CERN-THESIS-2022-165
08/04/2022

THESE DE DOCTORAT

NNT : 2022UPASP037

université PARIS-SACLAY

UNIVERSITY OF BIRMINGHAM

Search for semileptonic vector boson scattering with the ATLAS Detector using LHC Run-II data

Recherche de la diffusion de bosons vecteurs dans la désintégration semi-leptonique avec le détecteur ATLAS en utilisant la statistique du LHC Run-II

Thèse de doctorat de l'université Paris-Saclay et de l'université de Birmingham

École doctorale n°576 Particules, hadrons, énergie et noyau : instrumentation, imagerie, cosmos et simulation (PHENIICS)
Spécialité de doctorat: Physique des accélérateurs
Graduate School : Physique. Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche **IJCLab** (Université Paris-Saclay, CNRS), sous la direction de **Lydia FAYARD**, Directrice de Recherche, et la co-direction de **David CHARLTON**, Professeur, sous le co-encadrement de **Dimitris VAROUCHAS**, Chargé de Recherche, et le co-encadrement de **Konstantinos NIKOLOPOULOS**, Professeur

Thèse soutenue à Paris-Saclay, le 8 avril 2022, par

Tobias FITSCHEN

Composition du jury

Marie-Hélène SCHUNE Directrice de Recherche, IJCLab/IN2P3/CNRS	Présidente
Richard BATLEY Maître de Conférences (eq. HDR), Cavendish Laboratory, et Clayton Fellow, Christ's College, University of Cambridge	Rapporteur & Examinateur
Juan ALCARAZ MAESTRE Directeur de Recherche, CIEMAT, Madrid	Rapporteur & Examinateur
Koji NAKAMURA Maître de Conférences, Institute of Particle and Nuclear Studies (IPNS), High Energy Accelerator Research Organization (KEK)	Examinateur
Lydia FAYARD Directrice de Recherche, IJCLab/IN2P3/CNRS	Directrice de thèse

université PARIS-SACLAY

ÉCOLE DOCTORALE
Particules, hadrons, énergie et noyau : instrumentation, imagerie, cosmos et simulation (PHENIICS)

Titre: Recherche de la diffusion de bosons vecteurs dans la désintégration semi-leptonique avec le détecteur ATLAS en utilisant la statistique du LHC Run-II

Mots clés: ATLAS, LHC, Physique du Modéle Standard, diffusion de bosons vecteurs

Résumé: Les considérations sur les processus de diffusion de bosons vecteurs (VBS) violant l'unitarité étaient l'une des principales motivations du mécanisme de Higgs. Ce n'est qu'en raison de l'annulation précise de tels processus avec des interactions impliquant le boson de Higgs que l'unitarité est restaurée dans le modèle standard de la physique des particules. Cette interaction complexe fait de la mesure de la diffusion des bosons vecteurs un test particulièrement puissant pour le modèle standard et un excellent endroit pour rechercher de nouvelles contributions au-delà du modèle standard. L'analyse présentée est une recherche de la production électrofaible d'une paire de bosons vecteurs en association avec une paire de jets aux caractéristiques typiques des processus de diffusion des bosons vecteurs. La recherche est effectuée avec le détecteur ATLAS dans le canal semi-leptonique en état final du 0-lepton, et utilise la statistique complète du Run-II (énergie dans le centre de masse de 13 TeV). Dans l'analyse, des algorithmes d'étiquetage de bosons vecteurs sont utilisés pour l'identification des structures de désintégration hadronique à l'aide de variables de sous-structure de jet. Des études dédiées à de tels algorithmes sont présentées. En utilisant de nouvelles définitions de jets, des améliorations potentielles en termes d'efficacité sont présentées. Des méthodes potentielles de décorrélation de la décision d'étiquetage à partir de la masse du jet sont étudiées. La structure de désintégration caractéristique des processus de diffusion des bosons vecteurs produit des jets dans la région avant du détecteur. Le détecteur ATLAS ITk, une mise à niveau prévue pour le trajectographe interne, étend l'acceptation dans ces régions bien au-delà de la portée actuelle. Des mesures sur des capteurs prototypes pour l'ITk sont présentées. Ils comprennent des mesures géométriques, électriques et d'efficacité de reconstruction des traces dans une configuration de faisceau de test d'électrons. Enfin, les perspectives sur la manière dont des travaux supplémentaires dans ces deux domaines pourraient contribuer à améliorer la précision des futures recherches VBS sont discutées.

université PARIS-SACLAY

ÉCOLE DOCTORALE
Particules, hadrons, énergie et noyau : instrumentation, imagerie, cosmos et simulation (PHENIICS)

Title: Search for semileptonic vector boson scattering with the ATLAS Detector using LHC Run-II data
Keywords: ATLAS, LHC, Standard Model Physics, vector boson scattering

Abstract: Considerations about unitarity violating vector boson scattering (VBS) processes were one of the main motivations for the Higgs mechanism. Only due to the precise cancellation of such processes with interactions involving the Higgs boson is unitarity restored within the Standard Model of particle physics. This intricate interplay makes the measurement of vector boson scattering an especially potent test for the Standard Model and an excellent place to search for new contributions beyond it. The presented analysis is a search for the electroweak production of a pair of vector bosons in association with a pair of jets with typical characteristics of vector boson scattering processes. The search is performed in the semileptonic channel with a 0-lepton final state and uses the full Run-II dataset of the ATLAS detector recorded at a center of mass energy of 13 TeV. Within the analysis vector boson tagging algorithms are used for the identification of hadronic decay structures with the help of jet substructure variables. Dedicated studies on such algorithms are presented. Using novel jet definitions potential improvements with respect to efficiency are shown. Potential methods of decorrelating the tagger decision from jet mass are studied. The characteristic decay structure of vector boson scattering processes yields jets in the forward detector region. The ATLAS ITk detector, a planned upgrade of the inner tracking detector, will extend the acceptance in these regions far beyond the current reach. Measurements on prototype sensors for the ITk are presented. They include geometrical, electrical, as well as hit efficiency measurement in an electron test beam setup. Prospects on how the ITk and further work on vector boson tagging could help to enhance the precision of future VBS searches are discussed.

Acknowledgements

I would like to start by expressing my gratitude to Marie-Hélène Schune, Richard Batley, Juan Alcaraz Maestre, Koji Nakamura, Lydia Fayard, David Charlton, Dimitris Varouchas, and Konstantinos Nikolopoulos for accepting the role of being part of my jury. I am especially thankful for the very helpful remarks and comments I received for the manuscript and for the constructive discussions during the defense.

Of course I owe great thanks to my supervisors Dave, Lydia, Dimitris, and Kostas who guided my way through the various projects and tasks with great experience and expertise. I am aware that the special circumstances of collaboration and teaching within a pandemic often made this task much more difficult. Because of this I am especially thankful for their effort to do so in the best possible way. As a result I still felt closely connected to the scientific community and the collaborative aspect of the work even in times where we all mostly worked from home.

I would like to give special thanks to Jean-Francois Grivaz who helped relentlessly in improving the quality of the VBS analysis. I am sure that many aspects of the analysis would not be of such quality if not for his expert remarks.

I am thankful for the friendly and welcoming atmosphere that I was lucky to experience from the members of both the LAL/IJCLab and Birmingham groups. I am looking forward to seeing many of them again throughout the future. I am sure there will be many opportunities.

Throughout the collaboration within the VBS analysis, the ITk work, and the vector boson tagger studies I was always lucky to work with colleagues which made the experience enjoyable. I am thankful for each one of them.

Of course I would like to thank all of my friends. Being able to look forward to gatherings (even if often digital) was essential for getting through the special circumstances of working from home for months on end while still staying motivated and focused.

Lastly I would like to thank my parents and family who always supported me throughout my whole live and who I can always rely on. I will always be grateful for all of what they have done for me.

Contents

0 - Introduction

The Standard Model of particle physics is the most accurate theoretical description of the thus far discovered elementary particles and their interactions. It is therefore the baseline framework to which all new measurements or discoveries in the field of particle physics have to be compared to or integrated into. The discovery of a particle consistent with the Standard Model Higgs boson by the ATLAS and CMS collaborations at the Large Hadron Collider in 2012 finalised the decade-long pursuit of experimentally verifying the existence of each particle described by the Standard Model; But many questions about the way in which the Higgs boson interacts remain to be answered and studied with precision measurements. Deviations from the Standard Model predictions may hint at a way to integrate physical phenomena like dark matter or gravity into the model which are thus far separate or not compatible.

The analysis presented in this thesis aims to study the process of vector boson scattering, the scattering of W and Z bosons, the carriers of the weak force. Historically this process played an important role for the theoretical motivation of the Higgs mechanism: Calculations of the probability amplitude for the scattering of longitudinally polarised W bosons violate unitarity at high energies. This was solved by introducing additional interactions for this process including the Higgs boson. The intricate cancellation between the involved diagrams makes this process especially sensitive to possible variations in the way the Higgs boson interacts compared to the predictions of the Standard Model. The main analysis presented in this thesis focuses on the 0-lepton semileptonic process, a final state in which one of the vector bosons decays hadronically while the other decays leptonically in such a way that no charged lepton can be detected. This may either be the case if the leptons are neutrinos or charged leptons outside of the detector's acceptance. The goal of the analysis is to be the first to observe this process with the ATLAS detector and together with observations from other vector boson scattering analyses to constitute a potent test for the predictions of the Standard Model.

The full ATLAS dataset of proton-proton collisions in the LHC Run 2 amounting to an integrated luminosity of 139 fb^{-1} is utilised. A previous analysis [1] using a subset of this motivated this study hinting at the possibility of an observation when utilising the full dataset. Additionally to the increased dataset several improvements to the analysis strategy and methods were made.

An important aspect for the main analysis presented in this thesis is the identification of hadronically decaying vector bosons. For this purpose a tagger algorithm is used to perform a discrimination based on jet substructure variables. These variables describe the inner composition and structure of jets, the reconstructed detector signatures produced by hadronic decays of vector bosons, but which can also be initiated by other particles. A part of this thesis describes studies of sev-

eral jet taggers using neural networks among other techniques. It is shown that a higher classification strength can be achieved but the correlation of the resulting discriminant to the jet mass increases drastically. The utilisation of adversarial training is studied with the aim of decreasing the correlation while maintaining a high classification strength. Prospects for improvement from applying such tagger algorithms to a novel jet definition of unified flow objects (UFOs) are presented.

A precise measurement of vector boson scattering needs a large number of measurements of collisions as well as an excellent spatial resolution of decay products especially in the forward detector regions. The planned High-Luminosity Large Hadron Collider (HL-LHC) upgrade aims to improve on both of these aspects. A part of this thesis describes work on tracking sensors for the ITk, the planned upgrade to the inner tracker system of the ATLAS detector for the HL-LHC. Many different measurements where performed on prototype sensors and are presented in this thesis. This has been done within the framework of a market survey to qualify a number of vendors for the final production of the detector. This encompasses electrical and optical measurements on single sensors in a clean room environment as well as efficiency measurements on sensor modules performed during various test beam measurement periods within an electron beam line.

This document is divided into five chapters. Chapter 1 describes the Standard Model of particle physics with emphasis on the theoretical background of vector boson scattering as well as the ATLAS detector at the LHC. This serves as the common theoretical and experimental framework for the work presented in the following chapters. Chapter 2 starts with the phenomenology and basic principles of silicon tracking detectors and based on that presents measurements on prototype sensors for the ITk, a planned upgrade for the tracking detector of ATLAS. Chapter 3 begins with describing the phenomenon of hadronisation, and how it leads to the presence of jets in an event. Based on this, it presents work on several methods for vector boson taggers acting on jets. The main analysis of this thesis is presented in chapter 4: A search for semileptonic vector boson scattering in the 0-lepton channel. Finally the conclusion in chapter 5 summarises the results from each study presented in the previous three chapters and describes how the results from the ITk measurements and vector boson tagger studies may impact an eventual future reiteration of the vector boson scattering analysis presented as main focus of this thesis.

1 - Theoretical and Experimental Foundations

The Standard Model (SM) of particle physics provides a self-consistent theoretical description of all thus far experimentally discovered fundamental particles. The following sections provide an overview on its history as well as a summary of its theoretical foundation. Sec. 1.1 gives a brief overview on the experimental path that led to the particle content of the modern SM. Sec. 1.2 shows how the theoretical framework of the SM is based on the assumption of various symmetries. It goes on to describe how electroweak symmetry breaking and the Higgs mechanism allow for massive gauge bosons and fermions, how the existence of the Higgs boson emerges from this, and in which way considerations concerning vector boson scattering played a role in leading the way towards this theory. Sec. 1.3 points out several shortcomings of the Standard Model, the efforts to mend them with the help of extensions to it, and how a measurement of vector boson scattering could indicate which kind of extension may be realised in nature.

All studies presented throughout the thesis were performed with data or simulations from the ATLAS detector, or concern measurements on hardware prototypes for this detecor. Sec. 1.4 describes the Large Hadron Collider, the synchrotron accelerator at which the detector is situated. The following Sec. 1.5 describes the detector itself and its various subcomponents. Finally, Sec. 1.6 presents the planned High-Luminosity LHC upgrade. Special focus is set on the ITk which will replace the current inner detector of ATLAS, promising several improvements with respect to analyses like the one presented in this thesis.

1.1 . The Standard Model of Particle Physics

J. J. Thomson's discovery of the electron [2] at the end of the 19th century was arguably the initiation of elementary particle physics. Since then the field has grown immensely in complexity and accuracy. Fig. 1.1 shows all thus far discovered elementary particles. Their discoveries are summarised in the paragraphs below.

Even though not elementary particles, the discovery of the nucleus (and therefore the proton as the nucleus of hydrogen) by Rutherford's famous scattering experiment [3] and Chadwick's subsequent discovery of the neutron [4] opened up a way towards a theory describing matter as made up of particles. The discovery of the photon as the first fundamental force-carrier (i.e. the electromagnetic force) may be attributed to Planck's considerations on the quantisation of black body radiation [5] as well as Einstein's discovery of the photoelectric effect [6].

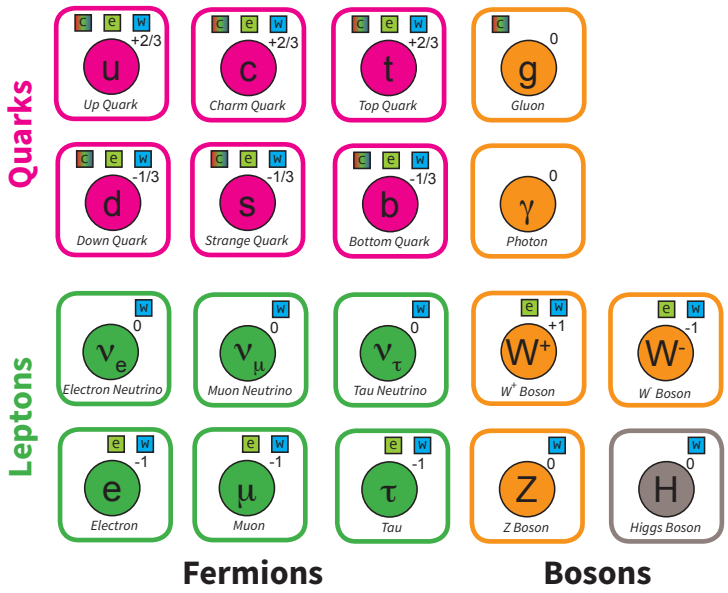


Figure 1.1: Particles of the Standard Model: Leptons (green), quarks (magenta), gauge bosons (orange) and the Higgs boson (grey). The numbers on the upper right of each particle are its electromagnetic charge. The boxes on the upper right indicate the fields to which the particle couples: colour (c), electromagnetic (e) and weak (w) (Fig. from Ref. [7]).

The Dirac equation which for the first time proposed a relativistically consistent quantised description of fermionic (i.e. spin $\frac{1}{2}$) particles suggested the existence of negative energy states which were regarded to be unphysical. Later these states were reinterpreted by Stückelberg [8] and Feynman [9] as positive-energy solutions of particles which are equal to fermions in all properties except their charge which is equal but opposite. The first candidate of these so called antiparticles, the positron (partner to the electron) was subsequently experimentally discovered by Anderson in cloud chamber experiments [10]. Antiprotons and antineutrons were later discovered at the Berkeley Bevatron by Chamberlain et al. and Cork et al. [11, 12].

The newly created theory of nuclei being made up of the electrically positively charged protons and neutral neutrons suggested the existence of an additional force acting against the electromagnetic repulsion and therefore holding the nuclei together. Yukawa proposed a quantised theoretical description of such force (now called the strong nuclear force) by introducing a new particle, the meson ('middle-weight') with a mass in-between that of the electron, and the proton and neutron [13]. The latter two were subsequently categorised as baryons ('heavy-weight') and the electron as lepton ('light-weight'). Mesons and baryons are jointly described

as hadrons. The independent measurements of cosmic radiation of Anderson and Neddermeyer [14], and Street and Stevenson [15] showed evidence for a particle in that mass range. Subsequent experiments by Powell et al. [16] identified two separate particles in cosmic radiation: The pion π which fits into Yukawa's description of the strong nuclear force and the muon μ which in later experiments was shown to be identical to the electron in all properties except mass ($m_\mu \approx 200\ m_e$). It is therefore now understood to be a lepton, like the electron, instead of a meson.

After the discovery of various new mesons and baryons, Gell-Mann proposed a unified description of them by proposing the existence of quarks as well as gluons [17]. Experimentally, the composite nature of hadrons was shown by deep-inelastic scattering experiments at SLAC [18]. Mesons are therefore combined states of two and baryons of three quarks held together by the strong nuclear force which is mediated by gluons. It was shown that the pion, initially thought to be responsible for the nuclear force, is an example of such a composite meson, while gluons are the actual fundamental force carriers. The predicted properties of gluons were experimentally verified at the DORIS and PETRA colliders at DESY [19, 20].

The six different quark flavours were discovered in various experiments. At first the physical community was hesitant to describe quarks as physical particles, thinking of them as not more than an accounting tool for conserved quantities in particle reactions. This changed in 1974 with the simultaneous discovery of the J/Ψ, a bound state of two charm quarks, by Richter and Ting in experiments at SLAC [21] and the Brookhaven National Laboratory [22]. Together with the strange quark containing kaons, which earlier have been discovered in cosmic ray experiments, this constituted the four quarks of the first two generations (columns in Fig. 1.1) of the Standard Model. The existence of a third generation of quarks was proposed by Kobayashi and Maskawa to explain the existence of charge-parity violating processes in the Standard Model [23]. The τ lepton, being a member of the third generation of leptons, was discovered shortly after [24]. The corresponding two quarks, bottom and top, were later discovered by the E288 [25] and DØ [26] and CDF [27] experiments at Fermilab.

The introduction of the neutrino as an additional particle was prompted by observations of radioactive beta decay: In this process, a neutron is converted into a proton. Electromagnetic charge conservation is not broken since a negatively charged electron is ejected which cancels out the positive charge of the proton. Due to energy conservation, the kinetic energy of the ejected electron should be fully determined by the difference in the energy levels of the nucleus before and after conversion. It should therefore be identical for each individual decay. The experimentally observed kinematic spectrum however showed a wide variation in the electrons' energies. Pauli proposed an additional neutral particle, the neutrino, to be ejected during these reactions [28]. This explained the variation in the electron spectrum since the additional energy may be carried away by the neutrino. Despite its extremely weak interaction with matter and the resulting difficulty in detecting

it, the existence of the electron neutrino was finally experimentally confirmed by Cowan and Reines at the Savannah River nuclear reactor [29]. The muon neutrino and the tau neutrino were subsequently discovered by Ledermann et al. [30] and the DONUT collaboration [31].

The W and Z bosons, the force carriers of the weak nuclear force, have been theorised by Glashow, Weinberg, and Salam [32] by unifying the weak with the electromagnetic force. They were later discovered by the UA1 and UA2 collaborations [33, 34, 35, 36]. In 2012 the discovery of a boson compatible with its properties of the previously proposed Higgs boson was announced by the ATLAS and CMS collaborations [37, 38]. The underlying mechanism is described in Sec. 1.2.2.

Nowadays the Standard Model of particle physics can be divided into a bosonic- and a fermionic sector containing all of the particles described above. The bosonic sector includes the force carrying gauge bosons (photon γ, $W^{\pm}$ and Z^0 bosons, and gluons g) and the Higgs boson H, while the fermionic sector includes all leptons and quarks.

1.2 . Theoretical Framework

Mathematically, the Standard Model is formulated as a quantum field theory. This allows for a quantised description while staying consistent with special relativity. The following sections describe some of the theoretical fundamentals of this theory. Sec. 1.2.1 starts by deriving basic properties of this model from considerations about the symmetries realised in it. Based on this, the following Sec. 1.2.2 describes the phenomenon of electroweak symmetry breaking and the role of the Higgs mechanism in the Standard Model. Finally, Sec. 1.2.3 describes vector boson scattering, the process at focus in the main analysis of this thesis. It points out how this process initially appears to break unitarity but how physical results can be recovered by the introduction of the Higgs mechanism. Vector boson scattering therefore served as one of the main motivations for the Higgs mechanism long before experimental tests for it were possible.

1.2.1 . Symmetries of the Standard Model

According to the Noether theorem [39] any generator of a continuous symmetry group leads to a conserved quantity, in the following called Noether charge. Every necessarily conserved quantity in the SM is a consequence of a symmetry that can be summarised either in the Poincaré group P or the SM gauge group $SU(3) \times SU(2) \times U(1)$. The former are called external- while the latter are called internal symmetries.

The underlying principle of the external symmetries are derived from the assumption of causality. Simply stated, this means that an event x cannot be influenced by any event y that is not causally connected, i.e. separated by a distance in space and a time interval which would not allow light (or information in general) to

reach x from y. In the framework of special relativity x and y become four-vectors and the causal connectivity is described by space-like separation: $(x-y)^2 < 0$. In quantum field theory the causality principle can be elegantly expressed in the following way: The field strengths $\Phi(x)$ and $\Phi(y)$ of a field at x and y with space-like separation commute, i.e. have no influence on each other:

$$[\Phi(x), \Phi(y)] = 0 \; \forall \; (x-y)^2 < 0. \tag{1.1}$$

It is possible to show that the Poincaré group P is the group including all Minkowski space time isometries, i.e. intervals $(x-y)^2$ are invariant under P. Hence, causality is preserved under Poincaré transformations. P has ten generators: Four translations $x_\mu \to x_\mu + a_\mu$ in space-time and six Lorentz transformations $x_\mu \to M^\nu_\mu x_\nu$. The corresponding Noether charges to the translations can be identified as energy (first component of a_μ) and momentum (the other three components of a_μ). The components of the Lorentz transformation yield the four-dimensional equivalent of the centre of mass theorem and conservation of angular momentum (including spin). Since the Poincaré group P includes all Minkowski space time isometries, i.e. is the complete group of special relativity, there cannot be any other external symmetries of the SM outside of the 10 described above corresponding to its 10 generators.

While the external symmetries follow necessarily and elegantly from fundamental physical principles just by imposing causality, the internal symmetries in the SM gauge group $SU(3) \times SU(2) \times U(1)$ were included successively to comply with experimental results. By imposing gauge symmetry of the theory, i.e. the equations of motion do not change under transformations of the SM gauge group, the conserved charges corresponding to the internal symmetries may be derived. In fact, an even stronger criterion is applied: The invariance has to hold not only for global (i.e. location-independent) transformations but for local transformations. In other words, the fields Φ may be transformed by an arbitrary extent θ at any point x in space-time:

$$\Phi(x) \to T(x;\theta)\Phi(x) \tag{1.2}$$

The SM gauge group is represented by the set of transformations $T(x;\vec{\theta})$ corresponding to its three individual Lie-groups: The unitary group of degree 1, $U(1)$ transforms with

$$T_{U(1)} = \mathrm{e}^{\frac{i\theta(x)}{2}} \tag{1.3}$$

and can be understood as a rotation around a unitary circle in the complex plane. The special unitary group of degree 2, $SU(2)$ contains transformations of the form

$$T_{SU(2)} = \mathrm{e}^{\frac{i\theta^k(x)\sigma_k}{2}} \tag{1.4}$$

where σ_k are the three Pauli matrices. The special unitary group of degree 3, $SU(3)$ contains transformations of the form

$$T_{SU(3)} = \mathrm{e}^{\frac{i\theta^k(x)\lambda_k}{2}} \tag{1.5}$$

where λ_k are the eight Gell-Mann matrices [40]. Each fundamental force of the SM can be derived from these groups. While $U(1)$ and $SU(2)$ correspond to the electromagnetic and weak force, $SU(3)$ describes the strong force.

The properties of a quantum field theory can be elegantly formulated by stating a Lagrangian density $\mathcal{L}$ from which the equations of motion of the fields ϕ_i may be derived by the action principle

$$\frac{\delta \mathcal{S}}{\delta \phi_i} = 0 \tag{1.6}$$

where the action is defined as the functional

$$\mathcal{S}[\phi_i] = \int \mathcal{L}\left(\phi_i(x),\ \frac{\partial \phi_i(x)}{\partial x^\alpha},\ x^\alpha\right) \mathrm{d}^n x. \tag{1.7}$$

Here $n = 4$ corresponds to the four space-time coordinates with index α. As alluded to in Sec. 1.1, the Dirac equation is the relativistically consistent (i.e. invariant under Poincaré transformations) equation of motion for fermionic particles. It can be derived from the Dirac Lagrangian

$$\mathcal{L}_D = i\bar{\Psi}\gamma^\mu \partial_\mu \Psi - m\bar{\Psi}\Psi. \tag{1.8}$$

Here Ψ is the spinor doublet of a fermionic field with mass m. γ^μ are the four Dirac matrices. It is easy to verify that $\mathcal{L}$ is already symmetric under global transformations of $SU(2) \times U(1)$ (Eq. 1.3 and 1.4). It may even be made invariant under local transformations of such by replacing the four-gradient ∂_μ by the covariant derivative

$$D_\mu^{SU(2)\times U(1)} = \partial_\mu + i g_W \frac{\vec{\sigma}}{2}\vec{W}_\mu + i g' \frac{Y}{2} B_\mu. \tag{1.9}$$

In this way four additional bosonic fields are introduced: B_μ and the three components of $\vec{W}_\mu$. The corresponding coupling strengths are given by g_W and g'. $\vec{\sigma}$ is a three-vector with the Pauli matrices as elements and the free parameter Y is called the weak hypercharge. D_μ is constructed in such a way that, due to its transformation properties, the additional terms exactly cancel those terms that would spoil local gauge symmetry. The newly introduced bosonic fields are related to the experimentally discovered W and Z boson, and the photon γ, i.e. the mediators of the electroweak force; how exactly will become clear when electroweak symmetry breaking and the Higgs Mechanism are introduced in Sec. 1.2.2. This unification of electromagnetic and weak force into one description of a electroweak force was first proposed in the form of the Glashow-Weinberg-Salam model [41].

The remaining part of the SM gauge group, $SU(3)$, is related to the strong nuclear force. Again the four gradient ∂_μ in the Dirac Lagrangian (Eq. 1.8) is replaced by a covariant derivative [1]

$$D_\mu^{SU(3)} = \partial_\mu - ig_s G_\mu^a T^a \tag{1.10}$$

yielding local gauge invariance under the symmetries of $SU(3)$. T^a are the eight generators of $SU(3)$. The eight newly introduced bosonic fields G^a can be related to the experimentally observed gluons. They mediate the strong nuclear force with a coupling strength of g_s.

1.2.2 . Electroweak Symmetry Breaking and the Higgs Mechanism

The Higgs mechanism, independently proposed by Brout and Englert [42], Kibble, Hagen, and Guralnik [43], and Higgs [44], together with the principle of electroweak symmetry breaking extends the Glashow-Weinberg-Salam model (as described in Sec. 1.2). It provides a way to include mass terms for gauge bosons and for fermions into the theory while maintaining the local gauge symmetry of $SU(2) \times U(1)$.

A new bosonic field is introduced in the form of a weak isospin doublet

$$\Phi = \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{bmatrix}. \tag{1.11}$$

This results in the following additional terms for the Lagrange density of the Standard Model:

$$\mathcal{L}_H = |\partial_\mu \Phi|^2 - V(\Phi) \text{ with } V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \tag{1.12}$$

$\lambda \in \mathbb{R}$ and $\mu \in \mathbb{C}$ are free parameters of the potential $V(\Phi)$. For $\mu^2 < 0$, $V(\Phi)$ has a set of infinite global minima at $(\phi^0)^2 + (\phi^+)^2 = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2}$ and a local maximum at $\phi^0 = \phi^+ = 0$. v is called the vacuum expectation value of the field Φ. In the local maximum, $V(0,0)$ is invariant under local $SU(2) \times U(1)$ symmetry transformations. Transitioning from this state into one of the global minima constitutes the breaking of this symmetry. The set of global minima corresponds to an infinite set of possible vacuum states of the field. Without loss of generality, the state

$$\Phi_v = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ v \end{bmatrix} \tag{1.13}$$

is chosen and the potential is expanded around it:

$$\Phi = \frac{1}{\sqrt{2}} \begin{bmatrix} \phi_1(x) + i\phi_2(x) \\ v + \phi_3(x) + i\phi_4(x) \end{bmatrix}. \tag{1.14}$$

[1] Here the Dirac Lagrangian only acts on quarks but not on leptons because the latter do not possess colour charge.

By applying the unitary gauge (i.e. $\phi_1 = \phi_2 = \phi_4 = 0$ and ϕ_3 renamed to h for 'Higgs field'), replacing the four-gradient ∂_μ in Eq. 1.12 by the covariant derivative of the Glashow-Weinberg-Salam theory (Eq. 1.9), and by identifying the mass eigenstates of the experimentally discovered W, Z and γ bosons as

$$W^{\pm} = \frac{1}{\sqrt{2}}(W_1 \mp iW_2) \tag{1.15}$$

$$\begin{bmatrix}\gamma \\ Z\end{bmatrix} = \begin{bmatrix}\cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W\end{bmatrix}\begin{bmatrix}B \\ W_3\end{bmatrix}, \tag{1.16}$$

Eq. 1.12 transforms into

$$\begin{aligned}\mathcal{L}_H = &\underbrace{\frac{g_W^2 v}{2}W_\mu^- W^{+\mu}h + \frac{v}{4(g_W^2+g'^2)}Z_\mu Z^\mu h + \frac{g_W^2}{4}W_\mu^- W^{+\mu}h^2 + \frac{1}{8(g_W^2+g'^2)}Z_\mu Z^\mu h^2}_{\text{couplings of } h \text{ to } W \text{ and } Z} \\ &\underbrace{-\lambda v^2 h^2 + \frac{g_W^2 v^2}{4}W_\mu^- W^{+\mu} + \frac{v^2}{8(g_W^2+g'^2)}Z_\mu Z^\mu}_{h,\ W, \text{ and } Z \text{ mass terms}} \\ &\underbrace{+\frac{1}{2}\partial_\mu h \partial^\mu h - \frac{1}{4}\lambda h^4 - \lambda v h^3}_{h \text{ self-interactions}} \\ &\underbrace{-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{W,\ Z, \text{ and } \gamma \text{ self-interactions}} \quad .\end{aligned} \tag{1.17}$$

The Weinberg angle θ_W, a free parameter of the theory, is defined by the ratio $\tan\theta_W = \frac{g'}{g_W}$ of the two coupling strengths inherent to the electroweak force. $F_{\mu\nu}$ denotes the field strength tensor of the bosonic fields (W, Z, and γ) with respect to the covariant derivative. The W , Z, and γ self-interaction term includes among others triple (VVV) and quartic ($VVVV$) interactions between vector bosons $V = W/Z$ which are of special importance for the vector boson scattering processes at focus in this thesis. They will be described further in Sec. 1.2.3. Without the introduction of the Higgs field, mass terms for the W and Z bosons would spoil the $SU(2) \times U(1)$ gauge symmetry. In the Lagrangian above however, such mass terms are recovered with the corresponding masses

$$m_W = \frac{g_W v}{2} \text{ and } m_Z = \frac{v}{2\sqrt{g_W^2 + g'^2}} \tag{1.18}$$

while the photon γ remains massless. It can be shown that in this form $\mathcal{L}_H$ is invariant under local transformations of $SU(2) \times U(1)$. The remaining terms from the W and Z mass terms that before spoiled this symmetry, are now exactly cancelled by corresponding terms involving the Higgs boson. The Higgs field h itself also acquired a mass term with $m_H = \sqrt{2\lambda v^2}$ corresponding to the Higgs

boson's mass. Fermion masses are introduced via additional Yukawa coupling terms

$$\mathcal{L}_Y = -\frac{\lambda_f}{\sqrt{2}}(v+h)(\bar{f_L}f_R + \bar{f_R}f_L) \tag{1.19}$$

between fermions f and the Higgs field h. f_L and f_R represent the left- and right-handed components of the fermionic fields. This results in fermion masses of $m_f = \frac{\lambda_f v}{2}$ where the Yukawa coupling constants λ_f are free parameters of the model.

Note that the considerations above to derive the Lagrangian of the Higgs sector $\mathcal{L}_H$ in the form given in Eq. 1.17 did not include the strong force. This can be done by extending the covariant derivative $D_\mu^{SU(2)\times U(1)}$ (Eq. 1.9) used above by the corresponding derivative for the strong sector $D_\mu^{SU(3)}$ (Eq. 1.10). Substituting the complete derivative for the Standard Model

$$D_\mu^{\mathsf{SM}} = D_\mu^{SU(3)} + D_\mu^{SU(2)\times U(1)} \tag{1.20}$$

into the Higgs Lagrangian (Eq. 1.12) in addition to the Dirac-terms for the fermions (Eq. 1.8) yields the complete Lagrangian density for the Standard Model of particle physics:

$$\begin{aligned}
\mathcal{L}_{\mathsf{SM}} = &-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} && \rightarrow \text{gauge boson interaction terms}\\
&+ i\bar{\Psi}\gamma^\mu D_\mu^{\mathsf{SM}}\Psi && \rightarrow \text{gauge boson - fermion couplings}\\
&+ \Psi_i y_{ij}\Psi_j\Phi + \text{h.c.} && \rightarrow \text{Yukawa couplings (fermion mass terms)}\\
&+ |D_\mu^{\mathsf{SM}}\Phi|^2 - V(\Phi) && \rightarrow \text{gauge boson - Higgs interactions and mass terms}
\end{aligned}$$

where $F_{\mu\nu}$ is the field strength tensor of the bosonic fields with respect to the covariant derivative D_μ^{SM} and 'h.c.' denotes the hermitian conjugate of the previous term. The last term ($|D_\mu^{\mathsf{SM}}\Phi|^2 - V(\Phi)$) in this equation can be translated into the form given above in Eq. 1.17 again to show the bosonic mass terms explicitly. Every interaction within the Standard Model can in principle be derived from this single equation by applying the action principle (Eq. 1.7), given that the values of the free parameters are known. They cannot be determined solely from theoretical considerations and have to be measured in experiments.

In total there are 19 such parameters in the Standard Model: The 9 Yukawa couplings corresponding to the massive fermions, 3 mixing angles and 1 complex phase parametrising quark flavour mixing (CKM matrix), 3 gauge couplings g (corresponding to the 3 gauge groups $SU(3)\times SU(2)\times U(1)$), a strong CP violating phase (measured to be ≈ 0), the Higgs mass m_H, and the vacuum expectation value v of the Higgs field. By extending the SM to introduce neutrino masses in the simplest possible form, 7 additional parameters are added (the 3 neutrino masses, 3 mixing angles and 1 complex phase parametrising neutrino flavour mixing (PMNS matrix)) yielding 26 total parameters.

1.2.3 . Vector Boson Scattering

The motivation for the introduction of the Higgs mechanism into the Standard Model was twofold: First, while mass terms of the form $m_V^2 V_\mu V^\mu$ in the SM Lagrangian without this mechanism would spoil the local $SU(2) \times U(1)$ symmetry, such terms are naturally introduced in Eq. 1.17 in a form that obeys the symmetry and hence yields a consistent way of introducing mass into the SM[2]. This was described in the previous section.

The second motivation concerns the concept of unitarity. A demonstration of this principle in the context of vector boson scattering can be found in Ref. [46]. The main points are summarised below.

Simply stated, unitarity means that the sum of probabilities of all final states evolving from a given initial state must be equal to one. In principle, the verification of this for a given process would involve the calculation of all possible amplitudes to infinite order in perturbative theory. This of course is not feasible. Instead, the principle of tree unitarity [47] is invoked. It states that for any given $2 \to 2$ process its amplitude with respect to increasing energy cannot grow indefinitely. It has to be asymptotically flat or decreasing. It can be shown that in the Standard Model this would be the case for scattering processes of longitudinally polarised vector bosons if interactions involving the Higgs boson were not considered.

Massless vector bosons do not have a longitudinal component. This is the case because there is no frame of reference in which a massless particle can be at rest. Hence at least one component of its momentum p_μ must be non-zero. A longitudinal polarisation vector ϵ_μ would have a non-zero value in the same component. This would violate the condition $p_\mu \epsilon^\mu = 0$ imposed by its equation of motion (in this case the Maxwell equations).

The only massive vector bosons in the Standard Model are the W and Z bosons, the force carriers of the weak force. Hence, considerations about unitarity violation in the scattering of longitudinal components only pose a problem for them, but not for photons or gluons. Massive vector bosons B^μ with mass m are described by the Proca-Lagrangian[3]

$$\mathcal{L}_{\text{Proca}} = -\frac{1}{2} F_{\mu\nu} F^{\nu\mu} + m^2 B_\nu B^\nu \tag{1.21}$$

where the field strength tensor is defined as

$$F_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{1.22}$$

Applying the action principle (Eq. 1.7) the equation of motion, the Proca equation,

[2] It was shown by t'Hooft and Veltman [45] that, additionally to conserving the symmetry, the Higgs mechanism's way of introducing masses into the SM is renormalisable. This was one additional important reason for the popularity of this theory before the eventual discovery of the Higgs boson in 2012.

[3] Note that all terms of this Lagrangian are included in Eq. 1.17 for $B = W/Z$.

is recovered:

$$\partial_\mu F^{\mu\nu} + m^2 B^\nu = 0. \tag{1.23}$$

It can be shown that this is equivalent to the Klein-Gordon equation

$$(\partial_\mu \partial^\mu - m^2) B^\nu = 0 \tag{1.24}$$

when imposing the Lorenz condition

$$\partial_\mu B^\mu = 0. \tag{1.25}$$

A plane-wave ansatz

$$B^\mu(x) = C\epsilon^\mu(p)\mathrm{e}^{-ipx} \tag{1.26}$$

with polarisation vector ϵ^μ and arbitrary constant C can be used to find a solution for the Proca equation. In the particle's rest frame the four components μ of the polarisation vector have only three degrees of freedom corresponding to its Cartesian components:

$$\begin{aligned}\epsilon_x &= (0;1,0,0)\\ \epsilon_y &= (0;0,1,0)\\ \epsilon_z &= (0;0,0,1)\end{aligned}$$

Its zeroth component is required to vanish because of the Lorenz condition (Eq. 1.25). Choosing a direction of motion of the particle in z-direction, an alternative basis for ϵ_μ can be chosen: The components

$$\begin{aligned}\epsilon_+ &= -\frac{1}{\sqrt{2}}(0;1,i,0)\\ \epsilon_- &= \frac{1}{\sqrt{2}}(0;1,-i,0)\end{aligned}$$

are constructed as linear combinations of ϵ_x and ϵ_y. Physically they can be identified as circular polarisation states. The remaining degree of freedom, the longitudinal component ϵ_L has to point into the direction of motion $\vec{p}$, in this case the z-direction, hence:

$$\epsilon_L = (a\frac{|\vec{p}|}{E}; a\frac{\vec{p}}{|\vec{p}|}) = (a\frac{p_z}{E}; 0, 0, a) \tag{1.27}$$

with an arbitrary parameter $a > 0$. The normalisation condition yields $a = \frac{E}{m}$ and hence

$$\epsilon_L = \frac{1}{m}(p_z; 0, 0, E). \tag{1.28}$$

Substituting these polarisation vectors into the plane-wave solution (Eq. 1.26) and then into the Proca equation (Eq. 1.21) shows that they constitute valid solutions.

It is evident from Eq. 1.28 that amplitudes containing the longitudinal component $\epsilon_L \propto E$ grow indefinitely as a function of energy. This is in direct violation to the tree unitarity principle. In the context of a scattering process (shown in Fig. 1.2) of longitudinally polarised W bosons (W_L) in the absence of the Higgs mechanism this leads to amplitudes of the form [48]:

$$\mathcal{A}(W_L W_L \to W_L W_L) \propto g_W^2(-s-t) \tag{1.29}$$

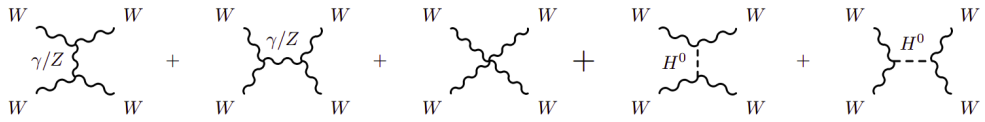


Figure 1.2: Tree-level diagrams for the process $W_L W_L \to W_L W_L$ [48].

with the Mandelstam variables s and t, where s corresponds to the centre of mass energy squared. This amplitude evidently diverges with growing centre of mass energy. By introducing the Higgs boson as another possible mediator for the $t-$ and $s-$channel interactions, therefore making the two rightmost diagrams of Fig. 1.2 possible, this amplitude becomes

$$\mathcal{A}(V_L V_L \to V_L V_L) \propto g_W^2(-s-t+\frac{s^2}{s-m_H^2}+\frac{t^2}{t-m_H^2}) \tag{1.30}$$

which cancels the divergent behaviour if the mass m_H of the Higgs boson is small. As a consequence, considerations from unitarity yielded an upper bound on the Higgs mass long before its discovery [49].

Similar considerations as shown above for $W_L W_L \to W_L W_L$ processes also hold for other vector boson scattering processes like $W_L Z_L \to W_L Z_L$ [46]. Consequently, due to the delicate nature of the mutual cancellation of individual amplitudes described above, vector boson scattering is a key process to study the Standard Model nature of electroweak symmetry breaking.

1.3 . Shortcomings of the Standard Model

While the Standard Model proved successful at describing all interactions of fundamental particles that so far have been observed within particle collider experiments, there are still physical phenomena that have been observed in other experiments but are not yet integrated into it.

Due to being a quantum field theory, the SM naturally unifies quantum mechanics with special relativity. But it has so far not been possible to unify general

relativity with it. As a consequence the SM does not include gravity. Efforts to introduce gravity into the theory in the form of an additional force carrier boson called graviton, like for the other fundamental forces, have so far failed due to the non-renormalisability of such theories. The relative weakness of gravity with respect to the other fundamental forces on energy scales probed by accelerator experiments make an experimental approach difficult.

The ΛCDM (Λ-Cold Dark Matter) model is currently the cosmological model which best matches the experimental observations. Out of the three components included in this model: dark matter, dark energy (from a cosmological constant Λ), and baryonic matter; only the latter is described by the Standard Model of particle physics. Many extensions of the SM offer possibilities of integrating the other two components by introducing additional fields with the appropriate properties. Thus far none have been verified experimentally.

Many aspects of beyond-the-Standard-Model (BSM) physics can be probed by vector boson scattering processes. An extensive overview on this topic can be found in [50]. Generally two approaches are possible when searching for BSM contributions in VBS processes: First, measurements can be made with respect to explicit BSM models of various kinds. In this scenario the likelihood of a given analysis outcome to be consistent with a given BSM model rather than the Standard Model process is derived. The other scenario takes a more generalised approach, being independent of specific models. In this approach the possible impact from BSM physics is parametrised in the form of a Standard Model Effective Field Theory (SMEFT). The Standard Model Lagrangian density is extended in a systematic manner by additional operators. Especially operators concerning anomalous quartic gauge couplings, alterations of the self-interaction vertices of four gauge bosons, are of interest in the context of VBS processes. An example for a VBS measurement setting limits on SMEFT parameters can be found in Ref. [51].

Specific models for BSM contributions to the couplings of vector boson scattering often are realised in the alteration of VBS cross-sections especially in the high-energy region. This can be understood as a hint for BSM particles within or just out of the reach of the accelerator's centre of mass capabilities. Such particles would contribute in the form of additional mediators for scattering diagrams beyond those shown in Fig. 1.2. In the SM case the Higgs boson processes interfered destructively with the gauge boson mediated processes, reducing the cross-section in the high-energy region with the consequence of preserving tree unitarity. Depending on the property of the BSM particle this can also lead to enhanced cross-sections through constructive interference. Typical examples of models containing candidate particles for this are various extensions to the Higgs sector, supersymmetric models, or extra-dimensional models like Randall-Sundrum gravity [50].

1.4 . The Large Hadron Collider

The Large Hadron Collider (LHC) [52] is a synchrotron particle accelerator which is able to provide a centre of mass energy of $\sqrt{s} = 13$ TeV in proton-proton collisions. Alternatively it is possible to operate one or both of the LHC's beams with heavy ions (e.g. lead). Fig. 1.3 shows the 26.659 km long main accelerator ring as well as its preaccelerator facilities and the four interaction points at which the four largest detector experiments, ATLAS [53], CMS [54], ALICE [55], and LHCb [56] are situated. Both ATLAS and CMS are multipurpose particle detectors designed to probe a wide range of physical processes. ALICE and LHCb are designed for a more specialised research programme with focus on quark-gluon plasma and b-physics.

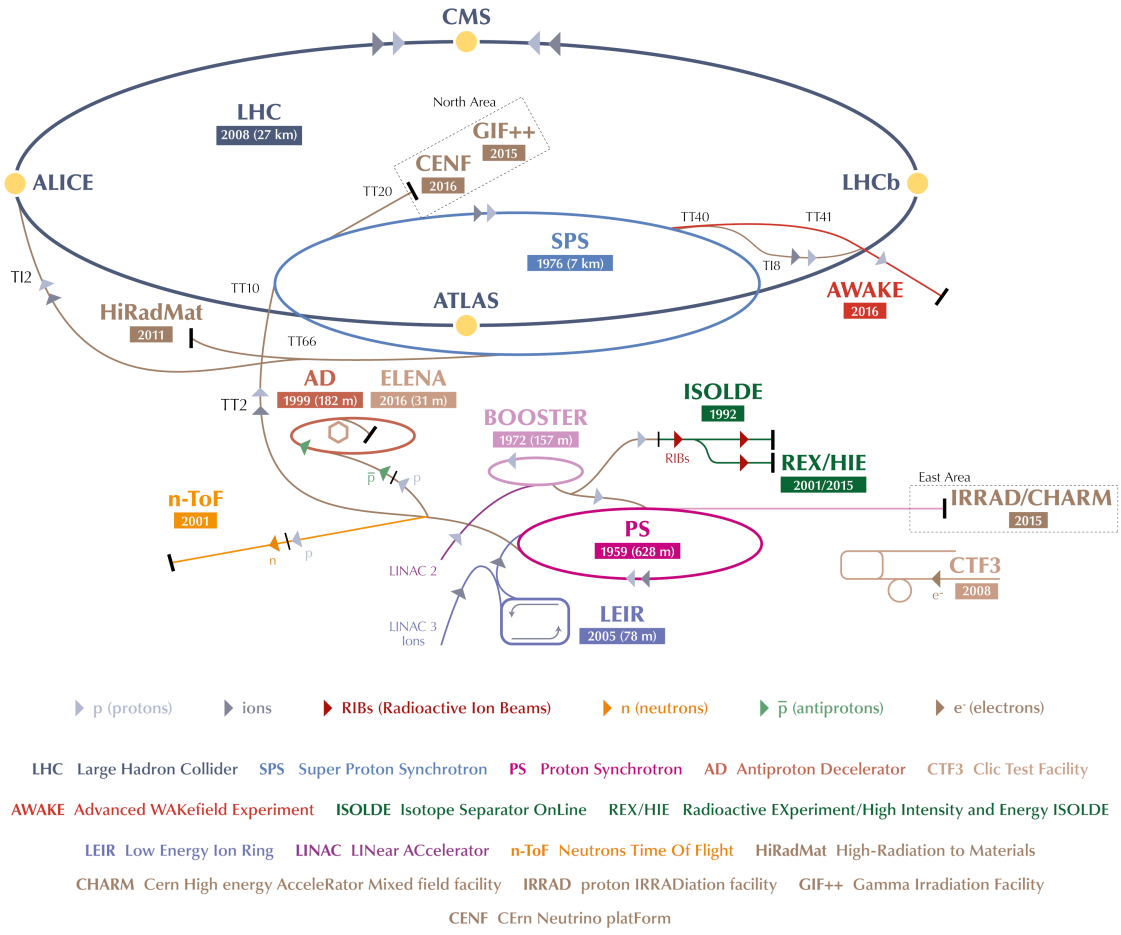


Figure 1.3: Schematic drawing of the Large Hadron Collider. The locations of several experiments within its main ring as well as its preaccelerator facilities are indicated [57].

1.5 . The ATLAS Detector

The ATLAS detector is a particle detector situated at one of the interaction points of the Large Hadron Collider (LHC). All studies presented in this thesis are made either for or with this detector. In order to study the physical signatures left by the particle collisions provided by the LHC, the ATLAS detector encompasses several specialised subdetectors the information of which can be combined during reconstruction. Their position in the detector can be seen in Fig. 1.4. A detailed description of the whole detector is given in [53]. The paragraphs below provide a brief summary.

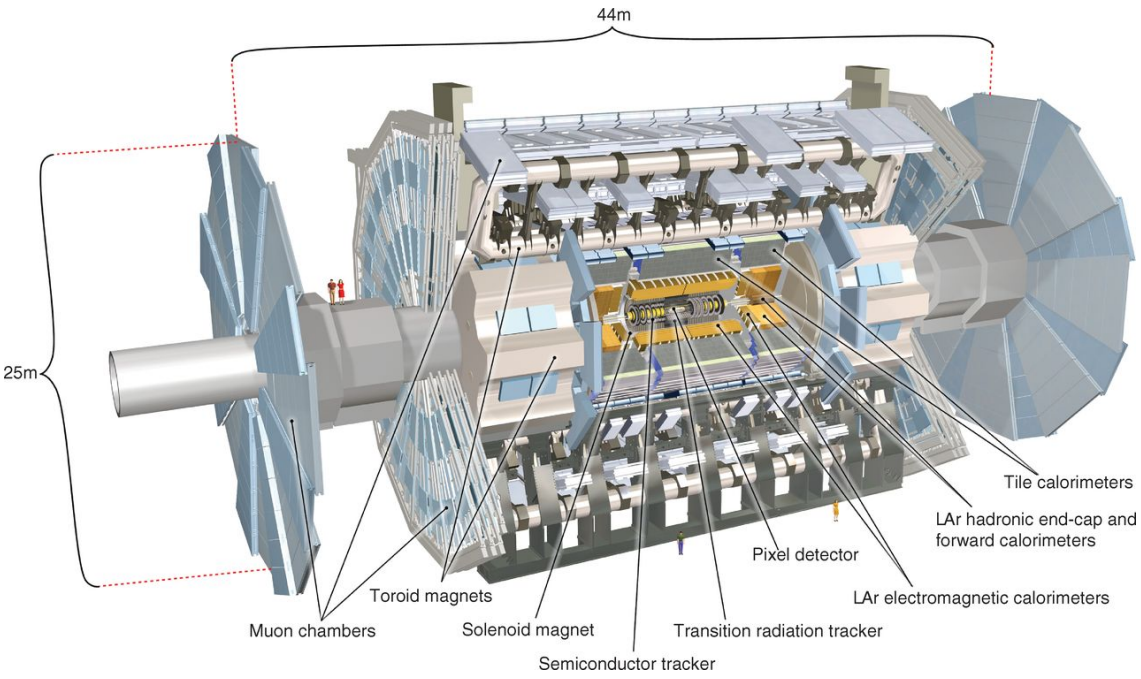


Figure 1.4: Schematic drawing of the ATLAS detector. Several subdetector components are labelled (Fig. from [58]).

The usual right-handed ATLAS coordinate system is used throughout this thesis: The z-axis is considered to point along the beam pipe with its origin at the interaction point. The perpendicular x, y-plane is referred to as the transverse plane where the positive x-direction points towards the centre of the LHC ring. The azimuthal angle Φ within the transverse plane is measured from the x-axis. The polar angle θ in the z, y-plane is measured from the z-axis. Often the pseudorapidity $\eta = -\ln \tan(\frac{\theta}{2})$ is used instead. The angular distance is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\Phi)^2}$.

The protons in the LHC beam are grouped in bunches which cross every 25 ns. This results in multiple pp interactions per event. Usually one is only interested

in one event in a given bunch (called the hard scatter event) and the effects of the other collisions (pileup) are desired to be supressed. The rapid bunch crossing interval results in an interaction rate of up to 40 MHz which makes it impossible to read out and store all information collected by the ATLAS detector. Therefore a triggering system is installed which reduces the amount of saved data by prioritising events with signatures hinting at scientifically interesting physical processes. It is divided into a hardware-based level one (L1) trigger and a software-based high level trigger (HLT). Together they reduce the rate to only a few hundred Hz.

The outermost subdetector is the Muon Spectrometer. It is designed to measure signatures from muons which usually penetrate the inner detector parts without significant loss in momentum. Three superconducting toroidal magnets bend the paths of the traversing muons which can be reconstructed in dedicated tracking chambers. The curvature induced by the magnetic field allows a calculation of the muon's transverse momenta up to a pseudorapidity of $|\eta| < 2.7$.

The muon chambers surround the calorimeter subdetector. This is divided into electronic (ECAL) and hadronic (HCAL) parts. Both measure the deposited energy of traversing and stopping particles. The former is specifically designed for the measurement of photons and electrons, the latter for the decays of hadronic particles. The outermost section is the hadronic tile calorimeter which uses plastic scintillators in combination with steel absorbers. It spans a pseudorapidity region of $|\eta| < 1.7$ with a gap for services and cabling at $1.0 < |\eta| < 1.2$. Degrading effects of the latter on the calorimeter's resolution are partially compensated by the installation of several gap scintillators in that region. The innermost section of the calorimeter is given by the Liquid-Argon Calorimeter (LAr). It can be subdivided into a central barrel region as well as two endcap regions. In the forward region ($1.5 < |\eta| < 3.2$) of the LAr the hadronic endcap calorimeters are installed. The rest of the LAr is devoted for electromagnetic calorimetry (i.e. the ECAL). It consists of a barrel ($|\eta| < 1.475$), endcap ($1.375 < |\eta| < 3.2$), and a forward region ($3.1 < |\eta| < 4.9$). Additional scintillators are installed in the transition region ($1.2 < |\eta| < 1.6$) in between barrel and endcap.

The innermost part of the ATLAS detector is called the Inner Detector (ID). This tracking system is immersed in a 2 T magnetic field. As in the muon-spectrometers, the resulting curvature of charged particles' paths is used to measure their momenta. The ID consists of a barrel and endcap and extends up to $\eta < 2.5$. Different tracking technologies are used depending on the distance to the interaction point. The innermost layer of the ID consists of silicon pixel detectors allowing a high resolution. They are surrounded by silicon microstrip detectors which in turn are surrounded by straw-tube transition radiation chambers. The latter two technologies allow to cover the whole area of the ID while being considerably less expensive than silicon pixels.

1.6 . The ITk at the High Luminosity LHC

The High-Luminosity LHC (HL-LHC) [59] is a planned upgrade for the Large Hadron Collider. The time schedule for this upgrade can be seen in Fig. 1.5.

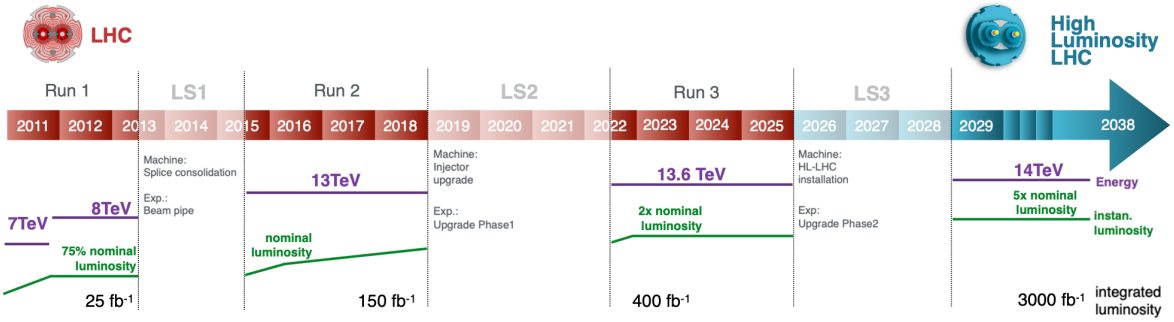
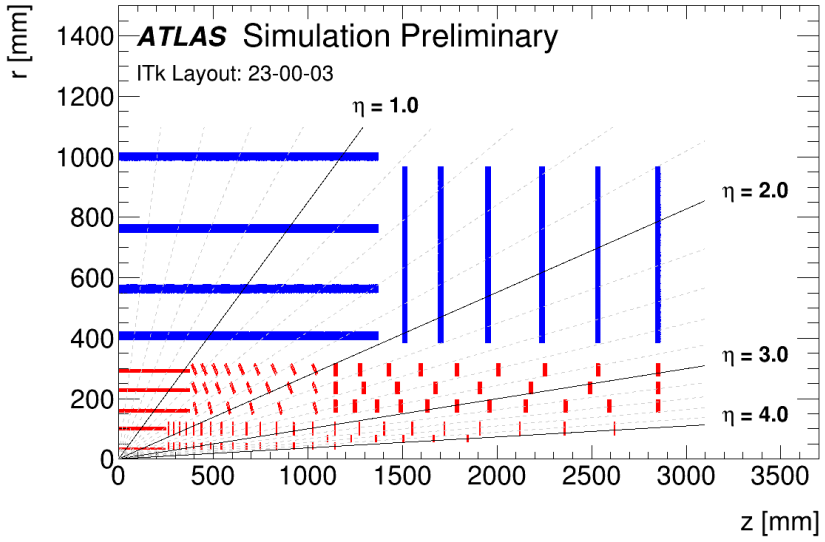


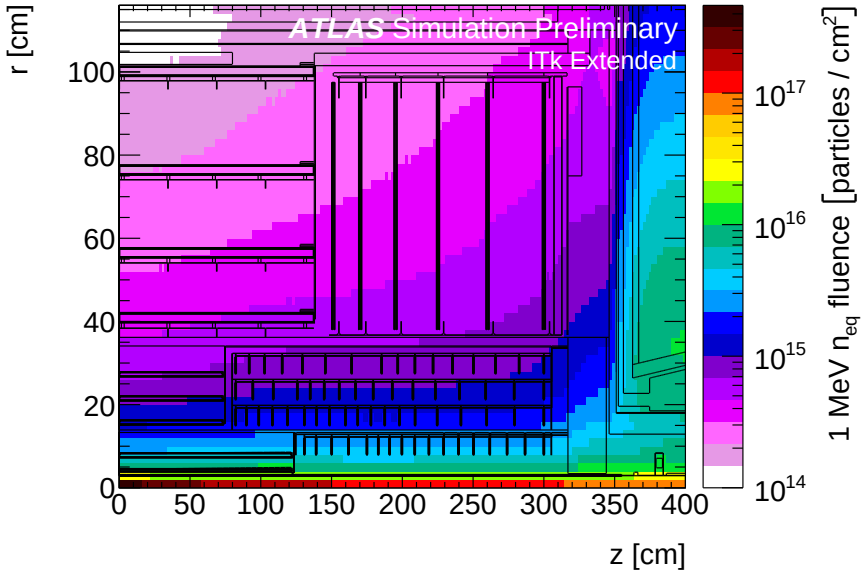
Figure 1.5: Time schedule for LHC Run1-5 including the HL-LHC period [60].

First data is expected to be taken in 2029. It is anticipated to increase the LHC's instantaneous luminosity by a factor of around 3.5. The resulting higher number of collisions per bunch crossing is expected to improve sensitivities of physics analyses which at the moment are limited by statistical uncertainties. On the other hand this increase poses hard constraints on the radiation hardness, especially of the innermost layers of the ATLAS detector. A simulation of the expected radiation dose can be seen in Fig. 1.6(b).

Because of this, the ID will be replaced by the ITk which is a tracking system completely consisting of radiation-hard silicon pixels and strips (see Fig. 1.6(a)). It is planned to replace the two innermost layers, that suffer most of the irradiation damage, once during the HL-LHC runtime, which will be after 5 years of continuous operation. A detailed description of the ITk's properties is given in the separate Technical Design Reports (TDR) of the pixel- and strip detectors [61, 62]. An active area of 13 m^2 in the form of 5 layers around the beam pipe will be instrumented with pixel sensors with 5 billion readout channels. The strip detector surrounding the pixel detector will cover an active area of 160 m^2 with 500 million readout channels in 4 layers and 6 end-caps on both sides. The ITk is designed to provide tracking information up to $|\eta| < 4$ far beyond the current ID's acceptance of $|\eta| < 2.7$. This is especially needed because of the enhanced importance of track-based pileup rejection due to the increased luminosity of the HL-LHC. Analyses, such as the vector boson scattering analysis presented in this thesis, that rely on decay structures in the forward detector region, may especially profit from this upgrade.



(a)



(b)

Figure 1.6: (a): Physical positions of the pixel (red) and strip (blue) sensors in the ITk detector [63]. (b): 1 MeV neutron equivalent fluence in the ITk region from a FLUKA 2011.2c.4 simulation of the Phase-2 Extended@4 Step 1.5 ITk layout normalised to 3000 fb^{-1} of pp collisions at 14 TeV corresponding to the expected integrated luminosity of the HL-LHC [64]. Note that the layout of the ITk has been revised in the meantime. The current layout can be seen in (a).

The innermost layer of the ITk uses 3D pixel sensors which, while being more expensive than planar sensors, are especially resistant to high radiation doses. The sensors closest to the interaction point (34 mm from the beam) are expected to be exposed to a radiation dose of 2×10^{16} $\frac{\mathrm{n_{eq}}}{\mathrm{cm}^2}$ [65]. The 3D sensors have two different layouts: 50×50 μm^2 and 25×100 μm^2 and an active thickness of 150 μm on top of an inactive support wafer of 100 μm. Each pixel has a single collection electrode in its centre consisting of a 3D column etched into the active material. The single layer of 3D sensors is surrounded by four layers of planar pixel sensors. They have a pixel size of 50×50 μm^2 with an active thickness of 100 μm in the innermost and 150 μm in the outer layers.

The final production process for both planar and 3D sensors is expected to start within 2022. Both passed the final design review and are at the time of writing in a pre-production phase in which the first 10% of sensors are made. As a final step before the full production can commence, a production readiness report is foreseen based on the sensors produced during this period. Due to the large number of planar sensors that have to be produced and the resultingly high financial risk for the project, special care was taken to qualify potential producers even before the pre-production. For this reason, an extensive market survey has been performed on prototype sensors. The vendors were chosen based on the results. The following Sec. 2 presents measurements performed for this survey.

2 - ITk Planar Pixel Sensors

This chapter presents work on the ITk Planar Pixel Market Survey, an effort to qualify a number of vendors for the final production of sensors for the ITk detector.

Sec. 2.1 describes the phenomenology of silicon detectors in general. The band structure model, the concept of doping as well as p-n junctions are introduced in Sec. 2.1.1. Based on this, usual characteristics of silicon detectors like depletion voltage and leakage current are introduced in Sec. 2.1.2. Finally, Sec. 2.1.3 describes the effects of long-term radiation exposure on silicon detectors.

Sec. 2.2 presents the results of various measurements that were performed in the frame of the Market Survey. First, the sensors underwent a visual inspection followed by geometrical measurements of thickness and bow as described in Sec. 2.2.1. Measurements of depletion voltage and leakage current are presented in Sec. 2.2.2. Sec. 2.2.3 presents hit efficiency measurements performed in a test beam setup. Finally, Sec. 2.2.4 summarises the results of the Market Survey and gives an outlook on the next steps for the ITk production.

2.1 . Silicon Detectors

Since the 1960s silicon has been used for the construction of particle detectors [66]. The electric conductivity of silicon can be explained by its properties as a semiconductor: Electrons bound in the crystal lattice can be excited by inducing an amount of energy above a certain threshold. In the following, this threshold will be called gap energy and the freed up place of energy levels in the crystal lattice will be called hole. In silicon, for every 3.6 eV of energy one electron-hole pair is created. This low threshold makes silicon a promising candidate for particle detectors because it leads to a narrow energy resolution. Electron-hole pair creation can be induced by various means, including the temperature of the material itself (leading to noise in the detector). The effect utilised in particle detectors is due to the energy introduced by traversing charged particles. Additionally to this property, silicon is widely used in particle experiments due to its high abundance in the Earth's crust, as well as the maturity of silicon-related production processes in industry. The following provides a brief summary of the main concepts of silicon sensors which are needed for the measurements performed during the Market Survey. For a more detailed treatment the reader is referred to a wide variety of introductory books on the topic, e.g. Ref. [67].

2.1.1 . Band Structure and Doping

A commonly used model to explain the electrical properties of semiconductors like silicon is the so called band gap: The structure of the energy levels is divided into three regions (from lowest to higher energy): (1) The valence band which

is completely filled in the ground state, (2) the gap region which is energetically forbidden, and (3) the conductivity band which initially is empty. Excited electrons in the semiconductor go up into the conductivity region where they can traverse the material as free electrons while leaving a hole (in the energy level structure) in the valence band.

By introducing impurities (i.e. atoms of other elements than silicon) into the crystal lattice, it is possible to add additional energy levels into the band structure according to need. This process is called doping and the introduced element is called dopant. In its pure state, the number of electrons n or holes p acting as charge carriers is in equilibrium: $n = p$. Since a silicon atom has four valence electrons, doping it with an atom with five valence electrons (a common choice is phosphorus), an excess of electrons as majority charge carriers can be achieved: $n > p$. The resulting material is called n-type silicon. A dopant with three valence electrons (e.g. boron) on the other hand makes the holes the majority charge carriers: $p > n$. In this way p-type silicon is created.

In a pure silicon substrate at room temperature the number of free charge carriers is several orders of magnitude larger than the typical number of additional charge carriers produced by interacting particles. In this state it would not be possible to separate a potential signal from the vast background of thermal charge carriers. In a large scale silicon detector experiment it is not practically feasible to cool the substrate down to temperatures at which this effect would be negligible. Instead, the number of free charge carriers is reduced by introducing a reverse biased pn-junction: A contact is made between n-type and p-type silicon while applying a sufficiently high voltage V_{bias} with higher potential (positive) on the side of the n-type. Electrons (negative charge carriers) are moving from the n-region into the p-region while holes (positive charge carriers) are moving in the opposite direction until an equilibrium is reached.

As a result, the substrate gets depleted of charge carriers in a region around the junction. Charges created by particles traversing the non-depleted zone recombine with the free charge carriers in the substrate and are therefore lost. Hence, for a detector it is desirable to increase the depleted region as much as possible. The extent of this region can be increased by raising the bias voltage V_{bias}. The voltage necessary to deplete the complete volume of the substrate of charge carriers is called depletion voltage V_{dep}. Usually, detectors are operated at a bias voltage above V_{dep}.

2.1.2 . Depletion Voltage, Leakage Current, and Breakdown

Regarding the two edges of the depletion region as capacitor plates, and the silicon within as dielectricum, the capacitance of the detector can be measured. The depletion voltage of a given silicon detector can be determined by measuring

the capacitance C as a function of the applied bias voltage V_{bias}:

$$C = \begin{cases} \sqrt{A\frac{\epsilon_0\epsilon_{\text{Si}}}{2\rho\mu V_{\text{bias}}}} & V_{\text{bias}} \leq V_{\text{dep}} \\ A\frac{\epsilon_0\epsilon_{\text{Si}}}{D} = \text{const} & V_{\text{bias}} > V_{\text{dep}} \end{cases} \tag{2.1}$$

Here ϵ_0 and ϵ_{Si} are the vacuum permittivity and the relative permittivity of silicon, ρ is the specific resistivity of the bulk, μ is the mobility of the majority charge carrier (here: electrons), D is the thickness and A the surface area of the sensor [66]. Due to the inverse proportionality $C \propto V_{\text{bias}}^{-\frac{1}{2}}$ of Eq. 2.1 for $V_{\text{bias}} \leq V_{\text{dep}}$, often $\frac{1}{C}$ is graphed as a function of V_{bias}. The resulting curve increases with V_{bias} until the complete volume of the substrate is fully depleted ($V_{\text{bias}} > V_{\text{dep}}$). Beyond this depletion voltage, the capacitance will stay constant.

At non-zero temperatures, electron hole pairs are constantly created within the depletion region. Electrons tend to drift towards the p- while holes tend to drift towards the n-regions where they recombine. The resulting leakage current is an undesired effect and can be described by

$$I_{\text{L}} \propto \frac{e\omega n_i}{2\tau_{\text{L}}}, \tag{2.2}$$

where e is the elementary charge of the electron, ω is the thickness of the depletion region, n_i is the concentration of charge carriers in the material, τ_{L} is the average generation life time of charge carriers, and A is the surface area of the sensor. The temperature T dependence of the leakage current enters via the relation

$$I_{\text{L}} \propto n_i \propto T^{\frac{3}{2}} \text{e}^{-\frac{E_g}{2k_{\text{B}}T}}, \tag{2.3}$$

where k_{B} is the Boltzmann constant and $E_g = 1.11$ eV is the energy gap of Silicon at room temperature [66]. A decrease in leakage current of a given sensor can therefore be achieved by cooling it down. When increasing the bias voltage, the leakage current increases linearly with the thickness of the depletion region [66]

$$\omega = \sqrt{2\epsilon_0\epsilon_{\text{Si}}\mu\rho V_{\text{bias}}} \propto \sqrt{V_{\text{bias}}} \tag{2.4}$$

until full depletion is achieved, i.e. ω approaches the thickness of the substrate.

At even higher voltages an electrical breakdown can occur. In this case the measured current increases dramatically. The two main reasons for this behaviour are avalanche breakdown and the Zener effect (see e.g. [67]).

2.1.3 . Effects of Radiation

Both, the leakage current and the depletion voltage of a given sensor can be affected by defects introduced by exposure to particle radiation. The detection mechanism in silicon detectors makes use of the ionising effect of such particles, i.e. their interaction with the outer-shell electrons of the silicon atoms. Less common interactions with the atomic nuclei can produce defects in the lattice

structure of the substrate which may impair its functionality. The displacement of a Si atom may lead to vacancies (atoms missing in the grid) or interstitials (atoms at irregular positions in the grid). Additionally impurities by atoms other than Si may be introduced.

As a result, new energy levels are introduced. Their effect on the detectors characteristics is mainly determined by their position in the band structure. Mid-gap defects increase the probability of electron-hole creation within the depletion zone. According to Eq. 2.2 the resultingly decreased average generation life time τ_{L} leads to increased leakage currents. When operating silicon sensors in a detector the temperature has to be decreased accordingly in order to lower the leakage current and the associate power drain. Shallow levels in-between the mid-gap region and the valence or conduction bands may function as charge traps for electrons or holes. As a result, the charge collection efficiency of the detector degrades. Levels close to the valence or conduction band create charge defects which may shift the depletion voltage.

At non-zero temperatures, short-range structures in the lattice are mobile. As a result, some defects may either fully heal or develop into less severe defects over time. This process is called annealing. The simplest example of this is Frenkel-pair recombination ($V + I \rightarrow Si$) in which a vacancy V is filled up by an interstitial I. Other examples include the combination of two vacancies or interstitials into multi-vacancies or -interstitials ($V + V \rightarrow V_2$ or $I + I \rightarrow I_2$) or more complex combination of impurities.

2.2 . ITk Planar Pixel Market Survey Measurements

The ITk Planar Pixel Market Survey (MS) was a procedure to qualify vendors for the production of the planar pixel sensors of the ITk. For this purpose, prototypes of sensors from several companies were produced. They were tested with respect to several physical and technical properties the requirements of which are documented in the Technical Specification Report [68]. The identity of the competing vendors remains confidential. Hence, only examples of results are shown in the following without explicitly mentioning the vendors which produced the corresponding modules. A small subset of the measurement results shown below were later publicised in Ref. [69].

Fig. 2.1 summarises the measurement steps for the MS. The individual steps are described in the sections below. Some of the required measurements are repeated after irradiating the modules to fluences of 2 and $5 \times 10^{15} \frac{n_{\mathsf{eq}}}{\mathsf{cm}^2}$ [1]. This corresponds roughly to the expected fluences (see Fig. 1.6(b)) the innermost planar sensors would encounter inside of the ITk after the HL-LHC's full run cor-

[1] n_{eq} stands for neutron equivalents. The irradiations were performed at the University of Birmingham (UK) and the CYRIC irradiation facility (Japan) with proton beams and the corresponding dose converted into n_{eq} was calculated.

responding to an integrated luminosity of 3000 fb^{-1}. Except for the test beam hit efficiency measurement presented in Sec. 2.2.3, all results presented below were obtained with non-irradiated sensors. The corresponding tests on irradiated sensors were performed by other institutes.

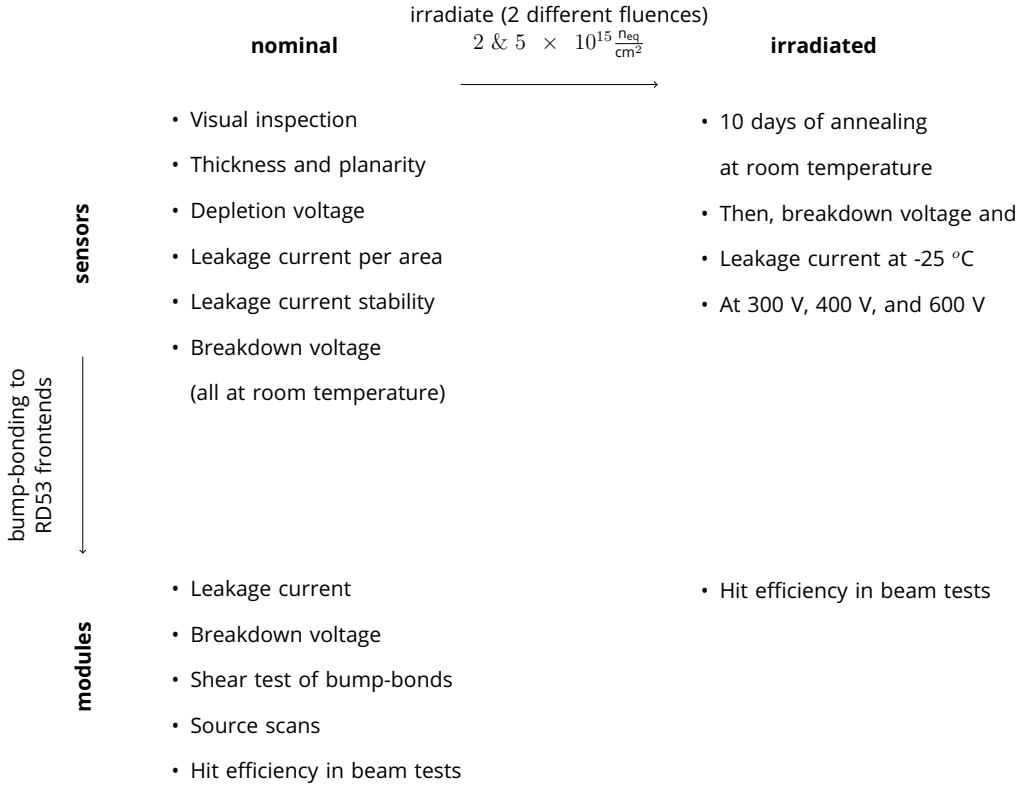
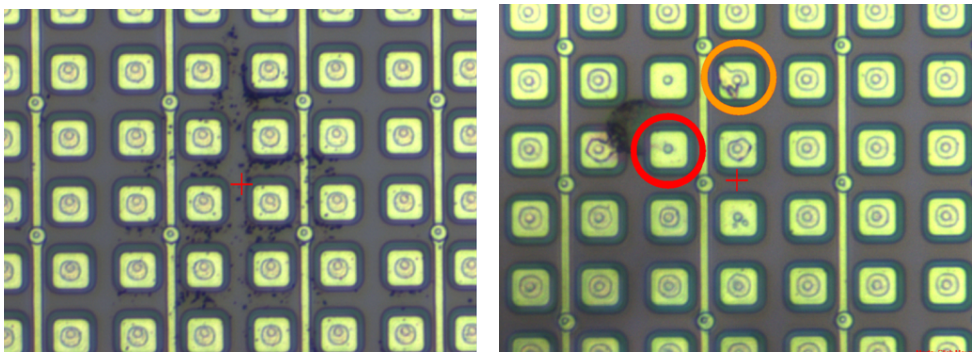


Figure 2.1: Schematic overview of the various measurement steps of the ITk Planar Pixel Market Survey. Measurements on bare sensors are indicated on the top of the diagram. Measurements performed on modules, i.e. sensors flip-chipped onto RD53a [70] front end readout chips are shown on the bottom. Some of the measurements have to be performed after irradiating the devices in order to make predictions on their radiation hardness. Such measurements are shown on the right side of the diagram.

2.2.1 . Visual Inspection and Geometrical Measurements

The first action performed after receiving new sensors or wafers at one of the ITk institutes is a visual inspection. For this purpose, a microscope is used to scan manually over each sensor. It is required that no visible scratches or residues from the production are visible (see Fig. 2.2(a)). In particular, the implants and the aluminium layer must not create any visible shorts between individual pixels. In

the final product, a reliable electrical contact between the silicon of the sensor and the solder bumps (added in a later process) will be ensured by pads of under-bump metallisation (UBM) material on each individual pixel. For the Market Survey it is required that not more than 0.1 % of these pads show defects. An example of such defects can be seen in Fig. 2.2(b).



(a) (b)

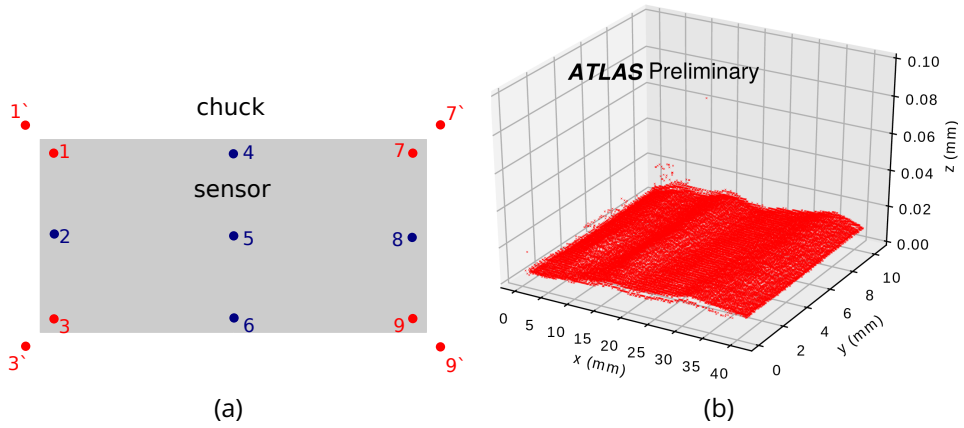
Figure 2.2: Examples of irregularities encountered during the visual inspection of bare sensors for the ITk Planar Pixel Market Survey. Fig. 2.2(a) shows grainy residue (black dots) while Fig. 2.2(b) shows some deformed (orange) or missing (red) UBM pads.

At the time of the Market Survey the final geometry of the planar pixel sensors was not finalised yet. Additionally to the 50 μm $\times$ 50 μm layout, that later on was chosen for the final production of the ITk, sensors with dimensions of 25 μm $\times$ 100 μm were tested. The pictures here show a sensor with the 50 μm $\times$ 50 μm layout. In addition to the two separate pixel layouts, sensors with two different thicknesses were tested: 100 μm and 150 μm. The former are foreseen to be installed into layer 1 while the latter are installed in all other layers containing planar sensors. 3D pixel sensors, which are not part of this but a separate Market Survey, are foreseen for layer 0 and will have a geometry of 25 μm $\times$ 100 μm in the barrel and 50 μm $\times$ 50 μm in the end-cap regions.

Parallel rails can be seen in between columns of pixels. These rails serve the purpose of electrically biasing the whole sensor area for electrical measurements that are described further below. The circular structures on these rails that can be seen in regular distances of 100 μm are the punch through bias dots. They are used to apply an electric potential similar to the one that the front-end chip would apply to each individual pixel after front-end and sensor are bonded together. It is known that this type of biasing structure introduces small inefficiencies in the region of the bias dots. This will be investigated in detail in Sec. 2.2.3 where measurements of the hit efficiency of such sensors are presented. Depending on the vendor alternative types of biasing structure (namely poly-silicon, or temporary

metal) may be used, avoiding this source of inefficiency.

Geometrical measurements are necessary to ensure that the sensors geometrical properties are suitable for the further production steps towards the final modules. Especially the bump-bonding and subsequent flip-chipping process, connecting the sensor itself to the front-end chip for readout, require a uniform thickness and minimal bow. The limits defined for the Market Survey require a deviation of < 15 μm of the measured thickness with respect to the desired value (either 100 or 150 μm) plus 50 μm additional handle wafer thickness. The measured bow of each sensor is required to be < 25 μm. The measurements are performed with the microscope-focus method: A microscope at height h_i is calibrated to be in focus on a point i on the sensor (see Fig. 2.3(a)).



(a) (b)

Figure 2.3: (a): Measurement points for the microscope-focus method used to determine thickness and bow of sensors produced for the ITk Planar Pixel Market Survey. The thickness is calculated from the measured heights of four points on and four points next to the sensor (marked in red). The nine points on the sensor are used to determine the bow of the sensor. (b): 3D profile of a sensor produced for the ITk Planar Pixel Market Survey obtained by a laser scan performed by the Oxford working group [69]. The results from the microscope-focus method have been compared with this alternative method.

Subsequently, the microscope is moved horizontally to the point i' outside of the sensor. The microscope height is altered until at $h_{i'}$ the surface of the chuck is in focus. Each of these measurements is performed several times to estimate the systematic uncertainty which is dominated by the ability of the experimenter to judge whether the optimal focus is achieved. The thickness T can be calculated by taking the average over the height differences at various points:

$$T = \langle h_i - h_{i'} \rangle_{i \in \{1,3,7,9\}}, \tag{2.5}$$

while the bow B is defined as

$$B = \frac{h_4 + h_5 + h_6}{3} - \frac{h_1 + h_2 + h_3 + h_7 + h_8 + h_9}{6}, \tag{2.6}$$

only using points on the sensor. Tab. 2.1 and 2.2 show examples of measurements of one of the sensors for the ITk Planar Pixel Market Survey. The resulting thickness was calculated to be 192 ± 8 μm with a bow of -8.3 μm[2]. For the thickness measurement the sensors are fixated onto the chuck by applying negative pressure onto multiple small holes in the chuck. So as to not distort the curvature of the sensor, no pressure is applied during the bow measurement. The thickness measurement is performed ten times per sensor. The nominal value (here 192 μm) corresponds to the mean, the uncertainty (here 8 μm) to the standard deviation of these measurements. Just one bow measurement is performed per sensor.

	h_1	$h_{1'}$	h_3	$h_{3'}$	h_7	$h_{7'}$	h_9	$h_{9'}$	T
1	9843.0	10058.0	9864.5	10069.3	9921.6	10085.3	9862.8	10064.0	196.2
2	9883.9	10022.6	9875.3	10069.0	9915.3	10094.8	9843.3	10051.2	180.0
3	9861.9	10077.8	9835.2	10054.4	9913.8	10073.4	9857.6	10052.1	197.3
4	9879.8	10064.9	9842.7	10063.0	9917.0	10071.1	9861.1	10048.2	186.7
5	9875.1	10069.9	9837.0	10074.2	9907.1	10103.2	9864.9	10050.0	203.3
6	9878.1	10054.8	9834.1	10058.9	9919.5	10075.3	9882.2	10043.2	179.6
7	9877.6	10051.5	9836.4	10046.2	9913.1	10091.9	9881.3	10075.4	189.2
8	9867.5	10070.7	9852.3	10076.6	9910.1	10091.9	9854.3	10060.6	203.9
9	9885.4	10085.9	9837.0	10060.6	9917.8	10098.7	9867.3	10047.9	196.4
10	9872.3	10057.2	9839.2	10068.0	9915.6	10089.5	9864.2	10035.5	189.7

Table 2.1: Thickness T (Eq. 2.5) measurements of a sensor for the ITk Planar Pixel Sensor Market Survey using the microscope focus method. The positions of the various measurement points are specified in Fig. 2.3(a). Each measurement was performed 10 times. All values are given in μm.

h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9	B
10096.0	10079.7	10079.9	10090.8	10089.0	10054.6	10085.0	10089.7	10088.2	-8.3

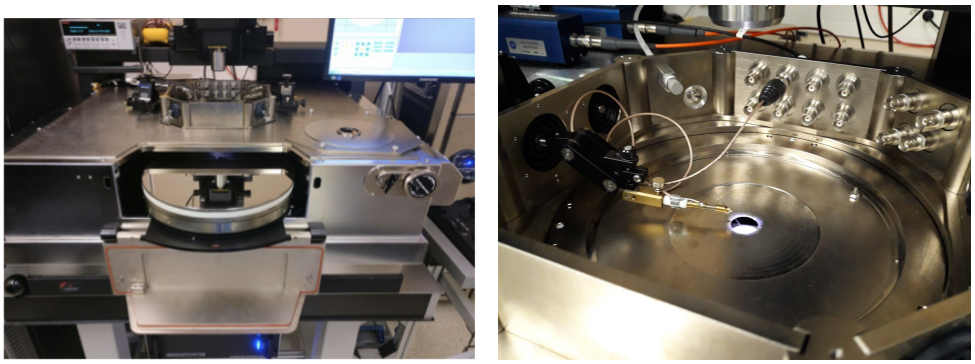
Table 2.2: Bow B (Eq. 2.6) measurements of a sensor for the ITk Planar Pixel Sensor Market Survey using the microscope focus method. The positions of the various measurement points are specified in Fig. 2.3(a). All values are given in μm.

[2]It should be noted that in this example the bow is dominated by just one measurement (h_6) which deviates much more from the other measurements then they do from each other

Similar geometrical measurements were performed on only one other sensor by the same vendor in order to validate the use of the microscope focus method[3]. The other sensor showed a thickness of 218 ± 22 μm and a bow of 12 μm. The corresponding tables can be found in the appendix: Tab. A.1.1 and Tab. A.1.2. If the setup is available at the institute, these measurements can alternatively be made with a laser scan profile meter (see Fig. 2.3(b)). The thickness and bow measurements obtained by this method were comparable to those made by the microscope focus method. The results from both methods were determined to be generally compatible which validated the use of the microscope focus method that could be performed by all institutes.

2.2.2 . Electrical Characterisation

After geometrical and visual inspection, several electrical tests were performed in the clean room. For this purpose, a probe station was used which allows for the regulation of temperature and humidity. Fig. 2.4(a) shows the chamber which can be closed off to ensure stable temperature and humidity and to shield off external electromagnetic radiation (including visible light). Fig. 2.4(b) shows the needle of the station (in the centre) which was used to make electrical contact with the sensor. The sensor is grounded by lying on the chuck of the probe station. This setup allows for measurements of depletion voltage and leakage current.

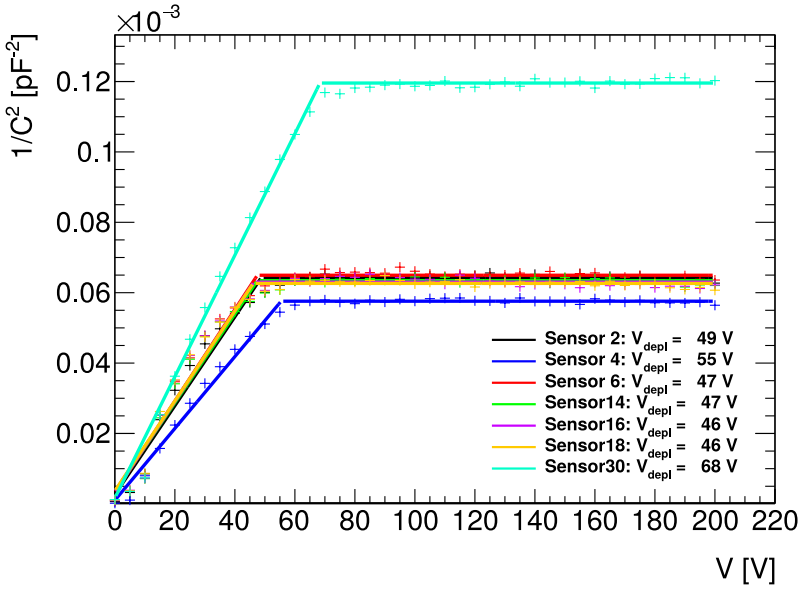


(a) (b)

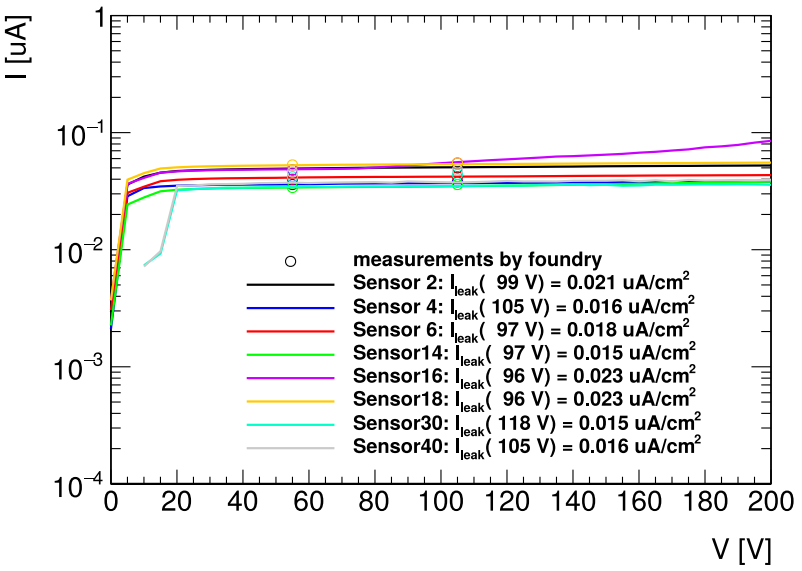
Figure 2.4: RF Wavelink WL350 probe station in the clean room at the IJCLab.

The depletion voltage of a given sensor is measured by applying various values of alternating bias voltage V_{bias} with a fixed frequency of $f = 10$ kHz. The resulting curve of measured capacitance C with respect to V_{bias} is expected to follow Eq. 2.1. A plot from measurements for several sensors can be seen in Fig. 2.5(a).

[3]The subsequent geometrical measurements on sensors on the vendor under study here were later performed by the LPNHE institute in Paris due to a more efficient setup.

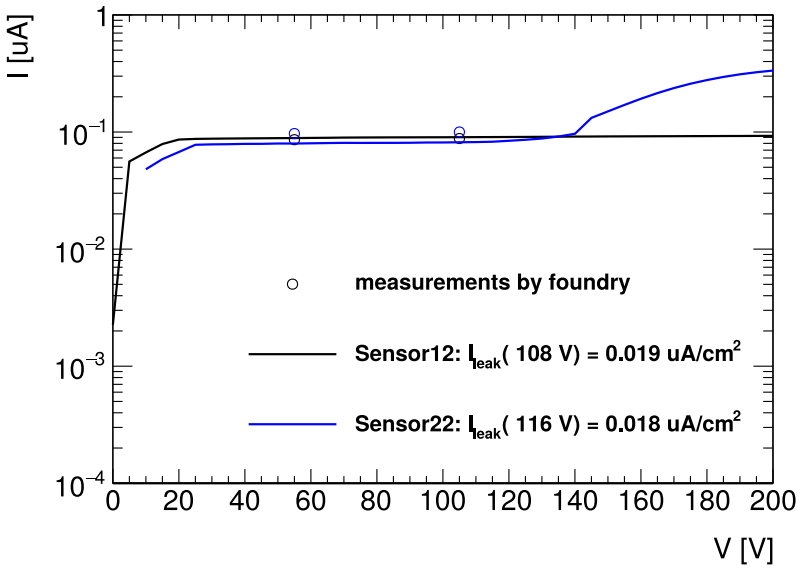


(a)

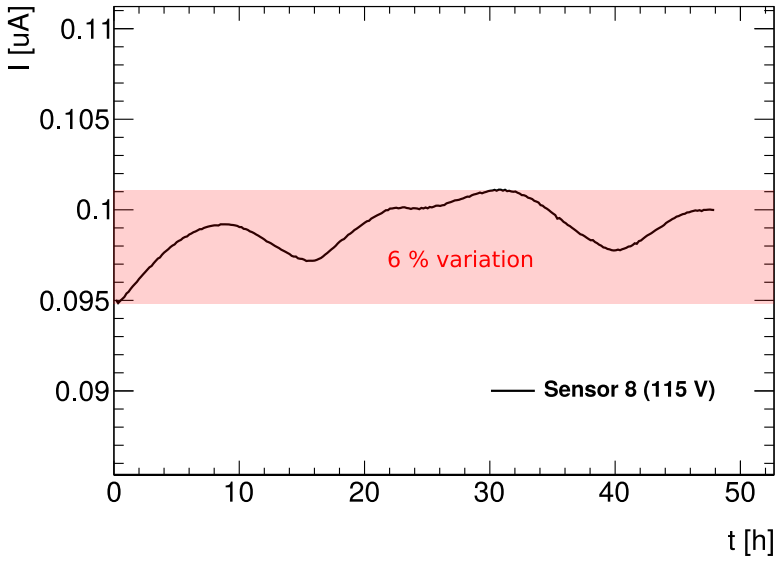


(b)

Figure 2.5: Capacitance C (a) and current I (b) scan as a function of applied bias voltage V for several sensors from a single wafer produced for the ITk Planar Pixel Market Survey. Leakage current measurements performed by the producing foundry are shown for comparison. They are plotted as circles.



(a)



(b)

Figure 2.6: (a): Current I scan as a function of applied bias voltage V for two sensors from a single wafer produced for the ITk Planar Pixel Market Survey where one shows a visible breakdown at 145 V. (b): Current I stability measurement over a time period of 48 h.

The requirement for the Market Survey states that the depletion voltage has to be below 100 V. All measurements shown in that figure are of non-irradiated 150 μm thick 50 μm$\times 50$ μm planar pixel sensors with punch-through bias structure and were performed at room temperature. A linear fit is applied to the rising section and a constant fit is applied to the stable fully depleted region to extract the depletion voltage. The depletion voltage V_{dep} (shown in the legend) corresponds to the intersection of the fitted lines.

Subsequently the leakage current was measured by applying various values of direct current bias voltage V_{bias}. A plot from measurements for several sensors can be seen in Fig. 2.5(b). Measurement points performed by the vendor are also shown. It is evident that they agree with the measurements performed at the institute. The requirement for the Market Survey concerns the leakage current at a voltage of $V_{\text{bias}} = V_{\text{depl}} + 50$ V. In this state the sensor is fully depleted which corresponds to how it would be operated if placed in a detector. The leakage current per sensor area must be below a threshold of 0.75 $\mu\frac{\text{A}}{\text{cm}^2}$. All sensors from this wafer showed values within limits specified for the Market Survey.

By applying even higher voltages, the breakdown voltage of sensors can be measured. In the scope of the Market Survey it is defined as the voltage at which the leakage current I_{leak} increases by more than 20% over a voltage step of 5 V. An example of this can be seen in Fig. 2.6(a). The stability of the leakage current is measured over a time period of 48 h (see Fig. 2.6(b)). For this purpose a constant bias voltage of $V_{\text{bias}} = V_{\text{depl}} + 50$ V is applied. The variation of the current is required to stay within 25%. While a 24 h day-night cycle may be identified in the shown current stability measurement, the variation is within the specified limit.

2.2.3 . Test Beam Hit Efficiency Measurements

The hit efficiency of several sensors was measured using a 5 GeV electron beam available at the DESY II Test Beam Facility [71]. While all previous measurements were made on bare sensors, test beam measurements were performed on modules. These consist of a sensor bumpbonded to a front end chip (RD53a [70]). The purpose of these measurements, as for all other measurements of the Market Survey, was to ascertain the efficiency of the sensors. The quality of the front end chips was studied in separate measurement campaigns within ATLAS.

A beam telescope consisting of six Mimosa [72] pixel detector plates, four scintillators for triggering, and an FE-I4 [73] pixel detector acting as reference plane was used (see Fig. 2.7). The Mimosa-26 pixel sensors consist of 576×1152 (10.6 mm $\times$ 21.2 mm) individual square pixels with a pitch of 18.5 μm, a sensor-thickness of 50μm, and a readout-speed of 115 MHz [72, 74]. Four scintillators in the beam line were used to trigger on individual electrons passing the telescope. Corresponding hits in the Mimosa detectors coinciding with a trigger signal were used to reconstruct the beam path using the General Broken Line (GBL) [75] algorithm in EuTelescope [76]. The thus reconstructed tracks are matched to corresponding hits in the two devices under test (DUTs) (i.e. the ITk planar pixel

sensors) in the middle of the telescope. During measurements, a styrofoam box (not shown in the picture) encloses the DUTs to shield them from light and, in the case of irradiated modules, to isolate the dry ice used for cooling.

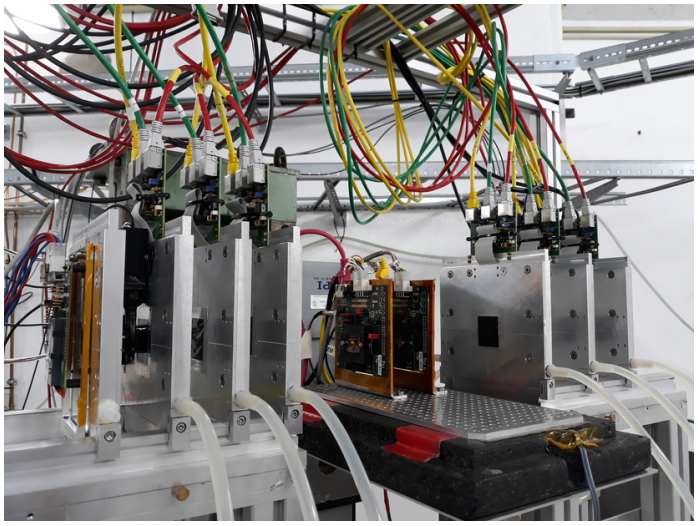


Figure 2.7: Test beam telescope setup used for the hit efficiency measurements of sensor modules for the ITk Planar Pixel Market Survey at the DESY II Test Beam Facility. Two devices under test (DUT) are situated in the middle in between two triplets of Mimosa sensors. The FE-I4 reference sensor is situated at the downstream (left) end of the telescope just after two of the scintillators placed at a right angle. The other two scintillators are placed at the upstream (right) end and are not visible in the picture.

The hit efficiency is calculated as the number of hits registered by the DUT divided by the number of reconstructed tracks as measured by the beam telescope which pass the active area of a DUT. The threshold for an acceptable hit efficiency depends on the thickness of the active substrate of the sensor as well as the radiation dose the sensor had received before measurement. The applied bias voltage during the measurement also depends on these criteria. For non-irradiated sensors it is a function of the measured depletion voltage V_{dep} while for irradiated sensors constant values are defined. The requirements are summarised in Tab. 2.3.

Schematic representations of the two sensor geometries ($50\ \mu$m $\times$ $50\ \mu$m and $25\ \mu$m $\times$ $100\ \mu$m) under study in the following measurements can be seen in Fig. 2.8. Due to expected inefficiencies around punch through bias dots fiducial regions excluding these areas have been defined depending on the pixel geometry. They are indicated in green. Punch through bias dots are part of the biasing structure and constitute holes in the active material of the sensor. Not all sensors under measurement were equipped with a biasing structure of this kind. In some cases

alternative biasing methods (poly-silicon or temporary metal) avoiding this effect were used. Two efficiencies have been defined: A global efficiency ϵ_{glob} taking the full sensor area into account and a fiducial efficiency ϵ_{fid} only considering tracks passing the fiducial regions. The latter is only used in the case of a punch through biasing structure.

thickness [μm]	bias voltage V_{bias} [V]	fluence [$10^{15}\frac{\text{neq}}{\text{cm}^2}$]	efficiency [%]
100 and 150	$V_{\text{dep}} + 50$	0	98.5
100	300	2	97.0
100	400	5	97.0
150	400	2	97.0
150	600	5	97.0

Table 2.3: Criteria for hit efficiency and applied bias voltage for test beam sensor measurements in the ITk Planar Pixel Market Survey depending on thickness and neutron equivalent (neq) radiation fluence.

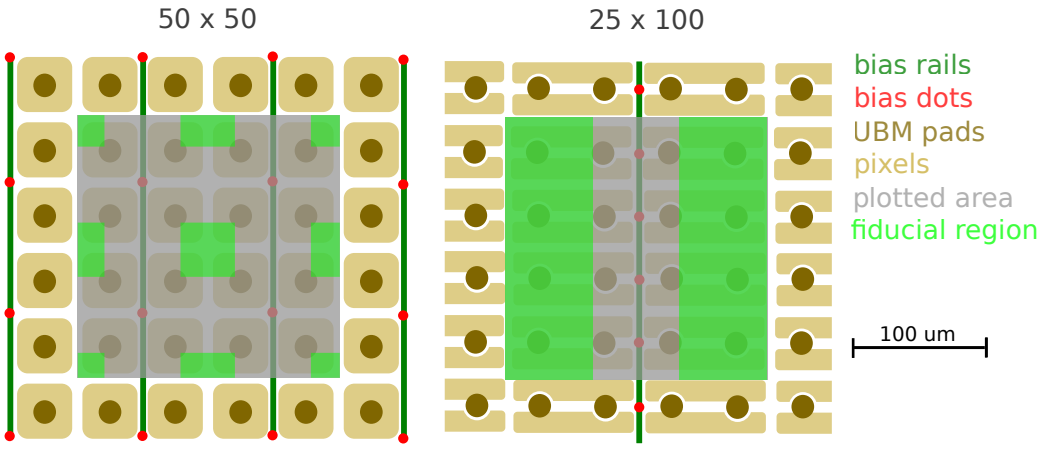
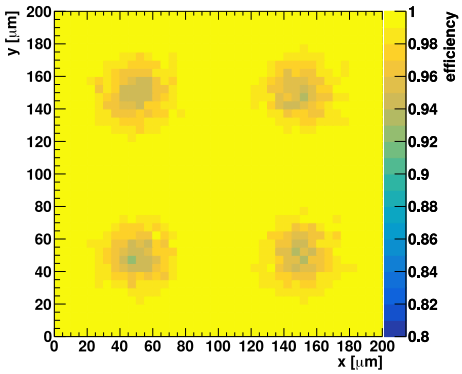


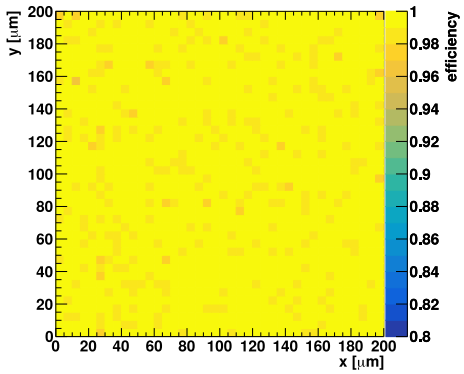
Figure 2.8: Schematic layout of the two different geometries of planar pixel sensors. The punch through biasing structure as well as the fiducial analysis regions defined to avoid the biasing dots are shown. The connection points to the frond-end readout chip correspond to the under-bump-metallisation (UBM) areas. The area corresponding to the plotted area in hit efficiency map plots (e.g. Fig. 2.9) is indicated in grey.

Fig. 2.9 shows the measured hit efficiency of two sensors corresponding to the two different geometries before and after irradiation. The plotted areas correspond to the grey areas indicated in Fig. 2.8. The hit efficiencies were measured over the full sensor area and then folded onto the indicated area according to the layout's symmetry. For the sensor with $50\ \mu\text{m} \times 50\ \mu\text{m}$ geometry (Fig. 2.9 a) four circular

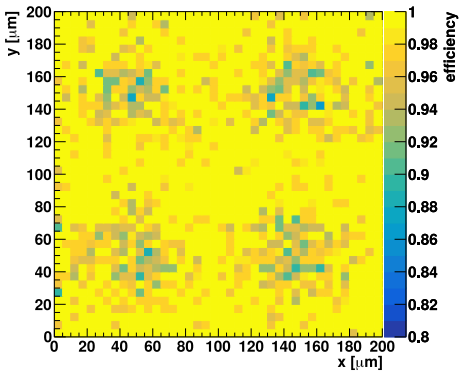
regions with reduced efficiency are visible. The localised drop in efficiency in these regions becomes more prominent after irradiation (Fig. 2.9 c). The corresponding regions of inefficiency for the sensor with $25\ \mu$m $\times\ 100\ \mu$m geometry (Fig. 2.9 b) are only visible after irradiation (Fig. 2.9 d). As expected, for both sensors a small reduction in overall hit efficiency (both global and fiducial) is noticeable after irradiation, indicating mild radiation damage as described in Sec. 2.1.3.



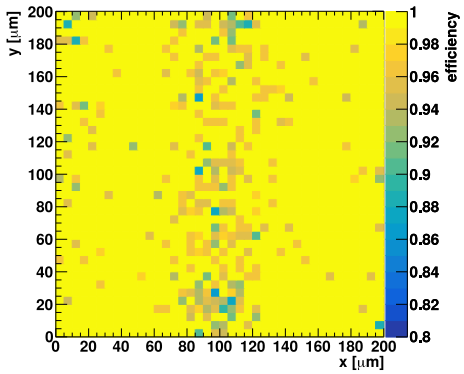
(a) V1S02 before irradiation:
$\epsilon_{\text{global}} = 99.2\% \rightarrow \epsilon_{\text{fiducial}} = 99.8\%$



(b) V2S03 before irradiation:
$\epsilon_{\text{global}} = 99.7\% \rightarrow \epsilon_{\text{fiducial}} = 99.7\%$



(c) V1S02 at $5 \times 10^{15}\ \frac{\text{neq}}{\text{cm}^2}$:
$\epsilon_{\text{global}} = 98.7\% \rightarrow \epsilon_{\text{fiducial}} = 99.7\%$



(d) V2S03 at $5 \times 10^{15}\ \frac{\text{neq}}{\text{cm}^2}$:
$\epsilon_{\text{global}} = 99.1\% \rightarrow \epsilon_{\text{fiducial}} = 99.5\%$

Figure 2.9: Hit efficiency maps of two sensors. Sensor V1S02 (a) has a $50\ \mu$m $\times\ 50\ \mu$m pixel geometry while sensor V2S03 (b) has a $25\ \mu$m $\times\ 100\ \mu$m pixel geometry. Results before (a,b) and after (c,d) irradiation to the indicated fluence in neutron equivalents (neq) are shown. The plotted areas correspond to the grey areas indicated in Fig. 2.8. The overall global efficiencies ϵ_{glob} measured over the complete maps as well as the fiducial efficiencies ϵ_{fid} measured only in the fiducial region are shown.

Hit efficiency measurements like these were performed on various sensors during four different test beam campaigns over two years. The results of this are summarised in Tab. 2.4. Besides the global efficiency ϵ_{glob}, the fiducial efficiency ϵ_{fid} is shown for sensors with punch through biasing structure (indicated in column 'PT'). Measurements were performed before irradiation and after irradiation to two different fluences given in neutron equivalents (neq). The pixel geometry is given in the format 'sensor thickness, pixel-length$\times$pixel-width'. Also shown is the absolute number and fraction of pixels that have been masked either due to being noisy or dead. Note that the total number of masked pixels for the analysis is higher since also direct neighbours of these pixels were masked. Efficiencies passing the Market Survey requirements given in Tab. 2.3 are indicated in green, failing in red. The bias voltages V_{bias} applied during measurement correspond if possible to the requirements given in Tab. 2.3. Deviations from this rule are marked in red and are due to too high leakage currents making operation at higher voltage not possible. It is evident that most sensors pass the hit efficiency requirements defined above (Tab. 2.3). In total, five measurements showed global efficiencies ϵ_{glob} falling below the threshold. They will be studied in detail in the following paragraphs.

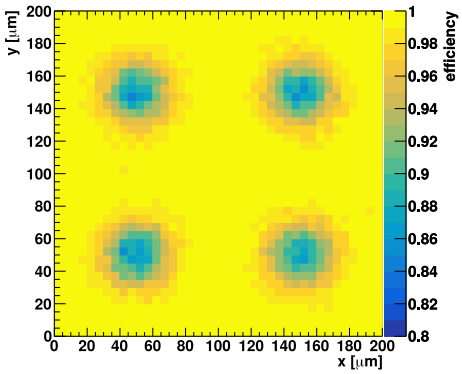
The measurements of V6S02 before irradiation and of V3S1x1028 at $5 \times 10^{15} \frac{\text{neq}}{\text{cm}^2}$ show global efficiencies ϵ_{glob} below, but fiducial efficiencies ϵ_{fid} above the threshold. It should be noted that for V6S02 ϵ_{glob} is just $\approx 0.7\%$ below threshold and shows both a global and fiducial efficiency above the threshold in a later measurement at $5 \times 10^{15} \frac{\text{neq}}{\text{cm}^2}$. The hit efficiency maps for both sensors can be seen in Fig. 2.10. For V6S02 measurements before and after irradiation are shown while V3S1x1028 has only been measured after irradiation. It is evident that the detrimental effect of the punch through (PT) biasing structure on hit efficiency is especially strong for both sensors. This becomes clear when comparing these to the maps of other sensors with PT that have been shown above in Fig. 2.9.

The other three sensors that show an efficiency below threshold show unusual masking patterns: V2S11 and V6S02 at $2 \times 10^{15} \frac{\text{neq}}{\text{cm}^2}$ and V3S14 at $5 \times 10^{15} \frac{\text{neq}}{\text{cm}^2}$. Masking is done to compensate for either dead or noisy pixels. Noisy pixels are pixels that fire constantly irrespective of the presence of an actual physical signal in the substrate of the sensor that could be initiated from a traversing particle. When integrated into a detector-setup, a high number of noisy pixels would quickly overwhelm the available readout bandwidth. Hence, as part of the pre-measurement tuning procedure of sensor modules, noisy pixels are identified and disabled, i.e. masked. Dead pixels are pixels that do not fire even if an ionising particle traverses nearby. They are identified during measurements if they never fire over a long-enough run period and based on that masked for the analysis of the run data. The number of both noisy and dead pixels usually increases drastically after irradiation due to damage of the substrate.

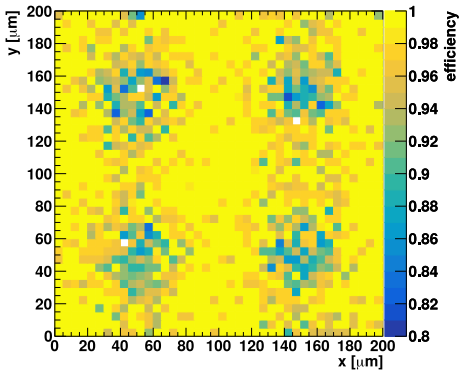
fluence $[10^{15}\frac{\text{neq}}{\text{cm}^2}]$	sensor ID	geometry [μm]	PT	V_{bias} [V]	ϵ_{glob} [%]	ϵ_{fid} [%]	masked pixels		comment
0	V1S01	$150,\ 50 \times 50$	yes	90	99.3685	99.7424	91	0.36%	
0	V1S02	$100,\ 50 \times 50$	yes	35	99.2435	99.8837	25212	100.00%	
0	V1S03	$150,\ 50 \times 50$	yes	90	99.2962	99.7780	25212	100.00%	
0	V2S01	$150,\ 25 \times 100$	yes	150	99.7347	99.7222	4	0.02%	
0	V2S03	$150,\ 25 \times 100$	yes	150	99.7860	99.7779	0	0.00%	
0	V2S11	$100,\ 25 \times 100$	-	75	99.6036	-	0	0.00%	
0	V2S12	$100,\ 50 \times 50$	-	150	99.8134	-	1	0.00%	
0	V2S13	$100,\ 25 \times 100$	-	100	99.9166	-	318	1.26%	
0	V2S14	$100,\ 25 \times 100$	-	75	99.9347	-	1	0.00%	
0	V3S05	$150,\ 25 \times 100$	yes	80	99.4554	99.8802	112	0.44%	
0	V3S06	$150,\ 25 \times 100$	yes	80	99.9494	99.9180	6191	24.56%	
0	V4S01	$150,\ 100 \times 25$	-	150	99.7371	-	6068	24.07%	
0	V4S02	$150,\ 50 \times 50$	-	150	99.6611	-	2556	10.14%	
0	V4S03	$150,\ 100 \times 25$	-	150	99.8133	-	1306	5.18%	
0	V4S06	$150,\ 100 \times 25$	-	150	99.2934	-	8432	33.44%	
0	V6S02	$150,\ 25 \times 100$	yes	130	98.8287	99.8658	9	0.04%	
0	V6S04	$150,\ 50 \times 50$	yes	130	98.6014	99.7185	3	0.01%	
0	V6S05	$150,\ 50 \times 50$	yes	130	98.3357	99.7831	2	0.01%	† Fig. 2.10(a)
0	V6S12	$100,\ 50 \times 50$	-	80	99.1251	-	73	0.29%	
0	V6S13	$100,\ 50 \times 50$	-	80	99.6120	-	8562	33.96%	
2	V1S01	$150,\ 50 \times 50$	-	400	98.9955	-	212	0.84%	
2	V2S05	$150,\ 25 \times 100$	-	400	99.4303	-	1394	5.53%	
2	V2S11	$100,\ 25 \times 100$	-	300	29.6496	-	14655	58.13%	‡ Fig. 2.11(d)
2	V3S01	$150,\ 50 \times 50$	-	400	99.8105	-	4402	17.46%	
2	V3S12	$100,\ 50 \times 50$	-	300	99.6079	-	188	0.75%	
2	V4S01	$150,\ 50 \times 50$	-	400	98.9362	-	22315	88.51%	
2	V6S02	$150,\ 25 \times 100$	-	400	91.7406	-	8041	31.89%	‡ Fig. 2.11(a)
2	V6S11	$100,\ 50 \times 50$	-	300	99.3817	-	669	2.65%	
5	V1S02	$100,\ 50 \times 50$	yes	400	98.7103	99.7791	742	2.94%	
5	V1S03	$150,\ 50 \times 50$	-	550	98.5852	-	6590	26.14%	
5	V2S03	$150,\ 25 \times 100$	-	600	99.3142	-	2447	9.71%	
5	V2S13	$100,\ 25 \times 100$	-	350	99.2870	-	11775	46.70%	
5	V3S02	$150,\ 25 \times 100$	-	600	99.7043	-	844	3.35%	
5	V3S03	$150,\ 50 \times 50$	-	600	99.6966	-	156	0.62%	
5	V3S11	$100,\ 25 \times 100$	-	400	99.0070	-	1610	6.39%	
5	V3S14	$100,\ 50 \times 50$	yes	350	92.5727	96.7016	6612	26.23%	‡ Fig. 2.11(b)
5	V3S1x1026	$100,\ 25 \times 100$	yes	400	98.8372	99.6258	2985	11.84%	
5	V3S1x1028	$100,\ 50 \times 50$	yes	400	96.2227	99.0838	995	3.95%	† Fig. 2.10(c)
5	V3xx1026	$100,\ 25 \times 100$	yes	400	97.8301	98.0776	0	0.00%	
5	V4S03	$150,\ 50 \times 50$	-	600	99.8658	-	8174	32.42%	
5	V6S03	$150,\ 25 \times 100$	-	600	97.5980	-	2034	8.07%	
5	V6S05	$150,\ 50 \times 50$	yes	600	97.7773	99.6792	141	0.56%	
5	V6S16	$100,\ 50 \times 50$	-	400	99.1557	-	630	2.50%	

†: only ϵ_{fid} above threshold
‡: pattern in mask

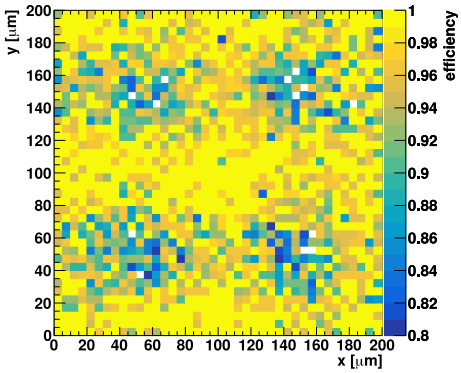
Table 2.4: Hit efficiency measurements of various sensors for the Planar Pixel Market Survey.



(a) V6S05 before irradiation:
$\epsilon_{\text{global}} = 98.2\% \rightarrow \epsilon_{\text{fiducial}} = 99.6\%$



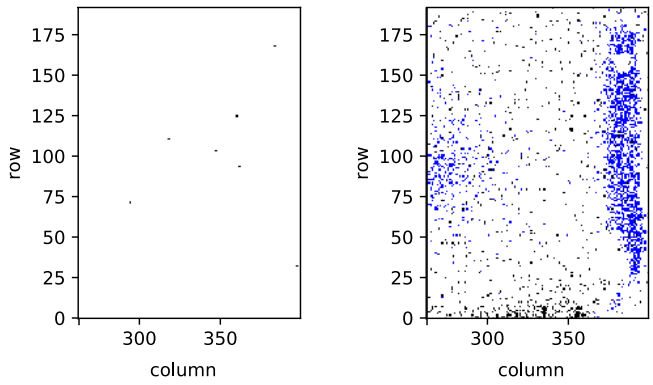
(b) V6S05 at $5 \times 10^{15} \frac{\text{neq}}{\text{cm}^2}$:
$\epsilon_{\text{global}} = 97.8\% \rightarrow \epsilon_{\text{fiducial}} = 99.7\%$



(c) V3S1x 1028 at $5 \times 10^{15} \frac{\text{neq}}{\text{cm}^2}$:
$\epsilon_{\text{global}} = 96.2\% \rightarrow \epsilon_{\text{fiducial}} = 99.1\%$

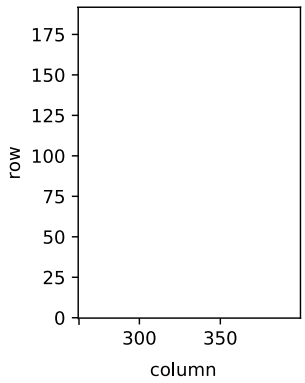
Figure 2.10: Hit efficiency maps of two sensors where the detrimental effect of the punch through biasing structure on the hit efficiency is especially pronounced. As a result the global efficiency ϵ_{glob} is significantly lower than the fiducial efficiency ϵ_{fid}. Both sensors have $50\ \mu$m $\times\ 50\ \mu$m pixel geometry.

Fig. 2.11 shows the masking pattern for the three sensors in question. For V2S11 and V6S02 the patterns are shown before and after irradiation. It is clearly visible that patterns of noisy and dead pixels emerge only after irradiation. Due to the localised nature of these patterns it is not clear if, instead of being caused by the irradiation alone, they may alternatively originate from a mechanical failure like disconnected bump bonds. This could be caused from stress when handling and shipping the sensors in-between measurement and irradiation periods. V3S14 has not been measured before irradiation but also here distinct patterns are visible after irradiation.

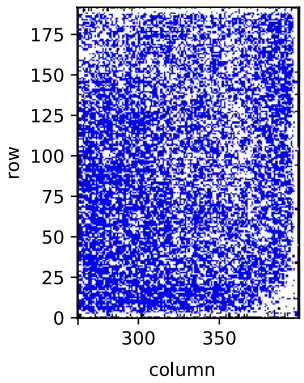


(a) V6S02 before irradiation

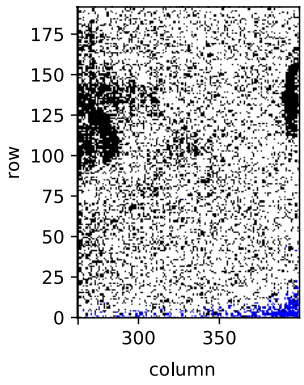
(b) V6S02 at $2 \times 10^{15} \frac{\text{neq}}{\text{cm}^2}$



(c) V2S11 before irradiation



(d) V2S11 at $2 \times 10^{15} \frac{\text{neq}}{\text{cm}^2}$



(e) V3S14 at $5 \times 10^{15} \frac{\text{neq}}{\text{cm}^2}$

Figure 2.11: Masking maps of three sensors showing unusual patterns. Pixels masked as dead pixels are marked in blue. Pixels marked in black correspond to noisy pixels. For all measurements of the Market Survey only the rightmost third of the sensor was used (starting at column 264). This is because the prototype RD53A readout chip [70] is built with three different analogue frontends (left-to-right): synchrotron, linear, differential. The final ITk readout will use a differential front end.

2.2.4 . Conclusion

The ITk Planar Pixel Market Survey was an effort to qualify a number of vendors for the final production of the ITk sensors. For this purpose various types of measurements were performed on prototype sensors.

The visual inspection of the sensors verified that defects like missing or deformed bump pads are a rare occurrence and the necessary yield for the ITk production can be achieved by all vendors. The presented measurements of electrical characteristics (depletion voltage, breakdown voltage, and leakage current) on non-irradiated sensors showed overall good performance from the vendor that was under study in the presented results. Other institutes performed these tests also on irradiated sensors verifying that radiation damages significantly affect the electrical characteristics of only a small number of sensors. Geometrical test of thickness and bow verified that the sensors were uniform enough to be bonded to front end chips.

After bump-bonding the resulting modules were tested in a test beam environment. Hit efficiency was measured for modules before and after irradiation to two different fluences. It was shown that most modules passed the efficiency requirements. A small number of modules were strongly affected by efficiency degradations in an area around the punch through bias structure. Efficiency measurements in a fiducial region excluding such structures showed that they otherwise performed within the requirements. Some modules showed significant patterns in the noisy- and dead-pixel masks after they had been irradiated. These effects resulted in lower efficiencies. The test beam measurements showed overall that degrading effects from irradiation affect only a small number of sensors to such an extent that they fall below the required threshold.

In summary, the Market Survey demonstrated that the sensor prototypes show the necessary requirements for the ITk. The Final Design Review (FDR) for the planar pixels has been held and passed in September 2020 based on the measurements presented above. As a result three vendors where chosen, the pre-production of the planar sensors has started, and the full production is expected to commence soon.

3 - Jet Reconstruction and Classification

This chapter presents work on vector boson taggers, classification algorithms with the goal of discerning whether a detector signature has been initiated by a vector boson (W or Z) or a different particle. One of the taggers described and studied in this chapter, the 3-variable tagger, has been used in the main analysis of this thesis (Sec. 4), the search for semileptonic vector boson scattering. Based on the studies presented in this chapter it may be possible to integrate one of the alternative methods in an eventual follow-up analysis with the prospect of increasing the analysis sensitivity.

The chapter is divided into two parts, starting with Sec. 3.1 which describes the necessary theoretical and phenomenological concepts for this task. The process of reconstructing hadronic decay structures and the resulting particle showers from raw detector signatures in the form of so called jets is presented. The following Sec. 3.2 presents work leading up to a study in which the performance of different tagging methods on jets formed from unified flow objects (UFOs) are investigated.

UFOs are the result of a recently proposed scheme of combining signatures from different detector components which in previous studies have been proven to yield superior resolution of the substructure of jets. It is investigated to what extent this may translate into improved efficiencies of various vector boson tagging methods. The final results recently have been published in the form of a public note [77]. It is hypothesised that, using the novel UFO jets, the improved tagging efficiency may lead to higher sensitivities in studies like the vector boson scattering analysis of this thesis.

3.1 . Hadronic Decay Structures

The reconstruction of hadronic decay structures is especially important in hadron collider experiments like ATLAS. The following sections give an overview on the underlying physical phenomena and reconstruction methods. Sec. 3.1.1 describes how hadrons are formed during a process called hadronisation. Sec. 3.1.2 shows how the detector signatures can be used to form and calibrate jets, a representation of the resulting particle showers. Finally, Sec. 3.1.3 describes the unique challenges when working with highly boosted decay structures and how jet substructure can help with their classification.

3.1.1 . Hadronisation

Due to colour confinement all particles with colour charge cannot be detected in isolation[1]. Hence, quarks produced in particle collisions tend to be difficult

[1]This is with the exception of the top quark which has a lifetime shorter than the average hadronisation time scale [78].

to reconstruct from raw detector signatures. As a consequence the simulation of processes involving the strong force is unique among all interactions of the Standard Model and especially challenging. This phenomenon is a consequence of the non-abelian structure of the $SU(3)$ group which, in the Standard Model, describes the underlying symmetry of the strong interaction (see Sec. 1.2.1). Consequently, gluons, being the force carriers of this interaction, possess colour charge themselves, leading to self-interaction.

The result of this can phenomenologically be understood as the formation of so called QCD flux tubes. These string-like excitation states of the gluon field form in between colour-charged particles if they are separated. The tension of these tubes increases the further the separation grows. New quark pairs are formed in between the original particles once a high enough energy from tension is reached. They must then combine into colour neutral states. This phenomenon is called hadronisation. This process continues until the resulting final state consists exclusively of colour-neutral states.

Two phenomenological models for this process are widely used in Monte Carlo generators. The PYTHIA generator uses the Lund Fragmentation model [79]. In this model, the flux tubes are described as strings gaining in tension until they eventually break down forming colour-neutral hadrons. The hadrons are chosen in such a way that their combined masses and momenta correspond to the energy stored in the tension of the string as well as the string's momentum at the time of breakdown. The HERWIG generator uses the cluster fragmentation model [80] in which gluons are first split into $q\bar{q}$ pairs which then are collected into overall colourless clusters and combined into hadrons.

3.1.2 . Jets

The hadrons formed by hadronisation processes may further decay before their products eventually produce a detector signature. In an analysis situation, one is usually interested in the particles at the very beginning of this chain. It is therefore necessary to define algorithms acting on the resulting detector signatures to deduce the properties of the initial particles and form objects, called jets, with properties as close as possible to those of the particle itself. Ref. [81] gives an overview on these objects, the main points of which are presented in the following.

Two components are necessary for the task of building jets. A jet-finding algorithm and a recombination scheme. The former specifies the order in which objects should successively be combined until the final objects, the jets remain. The latter defines the method of combination at each individual step. The most widely used recombination scheme is the four-vector sum, simply combining the momenta of individual objects. It is used in all applications throughout this thesis.

Jet algorithms fall into two general categories: Cone algorithms and sequential recombination algorithms. The iterative cone algorithm with progressive removal (IC-PR) is an example of the former. The object with highest transverse momentum is used as a seed. Subsequently, all objects with an angular separation ΔR

smaller than a specified parameter R are combined with the seed. This parameter is called the jet radius. The resulting object, now with a slightly different direction of momentum, is used as the new seed. This procedure is repeated until the direction does not change significantly anymore, in which case the resulting object is defined to be a jet. Then all contained objects are removed from the initial set. This is repeated on all remaining objects until none are left, leaving a list of jets instead.

Widely used examples of sequential recombination algorithms are the closely related Cambridge-Aachen [82], k_T [83], and anti-k_T [84] algorithms, the latter of which is used throughout this thesis. They are defined on the set $\{i\}$ of input objects by the following rule-set:

1 Find the minimum of the parameters d_{ij} (for all pairs of objects) and d_i (for all objects) defined as:

$$d_{ij} = \min(p_{\mathrm{T}\,i}^{2\lambda},\ p_{\mathrm{T}\,j}^{2\lambda})\frac{\Delta R_{ij}}{R^2}, \tag{3.1}$$

$$d_i = p_{\mathrm{T}\ i} \tag{3.2}$$

2a If one of the d_i is smallest (out of all d_i and d_{ij}), define object i as a jet, remove it from the set of objects, and return to step 1.

2b If not, i.e. if one of the d_{ij} is smallest, recombine i and j and return to step 1.

3 Stop when no objects are left.

The only difference between the three algorithms lies in the value of the parameter $\lambda = 0, 1,$ or -1 for Cambridge-Aachen, k_T, or anti-k_T. Since this parameter is in the exponent of the transverse momenta in Eq. 3.1 but does not influence the angular separation part (ΔR_{ij}) of that equation, it specifies the algorithm's recombination strategy with respect to the constituents' momenta. As a consequence, the k_T algorithm ($\lambda = 1$) tends to recombine soft (i.e. low p_{T}) constituents earlier than the anti-k_T algorithm ($\lambda = -1$). This property of the anti-k_T algorithm is called 'soft-resiliance' and makes the algorithm more robust with respect to contamination from underlying-event or pileup [84]. The parameter R, like for the IC-PR algorithm, is called the radius of the jet. In contrast to most cone algorithms, jets formed by these sequential recombination algorithms are not guaranteed to be strictly cone-shaped, meaning R does not correspond to a geometrical radius in the strictest sense. Nethertheless, jets formed with the anti-k_{T} algorithm are known to be approximately cone symmetric most of the time.

The energy deposited in the calorimeters of the detector as well as reconstructed tracks (in the case of charged hadrons) can be used as inputs for jet algorithms. Jets formed from these objects are called reconstructed jets. It is also possible to perform jet algorithms directly on simulated final state particles from

Monte Carlo generation. Jets formed from these objects are called truth jets within the ATLAS collaboration. With the help of truth jets it is possible to define a calibration scheme correcting mass and energy of reconstructed jets. In this way the parameters of reconstructed jets are expected to be similar to the corresponding parameters of the initial partons.

3.1.3 . Boosted Decay Structures

Hadronically decaying W and Z bosons decay into two quarks: $W/Z \to qq'$. In the mother particle's rest frame the decay occurs back to back yielding clearly separated remnants of q and q'. The transverse momentum $p_{\mathrm{T}}^{W/Z}$ of the mother particle leads to a boost of the system from its rest frame into the laboratory frame. Small boosts change the decay structure only mildly and the back to back detector signatures of q and q' may be reconstructed as two clearly separated jets. A common choice of jet radius in such cases is $R = 0.4$. Throughout this thesis, this will be called the resolved regime.

However, in the case of large transverse momentum $p_{\mathrm{T}}^{W/Z}$, this may result in a structure in which the decay products of q and q' are very collimated and may even overlap (see Fig. 3.1). In this case it may be advantageous to reconstruct the decay products with a single large-radius jet instead. A common choice of jet radius here is $R = 1.0$. This is called the merged regime.

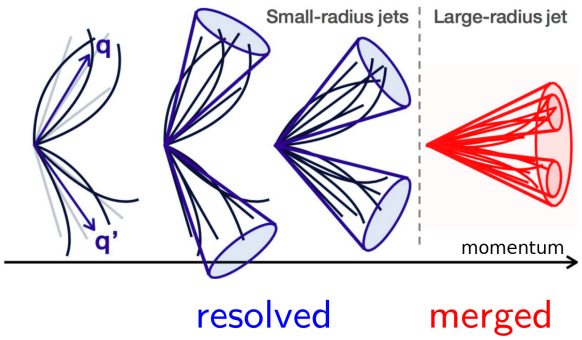


Figure 3.1: Illustration of a hadronic vector boson decay in the resolved and merged regime.

Studies [85] on simulated W boson decays resulted in the approximate relation

$$\Delta R(q, q') \approx \frac{2m_W}{p_{\mathrm{T}}^W} \tag{3.3}$$

for the angular separation of the daughter particles based on mass and transverse momentum of the mother particle (see Fig 3.2).

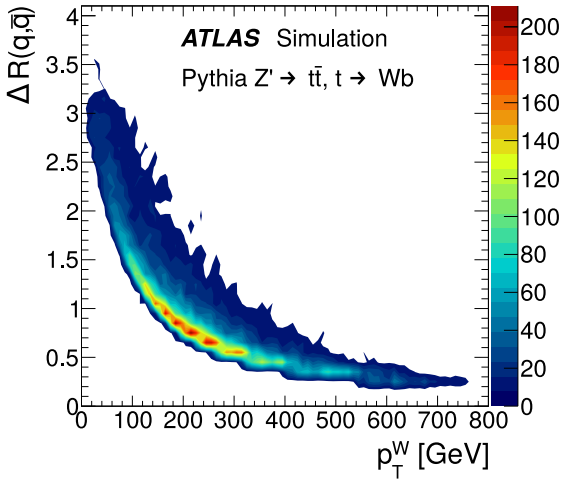


Figure 3.2: Angular separation ΔR between the quarks q, q' produced in simulated W decays in the ATLAS detector as a function of transverse momentum [85].

Based on this, a W boson ($m_W \approx 80$ GeV) with a transverse momentum greater than 200 GeV is likely to produce daughter particles that, if reconstructed as $R = 0.4$ jets, would overlap ($\Delta R(q, q') < 0.8$) and would therefore be better described by a single large-radius jet.

In the merged regime, information from the substructure of the large-radius jet can be used to identify the type of the mother particle. Various different substructure variables have been proposed for this purpose. Most of them fundamentally rely on the same principle: Hadronically decaying vector bosons decay into two quarks. This two-pronged decay leads to distinctive structures in both the distribution of energy inside of the jet cone, as well as the order in which its constituents are clustered within conventional jet finding algorithms.

3.2 . Vector Boson Tagging Studies

Many processes under study at the LHC, including the vector boson scattering process forming the focus of this thesis, include vector boson (W or Z) decays in their final state. In many analysis situations such events may constitute the signal while events with similar kinematics initiated not by vector bosons but by gluons or quarks are considered background. Hence, a reliable classification of such events is desirable. In the case of a decay into leptons, the decay signatures of these processes differ fundamentally. For hadronically decaying vector bosons however, it is more challenging to perform a reliable separation from background. This is because they are reconstructed as jets which is also the case for background QCD multijet production initiated by quarks and gluons. The classification of

jets to originate from either vector bosons or other particles is done with tagging algorithms.

The semileptonic vector boson scattering analysis described in this thesis greatly profits from the classification strength of the vector boson tagger which is applied in the merged regime. The tagger implemented in the analysis is a 3-variable cut-based tagger using jet mass m, number of tracks n_{trk} associated with the jet, and a substructure variable D_2 as input.

In this context the jet mass is a potent discriminant. For jets initiated by vector bosons it tends to be in the vicinity of the boson's rest mass. Unfortunately this leads to an effect that may be unwanted within an analysis situation: If the classification by a tagging algorithm strongly relies on the jet mass, the probability of background events passing the tagger (i.e. being misidentified as signal) will be highly correlated to its jet mass. Consequently, the distribution of the jet mass for such background will be enhanced in the vicinity of the vector boson mass. For the signal this is naturally the case. Hence this can be viewed as a shaping of the background distribution according to the signal distribution. This may even be the case if the jet mass is not explicitly used as an input for the tagging algorithm. It is sufficient that the input variables are correlated with the jet mass. Since this correlation can be non-linear, accounting for it may be challenging. The goal of a mass-decorrelated tagger as presented in the following sections is to achieve good classification strength while not allowing for too much shaping of the background distribution; and therefore still allowing for the use of a data-driven background estimate as described above. In many analyses one expects a smoothly falling mass distribution in the background and an excess (bump) around the W/Z mass which is caused only by the signal. This makes a data-driven approach for the estimation of the background in the signal region possible: A smoothly falling distribution can be fitted to data from the low- to the high-mass sideband control regions left and right of the signal region. The resulting fit function in the signal region is then the data-driven background estimate. It would be beneficial to apply a vector boson tagger before this to enhance the signal-to-background ratio. But if this tagger is correlated to the jet mass it would spoil the smoothly falling background distribution, introducing a bump in the signal region. Then this method would no longer be valid.

In the following three different methods of decorrelation are presented: Sec. 3.2.1 presents an analytical method decorrelating one of the variables (D_2) of the baseline 3-variable tagger, Sec. 3.2.2 describes the use of an adversarial neural network architecture to decorrelate a deep learning based tagger, and Sec. 3.2.3 describes a genetic training method. The following Sec. 3.2.4 compares the newly developed genetic method with the established adversarial method with respect to classification power and mass-decorrelation. Sec. 3.2.5-3.2.8 apply these methods to jets constructed from unified flow objects (UFO), a recently invented type of jet constituents, comparing it to results using established jet definitions as well as

comparing the taggers with each other. Finally, Sec. 3.2.9 presents conclusions and an outlook of potential further work based on the studies presented.

The studies presented in the following were performed on Monte Carlo simulated samples with a centre of mass energy of $\sqrt{s} = 13$ TeV which have been passed through the full ATLAS detector simulation based on GEANT4 [86]. The background samples correspond to the production of jets originating from light quarks or gluons and are referred to as multijet. The signal samples contain jets originating from W bosons. The W bosons were produced in high-mass sequential Standard Model $W' \to WZ \to q\bar{q}q\bar{q}$ decays [87] with a W' mass of $m_{W'} = 2$ TeV. More detailed information on the signal and background samples can be found in the public note [77] in which a part of these studies were published.

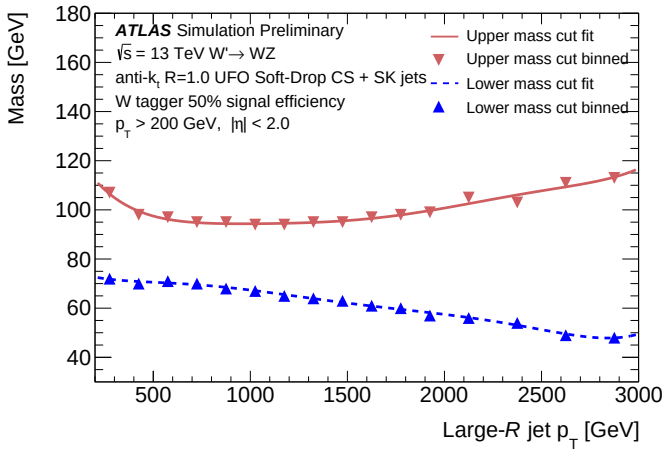
3.2.1 . 3-Variable Tagger

The baseline vector boson tagger within ATLAS is a 3-variable tagger. It is also the tagger used for the main focus of this thesis, the vector boson scattering analysis presented in Sec. 4. Its selection criterion is based on an upper and lower cut on the jet mass m reconstructed from the four vector sum of the jets' constituents, as well as lower cuts on the substructure variables D_2 and n_{trk}. Here D_2 is an energy-correlation function ratio as defined in Ref. [88] (with $\beta = 1$) and n_{trk} is the number of tracks with $p_{\text{T}} > 500$ GeV associated[2] with the jet. While D_2 makes use of the tendency of vector bosons to decay into a two-pronged structure [88], the number of tracks within a jet is known to be a good discriminant between quark- or gluon-initiated jets [91].

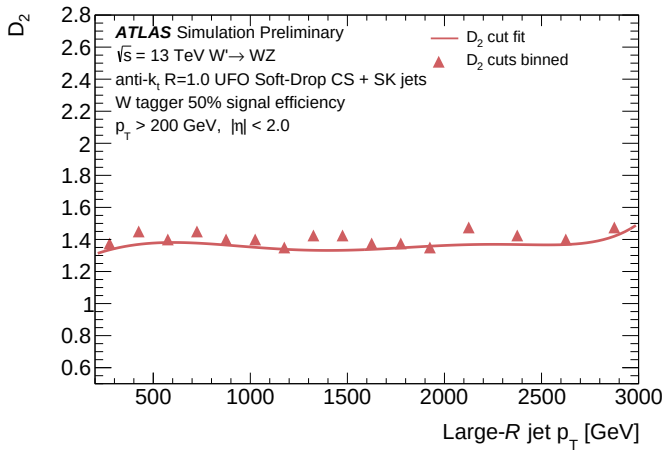
The cut value for each variable is chosen to yield the highest background rejection $1/\epsilon^{\text{bkg}}$ at a signal efficiency of $\epsilon^{\text{sig}} = 50\%$ independently of the transverse momentum p_{T} of the jet. They are derived in bins of p_{T} and interpolated with a 6th order polynomial. Signal and background efficiency in this context are defined as $\epsilon^{\text{sig}} = N^{\text{sig}}_{\text{pass}}/N^{\text{sig}}_{\text{tot}}$ and $\epsilon^{\text{bkg}} = N^{\text{bkg}}_{\text{pass}}/N^{\text{bkg}}_{\text{tot}}$ where $N^{\text{sig/bkg}}_{\text{pass}}$ are the number of signal/background events satisfying the tagger requirement and N_{tot} are the total number of events passing the event selection.

The resulting cuts (here derived for the UFO jets) in the presented study can be seen in Fig. 3.3. Compared to an earlier study [92], which derived the optimal cuts for LCTopo [93] jets, the optimal cut on D_2 (derived here for UFO jets) is much less dependent on p_{T} (compare Fig. 3.3(b) with Fig. 3.4). This is attributed to the known better angular resolution of UFO jets [94]. The tagging efficiency of this tagger will be presented in Sec. 3.2.7 where its performance is compared to the performance of the other tagging methods which are defined in the following.

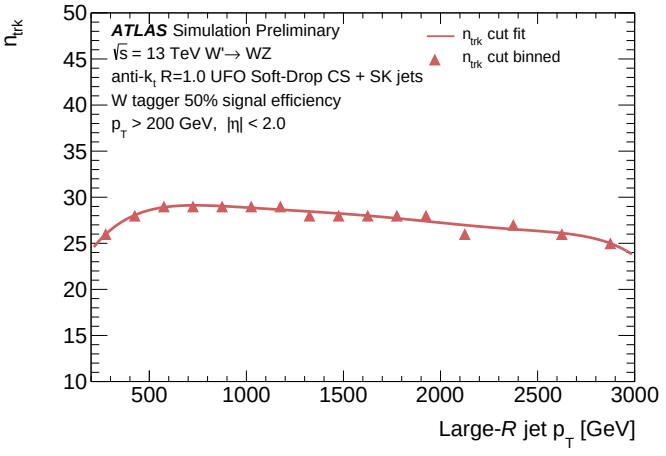
[2] Ghost association is used here. See [89, 90] for a definition.



(a) m



(b) D_2



(c) n_{trk}

Figure 3.3: Optimal cuts on m, D_2, and n_{trk} derived for UFO jets [77].

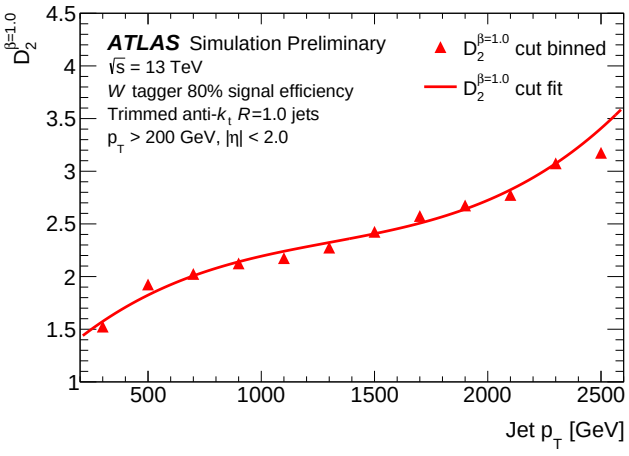


Figure 3.4: Optimal cut on D_2 for the 3-variable tagger on LCTopo jets as derived in an earlier study (Ref. [92]).

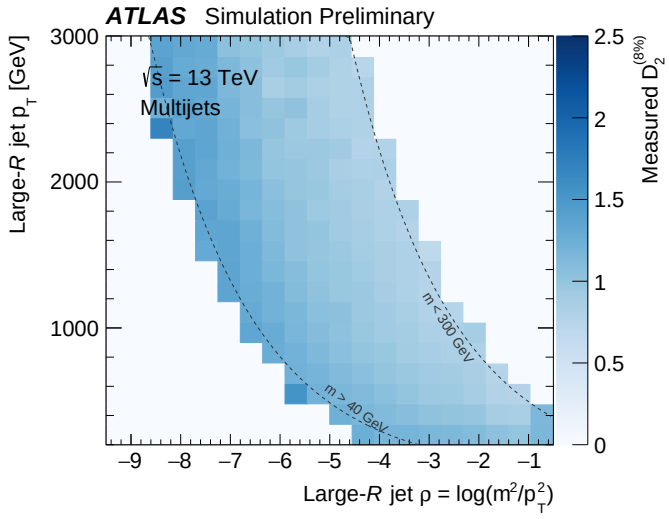
Since the 3-variable tagger includes cuts on the jet mass it would not be constructive to try to decorrelate the tagger as a whole from this variable. Instead the cut on D_2 alone was studied in the mass window provided by the tagger. In this way it is possible to ascertain the strength of 1-variable decorrelation methods in the region where mass-decorrelation is most needed, i.e. in the vicinity of the vector bosons' masses.

An analytical decorrelation method was chosen making use of fixed-efficiency regression: Cut values on D_2 corresponding to a fixed background efficiency of $\epsilon_{\text{bkg}}^{\text{rel}} = 8\%$ were determined in bins of p_{T} and $\rho = \log(m^2/p_{\text{T}}^2)$. This background efficiency was chosen since it was determined to correspond roughly to a signal efficiency of $\epsilon_{\text{sig}} = 50\%$. The resulting two-dimensional histogram (seen in Fig. 3.5(a)) is fitted using the k-nearest neighbour (k-nn) method. The result of this procedure can be seen in Fig. 3.5(b).

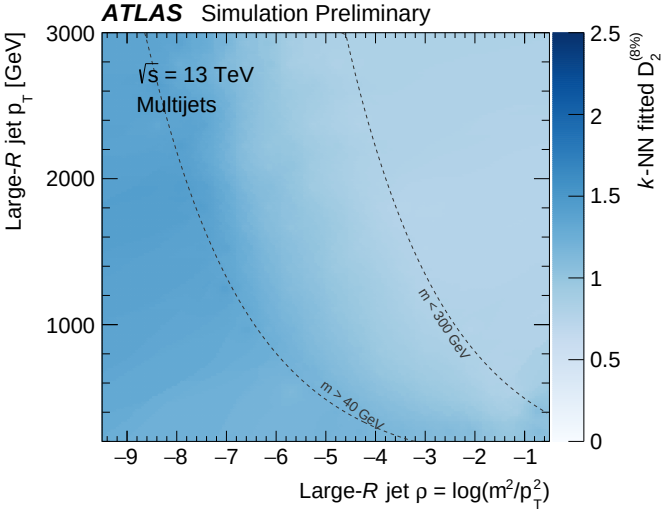
The resulting fit is a function $D_2^{(8\%)}(\rho(p_{\text{T}}, m), p_{\text{T}})$ parametrising the mass-dependence of the cut on D_2 over the p_{T} spectrum. Subtracting this function from D_2 yields the parameter

$$D_2^{\text{k-nn}} = D_2 - D_2^{(8\%)} \tag{3.4}$$

which retains some of the signal/background separation power while having most of the correlation to the jet mass removed. This will be quantified in Sec. 3.2.7 when comparing the performance of different tagging methods. D_2 will be used as a single variable tagger and $D_2^{\text{k-nn}}$ as its mass-decorrelated version.



(a)



(b)

Figure 3.5: Cut values on D_2 to achieve a background efficiency of $\epsilon_{\text{bkg}}^{\text{rel}} = 8\%$. (a): Histogram in bins of ρ and p_{T}. (b): Resulting k-nn fit [77].

3.2.2 . Adversarial Neural Networks

In many cases multivariate analysis (MVA) methods like boosted decision trees or neural networks tend to outperform cut-based taggers like the one presented in the previous section. This can be attributed mostly due to two effects: First, MVA methods are able to combine information from many input variables. Secondly, they are able to make use of higher order correlations among the variables.

This however also makes these methods more susceptible to unwanted corre-

lations to parameters like the jet mass, even if it is not explicitly used as an input. Since the jet mass is a good discriminant for the decision task at hand, it is likely that a MVA method learns to determine this parameter from its input parameters and bases the tagger decision on it. Hence decorrelation is especially important when using methods like these.

A study on the use of adversarial neural networks to define a mass-decorrelated vector boson tagger was presented in a previous ATLAS public note [95]. In that study, the performance of this multivariate method was compared to several analytical methods, as well as an alternative multivariate method using adaptive boosting of decision trees. The adversarial method was determined to perform best closely followed by the decision trees. It was therefore chosen for the following study.

A neural network architecture was utilised which can be divided into two separate parts: the classifier and the adversarial (see Fig. 3.6). The classifier takes a set of substructure variables $\vec{x}$ from the large-radius jet in question as input and has a single output node with a continuous value z. The set $\vec{x}$ chosen for the study at hand is summarised in Tab. 3.1. The adversary is an additional network taking the output of the classifier as input. Using a Gaussian Mixture Model it estimates the jet mass purely based on the output of the classifier. The estimated mass is compared to the actual reconstructed mass of the large-radius jet yielding the single continuous output of the adversary.

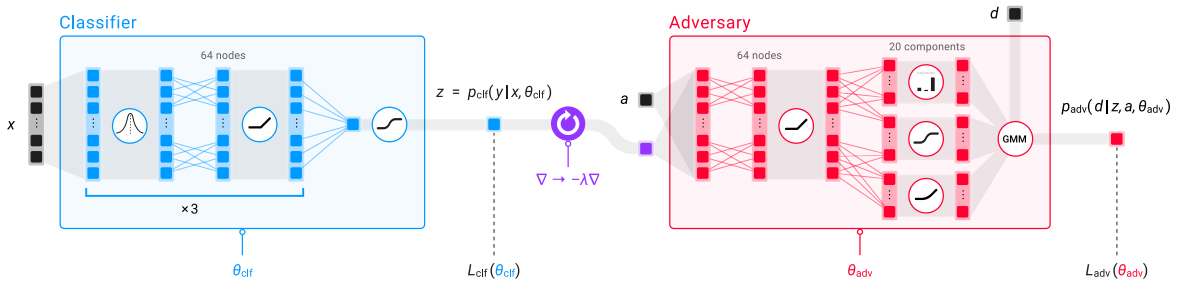


Figure 3.6: Schematic architecture of the networks used for the adversarial tagger. The classifier output is connected to the adversarial input via a gradient reversal layer ($\nabla \to -\lambda\nabla$). A Gaussian Mixture Model (GMM) is applied to model the distribution of the decorrelation parameter d (here: jet mass). Auxiliary variables a may be used to aid the adversarial network's prediction of d. (Fig. from [95]).

Variable	Description	Reference
D_2, C_2	Energy correlation ratios	[88]
τ_{21}	N-subjettiness	[96]
R_2^{FW}	Fox-Wolfram moment	[97]
$\mathcal{P}$	Planar flow	[98]
a_3	Angularity	[99]
A	Aplanarity	[100]
Z_{cut}, $\sqrt{d_{12}}$	Splitting scales	[101, 102]
$Kt\Delta R$	k_t-subjet ΔR	[103]

Table 3.1: List of substructure variables used in the neural network based tagger training [77].

The training process of the networks is done in three consecutive phases: First, the classifier is trained without considering the adversary. Its output is optimised for the binary prediction as to whether a jet is initiated by a vector boson or not. In the second stage only the adversary is trained while all properties (i.e. weights and biases) of the classifier are held constant. In this way the adversary is optimised to predict a close approximation to the actually reconstructed jet mass. In the last phase classifier and adversary are trained together optimising a linear combination of the two loss functions at the same time. For this purpose a gradient reversal layer is introduced between classifier and adversary. During backpropagation it multiplies the gradients propagated from the adversary into the classifier by the factor $-\lambda$, effectively reversing the optimisation goal. Here $\lambda > 0$ is a chosen parameter indicating the emphasis on mass-decorrelation. It is only applied for background events.

To understand the utility of this layer it may help to think about the training process of the combined networks as a multidimensional minimisation problem: The weights and biases (each constituting one dimension) of each node in the nets are optimised in such a way that the loss function is minimised. The gradients with respect to the loss function are calculated at each node. The weights and biases are then updated according to their gradients, each descending along the direction of their respective gradient. In this way, during backpropagation, the properties of the classifier are updated in such a way that they are more likely to lead to a better mass estimate in the adversary. However, due to the negative sign in front of the factor λ of the gradient reversal layer, the opposite is the case. Hence, this leads to a decorrelation of the classifier's output with respect to the jet mass. For validation or when the tagger is applied in an actual analysis situation the adversary part is not needed anymore. Then the classifier network constitutes the whole tagger.

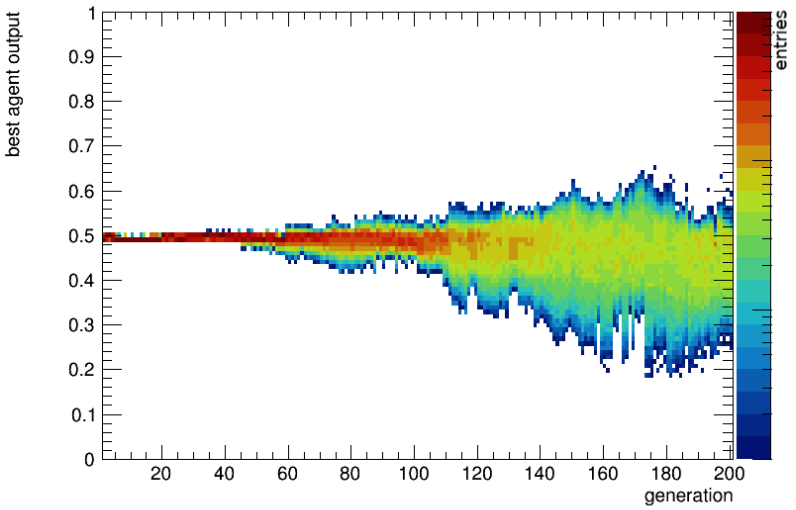
3.2.3 . Genetic Training

An alternative method of tagging, the genetic method, was studied. It combines the tagging and decorrelation tasks into a single task requiring only one optimisation. This is in contrast to the adversarial method described in the previous section which first trained the tagger and then decorrelated it in a later step. The goal of this was to determine if this two-step approach might yield suboptimal results, optimising for a local minimum rather than an even lower global minimum.

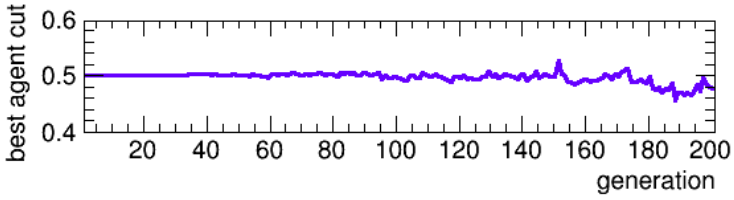
The main difference in the genetic method with respect to the adversarial method is that the training process does not rely on gradient descent. This has the advantage, that the evaluation metric during the training process, i.e. the loss function in the case of the adversarial method, does not have to be differentiable. This leaves more freedom to its definition. Instead of a conventional loss function a score s is calculated which takes into account both the classification power and the mass-decorrelation:

$$s = \frac{(1-\epsilon_{\text{bkg}}) + \beta[(1-J) + (1-b) + (1-\bar{b})]}{1+\beta}\Big|_{\epsilon_{\text{sig}}=50\%} \tag{3.5}$$

The factor β is a measure for the emphasis on mass-decorrelation of the training process and has therefore a similar role to the factor λ in the adversarial method. J constitutes the Jensen–Shannon divergence (JSD) [104] between the jet-mass histogram of the background passing the tagger with respect to the total background. It is a measure for how much two histograms differ bin by bin. b and $\bar{b}$ are the result of a bump hunter algorithm [105] on these distributions either searching for a local excess in the former with respect to the latter distribution (b) or the other way around ($\bar{b}$). All of the three latter variables are encoding the mass-decorrelation. JSD is a good estimator for the global difference between the histograms while the bump-hunter variables are designed in order to prevent local excesses or deficiencies (i.e. bumps) in one histogram with respect to the other. The background efficiency ϵ_{bkg} is a measure for the classification power of the tagger. Since the building of histograms is not a differentiable task, it would not have been possible to use this score function for the adversarial neural networks. All of these metrics are evaluated in the case of a signal efficiency of $\epsilon_{\text{sig}} = 50\%$. The latter is achieved by continuously adapting a cut on the network's output at each training step in order to achieve the wanted working point (see Fig. 3.7).



(a)



(b)

Figure 3.7: Histogram of the network output values (a) of the highest-scoring model in each generation during a training process with the genetic method and a mass-decorrelation factor of $\beta = 10$. The corresponding value to achieve a signal efficiency of $\epsilon_{\text{sig}} = 50\%$ is shown in (b). This cut value is used to calculate all evaluation metrics as shown in Fig. 3.8.

To train the network with this method, first a population of N_{pop} identical networks is created. In the following they are applied to the input data set and the score s is calculated. From there on in each training step, the next generation of N_{pop} nets is created by randomly choosing nets from the previous generation with a higher probability for those with higher score. In between each training step, the weights and biases of all nets are varied by a random value. In this way, after a high enough number of generations, the resulting nets tend to achieve higher scores. In the last step the net with the highest score is chosen as the final one for the tagger. Fig. 3.8 shows the evolution of the score and the evaluation metrics used to calculate it for such a training process. It is apparent that while the signal efficiency remains constant at $\epsilon_{\text{sig}} = 50\%$, background efficiency as well as JSD are decreasing throughout the training process. This results in a growing score.

The bump hunter metrics b and $\bar{b}$ appear to remain largely constant throughout the training after a sharp decrease in the very first few generations.

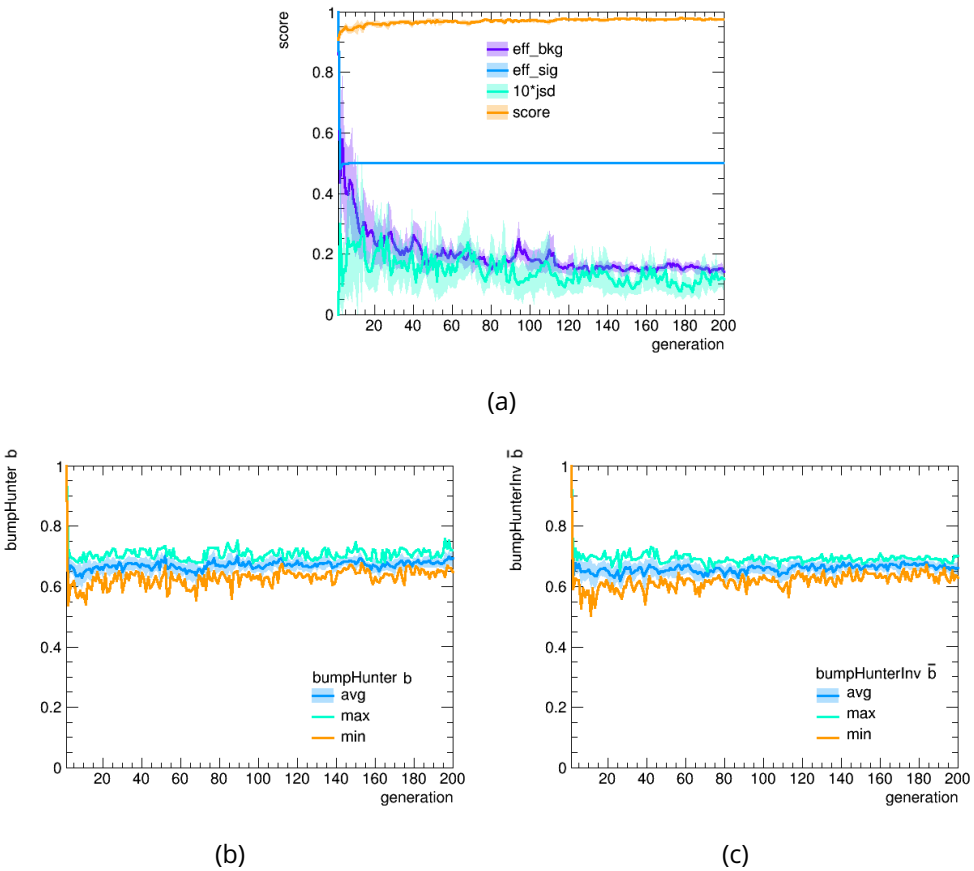
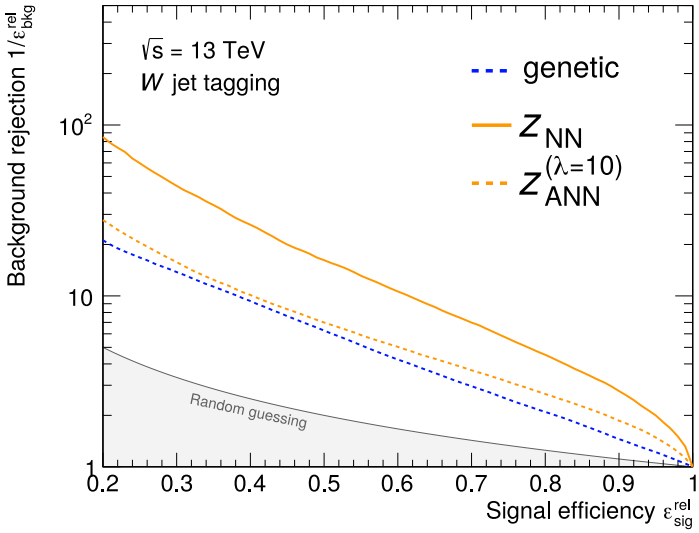


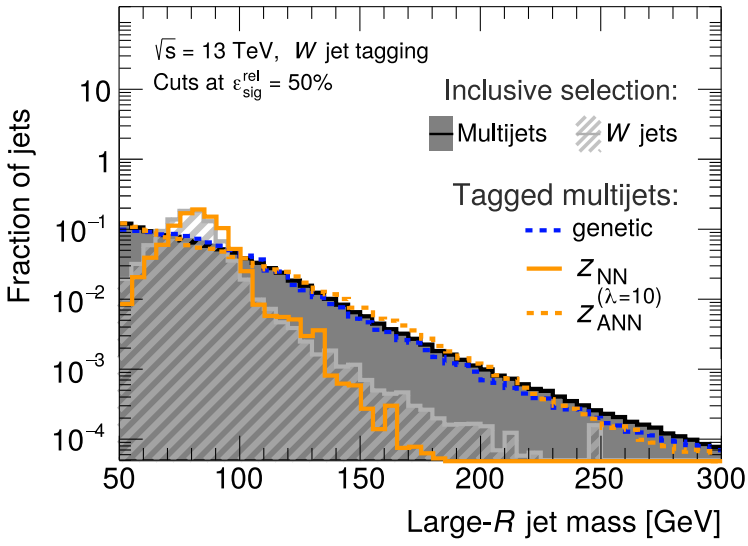
Figure 3.8: (a): Evolution of the score (Eq. 3.5), signal and background efficiencies, and Jensen–Shannon divergence (JSD) of neural nets trained with the genetic method evaluated on a testing data set. (b-c): Bump-hunter b and inverse bump-hunter $\bar{b}$ metric. Solid lines denote the mean of all models in a given generation, transparent bands denote the corresponding standard deviation. For b and $\bar{b}$ the minimum and maximum values per generation are shown as well, indicating the spread.

3.2.4 . Comparison of Tagger Performance

The evaluation of the taggers trained with the various methods described above was done by comparing their results with respect to classification efficiency and mass-decorrelation. The former is evaluated by the receiver operating characteristic (ROC) curve shown in Fig. 3.9(a) while the latter can be accessed by comparing the histograms of jet mass for the background passing the tagger and the total background which is shown in Fig. 3.9(b).



(a)



(b)

Figure 3.9: Receiver Operator Characteristic (ROC) curve (a) and jet mass distribution (b) obtained with the adversarial neural network method (ANN) and the genetic method (gen). Both methods were trained with a mass-decorrelation factor of $\lambda = 10$ in case for the ANN ($Z_{\text{ANN}}^{(\lambda=10)}$) and $\beta = 10$ (gen 10.0). The neural network method without any decorrelation applied, i.e. $\lambda = 0$ is shown for comparison as Z_{NN}. The mass distributions (b) shown as lines in the plots correspond to background passing the corresponding tagger. The solid (shaded) histogram is the true background (signal) without any tagger applied.

The mass-decorrelated version of the adversarial method performs better according to both goals compared to the genetic method. By observing the ROC curve for the non-decorrelated versions of the taggers it becomes apparent that the genetic training method does not achieve the same background rejection as the other method. This may be caused by the significantly less efficient training method. While gradient descent (used in the adversarial method) promises a reasonably quick convergence of the model towards a minimisation of the loss function due to the directed updates along the gradient, the genetic method relies on randomly varying the properties of the model. During the training process it was observed that reasonable results can be achieved with the adversarial method within a few minutes of training on a consumer CPU while training on powerful GPUs for several hours was needed for the genetic method.

The genetic method, as presented here, performs significantly worse with respect to both decorrelation and tagging. Further studies would be needed to obtain comparable results to the adversarial neural methods technique. It was therefore decided to abandon the genetic method for the studies presented in the following sections and instead compare only the adversarial method (defined in Sec. 3.2.2) with the single-variable method using $D_2^{\text{k-nn}}$ (defined in Sec. 3.2.1).

3.2.5 . Tagging with Unified Flow Objects

A new type of jet input object called Unified Flow Objects (UFOs) has recently been developed within ATLAS. It has been shown that jets formed from these constituents exhibit overall better resolution of jet substructure variables compared to jets formed from established jet inputs [94]. This has been achieved by combining several of the already established jet inputs based on the jet's transverse momentum resulting in better performance over the whole p_{T} range.

While jets formed solely from topological clusters of energy depositions in the calorimeters (LCTopo jets) [93] are commonly used in ATLAS analyses, it has been shown that combining the information from the calorimeters with information from the inner tracker can be beneficial. Track Calorimeter Cluster (TCC) jets [106] are composed from the energy information of the calorimeters and the angular information of the tracker. Particle Flow Objects (PFO) combine the expected energy deposition of particles reconstructed from tracks with topological clusters from the calorimeter. While PFO outperform LCTopo jets over the whole p_{T} range with respect to substructure resolution, TCC outperforms PFO in the highly boosted regime but shows less performance in the low p_{T} range [94]. UFO jets combine LCTopo and TCC jets to achieve an optimal performance over the full range.

In the following (Sec. 3.2.6 - 3.2.8), a study on the performance of different vector boson taggers using this novel jet collection is presented. It has been published in the form of a public note [77]. The goal of this study was to determine if the enhanced substructure resolution of UFO jets leads to better performance of jet tagging algorithms. The results of the 3-variable tagger as well as the deep

learning based tagger on UFO jets are compared to results with established jet collections as well as with each other. Additionally a 1-variable tagger, only using the D_2 cut of the 3-variable tagger is studied. The performance of decorrelating these taggers from the jet mass is shown as well. The methods used for this are those described above in Sec. 3.2.1 and 3.2.2, namely an analytical method for the 3-variable tagger and adversarial training for the deep learning based tagger.

3.2.6 . Truth Labelling

A detailed description of the Monte Carlo generated samples for this study can be found in the public note itself [77]. The study was performed with signal samples in which either Z or W bosons decay hadronically and are reconstructed as jets. Background samples contain quark- or gluon-initiated jets.

Truth labelling in this context refers to the way in which jets built from reconstructed Monte Carlo objects, in this case UFOs, have been either labelled as signal or background jets based on information from jets built from stable particles within the Monte Carlo simulation (truth jets). Both UFO and truth jets were built with the anti-k_T algorithm with $R = 1.0$. Constituent Subtraction + SoftKiller (CS+SK) [94] and the Soft-Drop [107] jet grooming algorithm with $Z_{\text{cut}} = 0.1$ and $\beta = 1.0$ were used to remove contributions from pileup and soft QCD from the reconstructed jets. No grooming is applied to truth level jets. The usual Monte Carlo based ATLAS calibration scheme [108] is performed on the reconstructed jets to apply average correction factors for energy and mass scale to particle level. No in-situ calibration based on data was available for UFO jets at the time of performing these studies. Once available, scale factors will be derived for the taggers to finalise the for now purely Monte Carlo based studies presented in the following.

For this study truth jets containing the products of a $W/Z \to qq'$ decay are considered signal while all others are considered background. More specifically, a signal truth jet J has to fulfill the following requirements:

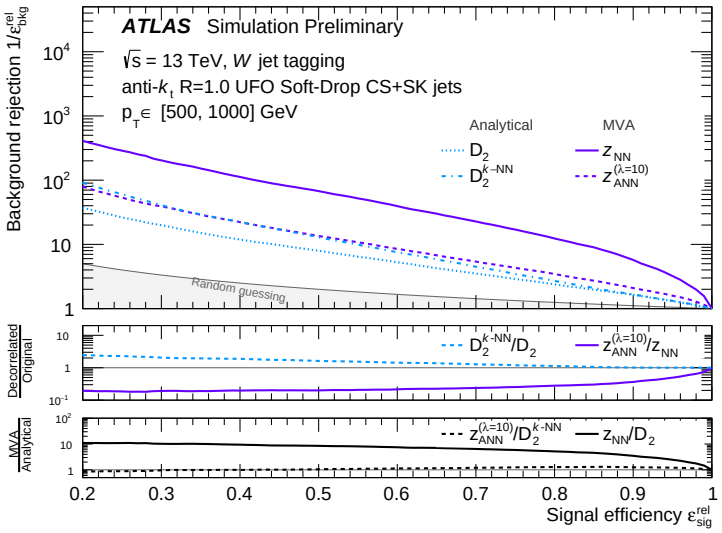
- W or Z boson within $\Delta R < 0.75$
- Jet mass $m_J > 50$ GeV
- Energy scale of the first k_T-declustering [101] $\sqrt{d_{12}} > 55.25\text{exp}\left(\frac{-2.34\times10^{-3}}{\text{GeV}} p_{\text{T}}^{J}\right)$ GeV
- No B hadrons associated[3] to W jets (to reduce top contamination)

Truth jets are matched to UFO jets with $\Delta R < 0.75$ and are labelled as signal/background accordingly. With this labelling scheme, a fraction of 98% (96%) of the UFO jets matched to a particle-level W (Z) contained both particle-level daughter quarks from the simulated $W/Z \to qq'$ decay.

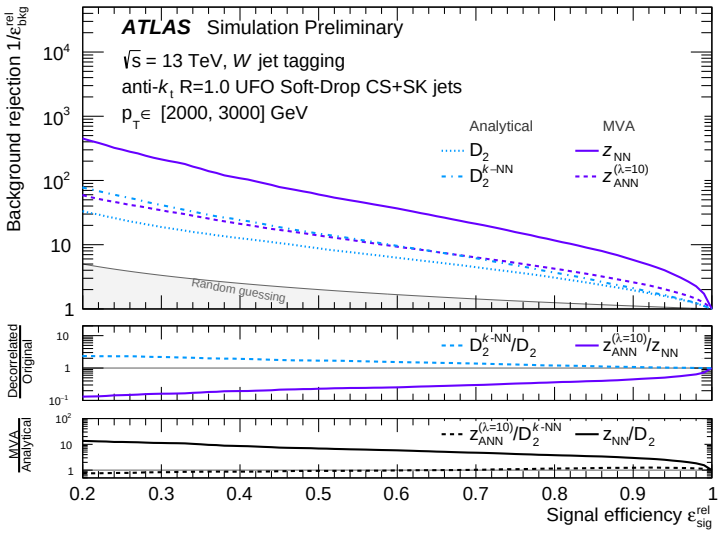
[3] See [89, 90] for a definition of ghost association which is used here.

3.2.7 . Tagger Comparison

The performance of the different taggers was evaluated with respect to classification power and decorrelation to the jet mass. The former criterion is characterised by ROC curves which can be seen in Fig. 3.10.

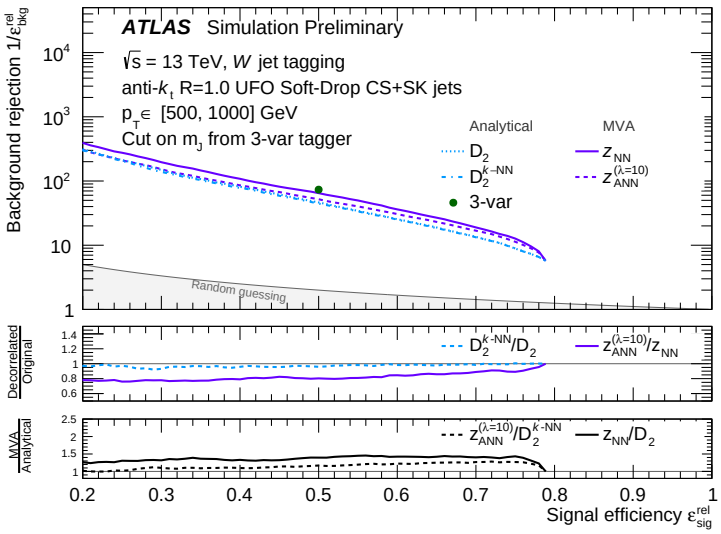


(a)

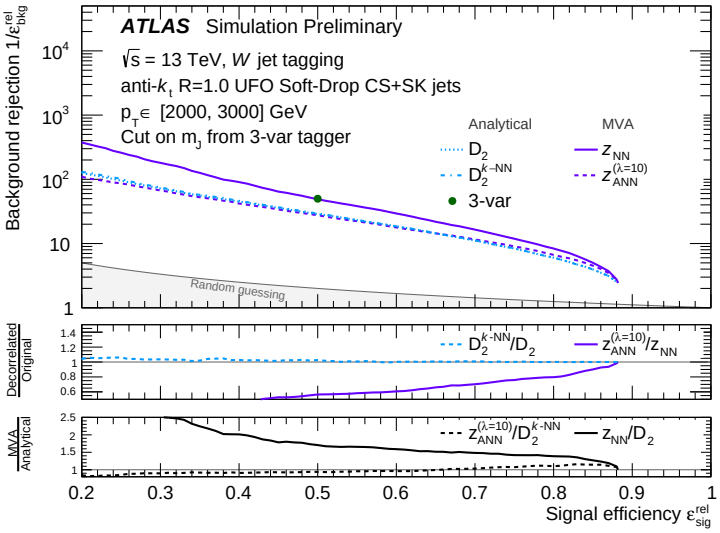


(b)

Figure 3.10: ROC curves for the 1-variable (D_2), and the NN (Z_{NN}) tagger as well as their decorrelated versions $D_2^{k-\text{nn}}$, and $Z_{\text{ANN}}^{(\lambda=10)}$ in the low (a) and high (b) p_{T} range [77].



(a)



(b)

Figure 3.11: ROC curves for the 1-variable (D_2), and the NN (Z_{NN}) tagger as well as their decorrelated versions $D_2^{k-\text{nn}}$, and $Z_{\text{ANN}}^{(\lambda=10)}$ in the low (a) and high (b) p_{T} range with a jet-mass window cut corresponding to the 3-variable taggers mass criterion. A single point for the 3-variable tagger at a signal efficiency of $\epsilon_{\text{sig}}^{\text{rel}} = 50\%$ is shown for comparison [77].

For both low (Fig. 3.10(a)) and high p_{T} jets (Fig. 3.10(b)), it is apparent that a cut on the MVA neural network (NN) score Z_{NN} performs significantly better than a cut on the single variable D_2. However the decorrelated versions

of these taggers, $D_2^{k-\text{nn}}$ and $Z_{\text{ANN}}^{(\lambda=10)}$ perform similarly well. This suggests that a large part of the increase in performance of the neural network tagger compared to the single variable tagger is caused by it learning to reconstruct the jet mass from its inputs. The neural network based tagger loses a significant amount of classification strength after decorrelation. Counterintuitively however $D_2^{k-\text{NN}}$, the decorrelated version of D_2, performs better than D_2 itself. This is assumed to be the case because of an improved rejection of mass-sideband events especially. Fig. 3.11(a) and 3.11(b) shows the ROC curves in the mass window as defined by the 3-variable tagger (Fig. 3.3(a)). This enables comparison of the taggers under study with the 3-variable tagger which is widely used within ATLAS analyses. The NN tagger shows similar performance to the 3-variable tagger while outperforming the others. It is assumed that the 3-variable tagger profits from its cut on the number of tracks n_{trk} associated to the jet which is not part of any of the taggers under study. A further study, including n_{trk} as an additional feature for the NN tagger would show if it would then be able to achieve better performance than the 3-variable tagger.

The dependence of background rejection to the transverse momentum of the jet is shown in Fig. 3.12.

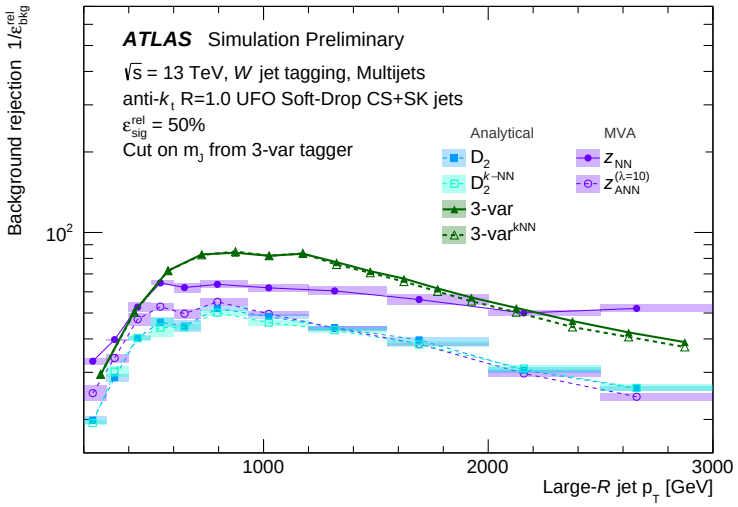
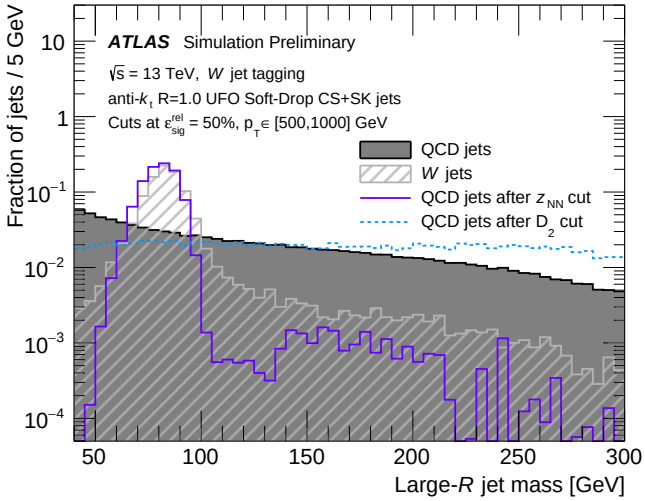
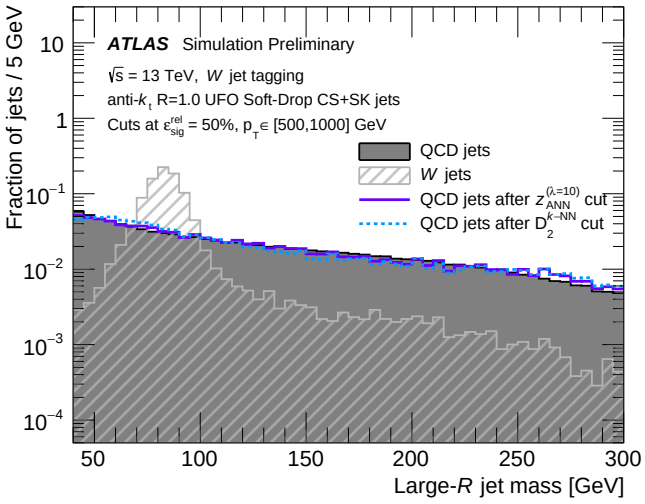


Figure 3.12: Background rejection $1/\epsilon_{\text{bkg}}^{\text{eff}}$ at a fixed signal efficiency of $\epsilon_{\text{sig}}^{\text{eff}} = 50\%$ as a function of transverse momentum for the 1-variable (D_2), and the NN (Z_{NN}) tagger as well as their decorrelated versions $D_2^{k-\text{nn}}$ and $Z_{\text{ANN}}^{(\lambda=10)}$, as well as the 3-variable tagger (3-var) and a variation of this tagger (3-var$^{\text{kNN}}$) in which D_2 has been replaced by $D_2^{k-\text{nn}}$ [77].

It can be seen that the 3-variable tagger outperforms the NN tagger in the low- and medium-p_{T} range while the NN tagger shows higher background rejection in the highly boosted regime. Since this only relies on one point in the graph, further studies with higher statistics in this region would have to be made to obtain a definite result. It should be noted that the ANN and the 1-variable $D_2^{k-\mathrm{nn}}$ tagger show consistently similar results over the whole spectrum.

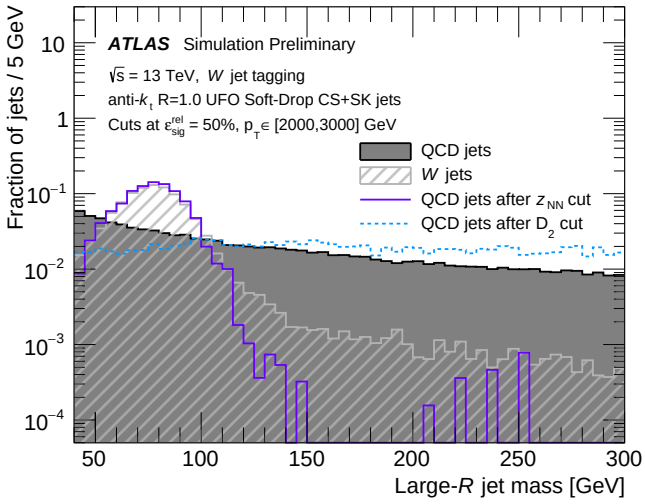


(a) nominal

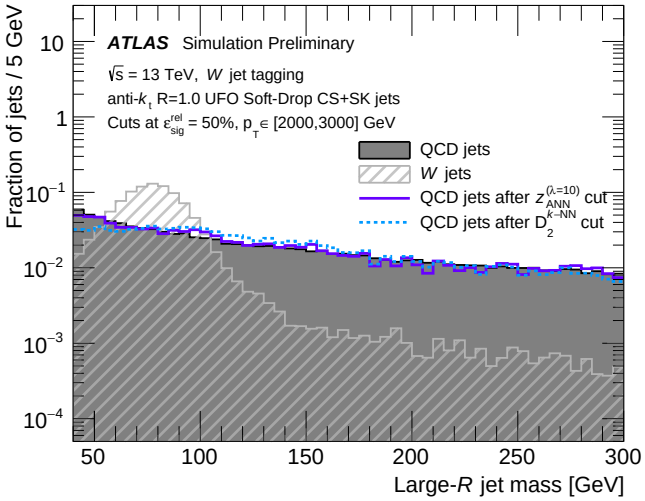


(b) decorrelated

Figure 3.13: Signal and background jet mass distributions before and after a cut corresponding to a signal efficiency of $\epsilon_{\mathrm{sig}}^{\mathrm{rel}} = 50\%$. Shown are the 1-variable (D_2), and the NN (Z_{NN}) tagger (a) as well as their decorrelated versions $D_2^{k-\mathrm{nn}}$, and $Z_{\mathrm{ANN}}^{(\lambda=10)}$ (b) [77].



(a) nominal



(b) decorrelated

Figure 3.14: Mass distributions as shown in Fig. 3.13 but in a high p_{T} region ($2 - 3$ TeV) [77].

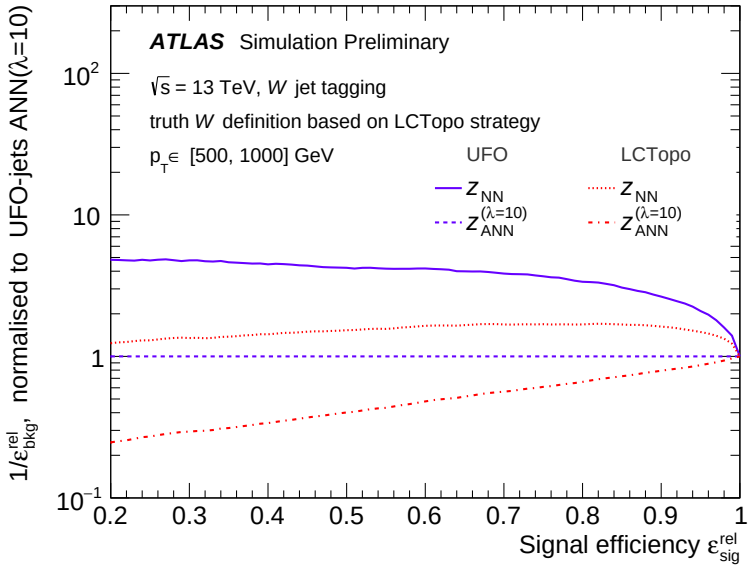
The correlation of the tagger score to the jet mass was evaluated by comparing the jet mass distributions before the application of the tagger with the corresponding distributions after a cut on the tagger score. This can be seen in Fig. 3.13. A cut on the tagger score corresponding to a signal efficiency of $\epsilon_{\mathrm{sig}}^{\mathrm{rel}} = 50\%$ was applied. Fig. 3.13(a) shows these distributions for the D_2 and NN tagger, i.e. without any decorrelation method applied. It is apparent that the background jet mass distributions after the application of the tagger deviate from the distributions

before the application of the tagger. This is especially visible for the NN tagger where the shape of the background distribution after the cut on the tagger closely resembles the corresponding distribution for signal jets. This confirms the assumption that the NN tagger learns to derive the jet mass from its inputs, leading to a strong correlation of its score to the mass. Fig. 3.13(b) shows these type of distributions for the $D_2^{k-\mathsf{nn}}$ and the ANN tagger, i.e. the decorrelated taggers. It is apparent that the background distributions with the tagger applied resemble the corresponding distributions before the tagger has been applied. No significant shaping of the background according to the signal can be seen. While this is generally the case for both decorrelated taggers, it should be noted that especially in the high-p_{T} range (Fig. 3.14(b)) for the $D_2^{k-\mathsf{nn}}$ tagger there is still some deviation visible for low masses (up to ≈ 70 GeV)

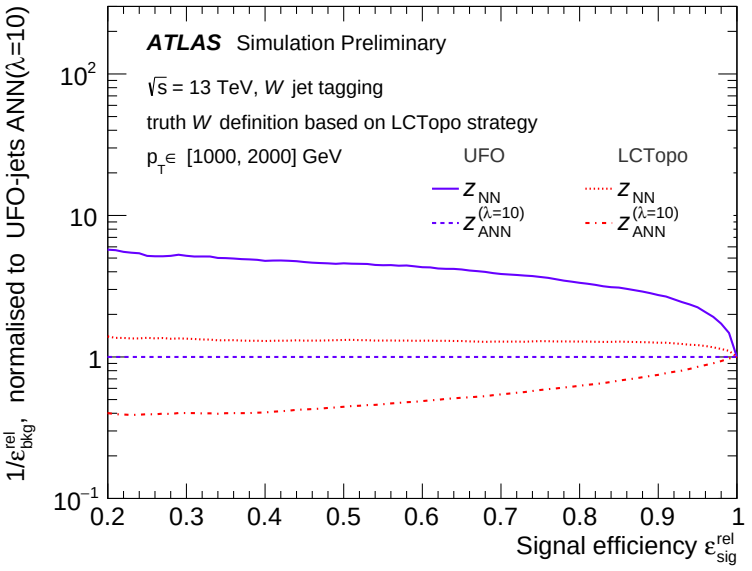
3.2.8 . Comparison with UFO and LCTopo Jets

Fig. 3.15 compares the performance of the NN and ANN taggers on UFO jets with their performance on LCTopo jets. For the latter, the models have been trained in an earlier study [109]. The ATLAS recommendations for truth labelling have changed since the study performed in [109]. As a consequence the truth labelling described in Sec. 3.2.6 diverges. To ensure a fair comparison the event selection and truth labelling strategy of Ref. [109] was used for both UFO and LCTopo jets for this set of plots. Note that for all other plots shown earlier in this thesis, the new truth labelling of Sec. 3.2.6 has been used.

It is apparent that the improved resolution of substructure variables [94] of UFO jets compared to LCTopo jets translates into better tagging performance. Especially in the high-p_{T} range (Fig. 3.15(b)) where better resolution is more important this leads to a significantly better background rejection with UFO jets compared to LCTopo jets. Even in the low-p_{T} range (Fig. 3.15(a)) there is still an improvement by a factor of approximately two (at $\epsilon_{\mathsf{sig}}^{\mathsf{rel}} \approx 80\%$) visible. This is the case for both NN and its decorrelated version ANN.



(a)



(b)

Figure 3.15: ROC curves for the neural network (Z_{NN}) tagger and its decorrelated version, the adversarial neural network tagger $Z_{\text{ANN}}^{(\lambda=10)}$ in the low [left] and high [right] p_{T} range acting on UFO and LCTopo jets. The curves are normalised to the ANN UFO tagger [77].

3.2.9 . Conclusion

Various taggers have been studied with respect to tagging efficiency and mass decorrelation in a vector boson tagging task.

In a first step, the adversarial neural networks (ANN) decorrelation method for neural network (NN) taggers has been compared to a genetic tagger including mass decorrelation within its score function. Extensive further studies would have been needed to obtain similar results with the latter method compared to the former. It is assumed that this stems among others from the superior training method of the ANN, making use of gradient descent, leading to more robustness with respect to starting parameters as well as significantly faster convergence of the optimisation task. As a result, only the ANN tagger was compared to the established 3-variable tagger in the following study on UFO jets.

The background rejection of the NN and ANN tagger on UFO jets improved by a factor of approximately 2-3 with respect to LCTopo jets. ANN, the mass-decorrelated version of the NN tagger showed a similar performance to the mass-decorrelated single-variable tagger D_2^{k-nn}, suggesting the NN's advantage without decorrelation stems mostly from it learning to construct the jet mass from its inputs. The 3-variable tagger showed a comparable background rejection to the NN tagger despite using a much smaller set of input parameters and only accounting for simple cuts, not being able to take higher order correlations into account. It is assumed that this may be the case due to its use of the number of tracks associated with the jet. This variable is known to be a strong discriminant between quark or gluon initiated jets. The NN lacks this kind of input. A follow-up study should add this to the list of NN inputs to make a fair comparison.

A moderately improved performance of the NN over the 3-variable tagger has been observed in the highly boosted regime. A targeted study with higher statistics in this regime would be needed to make definite conclusions. It should also be noted that it may be possible to adjust the 3-variable tagger for this regime specifically if the need arises.

The significantly improved separation power of the NN and ANN tagger on UFO jets compared with LCTopo jets suggests that it may be advisable to use this new type of jet definition in analyses that make extensive use of vector boson tagging. This should therefore be considered for an eventual follow-up study of the main analysis presented in this thesis, the search for semileptonic vector boson scattering. The lack of improvement of the NN tagger (without n_{trk} as feature) compared to the established 3-variable tagger suggest that in its current form however, the latter should be preferred over the former.

4 - Semileptonic Vector Boson Scattering

The following section presents a search for an electroweak (EW) diboson system (i.e. WW, ZZ, or WZ) in association with a high-mass dijet system (jj) in the semileptonic 0-lepton channel. This process is sensitive to vector boson scattering (VBS) and the event selection is defined in such a way as to enhance the contribution from VBS. It is the main analysis of this thesis. The full Run II (2015-2018) dataset of ATLAS is used, corresponding to an integrated luminosity of 139 fb^{-1} recorded in pp collisions with a centre of mass energy of $\sqrt{s} = 13$ TeV. A previous iteration of this analysis channel using a subset of this, corresponding to 35.5 fb^{-1}, achieved an observed (expected) significance of 2.7 σ $(2.5$ $\sigma)$ for all (0,1,2-lepton) channels combined and 1.43 σ $(1.35$ $\sigma)$ in the 0-lepton channel alone [1]. Aspects of the following 0-lepton channel analysis are based on this previous iteration. Improvements with respect to several analysis methods were made to achieve a higher sensitivity. This analysis will eventually be combined with similar analyses of the other two channels that are in preparation in parallel.

Sec. 4.1 describes the experimental signature for the process under study. The following Sec. 4.2 lists the Monte Carlo generated samples as well as the data that was used for the analysis. The reconstructed objects used throughout it are defined in Sec. 4.3. Sec. 4.4 defines the event selection and analysis regions. Sec. 4.5 presents a reweighting procedure that was performed in order to compensate for poorly modelled background. The following Sec. 4.6 presents the multivariate analysis (MVA) strategy defining several final discriminants and compares them with each other. Besides the nominal method, two alternative MVAs are presented in order to characterise the sensitivity of the MVA with respect to a known mismodelling of jet multiplicity in V+jets background processes. Sec. 4.7 investigates this aspect in detail. The way in which this and other systematic uncertainties are implemented into the analysis is described in Sec. 4.8. Finally Sec. 4.9 presents a statistical interpretation of the analysis in the form of a binned maximum likelihood method and derives values for the signal strength as well as the observed significance with respect to the semileptonic EW $VV + jj$ process under study.

4.1 . Signature of the Process

A diagram at tree level for the process in question (electroweak diboson production in association with a pair of jets) can be seen in Fig. 4.1. The process of physical interest, containing vector boson scattering, is shown on the left. Due to having the same final state in terms of particle content the other two processes can not easily be separated from this one. Only the electroweak processes (left and middle) are considered signal while the strong process (right) is considered to

be background in this analysis. This is because the electroweak processes cannot be separated from each other in a gauge invariant way [110].

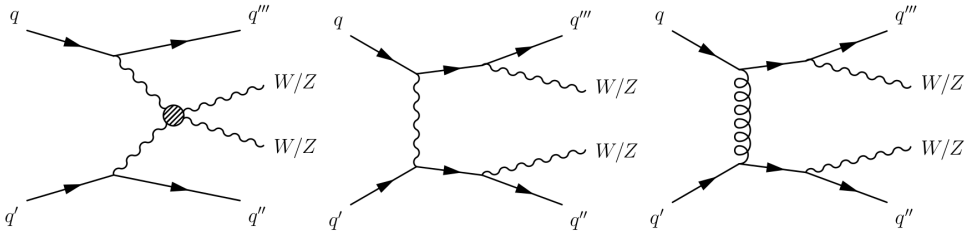


Figure 4.1: Feynman diagram at tree level for the vector boson scattering (VBS) process (left) and the electroweak (middle) and QCD (right) non-VBS production processes with the same final state [1].

The final state in the semileptonic channel under study can be expressed as

$$qq + V_{\text{had}} + V_{\text{lep}} \tag{4.1}$$

where qq are the two quarks stemming from the VBS (or VBS-like) production and V_{had} (V_{lep}) is a hadronically (leptonically) decaying vector boson (i.e. W or Z). These decay products are reconstructed as

$$qq \rightarrow (jj)^{\text{tag}} \tag{4.2}$$

$$V_{\text{had}} \rightarrow (jj)^{\text{sig}} \text{ or } J^{\text{sig}} \tag{4.3}$$

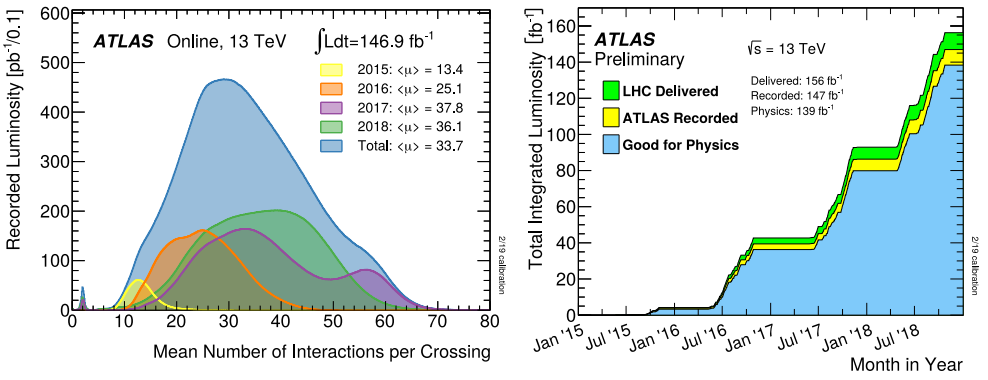
$$V_{\text{lep}} \rightarrow \nu\nu \text{ or } \nu\ell \text{ or } \ell\ell \tag{4.4}$$

where $(jj)^{\text{tag}}$ is the dijet system initiated by the quarks qq from VBS production. In the following these are referred to as tagging jets. The signal dijet system $(jj)^{\text{sig}}$ stems from the hadronically decaying vector boson V_{had}. Alternatively it can be reconstructed as a single large-radius jet indicated as an upper case J^{sig}. The latter case is referred to as the merged selection while the former is called resolved selection. The merged selection targets events in which the hadronicallly decaying vector boson is highly boosted (see Sec. 3.1.3 on boosted decay structures). $Z \rightarrow \nu\nu$, $W \rightarrow \nu\ell$, or $Z \rightarrow \ell\ell$ are the decay products of the leptonically decaying vector boson $V_{\text{lep}} = W/Z$. Depending on the number of charged leptons ℓ, the decay channels are called 0-, 1-, or 2-lepton. In the case of $W \rightarrow \nu\ell$ decay the charged lepton can be any of $\ell = e,\ \mu$, or τ while for $Z \rightarrow \ell\ell$ decays it is either $\ell = e,$ or μ. The analysis presented below focuses on the 0-lepton channel in which no charged lepton is detected. This can either be the case if both leptons are neutrinos ν or if a charged lepton is produced outside of the detector acceptance. This analysis has been developed in parallel with analyses of the other (1-, and 2-lepton) channels. A statistical combination of all channels is foreseen to be published in the near future.

Relevant background processes include the production of a single vector boson in association with jets (W/Z+jets), single top (t) and top pair ($t\bar{t}$) production, QCD diboson production, as well as QCD multijet production. For the 0-lepton channel all of these backgrounds play a significant role with the largest contribution from V+jets and top pair production. While in the 1-lepton channel W+jets and in the 2-lepton channel Z+jets dominates, both backgrounds have comparable yields in the 0-lepton selection.

4.2 . Data and Monte Carlo Samples

The Monte Carlo (MC) generated samples used for the modelling of signal and background processes have been centrally produced within the ATLAS collaboration. They are scaled to correspond to the integrated luminosity of 139 fb^{-1} of data recorded by the ATLAS detector in the full Run 2 during the years 2015-2018. Pileup effects are modelled by overlaying additional pp collisions simulated with Pythia 8.186 [111] in each MC sample with a mean number of interactions per bunch crossing $\langle\mu\rangle$ corresponding to the measured value at that time (see Fig. 4.2). A full simulation of the ATLAS detector [112] using the GEANT4 simulation toolkit [86] has been used in each instance.



(a) interactions per bunch crossing (b) cumulative luminosity

Figure 4.2: (a): Mean number of interactions per bunch crossing $\langle\mu\rangle$ for the 2015-2018 pp collision data at $\sqrt{s} = 13$ TeV centre-of-mass energy in ATLAS. (b): Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during that time period. Figures are taken from Ref. [113].

The signal sample, modelling semileptonic EW $VV + jj$ processes, has been generated using MadGraph5_aMC@NLO v2.3.3 [114] with Pythia 8 [115] for fragmentation. The NNPDF30LO PDF set is used [116]. Two on-shell vector bosons ($V = W$ or Z) are required where one decays into a pair of leptons (V_{lep}) and the

other one decays hadronically (V_{had}), i.e. into a pair of quarks. All purely electroweak, i.e. $\mathcal{O}(\alpha_{EW}^6)$, diagrams are included. The cross-sections as well as the number of generated events for each of the three Monte Carlo campaigns (mc16-a,d, and e corresponding to data taking campaigns of 2015-2016, 2017, and 2018) are listed in Tab. 4.1.

process	events - mc16a	events - mc16d	events - mc16e	cross-section [pb]
$W(\ell\nu)W(qq')jj$, b-veto	1958000	2296000	3320000	1.9994
$W(\ell\nu)W(qq')jj$, b-filter	1996000	2400000	3388000	1.9777
$W(\ell\nu)Z(qq)jj$	1994000	2394000	3392000	0.2571
$Z(\nu\nu)W(qq')jj$	1986000	2394000	3356000	0.15532
$Z(\ell\ell)W(qq')jj$	1996000	2374000	3390000	0.045609
$Z(\nu\nu)Z(qq)jj$	1998000	2396000	3390000	0.032238
$Z(\ell\ell)Z(qq)jj$	1990000	2388000	3396000	0.0096553

Table 4.1: List of signal samples used in the analysis. $b-$veto/filter correspond to samples without/enriched in decays containing b-quarks.

The $W/Z+$jets background is modelled using the Sherpa 2.2.1 generator [117] at next-to-leading order (NLO) for 2 partons and leading order (LO) for 4 partons with the Comix [118] and Open Loops [119] matrix element generators, Sherpa parton shower [120] using the ME+PS@NLO [121] prescription and the NNPDF3.0NNLO PDF set [122]. The resulting events were normalised to the next-to-next-to-leading order (NNLO) cross-sections. The QCD diboson background has also been generated with Sherpa 2.2.1.

The top ($t\bar{t}$ and single t) backgrounds are generated with Powheg-Box [123] using the NNPDF3.0NLO PDF set [122] for matrix element calculation. The top mass is set to 172.5 GeV. Top quark decay is simulated through MadSpin [124], assuring that spin correlations are preserved. Parton shower, fragmentation, and underlying event were simulated using Pythia8.230 [111] with the A14 tune set [125]. The cross section of the simulated top processes are known to NNLO in QCD including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [126, 127, 128, 129]. To ensure a good data/MC agreement especially in the high-p_{T} region, the parameter H_{damp} in the Powheg simulation is set to $1.5\ m_t$ in accordance with Ref. [130].

In addition to the nominal background sample described above, alternative samples produced with alternative Monte Carlo generators are used in various aspects of the analysis. This includes the implementation of systematic uncertainties on the modelling of the shape of discriminants with respect to such backgrounds. It also includes the derivation of priors for uncertainties on the normalisation based on differences in event yields. Both aspects are described in Sec. 4.8. These are then utilised in the final statistical interpretation of the analysis. The alternative generators used are: MadGraph+Pythia for $W+$jets and $Z+$jets, Powheg for diboson, and Powheg+Herwig for $t\bar{t}$ events.

The data used throughout the analysis have been collected with the ATLAS detector in pp collisions at a centre of mass energy of $\sqrt{s} = 13$ TeV between the years 2015 and 2018. The corresponding integrated luminosity $\mathcal{L} = 139$ fb^{-1} has an experimental uncertainty of 1.7% obtained from primary-luminosity measurements with the LUCID-2 detector [131] . Tab. 4.2 summarises the integrated luminosity recorded in the individual years of this campaign.

year	$\mathcal{L}$ [fb^{-1}]
2015	3.21
2016	32.88
2017	44.31
2018	58.45
total	139.0

Table 4.2: Integrated luminosity used in this analysis.

4.3 . Object Definition

The object definition of the analysis has the goal to define objects that encapsulate the typical final state structure of the semileptonic EW $VV + jj$ signal process and to aid with separating it from the various background processes. All objects are defined in the same way on all MC samples and on data. They are summarised below. Fig. 4.3 shows a diagram of a signal VBS process in which the main objects are highlighted indicating from which particles in the process they are supposed to originate from.

Small radius jets j are reconstructed from EMPFlow [132] constituents using the anti-k_{T} algorithm [84] (see Sec. 3.1.2) with a radius parameter of $R = 0.4$. EMPFlow uses particle flow to combine information from calorimeter and tracking system of the detector. In the analysis they are either used as tagging j^{tag} or signal jets j^{sig}.

Tagging jets j^{tag} are defined with the goal of encapsulating the pair of quarks qq which are characteristic for VBS-like final structures. Because of their origin from the scattering process they tend to have only a small change in direction with respect to the initial momenta of the partons in the beams. Therefore they are usually found in the forward detector regions in opposite hemispheres leading to a large invariant mass of the combined system. This is reflected in their selection criteria. They are selected as the two small radius ($R = 0.4$) jets up to $|\eta| < 4.5$ with the highest mass of the combined (sum of four momenta) system $m(jj)$. Both jets have to fulfill the requirement $p_{\mathrm{T}} > 20$ GeV. Note that in the event selection (Sec. 4.4) a tighter requirement of $p_{\mathrm{T}} > 30$ GeV will be applied. They are required not to be tagged as $b-$jets by the recurrent neural network based DL1r $b-$tagger [133] at a working point of 70%. The jet vertex tagger JVT [134]

is applied at the medium working point in order to mitigate effects from pileup. Additionally, both jets are required to pass the forward jet vertex tagger fJVT [135] at the loose working point which is especially optimised for the very-forward region that is characteristic for these jets.

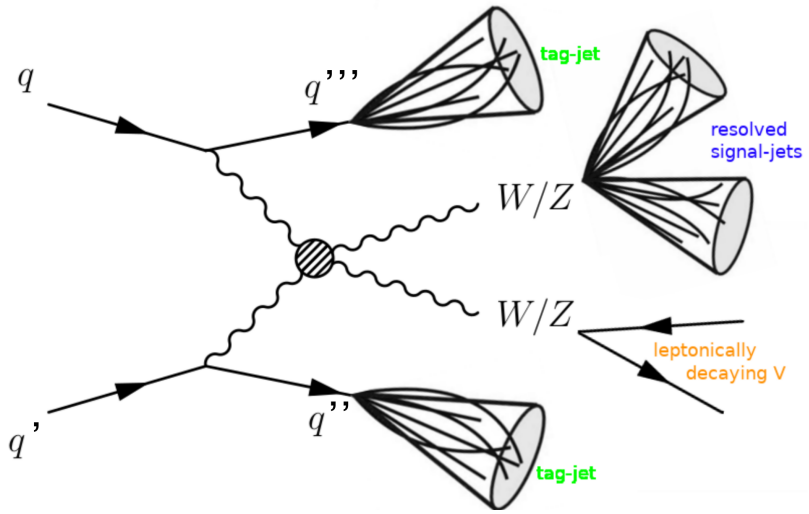


Figure 4.3: Schematic illustration of a VBS process. Reconstructed objects like tagging jets j^{tag}, signal jets $(jj)^{\text{sig}}$ from the hadronically decaying vector boson V_{had} and the leptonically decaying vector boson V_{lep} which, in the zero-lepton channel, is reconstructed from missing transverse momentum $E_{\text{T}}^{\text{miss}}$, are indicated. In the merged regime (not shown), the resolved signal jet pair $(jj)^{\text{sig}}$ is replaced by a single merged signal jet J^{sig}.

Signal jets j^{sig} are defined with the goal of capturing the decay structure of the hadronically decaying vector boson V_{had} into a pair of quarks. They are required to have a transverse momentum of $p_{\text{T}}(j^{\text{sig}}) > 20$ GeV and to be within $|\eta| < 2.5$. They are selected to be the two jets with the highest transverse momenta p_{T} except for the tagging jets which have been selected earlier. Similarly to the tagging jets, the JVT at the medium working point is used but neither $b-$tagging nor the fJVT are applied.

Alternatively to the reconstruction with a pair of small radius jets $(jj)^{\text{sig}}$ the decay of V_{had} can also be parametrised by a single large radius jet J^{sig}. In the context of this analysis the former is called the resolved and the latter the merged regime. Which regime is applied and therefore which of these objects is used is decided by the event selection (Sec. 4.4). Large radius jets J are reconstructed from clusters of topologically connected calorimeter cell signals using the local hadronic cell weighting scheme (called LCTopo [93]) and the anti-k_{T} algorithm [84] with a radius parameter of $R = 1.0$. Trimming [136] is applied to reduce

effects from pileup and soft radiation. The large radius jet J with the highest transverse momentum p_{T} is considered to be the signal jet J^{sig}.

The effect of particles that cannot directly be detected in any of the subdetectors is estimated in the form of missing transverse momentum. In the 0-lepton channel, which is the focus of this analysis, this is the case for the leptonically decaying vector boson V_{lep} decaying either into neutrinos ν or into charged leptons $\ell = e, \mu, \tau$ outside of the detector acceptance. Two separate definitions for this variable are used: $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss,track}}$. The missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined by the negative vector sum of the transverse momenta of hard objects like leptons, jets, photons, electrons, muons, and τ leptons in combination with a soft term arising from tracks which are associated to the primary interaction vertex but not included in any of the hard objects. The track-based transverse momentum $E_{\mathrm{T}}^{\mathrm{miss,track}}$ is derived from reconstructed tracks in the inner detector (ID) only. $E_{\mathrm{T}}^{\mathrm{miss,track}}$ is more robust against pileup effects than $E_{\mathrm{T}}^{\mathrm{miss}}$ but does not include contributions from neutral particles which do not leave tracks in the ID.

In the channel under study (i.e. 0-lepton) charged leptons are defined in order to veto events that fall into the other (1-, and 2-lepton) analysis channels. These veto-leptons ℓ^{veto} can either be electrons e^{veto} or muons μ^{veto}. They are required to have a transverse momentum $p_{\mathrm{T}} > 7$ GeV and pseudorapidity $|\eta < 2.47|$ ($|\eta < 2.5|$) for e^{veto} (μ^{veto}). The loose working point [137] is used for object identification as well as isolation.

With respect to the previous 35.5 fb^{-1} analysis [1], this object definition constitutes a number of improvements: Small radius ($R = 0.4$) jets are constructed from PFlow objects instead of EMTopo, now making use of the combination of information from calorimeters and tracker with particle flow. This has been shown to yield improved pileup stability and resolution [132]. LCTopo [93] constituents are used for large radius ($R = 1.0$) jets instead of track-calo-clusters (TCC) [106]. The DL1r tagger is used for the $b-$tagging of the tagging jets $(jj)^{\mathrm{tag}}$. It has been shown in Ref. [138] that, especially in the high-p_{T} region, it outperforms the MV2 algorithm which was used in the previous analysis.

4.4 . Event Selection

An event selection with the goal of retaining a large number of signal events while reducing the number of background events has been applied. Its form and implementation are motivated by the selection which has been applied in the previous analysis [1].

Several event-cleaning criteria are applied: Events with bad status or incomplete information from one of the subdetectors as well as events with noise bursts in the liquid argon calorimeter are excluded. Jet cleaning removes events with jets built from noisy calorimeter cells or non-collision backgrounds. Events are required to have a primary vertex with at least two associated tracks. The primary vertex

is defined as the vertex with the largest sum $\sum p^2_{\mathsf{T},\,\mathsf{track}}$ of transverse momenta squared of all tracks with $p_{\mathsf{T},\,\mathsf{track}} > 0.5$ GeV associated to it.

Before the main analysis selection, a loose selection has been applied to reduce the file size and to ensure orthogonality between analysis channels (0-, 1-, and 2-lepton). For the 0-lepton channel presented in this thesis this comprises the $E_{\mathsf{T}}^{\mathsf{miss}}$ trigger, a transverse momentum minimum of $E_{\mathsf{T}}^{\mathsf{miss}} > 140$ GeV, a lepton veto requiring no isolated leptons ℓ^{veto}, as well as at least two small- or one large-radius jet in the event. The lepton veto makes the 0-lepton channel orthogonal to the two other analysis channels.

The actual analysis work and optimisation was performed after the event-cleaning and channel selections have already been applied. Tab. 4.3 summarises the individual cuts for the main analysis selection.

merged regime	**resolved regime**
preselection	
$\mathsf{N}(j^{\mathsf{tag}}) \geqq 2$	
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{lead}}) > 30$ GeV and $p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{sublead}}) > 30$ GeV	
$E_{\mathsf{T}}^{\mathsf{miss}} > 200$ GeV	
$E_{\mathsf{T}}^{\mathsf{miss,track}} > 50$ GeV	
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, E_{\mathsf{T}}^{\mathsf{miss,track}})\rvert/\pi < \frac{1}{2}$	
anti-QCD	
$\mathsf{min}\{\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, j_i^{\mathsf{any}})\rvert\}_i/\pi > \frac{1}{6}$	
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, J^{\mathsf{sig}})\rvert/\pi > \frac{1}{9}$	$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, (jj)^{\mathsf{sig}})\rvert/\pi > \frac{1}{9}$
signal regions	
$m(j^{\mathsf{tag}}_{\mathsf{lead}} + j^{\mathsf{tag}}_{\mathsf{sublead}}) > 400$ GeV	
$m(J) > 50$ GeV $\mathsf{tagger}_V^{\epsilon=50\%}$ || $\mathsf{tagger}_V^{\epsilon=80\%}$ **merged HP SR** || **merged LP SR**	64 GeV $< m(jj) < 106$ GeV $m(jjj) > 220$ GeV **resolved SR**

Table 4.3: Cuts from preselection to signal regions in the merged and resolved analysis regime.

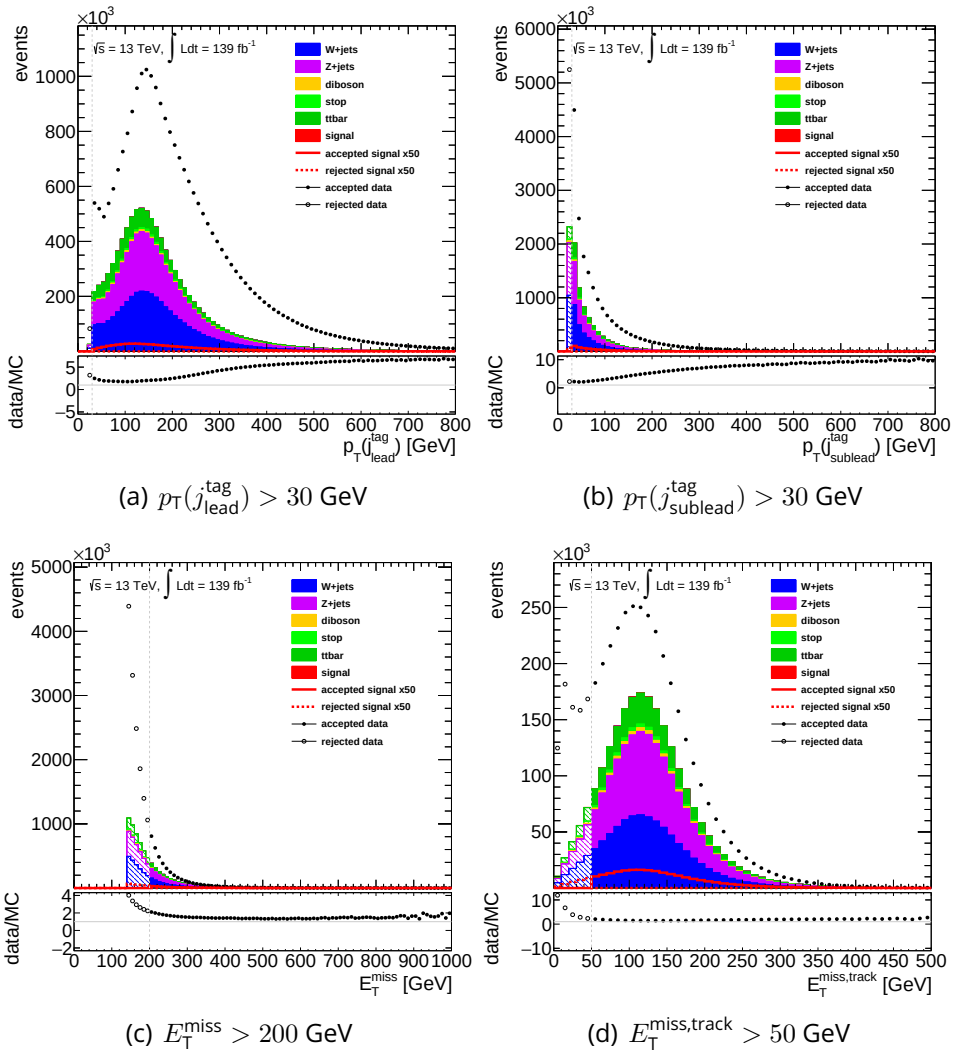
The event selection can be divided into three sections: Preselection, anti-QCD, and signal-region cuts. This results in three signal regions (SR): one for the resolved regime (resolved SR) and two for the merged regime. The latter is split in two by different working point requirements on the vector boson tagger that is applied as the last cut in this regime: high purity (merged HP SR), and low purity (merged LP SR). The individual cuts are described in the order of their application in the following.

The preselection cuts select events based on the expected typical final-state structure of the signal process. The typical structure of vector boson scattering yields a pair of tagging jets in the forward detector region, each with high transverse momentum. The first cut requires that a pair $(jj)^{\text{tag}}$ is found that fulfills the tagging jet requirements specified in the object definition (Sec. 4.3). They are each required to have a transverse momentum $p_{\text{T}}(j^{\text{tag}}) > 30$ GeV. The effect of this cut can be seen in Fig. 4.4 a and b. Note the large discrepancy between Monte Carlo and data in these figures. This discrepancy will be successively reduced throughout the following selection cuts, especially from the anti-QCD selection (described further below) which excludes QCD background in data which is not simulated in Monte Carlo. It is evident that, because of the p_{T} requirement in the tagging jet object definition, this additional p_{T} requirement only excludes a small number of events.

In the 0-lepton channel, a high amount of missing transverse momentum is expected stemming from the two neutrinos (or charged leptons outside of the detector acceptance) in the final state. Cuts on both $E_{\text{T}}^{\text{miss}}$ and $E_{\text{T}}^{\text{miss,track}}$ are applied. Fig. 4.4 c and d show that due to their different definitions both types of missing transverse momentum exclude a different phase space of events. The cut on $E_{\text{T}}^{\text{miss}}$ excludes more than two thirds of the background (see later cut flow table Tab. 4.5 a) while the subsequent cut on $E_{\text{T}}^{\text{miss,track}}$ excludes a smaller number of events. Since $E_{\text{T}}^{\text{miss,track}}$ is reconstructed from charged particles' tracks in the inner detector only, it is less susceptible to pileup. The goal of this selection is to exclude background events that, because of pileup effects, have a significantly large amount of missing transverse momentum and therefore pass the $E_{\text{T}}^{\text{miss}}$ selection but consequently not the $E_{\text{T}}^{\text{miss,track}}$ selection. Additionally, a cut on the relative directions in the transverse plane of the two definitions of missing transverse momenta is applied (see Fig. 4.5 a). The reasoning behind this is that background events with missing transverse momentum dominated by pileup tend to show different directions in the two variables.

The anti-QCD selection is a measure to reduce the contribution of the QCD multijet background in data. This is necessary because no well-enough modelled Monte Carlo samples of this type are available. It is therefore the only main background in the analysis that is not modelled by Monte Carlo. Hence, it has to be suppressed as much as possible in data in order to arrive at a good description of the final selected data sample by the available Monte Carlo samples. Such background

events typically show large missing transverse momentum (which is necessary to pass the preselection) only if the energy of one jet in the event is significantly mismeasured. In that case the direction of the missing transverse momentum is dominated by the direction of that jet. Hence, a cut on the minimum angular distance ΔR between any $R = 0.4$ jet j^{any} and the missing transverse momentum $E_{\text{T}}^{\text{miss}}$ is applied (see Fig. 4.5 b).



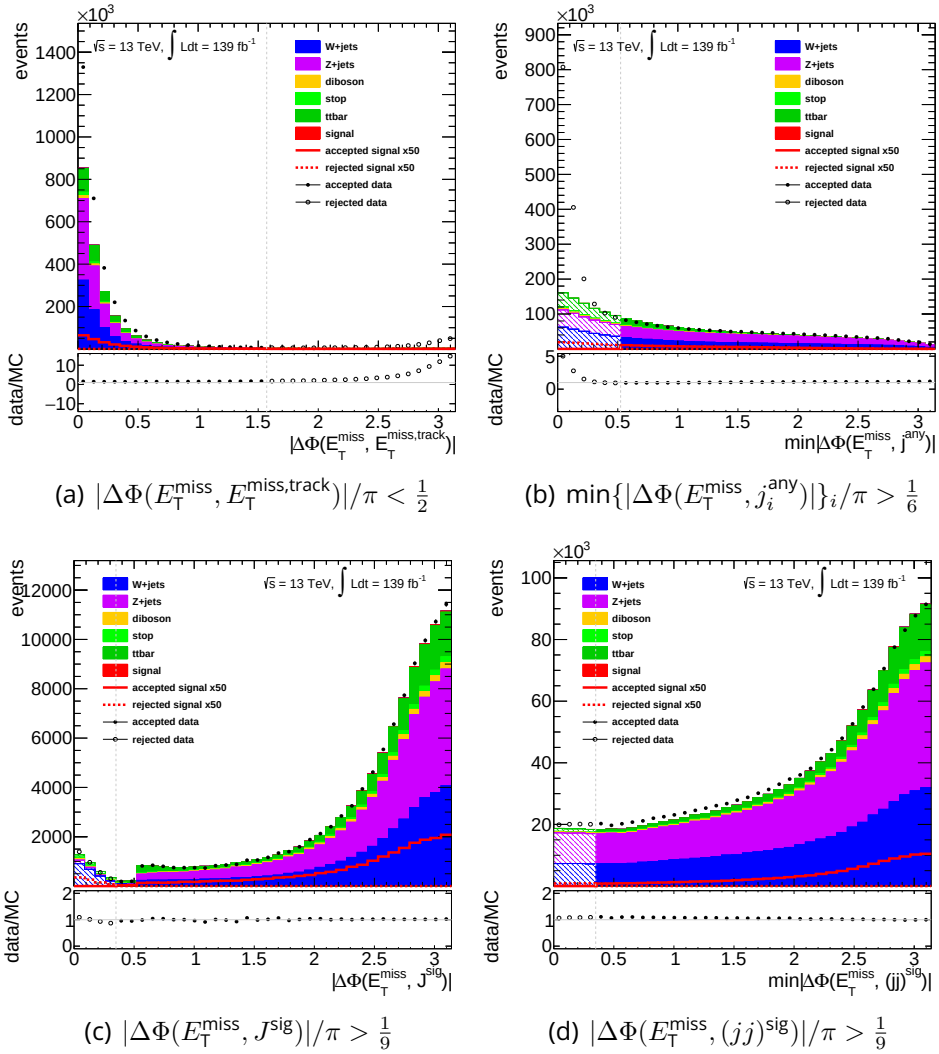
(a) $p_{\text{T}}(j_{\text{lead}}^{\text{tag}}) > 30$ GeV

(b) $p_{\text{T}}(j_{\text{sublead}}^{\text{tag}}) > 30$ GeV

(c) $E_{\text{T}}^{\text{miss}} > 200$ GeV

(d) $E_{\text{T}}^{\text{miss,track}} > 50$ GeV

Figure 4.4: Preselection cuts. Shown are all events passing the event selection defined in Tab. 4.3 up to the cut that is shown. Dashed vertical lines indicate the cut boundaries, the excluded region is indicated by shaded histograms.



(a) $|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}, E_{\mathrm{T}}^{\mathrm{miss,track}})|/\pi < \frac{1}{2}$

(b) $\min\{|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}, j_i^{\mathrm{any}})|\}_i/\pi > \frac{1}{6}$

(c) $|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}, J^{\mathrm{sig}})|/\pi > \frac{1}{9}$

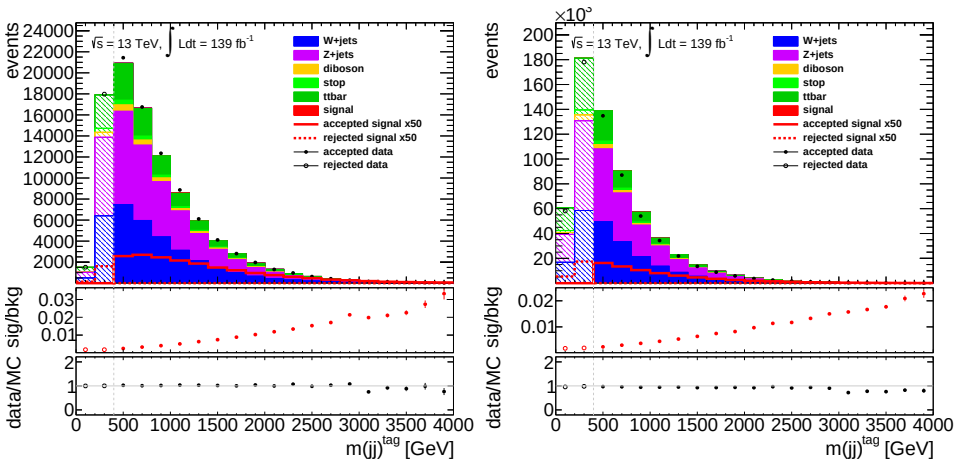
(d) $|\Delta\Phi(E_{\mathrm{T}}^{\mathrm{miss}}, (jj)^{\mathrm{sig}})|/\pi > \frac{1}{9}$

Figure 4.5: Last preselection (a) and all Anti-QCD selection cuts (b-d). Shown are all events passing the event selection defined in Tab. 4.3 up to the cut that is shown. Dashed vertical lines indicate the cut boundaries, the excluded region is indicated by shaded histograms.

It is evident that this cut excludes a kinematic phase space with a data over Monte Carlo ratio significantly above one while retaining a phase space with a ratio close to one. The excess above one is caused by the QCD multijet background only being part of the data but not the Monte Carlo simulation. This shows that the cut is effective with respect to its intended purpose. The remaining background appears to be well modelled by the Monte Carlo simulation. Additional cuts (Fig. 4.5 c and d) with the same purpose are implemented in the form of angular distance between missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ and the reconstructed remnants of

the hadronically decaying vector boson V_{had} . Depending on the merged/resolved regime the latter is reconstructed either as $(jj)^{\text{sig}}$ or J^{sig}. It can be seen that these additional cuts only have a small additional effect compared to the initial anti-QCD cut. Since they exclude only a small amount of signal they are kept.

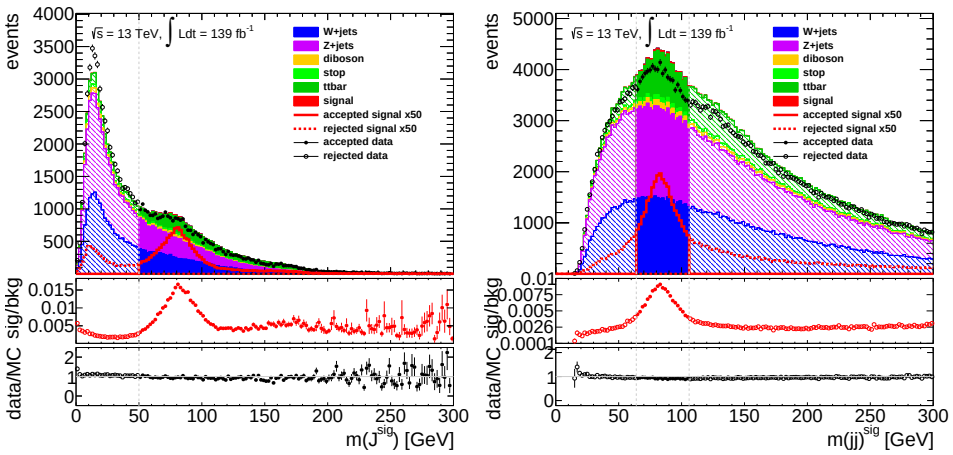
The last set of cuts, the signal-region cuts, are in place to enhance the ratio of signal to background at the end of the selection and hence to increase the sensitivity of the study to the signal process. The first cut of this set is a requirement on the combined (sum of four momenta) mass of the tagging jet system $m(jj)^{\text{tag}}$. It can be seen in Fig. 4.6. It is evident that the signal over background ratio in Monte Carlo increases monotonically with this variable. Events with $m(jj)^{\text{tag}} < 400$ GeV are excluded rejecting a significant amount of background while retaining most of the signal.



(a) merged (b) resolved

Figure 4.6: Tagging jet mass $m(jj)^{\text{tag}}$ selection. The distributions are shown for the merged and resolved regime. Shown are all events passing the event selection defined in Tab. 4.3 up to the cut that is shown. The excluded region is indicated by shaded histograms.

The second part of the signal-region cuts concerns the hadronically decaying vector boson V_{had}. In the resolved regime (Fig. 4.7(b)) an upper and lower cut on the combined mass of the signal jets system $m(jj)^{\text{sig}}$ is applied. Since these jets are intended to encapsulate the remnants of the decay of either a W or Z boson, the mass window is chosen to be in a region around their respective masses: $m_W = 80.38$ GeV, $m_Z = 91.19$ GeV [139]. A clear peak in the signal over background ratio can be seen in this region. In the merged regime only a lower cut on the mass of the signal jet $m(J^{\text{sig}})$ is applied. Fig. 4.7(a) shows that it successfully excludes events below the mass peak.



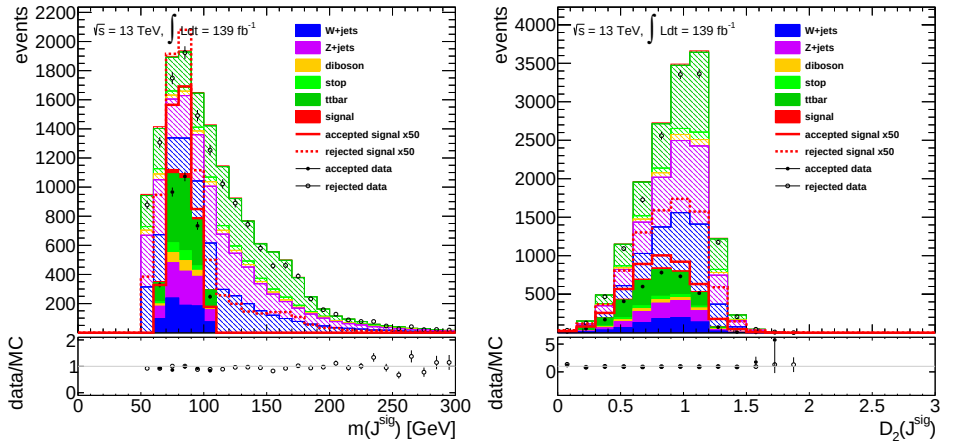
(a) merged (b) resolved

Figure 4.7: Signal jet mass $m(J)^{\text{sig}}$ ($m(jj)^{\text{sig}}$) selection in the merged (resolved) regime. Shown are all events passing the event selection defined in Tab. 4.3 up to the cut that is shown. The excluded region is indicated by shaded histograms.

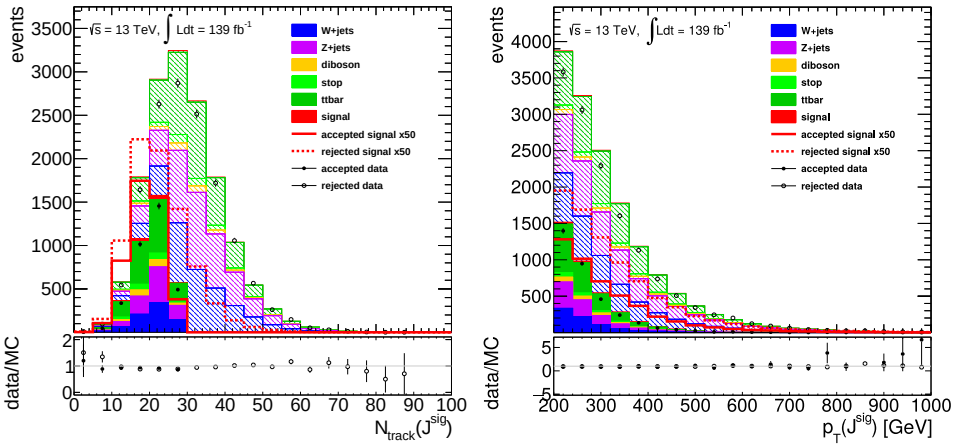
Additionally to this lower cut on the jet mass in the merged regime, a vector boson tagger is applied to determine if the signal jet J^{sig} is likely to be initiated by a vector boson. The tagger used is the 3-variable tagger described in Sec. 3.2.1. It is comprised of cuts on the large radius jet's mass, the number of tracks associated to it, as well as the substructure variable D_2. All of these cuts are defined as functions of the transverse momentum p_{T}. Based on these variables a binary decision is made whether the jet is likely to be initiated by a vector boson. Two different working points of the tagger are used with a signal efficiency of $\epsilon = 50\%$ or $\epsilon = 80\%$. Depending on the jet passing either the former, or not the former but the latter, the event is categorised into the high purity (HP) or low purity (LP) merged signal region.

Fig. 4.8 and 4.9 show the phase space of events with respect to the four variables (3 variables and p_{T}) passing or failing the tagger requirement at the two working points. The two variables that based on these plots appear to have the highest impact on the tagger decision are the jet mass and number of tracks. For events passing the tagger the former tends to be in the vicinity of m_W or m_Z which is in accordance with the assumption that the four momentum of J^{sig} corresponds to the hadronically decaying vector boson V_{had}. The number of tracks N_{trk} in the jet appear to be generally lower for jets passing the tagger. This can be explained by the excellent quark-gluon discrimination capabilities of N_{trk}. Jets with many associated tracks are more likely to be initiated by gluons instead of quarks and therefore are less likely to stem from the decay of V_{had} into a pair of quarks. The

energy correlation ratio D_2 [88] parametrises the two-pronged structure of the $V_{\text{had}} \to qq$ decay and therefore also generally selects for similar signatures.

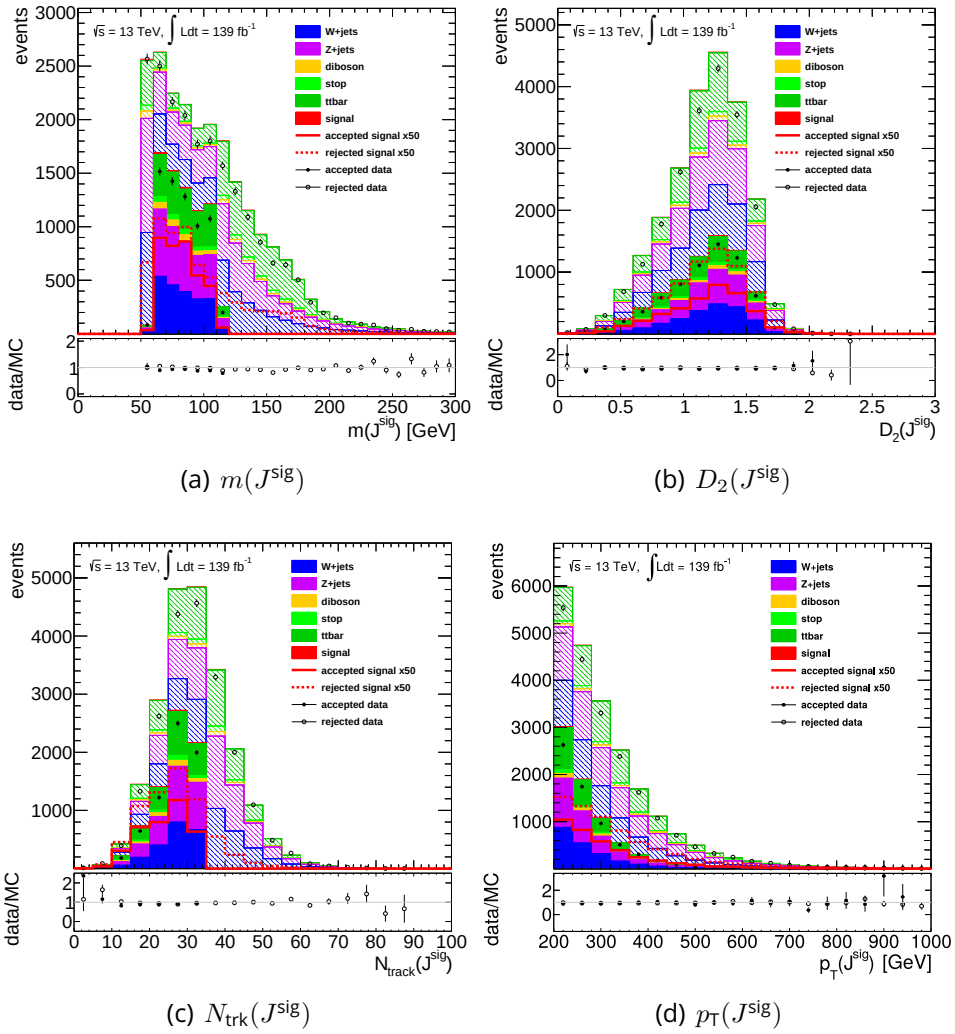


(a) $m(J^{\text{sig}})$ (b) $D_2(J^{\text{sig}})$



(c) $N_{\text{trk}}(J^{\text{sig}})$ (d) $p_{\text{T}}(J^{\text{sig}})$

Figure 4.8: Vector boson tagger (tagger$_V^{\epsilon=50\%}$) selection in the merged high purity (HP) signal region. All four features (3 variables and p_{T}) of the 3-variable tagger are presented. The region failing the tagger requirement corresponding to $\epsilon = 50\%$ is indicated by shaded histograms.



(a) $m(J^{\text{sig}})$ (b) $D_2(J^{\text{sig}})$

(c) $N_{\text{trk}}(J^{\text{sig}})$ (d) $p_{\text{T}}(J^{\text{sig}})$

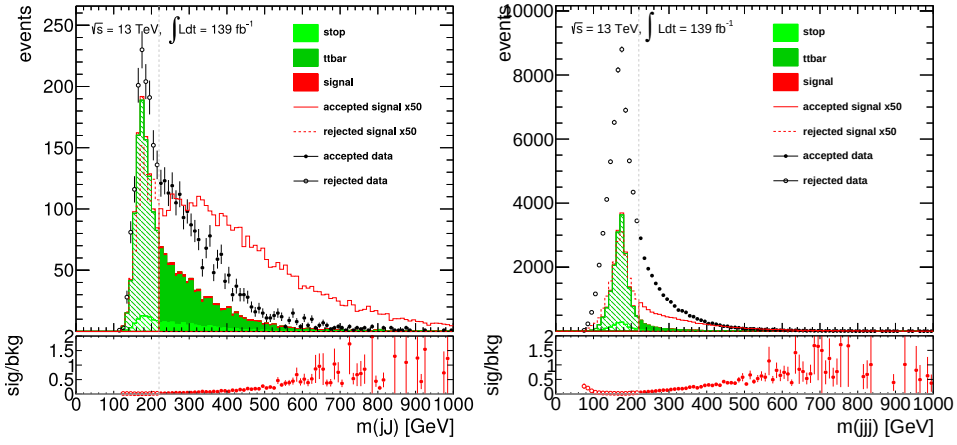
Figure 4.9: Vector boson tagger (tagger$_V^{\epsilon=80\%}$) selection in the merged low purity (LP) signal region. All four features (3 variables and p_{T}) of the 3-variable tagger are presented. The region failing the tagger requirement corresponding to $\epsilon = 80\%$ is indicated by shaded histograms.

In the resolved regime an additional cut on the 3-jet mass $m(jjj)$ is applied to reduce the fraction of EW $VV + jj$ events with a top quark in its decay structure. These events are part of the signal definition in the analysis but are intended to be reduced in order to enhance the vector boson scattering component of the EW $VV + jj$ process. It is defined as the combined (four momentum sum) mass $m(jjj)$ of the 3-jet system comprised of the two signal jets $(jj)^{\text{sig}}$ and any extra jet j^{extra} $(R = 0.4)$ where $m(jjj)$ is closest to the SM top mass of 172.76 GeV. This is motivated by top quarks most commonly decaying into a bottom

quark in association with a W boson where the W boson decays hadronically into two further quarks. This results in a three-pronged decay structure that can be reconstructed in the form of three jets. The 3-jet mass can therefore be understood as the mass of a top-candidate. A similar definition can be made in the merged regime where $m(jJ)$ is the mass calculated from the combination of the signal large radius ($R = 1.0$) jet J^{sig} with any extra $R = 0.4$ jet j^{extra} outside of the signal jet's cone: $\Delta R(j^{\text{extra}}, J^{\text{sig}}) > 1.4$. Tab. 4.4 shows the effect of such cuts ($m(jjj)$ or $m(jJ) > 220$ GeV) on the signal composition, i.e. the ratio of events with or without a top out of all signal Monte Carlo events.

	merged HP SR		resolved SR	
	before cut	after cut	before cut	after cut
top ($t\bar{t}$ + single t) MC events	1786	819	18208	1642
signal MC events	94	77	594	219
truth top fraction in signal MC	17%	13%	33%	9%

Table 4.4: Number of top ($t\bar{t}$ and single t) background and signal Monte Carlo events before and after applying the 3-jet mass $m(jjj)$ ($m(jJ)^{\text{sig}}$) > 220 GeV selection in the resolved (merged HP) SR. Also shown is the fraction of signal Monte Carlo events containing a top quark in the decay structure.



(a) merged HP SR (b) resolved SR

Figure 4.10: 3-jet mass $m(jjj)$ ($m(jJ)^{\text{sig}}$) selection in the resolved (merged HP) SR. Shown are top ($t\bar{t}$ and single t) background and signal Monte Carlo as well as data.

It can be seen that the additional cut significantly reduces the top contribution in the resolved regime while the contribution is already reasonably low without the cut in the merged regime. As a result it was decided only to apply this cut in the resolved regime. This avoids the loss of too many signal events in the merged regime while achieving an overall low top contribution which is now similar in both regimes. Fig. 4.10 shows the effect that cutting on this variable has on the number of signal and background Monte Carlo events. Shown is the signal compared to single top and $t\bar{t}$ backgrounds. An additional effect of this cut can be seen in the increased signal over background ratio in the accepted phase space: Additionally to excluding signal events with top contribution, the application of this cut also efficiently excludes top background events.

(a) merged SR HP

cut	signal	background	all MC	data	data/MC
All	15991	15306269	15322260	40529219	2.65
$\mathsf{N}(j^{\mathsf{tag}}) \geq 2$	12688	9353701	9366389	23971331	2.56
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{lead}}) > 30$ GeV	12677	9327811	9340487	23888467	2.56
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{sublead}}) > 30$ GeV	11289	7005564	7016853	18640386	2.66
$E_{\mathsf{T}}^{\mathsf{miss}} > 200$ GeV	5084	2327644	2332728	4134105	1.77
$E_{\mathsf{T}}^{\mathsf{miss,track}} > 50$ GeV	4559	2121651	2126210	3339761	1.57
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, E_{\mathsf{T}}^{\mathsf{miss,track}})\rvert/\pi < \frac{1}{2}$	4378	2072848	2077226	3099772	1.49
$\min\{\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, j_i^{\mathsf{any}})\rvert\}_i/\pi > \frac{1}{6}$	2567	1319134	1321701	1365047	1.03
$\mathsf{N}(J) \geq 1$	425	99296	99720	101238	1.02
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, J)\rvert/\pi > \frac{1}{9}$	408	96153	96561	98031	1.02
$m(jj)^{\mathsf{tag}} > 400$ GeV	373	76735	77109	78537	1.02
$m(J) > 50$ GeV	271	33166	33437	31665	0.95
tagger $\epsilon^{=50\%}_{V}(D_2$only)	166	15006	15172	14156	0.93
tagger $\epsilon^{=50\%}_{V}$	92	3557	3649	3349	0.92

(b) merged SR LP

cut	signal	background	all MC	data	data/MC
All	15899	15302712	15318611	40525870	2.65
$\mathsf{N}(j^{\mathsf{tag}}) \geq 2$	12596	9350144	9362740	23967982	2.56
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{lead}}) > 30$ GeV	12584	9324254	9336838	23885118	2.56
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{sublead}}) > 30$ GeV	11197	7002007	7013204	18637037	2.66
$E_{\mathsf{T}}^{\mathsf{miss}} > 200$ GeV	4992	2324087	2329079	4130756	1.77
$E_{\mathsf{T}}^{\mathsf{miss,track}} > 50$ GeV	4467	2118094	2122560	3336412	1.57
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, E_{\mathsf{T}}^{\mathsf{miss,track}})\rvert/\pi < \frac{1}{2}$	4286	2069291	2073577	3096423	1.49
$\min\{\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, j_i^{\mathsf{any}})\rvert\}_i/\pi > \frac{1}{6}$	2475	1315577	1318051	1361698	1.03
$\mathsf{N}(J) \geq 1$	333	95738	96071	97889	1.02
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, J)\rvert/\pi > \frac{1}{9}$	316	92596	92912	94682	1.02
$m(jj)^{\mathsf{tag}} > 400$ GeV	281	73178	73459	75188	1.02
$m(J) > 50$ GeV	179	29609	29788	28316	0.95
tagger $\epsilon^{=80\%}_{V}(D_2$only)	136	21877	22013	20685	0.94
tagger $\epsilon^{=80\%}_{V}$	73	7206	7280	6585	0.90

(c) resolved SR

cut	signal	background	all MC	data	data/MC
All	15826	15295505	15311331	40519285	2.65
$\mathsf{N}(j^{\mathsf{tag}}) \geq 2$	12522	9342938	9355460	23961397	2.56
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{lead}}) > 30$ GeV	12511	9317047	9329558	23878533	2.56
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{sublead}}) > 30$ GeV	11124	6994801	7005925	18630452	2.66
$E_{\mathsf{T}}^{\mathsf{miss}} > 200$ GeV	4918	2316881	2321799	4124171	1.78
$E_{\mathsf{T}}^{\mathsf{miss,track}} > 50$ GeV	4393	2110887	2115281	3329827	1.57
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, E_{\mathsf{T}}^{\mathsf{miss,track}})\rvert/\pi < \frac{1}{2}$	4213	2062084	2066297	3089838	1.50
$\min\{\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, j_i^{\mathsf{any}})\rvert\}_i/\pi > \frac{1}{6}$	2401	1308370	1310772	1355113	1.03
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, jj)\rvert/\pi > \frac{1}{9}$	2331	1234139	1236471	1274987	1.03
$\mathsf{N}(j^{\mathsf{sig}}) \geq 2$	1990	734182	736172	707381	0.96
$p_{\mathsf{T}}(j^{\mathsf{sig}}_{\mathsf{lead}}) > 40$ GeV	1902	628883	630785	607528	0.96
$p_{\mathsf{T}}(j^{\mathsf{sig}}_{\mathsf{sublead}}) > 20$ GeV	1902	628883	630785	607528	0.96
$m(jj)^{\mathsf{tag}} > 400$ GeV	1447	386978	388424	371275	0.96
$64 < m(jj)^{\mathsf{sig}} < 106$ GeV	573	85690	86263	79869	0.93
$m(jjj) > 220$ GeV	218	21815	22033	21789	0.99

(d) merged CR

cut	signal	background	all MC	data	data/MC
All	15253	15209816	15225068	40439416	2.66
$\mathsf{N}(j^{\mathsf{tag}}) \geq 2$	11949	9257248	9269197	23881528	2.58
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{lead}}) > 30$ GeV	11938	9231358	9243296	23798664	2.57
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{sublead}}) > 30$ GeV	10551	6909111	6919662	18550583	2.68
$E_{\mathsf{T}}^{\mathsf{miss}} > 200$ GeV	4345	2231191	2235536	4044302	1.81
$E_{\mathsf{T}}^{\mathsf{miss,track}} > 50$ GeV	3820	2025198	2029018	3249958	1.60
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, E_{\mathsf{T}}^{\mathsf{miss,track}})\rvert/\pi < \frac{1}{2}$	3640	1976395	1980034	3009969	1.52
$\min\{\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, j_i^{\mathsf{any}})\rvert\}_i/\pi > \frac{1}{6}$	1828	1222681	1224509	1275244	1.04
$\mathsf{N}(J) \geq 1$	208	81297	81506	84295	1.03
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, J)\rvert/\pi > \frac{1}{9}$	195	78428	78623	81354	1.03
$m(jj)^{\mathsf{tag}} > 400$ GeV	160	59010	59171	61860	1.05
$m(J) > 50$ GeV	70	17213	17283	16833	0.97
fail tagger $\epsilon^{=80\%}_{V}$	70	17213	17283	16833	0.97

(e) resolved CR

cut	signal	background	all MC	data	data/MC
All	15182	15192603	15207786	40422583	2.66
$\mathsf{N}(j^{\mathsf{tag}}) \geq 2$	11879	9240035	9251915	23864695	2.58
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{lead}}) > 30$ GeV	11868	9214145	9226013	23781831	2.58
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{sublead}}) > 30$ GeV	10480	6891899	6902379	18533750	2.69
$E_{\mathsf{T}}^{\mathsf{miss}} > 200$ GeV	4275	2213979	2218254	4027469	1.82
$E_{\mathsf{T}}^{\mathsf{miss,track}} > 50$ GeV	3750	2007985	2011735	3233125	1.61
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, E_{\mathsf{T}}^{\mathsf{miss,track}})\rvert/\pi < \frac{1}{2}$	3570	1959182	1962751	2993136	1.52
$\min\{\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, j_i^{\mathsf{any}})\rvert\}_i/\pi > \frac{1}{6}$	1758	1205468	1207226	1258411	1.04
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, jj)\rvert/\pi > \frac{1}{9}$	1689	1131398	1133086	1178463	1.04
$\mathsf{N}(j^{\mathsf{sig}}) \geq 2$	1352	631861	633213	611368	0.97
$p_{\mathsf{T}}(j^{\mathsf{sig}}_{\mathsf{lead}}) > 40$ GeV	1264	526564	527828	511515	0.97
$p_{\mathsf{T}}(j^{\mathsf{sig}}_{\mathsf{sublead}}) > 20$ GeV	1264	526564	527828	511515	0.97
$m(jj)^{\mathsf{tag}} > 400$ GeV	808	284659	285467	275262	0.96
not $64 < m(jj)^{\mathsf{sig}} < 106$ GeV	808	284659	285467	275262	0.96
$m(jjj) > 220$ GeV	516	177607	178122	175982	0.99

Table 4.5: Cut flow of the complete event selection of the analysis. The selections are applied subsequently, meaning only events failing the first (merged SR HP) selection enter the second (merged SR LP) selection etc.

For each signal region (SR) there is a control region (CR) defined by inverting the requirement with respect to the hadronically decaying vector boson V_{had}: In the resolved case the $m(jj)^{\mathsf{sig}}$ mass cut is inverted. In the merged case, the event has to fail the vector boson tagger's lower working point of $\epsilon = 80\%$. In total this yields five analysis regions: Three signal and two control regions.

Tab. 4.5 shows the complete cutflow for all events entering these analysis regions. The cuts are applied successively and only events failing the selection of the previous region are considered for the following region. There is one exception to this rule: Events failing the resolved SR selection only because of the extra 3-jet mass $m(jjj)$ cut are not considered for the merged CR. The order of selection is

as follows: High- (HP) and low-purity (LP) merged signal region, resolved signal region, and merged and resolved control regions (see Fig. 4.11).

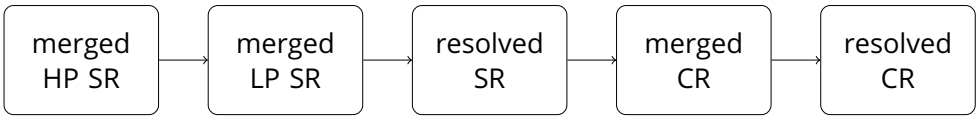


Figure 4.11: Order of selection for the five analysis regions defined by the event selection.

4.5 . $m(jj)^{\text{tag}}$ Reweighting

The invariant mass of the two selected tagging jets $m(jj)^{\text{tag}}$ is well known to be mismodelled in the nominal Sherpa W/Z+jets background samples. An investigation of this can for example be found in a recent ATLAS vector boson fusion measurement on EW Zjj processes [140]. This variable is sensitive to the forward topology of the signal process. Hence, a good modelling is required. The mismodelling is partly compensated for by a simple linear reweighting procedure. For this purpose the sum of W+jets and Z+jets background in the analysis control regions was compared with the data minus all other backgrounds, the result of which can be seen in Fig. 4.12. Linear fits were performed to the ratios of these distributions with each other. This procedure was done separately for the three Monte Carlo campaigns mc16a, -d, and -e. Tab. 4.6 shows the resulting factors, i.e. the slopes and constants of the linear fits. The uncertainties given there correspond to the uncertainties of the linear fits alone. No systematic uncertainties are considered here.

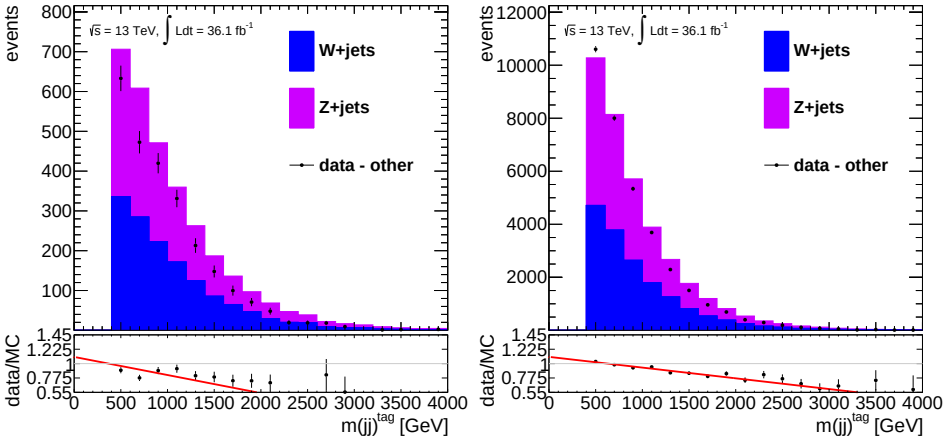
campaign	offset merged	slope merged [GeV^{-1}]	offset resolved	slope resolved [GeV^{-1}]
mc16a	1.12 ± 0.04	$(2.9 \pm 0.2) \times 10^{-4}$	1.11 ± 0.01	$(1.7 \pm 0.1) \times 10^{-4}$
mc16d	1.07 ± 0.04	$(1.4 \pm 0.3) \times 10^{-4}$	1.06 ± 0.01	$(1.0 \pm 0.1) \times 10^{-4}$
mc16e	1.01 ± 0.03	$(1.3 \pm 0.3) \times 10^{-4}$	1.07 ± 0.01	$(1.1 \pm 0.09) \times 10^{-4}$

Table 4.6: 0-lepton $m(jj)^{\text{tag}}$ reweighting factors derived by the linear fits shown in Fig. 4.12.

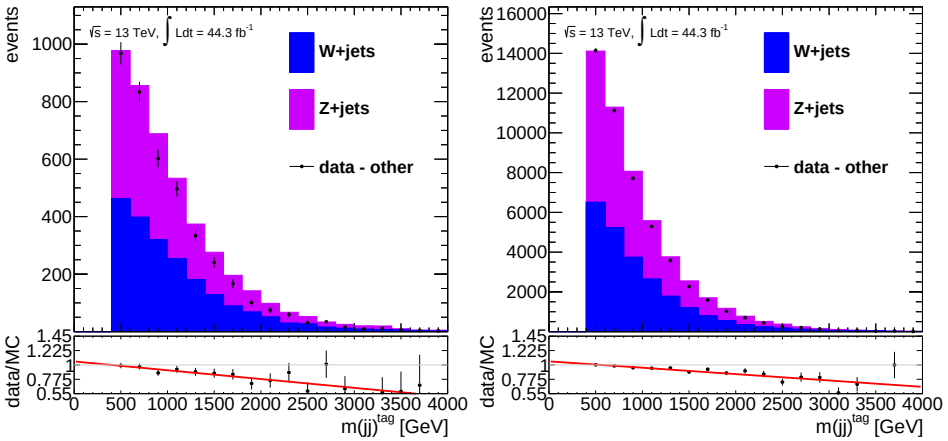
Fig. 4.13 shows the difference of the $m(jj)^{\text{tag}}$ distribution before (left column) and after (middle column) $m(jj)^{\text{tag}}$ reweighting in the analysis signal regions[1]. The corresponding distributions for the control regions are shown in Fig. 4.14.

The necessity for the reweighting process is especially apparent in the resolved signal region. It is apparent that the reweighting procedure is successful in compensating for the slope in the $m(jj)^{\text{tag}}$ distribution. The $m(jj)^{\text{tag}}$ distributions of the alternative (MadGraph) W/Z+jets samples are also shown in 4.15. The alternative samples show a smaller slope compared to the nominal samples before reweighting but a considerable effect is visible nonetheless. Since at a later stage (Sec. 4.7) these alternative samples are used to derive a systematic uncertainty on the modelling of the W/Z+jets background with respect to the shape of the multivariate analysis discriminant, and a separate systematic uncertainty for the reweighting of the nominal samples is derived, the alternative samples were not reweighted.

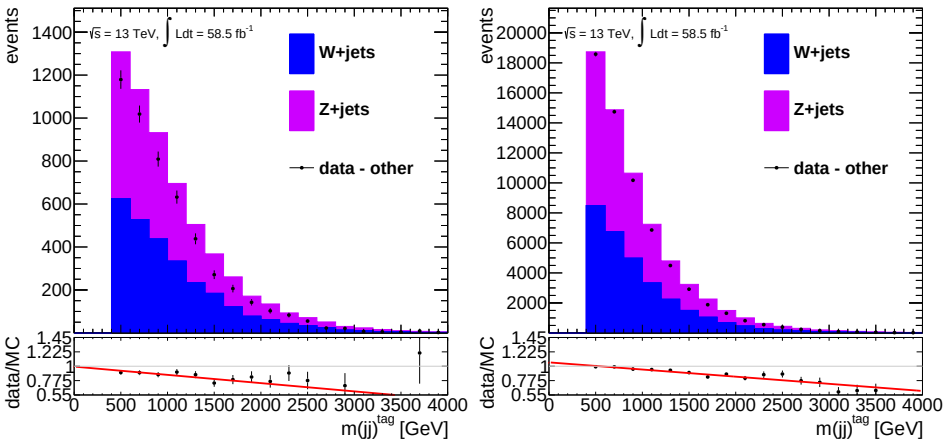
[1] In appendix Sec. A.2 the corresponding figures for each of the Monte Carlo campaigns are shown separately.



(a) mc16a merged CR (b) mc16a resolved CR

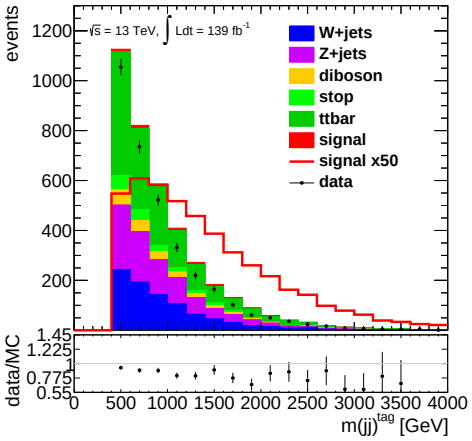


(c) mc16d merged CR (d) mc16d resolved CR

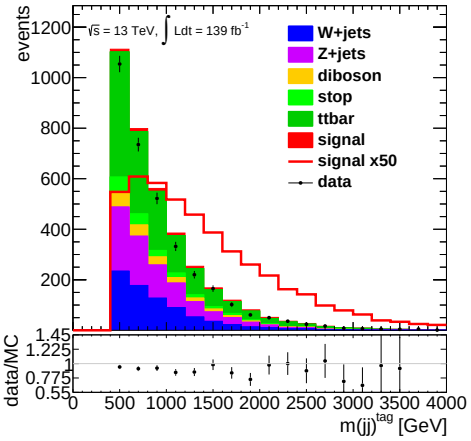


(e) mc16e merged CR (f) mc16e resolved CR

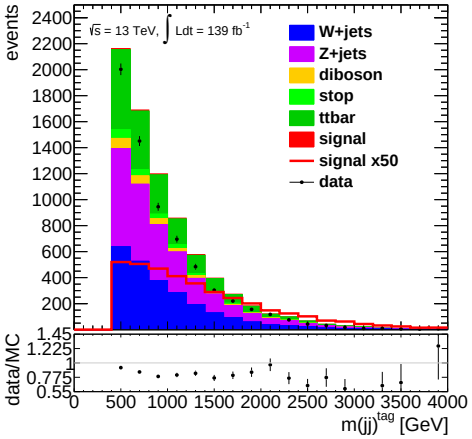
Figure 4.12: Derivation of the $m(jj)^{\text{tag}}$ reweighting factors from data over Monte Carlo ratios.



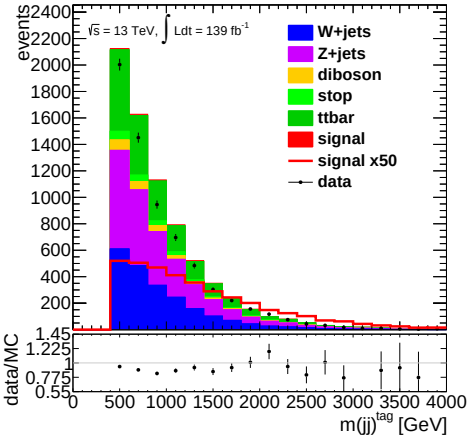
(a) merged HP SR before



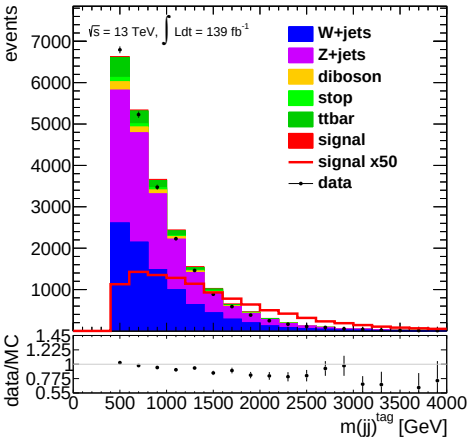
(b) merged HP SR after



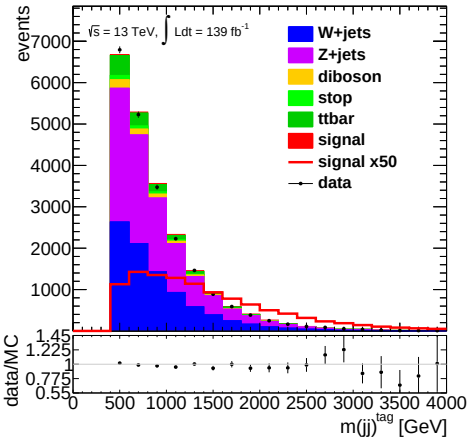
(c) merged LP SR before



(d) merged LP SR after

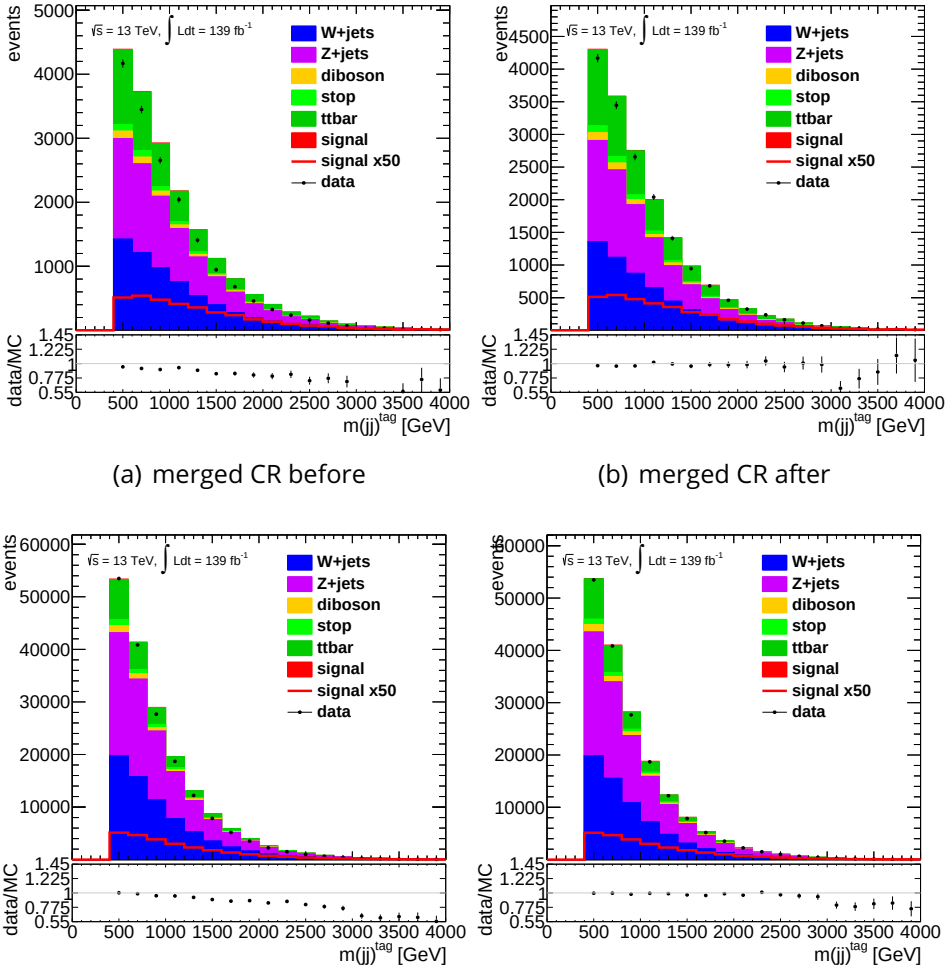


(e) resolved SR before



(f) resolved SR after

Figure 4.13: $m(jj)^{\text{tag}}$ distributions in the signal regions before (left) and after (right) reweighting.



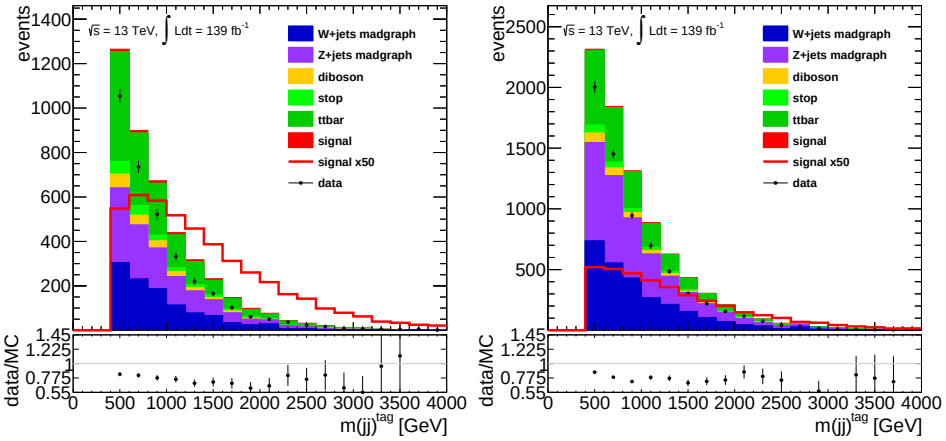
(a) merged CR before

(b) merged CR after

(c) resolved CR before

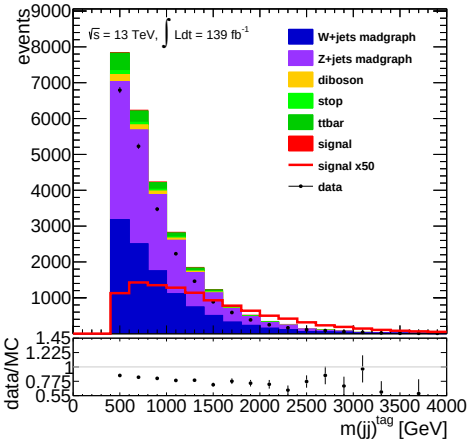
(d) resolved CR after

Figure 4.14: $m(jj)^{\text{tag}}$ distributions in the control regions before (left) and after (right) reweighting.

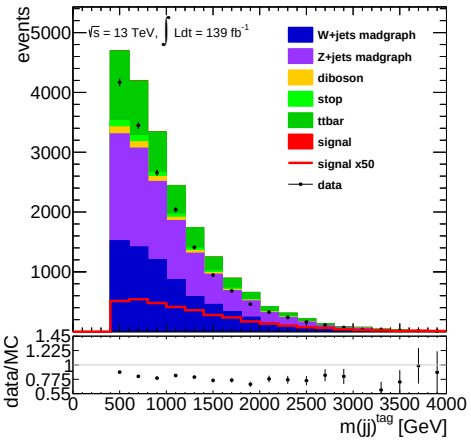


(a) merged HP SR

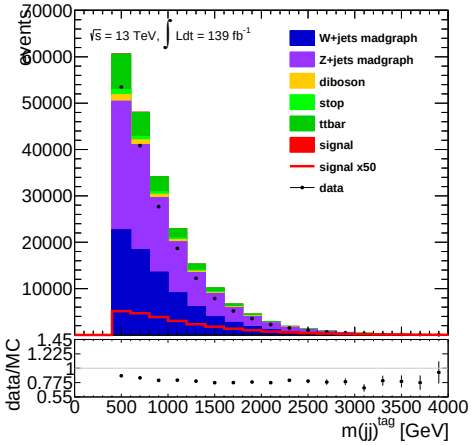
(b) merged LP SR



(c) resolved SR



(d) merged CR



(e) resolved CR

Figure 4.15: $m(jj)^{\text{tag}}$ distributions using alternative Madgraph $W/Z+$ jets samples without reweighting.

4.6 . Multivariate Analysis

The goal of the multivariate analysis (MVA) was to construct a potent discriminant between signal and background events that is more efficient than any one of the features used in its construction alone. The final fit of the analysis, which is used in the statistical interpretation of the results (Sec. 4.9), uses the distribution of the MVA output in the signal regions. Hence, besides a good discrimination power, also a relatively low susceptibility to eventual modelling uncertainties of the Monte Carlo generated background is of great importance. For this reason, in addition to the nominal MVA approach using a relatively ambitious method of recurrent neural networks (RNN), a more conservative approach using a simple feed forward neural network (NN) has been studied to serve as a comparison point with respect to eventual discrepancies. Both the RNN and NN were trained with Keras [141] with TensorFlow backend [142]. The trained models were implemented into the analysis framework with the help of the light weight neural network (lwtnn) package [143].

The NN approach comprises a neural network with two hidden layers of 16 nodes each. ReLu (rectified linear unit) activation functions were used for the hidden layers. The single output node was assigned a sigmoid activation function yielding a continuous output of the NN MVA in the range from zero to one. A different set of input features was used for the merged and resolved regime. They are listed in Tab. 4.7.

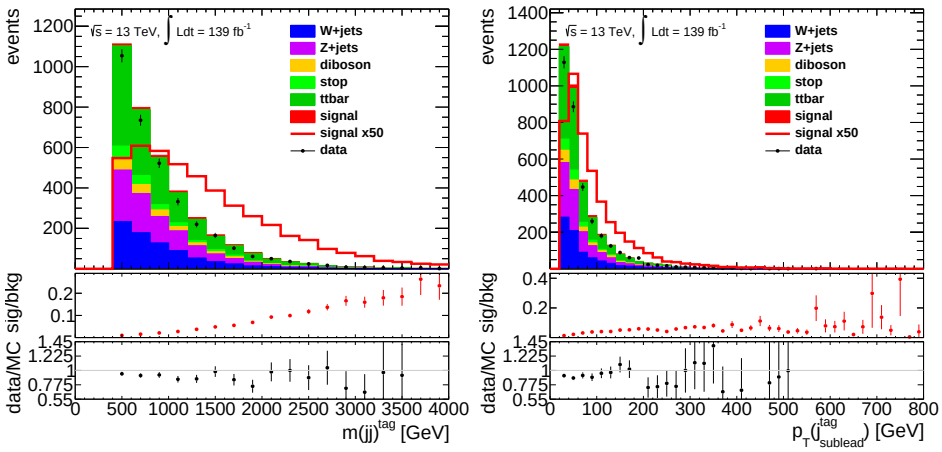
feature	merged	resolved
signal jet mass	$m(J^{\text{sig}})$	$m(jj^{\text{sig}})$
signal jet transverse momentum	$p_{\text{T}}(J^{\text{sig}})$	$p_{\text{T}}(j^{\text{sig}}_{\text{lead}})$ and $p_{\text{T}}(j^{\text{sig}}_{\text{sublead}})$
signal jet separation	-	$\Delta\eta(j^{\text{sig}}_{\text{lead}}, j^{\text{sig}}_{\text{sublead}})$
number of tracks associated to the signal jets	-	$N_{\text{trk}}(j^{\text{tag}}_{\text{lead}})$ and $N_{\text{trk}}(j^{\text{tag}}_{\text{sublead}})$
signal jet - $E_{\text{T}}^{\text{miss}}$ separation	$\Delta\Phi(J^{\text{sig}}, E_{\text{T}}^{\text{miss}})$	-
tagging jet transverse momentum	$p_{\text{T}}(j^{\text{tag}}_{\text{sublead}})$ only[†]	$p_{\text{T}}(j^{\text{tag}}_{\text{lead}})$ and $p_{\text{T}}(j^{\text{tag}}_{\text{sublead}})$
tagging jet mass	$m(jj)^{\text{tag}}$	
number of tracks associated to the tagging jets	$N_{\text{trk}}(j^{\text{tag}}_{\text{lead}})$ and $N_{\text{trk}}(j^{\text{tag}}_{\text{sublead}})$	
missing transverse momentum	$E_{\text{T}}^{\text{miss}}$	

Table 4.7: Features used in the NN MVA. A different set of features is used in the merged and resolved regimes. The last 4 rows show features that are defined in the same way in both regimes. '-' indicates that no such feature is defined in the respective regime. In total 8 features are used in the merged and 12 features are used in the resolved regime.

[†] : Note that in the merged regime only the subleading tagging jet p_T is used. It is expected that including the leading tagging jet p_{T} would increase the separation power. But when optimising the MVA it was decided to stay consistent with what was done in the MVA of the previous 36.1 fb^{-1} analysis [1] which also only uses the subleading jet.

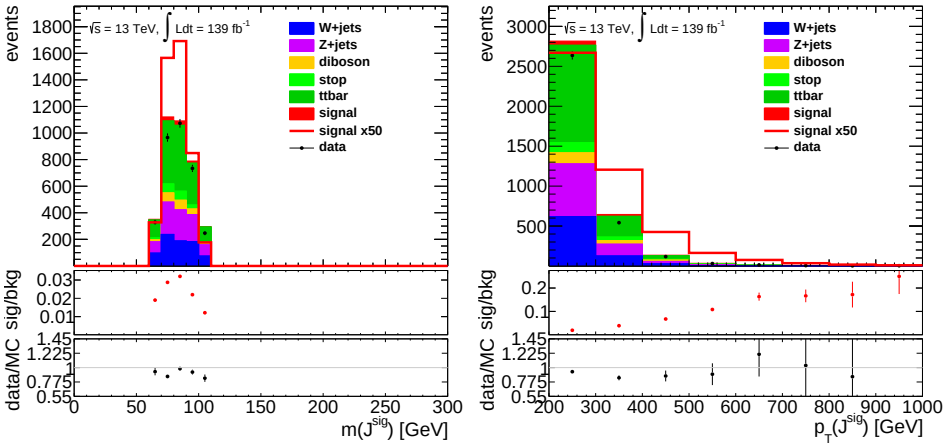
Separate models were trained in the three signal regions. In each region the samples were split in half (by event number) and one model each was trained on each half. The other half was used for validation. After training, the models were deployed in such a way that they are only applied to the validation set. An early stopping criterion was applied with a patience of 50 epochs, i.e. if no improvement in the loss function with respect to the validation set has been achieved for this number of epochs the training was stopped. These measures were performed in order to avoid overtraining. Before training (and application of the network) all features were scaled in such a way that the mean of their respective distribution lies at zero with a standard deviation scaled to one. The scaling factors were derived from the training set and applied to both training and validation set. The training was performed with a batch size of 2^8 events and a learning rate of 0.001 using the Adam optimiser [144]. Binary crossentropy was chosen as loss function. Signal EW $VV + jj$ events were labelled as $'1'$ and background events (all other Monte Carlo) were labelled as $'0'$. Hence the NN was trained in such a way that a high output corresponds to a high likelihood that the event is signal rather than background.

Fig. 4.16, 4.17, and 4.18 show the distributions of all features of the NN in the three signal regions of the analysis. From comparing the signal over background ratio panels shown there with the corresponding panels of the NN output in Fig. 4.19 it is evident that the NN succeeds in its goal of defining a discriminant with much higher separation power than any of the features alone. The data over Monte Carlo panels show that overall the features are reasonably well modelled without any strong biases (e.g. expressed in the form of a slope).

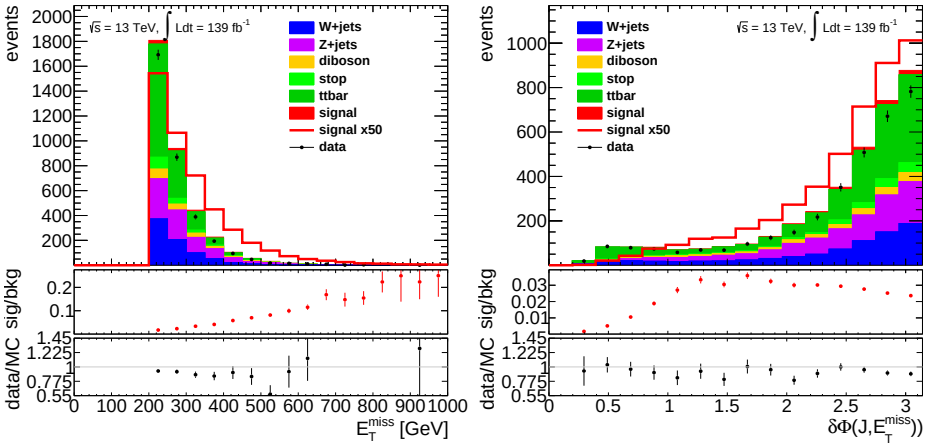


(a) tagging jet mass (b) tagging jet p_{T}

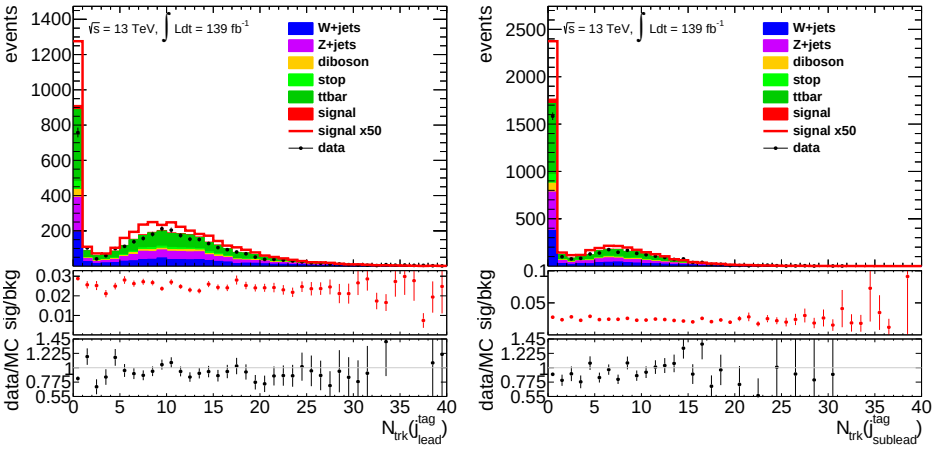
Figure 4.16: Features of the NN MVA in the merged HP signal region (continued on next page).



(a) signal jet mass (b) signal jet p_{T}

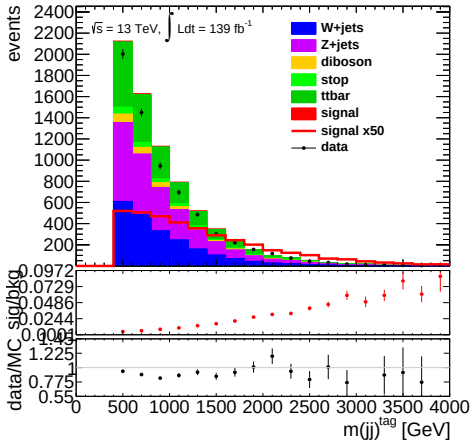


(c) $E_{\mathrm{T}}^{\mathrm{miss}}$ (d) signal jet-$E_{\mathrm{T}}^{\mathrm{miss}}$ separation

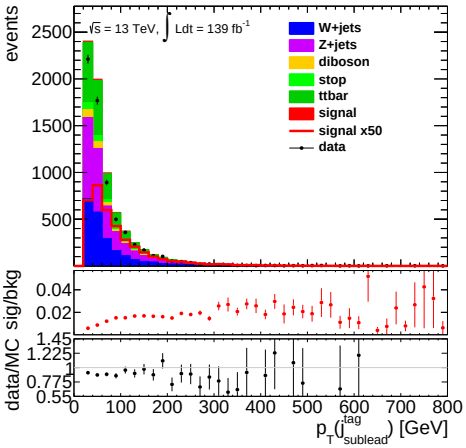


(e) tracks assoc. to tagging jet (f) tracks assoc. to tagging jet

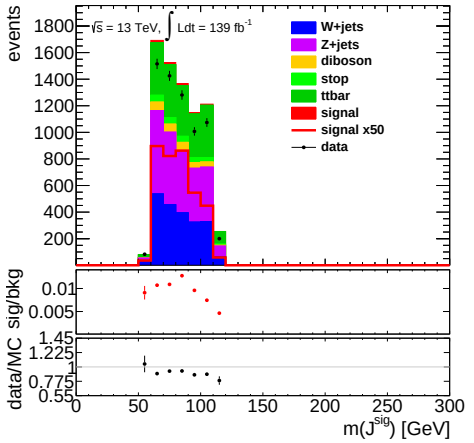
Figure 4.16: Continuation from previous page.



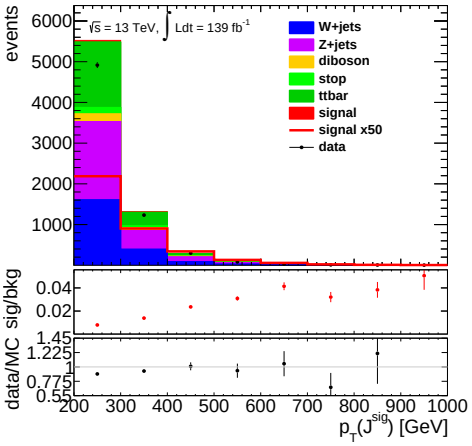
(a) tagging jet mass



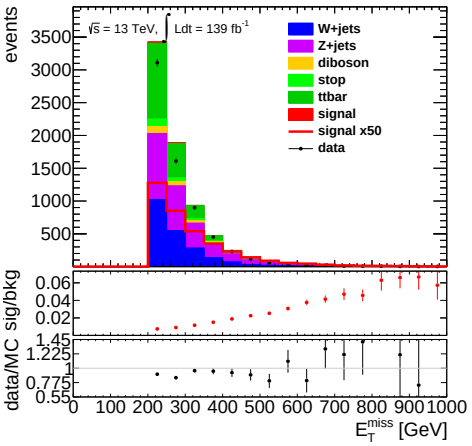
(b) tagging jet p_{T}



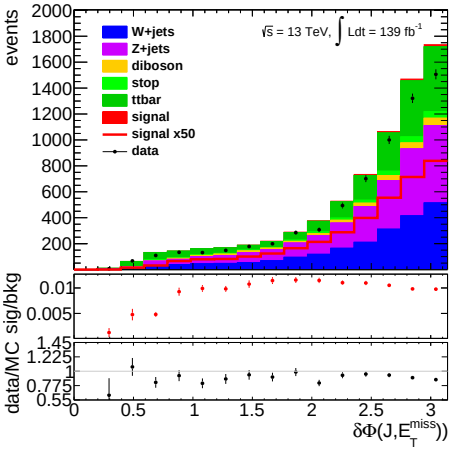
(c) signal jet mass



(d) signal jet p_{T}

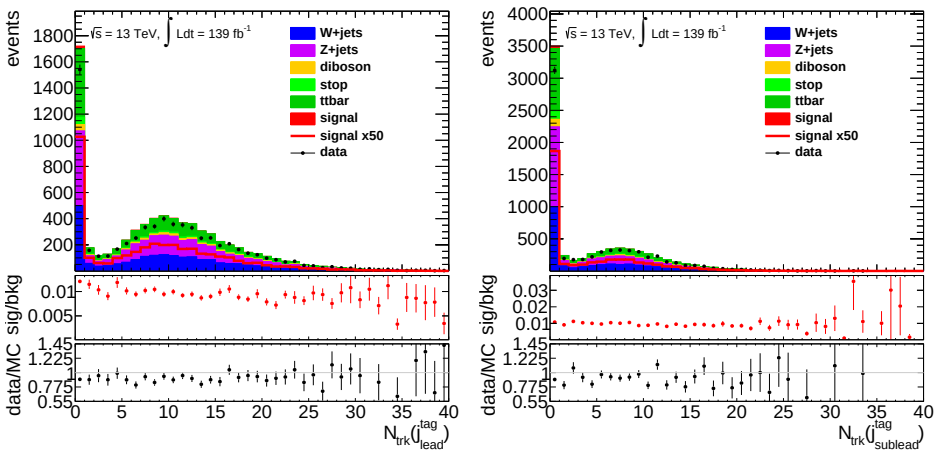


(e) $E_{\mathrm{T}}^{\mathrm{miss}}$



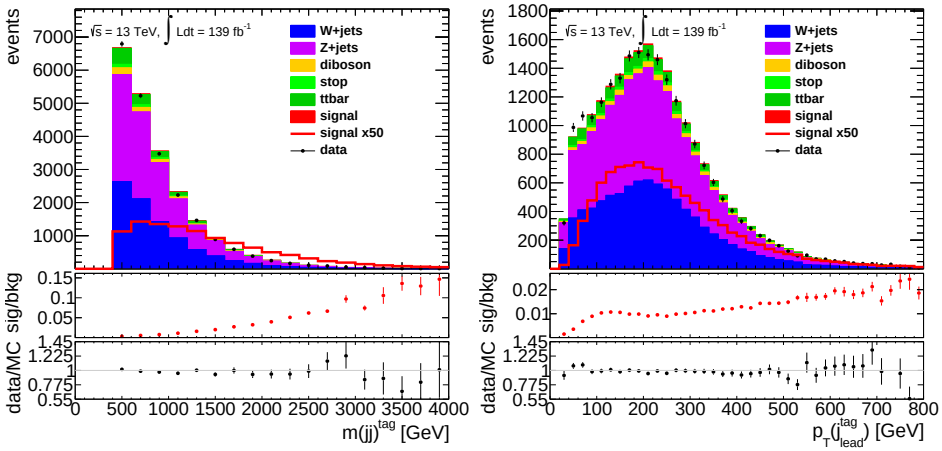
(f) signal jet-$E_{\mathrm{T}}^{\mathrm{miss}}$ separation

Figure 4.17: Features of the NN MVA in the merged LP signal region (continued on next page).



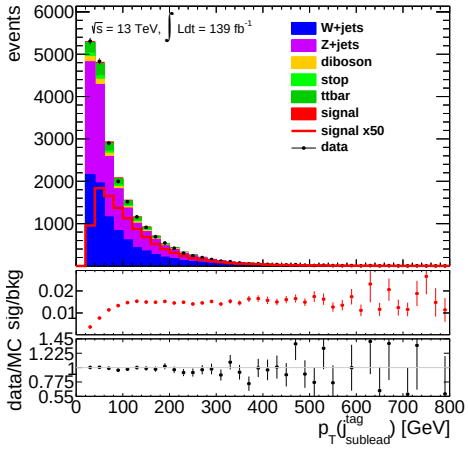
(a) tracks assoc. to tagging jet (b) tracks assoc. to tagging jet

Figure 4.17: Continuation from previous page.

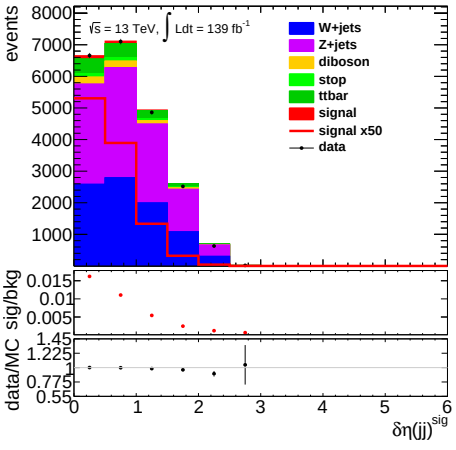


(a) tagging jet mass (b) tagging jet p_{T}

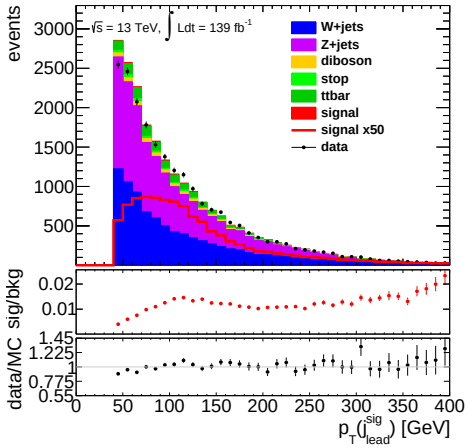
Figure 4.18: Features of the NN MVA in the resolved signal region (continued on next two pages).



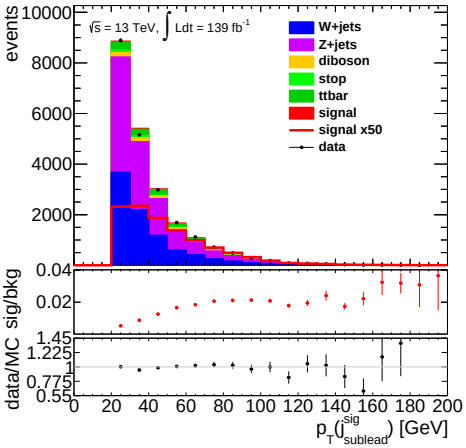
(a) tagging jet p_{T}



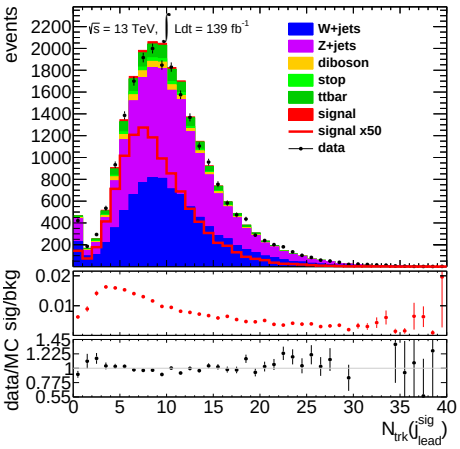
(b) DeltaEtaSignalJets



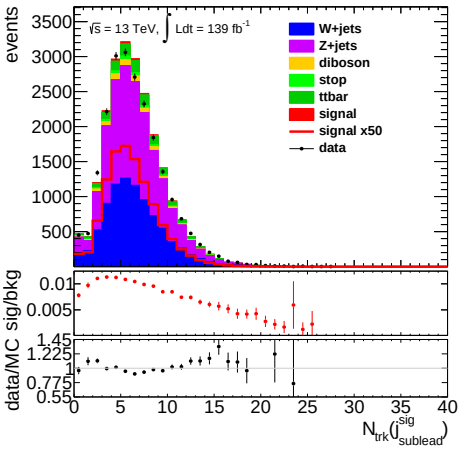
(c) signal jet p_{T}



(d) signal jet p_{T}

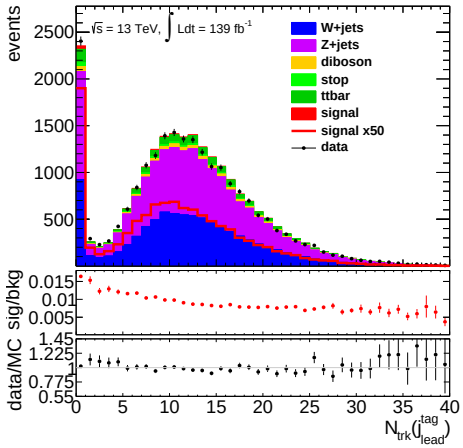


(e) tracks assoc. to signal jet

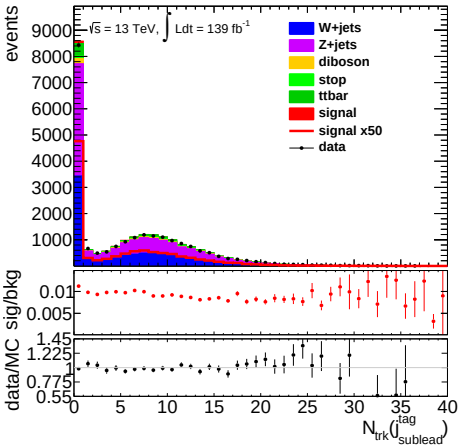


(f) tracks assoc. to signal jet

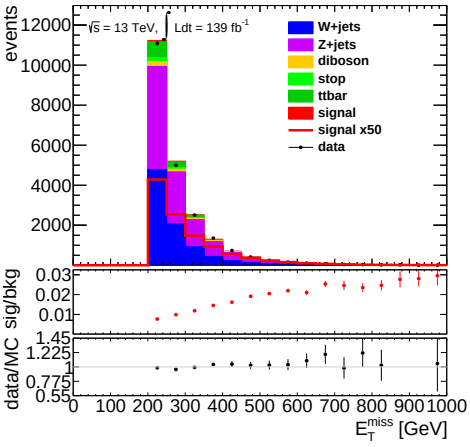
Figure 4.18: Continuation from previous page.



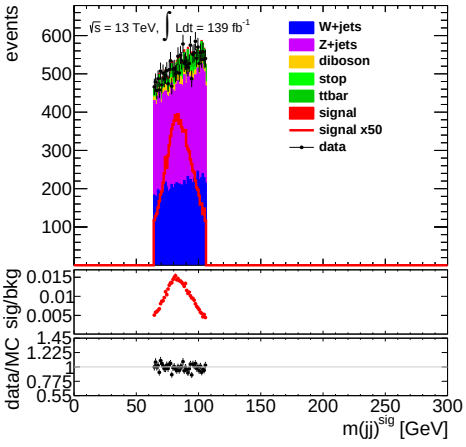
(a) tracks assoc. to tagging jet



(b) tracks assoc. to tagging jet



(c) $E_{\mathrm{T}}^{\mathrm{miss}}$



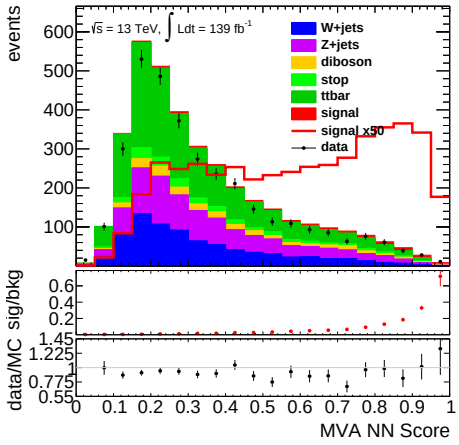
(d) signal jet mass

Figure 4.18: Continuation from previous page.

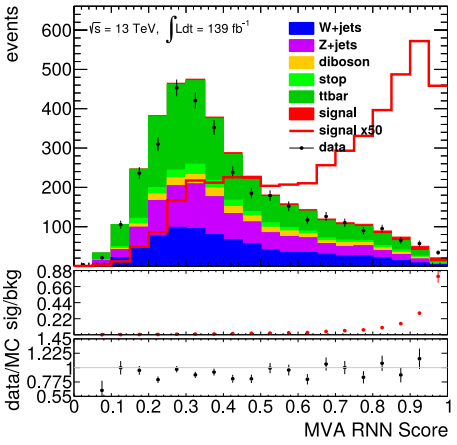
The baseline MVA method of this analysis is a recurrent neural network (RNN) with features corresponding to all momentum components of the five $R = 0.4$ p_{T}-leading jets in the event. While the configuration with five jets (called RNN($5j$) in the following) is the baseline, for comparisons a configuration using only four jets (RNN($4j$)) has also been trained independently. The corresponding plots produced with RNN($4j$) instead of RNN($5j$) can be seen in the appendix Sec. A.8. Compared to the NN method, the RNN method utilises a large number of low-level features which are general for any event while the NN method utilises a smaller number of high-level features which are especially engineered for the signal process. The goal of this method was to utilise additional information that is not captured by these high-level features. An RNN architecture was chosen because not all events necessarily have five jets defined. While a standard feed forward neural network takes a set number of features as input, the architecture of recurrent neural networks allows for variable feature size. In addition to the four momenta, the number N_{trk} of tracks with $p_{\mathrm{T}} > 0.5$ GeV associated to each jet are added as additional features. In the merged regime, both momentum and N_{trk} of the signal large radius ($R = 1.0$) jet J^{sig} are used as well. Since this MVA method has been developed centrally for all analysis channels it does not use channel-specific features like missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$. It is assumed that the RNN is able to somewhat deduce the direction and amplitude of $E_{\mathrm{T}}^{\mathrm{miss}}$ from the four momenta of the p_{T}-leading jets since they dominate the sum of p_{T}.

Previous analyses searching for EW $VV + jj$ structures rely heavily on the characteristic structure of the tagging jet system. Since the RNN is not explicitly given the information which jets are selected to be tagging jets, it is assumed that it uses a similar heuristic based on the four momenta of the given jets. Most of the information of the tagging jet object definition given in Sec. 4.3 are available for the RNN: The four momenta indicate if jets are in opposite hemispheres and their combined mass can be deduced. But the RNN is not given any specific information from a $b-$tagging algorithm which is used to exclude jets from being selected as tagging jets. While this is the case for the features of the RNN, it should be noted that both RNN and NN are applied after the same event selection as defined in Sec. 4.4. The event selection does rely on the characteristic structure of the tagging jet system by utilising cuts on various aspects of the tagging jets.

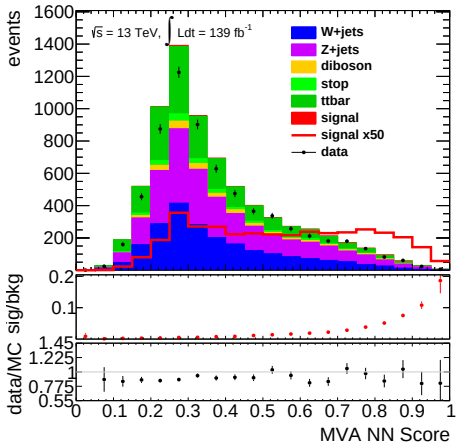
Fig. 4.19 shows the MVA scores of the NN and RNN. Fig. 4.20 compares the corresponding ROC curves of the MVAs. It is evident that RNN($5j$) shows the highest separation power, followed by RNN($4j$), while NN shows the lowest separation power. This is partly due to the higher number of features the RNN can work with. In the merged regime the RNN($5j$) uses the four momenta components and number of tracks associated to five $R = 0.4$ jets and the signal $R = 1.0$ jet amounting to 30 features, compared to the 9 features of the NN.



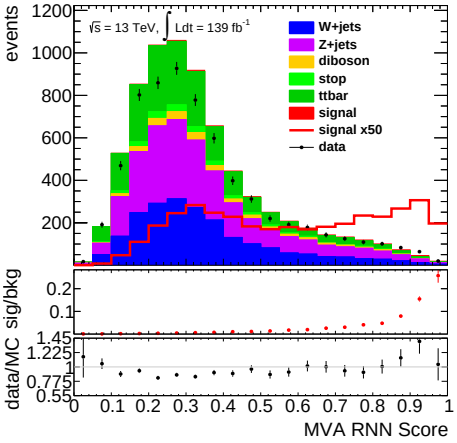
(a) merged HP SR, NN



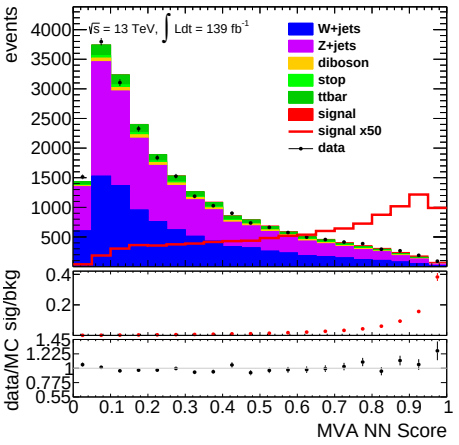
(b) merged HP SR, RNN($5j$)



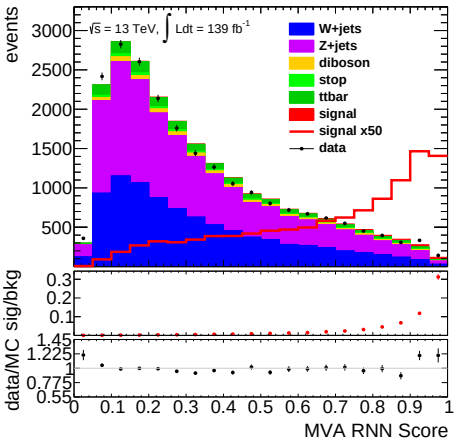
(c) merged LP SR, NN



(d) merged LP SR, RNN($5j$)

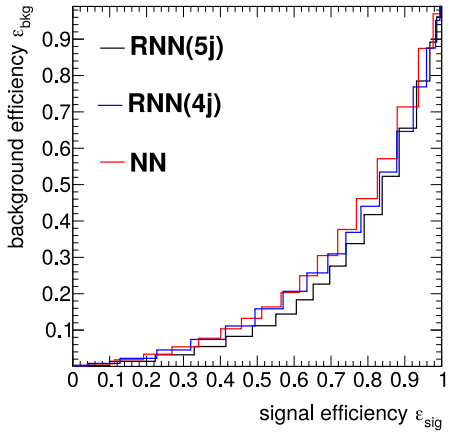


(e) resolved SR, NN

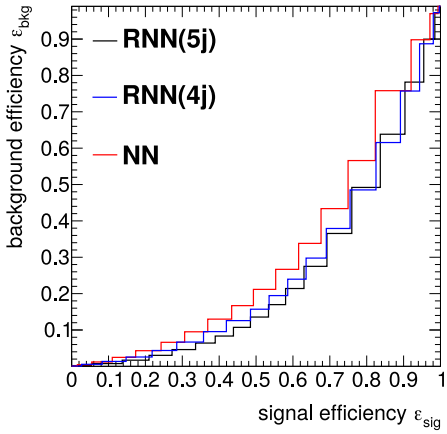


(f) resolved SR, RNN($5j$)

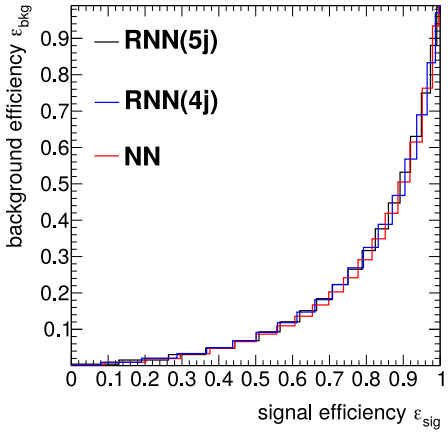
Figure 4.19: MVA scores in all signal regions of either NN or RNN($5j$) discriminant after training (see Fig. A.8.22 for RNN($4j$)).



(a) merged HP SR



(b) merged LP SR



(c) resolved SR

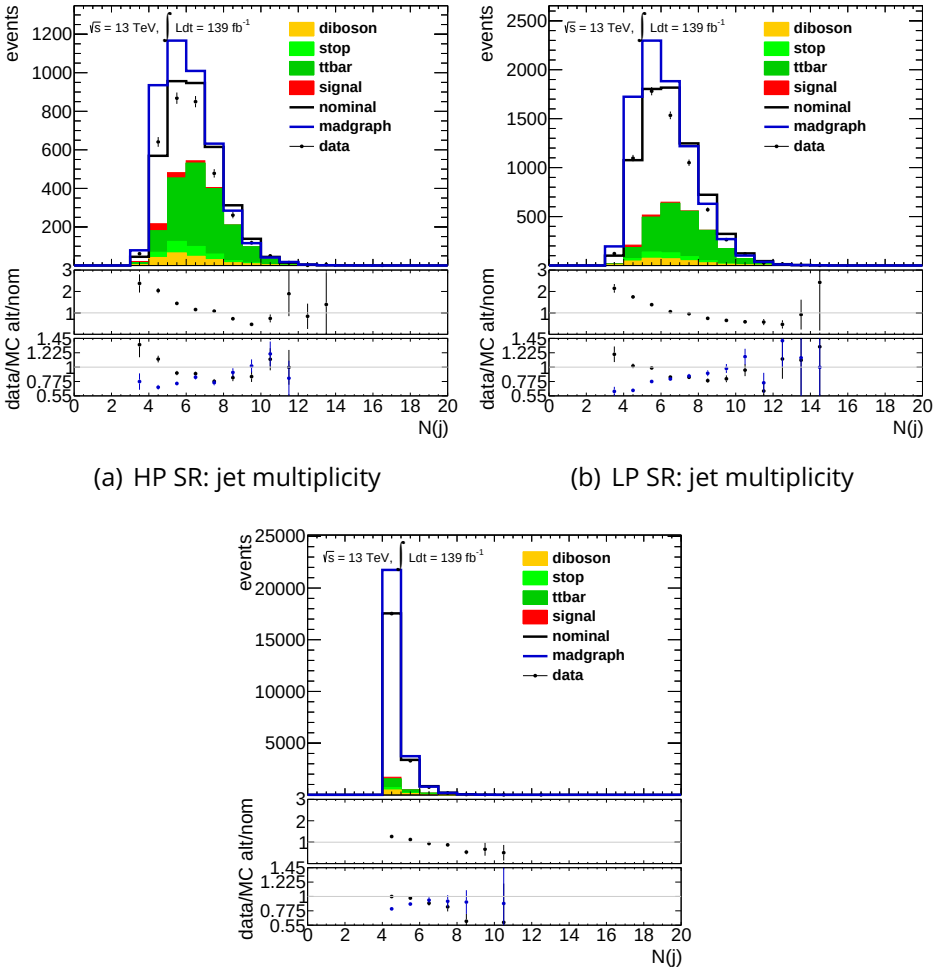
Figure 4.20: ROC curves in all signal regions of the three different MVA methods after training.

4.7 . Sensitivity to Physics Modelling

In addition to the features themselves, the RNN also has some limited information on how many $R = 0.4$ jets are in the event, i.e. the jet multiplicity N_j. Since jet multiplicity is a relatively poorly modelled variable in Monte Carlo (as discussed further below), this is a somewhat unwanted side effect inherent to the RNN's architecture. If there are fewer than 5 jets in the event, the RNN($5j$) has fewer features to use than otherwise and can therefore differentiate between events with 2, 3, 4, or 5 jets[2].

[2]Due to the event selection of the signal regions requiring a pair of tagging jets, there are no events with fewer than 2 jets.

In contrast to this, the NN has no feature that allows to assess the jet multiplicity directly. All features concerning jets are either of tagging or signal jets These are guaranteed to exist in any event in the signal regions which the NN is trained and applied on. This was a conscious decision in contrast to the previous analysis [1] in which the track-jet multiplicity, a similarly poorly modelled variable with high correlation to the jet multiplicity, was used as one of the features. While without jet multiplicity as feature the NN is expected to perform worse with respect to signal/background separation it is more robust against the modelling uncertainty on this variable. In this way, when compared with the RNN, the NN can be used to isolate the effects from this uncertainty and to investigate its impact on the RNN.



(a) HP SR: jet multiplicity

(b) LP SR: jet multiplicity

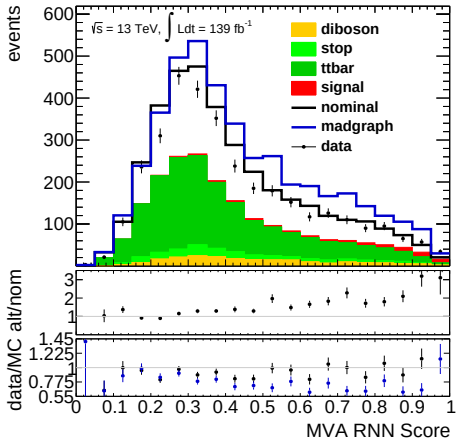
(c) resolved SR: jet multiplicity

Figure 4.21: Modelling differences of jet multiplicity between Sherpa nominal W/Z+jets samples and MadGraph alternative W/Z+jets samples.

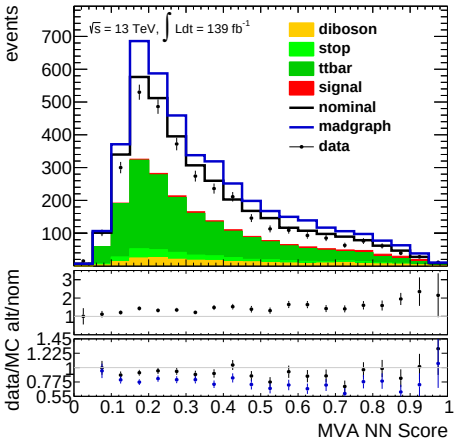
Fig. 4.21 shows the $R = 0.4$ jet multiplicity in the three signal regions. Data is compared to Monte Carlo in which the dominant W/Z+jets background is either modelled using the nominal Sherpa or the alternative MadGraph samples. It is evident that Sherpa generally overestimates while MadGraph underestimates the number of jets. It can also be seen that signal events are most likely to have a jet multiplicity of $N_j = 4$. This is in accordance with the resolved analysis strategy which consists of two tagging and two signal jets [3]. However, this also indicates that the RNN may learn that events with $N_j = 4$ are more signal-like while events with $N_j = 3$ or 5 are less likely to be signal. The effect of this on the RNN can be seen in Fig. 4.22.

Here the RNN score is shown for nominal Sherpa and alternative MadGraph background. MadGraph W/Z+jets background, generally having lower jet multiplicity, appears to be more likely to be misidentified as signal, i.e. assigned an RNN score close to 1. This can be seen in the rise of the alternative over nominal (i.e. MadGraph over Sherpa) ratio panels showing a positive slope towards higher values of the RNN score. It is evident that this effect is most pronounced in RNN($5j$), less so in RNN($4j$), and the least in NN. This is in accordance with the expectation since RNN($5j$) is able to discern the jet multiplicity in events with up to $N_j = 5$, while RNN($4j$) can only do so up until $N_j = 4$, and NN has no features that allow it to discern the jet multiplicity directly. As a consequence, the higher separation power of the RNN comes with the caveat of a stronger systematic modelling uncertainty. In the final statistical analysis this phenomenon is accounted for by introducing a systematic background shape modelling uncertainty. It is derived from the difference of the MVA score distributions when modelled with MadGraph or Sherpa (see Sec. 4.8).

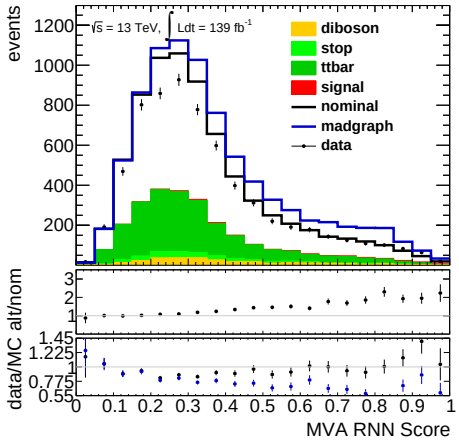
[3] Since there is no overlap removal between $R = 0.4$ and $R = 1.0$ jets, this also applies in the merged regime: While V_{had} is reconstructed as a single $R = 1.0$ signal jet J^{sig} instead two $R = 0.4$ signal jets $(jj)^{\text{sig}}$, the latter still count towards the jet multiplicity.



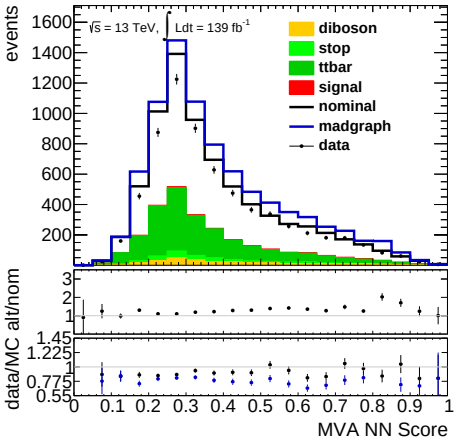
(a) merged HP SR, RNN($5j$)



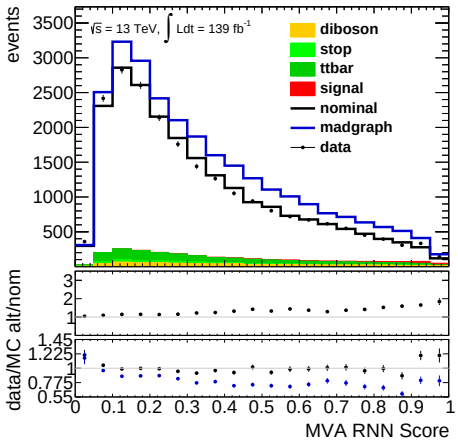
(b) merged HP SR, NN



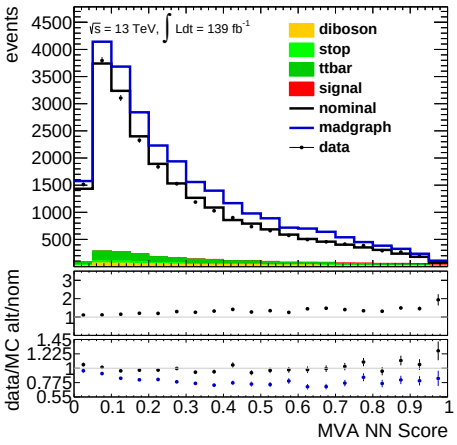
(c) merged LP SR, RNN($5j$)



(d) merged LP SR, NN



(e) resolved SR, RNN($5j$)



(f) resolved SR: NN score

Figure 4.22: Modelling differences of MVA scores between Sherpa nominal W/Z+jets samples and MadGraph alternative W/Z+jets samples (see Fig. A.8.23 for RNN($4j$)).

4.8 . Systematic Uncertainties

Various systematic uncertainties had to be considered for the analysis. In the statistical analysis each systematic uncertainty is treated as a nuisance parameter and is estimated with respect to the final discriminants. Wherever possible, the latest recommendations from the combined ATLAS performance groups were followed.

Each reconstructed object has several sources of uncertainties, each of which are evaluated separately. The uncertainty on $E_{\mathrm{T}}^{\mathrm{miss}}$ stems partly from a combination of the uncertainties on its constituents (dominantly jets) and partly from a soft term arising from unassociated tracks. The impact of small variations of the value of the track based transverse momentum $E_{\mathrm{T}}^{\mathrm{miss,track}}$ on the event yields was found to be negligible. Since it is only used in the preselection cuts based on this no systematic uncertainty had to be implemented for it.

Various systematic uncertainties concerning both small ($R = 0.4$) and large ($R = 1.0$) radius jets are estimated centrally within ATLAS and implemented using the usual tool chain. They are combined into a set of effective nuisance parameters (EffectiveNP). Additional systematics are estimated for pileup effects and for the forward jet vertex tagger (fJVT).

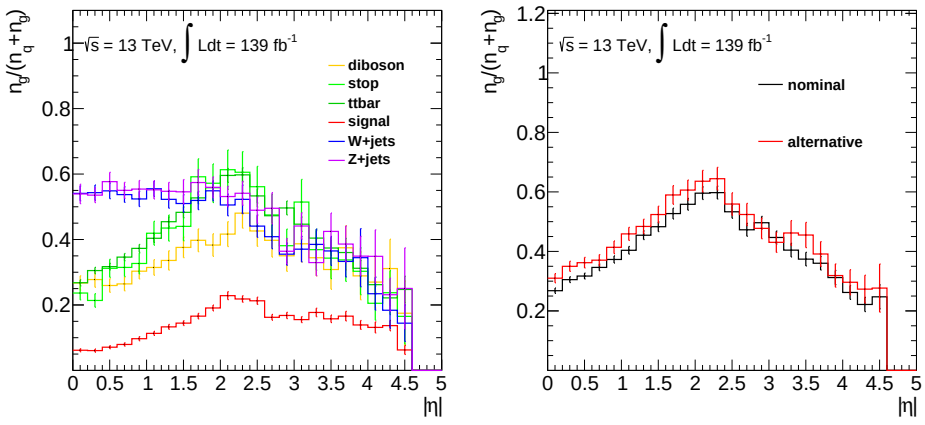
Jet flavour tagging with the DL1r $b-$tagger [133] is used in the tagging jet selection of the analysis. Uncertainties are implemented on the scale factors taking into account differences between the tagger performance in Monte Carlo and data.

The uncertainty in the combined 2015-2018 integrated luminosity ($\mathrm{LUMI\_2015\_2018}$) is 1.7% [131], obtained using the LUCID-2 detector for the primary luminosity measurements. The pileup reweighting accounts for the difference in the distribution of vertex multiplicity used to produce the Monte Carlo samples to that distribution measured for the actual data. The pileup reweighting scale factors as well as the associated uncertainties are provided by the ATLAS PileupReweighting analysis tool [145] and are implemented as $\mathrm{PRW\_DATASF}$. The uncertainty of the missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger has been estimated from contributions from statistics and the efficiency discrepancy between Monte Carlo samples.

Jet flavour response and composition uncertainties account for the different response in quark- or gluon-initiated jets. The response is derived centrally within ATLAS from differences using alternative Monte Carlo samples. The flavour composition uncertainty, corresponding to the relative abundance of quark- or gluon-initiated jets, has been derived within the analysis itself using the procedure described below. This is in contrast to the previous analysis [1] where a default value for the composition of 50% with a very conservative uncertainty of 100% was chosen.

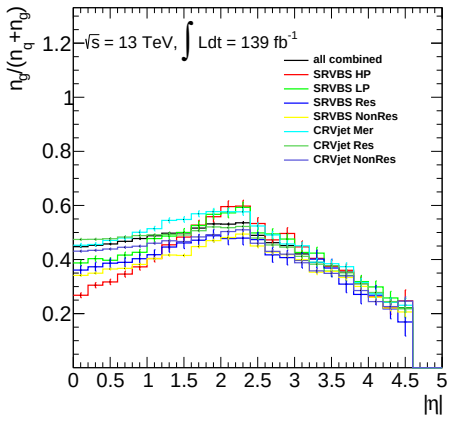
For the derivation of this uncertainty, the quark-gluon fraction was measured in the signal and background Monte Carlo samples of the analysis. The number of quark- and gluon initiated jets was calculated based on parton-level information.

For this purpose quark initiated jets are defined as jets that are initiated by any quark lighter than a b quark. Two dimensional histograms as a function of jet transverse momentum p_{T} and pseudorapidity η were produced for each sample in each analysis region. A variable bin size was chosen in such a way that all bins in all regions have a statistical uncertainty of less than 5%. One dimensional profiles in η of these two dimensional histograms can be seen in Fig. 4.23. The fractions are compared by sample (i.e. process), by alternative MC generator, and by analysis region. The latter two comparisons are shown only for the $t\bar{t}$ sample but similar comparisons for the other background and signal samples can be found in the appendix (Sec. A.6). The corresponding figures in the other analysis regions can also be found there.



(a) difference by sample

(b) difference by generator



(c) difference by region

Figure 4.23: Quark gluon fraction in the merged HP signal region as a function of jet pseudorapidity.

It is clear that the differences by sample are dominant while the differences by generator are of the same order as the statistical uncertainties per bin. The differences by region are slightly higher and should therefore also be considered. The nominal value for the quark gluon fraction used in the implementation of the corresponding systematic was calculated for each sample individually from the combination of all analysis regions. The differences in region were taken into account by taking the maximum difference in the outcome between any two regions. Where alternative generators were available, an additional generator uncertainty was calculated from the difference of the fraction with nominal and alternative sample. The total uncertainty used in the systematic implementation was calculated by combining generator and region uncertainty bin by bin:

$$\sigma_{\text{total}} = \sqrt{\sigma^2_{\text{generator}} + \sigma^2_{\text{region}}}. \tag{4.5}$$

The resulting two dimensional histogram of quark-gluon fractions as well as the associated uncertainties that were eventually used for the flavour composition systematic uncertainty in the final fit can be found in the appendix Sec. A.6, Fig. A.6.15 and Fig. A.6.16. From them it is evident that the total uncertainty is below 20% in each bin and for each Monte Carlo sample. This is a significant improvement over the very conservative default value of 100% that was used in the previous analysis [1].

4.9 . Statistical Interpretation

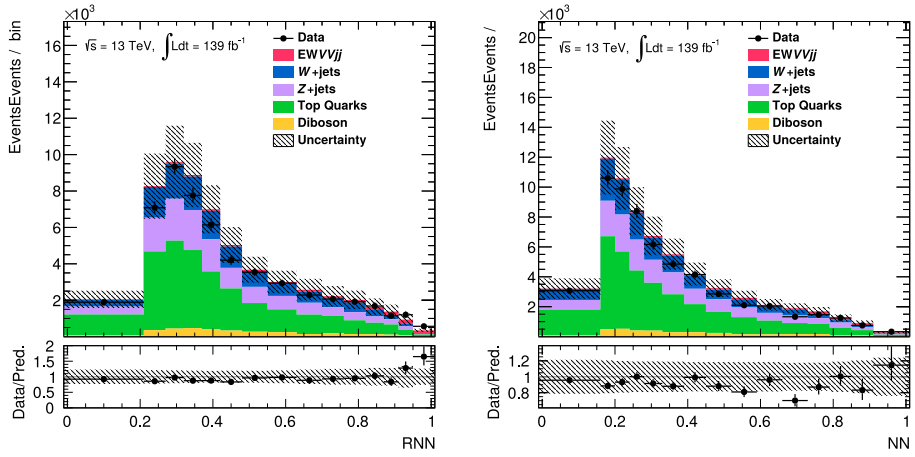
A statistical interpretation of the analysis has been performed in the form of a binned maximum likelihood fit of the Monte Carlo simulated signal and background with respect to the observed data. The likelihood

$$L(\mu^{\text{semileptonic}}_{\text{EW } VV+jj}, \vec{\theta}) = L_{\text{EML}}(\mu^{\text{semileptonic}}_{\text{EW } VV+jj}, \vec{\alpha}, \vec{\gamma}, \vec{\tau}) L_{\text{syst}}(\vec{\alpha}) L_{\text{stat}}(\vec{\gamma}) \tag{4.6}$$

is a function of the parameter of interest $\mu^{\text{semileptonic}}_{\text{EW } VV+jj} = \frac{\sigma_{\text{obs}}}{\sigma_{\text{SM}}}$, called the signal strength, and a set of nuisance parameters $\vec{\theta} = \{\vec{\alpha}, \vec{\gamma}, \vec{\tau}\}$. The full list of nuisance parameters considered for the fit can be found in the appendix Sec. A.9. The signal strength is the ratio of the observed process cross-section over the expected cross-section from Monte Carlo (MC) simulation. The extended maximum likelihood term L_{EML} is constructed as a product of Poisson terms over all n_{bins} bins of histograms in the five analysis regions:

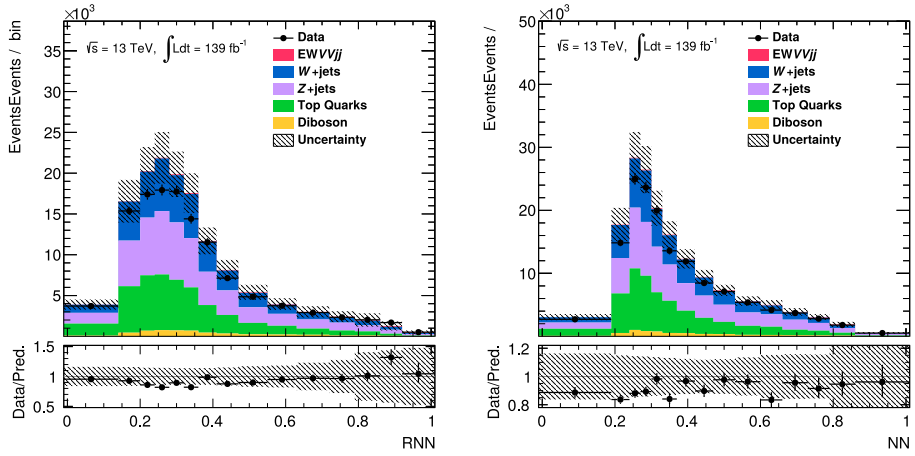
$$L_{\text{EML}} = \prod_{i=1}^{n_{\text{bins}}} \text{Poisson}\left(N_i | \mu^{\text{semileptonic}}_{\text{EW } VV+jj} s_i(\vec{\alpha}) + b_i(\vec{\alpha}, \vec{\gamma}, \vec{\tau})\right) \tag{4.7}$$

where s_i and b_i are the expected number of signal and background events from Monte Carlo simulation in a specific bin and N_i are the observed number of events in data. In the analysis signal regions the fitted distributions are the MVA score (Fig. 4.24) while in the control regions the mass of the tagging jet system $m(jj)^{\text{tag}}$ was chosen (Fig. 4.25).



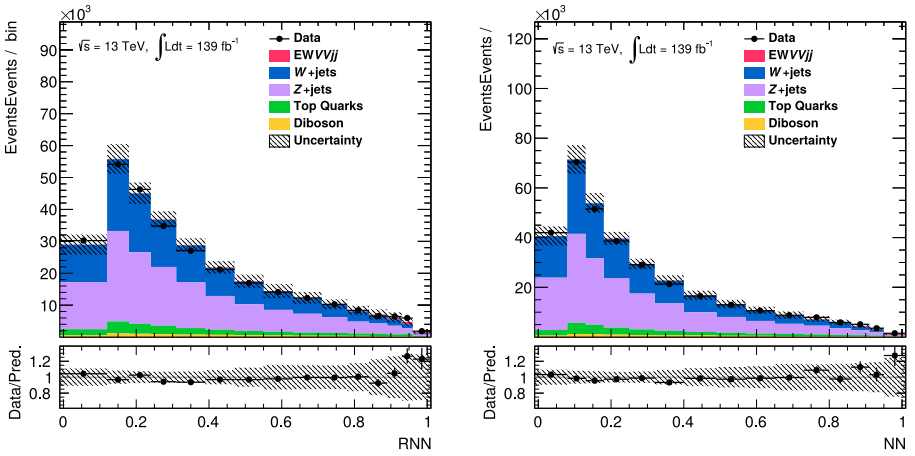
(a) RNN($5j$) merged HP SR

(b) NN merged HP SR



(c) RNN($5j$) merged LP SR

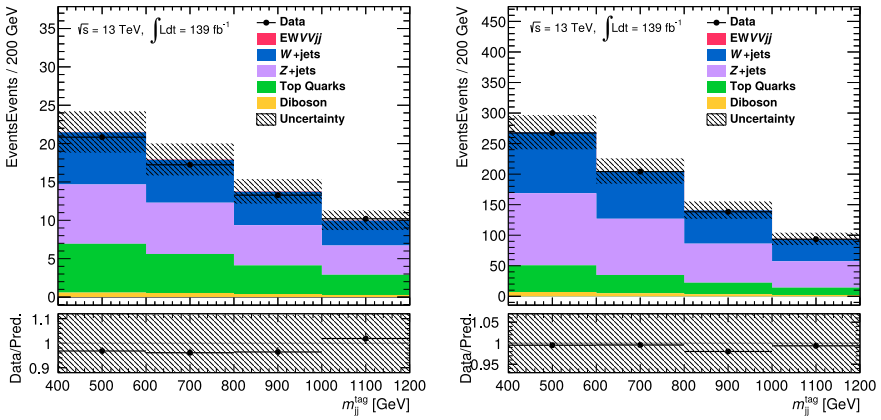
(d) NN merged LP SR



(e) RNN($5j$) resolved SR

(f) NN resolved SR

Figure 4.24: Signal region prefit distributions of either NN or RNN($5j$) discriminant used in the final fit of the statistical interpretation (see Fig. A.8.24 for RNN($4j$)).



(a) merged CR (b) resolved CR

Figure 4.25: Control region prefit distributions of the tagging jet mass $m(jj)^{tag}$ used in the final fit of the statistical interpretation.

The binning was optimised with the Transformation-D algorithm (further described in [146]). $\vec{\alpha}$ are nuisance parameters (NP) corresponding to systematic uncertainties (see Sec. A.9 for the complete list of NPs considered in the fit), $\vec{\gamma}$ parametrise the statistical uncertainty of the number of Monte Carlo events in each bin, and $\vec{\tau}$ are floating normalisations of the Monte Carlo background. The n_{syst} systematic uncertainties $\vec{\alpha}$ are generally constrained by Gaussian terms

$$L_{\text{syst}} = \prod_{j=1}^{n_{\text{syst}}} \text{Gauss}(\alpha_j|0,1) \tag{4.8}$$

where $\alpha_j \to 0$ corresponds to the nominal case and $\alpha_j \to \pm 1$ corresponds to a variation of a standard deviation, i.e. $\pm 1\ \sigma$. An exception to this are those Monte Carlo background normalisations that are not left floating. They are also counted towards the systematic uncertainties $\vec{\alpha}$ but are constrained by log-normal terms instead. The n_{bins} statistical uncertainties $\vec{\gamma}$ corresponding to each individual bin throughout all analysis regions are constrained by Gamma functions

$$L_{\text{stat}} = \prod_{k=1}^{n_{\text{bins}}} \text{Gamma}\left(\gamma_k|\left(\frac{b_i}{\sigma_{b_i}}\right)^2, n_i - 1\right). \tag{4.9}$$

The norms of the two dominant background processes W+jets and Z+jets were left floating with a single parameter τ acting on both in the same way. This is implemented separately for merged and resolved analysis regions, resulting in a total of two floating norms. This approach has been chosen because the introduction of more floating norms, either for the other background processes or separate for W+jets and Z+jets, did not yield a converging fit result. The reason for this lies in the very similar shape of W+jets and Z+jets background

which results in a degeneracy. This similarity can be seen in the prefit distributions shown in Fig. 4.24 and 4.25. Since there is no dedicated control region for the top background, the fit would not be able to constrain a floating norm properly on this background either. This is only the case for the 0-lepton channel-only fit which is presented here. When performing a combined fit of all analysis channels (including 1-, and 2-lepton) it is possible to introduce more floating norms. This is the case because the 1-lepton channel is dominated by W+jets while the 2-lepton channel is dominated by Z+jets background. Hence their corresponding V+jets control regions can be used to constrain these backgrounds separately. The top background can be constrained by a dedicated top control region in the 1-lepton channel which has a higher yield of this background than the other analysis channels.

Constraints on the non-floating Monte Carlo background normalisations were derived from the difference in prefit event yield of the nominal MC samples compared to alternative samples. They can be seen in Tab. 4.8. Based on this, a constraint of 40% and 30% was chosen for the diboson and $t\bar{t}$ backgrounds, respectively. For the single top background the same (30%) constraint as for $t\bar{t}$ was chosen.

$t\bar{t}$ yields	merged SR HP	merged SR LP	resolved SR	merged CR	resolved CR
nominal	1533	2009	1284	4361	21802
alternative	1064	1678	941	3997	16184
(nom-alt)/nom	31%	16%	27%	8%	26%
diboson yields	merged SR HP	merged SR LP	resolved SR	merged CR	resolved CR
nominal	208	275	586	494	4376
alternative	142	164	512	260	2652
(nom-alt)/nom	32%	40%	13%	47%	39%

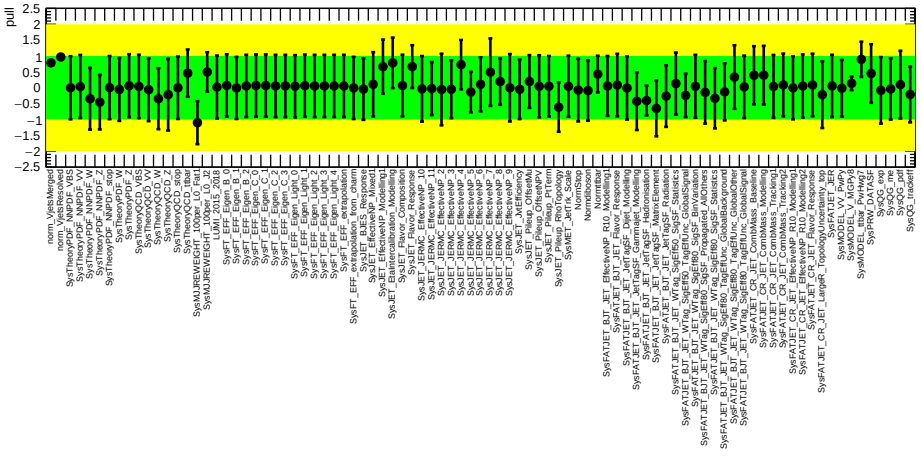
Table 4.8: Difference in prefit event yields of the $t\bar{t}$ and diboson background when simulated with nominal and alternative Monte Carlo samples (defined in Sec. 4.2).

After the fit has been thoroughly studied throughout the unblinding procedure (described in appendix A.7) the final fits were performed on data. In the following, the behaviour of various nuisance parameters are discussed. If not explicitly stated otherwise, all of the following considerations and figures correspond to the nominal case in which RNN($5j$) is used as MVA. Similar fits on RNN($4j$) and NN are used for comparison when necessary.

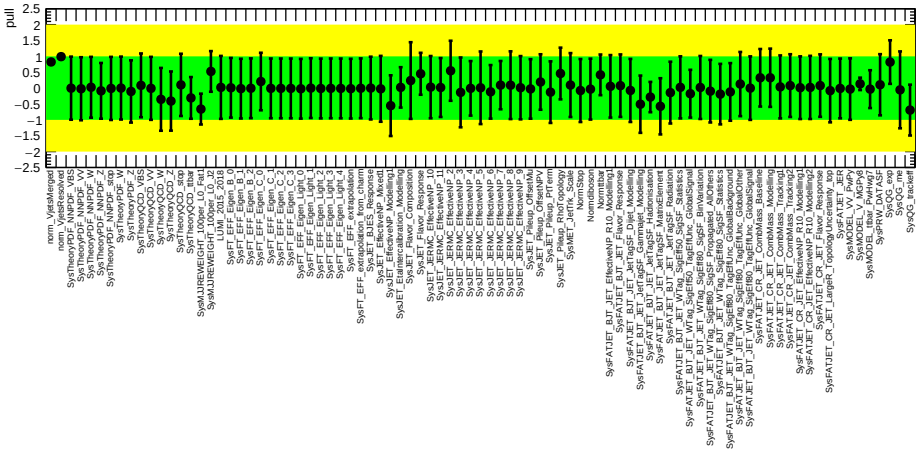
The pull of a nuisance parameter (NP) is defined as

$$p = \frac{\hat{\theta} - \theta}{\sigma_\theta^2}, \tag{4.10}$$

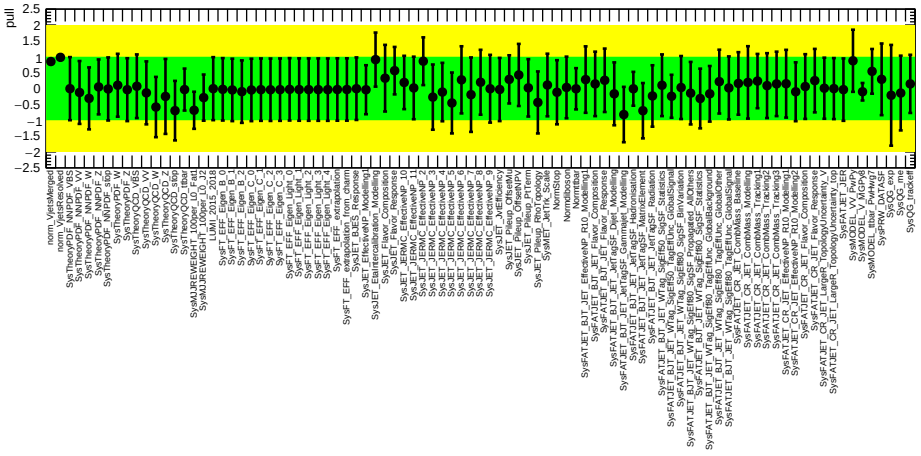
where θ and $\hat{\theta}$ are the pre- and postfit values of the NP and σ_θ^2 is the variance of the corresponding uncertainty. The resulting pulls of all nuisance parameters considered in the fits are shown in Fig. 4.26



(a) RNN($5j$)



(b) RNN($4j$)



(c) NN

Figure 4.26: Pulls of the final unblinded fits for the three different options of MVA.

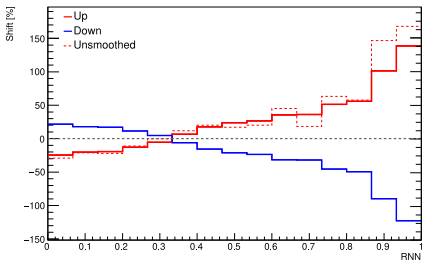
Pulls with significant deviations from zero indicate that the postfit value of such NP is significantly different from its prefit value. Nuisance parameters that are visibly under- or overconstrained also need further investigation. An underconstrained NP has a postfit uncertainty much larger than its prefit uncertainty. This is indicated by uncertainty bars much larger than ± 1. This can be the case if the uncertainty with respect to the NP has been significantly underestimated which may lead to too small uncertainties for the parameter of interest in the final fit result. Overconstrained NPs have a much smaller post- than prefit uncertainty which is indicated by uncertainty bars much smaller than ± 1. This can be caused by the fit being able to constrain the uncertainty parametrised by the NP. Hence, this is not necessarily an indication of a problematic fit but the source of the constraint needs to be understood in each case.

The modelling systematic for the V+jets background $\mathrm{SysMODEL\_V\_MGPy8}$ is the most constrained. This is expected especially for the RNN($5j$) MVA which is very sensitive to differences in the modelling of this background. This was discussed in Sec. 4.7. The corresponding pulls for the other MVAs substantiate this since they show weaker constraints. This effect can be seen directly in the prefit $\pm 1\sigma$ up/down variation of this nuisance parameter given in terms of a shift in each bin. Fig 4.27 shows this for the Z+jets background. The corresponding shapes for the W+jets background are very similar and can be found in the appendix Fig. A.5.8.

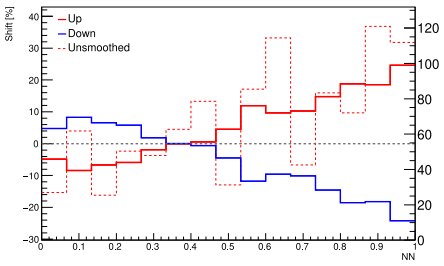
It is evident that this uncertainty has the highest impact in the high purity (HP) signal region and grows towards larger MVA scores, which coincides with a higher abundance of signal events, and therefore a larger impact on the parameter of interest. While for the NN the shift is in a range of 20% in the right-most bin, it is at 40% for the RNN($4j$) and nearly 150% for RNN($5j$). This is in accordance with the observation that was made in Sec. 4.7: Due to the high susceptibility of the RNN($5j$) on the mismodelled jet multiplicity it shows a higher effect when compared to RNN($4j$) and NN which have less information with respect to this parameter. When deriving this systematic from the difference between nominal and alternative Monte Carlo sample, the resulting variation is one-sided. The systematic was then implemented in a symmetrised way, mirroring the single variation around the nominal value. As a result this might be an overly conservative approach. For example, while a 150% up variation may be conceptually sensible, a 150% down variation stemming from such symmetrisation would in its most extreme form correspond to a negative number of V+jets events in the right-most bins. Because of this it is expected that the associated nuisance parameter is heavily constrained by the fit.

Another notably constrained pull is the uncertainty associated to the tagging jet mass $m(jj)^{\text{tag}}$ reweighting: $\mathrm{SysMJJREWEIGHT\_100per\_L0\_Fat1}$ for the reweighting factors in the merged regime and $\mathrm{SysMJJREWEIGHT\_100per\_L0\_J2}$ for the resolved regime. Their shapes are shown in Fig. 4.28 for the

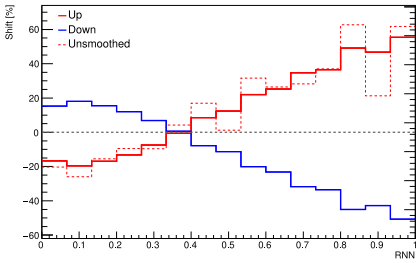
Z+jets background and A.5.9 (appendix) for the W+jets background. This was expected due to the over-conservative approach when assigning these uncertainties. A variation of 100% with respect to the nominal values of the reweighting factors was used.



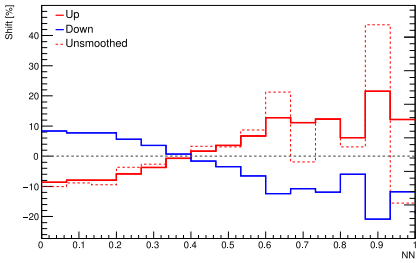
(a) RNN($5j$) merged HP SR



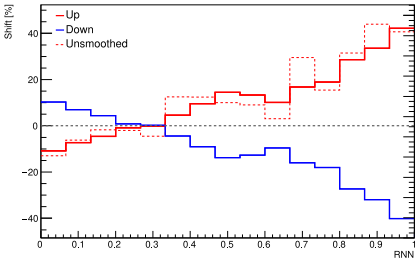
(b) NN merged HP SR



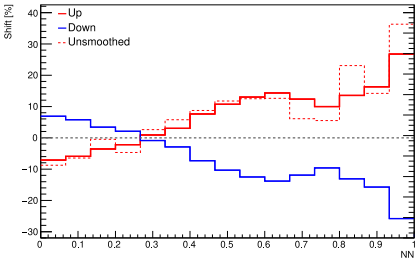
(c) RNN($5j$) merged LP SR



(d) NN merged LP SR

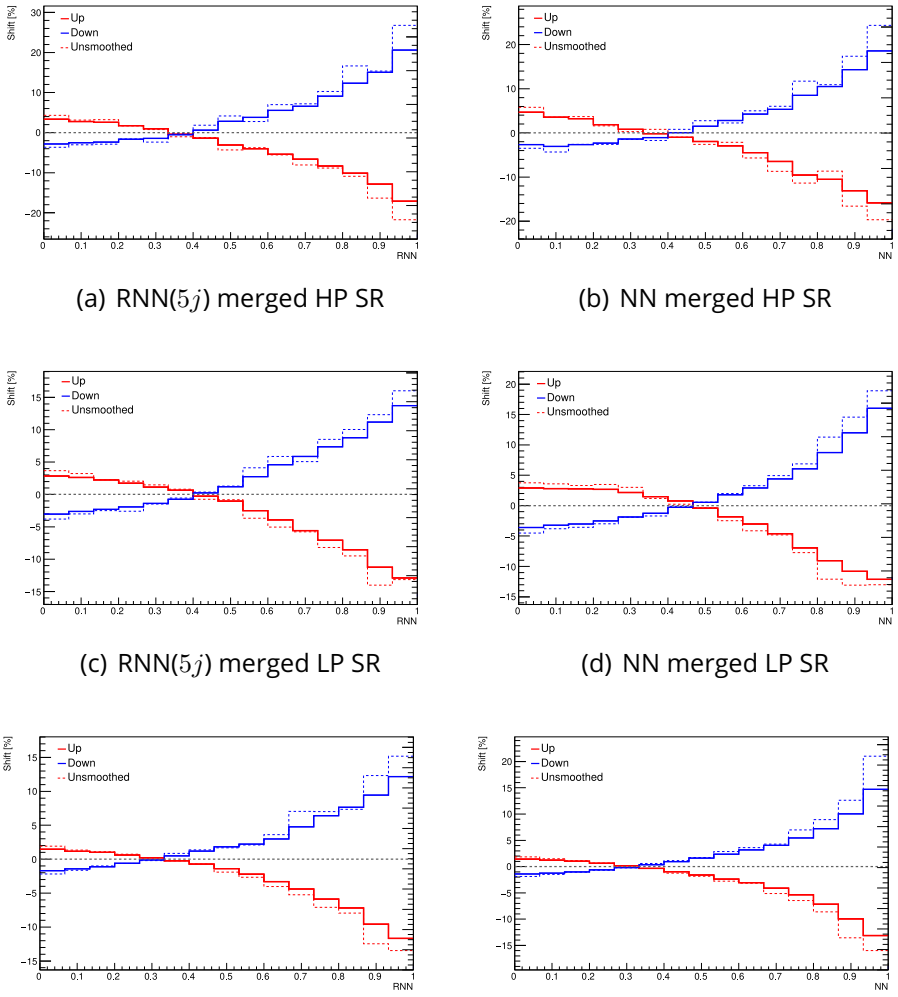


(e) RNN($5j$) resolved SR



(f) NN resolved SR

Figure 4.27: $\pm 1\ \sigma$ variation of the V+jets modelling systematic uncertainty $\mathrm{SysMODEL\_V\_MGPy8}$ on the Z+jets background sample for the NN and RNN($5j$) discriminant (see Fig. A.8.25 for RNN($4j$)).



(a) RNN($5j$) merged HP SR

(b) NN merged HP SR

(c) RNN($5j$) merged LP SR

(d) NN merged LP SR

(e) RNN($5j$) resolved SR

(f) NN resolved SR

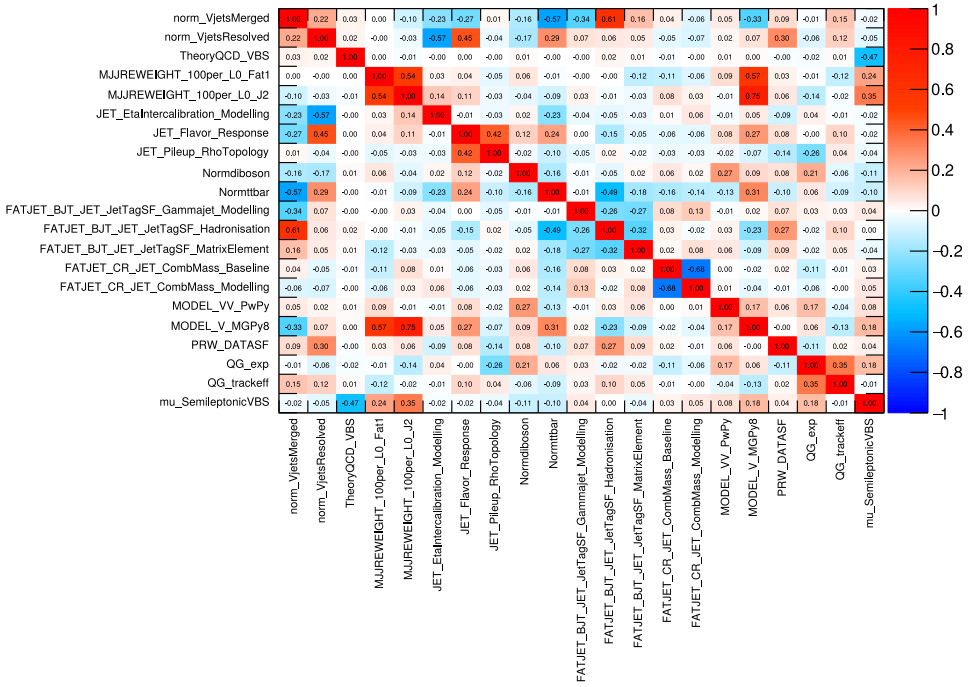
Figure 4.28: ± 1 σ variation of the $m(jj)^{\text{tag}}$ reweighting systematic uncertainty $\mathrm{SysMJJREWEIGHT\_100per\_L0\_Fat1}$ ($\mathrm{SysMJJREWEIGHT\_100per\_L0\_J2}$) in the merged (resolved) regime on the Z+jets background sample for the NN and RNN($5j$) discriminant (see Fig. A.8.26 for RNN($4j$).)

An alternative would have been to use the uncertainties associated to the linear fits that were used to derive these factors. They can be found in Tab. 4.6 and are of the order of $10 - 20\%$. In the merged regime this NP shows a relatively strong downward pull while the corresponding NP in the resolved regime is slightly pulled upward. This is in accordance with the relatively high correlation factor of these

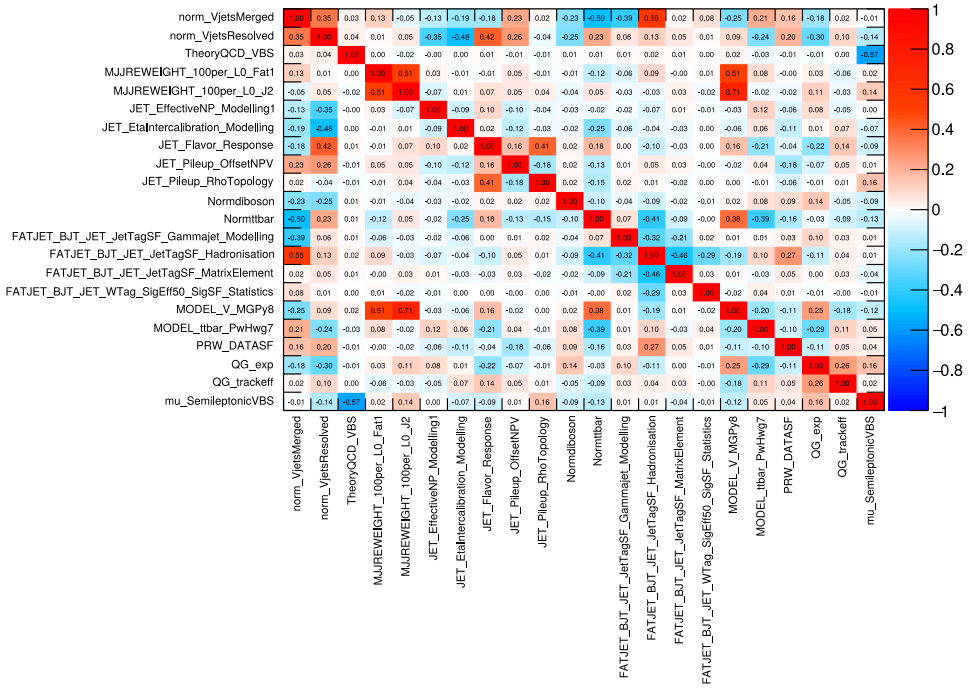
NPs with each other. The correlation matrix in Fig. 4.29 shows it to be 51%. They are also highly correlated to the V+jets background modelling uncertainty $\mathrm{SysMODEL\_V\_MGPy8}$ which was discussed further above. The correlation factors to it are 51% in the merged and 71% in the resolved regime. Since both types of nuisance parameter act exclusively on the W/Z+jets backgrounds and are especially strong in the high MVA score regions it is understandable that they are highly correlated. Besides their relatively high constraints, the $m(jj)^{\mathsf{tag}}$ reweighting systematics show relatively strong pulls in the RNN($5j$) fit: The merged regime NP is pulled by approximately 1σ down, the resolved regime NP is pulled up by a lesser extent. This is much less pronounced in the RNN($4j$) and NN fits. It may be assumed that this is in part due to their high correlation to the V+jets modelling NPs which are much stronger in RNN($5j$) as well. It should also be noted that the $m(jj)^{\mathsf{tag}}$ reweighting uncertainties themselves have a slightly larger variation on the RNN($5j$) (Fig. 4.28) than on the other MVAs.

Another pair of slightly pulled and strongly constrained nuisance parameters that have a high correlation (39%) with each other are the normalisation (Normttbar) and modelling uncertainty ($\mathrm{MODEL\_ttbar\_PwgHwg7}$) of the $t\bar{t}$ background. The high correlation is understandable since they both concern the same background. While the former parametrises the normalisation, the latter concerns the shape. The fact that the normalisation NP is significantly constrained suggests that the 30% prior derived from the yield difference in nominal and alternative samples may have been too conservative. It should be noted that Normttbar has a high correlation to the floating norms of the dominant V+jets background: 50% in the merged regime ($\mathrm{norm\_VjetsMerged}$) and 23% in the resolved regime ($\mathrm{norm\_VjetsResolved}$).

As expected, the highest correlation (51%) of the parameter of interest ($\mathrm{mu\_SemileptonicVBS}$) is to the QCD scale theory uncertainty on the signal Monte Carlo Sample ($\mathrm{TheoryQCD\_VBS}$). It is followed by a 16% correlation of the ρ-topology term of the $R = 0.4$ jet pileup calibration ($\mathrm{JET\_Pileup\_RhoTopology}$). This is also reflected by the fact that these NPs are in place 1 and 2 of the nuisance parameter ranking (Fig. 4.30) which ranks all nuisance parameters by their impact on the final result of the parameter of interest in the fit. The pileup term ranks much lower for the RNN($4j$) and NN fits. The RNN($5j$) takes the five leading-p_{T} jets of the event as input features while the signal process is only excepted to produce four jets (two tagging, two signal jets) at most. The additional jets the RNN($5j$) relies on are therefore very likely to be from pileup. Following this assumption, it is understandable that the RNN($5j$) is very susceptible to this nuisance parameter while the other MVAs are more robust against it.

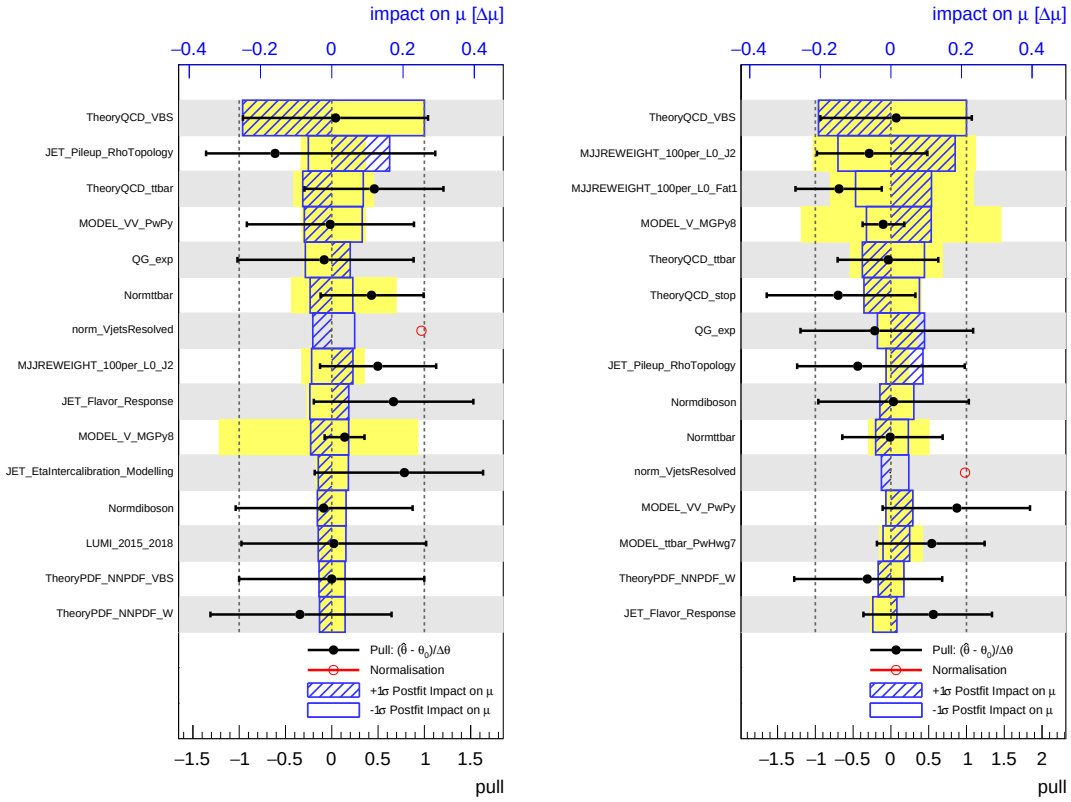


(a) NN



(b) RNN($5j$)

Figure 4.29: Correlations of nuisance parameters in the final unblinded fit for either the NN or RNN($5j$) MVA (see Fig. A.8.29 for RNN($4j$)).

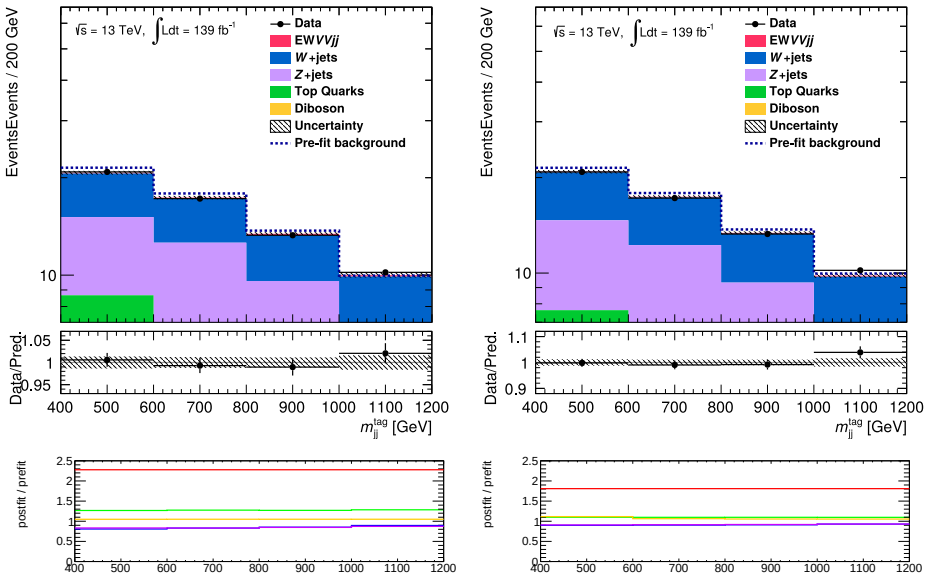


(a) RNN($5j$)

(b) NN

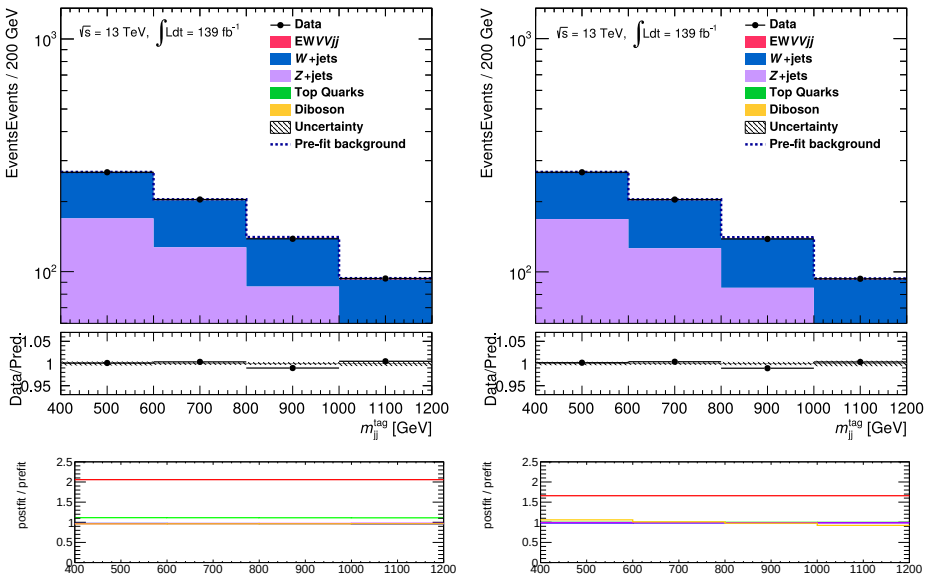
Figure 4.30: Nuisance parameter (NP) ranking of the final unblinded fits using either NN or RNN($5j$). The impact of each individual NP on the parameter of interest μ is determined by fixing the NP to a value corresponding to $\pm 1\sigma$ of its uncertainty and redoing the fit with all but that NP (see Fig. A.8.30 for RNN($4j$)).

Fig. 4.31 and 4.32 show the postfit distributions of the fitted discriminants (MVA score in signal-, $m(jj)^{\text{tag}}$ in control regions) in all analysis regions. The panels below each plot show the pre-to-postfit ratio of the number of events from each Monte Carlo sample in each bin. It is evident that the signal shows the highest pre-to-postfit difference with respect to normalisation. The strongest shape in these ratios can be seen for the W+jets and Z+jets backgrounds in the merged HP SR. This is an effect of the shapes of the V+jets modelling and (to a lesser extent) the $m(jj)^{\text{tag}}$ reweighting systematics (Fig. 4.27 and 4.28) which give the fit enough freedom for this.



(a) RNN($5j$) merged CR

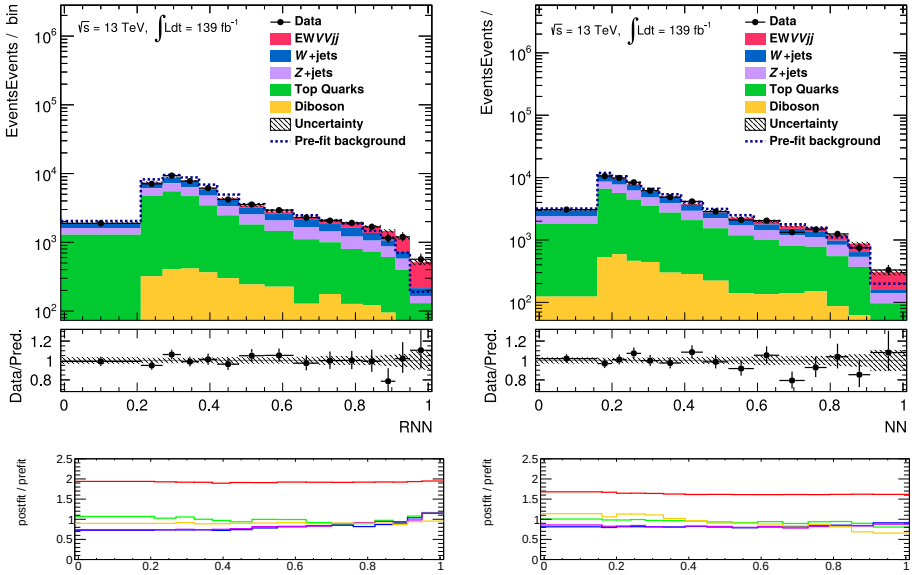
(b) NN merged CR



(c) RNN($5j$) resolved CR

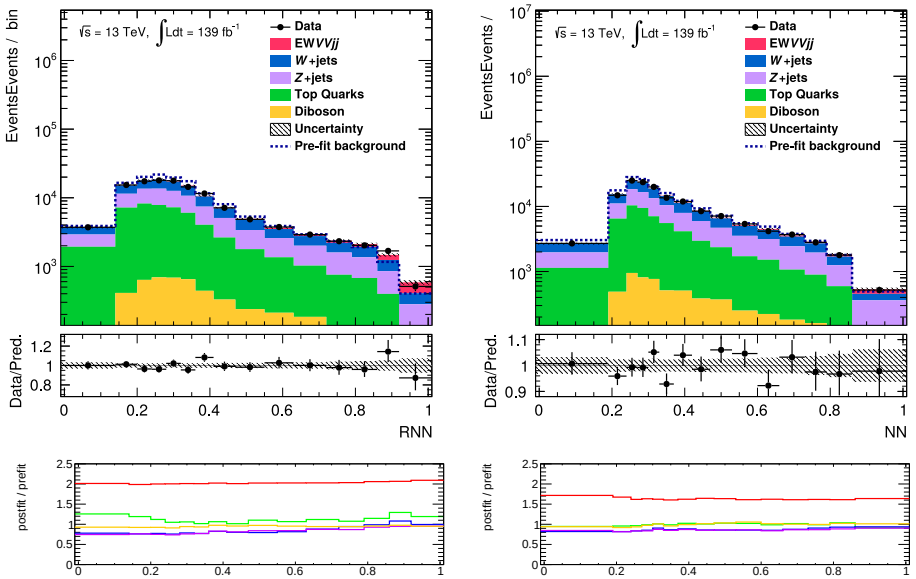
(d) NN resolved CR

Figure 4.31: Control region postfit distributions of the tagging jet mass $m(jj)^{tag}$ after performing fits using either the NN or RNN($5j$) discriminant (see Fig. A.8.28 for RNN($4j$)).



(a) RNN($5j$) merged HP SR

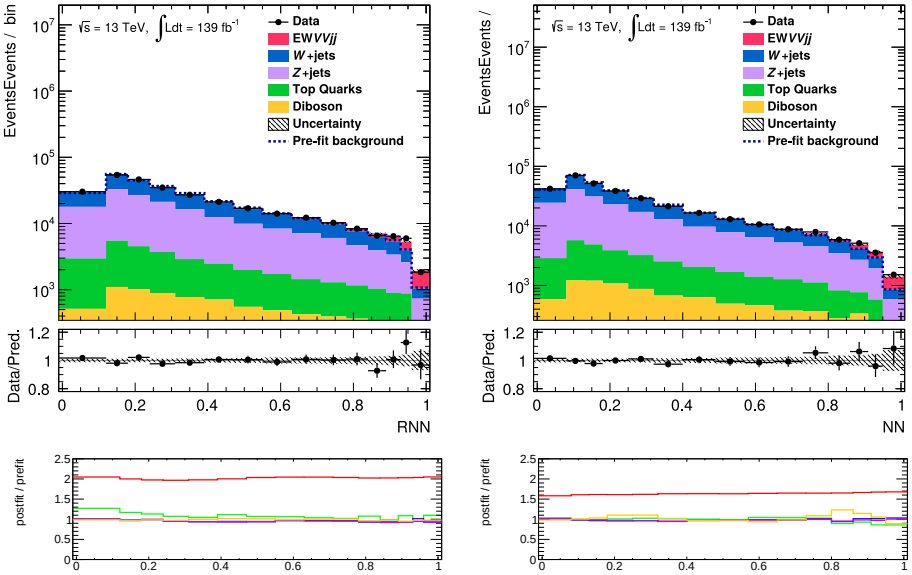
(b) NN merged HP SR



(c) RNN($5j$) merged LP SR

(d) NN merged LP SR

Figure 4.32: Signal region postfit distributions after performing fits using either the NN or RNN($5j$) discriminant (continued on next page) (see Fig. A.8.27 for RNN($4j$)).



(a) RNN($5j$) resolved SR (b) NN resolved SR

Figure 4.32: Continuation from previous page.

The relatively large pre-to-postfit difference of the signal normalisation is also reflected in the postfit observed signal strength $\mu^{\text{semileptonic}}_{\text{EW } VV+jj} = 2.02 \pm 0.46$ (see Tab. 4.9) being larger than one. While the Standard Model value of one is outside the 1σ uncertainty interval of this value, this does not constitute evidence for a significant deviation from the Standard Model expectation. Due to the signal strength being larger than one, the observed significance (see Tab. 4.9) with a deviation of $5.65\ \sigma$ from the background-only hypothesis is larger than the expected significance of $3.03\ \sigma$. Due to its better separation power and consequentially better sensitivity this effect is largest in the RNN($5j$) compared to the other two MVA methods. All three methods consistently show signal strengths above one confirming the result obtained with the RNN($5j$).

(a) significance

	expected		observed
	pre-fit	post-fit	
RNN(5j)	3.04	3.03	5.65
RNN(4j)	2.77	2.80	4.86
NN	2.57	2.73	4.27

(b) signal strength

	pre-fit expected	post-fit observed
RNN(5j)	1.00 ± 0.36	2.02 ± 0.46
RNN(4j)	1.00 ± 0.40	1.94 ± 0.48
NN	1.00 ± 0.42	1.63 ± 0.45

Table 4.9: Significance [σ] and signal strength from full-range 1POI fits for $\mu^{\text{semileptonic}}_{\text{EW } VV+jj}$.

Separate fits were performed to investigate if the small observed deviation from the expected value is dominated either by the merged or resolved regime. For this

purpose, instead of a combined parameter of interest $\mu^{\text{semileptonic}}_{\text{EW } VV+jj}$, two separate POIs $\mu^{\text{semileptonic}}_{\text{EW } VV+jj\ merged}$ and $\mu^{\text{semileptonic}}_{\text{EW } VV+jj\ resolved}$ were fitted simultaneously. While the former concerns the distributions in the merged regime (merged HP and LP SR, and merged CR) the latter concerns the distributions in the resolved regime (resolved SR and CR). The resulting signal strength and significance can be seen in Tab. 4.10. It is evident that both regimes contribute in a similar way.

(a) significance

	expected		observed
	pre-fit	post-fit	
$\mu^{\text{semileptonic}}_{\text{EW VV+jj},\ merged}$			
RNN(5j)	2.28	2.21	3.90
RNN(4j)	1.95	1.96	2.59
NN	1.56	1.60	1.44
$\mu^{\text{semileptonic}}_{\text{EW VV+jj},\ resolved}$			
RNN(5j)	2.34	2.38	4.72
RNN(4j)	2.07	2.24	4.54
NN	2.15	2.22	4.43

(b) signal strength

	pre-fit expected	post-fit observed
$\mu^{\text{semileptonic}}_{\text{EW VV+jj},\ merged}$		
RNN(5j)	1.00 ± 0.48	1.88 ± 0.57
RNN(4j)	1.00 ± 0.55	1.35 ± 0.57
NN	1.00 ± 0.66	0.88 ± 0.62
$\mu^{\text{semileptonic}}_{\text{EW VV+jj},\ resolved}$		
RNN(5j)	1.00 ± 0.46	2.14 ± 0.56
RNN(4j)	1.00 ± 0.50	2.36 ± 0.58
NN	1.00 ± 0.50	2.21 ± 0.62

Table 4.10: Significance [σ] and signal strength from full-range 2POI fits for $\mu^{\text{semileptonic}}_{\text{EW VV+jj},\ merged}$ and $\mu^{\text{semileptonic}}_{\text{EW VV+jj},\ resolved}$.

Tab. 4.11 shows a breakdown of the total uncertainty on $\mu^{\text{semileptonic}}_{\text{EW } VV+jj}$ into various categories. It is evident that the analysis is dominated by systematic uncertainties (80%) but the statistical uncertainty on the data is not negligible (20%). The highest contribution out of all systematic uncertainties is expectedly given by the theory uncertainty of the QCD scale of the Monte Carlo simulated signal process. The corresponding uncertainty for the background plays a significant but much smaller role. The $R = 0.4$ jets dominate the experimental uncertainties. While the $m(jj)^{\text{tag}}$ reweighting uncertainty plays a small role for the baseline RNN($5j$) MVA, it plays a larger role for RNN($4j$), and especially NN. This may be the case because it acts on the same backgrounds (W/Z+jets) and has a similar shape to the V+jets modelling systematic (compare Fig. A.5.8 with Fig. A.5.9) which is much stronger for RNN($5j$) than for the other MVAs.

Tab. 4.12 shows the signal strength and significance obtained from various previous analyses searching for similar processes containing vector boson scattering contributions. The result obtained in this analysis should be set into context with the results that are shown there. For the EW $WWjj$ same sign analysis [147] two separate values for the expected significance are given: 4.4^{Sherpa} and 6.5^{Powheg}. The former was derived with the nominal signal Monte Carlo sample produced with Sherpa. For the latter an alternative sample produced with Powheg was chosen. The value obtained with Sherpa shows a deviation from the observed significance of $6.5\ \sigma$ of a similar order to the deviation found in the analysis presented in this

thesis. The value obtained with the alternative Powheg generator on the other hand shows a negligible deviation from the observed value. In an eventual future study it may be of interest to investigate the source of the difference with respect to the different generators of the signal sample.

In summary, the fit on the RNN($5j$) shows overall higher sensitivity to the process under study than the fits to the control MVAs (RNN($4j$) and NN). This is an effect of its higher separation power caused by the use of more features. The impact of various systematic uncertainties on the MVAs was investigated. While it is clear that RNN($5j$) is significantly more susceptible to some variations, it has been shown that none of the systematic uncertainties are strongly underestimated and all nuisance parameters of the fit are under control. Here this was demonstrated on the final unblinded fit results, during the analysis itself however, this was thoroughly studied before unblinding and throughout the various steps of this procedure as they are described in the appendix A.7. This justifies the use of RNN($5j$) rather then RNN($4j$) or NN as the nominal MVA of the analysis in order to achieve a higher significance.

	RNN($5j$)				RNN($4j$)				NN			
	Expected		Observed		Expected		Observed		Expected		Observed	
Total	0.365	-	0.456	-	0.397	-	0.477	-	0.425	-	0.460	-
Systematic	0.313	73%	0.408	80%	0.340	73%	0.427	80%	0.372	77%	0.411	80%
Statistical	0.189	27%	0.203	20%	0.204	26%	0.212	20%	0.205	23%	0.206	20%
Normalisations												
Floating (V+jets) normalisations	0.050	2%	0.066	2%	0.028	0%	0.106	5%	0.043	1%	0.033	1%
Normalisations with priors	0.082	5%	0.079	3%	0.087	5%	0.113	6%	0.117	8%	0.074	3%
Experimental Uncertainties												
$m(jj)^{\text{tag}}$ reweighting	0.059	3%	0.067	2%	0.019	11%	0.133	8%	0.173	17%	0.169	14%
$E_{\text{T}}^{\text{miss}}$	0.018	0%	0.024	0%	0.018	0%	0.068	2%	0.010	0%	0.009	0%
$R = 0.4$ jets	0.126	12%	0.141	10%	0.106	7%	0.168	12%	0.103	6%	0.089	4%
$R = 1.0$ jets	0.034	1%	0.039	1%	0.041	1%	0.108	5%	0.058	2%	0.069	2%
Flavour tagging	0.028	1%	0.039	1%	0.023	0%	0.019	0%	0.012	0%	0.014	0%
Track uncertainty	0.029	1%	0.075	3%	0.037	1%	0.059	2%	0.058	2%	0.104	5%
Pileup reweighting	0.009	0%	0.017	0%	0.019	0%	0.010	0%	0.024	0%	0.013	0%
Luminosity	0.022	0%	0.039	1%	0.022	0%	0.073	2%	0.024	0%	0.031	0%
MC statistical	0.100	8%	0.105	5%	0.116	9%	0.108	5%	0.120	8%	0.117	6%
Background Modelling												
V+jets modelling	0.060	7%	0.055	1%	0.049	2%	0.022	0%	0.077	3%	0.092	4%
Diboson modelling	0.073	3%	0.083	3%	0.006	0%	0.016	0%	0.044	1%	0.044	1%
$t\bar{t}$ modelling	0.031	4%	0.023	0%	0.015	0%	0.052	1%	0.041	1%	0.037	1%
Theory Uncertainties												
Signal QCD scale	0.128	12%	0.213	22%	0.127	10%	0.219	21%	0.130	9%	0.212	21%
Signal PDF	0.019	0%	0.038	1%	0.019	0%	0.073	2%	0.019	0%	0.032	0%
Background QCD scale	0.076	4%	0.092	4%	0.093	6%	0.099	4%	0.122	8%	0.122	7%
Background PDF	0.037	1%	0.040	1%	0.050	2%	0.058	1%	0.049	1%	0.047	1%

Table 4.11: Symmetrised uncertainty σ_μ from from each source in the best fit of the signal strength parameter $\mu_{\text{EW}\,VV+jj}^{\text{semileptonic}}$. The fractional uncertainty (in percent) is given as one minus the ratio of the total uncertainty of the fit without and with the corresponding source of uncertainty.

process	signal strength μ	signif. σ_{obs} (σ_{exp})	int. lumin. [fb^{-1}]	publication
EW $VVjj$ semilep	$1.05^{+0.42}_{-0.40}$	2.7 (2.5)	36.1	ATLAS [148]
EW $VVjj$ semilep, 0lep	$2.47^{+1.33}_{-1.22}$	1.43(1.35)	36.1	ATLAS [148]
EW $VVjj$ semilep, 1lep	$0.33^{+0.53}_{-0.52}$	0.53(1.77)	36.1	ATLAS [148]
EW $VVjj$ semilep, 2lep	$1.97^{+0.83}_{-0.77}$	2.07(1.34)	36.1	ATLAS [148]
EW $WZjj$	$1.77^{+0.51}_{-0.45}$	5.3 (3.2)	36.1	ATLAS [149]
EW $WWjj$ same sign	$1.44^{+0.26}_{-0.24}(\text{stat})^{+0.28}_{-0.22}(\text{syst})$	6.5 (4.4^{Sherpa}, 6.5^{Powheg})	36.1	ATLAS [147]
EW $ZZjj$	1.35 ± 0.34	5.5 (4.3)	139	ATLAS [150]
EW $ZZjj\ \ell\ell\ell\ell jj$	1.5 ± 0.4	5.5 (3.9)	139	ATLAS [150]
EW $ZZjj\ \ell\ell\nu\nu jj$	0.7 ± 0.7	1.2 (1.8)	139	ATLAS [150]
EW $WWjj$ same sign	0.90 ± 0.22	5.5 (5.7)	35.9	CMS [151]
EW $ZZjj\ \ell\ell\ell\ell jj$	$1.39^{+0.86}_{-0.65}$	2.7 (1.6)	35.9	CMS [51]
EW $WZjj\ \ell\nu\ell\ell jj$	$0.82^{+0.51}_{-0.43}$	2.2 (2.5)	35.9	CMS [152]
EW $WVjj\ \ell\nu jj$	$0.85 \pm 0.12(\text{stat})^{+0.19}_{-0.17}(\text{syst})$	4.4 (5.1)	138	CMS [153]

Table 4.12: Searches of vector boson scattering processes with the ATLAS and CMS detectors at $\sqrt{s} = 13$ TeV. In each case, The experimental signature comprises the electroweak (EW) production of massive diboson (VV, $V = W/Z$) pairs in association with a pair of jets (jj).

5 - Conclusion and Outlook

Presented is a search for the electroweak production of a pair of vector bosons in association with a pair of jets with typical characteristics of vector boson scattering processes. The semileptonic channel with zero charged leptons in the final state is investigated using the full Run-II dataset of the ATLAS detector corresponding to an integrated luminosity of 139 fb^{-1}. A previous analysis [1] using a subset of this dataset of 36.1 fb^{-1} found a signal strength of $2.47^{+1.33}_{-1.22}$ and an observed (expected) significance of $1.43\ \sigma$ ($1.35\ \sigma$) for the exclusion of the background-only hypothesis. The presented analysis resulted in a signal strength of 2.02 ± 0.46 with an observed (expected) significance of $5.65\ \sigma$ ($3.03\ \sigma$).

This significant increase in the sensitivity of the study stems only partly from the increased integrated luminosity. Several improvements to the analysis strategy were made. A novel multivariate analysis approach making use of a recurrent neural network (RNN) was used in order to enhance the signal-to-background separation significantly. The associated increase in susceptibility to various systematic uncertainties has been studied thoroughly. The behaviour of the associated nuisance parameters was studied by comparing the results obtained with this method with the results obtained from a more conservative method using a feed-forward neural network (NN). The set of features used in the NN was chosen specifically to avoid correlations to the not well modelled jet multiplicity in the dominant V+jets background. This was identified as the main reason for the high systematic uncertainty associated to the RNN method. With this behaviour understood, the use of the more ambitious RNN method in order to achieve a higher sensitivity appears to be justified.

The presented study on the zero lepton channel has been developed in parallel with similar studies in the 1-, and 2-lepton channels. A combined statistical interpretation is expected to be published soon. In addition to the combined statistics this analysis will profit from synergies between channels despite their orthogonal kinematic phase spaces. The degeneracy caused by the similar shapes of the two dominant backgrounds (W+jets and Z+jets) in the 0-lepton channel made it necessary to treat various associated nuisance parameters together rather than defining separate nuisance parameters for each of the two. This will not be necessary in the combined analysis. The V+jets control region of the W+jets dominated 1-lepton channel and the corresponding control region in the Z+jets dominated 2-lepton channel will make it possible to constrain both backgrounds separately. The abundance of top background in the 1-lepton channel makes it possible to define a dedicated top control region which will be able to constrain the top backgrounds (single t and $t\bar{t}$). Many other nuisance parameters have in general similar effects in all three analysis channels and are therefore well constrained in the combined analysis. As a result, a much higher sensitivity is expected in the combined fit

compared to the already high sensitivity, which here is presented for the 0-lepton channel alone.

The two other projects (vector boson taggers, ITk planar pixels) presented in this thesis give prospects for further improvements of sensitivity in eventual future analyses searching for vector boson scattering processes. The measurements on prototypes of the future ITk planar pixel sensors presented in Sec. 2 show that the ambitious goal of a tracking detector withstanding the harsh radiation requirements imposed by the HL-LHC is possible. Due to the characteristic tagging jet system, analyses searching for vector boson scattering are especially sensitive to decay structures in the forward detector region. As a consequence, such analyses not only profit from the vastly increased integrated luminosity foreseen to be achieved with the HL-LHC, but also from the significantly increased acceptance for tracks up to a pseudorapidity of $|\eta| = 4$ beyond the current ID's acceptance of $|\eta| = 2.7$. The presented market survey measurements served as the basis for the selection of vendors for the final production of the ITk. The planar sensor Final Design Review (FDR) was passed in September 2020 which paved the way towards the sensor pre-production which is currently ongoing.

The vector boson tagger studies presented in Sec. 3 showed that the known improvement in jet substructure resolution from using the novel UFO jet definition translates into a significantly higher identification efficiency of vector-boson-initiated jets. This could significantly improve the event selection in the merged analysis regime. Additionally, the 3-variable cut-based tagger, which is used in the VBS analysis presented in this thesis, may be replaced by the neural network based tagger; or even by its mass-decorrelated version based on adversarial neural networks. The latter would compensate for the very apparent shaping of the signal jet mass $m(J^{\text{sig}})$ (Fig. 4.8 a and 4.9 a) in the background distributions after the application of the tagger. While an inclusion of a fully calibrated version of these taggers into the main analysis of this thesis was not feasible within the given time frame, such a version is now available for future analyses to use [154].

The points mentioned above show excellent prospects for future vector boson scattering analyses. The expected further improvement in sensitivity makes searches for these processes even more potent for precise tests of the Higgs sector. If contributions from beyond the Standard Model processes are realised in nature, the intricate cancellation of VBS and Higgs diagrams suggests they may first be found in measurements like these. Even before future improvements from the vector boson tagger studies and the ITk come into play, the soon to be finalised combination of this (0-lepton) analysis channel with the other two (1-, and 2-lepton) channels will be a result to be excited about.

Appendix

A.1 . Geometrical Measurements

Tab. A.1.1 and Tab. A.1.2 show the results of a thickness and bow measurement of an additional sensor from the ITk Planar Pixel Market Survey similar to those that were discussed in Sec. 2.2.1 (Tab. 2.1 and Tab. 2.2). The sensor shows a thickness of 218 ± 22 μm and a bow of 12 μm.

	h_1	$h_{1'}$	h_3	$h_{3'}$	h_7	$h_{7'}$	h_9	$h_{9'}$	T
1	9438.4	9707.2	9479.4	9628.4	9469.8	9648.7	9474.2	9722.6	211.3
2	9496.7	9657.6	9494.1	9698.1	9440.4	9628.6	9473.2	9688.8	192.2
3	9478.3	9639.1	9463.1	9697.7	9487.0	9725.1	9498.9	9744.9	219.9
4	9503.7	9702.5	9447.5	9747.6	9508.2	9696.8	9506.0	9691.3	218.2
5	9501.1	9727.5	9463.6	9732.7	9509.8	9688.8	9500.4	9698.8	218.2
6	9491.3	9719.7	9450.4	9735.0	9526.7	9634.8	9515.2	9718.8	206.2
7	9428.8	9693.3	9437.8	9729.7	9516.1	9687.8	9572.6	9756.0	227.9
8	9467.9	9723.2	9467.6	9708.7	9451.5	9717.3	9492.7	9740.2	252.4
9	9484.9	9686.9	9444.6	9732.1	9509.7	9764.5	9537.1	9677.9	221.3
10	9528.3	9700.5	9446.3	9729.8	9540.8	9730.3	9450.0	9783.1	244.6

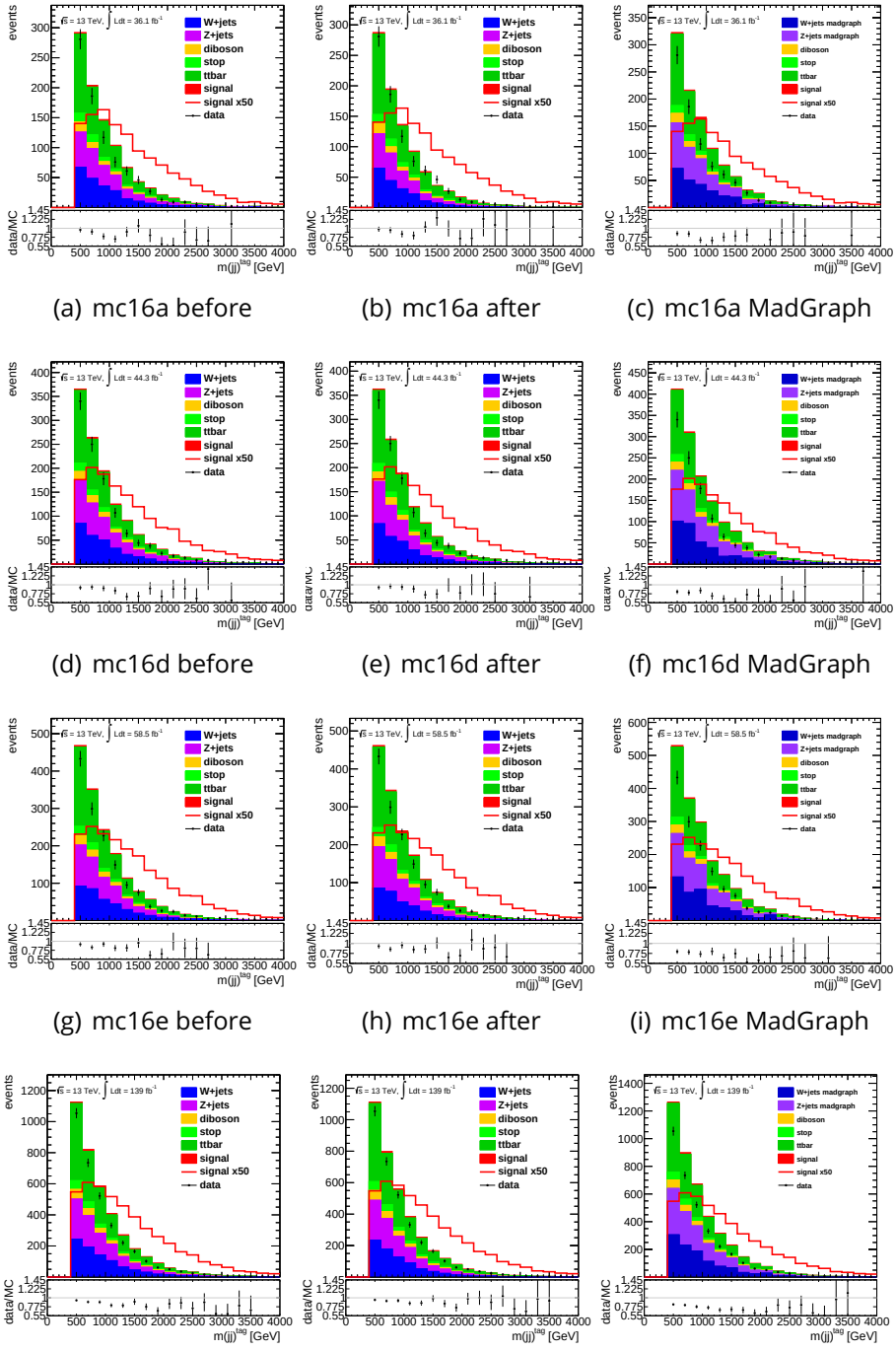
Table A.1.1: Thickness T (Eq. 2.5) measurements of a sensor for the ITk Planar Pixel Sensor Market Survey using the microscope focus method. The positions of the various measurement points are specified in Fig. 2.3(a). Each measurement was performed 10 times. All values are given in μm.

h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9	B
9724.5	9696.8	9717.7	9742.6	9709.8	9724.9	9703.7	9695.7	9745.1	11.9

Table A.1.2: Bow B (Eq. 2.6) measurements of a sensor for the ITk Planar Pixel Sensor Market Survey using the microscope focus method. The positions of the various measurement points are specified in Fig. 2.3(a). All values are given in μm.

A.2 . $m(jj)^{\text{tag}}$ Reweighting

Fig. A.2.1 - Fig. A.2.5 show the tagging jet mass $m(jj)^{\text{tag}}$ distributions before (left columns) and after (middle columns) applying the reweighting procedure as defined in Sec. 4.5. The corresponding distributions (right columns) using the alternative (MadGraph) Monte Carlo samples for the W/Z+jets background are presented as well. No reweighting is applied to them. The figures show the results for each individual Monte Carlo campaign (mc16a, -d, and -e) as well as their combination amounting to 139 fb^{-1}.



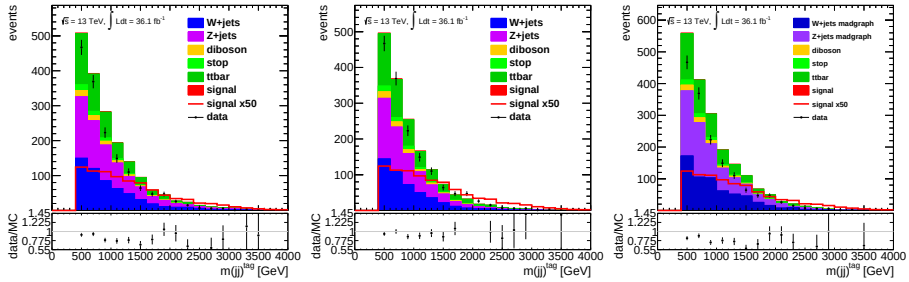
(a) mc16a before (b) mc16a after (c) mc16a MadGraph

(d) mc16d before (e) mc16d after (f) mc16d MadGraph

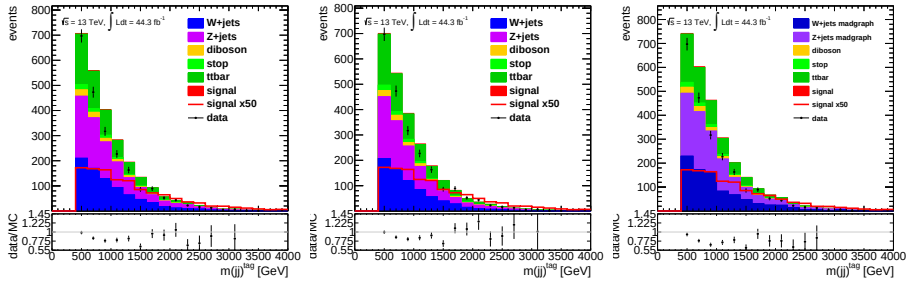
(g) mc16e before (h) mc16e after (i) mc16e MadGraph

(j) combined before (k) combined after (l) combined MadGraph

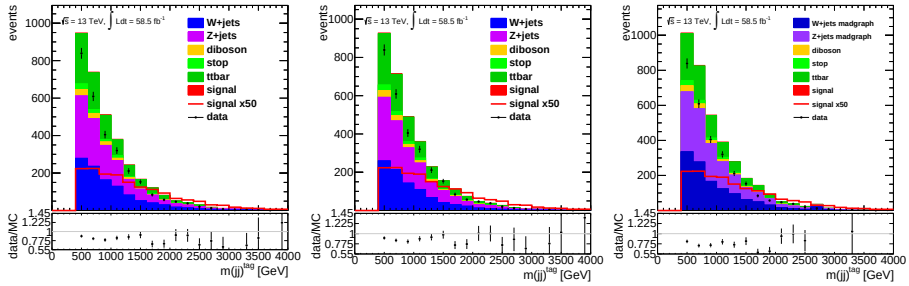
Figure A.2.1: $m(jj)^{\text{tag}}$ distributions in the 0-lepton merged HP signal region before (left) and after (middle) reweighting as well as using alternative $W/Z+$ jets samples without reweighting (right).



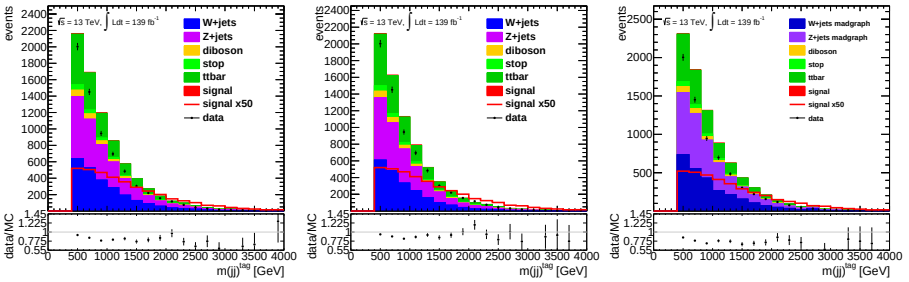
(a) mc16a before (b) mc16a after (c) mc16a MadGraph



(d) mc16d before (e) mc16d after (f) mc16d MadGraph

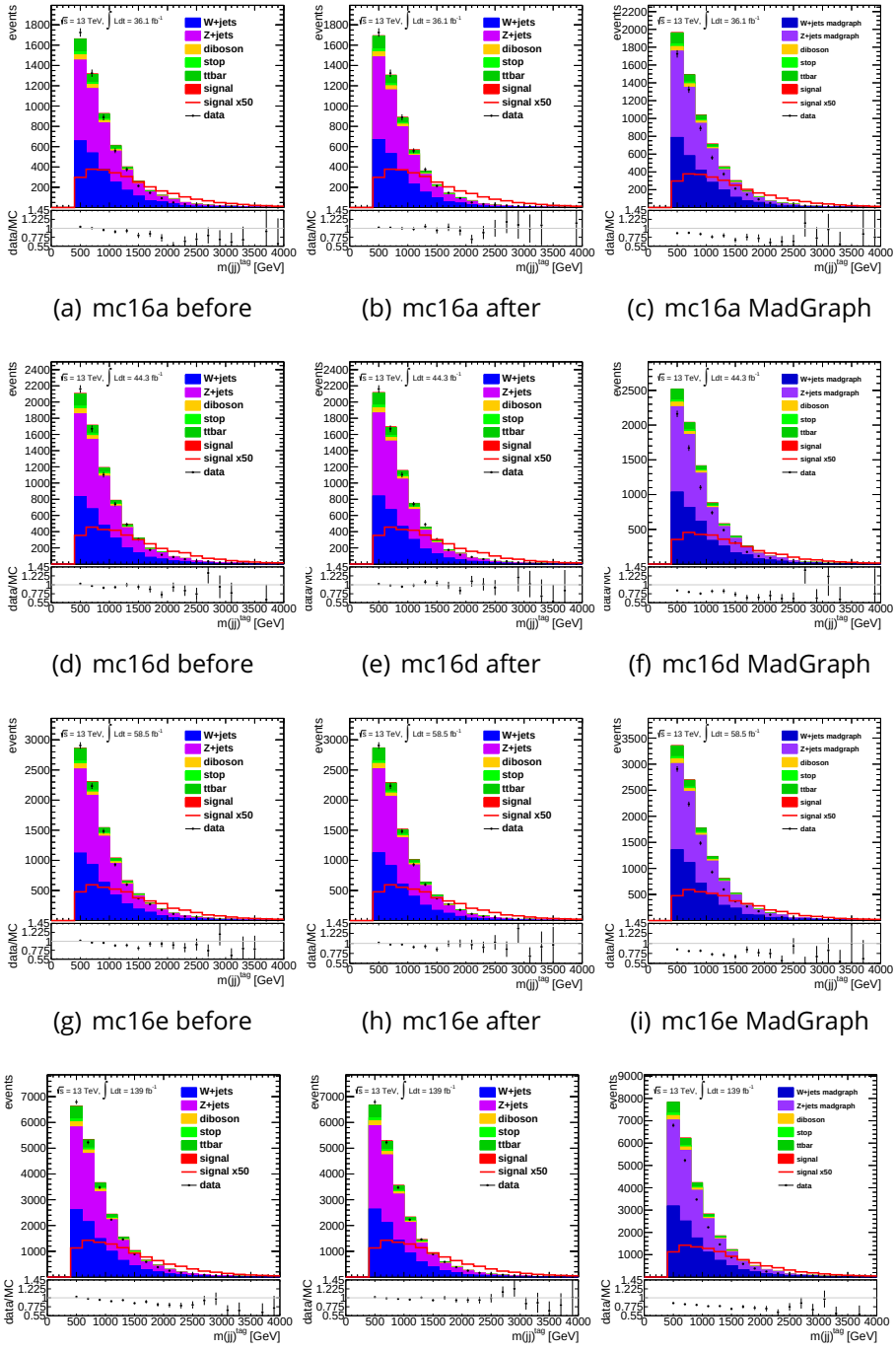


(g) mc16e before (h) mc16e after (i) mc16e MadGraph



(j) combined before (k) combined after (l) combined MadGraph

Figure A.2.2: $m(jj)^{\text{tag}}$ distributions in the 0-lepton merged LP signal region before (left) and after (middle) reweighting as well as using alternative $W/Z+$ jets samples without reweighting (right).



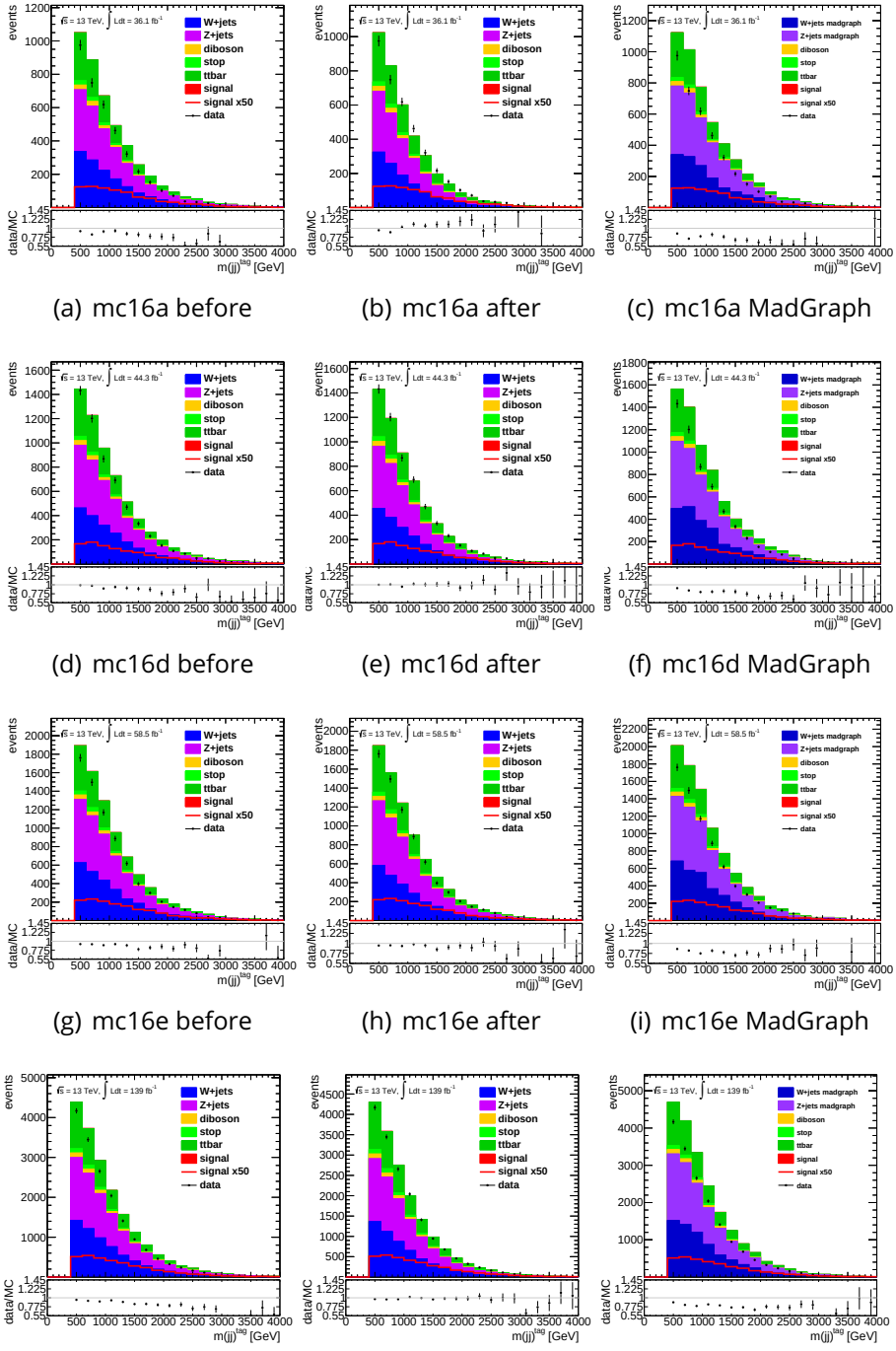
(a) mc16a before (b) mc16a after (c) mc16a MadGraph

(d) mc16d before (e) mc16d after (f) mc16d MadGraph

(g) mc16e before (h) mc16e after (i) mc16e MadGraph

(j) combined before (k) combined after (l) combined MadGraph

Figure A.2.3: $m(jj)^{\text{tag}}$ distributions in the 0-lepton resolved signal region before (left) and after (middle) reweighting as well as using alternative $W/Z+$ jets samples without reweighting (right).



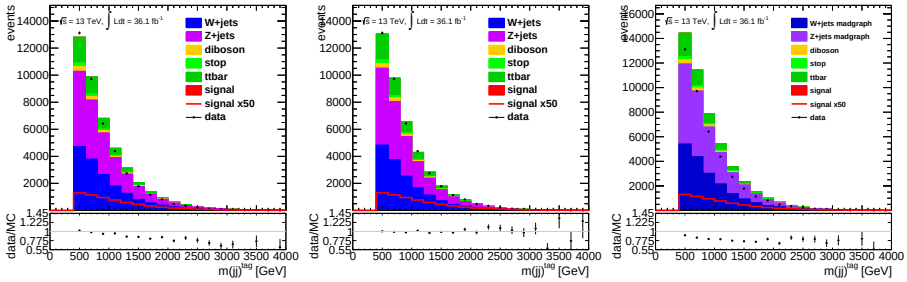
(a) mc16a before (b) mc16a after (c) mc16a MadGraph

(d) mc16d before (e) mc16d after (f) mc16d MadGraph

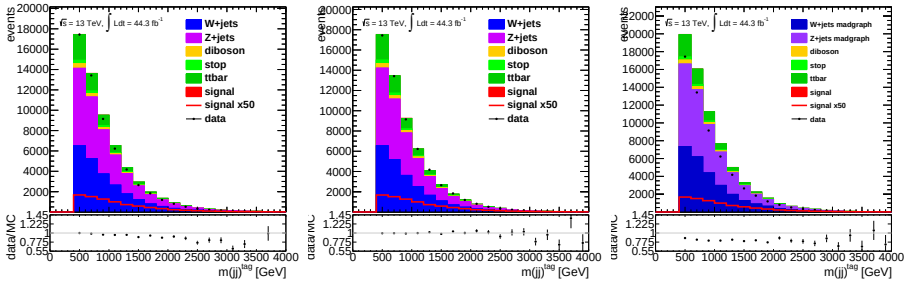
(g) mc16e before (h) mc16e after (i) mc16e MadGraph

(j) combined before (k) combined after (l) combined MadGraph

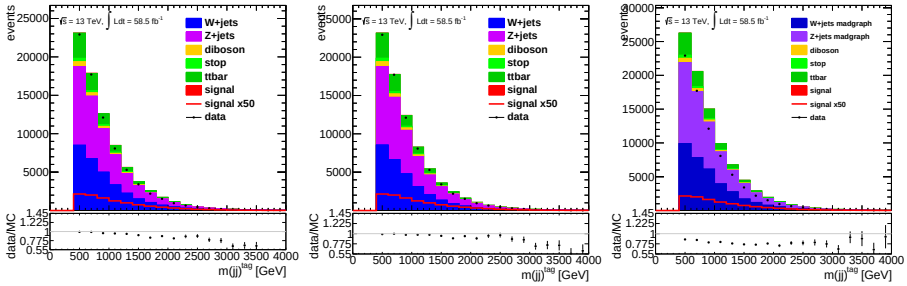
Figure A.2.4: $m(jj)^{\text{tag}}$ distributions in the 0-lepton merged control region before (left) and after (middle) reweighting as well as using alternative $W/Z+$ jets samples without reweighting (right).



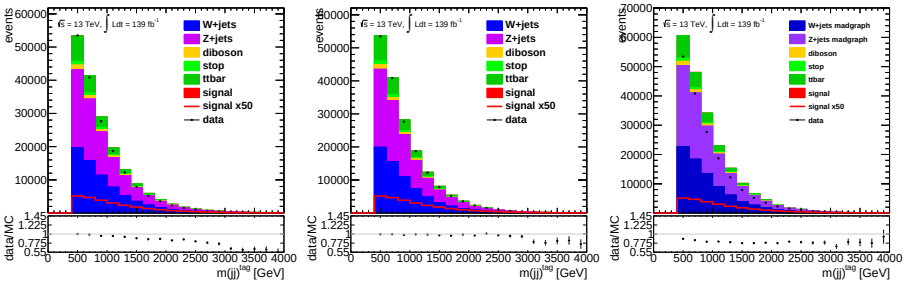
(a) mc16a before (b) mc16a after (c) mc16a MadGraph



(d) mc16d before (e) mc16d after (f) mc16d MadGraph



(g) mc16e before (h) mc16e after (i) mc16e MadGraph



(j) combined before (k) combined after (l) combined MadGraph

Figure A.2.5: $m(jj)^{\text{tag}}$ distributions in the 0-lepton resolved control region before (left) and after (middle) reweighting as well as using alternative $W/Z+$ jets samples without reweighting (right).

A.3 . Forward Jet Vertex Tagger (fJVT)

The tagging jet selection requires both jets to pass the forward jet tagger (fJVT) in addition to the baseline JVT tagger (see object definition in Sec. 4.3). The reason for applying the fJVT for these but not generally for the other $R = 0.4$ jets in the event is that tagging jets tend to be in the forward detector region. Tab. A.3.3 shows what the cut-flow of the analysis would look like without the fJVT requirement. This should be compared with Tab. 4.5 (Sec. 4.4) which shows the nominal cut-flow. It can be seen that the number of signal events in the signal analysis regions is not significantly reduced by this additional requirement. While the signal over background ratio is not significantly improved by the application of the fJVT, it was decided to apply it to be more robust against pileup effects in the forward region that are not well excluded by the JVT alone.

(a) merged SR HP

cut	signal	background	all MC	data	data/MC
All	15945	15275273	15291218	40529219	2.65
$\mathsf{N}(j^{\mathsf{tag}}) \geq 2$	12874	9747181	9760056	25218566	2.58
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{lead}}) > 30$ GeV	12862	9717425	9730287	25162070	2.59
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{sublead}}) > 30$ GeV	11398	7216204	7227602	19396872	2.68
$E_{\mathsf{T}}^{\mathsf{miss}} > 200$ GeV	5126	2391625	2396751	4283204	1.79
$E_{\mathsf{T}}^{\mathsf{miss,track}} > 50$ GeV	4596	2178090	2182686	3407864	1.56
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, E_{\mathsf{T}}^{\mathsf{miss,track}})\rvert/\pi < \frac{1}{4}$	4414	2128930	2133344	3165470	1.48
$\min\{\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, j_i^{\mathsf{any}})\rvert\}_i/\pi > \frac{1}{6}$	2587	1358278	1360866	1403011	1.03
$\mathsf{N}(J) \geq 1$	445	109794	110239	110911	1.01
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, J)\rvert/\pi > \frac{1}{9}$	428	106604	107032	107644	1.01
$m(jj)^{\mathsf{tag}} > 400$ GeV	392	86001	86393	87169	1.01
$m(J) > 50$ GeV	284	36862	37146	34922	0.94
tagger $\mathfrak{f}_V^{=50\%}(D_2\mathsf{only})$	173	16480	16653	15435	0.93
tagger $\mathfrak{f}_V^{=50\%}$	95	3977	4073	3683	0.90

(b) merged SR LP

cut	signal	background	all MC	data	data/MC
All	15850	15271296	15287146	40525536	2.65
$\mathsf{N}(j^{\mathsf{tag}}) \geq 2$	12779	9743204	9755983	25214883	2.58
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{lead}}) > 30$ GeV	12767	9713448	9726214	25158387	2.59
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{sublead}}) > 30$ GeV	11303	7212226	7223529	19393189	2.68
$E_{\mathsf{T}}^{\mathsf{miss}} > 200$ GeV	5031	2387648	2392678	4279521	1.79
$E_{\mathsf{T}}^{\mathsf{miss,track}} > 50$ GeV	4501	2174113	2178613	3404181	1.56
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, E_{\mathsf{T}}^{\mathsf{miss,track}})\rvert/\pi < \frac{1}{4}$	4319	2124952	2129271	3161787	1.48
$\min\{\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, j_i^{\mathsf{any}})\rvert\}_i/\pi > \frac{1}{6}$	2492	1354301	1356793	1399328	1.03
$\mathsf{N}(J) \geq 1$	349	105817	106166	107228	1.01
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, J)\rvert/\pi > \frac{1}{9}$	333	102626	102959	103961	1.01
$m(jj)^{\mathsf{tag}} > 400$ GeV	297	82024	82321	83486	1.01
$m(J) > 50$ GeV	189	32884	33073	31239	0.94
tagger $\mathfrak{f}_V^{=80\%}(D_2\mathsf{only})$	144	24152	24297	22694	0.93
tagger $\mathfrak{f}_V^{=80\%}$	76	8073	8149	7334	0.90

(c) resolved SR

cut	signal	background	all MC	data	data/MC
All	15773	15263223	15278996	40518202	2.65
$\mathsf{N}(j^{\mathsf{tag}}) \geq 2$	12703	9735131	9747834	25207549	2.59
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{lead}}) > 30$ GeV	12690	9705375	9718065	25151053	2.59
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{sublead}}) > 30$ GeV	11226	7204154	7215380	19385855	2.69
$E_{\mathsf{T}}^{\mathsf{miss}} > 200$ GeV	4955	2379575	2384529	4272187	1.79
$E_{\mathsf{T}}^{\mathsf{miss,track}} > 50$ GeV	4424	2166040	2170464	3396847	1.57
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, E_{\mathsf{T}}^{\mathsf{miss,track}})\rvert/\pi < \frac{1}{4}$	4243	2116880	2121122	3154453	1.49
$\min\{\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, j_i^{\mathsf{any}})\rvert\}_i/\pi > \frac{1}{6}$	2416	1346228	1348644	1391994	1.03
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, jj)\rvert/\pi > \frac{1}{9}$	2346	1269822	1272168	1309713	1.03
$\mathsf{N}(j^{\mathsf{sig}}) \geq 2$	2002	748713	750714	722561	0.96
$p_{\mathsf{T}}(j^{\mathsf{sig}}_{\mathsf{lead}}) > 40$ GeV	1916	644659	646575	623828	0.96
$p_{\mathsf{T}}(j^{\mathsf{sig}}_{\mathsf{sublead}}) > 20$ GeV	1916	644659	646575	623828	0.96
$m(jj)^{\mathsf{tag}} > 400$ GeV	1472	408265	409736	391247	0.95
$64 < m(jj)^{\mathsf{sig}} < 106$ GeV	582	91051	91633	84479	0.92
$m(jjj) > 220$ GeV	220	23542	23763	23292	0.98

(d) merged CR

cut	signal	background	all MC	data	data/MC
All	15192	15172172	15187364	40433723	2.66
$\mathsf{N}(j^{\mathsf{tag}}) \geq 2$	12121	9644080	9656201	25123070	2.60
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{lead}}) > 30$ GeV	12108	9614324	9626432	25066574	2.60
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{sublead}}) > 30$ GeV	10645	7113102	7123747	19301376	2.71
$E_{\mathsf{T}}^{\mathsf{miss}} > 200$ GeV	4373	2288524	2292896	4187708	1.83
$E_{\mathsf{T}}^{\mathsf{miss,track}} > 50$ GeV	3842	2074989	2078831	3312368	1.59
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, E_{\mathsf{T}}^{\mathsf{miss,track}})\rvert/\pi < \frac{1}{4}$	3661	2025828	2029489	3069974	1.51
$\min\{\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, j_i^{\mathsf{any}})\rvert\}_i/\pi > \frac{1}{6}$	1834	1255177	1257011	1307515	1.04
$\mathsf{N}(J) \geq 1$	220	89617	89837	92077	1.02
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, J)\rvert/\pi > \frac{1}{9}$	207	86709	86916	89082	1.02
$m(jj)^{\mathsf{tag}} > 400$ GeV	171	66107	66278	68607	1.04
$m(J) > 50$ GeV	76	19023	19099	18477	0.97
fail tagger $\mathfrak{f}_V^{=80\%}$	76	19023	19099	18477	0.97

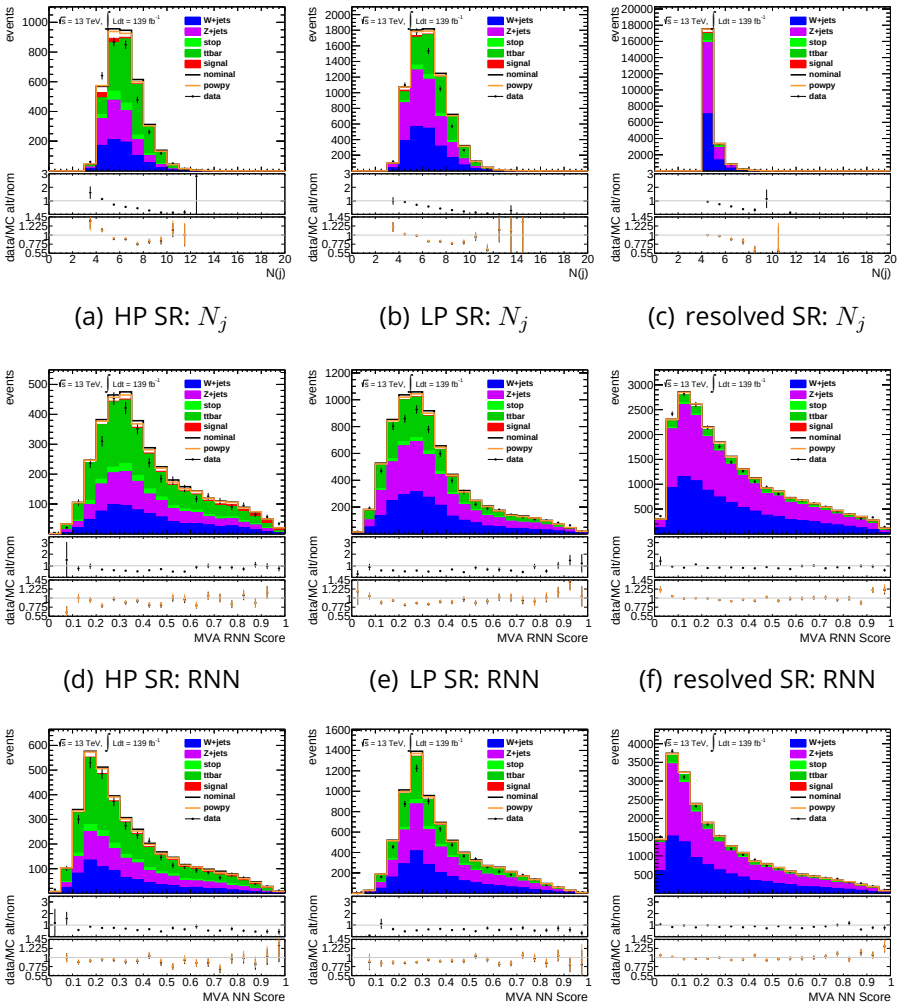
(e) resolved CR

cut	signal	background	all MC	data	data/MC
All	15116	15153149	15168265	40415246	2.66
$\mathsf{N}(j^{\mathsf{tag}}) \geq 2$	12046	9625057	9637102	25104593	2.60
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{lead}}) > 30$ GeV	12033	9595301	9607334	25048097	2.61
$p_{\mathsf{T}}(j^{\mathsf{tag}}_{\mathsf{sublead}}) > 30$ GeV	10569	7094079	7104648	19282899	2.71
$E_{\mathsf{T}}^{\mathsf{miss}} > 200$ GeV	4297	2269500	2273798	4169231	1.83
$E_{\mathsf{T}}^{\mathsf{miss,track}} > 50$ GeV	3767	2055966	2059733	3293891	1.60
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, E_{\mathsf{T}}^{\mathsf{miss,track}})\rvert/\pi < \frac{1}{4}$	3585	2006805	2010391	3051497	1.52
$\min\{\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, j_i^{\mathsf{any}})\rvert\}_i/\pi > \frac{1}{6}$	1759	1236154	1237912	1289038	1.04
$\lvert\Delta\Phi(E_{\mathsf{T}}^{\mathsf{miss}}, jj)\rvert/\pi > \frac{1}{9}$	1689	1159923	1161612	1206948	1.04
$\mathsf{N}(j^{\mathsf{sig}}) \geq 2$	1349	639318	640667	620403	0.97
$p_{\mathsf{T}}(j^{\mathsf{sig}}_{\mathsf{lead}}) > 40$ GeV	1264	535266	536530	521670	0.97
$p_{\mathsf{T}}(j^{\mathsf{sig}}_{\mathsf{sublead}}) > 20$ GeV	1264	535266	536530	521670	0.97
$m(jj)^{\mathsf{tag}} > 400$ GeV	819	298872	299691	289089	0.96
not $64 < m(jj)^{\mathsf{sig}} < 106$ GeV	819	298872	299691	289089	0.96
$m(jjj) > 220$ GeV	521	185796	186317	184287	0.99

Table A.3.3: Cut flow of the complete event selection of the analysis but without applying the fJVT for the tagging jet selection. The nominal case with fJVT applied can be seen in table 4.5.

A.4 . Monte Carlo Modelling Uncertainties

The difference in the modelling of the dominant W/Z+jets background from using alternative MC samples (MadGraph) were discussed in Sec. 4.7 (Fig. 4.21 and Fig. 4.22). Especially the impact of the considerable differences in the jet multiplicity (N_j) modelling on the modelling of the final discriminants, i.e. MVA scores was studied there. Here, in addition to that, such differences are shown for the diboson (Fig. A.4.6) and $t\bar{t}$ (Fig. A.4.7) backgrounds using the alternative PowhegPythia and PowhegHerwig MC samples respectively. The samples are defined in Sec. 4.2. It is evident that such differences are much smaller here than what was seen for the W/Z+jets backgrounds.

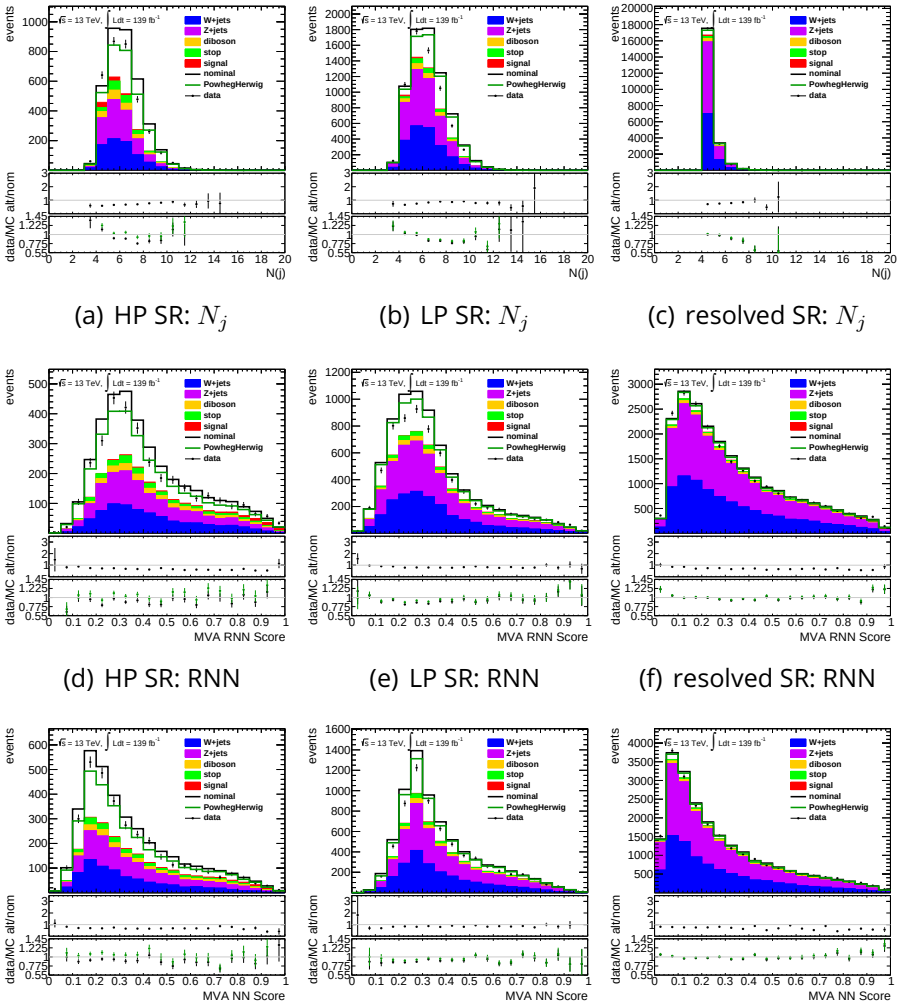


(a) HP SR: N_j (b) LP SR: N_j (c) resolved SR: N_j

(d) HP SR: RNN (e) LP SR: RNN (f) resolved SR: RNN

(g) HP SR: NN score (h) LP SR: NN score (i) resolved SR: NN score

Figure A.4.6: Modelling differences between Sherpa nominal diboson samples and PowhegPythia alternative diboson samples in the 0-lepton channel. Shown are jet multiplicity, score of the RNN($5j$) and NN score.



(a) HP SR: N_j (b) LP SR: N_j (c) resolved SR: N_j

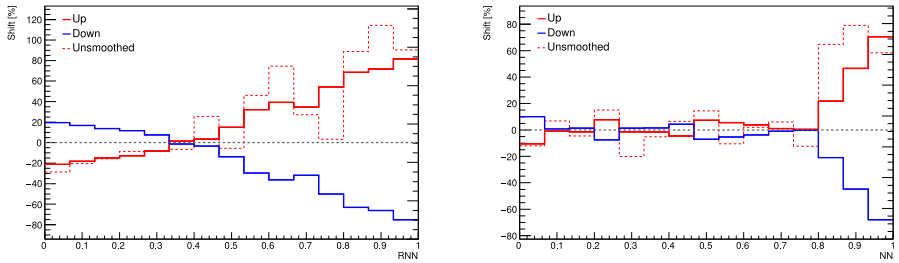
(d) HP SR: RNN (e) LP SR: RNN (f) resolved SR: RNN

(g) HP SR: NN score (h) LP SR: NN score (i) resolved SR: NN score

Figure A.4.7: Modelling differences between PowhegPythia nominal ttbar samples and PowhegHerwig alternative ttbar samples in the 0-lepton channel. Shown are jet multiplicity, score of the RNN($5j$) and NN score.

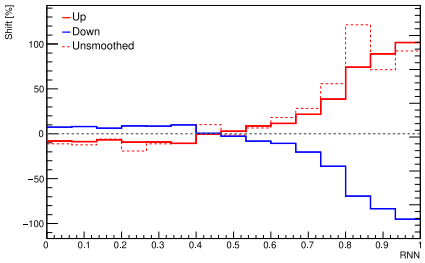
A.5 . W+jets Modelling Uncertainties

Fig. A.5.8 and Fig. A.5.9 show the variation of the V+jets and $m(jj)^{\text{tag}}$ reweighting modelling systematic uncertainties on the W+jets background corresponding to such figures for the Z+jets background that were shown in Sec. 4.9.

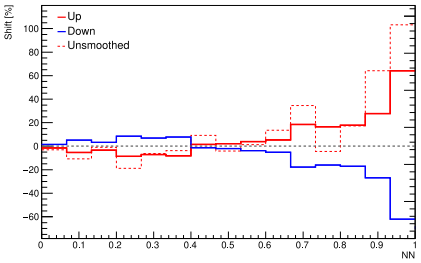


(a) RNN merged HP SR

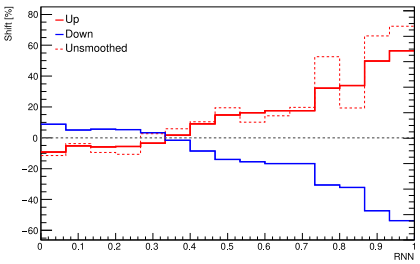
(b) NN merged HP SR



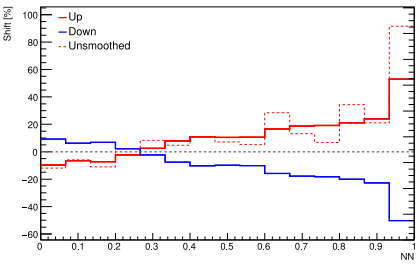
(c) RNN merged LP SR



(d) NN merged LP SR

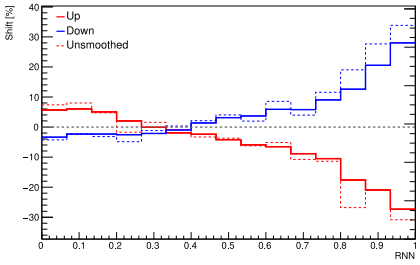


(e) RNN resolved SR

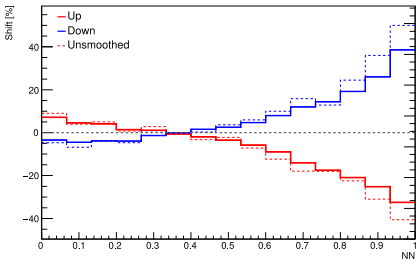


(f) NN resolved SR

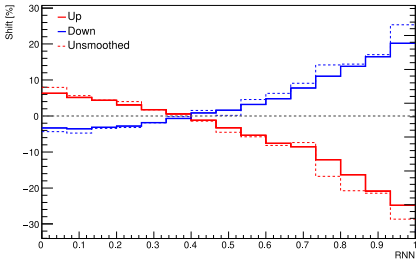
Figure A.5.8: $\pm 1\,\sigma$ variation of the V+jets modelling systematic uncertainty $\mathrm{SysMODEL\_V\_MGPy8}$ on the W+jets background sample for the three different options of MVA in the three different signal regions.



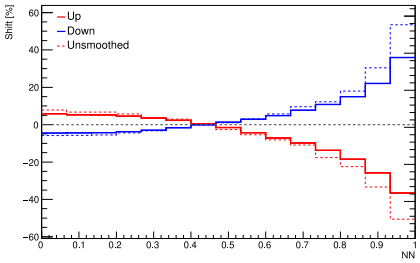
(a) RNN merged HP SR



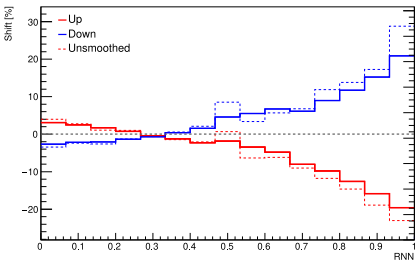
(b) NN merged HP SR



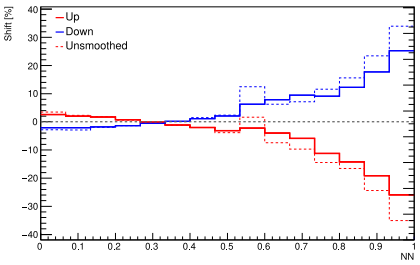
(c) RNN merged LP SR



(d) NN merged LP SR



(e) RNN resolved SR



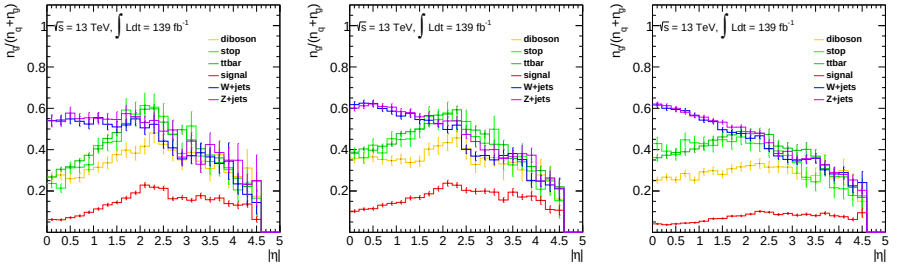
(f) NN resolved SR

Figure A.5.9: ± 1 σ variation of the $m(jj)^{\text{tag}}$ reweighting systematic uncertainty $\mathrm{SysMJJREWEIGHT\_100per\_L0\_Fat1(SysMJJREWEIGHT\_100per\_L0\_J2)}$ in the merged (resolved) regime on the W+jets background sample for the three different options of MVA in the three different signal regions.

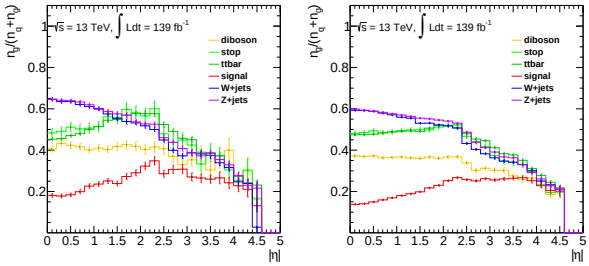
A.6 . Jet Quark/Gluon Fraction Uncertainties

The jet flavour composition and responce uncertainties were derived based on the quark gluon fractions of the MC samples used to model the signal and background processes. This was described in Sec. 4.8. Fig. A.6.10 - Fig. A.6.14 show additional η and p_{T} profiles in addition to those that were discussed there (Fig. 4.23). Fig. A.6.15 and Fig. A.6.16 show the final two-dimensional histograms of the fractions and their associated uncertainties that were used for the derivation

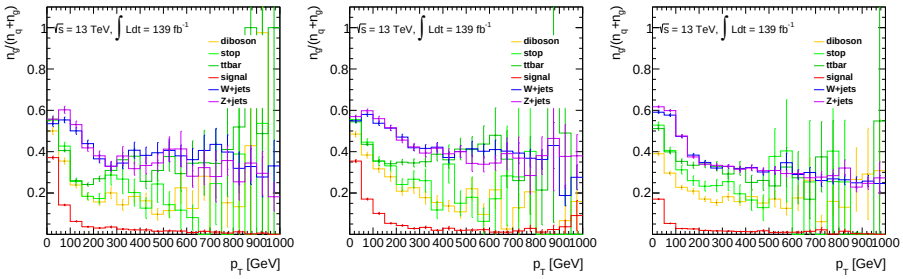
of the systematics as they were used in the fits of the statistical interpretation.



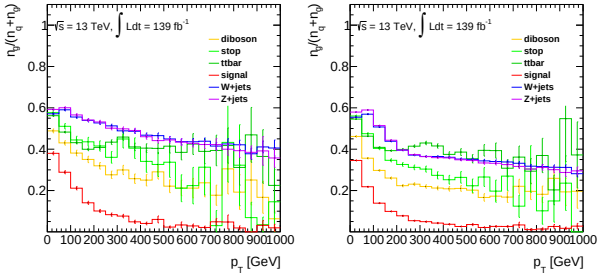
(a) merged HP SR (b) merged LP SR (c) resolved SR



(d) merged CR (e) resolved CR

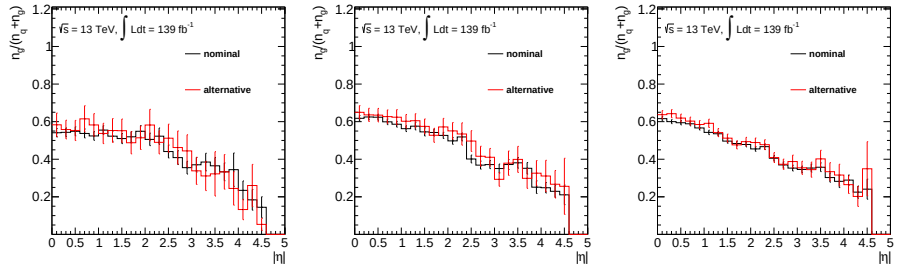


(f) merged HP SR (g) merged LP SR (h) resolved SR

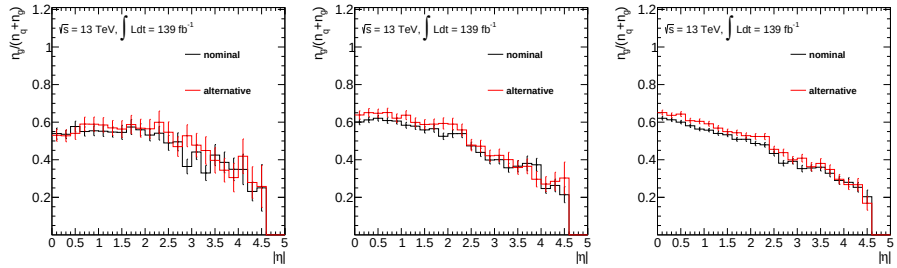


(i) merged CR (j) resolved CR

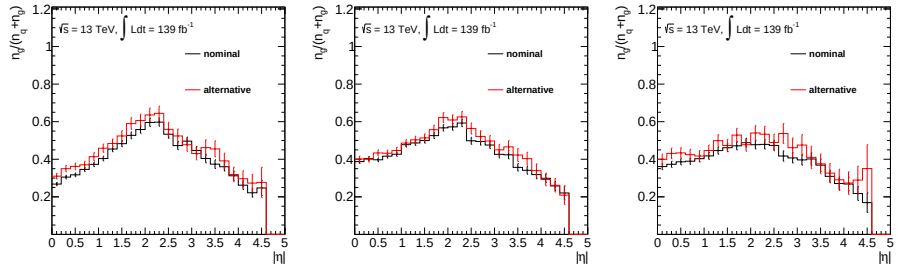
Figure A.6.10: Quark gluon fraction as a function of jet η and p_{T} for the different background and signal samples.



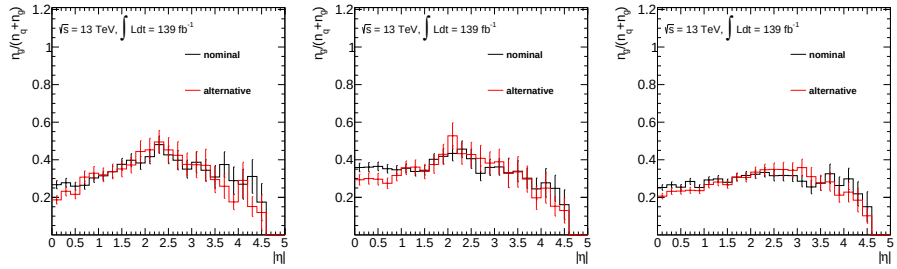
(a) merged HP SR W+jet (b) merged LP SR W+jet (c) resolved SR W+jet



(d) merged HP SR Z+jet (e) merged LP SR Z+jet (f) resolved SR Z+jet

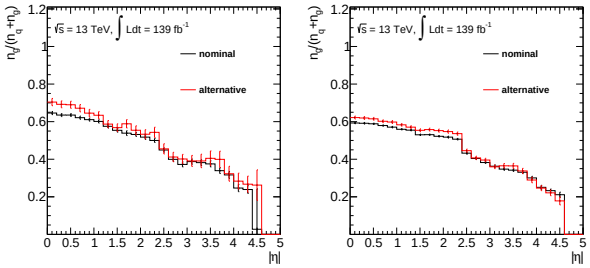


(g) merged HP SR $t\bar{t}$ (h) merged LP SR $t\bar{t}$ (i) resolved SR $t\bar{t}$

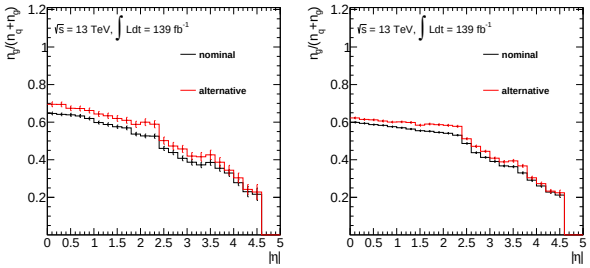


(j) merged HP SR diboson (k) merged LP SR diboson (l) resolved SR diboson

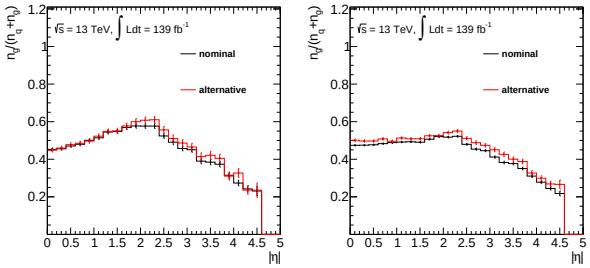
Figure A.6.11: Quark gluon fraction as a function of jet η for the different background and signal samples in the signal regions. Shown are comparisons of nominal and alternative models for individual samples.



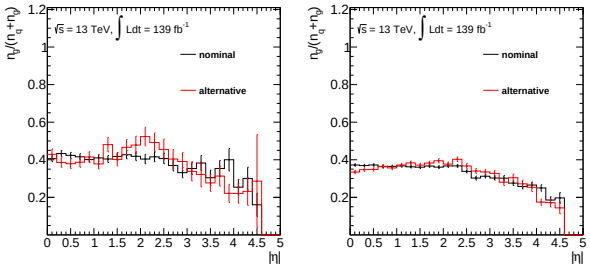
(a) merged HP CR W+jet (b) resolved CR W+jet



(c) merged HP CR Z+jet (d) resolved CR Z+jet

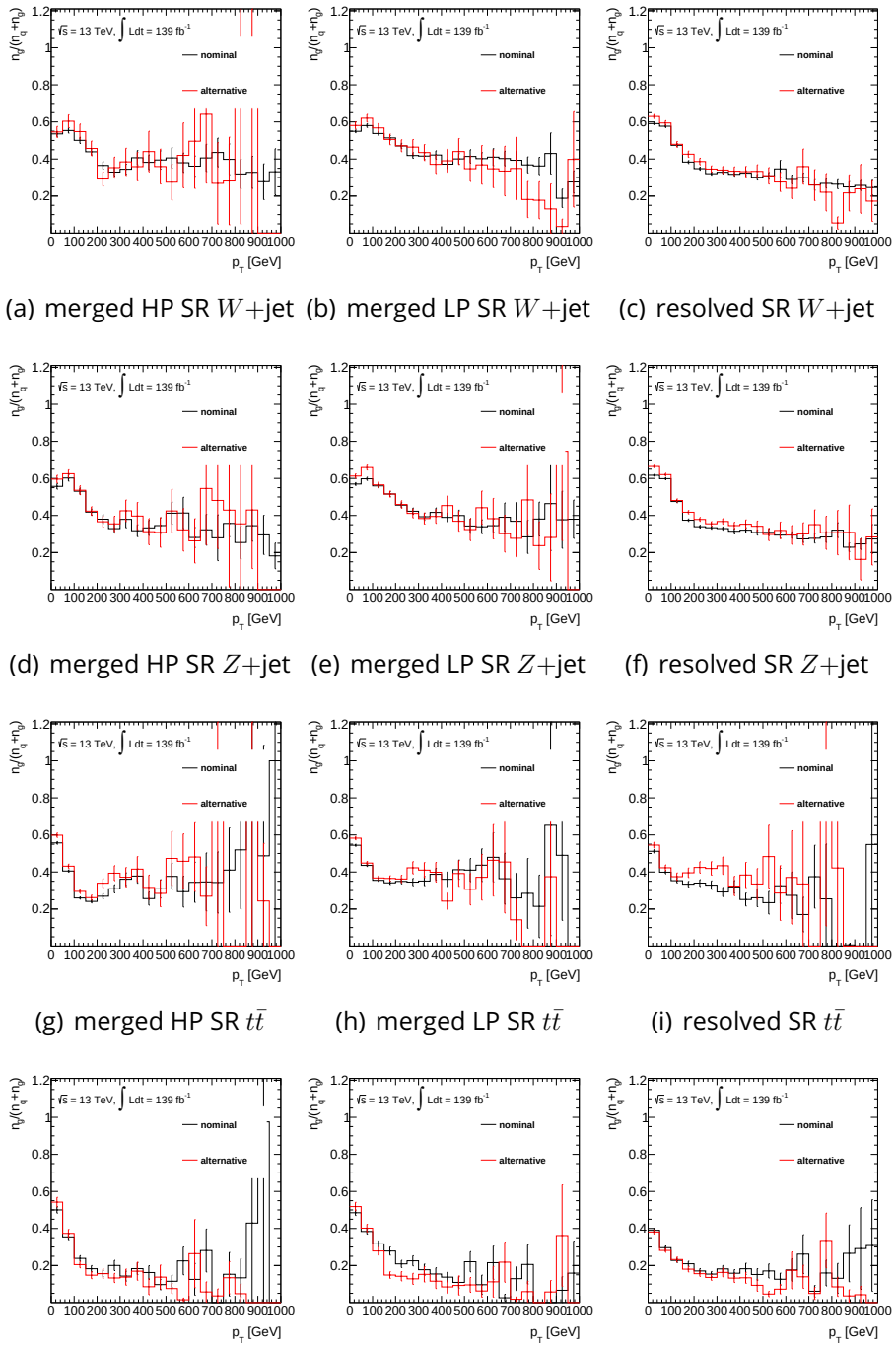


(e) merged HP CR $t\bar{t}$ (f) resolved CR $t\bar{t}$



(g) merged HP CR diboson (h) resolved CR diboson

Figure A.6.12: Quark gluon fraction as a function of jet η for the different background and signal samples in the control regions. Shown are comparisons of nominal and alternative models for individual samples.



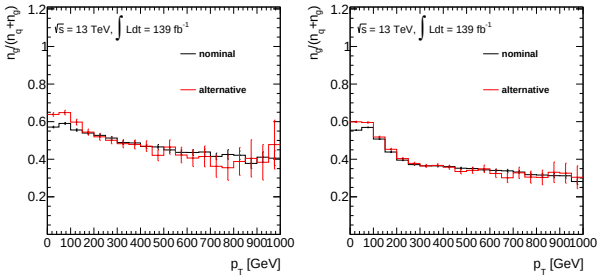
(a) merged HP SR W+jet (b) merged LP SR W+jet (c) resolved SR W+jet

(d) merged HP SR Z+jet (e) merged LP SR Z+jet (f) resolved SR Z+jet

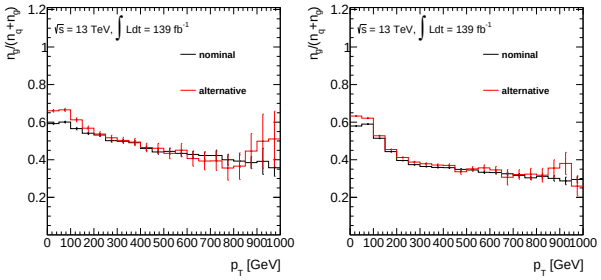
(g) merged HP SR $t\bar{t}$ (h) merged LP SR $t\bar{t}$ (i) resolved SR $t\bar{t}$

(j) merged HP SR diboson (k) merged LP SR diboson (l) resolved SR diboson

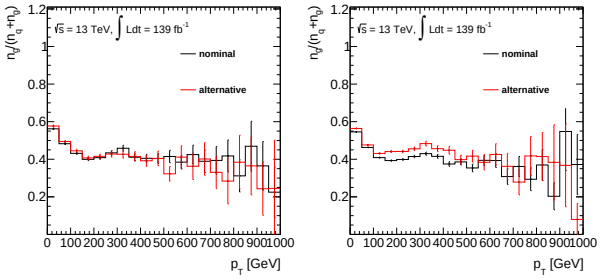
Figure A.6.13: Quark gluon fraction as a function of jet p_{T} for the different background and signal samples in the signal regions. Shown are comparisons of nominal and alternative models for individual samples.



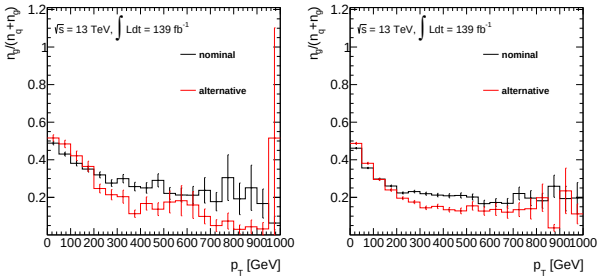
(a) merged HP CR W+jet (b) resolved CR W+jet



(c) merged HP CR Z+jet (d) resolved CR Z+jet

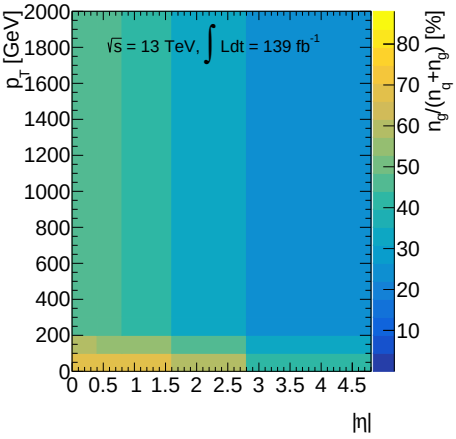


(e) merged HP CR $t\bar{t}$ (f) resolved CR $t\bar{t}$

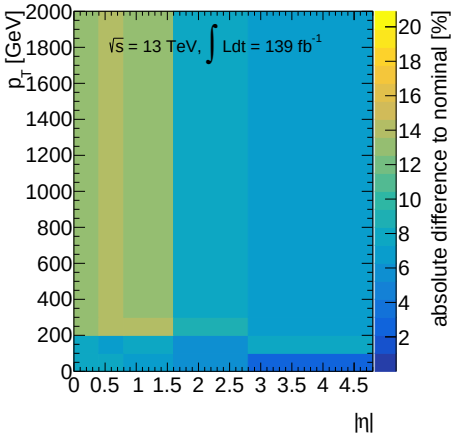


(g) merged HP CR diboson (h) resolved CR diboson

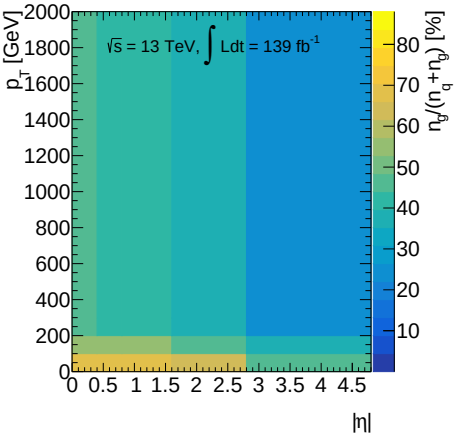
Figure A.6.14: Quark gluon fraction as a function of jet p_{T} for the different background and signal samples in the control regions. Shown are comparisons of nominal and alternative models for individual samples.



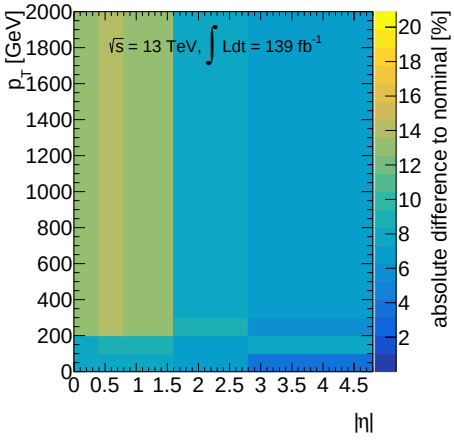
(a) fractions W+jets



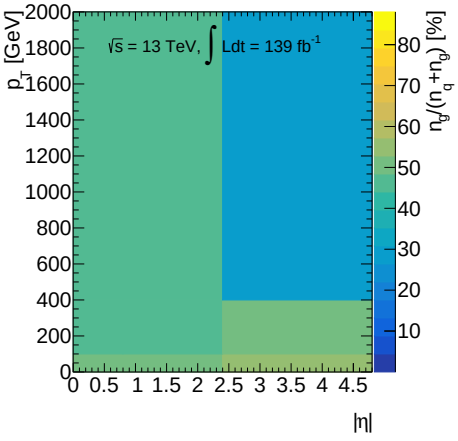
(b) total uncertainties W+jets



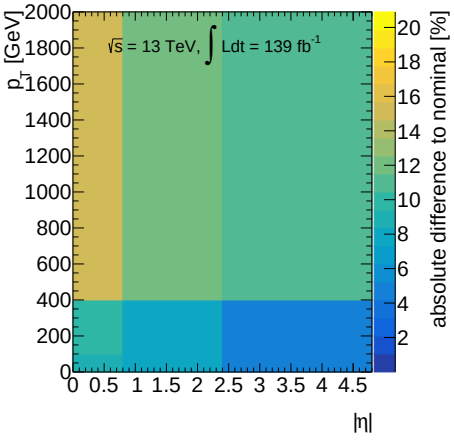
(c) fractions Z+jets



(d) total uncertainties Z+jets

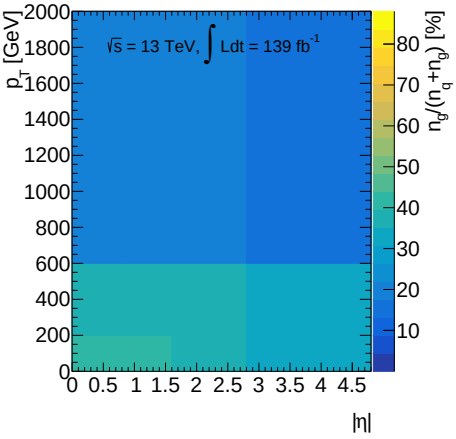


(e) fractions $t\bar{t}$

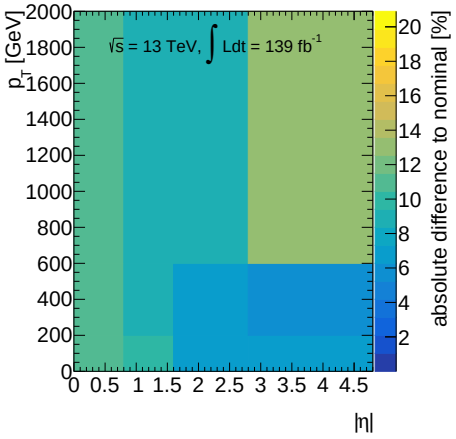


(f) total uncertainties $t\bar{t}$

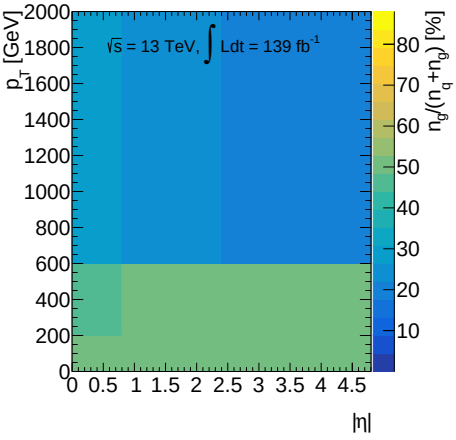
Figure A.6.15: Quark gluon fraction and associated total uncertainty as a function of jet η and p_{T} (continued on next page).



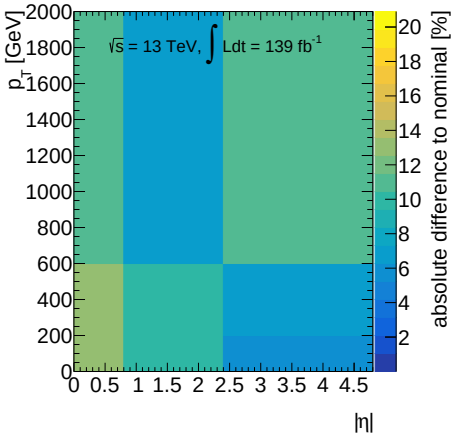
(a) fractions diboson



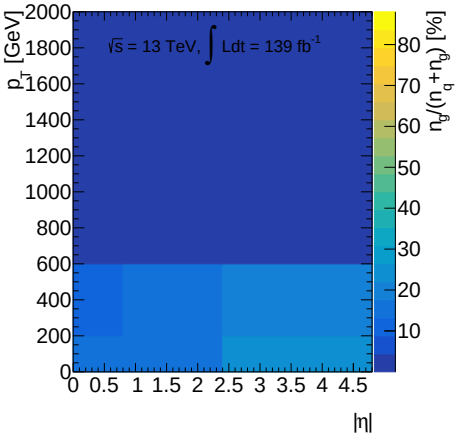
(b) total uncertainties diboson



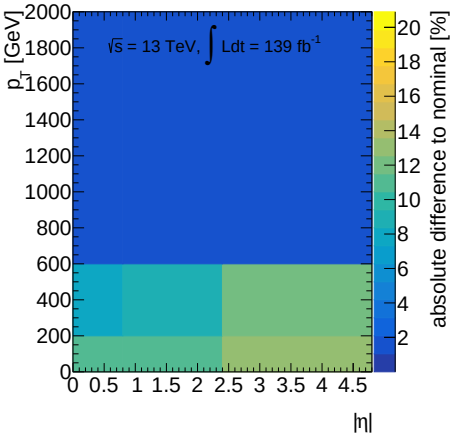
(c) fractions single-top



(d) total uncertainties single-top

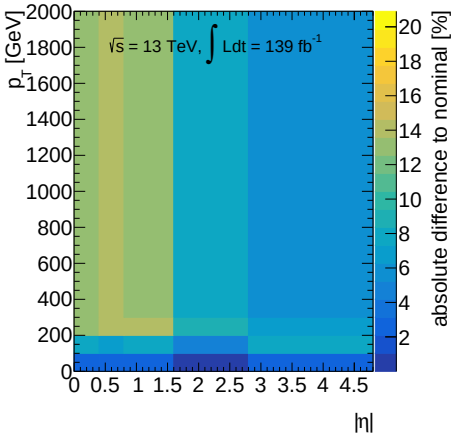


(e) fractions signal

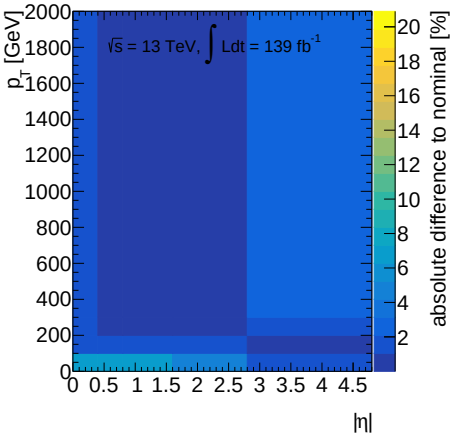


(f) total uncertainties signal

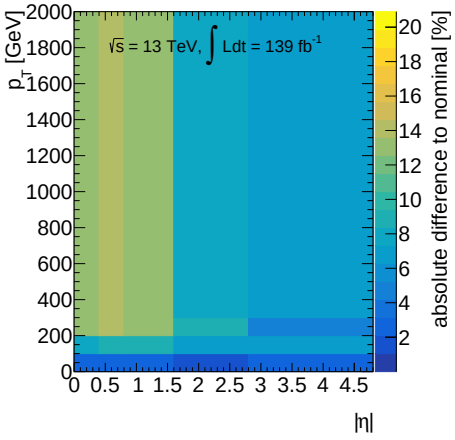
Figure A.6.15: Continuation from previous page.



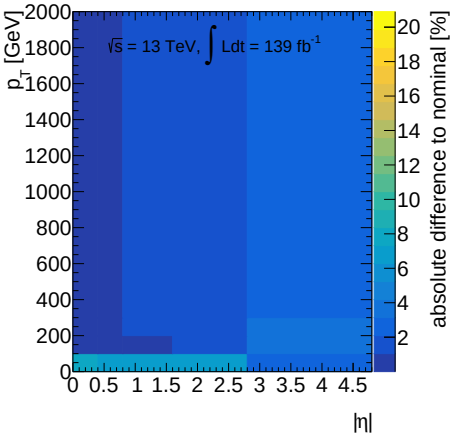
(a) region uncertainties W+jets



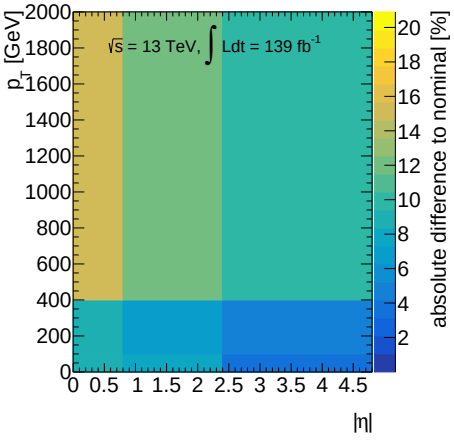
(b) generator uncertainties W+jets



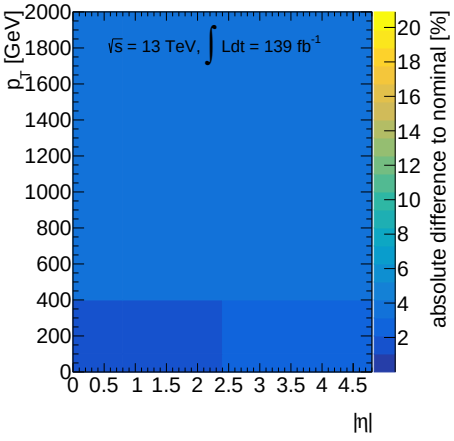
(c) region uncertainties Z+jets



(d) generator uncertainties Z+jets

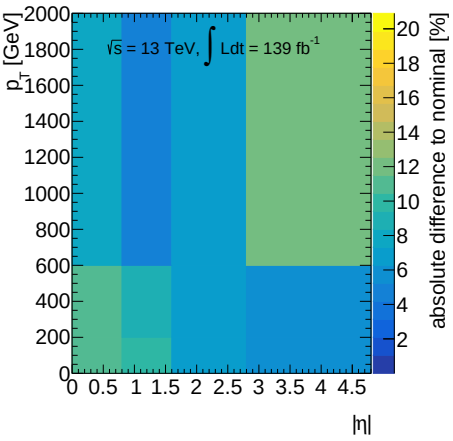


(e) region uncertainties $t\bar{t}$

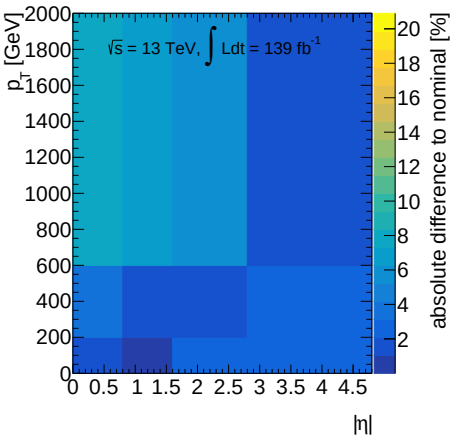


(f) generator uncertainties $t\bar{t}$

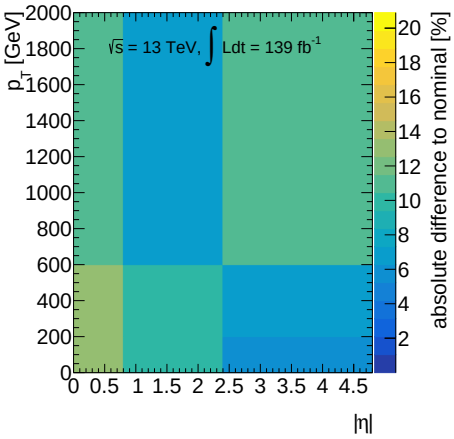
Figure A.6.16: Region- and generator-uncertainies used to derive the total uncertainty on the Quark gluon fraction as a function of jet η and p_{T} (continued on next page).



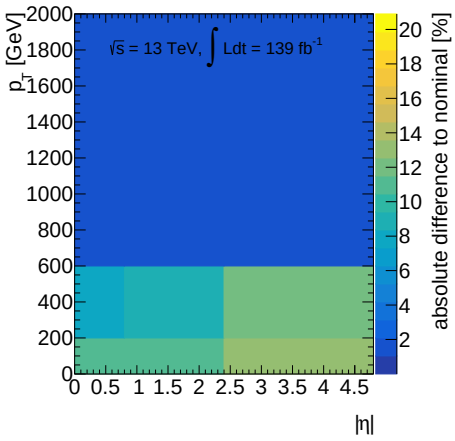
(a) region uncertainties diboson



(b) generator uncertainties diboson



(c) region uncertainties single-top



(d) region uncertainties signal

Figure A.6.16: Continuation from previous page.

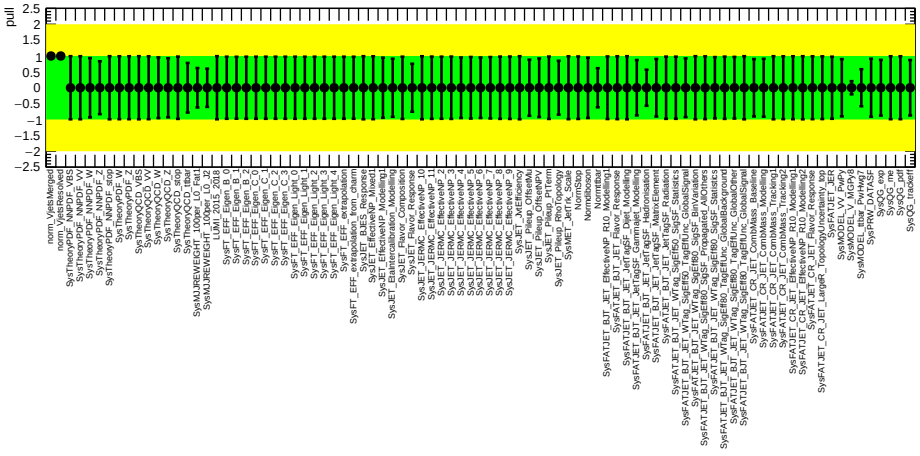
A.7 . Unblinding Procedure

A careful unblinding procedure has been implemented throughout the whole work with the analysis data. As usual this was done to avoid biasing the analysis configuration even unconsciously towards a wanted outcome. All steps before the statistical interpretation were performed in such a way that data in the most sensitive regions was neither used nor looked at. For sensitive discriminants, the data in the signal regions was only looked at in bins with lower signal to background ratio while in the control regions all data was available. Analysis steps requiring the use of data (e.g. $m(jj)^{\text{tag}}$ reweighting in Sec. 4.5) were derived in the control regions and applied in the signal regions.

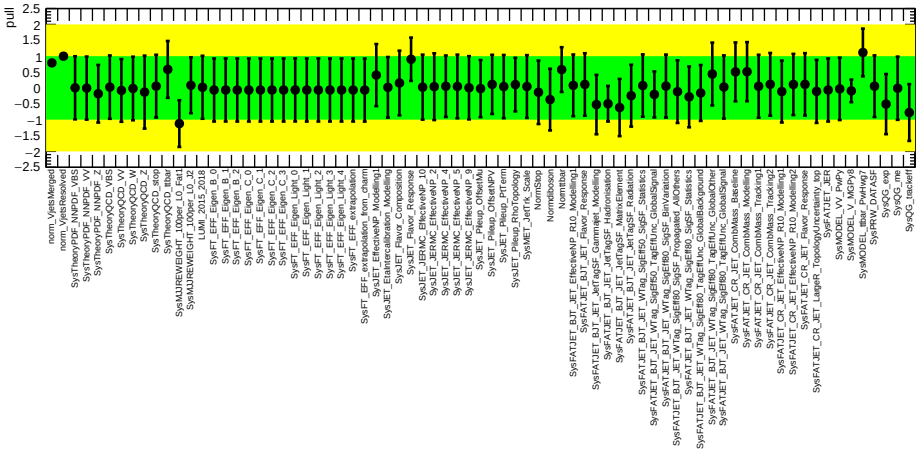
For the statistical interpretation itself an unblinding procedure in several steps was chosen: In the following they will be called 'Asimov', 'data left bins', and 'data conditional'. First, fits were made on an Asimov set generated from Monte Carlo simulation with the Standard Model signal strength $\mu^{\text{semileptonic}}_{\text{EW } VV+jj} = 1$. All nuisance parameters were investigated and checked for unusual or unexpected behaviour. Once the fit configuration was finalised based on this, fits were performed on data but first only on the left side of the signal region distributions[1]. Since the MVA is built in such a way that it concentrates signal in the right-most bins (toward a score of 1), bins from the right of the distribution were excluded from the fit until only 25% of the initial number of Monte Carlo signal events remained. The complete distributions of $m(jj)^{\text{tag}}$ in the control regions were used. The behaviour of all nuisance parameters was investigated once more and compared to the previous behaviour in the Asimov fits. All other fits up to this point were unconditional, i.e. with floating POI $\mu^{\text{semileptonic}}_{\text{EW } VV+jj}$. As a final step before complete unblinding, a conditional fit with a signal strength fixed to the Standard Model value $\mu^{\text{semileptonic}}_{\text{EW } VV+jj} = 1$ was performed using data in all bins. After verifying the integrity of the fit also in this iteration, the analysis was unblinded and the final unconditional fit to all data was performed.

The pulls of the fits associated to the three steps of the unblinding procedure (Asimov, data left bins, data conditional) can be seen in Fig. A.7.17 - Fig. A.7.19. The corresponding 'data left bins' and 'data conditional' postfit distributions of MVA discriminant ($m(jj)^{\text{tag}}$) in the signal (control) regions are shown in Fig. A.7.20 and Fig. A.7.21.

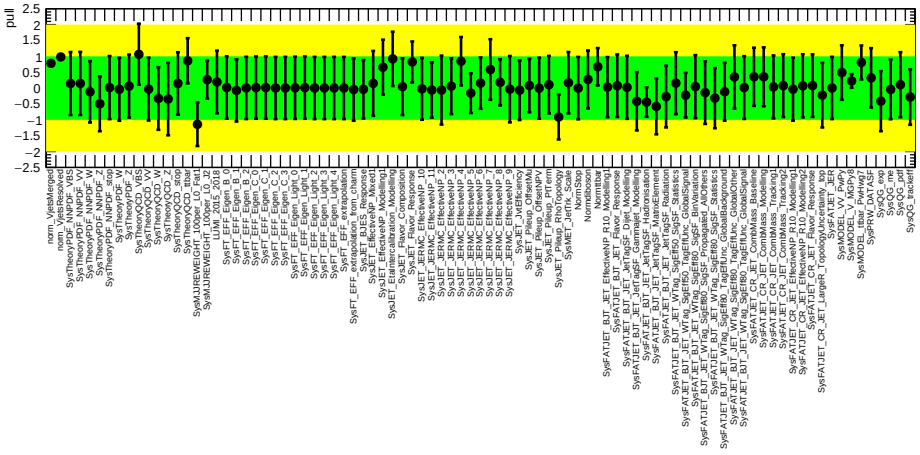
[1] As a intermediate step this was also done using an Asimov set. This step is omitted in the following to keep the discussion brief.



(a) Asimov

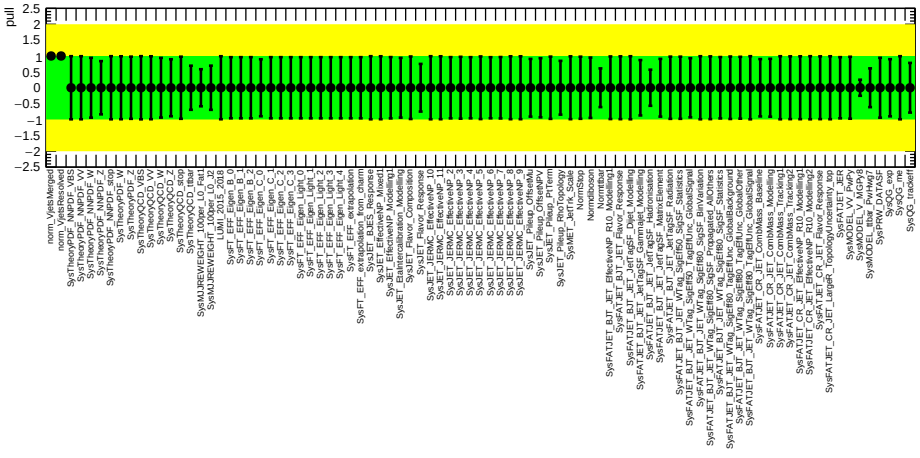


(b) data left bins

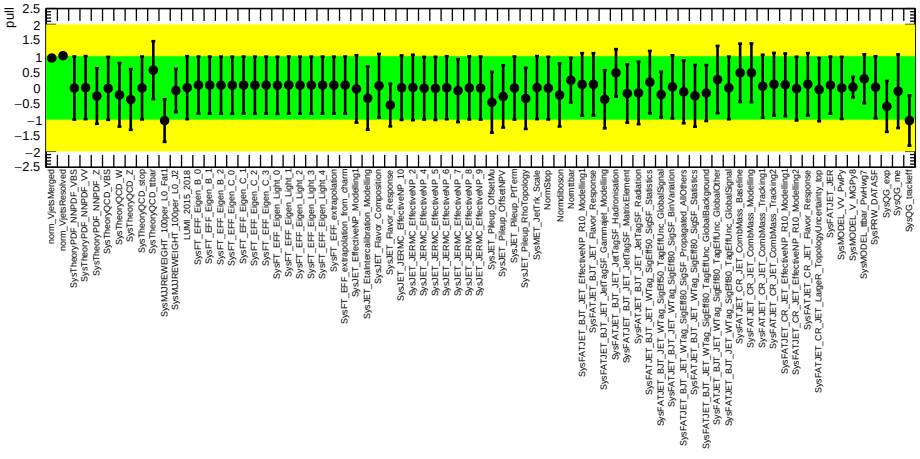


(c) data conditional

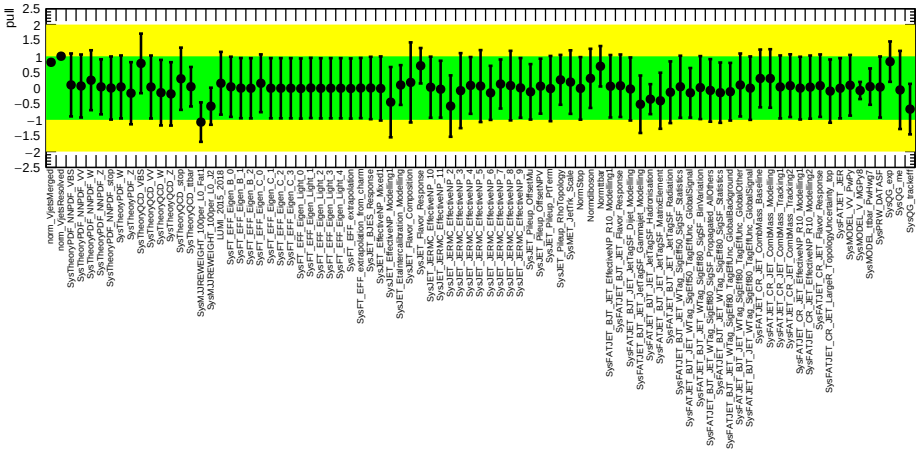
Figure A.7.17: Pulls of the RNN($5j$) fits at the three different steps of the unblinding procedure.



(a) Asimov

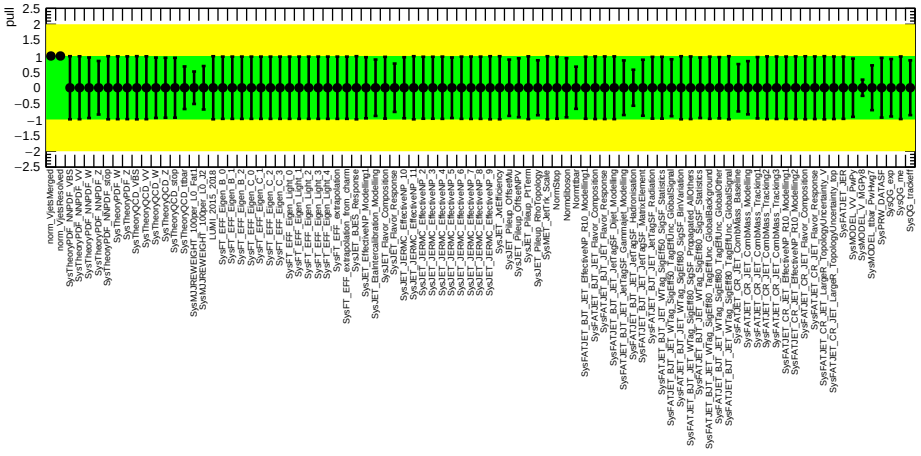


(b) data left bins

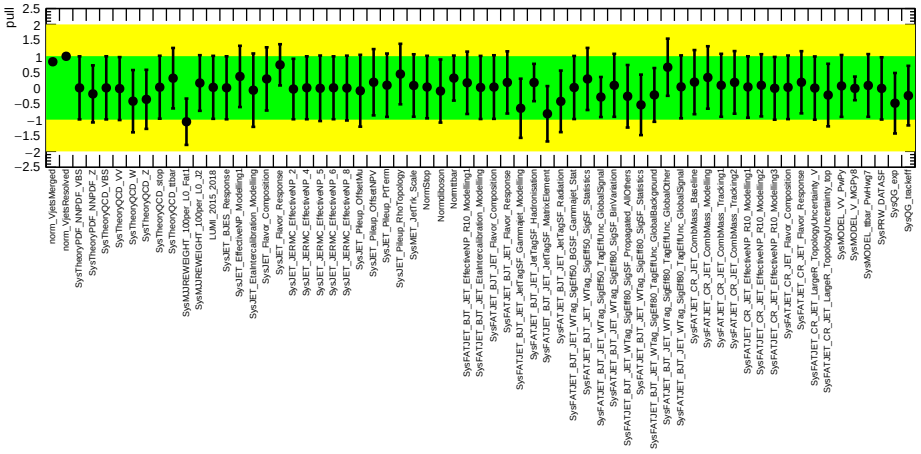


(c) data conditional

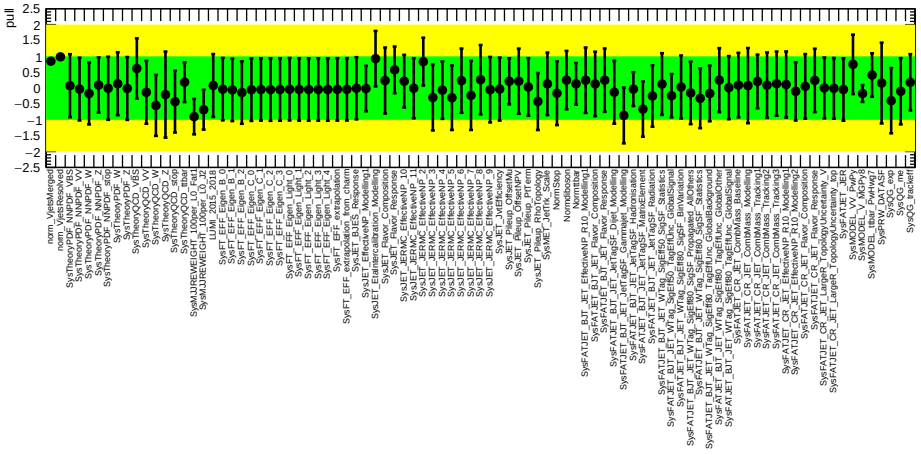
Figure A.7.18: Pulls of the RNN($4j$) fits at the three different steps of the unblinding procedure.



(a) Asimov

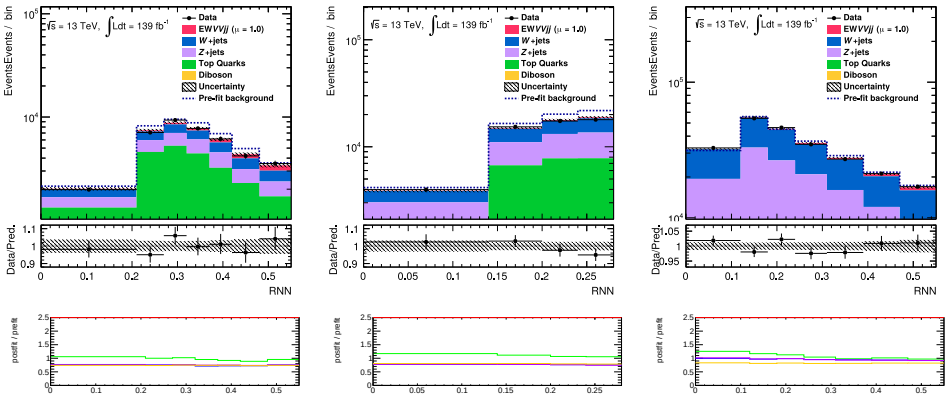


(b) data left bins

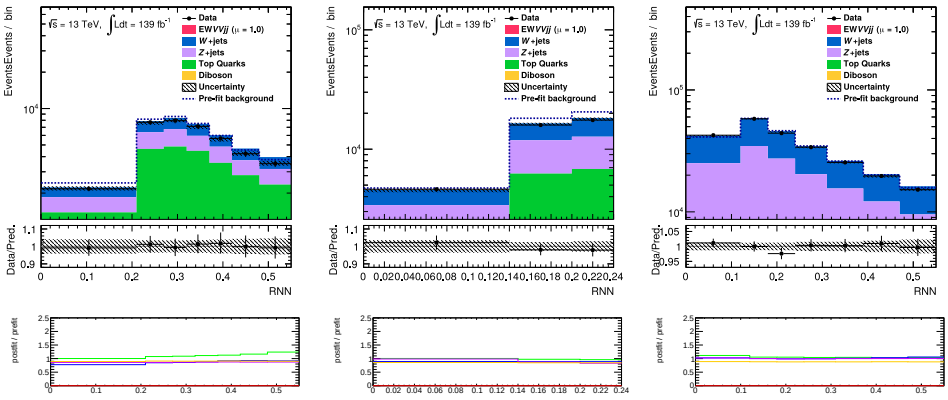


(c) data conditional

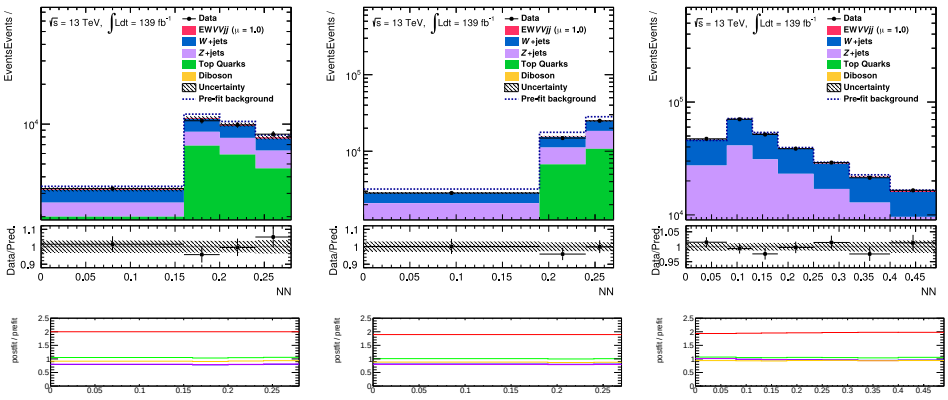
Figure A.7.19: Pulls of the NN fits at the three different steps of the unblinding procedure.



(a) RNN($5j$) HP SR (b) RNN($5j$) LP SR (c) RNN($5j$) resolved SR

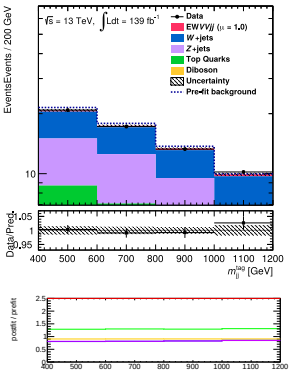


(d) RNN($4j$) HP SR (e) RNN($4j$) LP SR (f) RNN($4j$) resolved SR

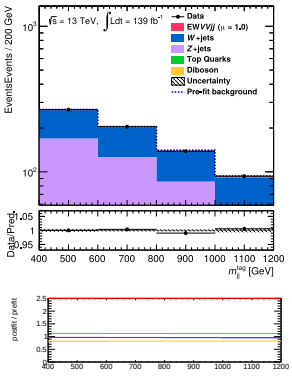


(g) NN HP SR (h) NN LP SR (i) NN resolved SR

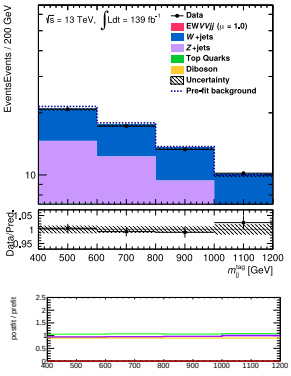
Figure A.7.20: Signal region postfit distributions of the three options for the MVA discriminants for a 'data left bins' fit.



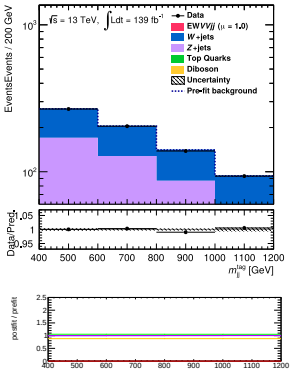
(a) RNN($5j$) CR



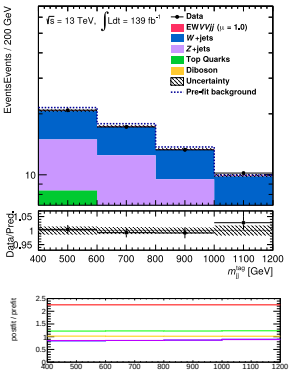
(b) RNN($5j$) resolved CR



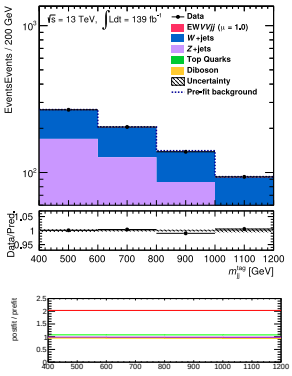
(c) RNN($4j$) CR



(d) RNN($4j$) resolved CR

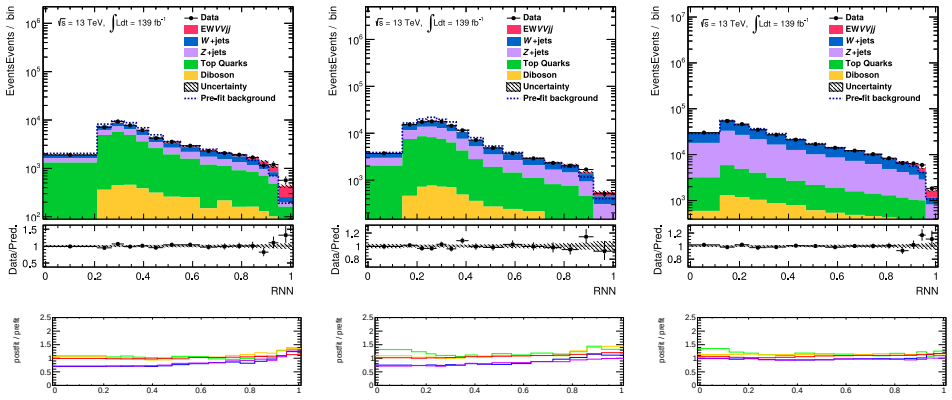


(e) NN CR

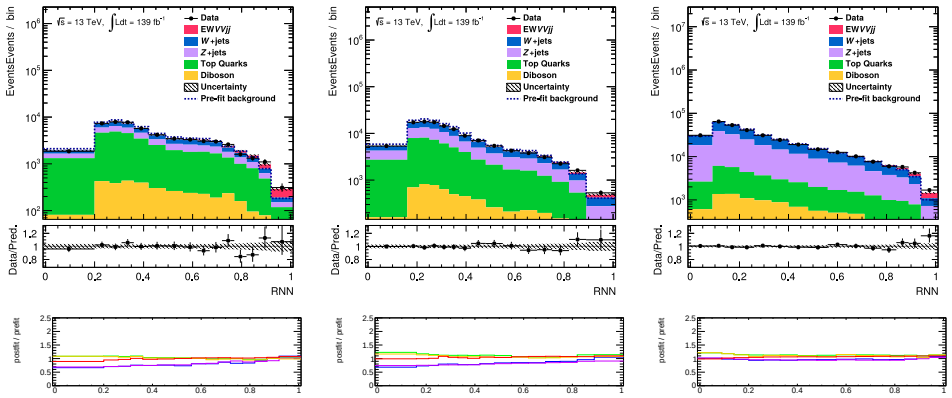


(f) NN resolved CR

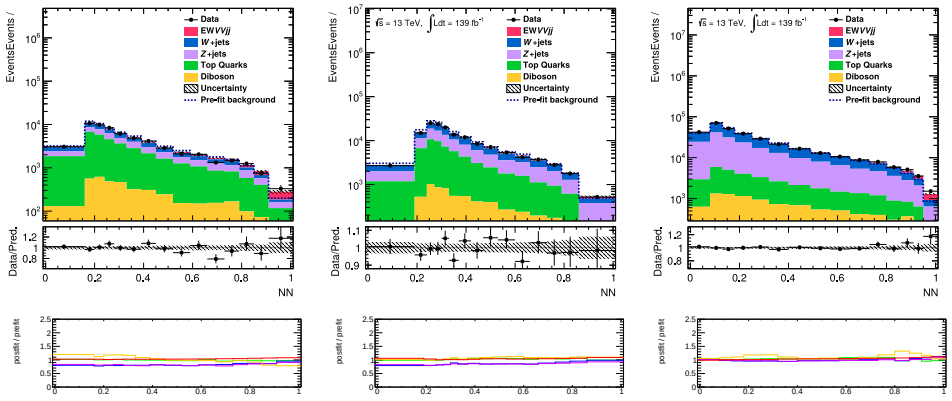
Figure A.7.20: Continuation from previous page.



(a) RNN($5j$) HP SR (b) RNN($5j$) LP SR (c) RNN($5j$) resolved SR

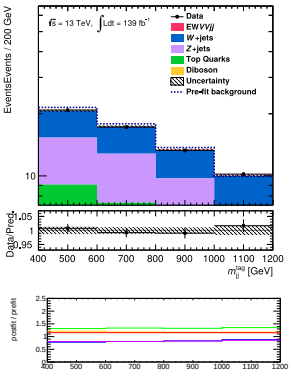


(d) RNN($4j$) HP SR (e) RNN($4j$) LP SR (f) RNN($4j$) resolved SR

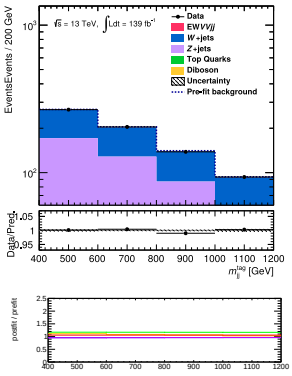


(g) NN HP SR (h) NN LP SR (i) NN resolved SR

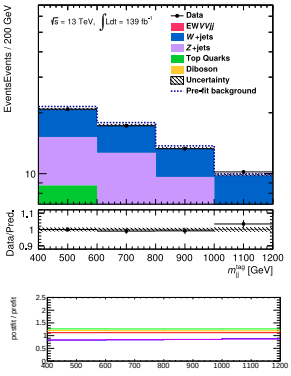
Figure A.7.21: Signal region postfit distributions of the three options for the MVA discriminants for a conditional $\mu_{\text{EW}\,VV+jj}^{\text{semileptonic}} = 1$ fit.



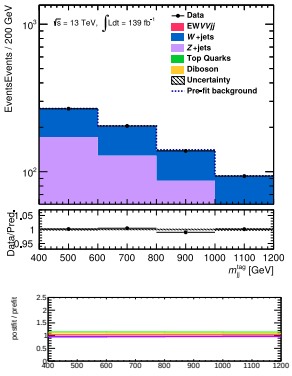
(a) RNN($5j$) CR



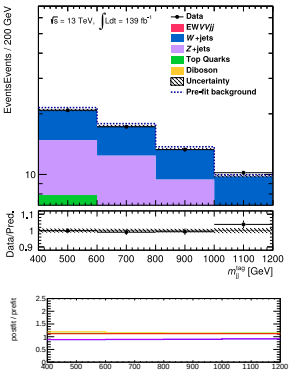
(b) RNN($5j$) resolved CR



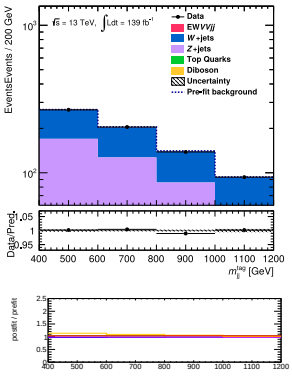
(c) RNN($4j$) CR



(d) RNN($4j$) resolved CR



(e) NN CR

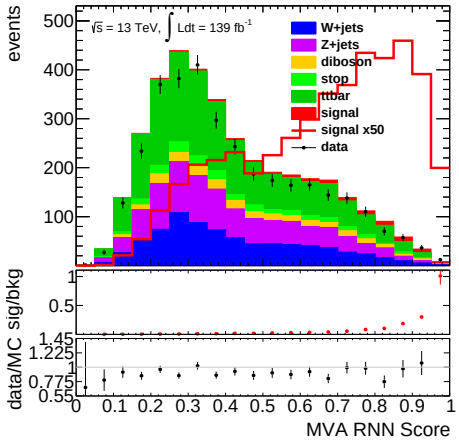


(f) NN resolved CR

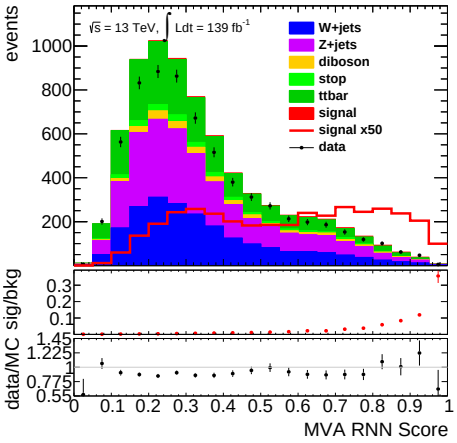
Figure A.7.21: Continuation from previous page.

A.8 . RNN($4j$)

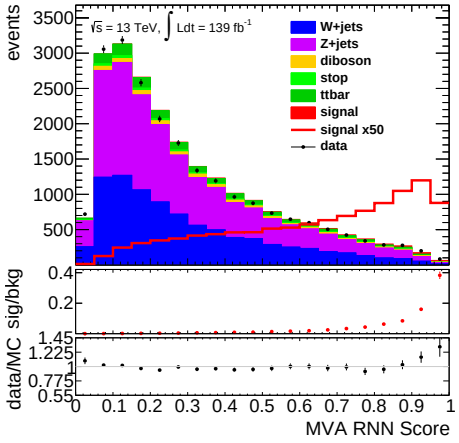
In addition to the baseline MVA of the anlysis (RNN($5j$)), all steps were also performed with a version of the RNN that uses only the leading 4 instead of 5 jets in the event (RNN($4j$)). This section of the appendix contains plots produces with the RNN($4j$) that correspond to those produces with the RNN($5j$) which were shown in Sec. 4.6 - Sec. 4.9.



(a) merged HP SR, RNN(4j)

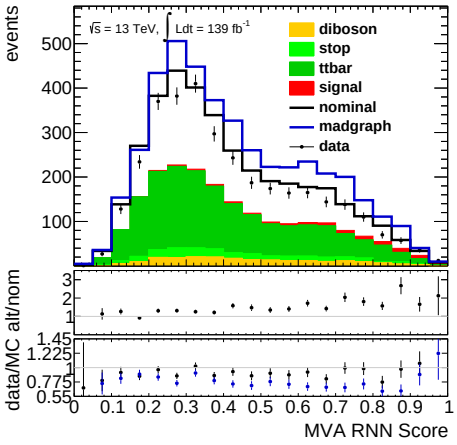


(b) merged LP SR, RNN(4j)

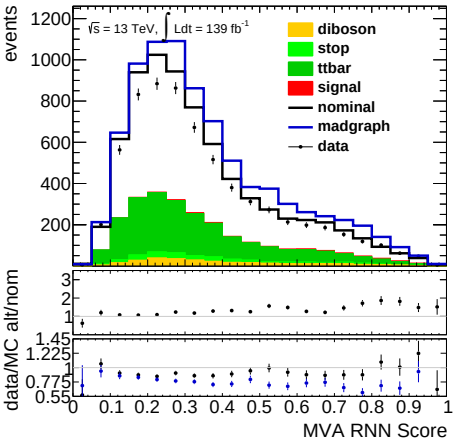


(c) resolved SR, RNN(4j)

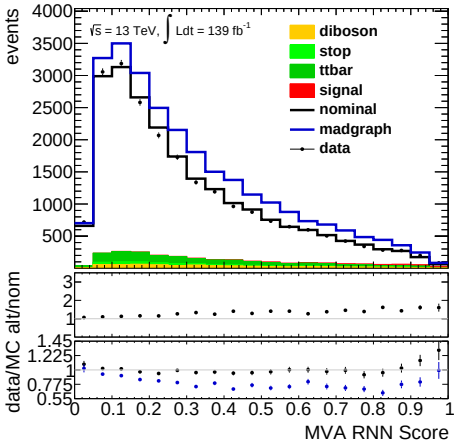
Figure A.8.22: MVA scores in all signal regions after training (see Fig. 4.19 for NN and RNN($5j$)).



(a) merged HP SR, RNN($4j$)

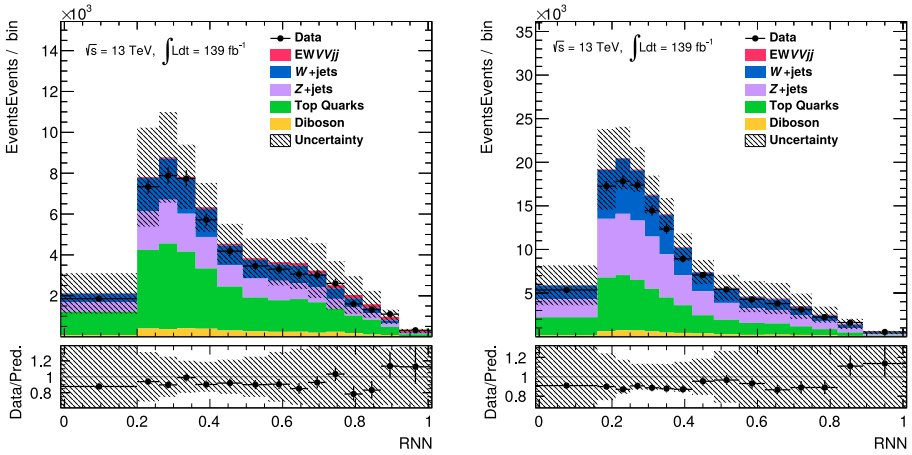


(b) merged LP SR, RNN($4j$)



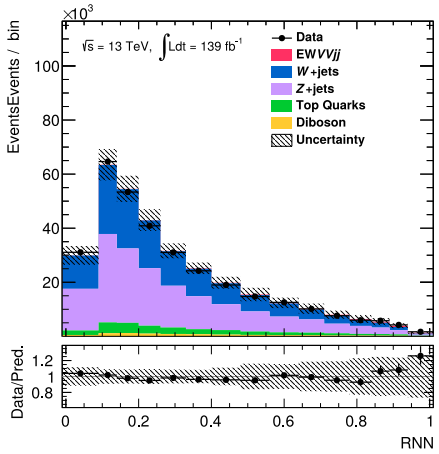
(c) resolved SR, RNN($4j$)

Figure A.8.23: Modelling differences of MVA scores between Sherpa nominal W/Z+jets samples and MadGraph alternative W/Z+jets samples (see Fig. 4.22 for NN and RNN($5j$)).



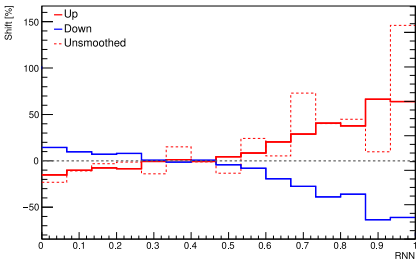
(a) RNN($4j$) merged HP SR

(b) RNN($4j$) merged LP SR

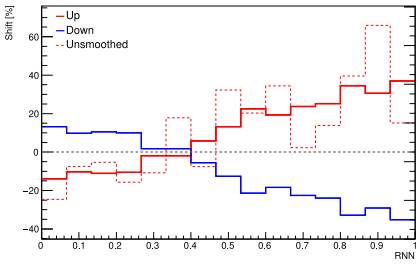


(c) RNN($4j$) resolved SR

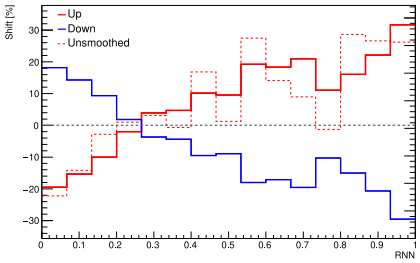
Figure A.8.24: Signal region prefit distributions of the three options for the MVA discriminants used in the final fit of the statistical interpretation (see Fig. 4.24 for NN and RNN($5j$)).



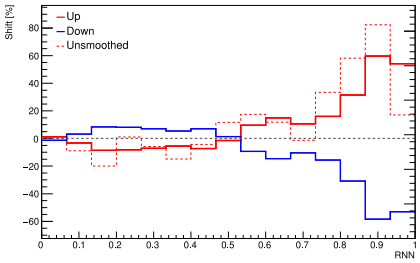
(a) Z+jets, merged HP SR



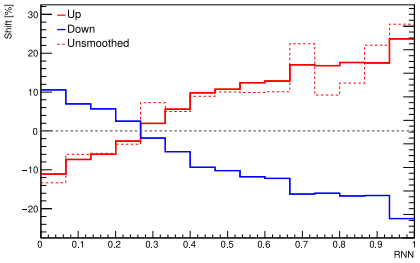
(b) W+jets, merged HP SR



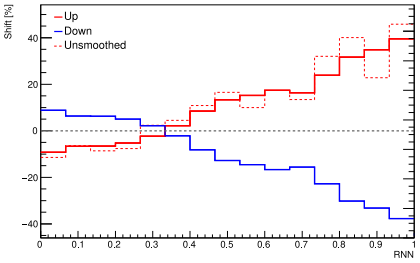
(c) Z+jets, merged LP SR



(d) W+jets, merged LP SR

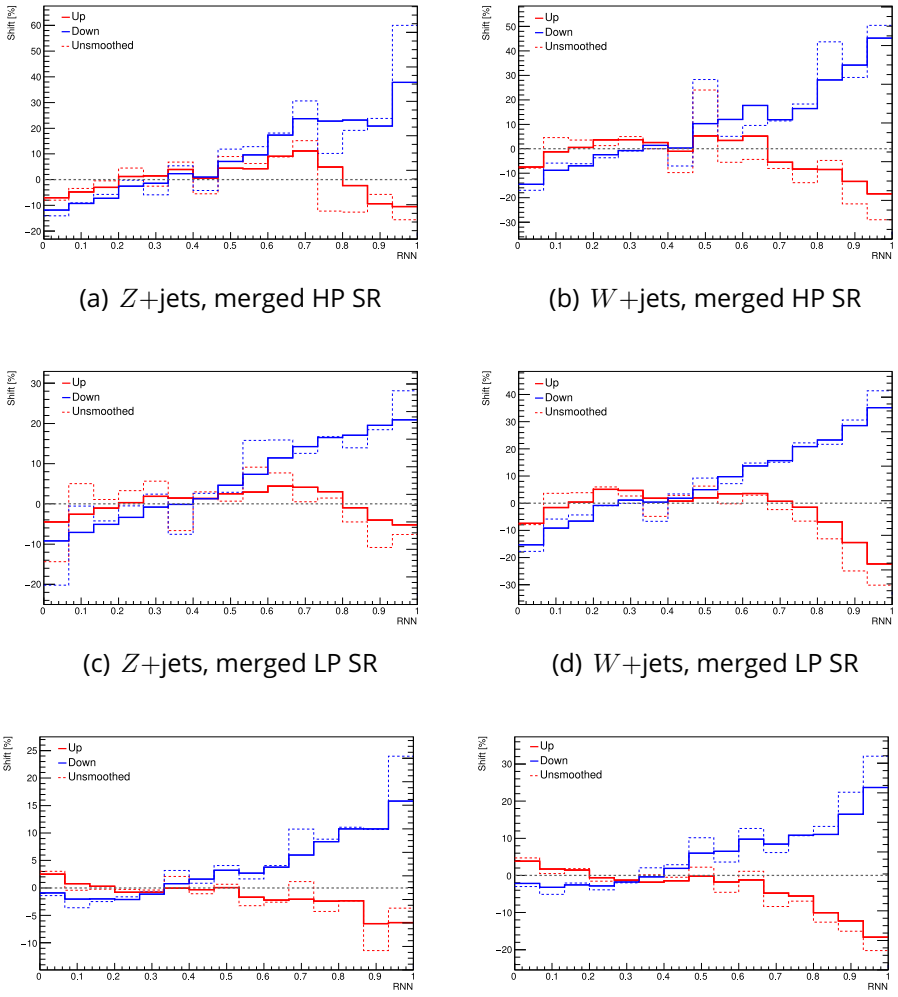


(e) Z+jets, resolved SR



(f) W+jets, resolved SR

Figure A.8.25: $\pm 1\ \sigma$ variation of the V+jets ($V = Z, W$) modelling systematic uncertainty $\mathrm{SysMODEL\_V\_MGPy8}$ on the Z+jets and W+jets background samples for the RNN($4j$) discriminant (see Fig. 4.27 (Z+jets) and A.5.8 (W+jets) for NN and RNN($5j$)).



(a) Z+jets, merged HP SR

(b) W+jets, merged HP SR

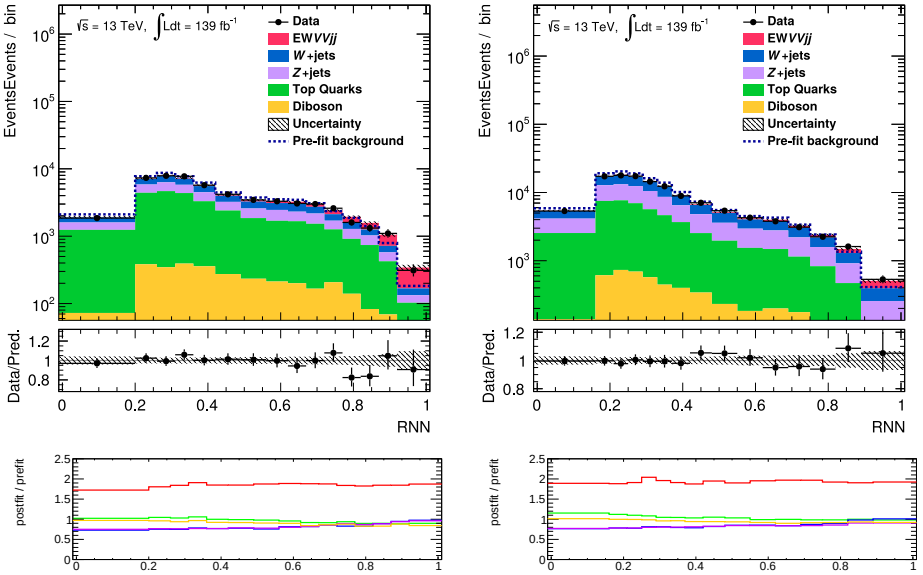
(c) Z+jets, merged LP SR

(d) W+jets, merged LP SR

(e) Z+jets, resolved SR

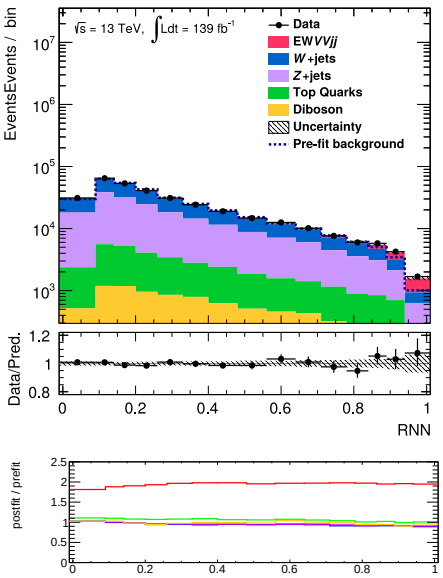
(f) W+jets, resolved SR

Figure A.8.26: ± 1 σ variation of the $m(jj)^{\text{tag}}$ reweighting systematic uncertainty $\mathrm{SysMJJREWEIGHT\_100per\_L0\_Fat1}$ ($\mathrm{SysMJJREWEIGHT\_100per\_L0\_J2}$) in the merged (resolved) regime on the Z+jets and W+jets background samples for the RNN($4j$) discriminant (see Fig. 4.28 (Z+jets) and A.5.9 (W+jets) for NN and RNN($5j$)).



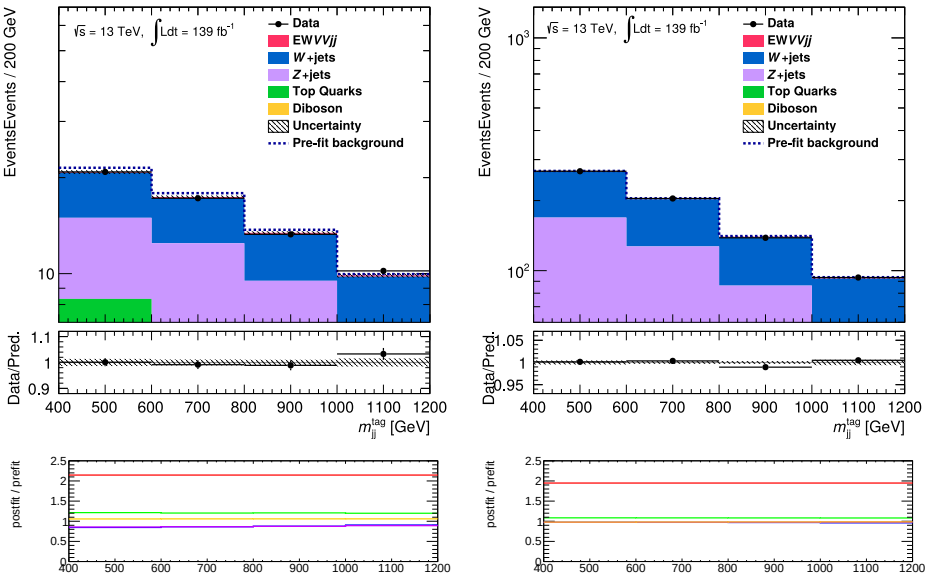
(a) RNN($4j$) merged HP SR

(b) RNN($4j$) merged LP SR



(c) RNN($4j$) resolved SR

Figure A.8.27: Signal region postfit distributions after performing fits using the RNN($4j$) discriminant (see Fig. 4.32 for NN and RNN($5j$)).



(a) RNN($4j$) merged CR (b) RNN($4j$) resolved CR

Figure A.8.28: Control region postfit distributions of the tagging jet mass $m(jj)^{tag}$ after performing fits using the RNN($4j$) discriminant (see Fig. A.8.28 for NN and RNN($5j$)).

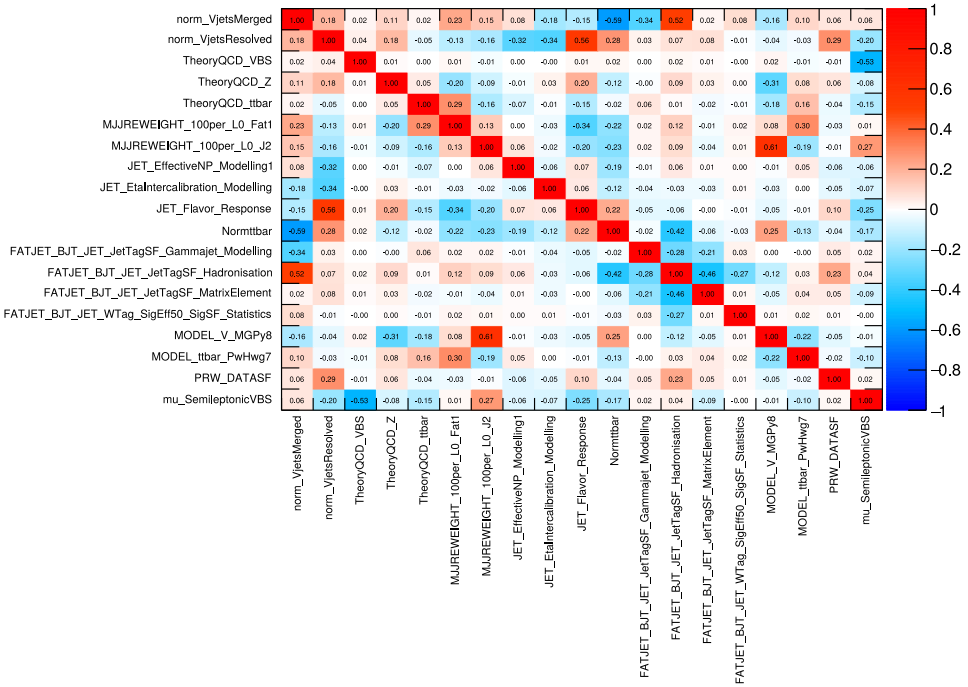


Figure A.8.29: Correlations of nuisance parameters in the final unblinded fit for the RNN($4j$) MVA (see Fig. 4.29 for NN and RNN($5j$)).

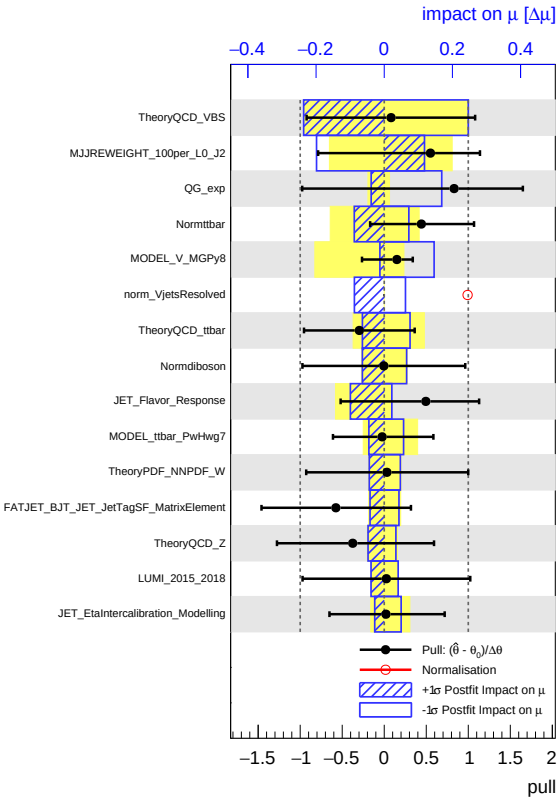


Figure A.8.30: Nuisance parameter (NP) ranking of the final unblinded fit using RNN($4j$). The impact of each individual NP on the parameter of interest μ is determined by fixing the NP to a value corresponding to $\pm 1\sigma$ of its uncertainty and redoing the fit with all but that NP (see Fig. 4.30 for NN and RNN($5j$)).

Source	Name in fit
Normalisations with priors	NormStop
	Normttbar
	Normdiboson
Modelling from alternative MC	SysMODEL_V_MGPy8
	SysMODEL_ttbar_PwHwg7
	SysMODEL_VV_PwPy
QCD scale	SysTheoryQCD_Z
	SysTheoryQCD_W
	SysTheoryQCD_ttbar
	SysTheoryQCD_stop
	SysTheoryQCD_VV
	SysTheoryQCD_VBS
PDF set	SysTheoryPDF_NNPDF_Z
	SysTheoryPDF_NNPDF_W
	SysTheoryPDF_NNPDF_ttbar
	SysTheoryPDF_NNPDF_stop
	SysTheoryPDF_NNPDF_VV
	SysTheoryPDF_NNPDF_VBS
	SysTheoryPDF_Z
	SysTheoryPDF_W
Flavour tagging	SysFT_EFF_Eigen_B_0
	SysFT_EFF_Eigen_B_1
	SysFT_EFF_Eigen_B_2
	SysFT_EFF_Eigen_C_0
	SysFT_EFF_Eigen_C_1
	SysFT_EFF_Eigen_C_2
	SysFT_EFF_Eigen_C_3
	SysFT_EFF_Eigen_Light_0
	SysFT_EFF_Eigen_Light_1
	SysFT_EFF_Eigen_Light_2
	SysFT_EFF_Eigen_Light_3
	SysFT_EFF_Eigen_Light_4
	SysFT_EFF_extrapolation_from_charm
	SysFT_EFF_extrapolation
$R = 0.4$ jets	SysJET_JERMC_DataVsMC_MC16
	SysJET_JERMC_EffectiveNP_1
	SysJET_JERMC_EffectiveNP_2
	SysJET_JERMC_EffectiveNP_3
	SysJET_JERMC_EffectiveNP_4
	SysJET_JERMC_EffectiveNP_5
	SysJET_JERMC_EffectiveNP_6
	SysJET_JERMC_EffectiveNP_7
	SysJET_JERMC_EffectiveNP_8
	SysJET_JERMC_EffectiveNP_9
	SysJET_JERMC_EffectiveNP_10
	SysJET_JERMC_EffectiveNP_11
	SysJET_JERMC_EffectiveNP_12restTerm
	SysJET_BJES_Response
	SysJET_EffectiveNP_Detector1
	SysJET_EffectiveNP_Detector2
	SysJET_EffectiveNP_Mixed1
	SysJET_EffectiveNP_Mixed2
	SysJET_EffectiveNP_Mixed3
	SysJET_EffectiveNP_Modelling1
	SysJET_EffectiveNP_Modelling2
	SysJET_EffectiveNP_Modelling3
	SysJET_EffectiveNP_Modelling4
	SysJET_EffectiveNP_Statistical1
	SysJET_EffectiveNP_Statistical2
	SysJET_EffectiveNP_Statistical3
	SysJET_EffectiveNP_Statistical4
	SysJET_EffectiveNP_Statistical5
	SysJET_EffectiveNP_Statistical6
	SysJET_EtaIntercalibration_Modelling
	SysJET_EtaIntercalibration_NonClosure_highE
	SysJET_EtaIntercalibration_NonClosure_negEta
	SysJET_EtaIntercalibration_NonClosure_posEta
	SysJET_EtaIntercalibration_TotalStat
	SysJET_Flavor_Composition
	SysJET_Flavor_Response
	SysJET_JvtEfficiency
	SysJET_Pileup_OffsetMu
	SysJET_Pileup_OffsetNPV
	SysJET_Pileup_PtTerm
	SysJET_Pileup_RhoTopology
	SysJET_PunchThrough_MC16
	SysJET_SingleParticle_HighPt

Table A.9.4: List of all nuisance parameters considered for the fits ordered by category (continued on next page).

Source	Name in fit
Pileup reweighting	SysPRW_DATASF
Luminosity	LUMI_2015_2018
Missing transverse Energy	SysMET_SoftTrk_Scale
	SysMET_JetTrk_Scale
$m(jj)^{\text{tag}}$ reweighting	SysMJJREWEIGHT_100per_L0_J2
	SysMJJREWEIGHT_100per_L0_Fat1
Track uncertainty	SysQG_exp
	SysQG_me
	SysQG_pdf
	SysQG_trackeff
	SysQG_fake
$R = 1.0$ jets	SysFATJET_CR_JET_CombMass_Baseline
	SysFATJET_CR_JET_CombMass_Modelling
	SysFATJET_CR_JET_CombMass_TotalStat
	SysFATJET_CR_JET_CombMass_Tracking1
	SysFATJET_CR_JET_CombMass_Tracking2
	SysFATJET_CR_JET_CombMass_Tracking3
	SysFATJET_CR_JET_EffectiveNP_R10_Detector1
	SysFATJET_CR_JET_EffectiveNP_R10_Detector2
	SysFATJET_CR_JET_EffectiveNP_R10_Mixed1
	SysFATJET_CR_JET_EffectiveNP_R10_Mixed2
	SysFATJET_CR_JET_EffectiveNP_R10_Mixed3
	SysFATJET_CR_JET_EffectiveNP_R10_Mixed4
	SysFATJET_CR_JET_EffectiveNP_R10_Modelling1
	SysFATJET_CR_JET_EffectiveNP_R10_Modelling2
	SysFATJET_CR_JET_EffectiveNP_R10_Modelling3
	SysFATJET_CR_JET_EffectiveNP_R10_Modelling4
	SysFATJET_CR_JET_EffectiveNP_R10_Statistical1
	SysFATJET_CR_JET_EffectiveNP_R10_Statistical2
	SysFATJET_CR_JET_EffectiveNP_R10_Statistical3
	SysFATJET_CR_JET_EffectiveNP_R10_Statistical4
	SysFATJET_CR_JET_EffectiveNP_R10_Statistical5
	SysFATJET_CR_JET_EffectiveNP_R10_Statistical6
	SysFATJET_CR_JET_EtaIntercalibration_Modelling
	SysFATJET_CR_JET_EtaIntercalibration_NonClosure_2018data
	SysFATJET_CR_JET_EtaIntercalibration_R10_TotalStat
	SysFATJET_CR_JET_Flavor_Composition
	SysFATJET_CR_JET_Flavor_Response
	SysFATJET_CR_JET_LargeR_TopologyUncertainty_V
	SysFATJET_CR_JET_LargeR_TopologyUncertainty_top
	SysFATJET_CR_JET_SingleParticle_HighPt
	SysFATJET_JER
	SysFATJET_JMR
	SysFATJET_SubR
Vector boson tagger	SysFATJET_BJT_bTag_B_0
	SysFATJET_BJT_bTag_Light_0
	SysFATJET_BJT_bTag_Light_1
	SysFATJET_BJT_JET_EffectiveNP_R10_Mixed1
	SysFATJET_BJT_JET_EffectiveNP_R10_Modelling1
	SysFATJET_BJT_JET_EtaIntercalibration_Modelling
	SysFATJET_BJT_JET_Flavor_Composition
	SysFATJET_BJT_JET_Flavor_Response
	SysFATJET_BJT_JET_JetTagSF_Dijet_Modelling
	SysFATJET_BJT_JET_JetTagSF_Gammajet_Modelling
	SysFATJET_BJT_JET_JetTagSF_Hadronisation
	SysFATJET_BJT_JET_JetTagSF_MatrixElement
	SysFATJET_BJT_JET_JetTagSF_Radiation
	SysFATJET_BJT_JET_WTag_SigEff50_BGSF_Dijet_Stat
	SysFATJET_BJT_JET_WTag_SigEff50_BGSF_Gammajet_Stat
	SysFATJET_BJT_JET_WTag_SigEff50_BGSF_Propagated_AllOthers
	SysFATJET_BJT_JET_WTag_SigEff50_SigSF_BinVariation
	SysFATJET_BJT_JET_WTag_SigEff50_SigSF_ExtrapolationPt
	SysFATJET_BJT_JET_WTag_SigEff50_SigSF_ExtrapolationZ
	SysFATJET_BJT_JET_WTag_SigEff50_SigSF_Propagated_AllOthers
	SysFATJET_BJT_JET_WTag_SigEff50_SigSF_Statistics
	SysFATJET_BJT_JET_WTag_SigEff50_TagEffUnc_GlobalBackground
	SysFATJET_BJT_JET_WTag_SigEff50_TagEffUnc_GlobalOther
	SysFATJET_BJT_JET_WTag_SigEff50_TagEffUnc_GlobalSignal
	SysFATJET_BJT_JET_WTag_SigEff80_BGSF_Dijet_Stat
	SysFATJET_BJT_JET_WTag_SigEff80_BGSF_Gammajet_Stat
	SysFATJET_BJT_JET_WTag_SigEff80_BGSF_Propagated_AllOthers
	SysFATJET_BJT_JET_WTag_SigEff80_SigSF_BinVariation
	SysFATJET_BJT_JET_WTag_SigEff80_SigSF_ExtrapolationPt
	SysFATJET_BJT_JET_WTag_SigEff80_SigSF_ExtrapolationZ
	SysFATJET_BJT_JET_WTag_SigEff80_SigSF_Propagated_AllOthers
	SysFATJET_BJT_JET_WTag_SigEff80_SigSF_Statistics
	SysFATJET_BJT_JET_WTag_SigEff80_TagEffUnc_GlobalBackground
	SysFATJET_BJT_JET_WTag_SigEff80_TagEffUnc_GlobalOther
	SysFATJET_BJT_JET_WTag_SigEff80_TagEffUnc_GlobalSignal

Table A.9.4: Continuation from previous page.

A.10 . Summary in English

The Standard Model of particle physics is as of yet the most accurate theory to describe the interactions of the thus far discovered fundamental particles. Its properties have been thoroughly tested in various experiments. However, many open questions remain: A large number of observations from cosmology and astronomy motivate the existence of dark matter, which as of today is not described within the Standard Model. While the strong, weak, and electromagnetic force are incorporated, the gravitational force is not represented in the Standard Model.

A large number of theories have been proposed over the past decades with the aim of extending the Standard Model to incorporate these and other phenomena. Many of these theories propose additional particles that would act in a similar way as the Standard Model Higgs boson. To obtain a hint as to which of these theories may be realised in nature it is important to study phenomena which are especially sensitive to small deviation introduced by such additions. One especially promising process of this kind is vector boson scattering (VBS), the scattering of W and Z bosons with each other. The main analysis presented in this thesis is a search for semileptonic vector boson scattering using the full Run-II data set of the ATLAS detector.

In addition to the main analysis, this thesis presents two additional projects. They concern possible improvements to certain aspects of the analysis which may improve the significance of an eventual follow-up analysis in the future. The first project concerns a study of a vector boson tagger similar to the one used in the analysis. It presents improvements in efficiency from the use of the novel Unified Flow Object (UFO) jet definition. The second project concerns various measurements of prototype sensors for the ITk, an upcoming upgrade of the inner detector (ID) which will greatly enhance the tracking capabilities of ATLAS.

Search for Semileptonic Vector Boson Scattering

The scattering of longitudinal polarisation components of W bosons played an important role for the initial motivation of the Higgs mechanism: Without considering the Higgs sector, calculations at tree level yield an amplitude of this process which is proportional to the centre of mass energy. It hence grows indefinitely with increasing energy in the collision. This eventually yields unphysical probabilities larger than one, a phenomenon known as unitarity violation. Only after introducing the Higgs mechanism unitarity is restored in the Standard Model: The diagrams mediated by vector bosons interfere with similar diagrams mediated by the Higgs boson. The latter exhibit a similarly diverging amplitude as the former. Both divergencies precisely cancel each other leading to a non-diverging amplitude. Hence, even small deviations from the Standard Model values of the involved coupling strengths of vector bosons to each other or to the Higgs boson have a large effect on the resulting amplitude. VBS processes are therefore an ideal probe of

the interplay between the electroweak and the Higgs sector of the Standard Model and for searches of eventual effects from physics beyond the Standard Model.

The semileptonic decay channel of vector boson scattering, which is studied in this thesis, has the advantage of the clean signature of the leptonically decaying vector boson in combination with the large branching fraction of the hadronic decay of the other vector boson. The analysis is split into three channels (0-, 1-, and 2-lepton) depending on the number of charged leptons in the final state. A combined analysis combining all channels is foreseen to be published in the near future. The study presented in this thesis however focuses on the 0-lepton channel alone. This channel is realised in the leptonically decaying vector boson decaying either into two neutrinos ($Z \to \nu\nu$) or into a neutrino and a charged lepton outside of the detector's acceptance ($W \to \nu\ell$).

The event selection makes use of the characteristic decay structure of vector boson scattering processes: The partons scattering off each other via the VBS process tend to leave detector signatures in forward regions (i.e. close to the beam pipe) in opposite hemispheres of the detector. They are reconstructed in the form of jets which in the context of the analysis are called tagging jets $(jj)^{\text{tag}}$. The hadronically decaying vector boson V_{had} decays into a pair of quarks which can either be reconstructed as two separate jets (called the resolved regime) or, in the case that they are strongly collimated, in the form of one large radius jet encompassing the full system (called the merged regime). In the 0-lepton channel, the leptonically decaying vector boson V_{lep} does not leave a direct signature in the detector and is instead reconstructed in the form of the missing transverse momentum in the event.

It is not possible to measure VBS processes alone. This is due to the interference with other electroweak processes with the same initial and final state. As a consequence the analysis selects for the electroweak (EW) production of a diboson system ($V_{\text{had}}V_{\text{lep}}$) in association with two jets ($(jj)^{\text{tag}}$): EW VVjj. A previous analysis using a 35.5fb^{-1} subset of the Run-II dataset obtained an observed (expected) significance of this process of $\sigma_{\text{obs}}^{\text{EW VVjj}} = 2.7$, $(\sigma_{\text{exp}}^{\text{EW VVjj}} = 2.5)$. The 0-lepton channel analysis presented in this thesis observes the process with $\sigma_{\text{obs}}^{\text{EW VVjj}} = 5.65$, $(\sigma_{\text{exp}}^{\text{EW VVjj}} = 3.03)$. This constitutes the first ever observation of the semileptonic EW VVjj process.

This result was obtained by a simultaneous binned maximum likelihood fit to a multivariate analysis (MVA) discriminant in three signal regions as well as the mass of the tagging jet system in two control regions. The control regions are designed to constrain the dominant $V+$ jets ($V = W/Z$) background which is modelled with Monte Carlo (MC) samples. Background distributions from $t\bar{t}$, single top, as well as QCD diboson production are modelled as well. The event selection is designed to reject QCD multijet background fully in data. As a consequence this background does not have to be modelled with MC samples in order to achieve a good agreement between data and simulation.

The MVA of the analysis makes use of a recurrent neural network (RNN). It uses a set of very low level features consisting of the four momentum components of the five p_{T}-leading jets in the event (and large radius jet in the merged regime) as well as the number of tracks associated to them. An alternative more conservative MVA is constructed as a simple feed forward neural network (NN) using a set of high-level features defined to be sensitive to the expected decay structure of the signal process. It is shown that the RNN method yields a higher signal/background separation due to the larger amount of information and high flexibility in combining it. The NN shows a smaller separation power but serves as a control for the RNN due to its smaller susceptibility to poorly modelled features. It was especially designed to be agnostic towards the jet multiplicity which exhibits an especially high discrepancy between data and MC simulation.

The statistical interpretation of the analysis was repeated using the NN score instead of RNN score as final discriminant. The resulting significance of $\sigma_{\mathrm{obs}}^{\mathrm{EW\ VVjj}} = 4.27$, $(\sigma_{\mathrm{exp}}^{\mathrm{EW\ VVjj}} = 2.73)$ is in accordance with the lower separation power of the NN compared to the RNN which achieved $\sigma_{\mathrm{obs}}^{\mathrm{EW\ VVjj}} = 5.65$, $(\sigma_{\mathrm{exp}}^{\mathrm{EW\ VVjj}} = 3.03)$. The signal strength of $\mu_{\mathrm{obs}}^{\mathrm{EW\ VVjj}} = 2.02 \pm 0.46$ obtained with the RNN shows a small deviation from the Standard Model value on the order of $2\ \sigma$. This is consistent with a similar upward pull of $\mu_{\mathrm{obs}}^{\mathrm{EW\ VVjj}} = 1.63 \pm 0.45$ which was obtained with the NN. The analysis is dominated by systematics: 80% of the uncertainty on the value of $\mu_{\mathrm{obs}}^{\mathrm{EW\ VVjj}}$ stem from various systematic uncertainties while only 20% stem from the statistical uncertainty on the observed data. This is the case for both the NN and RNN. The largest contribution to the systematic uncertainty is given by the QCD scale uncertainty on the signal MC sample. Especially using the baseline RNN, the experimental uncertainty is dominated by various uncertainties associated to the jets in the event.

Vector Boson Tagging

A vector boson tagger is used in the VBS analysis presented in this thesis in order to identify events in which the hadronically decaying vector boson candidate V_{had} is actually initiated by a vector boson and not e.g. by quarks or gluons. A separate study on such a taggers is presented. It is demonstrated that by utilizing the novel Unified Flow Objects (UFOs) jets, a significantly higher background separation by a factor of $2-3$ may be achieved compared to the use of locally reweighted topological cluster (LCTopo) jets which were utilised in the VBS analysis.

LCTopo jets only make use of calorimetry information while UFO jets combine the information of calorimeter and tracking detector. This process is called particle flow. While various particle flow techniques have been proposed in the past, UFO jets can be understood as a combination of such techniques in order to achieve the best possible performance over a large spectrum of transverse momentum. Previous studies have shown that UFO jets achieve a significantly higher resolution of various jet substructure variables which may be used for jet tagging. The presented study

demonstrates how this translates into a better efficiency of such taggers.

A neural network based tagger using a large list of such substructure variables as features is studied. The results are compared to a 3-variable tagger similar to the one used in the main analysis of this thesis.

Additionally the possibility of mass decorrelation was studied for the neural network based tagger. It was demonstrated that by introducing an adversarial neural network in the training process, decorrelation of the tagger score to the reconstructed jet mass can be achieved. This is important for analyses that aim to perform a data driven background estimate of the jet mass distribution and are therefore sensitive to a bias introduced by a tagger which is correlated to this variable. Expectedly a significant increase in tagging efficiency is observed when introducing mass decorrelation to the tagger.

As a result of the studies presented in this thesis, the neural network based tagger for UFO jets is now available to be used in the usual software framework for the ATLAS collaboration. Once a complete calibration of UFO jets including in-situ corrections from data is available, scale factors for the tagger will be derived. This is foreseen for the near future, after which such tagger may be used in an eventual follow-up analysis to the main analysis of this thesis.

ITk Planar Pixel Sensors

The tagging jets $(jj)^{\text{tag}}$ are a crucial part for the identification of VBS processes and, as mentioned above, play an important role in the event selection of the VBS analysis presented in this thesis. Due to their characteristic position in the forward region of the detector, an eventual follow-up analysis would greatly benefit from better instrumentation in this region. The ITk, an upgrade of the current inner detector (ID) foreseen for the High-Luminosity LHC (HL-LHC) phase starting in 2029, promises such improvement. It will increase the tracking acceptance of the ATLAS detector in the high-forward region up to a pseudorapidity of $\eta < 4$ compared to the ID which currently achieves an acceptance of $\eta < 2.5$. The ITk will be instrumented with pixel and strip silicon sensors.

The third project presented in this thesis concerns various measurements undertaken on prototypes of planar pixel sensors for the ITk. They were performed in the frame of the ITk Planar Pixel Sensor Market Survey, an effort to qualify vendors for the final production of such sensors for the ITk. The presented measurements can be divided into two groups:

The first kind of measurements concern bare sensors and were performed at a probe station in a clean room at the IJCLab. Results of visual inspections, geometrical measurements of thickness and bow, as well as electrical characteristics are presented. The electrical characteristics include typical measurements performed on silicon-based particle detectors such as leakage current, breakdown voltage, and depletion voltage.

The second kind of measurements were performed on modules, i.e. sensors

which are connected to front-end readout chips which can be operated as fully functioning particle detectors. The hit efficiency of such modules was measured using the electron beam of the DESY II test beam facility in Hamburg. A beam telescope was used to reconstruct the tracks of individual electrons in the beam. The reconstructed tracks were compared with the measured particle hits registered by the modules under test. These measurements were performed on non-irradiated modules as well as modules after they have been irradiated with a fluence of either 2 or 5×10^{15} $\frac{\text{neq}}{\text{cm}^2}$. The latter corresponds to the maximum expected fluence such modules will be exposed to after the full run of the HL-LHC.

The presented results demonstrate that most of the measured modules are able to fulfill the efficiency requirements for the ITk even after irradiation. A few exceptions are presented for which the efficiency falls below the requirement. Individual investigations of such exceptions are presented demonstrating possible causes for the inefficiency.

The various presented measurements on the sensor prototypes were taken as a basis for the final decision of the Market Survey. As a consequence three vendors were chosen for the final production. In general the prototypes showed the required properties to allow an operation of the ITk pixel detector with the expected performance. A pre-production, in which 10% of the total number of sensors are produced, is ongoing at the time of writing and the full production is expected to commence soon after.

Outlook

The main analysis presented in this thesis constitutes the first observation of the electroweak production of a diboson system in association with a high-mass dijet system in the semileptonic decay channel. The associated observed (expected) significance of $\sigma_{\text{obs}}^{\text{EW VVjj}} = 5.65$, $(\sigma_{\text{exp}}^{\text{EW VVjj}} = 3.03)$ was achieved in the 0-lepton channel alone. A combination of this with the two remaining analysis channels (1-, and 2-lepton) is foreseen for the near future. It is expected that this combination may achieve a significantly higher expected significance.

The presented observed signal strength of $\mu_{\text{obs}}^{\text{EW VVjj}} = 2.02 \pm 0.46$ constitutes an upward pull on the order of two sigma with respect to the Standard Model prediction. In the case that no such pull or even a downward pull is observed in the 1-, and 2-lepton channel, the observed significance of the combined analysis may decrease with respect to the presented 0-lepton only analysis. If however a similar upward pull with high enough significance were observed in all channels this may be interpreted as an indication for physics beyond the Standard Model. An effective field theory (EFT) study with respect to extensions to the Standard Model related to anomalous quartic gauge couplings between W and Z bosons is ongoing in parallel and will be included in the combined analysis.

Both the vector boson tagging studies as well as the ITk planar pixel sensor measurements indicate that eventual follow-up analyses of the presented VBS anal-

ysis may greatly benefit from these projects. The increase in tagging efficiency by using UFO jets instead of LCTopo jets for the vector boson tagger may lead to a significantly higher signal/background ratio in the signal regions of the analysis. This may already come into effect in an eventual follow-up analysis with Run-III data since a fully calibrated version of the UFO based tagger is expected to be available in the near future. The greatly increased tracking acceptance in the forward region of the ITk with respect to the ID is especially important for the characteristic tagging jets of the VBS process. The various presented measurements on the planar pixel sensor prototypes demonstrated their capability to achieve the necessary requirements for this even after being irradiated to the maximum expected fluence of the full HL-LHC run.

A.11 . Résumé en Français

Le modèle standard de la physique des particules est à ce jour la théorie la plus précise pour décrire les interactions des particules fondamentales découvertes jusqu'à présent. Ses propriétés ont été minutieusement testées dans diverses expériences. Cependant, de nombreuses questions ouvertes demeurent: Un grand nombre d'observations issues de la cosmologie et de l'astronomie motivent l'existence de la matière noire, qui à ce jour n'est pas décrite dans le modèle standard. Alors que les forces fortes, faibles et électromagnétiques sont incorporées, la force gravitationnelle n'est pas représentée dans le modèle standard.

Un grand nombre de théories ont été proposées au cours des dernières décennies dans le but d'étendre le modèle standard pour incorporer ces phénomènes et d'autres. Beaucoup de ces théories proposent des particules supplémentaires qui agiraient de la même manière que le boson de Higgs du modèle standard. Pour obtenir une indication quant à laquelle de ces théories peut être réalisée dans la nature, il est important d'étudier les phénomènes qui sont particulièrement sensibles aux petites déviations introduites par de telles additions. Un processus particulièrement prometteur de ce type est la diffusion des bosons vecteurs W et Z (vector boson scattering : VBS). L'analyse principale présentée dans cette thèse est une recherche de la diffusion de bosons vecteurs en état final semi-leptonique en utilisant l'ensemble de données Run-II enregistré par le détecteur ATLAS.

En plus de l'analyse principale, cette thèse présente deux projets supplémentaires concernant les possibilités d'amélioration des futures analyses. Le premier projet concerne l'étude d'un tagger de bosons vecteurs similaire à celui utilisé dans l'analyse. Il présente une meilleure efficacité grâce à l'utilisation de la nouvelle définition de jets qui utilisent des Unified Flow Objects (UFOs). Le deuxième projet concerne diverses mesures de prototypes de capteurs pour l'ITk, une prochaine mise à niveau du détecteur interne (ID) qui améliorera considérablement les capacités d'ATLAS.

Recherche de la Diffusion des Bosons Vecteurs en État Final Semi-Leptonique

La diffusion des composantes de polarisation longitudinale des bosons W a joué un rôle important dans la motivation initiale du mécanisme de Higgs: Sans considérer le secteur de Higgs, les calculs au niveau de l'arbre donnent une amplitude de ce processus qui est proportionnelle à l'énergie du centre de masse. Il croît donc indéfiniment avec l'augmentation de l'énergie dans la collision. Cela donne finalement des probabilités non physiques supérieures à 100%. Ce phénomène est appelé violation d'unitarité. Ce n'est qu'après l'introduction du mécanisme de Higgs que l'unitarité est restaurée dans le modèle standard: Les diagrammes médiés par les bosons vecteurs interfèrent avec des diagrammes similaires médiés par le boson de Higgs. Ces derniers présentent une amplitude tout aussi divergente que les pre-

miers. Les deux divergences s'annulent précisément, conduisant à une amplitude non divergente. Par conséquent, même de petites différences par rapport aux prédictions du modèle standard ont un effet important sur l'amplitude résultante. Les processus VBS sont donc une sonde idéale de l'interaction entre le secteur électrofaible et le secteur de Higgs du Modèle Standard et pour la recherche d'éventuels effets de la physique au-delà du Modèle Standard.

Le canal semileptonique de la diffusion de bosons vecteurs, qui est étudié dans cette thèse, présente l'avantage de la signature claire de la désintégration leptonique d'un boson vecteur (V_{lep}) en combinaison avec la grande probabilité d'embranchement de la désintégration hadronique de l'autre boson vecteur (V_{had}). L'analyse est divisée en trois canaux (0-, 1- et 2-lepton) en fonction du nombre de leptons chargés dans l'état final. Une analyse combinée combinant tous les canaux devrait être publiée prochainement. Mais l'étude présentée dans cette thèse se concentre sur le seul canal 0-lepton. Ce canal est réalisé dans la désintégration de V_{lep} en deux neutrinos ($Z \to \nu\nu$) ou en un neutrino et un lepton chargé en dehors de l'acceptation du détecteur ($W \to \nu\ell$) .

La sélection des événements utilise la structure de désintégration caractéristique des processus de diffusion des bosons vecteurs: Les partons inter agissant via le processus VBS ont tendance à laisser des signatures de détecteur dans les régions avant (c'est-à-dire à proximité du tube de faisceau) dans les hémisphères opposés du détecteur. Ils sont reconstruits sous forme de jets qui dans le cadre de l'analyse sont appelés tagging jets $(jj)^{\text{tag}}$. Le boson vecteur à désintégration hadronique V_{had} se désintègre en une paire de quarks qui peuvent soit être reconstruits sous la forme de deux jets séparés (appelés régime résolu), soit, dans le cas où ils sont fortement collimatés, sous la forme d'un jet à grand rayon englobant le système complet (appelé régime fusionné). Dans le canal 0-lepton, le boson vecteur à désintégration leptonique V_{lep} ne laisse pas de signature directe dans le détecteur. Il est plutôt reconstruit comme l'impulsion transverse manquante dans l'événement.

Il n'est pas possible de mesurer les processus VBS seuls. Cela est dû à l'interférence avec d'autres processus électrofaibles avec le même état initial et final. En conséquence, l'analyse sélectionne la production électrofaible (electroweak: EW) d'un système de dibosons ($V_{\text{had}}V_{\text{lep}}$) en association avec deux jets ($(jj)^{\text{tag}}$): EW VVjj. Une analyse précédente utilisant un sous-ensemble 35.5fb^{-1} de Run-II a obtenu une signification observée (attendue) de ce processus de $\sigma_{\text{obs}}^{\text{EW VVjj}} = 2.7$, $(\sigma_{\text{att}}^{\text{EW VVjj}} = 2.5)$. L'analyse du canal 0-lepton présentée dans cette thèse observe le processus avec $\sigma_{\text{obs}}^{\text{EW VVjj}} = 5.65$, $(\sigma_{\text{att}}^{\text{EW VVjj}} = 3.03)$. Ceci constitue la toute première observation du processus semileptonique EW VVjj.

Ce résultat a été obtenu par un fit simultané du maximum de vraisemblance à un discriminant d'analyse multivariée dans trois régions de signal et de la masse du système de tagging jets ($m(jj)^{\text{tag}}$) dans deux régions de contrôle. Les régions de contrôle sont conçues pour contraindre le bruit de fond dominant: V+jets ($V = W/Z$) qui est modélisé avec la simulation de Monte Carlo (MC). Les distri-

butions de fond de $t\bar{t}$, du production simple t, ainsi que la production de dibosons QCD sont également modélisées. La sélection d'événement est conçue pour rejeter entièrement le bruit de fond QCD multijet dans les données. En conséquence, ce fond n'a pas besoin d'être modélisé.

L'analyse multivariée (multivariate analysis: MVA) utilise un réseau neuronal récurrent (recurrent neural network: RNN). Il utilise un ensemble de caractéristiques de base composé des quatre composantes d'impulsion des cinq jets de tête p_{T} dans l'événement (et jet à grand rayon dans le régime fusionné) ainsi que le nombre de traces associées pour eux. Un MVA alternatif plus conservateur est construit comme un simple réseau de neurones à action directe (feed forward neural network: NN) utilisant un ensemble de variables plus élaborées, définies pour être sensibles à la structure de processus de signal. Il est démontré que la méthode RNN produit une séparation signal/fond plus élevée en raison de la plus grande quantité d'informations et de la grande flexibilité de leur combinaison. Le NN montre une moindre séparation mais sert de contrôle pour le RNN en raison de sa plus faible sensibilité aux caractéristiques mal modélisées. Il a été spécialement conçu pour être agnostique vis-à-vis de la multiplicité des jets qui présente un écart particulièrement élevé entre les données et la simulation MC.

L'interprétation statistique de l'analyse a été répétée en utilisant le score NN au lieu du score RNN comme discriminant final. La signification résultante de $\sigma_{\mathrm{obs}}^{\mathrm{EW\ VVjj}} = 4.27$, $(\sigma_{\mathrm{att}}^{\mathrm{EW\ VVjj}} = 2.73)$ est conforme à au pouvoir de séparation plus faible du NN par rapport au RNN qui atteint $\sigma_{\mathrm{obs}}^{\mathrm{EW\ VVjj}} = 5.65$, $(\sigma_{\mathrm{att}}^{\mathrm{EW\ VVjj}} = 3.03)$. La force du signal de $\mu_{\mathrm{obs}}^{\mathrm{EW\ VVjj}} = 2.02 \pm 0.46$ obtenue avec le RNN montre un petit écart par rapport à la valeur du modèle standard de l'ordre de $2\ \sigma$. Ceci est cohérent avec une tendance vers le haut similaire de $\mu_{\mathrm{obs}}^{\mathrm{EW\ VVjj}} = 1.63 \pm 0.45$ qui a été obtenue avec le NN. L'analyse est dominée par la systématique : 80 $\%$ de l'incertitude sur la valeur de $\mu_{\mathrm{obs}}^{\mathrm{EW\ VVjj}}$ proviennent de diverses incertitudes systématiques, tandis que seulement 20% proviennent de l'incertitude statistique sur les données observées. C'est le cas pour les deux discriminant, NN et RNN. La plus grande contribution à l'incertitude systématique est donnée par l'incertitude de la QCD scale sur le simulation de signal MC. En particulier en utilisant le RNN, l'incertitude expérimentale est dominée par diverses incertitudes associées aux jets dans l'événement.

Taggers des Bosons Vecteurs

Un tagger de boson vecteur est utilisé dans l'analyse VBS présentée dans cette thèse afin d'identifier les événements dans lesquels le candidat boson vecteur à désintégration hadronique V_{had} est réellement initié par un boson vecteur et non par example par des quarks ou des gluons. Une étude distincte sur ces taggers est présentée. Il est démontré qu'en utilisant les nouveaux jets d'Unified Flow Objects (UFO), une séparation de fond significativement plus élevée d'un facteur de 2 à 3 peut être obtenue par rapport à l'utilisation de jets de Locally Reweighted

Topological Clusters (LCTopo) qui ont été utilisés dans le analyse VBS.

Les jets LCTopo utilisent uniquement les informations du calorimètre. Mais les jets UFO combinent les informations du calorimètre et du détecteur de traces. Ce processus est appelé flux de particules (particle flow: PFlow). Alors que diverses techniques de PFlow ont été proposées dans le passé, les jets UFO peuvent être compris comme une combinaison de ces techniques afin d'obtenir les meilleures performances possibles sur un large spectre d'impulsion transverse. Des études antérieures ont montré que les jets UFO atteignent une résolution significativement plus élevée de diverses variables de sous-structure de jet qui peuvent être utilisées pour le tagging des jets. L'étude présentée montre comment cela se traduit par une meilleure efficacité de ces taggers.

Un tagger basé sur un réseau neuronal utilisant une longue liste de variables de sous-structure est étudié. Les résultats sont comparés à un tagger à 3 variables similaire à celui utilisé dans l'analyse principale de cette thèse.

De plus, la possibilité d'une décorrélation de masse a été étudiée pour le tagger. Il a été démontré qu'en introduisant un réseau neuronal contradictoire (adversarial neural network: ANN) dans le processus, la décorrélation du score de tagger à la masse de jet reconstruite peut être obtenue. Ceci est important pour les analyses qui visent à effectuer une estimation de fond basée sur les données de la distribution de la masse du jet et sont donc sensibles à un biais introduit par un tagger qui est corrélé à cette variable. On s'attend à ce qu'une diminution significative de l'efficacité du marquage soit observée lors de l'introduction d'une décorrélation de masse.

À la suite des études présentées dans cette thèse, le tagger basé sur un réseau de neurones pour les jets UFO est maintenant disponible pour être utilisé dans le software habituel de la collaboration ATLAS. Une fois qu'un étalonnage complet des jets UFO, y compris des corrections in situ à partir des données, sera disponible, les facteurs d'échelle (scale factors) pour le tagger seront dérivés. Ceci est prévu dans un avenir proche, après quoi un tel tagger pourra être utilisé dans une éventuelle analyse de suivi de l'analyse principale de cette thèse.

Capteurs de Pixels Planaires ITk

Les tagging jets $(jj)^{\text{tag}}$ sont un élément crucial pour l'identification des processus VBS et, comme mentionné ci-dessus, jouent un rôle important dans la sélection des événements de l'analyse VBS présentée dans cette thèse. En raison de leur position caractéristique dans la région avant du détecteur, cette analyse bénéficierait grandement d'une meilleure instrumentation dans cette région. L'ITk, une mise à niveau du détecteur interne (ID) actuel prévue pour la phase du LHC à haute luminosité (High-Luminosity LHC: HL-LHC) à partir de 2029, promet une telle amélioration. Il augmentera l'acceptance géométrique du détecteur ATLAS jusqu'à une pseudorapidité de $\eta < 4$ par rapport à l'ID qui atteint actuellement $\eta < 2.5$. L'ITk sera instrumenté avec des capteurs de silicium à pixels et à strips.

Le troisième projet présenté dans cette thèse concerne différentes mesures réalisées sur des prototypes de capteurs planaires à pixels pour l'ITk. Ils ont été réalisés dans le cadre de l'étude de marché ITk Planar Pixel Sensor (ITk Planar Pixel Market Survey: MS), un effort visant à qualifier les fournisseurs pour la production finale de tels capteurs pour l'ITk. Les mesures présentées peuvent être divisées en deux groupes:

Le premier type de mesures concerne des capteurs nus et a été effectué avec une machine à pointe dans une salle blanche à l'IJCLab. Des résultats d'inspections visuelles, des mesures géométriques d'épaisseur et de planéité, ainsi que des caractéristiques électriques sont présentés. Les caractéristiques électriques comprennent des mesures typiques effectuées sur des détecteurs de particules à base de silicium telles que le courant de fuite, la tension de claquage et la tension de dépletion de charge.

Le deuxième type de mesures a été effectué sur des modules, c'est-à-dire des capteurs connectés à des puces de lecture frontale pouvant être utilisées comme des détecteurs de particules entièrement fonctionnels. L'efficacité d'impact de ces modules a été mesurée à l'aide du faisceau d'électrons de l'installation de faisceaux d'essai au laboratoire DESY à Hambourg. Un télescope à faisceau a été utilisé pour reconstruire les traces d'électrons individuels dans le faisceau. Les pistes reconstruites ont été comparées aux impacts de particules mesurés enregistrés par les modules testés. Ces mesures ont été effectuées sur des modules non irradiés ainsi que sur des modules après qu'ils aient été irradiés avec une fluence de 2 ou 5×10^{15} $\frac{\text{neq}}{\text{cm}^2}$. Cette dernière correspond à la fluence maximale attendue à laquelle ces modules seront exposés après l'exploitation complète du HL-LHC.

Les résultats présentés démontrent que la plupart des modules mesurés sont capables de répondre aux exigences d'efficacité pour l'ITk même après irradiation. Quelques exceptions sont présentes pour lesquelles l'efficacité est inférieure à l'exigence. Des enquêtes individuelles sur ces exceptions sont présentées, démontrant les causes possibles de l'inefficacité.

Les différentes mesures présentées sur les prototypes de capteurs ont servi de base à la décision finale de l'étude de marché. En conséquence, trois fournisseurs ont été choisis pour la production finale. En général, les prototypes ont montré les propriétés requises pour permettre un fonctionnement du détecteur de pixels ITk avec les performances attendues. Une pré-production, dans laquelle 10% du nombre total de capteurs sont produits, est en cours au moment de la rédaction et la production complète devrait commencer peu de temps après.

Perspectives

L'analyse principale présentée dans cette thèse constitue la première observation de la production électrofaible d'un système de dibosons en association avec un système de dijet de masse élevée dans le canal de désintégration semileptonique. La signification associée observée (attendue) de $\sigma_{\text{obs}}^{\text{EW VVjj}} = 5.65$, $(\sigma_{\text{att}}^{\text{EW VVjj}} = 3,03)$

a été obtenue en le canal 0-lepton seul. Une combinaison de cela avec les deux canaux d'analyse restants (1- et 2-lepton) est prévue dans un proche avenir. On s'attend à ce que cette combinaison puisse atteindre une signification attendue significativement plus élevée.

L'intensité du signal observé $\mu_{\text{obs}}^{\text{EW VVjj}} = 2.02 \pm 0.46$ constitue une fluctuation vers le haut de l'ordre de deux sigma par rapport à la prédiction du modèle standard. Dans le cas où aucune déviation similaire n'est observée dans le canal 1 et 2-lepton, la signification observée de l'analyse combinée peut diminuer par rapport à l'analyse 0-lepton présentée. Dans le cas contraire, cela pourrait être interprété comme une indication de la physique au-delà du modèle standard. Une étude de la théorie des champs effectifs (effective field theory: EFT) concernant les extensions du modèle standard liées aux couplages anormaux de jauge quartique entre les bosons W et Z est en cours en parallèle et sera incluse dans l'analyse combinée.

Les études de taggers des bosons vecteurs ainsi que les mesures des capteurs de pixels planaires ITk indiquent que les analyses de suivi éventuelles de l'analyse VBS présentée pourraient grandement bénéficier de ces projets. L'augmentation de l'efficacité de tagging en utilisant des jets UFO au lieu de jets LCTopo pour le tagger de boson vecteur peut conduire à un rapport signal/fond significativement plus élevé dans les régions de signal de l'analyse. Cela pourrait déjà entrer en vigueur dans une éventuelle analyse de suivi avec les données du Run-III puisqu'une version entièrement calibrée du tagger basé sur les UFOs devrait être disponible dans un avenir proche. L'acceptance de suivi considérablement accrue dans la région avant de l'ITk par rapport à l'ID est particulièrement importante pour les tagging jets caractéristiques du processus VBS. Les différentes mesures présentées sur les prototypes de capteurs à pixels plans ont démontré leur capacité à répondre aux exigences nécessaires, même après avoir été irradiés à la fluence maximale attendue de l'ensemble de l'exploitation du HL-LHC.

Bibliography

[1] The ATLAS Collaboration. "Search for the electroweak diboson production in association with a high-mass dijet system in semileptonic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. D* 100.3 (2019), p. 032007. doi: 10.1103/PhysRevD.100.032007. arXiv: 1905.07714 [hep-ex].

[2] J. J. Thomson. "Cathode Rays". In: *Phil. Mag. Ser. 5* (1897).

[3] E. Rutherford. "The scattering of alpha and beta particles by matter and the structure of the atom". In: *Phil. Mag. Ser. 6* 21 (1911), pp. 669–688. doi: 10.1080/14786440508637080.

[4] J. Chadwick. "Possible Existence of a Neutron". In: *Nature* 129 (1932), p. 312. doi: 10.1038/129312a0.

[5] Max Planck. "The Theory of Heat Radiation". Philadelphia, P. Blakiston's Son & Co, 1914.

[6] Albert Einstein. "Concerning an heuristic point of view toward the emission and transformation of light". In: *Annalen Phys.* 17 (1905), pp. 132–148.

[7] Boris Lemmer. "Measurement of Spin Correlations in $t\bar{t}$ Events from pp Collisions at $\sqrt{s}$ = 7 TeV in the Lepton + Jets Final State with the ATLAS Detector". II.PHYSIK-UNIGÖ-DISS-2014-02. 2014. arXiv: 1410.1791 [hep-ex].

[8] Ernst Carl Gerlach Stückelberg. "La signification du temps propre en mécanique ondulatoire". In: *Helv. Phys.* 14 (1941), pp. 322–323. doi: 10.1007/978-3-7643-8878-2_18.

[9] R. P. Feynman. "Space-Time Approach to Non-Relativistic Quantum Mechanics". In: *Rev. Mod. Phys.* 20 (2 Apr. 1948), pp. 367–387. doi: 10.1103/RevModPhys.20.367.

[10] C.D. Anderson. "The Positive Electron". In: *Phys. Rev.* 43 (1933), pp. 491–494. doi: 10.1103/PhysRev.43.491.

[11] O. Chamberlain et al. "Observation of Anti-protons". In: *Phys. Rev.* 100 (1955), pp. 947–950. doi: 10.1103/PhysRev.100.947.

[12] B. Cork et al. "Anti-neutrons Produced From Anti-protons in Charge Exchange Collisions". In: *Phys. Rev.* 104 (1957), pp. 1193–1197. doi: 10.1103/PhysRev.104.1193.

[13] Hideki Yukawa. “On the Interaction of Elementary Particles I”. In: *Proc. Phys. Math. Soc. Jap.* 17 (1935), pp. 48–57. doi: 10.1143/PTPS.1.1.

[14] S.H. Neddermeyer and C.D. Anderson. “Note on the Nature of Cosmic Ray Particles”. In: *Phys. Rev.* 51 (1937), pp. 884–886. doi: 10.1103/PhysRev.51.884.

[15] J. C. Street Street and Stevenson Edgar. “New Evidence for the Existence of a Particle of Mass Intermediate Between the Proton and Electron”. In: *Phys. Rev.* (1937). doi: 10.1103/PhysRev.52.1003.

[16] C.M.G. Lattes et al. “Processes Involving Charged Mesons”. In: *Nature* 159 (1947), pp. 694–697. doi: 10.1038/159694a0.

[17] Murray Gell-Mann. “A Schematic Model of Baryons and Mesons”. In: *Phys. Lett.* 8 (1964), pp. 214–215. doi: 10.1016/S0031-9163(64)92001-3.

[18] Elliott D. Bloom et al. “High-Energy Inelastic e p Scattering at 6-Degrees and 10-Degrees”. In: *Phys. Rev. Lett.* 23 (1969), pp. 930–934. doi: 10.1103/PhysRevLett.23.930.

[19] Bruno R. Stella and Hans-Jurgen Meyer. “Y(9.46 GeV) and the gluon discovery (a critical recollection of PLUTO results)”. In: *Eur. Phys. J. H* 36 (2011), pp. 203–243. doi: 10.1140/epjh/e2011-10029-3. arXiv: 1008.1869 [hep-ex].

[20] P. Söding. “On the discovery of the gluon”. In: *The European physical journal / H* 35 (2010), pp. 205–230. issn: 2102-6459. doi: 10.1140/epjh/e2010-00002-5. url: https://bib-pubdb1.desy.de/record/92607.

[21] The SLAC-SP-017 Collaboration. “Discovery of a Narrow Resonance in e^+e^- Annihilation”. In: *Phys. Rev. Lett.* 33 (1974), pp. 1406–1408. doi: 10.1103/PhysRevLett.33.1406.

[22] The E598 Collaboration. “Experimental Observation of a Heavy Particle J”. In: *Phys. Rev. Lett.* 33 (1974), pp. 1404–1406. doi: 10.1103/PhysRevLett.33.1404.

[23] Makoto Kobayashi and Toshihide Maskawa. “CP Violation in the Renormalizable Theory of Weak Interaction”. In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. doi: 10.1143/PTP.49.652.

[24] Martin L. Perl et al. “Evidence for Anomalous Lepton Production in e+ - e- Annihilation”. In: *Phys. Rev. Lett.* 35 (1975), pp. 1489–1492. doi: 10.1103/PhysRevLett.35.1489.

[25] S. W. Herb et al. “Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions”. In: *Phys. Rev. Lett.* 39 (1977), pp. 252–255. doi: 10.1103/PhysRevLett.39.252.

[26] The D0 Collaboration. “Observation of the top quark”. In: *Phys. Rev. Lett.* 74 (1995), pp. 2632–2637. doi: 10.1103/PhysRevLett.74.2632. arXiv: hep-ex/9503003.

[27] The CDF Collaboration. “Observation of top quark production in $\bar{p}p$ collisions”. In: *Phys. Rev. Lett.* 74 (1995), pp. 2626–2631. doi: 10.1103/PhysRevLett.74.2626. arXiv: hep-ex/9503002.

[28] W. Pauli. “Dear radioactive ladies and gentlemen”. In: *Phys. Today* 31N9 (1978), p. 27.

[29] Frederick Reines and Clyde L. Cowan. “The neutrino”. In: *Nature* 178 (1956), pp. 446–449. doi: 10.1038/178446a0.

[30] G. Danby et al. “Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos”. In: *Phys. Rev. Lett.* 9 (1962), pp. 36–44. doi: 10.1103/PhysRevLett.9.36.

[31] The DONUT Collaboration. “Observation of tau neutrino interactions”. In: *Phys. Lett. B* 504 (2001), pp. 218–224. doi: 10.1016/S0370-2693(01)00307-0. arXiv: hep-ex/0012035.

[32] Samoil M. Bilenky and J. Hosek. “Glashow-Weinberg-Salam Theory of Electroweak Interactions and the Neutral Currents”. In: *Phys. Rept.* 90 (1982), pp. 73–157. doi: 10.1016/0370-1573(82)90016-3.

[33] The UA1 Collaboration. “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s}$ = 540 GeV”. In: *Phys. Lett. B* 122 (1983), pp. 103–116. doi: 10.1016/0370-2693(83)91177-2.

[34] The UA2 Collaboration. “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider”. In: *Phys. Lett. B* 122 (1983), pp. 476–485. doi: 10.1016/0370-2693(83)91605-2.

[35] The UA1 Collaboration. “Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c^2 at the CERN SPS Collider”. In: *Phys. Lett. B* 126 (1983), pp. 398–410. doi: 10.1016/0370-2693(83)90188-0.

[36] The UA2 Collaboration. “Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ Collider”. In: *Phys. Lett. B* 129 (1983), pp. 130–140. doi: 10.1016/0370-2693(83)90744-X.

[37] The ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 1–29. doi: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[38] The CMS Collaboration. “Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 30–61. doi: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[39] Emmy Noether. “Invariant Variation Problems”. In: *Gött. Nachr.* 1918 (1918), pp. 235–257. doi: 10.1080/00411457108231446. arXiv: physics/0503066.

[40] Michael E. Peskin and Daniel V. Schroeder. “An Introduction to quantum field theory”. Reading, USA: Addison-Wesley, 1995. isbn: 978-0-201-50397-5.

[41] Samoil M. Bilenky and J. Hosek. “Glashow-Weinberg-Salam Theory of Electroweak Interactions and the Neutral Currents”. In: *Phys. Rept.* 90 (1982), pp. 73–157. doi: 10.1016/0370-1573(82)90016-3.

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

[43] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. “Global Conservation Laws and Massless Particles”. In: *Phys. Rev. Lett.* 13 (1964), pp. 585–587. doi: 10.1103/PhysRevLett.13.585.

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

[45] G. ’t Hooft and M. Veltman. “Regularization and renormalization of gauge fields”. In: *Nuclear Physics B* 44.1 (1972), pp. 189–213. issn: 0550-3213. doi: https://doi.org/10.1016/0550-3213(72)90279-9.

[46] Michał Szleper. “The Higgs boson and the physics of WW scattering before and after Higgs discovery”. Dec. 2014. arXiv: 1412.8367 [hep-ph].

[47] J. Horejsi. “Introduction to electroweak unification: Standard model from tree unitarity”. 1994. doi: https://doi.org/10.1142/2445.

[48] Anja Vest. “Vector Boson Scattering and Quartic Gauge Couplings at the LHC”. 2014. url: https://www.phy.bnl.gov/~partsem/fy14/Seminar_BNL_Vest.pdf.

[49] Benjamin W. Lee, C. Quigg, and H. B. Thacker. “The Strength of Weak Interactions at Very High-Energies and the Higgs Boson Mass”. In: *Phys. Rev. Lett.* 38 (1977), pp. 883–885. doi: 10.1103/PhysRevLett.38.883.

[50] Michele Gallinaro et al. “Beyond the Standard Model in Vector Boson Scattering Signatures”. In: *International Workshop on BSM models in Vector Boson Scattering processes*. May 2020. arXiv: 2005.09889 [hep-ph].

[51] The CMS Collaboration. “Measurement of vector boson scattering and constraints on anomalous quartic couplings from events with four leptons and two jets in proton–proton collisions at $\sqrt{s} = 13$ TeV”. In: *Phys. Lett. B* 774 (2017), pp. 682–705. doi: 10.1016/j.physletb.2017.10.020. arXiv: 1708.02812 [hep-ex].

[52] CERN. “LHC Design Report Vol.1: The LHC Main Ring”. June 2004. doi: 10.5170/CERN-2004-003-V-1.

[53] ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08003. doi: 10.1088/1748-0221/3/08/S08003.

[54] The CMS Collaboration. “The CMS Experiment at the CERN LHC”. In: *JINST* 3 (2008), S08004. doi: 10.1088/1748-0221/3/08/S08004.

[55] ALICE Collaboration. “The ALICE experiment at the CERN LHC”. In: *JINST* 3 (2008), S08002. doi: 10.1088/1748-0221/3/08/S08002.

[56] The LHCb Collaboration. “The LHCb experiment”. In: *Proceedings, 22nd Lake Louise Winter Institute: Fundamental Interactions (LLWI 2007)*. 2007, pp. 60–79. doi: 10.1142/9789812776105_0003.

[57] Esma Mobs. “The CERN accelerator complex. Complexe des accélérateurs du CERN”. General Photo. July 2016. url: http://cds.cern.ch/record/2197559.

[58] Joao Pequenao. “Computer generated image of the whole ATLAS detector”. 2008. url: http://www.atlas.ch/photos/full-detector-cgi.html.

[59] G. Apollinari et al. “High Luminosity Large Hadron Collider HL-LHC”. In: *CERN Yellow Rep.* 5 (2015). Ed. by G Apollinari et al., pp. 1–19. doi: 10.5170/CERN-2015-005.1. arXiv: 1705.08830 [physics.acc-ph].

[60] CERN. “HL-LHC Project Schedule”. url: https://project-hl-lhc-industry.web.cern.ch/content/project-schedule.

[61] ATLAS Collaboration. “Technical Design Report for the ATLAS Inner Tracker Pixel Detector”. Sept. 2017. url: https://cds.cern.ch/record/2285585.

[62] ATLAS Collaboration. “Technical Design Report for the ATLAS Inner Tracker Strip Detector”. Apr. 2017. url: https://cds.cern.ch/record/2257755.

[63] The ATLAS Collaboration. “Expected tracking and related performance with the updated ATLAS Inner Tracker layout at the High-Luminosity LHC”. 2021. url: https://cds.cern.ch/record/2776651.

[64] The ATLAS ITk Project. “Fluence distributions for ITk extended@4 layout”. url: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/ITK-2016-002/.

[65] Lingxin Meng. “ATLAS ITk Pixel Detector Overview”. In: *International Workshop on Future Linear Colliders*. May 2021. arXiv: 2105.10367 [physics.ins-det].

[66] Frank Hartmann. “Evolution of Silicon Sensor Technology in Particle Physics”. In: *Springer Tracts Mod. Phys.* 275 (2017), pp.1–372. doi: 10.1007/978-3-319-64436-3.

[67] Simon Sze. “Physics of Semiconductor Devices”. John Wiley & Sons, 2006.

[68] The ATLAS ITk Project. “Price Enquiry - Technical Specification - Production of Planar Pixel Sensors for the ATLAS ITk”. url: https://edms.cern.ch/ui/file/2040673/2/DO-31790-EP_Technical_Specification_V5.pdf.

[69] The ATLAS ITk Project. “ITk Planar Pixel Sensor MS Public Plots”. url: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/ITK-2020-001/.

[70] Maurice Garcia-Sciveres. “The RD53A Integrated Circuit”. Oct. 2017. url: https://cds.cern.ch/record/2287593.

[71] R. Diener et al. “The DESY II Test Beam Facility”. In: *Nucl. Instrum. Meth. A* 922 (2019), pp. 265–286. doi: 10.1016/j.nima.2018.11.133. arXiv: 1807.09328 [physics.ins-det].

[72] A Himmi et al. “A Zero Suppression Micro-Circuit for Binary Readout CMOS Monolithic Sensors”. 2009. doi: 10.5170/CERN-2009-006.426. url: http://cds.cern.ch/record/1235850.

[73] Marlon Barbero et al. “FE-I4 ATLAS pixel chip design”. In: *PoS* VERTEX2009 (2009), p. 027. doi: 10.22323/1.095.0027.

[74] J. Baudot et al. “First test results Of MIMOSA-26, a fast CMOS sensor with integrated zero suppression and digitized output”. In: *Proceedings, 2009 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC 2009)*. 2009, pp. 1169–1173. doi: 10.1109/NSSMIC.2009.5402399.

[75] Claus Kleinwort. “General Broken Lines as advanced track fitting method”. In: *Nucl. Instrum. Meth. A* 673 (2012), pp. 107–110. doi: 10.1016/j.nima.2012.01.024. arXiv: 1201.4320 [physics.ins-det].

[76] Hanno Perrey. “EUDAQ and EUTelescope: Software Frameworks for Test Beam Data Acquisition and Analysis”. In: *PoS* TIPP2014 (2014), p. 353. doi: 10.22323/1.213.0353.

[77] ATLAS Collaboration. “Performance of W/Z taggers using UFO jets in ATLAS”. July 2021. url: http://cds.cern.ch/record/2777009.

[78] Ikaros I.Y. Bigi et al. “Production and Decay Properties of Ultra-heavy Quarks”. In: *Phys.Lett.* B181 (1986), p. 157. doi: 10.1016/0370-2693(86)91275-X.

[79] Torbjörn Sjöstrand et al. “An Introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. doi: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[80] M. Bahr et al. “Herwig++ Physics and Manual”. In: *Eur. Phys. J.* C58 (2008), pp. 639–707. doi: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[81] Gavin P. Salam. “Towards Jetography”. In: *Eur. Phys. J.* C67 (2010), pp. 637–686. doi: 10.1140/epjc/s10052-010-1314-6.

[82] Yuri L. Dokshitzer et al. “Better jet clustering algorithms”. In: *JHEP* 08 (1997), p. 001. doi: 10.1088/1126-6708/1997/08/001. arXiv: hep-ph/9707323 [hep-ph].

[83] Stephen D. Ellis and Davison E. Soper. “Successive combination jet algorithm for hadron collisions”. In: *Phys. Rev.* D48 (1993), pp. 3160–3166. doi: 10.1103/PhysRevD.48.3160. arXiv: hep-ph/9305266 [hep-ph].

[84] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. doi: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[85] The ATLAS Collaboration. "Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector". In: *JHEP* 09 (2013), p. 076. doi: 10.1007/JHEP09(2013)076. arXiv: 1306.4945 [hep-ex].

[86] S. Agostinelli et al. "GEANT4–a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. doi: 10.1016/S0168-9002(03)01368-8.

[87] Guido Altarelli, B. Mele, and M. Ruiz-Altaba. "Searching for New Heavy Vector Bosons in $p\bar{p}$ Colliders". In: *Z. Phys. C* 45 (1989). [Erratum: Z.Phys.C 47, 676 (1990)], p. 109. doi: 10.1007/BF01556677.

[88] Ian Moult, Lina Necib, and Jesse Thaler. "New Angles on Energy Correlation Functions". In: *JHEP* 12 (2016), p. 153. doi: 10.1007/JHEP12(2016)153. arXiv: 1609.07483 [hep-ph].

[89] Matteo Cacciari and Gavin P. Salam. "Pileup subtraction using jet areas". In: *Physics Letters B* 659.1 (2008), pp. 119–126. issn: 0370-2693. doi: https://doi.org/10.1016/j.physletb.2007.09.077. url: https://www.sciencedirect.com/science/article/pii/S0370269307011094.

[90] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The catchment area of jets". In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 005–005. doi: 10.1088/1126-6708/2008/04/005. url: https://doi.org/10.1088/1126-6708/2008/04/005.

[91] The ATLAS Collaboration. "Light-quark and gluon jet discrimination in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 74.8 (2014), p. 3023. doi: 10.1140/epjc/s10052-014-3023-z. arXiv: 1405.6583 [hep-ex].

[92] The ATLAS Collaboration. "Boosted hadronic vector boson and top quark tagging with ATLAS using Run 2 data". July 2020. url: https://cds.cern.ch/record/2724149.

[93] The ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J. C* 77 (2017), p. 490. doi: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

[94] The ATLAS Collaboration. "Optimisation of large-radius jet reconstruction for the ATLAS detector in 13 TeV proton–proton collisions". In: *Eur. Phys. J. C* 81.4 (2021), p. 334. doi: 10.1140/epjc/s10052-021-09054-3. arXiv: 2009.04986 [hep-ex].

[95] The ATLAS Collaboration. "Performance of mass-decorrelated jet substructure observables for hadronic two-body decay tagging in ATLAS". In: ATL-PHYS-PUB-2018-014 (July 2018). url: https://cds.cern.ch/record/2630973.

[96] Jesse Thaler and Ken Van Tilburg. "Identifying Boosted Objects with N-subjettiness". In: *JHEP* 03 (2011), p. 015. doi: 10.1007/JHEP03(2011)015. arXiv: 1011.2268 [hep-ph].

[97] Geoffrey C. Fox and Stephen Wolfram. "Observables for the Analysis of Event Shapes in e+ e- Annihilation and Other Processes". In: *Phys. Rev. Lett.* 41 (1978), p. 1581. doi: 10.1103/PhysRevLett.41.1581.

[98] Leandro G. Almeida et al. "Top Jets at the LHC". In: *Phys. Rev. D* 79 (2009), p. 074012. doi: 10.1103/PhysRevD.79.074012. arXiv: 0810.0934 [hep-ph].

[99] The ATLAS Collaboration. "ATLAS Measurements of the Properties of Jets for Boosted Particle Searches". In: *Phys. Rev. D* 86 (2012), p. 072006. doi: 10.1103/PhysRevD.86.072006. arXiv: 1206.5369 [hep-ex].

[100] Chunhui Chen. "New approach to identifying boosted hadronically-decaying particle using jet substructure in its center-of-mass frame". In: *Phys. Rev. D* 85 (2012), p. 034007. doi: 10.1103/PhysRevD.85.034007. arXiv: 1112.2567 [hep-ph].

[101] Jesse Thaler and Lian-Tao Wang. "Strategies to Identify Boosted Tops". In: *JHEP* 07 (2008), p. 092. doi: 10.1088/1126-6708/2008/07/092. arXiv: 0806.0023 [hep-ph].

[102] The ATLAS Collaboration. "Measurement of kT splitting scales in W->lv events at sqrt(s)=7 TeV with the ATLAS detector". In: *Eur. Phys. J. C* 73.5 (2013), p. 2432. doi: 10.1140/epjc/s10052-013-2432-8. arXiv: 1302.1415 [hep-ex].

[103] S. Catani et al. "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions". In: *Nucl. Phys. B* 406 (1993), pp. 187–224. doi: 10.1016/0550-3213(93)90166-M.

[104] Frank Nielsen. "On the Jensen–Shannon Symmetrization of Distances Relying on Abstract Means". In: *Entropy* 21.5 (2019). issn: 1099-4300. url: https://www.mdpi.com/1099-4300/21/5/485.

[105] Georgios Choudalakis. “On hypothesis testing, trials factor, hypertests and the BumpHunter”. In: *PHYSTAT 2011*. Jan. 2011. arXiv: 1101.0390 [physics.data-an].

[106] The ATLAS Collaboration. “Improving jet substructure performance in ATLAS using Track-CaloClusters”. Apr. 2017. url: https://cds.cern.ch/record/2275636.

[107] Andrew J. Larkoski et al. “Soft Drop”. In: *JHEP* 05 (2014), p. 146. doi: 10.1007/JHEP05(2014)146. arXiv: 1402.2657 [hep-ph].

[108] The ATLAS Collaboration. “In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector”. In: *Eur. Phys. J. C* 79.2 (2019), p. 135. doi: 10.1140/epjc/s10052-019-6632-8. arXiv: 1807.09477 [hep-ex].

[109] The ATLAS Collaboration. “Performance of mass-decorrelated jet substructure observables for hadronic two-body decay tagging in ATLAS”. July 2018. url: https://cds.cern.ch/record/2630973.

[110] V. Barger et al. “Strong W^+W^+ scattering signals at pp supercolliders”. In: *Phys. Rev. D* 42 (9 Nov. 1990), pp. 3052–3077. doi: 10.1103/PhysRevD.42.3052. url: https://link.aps.org/doi/10.1103/PhysRevD.42.3052.

[111] Torbjorn Sjostrand. “PYTHIA 8 Status Report”. In: *HERA and the LHC: 4th Workshop on the Implications of HERA for LHC Physics*. Sept. 2008, pp. 726–732. doi: 10.3204/DESY-PROC-2009-02/41. arXiv: 0809.0303 [hep-ph].

[112] The ATLAS Collaboration. “The ATLAS Simulation Infrastructure”. In: *Eur. Phys. J. C* 70 (2010), pp. 823–874. doi: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [physics.ins-det].

[113] The ATLAS Collaboration. “Luminosity Public Results Run 2”. In: (). url: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[114] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079. doi: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[115] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. “A Brief Introduction to PYTHIA 8.1”. In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. doi: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[116] Richard D. Ball et al. “Parton distributions with LHC data”. In: *Nucl. Phys. B* 867 (2013), pp. 244–289. doi: 10.1016/j.nuclphysb.2012.10.003. arXiv: 1207.1303 [hep-ph].

[117] T. Gleisberg et al. “Event generation with SHERPA 1.1”. In: *JHEP* 02 (2009), p. 007. doi: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[118] Tanju Gleisberg and Stefan Hoeche. “Comix, a new matrix element generator”. In: *JHEP* 12 (2008), p. 039. doi: 10.1088/1126-6708/2008/12/039. arXiv: 0808.3674 [hep-ph].

[119] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini. “Scattering Amplitudes with Open Loops”. In: *Phys. Rev. Lett.* 108 (2012), p. 111601. doi: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph].

[120] Steffen Schumann and Frank Krauss. “A Parton shower algorithm based on Catani-Seymour dipole factorisation”. In: *JHEP* 03 (2008), p. 038. doi: 10.1088/1126-6708/2008/03/038. arXiv: 0709.1027 [hep-ph].

[121] Stefan Hoeche et al. “QCD matrix elements + parton showers: The NLO case”. In: *JHEP* 04 (2013), p. 027. doi: 10.1007/JHEP04(2013)027. arXiv: 1207.5030 [hep-ph].

[122] Richard D. Ball et al. “Parton distributions for the LHC Run II”. In: *JHEP* 04 (2015), p. 040. doi: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[123] Simone Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *JHEP* 06 (2010), p. 043. doi: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[124] Pierre Artoisenet et al. “Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations”. In: *JHEP* 03 (2013), p. 015. doi: 10.1007/JHEP03(2013)015. arXiv: 1212.3460 [hep-ph].

[125] The ATLAS Collaboration. “ATLAS Pythia 8 tunes to 7 TeV data”. Nov. 2014. eprint: ATL-PHYS-PUB-2014-021. url: https://cds.cern.ch/record/1966419.

[126] Michal Czakon and Alexander Mitov. “Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders”. In: *Comput. Phys. Commun.* 185 (2014), p. 2930. doi: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph].

[127] Nikolaos Kidonakis. “Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production”. In: *Phys. Rev. D* 83 (2011), p. 091503. doi: 10.1103/PhysRevD.83.091503. arXiv: 1103.2792 [hep-ph].

[128] Nikolaos Kidonakis. “NNLL resummation for s-channel single top quark production”. In: *Phys. Rev. D* 81 (2010), p. 054028. doi: 10.1103/PhysRevD.81.054028. arXiv: 1001.5034 [hep-ph].

[129] Nikolaos Kidonakis. “Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-”. In: *Phys. Rev. D* 82 (2010), p. 054018. doi: 10.1103/PhysRevD.82.054018. arXiv: 1005.4451 [hep-ph].

[130] The ATLAS Collaboration. “Studies on top-quark Monte Carlo modelling for Top2016”. Oct. 2016. eprint: ATL-PHYS-PUB-2016-020. url: http://cdsweb.cern.ch/record/2216168.

[131] The ATLAS Collaboration. “Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC”. June 2019. eprint: ATLAS-CONF-2019-021. url: http://cdsweb.cern.ch/record/2677054.

[132] The ATLAS Collaboration. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In: *Eur. Phys. J. C* 77.7 (2017), p. 466. doi: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485 [hep-ex].

[133] The ATLAS Collaboration. “Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment”. Mar. 2017. url: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-003/.

[134] The ATLAS Collaboration. “Tagging and suppression of pileup jets”. May 2014. url: https://cds.cern.ch/record/1700870.

[135] The ATLAS Collaboration. “Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector”. In: *Eur. Phys. J. C* 77.9 (2017). [Erratum: Eur.Phys.J.C 77, 712 (2017)], p. 580. doi: 10.1140/epjc/s10052-017-5081-5. arXiv: 1705.02211 [hep-ex].

[136] David Krohn, Jesse Thaler, and Lian-Tao Wang. “Jet Trimming”. In: *JHEP* 02 (2010), p. 084. doi: 10.1007/JHEP02(2010)084. arXiv: 0912.1342 [hep-ph].

[137] The ATLAS Collaboration. “Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data”. June 2016. url: https://cds.cern.ch/record/2157687.

[138] The ATLAS Collaboration. "Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run". July 2017. url: https://cds.cern.ch/record/2273281.

[139] Particle Data Group. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), p. 083C01. doi: 10.1093/ptep/ptaa104.

[140] The ATLAS Collaboration. "Differential cross-section measurements for the electroweak production of dijets in association with a Z boson in proton–proton collisions at ATLAS". In: *Eur. Phys. J. C* 81.2 (2021), p. 163. doi: 10.1140/epjc/s10052-020-08734-w. arXiv: 2006.15458 [hep-ex].

[141] Francois Chollet et al. "Keras". 2015. url: https://github.com/fchollet/keras.

[142] Martín Abadi et al. "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems". Software available from tensorflow.org. 2015. url: https://www.tensorflow.org/.

[143] Daniel Hay Guest et al. *lwtnn: Version 2.12.1*. July 2021. doi: 10.5281/zenodo.5082190.

[144] Diederik P. Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization". 2017. arXiv: 1412.6980 [cs.LG].

[145] William Buttinger. "Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data". May 2015. url: https://cds.cern.ch/record/2014726.

[146] The ATLAS Collaboration. "Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector". In: *JHEP* 01 (2015), p. 069. doi: 10.1007/JHEP01(2015)069. arXiv: 1409.6212 [hep-ex].

[147] The ATLAS Collaboration. "Observation of electroweak production of a same-sign W boson pair in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. Lett.* 123.16 (2019), p. 161801. doi: 10.1103/PhysRevLett.123.161801. arXiv: 1906.03203 [hep-ex].

[148] The ATLAS Collaboration. "Search for the electroweak diboson production in association with a high-mass dijet system in semileptonic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. D* 100.3 (2019), p. 032007. doi: 10.1103/PhysRevD.100.032007. arXiv: 1905.07714 [hep-ex].

[149] The ATLAS Collaboration. "Observation of electroweak $W^{\pm}Z$ boson pair production in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett. B* 793 (2019), pp. 469–492. doi: 10.1016/j.physletb.2019.05.012. arXiv: 1812.09740 [hep-ex].

[150] The ATLAS Collaboration. "Observation of electroweak production of two jets and a Z-boson pair with the ATLAS detector at the LHC". Apr. 2020. arXiv: 2004.10612 [hep-ex].

[151] The CMS Collaboration. "Observation of electroweak production of same-sign W boson pairs in the two jet and two same-sign lepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. Lett.* 120.8 (2018), p. 081801. doi: 10.1103/PhysRevLett.120.081801. arXiv: 1709.05822 [hep-ex].

[152] The CMS Collaboration. "Measurement of electroweak WZ boson production and search for new physics in WZ + two jets events in pp collisions at $\sqrt{s} = 13$TeV". In: *Phys. Lett. B* 795 (2019), pp. 281–307. doi: 10.1016/j.physletb.2019.05.042. arXiv: 1901.04060 [hep-ex].

[153] The CMS Collaboration. "Evidence for WW/WZ vector boson scattering in the decay channel $\ell\nu$qq produced in association with two jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV". Dec. 2021. arXiv: 2112.05259 [hep-ex].

[154] The ATLAS Collaboration. "Boosted JetTagging Recommendation Full Run2". url: twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/BoostedJetTaggingRecommendationFullRun2.