

Sede Amministrativa: Università degli studi di Padova

Dipartimento di Fisica e Astronomia

Corso di Dottorato di Ricerca in Fisica
XXX ciclo

**Search for heavy resonances decaying
into a Z boson and a vector boson in
the $\nu\bar{\nu} q\bar{q}$ final state at CMS**

Coordinatore: Ch.mo Prof. Gianguido Dall'Agata

Supervisore: Ch.mo Prof. Franco Simonetto

Dottoranda: Lisa Benato

"I have no special talent. I am only passionately curious."

(A. Einstein)

Contents

1	Introduction	1
2	Theoretical motivation	3
2.1	Beyond standard model theories	3
2.2	Heavy Vector Triplet	5
2.2.1	Simplified Lagrangian	5
2.2.2	Mass eigenstates, mixing parameters and decay widths	6
2.2.2.1	Decay widths into fermions	8
2.2.2.2	Decay widths into bosons	9
2.2.2.3	Decays into fermions and bosons: concluding remarks	9
2.2.3	Benchmark model A: weak coupling scenario	10
2.2.4	Benchmark model B: strong coupling scenario	10
2.2.5	HVT production	11
2.2.6	Search for HVT resonances at LHC	13
2.3	Warped extra-dimension	17
2.3.1	Randall-Sundrum original model (RS1)	17
2.3.2	Bulk extension of RS1: graviton production and decays	20
2.3.3	Search for KK bulk gravitons at LHC	22
3	The Large Hadron Collider and the CMS experiment	25
3.1	The Large Hadron Collider	25
3.1.1	Proton-proton interactions	28
3.2	The CMS detector	29
3.2.1	The coordinate system	30
3.2.2	The magnet	32
3.2.3	The tracking system	32
3.2.3.1	The pixel detector	32
3.2.3.2	The strip detector	33
3.2.4	The electromagnetic calorimeter	33
3.2.5	The hadronic calorimeter	34
3.2.6	The muon system	35
3.2.6.1	The Drift Tubes	35
3.2.6.2	The Cathode Strip Chambers	35

3.2.6.3	The Resistive Plate Chambers	36
3.2.7	The trigger system and data acquisition	36
3.2.7.1	The Level-1 trigger	37
3.2.7.2	The High Level Trigger	37
3.2.7.3	Data acquisition, computing and storage	37
3.2.8	Particle Flow event reconstruction	37
3.2.9	Physics objects	38
3.2.9.1	Track reconstruction	38
3.2.9.2	Vertices reconstruction	38
3.2.9.3	Electrons and photons reconstruction	39
3.2.9.4	Muon reconstruction	40
3.2.9.5	Jet reconstruction	40
3.2.9.6	Tau reconstruction	44
3.2.9.7	b-jets tagging	45
3.2.9.8	Missing transverse energy reconstruction	45
3.3	ATLAS, ALICE, LHCb detectors	48
3.3.1	ATLAS	48
3.3.2	ALICE	49
3.3.3	LHCb	49
4	Search for diboson resonances in the $VZ \rightarrow q\bar{q} \nu\bar{\nu}$ final state	51
4.1	Analysis overview	51
4.2	Data and Monte Carlo simulations samples	52
4.2.1	Signal samples	52
4.2.2	Signal characterization	53
4.2.3	Background samples	60
4.2.4	Vector boson momentum corrections	62
4.2.5	Data samples	63
4.2.6	Trigger	63
4.3	Event selection	67
4.3.1	Vertex and Pile-up	67
4.3.2	Electrons	67
4.3.3	Photons	68
4.3.4	Muons	69
4.3.5	Taus	69
4.3.6	Jets	69
4.3.7	Jet mass	71
4.3.8	Jet substructure	75
4.3.8.1	Corrections induced by jet substructure variables	78
4.3.9	b-tagging	79
4.3.10	Missing Energy	79
4.3.11	Diboson candidate reconstruction	79
4.3.11.1	$V \rightarrow q\bar{q}$ reconstruction	79
4.3.11.2	$Z \rightarrow \nu\bar{\nu}$ reconstruction	80
4.3.11.3	Composite VZ candidate reconstruction	80
4.3.12	Final analysis selections	80
4.3.12.1	Z candidate selections	81
4.3.12.2	V candidate selections	81

CONTENTS

4.3.12.3	Topology and event cleaning	81
4.4	Data and simulations comparison	84
4.5	Background estimation technique	95
4.5.1	Background normalization	95
4.5.2	Background shape	98
4.5.3	Alpha method validations	100
4.5.4	Signal modeling	107
4.5.4.1	Signal parametrization	107
4.6	Systematic uncertainties	112
4.6.1	Uncertainties affecting the data-driven main background estimation	112
4.6.1.1	Normalization	112
4.6.1.2	Shape	112
4.6.2	Uncertainties affecting the signal and the sub-dominant backgrounds	112
4.6.2.1	Trigger uncertainty	112
4.6.2.2	Jet momentum uncertainties	112
4.6.2.3	Jet mass uncertainties	113
4.6.2.4	V-tagging uncertainties	113
4.6.2.5	b-tagging uncertainties	114
4.6.2.6	Missing Energy uncertainties	115
4.6.2.7	Pile-up uncertainty	116
4.6.2.8	QCD renormalization and factorization scale uncertainties	116
4.6.2.9	PDF	117
4.6.3	Summary	117
4.7	Results and interpretation	121
4.7.1	Statistical approach	121
4.7.1.1	The modified frequentist approach: asymptotic formulae to extract an upper limit on signal strength	121
4.7.1.2	Treatment of the systematic uncertainties	122
4.7.1.3	Computation of local p-values	123
4.7.2	Signal extraction strategy for the analysis	123
4.7.2.1	Fit diagnostics: nuisances pulls and impacts	123
4.7.2.2	Results: expected and observed limits	128
4.7.2.3	Results: local p-value scan	128
4.7.3	Interpretation of the results in the HVT model	133
5	Conclusions	135

CONTENTS

1

Abstract

2

3

4 This thesis presents a search for potential signals of new heavy resonances decaying into a pair of
5 vector bosons, with masses between 1 TeV and 4 TeV, predicted by beyond standard model the-
6 ories. The signals probed are spin-1 W' , predicted by the Heavy Vector Triplet model, and spin-2
7 bulk gravitons, predicted by warped extra-dimension models. The scrutinized data are produced by
8 LHC proton-proton collisions at a center-of-mass energy $\sqrt{s} = 13$ TeV during the 2016 operations,
9 and collected by the CMS experiment, corresponding to an integrated luminosity of 35.9 fb^{-1} . One
10 of the boson should be a Z , and it is identified through its invisible decay into neutrinos, while
11 the other electroweak boson, consisting either into a W or into a Z boson, is required to decay
12 hadronically into a pair of quarks. The decay products of heavy resonances are produced with large
13 Lorentz boosts; as a consequence, the decay products of the bosons (quarks and neutrinos) are
14 expected to be highly energetic and collimated. The couple of neutrinos, escaping undetected, is
15 reconstructed as missing momentum in the transverse plane of the CMS detector. The couple of
16 quarks is reconstructed as one large-cone jet, with high transverse momentum, recoiling against
17 the couple of neutrinos. Grooming algorithms are adopted in order to improve the jet mass reso-
18 lution, by removing soft radiation components and spectator events from the particles clustered as
19 the large-cone jet. The groomed jet mass is used to tag the hadronically decaying vector boson, to
20 define the signal region of the search (close to the nominal mass of the W and Z bosons, between
21 65–105 GeV) and a signal-depleted control region, that is used for the background estimation. An hy-
22 brid data-simulation approach predicts the normalization and the shape of the main background,
23 represented by a vector boson produced in association with jets, by taking advantage of the distri-
24 bution of data in the signal-depleted control regions. Secondary backgrounds are predicted from
25 simulations. Jet substructure techniques are exploited, in order to classify events into two exclusive
26 purity categories, by distinguishing the couple of quarks inside the large-cone jet. This approach
27 improves the background rejection and the discovery reach. The search is performed by scanning
28 the distribution of the reconstructed mass of the resonance, looking for a local excess in data with
29 regards to the prediction. Depending on the mass, upper limits on the cross-section of heavy spin-1
30 and spin-2 narrow resonances, multiplied by the branching fraction of the resonance decaying into
31 Z and a W boson for a spin-1 signal, and into a pair of Z bosons for spin-2, are set in the range
32 $0.9 - 63 \text{ fb}$ and in the range $0.5 - 40 \text{ fb}$ respectively. A W' hypothesis is excluded up to 3.11 TeV, in
33 the Heavy Vector Triplet benchmark A scenario, and up to 3.41 TeV, considering the benchmark B
34 scenario. A bulk graviton hypothesis, given the curvature parameter of the extra-dimension $\tilde{k} = 1.0$,
35 is excluded up to 1.14 TeV.

CONTENTS

Riassunto

39 Questa tesi presenta una ricerca di potenziali segnali di nuove risonanze pesanti, che decadono in
40 una coppia di bosoni vettori, con masse comprese tra 1 TeV e 4 TeV, predette da teorie oltre il mod-
41 ello standard. I segnali indagati sono W' di spin 1, predette dal modello Heavy Vector Triplet, e
42 gravitoni di spin 2, predetti da modelli che prevedono extra dimensioni ripiegate. I dati esaminati
43 sono prodotti dalle collisioni protone-protone di LHC ad un'energia del centro di massa di $\sqrt{s} = 13$
44 TeV durante le operazioni del 2016, e raccolti dall'esperimento CMS, per una luminosità integrata
45 di 35.9 fb^{-1} . Uno dei bosoni dev'essere una Z , che viene identificata dal suo decadimento invisibile
46 in neutrini, mentre l'altro bosone elettrodebole, sia una W che una Z , deve decadere nel canale
47 adronico in una coppia di quark. I prodotti di decadimento di risonanze pesanti sono generati con
48 significativi boost di Lorentz; di conseguenza, ci si aspetta che i prodotti di decadimento dei bosoni
49 (i quark e i neutrini) abbiano elevate energie e siano collimati. La coppia di neutrini, che sfugge alla
50 rivelazione, viene ricostruita come momento mancante nel piano trasverso del rivelatore CMS. La
51 coppia di quark viene ricostruita come un jet a largo cono, con elevato momento trasverso, che rin-
52 cula contro la coppia di neutrini. Algoritmi di grooming sono impiegati per migliorare la risoluzione
53 della massa del jet, rimuovendo la radiazione soffice e gli eventi spettatori dalle particelle clusteriz-
54 zate come jet a largo cono. La massa ripulita del jet viene utilizzata per identificare il bosone vettore
55 che decade in adroni, per definire la regione di segnale della ricerca (vicina alla massa nominale dei
56 bosoni W e Z , nell'intervallo 65-105 GeV) e una regione di controllo svuotata dal segnale, che viene
57 utilizzata per la stima dei fondi. Un approccio ibrido dati-simulazione predice la normalizzazione
58 e la forma del fondo principale, rappresentato da un bosone vettore prodotto in associazione con
59 jet, sfruttando la distribuzione dei dati nelle regioni di controllo svuotate dal segnale. I fondi sec-
60 ondari sono predetti completamente con le simulazioni. Tecniche di sottostruttura del jet sono
61 adoperate per classificare gli eventi in due categorie esclusive di purezza, distinguendo le coppie di
62 quark dentro al jet a largo cono. Questo approccio migliora la soppressione del fondo e la poten-
63 zialità di scoperta. La ricerca viene fatta scansionando la distribuzione della massa ricostruita della
64 risonanza, cercando un eccesso locale nei dati rispetto alle predizioni. In funzione della massa,
65 limiti superiori sulla sezione d'urto per risonanze pesanti e strette di spin 1 e spin 2, moltiplicate
66 per il rapporto di diramazione della risonanza che decade in Z e W per il segnale di spin 1, e in
67 una coppia di bosoni Z per lo spin 2, sono fissati nell'intervallo $0.9 - 63 \text{ fb}$ e nell'intervallo $0.5 -$
68 40 fb rispettivamente. Un'ipotesi di W' è esclusa fino ad una massa di 3.11 TeV, nello scenario A di
69 riferimento dell'Heavy Vector Triplet, e fino a 3.41 TeV, nello scenario B di riferimento. Un'ipotesi
70 di gravitone, dato il parametro di curvatura della dimensione addizionale $\tilde{k} = 1.0$, è esclusa fino ad
71 una massa di 1.14 TeV.

CONTENTS

Introduction

- 75 The discovery of the Higgs boson at the CERN Large Hadron Collider represents a milestone in the
 76 knowledge of the particle physics. The Higgs mechanism connects the theoretical formulation of
 77 the standard model of the particles to the current picture of the universe, as it is known: spin-1 weak
 78 bosons and standard model fermions are allowed to acquire masses, constituting the fundamental
 79 bricks of the known matter. Despite this successful achievement, some questions are still left unan-
 80 swered; in order to solve the open problems, a plethora of new beyond standard model theories has
 81 been built.
- 82 Many of these theories hypothesize the existence of larger symmetries in the universe, or new extra-
 83 dimensions, that will result into the appearance of new heavy particles, expected to have masses
 84 around the TeV scale. The Large Hadron Collider (LHC) is the ideal tool to investigate this unknown
 85 phase-space, given the fact that during the so-called LHC Run 2 era (started in 2015), the unprece-
 86 dented center-of-mass energy of 13 TeV has been reached in the proton-proton collisions.
- 87 The CMS experiment, located in the northern part of the LHC ring, is a multi-purpose detector, suit-
 88 able to study highly energetic new phenomena. Its intense magnetic field, its sharp segmentation,
 89 its hermeticity and the interplay of many sophisticated reconstruction algorithms allow to measure
 90 with a very high precision the trajectories, the momenta and the energy deposits left by energetic
 91 particles.
- 92 This thesis presents a search for signals of heavy resonances that decay into a pair of vector bosons.
 93 The search is performed by using the 2016 data produced by proton-proton collisions of the LHC,
 94 and collected by the CMS detector. One Z boson is identified through its invisible decay in neutrino-
 95 nos, while the other vector boson is required to decay hadronically into a pair of quarks. Given the
 96 fact that the searched resonances have masses around the TeV, their decay products are expected
 97 to be produced with large Lorentz boosts. This leads to a non-trivial identification of the couple of
 98 quarks or leptons, coming from the vector bosons decays. In fact, they are expected to lie very close
 99 in angle. Dedicated algorithms and substructure techniques allow to distinguish a pair of quarks
 100 originating from a vector boson from the background processes, initiated by the strong interaction.
 101 The search is performed by scanning the distribution of the reconstructed mass of the resonance,
 102 looking for a local excess in data with regards to the predictions. The background estimation is
 103 performed with an hybrid data-simulation approach, by using the distribution of data in signal-
 104 depleted control regions.

105 The thesis is organized as follows.

106 In chapter 2, an overview of the theoretical motivations is presented. Two beyond standard model
107 theories are considered: the Heavy Vector Triplet model and the bulk graviton model, a particu-
108 lar scenario included in warped extra-dimensions theories. In both cases, new heavy particles are
109 expected to decay into vector bosons with a sizeable rate.

110 In chapter 3, the CMS detector is briefly described, along with the physics objects exploited for the
111 purpose of this search.

112 Chapter 4 is dedicated to the analysis: after a general introduction (sec. 4.1), the features of the
113 data, signal and background samples used in the analysis are described in detail (sec. 4.2). Sec. 4.3
114 is dedicated to the selections applied, in order to reach the best signal-to-noise efficiency and to
115 properly build the resonance candidate. The very first data-simulation comparison is performed in
116 sec. 4.4. The background estimation technique, the final data-predicted background comparison
117 and the signal modelling are included in sec. 4.5. Systematic uncertainties are listed in sec. 4.6. The
118 final results, the statistical analysis and the physics interpretation are shown in sec. 4.7.

119 Chapter 5 summarizes the conclusions.

Theoretical motivation

121

122

123 The standard model (SM) of elementary particles represents, so far, the best available description
 124 of the fundamental particles and their interactions. It is the summation of two gauge theories: the
 125 electroweak interaction, that portrays the weak and electromagnetic interactions together, and the
 126 strong interaction, or quantum chromodynamics (QCD). Particles, namely quarks and leptons, are
 127 described as spin-1/2 fermions, whilst interactions are mediated by spin-1 bosons. The symmetry
 128 group of the standard model is:

$$SU_C(3) \times SU_L(2) \times U_Y(1), \quad (2.1)$$

129 where the first factor is related to strong interaction, whose mediators are eight gluons, while $SU_L(2) \times$
 130 $U_Y(1)$ is the electroweak symmetry group, whose mediators are the photon and the Z - W^\pm bosons.
 131 In renormalizable theories, with no anomalies, all gauge bosons are expected to be massless, in
 132 contrast with our experimental knowledge. This inconsistency is solved by introducing a new scalar
 133 particle, the Higgs boson, that gives mass to weak bosons and fermions via the spontaneous sym-
 134 metry breaking mechanism.

135 In the last decades, the standard model has been accurately probed by many experimental facilities
 136 (LEP, Tevatron, LHC), and the results lead to an impressive agreement between theoretical predic-
 137 tions and experiments [1]. The discovery of the Higgs boson at the CERN Large Hadron Collider,
 138 measured by both the CMS and the ATLAS collaborations [2–8], represents not only an extraordinary
 139 confirmation of the model, but also the latest biggest achievement in particle physics as a whole.

141 2.1 Beyond standard model theories

142 Even though the SM is the most complete picture of the universe of the particles, many questions
 143 are still left open. From a phenomenological point of view, some experimental observations are not
 144 included in the theory:

- 145 • in SM, neutrinos are massless (whilst the well established observation of neutrino flavour os-
 146 cillation proves that neutrinos do carry mass);
- 147 • no candidates for dark matter are predicted;

- 148 • no one of the fields included in the SM can explain the cosmological inflation;

- 149 • the CP asymmetry embedded in the SM is not sufficient to explain the matter-antimatter
150 asymmetry in the universe.

151 From a purely theoretical perspective, some issues are still relevant in the formulation of the model:

- 152 • *Flavour problem.*

153 The standard model has 18 free parameters: 9 fermionic masses; 3 angular parameters in the
154 Cabibbo-Kobayashi-Maskawa matrix, plus 1 phase parameter; electromagnetic coupling α ;
155 strong coupling α_{strong} ; weak coupling α_{weak} ; Z mass; the mass of the Higgs boson. Such a
156 huge number of degrees of freedom marks the SM as weakly predictive in the flavour sector.

- 157 • *Unification.*

158 There is not a “complete” unification of strong, weak and electromagnetic interactions, since
159 each one has its own coupling constant, behaving differently at different energy scales; not to
160 mention the fact that gravitational interaction is completely excluded from the SM.

- 161 • *Hierarchy problem.*

162 From quantum field theory, it is known that perturbative corrections to the mass of the scalar
163 bosons included in the theory tend to make it increase towards the energy scale at which the
164 considered theory still holds [9]. If the standard model is seen as a low-mass approximation
165 of a more general theory valid up to the Planck mass scale (*i.e.*, $\sim 1.2 \times 10^{19}$ GeV), a fine-
166 tuning cancellation of the order of 1 over 10^{34} is needed to protect the Higgs boson mass at the
167 electroweak scale (~ 100 GeV). Such an astonishing correction is perceived as very unnatural.

168 Numerous beyond standard model theories (BSM) have been proposed in order to overcome the
169 limits of the SM.

170 Grand Unified Theories (GUT) aim at extending the symmetry group of the SM (eq. 2.1) into largest
171 candidates, such as $SO(10)$, $SU(5)$ and $E(6)$. At GUT scale, approximately at 10^{16} GeV, non-gravitational
172 interactions are expected to be ruled by only one coupling constant, α_{GUT} .

173 Super symmetric (SUSY) models [10] state that every fermion (boson) of the SM has a bosonic
174 (fermionic) superpartner, with exactly the same quantum numbers, except the spin. If SUSY is not
175 broken, each couple of partners and superpartners should have the same masses, hypothesis ex-
176 cluded by the non-observation of the s-electron. Super symmetry represents a very elegant solu-
177 tion of the hierarchy problem of the Higgs boson mass, since the perturbative corrections brought
178 by the new SUSY particles exactly cancel out the divergences caused by the SM particles corrections.

179 A particular sub-class of SUSY models, minimal super symmetric standard models [11–13], is char-
180 acterized by the introduction of a new symmetry, the R-parity, that guarantees the proton stability
181 and also the stability of the lightest SUSY particle, a possible good candidate for dark matter.

182
183 Two other possible theoretical pictures are extensively described in sec. 2.2-2.3.

184 2.2 Heavy Vector Triplet

185 The heavy vector triplet model [14] provides a general framework aimed at studying new physics
 186 beyond the standard model, that can manifest into the appearance of new resonances.

187 The adopted approach is that of the simplified model, in which an effective Lagrangian is intro-
 188 duced, in order to describe the properties and interactions of new particles (in this case, a triplet of
 189 spin-1 bosons) by using a limited set of parameters, that can be easily linked to the physical observ-
 190 ables at the LHC experiments. These parameters can describe many physical motivated theories
 191 (such as sequential extensions of the SM [15, 16] or composite Higgs [17, 18]), built to solve the hier-
 192 archy problem of the SM.

193 Since a simplified model is not a complete theory, its validity is restricted to the on-shell quanti-
 194 ties related to the production and decay mechanisms of the new resonances, that is how most of
 195 the LHC BSM searches are performed. Given these conditions, experimental results in the resonant
 196 region are sensitive to a limited number of the phenomenological Lagrangian parameters (or to a
 197 combination of those), whilst the remaining parameters tend to influence the tail of the distribu-
 198 tions.

199 Limits on production cross-section times branching fraction ($\sigma \mathcal{B}$), as a function of the invariant
 200 mass spectrum of the probed resonance, can be extracted from experimental data. Given that $\sigma \mathcal{B}$
 201 are functions of the simplified model parameters and of the parton luminosities, it is then possible
 202 to interpret the observed limits in the parameter space.

203 2.2.1 Simplified Lagrangian

204 The heavy vector triplet framework assumes the existence of an additional vector triplet, V_μ^a , $a =$
 205 1, 2, 3, in which two spin-1 particles are charged and one is neutral:

$$V_\mu^\pm = \frac{V_\mu^1 \mp i V_\mu^2}{\sqrt{2}}; \\ V_\mu^0 = V_\mu^3. \quad (2.2)$$

206 The triplet interactions are described by a simplified Lagrangian, that is invariant under SM gauge
 207 and CP symmetry, and accidentally invariant under the custodial symmetry $SU(2)_L \times SU(2)_R$:

$$\begin{aligned} \mathcal{L}_V = & -\frac{1}{4} \left(D_\mu V_\nu^a - D_\nu V_\mu^a \right) \left(D^\mu V^\nu{}^a - D^\nu V^\mu{}^a \right) + \frac{m_V^2}{2} V_\mu^a V^\mu{}^a \\ & + i g_V c_H V_\mu^a (H^\dagger \tau^a D^\mu H - D^\mu H^\dagger \tau^a H) + \frac{g^2}{g_V} c_F V_\mu^a \sum_f \bar{f}_L \gamma^\mu \tau^a f_L \\ & + \frac{g_V}{2} c_{VVV} \epsilon_{abc} V_\mu^a V_\nu^b (D^\mu V^\nu{}^c - D^\nu V^\mu{}^c) + g_V^2 c_{VHH} V_\mu^a V^\mu{}^a H^\dagger H - \frac{g}{2} c_{VWW} \epsilon_{abc} W^{\mu\nu}{}^a V_\mu^b V_\nu^c. \end{aligned} \quad (2.3)$$

208 In the first line of the formula 2.3, V mass and kinematic terms are included, described with the co-
 209 variant derivative $D_\mu V_\nu^a = \partial_\mu V_\nu^a + g \epsilon^{abc} W_\mu^b V_\nu^c$, where W_μ^a are the fields of the weak interaction and
 210 g is the weak gauge coupling. V_μ^a are not mass eigenstates, since they mix with the electroweak fields
 211 after the spontaneous symmetry breaking, therefore m_V isn't the physical mass of the V bosons.
 212 The second line describes the interaction of the triplet with the Higgs field and the SM left-handed
 213 fermions; c_H describes the vertices with the physical Higgs and the three unphysical Goldstone
 214 bosons that, for the Goldstone equivalence theorem, are equivalent to the longitudinal polarization

215 of W and Z bosons at high-energy; hence, c_H is related to the bosonic decays of the resonances. c_F
 216 is the analogous parameter describing the V interaction with fermions, that can be generalized as
 217 a flavour dependent coefficient, once defined $J_F^{\mu a} = \sum_f \bar{f}_L \gamma^\mu \tau^a f_L$:

$$c_F V_\mu^a J_F^{\mu a} = c_\ell V_\mu^a J_\ell^{\mu a} + c_q V_\mu^a J_q^{\mu a} + c_3 V_\mu^a J_3^{\mu a}. \quad (2.4)$$

218 The last part of the equation includes terms that are relevant only in strongly coupled scenarios (see
 219 sec. 2.2.2.2) through the V-W mixing, but it does not include vertices of V with light SM fields, hence
 220 it can be neglected while describing the majority of the LHC phenomenology, under the assump-
 221 tions previously stated. Additional dimension-four quadrilinear V interactions are non relevant for
 222 the processes discussed, otherwise their effects would be appreciated in electroweak precision tests
 223 and precise Higgs coupling measurements [19].

224
 225 The parameters in the Lagrangian can be interpreted as follows: g_V describes the strength of the
 226 interaction, that is weighted by c parameters. g_V ranges from $g_V \sim 1$ when the coupling is weak
 227 (sec. 2.2.3), to $g_V \sim 4\pi$ when the coupling is strong (sec. 2.2.4). c parameters are expected to be
 228 $c \sim 1$, except to c_H , that can be smaller for weak couplings. The combinations describing the ver-
 229 tices, $g_V c_H$ and $g^2 c_F / g_V$, can be considered as the fundamental parameters, used to interpret the
 230 experimental results.

231 2.2.2 Mass eigenstates, mixing parameters and decay widths

232 The newly introduced $SU(2)_L$ triplet is expected to mix with the weak SM fields. The $U(1)_{em}$ sym-
 233 metry is left unbroken by the new interaction, hence the massless combination of the electroweak
 234 fields, namely the photon, is the same as the SM:

$$A_\mu = B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W, \quad (2.5)$$

235 with the usual definitions of the electroweak parameters:

$$\begin{aligned} \tan \theta_W &= \frac{g'}{g}, \\ e &= \frac{gg'}{\sqrt{g^2 + g'^2}}, \\ g &= e / \sin \theta_w, \\ g' &= e / \cos \theta_w. \end{aligned} \quad (2.6)$$

236 The Z boson, on the other hand, mixes with the neutral component of the triplet, V^0 , with a rotation
 237 parametrized with the angle θ_N :

$$\begin{pmatrix} \cos \theta_N & \sin \theta_N \\ -\sin \theta_N & \cos \theta_N d \end{pmatrix} \begin{pmatrix} Z \\ V^0 \end{pmatrix}. \quad (2.7)$$

238 The mass matrix of the rotated system is given by:

$$\mathbb{M}_N^2 = \begin{pmatrix} \hat{m}_Z^2 & c_H \zeta \hat{m}_Z \hat{m}_V \\ c_H \zeta \hat{m}_Z \hat{m}_V & \hat{m}_V^2 \end{pmatrix}, \quad (2.8)$$

2.2 Heavy Vector Triplet

²³⁹ where the parameters are defined as:

$$\begin{cases} \hat{m}_Z = \frac{e}{2 \sin \theta_W \cos \theta_W} \hat{v} \\ \hat{m}_V^2 = m_V^2 + g_V^2 c_{VVHH} \hat{v}^2 \\ \zeta = \frac{g_V \hat{v}}{2 \hat{m}_V} \\ \frac{\hat{v}^2}{2} = \langle H^\dagger H \rangle \end{cases}, \quad (2.9)$$

²⁴⁰ and \hat{v} , the vacuum expectation value of the Higgs field, can be different from the SM $v = 246$ GeV.
²⁴¹ The physical masses of Z and V^0 , m_Z and M_0 , and θ_N come from the matrix relations:

$$\begin{aligned} \text{Tr}(\mathbb{M}_N^2) &= \hat{m}_Z^2 + \hat{m}_V^2 = m_Z^2 + M_0^2, \\ \|\mathbb{M}_N\| &= \hat{m}_Z^2 \hat{m}_V^2 (1 - c_H^2 \zeta^2) = m_Z^2 M_0^2, \\ \tan 2\theta_N &= \frac{2 c_H \zeta \hat{m}_Z \hat{m}_V}{\hat{m}_V^2 - \hat{m}_Z^2}. \end{aligned} \quad (2.10)$$

²⁴² The W^\pm bosons mix with the charged components of the triplet, V^\pm , leading to a mass matrix anal-
²⁴³ ogous to eq. 2.11:

$$\mathbb{M}_C^2 = \begin{pmatrix} \hat{m}_W^2 & c_H \zeta \hat{m}_W \hat{m}_V \\ c_H \zeta \hat{m}_W \hat{m}_V & \hat{m}_V^2 \end{pmatrix}, \quad (2.11)$$

²⁴⁴ where \hat{m}_W is defined as:

$$\left\{ \hat{m}_W = \frac{e}{2 \sin \theta_W} \hat{v} = \hat{m}_Z \cos \theta_W \right.; \quad (2.12)$$

²⁴⁵ the physical masses of W and V^\pm , m_W and M_\pm , and the angle θ_C parametrizing the rotation of the
²⁴⁶ charged sector are described by:

$$\begin{aligned} \text{Tr}(\mathbb{M}_C^2) &= \hat{m}_W^2 + \hat{m}_V^2 = m_W^2 + M_\pm^2, \\ \|\mathbb{M}_C\| &= \hat{m}_W^2 \hat{m}_V^2 (1 - c_H^2 \zeta^2) = m_W^2 M_\pm^2, \\ \tan 2\theta_C &= \frac{2 c_H \zeta \hat{m}_W \hat{m}_V}{\hat{m}_V^2 - \hat{m}_W^2}. \end{aligned} \quad (2.13)$$

²⁴⁷ The custodial symmetry of eq. 2.3 guarantees that:

$$\mathbb{M}_C^2 = \begin{pmatrix} \cos \theta_W & 0 \\ 0 & 1 \end{pmatrix} \mathbb{M}_N^2 \begin{pmatrix} \cos \theta_W & 0 \\ 0 & 1 \end{pmatrix}. \quad (2.14)$$

²⁴⁸ The determinant of these matrices allows to extract a custodial relation among the masses:

$$m_W^2 M_\pm^2 = \cos^2 \theta_W m_Z^2 M_0^2. \quad (2.15)$$

²⁴⁹ The HVT model predicts the existence of new particles at the TeV scale, but it has also to reproduce
²⁵⁰ the SM parameters up to the current experimental accuracy. The scale of the electroweak masses
²⁵¹ ($m_Z \sim m_W \sim 100$ GeV) can be preserved in the model, without fine-tuning cancellations, if there is
²⁵² a very natural hierarchy among $\hat{m}_{(W,Z)}$ and \hat{m}_V :

$$\frac{\hat{m}_{(W,Z)}}{\hat{m}_V} \sim \frac{m_{(W,Z)}}{M_{(\pm,0)}} \ll 1. \quad (2.16)$$

253 No constraints on the strength of the interaction are necessary to guarantee the natural hierarchy,
 254 hence the parameter $\zeta = g_V \hat{v} / 2\hat{m}_V$ can either be very small or close to unity (strong coupling). If
 255 the hierarchy applies, the second lines in eq. 2.10 and eq. 2.13 can be approximated as follows:

$$\begin{aligned} m_Z^2 &= \hat{m}_Z^2 (1 - c_H^2 \zeta^2) (1 + \mathcal{O}(\hat{m}_Z^2 / \hat{m}_V^2)) \\ m_W^2 &= \hat{m}_W^2 (1 - c_H^2 \zeta^2) (1 + \mathcal{O}(\hat{m}_W^2 / \hat{m}_V^2)). \end{aligned} \quad (2.17)$$

256 By definition (eq. 2.12), $\hat{m}_W = \hat{m}_Z \cos \theta_W$, hence the following relation holds to percent accuracy:

$$\frac{m_W^2}{m_Z^2} \approx \cos^2 \theta_W. \quad (2.18)$$

257 The SM tree-level relation, $\rho = \frac{m_W^2}{m_Z^2 \cos^2 \theta_W} = 1$, is then reproduced if $\cos^2 \theta_W$ is equivalent to the
 258 experimental measurement of the weak mixing angle, within 1% accuracy:

$$\cos^2 \theta_W \approx 1 - 0.23. \quad (2.19)$$

259 By combining the custodial relation with the mass hierarchy required to naturally reproduce the
 260 SM, another fundamental consequence can be derived, namely the mass degeneracy of the triplet
 261 (to percent accuracy):

$$M_{\pm}^2 = M_0^2 (1 + \mathcal{O}(\%)). \quad (2.20)$$

262 The degenerate mass will be called $M_V \approx M_{\pm} \approx M_0$; given 2.16, $M_V = \hat{m}_V$. The neutral and charged
 263 components will therefore have similar production cross-sections.
 264 Another implication of the mass hierarchy (2.16) is that the mixing angles $\theta_{(N,C)}$ between the elec-
 265 troweak fields and the triplet are small:

$$\theta_{(N,C)} \approx c_H \zeta \frac{\hat{m}_{(W,Z)}}{\hat{m}_V} \ll 1, \quad (2.21)$$

266 hence the couplings among SM particles are very close to the couplings predicted by the SM.

2.2.2.1 Decay widths into fermions

268 The couplings among the triplet and SM fermions are expressed as a function of the rotation angles
 269 $\theta_{(C,N)}$ and SM couplings (omitting the CKM matrix elements for quarks):

$$\begin{cases} g_L^N = \frac{g^2}{g_V} \frac{c_F}{\sqrt{2}} \cos \theta_N + (g_L^Z)_{SM} \sin \theta_N \approx \frac{g^2}{g_V} \frac{c_F}{\sqrt{2}}, \\ g_R^N = (g_R^Z)_{SM} \sin \theta_N \approx 0 \end{cases}, \quad (2.22)$$

$$\begin{cases} g_L^C = \frac{g^2}{g_V} \frac{c_F}{\sqrt{2}} \cos \theta_C + (g_L^W)_{SM} \sin \theta_N \approx \frac{g^2}{g_V} \frac{c_F}{\sqrt{2}}, \\ g_R^C = 0 \end{cases},$$

270 where $g_L^W = g/\sqrt{2}$; $g_{L,R}^{W,Z}$ are those predicted by the standard model. The V bosons interact with SM
 271 left fermions, and the strength of the couplings with fermions is determined by $g^2 c_F / g_V$, as stated
 272 in sec. 2.2.1. The decay width into fermions is then given by:

$$\Gamma_{V^{\pm} \rightarrow f \bar{f}} \approx 2\Gamma_{V^0 \rightarrow f \bar{f}} \approx N_c \left(\frac{g^2 c_F}{g_V} \right)^2 \frac{M_V}{48\pi}, \quad (2.23)$$

273 where N_c is the number of colours (3 for quarks, 1 for leptons).

2.2 Heavy Vector Triplet

274 2.2.2.2 Decay widths into bosons

275 As a starting point, a proper choice of the gauge makes the derivation of the approximate decay
 276 widths easier. While the unitary gauge is very convenient in discussing the electroweak symme-
 277 try breaking mechanism, since it provides a basis in which the Goldstone components of the scalar
 278 fields of the theory are set to zero, it does not properly describe the longitudinally polarized bosons in
 279 high-energy regimes, since it introduces a dependence of the type E/m in the longitudinal polarization
 280 vector, not corresponding to the experimental results [20, 21]. This pathological behaviour can
 281 be overcome profiting of the equivalence theorem: while calculating the scattering amplitude of
 282 an high-energy process, the longitudinally polarized vectors are equivalent to their corresponding
 283 Goldstone scalars. The scattering amplitude can therefore be calculated with Goldstone diagrams.
 284 In the so-called equivalent gauge [22], the Higgs doublet is then parametrized as:

$$H = \begin{pmatrix} i\pi_+ \\ \frac{\hat{v} + h - i\pi_0}{\sqrt{2}} \end{pmatrix}, \quad (2.24)$$

285 and the Goldstones π_0 and π_+ describe respectively W and Z longitudinal bosons; h is the physical
 286 Higgs boson. Rewriting the simplified Lagrangian 2.3 with 2.24 parametrization, two terms hold
 287 the information of the interaction of the V particles with the Goldstones:

$$\mathcal{L}_\pi = \dots + c_H \zeta \hat{m}_V V_\mu^a \partial^\mu \pi^a + \frac{g_V c_H}{2} V_\mu^a (\partial^\mu h \pi^a - h \partial^\mu \pi^a + \epsilon^{abc} \pi^b \partial^\mu \pi^c) + \dots, \quad (2.25)$$

288 that are ruled by the $c_H g_V$ parameters combination. When ζ parameter is $\zeta \approx 1$, the first term in
 289 eq. 2.25 becomes important, and it is absorbed by a redefinition of the V_μ^a and π^a fields,

$$\begin{aligned} V_\mu^a &\rightarrow V_\mu^a + \frac{c_H \zeta}{\hat{m}_V} \partial_\mu \pi^a, \\ \pi^a &\rightarrow \frac{1}{\sqrt{1 - c_H^2 \zeta^2}} \pi^a; \quad c_H^2 \zeta^2 < 1, \end{aligned} \quad (2.26)$$

290 By properly taking into account all the terms of the simplified lagrangian in the equivalent gauge,
 291 the partial widths of the dibosonic decays are ($\hat{m}_V = M_V$):

$$\begin{aligned} \Gamma_{V^0 \rightarrow W_L^+ W_L^-} &\approx \Gamma_{V^\pm \rightarrow W_L^\pm Z_L} \approx \frac{g_V^2 c_H^2 M_V}{192\pi} \frac{(1 + c_H c_{VVV} \zeta^2)^2}{(1 - c_H^2 \zeta^2)^2} = \frac{g_V^2 c_H^2 M_V}{192\pi} (1 + \mathcal{O}(\zeta^2)), \\ \Gamma_{V^0 \rightarrow Z_L h} &\approx \Gamma_{V^\pm \rightarrow W_L^\pm h} \approx \frac{g_V^2 c_H^2 M_V}{192\pi} \frac{(1 - 4c_H c_{VVV} \zeta^2)^2}{(1 - c_H^2 \zeta^2)} = \frac{g_V^2 c_H^2 M_V}{192\pi} (1 + \mathcal{O}(\zeta^2)). \end{aligned} \quad (2.27)$$

292 2.2.2.3 Decays into fermions and bosons: concluding remarks

293 From eq. 2.23-2.27, some important conclusions can be extracted.

- 294 • When the ζ parameter is small, all the triplet decays (both in fermions and in dibosons),
 295 branching fractions and productions are completely determined by $g^2 c_F/g_V$, $g_V c_H$, and the
 296 degenerate mass of the triplet M_V ;
- 297 • c_{VVV} , c_{VVHH} , c_{VWW} can be neglected, as long as the interest is focused in narrow resonances.

298 The couplings of the new resonances to fermions and bosons depend in fact by several parameters;
 299 in the following paragraphs two simplified scenarios are discussed.

300 2.2.3 Benchmark model A: weak coupling scenario

301 Model A scenario aims at reproducing a simple generalization of the SM [15], obtained by extending
 302 the gauge symmetry group with an additional $SU(2)'$. The low-energy phenomena are expected to
 303 be dominated by the SM, while the high-energy processes are relevant for the additional symmetry,
 304 bringing additional light vector bosons in play.
 305 It can be shown that this kind of picture is portrayed by HVT when $c_H \sim -g^2/g_V^2$ and $c_F \sim 1$. This
 306 implies that:

$$g_V c_H \approx g^2/g_V, \\ g^2 c_F/g_V \approx g^2/g_V, \quad (2.28)$$

307 hence the partial decay widths into fermions (eq. 2.23) and bosons (eq. 2.27) differ only by a factor
 308 2 and the colour factor (N_c). Branching fractions for the model A benchmark scenario ($g_V = 1$) are
 309 shown in fig. 2.1 (left); total widths are reported in fig. 2.1 (right) for different coupling parameters
 310 g_V .

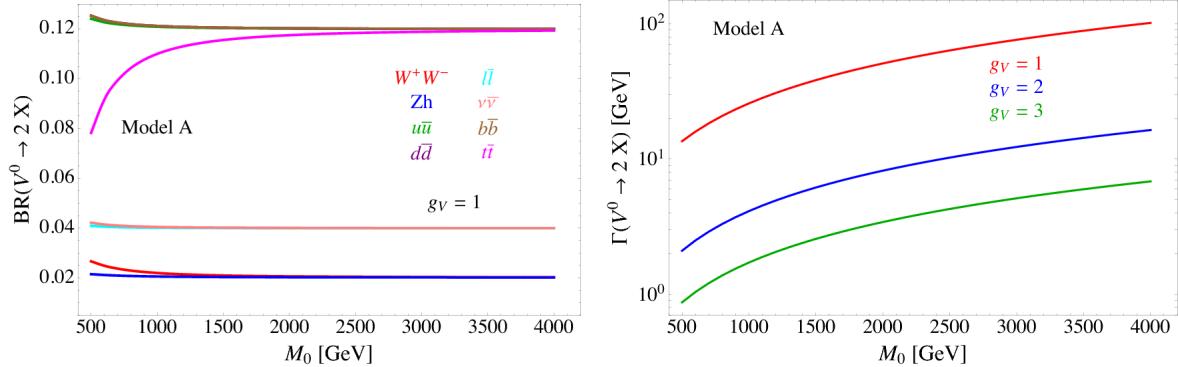


Figure 2.1: HVT model A scenario: branching fractions for fermionic and bosonic decays when $g_V = 1$ (left) as a function of the mass of the resonance M_0 ; total width of the resonance, as a function of its mass, considering different values of the parameter g_V (right).

311 2.2.4 Benchmark model B: strong coupling scenario

312 In composite Higgs models [17], the Higgs boson is the result of the spontaneous symmetry breaking
 313 of an $SO(5)$ symmetry to a $SO(4)$ group. New vector bosons are expected to appear, and the lightest
 314 ones can be represented by HVT model B when $c_H \sim c_F \sim 1$.
 315 In this case:

$$g_V c_H \approx -g_V, \\ g^2 c_F/g_V \approx g^2/g_V, \quad (2.29)$$

316 hence the decay into bosons is not suppressed by g_V parameter. In the benchmark scenario $g_V = 3$,
 317 decays into dibosons are largely dominant, as it can be seen in fig. 2.2 (left); the total decay width
 318 increases for larger g_V (fig. 2.2, right). When the resonances start to be broad, *i.e.* $\Gamma/M_V \gg 10\%$, the
 319 assumptions leading to the simplified model are no longer valid, hence higher order, non-resonant
 320 effects must be taken into account.

2.2 Heavy Vector Triplet

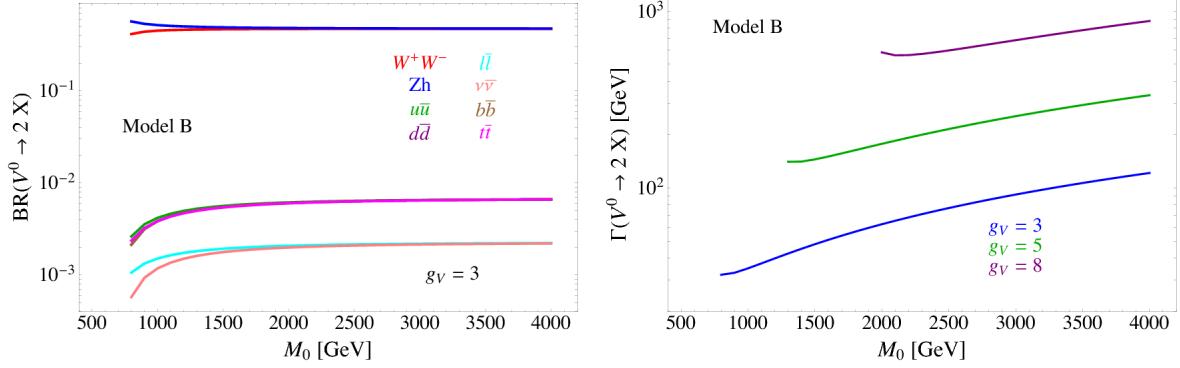


Figure 2.2: HVT model B scenario: branching fractions for fermionic and bosonic decays when $g_V = 3$ (left) as a function of the mass of the resonance M_0 ; total width of the resonance, as a function of its mass, considering different values of the parameter g_V (right).

2.2.5 HVT production

For resonance masses in the range of interest (~ 1 TeV), the production mechanisms expected to be relevant are Drell-Yan (fig. 2.3) and Vector Boson Fusion (VBF) (fig. 2.4).

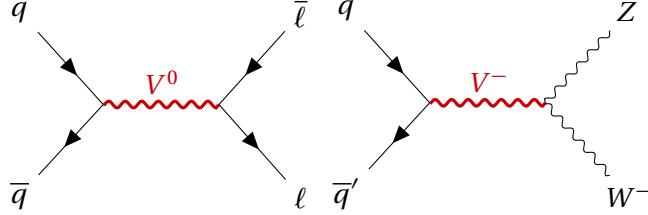


Figure 2.3: Examples of Drell-Yan production mechanism of a heavy V HVT boson: $q - \bar{q}$ quark scattering producing a neutral V^0 that decays leptonically (left); $q - \bar{q}'$ scattering producing a charged V^- that decays into a W and Z bosons (right).

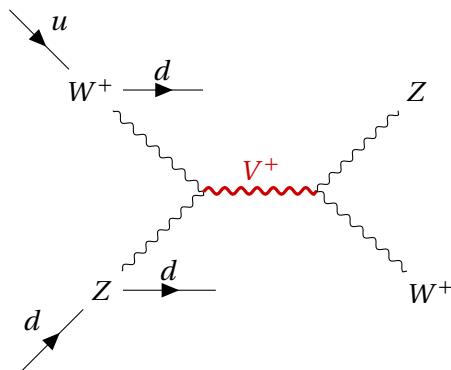


Figure 2.4: Example of VBF production mechanism of a heavy V HVT boson: a charged V^+ boson is produced by a couple of W and Z bosons, as a result of electroweak interactions of initial state u and d quarks. V^+ decays into a Z boson and a W^+ boson. The final state signature includes the presence of a pair of quarks, due to the primary interactions.

³²⁴ The cross-section of the production mechanisms is given by:

$$\sigma(pp \rightarrow V + X) = \sum_{i,j \in p} \frac{\Gamma_{V \rightarrow ij}}{M_V} f(J, S_i, S_j) g(C_i, C_j) \left. \frac{dL_{ij}}{d\hat{s}} \right|_{\hat{s}=M_V^2}, \quad (2.30)$$

³²⁵ where i, j are the partons involved in the hard interaction, $\Gamma_{V \rightarrow ij}$ is the partial width of the process
³²⁶ $V \rightarrow ij$, $f(J, S_i, S_j)$ is a function of the spin of the resonance and of the partons, $g(C_i, C_j)$ is a func-
³²⁷ tion of the colour factors of each parton, \hat{s} is the center-of-mass energy at parton level and $\frac{dL_{ij}}{d\hat{s}}$ are
³²⁸ the parton luminosities, that are independent from HVT model (that enters only in $\Gamma_{V \rightarrow ij}$).
³²⁹ Parton luminosities, calculated for a center-of-mass energy of 14 TeV starting from quark and anti-
³³⁰ quark parton distribution functions (PDF), are displayed in fig. 2.5 (Drell-Yan mechanism) and 2.6
³³¹ (VBF mechanism). VBF luminosities are suppressed by the α_{weak} factor, therefore the process is
 relevant only when the bosonic decays of the triplet are dominant (strongly coupled scenario).

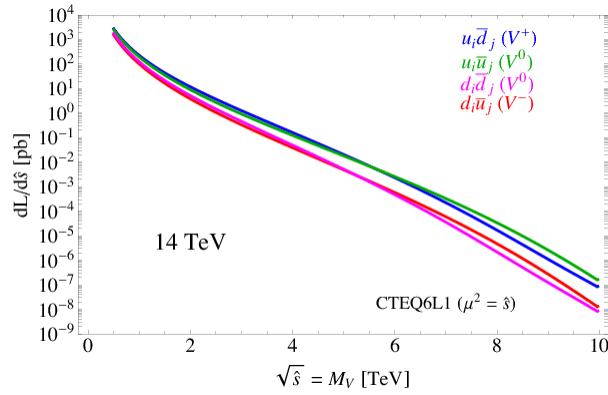


Figure 2.5: Parton luminosities for Drell-Yan process between i and j partons, as a function of the parton center-of-mass energy, for the LHC proton-proton collisions performed at 14 TeV.

³³²

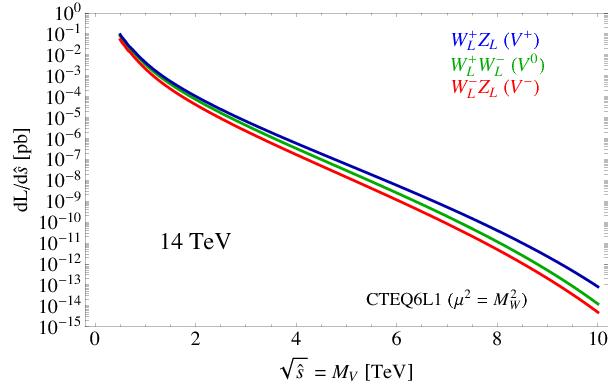


Figure 2.6: Parton luminosities for VBF process between i and j partons, as a function of the parton center-of-mass energy, for the LHC proton-proton collisions performed at 14 TeV.

2.2 Heavy Vector Triplet

2.2.6 Search for HVT resonances at LHC

No evidence of HVT resonances has been observed so far at the LHC experiments. Data collected by the ATLAS and CMS detectors are used to set limits on the HVT resonance masses and coupling parameters. Experimental results from proton-proton collisions performed at a center-of-mass energy of 8 TeV (Run 1 era) at the LHC brought to the following conclusions. A weakly coupled resonance, in the context of benchmark model A ($g_V = 1$) was excluded up to 3 TeV by Run 1 data. By looking at parton luminosities in fig.2.5, in data produced by LHC proton-proton collision at 14 TeV, collected for an integrated luminosity of 300 fb^{-1} , the sensitivity is expected to increase up to $m_V \approx 6 \text{ TeV}$. A strongly coupled resonance, in the context of benchmark model B ($g_V = 3$) is excluded up to 2 TeV by Run 1 data. Data produced by LHC at 14 TeV should increase the sensitivity up to $m_V \approx 3-4 \text{ TeV}$. The most stringent limits are provided by the latest data produced by LHC at a center-of-mass energy of 13 TeV (Run 2 era).

Numerous searches for HVT triplet have been performed at the CMS experiment in different final states: the most sensitive ones were those in all-hadronic topology. The analysis searching for WW , WZ , ZZ resonances in the $q\bar{q}q\bar{q}$ final state [23, 24] excludes a W' with mass below 3.6 TeV and a Z' with mass below 2.7 TeV in the model B scenario (fig. 2.7). The analysis searching for WH , ZH resonances in the $q\bar{q}b\bar{b}$ final state [25, 26] excludes a W' lighter than 2.97 (3.15) TeV in the HVT model A (model B), and a Z' up to 1.67 (2.26) TeV in HVT model A (model B) (fig. 2.8). In fig. 2.9, results of [23, 24] (left) and [25, 26] (right) searches are interpreted as exclusion contours in the coupling parameters plane of the HVT model ($g_V c_H$ and $g^2 c_F / g_V$). In the gray shaded area, the narrow width approximation fails. The coloured curves display the parameter exclusion for different mass hypotheses of the triplet. Coloured markers show the model A and B benchmark scenarios.

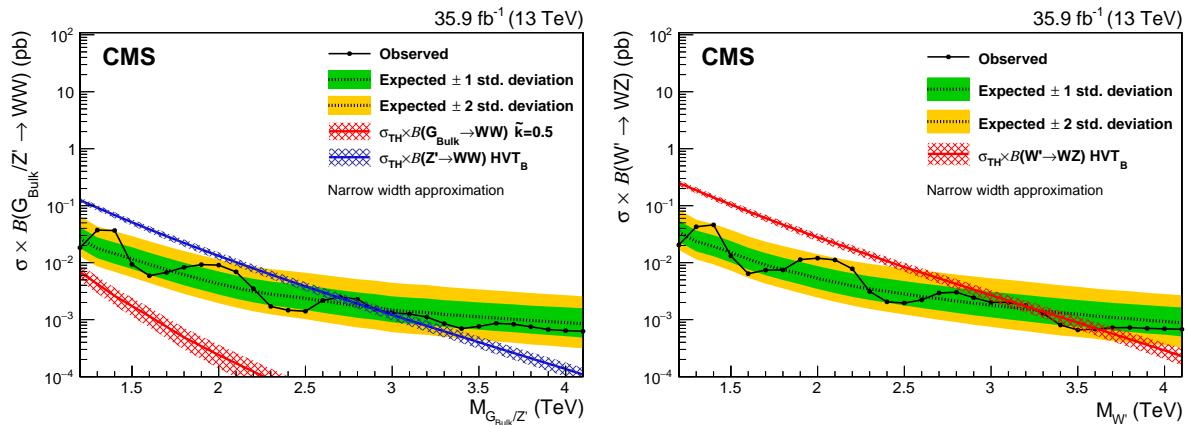


Figure 2.7: The observed and expected limits, with 68% and 95% uncertainty bands, on the product of the cross-section and branching fraction $\sigma \mathcal{B}(Z' \rightarrow WW)$ for a spin-1 Z' (left) and $\sigma \mathcal{B}(W' \rightarrow WZ)$ for a spin-1 W' (right), as a function of the reconstructed mass of the diboson resonance. The coloured lines show the theoretical predictions for the HVT model B. [23, 24]

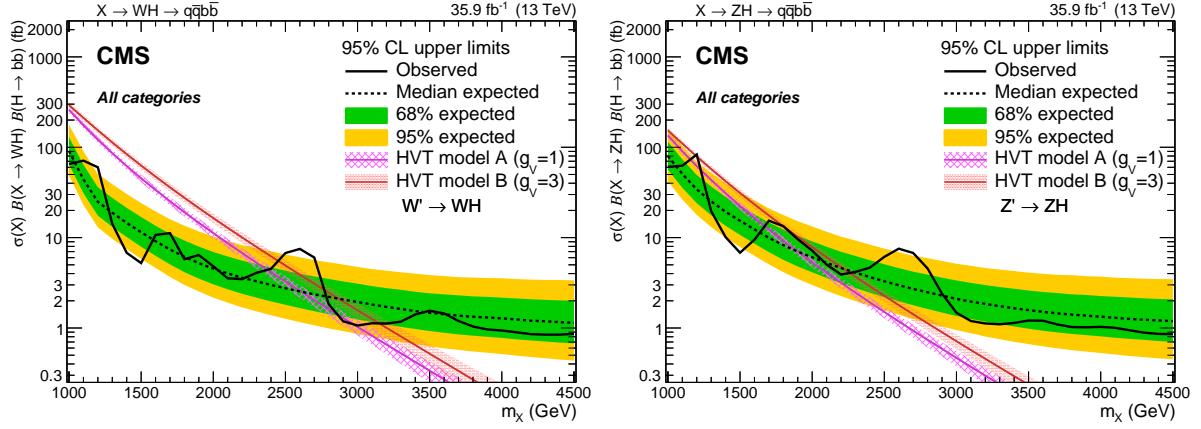


Figure 2.8: The observed and expected limits, with 68% and 95% uncertainty bands, on the product of the cross-section and branching fraction $\sigma \mathcal{B}(W' \rightarrow WH)$ for a spin-1 W' (left) and $\sigma \mathcal{B}(Z' \rightarrow ZH)$ for a spin-1 Z' (right), as a function of the reconstructed mass of the diboson resonance. The coloured lines show the theoretical predictions for the HVT model A and B. [25, 26]

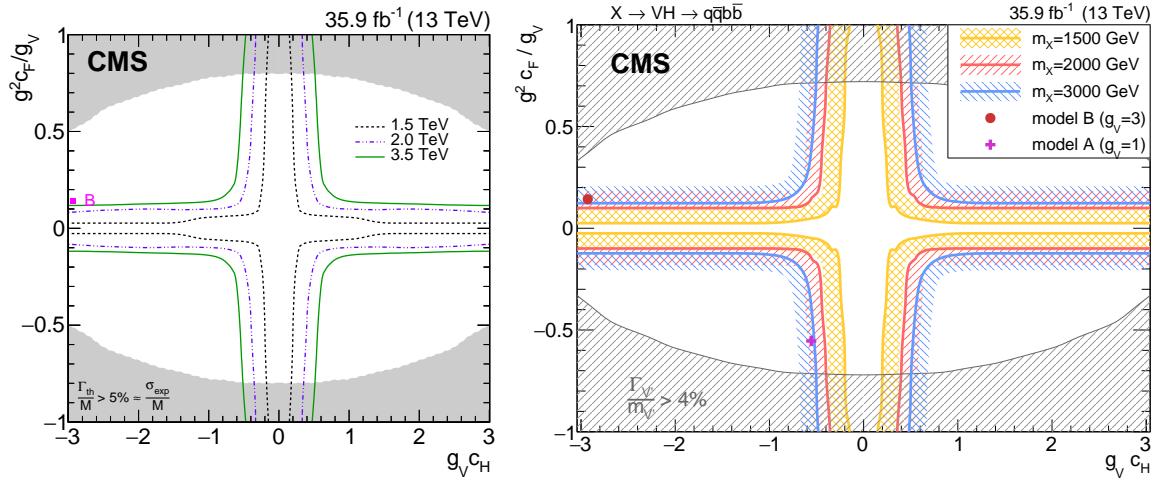


Figure 2.9: Exclusion contours in the coupling parameters plane of the HVT model ($g_V c_H$ and $g^2 c_F / g_V$), for [23, 24] analysis (left), and [25, 26] analysis (right).

2.2 Heavy Vector Triplet

355 Many other final states have been exploited at CMS: $ZW, ZZ \rightarrow \ell\bar{\ell}q\bar{q}$ [27]; $WH, ZH \rightarrow (\ell\bar{\ell}, \ell\nu, \nu\bar{\nu})b\bar{b}$ [28];
 356 $WZ, WW \rightarrow \ell\nu q\bar{q}$ [29]. Finally, $ZW, ZZ \rightarrow \nu\bar{\nu}q\bar{q}$ [30] results will be extensively described in this
 357 thesis.

358 The results (or preliminary results) on HVT searches in diboson final states, performed with 2016
 359 data and published by the CMS Collaboration so far, are summarized in fig. 2.10. The dark orange
 360 curve corresponds to the $(WZ, WW) \rightarrow q\bar{q}q\bar{q}$ analysis [23, 24]; the light blue curve corresponds
 361 to the $(WH, ZH) \rightarrow q\bar{q}b\bar{b}$ analysis [25, 26]; the light orange curve corresponds to the analysis dis-
 362 cussed in this thesis [30].

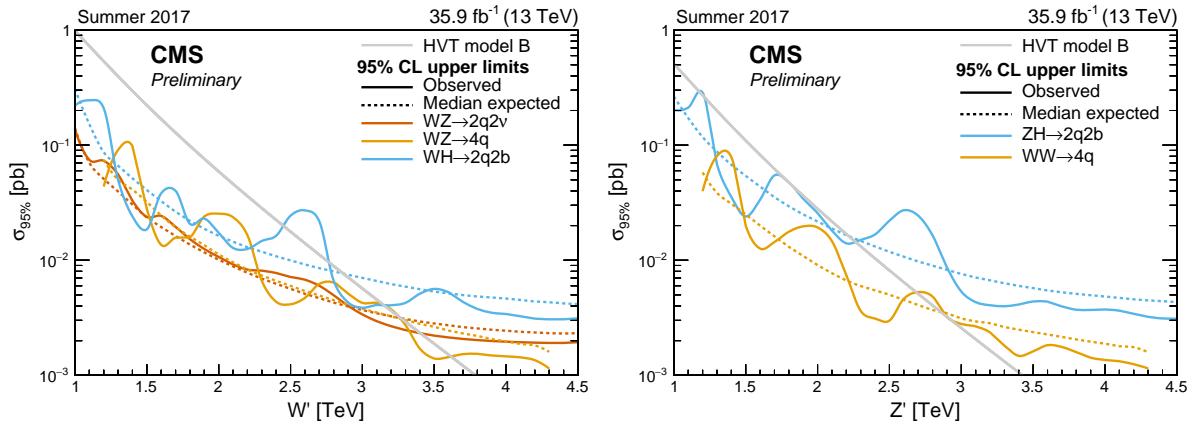


Figure 2.10: The observed and expected limits on the product of the cross-section and branching fraction $\sigma\beta(W' \rightarrow (WZ, WH))$ for a spin-1 W' (left) and $\sigma\beta(Z' \rightarrow (ZH, WW))$ for a spin-1 Z' (right), as a function of the reconstructed mass of the diboson resonance. The gray line shows the theoretical prediction for the HVT model B.

363 Searches for HVT model B resonances have been performed at the ATLAS experiment as well. Re-
 364 sults for a $W' \rightarrow WZ$ reported in fig. 2.11 include the searches performed in $WW, WZ, ZZ \rightarrow q\bar{q}q\bar{q}$
 365 final state [31]; $WZ, WW \rightarrow \ell\nu q\bar{q}$ final state [32]; $ZW, ZZ \rightarrow (\ell\bar{\ell}, \ell\nu, \nu\bar{\nu})q\bar{q}$ final state [33]. The all-
 366 hadronic final state has the best sensitivity and it excludes a W' resonance up to 3.3 TeV (model B
 367 scenario). Results for a $W' \rightarrow WH$ and for a $Z' \rightarrow ZH$ are displayed in fig. 2.12 (left and right respec-
 368 tively), and they include searches performed in $WH, ZH \rightarrow q\bar{q}b\bar{b}$ final state [34], and $WH, ZH \rightarrow$
 369 $\ell\bar{\ell}, \ell\nu, \nu\bar{\nu}b\bar{b}$ [35]. A W' is excluded up to 2.9 TeV and a Z' is excluded up to 2.8 TeV (in the model B
 370 scenario).

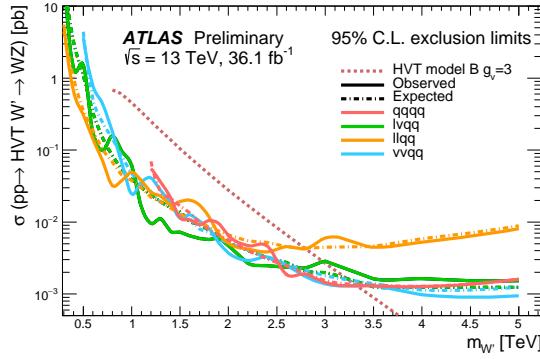


Figure 2.11: The observed and expected limits on the product of the cross-section and branching fraction $\sigma \mathcal{B}(W' \rightarrow WZ)$ for a spin-1 W' , as a function of the reconstructed mass of the diboson resonance. The dark red dotted line shows the theoretical predictions for the HVT model B.

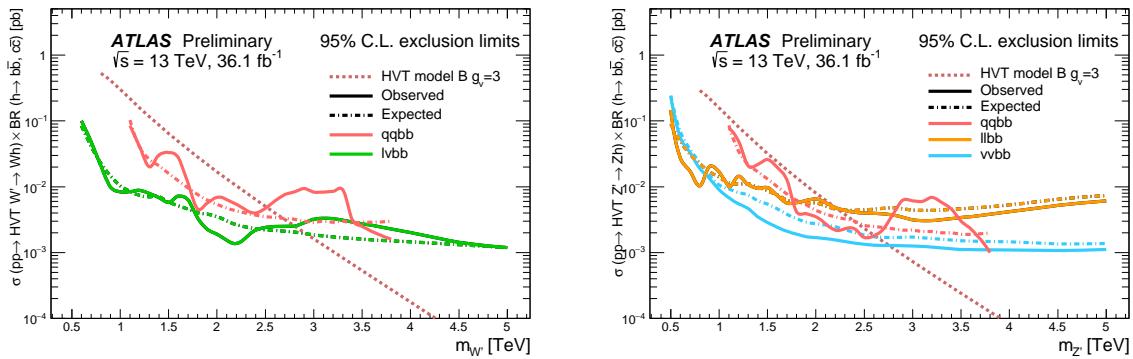


Figure 2.12: The observed and expected limits on the product of the cross-section and branching fraction $\sigma \mathcal{B}(W' \rightarrow WH)$ for a spin-1 W' (left) and $\sigma \mathcal{B}(Z' \rightarrow ZH)$ for a spin-1 Z' (right), as a function of the reconstructed mass of the diboson resonance. The dark red dotted lines show the theoretical predictions for the HVT model B.

371 2.3 Warped extra-dimension

372 The Randall-Sundrum model [36, 37] (RS1) proposes the introduction of one additional warped di-
 373 mension in order to solve the hierarchy problem. The metric of the 5-dimensional space (a slice of
 374 AdS_5) generates an exponential hierarchy between the electroweak and Planck scales, associated re-
 375 spectively to the TeV three-brane, where the SM particles are confined, and the Planck three-brane.
 376 As a consequence of the new geometry, spin-2 massive gravitons are predicted to exist.
 377 The bulk extension of the Randall-Sundrum model [38, 39] states that the SM fields can propagate in
 378 the extra-dimension. Light fermions are near the Planck brane, heavy fermions are close to the TeV
 379 brane, while the Higgs sector is confined in the TeV brane. Higgs couplings to the heavy fermions
 380 are therefore expected to be stronger: this naturally arising hierarchy of the masses of the SM fields
 381 gives a solution to the flavour problem. In this scenario, the fermionic decays of the bulk gravitons
 382 are suppressed, while the bosonic decays are preferred.

383 2.3.1 Randall-Sundrum original model (RS1)

384 The existence of additional n -dimensions implies that the effective Planck scale observed in 4-
 385 dimensions, $M_{PL} = 1.220910^{19}$ GeV, is related to the fundamental $4+n$ -dimensional Planck scale,
 386 M , via the geometry. If the 4-dimensional and the n additional metrics are factorizable, the re-
 387duced Planck scale $\overline{M}_{PL} = M_{PL}/2\pi$ can be seen as the product of M and the volume of the compact
 388 space V_n :

$$\overline{M}_{PL}^2 = V_n M^{2+n}. \quad (2.31)$$

389 If $M \sim$ TeV, this implies that V_n must be very large, hence the compactification scale $\mu \sim 1/V_n^{1/n}$
 390 is necessarily small (eV – MeV for $n=2 - 7$). Given the smallness of μ when compared to the elec-
 391 troweak scale, the effects of the n extra-dimensions should be evident in SM processes. Since they
 392 are not observed, SM particles are assumed to be confined in a 4-dimensional space, the TeV three-
 393 brane, while only gravity is allowed to propagate into the $4+n$ -dimensional space, the bulk. This
 394 mechanism solves the hierarchy of the Higgs scale, but on the other hand it introduces a new hier-
 395 archy between μ and M .

396 In the Randall-Sundrum model [36, 37], only one additional dimension is added. The geometry of
 397 the 5-dimensional bulk is non-factorizable, and it is a slice of AdS_5 spacetime.¹ The 4-dimensional
 398 metric is multiplied by an exponential function of the fifth dimension (the "warp" factor):

$$ds^2 = e^{-2kr_c\varphi} \eta_{\mu\nu} dx^\mu dx^\nu + r_c^2 d\varphi^2; \quad (2.32)$$

399 x^μ are the usual 4-dimensional coordinates, $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$ is the Minkowski metric, k is a
 400 scale of order of \overline{M}_{PL} , φ is the coordinate of the extra-dimension, $0 < |\varphi| < \pi$, and r_c is the com-
 401 pactification radius of this finite interval. 4-dimensional mass scales are obtained by multiplying
 402 the bulk masses by $e^{-2kr_c\varphi}$: given the exponential form of the warp factor, a small r_c suffices for
 403 generating a large hierarchy between Planck and Higgs scales.

404 Two 4-dimensional three-branes are located at the boundaries of the fifth dimension: the visible
 405 brane at $\varphi = \pi$; the hidden brane at $\varphi = 0$, and their metrics are obtained starting from the bulk

¹An n -dimensional anti-de Sitter space (AdS_n) is a maximally symmetric Lorentzian manifold, that solves the Einstein equation with a negative curvature (negative cosmological constant).

406 5-dimensional metric G_{MN} , where $M, N = \mu, \varphi$:

$$\begin{aligned} g_{\mu\nu}^{\text{vis}}(x^\mu) &= G_{\mu\nu}(x^\mu, \varphi = \pi), \\ g_{\mu\nu}^{\text{hid}}(x^\mu) &= G_{\mu\nu}(x^\mu, \varphi = 0). \end{aligned} \quad (2.33)$$

407 The classical action is given by:

$$\begin{aligned} S &= S_{\text{gravity}} + S_{\text{vis}} + S_{\text{hid}}, \\ S_{\text{gravity}} &= \int d^4x \int_{-\pi}^{+\pi} d\varphi \sqrt{-G} (-\Lambda + 2M^3 \mathcal{R}), \\ S_{\text{vis}} &= \int d^4x \sqrt{-g_{\text{vis}}} (\mathcal{L}_{\text{vis}} - V_{\text{vis}}), \\ S_{\text{hid}} &= \int d^4x \sqrt{-g_{\text{hid}}} (\mathcal{L}_{\text{hid}} - V_{\text{hid}}), \end{aligned} \quad (2.34)$$

408 where G (g) is the trace of the G_{MN} ($g_{\mu\nu}$) metric, Λ is the cosmological constant in the bulk, \mathcal{R} is
409 the 5-dimensional Ricci scalar, \mathcal{L} and V are the lagrangian and the vacuum energy of the hidden
410 and visible branes.

411 A 5-dimensional metric that preserves the 4-dimensional Poincaré invariance has the form:

$$ds^2 = e^{-2\sigma(\varphi)} \eta_{\mu\nu} dx^\mu dx^\nu + r_c^2 d\varphi^2. \quad (2.35)$$

412 The Poincaré invariance guarantees that r_c does not depend on x^μ . Given 2.35, the solution of the
413 5-dimensional Einstein's equations simplifies into:

$$\sigma = r_c |\varphi| \sqrt{\frac{-\Lambda}{24M^3}}. \quad (2.36)$$

414 Furthermore, the Poincaré invariance imposes constraints to the vacuum energies and cosmological
415 constant:

$$\begin{aligned} V_{\text{hid}} &= -V_{\text{vis}} = 24M^3 k \\ \Lambda &= -24M^3 k^2. \end{aligned} \quad (2.37)$$

416 The final 5-dimensional metric is then:

$$ds^2 = e^{-2kr_c|\varphi|} \eta_{\mu\nu} dx^\mu dx^\nu + r_c^2 d\varphi^2. \quad (2.38)$$

417 A small r_c is considered, so the effects of the fifth dimension on the 4-dimensional spacetime can't
418 be appreciated. A 4-dimensional effective field theory approach is therefore motivated, and its mass
419 parameters are related to the bulk parameters, M , k and r_c . In the Randall-Sundrum model, the SM
420 matter fields are confined in the TeV brane.

421 The massless gravitons, the mediators of the gravitational interaction in the effective field theory,
422 are the zero modes ($h_{\mu\nu}$) of the quantum fluctuations of the classical solution (2.38):

$$ds^2 = e^{-2kT(x)|\varphi|} (\eta_{\mu\nu} + h_{\mu\nu}(x)) dx^\mu dx^\nu + T^2(x) d\varphi^2, \quad (2.39)$$

423 where the usual Minkowski metric has been replaced by $\bar{g}_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}$; $h_{\mu\nu}$ are the tensor fluc-
424 tuations around the Minkowski space, and represent both the physical graviton in 4-dimensions

2.3 Warped extra-dimension

425 and the massless mode of the Kaluza-Klein decomposition of the bulk metric. r_c is the vacuum ex-
426 pectation value of $T(x)$.

427 By substituting eq. 2.39 in the classical action 2.34, an effective action can be extracted, and in par-
428 ticular the curvature term holds:

$$S_{\text{eff}} \sim \int d^4x \int_{-\pi}^{+\pi} d\varphi 2M^3 r_c e^{-2kr_c|\varphi|} \overline{\mathcal{R}} \sqrt{-\overline{g}}, \quad (2.40)$$

429 where \overline{g} is the trace of $\overline{g}_{\mu\nu}$, and $\overline{\mathcal{R}}$ is the 4-dimensional Ricci scalar of $\overline{g}_{\mu\nu}$ metric. In this effective
430 4-dimensional action, the φ dependence can be integrated out, and the 4-dimensional Planck mass
431 can be calculated:

$$\overline{M}_{PL}^2 = M^3 r_c \int_{-\pi}^{+\pi} d\varphi e^{-2kr_c|\varphi|} = \frac{M^3}{k} (1 - e^{-2kr_c\pi}). \quad (2.41)$$

432 It can be shown [36] that a field with a fundamental mass parameter m_0 in the bulk manifests in the
433 visible three-brane with a physical mass m :

$$m = e^{-kr_c\pi} m_0. \quad (2.42)$$

434 Scales $m \sim \text{TeV}$ are generated from $m_0 \sim \overline{M}_{PL}$ if $e^{kr_c\pi} \sim 10^{15}$. This relation stands still when Higgs
435 field is introduced and confined in the visible three-brane:

$$v = e^{-kr_c\pi} v_0, \quad (2.43)$$

436 where v is the Higgs vacuum expectation value in the TeV brane and v_0 is the 5-dimensional Higgs
437 v.e.v.

438 The hierarchy problem is then solved by the exponential warp factor. The weakness of gravity in the
439 TeV three-brane is motivated by the small overlap with the graviton wave function.

440 In order to calculate the mass spectrum of the graviton in the TeV brane, the tensor fluctuations of
441 the Minkowski metric are expanded into a Kaluza-Klein (KK) tower $h_{\mu\nu}^{(n)}$:

$$h_{\mu\nu}(x, \varphi) = \sum_{n=0}^{\infty} h_{\mu\nu}^{(n)}(x) \frac{\chi^{(n)}(\varphi)}{\sqrt{r_c}}. \quad (2.44)$$

442 Once a suitable gauge is chosen, i.e. $\eta^{\mu\nu} \partial_\mu h_{\nu a}^{(n)} = \eta^{\mu\nu} h_{\mu\nu}^{(n)} = 0$, the equation of motion of $h_{\mu\nu}^{(n)}$ becomes
443 the Klein-Gordon relation, where $m_n^G \geq 0$:

$$(\eta^{\mu\nu} \partial_\mu \partial_\nu - (m_n^G)^2) h_{\mu\nu}^{(n)}(x) = 0. \quad (2.45)$$

444 By substituting eq. 2.44 into Einstein's equation, the solutions for $\chi^{(n)}(\varphi)$ (commonly called "pro-
445 files") are [40, 41]:

$$\chi^{(n)}(\varphi) = \frac{e^{2\sigma}}{N} [J_2(z_n^G) + \alpha_n Y_2(z_n^G)], \quad (2.46)$$

446 where J_2 and Y_2 are second order Bessel functions, N is the normalization of the wavefunction, α_n
447 are coefficients and $z_n^G = m_n^G e^{\sigma(\varphi)}/k$. m_n^G is the mass of the n -mode, and it depends on the roots
448 of the Bessel functions $z_n^G = (3.83, 7.02, 10.17, 13.32, \dots)$. In the limit $m_n^G/k \ll 1$ and $e^{kr_c\pi} \gg 1$:

$$m_n^G = k z_n^G(\pi) e^{-kr_c\pi}. \quad (2.47)$$

449 The interactions between the graviton KK modes and the matter fields in the TeV brane can be de-
 450 rived from the 4-dimensional effective Lagrangian, once $h_{\mu\nu}$ is replaced by its KK decomposition:

$$\mathcal{L} = -\frac{1}{M_{PL}} T^{\mu\nu}(x) h_{\mu\nu}^{(0)} - \frac{1}{e^{-kr_c\pi} M_{PL}} T^{\mu\nu}(x) \sum_{n=1}^{\infty} h_{\mu\nu}^{(n)}(x); \quad (2.48)$$

451 $T^{\mu\nu}$ is the space energy-momentum tensor of the matter fields. The zero mode of the gravitons cou-
 452 pling is $1/M_{PL}$, while higher order KK modes couplings to all SM fields are suppressed by $e^{-kr_c\pi} M_{PL}$,
 453 that is of the order of the TeV scale. Spin-2 KK masses and couplings are hence determined by the
 454 TeV scale, or, equivalently, KK gravitons are close to the TeV brane. This implies that KK gravitons
 455 can be produced via $q\bar{q}$ or gluon fusion, and that a leptonic decay of the resonance could represent
 456 a very clear signal signature.

457 2.3.2 Bulk extension of RS1: graviton production and decays

458 An extension of the original RS1 formulation has been proposed. It states that the usual SM fields are
 459 no longer confined in the TeV brane, but they are the zero modes of the corresponding 5-dimensional
 460 SM fields. If first and second generation fermions are close to the Planck brane, contribution to
 461 flavour changing neutral currents by higher-dimensional operators are suppressed. These contrib-
 462 utions are excluded by electroweak precision tests, but they were not prevented in original RS1.
 463 The second motivation behind the choice is, as mentioned previously, the naturally arising flavour
 464 hierarchy: first and second generation quarks have small Yukawa couplings to the Higgs sector, con-
 465 fined in the TeV brane, while top quark and bosons have stronger Yukawa couplings.
 466 In this picture, couplings between higher-order KK gravitons and light fermions are strongly sup-
 467 pressed, resulting into a negligible KK gravitons production via $q\bar{q}$, whilst gluon fusion production
 468 becomes dominant. KK gravitons decays into top quarks and Higgs bosons are dominant, given
 469 that both their profiles are near the TeV brane, while leptonic decays are negligible. Via the equi-
 470 valence theorem, the Goldstone bosons are equivalent to the longitudinally polarized weak bosons,
 471 W_L^\pm and Z_L , that have profiles close to the TeV brane. Decays of KK gravitons into weak dibosons
 472 (and production in VBF) are comparable to di-top and di-Higgs decays.

473

474 The KK decomposition and the KK mass spectrum of the graviton have already been presented in
 475 sec. 2.3.1. The KK decomposition of a massless 5-dimensional gauge field $A_M(x, \varphi)$ is similarly per-
 476 formed [42]:

$$A_\mu(x, \varphi) = \sum_{n=0}^{\infty} A_\mu^{(n)}(x) \frac{\chi_A^{(n)}(\varphi)}{\sqrt{r_c}}. \quad (2.49)$$

477 The profiles for the gauge fields are:

$$\chi_A^{(n)}(\varphi) = \frac{e^\sigma}{N_A} [J_1(z_n^A) + \alpha_n^A Y_1(z_n^A)], \quad (2.50)$$

478 where J_1 and Y_1 are first order Bessel functions. Similarly to eq. 2.51, the mass spectrum of the gauge
 479 field is:

$$m_n^A = k z_n^A(\pi) e^{-kr_c\pi}; \quad (2.51)$$

480 the first roots of the Bessel functions are $z_n^A = (2.45, 5.57, 8.70, 11.84, \dots)$.

2.3 Warped extra-dimension

- 481 The Lagrangian expressing the interaction between the m and n modes of the bulk field F to the q
 482 KK gravitons mode G is [42]:

$$\mathcal{L}_{G-F} = \sum_{m,n,q} C_{mnq}^{FFG} \frac{1}{M_{PL}} \eta^{\mu\alpha} \eta^{\nu\beta} h_{\alpha\beta}^{(q)}(x) T_{\mu\nu}^{(m,n)}(x), \quad (2.52)$$

- 483 C_{mnq}^{FFG} is the overlap integral of the profiles:

$$C_{mnq}^{FFG} = \int \frac{d\varphi}{\sqrt{k}} e^{t\sigma} \frac{\chi_F^{(m)} \chi_F^{(n)} \chi_G^{(q)}}{\sqrt{r_c}}; \quad (2.53)$$

- 484 t depends on the type of field considered.
 485 The coupling between gluons and the q KK graviton mode is given by:

$$C_{00q}^{AAG} = e^{k\pi r_c} \frac{2[1 - J_0(x_q^G)]}{k\pi r_c (x_q^G)^2 |J_2(x_q^G)|}. \quad (2.54)$$

- 486 Once eq. 2.54 is put in eq. 2.52, the most significant partial decay widths into the q KK graviton mode
 487 are:

$$\begin{aligned} \Gamma(G \rightarrow t_R \bar{t}_R) &\sim N_c \frac{[\tilde{k} x_q^G]^2 m_q^G}{320\pi}, \\ \Gamma(G \rightarrow hh) &\sim \frac{[\tilde{k} x_q^G]^2 m_q^G}{960\pi}, \\ \Gamma(G \rightarrow W_L^+ W_L^-) &\sim \frac{[\tilde{k} x_q^G]^2 m_q^G}{480\pi}, \\ \Gamma(G \rightarrow Z_L Z_L) &\sim \frac{[\tilde{k} x_q^G]^2 m_q^G}{960\pi}, \end{aligned} \quad (2.55)$$

- 488 where $\tilde{k} = k/\overline{M}_{PL}$; the total decay width is:

$$\Gamma_G = \frac{13 [\tilde{k} x_q^G]^2 m_q^G}{960\pi}. \quad (2.56)$$

- 489 Calculations, so far, have been performed considering $M \sim \overline{M}_{PL}$ and $k < M$, hypotheses under which
 490 the solution for the bulk metric (eq. 2.38) is valid. Hence, $\tilde{k} = k/\overline{M}_{PL} \leq 1$ is taken as a reference
 491 interval. This has also phenomenological consequences on the width of the resonance, as stated in
 492 eq. 2.56. The total decay width of the lightest KK graviton mode, compared to its mass, is shown as
 493 a function of \tilde{k} in fig. 2.13 [43]. At $\tilde{k} = 1$, in the bulk scenario, the KK graviton width is expected to
 494 be few % of its mass, up to 4 TeV (dotted red curve). The narrow width approximation holds, hence
 495 the resonance properties can be probed at the peak, neglecting the effects in the tails of the mass
 496 distribution.

- 497 The total cross-section of a bulk graviton, produced at the LHC in proton-proton interactions via
 498 gluon fusion (displayed in fig. 2.14), decaying into a couple of vector bosons (for the purpose of this

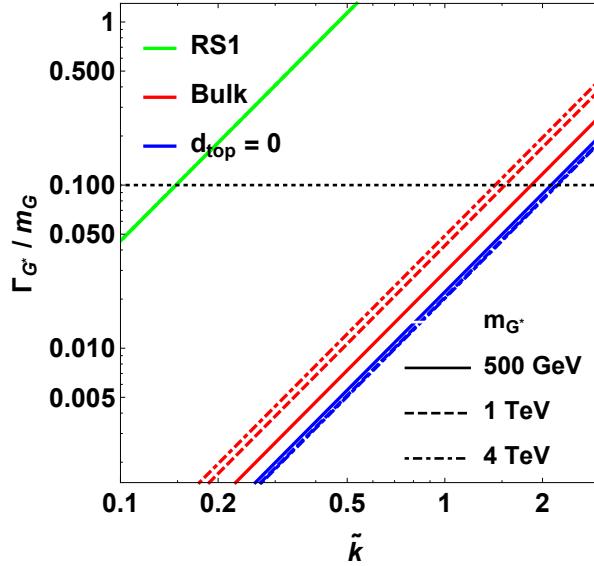


Figure 2.13: Width of the KK gravitons, in units of the mass of the resonance, as a function of the curvature parameter \tilde{k} . The red curves represent the bulk extension of RS1 original model for different mass hypotheses (from 500 GeV up to 4 TeV).

499 thesis, a final state with two longitudinally polarized Z bosons is considered) is expressed as a func-
 500 tion of the parton level cross-section $\hat{\sigma}$, the gluon parton distribution functions f_g , the momentum
 501 transfer $Q^2 \sim (m_q^G)^2$ and the center-of-mass energy s :

$$\sigma(pp \rightarrow ZZ) = \int dx_1 dx_2 f_g(x_1, Q^2) f_g(x_2, Q^2) \hat{\sigma}(x_1 x_2 s). \quad (2.57)$$

502 The differential parton level cross-section, averaged over colors and initial spin states, is (hatted
 503 quantities are calculated in the center-of-mass frame):

$$\frac{d\hat{\sigma}(gg \rightarrow ZZ)}{d \cos \theta} \approx \frac{|\mathcal{M}_{+-00}|^2}{1024\pi \hat{s}}, \quad (2.58)$$

504 where $|\mathcal{M}_{+-00}|$ is the matrix element of the dominant contribution in $gg \rightarrow VV$ process (Γ_G is de-
 505 fined in eq. 2.56, a, b are colour factors):

$$\mathcal{M}_{+-00}(g^a g^b \rightarrow VV) = -C_{00q}^{AAG} e^{-k\pi r_c} \left(\frac{x_n^G \tilde{k}}{m_n^G} \right)^2 \sum_n \frac{\delta_{ab} \mathcal{A}_{+-00}}{\hat{s} - m_n^G + i\Gamma_G m_n^G}. \quad (2.59)$$

506 The relevant amplitudes taken account in the matrix element calculation are [38]:

$$\mathcal{A}_{+-00} = \mathcal{A}_{-+00} = \frac{(1 - 1/\beta_Z^2)(\beta_Z^2 - 2)[(\hat{t} - \hat{u})^2 - \beta_Z^2 \hat{s}^2] \hat{s}}{8M_Z^2}, \quad (2.60)$$

507 where $\beta_Z^2 = 1 - 4M_Z^2/\hat{s}$ and M_Z is the mass of the Z boson.

508 2.3.3 Search for KK bulk gravitons at LHC

509 No evidence of spin-2 bulk graviton resonances has been observed so far at the LHC experiments.
 510 Data collected by the ATLAS and CMS detectors are used to set limits on the graviton masses, gen-

2.3 Warped extra-dimension

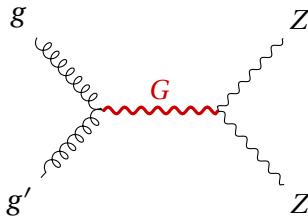


Figure 2.14: Gluon fusion production mechanism for a KK graviton that decays into a couple of Z bosons.

511 erally considering different curvature parameter \tilde{k} hypotheses, once assured the narrow width ap-
 512 proximation is still valid (up to $\tilde{k} \sim 1$). The most stringent limits have been set with Run 2 data.

513 Many results of the diboson searches performed at CMS and already presented in sec. 2.2.6 are
 514 interpreted in the context of the bulk graviton models, together with the additional final states
 515 $ZZ \rightarrow \ell\bar{\ell}\nu\bar{\nu}$ [44] and $HH \rightarrow b\bar{b}b\bar{b}$ [45]. The most interesting limit is provided by the search for
 516 $ZZ \rightarrow \ell\bar{\ell}\nu\bar{\nu}$ resonances [44], that, under the hypothesis $\tilde{k} = 0.5$, excludes a spin-2 bulk graviton with
 517 a mass lower than 800 GeV (fig. 2.15).

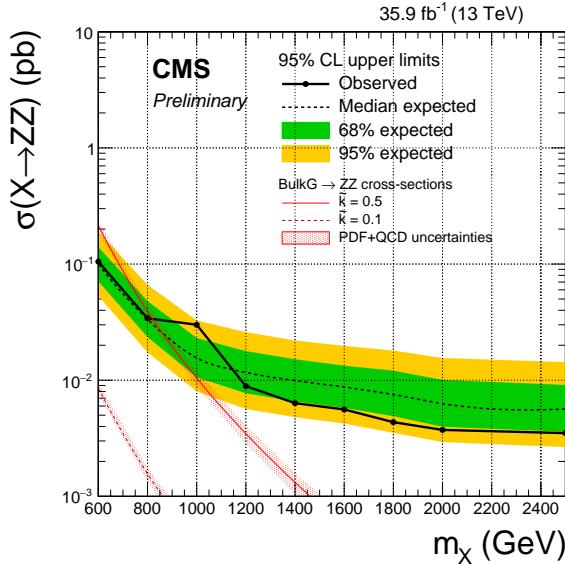


Figure 2.15: The observed and expected limits, with 68% and 95% uncertainty bands, on the product of the cross-section and branching fraction $\sigma \mathcal{B}(G \rightarrow ZZ)$ for a spin-2 bulk graviton, as a function of the reconstructed mass of the diboson resonance. The coloured lines show the theoretical predictions for $\tilde{k} = 0.1$ and 0.5 . [44]

518 The results (or preliminary results) on bulk graviton searches in diboson final states, performed with
 519 2016 data and published by the CMS Collaboration so far, are summarized in fig. 2.16. The dark
 520 orange curve and the pink curve correspond to two possible final states included the $(ZZ, WW) \rightarrow$
 521 $q\bar{q}q\bar{q}$ analysis [23, 24]; the light blue curve corresponds to the $ZZ \rightarrow \ell\bar{\ell}\nu\bar{\nu}$ analysis [44]; the green
 522 curve corresponds to the $HH \rightarrow b\bar{b}b\bar{b}$ analysis [45]; the light orange curve corresponds to the
 523 analysis discussed in this thesis [30].

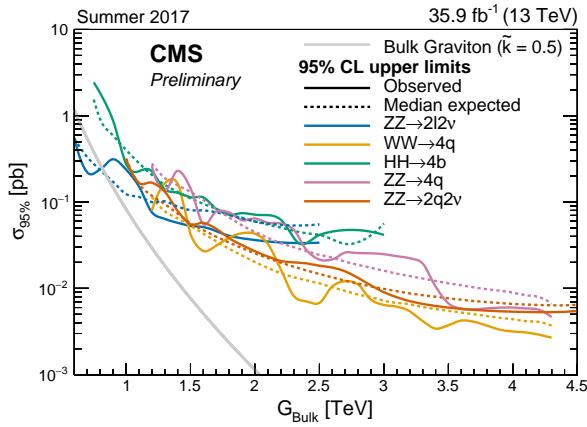


Figure 2.16: The observed and expected limits on the product of the cross-section and branching fraction $\sigma \mathcal{B}(G \rightarrow (WW, ZZ))$ for a spin-2 bulk graviton, as a function of the reconstructed mass of the diboson resonance. The gray line shows the theoretical prediction for the bulk graviton model, once assumed a curvature parameter $\tilde{k} = 0.5$.

524 Similarly for the ATLAS experiment, searches for diboson resonances in sec. 2.2.6 have been inter-
 525 preted in the graviton context. The most stringent limit is given by [32], where, under the assump-
 526 tion $\tilde{k} = 1$, a spin-2 bulk graviton with mass lower than 1.76 TeV is excluded (fig. 2.17).

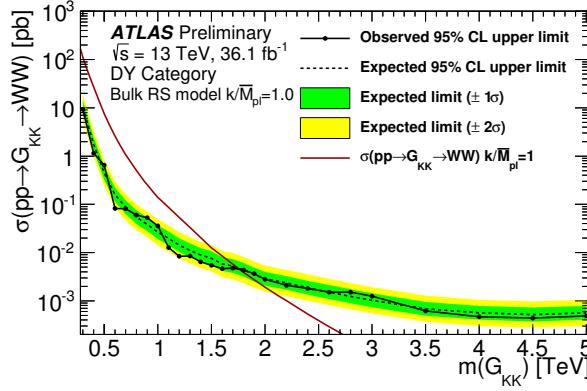


Figure 2.17: The observed and expected limits, with 68% and 95% uncertainty bands, on the product of the cross-section and branching fraction $\sigma \mathcal{B}(G \rightarrow ZZ)$ for a spin-2 bulk graviton, as a function of the reconstructed mass of the diboson resonance. The coloured lines show the theoretical predictions for $\tilde{k} = 1$.

The Large Hadron Collider and the CMS experiment

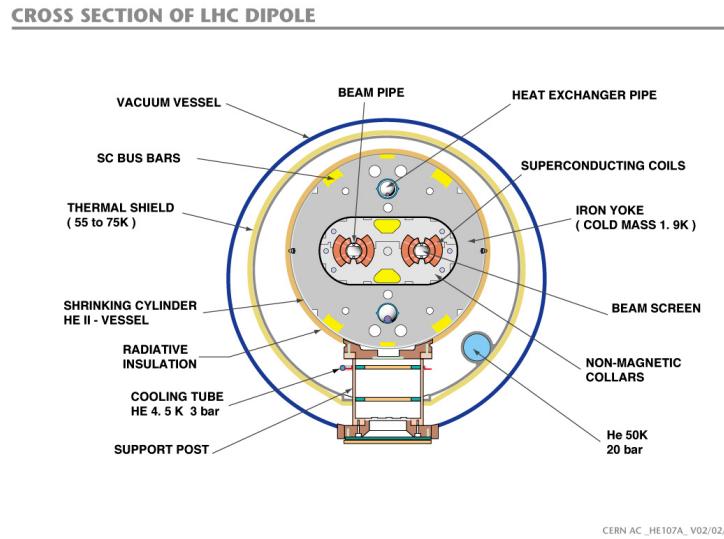
530 3.1 The Large Hadron Collider

531 The Large Hadron Collider (LHC) [46] is a 27 km ring structure designed for the acceleration and col-
 532 lision of protons and heavy ions. It is situated approximately 100 m underground, between France
 533 and Switzerland, in the Geneva area, and it is part of the CERN research facilities. In order to re-
 534 duce the cost of the project, approved in 1996, the LHC has been designed to fit the pre-existing
 535 underground tunnel of the Large Electron-Positron collider (LEP), built to accelerate electrons and
 536 positrons and operating until the year 2000.

537 Moving from an electron-positron collider to an hadron collider allowed to reach higher energies
 538 in the center-of-mass frame, since the synchrotron radiation loss is inversely proportional to the
 539 fourth power of the mass of the particle involved: hence, the radiation is reduced by a factor m_p/m_e
 540 $\sim 10^3$. The choice of a proton-proton collider was driven also by the possibility to collect higher lu-
 541 minosities (and hence more statistics) with regards to, for example, a proton-antiproton collider,
 542 like Tevatron at Fermilab, in the USA.

543 In the LHC two identical beam pipes are designed to let protons circulate in opposite directions, in
 544 ultra-high vacuum conditions (10^{-11} – 10^{-10} mbar), to avoid spurious collisions with gas molecules.
 545 Given the reduced diameter of the tunnel (4 m), the two proton beams are magnetically coupled.
 546 The collider is composed by 8 arc sections (48 km) driving protons around the ring, and straight
 547 sections (6 km) where beam control systems and detectors are inserted. Proton beams collide in
 548 four interaction points, where the main LHC experiments are installed: ALICE, ATLAS, CMS, LHCb.

549 In fig. 3.1, a slice of the arc section is displayed. Around the beam pipes, two superconducting
 550 magnetic dipoles are located: they generate vertical magnetic fields in opposite directions. The
 551 superconducting coils are made of niobium-titanium, materials that are superconducting at very
 552 low temperature. At the LHC, they are kept at a temperature of 1.9 K (-271.3°C) by a closed liq-
 553 uid helium circuit. A current of 11850 A flows through the magnets, without any energy loss due
 554 to electrical resistance, generating a magnetic field of 8.33 T. Magnets of higher order in multipole
 555 expansion (quadrupoles, sextupoles, octupoles, etc.) are employed to optimize the proton trajec-
 556 tories; in particular, quadrupoles allow to focus and squeeze the beams. Along the LHC ring there are
 557 9593 magnets; 1232 are dipoles, 392 are quadrupoles.



CERN AC _HE107A_ V02/02/98

Figure 3.1: Section of the LHC dipole magnet structure.

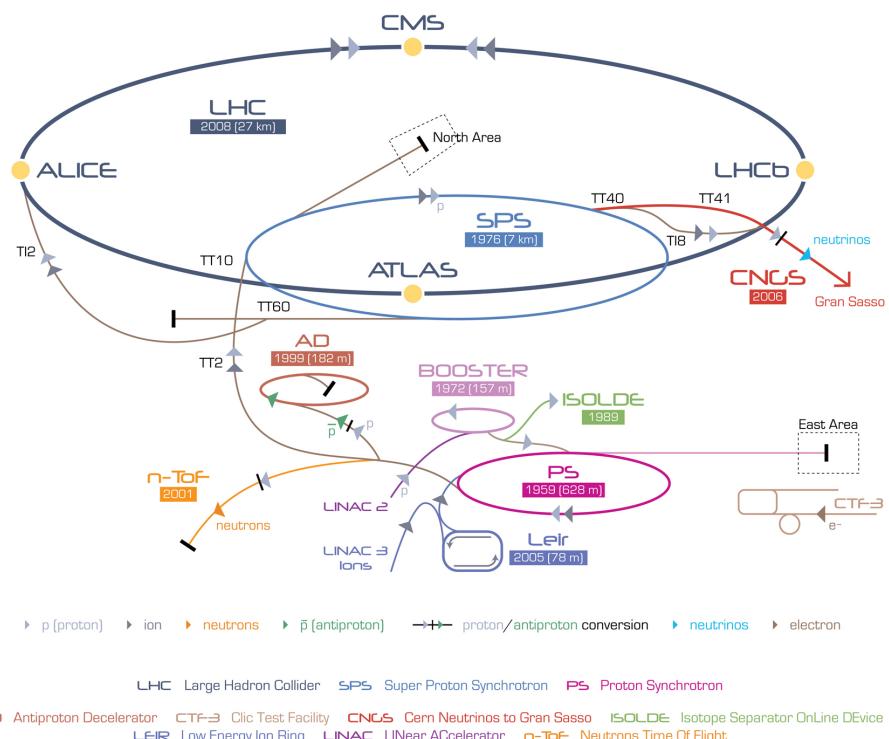


Figure 3.2: The CERN accelerator complex.

558 The LHC represents the final step of the CERN accelerator complex, shown in fig. 3.2. Protons are
 559 extracted from hydrogen atoms and inserted in the linear accelerator Linac2, that brings them to an
 560 energy of 50 MeV. They circulate around a little synchrotron, the Proton Synchrotron Booster, reach-
 561 ing an energy of 1.4 GeV, and then in the Proton Synchrotron (PS), where their energy is increased to
 562 25 GeV. The second to last step is the Super Proton Synchrotron, SPS, accelerating protons up to 450
 563 GeV. They are finally injected in the Large Hadron Collider, where sixteen radiofrequency cavities
 564 (RF) accelerate protons inside each beam up to an energy of 6.5 TeV, corresponding to a center-of-

3.1 The Large Hadron Collider

mass energy of 13 TeV when colliding. The RF cavities provide an accelerating electromagnetic field up to 5 MV/m (maximum voltage of 2 MV), that oscillates with a frequency of 400 MHz. Like the magnets, the cavities are kept at low temperature (4.5 K, or -268.7°C) in order to allow superconducting conditions. The maximum beam energy can be reached in 15 minutes. After several hours of collisions (~ 10 hours), the quality of the beams deteriorates and they are extracted from the machine and dumped.

Protons circulate inside the LHC ring in bunches of $\sim 10^{11}$ particles each, 80 mm long. Focusing magnets allow to reduce the bunch diameter down to 16 μm . Different bunches are separated by 25 ns (or, ~ 7.5 m), corresponding to a bunch collision frequency of 40 MHz and an instantaneous (peak) luminosity (defined in eq. 3.1) of $1.2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. Given the structure of the beams, at every bunch crossing many protons interact simultaneously: this phenomenon is called pile-up. The designed maximum number of bunches per fill is 2808.

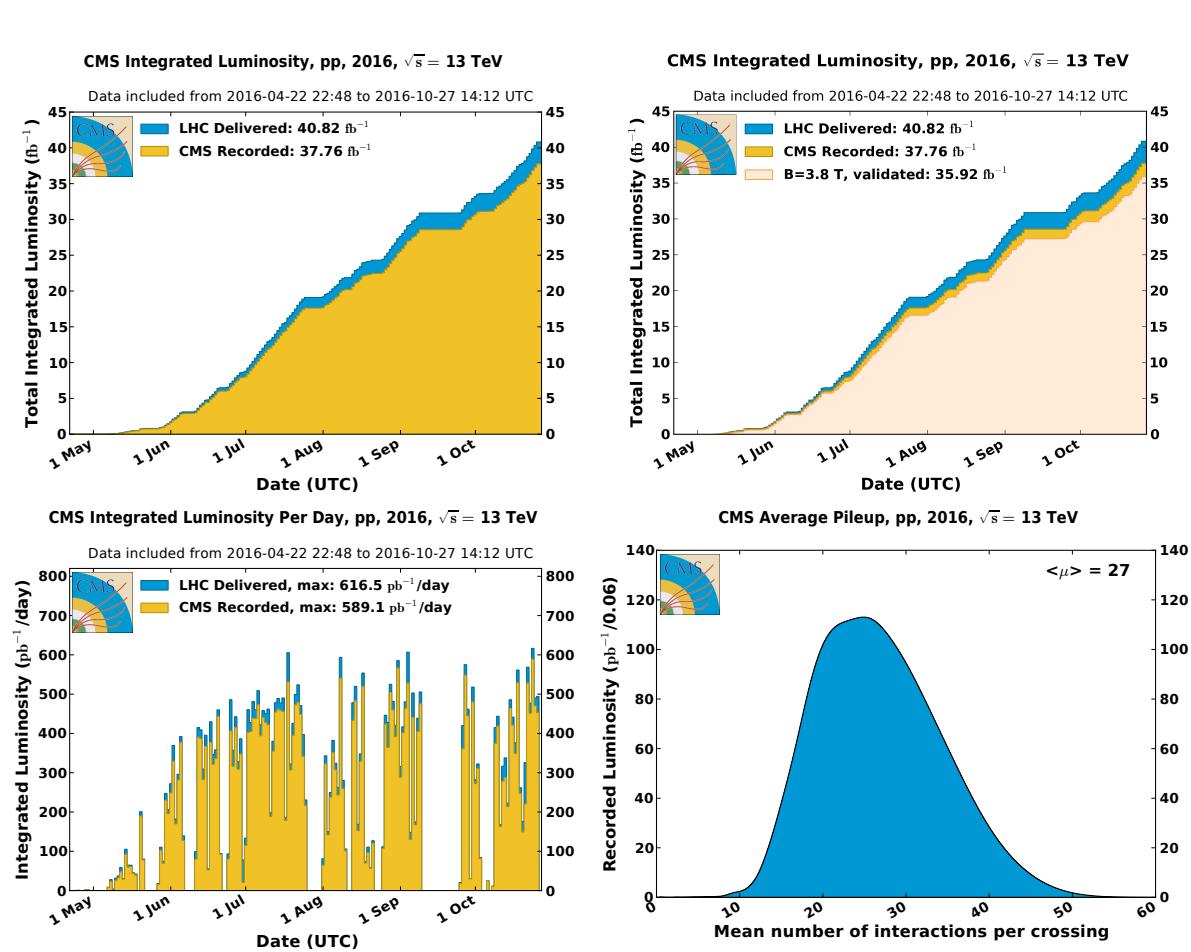


Figure 3.3: Luminosity in 2016 LHC data. Top-left plot: the cumulative integrated luminosity delivered by the LHC (in blue) and recorded by the CMS detector (in orange), as a function of the data-taking period. Top-right plot: data recorded by CMS and declared as optimal for the physics analyses (in light orange), corresponding to a total integrated luminosity of 35.9 fb^{-1} . Bottom-left plot: maximum integrated luminosity per day. Bottom-right plot: number of proton interactions per bunch crossing (pile-up).

The main parameters characterizing an hadronic collider are the center-of-mass energy, correspond-

580 ing to the sum of the energies of the beams, and the instantaneous luminosity, that describes the
 581 frequency of the interactions among the bunches in the beams. If the bunches in the first beam
 582 contain n_1 protons, the bunches in the second beam contain n_2 protons, the colliding area is Σ , and
 583 the frequency of complete turns around the ring is f , the instantaneous luminosity $\mathcal{L}_{\text{inst}}$ is:

$$\mathcal{L}_{\text{inst}} = f \frac{n_1 n_2}{\Sigma}. \quad (3.1)$$

584 If a generic physics process i has a cross-section of σ_i , the interaction rate R_i is:

$$R_i = \frac{dN_i}{dt} = \sigma_i \mathcal{L}_{\text{inst}}; \quad (3.2)$$

585 the number of events N_i recorded in the time interval $(0, \tau)$ is obtained from the integrated lumi-
 586 nosity $\mathcal{L} = \int_0^\tau \mathcal{L}_{\text{inst}} dt$:

$$N_i = \sigma_i \int_0^\tau \mathcal{L}_{\text{inst}} dt. \quad (3.3)$$

587 In fig. 3.3, a summary of the luminosity measurements in 2016 data is presented. The luminosity
 588 delivered by the LHC is represented in blue, the luminosity recorded by the CMS is displayed in
 589 orange. The mean number of interactions per bunch crossing (pile-up) is presented as well. The
 590 average number of interactions per collision is 27, the maximum is generally around 50 (in fig. 3.4,
 591 a record of 78 pile-up collisions was detected).

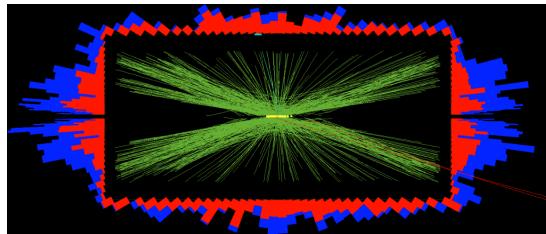


Figure 3.4: CMS collision event, where a record of 78 interactions per single bunch crossing were taking place simultaneously.

592 3.1.1 Proton-proton interactions

593 Proton-proton collisions allow to reach high energies and luminosities, but the drawback is the
 594 complexity of the events when compared to electron-positron collisions: not only because of the
 595 increasing backgrounds due to strong interaction among partons, but also because the momenta of
 596 the proton partons taking part in the interaction are unknown; not to mention the problem of disen-
 597 tangling the tracks of the particles coming from the interesting hard interactions from the spectator
 598 pile-up interactions.

599 The majority of the LHC events is represented by soft interactions, with low transverse momen-
 600 tum transfer, namely elastic and diffractive scatterings. In the so-called hard interactions, on the
 601 other hand, the transferred momentum among particles is high, allowing to produce massive reso-
 602 nant phenomena. These events manifest in peculiar final state signatures that can be distinguished
 603 from the soft background interactions.

604 At high momentum transfer (perturbative regime), a proton can be described as a collection of par-
 605 tons, each bringing a fraction x of the initial beam momentum, whose distributions are described
 606 by the parton distribution functions (PDF), $f(x, Q^2)$, as a function of the Bjorken's variable x and of

3.2 The CMS detector

the momentum transfer Q^2 . At very high center-of-mass energies (13 TeV), the proton mass can be neglected; the available energy in the parton 1 – parton 2 scattering is unknown, $\sqrt{x_1 x_2 s}$. The total cross-section of any interaction is given by:

$$\sigma = \int dx_1 f_1(x_1, Q^2) \int dx_2 f_2(x_2, Q^2) \sigma_{12}(x_1 p_1, x_2 p_2, Q^2), \quad (3.4)$$

where σ_{12} is the cross-section at parton level, and f_1, f_2 are the parton PDFs. In fig. 3.5, parton cross-sections of the main standard model processes are displayed, as a function of the center-of-mass energy.

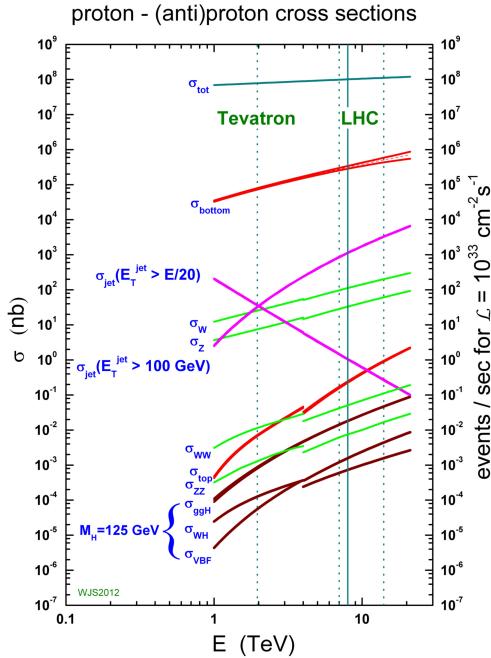


Figure 3.5: Cross-sections and number of expected events in proton-proton collisions, as a function of the center-of-mass energy. Rare phenomena, such as the Higgs boson production, can be observed at the LHC.

3.2 The CMS detector

The Compact Muon Solenoid (CMS) is a multi-purpose detector built in the LHC ring. It is situated in a cavern 100 m underground, near Cessy, in France. It is a cylinder 22 m long, with a diameter of 15 m, and a weight of 12500 tons. Its physics programme includes the search for the Higgs boson (discovered in 2012), precision measurements of the standard model parameters and rare decays (physics of bottom quark), and search for new physics beyond the SM (SUSY, exotic phenomena, dark matter, extra dimensions).

The CMS detector is structured in many layers of sub-detectors, giving different responses depending on the nature and the momentum of the particles passing through. The inner detectors have been finely segmented in order to afford the high radiation levels and particle multiplicity at the interaction point, so that the reduced occupancy of each layer allows to measure and distinguish precisely the primary vertices of the hard interactions from the pile-up events. A very accurate time

625 resolution is necessary to synchronize all the subsystems together.

626

CMS Detector

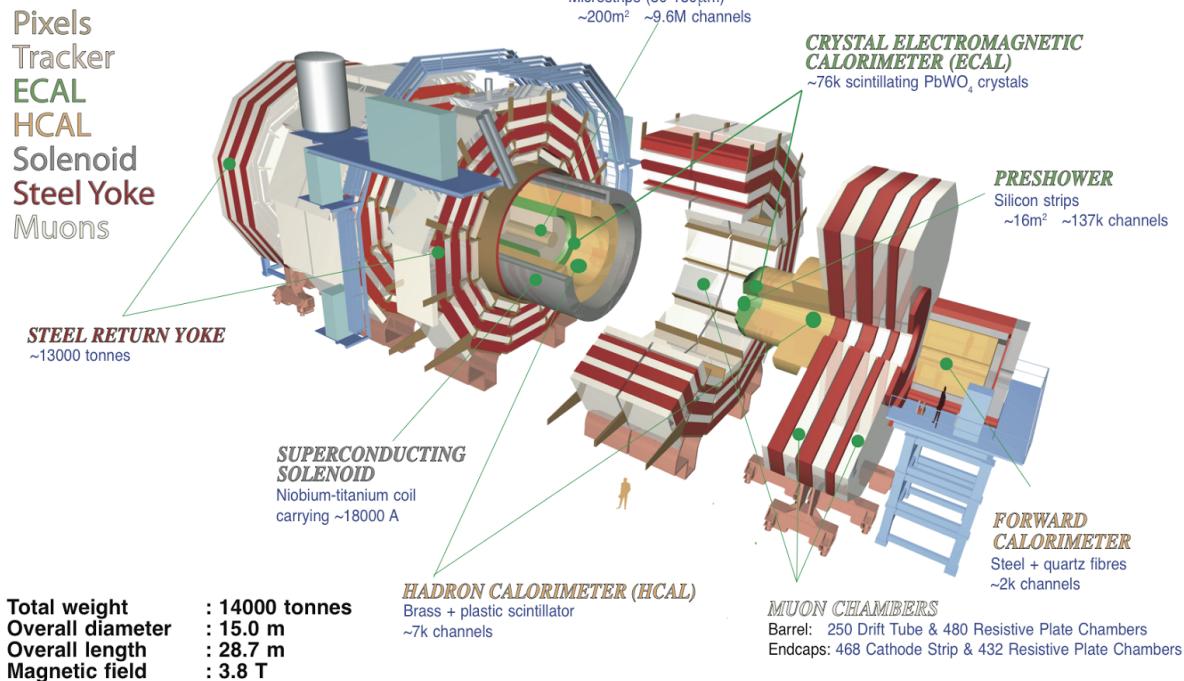


Figure 3.6: The CMS experiment.

627 Fig. 3.6 shows a sketch of the CMS detector. It is longitudinally segmented in the barrel region and
 628 two endcaps. In the forward region (over the endcaps), where the beam radiation is very intense,
 629 additional calorimeters have been placed. In fig. 3.7, the mean path of a specific particle through
 630 the sub-detectors is represented, depending on its flavour.
 631 A detailed description of the CMS detector can be found in [47].

632 3.2.1 The coordinate system

633 The CMS coordinate system is depicted in fig. 3.8. x and y are the coordinates in the transverse
 634 plane, z is the longitudinal coordinate. The x axis points at the center of the LHC ring, the y axis
 635 points upward, the z axis is along the beam direction. The azimuthal angle φ lies in the transverse
 636 plane, and it is measured starting from the x axis; the radial coordinate is r . The polar angle θ lies in
 637 the plane (r, z). The transverse component of the 3-momentum, \vec{p}_T , is orthogonal to the beam axis
 638 and lies in the plane (x, y). The transverse energy is defined as the magnitude of \vec{p}_T : $E_T = E \sin \theta$.
 639 Two other commonly used variables are the rapidity, \mathcal{Y} , and pseudorapidity, η , defined as functions
 640 of the particle energy E , the longitudinal component of the momentum p_z and the 3-momentum
 641 modulus:

$$\begin{aligned}\mathcal{Y} &= \frac{1}{2} \log \frac{E + p_z}{E - p_z} \\ \eta &= \frac{1}{2} \log \frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} = -\log \tan \frac{\theta}{2}\end{aligned}\tag{3.5}$$

3.2 The CMS detector

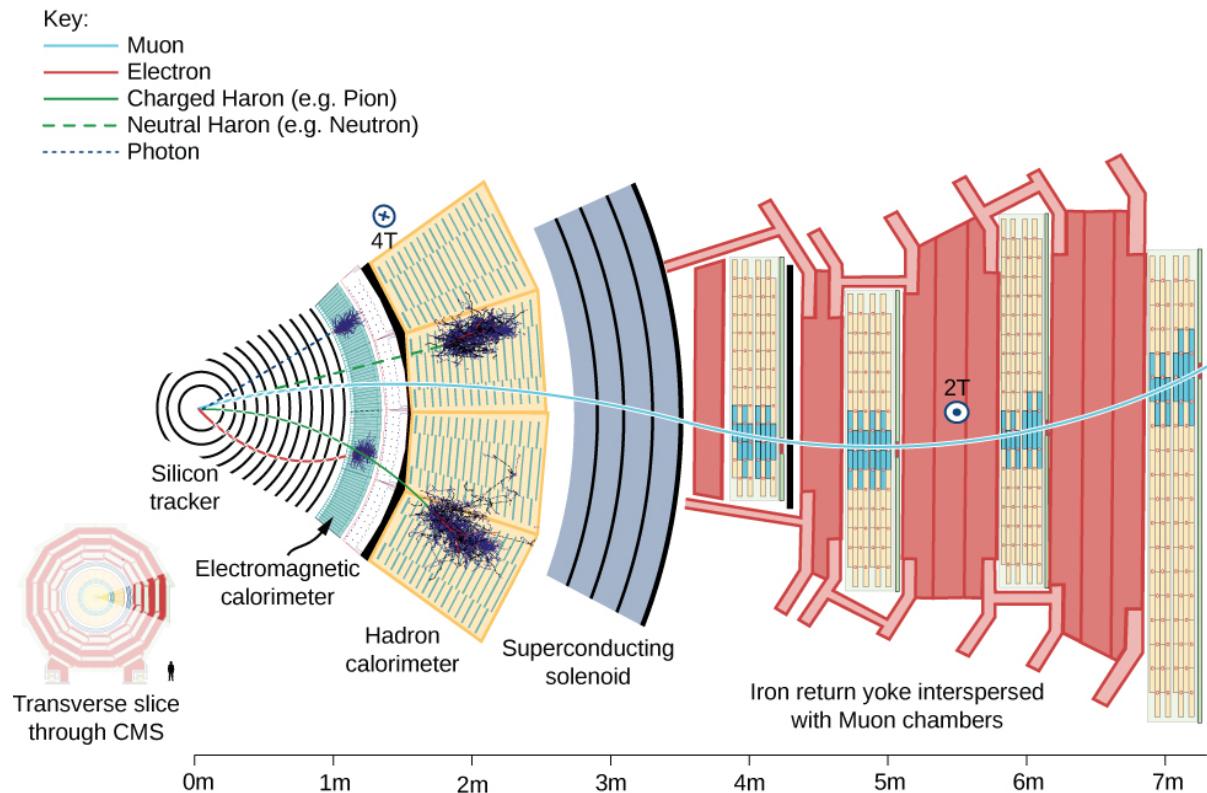


Figure 3.7: Mean path of a particle through the CMS detector. A muon, in light blue, passes through with a bended trajectory, depending on its momentum and charge, triggering signals in all the subsystems. An electron, in red, leaves a track in the silicon tracker and is absorbed by the electromagnetic calorimeter. A neutral or charged hadron, in green, stops inside the hadronic calorimeter. A photon, dotted blue line, showers in the electromagnetic calorimeter, without leaving any track in the silicon detector.

- 642 When the considered particle is produced in the forward region, hence at $\theta = 0$, it means that
 643 $\eta \rightarrow \infty$. When the particle is produced in the transverse plane, hence $\theta = \pi/2$, $\eta = 0$. At high
 644 energies, when the masses can be neglected, rapidity and pseudorapidity coincide; these variables
 645 are largely used at colliders because they are invariant under Lorentz boosts along the beam direc-
 646 tion.

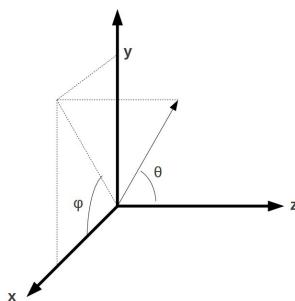


Figure 3.8: The CMS coordinate system.

647 3.2.2 The magnet

648 The CMS superconducting magnet is an hollow cylinder (13 m long, 6 m of diameter, shown in
 649 fig. 3.9). An electrical current of 19 kA flows through the niobium and titanium fibers that consti-
 650 tute the solenoid, providing a maximum magnetic field of 3.8 T and storing a maximum energy of
 651 2.6 GJ. Superconducting conditions are mantained by a liquid helium cooling system, keeping the
 652 solenoid temperature at 4.5 K. In order to avoid stray fields, the magnetic field lines are closed by
 653 the return yoke, composed by 10 ktons of magnetized iron blocks, located in the outer part of CMS
 654 and alternated to the muon chambers. The homogeneus magnetic field inside the detector bends
 655 the trajectories of the charged particles, allowing the measurement of their momenta p , given the
 656 relation with the magnetic field strength B and the radial coordinate r of the trajectory:

$$p[\text{GeV}] = 0.3 \times B[\text{T}] \times r[\text{m}]. \quad (3.6)$$

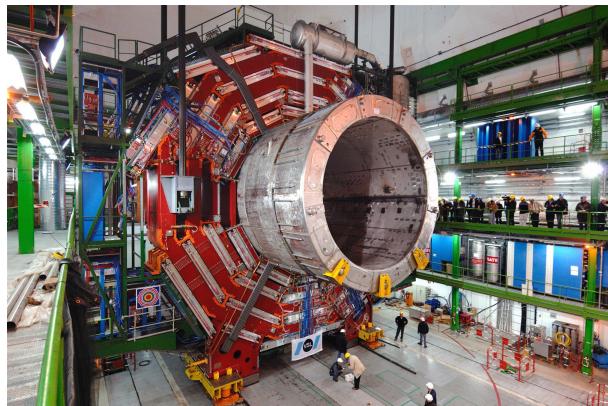


Figure 3.9: Installation of the superconducting solenoid in the CMS cavern.

657 3.2.3 The tracking system

658 The CMS tracking system [48, 49] is composed by a cylinder of silicon detectors (2.5 m of diameter
 659 and 5.8 m of length). Their design guarantees a precise reconstruction of the tracks left by charged
 660 particles and of the interaction vertices, a fundamental tool to identify heavy quarks (charm, beauty)
 661 and leptons (taus). Tracker detectors cover a pseudorapidity region of $|\eta| < 2.5$ and have an active
 662 area of 210 m^2 . The two sub-detectors of the tracking system are the pixel detector, installed close
 663 to the interaction point, and the strip detector, covering a radius of $0.2 - 1.2 \text{ m}$. The high granularity
 664 of the pixels and strips allows to keep the occupancy at acceptable levels, given the high multiplicity
 665 of the tracks ($\sim 1 \text{ MHz/mm}^2$). The silicon detectors and the electronic cables are cooled down to a
 666 temperature of $\sim 10^\circ \text{ C}$. The structure of the tracking system is shown in fig. 3.10.

667 3.2.3.1 The pixel detector

668 The pixel detector is composed by 66 millions of silicon cells, whose dimensions are $100 \times 150 \mu\text{m}^2$,
 669 $285 \mu\text{m}$ of thickness, placed in 1440 modules. Silicon cells are set in three layers in the barrel re-
 670 gion and in two disks at each endcap. Barrel modules are disposed parallel to the magnetic filed,
 671 whilst at the endcap they are tilted by about 20° . Pixels allow a spatial resolution of $10 \mu\text{m}$ in the
 672 transverse plane, and of $\sim 20 \mu\text{m}$ along the longitudinal coordinate. Their reduced size guarantees
 673 an occupancy of 10^{-4} per pixel at each bunch crossing, in high luminosity regime.

3.2 The CMS detector

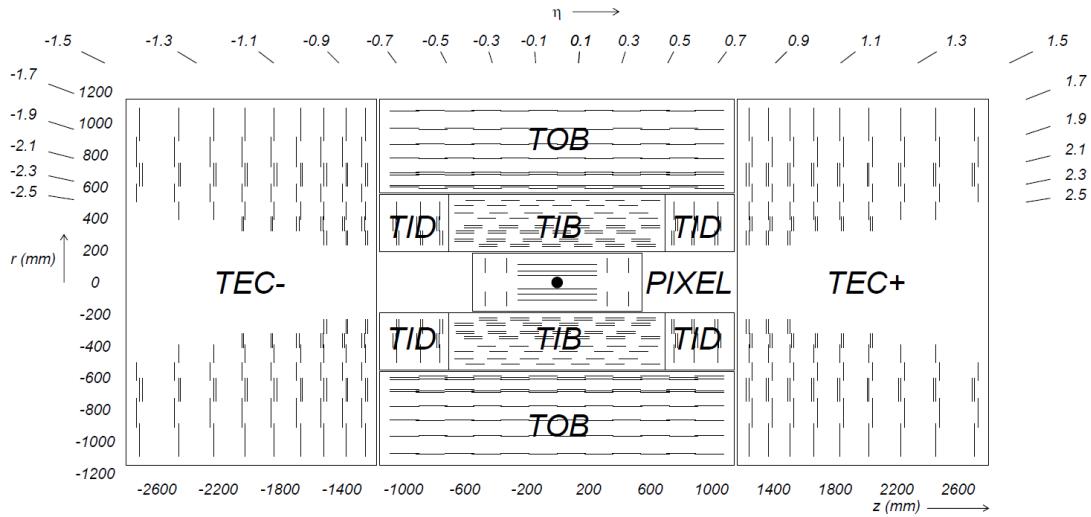


Figure 3.10: The CMS tracking system: the inner pixel detector, close to the interaction point, and the outer strip detector.

674 **3.2.3.2 The strip detector**

675 The strip system is divided in the four-layered tracker inner barrel (TIB), covering a region $20 < r <$
 676 55 cm with respect to the interaction point, the six-layered tracker outer barrel (TOB), located at
 677 $55 < r < 110$ cm, the three tracker inner disks (TID) and the nine tracker endcaps (TEC) at each
 678 cylinder base. Given the lower radiation level at higher radii (and hence a lower occupancy, around
 679 few percent), strips are bigger than pixels. Silicon strips in TIB and TID are $320 \mu\text{m}$ thick, 10 cm long,
 680 and with a pitch ranging from 80 to $120 \mu\text{m}$; strips in TOB and TEC are 25 cm long, with a different
 681 thickness ($320 \mu\text{m}$ for TID, $500 \mu\text{m}$ for TEC) and pitch (97–184 μm). There are 15148 strip modules,
 682 and 9.3 million readout channels. The strip spatial resolution is about $20 – 50 \mu\text{m}$ in the transverse
 683 plane and about $200 – 500 \mu\text{m}$ along the longitudinal coordinate.

684 **3.2.4 The electromagnetic calorimeter**

685 The CMS electromagnetic calorimeter (ECAL, shown in fig. 3.11) [50] is a homogeneous detector
 686 composed by lead tungstate (PbWO_4) scintillating crystals, designed to measure the energy de-
 687 posits of photons and electrons through their electromagnetic showers. PbWO_4 is transparent and
 688 dense ($8.3 \text{ gr}/\text{cm}^3$); it has a fast time response (the 85% of the scintillating light is emitted at every
 689 bunch crossing), high scintillating efficiency and radiation resistance; it has a radiation length of
 690 $X_0 = 0.89 \text{ cm}$ and a Molière radius of 2.19 cm. The ECAL is divided in the barrel region ($\eta < 1.479$,
 691 at a radius of 1.3 m) and the endcaps ($1.479 < \eta < 3$). The 61200 crystals employed in the barrel
 692 region, whose size is $(22 \times 22) \text{ mm}^2 \times 23 \text{ cm}$, have a radiation length of $25.8X_0$; the 7324 crystals in
 693 the endcaps, of size $28.6 \times 28.6 \text{ mm}^2 \times 22 \text{ cm}$, have a radiation length of $24.7X_0$. Before the endcaps,
 694 on each side, a pre-shower detector is installed: it is composed by two disks of lead absorber and
 695 two layers of silicon strips, of radiation lengths up to $3X_0$. The pre-shower calorimeter has been
 696 designed to distinguish the photons coming from the π^0 decay, from the photons produced in the
 697 rare Higgs decay $H \rightarrow \gamma\gamma$. The readout and amplification of the scintillating light, performed by
 698 avalanche photodiodes in the barrel and by vacuum phototriodes in the endcaps, requires a stable
 699 temperature of 18° C , maintained by a water cooling system.

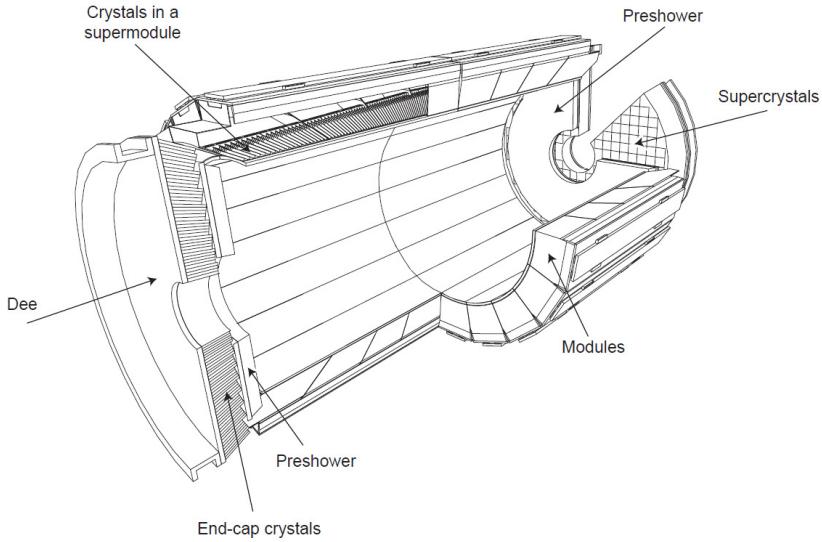


Figure 3.11: The CMS electromagnetic calorimeter.

700 The energy resolution of the calorimeter is parametrized as:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \quad (3.7)$$

701 where $S = 0.018 \text{ GeV}^{1/2}$ is the stochastic term, $N = 0.04 \text{ GeV}$ is related to noise contribution, and
702 $C = 0.005$ is a constant term depending on the calibration.

703 3.2.5 The hadronic calorimeter

704 The hadronic calorimeter (HCAL, displayed in fig. 3.12) [51] is a sampling calorimeter, composed of
705 brass and plastic scintillator layers. It has been designed to guarantee a good hermeticity, allowing
706 to perform a precise measurement of the missing transverse energy. It is located within the electro-
707 magnetic calorimeter and the solenoid, covering a region of $|\eta| < 1.3$ in the barrel, and $1.3 < |\eta| < 3$ in
708 the endcaps. Brass is non-magnetic and has short interaction length (16.4 cm): the 60 mm thick ab-
709 sorber layers used in the barrel reach 5.6 interaction lengths at $\eta = 0$ and 10.8 interaction lengths at
710 $\eta = 1.3$; the 80 mm thick layers in the endcaps reach 11 interaction lengths. An additional calorimet-
711 ric layer has been installed out of the solenoid, in order to reach 11.8 interaction lengths in the barrel
712 region. The scintillation light, typically in the blue-violet region of the electromagnetic spectrum, is
713 collected by wavelength-shifter fibers, translated and amplified by multi-channel hybrid photodi-
714 odes, proportionally to the magnitude of the energy deposits. An additional hadronic calorimeter
715 (HF) has been placed in the forward region, $3 < |\eta| < 5.2$, at 11.2 m from the interaction point. It has
716 been studied to afford the high levels of radiation: it is composed by 55 mm thick absorber layers
717 of stainless-steel, and quartz fibers, able to detect the Cherenkov scintillating light of the charged
718 particles of the hadronic showering. A longitudinally segmentation allow to distinguish hadronic
719 particles from electromagnetic components. The energy resolution of the hadronic calorimeter is:

$$\left(\frac{\sigma}{E}\right) \approx \frac{a}{\sqrt{E}} \oplus b\%, \quad (3.8)$$

720 where $a = 65\%$ in the barrel region, 85% in the endcaps, 100% in the forward region, and $b = 5\%$.

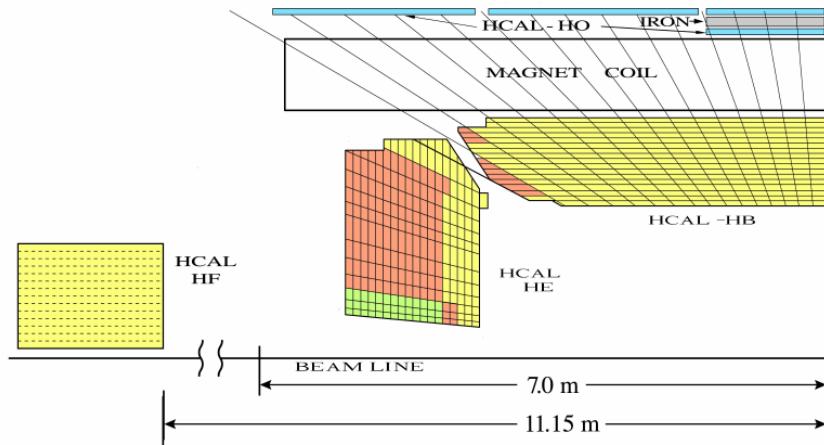


Figure 3.12: The CMS hadronic calorimeter.

3.2.6 The muon system

The outer system of the CMS experiment consists into gas detectors for identifying muons [52], that are located between the iron return yokes, designed to close the magnetic field generated by the solenoid. In the barrel region, where a smaller number of muons is expected and the magnetic field is less strong, Drift Tubes (DT) detectors are installed. In the endcaps, where the flux of particles is larger, Cathod Strip Chambers (CSC) are used, and disposed in three disks. CSCs are designed to allow faster responses, higher granularity and radiation resistance. Resistive Plate Chambers (RPC) are installed both in the barrel and in the endcaps as additional triggering system. The geometry of the muon system is shown in fig. 3.13; it consists of 250 DTs, 530 CSCs, 610 RPCs, and it covers a region $|\eta| < 2.4$.

3.2.6.1 The Drift Tubes

Drift Tube detectors cover a region of $|\eta| < 1.2$ and are arranged in four stations, segmented along the beam line in five wheels. The basic element of the detector is the cell, that has a size of $42 \times 13 \text{ mm}^2$. Each cell is filled with a gas mixture (85% argon, 15% CO_2), in which the process of ionization takes places; the ionization electrons drift from the $50 \mu\text{m}$ thick steel anodic wire, located in the center of the cell, towards the aluminium cathodic strips, located at its edge. Additional electrodes on the surface of the cells allows to shape the electric field, in order to make the drift speed of the electrons uniform: the muon position is then extrapolated from the measurement of the drift time. Every station is composed by three cells superlayers. In the inner and the outer superlayers, the cells are oriented such in a way that the anodic wire is located along the z axis, to measure the φ coordinate. In the intermediate superlayer, wires are parallel to the radial coordinate, hence they can measure the z position. The spatial resolution of the system is $100 \mu\text{m}$ in the (r, φ) plane, 1 mrad in the φ coordinate, and $150 \mu\text{m}$ in the longitudinal z coordinate.

3.2.6.2 The Cathode Strip Chambers

Cathode Strip Chambers cover a region of $0.9 < |\eta| < 2.4$, overlapping with the DTs in the pseudorapidity range $0.9 < |\eta| < 1.2$. The anodic wires inside each CSC are installed in six planes, with the aim of measuring the radial coordinate; the wire planes are perpendicularly crossed by cathodic strips, disposed along the radial direction to measure the φ coordinate. Ionization electrons produced

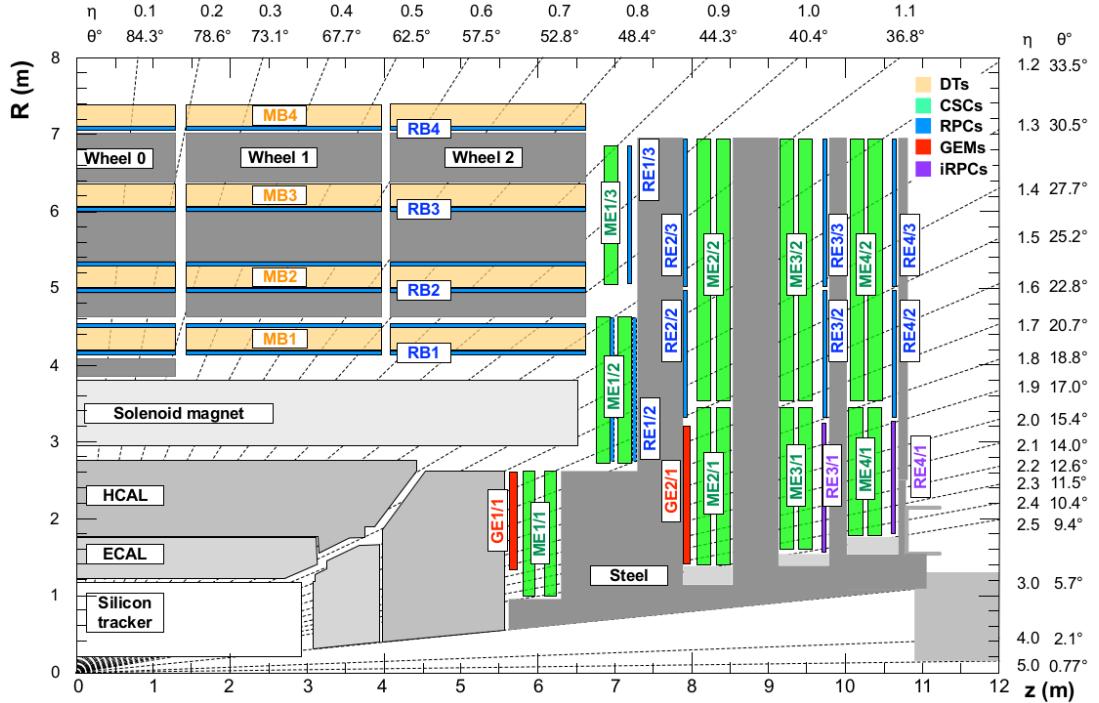


Figure 3.13: Section of CMS detector, in the plane (r, z), parallel to the beamline, that emphasizes the location of the muon detectors, in particular: Drift Tubes (DT, in yellow); Cathode Strip Chambers (CSC, in green); Resistive Plate Chambers (RPC, in blue).

749 by muons passing through the gas mixture in the chambers migrate from the anodes, inducing a
 750 charge distribution on the cathodes, from which the azimuthal coordinate can be reconstructed.
 751 The spatial resolution is $75 - 150 \mu\text{m}$ in the (r, φ) plane. CSCs are arranged in four disks and in three
 752 concentric rings.

753 3.2.6.3 The Resistive Plate Chambers

754 Resistive Plate Chambers are located both in the barrel (disposed in six layers) and in the endcap
 755 region (three layers), up to a pseudorapidity of $|\eta| < 1.6$. These gas detectors are charged at very
 756 high voltages, in order to work in the avalanche ionization mode. The plastic resistive plates are
 757 equipped with readout strips. The spatial resolution of the detector is low (1-2 cm), but the fast
 758 timing response (2-3 ns) and good time resolution (1 ns) allow to employ RPCs as an additional
 759 trigger system and to profit of a precise measurement of the bunch-crossing time.

760 3.2.7 The trigger system and data acquisition

761 The CMS trigger system [53] has been designed considering the high instantaneous luminosity, such
 762 that it can provide a fast response and it allows to reduce the nominal event rate of 40 MHz in pro-
 763 ton proton collision. The complexity of the CMS detector and the very high number of readout
 764 channels result into a huge amount of data per event, approaching the order of few MB per bunch
 765 crossing, hence 40 TB per second. The processes of handling and recording data are currently lim-
 766 ited by the employed technology to a frequency of ~ 100 Hz. Applying online selections to skim the
 767 events that are going to be written on tape, without rejecting interesting signals of hard processes
 768 and rare phenomena, is therefore a crucial and challenging point for every data analysis. Events are

3.2 The CMS detector

769 filtered by trigger selections at different levels: the Level-1 (L1) trigger is an hardware device, that
770 allows to reduce the event rate from 40 MHz to the order of 100 kHz; the High Level Trigger (HLT)
771 is a set of software algorithms that skims the event rate down to few hundred Hz. Once the trigger
772 decisions are taken, the final events are handled by the Data Acquisition System (DAQ), that collects
773 the informations coming from the sub-detectors and sends them to the storage unities.

774 **3.2.7.1 The Level-1 trigger**

775 The L1 trigger is an hardware device composed by customized electronics, and it accesses the in-
776 formations coming from the calorimeters and the muon system, while the tracker is not considered
777 given the excessively large bandwidth needed by its readout channels. The L1 trigger performs a
778 first raw local reconstruction of each object, called “trigger primitive”. The L1 trigger is composed
779 by three subsystems: the calorimeter trigger, the muon trigger (divided in three independent sub-
780 subsystems for each muon sub-detector, DTs, RPCs and CSCs), and the global trigger, that combines
781 the informations of the former subsystems. The best quality trigger primitives reconstructed by the
782 calorimeters and muon detectors (namely, roughly reconstructed electrons, photons, muons, jets,
783 jets coming from the hadronic decays of tau leptons, and missing transverse energy) are handled
784 by the global trigger, which takes the decision of discarding or keeping the event every 3.2 μ s. The
785 simplest trigger selections require the presence of a single object, whose energy or transverse mo-
786 mentum is higher than a certain threshold; more complicated triggers involve multiple objects or
787 geometrical selections, that can be performed in parallel up to 128 simultaneous requirements.

788 **3.2.7.2 The High Level Trigger**

789 The HLT skims the L1 output rate down to few hundreds Hz by applying a set of algorithms, im-
790 plemented in the same software used for the offline analyses, consisting in an event reconstruction
791 performed by exploiting the whole informations coming from all sub-detectors. The computing
792 time is still a crucial factor, hence selections applied to HLT physics objects are generally less accu-
793 rate than those of the offline analyses; furthermore, HLT can discard the event even before its full
794 reconstruction (*i.e.* by looking only at certain region of the detectors). Events filtered by the HLT
795 decisions are assigned to precise trigger paths and recorded in different categories of datasets.

796 **3.2.7.3 Data acquisition, computing and storage**

797 The DAQ system deals with the storage, transfer and handling of the data collected by CMS; it also
798 supports and stores the data simulations and calibrations of the sub-detectors. The CMS compu-
799 tational resources are located in worldwide distributed data nodes, called Tiers. The CMS software
800 (CMSSW) is based on an object oriented architecture (mainly C++). The basic unity of every data,
801 both real and simulated ones, is the Event, that could contain very rough informations (RAW data
802 format) or higher level refined objects (AOD, Analysis Object Data) where all the calibrations and
803 corrections needed to properly deal with the final physics objects are already in place. Data are
804 handled by C++ or python modules, and the outputs are written in ROOT [54] files.

805 **3.2.8 Particle Flow event reconstruction**

806 The particle flow (PF) algorithm [55] aims at identifying and reconstructing each particle produced
807 by the proton-proton collisions, combining the informations coming from all the CMS sub-detectors.
808 It is particularly suitable to improve the reconstruction of jets, missing transverse momentum (used
809 to identify neutrinos) and hadronically decaying tau leptons.
810 The association of the informations is performed at different stages. The reconstruction of the

811 charged particles in the silicon detector is executed with an iterative algorithm, and the recon-
 812 structed object is called a tracker track. Then, a clustering algorithm is performed to collect and
 813 combine the energy deposits in the calorimeters, in such a way to distinguish neutral from charged
 814 particles, reconstruct their directions, and improve the energy measurement of the very energetic
 815 charged particles, whose tracks are less bended by the magnet and hence less precisely determined.
 816 The last informations are provided by the hits collected in the muon system. The three sets of re-
 817 constructions are then combined with a link algorithm, that aims at associating tracker tracks to
 818 calorimeter clusters and muon hits with geometrical criteria. A track in the silicon detector is linked
 819 to a calorimeter cluster if the extrapolated position lies in the cluster itself. Similarly, clusters in dif-
 820 ferent calorimeters are linked when the position calculated in the more granular calorimeter (*i.e.*
 821 ECAL) lies in the envelope of the clusters in the less granular calorimeter (*i.e.* HCAL). The decision
 822 of linking a tracker track to a muon track is based on the χ^2 of a global fit between the two tracks.
 823 The particle flow algorithm then interprets the collected and linked informations as particles. Muons
 824 are identified by the combination of a track in the silicon detectors and a track in the muon cham-
 825 bers. Photons are determined directly by ECAL clusters. Electrons energies and positions are mea-
 826 sured by ECAL clusters, linked to a corresponding tracker track, and considering all the energy clus-
 827 ters produced by the bremsstrahlung photons radiated while interacting with the detector mate-
 828 rial. The hadrons are identified by the tracks (if charged) linked to the corresponding ECAL and
 829 HCAL clusters. The hadron energy resolution, 10% at 100 GeV combining ECAL and HCAL, is such
 830 that neutral hadrons can be distinguished as an energy calorimetric excess when overlapped by a
 831 charged hadron occupying the same calorimetric towers. Finally, the missing transverse momen-
 832 tum is defined as the negative sum of the transverse momenta of all the particles identified by the
 833 PF algorithm.

834

835 **3.2.9 Physics objects**

836 **3.2.9.1 Track reconstruction**

837 The reconstruction of the trajectories of the charged particles passing through the CMS detector is
 838 performed by multiple iterations of the Combined Track Finder (CTF) algorithm, that is based on a
 839 Kalman filter approach [56]; given the high multiplicity of particles produced at each bunch cross-
 840 ing and the multiple scatterings in the detector materials, tracking represents a challenging task.
 841 The CTF algorithm builds a track starting from the so-called seeds, namely triplets of hits collected
 842 in the pixel detector inner layers, or couples of hits if the track originates from the interaction point.
 843 The initial guess of the track given by the seeds is then extrapolated to the outer layers: if other hits
 844 are found to be compatible with the trajectory hypothesis (χ^2 -based hypothesis test), they are added
 845 to the track. Once the outer layers are reached, another reconstruction is performed backward, in
 846 order to clean the track from spurious hits and enhance the tracking efficiency. The final collected
 847 hits are re-fitted with Kalman filter and more precise algorithms, in order to improve the quality of
 848 the measurement. If two tracks share more than a half of their hits, the worst quality track is re-
 849 jected. The track reconstruction efficiency for particles with $p_T > 0.9$ GeV is 94% in the barrel and
 850 85% in the endcap region [49].

851

852 **3.2.9.2 Vertices reconstruction**

853 The reconstruction of the vertices at each bunch crossing is performed in steps. Primary vertices
 854 originate from the proton-proton collisions, whilst secondary vertices are due to long-lived particles

(heavy quarks and τ leptons). The starting point of the procedure is clustering the reconstructed tracks originating from the primary vertex; the decision is taken by the deterministic annealing algorithm [57], that uses the longitudinal impact parameters of each track as inputs. The algorithm allows to distinguish vertices further than 1 mm. The second step is run by the adaptive vertex fitter [58], that measures the position of the vertex for the chosen set of tracks. The algorithm is based on an iterative re-weighted Kalman filter, that down-weights the wrongly associated tracks not compatible with the considered vertex. The primary vertex is selected as the vertex where the sum of the p_T^2 of the associated tracks is the largest. The spatial resolution on the vertex position is 10-40 μm in the (r, φ) plane, and 15-50 μm in the longitudinal coordinate.

3.2.9.3 Electrons and photons reconstruction

Electrons are reconstructed [59] combining a track with the energy deposits clustered in the ECAL, due to the showering of the electron through the detector and the emission of bremsstrahlung photons. The combination can proceed both from the silicon detector in the outgoing direction, and in the opposite way: the tracker seeding as starting point is suitable for low energy electrons, whose trajectories are more bended (smaller curvature radii) and hence more accurately measured by the tracker system; the grouping of ECAL clusters (called superclusters) followed by a consecutive track extrapolation, performed by taking into account the electron interaction with the detector material, is more efficient in case of high energetic electrons, due to the higher resolution of the ECAL scintillating crystals. A Gaussian-sum filter algorithm (GSF) [60] allows to properly take into account the effects of the bremsstrahlung radiation, that is distributed not as a single Gaussian (standard Kalman filters) but rather as a sum of Gaussian functions.

The identification of an electron relies on three groups of variables: observables built by combining measurements performed in the silicon detectors and in the calorimeter; purely calorimetric observables; purely tracking informations. Different selections are used for electron candidates found in the barrel or in the endcaps, and they can vary from loose criteria (high detection efficiency but less purity, namely more contamination from objects mis-identified as electrons) to tight criteria. Data and Monte Carlo simulations reproducing Z , Υ and J/Ψ decays into e^+e^- are used to study the optimal working points, each one targetting at a different purity.

The electron energy is determined correcting the raw energy measurement of the ECAL superclusters by taking into account the effects of the losses due to radiation or gaps between the calorimeter modules, and the pile-up contribution. The electron momentum resolution has been measured in $Z \rightarrow e^+e^-$ decays in Run 1 LHC data, and it varies from 1.7% to 4.5% depending on the pseudorapidity range [61]. The electron isolation variable is defined as the p_T sum of the charged and neutral particles lying in a cone of $\Delta R = 0.3$ around the electron trajectory, divided by the transverse momentum of the electron itself:

$$I_{\Delta R=0.3}^e = \frac{\sum_{\text{char. hadrons}} p_T + \max[0, \sum_{\text{neut. hadrons}} p_T + \sum_{\text{photons}} p_T - 0.5 \sum_{\text{pile-up char. hadrons}} p_T]}{p_T^e}; \quad (3.9)$$

the contribution of the pile-up charged particles is removed. The isolation variable is used to distinguish electrons coming from the leptonic decays of electroweak bosons (low $I_{\Delta R=0.3}^e$) from electrons coming from the decays of heavy fermions, in which case they are more likely produced in association with light flavour jets and hence topologically close to calorimetric deposits due to hadrons (high $I_{\Delta R=0.3}^e$).

Photons are reconstructed with the ECAL clusters only. Given their importance in the discovery of the Higgs boson, dedicated studies have been performed both in data and in Monte Carlo simula-

897 tions reproducing the $H \rightarrow \gamma\gamma$ process. Particular care has been taken in the treatment of the photon
 898 conversions into electron-positron pairs while interacting with the tracker detector. Dedicated se-
 899 lections allow to define different photon identification working points. Similarly to the case of the
 900 electrons, the photon isolation variable can be defined. The photon energy resolution varies from
 901 1% to 3%, depending on the η range [62].

902 **3.2.9.4 Muon reconstruction**

903 A muon candidate can be built exploiting the hits collected in the silicon tracker (track) and in the
 904 muon system (standalone muon) [63]. Each muon sub-detector (DTs, RPCs and CSCs) performs a
 905 local reconstruction of the particle candidate; the informations from the three muon chambers are
 906 combined with a Kalman filter approach.

907 Three different strategies are adopted to define a muon candidate in the CMS detector. A stan-
 908 dalone muon is reconstructed by using only the local reconstruction in the muon chambers. A
 909 tracker muon is built starting from a track in the silicon detector, that is extrapolated up to the muon
 910 chambers, taking into account the multiple scattering and the energy loss through the material.
 911 The tracker muon is defined if at least one segment, *i.e.* a short track built with CSCs or DTs hits,
 912 is matched to the starting track. This technique is the most efficient for the reconstruction of low
 913 energetic muons. A global muon is built starting from a standalone muon, and then its trajectory is
 914 extrapolated towards the inner layers of the silicon detector and eventually matched to a track; this
 915 approach is suitable for high energetic muons ($p_T > 200$ GeV).

916 Different algorithms are used to assign a momentum to the muon candidate, in order to mitigate
 917 the effects of bremsstrahlung, that becomes significant when the muon approaches energies of the
 918 order of 1 TeV. The radiated photons generate spurious hits in the chambers and larger occupancy,
 919 significantly deteriorating the momentum measurement.

920 Starting from 2016 LHC Run, the muon reconstruction takes into account the alignment position
 921 errors, namely the uncertainties due to the position of the muon chambers with respect to the sil-
 922 icon detectors. The final resolution on the muon momentum measurement depends on the p_T
 923 and η of the candidate, and ranges from 1% for very low momenta, up to $\sim 7\%$ ($|\eta| < 0.9$) – 10%
 924 ($1.2 < |\eta| < 2.4$) [64].

925 The muon isolation $I_{\Delta R=0.4}^\mu$ is defined similarly to the electron isolation, but by taking into account
 926 a larger cone $\Delta R = 0.4$ around the muon direction.

927 **3.2.9.5 Jet reconstruction**

928 The nature of the strong interaction is such that coloured partons, namely quarks and gluons, are
 929 forced to aggregate to form a color-neutral hadron, in the process called hadronization. Therefore,
 930 partons cannot be observed as free particles in a detector, but rather as collimated jets of hadronic
 931 particles.

932 Jets are reconstructed starting by the PF candidates in the event. The charged hadron subtraction al-
 933 gorithm (CHS) removes candidates not associated to the primary vertex in order to suppress pile-up
 934 contributions [65]. The remaining particles are used as input to jet clustering algorithms to recon-
 935 struct particle flow jets. The jets are clustered using the FASTJET package [66] with the anti- k_T jet
 936 sequential clustering algorithm [67]. A sequential clustering algorithm is designed to be infrared
 937 and collinear safe, namely, if the final state particles undergo a soft emission or a collinear gluon
 938 splitting, the number and shapes of the jets should not change. The starting point of a sequen-
 939 tial clustering algorithm is the definition of the distances between two particles i and j , and the

3.2 The CMS detector

940 distance of a given particle i from the beam-spot B :

$$d_{ij} = \min(p_{T,i}^{2a} p_{T,j}^{2a}) \frac{R_{ij}^2}{R_0^2}, \\ d_{iB} = p_{T,i}^{2a}, \quad (3.10)$$

941 where $p_{T(i,j)}$ are the transverse momenta of the particles, $R_{ij}^2 = (\mathcal{Y}_i - \mathcal{Y}_j)^2 + (\varphi_i - \varphi_j)^2$ is the angular
942 distance between the particles, a is an exponent depending on the clustering algorithm chosen,
943 and R_0 is the clustering parameter. The algorithm then operates as follows:

- 944 • it computes all the possible combinations of distances d_{ij} and d_{iB} and it finds the minimum;
- 945 • if the minimum is d_{ij} , the four-momenta of the particles i and j are summed up in one can-
946 didate jet; i and j are removed from the list of available particles, the distances are updated,
947 and the algorithm proceeds to re-calculate all the possible remaining d_{ij} ;
- 948 • the clustering stops when the smallest quantity is d_{iB} : i particle is defined as one jet, and it is
949 removed from the list of particles;
- 950 • this process is repeated until all the particles are assigned to a jet, that must be separated from
951 another jet at least by a distance $R_{ij} > R_0$.

952 If the anti- k_T algorithm is applied, the exponent $a = -1$. This means that it tends to cluster high p_T
953 particles first, given that the hard term dominates d_{ij} in equation 3.10. Since the soft particles have
954 lower impacts, the shape of the jet is not sensitive to the soft radiation and rather stable against the
955 softer pile-up contributions.

956 In this analysis, clustering parameters of $R_0 = 0.8$ and $R_0 = 0.4$ will be used to define the “large-
957 cone” jets or AK8 jets, and the “standard” jets or AK4 jets. In order to avoid double-counting of PF
958 candidates, AK4 jets are considered only if the angular separation from the leading AK8 jet is larger
959 than $R_0 > 0.8$.

960 Since the detector response to different particles is non-linear, particular care should be taken in
961 the assignment of the measured momentum of the clustered jet to the corresponding true value
962 of the original parton [68]. A set of jet energy corrections (JECs) are applied sequentially and with
963 a fixed order. Each correction consists in a rescaling of the jet four-momentum, and it takes into
964 account different effects that are factorized.

- 965 • The L1 JECs remove the effect of the pile-up; they consist into an offset correction of the jet
966 p_T . They are determined from Monte Carlo (MC) simulations of di-jet events produced by
967 strong interaction with and without pile-up events on top, and parametrized as a function of
968 kinematical variables (jet area, pseudorapidity and p_T) and of the average p_T density per unit
969 area, ρ . Residual differences between data and the detector simulation are evaluated in data
970 collected with a random trigger, called zero bias, applying the only requirement of the beam
971 crossing happening. Pile-up offset corrections are displayed in fig. 3.14 (top left), as a function
972 of the jet pseudorapidity.
- 973 • The simulated response of the detector is not uniform over jet p_T and η . This effect is mit-
974 iated by the L2L3 MC-truth corrections. They are calculated in MC simulations of di-jet
975 events, by taking into account the discrepancy between the reconstructed p_T of the jet and the
976 true p_T at particle generator level (*i.e.*, before simulating the interaction of the parton showers
977 with the detector), as a function of jet p_T and η . L2L3 scale factors describing the simulated
978 jet response are reported in fig. 3.14 (top right), as a function of the jet pseudorapidity.

- 979 • The small data-MC discrepancies ($\sim 1\%$) left after applying the previous set of JECs are cor-
 980 rected by the L2 and L3 residual corrections. The L2Residuals are calculated in di-jet events,
 981 as a function of p_T . The L3Residuals are calculated in $Z \rightarrow (\mu\mu, e e) + \text{jet}$ events, photon
 982 + jet events and multi-jet events, as a function of η and p_T [68]. Data-MC scale factors for
 983 L2L3Residuals are displayed in fig. 3.14 (bottom), as a function of the jet η and p_T .
- 984 • An optional correction, not used in this analysis, is the L5 flavour-dependent correction, that
 985 is extracted from MC simulations.

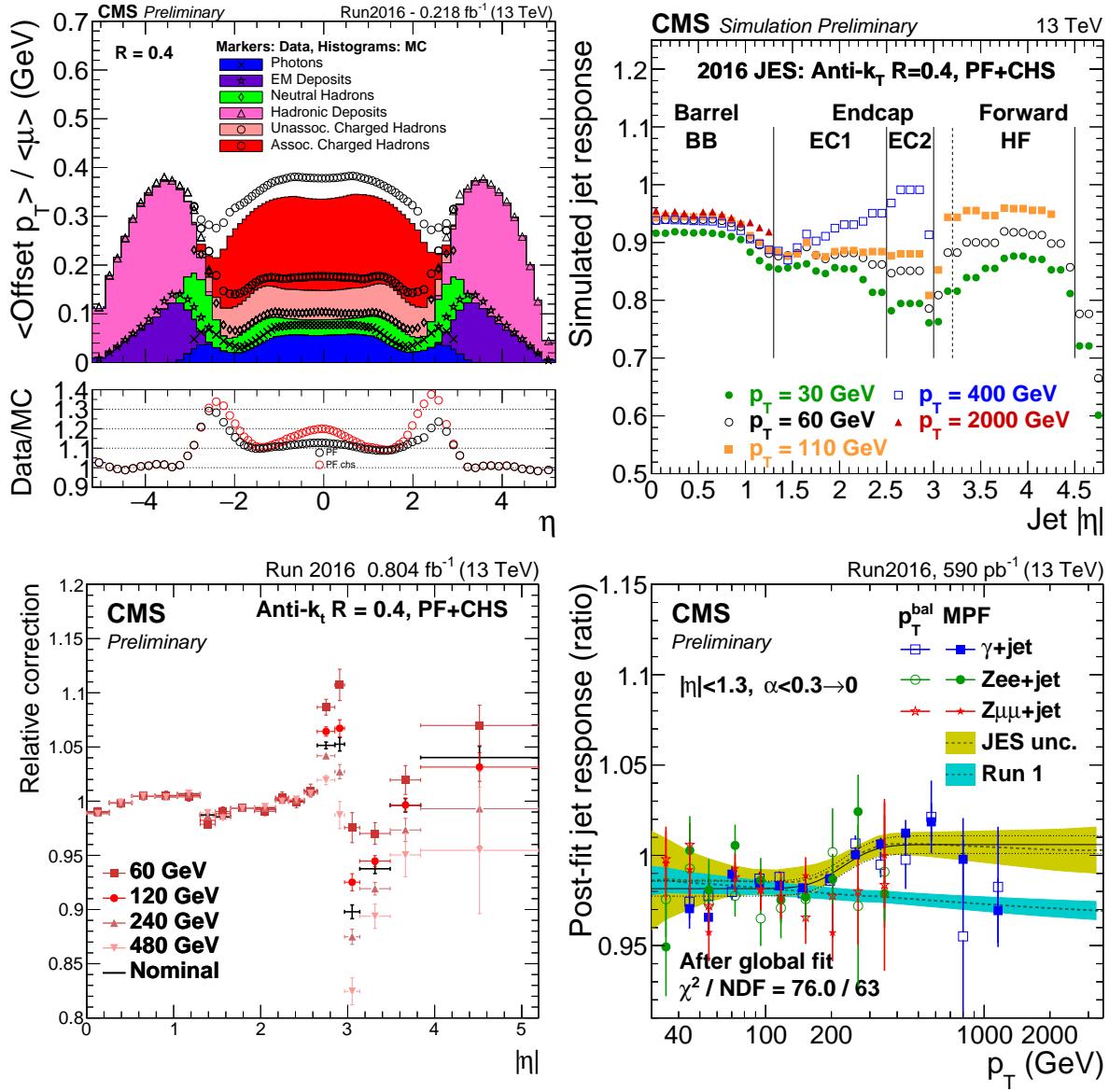


Figure 3.14: Top left: average p_T offset due to additional pile-up events, measured both in data and MC simulations, as a function of the jet pseudorapidity. Top right: simulated jet response (L2L3 MC-truth corrections), as a function of the jet pseudorapidity. Bottom left: L2L3 residual data-MC corrections, evaluated on di-jet events, as a function of the jet η . Bottom right: L2L3 residual data-MC corrections, evaluated on di-jet and $Z/\gamma + \text{jet}$ events, as a function of the jet p_T .

3.2 The CMS detector

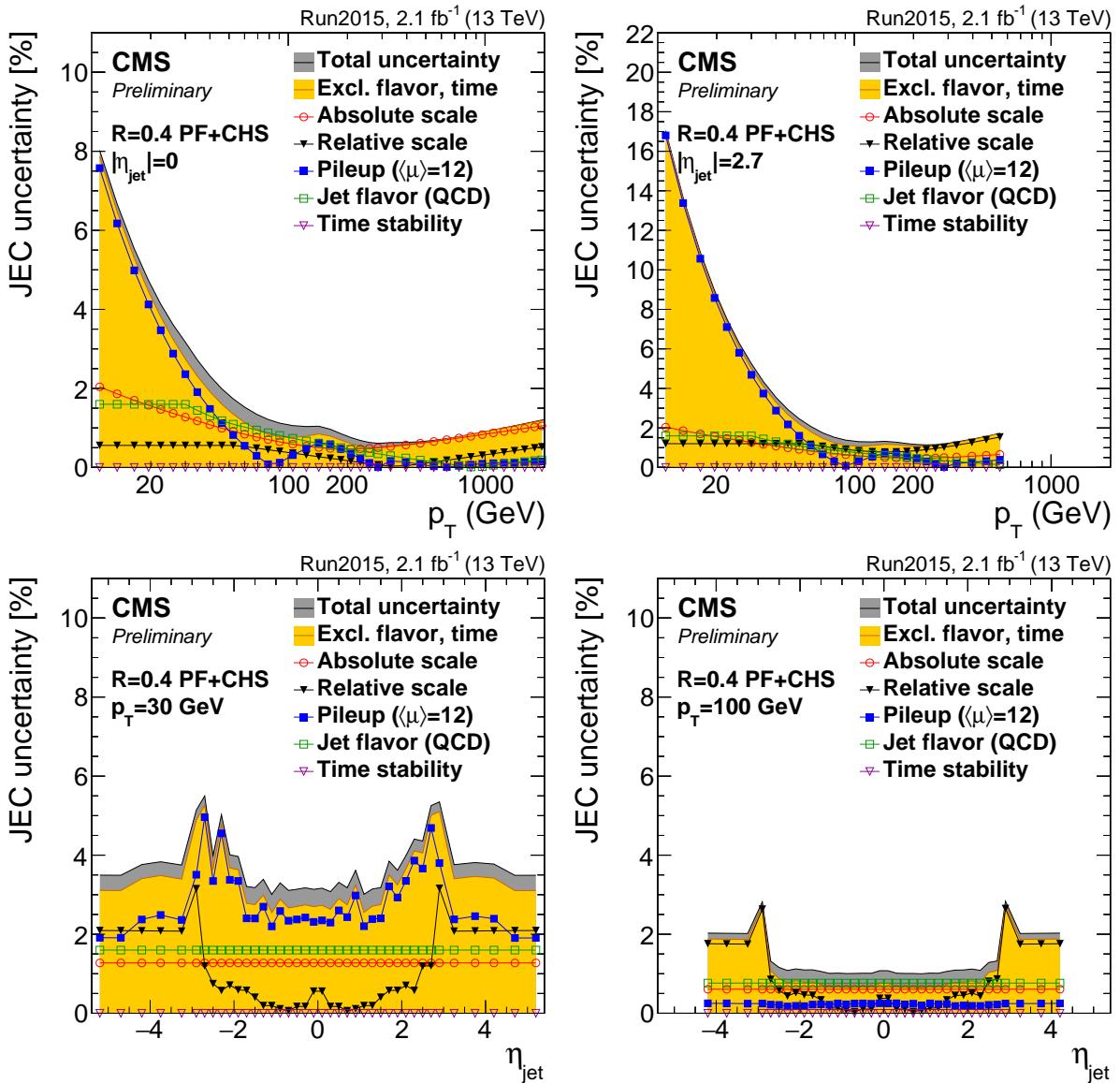


Figure 3.15: Jet energy corrections uncertainties, as a function of jet p_T (top) and η (bottom), calculated in 2015 data. The yellow histograms report the convolution of the uncertainties applied in the analysis.

986 Each jet energy correction is determined with an uncertainty, and reported in fig. 3.15 for 2015 data,
987 as a function of p_T and η of the jet. The total uncertainty for jets with p_T larger than 30 GeV (100
988 GeV) is smaller than 3% (1%) in the barrel, and up to 5% (3%) in the endcaps [69].

989 An additional effect that must be taken in account in the analysis is the discrepancy in the jet energy
990 resolution (JER) observed in data and in Monte Carlo samples. A smearing procedure is applied in
991 MC simulations (described in detail in sec. 4.3.6), in order to restore a better agreement. Jet energy
992 resolutions in Monte Carlo simulations are displayed in fig. 3.16 (top), as a function of the jet p_T
993 and the average number μ of reconstructed primary vertices, considering central (left) and forward
994 (right) jets. The resolution is stable against the pile-up for jet $p_T > 100$ GeV, and it ranges from 10% at
995 100 GeV, down to 4% at 1 TeV [69]. In fig. 3.16 (bottom), data-MC smearing scale factors are reported
996 as a function of η .

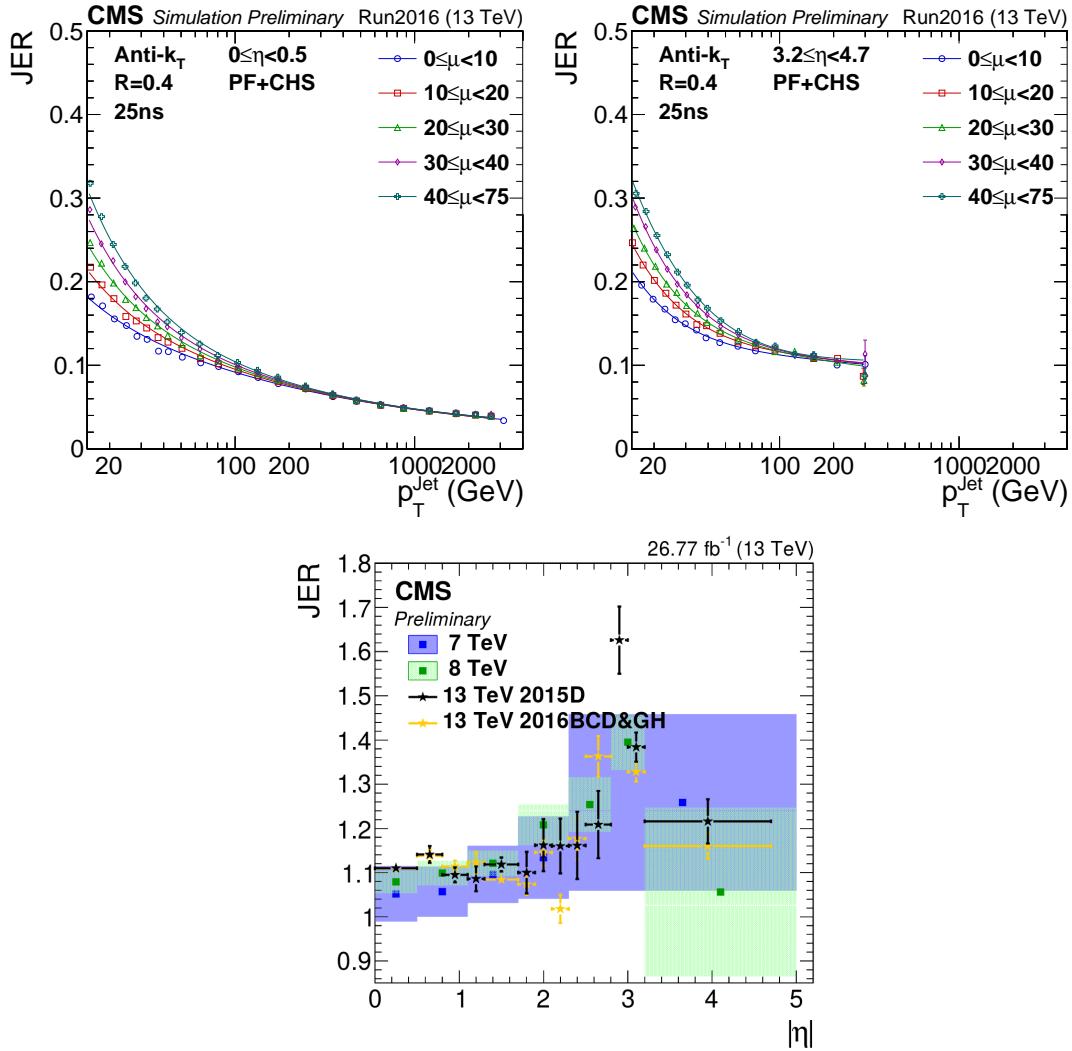


Figure 3.16: Top: jet energy resolution in MC simulations, as a function of the jet p_T . Different curves represent a different average number of primary vertices per event (μ). Bottom: data-MC scale factors, as a function of the jet η , measured in 2016 data (yellow dots).

997 3.2.9.6 Tau reconstruction

998 Tau leptons have a very small lifetime ($\sim 3 \times 10^{-13}$ s), hence they decay before reaching the pixel
 999 detector and they can only be reconstructed through their decay products. Approximately 60%
 1000 of the times, τ leptons decay into hadrons, hence they are reconstructed as small collimated jets
 1001 in the CMS detector. The main decay modes of the hadronic tau, τ_h , are one or three charged
 1002 mesons (mainly π^\pm), also in association with a π^0 decaying into a couple of photons, and a τ neu-
 1003 trino. Hence, photons and charged hadrons are the main ingredients of dedicated algorithms to per-
 1004 form the τ_h reconstruction and identification, in order to distinguish them from quark and gluon-
 1005 initiated jets. The main CMS τ_h reconstruction algorithm, Hadron Plus Strips (HPS) [70], is particle
 1006 flow based. HPS builds the tau candidate from a PF jet, clustered with the anti- k_T algorithm with
 1007 $R_0 = 0.5$, and it reconstructs the $\pi^0 \rightarrow \gamma\gamma$ decays within the jet cone, by taking into account the
 1008 photon conversions in the silicon detector. The exploitation of the PF informations is such that the
 1009 HPS algorithm shows stable performances in the reconstruction of the τ_h energy as a function of

3.2 The CMS detector

1010 the energy itself. The τ_h candidate is required to be isolated, namely no energy deposits other than
1011 the τ decay products should be present in the tau cone. Depending on the low threshold set to con-
1012 sider the surrounding particles as included in the cone, different isolation working points can be
1013 defined. With the looser working point, the probability of mis-identifying a quark or gluon jet as a
1014 tau is around 1% [70].

1015 **3.2.9.7 b-jets tagging**

1016 The bottom quark plays a fundamental role in numerous standard model processes, *i.e.* the physics
1017 related to the top quark (that decays into a W boson and a bottom quark, or b-quark, with a branch-
1018 ing fraction of 100%) and the Higgs boson (decaying into $b\bar{b}$ with a branching fraction $\sim 60\%$). Many
1019 algorithms have been exploited by the CMS collaboration, with the aim of distinguishing a b-quark
1020 initiated jet and jets originating from light quarks or gluons [71]. The most remarkable feature of the
1021 b-quark is the long lifetime (~ 1.5 ps), that has the experimental consequence of a displaced decay
1022 (few mm) with respect to the primary vertex. The direct leptonic decays of the b-quark (into μ and
1023 e) or the cascade leptonic decays involving charm quarks give an additional handle to its identifi-
1024 cation.

1025 Given the high spatial resolution of the silicon detector, track reconstruction is a key point of the b-
1026 tagging procedure. Tracks inside a jet candidate must satisfy criteria related not only to their quality
1027 but also to their distance from the interaction point. The track impact parameter is the distance
1028 between the primary vertex and the coordinate of closest approach. Tracks that are too far from
1029 the interaction point are discarded, in order to suppress the pile-up contributions. The Combine
1030 Secondary Vertex (CSV) algorithm [72] sorts jet candidates in categories, based on the number of
1031 reconstructed secondary vertices (one reconstructed secondary vertex, no secondary vertices but
1032 two tracks with large impact parameters, and the remaining cases). A multivariate approach allows
1033 to train the algorithm over the categories, considering as discriminating variables both tracking in-
1034 formations (numbers and properties of the tracks) and their relations with the secondary vertex re-
1035 construction (impact parameters; angular, linear, 2D and 3D distances of the vertex from the tracks
1036 and the jet axis; invariant mass of the charged particles associated to the secondary vertex).
1037 By tuning the selections, working points with different efficiencies have been set. The loose work-
1038 ing point, used in this analysis, has a 90% signal efficiency and a 40% mis-identification rate. The
1039 b-tagging efficiency is different in data and in simulations. Multiplicative scale factors are calcu-
1040 lated in events enriched in b-quark jets.

1041

1042 **3.2.9.8 Missing transverse energy reconstruction**

1043 Neutrinos can interact with the other particles only via the weak interaction; hence, when a neutrino
1044 is produced in the proton-proton collisions, it passes through the CMS experiment, undetected. Its
1045 only experimental signature is the momentum imbalance (\vec{p}_T^{miss}) in the transverse plane (r, φ). The
1046 magnitude of \vec{p}_T^{miss} vector is also called missing transverse energy, E_T^{miss} . Given its definition, it is
1047 evident that E_T^{miss} is a delicate variable to deal with, since it depends on all the other objects, on
1048 their imperfect measurements, on the detector noise and the pile-up events.
1049 The PF E_T^{miss} is the negative sum of the transverse momenta of the PF candidates reconstructed in
1050 the event. Inefficiencies in the tracker reconstruction and non-linear responses of the calorimeters
1051 can be corrected by propagating the jet energy corrections to \vec{p}_T^{miss} [73]:

$$\vec{p}_T^{\text{miss,corr}} = \vec{p}_T^{\text{miss}} - \sum_{j \in \text{jets}} (\vec{p}_{T,j}^{\text{corr}} - \vec{p}_{T,j}^{\text{raw}}), \quad (3.11)$$

1052 where "corr" ("raw") is related to the corrected (raw) p_T of the considered jet. This correction is
 1053 known as the "Type-I" correction to E_T^{miss} . Jets included in the calculation are AK4 jets with CHS
 1054 algorithm applied to remove the pile-up contributions, they must have $p_T > 15 \text{ GeV}$ and less than
 1055 90% of their energy deposited in the electromagnetic calorimeter. If a muon lies in the jet cone, it
 1056 is subtracted from the jet and added after the p_T correction. A similar correction is performed to
 1057 correct \vec{p}_T^{miss} at trigger level; in this case, a jet p_T threshold of 35 GeV is chosen.

1058 The E_T^{miss} uncertainty depends on the topology of the final state. It is calculated per-event by fac-
 1059 torizing \vec{p}_T^{miss} in components: electrons, photons, muons, taus, jets, jets with $p_T < 10 \text{ GeV}$ and all
 1060 the remaining PF candidates that are not clustered inside jets, called unclustered energy. The mo-
 1061 mentum of every object is varied within its uncertainties (namely, the energy scale and resolution),
 1062 and the effects are propagated to \vec{p}_T^{miss} . The most significant contributions to the unclustered en-
 1063 ergy is due to neutral PF hadrons and hadrons reconstructed in the forward hadronic calorimeter.
 1064 The effects related to jet energy scale and unclustered energy scale are measured on simulation, in
 1065 events with a top and an anti-top quarks, and amounts to 5% and 30% respectively [73].

1066

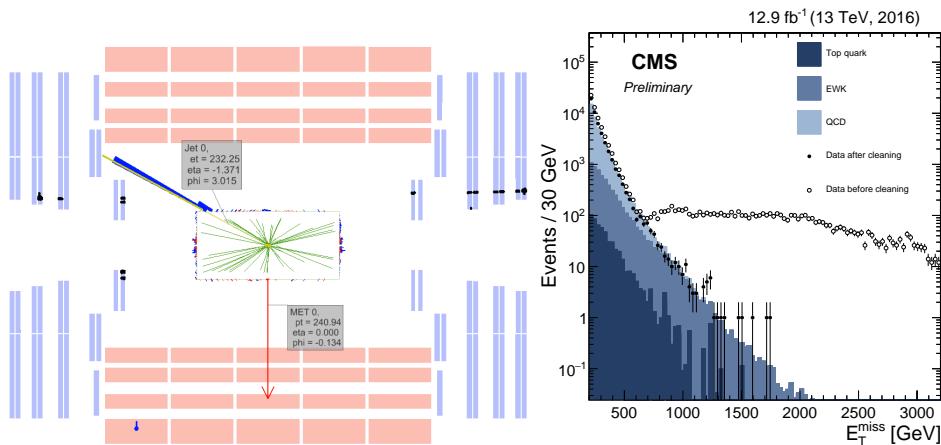


Figure 3.17: Left: event display of beam halo particles hitting the CSC detector. Right: comparison of data and simulations (histograms) when di-jet events are selected, before (open markers) and after (filled markers) anomalous E_T^{miss} cleaning algorithms have been applied on data.

1067 Many instrumental effects can give rise to anomalous E_T^{miss} determination: they have been studied
 1068 in detail during Run1 [74, 75] and Run2 [73], and they are mainly caused by ECAL and HCAL. In ECAL,
 1069 anomalous \vec{p}_T^{miss} is caused by particles hitting the sensors of the photodetectors, or by beam halo
 1070 particles (namely, particles produced in spurious proton interactions before reaching the interac-
 1071 tion point in the detector) showering inside the calorimeter, or by losses due to ECAL dead cells. An
 1072 event display representing beam halo muons hitting the CSC detector is shown in fig. 3.17 (left). In
 1073 HCAL, spurious \vec{p}_T^{miss} can be related to noise in the hybrid photodiodes and readout frontend. In
 1074 HF, missing p_T can be related to particles lost in the light guides and photomultipliers. Additional
 1075 anomalous E_T^{miss} can be produced by low quality muon tracks, that are not linked to segments re-
 1076 constructed in the muon chambers by the PF algorithm. These tracks are then classified as charged
 1077 hadrons, taken into account in the \vec{p}_T^{miss} calculation, and result into a large amount of fake E_T^{miss} .
 1078 Dedicated algorithms have been designed to identify and reject events with anomalous E_T^{miss} , and
 1079 they are consistently applied on data and simulations. In fig. 3.17 (right), Monte Carlo simulations
 1080 (coloured histograms) are compared to data before the algorithms removing the anomalous E_T^{miss}
 1081 have been applied (open markers) and after the cleaning (filled markers): the spurious high- \vec{p}_T^{miss}

3.2 The CMS detector

tail has been suppressed.

Performances of E_T^{miss} reconstruction are studied in events with a leptonic decay of a Z boson (into two muons or into two electrons) or an isolated photon. The distributions of E_T^{miss} are shown in fig. 3.18, separately for the three event categories. The hadronic recoil \vec{u}_T is defined in the transverse plane as the vectorial sum of all the particle transverse momenta, except the momentum \vec{q}_T of the vector boson considered (Z or γ). From the momentum conservation, the following relation holds:

$$\vec{q}_T + \vec{p}_T^{\text{miss}} + \vec{u}_T = 0. \quad (3.12)$$

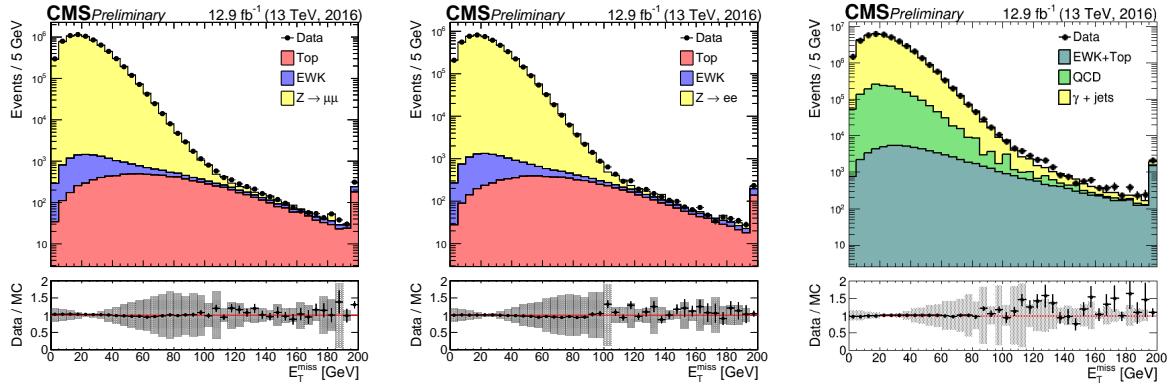


Figure 3.18: Data (black markers) and Monte Carlo (histograms) distributions of E_T^{miss} variable, in events reconstructing respectively a $Z \rightarrow \mu\mu$ decay (left), a $Z \rightarrow ee$ decay (center), an isolated photon (right).

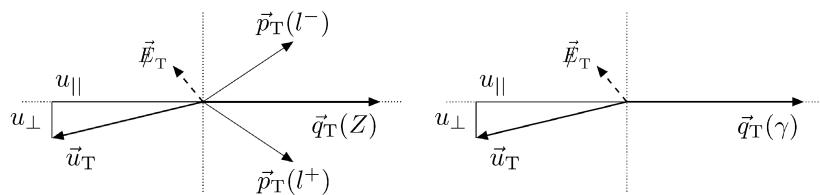


Figure 3.19: Left (Right): kinematics of $Z \rightarrow \ell\ell$ (photon) events in the (r, φ) plane; \vec{u}_T is the hadronic recoil, \vec{q}_T is the transverse momentum of the considered boson.

The hadronic recoil is projected in the parallel and perpendicular directions with regards to \vec{q}_T . The components $u_{||}$ and u_{\perp} , along with the vectors described in eq.3.12, are schematically represented in fig. 3.19. The E_T^{miss} response, defined as $-\langle u_{||} \rangle / \langle q_T \rangle$, is calculated as a function of q_T in data and simulations (fig. 3.20, left). The distributions of the two components of the hadronic recoil, $u_{||} + q_T$ and u_{\perp} , are modelled as Voigtian functions (the convolution of a Gaussian with a Breit-Wigner). The resolution of each component is calculated as the full width at half maximum of the corresponding Voigtian, and it is displayed in fig. 3.20 (center and right plots), as a function of q_T .

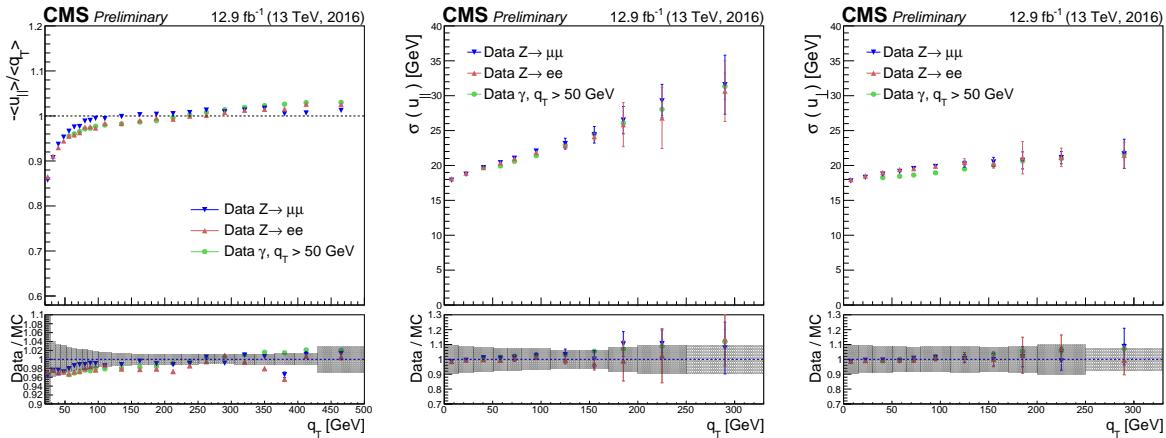


Figure 3.20: Left: E_T^{miss} response, as a function of the transverse momentum q_T of the vector boson considered in the event (Z decaying into $\mu\mu$, Z in ee or a photon), measured on data. Center and right: resolution on the measurement of the parallel and perpendicular hadronic recoil in data, as a function of q_T .

1095 3.3 ATLAS, ALICE, LHCb detectors

1096 3.3.1 ATLAS

1097 ATLAS (A Toroidal LHC ApparatuS) [76] is a multi-purpose experiment, that shares the same scientific goals of CMS. The simultaneous observation of an Higgs boson-like particle at the two experimental facilities represented an irrefutable proof of the discovery of the Higgs boson.

1100 ATLAS has a cylindrical shape (diameter of 25 m, length of 46 m) and weights 7000 tons. Like CMS, 1101 ATLAS is composed by many sub-detectors: trackers, calorimeters and muon system. The ATLAS 1102 magnetic field is provided by a solenoid, located inside the cylinder, and a big toroid, located outside 1103 the sub-detectors, able to reach a magnetic field of 2 T at the interaction point. The main differences 1104 among the two experiments are listed below.

- 1105 • *Tracker* – the CMS tracker has a better p_T resolution (mainly due to the higher magnetic field):
 $\sigma_{p_T}/p_T \approx 5 \cdot 10^{-4} p_T + 0.01$ at ATLAS; $\sigma_{p_T}/p_T \approx 1.5 \cdot 10^{-4} p_T + 0.005$ at CMS.
- 1106 • *Electromagnetic calorimeter* – the CMS electromagnetic calorimeter is completely enclosed
 inside the solenoid, whilst the ATLAS calorimeter is outside the solenoid. The particles going
 through the solenoid suffer an energy loss and a consequent deterioration of the energy reso-
 lution. The CMS ECAL has an energy resolution of $\sigma_E/E \approx 3\%/\sqrt{E}$; the ATLAS calorimeter
 has a sandwich structure (liquid argon and lead layers) and a resolution of $\sigma_E/E \approx 10\%/\sqrt{E}$.
- 1107 • *Hadronic calorimeter* – the CMS HCAL is partly inside the solenoid, partly outside, depauper-
 ating the resolution. The ATLAS hadronic calorimeter (made of iron and plastic scintillator
 tiles) has an energy resolution $\sigma_E/E \approx 50\%/\sqrt{E} + 0.03$ GeV; the CMS HCAL has a resolution
 of $\sigma_E/E \approx 100\%/\sqrt{E} + 0.05$ GeV.
- 1108 • *Muon system* – the peculiar geometry of the ATLAS muon system allows a better resolution of
 the standalone measurement of the muon momenta (*i.e.*, without using tracker and calorime-
 ters), that is around 10% at 1 TeV. CMS reaches better performances when combining the in-
 formations coming from the inner detectors (7% at 1 TeV against the 35% for the standalone
 measurement).

1121 **3.3.2 ALICE**

1122 ALICE (A Large Ion Collider Experiment) [77] studies the heavy ion collisions (lead-lead) or proton-
1123 ion collisions, in order to explore the physics of the hadrons in high density (or temperature) regimes,
1124 when a new state of matter appears, the so-called quark-gluon plasma (QGP). The QGP played a
1125 crucial role in the very first instants of the life of the universe.

1126 **3.3.3 LHCb**

1127 LHCb (Large Hadron Collider beauty) [78] is a detector designed to study the b-quark properties, in
1128 particular the CP violation and other rare phenomena related to B hadrons. The final aim of these
1129 measurements is trying to solve the matter-antimatter asymmetry problem.

1130 The three detectors are depicted in fig. 3.21.

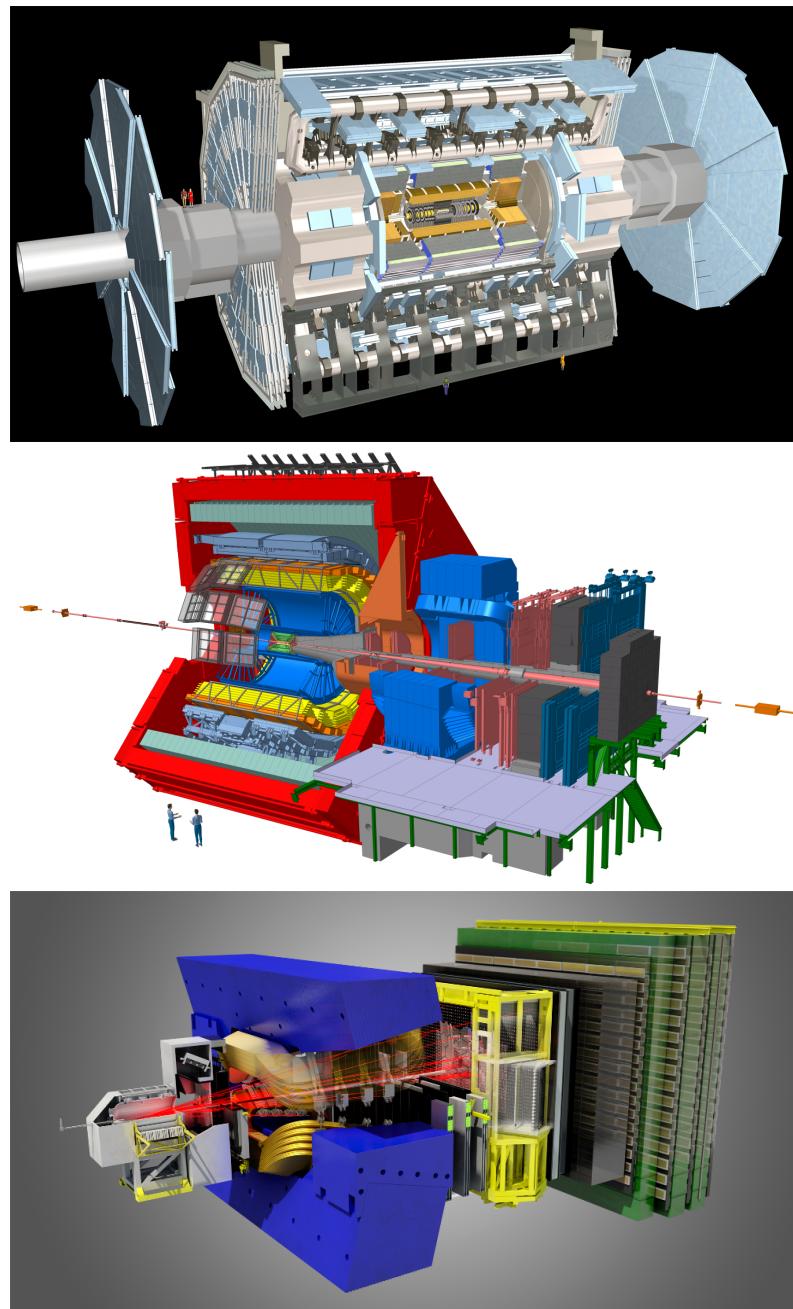


Figure 3.21: Top: the ATLAS experiment. Center: the ALICE experiment. Bottom: the LHCb experiment.

Search for diboson resonances in the $VZ \rightarrow q\bar{q} \nu\bar{\nu}$ final state

1132

1133

1134 4.1 Analysis overview

1135 This analysis searches for potential signals of heavy resonances decaying into a pair of vector bosons,
 1136 using the data collected by the CMS experiment during 2016, corresponding to an integrated lumi-
 1137 nosity of $\mathcal{L} = 35.9 \text{ fb}^{-1}$. One of the boson should be a Z , and it is identified through its invisible de-
 1138 cay into a couple of neutrinos ($\nu\bar{\nu}$), while the other electroweak boson, labelled as V and consisting
 1139 either in a W or in a Z boson, is required to decay hadronically into a pair of quarks ($q\bar{q}$). The decay
 1140 products (the bosons) of heavy (around the TeV scale) resonances are produced with large Lorentz
 1141 boosts; as a consequence, the decay products of the bosons (quarks and neutrinos) are expected to
 1142 be highly energetic and collimated. In this regime, the standard jet reconstruction algorithms fail in
 1143 distinguishing the two jets from the quarks, suggesting to look for a signature composed of a large-
 1144 cone high- p_T jet, in which both q and \bar{q} lie, recoiling against a large amount of missing transverse
 1145 momentum (\vec{p}_T^{miss}) due to the neutrinos escaping the detector. The hadronically decaying boson
 1146 (Z , W) is then reconstructed as one large-cone jet, whose mass is used to define the signal region
 1147 and signal-depleted control regions, the sidebands. Jet substructure techniques are exploited in or-
 1148 der to suppress background contamination and to classify the events in two exclusive signal purity
 1149 categories, allowing to improve the discovery reach.

1150 A general ZZ decay, predicted by the bulk graviton model (sec. 2.3.2), can be reconstructed both in
 1151 final states with high signal purity but limited statistics (four charged leptons) and large statistics
 1152 but overwhelming backgrounds (no charged leptons). The choice to look for one boson decaying
 1153 hadronically and the other Z into neutrinos represents the best compromise between these two ex-
 1154 tremes. This topology can be also utilized to reconstruct a charged spin-1 vector boson W' decaying
 1155 into an invisible Z and an hadronic W , predicted by the HVT model (sec. 2.2), making this analysis
 1156 sensitive to a generic VZ final state.

1157 Signal events are collected with trigger paths requiring high \vec{p}_T^{miss} recoiling against jet activity. This
 1158 signature is clearly a very challenging one in an environment with more than 50 primary collisions
 1159 per bunch crossing. For this reason, the Particle-Flow algorithm is run at trigger level to obtain the
 1160 highest possible resolution on the jets and thus on the \vec{p}_T^{miss} .

1161 The search is performed by examining the distribution of the diboson reconstructed transverse

mass of the resonance VZ (m_{VZ}^T) for a localized excess. The shape and normalization of the main background of the analysis (namely, the production of an electroweak boson in association with jets) are estimated with a data-simulation hybrid approach using the distribution of data in the sidebands, corrected for a function accounting for potential differences between the signal region and the sidebands. The predictions of the secondary background sources completely rely on simulations.

In fig. 4.1, a typical signal event of the $W' \rightarrow WZ \rightarrow q\bar{q}' \nu\bar{\nu}$ process detected by the CMS experiment is displayed; the mass of the W' is 2.5 TeV. The muon chambers in the barrel (DTs, in light red) and in the endcaps (CSCs, in light blue), along with the tracker detector (green) are shown in the (r, φ) transverse plane (left) and the (r, z) longitudinal plane (right). The large-cone jet, identifying the W hadronic decay, is displayed in red; the energy deposits in ECAL (light orange) and in HCAL (in violet) can be seen in the picture. The missing transverse energy, signature of the Z invisible decay, is represented as a blue arrow, lying in the transverse plane. The track multiplicity (green tracks) is shown in the center of the detector, where the tracker is installed.

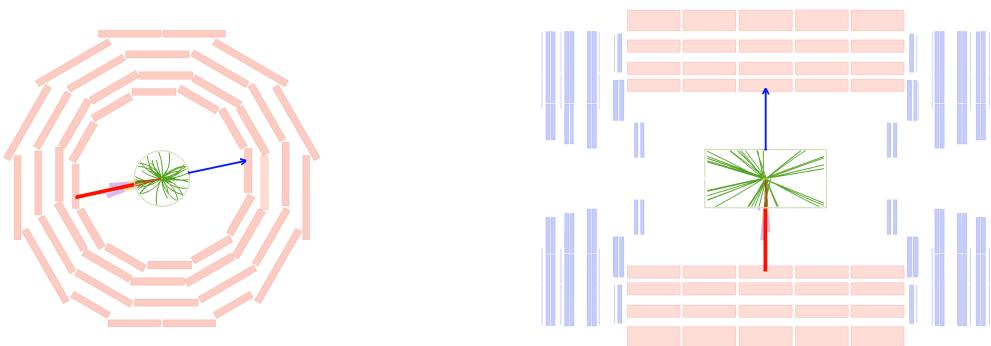


Figure 4.1: Left: representation of the decay of a W' of mass 2.5 TeV, in the transverse plane of the CMS detector. Right: representation of the decay of a W' of mass 2.5 TeV, in the longitudinal plane of the CMS detector.

4.2 Data and Monte Carlo simulations samples

4.2.1 Signal samples

Signal samples of a spin-2 (bulk graviton) decaying into a pair of Z bosons are exploited in the analysis. To target the final state, one of the two Z bosons is forced to decay into neutrinos, while the other Z is forced to decay hadronically. The signal samples are produced in the narrow-width approximation by setting the resonance width to 0.1% of its mass. Twelve mass points with 100000 events each are simulated, with a m_G ranging from 600 GeV up to 4500 GeV.

Additionally, samples of a spin-1 HVT-like W' resonance decaying into a Z boson and a W boson are studied. The Z boson is forced to decay into neutrinos, and the W boson is forced to decay hadronically. Also in this case the signal samples are produced in the narrow-width approximation by setting the resonance width to 0.1% of its mass. Twelve mass points with 100000 events each are simulated, with a $m_{W'}$ ranging from 600 GeV up to 4500 GeV.

The signal samples are generated at leading-order (LO) with the MADGRAPH5_AMCATNLO v2.2.2[79] matrix element generator, while hadronization and fragmentation are handled by PYTHIA 8 [80] version 8.2121 with CUETP8M1 [81] tuning. A full detector simulation and event reconstruction has

4.2 Data and Monte Carlo simulations samples

been performed with GEANT4 [82] and CMSSW. The detector alignment scenario, calibrations and pile-up distributions are generated according to the expectations in 2016 data.
All the signal samples used in the analysis and the related properties are reported in Tables 4.1-4.2.

Table 4.1: Spin-2 (bulk graviton) signal samples and production cross sections (assumed to be 1 pb) multiplied by the respective branching fractions of the Z decays considered ($\mathcal{B}(Z \rightarrow \nu\bar{\nu}) = 0.20$, $\mathcal{B}(Z \rightarrow q\bar{q}) = 0.6991$). A combinatorial factor of 2 is included in the cross-section calculation.

Signal process	m_G	Events	$\sigma \times \mathcal{B}$ (pb)
$G \rightarrow ZZ \rightarrow q\bar{q}\nu\bar{\nu}$	600 GeV	100000	0.27964
$G \rightarrow ZZ \rightarrow q\bar{q}\nu\bar{\nu}$	800 GeV	100000	0.27964
$G \rightarrow ZZ \rightarrow q\bar{q}\nu\bar{\nu}$	1000 GeV	100000	0.27964
$G \rightarrow ZZ \rightarrow q\bar{q}\nu\bar{\nu}$	1200 GeV	100000	0.27964
$G \rightarrow ZZ \rightarrow q\bar{q}\nu\bar{\nu}$	1400 GeV	100000	0.27964
$G \rightarrow ZZ \rightarrow q\bar{q}\nu\bar{\nu}$	1800 GeV	100000	0.27964
$G \rightarrow ZZ \rightarrow q\bar{q}\nu\bar{\nu}$	2000 GeV	100000	0.27964
$G \rightarrow ZZ \rightarrow q\bar{q}\nu\bar{\nu}$	2500 GeV	100000	0.27964
$G \rightarrow ZZ \rightarrow q\bar{q}\nu\bar{\nu}$	3000 GeV	100000	0.27964
$G \rightarrow ZZ \rightarrow q\bar{q}\nu\bar{\nu}$	3500 GeV	100000	0.27964
$G \rightarrow ZZ \rightarrow q\bar{q}\nu\bar{\nu}$	4000 GeV	100000	0.27964
$G \rightarrow ZZ \rightarrow q\bar{q}\nu\bar{\nu}$	4500 GeV	100000	0.27964

Table 4.2: Spin-1 (W') signal samples and production cross sections (assumed to be 1 pb) multiplied by the Z and W branching fraction ($\mathcal{B}(Z \rightarrow \nu\bar{\nu}) = 0.2$, $\mathcal{B}(W \rightarrow q\bar{q}) = 0.6760$).

Signal process	$m_{W'}$	Events	$\sigma \times \mathcal{B}$ (pb)
$W' \rightarrow WZ \rightarrow q\bar{q}'\nu\bar{\nu}$	600 GeV	100000	0.13482
$W' \rightarrow WZ \rightarrow q\bar{q}'\nu\bar{\nu}$	800 GeV	100000	0.13482
$W' \rightarrow WZ \rightarrow q\bar{q}'\nu\bar{\nu}$	1000 GeV	100000	0.13482
$W' \rightarrow WZ \rightarrow q\bar{q}'\nu\bar{\nu}$	1200 GeV	100000	0.13482
$W' \rightarrow WZ \rightarrow q\bar{q}'\nu\bar{\nu}$	1400 GeV	100000	0.13482
$W' \rightarrow WZ \rightarrow q\bar{q}'\nu\bar{\nu}$	1800 GeV	100000	0.13482
$W' \rightarrow WZ \rightarrow q\bar{q}'\nu\bar{\nu}$	2000 GeV	100000	0.13482
$W' \rightarrow WZ \rightarrow q\bar{q}'\nu\bar{\nu}$	2500 GeV	100000	0.13482
$W' \rightarrow WZ \rightarrow q\bar{q}'\nu\bar{\nu}$	3000 GeV	100000	0.13482
$W' \rightarrow WZ \rightarrow q\bar{q}'\nu\bar{\nu}$	3500 GeV	100000	0.13482
$W' \rightarrow WZ \rightarrow q\bar{q}'\nu\bar{\nu}$	4000 GeV	100000	0.13482
$W' \rightarrow WZ \rightarrow q\bar{q}'\nu\bar{\nu}$	4500 GeV	100000	0.13482

4.2.2 Signal characterization

This analysis is performed in a high mass region (from 1 TeV to 4.5 TeV). The MADGRAPH algorithm generates the hard process production in the collision with $p_T = 0$. In the next step of the simulation, during the hadronization, PYTHIA adds the QCD ISR (initial state radiation) and consequently a resonance p_T different from 0. Kinematical distributions at generator level are shown in fig. 4.2-4.4 for spin-2 bulk graviton signal, and in fig. 4.5-4.7 for spin-1 HVT W' signal.

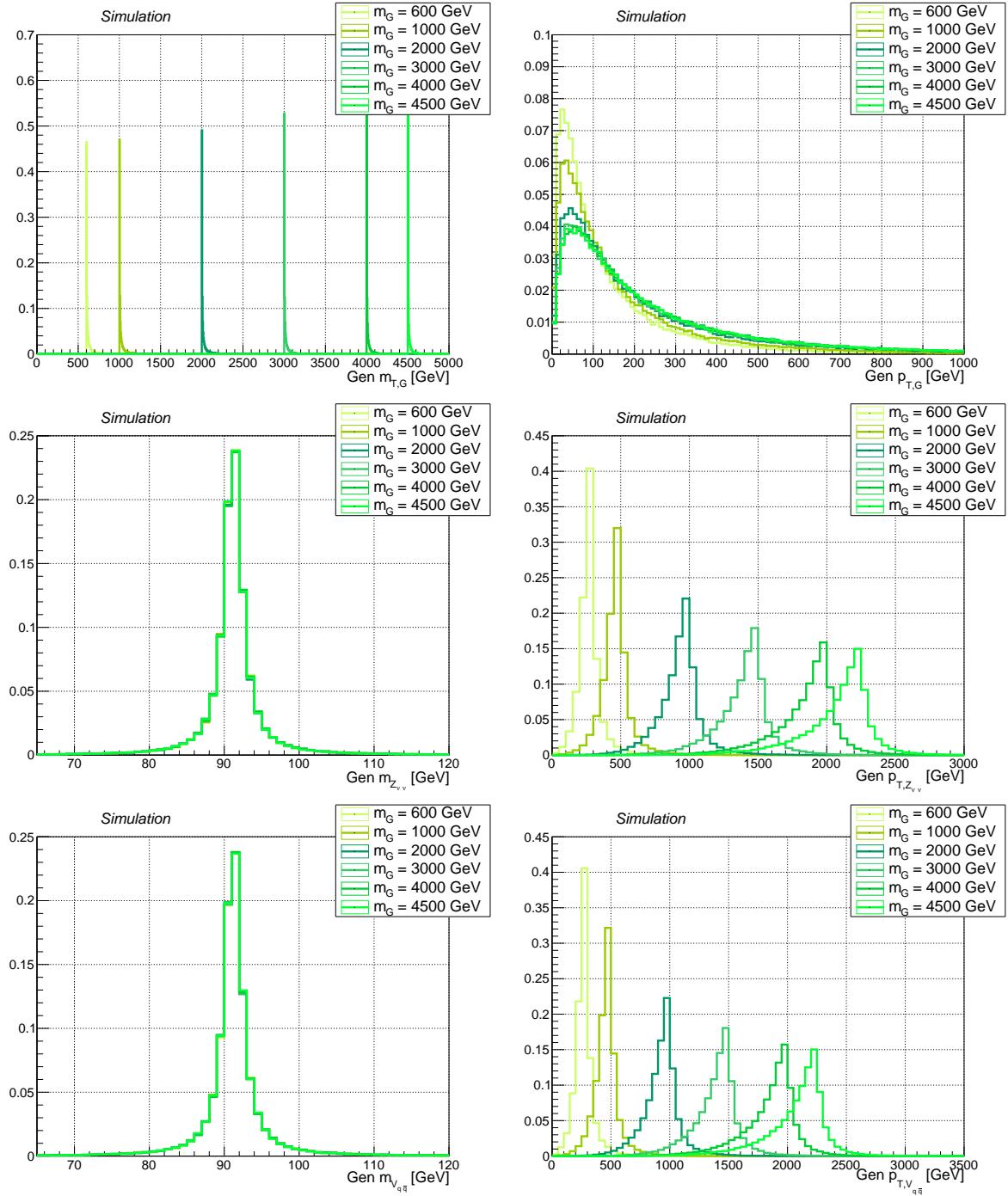


Figure 4.2: Main signal kinematic quantities at generation level after parton showering, for spin-2 bulk graviton signal, considering different mass hypotheses ($m_G = 0.6, 1, 2, 3, 4, 4.5$ TeV). Top: graviton transverse mass and p_T distributions. Center: invisibly decaying Z mass and p_T . Bottom: hadronically decaying Z mass and p_T .

4.2 Data and Monte Carlo simulations samples

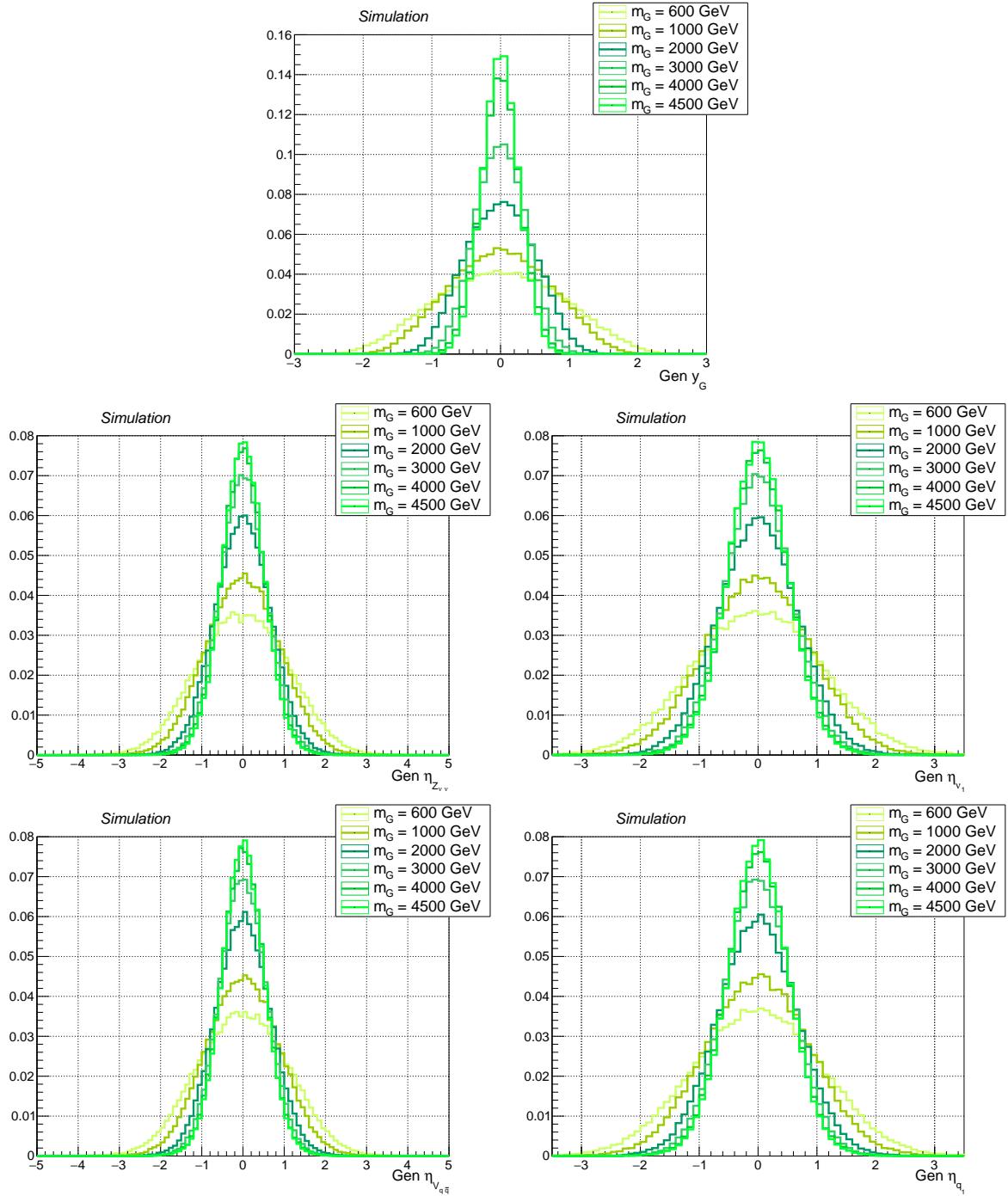


Figure 4.3: Main signal kinematic quantities at generation level after parton showering, for spin-2 bulk graviton signal, considering different mass hypotheses ($m_G = 0.6, 1, 2, 3, 4, 4.5$ TeV). Top: graviton rapidity \mathcal{Y} . Center: pseudorapidity η of the invisibly decaying Z , and pseudorapidity of the leading neutrino. Bottom: pseudorapidity η of the hadronically decaying Z , and pseudorapidity of the leading quark.

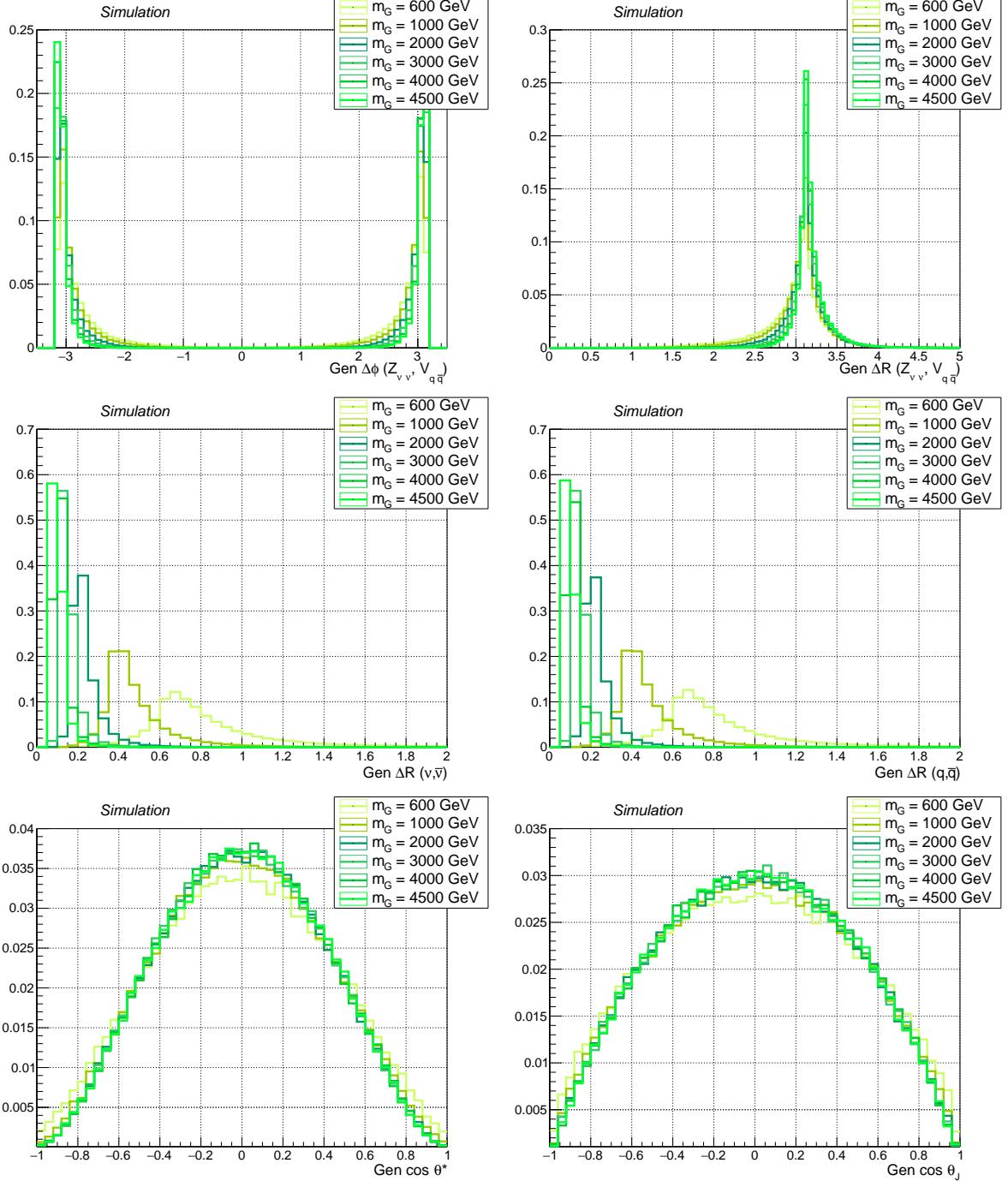


Figure 4.4: Main signal kinematic quantities at generation level after parton showering, for spin-2 bulk graviton signal, considering different mass hypotheses ($m_G = 0.6, 1, 2, 3, 4, 4.5 \text{ TeV}$). Top: angular separation in the transverse plane $\Delta\varphi$ (left) and solid angle ΔR (right) between leptonic Z and hadronic Z . Center: solid angle between the neutrinos and the quarks. Bottom: distribution of $\cos \theta^*$ and $\cos \theta_j$ (described in text).

4.2 Data and Monte Carlo simulations samples

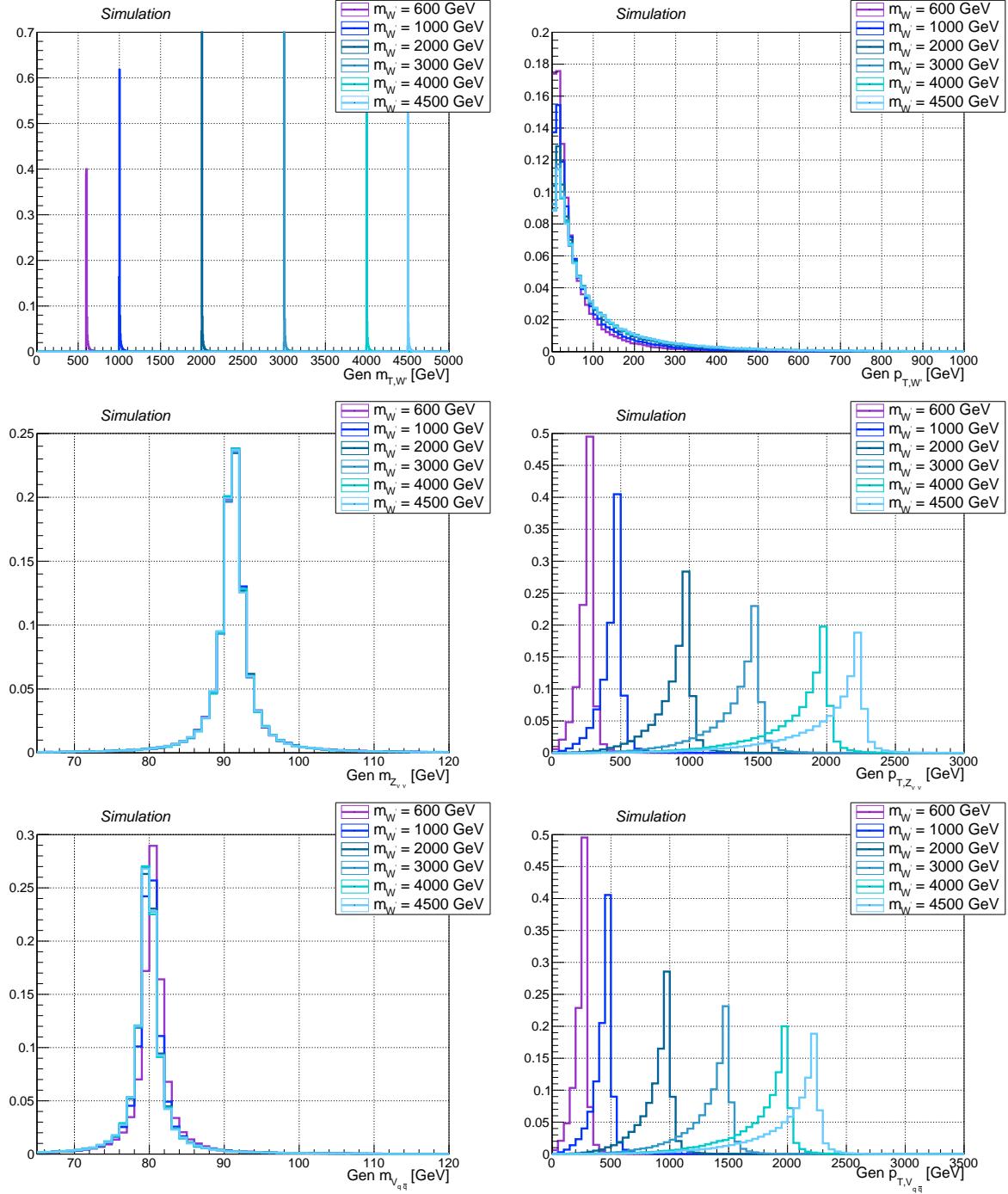


Figure 4.5: Main signal kinematic quantities at generation level after parton showering, for spin-1 W' signal, considering different mass hypotheses ($m_{W'} = 0.6, 1, 2, 3, 4, 4.5 \text{ TeV}$). Top: W' transverse mass and p_T distributions. Center: invisibly decaying Z mass and p_T . Bottom: hadronically decaying W mass and p_T .

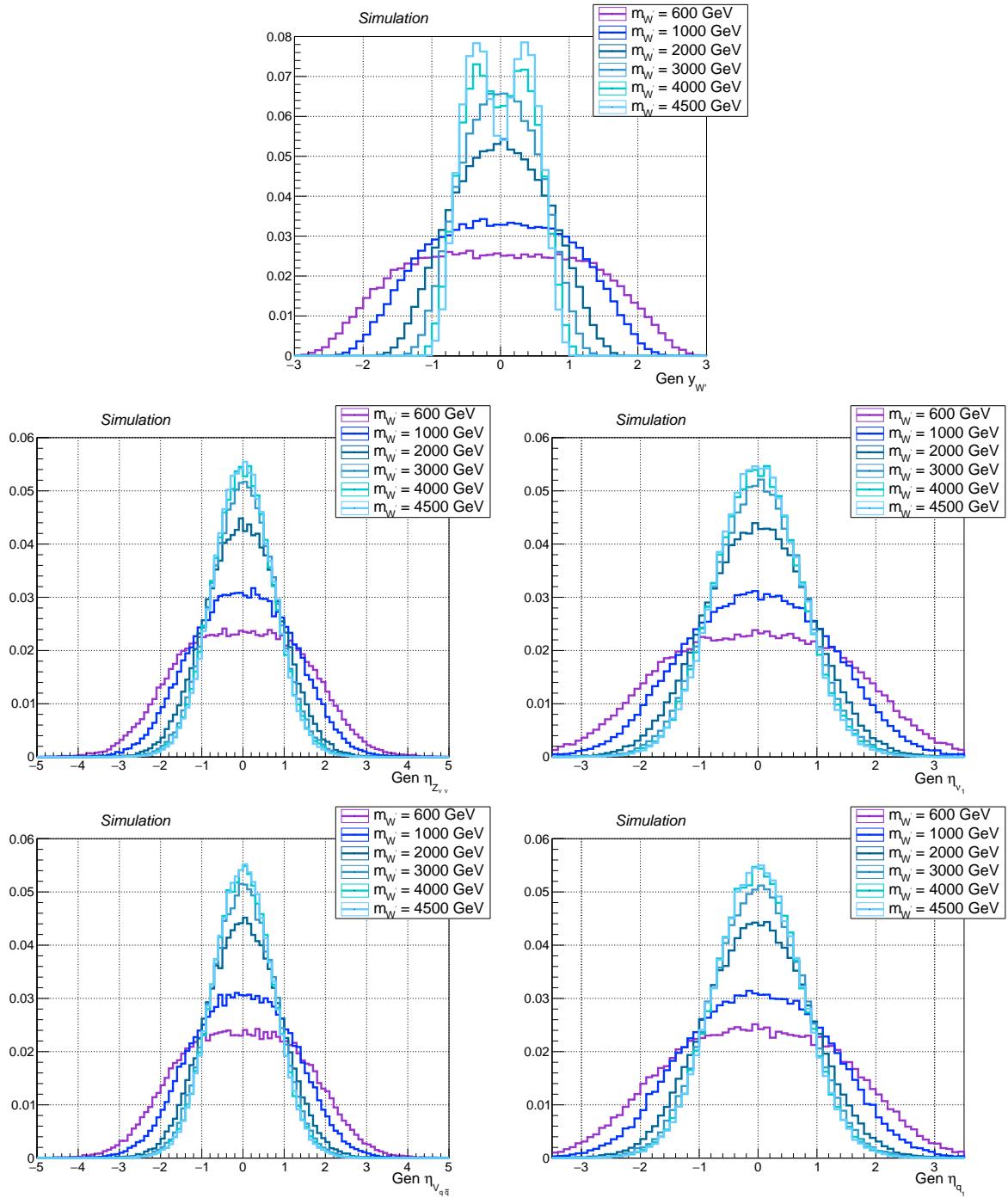


Figure 4.6: Main signal kinematic quantities at generation level after parton showering, for spin-1 W' signal, considering different mass hypotheses ($m_{W'} = 0.6, 1, 2, 3, 4, 4.5$ TeV). Top: W' rapidity \mathcal{Y} . Center: pseudorapidity η of the invisibly decaying Z , and pseudorapidity of the leading neutrino. Bottom: pseudorapidity η of the hadronically decaying W , and pseudorapidity of the leading quark.

4.2 Data and Monte Carlo simulations samples

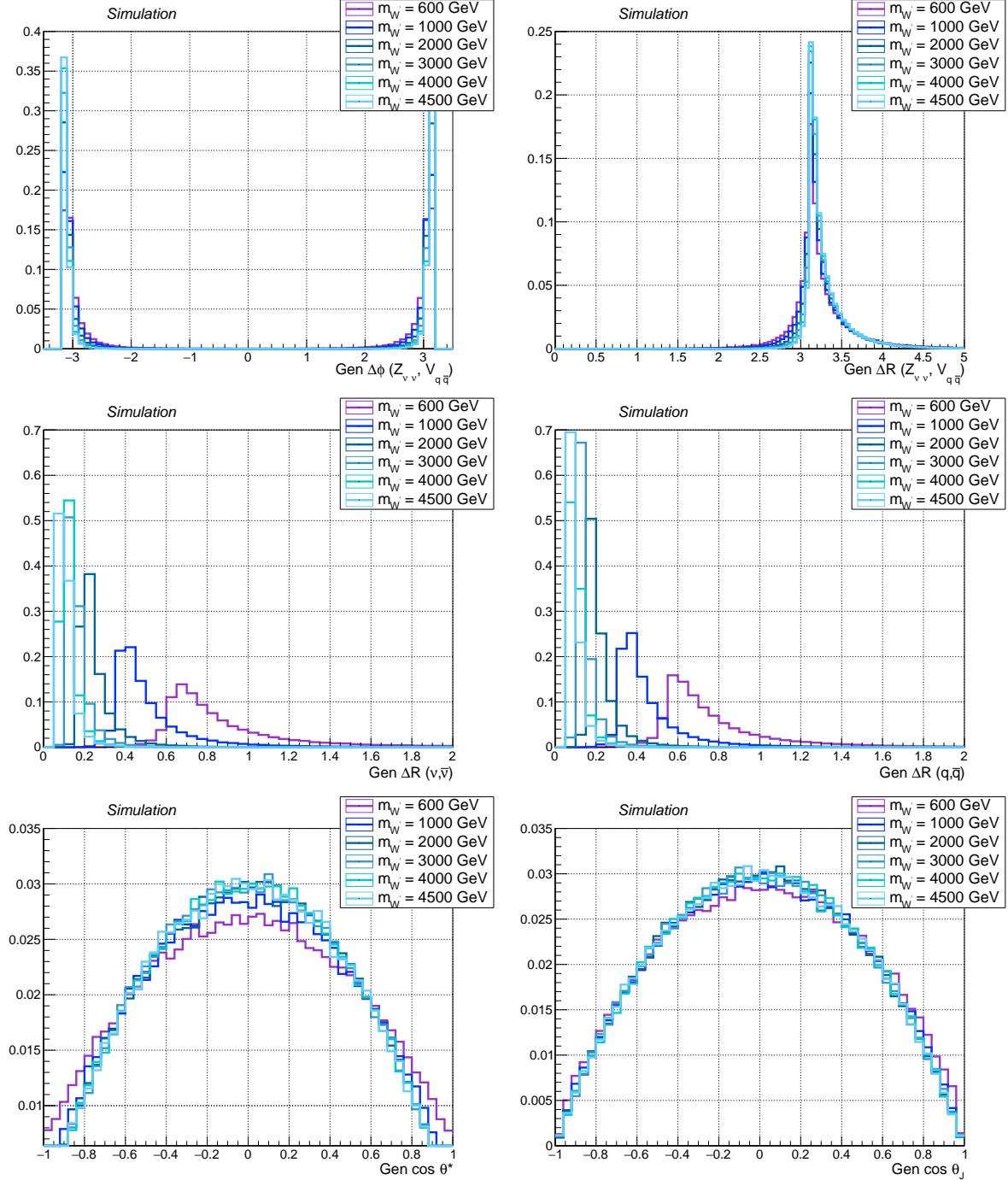


Figure 4.7: Main signal kinematic quantities at generation level after parton showering, for spin-1 W' signal, considering different mass hypotheses ($m_{W'} = 0.6, 1, 2, 3, 4, 4.5 \text{ TeV}$). Top: angular separation in the transverse plane $\Delta\varphi$ (left) and solid angle ΔR (right) between leptonic Z and hadronic W . Center: solid angle between the neutrinos and the quarks. Bottom: distribution of $\cos \theta^*$ and $\cos \theta_J$ (described in text).

1200 Angular distributions are related to the spin, the polarization and the kinematics of the produced
 1201 resonance; in particular:

- 1202 • the ΔR among neutrinos and quarks reflect the boosted nature of the electroweak bosons:
 1203 the more massive the resonance, the larger the boost, and hence the closer the fermions. By
 1204 looking at fig. 4.4-4.7, with a jet clustering parameter of 0.8 (AK8 jet) it is possible to enclose
 1205 the quarks produced by the decay of the V boson, for a resonance mass over 1 TeV;
- 1206 • the $\cos \theta^*$, namely the cosine of the angle between the momentum of the V boson, calculated
 1207 in the resonance rest frame, and the flight direction of the resonance itself in the laboratory
 1208 frame. This variable depends on the spin of the diboson resonance (spin-2 and spin-1 distri-
 1209 butions are different, fig. 4.4-4.7).
- 1210 • the $\cos \theta_J$, the cosine of the angle between the momentum of the leading quark, calculated in
 1211 the V rest frame, and the flight direction of the V boson in the laboratory frame. This variable
 1212 depends on the polarization state of the decay bosons [83]; in both HVT and bulk graviton
 1213 model, electroweak bosons are expected to be longitudinally polarized. When $\cos \theta_J \rightarrow 0$,
 1214 quarks are produced very close in angle and hence it is difficult to disentangle the two sub-
 1215 structures in the large-cone jet (sec. 4.3.8); when $\cos \theta_J \rightarrow \pi$ the quarks are emitted asymmet-
 1216 rically (one is softer than the other).

1217 4.2.3 Background samples

1218 The physics processes yielding final states with two neutrinos in association with a pair of quarks
 1219 are considered as sources of background; they are listed in tab. 4.3, along with the expected cross-
 1220 sections at next-to-leading order (NLO) or next-to-next-to leading (NNLO). A summary of the stan-
 1221 dard model cross-sections, measured by CMS, and their theoretical predictions is included in fig. 4.8-4.9.

- 1222 • **$Z + \text{jets}$** : this process represents the main irreducible background for the signal. The produc-
 1223 tion of a Z boson in association with one or more partons in the final state has topology that
 1224 is similar to the signal. This $Z + \text{jets}$ background is produced in samples binned in p_T of the
 1225 Z boson, starting from 100 GeV, with the AMC@NLO generator, with FXFX merging [84]. The
 1226 contribution from events with $p_T < 100$ GeV is negligible after the requirement on the \vec{p}_T^{miss}
 1227 to be greater than 200 GeV (sec. 4.3.12).
- 1228 • **$W + \text{jets}$** : the leptonic decay of a W boson can be an irreducible background if the charged
 1229 lepton escapes undetected (*i.e.* outside the detector acceptance) or fails the lepton identifi-
 1230 cation requirements. The production of a W boson has a cross section larger by an order
 1231 of magnitude with respect to the Z , and this makes the $W + \text{jets}$ a relevant background also
 1232 when a lepton veto is applied. This $W + \text{jets}$ background is produced in samples binned in
 1233 p_T of the W boson, starting from 100 GeV, with the AMC@NLO generator.
- 1234 • **Top**: the pair and single production of top quarks represent a source of background, given
 1235 the decay chain $t \rightarrow b W$, always including an electroweak W boson and a b-quark. The $t \bar{t}$
 1236 pair production manifests as two b-jets and two W bosons in the final state, that can decay to
 1237 leptons that escape the detector or fail to be identified as leptons. This analysis makes use of
 1238 $t \bar{t}$ inclusive decays samples based on POWHEG v2 [85] NLO generator. Single-top and single-
 1239 antitop samples are produced in the 5-flavours scheme using POWHEG v2 [86] NLO generator.
 1240 Different production mechanisms are considered: $t W$ channel, when a top quark is produced
 1241 in association with a W boson, due to a gluon-bottom quark scattering; s-channel, due to

4.2 Data and Monte Carlo simulations samples

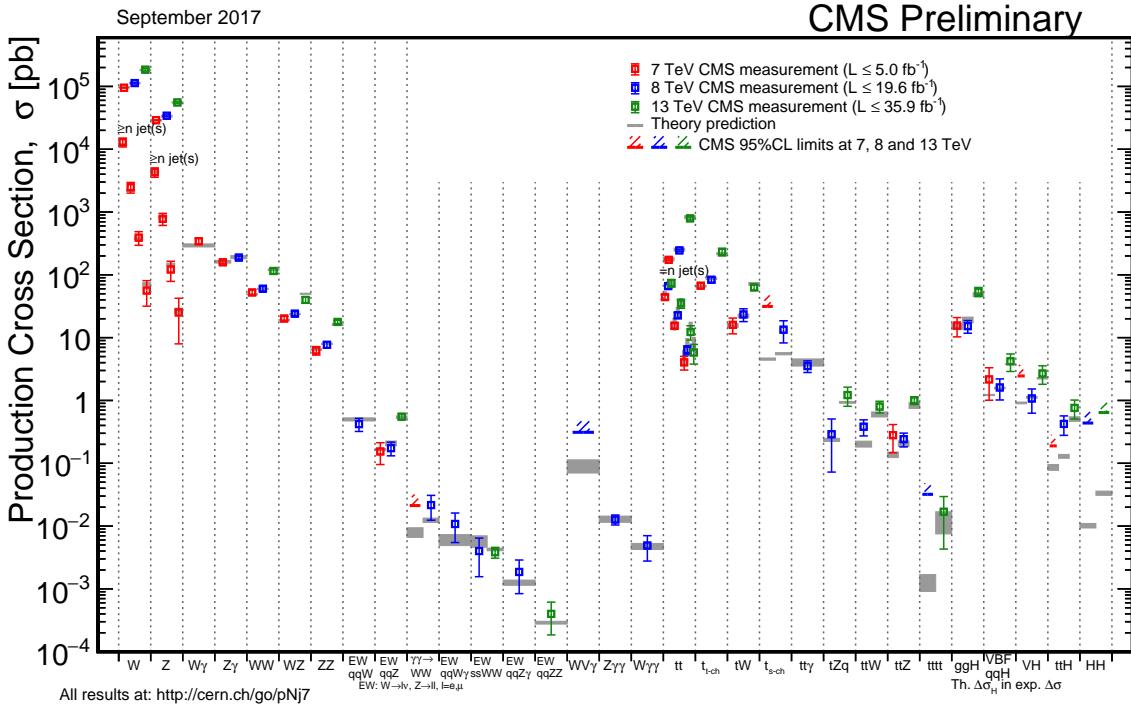


Figure 4.8: Production cross-sections of the main standard model processes, as measured by CMS, and theoretical predictions.

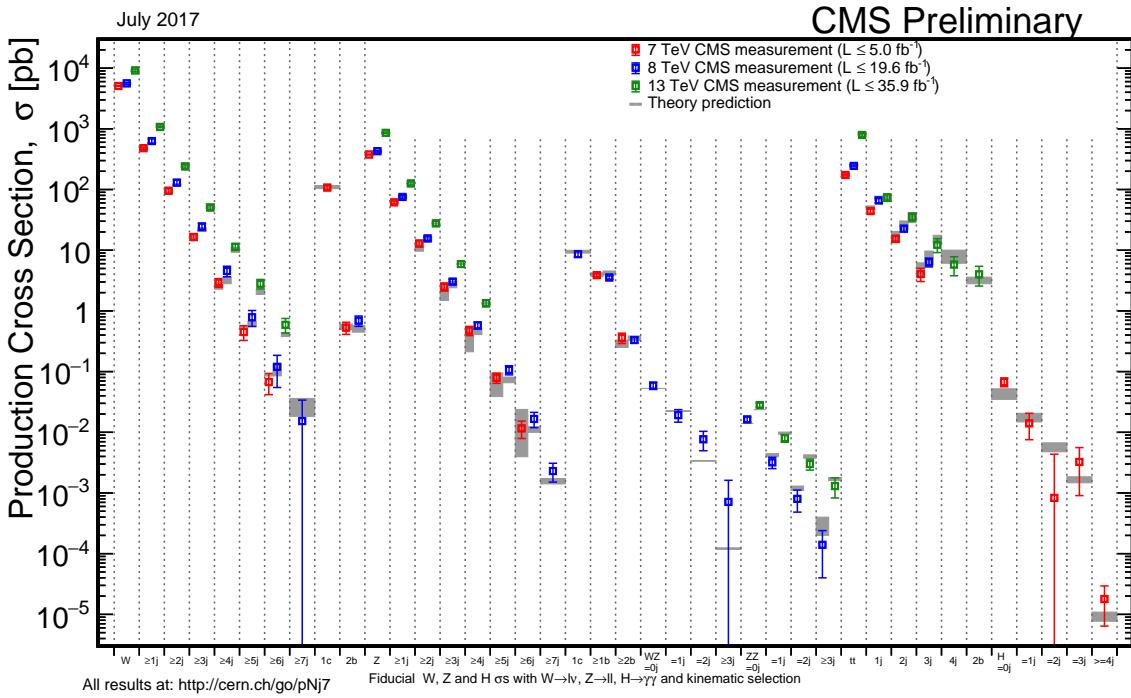


Figure 4.9: Production cross-sections of the standard model processes involving a vector boson in association with jets, as measured by CMS, and theoretical predictions. These phenomena represent the main background sources for the analysis.

1242 quark-antiquark scattering, producing a top and an anti-bottom quark in the final state; t-
 1243 channel, via a virtual W in a quark-b-quark scattering, resulting in a top quark and a quark
 1244 jet in the final state.

- 1245 • **Diboson:** the SM production of a pair of vector bosons is topologically close to the searched
 1246 signal, by the way the cross-section of the process is low. The WW production is the most
 1247 probable process, that imitates the signal when one of the W decays leptonically and the
 1248 charged lepton fall outside the detector acceptance or it is mis-identified; the WZ and ZZ
 1249 processes have smaller cross-sections but are topologically identical to the signal, except for
 1250 the fact that the invariant mass of the diboson system has a smoothly falling spectrum, in
 1251 contrast to the resonant signal distribution. Inclusive diboson production processes (WW ,
 1252 WZ , ZZ) are considered; they are simulated at LO by PYTHIA generator.
 1253
- 1254 • **Multi-jet:** despite the production of multi-jet events, initiated by the strong interactions, has
 1255 a very large cross-section, this source of background is suppressed by a dedicated selection
 and hence negligible for the analysis (sec. 4.3.12).

Table 4.3: Simulated Monte Carlo samples. The cross-section \times branching fraction for each process is shown in pb.

Signal process	Kinematical cuts	Generator	$\sigma \times \mathcal{B}$ [pb]	N. of events
$Z \rightarrow \nu\nu + \text{jets}$	$100 < p_{T,Z} < 250 \text{ GeV}$	amcatnloFXFX – Pythia8	170.4	10710313
$Z \rightarrow \nu\nu + \text{jets}$	$250 < p_{T,Z} < 400 \text{ GeV}$	amcatnloFXFX – Pythia8	6.636	2112619
$Z \rightarrow \nu\nu + \text{jets}$	$400 < p_{T,Z} < 650 \text{ GeV}$	amcatnloFXFX – Pythia8	0.9372	1101297
$Z \rightarrow \nu\nu + \text{jets}$	$p_{T,Z} > 650 \text{ GeV}$	amcatnloFXFX – Pythia8	0.1042	2047215
$W \rightarrow \ell \nu + \text{jets}$	$100 < p_{T,W} < 250 \text{ GeV}$	amcatnloFXFX – Pythia8	676.3	20178260
$W \rightarrow \ell \nu + \text{jets}$	$250 < p_{T,W} < 400 \text{ GeV}$	amcatnloFXFX – Pythia8	23.94	2001382
$W \rightarrow \ell \nu + \text{jets}$	$400 < p_{T,W} < 650 \text{ GeV}$	amcatnloFXFX – Pythia8	3.031	1939947
$W \rightarrow \ell \nu + \text{jets}$	$p_{T,W} < 650 \text{ GeV}$	amcatnloFXFX – Pythia8	0.4524	1974609
$t\bar{t}$ inclusive	-	Powheg – Pythia8	831.76	77229341
$t(\bar{t}W \text{ channel})$	-	Powheg – Pythia8	35.85	6952830
5f inclusive				
$\bar{t}(\bar{t}W \text{ channel})$	-	Powheg – Pythia8	35.85	6933094
5f inclusive				
t (s-channel)	-	amcatnloFXFX – Pythia8	3.344	622990
4f lepton decays				
t (t-channel)	-	Powheg – Madspin – – Pythia8	136.02	67240808
4f inclusive				
\bar{t} (t-channel)	-	Powheg – Madspin – – Pythia8	80.95	38811017
4f inclusive				
WW inclusive	-	Pythia8	118.7	7981136
WZ inclusive	-	Pythia8	47.2	3995828
ZZ inclusive	-	Pythia8	16.6	1988098

1256 4.2.4 Vector boson momentum corrections

1257 Corrections to the p_T spectrum of the V boson, due to NLO electroweak contributions, are en-
 1258 hanced at TeV scale [87], and they become significant for the purpose of this search. These cor-

4.2 Data and Monte Carlo simulations samples

reactions are effectively applied on a per-event basis, depending on the p_T of the vector boson at generation level. Figure 4.10 shows the amount of the corrections for the W and Z bosons.

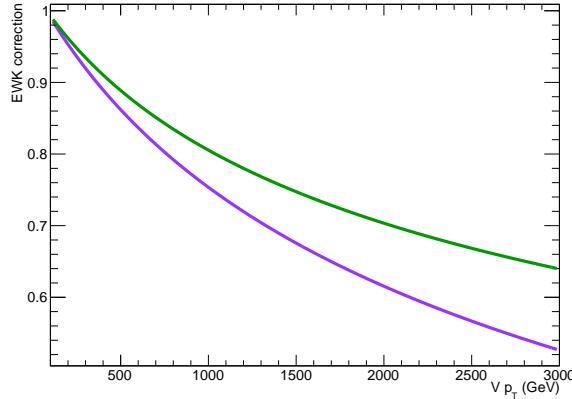


Figure 4.10: Electroweak corrections for the Z (green line) and W boson (purple line) as a function of the transverse momentum of the boson [87].

4.2.5 Data samples

Data samples used in this analysis have been collected during proton-proton collisions produced at LHC in 2016, at a center-of-mass energy of 13 TeV, with colliding bunches spaced by 25 ns, and with the magnetic field enabled. Three group of datasets have been considered:

- the MET dataset, where the analysis is performed, is collected by triggers requiring a large amount of \vec{p}_T^{miss} at HLT level in the event;
- the SingleMuon dataset, used to perform an unbiased trigger efficiency estimate, is collected by triggers requiring at least one well defined muon at HLT level;
- the SingleElectron dataset, used as cross-check for the trigger efficiency estimation, is collected by triggers requiring at least one well defined electron at HLT level.

Data selected for the analysis include all the runs certified as “good” for all subsystems. The corresponding integrated luminosity amounts to $35.9 \pm 0.9 \text{ fb}^{-1}$ [88]. In order to remove problematic or noise-dominated events, dedicated E_T^{miss} filters have been applied on data (and simulations).

4.2.6 Trigger

The most remarkable feature of the signal topology is the presence of a boosted Z decaying into neutrinos; the natural choice for the trigger requirement is to filter data firing at least one of the \vec{p}_T^{miss} trigger HLT paths listed in tab. 4.4, along with their corresponding L1 missing energy or jet seeds. PFMETNoMu indicates the E_T^{miss} (no μ) quantity, defined as the magnitude of the missing transverse momentum, reconstructed with the Particle-Flow algorithm at HLT, removing the muon candidates from the vector sum. PFMHTNoMu indicates the missing hadronic activity H_T^{miss} (no μ), defined as the magnitude of the vector sum of the transverse momenta of the jets, reconstructed with the Particle-Flow algorithm at HLT, once the muon candidates have been removed. PFMET indicates the pure E_T^{miss} calculated with Particle-Flow algorithm at HLT; different filters are applied

Table 4.4: HLT trigger paths used in the analysis.

HLT path	L1 seeds
HLT_PFMETNoMu90_PFMHTNoMu90_IDTight	L1_ETM70 OR L1_DoubleJetC56_ETM60 OR L1_ETM60 OR L1_ETM50
HLT_PFMETNoMu110_PFMHTNoMu110_IDTight	L1_ETM70 OR L1_DoubleJetC56_ETM60 OR L1_ETM60 OR L1_ETM50
HLT_PFMETNoMu120_PFMHTNoMu120_IDTight	L1_ETM70 OR L1_DoubleJetC56_ETM60 OR L1_ETM60 OR L1_ETM50
HLT_PFMET170_NoiseCleaned or	L1_ETM60 OR L1_ETM70
HLT_PFMET170_JetIdCleaned or	L1_ETM60 OR L1_ETM70
HLT_PFMET170_HBHECleaned	L1_ETM60 OR L1_ETM70

1284 at HLT (cleaning events from noise in the detector). Different thresholds are applied to E_T^{miss} (no μ)
 1285 and H_T^{miss} (no μ).

1286 The approach adopted in this analysis consists in calculating the trigger efficiency on data, and
 1287 applying the measured efficiency to Monte Carlo samples. Therefore the trigger is not required to
 1288 have been fired in MC.

1289 Given that the final state probed by the analysis consists into one AK8 jet, large E_T^{miss} and no charged
 1290 leptons, an unbiased measurement of the E_T^{miss} trigger efficiency can be performed in an orthogo-
 1291 nal dataset, collected with different triggers, and requiring events where a $W \rightarrow \ell \nu$ leptonic decay
 1292 is taking place. This guarantees the presence of real \vec{p}_T^{miss} in the event, due to the neutrino; fur-
 1293 thermore, the presence of a charged lepton guarantees that the leptonic W -like events are not over-
 1294 lapped with the search region. The additional requirement to have at least one AK8 jet is applied in
 1295 the trigger measurement, in order to probe a kinematical region similar to that of the signal region
 1296 of the analysis.

1297 The efficiency of the E_T^{miss} triggers is measured on SingleMuon dataset by selecting $W \rightarrow \mu \nu$ events
 1298 using a logic or of single muon triggers HLT_IsoMu24 OR HLT_IsoTkMu24_v, namely, triggers
 1299 asking for a PF muon reconstructed at HLT, with a p_T threshold of 24 GeV, that is isolated (in the
 1300 whole reconstruction or at tracker level only). Offline selections consist in asking to have one iso-
 1301 lated muon, with a suitable p_T threshold to be in the plateau of the muon trigger. The efficiency has
 1302 been calculated as a function of the minimum quantity between the offline reconstructed E_T^{miss} (no
 1303 μ):

$$E_T^{\text{miss}} (\text{no } \mu) = \left| \vec{p}_T^{\text{miss}} + \sum_i \vec{p}_T^{\mu,i} \right|, \quad (4.1)$$

1304 where the contribution of all the offline PF muons is removed from the \vec{p}_T^{miss} computation as in the
 1305 online algorithm, and the offline H_T^{miss} , defined as

$$H_T^{\text{miss}} = \left| \sum_j^{\text{n. of AK4 jets}} p_T^j \right|. \quad (4.2)$$

1306 This approach guarantees to mimic the behaviour of the online L1 trigger seeds. The detailed se-
 1307 lections are listed below:

- 1308 • HLT_IsoMu24_v OR HLT_IsoTkMu24_v,

4.2 Data and Monte Carlo simulations samples

- 1309 • 1 isolated muon $p_T > 35$ GeV, identified with tight requirements,
- 1310 • at least one AK8 jet, $p_T > 170$ GeV, $|\eta| < 2.5$, identified with loose requirements,
- 1311 • AK4 jets included in H_T^{miss} : $p_T > 30$ GeV, $|\eta| < 2.5$, identified with loose requirements.

1312 The efficiency of the E_T^{miss} triggers has independently been measured also on SingleElectron
1313 dataset, by selecting $W \rightarrow e \nu$ events using a single electron trigger (HLT_Ele27_WP Loose_Gsf
1314 OR HLT_Ele27_WP Tight_Gsf OR HLT_Ele32_WP Tight_Gsf), asking to have one well identified
1315 electron, with a suitable p_T threshold, and asking to the electron and the \vec{p}_T^{miss} to be separated in
1316 the transverse plane (hence, in φ) in order to suppress fake jet events mis-identified as electrons at
1317 trigger level ($\Delta\varphi > 0.5$). The detailed selections are listed below:

- 1318 • HLT_Ele27_WP Loose_Gsf OR HLT_Ele27_WP Tight_Gsf OR HLT_Ele32_WP Tight_Gsf,
- 1319 • 1 electron, $p_T > 35$ GeV, identified with tight requirements,
- 1320 • at least one AK8 jet, $p_T > 170$ GeV, $|\eta| < 2.5$, identified with loose requirements,
- 1321 • AK4 jets included in H_T^{miss} : $p_T > 30$ GeV, $|\eta| < 2.5$, identified with loose requirements.

1322 All the available data have been employed to derive the efficiency. The final turn-on curves for the
1323 E_T^{miss} triggers are shown in fig.4.11-4.12, measured in muon and electron dataset respectively. The
1324 PF MET No Mu trigger efficiencies are displayed separately, together with their logic OR. The trigger
1325 efficiency measured on SingleMuon dataset amounts to 96% at $E_T^{\text{miss}} = 200$ GeV; the trigger effi-
1326 ciency measured on SingleElectron dataset amounts to 95% at $E_T^{\text{miss}} = 200$ GeV. The difference
1327 needed to cover the gap between the two independent measurements is taken as trigger systematic
1328 uncertainty, and it amounts to 1% at 200 GeV.

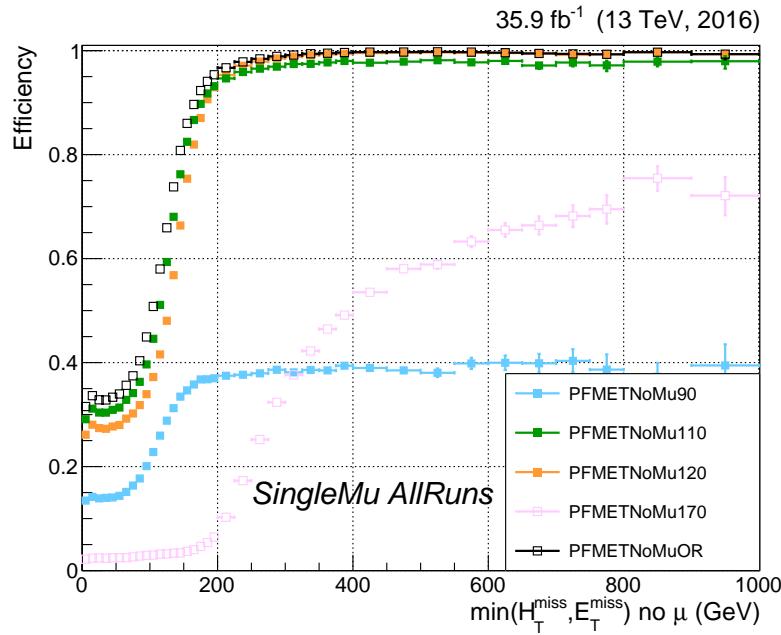


Figure 4.11: \vec{p}_T^{miss} trigger efficiency for the \vec{p}_T^{miss} trigger paths used in this analysis, calculated on SingleMuon dataset, as a function of the minimum of the variables E_T^{miss} (no μ) (eq. 4.1) and H_T^{miss} (eq. 4.2).

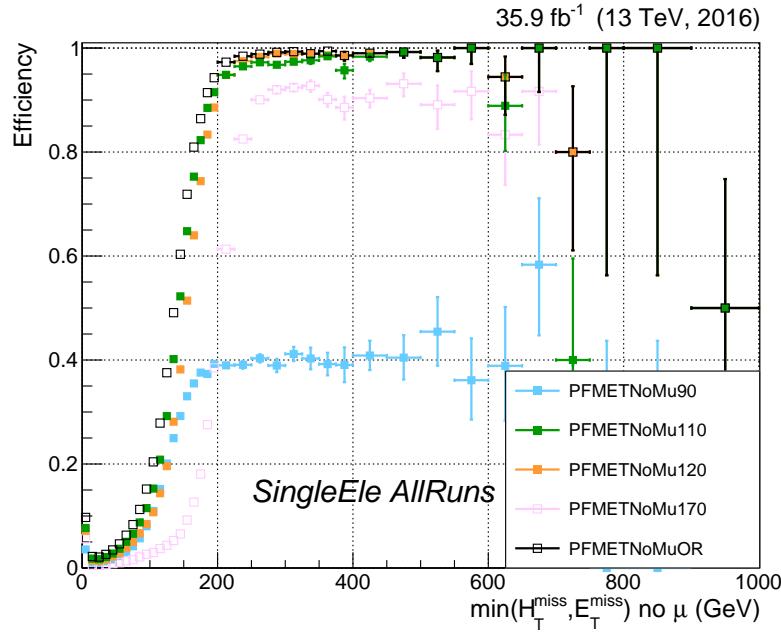


Figure 4.12: \vec{p}_T^{miss} trigger efficiency for the \vec{p}_T^{miss} trigger paths used in this analysis, calculated on SingleElectron dataset, as a function of the minimum of the variables E_T^{miss} (no μ) (eq. 4.1) and H_T^{miss} (eq. 4.2).

4.3 Event selection

4.3 Event selection

In this section, the selections applied to the physics objects used in the analysis are presented and motivated by performance and validation plots. Background events are represented as coloured histograms: $Z + \text{jets}$ events in light blue, $W + \text{jets}$ events in violet, $t\bar{t}$ events in yellow, single-top events in orange, diboson (or VV) events in blue, multi-jet (QCD) events in gray. Background uncertainties are displayed as black shaded areas. Signal samples are represented as coloured shaded histograms: the kind of signal (graviton or W'), the mass and cross-section of the considered resonance are reported in the legend. Data are represented with black markers, with their corresponding Poissonian uncertainty bars. If data are displayed, the data-MC ratio is reported per each bin in the bottom panel, along with the overall data-MC ratio calculated in the whole spectrum and the scores of χ^2 and Kolmogorov-Smirnov goodness-of-fit tests.

4.3.1 Vertex and Pile-up

Due to the pile-up effect, several vertices are typically reconstructed in one event. The primary vertex of the event is defined as the one with the highest sum of transverse momenta $\sum p_T^2$ of clustered physics objects associated to it, which passes the following selections:

- number of degrees of freedom $N_{DoF} > 4$
- vertex position along the beampipe $|z_{vtx}| < 24\text{cm}$
- vertex distance with respect the beam pipe $d_0 < 2\text{cm}$

where z_{vtx} and d_0 are the distance along and perpendicular to the beam line of the vertex with respect the nominal interaction point $(0, 0, 0)$.

The Monte Carlo samples listed in sec. 4.2 are generated simulating the pile-up conditions, as expected in the 25 ns bunch crossing pile-up scenario. Nevertheless, the MC pile-up description does not match exactly the conditions in data, and there is therefore the need to reweight the simulated events in order to improve the agreement with the data.

The MC samples are reweighted assuming a total inelastic cross section of $\sigma_{in} = 69\,200\mu b$. The comparison between the distributions of primary vertices in data and MC after the pile-up reweighting is applied is shown in fig. 4.13 for an event selection (called inclusive selection, described in sec. 4.4) requiring large amount of \vec{p}_T^{miss} recoiling against an AK8 fat jet (tab. 4.12).

4.3.2 Electrons

Electrons considered in this analysis, reconstructed from energy deposits in the ECAL matched to tracks reconstructed in the silicon tracker, are required to pass the particle flow criteria, and to fall in the ECAL pseudorapidity fiducial range ($|\eta| < 2.5$). The electron identification is defined with a “cut-based” approach. In the isolation definition, the effect of neutral pile-up contributions is considered by taking into account the energy deposits in the calorimeter, estimated through the so-called ρ -area method, by subtracting the median energy density in the event ρ multiplied by the electron energy deposits effective area. The isolation value is computed in a ΔR cone of 0.3 centered along the lepton direction.

Since in this analysis aims at a final state without any lepton, every electron identified with the looser cut-based criteria (*veto Id*) and transverse momentum $p_T > 10\text{ GeV}$ is vetoed. The detailed set of cuts to define a *veto Id* cut-based electron are reported in tab. 4.5; this set of selections allow to identify an electron with an efficiency of $\sim 95\%$. The supercluster width is indicated as $\sigma_{inj\eta}$; $\Delta\eta_{in}^{seed}$ and $\Delta\varphi_{in}$ are the difference in η and φ between the track position as it is measured in the inner layer,

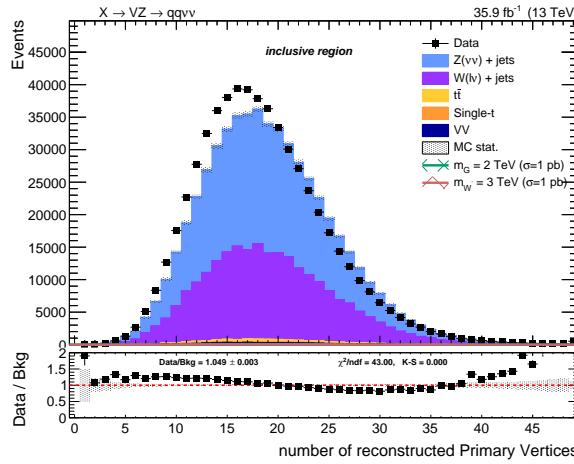


Figure 4.13: Primary vertices distributions in data and MC samples, after reweighting.

and then extrapolated to the interaction vertex and to the calorimeter, and the η of the seed cluster or the φ of the supercluster; H/E is the hadronic leakage, *i.e.* the ratio of the hadronic energy of the calorimetric towers to the electromagnetic energy of the electron supercluster; Iso indicates the particle flow isolation corrected with the effective area approach; $1/E - 1/p$ is the difference of the inverse of the energy and the momentum; d_0 and d_z are the transverse and longitudinal impact parameters. A dedicated conversion veto is applied to mitigate the effects of electrons undergoing bremsstrahlung in the silicon detector.

Table 4.5: Electron cut-based selection for 25 ns bunch spacing conditions. EB: barrel cuts ($|\eta_{\text{supercluster}}| \leq 1.479$); EE: endcap cuts ($|\eta_{\text{supercluster}}| > 1.479$)

Electrons	<i>Veto Id</i>	
	EB	EE
$\sigma_{in in}$	<	0.0115
$\Delta\eta_{in}^{seed}$	<	0.00749
$\Delta\varphi_{in}$	<	0.228
H/E	<	0.356
Iso (Effective Area)	<	0.175
$ 1/E - 1/p $	<	0.299
$ d_0 $	<	0.05
$ d_z $	<	0.10
missing hits	\leq	0.10
conversion veto	yes	yes

4.3.3 Photons

As in the case of electrons, a photon veto is applied in the analysis both for the signal and the control regions. Events are rejected if they contain one (or more) photon with $p_T > 15$ GeV, $|\eta| < 2.5$, passing the *loose* cut-based photon Id, whose definition is reported in tab. 4.6. The isolation cuts (using the ρ -area method for the mitigation of the pile-up) and conversion-safe veto are applied. The isolation value is computed in a ΔR cone of 0.3 and it is corrected for pile-up by subtracting the event-by-event energy density (ρ) times the photon energy deposits effective area.

4.3 Event selection

Table 4.6: Photon cut-based selection for 25 ns bunch spacing conditions. EB: barrel cuts ($|\eta_{\text{supercluster}}| \leq 1.479$); EE: endcap cuts ($|\eta_{\text{supercluster}}| > 1.479$)

Photons	<i>Loose Id</i>	
	EB	EE
H/E	< 0.0597	0.0481
$\sigma_{i\eta i\eta}$	< 0.01031	0.03013
PF ch.had.iso.(ρ -corr)	< 1.295	1.011
PF neu.had.iso.(ρ -corr)	< $10.910 + 0.0148 p_T + 0.000017 p_T^2$	$5.931 + 0.0163 p_T + 0.000014 p_T^2$
PF photon iso.(ρ -corr)	< $3.630 + 0.0047 p_T$	$6.641 + 0.0034 p_T$
conversion veto	yes	yes

1385 **4.3.4 Muons**

1386 The minimal criteria to define a muon is that it must be identified by the particle flow algorithm,
 1387 and should be reconstructed either as a global muon or as a tracker muon (sec. 3.2.9). The muon
 1388 isolation is defined in a cone with a radius of $\Delta R = 0.4$ centered along the lepton direction. In the
 1389 analysis event selection, all muons identified with the loosest criteria previously described, p_T over
 1390 10 GeV, PF isolation below 0.25, $\eta < |2.4|$ are vetoed.

1391 **4.3.5 Taus**

1392 The presence of hadronically decaying taus acts as a veto for the events both in the signal and in
 1393 the control regions, in order to suppress electroweak backgrounds. The selection criteria for taus
 1394 are $p_T > 18$ GeV and $|\eta| < 2.3$. Loose identification criteria of the hadronic tau reconstruction algo-
 1395 rithms are required and applied in order to identify possible tau candidates.

1396 **4.3.6 Jets**

1397 In this analysis, jets are considered if the corrected p_T is larger than 30 GeV for AK4 jets, and larger
 1398 than 200 GeV for AK8 jets, and lie in the tracker acceptance ($|\eta| < 2.4$). The requirement on AK8
 1399 jets transverse momentum is motivated by the fact that $p_T = 200$ GeV is the minimum kinematical
 1400 threshold ensuring to enclose the lighter hadronically decaying vector boson (namely, the W bo-
 1401 son) in the jet cone. Additionally, AK4 jets are required to pass *loose* jet identification requirements,
 1402 AK8 are required to pass *tight* jet identification requirements defined in tab. 4.7. AK8 jets are used
 1403 to reconstruct the hadronically decaying electroweak boson candidate, whilst AK4 jets are used to
 1404 suppress the contribution of top and QCD background events. Jet energy corrections are applied
 1405 to AK4 and AK8 CHS jets. Fig. 4.14- 4.16 show the data/simulation comparison after the analysis
 1406 selections (tab. 4.12 without Top cleaning and Event cleaning).

1407 Since it has been measured that the jet energy resolution (JER) is not the same in data and MC,
 1408 an additional smearing is applied in simulation, in order to get a better agreement. There are two
 1409 independent ways to get the smearing. The scaling method rescales the corrected four-momentum
 1410 of a reconstructed jet by a factor

$$c_{\text{JER}} = 1 + (s_{\text{JER}} - 1) \frac{p_T - p_T^{\text{gen}}}{p_T}, \quad (4.3)$$

1411 where p_T is the transverse momentum of the jet, p_T^{gen} is the transverse momentum of the genera-
 1412 tor level particle corresponding to the reconstructed jet, and s_{JER} is the data-simulation resolution

Table 4.7: *Loose* and *Tight* jet identification requirements for 25 ns bunch spacing conditions.

Particle flow jet ID	<i>Loose</i>	<i>Tight</i>
Neutral Hadron Fraction	< 0.99	< 0.90
Neutral EM Fraction	< 0.99	< 0.90
Number of Constituents	> 1	> 1
Muon Fraction	-	-
Additionally, for $ \eta < 2.4$		
Charged Hadron Fraction	> 0	> 0
Charged Multiplicity	> 0	> 0
Charged EM Fraction	< 0.99	< 0.99

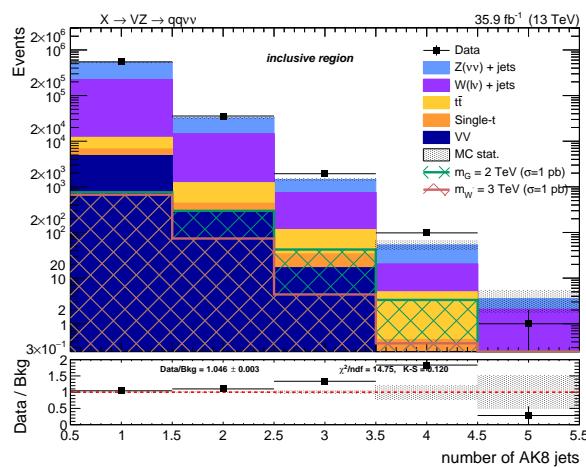
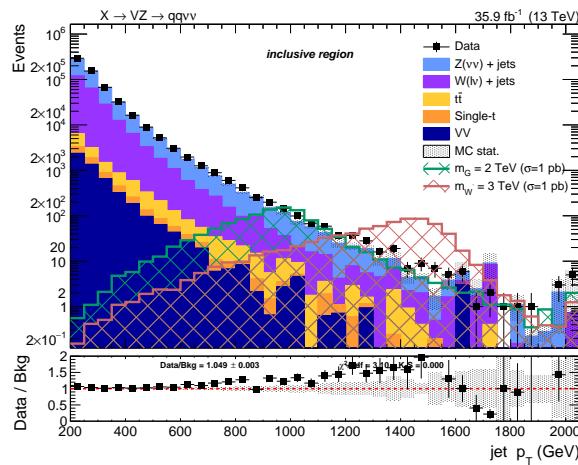


Figure 4.14: Number of reconstructed AK8 jets after inclusive selections.


 Figure 4.15: Leading AK8 jet p_T spectrum after inclusive selections.

1413 scale factor. The factor c_{JER} is positively defined, hence, when negative, it is set equal to zero. The

4.3 Event selection

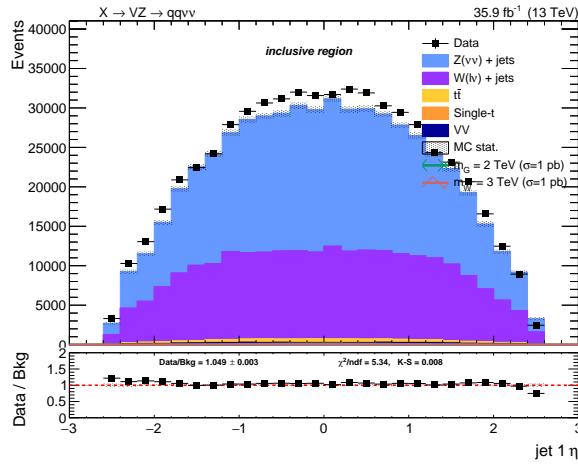


Figure 4.16: Leading AK8 jet η spectra after inclusive selections.

1414 generator level particle and a reconstructed jet are defined as matched if:

$$\Delta R < R_0/2, \quad (4.4)$$

$$|p_T - p_T^{\text{gen}}| < 3 \times \sigma_{\text{JER}} \times p_T,$$

1415 where R_0 is the jet clustering parameter and σ_{JER} is the relative p_T resolution measured in simula-
1416 tion.

1417 The alternative approach is the stochastic smearing, and it does not require the matching with the
1418 generator level particle. The jet four-momentum is rescaled by a factor

$$c_{\text{JER}} = 1 + \mathcal{N}(0, \sigma_{\text{JER}}) \sqrt{\max(s_{\text{JER}}^2 - 1, 0)}, \quad (4.5)$$

1419 where σ_{JER} is the relative p_T resolution in simulation, s_{JER} is the data-simulation scale factor, and
1420 $\mathcal{N}(0, \sigma)$ is a random number extracted from a gaussian normal distribution, whose mean is zero
1421 and variance σ^2 . Scaling factor c_{JER} is positively defined.

1422 The smearing procedure adopted in this analysis is the hybrid method: when a matching jet at
1423 generator level is found, the scaling method is adopted, else the stochastic smearing is chosen. The
1424 smearing coefficients (scale factors, SF) as a function of the jet η and their uncertainties are reported
1425 in tab. 4.8 for 2016 data [69].

1426 4.3.7 Jet mass

1427 The jet mass is the main observable in distinguishing a V jet from a jet produced by colour inter-
1428 action (QCD jets). Jet grooming procedure consists in the suppression of uncorrelated underlying
1429 event, pile-up and soft radiation from the jet: it improves the signal and background discrimina-
1430 tion, by pushing the jet mass for QCD jets towards lower values of the spectrum, while maintaining
1431 the jet mass for V -jets around the electroweak boson mass window.

1432 The grooming technique of the analysis relies on the “soft drop declustering” algorithm, a jet sub-
1433 structure technique that recursively removes soft wide-angle radiation from a jet [89], in order to
1434 mitigate the contaminations from initial state radiation, along with pile-up and multiple scatter-
1435 ings.

1436 The soft drop algorithm starts with a jet clustered with the anti- k_T algorithm with a parameter R_0 ;
1437 the jet is then reclustered with the Cambridge-Aachen method [90], whose definition is included in

Table 4.8: Data-simulation jet smearing coefficients and their corresponding uncertainties.

Jet η	Smearing SF
0.0–0.5	1.109 ± 0.008
0.5–0.8	1.138 ± 0.013
0.8–1.1	1.114 ± 0.013
1.1–1.3	1.123 ± 0.024
1.3–1.7	1.084 ± 0.011
1.7–1.9	1.084 ± 0.011
1.9–2.1	1.140 ± 0.047
2.1–2.3	1.067 ± 0.053
2.3–2.5	1.177 ± 0.041
2.5–2.8	1.364 ± 0.039
2.8–3.0	1.857 ± 0.071
3.0–3.2	1.328 ± 0.022
3.2–5.0	1.16 ± 0.029

1438 eq. 3.10, once the exponent a is set $a = 0$. The soft drop algorithm is ruled by two parameters, a soft
 1439 threshold z_{cut} , that cuts on the energy fraction of soft radiation, and an angular exponent β . The
 1440 procedure is the following:

- 1441 • the jet is declustered into two subjets, j_1 and j_2 , by reverting the final step of Cambridge-
 1442 Aachen algorithm;
- 1443 • if j_1 and j_2 respect the soft drop condition (eq. 4.6), j is defined as the groomed jet;
- 1444 • if they don't pass the condition, the leading subjet in p_T is redefined as the new j ;
- 1445 • if j can't be declustered anymore, it is defined as the groomed jet.

1446 The parameters $z_{\text{cut}} = 0.1$ and $\beta = 0$ are set in the soft drop condition:

$$\frac{\min(p_T^1, p_T^2)}{p_T^1 + p_T^2} > z_{\text{cut}} \left(\frac{\Delta R_{12}}{R_0} \right)^\beta, \quad (4.6)$$

1447 where p_T^1 and p_T^2 are the momenta of the constituents, ΔR_{12} is their angular distance. z_{cut} and β
 1448 parameters affect the degree of jet grooming: if $\beta \rightarrow \infty$ the jet remains ungroomed, while the more
 1449 β approaches zero, the more soft collinear radiation is removed.

1450 The net effect of the soft drop algorithm is studied in Monte Carlo simulations of a W hadronic
 1451 decay process (signal), in association with jets, and of a multi-jet QCD process (background). Jets
 1452 are clustered with the anti- k_T algorithm with a parameter $R_0 = 1$ and asked to have $p_T > 500$ GeV
 1453 and $|\mathcal{Y}| < 4$. The parameter z_{cut} is chosen such in a way that the number of events falling in the W
 1454 mass window ($[70, 90]$ GeV) is the 35% of the total number of events. The results before (black curve)
 1455 and after the application of soft drop algorithm (coloured curves, depending on the value of β) are
 1456 presented in fig. 4.17 [89]. In particular, by comparing the ungroomed jet mass (in black) with the
 1457 mass groomed with a parameter $\beta = 0$ (adopted in this analysis and displayed with a green curve),
 1458 the soft drop mass of the leading jet is a very narrow distribution peaking around the nominal W
 1459 window in the signal sample, whilst it is pushed at lower values in the background sample.

1460 The soft drop algorithm is used in association with the Pile Up Per Particle Identification algorithm
 1461 (PUPPI) [91], designed to combine detector informations in order to compute a local metric α , that

4.3 Event selection

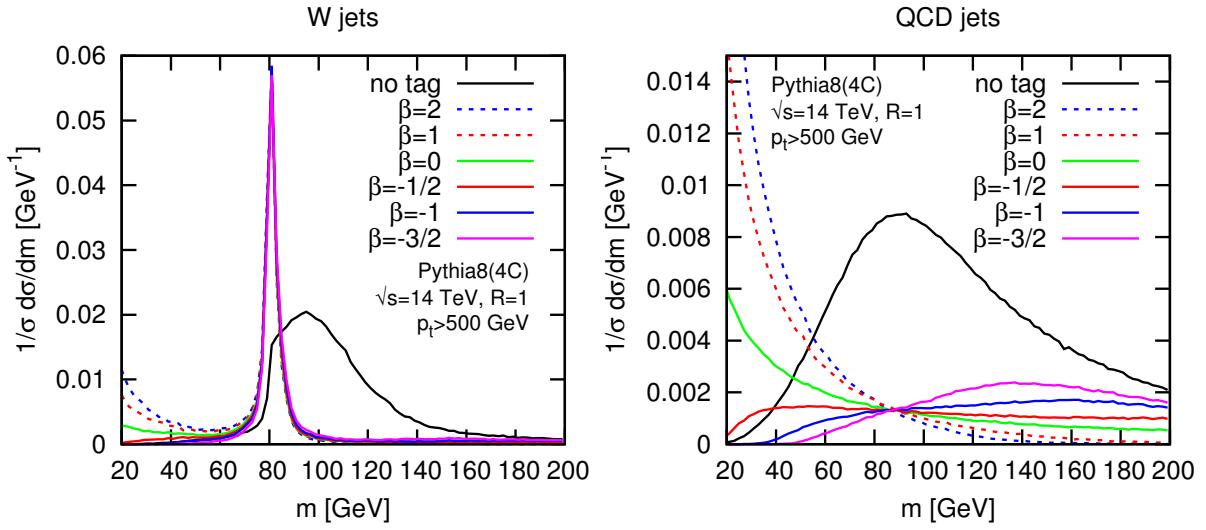


Figure 4.17: Distributions of the jet mass in $W +$ jeta signal simulations (left) and multi-jet QCD background (right), before (in black) and after applying soft drop algorithm. Each curve corresponds to a different value of the parameter β . [89]

describes with a weight how likely it is that one particle is coming from the primary vertex or from a pile-up event. A fundamental feature exploited by the algorithm is the p_T spectrum of the primary vertex particles, expected to be harder than that of the pile-up ones.

The local shape α is defined as:

$$\alpha_i = \log \sum_{j \in \text{event}} \frac{p_{T,j}}{\Delta R_{ij}} \Theta(R_{\min} \leq \Delta R_{ij} \leq R_0), \quad (4.7)$$

where Θ is the Heaviside step function, ΔR_{ij} is the angular distance between the considered i particle and the neighbour j particle, laying in a cone $R_0 = 0.4$ centered around i direction, within a minimum distance $R_{\min} = 0.0001$. Given the softer p_T spectra of pile-up particles, α_i is smaller when i particle does not originate from the primary vertex.

The function

$$\chi^2_i = \Theta(\alpha_i - \bar{\alpha}_{PU}) \frac{(\alpha_i - \bar{\alpha}_{PU})^2}{\sigma_{PU}^2} \quad (4.8)$$

estimates how much α_i fluctuates from the median of the pile-up local shape $\bar{\alpha}_{PU}$ (that has a variance σ_{PU}^2), and it is distributed like a χ^2 with 1 degree of freedom. The PUPPI weight is defined as the cumulative χ^2 distribution $F_{\chi^2, 1 \text{ d.o.f.}}$,

$$w_i = F_{\chi^2, 1 \text{ d.o.f.}}(\chi^2_i). \quad (4.9)$$

If the local metric of a particle is distributed closely to the expected distribution of the pile-up, its weight is $w = 0$. Large fluctuations are more likely related to non pile-up particles, and they receive a weight close to 1. All the particles whose weights are smaller than 0.01 are removed from the jet clustering procedure.

The default soft drop PUPPI jet mass suffers from a systematic shift from the expected value of about $\sim 10\%$, and from some residual dependence on the jet p_T . Further corrections to the jet mass have been applied:

- 1481 1. a p_T -dependent correction to account for a small shift in the generated vector boson mass,
 1482 applied only on simulated samples,
- 1483 2. a p_T - η -dependent correction to the reconstructed jet mass, applied separately for jets in the
 1484 barrel and endcaps regions.

1485 Fig. 4.18- 4.19 show the jet mass for hadronically decaying W or Z bosons in bulk graviton and
 1486 W' signal samples, before and after the correction, without applying any selections. In fig. 4.20,
 1487 the distribution of soft drop PUPPI jet mass is shown for the expected backgrounds of the analysis
 1488 (coloured histograms) and data (black markers), before and after the corrections.

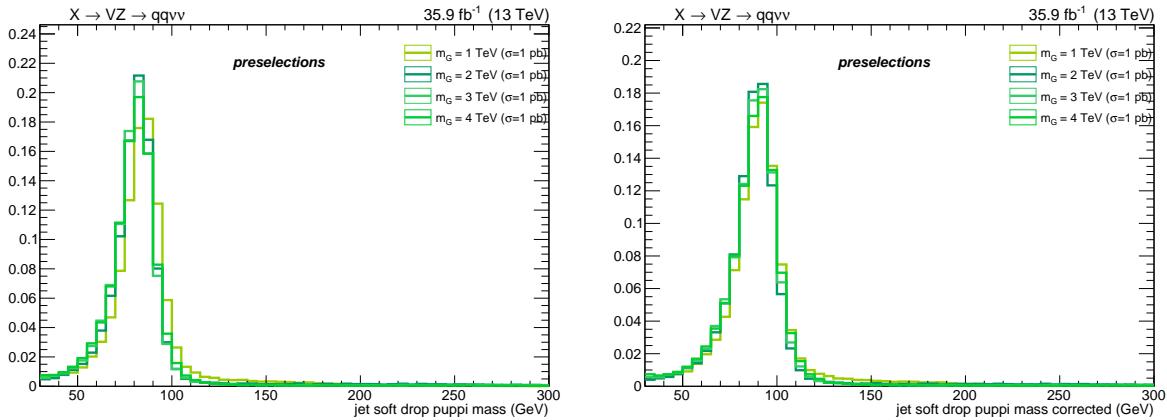


Figure 4.18: Soft drop PUPPI mass of AK8 jet reconstructed for different bulk graviton signal samples, before corrections (left) and after corrections (right).

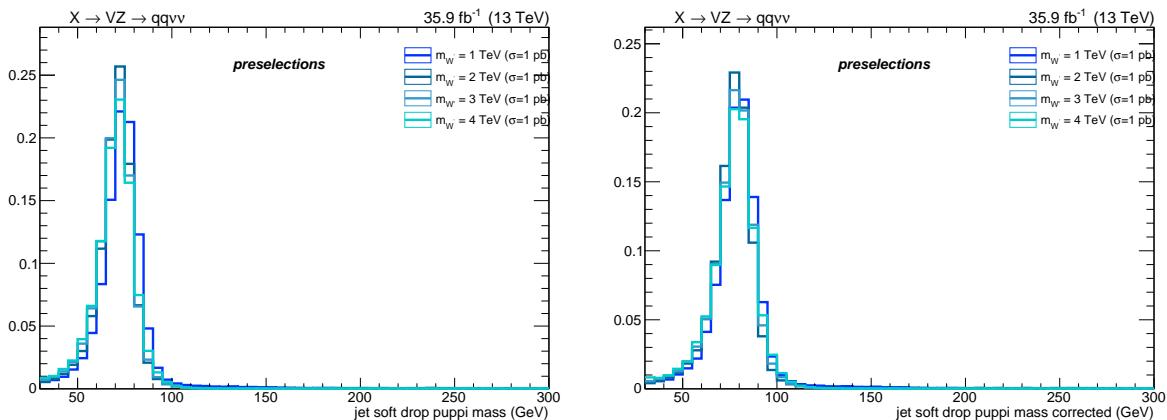


Figure 4.19: Soft drop PUPPI mass of AK8 jet reconstructed for different W' signal samples, before corrections (left) and after corrections (right).

1489 In order to obtain a better data-Monte Carlo agreement, a smearing procedure has been applied to
 1490 the soft drop PUPPI jet mass, by using the stochastic method, with a constant smearing coefficient
 1491 (1.00 ± 0.20), that does not depend on jet pseudorapidity, if it is restricted to $|\eta| < 2.5$.

1492 The selection applied on the jet mass is a crucial step of the analysis, and it has to fulfill three pur-

4.3 Event selection

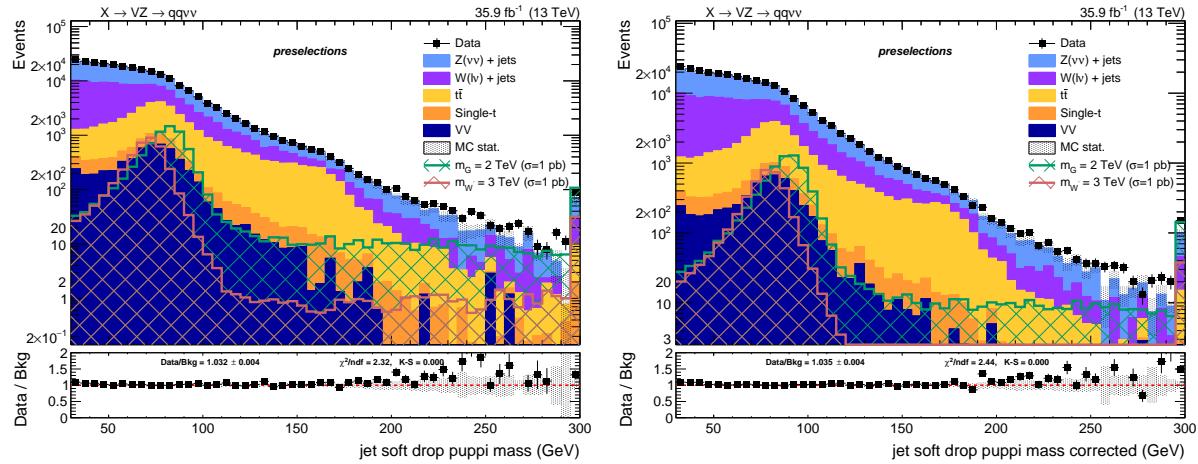


Figure 4.20: Soft drop PUPPI mass of AK8 jet; left: before corrections. right: after corrections.

1493 poses: it has to provide the maximum signal significance (best compromise between signal efficiency and background reduction), it has to avoid overlaps with the Higgs boson mass window, and
 1494 it has to provide a sufficient data and simulation statistics for the control regions (the regions outside
 1495 the mass cut). The soft drop PUPPI mass variable is used to define the following regions:
 1496

Table 4.9: Mass regions defined for the analysis.

	low-sideband	V -region	H -region	high-sideband
M_J	30-65 GeV	65-105 GeV	105-135 GeV	> 135 GeV

1497 The "signal region" (SR) refers to the V -region, where the largest signal yield is expected. The "sidebands" (SB) refer to the low-sideband and high-sideband, where a negligible amount of signal is
 1498 expected. Events with a jet mass value lower than 30 GeV are discarded, because of the high back-
 1499 ground contamination. The jet mass distribution of the V candidate, in the sidebands and in the
 1500 signal region, is shown in fig. 4.21. If the soft drop, PUPPI corrected mass of a large-con jet falls into
 1501 the V -region, the jet is defined as V -tagged.
 1502

4.3.8 Jet substructure

1503 In order to further discriminate signal from background, the inner structure of the jet is investigated.
 1504 Studying the distribution of the jet constituents with respect to the jet axis allows to test the hypoth-
 1505 esis of the existence of multiple substructures, that could be an evidence of jets originated by more
 1506 than one parton. The constituents of the considered jet are clustered again with the k_T algorithm,
 1507 and it is forced to return n subjets. The n -subjettiness [92], τ_n , is defined as
 1508

$$\tau_n = \frac{1}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}^\beta, \Delta R_{2,k}^\beta, \dots, \Delta R_{n,k}^\beta), \quad (4.10)$$

1509 where k labels the particles included in the jet, $p_{T,k}$ is the corresponding transverse momentum of
 1510 the k constituent, and $\Delta R_{i,k}$ is the solid angle between the k constituent and the i subjet candidate.

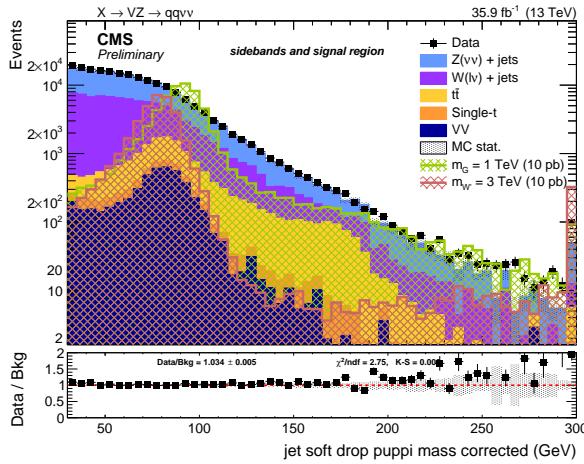


Figure 4.21: Distribution of the soft drop PUPPI corrected mass of the leading AK8 jet, selected as the hadronically decaying V candidate, in the sidebands and control region of the analysis, for expected SM background, bulk graviton signal, W' signal, and data.

1511 The parameter d_0 is a normalization factor:

$$d_0 = \sum_k p_{T,k} R_0, \quad (4.11)$$

1512 where R_0 is the clustering parameter of the considered jet. The τ_n variable describes to what degree
1513 a jet can be considered as composed by n substructures; smaller values of τ_n correspond to higher
1514 compatibility to the n -prong hypothesis. A large-cone jet generated by the hadronic decay of an
1515 electroweak boson is expected to be a 2-prong object, whilst light flavour and gluon jets generated
1516 by colour interaction have a 1-prong monolithic structure. The τ_2 or the τ_1 alone, by the way, do
1517 not provide an optimal signal and background discrimination, as shown in fig. 4.22 (left and center);
1518 by looking at fig. 4.22 (right), it is clear that the most powerful discriminating variable is their ratio
1519 $\tau_{21} = \tau_2 / \tau_1$:

$$\tau_{21} = \frac{\frac{1}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}, \Delta R_{2,k})}{\frac{1}{d_0} \sum_k p_{T,k} \Delta R_{1,k}}. \quad (4.12)$$

1520 In fig. 4.23, the distributions of the τ_{21} variable are displayed for background and data, after applying
1521 the PUPPI algorithm (left), and for different bulk graviton mass hypotheses (right). The signal
1522 distribution is expected to peak at low values of the τ_{21} subjettness variable.

1523 The τ_{21} variable is used to classify the events into two exclusive categories, in order to improve
1524 the signal discovery reach. Events are included in either the high-purity ($\tau_{21} < 0.35$) or low-purity
1525 ($0.35 < \tau_{21} < 0.75$) category.

1526 The choice of the τ_{21} categorization listed above is based on a study of the analysis sensitivity. An-
1527 other τ_{21} categorization is probed, according to which events are grouped into different high-purity
1528 ($\tau_{21} < 0.40$) and low-purity ($0.40 < \tau_{21} < 0.75$) categories. This different set of τ_{21} cuts has been
1529 tested, along with that chosen for this analysis. Two figures of merit are considered: the discovery
1530 reach, namely the bulk graviton signal significance (displayed in fig. 4.24), and the expected exclu-
1531 sion limit on cross-section times branching fraction at 95% CL (displayed in fig. 4.25), as a function
1532 of the reconstructed transverse mass of the resonance. To this purpose, the entire analysis workflow

4.3 Event selection

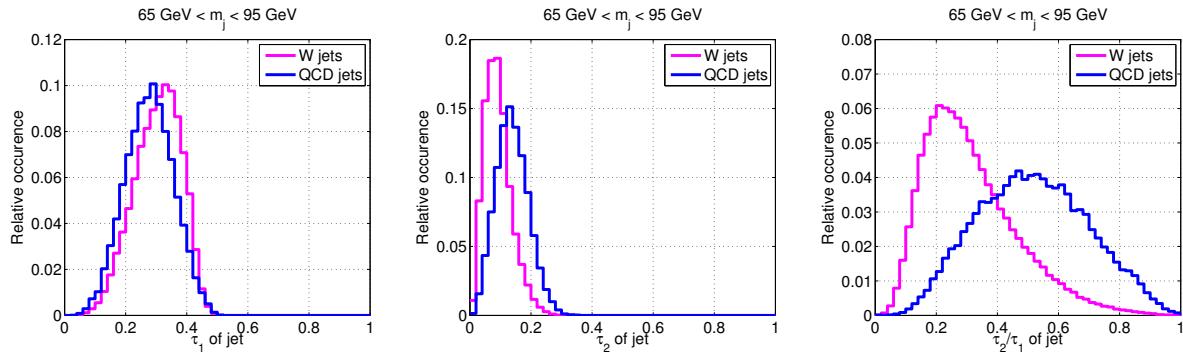


Figure 4.22: Distribution of τ_1 (left), τ_2 (center), and τ_{21} (right) variables, in simulations of a W plus jets process (in pink) and for a multi-jet QCD originated process (in blue). A selection on the leading jet mass is applied: $65 < m_j < 95 \text{ GeV}$; jets are clustered with a parameter $R_0 = 0.6$, $p_T > 300 \text{ GeV}$, $|\eta| < 1.3$ [92].

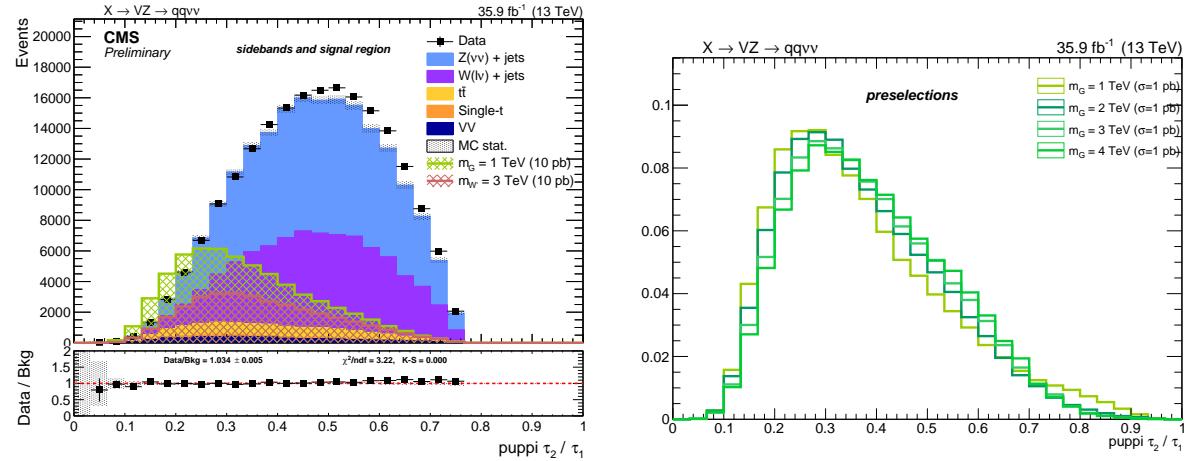


Figure 4.23: Distribution of the τ_{21} subjetiness of the leading AK8 jet, selected as the hadronically decaying V candidate, for expected SM background and data (left), and for bulk graviton signal (right).

1533 has been applied, performing an unbinned shape analysis with the analysis background estimation
 1534 method, taking into account all the systematic uncertainties. In each figure, on the left, the figure of
 1535 merit is plotted separately for each purity category, while in the right part of the figures the low and
 1536 purity categories are combined together. Significance has been computed with a limited number of
 1537 toys (100), hence the curves are non perfectly smooth, while the exclusion limit has been computed
 1538 with the asymptotic formula. The procedures to extract signal significance and exclusion limits are
 1539 described in sec. 4.7. Considering that the search region is 1-4 TeV, the choice of 0.35-0.75 τ_{21} work-
 1540 ing points is legitimated.

1541 When doing the τ_{21} categorization, V -tagging scale factors have been taken into account to cor-
 1542 rect data and simulation discrepancies introduced by the n -subjettiness. They are described in
 1543 sec. 4.3.8.1.

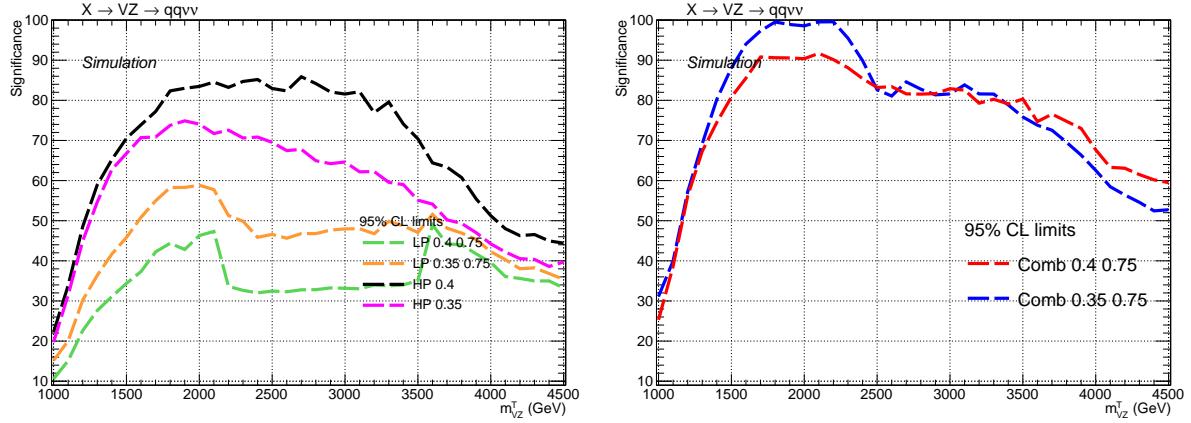


Figure 4.24: Analysis sensitivity to bulk graviton signals, computed by applying different τ_{21} categorizations, considering the categories separately (left) and combining them together (right), as a function of the resonance mass.

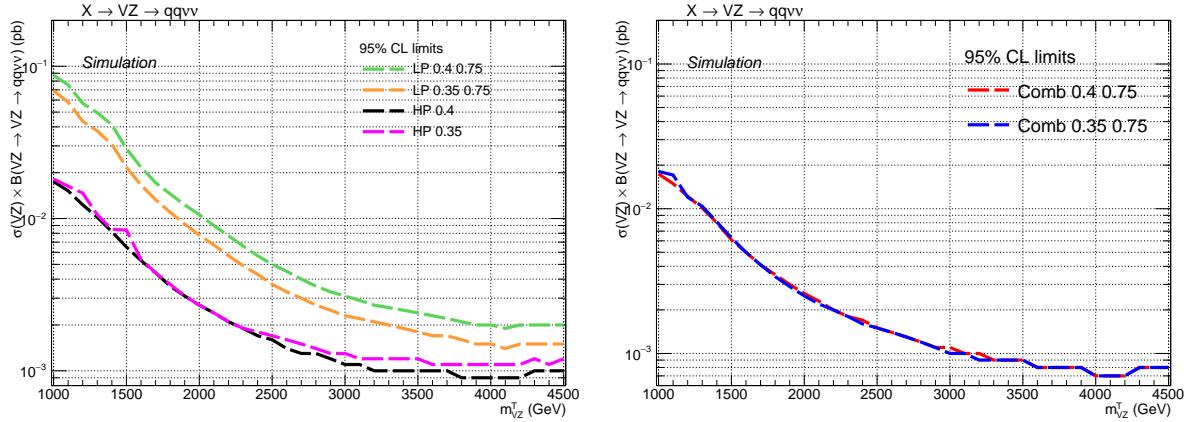


Figure 4.25: Exclusion limit on cross-section time branching fraction at 95% CL of bulk graviton signals, computed by applying different τ_{21} categorizations, considering the categories separately (left) and combining them together (right), as a function of the resonance mass.

1544 4.3.8.1 Corrections induced by jet substructure variables

1545 By applying a selection on the jet τ_{21} , the jet mass spectrum is sculpted, hence the effects of the
 1546 V -tagging procedure shall take into account both the selections on mass and on substructure si-
 1547 multaneously. The distributions of the groomed jet mass and τ_{21} subjetiness have been compared
 1548 in data and simulations, by selecting samples of di-jet, $t\bar{t}$ and $W +$ jets events, and a significative
 1549 discrepancy has been observed (10%) [83]. Scale factors are extracted by selecting a $t\bar{t}$ sample in
 1550 data, because an high p_T W boson is produced by the top quark decay. The hadronically decaying
 1551 W boson is tagged by choosing events where the soft drop mass of a large-cone jet lies in a win-
 1552 dow centered around the nominal W mass. The jet mass distributions of events passing and failing
 1553 the selection on the τ_{21} variable ($\tau_{21} < 0.35$ and $0.35 < \tau_{21} < 0.75$, considered separately) are fit-
 1554 ted simultaneously, both in data and in simulations. The V -tagging scale factors are defined as the
 1555 ratio of the τ_{21} categorization efficiencies in data and MC, and they are summarized in tab. 4.10.

4.3 Event selection

1556 The systematic uncertainties depend on the simulation of the $t\bar{t}$ process, they cover the discrepancies
1557 observed while using different Monte Carlo simulations, and due to the choice of the fitting
1558 function.

Table 4.10: Data-simulation scale factors, calculated on $t\bar{t}$ samples, that correct the discrepancies related to the τ_{21} categorization.

τ_{21} selection	Purity category	Data-MC scale factor
$\tau_{21} < 0.35$	high-purity	0.99 ± 0.11
$0.35 < \tau_{21} < 0.75$	low-purity	1.03 ± 0.23

1559 **4.3.9 b-tagging**

1560 The presence of a b-tagged quark can be an hint to identify the top quark decays, representing a
1561 potential background to the search. The CSV b-tagging algorithm [72] is applied to the AK4 jets.
1562 The jet is considered as tagged if the CSV discriminator value is above a threshold value; the b-tag
1563 efficiency is defined as the number of jets fulfilling this requirement, divided by the total number of
1564 jets. Since the purpose of the b-tagging is to reject the top quark events, the working point with the
1565 largest efficiency is chosen; the threshold of the CSV multivariate discriminant is listed in tab. 4.3.9.

Table 4.11: Working point for CSV b-tagging algorithm.

Working point	CSV discriminant threshold	tagging efficiency	mis-tag probability
CSVLoose	> 0.5426	~ 85%	~ 10%

1566 Events where an AK4 jet, not laying in the AK8 jet cone, is b-tagged with the *loose* working point
1567 threshold, are rejected. This veto allows to suppress the single-top events and $t\bar{t}$ events by one half.
1568 The b-tagging efficiency is not the same in data and MC. In order to take into account this difference,
1569 b-tagging scale factors for b-jets and mis-tagged light jets, measured for different physics processes,
1570 are calculated. A weight is extracted on a per-event basis, as a function of the b-tagging status of the
1571 jets and their kinematic variables [93].

1572 **4.3.10 Missing Energy**

1573 As pointed out in sec. 3.2.9.8, Type-I corrected E_T^{miss} is used in the analysis, along with dedicated
1574 filters to remove detector noise and events with bad reconstruction. In order to lie in the plateau
1575 of the trigger efficiency, $E_T^{\text{miss}} > 200\text{GeV}$. Fig. 4.26 shows the E_T^{miss} distribution for data and Monte
1576 Carlo after the corrections and filters.

1577 **4.3.11 Diboson candidate reconstruction**

1578 **4.3.11.1 $V \rightarrow q\bar{q}$ reconstruction**

1579 The identification of jets produced by the hadronic decays of one vector boson is based on the two
1580 concepts:

- 1581 • *Jet mass*: jets produced by the decay of a massive particle should have an invariant mass
1582 around the nominal mass of the original particle. Oppositely, jets originated by QCD radi-
1583 ation are produced by the emission of quarks or gluons and typically have smaller invariant
1584 mass. This effect is further enhanced by the grooming techniques (sec. 4.3.7).

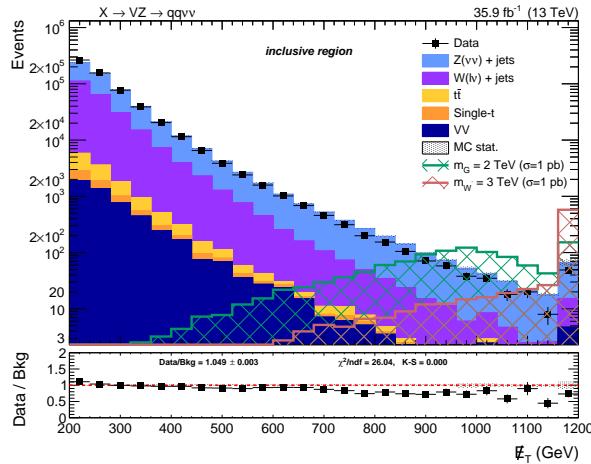


Figure 4.26: Type-1 corrected E_T^{miss} distribution after inclusive selections.

- *Jet substructure:* looking inside the structure of jets gives an handle in discriminating the original seed of the jet. Z and W -jets are produced by two partons merged into a single large-cone jet.
- 1588 The leading AK8 jet respecting the jet mass and jet substructure selections is tagged as the V candidate.

1590 4.3.11.2 $Z \rightarrow \nu\bar{\nu}$ reconstruction

1591 If the Z boson decays into a pair of neutrinos, no product is visible in the detector, hence the
1592 invisible decay of the Z boson is determined only by its transverse component, namely by the E_T^{miss} .

1593 4.3.11.3 Composite VZ candidate reconstruction

1594 Given that the longitudinal component of the Z boson is unknown, a simple and effective solution
1595 is to consider the transverse mass of the VZ candidate, using the jet and \vec{p}_T^{miss} kinematics, defined
1596 by the following formula:

$$m_{VZ}^T = \sqrt{2E_T^V E_T^{\text{miss}} \cdot (1 - \cos \Delta\varphi(V, \vec{p}_T^{\text{miss}}))}, \quad (4.13)$$

1597 where E_T^V is the transverse energy of the V candidate (defined in sec. 3.2.1), and $\Delta\varphi$ is the angle
1598 between the V and the Z candidates in the transverse plane.

1599 4.3.12 Final analysis selections

1600 Events considered in this analysis have to pass a certain number of selections before being con-
1601 sidered as suitable signal candidates, both in data and in simulations. The selections are reported
1602 below and in tab. 4.12. The selections applied to group the events in purity category, defined on the
1603 PUPPI corrected τ_{21} subjetiness variable (sec. ??), and into signal or control region, defined on the
1604 PUPPI corrected soft drop mass (sec. 4.3.7) are reported in tab. 4.13. The final signal efficiency is
1605 shown separately in purity categories in fig. 4.3.12.3, for both spin-2 and spin-1 signal hypotheses.

4.3 Event selection

4.3.12.1 Z candidate selections

- *Trigger:* HLT_PFMETNoMu90_PFMHTNoMu90_IDTight or HLT_PFMETNoMu110_PFMHTNoMu110_IDTight or HLT_PFMETNoMu120_PFMHTNoMu120_IDTight or HLT_PFMET170_NoiseCleaned or HLT_PFMET170_JetIdCleaned or HLT_PFMET170_HBHECleaned (required in data only);
- $E_T^{miss} > 200 \text{ GeV};$
- *Corrections:* Type-I, noise filters.

4.3.12.2 V candidate selections

- p_T : at least one AK8 Particle-Flow jet with $p_T > 200 \text{ GeV}$;
- η : $|\eta| < 2.4$;
- *Identification:* tight Particle-Flow Id;
- *charged hadron fraction:* chf > 0.2 ;
- *neutral hadron fraction:* nhf < 0.9 ;
- *Mass:* soft drop PUPPI corrected mass $> 30 \text{ GeV}$;
- *Substructure:* PUPPI corrected τ_{21} subjetttines, depending on the category $\tau_{21} < 0.35$ for high-purity, $0.35 < \tau_{21} < 0.75$ for low-purity.

4.3.12.3 Topology and event cleaning

Minimal requirements are applied to objects that are vetoed:

- *Veto on electrons:*
 - p_T : $p_T > 10 \text{ GeV}$;
 - η : $|\eta| < 2.5$;
 - *Id: veto* cut-based working point;
- *Veto on muons:*
 - p_T : $p_T > 10 \text{ GeV}$;
 - η : $|\eta| < 2.4$;
 - *Id: loose* Id;
 - *Isolation:* Particle-Flow Isolation < 0.25 ;
- *Veto on hadronic taus:*
 - p_T : $p_T > 18 \text{ GeV}$;
 - η : $|\eta| < 2.4$;
 - *Id: loose* Id;
- *Veto on photons:*
 - p_T : $p_T > 15 \text{ GeV}$;

- 1638 – $\eta: |\eta| < 2.5$;
- 1639 – *Id: loose cut-based working point.*
- 1640 Further selections are applied to suppress spurious events.
- 1641 • *Event cleaning:* events where the V and the Z candidates are collinear are rejected:
 1642 $\Delta\varphi(V, \vec{p}_T^{\text{miss}}) > 2$.
- 1643 • *Top rejection:* as discussed in sec. 4.3.9, a b-tag veto is imposed on AK4 jets lying outside the
 1644 AK8 cone; this reduces the top quark background contamination by 50%.
- 1645 • *QCD rejection:* a minimum angular separation $\Delta\varphi > 0.5$ is imposed in the transverse plane be-
 1646 tween the \vec{p}_T^{miss} vector and the momenta of all the AK4 jets in the event, lying outside the AK8
 1647 cone and not tagged as b-quark initiated jets. The effect of this cut is to suppress the multi-jet
 1648 QCD background: it has been studied by considering additional QCD simulated samples to
 1649 the analysis backgrounds. As it can be inferred by looking at the distribution of the minimum
 1650 azimuthal separation between \vec{p}_T^{miss} and the AK4 jets, shown in fig. 4.27 (where looser selec-
 1651 tions are applied w.r.t. the nominal selections of the analysis, *i.e.*, no QCD event cleaning is
 1652 performed), if a minimum $\Delta\varphi = 0.5$ threshold is imposed, the QCD contribution is reduced
 1653 from 32% to 5%. In the final signal region, the QCD event yield amounts to 2%, and hence it
 1654 is negligible (3% in low-purity, less than 1% in high-purity).

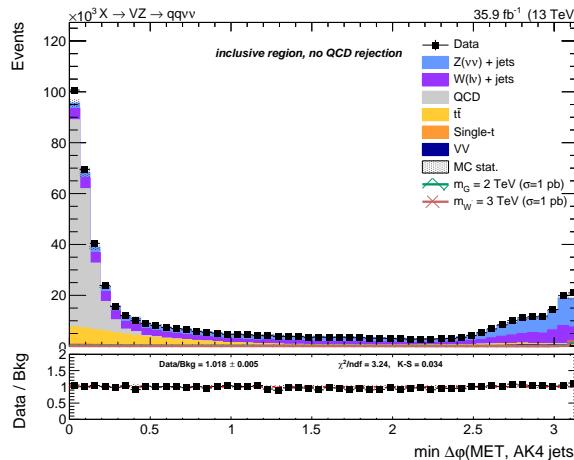


Figure 4.27: Distribution of the minimum azimuthal separation bewteen \vec{p}_T^{miss} and the momenta of all the AK4 jets present in each event. By imposing $\min\Delta\varphi > 0.5$, the QCD background (in gray) is suppressed.

- 1655 The final selections of the analysis are summarized in tab. 4.12-4.13. The detection efficiencies due
 1656 to each cut sequentially applied to bulk graviton signal samples (fig. 4.3.12.3, left) and W' signal
 1657 samples (fig. 4.3.12.3, right) are shown. The signal efficiency for bulk graviton ranges from $\sim 30\%$
 1658 at 1 TeV, up to 20% at 4.5 TeV for low-purity category, whilst it's around 20% for the high-purity
 1659 category in the whole mass range. The signal efficiency for W' ranges from $\sim 40\%$ at 1 TeV, up to 25%
 1660 at 4.5 TeV for low-purity category, whilst it's around 25% for the high-purity category in the whole
 1661 mass range. The different detection efficiencies for the two signals are related to their production
 1662 mechanisms: the graviton is produced in gluon fusion, hence more hadronic activity is expected

4.3 Event selection

Table 4.12: Summary of the selection cuts for the $VZ \rightarrow \nu\nu qq$ analysis.

	$Z \rightarrow \nu\nu$
Trigger	HLT_PFMETNoMu90_PFMHTNoMu90_IDTight or HLT_PFMETNoMu110_PFMHTNoMu110_IDTight or HLT_PFMETNoMu120_PFMHTNoMu120_IDTight or HLT_PFMET170_NoiseCleaned or HLT_PFMET170_JetIdCleaned or HLT_PFMET170_HBHECleaned
E_T^{miss}	Type-I corrected $> 200 \text{ GeV}$
Veto	e, μ, τ, γ
V	$p_T > 200 \text{ GeV, tight Id}$ $\text{nhf} < 0.8; \text{chf} > 0.2$
QCD cleaning	$\min\Delta\varphi(\text{AK4 jets}, \vec{p}_T^{\text{miss}}) > 0.5$
Top cleaning	veto on b-tagged AK4 jets outside the AK8 cone, <i>loose</i> working point (< 0.460)
Event cleaning	$\Delta\varphi(V, \vec{p}_T^{\text{miss}}) > 2$

Table 4.13: Cuts to categorize the $VZ \rightarrow \nu\nu qq$ analysis events into low- and high-purity categories, and into signal region and sidebands.

	$Z \rightarrow \nu\nu$
V mass	Signal Region: $65 < m_V < 105$ Side Bands: $30 < m_V < 65, m_V > 135 \text{ GeV}$
$V \tau_{21}$	$0.35 < \tau_{21} < 0.75$ for low-purity $\tau_{21} < 0.35$ for high-purity

around the VZ decay process, and this results as a loss of efficiency when the QCD rejection cut is applied.

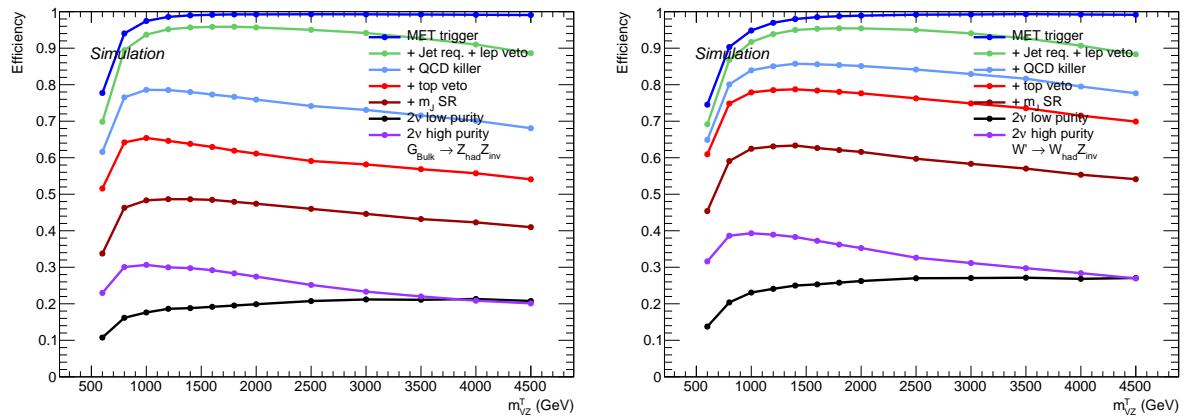


Figure 4.28: Signal efficiency for a spin-2 bulk graviton decaying into a pair of Z bosons (left), and for a spin-1 W' decaying into a W and a Z bosons (right), as a function of the mass of the heavy particle. The efficiencies are separated by purity category after the signal region selections.

1665 4.4 Data and simulations comparison

1666 In this section, a comparison between data and simulation is reported for various kinematic ob-
 1667 servables. It can be seen that the dominant background contribution comes from the $Z + \text{jets}$ and
 1668 $W + \text{jets}$ production, while sub-leading contributions from top ($t\bar{t}$ and single-top) production and
 1669 dibosons can be minor yet non-negligible.

1670 In the following plots (fig. 4.29-4.38), the comparison is performed in three different regions. On
 1671 top of the selections defined above, additional criteria are defined:

- 1672 • *Inclusive*: no selection is performed on top of the selections, except for a veto on the jet mass
 1673 $65 < m_V < 135 \text{ GeV}$ to avoid potential signal contamination from VZ signals;
- 1674 • *Sidebands (SB)*: only events in the sidebands, defined the interval between $30 < m_V < 65 \text{ GeV}$
 1675 and $m_V > 135 \text{ GeV}$ are considered. This region can be considered as signal-depleted. The
 1676 main difference with the previous regions is that the bulk of the jet mass distribution, peaking
 1677 at $m_V \sim 20 \text{ GeV}$, is not included. The region selected is thus much closer kinematically to the
 1678 signal region.
- 1679 • *Signal region (SR)*: it represents the phase space where is signal is expected.

1680 A summary of the number of expected events from Monte Carlo simulations, per each sample, along
 1681 with the number of events observed in data in each category is reported in tab. 4.14. No significant
 1682 excess is observed in data distributions with regards to simulation predictions in signal region.

Table 4.14: Expected background yields and number of events observed in data.

cut	inclusive	SB low-purity	SB high-purity	SR low-purity	SR high-purity
data	586318.00	107363.00	13967.00	44989.00	23074.00
$Z + \text{jets}$	320996.11	57551.99	7774.40	22933.14	10763.87
	57%	56%	56%	53%	45%
$W + \text{jets}$	224607.51	40447.51	5197.74	16248.78	7428.42
	40%	40%	37%	38%	31%
$t\bar{t}$	6308.09	2599.53	670.29	2482.38	3035.21
	1%	3%	5%	6%	13%
VV	5168.06	1075.75	206.54	1283.63	2053.19
	1%	1%	1%	3%	9%
single-top	1968.65	431.28	79.27	329.71	461.84
	<1%	<1%	1%	1%	2%
BkgSum	559048.42	102106.07	13928.25	43277.64	23742.54

4.4 Data and simulations comparison

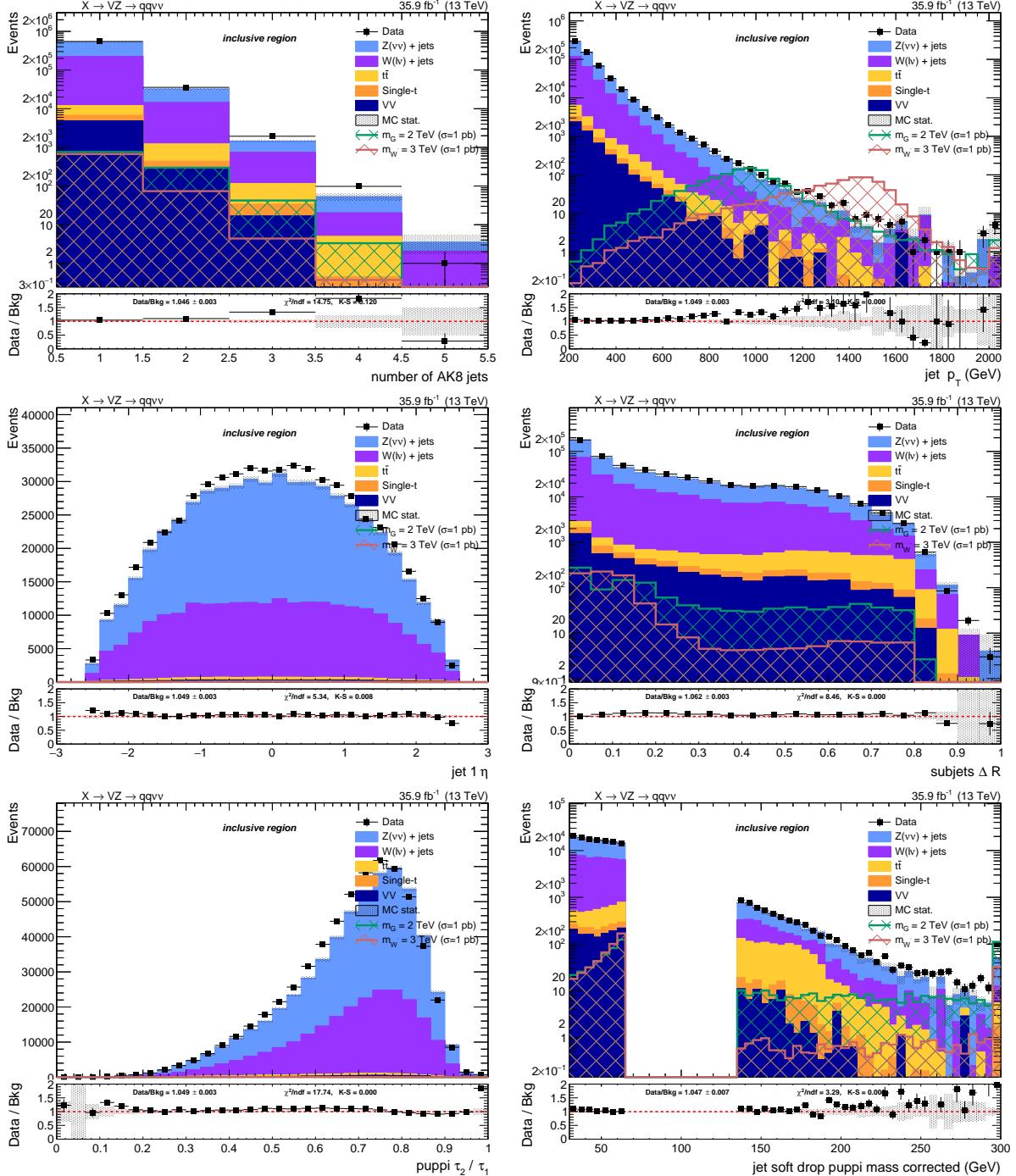


Figure 4.29: Top: number of AK8 jets in the event (left) and V jet candidate p_T (right). Center: V jet candidate η (left) and angular separation ΔR between the constituents leading subjets (right). Bottom: V jet candidate τ_{21} subjetiness after PUPPI correction (left) and V jet candidate soft drop PUPPI mass (right). Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

Search for diboson resonances in the $VZ \rightarrow q\bar{q} \nu\bar{\nu}$ final state

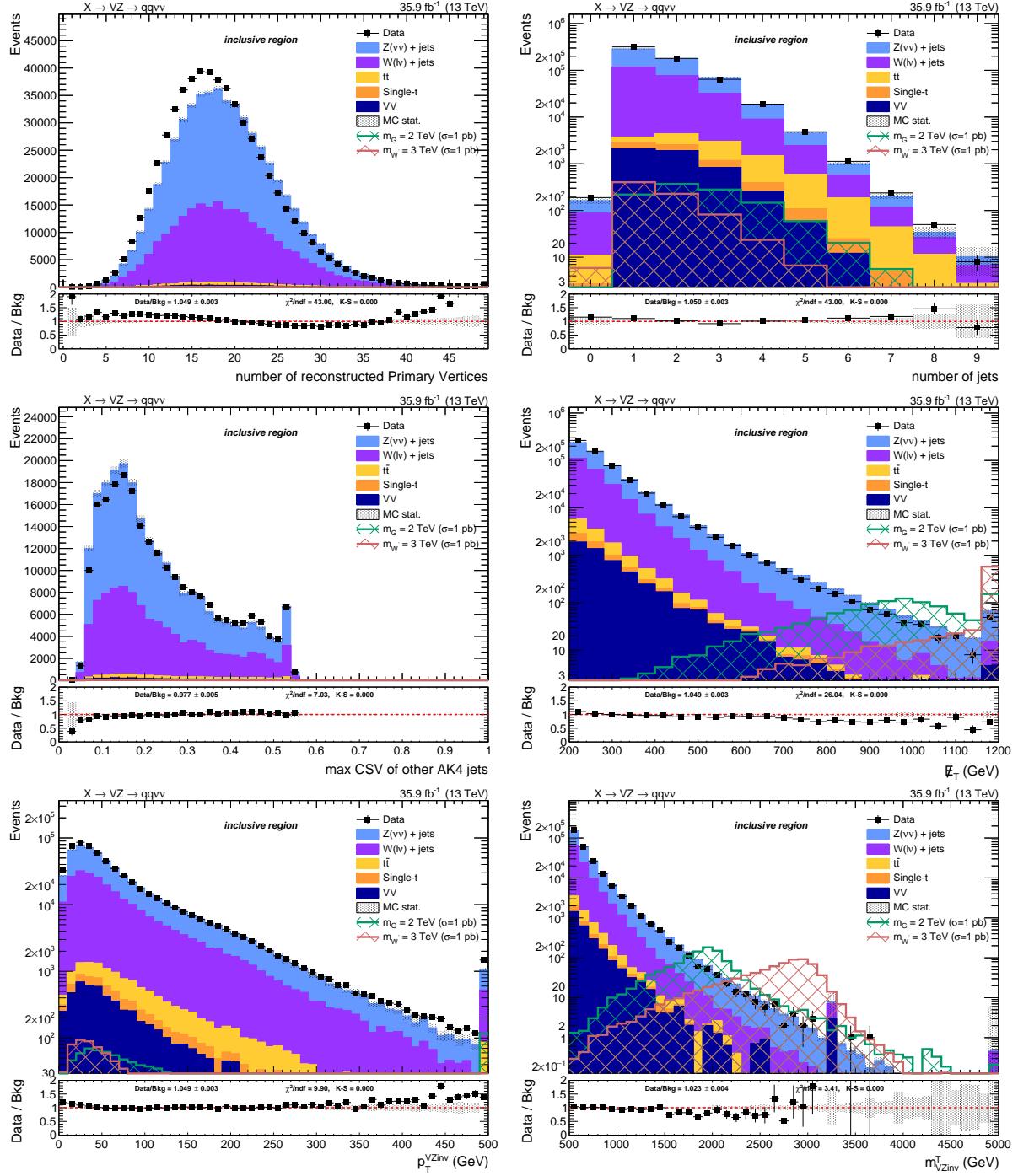


Figure 4.30: Top: number of reconstructed primary vertices (left) and number of AK4 jets in the event (right). Center: distribution of the b-tagging multivariate discriminant for the AK4 jets not included in the V jet cone (left) and E_T^{miss} distribution (right). Bottom: p_T of the VZ candidate (left) and transverse mass of the VZ candidate (right). Events are selected with the *inclusive* selection, and simulated backgrounds are normalized to luminosity.

4.4 Data and simulations comparison

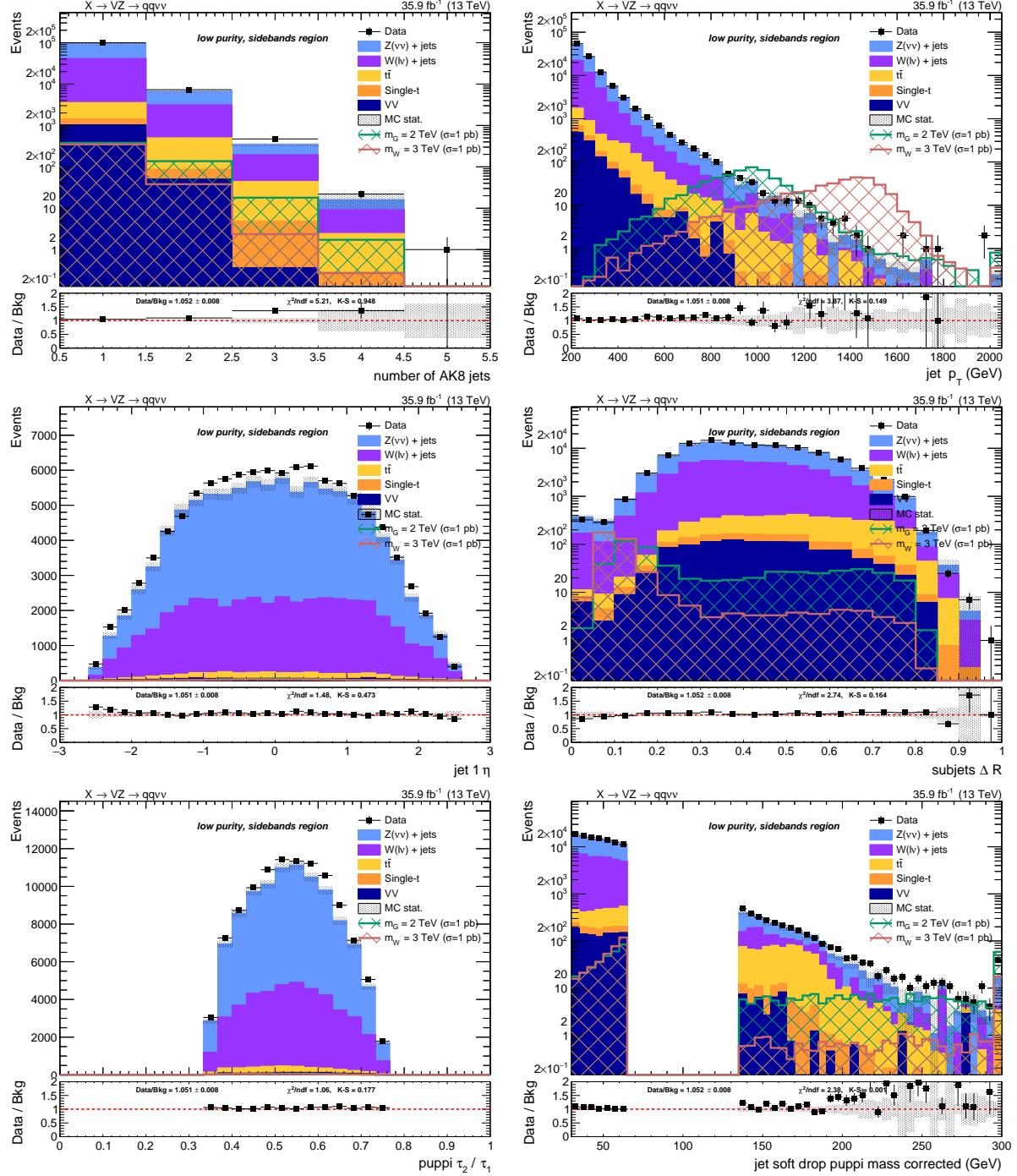


Figure 4.31: Top: number of AK8 jets in the event (left) and V jet candidate p_T (right). Center: V jet candidate η (left) and angular separation ΔR between the constituents leading subjets (right). Bottom: V jet candidate τ_{21} subjettiness after PUPPI correction (left) and V jet candidate soft drop PUPPI mass (right). Events are selected with the *low-purity sidebands* selection, and simulated backgrounds are normalized to luminosity.

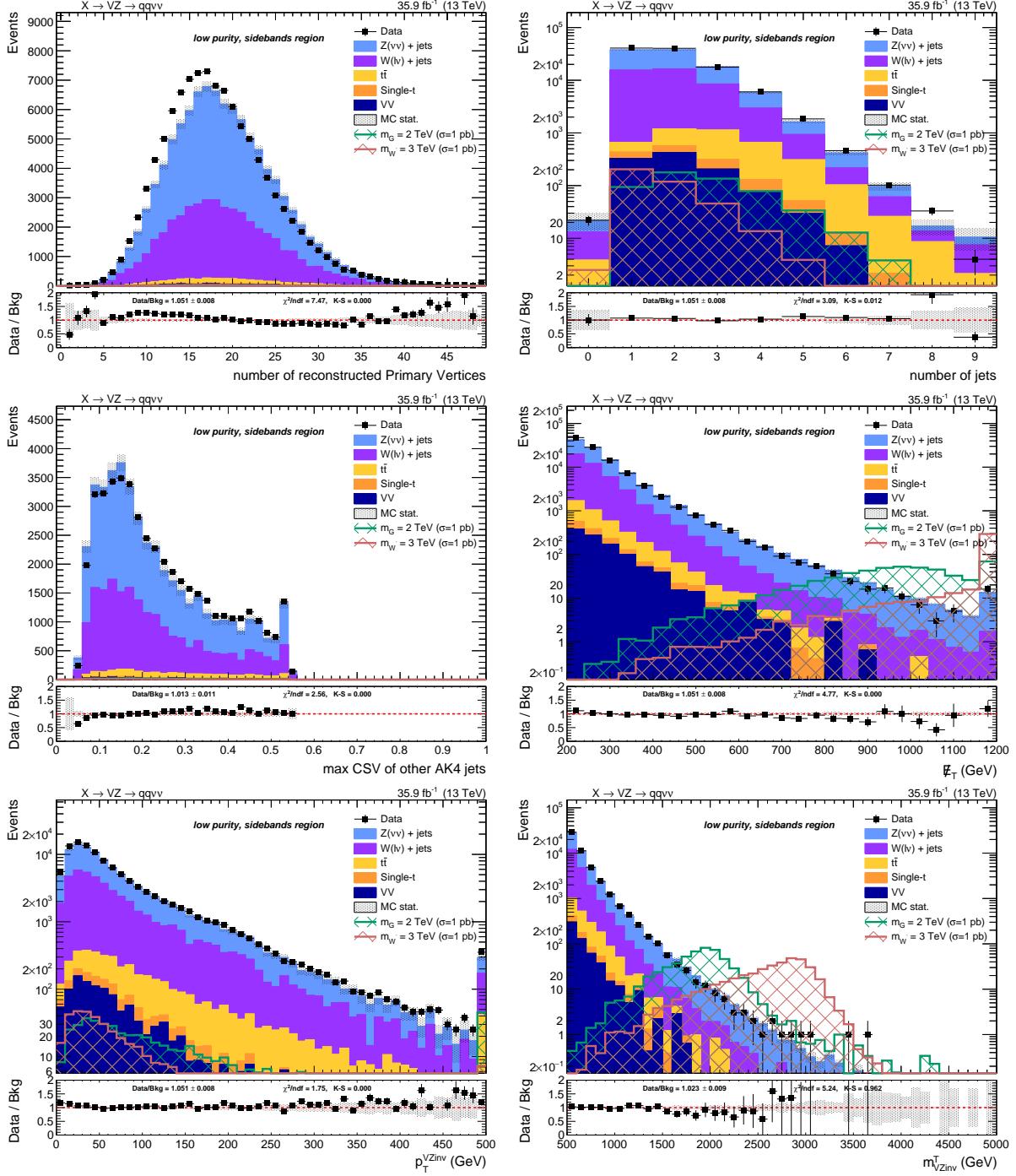


Figure 4.32: Top: number of reconstructed primary vertices (left) and number of AK4 jets in the event (right). Center: distribution of the b-tagging multivariate discriminant for the AK4 jets not included in the V jet cone (left) and E_T^{miss} distribution (right). Bottom: p_T of the VZ candidate (left) and transverse mass of the VZ candidate (right). Events are selected with the *low-purity sidebands* selection, and simulated backgrounds are normalized to luminosity.

4.4 Data and simulations comparison

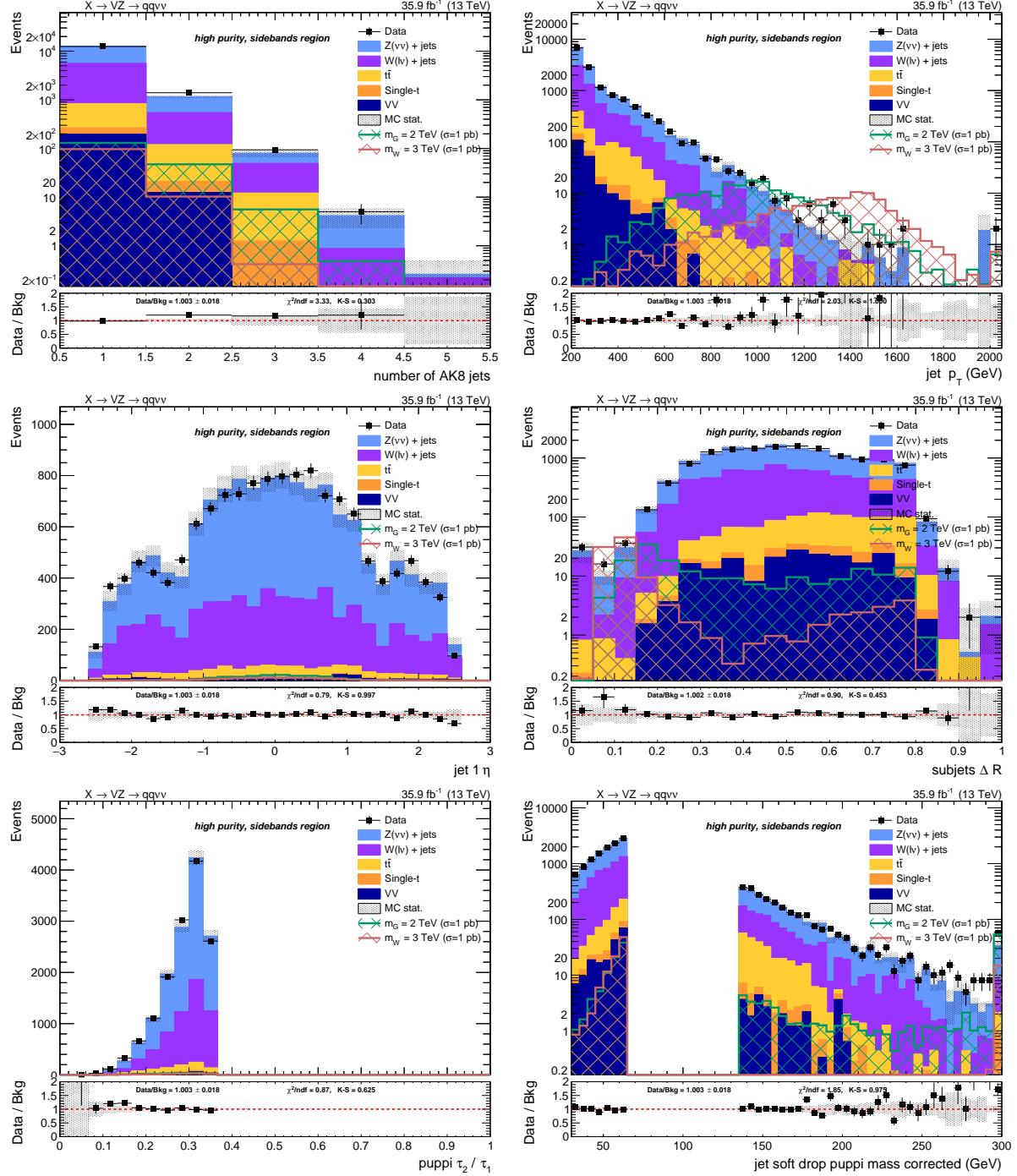


Figure 4.33: Top: number of AK8 jets in the event (left) and V jet candidate p_T (right). Center: V jet candidate η (left) and angular separation ΔR between the constituents leading subjets (right). Bottom: V jet candidate τ_{21} subjetiness after PUPPI correction (left) and V jet candidate soft drop PUPPI mass (right). Events are selected with the *high-purity sidebands* selection, and simulated backgrounds are normalized to luminosity.

Search for diboson resonances in the $VZ \rightarrow q\bar{q} \nu\bar{\nu}$ final state

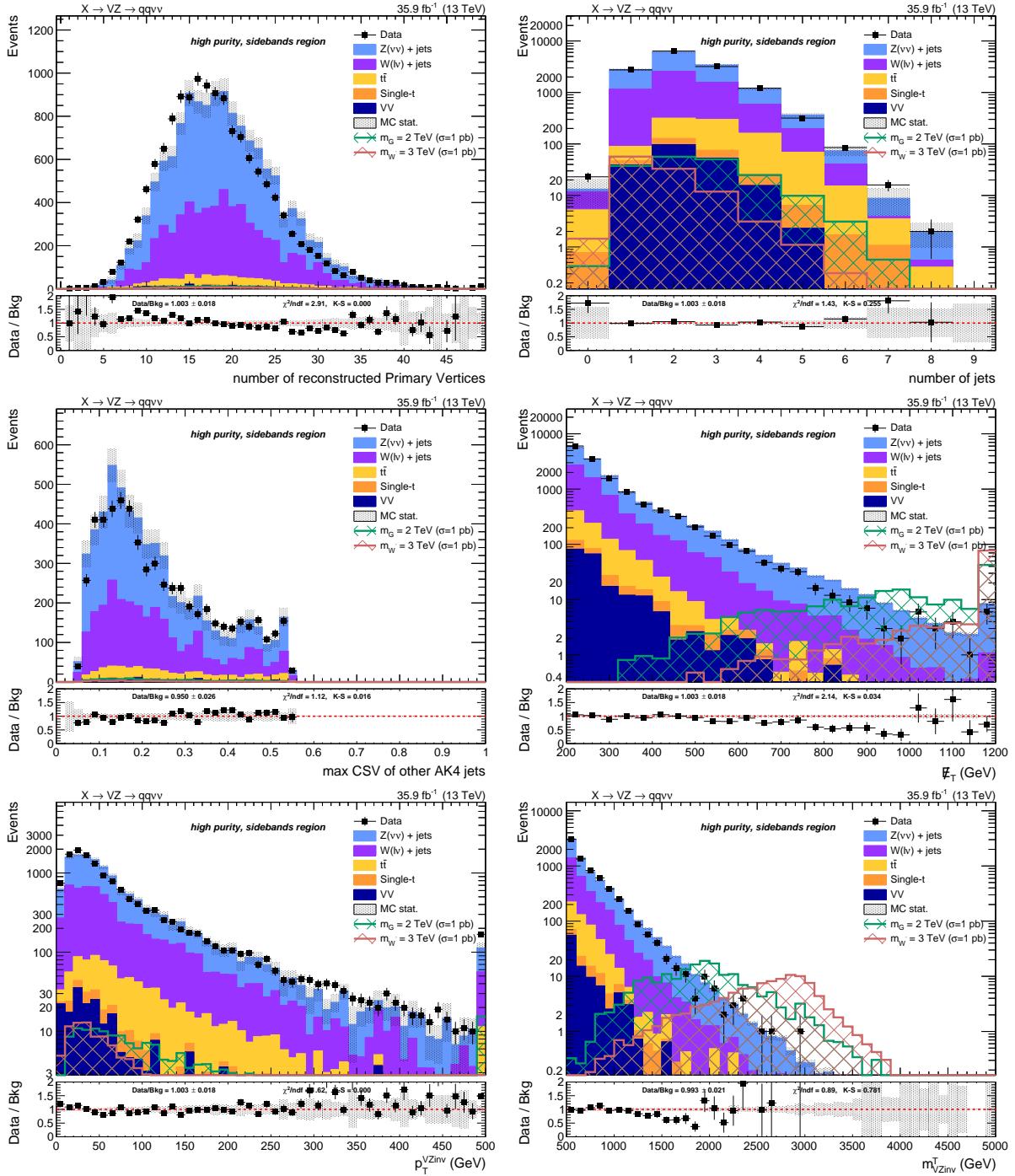


Figure 4.34: Top: number of reconstructed primary vertices (left) and number of AK4 jets in the event (right). Center: distribution of the b-tagging multivariate discriminant for the AK4 jets not included in the V jet cone (left) and E_T^{miss} distribution (right). Bottom: p_T of the VZ candidate (left) and transverse mass of the VZ candidate (right). Events are selected with the *high-purity sidebands* selection, and simulated backgrounds are normalized to luminosity.

4.4 Data and simulations comparison

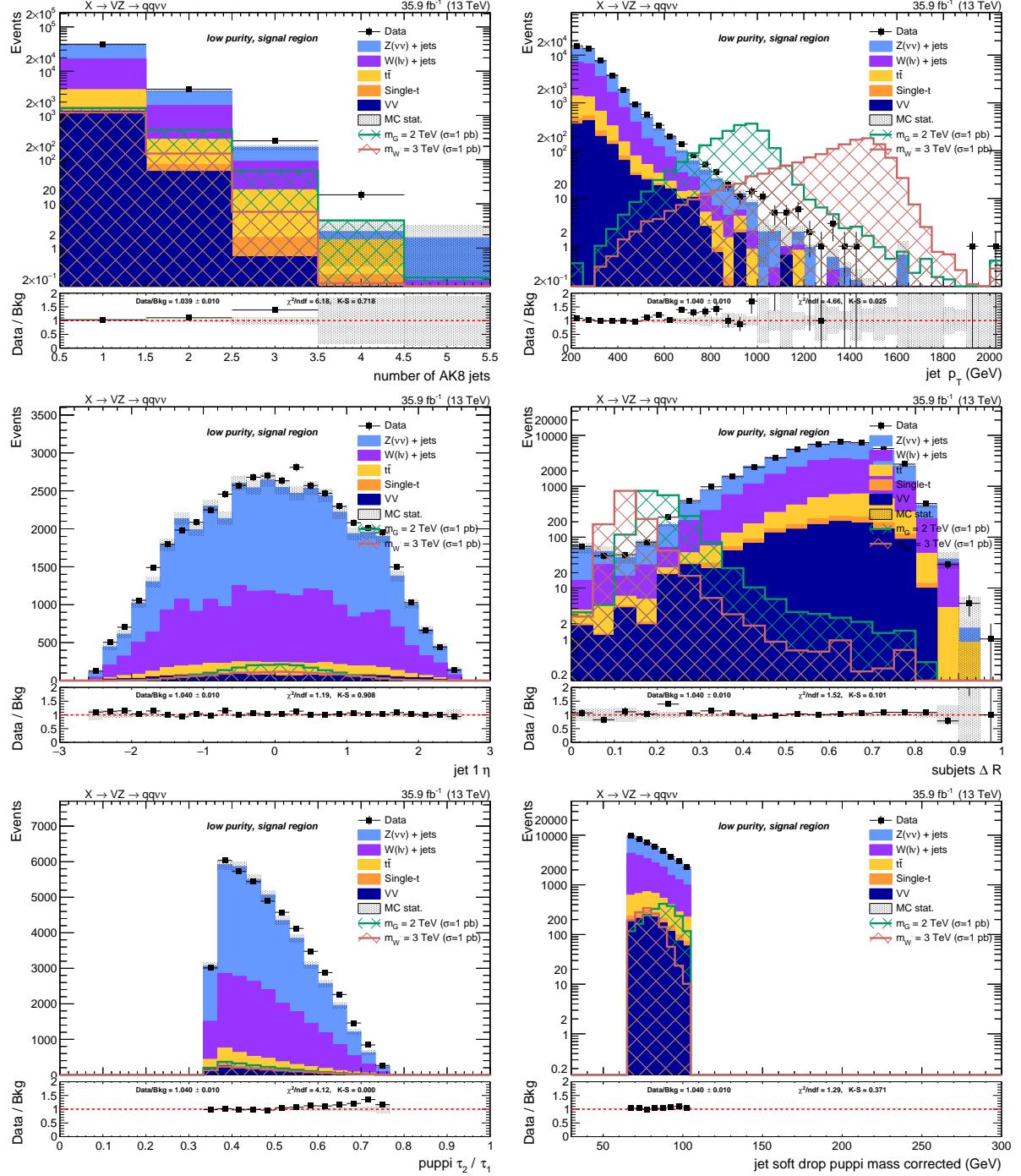


Figure 4.35: Top: number of AK8 jets in the event (left) and V jet candidate p_T (right). Center: V jet candidate η (left) and angular separation ΔR between the constituents leading subjets (right). Bottom: V jet candidate τ_{21} subjetiness after PUPPI correction (left) and V jet candidate soft drop PUPPI mass (right). Events are selected with the *low-purity signal region* selection, and simulated backgrounds are normalized to luminosity.

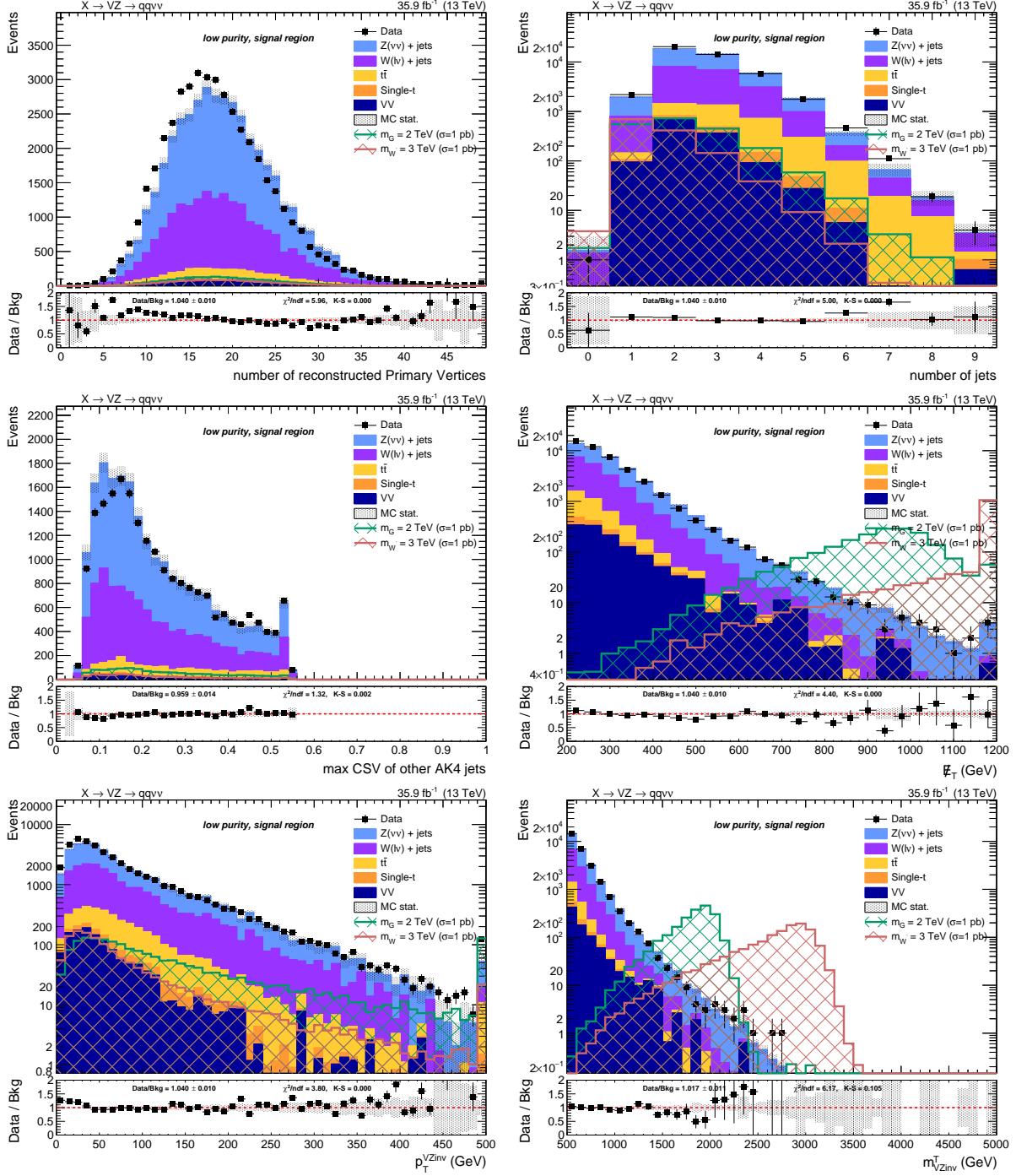


Figure 4.36: Top: number of reconstructed primary vertices (left) and number of AK4 jets in the event (right). Center: distribution of the b-tagging multivariate discriminant for the AK4 jets not included in the V jet cone (left) and E_T^{miss} distribution (right). Bottom: p_T of the VZ candidate (left) and transverse mass of the VZ candidate (right). Events are selected with the *low-purity signal region* selection, and simulated backgrounds are normalized to luminosity.

4.4 Data and simulations comparison

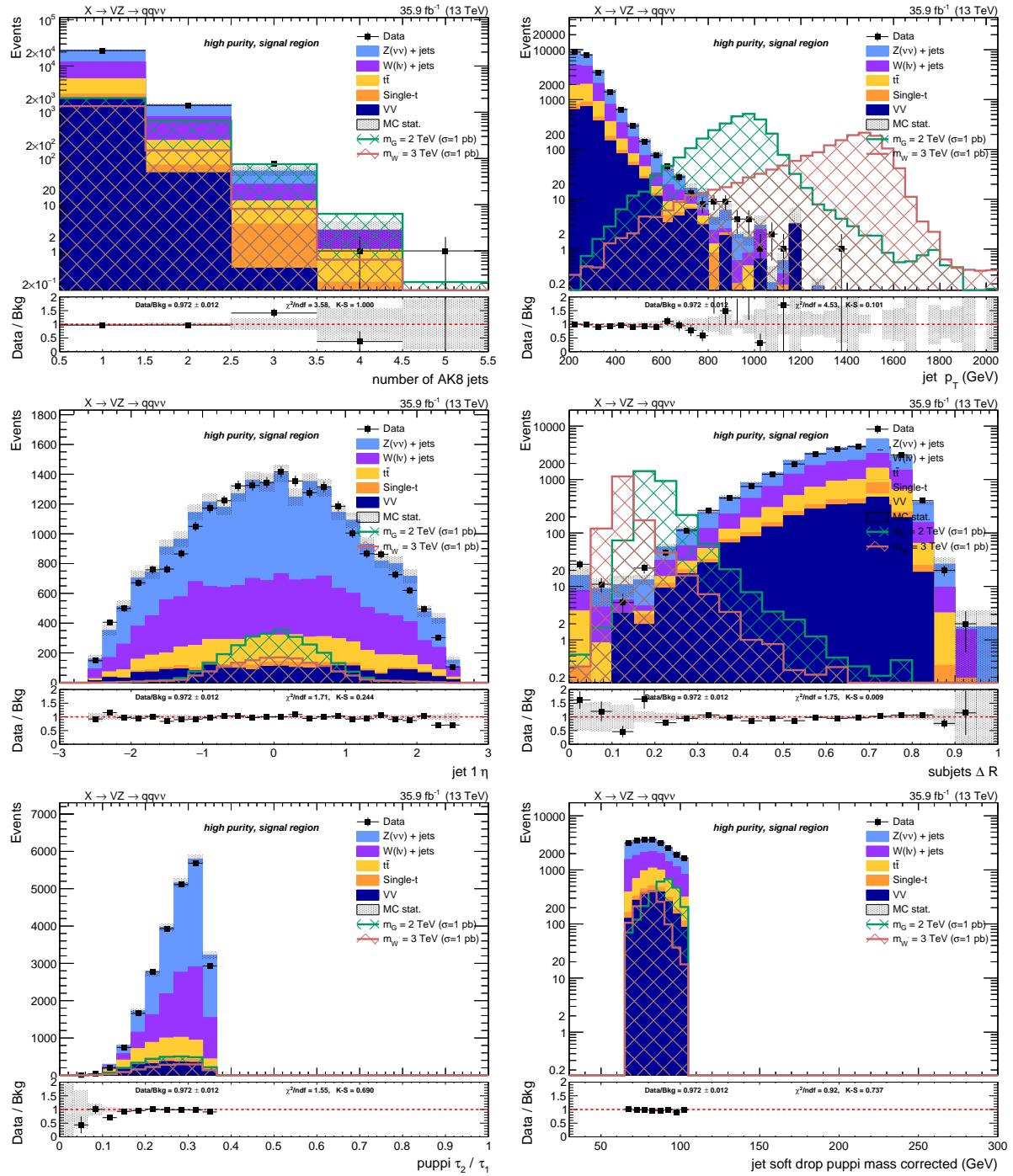


Figure 4.37: Top: number of AK8 jets in the event (left) and V jet candidate p_T (right). Center: V jet candidate η (left) and angular separation ΔR between the constituents leading subjets (right). Bottom: V jet candidate τ_{21} subjetiness after PUPPI correction (left) and V jet candidate soft drop PUPPI mass (right). Events are selected with the *high-purity signal region* selection, and simulated backgrounds are normalized to luminosity.

Search for diboson resonances in the $VZ \rightarrow q\bar{q} \nu\bar{\nu}$ final state

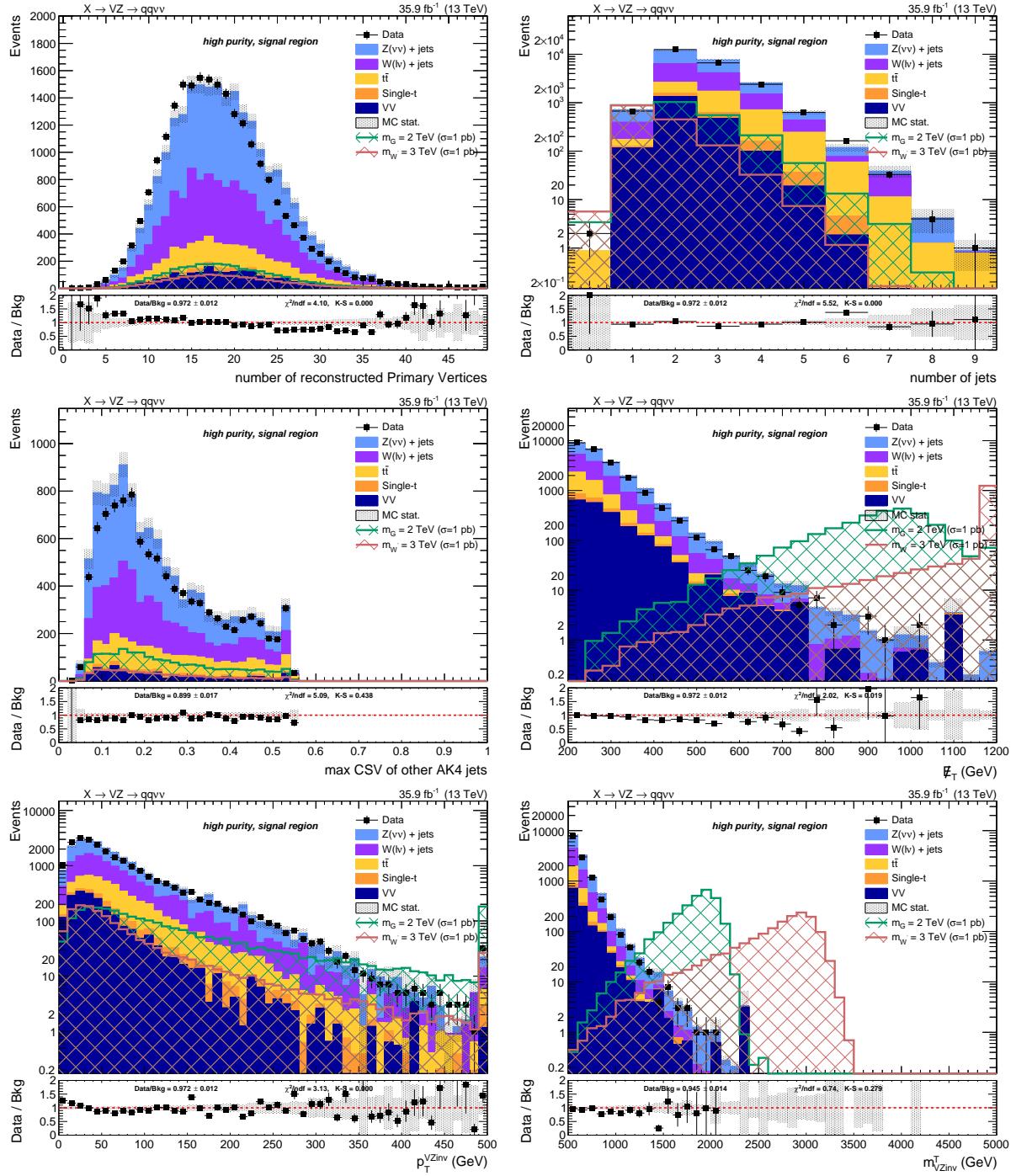


Figure 4.38: Top: number of reconstructed primary vertices (left) and number of AK4 jets in the event (right). Center: distribution of the b-tagging multivariate discriminant for the AK4 jets not included in the V jet cone (left) and E_T^{miss} distribution (right). Bottom: p_T of the VZ candidate (left) and transverse mass of the VZ candidate (right). Events are selected with the *high-purity signal region* selection, and simulated backgrounds are normalized to luminosity.

4.5 Background estimation technique

The goal of the analysis is to look for localized excesses in the m_{VZ}^T spectrum. The α method is used in searches for heavy resonances since Run 1 [94], and it has been introduced to be less dependent on the MC simulation for the background m_{VZ}^T estimation, due to the many sources of systematic uncertainties that are hard to understand and control. The two exclusive regions, *signal region* (SR) and *sidebands region* (SB), define a signal enriched or signal depleted phase space, respectively. First, the background normalization is extracted from data in the SB. Then, the α method extracts a predicted shape from the data in the SB to the SR using a transfer function (the α function) derived from simulation. The method relies on the assumption that the correlation between m_{VZ}^T and the groomed jet mass is reasonably well reproduced by the MC. The α -ratio is deemed to be more trustworthy since many systematic uncertainties would approximately cancel in the ratio.

Let's assume that, in the simplest case, only one dominant background is present. The α function is defined as the ratio of the two functions describing the simulated m_{VZ}^T shape in the SR and SB:

$$\alpha(m_{VZ}^T) = \frac{f_{\text{SR}}^{\text{MC,bkg}}(m_{VZ}^T)}{f_{\text{SB}}^{\text{MC,bkg}}(m_{VZ}^T)}, \quad (4.14)$$

and the background distribution in the SR is thus estimated as the product of $\alpha(m_{VZ}^T)$ with the shape in the data SB:

$$f_{\text{bkg}}(m_{VZ}^T) = f_{\text{SB}}(m_{VZ}^T) \times \alpha(m_{VZ}^T) \quad (4.15)$$

In the above description, no definition of the SB and SR is included. Ideally, the best choice would be a variable such that the distribution of m_{VZ}^T in the signal region and sidebands are similar. In this analysis, the soft drop PUPPI corrected jet mass m_V (sec. 4.3.7) is chosen as the control variable, and the cut values are those reported in tab. 4.3.7. All the selections used in the α method background prediction are the same reported in sec. 4.3.11.

In a real case scenario, the background is not purely composed of one single process neither in the SR nor in the SB. As already pointed out in sec. 4.2.3 and confirmed in sec. 4.4, the background composition is dominated by two processes, $Z + \text{jets}$ ($\sim 50\%$ in the whole SR) and $W + \text{jets}$ ($\sim 35\%$ in the whole SR), grouped together as $V + \text{jets}$, whose modeling in simulation is considered not to be trustworthy. Other subdominant backgrounds, $t\bar{t}$ and single-t production, grouped as Top, and diboson (VV), generally have smaller contributions (of the order of 5% for VV , and 9% for Top, in the whole SR), and are considered quite well understood and modeled by MC generators. The justification of merging $W + \text{jets}$ and $Z + \text{jets}$ together as a single $V + \text{jets}$ background is provided in sec. 4.5.3

The shape and normalization of the VV and Top production are taken from the simulation. The shape and normalization of the main background are evaluated with the α approach. The V candidate mass variable is used to perform the normalization prediction, the VZ candidate transverse mass variable is used for the shape prediction.

A different background prediction is derived for each category separately, thus dividing low- and high-purity categories, and it is calculated in a transverse mass range $950 < m_{VZ}^T < 4750$ GeV.

4.5.1 Background normalization

The first step in the background prediction consists in a proper estimation of the background normalization. The jet mass distributions of the three backgrounds ($V + \text{jets}$, Top, and VV) are described with functional forms determined by fits on the simulated backgrounds. The so-built templates are summed together, maintaining the relative weights between the three, and finally fitted

1723 to the data in the jet mass sidebands. During the fit to data SB, the parameters of the $V +$ jets back-
 1724 ground are left free to float and adapt to the data distribution. The integral of the final sum of the
 1725 fitted functions over the SR jet mass range represents the background yield prediction in the SR.
 1726 The empirical functional forms for each background are chosen to reflect the physics properties of
 1727 the samples. In the low-purity category, the $V +$ jets background is a falling background with no
 1728 peaks, modelled as a power law, while in the high-purity category the $V +$ jets background compo-
 1729 nent is characterized by a broad distribution roughly centered at m_V , modelled as a gaussian, with
 1730 an exponential tail at high mass values. The exponential falling $VV m_V$ spectrum shows a peak,
 1731 corresponding to the reconstruction of a vector boson hadronic decay. The hadronic decays of W
 1732 and Z bosons cannot be distinguished, hence they are modelled together as a gaussian. For the
 1733 jet mass spectrum of the Top backgrounds, two peaks corresponding to the W and top quark mass
 1734 can be observed; they are modelled as gaussian functions, superimposed to a falling exponential
 1735 background.
 1736 An extended likelihood fit is performed, hence the functional forms chosen to build the jet mass
 1737 templates are normalized to unity (becoming probability density functions) through normalization
 1738 factors (f_0, f_1):

- ErfPow2: an error function (Erf) multiplied by a power law, that is a function of the center-of-mass energy $\sqrt{s} = 13$ TeV. It depends on 4 parameters (the power law parameters c_0, c_1 , and the error function offset o and width w):

$$F_{\text{ErfPow2}}(x) = \left(\frac{x}{\sqrt{s}}\right)^{-c_0+c_1 \log(x/\sqrt{s})} \cdot \frac{1 + \text{Erf}((x-o)/w)}{2};$$

- ExpGaus: an exponential plus one gaussian. It depends on 4 parameters (the normalization f_0 , the exponential parameter a , the gaussian mean b and variance c):

$$F_{\text{ExpGaus}}(x) = f_0 \cdot e^{ax} + (1-f_0) \cdot e^{2(x-b)^2/c};$$

- ErfExpGaus: an error function, multiplied to an exponential, plus one gaussian. It depends on 6 parameters (the normalization f_0 , the exponential parameter a , the gaussian mean b and variance c , the error function offset o and width w):

$$F_{\text{ErfExpGaus}}(x) = f_0 \cdot e^{ax} \cdot \frac{1 + \text{Erf}((x-o)/w)}{2} + (1-f_0) \cdot e^{2(x-b)^2/c};$$

- ErfExpGaus2: an error function, multiplied to an exponential, plus two gaussians. It depends on 9 parameters (the normalization factors f_0 and f_1 , the exponential parameter a , the two gaussians means $b-d$ and variances $c-e$, the error function offset o and width w):

$$F_{\text{ErfExpGaus2}}(x) = f_0 \cdot e^{ax} \cdot \frac{1 + \text{Erf}((x-o)/w)}{2} + f_1 \cdot e^{2(x-b)^2/c} + (1-f_0-f_1) \cdot e^{2(x-d)^2/e}.$$

1739 The choice of the functions is category-dependent, and it is summarized in tab. 4.15. In order to
 1740 make the background evaluation less dependent as possible from the choice of the function de-
 1741 scribing the jet mass of the main $V +$ jets background, an alternative function has been used to
 1742 fit the $V +$ jets mass spectrum. The absolute difference bewteen the number of expected events
 1743 calculated with the main $V +$ jets function and the alternative is taken as systematic uncertainty.
 1744 The following plots (fig. 4.39-4.40) show the fits to the jet mass distributions in Monte Carlo samples,
 1745 in the different categories; the alternative functions for the main background are displayed with

4.5 Background estimation technique

Table 4.15: Chosen functions to fit the jet mass distributions for each category.

Category	$V + \text{jets}$	alt. $V + \text{jets}$	Top	VV
low-purity	ErfPow2	ExpGaus	ErfExpGaus2	ExpGaus
high-purity	ExpGaus	ErfExpGaus	ErfExpGaus2	ExpGaus

Table 4.16: Expected background yield in the SB ($30 < m_V < 65 \text{ GeV}$, $m_V > 135 \text{ GeV}$) and in the SR ($65 < m_V < 105 \text{ GeV}$) and the respective systematic and statistical uncertainties.

Region	Category	Expected events	Statistical uncertainty	Systematic uncertainty	Alternative function uncertainty	Observed events
SB	low-purity	2356.6	± 52.5	± 16.0	± 1.1	2314
SR	low-purity	1093.2	± 48.1	± 16.4	± 49.1	1153
SB	high-purity	779.8	± 29.1	± 13.1	± 0.3	774
SR	high-purity	254.4	± 15.3	± 17.9	± 7.8	271

1746 dotted lines. The background estimation, after the fit to data SB, is shown in fig. 4.41. The bottom
 1747 panels of each plot display the fit pulls (per-bin), namely, the number of events observed in data
 1748 (or in Monte Carlo simulations) minus the number of events predicted by the fit, divided by the
 1749 uncertainty in the data (or simulations). Table 4.16 summarizes the expected background yield in
 1750 the signal region, that is in agreement with observation in both the purity categories. The quoted
 1751 uncertainties are calculated as follows:

- 1752 • the statistic uncertainty is the uncertainty of the fit to the $V + \text{jets}$ background performed on
 1753 data SB;
 1754 • the systematic uncertainty is the propagation of the uncertainties of the fits to the VV and
 1755 Top backgrounds performed on simulations, to the fit performed on data SB to extract the V
 1756 + jets functional parameters;
 1757 • the alternative function uncertainty describes the discrepancy in the background yield in SR
 1758 depending on the choice of the function to describe the $V + \text{jets}$ background.

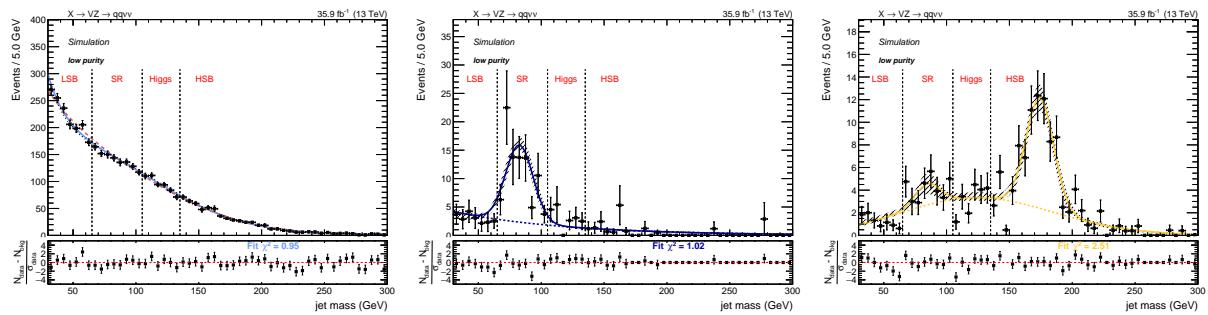


Figure 4.39: Fit to the simulated m_V in the low-purity category for the three backgrounds: $V + \text{jets}$ (left), VV (center), Top (right). For the main background prediction, the alternative function is displayed with a dotted red line, superimposed to the main choice (continuous light blue curve).

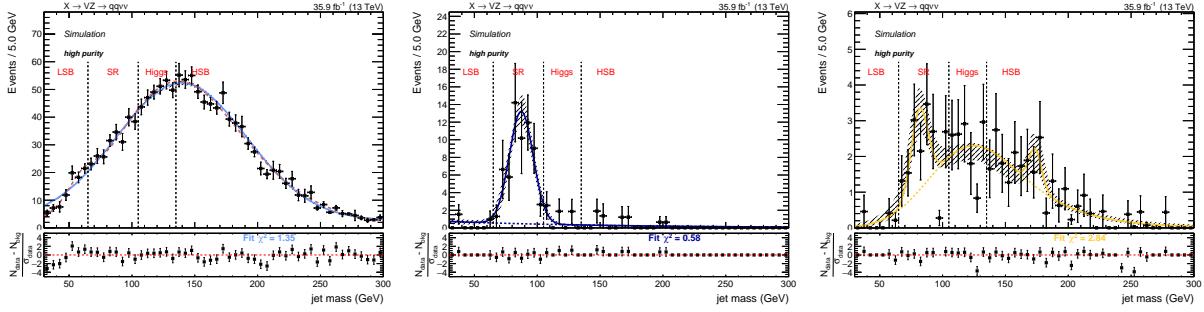


Figure 4.40: Fit to the simulated m_V in the high-purity category for the three backgrounds: $V +$ jets (left), VV (center), Top (right). For the main background prediction, the alternative function is displayed with a dotted red line, superimposed to the main choice (continuous light blue curve).

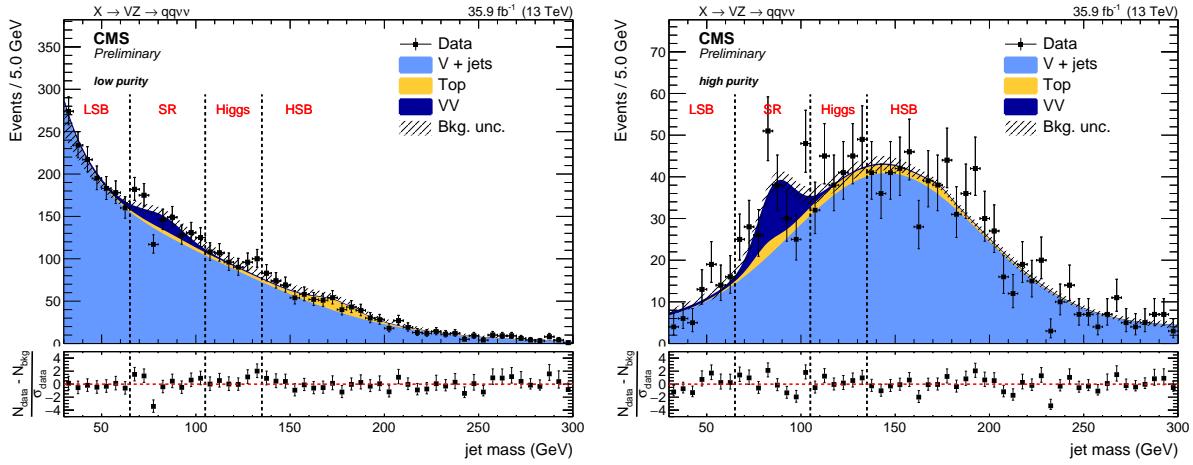


Figure 4.41: Background yield prediction in the signal region, after the fit to data sidebands, in the low- (left) and high-purity category (right). Data and predictions are in agreement.

1759 4.5.2 Background shape

1760 The second part of the background prediction consists in estimating the background shape of the
 1761 transverse mass of the diboson candidate, m_{VZ}^T . Each transverse mass spectrum is parametrized
 1762 separately for the $V +$ jets background ($f_{SR}^{MC, V + jets}(m_{VZ}^T)$, $f_{SB}^{MC, V + jets}(m_{VZ}^T)$), Top production ($f_{SR}^{MC, Top}(m_{VZ}^T)$,
 1763 $f_{SB}^{MC, Top}(m_{VZ}^T)$), and diboson background ($f_{SR}^{MC, VV}(m_{VZ}^T)$, $f_{SB}^{MC, VV}(m_{VZ}^T)$). The parameters of these
 1764 functions are extracted by fitting the simulated m_{VZ}^T spectra in SR and SB, respectively. The top and
 1765 the diboson spectra are normalized to luminosity; the $V +$ jets spectrum is normalized according
 1766 to the data-driven prediction obtained in sec. 4.5.1. The functions describing the $V +$ jets back-
 1767 ground, calculated from simulations, are used to define the α function, that has the purpose to take
 1768 into account the kinematical differences of the SR compared to SB:

$$1769 \quad \alpha(m_{VZ}^T) = \frac{f_{SR}^{MC, V + jets}(m_{VZ}^T)}{f_{SB}^{MC, V + jets}(m_{VZ}^T)}. \quad (4.16)$$

1770 The parameters describing the main background are then left free to float and extracted through a
 1771 fit to data in the SB, after subtracting the corresponding Top and VV contribution from data. The
 1772 resulting shape is then multiplied by the α function in order to get the main background expectation

4.5 Background estimation technique

1772 in the SR. Finally, the Top and diboson contributions in the SR are added to the main background
 1773 estimation.

1774 In formulas, the procedure used to extract the total background prediction is the following:

$$f_{\text{SR}}^{\text{data}}(m_{VZ}^T) = \left(f_{\text{SB}}^{\text{data}}(m_{VZ}^T) - f_{\text{SB}}^{\text{MC, Top}}(m_{VZ}^T) - f_{\text{SB}}^{\text{MC, VV}}(m_{VZ}^T) \right) \times \left[\frac{f_{\text{SR}}^{\text{MC, } V + \text{jets}}(m_{VZ}^T)}{f_{\text{SB}}^{\text{MC, } V + \text{jets}}(m_{VZ}^T)} \right] + f_{\text{SR}}^{\text{MC, Top}}(m_{VZ}^T) + f_{\text{SR}}^{\text{MC, VV}}(m_{VZ}^T), \quad (4.17)$$

1775 where the expression in brackets represents the main background evaluation in data SB; the α -ratio
 1776 is the expression enclosed in square brackets.

1777 The functions probed to parametrize the m_{VZ}^T distributions are smoothly falling exponential func-
 1778 tions:

- ExpN: a product of two exponentials. It depends on two parameters a, b :

$$F_{\text{ExpN}}(x) = e^{ax+b/x}$$

- ExpTail: a modified exponential function with an additional parameter to model the expo-
 1779 nential tails. It depends on two parameters a, b :

$$F_{\text{ExpTail}}(x) = e^{-x/(a+bx)}$$

Table 4.17: Main and alternative functions chosen to parametrize the background contributions in the m_{VZ}^T distribution for each category.

Category	Main bkg function	Main bkg alternative	Diboson	Top
low-purity	ExpN	ExpTail	ExpTail	ExpTail
high-purity	ExpTail	ExpN	ExpTail	ExpTail

1779 The functions chosen to parametrize the backgrounds and extract the α function are reported in
 1780 tab. 4.17, for each category. As a cross-check for the main α function used in the background esti-
 1781 mation, an additional α function is extracted with alternative function choices for the $V + \text{jets}$ back-
 1782 ground. Table 4.17 reports both the main function and the alternative function. In fig. 4.42 (4.44),
 1783 the fits to each simulated background are reported for sidebands and signal region respectively, for
 1784 low- (high-) purity categories. In fig. 4.43 (4.45), the results of the fit to data SB are presented for
 1785 the low- (high-) purity categories: the expected background distribution in SB, where parameters
 1786 describing the $V + \text{jets}$ background are extracted according to data distribution (left); the α -ratio
 1787 function, calculated with the main function to describe the $V + \text{jets}$ background (black solid line)
 1788 and the alternative function (gray dotted line) (center); the full background estimation performed
 1789 with the main and alternative functions for describing the $V + \text{jets}$ background: the background
 1790 shape in SB (blue solid curve for the main function, light blue dotted curve for the alternative) and
 1791 the final background shape in SR (red solid line for the main function, green dotted line for the al-
 1792 ternative) (right). A proof to the compatibility of the two predictions in SR is presented in sec. 4.5.3.
 1793 The bottom panels in the plots display the fit pulls (per-bin), namely, the number of events observed
 1794 in data (or in Monte Carlo simulations) minus the number of events predicted by the fit, divided by
 1795 the uncertainty in the data (or simulations).

1796 Fig. 4.46 summarizes the final background predictions as a function of the search variable, the trans-
 1797 verse mass. Data and predictions are in agreement in both the categories.

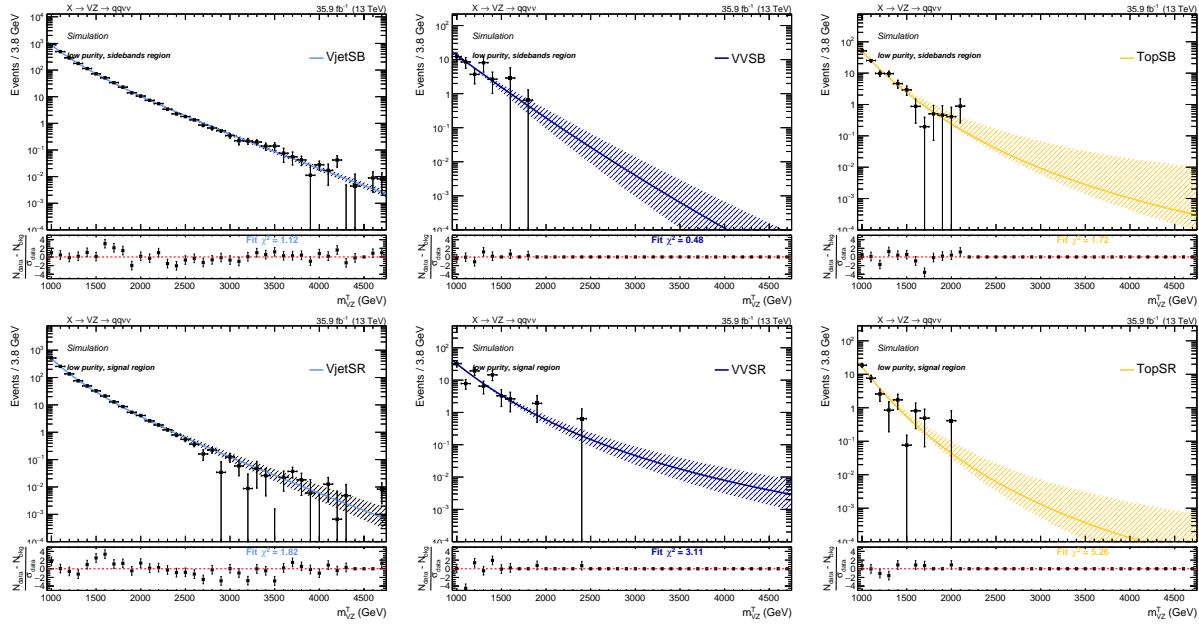


Figure 4.42: Low-purity category. Top: fits to the simulated background components $V + \text{jets}$ (left), VV (center), Top (right) in the sidebands (SB). Bottom: fits to the simulated background components $V + \text{jets}$ (left), VV (center), Top (right) in the signal region (SR).

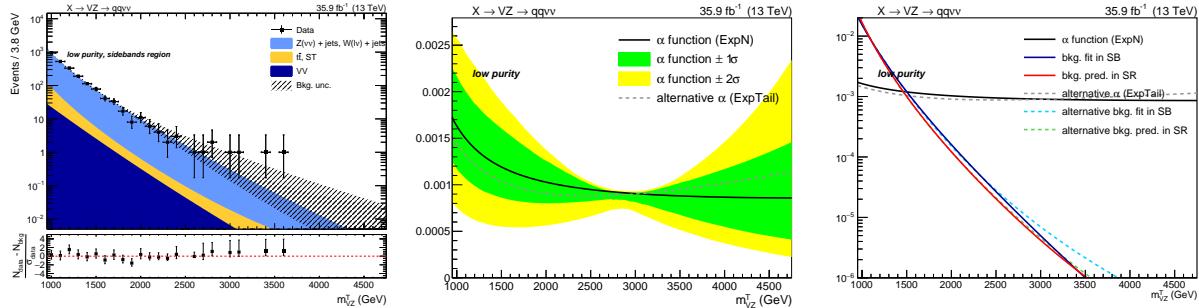


Figure 4.43: Low-purity category. Result of the fit to data in the SB (left), α -ratio function (center), and α function compared to the background shape in both SB and SR (right). The black line, with the corresponding 1σ (green) and 2σ (yellow) uncertainty bands, represents the α function. The gray line is the alternative α function. The blue and red solid lines represent the estimated background in the SB and SR, respectively, with both the main (solid line) and alternative (dotted line) parametrizations.

1798 4.5.3 Alpha method validations

1799 The first required validation is performed in order to legitimate the choice of putting the $Z + \text{jets}$
 1800 and the $W + \text{jets}$ backgrounds together while performing the background estimation. The full pro-
 1801 cedure has been repeated, by keeping the two background contributions separated. Fit results per-
 1802 formed in SB (top plots) and SR (bottom plots) in MC samples are displayed in fig. 4.47 (4.48) for
 1803 low- (high-) purity category, for $Z + \text{jets}$ and W jets background, separately, and for the combina-
 1804 tion of the two. In fig. 4.49, the α functions calculated for $Z + \text{jets}$ background (red dotted line) and
 1805 for $W + \text{jets}$ background (blue dotted line) are in agreement with the α function used in the analysis
 1806 (black solid line), calculated by merging together the two backgrounds, both in low- (left plot) and

4.5 Background estimation technique

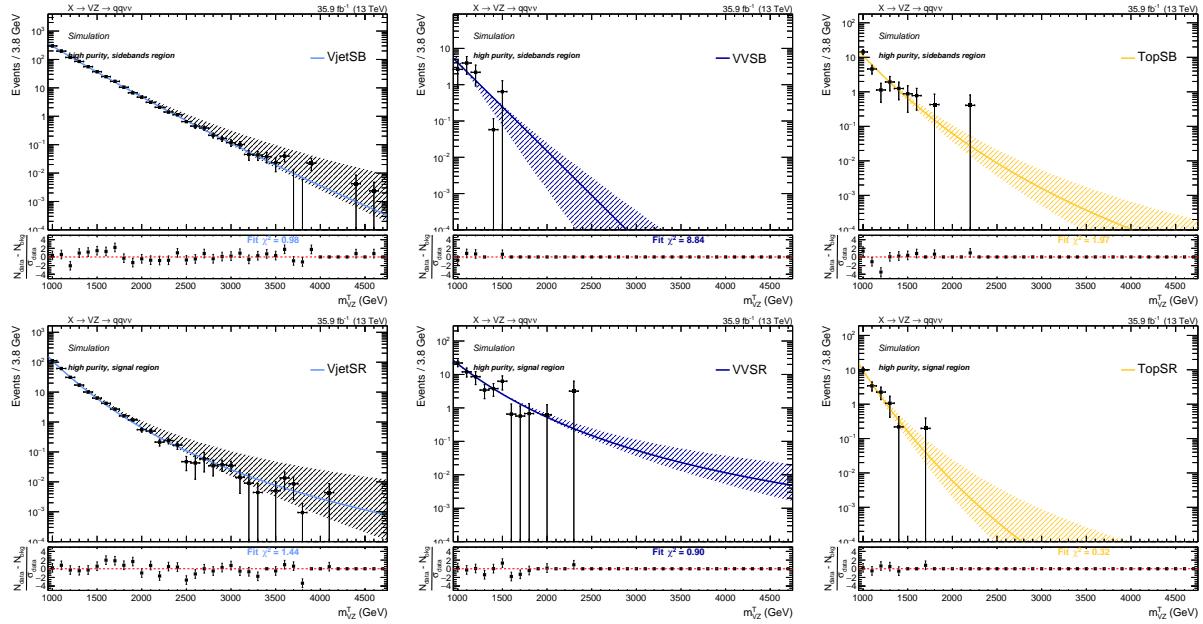


Figure 4.44: High-purity category. Top: fits to the simulated background components $V + \text{jets}$ (left), VV (center), Top (right) in the sidebands (SB). Bottom: fits to the simulated background components $V + \text{jets}$ (left), VV (center), Top (right) in the signal region (SR).

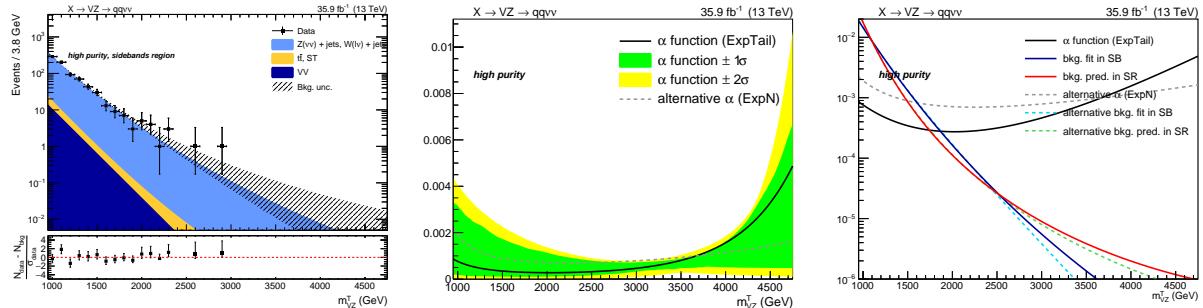


Figure 4.45: High-purity category. Result of the fit to data in the SB (left), α -ratio function (center), and α function compared to the background shape in both SB and SR (right). The black line, with the corresponding 1σ (green) and 2σ (yellow) uncertainty bands, represents the α function. The gray line is the alternative α function. The blue and red solid lines represent the estimated background in the SB and SR, respectively, with both the main (solid line) and alternative (dotted line) parametrizations.

1807 high-purity category (right plot).

1808 As a robustness check of the α -ratio method, a closure test is performed on data. Instead of pre-
1809 dicting the background in the real SR from both the lower and the upper jet mass sidebands, the SB
1810 and SR are redefined for the purposes of this test. The low sideband is splitted into two sub-regions:
1811 30 – 50 GeV (LSB) and 50 – 65 GeV (SR). The former is considered as the new low sideband, while
1812 the latter is exploited as a pseudo-signal region. The high sideband is instead effectively used in the
1813 fit without any modifications with respect to the standard α -ratio method. With this configuration,
1814 the prediction of the background in the SR region is estimated from the fit to the LSB region and

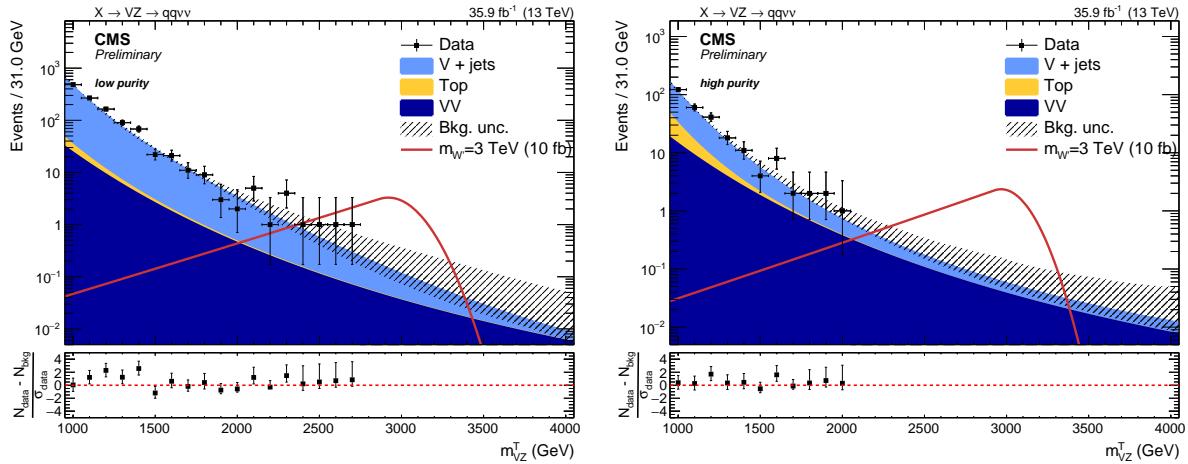


Figure 4.46: Expected background predicted with the α method in the low- (left) and high-purity category (right), compared to observations (black markers) and a signal hypothesis of a spin-1 W' of mass 3 TeV.

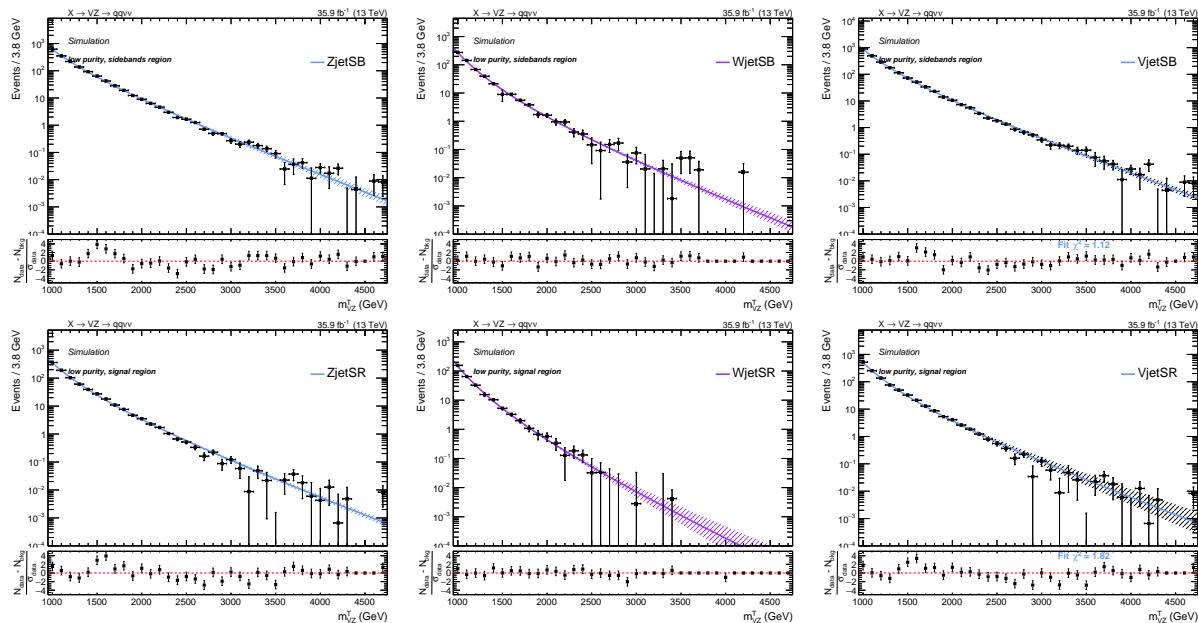


Figure 4.47: Validation of the α method, low-purity category. Top: fits to the simulated background components $Z + \text{jets}$ (left), $W + \text{jets}$ (center), and their combination $V + \text{jets}$ (right), in the sidebands (SB). Bottom: fits to the simulated background components $Z + \text{jets}$ (left), $W + \text{jets}$ (center), and their combination $V + \text{jets}$ (right), in the signal region.

1815 the high-sidebands, and checked with data for both shape and normalization. This test has been
 1816 performed before the unblinding of the signal region of the analysis.

1817 In fig. 4.50 and tab. 4.18, the predicted shapes and normalizations are compared to the observed
 1818 ones in data. A good overall agreement both in normalization and shape is obtained. There is a
 1819 bit of tension in normalization for high-purity category, due to an upper fluctuation in data around
 1820 60 GeV. This cross check confirms that the method to extract the $V + \text{jets}$ background is reliable
 1821 and can be used to model the background in the search for potential excesses in the signal region

4.5 Background estimation technique

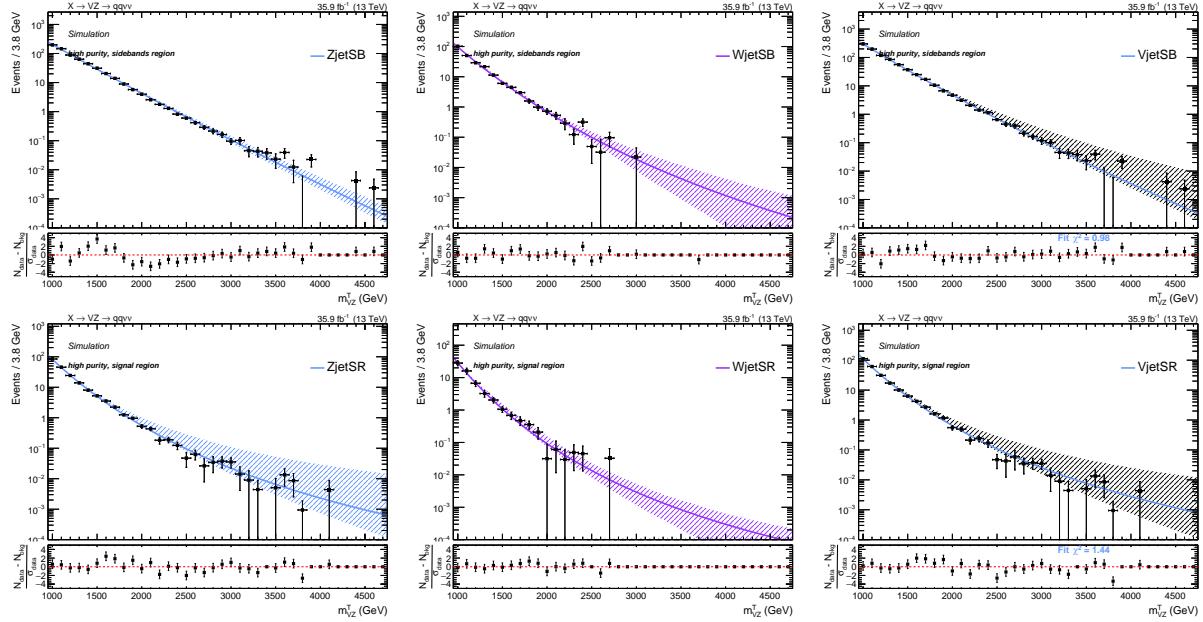


Figure 4.48: Validation of the α method, high-purity category. Top: fits to the simulated background components $Z + \text{jets}$ (left), $W + \text{jets}$ (center), and their combination $V + \text{jets}$ (right), in the sidebands (SB). Bottom: fits to the simulated background components $Z + \text{jets}$ (left), $W + \text{jets}$ (center), and their combination $V + \text{jets}$ (right), in the signal region.

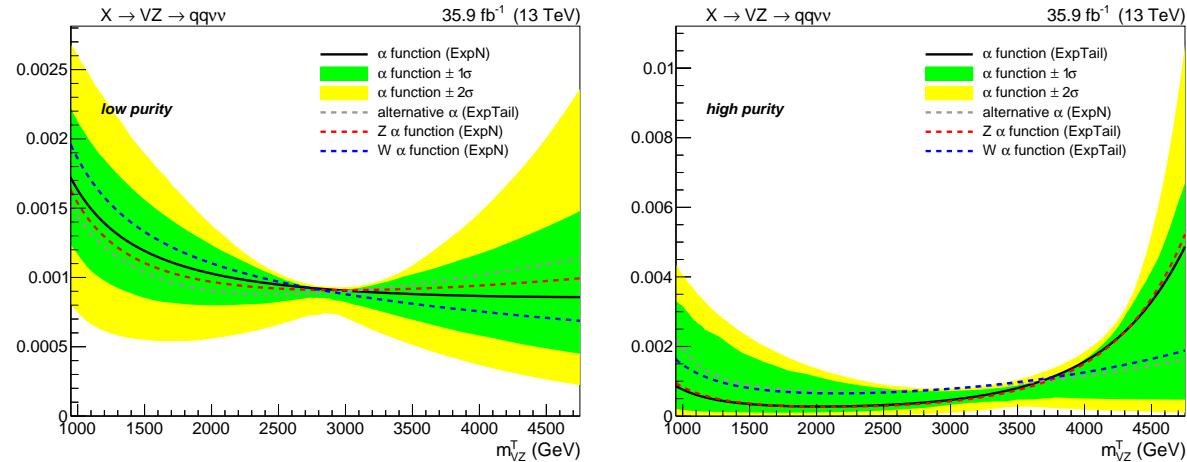


Figure 4.49: Validation of the α method: α functions calculated for $Z + \text{jets}$ background (red dotted line) and for $W + \text{jets}$ background (blue dotted line) separately, and α function for the total $V + \text{jets}$ background (black solid line). Left: low-purity category; right: high-purity category.

1822 defined in the analysis.

1823 The last check performed is a study of the impact of the choice of the function to describe the V
 1824 + jets background on the very last result of the analysis, namely the exclusion limit on the signal
 1825 cross-section times branching fraction. The procedure of the limit extraction is discussed in detail
 1826 in sec. 4.7. The main and alternative functions chosen to parametrize the dominant background
 1827 depend on the purity category and are listed in tab. 4.17. In fig. 4.51 (top), the fit results of the

Table 4.18: Expected and observed background yield in the pseudo-SR jet mass region ($50 < m_V < 65$ GeV), predicted from the LSB one ($30 < m_V < 50$ GeV) and high-sideband ($m_V > 135$ GeV).

Region	Category	Expected	Observed
SB	low-purity	1841.3 ± 45.7	1793
pseudo-SR	low-purity	529.9 ± 37.8	521
SB	high-purity	728.5 ± 29.9	725
pseudo-SR	high-purity	39.3 ± 5.2	49

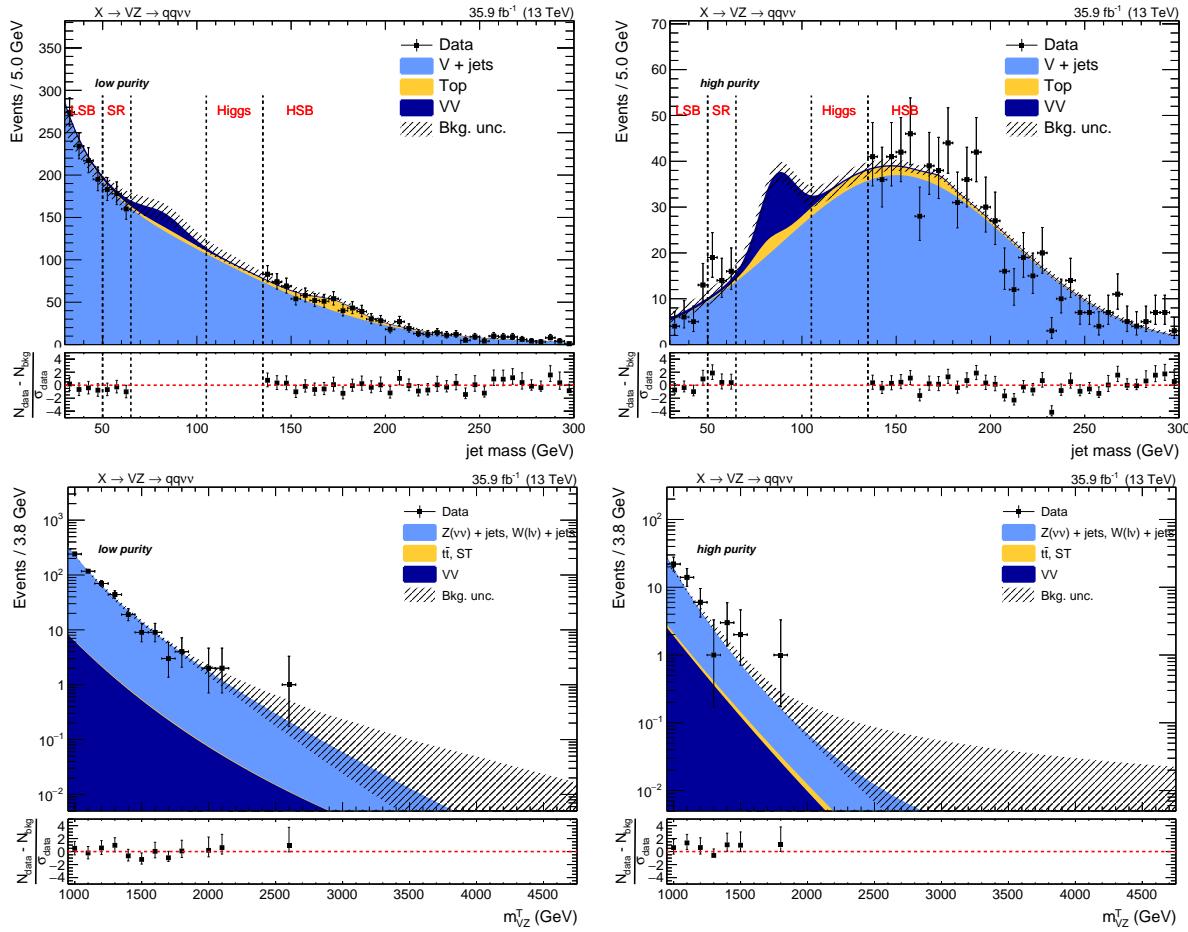


Figure 4.50: Top: results of the fit to the m_V spectrum in data, in the sidebands defined for the α method validation: low-sideband ($30 < m_V < 50$ GeV) and high-sideband ($m_V > 135$ GeV) (left: low-purity category, right: high-purity category). Bottom: results of the fits to the m_{VZ}^T spectrum, in the pseudo-signal region ($50 < m_V < 65$ GeV) defined for the α method validation (left: low-purity category, right: high-purity category). Both the true signal region and the Higgs regions are kept blind.

background shape prediction of the transverse mass are displayed. They are obtained by choosing the main function to describe the main background (red curve) and the alternative function (green curve); the two predictions are in agreement and very close to each other, for both low- (left) and high- (right) purity category. In fig. 4.51 (center), the 95% CL exclusion limits on cross-section times branching fraction are displayed for a spin-2 bulk graviton hypothesis, as a function of the mass of the

4.5 Background estimation technique

resonance. The same figure of merit is shown in fig. 4.51 (bottom), considering a spin-1 W' hypothesis. In the plots, the exclusion limits are calculated by choosing the main function to describe the $V + \text{jets}$ background (left plots: green curve for low-purity category alone and black curve for high-purity category alone, right plot: red curve for the combination of the categories) or the alternative function (left plots: orange curve for low-purity category alone and pink curve for high-purity category alone, right plot: blue curve for the combination of the categories). The impact of the choice of the function is negligible ($<< 1\%$).

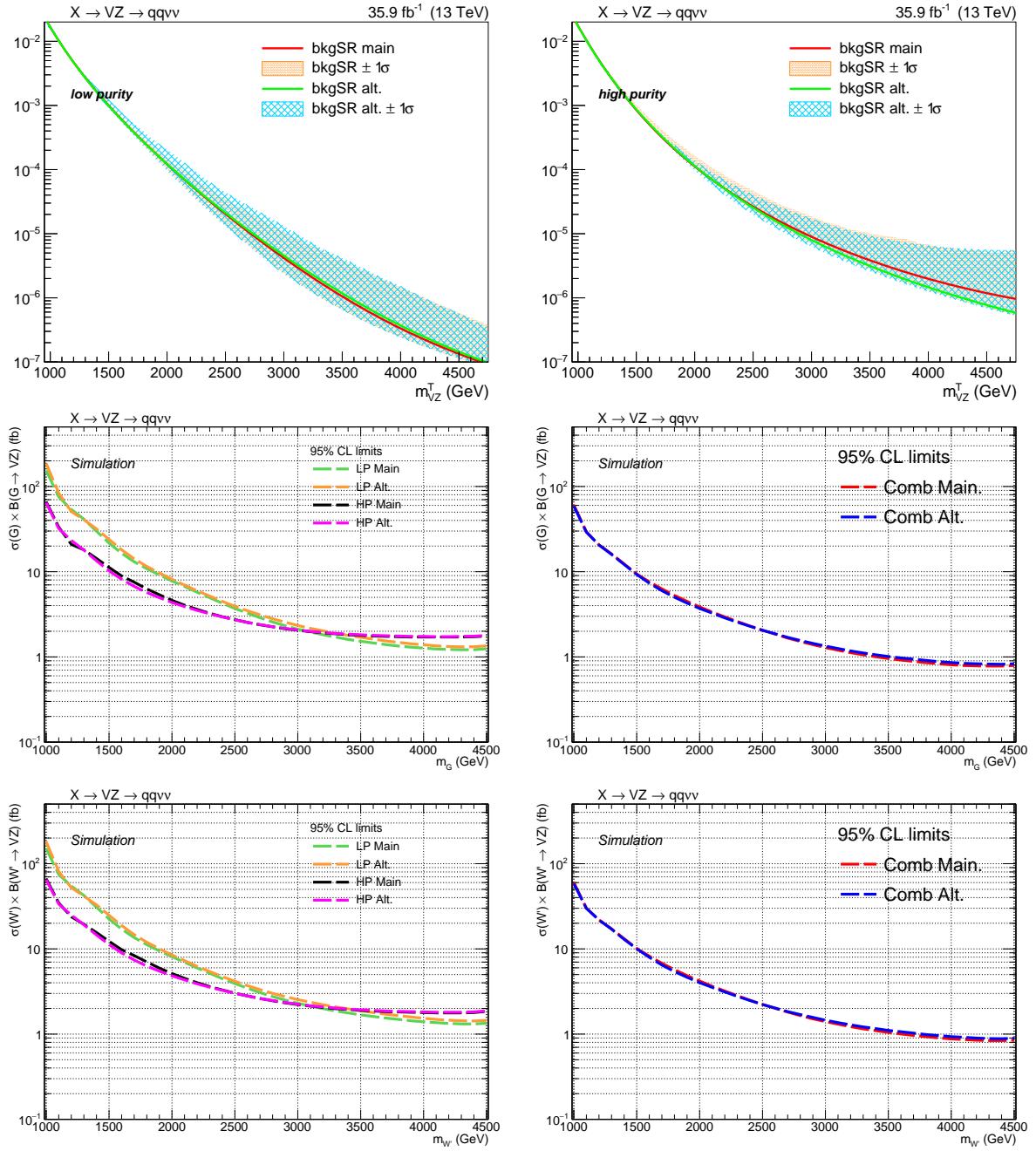


Figure 4.51: Validation of the α method: impact of the choice of the function to describe the dominant $V + \text{jets}$ background. Top: fit results of the background shape prediction in the SR obtained with the main function (red curve) and the alternative function (green curve), for low- (left) and high- (right) purity categories. Center: exclusion limits on cross-section times branching fraction for a spin-2 bulk graviton hypothesis, as a function of the mass of the resonance, calculated by choosing the main function (left plots: green curve for low-purity category alone and black curve for high-purity category alone, right plot: red curve for the combination of the categories) or the alternative function (left plots: orange curve for low-purity category alone and pink curve for high-purity category alone, right plot: blue curve for the combination of the categories). Bottom: exclusion limits on cross-section times branching fraction for a spin-1 W' hypothesis.

4.5 Background estimation technique

1840 4.5.4 Signal modeling

1841 The simulated signal samples, with different resonance mass hypotheses, are fitted in the SR with an
 1842 empirical function in order to be able to perform an unbinned likelihood fit for the signal extrac-
 1843 tion. The function chosen to model the signal samples is a *Crystal Ball* function [95, 96], which is
 1844 composed by a gaussian-like core convoluted to two power-law tails. Both spin-2 (fig. ??) and spin-1
 1845 (fig. 4.53) signal samples are fitted.

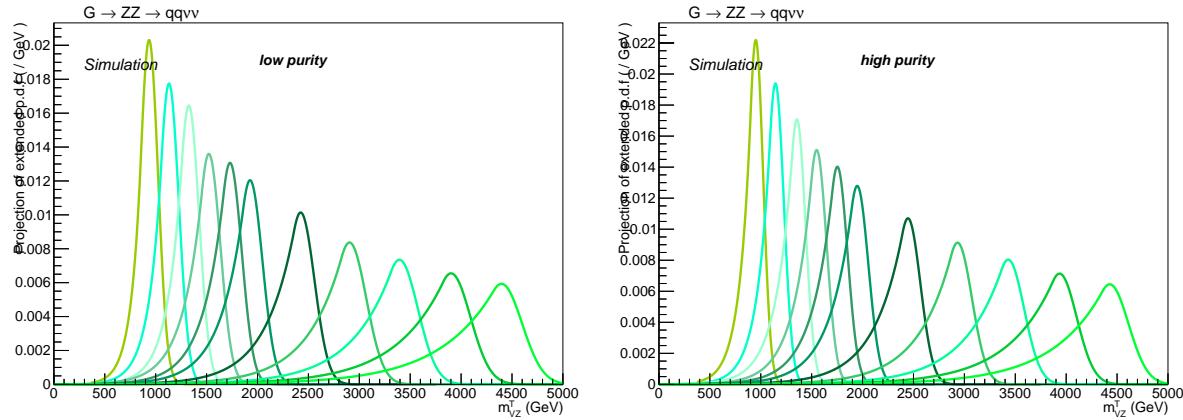


Figure 4.52: Interpolation of the signal as a function of the resonance transverse mass m_{VZ}^T , for a spin-2 (bulk graviton) signal hypothesis with an arbitrary cross section of 1 pb in the low- (left) and high-purity category (right).

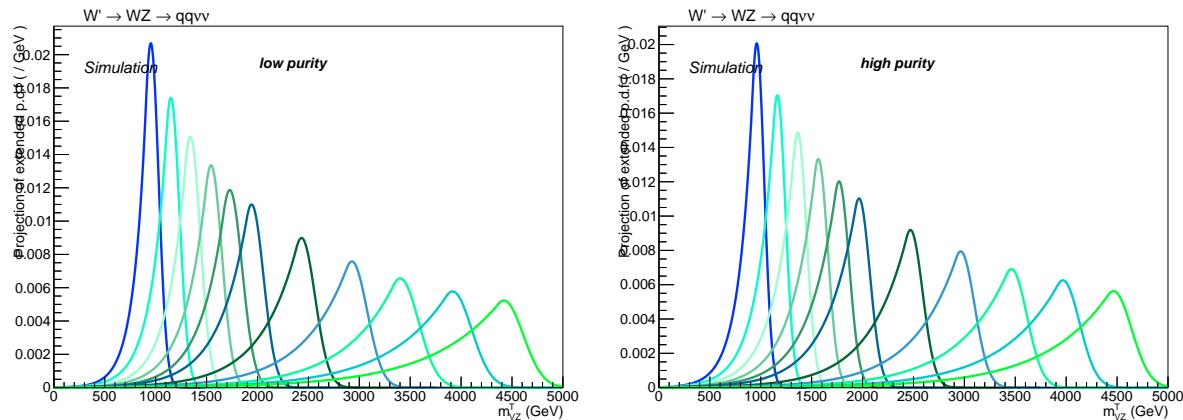


Figure 4.53: Interpolation of the signal as a function of the resonance transverse mass m_{VZ}^T , for a spin-1 (W') signal hypothesis with an arbitrary cross section of 1 pb in the low- (left) and high-purity category (right).

1846 4.5.4.1 Signal parametrization

1847 The signal is parametrized by interpolating the fitted parameters separately for each category in or-
 1848 der to have a continuous variation of the signal shape for every possible m_{VZ}^T value within the range.
 1849 A linear fit is performed on the mean and the width of the gaussian core of the Crystal Ball functions.
 1850 The interpolations are shown in fig. 4.54- 4.55 for the spin-2 signal model, and in fig. 4.56- 4.57 for

1851 the spin-1 signal model. Shape systematic uncertainties, as described in sec. 4.6, are taken into ac-
 1852 count while describing the mean and sigma of the gaussian core, and they are related to the effects
 1853 of the jet mass scale and resolution. Other shape parameters describing the tails of the Crystal Ball
 1854 are fitted as 3rd degree polynomial.

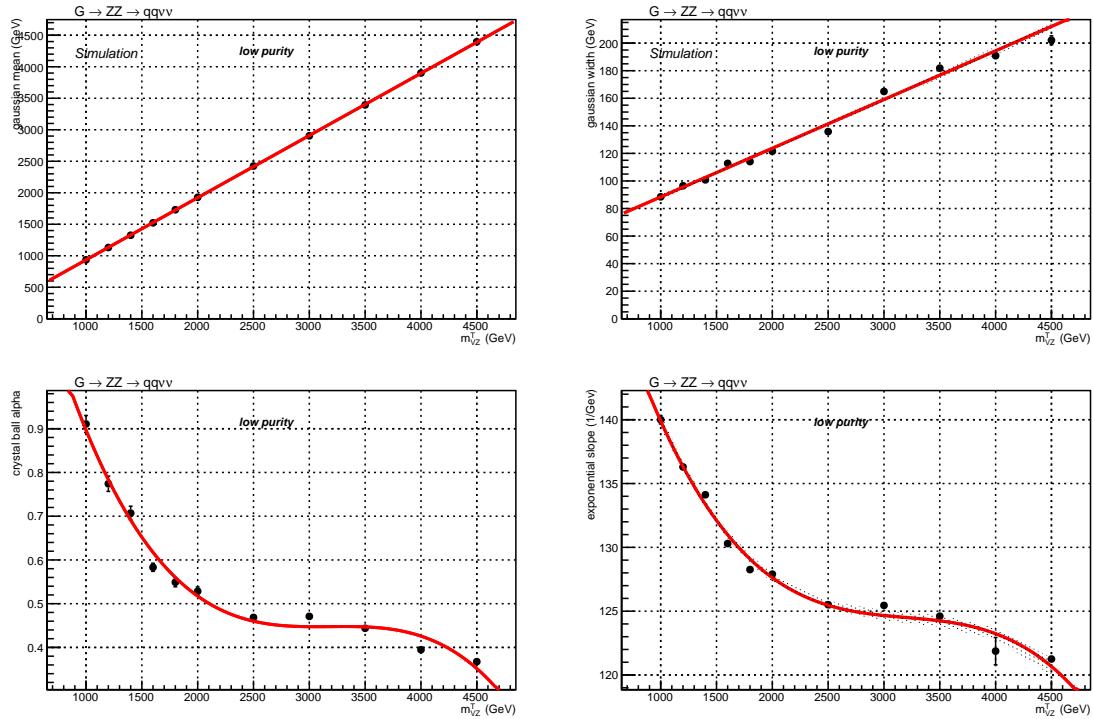


Figure 4.54: Interpolation of the fitted parameters as a function of the resonance transverse mass m_{VZ}^T , for a spin-2 (bulk graviton) signal hypothesis, low-purity category.

1855 The normalization of the signal samples is extrapolated from the fitted integral of the Crystal Ball
 1856 functions. The points are then connected with a line, in order to have an acceptable description
 1857 of the normalization as a function of m_{VZ}^T . The interpolations are shown in fig. 4.58 for the spin-2
 1858 signal model, and in fig. 4.59 for the spin-1 signal model.

4.5 Background estimation technique

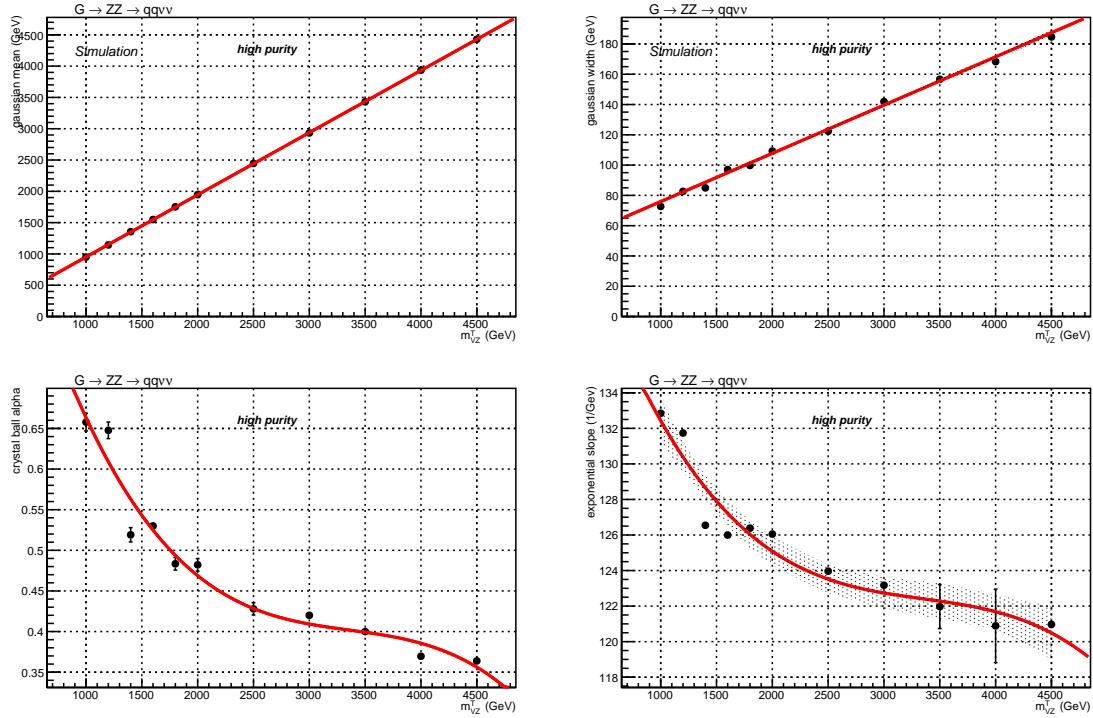


Figure 4.55: Interpolation of the fitted parameters as a function of the resonance transverse mass m_{VZ}^T , for a spin-2 (bulk graviton) signal hypothesis, high-purity category.

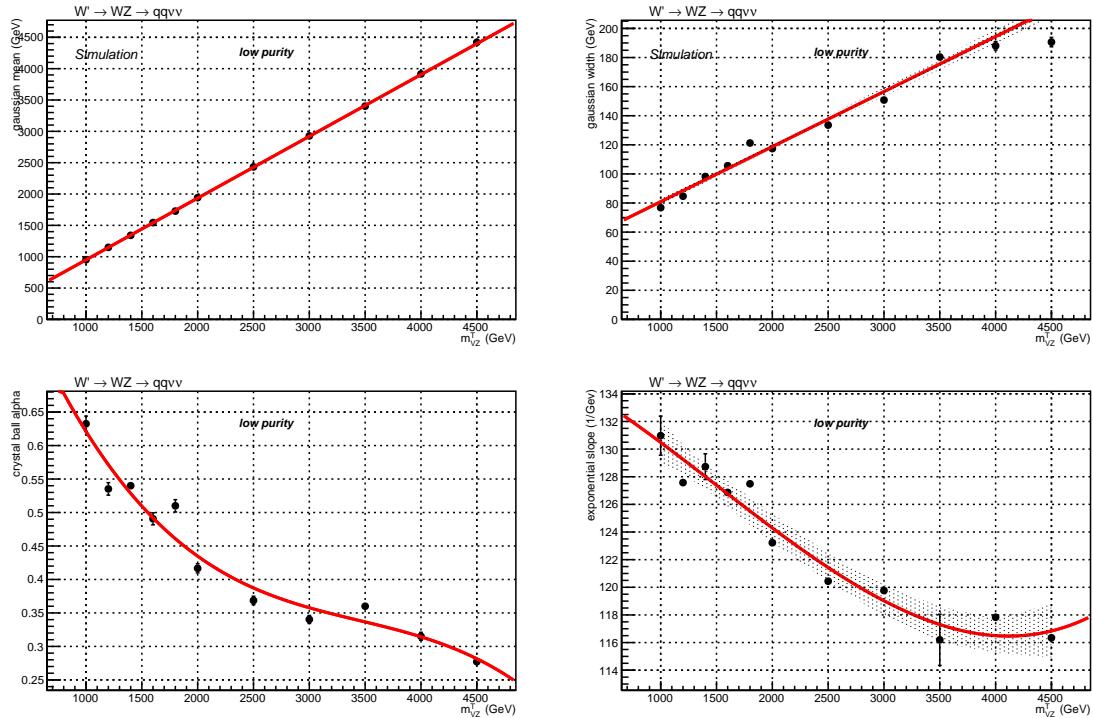


Figure 4.56: Interpolation of the fitted parameters as a function of the resonance transverse mass m_{VZ}^T , for a spin-1 (W') signal hypothesis, low-purity category.

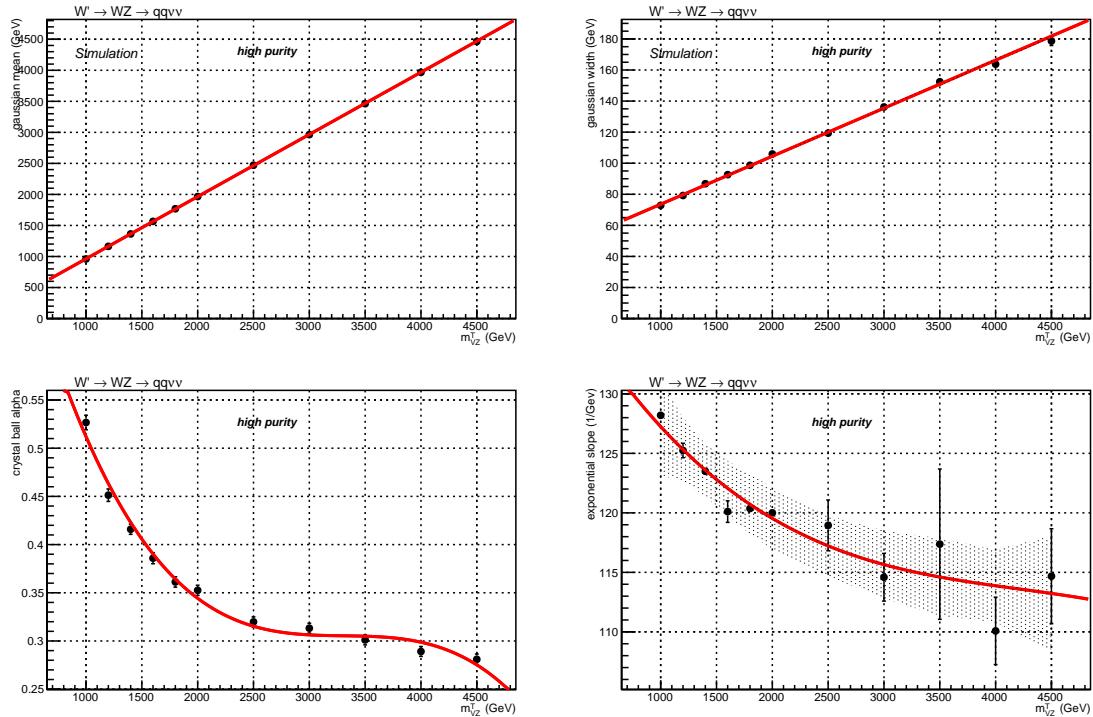


Figure 4.57: Interpolation of the fitted parameters as a function of the resonance transverse mass m_{VZ}^T , for a spin-1 (W') signal hypothesis, high-purity category.

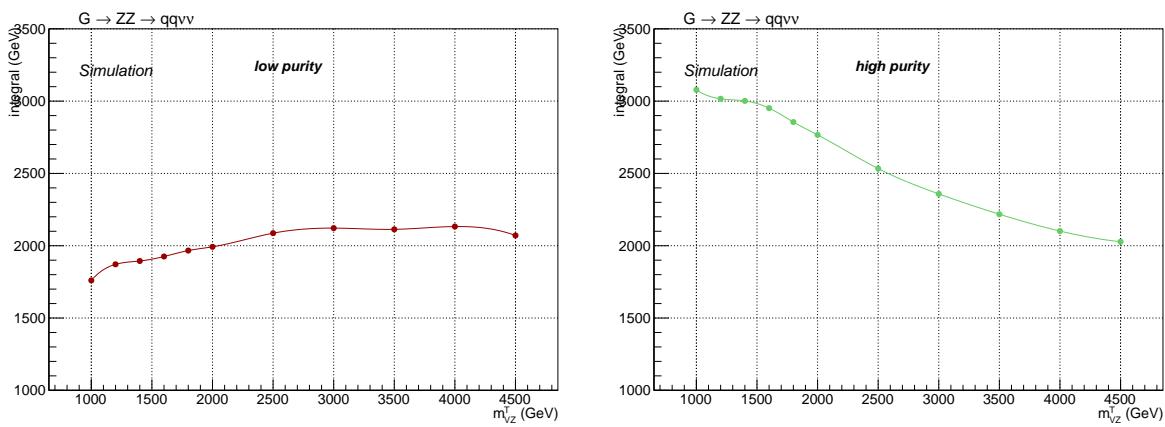


Figure 4.58: Interpolation of the signal normalization as a function of the resonance transverse mass m_{VZ}^T , for a spin-2 (bulk graviton) signal hypothesis. From left to right: low-purity, high-purity.

4.5 Background estimation technique

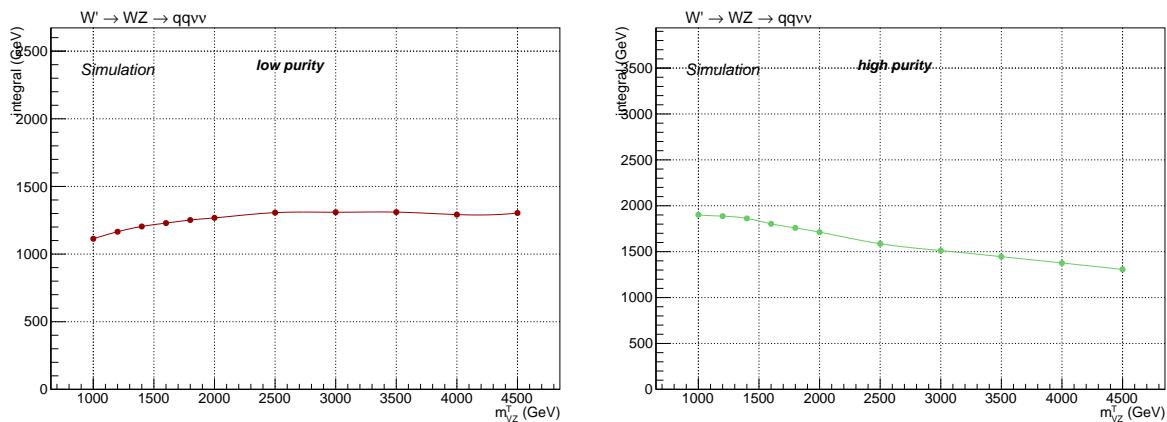


Figure 4.59: Interpolation of the signal normalization as a function of the resonance transverse mass m_{VZ}^T , for a spin-1 (W') signal hypothesis. From left to right: low-purity, high-purity.

1859 4.6 Systematic uncertainties

1860 The background and signal predictions are affected by systematic uncertainties that have to be es-
1861 timated and taken into account in the signal extraction procedure. This section includes a list of the
1862 relevant systematic uncertainties for this analysis and how they are estimated.

1863 4.6.1 Uncertainties affecting the data-driven main background estimation**1864 4.6.1.1 Normalization**

1865 The prediction of the normalization and shape of main background, $V + \text{jets}$, are both taken from
1866 data. The normalization is extracted from fits to the jet mass sidebands with arbitrary functions
1867 tested on simulation. The effects related to the contribution of the sub-dominant backgrounds are
1868 also taken into account, for both the normalization and the shape.

1869 The uncertainties on the sub-dominant backgrounds normalization, namely the uncertainties on
1870 the parameters describing the jet mass spectra obtained with the fits performed on simulations,
1871 are propagated to the main background yield prediction. An additional uncertainty on the main
1872 background yield comes from the fit with the alternative function. In this case, the difference in
1873 the predicted number of events due to the function choice is taken as a systematic uncertainty.
1874 The limited number of events in data in the sidebands is treated separately as a source of statistical
1875 uncertainty. Numerical values are reported in tab. 4.16.

1876 4.6.1.2 Shape

1877 The shape uncertainties on the main background are determined with the α method, discussed in
1878 sec. 4.5.2. The uncertainties on the parameters of the main background prediction in the signal
1879 regions are affected by the parameter uncertainties of the fit to m_{VZ}^T in data in the jet mass side-
1880 bands, and by the parameter uncertainties of the two components of the α function (numerator
1881 and denominator), that are the m_{VZ}^T fits to the simulated $V + \text{jets}$ distributions in SR and SB. These
1882 uncertainties are propagated to the shape of the main background in the signal region. Before being
1883 provided to the likelihood fit, these parameters are decorrelated through a linear transformation.

1884 4.6.2 Uncertainties affecting the signal and the sub-dominant backgrounds**1885 4.6.2.1 Trigger uncertainty**

1886 Trigger uncertainty is evaluated shifting by one standard deviation (*i.e.* 1%, as discussed in sec. 4.2.6)
1887 the E_T^{miss} trigger efficiency calculated on data, that is applied as per-event weight to MC samples.
1888 The impact has been studied in signal and secondary background samples: it amounts to 0.7–0.5%
1889 for signal samples, depending on the mass hypothesis, whilst it affects by 1% the top and diboson
1890 normalization. No effect can be appreciated in signal and background shapes.

1891 4.6.2.2 Jet momentum uncertainties

1892 Jet uncertainties are evaluated in the signal regions by moving up and down by one standard devia-
1893 tion the source of the uncertainty. The two sources are the uncertainty on the jet energy correction,
1894 also identified as jet energy scale (JES) [68], [69], and the uncertainty due to the different jet mo-
1895 mentum resolution (JER) [69].

1896 Considering the jet energy scale, the transverse momenta of the jets are shifted by the uncertainty
1897 value of the corresponding jet energy correction. The impact on the normalization due to the jet

4.6 Systematic uncertainties

- 1898 energy correction is evaluated in the signal region, by taking into account its effect on jets and on
 1899 E_T^{miss} simultaneously, in a correlated fashion.
 1900 The JER effect is evaluated (together with its impact on E_T^{miss}) by smearing the jet p_T by the η -
 1901 dependent coefficients listed in tab. 4.8, up and down by one standard deviation, using the hybrid-
 1902 method (sec. 4.3.6).
 1903 The impact of JEC uncertainties is evaluated also on the signal and background shapes. The result-
 1904 ing normalization and shape uncertainties are reported in sec. 4.6.2.6.

1905 **4.6.2.3 Jet mass uncertainties**

- 1906 The soft drop PUPPI corrected jet mass is affected by two different uncertainties sources.
 1907 Soft drop jet mass calibration is varied within $\pm \sqrt{(\text{JES}_{\text{unc.}}^2 + \text{JMS}_{\text{unc.}}^2)}$, where $\text{JES}_{\text{unc.}}$ is the uncertainty
 1908 of the JES, described above, and $\text{JMS}_{\text{unc.}} = 0.0094$ is a constant coefficient (4.3.7, [68], [69]). The
 1909 impact is calculated on signal and secondary backgrounds, both in normalization and shape.
 1910 As regarding the smearing, the soft drop PUPPI corrected jet mass of the signal samples and sub-
 1911 dominant backgrounds has been smeared up or down by a smearing coefficient (described in sec. 4.3.7),
 that is $\text{JMR} = 1.00 \pm 0.20$.

Table 4.19: Summary of jet mass energy corrections systematic uncertainties (JMS). The symbol Δ indicates the variation for each variable, due to the considered systematic uncertainty shift.

m_{VZ}^T	1 TeV	4 TeV
Δ events	1.0%	1.0%
Δ mean	0.1%	0.1%
Δ RMS	<0.1%	0.4%
secondary background	VV	Top
Δ events	0.1%	0.7%
Δ slope	<0.1%	0.2%

1912

Table 4.20: Summary of et mass resolution corrections systematic uncertainties (JMR). The symbol Δ indicates the variation for each variable, due to the considered systematic uncertainty shift.

m_{VZ}^T	1 TeV	4 TeV
Δ events	5.2%	4.9%
Δ mean	0.1%	0.1%
Δ RMS	0.4%	0.3%
secondary background	VV	Top
Δ events	2.0%	3.1%
Δ slope	1.0%	4.0%

- 1913 Results are presented in detail in tab. 4.19-4.20, for JMS and JMR uncertainties. Shape uncertainties
 1914 on signal are evaluated as the variation in the mean and variance of the transverse mass distribution.
 1915 Shape uncertainties on top and diboson backgrounds are quoted as the relative variation in the
 1916 slope of the exponential falling distribution of m_{VZ}^T , and their effects are shown in fig. 4.60-4.61.

1917 **4.6.2.4 V-tagging uncertainties**

- 1918 Data-Monte Carlo V-tagging scale factors are applied to the signal and secondary background yields
 1919 (sec. 4.3.8.1), and their uncertainty is taken as systematic. The contribution of the uncertainty is 11%

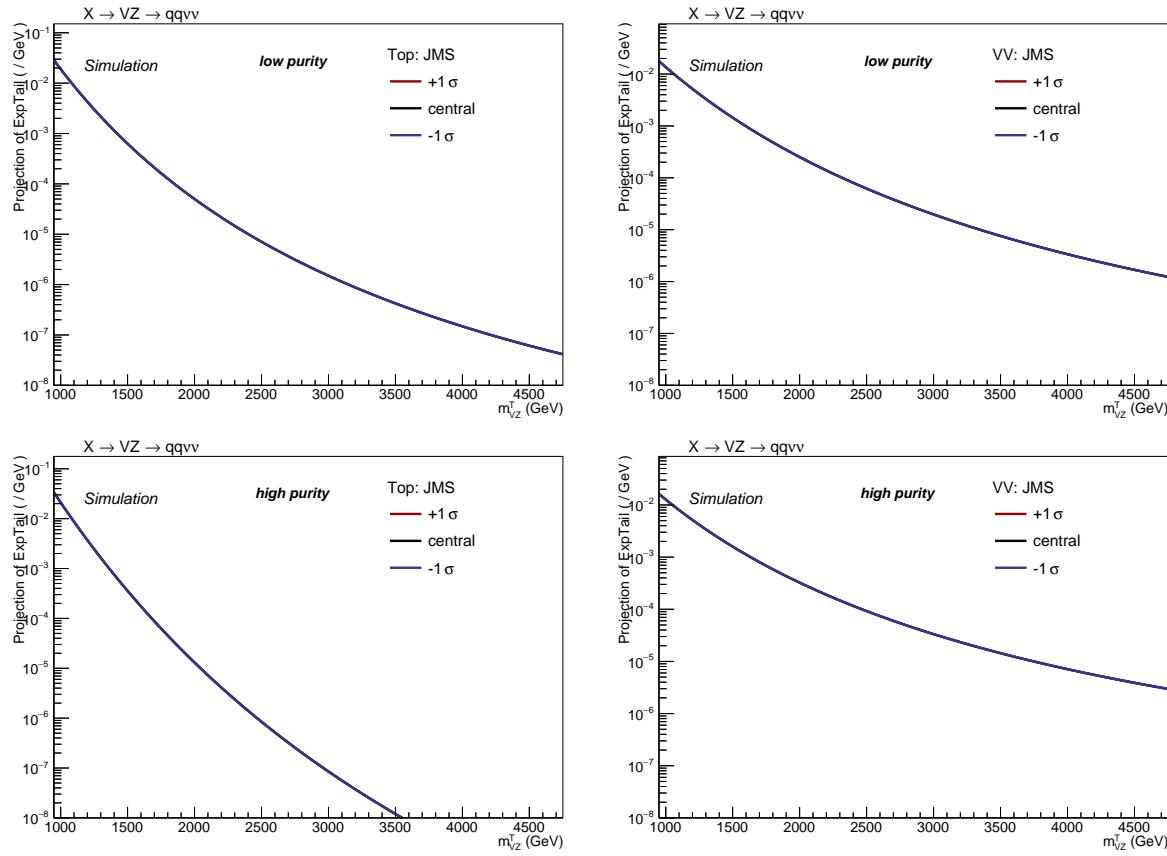


Figure 4.60: Shape variations due to jet mass calibration corrections obtained in the Top (left) and diboson (right) backgrounds, in the low-purity (top) and high-purity (bottom) category.

for the high-purity and 23% for low-purity category, applied on signal and secondary backgrounds. While combining the categories, V -tagging uncertainties are considered as anti-correlated. The V -tagging scale factors are measured in $t\bar{t}$ samples, hence at p_T values generally not larger than 200–300 GeV. An uncertainty due to the V -tagging extrapolation at higher momenta is considered by using an alternative showering scheme (HERWIG [97]). It is parametrized as a function of the jet p_T : $X \times \log(p_T/200\text{GeV})$, where $X = 0.085$ for the high-purity category and $X = 0.039$ for low-purity category. It amounts to 9–20%, depending on the mass of the signal sample considered, to 2–3% for VV and Top backgrounds in high-low purity category. While combining the categories, V -tagging extrapolation uncertainties are considered as correlated.

4.6.2.5 b-tagging uncertainties

The assigned b-tagging uncertainty, related to the b-tag veto applied to AK4 jets that lie outside the V jet cone, with the aim of suppressing the top quark induced background, is the relative difference in shape and normalization, calculated in signal and secondary background events, obtained by shifting up or down the event weight through the envelope of the data-MC b-tagging scale factors uncertainties [93]. The impact of this systematic uncertainty on signal normalization ranges from 0.7% at 1 TeV, up to 1.0% at 4 TeV. The impact on VV background normalization is 0.3%, whilst on Top it is 2.2%. Effects on signal and background shapes are negligible.

4.6 Systematic uncertainties

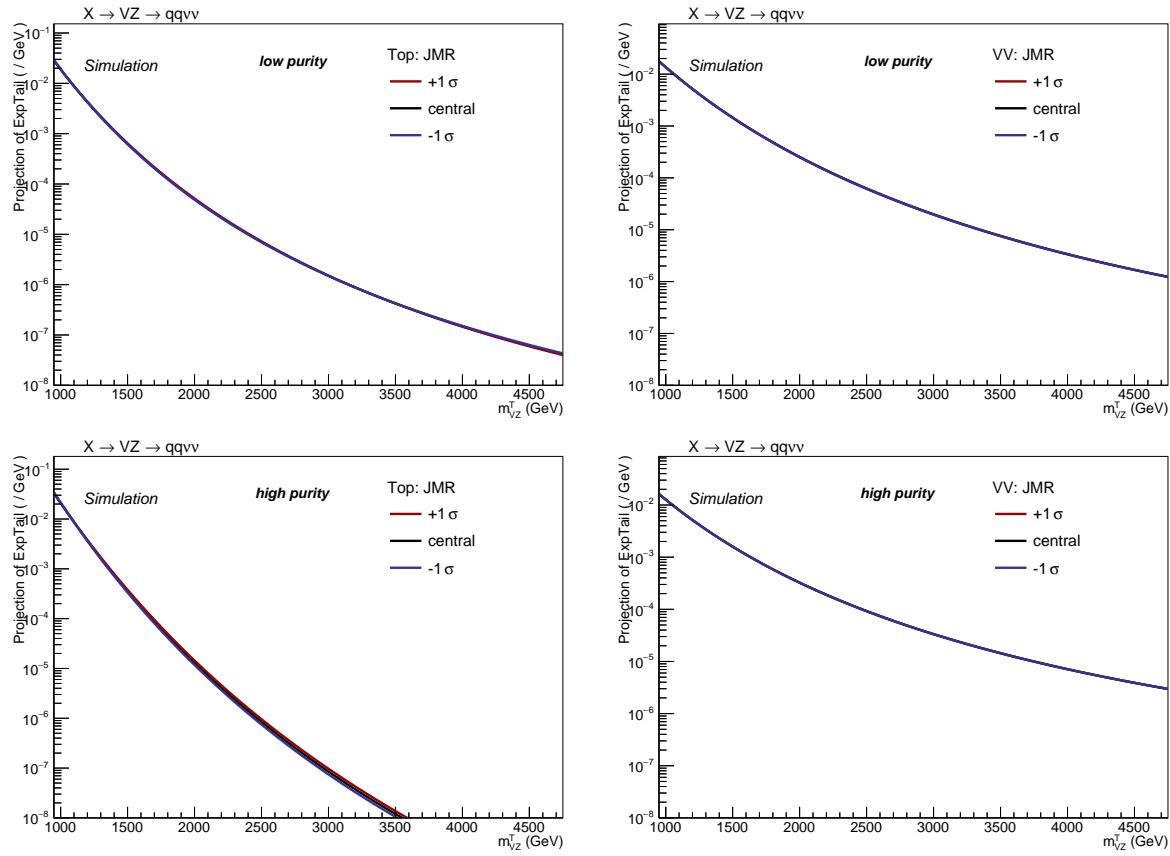


Figure 4.61: Shape variations due to jet mass resolution obtained in the Top (left) and diboson (right) backgrounds, in the low-purity (top) and high-purity (bottom) category.

1938 4.6.2.6 Missing Energy uncertainties

- 1939 As described in sec. 3.2.9.8, the E_T^{miss} evaluation depends on all the reconstructed particles in the
1940 event, and on their uncertainties. Missing energy uncertainties are calculated by factorizing \vec{p}_T^{miss}
1941 in components: electrons, photons, muons, taus, jets and unclustered energy. Dedicated uncer-
1942 tainties are derived by propagating the original object scales and resolutions to the E_T^{miss} itself.
1943 In this analysis, a leptonic veto is applied, hence the E_T^{miss} uncertainties are due to jets and unclus-
1944 tered energy. The effect of JES is evaluated on E_T^{miss} in a correlated way with jets, by scaling up or
1945 down the central value of JES by one sigma, both on E_T^{miss} and on jets p_T . The result is a negli-
1946 gible uncertainty on signal normalization, 0.2% and less than 0.1% uncertainty on top and diboson
1947 normalizations, negligible impact on signal, top and diboson shapes.
1948 The same procedure applies for the uncertainties related to jet JER, that are varied up and down by
1949 one sigma in both jets and \vec{p}_T^{miss} at the same time. The result is a negligible uncertainty on signal and
1950 diboson normalizations, 0.3% uncertainty on top normalization, and negligible effects on signal
1951 and background shapes.
1952 The last contribution in E_T^{miss} uncertainty is related to unclustered energy, whose impact is eval-
1953 uated scaling up or down the central value by its own resolution, depending on the particle type. The
1954 uncertainty is negligible on signal and background normalizations and shape.

1955 4.6.2.7 Pile-up uncertainty

1956 An additional source of systematic uncertainty is the limited knowledge of the total proton-proton
 1957 inelastic cross-section at 13 TeV, used to get the expected number of vertices distribution for the
 1958 pile-up reweighting procedure. A 4.6% uncertainty is assumed for the default value of 69 200 mb,
 1959 and the vertices distributions are varied accordingly (fig. 4.62). Changing the pile-up weight varies
 1960 also the MC normalizations in the signal region, and the relative difference is estimated to be 0.2%
 1961 for the diboson background, 0.3% for top processes, and 0.4-0.7% for signal samples. Pile-up im-
 1962 pacts on signal shapes are negligible, and it affects by 0.8% and 0.4% the diboson and top shapes
 1963 (fig. 4.63).

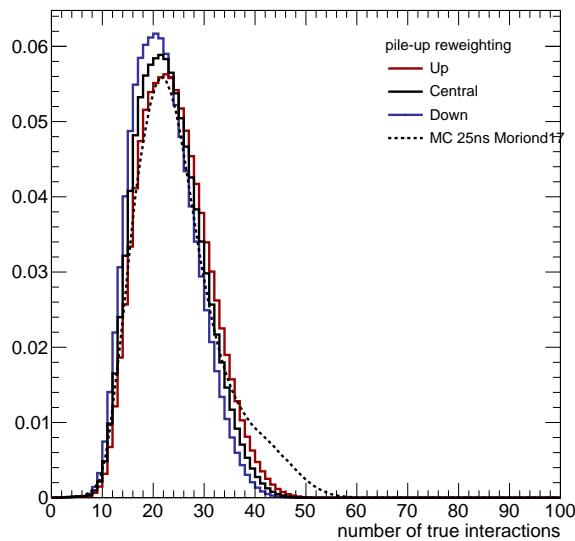


Figure 4.62: Pile-up scenario in 2016 data (black curve), and scenarios obtained by shifting up (red curve) or down (blue curve) the central value of the total inelastic cross-section (69 200 mb), compared to pile-up distribution simulated in Monte Carlo samples (dotted curve).

1964 4.6.2.8 QCD renormalization and factorization scale uncertainties

1965 Divergencies appearing in perturbative QCD calculation, used to predict the cross-sections and the
 1966 spectra of the observables in Monte Carlo simulations, are absorbed in the renormalization and
 1967 factorization scales, μ_R and μ_F . Per-event weights are calculated for a variation of these scales by
 1968 a factor 2. The two scales can be varied separately and independently, or together assuming 100%
 1969 correlation; the first approach is adopted. The weight is propagated up to the final distributions,
 1970 accounting for normalization and shape uncertainties.
 1971 The QCD variations have negligible effect on signal acceptance and on the mean and sigma of the
 1972 gaussian core of the Crystal Ball functions. The QCD factorization has an impact on top background
 1973 shape (1.1%) and normalization (3.1%), and on diboson normalization (0.9%). The QCD renorma-
 1974 lization affects the top normalization (7.3%) and diboson normalization (1.3%).

4.6 Systematic uncertainties

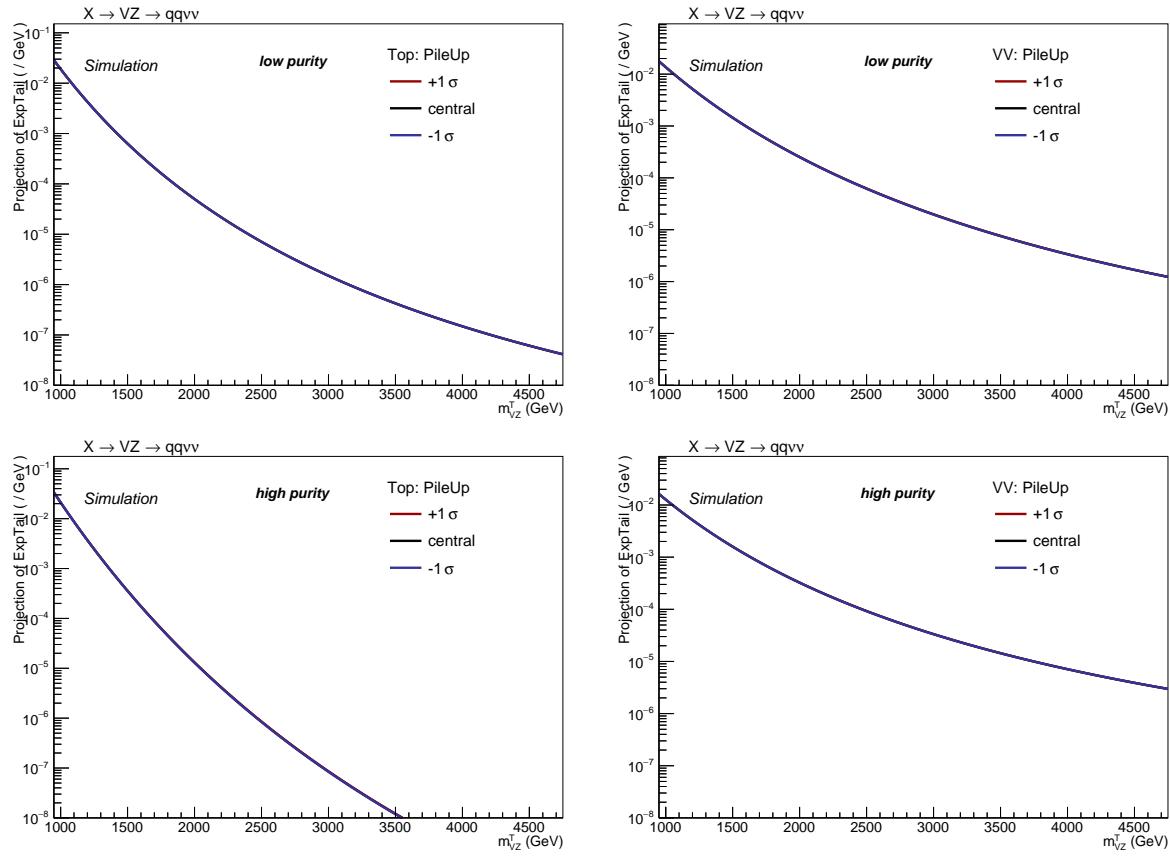


Figure 4.63: Shape variations due to pile-up uncertainty obtained in the Top (left) and diboson (right) backgrounds, in the low-purity (top) and high-purity (bottom) category.

1975 **4.6.2.9 PDF**

1976 Systematic uncertainties related to the PDFs parameters are estimated according to the PDF4LHC
 1977 prescriptions [98], and using the NNPDF3.1 [99] set. Each parameter describing the PDFs is varied
 1978 within its uncertainty, resulting in a set of per-event weights. The 100 shifted weights have been
 1979 considered together, by calculating the effect of their envelope, compared to their central values,
 1980 on the expected event yield and on the m_{VZ}^T distributions, and propagated as a normalization or
 1981 shape uncertainty. The effect of the PDF uncertainty on the signal acceptance is found to be negli-
 1982 gible, and it amounts to 10.3% for top background normalization and 2.1% for diboson background
 1983 normalization. PDF uncertainties affect top background shape by 1.2%.

1984 **4.6.3 Summary**

1985 A summary of all the systematic uncertainties is listed in tab. 4.21. In addition to those described in
 1986 the previous sections, an uncertainty of 10% on top background normalization is assumed, that is
 1987 the uncertainty on the top production cross-sections obtained from CMS measurements (sec. 4.2.3),
 1988 and an uncertainty of 15% is assigned to the diboson background normalization, due to the uncer-
 1989 tainty on the cross-section measurements performed by CMS. An additional 3% covers the uncer-
 1990 tainty related to the tau veto, and an uncertainty of 2.5% is assigned to the data integrated luminos-
 1991 ity [88].

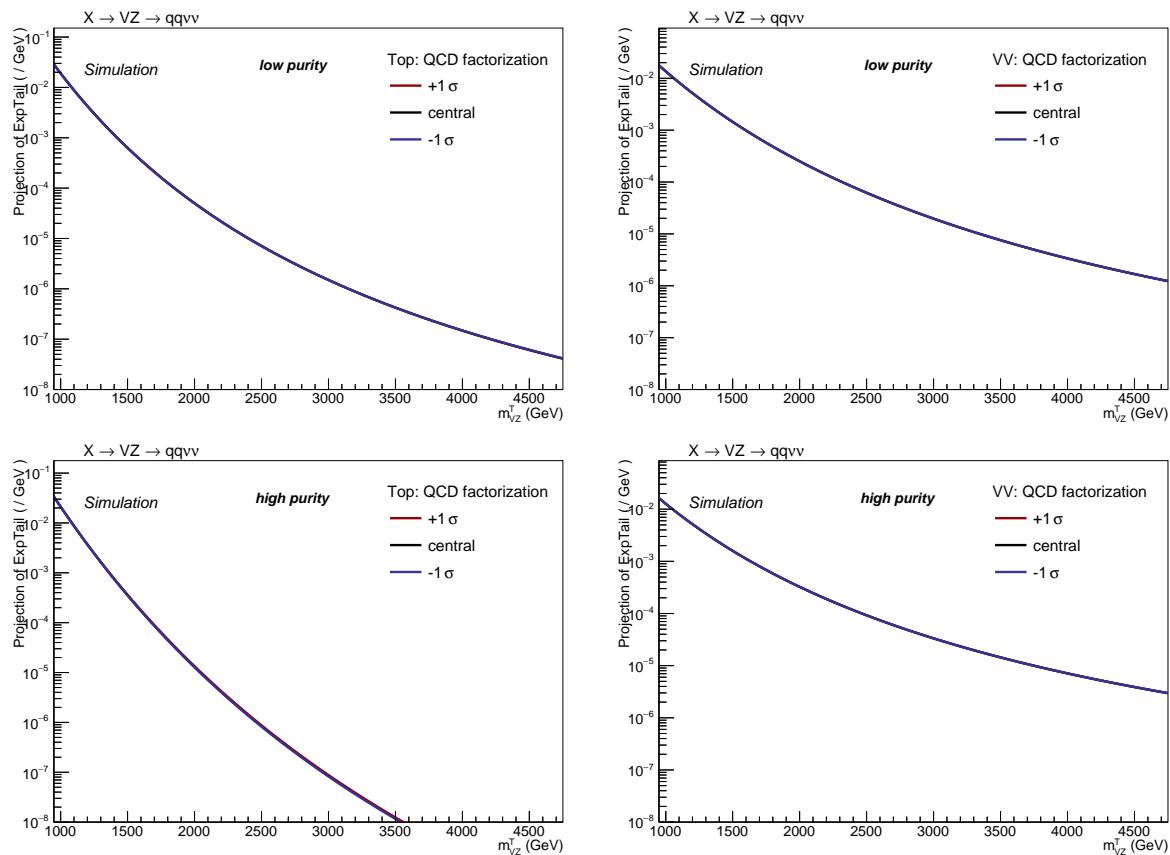


Figure 4.64: Shape variations due to QCD factorization in the Top (left) and diboson (right) backgrounds, in the low-purity (top) and high-purity (bottom) category.

4.6 Systematic uncertainties

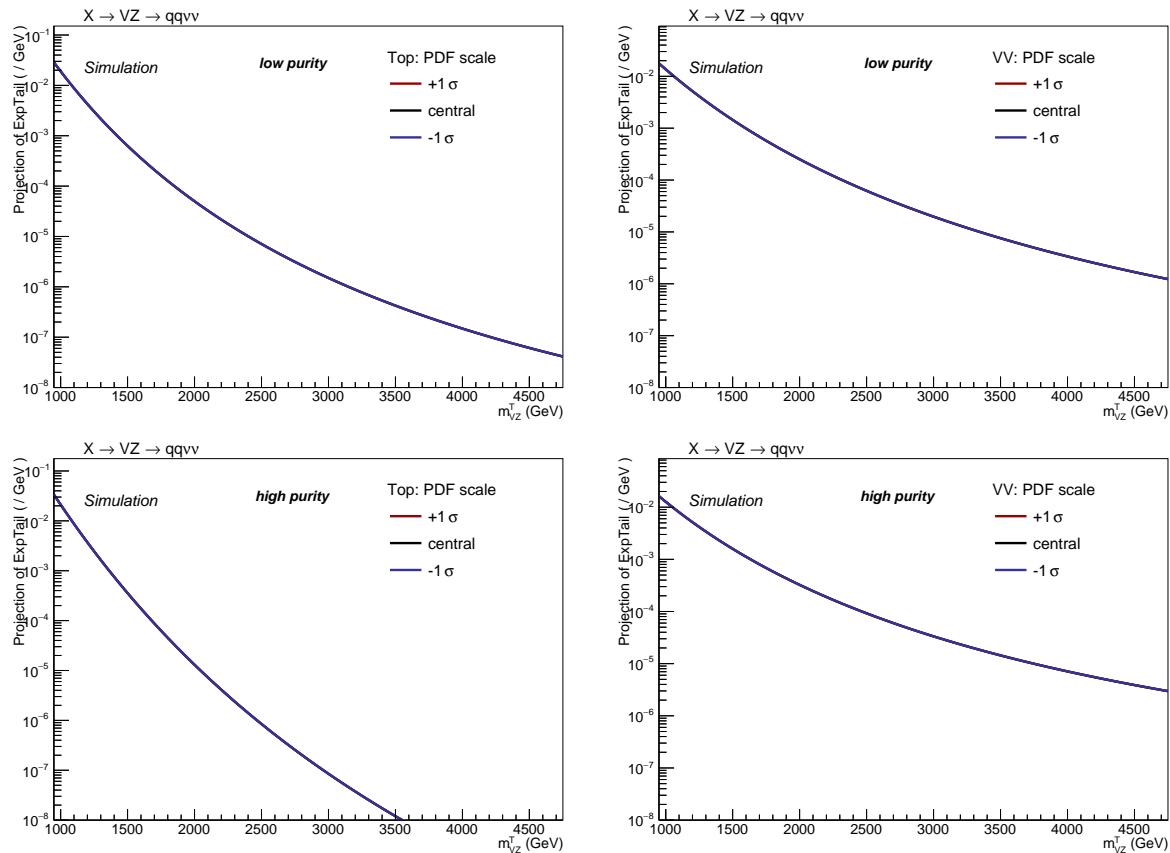


Figure 4.65: Shape variations due to PDF scale in the Top (left) and diboson (right) backgrounds, in the low-purity (top) and high-purity (bottom) category.

Table 4.21: Summary of the systematic uncertainties for the backgrounds and signal samples. LP and HP indicate the uncertainty assigned for each purity category, low- and high-purity, respectively.

	shape	$V + \text{jets}$	Top	VV	Signal
α -function	✓	✓	-	-	-
Bkg. normalization (fit)		4.8%(LP) 14.7%(HP)	68.2%(LP) 47.7%(HP)	11.4%(LP) 19.1%(HP)	-
Bkg. normalization (alternative function)		4.9%(LP) 4.4%(HP)	- -	- -	-
jet energy scale	-	-	0.2%	0.1%	<0.1%
jet energy resolution	-	-	0.3%	<0.1%	<0.1%
unclustered energy	-	-	<0.1%	<0.1%	<0.1%
jet mass scale	✓	-	0.7%	0.1%	1.8%
jet mass resolution	✓	-	3.1%	2.0%	5.1%
trigger	-	-	1.0%	0.9%	0.7-0.5%
V boson tagging (τ_{21})	-	-		11% (HP), 23% (LP)	
V tagging extrapolation	-	-	1.4% (LP) 2.8% (HP)	1.7% (LP) 3.3% (HP)	3.2-9.4% (LP) 6.9-20.6% (HP)
b-tag veto	-	-	2.2%	0.3%	0.7-1.0%
pile-up	✓	-	0.3%	0.2%	0.4-0.7%
QCD renormalization	✓	-	7.3%	1.3%	<0.1%
QCD factorization	✓	-	3.1%	0.9%	<0.1%
PDF	✓	-	10.3%	2.1%	10.4-18.9% (scale)
luminosity	-	-	2.5%	2.5%	2.5%
cross section	-	-	10%	15%	-
tau veto	-	-	3%	3%	3%

1992 4.7 Results and interpretation

1993 4.7.1 Statistical approach

1994 4.7.1.1 The modified frequentist approach: asymptotic formulae to extract an upper limit on 1995 signal strength

1996 The modified frequentist approach, also known as CL_s criterion [100–102], is used to determine the
1997 95% confidence level upper limit on the signal contribution in the data.

1998 The parameters used to model the data distribution are the background event yield, b , the sig-
1999 nal event yield s , predicted by the theoretical model, the signal strength modifier μ , parametrizing
2000 how much the signal yield deviates from the model expectation s , and the nuisance parameters
2001 θ , namely, the uncertainties affecting the signal and background yields, that can be seen as func-
2002 tions of the nuisances: $b(\theta)$, $s(\theta)$. In this approach, the uncertainties are considered either as fully
2003 correlated (100%) or uncorrelated.

2004 The likelihood function is built starting from a Poissonian probability density function:

$$\mathcal{L}(\text{data} | \mu, \theta) = \text{Poisson}(\text{data} | \mu \cdot s(\theta) + b(\theta)) \cdot p(\tilde{\theta} | \theta), \quad (4.18)$$

2005 where "data" can either be real or generated pseudo-data, whilst $p(\tilde{\theta} | \theta)$ is the probability distri-
2006 bution of the nuisance parameters, inferred through an independent dataset $\tilde{\theta}$. Considering an
2007 unbinned likelihood, where k events have been observed,

$$\text{Poisson}(\text{data} | \mu \cdot s(\theta) + b(\theta)) = \frac{1}{k} \prod_i (\mu S f_s(x_i) + B f_b(x_i)) \times e^{-(\mu S + B)}, \quad (4.19)$$

2008 where f_s and f_b are the probability density functions for signal and background for an observable
2009 x , and S and B are the total expected signal and background event yields.

2010 The measurement of the compatibility of data with the signal plus background or the background-
2011 only hypotheses is performed by defining a likelihood ratio test statistics \tilde{q}_μ [103],

$$\tilde{q}_\mu = -2 \log \frac{\mathcal{L}(\text{data} | \mu, \hat{\theta}_\mu)}{\mathcal{L}(\text{data} | \hat{\mu}, \hat{\theta})}, \quad (4.20)$$

$$0 \leq \hat{\mu} \leq \mu.$$

2012 The quantities $\hat{\mu}$ and $\hat{\theta}$ are global maxima of the likelihood, while $\hat{\theta}_\mu$ is the conditional maximum,
2013 given μ . The signal strength $\hat{\mu}$ is defined positive, the upper boundary $\hat{\mu} \leq \mu$ is set in order to avoid
2014 to consider upward fluctuations in data (namely, when the global maximum is larger than the hy-
2015 potesis μ) as an incompatibility with the signal hypothesis (μ).

2016 Given the μ hypothesis, the test statistic value is measured on data, and labelled as $\tilde{q}_\mu^{\text{obs}}$. Param-
2017 eters $\hat{\theta}_0^{\text{obs}}$ and $\hat{\theta}_\mu^{\text{obs}}$ are calculated by maximizing the likelihood function 4.18. Toy Monte Carlo
2018 pseudo-data are then generated to build the probability density functions $f(\tilde{q}_\mu | \mu, \hat{\theta}_\mu^{\text{obs}})$ (signal with
2019 μ strength hypothesis) and $f(\tilde{q}_\mu | 0, \hat{\theta}_0^{\text{obs}})$ (background-only hypothesis). Nuisance parameters are fixed
2020 to their values measured on data, $\hat{\theta}_\mu^{\text{obs}}$ and $\hat{\theta}_0^{\text{obs}}$, but left free to float in fits that are required to eval-
2021 uate \tilde{q}_μ .

2022 The p-values associated to signal plus background and background-only hypotheses are defined as:

$$p_\mu = \mathcal{P}(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}} | \text{signal + background}) = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} f(\tilde{q}_\mu | \mu, \hat{\theta}_\mu^{\text{obs}}) d\tilde{q}_\mu, \quad (4.21)$$

$$1 - p_b = \mathcal{P}(\tilde{q}_\mu \geq \tilde{q}_\mu^{\text{obs}} | \text{background-only}) = \int_{\tilde{q}_\mu^{\text{obs}}}^{\infty} f(\tilde{q}_\mu | 0, \hat{\theta}_0^{\text{obs}}) d\tilde{q}_\mu.$$

2023 The CL_s is defined as the ratio of the above p-values:

$$CL_s = \frac{p_\mu}{1 - p_b}. \quad (4.22)$$

2024 Given the a-priori confidence level α , if $CL_s \leq \alpha$, a model with signal strength μ is excluded at
2025 $(1 - \alpha)$ confidence level (C.L.). The 95% C.L. *observed* upper limit on the theoretical model is set by
2026 extracting μ from the equation $CL_s = 0.05$.

2027 Similarly to the observed limit, an upper *expected* limit, along with the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty
2028 bands, can be extracted by generating pseudo-data under the background-only hypothesis, and by by
2029 calculating the CL_s and 95% upper limit for each of the pseudo-data. A cumulative distribution
2030 of the calculated upper limits is then constructed: the 50% quantile corresponds to the median
2031 expected, the 2.5%, 16%, 84%, 95.5% quantiles correspond respectively to -2σ , -1σ , $+1\sigma$, $+2\sigma$
2032 uncertainty bands.

2033 Generating a large number of pseudo-data, however, can be a very expensive computational effort.
2034 This problem is overcome by profiting of asymptotic formulae [103], derived through Wilk's [104]
2035 and Wald's [105] theorems. The set of pseudo-data is replaced by only one dataset, the Asimov
2036 dataset: it corresponds to a dataset where the statistical fluctuations are suppressed, and hence
2037 every parameter is set to its expectation value. These values are then equivalent to the outcomes of
2038 a large sample of Monte Carlo simulations. The *expected* limit can therefore be calculated from the
2039 Asimov dataset.

2040 By using the asymptotic formulae, the distribution of the test statistic \tilde{q}_μ is given by:

$$f(\tilde{q}_\mu | \mu) = \frac{1}{2} \delta(\tilde{q}_\mu) + \begin{cases} \frac{1}{2\sqrt{2\pi}} \frac{1}{\tilde{q}_\mu} e^{-\tilde{q}_\mu/2} & 0 < \tilde{q}_\mu \leq \mu^2/\sigma^2 \\ \frac{1}{2\sqrt{2\pi}} \frac{1}{2\mu/\sigma} e^{-\frac{1}{2} \frac{(\tilde{q}_\mu + \mu^2/\sigma^2)^2}{(2\mu/\sigma)^2}} & \tilde{q}_\mu > \mu^2/\sigma^2 \end{cases}; \quad (4.23)$$

$$\sigma^2 = \frac{\mu^2}{\tilde{q}_{\mu,A}},$$

2041 where the test statistic $\tilde{q}_{\mu,A}$ is evaluated in the Asimov dataset. Once defined the inverse of the cu-
2042 mulative Gaussian distribution Φ , the asymptotic expression of the CL_s simplifies into:

$$CL_s = \frac{1 - \Phi(\sqrt{\tilde{q}_\mu})}{\Phi(\sqrt{\tilde{q}_{\mu,A}} - \sqrt{\tilde{q}_\mu})}. \quad (4.24)$$

2043 The expected upper limit and its N uncertainty bands are given by:

$$\mu_{up} = \sigma \cdot \Phi^{-1}(1 - 0.5\alpha), \quad (4.25)$$

$$\mu_{up+N} = \sigma \cdot [\Phi^{-1}(1 - \alpha\Phi(N)) + N].$$

2044 4.7.1.2 Treatment of the systematic uncertainties

2045 The nuisance parameters θ , introduced to describe the systematic uncertainties, are expected to
2046 have their own probability density function, $\rho(\theta)$, called *prior*, that is inferred by an additional set
2047 of measurements $\tilde{\theta}$, used to define the mean, the shape and the width of each uncertainty. The
2048 distribution of the priors depends on the type of uncertainty considered. Flat priors (namely, a
2049 constant value) are assigned to nuisances unconstrained a-priori; gaussian priors are assigned to
2050 nuisances allowed to assume both negative and positive values; log-normal priors are used for pos-
2051 itively defined nuisances (such as cross-sections, efficiencies, luminosity, scale factors). For the
2052 purpose of this search, log-normal priors are being adopted. Partially correlated uncertainties, *i.e.*
2053 those associated to the α method parameters, are decorrelated through linear transformations.

4.7 Results and interpretation

2054 4.7.1.3 Computation of local p-values

2055 The discovery of a signal can be inferred from data if a p-value that is incompatible with the background-only hypothesis is observed. The discovery test statistics is defined as:

$$q_0 = -2 \log \frac{\mathcal{L}(\text{data} | 0, \hat{\theta}_0)}{\mathcal{L}(\text{data} | \hat{\mu}, \hat{\theta})}, \quad (4.26)$$
$$\hat{\mu} \geq 0.$$

2057 The boundary $\hat{\mu} \geq 0$ is motivated by the fact that an underfluctuation of the background is not
2058 considered as an evidence against the background-only hypothesis. The distribution $f(q_0 | 0, \hat{\theta}_0^{obs})$
2059 is again built with pseudo-data, generated under the background-only hypothesis with nuisances
2060 $\hat{\theta}_0^{obs}$. The exact p-value is therefore:

$$p_0 = \mathcal{P}(q_0 \geq q_0^{obs} | \text{background-only}) = \int_{q_0^{obs}}^{\infty} f(q_0 | 0, \hat{\theta}_0^{obs}) dq_0, \quad (4.27)$$

2061 that can be converted into a significance Z , once the convention of the one-sided Gaussian tail is
2062 adopted:

$$p_0 = \int_Z^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx. \quad (4.28)$$

2063 By taking advantage of the Wilk's theorem, the p-value can be approximated as:

$$p_0^{\text{appr}} = \frac{1}{2} \left[1 - \text{erf} \left(\sqrt{q_0^{obs}/2} \right) \right]. \quad (4.29)$$

2064 Since the p-value depends on the phase-space considered (specifically, on the resonance mass hy-
2065 potesis), eq. 4.27 is known as the *local p-value*. A scan of the local p-values is a measurement of
2066 a local departure from the background-only hypothesis. In case of a local excess, the global signifi-
2067 cance is computed by correcting the local significance with trial factors, that take into account the
2068 so-called *look-elsewhere* effect [106], namely, the probability to observe the same excess anywhere
2069 in the whole mass range.

2070 4.7.2 Signal extraction strategy for the analysis

2071 The background prediction, estimated with the α method (sec. 4.5), the signal parametrization
2072 (sec. 4.5.4.1), and the observed data are used as inputs for the signal extraction procedure. An un-
2073 binned maximum likelihood fit is performed on each purity category, and on the combination of
2074 the categories, in order to present, for each theoretical model taken into account, a global limit on
2075 the production cross-section times branching ratio, that is the parameter describing the signal yield
2076 and defining the signal strength r (equivalent to the signal strength μ discussed in the previous sec-
2077 tion).

2078 4.7.2.1 Fit diagnostics: nuisances pulls and impacts

2079 The systematic uncertainties, treated as log-normal nuisance parameters, are allowed to vary around
2080 their nominal values and are profiled during the maximum likelihood estimation of the signal strength.
2081 As a diagnostic, the profiled values (post-fit) of the nuisance parameters $\hat{\theta}$ are compared to their
2082 a-priori expectations (pre-fit) θ_0 , in unities of the width of the gaussian core of the nuisance par-
2083 meter $\Delta\theta$. The quantities $(\hat{\theta} - \theta_0)/\Delta\theta$ are called nuisance pulls, and they have been computed both

2084 in the background-only hypothesis (blue bars) and in the signal plus background hypothesis (green
2085 bars), for the low- (fig. 4.66) and high-purity (fig. 4.67) categories. In fig. 4.66-4.67, the signal of a
2086 spin-2 bulk graviton with a mass of 3 TeV is considered. The distribution of pulls does not show any
2087 anomaly, since pulls are centered around zero (no discrepancies with the a-priori expectations) and
2088 their widths are around one (no strong deviations from the original assumption on the width of the
2089 nuisance distributions), for both the background-only and signal plus background hypotheses. The
2090 only pulls with mean values a bit shifted from zero or with widths smaller than one are related to α
2091 method parameters, that are under control.

2092 The impacts of a nuisance parameter θ are defined as the shifts induced in the signal strength (r ,
2093 the cross-section times branching fraction in this case) as θ is fixed and brought to its $+1\sigma$ or -1σ
2094 post-fit values, while all the other nuisance parameters are simultaneously profiled as log-normal.
2095 In fig. 4.68, impacts are calculated by combining the two purity categories, assuming a signal hy-
2096 potesis of a spin-2 bulk graviton of mass 2.5 TeV. As expected a-priori (sec. 4.6), the most relevant
2097 systematic impacting the determination of the signal strength is represented by the uncertainty on
2098 the V -tagging procedure. No pathological behaviour can be observed.

4.7 Results and interpretation

Figure 4.66: Nuisance pulls for the low-purity category, calculated under both the background-only (blue bars) and signal plus background hypotheses (green bars). A signal hypothesis of a spin-2 bulk graviton of mass 3 TeV is considered.

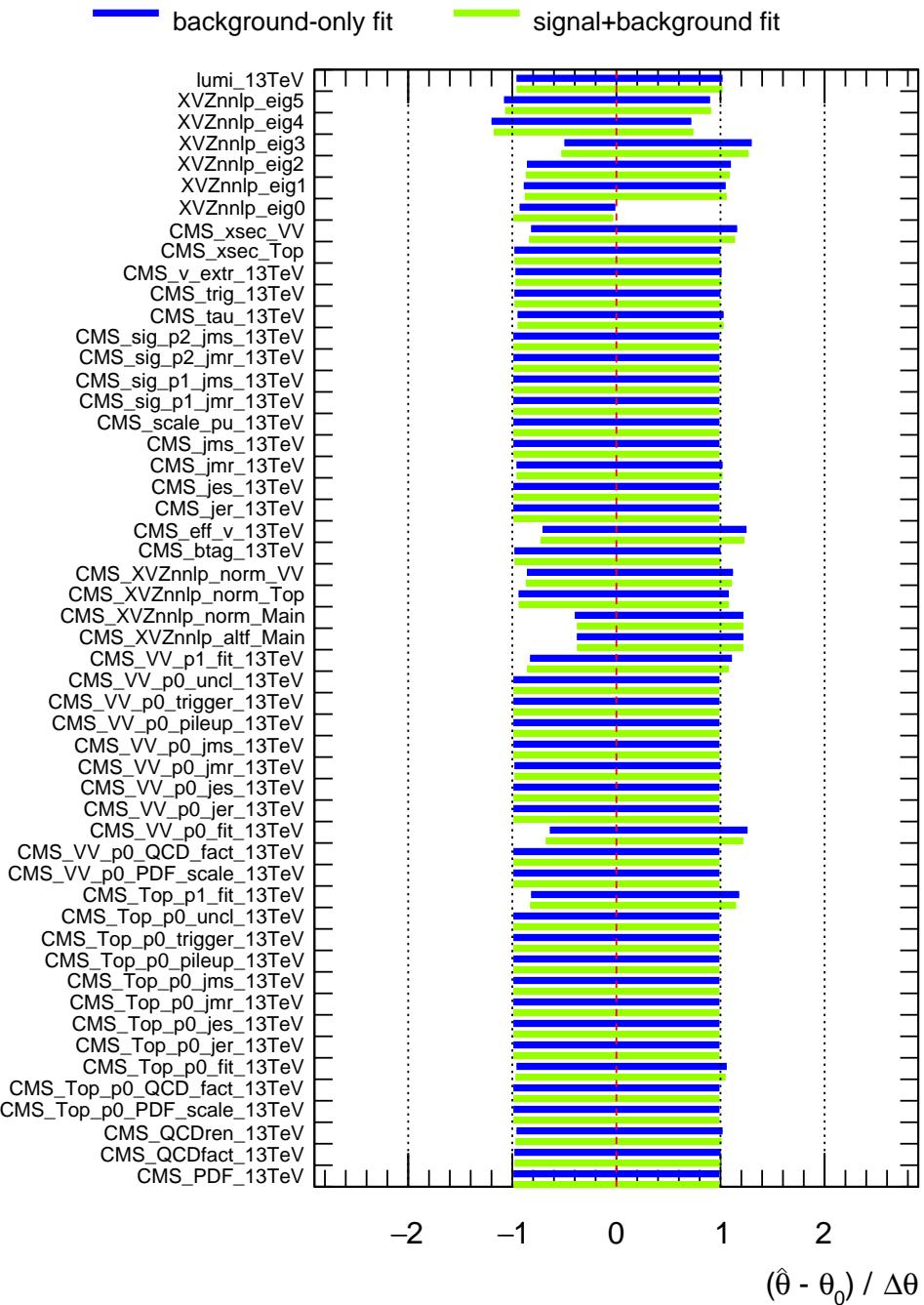
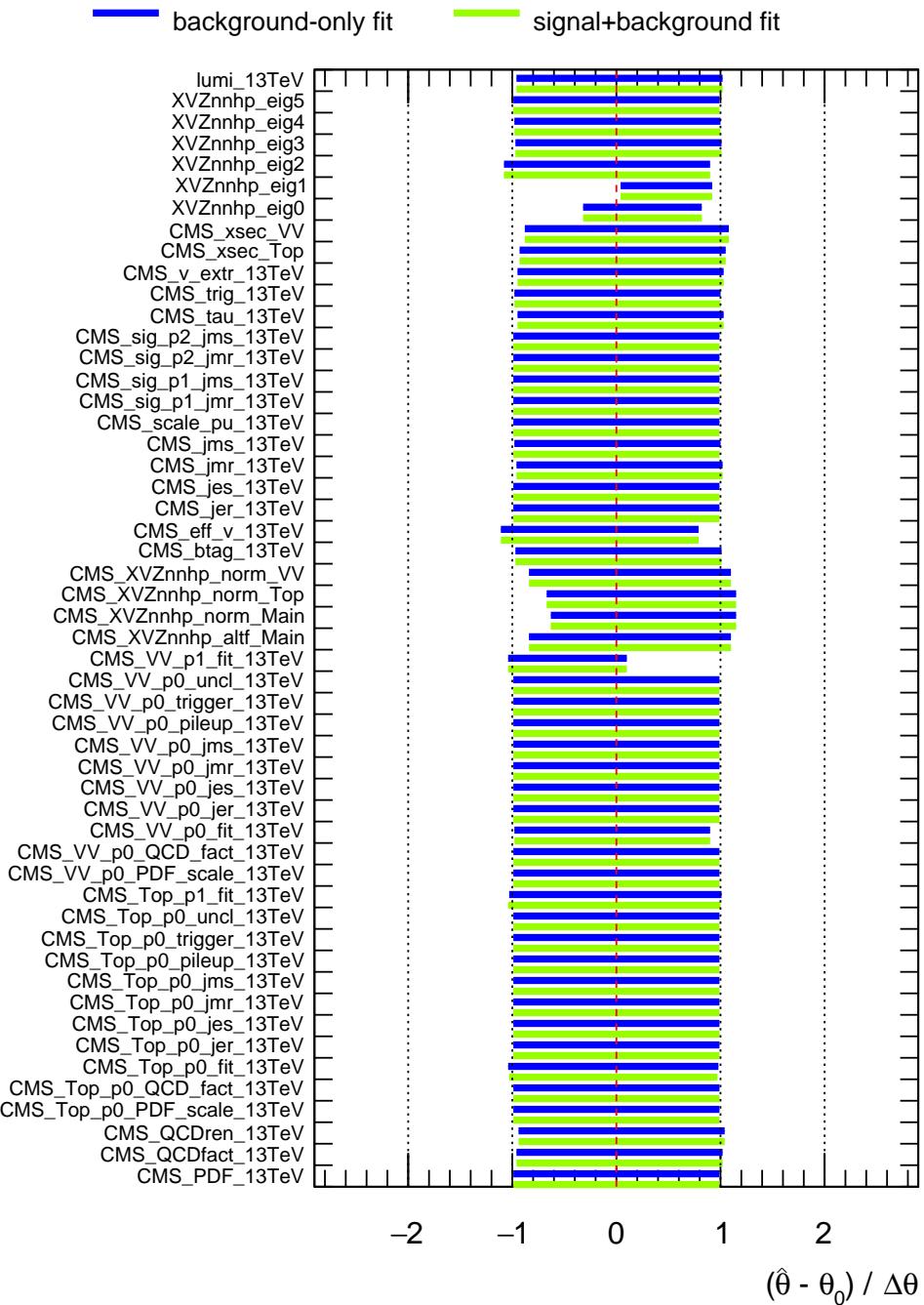
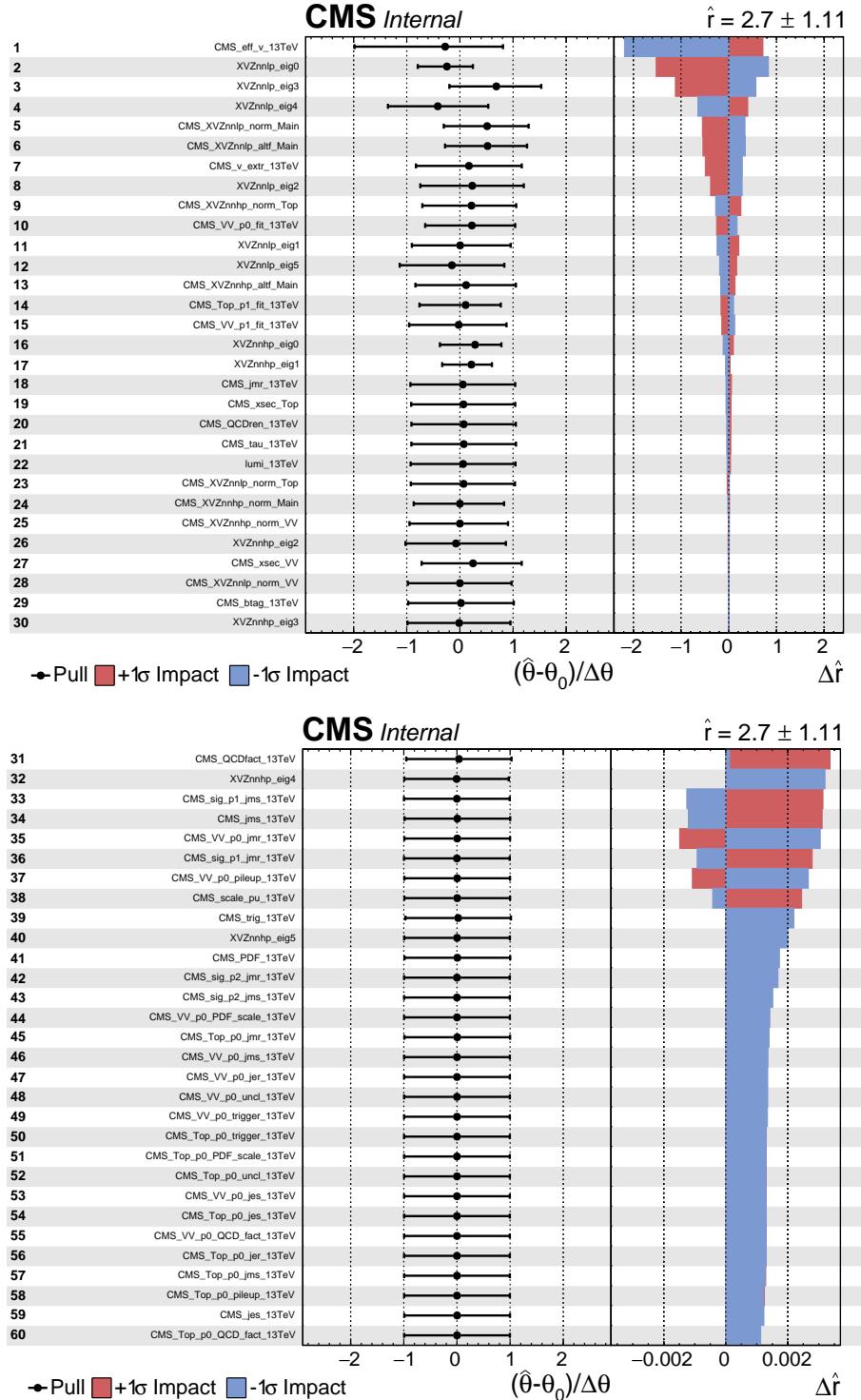


Figure 4.67: Nuisance pulls for the high-purity category, calculated under both the background-only (blue bars) and signal plus background hypotheses (green bars). A signal hypothesis of a spin-2 bulk graviton of mass 3 TeV is considered.



4.7 Results and interpretation

Figure 4.68: Impacts of the nuisance parameters on the signal strength estimation, for the combination of the low- and high-purity categories. A signal hypothesis of a spin-2 bulk graviton of mass 2.5 TeV is considered. θ_0 is the pre-fit value of the nuisance parameter taken into account; $\hat{\theta}$ is the value of the nuisance parameter after the maximum likelihood fit; $\Delta\hat{r}$ represents the impact, *i.e.* the shift induced in the parameter of interest (in this case, r , the cross-section times branching fraction, describing the signal strength) as the θ parameter is fixed and brought to its $+1\sigma$ or -1σ post-fit values, with all the other nuisance parameters profiled as log-normal.



2099 4.7.2.2 Results: expected and observed limits

2100 The observed upper limit on the resonance cross-section times branching fraction $\sigma \mathcal{B}(X \rightarrow V_{\text{had}} Z_{\text{inv}})$,
 2101 as well as the expected limit and its relative 68% and 95% uncertainty bands, are reported as a func-
 2102 tion of the resonance mass. The limits are obtained by considering separately a spin-2 bulk gravi-
 2103 ton and a spin-1 (W') heavy resonances in the narrow-width approximation. For the spin-2 case
 2104 (fig. 4.69), data are compared to theoretical predictions on $\sigma \mathcal{B}(G \rightarrow Z_{\text{had}} Z_{\text{inv}})$, obtained by impos-
 2105 ing a curvature parameter of the fifth extra-dimension $\tilde{k} = 0.5$ (red curve) and $\tilde{k} = 1.0$ (blue curve).
 2106 In case of spin-1 hypothesis (fig. 4.70), HVT model A (red curve) and model B (blue curve) theoretical
 2107 predictions on $\sigma \mathcal{B}(W' \rightarrow W_{\text{had}} Z_{\text{inv}})$ are reported.
 2108 Given the fact that the background prediction is performed in a transverse mass range $950 < m_{VZ}^T <$
 2109 4750 GeV of the resonance, and given that the higher the nominal mass of the resonance, the more
 2110 the Crystal Ball functions, parametrizing the m_{VZ}^T distributions of both spin-1 and spin-2 signals,
 2111 tend to have low-mass tails (sec. 4.5.4.1), a safe conservative criterion is to set limits in the resonance
 2112 mass range 1 TeV–4 TeV.
 2113 No significant excess is observed in data with respect to the background-only hypothesis, neither in
 2114 the low-purity, nor in the high-purity category. As it can be inferred by fig. 4.69–4.70, low-purity
 2115 category has a larger sensitivity to both spin-1 and spin-2 signals in the high mass region, whilst
 2116 high-purity category is more sensitive at low masses. This reflects the different signal efficiencies of
 2117 the two categories, as discussed in sec. 4.3.12 (fig. 4.3.12.3). By combining the two categories, the
 2118 best exclusion limits can be determined. Upper limits on the $\sigma \mathcal{B}(X \rightarrow V_{\text{had}} Z_{\text{inv}})$ of heavy spin-2 and
 2119 spin-1 narrow resonances are set in the range 0.5 – 40 fb and in the range 0.9 – 63 fb respectively.
 2120 A spin-2 bulk-graviton, once assumed a curvature parameter $\tilde{k} = 1.0$, is excluded up to 1.14 TeV. A
 2121 spin-1 W' , predicted by the model A scenario ($g_V = 1$), is excluded up to a mass of 3.11 TeV. A spin-1
 2122 W' , predicted by the model B scenario ($g_V = 3$), is excluded up to a mass of 3.41 TeV.

2123 4.7.2.3 Results: local p-value scan

2124 Scans of the local significance (left plots) and of the local p-values (right plots), as a function of
 2125 the resonance mass, are presented in fig. 4.71 (spin-2 signal) and in fig. 4.72 (spin-1 signal). No
 2126 significant deviation is observed with regards to the background-only hypothesis. The maximum
 2127 deviation is observed in the low-purity category, around 1.3 and 2.5 TeV, and it amounts to $\sim 2\sigma$.
 2128 For the combination of the categories, data are compatible with the background-only hypothesis
 2129 within 1σ in the whole mass spectrum.

4.7 Results and interpretation

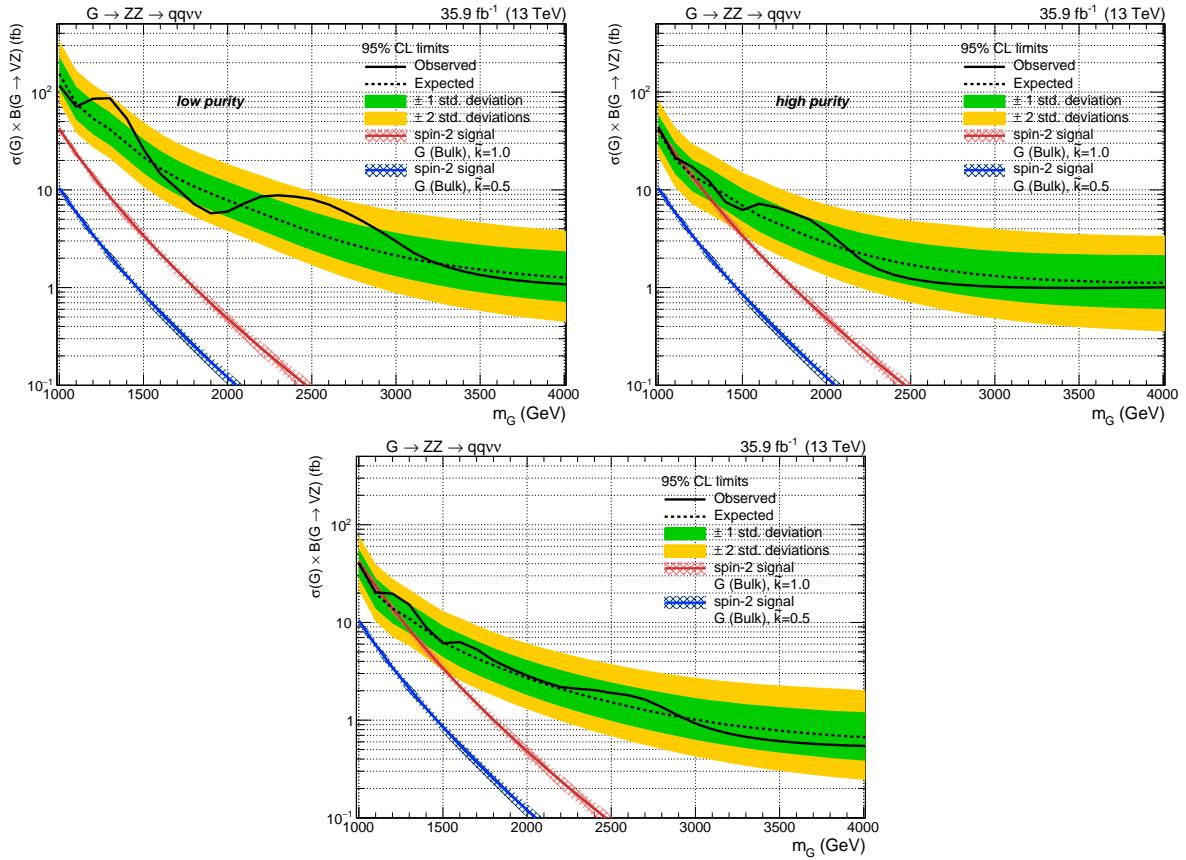


Figure 4.69: Top: observed and expected (with $\pm 1(2)\sigma$ band) 95% C.L. upper limit on $\sigma \mathcal{B}(G \rightarrow Z_{\text{had}}Z_{\text{inv}})$ for a spin-2 (bulk graviton) signal, for low-purity (left) and high-purity (right) categories, including all statistical and systematics uncertainties. Background predictions are extracted with the α method. Bottom: observed and expected (with $\pm 1(2)\sigma$ band) 95% C.L. upper limit on $\sigma \mathcal{B}(G \rightarrow Z_{\text{had}}Z_{\text{inv}})$ for a spin-2 (bulk graviton) signal, combining the two purity categories.

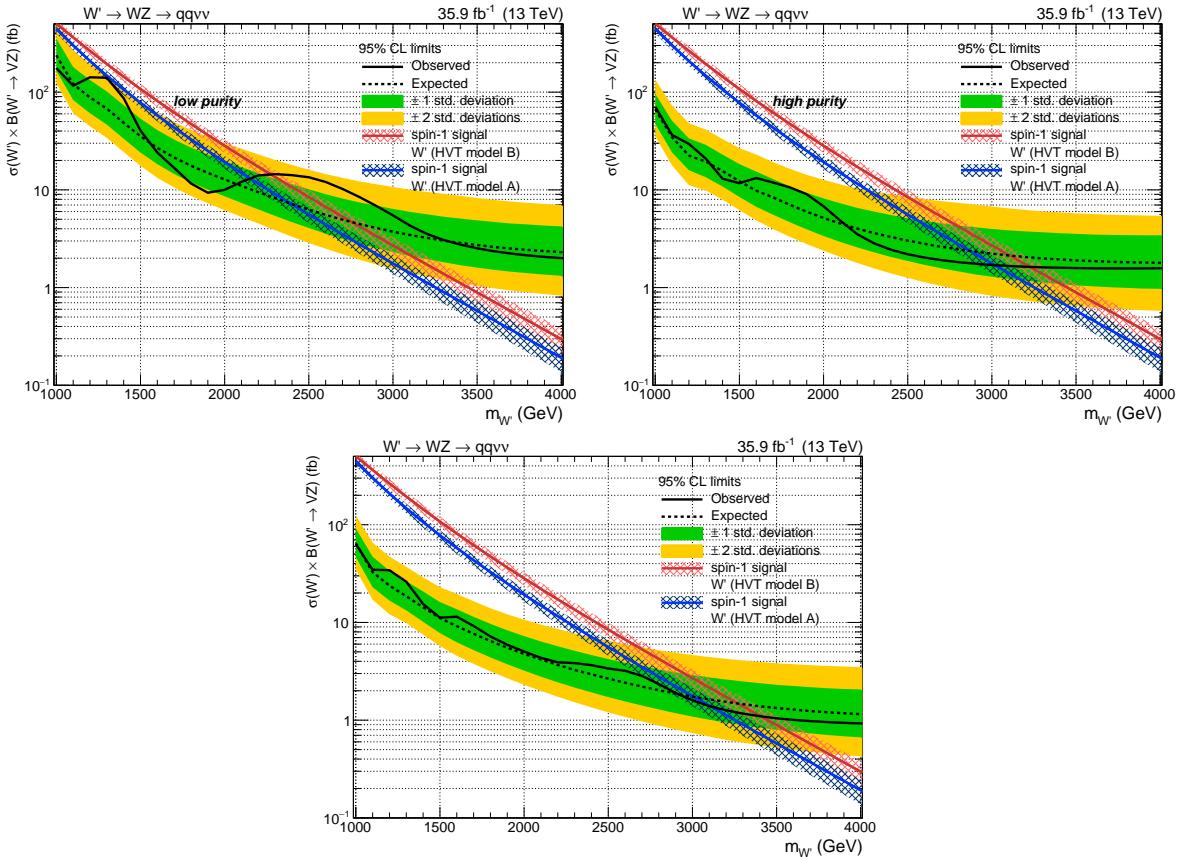


Figure 4.70: Top: observed and expected (with $\pm 1(2)\sigma$ band) 95% C.L. upper limit on $\sigma \mathcal{B}(W' \rightarrow W_{\text{had}} Z_{\text{inv}})$ for a spin-1 (HVT) signal, for low-purity (left) and high-purity (right) categories, including all statistical and systematics uncertainties. Background predictions are extracted with the α method. Bottom: observed and expected (with $\pm 1(2)\sigma$ band) 95% C.L. upper limit on $\sigma \mathcal{B}(W' \rightarrow W_{\text{had}} Z_{\text{inv}})$ for a spin-1 (HVT) signal, combining the two purity categories.

4.7 Results and interpretation

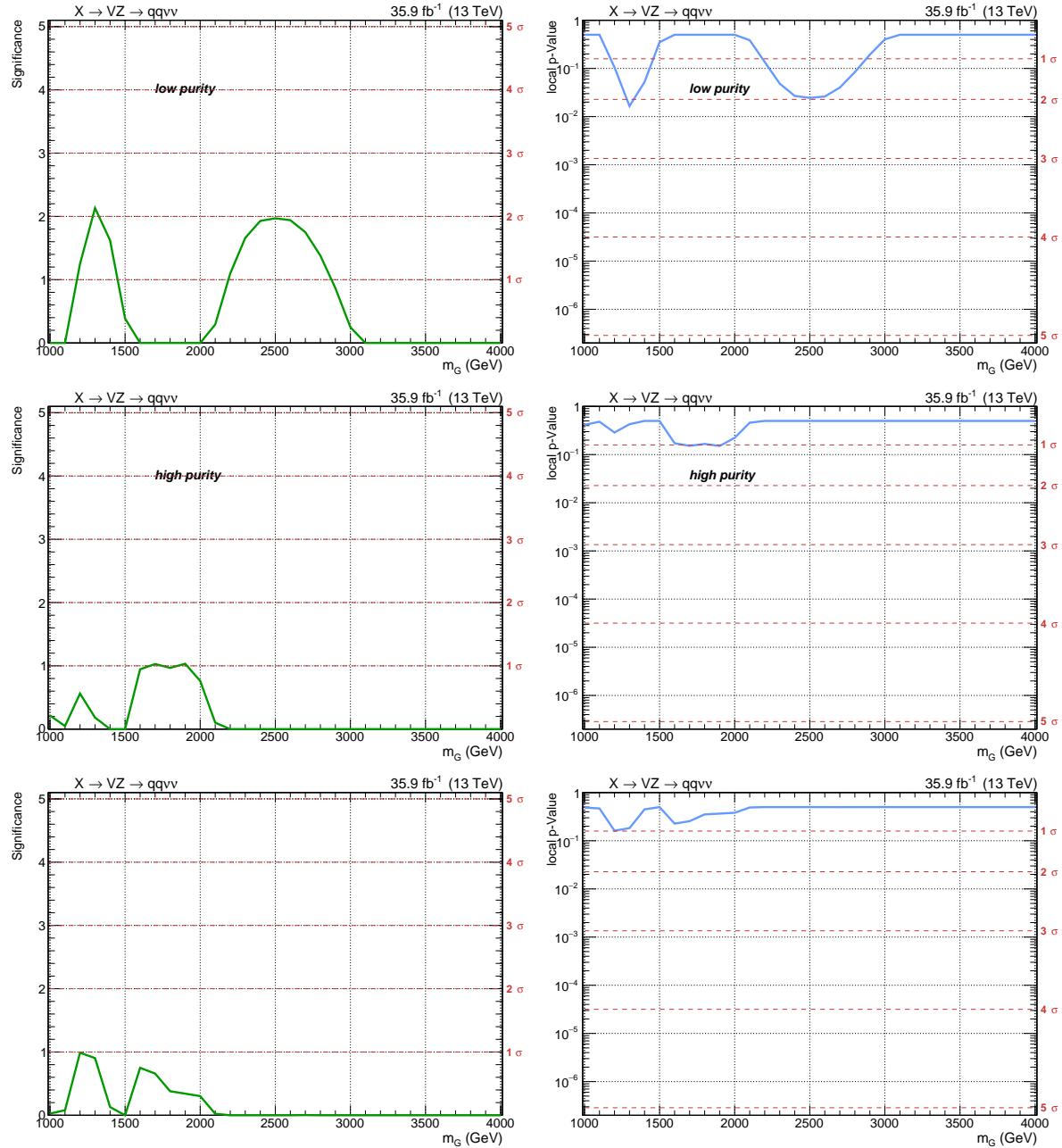


Figure 4.71: Local significances (left plots) and local p-values (right plots) as a function of the resonance mass, for a spin-2 bulk graviton hypothesis, in the low- (top), high-purity categories (center), and in the combination of the categories (bottom).

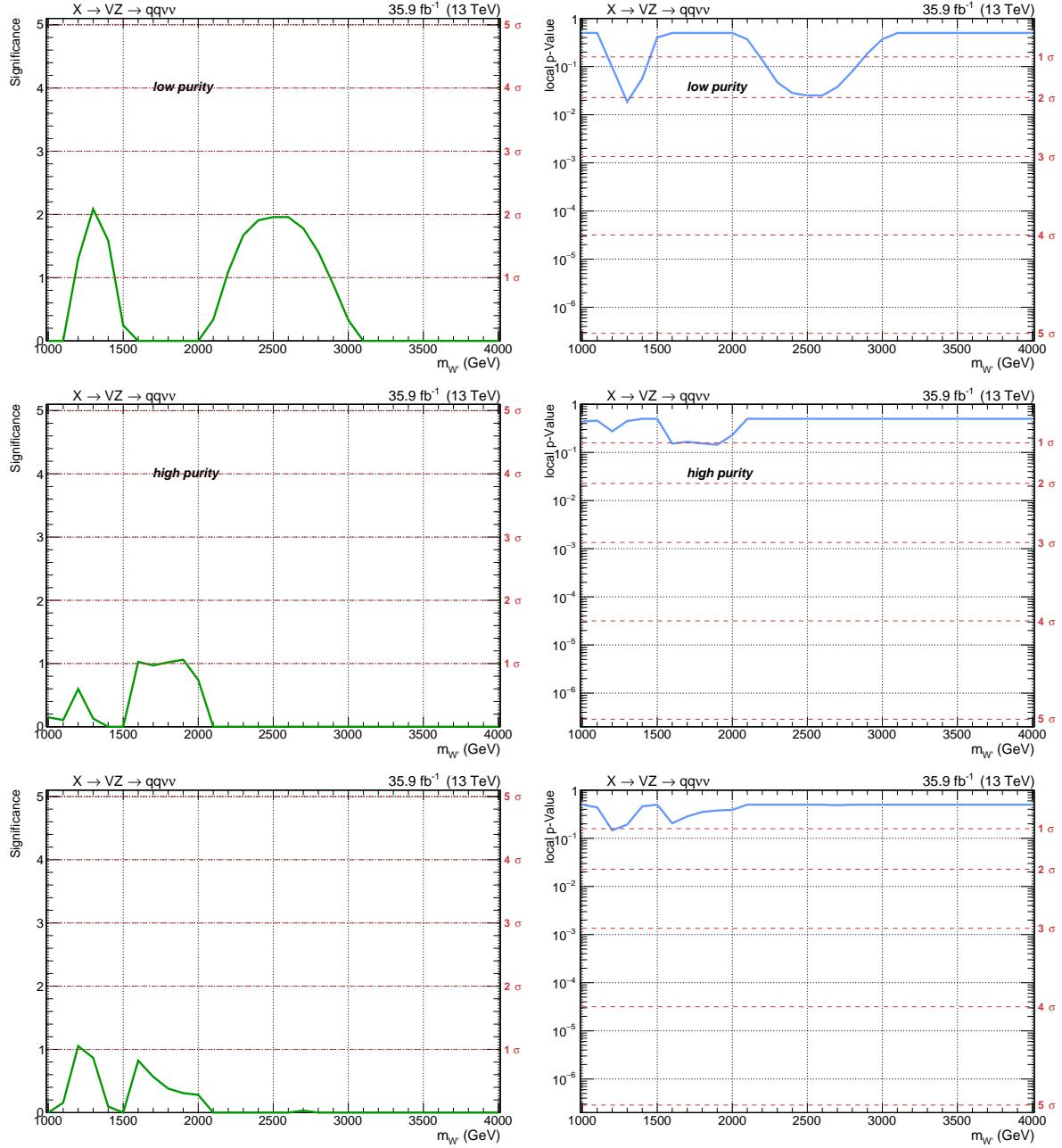


Figure 4.72: Local significances (left plots) and local p-values (right plots) as a function of the resonance mass, for a spin-1 W' hypothesis, in the low- (top), high-purity categories (center), and in the combination of the categories (bottom).

4.7.3 Interpretation of the results in the HVT model

For the HVT signal models, upper limits on the cross-section times branching fraction can be interpreted in the parameter space of the model (sec. 2.2), $(g_V c_H, g^2 c_F/g_V)$, where c_H describes the coupling of the heavy triplet to SM bosons, c_F the coupling of the triplet to SM fermions, g_V is the strength of the interaction, and g is the weak gauge coupling (sec. 2.2.1). The benchmark model A is realized when $(g_V = 1, c_H = -0.556, c_F = -1.316)$; benchmark model B scenario is realized when $(g_V = 3, c_H = 0.976, c_F = 1.024)$ [14]. This search is sensitive to the charged components of the vector triplet, namely to (W^{+}, W^{-}) . The excluded parameter space is shown in fig. 4.73. Since in the benchmark model A and model B all parameters are fixed, they are represented as a blue and a red marker respectively. The coloured curves represent the contours of the parameter space excluded by the observations in data, by considering a signal hypothesis of mass 1.5 TeV (in orange), 2 TeV (in green), 3 TeV (in violet). Currently, upper limits suggest an exclusion up to 3 TeV. The shaded gray area indicates the parameter space where the narrow width approximation fails; namely, the resonance intrinsic width becomes comparable to the experimental resolution, that amounts to 6% in this analysis (sec. 4.5.4.1).

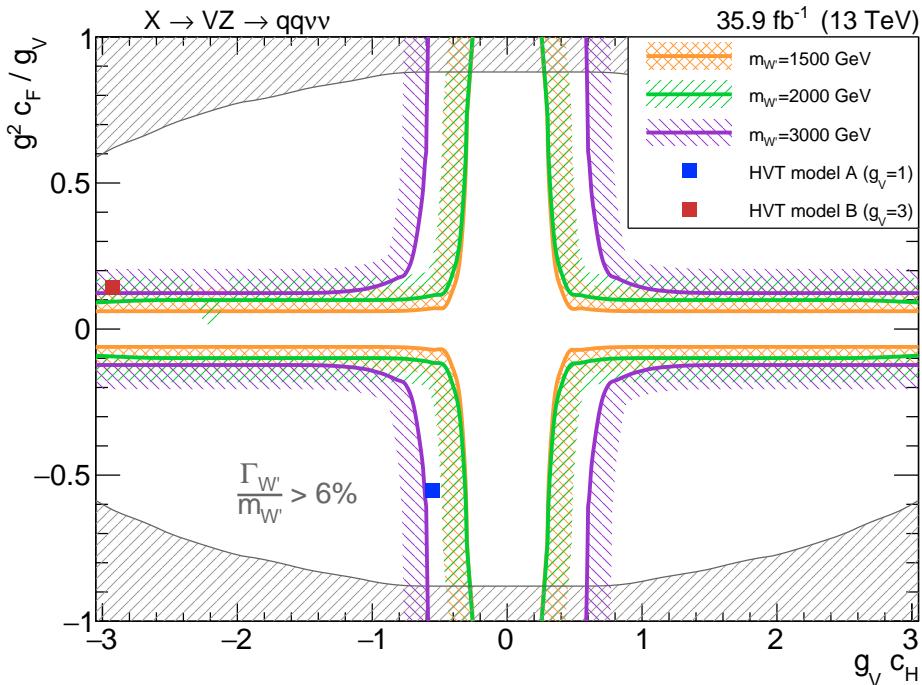


Figure 4.73: Exclusion limits on the parameter space of the HVT model. Coloured curves represent the contours of the parametric region excluded by observations in data, considering a spin-1 W' resonance of mass 1.5 TeV (in orange), 2 TeV (in green), 3 TeV (in violet). Benchmark model A and model B are represented as blue and red markers. The shaded gray area indicates the parameter space where the narrow width approximation fails.

Conclusions

2146

2147

2148 This thesis presented a search for heavy resonances with masses between 1 TeV and 4 TeV, decaying
 2149 into a pair of vector bosons, predicted by beyond standard model theories. The data produced by
 2150 LHC proton-proton collisions, at a center-of-mass energy $\sqrt{s} = 13$ TeV during the 2016 operations,
 2151 and collected by the CMS experiment, corresponding to an integrated luminosity of 35.9 fb^{-1} , are
 2152 analyzed. The probed final state includes the invisible decay modes of one Z boson, reconstructed
 2153 as a large amount of missing transverse momentum, and the hadronic decay of the other vector
 2154 boson (Z , W), reconstructed as a large-cone jet. The collected events are divided into two purity
 2155 categories, based on the substructure of the hadronically decaying V boson. No significant excesses
 2156 over the expected background are observed in the entire mass range probed by the analysis.

2157 Depending on the resonance mass, 95% C.L. upper limits on the cross-section of heavy spin-1 and
 2158 spin-2 narrow resonances, multiplied by the branching fraction of the resonance decaying into Z
 2159 and a W boson for a spin-1 signal, and into a pair of Z bosons for spin-2, are set in the range $0.9 - 63$
 2160 fb and in the range $0.5 - 40 \text{ fb}$ respectively. A W' hypothesis is excluded up to 3.11 TeV, in the context
 2161 of the Heavy Vector Triplet model A scenario, and up to 3.41 TeV, considering the model B scenario.
 2162 A bulk graviton hypothesis, given the curvature parameter $\tilde{k} = 1.0$, is excluded up to 1.14 TeV.

2163 This is the first search for $VZ \rightarrow q\bar{q} \nu\bar{\nu}$ performed by the CMS Collaboration at $\sqrt{s} = 13$ TeV. This
 2164 analysis is part of a set of searches for heavy resonances decaying into dibosons. The future perspec-
 2165 tives of the analysis consist both in the combination of this final state with other diboson searches
 2166 sharing the same treatment of one boson hadronic decay (namely, the same definition of the side-
 2167 bands and signal regions), and in the combination of the 2016 data with the newly collected 2017
 2168 data. The luminosity planned to be delivered by the LHC collider in 2017 is comparable to what
 2169 was collected in 2016 ($\sim 40 \text{ fb}^{-1}$). By doubling the statistics, marginal improvements are foreseen;
 2170 hence, a larger enhancement can be achieved by decreasing the impacts of the systematic uncer-
 2171 tainties. New jet substructure techniques are currently being tested, in order to improve the jet mass
 2172 resolution (recursive soft drop), suppress the pile-up contribution (PUPPI associated to SoftKiller
 2173 algorithm [107]), exploit the jet substructure and tag the nature of a large-cone jet (originating from
 2174 W , Z , Higgs boson or top quark) with machine learning techniques. Preliminary results on these
 2175 new methods seem to be promising.

2176

Bibliography

2177

2178

- 2179 [1] M. Baak and R. Kogler, *The global electroweak Standard Model fit after the Higgs discovery*, in
2180 *Proceedings, 48th Rencontres de Moriond on Electroweak Interactions and Unified Theories:*
2181 *La Thuile, Italy, March 2-9, 2013*, pp. 349–358, 2013, 1306 .0571,
2182 <https://inspirehep.net/record/1236809/files/arXiv:1306.0571.pdf>.
- 2183 [2] ATLAS Collaboration, G. Aad et al., *Observation of a new particle in the search for the*
2184 *Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1,
2185 [1207 .7214].
- 2186 [3] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV*
2187 *with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30, [1207 .7235].
- 2188 [4] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson with mass near 125 GeV*
2189 *in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, *JHEP* **06** (2013) 081, [1303 .4571].
- 2190 [5] ATLAS Collaboration, G. Aad et al., *Evidence for the spin-0 nature of the Higgs boson using*
2191 *ATLAS data*, *Phys. Lett. B* **726** (2013) 120, [1307 .1432].
- 2192 [6] CMS Collaboration, V. Khachatryan et al., *Precise determination of the mass of the Higgs*
2193 *boson and tests of compatibility of its couplings with the standard model predictions using*
2194 *proton collisions at 7 and 8 TeV*, *Eur. Phys. J. C* **75** (2015) 212, [1412 .8662].
- 2195 [7] ATLAS Collaboration, G. Aad et al., *Measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$*
2196 *and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with*
2197 *the ATLAS detector*, *Phys. Rev. D* **90** (2014) 052004, [1406 .3827].
- 2198 [8] CMS and ATLAS Collaborations, “Combined Measurement of the Higgs Boson Mass in pp
2199 *Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments.”* 2015.
- 2200 [9] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, G. Isidori et al., *Higgs mass*
2201 *and vacuum stability in the Standard Model at NNLO*, *JHEP* **08** (2012) 098, [1205 .6497].
- 2202 [10] S. P. Martin, *A Supersymmetry primer*, hep-ph/9709356.
- 2203 [11] C. Csaki, *The Minimal supersymmetric standard model (MSSM)*, *Mod. Phys. Lett. A* **11** (1996)
2204 599, [hep-ph/9606414].

- 2205 [12] D. J. Castano, E. J. Piard and P. Ramond, *Renormalization group study of the Standard Model*
2206 *and its extensions. 2. The Minimal supersymmetric Standard Model*, *Phys. Rev.* **D49** (1994)
2207 4882–4901, [[hep-ph/9308335](#)].
- 2208 [13] H. E. Haber, *Introductory low-energy supersymmetry*, in *Proceedings, Theoretical Advanced*
2209 *Study Institute (TASI 92): From Black Holes and Strings to Particles: Boulder, USA, June 1-26,*
2210 1992, pp. 589–686, 1993, [hep-ph/9306207](#).
- 2211 [14] D. Pappadopulo, A. Thamm, R. Torre and A. Wulzer, *Heavy vector triplets: bridging theory*
2212 *and data*, *Journal of High Energy Physics* **2014** (2014) 1–50.
- 2213 [15] V. D. Barger, W.-Y. Keung and E. Ma, *A Gauge Model With Light W and Z Bosons*, *Phys. Rev.*
2214 **D22** (1980) 727.
- 2215 [16] C. Grojean, E. Salvioni and R. Torre, *A weakly constrained W' at the early LHC*, *JHEP* **07** (2011)
2216 002, [[1103.2761](#)].
- 2217 [17] R. Contino, D. Pappadopulo, D. Marzocca and R. Rattazzi, *On the effect of resonances in*
2218 *composite higgs phenomenology*, *Journal of High Energy Physics* **2011** (2011) 1–50.
- 2219 [18] B. Bellazzini, C. Csáki and J. Serra, *Composite Higgsses*, *Eur. Phys. J.* **C74** (2014) 2766,
2220 [[1401.2457](#)].
- 2221 [19] G. F. Giudice, C. Grojean, A. Pomarol and R. Rattazzi, *The Strongly-Interacting Light Higgs*,
2222 *JHEP* **06** (2007) 045, [[hep-ph/0703164](#)].
- 2223 [20] ATLAS Collaboration, G. Aad et al., *Evidence for Electroweak Production of $W^\pm W^\pm jj$ in pp*
2224 *Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, *Phys. Rev. Lett.* **113** (2014) 141803,
2225 [[1405.6241](#)].
- 2226 [21] CMS Collaboration, V. Khachatryan et al., *Study of vector boson scattering and search for new*
2227 *physics in events with two same-sign leptons and two jets*, *Phys. Rev. Lett.* **114** (2015) 051801,
2228 [[1410.6315](#)].
- 2229 [22] A. Wulzer, *An Equivalent Gauge and the Equivalence Theorem*, *Nucl. Phys.* **B885** (2014)
2230 97–126, [[1309.6055](#)].
- 2231 [23] CMS Collaboration, A. M. Sirunyan et al., *Search for massive resonances decaying into WW,*
2232 *WZ, ZZ, qW, and qZ with dijet final states at $\sqrt{s} = 13$ TeV*, [1708.05379](#).
- 2233 [24] CMS Collaboration, *Search for massive resonances decaying into WW, WZ, ZZ, qW and qZ in*
2234 *the dijet final state at $\sqrt{s} = 13$ TeV*, CMS Physics Analysis Summary CMS-PAS-B2G-17-001,
2235 CERN, Geneva, 2017. <http://cds.cern.ch/record/2256663>.
- 2236 [25] CMS Collaboration, CMS Collaboration, *Search for heavy resonances decaying into a vector*
2237 *boson and a Higgs boson in hadronic final states with 2016 data*, .
- 2238 [26] CMS Collaboration, *Search for heavy resonances decaying into a vector boson and a Higgs*
2239 *boson in hadronic final states with 2016 data*, CMS Physics Analysis Summary
2240 CMS-PAS-B2G-17-002, CERN, Geneva, 2017. <http://cds.cern.ch/record/2256742>.
- 2241 [27] CMS Collaboration, *Search for heavy resonances decaying into a Z boson and a W boson in*
2242 *the $\ell^+\ell^-q\bar{q}$ final state*, CMS Physics Analysis Summary CMS-PAS-B2G-16-022, CERN,
2243 Geneva, 2017. <http://cds.cern.ch/record/2242955>.

BIBLIOGRAPHY

- [2244] [28] CMS Collaboration, V. Khachatryan et al., *Search for heavy resonances decaying into a vector boson and a Higgs boson in final states with charged leptons, neutrinos, and b quarks*, *Phys. Lett.* **B768** (2017) 137–162, [[1610.08066](#)].
- [2245]
- [2246]
- [2247] [29] CMS Collaboration, *Search for new resonances decaying to WW/WZ → ℓ νqq*, CMS Physics Analysis Summary CMS-PAS-B2G-16-020, CERN, Geneva, 2016.
<http://cds.cern.ch/record/2205880>.
- [2248]
- [2249]
- [2250] [30] CMS Collaboration, *Search for heavy resonances decaying into a Z boson and a vector boson in the νν q̄q̄ final state*, CMS Physics Analysis Summary CMS-PAS-B2G-17-005, CERN, Geneva, 2017. <http://cds.cern.ch/record/2273910>.
- [2251]
- [2252]
- [2253] [31] ATLAS Collaboration, M. Aaboud et al., *Search for diboson resonances with boson-tagged jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [1708.04445](#).
- [2254]
- [2255] [32] ATLAS Collaboration, *Search for WW/WZ resonance production in ℓ νqq final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATLAS Conference Note ATLAS-CONF-2017-051, CERN, Geneva, Jul, 2017. <http://cds.cern.ch/record/2273867>.
- [2256]
- [2257]
- [2258] [33] ATLAS Collaboration, *Searches for heavy ZZ and ZW resonances in the llqq and vvqq final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATLAS Conference Note ATLAS-CONF-2016-082, CERN, Geneva, Aug, 2016. <https://cds.cern.ch/record/2206275>.
- [2259]
- [2260]
- [2261] [34] ATLAS Collaboration, M. Aaboud et al., *Search for heavy resonances decaying to a W or Z boson and a Higgs boson in the q̄q̄(ℓ) b̄b final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [1707.06958](#).
- [2262]
- [2263]
- [2264] [35] ATLAS Collaboration, *Search for heavy resonances decaying to a W or Z boson and a Higgs boson in final states with leptons and b-jets in 36.1 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATLAS Conference Note ATLAS-CONF-2017-055, CERN, Geneva, Jul, 2017. <http://cds.cern.ch/record/2273871>.
- [2265]
- [2266]
- [2267]
- [2268] [36] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, *Phys. Rev. Lett.* **83** (1999) 3370–3373, [[hep-ph/9905221](#)].
- [2269]
- [2270] [37] L. Randall and R. Sundrum, *An Alternative to compactification*, *Phys. Rev. Lett.* **83** (1999) 4690–4693, [[hep-th/9906064](#)].
- [2271]
- [2272] [38] K. Agashe, H. Davoudiasl, G. Perez and A. Soni, *Warped Gravitons at the LHC and Beyond*, *Phys. Rev.* **D76** (2007) 036006, [[hep-ph/0701186](#)].
- [2273]
- [2274] [39] A. L. Fitzpatrick, J. Kaplan, L. Randall and L.-T. Wang, *Searching for the Kaluza-Klein Graviton in Bulk RS Models*, *JHEP* **09** (2007) 013, [[hep-ph/0701150](#)].
- [2275]
- [2276] [40] H. Davoudiasl, J. L. Hewett and T. G. Rizzo, *Phenomenology of the Randall-Sundrum Gauge Hierarchy Model*, *Phys. Rev. Lett.* **84** (2000) 2080, [[hep-ph/9909255](#)].
- [2277]
- [2278] [41] T. Gherghetta and A. Pomarol, *Bulk fields and supersymmetry in a slice of AdS*, *Nucl. Phys.* **B586** (2000) 141–162, [[hep-ph/0003129](#)].
- [2279]
- [2280] [42] H. Davoudiasl, J. L. Hewett and T. G. Rizzo, *Experimental probes of localized gravity: On and off the wall*, *Phys. Rev.* **D63** (2001) 075004, [[hep-ph/0006041](#)].
- [2281]

- 2282 [43] A. Oliveira, *Gravity particles from Warped Extra Dimensions, predictions for LHC*,
2283 1404.0102.
- 2284 [44] CMS Collaboration, *Search for diboson resonances in the $2l2\nu$ final state*, CMS Physics
2285 Analysis Summary CMS-PAS-B2G-16-023, CERN, Geneva, 2017.
2286 <http://cds.cern.ch/record/2264700>.
- 2287 [45] CMS Collaboration, *Search for heavy resonances decaying to a pair of Higgs bosons in the four*
2288 *b quark final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, CMS Physics Analysis
2289 Summary CMS-PAS-B2G-16-026, CERN, Geneva, 2017. <http://cds.cern.ch/record/2264684>.
- 2290 [46] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.
- 2291 [47] CMS Collaboration, S. Chatrchyan et al., *The CMS experiment at the CERN LHC*, JINST **3**
2292 (2008) S08004.
- 2293 [48] V. Karimäki, M. Mannelli, P. Siegrist, H. Breuker, A. Caner, R. Castaldi et al., *The CMS tracker*
2294 *system project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997,
2295 <https://cds.cern.ch/record/368412>.
- 2296 [49] CMS Collaboration, S. Chatrchyan et al., *Description and performance of track and*
2297 *primary-vertex reconstruction with the CMS tracker*, JINST **9** (2014) P10009, [1405.6569].
- 2298 [50] CMS Collaboration, *The CMS electromagnetic calorimeter project: Technical Design Report*.
2299 Technical Design Report CMS. CERN, Geneva, 1997, <https://cds.cern.ch/record/349375>.
- 2300 [51] CMS Collaboration, *The CMS hadron calorimeter project: Technical Design Report*. Technical
2301 Design Report CMS. CERN, Geneva, 1997, <https://cds.cern.ch/record/357153>.
- 2302 [52] CMS Collaboration, *The CMS muon project: Technical Design Report*. Technical Design
2303 Report CMS. CERN, Geneva, 1997, <https://cds.cern.ch/record/343814>.
- 2304 [53] CMS Collaboration, *CMS The TriDAS Project: Technical Design Report, Volume 2: Data*
2305 *Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical
2306 Design Report CMS. CERN, Geneva, 2002, <https://cds.cern.ch/record/578006>.
- 2307 [54] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl.
2308 Instrum. Meth. **A389** (1997) 81–86.
- 2309 [55] CMS Collaboration, A. M. Sirunyan et al., *Particle-flow reconstruction and global event*
2310 *description with the cms detector*, JINST **12** (2017) P10003, [1706.04965].
- 2311 [56] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth.
2312 **A262** (1987) 444–450.
- 2313 [57] K. Rose, *Deterministic annealing for clustering, compression, classification, regression, and*
2314 *related optimization problems*, in *Proceedings of the IEEE*, vol. 86, pp. 2210 – 2239, 12, 1998.
- 2315 [58] R. Fruhwirth, W. Waltenberger and P. Vanlaer, *Adaptive vertex fitting*, J. Phys. **G34** (2007)
2316 N343.
- 2317 [59] CMS Collaboration, S. Chatrchyan et al., *Energy Calibration and Resolution of the CMS*
2318 *Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV*, JINST **8** (2013) P09009,
2319 [1306.2016].

BIBLIOGRAPHY

- [60] W. Adam, R. Fröhwirth, A. Strandlie and T. Todorov, *RESEARCH NOTE FROM COLLABORATION: Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC*, *Journal of Physics G Nuclear Physics* **31** (Sept., 2005) N9–N20, [physics/0306087].
- [61] CMS Collaboration, V. Khachatryan et al., *Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV*, *JINST* **10** (2015) P06005, [1502.02701].
- [62] CMS Collaboration, V. Khachatryan et al., *Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV*, *JINST* **10** (2015) P08010, [1502.02702].
- [63] CMS Collaboration, S. Chatrchyan et al., *Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV*, *JINST* **7** (2012) P10002, [1206.4071].
- [64] CMS Collaboration, *Performance of muon reconstruction including Alignment Position Errors for 2016 Collision Data*, CMS Performance Note CMS-DP-2016-067, CERN, Geneva, Nov, 2016. <http://cds.cern.ch/record/2229697>.
- [65] CMS Collaboration, *Pileup Removal Algorithms*, Tech. Rep. CMS-PAS-JME-14-001, CERN, Geneva, 2014. <https://cds.cern.ch/record/1751454>.
- [66] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896, [1111.6097].
- [67] M. Cacciari, G. P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, [0802.1189].
- [68] CMS Collaboration, S. Chatrchyan et al., *Determination of jet energy calibration and transverse momentum resolution in cms*, *Journal of Instrumentation* **6** (2011) P11002.
- [69] CMS Collaboration, *Jet energy scale and resolution performances with 13TeV data*, CMS Performance Note CMS-DP-2016-020, CERN, Geneva, Jun, 2016. <http://cds.cern.ch/record/2160347>.
- [70] CMS Collaboration, S. Chatrchyan et al., *Performance of tau-lepton reconstruction and identification in CMS*, *JINST* **7** (2012) P01001, [1109.6034].
- [71] CMS Collaboration, *Identification of b quark jets at the CMS Experiment in the LHC Run 2*, CMS Physics Analysis Summary CMS-PAS-BTV-15-001, CERN, Geneva, 2016. <https://cds.cern.ch/record/2138504>.
- [72] CMS Collaboration, S. Chatrchyan et al., *Identification of b-quark jets with the CMS experiment*, *JINST* **8** (2013) P04013, [1211.4462].
- [73] CMS Collaboration, *Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector*, CMS Physics Analysis Summary CMS-PAS-JME-16-004, CERN, Geneva, 2016. <https://cds.cern.ch/record/2205284>.
- [74] CMS Collaboration, S. Chatrchyan et al., *Missing transverse energy performance of the CMS detector*, *JINST* **6** (2011) P09001, [1106.5048].

- [75] CMS Collaboration, *Performance of Missing Transverse Momentum Reconstruction Algorithms in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV with the CMS Detector*, CMS Physics Analysis Summary CMS-PAS-JME-12-002, CERN, Geneva, 2012.
<https://cds.cern.ch/record/1543527>.
- [76] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [77] ALICE Collaboration, K. Aamodt et al., *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.
- [78] LHCb Collaboration, A. A. Alves, Jr. et al., *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005.
- [79] 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, [1405.0301].
- [80] Torbjorn Sjostrand and Stephen Mrenna and Peter Skands, *A brief introduction to pythia 8.1*, *Comput.Phys.Commun.* **178** (2008) 852–867, [0710.3820].
- [81] CMS Collaboration, V. Khachatryan et al., *Event generator tunes obtained from underlying event and multiparton scattering measurements*, *Eur. Phys. J. C* (2016) 155, [1512.00815].
- [82] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce et al., *Geant4 - a simulation toolkit*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003) 250 – 303.
- [83] CMS Collaboration, V. Khachatryan et al., *Identification techniques for highly boosted W bosons that decay into hadrons*, *JHEP* **12** (2014) 017, [1410.4227].
- [84] Rikkert Frederix and Stefano Frixione, *Merging meets matching in mc@nlo*, *JHEP* **12** (2012) 061, [1209.6215].
- [85] Emanuele Re, *Single-top wt-channel production matched with parton showers using the powheg method*, *Eur.Phys.J.* **C71** (2011) 1547, [1009.2450].
- [86] John M. Campbell, R. Keith Ellis, Paolo Nason, Emanuele Re, *Top-pair production and decay at nlo matched with parton showers*, *JHEP* **04** (2015) 114, [1412.1828].
- [87] S. Kallweit, J. M. Lindert, S. Pozzorini, M. Schönherr and P. Maierhöfer, *NLO QCD+EW automation and precise predictions for V+multijet production*, in *Proceedings, 50th Rencontres de Moriond, QCD and high energy interactions*, pp. 121–124, 2015, 1505.05704, <http://inspirehep.net/record/1372103/files/arXiv:1505.05704.pdf>.
- [88] CMS Luminosity Measurements for the 2016 Data Taking Period, CMS Physics Analysis Summary CMS-PAS-LUM-17-001, CERN, Geneva, 2017.
<https://cds.cern.ch/record/2257069>.
- [89] A. J. Larkoski, S. Marzani, G. Soyez and J. Thaler, *Soft Drop*, *JHEP* **05** (2014) 146, [1402.2657].
- [90] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001, [[hep-ph/9707323](#)].

BIBLIOGRAPHY

- [91] D. Bertolini, P. Harris, M. Low and N. Tran, *Pileup per particle identification*, *Journal of High Energy Physics* **2014** (2014) 59.
- [92] J. Thaler and K. Van Tilburg, *Identifying boosted objects with n-subjettiness*, *Journal of High Energy Physics* **2011** (2011) 15.
- [93] *Performance of b-Tagging Algorithms in Proton Collisions at 13 TeV using the 2016 Data*, Physics Analysis Summary CMS-DP-2016-042, CERN, Geneva, Jul, 2016.
<https://cds.cern.ch/record/2202967>.
- [94] CMS Collaboration, V. Khachatryan et al., *Search for a higgs boson in the mass range from 145 to 1000 GeV decaying to a pair of W or Z bosons*, *JHEP* **10** (2015) 144, [1504.00936].
- [95] M. Oreglia, *A Study of the Reactions $\psi' \rightarrow \gamma\gamma\psi$* , Ph.D. thesis, SLAC, 1980.
- [96] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*, Ph.D. thesis, Cracow, INP, 1986.
- [97] M. Bahr, S. Gieseke, M. A. Gigg, D. Grellscheid, K. Hamilton, O. Latunde-Dada, S. Platzer, P. Richardson, M. H. Seymour, A. Sherstnev, J. Tully, B. R. Webber, *Herwig++ physics and manual*, *Eur.Phys.J.* **C58** (2008) 639–707, [0803.0883].
- [98] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys.* **G43** (2016) 023001, [1510.03865].
- [99] NNPDF Collaboration, R. D. Ball et al., *Parton distributions from high-precision collider data*, *Eur. Phys. J.* **C77** (2017) 663, [1706.00428].
- [100] A. L. Read, *Presentation of search results: the CL_s technique*, *J. Phys.* **G28** (2002) 2693.
- [101] T. Junk, *Confidence level computation for combining searches with small statistics*, *Nucl. Instrum. Meth.* **A434** (1999) 435, [9902006].
- [102] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group, *Procedure for the LHC Higgs boson search combination in Summer 2011*, Tech. Rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug, 2011.
<https://cds.cern.ch/record/1379837>.
- [103] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *The European Physical Journal C* **71** (2011) .
- [104] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, *The Annals of Mathematical Statistics* **9** (03, 1938) 60–62.
- [105] A. Wald, *Tests of statistical hypotheses concerning several parameters when the number of observations is large*, *Transactions of the American Mathematical Society* **54** (1943) 426–482.
- [106] E. Gross and O. Vitells, *Trial factors for the look elsewhere effect in high energy physics*, *The European Physical Journal C* **70** (Nov, 2010) 525–530.
- [107] M. Cacciari, G. P. Salam and G. Soyez, *SoftKiller, a particle-level pileup removal method*, *Eur. Phys. J.* **C75** (2015) 59, [1407.0408].