\hat{\boldsymbol{\theta}}}$ being the set of values that maximise the likelihood, for that given μ.

Expected upper limits are obtained using an Asimov dataset, which is defined as the dataset generated from the defined model, for a given set of parameter values, such that when a maximum likelihood fit is performed using the same model, the ML estimators of the parameters are their true values. An initial validation of the fitting framework was performed using an Asimov dataset built from the pre-fit background expectation. Likelihood scans of the signal strength parameters were performed for each independent final state fit configuration and for the simultaneous fit, and preliminary upper limits on the branching fractions were derived including different systematics, allowing to study their effect on the analysis sensitivity. A background-only fit on sideband data was then performed to determine the values of the background nuisance parameters. A post-sideband fit Asimov was constructed with the background parameters fixed to their post-sideband fit values and the expected upper limits were re-derived. The

following subsections detail the results from the initial validation using the pre-sideband fit Asimov, the results of the fit to sideband data and the expected upper limits obtained using the post-sideband fit Asimov. Finally the fit to the full, unblinded, dataset is presented, as well as the observed 95% CL upper limits on the branching fractions.

5.12.1 Fit validation studies with pre-fit background Asimov dataset

Before fitting the data, the fitting framework was validated by performing a signal-plus-background fit to an Asimov dataset constructed from the pre-fit background expectation, with the background nuisance parameters set to their nominal values: $\mu_s = 0$, $\mu_b = 1$ and $\{\alpha(p_{\mathrm{T}}(\gamma)), \alpha(\Delta\phi(\mathrm{M}, \gamma)), \alpha(\mathrm{tilt})\} = \{0, 0, 0\}$. A fit to this dataset should return exactly these parameter values. An independent fit was performed for each final state, as well as a simultaneous fit to both final states. The results of the signal-plus-background fits can be found in Table 5.20. Asymmetric uncertainties on the parameters are presented, obtained using `MINOS` [133]. As expected, no significant pulls are observed from the generated values of each parameter, verifying the consistency of the fit. The post-fit uncertainties of the background shape parameters show that the pre-fit size of the variations was adequate, and that the fit has power to constrain their size. Due to the large number of parameters associated with the signal systematics, these were omitted from the table. For all of them the fit yielded a post fit value of 0 ± 1, as expected, since they are fully constrained by the auxiliary measurements from which their pre-fit value was derived, and the fit has no power to constrain them. The results of the combined fit can be seen in Fig. 5.41. The signal contributions are scaled to a fixed branching fraction and overlaid. The correlations between the fit parameters in the simultaneous fit are shown in Fig. 5.42. Negligible correlations are omitted from the matrix. Sizeable correlations are observed for the background nuisance parameters. These somewhat large correlations do not affect the convergence and stability of the fit. Appendix E includes a discussion on the dependence of the size of the correlations on the size of the sideband.

Likelihood scans were performed for the signal strength parameters in the considered fit configurations, and are presented in Figs. 5.43 to 5.45. The value of $-2\ln\lambda$ is displayed in the y axis of these figures, in which λ is the profile likelihood ratio,

described by Eq. (5.19). These scans verify the stability of the fit, showing a likelihood which is parabolic near the minimum, and with no local minima. These likelihood scans are also in agreement with the results in Table 5.20, both in terms of the best fit value and the uncertainties on the signal strengths, which correspond to $-2\ln\lambda = 1$ in the scans.

The 95% CL upper limits on $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma)$, $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma)$ and $\mathcal{B}(W^{\pm} \to \rho^{\pm}\gamma)$, obtained using the pre-fit background Asimov dataset are presented in Table 5.21, as well as the corresponding $\pm 1\sigma$ bands. The $W^{\pm} \to K^{\pm}\gamma$ and $W^{\pm} \to \pi^{\pm}\gamma$ processes cannot be distinguished, so only one of these signals is considered at a time, while the contribution of the other is assumed to be zero. When calculating the upper limit on $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma)$, $\mu(W^{\pm} \to \rho^{\pm}\gamma)$ is profiled in the fit, and vice-versa. The better limit on $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma)$ compared to the limit on $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma)$ reflects the higher SR efficiency for the $W^{\pm} \to K^{\pm}\gamma$ signal. Limits were calculated with and without the inclusion of background shape and signal normalisation systematics in the likelihood. One can see that including the signal normalisation systematics in the combined fit only deteriorates the limits on $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma)$ and $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma)$ by 1% and on $\mathcal{B}(W^{\pm} \to \rho^{\pm}\gamma)$ by 6%. The inclusion of the background shape systematics in the fit has a more substantial effect on the limit, deteriorating the limits on $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma)$ and $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma)$ by 47% and on $\mathcal{B}(W^{\pm} \to \rho^{\pm}\gamma)$ by 41%. Comparing the combined track-photon and tau-photon fit with the independent fits on each final state (including all systematics) one concludes that the combination improves the expected limit on $\mathcal{B}(W^{\pm} \to \rho^{\pm}\gamma)$ by 7%. No improvement of the limits on $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma)$ and $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma)$ is observed in the combined configuration when compared to the track-photon only fit, when all the systematics are included in the fit. This is understood when looking at the correlation matrix in Fig. 5.42, since the correlation between $\mu(W^{\pm} \to \rho^{\pm}\gamma)$ and $\mu(W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma)$ is negligible. The same is not true when no background shape systematics are included in the fit. In this configuration the correlation between $\mu(W^{\pm} \to \rho^{\pm}\gamma)$ and $\mu(W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma)$ becomes more substantial and the combination improves the limits on $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma)$ and $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma)$ by 18%. It was also verified that fixing one of the signal strengths to zero yields a negligible improvement in the upper limit, when all systematics are

included, as one would expect given the negligible correlation between $\mu(W^{\pm} \to \rho^{\pm}\gamma)$ and $\mu(W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma)$. One extra validation of the fit was performed by injecting signal contributions of different size in the Asimov dataset and then performing a signal-plus-background fit. This test was done in the track-photon and combined configurations. The study is documented in Appendix F and shows that the post-fit value of the signal strength corresponds to the amount of injected signal in all tested cases ($\mu = 1, 2, ..., 10$).

Table 5.20: Values of the fit parameters after fit to pre-fit background Asimov dataset, for each independent final state fit configuration and for the simultaneous fit. Due to the large number of parameters associated with the signal systematics, these were omitted from the table, as explained in the text.

Parameter	Track-photon	Tau-photon	Combined	
			Track-photon	Tau-photon
$\alpha(p_{\mathrm{T}}^{\gamma})$	$0.00^{+0.03}_{-0.24}$	$0.00^{+0.15}_{-0.25}$	$0.00^{+0.03}_{-0.21}$	$0.00^{+0.15}_{-0.24}$
$\alpha(\Delta\phi)$	$0.00^{+0.02}_{-0.01}$	$0.00^{+0.34}_{-0.22}$	$0.00^{+0.02}_{-0.01}$	$0.00^{+0.33}_{-0.22}$
$\alpha(\mathrm{tilt})$	$0.00^{+0.38}_{-0.30}$	$0.00^{+0.74}_{-0.37}$	$0.00^{+0.30}_{-0.22}$	$0.00^{+0.72}_{-0.36}$
$\mu(Z \to e^{-}e^{+})$	$1.00^{+1.05}_{-0.90}$	$0.99^{+1.86}_{-2.15}$	$0.99^{+0.82}_{-0.73}$	$0.99^{+1.83}_{-2.09}$
$\mu(\mathrm{Multijet})$	$1.000^{+0.005}_{-0.005}$	$1.00^{+0.01}_{-0.01}$	$1.000^{+0.003}_{-0.004}$	$1.00^{+0.01}_{-0.01}$
$\mu(W^{\pm} \to \pi^{\pm}\gamma)$	0^{+163}_{-155}	—	0^{+164}_{-154}	
$\mu(W^{\pm} \to \rho^{\pm}\gamma)$	2^{+1009}_{-935}	1^{+355}_{-359}	2^{+318}_{-333}	

5.12.2 Expected sensitivity with post-sideband fit Asimov dataset

The pre-fit background Asimov dataset allowed to perform preliminary studies of the sensitivity of the analysis, presented in the previous section, but we must stress that this "nominal" background expectation was not constrained in anyway using sideband data. The current section details the results of the fit performed to sideband data and the expected upper limits derived using a post-sideband fit background Asimov dataset, which is a more accurate description of the background data. The post-sideband fit upper limits presented in this section are the ones that should be compared to the analysis observed upper limits (presented in the next section).

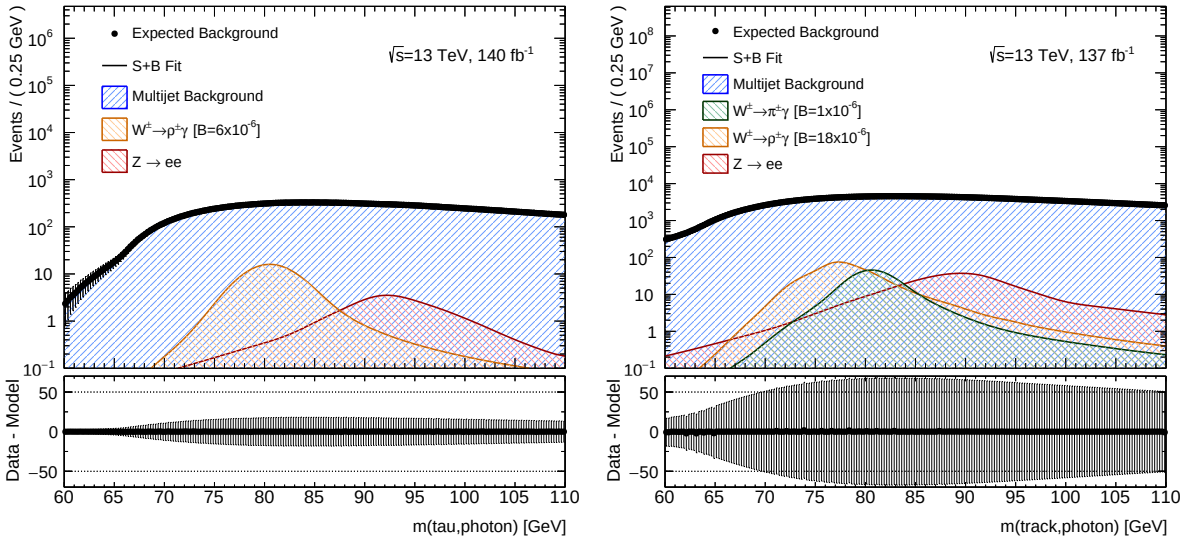


Figure 5.41: Combined fit to pre-fit background Asimov dataset. The results are shown in the tau-photon final state (left), and the track-photon final state (right). The signal contributions are scaled to a fixed branching fraction and overlaid.

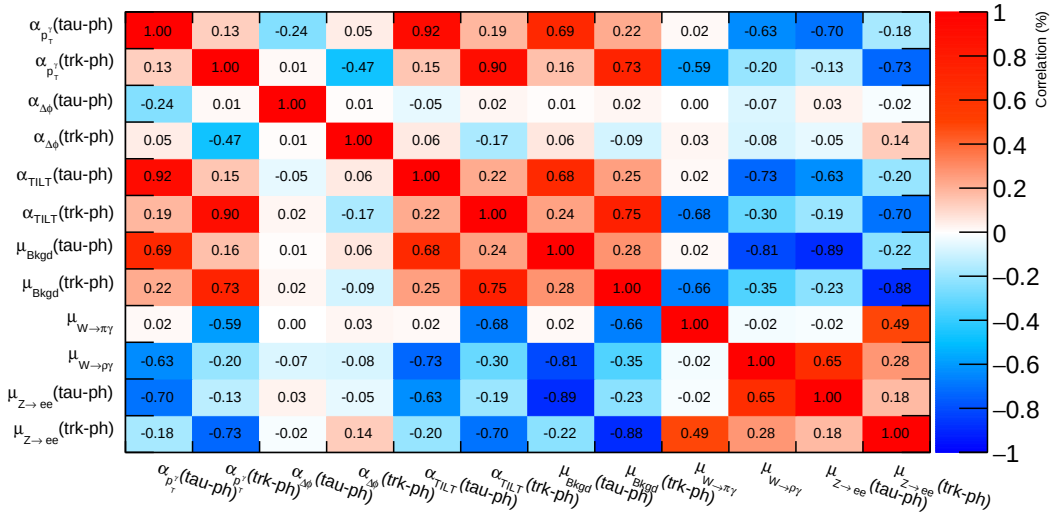


Figure 5.42: Correlation matrix between parameters in combined fit to pre-fit background Asimov dataset. Given the large number of parameters present in the fit, only parameters with sizeable correlations were included in the matrix.

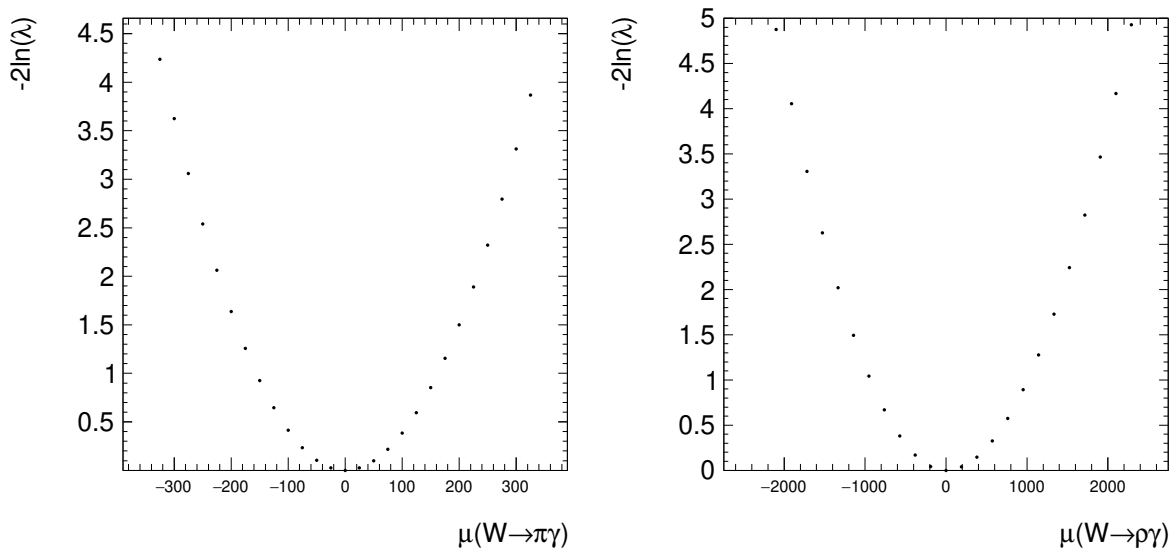


Figure 5.43: Likelihood scans of $\mu(W^{\pm} \to \pi^{\pm}\gamma)$ (left) and $\mu(W^{\pm} \to \rho^{\pm}\gamma)$ (right) in the track-photon final state.

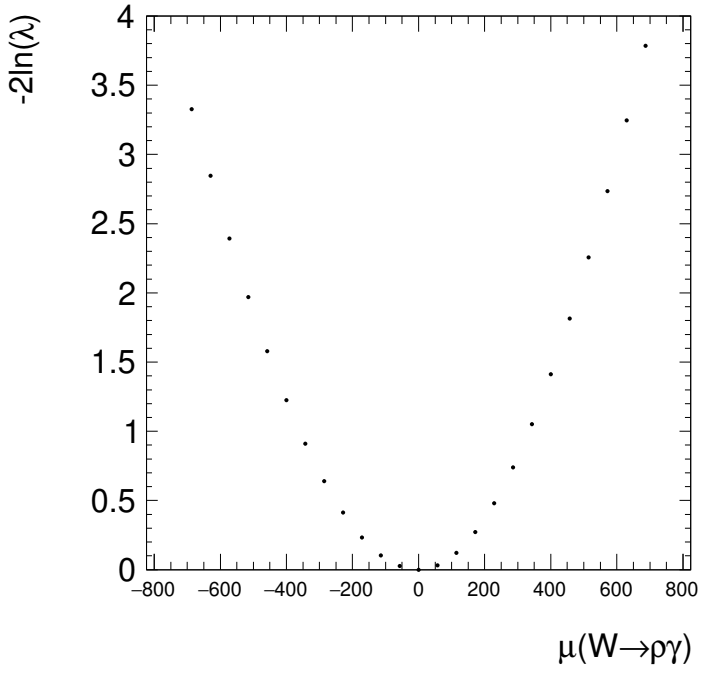


Figure 5.44: Likelihood scan of $\mu(W^{\pm} \to \rho^{\pm}\gamma)$ in the tau-photon final state.

Table 5.21: Expected branching fraction upper limits at 95% CL obtained using the pre-fit background Asimov dataset for each final state independently and for the combination. The limits were calculated including different sets of systematics. The resulting 1σ and 2σ bands are also provided.

		Expected ($\times 10^{-6}$)	$\pm 1\sigma$	$\pm 2\sigma$
		track-photon		
$W^{\pm} \to \pi^{\pm}\gamma$	No systematics	1.06	1.48/0.77	1.98/0.57
	Background shape	1.28	1.78/0.92	2.38/0.69
	Signal normalisation	1.08	1.50/0.78	2.01/0.58
	All systematics	1.30	1.80/0.93	2.42/0.70
$W^{\pm} \to K^{\pm}\gamma$	No systematics	0.96	1.33/0.69	1.79/0.51
	Background shape	1.15	1.60/0.83	2.15/0.62
	Signal normalisation	0.97	1.35/0.70	1.81/0.52
	All systematics	1.17	1.62/0.84	2.18/0.63
$W^{\pm} \to \rho^{\pm}\gamma$	No systematics	11.64	16.19/8.38	21.71/6.25
	Background shape	17.34	24.31/12.49	32.35/9.31
	Signal normalisation	11.80	16.43/8.50	22.02/6.33
	All systematics	17.59	24.47/12.67	32.81/9.44
		tau-photon		
$W^{\pm} \to \rho^{\pm}\gamma$	No systematics	3.98	5.54/2.87	7.43/2.14
	Background shape	5.62	7.82/4.05	10.49/3.02
	Signal normalisation	4.29	5.97/3.09	8.00/2.30
	All systematics	6.05	8.42/4.36	11.29/3.25
		Combined		
$W^{\pm} \to \pi^{\pm}\gamma$	No systematics	0.87	1.21/0.63	1.63/0.47
	Background shape	1.28	1.78/0.92	2.38/0.69
	Signal normalisation	0.88	1.23/0.64	1.64/0.48
	All systematics	1.30	1.80/0.93	2.42/0.70
$W^{\pm} \to K^{\pm}\gamma$	No systematics	0.78	1.09/0.57	1.46/0.42
	Background shape	1.15	1.60/0.83	2.15/0.62
	Signal normalisation	0.80	1.11/0.57	1.48/0.43
	All systematics	1.17	1.62/0.84	2.18/0.63
$W^{\pm} \to \rho^{\pm}\gamma$	No systematics	3.78	5.24/2.71	7.03/2.02
	Background shape	5.34	7.43/3.85	9.97/2.87
	Signal normalisation	4.00	5.57/2.88	7.46/2.15
	All systematics	5.67	7.89/4.09	10.58/3.04

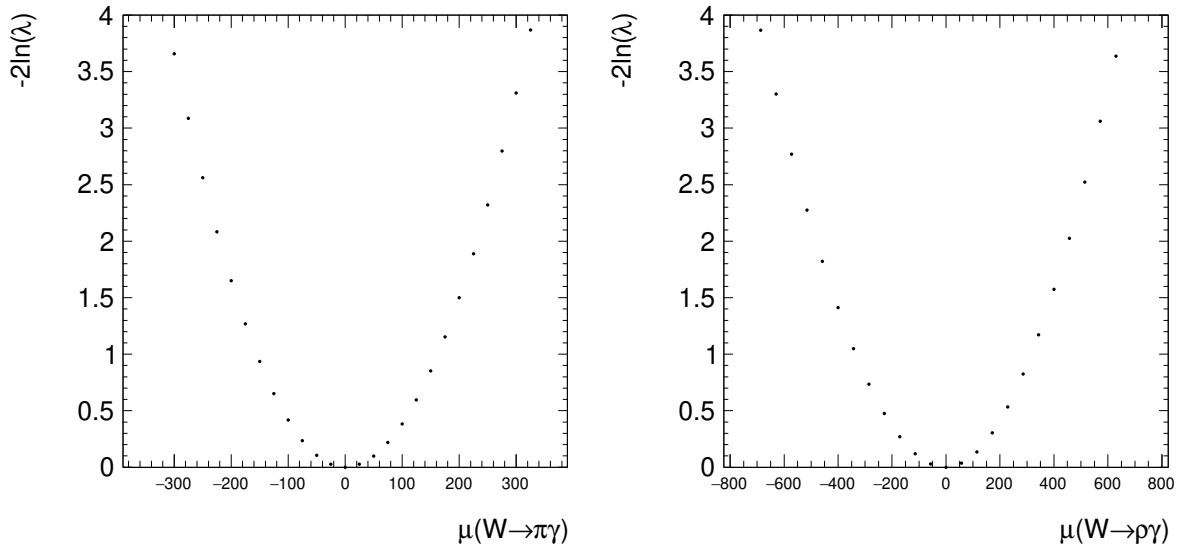


Figure 5.45: Likelihood scans of $\mu(W^{\pm} \to \pi^{\pm}\gamma)$ (left) and $\mu(W^{\pm} \to \rho^{\pm}\gamma)$ (right) in simultaneous fit to the track-photon and tau-photon SRs.

Binned maximum likelihood fits were performed to sideband data in the track-photon and tau-photon SRs, excluding the W boson invariant mass range between 76.5 and 84.5 GeV. The signal strengths $\mu(W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma)$ and $\mu(W^{\pm} \to \rho^{\pm}\gamma)$, and nuisance parameters controlling the signal normalisation uncertainties were fixed to zero in this fit. The results of the fit are presented in Fig. 5.46, which shows good agreement between the post-fit background model and the data. Table 5.22 lists the post-fit values of the background normalisations and the background shape nuisance parameters. The non-zero values of the background shape parameters show their capability to absorb the residual mismodelling of the multijet background in SR.

Table 5.22: Post-sideband fit value of the track-photon and tau-photon background nuisance parameters.

Parameter	Track-photon	Tau-photon
$\alpha(p_{\mathrm{T}}^{\gamma})$	$0.06^{+0.02}_{-0.02}$	$0.09^{+0.11}_{-0.15}$
$\alpha(\Delta\phi)$	$0.10^{+0.21}_{-0.02}$	$-0.14^{+0.29}_{-0.24}$
α(tilt)	$-1.20^{+0.20}_{-0.20}$	$0.27^{+0.58}_{-0.49}$
$\mu(Z \to e^{-}e^{+})$	$3.09^{+0.69}_{-0.69}$	$-0.20^{+1.96}_{-1.96}$
μ(Multijet)	$0.977^{+0.003}_{-0.003}$	$1.00^{+0.01}_{-0.01}$

A post-sideband fit Asimov dataset was constructed from the post-sideband fit

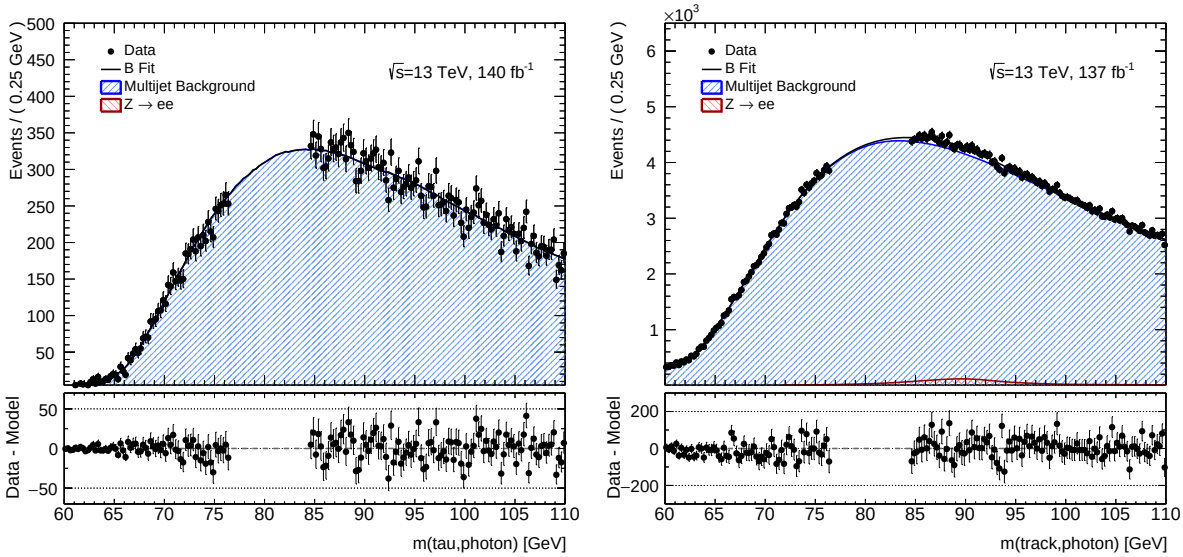


Figure 5.46: Background only fit to track-photon (left) and tau-photon (right) invariant mass distribution of sideband data. The $Z \to e^-e^+$ and multijet are drawn separately in red and blue, respectively. The black full line shows the post-fit background model (multijet + $Z \to e^-e^+$). The $Z \to e^-e^+$ component in the tau-photon final state is not visible in the plot since it has a negative post-fit value. The signal components are fixed at a value of 0, and as such are not represented in the plot. The residuals panels show the differences between the data and total post-fit background.

background prediction, with the nuisance parameters fixed to the values listed in Table 5.22. A signal-plus-background fit was performed to this Asimov dataset, yielding the post-fit values presented in Table 5.23. This was done for each final state independently and for the combination of both track-photon and tau-photon SRs. Once again, the parameters associated with the signal systematics were omitted from the table, since all of them presented a post fit value of 0 ± 1. No significant pulls are observed, with respect to the expected values (the post-sideband fit values, listed in Table 5.22). The post-fit values of the signal strengths are consistent with zero, as expected, since no signal was included in the Asimov.

Expected upper limits at 95% CL were derived and can be found in Table 5.25 ("post-fit expected"), compared to the preliminary limits obtained using the pre-fit background Asimov dataset ("pre-fit expected"). This comparison shows that the effect of the pulls of the background nuisance parameters on the upper limits is small - 6% difference for $\mathcal{B}(W^\pm \to \pi^\pm\gamma)$ and $\mathcal{B}(W^\pm \to K^\pm\gamma)$ and 5% difference for $\mathcal{B}(W^\pm \to \rho^\pm\gamma)$. The effect of the systematic uncertainties on the post-sideband fit expected upper limits was

also evaluated for the combined configuration, and is detailed in Table 5.24. The same overall behaviour that was observed in the previous section is observed. The signal systematic uncertainties are responsible for a 1% deterioration of the expected limits on $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma)$ and $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma)$ and 6% of the limit on $\mathcal{B}(W^{\pm} \to \rho^{\pm}\gamma)$. The effect of the background shape uncertainties is, once again, the dominant effect, causing an approximately 40% deterioration of the three upper limits.

Table 5.23: Values of the fit parameters after fit to Asimov dataset created after a background-only fit to data sidebands, for each independent final state fit configuration and for the simultaneous fit. Due to the large number of parameters associated with the signal systematics, these were omitted from the table.

Parameter	Track-photon	Tau-photon	Combined	
			Track-photon	Tau-photon
$\alpha(p_{\mathrm{T}}^{\gamma})$	$0.06^{+0.03}_{-0.04}$	$0.09^{+0.15}_{-0.20}$	$0.06^{+0.03}_{-0.03}$	$0.09^{+0.14}_{-0.19}$
$\alpha(\Delta\phi)$	$-0.10^{+0.01}_{-0.02}$	$-0.13^{+0.28}_{-0.23}$	$-0.10^{+0.01}_{-0.01}$	$-0.13^{+0.28}_{-0.23}$
$\alpha(\mathrm{tilt})$	$-1.20^{+0.33}_{-0.34}$	$0.26^{+0.75}_{-0.54}$	$-1.20^{+0.24}_{-0.25}$	$0.26^{+0.72}_{-0.53}$
$\mu(Z \to e^{-}e^{+})$	$3.08^{+0.95}_{-0.95}$	$-0.18^{+2.14}_{-2.15}$	$3.09^{+0.73}_{-0.73}$	$-0.18^{+2.09}_{-2.10}$
$\mu(\mathrm{Multijet})$	$0.976^{+0.005}_{-0.005}$	$1.00^{+0.01}_{-0.01}$	$0.977^{+0.003}_{-0.003}$	$1.00^{+0.01}_{-0.01}$
$\mu(W^{\pm} \to \pi^{\pm}\gamma)$	-1^{+155}_{-154}	—	-1^{+155}_{-154}	
$\mu(W^{\pm} \to \rho^{\pm}\gamma)$	-8^{+1019}_{-1016}	0^{+361}_{-359}	0^{+339}_{-338}	

5.12.3 Fit to full dataset and observed upper limits

This section details the analysis results, obtained using the full unblinded dataset. A signal-plus-background fit was performed to the full dataset in each final state independently, and also combining the track-photon and tau-photon final states. The results of the fit to the track-photon dataset can be found in Fig. 5.47. Both the $W^{\pm} \to \pi^{\pm}\gamma/W^{\pm} \to K^{\pm}\gamma$ and the $W^{\pm} \to \rho^{\pm}\gamma$ signal strength parameters are free to float in the fit. The dotted red line shows the post-fit model (including signal and background). The signal processes are shown scaled to the same branching fraction of $\mathcal{B} = 5 \times 10^{-5}$. On the left the $W^{\pm} \to \pi^{\pm}\gamma$ signal is displayed, while the plot on the right shows the $W^{\pm} \to K^{\pm}\gamma$ signal. The only difference between the two plots is the normalisation of the

Table 5.24: Expected branching fraction upper limits at 95% CL obtained using the post-sideband fit background Asimov dataset for the combined track-photon and tau-photon fit. The limits were calculated including different sets of systematics. The resulting 1σ and 2σ bands are also provided.

		Expected ($\times 10^{-6}$)	$\pm 1\sigma$	$\pm 2\sigma$
		Combined		
$W^{\pm} \to \pi^{\pm}\gamma$	No systematics	0.86	1.19/0.62	1.60/0.46
	Background shape	1.20	1.68/0.87	1.62/0.47
	Signal normalisation	0.87	1.21/0.63	1.62/0.47
	All systematics	1.22	1.70/0.88	2.28/0.66
$W^{\pm} \to K^{\pm}\gamma$	No systematics	0.77	1.07/0.55	1.44/0.41
	Background shape	1.08	1.51/0.78	2.02/0.58
	Signal normalisation	0.78	1.09/0.56	1.46/0.42
	All systematics	1.10	1.53/0.79	2.05/0.59
$W^{\pm} \to \rho^{\pm}\gamma$	No systematics	3.75	5.22/2.70	7.00/2.01
	Background shape	5.60	7.79/4.03	10.44/3.00
	Signal normalisation	3.98	5.54/2.87	7.42/2.14
	All systematics	5.95	8.28/4.29	11.10/3.19

Table 5.25: Expected branching fraction upper limits at 95% CL obtained by using the pre-sideband fit background Asimov dataset ("Pre-fit") and the Asimov dataset built using the results of the sideband fit ("Post-fit").

		Pre-fit expected ($\times 10^{-6}$)	Post-fit expected ($\times 10^{-6}$)
$W^{\pm} \to \pi^{\pm}\gamma$	track-photon	$1.30^{+0.50}_{-0.37}$	$1.22^{+0.48}_{-0.34}$
	combined	$1.30^{+0.50}_{-0.37}$	$1.22^{+0.48}_{-0.34}$
$W^{\pm} \to K^{\pm}\gamma$	track-photon	$1.17^{+0.45}_{-0.33}$	$1.10^{+0.43}_{-0.31}$
	combined	$1.17^{+0.45}_{-0.33}$	$1.10^{+0.43}_{-0.31}$
$W^{\pm} \to \rho^{\pm}\gamma$	track-photon	$17.59^{+6.88}_{-4.92}$	$17.57^{+6.88}_{-4.91}$
	tau-photon	$6.05^{+2.37}_{-1.69}$	$6.38^{+2.48}_{-1.78}$
	combined	$5.67^{+2.22}_{-1.58}$	$5.95^{+2.33}_{-1.66}$

$W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma$ signal, which comes from the efficiency difference between the two signals. The results of the fit to the full tau-photon SR are seen in Section 5.12.3. In this case the $Z \to e^{-}e^{+}$ is not visible since it has a negative post-fit value. On Fig. 5.49 the results of the combined fit to the track-photon (right) and tau-photon (left) SRs are shown. The post-fit values of the parameters for all fit configurations are listed in Table 5.26. Table 5.27 lists the post-fit number of events for each contribution. For all configurations, no significant deviation from the background prediction was observed: in the track-photon final state the post-fit signal strength is consistent with zero within 1σ and 1.3σ for $W^{\pm} \to \rho^{\pm}\gamma$ and $W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma$, respectively; in the tau-photon final state $\mu(W^{\pm} \to \rho^{\pm}\gamma)$ is consistent with 0 within 0.04σ; and in the combined fit the signal strengths are compatible with zero within 0.37σ and 1.3σ, respectively. The signal systematics nuisance parameters, which were once again omitted from the table have a post-fit value of 0 ± 1. When comparing Table 5.26 with Table 5.22, no significant deviations are observed for the background nuisance parameters, showing that the the background model derived from the fit to sideband data provides a good prediction of the full SR background.

The expected and observed upper limits on the branching fractions, at 95% CL are reported in Table 5.28, for each final state independently and for the combination. The limits on $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma)$ and $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma)$ for the combined and track-photon only fits are numerically equivalent. Including the track-photon SR in the fit improves the observed (expected) limit on $\mathcal{B}(W^{\pm} \to \rho^{\pm}\gamma)$ by 18% (7%), with respect to the tau-photon only limit. These results correspond to the first reported upper limits set on $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma)$ and $\mathcal{B}(W^{\pm} \to \rho^{\pm}\gamma)$ and the most stringent upper limit on $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma)$, improving the previous limit set by CDF by a factor of 4. This analysis is a step towards precision tests of QCD factorisation using exclusive hadronic decays of the W boson. A W boson mass measurement using fully reconstructed decays might also become possible as larger datasets become available.

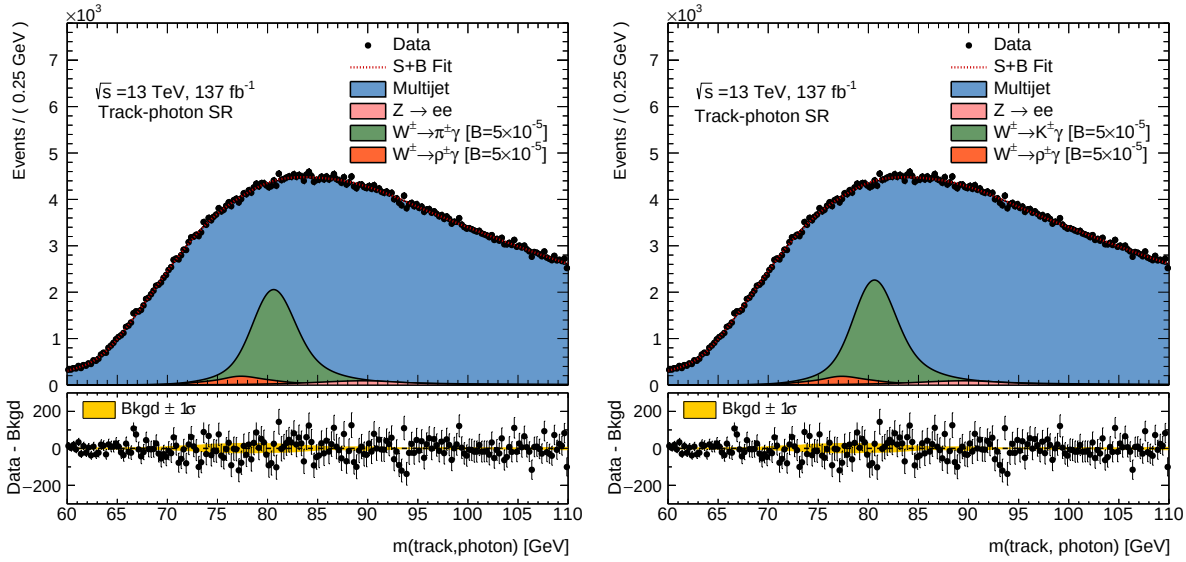


Figure 5.47: Signal-plus-background fit to the full dataset in the track-photon final state. Both the $W^{\pm} \to \pi^{\pm}\gamma / W^{\pm} \to K^{\pm}\gamma$ and the $W^{\pm} \to \rho^{\pm}\gamma$ signal strengths are free to float in the fit. The $Z \to e^{-}e^{+}$ and multijet are drawn separately in red and blue, respectively.The $Z \to e^{-}e^{+}$ contribution is stacked on top of the multijet background. The dotted red line shows the post-fit signal-plus-background model. The $W^{\pm} \to \pi^{\pm}\gamma$ (left) and $W^{\pm} \to K^{\pm}\gamma$ (right), as well as the $W^{\pm} \to \rho^{\pm}\gamma$ are scaled to the same branching fraction, $\mathcal{B} = 5 \times 10^{-5}$, for visibility. showing the differences in efficiency. The bottom panel shows the differences between the data and the fitted background. The black error bars include only the statistical uncertainty on the data. The yellow band represents the post-fit $\pm 1\sigma$ uncertainty band on the background estimation.

Table 5.26: Post-fit values of the parameters after signal-plus-background fit to full dataset, for each independent final state fit configuration and for the simultaneous fit. Due to the large number of parameters associated with the signal systematics, these were omitted from the table.

Parameter	Track-photon	Tau-photon	Combined	
			Track-photon	Tau-photon
$\alpha(p_{\mathrm{T}}^{\gamma})$	$0.10^{+0.03}_{-0.03}$	$0.15^{+0.15}_{-0.15}$	$0.08^{+0.03}_{-0.03}$	$0.17^{+0.14}_{-0.14}$
$\alpha(\Delta\phi)$	$-0.08^{+0.01}_{-0.01}$	$-0.15^{+0.27}_{-0.23}$	$-0.08^{+0.01}_{-0.01}$	$-0.14^{+0.28}_{-0.27}$
$\alpha(\mathrm{tilt})$	$-0.85^{+0.33}_{-0.33}$	$0.44^{+0.74}_{-0.64}$	$-1.06^{+0.24}_{-0.25}$	$0.61^{+0.72}_{-0.68}$
$\mu(Z \to e^{-}e^{+})$	$2.40^{+0.94}_{-0.94}$	$-0.45^{+2.15}_{-2.02}$	$2.97^{+0.72}_{-0.72}$	$-0.89^{+2.10}_{-2.10}$
$\mu(\mathrm{Multijet})$	$0.978^{+0.005}_{-0.005}$	$1.00^{+0.01}_{-0.01}$	$0.975^{+0.003}_{-0.003}$	$1.00^{+0.01}_{-0.01}$
$\mu(W^{\pm} \to \pi^{\pm}\gamma)$	196^{+158}_{-153}	—	195^{+157}_{-154}	
$\mu(W^{\pm} \to \rho^{\pm}\gamma)$	-1015^{+1007}_{-1033}	-1^{+359}_{-360}	-122^{+334}_{-350}	

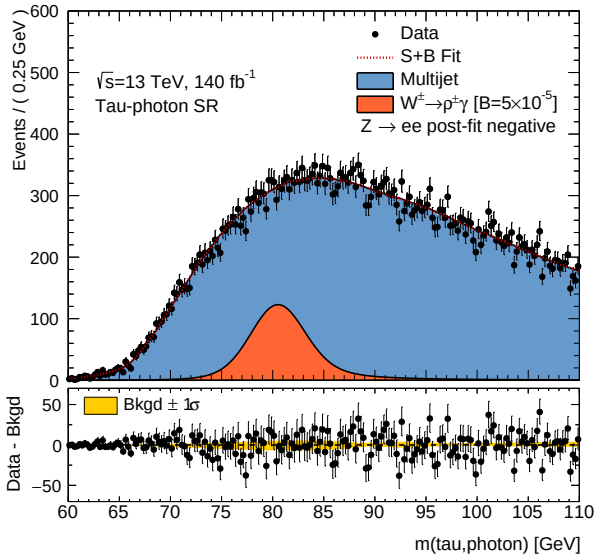


Figure 5.48: Signal-plus-background fit to the full dataset in the tau-photon final state. The dotted red line shows the post-fit signal-plus-background model. The multijet background is stacked on top of the $Z \to e^-e^+$ contribution, which is not visible since it has a negative post-fit value. The $W^\pm \to \rho^\pm\gamma$ is scaled to a branching fraction of $\mathcal{B} = 5 \times 10^{-5}$ for visibility. The bottom panel shows the differences between the data and the fitted background.The black error bars include only the statistical uncertainty on the data. The yellow band represents the post-fit $\pm 1\sigma$ uncertainty band on the background estimation

Table 5.27: Post-fit number of events.

Process	Track-photon	Tau-photon	Combined	
			Track-photon	Tau-photon
$W^\pm \to \pi^\pm\gamma$	978 ± 770	—-	971 ± 768	—-
$W^\pm \to K^\pm\gamma$	979 ± 766	—-	971 ± 767	—-
$W^\pm \to \rho^\pm\gamma$	-1196 ± 1191	-9 ± 252	-144 ± 396	-87 ± 239
$Z \to e^-e^+$	4946 ± 1961	-80 ± 385	6126 ± 1495	-160 ± 377
Multijet	634230 ± 3280	43007 ± 618	632008 ± 2242	43166 ± 598
Data	638962	42918	638963	42918

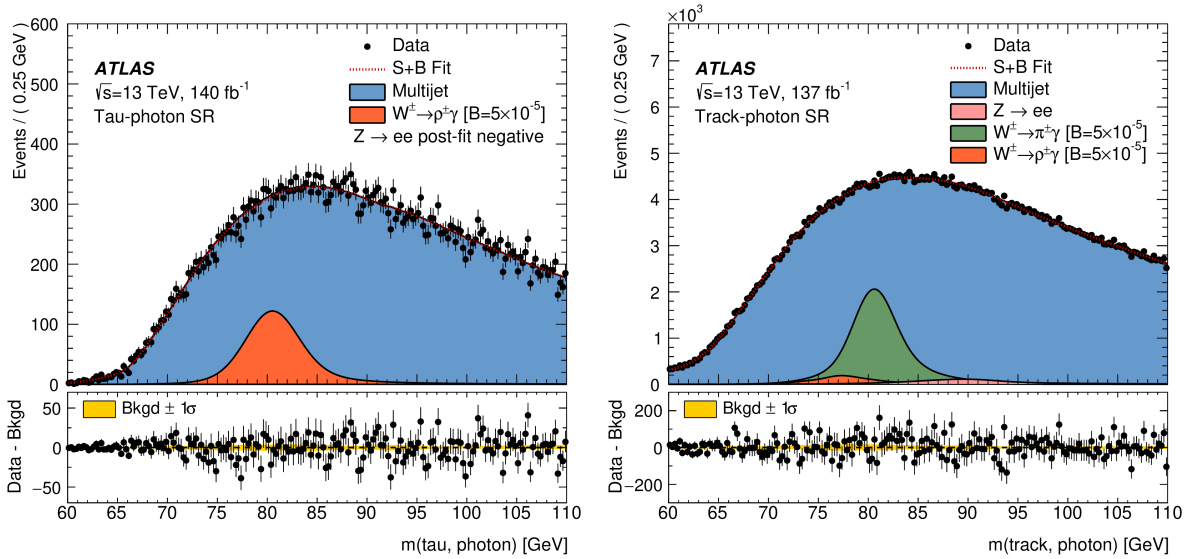


Figure 5.49: Signal-plus-background combined fit to full dataset in the tau-photon final state (left), and the track-photon final state (right). The dotted red line shows the post-fit signal-plus-background model. The $Z \to e^-e^+$ contribution is stacked on top of the multijet background. On the right, the $Z \to e^-e^+$ contribution can be seen on the bottom, while for the tau-photon final state (left) this contribution is not visible due to its negative post-fit value. The signal processes are shown scaled to the same branching fraction of $\mathcal{B} = 5 \times 10^{-5}$. The bottom panels show the differences between the data and the fitted background. The black error bars include only the statistical uncertainty on the data. The yellow band represents the post-fit $\pm 1\sigma$ uncertainty band on the background estimation.

Table 5.28: Observed upper limits, compared to the expected upper limits estimated using the post sideband fit Asimov.

		Expected limit ($\times 10^{-6}$)	Observed limit ($\times 10^{-6}$)
$W^{\pm} \to \pi^{\pm}\gamma$	track-photon	$1.22^{+0.48}_{-0.34}$	1.88
	combined	$1.22^{+0.48}_{-0.34}$	1.87
$W^{\pm} \to K^{\pm}\gamma$	track-photon	$1.10^{+0.43}_{-0.31}$	1.67
	combined	$1.10^{+0.43}_{-0.31}$	1.68
$W^{\pm} \to \rho^{\pm}\gamma$	track-photon	$17.57^{+6.88}_{-4.91}$	12.58
	tau-photon	$6.38^{+2.48}_{-1.78}$	6.29
	combined	$5.95^{+2.33}_{-1.66}$	5.17

Chapter 6

Conclusion

In the last decades, the SM has been tested extensively by the LHC experiments at unprecedented centre-of-mass energies, allowing to survey larger areas of the phase space; and using increasingly larger datasets, achieving unprecedented levels of statistical precision. Time and time again, physics analyses have yielded results consistent with the SM predictions. Nevertheless, there are still gaps which can be explored in the hunt for physics beyond the SM. Many rare processes which have been predicted by the SM are yet to be observed. Additionally, the theoretical precision on some SM parameters is yet to be matched by experiments. Any observed deviation from theory predictions could arise from physics beyond the SM, and would open a new path for experimental exploration.

New experimental challenges arise from pushing the luminosity and centre-of-mass energy frontiers at the LHC. As the luminosity increases, uncertainties associated with the modelling of background processes through simulation become limiting factors in analyses: in many cases the available Monte Carlo simulations have large associated theoretical uncertainties, as is the case for processes involving non-perturbative QCD effects; moreover, it is often computationally expensive to generate large samples, leading to large associated statistical uncertainties. Another challenge comes as higher centre-of-mass energies are reached, and dense charged particle environments, such as the cores of highly energetic jets, become more prevalent in analyses. In these environments, the level of separation between charged particles becomes of the order of the granularity of the ATLAS inner detector, negatively impacting the track reconstruction

efficiency. It is important to fully characterise the tracking performance in these challenging environments and to validate its modelling in simulation.

The work presented in this thesis addresses these experimental challenges and introduces new experimental constrains on rare SM processes. A data-driven methodology for measuring the track reconstruction efficiency in dense environments, which had been previously introduced in Ref. [2], is developed further. Several improvements to the method are introduced in this thesis: the multiple-track template is now derived from sampling the single-track template, and consequently the contamination from clusters wrongly identified as merged is removed; two and three-track templates are built separately allowing to quantify two and three track contributions in the fits to the measurement distributions; the F_{lost} calculation is extended to account for three track clusters; the single-track templates are smoothed reducing the impact of statistical fluctuations, and eliminating the need to use the templates from the lower jet p_{T} bin in the fits corresponding to higher p_{T} bins; the templates are allowed to shift in $\text{d}E/\text{d}x$ in order to account for the discrepancies observed due to momentum difference between clusters from inside and outside the jet core; and finally the fit range is extended. These improvements eliminate most of the sources of systematic uncertainties listed in Ref. [2]. A preliminary measurement of F_{lost} as a function of $p_{\text{T}}(\text{jet})$ and $|\eta(\text{trk})|$, using 2018 ATLAS pp collision data at $\sqrt{s} = 13$ TeV is presented. The results are comparable to what was previously observed for 2015 ATLAS data. F_{lost} is shown to increase with higher values of $p_{\text{T}}(\text{jet})$ and $|\eta(\text{trk})|$.

A novel, non-parametric data-driven modelling technique was presented. This technique relies on obtaining a model of conditional probabilities of relevant variables which describes the data, using a region defined by a looser version of the analysis event selection. By employing the method, one can obtain a description of the shape of the background in a discriminant variable, after re-applying the analysis signal region requirements. An implementation of this method which uses ancestral sampling is presented in detail, and its application was exemplified using the search for $H \to \phi(K^+K^-)\gamma$ as a case study. A signal injection test was performed, showing the robustness of the technique under signal contamination. Furthermore, an ensemble of independent statistical tests was performed, demonstrating that the method introduces no significant

bias to signal extraction. A second implementation of the method, relying on conditional adversarial networks was also briefly discussed. The two implementations of this novel background modelling technique are general enough to be applicable to a broad spectrum of analyses.

Finally, the search for exclusive hadronic decays of the W boson, using pp collision data collected by the ATLAS experiment, at $\sqrt{s} = 13$ TeV, was presented. Three decays were targeted: $W^{\pm} \to \pi^{\pm}\gamma$, $W^{\pm} \to K^{\pm}\gamma$ and $W^{\pm} \to \rho^{\pm}\gamma$. This search was enabled by the implementation of dedicated track-photon triggers. The use of standard di-photon triggers, taking advantage of the decay of the π^0 in the final state into collimated photons, was also instrumental in the search for $W^{\pm} \to \rho^{\pm}\gamma$, significantly increasing the analysis sensitivity to this specific decay channel. Moreover, the data-driven background modelling discussed in this thesis allowed to model the dominant multijet background, for which no reliable MC simulation samples were available. Upper limits are set on the branching fractions of the decays, at 95% CL: $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma) < 1.9 \times 10^{-6}$, $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma) < 1.7 \times 10^{-6}$ and $\mathcal{B}(W^{\pm} \to \rho^{\pm}\gamma) < 5.4 \times 10^{-6}$. These results correspond to the first experimental constraints on $\mathcal{B}(W^{\pm} \to K^{\pm}\gamma)$ and $\mathcal{B}(W^{\pm} \to \rho^{\pm}\gamma)$, and the most stringent upper limit on $\mathcal{B}(W^{\pm} \to \pi^{\pm}\gamma)$ to date, improving the limit previously set by the CDF experiment, by a factor of 4. Future sensitivity improvements could come from further trigger optimisation. Most of the efficiency loss coming from the track-photon trigger originates from the $E_T(\text{tau})/p_T(\text{trk})$ requirement, so a loosened higher limit would be ideal. Nevertheless, one must consider the total trigger rates, so it might not be trivial to loosen this requirement without introducing a new one. The search for the $W^{\pm} \to \rho^{\pm}\gamma$ decay would greatly benefit from the development of a dedicated trigger. Finally, one must not rule out potential improvements to the sensitivity from multivariate analysis techniques, which were not explored during the development of the analysis, but could be studied in the future.

This work provides relevant input for the design of future collider experiments, where exclusive hadronic decays of the W boson could potentially be observed for the first time. Future e^+e^- colliders are expected to deliver a clean sample of $\mathcal{O}(10^8)$ W^+W^- events [134], which would allow access to the $W^{\pm} \to D_s^{\pm}\gamma$ and $W^{\pm} \to \pi^{\pm}\pi^{\mp}\pi^{\pm}$ decays according to current SM predictions [28, 31, 35, 135]. Future hadron colliders promise to

produce $\mathcal{O}(10^{12})$ W bosons [136], which would translate to thousands of $W^{\pm} \to \pi^{\pm}\gamma$ and $W^{\pm} \to \rho^{\pm}\gamma$ decays and hundreds of $W^{\pm} \to K^{\pm}\gamma$ decays. In both cases, the observation of these channels pose significant experimental challenges both in terms of trigger strategy and background discrimination. Thus, careful detector optimization is required in order to access these signatures, and exploit them as a new way to measure the W boson properties and to probe the QCD factorization formalism. The novel experimental techniques presented in this thesis are a first step towards the observation of these decays in future facilities, which are currently being planned.

Bibliography

[1] ATLAS Collaboration, *Measurement of the $t\bar{t}$ cross section and its ratio to the Z production cross section using pp collisions at $\sqrt{s} = 13.6$ TeV with the ATLAS detector*, (2023), arXiv: 2308.09529 [hep-ex] (cit. on p. ii).

[2] ATLAS Collaboration, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv: 1704.07983 [hep-ex] (cit. on pp. ii, 36–38, 48–56, 58, 60, 175).

[3] ATLAS Collaboration, *Search for exclusive Higgs and Z boson decays to $\phi\gamma$ and $\rho\gamma$ with the ATLAS detector*, JHEP **07** (2018) 127, arXiv: 1712.02758 [hep-ex] (cit. on pp. iii, 48, 70, 75).

[4] A. Chisholm et al., *Non-parametric data-driven background modelling using conditional probabilities*, Journal of High Energy Physics **2022** (2022), URL: https://doi.org/10.1007%2Fjhep10%282022%29001 (cit. on pp. iii, 70, 95, 96).

[5] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321, ed. by J. C. Taylor (cit. on pp. 2, 5).

[6] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (16 1964) 508, URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508 (cit. on pp. 2, 5).

[7] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex] (cit. on pp. 2, 71).

[8] CMS Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: 1207.7235 [hep-ex] (cit. on p. 2).

[9] *Standard Model of Elementary Particles*, Accessed: 2023-09-29, URL: https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg (cit. on p. 3).

[10] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579 (cit. on p. 2).

[11] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264 (cit. on p. 2).

[12] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367 (cit. on p. 2).

[13] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652 (cit. on p. 4).

[14] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (12 1963) 531, URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531 (cit. on p. 4).

[15] R. L. Workman et al., *Review of Particle Physics*, PTEP **2022** (2022) 083C01 (cit. on pp. 4, 49, 50, 102, 115, 144, 147).

[16] J. Goldstone, A. Salam, and S. Weinberg, *Broken Symmetries*, Phys. Rev. **127** (1962) 965 (cit. on p. 6).

[17] G. Barr, R. Devenish, R. Walczak, and T. Weidberg, *Particle Physics in the LHC era*, Oxford master series in particle physics, astrophysics and cosmology, Oxford University Press, 2016, ISBN: 978-0-19-874855-7 (cit. on pp. 6, 8).

[18] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (20 1964) 585, URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585 (cit. on p. 6).

[19] J. Campbell, J. Huston, and F. Krauss, *The Black Book of Quantum Chromodynamics : a Primer for the LHC Era*, Oxford University Press, 2018, ISBN: 978-0-19-965274-7 (cit. on pp. 9, 11, 12).

[20] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438 (cit. on p. 10).

[21] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, Sov. Phys. JETP **46** (1977) 641 (cit. on p. 10).

[22] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. B **126** (1977) 298 (cit. on p. 10).

[23] S. D. Drell and T.-M. Yan, *Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies*, Phys. Rev. Lett. **25** (5 1970) 316, URL: https://link.aps.org/doi/10.1103/PhysRevLett.25.316 (cit. on p. 10).

[24] Z. Bern et al., *Left-Handed W Bosons at the LHC*, Phys. Rev. D **84** (2011) 034008, arXiv: 1103.5445 [hep-ph] (cit. on pp. 11, 13).

[25] ATLAS Collaboration, *Measurement of $W^{\pm}$ and Z-boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **759** (2016) 601, arXiv: 1603.09222 [hep-ex] (cit. on pp. 12, 102, 149).

[26] C. Collaboration, *Measurements of the W boson rapidity, helicity, double-differential cross sections, and charge asymmetry in pp collisions at $\sqrt{s}$ = 13 TeV*, Phys. Rev. D **102** (2020) 092012, arXiv: 2008.04174 [hep-ex] (cit. on pp. 12, 14).

[27] ATLAS Collaboration, *Standard Model Summary Plots February 2022*, tech. rep., CERN, 2022, URL: https://cds.cern.ch/record/2804061 (cit. on p. 14).

[28] Y. Grossman, M. König, and M. Neubert, *Exclusive Radiative Decays of W and Z Bosons in QCD Factorization*, JHEP **04** (2015) 101, arXiv: 1501.06569 [hep-ph] (cit. on pp. 15–17, 100, 133, 176).

[29] R. L. Workman et al., *Review of Particle Physics*, PTEP **2022** (2022) 083C01 (cit. on pp. 15, 20, 132).

[30] LHCb Collaboration, *Search for the rare decays $W^+ \to D_s^+\gamma$ and $Z \to D^0\gamma$ at LHCb*, Chin. Phys. C **47** (2023) 093002, arXiv: 2212.07120 [hep-ex] (cit. on pp. 15, 16).

[31] M. Mangano and T. Melia, *Rare exclusive hadronic W decays in a $t\bar{t}$ environment*, Eur. Phys. J. C **75** (2015) 258, arXiv: 1410.7475 [hep-ph] (cit. on pp. 16, 21, 176).

[32] L. Arnellos, W. J. Marciano, and Z. Parsa, *Radiative decays: $W^{\pm} \to P^{\pm} + \gamma$ and $Z \to P^0 + \gamma$*, Nuclear Physics B **196** (1982) 378, ISSN: 0550-3213, URL: https://www.sciencedirect.com/science/article/pii/0550321382904965 (cit. on p. 16).

[33] Y. Y. Keum and X.-Y. Pham, *Possible huge enhancement in the radiative decay of the weak W boson into the charmed D_s meson*, Mod. Phys. Lett. A **9** (1994) 1545, arXiv: hep-ph/9303300 (cit. on p. 16).

[34] CDF Collaboration, *Search for the Rare Radiative Decay: $W \to \pi\gamma$ in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. D **85** (2012) 032001, arXiv: 1104.1585 [hep-ex] (cit. on pp. 16, 17, 19, 100).

[35] CMS Collaboration, *Search for W Boson Decays to Three Charged Pions*, Phys. Rev. Lett. **122** (2019) 151802, arXiv: 1901.11201 [hep-ex] (cit. on pp. 16, 176).

[36] ATLAS Collaboration, *Summary of ATLAS searches for Higgs and Z boson decays to a meson and a photon*, tech. rep., CERN, 2023, URL: https://cds.cern.ch/record/2851888 (cit. on p. 17).

[37] CMS Collaboration, *Search for the rare decay of the W boson into a pion and a photon in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, Physics Letters B **819** (2021) 136409, ISSN: 0370-2693, URL: https://www.sciencedirect.com/science/article/pii/S037026932100349X (cit. on pp. 17, 19).

[38] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section and lepton differential distributions in $e\mu$ dilepton events from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, (2019) (cit. on p. 18).

[39] M. E. Peskin, *Concepts of Elementary Particle Physics*, Oxford Master Series in Physics, Oxford University Press, 2019, ISBN: 978-0-19-881218-0, 978-0-19-881219-7 (cit. on p. 18).

[40] S. Abachi et al., *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632, arXiv: hep-ex/9503003 (cit. on p. 18).

[41] J. Haller, A. Hoecker, R. Kogler, K. Mönig, and J. Stelzer, *Status of the global electroweak fit with Gfitter in the light of new precision measurements*, PoS **ICHEP2022** (2022) 897, arXiv: 2211.07665 [hep-ph] (cit. on p. 20).

[42] CDF Collaboration, *High-precision measurement of the W boson mass with the CDF II detector*, Science **376** (2022) 170, eprint: https://www.science.org/doi/pdf/10.1126/science.abk1781, URL: https://www.science.org/doi/abs/10.1126/science.abk1781 (cit. on p. 20).

[43] ATLAS Collaboration, *Improved W boson Mass Measurement using 7 TeV Proton-Proton Collisions with the ATLAS Detector*, tech. rep., CERN, 2023, URL: https://cds.cern.ch/record/2853290 (cit. on pp. 20, 21).

[44] E. Jones and W. J. Murray, *Mass biases in exclusive radiative hadronic decays of W bosons at the LHC*, New Journal of Physics **23** (2021) 113035, URL: https://dx.doi.org/10.1088/1367-2630/ac3572 (cit. on p. 21).

[45] O. S. Brüning et al., *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, 2004, URL: https://cds.cern.ch/record/782076 (cit. on pp. 23, 24).

[46] E. Mobs, *The CERN accelerator complex in 2019. Complexe des accélérateurs du CERN en 2019*, (2019), General Photo, URL: https://cds.cern.ch/record/2684277 (cit. on p. 24).

[47] *Luminosity determination in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector at the LHC*, (2022), arXiv: 2212.09379 [hep-ex] (cit. on pp. 25, 149).

[48] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det] (cit. on p. 25).

[49] *Public ATLAS Luminosity Results for Run-2 of the LHC*, Accessed: 2023-09-30, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (cit. on p. 26).

[50] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08003, URL: https://doi.org/10.1088/1748-0221/3/08/s08003 (cit. on pp. 25, 29, 31, 32, 34).

[51] J. Pequenao, *Computer generated image of the whole ATLAS detector*, (2008), URL: https://cds.cern.ch/record/1095924 (cit. on p. 28).

[52] *ATLAS Inner Detector Alignment Performance with February 2015 Cosmic Rays Data*, tech. rep., CERN, 2015, URL: https://cds.cern.ch/record/2008724 (cit. on p. 28).

[53] ATLAS Collaboration, *Study of the material of the ATLAS inner detector for Run 2 of the LHC*, JINST **12** (2017) P12009, arXiv: 1707.02826 [hep-ex] (cit. on pp. 30, 152).

[54] H. Kolanoski and N. Wermes, *Particle Detectors: Fundamentals and Applications*, Oxford University Press, 2020 (cit. on pp. 29, 31).

[55] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, tech. rep., 2010, URL: https://cds.cern.ch/record/1291633 (cit. on p. 29).

[56] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector in Run-2*, Eur. Phys. J. C **80** (2020) 1194, arXiv: 2007.07624 [hep-ex] (cit. on pp. 30, 31).

[57] ATLAS Collaboration, *ATLAS pixel detector electronics and sensors*, Journal of Instrumentation **3** (2008) P07007, URL: https://doi.org/10.1088/1748-0221/3/07/p07007 (cit. on p. 30).

[58] A. Vogel, *ATLAS Transition Radiation Tracker (TRT): Straw Tube Gaseous Detectors at High Rates*, tech. rep., CERN, 2013, URL: https://cds.cern.ch/record/1537991 (cit. on p. 31).

[59] ATLAS Collaboration, *Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties*, JINST **12** (2017) P05002, arXiv: 1702.06473 [hep-ex] (cit. on p. 31).

[60] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex] (cit. on pp. 32, 40–44, 121).

[61] J. Pequenao, *Computer Generated image of the ATLAS calorimeter*, (2008), URL: https://cds.cern.ch/record/1095927 (cit. on p. 33).

[62] J. Pequenao, *Computer generated image of the ATLAS Muons subsystem*, (2008), URL: https://cds.cern.ch/record/1095929 (cit. on p. 34).

[63] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004, arXiv: 2007.12539 [physics.ins-det] (cit. on p. 35).

[64] ATLAS Collaboration, *dE/dx measurement in the ATLAS Pixel Detector and its use for particle identification*, ATLAS-CONF-2011-016, 2011, URL: https://cds.cern.ch/record/1336519 (cit. on p. 36).

[65] T. Cornelissen et al., *The new ATLAS track reconstruction (NEWT)*, J. Phys. Conf. Ser. **119** (2008) 032014, ed. by R. Sobie, R. Tafirout, and J. Thomson (cit. on p. 36).

[66] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444 (cit. on p. 36).

[67] ATLAS Collaboration, *A neural network clustering algorithm for the ATLAS silicon pixel detector*, JINST **9** (2014) P09009, arXiv: 1406.7690 [hep-ex] (cit. on p. 37).

[68] ATLAS Collaboration, *Training and validation of the ATLAS pixel clustering neural networks*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.cl PHYS-PUB-2018-002: CERN, 2018, URL: http://cds.cern.ch/record/2309474 (cit. on p. 37).

[69] P. F. Åkesson et al., *ATLAS Tracking Event Data Model*, tech. rep., CERN, 2006, URL: https://cds.cern.ch/record/973401 (cit. on p. 38).

[70] T. G. Cornelissen et al., *Updates of the ATLAS Tracking Event Data Model (Release 13)*, tech. rep., CERN, 2007, URL: https://cds.cern.ch/record/1038095 (cit. on p. 39).

[71] ATLAS Collaboration, *Early Inner Detector Tracking Performance in the 2015 Data at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-051, 2015, URL: https://cds.cern.ch/record/2110140 (cit. on pp. 38, 40, 113).

[72] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on p. 44).

[73] ATLAS Collaboration, *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*, tech. rep., CERN, 2015, URL: https://cds.cern.ch/record/2037702 (cit. on pp. 44, 45).

[74] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, tech. rep., CERN, 2014, URL: https://cds.cern.ch/record/1700870 (cit. on p. 45).

[75] ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV*, Eur. Phys. J. C **75** (2015) 303, arXiv: 1412.7086 [hep-ex] (cit. on p. 45).

[76] ATLAS Collaboration, *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, Eur. Phys. J. C **76** (2016) 295, arXiv: 1512.05955 [hep-ex] (cit. on p. 46).

[77] ATLAS Collaboration, *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, (2019), URL: https://cds.cern.ch/record/2688062 (cit. on pp. 46, 47, 115).

[78] ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, ATL-PHYS-PUB-2015-045, 2015, URL: https://cds.cern.ch/record/2064383 (cit. on pp. 47, 117).

[79] K. Cranmer, *Kernel estimation in high-energy physics*, Computer Physics Communications **136** (2001) 198, ISSN: 0010-4655, URL: http://dx.doi.org/10.1016/S0010-4655(00)00243-5 (cit. on pp. 58, 133).

[80] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, ed. by L. Lyons and M. Karagoz, arXiv: physics/0306116 (cit. on p. 58).

[81] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke, *Interpolation between multi-dimensional histograms using a new non-linear moment morphing method*, Nucl. Instrum. Meth. A **771** (2015) 39, arXiv: 1410.7388 [physics.data-an] (cit. on pp. 61, 83, 85, 141).

[82] ATLAS Collaboration, *Searches for exclusive Higgs and Z boson decays into $J/\psi\gamma$, $\psi(2S)\gamma$, and $\Upsilon(nS)\gamma$ at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **786** (2018) 134, arXiv: 1807.00802 [hep-ex] (cit. on pp. 70, 75, 92).

[83] ATLAS Collaboration, *Search for Higgs and Z Boson Decays to $\phi\gamma$ with the ATLAS Detector*, Phys. Rev. Lett. **117** (2016) 111802, arXiv: 1607.03400 [hep-ex] (cit. on pp. 70, 75, 92).

[84] ATLAS Collaboration, *Search for Higgs and Z Boson Decays to $J/\psi\gamma$ and $\Upsilon(nS)\gamma$ with the ATLAS Detector*, Phys. Rev. Lett. **114** (2015) 121801, arXiv: 1501.03276 [hep-ex] (cit. on pp. 70, 75, 92).

[85] ATLAS Collaboration, *Direct constraint on the Higgs-charm coupling from a search for Higgs boson decays into charm quarks with the ATLAS detector*, Eur. Phys. J. C **82** (2022) 717, arXiv: 2201.11428 [hep-ex] (cit. on p. 71).

[86] ATLAS Collaboration, *Measurement of Higgs boson decay into b-quarks in associated production with a top-quark pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **06** (2022) 097, arXiv: 2111.06712 [hep-ex] (cit. on p. 71).

[87] ATLAS Collaboration, *A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector*, Phys. Lett. B **812** (2021) 135980, arXiv: 2007.07830 [hep-ex] (cit. on pp. 71, 92).

[88] ATLAS Collaboration, *Measurement of the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ cross-sections in pp collisions at $\sqrt{s} = 13.6$ TeV with the ATLAS detector*, 2023, arXiv: 2306.11379 [hep-ex] (cit. on p. 72).

[89] ATLAS Collaboration, *Measurement of the properties of Higgs boson production at $\sqrt{s} = 13$ TeV in the $H \to \gamma\gamma$ channel using 139 fb^{-1} of pp collision data with the ATLAS experiment*, JHEP **07** (2023) 088, arXiv: 2207.00348 [hep-ex] (cit. on p. 72).

[90] ATLAS Collaboration, *Measurements of the Higgs boson inclusive and differential fiducial cross-sections in the diphoton decay channel with pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **08** (2022) 027, arXiv: 2202.00487 [hep-ex] (cit. on pp. 72, 92).

[91] P. Dauncey, M. Kenzie, N. Wardle, and G. Davies, *Handling uncertainties in background shapes: the discrete profiling method*, JINST **10** (2015) P04015, arXiv: 1408.6865 [physics.data-an] (cit. on p. 72).

[92] F. Abe et al., *A Measurement of $\sigma B(W \to e\nu)$ and $\sigma B(Z^0 \to e^+e-)$ in $\bar{p}p$ collisions at $\sqrt{s} = 1800$ GeV*, Phys. Rev. D **44** (1991) 29 (cit. on p. 73).

[93] M. Frate, K. Cranmer, S. Kalia, A. Vandenberg-Rodes, and D. Whiteson, *Modeling Smooth Backgrounds and Generic Localized Signals with Gaussian Processes*, (2017), arXiv: 1709.05681 [physics.data-an] (cit. on p. 73).

[94] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag, 2006, ISBN: 0387310738 (cit. on p. 73).

[95] M. König and M. Neubert, *Exclusive Radiative Higgs Decays as Probes of Light-Quark Yukawa Couplings*, JHEP **08** (2015) 012, arXiv: 1505.03870 [hep-ph] (cit. on p. 75).

[96] A. L. Kagan et al., *Exclusive Window onto Higgs Yukawa Couplings*, Phys. Rev. Lett. **114** (2015) 101802, arXiv: 1406.1722 [hep-ph] (cit. on p. 75).

[97] G. Perez, Y. Soreq, E. Stamou, and K. Tobioka, *Prospects for measuring the Higgs boson coupling to light quarks*, Phys. Rev. D **93** (2016) 013001, arXiv: 1505.06689 [hep-ph] (cit. on p. 75).

[98] CMS Collaboration, *Search for a Higgs boson decaying into $\gamma^*\gamma \to \ell\ell\gamma$ with low dilepton mass in pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. B **753** (2016) 341, arXiv: 1507.03031 [hep-ex] (cit. on p. 75).

[99] CMS Collaboration, *Search for rare decays of Z and Higgs bosons to J/ψ and a photon in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 94, arXiv: 1810.10056 [hep-ex] (cit. on p. 75).

[100] A. Collaboration, *Search for exclusive Higgs and Z boson decays to $\phi\gamma$ and $\rho\gamma$ with the ATLAS detector*, JHEP **07** (2018) 127, arXiv: 1712.02758 [hep-ex] (cit. on pp. 75, 76, 92).

[101] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph] (cit. on p. 75).

[102] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. D **93** (2016) 033006, arXiv: 1506.07443 [hep-ph] (cit. on p. 75).

[103] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on p. 75).

[104] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on p. 75).

[105] J. de Favereau et al., *DELPHES 3, A modular framework for fast simulation of a generic collider experiment*, JHEP **02** (2014) 057, arXiv: 1307.6346 [hep-ex] (cit. on p. 75).

[106] ATLAS Collaboration, *Searches for exclusive Higgs and Z boson decays into a vector quarkonium state and a photon using 139 fb^{-1} of ATLAS $\sqrt{s} = 13$ TeV proton-proton collision data*, The European Physical Journal C **83** (2023) 781 (cit. on p. 92).

[107] I. J. Goodfellow et al., *Generative Adversarial Networks*, (2014), arXiv: 1406.2661 [stat.ML] (cit. on p. 93).

[108] Y. Alanazi et al., *A survey of machine learning-based physics event generation*, (2021), arXiv: 2106.00643 [hep-ph] (cit. on p. 93).

[109] A. Butter and T. Plehn, *Generative Networks for LHC events*, (2020), arXiv: 2008.08558 [hep-ph] (cit. on p. 93).

[110] M. Mirza and S. Osindero, *Conditional Generative Adversarial Nets*, (2014), arXiv: 1411.1784 [cs.LG] (cit. on p. 93).

[111] D. Curtin et al., *Exotic decays of the 125 GeV Higgs boson*, Phys. Rev. D **90** (2014) 075004, arXiv: 1312.4992 [hep-ph] (cit. on p. 96).

[112] G. Aad et al., *Search for Higgs Boson Decays into a Z Boson and a Light Hadronically Decaying Resonance Using 13 TeV pp Collision Data from the ATLAS Detector*, Phys. Rev. Lett. **125** (2020) 221802, arXiv: 2004.01678 [hep-ex] (cit. on p. 96).

[113] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, http://www.deeplearningbook.org, MIT Press, 2016 (cit. on p. 96).

[114] L. G. Almeida et al., *Substructure of high-p_T Jets at the LHC*, Phys. Rev. D **79** (2009) 074017, arXiv: 0807.0234 [hep-ph] (cit. on p. 96).

[115] I. Moult, L. Necib, and J. Thaler, *New Angles on Energy Correlation Functions*, JHEP **12** (2016) 153, arXiv: 1609.07483 [hep-ph] (cit. on p. 96).

[116] ATLAS Collaboration, *Search for the exclusive W boson hadronic decays $W^{\pm} \to \pi^{\pm}\gamma$, $W^{\pm} \to K^{\pm}\gamma$ and $W^{\pm} \to \rho^{\pm}\gamma$ with the ATLAS detector*, (2023), arXiv: 2309.15887 [hep-ex] (cit. on p. 100).

[117] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph] (cit. on p. 102).

[118] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv: 1007.2241 [hep-ph] (cit. on p. 102).

[119] E. Boos et al., **?**Generic User Process Interface for Event Generators**?**, *2nd Les Houches Workshop on Physics at TeV Colliders*, 2001, arXiv: hep-ph/0109068 (cit. on p. 102).

[120] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph] (cit. on p. 102).

[121] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv: hep-ph/0201195 [hep-ph] (cit. on p. 102).

[122] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **09** (2014) 145, arXiv: 1406.3660 [hep-ex] (cit. on p. 102).

[123] S. Agostinelli et al., *GEANT4–a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 102).

[124] P. Faccioli and C. Lourenço, *Particle Polarization in High Energy Physics*, Springer Cham, 2022, URL: https://link.springer.com/book/10.1007/978-3-031-08876-6 (cit. on pp. 104, 108).

[125] R. Kutschke, "*An Angular Distribution Cookbook*, (1996), unpublished, URL: https://home.fnal.gov/~kutschke/Angdist/angdist.ps (cit. on p. 105).

[126] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex] (cit. on pp. 113, 115).

[127] P. Krivkova and R. Leitner, *Measurement of the interaction length of pions and protons in the TileCal Module 0*, tech. rep., CERN, 1999, URL: https://cds.cern.ch/record/683812 (cit. on p. 126).

[128] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data*, Journal of Instrumentation **14** (2019) P03017, URL: https://dx.doi.org/10.1088/1748-0221/14/03/P03017 (cit. on p. 151).

[129] G. Aad et al., *Measurement of the energy response of the ATLAS calorimeter to charged pions from $W^{\pm} \to \tau^{\pm}(\to \pi^{\pm}\nu_{\tau})\nu_{\tau}$ events in Run 2 data*, Eur. Phys. J. C **82** (2022) 223, arXiv: 2108.09043 [hep-ex] (cit. on pp. 153, 154).

[130] O. Behnke, K. Kröninger, T. Schörner-Sadenius, and G. Schott, *Data analysis in high energy physics: A practical guide to statistical methods*, Wiley-VCH, 2013, ISBN: 978-3-527-41058-3, 978-3-527-65344-7, 978-3-527-65343-0 (cit. on p. 157).

[131] A. L. Read, *Presentation of search results: the CL_S technique*, J. Phys. G **28** (2002) 2693 (cit. on p. 159).

[132] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, The European Physical Journal C **71** (2011), ISSN: 1434-6052, URL: http://dx.doi.org/10.1140/epjc/s10052-011-1554-0 (cit. on p. 159).

[133] F. James, *MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1*, (1994) (cit. on p. 160).

[134] A. Blondel and P. Janot, *FCC-ee overview: new opportunities create new challenges*, Eur. Phys. J. Plus **137** (2022) 92, arXiv: 2106.13885 [hep-ex] (cit. on p. 176).

[135] T. Melia, *Exclusive hadronic W decay: $W \to \pi\gamma$ and $W \to \pi^+\pi^+\pi^-$*, Nucl. Part. Phys. Proc. **273-275** (2016) 2102, ed. by M. Aguilar-Benítez, J. Fuster, S. Martí-García, and A. Santamaría (cit. on p. 176).

[136] M. Aleksa et al., *Conceptual design of an experiment at the FCC-hh, a future 100 TeV hadron collider*, **2/2022** (2022), ed. by M. L. Mangano and W. Riegler (cit. on p. 177).

Appendix A

Cluster dE/dx fit results

Figs. A.1 to A.8 show the results of the binned maximum likelihood fit performed to the d$E/\mathrm{d}x$ distribution of single track clusters from inside the jet core for the different $p_{\mathrm{T}}(\mathrm{jet})$ and $|\eta(\mathrm{trk})|$ bins considered in the analysis. Figs. A.9 to A.16 show the results of the fit to the d$E/\mathrm{d}x$ distribution of two track clusters from inside the jet core, in the same bins.

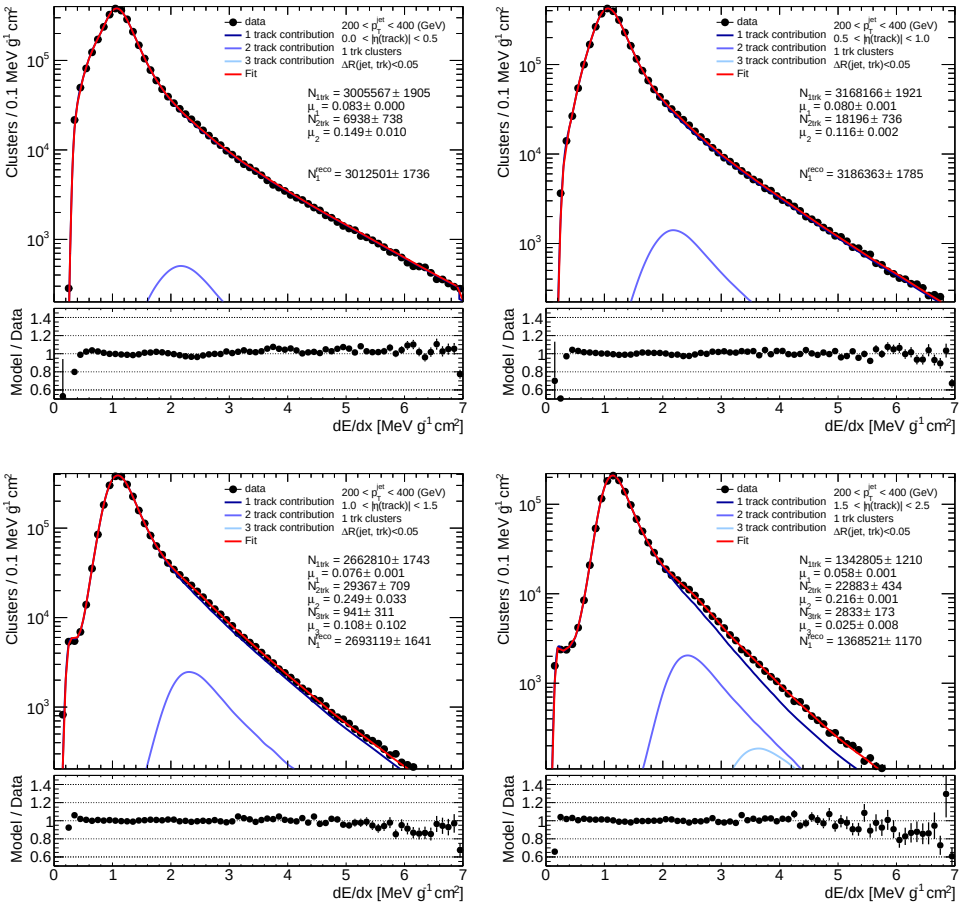


Figure A.1: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of single track clusters from inside the jet core, for $200 < p_{\mathrm{T}}(\mathrm{jet}) < 400$ GeV. The bottom panel displays the ratio between the post-fit model and the data.The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_1^{\mathrm{reco}}$).

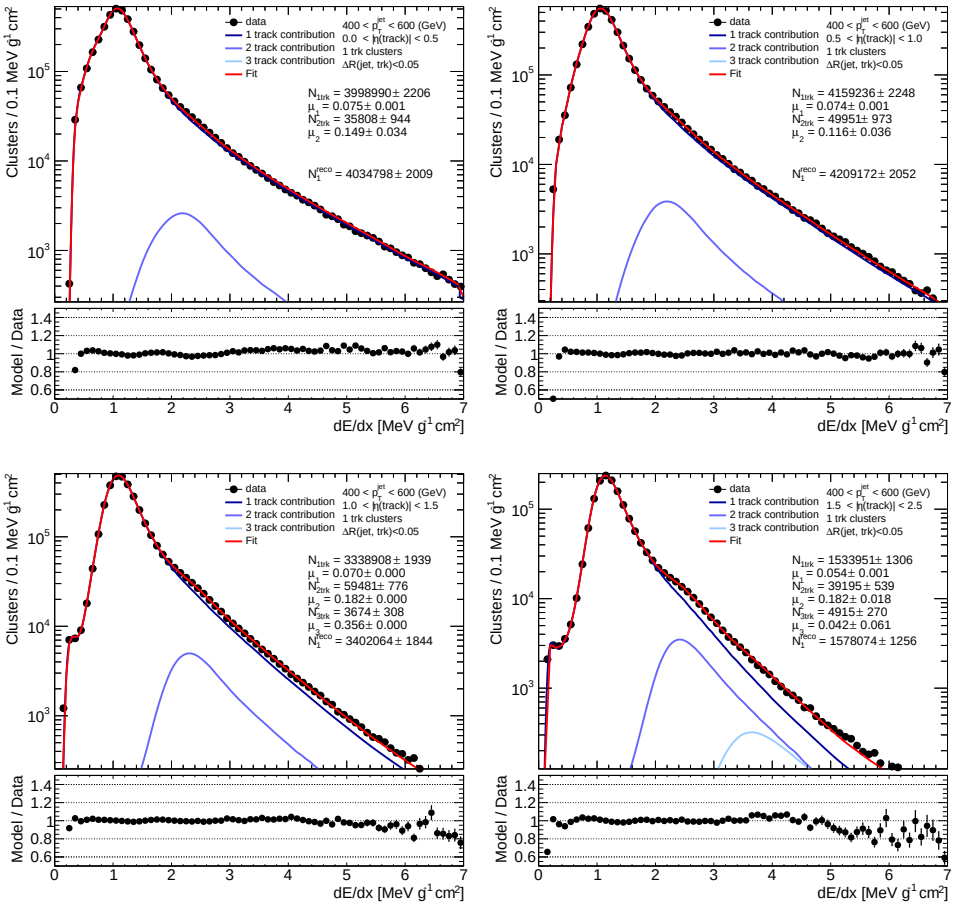


Figure A.2: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of single track clusters from inside the jet core, for $400 < p_{\mathrm{T}}(\mathrm{jet}) < 600$ GeV. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_1^{\mathrm{reco}}$).

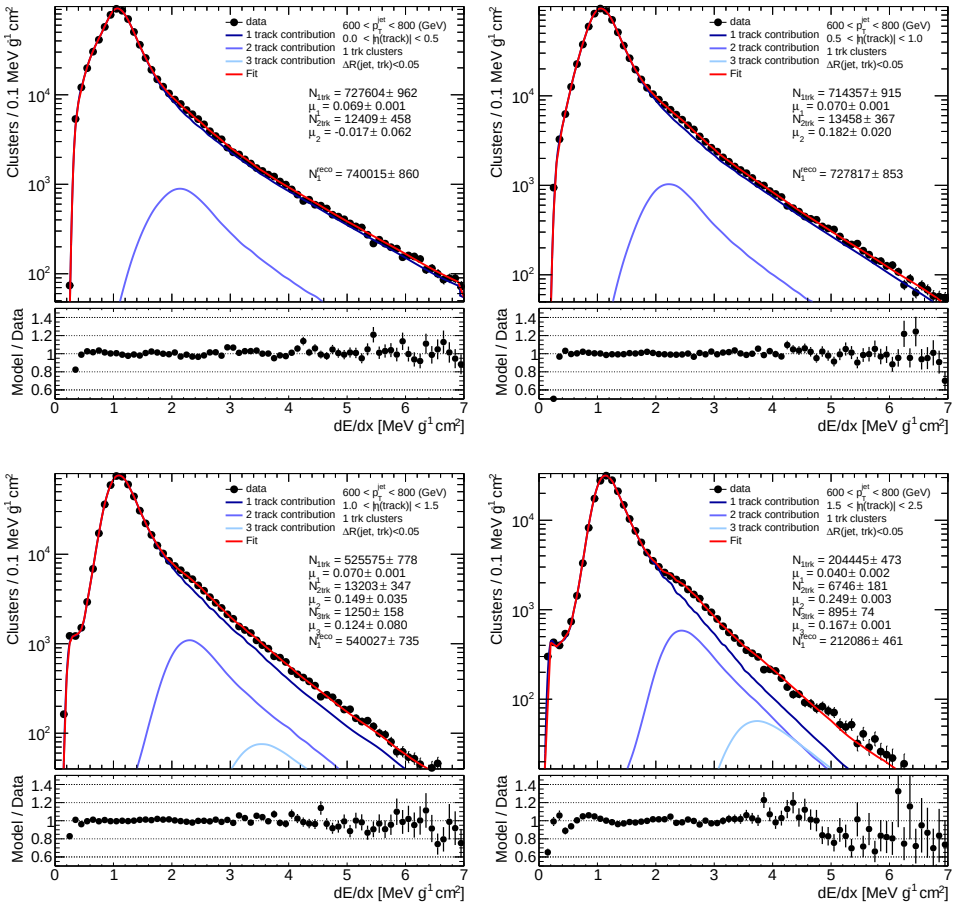


Figure A.3: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of single track clusters from inside the jet core, for $600 < p_{\mathrm{T}}(\mathrm{jet}) < 800$ GeV. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_1^{\mathrm{reco}}$).

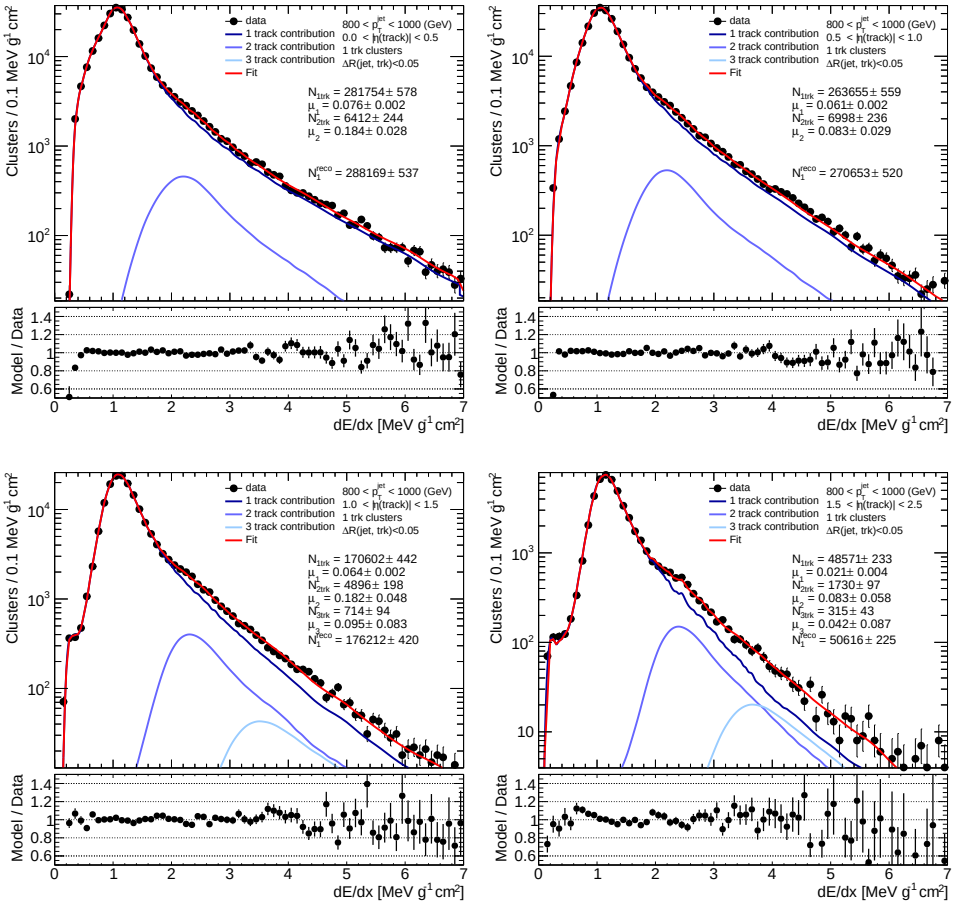


Figure A.4: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of single track clusters from inside the jet core, for $800 < p_{\mathrm{T}}(\mathrm{jet}) < 1000$ GeV. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_1^{\mathrm{reco}}$).

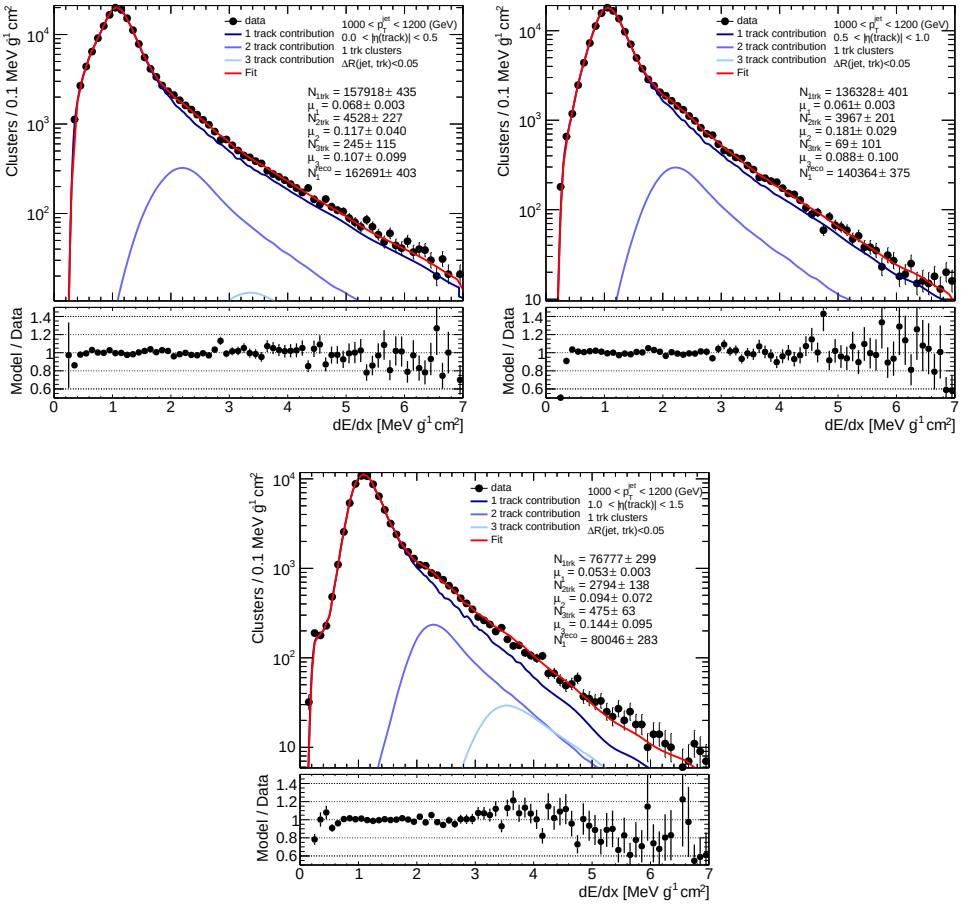


Figure A.5: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of single track clusters from inside the jet core, for $1000 < p_{\mathrm{T}}(\mathrm{jet}) < 1200$ GeV and $0 < |\eta| < 1.5$. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_1^{\mathrm{reco}}$).

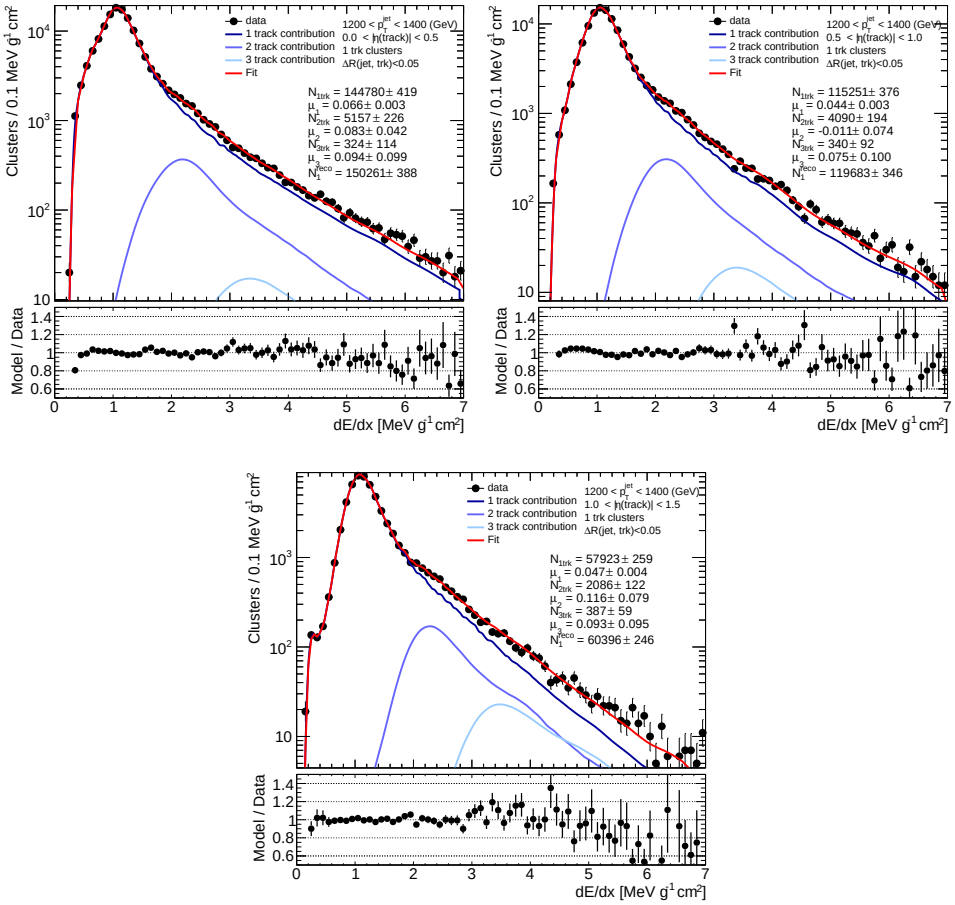


Figure A.6: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of single track clusters from inside the jet core, for $1200 < p_{\mathrm{T}}(\mathrm{jet}) < 1400$ GeV and $0 < |\eta| < 1.5$. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_1^{\mathrm{reco}}$).

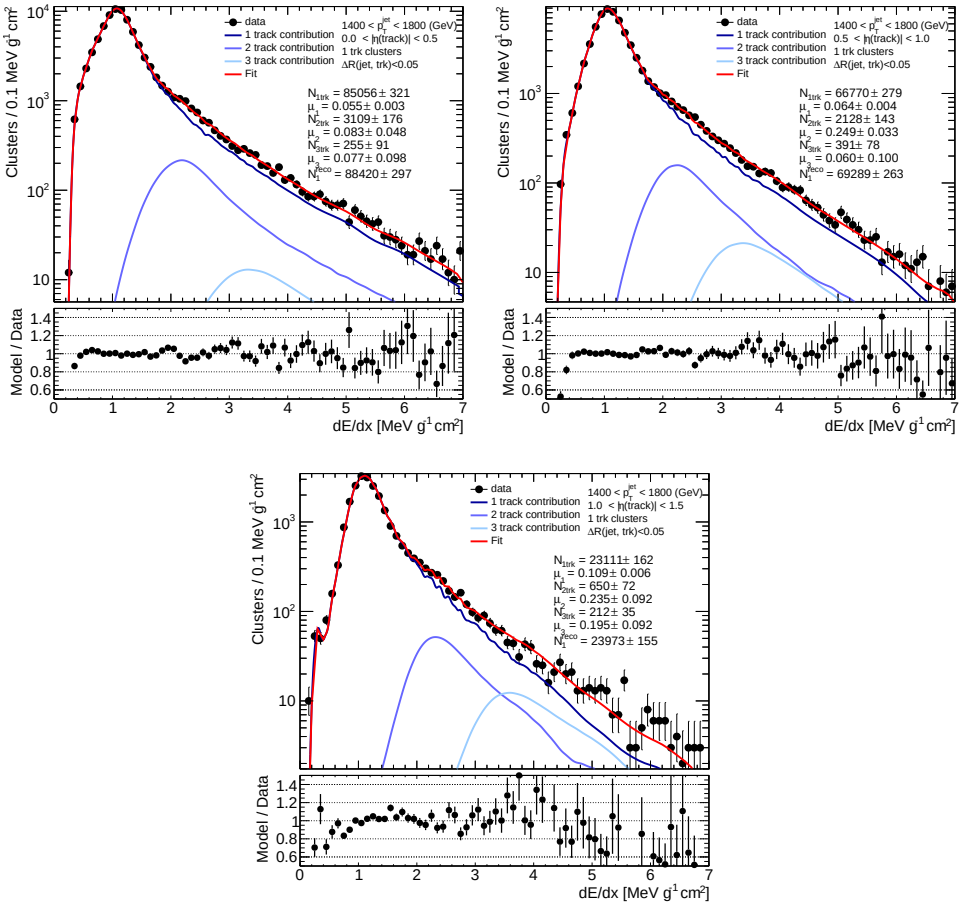


Figure A.7: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of single track clusters from inside the jet core, for $1400 < p_{\mathrm{T}}(\mathrm{jet}) < 1800$ GeV and $0 < |\eta| < 1.5$. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_1^{\mathrm{reco}}$).

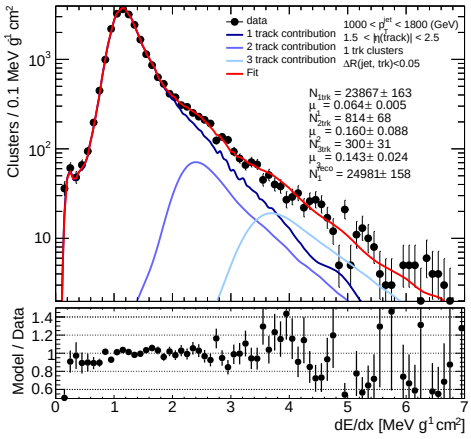


Figure A.8: Result of the binned maximum likelihood fits to the d$E/$dx distribution of single track clusters from inside the jet core, for $1000 < p_{\mathrm{T}}(\mathrm{jet}) < 1800$ GeV and $1.5 < |\eta| < 2.5$. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_1^{\mathrm{reco}}$).

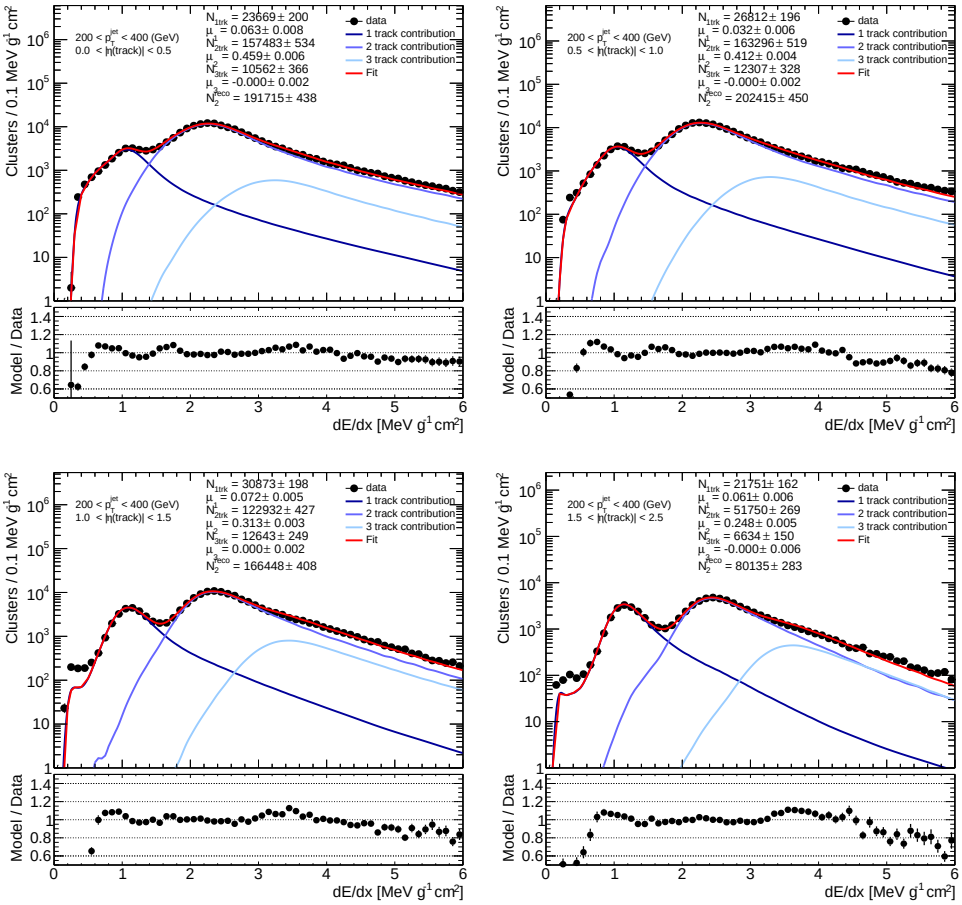


Figure A.9: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of two track clusters from inside the jet core, for $200 < p_{\mathrm{T}}(\mathrm{jet}) < 400$ GeV. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_2^{\mathrm{reco}}$).

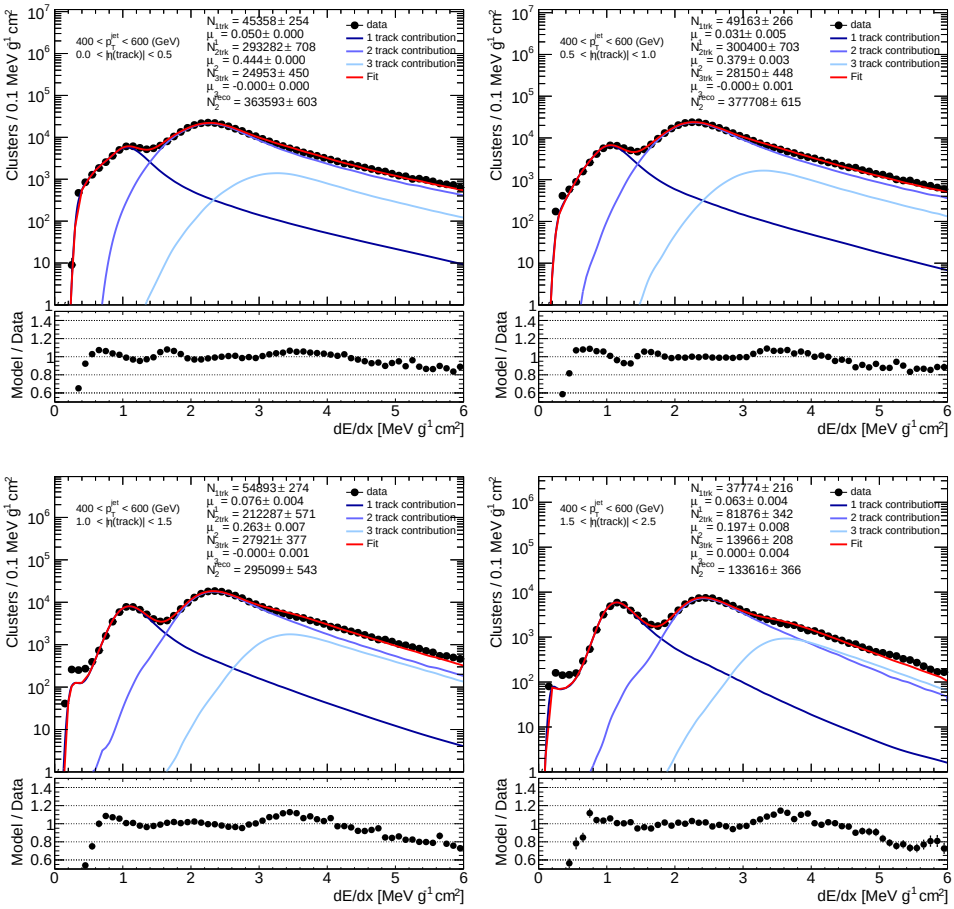


Figure A.10: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of two track clusters from inside the jet core, for $400 < p_{\mathrm{T}}(\mathrm{jet}) < 600$ GeV. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_2^{\mathrm{reco}}$).

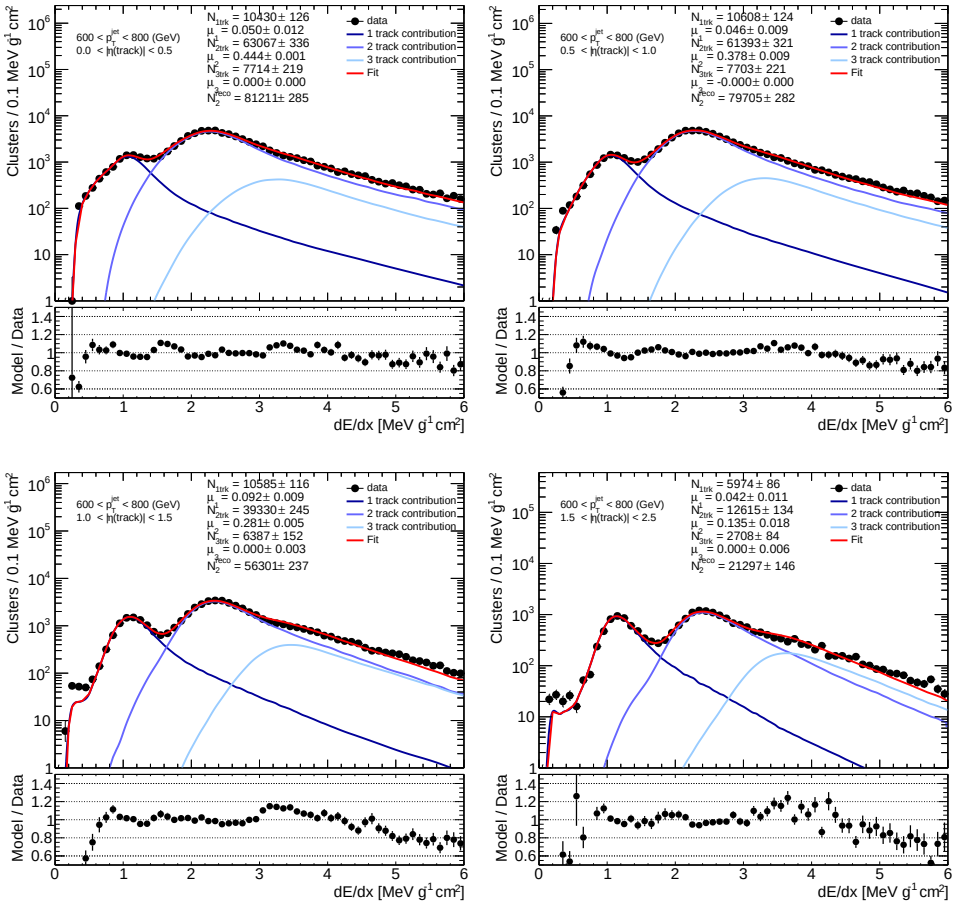


Figure A.11: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of two track clusters from inside the jet core, for $600 < p_{\mathrm{T}}(\mathrm{jet}) < 800$ GeV. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_2^{\mathrm{reco}}$).

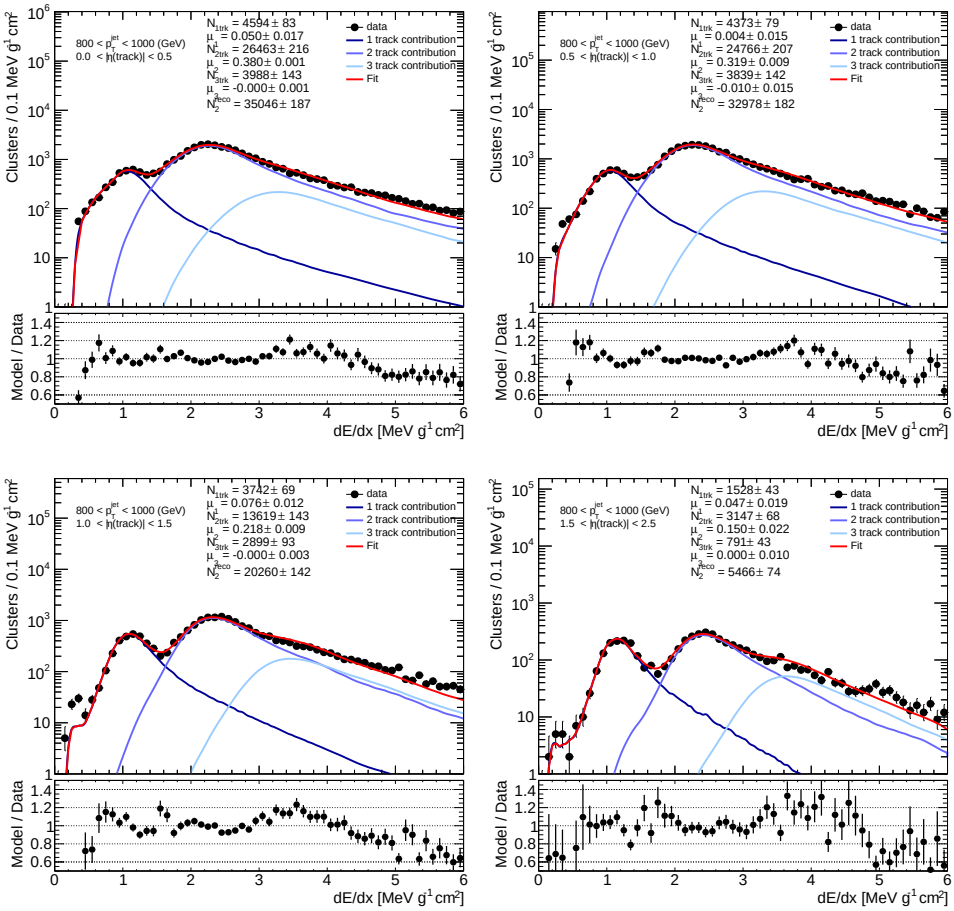


Figure A.12: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of two track clusters from inside the jet core, for $800 < p_{\mathrm{T}}(\mathrm{jet}) < 1000$ GeV. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_2^{\mathrm{reco}}$).

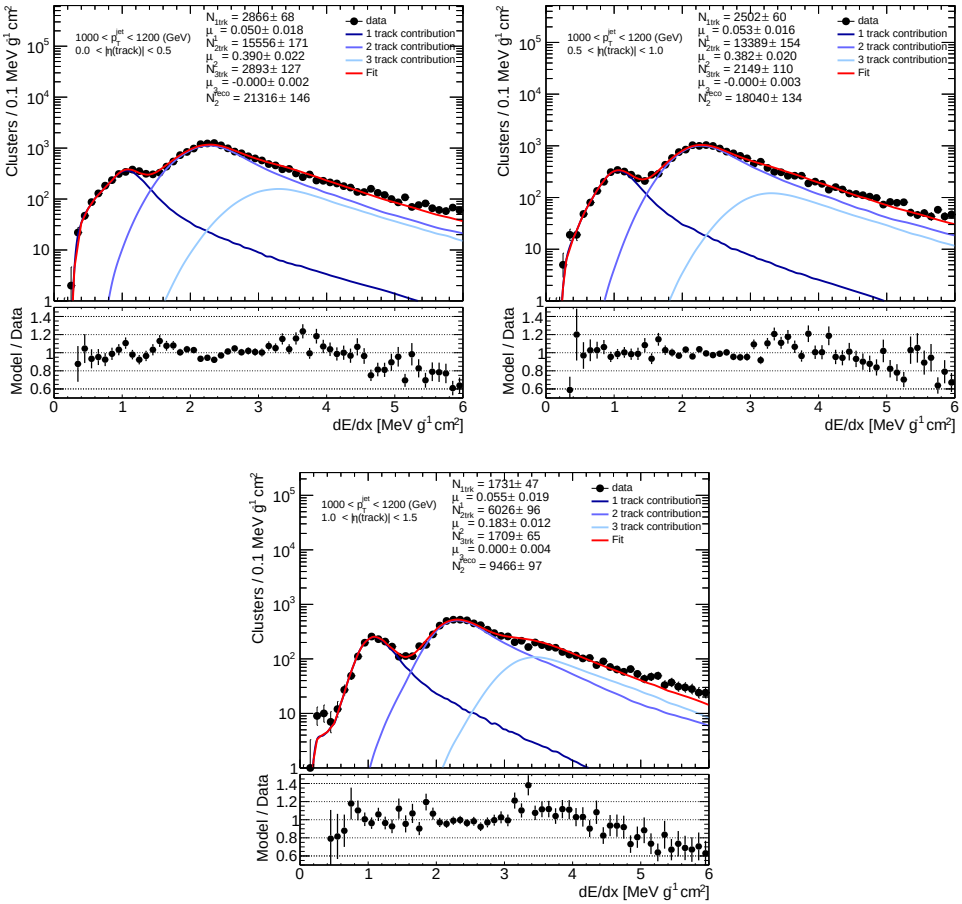


Figure A.13: Result of the binned maximum likelihood fits to rhe $\mathrm{d}E/\mathrm{d}x$ distribution of two track clusters from inside the jet core, for $1000 < p_{\mathrm{T}}(\mathrm{jet}) < 1200$ GeV and $0 < |\eta| < 1.5$. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_2^{\mathrm{reco}}$).

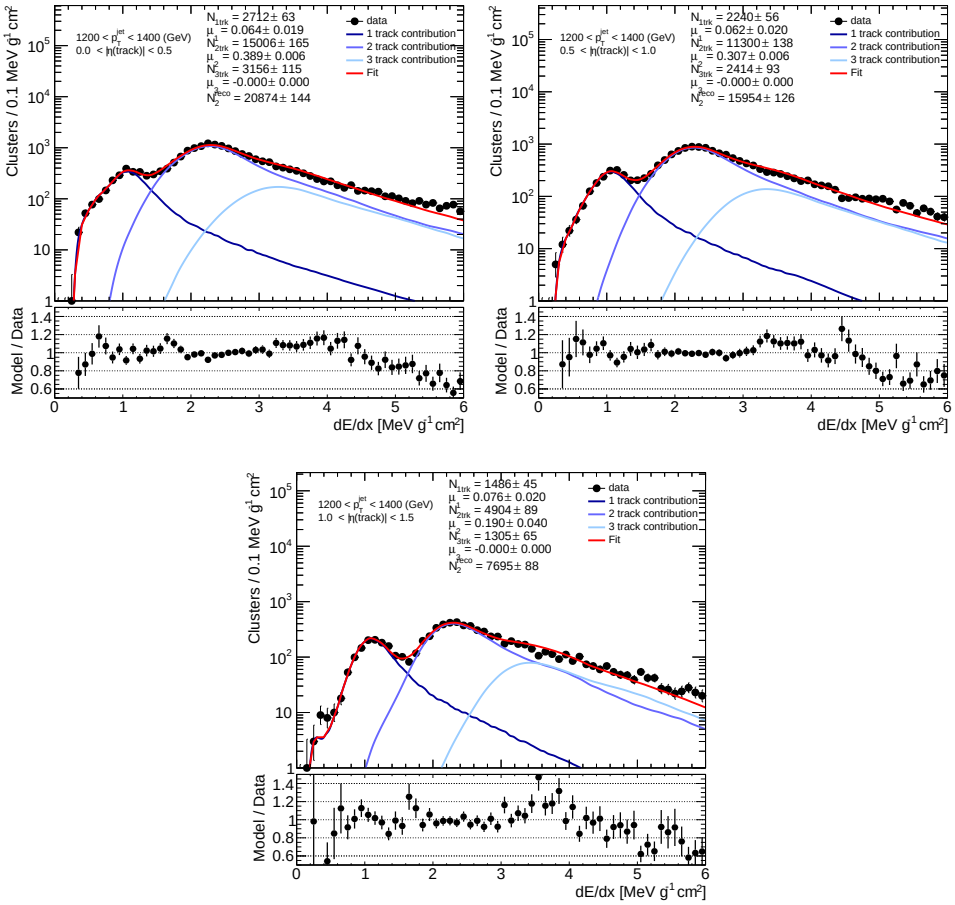


Figure A.14: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of two track clusters from inside the jet core, for $1200 < p_{\mathrm{T}}(\mathrm{jet}) < 1400$ GeV and $0 < |\eta| < 1.5$. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_2^{\mathrm{reco}}$).

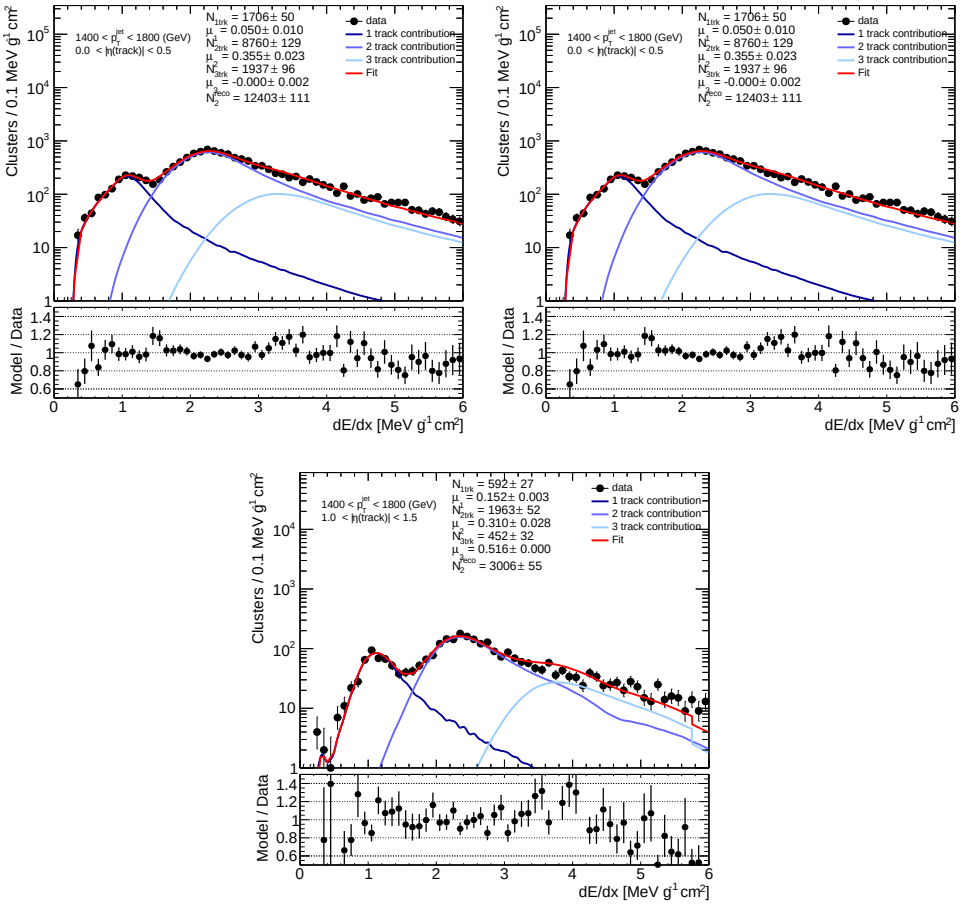


Figure A.15: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of two track clusters from inside the jet core, for $1400 < p_{\mathrm{T}}(\mathrm{jet}) < 1800$ GeV and $0 < |\eta| < 1.5$. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_2^{\mathrm{reco}}$).

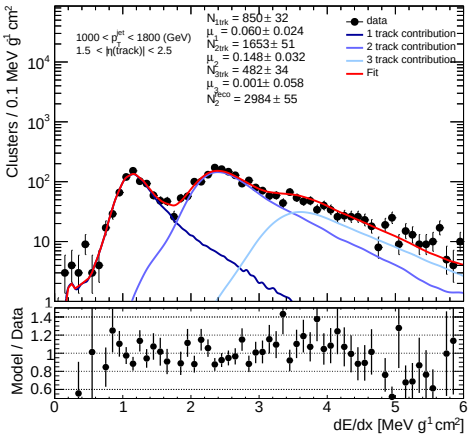


Figure A.16: Result of the binned maximum likelihood fits to the $\mathrm{d}E/\mathrm{d}x$ distribution of two track clusters from inside the jet core, for $1000 < p_{\mathrm{T}}(\mathrm{jet}) < 1800$ GeV and $1.5 < |\eta| < 2.5$. The bottom panel displays the ratio between the post-fit model and the data. The post-fit values of the fit parameters are reported, as well as the total number of clusters in the measurement distribution ($\mathrm{N}_2^{\mathrm{reco}}$).

Appendix B

KDE smoothing of single-track templates

The single track templates obtained after KDE smoothing can be found in Figs. B.1 to B.7, for the different$p_{\mathrm{T}}(\mathrm{jet})$ and $|\eta(\mathrm{trk})|$ bins considered in the analysis.

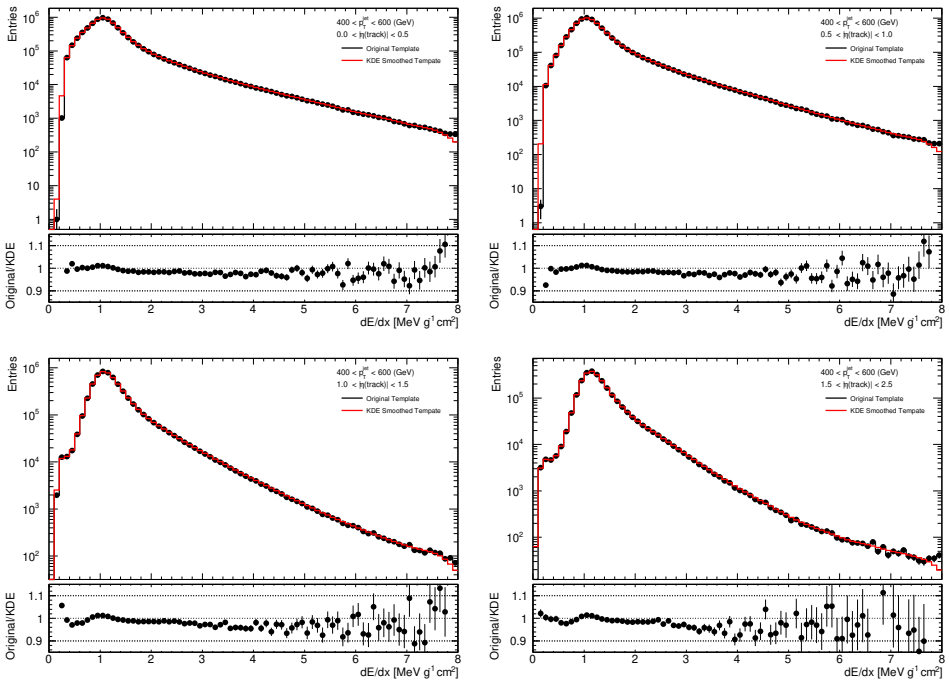


Figure B.1: Single track template before and after KDE smoothing, for $400 < p_{\mathrm{T}}(\mathrm{jet}) < 600$ GeV.

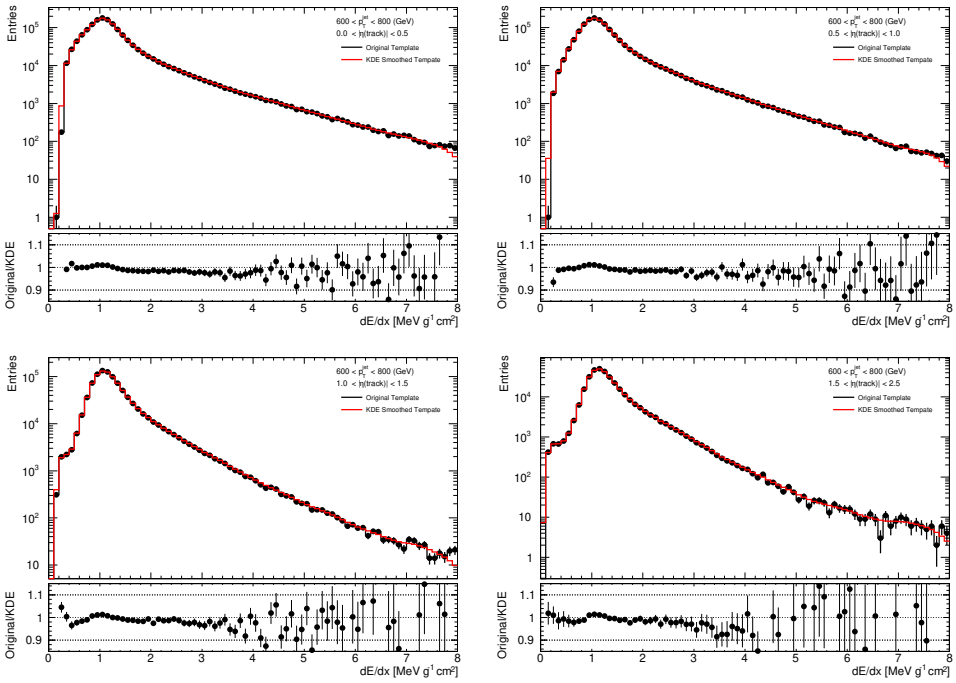


Figure B.2: Single track template before and after KDE smoothing, for $600 < p_{\mathrm{T}}(\mathrm{jet}) < 800$ GeV.

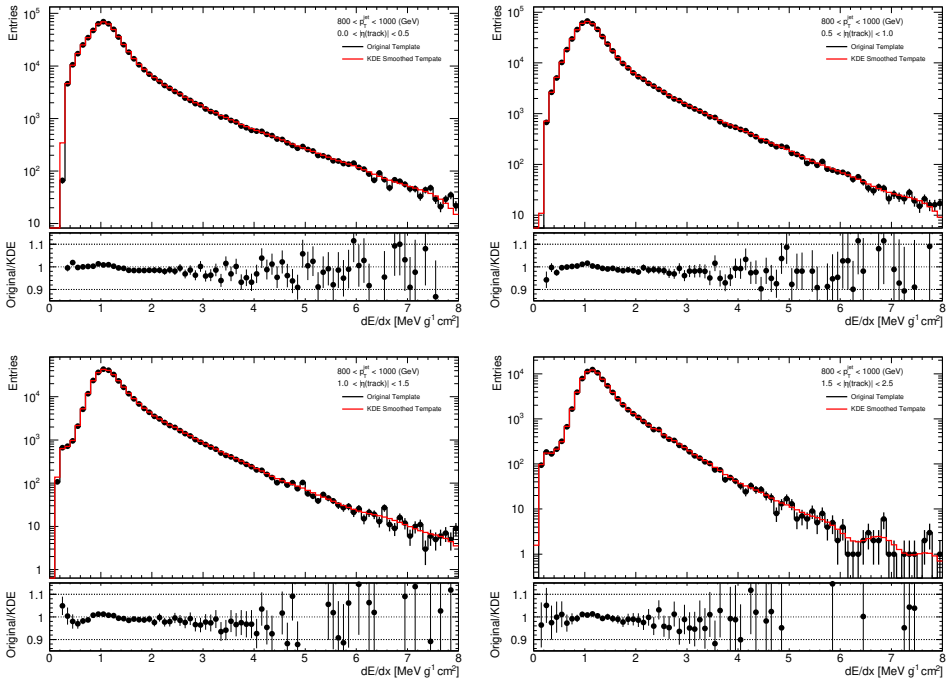


Figure B.3: Single track template before and after KDE smoothing, for $800 < p_{\mathrm{T}}(\mathrm{jet}) < 1000$ GeV.

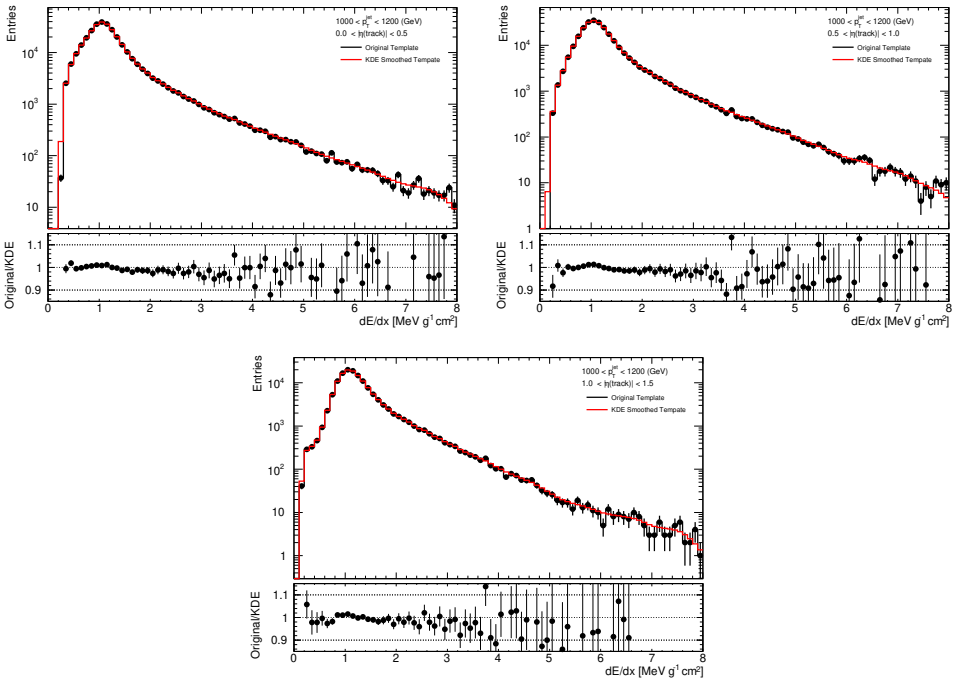


Figure B.4: Single track template before and after KDE smoothing, for $1000 < p_{\mathrm{T}}(\mathrm{jet}) < 1200$ GeV and $0 < |\eta(\mathrm{trk})| < 1.5$.

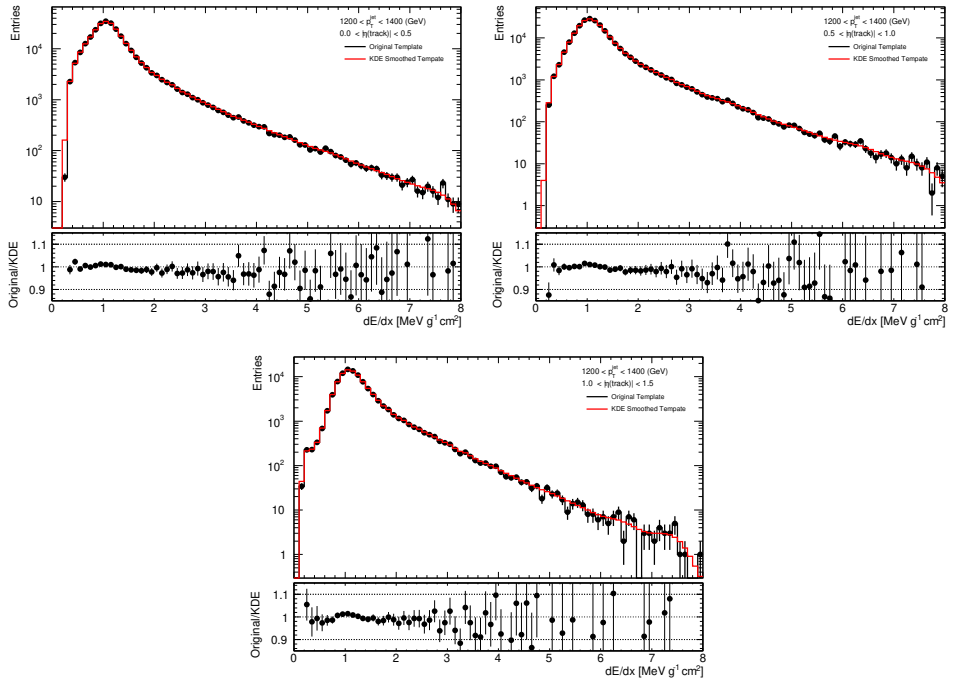


Figure B.5: Single track template before and after KDE smoothing, for $1200 < p_{\mathrm{T}}(\mathrm{jet}) < 1400$ GeV and $0 < |\eta(\mathrm{trk})| < 1.5$.

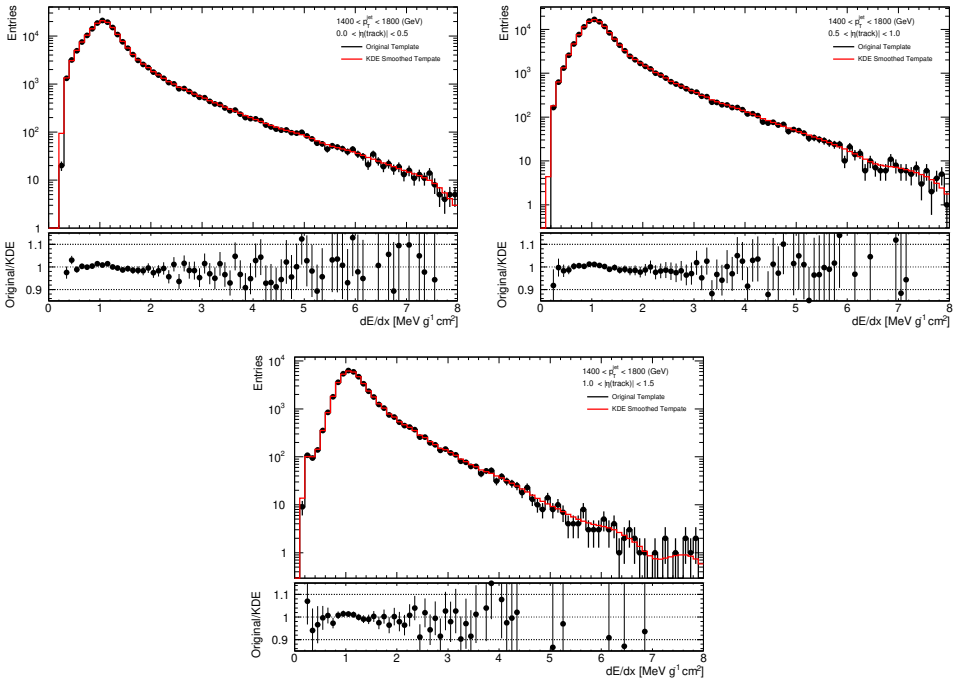


Figure B.6: Single track template before and after KDE smoothing, for $1400 < p_{\mathrm{T}}(\mathrm{jet}) < 1800$ GeV and $0 < |\eta(\mathrm{trk})| < 1.5$.

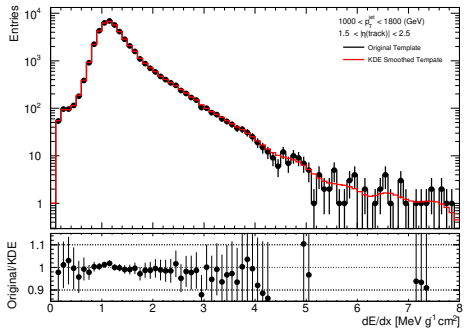
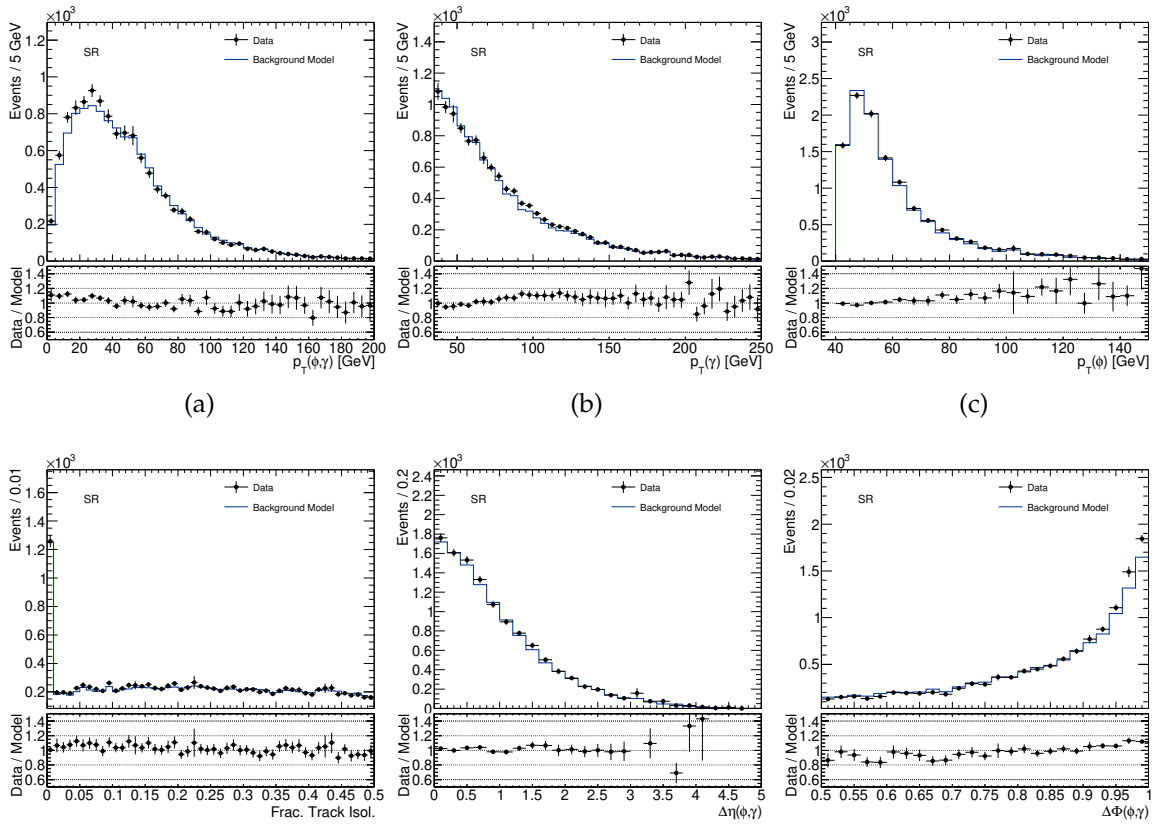


Figure B.7: Single track template before and after KDE smoothing, for $1000 < p_{\mathrm{T}}(\mathrm{jet}) < 1800$ GeV and $1.5 < |\eta(\mathrm{trk})| < 2.5$.

Appendix C

Data-model comparisons in SR

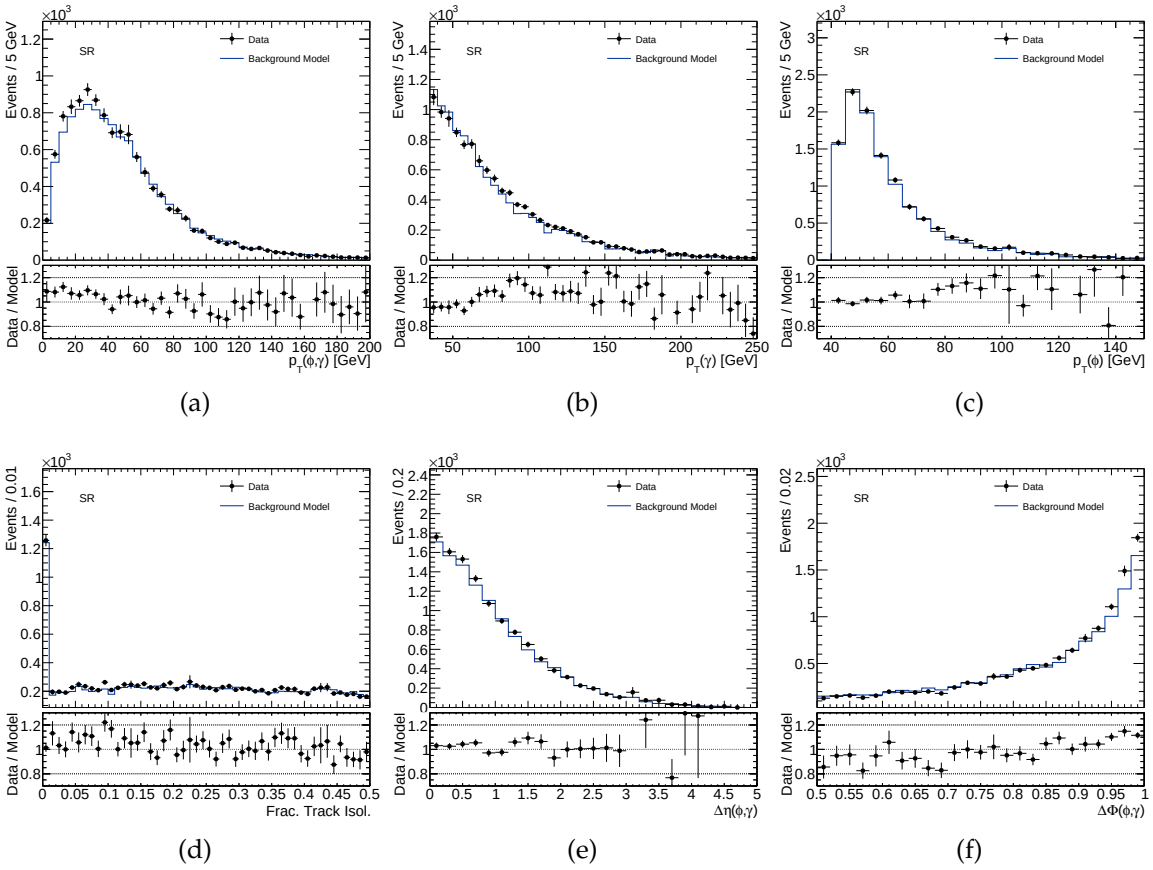
Fig. C.1 and Fig. C.2 show the SR distributions of the primary variables of the modelling and of another compound variable, $p_{\mathrm{T}}(\phi,\gamma)$, once again for both realisations of the sampling. Reasonable data-model agreement is seen for all variables.



(a) (b) (c)

(d) (e) (f)

Figure C.1: Distributions of kinematic and isolation variables in the SR, for the simulated data and background model built using the events in GR directly.



(a) (b) (c)

(d) (e) (f)

Figure C.2: Distributions of kinematic and isolation variables in the SR, for the simulated data and background model built after sampling with replacement from the GR.

Appendix D

Control plots for track-photon final state

Figs. D.1 to D.8 show the background prediction compared to data in all regions, for the track-photon final state. The $W^{\pm} \to \pi^{\pm}\gamma$, $W^{\pm} \to K^{\pm}\gamma$ and $W^{\pm} \to \rho^{\pm}\gamma$ signal and $Z \to e^{-}e^{+}$ background distributions are also shown. The signal distributions are scaled to branching ratios that improve their visualisation.

The distributions of the kinematic and isolation variables directly included in the modelling are shown. Even though the agreement is not perfect for some of the variables, the model is still able to describe compound variables like the $m(\text{track}, \gamma)$, as can be seen in **??** and $p_{\text{T}}(\text{track}, \gamma)$.

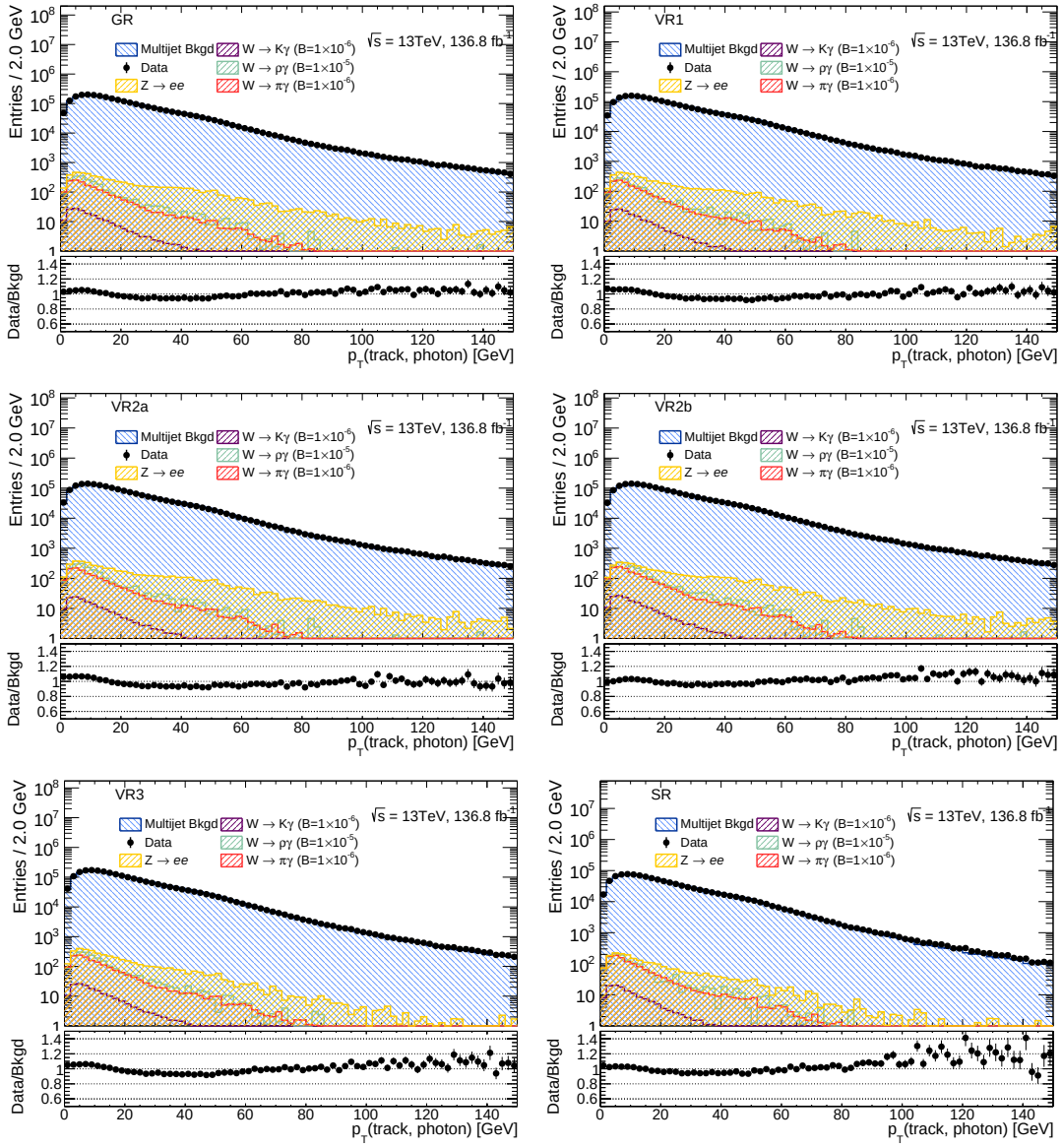


Figure D.1: p_{T}(track, photon) distribution for data, background and signal in the track-photon GR, VR1, VR2a, VR2b, VR3 and SR.

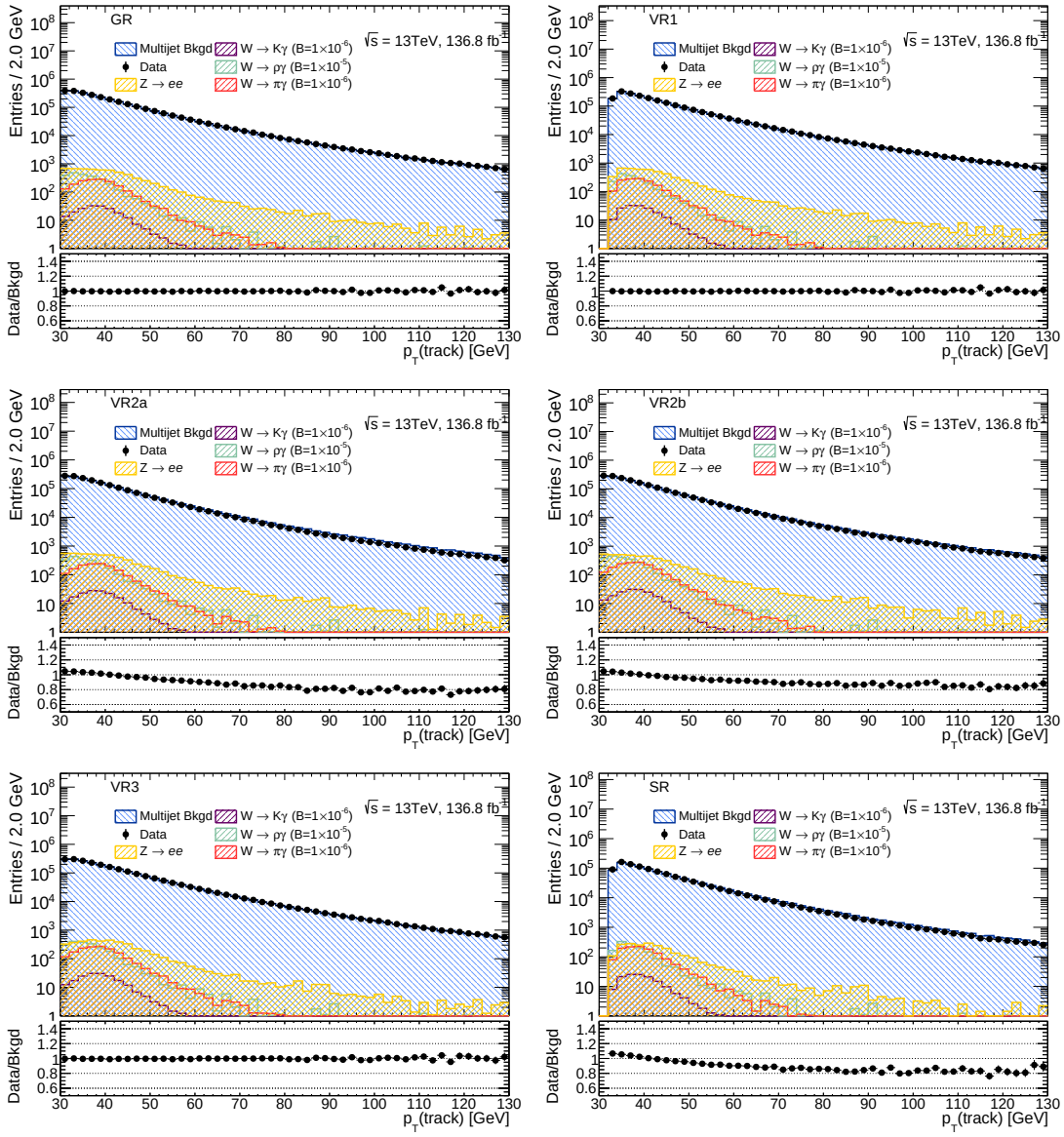


Figure D.2: p_{T}(track) distribution for data, background and signal in the track-photon GR, VR1, VR2a, VR2b, VR3 and SR.

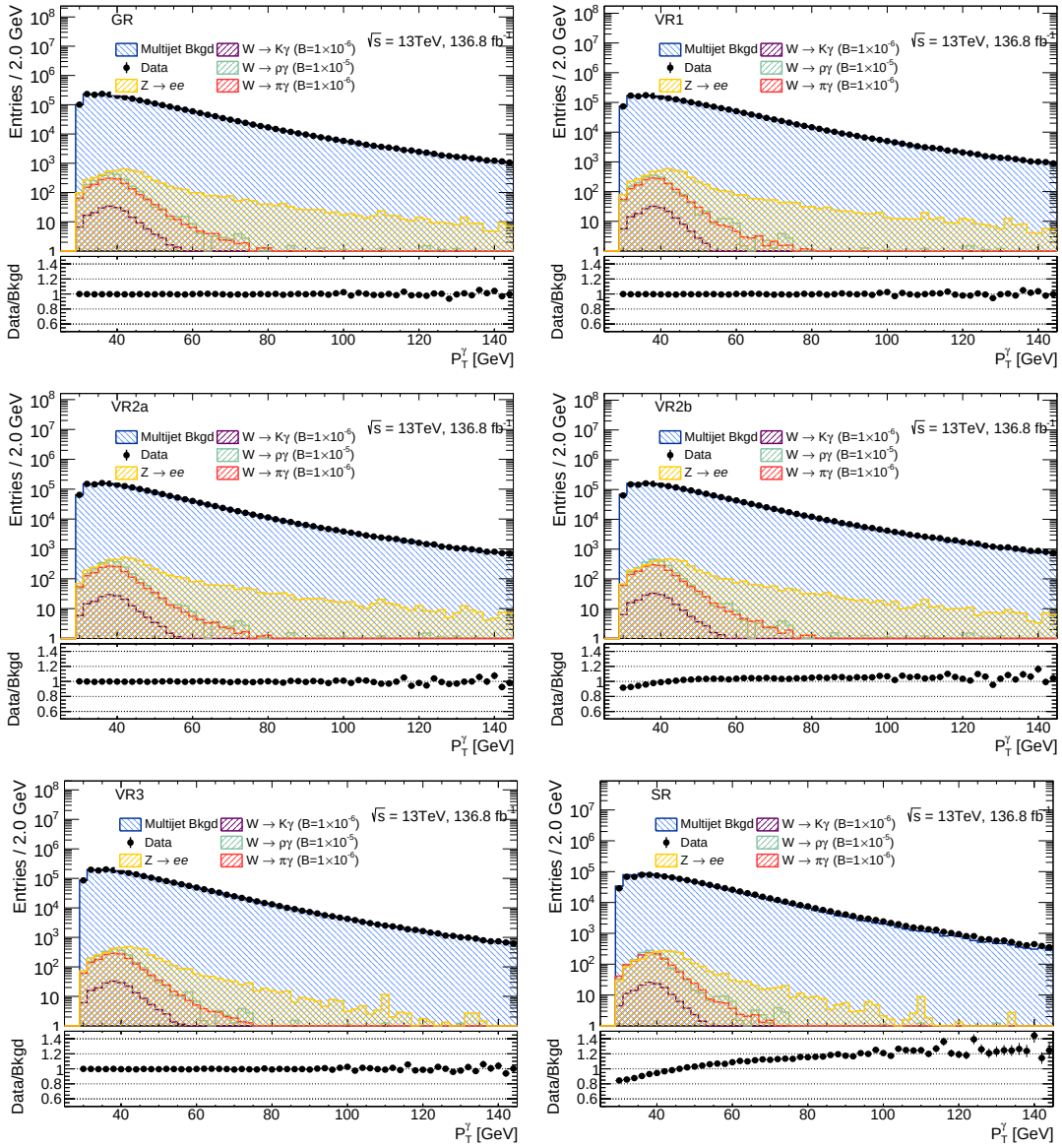


Figure D.3: $p_{\mathrm{T}}(\gamma)$ distribution for data, background and signal in the track-photon GR, VR1, VR2a, VR2b, VR3 and SR.

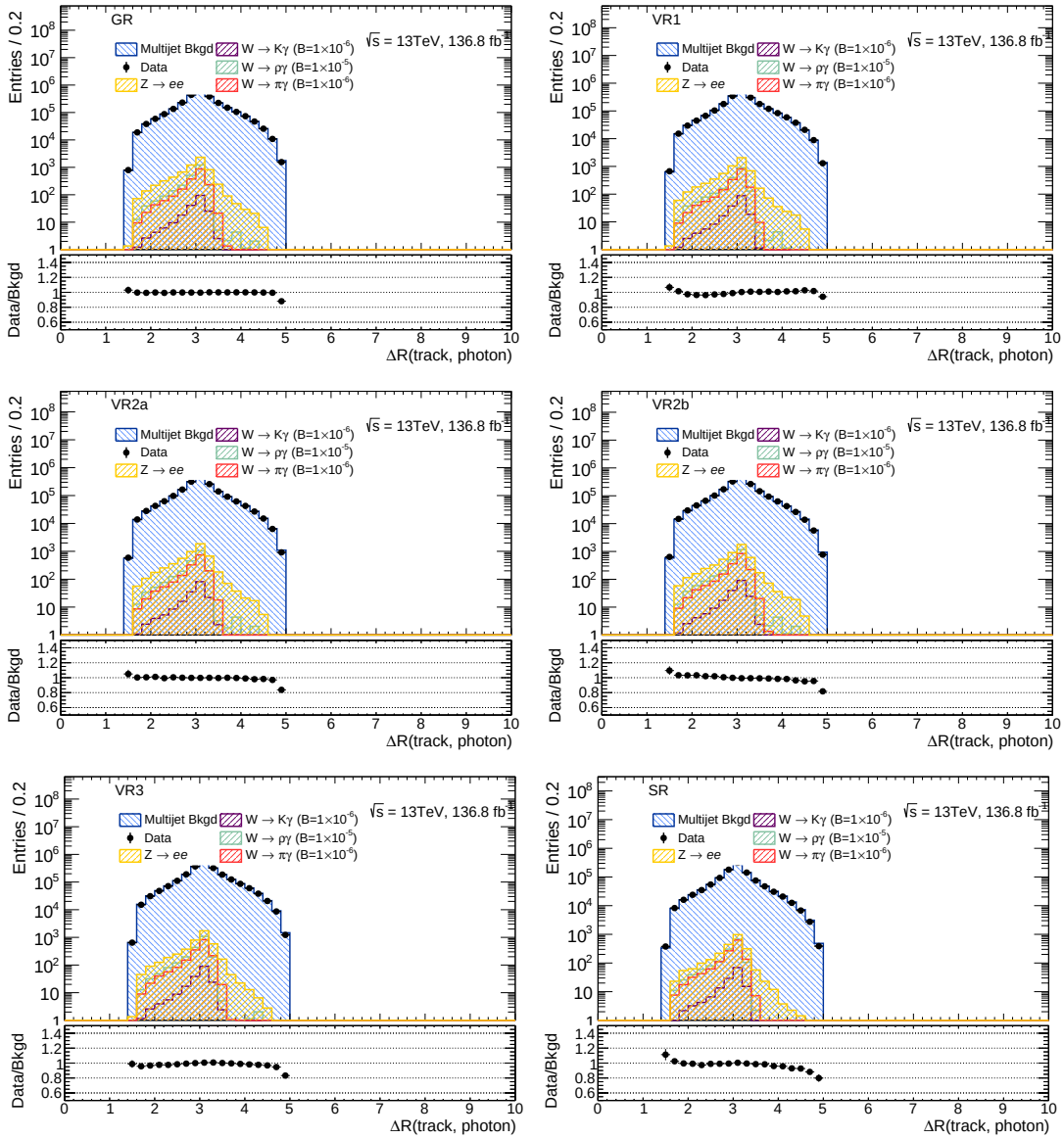


Figure D.4: ΔR(track, photon) distribution for data, background and signal in the track-photon GR, VR1, VR2a, VR2b, VR3 and SR.

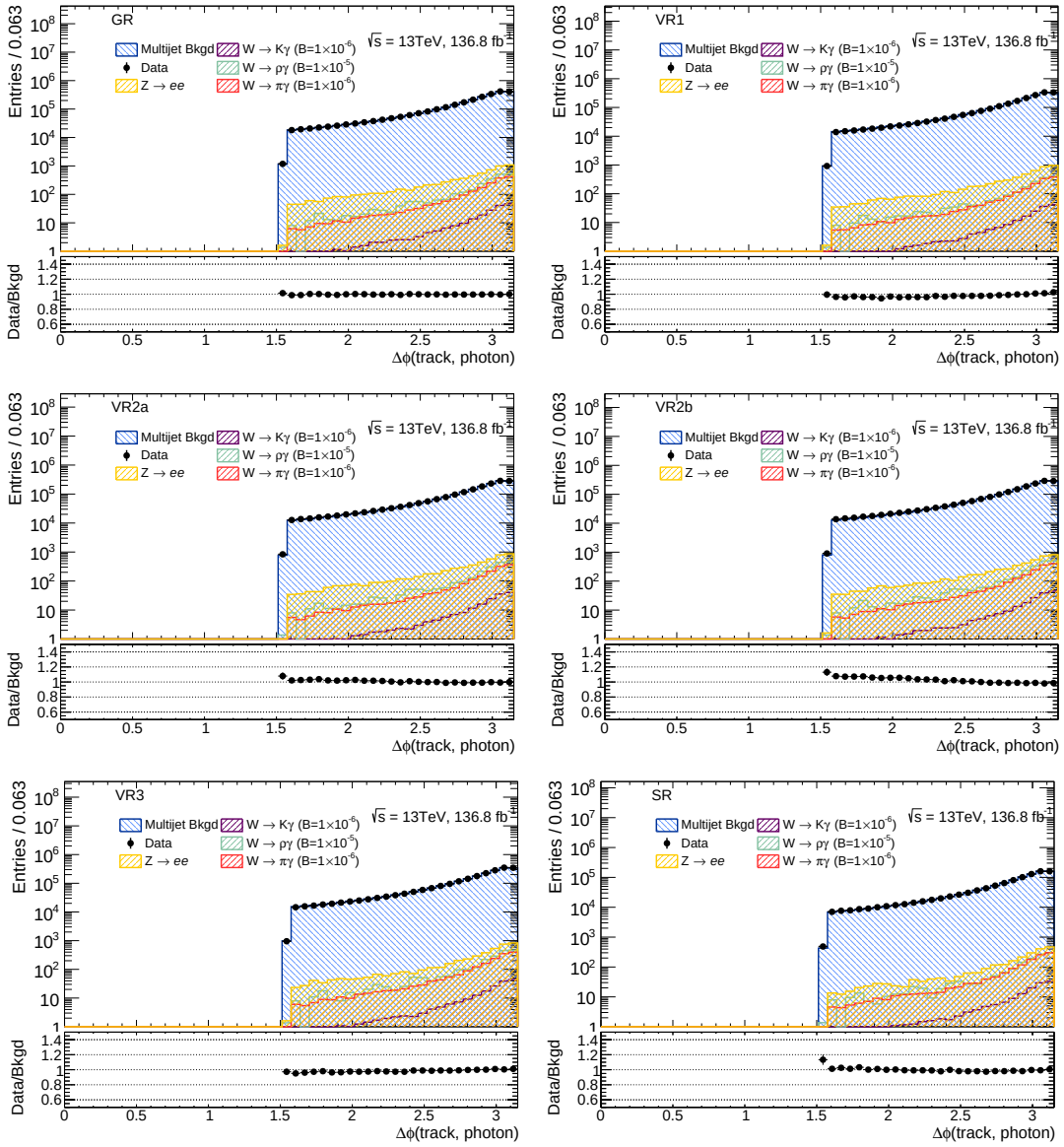


Figure D.5: $\Delta\phi$(track, photon) distribution for data, background and signal in the track-photon GR, VR1, VR2a, VR2b, VR3 and SR.

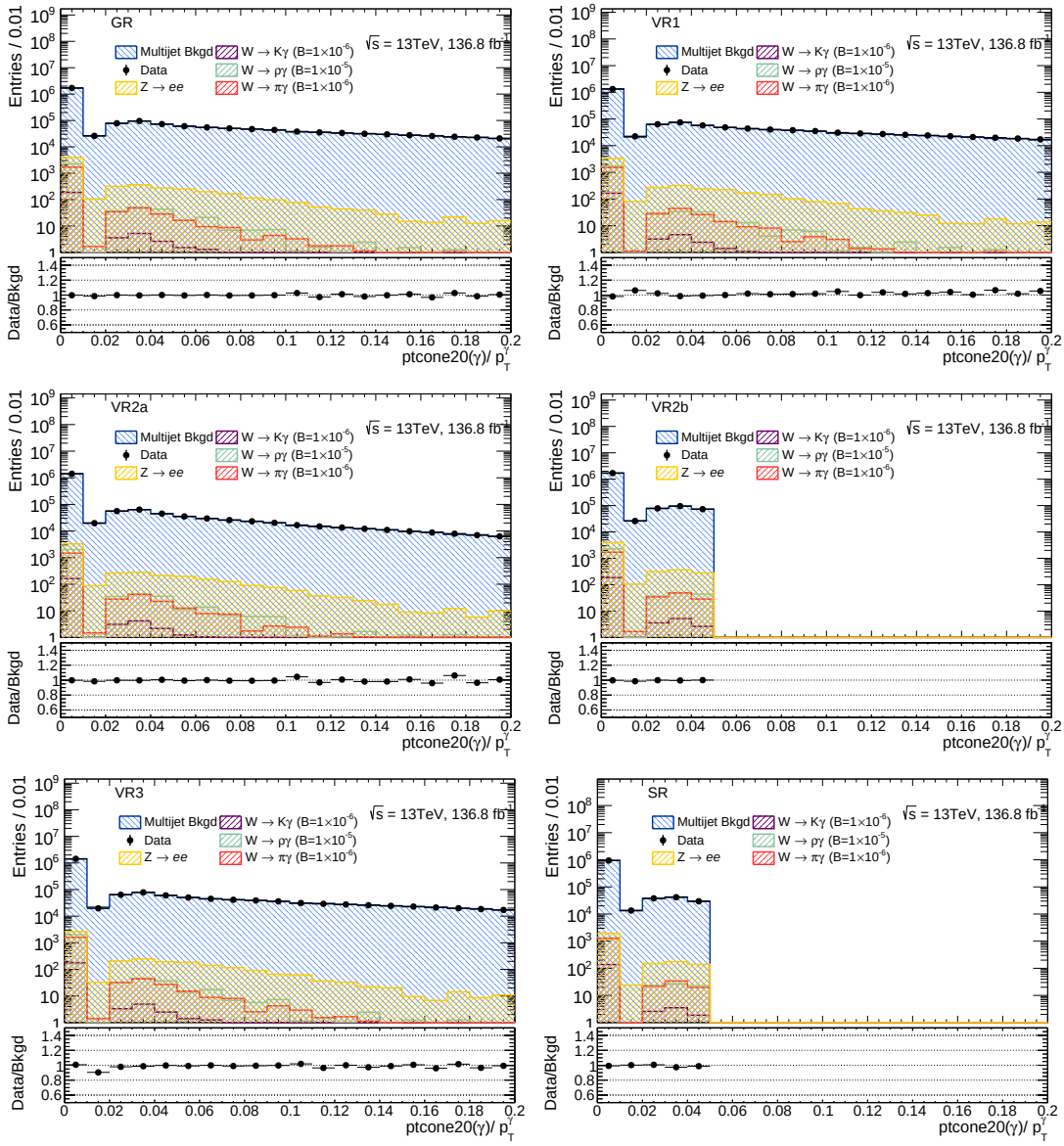


Figure D.6: Photon track isolation distribution for data, background and signal in the track-photon GR, VR1, VR2a, VR2b, VR3 and SR.

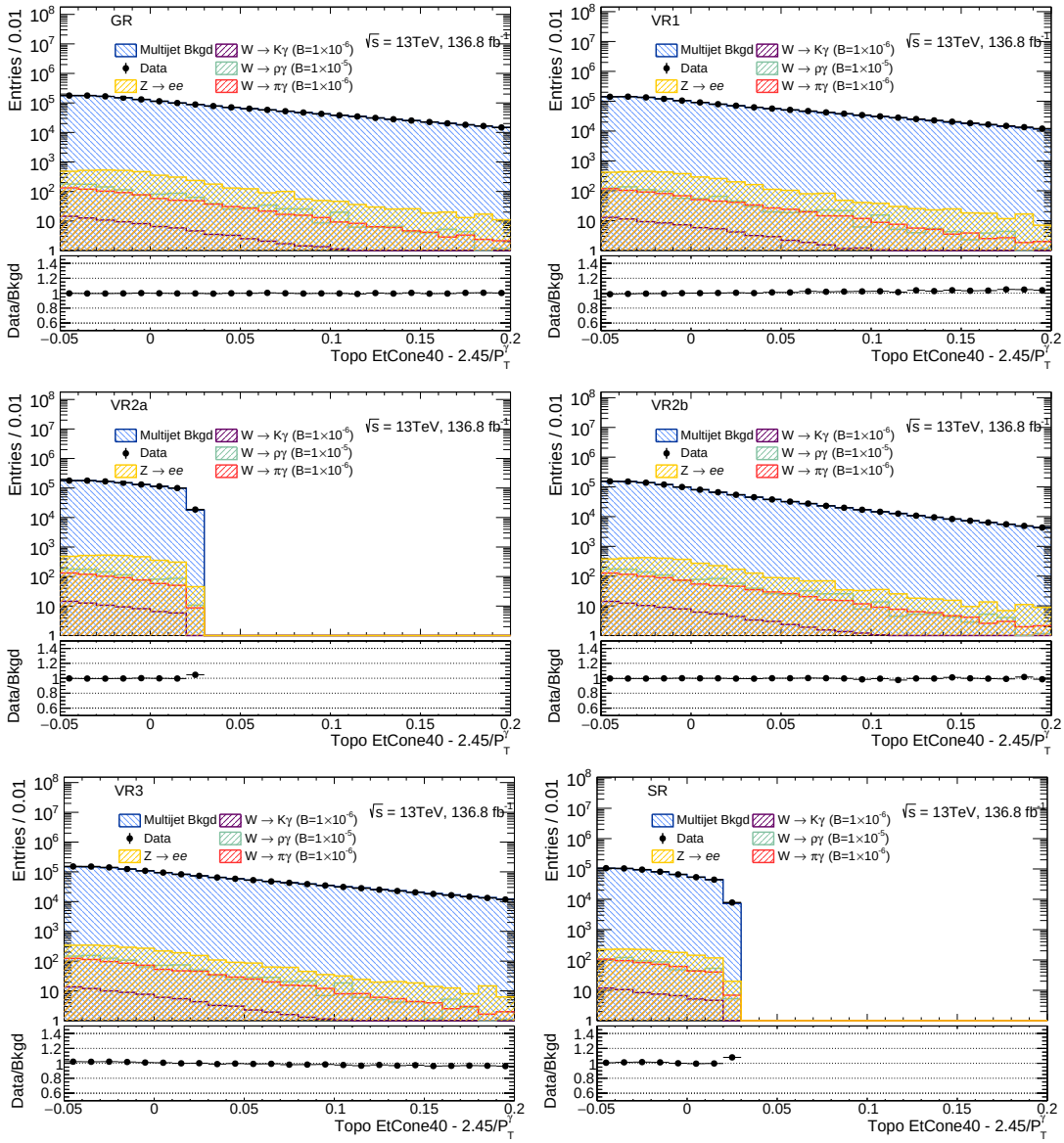


Figure D.7: Photon calo isolation distribution for data, background and signal in the track-photon GR, VR1, VR2a, VR2b, VR3 and SR.

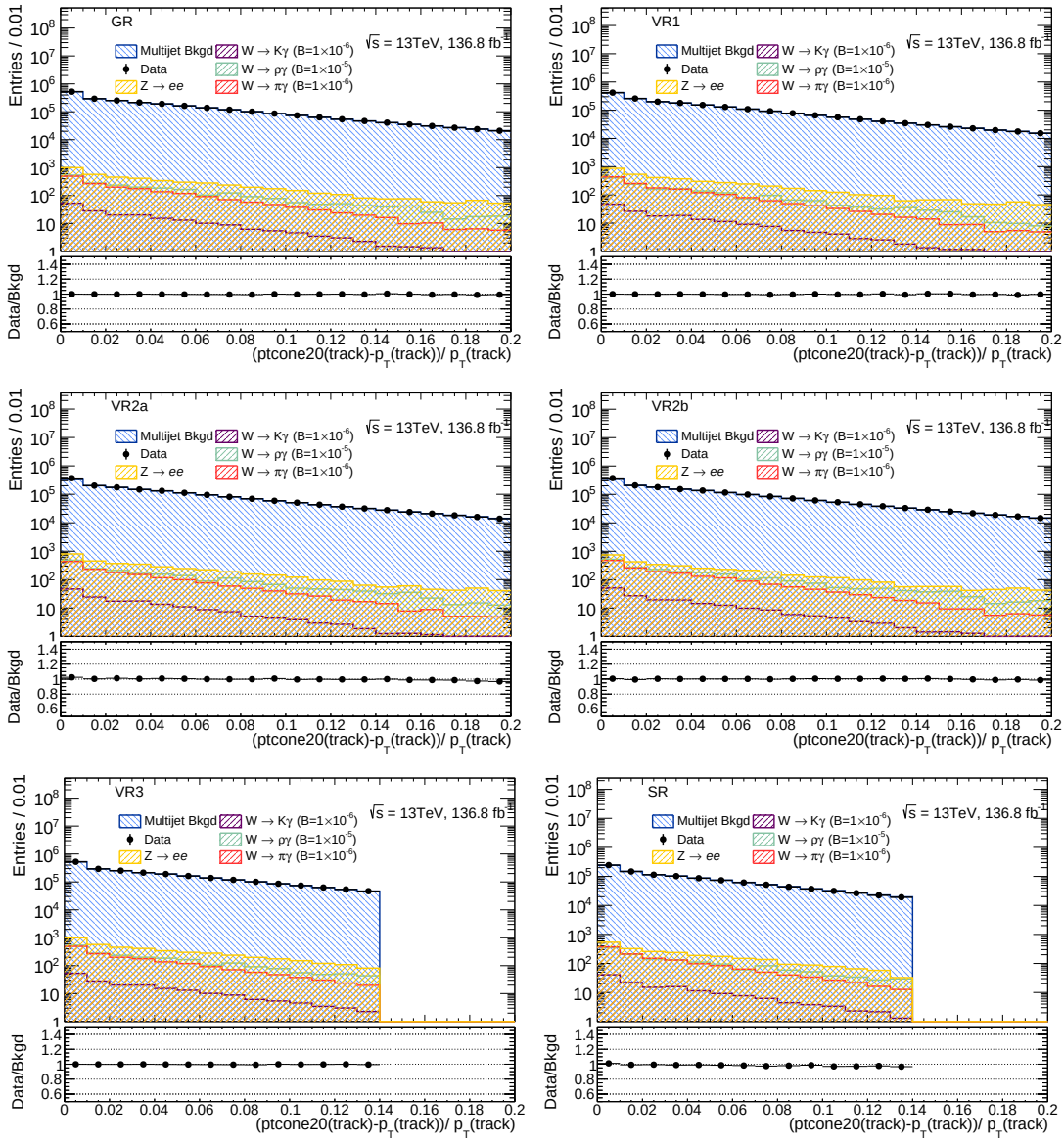


Figure D.8: Track isolation distribution for data, background and signal in the track-photon GR, VR1, VR2a, VR2b, VR3 and SR.

Appendix E

Fit Range Studies

The fit was initially performed from 0 to 300 GeV. However, it was found that the background model described the data better in a range closer to the signal window. This appendix details the results obtained when using different ranges in the track-photon final state.s

Fits were performed to Asimov datasets built from the pre-fit expected backgrounds in the different mass ranges. The results are detailed in Table E.1. The correlations between the fit parameters can be found in Figs. E.1 to E.4. One can see that the correlations between the background parameters are more sizeable for the reduced ranges. Table E.3 presents the expected limits on the $W^{\pm} \rightarrow \pi^{\pm}\gamma$ and the $W^{\pm} \rightarrow \rho^{\pm}\gamma$ channels, obtained while fixing the other signal to zero. Comparing Table E.3 with Table E.2 one can see that in the lowest range considered there is no appreciable gain in sensitivity, while for the increased fit ranges the sensivity is improved by fixing the other signal μ. This is understood by looking at Fig. E.1, where it can be seen that for the reduced range of 60-110 GeV the correlation between the two signals becomes negligible.

Fits to sideband data were also repeated for different mass ranges and can be seen in Fig. E.5. The final values and uncertainties of the fit parameters are reported in Table E.4. Upper limits at 95% CLs were derived using the post-sideband fit Asimov dataset, with the background normalisations and background shape parameters set to their post-sideband fit values. The upper limit results can be found in Table E.5 (post-fit expected limits), and can be compared to the limits obtained using pre-sideband fit

Asimov datasets (Asimovs built with background nuisance parameters fixed to their pre-fit values). From Table E.4 it can be seen that the post-fit values of the parameters vary depending on the considered range, with $\alpha(\text{tilt})$ and $\mu(Z \to e^-e^+)$ varying 3.5σ and $2\ \sigma$, respectively, from the smallest to the largest range considered. However, as seen from the results of Table E.5 the effect of the background parameter pulls on the limits is not significant - it ranges from 0.06σ to 0.22σ for $W^\pm \to \pi^\pm\gamma$ and from 0.002σ to 0.04σ for $W^\pm \to \rho^\pm\gamma$. Furthermore, it can be seen that the effect of increasing the fitting window between 60 - 110 GeV and 60 - 140 GeV is also not considerable - 0.35σ for $W^\pm \to \pi^\pm\gamma$ and 0.86σ for $W^\pm \to \rho^\pm\gamma$.

As seen from the fit results and the expected limits (Table E.2), increasing the fit range improves the sensitivity of the search, mainly due to better data-driven constraints on the inclusive background shape parameters. At the same time, by looking at the residuals in Fig. E.5, once can start seeing a slight deterioration of the background modelling near the blinded region. Therefore, it is preferable to use a reduced fitting range to utilise the area where the background model can better accommodate the data. Thus, the decision was made to use the $60 - 110$ GeV range, the most conservative choice. Furthermore, the expected improvement in the sensitivity by enlarging the fit range counteracts any potential issues arising from the background modelling, and thus no appreciable fit range systematics are expected to arise from this choice.

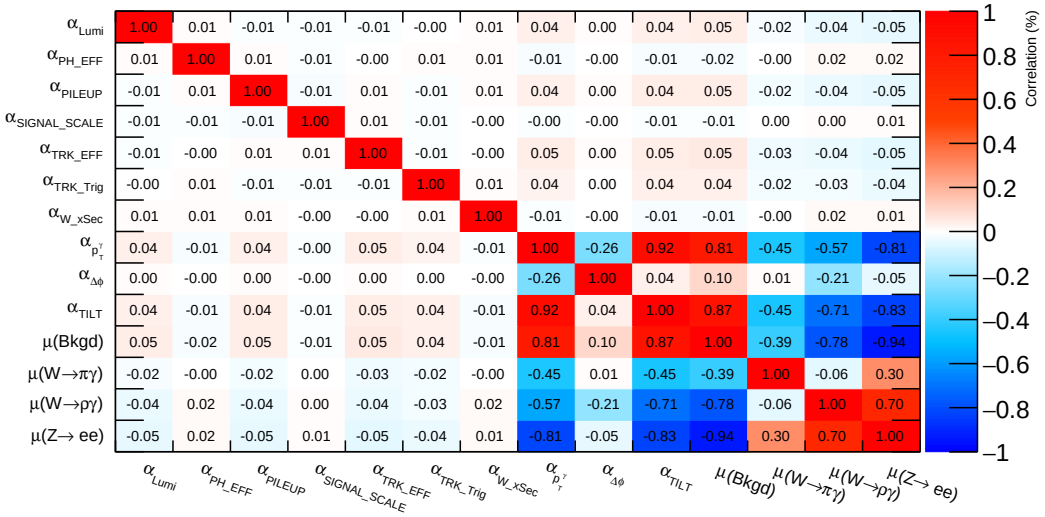


Figure E.1: Correlation matrix between parameters in fit to Asimov dataset constructed from the expected background in the track-photon final state, with a considered mass range of 60-110 GeV.

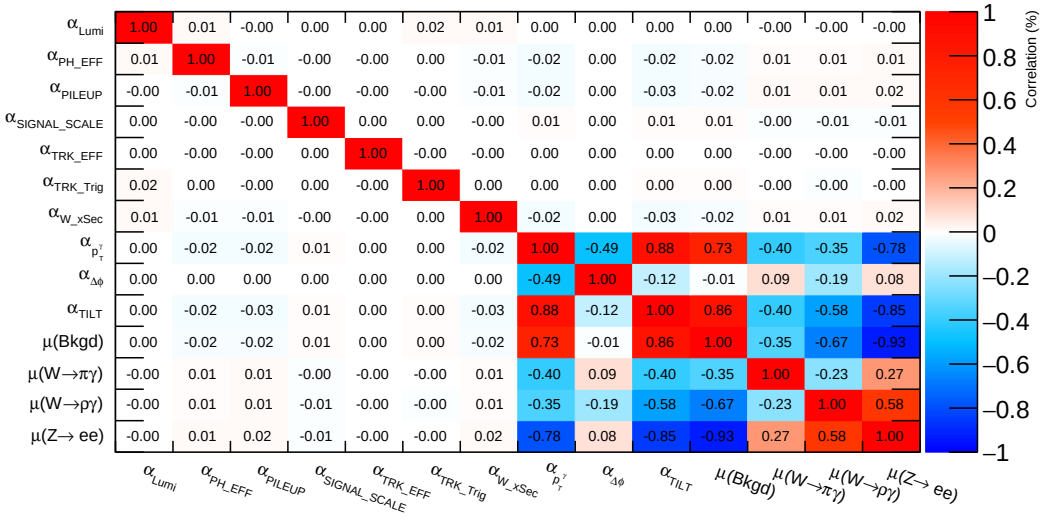


Figure E.2: Correlation matrix between parameters in fit to Asimov dataset constructed from the expected background in the track-photon final state, with a considered mass range of 60-120 GeV.

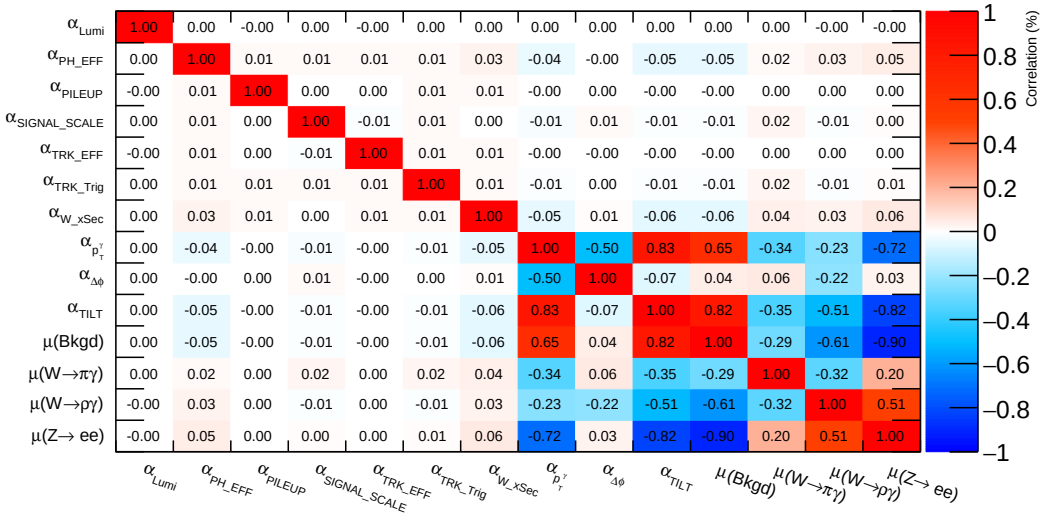


Figure E.3: Correlation matrix between parameters in fit to Asimov dataset constructed from the expected background in the track-photon final state, with a considered mass range of 60-130 GeV.

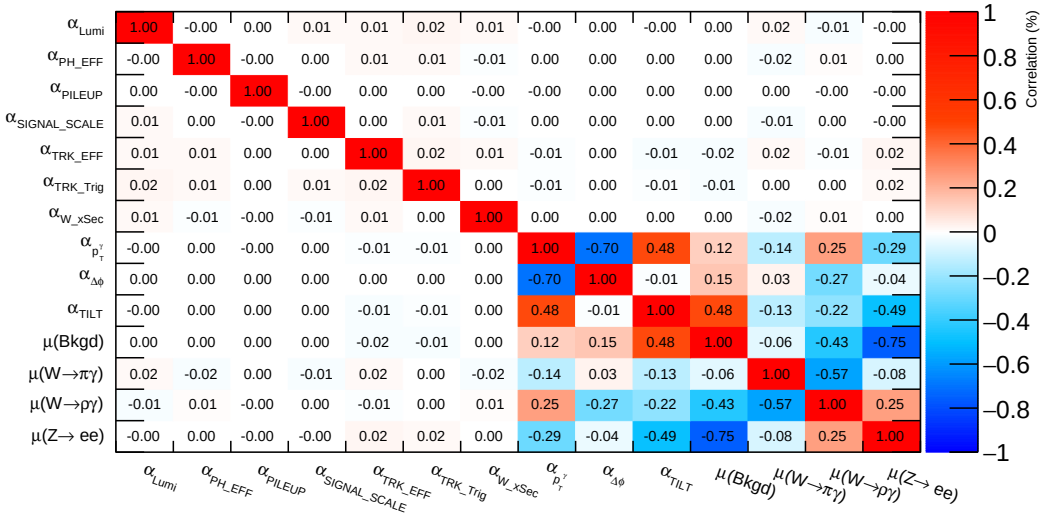


Figure E.4: Correlation matrix between parameters in fit to Asimov dataset constructed from the expected background in the track-photon final state, with a considered mass range of 60-140 GeV.

Table E.1: Values of the fit parameters after fit to Asimov built from the pre-fit expected backgrounds. The signal components are scaled to a $\mathcal{B} = 1 \times 10^{-6}$.

Parameter	60-110 GeV	60-120 GeV	60-130 GeV	60-140 GeV
$\alpha(p_{\mathrm{T}}^{\gamma})$	0.000 (+0.031/-0.242)	0.000 (+0.024/-0.178)	0.000 (+0.020/-0.139)	0.000 (+0.018/-0.122)
$\alpha(\Delta\phi)$	0.000 (+0.024/-0.014)	0.000 (+0.024/-0.014)	0.000 (+0.024/-0.014)	0.000 (+0.024/-0.014)
α(tilt)	0.002 (+0.379/-0.298)	0.001 (+0.233/-0.180)	0.000 (+0.159/-0.113)	0.000 (+0.116, -0.082)
α(luminosity)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)
α(x-section)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)
α(pileup)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)
α(theory scale)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)
$\alpha(\gamma$eff)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)
α(trk eff)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)
α(trigger)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)	0.000 (+1.000/-1.000)
μ(Background)	1.000 (+0.005/-0.005)	1.000 (+0.004/-0.004)	1.000 (+0.003/-0.003)	1.000 (+0.003/-0.003)
$\mu(Z \to e^{-}e^{+})$	0.996 (+1.060/-0.904)	0.995 (+0.895/-0.798)	0.997 (+0.755/-0.716)	0.999 (+0.682/-0.655)
$\mu(W^{\pm} \to \pi^{\pm}\gamma)$	-0.003 (+0.658/-0.626)	-0.002 (+0.608/-0.590)	-0.001 (+0.581/-0.574)	-0.000 (+0.570/-0.566)
$\mu(W^{\pm} \to \rho^{\pm}\gamma)$	0.008 (+8.893/-8.235)	0.005 (+7.724/-7.405)	0.009 (+7.109/-6.981)	0.011 (+6.846/-6.752)

Table E.2: Expected branching fraction upper limits at 95% CL, estimated using the pre-sideband fit Asimov dataset . Upper limits are estimated including different sets of uncertainties.

	60-110 GeV	60-120 GeV	60-130 GeV	60-140 GeV
$W^{\pm} \to \pi^{\pm}\gamma$ [$\times 10^{-6}$] ($W^{\pm} \to \rho^{\pm}\gamma$ profiled, $W^{\pm} \to K^{\pm}\gamma$ set to 0)				
No systematics	1.07	1.07	1.07	1.07
Shape	1.29	1.18	1.13	1.11
Norm	1.09	1.09	1.09	1.09
Shape + Norm	1.31	1.20	1.15	1.13
$W^{\pm} \to \rho^{\pm}\gamma$ [$\times 10^{-6}$] ($W \to \pi/K\gamma$ profiled)				
No systematics	11.72	11.35	11.18	11.08
Shape	17.47	15.14	13.92	13.41
Norm	11.89	11.52	11.34	11.24
Shape + Norm	17.72	15.36	14.11	13.61

Table E.3: Expected branching fraction upper limits at 95% CLs for $W^{\pm} \to \pi^{\pm}\gamma$, $W^{\pm} \to K^{\pm}\gamma$, $W^{\pm} \to \rho^{\pm}\gamma$ signal considering one process at a time (the other signal is fixed at zero). These limits were obtained using the pre-sideband fit Asimov dataset

	60-110	60-120 GeV	60-130 GeV	60-140 GeV
$W^{\pm} \to \pi^{\pm}\gamma$ only [$\times 10^{-6}$]				
Shape + Norm	1.31	1.17	1.06	1.00
$W^{\pm} \to \rho^{\pm}\gamma$ only [$\times 10^{-6}$]				
Shape+Norm	17.72	14.94	12.91	11.98

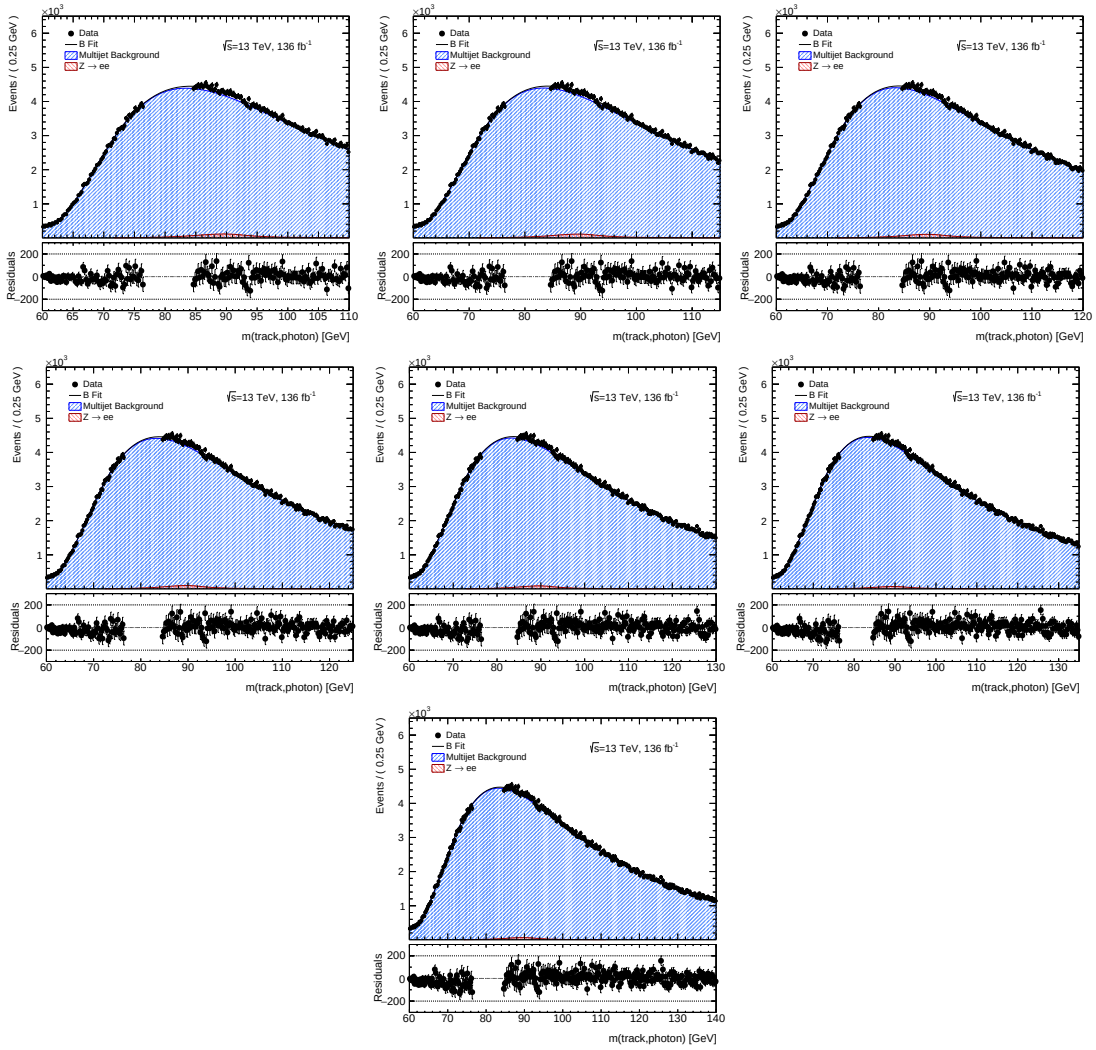


Figure E.5: Background only fit to the sideband dataset in the track-photon final state, for different ranges, from top to bottom, left to right: 60-110 GeV, 60-115 GeV, 60-120 GeV, 60-125 GeV, 60-130 GeV, 60-135 GeV and 60-140 GeV.

Table E.4: Values of the fit parameters after fit to sideband data.

Parameter	60-110 GeV	60-115 GeV	60-120 GeV	60-125 GeV	60-130 GeV	60-135 GeV	60-140 GeV
$\alpha(p_{\mathrm{T}}^{\gamma})$	0.057 (-0.021/+0.021)	0.064 (-0.020/+0.020)	0.073 (-0.018/+0.018)	0.082 (-0.017/+0.017)	0.091 (-0.016/+0.016)	0.106 (-0.015/+0.015)	0.110 (-0.015/+0.015)
$\alpha(\Delta\phi)$	-0.098 (-0.017/+0.017)	-0.098 (-0.017/+0.017)	-0.099 (-0.017/+0.017)	-0.100 (-0.017/+0.017)	-0.101 (-0.017/+0.017)	-0.103 (-0.017/+0.017)	-0.103 (-0.017/+0.017)
α(tilt)	-1.200 (-0.198/+0.197)	-1.084 (-0.174/+0.173)	-0.967 (-0.154/+0.150)	-0.855 (-0.130/ +0.129)	-0.750 (-0.113/+0.113)	-0.583 (-0.099/+0.099)	-0.542 (-0.089/0.088)
μ(Background)	0.977 (-0.003/+0.003)	0.981 (-0.003/+0.003)	0.984 (-0.003/+0.003)	0.987 (+0.003/-0.003)	0.990 (-0.003/+0.003)	0.993 (-0.003/+0.003)	0.994 (+0.002/-0.002)
$\mu(Z \to e^{-}e^{+})$	3.115 (-0.698 /+0.697)	3.020 (-0.699/+0.699)	2.843 (-0.693/0.692)	2.622 (-0.679/+0.678)	2.350 (-0.662/+0.662)	1.871 (-0.646/+0.646)	1.728 (-0.631/-0.631)

Table E.5: Expected branching fraction upper limits at 95% CL, with corresponding 1σ intervals, for different ranges considered. The limits are derived using Asimov datasets built from the expected backgrounds before and after the fit to sideband data (i.e. with the pre-fit and post-fit values of the background normalisations and background shape parameters).

	60-110 GeV	60-115 GeV	60-120 GeV	60-125 GeV	60-130 GeV	60-135 GeV	60-140 GeV
Pre-sideband Fit Expected Limits [$\times 10^{-6}$]							
$W^{\pm} \to \pi^{\pm}\gamma$	$1.31^{+0.51}_{-0.37}$	$1.25^{+0.49}_{-0.35}$	$1.20^{+0.48}_{-0.33}$	$1.17^{+0.46}_{-0.33}$	$1.15^{+0.45}_{-0.32}$	$1.14^{0.45}_{0.32}$	$1.13^{+0.44}_{-0.31}$
$W^{\pm} \to \rho^{\pm}\gamma$	$17.72^{+6.95}_{-4.95}$	$16.45^{+6.44}_{-4.60}$	$15.36^{+6.02}_{-4.29}$	$14.35^{+5.62}_{-4.01}$	$14.11^{+5.53}_{-3.94}$	$13.83^{+5.41}_{-3.87}$	$13.61^{+5.33}_{-3.80}$
Post-sideband Fit Expected Limits [$\times 10^{-6}$]							
$W^{\pm} \to \pi^{\pm}\gamma$	$1.23^{+0.48}_{-0.34}$	$1.20^{+0.46}_{-0.34}$	$1.16^{+0.46}_{-0.32}$	$1.14^{+0.45}_{-0.32}$	$1.12^{+0.45}_{-0.31}$	$1.12^{+0.44}_{-0.31}$	$1.11^{+0.44}_{-0.31}$
$W^{\pm} \to \rho^{\pm}\gamma$	$17.71^{+6.93}_{-4.95}$	$16.46^{+6.45}_{-4.60}$	$15.37^{+6.02}_{-4.30}$	$14.59^{+5.71}_{-4.08}$	$14.05^{+5.50}_{-3.93}$	$13.70^{+5.37}_{-3.83}$	$13.45^{+5.27}_{-3.76}$

Appendix F

Signal Injection Tests

F.1 Fit in track-photon final state

As an additional check, different levels of $W^{\pm} \to \pi^{\pm}\gamma$ signal were injected into the pre-sideband fit Asimov dataset of the expected backgrounds in the track-photon final state, and the fit was repeated for each configuration. The fit was found to be robust under signal injections, with the signal strength parameter reproducing the amount of signal injected in each test. Table F.1 lists the values of fitted $\mu(W^{\pm} \to \pi^{\pm}\gamma)$ after injection of different amounts of signal. Fig. F.1 shows the fit results for a signal injection of $\mu = 5$ and $\mu = 10$.

Table F.1: Fitted values of $\mu(W^{\pm} \to \pi^{\pm}\gamma)$ for different amounts of injected signal in the track-photon SR. In this case, $\mu(W^{\pm} \to \pi^{\pm}\gamma) = 1$ corresponds to a branching fraction of 1×10^{-6}.

μ_{INJ}	μ_{FIT}	σ_{μ}
1	1.000	+0.654/-0.612
2	2.000	+0.680/-0.625
3	3.000	+0.717/-0.648
4	4.000	+0.762/-0.682
5	5.000	+0.816/-0.724
6	6.000	+0.875/-0.773
7	7.000	+0.940-0.828
8	8.000	+1.009/-0.887
9	9.000	+1.081/-0.950
10	10.000	+1.156/-1.015

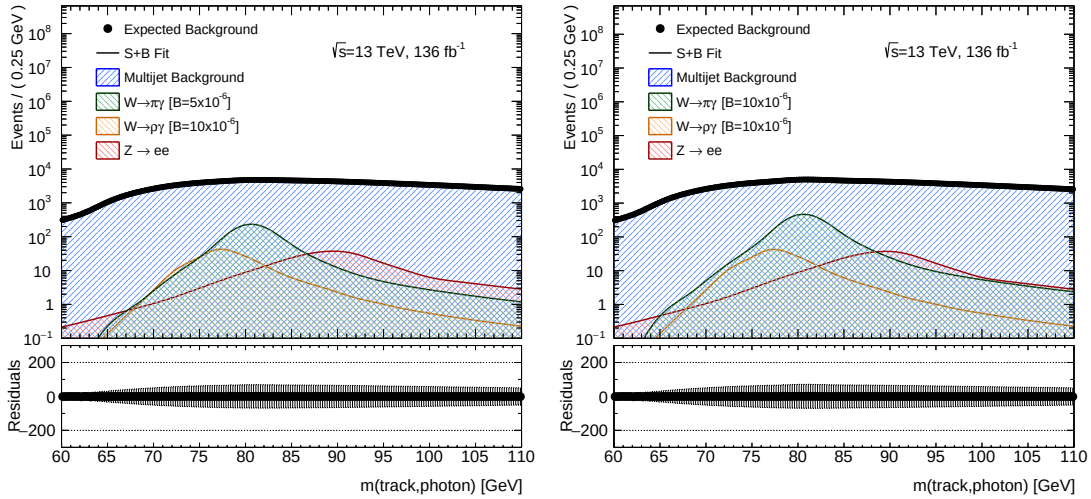


Figure F.1: Results of the fit to the pre-sideband fit Asimov dataset in the track-photon final state with signal injected equivalent to $\mu(W^{\pm} \to \pi^{\pm}\gamma) = 5$ (left) and $\mu(W^{\pm} \to \pi^{\pm}\gamma) = 10$ (right). In this case, $\mu = 1$ corresponds to a branching fraction of 1×10^{-6}.

F.2 Combined Fit

A signal corresponding to $\mu(W^{\pm} \to \rho^{\pm}\gamma) = 10$ (corresponding to $\mathcal{B} = 10 \times 10^{-6}$) was injected into the pre-sideband fit Asimov dataset for the combination between track-photon and tau-photon SRs. As for the fit including only the track-photon final state, the result reproduces the level of signal injected, with $\mu_{\text{FIT}}(W^{\pm} \to \rho^{\pm}\gamma) = 10.000$ (+3.424, -3.300). Fig. F.2 shows the result of this fit on both categories.

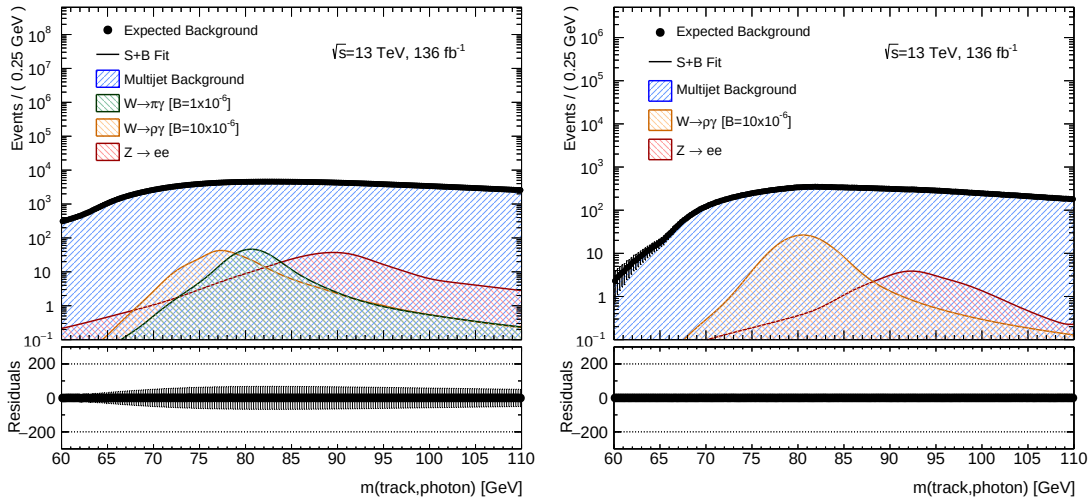


Figure F.2: Results of the fit to the pre-sideband fit Asimov dataset in the combined track-photon (left) and tau-photon final state (right) with signal injected $\mu(W^{\pm} \to \rho^{\pm}\gamma) = 10$.