
**Searching for diboson resonances in the
all-hadronic final state
and
a Lorentz invariance based deep neural network for
W-tagging**

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Thea Klaeboe Arrestad
aus
Norwegen

Promotionskommission:
Prof. Dr. Benjamin Kilminster
Prof. Dr. Florencia Canelli
Prof. Dr. Jesse Thaler
Dr. Andreas Hinzmann

Zürich 2018

Abstract

Zusammenfassung

Contents

1	The Standard Model and Beyond	2
1.1	The Standard Model Lagrangian	3
1.1.1	Particles and fields	3
1.1.2	Electroweak theory	4
1.1.3	The Higgs Mechanism	4
1.1.4	Quantum Chromodynamics	4
1.2	Beyond Standard Model Physics	4
1.2.1	The hierarchy problem and the gravitational force	4
1.2.2	Theories of New Physics	4
1.2.3	Heavy Vector Triplet formalism	4
2	Statistical Methods	5
2.1	Statistical methods: Confidence Limits	6
3	Experimental setup	7
3.1	The Large Hadron Collider	8
3.2	The CMS detector	10
3.2.1	Coordinate system	12
3.2.2	Tracking detectors	12
3.2.3	Electromagnetic calorimeter	13
3.2.4	Hadron calorimeter	15
3.2.5	Muon chambers	18
3.3	Trigger system: From collision to disk	18
4	Event reconstruction	20
4.1	Track and primary vertex reconstruction	21
4.2	The Particle Flow Algorithm	22
4.2.1	Reconstruction of the Particle Flow inputs	22
4.2.2	Particle Flow identification	24
4.3	Pile-up removal	27
4.3.1	Charged Hadron Subtraction	27
4.3.2	Pile up per particle identification (PUPPI)	27
4.4	Jet reconstruction	29
4.4.1	Jet clustering	29
4.4.2	PF jets in CMS	31
4.4.3	Jet energy corrections	32
4.5	Jet substructure reconstruction	33
4.5.1	Grooming	34
4.5.2	N-subjettiness	36

4.5.3	Vector boson tagging	37
4.6	Monte Carlo Event Generators	41
5	Diboson resonance searches in CMS	42
5.1	Search I: First search for diboson resonances at 13 TeV	43
5.1.1	A small bump	44
5.1.2	Analysis strategy	45
5.1.3	Data and simulated samples	47
5.1.4	Event selection	47
5.1.5	Background modeling	60
5.1.6	Signal modeling	63
5.1.7	V-tagging scale factors	64
5.1.8	Systematic uncertainties	72
5.1.9	Results	74
5.1.10	Limits: All-hadronic analysis	74
5.1.11	Limits: Semi-leptonic and all-hadronic combination	76
5.2	Search II: A new pileup resistant, perturbative safe tagger	80
5.2.1	Towards robust boosted jet tagging	81
5.2.2	Analysis strategy	82
5.2.3	Data and simulated samples	83
5.2.4	Event selection	83
5.2.5	Developing a new W-tagger	86
5.2.6	W-tagging mistagging rate measurement	94
5.2.7	Mass and purity categorization	97
5.2.8	Background modeling: F-test	97
5.2.9	Signal modeling	99
5.2.10	Systematic uncertainties	99
5.2.11	Results	99
5.3	Search III: A novel framework for multi-dimensional searches	103
5.3.1	Small bumps and tri-bosons	104
5.3.2	Analysis strategy	105
5.3.3	Data and simulated samples	106
5.3.4	Event selection	106
5.3.5	Triggering	107
5.3.6	A mass and p_T decorrelated tagger	110
5.3.7	The multidimensional fit	115
5.3.8	Modeling of the signal	115
5.3.9	Modeling of the non-resonant background	117
5.3.10	Modeling of the resonant background	122
5.3.11	Systematic uncertainties	126
5.3.12	Results	128
5.4	Summary and outlook	132
6	A Lorentz Invariance Based Deep Neural Network for W-tagging	134
6.1	Infusing deep neural networks with physics	135
6.2	LoLa	136
6.2.1	Architecture	136
6.2.2	Input	137
6.2.3	The Combination Layer	138
6.2.4	The Lorentz Layer	139

6.3	Performance	141
6.4	p_T and mass dependence	141
6.4.1	p_T decorrelation	142
6.4.2	Mass sculpting	143
6.5	Validation on an independent sample	145
6.6	Summary and outlook	148
7	Summary	149
A	Search I: Limits per mass category	150
B	2016 W-tagging scale factor measurement	155
B.1	Efficiency scale factors for 12.9 and 35.9 fb^{-1}	155
C	2015 cross-check analysis	158
D	Background fit checks for 2016 analysis	159
D.0.1	Background fit checks in data sideband	160
D.0.2	Background fit checks in QCD MC	162
D.0.3	F-test in signal region	164
E	Search III: Multidimensional fit	170
E.1	Signal fits	170
E.2	2016 kernels	171
E.3	Resonant background	171
	Bibliography	173

Introduction

CHAPTER 1

The Standard Model and Beyond

1.1 The Standard Model Lagrangian

Everything this thesis is built on, has its roots in the Standard Model. The Standard Model of particle physics (SM) addresses the question *What is matter made of?* on the smallest possible scale. It maps the fundamental constituents of the universe together through the forces that bind them, hoping to provide a complete picture of the laws of nature.

1.1.1 Particles and fields

It appears that all matter in the universe can be described by a very small collection of fundamental particles, six leptons and six quarks. These are collectively called fermions and are, as far as we can tell, truly elementary (not composed of any other particles). Leptons are particles with integer or zero electric charge (defined in units of electron charge). They come in three flavours, or generations, and their mass increases with generation. Each generation of leptons consists of two particles: one charged lepton and one neutrally charged particle denoted *neutrino* (ν). The three generations can be arranged in a doublet structure, and are as follows

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix} \quad \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \quad \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \quad (1.1)$$

The leptons come in two states; positively charged and negatively charged, where charged is defined in unit of electron charge e . The base state is negatively charged, e^- , μ^- , and τ^- , whereas the positively charged leptons, e^+ , μ^+ , and τ^+ are considered their anti-particles states. Each lepton generation is assigned its own quantum number L which must be conserved after any process. A summary of the lepton properties is listed in Table 1.1. Leptons interact

Lepton	Mass	Charge	L_e	L_μ	L_τ
e^-	0.5 MeV	e	1	0	0
μ^-	106 MeV	e	0	1	0
τ^-	1777 MeV	e	0	0	1
ν_e	< 3 eV	0	1	0	0
ν_μ	< 0.19 MeV	0	0	1	0
ν_τ	< 18.2 MeV	0	0	0	1

Table 1.1: Lepton Properties

with one another through the *electroweak force*, which will be explained in more detailed in Section 1.1.2.

The other six fundamental matter particles are the *quarks*. They are distinguished from the leptons in that they interact with one another through the strong force, described in Section 1.1.4. This force binds the quarks together to form baryons (like protons and neutrons) or mesons (like pions), and in addition keeps the quarks from being observed as free particles (they are only visible through their baryon/meson bound states) Also organized in three generations, the six quarks are called *up*, *down*, *charm*, *strange*, *top* and *bottom*, and are organized in flavour doublets as follow

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix} \quad (1.2)$$

Each quark comes with a fractional charge of $-\frac{1}{3}$ (u, c and t) and $\frac{2}{3}$ (d, s and b) of one electron charge. As with the leptons, they also come with distinct particles of opposite charge,

Quark	Mass	Charge	Properties
u	1 – 5 MeV	$\frac{2}{3}e$	$I_z = \frac{1}{2}$
d	3 – 9 MeV	$-\frac{1}{3}e$	$I_z = -\frac{1}{2}$
c	1.15 – 1.35 GeV	$\frac{2}{3}e$	Charm = +1
s	75 – 170 MeV	$-\frac{1}{3}e$	Strangeness = -1
t	≈ 174 GeV	$\frac{2}{3}e$	Top = +1
b	4.0 – 4.4 GeV	$-\frac{1}{3}e$	Bottom = -1

Table 1.2: Quark Properties

anti-quarks. As mentioned above, quarks can interact with one another through the strong force. However, they also interact through the weak and electro-magnetic forces. Some of the quark properties are listed in Table 1.2.

1.1.2 Electroweak theory

1.1.3 The Higgs Mechanism

1.1.4 Quantum Chromodynamics

1.2 Beyond Standard Model Physics

1.2.1 The hierarchy problem and the gravitational force

1.2.2 Theories of New Physics

Warped extra dimensions

Compositeness

1.2.3 Heavy Vector Triplet formalism

CHAPTER 2

Statistical Methods

2.1 Statistical methods: Confidence Limits

CHAPTER 3

Experimental setup

3.1 The Large Hadron Collider

In March 1984, the European Organization for Nuclear Research (CERN) and the European Committee for Future Accelerators (ECFA) held a workshop in Lausanne entitled ‘Large Hadron Collider in the LEP Tunnel’. This is history’s first written mention of the Large Hadron Collider (LHC) and the topic under discussion was exactly how and where to build a new type of high-energy collider, capable of bringing hadrons to collide rather than leptons. The LHC would be housed in a tunnel which, at the time, was under excavation to host the Large Electron-Positron Collider (LEP) designed to collide leptons with center-of-mass energies up to around 200 GeV. LEP was a circular collider with a circumference of 27 km and the tunnel hosting it was located roughly 100 meters underground on the border between France and Switzerland, at the outskirts of Geneva. The justification for building a machine like the LHC, was that once LEP got to maximum reach, a new and more powerful collider would be needed in its place in order to probe higher energies. While collisions of electrons with positrons provided exceptionally clean and precise measurements due to them being point particles, their lightness prevent them from being accelerated to higher energies. Collisions of hadrons, however, would allow for center-of-mass energies two orders of magnitude higher than that of LEP. Therefore, after running a while at two times the W mass (160 GeV) and reaching a maximum center-of-mass energy of 209 GeV, LEP was dismantled in 2000 in order to make room for the LHC.

The Large Hadron Collider started up in September 2008 and, while having the same 27-kilometer radius as the LEP collider, is capable of accelerating protons up to a center-of-mass energy of around 14 TeV, 70 times that of LEP. The accelerator consists of two oppositely going proton beams, isolated from each other and under ultrahigh vacuum, which are accelerated up to speeds close to that of the speed of light through radio frequency (RF) cavities, before being brought to collide at four different interaction points along the ring. These four collision points correspond to the location of the four LHC particle detectors; ATLAS, CMS, LHCb and ALICE. While ATLAS and CMS are general-purpose detectors built in order to study a large range of different physics processes, LHCb and ALICE are built for dedicated purposes; LHCb for b-physics processes and ALICE for heavy ion collision. A proton’s journey from gas to one of the LHC collision points is as follows: First, hydrogen nuclei are extracted from a small tank of compressed hydrogen gas and stripped of their electrons. The remaining protons are then injected into the LINAC2, a linear accelerator responsible for increasing the proton energy to about 50 MeV through RF cavities that push charged particles forward by switching from positive to negative electric fields. LINAC2 additionally divides the constant stream of particles into equally spaced ‘bunches’ by careful tuning of the frequency of the field switch. The accelerated protons are then injected into the Proton Synchrotron Booster (PSB), where their energy is increased thirty folds more, to an energy of roughly 1.4 GeV. The two final acceleration stages before the protons reach the LHC ring are the Proton Synchrotron and Super Proton Synchrotron, eventually leaving the protons with a total energy of 450 GeV. The protons are now ready for the final stage of their travel and are injected into the two beam pipes of the LHC in oppositely going direction. They are injected in trains of 144 bunches each (with an order of 10^{11} protons per bunch), where each bunch is roughly 7.5 meters apart (or 25 ns). There are some larger beam gaps present in each beam in order to give the beam dump and injection kickers sufficient time to reach full voltage, where the largest one, the beam abort gap, is roughly 3 ms or 900 m long. The ring is filled with proton bunches until these are equally distributed throughout the two rings, a process taking roughly 4 minutes. This is called a ‘fill’. Here, the protons are accelerated to their maximum energy of 6.5 TeV, a process taking roughly 20 minutes, through eight

RF cavities. These RF cavities are also responsible of keeping the proton bunches tightly bunched, ensuring maximum luminosity at the four collision points. A complete sketch of the CERN accelerator complex is shown in Figure 3.1.

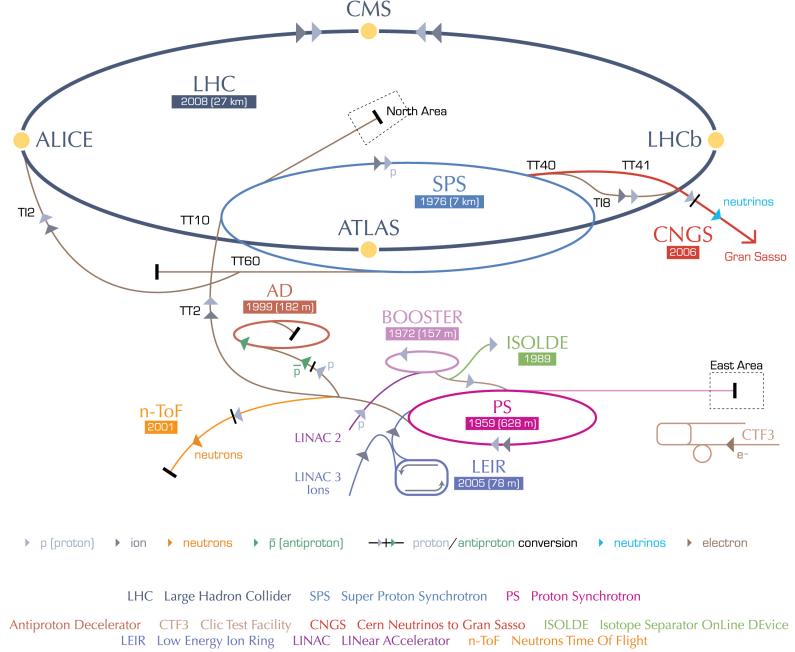


Figure 3.1: The Large Hadron Collider accelerator complex. The four collision points along the ring correspond to the location of the LHC particle detectors CMS, LHCb, ATLAS and ALICE [1].

After the beams have reached their maximum energy and are stably circulating in the LHC ring, they are brought to collide. The goal of such a collision, which occurs every 25 nano seconds, is that some of the protons will undergo an inelastic collision, allowing the quark/gluon constituents of each proton to interact with one another and produce new and interesting particles. The number of times such an interaction will take place inside a detector per area and time is quantified through the luminosity, \mathcal{L} , which is the proportionality factor between the number of observable events per second, and the cross section σ of the process you are interested in

$$\frac{dN_{\text{events}}}{dt} = \mathcal{L}\sigma. \quad (3.1)$$

The cross section is the probability that an event (like one which would produce new and interesting particles) will occur and is measured in barns, where 1 barn = 10^{-28} m^2 . This proportionality factor should therefore be as high as possible. It depends only on parameters of the detector and can, in the case of LHC, be defined through the following accelerator quantities

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta_*} F, \quad (3.2)$$

Parameter	Units	Nominal	2015	2016	2017
Energy	[TeV]	7.0	6.5	6.5	6.5
Bunch spacing	[ns]	25	25	25	25
Bunch intensity	$\times 10^{11}$ [protons/bunch]	1.15	1.15	1.15	1.2-1.45
Bunches per train		144	144	96	144
Total number of bunches		2808	2244	2220	2556
β^*	[cm]	55	80	40	27/25
Peak luminosity	$\times 10^{34}$ [cm $^{-2}$ s $^{-1}$]	1.0	0.5	1.4	2.0
Integrated luminosity			4.2	39.7	50.2

Table 3.1: Some key LHC detector parameters achieved during the first years of 13 TeV data taking

where N_b is the number of particles per bunch, n_b is the number of bunches, f_{rev} is their revolution frequency, γ_r is the relativistic gamma factor, ϵ_n is the transverse beam emittance (how confined the particles are in space and momentum), β^* is the beta function at the collision point (how narrow, or ‘squeezed’, the beam is) and F is a reduction factor to account for a constellation where the beams do not collide heads-on but at slight crossing angles. From this, it becomes clear that the main goal of the LHC is to; maximize the number of particles (N_b, n_b), their frequency (f_{rev}) and their energy (γ_r), while at the same time ensuring the protons are packed together as tightly as possible (lower ϵ_n and β^*). Using the nominal values of the LHC, the peak luminosity is roughly $\mathcal{L} \sim 10^{34}\text{cm}^{-2}\text{s}^{-1}$.

The peak luminosity of the LHC by the end of Run 2 in 2018 was grazing around $2.0 \cdot 10^{34}\text{cm}^{-2}\text{s}^{-1}$, corresponding to 2 times the nominal design luminosity.

To quantify the size and statistical power of a given LHC dataset, the integrated luminosity is used. This is the integral of the instantaneous luminosity over time and is defined as

$$\mathcal{L}_{int} = \int \mathcal{L} dt. \quad (3.3)$$

It is usually defined in units of inverse cross section, b $^{-1}$.

Despite the LHC starting up in 2008, there would be another year before data taking began. In March 2010, the LHC saw its first collision with a center-of-mass energy of 7 TeV, and continued running at this energy collecting around 5 inverse femtobarns of data by the end of 2011. In 2012, the energy was increased to 8 TeV and the LHC continued running until a planned long shutdown scheduled to begin in February 2013, collecting a total of $\sim 20\text{ fb}^{-1}$ and discovering the Higgs boson. This marked the end of Run 1 and the beginning of a two-year maintenance project intended to prepare the LHC for running at a center-of-mass energy of 13 TeV; Run 2.

Run 2, and where this thesis begins, started in June 2015. With the accelerator now running at 90% of its nominal energy, and with a peak luminosity between 1-2 times the design luminosity, the LHC managed to collect an impressive $\sim 160\text{ fb}^{-1}$ at this energy until its planned shutdown at the end of 2018. Some key LHC accelerator parameters that were in use for the datasets analyzed in this thesis, are quoted in Table 3.1

3.2 The CMS detector

The Compact Muon Solenoid (CMS) detector is true to its name; with a diameter of 15 meters and a weight of 14000 tons, it is 60 % smaller but two times heavier than its general purpose counterpart, the ATLAS detector. Its large weight is due to the CMS housing the world's largest and most powerful solenoid: A superconducting niobium titanium coil

circulating 18500 Amps and capable of generating a magnetic field of 3.8 Tesla. Together with its corresponding iron return yoke, responsible for reflecting the escaping magnetic flux, it accounts for 90% of the total detector weight. The CMS detector is cylindrically symmetric and organized in the following way: closest to the beam pipe and at a radius of about 3 cm, the inner tracking system begins. It consists of an inner silicon pixel detector and an outer silicon strip tracker, stretching out to a radius of roughly 1.2 meters. Following the tracker are two calorimeter layers: the electromagnetic calorimeter (ECAL) consisting of lead tungstate scintillating crystals and responsible for measuring the energy of electromagnetically interacting particles, followed by the hermetic hadronic calorimeter (HCAL) measuring the energy of hadrons. Contrary to ‘standard’ configurations for general purpose detectors, the CMS calorimeters are located inside the superconducting solenoid. This allows the detector to be rather compact, by reducing the necessary radius of the calorimeters, and additionally for the magnet to be strong enough (the magnetic field strength depends on the coil radius) to allow muon detectors to be located within the magnetic field so their momentum can be measured. The muon detectors are alternated with three layers of steel return yoke responsible for containing and reflecting the magnetic field and which only allows muons and weakly interacting particles to pass. A schematic overview of the CMS detector is shown in Figure 3.2. In the following, the different sub-detectors will be described in detail.

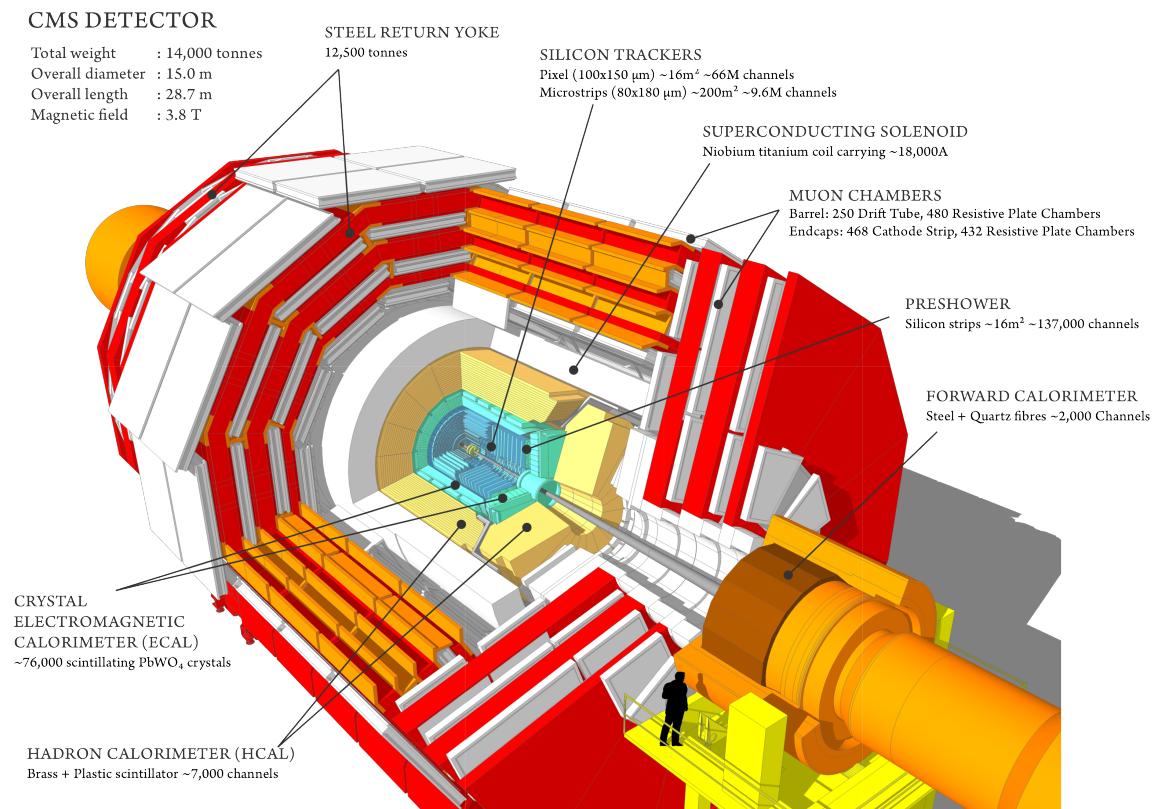


Figure 3.2: The CMS detector and its subsystems: The silicon tracker, electromagnetic and hadron calorimeters, the superconducting solenoid and the muon chambers inter-layered with the steel return yoke [2].

3.2.1 Coordinate system

To describe locations within the CMS detector, a Euclidian space coordinate system is used. Here, the positive z axis points along the beam pipe towards the west, the positive x axis points towards the center of the LHC ring, and the positive y axis upw towards the earths surface. Due to the cylindrical symmetry of the detector, polar coordinates are more convenient and most frequently encountered. In this scheme, the azimuthal angle ϕ is measured in the xy-plane, where $\phi = 0$ correspond to the positive x axis and $\phi = \pi/2$ correspond to the positive y axis. The polar angle θ is measured with respect to the z axis, $\theta = 0$ aligning with the positive and $\theta = \pi$ with the negative z axis. To define a particles angle with respect to the beam line, the pseudorapidity $\eta = -\ln \tan(\theta/2)$ is preferred over θ . This is due to the fact that particle production is approximately constant as a function of pseudorapidity and, more importantly, because differences in pseudorapidity are Lorentz invariant under boosts along the z-axis when assuming massless particles. To measure angular difference between particles in the detector, the variable $R = \sqrt{\eta^2 + \phi^2}$ is used, again Lorentz invariant under longitudinal boosts. A summary of the CMS coordinate system together with some example values are shown in Figure 3.3.

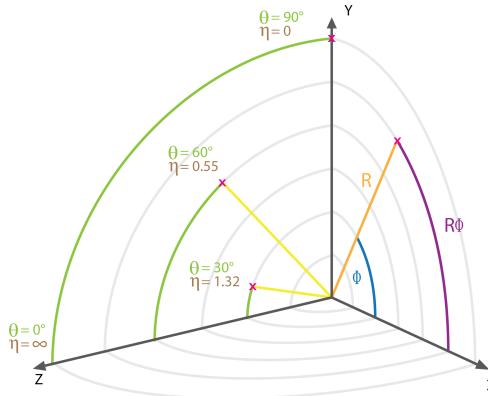


Figure 3.3: The CMS coordinate system [3]

3.2.2 Tracking detectors

The CMS tracker is responsible for accurately reconstructing the momentum of charged particles and consists of two sub-detectors. Closest to the interaction point, and where the particle intensity is the highest, the silicon pixel detector is located. Upgraded in 2017, from the so-called Phase-0 to the Phase-1 layout, it is structured in four cylindrical barrel layers at radii 2.9, 6.8, 10.9 and 16.0 cm (the barrel pixel) and three disks in each of the forward regions placed at a distance from the nominal interaction point of 29.1, 39.6 and 51.6 cm (the forward pixel). A sketch of the current Phase-I pixel detector compared to the Phase-0 detector is shown in Figure 3.4. The sensors located closest to the beam pipe are subject to hit intensities of $\mathcal{O}(\text{MHz/mm}^2)$ which puts strict constraints on the maximum sensor size in order to minimize occupancy in the detector. The pixel sensors are $100 \mu\text{m} \times 150 \mu\text{m}$ with a thickness of $285 \mu\text{m}$, and when counting both barrel and pixel sensors, sum up to a total of 79 million. The pixel sensors are mounted on detector modules with 16 read-out chips each, where the type of read-out chip depends on how close the module is to the beam pipe: the inner layer uses read-out chips with a rate capability of 600 MHz/cm^2 while for the outer layers, read-out chips with a rate capability of up to 200 MHz/cm^2 are sufficient.

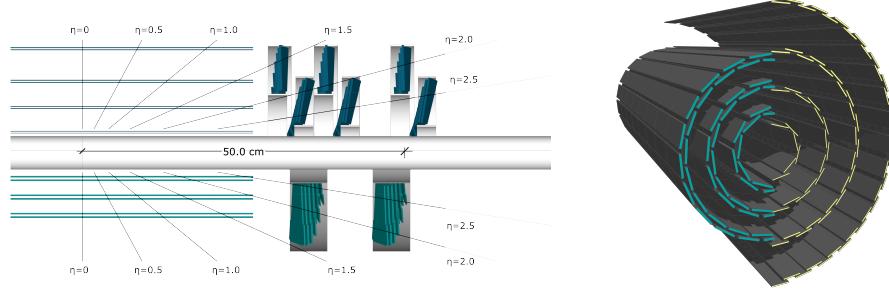


Figure 3.4: Left: The forward pixel detector layout before (bottom) and after (top) the Phase-I upgrade. Right: The barrel pixel detector before (left) and after (right) the Phase-I upgrade [4].

As the hit intensity reduces as you go further away from the beam pipe, the pixel sensors are replaced by silicon strip sensors, making up the second of the two tracker sub-systems, the silicon strip tracker. There are ten strip layers in total, stretching out to a radius of roughly 130 cm. These are divided into four sections: The inner barrel (TIB) with four strip layers, the two inner endcaps (TID) consisting of three disks each, the outer barrel (TOB) consisting of 6 cylindrical layers and the two endcaps (TEC) with 9 strip layers each. A schematic overview of the strip tracker layout is shown in Figure 3.5. The strips in the TIB and TID are 10 cm long, with a width of 80 μm and a thickness of 320 μm . The TOB and TEC sections consist of slightly larger strips of 25 cm x 180 μm and a thickness of 500 μm . The strip tracker has a total of 10 million detector strips and covers an area of $\sim 200 \text{ m}^2$. To prolong the silicon detector lifetime, the entire tracker (pixel and strip) is kept at a temperature of -20°C through a liquid cooling system. The tracker has a coverage up to $|\eta| < 2.6$ and a resolution of roughly $\sigma/p_T \approx 1.5 \times 10^{-5} p_T + 0.005$.

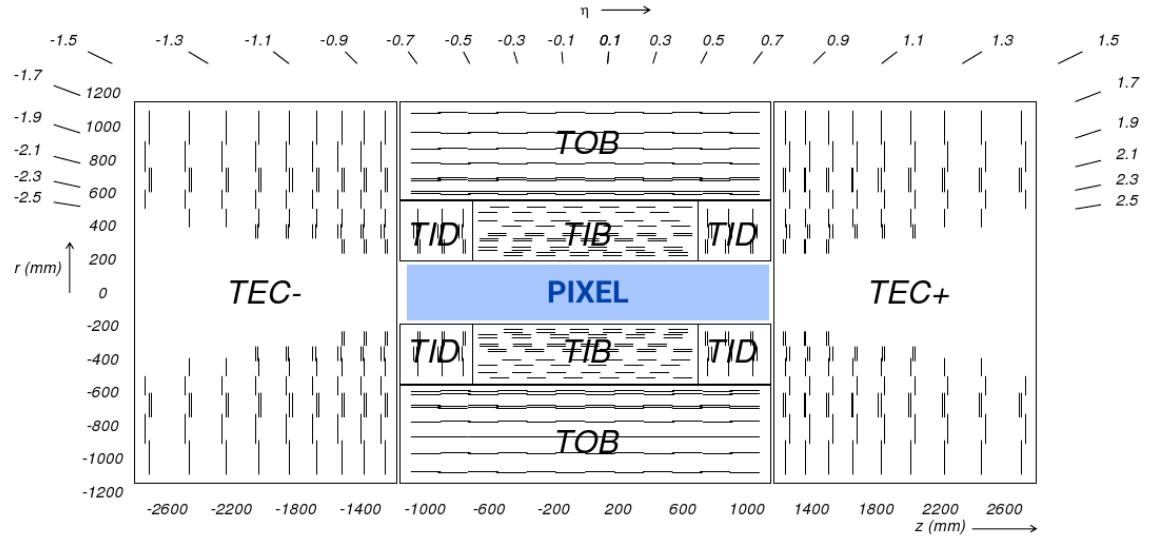


Figure 3.5: Schematic of the CMS silicon strip tracker and its four subsections: The inner barrel (TIB), inner endcaps (TID), the outer barrel (TOB) and the two endcaps (TEC) [5].

3.2.3 Electromagnetic calorimeter

Following the tracking detectors is the electromagnetic crystal calorimeter (ECAL). Consisting of 75 848 laterally segmented scintillating lead tungstate (PbWO_4) crystals, it was designed

to have the highest possible photon energy and position resolution in order to resolve a Higgs boson decaying into two photons, the cleanest of the Higgs discovery channels. With a goal of a photon/electron energy resolution of 0.5% above 100 GeV, the choice of detector material for the ECAL has been its most crucial design feature. In order to withstand the high doses of radiation and the high magnetic field present within the detector, while at the same time generating well-defined signal responses within the 25 nanoseconds between particle collisions, an extremely dense and transparent material capable of producing fast and clean photon bursts when hit, is required. The choice eventually fell on metal-heavy lead tungstate crystals, each taking roughly two days to artificially grow (and a total of about ten years to grow all of them). With a density of $\delta = 8.28 \text{ g/cm}^3$ (slightly higher than for stainless steel), the crystals are compact enough to yield excellent performance without taking up too much volume, allowing the ECAL to sit within the CMS superconducting solenoid. The homogeneous medium allows for a better energy resolution as it minimizes sampling fluctuation effects and it additionally contains enough oxygen in crystalline form to make it highly transparent to their entire scintillation emission spectrum. With an extremely short radiation length and small Moli  re radius ($X_0 = 0.85 \text{ cm}$,

$$R_M = 2.19$$

cm), the required homogeneity, granularity and compactness is obtained while at the same time emitting 80% of generated light within the 25 ns timeframe required. The largest drawbacks with a lead tungstate detector is the low light yield (100 γ per MeV), requiring dedicated avalanche photodiodes to increase the gain, as well as a light yield which strongly depends on the temperature. The detector response to an incident electron changes by $3.8 \pm 0.4 \%$ per degree Celsius which requires the ECAL temperature to be kept stable around $18.18.0(5)^\circ\text{C}$, obtained through an intricate water cooling system. The ECAL is completely hermetic and sorted into a barrel part (EB), covering pseudorapidities up to $|\eta| < 1.48$, and two endcap parts (EE) extending the total coverage to $|\eta| < 3.0$ in order to match the tracker coverage of $|\eta| < 2.5$. In order to improve the γ/π_0 separation power, a pre-shower detector (ES) using lead absorbers and silicon sensors covers the forward region between $1.65 < |\eta| < 2.6$. The crystals in the barrel are organized into supermodules, each consisting of about 1700 crystals, while the endcap is divided into half disks consisting of 3662 crystals each (so-called ‘Dees’). Each PbWO₄ crystal weighs around 1.5 kilogram and has a slightly tapered shape with a front face of $2.2 \times 2.2 \text{ cm}^2$ in the barrel and $2.86 \times 2.86 \text{ cm}^2$ in the endcaps. The crystals are 2.3 and 2.2 cm long in the barrel and endcaps, respectively. The total volume of the calorimeter including barrel and endcaps is 11 m^2 and weighs a total of 92 tons. The ECAL detector layout is illustrated in Figure 3.6.

Having no longitudinal segmentation, the ECAL relies on an accurate reconstruction of the event primary vertex, provided by the tracker, in order to reconstruct the photon angle correctly.

The obtained energy resolution of the ECAL can be parametrized in three parts: a stochastic, a noise and a constant term [6]. It is given as

$$\frac{\sigma E}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{0.128 \text{ GeV}}{E} \oplus 0.3\%$$

where the constant values were estimated in an electron test beam. The constant term of 0.3% is dominated by the non-uniformity in longitudinal light collection [7], and one of the main goals of the detector design was to get this term below 1%. The energy resolution as a function of electron energy is shown in Figure 3.7.

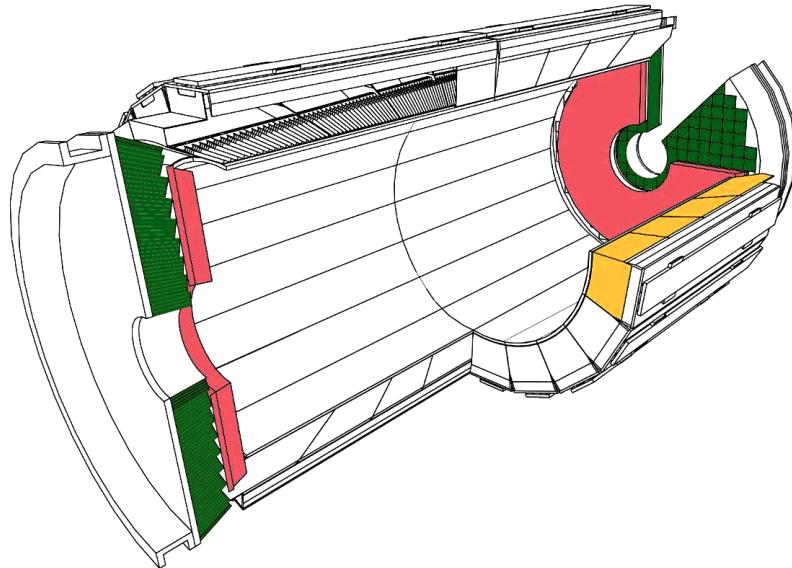


Figure 3.6: A schematic of the CMS electromagnetic calorimeter showing the barrel supermodules (yellow), the individual barrel crystals (black,top left), the endcap modules (green) and the pre-shower detectors (pink) [5].

3.2.4 Hadron calorimeter

Outside the crystal calorimeter is the hadron calorimeter (HCAL). It is the combined response of the ECAL and the HCAL that are responsible for measuring the energy of quarks, gluons and neutrinos through the reconstruction of particle jet energy and missing transverse energy. The HCAL is a sampling calorimeter, meaning it consists of alternating layers of dense brass absorber material and plastic scintillators. When a particle hits an absorber plate, it interacts with the absorber material and generates a shower of secondary particles which themselves generate new particle showers. These particles then generate light in the scintillating material which is proportional to their energy, and summing up the total amount of generated light over consecutive layers within a region, called a ‘tower’, is representative of the initial particles energy. The hadron calorimeter is split into four regions: the inner (HB) and outer (HO) barrel, the endcap (HE) and the forward region (HF). A schematic of the CMS HCAL is shown in Figure 3.8.

The inner barrel lies within the superconducting solenoid volume and covers the pseudorapidity range $|\eta| < 1.3$. It consists of 36 identical wedges, each of which weighing 26 tonnes, split into two half barrels (HB+ and HB-). A photograph of the wedges taken during installation is shown in Figure 3.9.

The wedges are made up of flat brass absorber plates oriented parallel to the beam axis. These plates consist of a 4 cm thick front steel plate followed by eight 5 cm thick brass plates, six 5.6 cm thick brass plates and ending with a 7.5 cm thick steel back plate. The absorber plates are then alternated by 4 mm thick plastic scintillator tiles, the detectors active medium, which are read out using wavelength-shifting plastic fibers. The effective thickness of the barrel hadron calorimeter in terms of interaction lengths increases with the polar angle θ , starting out at about $5.8 \lambda_I$ at an angle of 90 degrees, and increases to $10.6 \lambda_I$ at $|\eta| < 1.3$. As the energy resolution of the calorimeter depends on how much of the particles shower can be absorbed by the calorimeter, the quality of the energy measurement depends on its thickness. Due to the CMS design, the HB is confined to the volume between the ECAL

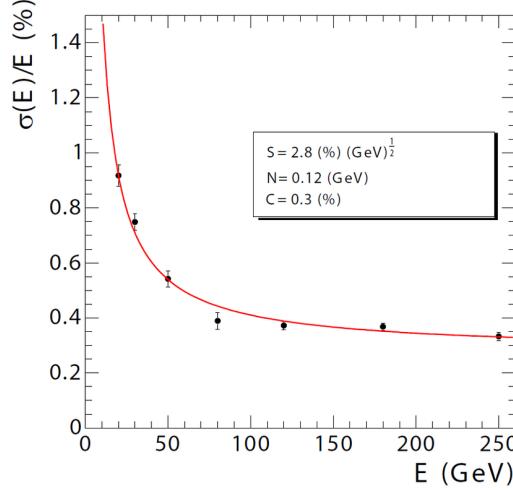


Figure 3.7: The ECAL energy resolution as a function of electron energy as measured in an electron test beam. [6]

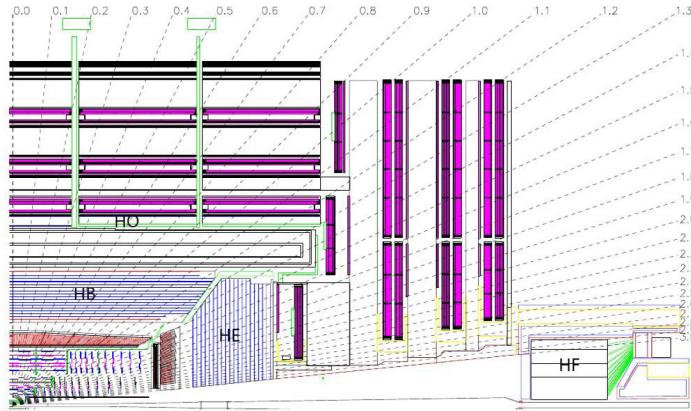


Figure 3.8: The four regions of the CMS hadron calorimeter: the inner (HB) and outer (HO) barrel, the endcap (HE) and the forward region (HF) [5]

(ending at a radius of 1.77 m) and the magnetic coil (starting at a radius of 2.95 m). In the central η region, the combined ECAL and HCAL interaction length is too small to sufficiently contain hadron showers. In order to ensure adequate sampling, especially of late starting showers, an additional layer of scintillator has therefore been added outside of the solenoid coil. This is the outer barrel (HO). It uses the coil itself as absorbing material and increases the total barrel calorimeter interaction length to $11.8 \lambda_I$. The hadron calorimeter endcaps (HE) are located in the forward region close to the beam pipe and covers the pseudorapidity range $1.3 < |\eta| < 3.0$, a region containing about 35 % of the particles produced in collisions. Due to its close proximity to the beam pipe, the endcaps need to handle extremely high rates as well as have a high radiation tolerance. As the resolution in the endcap region anyways is limited by pile-up and magnetic field effects, the hadron calorimeter endcaps were designed to minimize the cracks between HB and HE rather than having the best single-particle resolution (as is the case for the barrel). The absorber plates in the endcaps are mounted in a staggered geometry rather than on top of each other as is done in the barrel, in order to contain no dead material and provide a hermetic self-supporting construction. The HCAL is read out in individual towers with a size $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ in the barrel, and 0.17×0.17 at larger pseudorapidities. In order to obtain a completely hermetic calorimeter, an additional



Figure 3.9: The installation of the barrel HCAL wedges consisting of alternating layers of brass absorber plates and plastic scintillator, each weighing roughly 26 tonnes [8].

hadron forward calorimeter (HF) is added in the very forward region. Stretching out to a pseudorapidity of $|\eta| = 5.2$, this detector is located so close to the beam pipe that the particle rate exceeds 10^{11} per cm^2 , receiving roughly 760 GeV per proton-proton collision compared to an average of 100 GeV for the rest of the detector. It consists of a cylindrical steel structure with an outer radius of 130 cm and inner radius of 12.5 cm, located 11.2 meters from the interaction points. Also a sampling calorimeter, it consists of grooved 5 mm thick steel absorber plates, where the quartz fiber active medium is inserted into these grooves. The energy resolution of the CMS ECAL and HCAL for pions is measured in a test beam as a function of energy and is shown in Figure 3.10.

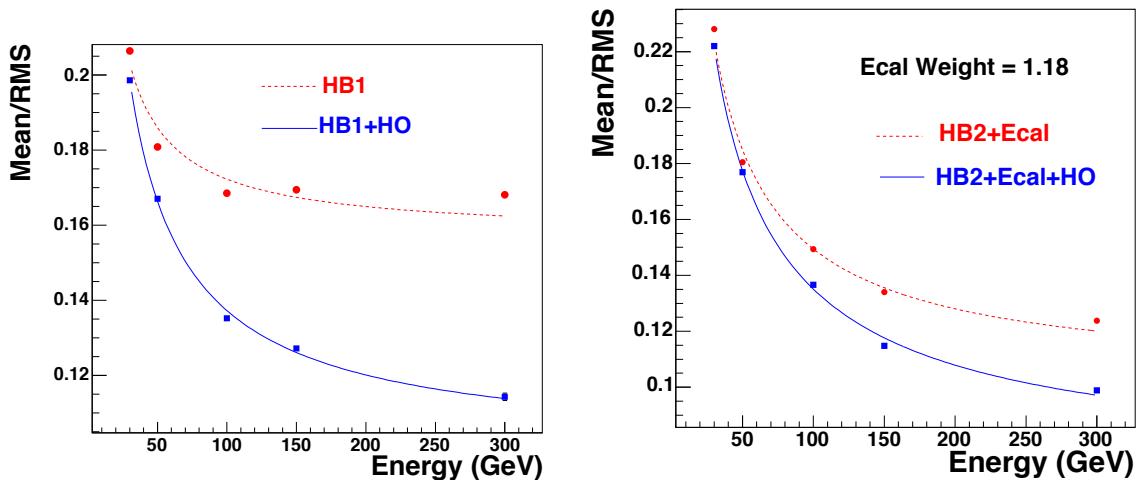


Figure 3.10: The calorimeter energy resolution as a function of pion energy using the HB only or HB+HO (left) and when adding ECal and HCal measurements (right) [9].

The typical HCAL electronics noise is measured to be 200 MeV per tower. The inclusion of the HO increases the resolution by 10 % for a pion energy of 300 GeV. The final energy resolution parametrization when using ECal+HB+HO is given by a stochastic and a constant

term, as for ECAL, and is

$$\frac{\sigma E}{E} = \frac{84.7\%}{\sqrt{E}} \oplus 7.4\%.$$

3.2.5 Muon chambers

The outer part of the CMS detector is dedicated to performing muon identification, momentum measurement and triggering. In order to do so, the muon system is made up of three types of gaseous particle detectors: drift tube (DT) chambers, cathode strip chambers (CSCs) and resistive plate chambers (RPCs), all integrated into the magnetic return yoke structure. In the barrel region, where particle rates are low and the magnetic field uniform, DT chambers are used and cover the pseudorapidity region $|\eta| < 1.2$. In the endcap regions, however, the muon rates and background levels are considerably higher and the magnetic field itself is large and non-uniform. Here; faster, finer segmented and more radiation hard CSCs are used, covering the region $0.9 < |\eta| < 2.4$. To ensure accurate muon triggering, a complimentary dedicated muon triggering system has been added both in the barrel and in the endcaps. Made out of RPCs, they provide an excellent time resolution at a sharp p_T threshold and cover the region $|\eta| < 1.6$. These chambers also assist in resolving ambiguities if multiple hits are present within a DT/CS chamber. A schematic overview of the muon system is shown in Figure 3.11.

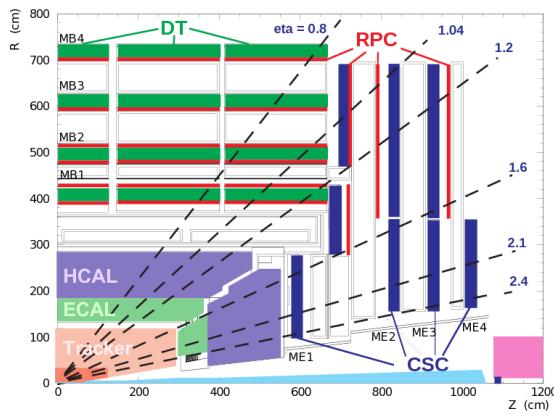


Figure 3.11: A schematic overview of the muon chambers: the DT chambers in the barrel, the CSCs in the endcaps and the redundant RPC system stretching out to $|\eta| < 1.6$ and used for triggering purposes [10].

3.3 Trigger system: From collision to disk

With protons in CMS colliding at a rate of 40 MHz, there are only 25 nanoseconds between collisions available to process event data. One billion collisions take place every second, and with an event size of roughly 1 MB, it is impossible for all of these events to be read out and stored to disk. The CMS triggering system is therefore designed to make ultra fast high-quality decisions of which events are interesting and which events are not. The first stage of triggering, called Level 1 (L1), is designed to reduce the event rate to a maximum of 100 kHz through custom-designed hardware. It uses coarse data from the muons system and calorimeters in order to make a decision on whether the event should be recorded or not, a decision that needs to happen within 3.2 micro seconds. In the mean time, the full granularity data is stored in detector front-end electronics awaiting the L1 decision. The information

used by L1 is gathered in three steps. First, trigger primitives are created. For the muon system, these consist of track segments from each of the three types of muon detectors. For the calorimeter, trigger primitives are generated by calculating the transverse energy of a trigger tower (energy deposits with an $\eta - \phi$ coverage of 0.087×0.087) and assigning it to the correct bunch crossing. Trigger primitives from the calorimeter information is then passed on to a regional trigger which defines electron, muon and jet candidates. Some of this information is passed to the muon trigger (is particle a minimum ionizing particle?). The muon trigger combines the track information with the calorimeter information and selects a maximum of four muon candidates and calculates their momentum, position, charge and quality. This is done in the global muon trigger. The output from the regional calorimeter trigger is also passed to a global calorimeter trigger which provides information about the jets, total transverse energy and missing energy in the event. Combining the information from the global muon trigger and the global calorimeter trigger, the L1 decides whether to keep the event or not by combining several decisions by simple logic operations (AND/OR/NOT) to form up to 128 algorithms.

If the events is accepted, the full event information is read out at a rate of 100 kHz and passed to the so-called ‘event filter farm’, a single processor farm made out of commodity computers. Here, the full precision of the detector data is used in order to take decisions based on offline-quality reconstruction algorithms. The goal of the HLT is to eventually reduce the event rate to an average rate of 400 Hz for offline event storage.

CHAPTER 4

Event reconstruction

4.1 Track and primary vertex reconstruction

The CMS tracker gets traversed by \mathcal{O} 1000 charged particles at each bunch crossing, produced by an average of roughly 34 proton-proton interactions happening simultaneously. This makes track reconstructions extremely challenging, and is the reason why a high granularity of the tracker is vital. The average number of vertices per event for the whole Run 2 is shown in Figure 4.1, with a combined average of 34 number of interactions per bunch crossing.

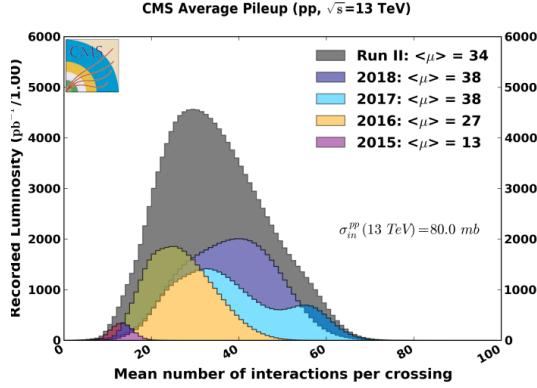


Figure 4.1: The average number of vertices per event in CMS during the Run 2 datataking [11].

Track reconstruction describes the process of taking hits from the pixel and strip detectors, combining them and estimating the momentum and flight direction of the charged particle responsible for producing the hits. It is an extremely computationally heavy process and is based on what is called a combinatorial Kalman filter [12]. A Kalman filter is an algorithm that uses time-dependent observations in order to estimate unknown variables, by proceeding progressively from one measurement to the next, improving the knowledge of the trajectory with each new measurement. The track reconstruction software in CMS (called the Combinatorial Track Finder (CTF)) constructs its collection of tracks by iteratively looping over the hits and reconstructing tracks, then removing those which are already used as inputs for a previous track. It starts from a seed in the inner most tracker layers, usually two or three hits, and then extrapolates the seed trajectories searching for additional hits to associate to that candidate. It then disregards tracks that fail certain criteria based on a χ^2 calculation taking both hit and trajectory uncertainties into account, as well as the number of missing hits. The track reconstruction algorithm is effective over the full tracker coverage range up to $|\eta| < 2.5$ and can reconstruct particles with momenta as low as 0.1 GeV or particles which are produced up to 60 cm from the beam line. In the central region, particles with a momentum of 100 GeV have a p_T -resolution of roughly 2.8 %, a transverse impact parameter resolution of 10 μm and a longitudinal impact parameter of 30 μm .

In order to define the location and uncertainty of every proton-proton interaction in an event, primary-vertex reconstruction is performed. Primary vertices lie within a radius of a few millimeters of the beam axis and are defined as the common origin of groups of tracks. The reconstruction algorithm takes as input the reconstructed tracks from the previous step which pass certain selection criteria, clusters the tracks that share a common origin and then fit for the position of each vertex. Each track must have at least 2 hits in the pixel layers and no less than 5 hits in the pixel+strip as well as a $\chi^2 < 20$ from a fit to the particle trajectory to be considered as input for the vertex finder. The primary vertex resolution is around 12 μm in x and 10 μm in z for vertices with at least 50 tracks.

Offline, all events are required to have at least one primary vertex reconstructed within a

24 cm window along the beam axis, with a transverse distance from the nominal interaction region of less than 2 cm. The reconstructed vertex with the largest value of summed physics object p_T^2 is selected as the primary interaction vertex where the hard scattering process occurred.

4.2 The Particle Flow Algorithm

After track reconstruction, what remains is an incoherent collection of tracks, calorimeter clusters and hits in the muon chambers. In order to connect these, CMS uses an algorithm called Particle Flow (PF) [13] to combine the information obtained from all sub-detectors in order to infer which particles were actually produced in the event. The reconstructed physics object in the order of which they are reconstructed are

- Muons, through hits in the tracker and in the muon chambers
- Charged hadrons, through hits in the tracker and energy deposits in the calorimeters
- Neutral hadrons, through energy deposits in the calorimeters but no hits in the tracker
- Photons, through energy deposits in the ECAL but not in the HCAL and no hits in the tracker
- Electrons, through hits in the tracker and energy deposits in the ECAL

How these different particles propagate through the CMS detector is illustrated in Figure 4.2.

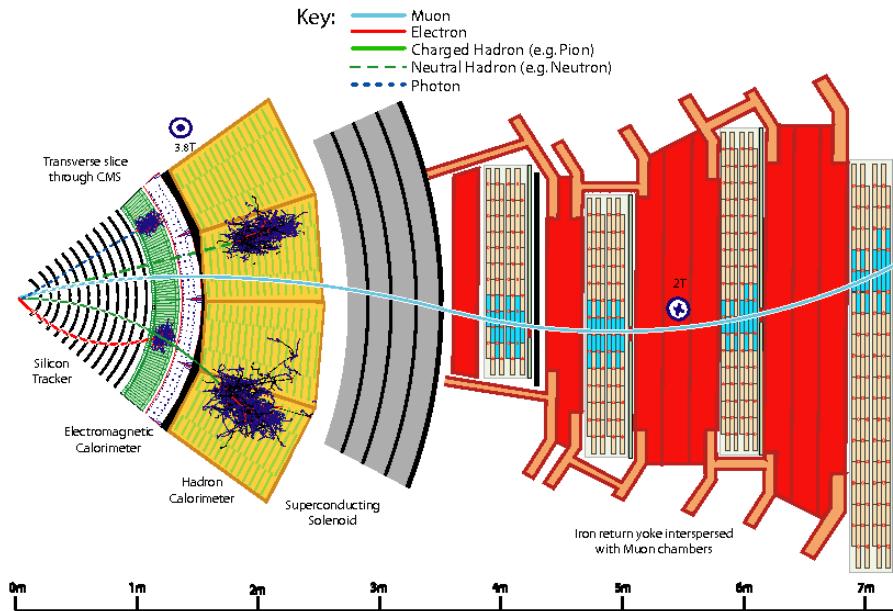


Figure 4.2: Particle interactions in the different subdetectors for a transverse slice through the CMS detector [13].

4.2.1 Reconstruction of the Particle Flow inputs

Electron tracking

Electron seeding is done in two different ways: ECAL-based or tracker-based electron seeding. In the ECAL-based method, electrons are seeded from ECAL clusters with $E_T > 4 \text{ GeV}$, where the position of the cluster is used to infer which hits in the inner tracker belongs to a given electron or positron. As a large fraction of the electron/positron energy is emitted through bremsstrahlung before even reaching the ECAL, ECAL superclusters covering a small window in η and a larger window in ϕ are defined in order to fully contain the electron as well as its bremsstrahlung photons. As these superclusters are prone to contamination, tight isolation requirements need to be applied, leading to reconstruction inefficiencies. Therefore, an additional tracker-based seeding approach has been developed. All tracks with $p_T > 2 \text{ GeV}$ are used as potential electron seeds. These tracks are then extrapolated to the ECAL and matched to the closest ECAL cluster. The ratio of the cluster energy to the track momentum is required to be 1. The electron candidates are then fit with a Gaussian-sum filter (GSF) [14] and required to pass certain criteria based on the score of a boosted-decision-tree (BDT) which combines the number of tracker hits, the χ^2 of the GSF track, the energy loss along the track, and the distance between the extrapolated track to the closest ECAL cluster.

Muon tracking

Muon tracking consists of two part: the muon spectrometer allows muons to be identified with high efficiency over the full pseudorapidity range, while maintaining a low background due to the absorbing calorimeter layers upstream. The inner tracker on the other hand, provides an accurate measurement of the muon momentum. Three muon quality flags are defined

- Standalone muon: Muon tracks based on hits in the DT or CSC only
- Global muon: A standalone muon track matched to a track in the tracker if the track parameters of the two are compatible
- Tracker muon: An inner track with $p_T > 0.5 \text{ GeV}$, a total momentum greater than 2.5 GeV and at least one muon segment matching the extrapolated inner track

Around 99% of muons produced within $|\eta| < 2.4$ are reconstructed as a global muon or a tracker muon, and very often as both. If the global and tracker muon share the same inner tracker segment, the two are combined.

Calorimeter clusters

The calorimeter clustering is performed separately for each calorimeter subdetector (ECAL barrel and endcaps, HCAL barrel and endcaps and the preshower layers). The first step is to define cluster seeds from cells with an energy exceeding some predefined threshold and in addition is larger than the energy in its neighboring cells. Topological clusters are then formed by adding cells to the seed which has at least one corner in common with a cell already in the cluster, and that has an energy which is at least twice the noise level of the detector. In Figure 4.3, an example of calorimeter clustering for a five-particle jet is shown for the HCAL (left) and ECAL (right). In the HCAL (left), two seeds have been identified (gray filled areas) inside a topological cluster consisting of 9 cells. These are then defined as two HCAL clusters, with a position as indicated by the red circles. The green solid lines correspond to charged tracks reconstructed in the tracker, both pointing to the center of the HCAL cluster seeds. The observed deposits left by the same particles are shown on the right in Figure 4.3, where the K_L^0 , π^- and the two photons from the decay of a π^0 leave distinct clusters in the ECAL. The π^+ leaves no energy deposit in this case.

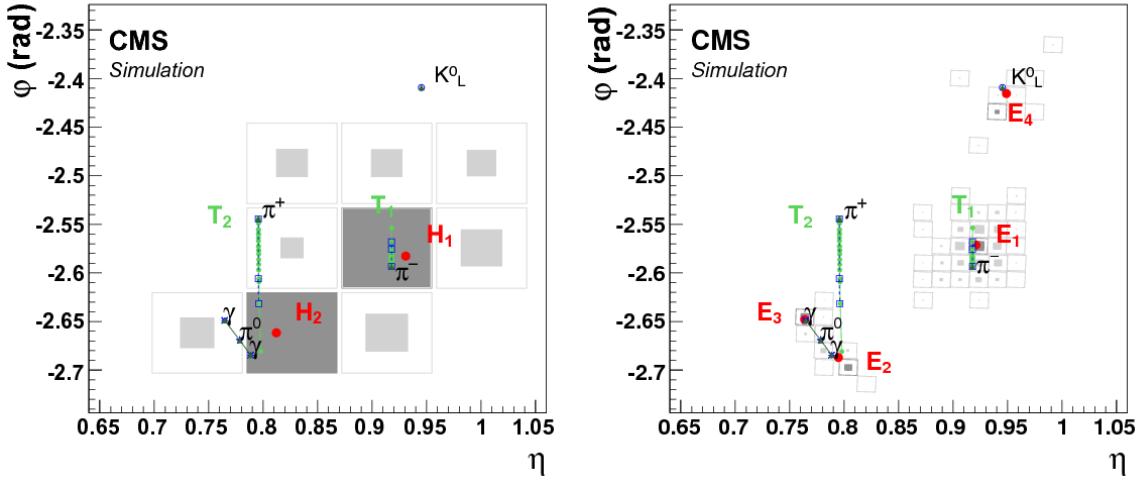


Figure 4.3: The $\eta - \phi$ vies of calorimeter clusters generated by a five-particle jet in the HCAL (left) and in the ECAL (right). The squares correspond to calorimeter cells, where the inner area is proportional to the logarithm of the cell energy. Cluster seeds are depicted in dark gray. The dotted blue lines correspond to the simulated particle trajectory, while the green lines correspond to charged tracks reconstructed in the tracker [13].

4.2.2 Particle Flow identification

The link algorithm

The link algorithm is the algorithm responsible for combining the particle flow elements from different subdetectors. It can test any pair of elements in the event based on specific requirements depending on the nature of the element, but is restricted to the nearest neighbors in the $\eta - \phi$ plane. The outputs of the link algorithm are so-called *PF blocks* of linked elements, either directly linked or linked through having common elements.

- **Inner track - calorimeter cluster link:** The track is interpolated from its last hit, through the preshower layers, the ECAL and ending in the HCAL at an interaction length depth of 1. A link is made if the track is within the cluster area, where the areas is enlarged by up to a cell in each direction to account for detector gaps. In case several ECAL/HCAL clusters are linked to the same track, only the one with the smallest distance in $\eta - \phi$ is kept.
- **Calorimeter cluster - cluster link:** A link between ECAL and HCAL clusters as well as between ECAL and preshower clusters is made when the cluster position of the more granular calorimeter is within the cluster envelope in the less granular calorimeter. Also here, if there is link overlap, only the link with the smallest distance is kept
- **Inner tracker -muon chamber link:** As described in Section 4.2.1

For each PF block, the reconstruction proceeds in the following order. First, muons are reconstructed and their corresponding PF elements removed from the PF block. Then the electrons are reconstructed, with the hopes of removing their corresponding bremsstrahlung photons from the list of PF elements. Energetic photons are reconstructed in the same step. Finally, neutral and charged hadrons are reconstructed.

Muons

First, isolated global muons are selected by requiring the sum of track p_T and calorimeter energy deposits within a cone of $\Delta R = 0.3$ not belonging to the muon track, to be smaller than 10 % of the muon p_T . If the muons are non-isolated, they are required to pass the tight muon requirement [15] and have at least three matching track segments in the muon detector or have matched calorimeter deposits compatible with being a minimum ionizing particle. Muons failing both the requirements above are kept if the standalone muon track is of high quality and have a lot of hits in the muon detectors, otherwise they are discarded. The muon momentum is defined from the inner tracker measurement if the muon p_T is less than 200 GeV. Otherwise, its chosen according to the track fit with the smallest χ^2 probability.

Muons used in this thesis are required to pass an isolation requirement in order to suppress the background from QCD multijet events where jet constituents are identified as muons. For this, a cone of radius $\Delta R = 0.3$ is constructed around the muon direction. The isolation parameter is defined as the scalar sum of the transverse momenta of all additional reconstructed tracks within the cone, divided by the muon p_T . Muon candidates with an isolation parameter less than 0.1 are considered isolated and are used for further analysis. Further, the following selection criteria are applied:

- The χ^2 of the global muon track fit must be less than 10
- At least one muon-chamber hit is included in the global-muon track fit. The global muon track fit must include at least one muon chamber hit
- Muon segments in at least two of the muon stations must be matched to the muon tracker track
- The inner tracker track must be no more than 2 millimeters from the primary vertex in the xy plane and no more than 5 millimeters in the longitudinal direction
- At least one hit in the pixel detector.
- At least six layers of the inner tracker must contain hits.
- At least three matching track segments must be found in the muon detectors

Electrons

The electrons are seeded from a GSF track, as described in Section 4.2.1. To differentiate electrons from charged hadrons, the energy deposit in the HCAL within a distance of 0.15 in the $\eta - \phi$ plane of the supercluster , is required to be less than 10 % of that of the supercluster. The electron candidate must further pass a requirement on the output of a dedicated electron-identification BDT, using inputs such as track-cluster distance, track χ^2 and number of hits as input. In this step, isolated photons are also reconstructed, seeded from ECAL superclusters with $|E_T| > 10 \text{ GeV}$ and no link to a GSF track. All the tracks and calorimeter deposits used to reconstruct electrons and isolated photons are further removed from the list of PF blocks.

Only electrons passing certain quality requirements, corresponding to the CMS electron HEEP ID, are used in this thesis. These requirements are listed in Table 4.1, with the following variable definitions:

- E_T : The supercluster energy $x \sin(\theta_{track})$ where θ_{track} is the electron track polar angle measured in the inner tracker layer and extrapolated to the interaction vertex.

- η^{sc} : η of the electron supercluster.
- **isEcalDriven**: Electron is found through ECAL requirements rather than through Particle Flow and the tracker.
- $\Delta\eta_{in}^{seed}$: η difference between the track position measured in the inner layer, extrapolated to the interaction vertex and to the calorimeter, and the η of the supercluster.
- $\Delta\phi_{in}$: ϕ difference between the track position measured in the inner layer, extrapolated to the interaction vertex and to the calorimeter, and the ϕ of the supercluster.
- **H/E**: Ratio of hadronic energy in the calorimeter towers within a cone of radius 0.15 centered at the electrons calorimeter position, to the electromagnetic energy of the supercluster.
- $\sigma_{inj\eta}$: Measure of the energy spread in η in units of crystals of electron energy in a 5×5 block centered on the seed crystal.
- **ECAL Isolation**: The transverse electromagnetic energy of all reconstructed hits (with $E > 0.08$ GeV) in a cone of radius 0.3 centered at the electron calorimeter position, excluding those in an inner cone with a radius of 3 crystals and an η strip with a width of 3 crystals.
- **Hadronic Depth Isolation**: Defined as the transverse depth of the hadronic energy in the HCAL inside a cone of 0.3 centered on the electron calorimeter position, excluding towers in a cone of 0.15 radius.
- **Track p_T Isolation**: The sum p_T of the tracks in a ΔR cone of 0.04 to 0.3, excluding an η region of 0.015.
- d_{xy} : Transverse distance between the electron track and the primary vertex.

Variable	Barrel	Endcap
E_T	> 35 GeV	> 35 GeV
η range	$ \eta_{sc} < 1.4442$	$1.566 < \eta_{sc} < 2.5$
isEcalDriven	yes	yes
$\Delta\eta_{in}^{seed}$	< 0.004	< 0.006
$\Delta\phi_{in}$	< 0.06	< 0.06
H/E	$< 1/E + 0.05$	$< 5/E + 0.05$
full 5×5 $\sigma_{inj\eta}$	n/a	< 0.03
full 5×5 $E^{1 \times 5}/E^{5 \times 5}$	> 0.94 OR $E^{1 \times 5}/E^{5 \times 5} > 0.83$	n/a
EM+Had. Depth Iso.	$< 2 + 0.03 \times E_T + 0.28 \times \rho$	For $E_T < 50$ GeV: $< 2.5 + 0.28 \times \rho$ else: $< 2.5 + 0.03 \times (E_T - 50) + 0.28 \times \rho$
Track p_T iso.	For $E_T < 100$ GeV: < 5 GeV / $< 5 + 1.5 \times \rho$ else: $< 5 + 1.5 \times \rho$	$< 5 + 0.5 \times \rho$
Inner Layer Lost Hits	≤ 1	≤ 1
d_{xy}	< 0.02	< 0.05

Table 4.1: Summary of the electron requirements allied to all electrons used in this analysis.

Hadrons

Finally, after the removal of muons and electrons, the remaining hadrons and non-isolated photons are identified. HCAL clusters with no track link are defined as neutral hadrons, while

ECAL clusters with no track link are defined as photons (photons are exclusively associated to the ECAL deposits as neutral hadrons leave only 3 % of their energy in the ECAL). The remaining HCAL clusters are then linked to one or more tracks from the inner tracker. In order to determine the particle content within a cluster, the sum of track momenta and the calorimeter energy is compared. If the calorimeter energy is compatible with the sum of track momenta, a particle for each track is inferred, with its corresponding energy taken from the track momentum. If the calorimeter energy is larger than the sum of track momenta, a photon or a neutral hadron is added, together with one charged hadron for each track within the cluster area.

Missing transverse energy

Neutrinos (and other predicted non-SM weakly interacting particles) do not interact in the detector and are instead inferred from the presence of a momentum imbalance in the detectors transverse plane. The missing transverse momentum is defined as the negative p_T vector sum of all reconstructed PF candidates in the event

$$\vec{p}_T^{\text{miss}} = - \sum_i^N \vec{p}_{T,i}. \quad (4.1)$$

and its magnitude, $|\vec{p}_T^{\text{miss}}|$, is denoted the missing transverse energy E_T^{miss} (which is used as a proxy for the neutrino p_T).

4.3 Pile-up removal

Particles originating from proton-proton interactions not associated with the hardest primary vertex, are denoted pileup events. These distorts observables of interest from the hard scattering event and must be mitigated through dedicated pileup removal techniques

4.3.1 Charged Hadron Subtraction

As mentioned previously, primary vertices are reconstructed using tracks from charged hadrons. If a primary vertex does not correspond to the hard scattering vertex of the event, the charged hadrons (as reconstructed through Particle Flow) associated to this vertex (called pileup vertex) are removed from the event collection of particles and will not participate in any further object reconstruction. This method is denoted charged hadron subtraction (CHS).

4.3.2 Pile up per particle identification (PUPPI)

CHS was the default pileup removal algorithm in CMS until very recently. In 2014, a new pileup removal algorithm with improved performance was proposed; the pileup per particle identification (PUPPI) [16] algorithm. PUPPI uses a combination of local shape information, event pileup properties and tracking information to compute a weight describing the degree of ‘pileup-likeness’ of a given particle. First, a variable denoted α is computed based on the difference between soft radiation coming from pileup and the harder collinear QCD pattern. The shape of α for charged particles is then used as a proxy for all pileup particles and is used on an event-by-event basis to calculate a weight for each particle. This weight in turn describes the degree to which particles are pileup-like and are used to rescale the particle four-momenta.

The shape variable for a given particle i is defined as

$$\alpha_i = \log \sum_{\substack{j \in \text{Ch, PV} \\ j \neq i}} \left(\frac{p_{T,j}}{\Delta R_{ij}} \right)^2 \Theta(R_0 - \Delta R_{ij}), \quad (4.2)$$

where Θ is the step function and j refers to the neighboring charged particles from the primary vertex within a cone of radius $R_0 = 0.4$. Charged particles are defined as coming from the primary vertex if they are associated to the leading vertex of the event or are within a distance of $d_z < 0.3$ cm from the leading vertex.

In order to determine the probability that a particle comes from pileup, a χ^2 calculation is performed. The probability is defined as

$$\chi_i^2 = \frac{(\alpha_i - \bar{\alpha}_{PU})^2}{RMS_{PU}^2}, \quad (4.3)$$

where $\bar{\alpha}_{PU}$ is the median value of the α_i distribution for pileup particles in the given event and RMS_{PU} is its RMS.

Each particle (neutral and charged) is then assigned a weight $w_i = F_{\chi^2, NDF=1}(\chi_i^2)$, where $F_{\chi^2, NDF=1}$ is the cumulative distribution function of the χ^2 distribution with one degree of freedom. Particles with $w_i < 0.01$ are rejected. In addition, a cut on the weighted p_T of neutral particles of $w_i \cdot p_{T,i} > (A + B \cdot N_{PV})$ GeV is applied, where N_{PV} correspond to the number of reconstructed vertices in the event and A and B are tunable parameters.

The performance of the PUPPI algorithm compared to CHS for jet observables is shown in Figure 4.4.

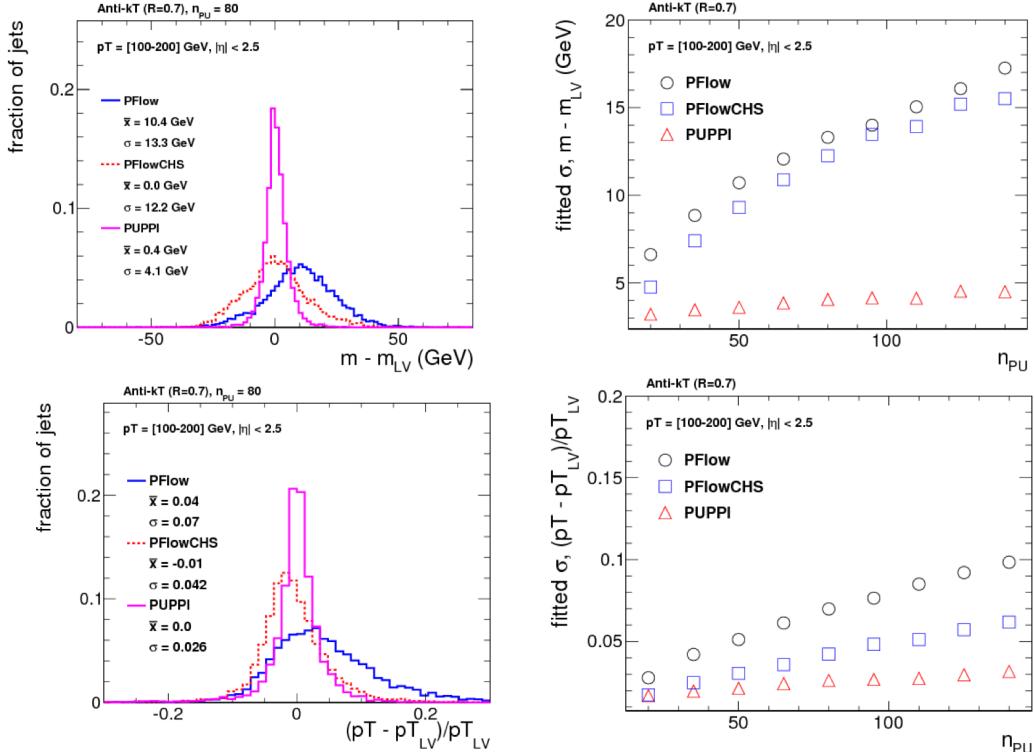


Figure 4.4: The mass (top) and p_T (bottom) resolution comparing PF only (blue), PF+CHS (red) and PUPPI (pink) jets. The absolute resolution (left) as well as the resolution as a function of the number of reconstructed primary vertices in the event (right) is shown [16].

The top row shows the absolute mass resolution (left) as well as the mass resolution as a function of N_{PV} for CHS jets (red) and PUPPI (pink) jets. The bottom row shows the corresponding quantities but for jet transverse momentum. A significantly better resolution on jet observables can be achieved using PUPPI compared to CHS.

4.4 Jet reconstruction

As explained in Section 1.1.4, quarks and gluons are never themselves visible in a detector. Within 10^{-23} seconds, the timescale of the strong interactions, they fragment and hadronize into a collimated spray of hadrons, a so-called jet. In order to infer the properties of the original parton generating the jet, the properties of the full particle spray needs to be evaluated. Combining these particles algorithmically is non-trivial, and several algorithms designed to do, called jet clustering algorithms, exist. These provide a set of rules for grouping particles together into jets and are usually based on certain distance requirements between particles as well as rules for how to recombine their momenta. Thanks to Particle Flow, objects like charged hadrons, neutral hadrons and photons together with their estimated energy and direction are already defined, and jet clustering in CMS therefore consists of associating these particles to one common origin.

4.4.1 Jet clustering

The most common jet clustering algorithms used in hadron colliders are the Cambridge/Aachen algorithm [17], the k_T algorithm [18] and the anti- k_T algorithm [19]. These are all sequential recombination algorithms, meaning they systematically go through each particle pair in the event and recombines them into one particle if the combination satisfies certain criteria. The rules, shared by all three algorithms, are as follows:

1. For each pair of particles i and j , compute the longitudinally invariant distances

$$d_{ij} = \min(p_{ti}^{2p}, p_{tj}^{2p}) \frac{\Delta R_{ij}^2}{R^2}, \text{ with } \Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 \quad (4.4)$$

$$d_{iB} = p_{ti}^{2p}, \quad (4.5)$$

where d_{ij} is a measure of the relative transverse momenta between the particles, ΔR_{ij}^2 is the distance between them in the $\eta - \phi$ plane (which can be roughly translated into a jet radius), ΔR^2 corresponds to a distance parameter which controls the extension of the jet and d_{iB} is the distance between the particle and the beam. The parameter p is what separates the three algorithms from one another and controls the relative power of energy versus geometrical scales. For the anti- k_T algorithm, it is defined as $p = -1$, for the k_T algorithm $p = 1$ and in the case of the C/A algorithm, $p = 0$. The consequences of these choices are explained in detail below.

2. Find the minimum distance of d_{ij} and d_{iB} .
3. If this is d_{ij} , recombine particles i and j and return to step 1.
4. If it is d_{iB} , the particle is defined to be a final state jet, and is removed from the list of particles. The algorithms proceeds back to step 1.
5. Repeat until no particles remain.

Infrared and collinear safety

There are two requirements that are extremely important when defining jet algorithms: They must be 1) *infrared* (IR) and 2) *collinear* (C) safe. *Infrared* safety corresponds to the requirement that if the final state particles are modified by the presence of a soft emission, and there are always soft emission in QCD events (both perturbative and non-perturbative), then the set of hard jets should remain unchanged. This is illustrated by the two left figures in Figure 4.5.

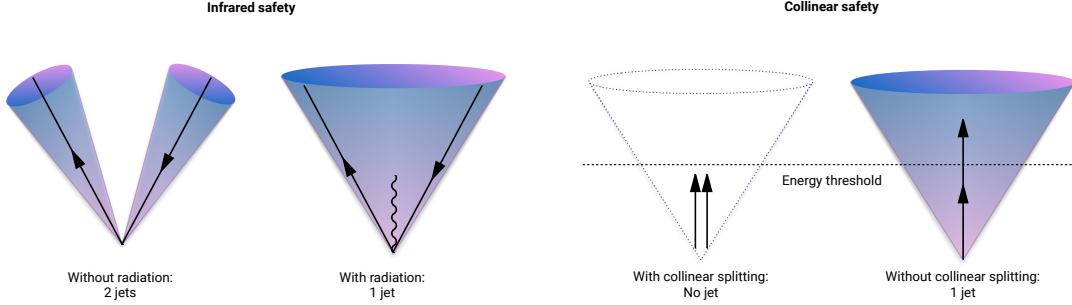


Figure 4.5: An illustration of what would happen for an infrared (left) and collinear (right) unsafe jet algorithm. If an algorithm is infrared unsafe, the presence of a soft emission changes the jet configuration. If an algorithm is collinear unsafe, then if a parton undergoes a collinear splitting this will change the configuration of the jet

Here, the algorithm is infrared unsafe: the presence of an additional soft gluon changes the jet configuration from 2 to 1 jets. If an algorithm is *collinear* unsafe, it means that the jet configuration would change if the hard parton undergoes collinear splitting (which a hard parton often does as part of the fragmentation process and which are also part of non-perturbative dynamics, like the decay of highly energetic hadrons). This is shown in the two left figures of Figure 4.5, where a hard parton undergoing collinear splitting fails to be reconstructed due to its daughters being below the energy threshold of the algorithm.

All sequential recombination algorithms are trivially infrared safe.

The k_T algorithm

The k_T algorithm is the oldest of the sequential recombination algorithms and, due to its $p = 1$ definition in the distance measures, follows the QCD branching structure in both p_T and in angle (in reverse). Soft particles are clustered together first, and the final step is the clustering of the two hardest particles. A consequence of this definition is that there is nothing that keeps arbitrarily soft particles from being defined as jets, and a minimum cut on the jet p_T should be introduced. Despite several favorable qualities, the k_T algorithm is not the algorithm of choice in most hadron collider experiments due to the irregular jets it produces, a consequence of clustering soft particles first.

The Cambridge/Aachen algorithm

The Cambridge/Aachen algorithm, with $p = 0$ in the distance measures, follows the QCD branching structure only in angle as the clustering order is based solely on spatial separation. The simplest of the algorithms, it recombines all pairs close in ΔR until $\Delta R_{ij} > R$. The benefits of this is that the clustering history contains information about the presence of any geometrical substructure within a jet, a feature that will become important in Section 4.5.

The anti- k_T algorithm

The default jet clustering algorithm in CMS is the anti- k_T algorithm [19], which follows the rules and distance measures above with $p = -1$. The algorithm favors the clustering of high- p_T - high- p_T and high- p_T – low- p_T particles first, disfavoring clustering between soft particles. That means the algorithm grows around a hard core, yielding jets with a well-defined cone shaped area. Together with being IRC-safe and insensitive to the underlying event (any event not arising from the primary hard scattering process) and pileup, makes it the main jet algorithm in CMS.

A comparison of the resulting jet area in the $\phi - \eta$ plane after clustering with either k_T , C/A and anti- k_T , is shown in Figure 4.6. The z-axis correspond to the parton p_T . One can clearly see that when clustering with the anti- k_T algorithm, the produced jets are circular, with a radius set by R , around the hardest parton.

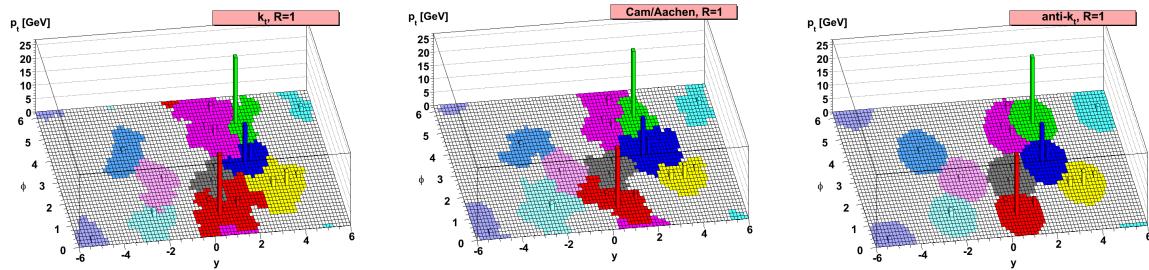


Figure 4.6: A comparison of the resulting jet cone area in the $\phi - \eta - p_T$ plane after clustering the same event with three different jet algorithms: k_T , C/A and anti- k_T . [19]

4.4.2 PF jets in CMS

Jet algorithms in CMS mainly use PF candidate four-vectors as input and a pileup removal algorithm is usually applied before clustering occurs. If using CHS (Section 4.3.1), charged hadrons not associated to the primary vertex are discarded before clustering. If PUPPI is used (Section 4.3.2), all the PF candidates are reweighted based on how likely they are to have originated from pileup. For the anti- k_T algorithm, CMS by default uses two jet cone sizes: $R=0.4$ and $R=0.8$. Jets with $R=0.4$, called PFAK4, are used for single-prong jets while the larger $R=0.8$ jets, PFAK8, are more often used when looking for jets containing multiple hard quarks/gluons in order to contain all the hadronization products.

These jets are further required to pass certain jet identification requirements provided by the JetMET POG [20], in order to distinguish them from fake jets. All jets used in this analysis are required to pass the *tight ID* requirements which are as follows:

- The jet must contain at least two PF constituents
- At least one of these constituents must be a charged hadron
- The fraction of jet energy coming from neutral hadrons must be < 0.90
- The fraction of jet energy coming from neutral electromagnetic energy must be < 0.90
- The fraction of jet energy coming from charged electromagnetic energy must be < 0.99

4.4.3 Jet energy corrections

All jets are further corrected for nonlinearities in p_T and rapidity using standard CMS jet energy corrections (JEC), as described in Ref. [21]. These are intended to bring the measured jet energy closer to the true jet energy by correcting the jet energy scale (JES) and jet energy resolution (JER). The energy corrections are derived in three steps:

- L1: Energy offset corrections intended to remove pileup and electronic noise, both for data and simulation
- L2L3: A relative (L2) and absolute (L3) correction to particle level jet response for simulation only
- Residual: A correction for data only meant to correct for residual differences between data and simulation

These are illustrated in Figure 4.7.

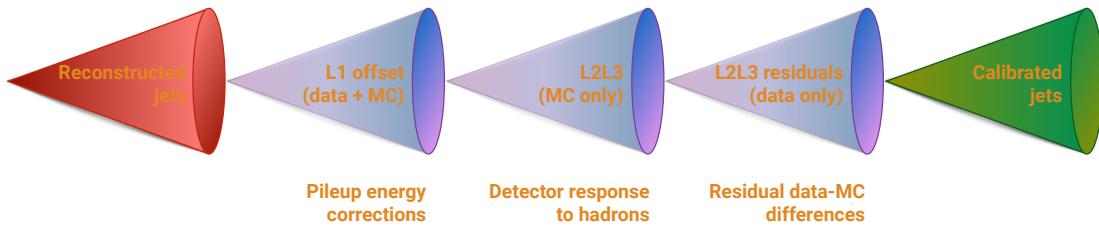


Figure 4.7: The CMS jet energy corrections are derived in three steps: A correction due to offset energy coming from pileup, applied to data and MC, a correction due to the particle level jet response, also applied to data and MC and finally a correction to account for residual differences between data and MC.

L1 offset correction

The largest correction is the L1 pileup offset correction, which are meant to subtract the additional energy in a jet due to pileup. This is done on an event-by-event basis through the *jet area method* which uses the jet effective area multiplied by the average event energy density to calculate the size of the offset energy to be subtracted from each jet. An additional p_T - and η -dependent term is added in order to account for different pileup densities in different parts of the detector and for different jet energies. For data, an additional scalefactor to account for data and simulation differences is computed. This is done by constructing a *Random Cone* (*RC*) centered at a given η, ϕ and dividing the energy density within that cone in data, evaluated in a dataset with no hard interactions (*Zero Bias*), by that of the true energy offset in simulation

L2 relative and L3 absolute corrections

After L1 corrections are applied, corrections to account for the detector response to hadrons are derived based on the true detector response in QCD MC. The simulated particle response is defined as the ratio

$$R_{\text{particle}} = \frac{p_{T,\text{reco}}}{p_{T,\text{particle}}} \quad (4.6)$$

These are derived in bins of particle level p_T and reconstructed η : The L2 relative corrections are intended to uniform the detector response and are derived as a function of η , while the L3 absolute corrections are derived as a function of jet p_T . These corrections are applied both to data and to MC.

Residual data corrections

After L1 and L2L3 corrections are applied, two additional corrections are derived for data only in order to account for any residual discrepancies between data and MC. This is done by looking at the transverse momentum balance between a jet which is to be calibrated, and some reference object (either another jet, a Z boson or a photon). If the jet energy scale is not equal to one, a p_T imbalance will be visible. The measurements are performed in a data dijet sample, where the statistical uncertainty is small but the energy of the reference object poorly measured, as well as in $Z(\mu\mu)$ +jet, $Z(ee)$ +jet and γ +jet samples, where the energy of the Z and γ is very well known but the statistics are small.

The ‘L2 relative’ residual correction is measured in dijet events by comparing the measured p_T of the reference jet, required to be central with $\eta < 1.3$, to that of the calibration jet, with an unconstrained η . This is done as a function of jet η , in bins of average jet p_T . The ‘L3 relative’ residual correction, is instead measured in W/γ + jet events by comparing the measured jet p_T to the p_T of the precisely measured Z/γ , as a function of jet p_T . The response,

$$R_{\text{jet}, p_T} = \frac{p_{T,\text{jet}}}{p_{T,\text{ref}}}$$

is then evaluated in data and in simulation. The ratio of the two, $R_{\text{data}}/R_{\text{MC}}$, defines the residual corrections.

The above description of jet energy corrections in CMS is meant as a rough, instructive summary only. A full description of the measurement techniques used in CMS, can be found in [21].

4.5 Jet substructure reconstruction

In analyses looking for highly energetic (‘boosted’) vector bosons like, a main theme of this thesis, the opening angle between the vector boson quark decay products become so small that the highly boosted boson appears as a single large jet instead of two well-separated smaller jets. The distance between the two quarks, in the case of an hadronic decay, depends on the mass of the vector boson and its p_T and goes as

$$\Delta R = \frac{2M_V}{p_{T,V}}. \quad (4.7)$$

Above a W boson p_T of 200 GeV, the two quarks are therefore merged into a single large cone jet of size $R = 0.8$. A sketch of the two different situation is shown in Figure 4.8. If the W p_T is well below 200 GeV, its decay products are well-defined jets in their own right (left). However, once the W transverse momenta starts exceeding 200 GeV, both the quarks are completely contained within a single jet (right).

In order to distinguish hadronically decaying vector boson from QCD quark/gluon jets, the jet mass would in principle be a good discriminant as we know the W has a mass of around 80 GeV while the quark/gluon mass is close to zero. At very high transverse momenta,

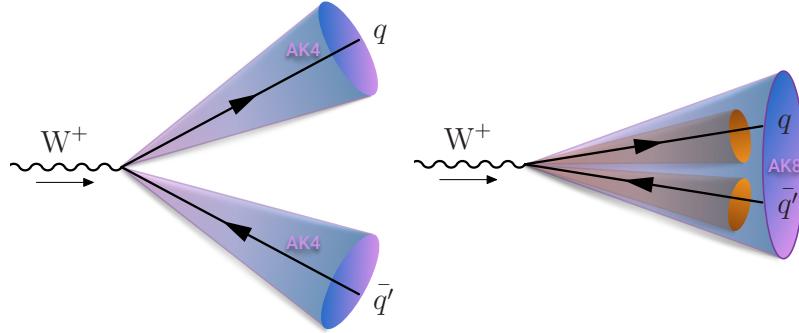


Figure 4.8: If the mass of the resonance is low enough, the quark decay products of each vector boson are well separated and clustered into distinguishable AK4 jets (left). If the transverse momentum of the vector boson is greater than 200 GeV, the vector boson decay products are merged into one single large cone AK8 jet.

however, the width (and therefore the mass) of QCD jets may become equally large. In addition, diffuse radiation caused by the Underlying Event and pileup give rise to a significant number of additional particles in the event contributing to the total jet mass. Therefore, being able to accurately and efficiently separate highly boosted QCD jets from highly boosted vector bosons, requires other methods. In order to get rid of UE and pileup, algorithms like PUPPI and CHS can be used. Then, to improve the mass resolution further, dedicated grooming algorithms must be applied.

4.5.1 Grooming

Grooming was introduced as a tool to improve the signal, most often $W/Z/$, mass resolution without significantly changing the background and signal event numbers. It consists of removing the softest parts of a jet in order to resolve its ‘true’ mass, by means of reclustering and identifying soft particles within the jet which then are removed.

Trimming

The trimming algorithm [22] is a grooming algorithm mostly used at trigger level in CMS (also where it is used in this thesis) due to it being less aggressive than other grooming algorithms. It works in the following way: Starting from a large jet clustered with either anti- k_T or C/A (in the case of CMS), it reclusters the jet using the k_T algorithm in order to create subjets of some size R_{sub} . It then proceeds to check whether each subjet has a momentum fraction above a certain threshold,

$$p_{T,i}/p_{T,jet} > p_{T,frac}.$$

If the subjet fails this requirement, it is removed. The remaining subjets are then assembled into a new ‘trimmed’ jet. The effect of trimming on real W jets and QCD quark/gluon jets for different values of r_{sub} and $p_{T,frac}$ is shown in Figure 4.9. The best signal mass resolution is obtained with $r_{sub} = 0.2$ and $p_{T,frac} = 0.03$, which is also the parameter setting that provides the best signal and background discrimination by pushing the QCD jet mass closer to zero. These are the default values of the tuned parameters of the trimming algorithm in CMS ($r_{sub} = 0.2$ and $p_{T,frac} = 0.03$).

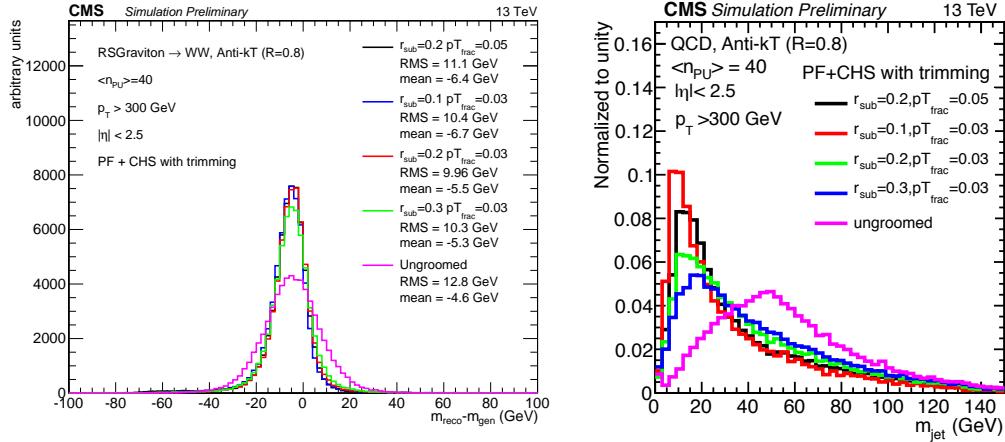


Figure 4.9: The effect of trimming on a signal jet (left) and a background jet (right) for different values of the tuned parameters r_{sub} and $p_{T,frac}$ [23].

Pruning

The pruning algorithm, in addition to removing soft particles, has an additional requirement on the distance between any recombination that are at wide angle. It proceeds by reclustering the jet with the C/A algorithm, requiring at each step that

$$\frac{\min(p_{T,i}, p_{T,j})}{p_{T,P}} > z_{cut} \quad \text{and} \quad \Delta R_{i,j} < D_{cut} = \frac{2r_{cut}m_{jet}}{p_T}.$$

The first requirement is a requirement on the hardness of the combination. $p_{T,i}$ and $p_{T,j}$ correspond to the transverse momenta of each protojet (single particle or group of particles already combined in a previous step) and $p_{T,P}$ is the combined p_T of the two. The protojet with the lowest transverse momenta is removed if its hardness is below z_{cut} , or if it forms an angle wider than D_{cut} relative to the axis of the recombination of the two protojets. In CMS, the tuned parameters are set to $r_{cut} = 0.5$ and where $z_{cut} = 0.1$. Figure 4.10 shows the ungroomed as well as the pruned jet mass distribution for signal (left) and background jets. The highest amount of signal and background separation in CMS, is achieved with $r_{cut} = 0.5$ and $z_{cut} = 0.1$.

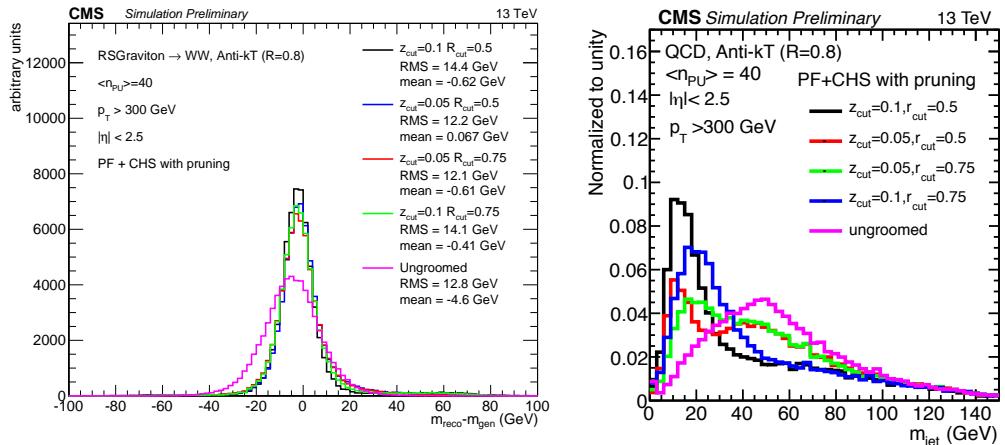


Figure 4.10: The effect of pruning on a signal jet (left) and a background jet (right) for different values of the tuned parameters z_{cut} and r_{cut} [23].

Modified Mass Drop Tagger and Soft Drop

The modified mass drop tagger (mMDT) [24] (a modified version of the originally suggested mass drop tagger [25]) is based on the idea that a W/Z jet is formed by two quark subjets and that, therefore, the mass of each subjet is much smaller than their combined mass (and much smaller than the mass of the boson itself). A QCD jet is, on the other hand, formed by continuous soft radiation, meaning that its heaviest subjet should be close to the mass of the jet itself. The mMDT tagger therefore starts from an already clustered jet, reclusters it with the C/A algorithm and then declusters it again defining subjets s_1 and s_2 . It then looks for a significant mass drop going from total jet mass to the mass of each subjet, and checks that the splitting is not too asymmetric. The modified mass drop condition is generalized through the soft drop declustering method [26], simply called Soft Drop, which allows for different types of angular requirements to enter the condition. The Soft Drop condition is the following

$$\frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > z_{cut} \frac{\Delta R_{12}}{R_0}^\beta.$$

If no significant mass drop occurred and the splitting is not too asymmetric, the condition is met and the full jet is deemed the softdrop jet. Otherwise only the highest- p_T subjet is kept and the declustering continues. If the jet can not be declustered any further, it can either be removed from consideration, so-called ‘tagging’-mode, or deemed the final soft-dropped jet, ‘grooming’-mode. A $\beta = 0$ corresponds to the modified mass drop tagger and removes all soft emission from the jet. For $\beta > 0$, soft radiation is removed, but some fraction of soft-collinear radiation is kept. Lastly, with $\beta < 0$, Soft Drop can remove soft as well as collinear radiation. The performance of Soft Drop on W jets and QCD quark/gluon jets for different values of β is shown in Figure 4.11. The modified mass drop tagger (Softdrop with $\beta=0$) with $z_{cut} = 0.1$ is the default Soft Drop settings in CMS, due to it providing the best signal/background discrimination while maintaining an excellent signal mass resolution.

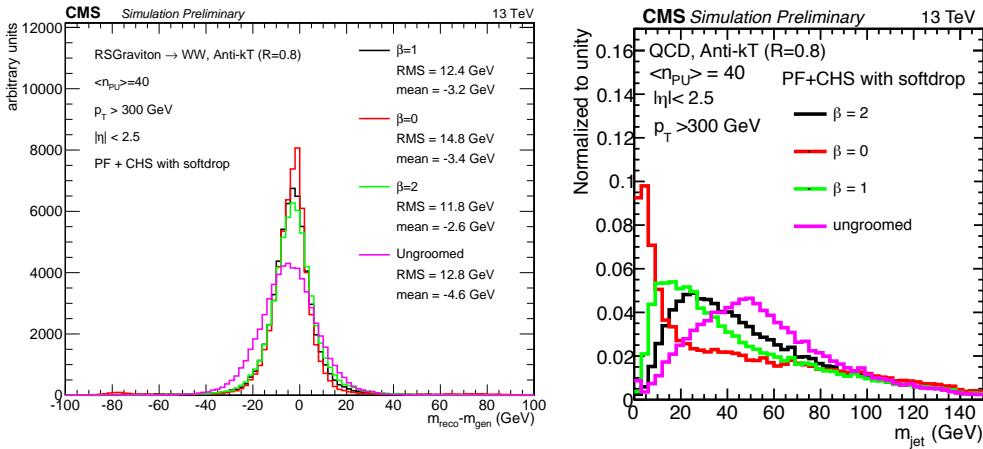


Figure 4.11: The effect of softdrop on a signal jet (left) and a background jet (right) for different values of the tuned parameters β . $\beta = 0$ corresponds to the Modified Mass Drop Tagger, which is the default Softdrop setting in CMS [23].

4.5.2 N-subjettiness

After hopefully having resolved the particle mass with one of the grooming algorithms above, there is still discriminating information to be gathered from the jet structure itself. A W or Z

jet consists of two well-defined high- p_T subjets. A quark/gluon jet on the other hand, made from a single parton, consists of several large angle, asymmetric splittings, as illustrated in Figure 4.12.

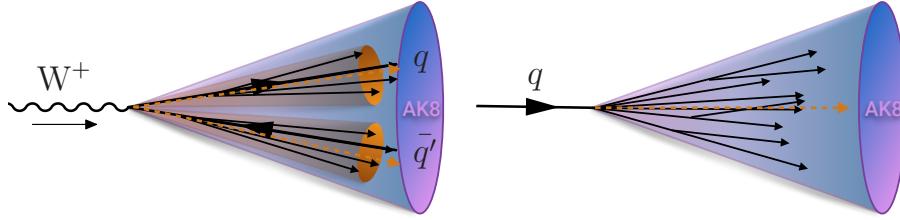


Figure 4.12: A jet stemming from the decay of a W will usually have two well-separated high- p_T subjets, while a jet with a single-prong origin consists of several large angle splittings.

The N-subjettiness algorithm [27] takes advantage of this fact by attempting to count the number of hard subelements within a jet. This is quantified through the n-subjettiness variable, τ_N , defined as

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}) \quad (4.8)$$

where k runs over all the jet constituents, $p_{T,k}$ is the constituent transverse momentum and $\Delta R_{i,k}$ is the distance between the constituent and candidate subjet axes. These subjet axes are obtained through a one-pass optimization procedure which minimizes N-subjettiness [28]. The normalization factor in front is given as

$$d_0 = \sum_k p_{T,k} R_0 \quad (4.9)$$

where R_0 corresponds to the cone size of the initial jet. With this definition, jets with $\tau_N = 0$ have most of its constituents aligned along the subjet axes. However, if $\tau_N \gg 0$, a large fraction of the energy is radiated away from the subjet directions and are more likely to have more than N subjets. In CMS, and as recommended by the authors in [27], the ratio τ_2/τ_1 is used to discriminate W jets from QCD jets. The reason for this is that, while signal jets are expected to have a large τ_1 , quark/gluon can similarly have large τ_1 due to the diffuse radiation present. However, QCD jets with a large τ_1 tend to have an equally large τ_2 , while signal jets do not, hence the ratio of the two provides greater separation power. In CMS, the n-subjettiness algorithm is by default applied to ungroomed jets. The distribution of τ_{21} for signal and background jets with different pileup subtraction algorithms applied are shown in Figure 4.13, where τ_{21} in combination with PF+PUPPI (green) yields a distribution most similar to the generated one (black).

4.5.3 Vector boson tagging

In order to discriminate W and Z bosons from quark/gluon jets a combination of a groomer and shape-tagger (like n-subjettiness) is usually used. Typical values for a W-tagger are jet groomed mass between 60 and 100 GeV and $\tau_{21} < 0.5$. Which combination and which cuts to use is analysis dependent, and have been optimized for each search presented in this thesis. The details are thoroughly explained in each section.

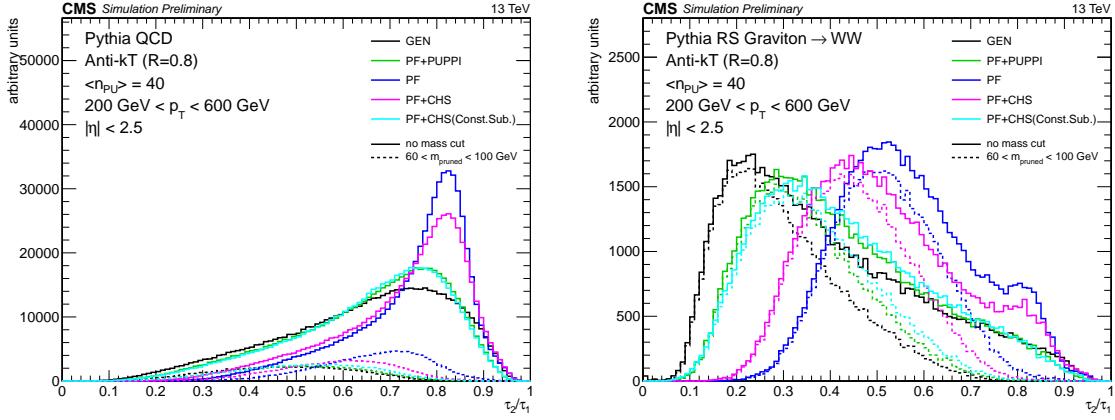


Figure 4.13: The distribution of the n-subjettiness ratio τ_{21} for signal jets (left) and background jets (right) with different combinations of pileup subtraction algorithms applied. The solid lines corresponds to the τ_{21} distribution with no mass cut applied, while the dotted lines are within a mass window of 60–100 GeV [23].

Polarization effects

The vector boson polarization has a significant effect on the W-tagging efficiency. The helicity angle θ , defined as the angle between the outgoing quark daughters of the W in the W bosons rest frame relative to its direction of motion [29], is very different for longitudinally polarized vector bosons, W_L , and transversely polarized vector bosons W_T [30]. Figure 4.14 shows the $\cos \theta^*$ distribution for the outgoing quarks from a $W_L \rightarrow q\bar{q}$ (black) and $W_T \rightarrow q\bar{q}$ (red) decay, and it can be observed that transversely polarized W bosons decay with the quarks emitted closer to the vector boson direction of motion.

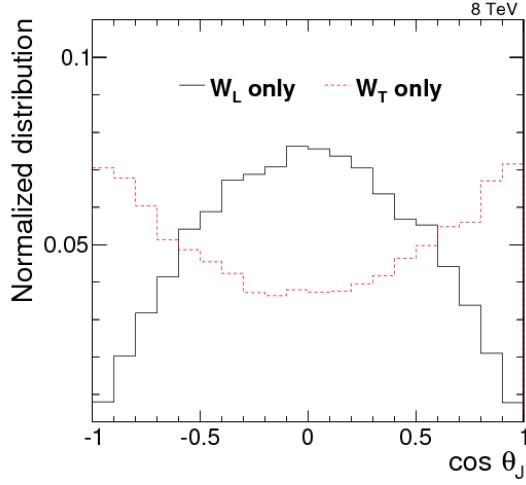


Figure 4.14: The helicity angle for generated quarks from $W_L \rightarrow q\bar{q}$ (black) and $W_T \rightarrow q\bar{q}$ (red) decays [30].

The consequence of this, is that there is a higher asymmetry in the transverse momenta of the two quarks from a W_T decay. This in turn makes grooming algorithms, designed to remove soft constituents of a jet, tend to reject particles coming from the softer quark, resulting in a lower jet mass and a drop in tagging efficiency. Figure 4.15 shows the W-jet

tagging efficiency versus q/g jet mistagging rate for a selection on the jet pruned mass of $60 \text{ GeV} < m_{\text{pruned}} < 100 \text{ GeV}$, scanning τ_{21} cuts (here for CA R=0.8 jets).

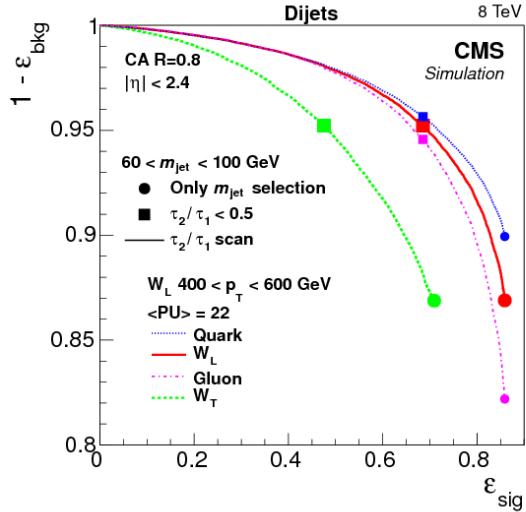


Figure 4.15: The helicity angle for generated quarks from $W_L \rightarrow q\bar{q}$ (black) and $W_T \rightarrow q\bar{q}$ (red) decays [30].

The tagging efficiency for transversely polarized W bosons (green) is significantly lower than the tagging efficiency for longitudinally polarized bosons (red). This can be explained by looking at the $\cos \theta^*$ distribution on reconstructed level, using the C/A subjets, with a cut on the jet pruned mass of $60 \text{ GeV} < m_{\text{pruned}} < 100 \text{ GeV}$, shown in Figure 4.16. When comparing to the distribution at generator level with no groomed mass window applied, Figure 4.14, one can see that the W_T jets with $\cos \theta^* \approx 1$ are completely removed.

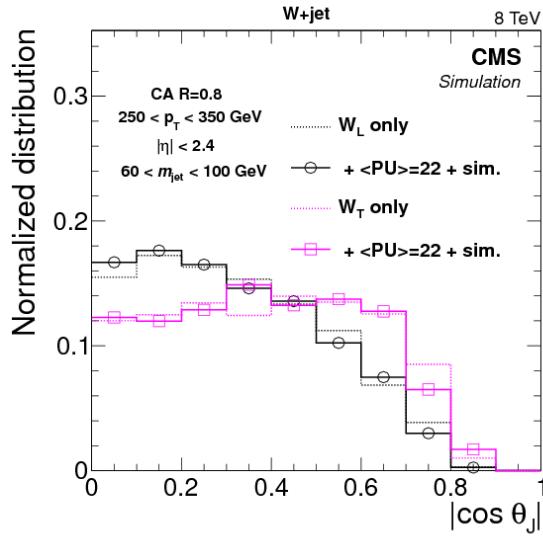


Figure 4.16: The helicity angle for subjets from $W_L \rightarrow q\bar{q}$ (black) and $W_T \rightarrow q\bar{q}$ (pink) decays. [30].

This is due to two effects: the p_T -asymmetry explained above and the fact that the ΔR distribution between the two quarks is much smaller in the case of W_L , making them more

likely to be fully contained within a jet cone of R=0.8.

4.6 Monte Carlo Event Generators

Monte Carlo event generators offer a realistic estimate of high-energy collisions on an event-by-event basis, allowing us to estimate signal and background processes accurately. Simulated events are usually produced in three steps, describing a process from very short timescales up until hadronization and decay. First, a matrix element generator simulates the hard scattering process and subsequent decays. Secondly, the showering and hadronization of unstable particles is performed and, lastly, the final state particles are passed through a full detector simulation in order to reproduce a range of experimental effects.

General-purpose Monte Carlo (GPMC) generators, like HERWIG ++ [31] and PYTHIA 8 [32], deal with both perturbative as well as hadronization phenomena, simulating an event all the way up until detector simulation. In HERWIG ++ and PYTHIA 8, the hardest processes are only simulated at the lowest order of perturbative expansion, meaning $2 \rightarrow 2$ or $2 \rightarrow 3$ scatterings. In order to have tree-level matrix elements with an arbitrary final-state multiplicity, they can be combined with programs used to generate parton-level events at higher accuracy, which are then processed through showering and hadronization with the GPMC generators. One popular program for generating matrix elements is MADGRAPH [33]. This, however, still correspond to a tree-level (leading order) approach. To go to next-to-leading-order (NLO), meaning the inclusion of virtual corrections, two methods exist: MC@NLO [34, 35] and POWHEG [36]. These combine the full next-to-leading-order prediction for inclusive processes with the subsequent parton showers, either by a subtraction method regularizing the real contributions, or by a matrix-element correction of the parton shower branching probability. After hadronization, all final state particles are passed through a full simulation of the CMS detector. This is done with GEANT4 [37], where experimental effects like object reconstruction and detector resolution are accounted for.

For the work presented in this thesis, simulated samples of the Standard Model background processes are used to optimize the analysis and in some cases provide flexible background templates. QCD multijet production is simulated with four generator configurations: 1. PYTHIA standalone, 2. the LO mode of MADGRAPH matched with PYTHIA, 3. POWHEG matched with PYTHIA and 4. HERWIG++ 2.7.1 with tune CUETHS1 [38]. Top quark pair production is modeled with POWHEG and showered with PYTHIA unless otherwise stated. W+jets and Z+jets production are simulated with the leading-order (LO) mode of MADGRAPH matched with PYTHIA. Signal samples are generated with standalone PYTHIA.

All samples are processed through a GEANT4-based simulation of the CMS detector. To simulate the effect of additional proton-proton collisions within the same or adjacent bunch crossings (pileup), additional inelastic events are generated using PYTHIA and superimposed on the hard-scattering events. The MC simulated events are finally weighted to reproduce the distribution of the number of pileup interactions observed in data.

CHAPTER 5

Diboson resonance searches in CMS

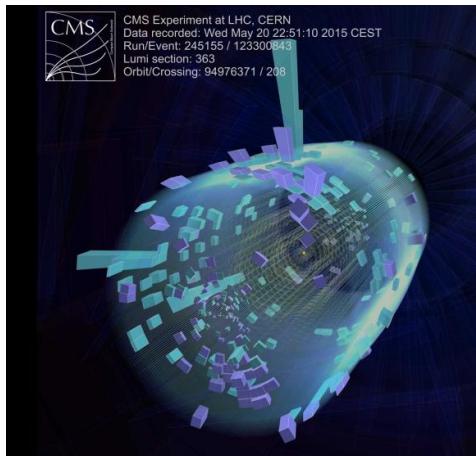
5.1 Search I: First search for diboson resonances at 13 TeV

When the LHC started its Run II data taking period in summer 2015, it would be the first time ever for a particle collider to produce collisions with center-of-mass energies of 13 TeV. The Higgs boson for which the LHC was designed to find had been discovered at the end of the previous data taking era, leaving us with a Standard Model that we know is, in the best case, in need of extensions and, in the worst case, an effective theory valid only in a certain energy domain. The Run II search program would therefore be oriented around two main efforts: Precision measurements of the newly discovered Higgs boson and searches for Beyond Standard Model physics.

I started my PhD four months before the first 13 TeV collisions took place and had to consider the following: What was the most interesting search that could be done on a short time scale (to be presented 6 months after first collisions, at the CERN end-of-year ‘Jamboree’), would be manageable for a student with no previous analysis experience and would be robust enough incase there were issues with the never-before-validated 13 TeV Monte Carlo?

The attention of the high-energy physics community has in the past years been focused on certain ‘hot topics’: In 2018, this was most certainly leptoquarks (driven by a dimuon excess around 30 GeV), in 2016 and 2017 it was diphoton resonances (with $> 3\sigma$ excesses observed both in ATLAS and in CMS). And in 2015 during the 13 TeV LHC start-up, attention was centered on diboson resonances in the all-hadronic final state. The choice was therefore clear: My first analysis would be a search for diboson resonances in the boosted dijet final state. With a background model based on a smooth fit to data in the signal region, eliminating the need for accurate QCD MC predictions, this was a simple one-background only (QCD) analysis, feasible for a first-year PhD student to finalize within a year. Despite its straightforwardness, due to observed 8 TeV excesses, it was in addition considered a high-profile analysis.

This search became one of the first ‘boosted’ searches published with 13 TeV data as well as the first search to take advantage of dedicated ‘grooming’ triggers. Published with 2.7 fb^{-1} of 2015 data.



5.1.1 A small bump

On June 2nd 2015, the day before CMS recorded its first ever 13 TeV event, a pre-print appeared on the arXiv ‘Search for high-mass diboson resonances with boson-tagged jets in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector’ [33]. It was an analysis of the full ATLAS Run 1 dataset, corresponding to 20.3 fb^{-1} , searching for heavy resonances decaying to vector bosons in the all-hadronic state. The analysis documented a 3.4σ excess for a heavy resonance decaying to W Z around 2 TeV. The corresponding CMS analysis, published the previous year, had a 1.3σ excess at roughly the same resonance mass, but mostly compatible with a W W final state hypothesis [39]. Figure 5.1 shows the corresponding dijet invariant mass spectrum as seen by ATLAS (left) and the upper limit on the production times the cross section for a G_{Bulk} decaying to W W (right) as documented by CMS.

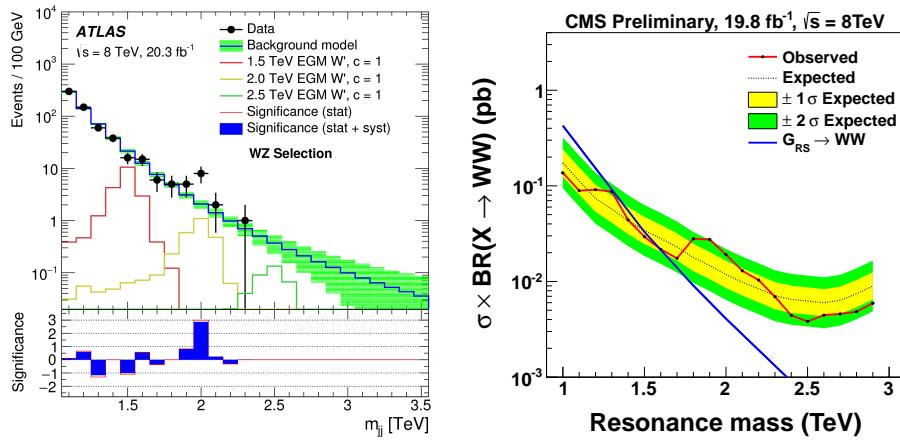


Figure 5.1: A ‘bump’ corresponding to 3.4σ in the dijet invariant mass spectrum around 2 TeV (left) observed by ATLAS when analyzing the full 8 TeV dataset [33], together with a similar excess (1.3σ) observed in the corresponding CMS analysis [39].

The two measurements were found to be compatible, favoring a heavy resonance with a production cross section of around 5 fb^{-1} and a mass between 1.9 and 2.0 TeV decaying to either W W, W Z or Z Z [40]. Figure 5.2 show the obtained p-value of the ATLAS (red) and CMS (blue) search as well as their combination (black).

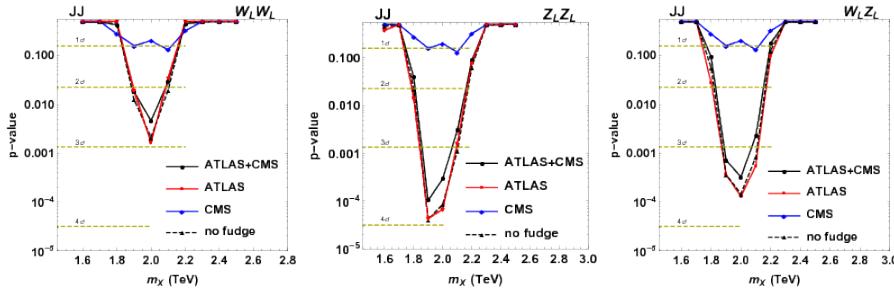


Figure 5.2: p-values as a function of resonance mass obtained with an emulation of the ATLAS (red) and CMS (blue) searches as well as the combination of the two (black). Here for a W W (left), W Z (middle) and Z Z (right) hypothesis [40].

The combination of the two excesses and the timing of the ATLAS paper, naturally lead to some excitement and in the coming weeks, the arXiv was flooded with theory papers seeking an explanation for the measurements.

In addition, one of the main benefits of increasing the LHC center-of-mass energy from 8 to 13 TeV, was that the partonic luminosity would increase. That meant that you could expect them same number of signal events as you would expect for the full 8 TeV dataset ($\sim 20 \text{ fb}^{-1}$), for a considerably smaller 13 TeV dataset. Figure 5.3 shows the system mass that can be probed with 3 fb^{-1} of 13 TeV data (y-axis), the expected 2015 integrated luminosity, as a function of the probe-able system mass with 20 fb^{-1} of 8 TeV data (x-axis) for different partonic channels. The probable 13 TeV mass is defined by finding the system mass which results in the same number of expected events at 8 TeV, if assuming cross sections scale with partonic luminosity and $1/m^2$. Three different partonic scattering channels are considered: qq, qg and gg. We see that, for instance for a resonance with a mass of 2000 GeV, the reach at 13 TeV is 2241(gg), 2091(qq), 1851(qq, one type) and 2046(qq, all types) GeV.

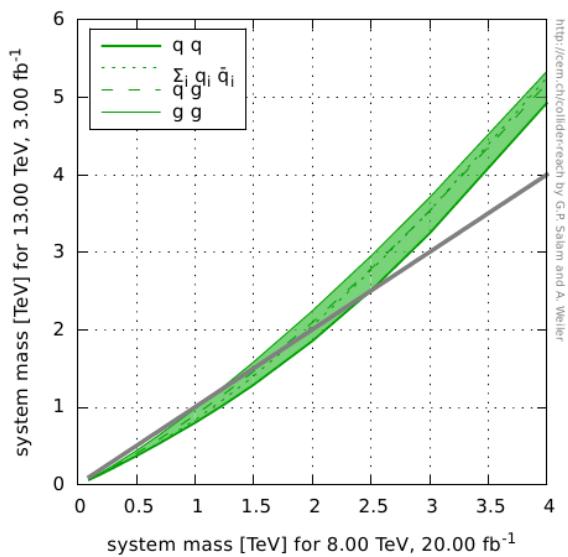


Figure 5.3: The system mass that can be probed with 3 fb^{-1} of 13 TeV data (y-axis) as a function of the probe-able system mass with 20 fb^{-1} of 8 TeV data (x-axis) for different partonic channels (Generated with [41]).

What this meant was that, if we saw hints of a 2 TeV G_{bulk} (mainly produced through gluon fusion, gg) with the 8 TeV dataset, we should be able to confirm it with nothing but the expected 3 fb^{-1} of data expected in 2015. The pressure on seeing early results with 13 TeV data in the VV all-hadronic final state was therefore extremely high, and it was agreed with CMS Physics Coordination that a preliminary analysis would be ready in December that same year, only 6 months after the first 13 TeV collision.

5.1.2 Analysis strategy

When a resonance X with a mass above 1 TeV decays into a vector boson pair, the bosons have a very high energy ($\tilde{p}_T = M_X/2 = 500 \text{ GeV}$, assuming X is produced at rest). The boson is co-called ‘boosted’. The decay products of a hadronically decaying boosted vector boson, will therefore not appear as back-to-back in the lab frame but rather be very collimated, as described in Section 4.5. This results in a final state with two large high- p_T jets, where an AK R=0.8 jet is expected to fully contain the two quarks coming from the vector boson decay. This is illustrated in Figure 5.4.

The two jets are both expected to have a mass around the W or Z mass, and some intrinsic

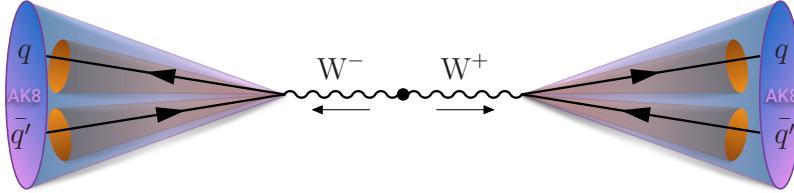


Figure 5.4: If a heavy ($> 1 \text{ TeV}$) resonance decays into vector bosons, the transverse momentum of each boson will be large and its decay products are merged into one single large cone AK8 jet.

substructure stemming from their two-prong origin. The invariant mass of the dijet system, m_{jj} , should be roughly equal to the resonance mass M_x . This dijet system is the final state under scrutiny and the dijet invariant mass is the parameter of interest. Both WW and ZZ, as well as WZ final states are of interest.

The main background for such an analysis, is QCD multijet events. As mentioned in Section 4.5, quark/gluon jets can obtain a high mass due to diffuse radiation and QCD processes have such a large cross section that the number of QCD jets with a mass compatible with the W mass can be large. In order to discriminate between the two, we take advantage of three properties: 1. The groomed mass of signal and background jets should be very different, 2. signal jets should appear two-prong like, quark/gluon jets not, and 3. the dijet invariant mass for a signal process should peak around the resonance mass while the QCD spectrum is predicted to be smoothly falling (we will get back to why this assumption is justified in Section 5.1.5). The strategy therefore consists of performing a smoothness test on m_{jj} of the observed data, a so-called ‘bump-hunt’, by assuming that the signal will appear as a bump on top of a smooth distribution. This is illustrated in Figure 5.5.

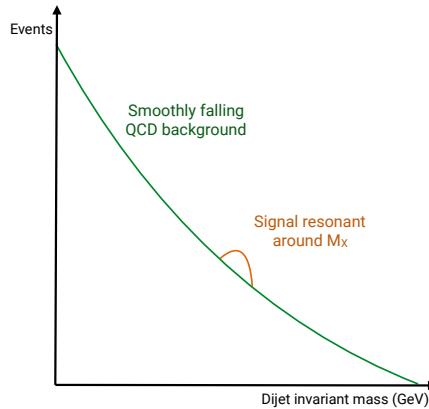


Figure 5.5: The search strategy consists of looking for signal ‘bumps’ in the dijet invariant mass on top of a smoothly falling QCD multijet background.

The benefit of such a method is that there is no need for any background simulation and the strategy is simple and robust. The disadvantage is that the analysis is intrinsically limited to regions where the dijet invariant mass spectrum is smooth, hence must avoid regions with continuities due to trigger turn-ons or kinematic selections.

5.1.3 Data and simulated samples

The data analyzed in this search correspond to a total integrated luminosity of 2.7 fb^{-1} collected at a center-of mass energy of 13 TeV between June and December 2015. The instantaneous luminosity of the LHC during this run was around half of the design luminosity ($0.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$), with an average number of primary vertices per event of $\langle \mu \rangle = 13$.

The bulk graviton model (see Section 1.2.2) and the HVT model (W' and Z' , see Section 1.2.3) are used as benchmark signal processes. In these models, the vector gauge bosons are produced with a longitudinal polarization in more than 99% of the cases, which leads to a 24% higher acceptance per boson for reasons explained in Section 4.5.3. For the HVT model, a scenario (model B) with $g_V = 3$, $c_H = -0.976243$, and $c_F = 1.02433$ is chosen, where the heavy resonance predominantly couple to bosons and the coupling to fermions is suppressed. The bulk graviton samples were generated with $\tilde{k} = 0.5$. The resonance masses considered lie in the range 1.2 to 4 TeV and has a width of 0.1% of the resonance mass. The narrow width allows us to neglect detector effect as the natural width of the resonance is smaller than the detector resolution, making the modeling of detector effects on the signal shape independent of the model. All signal samples are generated at leading order with MADGRAPH5_AMC@NLO v2.2.2 [42]

Simulated samples of the production of QCD multijet events are generated to leading order using PYTHIA version 8.205 [43] with the CUETP8M1 tune [38].

5.1.4 Event selection

Triggering

The first selection to be confronted in any analysis, is the trigger selection. Due to an overwhelming QCD background in all-hadronic final states, the threshold for fully-hadronic triggers is very large in order to keep the trigger rate low (preferably around 10-30 Hertz). In this analysis, we therefore decided to take advantage of triggers that place requirements on the jet groomed mass in addition to the ‘standard’ triggers based on the scalar sum of jet transverse energy H_T . These ‘boosted’ triggers were never before tested in data, and this analysis was the first published result taking advantage of grooming at the trigger level in CMS. The following H_T -based triggers (called inclusive triggers in the following) are used

- HLT_PFHT650_WideJetMJJ900DEtaJJ1p5
- HLT_PFHT650_WideJetMJJ950DEtaJJ1p5,
- HLT_PFHT800

where the two first triggers apply an additional cut on the $|\Delta\eta|$ between the two jets for reasons that will be explained below. In addition, two grooming triggers cutting on the jet trimmed mass (see Section 4.5.1) of 30 and 50 GeV are used

- HLT_AK8PFJet360_TrimMass30
- HLT_AK8PFHT700_TrimR0p1PT0p03Mass50.

The tuneable parameters for the trimming algorithm at HLT are $r_{sub} = 0.2$ and $p_{T,frac} = 0.03$. The HLT_AK8PFJet360_TrimMass30 trigger is seeded by single-object Level 1 triggers with jet p_T thresholds of 176 or 200 GeV (L1_SingleJet176 or L1_SingleJet200), and the remaining triggers requires an online $H_T > 150$ or 175 GeV (L1_HTT150 or L1_HTT175).

In order to avoid any kinks in the dijet invariant mass spectrum due to the presence of a trigger turn-on, we need to define for which dijet invariant mass the analysis triggers are fully efficient (> 99%), then cut away everything below that point.

In order to estimate the trigger efficiency, we use a lower threshold prescaled H_T trigger `HLT_PFHT650` as reference trigger. This trigger has a prescale of 40, meaning it only stores information for every 40 events that trigger it, and is seeded by L1 triggers `L1_HTT150` or `L1_HTT175`. We then define the efficiency as

$$\text{Efficiency} = \frac{N_{\text{trigger+ref}}}{N_{\text{ref}}}$$

where $N_{\text{trigger+ref}}$ corresponds to events passing the trigger under study as well as the reference trigger and N_{ref} corresponds to all events passing the reference trigger. Figure 5.6 shows the trigger turn-on curves as a function of dijet invariant mass for jets where one of the jets is required to have a pruned mass larger than 65 GeV (in other words, compatible with a W jet). A sharp turn-on for the inclusive triggers (top left) is observed, reaching the 100% efficiency plateau for dijet masses of around 1.0–1.1 TeV. The grooming triggers, however, turn on more slowly and are not fully efficient before dijet invariant masses of around 1.2 TeV (top right). The real power of the grooming triggers become clear when adding them in OR with the H_T -based triggers. The bottom plot in Figure 5.6 compares the trigger turn-on curves as a function of dijet invariant mass for jets passing one of the three inclusive triggers only, one of the grooming triggers only and when combining all of them. Here, one can see that the 99% efficiency threshold is lowered by 75 GeV when including the substructure triggers, once substructure is required at analysis level. This allowed for the analysis to start at a dijet invariant mass of 1 TeV.

As a measure of the performance of the grooming triggers, we have in addition looked at the trigger efficiencies as a function of the offline groomed mass (pruned and softdrop, see Sections 4.5.1 and 4.5.1), for the grooming trigger with the lowest mass threshold (30 GeV). This is shown in Figure 5.7, where an additional cut on the jet transverse momentum of one of the jets of 600 GeV is required and no other mass cut is applied. The trigger plateau is reached for offline groomed-jet masses around 50 GeV, an impressively sharp turn-on for a trigger paths first test in data (as reference trigger for this study, the prescaled trigger `HLT_PFTJet320` was used).

Preselection

After trigger selections and the corresponding requirement of a dijet invariant mass above 1 TeV to ensure a smooth falling background, the process of maximizing the signal significance while keeping the background low can begin. This is done through a set of jet requirements. The jets used in this analysis are clustered with the anti- k_T jet clustering algorithm with a clustering parameter of $R = 0.8$ (see Section 4.4) to allow containment of the full vector boson decay products. As we know that a minimum transverse of 200 GeV is required for the decay products of a W/Z to be fully contained within an $R=0.8$ jet, events are further selected by requiring at least two jets with $p_T > 200$ GeV. These are in addition required to be central, with an $|\eta| < 2.4$.

The two highest p_T jets in the event passing these criteria are selected as potential vector boson candidates. As our main background is QCD multijet events, we further take advantage of the fact that the angular distribution between these, mainly t-channel, processes are very different from the s-channel signal processes under study. The crosssection for QCD t-channel processes as a function of the opening angle with respect to the beam axis (θ^*), exhibit a

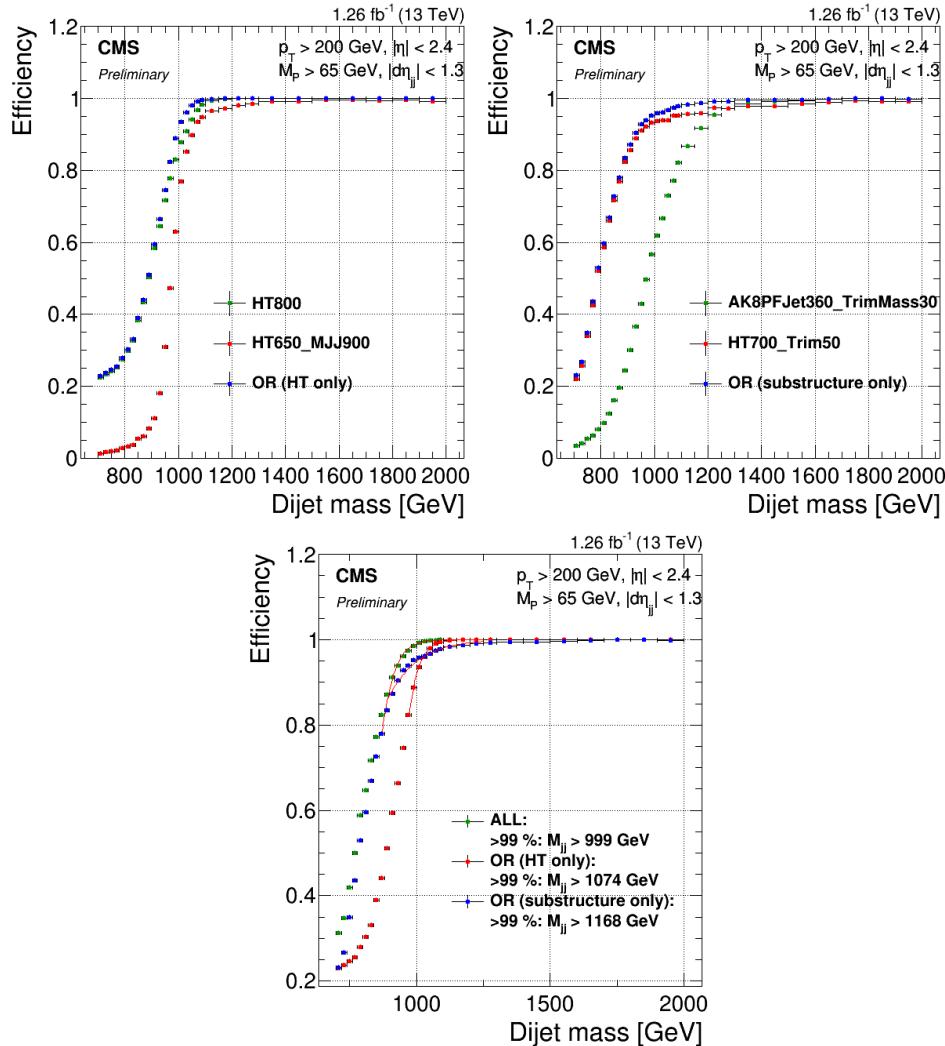


Figure 5.6: Top: Efficiency for the inclusive triggers (top left) and the grooming triggers (top right) as a function of dijet invariant mass for jet pairs where one jet has a pruned mass larger than 65 GeV. Bottom: Comparison of trigger efficiencies for jets passing one of the HT-triggers only (red), for jets passing one of the grooming-triggers only (blue) and for jets passing one of the HT-triggers or one of the grooming triggers (green). Here as a function of dijet invariant mass for all jet pairs passing loose selections and where one jet has a pruned mass larger than 65 GeV. The 99% efficiency threshold is lowered by 75 GeV when including substructure taggers.

pole around $\cos \theta^* = 1$, meaning QCD t-channel jets are mostly forwardly produced, with an opening angle with respect to the beam axis close to zero. The signal jets on the other hand, produced through an s-channel process, are concentrated in the barrel region. We therefore require the jets to have a separation of $|\Delta\eta| < 1.3$ in order to reduce the QCD multijets background. The distribution of $|\Delta\eta|$ between the two highest- p_T jets for QCD as well as for different signal scenarios, is shown in Figure 5.8

A cut of $|\Delta\eta|_{jj} < 1.3$ makes sure to remove the t-channel pole at $\cos \theta^* = 1$ and is in addition found to yield the best separation between signal and the QCD background.

In addition to these requirements on the jets themselves, an overlap veto with leptons in the event is applied. Here the overlap $\Delta R(\text{jet}, \text{lepton})$ between the jet candidate and a lepton is required to be larger than 0.8. Leptons used for this veto are required to pass the identification requirements described in Section 4.2.2 and 4.2.2, have a transverse momentum

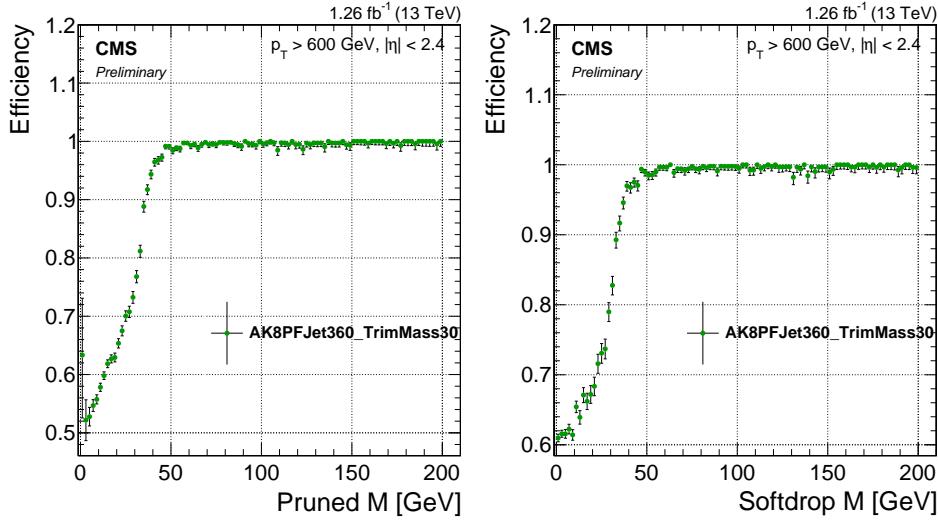


Figure 5.7: Efficiency for the lowest threshold grooming trigger as a function of pruned-jet (left) and softdrop-jet (right) mass for jets with $p_T > 600 \text{ GeV}$.

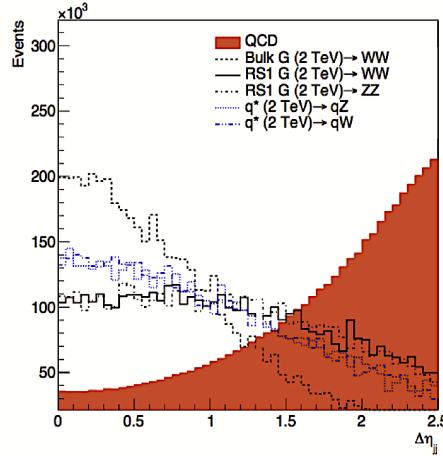


Figure 5.8: $|\Delta\eta|$ between the two highest- p_T jets for QCD jets and jets stemming from different signal scenarios.

larger than 35 (30) GeV and a pseudorapidity smaller than 2.5 (2.4) in case of electrons (muons).

The p_T , η , dijet invariant mass and $|\Delta\eta|_{jj}$ distribution for the two leading jets in the event after the above preselections have been applied is shown in Figure 5.67.

Vector boson tagging

After preselections, we take advantage of the jet substructure algorithms described in Section 4.5 to further separate boosted W/Z jets from the QCD multijets background. In the 8 TeV analysis [39] published the previous year, the pruning algorithm was the groomer of choice. However, recent progress had been made in the development of alternative groomers which had favorable properties from a theoretical point of view (see Sections 4.5.1 and 5.2.5). We therefore studied two different grooming algorithms: pruning and softdrop (with $\beta = 0$ and $z_{cut} = 0.1$). A comparison of the softdrop (dotted lines) and pruned (solid lines) jet mass

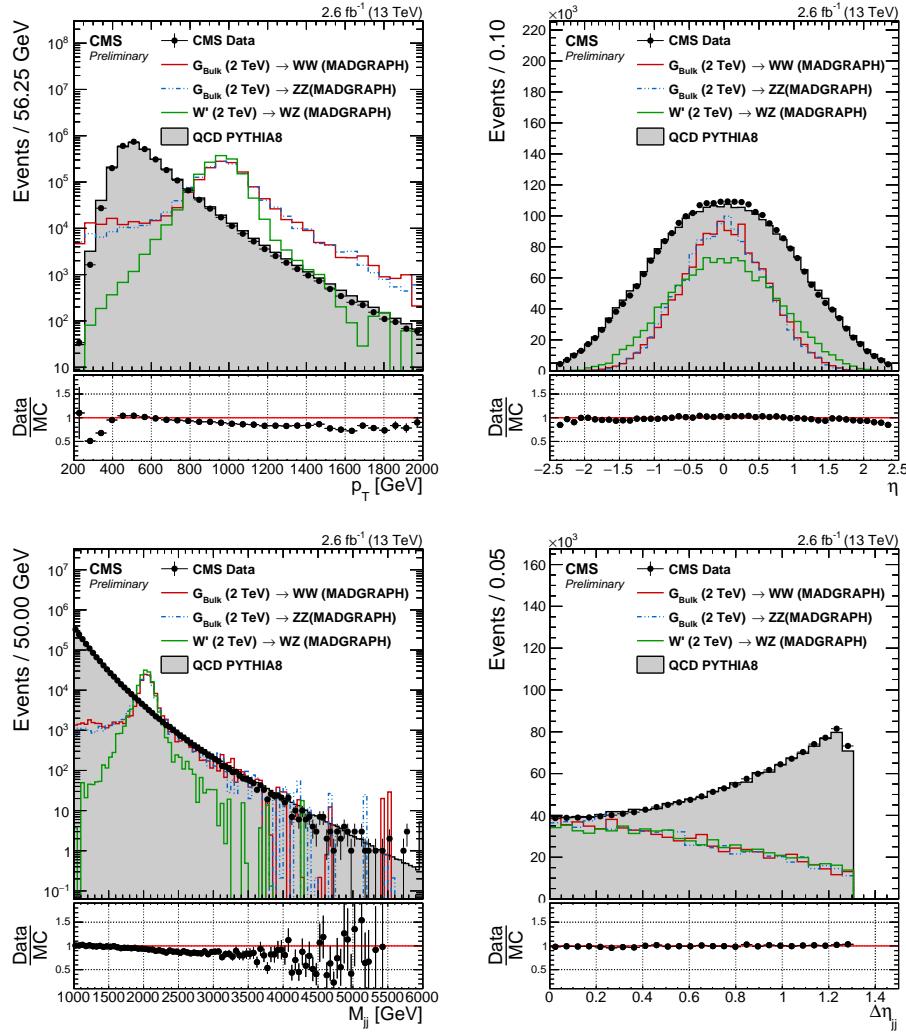


Figure 5.9: Jet p_T (top left), η (top right), dijet invariant mass (bottom left) and $|\Delta\eta|_{jj}$ (bottom right) distribution for the two leading jets in the event after loose preselections are applied. The signal is scaled by an arbitrary number.

for W, Z and H jets is shown in Figure 5.10.

One of the first observations we made comparing the two groomers, was that there appeared to be a strong dependence of softdrop mass on the jet p_T . Figure 5.11 shows the pruned (left) and softdrop (right) mass distributions for W jets coming from the decay of a G_{bulk} with a resonance mass of $0.8 \text{ TeV} < M_X < 4 \text{ TeV}$. While the pruned jet mass mean appeared stable as the jet transverse momenta of the jet increased ($p_T \sim M_X/2$), the softdrop jet mass mean shifted towards lower values as jet p_T increased.

In order to investigate whether this was a reconstruction effect or an algorithmic effect, we additionally looked at the pruned and softdrop mass for generator level jets (jets clustered with generator level particles before they are passed through the detector simulation). Figure 5.12 shows the reconstructed (solid line) and generator level (dotted line) jet mass distributions after pruning (left) or softdrop (right) have been applied. Again, the distributions are compared for jets with very different p_T profiles, here for W jets coming from a $G_{\text{bulk}} \rightarrow WW$ of mass $M_X = 0.8 \text{ TeV}$ (red), roughly $p_T \sim 400 \text{ GeV}$, and $M_X = 2.0 \text{ TeV}$ (blue), $p_T \sim 1 \text{ TeV}$. Interestingly, we observe a p_T -dependent mass shift already for generator level softdrop jets

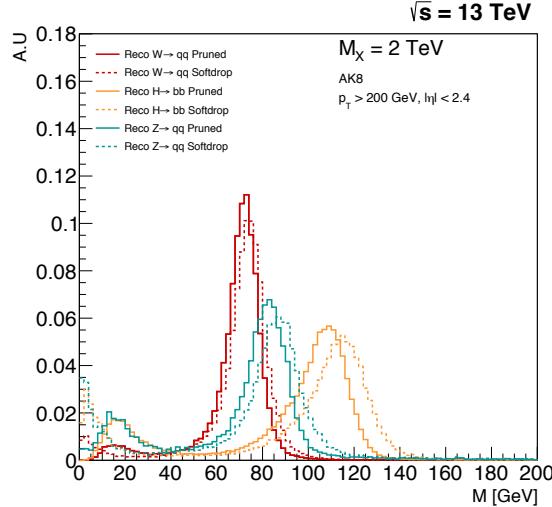


Figure 5.10: The softdrop (dotted lines) and the pruned (solid lines) jet mass for W, Z and H jets.

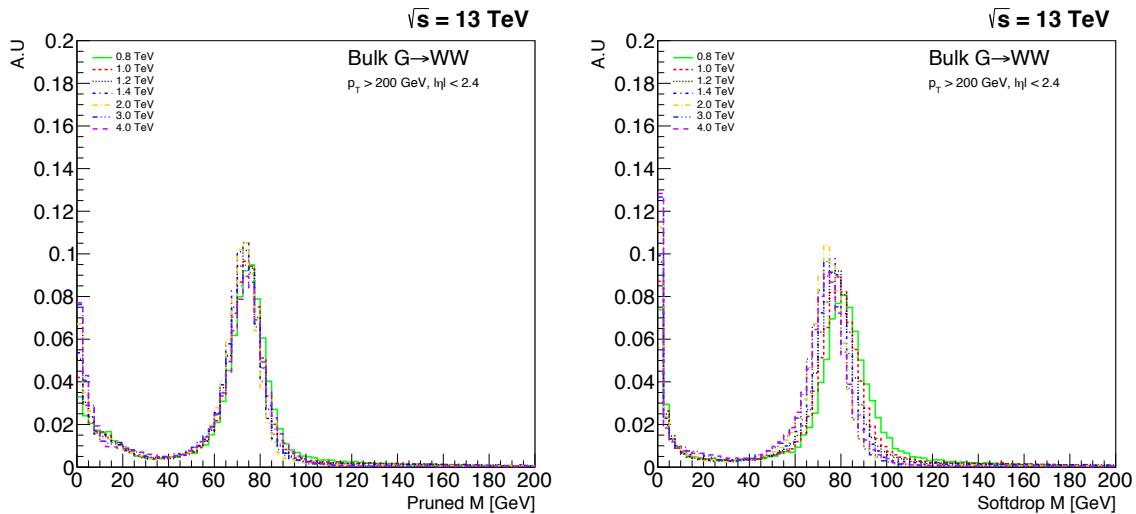


Figure 5.11: The jet mass distribution for W jets coming from a G_{bulk} of masses in the range $0.8 \text{ TeV} < M_X < 4 \text{ TeV}$ decaying to WW, here with pruning applied (left) and softdrop (right). A strong shift in the jet mass mean as a function of p_T ($\sim M_X/2$), is observed for jets groomed with the softdrop algorithm. Charge hadron subtraction is applied to all jets before clustering.

(comparing the dotted lines in the right plot); an effect further enhanced at reconstruction level. This effect is not present for pruned jets, neither at generator level nor reconstruction level.

The observed softdrop mass p_T -dependence was problematic, due to the fact that it would require a p_T dependent mass window. This would again require several different measurements of data to simulation tagging efficiency scale factors, for the respective mass windows, or a significantly higher uncertainty on the signal yield. Due to these observations, the grooming algorithm of choice for this analysis is pruning, with $65 \text{ GeV} < m_{\text{pruned}} < 105 \text{ GeV}$. However, this would be a study we would return to in Search II (Section 5.2).

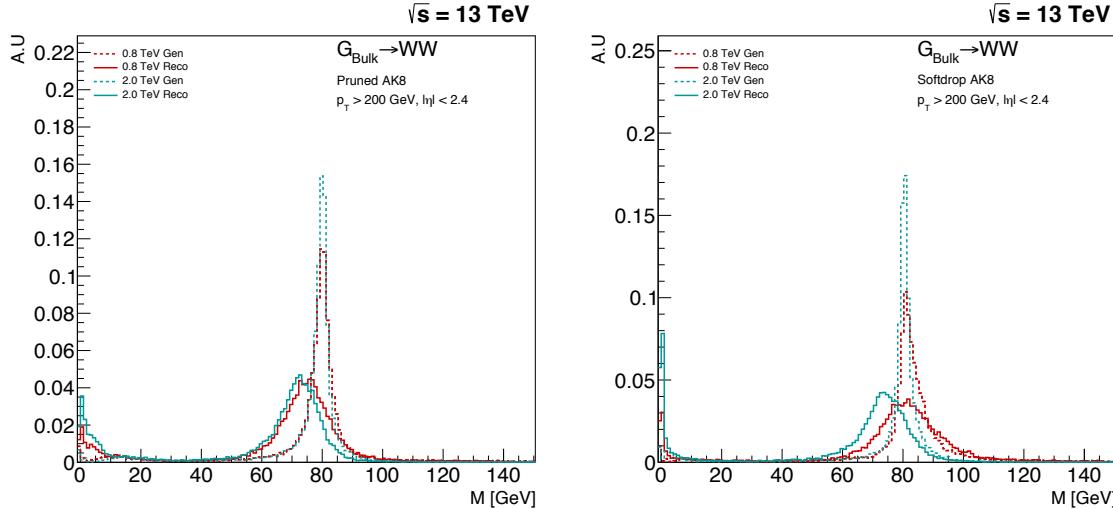


Figure 5.12: The reconstructed (solid line) and generator level (dotted line) jet mass distribution for W jets coming from a $G_{\text{bulk}} \rightarrow WW$ of mass $M_X = 0.8 \text{ TeV}$ (red), roughly $p_T \sim 400 \text{ GeV}$, and $M_X = 2.0 \text{ TeV}$ (blue), $p_T \sim 1 \text{ TeV}$. Here for the pruned (left) and softdrop (right) jet mass.

The shape tagger we chose for this analysis was the n-subjettiness ratio τ_{21} . τ_{21} is strongly correlated to the pruned jet mass, and the discriminating power of the variable is reduced when applying a pruned mass cut. The τ_{21} distribution for the QCD background and W jets from a signal decay before (left) and after (right) a pruned mass cut of $65 \text{ GeV} < m_{\text{pruned}} < 105 \text{ GeV}$ have been applied, is shown in Figure 5.13.

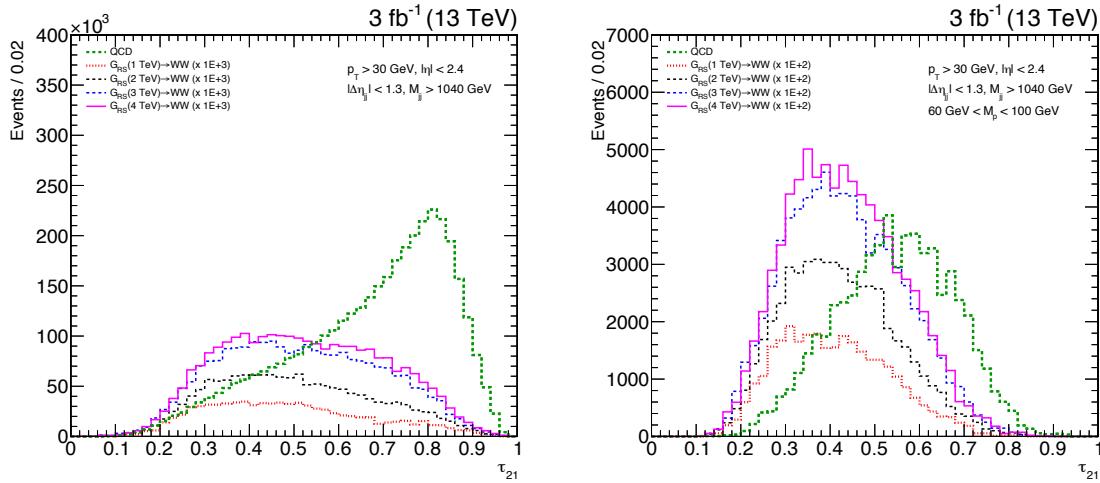


Figure 5.13: The τ_{21} distribution for QCD background and signal jets before (left) and after (right) a pruned mass window is applied. The discriminating power of τ_{21} is strongly reduced after grooming.

We therefore perform a cut optimization on τ_{21} after all analysis selections, including a pruned mass window of $65 \text{ GeV} < m_{\text{pruned}} < 105 \text{ GeV}$, have been applied. This is done by scanning the τ_{21} cut, and for each cut computing the Punzi significance [44] defined as

$$S = \frac{\epsilon_S}{1 + \sqrt{B}}$$

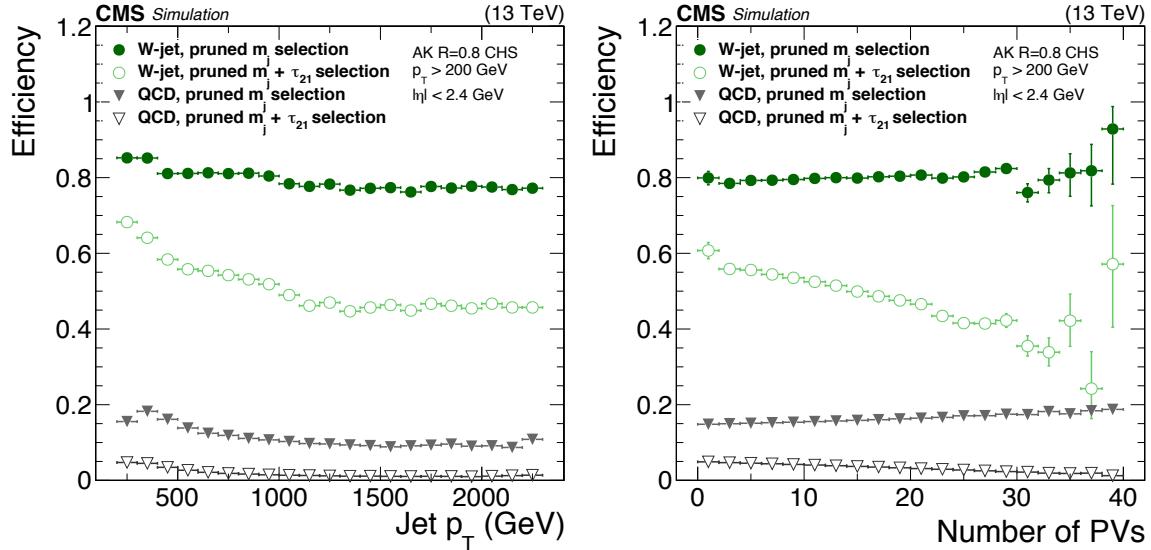


Figure 5.15: The W-tagging efficiency (green) and light jet mistag rate (grey) for a pruned jet mass cut only and pruned jet mass + τ_{21} cut as a function of p_T (left) and number of primary vertices (right).

and the n-subjettiness τ_{21} distribution (right) for signal and background Monte Carlo, as well as the distributions measured in data.

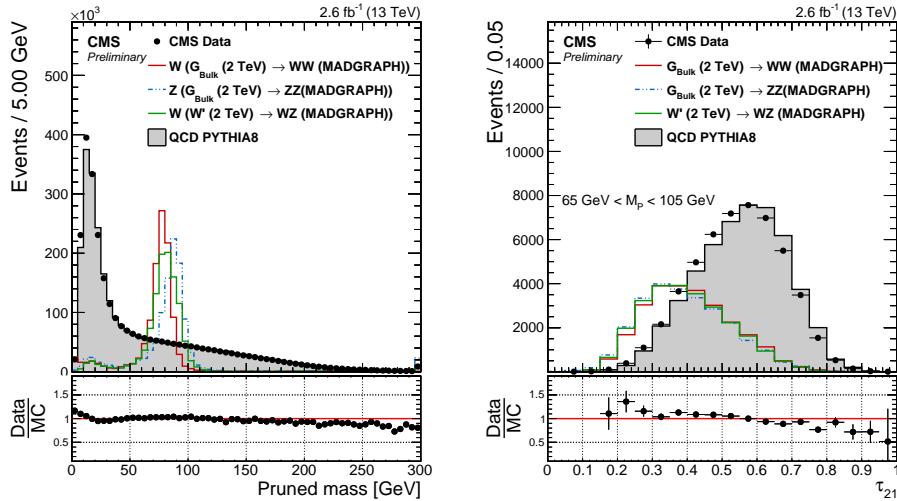


Figure 5.16: Pruned jet mass distribution (left) and n-subjettiness τ_{21} (right) distribution for data and simulated samples. Simulated samples are scaled to match the distribution in data. The τ_{21} distribution is shown for jets after a cut of $65 \text{ GeV} < M_p < 105 \text{ GeV}$ has been applied.

Analysis categorization

As the analysis requires two W/Z-tags, we always require one HP tagged jet and then divide into LP and HP categories depending on whether the other jet is of high or low purity. In addition, in order to further enhance the analysis sensitivity, we further split the pruned jet mass window into a W and a Z boson window where the W window is defined as

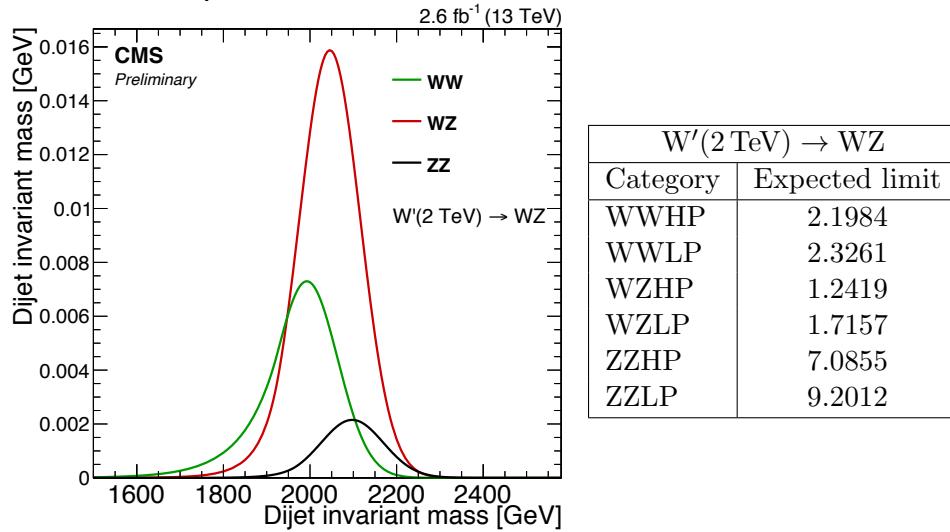


Figure 5.17: The expected signal yield per mass category for W' (2 TeV) decaying to a W and Z (left) together with the expected limit per mass category for the same signal (right).

$65 \text{ GeV} < m_{\text{pruned}} < 85 \text{ GeV}$ and the Z boson window as $85 \text{ GeV} < m_{\text{pruned}} < 105 \text{ GeV}$. This has the added benefit of allowing us to make a discrimination between a G_{bulk} decaying to WW or ZZ, and a W' decaying into WZ through counting events in each category. We, for instance, expect a higher signal yield in WZ category for a W' decaying to a W and Z boson than for a G_{bulk} decaying to WW or ZZ. Figure 5.17 shows the relative expected signal yield (left) and expected limits (left) in the different mass categories for a 2 TeV W' .

All categories are combined in the end, leading to the same or better sensitivity than when using the whole pruned mass window. Figure 5.18 shows the expected 95% CL upper limits on the production cross section of a W' decaying to WZ (left) and a G_{bulk} decaying to WW (right) as function of the resonance mass in the HP category. The blue line corresponds to the expected limits obtained when not splitting into mass categories and the red line corresponds to the limit using the combination of two categories. The dotted and solid black lines are the limits in the W and Z categories, respectively. The combination of two mass categories leads to a slightly better (10%) or to the same sensitivity as when using one large mass window.

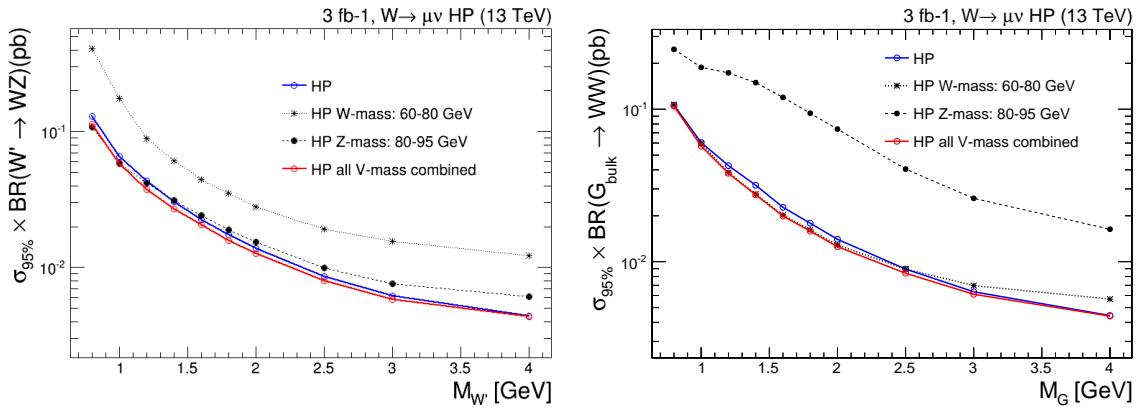


Figure 5.18: Expected 95% CL upper limits on the production cross section of a W' (left) and G_{bulk} (right) signal as function of the resonance mass for the different mass categories for events passing the high-purity τ_{21} selections.

The real benefit of splitting into mass categories becomes obvious when defining a test statistics based on the likelihood ratios of each signal hypothesis, $q = -2 \ln(L_{G_{\text{bulk}}}/L_{W'})$, shown in Figure 5.19. For a signal with a signal strength corresponding to a $3\text{-}4\sigma$ excess, the test statistics for each signal hypothesis are well separated ($\sim 3.5\sigma$), allowing us to make a statement of how G_{bulk} or W' like a possible signal is. With the high-purity and low-purity

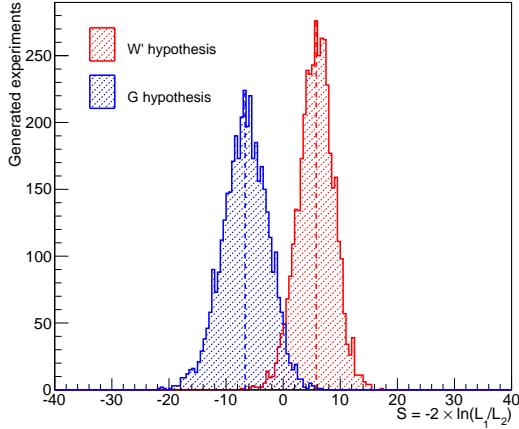


Figure 5.19: Distribution of the test statistic $q = -2 \ln(L_{G_{\text{bulk}}}/L_{W'})$ for a G_{bulk} (blue) and W' signal hypothesis.

categories as defined above for each mass window combination, this leaves us with six different signal categories. They are as follows:

- High-purity, 3 mass categories: WW, ZZ and WZ
- Low-purity , 3 mass categories: WW, ZZ and WZ

In parallel to the mass-category based analysis, we perform an analysis without categorization in mass (similar to the 8 TeV analysis) as a cross-check. These studies can be found in Appendix C. The final tagging efficiency for different signal hypothesis (top) together with the QCD mistag rate (bottom) in the different signal categories is shown in Figure 5.20. The solid lines represent the tagging efficiency in the full mass window ($65 \text{ GeV} < M_p < 105 \text{ GeV}$) before splitting into mass categories. A lower signal efficiency the ZZ mass category is observed in all cases. This can be explained from the pruned jet mass distribution on the left in Figure 5.1.7, where a cut at 85 GeV leaves a large fraction of the Z peak in the W mass window. As the main benchmark models under consideration preferably decays to W bosons (in the Bulk Graviton model the branching ratio $\text{BR}(G_{\text{Bulk}} \rightarrow W W) = 2 * \text{BR}(G_{\text{Bulk}} \rightarrow ZZ)$, and in the HVT model $W'/Z' \rightarrow WZ/WW$ (but not ZZ)), a high tagging efficiency for the W boson is preferred. In the limit-setting procedure all the categories are combined and the overall signal efficiency is conserved. For the combined mass-categories (solid line) the signal efficiency is between 16 and 23 % in the double-tag categories, and between 20 and 34 % in the single-V tag categories. The mistagging rate in the double-V tag categories is below 1 % in the high-purity category. The full analysis selections and final categories are listed in Table 5.1.

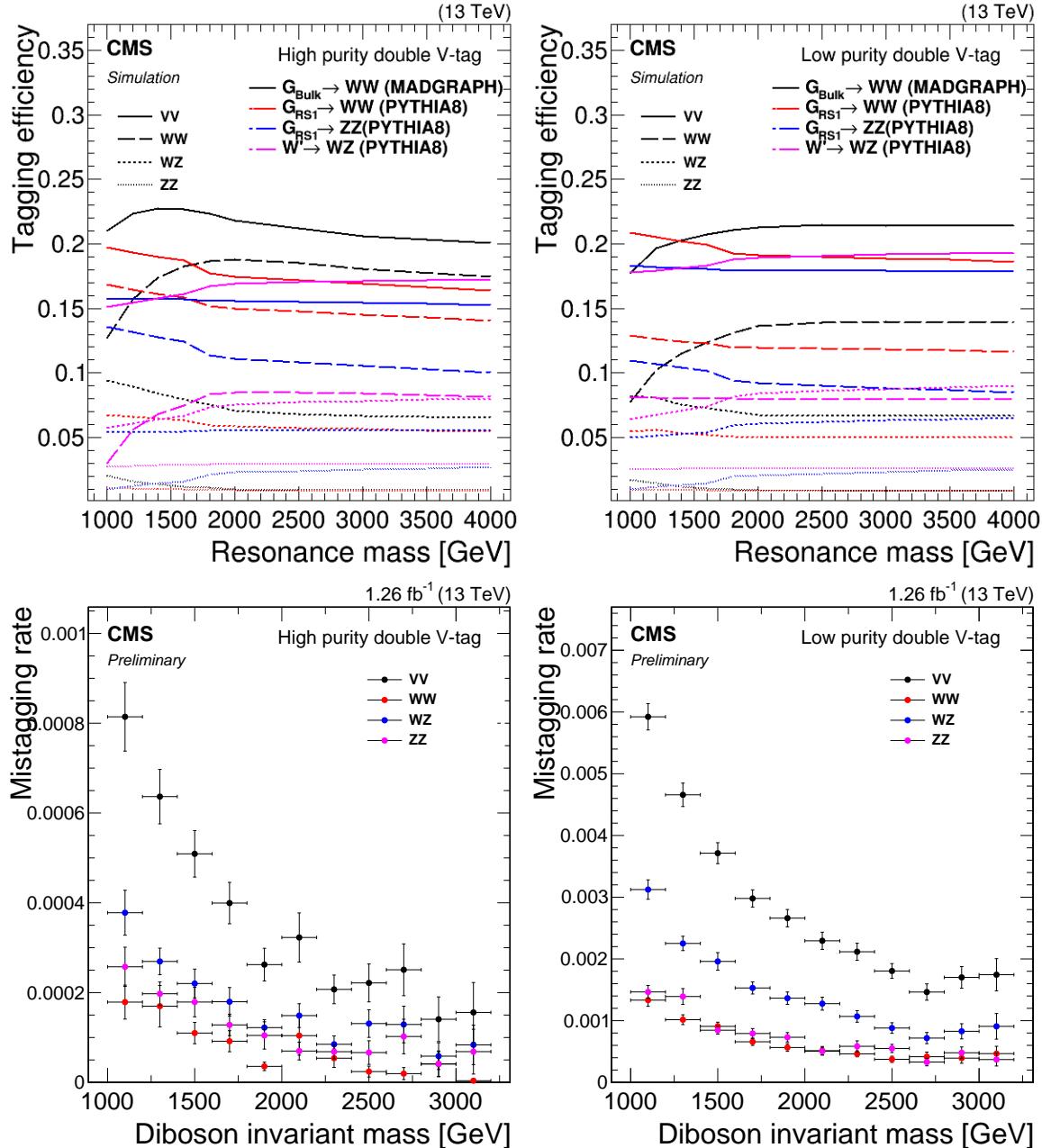


Figure 5.20: Tagging efficiency (top) and mistagging rate (bottom) in the different pruned mass categories in the high-purity category (left) and in the low-purity category (right)

Selection	Value
<hr/>	
Boson selections	
$V \rightarrow q\bar{q}$ (2 AK8 jets)	$p_T > 200 \text{ GeV}$ $ \eta < 2.4$
Pruned jet mass	$65 < m_{\text{jet}_1}, m_{\text{jet}_2} < 105 \text{ GeV}$
Topology	$ \Delta\eta_{jj} < 1.3$
Dijet invariant mass	$m_{jj} > 1 \text{ TeV}$
2- to 1-subjettiness ratio	$\tau_{21} < 0.75$
<hr/>	
<i>m_{jet}</i> categories	
WW	$65 < m_{\text{jet}_1} < 85 \text{ GeV}, 65 < m_{\text{jet}_2} < 85 \text{ GeV}$
WZ	$65 < m_{\text{jet}_1} < 85 \text{ GeV}, 85 < m_{\text{jet}_2} < 105 \text{ GeV}$
ZZ	$85 < m_{\text{jet}_1} < 105 \text{ GeV}, 85 < m_{\text{jet}_2} < 105 \text{ GeV}$
<hr/>	
τ_{21} categories	
High-purity	$\tau_{21,\text{jet}1} < 0.45, \tau_{21,\text{jet}2} < 0.45$
Low-purity	$\tau_{21,\text{jet}1} < 0.45, 0.45 < \tau_{21,\text{jet}2} < 0.75$

Table 5.1: The full analysis selections, mass and τ_{21} categories.

5.1.5 Background modeling

The background modeling in this analysis is based on a smoothness test performed directly on unblinded data, similar to what is done in previous CMS analyses looking for bumps in the dijet invariant mass spectrum [45, 46]. We assume that the QCD multijet background in the different analysis categories can be described by smooth, monotonically decreasing functions of 2 or 3 parameters

$$\frac{dN}{dm_{jj}} = \frac{P_0}{(m_{jj}/\sqrt{s})^{P_2}} \quad \text{and} \quad \frac{dN}{dm_{jj}} = \frac{P_0(1 - m_{jj}/\sqrt{s})^{P_1}}{(m_{jj}/\sqrt{s})^{P_2}} , \quad (5.1)$$

where m is the dijet invariant mass, \sqrt{s} the centre of mass energy and P_0 is a normalization parameter for the probability density function and P_1 and P_2 describe the shape. The number of fit parameters is decided through a Fishers F-test [47]. In this test, we start from the 2 parameter function and compare the goodness of fit (χ^2 divided by degrees of freedom) when fitting the data signal region with a 2, 3, 4 and 5 parameter function. We then check at 10% confidence level (CL) if additional parameters are needed to model the background distribution. The 4 and 5 parameter functions are

$$\frac{dN}{dm_{jj}} = \frac{P_0(1 - m_{jj}/\sqrt{s})^{P_1}}{(m_{jj}/\sqrt{s})^{P_2 + P_3 \times \log(m_{jj}/\sqrt{s})}} \quad (5.2)$$

$$\frac{dN}{dm_{jj}} = \frac{P_0(1 - m_{jj}/\sqrt{s})^{P_1}}{(m_{jj}/\sqrt{s})^{P_2 + P_3 \times \log(m_{jj}/\sqrt{s}) + P_4 \times \log(m_{jj}/\sqrt{s})^2}} \quad (5.3)$$

where P_3 and P_4 are additional free parameters. As an additional cross check, an alternative fit function is also tested:

$$\frac{dN}{dm_{jj}} = \frac{P_0(1 - m_{jj}/\sqrt{s} + P_3(m_{jj}/\sqrt{s})^2)^{P_1}}{(m_{jj}/\sqrt{s})^{P_2}} . \quad (5.4)$$

The fit range is chosen such that it starts where the trigger efficiency has reached its plateau to avoid bias from trigger inefficiency, and extends to the bin after the highest m_{VV} mass point. The binning chosen for the fit follows the detector resolution as in [45, 46]. Before unblinding the signal region, we check that the QCD dijet invariant mass spectrum is expected to be smooth from the distribution in QCD MC as well as exercise the F-test in QCD MC and in a data sideband.

The fits to data in the signal region using the different fit functions, are shown in Figure 5.21, and the corresponding F-test output are given in Table 5.2 through Table 5.4. The findings can be summarized as follows: for the WW enriched category a 2 parameter fit is sufficient to describe the data in both the high- and low-purity categories. In the WZ category, a two parameter fit is sufficient in the high-purity category, while three parameters are needed for the low-purity category. For the ZZ category, a 3 parameter fit is needed for both purity categories. The 2 and 3 parameters fit functions as defined in Equation D.2 will therefore be used to model the background component in the simultaneous signal and background fit.

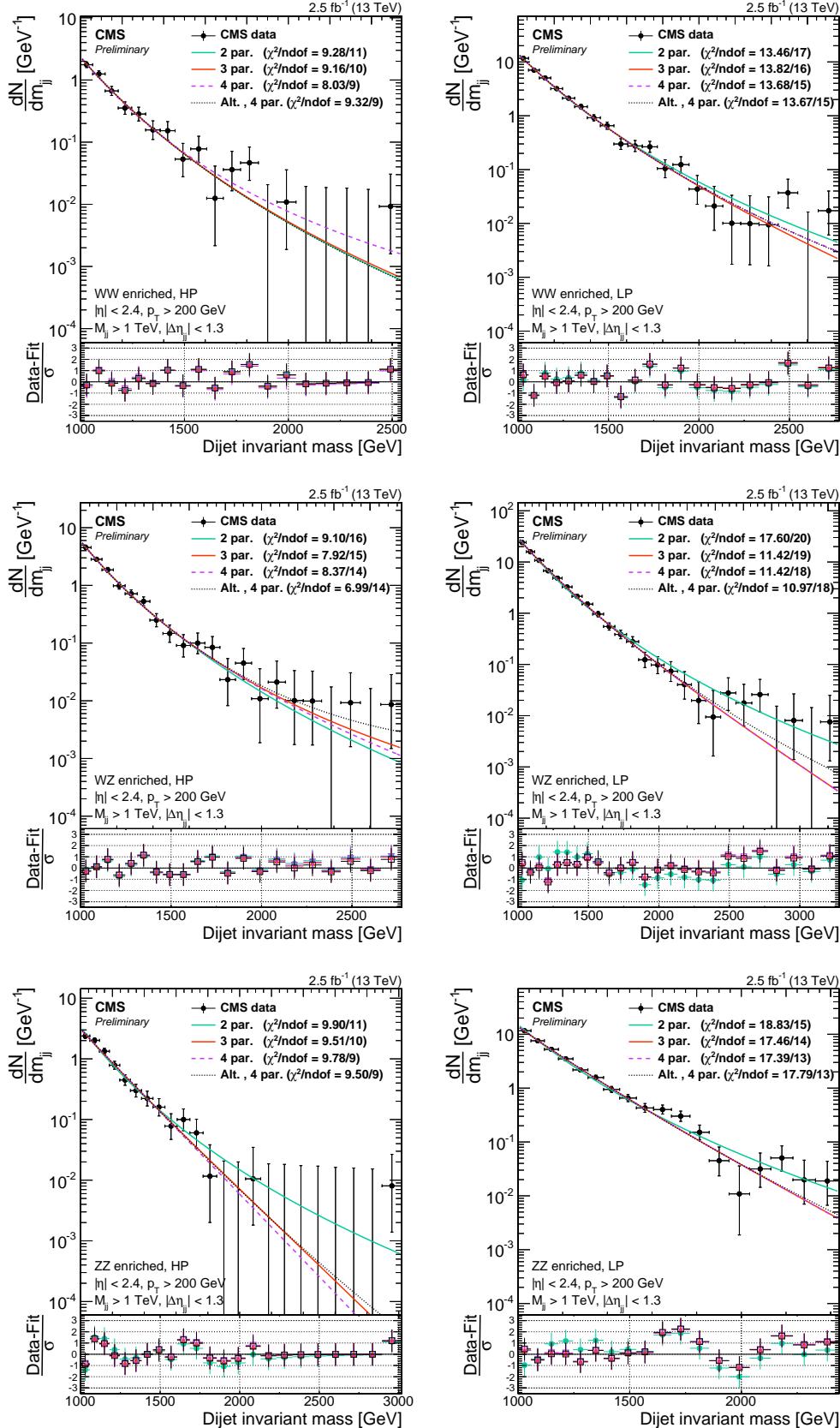


Figure 5.21: Fitted dijet mass spectrum in the different mass and purity categories in data for the double V-tag category. A 2 parameter fit is sufficient to describe the data for the WW (HP and LP) and WZ (LP) enriched categories. For the ZZ enriched (HP and LP) and WZ (HP) categories, a 3 parameter fit is needed.

5.1.6 Signal modeling

The signal shape is extracted from signal MC with masses in the range from 1 to 4 TeV. A linear interpolation provides shapes for the mass points in between in steps of 100 GeV. From these shapes, pdf models are constructed as composite models with a Gaussian core due to detector resolution and an exponential tail to account for parton distribution function effects. Parametric shape uncertainties due to jet energy scale and resolution uncertainties are inserted by variations of the Gaussian peak position and width. The dijet invariant mass shape for different benchmark model signals are shown in Figure 5.22. The signal and background components are then simultaneously fitted to the data points.

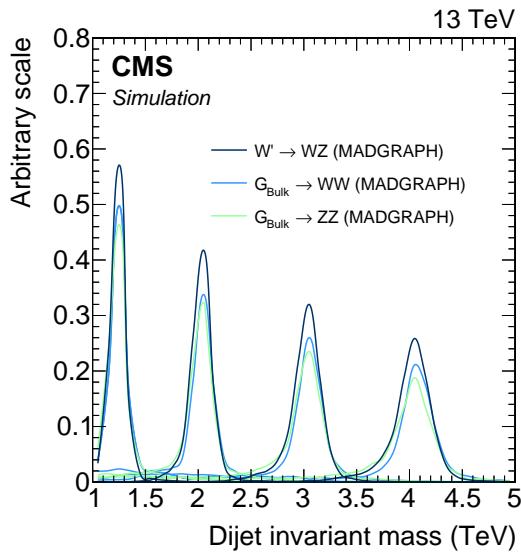


Figure 5.22: Dijet invariant mass from signal MC used to extract the signal shape. Here for 1.2, 2, 3 and 4 TeV resonances.

5.1.7 V-tagging scale factors

As seen in Figure , some discrepancy is observed in the τ_{21} distribution between data and MC. This can lead to a bias in the signal efficiency estimation and we therefore measure the real data signal efficiency in an orthogonal data sample. The W-tagging efficiency is measured using real boosted W-jets in a semi-leptonic $t\bar{t}$ enriched data sample. This region is mainly quark-enriched, as opposed to the QCD gluon-enriched region we saw previously, and substructure variables are better described here. The sample is obtained through requiring a final state compatible with two b-jets and two W bosons, where one of the bosons decay leptonically and the other one hadronically. There are several good reasons to use this channel: Top quark pair production events are plentifully produced at the LHC, we can ensure a high purity of the sample through high-energy lepton, b-tag and missing energy requirements and lastly we can ensure that the W jets are boosted by requiring the leptonic leg, together with the hadronic W candidate, to have high transverse momentum. The final state is illustrated in Figure 5.23, with the object of interest being the AK R=0.8 jet containing the two quark daughters of the hadronically decaying W.

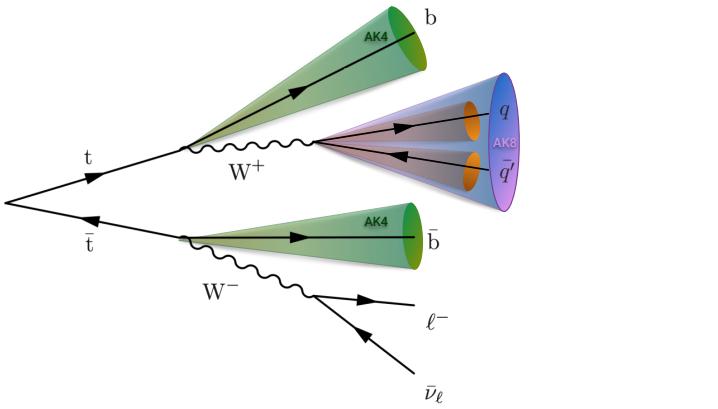


Figure 5.23: A top quark pair decaying into two b quarks and two W bosons, one of which decays leptonically and one on which decays hadronically

Event selection

The W can decay either to an electron or a muon, both final states ('channels') are used in the analysis. We select events through triggering and selections on the leptonic leg. First, we require a high-energy lepton at trigger level, with an online p_T above 45 GeV for the muon and 135 GeV for the electron. This requires an offline muon(electron) p_T threshold of 53(120) GeV. The leptons are further required to pass the lepton requirements defined in Section 4.2.2 and Section 4.2.2, and events containing additional leptons (passing the same ID requirements, but looser cuts as defined in Table 5.5) are vetoed. Offline, we further require a high missing energy of 40(80) GeV in the muon(electron) channel. To insure a high signal (boosted hadronic W) purity, the leptonic W four-vector is reconstructed such that we can put tight momentum requirements on the leptonic leg (ensuring that both tops, and therefore vector bosons, have a high momentum). The leptonic W is reconstructed in two steps: First, the unknown z component of the neutrino momentum must be solved for through a second order equation assuming the real W mass

$$M_W^2 = m_\ell^2 + 2(E_\ell E_\nu - p_{x_\ell} p_{x_\nu} - p_{y_\ell} p_{y_\nu} - p_{z_\ell} p_{z_\nu}) = (80.4)^2.$$

This results in a completely defined neutrino four-vector, which is then added to the lepton four-vector. The sum of the two defines the leptonic W and its momentum is required to be greater than 200 GeV.

Further, we require at least one AK R=0.4 jet to be b-tagged with the Combined Secondary Vertex (CSV) algorithm [48, 49]. This algorithm exploits the relatively long lifetime of b quarks leading to the presence of a displaced vertex, in order to distinguish between jets originating from b quarks to those originating from light flavor quarks. More information on the CSV algorithm can be found in [48, 49]. The reason for requiring only one b-tagged jet is to ensure a high selection efficiency.

Finally, we require at least one AK R=0.8 jet in the event with a momentum greater than 200 GeV which will be the hadronic W candidate. Its pruned jet mass is required to be between 40 GeV and 150 GeV. After reconstructing and selecting all our objects, a set of angular selections are applied to ensure a diboson like topology. These are the following:

- $\Delta R(l, W_{AK8}) > \pi/2$
- $\Delta\phi(W_{AK8}, E_T^{\text{miss}}) > 2$
- $\Delta\phi(W_{AK8}, W_{lep}) > 2$

With these requirements, we have a nearly pure sample of $t\bar{t}$ events, with a small contamination from single top, W+jets and VV events. A summary of the final selection criteria is presented in Table 5.5. The pruned jet mass and τ_{21} variables in data and in MC are shown in Figure 5.24.

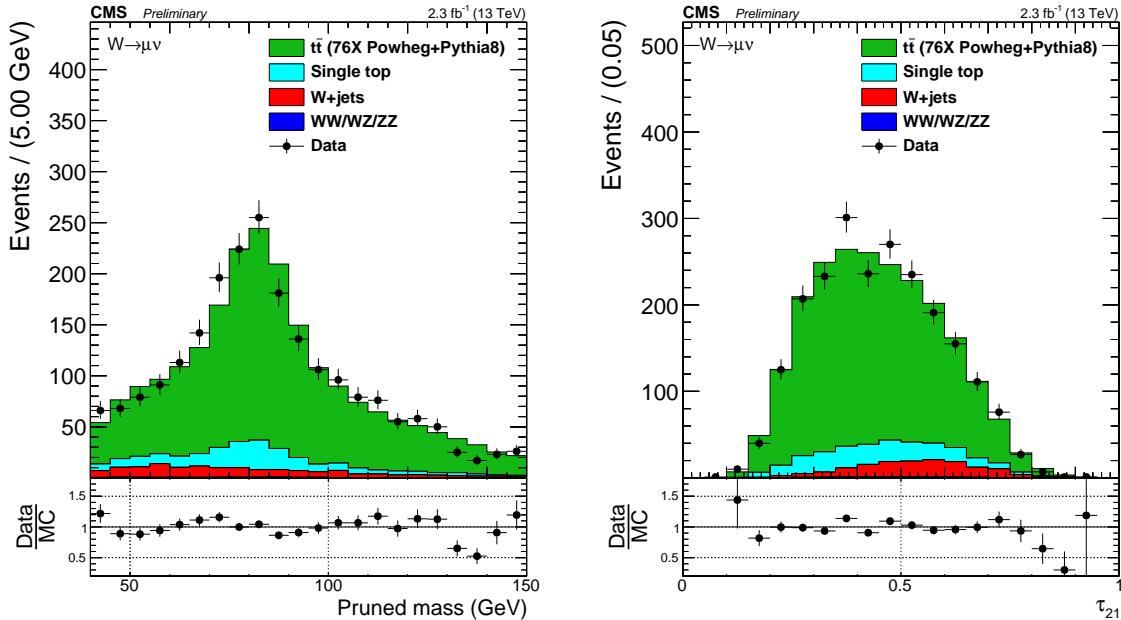


Figure 5.24: Distribution of pruned jet mass (left) and n-subjettiness (right) in the $t\bar{t}$ control sample.

Fitting procedure

For this measurement, what we are interested in is to extract and compare the W-tagging efficiency of the combined jet mass and τ_{21} selection in data and in MC. We are additionally interested in the difference in jet mass scale (mean of the W jet mass peak) and jet mass

Selection	Value	Comments
Tight Lepton selection		
Electron p_T	$p_T > 120 \text{ GeV}$	
Muon p_T	$p_T > 53 \text{ GeV}$	
Electron η	$ \eta _{\text{SC}} < 2.5$ except [1.4442, 1.566]	Veto ECAL barrel-endcap transition.
Muon η	$ \eta < 2.1$	
Loose Lepton selection		
Electron p_T	$p_T > 35 \text{ GeV}$	
Muon p_T	$p_T > 20 \text{ GeV}$	
Electron η	$ \eta _{\text{SC}} < 2.5$ except [1.4442, 1.566]	Veto ECAL barrel-endcap transition.
Muon η	$ \eta < 2.4$	
AK8 jet selections		
Jet p_T	$p_T > 200 \text{ GeV}$	
Jet η	$ \eta < 2.4$	For hadronic W reconstruction
AK4 jet selections		
Jet p_T	$p_T > 30 \text{ GeV}$	
Jet η	$ \eta < 2.4$	Used for b-tag jet selection
E_T^{miss} selections		
E_T^{miss} (electron channel)	$E_T^{\text{miss}} > 80 \text{ GeV}$	
E_T^{miss} (muon channel)	$E_T^{\text{miss}} > 40 \text{ GeV}$	
Boson selections		
Pruned jet mass	$40 < m_p < 150 \text{ GeV}$	
Leptonic W p_T	$p_T > 200 \text{ GeV}$	
Hadronic W p_T	$p_T > 200 \text{ GeV}$	
Veto		
Number of loose electrons	0	
Number of loose muons	0	
Number of b-tagged jets	> 0	CSV medium working point
Angular selections		
$\Delta R(l, W_{\text{AK8}})$	$> \pi/2$	
$\Delta\phi(W_{\text{AK8}}, E_T^{\text{miss}})$	> 2	
$\Delta\phi(W_{\text{AK8}}, W_{\text{lep}})$	> 2	

Table 5.5: Summary of the final semi-leptonic $t\bar{t}$ selections.

resolution (width of W jet mass peak), as this also affects the signal jet mass shape and therefore efficiency. In order to study these variables, we look at the pruned jet mass spectrum between 40 and 150 GeV in two regions:

- Pass region: $0 < \tau_{21} \leq 0.45 \sim \text{high purity}$
- Fail region: $0.45 < \tau_{21} \leq 0.75 \sim \text{low purity}$

Our goal is to understand what the real fraction of merged W jets is in the pass category and in the fail category, assuming that the sum of the two correspond to a 100% selection efficiency (the amount of W jets falling outside of this region is negligible). The strategy is the following: We first derive probability density functions (PDFs) which describe the distribution of fully merged W jets and non-W jets in $t\bar{t}$, both in the pass and in the fail region. The PDFs describing real W jets and non-W jets are added with a fraction which is left floating: the fit decides what the fraction of real W to non-W jets is in the pass and in the fail region. As simultaneous fit of pass and fail is then performed (using the two composite W +non-W PDFs), where the fraction of real W jets in both pass and fail is constrained such that, if the signal efficiency in pass is ϵ_S , the signal efficiency in fail is $(1 - \epsilon_S)$. This is done by letting the normalization of the PDF describing real W jets in the pass category, be defined as the *total* real W yield in pass and fail combined multiplied by some fraction, ϵ_S . The normalization of the PDF describing real W jets in the fail category is then the total real W yield multiplied by $(1 - \epsilon_S)$.

To understand which part of the $t\bar{t}$ jet mass distribution contains “real” merged Ws and which are only pure combinatorial background, non-PWs, we start from $t\bar{t}$ MC. By matching the AK8 jet with quarks coming from the hadronic W at generator level, in a cone of $\Delta R < 0.8$, we can access the real merged W and non-merged W shapes. The real W and non-W PDFs for jets that pass and fail the N-subjettiness selection $\tau_{21} < 0.45$, are found to be well described by the following functions:

$$f_{\text{bkg}}(m_j) = F_{\text{ExpErf}} = e^{c_0 m_j} \cdot \frac{1 + \text{Erf}((m_j - a)/b)}{2} \quad \sim \text{for non-W jets in both pass and fail}$$

$$f^{\text{sig}}(m_j) = F_{\text{Gaus}}(m_j) + F_{\text{ExpErf}}(m_j) \quad \sim \text{for real W jets in both pass and fail}$$

Figure 5.25 shows the fitted PUPPI softdrop mass spectrum for $t\bar{t}$ real W (top) and non-W (bottom) distributions for jets that passed (left) and failed (right column) the N-subjettiness selection PUPPI $\tau_{21} < 0.4$. The corresponding plots for the jet pruned mass can be found in Figure B.

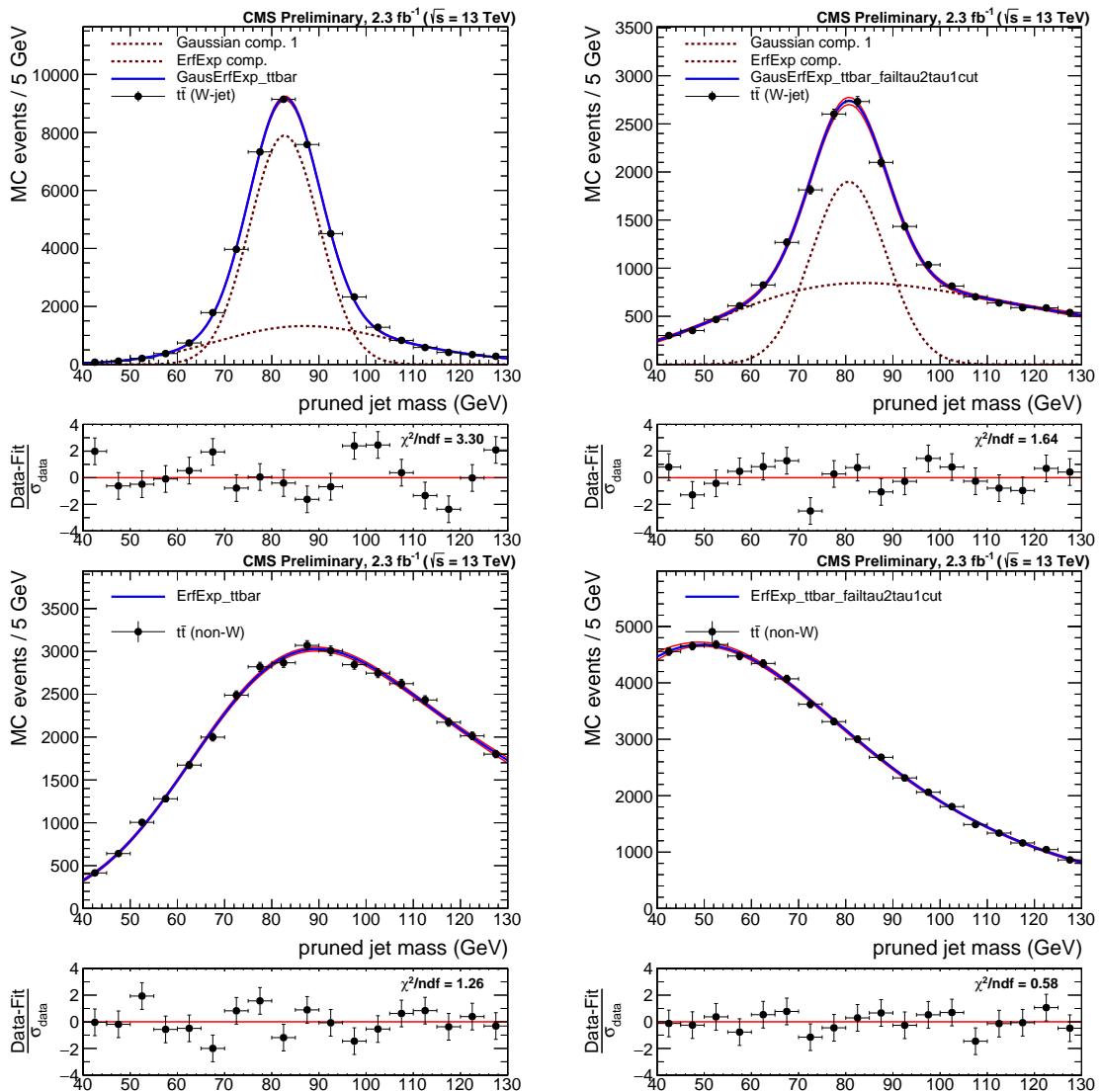


Figure 5.25: Fit to the real W (top) and non-W (bottom) pruned jet mass distribution for jets that pass (left) and fail (right) the cut on $\tau_{21} < 0.45$.

These shapes constitute the fit functions used for the simultaneous fit. As can be seen

from the fit to real W jets in the pass region, the distribution is not purely Gaussian and have a tail at higher groomed masses. This tail depends on the matching requirements used to define real merged W jets and is unphysical. We therefore assume that the distribution of real W-jets can be described by a Gaussian only, allowing the exponential error function used to describe non W-jets to cover the contribution from the tails, hereby taking the number of real W-jets as the integral of the Gaussian shape only. This eliminates two additional fit functions, corresponding to six free parameters from the fit. In older estimations of the W-tagging scale factor based on the same procedure [50]), the functions used to describe the tail of the real W-jet distributions were also taken into account as contributing to the real W-jet tagging efficiency. These two calculations tests two extremes: The new method assumes a Gaussian peak, absorbing the tails into the background function making the fit more robust, while the old method assumes a Gaussian peak with tails estimated from matched MC. The latter uses a more precise definition of real W jets, but a less robust fit. Both methods were investigated and we found that the absorption of tails into the background function resulted in a decrease in the relative uncertainty on the final scale factor of 50 % and an overall improvement on the fit quality, reducing the fit χ^2 by 15 %. The fit parameters of the functions used to describe non W-jets in both the pass and in the fail region, are further constrained using the values obtained from matched $t\bar{t}$ MC. The W-tagging scale factors (SF_{HP}), for the high purity selection ($\tau_{21} < 0.45$), are then extracted estimating the cut efficiency (ϵ_{HP}) on both data and simulated samples fitting, simultaneously, pass and fail samples:

$$L_{\text{pass}} = \prod_i^{N_{\text{evt}}^{\text{pass}}} \left[N_W \cdot \epsilon_{HP} \cdot f_{\text{pass}}^{\text{sig}}(m_j) + N_2 \cdot f_{\text{pass}}^{\text{bkg}}(m_j) + \sum_{j=\text{ST,VV,WJet}} N_{\text{pass}}^j \cdot f_{\text{pass}}^j \right]$$

$$L_{\text{fail}} = \prod_i^{N_{\text{evt}}^{\text{fail}}} \left[N_W \cdot (1 - \epsilon_{HP}) \cdot f_{\text{fail}}^{\text{sig}}(m_j) + N_3 \cdot f_{\text{fail}}^{\text{bkg}}(m_j) + \sum_{j=\text{ST,VV,WJet}} N_{\text{fail}}^j \cdot f_{\text{fail}}^j \right]$$

where N_W is the number of real W jets, N_2 and N_3 are the number of combinatorial background events passing and failing the τ_{21} cut respectively. N_j and f_j , with $j = \text{ST, VV, WJet}$, are the normalizations and shapes of the minor backgrounds (single top, VV, W+jets) which are fixed from simulation. The fit functions used are

$$f_{\text{pass}}^{\text{Top}} = F_{\text{ErfExpGaus}}(x) = \frac{1 + \text{Erf}((x - a)/b)}{2} \cdot e^{-(x - x_0)^2/2\sigma^2}$$

$$f_{\text{fail}}^{\text{Top}} = F_{\text{ExpGaus}}(x) = e^{ax} \cdot e^{-(x-b)^2/2s^2}$$

$$f_{\text{pass}}^{\text{VV}} = F_{\text{ExpGaus}}(x) = e^{ax} \cdot e^{-(x-b)^2/2s^2}$$

$$f_{\text{fail}}^{\text{VV}} = F_{\text{ExpGaus}}(x) = e^{ax} \cdot e^{-(x-b)^2/2s^2}$$

$$f_{\text{pass}}^{\text{wjet}} = F_{\text{ErfExp}}(x) = e^{c_0 x} \cdot \frac{1 + \text{Erf}((x - a)/b)}{2}$$

$$f_{\text{fail}}^{\text{wjet}} = F_{\text{ErfExp}}(x) = e^{c_0 x} \cdot \frac{1 + \text{Erf}((x - a)/b)}{2}$$

with the corresponding distributions shown in Figure 5.26. The floating parameters of the fit (besides the PDF shape parameters themselves) are the rates N_W , N_2 and N_3 , and the mean and sigma of the W-mass distribution defined in $f_{\text{pass}}^{\text{sig}}(m_j)$ and $f_{\text{fail}}^{\text{sig}}(m_j)$. The ratio between data and simulation efficiencies are then taken as the W-tagging scale factor:

$$SF_{HP} = \frac{\epsilon_{HP}(\text{data})}{\epsilon_{HP}(\text{sim})} \quad (5.5)$$

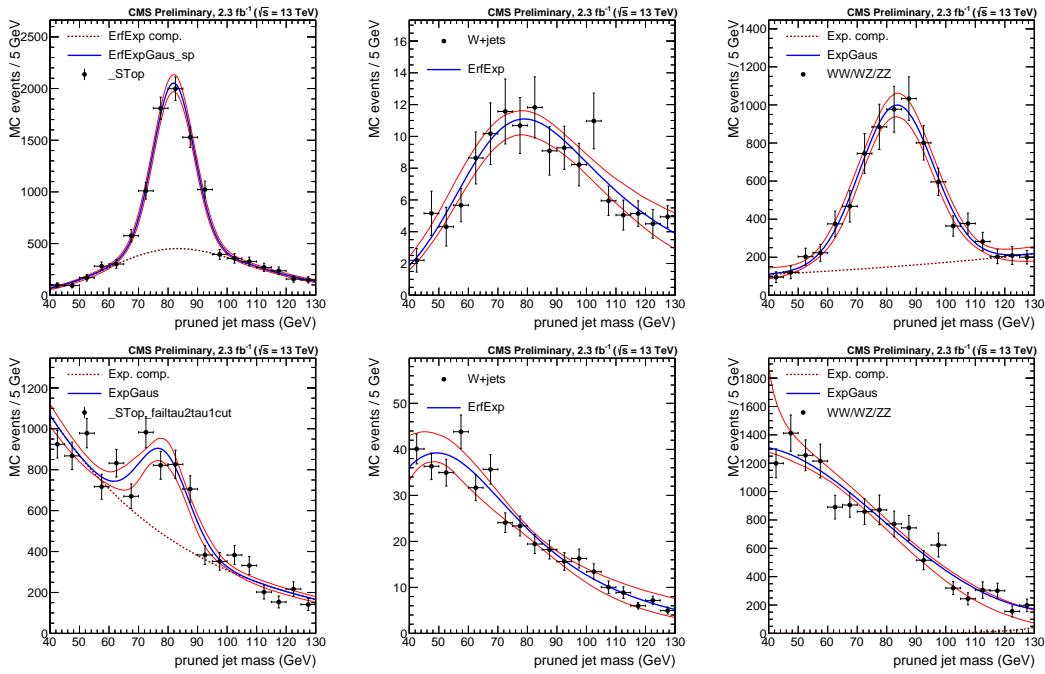


Figure 5.26: Fits to the pruned jet mass spectrum for the non-dominant backgrounds (Single top, W+jets and VV respectively) in the pass (top) and fail (bottom) regions.

Considering that, both for data and simulation, $\epsilon_{HP} + \epsilon_{LP} + \epsilon_{fail} = 1$, the scale factor for low purity category can be defined as:

$$SF_{LP} = \frac{1 - \epsilon_{HP}(\text{data}) - \epsilon_{fail}(\text{data})}{1 - \epsilon_{HP}(\text{sim}) - \epsilon_{fail}(\text{sim})}$$

where ϵ_{fail} is the ratio between the number of events with $\tau_2/\tau_1 > 0.75$ and the total number of events. As mentioned previously, the number of real W jets with $\tau_2/\tau_1 > 0.75$ is negligible and the definition of the low purity scale factor simplifies to

$$SF_{LP} = \frac{1 - \epsilon_{HP}(\text{data})}{1 - \epsilon_{HP}(\text{sim})} \quad (5.6)$$

Systematic uncertainties

As systematic uncertainties, we consider effects due to differences in $t\bar{t}$ simulation as well as effects due to choice of fit method. The former is evaluated by comparing the extracted scale factor when using $t\bar{t}$ MC samples produced with different matrix element (ME) and shower generators: POWHEG (NLO) interfaced with PYTHIA8, MADGRAPH (LO) QCD interfaced with HERWIG++ and POWHEG interfaced with HERWIG++. The uncertainty due to different ME generators (POWHEG versus MADGRAPH) correspond to 3(13)% and are listed in Table 5.6 as the first quoted systematic uncertainty. The uncertainty due to parton showering (PYTHIA8 versus HERWIG++) is 8.6%, but are not relevant for analyses where no HERWIG++ based simulation is used, as is the case for the search presented in this chapter. For the systematic uncertainty accounting for effects due to choice of fit method, we compare the estimated extracted efficiency in $t\bar{t}$ MC using the two different fit models described above: The new model, where the signal is modeled by a Gaussian peak and the tails of the distribution are absorbed in the background fit model, and the old model, including the tails when calculating

the fraction of real W jets. Figure 5.27 shows the fits obtained in the pass and fail regions using the two different models. With the new model only the Gaussian component of the fit contributes to the W-tagging efficiency while, with the old model, a Chebyshev component is additionally contributing to the total W-tagging efficiency. The estimated efficiencies

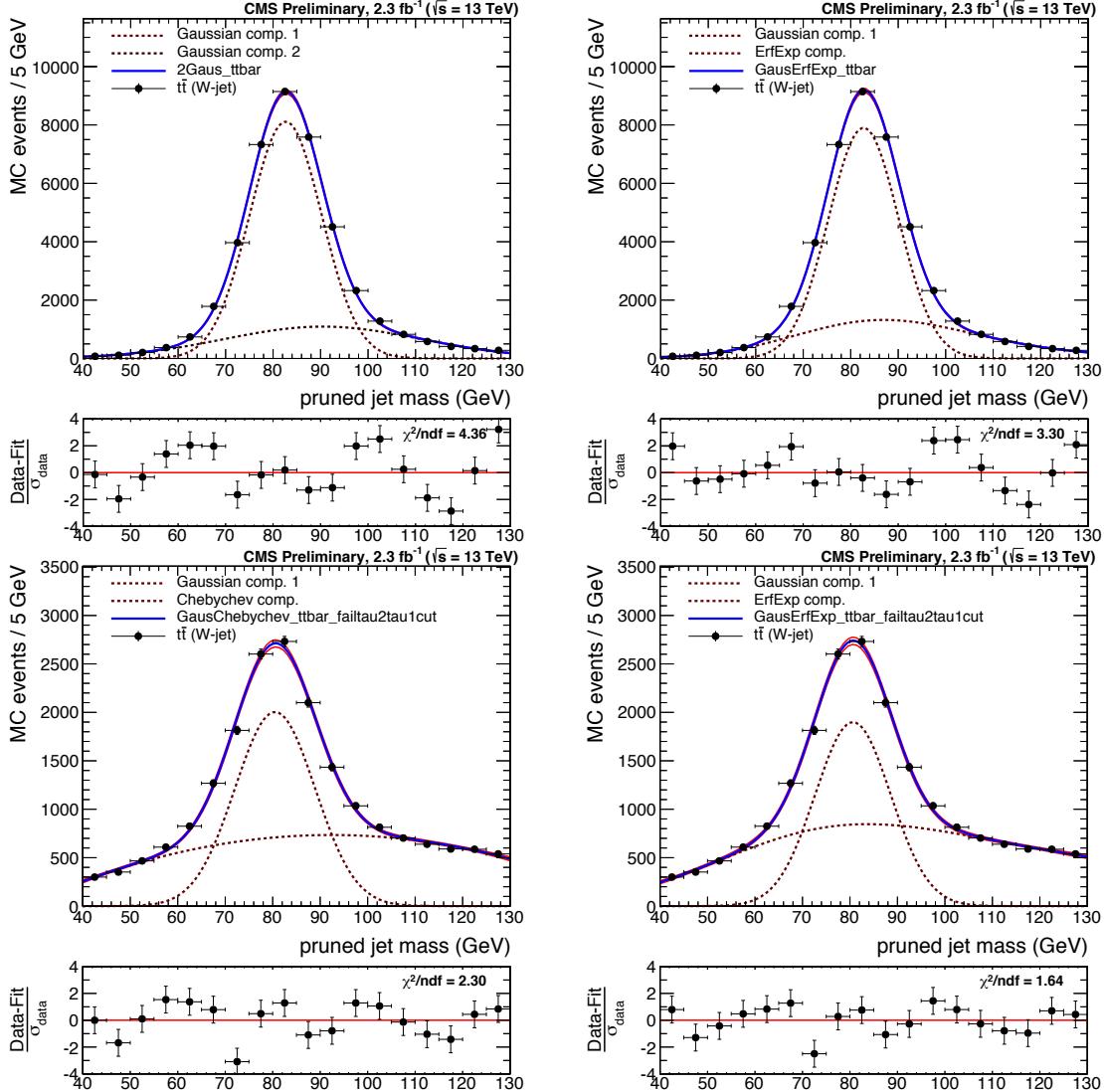


Figure 5.27: Fits obtained in the pass (top) and fail (bottom) regions using two different models: An alternative model with tails (top and bottom, left) where the tail component is contributing to the total W-tagging efficiency. When using the default model (top and bottom, right), only the Gaussian component of the fit contributes to the W-tagging efficiency.

obtained using both methods, after being corrected for the fraction of W jets in the tails, agree within 0.3(0.8)% and are listed as systematic uncertainty in Table 5.6.

One additional uncertainty is added. As the W-tagging scale factor is evaluated in a $t\bar{t}$ sample, the transverse momentum range is rather limited. When the W p_T reaches ~ 400 GeV, the AK8 jet becomes a fully merged top jet with a mass of 170 GeV and a scale factor measurement becomes impossible. However, the jets used in the analyses presented in this thesis have very high transverse momenta, up to 2-3 TeV, and we therefore need an estimate of how the uncertainty on the W-tagging scalefactor changes as a function of p_T . This is estimated by comparing the difference in tagging between $G_{\text{bulk}} \rightarrow WW$ signal MC showered

by PYTHIA8 and HERWIG++ as a function of p_T , relative to the difference in tagging efficiency between the two at a $p_T \sim 200$ GeV. This measurement was performed by a separate analysis team, and found to be $5.90\% \times \ln(p_T/200\text{ GeV})$.

Systematic uncertainties from other sources (lepton identification, b tagging etc.) are less than 0.5% and therefore negligible.

Results

The simultaneous fit as described above is then performed both for data and for simulation, where we take the ratio of data and MC efficiencies as efficiency scale factors. The corresponding fits are shown in Figure 5.51, with the corresponding extracted efficiencies and scale factors summarized in Table 5.6.

Category	Working point	Eff. data	Eff. simulation	Scale factor
HP	$\tau_2/\tau_1 < 0.45$	0.775 ± 0.041	0.822 ± 0.033	$0.94 \pm 0.05 \text{ (stat)} \pm 0.03 \text{ (sys)} \pm 0.003 \text{ (sys)}$
LP	$0.45 < \tau_2/\tau_1 < 0.75$	0.225 ± 0.041	0.178 ± 0.033	$1.27 \pm 0.25 \text{ (stat)} \pm 0.13 \text{ (sys)} \pm 0.008 \text{ (sys)}$

Table 5.6: Efficiencies in data and in MC together with the corresponding W-tagging scale factors for the high purity and low purity categories.

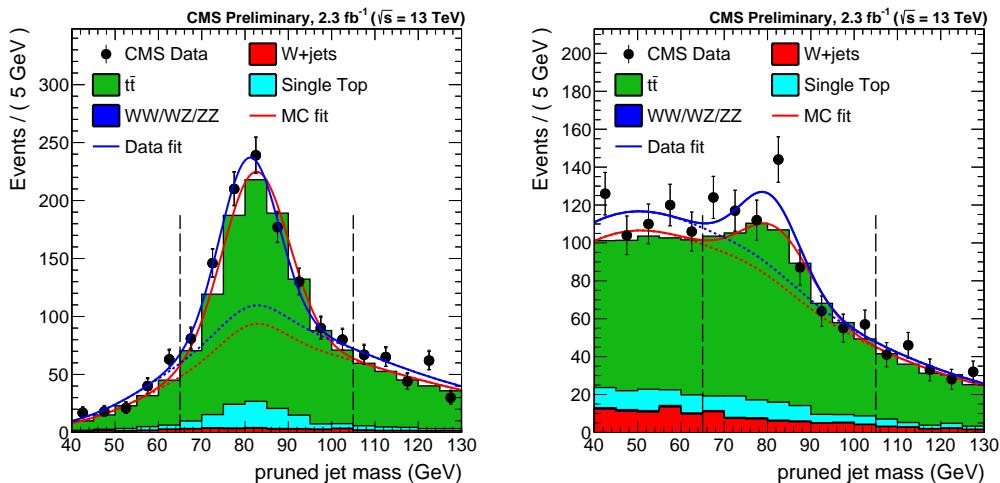


Figure 5.28: Pruned jet mass distribution that pass (left) and fail (right) the $\tau_2/\tau_1 < 0.45$ selection. Results of both the fit to data (blue) and simulation (red) are shown. The background components of the fit are shown as short-dashed lines.

We additionally extract the jet mass scale and jet mass resolution, used to scale and smear the jet mass signal shape in the limit setting procedure. These values are taken from the mean $\langle m \rangle$ and width σ of the Gaussian component of the simultaneous fit in the pass region and are summarized in Table 5.7. Both the jet mass scale as well as the jet mass resolution is larger in simulation than in data with a relative difference of 2 and 10%, respectively. However, the jet mass resolution scale factor has a large uncertainty attached to it and is statistically insignificant (in agreement with unity within uncertainty).

Impact on search variables

The obtained W-tagging scale factors are used as a scale of the signal yield. As we require two W-tagged jets, either HPHP or HPLP, the actual scale factors for the high-purity signal

Parameter	Data	Simulation	Data/Simulation
Pruning $\langle m \rangle$	80.9 ± 0.6 GeV	82.5 ± 0.1 GeV	0.980 ± 0.007
Pruning σ	6.7 ± 0.7 GeV	7.5 ± 0.3 GeV	0.89 ± 0.10

Table 5.7: Jet mass scale and resolution in data and in simulation together with the relevant data-simulation scale factors.

yield is $SF_{HP} \times SF_{HP}$ and for the low-purity category $SF_{HP} \times SF_{LP}$. The signal yields are then

$$N_S^{HP} = N_{\text{HP tot. yield}} \times SF_{HP} \times SF_{HP}$$

$$N_S^{LP} = N_{\text{LP tot. yield}} \times SF_{HP} \times SF_{LP}$$

The uncertainties on the scale factors are considered as anti-correlated between the HP and the LP categories. The jet mass scale and resolution are used to scale and smear the signal Monte Carlo. An uncertainty on the signal yield based on the uncertainty on jet mass scale and resolution is also considered by scaling and smearing the jet mass up and down within the quoted uncertainties and then recomputing the signal efficiency. The results are listed in Table 5.8.

5.1.8 Systematic uncertainties

The uncertainty on the background parametrization is statistical only and is taken as the covariance matrix of the dijet fit function. As demonstrated in the F-test, we study different background parameterizations and we have found these to be within the fit uncertainty of the nominal fit. The remaining uncertainties concern the signal shape and yield and are listed in Table 5.8. Jet reconstruction uncertainties affect both the signal yield and the signal shape. These are evaluated by rescaling the jet four-momenta according to uncertainties on the jet energy scale and resolution and recomputing the signal efficiency. The difference in efficiency with and without smearing/scaling is taken as systematic uncertainties, as described above. The jet mass/energy scale and resolution also affect the signal shape, and are added as uncertainties in the peak position and width of the Gaussian component of the signal PDFs.

Source	Relevant quantity	HP uncertainty (%)	LP uncertainty (%)
Jet energy scale	Resonance shape	2	2
Jet energy resolution	Resonance shape	10	10
Jet energy and m_{jet} scale	Signal yield	0.1–4	
Jet energy and m_{jet} resolution	Signal yield	0.1–1.4	
Pileup	Signal yield	2	
Integrated luminosity	Signal yield	2	
PDFs (W')	Signal yield	4–19	
PDFs (Z')	Signal yield	4–13	
PDFs (G_{bulk})	Signal yield	9–77	
Scales (W')	Signal yield	1–14	
Scales (Z')	Signal yield	1–13	
Scales (G_{bulk})	Signal yield	8–22	
Jet energy and m_{jet} scale	Migration	1–50	
V tagging τ_{21}	Migration	14	21
V tagging p_T -dependence	Migration	7–14	5–11

Table 5.8: Summary of systematic uncertainties and the quantities they affect. Migration uncertainties result in events switching between the purity/mass categories and changes the efficiency in each category, but do not affect the total signal efficiency.

5.1.9 Results

The background fits for each analysis category in the data signal region are shown in Figure 5.29. Here a background only fit is performed while, as described above, a simultaneous fit is used for the limit setting procedure. The filled area correspond to the 1 sigma error band of the background fit, obtained using linear error propagation.

We proceed by setting limits on the cross section of the process $X \rightarrow VV$, using the asymptotic CL_s method as described in Section 2.1. The binned likelihood is defined as

$$L = \prod_i \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!} \quad (5.7)$$

with

$$\mu_i = \sigma \cdot N_i(S) + N_i(B) \quad (5.8)$$

Here σ is the signal strength scaling the expected number of signal events in the i -th dijet invariant mass bin $N_i(S)$, $N_i(B)$ is the expected number of background events in dijet invariant mass bin i and n_i is the observed number of events in the i th dijet invariant mass bin. The background per bin $N_i(B)$ is estimated from the background component of the best signal+background fit to the data points with the signal cross section set to zero. The number of signal events in the i -th dijet invariant mass bin, $N_i(S)$, is then estimated from the signal templates, where only a dijet invariant mass in a 20% window around the resonance mass is considered, containing most of the signal contribution while making sure to keep a good description of the core.

5.1.10 Limits: All-hadronic analysis

As mentioned in Section 5.1.3, we set limits on three different signal scenarios: $G_{\text{bulk}} \rightarrow WW$, $G_{\text{bulk}} \rightarrow ZZ$ and $W' \rightarrow WZ$, with a $\tilde{k} = 0.5$ for the G_{bulk} . Figure 5.30 shows the asymptotic limits at 95 % confidence level on signal cross section as a function of the resonance mass obtained with 2.7 fb^{-1} of 13 TeV CMS data after combining all mass and purity categories (top). The corresponding p-values are shown in the bottom panel.

The statistics are too low to exclude the excess around 2 TeV observed in the corresponding Run 1 analysis and in addition an under-fluctuation in data is present in this region. The largest excess is observed for a $G_{\text{bulk}} \rightarrow ZZ$ hypothesis at a resonance mass of 2.8-3 TeV, around 2.3σ . This is driven by the ZZ high-purity category, the category with the lowest statistics, where one event at 3 TeV yields a local significance of 2.8σ . A 3 parameter fit is the default background fit function for this category, however, a 2 parameter fit could also be used to describe these data. In Figure 5.31 we compare the limits and p-values obtained using a 2 parameter and a 3 parameter fit to describe the background in this category. The significance at 3 TeV is reduced from 2.8 to 1.5σ with a 2 parameter fit, reflecting the fact that the fit is poorly constrained in the high mass tail due to low statistics. The fit to data using both a 2 and 3 parameter fit in the ZZHP category is shown in Figure 5.32 and we in addition see that the 2 parameter fit lies within the fit uncertainties of the nominal fit.

The lack of constraint on the fit in the dijet invariant mass tail when statistics are very low, is a drawback of a method relying fully on a parametric fit and reduces the analysis sensitivity in the high- m_{jj} region. In Search II (Section 5.2) we will keep taking advantage of the dijet fit, however, the integrated luminosity is ~ 15 times higher, resulting in more datapoints in the m_{jj} tail which further constrains the fit. In Search III (Section 5.3), we will

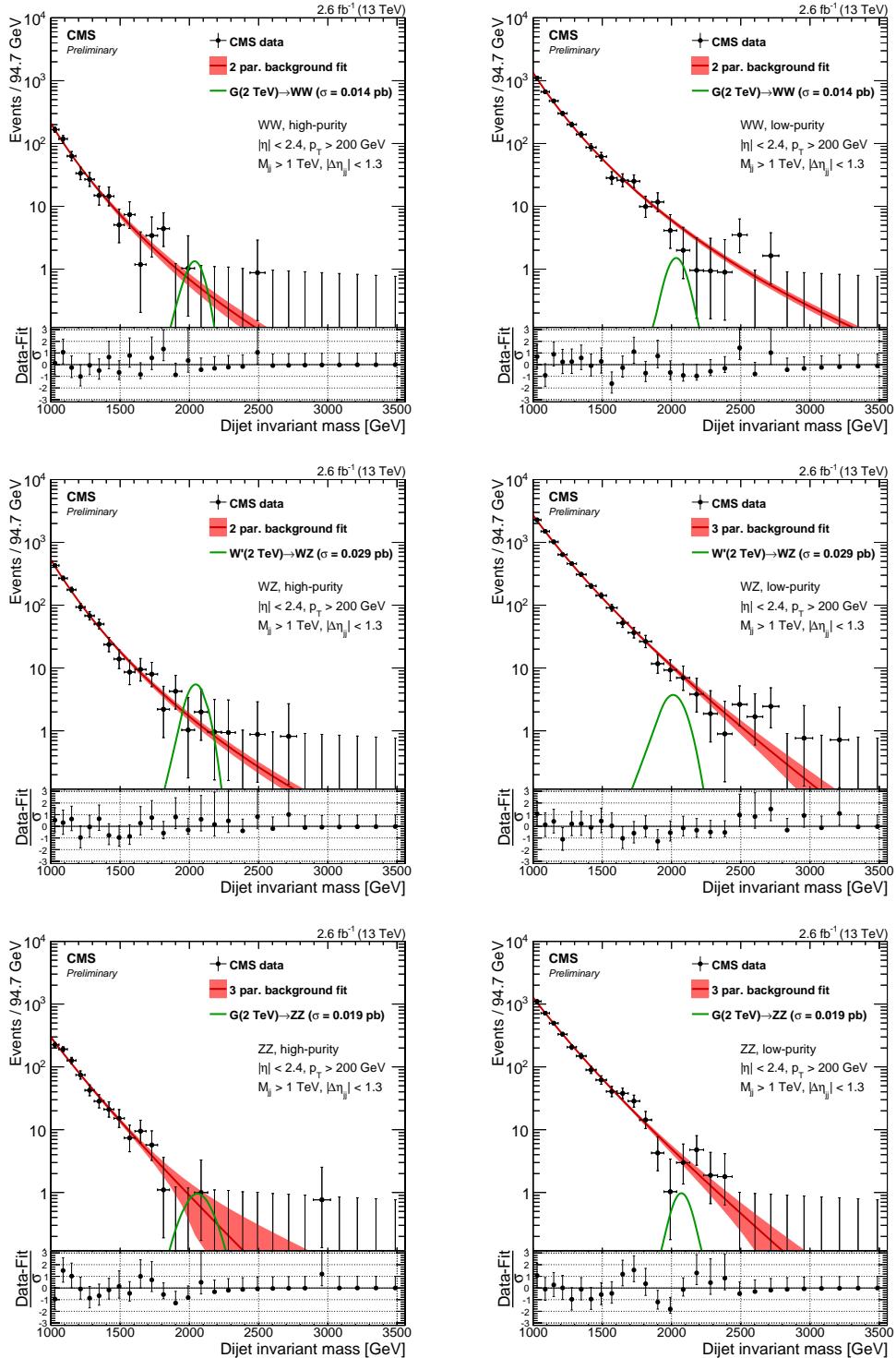


Figure 5.29: Fit to data in the signal region using the background fit only for the different mass and purity categories. The filled red area correspond to the 1 sigma statistical error of the fit.

explore alternate methods which allow more control over the background shape across the full mass spectrum.

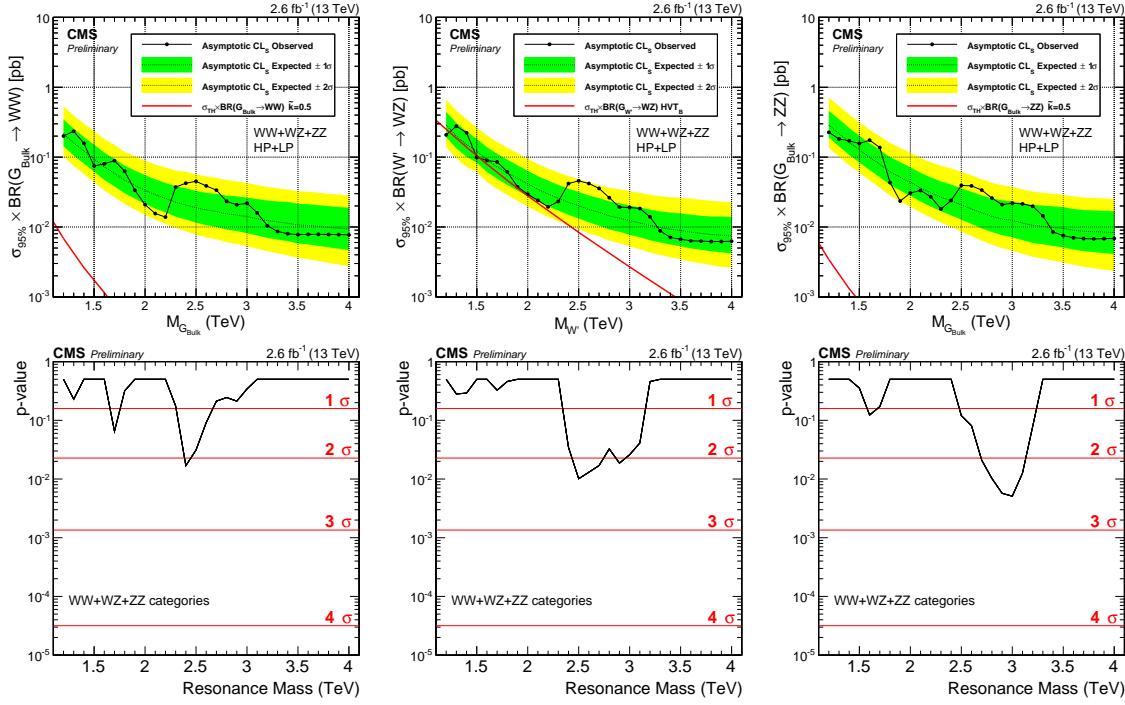


Figure 5.30: Expected and observed limits with corresponding p-values obtained using 2.6 fb^{-1} of CMS data after combining all mass and purity categories. Here for a Bulk $G \rightarrow WW$ (left), $W' \rightarrow WZ$ (middle) and $G \rightarrow ZZ$ (right) signal.

5.1.11 Limits: Semi-leptonic and all-hadronic combination

To maximize the search sensitivity, we combine the results obtained above with those of the corresponding semi-leptonic analysis. We assume the uncertainties on luminosity, V-tagging efficiency, jet mass scale and resolution to be fully correlated.

The obtained exclusion limits are shown in Figure 5.33 shows the resulting expected and observed exclusion limits. As before, we consider a scenario where only either a W' or Z' resonance is expected, called the singlet hypothesis (upper two plots). In addition, we set limits on the triplet hypothesis, assuming the W' and Z' bosons to be degenerate in mass (bottom left plot). Due to larger branching fraction, the all-hadronic analysis sets stronger upper limits than the semi-leptonic analysis above 1.7 TeV for Z' and $> 1.3 \text{ TeV}$ for W' ($\mathcal{B}(WW \rightarrow q\bar{q}q\bar{q}) = 44\%$, $\mathcal{B}(WW \rightarrow \ell\nu q\bar{q}) = 31\%$, $\mathcal{B}(WZ \rightarrow q\bar{q}q\bar{q}) = 46\%$, and $\mathcal{B}(WZ \rightarrow \ell\nu q\bar{q}) = 16\%$). The analysis sensitivity for G_{bulk} is too weak to set limits, but cross sections between 3–1200 fb are excluded. For the HVT model A and B, W' is excluded below < 2.0 and 2.2 TeV , respectively. Z' resonances are excluded below < 1.6 (1.7 TeV) for HVT model B(A). If assuming a HVT Model A(B) triplet hypothesis, resonances below < 2.3 ($< 2.4 \text{ TeV}$) are excluded.

The combined results would therefore just exclude a W' with a mass around 2 TeV, the favored candidate to explain the 8 TeV diboson excess. However, Bulk Graviton signals were still far from excluded and, with the expected ten times increase in luminosity in 2016, we were excited to keep on searching.

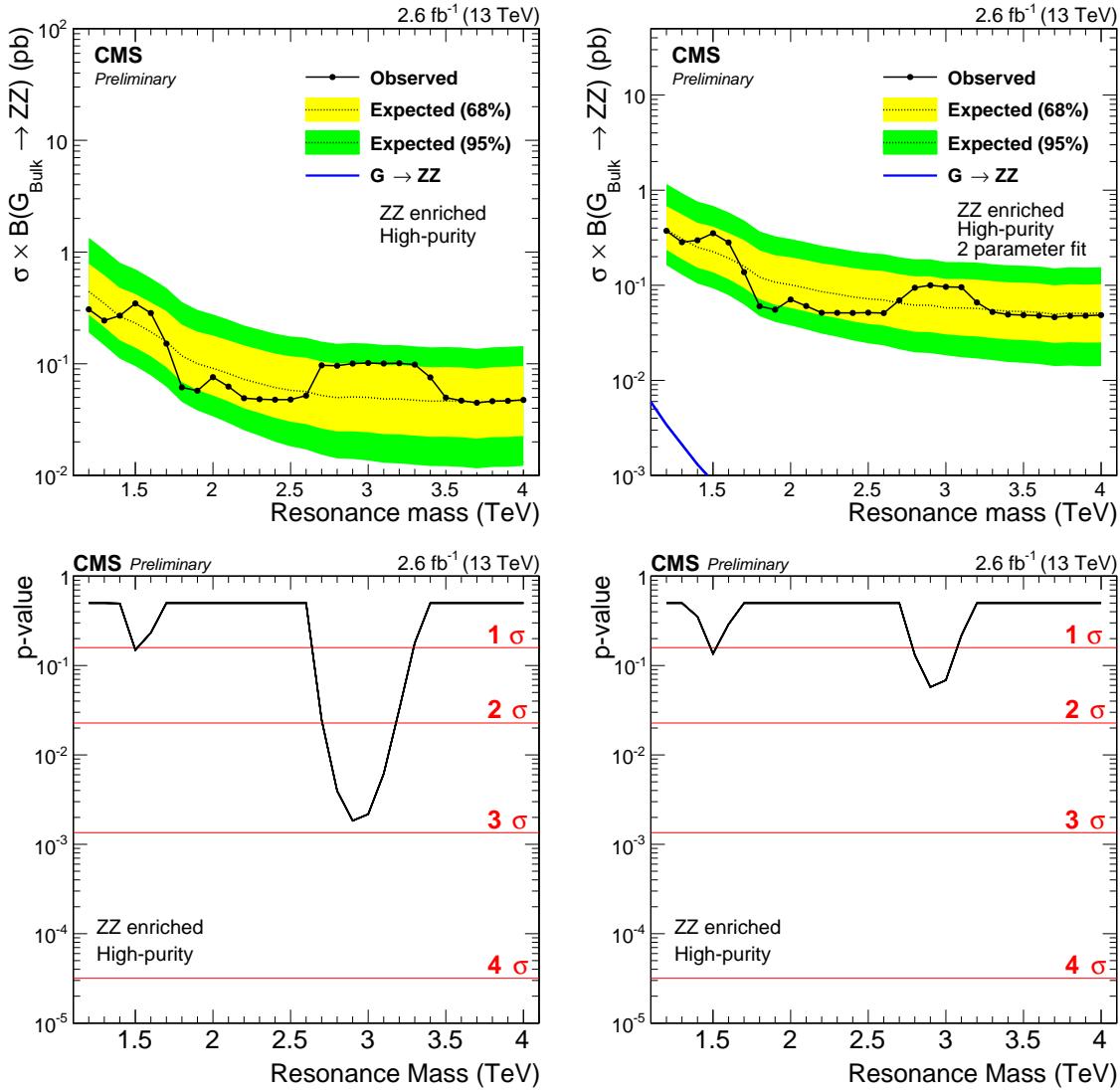


Figure 5.31: Expected/observed limits and corresponding p-values obtained in the ZZHP category using a 3 (left) and two (right) parameter fit to describe the background. The significance at 3 TeV is reduced from 2.8 to 1.5 σ .

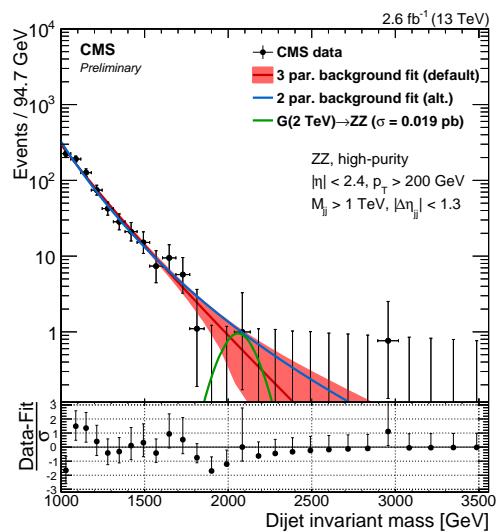


Figure 5.32: Background fit to data in the ZZHP category using the default 3 (red) and an alternate 2 (blue) parameter fit to describe the background.

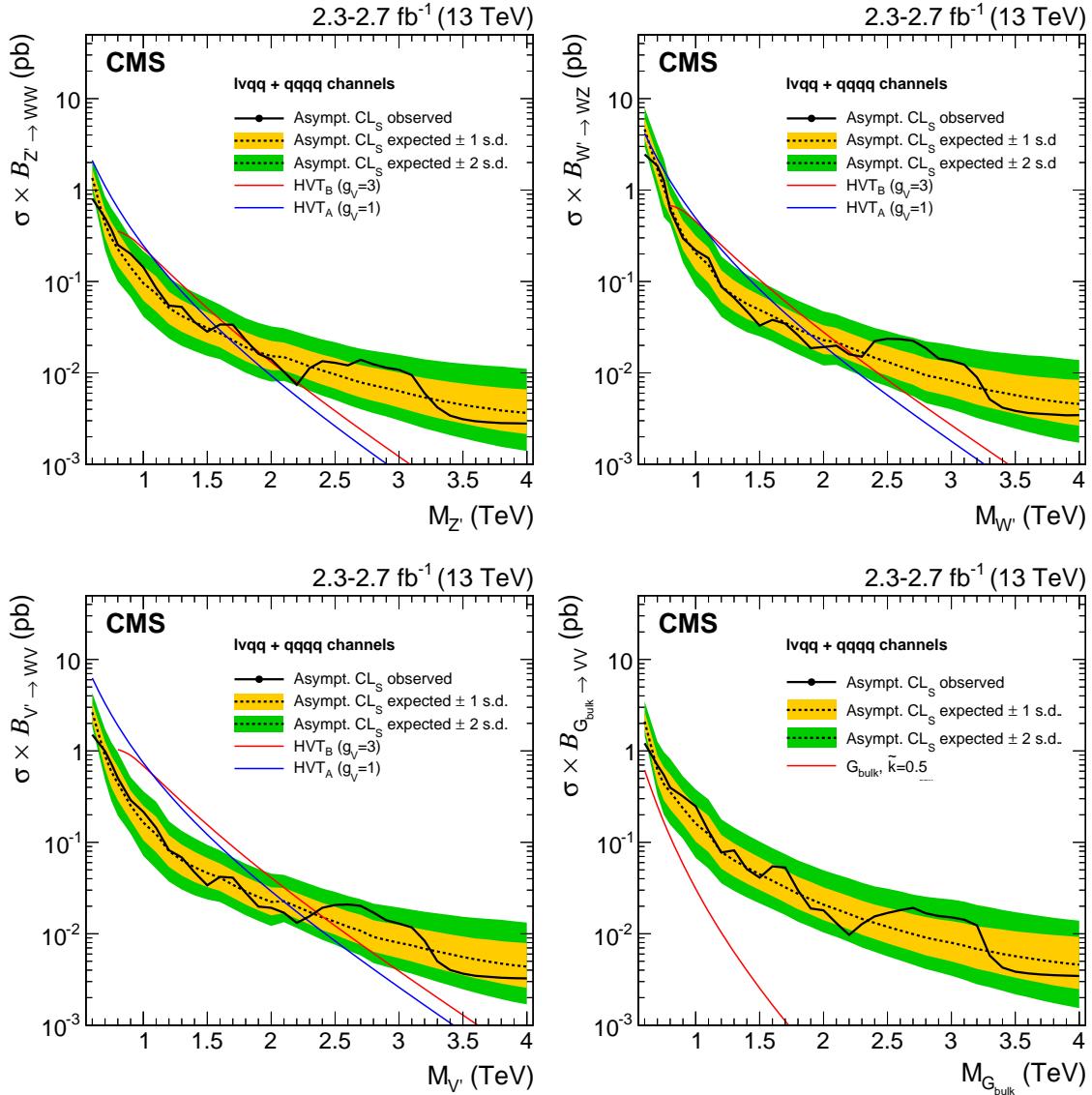


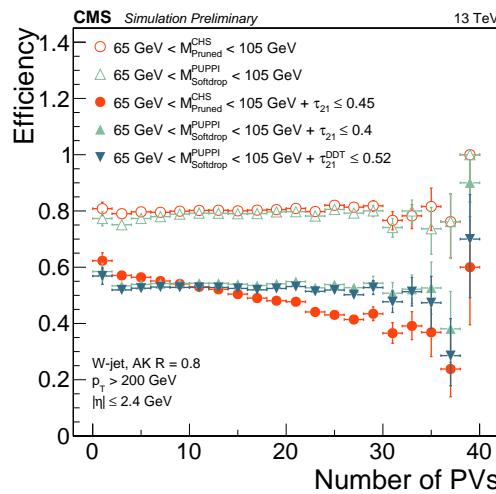
Figure 5.33: Observed (black solid) and expected (black dashed) 95% CL upper limits on the production of a narrow-width resonance decaying to a pair of vector bosons for different signal hypotheses. In the upper plots, limits are set in the context of a spin-1 neutral Z' (left) and charged W' (right) resonances, and compared with the prediction of the HVT Models A and B. In the lower left plot, limits are set in the same model under the triplet hypothesis (W' and Z'). In the lower right plot, limits are set in the context of a bulk graviton with $\tilde{k}=0.5$ and compared with the prediction.

5.2 Search II: A new pileup resistant, perturbative safe tagger

With the first 13 TeV diboson resonance search published, we could conclude that more data would be needed in order to fully exclude the observed Run 1 excess. Luckily, 2016 was right around the corner and, with the LHC planning to reduce β^* from 80 cm to 40, the machine was expected to deliver an instantaneous luminosity three times that of the 2015 peak luminosity. Higher instantaneous luminosity, however, meant double the pileup.

We knew that a novel pileup subtraction algorithm had been developed, which provided far better pileup and underlying event rejection than the current default (CHS). We also knew that there had been made progress on the theory side in the development of a groomer which was insensitive to the soft divergences of QCD and allowed to accomplish jet grooming in a theoretically calculable way, SoftDrop (mMDT). With more time at hand than in 2015, I therefore decided to pursue a novel W-tagger for this second search. This included work like optimization, development of dedicated jet mass corrections (in use today and recommended by the jet physics object group) as well as validation of the new tagger. The tagger, together with the mass corrections, afterwards became the default W-tagging algorithm in CMS.

Search II became the first published analysis to use the novel PUPPI+softdrop algorithm, now default for W-tagging in CMS. Through this search, the tagger was optimized, commissioned and validated, making it available for several analysis to come. In addition, the search was extended to setting limits on three additional signal hypothesis. Two of these were in a final state never before explored at 13 TeV, the $q^* \rightarrow qV$ single V-tag analyses. Published with $35.9(12.9) \text{ fb}^{-1}$ of 2016 data.



5.2.1 Towards robust boosted jet tagging

When we first studied W-tagging at 13 TeV in context with the analysis of the 2015 dataset, Section 5.2.5, two interesting correlations were observed:

- 1) A strong dependence of the AK8 CHS softdrop ($\beta = 0$) jet mass on jet p_T and
- 2) a strong dependence of the AK8 CHS τ_{21} cut efficiency on pileup.

The reason we studied the softdrop algorithm as an alternative to pruning in 2015 was, besides the possibility it would result in a higher signal efficiency, that we knew it had certain favorable qualities compared to other groomers: Softdrop removes all sensitivity to the soft divergences of QCD, by removing all soft emission, more specifically the non-global logarithmic terms (NGLs) in the jet mass [24]. These arise from constellations where, for instance, a soft gluon is radiated into the jet cone, as illustrated in Figure 5.34. The consequence of

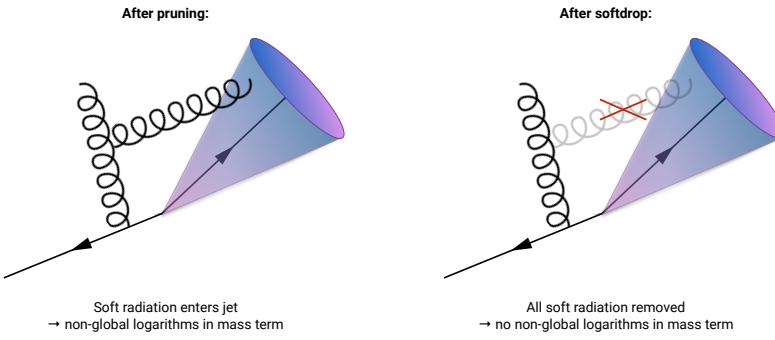


Figure 5.34: The pruning algorithm does not remove all soft emission and therefore has non-global logarithmic terms in the jet mass. Softdrop ($\beta = 0$) completely removes soft emissions and is therefore free of non-global logarithms.

this is that you can calculate the softdrop jet mass to way higher precision than what is possible for other grooming algorithms or for the plain jet mass (NGLs are the main reason a full resummation of the plain jet mass beyond NLL (considering e.g multiple-emission effects) accuracy does not exist). Despite this not being a precision measurement analysis, we had theoretically well-motivated reasons for wanting the baseline CMS V-tagger to be softdrop-based. However, despite being less sensitive to soft radiation for QCD jets, signal jets groomed with softdrop were found to be far more sensitive to the underlying event than pruned jets [51]. Figure 5.35 shows the signal efficiency for pruning (left) and softdrop (right) as a function of jet transverse momenta when including FSR only, FSR+ISR, hadronization and hadronization + underlying event. On parton level, as well as after hadronization, the two algorithms perform very similar as a function of p_T . However, once UE contamination is added, the softdrop tagging efficiency is severely affected. This can be explained by the larger effective radius considered by the softdrop algorithm ($\propto m_V/p_T \sqrt{z_{cut}(1-z_{cut})}$) in comparison to pruning ($\propto m_V/p_T$). This observation corresponds very well with the shift in jet mass we observed for softdrop as a function of p_T in Section 5.2.5: As the jet p_T decreases the softdrop effective radius increases and the jet mass mean shifts to higher values, due to absorbing more background radiation. If softdrop would be our new default tagger, a better rejection of pileup and UE contamination would be needed. In parallel to the ongoing theoretical work on groomers, a novel pileup removal algorithm had been proposed: Pileup per particle identification (PUPPI) [16]. Described in detail in Section 4.3.2, PUPPI considers not only charged pileup but rather reweights each particle in the jet with its probability of arising from pileup. PUPPI had proven it self far superior to the current CHS algorithm in

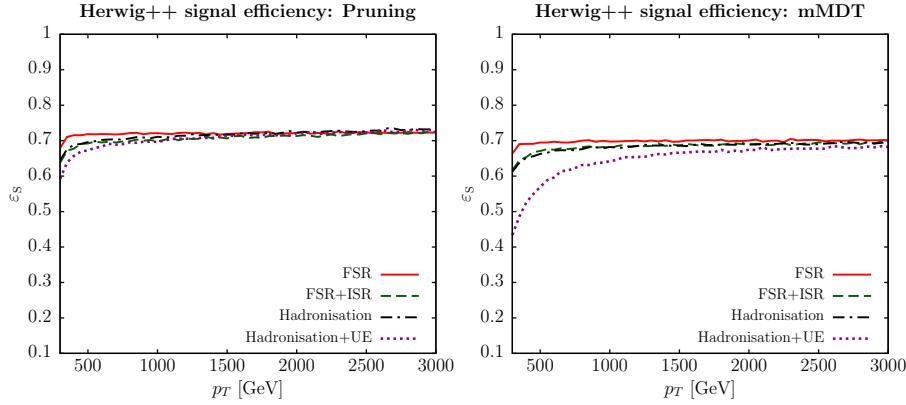


Figure 5.35: The signal efficiency for pruning (left) and softdrop (right) as a function of jet p_T when adding FSR, ISR, hadronization and UE. The UE has a severe impact on the softdrop efficiency for signal jets [51].

terms of jet observables for large radius jets, and therefore seemed like the obvious choice to address both issues listed above: The sensitivity of softdrop regarding UE contamination and the strong pileup dependence of τ_{21} . The focus of Search II would therefore be on the commissioning of a novel W-tagger. There are interesting changes and inclusions in the analysis strategy as well: The inclusion of a $Z' \rightarrow WW$ signal hypothesis and the addition of a completely new analysis, the single V-tag analysis.

5.2.2 Analysis strategy

The analysis strategy for this search is conceptually the same as for Search I. In addition, we'll take advantage of the n-subjettiness categorization and do an additional analysis in parallel: A search for excited quark resonances q^* [52, 53] decaying to qW or qZ . We call this the single V-tag analysis, and the analysis selection only differs in that one jet is not required to pass the V-tag selection (groomed mass and n-subjettiness). The VV analysis is hereby referred to as the double V-tag analysis. The difference between the two analyses is illustrated in Figure 5.36. In addition, limits are set on a $Z' \rightarrow WW$ signal hypothesis in the

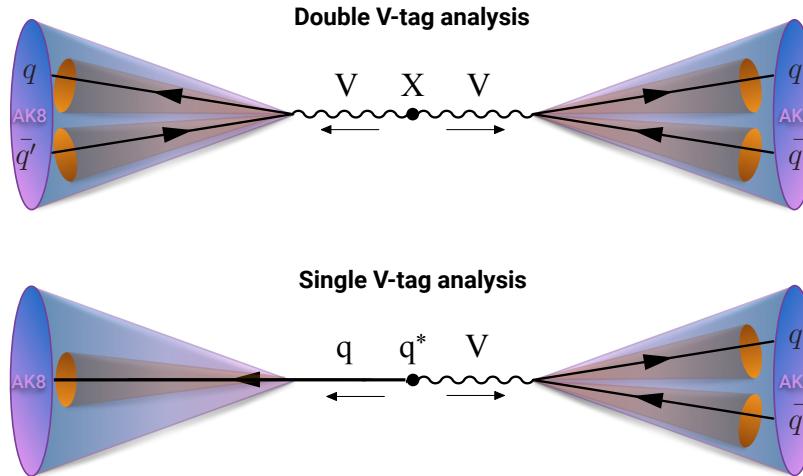


Figure 5.36: The double (top) and single (bottom) W/Z-tag analysis.

double V-tag analysis, another 13 TeV first.

This analysis was published in two steps: An early Physics Analysis Summary (PAS) based on 12.9 fb^{-1} of 2016 data [50], describing the new PUPPI+softdrop based V-tagger as well as the single V-tag analysis, and a second analysis topping up with the full 2016 data [54]. The commissioning of the new W Z-tagger has also been documented in a jet performance Physics Analysis Summary [55]. As the new V-tagger was developed and commissioned in the context of the early analysis, which was also where the single V-tag analysis was first published with 13 TeV data, the main emphasis will be on the work presented in CMS-PAS-B2G-16-021 [50]. The second part of the results chapter, Section 5.2.11, includes the results obtained using the full 2016 dataset of 35.9 fb^{-1} .

5.2.3 Data and simulated samples

As mentioned above, the analysis of the 2016 dataset was done in two steps: One analysis based on 12.9 fb^{-1} of early 2016 data, describing the new W-tagger and single V-tag category, and a second paper topping up with the full 2016 dataset, corresponding to 35.9 fb^{-1} .

The G_{bulk} and HVT signal samples are modeled in precisely the same way as in 2015. For the single V-tag q^* samples, we simulate unpolarized boson with a compositeness scale Λ set equal to the resonance mass. These are generated to leading order using PYTHIA version 8.212 [43].

The background Standard Model processes; QCD, W+jets and Z+jets are all simulated to leading order. V+jets is simulated with MADGRAPH5_AMC@NLO [42, 56], while three different combinations of matrix element and shower generators are used for QCD as these predictions are known to differ: PYTHIA only, the leading order mode of MADGRAPH5_AMC@NLO matched with PYTHIA, and HERWIG++ 2.7.1 [31] with tune CUETHS1 [38].

5.2.4 Event selection

Triggering

The triggers used in this analysis are the same ones as in 2015 (see Section 5.1.4), however, due to the new single V-tag analysis, the trigger turn-ons have this time been re-evaluated separately requiring either one or two jets to have an offline softdrop jet mass above 65 GeV.

Figure 5.37 shows the trigger turn-on curves as a function of dijet invariant mass for jets passing one of the three inclusive triggers only, one of the grooming triggers only and when combining all of them. The turn-on curves are shown for all jet pairs passing loose selections as described in Section 5.1.4. Zero, one or two of the two jets is further required to have a softdrop mass larger than 65 GeV.

Including grooming triggers lowers the 99% trigger efficiency threshold by around 50(80) GeV in the single (double) tag category once substructure is requested on the analysis level. Using the or of all triggers, we are safely on the trigger plateau for dijet invariant masses above 955(986) GeV in the double (single) tag category, setting the analysis threshold at a dijet invariant mass of 955 GeV for the double tag analysis and 990 GeV for the single tag analysis. For controlplots, where no groomed mass window is applied, a trigger threshold of 1020 GeV is used.

Trigger efficiencies as a function of the offline softdrop-jet mass for the `HLT_AK8PFJet360_TrimMass30` trigger are shown in Figure 5.38. Here the jet transverse momentum of one of the jets is required to be at least 600 GeV and no other mass cut is applied. This trigger requires one jet to have a trimmed mass above 30 GeV at HLT level

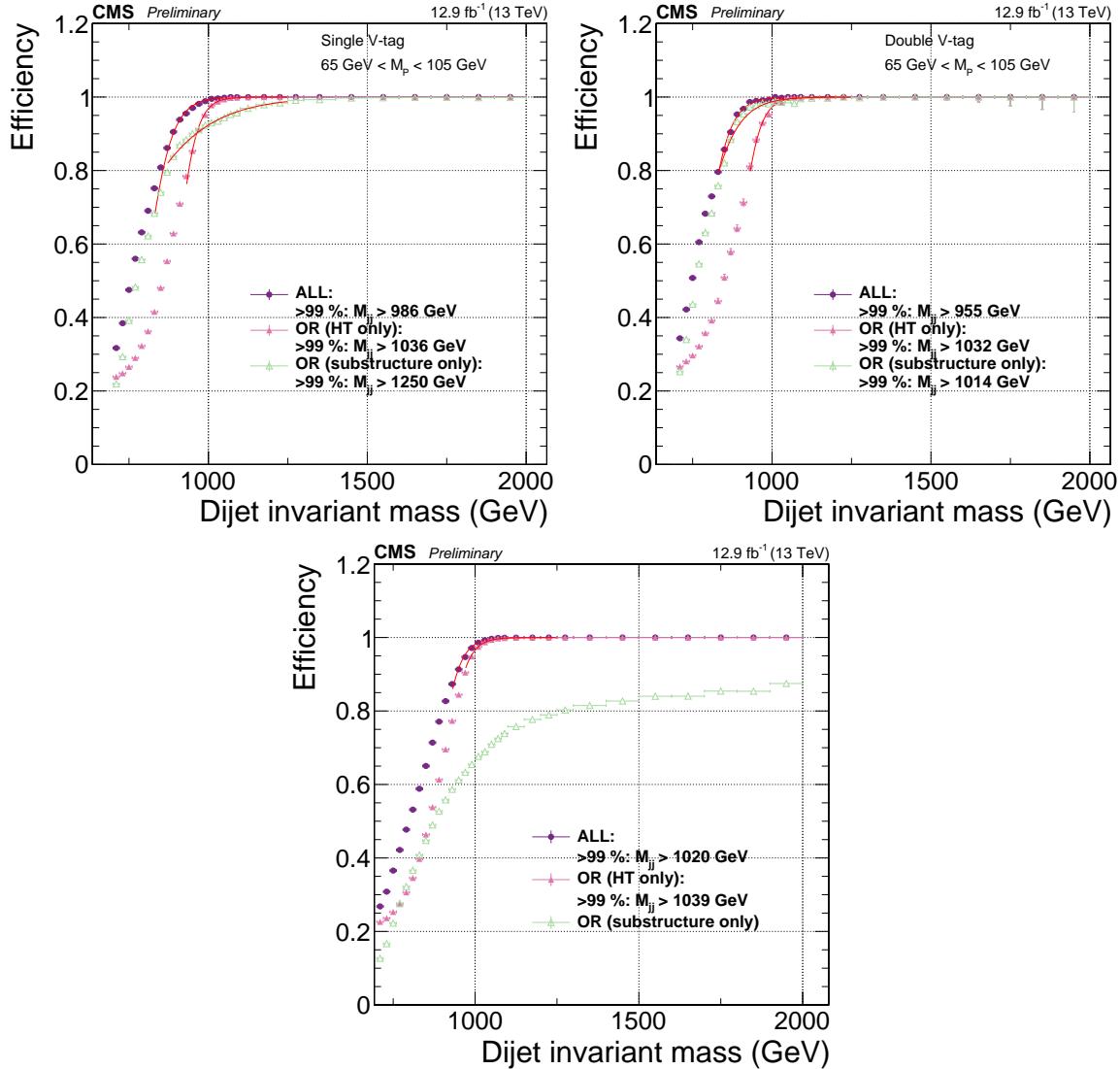


Figure 5.37: Comparison of trigger efficiencies for jets passing one of the HT-triggers only (pink), for jets passing one of the grooming-triggers only (green) and for jets passing one of the HT-triggers or one of the grooming triggers (purple). Here as a function of dijet invariant mass for all jet pairs passing loose selections and where one jet has a softdrop mass larger than 65 GeV (top left), both jets have a softdrop mass larger than 65 GeV (top right) and where no mass cut is applied (bottom).

and reaches the trigger plateau for groomed-jet masses around 50 GeV. As reference trigger, the prescaled trigger HLT_PFTJet320 is used.

Preselection

The same preselections as in Search I, described in , have been applied: We require two AK R=0.8 jets with CHS applied pre-clustering, required to pass the tight jet ID requirement, $p_T > 200 \text{ GeV}$ and $|\eta| < 2.5$. The same QCD t-channel suppressing cut of $|\Delta\eta| < 1.3$ is required together with the following trigger thresholds on the dijet invariant mass: $m_{jj} > 955 \text{ GeV}$ for the double V-tag and 990 GeV for the single V-tag analysis. The jet p_T (top left), η (top right), $\Delta\eta_{jj}$ and dijet invariant mass (bottom left) for the two leading jets in the event after loose preselections are applied is shown in Figure 5.39. A large difference in slope in the

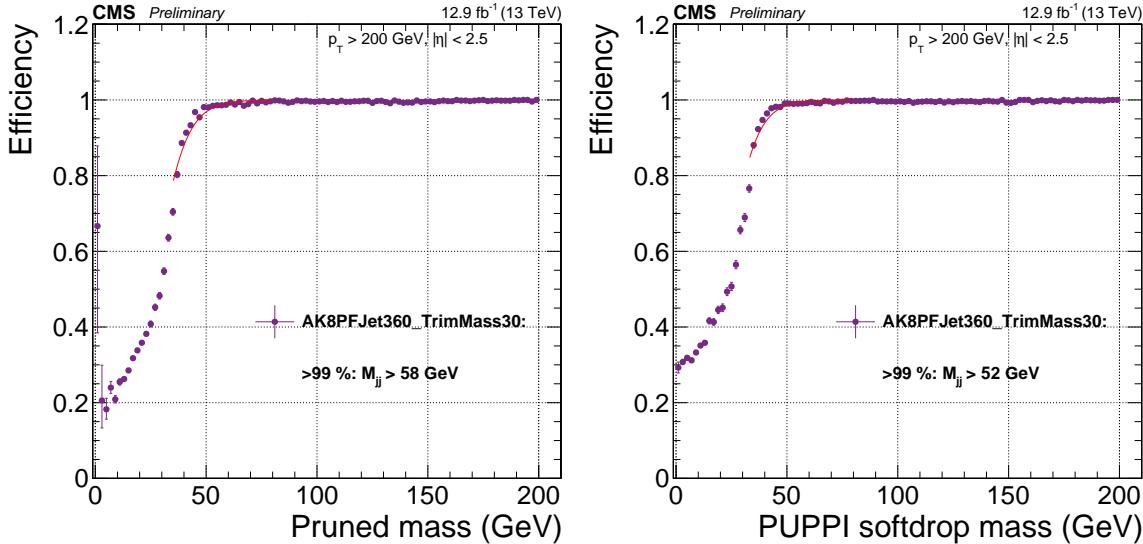


Figure 5.38: Efficiency for the HLT_AK8PFJet360_TrimMass30 trigger as a function of pruned-jet (left) and softdrop-jet (right) mass for jets with $p_T > 600$ GeV.

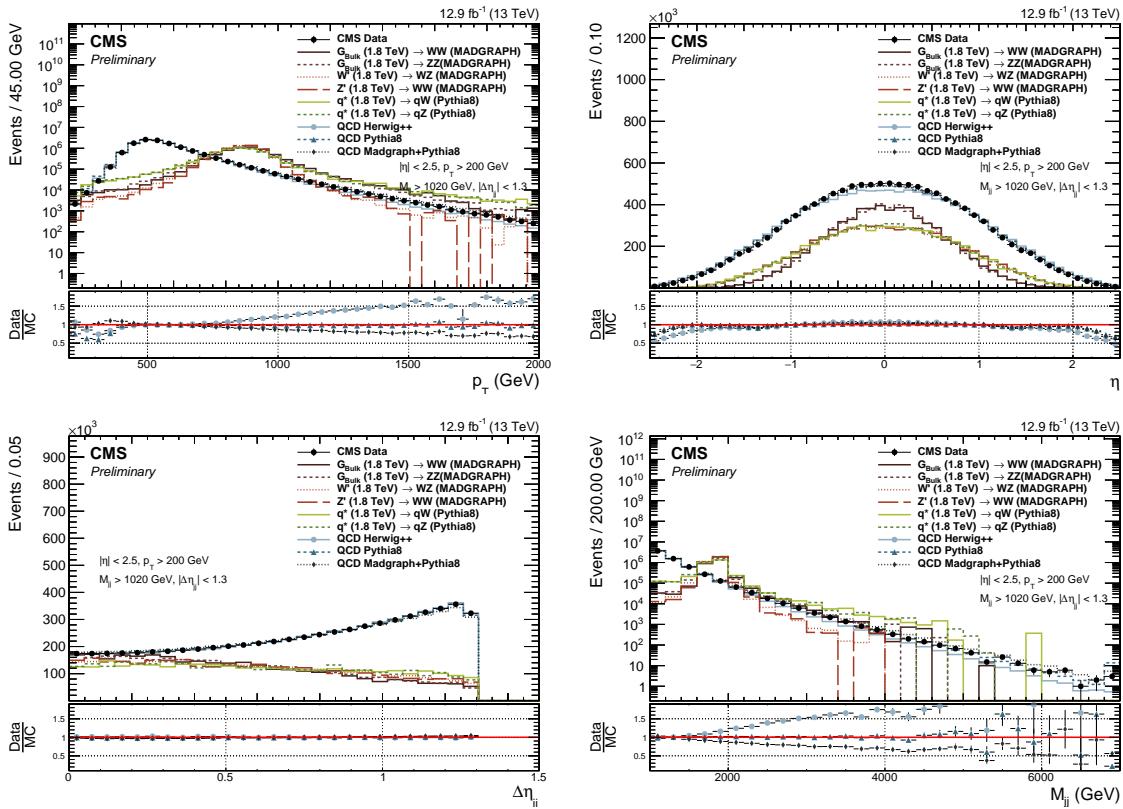


Figure 5.39: Jet p_T (top left), η (top right), $\Delta\eta_{jj}$ and dijet invariant mass (bottom left) for the two leading jets in the event after loose preselections are applied. The signal is scaled by an arbitrary number.

jet p_T and dijet invariant mass spectrum depending on the QCD matrix element or shower generator is observed. Pure PYTHIA QCD MC describes the data best, while HERWIG++ and MADGRAPH5_AMC@NLO+PYTHIA tend to under- or over-estimate the number of high p_T/m_{jj}

jets, respectively. Pure PYTHIA QCD MC is therefore used for all background checks in this analysis.

5.2.5 Developing a new W-tagger

As mentioned in the introduction to this chapter, early studies had shown that the PUPPI pileup subtraction algorithm yielded superior resolution on large-cone jet observables like the jet mass. We therefore wanted to check whether the softdrop jet mass, and its observed sensitivity to the Underlying Event and pileup, would be improved if a better pileup subtraction algorithm was applied pre-clustering.

Two interesting observations were made. Softdrop used together with PUPPI pileup subtraction displayed a much smaller p_T -dependent shift than CHS+Softdrop, as hoped. Figure 5.40 shows the PUPPI softdrop mass for W-jets from a 1 TeV ($p_T \sim 500$ GeV) and 4 TeV ($p_T \sim 2$ TeV) resonance, exhibiting the desired reduced p_T dependence in jet mass scale. However, when applying centrally provided L2 and L3 jet energy corrections (see Section 4.4.3)

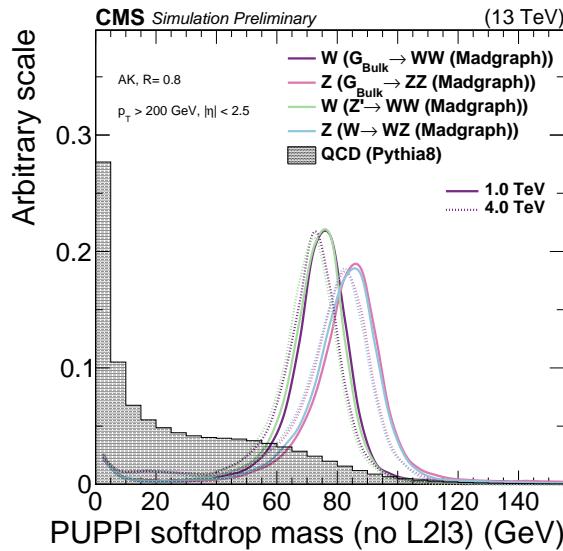


Figure 5.40: The PUPPI softdrop jet mass distribution with no jet energy corrections applied

to the jet groomed mass, as is recommended, a strong p_T dependence is re-introduced. This effect is not present for the pruned jet mass. Figure 5.41 show the softdrop (top left) and pruned (top right) jet mass distribution with recommended L2L3 corrections applied. Here, the PUPPI+softdrop jet mass shift is significantly increased with respect to what was observed for the uncorrected mass, while CHS+pruned jet mass is stable. This points to the PUPPI jet energy corrections not being optimal for scalar jet mass variables, while they may be good for correcting jet 4-vectors. The jet energy corrections derived for CHS and PUPPI jets as a function of jet p_T is shown in the bottom plot in Figure 5.41 . A significant slope in JEC as a function of p_T is measured for PUPPI, while not present for CHS.

Dedicated PUPPI softdrop mass corrections

In order to minimize p_T dependence in the PUPPI softdrop jet mass, all jet energy corrections to the softdrop jet mass are removed. However, this still leaves a residual p_T dependence and, in addition, the uncorrected mass does not peak at the correct W-mass of 80.4 GeV. Figure 5.42 shows the mean of a Gaussian fit to the uncorrected PUPPI softdrop mass as

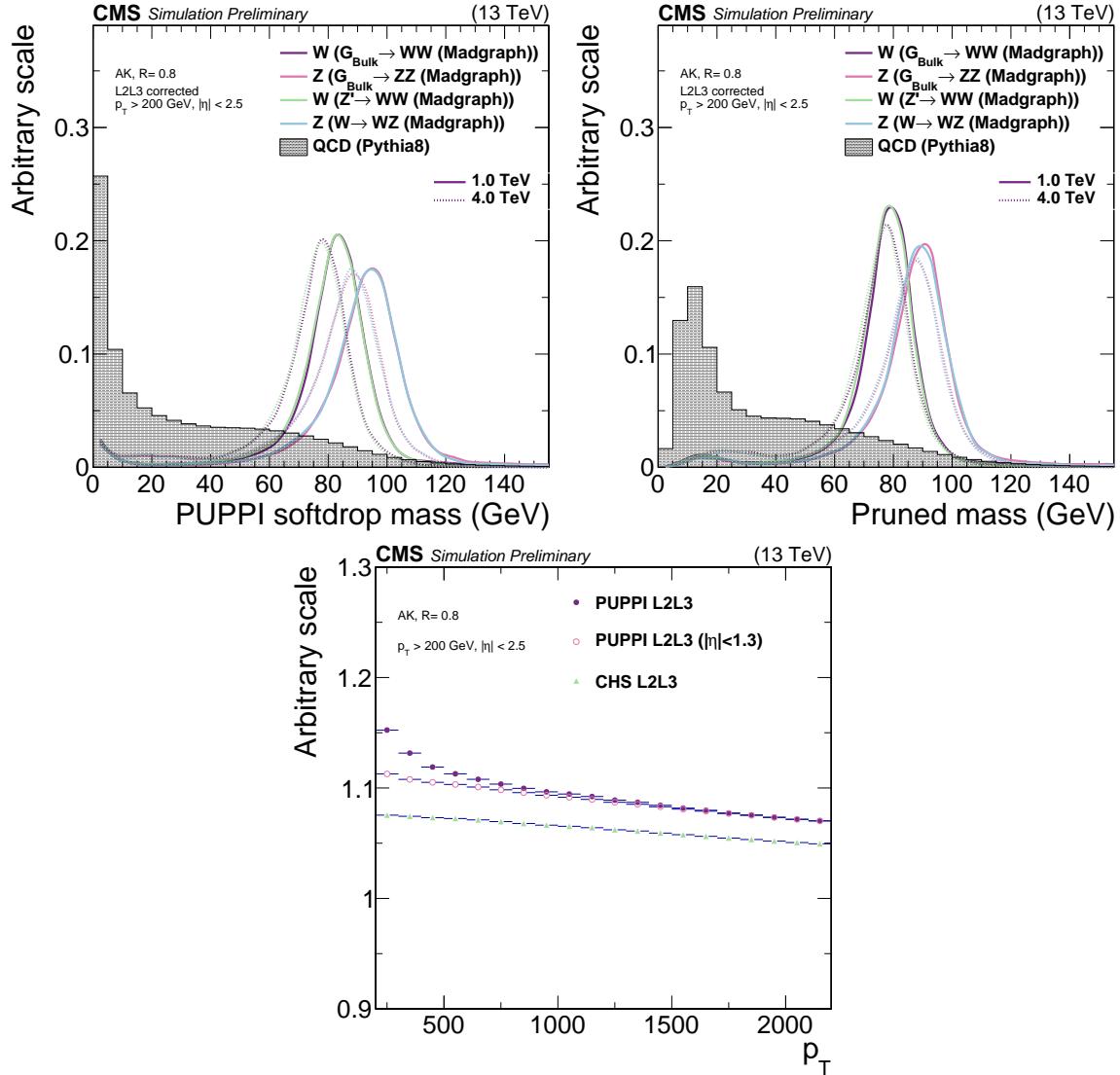


Figure 5.41: Top: PUPPI softdrop mass distribution (top left) and pruned jet mass distribution (top right) with L2 and L3 corrections applied. Bottom: The projection of CHS and PUPPI jet energy corrections versus jet p_T .

a function of jet p_T in two different η bins (smaller or greater than $|\eta| = 1.3$) for W-jets coming from a Bulk Graviton signal sample. A mass shift both as a function of η and p_T is observed, together with an average mean significantly lower than the W-mass. In order to use PUPPI+softdrop for W-tagging, we therefore derive dedicated jet mass corrections to compensate for two factors: A generator level p_T -dependence, as first observed in , and a reconstruction level p_T - and η -dependence, most likely caused by UE effects and the growing effective softdrop radius at low jet p_T . Figure 5.43 shows the mean of the generated softdrop mass (left) and the normalized difference in reconstructed and generated softdrop mass (right) as a function of jet p_T . The shift in generated softdrop mass at lower p_T is of the order of 2-3% while the difference between reconstructed and generated softdrop mass is a 5-10% effect. The mass shift introduced at generator level is corrected by a fit to $M_{\text{PDG}}/M_{\text{GEN}}$ as a function of jet p_T , where $M_{\text{PDG}} = 80.4$ GeV and M_{GEN} is the fitted mean of the generator level mass as shown in the left plot in Figure 5.43. To correct for the residual shift between generated and reconstructed softdrop mass, a fit to $(M_{\text{RECO}} - M_{\text{GEN}})/M_{\text{RECO}}$, where M_{RECO}

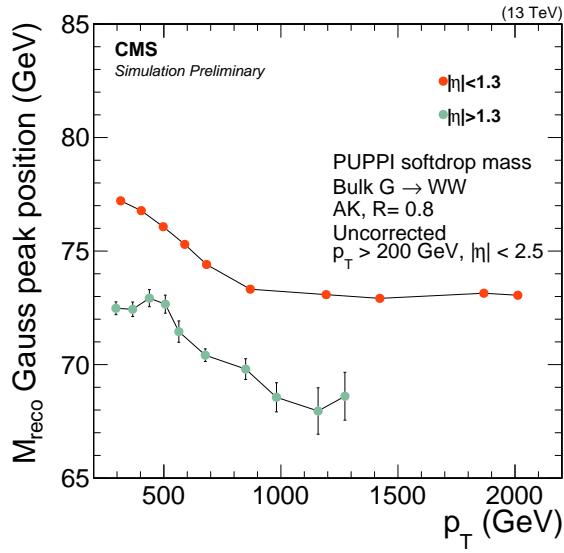


Figure 5.42: The mean of a Gaussian fit to the W-jet PUPPI softdrop mass peak as a function of jet p_T in two different η bins (smaller or greater than $|\eta| = 1.3$). No corrections have been applied to the softdrop mass.

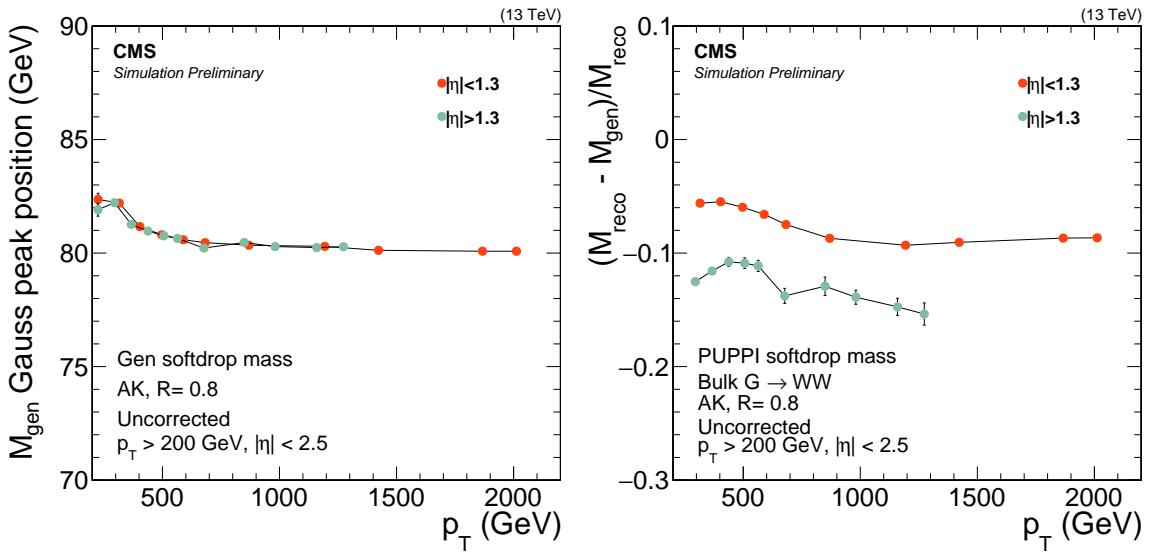


Figure 5.43: The mean of the fitted generator level W-jet softdrop mass distribution as a function of jet p_T (left) and the normalized difference in reconstructed and generated softdrop mass (right).

is the reconstructed mass shown in the right plot in Figure 5.43 and M_{GEN} is as defined above, as a function of jet p_T in two η bins (smaller or greater than $|\eta| = 1.3$) is performed. Polynomial fit functions of the following forms are used

$$\begin{aligned} w(p_T) &= A + B(x^2)^{-C} && \sim \text{"gen correction"} \\ w(p_T) &= A + Bx + Cx^2 + Dx^3 + Ex^4 + Fx^5 && \sim \text{"reco correction"} \end{aligned}$$

The distribution and corresponding fits for the two weights is shown in Figure 5.44 for the ‘gen correction’ (left) and ‘reco correction’ (right). The two corrections are then applied to

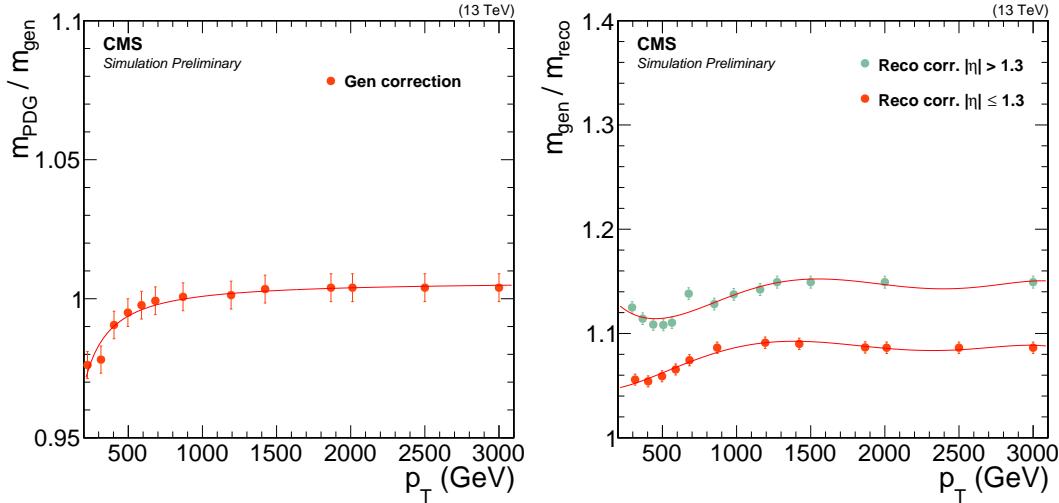


Figure 5.44: Fit to $M_{\text{PDG}}/M_{\text{GEN}}$ as a function of jet p_T (left), where $M_{\text{PDG}} = 80.4$ GeV and M_{GEN} is the fitted mean of the generator level mass and a fit to $(M_{\text{RECO}} - M_{\text{GEN}})/M_{\text{RECO}}$ (right), where M_{RECO} is the reconstructed softdrop mass, as a function of jet p_T in two η bins.

the uncorrected PUPPI softdrop mass both in data and in MC as

$$M_{SD} = M_{\text{SD,uncorr}} \times w_{\text{GEN}} \times w_{\text{RECO}} \quad (5.9)$$

where w_{GEN} and w_{RECO} correspond to the gen and reco corrections respectively and $M_{\text{SD,uncorr}}$ is the uncorrected PUPPI softdrop mass.

Finally, a closure test is performed in order to check that the corrected PUPPI+softdrop W-jet mass peaks at 80.4 GeV and is stable with p_T and η . The fitted mean of the corrected PUPPI softdrop mass peak as a function of jet p_T in two different η bins is shown in Figure 5.45. Good closure is observed, with the corrected mass peaking around 80 GeV independent of the jet p_T and η . The PUPPI softdrop jet mass peak for W/Z-jets from different signal samples after jet mass corrections have been applied is shown in Figure 5.46, for resonances with a mass of 1 and 4 TeV. The corrections applied to Z-jets yield a mass stable with p_T , peaking around the Z mass.

W-tagging performance

The new PUPPI+softdrop based W/Z-tagger uses a mass window of $65 \text{ GeV} < m_{SD} < 105 \text{ GeV}$ in combination with a cut of PUPPI $\tau_{21} < 0.4$. We compare its performance to that of the CHS+pruning based tagger used in Search I as well as to that of a ‘DDT-transformed’ τ_{21} based tagger [57]. The τ_{21}^{DDT} variable is a linear transformation of τ_{21} given as

$$\tau_{21}^{DDT} = \tau_{21} + M \times \log \left(\frac{m^2}{p_T \times 1 \text{ GeV}} \right) \quad (5.10)$$

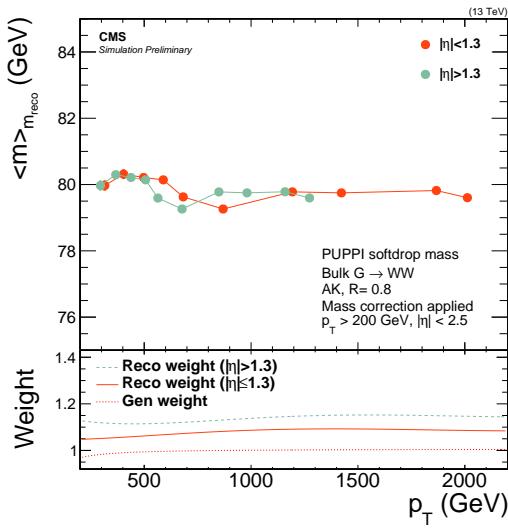


Figure 5.45: The mean of the fitted W-jet corrected PUPPI softdrop mass peak as a function of jet p_T in two different η bins.

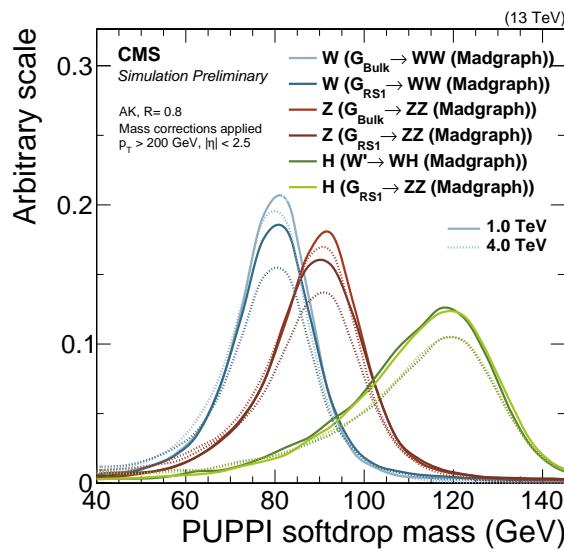


Figure 5.46: The W/Z/H-jet corrected PUPPI softdrop mass peak for jets from different signal samples with masses of 1 and 4 TeV.

where $M = -0.063$ is obtained from a fit of τ_{21} against the variable $\rho' = \log(m^2/p_T/\mu)$, where $\mu = 1 \text{ GeV}$. The purpose of this is to de-correlate τ_{21} from the softdrop mass and p_T , yielding a mass and dijet invariant mass spectrum minimally sculpted by a cut on the τ_{21}^{DDT} tagging variable. This is tagger that will be further explored and explained in detail in the context of Search III, Section 5.3.6.

The background rejection efficiency for QCD light flavor jets as a function of W-jet signal efficiency is shown in Figure 5.47. The efficiency is measured requiring a fixed jet mass window of 65-105 GeV, while scanning the cut on τ_2/τ_1 . The general performance of each tagger

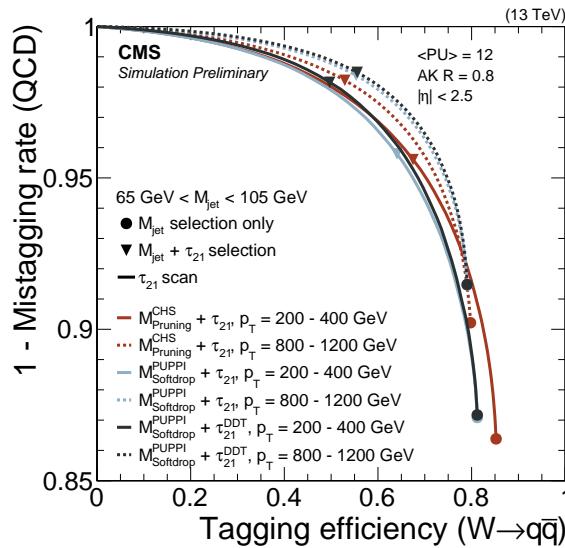


Figure 5.47: The background rejection efficiency for QCD light flavor jets as a function of W-jet signal efficiency. A cut on CHS pruned or PUPPI softdrop jet mass of $65 < m_{\text{jet}} < 105 \text{ GeV}$ is applied while scanning the cut on τ_{21} . The cuts corresponding to $\tau_2/\tau_1 < 0.45$ for CHS+pruning, PUPPI $\tau_2/\tau_1 < 0.4$ for PUPPI+softdrop or $\tau_{21}^{DDT} < 0.52$ are indicated with triangles, while the solid circles represent the efficiency and mistag rate for a mass cut only.

is very similar, with the PUPPI+softdrop based taggers displaying a slightly higher signal efficiency for a given mistag rate at high p_T and CHS+pruning slightly better at low p_T . To better understand the difference between each tagger, we look at the tagging performance as a function of jet p_T as well as pileup, shown in Figure 5.48 and 5.49.

Starting with the tagger p_T -dependence in Figure 5.48, we observe that the signal efficiency of a PUPPI+softdrop of CHS+pruned jet mass cut is flat as a function of p_T , at around 80%. The QCD mistag rate drops for both groomers, with a 1-3% lower mistag rate using PUPPI+softdrop than CHS+pruning. Once applying an n-subjettiness cut, the signal efficiency as well as the mistag rate for the PUPPI τ_{21} and CHS τ_{21} taggers drops as a function of p_T , with an average signal efficiency of around 50% for a $\sim 2\%$ mistag rate. An interesting behavior is observed for the τ_{21}^{DDT} tagger: While the mistag rate is flat as a function of p_T , as is the purpose of decorrelated taggers, the signal efficiency improves as the p_T increases, outperforming the other taggers above 1 TeV.

Turning to the tagger pileup dependence, shown in Figure 5.49, the expected benefit from using the PUPPI algorithm is observed: The tagging efficiency for the CHS+pruning (red solid circles) based tagger falls off steeply versus the number of primary vertices in the event, while the PUPPI+softdrop based taggers (light and dark blue solid circles) are more or less insensitive to pileup. Based on general performance, tagging stability versus pileup and due to theoretical considerations, PUPPI softdrop mass with dedicated mass corrections applied

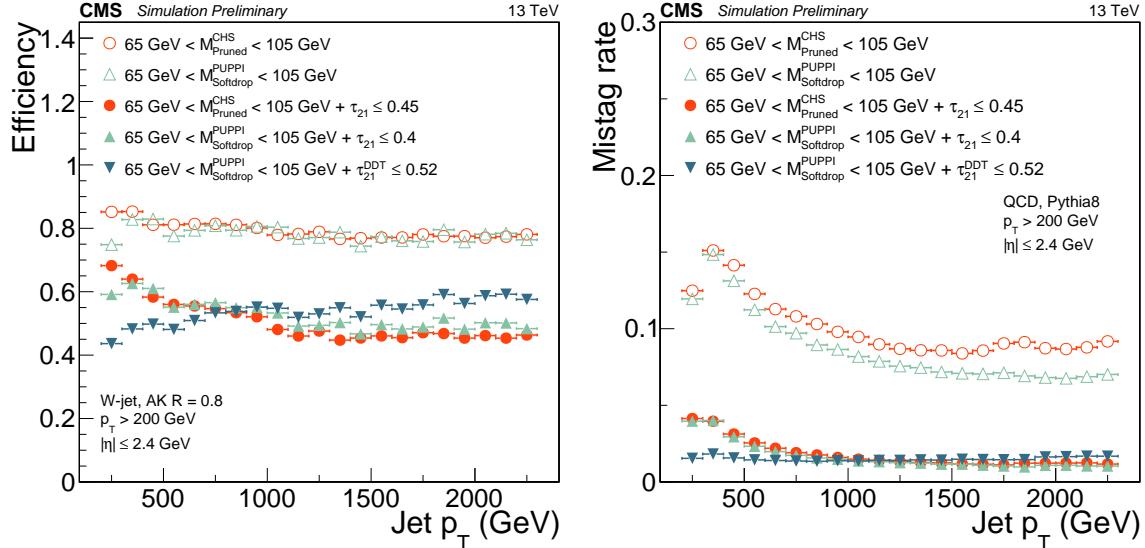


Figure 5.48: W-jet efficiency (left) and QCD light jet mistag rate (right) for a PUPPI+softdrop or CHS+pruned jet mass selection only (hollow circles) and the combined $m_{\text{jet}} + (\text{PUPPI}) \tau_2/\tau_1$ (DDT) selection (solid circles) as a function of jet p_T .

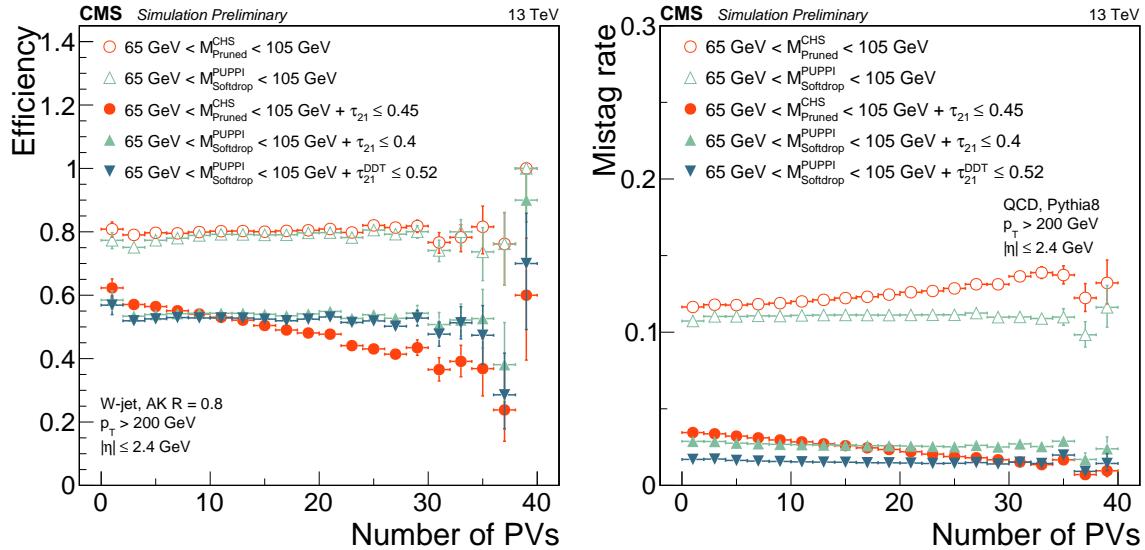


Figure 5.49: W-jet efficiency (left) and QCD light jet mistag rate (right) for a PUPPI+softdrop or CHS+pruned jet mass selection only (hollow circles) and the combined $m_{\text{jet}} + (\text{PUPPI}) \tau_2/\tau_1$ (DDT) selection (solid circles) as a function of jet pileup.

together with PUPPI τ_{21} is chosen as this analysis W-tagger. The per-jet efficiency is around 50-55% for a 1-2% mistag rate.

Efficiency scale factors and mass scale/resolution measurement

In order to measure the W-tagging efficiency, jet mass scale and resolution for the new PUPPI+softdrop based tagger, we use the same procedure as outlined in Section 5.1.7. We first did an early measurement of the efficiency using 2.3 fb^{-1} of data collected in 2015, which was published in

a jet algorithms performance note [55] and served as the first commissioning of the new tagger. We then redid the measurement with 12.9 and 35.9 fb^{-1} of 2016 data, respectively, for the two analyses presented in this chapter (the latter measurement performed by a separate analysis team). The results shown in the following will be those obtained during the commissioning of the tagger, while those used in the two analyses are listed in Appendix B. In order to better understand the differences between the CHS+pruning and PUPPI+softdrop based taggers, the first efficiency measurement was done in parallel for both algorithms, requiring either a softdrop or a pruned jet mass between 40 GeV and 150 GeV. The softdrop mass is computed after PUPPI and the jet mass corrections as described in Section 5.2.5 are applied, while the pruned mass is corrected with L2L3 jet energy corrections. The method is the same as the one outlined in detail in Section 5.1.7 and fits to matched $t\bar{t}$ MC and minor backgrounds for the PUPPI softdrop based tagger are skipped here and can be found in Appendix B. The PUPPI softdrop jet mass and PUPPI τ_{21} variables in data and in MC are shown in Figure 5.50 and can be compared to the corresponding plots for the CHS pruned jet mass and CHS τ_{21} distributions in Figure 5.50. The data to MC agreement as well as the observed spectra, is very similar between CHS pruning and PUPPI softdrop in this region.

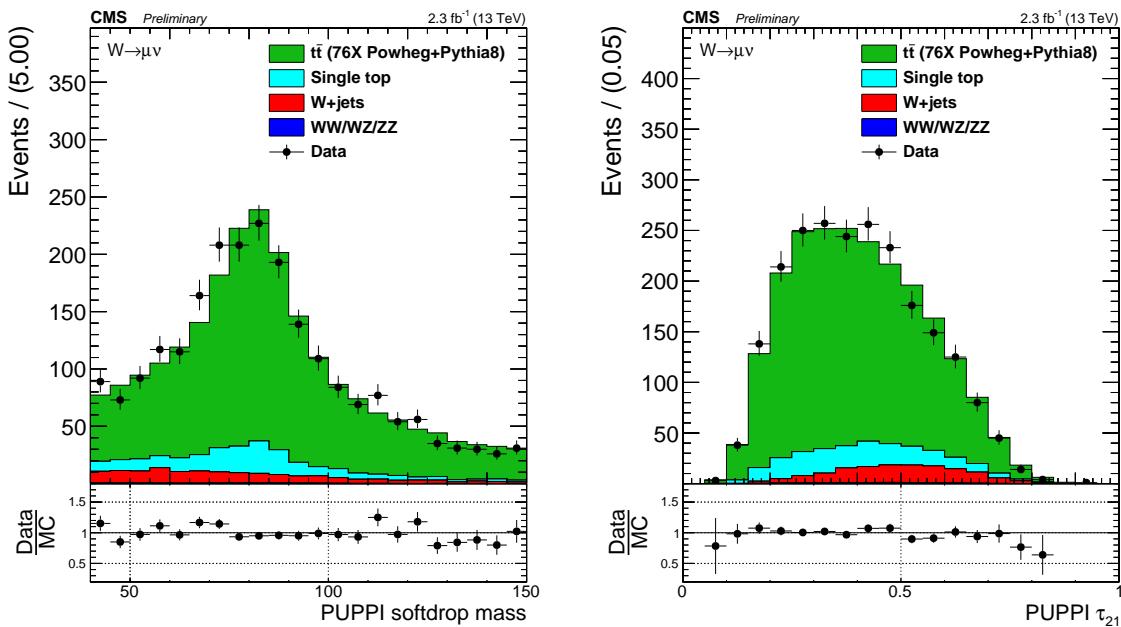


Figure 5.50: Distribution of the PUPPI softdrop mass (left) and PUPPI n-subjettiness (right) distribution in the $t\bar{t}$ control sample.

Following what was done in Section 5.1.7, we extract and compare the W-tagging efficiency, jet mass scale and resolution of the combined jet mass and τ_{21} selection in data and in MC. This is done through a simultaneous fit of the the softdrop jet mass spectrum between 40 and 150 GeV in two regions:

- Pass region: $0 < \tau_{21} \leq 0.40 \sim \text{high purity}$
- Fail region: $0.40 < \tau_{21} \leq 0.75 \sim \text{low purity}$

The corresponding fits are shown in Figure 5.51, with the corresponding extracted efficiencies from the Gaussian component of the total fit and scale factors summarized in Table 5.9. The quoted systematic uncertainties are evaluated the same was as described in Section 5.1.7 and correspond to the uncertainty due to ME generator and due to choice of fit method.

Category	Working point	Eff. data	Eff. simulation	Scale factor
HP	$\tau_2/\tau_1 < 0.4$	0.785 ± 0.045	0.81 ± 0.01	0.97 ± 0.06 (stat) ± 0.04 (sys) ± 0.06 (sys)
LP	$0.45 < \tau_2/\tau_1 < 0.75$	0.215 ± 0.057	0.204 ± 0.041	1.13 ± 0.24 (stat) ± 0.17 (sys) ± 0.12 (sys)

Table 5.9: W-tagging scale factors for both categories the high purity and low purity categories for two taggers: Pruned jet mass + τ_{21} and PUPPI softdrop jet mass + PUPPI τ_{21} .

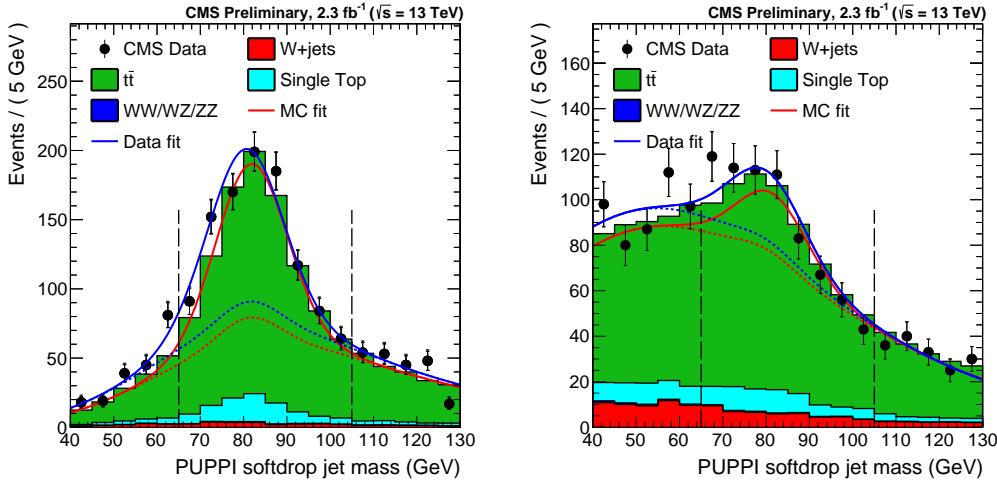


Figure 5.51: PUPPI softdrop jet mass distribution that pass (left) and fail (right) the PUPPI $\tau_2/\tau_1 < 0.40$ selection. Results of both the fit to data (blue) and simulation (red) are shown and the background components of the fit are shown as short-dashed lines. (!RATHER REPLACE BY 12.9INVFB MEASUREMENT)

Both scale factors are comparable to unity, within uncertainties. We additionally extract the jet mass scale and jet mass resolution from the mean and width, respectively, of the Gaussian component of the total fit in the pass region. These are summarized in Table 5.10. As for pruning (Table 5.7), we find that the W jet mass scale is larger in simulation than in data, of roughly 2%. The jet mass resolution, on the other hand, is larger in data, of roughly 7%, whereas for pruning the resolution is larger in simulation (11%). However, both are statistically insignificant and comparable with unity within uncertainties.

Parameter	Data	Simulation	Data/Simulation
PUPPI softdrop $\langle m \rangle$	80.3 ± 0.8 GeV	81.9 ± 0.01 GeV	0.98 ± 0.01
PUPPI softdrop σ	9.0 ± 0.9 GeV	8.5 ± 0.4 GeV	1.07 ± 0.12

Table 5.10: Summary of the fitted W-mass peak fit parameters.

The W-tagging efficiency scale factors, jet mass scale and resolution affects the signal yield and are included as described in Section 5.1.7: as a scale of the total signal yield and an uncertainty on the signal efficiency due to a shift and broadening of the W-jet mass peak.

5.2.6 W-tagging mistagging rate measurement

We additionally measure the W-tagging fake rate in data in a QCD dijet enriched region and compare this to the prediction from QCD MC using the three different combination of generators: HERWIG++, PYTHIA and MADGRAPH +PYTHIA. Figure 5.52 shows the mistag rate as a function of p_T for three different taggers: CHS pruning + τ_{21} , PUPPI softdrop +

PUPPI τ_{21} and PUPPI softdrop + τ_{21}^{DDT} . We find a substantial difference in the modeling

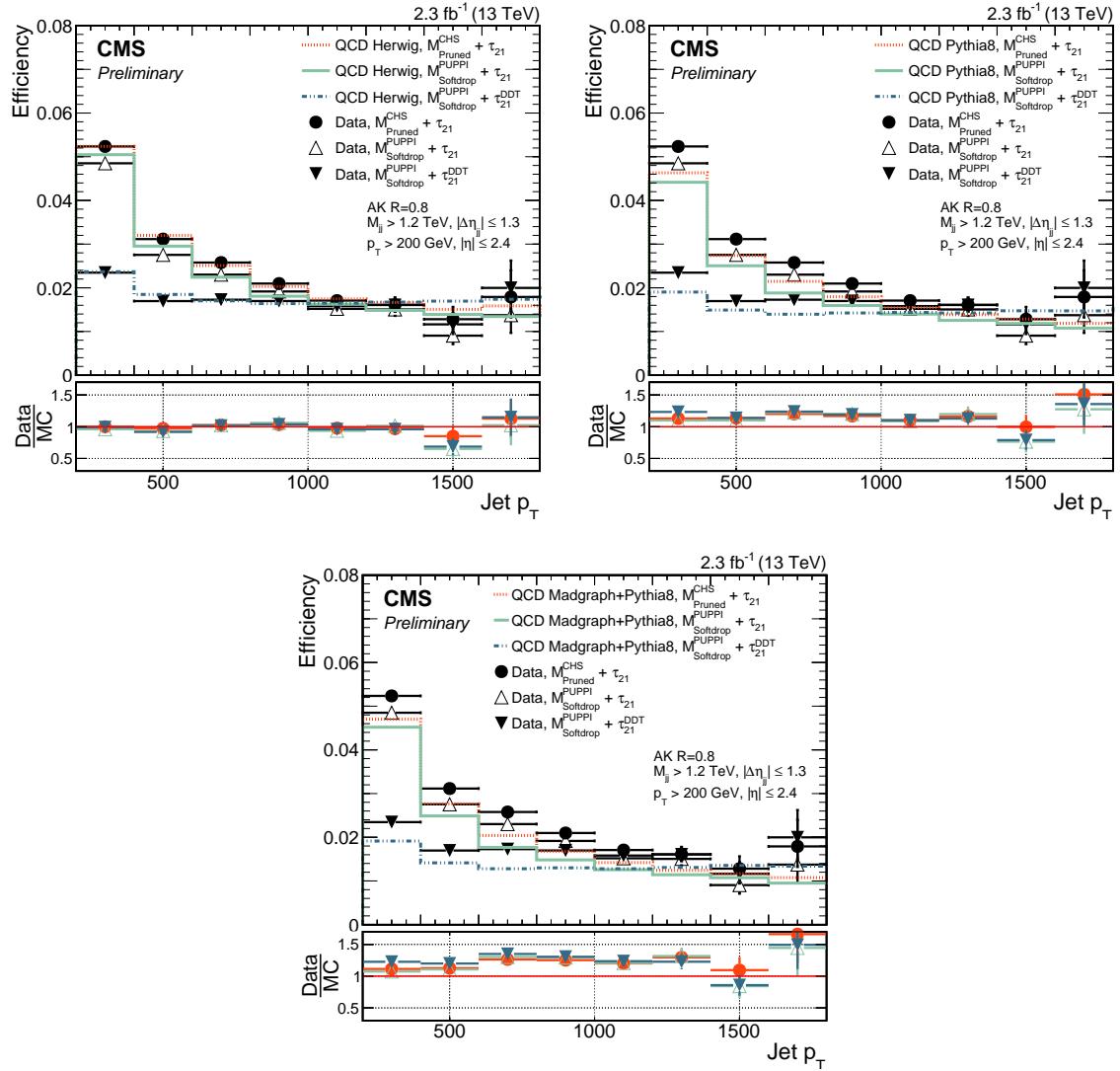


Figure 5.52: The fraction of jets that pass the $m_{\text{jet}} + \tau_2/\tau_1$ selections in a dijet enriched sample for data and for simulation as a function of jet p_T . Here comparing HERWIG++ (left), PYTHIA8 (right) and PYTHIA8 with MADGRAPH as matrix-element generator (left).

of substructure variables between the different generators, most likely coming from their very different description of gluon radiation (dominant in QCD multijet events). The best description is obtained with HERWIG++, while all three generators model the tagging p_T dependence well. We additionally study the difference in the total quark/gluon-content for the two PUPPI softdrop based taggers: τ_{21} and τ_{21}^{DDT} . Figure 5.53 shows the stacked relative q/g content in a Pythia 8 QCD dijet sample for a cut on PUPPI τ_{21} and τ_{21}^{DDT} . We see that the quark content increases as a function of jet p_T when cutting on the DDT, while it decreases when cutting on τ_{21} . This can be attributed to the fact that the m/p_T distribution for quark and gluon jets are very different from one another, and this difference increase as the jet p_T increases. Figure 5.54 shows the m/p_T for jets originating from a quark (blue) and a gluon (red) for a jet p_T of 200 GeV (left) and 1600 GeV. We see that the mass over p_T for gluon jets is significantly higher for gluon jets than for quark jets. With the τ_{21}^{DDT} tagger being defined as in Equation 5.10, the DDT will therefore act more aggressive on jets with a

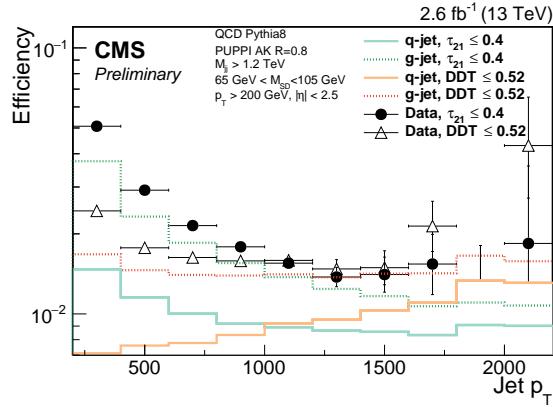


Figure 5.53: The fraction of jets that pass the PUPPI softdrop m_{jet} with τ_2/τ_1 (turquoise) or τ_{21}^{DDT} (orange) selections in a dijet enriched sample. The jets from QCD MC are split into two contributions: jets originating from gluons (dotted line) and jets originating from quarks (solid line).

high m/p_T , effectively removing more gluon jets.

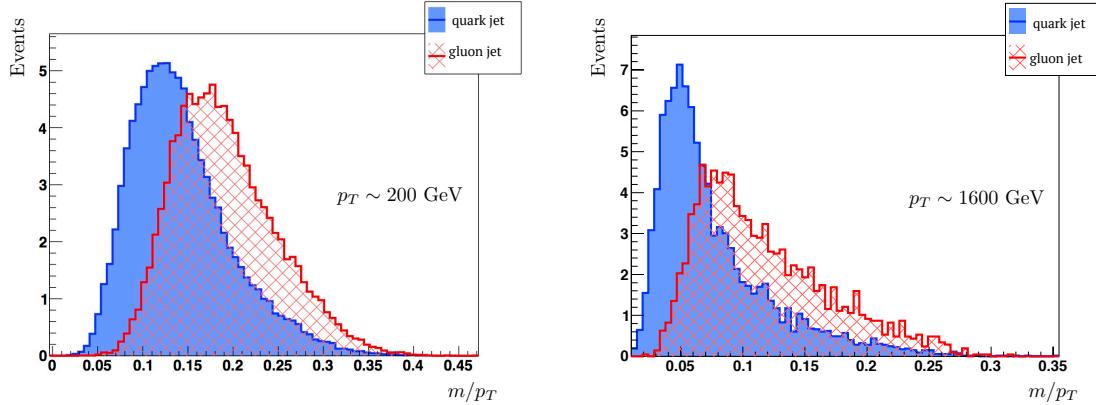


Figure 5.54: The jet mass divided by the jet p_T for quark (blue) and gluon (red) jets for a jet p_T of 200 (left) and 1600 GeV (right). Created with [58].

5.2.7 Mass and purity categorization

The PUPPI softdrop jet mass and PUPPI τ_{21} distribution after loose analysis preselections, as outlined in Section 5.2.3, are shown in Figure 5.55. We see some disagreement between data and MC, especially in the high-purity region ($\text{PUPPI } \tau_{21}^{DDT} < 0.4$), confirming what we observed in Section 5.2.6 As this analysis is sensitive to both heavy resonances decaying

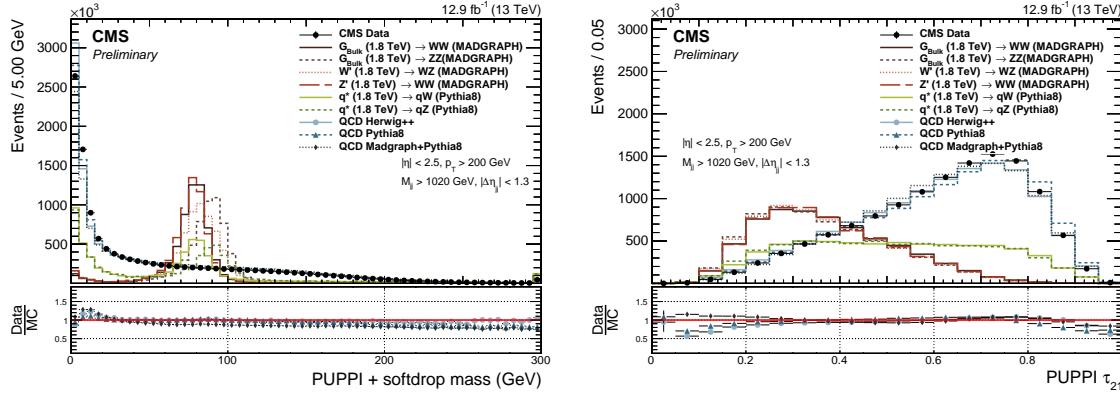


Figure 5.55: PUPPI softdrop jet mass distribution (left) and PUPPI n-subjettiness τ_{21} (right) distribution for data and simulated samples. Simulated samples are scaled to match the distribution in data.

into two vector bosons and excited quark resonances q^* decaying to qW and qZ, we look for events with both a single W/Z-tag and events with two W/Z-tags. Vector boson candidates are selected with a PUPPI softdrop jet mass of $65 < m_{\text{jet}} < 105$ GeV. Further, and similar to what was done in Search I, we select ‘high purity’ (HP) W/Z jets by requiring PUPPI $0 < \tau_{21} \leq 0.40$ and ‘low purity’ (LP) jets with $0.40 < \tau_{21} \leq 0.75$. The events with one W/Z-tag are classified in HP and LP events according to the two categories described previously. Events with two W/Z-tagged jets are always required to have one HP tagged jet, and are further divided into LP and HP categories depending on whether the other jet is of high or low purity. We additionally split into two mass categories in order to enhance the analysis sensitivity, with the W window defined as $65 \text{ GeV} < m_{\text{pruned}} < 85 \text{ GeV}$ and the Z boson window as $85 \text{ GeV} < m_{\text{pruned}} < 105 \text{ GeV}$. This results in ten different signal categories. They are as follows:

- High-purity double W/Z-tag, 3 mass categories: WW, ZZ and WZ
- Low-purity double W/Z-tag, 3 mass categories: WW, ZZ and WZ
- High-purity single W/Z-tag, 2 mass categories: qW and qZ
- Low-purity single W/Z-tag, 2 mass categories: qW and qZ

5.2.8 Background modeling: F-test

With the full analysis selections and categorization defined, we move to the determination of background fit function. Following the same strategy as in Section 5.1.5, we determine the number of necessary parameters in order to describe the background through a Fishers F-test, comparing the same fit functions as in Section 5.1.5. This test is first exercised in QCD MC and then in a data sideband before the final determination in the data signal region. As the F-test results were presented in detail in the context of Search I, only a brief summary and

the fits in the new single-tag categories will be presented here, while all fits and F-test results can be found in Appendix D.

A two or three parameter fit is sufficient to describe the background for all the double tag categories: a two parameter fit is sufficient for the ‘high-purity’ WZ and ZZ categories, as well as the ‘low-purity’ WW category, while the remaining analysis categories require a three parameter background fit. From the fits to the single tag categories, shown in Figure 5.56, a three parameter fit is sufficient for all categories except the ‘high-purity’ qW category. In this category the improvement in fit quality when increasing the number of parameters is so large adding that adding an additional fit parameter is justified, and we continue by using a 5 parameter fit for this category. A summary of what fit functions are used for each analysis

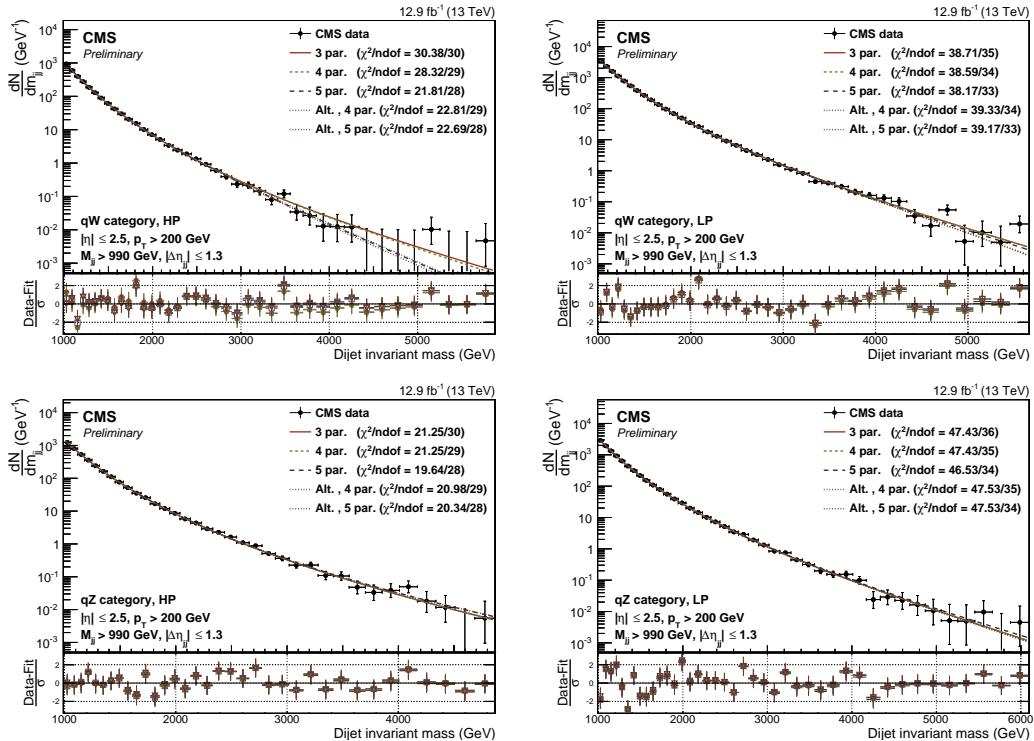


Figure 5.56: Background fit for the M_{jj} distribution in the data signal region for the single-tag analysis. Here for the high- (left) and low-purity (right) single W/Z-tag categories qW (top) and qZ (bottom).

category is listed in Table 5.11.

Mass category	N pars.	
	HP	LP
WW	3	2
WZ	2	3
ZZ	2	3
qW	5	3
qZ	3	3

Table 5.11: Fit parameters used in each analysis category

5.2.9 Signal modeling

The signal is modeled from signal MC in the same way as was done in Section 5.1.6, assuming a Gaussian core and an exponential tail. The interpolated signal shapes for $q^* \rightarrow qW$ and $q^* \rightarrow qZ$ in their most sensitive analysis categories (qW and qZ , respectively) are shown in Figure 5.57. The signal shapes for the double-tag category can be compared to those in Figure 5.22.

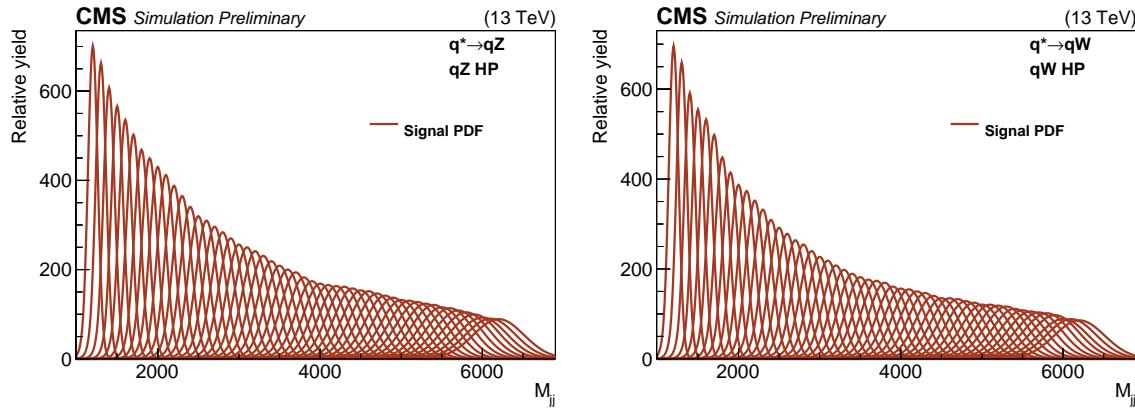


Figure 5.57: Interpolated signal shapes for a $q^* \rightarrow qZ$ (left) and $q^* \rightarrow qW$ (right) signal.

5.2.10 Systematic uncertainties

The largest sources of systematic uncertainty for this analysis is, as for Search I, related to the signal modeling and are due to the uncertainty in the tagging efficiency of the W/Z-tagger, the jet energy/mass scale, the jet energy/mass resolution and integrated luminosity. The W/Z tagging uncertainty is estimated in $t\bar{t}$ events, as described in Section 5.1.7, and yield uncertainties on the scale factors for the HP and LP tagging categories. The p_T - and η -dependent jet energy scale and resolution uncertainties on the resonance shape were approximated by a constant 2% and 10% uncertainty in Search I (Section 5.1.7) and are not expected to change for the 2016 analysis. The jet energy response and resolution uncertainty are taken into account as shape uncertainty by shifting and widening the signal resonance model, while all other signal uncertainties only affect the yield. The list of most relevant systematic uncertainties are listed in Table 5.12.

5.2.11 Results

As mentioned in the introduction to this chapter, the analysis of the 2016 dataset was done in two steps: One based on 12.9 fb^{-1} of early 2016 data, demonstrating the new PUPPI softdrop based tagger and single-tag analysis categories, and one topping up with the full 35.9 fb^{-1} dataset. The results from both will be presented in the following.

Early analysis

Exclusion limits are set in the context of the bulk graviton model, the HVT model B scenario and excited quark resonances, assuming the resonances to have a natural width negligible with respect to the experimental resolution (as in Search I).

Source	Relevant quantity	HP+HP unc. (%)	HP+LP unc. (%)
Jet energy scale	Resonance shape	2	2
Jet energy resolution	Resonance shape	10	10
Jet energy scale	Signal yield	<0.1–4.4	
Jet energy resolution	Signal yield	<0.1–1.1	
Jet mass scale	Signal yield	0.02–1.5	
Jet mass resolution	Signal yield	1.3–6.8	
Pileup	Signal yield	2	
Integrated luminosity	Signal yield	6.2	
PDFs (W')	Signal yield	4–19	
PDFs (Z')	Signal yield	4–13	
PDFs (G_{bulk})	Signal yield	9–77	
Scales (W')	Signal yield	1–14	
Scales (Z')	Signal yield	1–13	
Scales (G_{bulk})	Signal yield	8–22	
Jet mass scale	Migration	<0.1–16.8	
Jet mass resolution	Migration	<0.1–17.8	
W-tagging τ_{21}	Migration	15.6	21.9
W-tagging p_T -dependence	Migration	7–14	5–11

Table 5.12: Summary of the signal systematic uncertainties for the analysis and their impact on the event yield in the signal region and on the reconstructed dijet invariant mass shape (mean and width).

Figure 5.58 shows the 95% confidence level (CL) expected and observed exclusion limits on the signal cross section as a function of the resonance mass for the different signal hypotheses in the double-tag analysis. The limits are compared with the cross section times the branching fraction to WW and ZZ for a bulk graviton with $\tilde{k} = 0.5$, and with the cross section times the branching fraction to WZ and WW for spin-1 particles predicted by the HVT model B for both the singlet (W' or Z') and triplet (W' and Z') hypothesis. For the HVT model B, we exclude W' (Z') resonances with masses below 2.7 (2.6) TeV. The signal cross section uncertainties are displayed as a red checked band and result in an additional uncertainty on the resonance mass limits of 0.05 (0.04) TeV. The cross section limits for $Z' \rightarrow WW$ and $G_{\text{bulk}} \rightarrow WW$ are not identical due to the different acceptance for those two signal scenarios.

Figure ?? shows the corresponding exclusion limits for excited quarks decaying into qW and qZ . We exclude excited quark resonances decaying into qW and qZ with masses below 5.0 and 3.9 TeV, respectively. The signal cross section uncertainties are displayed as a red checked band and result in an additional uncertainty on the resonance mass limits of 0.1 TeV.

Full 2016 dataset

The results obtained with the full $\sim 36 \text{ fb}^{-1}$ of 2016 data are as follows: For a G_{bulk} we exclude production cross sections in a range from 36.0 fb, at a resonance mass of 1.3 TeV, to 0.6 fb at resonance masses above 3.6 TeV. W' (Z') resonances are excluded with masses below 3.2 (2.7) TeV for the HVT model B, in addition to W' resonances with a mass between 3.3 and 3.6 TeV. For excited quark resonances, we can exclude the production of q^* decaying to qW or qZ for masses below 5.0 and 4.7 TeV. Figure 5.60 and 5.61 show the resulting 95% confidence level expected and observed exclusion limits on the signal cross section as a function of the resonance mass for VV and QV resonances, respectively.

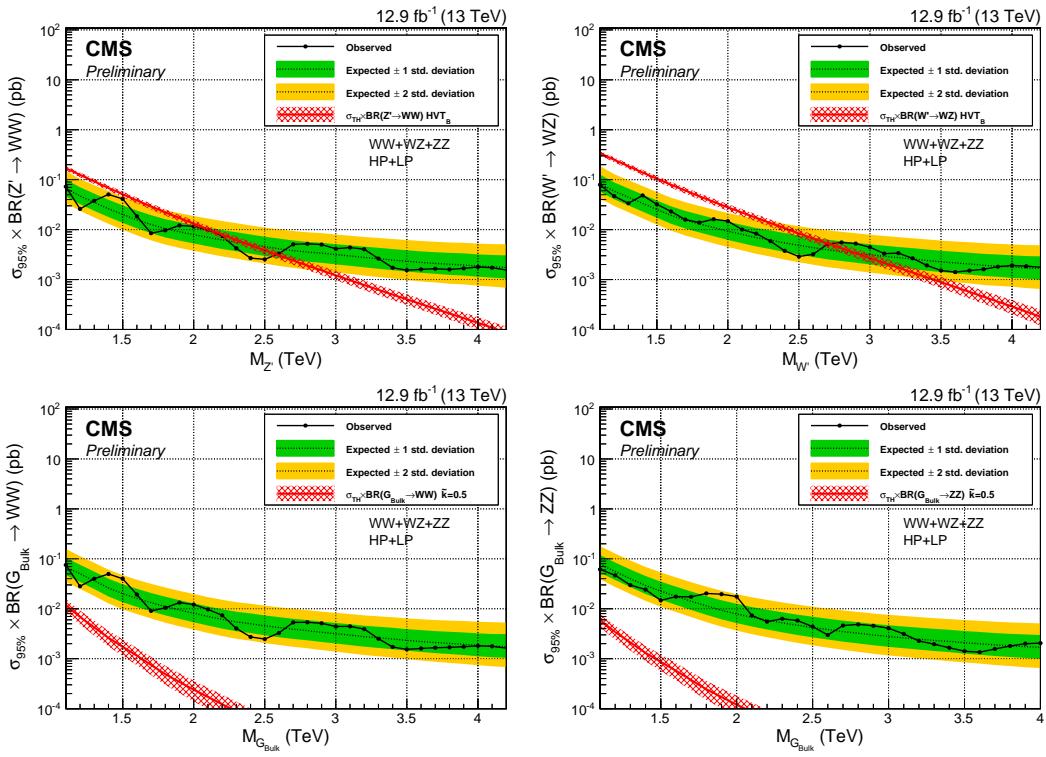


Figure 5.58: Observed (black solid) and expected (black dashed) 95% CL upper limits on the production of a narrow-width resonance decaying to a pair of vector bosons for different signal hypotheses. Limits are set in the context of a spin-1 neutral Z' (left) and charged W' (right) resonance, and compared with the prediction of the HVT model B. On the bottom, limits are set in the context of a bulk graviton decaying into WW (left) and ZZ (right) with $\tilde{k} = 0.5$ and compared with the model prediction. Signal cross section uncertainties are displayed as a red checked band.

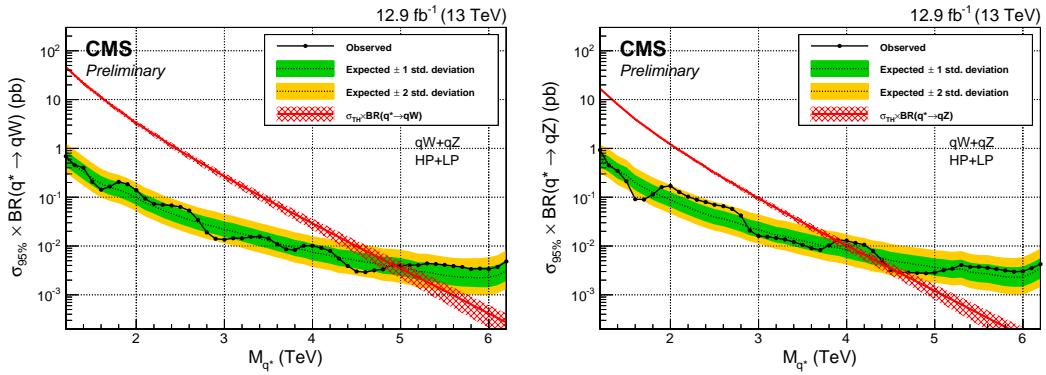


Figure 5.59: Observed (black solid) and expected (black dashed) 95% CL upper limits on the production of an excited quark resonance decaying into qW (left) or qZ (right). Signal cross section uncertainties are displayed as a red checked band.

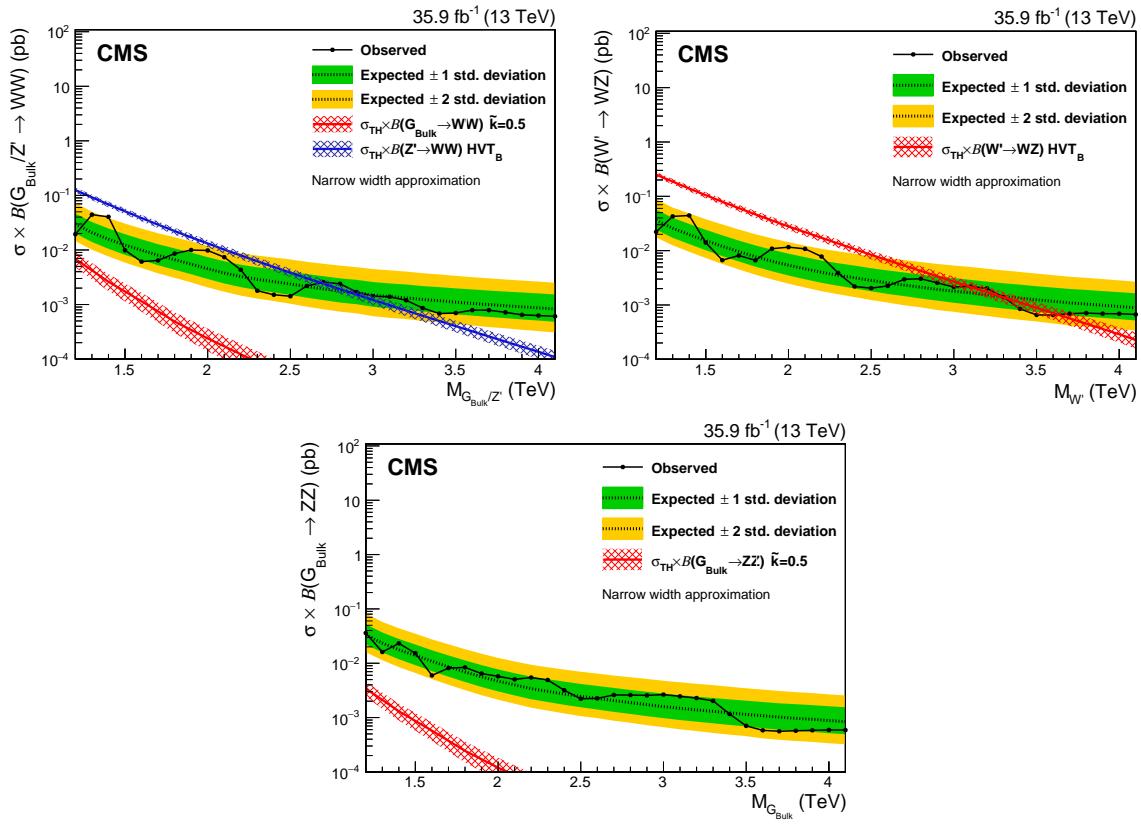


Figure 5.60: Observed (solid line) and expected (dashed line) 95% CL upper limits on the production cross section of a narrow resonance decaying into two vector bosons for different signal hypotheses: A Z' or G_{bulk} resonance decaying into WW (top left), a Z' decaying into WZ (top right) and a bulk graviton decaying into ZZ (bottom).

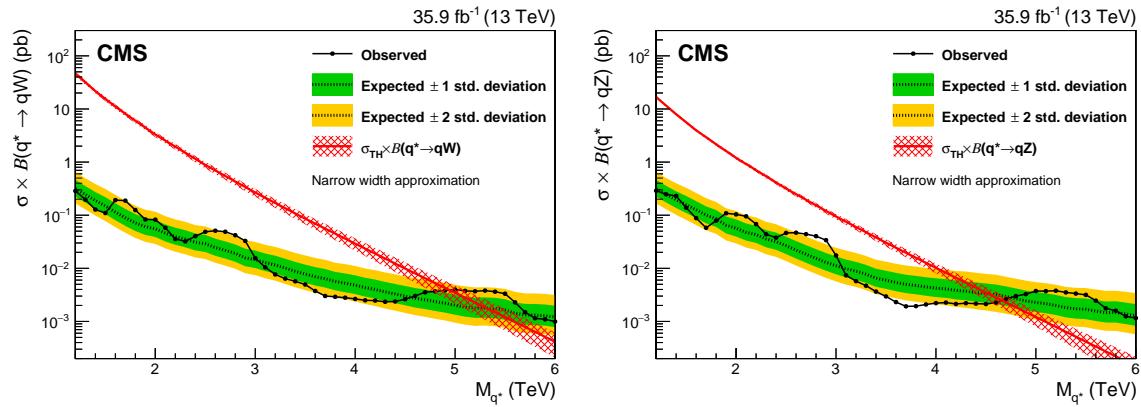
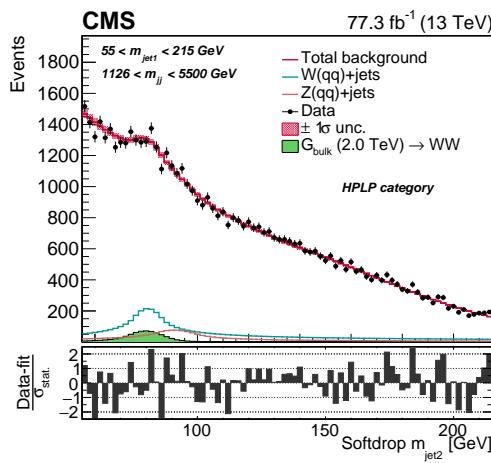


Figure 5.61: Observed (solid line) and expected (dashed line) 95% CL upper limits on the production of an excited quark resonance decaying into qW (left) or qZ (right).

5.3 Search III: A novel framework for multi-dimensional searches

After two successful analyses of 13 TeV data, no excess had confirmed the 8 TeV bump and the available phase space for New Physics to hide out was shrinking. However, this fact wouldn't disarm everybody. On the BSM theory front, ideas were simmering about whether it was possible that the small bumps we were observing here and there in the dijet invariant mass spectrum were due to us catching the tail of another type of boson with a mass slightly different from that of a W or a Z boson? And that perhaps these jets were not 2-prong objects, but in reality 4-prong? With Run 2 coming to an end, marking the beginning of a two year long shut-down, it was time to think about how we could probe alternative BSM models as effectively as possible. Our idea was therefore the following: We would build a novel framework capable of easily scanning the full softdrop jet mass and N-prong spectrum, and which, in addition, would lead to a gain in sensitivity for the standard VV all-hadronic search. We would do this by taking advantage of the fact that we were looking for bumps in a three-dimensional plane: the softdrop mass of the two jets as well as their invariant mass. As a validation of the method, the method would be demonstrated in the context of the VV all-hadronic search, replacing the dijet fit method. We would then extend this to simultaneously search for resonances decaying to $W(qq)$, $Z(qq)$ and $H(qq)$ and, finally, take full advantage of the framework and look for generic resonances peaking anywhere in the jet mass and dijet invariant mass spectrum.

Search III introduces a novel three-dimensional search method, allowing to simultaneously search for $W/Z/H$ signals, and eventually non-SM bosons, in the softdrop jet mass spectrum. It is the first analysis to measure the $V+jets$ cross section and the jet mass scale/resolution from a $W(qq)+jets$ and $Z(qq)+jets$ mass peak. Published with full 2016+2017 dataset, $\sim 80 \text{ fb}^{-1}$.



5.3.1 Small bumps and tri-bosons

Ever since the ATLAS observation of a 3.4σ excess in the search for VV resonances in the all-hadronic final state [33], several little bumps kept re-appearing near 2 TeV. These were not statistically insignificant, as we've already seen in Search I and Search II, but rather small elusive enhancements, illustrated by the collection of ATLAS/CMS observations in Figure 5.62. Due to their small size and the way the excesses seemed to slightly shift around,

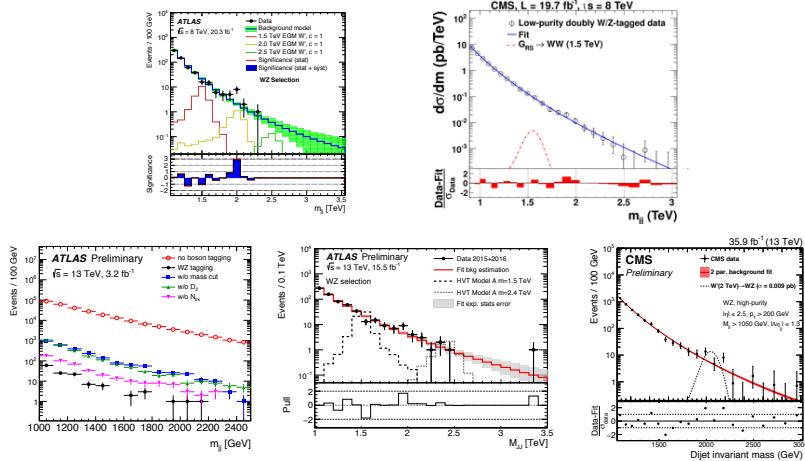


Figure 5.62: Several small bumps observed in VV resonance searches in the all-hadronic final state, both in ATLAS and in CMS [59].

these were obviously not diboson resonances. However, could they be caused by us catching the tails of some non-SM boson with a mass slightly different from that of a SM vector boson, as illustrated in Figure 5.63? Further, these could be 4-pronged objects rather than

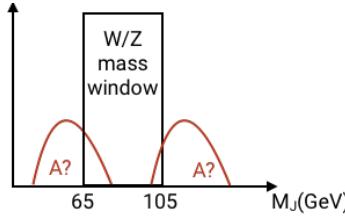


Figure 5.63: Slight excesses in diboson analyses could be caused by catching the tail of a non-SM object peaking at slightly higher/lower jet mass than at the W/Z mass.

2-prong, which would cause the excess to vary in size depending on the 4-prong efficiency of the analysis specific W-tagger used.

An explanation for the observed excesses was proposed in [60]. This paper pointed out that, if particles like W' and Z' exist, an extended scalar sector is needed in order to give mass to the vector bosons. These heavy scalars will decay to lighter bosons, if kinematically allowed, leading to multiboson signals from cascade decays. Some example signatures are illustrated in Figure 5.64. Signatures like these would peak in the groomed jet mass spectrum and, depending on what the final bosons decay into, have very different substructure profiles (4- and 4-prong, 2- and 4-prong etc.).

In order to effectively search for such types of signals, or any signal peaking in the softdrop jet mass spectrum, we therefore wanted to build a generic new framework allowing to look for peaks anywhere in the groomed mass - dijet invariant mass spectrum. Rather than selecting jets with a groomed mass between 65 and 105 GeV and look for resonances peaking in the

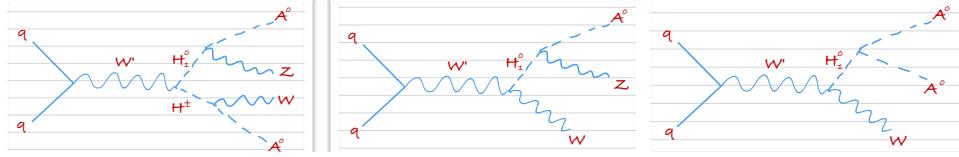


Figure 5.64: A W' decaying to a neutral H^0 and a charged H^\pm scalar particle leading to a quadriboson final state (left), and a W' decaying to a neutral scalar particle H^0 and a W leading to a triboson final state (middle and right) [60].

dijet invariant mass, we'd look for resonances peaking anywhere in the three dimensional plane formed by the groomed mass of each jet and their invariant mass. The benefits of this procedure, was that it would allow us to scan the full groomed mass spectrum in one analysis. We would first demonstrate the method through the VV all-hadronic analysis, which is the paper introduced here.

5.3.2 Analysis strategy

The background estimation used in Search I and Search II rely on a one dimensional fit of the dijet invariant mass signal region after a tight jet mass cut (65–105 GeV) has been applied. We now take advantage of the fact that the signal peaks in three dimensions; dijet invariant mass (M_{VV}) and the jet groomed mass of jet 1 and jet 2 ($M_{\text{jet}1}$ and $M_{\text{jet}2}$), and attempt to extract the signal from the three dimensional M_{VV} - $M_{\text{jet}1}$ - $M_{\text{jet}2}$ plane. The benefits of

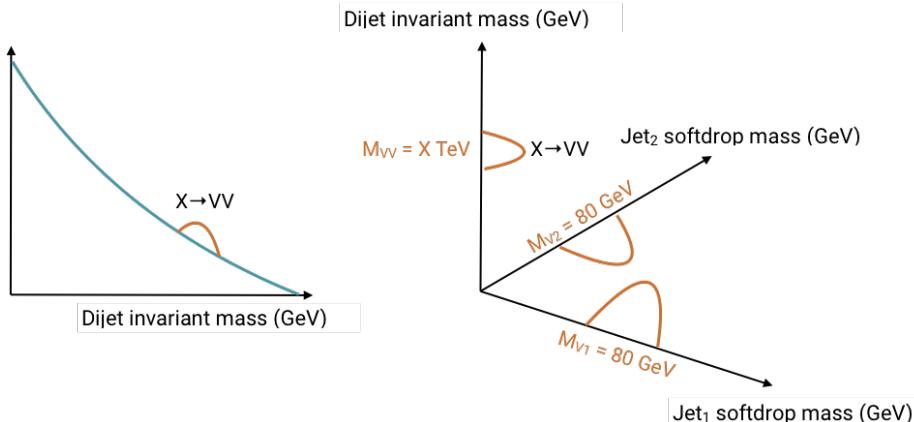


Figure 5.65: The one dimensional VV analysis versus the three dimensional fit.

doing so is that we now can perform different searches in the WW , WZ , ZZ , WH or XX final states encoded in the same analysis. Additionally, tight jet mass cuts are no longer needed as we fit the full jet mass line-shape to extract the signal. This effectively increases our signal statistics as a large fraction of the W and Z signal fall outside the above mass window. Fitting the jet groomed mass and resonance mass together also allow us to add nuisance parameters that simultaneously affect the jet groomed mass and the resonance mass, fully accounting for the correlation between the variables. We would model the background starting from simulation, rather than the dijet fit to data, which would allow us to model peaky distributions like a trigger turn-on. This could allow the search to go even lower in M_{VV} , something we will discuss further in Section 5.3.5.

This chapter, and its corresponding publication, is an analysis of the 2016 and 2017 dataset, corresponding to $\sim 80 \text{ fb}^{-1}$, and serves as the first documentation and demonstration of the

novel three-dimensional fit method. In Section 5.4, I'll discuss how we plan to take this framework further in searches for VH and HH as well as for generic resonances peaking in jet softdrop mass.

5.3.3 Data and simulated samples

The data analyzed in this search consists of 35.9 fb^{-1} of data collected in 2016 and 41.4 fb^{-1} of data collected in 2017, yielding a total of 77.3 fb^{-1} .

The simulated samples are the same as those described in Section 5.2.3, with specific detector conditions to match the 2016 and 2017 dataset.

5.3.4 Event selection

Events are selected following the same criteria as in Search I and Search II (see Section 5.1.4) and can be summarized as follows:

- PF jet Tight ID applied
- Jet $\eta < 2.5$
- Jet $p_T > 200 \text{ GeV}$
- $|\Delta\eta|_{jj} < 1.3$

The two jets with the highest groomed jet mass in the event are selected as potential vector boson candidates. In addition, the dijet invariant mass is required to be $> 1126 \text{ GeV}$ in order to be on the trigger plateau. As already mentioned in the introduction, the background modeling used in this analysis is capable of modeling turn-ons and is something we explored. However, while the background modeling was reliable, we found it difficult to extract a signal peaking on top of a turn-on and had to abandon the trigger modeling for this first demonstration of the method. More details will be given in Section 5.3.5.

The dijet invariant mass and $|\Delta\eta|_{jj}$ distribution for the two leading jets in the event after the above preselections have been applied is shown in Figure 5.66. The jet p_T and η distributions for signal and for background is shown in Figure 5.67.

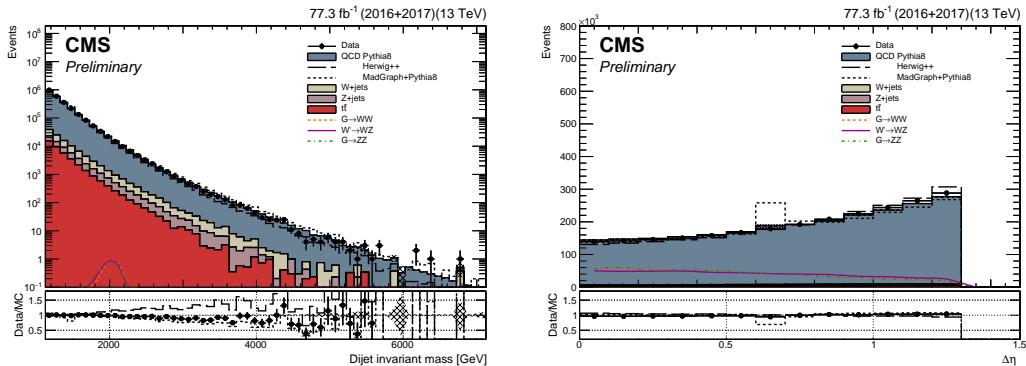


Figure 5.66: The dijet invariant mass (left) and $|\Delta\eta|_{jj}$ (right) for the two leading jets after preselections are applied. The signal is scaled by an arbitrary number.

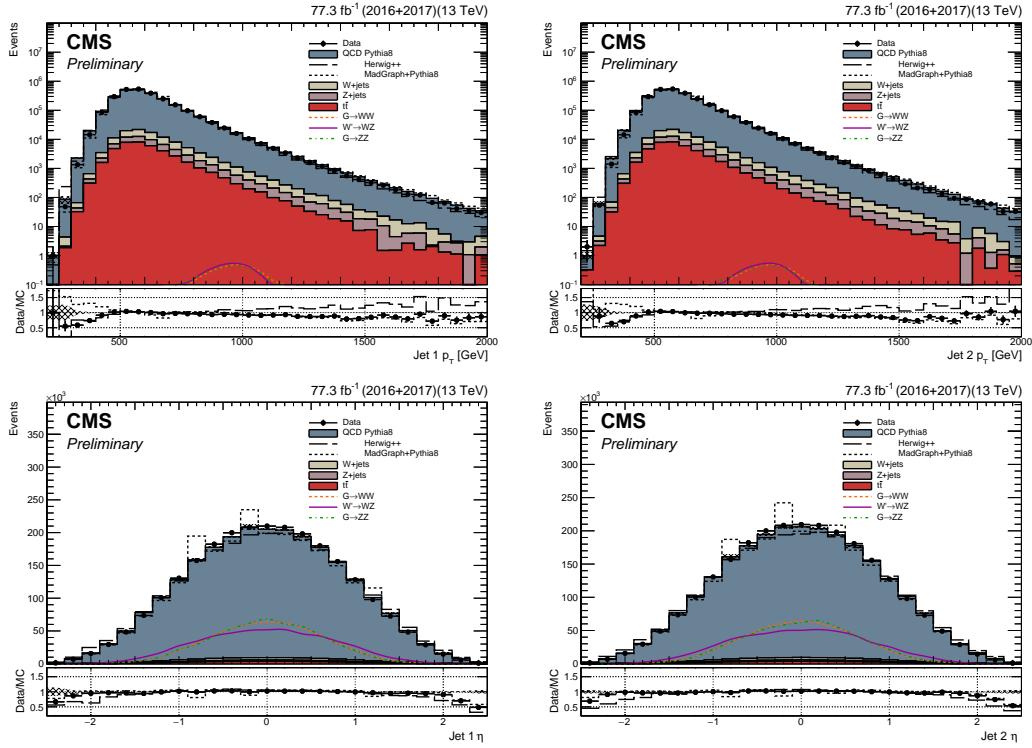


Figure 5.67: Jet p_T (top) and η (bottom) of the leading (left) and second leading (right) jet in the event. The signal is scaled by an arbitrary number.

5.3.5 Triggering

The triggers used for 2016 data are the same as in Section 5.1.4, while the thresholds in 2017 have increased in order to push the trigger rate to a level acceptable for the increased luminosity. The triggers used for 2017 data are

- HLT_PFHT1050
- HLT_AK8PFJet500
- HLT_AK8PFJet360/380/400/420_TrimMass30
- HLT_AK8PFHT750/800/850/900_TrimMass50

. For the results presented here, the analysis threshold is set by the trigger turn-on point (where the combination of all triggers are > 99 percent efficient). The trigger turn-on is evaluated in the Single Muon dataset, using the HLT_Mu50 and HLT_IsoMu27 triggers as reference triggers. The trigger turn-on curves as a function of dijet invariant mass and jet soft drop mass are shown in Figure 5.68. The combination of all triggers are $> 99\%$ efficient above a dijet invariant mass of 1126 GeV and this sets the analysis threshold.

Trigger turn-on modeling

The analysis threshold for searches depending on the dijet fit is set by the trigger turn-on point as the analysis relies on a background fit of the dijet invariant mass spectrum with a smoothly falling function. As the background modeling for this analysis does not depend on a smoothly falling spectrum (as will be described in detail in Section ??), and in order to compensate for

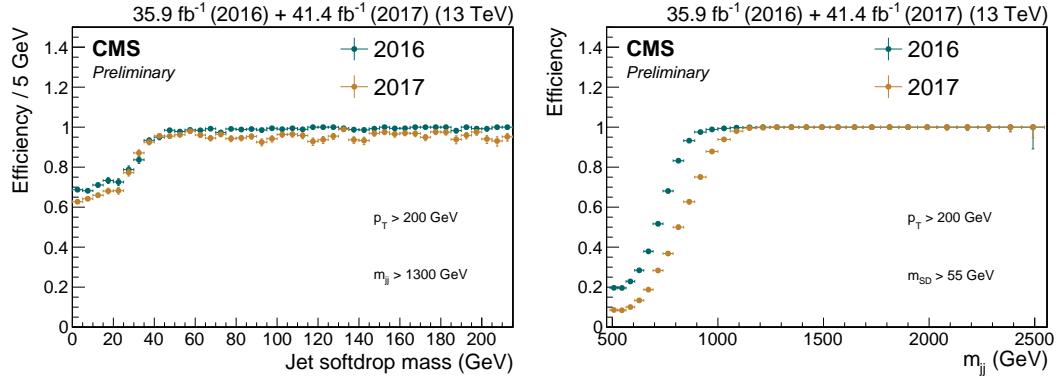


Figure 5.68: Trigger turn-on curves in the 2016 and 2017 datasets for the H_T based (left) and groomed mass based (right) trigger paths.

a loss in acceptance due to increased trigger thresholds, we attempted to model the trigger turn-on directly from data and apply a trigger weight in simulation. To do so, we first study the one dimensional trigger turn-ons versus dijet invariant mass and softdrop jet mass to understand where in M_{VV} and M_{jet1}/M_{jet2} the triggers are fully efficient. We then derive a three dimensional histogram of the trigger efficiency versus dijet invariant mass (M_{VV}) and the jet groomed mass of jet 1 and jet 2 (M_{jet1} and M_{jet2}), where each bin corresponds to the trigger efficiency for a given value of $M_{VV}-M_{jet1}-M_{jet2}$. The procedure is as follows: From the one dimensional histograms, the points of full efficiency versus M_{VV} , M_{jet1} and M_{jet2} are defined. For every bin above this threshold, the trigger efficiency is fixed to one (100 percent efficiency). For all bins below this trigger threshold, we fit slices of M_{VV} with a sigmoid function, evaluate the trigger weights from this function, and set the bin content of the three dimensional weight histogram accordingly. For every bin below this point, the trigger weight is extracted by fitting slices of M_{VV} in each bin of M_{jet1}/M_{jet2} . As the trigger efficiency falls below 50 percent around a dijet invariant mass of 800 GeV (Figure 5.68), searching for resonances with masses below this point is non-feasible. In addition, the full signal shape needs to be contained within the dijet invariant mass spectrum, excluding resonance masses of 0.8 and 0.9 TeV. The lowest mass point signal sample to set limits on for this analysis is therefore at 1 TeV and the analysis threshold is fixed at $M_{VV} = 900 \text{ GeV}$ in order to fully contain the signal shape. Starting from $M_{VV} = 893 \text{ GeV}$ (dijet bin closest to 900 GeV) and $M_{jet1}/M_{jet2} = 40 \text{ GeV}$, a coarsely binned three-dimensional histogram (10 GeV M_{jet1}/M_{jet2} binning and ‘dijet’ binning in M_{VV}) is filled with the fraction of events that pass one of the signal triggers,

$$w_{ijk}^{Bin} = \frac{\text{PASS}(m_{jj}^i - m_{j1}^j - m_{j2}^k)}{\text{ALL}(m_{jj}^i - m_{j1}^j - m_{j2}^k)}.$$

The resulting coarse histogram is then expanded in M_{VV} to 10 GeV dijet invariant mass bins, interpolating between the points using sigmoid fit for each $M_{jet1} - M_{jet2}$ bin (the fine binning in dijet invariant mass is sufficient enough to yield a smooth reweighted distribution so no expansion is done in M_{jet1}/M_{jet2}). From this histogram, each slice in M_{VV} is fitted with a sigmoid function,

$$s(x) = \frac{1}{1 + e^{-p_1(x-p_2)}}$$

the trigger weight of each bin $M_{VV}-M_{jet1}-M_{jet2}$ is extracted from the fit and set accordingly. Figure 5.69 shows the total projections on each axis for the full trigger weight histogram. The M_{jet} and M_{VV} spectra for the lowest mass-point signal sample and for the QCD background

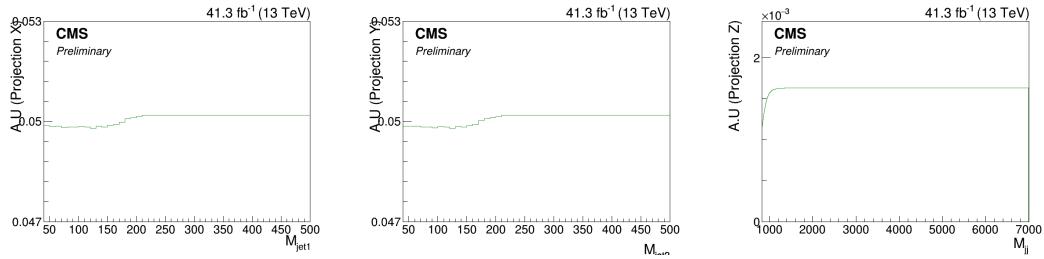


Figure 5.69: One-dimensional projections of the trigger weight histogram for $M_{\text{jet}1}$, $M_{\text{jet}2}$ and M_{VV} respectively.

before and after trigger weights are applied, are shown in Figure 5.70 and are compared to data. > 95% of the signal efficiency is retained, and reweighted QCD simulation agrees well with data. The modeling of the trigger turn-on was successful and implemented in

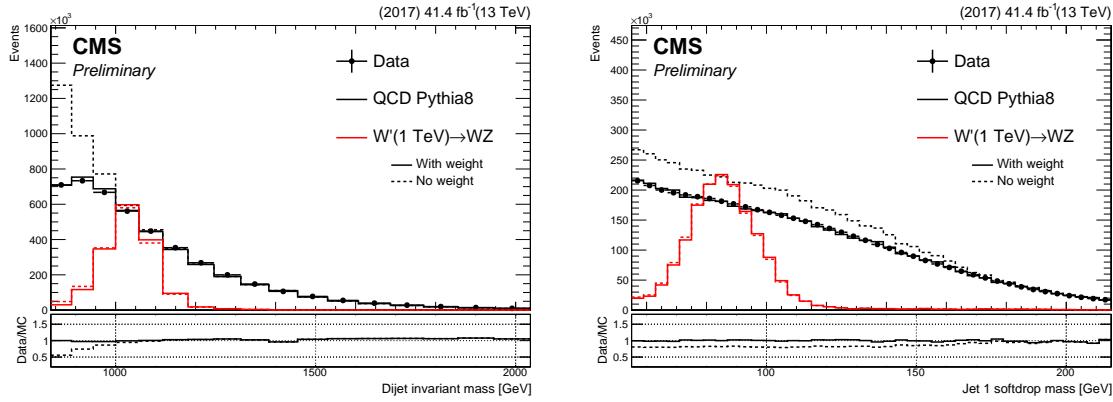


Figure 5.70: The M_{VV} (left) and M_{jet} (right) spectra for signal and background before and after trigger weights are applied.

the background fit method in the 3D analysis. We found that the method could model the turn-on well. However, when studying the bias on the extracted signal rate for a possible signal in this turn-on region, we found that this was large due to us attempting to fit a peak on top of a peaky background. As we wanted this analysis describing the 3D fit method to become available as soon as possible, we therefore abandoned modeling of the trigger turn-on for this paper. However, we still wish to pursue this strategy in the future.

5.3.6 A mass and p_T decorrelated tagger

In order to identify W and Z jet candidates, two algorithms are run on the AK8 PUPPI jet: softdrop [?] and the N-subjettiness ratio τ_{21} [?]. The softdropped jet mass is used to improve the mass resolution of the jet, while N-subjettiness serves as a discriminant by yielding a probability of how compatible the jet is with having N axes. For this search, we require the softdrop-jet mass to be in a window around the W/Z/H/top mass, between 55 and 215 GeV, something which we plan to extend in the future. In order to improve the statistical power of the jet substructure variable τ_{21} and ensure a minimal sculpting of the jet mass as a function of jet p_T , we decorrelate the variable from the jet softdrop mass and the jet p_T -scale dependence following what was done in Ref. [?]. This decorrelation is performed by flattening the τ_{21} profile dependence on $\rho' = \log(m^2/p_T/\mu)$, where $\mu = 1$ GeV. Figure 5.71 shows the profile distribution of τ_{21} as a function of $\rho' = \log(m^2/p_T/\mu)$, applying the preselections as listed above together with a softdrop mass cut of $55 \text{ GeV} < M_{\text{jet}} < 215 \text{ GeV}$ (right). A linear

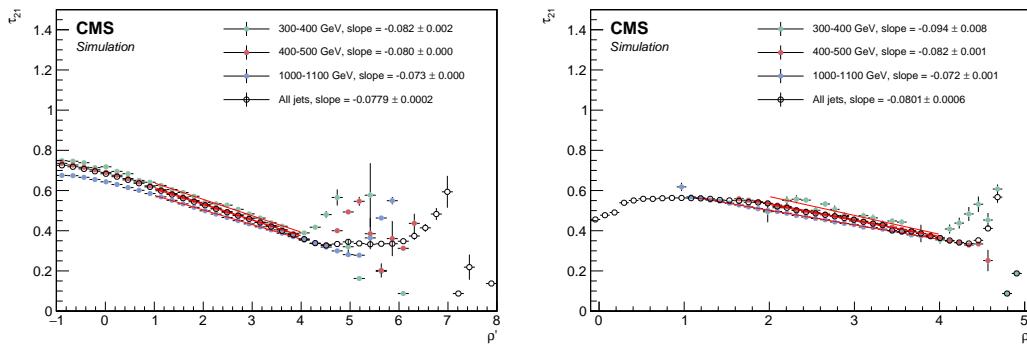


Figure 5.71: Profile distributions of τ_{21} as a function of $\rho' = \log(m^2/p_T/\mu)$, where $\mu = 1$ GeV (bottom), before applying a softdrop mass cut (left) and after applying a softdrop mass cut of $55 \text{ GeV} < M_{\text{jet}} < 215 \text{ GeV}$ (right).

transformation is then defined as

$$\tau_{21}^{DDT} = \tau_{21} - M \times \rho', \quad (5.11)$$

where the slope M is fitted from the linear part of the spectra of the τ_{21} profile versus ρ' with full selections (bottom left plot). For our purposes, the slope is extracted from fitting the inclusive p_T -spectrum ('All jets') with a mass window applied as it most closely corresponds to our full analysis selections. The resulting slope is $M = -0.080$, slightly steeper than the 2016 value of $M = -0.063$ [?]. The profile of the retuned τ_{21}^{DDT} versus ρ' is shown in Figure 5.72, exhibiting the desired flattened spectra. Working points for τ_{21}^{DDT} are chosen in the following way: First, we check which τ_{21}^{DDT} cut corresponds to the highest Punzi significance as a function of the resonance mass for different signal samples, shown in Figure 5.73. All other analysis cuts have been applied. The cut maximizing the Punzi significance at low resonance mass, where the background is highest, is chosen as the 'high purity' (HP) working point. This corresponds to $\tau_{21}^{DDT} \leq 0.43$. Second, we find the cut which, together with events falling in the HP region, contains at least 95 percent of the signal as well as optimizes the Punzi significance. This is found to be $0.43 < \tau_{21}^{DDT} \leq 0.79$, and is classified as the low purity (LP) category. The purpose of this category is to enhance the overall sensitivity, especially where the background is low. We observe a significant gain in signal efficiency at a fixed mistag rate with the retuned DDT tagger. The signal efficiency versus mistagging rate for all three taggers is shown in Figure 5.74, and we see the retuned

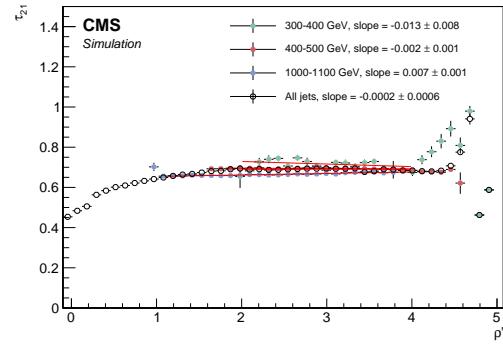


Figure 5.72: Profile distributions of τ_{21}^{DDT} as a function of $\rho' = \log(m^2/p_T/\mu)$, where $\mu = 1$ GeV.

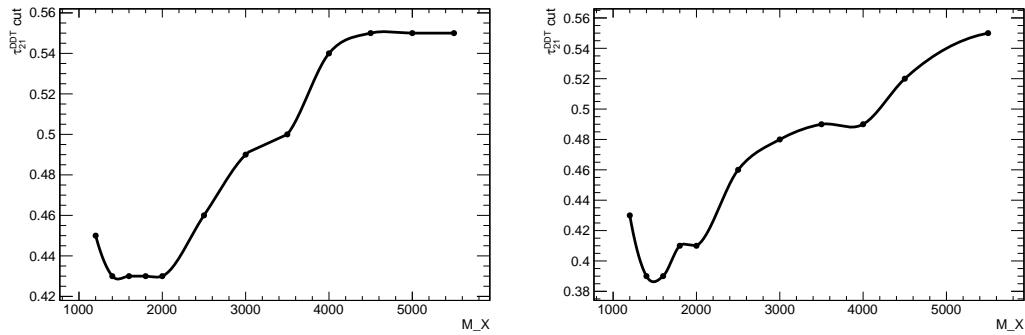


Figure 5.73: The τ_{21}^{DDT} cut corresponding to the highest Punzi significance for a given signal resonance mass, here for Bulk $G \rightarrow WW$ (left) and Bulk $G \rightarrow ZZ$.

τ_{21}^{DDT} performing better than τ_{21} and the version of τ_{21}^{DDT} using the old tune. In addition to

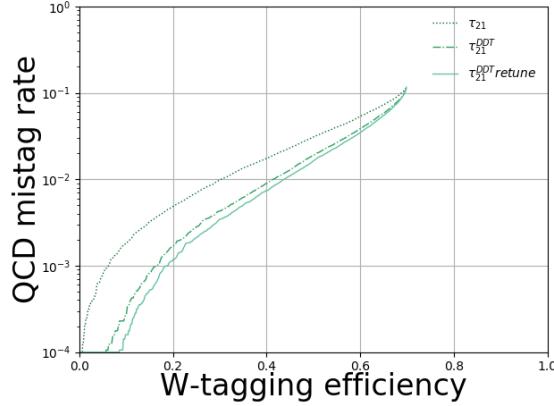


Figure 5.74: Performance of τ_{21} and τ_{21}^{DDT} (2016 and 2017 tune) in the background-signal efficiency plane.

cutting on τ_{21}^{DDT} a loose cut of $\rho = \log(m^2/p_T^2) < -1.8$ is applied. The reason for this is that, while the distribution of ρ is flat as a function of jet transverse momentum for QCD jets, this only holds in the region where perturbative contributions dominate and breaks down at around $\rho = \log(m^2/p_T^2) < -2.0$ due to the AK8 cone size being too small to contain the full jet at high masses. This has a negligible effect on the signal, which mainly peaks around 80 GeV and has a relatively high jet transverse momenta. Figure 5.75 shows the signal and background distribution for the PUPPI softdrop jet mass and τ_{21}^{DDT} . The signal softdrop mass distribution peaks nicely around the W mass, while the multijets background spectrum is peaked at lower softdrop masses. Also, in addition to having a higher signal efficiency for a given mistag rate, τ_{21}^{DDT} has the added benefit of being better modeled in MC than τ_{21} .

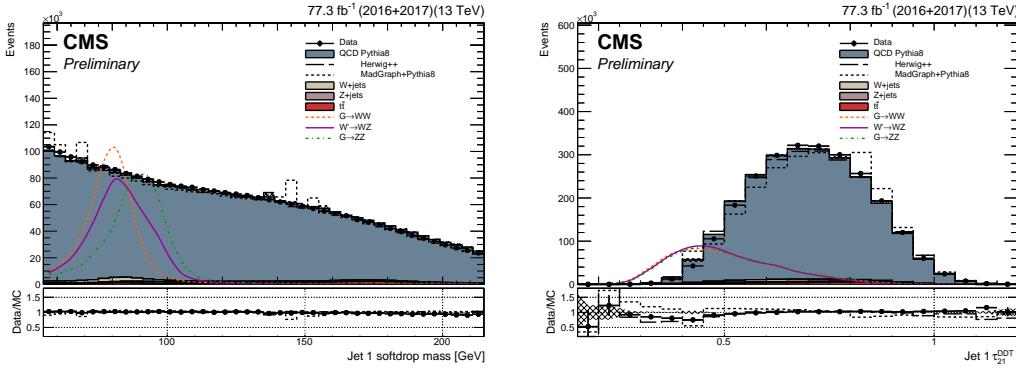


Figure 5.75: PUPPI softdrop jet mass distribution (left) and PUPPI N-subjettiness τ_{21}^{DDT} (right). Signal is scaled with an arbitrary number.

Data to simulation scale factors

Following what was done in Section 5.1.7 and 5.2.5, we derive W-tagging scale factors for the efficiency of the selection on τ_{21}^{DDT} by estimating the ratio of the selection efficiencies on data and simulation. The PUPPI softdrop mass range is extended to 55 to 215 GeV, and the two purity categories are

- Pass region: $0 < \tau_{21}^{DDT} \leq 0.43 \sim$ high purity
- Fail region: $0.43 < \tau_{21} \leq 0.79 \sim$ low purity

The obtained scale factors are listed in Tables 5.13 and 5.14 for 2016 and 2017 data, respectively, with the corresponding simultaneous fits shown in Fig. 5.76. The jet mass scale and resolution together with their are estimated in the same fits and also listed in Table 5.14. Two additional uncertainties are added: one due to generator differences and one due to NNLO corrections. These are evaluated by comparing the extracted efficiency with and without top p_T reweighting and when using $t\bar{t}$ simulation produced with different generators. The scale factors, jet mass scale and jet mass resolution with their total uncertainty after adding systematics, are listed in Table 5.15. As before, the scale factor is added as a scale of the signal yield and the jet mass scale and resolution are used to smear MC, and are additionally inserted as systematic uncertainties in the final fit (scale up/down).

	m [GeV]	σ [GeV]	W-tag efficiency
$\tau_{21}^{DDT} < 0.43$			
Data	81.999 ± 0.454 GeV	7.148 ± 0.544 GeV	0.080 ± 0.008
Simulation	80.890 ± 0.160 GeV	6.579 ± 0.149 GeV	0.085 ± 0.003
Data/simulation	1.014 ± 0.006	1.086 ± 0.086	0.937 ± 0.094
$0.43 < \tau_{21}^{DDT} < 0.79$			
Data			0.920 ± 0.008
Simulation			0.915 ± 0.003
Data/simulation			1.006 ± 0.009

Table 5.13: Jet mass scale, jet mass resolution and τ_{21}^{DDT} scale factors as evaluated in the full 2016 Single Muon dataset.

	m [GeV]	σ [GeV]	W-tag efficiency
$\tau_{21}^{DDT} < 0.43$			
Data	80.784 ± 0.391 GeV	7.694 ± 0.445 GeV	0.065 ± 0.006
Simulation	82.208 ± 0.293 GeV	7.127 ± 0.284 GeV	0.068 ± 0.005
Data/simulation	0.983 ± 0.006	1.080 ± 0.076	0.955 ± 0.113
$0.43 < \tau_{21}^{DDT} < 0.79$			
Data			0.935 ± 0.006
Simulation			0.932 ± 0.005
Data/simulation			1.003 ± 0.008

Table 5.14: Jet mass scale, jet mass resolution and τ_{21}^{DDT} scalefactors as evaluated in the full 2017 Single Muon dataset.

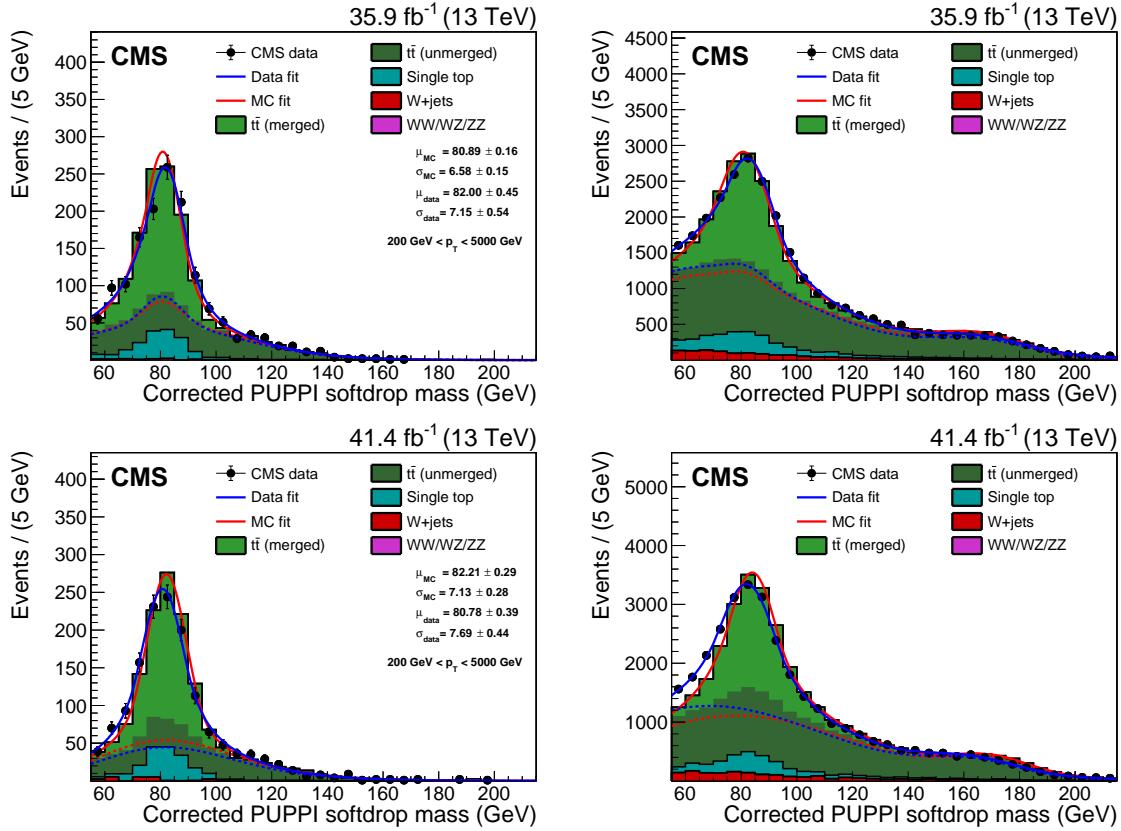


Figure 5.76: PUPPI softdrop jet mass distribution that pass (left) and fail (right) the $\tau_{21}^{DDT} 0.43$ selection in the $t\bar{t}$ control sample. The result of the fit to data and simulation are shown by the solid blue and solid red line, respectively. The background components of the fit are shown as dashed-dotted lines. The fit to 2016 data is shown in the upper panels and the fit to 2017 data in the lower panels.

	SF $\pm \sqrt{\text{Stat.} + \text{Sys}_{\text{Generator}} + \text{Sys}_{\text{NNLO}}}$	SF \pm Total Unc.
$HPSF_{DDT}^{2017}$	$0.955 \pm \sqrt{0.113^2 (\text{stat.}) + 0.003^2 (\text{sys.}) + 0.043^2 (\text{sys.})}$	0.955 ± 0.121
$HPSF_{DDT}^{2016}$	$0.937 \pm \sqrt{0.094^2 (\text{stat.}) + 0.003^2 (\text{sys.}) + 0.043^2 (\text{sys.})}$	0.937 ± 0.103
$LPSF_{DDT}^{2017}$	$1.003 \pm \sqrt{0.008^2 (\text{stat.}) + 0.003^2 (\text{sys.}) + 0.0^2 (\text{sys.})}$	1.003 ± 0.008
$LPSF_{DDT}^{2016}$	$1.006 \pm \sqrt{0.009^2 (\text{stat.}) + 0.003^2 (\text{sys.}) + 0.0^2 (\text{sys.})}$	1.006 ± 0.009
JMS^{2017}	$0.983 \pm \sqrt{0.006^2 (\text{stat.}) + 0.002^2 (\text{sys.}) + 0.001^2 (\text{sys.})}$	0.983 ± 0.007
JMS^{2016}	$1.014 \pm \sqrt{0.006^2 (\text{stat.}) + 0.002^2 (\text{sys.}) + 0.001^2 (\text{sys.})}$	1.014 ± 0.007
JMR^{2017}	$1.080 \pm \sqrt{0.076^2 (\text{stat.}) + 0.027^2 (\text{sys.}) + 0.001^2 (\text{sys.})}$	1.080 ± 0.081
JMR^{2016}	$1.086 \pm \sqrt{0.086^2 (\text{stat.}) + 0.027^2 (\text{sys.}) + 0.001^2 (\text{sys.})}$	1.086 ± 0.090

Table 5.15: Final jet mass scale, jet mass resolution and τ_{21}^{DDT} scalefactors.

5.3.7 The multidimensional fit

As mentioned in the introduction to this chapter, the three-dimensional fit method takes advantage of the fact that the signal peaks in three dimensions; dijet invariant mass (M_{VV}) and the jet groomed mass of jet 1 and jet 2 (M_{jet1} and M_{jet2}) and attempt to extract the signal from the three dimensional $M_{jet1}(x)$ - $M_{jet2}(y)$ - $M_{VV}(z)$ plane. In order to do so, four different types of PDFs need to be created in order to have a complete model:

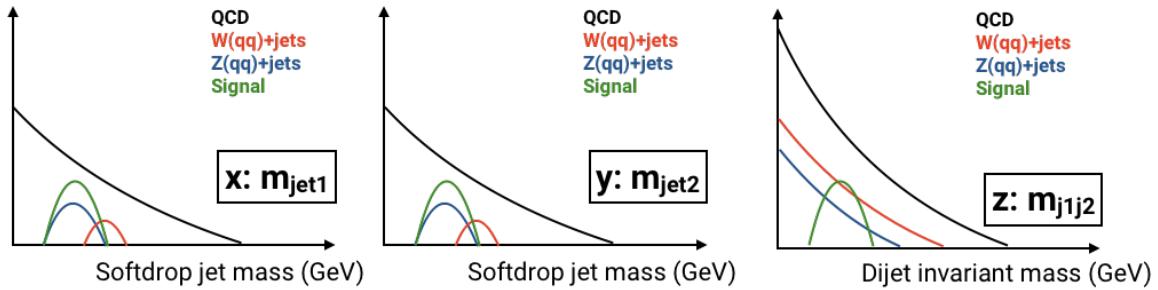


Figure 5.77: An illustration of the shape of the signal and the relative background contributions in the three relevant dimensions $M_{jet1}(x)$, $M_{jet2}(y)$ and $M_{VV}(z)$.

- **Signal 3D PDF:** Resonant in x, y and z. Parametrized as function of the resonance mass M_X
- **Non-resonant background:** Non-resonant in x, y and z and dominant background. Created through a forward folding kernel approach
- **Resonant background:** Mainly W/Z+jets (some $t\bar{t}$). Resonant in x and y, smoothly falling in z.
- **Alternative shapes:** 5 additional shape uncertainties implemented through vertical morphing

These are illustrated in Figure 5.77 and will be described in detail in the following.

5.3.8 Modeling of the signal

The signal shape in three dimensions is defined as a product of the shape of the resonance mass and the jet masses:

$$P_{sig}(M_{VV}, M_{jet1}, M_{jet2} | \theta(M_X)) = P_{VV}(M_{VV} | \theta_1(M_X)) \times P_{j1}(M_{jet1} | \theta_2(M_X)) \times P_{j2}(M_{jet2} | \theta_2(M_X)). \quad (5.12)$$

The shapes for M_{VV} , M_{jet1} and M_{jet2} all depend on the hypothesized mass of the new particle (M_X) and a set of parameters $\theta = (\theta_1, \theta_2)$ that in principle depend on M_X . The signal is parametrized by fitting the resonance mass and jet mass line shapes for each mass point, extracting the fitted parameters and then interpolating these as a function of the resonance mass hypothesis. For the resonance mass M_{VV} , the sum of a crystal-ball function and a Gaussian shape is used for each mass point, following the shapes used in Search II. Figure 5.78 shows the derived parameters and interpolation as a function of resonance mass. The final M_{VV} shapes as extracted from the parametrization are shown in Figure 5.79.

The same procedure is used to model the jet mass: The M_{jet1} spectrum for each resonance mass hypothesis is fitted using a double Crystal-ball function, the fitted parameters are

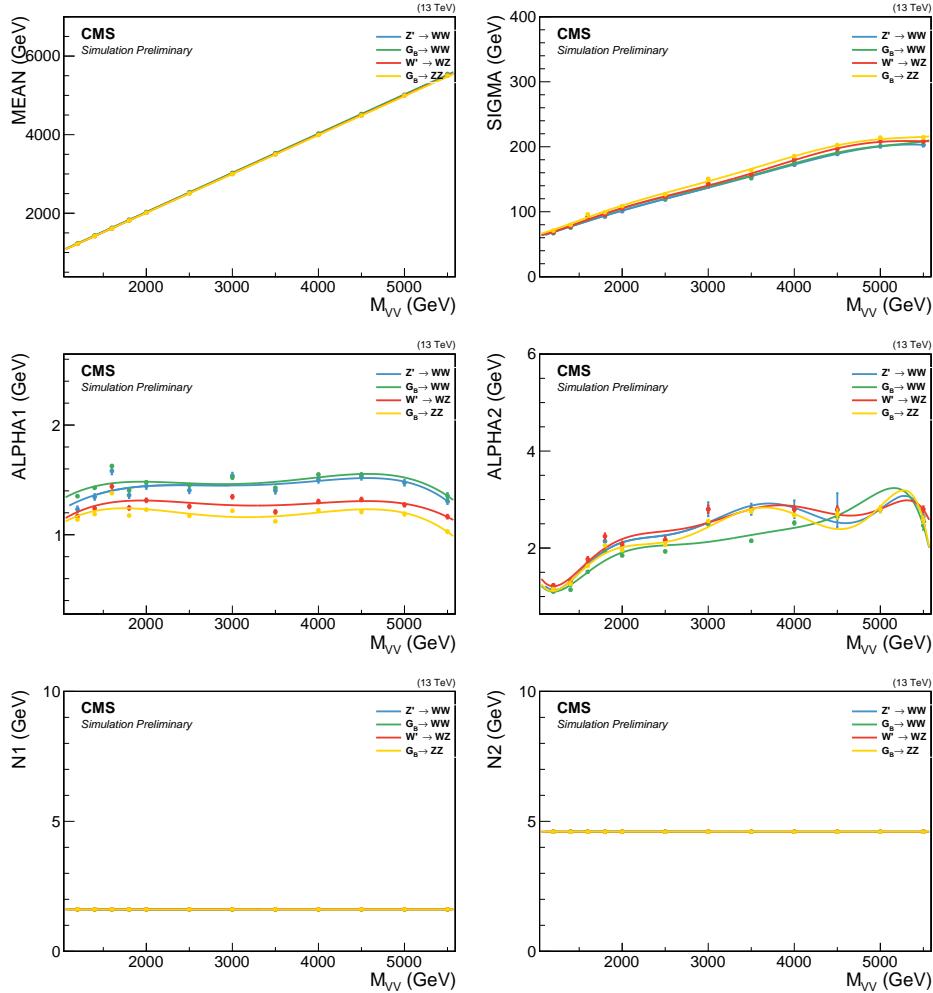


Figure 5.78: The interpolated Crystal-ball parameters for the dijet invariant mass as a function of M_X . The small variations for ALPHA2 have been shown to have no effect on the overall modeling.

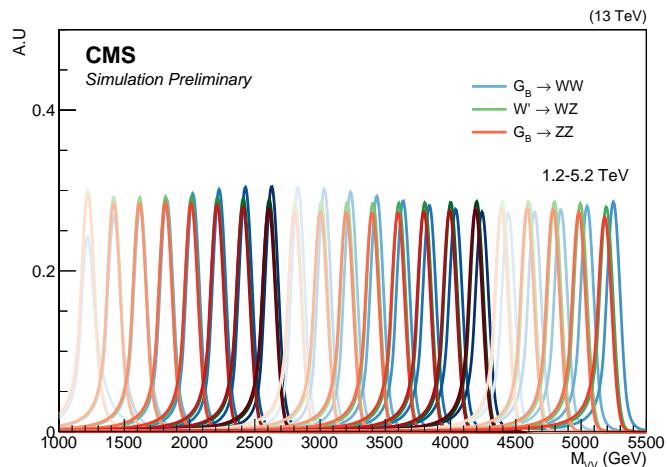


Figure 5.79: Final M_{VV} signal shapes extracted from the parametrization. Here for a G_{bulk} decaying to WW (blue) and ZZ (red) and for a W' decaying to WZ (green).

extracted and interpolated as functions of the resonance mass. This is done separately for $M_{\text{jet}1}$ and $M_{\text{jet}2}$. The fitted parameters and interpolations are shown in Figure 5.80 for $M_{\text{jet}1}$, the corresponding distributions for $M_{\text{jet}2}$ are in Appendix E.1. The final $M_{\text{jet}1}$ shapes as

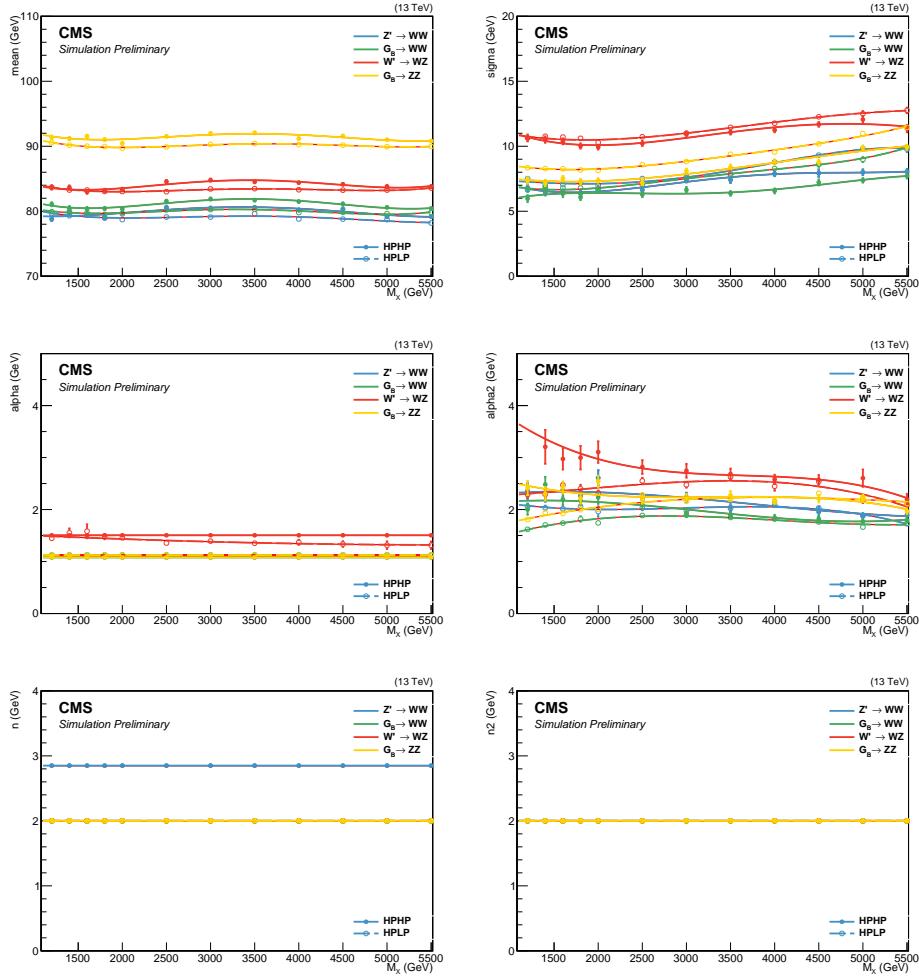


Figure 5.80: The interpolated double Crystal-ball parameters for the softdrop jet mass as a function of M_X . To improve the stability of the fit some parameters are set constant. Here for jet 1.

extracted from the parametrization are shown in Figures 5.81. Finally, the signal yield is parametrized as a function of the resonance mass. For each mass point M_X and each purity category, the signal yield per picobarn of cross section is calculated as the integral of the Monte Carlo histogram. The yields are then interpolated as a function of M_X . The signal efficiency as a function of resonance mass is shown in Figure 5.82.

5.3.9 Modeling of the non-resonant background

In order to model the QCD multijets background in the three-dimensional $M_{\text{VV}}\text{-}M_{\text{jet}1}\text{-}M_{\text{jet}2}$ plane, we use the following conditional product:

$$P(M_{\text{VV}}, M_{\text{jet}1}, M_{\text{jet}2}) = P_{\text{VV}}(M_{\text{VV}}|\theta_1) \times P_{\text{cond},1}(M_{\text{jet}1}|M_{\text{VV}}, \theta_2) \times P_{\text{cond},2}(M_{\text{jet}2}|M_{\text{VV}}, \theta_2). \quad (5.13)$$

This probability density requires a computation of the conditional two-dimensional shapes of $M_{\text{jet}1}/M_{\text{jet}2}$ given M_{VV} , as well as a one dimensional shape of the M_{VV} distribution.

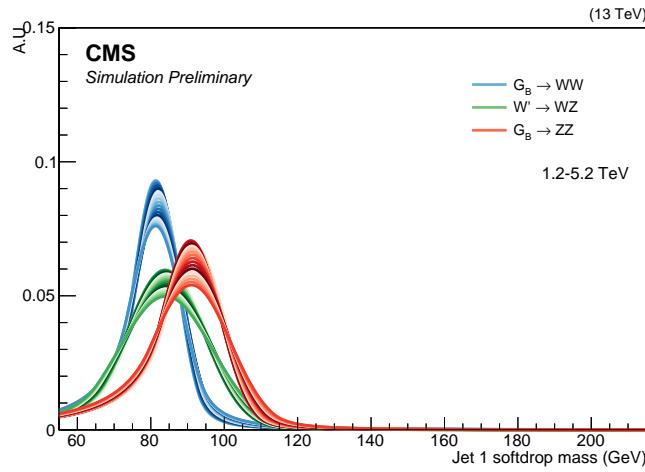


Figure 5.81: Final M_{jet} signal shapes extracted from the parametrization for a G_{bulk} decaying to ZZ, a G_{bulk} decaying to WW and for a W' decaying to WZ.

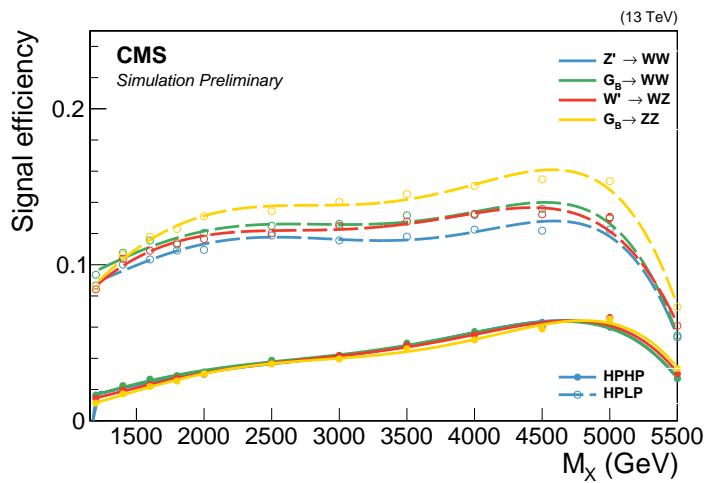


Figure 5.82: Signal efficiency as a function of resonance mass.

The following fit range and binning is used for the three axes: $M_{\text{jet}1}/M_{\text{jet}2}$ is fitted from 55 to 215 GeV using 2 GeV bins. M_{VV} is fitted from 1126 to 5500 GeV. The lower bound is chosen such that to avoid complications in the fitting procedure due to trigger turn-on effects, while the upper bound is chosen considering the highest dijet invariant mass found in data as well as avoiding mis-reconstruction effects at very large m_{jj} and low jet masses. For M_{VV} , the ‘dijet binning’ is used. This binning corresponds to the actual dijet mass resolution and is, in units of GeV:

Dijet binning = 1126, 1181, 1246, 1313, 1383, 1455, 1530, 1607, 1687, 1770, 1856, 1945, 2037, 2132, 2231, 2332, 2438, 2546, 2659, 2775, 2895, 3019, 3147, 3279, 3416, 3558, 3704, 3854, 4010, 4171, 4337, 4509, 4686, 4869, 5058, 5253, 5500

The background model is built starting from simulation and we encode sufficient nuisance parameters into the fit, allowing the shape to adapt itself to data. For this we use a ‘forward-folding’ approach. For each MC event in the two(one)-dimensional $M_{\text{VV}}-M_{\text{jet}}$ (M_{jet}) space, a 2D(1D) Gaussian kernel is built starting from generator level quantities. Each of these Gaussians then contribute to the total probability density of the final two(one)-dimensional probability density functions

First, the resonance mass and softdrop jet mass scale and resolution are derived. For this we use the anti- k_T generated jet collection matched to jets identified as V-jets on reconstruction level. We then derived the M_{jet} and M_{VV} scale and resolution from a Gaussian fit to $M_i(\text{reco})/M_i(\text{gen})$ ($i = M_{\text{jet}}$ or $i = M_{\text{VV}}$) in bins of generator jet p_T . Figure 5.83 shows the fit to $M_{\text{jet}}(\text{reco})/M_{\text{jet}}(\text{gen})$ (left) and $M_{jj}(\text{reco})/M_{jj}(\text{gen})$ (right) for an arbitrary bin. The Gaussian mean yields the mass scale and the Gaussian width the mass resolution. The

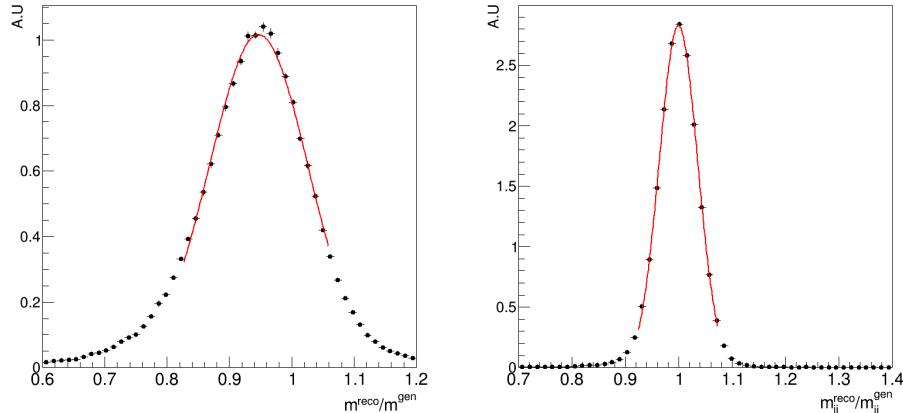


Figure 5.83: Fit to $M_{\text{jet}}(\text{reco})/M_{\text{jet}}(\text{gen})$ (left) and $M_{jj}(\text{reco})/M_{jj}(\text{gen})$. The mass resolution is taken as the width of the fitted Gaussian, while the Gaussian mean yields the mass scale.

resonance mass and softdrop jet mass scale and resolution as a function of generator jet p_T is shown in Figure 5.84. The projection of these resolution functions are shown in Figure 5.85. The mass scale and resolution are then used to populate the conditional 2D histogram as follows. Each generated event is smeared with a 2D Gaussian kernel

$$k(M_{\text{jet}}, M_{\text{VV}}) = \frac{w_i}{\sqrt{2\pi} r_{M_{\text{VV}},i} \cdot r_{M_{\text{jet}},i}} \exp \left(-\frac{1}{2} \left(\frac{M_{\text{VV}} - s_{M_{\text{VV}},i}}{r_{M_{\text{VV}},i}} \right)^2 - \frac{1}{2} \left(\frac{M_{\text{jet}} - s_{M_{\text{jet}},i}}{r_{M_{\text{jet}},i}} \right)^2 \right), \quad (5.14)$$

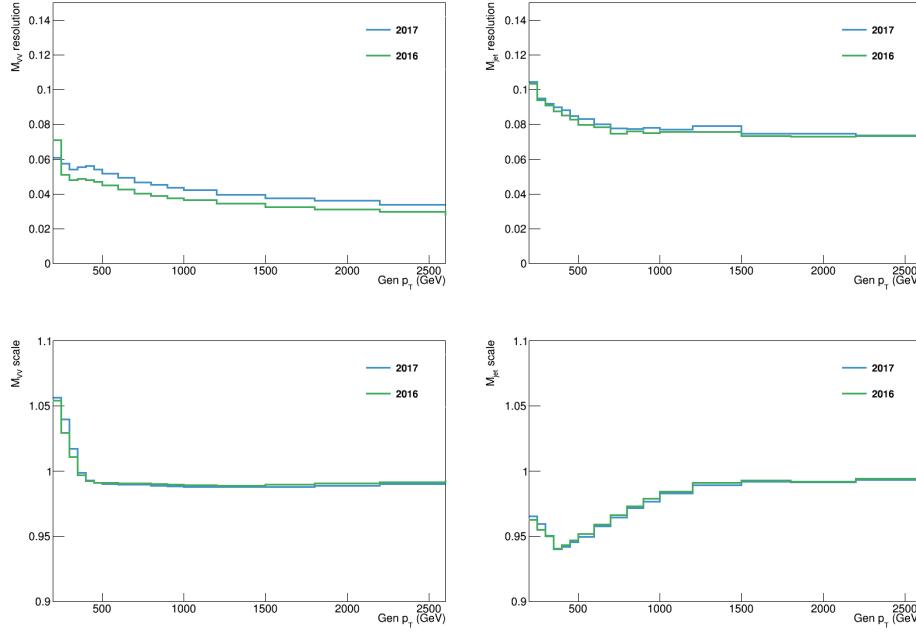


Figure 5.84: Resolution (top) and scale (bottom) for M_{VV} (left) and the M_{jet} (right) as a function of generator jet p_T .

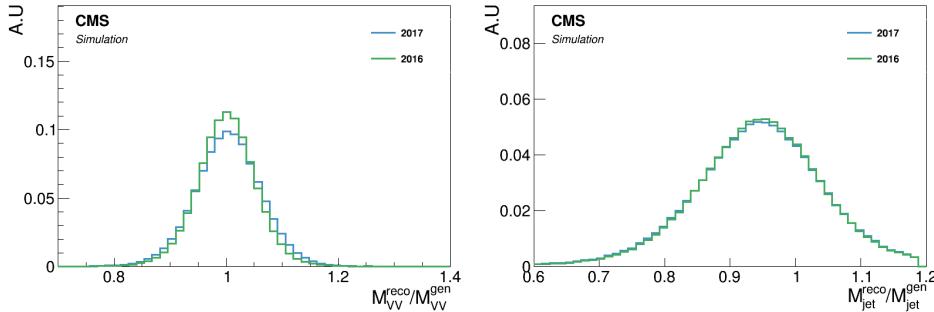


Figure 5.85: Projections of the resolution functions for all generator jet p_T bins for M_{VV} (left) and M_{jet} (right).

where s_i, r_i are the scale and the resolution derived in the previous step and w_i is the event weight product (e.g PU, cross sections etc.). The resulting kernel values are filled into a 2D histogram. This procedure is performed separately for M_{jet1} and M_{jet2} . To build the one-dimensional template for the dijet invariant mass the same procedure as above is used, with the exception that the smearing is done with one-dimensional Gaussian kernel only depending on M_{VV} . The templates are then added together to form a three-dimensional PDF. Finally, we fit this 3D PDF to QCD MC in order to remove any residual bias (mainly at the extreme ends of the spectra). The result is a full, smooth shape replacing the prediction from simulation.

As the high purity category is limited by statistics, we rather build this template starting from the low purity templates. This is done by fitting the low purity 3D template to QCD MC in the high purity category. Figure 5.86 show the final templates (solid lines) together with the QCD MC (data points) derived from the 2017 MC in the low purity (top) and high purity category (bottom). Good agreement between simulation and templates in all three

dimensions is observed, within statistical uncertainties. The corresponding distributions in 2016 MC can be found in Appendix E.2.

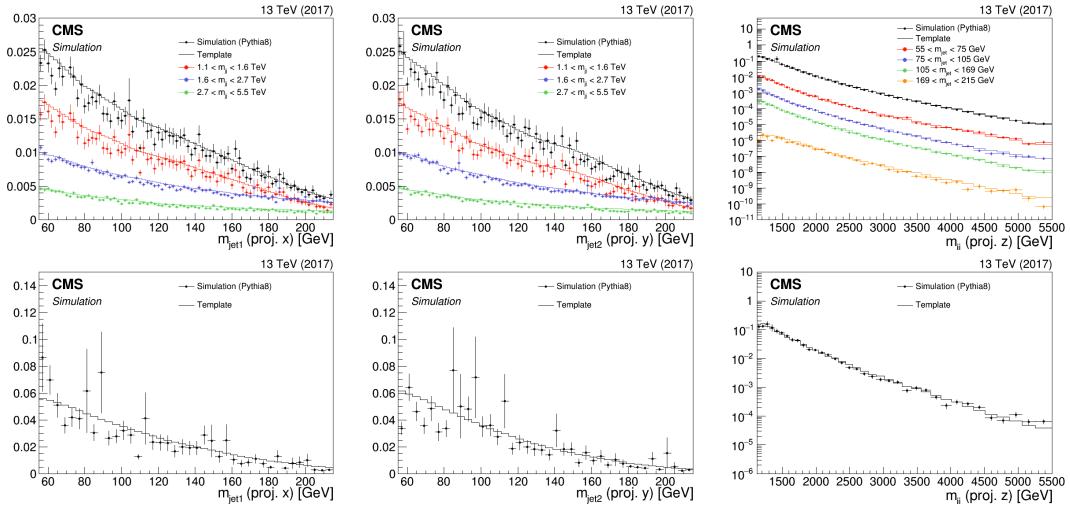


Figure 5.86: Comparison between QCD MC simulation (markers) and kernels derived from generator level quantities (lines) for the HPHP (top) and HPLP (bottom) categories using 2017 MC. The kernels are shown for $M_{\text{jet}1}$ (left), $M_{\text{jet}2}$ (middle) and M_{VVV} (right).

In order to validate the kernel transfer method, we check that we can fit a higher-statistics high purity region by loosening the τ_{21}^{DDT} cut to 0.49. This results in 12 times more background events, and should uncover whether any degeneracy is present in the fits themselves and whether the low purity kernel indeed is capable of modeling the high purity region, without being camouflaged by large error bars. The resulting kernel versus MC spectra are shown in Figure 5.87. Good closure is observed in all three dimensions, demonstrating that the HPLP kernels adapt well to the HPHP MC data points even when statistics are sufficient, and we consider the method sound.

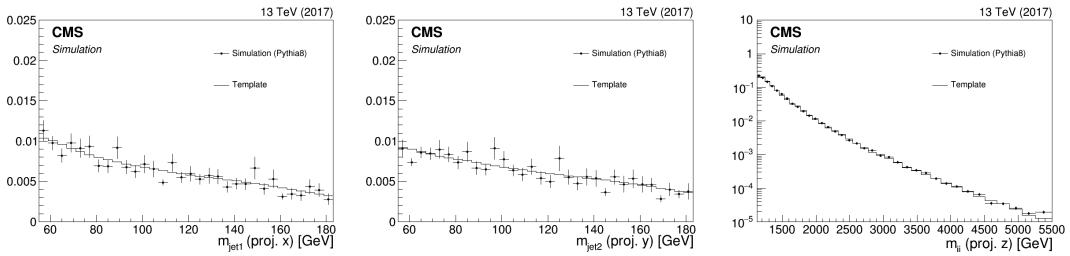


Figure 5.87: Comparison between QCD MC simulation (markers) and kernels derived from generator level quantities (lines) in the HPHP category, using a looser cut on τ_{21}^{DDT} .

5.3.10 Modeling of the resonant background

In addition to the QCD multijet background, there are a few sub-dominant processes to consider which contain one real vector boson and at least one QCD-jet. These are W+jets, Z+jets and events from $t\bar{t}$ processes. They are resonant around the W/Z mass in the two softdrop jet mass dimensions and must therefore be treated differently than the non-resonant QCD background. Figure 5.88 shows the projections on M_{jet1} (left), M_{jet2} (middle) and M_{VV} (right). As the jets are randomly sorted, each jet mass dimension contains two contributions:

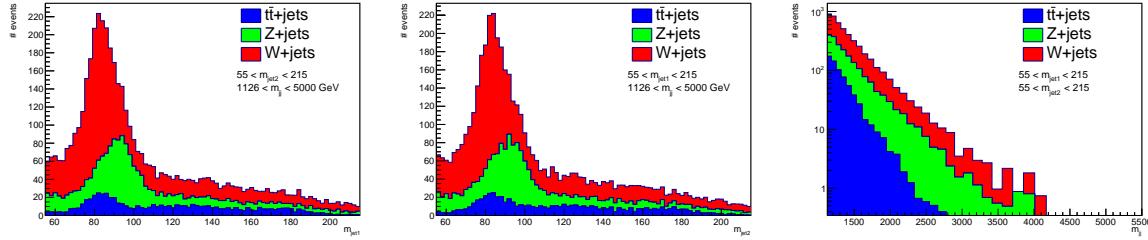


Figure 5.88: Projections of the sub-dominant backgrounds on the jet mass axes M_{jet1} (left) and M_{jet2} (middle), as well as on the dijet invariant mass M_{VV} (right). Here in the low purity category.

A resonant part consisting of real vector boson jets, peaking around the V mass, and a non-resonant part composed of jets originating from a quark or a gluon, similar to the QCD multijets background. These two contributions are modeled separately as we know that the non-resonant part of the jet mass spectrum is correlated with the dijet invariant mass (like QCD), while the resonant part is not. We additionally want to encode the fact that we know these backgrounds in reality only peaks in one jet mass axis (there is only one real vector boson) by requiring the PDFs to consist of a resonant part on one axis, and a non-resonant part on the other axis. A three dimensional PDF for the sub-dominant backgrounds is built as a product of three one dimensional pdf's as follows:

$$P_{res}(M_{jet1}, M_{jet2}, M_{VV}) = P_{VV}(M_{VV}) \times P_{jet1}(M_{jet1}, M_{jet2}) \times P_{jet2}(M_{jet2}, M_{jet1}) \quad (5.15)$$

where

$$P_{jet1}(M_{jet1}, M_{jet2}) = f \times R(M_{jet2}) \times P_{res}(M_{jet1}) + (1 - f)P_{non-res}(M_{jet1}) \quad (5.16)$$

$$P_{jet2}(M_{jet2}, M_{jet1}) = (1 - f) \times R(M_{jet1}) \times P_{res}(M_{jet2}) + fP_{non-res}(M_{jet2}) \quad (5.17)$$

Here, $R(M_{jet})$ parametrizes the correlation between M_{jet1} and M_{jet2} and f is a fit parameter that is used to adjust the fraction of real vector boson jets in M_{jet1} compared to M_{jet2} . Its value can vary 10% around $f = 0.5$, the expected median when using a random jet sorting. As the contribution of the $t\bar{t}$ background is much smaller than the one coming from V+jets (less than 2%), it is modeled together with the W+jets contribution. Therefore only two PDFs are built for the resonant background: One for the W+jets and $t\bar{t}$, and one for Z+jets. The available MC statistics in the high purity category is very low and therefore the parametrization of the resonant background is done for the low purity category only, and the resulting shapes are then used for both purity categories. The uncertainties for the different purity categories are, however, included as separate nuisance parameters in the fit.

The non-resonant M_{VV} PDF is constructed using the same kernel approach as is used to model the QCD multijet background with one minor difference: Due to the low statistics in the high- M_{VV} tail, an additional smoothing of the jet mass tail using a leveled exponential

$$\frac{dN}{dm_{jj}} = \frac{P_0(1 - m_{jj}/s)^{P_1}}{(m_{jj}/s)^{P_2}} \quad (5.18)$$

is performed. Here, s is the center of mass energy. The function is fitted to the spectrum starting from a dijet mass of 1.1 (2.1) TeV for high purity (low purity). Two uncertainties on this shape is added in order to accommodate for MC mis-modeling due to higher order QCD and electroweak corrections: One proportional to M_{VV} and one proportional to $1/M_{VV}$. The resulting M_{VV} kernels (solid lines) for the W+jets background are shown in Figure 5.89 and are compared to MC simulation (markers). The blue line corresponds to the nominal shape, while the red and green lines correspond to the uncertainties proportional to M_{VV} and $1/M_{VV}$, respectively. The corresponding distributions for the Z+jets background are shown in Appendix E.3.

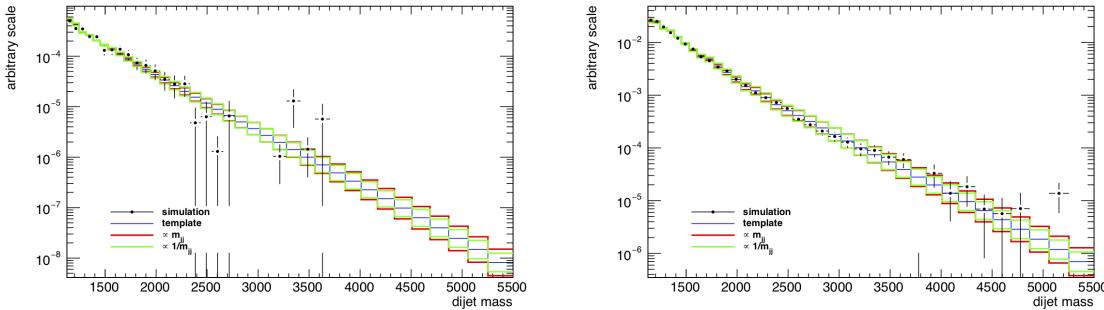


Figure 5.89: One dimensional M_{VV} kernels (solid line) compared to MC (markers) for the W+jets background in the HPHP (left) and HPLP (right) categories. The blue line corresponds to the nominal shape, while the red and green lines correspond to uncertainties proportional to M_{VV} and $1/M_{VV}$, respectively.

As mentioned above, the modeling of the M_{jet} spectrum is split into two different PDFs: One describing the resonant and one describing the non-resonant part. We distinguish between the two by requiring the resonant contribution to consist of jets matched to a generated boson within a cone of $\Delta R = 0.8$. A double-sided crystal-ball function is then fitted to the resonant spectrum for W+jets and Z+jets separately. This allows us to fully correlate uncertainties on the mean and width of the M_{jet} distribution for signal with the corresponding values for W/Z+jets, as these uncertainties affect all jets generated from real vector bosons in the same way. This effectively gives us a way of constraining these parameters directly from data. The parametrization of the resonant part of the jet mass spectrum for M_{jet1} is shown in Figure 5.90 for W+jets and $t\bar{t}$ (left) and Z+jets. The small enhancement around 170 GeV is caused by fully merged top jets and is so small ($< 2\%$ in the low purity and $< 0.5\%$ in the high purity category) that we do not take it into account in the final PDF. The non-resonant part of the V+jets background is modeled using a simple Gaussian fit to the spectrum of jets not matched to a real vector bosons, as shown in Figure 5.91. For the resonant modeling, correlations between the jet mass M_{jet} and dijet invariant mass M_{VV} have been found small enough to be neglected (the jet mass spectrum does not, within statistical uncertainties, depend on the jet p_T). However, there is a strong correlation between the softdrop jet mass of each jet due to the fact that when one of the two jets originates from a real boson peaking around the V mass, the other is bound to be a quark jet. Therefore, the fraction of real boson jets contained in M_{jet1} affects the fraction of real V jets in M_{jet2} . To account for this, the fraction of real V jets versus quark/gluon jets jets, is parametrized as a function of the jet mass. As the jets are randomly sorted, we define the parametrization as

$$R(M_{jet}) = \frac{N_{res,jet1}(M_{jet2}) + N_{res,jet2}(M_{jet1})}{N_{non-res,jet1}(M_{jet2}) + N_{non-res,jet2}(M_{jet1})} . \quad (5.19)$$

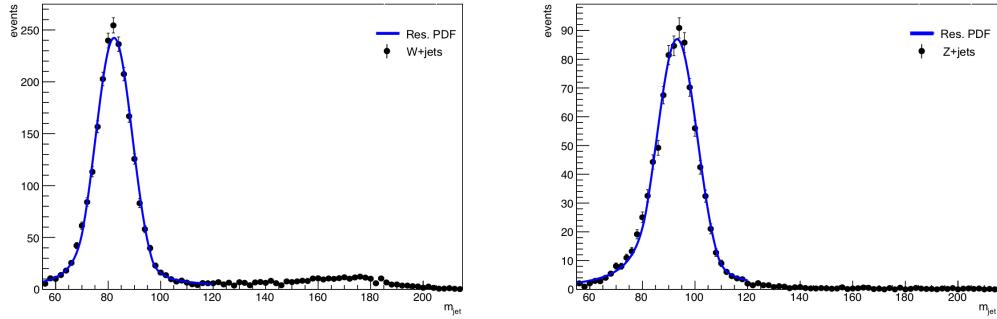


Figure 5.90: Fit to the resonant part of the $V + \text{jets}$ and $t\bar{t}$ spectrum for $W + \text{jets}$ and $t\bar{t}$ (left) and $Z + \text{jets}$ (right).

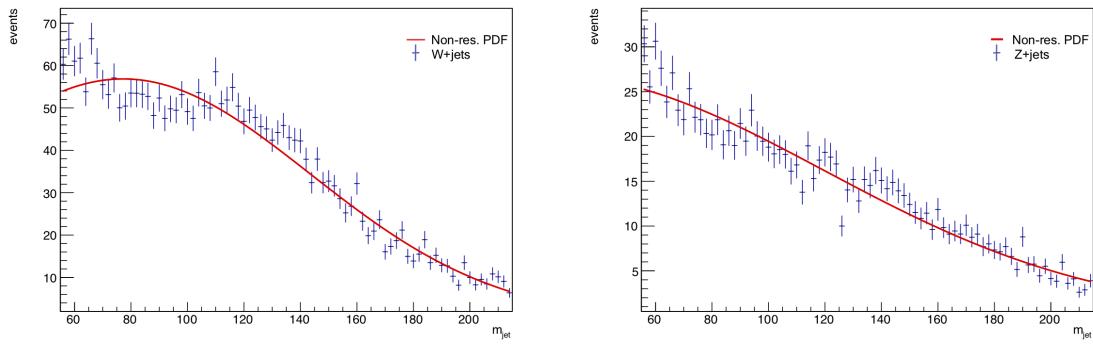


Figure 5.91: Fit to the non-resonant part of the $V + \text{jets}$ and $t\bar{t}$ spectrum for $W + \text{jets}$ and $t\bar{t}$ (left) and $Z + \text{jets}$ (right).

Where $N(M_{\text{jet}})$ denotes the number of events within a given M_{jet} window. The resulting ratio is then fitted with a polynomial function, as shown in Figure 5.92. As a closure test, the three dimensional kernel as defined in Equation 5.15 is fitted to the $V + \text{jets}$ and $t\bar{t}$ simulation. Figure 5.93 shows the fitted kernel (red) together with the MC data points (black markers) in the low (top) and high (bottom) purity categories. Mostly good agreement is observed along all dimensions, with some deviations in the very high M_{VV} tail in the high purity category. This is, however, a region dominated with very little statistics and is completely swallowed up by the QCD multijets background which has the same shape.

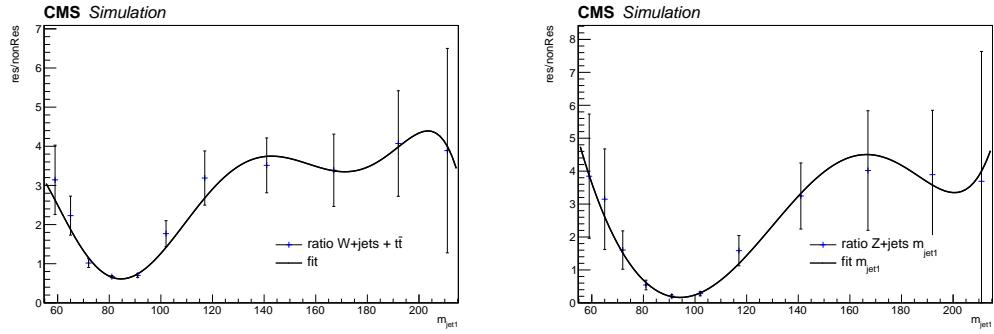


Figure 5.92: Ratio of resonant to non-resonant events in $W+jets$ (left) and $Z+jets$ (right) as a function of jet mass.

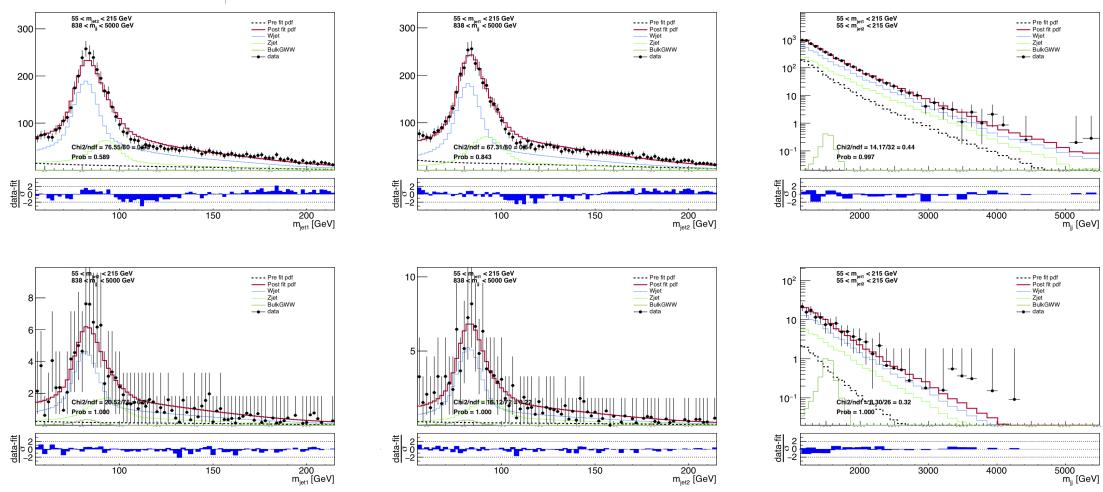


Figure 5.93: Final fits of the complete background model (red line) to the MC simulation of the $V+jets$ backgrounds (black markers) for the high purity category.

5.3.11 Systematic uncertainties

Systematic uncertainties are inserted as nuisance parameters in the fit and affect the normalization and shape of both signal and background.

Signal normalization uncertainties

As all background contributions are data driven, the largest systematic uncertainties affect only the signal.

- **luminosity** 2.6%(2.3%) for the 2016(2017) dataset
- τ_{21}^{DDT} **efficiency** 1-12% systematic uncertainty on the signal efficiency scale factor for the τ_{21}^{DDT} selection (listed in Table 5.15). Anti-correlated between the HP and LP categories.
- τ_{21}^{DDT} **p_T extrapolation** An additional uncertainty arising from the extrapolation of the V-tagging efficiency scale factors towards higher transverse momenta. Treated as correlated between the τ_{21}^{DDT} categories and is given as $3.9\% \times \ln(p_T/200 \text{ GeV})$ and $8.5\% \times \ln(p_T/200 \text{ GeV})$ for the LP and HP regions, respectively.
- **PDF and factorization/renormalization scale** 2% uncertainty on the signal acceptance due to choice of PDFs and factorization/renormalization scale.

Background normalization uncertainties

The QCD multijet background has a poorly known cross section and is therefore allowed to float within 50% of the yield expected by simulation. For the resonant V+jets background, the following uncertainties on the normalization is added: One due to cross section ($\sim 5\%$), one due to NLO QCD corrections ($\sim 10\%$) and one due NLO electroweak corrections ($\sim 15 - 35\%$).

Signal and resonant background shape uncertainties

There are two shape uncertainties correlated between the signal and V+jets background: Systematics due to jet mass scale (JMS) and jet mass resolution (JMR). These are evaluated in a $t\bar{t}$ control region, listed in Table 5.15, and affect the mean and width of the jet mass PDFs. Three additional shape uncertainties are added for the signal: Uncertainties due to jet energy scale (JES), jet energy resolution (JER) and PDF. The uncertainty due to JES/JER is evaluated by reweighting the transverse momentum of each MC event in signal MC, then fitting a Gaussian to the dijet invariant mass spectrum and calculating the change in the mean and variance with respect to no reweighting. These are then used as uncertainties on the signal M_{VV} shape (2% for the dijet mass scale, 5% for the resolution).

Uncertainties due to the PDF are evaluated by reweighting each event according to ≈ 100 PDF variations. The impact on the signal shapes are again evaluated by fitting a Gaussian to the m_{jet} and m_{jj} distributions before and after changing PDF weights. The obtained variation is $< 1\%$ for the 2017 signal MC, but slightly higher in 2016. An overall uncertainty on the acceptance of 3% is therefore applied.

Non-resonant background shape uncertainties

Alternate shapes for the non-resonant background components are added to the fit through vertical template morphing. This creates nuisance parameters for each additional shape

shape, allowing the derived nominal 3D PDF to vary within the respective nuisances to match the data. Shape uncertainties that simultaneously affect all three dimensions are used, totaling to 5 to take all possible effects into account. The first effect accounts for variation of the underlying transverse momentum spectrum and is obtained through an alternate shape produced by varying the jet masses M_{jet} and dijet mass M_{VV} by a quantity proportional to M_{VV} and M_{jet} . The second effect considered is a variation of the scale, where the corresponding alternate shape is obtained by simultaneously varying M_{jet} and M_{VV} up and down by a quantity proportional to $1/M_{\text{VV}}$ and $1/M_{\text{jet}}$. Finally, as we do not know a priori whether Nature behaves more like PYTHIA8, HERWIG++ or MADGRAPH+PYTHIA8, we account for differences due to MC generation and parton shower modeling by adding three additional shapes corresponding to the PDFs obtained using different QCD MC generators. Those five shape uncertainties are assigned very large pre-fit values (allowed to float within 33%), effectively allowing the simulation to take any value to describe the data. The alternate shapes described above are shown in Figure 5.94, here for the 2017 analysis.

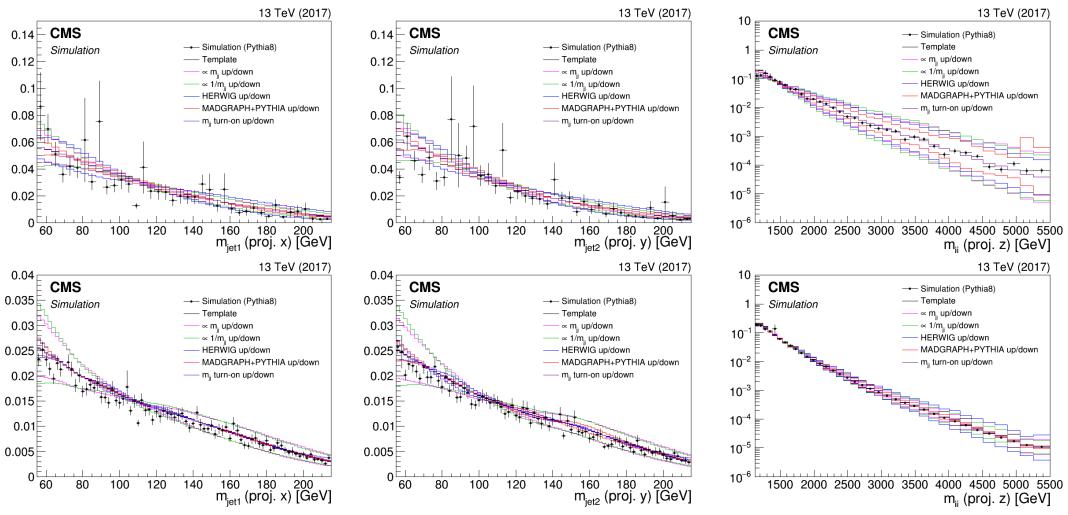


Figure 5.94: The nominal MC data (markers) and smooth nominal kernel obtained from PYTHIA8 (black line), together with the five alternate shapes added to the fit as shape nuisance parameters. Here for the high (top) and low (bottom) categories.

5.3.12 Results

The distributions obtained from a combined fit to the observed data in 2016 and 2017 are shown in Fig. 5.95, with the corresponding predicted and observed number of background events in the signal region summarized in Table 5.16.

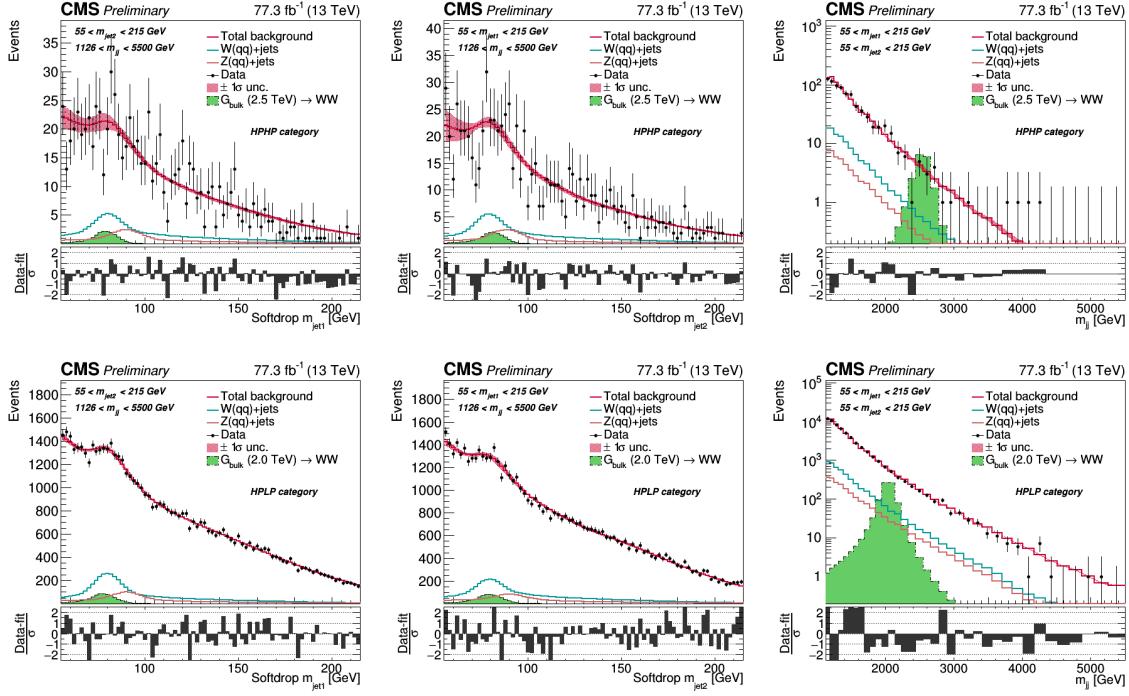


Figure 5.95: Postfit distribution after a fit to 2016 and 2017 data projected onto the $M_{\text{jet}1}$ (left), $M_{\text{jet}2}$ (middle), and M_{VV} (right) axis. Here for the HPHP (top) and HPLP (bottom) category. The background shape uncertainty is shown as a red shaded band, and the statistical uncertainties of the data are shown as vertical bars. The overlaid signal distribution is arbitrarily normalized.

Limits

As for Search I and Search II, exclusion limits on the cross section of the process $X \rightarrow VV$ are set in the context of the Bulk Graviton model and the HVT model B scenario (again

	HPHP	HPLP
W+jets	113.3 ± 18.1 100.4 (exp.)	4257.4 ± 257.0 4318.0 (exp.)
Z+jets	46.5 ± 8.3 50.2 (exp.)	1747.5 ± 163.7 2159.0 (exp.)
QCD	651.6 ± 4.0 684.4 (exp.)	51190.5 ± 313.1 53767.5 (exp.)
Observed yield	778 ± 28	57227 ± 239
Post-fit total background	811.4 ± 20.3	57195.5 ± 436.8

Table 5.16: Expected and observed yields and their total uncertainty (stat.+sys.) in the two purity categories.

obtained using the asymptotic CL_S method). Figure 5.96 show the resulting expected and observed 95% confidence level exclusion limits on the signal cross section as a function of the resonance mass for all signal models. The obtained limits are compared with the resonance production cross section times the branching fraction to WW, ZZ and WZ. To settle the

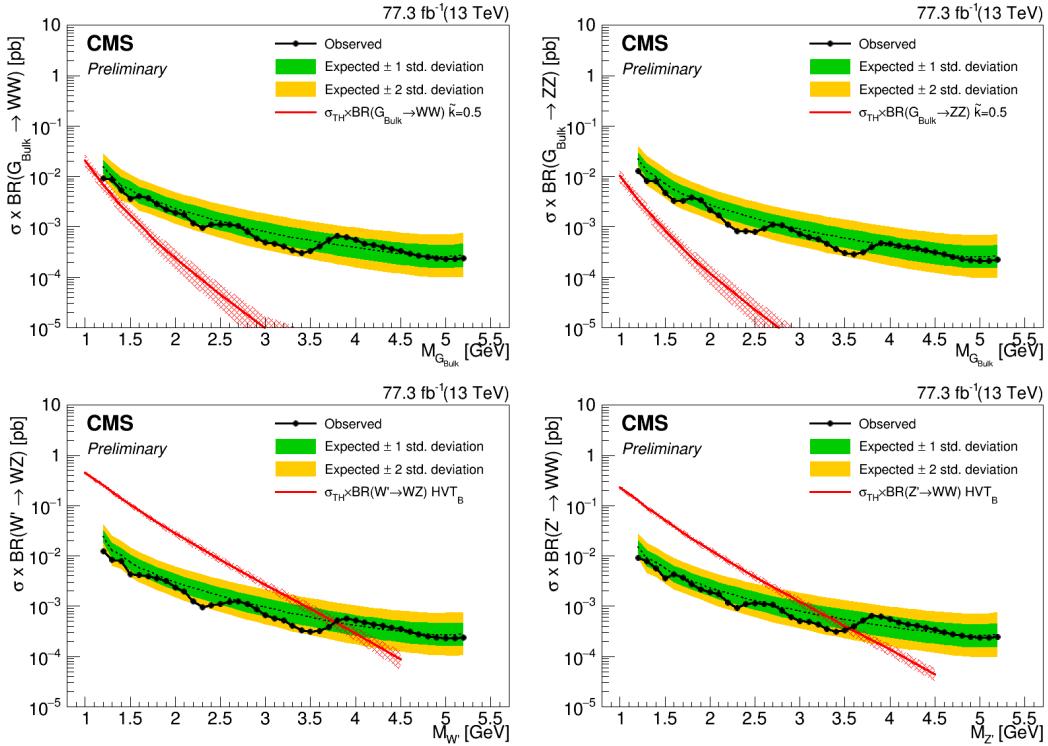


Figure 5.96: Expected limits obtained combining 35.9 fb^{-1} and 41.4 fb^{-1} of data of data after combining all purity categories for Bulk $G \rightarrow WW$ (top left), Bulk $G \rightarrow ZZ$ (top right), $W' \rightarrow WZ$ (bottom left) and $Z' \rightarrow WW$ (bottom right) signals.

question of whether the 3D fit method yields an improvement upon the 1D search method, we compare the limits obtained above to those obtained using the 1D fit method (the results from SearchII). In Figure 5.97 we see a 20–30% improvement for all signal hypothesis with the new method (comparing the 1D and 3D 2016 limits only), and a total gain of about 35–40% when combining the 2016 and 2017 datasets. In Figure 5.98, we additionally compare the 3D limits to those obtained by the ATLAS collaboration in a similar search [61] and find this search to be up to 35% more sensitive for the two signal scenarios considered (the G_{bulk} limits can not be compared due to different values of \tilde{k}). Finally, in Figure 5.99 we show a breakdown of the limits per purity category. As expected, the HPHP category dominates at low mass where background is high and the HPLP category dominates at high mass due to low background and high signal efficiency.

Pulls of nuisance parameters

As summarized in Section 5.3.11, we add a list of systematic uncertainties to the fit as nuisance parameters. To quantify how much the nuisances we insert differs from the ones preferred by the fit, we compute the pull

$$p_\theta = (\theta - \theta_{in})/\sigma_\theta \quad (5.20)$$

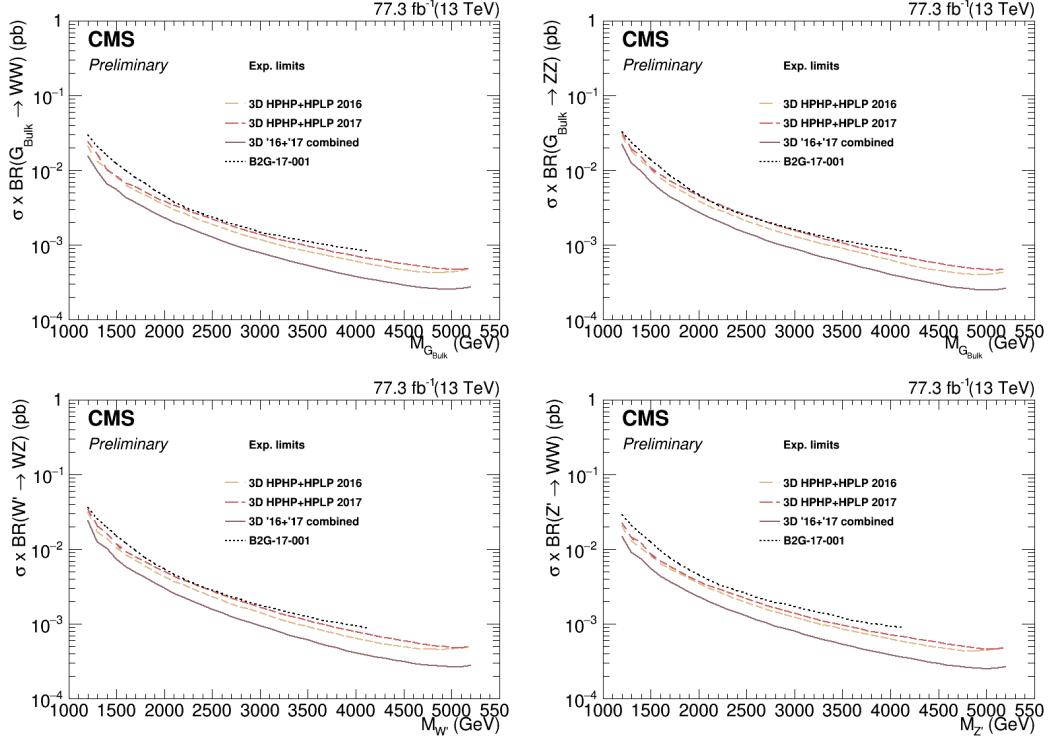


Figure 5.97: CA comparison of the 3D expected limits split into dataset (2016 and 2017), to that obtained with the 1D fit using 2016 data. Here for $G_{\text{bulk}} \rightarrow WW$ (top left), $G_{\text{bulk}} \rightarrow ZZ$ (top right), $W' \rightarrow WZ$ (bottom left) and $Z' \rightarrow WW$ (bottom right) signals.

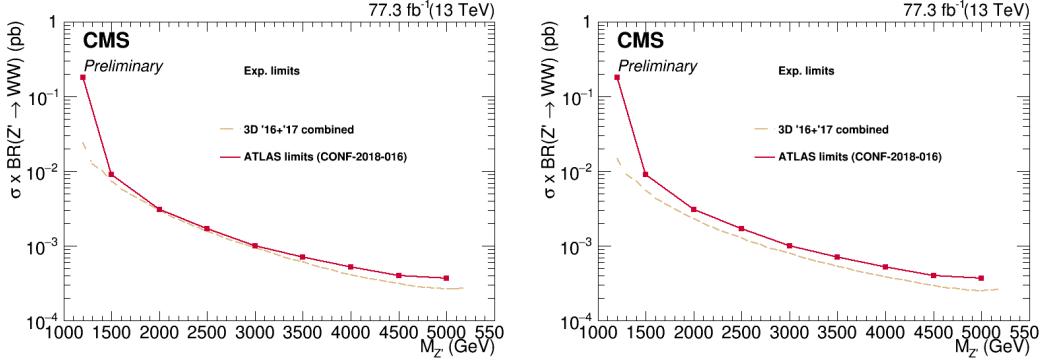


Figure 5.98: A comparison of the limits obtained above, to those by the ATLAS collaboration in a similar search [61]. Here for W' (left) and Z' (right) signal hypotheses.

SLAC where θ_{in} is the nuisance value pre-fit, θ the corresponding parameter post-fit and σ_θ its error, and its error bar calculated as the ratio between post- and pre-fit uncertainty. Figure 5.100 shows the pulls for a signal+background fit to the combined (2016+2017) dataset (left) and when fitting the two separately (right), here using a signal hypothesis corresponding to a 2 TeV G_{bulk} .

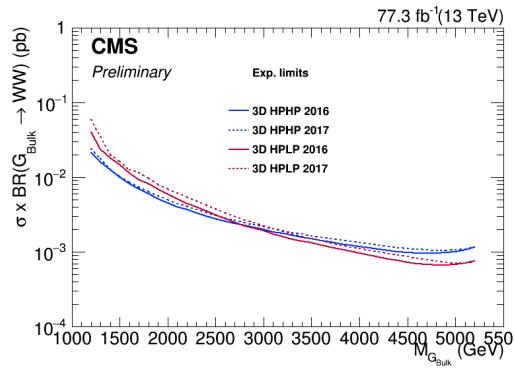


Figure 5.99: Limits split into purity categories (HPHP and HPLP) for a Bulk $G \rightarrow WW$ signal hypothesis.

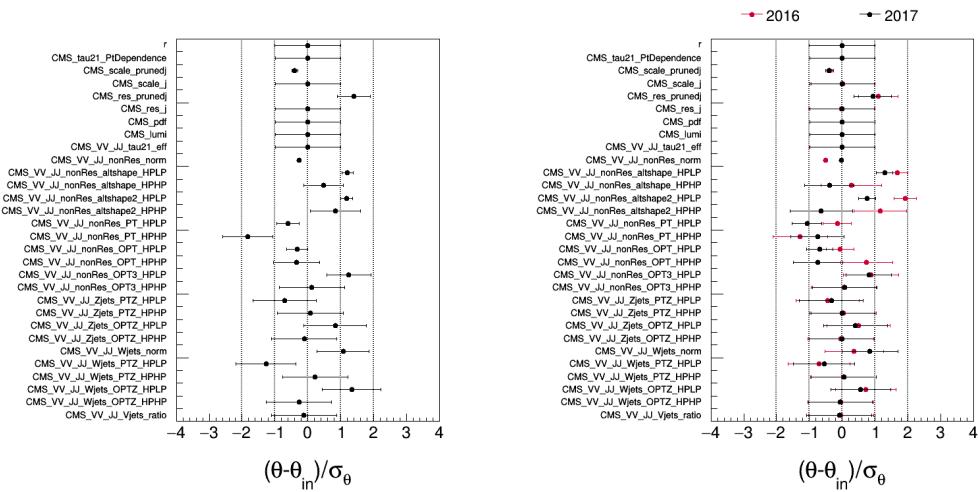


Figure 5.100: Pulls of each nuisance parameter for a combined signal+background fit to the combined 2016+2017 dataset (left) and when fitting the two separately.

5.4 Summary and outlook

In this chapter, we have followed the search for VV resonances in the all-hadronic final state through three stages and corresponding publications: From being one of the first ever analysis in the “boosted” final state at 13 TeV and the very first to take advantage of substructure at trigger level, through leading the development of a new W-tagging algorithm and mass corrections now default in CMS and finally ending with the development of a multi-dimensional fit for generic searches in jet groomed mass and dijet invariant mass.

Each analysis has built on significant improvements that came with the analysis before it: The substructure triggers and mass corrected softdrop jet mass are both used for the 3D fit, the early discovery of the softdrop signal efficiency dependence on p_T led us to derive corrections for it. Now the question which remains is: *What comes next?*.

A few ideas were already mentioned in the introduction to Search III, Section 5.3. The natural next step for this search is an incorporation of the VH(bb) and H(bb)H(bb) final states into the three-dimensional fit. Orthogonality between the three is guaranteed through b-tagging categories, as illustrated in Figure 5.101. This process is already underway, aiming

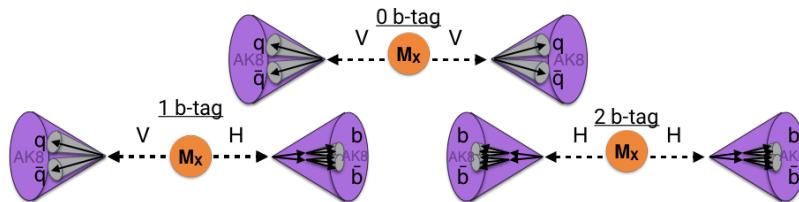


Figure 5.101: The VV, VH(bb) and H(bb)H(bb) analyses can all be incorporated into the multidimensional framework. Orthogonality between the analyses is ensured through b-tagging categories.

for a publication of the full Run 2 dataset (data collected in 2016, 2017 and 2018) in one common framework.

Secondly, after 14 TeV there will not be another increase in the collision center-of-mass energy at the LHC. That means that we want the best possible sensitivity when analyzing the dataset which is to come. One way of doing so is through changing the search method, as we did with the three dimensional fit, another is to work towards a better W-tagger.

Beyond that, and perhaps more interestingly, is the search for generic resonances peaking anywhere in the softdrop jet mass and dijet invariant mass spectrum, where the jets themselves could have other compositions than two subjets (for instance a scalar decaying to two vector bosons, whose decay products are merged into one jet; a 4-prong object). One caveat of the current setup is that it constrains the signal to be a 2-pronged object through its n-subjettiness cut. In order for the multi-dimensional framework to be truly generic, the tagger needs to be replaced by a generic anti-QCD tagger. Such a tagger works as an anomaly detector by encoding the probability density function for quark/gluon jets as a function of certain variables, variables for which signal jets are assumed to have a different probability density. In our case, good variables would for instance be groomed mass or substructure, as any generic signal is assumed to be peaking in softdrop mass and have some (unknown) substructure. The tagger itself is given no information about what a potential signal looks like and will only return the probability of any jet being a QCD jet. Such taggers are usually deep neural network (DNN) based where the quark/gluon jet PDF is obtained through training of the network, as demonstrated in [62, 63].

In order for such an encoding to work, the deep neural network needs access to the features distinguishing q/g jets from signal jets without these features being biased towards any signal in particular. The network must learn how to encode “non-substructure”.

As a side project in parallel to working on the multi-dimensional fit, I spent the last half year of my PhD working on a deep neural network capable of discriminating q/g jets from W jets in order to improve W-tagging performance and improve the search sensitivity of VV analyses to come. Based solely on jet constituent four-vectors, the idea is to let the neural network itself compute grooming and substructure like variables, without feeding it any high-level features (like softdrop mass and τ_{21}). This type of architecture is, in addition to improving W-tagging performance, ideal for the purpose described above: Encoding QCD in terms of substructure like features. The final chapter of this thesis is therefore dedicated to the two last points: How to improve W-tagging in CMS for future analyses and how to design a neural network capable of learning jet substructure in an unbiased manner.

CHAPTER 6

A Lorentz Invariance Based Deep Neural Network for W-tagging

6.1 Infusing deep neural networks with physics

The previous chapter ended by mentioning two ingredients that will become important for future searches with the multi-dimensional fit: A better vector boson tagger, and a generic anti-QCD tagger for signal independent searches. As a side project during my final PhD semester, I worked on a solution for the first, which has the added benefit of being a stepping stone towards the latter. This is what I will cover in the final chapter of this thesis.

When applying machine learning to particle physics problems, the input has historically consisted of pre-computed high-level features (quantities based on lower-level variables and certain theoretical assumptions). With the rise of deep learning however, computational graphs have achieved an increased capability to find even the smallest correlations in datasets, allowing them to construct complex features on their own. The deep neural network (DNN) I will present in the following is based on the assumption that, given sufficient instructions about the laws of Nature, a neural network should be capable of reconstructing its own high-level features based on lower-level variables only. In addition, if smartly designed, the network should be capable of finding novel correlations and physical features, *a-priori* unknown, by allocating a physical meaning to the training weights deep within the network. The deep neural network I will present here, is trained to discriminate quark/gluon jets from W -jets. However, as I will discuss in the final section of this chapter, it is also the perfect starting point for developing a generic anti-QCD tagger.

The work presented in the following has not been published and still qualifies as work in progress. However, I believe developing taggers such as these is of great importance for future versions of the searches presented here, and is something I hope to continue working on in the future..

COLA:

$$k_{M_{ij}} = \begin{bmatrix} E_1 & E_2 \\ P_1^1 & P_1^2 \\ P_2^1 & P_2^2 \end{bmatrix} \sim k_{M_{ij}} = k_{M_{ii}} \cdot L_{ij} = \begin{bmatrix} E_1 & E_2 \\ P_1^1 & P_1^2 \\ P_2^1 & P_2^2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & M=4 \end{bmatrix} = \begin{bmatrix} E_1 + E_2 & E_1 \\ P_1^1 + P_1^2 & P_1^1 \\ P_2^1 + P_2^2 & P_2^1 \\ P_3^1 + P_3^2 & P_3^1 \end{bmatrix}$$

(4,4)

(4,2)

(4,1)

linear combinations of momenta

$\sum E$
 $\sum P_x$
 $\sum P_y$
 $\sum P_z$

"What can we teach the machine?" → "What can we learn from the machine?" Work in progress

6.2 LoLa

LoLa is a deep neural network architecture which was first introduced for top tagging [?]. It is based on the idea that, given enough information about the laws of Nature, a neural network should be capable of calculating jet substructure observables on its own given only low-level information. The network is designed to discriminate between AK R=0.8 jets originating from W bosons from those originating from quarks or gluons, solely based on the jet constituent four-vectors (variables with little discriminating power on their own) as illustrated in Figure 6.1. Rather than being fed high-level features, the neural network is

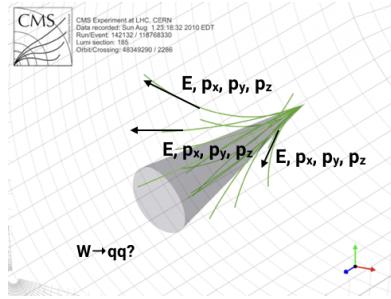


Figure 6.1: LoLa uses only jet constituent four-vectors as input to discriminate W from q/g jets.

given tools to perform calculations on Lorentz vectors using the Minkowski metric. Through two novel layers, linear combinations similar to jet clustering and jet substructure algorithms are performed, allowing the algorithm to create its own substructure variables. Additionally, training weights deep within the network correspond to physical quantities reconstructed by the algorithm; distance between particles, masses and energies, linear combinations of particle four-vectors etc. Besides the end goal of discriminating Ws from quarks and gluons, one could therefore hope to learn of new correlations separating QCD from vector boson jets.

6.2.1 Architecture

The LoLa architecture is designed as a four layer deep, feed-forward sequential network doing supervised learning on fixed size input vectors. Two novel layers are introduced, the Combination Layer (CoLa) and the Lorentz Layer (LoLa), which perform basic jet clustering and substructure calculations as well as implements the Minkowski metric. These two layers are then followed by two fully connected layers, consisting of 100 and 50 nodes respectively, before the final output is computed using a Softmax activation function, yielding two output probabilities between 0 and 1. The loss function to be minimized is ‘categorical crossentropy’ (or log loss) where the two categories in use are W versus non-W probabilities. Only the W jet probability is stored. The optimizer used in the training is the, now standard, ADAM optimizer [?], which adapts the learning rate of the model parameters during training. The code itself is written using the Keras [?] interface with a TensorFlow [?] backend. The full architecture with input and output dimension per layer is shown in Figure 6.2. The three first boxes are matrices, while the final four boxes correspond to vectors of different length. In the following, each layer will be explained in detail.

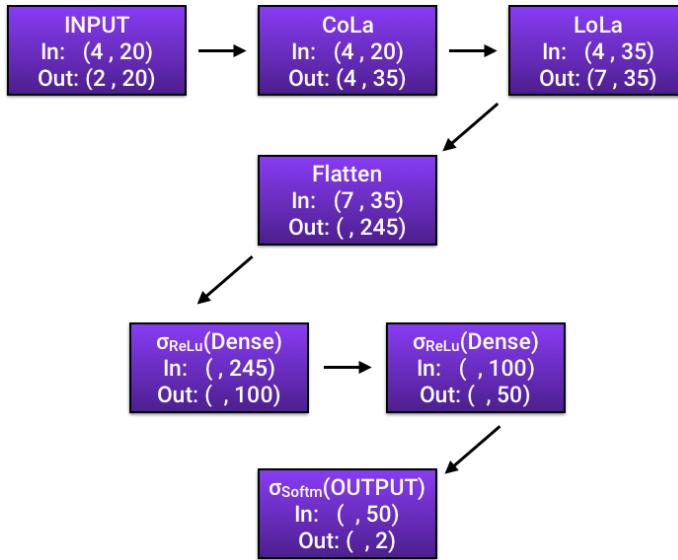


Figure 6.2: The full LoLa architecture. “In” denotes the dimension of the input tensor to the given layer, “Out” is the output tensors dimensions.

6.2.2 Input

This algorithm is trained to discriminate between fully merged hadronic W-jets coming from the process $G_{\text{bulk}} \rightarrow WW \rightarrow q\bar{q}q\bar{q}$ (where $M_{G_{\text{bulk}}} = 0.6 - 4.5 \text{ TeV}$), and quark/gluon jets from a QCD sample generated with PYTHIA8/Pythia 8. All jets are clustered with the anti- k_T algorithm with a distance parameter of $R=0.8$, with the PUPPI pileup removal algorithm applied. In addition, they are required to have $p_T > 200 \text{ GeV}$ and $|\eta| < 2.5$. Jets are defined as W-jets if they are matched to a generator level hadronically decaying W bosons, with the following matching criteria: The generated vector boson needs to be within $\Delta R < 0.6$ of the jet axis, and the quark decay products need to be within $\Delta R < 0.8$ of the jet axis. The p_T and η distribution of signal and background jets, is shown in Figure 6.3. From these signal and background jets, only the jet constituent four vectors of the 20 highest- p_T particles are used as input to the deep neural network: E , p_x , p_y and p_z . I use 20 constituents as any larger number has a negligible affect on the performance, while performance tends to drop once going below 15. The input is therefore a $4 \times N = 20$ matrix for each signal and background jet, one four-vector for each of the 20 jet constituents:

$$x_{\mu,i} = \begin{pmatrix} E^1 & E^2 & \dots & E^N \\ p_x^1 & p_x^2 & \dots & p_x^N \\ p_y^1 & p_y^2 & \dots & p_y^N \\ p_z^1 & p_z^2 & \dots & p_z^N \end{pmatrix} \quad (6.1)$$

The total number of jet constituents is shown in Figure 6.4, and the input variables (here for all constituents) is shown in Figure 6.5.

It is clear that the input variables provide little discriminating power on their own. Therefore, the network must learn how to derive other physical quantities where the signal and background PDFs differ to a larger extent. This is achieved through the two custom layers described in the following.

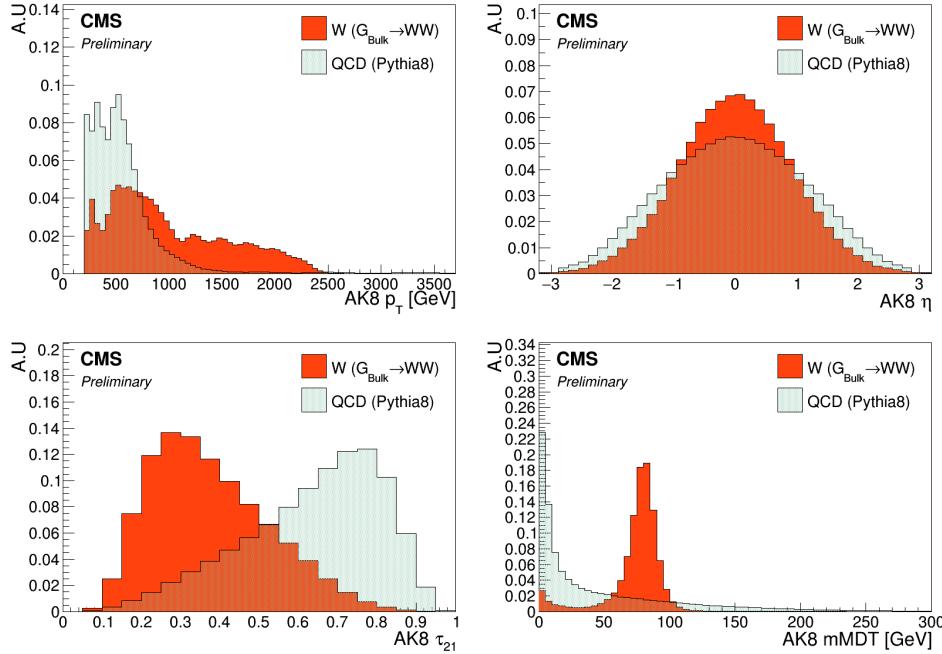


Figure 6.3: Jet p_T (top left), η (top right), τ_{21} (bottom left) and softdrop jet mass (bottom right) for signal and background jets.

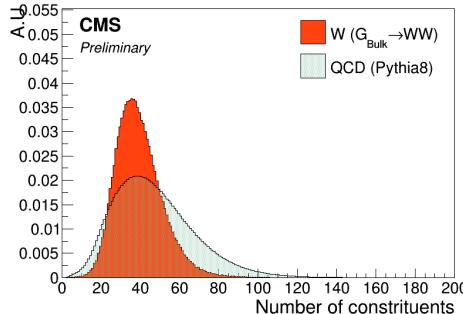


Figure 6.4: The number of jet constituents for signal (red) and background (blue). Only the 20 highest- p_T constituents are used during training.

6.2.3 The Combination Layer

The Combination Layer (CoLa) consists of a matrix which, when taking the scalar product with the input matrix, compute linear combinations of the jet constituents, similar to what is done in recombination jet algorithms. The main goal here is to create additional four-vectors as input for the next layer. The CoLa matrix is a concatenation of the following: A vector of 1's of length N , the $N \times N$ identity matrix ($N = 20$) and a matrix of $N \times M$ trainable weights.

$$C_{i,j} = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 & w_{1,N+2} & w_{1,N+3} & \dots & w_{1,(N+2)+M} \\ 1 & 0 & 1 & \dots & 0 & w_{2,N+2} & w_{2,N+3} & \dots & w_{2,(N+2)+M} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 0 & 0 & 0 & 1 & w_{N,N+2} & w_{N,N+3} & \dots & w_{N,(N+2)+M} \end{pmatrix} \quad (6.2)$$

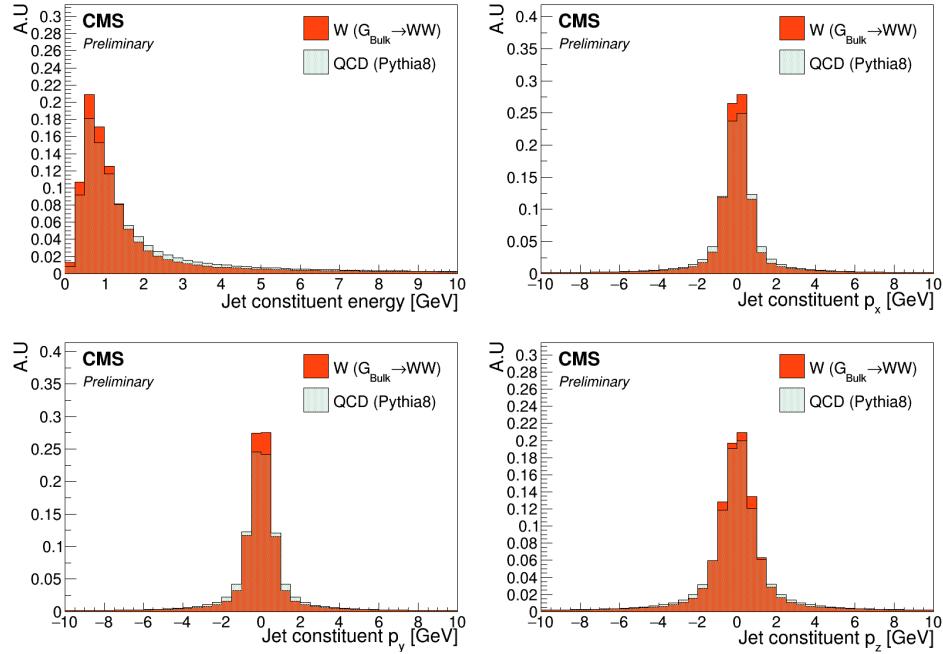


Figure 6.5: Energy (top left), p_x (top right), p_y (bottom left) and p_z (bottom right) for all jet constituents. These values are used as input to the neural network training.

When performing the following multiplication

$$x_{\mu,i}^C = x_{\mu,i} C_{i,j} \quad (6.3)$$

the resulting output matrix will have dimensions $4 \times (1 + N + M)$ and consists of the following: A first column containing the sum of all constituent momenta, the four-momenta of each individual constituent, and $M=14$ different linear combinations of particles with trainable weights. The first corresponds to the neural network computing the four-vector of the “full” jet, at least the full jet in terms of its 20 highest- p_T constituents. The second, simply passes each original constituent four-momentum to the next layer. The final, and most interesting part, lets the network construct alternative subjet four-vectors by letting it weigh constituents up and down as it sees fit in order to reach optimal discrimination power. As an example, lets look at the effect of CoLa in the simple case of only two input jet constituents and two trainable linear combinations:

$$\begin{pmatrix} E^1 & E^2 \\ p_x^1 & p_x^2 \\ p_y^1 & p_y^2 \\ p_z^1 & p_z^2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & w_{1,4} & w_{1,5} \\ 1 & 0 & 1 & w_{2,4} & w_{2,5} \end{pmatrix} = \begin{pmatrix} E^1 + E^2 & E^1 & E^2 & w_{1,4}E^1 + w_{2,4}E^2 & w_{1,5}E^1 + w_{2,5}E^2 \\ p_x^1 + p_x^2 & p_x^1 & p_x^2 & w_{1,4}p_x^1 + w_{2,4}p_x^2 & w_{1,5}p_x^1 + w_{2,5}p_x^2 \\ p_y^1 + p_y^2 & p_y^1 & p_y^2 & w_{1,4}p_y^1 + w_{2,4}p_y^2 & w_{1,5}p_y^1 + w_{2,5}p_y^2 \\ p_z^1 + p_z^2 & p_z^1 & p_z^2 & w_{1,4}p_z^1 + w_{2,4}p_z^2 & w_{1,5}p_z^1 + w_{2,5}p_z^2 \end{pmatrix}$$

In the two last columns, the neural network makes two “subjet” four-vectors by weighting the relative contribution of each particle as it sees fit. This is similar to jet grooming (Section 4.5.1) or PUPPI pileup subtraction (Section 4.3.2), and should allow the network to learn which constituents are part of the hard scatter and which are not. The $x_{\mu,i}^C$ matrix is finally passed on to the next layer, the Lorentz Layer.

6.2.4 The Lorentz Layer

The Lorentz Layer (LoLa) is responsible for encoding how particles move in space-time through a simple set of rules. Each column (four-vector) of $x_{\mu,i}^C$, is used to compute, and

afterwards is replaced by, the following $k = 7$ features:

$$x_{k,i}^L = \begin{pmatrix} m^2(x_{\mu,i}^C) \\ p_T(x_{\mu,i}^C) \\ w_{ij}^E E(x_{\mu,j}^C) \\ w_{ij}^{s1} \sum d^2(x_{\mu,i}^C, x_{\mu,j}^C) \\ w_{ij}^{s2} \sum d^2(x_{\mu,i}^C, x_{\mu,j}^C) \\ w_{ij}^{m1} \min d^2(x_{\mu,i}^C, x_{\mu,j}^C) \\ w_{ij}^{m2} \min d^2(x_{\mu,i}^C, x_{\mu,j}^C) \end{pmatrix} \quad (6.4)$$

Going through from top to bottom, these are:

- The invariant mass and p_T of each four-vector
- A linear combination of all four-vector energies where each is scaled by a trainable weight
- The sum of distances between the four-vector under consideration and every other column reweighted with a trainable weight
- The minimum distance between the four-vector under consideration and every other column where each distance again is reweighted by a trainable weight.

The Minkowski metric enters explicitly in the first and in the last four calculations, where the neural network is told to abide by the rules

$$m^2(x_{\mu,i}^C) = g^{\mu\nu} x_{\mu,i}^C x_{\nu,i}^C \quad (6.5)$$

and

$$d^2(x_{\mu,i}^C, x_{\mu,j}^C) = (x_{\mu,i}^C - x_{\mu,j}^C)_\mu g^{\mu\nu} (x_{\mu,i}^C - x_{\mu,j}^C)_\nu \quad (6.6)$$

with $g^{\mu,\nu} = [-1, 1, 1, 1]$, when calculating the invariant mass and distance between particles/subjets. This tells the neural network to use a space-time geometry in all its calculations to respect Lorentz Invariance. The four final rows of LoLa are the most interesting: Here the network computes quantities similar to n-subjettiness by summing up the distances between all constituents, the jet axis and the subjets produced by CoLa. If, for instance, the network has been capable of reconstructing two hard subjets in the final columns of CoLa, which do linear combinations of particles, it can create its own “ τ_2 ” variable by taking the distance between those subjets and all the jet constituents (and weighing down the column corresponding to the full jet four-vector, column one). Then it can do the same by calculating the distance between the full jet four-vector and all constituents (now weighing down the linear combinations) and compute “ τ_2 ”.

The two custom layers, CoLa and LoLa, therefore come together in order to encode jet clustering and substructure in a novel way. They provide the network with the necessary tools in order to create its own physical quantities, through linear combinations with trainable weights, which then again are used to produce other physical quantities with new trainable weights. This allows the network full freedom to explore all interesting particle correlations, where the resulting output features have a physical meaning that can be probed. LoLa turns the question “What can we teach the machine?” around to “What can we learn from the machine?”.

6.3 Performance

The deep neural network is trained on 320k signal and background jets for up to 100 epochs, but allow for an early stopping after ten epochs if the loss is stable. The test sample consists of 60k W jets and 60k quark/gluon jets. To quantify the performance we look at the signal efficiency versus mistagging rate comparing the performance of LoLa to that of the taggers used previously in this thesis: PUPPI softdrop with τ_{21} and PUPPI softdrop with τ_{21}^{DDT} . The performance of these three different taggers, is shown in Figure 6.6. The point where the blue curves end, represent the signal efficiency for a mass cut of $65 \text{ GeV} < \text{Softdrop jet mass} < 105 \text{ GeV}$, here roughly 70%. We clearly see that LoLa performs significantly better

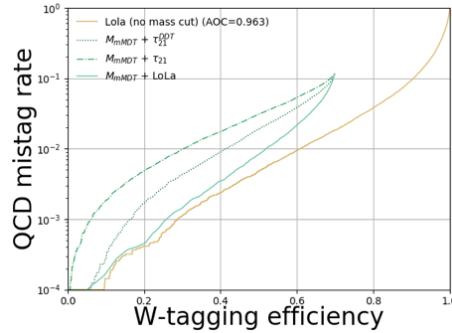


Figure 6.6: Performance of LoLa compared to other W-tagging discriminants in the background-signal efficiency plane: PUPPI softdrop + τ_{21} (dashed blue), PUPPI softdrop + τ_{21}^{DDT} (dotted blue), LoLa with a softdrop mass window applied (solid blue) and the nominal LoLa tagger with no mass cut applied.

than the current baseline W-taggers based on τ_{21} and τ_{21}^{DDT} , with a roughly 20% higher signal efficiency at a given mistagging rate. LoLa also has a higher signal acceptance, as it can be used without a mass window applied. If LoLa were to replace the tagger used in Search II (a better comparison than Search III as the latter uses a rather unconventional mass window), which has a signal efficiency of $\sim 42\%$ at a 2% mistagging rate for a single jet, the signal efficiency for the same mistagging rate would be 65%, a 55% increase. For an analysis requiring two tagged jets, that would imply going from an 18 to a 43% total signal efficiency, a significant gain.

6.4 p_T and mass dependence

Despite being a key feature, absolute performance is not all that quantifies how good one tagger is compared to another. Once a tagger is planned to be used in physics analysis, there are three key questions one needs to consider:

- Is the absolute performance better (compared to common methods)?
- Is the tagger p_T -dependent?
- Does the tagger sculpt the mass spectrum?

These three measures are equally important in quantifying performance and, in the following, I will attempt to explain why this is the case and which approaches are used here in order to tackle them.

Any deep neural network trained to distinguish W jets from q/g jets, will naturally learn that p_T and mass are discriminating features unless it is penalized for it. Figure 6.7 shows

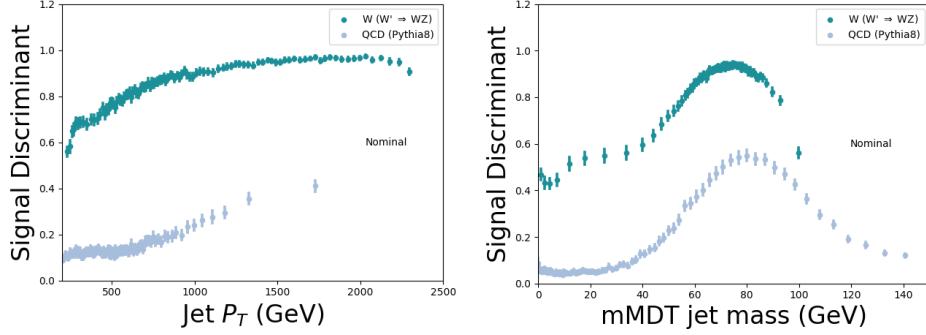


Figure 6.7: The LoLa discriminant as a function of jet p_T (left) and softdrop jet mass (right). A strong correlation with both variables is observed

the LoLa discriminant as a function of jet p_T and softdrop jet mass. A strong correlation is observed both for signal and for background jets (closer to 1 means more signal like), with a rising slope as a function of p_T (meaning the network interprets a higher jet p_T as more signal like) and a bump around the W mass for both signal and for background.

6.4.1 p_T decorrelation

A tagger which is p_T dependent is a problem for the following reasons: Firstly, the signal efficiency is variable, which requires a working point that scales with p_T . That in itself is not problematic and can easily be computed. However, it implies that when computing efficiency scale factors from data, a range of different scale factors for different working points is required. In addition, the performance is measured at low p_T , a region where the tagging efficiency can be substantially different from the analysis signal region due to the strong p_T correlation present. Finally, the dijet invariant mass is intrinsically linked to the p_T spectrum, meaning that any p_T dependence in addition can introduce sculpting of the dijet invariant mass spectrum.

From the top left plot in Figure 6.3, one clearly sees that the jet p_T distribution is very different for signal and for background. In order to avoid that the network learns jet p_T to be a discriminating feature, I therefore compute a jet-by-jet weight intended to flatten the jet p_T spectrum. This weight is passed as a sample weight to the training set, reweighting each jets contribution to the total loss (making high mass QCD jets and low mass signal jets count more). Figure 6.8 shows the jet p_T distribution before (solid lines) and after applying a p_T -weight (dashed lines). The training is then repeated,

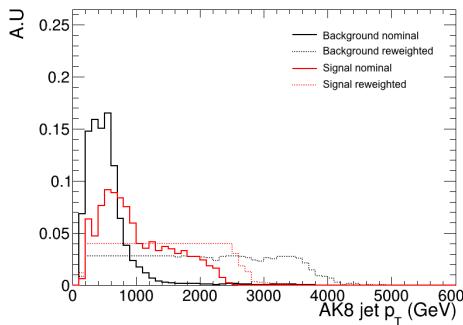


Figure 6.8: Jet p_T distribution before (solid lines) and after (dashed line) applying a weight intended to flatten the jet p_T spectrum.

this time passing a sample weight with each jet, and the final discriminant compared to the nominal training. Figure 6.9 shows the performance of the same taggers as above but with one additional line, LoLa p_T -reweighted. A clear drop in performance is observed, as

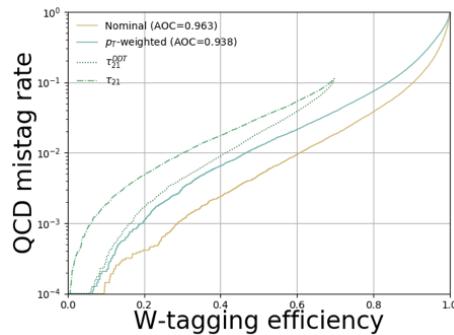


Figure 6.9: Performance of the p_T -reweighted LoLa tagger (solid blue) and the nominal LoLa tagger (solid yellow).

expected when removing information from the training. However, when we again look at the discriminant output as a function of jet p_T in Figure 6.10, the correlation we observed before has vanished and we are left with a tagger not depending on the jet p_T . For completeness,

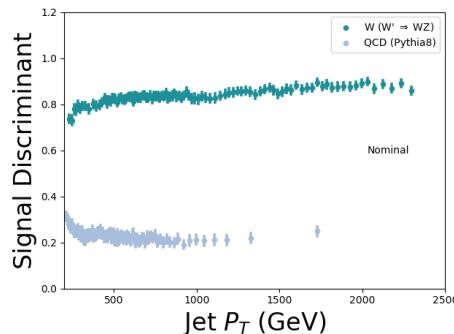


Figure 6.10: The LoLa discriminant as a function of jet p_T after training with a weight intended to flatten the sample p_T spectrum.

Figure 6.11 shows the τ_{21} and τ_{21}^{DDT} discriminants versus jet p_T . Whereas the nominal LoLa discriminant had a much larger correlation with jet p_T than the τ_{21} -based taggers, the p_T -reweighted version is as decorrelated from p_T as the τ_{21} variables while still exhibiting a better absolute performance than the baseline taggers. In summary, reweighting strategies as the one described above yield a loss in overall performance, as expected when removing information from the training. However, the p_T dependence of the tagger is strongly reduced, meaning that it might perform better overall in physics analysis when systematic uncertainties are taken into account. There is therefore no clear answer as to which method is better before running a full analysis including systematics for p_T -dependent tagging.

6.4.2 Mass sculpting

Any smart deep neural network intended to separate Ws from quarks and gluons, will inevitably learn the W mass as it clearly is very different from the q/g mass. Unfortunately, as these taggers are meant to be used in physics analysis where we often estimate the background in mass sidebands, this has some undesired side effects. If a deep neural network

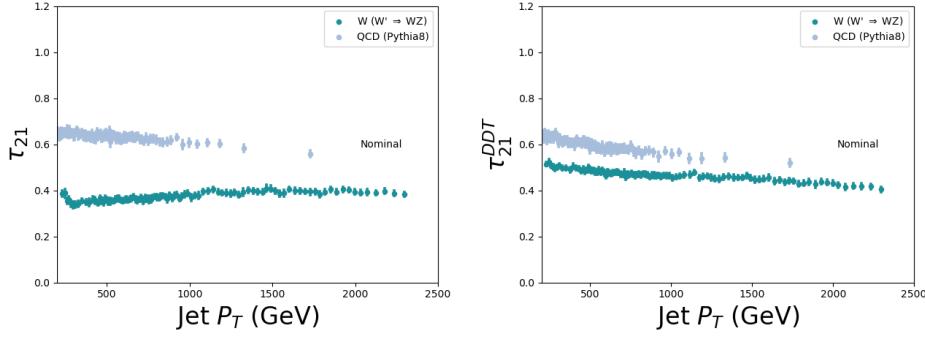


Figure 6.11: The τ_{21} (left) and τ_{21}^{DDT} (right) discriminant as a function of jet p_T .

has learned the mass then, after applying a cut on the discriminant, the background jet mass distribution becomes severely sculpted and difficult to constrain.

After applying a cut on the LoLa discriminant corresponding to a 1% mistagging rate, we see in the left plot in Figure 6.12 that the W jet signal shape is nicely retained. In addition, there are no QCD jets left at low mass so no jet mass window is needed when using this tagger, leading to a significantly higher signal acceptance. However, when looking more closely at the QCD background on the right plot of Figure 6.12, where all histograms are normalized to unit area, we see that the bulk of the remaining 1% QCD jets is right below the W mass peak and has been sculpted to look exactly like the signal. This mass sculpting is in and on its own not

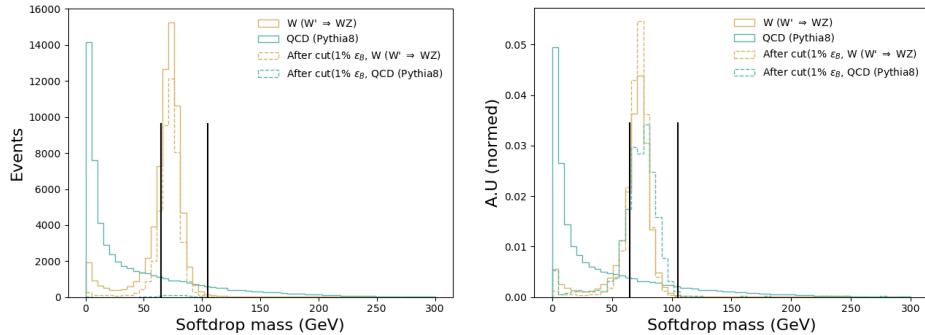


Figure 6.12: The softdrop jet mass distribution before (solid lines) and after (dotted lines) a cut on the LoLa discriminant corresponding to a 1% mistagging rate has been applied. The left plot shows the real number of events left after the cut, the right is normalized to area.

a problem, the tagger still manages to get rid of most of the background. However, in many physics analysis, in order to evaluate the background rate in the data signal region, mass sidebands are used. If the background distribution is peaky rather than smoothly falling, the shape and consequently the expected yield is very difficult to constrain. That leads to large uncertainties on the background rate and might eventually make an analysis less sensitive than when using a tagger with a worse absolute performance, but reduced mass correlation. In addition, if one were to search for peaks in the softdrop jet mass, as is the case for the multidimensional fit, this becomes increasingly difficult when attempting to fit a potential signal peak on top of a peaking background.

It should again be mentioned, that also for the baseline taggers based on τ_{21} , mass sculpting is a known feature. Figure 6.13 shown the same softdrop jet mass spectrum before and after a cut corresponding to a 1% mistagging rate on τ_{21} (left) and τ_{21}^{DDT} (right). Here τ_{21} clearly exhibits mass sculpting, but not as peaky as was the case for LoLa. τ_{21}^{DDT} exhibits the least

amount of sculpting, but is also the tagger with the worst absolute performance.

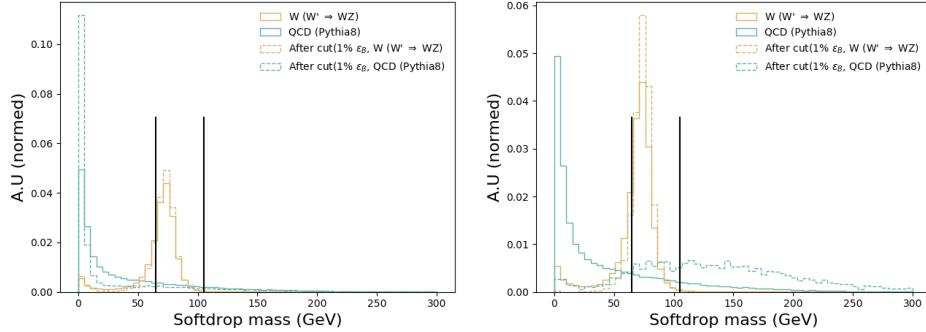


Figure 6.13: The softdrop jet mass distribution before (solid lines) and after (dotted lines) a cut on τ_{21} (left) and τ_{21}^{DDT} (right). All spectra are normalized to unit area.

I have not yet had the chance to implement a mass decorrelation strategy for LoLa, but I see two ways going forward: The first is, following the example of what was done to decorrelate LoLa from jet p_T , to pass a mass dependent sample weight to the training. LoLa would then be trained with a weight derived to flatten the two dimensional jet mass - jet p_T plane. Another option would be to train LoLa together with an adversarial, a dedicated deep neural network running in parallel to LoLa and attempting to learn the jet mass from the LoLa output. The total loss function would then be a sum of the two, where the better the adversarial is in learning the mass, the worse the total loss function gets. Both these options are something I'd like to explore in the future.

In summary, mass- and p_T -dependence are in their own right not a problem for a tagger. The problem occurs when using these taggers in actual physics analyses where background rate uncertainties and tagging p_T dependence uncertainties has a large impact on the final sensitivity. There is a trade-off between signal efficiency and (analysis-dependent) systematics. For LoLa, rather than choosing, I'd like to provide to different taggers: A nominal tagger, where no mass/ p_T -decorrelation is attempted, and a decorrelated version. Then both can be tested in a full analysis chain before deciding on which tagger to use when looking at data.

6.5 Validation on an independent sample

LoLa is additionally validated on independent samples as an unbiased measure of performance allowing to compare different CMS W-tagging algorithms to one another: A $Z' \rightarrow WW$ sample with $M_{Z'} = 3$ TeV produced with MadGraph and a QCD PYTHIA 8 sample in a p_T bin of 1000 to 1400 GeV. Here, only jets with $1000 \text{ GeV} < p_T < 1400 \text{ GeV}$ and $|\eta| < 1.5$ are used. The signal efficiency versus mistagging rate for LoLa compared to the baseline PUPPI Softdrop + τ_{21} tagger, is shown in Figure ???. As was pointed out in Section ??, a mass cut is not necessary when using LoLa, but has been added to this plot for completeness. A significant improvement in tagging efficiency is observed for LoLa compared to the default tagger, also when being validated on a sample completely independent from the training sample. The cut corresponding to a 30 % signal efficiency working point are used as reference working points when we will look at the tagging performance as a function of jet p_T and pileup in the following, and is marked by triangles in the plot. The signal efficiency and mistagging rate as a function of jet p_T , is shown in Figure 6.15. Again we observe the strong correlation between LoLa tagging efficiency and jet transverse momenta. There is, however, no point in the spectra where the τ_{21} tagger has a higher signal over background ratio than LoLa. LoLa

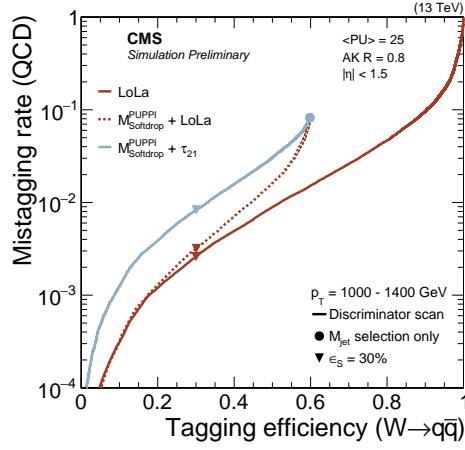


Figure 6.14: Performance of LoLa and PUPPI Softdrop + τ_{21} in the background-signal efficiency plane. The PUPPI softdrop jet mass selection of $65 < M_{SD} < 105 \text{ GeV}$, and the 30 percent efficiency points are indicated with symbols.

performs its worst at very high jet p_T , but in this region the background is very small (dijet invariant masses around 2.5-3 TeV) so the absolute performance here matters less than at lower p_T . Figure 6.16 shows the tagging efficiency and mistagging rate as a function of pileup.

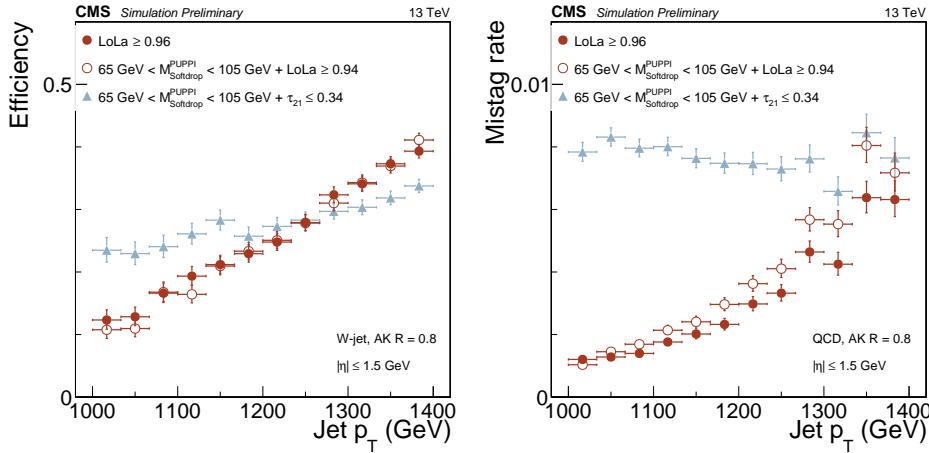


Figure 6.15: Efficiency (left) and mistagging rate (right) of the LoLa selection corresponding to a 30 percent signal efficiency as a function of jet p_T .

Both taggers under study are more or less decorrelated from pileup, with a flat efficiency up to 50 reconstructed primary vertices. In Run 3, this number is of course expected to be significantly higher, around 140-200, and the study should be redone up to a higher number of reconstructed primary vertices.

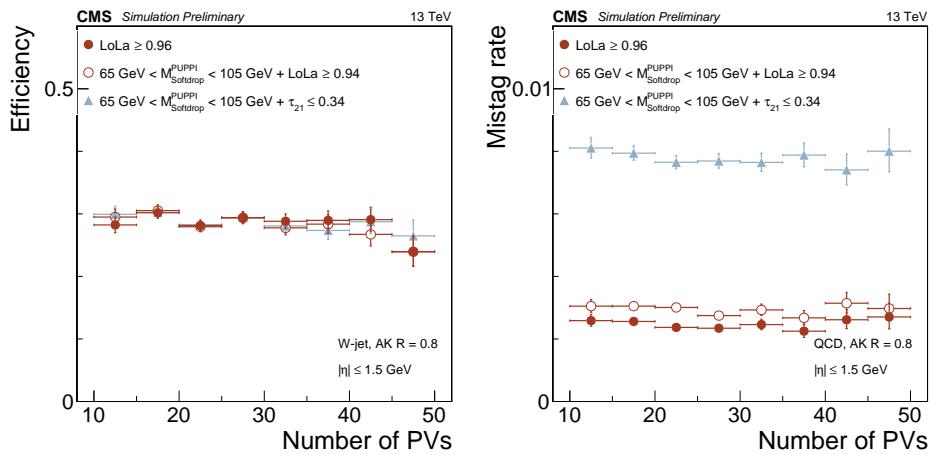


Figure 6.16: Efficiency (left) and mistagging rate (right) of the LoLa selection corresponding to a 30 percent signal efficiency as a function of the number of reconstructed vertices.

6.6 Summary and outlook

In this chapter, we have seen a promising new W-tagging algorithm for future VV searches. Its absolute performance is better than that of the baseline PUPPI softdrop + τ_{21} tagger up to a jet p_T of at least 1400 GeV, roughly corresponding to a dijet invariant mass of 2.5-3 TeV, and could lead to an increase in total signal efficiency from 18 to 43 % for the searches presented here. With a p_T decorrelation method already in place, it could already now be used for the one dimensional VV search presented in Search I and Search II. However, if to be used in the multidimensional search framework, a mass decorrelation method needs to be established. I have already outlined two possibilities of how to achieve this in Section 6.4.2, where one of these has already been shown to work in the context of p_T decorrelation. This is, as of this writing, left to future studies.

When discussing the future of the multidimensional search, I mentioned how a deep neural network such as the one presented here could be used to encode jet substructure in a way useful in order to develop a generic anti-QCD tagger. This has already been achieved by a parallel analysis team through the use of auto-encoders, published ten days before this writing and documented in [62], and has shown very promising results. However, this strategy is, to my knowledge after discussing with the authors, no longer pursued after observing that the auto-encoder version of LoLa was very difficult to decorrelate from the jet mass. It is my belief that this can be overcome by changing some of the features calculated in the Lorentz Layer (in [62], only the invariant mass is calculated and the other features listed in Equation 6.4 are stripped away) and this is something I would also like to pursue in future studies in order to achieve the truly generic search for boosted dijet resonances in the $M_{\text{jet}1}$ - $M_{\text{jet}2}$ - M_{VV} plane.

CHAPTER 7

Summary

APPENDIX A

Search I: Limits per mass category

The asymptotic limits obtained with 2.6 fb^{-1} of 13 TeV CMS data per mass and purity category.

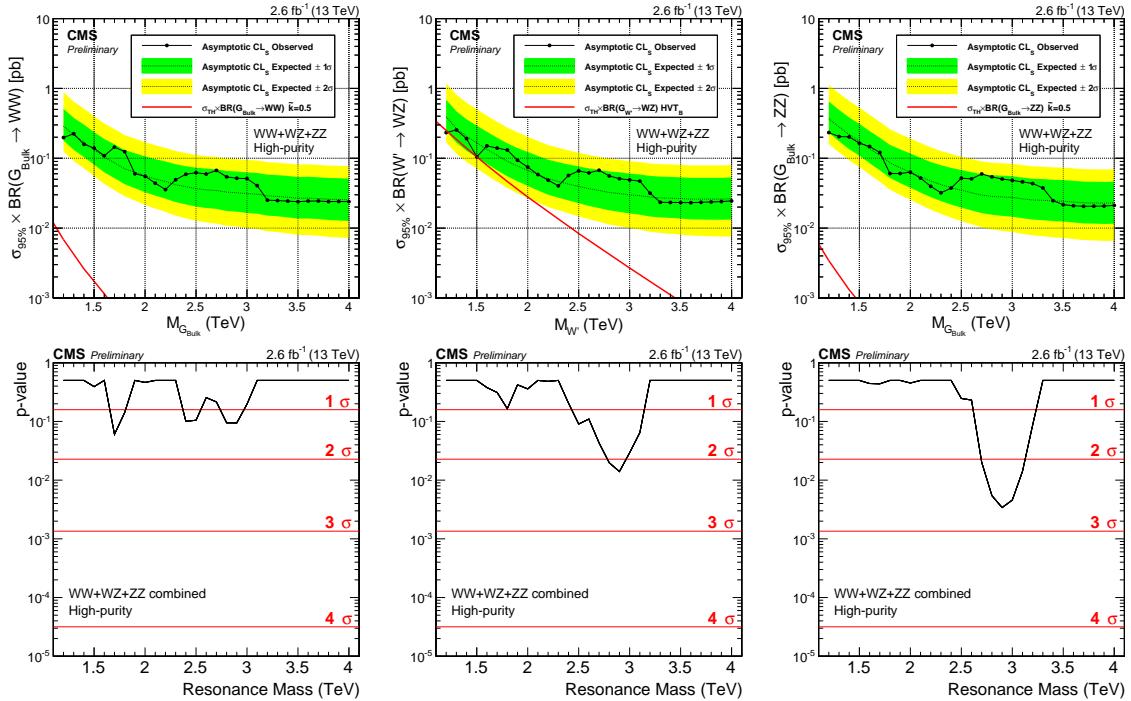


Figure A.1: Expected/observed limits and corresponding p-values obtained in the high purity category using 2.6 fb^{-1} of CMS data. Here for a Bulk $G \rightarrow WW$ (left), $W' \rightarrow WZ$ (middle) and $G \rightarrow ZZ$ (right) signal.

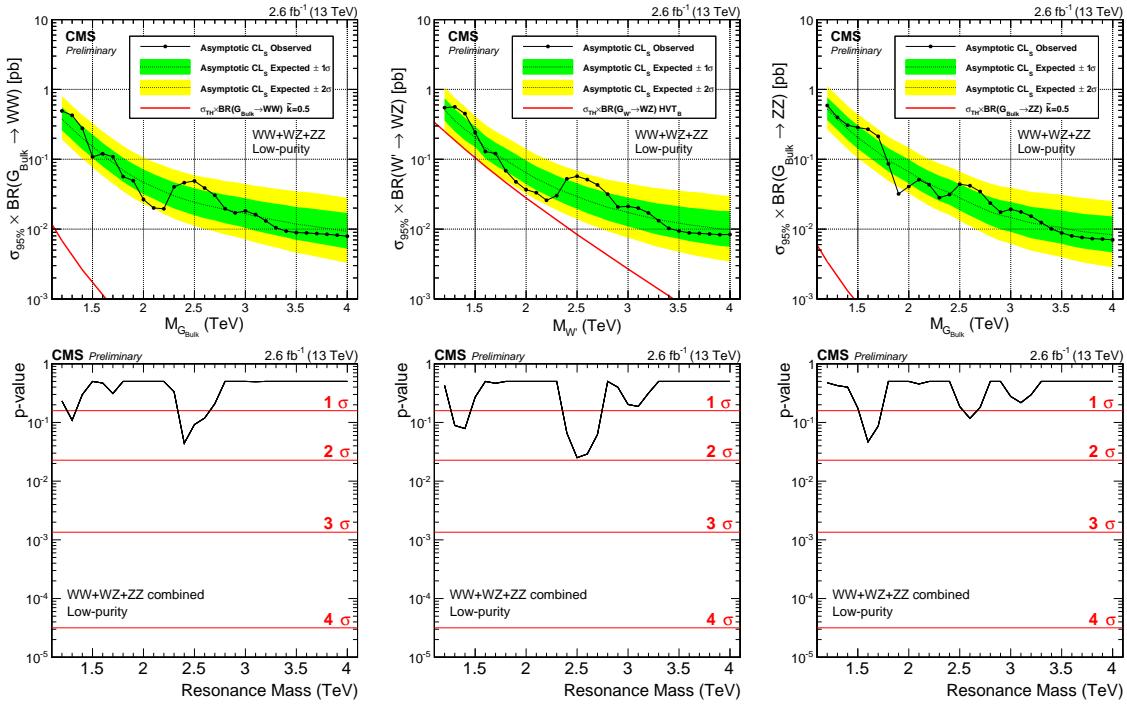


Figure A.2: Expected/observed limits and corresponding p-values obtained in the low purity category using 2.6 fb^{-1} of CMS data. Here for a Bulk $G \rightarrow WW$ (left), $W' \rightarrow WZ$ (middle) and $G \rightarrow ZZ$ (right) signal.

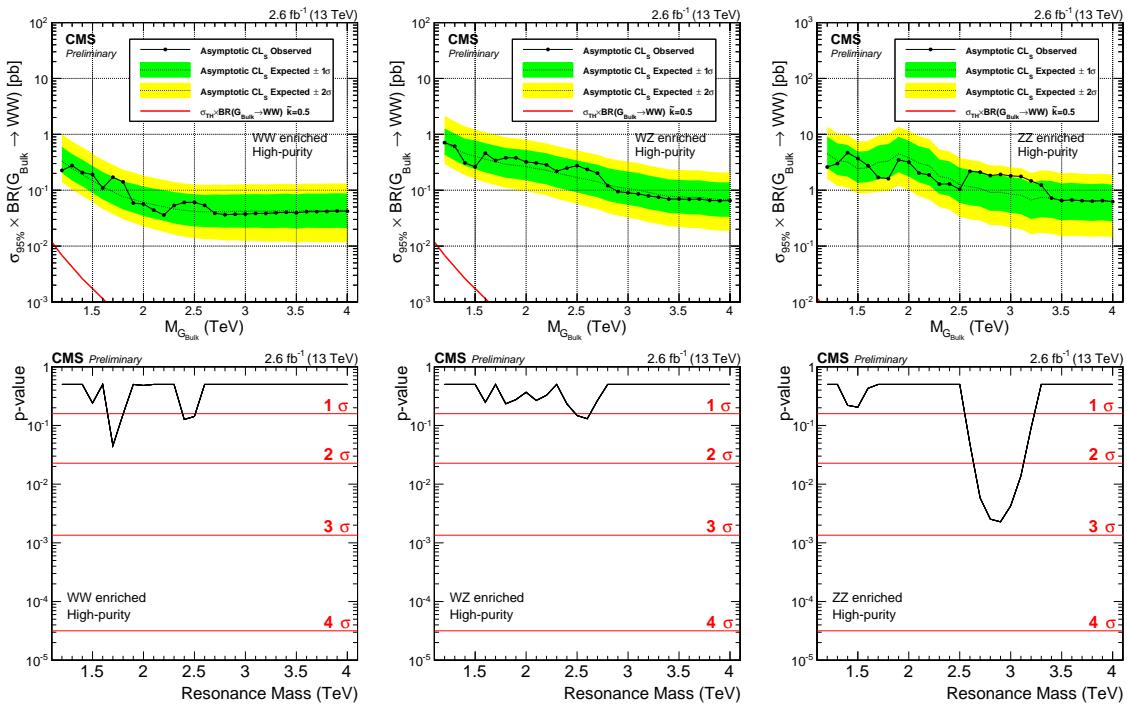


Figure A.3: Expected/observed limits and corresponding p-values obtained for the different mass categories using 2.6 fb^{-1} of CMS data. Here for a Bulk $G \rightarrow WW$ signal in the HP category

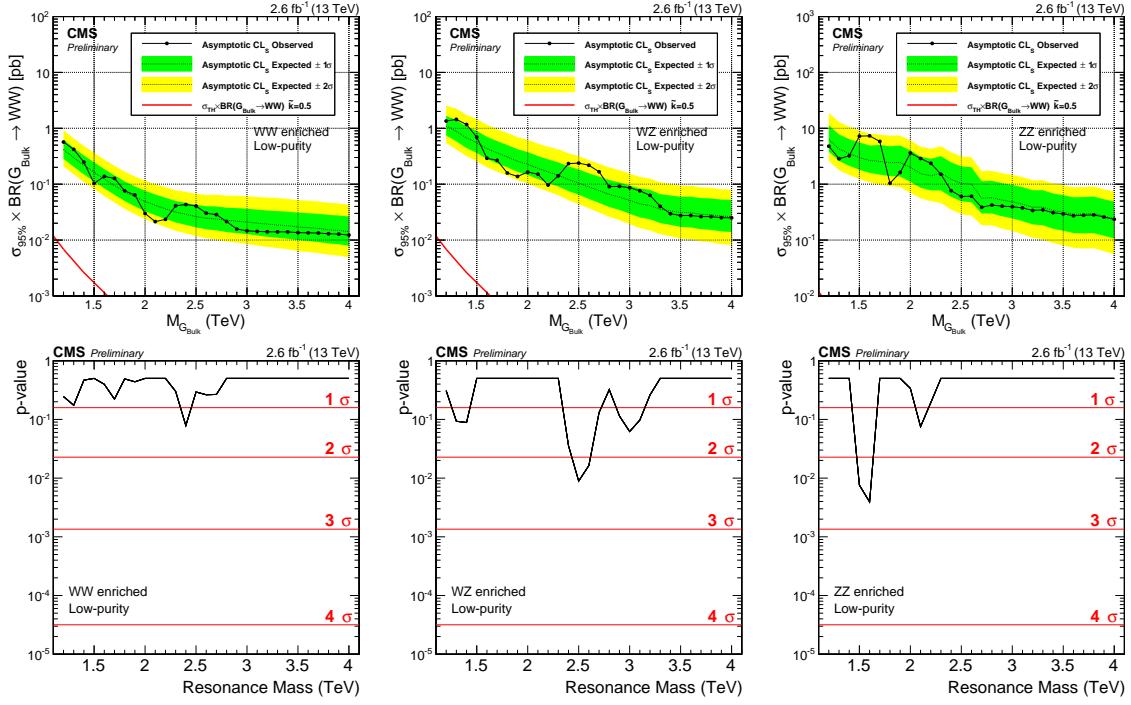


Figure A.4: Expected/observed limits and corresponding p-values obtained in the different mass categories using 2.6 fb^{-1} of CMS data. Here for a Bulk $G \rightarrow WW$ signal in the LP category.

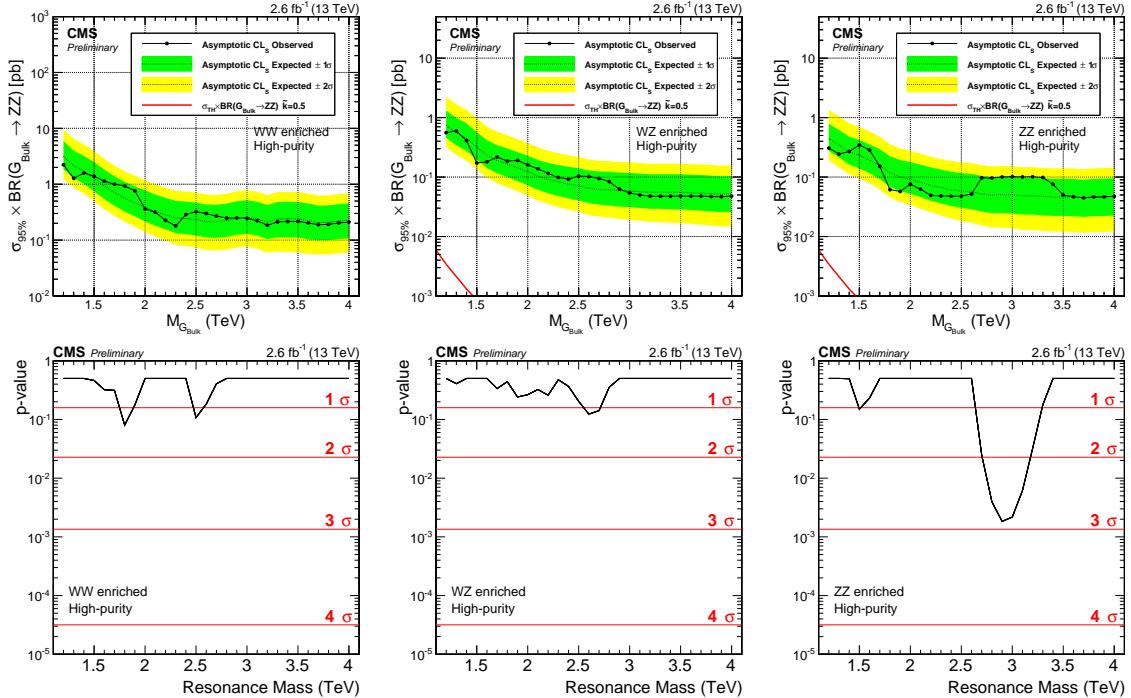


Figure A.5: Expected/observed limits and corresponding p-values obtained in the different mass categories. Here for a $G \rightarrow ZZ$ signal in the HP category.

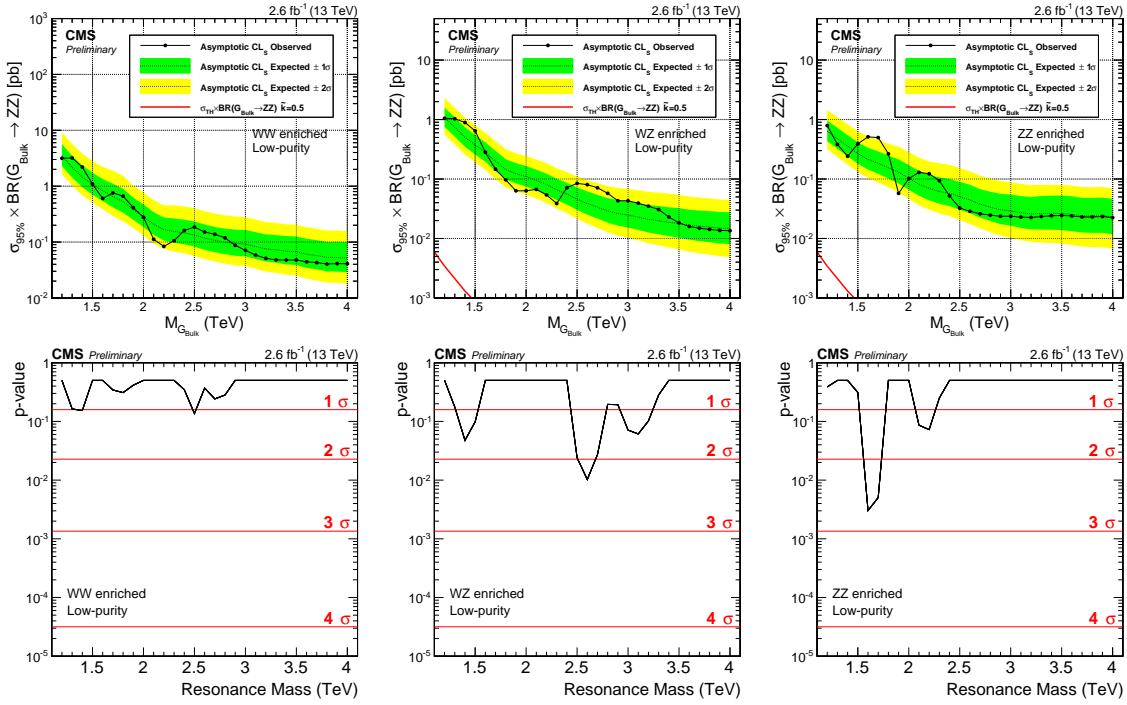


Figure A.6: Expected/observed limits and corresponding p-values obtained in the different mass categories. Here for a $G \rightarrow ZZ$ signal in the LP category.

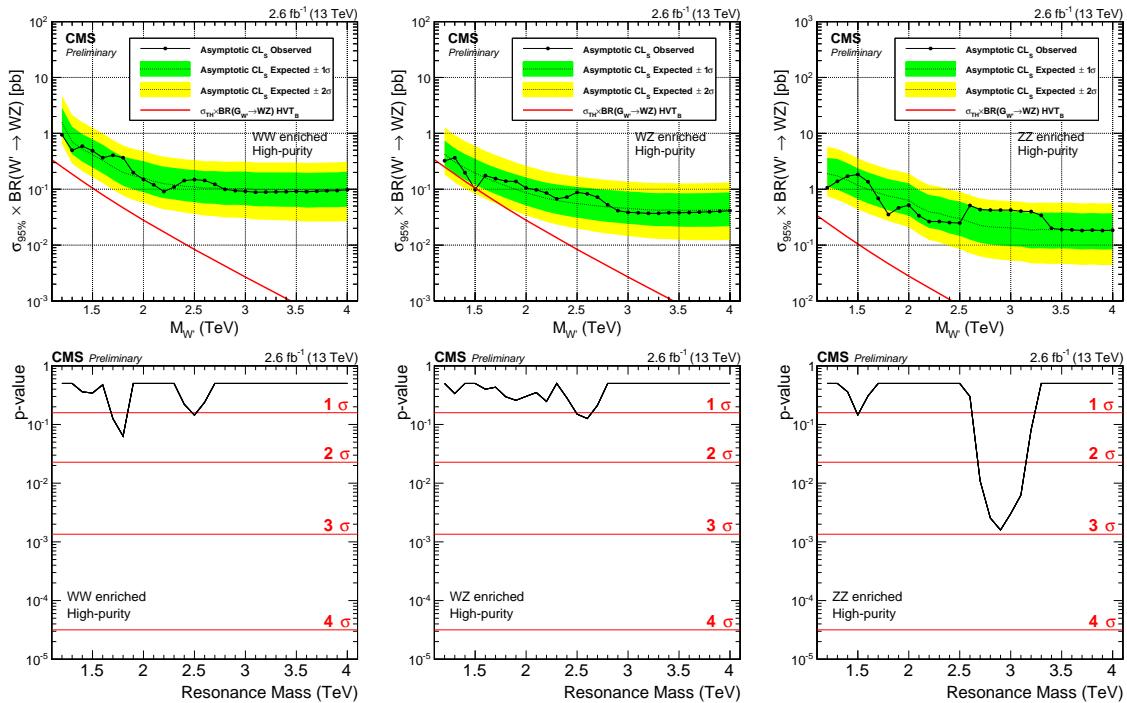


Figure A.7: Expected/observed limits and corresponding p-values obtained in the different mass categories. Here for a $W' \rightarrow WZ$ signal in the high-purity category.

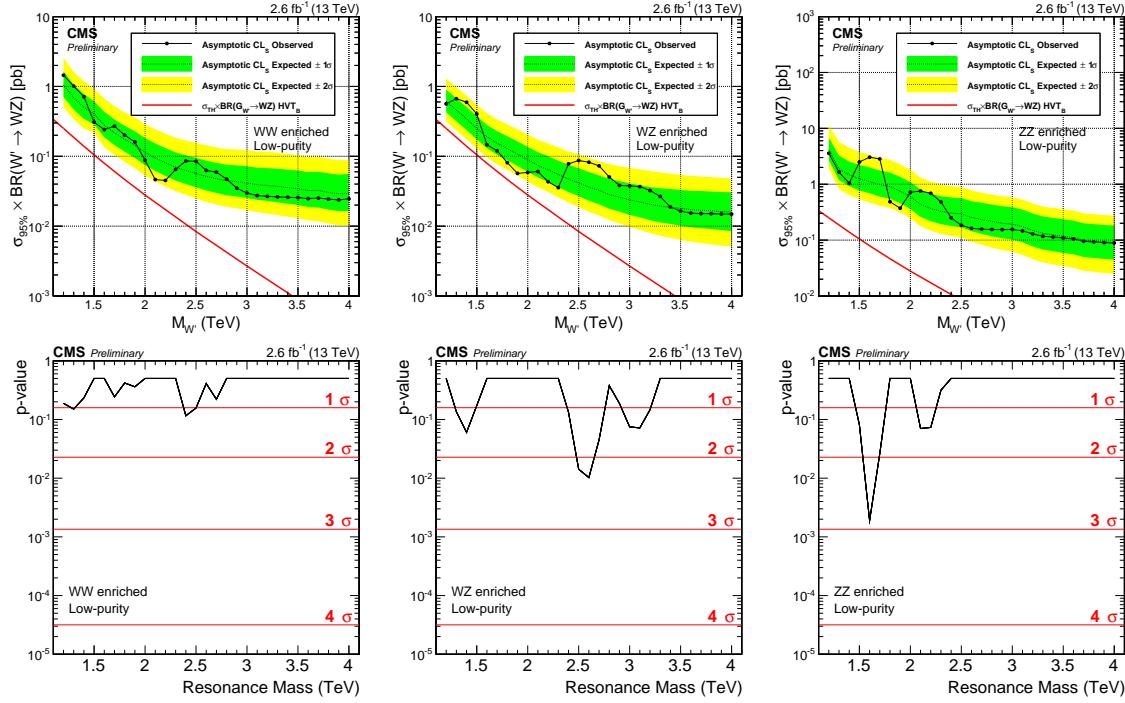


Figure A.8: Expected/observed limits and corresponding p-values obtained in the different mass categories. Here for a $W' \rightarrow WZ$ signal in the low purity category.

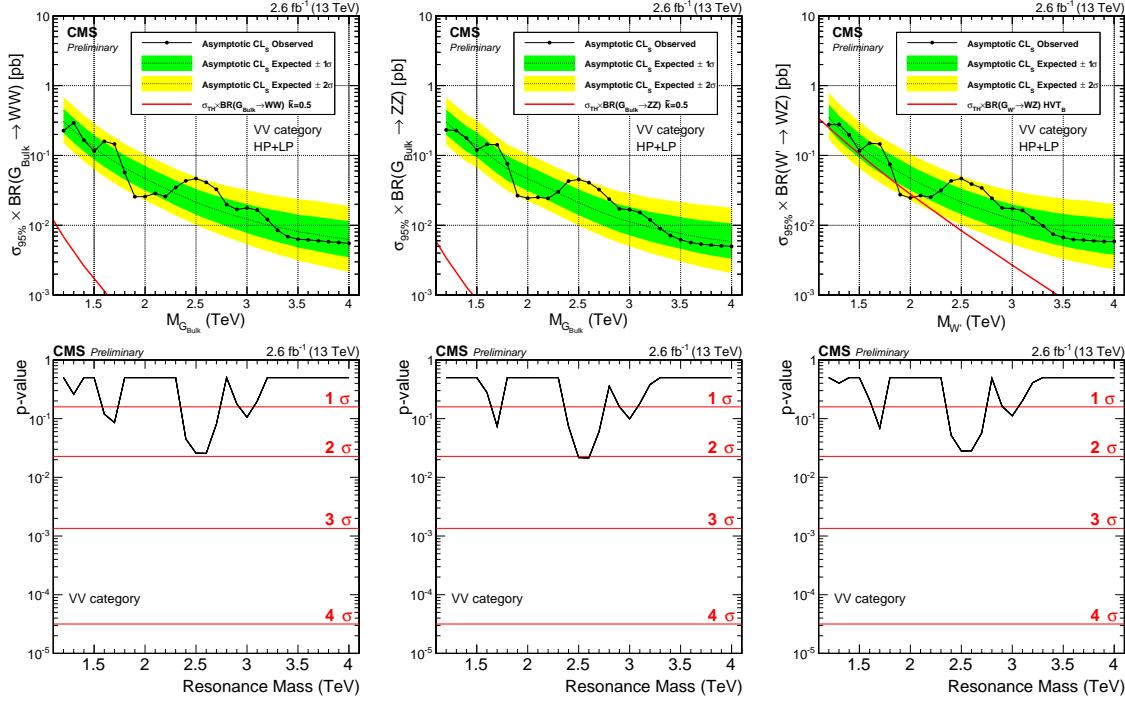


Figure A.9: Expected/observed limits and corresponding p-values obtained without splitting into mass categories. This analysis is performed as a cross check analysis and directly compares with the method used in the corresponding Run 1 analysis [?]. Here for a Bulk $G \rightarrow WW$ (left), $G \rightarrow ZZ$ (middle) and $W' \rightarrow WZ$ signal (right).

APPENDIX B

2016 W-tagging scale factor measurement

Figures B.1 show the $t\bar{t}$ real W (top) and non-W (bottom) PUPPI softdrop jet mass distributions for jets that passed (left) and failed (right column) the N-subjettiness selections $\text{PUPPI } \tau_{21} < 0.40$. Figures B.2 shows the fitted PUPPI softdrop jet mass distributions for the non-dominant backgrounds in the evaluation of the W-tagging scale factors. Here for jets that pass (top) and failed (bottom) the N-subjettiness selections $\text{PUPPI } \tau_{21} < 0.40$.

B.1 Efficiency scale factors for 12.9 and 35.9 fb^{-1}

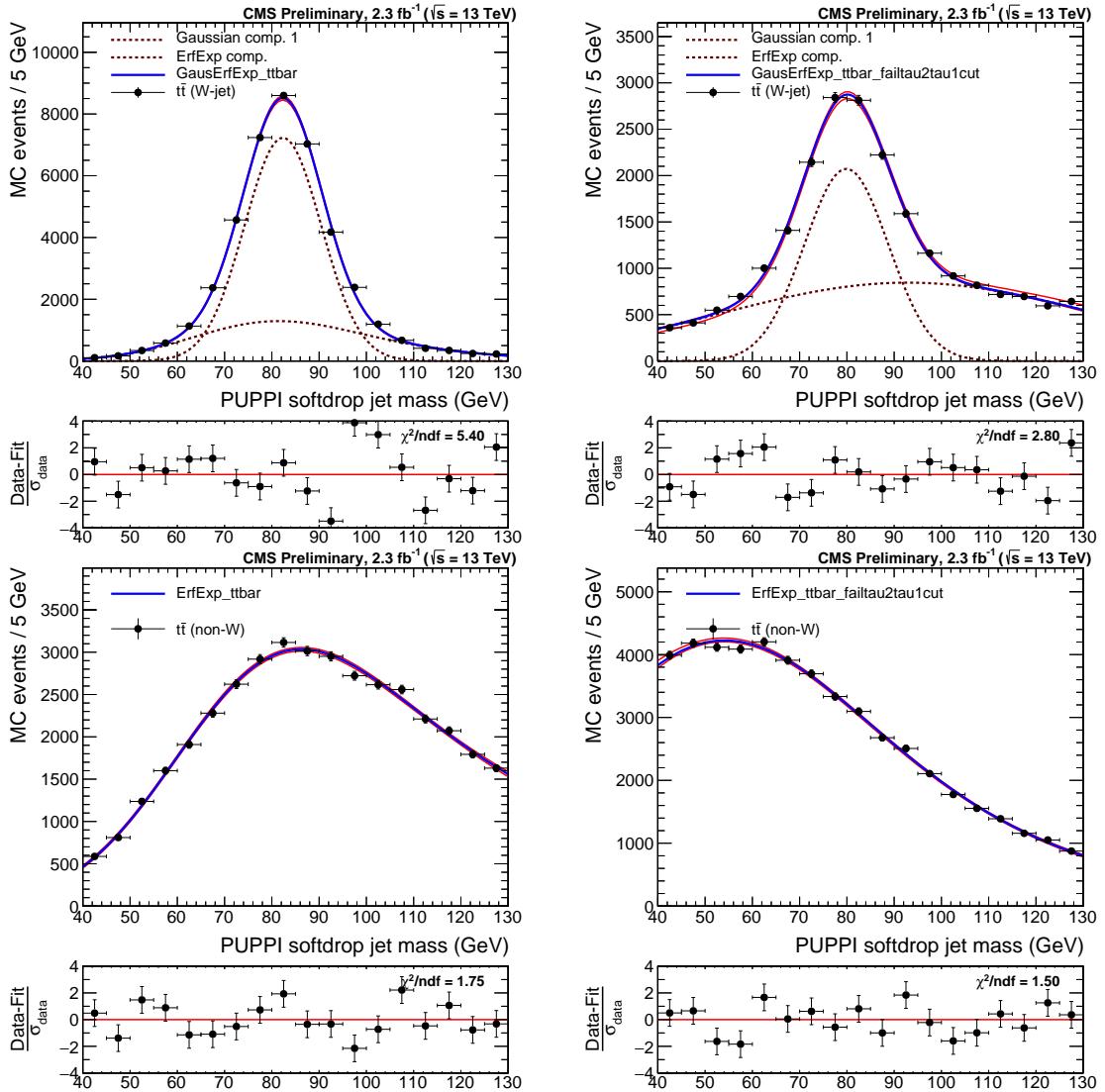


Figure B.1: Fit to the real W (top) and non-W (bottom) softdrop jet mass distribution for jets that pass (left) and fail (right) the cut on $\text{PUPPI } \tau_{21} < 0.4$.

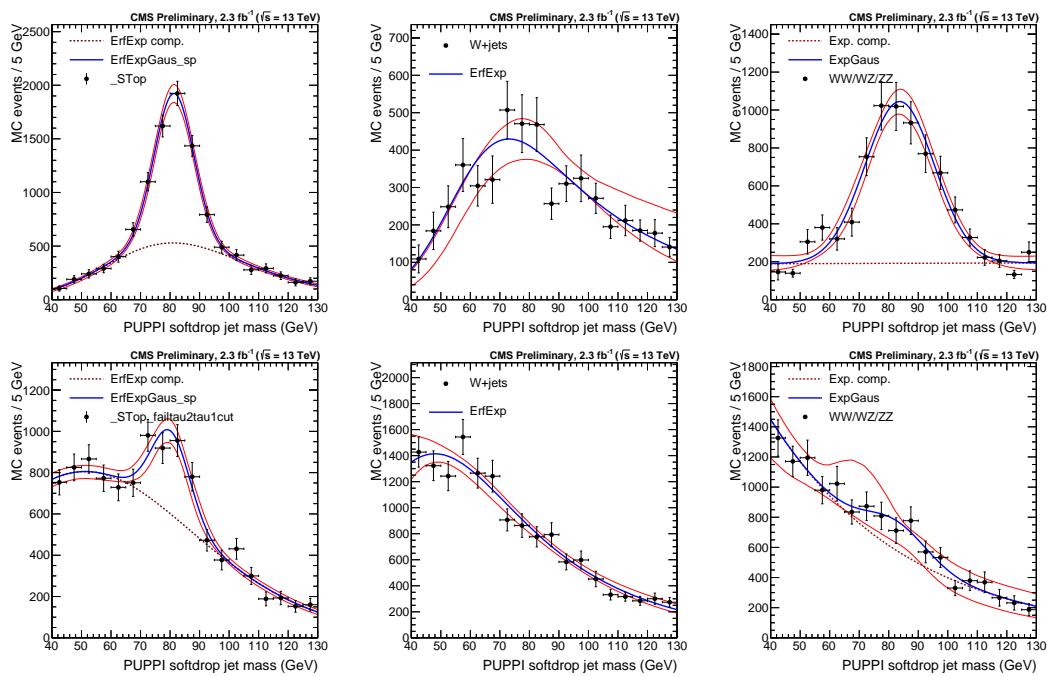


Figure B.2: Fits to the PUPPI softdrop jet mass spectrum for the non-dominant backgrounds (Single top, W+jets and VV respectively) in the pass (top) and fail (bottom) regions.

APPENDIX C

2015 cross-check analysis

APPENDIX D

Background fit checks for 2016 analysis

The background from QCD multijet events is modelled by a smoothly falling distribution in each analysis category. The method consists of a smoothness test of the observed data where the background is assumed to be described by the following empirical probability density function:

$$\frac{dN}{dm} = \frac{P_0(1 - m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2}} \quad (\text{D.1})$$

where m is the dijet invariant mass, \sqrt{s} the centre of mass energy, P_0 is a normalisation parameter for the probability density function and P_1 and P_2 describe the shape. To ensure that this function is sufficient to describe the data in all the different analysis categories, we first perform a test to check that no additional parameters are needed and to check the systematics due to choice of fit function. For these studies we use a data sideband, where one of the two jets is required to have a mass between $20 \text{ GeV} < M_{\text{Softdrop}} < 65 \text{ GeV}$. In order to quantify how many parameters are necessary, a Fishers F-test [47] is performed for the fits to data in the data sideband. The critical value that the test statistic must exceed is chosen to be $\alpha > 10\%$. If the returned Confidence Level is larger than α , the simpler fit is preferred. The three parameter fit is compared with the following 2, 4 and 5 parameter functions:

$$\frac{dN}{dm} = \frac{P_0}{(m/\sqrt{s})^{P_2}} \quad (\text{D.2})$$

$$\frac{dN}{dm} = \frac{P_0(1 - m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2 + P_3 \times \log(m/\sqrt{s})}} \quad (\text{D.3})$$

$$\frac{dN}{dm} = \frac{P_0(1 - m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2 + P_3 \times \log(m/\sqrt{s}) + P_4 \times \log(m/\sqrt{s})^2}} \quad (\text{D.4})$$

Additionally, fits with an alternative fit function has also been performed (for the single-tag categories we try both 4 and 5 parameter versions):

$$\frac{dN}{dm} = \frac{P_0(1 - m/\sqrt{s} + P_3(m/\sqrt{s})^2)^{P_1}}{(m/\sqrt{s})^{P_2}} \quad (\text{D.5})$$

$$\frac{dN}{dm} = \frac{P_0(1 - m/\sqrt{s} + P_3(m/\sqrt{s})^2)^{P_1}}{(m/\sqrt{s})^{P_2 + P_4 \times \log(m/\sqrt{s})}} \quad (\text{D.6})$$

D.0.1 Background fit checks in data sideband

We perform a test in a data sideband to make sure the fit functions work on real data and to exercise the estimation of number of necessary fit parameters via an F-test. The sideband is constructed by requiring one of the two jets two have a mass in the low softdrop jet mass sideband, between $20 \text{ GeV} < M_{\text{Softdrop}} < 65 \text{ GeV}$, while the full W/Z-tag selections are applied to the other jet. The low-mass jet is also required to pass the τ_{21} cut corresponding to the given category. For the single-tag category, the sideband is constructed by requiring one of the two jets to have a mass in the low softdrop jet mass sideband, between $20 \text{ GeV} < M_{\text{Softdrop}} < 65 \text{ GeV}$ and the other in a high-mass sideband, between $105 \text{ GeV} < M_{\text{Softdrop}} < 200 \text{ GeV}$. One of the jets is also required to pass the τ_{21} cut corresponding to the given category. We first check whether the sideband can be used to exercise the F test by checking whether or not there are features introduced in the dijet mass spectrum that may be hard to cover with the fit using QCD MC. To do so we look at the dijet invariant mass spectrum in the signal region divided by the distribution in the sideband. The obtained distributions are shown in Figure D.1 for three different generators, where the pure Pythia8 QCD samples (top left) has the highest statistics. The distributions are mostly smooth, but we do see features introduced in the ‘WW HP’ and ‘ZZ HP’ categories which might prove difficult to fit, as well as in the tail of the single-tag categories. These kinks shift around depending on what MC generator is used and do not seem to be a systematic feature caused by the cuts that have been applied, but rather due to limited statistics. We proceed with exercising the F-test in a data sideband, with the caveat that there might be features introduced in the spectrum where statistics are low.

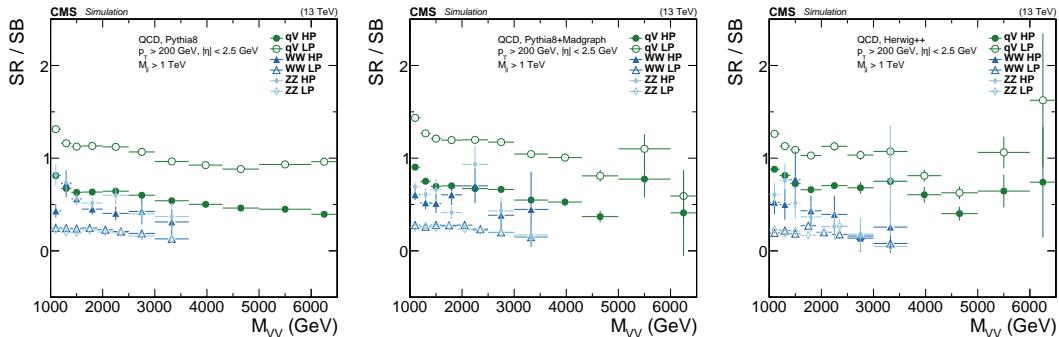


Figure D.1: Dijet mass spectrum in the signal region divided by the dijet mass spectrum in the sidebands using QCD Pythia8 (left), QCD Pythia8+Madgraph (middle) and QCD Herwig++ (right) simulated samples. Here for the double W/Z-tag and the single V-tagged HP and LP categories. Some jumps are observed in the high-mass tail of the dijet invariant mass distribution in the high-purity WW/ZZ categories, but otherwise no strange features seem to be induced by using the sideband.

Figure D.2 shows the fit to data in the data sideband for the WW and ZZ mass categories, both in the HP and in the LP n-subjettiness categories. The corresponding residuals, χ^2 and F-test results are shown i Table D.5 through D.8. For the double-tag categories, a two or three parameter function is sufficient to describe the data and we conclude that the function as defined in D.1 is sufficient for all mass categories. For the single-tag category a five parameter fit seems to be required in order to describe the data and the fit quality is not optimal. To ensure that the fit functions with sufficient number of parameters is able to describe the shape in the single-tag categories, we have additionally looked at the fit quality in QCD MC (see below). Here, the default dijet function seems sufficient to describe the distributions. The sideband in the single-tag categories in QCD MC do not show the same features as the data sideband as shown in Figure D.3.

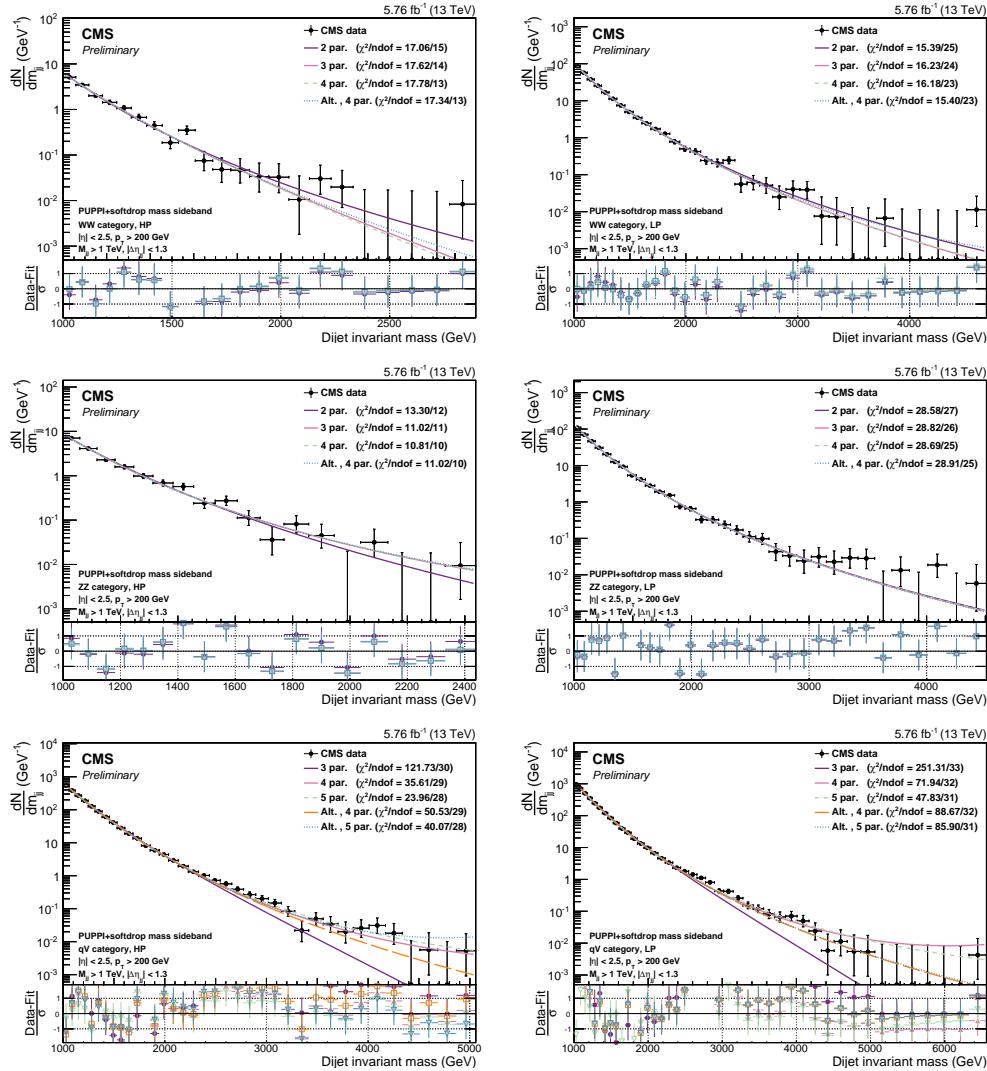


Figure D.2: Fitted dijet mass spectrum in the different mass and purity categories in a data sideband: WW high-purity (top left) and low-purity (top right), ZZ high-purity (middle left) and low-purity (middle right), qV high-purity (bottom left) and low-purity (bottom right).

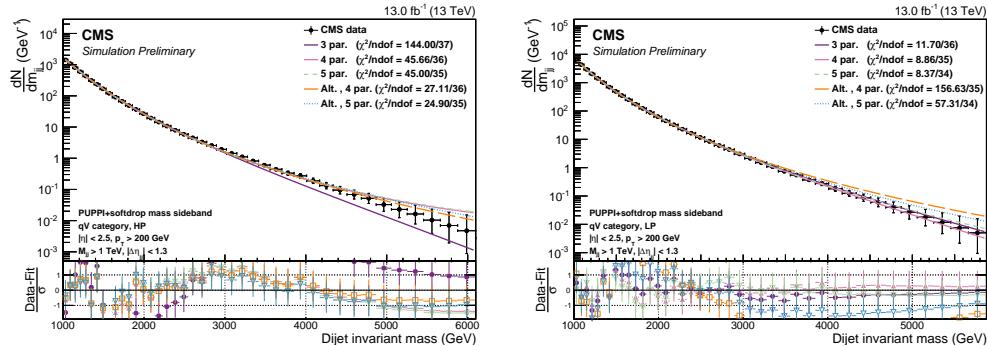


Figure D.3: Fitted dijet mass spectrum in the QCD MC sideband: qV high-purity (left) and low-purity (right).

WW category, HP			
Function	Residuals	χ^2	ndof
2 par	0.129	17.060	15
3 par	0.111	17.623	14
4 par	0.111	17.783	13
Fishers23	2.430	CL	0.140
Fishers34	0.012	CL	0.914

Table D.1: Residuals, χ^2 , and degrees of freedom for the WW category, HP category. A 2 parameter fit is needed to describe these data.

WW category, LP			
Function	Residuals	χ^2	ndof
2 par	0.908	15.388	25
3 par	0.279	16.225	24
4 par	0.263	16.178	23
Fishers23	56.395	CL	0.000
Fishers34	1.406	CL	0.247

Table D.2: Residuals, χ^2 , and degrees of freedom for the WW category, LP category. A 3 parameter fit is needed to describe these data.

ZZ category, HP			
Function	Residuals	χ^2	ndof
2 par	0.215	13.296	12
3 par	0.133	11.022	11
4 par	0.119	10.810	10
Fishers23	7.465	CL	0.018
Fishers34	1.304	CL	0.278

Table D.3: Residuals, χ^2 , and degrees of freedom for the ZZ category, HP category. A 3 parameter fit is needed to describe these data.

D.0.2 Background fit checks in QCD MC

As an additional check, we look at the fit functions in the different signal categories using QCD MC. This is shown in Figure D.4 for the double and Figure D.5 for the single tag categories. Here all fit functions and their pull distributions are plotted. The fits are performed in a mass range corresponding to the expected distribution in the different categories for 13 fb^{-1} of data. We have adapted the error bars to correspond to the maximum of the expected Poisson error for 13 fb^{-1} of data and the pure simulation error (accounting for the different weights assigned to the p_T -binned QCD MC sample). The reason for this choice is to get an estimate of whether the set of fit functions we plan to use to fit the background distribution in data, and plan to use in order to understand the systematic uncertainty due to our choice of fit function, are appropriate and do not produce fake bumps/kinks. As this distribution is the pure MC simulation curve, whose variation at high masses is much smaller than the expected poisson error for 13 fb^{-1} of data, we expect the χ^2/ndof to be lower than one. In order to protect against the fact that the MC simulation at lower dijet masses does not have

ZZ category, LP			
Function	Residuals	χ^2	ndof
2 par	2.459	28.583	27
3 par	2.363	28.817	26
4 par	2.175	28.694	25
Fishers23	1.107	CL	0.302
Fishers34	2.244	CL	0.146

Table D.4: Residuals, χ^2 , and degrees of freedom for the ZZ category, LP category. A 2 parameter fit is needed to describe these data.

qV category, HP			
Function	Residuals	χ^2	ndof
3 par	128.276	121.731	30
4 par	29.113	35.606	29
5 par	7.036	23.962	28
Alt. 4 par	37.232	50.528	29
Alt. 5 par	30.948	40.068	28
Fishers34	102.185	CL	0.000
Fishers45	90.988	CL	0.000
FishersAlt4Alt5	5.888	CL	0.022

Table D.5: Residuals, χ^2 , and degrees of freedom for the qV category, HP category. A 5 parameter fit is needed to describe these data.

more statistics than the expected data for 13 fb^{-1} of data, we use the largest of the Poisson and the MC error. The resulting errors are therefore a mixture of Poisson and the MC error and the χ^2/ndof for the QCD MC fits should not be considered. Fit quality in the form of χ^2/ndof should only be estimated from the fits to data sideband where pure Poisson errors are used. Overall the fits to QCD MC in the different categories describe the data well, with a two or three parameter function sufficient to describe the distributions. However, due to an under fluctuation of the first bin in the ‘ZZHP’ category, the higher parameter fits are steered by the first bin leading to discrepancies in the tail. This is the lowest statistics category and the danger for underfluctuations does exist. We have investigated the dijet invariant mass distribution down to a dijet invariant mass threshold of 800 GeV to make sure we are not

qV category, LP			
Function	Residuals	χ^2	ndof
3 par	671.341	251.311	33
4 par	171.593	71.942	32
5 par	80.801	47.830	31
Alt. 4 par	215.431	88.666	32
Alt. 5 par	214.766	85.896	31
Fishers34	96.109	CL	0.000
Fishers45	35.957	CL	0.000
FishersAlt4Alt5	0.099	CL	0.755

Table D.6: Residuals, χ^2 , and degrees of freedom for the qV category, LP category. A 5 parameter fit is needed to describe these data.

seeing a turn-on effect. This is shown in Figure D.6. The fits in the data sideband for the same category (Figure D.2, middle left) do not show the same trend.

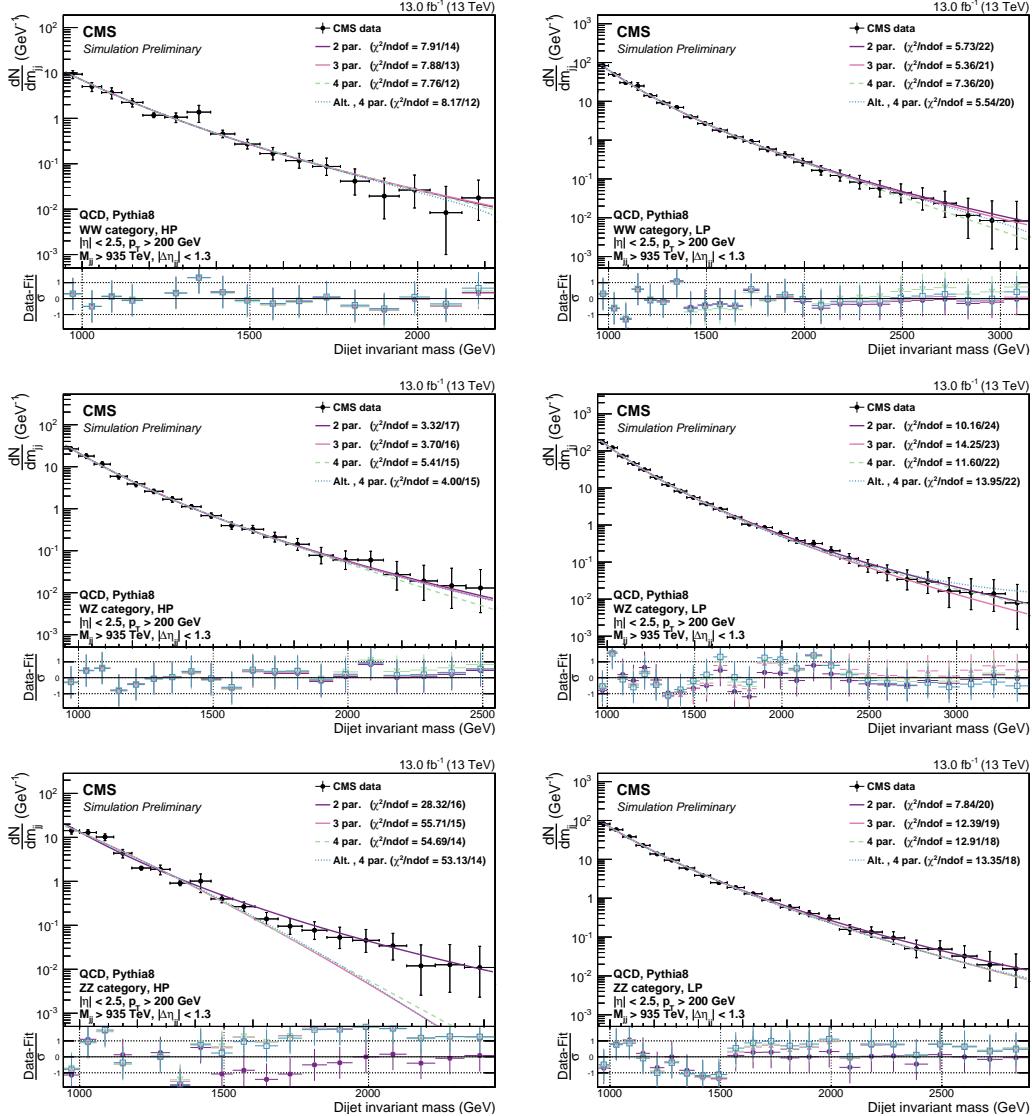


Figure D.4: Background fit for the M_{jj} distribution in QCD MC corresponding to 13 fb^{-1} . Here for the high- and low-purity double W/Z-tag category for the three different mass categories: WW category (top), WZ category (middle) and ZZ category (bottom).

D.0.3 F-test in signal region

The final F-test performed in order to define the number of fit parameters to be used to fit the background in each analysis category, is performed in the signal region. The resulting fits and F-test values for the double tag categories are shown in Figure D.7 and Tables D.7 to D.12.

The F-test results for the single-tag category are listed in Tables D.13 to D.16. Here, a three parameter fit is sufficient for all categories except the ‘high-purity’ qW category where a 5 parameter fit is preferred.

WW category, HP			
Function	Residuals	χ^2	ndof
2 par	0.251	17.673	16
3 par	0.187	14.863	15
4 par	0.183	14.618	14
Fishers23	5.454	CL	0.033
Fishers34	0.391	CL	0.541

Table D.7: Residuals, χ^2 , and degrees of freedom for the WW category, HP category. A 3 parameter fit is needed to describe these data.

WW category, LP			
Function	Residuals	χ^2	ndof
2 par	2.974	13.997	23
3 par	3.082	14.775	22
4 par	3.080	14.768	21
Fishers23	-0.805	CL	1.000
Fishers34	0.015	CL	0.905

Table D.8: Residuals, χ^2 , and degrees of freedom for the WW category, LP category. A 2 parameter fit is needed to describe these data.

WZ category, HP			
Function	Residuals	χ^2	ndof
2 par	2.333	17.562	17
3 par	2.158	16.952	16
4 par	2.114	16.842	15
Fishers23	1.372	CL	0.258
Fishers34	0.338	CL	0.569

Table D.9: Residuals, χ^2 , and degrees of freedom for the WZ category, HP category. A 2 parameter fit is needed to describe these data.

WZ category, LP			
Function	Residuals	χ^2	ndof
2 par	12.301	21.368	25
3 par	6.827	20.715	24
4 par	6.521	20.419	23
Fishers23	20.046	CL	0.000
Fishers34	1.126	CL	0.299

Table D.10: Residuals, χ^2 , and degrees of freedom for the WZ category, LP category. A 3 parameter fit is needed to describe these data.

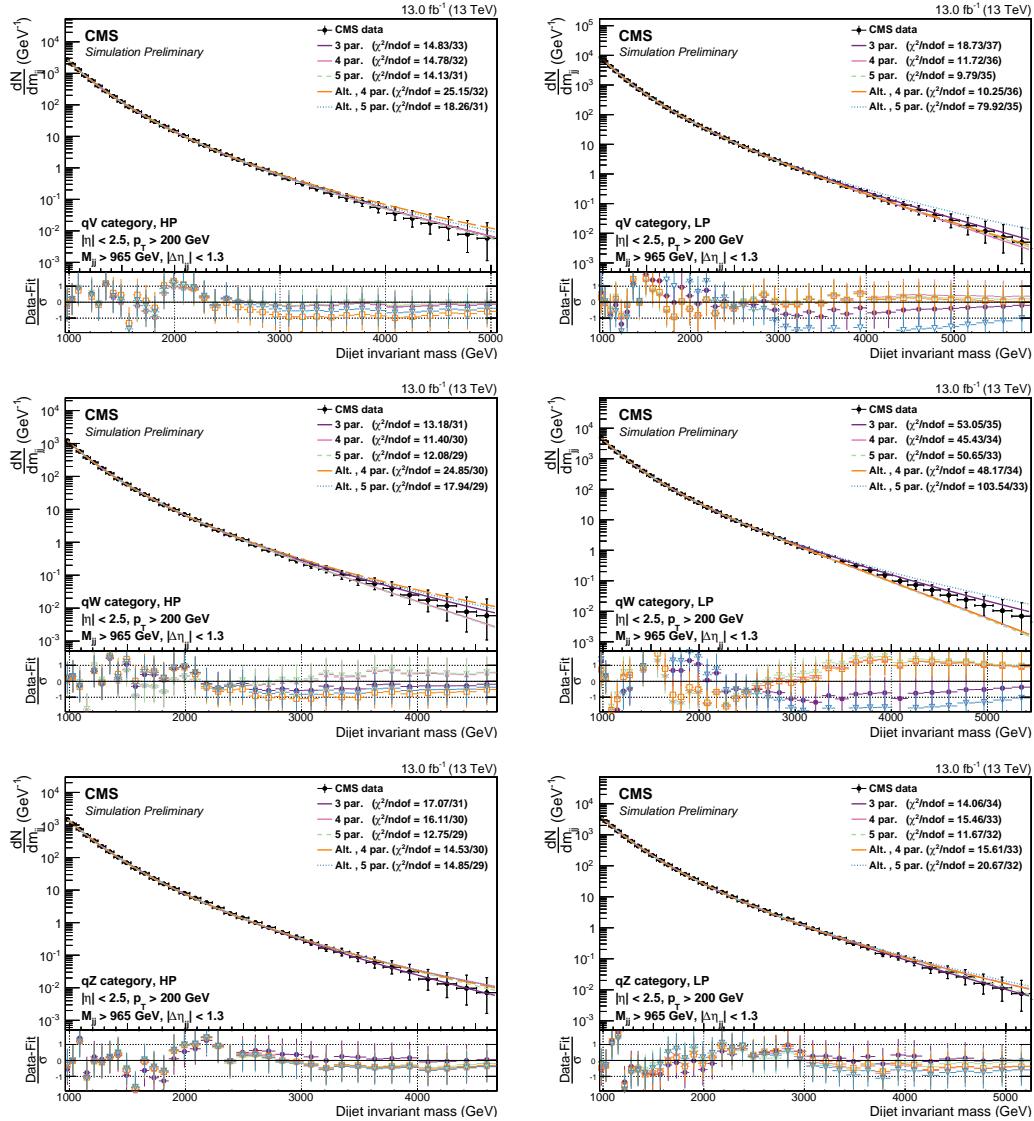


Figure D.5: Background fit for the M_{jj} distribution in QCD MC corresponding to 13 fb^{-1} . Here for the high- and low-purity single W/Z-tag category for the two different mass categories: wW category (top) and qZ category (bottom).

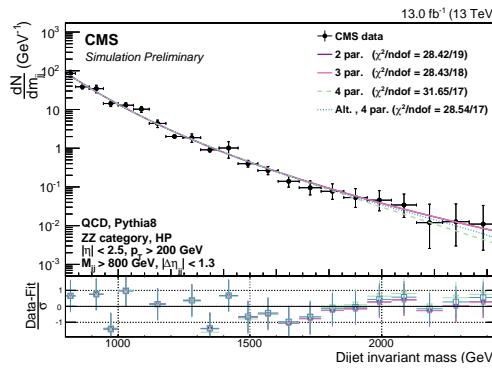


Figure D.6: Background fit for the M_{jj} distribution in QCD MC corresponding to 13 fb^{-1} . Here for the high-purity double Z-tag category using a dijet invariant mass threshold of 800 GeV. No turn-on effect at low invariant masses is observed.

ZZ category, HP			
Function	Residuals	χ^2	ndof
2 par	0.634	17.919	17
3 par	0.662	17.400	16
4 par	0.716	17.096	15
Fishers23	-0.720	CL	1.000
Fishers34	-1.197	CL	1.000

Table D.11: Residuals, χ^2 , and degrees of freedom for the ZZ category, HP category. A 2 parameter fit is needed to describe these data.

ZZ category, LP			
Function	Residuals	χ^2	ndof
2 par	9.293	19.452	22
3 par	6.884	20.118	21
4 par	6.598	20.076	20
Fishers23	7.701	CL	0.011
Fishers34	0.909	CL	0.351

Table D.12: Residuals, χ^2 , and degrees of freedom for the ZZ category, LP category. A 3 parameter fit is needed to describe these data.

qW category, HP			
Function	Residuals	χ^2	ndof
3 par	69.757	30.375	30
4 par	59.677	28.318	29
5 par	25.298	21.815	28
Alt. 4 par	35.610	22.810	29
Alt. 5 par	25.634	22.687	28
Fishers34	5.067	CL	0.032
Fishers45	39.409	CL	0.000
FishersAlt4Alt5	11.285	CL	0.002

Table D.13: Residuals, χ^2 , and degrees of freedom for the qW category, HP category. A 5 parameter fit is needed to describe these data.

qW category, LP			
Function	Residuals	χ^2	ndof
3 par	153.869	38.713	35
4 par	156.715	38.586	34
5 par	201.767	38.167	33
Alt. 4 par	189.434	39.327	34
Alt. 5 par	192.782	39.170	33
Fishers34	-0.636	CL	1.000
Fishers45	-7.592	CL	1.000
FishersAlt4Alt5	-0.590	CL	1.000

Table D.14: Residuals, χ^2 , and degrees of freedom for the qW category, LP category. A 3 parameter fit is needed to describe these data.

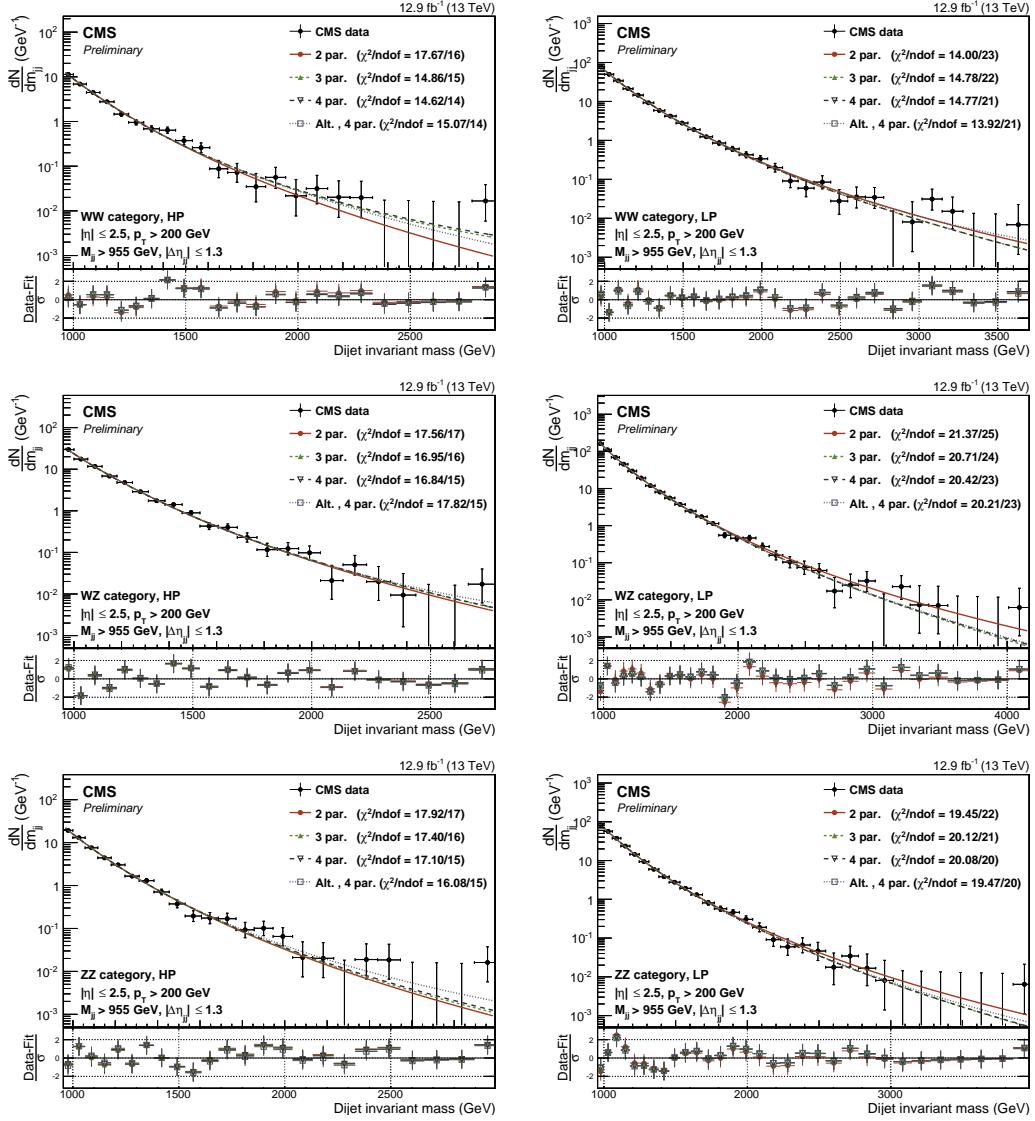


Figure D.7: Background fit for the M_{jj} distribution in the data signal region. Here for the high-(left) and low-purity (right) double W/Z-tag categories WW (top), WZ (middle) and ZZ (bottom).

qZ category, HP			
Function	Residuals	χ^2	ndof
3 par	12.963	21.252	30
4 par	12.961	21.252	29
5 par	9.256	19.644	28
Alt. 4 par	13.931	20.977	29
Alt. 5 par	9.739	20.344	28
Fishers34	0.004	CL	0.948
Fishers45	11.609	CL	0.002
FishersAlt4Alt5	12.484	CL	0.001

Table D.15: Residuals, χ^2 , and degrees of freedom for the qZ category, HP category. A 3 parameter fit is needed to describe these data.

qZ category, LP			
Function	Residuals	χ^2	ndof
3 par	369.554	47.426	36
4 par	369.554	47.426	35
5 par	298.358	46.525	34
Alt. 4 par	379.111	47.531	35
Alt. 5 par	379.120	47.531	34
Fishers34	0.000	CL	0.994
Fishers45	8.352	CL	0.007
FishersAlt4Alt5	-0.001	CL	0.000

Table D.16: Residuals, χ^2 , and degrees of freedom for the qZ category, LP category. A 3 parameter fit is needed to describe these data.

APPENDIX E

Search III: Multidimensional fit

E.1 Signal fits

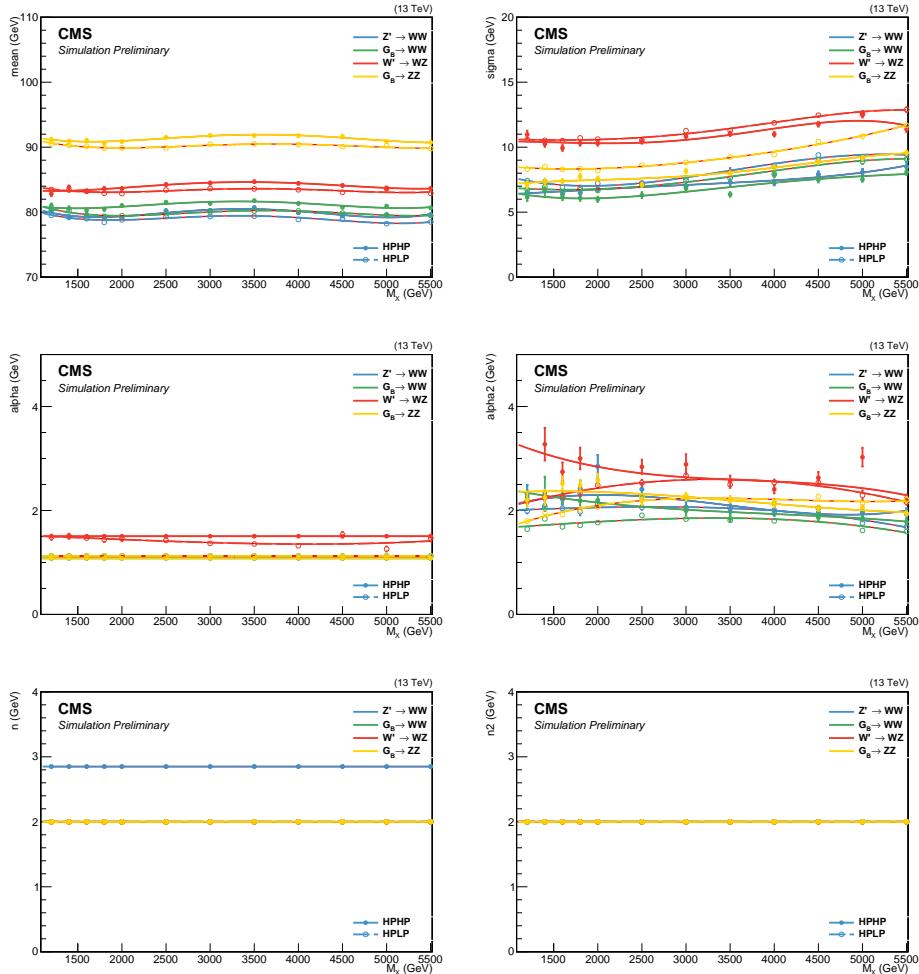


Figure E.1: The interpolated double Crystal-ball parameters for the softdrop jet mass as a function of M_X . To improve the stability of the fit some parameters are set constant. Here for jet 2.

E.2 2016 kernels

E.3 Resonant background

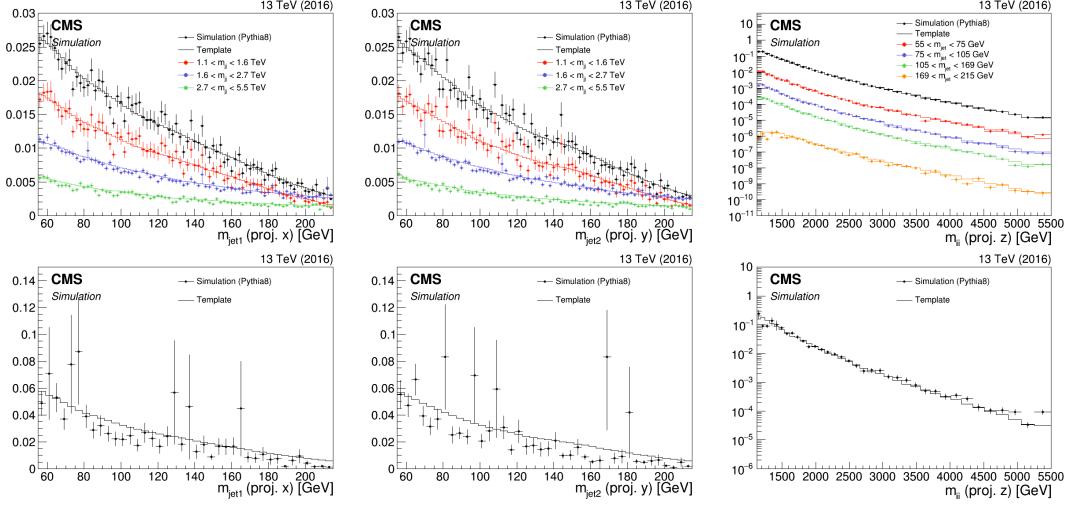


Figure E.2: Comparison between QCD MC simulation (markers) and kernels derived from generator level quantities (lines) for the HPHP category (top) and the HPLP category (bottom). The kernels are shown for $M_{\text{jet}1}$ (left), $M_{\text{jet}2}$ (middle) and M_{VV} (right).

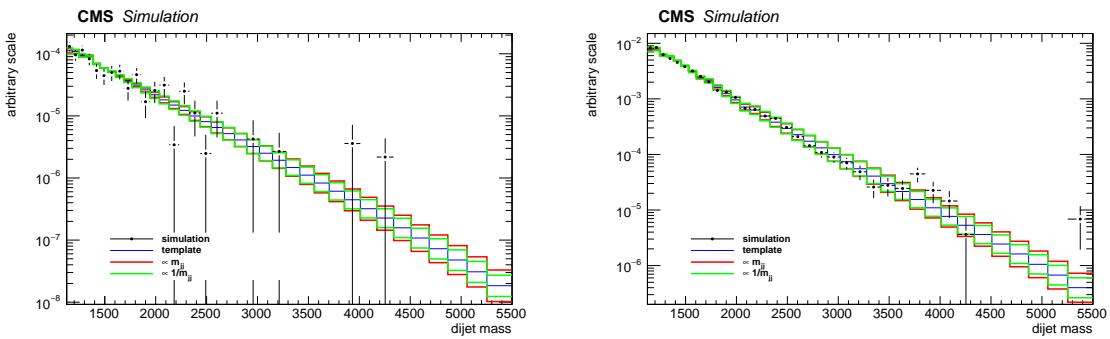


Figure E.3: One dimensional M_{VV} kernels (solid line) compared to MC (markers) for the Z+jets background in the HPHP (left) and HPLP (right) categories. The nominal shape derived from the smoothing procedure can be seen as blue line, alternative shapes derived from varying the slope of the M_{VV} spectrum are shown in green and red.

Bibliography

- [1] CERN, “Accelerators and Schedules”, 2018 (accessed November 12, 2018),
<https://beams.web.cern.ch/content/accelerators-schedules>.
- [2] CMS, “CMS Detector Design”, 2018 (accessed November 21, 2018),
<https://cms.cern/news/cms-detector-design>.
- [3] T. Lenzi, “Development and Study of Different Muon Track Reconstruction Algorithms for the Level-1 Trigger for the CMS Muon Upgrade with GEM Detectors”, Master’s thesis, U. Brussels (main), 2013.
- [4] A. Dominguez et al., “CMS Technical Design Report for the Pixel Detector Upgrade”, Technical Report CERN-LHCC-2012-016. CMS-TDR-11, CERN, Sep, 2012.
- [5] CMS Collaboration, “The CMS Experiment at the CERN LHC”, *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [6] P. Adzic et al., “Energy resolution of the barrel of the CMS electromagnetic calorimeter”, *JINST* **2** (2007) P04004, doi:10.1088/1748-0221/2/04/P04004.
- [7] C. Biino, “The CMS Electromagnetic Calorimeter: overview, lessons learned during Run 1 and future projections”, *Journal of Physics: Conference Series* **587** (2015), no. 1, 012001.
- [8] L. V. . EP, “Assembly of the 7th wedge. Assemblage du 7eme module”, (Aug, 2000). CMS Collection.
- [9] S. Sharma, “Understanding the performance of CMS calorimeter”, *Pramana* **69** (Dec, 2007) 1069–1074, doi:10.1007/s12043-007-0229-8.
- [10] on behalf of the CMS Collaboration Collaboration, “CMS reconstruction improvement for the muon tracking by the RPC chambers”, *PoS RPC2012* (Sep, 2012) 045. 9 p. Presented by Minsuk Kim at the XI workshop on Resistive Plate Chambers and Related Detectors - RPC2012, INFN Laboratori Nazionali di Frascati Italy, February 5-10, 2012.
- [11] CMS, “CMS Luminosity - public results”, 2018 (accessed December 11, 2018),
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [12] P. Billoir, “Progressive track recognition with a Kalman-like fitting procedure”, *Computer Physics Communications* **57** (1989), no. 1, 390 – 394, doi:[https://doi.org/10.1016/0010-4655\(89\)90249-X](https://doi.org/10.1016/0010-4655(89)90249-X).
- [13] A. S. et. Al., “Particle-flow reconstruction and global event description with the CMS detector”, *Journal of Instrumentation* **12** (2017), no. 10, P10003.

- [14] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, “Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC”, *Journal of Physics G: Nuclear and Particle Physics* **31** (2005), no. 9, N9.
- [15] T. C. collaboration, “Performance of CMS muon reconstruction in pp collision events at $s = 7$ TeV”, *Journal of Instrumentation* **7** (2012), no. 10, P10002.
- [16] D. Bertolini, P. Harris, M. Low, and N. Tran, “Pileup per particle identification”, *Journal of High Energy Physics* **2014** (2014), no. 10, doi:10.1007/JHEP10(2014)059.
- [17] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, “Better jet clustering algorithms”, *JHEP* **08** (1997) 001, doi:10.1088/1126-6708/1997/08/001, arXiv:hep-ph/9707323.
- [18] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions”, *Phys. Rev.* **D48** (1993) 3160–3166, doi:10.1103/PhysRevD.48.3160, arXiv:hep-ph/9305266.
- [19] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [20] J. POG, “Jet Identification: Recommendations for 13TeV data analysis”, 2018 (accessed December 18, 2018), <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID>.
- [21] CMS Collaboration, “Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS”, *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
- [22] D. Krohn, J. Thaler, and L.-T. Wang, “Jet Trimming”, *JHEP* **02** (2010) 084, doi:10.1007/JHEP02(2010)084, arXiv:0912.1342.
- [23] CMS Collaboration Collaboration, “Pileup Removal Algorithms”, Technical Report CMS-PAS-JME-14-001, CERN, Geneva, 2014.
- [24] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, “Towards an understanding of jet substructure”, *JHEP* **09** (2013) 029, doi:10.1007/JHEP09(2013)029, arXiv:1307.0007.
- [25] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, “Jet substructure as a new Higgs search channel at the LHC”, *Phys. Rev. Lett.* **100** (2008) 242001, doi:10.1103/PhysRevLett.100.242001, arXiv:0802.2470.
- [26] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, “Soft Drop”, *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.
- [27] J. Thaler and K. Van Tilburg, “Identifying Boosted Objects with N-subjettiness”, *JHEP* **03** (2011) 015, doi:10.1007/JHEP03(2011)015, arXiv:1011.2268.
- [28] J. Thaler and K. Van Tilburg, “Maximizing boosted top identification by minimizing N-subjettiness”, *Journal of High Energy Physics* **2012** (Feb, 2012) 93, doi:10.1007/JHEP02(2012)093.
- [29] S. Bolognesi et al., “Spin and parity of a single-produced resonance at the LHC”, *Phys. Rev. D* **86** (Nov, 2012) 095031, doi:10.1103/PhysRevD.86.095031.

- [30] CMS Collaboration, “Identification techniques for highly boosted W bosons that decay into hadrons”, *JHEP* **12** (2014) 017, doi:10.1007/JHEP12(2014)017, arXiv:1410.4227.
- [31] M. Bahr et al., “Herwig++ Physics and Manual”, *Eur. Phys. J. C* **58** (2008) 639–707, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.
- [32] T. Sjöstrand et al., “An Introduction to PYTHIA 8.2”, *Comput. Phys. Commun.* **191** (2015) 159–177, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.
- [33] The ATLAS collaboration, “Search for high-mass diboson resonances with boson-tagged jets in proton-proton collisions at $s=8$ TeV with the ATLAS detector”, *Journal of High Energy Physics* **2015** (Dec, 2015) 1–39, doi:10.1007/JHEP12(2015)055.
- [34] S. Frixione and B. R. Webber, “Matching NLO QCD computations and parton shower simulations”, *JHEP* **06** (2002) 029, doi:10.1088/1126-6708/2002/06/029, arXiv:hep-ph/0204244.
- [35] S. Frixione, P. Nason, and B. R. Webber, “Matching NLO QCD and parton showers in heavy flavor production”, *JHEP* **08** (2003) 007, doi:10.1088/1126-6708/2003/08/007, arXiv:hep-ph/0305252.
- [36] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [37] S. A. et. al., “Geant4—a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303, doi:[https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [38] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements”, *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [39] CMS Collaboration Collaboration, “Search for massive resonances in dijet systems containing jets tagged as W or Z boson decays in pp collisions at $\sqrt{s}=8$ TeV”, *JHEP* **08** (May, 2014) 173. 37 p. Comments: Replaced with published version. Added journal reference and DOI.
- [40] F. Dias et al., “Combination of Run-1 Exotic Searches in Diboson Final States at the LHC”, *JHEP* **04** (2016) 155, doi:10.1007/JHEP04(2016)155, arXiv:1512.03371.
- [41] CERN, “Collider Reach”, 2019 (accessed January 12, 2019), <http://collider-reach.web.cern.ch/>.
- [42] 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”, *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [43] T. Sjöstrand, S. Mrenna, and P. Z. Skands, “A brief introduction to PYTHIA 8.1”, *Comput. Phys. Commun.* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

- [44] G. Punzi, “Sensitivity of searches for new signals and its optimization”, *eConf C030908* (2003) MODT002, arXiv:physics/0308063. [,79(2003)].
- [45] CMS Collaboration, “Search for heavy resonances in the W/Z-tagged dijet mass spectrum in pp collisions at 7 TeV”, *Phys. Lett. B* **723** (2013) 280–301, doi:10.1016/j.physletb.2013.05.040, arXiv:1212.1910.
- [46] CMS Collaboration Collaboration, “Search for Narrow Resonances using the Dijet Mass Spectrum with 19.6fb-1 of pp Collisions at $\text{sqrt}s=8$ TeV”, Technical Report CMS-PAS-EXO-12-059, CERN, Geneva, 2013.
- [47] D. J. Hand, “Statistical Concepts: A Second Course, Fourth Edition by Richard G. Lomax, Debbie L. Hahs-Vaughn”, *International Statistical Review* **80** (2012), no. 3, 491–491.
- [48] T. C. collaboration, “Identification of b-quark jets with the CMS experiment”, *Journal of Instrumentation* **8** (2013), no. 04, P04013.
- [49] T. C. collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *Journal of Instrumentation* **13** (2018), no. 05, P05011.
- [50] CMS Collaboration Collaboration, “Search for massive resonances decaying into WW, WZ, ZZ, qW and qZ in the dijet final state at $\sqrt{s} = 13$ TeV using 2016 data”, Technical Report CMS-PAS-B2G-16-021, CERN, Geneva, 2016.
- [51] M. Dasgupta, A. Powling, and A. Siodmok, “On jet substructure methods for signal jets”, *JHEP* **08** (2015) 079, doi:10.1007/JHEP08(2015)079, arXiv:1503.01088.
- [52] U. Baur, I. Hinchliffe, and D. Zeppenfeld, “Excited quark production at hadron colliders”, *Int. J. Mod. Phys. A* **2** (1987) 1285, doi:10.1142/S0217751X87000661.
- [53] U. Baur, M. Spira, and P. M. Zerwas, “Excited-quark and -lepton production at hadron colliders”, *Phys. Rev. D* **42** (1990) 815, doi:10.1103/PhysRevD.42.815.
- [54] CMS Collaboration Collaboration, “Search for massive resonances decaying into WW, WZ, ZZ, qW, and qZ with dijet final states at $\sqrt{s} = 13$ TeV”, *Phys. Rev. D* **97** (Apr, 2018) 072006, doi:10.1103/PhysRevD.97.072006.
- [55] CMS Collaboration Collaboration, “Jet algorithms performance in 13 TeV data”, Technical Report CMS-PAS-JME-16-003, CERN, Geneva, 2017.
- [56] J. Alwall et al., “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions”, *Eur. Phys. J. C* **53** (2008) 473–500, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.
- [57] J. Dolen et al., “Thinking outside the ROCs: Designing Decorrelated Taggers (DDT) for jet substructure”, arXiv:1603.00027.
- [58] J. Gallicchio and M. D. Schwartz, “Quark and Gluon Tagging at the LHC”, *Phys. Rev. Lett.* **107** (2011) 172001, doi:10.1103/PhysRevLett.107.172001, arXiv:1106.3076.
- [59] J. A. Aguilar-Saavedra, “Stealth bosons and where to find them”, 2018, <https://indico.cern.ch/event/649482/contributions/2993323/attachments/1688557/2716149/SBAWTFD.pdf>.

- [60] J. A. Aguilar-Saavedra, “Running bumps from stealth bosons”, *Eur. Phys. J.* **C78** (2018), no. 3, 206, doi:10.1140/epjc/s10052-018-5717-0, arXiv:1801.08129.
- [61] ATLAS Collaboration Collaboration, “Search for diboson resonances in hadronic final states in 79.8 fb^{-1} of pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector”, Technical Report ATLAS-CONF-2018-016, CERN, Geneva, Jun, 2018.
- [62] T. Heimel, G. Kasieczka, T. Plehn, and J. M. Thompson, “QCD or What?”, arXiv:1808.08979.
- [63] J. A. Aguilar-Saavedra, J. H. Collins, and R. K. Mishra, “A generic anti-QCD jet tagger”, *JHEP* **11** (2017) 163, doi:10.1007/JHEP11(2017)163, arXiv:1709.01087.