

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

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Theoretical motivation</b>	<b>3</b>
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 . . . . .	8
2.2.2.3	Decays in fermions and bosons: concluding remarks . . . . .	9
2.2.3	HVT production . . . . .	9
2.2.4	Benchmark model A: weak coupling scenario . . . . .	11
2.2.5	Benchmark model B: strong coupling scenario . . . . .	12
2.2.6	Search for HVT resonances at LHC . . . . .	12
2.3	Warped extra dimension . . . . .	16
2.3.1	Randall-Sundrum original model (RS1) . . . . .	16
2.3.2	Bulk extension of RS1: graviton production and decays . . . . .	19
2.3.3	Search for KK bulk gravitons at LHC . . . . .	22
<b>3</b>	<b>The Large Hadron Collider and the CMS experiment</b>	<b>25</b>
3.1	The Large Hadron Collider . . . . .	25
3.1.1	Proton-proton interactions . . . . .	28
3.2	CMS detector . . . . .	29
3.2.1	The coordinate system . . . . .	30
3.2.2	The magnet . . . . .	31
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 . . . . .	36

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 . . . . .	38
3.2.9	Physics objects . . . . .	38
3.2.9.1	Track reconstruction . . . . .	38
3.2.9.2	Vertices reconstruction . . . . .	39
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 . . . . .	42
3.2.9.7	b-jets tagging . . . . .	42
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
<b>4</b>	<b>Search for diboson resonances in the <math>VZ \rightarrow q\bar{q} \nu\bar{\nu}</math> final state</b>	<b>51</b>
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 . . . . .	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 . . . . .	129
4.7.3	Interpretation of the results in the HVT model . . . . .	132
<b>5</b>	<b>Conclusions</b>	<b>133</b>

**CONTENTS**

---

---

## Abstract

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 theo-  
6 ries. The signals probed are spin-1  $W'$ , predicted by the Heavy Vector Triplet model, and spin-2 bulk  
7 gravitons, predicted by warped extra-dimension models. The scrutinized data are produced by LHC  
8 proton-proton collisions at a center-of-mass energy  $\sqrt{s} = 13$  TeV during the 2016 operations, and  
9 collected by the CMS experiment, corresponding to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . One of  
10 the boson should be a  $Z$ , and it is identified through its invisible decay into neutrinos ( $\nu\bar{\nu}$ ), 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 ( $q\bar{q}$ ). The decay products of heavy resonances are produced with  
13 large Lorentz boosts; as a consequence, the decay products of the bosons (quarks and neutrinos)  
14 are expected to be highly energetic and collimated. The couple of neutrinos, escaping undetected,  
15 is reconstructed as missing momentum in the transverse plane of the CMS detector. The couple  
16 of quarks is reconstructed as one large-cone jet, with high transverse momentum, recoiling against  
17 the couple of neutrinos. Grooming algorithm are adopted in order to improve the jet mass resolu-  
18 tion, 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 di  
45  $35.9 \text{ fb}^{-1}$ . Uno dei bosoni dev'essere una  $Z$ , che viene identificata dal suo decadimento invisibile in  
46 neutrini ( $\nu\bar{\nu}$ ), mentre l'altro bosone elettrodebole, sia una  $W$  che una  $Z$ , deve decadere nel canale  
47 adronico in una coppia di quark ( $q\bar{q}$ ). I prodotti di decadimento di risonanze pesanti sono prodotti  
48 con significativi boost di Lorentz; di conseguenza, i prodotti di decadimento dei bosoni (i quark e i  
49 neutrini) sono attesi avere elevate energie ed essere 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 in eccesso locale nei dati rispetto alle predizioni. In funzione della massa, lim-  
65 iti superiori sulla sezione d'urto per risonanze pesanti e strette di spin 1 e spin 2, moltiplicate per  
66 il rapporto di diramazione della risonanza che decade in  $Z$  e  $W$  per il segnale di spin 1, e in una  
67 coppia di bosoni  $Z$  per lo spin 2, sono fissati nell'intervallo  $0.9 - 63 \text{ fb}$  e nell'intervallo  $0.5 - 40 \text{ fb}$   
68 rispettivamente. Un'ipotesi di è esclusa fino ad una massa di 3.11 TeV, nello scenario A di riferi-  
69 mento dell'Heavy Vector Triplet, e fino a 3.41 TeV, nello scenario B di riferimento. Un'ipotesi di  
70 gravitone, dato il parametro di curvatura della dimensione addizionale  $\tilde{k} = 1.0$ , è esclusa fino ad  
71 una massa di 1.1 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 known matter.  
 79 Despite this successfull achievement, some questions are still left unanswered; in order to solve the  
 80 open problems, a plethora of new beyond standard model theories have been built.  
 81 Many of these theories predict the existence of larger symmetries in the universe, or new extra-  
 82 dimensions, that will result into the appearance of new heavy particles, expected to have a mass  
 83 around 1 TeV. The Large Hadron Collider (LHC) is the ideal tool to investigate this unknown phase-  
 84 space, given the fact that during the so-called LHC Run 2 era (started in 2015), the unprecedented  
 85 center-of-mass energy of 13 TeV has been reached in the proton-proton collisions.  
 86 The CMS experiment, located in the northern part of the LHC ring, is a multi purpose detector, suit-  
 87 able to study highly energetic new phenomena. Its intense magnetic field, its sharp segmentation,  
 88 its hermeticity and the interplay of many sophisticated reconstruction algorithms allow to measure  
 89 with a very high precision the trajectories, the momenta and the energy deposits left by energetic  
 90 particles.

91 This thesis presents a searche for signal of heavy resonances that decay into a pair of vector bosons.  
 92 The search is performed by used the 2016 data produced by proton-proton collisions of the LHC,  
 93 and collected by the CMS detector. One  $Z$  boson is identified through its invisible decay in neutrino-  
 94 nos, while the other vector boson is required to decay hadronically into a pair of quarks. Given the  
 95 fact that the searched resonances have masses around the TeV, their decay products are expected  
 96 to be produced with large Lorentz boosts. This leads to a non-trivial identification of the couple of  
 97 quarks or leptons, coming from the vector bosons decays. In fact, they are expected to lay very close  
 98 in angle. Dedicated algorithms and substructure techniques allow to distinguish a pair of quarks  
 99 originating from a vector boson from the background processes, initiated by the strong interaction.  
 100 The search is performed by scanning the distribution of the reconstructed mass of the resonance,  
 101 looking for a signal local excess in data with regards to the prediction. The background prediction  
 102 is performed with an hybrid data-simulation approach, by using the distribution of data in signal-  
 103 depleted control regions.

104 The thesis is organized as follows.  
105 In chapter 2, an overview of the theoretical motivations is presented. Two beyond standard model  
106 theories are particularly appealing: the Heavy Vector Triplet Model and the bulk graviton model.  
107 In chapter 3, the CMS detector is briefly described, along with the physics objects exploited for the  
108 purpose of this search.  
109 Chapter 4 is dedicated to the analysis: after a general introduction (sec. 4.1), the features of the  
110 data, signal and background samples used in the analysis are described in detail (sec. 4.2). Sec. 4.3  
111 is dedicated to the selections applied, in order to reach the best signal-to-noise efficiency and to  
112 properly build the resonance candidate. The very first data-simulation comparison is performed in  
113 sec. 4.4. The background estimation technique, the final data-predicted background comparison  
114 and the signal modelling are included in sec. 4.5. Systematic uncertainties are listed in sec. 4.6. The  
115 final results, the statistical analysis and the physics interpretation are shown in sec. 4.7. Chapter 5  
116 summarizes the conclusions.

# Theoretical motivation

120 The standard model (SM) of particles represents, so far, the best available description of the parti-  
 121 cles and their interactions. It is the summation of two gauge theories: the electroweak interaction,  
 122 that portrays the weak and electromagnetic interactions together, and the strong interaction, or  
 123 Quantum Chromodynamics (QCD). Particles, namely quarks and leptons, are described as spin 1/2  
 124 fermions, whilst interactions are represented by spin 1 bosons. The symmetry group of the standard  
 125 model is:

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

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

132 In the last decades, Standard Model has been accurately probed by many experimental facilities  
 133 (LEP, Tevatron, LHC), and the results lead to an impressive agreement between theoretical pre-  
 134 dictions and experiments. The discovery of the Higgs boson at the CERN Large Hadron Collider,  
 135 measured by both CMS and ATLAS collaborations [1–7], represents not only an extraordinary con-  
 136 firmation of the model, but also the latest biggest achievement in particle physics as a whole.

## 138 2.1 Beyond Standard Model theories

139 Even though the Standard Model is the most complete picture of the universe of the particles, many  
 140 questions are still left open. From a phenomenological point of view, some experimental observa-  
 141 tions are not included in the theory:

- 142 • in SM, neutrinos are massless (whilst experimentally their masses are confirmed to be non-  
 143 zero, i.e. by the neutrino oscillations);
- 144 • no candidates for dark matter are predicted;
- 145 • no one of the fields included in the SM can explain the cosmological inflation;

- 146     • SM can not justify the matter-antimatter asymmetry.

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

- 148     • *Flavour problem.*

149     The Standard Model has 18 free parameters: 9 fermionic masses; 3 angular parameters in  
150     Cabibbo-Kobayashi-Maskawa matrix, plus 1 phase parameter; electromagnetic coupling  $\alpha$ ;  
151     strong coupling  $\alpha_{\text{strong}}$ ; weak coupling  $\alpha_{\text{weak}}$ ;  $Z$  mass; the mass of the Higgs boson. Such a  
152     huge number of degrees of freedom marks the SM as weakly predictive in the flavour sector.

- 153     • *Unification.*

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

- 157     • *Hierarchy problem.*

158     From Quantum Field Theory, it is known that perturbative corrections to the mass of the scalar  
159     bosons included in the theory tend to make it increase towards the energy scale at which the  
160     considered theory still holds [8]. If the Standard Model is seen as a low-mass approximation of  
161     a more general theory valid up to the Planck mass scale (*i.e.*,  $\sim 1.2 \times 10^{19}$  GeV), a fine-tuning  
162     cancellation of the order of 1 over  $10^{34}$  is needed in order to protect the Higgs mass at the  
163     electroweak scale ( $\sim 100$  GeV). Such an astonishing correction is perceived as very unnatural.

164     Numerous Beyond Standard Model theories (BSM) have been proposed in order to overcome  
165     the limits of the Standard Model.

166     Grand Unified Theories (GUT) aim at extending the symmetry group of the SM (eq. 2) into largest  
167     candidates, such as  $S0(10)$ ,  $SU(5)$  and  $E(6)$ . At GUT scale, approximately at  $10^{16}$  GeV, non-gravitational  
168     interactions are expected to be ruled by only one coupling constant,  $\alpha_{GUT}$ .

169     Super Symmetry (SUSY) models state that every fermion (boson) of the Standard Model has a  
170     bosonic (fermionic) superpartner, with exactly the same quantum numbers, except the spin. If  
171     SUSY is not broken, each couple of partners and superpartners should have the same masses, hy-  
172     pothesis excluded by the non-observation of the s-electron. Super Symmetry represents a very ele-  
173     gant solution of the hierarchy problem of the Higgs boson mass, since the perturbative corrections  
174     brought by new SUSY particles exactly cancel out the divergences caused by SM particles correc-  
175     tions. A particular sub-class of SUSY models, Minimal Super Symmetric Standard Models, is char-  
176     acterized by the introduction of a new symmetry, the R-parity, that guarantees the proton stability  
177     and also the stability of the lightest SUSY particle, a possible good candidate for dark matter.

178  
179     Two other possible theoretical pictures are extensively described in sec. 2.2-2.3.

## 2.2 Heavy Vector Triplet

---

### 180 2.2 Heavy Vector Triplet

181 The heavy vector triplet model [9] provides a general framework aimed at studying new physics be-  
182 yond the standard model, that can manifest into the appearance of new resonances.  
183 The adopted approach is that of the simplified model, in which an effective Lagrangian is intro-  
184 duced, in order to describe the properties and interactions of new particles (in this case, a triplet of  
185 spin-1 bosons) by using a limited set of parameters, that can be easily linked to the physical observ-  
186 ables at the LHC experiments. These parameters can describe many physical motivated theories  
187 (such as sequential extensions of the SM [10, 11] or Composite Higgs [12, 13]).  
188 Since a simplified model is not a complete theory, its validity is restricted to the on-shell quanti-  
189 ties related to the production and decay mechanisms of the new resonances, that is how most of  
190 the LHC BSM searches are performed. Given these conditions, experimental results in the resonant  
191 region are sensitive to a limited number of the phenomenological Lagrangian parameters (or to a  
192 combination of those), whilst the remaining parameters tend to influence the tail of the distribu-  
193 tions.  
194 Limits on production cross-section times branching ratio ( $\sigma \mathcal{B}$ ), as a function of the invariant mass  
195 spectrum of the probed resonance, can be extracted from experimental data. Given that  $\sigma \mathcal{B}$  are  
196 functions of the simplified model parameters and of the parton luminosities, it is then possible to  
197 interpret the observed limits in the parameter space.

#### 198 2.2.1 Simplified Lagrangian

199 The heavy vector triplet framework assumes the existence of an additional vector triplet,  $V_\mu^a$ ,  $a =$   
200 1, 2, 3, in which two spin-1 particles are charged and one is neutral:

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

201  
202 The triplet interactions are described by a simplified Lagrangian, that is invariant under SM gauge  
203 and CP symmetry, and accidentally invariant under the custodial symmetry  $SU(2)_L \times SU(2)_R$ :

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

204  
205 In the first line of the formula 2.3,  $V$  mass and kinematic terms are included, described with the co-  
206 variant derivative  $D_\mu V_\nu^a = \partial_\mu V_\nu^a + g \epsilon^{abc} W_\mu^b V_\nu^c$ , where  $W_\mu^a$  are the fields of the weak interaction and  
207  $g$  is the weak gauge coupling.  $V_\mu^a$  are not mass eigenstates, since they mix with the electroweak fields  
208 after the spontaneous symmetry breaking, therefore  $m_V$  isn't the physical mass of the  $V$  bosons.  
209 The second line describes the interaction of the triplet with the Higgs field and the SM left-handed  
210 fermions;  $c_H$  describes the vertices with the physical Higgs and the three unphysical Goldstone  
211 bosons that, for the Goldstone equivalence theorem, are equivalent to the longitudinal polarization  
212 of  $W$  and  $Z$  bosons at high-energy; hence,  $c_H$  is related to the bosonic decays of the resonances.  $c_F$

213 is the analogous parameter describing the  $V$  interaction with fermions, that can be generalized as a  
 214 flavour dependent coefficient, once defined  $J_F^{\mu a} = \sum_f \bar{f}_L \gamma^\mu \tau^a f_L$ :  $c_F V_\mu^a J_F^{\mu a} = c_\ell V_\mu^a J_\ell^{\mu a} + c_q V_\mu^a J_q^{\mu a} +$   
 215  $c_3 V_\mu^a J_3^{\mu a}$ .

216 The last part of the equation includes terms that are relevant only in strongly coupled scenarios (see  
 217 sec. 2.2.2.2) through the  $V$ - $W$  mixing, but it does not include vertices of  $V$  with light SM fields, hence  
 218 it can be neglected while describing the majority of the LHC phenomenology, under the assump-  
 219 tions previously stated. Additional dimension four quadrilinear  $V$  interactions are non relevant for  
 220 the processes discussed, otherwise their effects would be appreciated in electroweak precision tests  
 221 and precise Higgs coupling measurements [14].

222 The parameters in the Lagrangian can be interpreted as follows:  $g_V$  describes the strength of  
 223 the interaction, that is weighted by  $c$  parameters.  $g_V$  ranges from  $g_V \sim 1$  when the coupling is  
 224 weak (sec. 2.2.4), to  $g_V \sim 4\pi$  when the coupling is strong (sec. 2.2.5).  $c$  parameters are expected  
 225 to be  $c \sim 1$ , except to  $c_H$ , that can be smaller for weak couplings. The combinations describing the  
 226 vertices,  $g_V c_H$  and  $g^2/g_V c_F$ , can be considered as the fundamental parameters, used to interpret  
 227 the experimental results.

## 229 2.2.2 Mass eigenstates, mixing parameters and decay widths

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

$$A_\mu = B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W, \quad (2.4)$$

233 with the usual definitions of the electroweak parameters:

$$\begin{aligned} \tan \theta_W &= \frac{g'}{g} \\ e &= \frac{gg'}{\sqrt{g^2 + g'^2}} \\ g &= e / \sin \theta_w \\ g' &= e / \cos \theta_w. \end{aligned} \quad (2.5)$$

234 The  $Z$  boson, on the other hand, mixes with the neutral component of the triplet,  $V^0$ , with a  
 235 rotation parametrized with the angle  $\theta_N$ :

$$\begin{pmatrix} \cos \theta_N & \sin \theta_N \\ -\sin \theta_N & \cos \theta_N d \end{pmatrix} \begin{pmatrix} Z \\ V^0 \end{pmatrix}. \quad (2.6)$$

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

$$\mathbb{M}_N^2 = \begin{pmatrix} \hat{m}_Z^2 & c_H \zeta \hat{m}_Z \hat{m}_V \\ c_H \zeta \hat{m}_Z \hat{m}_V & \hat{m}_V^2 \end{pmatrix}, \quad (2.7)$$

237 where the parameters are defined as:

$$\left\{ \begin{array}{l} \hat{m}_Z = \frac{e}{2 \sin \theta_W \cos \theta_W} \hat{v} \\ \hat{m}_V^2 = m_V^2 + g_V^2 c_{VVHH} \hat{v}^2 \\ \zeta = \frac{g_V \hat{v}}{2 \hat{m}_V} \\ \frac{\hat{v}^2}{2} = \langle H^\dagger H \rangle \end{array} \right., \quad (2.8)$$

## 2.2 Heavy Vector Triplet

---

- <sup>238</sup> and  $\hat{v}$ , the vacuum expectation value of the Higgs field, can be different from the SM  $v = 246$  GeV.  
<sup>239</sup> The physical masses of  $Z$  and  $V^0$ ,  $m_Z$  and  $M_0$ , and  $\theta_N$  come from the matrix relations:

$$\begin{aligned} \text{Tr}(\mathbb{M}_N^2) &= \hat{m}_Z^2 + \hat{m}_V^2 = m_Z^2 + M_0^2 \\ \|\mathbb{M}_N^2\| &= \hat{m}_Z^2 + \hat{m}_V^2 (1 - c_H^2 \zeta^2) = m_Z^2 M_0^2 \\ \tan 2\theta_N &= \frac{2c_H \zeta \hat{m}_Z \hat{m}_V}{\hat{m}_V^2 - \hat{m}_Z^2}. \end{aligned} \quad (2.9)$$

- <sup>240</sup> The  $W^\pm$  bosons mix with the charged components of the triplet,  $V^\pm$ , leading to a mass matrix  
<sup>241</sup> analogous to eq. 2.10:

$$\mathbb{M}_C^2 = \begin{pmatrix} \hat{m}_W^2 & c_H \zeta \hat{m}_W \hat{m}_V \\ c_H \zeta \hat{m}_W \hat{m}_V & \hat{m}_V^2 \end{pmatrix}, \quad (2.10)$$

- <sup>242</sup> where  $\hat{m}_W$  is defined as:

$$\left\{ \begin{array}{l} \hat{m}_W = \frac{e}{2 \sin \theta_W} \hat{v} = \hat{m}_Z \cos \theta_W \\ \end{array} \right. ; \quad (2.11)$$

- <sup>243</sup> the physical masses of  $W$  and  $V^\pm$ ,  $m_W$  and  $M_\pm$ , and the angle  $\theta_C$  parametrizing the rotation of the  
<sup>244</sup> charged sector are described by:

$$\begin{aligned} \text{Tr}(\mathbb{M}_C^2) &= \hat{m}_W^2 + \hat{m}_V^2 = m_W^2 + M_\pm^2 \\ \|\mathbb{M}_C^2\| &= \hat{m}_W^2 + \hat{m}_V^2 (1 - c_H^2 \zeta^2) = m_W^2 M_\pm^2 \\ \tan 2\theta_C &= \frac{2c_H \zeta \hat{m}_W \hat{m}_V}{\hat{m}_V^2 - \hat{m}_W^2}. \end{aligned} \quad (2.12)$$

- <sup>245</sup> The custodial symmetry of eq. 2.3 guarantees that:

$$\mathbb{M}_C^2 = \begin{pmatrix} \cos \theta_W & 0 \\ 0 & 1 \end{pmatrix} \mathbb{M}_N^2 \begin{pmatrix} \cos \theta_W & 0 \\ 0 & 1 \end{pmatrix}. \quad (2.13)$$

- <sup>246</sup>
- <sup>247</sup> By taking the determinant of these matrices, a custodial relation among the masses can be extracted:  
<sup>248</sup>

$$m_W^2 M_\pm^2 = \cos \theta_W m_Z^2 M_0^2, \quad (2.14)$$

- <sup>249</sup> that has some very important consequences.  
<sup>250</sup> Given that this model aims at searching new particles in the TeV scale and that the scale of the elec-  
<sup>251</sup> troweak interactions must lay at  $\sim 100$  GeV, a hierarchy of the physical masses seems very natural:

<sup>252</sup>

$$\frac{\hat{m}_{(W,Z)}}{\hat{m}_V} \sim \frac{m_{(W,Z)}}{M_{(\pm,0)}} \ll 1; \quad (2.15)$$

- <sup>253</sup>  $\zeta$  parameter can be  $\zeta \ll 1$  (weakly coupled scenario) or  $\zeta \sim 1$  (strongly coupled scenario). When  
<sup>254</sup> eq. 2.15 applies, the second lines in eq. 2.9 and eq. 2.12 can be approximated as follows:

$$\begin{aligned} m_Z^2 &= \hat{m}_Z^2 (1 - c_H^2 \zeta^2) (1 + \mathcal{O}(\hat{m}_Z^2 / \hat{m}_V^2)) \\ m_W^2 &= \hat{m}_W^2 (1 - c_H^2 \zeta^2) (1 + \mathcal{O}(\hat{m}_W^2 / \hat{m}_V^2)). \end{aligned} \quad (2.16)$$

- <sup>255</sup>
- <sup>256</sup> From eq. 2.11, the ratio of the physical masses of the charged and neutral electroweak bosons can  
<sup>257</sup> be approximated as:

$$\frac{m_W^2}{m_Z^2} \approx \cos \theta_W^2, \quad (2.17)$$

258 that satisfies the SM tree-level relation  $\rho = 1$  if  $\cos \theta_W^2 \approx 1. - 0.23$ . Adding this approximation into  
259 eq. 2.14, the  $V$  bosons are expected to have the same masses, hence the same production rates:

$$M_{\pm}^2 = M_0^2 (1 + \mathcal{O}(\%)). \quad (2.18)$$

260 The degenerate mass of the triplet will be called  $M_V \approx M_{\pm} \approx M_0$ ; given 2.15,  $M_V = \hat{m}_V$ .  
261 Another consequence of the mass hierarchy (2.15) is that the mixing angles  $\theta_{(N,C)}$  between the elec-  
262 troweak fields and the triplet are small:

$$\theta_{(N,C)} \approx c_H \zeta \frac{\hat{m}_{(W,Z)}}{\hat{m}_V} \ll 1, \quad (2.19)$$

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

#### 264 2.2.2.1 Decay widths into fermions

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

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

$$\begin{cases} g_L^C = \frac{g^2}{g_V} \frac{c_F}{2} \cos \theta_C + (g_L^W)_{SM} \sin \theta_N \approx \frac{g^2}{g_V} \frac{c_F}{2}, \\ g_R^C = 0 \end{cases}, \quad (2.20)$$

267 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  
268 left fermions, and the strength of the couplings with fermions is determined by  $g^2/g_V c_F$ , as stated  
269 in sec. 2.2.1. The decay width into fermions is then given by:

$$\Gamma_{V^{\pm} \rightarrow f \bar{f}} \approx 2\Gamma_{V^0 \rightarrow f \bar{f}} \approx N_c \left( \frac{g^2 c_F}{g_V} \right)^2 \frac{M_V}{48\pi}, \quad (2.21)$$

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

#### 271 2.2.2.2 Decay widths into bosons

272 As a starting point, a proper choice of the gauge makes the derivation of the approximate decay  
273 widths easier. While the unitary gauge is very convenient in discussing the electroweak symmetry  
274 breaking mechanism, since it provides a basis in which the Goldstone components of the scalar  
275 fields of the theory are set to zero, it does not properly describe the longitudinally polarized bosons  
276 in high-energy regimes, since it introduces a dependence of the type  $E/m$  in the longitudinal po-  
277 larization vector, not corresponding to the experimental results. This pathological behaviour can  
278 be overcome profiting of the equivalence theorem: while calculating the scattering amplitude of  
279 an high-energy process, the longitudinally polarized vectors are equivalent to their corresponding  
280 Goldstone scalars. The scattering amplitude can therefore be calculated with Goldstone diagrams.  
281 In the so-called equivalent gauge [15], the Higgs doublet is then parametrized as:

$$H = \begin{pmatrix} i\pi_+ \\ \frac{\hat{h} + h - i\pi_0}{\sqrt{2}} \end{pmatrix}, \quad (2.22)$$

## 2.2 Heavy Vector Triplet

---

and the Goldstones  $\pi_0$  and  $\pi_+$  describe respectively  $W$  and  $Z$  longitudinal bosons;  $h$  is the physical Higgs boson. Rewriting the simplified Lagrangian 2.3 with 2.22 parametrization, two terms hold the information of the interaction of the  $V$ s with the Goldstones:

$$\mathcal{L}_\pi = \dots + c_H \zeta \hat{m}_V V_\mu^a \partial^\mu \pi^a + \frac{g_V c_H}{2} V_\mu^a (\partial^\mu h \pi^a - h \partial^\mu \pi^a + \epsilon^{abc} \pi^b \partial^\mu \pi^c) + \dots, \quad (2.23)$$

that are ruled by the  $c_H g_V$  parameters combination. When  $\zeta$  parameter is  $\zeta \approx 1$ , the first term in eq. 2.23 becomes important, and it is absorbed by a redefinition of the  $V_\mu^a$  and  $\pi^a$  fields,

$$\begin{aligned} V_\mu^a &\rightarrow V_\mu^a + \frac{c_H \zeta}{\hat{m}_V} \partial_\mu \pi^a \\ \pi^a &\rightarrow \frac{1}{\sqrt{1 - c_H^2 \zeta^2}} \pi^a; \quad c_H^2 \zeta^2 < 1 \end{aligned} \quad (2.24)$$

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

$$\begin{aligned} \Gamma_{V^0 \rightarrow W_L^+ W_L^-} &\approx \Gamma_{V^\pm \rightarrow W_L^\pm Z_L} \approx \frac{g_V^2 c_H^2 M_V}{192\pi} \frac{(1 + c_H c_{VVV} \zeta^2)^2}{(1 - c_H^2 \zeta^2)^2} = \frac{g_V^2 c_H^2 M_V}{192\pi} (1 + \mathcal{O}(\zeta^2)) \\ \Gamma_{V^0 \rightarrow Z_L h} &\approx \Gamma_{V^\pm \rightarrow W_L^\pm h} \approx \frac{g_V^2 c_H^2 M_V}{192\pi} \frac{(1 - 4 c_H c_{VVV} \zeta^2)^2}{(1 - c_H^2 \zeta^2)^2} = \frac{g_V^2 c_H^2 M_V}{192\pi} (1 + \mathcal{O}(\zeta^2)). \end{aligned} \quad (2.25)$$

### 2.2.2.3 Decays in fermions and bosons: concluding remarks

From eq. 2.21-2.25, some important conclusions can be extracted.

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

### 2.2.3 HVT production

Given the mass scale of the resonances, the production mechanisms expected to be relevant are Drell-Yan (fig. 2.1) and Vector Boson Fusion (VBF) (fig. 2.2).

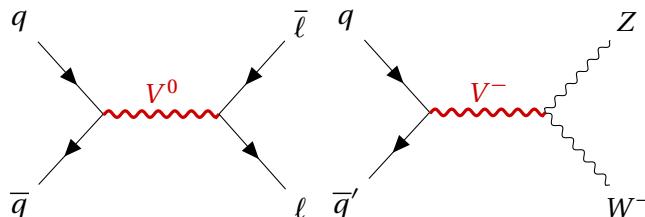


Figure 2.1: Examples of Drell-Yan production mechanism of a heavy  $V$  HVT boson:  $q - \bar{q}$  quark scattering producing a neutral  $V^0$  that decays leptonically (left);  $q - \bar{q}'$  scattering producing a charged  $V^-$  that decays in a  $W$  and  $Z$  bosons (right).

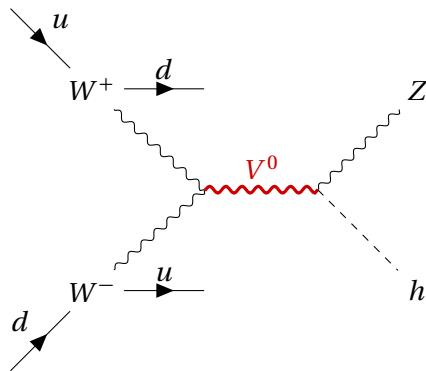


Figure 2.2: Example of VBF production mechanism of a heavy  $V$  HVT boson: a neutral  $V^0$  boson is produced by a couple of  $W$  bosons, as a result of electroweak interactions of initial state  $u$  and  $d$  quarks.  $V^0$  decays in a  $Z$  boson and a Higgs boson. The final state signature includes the presence of a pair of quarks, due to the primary interactions.

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

$$\sigma(pp \rightarrow V + X) = \sum_{i,j \in p} \frac{\Gamma_{V \rightarrow ij}}{M_V} f(J, S_i, S_j) g(C_i, C_j) \left. \frac{dL_{ij}}{ds} \right|_{s=M_V^2}, \quad (2.26)$$

299 where  $i, j$  are the partons involved in the hard interaction,  $\Gamma_{ij}$  is the partial width of the process  
 300  $V \rightarrow ij$ ,  $f(J, S_i, S_j)$  is a function of the spin of the resonance and of the partons,  $g(C_i, C_j)$  is a func-  
 301 tion of the colour factors of each parton,  $s$  is the center-of-mass energy and  $\frac{dL_{ij}}{ds}$  are the parton  
 302 luminosities, that are independent from HVT model (that enters only in  $\Gamma_{ij}$ ).  
 303 Parton luminosities, calculated for a center-of-mass energy of 14 TeV starting from quark and anti-  
 304 quark parton distribution functions (PDF), are displayed in fig. 2.3 (Drell-Yan mechanism) and 2.4  
 305 (VBF mechanism). VBF luminosities are suppressed by the  $\alpha_{EW}$  factor, therefore the process is rel-  
 evant only when the bosonic decays of the triplet are dominant (strongly coupled scenario).

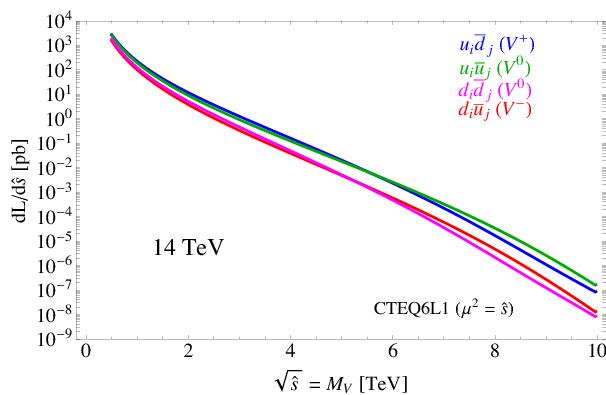


Figure 2.3: Parton luminosities for Drell-Yan process between  $i$  and  $j$  partons, as a function of the parton center-of-mass energy, for the LHC proton-proton collisions performed at 14 TeV.

## 2.2 Heavy Vector Triplet

---

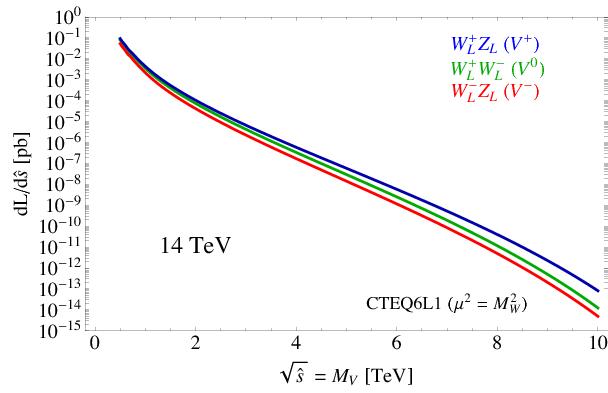


Figure 2.4: Parton luminosities for VBF process between  $i$  and  $j$  partons, as a function of the parton center-of-mass energy, for the LHC proton-proton collisions performed at 14 TeV.

### 2.2.4 Benchmark model A: weak coupling scenario

Model A scenario aims at reproducing a simple generalization of the SM [10], obtained by extending the gauge symmetry group with an additional  $SU(2)'$ . The low-energy phenomena are expected to be dominated by the SM, while the high-energy processes are relevant for the additional symmetry, bringing additional light vector bosons in play.

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 implies that:

$$\begin{aligned} g_V c_H &\approx g^2/g_V \\ g^2 c_F/g_V &\approx g^2/g_V, \end{aligned} \tag{2.27}$$

hence the partial decay widths into fermions (eq. 2.21) and bosons (eq. 2.25) differ only by a factor 2 and the colour factor ( $N_c$ ). Branching fractions for the model A benchmark scenario ( $g_V = 1$ ) are shown in fig. 2.5 (left); total widths are reported in fig. 2.5 (right) for different coupling parameters  $g_V$ .

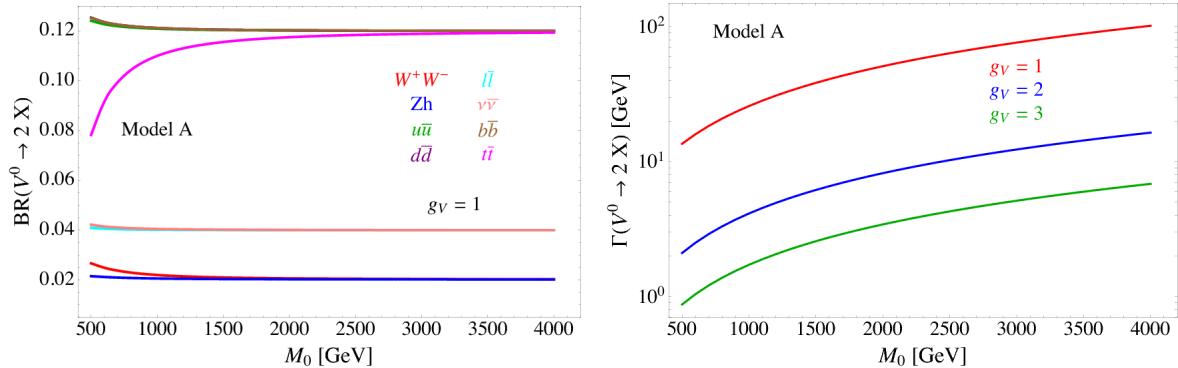


Figure 2.5: HVT model A scenario: branching fractions for fermionic and bosonic decays when  $g_V = 1$  (left) as a function of the mass of the resonance  $M_0$ ; total width of the resonance, as a function of its mass, considering different values of the parameter  $g_V$  (right).

### 318 2.2.5 Benchmark model B: strong coupling scenario

319 In composite Higgs models [12], the Higgs boson is the result of the spontaneous symmetry breaking  
 320 of an  $SO(5)$  symmetry to a  $SO(4)$  group. New vector bosons are expected to appear, and the lightest  
 321 ones can be represented by HVT model B when  $c_H \sim c_F \sim 1$ .

322 In this case:

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

323 hence the decay into bosons is not suppressed by  $g_V$  parameter. In the benchmark scenario  $g_V = 3$ ,  
 324 decays into dibosons are largely dominant, as it can be seen in fig. 2.6 (left); the total decay width  
 325 increases for larger  $g_V$  (fig. 2.6, right). When the resonances start to be very broad, *i.e.*  $\Gamma/M_V \gg$   
 326 0.1, the assumptions leading to the simplified model are no longer valid, hence higher order, non-  
 327 resonant effects must be taken into account.

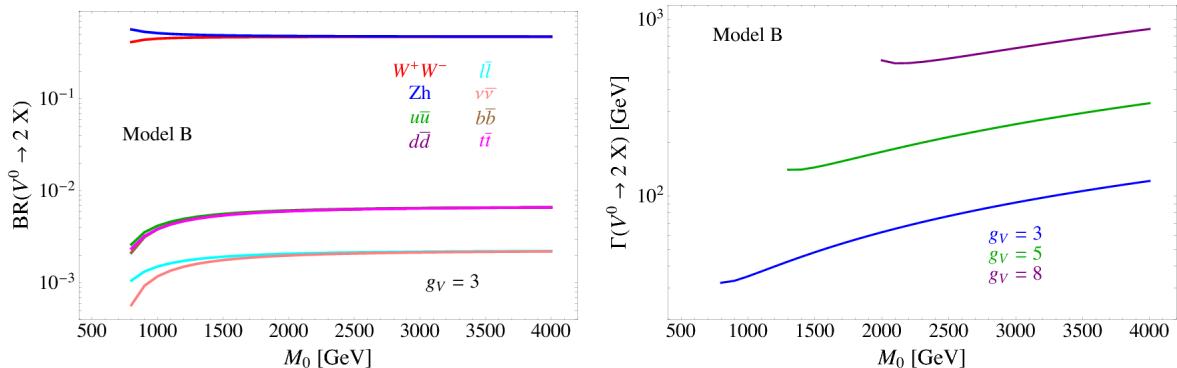


Figure 2.6: HVT model B scenario: branching fractions for fermionic and bosonic decays when  $g_V = 3$  (left) as a function of the mass of the resonance  $M_0$ ; total width of the resonance, as a function of its mass, considering different values of the parameter  $g_V$  (right).

### 328 2.2.6 Search for HVT resonances at LHC

329 No evidence of HVT resonances has been observed so far at LHC experiments. Data collected by  
 330 ATLAS and CMS detectors are used to set limits on the HVT resonance masses and coupling param-  
 331 eters. Experimental results from proton-proton collisions performed at a center-of-mass energy of  
 332 8 TeV (Run 1 era) at LHC brought to the following conclusions. A weakly coupled resonance, in the  
 333 context of benchmark model A ( $g_V = 1$ ) was excluded up to 3 TeV by Run 1 data. By looking at parton  
 334 luminosities in fig.2.3, in data produced by LHC proton-proton collision at 14 TeV, collected for an  
 335 integrated luminosity of  $300 \text{ fb}^{-1}$ , the sensitivity is expected to increase up to  $m_V \approx 6 \text{ TeV}$ . A strongly  
 336 coupled resonance, in the context of benchmark model B ( $g_V = 3$ ) is excluded up to 2 TeV by Run 1  
 337 data. Data produced by LHC at 14 TeV should increase the sensitivity up to  $m_V \approx 3-4 \text{ TeV}$ .  
 338 The most stringent limits are provided by the latest data produced by LHC at a center-of-mass en-  
 339 ergy of 13 TeV (Run 2 era).

340 Numerous searches for HVT triplet have been performed at CMS experiment in different final  
 341 states: the most sensitive ones were those in all-hadronic topology. [16, 17] (search for  $WW$ ,  $WZ$ ,  
 342  $ZZ$  resonances in the  $q\bar{q}q\bar{q}$  final state) excludes a  $W'$  with mass below 3.6 and a  $Z'$  with mass below

## 2.2 Heavy Vector Triplet

---

343 2.7 TeV in the model B scenario (fig. 2.7). [18, 19] (search for  $WH$ ,  $ZH$  resonances in the  $q\bar{q}b\bar{b}$  final  
 344 state) excludes a  $W'$  lighter than 2.97 (3.15) TeV in the HVT model A (model B), and a  $Z'$  up to 1.67  
 345 (2.26) TeV in HVT model A (model B) (fig. 2.8). In fig. 2.9, results of [16, 17] (left) and [18, 19] (right)  
 346 searches are interpreted as exclusion contours in the coupling parameter plane of the HVT model  
 347 ( $g_V c_H$  and  $g^2 c_F/g_V$ ). In the grey shaded area, the narrow width approximation fails. The colored  
 348 curves display the parameter exclusion for different mass hypotheses of the triplet. Colored dots  
 349 show the model A and B benchmark scenarios.

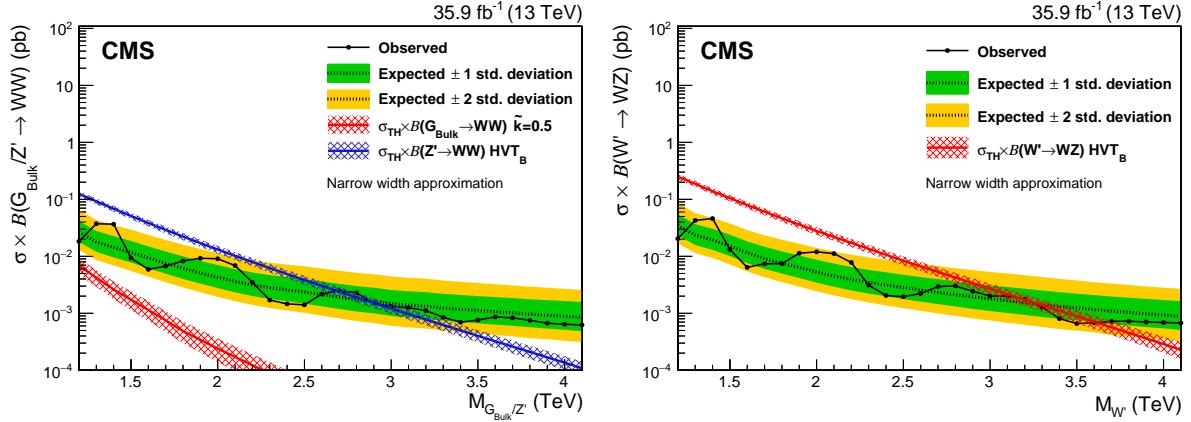


Figure 2.7: The observed and expected limits, with 68% and 95% uncertainty bands, on the product of the cross section and branching fraction  $\sigma \mathcal{B}(Z' \rightarrow WW)$  for a spin-1  $Z'$  (left) and  $\sigma \mathcal{B}(W' \rightarrow WZ)$  for a spin-1  $W'$  (right), as a function of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

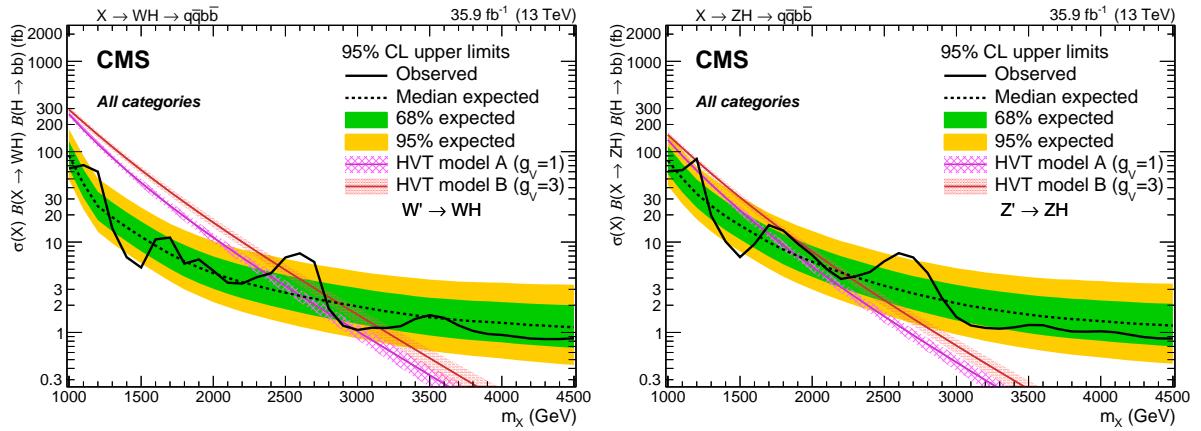


Figure 2.8: The observed and expected limits, with 68% and 95% uncertainty bands, on the product of the cross section and branching fraction  $\sigma \mathcal{B}(W' \rightarrow WH)$  for a spin-1  $W'$  (left) and  $\sigma \mathcal{B}(Z' \rightarrow ZH)$  for a spin-1  $Z'$  (right), as a function of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

350 Many other final states have been exploited at CMS:  $ZW, ZZ \rightarrow \ell\bar{\ell}q\bar{q}$  [20];  $WH, ZH \rightarrow (\ell\bar{\ell}, \ell\nu, \nu\bar{\nu})b\bar{b}$  [21];  
 351  $WZ, WW \rightarrow \ell\nu q\bar{q}$  [22]. Finally,  $ZW, ZZ \rightarrow \nu\bar{\nu}q\bar{q}$  [23] results will be extensively described in this  
 352 thesis.

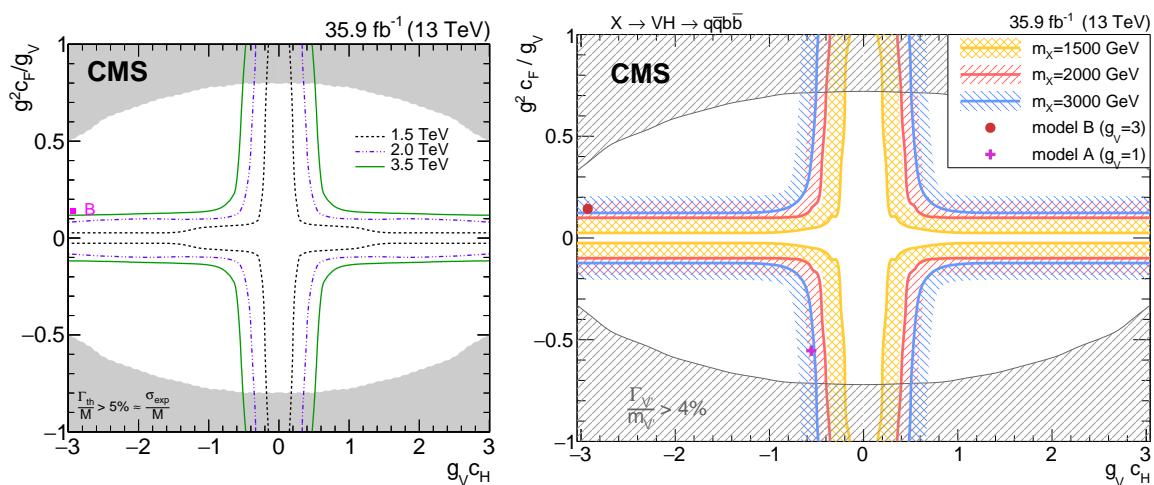


Figure 2.9: Exclusion contours in the coupling parameter plane of the HVT model ( $g_V c_H$  and  $g^2 c_F / g_V$ ).

## 2.2 Heavy Vector Triplet

---

353 Searches for HVT model B resonances have been performed at ATLAS experiment as well. Re-  
 354 sults for a  $W' \rightarrow WZ$  reported in fig. 2.10 include the searches performed in  $WW, WZ, ZZ \rightarrow q\bar{q}q\bar{q}$   
 355 final state [24];  $WZ, WW \rightarrow \ell\nu q\bar{q}$  final state [25];  $ZW, ZZ \rightarrow (\ell\bar{\ell}, \ell\nu, \nu\bar{\nu})q\bar{q}$  final state [26]. The all-  
 356 hadronic final state has the best sensitivity and it excludes a  $W'$  resonance up to 3.3 TeV (model B  
 357 scenario). Results for a  $W' \rightarrow WH$  and for a  $Z' \rightarrow ZH$  are displayed in fig. 2.11 (left and right respec-  
 358 tively), and they include searches performed in  $WH, ZH \rightarrow q\bar{q}b\bar{b}$  final state [27], and  $WH, ZH \rightarrow$   
 359  $\ell\bar{\ell}, \ell\nu, \nu\bar{\nu} b\bar{b}$  [28]. A  $W'$  is excluded up to 2.9 TeV and a  $Z'$  is excluded up to 2.8 TeV (in the model B  
 360 scenario).

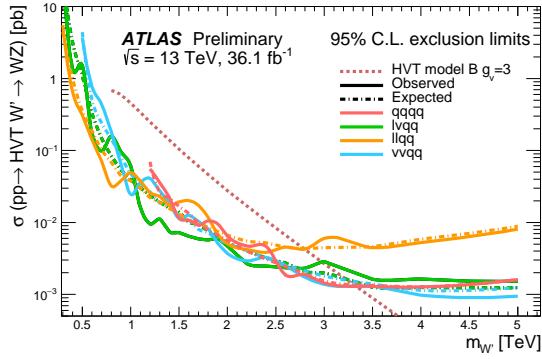


Figure 2.10: The observed and expected limits on the product of the cross section and branching fraction  $\sigma \mathcal{B}(W' \rightarrow WZ)$  for a spin-1  $W'$ , as a function of the reconstructed mass of the diboson resonance. The dotted line shows the theoretical predictions for the HVT model B.

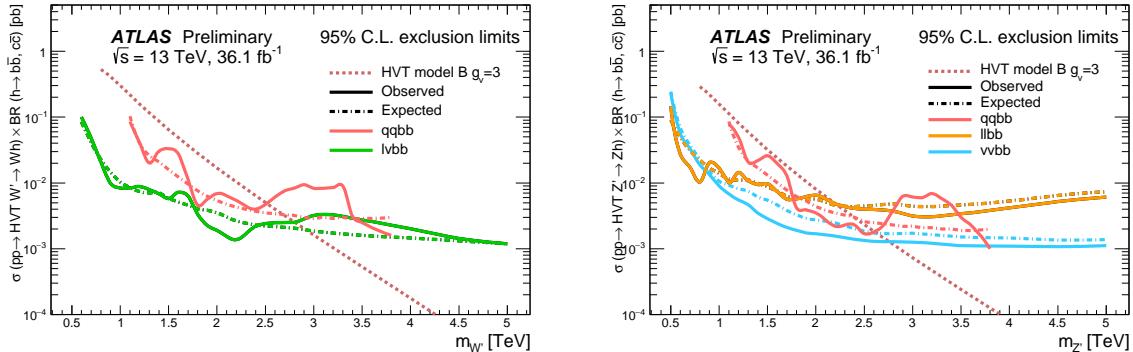


Figure 2.11: The observed and expected limits on the product of the cross section and branching fraction  $\sigma \mathcal{B}(W' \rightarrow WH)$  for a spin-1  $W'$  (left) and  $\sigma \mathcal{B}(Z' \rightarrow ZH)$  for a spin-1  $Z'$  (right), as a function of the reconstructed mass of the diboson resonance. The colored lines show the theoretical predictions for the HVT model B.

<sup>361</sup> **2.3 Warped extra dimension**

<sup>362</sup> The Randall-Sundrum model [29, 30] (RS1) proposes the introduction of one additional warped di-  
 363 mension in order to solve the hierarchy problem. The metric of the 5-dimensional space (a slice of  
 364  $AdS_5$ ) generates an exponential hierarchy between the electroweak and Planck scales, associated re-  
 365 spectively to the TeV three-brane, where the SM particles are confined, and the Planck three-brane.  
 366 As a consequence of the new geometry, spin-2 massive gravitons are predicted to exist.  
 367 The bulk extension of the Randall-Sundrum model [31, 32] states that the SM fields can propagate in  
 368 the extra dimension. Light fermions are near the Planck brane, heavy fermions are close to the TeV  
 369 brane, while the Higgs sector is confined in the TeV brane. Higgs couplings to the heavy fermions  
 370 are therefore expected to be stronger: this naturally arising hierarchy of the masses of the SM fields  
 371 gives a solution to the flavour problem. In this scenario, the fermionic decays of the bulk gravitons  
 372 are suppressed, while the bosonic decays are preferred.

<sup>373</sup> **2.3.1 Randall-Sundrum original model (RS1)**

<sup>374</sup> The existence of additional  $n$ -dimensions implies that the effective Planck scale observed in 4-  
 375 dimensions,  $M_{PL} = 1.220910^{19}$  GeV, is related to the fundamental  $4+n$ -dimensional Planck scale,  
 376  $M$ , via the geometry.  $M$  is expected to be of the order of the reduced  $\overline{M}_{PL} = M_{PL}/2\pi$ . If the 4-  
 377 dimensional and the  $n$  additional metrics are factorizable,  $\overline{M}_{PL}$  is the product of  $M$  and the volume  
 378 of the compact space  $V_n$ :

$$\overline{M}_{PL}^2 = V_n M^{2+n}. \quad (2.29)$$

<sup>379</sup> If  $M \sim$  TeV, this implies that  $V_n$  must be very large, hence the compactification scale  $\mu \sim 1/V_n^{1/n}$   
 380 is small (eV – MeV for  $n=2 - 7$ ). Given the smallness of  $\mu$  when compared to the electroweak scale,  
 381 the effects of the extra dimensions should be evident in SM processes. Since they are not observed,  
 382 SM particles are assumed to be confined in a 4-dimensional space, the TeV three-brane, while only  
 383 gravity is allowed to propagate into the  $4+n$ -dimensional space, the bulk. This mechanism solves  
 384 the hierarchy of the Higgs scale but introduces a new hierarchy between  $\mu$  and  $M$ .  
 385 In the Randall-Sundrum model [29, 30], only one additional dimension is added. The geometry of  
 386 the 5-dimensional bulk is non-factorizable, and it is a slice of  $AdS_5$  spacetime.<sup>1</sup> The 4-dimensional  
 387 metric is multiplied by an exponential function of the fifth dimension (the "warp" factor):

$$ds^2 = e^{-2kr_c\phi} \eta_{\mu\nu} dx^\mu dx^\nu + r_c^2 d\phi^2; \quad (2.30)$$

<sup>388</sup>  $x^\mu$  are the usual 4-dimensional coordinates,  $\eta_{\mu\nu} = diag(-1, 1, 1, 1)$  is the Minkowski metric,  $k$  is a  
 389 scale of order of  $\overline{M}_{PL}$ ,  $\phi$  is the coordinate of the extra dimension,  $0 < |\phi| < \pi$ , and  $r_c$  is the com-  
 390 pactification radius of this finite interval. 4-dimensional mass scales are obtained by multiplying  
 391 the bulk masses by  $e^{-2kr_c\phi}$ : given the exponential form of the warp factor, a small  $r_c$  suffices for  
 392 generating a large hierarchy between Planck and Higgs scales.  
 393 Two 4-dimensional three-branes are located at the boundaries of the fifth dimension: the visible  
 394 brane at  $\phi = \pi$ ; the hidden brane at  $\phi = 0$ , and their metrics are obtained starting from the bulk  
 395 5-dimensional metric  $G_{MN}$ , where  $M, N = \mu, \phi$ :

$$\begin{aligned} g_{\mu\nu}^{\text{vis}}(x^\mu) &= G_{\mu\nu}(x^\mu, \phi = \pi) \\ g_{\mu\nu}^{\text{hid}}(x^\mu) &= G_{\mu\nu}(x^\mu, \phi = 0). \end{aligned} \quad (2.31)$$

---

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

## 2.3 Warped extra dimension

---

396 The classical action is given by:

$$\begin{aligned} S &= S_{\text{gravity}} + S_{\text{vis}} + S_{\text{hid}} \\ S_{\text{gravity}} &= \int d^4x \int_{-\pi}^{+\pi} d\phi \sqrt{-G} (-\Lambda + 2M^3 \mathcal{R}) \\ S_{\text{vis}} &= \int d^4x \sqrt{-g_{\text{vis}}} (\mathcal{L}_{\text{vis}} - V_{\text{vis}}) \\ S_{\text{hid}} &= \int d^4x \sqrt{-g_{\text{hid}}} (\mathcal{L}_{\text{hid}} - V_{\text{hid}}), \end{aligned} \quad (2.32)$$

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

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

$$ds^2 = e^{-2\sigma(\phi)} \eta_{\mu\nu} dx^\mu dx^\nu + r_c^2 d\phi^2. \quad (2.33)$$

401 The Poincaré invariance guarantees that  $r_c$  does not depend on  $x^\mu$ . Given 2.33, the solution of the  
402 5-dimensional Einstein's equations simplifies into:

$$\sigma = r_c |\phi| \sqrt{\frac{-\Lambda}{24M^3}}. \quad (2.34)$$

403 Furthermore, the Poincaré invariance imposes constraints to the vacuum energies and cosmological  
404 constant:

$$\begin{aligned} V_{\text{hid}} &= -V_{\text{vis}} = 24M^3 k \\ \Lambda &= -24M^3 k^2. \end{aligned} \quad (2.35)$$

405 The final 5-dimensional metric is then:

$$ds^2 = e^{-2kr_c|\phi|} \eta_{\mu\nu} dx^\mu dx^\nu + r_c^2 d\phi^2. \quad (2.36)$$

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

410 The massless gravitons, the mediators of the gravitational interaction in the effective field theory, are  
411 the zero modes ( $h_{\mu\nu}$ ) of the quantum fluctuations of the classical solution (2.36):

$$ds^2 = e^{-2kT(x)|\phi|} (\eta_{\mu\nu} + h_{\mu\nu}(x)) dx^\mu dx^\nu + T^2(x) d\phi^2, \quad (2.37)$$

412 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 fluctuations around the Minkowski space, and represent both the physical graviton in 4-dimensions  
413 and the massless mode of the Kaluza-Klein decomposition of the bulk metric.  $r_c$  is the vacuum ex-  
414 pectation value of  $T(x)$ .

415 By substituting eq. 2.37 in the classical action 2.32, an effective action can be extracted, and in par-  
417 ticular the curvature term holds:

$$S_{\text{eff}} \sim \int d^4x \int_{-\pi}^{+\pi} d\phi 2M^3 r_c e^{-2kr_c|\phi|} \bar{\mathcal{R}} \sqrt{-\bar{g}}, \quad (2.38)$$

418 where  $\bar{g}$  is the trace of  $\bar{g}_{\mu\nu}$  and  $\bar{\mathcal{R}}$  is the 4-dimensional Ricci scalar of  $\bar{g}_{\mu\nu}$  metric. In this effective  
 419 4-dimensional action, the  $\phi$  dependence can be integrated out, and the 4-dimensional Planck mass  
 420 can be calculated:

$$\bar{M}_{PL}^2 = M^3 r_c \int_{-\pi}^{+\pi} d\phi e^{-2kr_c|\phi|} = \frac{M^3}{k} (1 - e^{-2kr_c\pi}). \quad (2.39)$$

421 It can be shown [29] that a field with a fundamental mass parameter  $m_0$  in the bulk manifests in the  
 422 visible three-brane with a physical mass  $m$ :

$$m = e^{-2kr_c\pi} m_0. \quad (2.40)$$

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

$$v = e^{-2kr_c\pi} v_0, \quad (2.41)$$

425 where  $v$  is the Higgs vacuum expectation value in the TeV brane and  $v_0$  is the 5-dimensional Higgs  
 426 v.e.v.

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

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

$$h_{\mu\nu}(x, \phi) = \sum_{n=0}^{\infty} h_{\mu\nu}^{(n)}(x) \frac{\chi^{(n)}(\phi)}{\sqrt{r_c}}. \quad (2.42)$$

431 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  
 432 the Klein-Gordon relation, where  $m_n^G \geq 0$ :

$$(\eta^{\mu\nu} \partial_\mu \partial_\nu - (m_n^G)^2) h_{\mu\nu}^{(n)}(x) = 0. \quad (2.43)$$

433 By substituting eq. 2.42 into Einstein's equation, the solutions for  $\chi^{(n)}(\phi)$  (commonly called "pro-  
 434 files") are [33, 34]:

$$\chi^{(n)}(\phi) = \frac{e^{2\sigma}}{N} [J_2(z_n^G) + \alpha_n Y_2(z_n^G)], \quad (2.44)$$

435 where  $J_2$  and  $Y_2$  are second order Bessel functions,  $N$  is the normalization of the wavefunction,  $\alpha_n$   
 436 are coefficients and  $z_n^G = m_n^G e^{\sigma(\phi)}/k$ .  $m_n^G$  is the mass of the  $n$ -mode, and it depends on the roots  
 437 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  $e^{kr_c\pi} \gg 1$ :

$$m_n^G = k z_n^G(\pi) e^{-kr_c\pi}. \quad (2.45)$$

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

$$\mathcal{L} = -\frac{1}{\bar{M}_{PL}} T^{\mu\nu}(x) h_{\mu\nu}^{(0)} - \frac{1}{e^{-kr_c\pi} \bar{M}_{PL}} T^{\mu\nu}(x) \sum_{n=1}^{\infty} h_{\mu\nu}^{(n)}(x); \quad (2.46)$$

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

## 2.3 Warped extra dimension

---

### 447 2.3.2 Bulk extension of RS1: graviton production and decays

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

463

464 The KK decomposition and the KK mass spectrum of the graviton have already been presented  
 465 in sec. 2.3.1. The KK decomposition of a massless 5-dimensional gauge field  $A_M(x, \phi)$  is similarly  
 466 performed [35]:

$$A_\mu(x, \phi) = \sum_{n=0}^{\infty} A_\mu^{(n)}(x) \frac{\chi^{(n)_A}(\phi)}{\sqrt{r_c}}. \quad (2.47)$$

467 The profiles for the gauge fields are:

$$\chi_A^{(n)}(\phi) = \frac{e^\sigma}{N_A} [J_1(z_n^A) + \alpha_n^A Y_1(z_n^A)], \quad (2.48)$$

468 where  $J_1$  and  $Y_1$  are first order Bessel functions. Similarly to eq. 2.49, the mass spectrum of the gauge  
 469 field is:

$$m_n^A = k z_n^A(\pi) e^{-k r_c \pi}; \quad (2.49)$$

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

471 The Lagrangian expressing the interaction between the  $m$  and  $n$  modes of the bulk field  $F$  to  
 472 the  $q$  KK gravitons mode  $G$  is [35]:

$$\mathcal{L}_{G-F} = \sum_{m,n,q} C_{mnq}^{FFG} \frac{1}{M_{PL}} \eta^{\mu\alpha} \eta^{\nu\beta} h_{\alpha\beta}^{(q)}(x) T_{\mu\nu}^{(m,n)}(x), \quad (2.50)$$

473  $C_{mnq}^{FFG}$  is the overlap integral of the profiles:

$$C_{mnq}^{FFG} = \int \frac{d\phi}{\sqrt{k}} e^{t\sigma} \frac{\chi_F^{(m)} \chi_F^{(n)} \chi_G^{(q)}}{\sqrt{r_c}}; \quad (2.51)$$

474  $t$  depends on the type of field considered.

475 The coupling between gluons and the  $q$  KK graviton mode is given by:

$$C_{00q}^{AAG} = e^{k\pi r_c} \frac{2[1 - J_0(x_n^G)]}{k\pi r_c (x_n^G)^2 |J_2(x_n^G)|}. \quad (2.52)$$

476 Once eq. 2.52 is put in eq. 2.50, the most significant partial decay widths into the  $q$  KK graviton mode  
 477 are:

$$\begin{aligned}\Gamma(G \rightarrow t_R \bar{t}_R) &\sim N_c \frac{\left[\tilde{k} x_q^G\right]^2 m_q^G}{320\pi} \\ \Gamma(G \rightarrow hh) &\sim \frac{\left[\tilde{k} x_q^G\right]^2 m_q^G}{960\pi} \\ \Gamma(G \rightarrow W_L^+ W_L^-) &\sim \frac{\left[\tilde{k} x_q^G\right]^2 m_q^G}{480\pi} \\ \Gamma(G \rightarrow Z_L Z_L) &\sim \frac{\left[\tilde{k} x_q^G\right]^2 m_q^G}{960\pi},\end{aligned}\tag{2.53}$$

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

$$\Gamma_G = \frac{13 \left[\tilde{k} x_q^G\right]^2 m_q^G}{960\pi}.\tag{2.54}$$

479 Calculations, so far, have been performed considering  $M \sim \overline{M}_{PL}$  and  $k < M$ , hypotheses under  
 480 which the solution for the bulk metric (eq. 2.36) is valid. Hence,  $\tilde{k} = k/\overline{M}_{PL} \leq 1$  is taken as a ref-  
 481 erence interval. This has also phenomenological consequences on the width of the resonance, as  
 482 stated in eq. 2.54. The total decay width of the lightest KK graviton mode, compared to its mass,  
 483 is shown as a function of  $\tilde{k}$  in fig. 2.12 [36]. At  $\tilde{k} = 1$ , in the bulk scenario, the KK graviton width is  
 484 expected to be few % of its mass, up to 4 TeV (dotted red curve). The narrow width approximation  
 485 holds, hence the resonance properties can be probed at the peak, neglecting the effects in the tails  
 486 of the mass distribution.

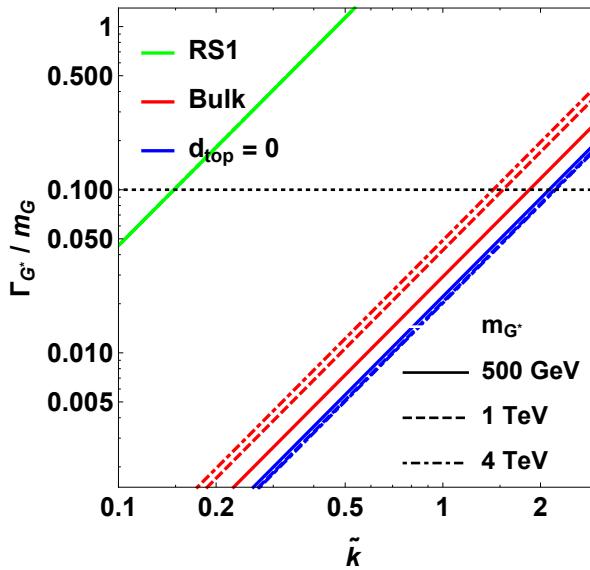


Figure 2.12: Width of the KK gravitons, in units of the mass of the resonance, as a function of the curvature parameter  $\tilde{k}$ . The red curves represent the bulk extension of RS1 original model for different mass hypotheses (from 500 GeV up to 4 TeV).

487 The total cross-section of a bulk graviton, produced at LHC in proton-proton interactions via  
 488 gluon fusion (displayed in fig. 2.13), decaying into a couple of vector bosons (for the purpose of this

## 2.3 Warped extra dimension

---

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

$$\sigma(pp \rightarrow ZZ) = \int dx_1 dx_2 f_g(x_1, Q^2) f_g(x_2, Q^2) \hat{\sigma}(x_1 x_2 s). \quad (2.55)$$

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

$$\frac{d\hat{\sigma}(gg \rightarrow ZZ)}{d \cos \hat{\theta}} \approx \frac{|\mathcal{M}_{+00}|^2}{1024\pi \hat{s}}, \quad (2.56)$$

494 where  $|\mathcal{M}_{+00}|$  is the matrix element of the dominant contribution in  $gg \rightarrow VV$  process ( $\Gamma_G$  is de-  
 495 fined in eq. 2.54,  $a, b$  are color factors):

$$\mathcal{M}_{+00}(g^a g^b \rightarrow VV) = -C_{00q}^{AAG} e^{-k\pi r_c} \left( \frac{x_n^G \tilde{k}}{m_n^G} \right)^2 \sum_n \frac{\delta_{ab} \mathcal{A}_{+00}}{\hat{s} - m_n^G + i\Gamma_G m_n^G}. \quad (2.57)$$

496 The relevant amplitudes taken account in the matrix element calculation are [31]:

$$\mathcal{A}_{+00} = \mathcal{A}_{-00} = \frac{(1 - 1/\beta_Z^2)(\beta_Z^2 - 2)[(\hat{t} - \hat{u})^2 - \beta_Z^2 \hat{s}^2] \hat{s}}{8M_Z^2}, \quad (2.58)$$

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

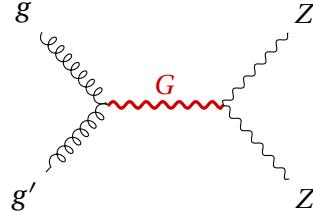


Figure 2.13: Gluon fusion production mechanism for a KK graviton that decays in a couple of  $Z$  bosons.

---

**498 2.3.3 Search for KK bulk gravitons at LHC**

499 No evidence of spin-2 bulk graviton resonances has been observed so far at LHC experiments. Data  
500 collected by ATLAS and CMS detectors are used to set limits on the graviton masses, generally con-  
501 sidering different curvature parameter  $\tilde{k}$  hypotheses, once assured the narrow width approximation  
502 is still valid (up to  $\tilde{k} \sim 1$ ). The most stringent limits have been set with Run 2 data.

503 Many results of the diboson searches performed at CMS and already presented in sec. 2.2.6 are  
504 interpreted in the context of the bulk gravitons, together with the additional final states  $WZ, ZZ \rightarrow$   
505  $\ell\bar{\ell}\nu\bar{\nu}$  [37] and  $HH \rightarrow b\bar{b}b\bar{b}$  [38]. The most interesting limit is provided by [37], that, under the  
506 hypothesis  $\tilde{k} = 0.5$ , excludes a spin-2 bulk graviton with a mass lower than 800 GeV (fig. 2.14).

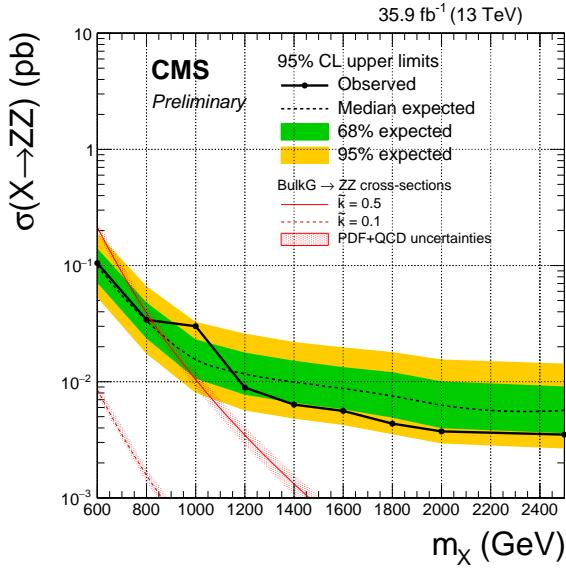


Figure 2.14: The observed and expected limits, with 68% and 95% uncertainty bands, on the product of the cross section and branching fraction  $\sigma \mathcal{B}(G \rightarrow ZZ)$  for a spin-2 bulk graviton, as a function of the reconstructed mass of the diboson resonance. The colored lines show the theoretical predictions for  $\tilde{k} = 0.1$  and  $0.5$ .

507 Similarly for ATLAS experiment, searches for diboson resonances in sec. 2.2.6 have been inter-  
508 preted in the graviton context. The most stringent limit is given by [25], where, under the assump-  
509 tion  $\tilde{k} = 1$ , a spin-2 bulk graviton with mass lower than 1.76 TeV is excluded (fig. 2.15).

## 2.3 Warped extra dimension

---

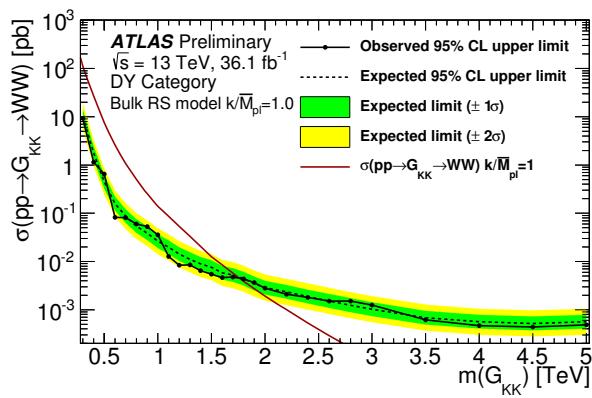


Figure 2.15: The observed and expected limits, with 68% and 95% uncertainty bands, on the product of the cross section and branching fraction  $\sigma \mathcal{B}(G \rightarrow ZZ)$  for a spin-2 bulk graviton, as a function of the reconstructed mass of the diboson resonance. The colored lines show the theoretical predictions for  $\tilde{k} = 1$ .



# The Large Hadron Collider and the CMS experiment

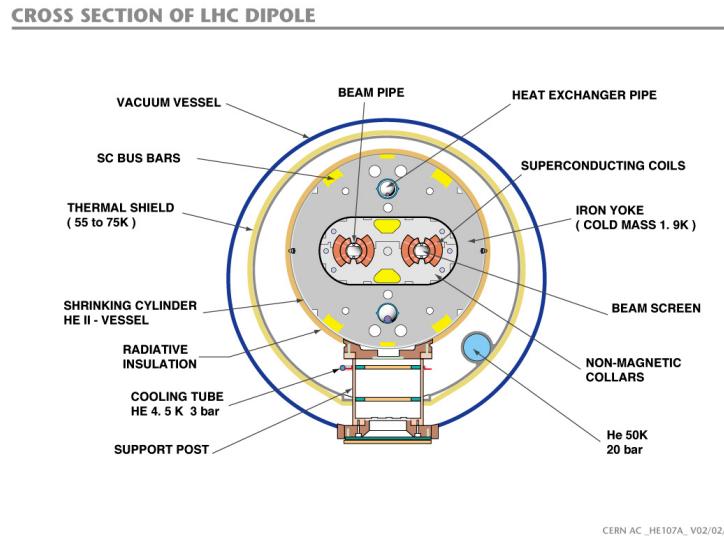
## 513 3.1 The Large Hadron Collider

514 The Large Hadron Collider (LHC) [39] is a 27 km ring structure designed for the acceleration and  
 515 collision of protons and heavy ions. It is situated approximatively 100 m underground, between  
 516 France and Switzerland, in the Geneva area, and it is part of the CERN research facilities. In order to  
 517 reduce the cost of the project, approved in 1996, the LHC has been designed to fit the pre-existing  
 518 underground tunnel of the Large Electron-Positron collider (LEP), built to accelerate electrons and  
 519 positrons and running until the year 2000.

520 Moving from an electron-positron collider to an hadron collider allowed to reach higher energies  
 521 in the center-of-mass frame, since the synchrotron radiation loss is inversely proportional to the  
 522 fourth power of the mass of the particle involved: hence, the radiation is reduced by a factor  $m_p/m_e \sim$   
 523  $10^3$ . The choice of a proton-proton collider was driven also by the possibility to collect higher lumi-  
 524 nosities (and hence more statistics) with regards to, for example, a proton-antiproton collider, like  
 525 Tevatron at Fermilab, in the USA.

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

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



CERN AC \_HE107A\_ V02/02/98

Figure 3.1: Section of the LHC dipole magnet structure.

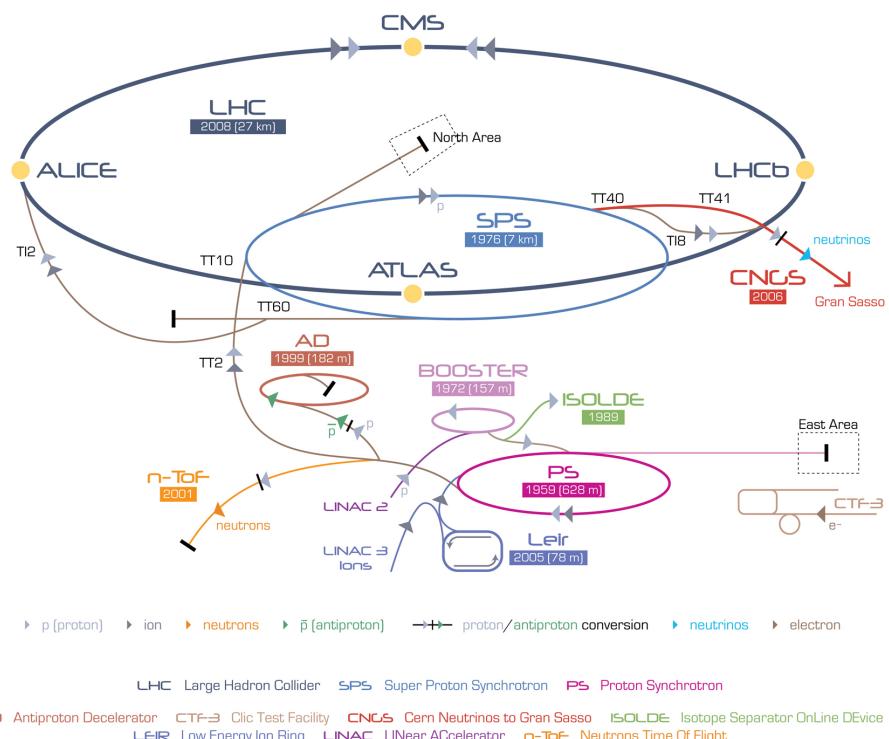


Figure 3.2: The CERN accelerator complex.

541 The LHC represents the final step of the CERN accelerator complex, shown in fig. 3.2. Protons are  
 542 extracted from hydrogen atoms and inserted in the linear accelerator Linac2, that brings them to an  
 543 energy of 50 MeV. They circulate around a little synchrotron, Proton Synchrotron Booster, reaching  
 544 an energy of 1.4 GeV, and then in the Proton Synchrotron (PS), where their energy is increased to 25  
 545 GeV. The second to last step is the Super Proton Synchrotron, SPS, accelerating protons up to 450  
 546 GeV. They are finally injected in the Large Hadron Collider, where sixteen radiofrequency cavities  
 547 (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 ( $\sim 10$  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  $\mu\text{m}$ . Different bunches are separated by 25 ns (or,  $\sim 7.5$  m), corresponding to a 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.

561

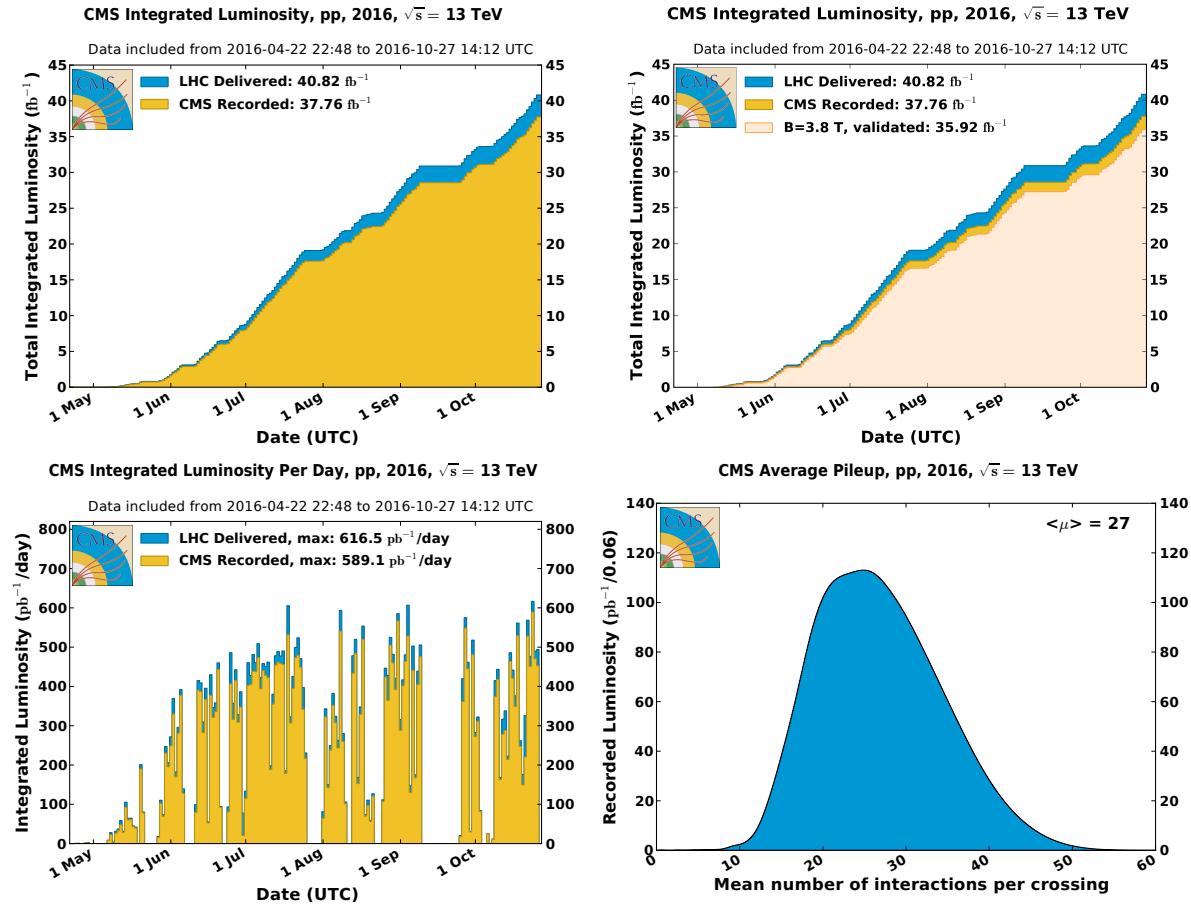


Figure 3.3: Luminosity in 2016 LHC data. Top-left plot: the cumulative integrated luminosity delivered by LHC (in blue) and recorded by CMS (in orange), as a function of the data taken period. Top-right plot: data recorded by CMS and declared as optimal for the physics analyses (in light orange), corresponding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

562 The main parameters characterizing an hadronic collider are the center-of-mass energy, correspond-  
563 ing to the sum of the energies of the beams, and the instantaneous luminosity, that describes the

frequency of the interactions among the bunches in the beams. If the bunches in the first beam contain  $n_1$  protons, and the bunches in the second beam contain  $n_2$  protons, and if the colliding area is  $\Sigma$ , the frequency of complete turns around the ring is  $f$ , the instantaneous luminosity  $\mathcal{L}_{\text{inst}}$  is:

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

If a generic physics process  $i$  has a cross-section of  $\sigma_i$ , the interaction rate  $R_i$  is:

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

and the number of events recorded in the time interval  $(0, \tau)$  is obtained from the integrated luminosity  $\mathcal{L} = \int_0^\tau \mathcal{L}_{\text{inst}} dt$ :

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

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

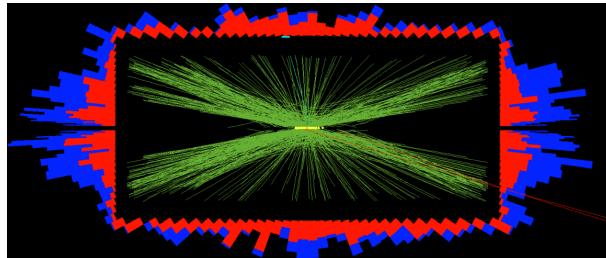


Figure 3.4: CMS collision event, where a record of 78 interactions per single bunch crossing were taking place simultaneously.

### 3.1.1 Proton-proton interactions

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

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

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

### 3.2 CMS detector

---

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

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

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

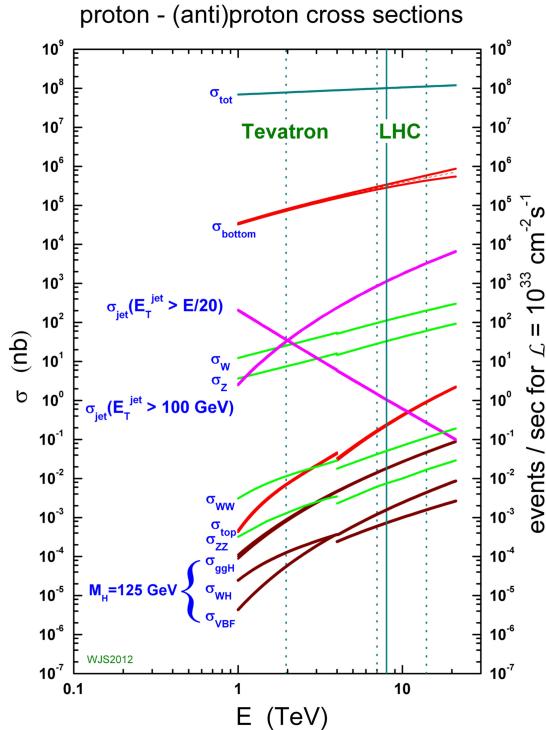


Figure 3.5: Cross-sections and number of expected events in proton-proton collisions, as a function of the center-of-mass energy. Rare phenomena, such as the Higgs boson production, can be observed at the LHC.

### 596 3.2 CMS detector

597 The Compact Muon Solenoid (CMS) is a multi-purpose detector built in the LHC ring. It is situated  
 598 in a cavern 100 m underground, near Cessy, in France. It is a cylinder 22 m long, with a diameter of  
 599 15 m, and a weight of 12500 tons. Its physics programme includes the search for the Higgs boson  
 600 (discovered in 2012), precision measurements of the Standard Model parameters and rare decays  
 601 (physics of beauty quark), and search for new physics beyond the standard model (SUSY, exotic  
 602 phenomena, dark matter, extra dimensions).  
 603 The CMS detector is structured in many layers of sub-detectors, giving different responses depend-  
 604 ing on the nature and the momentum of the particles passing through. The inner detectors have  
 605 been finely segmented in order to afford the high radiation levels and particle multiplicity at the

606 interaction point, so that the reduced occupancy of each layer allows to measure and distinguish  
 607 precisely the primary vertices of the hard interactions from the pile-up events. A very accurate time  
 608 resolution is necessary to synchronize all the subsystems together.

609

## CMS Detector

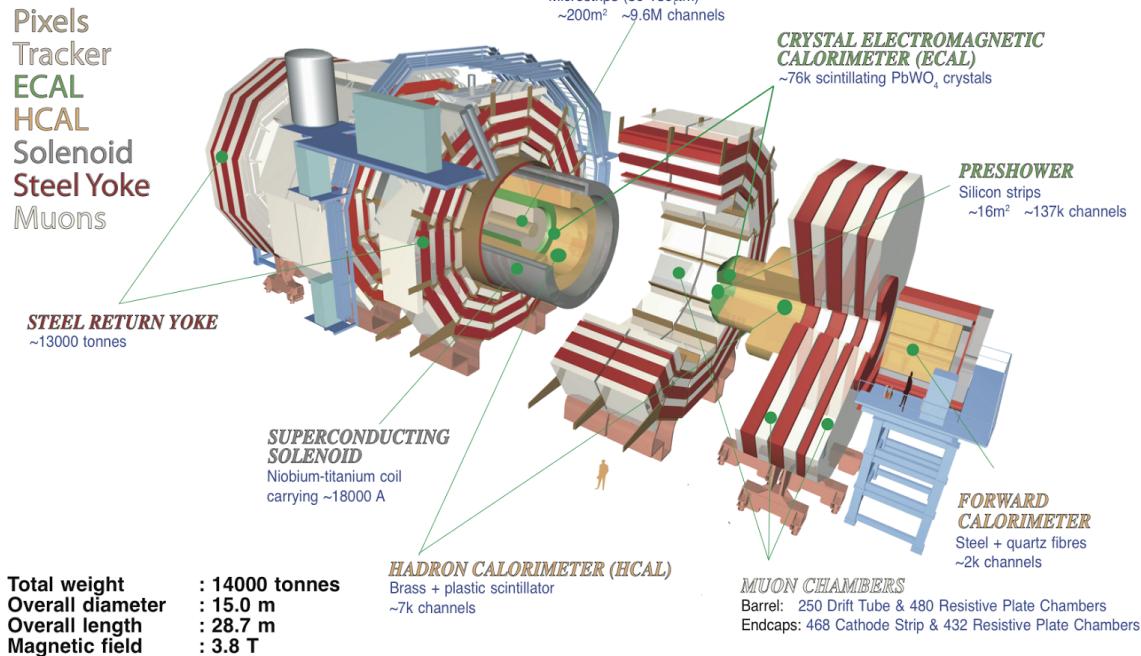


Figure 3.6: The CMS experiment.

610 Fig. 3.6 shows a sketch of the CMS detector. It is longitudinally segmented in the barrel region and  
 611 two endcaps. In the forward region (over the endcaps), where the beam radiation is very intense,  
 612 additional calorimeters have been placed. In fig. 3.7, the mean path of a specific particle through  
 613 the sub-detectors is represented, depending on its flavour.  
 614 A detailed description of the CMS detector can be found in [40].

### 615 3.2.1 The coordinate system

616 The CMS coordinate system is depicted in fig. 3.8.  $x$  and  $y$  are the coordinates in the transverse  
 617 plane,  $z$  is the longitudinal coordinate. The  $x$  axis points at the center of the LHC ring, the  $y$  axis  
 618 points upward, the  $z$  axis is along the beam direction. The azimuthal angle  $\varphi$  lies in the transverse  
 619 plane, and it is measured starting from the  $x$  axis; the radial coordinate is  $r$ . The polar angle  $\theta$  lies  
 620 in the plane  $rz$ . The transverse component of the 3-momentum,  $\vec{p}_T$ , is orthogonal to the beam axis  
 621 and lies in the plane  $xy$ . The transverse energy is defined as the magnitude of  $\vec{p}_T$ :  $E_T = E \sin \theta$ .  
 622 Two other commonly used variables are the rapidity,  $\mathcal{Y}$ , and pseudorapidity,  $\eta$ , defined as functions  
 623 of the particle energy  $E$ , the longitudinal component of the momentum  $p_z$  and the 3-momentum

### 3.2 CMS detector

---

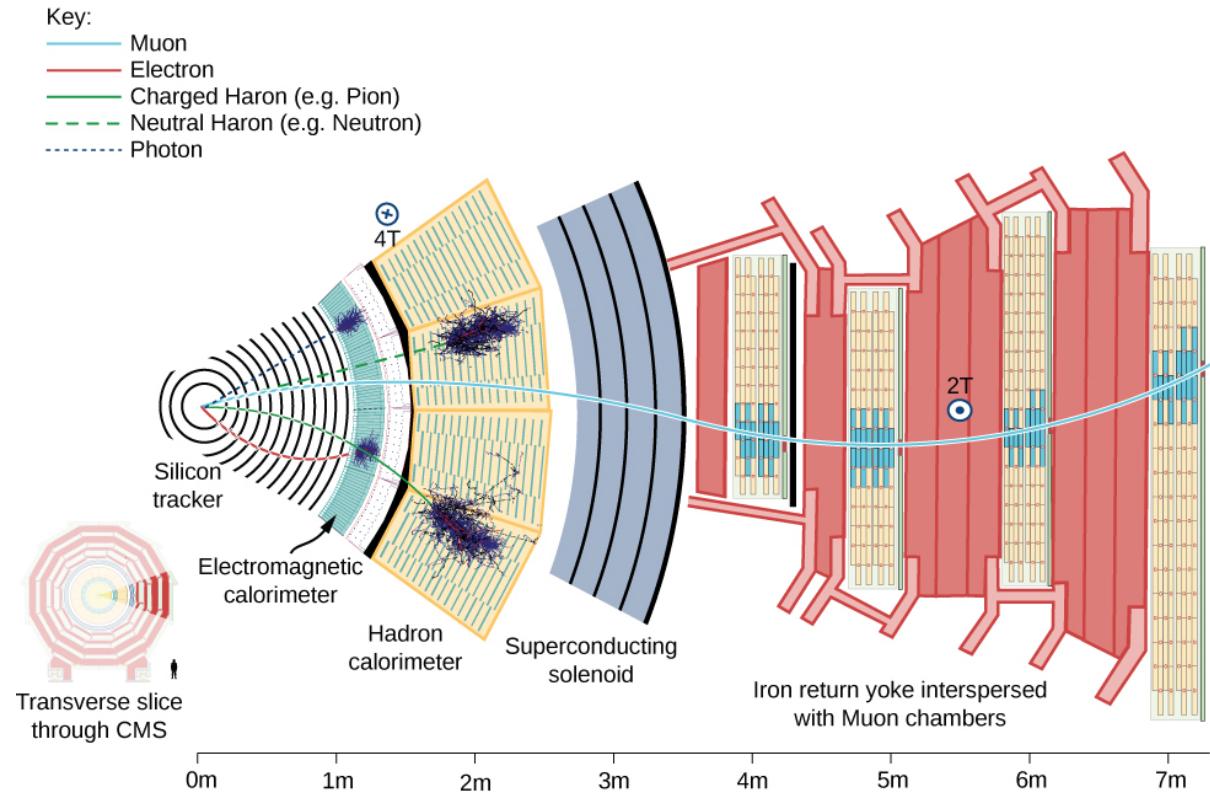


Figure 3.7: Mean path of a particle through the CMS detector. A muon, in light blue, passes through with a bended trajectory, depending on its momentum and charge, triggering signals in all the subsystems. An electron, in red, leaves a track in the silicon tracker and is absorbed by the electromagnetic 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.

624 modulus:

$$\begin{aligned} \gamma &= \frac{1}{2} \log \frac{E + p_z}{E - p_z} \\ \eta &= \frac{1}{2} \log \frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} = -\log \tan \frac{\theta}{2}. \end{aligned} \quad (3.5)$$

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

#### 630 3.2.2 The magnet

631 The CMS superconducting magnet is an hollow cylinder (13 m long, 6 m of diameter, shown in  
 632 fig. 3.9). An electrical current of 19 kA flows through the niobium and titanium fibers that consti-  
 633 tute the solenoid, providing a maximum magnetic field of 3.8 T and storing a maximum energy of  
 634 2.6 GJ. Superconducting conditions are mantained by a liquid helium cooling system, keeping the  
 635 solenoid temperature at 4.5 K. In order to avoid stray fields, the magnetic field lines are closed by

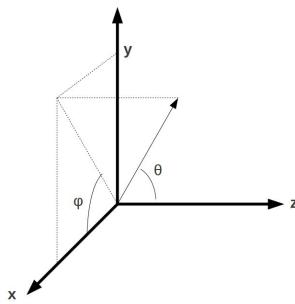


Figure 3.8: CMS coordinate system.

636 the return yoke, composed by 10 ktons of magnetized iron blocks, located in the outer part of CMS  
 637 and alternated to the muon chambers. The homogeneous magnetic field inside the detector bends  
 638 the trajectories of the charged particles, allowing the measurement of their momenta  $p$ , given the  
 639 relation with the magnetic field strength  $B$  and the radial coordinate  $r$  of the trajectory:

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

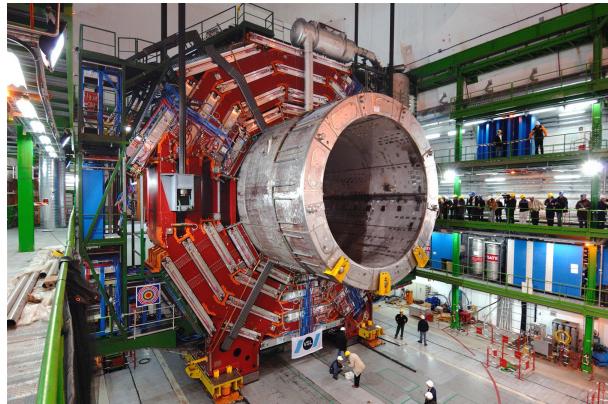


Figure 3.9: Installation of the superconducting solenoid in the CMS cavern.

### 640 3.2.3 The tracking system

641 The CMS tracking system [41, 42] is composed by a cylinder of silicon detectors (2.5 m of diameter  
 642 and 5.8 m length). Their design guarantees a precise reconstruction of the tracks left by charged  
 643 particles and of the interaction vertices, a fundamental tool to identify heavy quarks (charm, beauty)  
 644 and leptons (taus). Tracker detectors cover a pseudorapidity region of  $|\eta| < 2.5$  and have an active  
 645 area of  $210 \text{ m}^2$ . The two sub-detectors of the tracking system are the pixel detector, installed close  
 646 to the interaction point, and the strip detector, covering a radius of  $0.2 - 1.2 \text{ m}$ . The high granularity  
 647 of the pixels and strips allows to keep the occupancy at acceptable levels, given the high multiplicity  
 648 of the tracks ( $\sim 1 \text{ MHz/mm}^2$ ). The silicon detectors and the electronic cables are cooled down to a  
 649 temperature of  $\sim 10^\circ \text{ C}$ . The structure of the tracking system is shown in fig. 3.10.

#### 650 3.2.3.1 The pixel detector

651 The pixel detector is composed by 66 millions of silicon cells, whose dimensions are  $100 \times 150 \mu\text{m}^2$ ,  
 652  $285 \mu\text{m}$  of thickness, placed in 1440 modules. Silicon cells are set in three layers in the barrel re-

### 3.2 CMS detector

---

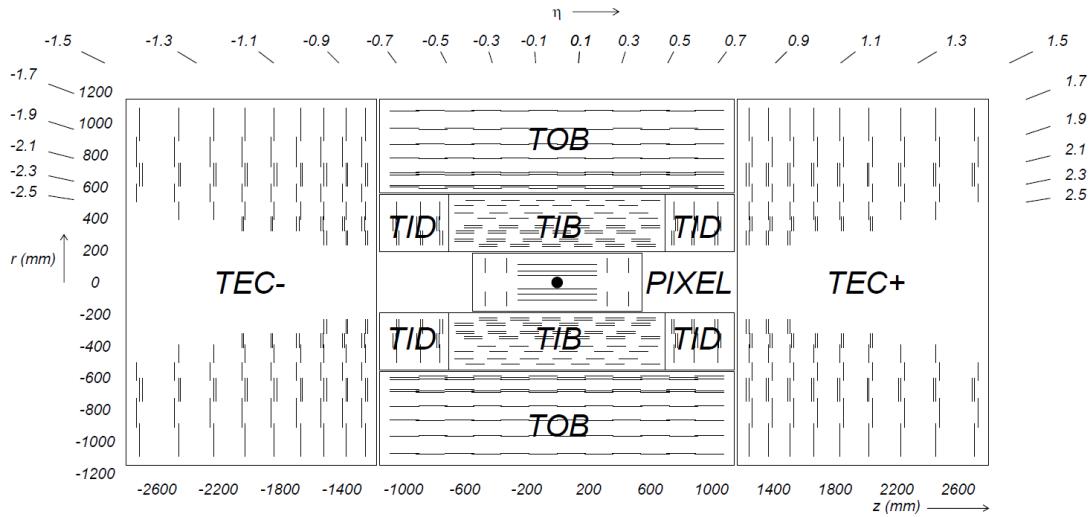


Figure 3.10: The CMS tracking system: the inner pixel detector, close to the interaction point, and the outer strip detector.

653 gion and in two disks at each endcap. Barrel modules are disposed parallel to the magnetic field,  
 654 whilst at the endcap they are tilted by about  $20^\circ$ . Pixels allow a spatial resolution of  $10\ \mu\text{m}$  in the  
 655 transverse plane, and of  $\sim 20\ \mu\text{m}$  along the longitudinal coordinate. Their reduced size guarantees  
 656 an occupancy of  $10^{-4}$  per pixel at each bunch crossing, in high luminosity regime.

657 **3.2.3.2 The strip detector**

658 The strip system is divided in the four-layered tracker inner barrel (TIB), covering a region  $20 < r <$   
 659  $55\ \text{cm}$  with respect to the interaction point, the six-layered tracker outer barrel (TOB), located at  
 660  $55 < r < 110\ \text{cm}$ , the three tracker inner disks (TID) and the nine tracker endcaps (TEC) at each  
 661 cylinder base. Given the lower radiation level at higher radii (and hence a lower occupancy, around  
 662 few percent), strips are bigger than the pixels. Silicon strips in TIB and TID are  $320\ \mu\text{m}$  thick,  $10\ \text{cm}$  long,  
 663 and with a pitch ranging from  $80$  to  $120\ \mu\text{m}$ ; strips in TOB and TEC are  $25\ \text{cm}$  long, with a  
 664 different thickness ( $320\ \mu\text{m}$  for TID,  $500\ \mu\text{m}$  for TEC) and pitch ( $97$ – $184\ \mu\text{m}$ ). There are 15148 strip  
 665 modules, and 9.3 million readout channels. The strip spatial resolution is about  $20$ – $50\ \mu\text{m}$  in the  
 666 transverse plane and about  $200$ – $500\ \mu\text{m}$  along the longitudinal coordinate.

667 **3.2.4 The electromagnetic calorimeter**

668 The CMS electromagnetic calorimeter (ECAL, shown in fig. 3.11) [43] is a homogeneous detector  
 669 composed by lead tungstate ( $\text{PbWO}_4$ ) scintillating crystals, designed to measure the energy de-  
 670 posits of photons and electrons through their electromagnetic showers.  $\text{PbWO}_4$  is transparent and  
 671 dense ( $8.3\ \text{gr}/\text{cm}^3$ ); it has a fast time response (the 85% of the scintillating light is emitted at every  
 672 bunch crossing, namely 24 ns), high scintillating efficiency and radiation resistance; it has a radia-  
 673 tion length is  $X_0 = 0.89\ \text{cm}$  and a Molière radius of  $2.19\ \text{cm}$ . The ECAL is divided in the barrel region  
 674 ( $\eta < 1.479$ , at a radius of  $1.3\ \text{m}$ ) and the endcaps ( $1.479 < \eta < 3$ ). The 61200 crystals employed in the  
 675 barrel region, whose size is  $(22 \times 22)\ \text{mm}^2 \times 23\ \text{cm}$ , have a radiation length of  $25.8X_0$ ; the 7324 crys-  
 676 tals in the endcaps,  $28.6 \times 28.6\ \text{mm}^2 \times 22\ \text{cm}$ , have a radiation length of  $24.7X_0$ . Before the endcaps,  
 677 on each side, a pre-shower detector is installed: it is composed by two disks of lead absorber and  
 678 two layers of silicon strips, of a radiation length up to  $3X_0$ . The pre-shower calorimeter has been

679 designed to distinguish the photons coming from the  $\pi^0$  decay, from the photons produced in the  
 680 rare Higgs decay  $H \rightarrow \gamma\gamma$ . The readout and amplification of the scintillating light, performed by  
 681 avalanche photodiodes in the barrel and by vacuum phototriodes in the endcaps, requires a stable  
 682 temperature of 18° C, mantained by a water cooling system.

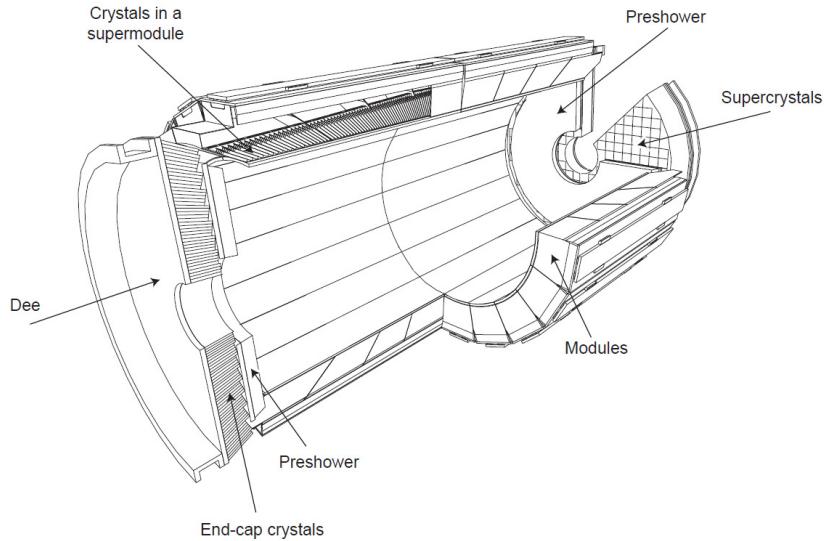


Figure 3.11: The CMS electromagnetic calorimeter.

683 The energy resolution of the calorimeter is parametrized as:

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

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

### 686 3.2.5 The hadronic calorimeter

687 The hadronic calorimeter (HCAL, displayed in fig. 3.12) [44] is a sampling calorimeter, composed by  
 688 brass and plastic scintillator layers. It has been designed to guarantee a good hermeticity, allowing  
 689 to perform a precise measurement of the missing transverse energy. It is located within the electro-  
 690 magnetic calorimeter and the solenoid, covering a region of  $|\eta| < 1.3$  in the barrel, and  $1.3 < |\eta| < 3$  in  
 691 the endcaps. Brass is non-magnetic and has short interaction length (16.4 cm): the 60 mm thick ab-  
 692 sorber layers used in the barrel reach 5.6 interaction lengths at  $\eta = 0$  and 10.8 interaction lenghts at  
 693  $\eta = 1.3$ ; the 80 mm thick layers in the endcaps reach 11 interaction lenghts. An additional calorimet-  
 694 ric layer has been installed out of the solenoid, in order to reach 11.8 interaction lenghts in the barrel  
 695 region. The scintillation light, typically in the blue-violet region of the electromagnetic spectrum, is  
 696 collected by wavelength-shifter fibers, translated and amplified by multi-channel hybrid photodi-  
 697 odes, proportionally to the magnitude of the energy deposits. An additional hadronic calorimeter  
 698 (HF) has been placed in the forward region,  $3 < |\eta| < 5.2$ , at 11.2 m from the interaction point. It has  
 699 beeен studied to afford the high levels of radiations: it is composed by 55 mm thick absorber layers  
 700 of stainless-steel, and quartz fibers, able to detect the Cherenkov scintillating light of the charged  
 701 particles of the hadronic showering. A longitudinally segmentation allow to distinguish hadronic

### 3.2 CMS detector

---

702 particles from electromagnetic components. The energy resolution of the hadronic calorimeter is:

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

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

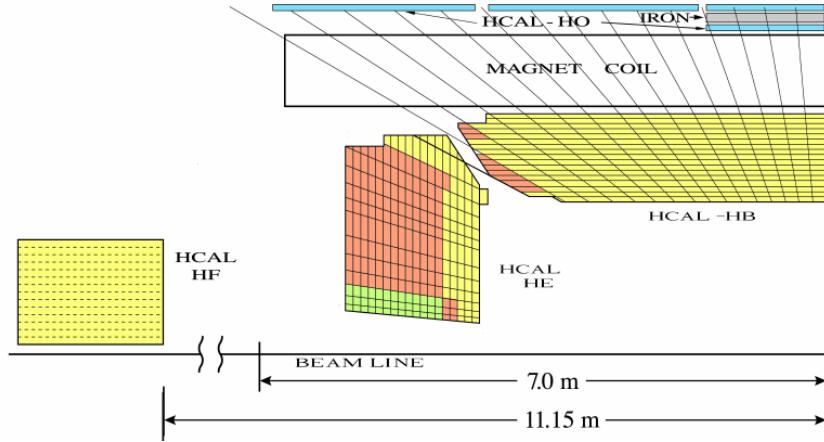


Figure 3.12: The CMS hadronic calorimeter.

#### 704 3.2.6 The muon system

705 The outer system of the CMS experiment consists into gas detectors for identifying muons [45], that  
706 are located between the iron return yokes, designed to close the magnetic field generated by the  
707 solenoid. In the barrel region, where a smaller number of muons is expected and the magnetic field  
708 is less strong, Drift Tubes (DT) detectors are installed. In the endcaps, where the flux of particles is  
709 larger, Cathod Strip Chambers (CSC) are used, and disposed in three disks. CSCs are designed to  
710 allow faster responses, higher granularity and radiation resistance. Resistive Plate Chambers (RPC)  
711 are installed both in the barrel and in the endcaps as additional triggering system. The geometry  
712 of the muon system is shown in fig. 3.13; it consists of 250 DTs, 530 CSCs, 610 RPCs, and it covers a  
713 region  $|\eta| < 2.4$ .

##### 714 3.2.6.1 The Drift Tubes

715 Drift Tube detectors cover a region of  $|\eta| < 1.2$  and are arranged in four stations, segmented along the  
716 beam line in five wheels. The basic element of the detector is the cell, that has a size of  $42 \times 13 \text{ mm}^2$ .  
717 Each cell is filled with a gas mixture (85% argon, 15%  $\text{CO}_2$ ), in which the process of ionization takes  
718 places; the ionization electrons drift from the  $50 \mu\text{m}$  thick steel anodic wire, located in the center  
719 of the cell, towards the aluminium cathodic strips, located at its edge. Additional electrodes on the  
720 surface of the cells allows to shape the electric field, in order to make the drift speed of the electrons  
721 uniform: the muon position is then extrapolated from the measurement of the drift time. Every  
722 station is composed by three cells superlayers. In the inner and the outer superlayers, the cells are  
723 oriented such in a way that the anodic wire is located along the  $z$  axis, to measure the  $\varphi$  coordinate.  
724 In the intermediate superlayer, wires are parallel to the radial coordinate, hence they can measure  
725 the  $z$  position. The spatial resolution of the system is  $100 \mu\text{m}$  in the  $(r, \varphi)$  plane,  $1 \text{ mrad}$  in the  $\varphi$   
726 coordinate, and  $150 \mu\text{m}$  in the longitudinal  $z$  coordinate.

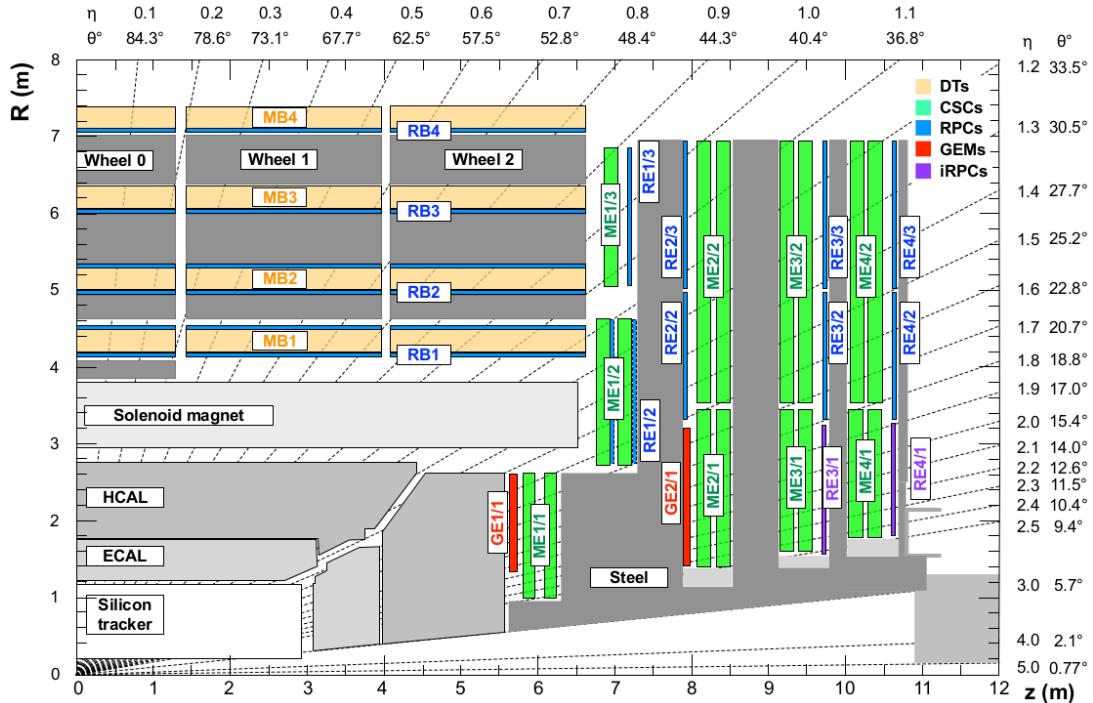


Figure 3.13: Section of CMS detector, in the plane ( $r, z$ ), parallel to the beamline, that emphasizes the location of the muon detectors, in particular: Drift Tubes (DT, in yellow); Cathode Strip Chambers (CSC, in green); Resistive Plate Chambers (RPC, in blue).

### 3.2.6.2 The Cathode Strip Chambers

Cathode Strip Chambers cover a region of  $0.9 < |\eta| < 2.4$ , overlapping with the DT 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  $\varphi$  coordinate. Ionization electrons produced by muons passing through the gas mixture in the chambers migrate from the anodes, inducing a charge distribution on the cathodes, from which the azimuthal coordinate can be reconstructed. The spatial resolution in the  $r$  coordinate is  $200 \mu\text{m}$ , and it is  $75 - 150 \mu\text{m}$  in the  $(r, \varphi)$  plane. CSCs are arranged in four disks and in three concentric rings.

### 3.2.6.3 The Resistive Plate Chambers

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

### 3.2.7 The trigger system and data acquisition

The CMS trigger system [46] has been designed considering the high instantaneous luminosity, such that it can provide a fast response and it allows to reduce the nominal event rate of 40 MHz in pro-

## 3.2 CMS detector

---

746 ton proton collision. The complexity of the CMS detector and the very high number of readout  
747 channels result into a huge amount of data per event, approaching the order of few MB per bunch  
748 crossing, hence 40 TB per second. The processes of handling and recording data are currently lim-  
749 ited by the employed technology to a frequency of  $\sim$ 100 Hz. Applying online selections to skim the  
750 events that are going to be written on tape, without rejecting interesting signals of hard processes  
751 and rare phenomena, is therefore a crucial and challenging point for every data analysis. Events are  
752 filtered by trigger selections at different levels: the Level-1 (L1) trigger is an hardware device, that  
753 allows to reduce the event rate from 40 MHz to the order of 100 kHz; the High Level Trigger (HLT)  
754 is a set of software algorithms that skims the event rate down to few hundred Hz. Once the trigger  
755 decisions are taken, the final events are handled by the Data Acquisition System (DAQ), that collects  
756 the informations coming from the sub-detectors and sends them to the storage unities.

757 **3.2.7.1 The Level-1 trigger**

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

771 **3.2.7.2 The High Level Trigger**

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

779 **3.2.7.3 Data acquisition, computing and storage**

780 The DAQ system deals with the storage, transfer and handling of the data collected by CMS; it also  
781 supports and stores the data simulations and calibrations of the sub-detectors. The CMS compu-  
782 tational resources are located in worldwide distributed data nodes, called Tiers. The CMS software  
783 (CMSSW) is based on an object oriented architecture (mainly C++). The basic unity of every data,  
784 both real and simulated ones, is the Event, that could contain very rough informations (RAW data  
785 format) or higher level refined objects (AOD, Analysis Object Data) where all the calibrations and  
786 corrections needed to properly deal with the final physics objects are already in place. Data are  
787 handled by C++ or python modules, and the outputs are written in ROOT [47] files.

---

### 788 3.2.8 Particle Flow event reconstruction

789 The particle flow (PF) algorithm [48] aims at identifying and reconstructing each particle produced  
 790 by the proton-proton collisions, combining the informations coming from all the CMS sub-detectors.  
 791 It is particularly suitable to improve the reconstruction of jets, missing transverse momentum (used  
 792 to identify neutrinos) and hadronically decaying tau leptons.  
 793 The association of the informations is performed at different stages. The reconstruction of the  
 794 charged particles in the silicon detector is performed with an iterative algorithm, and the recon-  
 795 structed object is called a tracker track. Then, a clustering algorithm is performed to collect and  
 796 combine the energy deposits in the calorimeters, in such a way to distinguish neutral from charged  
 797 particles, reconstruct their directions, and improve the energy measurement of the very energetic  
 798 charged particles, whose tracks are less bended by the magnet and hence less precisely determined.  
 799 The last informations are provided by the hits collected in the muon system. The three sets of re-  
 800 constructions are then combined with a link algorithm, that aims at associating tracker tracks to  
 801 calorimeter clusters and muon hits with geometrical criteria. A track in the silicon detector is linked  
 802 to a calorimeter cluster if the extrapolated position lies in the cluster itself. Similarly, clusters in dif-  
 803 ferent calorimeters are linked when the position in the more granular calorimeter (*i.e.* ECAL) lies in  
 804 the envelope of the clusters in the less granular calorimeter (*i.e.* HCAL). The decision of linking a  
 805 tracker track to a muon track is based on the  $\chi^2$  of a global fit between the two tracks.  
 806 The particle flow algorithm then interprets the collected and linked informations as different par-  
 807 ticles. Muons are identified by the combination of a track in the silicon detectors and a track in the  
 808 muon chambers. Photons are determined directly by ECAL clusters. Electrons energies and posi-  
 809 tions are measured by ECAL clusters, linked to a corresponding tracker track, and considering all the  
 810 energy clusters produced by the bremsstrahlung photons radiated while interacting with the detec-  
 811 tor material. The hadrons are identified by the tracks (if charged) linked to the corresponding ECAL  
 812 and HCAL clusters. The hadron energy resolution, 10% at 100 GeV combining ECAL and HCAL, is  
 813 such that neutral hadrons can be distinguished as an energy calorimetric excess when overlapped  
 814 by a charged hadron occupying the same calorimetric towers. Finally, the missing transverse mo-  
 815 mentum is defined as the negative sum of the transverse momenta of all the particles identified by  
 816 the PF algorithm.

817

### 818 3.2.9 Physics objects

#### 819 3.2.9.1 Track reconstruction

820 The reconstruction of the trajectories of the charged particles passing through the CMS detector is  
 821 performed by multiple iterations of the Combined Track Finder algorithm, that is based on a Kalman  
 822 filter approach [49]; given the high multiplicity of particles produced at each bunch crossing and the  
 823 multiple scatterings in the detector materials, tracking represents a challenging task. The CTF al-  
 824 gorithm builds a track starting from the so-called seeds, namely triplets of hits collected in the pixel  
 825 detector inner layers, or couples of hits if the track originates from the interaction point. The ini-  
 826 tial guess of the track given by the seeds is then extrapolated to the outer layers: if other hits are  
 827 found to be compatible with the trajectory hypothesis ( $\chi^2$ -based hypothesis test), they are added to  
 828 the track. Once the outer layers are reached, another reconstruction is performed backward, in or-  
 829 der to clean the track from spurious hits and enhance the tracking efficiency. The final collected hits  
 830 are re-fitted with Kalman filter and more precise algorithms, in order to improve the quality of the  
 831 measurement. If two tracks share more than a half of their hits, the worst quality track is rejected.  
 832 The track reconstruction efficiency for particles with  $p_T > 0.9$  GeV is 94% in the barrel and 85% in

833 the endcap region [42].

834

835 **3.2.9.2 Vertices reconstruction**

836 The reconstruction of the vertices at each bunch crossing is performed in steps. Primary vertices  
 837 are originating from the proton proton collisions, whilst secondary vertices are due to long-lived  
 838 particles (heavy quarks and  $\tau$  leptons). The starting point of the procedure is clustering the re-  
 839 constructed tracks originating from the primary vertex; the decision is taken by the deterministic  
 840 annealing algorithm [50], taking as input the longitudinal impact parameter. The algorithm allows  
 841 to distinguish vertices separated down to 1 mm. The second step is run by the adaptive vertex fit-  
 842 ter [51], that measures the position of the vertex for the chosen set of tracks. The algorithm is based  
 843 on an iterative re-weighted Kalman filter, that down-weights the wrongly associated tracks not com-  
 844 patible with the considered vertex. The primary vertex is selected as the vertex where the sum of the  
 845  $p_T^2$  of the associated tracks is the largest. The spatial resolution on the vertex position is 10-40  $\mu\text{m}$   
 846 in the  $(r, \varphi)$  plane, and 15-50  $\mu\text{m}$  in the longitudinal coordinate.

847 **3.2.9.3 Electrons and photons reconstruction**

848 Electrons are reconstructed [52] combining a track with the energy deposits clustered in the ECAL,  
 849 due to the showering of the electron through the detector and the emission of bremsstrahlung pho-  
 850 tons. The combination can proceed both from the silicon detector in the outgoing direction and in  
 851 the opposite way: the tracker seeding as starting point is suitable for low energy electrons, whose  
 852 trajectories are less bended and hence more accurately measured by the tracker system; the group-  
 853 ing of ECAL clusters (called superclusters) followed by a consecutive track extrapolation, performed  
 854 by taking into account the electron interaction with the detector material, is more efficient in case of  
 855 high energetic electrons, due to the higher resolution of the ECAL scintillating crystals. A Gaussian-  
 856 sum filter algorithm (GSF) [53] allows to properly take into account the effects of the bremsstrahlung  
 857 radiation, that is distributed not as a single Gaussian (standard Kalman filters) but rather as a sum  
 858 of Gaussian functions.

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

866 The electron energy is determined correcting the raw energy measurement of the ECAL superclus-  
 867 ters by taking into account the effects of the losses due to radiation or gaps between the calorimeter  
 868 modules, and the pile-up contribution. The electron momentum resolution has been measured  
 869 in  $Z \rightarrow e^+e^-$  decays in Run 1 LHC data, and it varies from 1.7% to 4.5% depending on the pseu-  
 870 dorapidity range [54]. The electron isolation variable is defined as the  $p_T$  sum of the charged and  
 871 neutral particles lying in a cone of  $\Delta R = 0.3$  around the electron trajectory, divided by the transverse  
 872 momentum of the electron itself:

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

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

878 Photons are reconstructed with the ECAL clusters only. Given their importance in the discovery of  
 879 the Higgs boson, dedicated studies have been performed both in data and in Monte Carlo simula-  
 880 tions reproducing the  $H \rightarrow \gamma\gamma$  process. Particular care has been taken in the treatment of the photon  
 881 conversions into electron-positron pairs while interacting with the tracker detector. Dedicated se-  
 882 lections allow to define different photon identification working points. Similarly to the case of the  
 883 electrons, the photon isolation variable can be defined. The photon energy resolution varies from  
 884 1% to 3%, depending on the  $\eta$  range [55].

#### 885 3.2.9.4 Muon reconstruction

886 A muon candidate can be built exploiting the hits collected in the silicon tracker (track) and in the  
 887 muon system (standalone muon) [56]. Each muon sub-detector (DTs, RPCs and CSCs) performs a  
 888 local reconstruction of the particle candidate; the informations from the three muon chambers are  
 889 combined with a Kalman filter approach.

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

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

903 Starting from 2016 LHC Run, the muon reconstruction takes into account the Alignment Position  
 904 Errors, namely the uncertainties due to the position of the muon chambers with respect to the sil-  
 905 icon detectors. The final resolution on the muon momentum measurement depends on the  $p_T$   
 906 and  $\eta$  of the candidate, and ranges from 1% for very low momenta, up to  $\sim 7\%$  ( $|\eta| < 0.9$ ) – 10%  
 907 ( $1.2 < |\eta| < 2.4$ ) [57].

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

#### 910 3.2.9.5 Jet reconstruction

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

915 Jets are reconstructed starting by the PF candidates in the event. The charged hadron subtraction al-  
 916 gorithm (CHS) removes candidates not associated to the primary vertex in order to suppress pile-up

### 3.2 CMS detector

---

917 contributions [58]. The remaining particles are used as input to jet clustering algorithms to recon-  
918 struct particle-flow jets. The jets are clustered using the FASTJET package [59] with the anti- $k_T$  jet  
919 sequential clustering algorithm [60]. A sequential clustering algorithm is designed to be infrared  
920 and collinear safe, namely, if the final state particles undergo a soft emission or a collinear gluon  
921 splitting, the number and shapes of the jets should not change. The starting point of a sequen-  
922 tial clustering algorithm is the definition of the distances between two particles  $i$  and  $j$ , and the  
923 distance of a given particle  $i$  from the beam-spot  $B$ :

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

924 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  
925 distance between the particles,  $a$  is an exponent depending on the clustering algorithm chosen,  
926 and  $R_0$  is the clustering parameter. The algorithm then operates as follows:

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

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

939 In this analysis, clustering parameters of  $R_0 = 0.8$  and  $R_0 = 0.4$  will be used to define the “fat”-jets or  
940 AK8 jets, and the “standard”-jets or AK4 jets. In order to avoid double-counting of PF candidates,  
941 AK4 jets are considered only if the angular separation from the leading AK8 jet is larger than  $R_0 > 0.8$ .  
942 Since the detector response to different particles is non-linear, particular care should be taken in the  
943 assignement of the measured momentum of the clustered jet to the corresponding true value of the  
944 original parton [61]. A set of jet energy corrections (JECs) are applied sequentially and with a fixed  
945 order. Each correction consists in a rescaling of the jet four-momentum, and it takes into account  
946 different effects that are factorized.

- 947 • The L1 JECs remove the effect of the pile-up; they consist into an offset correction of the jet  
948  $p_T$ . They are determined from Monte Carlo (MC) simulations of di-jet events produced by  
949 strong interaction with and without pile-up events on top, and parametrized as a function of  
950 kinematical variables (jet area, pseudorapidity and  $p_T$ ) and of the average  $p_T$  density per unit  
951 area,  $\rho$ . Residual differences between data and the detector simulation are evaluated in data  
952 collected with a random trigger, called zero bias, applying the only requirement of the beam  
953 crossing happening. Pile-up offset corrections are displayed in fig. 3.14 (top left), as a function  
954 of the jet pseudorapidity.

- The simulated response of the detector is not uniform over jet  $p_T$  and  $\eta$ . This effect is mitigated by the L2L3 MC-truth corrections. They are calculated in MC simulations of di-jet events, by taking into account the discrepancy between the reconstructed  $p_T$  of the jet and the true  $p_T$  at particle generator level (*i.e.*, before simulating the interaction of the parton showers with the detector), as a function of jet  $p_T$  and  $\eta$ . L2L3 scale factors describing the simulated jet response are reported in fig. 3.14 (top right), as a function of the jet pseudorapidity.
  - The small data-MC discrepancies ( $\sim 1\%$ ) left after applying the previous set of JECs are corrected by the L2 and L3 residual corrections. The L2Residuals are calculated in di-jet events, as a function of  $p_T$ . The L3Residuals are calculated in  $Z \rightarrow (\mu\mu, ee) + \text{jet}$  events, photon + jet events and multijet events, as a function of  $\eta$  and  $p_T$  [61]. Data-MC scale factors for L2L3Residuals are displayed in fig. 3.14 (bottom), as a function of the jet  $\eta$  and  $p_T$ .
  - An optional correction, not used in this analysis, is the L5 flavour-dependent correction, that is extracted from MC simulations.
- Each jet energy correction is determined with an uncertainty, and reported in fig. 3.15 for 2015 data, as a function of  $p_T$  and  $\eta$  of the jet. The total uncertainty for jets with  $p_T$  larger than 30 GeV (100 GeV) is smaller than 3% (1%) in the barrel, and up to 5% (3%) in the endcaps [62].
- An additional effect that must be taken in account in the analysis is the discrepancy in the jet energy resolution (JER) observed in data and in Monte Carlo samples. A smearing procedure is applied in MC simulations (described in detail in sec. 4.3.6), in order to restore a better agreement. Jet energy resolutions in Monte Carlo simulations are displayed in fig. 3.16 (top), as a function of the jet  $p_T$  and the average number  $\mu$  of reconstructed primary vertices, considering central (left) and forward (right) jets. The resolution is stable against the pile-up for jet  $p_T > 100$  GeV, and it ranges from 10% at 100 GeV, down to 4% at 1 TeV [62]. In fig. 3.16 (bottom), data-MC smearing scale factors are reported as a function of  $\eta$ .

### 3.2.9.6 Tau reconstruction

Tau leptons have a very small lifetime ( $\sim 3 \times 10^{-13}$  s), hence they decay before reaching the pixel detector and they can only be reconstructed through their decay products. Approximatively 60% of the times,  $\tau$  leptons decay in hadrons, hence they are reconstructed as small collimated jets in the CMS detector. The main decay modes of the hadronic tau,  $\tau_h$ , are one or three charged mesons ( $\text{mainly } \pi^\pm$ ), also in association with a  $\pi^0$  decaying in a couple of photons, and a  $\tau$  neutrino. Hence, photons and charged hadrons are the main ingredients of dedicated algorithms to perform the  $\tau_h$  reconstruction and identification, in order to distinguish them from quark and gluon-initiated jets. The main CMS  $\tau_h$  reconstruction algorithm, Hadron Plus Strips (HPS) [63], is particle-flow based. HPS builds the tau candidate from a PF jet, clustered with the anti- $k_T$  algorithm with  $R_0 = 0.5$ , and it reconstructs the  $\pi^0 \rightarrow \gamma\gamma$  decays within the jet cone, by taking into account the photon conversions in the silicon detector. The exploitation of the PF informations is such that the HPS algorithm shows stable performances in the reconstruction of the  $\tau_h$  energy as a function of the energy itself. The  $\tau_h$  candidate is required to be isolated, namely no energy deposits other than the  $\tau$  decay products should be present in the tau cone. Depending on the low threshold set to consider the surrounding particles as included in the cone, different isolation working points can be defined. With the looser working point, the probability of mis-identifying a quark or gluon jet as a tau is around 1% [63].

### 3.2.9.7 b-jets tagging

The bottom quark plays a fundamental role in numerous standard model processes, *i.e.* the physics related to the top quark (that decays into a W and a b-quark with a branching fraction of 100%) and

### 3.2 CMS detector

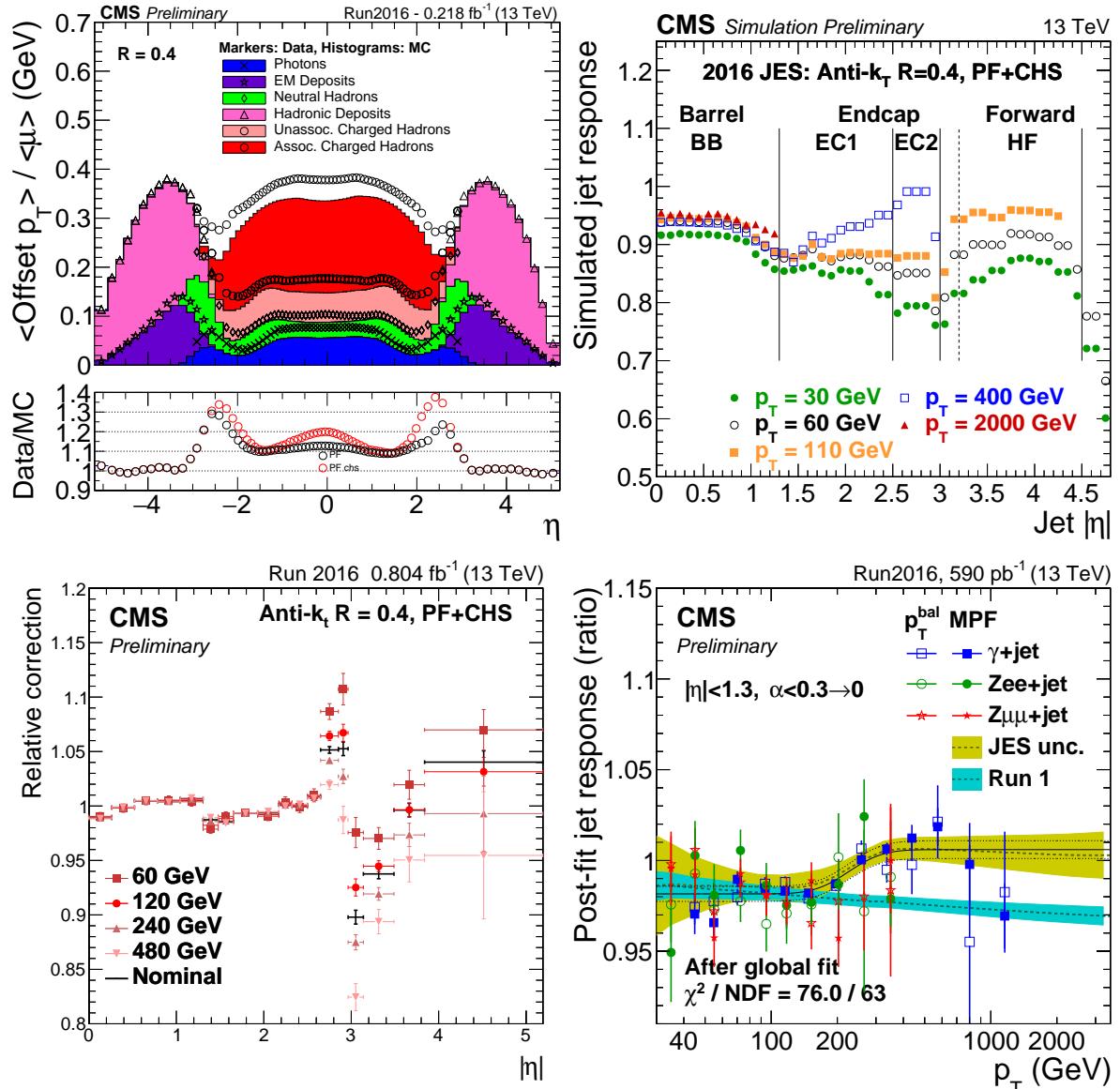


Figure 3.14: Top left: average  $p_T$  offset due to additional pile-up events, measured both in data and in MC simulations, as a function of the jet pseudorapidity. Top right: simulated jet response (L2L3 MC-truth corrections), as a function of the jet pseudorapidity. Bottom left: L2L3 residual data-MC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

the Higgs boson (decaying into  $b\bar{b}$  with a branching fraction  $\sim 60\%$ ). Many algorithms have been exploited by the CMS collaboration, with the aim of distinguishing a b-quark initiated jet and jets originating from light quarks or gluons [64]. The most remarkable feature of the b-quark is the long lifetime ( $\sim 1.5$  ps), that has the experimental consequence of a displaced decay (few mm) with respect to the primary vertex. The direct leptonic decays of the b-quark (into  $\mu$  and  $e$ ) or the cascade leptonic decays involving c-quarks give an additional handle to its identification. Given the high spatial resolution of the silicon detector, track reconstruction is a key point of the b-tagging procedure. Tracks inside a jet candidate must satisfy criteria related not only to their quality but also on their distance from the interaction point. The track impact parameter is the distance

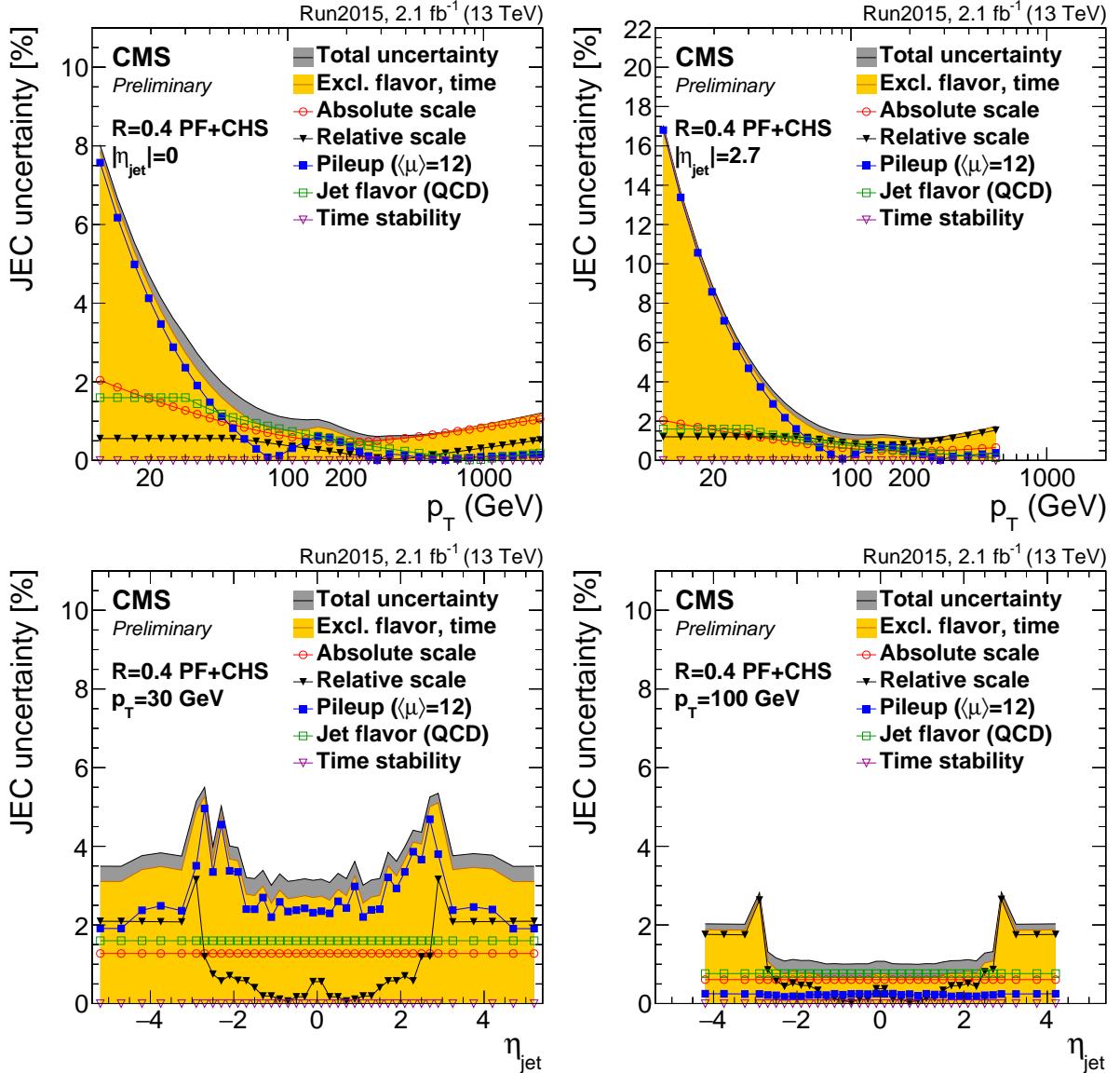


Figure 3.15: Jet energy corrections uncertainties, as a function of jet  $p_T$  (top) and  $\eta$  (bottom), calculated in 2015 data. The yellow histograms report the convolution of the uncertainties applied in the analysis.

between the primary vertex and the coordinate of closest approach. Tracks that are too far from the interaction point are discarded, in order to suppress the pile-up contributions. The Combine Secondary Vertex (CSV) algorithm [65] sorts jet candidates in categories, based on the number of reconstructed secondary vertices (one reconstructed secondary vertex, no secondary vertices but two tracks with high impact parameters, and the remaining cases). A multivariate approach allows to train the algorithm over the categories, considering as discriminating variables both tracking informations (numbers and properties of the tracks) and their relations with the secondary vertex reconstruction (impact parameters; angular, linear, 2D and 3D distances of the vertex from the tracks and the jet axis; invariant mass of the charged particles associated to the secondary vertex). By tuning the selections, working points with different efficiencies have been set. The loose working point, used in this analysis, has a 90% signal efficiency and a 40% mis-identification rate. The

### 3.2 CMS detector

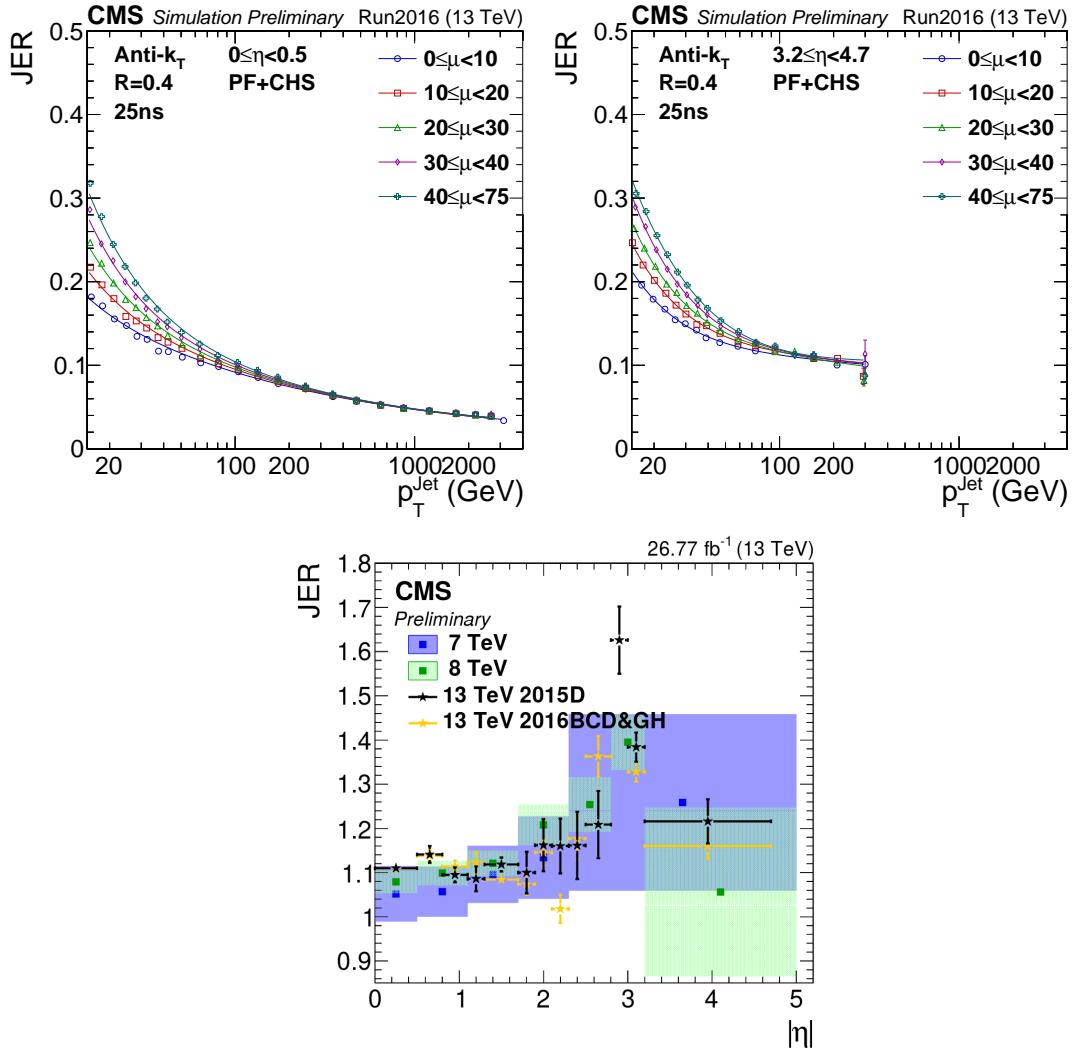


Figure 3.16: Top: jet energy resolution in MC simulations, as a function of the jet  $p_T$ . Different curves represent a different average number of primary vertices per event ( $\mu$ ). Bottom: data-MC scale factors, as a function of the jet  $\eta$ , measured in 2016 data (yellow dots).

1019 b-tagging efficiency is different in data and in simulations. Multiplicative scale factors are calcu-  
 1020 lated in events enriched in b-quark jets.

1021

1022 **3.2.9.8 Missing transverse energy reconstruction**

1023 Neutrinos can interact with the other particles only via the electroweak interactions; hence, when a  
 1024 neutrino is produced in the proton-proton collisions, it passes through the CMS experiment, unde-  
 1025 tected. Its only experimental signature is the momentum imbalance ( $\vec{p}_T^{\text{miss}}$ ) in the transverse plane  
 1026 ( $r, \varphi$ ). The magnitude of  $\vec{p}_T^{\text{miss}}$  vector is also called missing transverse energy,  $E_T^{\text{miss}}$ . Given its def-  
 1027 initition, it is evident that  $E_T^{\text{miss}}$  is a delicate variable to deal with, since it depends on all the other  
 1028 objects, on their imperfect measurements, on the detector noise and the pile-up events.

1029 The PF  $E_T^{\text{miss}}$  is the negative sum of the transverse momenta of the PF candidates reconstructed in  
 1030 the event. Inefficiencies in the tracker reconstruction and non-linear responses of the calorimeters

1031 can be corrected by propagating the jet energy corrections to  $\vec{p}_T^{\text{miss}}$  [66]:

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

1032 where "corr" ("raw") is related to corrected (raw)  $p_T$  of the considered jet. This correction is known  
 1033 as the "Type-I" correction to  $E_T^{\text{miss}}$ . Jets included in the calculation are AK4 with CHS algorithm applied  
 1034 to remove the pile-up contribution, they must have  $p_T > 15$  GeV and less than 90% of their  
 1035 energy deposited in the electromagnetic calorimeter. If a muon lies in the jet cone, it is subtracted  
 1036 from the jet and added after the  $p_T$  correction. A similar correction is performed to correct  $\vec{p}_T^{\text{miss}}$  at  
 1037 trigger level; in this case, a jet  $p_T$  threshold of 35 GeV is chosen.

1038 The  $E_T^{\text{miss}}$  uncertainty depends on the topology of the final state. It is calculated per-event by factorizing  
 1039  $\vec{p}_T^{\text{miss}}$  in components: electrons, photons, muons, taus, jets, jets with  $p_T < 10$  GeV and all  
 1040 the remaining PF candidates that are not clustered inside jets, called unclustered energy. The mo-  
 1041 mentum of every object is varied within its uncertainties (namely, the energy scale and resolution),  
 1042 and the effects are propagated to  $\vec{p}_T^{\text{miss}}$ . The most significant contributions in the unclustered en-  
 1043 ergy is due to neutral PF hadrons and hadrons reconstructed in the forward hadronic calorimeter.  
 1044 The effects related to jet energy scale and unclustered energy scale are measured on simulation, in  
 1045 events with a top and an anti-top quarks, and amounts to 5% and 30% respectively [66].

1046

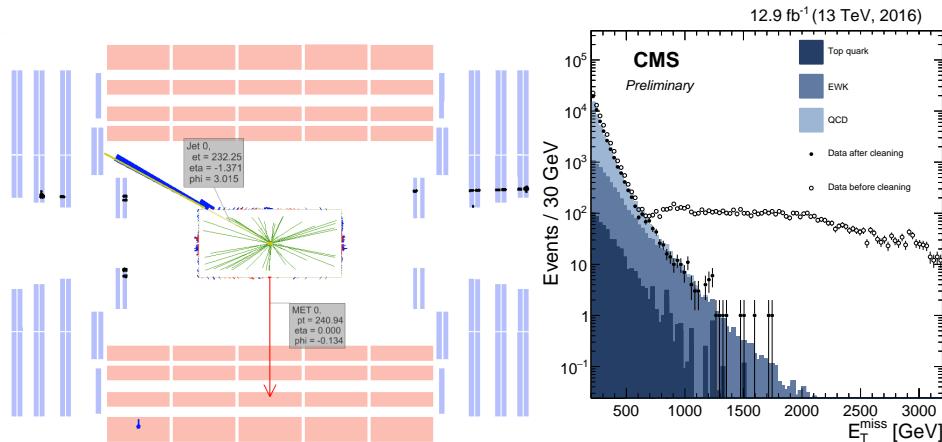


Figure 3.17: Left: event display of beam halo particles hitting the CSC detector. Right: comparison of data and simulations (histograms) when di-jet events are selected, before (open markers) and after (filled markers) anomalous  $E_T^{\text{miss}}$  cleaning algorithms have been applied on data.

1047 Many instrumental effects can give rise to anomalous  $E_T^{\text{miss}}$  determination: they have been studied  
 1048 in detail during Run1 [67,68] and Run2 [66], and they are mainly caused by ECAL and HCAL. In ECAL,  
 1049 anomalous  $\vec{p}_T^{\text{miss}}$  is caused by particles hitting the sensors of the photodetectors, or by beam halo  
 1050 particles (namely, particles produced in spurious proton interactions before reaching the interac-  
 1051 tion point in the detector) showering inside the calorimeter, or by losses due to ECAL dead cells. An  
 1052 event display representing beam halo muons hitting the CSC detector is shown in fig. 3.17 (top). In  
 1053 HCAL, spurious  $\vec{p}_T^{\text{miss}}$  can be related to noise in the hybrid photodiodes and readout frontend. In  
 1054 HF, missing  $p_T$  can be related to particles lost in the light guides and photomultipliers. Additional  
 1055 anomalous  $E_T^{\text{miss}}$  can be produced by low quality muon tracks, that are not linked to segments re-  
 1056 constructed in the muon chambers by the PF algorithm. These tracks are then classified as charged  
 1057 hadrons, taken into account in the  $\vec{p}_T^{\text{miss}}$  calculation, and result into a large amount of fake  $E_T^{\text{miss}}$ .

### 3.2 CMS detector

---

1058 Dedicated algorithms have been designed to identify and reject events with anomalous  $E_T^{\text{miss}}$ , and  
 1059 they are consistently applied on data and simulations. In fig. 3.17 (right), Monte Carlo simulations  
 1060 (coloured histograms) are compared to data before the algorithms removing the anomalous  $E_T^{\text{miss}}$   
 1061 have been applied (open markers) and after the cleaning (filled markers): the spurious high- $\vec{p}_T^{\text{miss}}$   
 1062 tail has been suppressed.

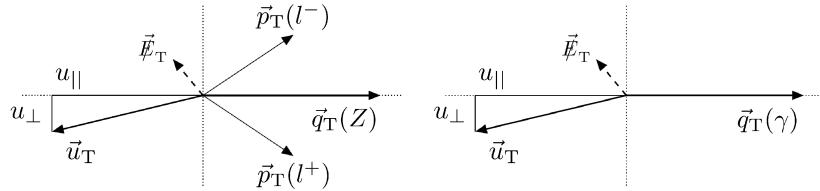


Figure 3.18: Left (Right): kinematics of  $Z \rightarrow \ell\ell$  (photon) events in the  $(r, \varphi)$  plane;  $\vec{u}_T$  is the hadronic recoil,  $\vec{q}_T$  is the transverse momentum of the considered boson.

1063 Performance of  $E_T^{\text{miss}}$  reconstruction are studied in events with a leptonic decay of a  $Z$  boson (in two  
 1064 muons or in two electrons) or an isolated photon. The distributions of  $E_T^{\text{miss}}$  are shown in fig. 3.19,  
 1065 separately for the three event categories. The hadronic recoil  $\vec{u}_T$  is defined in the transverse plane  
 1066 as the vectorial sum of all the particle transverse momenta, except the momentum  $\vec{q}_T$  of the vector  
 1067 boson considered ( $Z$  or  $\gamma$ ). From the momentum conservation, the following relation holds:

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

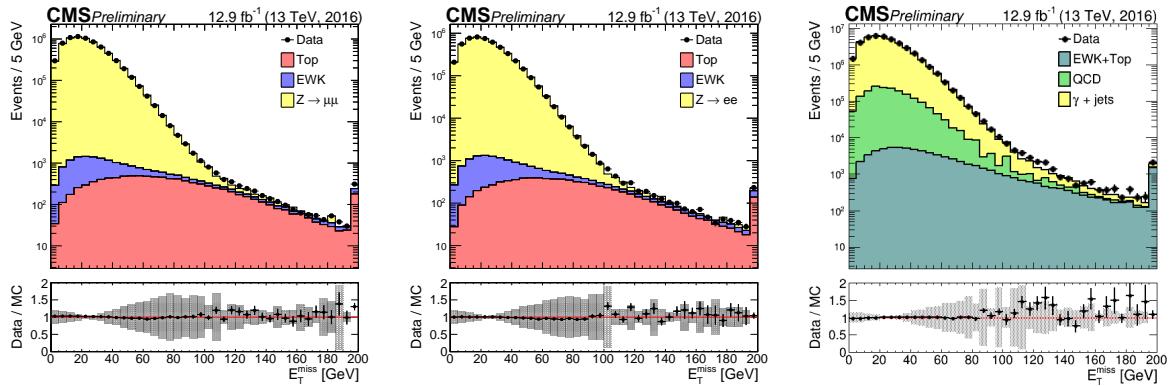


Figure 3.19: Data (black markers) and Monte Carlo (histograms) distributions of  $E_T^{\text{miss}}$  variable, in events reconstructing respectively a  $Z \rightarrow \mu\mu$  decay (left), a  $Z \rightarrow ee$  decay (center), an isolated photon (right).

1068 The hadronic recoil is projected in the parallel and perpendicular directions with regards to  $\vec{q}_T$ . The  
 1069 components  $u_{||}$  and  $u_{\perp}$ , along with the vectors described in eq.3.12, are schematically represented  
 1070 in fig. 3.18. The  $E_T^{\text{miss}}$  response, defined as  $-\langle u_{||} \rangle / \langle q_T \rangle$ , is calculated as a function of  $q_T$  in data and  
 1071 simulations (fig. 3.20, left). The distributions of the two components of the hadronic recoil,  $u_{||} + q_T$   
 1072 and  $u_{\perp}$ , are modelled as Voigtian functions (the convolution of a Gaussian with a Breit-Wigner). The  
 1073 resolution of each component is calculated as the full width at half maximum of the corresponding  
 1074 Voigtian, and it is displayed in fig. 3.20 (center and right plots), as a function of  $q_T$ .

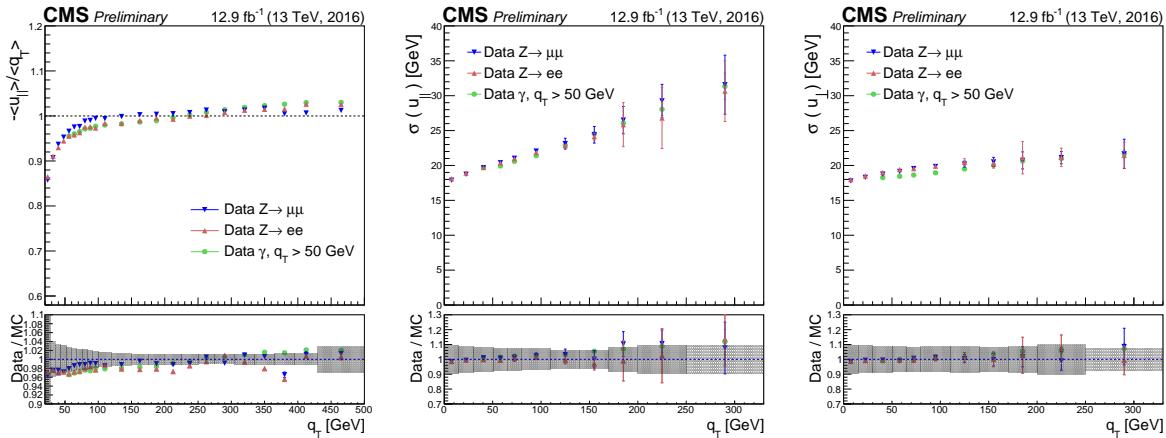


Figure 3.20: Left:  $E_T^{\text{miss}}$  response, as a function of the transverse momentum  $q_T$  of the vector boson considered in the event (Z decaying in  $\mu\mu$ , Z in  $ee$  or a photon), measured on data. Center and right: resolution on the measurement of the parallel and perpendicular hadronic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

### 1075 3.3 ATLAS, ALICE, LHCb detectors

#### 1076 3.3.1 ATLAS

1077 ATLAS (A Toroidal LHC ApparatuS) [69] is a multi-purpose experiment, that shares the same scientific aims 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.

1080 ATLAS has a cylindrical shape (diameter of 25 m, length of 46 m) and weights 7000 tons. Like CMS, 1081 ATLAS is composed by many sub-detectors: trackers, calorimeters and muon system. The ATLAS 1082 magnetic field is provided by a solenoid, located inside the cylinder, and a big toroid, located outside 1083 the sub-detectors, able to reach a magnetic field of 2 T at the interaction point. The main differences 1084 among the two experiments are listed below.

- 1085 • *Tracker* – the CMS tracker has a better  $p_T$  resolution (mainly due to the higher magnetic field):  
 1086  $\sigma_{p_T}/p_T \approx 5 \cdot 10^{-4} p_T + 0.01$  at ATLAS;  $\sigma_{p_T}/p_T \approx 1.5 \cdot 10^{-4} p_T + 0.005$  at CMS.
- 1087 • *Electromagnetic calorimeter* – the CMS electromagnetic calorimeter is completely enclosed  
 1088 inside the solenoid, whilst ATLAS calorimeter is outside of the solenoid. The particles going  
 1089 through the solenoid suffer an energy loss and a consequent deterioration of the energy reso-  
 1090 lution. The CMS ECAL has an energy resolution of  $\sigma_E/E \approx 3\%/\sqrt{E}$ ; the ATLAS calorimeter  
 1091 has a sandwich structure (liquid argon and lead layers) and a resolution of  $\sigma_E/E \approx 10\%/\sqrt{E}$ .
- 1092 • *Hadronic calorimeter* – the CMS HCAL is partly inside the solenoid, partly outside, depauper-  
 1093 ating the resolution. The ATLAS hadronic calorimeter (made of iron and plastic scintillator  
 1094 tiles) has an energy resolution  $\sigma_E/E \approx 50\%/\sqrt{E} + 0.03$  GeV; CMS HCAL has a resolution of  
 1095  $\sigma_E/E \approx 100\%/\sqrt{E} + 0.05$  GeV.
- 1096 • *Muon system* – the peculiar geometry of the ATLAS muon system allows a better resolution of  
 1097 the standalone measurement of the muon momenta (*i.e.*, without using tracker and calorime-  
 1098 ters), that is around 10% at 1 TeV. CMS has better performance when combining the informa-  
 1099 tions coming from the inner detectors (7% at 1 TeV against the 35% for the standalone mea-  
 1100 surement).

#### **3.3.2 ALICE**

ALICE (A Large Ion Collider Experiment) [70] studies the heavy ion collisions (lead-lead) or proton-ion in order to explore the physics of the hadrons in high density (or temperature) regimes, when a new state of matter appears, the so-called quark-gluon plasma (QGP). The QGP played a crucial role in the very first instants of the life of the universe.

#### **3.3.3 LHCb**

LHCb (Large Hadron Collider beauty) [71] is a detector designed to study the b-quark properties, in particular the CP violation and other rare phenomena related to B hadrons. The final aim of these measurements is trying to solve the matter-antimatter asymmetry problem.

The three detectors are depicted in fig. 3.21.

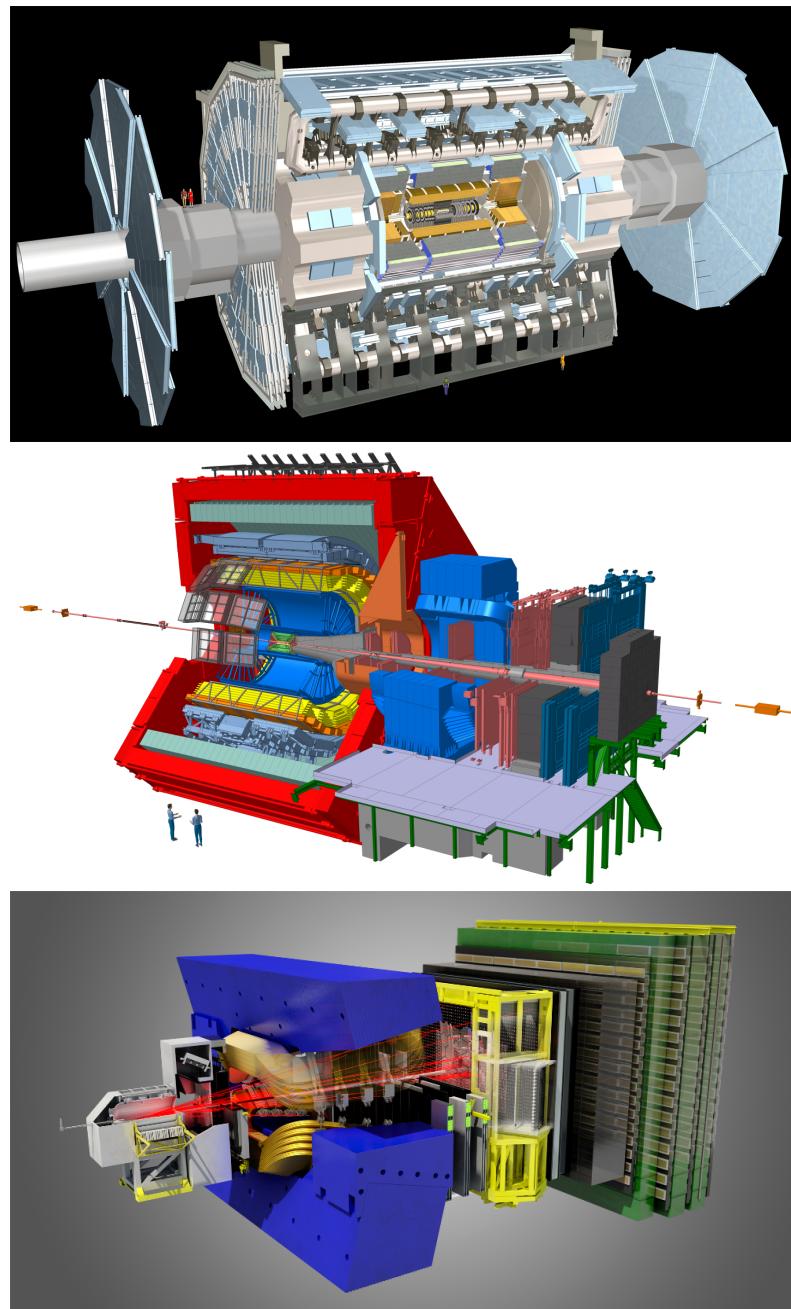


Figure 3.21: Top: the ATLAS experiment. Center: the ALICE experiment. Bottom: the LHCb experiment.

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

1112

1113

## 1114 4.1 Analysis overview

1115 This analysis searches for potential signals of heavy resonances decaying into a pair of vector bosons,  
 1116 using the data collected by the CMS experiment during 2016, corresponding to an integrated lumi-  
 1117 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-  
 1118 cay into a couple of neutrinos ( $\nu\bar{\nu}$ ), while the other electroweak boson, labelled as  $V$  and consisting  
 1119 either in a  $W$  or in a  $Z$  boson, is required to decay hadronically into a pair of quarks ( $q\bar{q}$ ). The decay  
 1120 products (the bosons) of heavy (around the TeV scale) resonances are produced with large Lorentz  
 1121 boosts; as a consequence, the decay products of the bosons (quarks and neutrinos) are expected to  
 1122 be highly energetic and collimated. In this regime, the standard jet reconstruction algorithms fail in  
 1123 distinguishing the two jets from the quarks, suggesting to look for a signature composed of a large-  
 1124 cone high- $p_T$  jet, in which both  $q$  and  $\bar{q}$  lie, recoiling against a large amount of missing transverse  
 1125 momentum ( $\vec{p}_T^{\text{miss}}$ ) due to the neutrinos escaping the detector. The hadronically decaying boson  
 1126 ( $Z$ ,  $W$ ) is then reconstructed as one large-cone jet, whose mass is used to define the signal region  
 1127 and signal-depleted control regions, the sidebands. Jet substructure techniques are exploited in or-  
 1128 der to suppress background contamination and to classify the events in two exclusive signal purity  
 1129 categories, allowing to improve the discovery reach.

1130 A general  $ZZ$  decay, predicted by the bulk graviton model (sec. 2.3.2), can be reconstructed both in  
 1131 final states with high signal purity but limited statistics (four charged leptons) and large statistics  
 1132 but overwhelming backgrounds (no charged leptons). The choice to look for one boson decaying  
 1133 hadronically and the other  $Z$  into neutrinos represents the best compromise between these two ex-  
 1134 tremes. This topology can be also utilized to reconstruct a charged spin-1 vector boson  $W'$  decaying  
 1135 into an invisible  $Z$  and an hadronic  $W$ , predicted by the HVT model (sec. 2.2), making this analysis  
 1136 sensitive to a generic  $VZ$  final state.

1137 Signal events are collected with trigger paths requiring high  $\vec{p}_T^{\text{miss}}$  recoiling against jet activity. This  
 1138 signature is clearly a very challenging one in an environment with more than 50 primary collisions  
 1139 per bunch crossing. For this reason, the Particle-Flow algorithm is run at trigger level to obtain the  
 1140 highest possible resolution on the jets and thus on the  $\vec{p}_T^{\text{miss}}$ .

1141 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  $(r, \varphi)$  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 in 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[72] matrix element generator, while hadronization and fragmentation are handled by PYTHIA 8 [73] version 8.2121 with CUETP8M1 [74] tuning. A full detector simulation and event reconstruction has

## 4.2 Data and Monte Carlo simulations samples

---

been performed with GEANT4 [75] 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 ( $\mathcal{B}(Z \rightarrow \nu\nu) = 0.20$ ,  $\mathcal{B}(Z \rightarrow qq) = 0.6991$ ). 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 ( $\mathcal{B}(Z \rightarrow \nu\nu) = 0.2$ ,  $\mathcal{B}(W \rightarrow qq) = 0.6760$ ).

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.

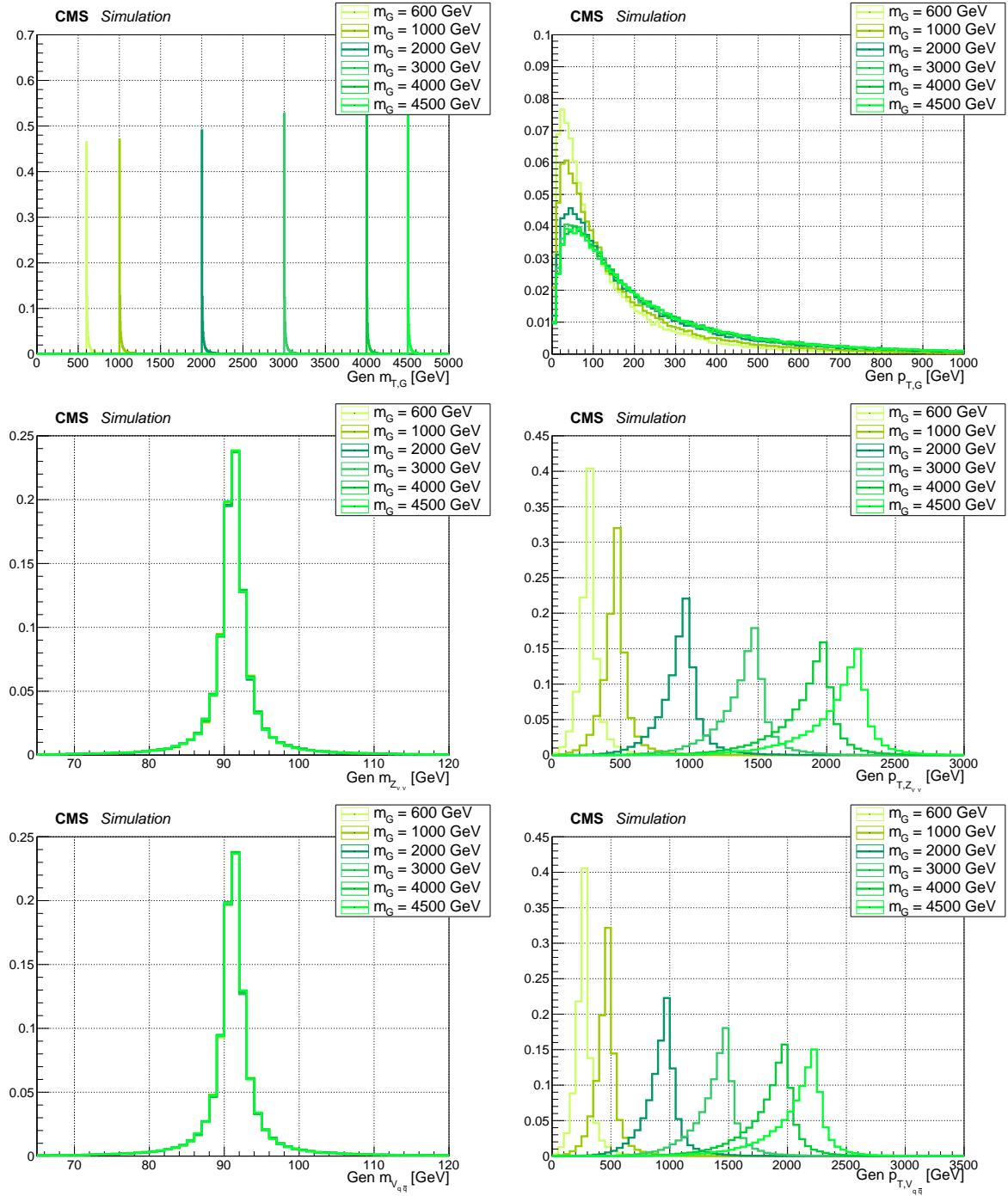


Figure 4.2: Main signal kinematic quantities at generation level after parton showering, for spin-2 bulk graviton signal, considering different mass hypotheses ( $m_G = 0.6, 1, 2, 3, 4, 4.5 \text{ TeV}$ ). Top: graviton transverse mass and  $p_T$  distributions. Center: invisibly decaying  $Z$  mass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

## 4.2 Data and Monte Carlo simulations samples

---

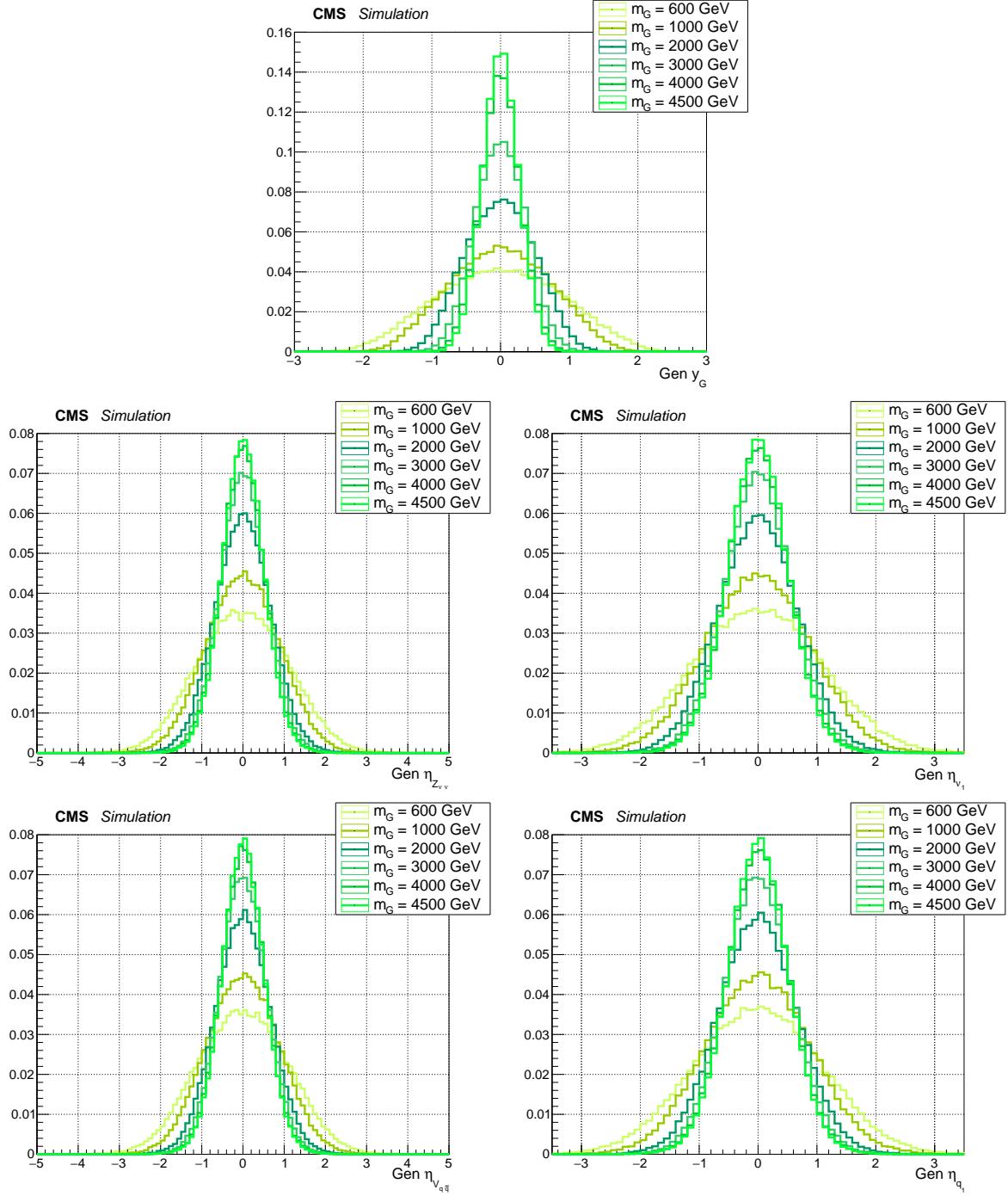


Figure 4.3: Main signal kinematic quantities at generation level after parton showering, for spin-2 bulk graviton signal, considering different mass hypotheses ( $m_G = 0.6, 1, 2, 3, 4, 4.5 \text{ TeV}$ ). Top: graviton rapidity  $\mathcal{Y}$ . Center: pseudorapidity  $\eta$  of the invisibly decaying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

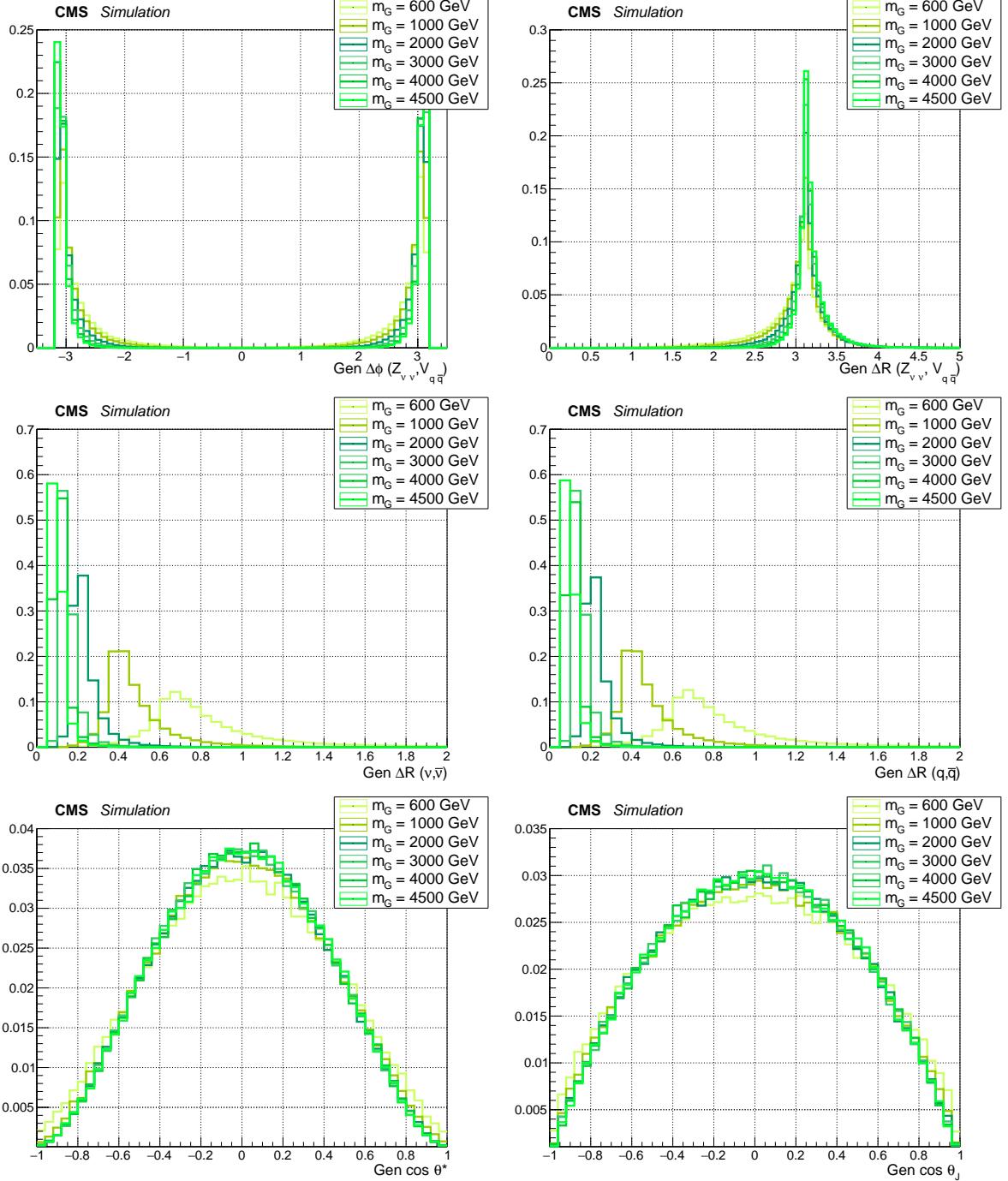


Figure 4.4: Main signal kinematic quantities at generation level after parton showering, for spin-2 bulk graviton signal, considering different mass hypotheses ( $m_G = 0.6, 1, 2, 3, 4, 4.5$  TeV). Top: angular separation in the transverse plane  $\Delta\varphi$  (left) and solid angle  $\Delta R$  (right) 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

---

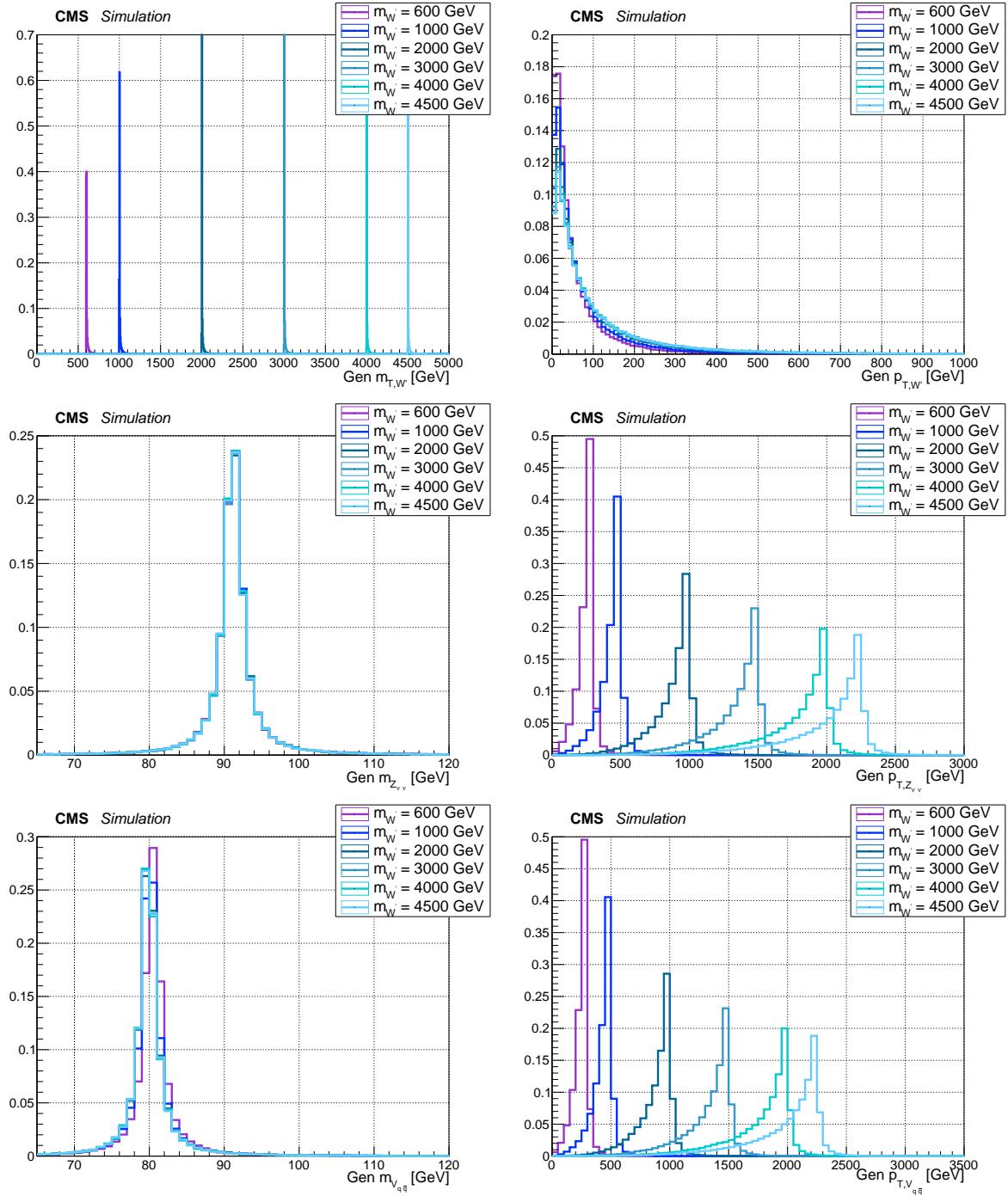


Figure 4.5: Main signal kinematic quantities at generation level after parton showering, for spin-1  $W'$  signal, considering different mass hypotheses ( $m_{W'} = 0.6, 1, 2, 3, 4, 4.5 \text{ TeV}$ ). Top:  $W'$  transverse mass and  $p_T$  distributions. Center: invisibly decaying  $Z$  mass and  $p_T$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

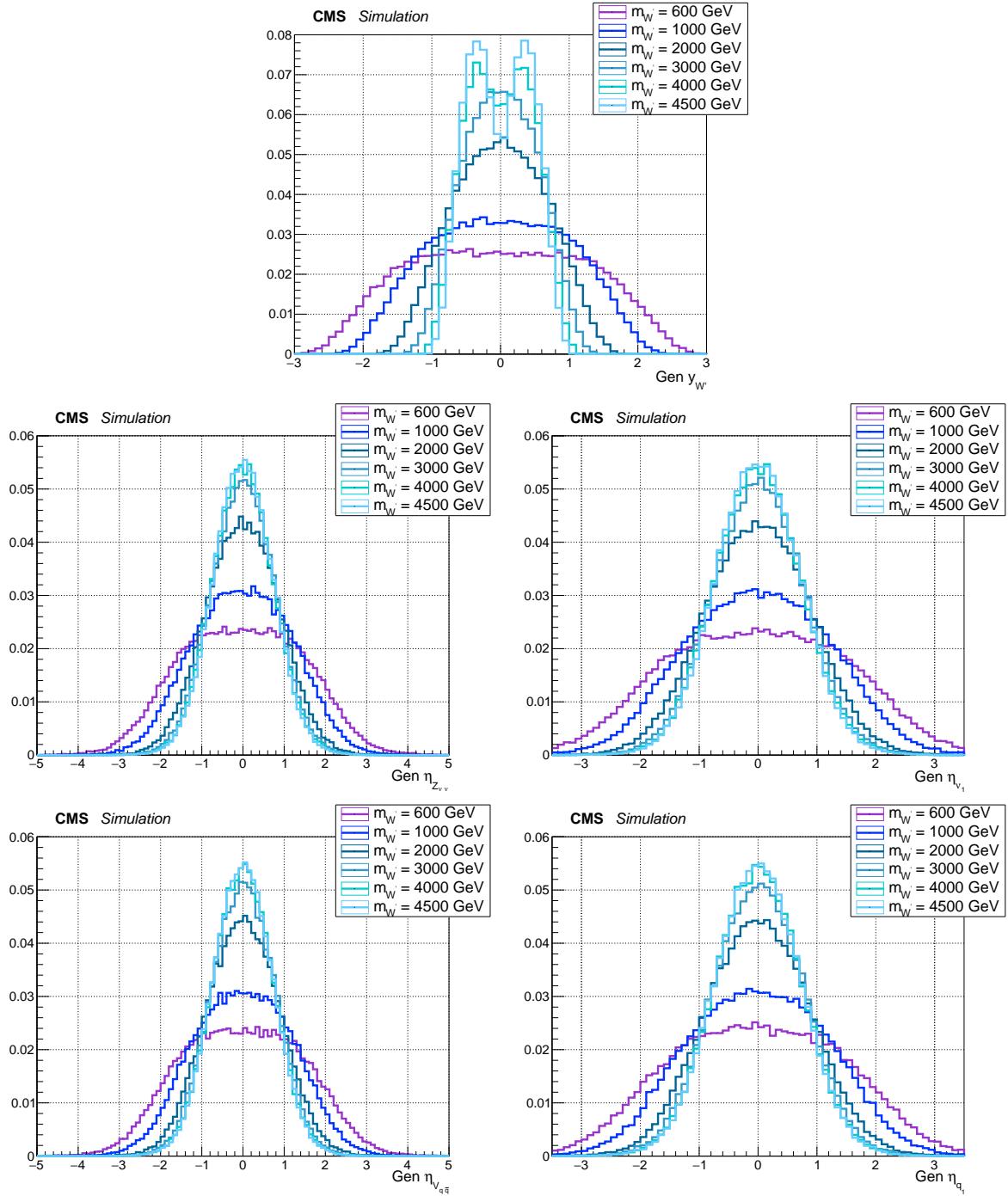


Figure 4.6: Main signal kinematic quantities at generation level after parton showering, for spin-1  $W'$  signal, considering different mass hypotheses ( $m_{W'} = 0.6, 1, 2, 3, 4, 4.5$  TeV). Top:  $W'$  rapidity  $\mathcal{Y}$ . Center: pseudorapidity  $\eta$  of the invisibly decaying  $Z$ , and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

## 4.2 Data and Monte Carlo simulations samples

---

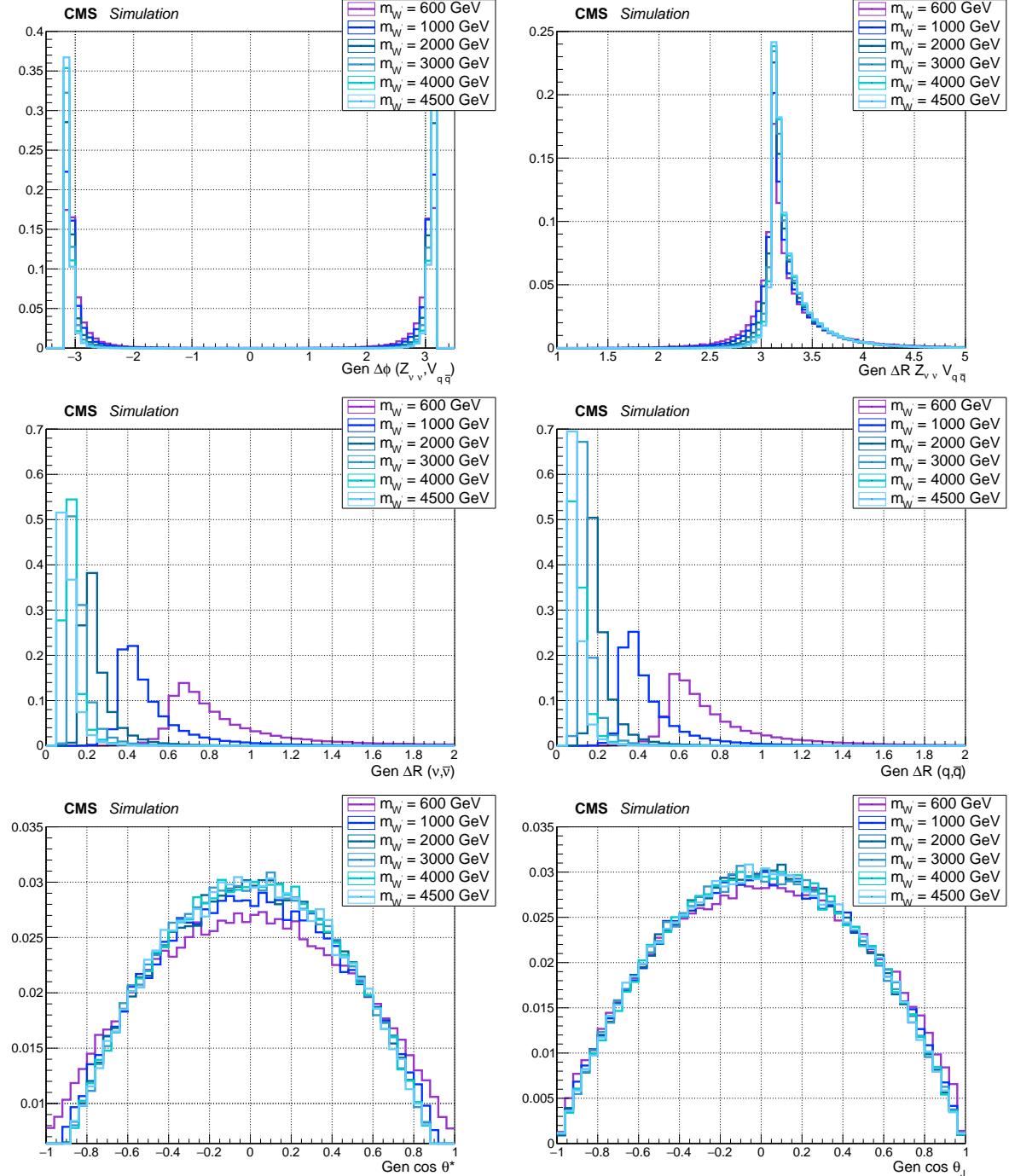


Figure 4.7: Main signal kinematic quantities at generation level after parton showering, for spin-1  $W'$  signal, considering different mass hypotheses ( $m_{W'} = 0.6, 1, 2, 3, 4, 4.5 \text{ TeV}$ ). Top: angular separation in the transverse plane  $\Delta\varphi$  (left) and solid angle  $\Delta R$  (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).

1180 Angular distributions are related to the spin, the polarization and the kinematics of the produced  
 1181 resonance; in particular:

- 1182 • the  $\Delta R$  among neutrinos and quarks reflect the boosted nature of the electroweak bosons:  
 1183 the more massive the resonance, the larger the boost, and hence the closer the fermions. By  
 1184 looking at fig. 4.4-4.7, with a jet clustering parameter of 0.8 (AK8 jet) it is possible to enclose  
 1185 the quarks produced by the decay of the  $V$  boson, for a resonance mass over 1 TeV;
- 1186 • the  $\cos \theta^*$ , namely the cosine of the angle between the momentum of the  $V$  boson, calculated  
 1187 in the resonance rest frame, and the flight direction of the resonance itself in the laboratory  
 1188 frame. This variable depends on the spin of the diboson resonance (spin-2 and spin-1 distri-  
 1189 butions are different, fig. 4.4-4.7).
- 1190 • the  $\cos \theta_J$ , the cosine of the angle between the momentum of the leading quark, calculated in  
 1191 the  $V$  rest frame, and the flight direction of the  $V$  boson in the laboratory frame. This variable  
 1192 depends on the polarization state of the decay bosons [76]; in both HVT and bulk graviton  
 1193 model, electroweak bosons are expected to be longitudinally polarized. When  $\cos \theta_J \rightarrow 0$ ,  
 1194 quarks are produced very close in angle and hence it is difficult to disentangle the two sub-  
 1195 structures in the large-cone jet (sec. 4.3.8); when  $\cos \theta_J \rightarrow \pi$  the quarks are emitted asymmet-  
 1196 rically (one is softer than the other).

### 1197 4.2.3 Background samples

1198 The physics processes yielding final states with two neutrinos in association with a pair of quarks  
 1199 are considered as sources of background; they are listed in tab. 4.3, along with the expected cross-  
 1200 sections at next-to-leading order (NLO) or next-to-next-to leading (NNLO). A summary of the stan-  
 1201 dard model cross-sections, measured by CMS, and their theoretical predictions is included in fig. 4.8-4.9.

- 1202 •  **$Z + \text{jets}$** : this process represents the main irreducible background for the signal. The produc-  
 1203 tion of a  $Z$  boson in association with one or more partons in the final state has topology that  
 1204 is similar to the signal. This  $Z + \text{jets}$  background is produced in samples binned in  $p_T$  of the  
 1205  $Z$  boson, starting from 100 GeV, with the AMC@NLO generator, with FXFX merging [77]. The  
 1206 contribution from events with  $p_T < 100$  GeV is negligible after the requirement on the  $\vec{p}_T^{\text{miss}}$   
 1207 to be greater than 200 GeV (sec. 4.3.12).
- 1208 •  **$W + \text{jets}$** : the leptonic decay of a  $W$  boson can be an irreducible background if the charged  
 1209 lepton escapes undetected (*i.e.* outside the detector acceptance) or fails the lepton identifi-  
 1210 cation requirements. The production of a  $W$  boson has a cross section larger by an order  
 1211 of magnitude with respect to the  $Z$ , and this makes the  $W + \text{jets}$  a relevant background also  
 1212 when a lepton veto is applied. This  $W + \text{jets}$  background is produced in samples binned in  
 1213  $p_T$  of the  $W$  boson, starting from 100 GeV, with the AMC@NLO generator.
- 1214 • **Top**: the pair and single production of top quarks represent a source of background, given  
 1215 the decay chain  $t \rightarrow b W$ , always including an electroweak  $W$  boson and a b-quark. The  $t \bar{t}$   
 1216 pair production manifests as two b-jets and two  $W$  bosons in the final state, that can decay to  
 1217 leptons that escape the detector or fail to be identified as leptons. This analysis makes use of  
 1218  $t \bar{t}$  inclusive decays samples based on POWHEG v2 [78] NLO generator. Single-top and single-  
 1219 antitop samples are produced in the 5-flavours scheme using POWHEG v2 [79] NLO generator.  
 1220 Different production mechanisms are considered:  $t W$  channel, when a top quark is produced  
 1221 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

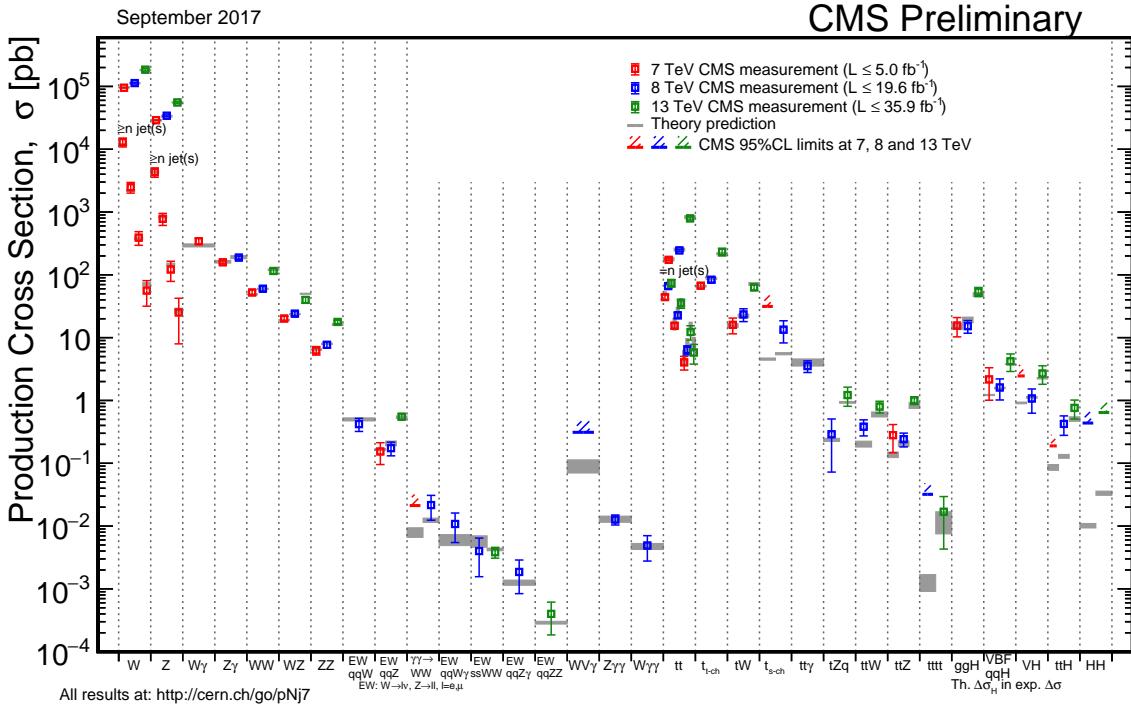


Figure 4.8: Production cross-sections of the main standard model processes, as measured by CMS, and theoretical predictions.

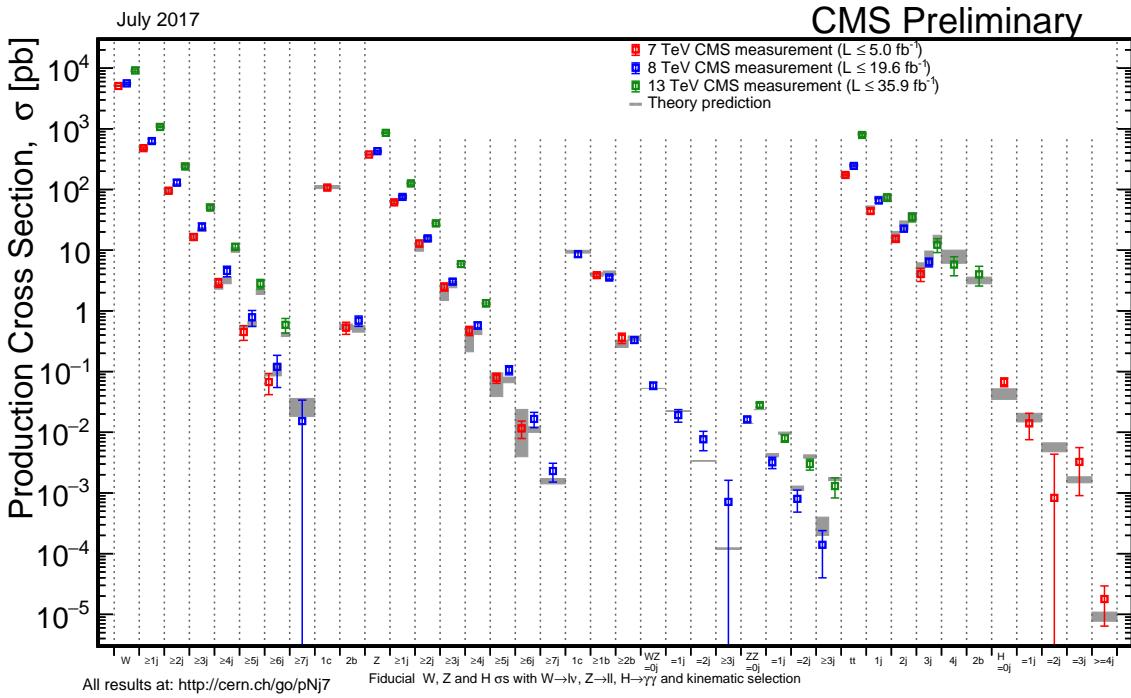


Figure 4.9: Production cross-sections of the standard model processes involving a vector boson in association with jets, as measured by CMS, and theoretical predictions. These phenomena represent the main background sources for the analysis.

1222 quark-antiquark scattering, producing a top and an anti-bottom quark in the final state; t-  
 1223 channel, via a virtual  $W$  in a quark-b-quark scattering, resulting in a top quark and a quark  
 1224 jet in the final state.

- 1225 • **Diboson:** the SM production of a pair of vector bosons is topologically close to the searched  
 1226 signal, by the way the cross-section of the process is low. The  $WW$  production is the most  
 1227 probable process, that imitates the signal when one of the  $W$  decays leptonically and the  
 1228 charged lepton fall outside the detector acceptance or it is mis-identified; the  $WZ$  and  $ZZ$   
 1229 processes have smaller cross-sections but are topologically identical to the signal, except for  
 1230 the fact that the invariant mass of the diboson system has a smoothly falling spectrum, in  
 1231 contrast to the resonant signal distribution. Inclusive diboson production processes ( $WW$ ,  
 1232  $WZ$ ,  $ZZ$ ) are considered; they are simulated at LO by PYTHIA generator.
- 1233 • **Multi-jet:** despite the production of multi-jet events, initiated by the strong interactions, has  
 1234 a very large cross-section, this source of background is suppressed by a dedicated selection  
 1235 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

#### 1236 4.2.4 Vector boson momentum corrections

1237 Corrections to the  $p_T$  spectrum of the  $V$  boson, due to NLO electroweak contributions, are en-  
 1238 hanced at TeV scale [80], 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.

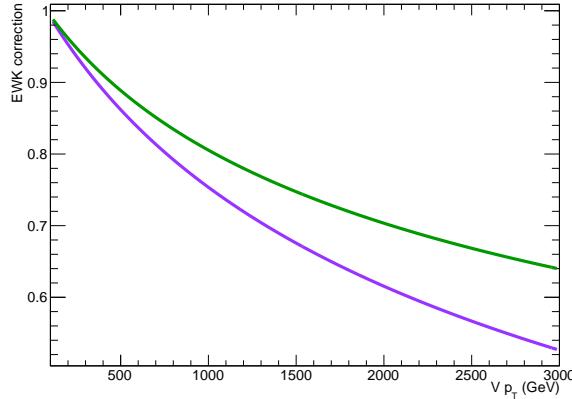


Figure 4.10: Electroweak corrections for the  $Z$  (green line) and  $W$  boson (purple line) as a function of the transverse momentum of the boson [80].

### 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  $\vec{p}_T^{\text{miss}}$  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  $35.9 \pm 0.9 \text{ fb}^{-1}$  [81]. In order to remove problematic or noise-dominated events, dedicated  $E_T^{\text{miss}}$  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  $\vec{p}_T^{\text{miss}}$  trigger HLT paths listed in tab. 4.4, along with their corresponding L1 missing energy or jet seeds. PFMETNoMu indicates the  $E_T^{\text{miss}}$  (no  $\mu$ ) 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  $H_T^{\text{miss}}$  (no  $\mu$ ), 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  $E_T^{\text{miss}}$  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

1264 at HLT (cleaning events from noise in the detector). Different thresholds are applied to  $E_T^{\text{miss}}$  (no  $\mu$ )  
 1265 and  $H_T^{\text{miss}}$  (no  $\mu$ ).

1266 The approach adopted in this analysis consists in calculating the trigger efficiency on data, and  
 1267 applying the measured efficiency to Monte Carlo samples. Therefore the trigger is not required to  
 1268 have been fired in MC.

1269 Given that the final state probed by the analysis consists into one AK8 jet, large  $E_T^{\text{miss}}$  and no charged  
 1270 leptons, an unbiased measurement of the  $E_T^{\text{miss}}$  trigger efficiency can be performed in an orthogo-  
 1271 nal dataset, collected with different triggers, and requiring events where a  $W \rightarrow \ell \nu$  leptonic decay  
 1272 is taking place. This guarantees the presence of real  $\vec{p}_T^{\text{miss}}$  in the event, due to the neutrino; fur-  
 1273 thermore, the presence of a charged lepton guarantees that the leptonic  $W$ -like events are not over-  
 1274 lapped with the search region. The additional requirement to have at least one AK8 jet is applied in  
 1275 the trigger measurement, in order to probe a kinematical region similar to that of the signal region  
 1276 of the analysis.

1277 The efficiency of the  $E_T^{\text{miss}}$  triggers is measured on SingleMuon dataset by selecting  $W \rightarrow \mu \nu$  events  
 1278 using a logic or of single muon triggers HLT\_IsoMu24 OR HLT\_IsoTkMu24\_v, namely, triggers  
 1279 asking for a PF muon reconstructed at HLT, with a  $p_T$  threshold of 24 GeV, that is isolated (in the  
 1280 whole reconstruction or at tracker level only). Offline selections consist in asking to have one iso-  
 1281 lated muon, with a suitable  $p_T$  threshold to be in the plateau of the muon trigger. The efficiency has  
 1282 been calculated as a function of the minimum quantity between the offline reconstructed  $E_T^{\text{miss}}$  (no  
 1283  $\mu$ ):

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

1284 where the contribution of all the offline PF muons is removed from the  $\vec{p}_T^{\text{miss}}$  computation as in the  
 1285 online algorithm, and the offline  $H_T^{\text{miss}}$ , defined as

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

1286 This approach guarantees to mimic the behaviour of the online L1 trigger seeds. The detailed se-  
 1287 lections are listed below:

- 1288 • HLT\_IsoMu24\_v OR HLT\_IsoTkMu24\_v,

## 4.2 Data and Monte Carlo simulations samples

---

- 1289     • 1 isolated muon  $p_T > 35$  GeV, identified with tight requirements,
- 1290     • at least one AK8 jet,  $p_T > 170$  GeV,  $|\eta| < 2.5$ , identified with loose requirements,
- 1291     • AK4 jets included in  $H_T^{\text{miss}}$ :  $p_T > 30$  GeV,  $|\eta| < 2.5$ , identified with loose requirements.

1292 The efficiency of the  $E_T^{\text{miss}}$  triggers has independently been measured also on SingleElectron  
1293 dataset, by selecting  $W \rightarrow e \nu$  events using a single electron trigger (HLT\_Ele27\_WP Loose\_Gsf  
1294 OR HLT\_Ele27\_WP Tight\_Gsf OR HLT\_Ele32\_WP Tight\_Gsf), asking to have one well identified  
1295 electron, with a suitable  $p_T$  threshold, and asking to the electron and the  $\vec{p}_T^{\text{miss}}$  to be separated in  
1296 the transverse plane (hence, in  $\varphi$ ) in order to suppress fake jet events mis-identified as electrons at  
1297 trigger level ( $\Delta\varphi > 0.5$ ). The detailed selections are listed below:

- 1298     • HLT\_Ele27\_WP Loose\_Gsf OR HLT\_Ele27\_WP Tight\_Gsf OR HLT\_Ele32\_WP Tight\_Gsf,
- 1299     • 1 electron,  $p_T > 35$  GeV, identified with tight requirements,
- 1300     • at least one AK8 jet,  $p_T > 170$  GeV,  $|\eta| < 2.5$ , identified with loose requirements,
- 1301     • AK4 jets included in  $H_T^{\text{miss}}$ :  $p_T > 30$  GeV,  $|\eta| < 2.5$ , identified with loose requirements.

1302 All the available data have been employed to derive the efficiency. The final turn-on curves for the  
1303  $E_T^{\text{miss}}$  triggers are shown in fig.4.11-4.12, measured in muon and electron dataset respectively. The  
1304 PF MET No Mu trigger efficiencies are displayed separately, together with their logic OR. The trigger  
1305 efficiency measured on SingleMuon dataset amounts to 96% at  $E_T^{\text{miss}} = 200$  GeV; the trigger effi-  
1306 ciency measured on SingleElectron dataset amounts to 95% at  $E_T^{\text{miss}} = 200$  GeV. The difference  
1307 needed to cover the gap between the two independent measurements is taken as trigger systematic  
1308 uncertainty, and it amounts to 1% at 200 GeV.

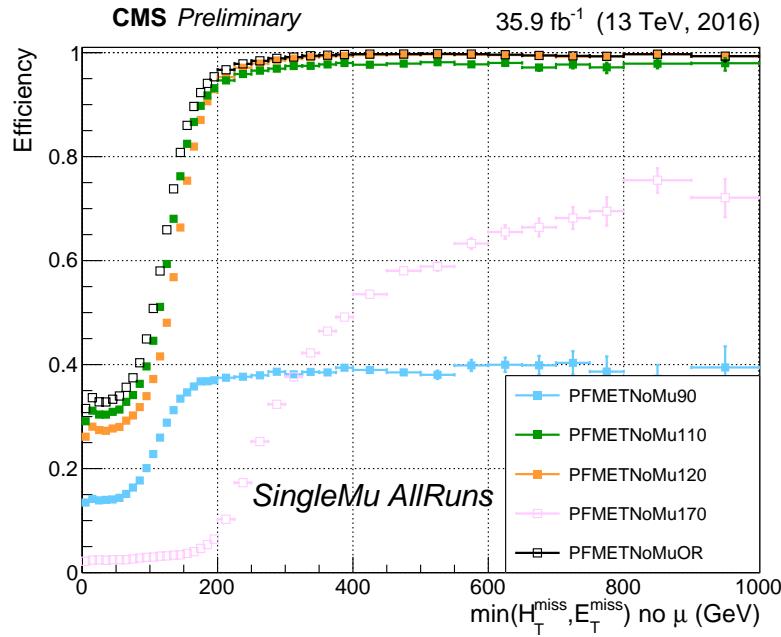


Figure 4.11:  $\vec{p}_T^{\text{miss}}$  trigger efficiency for the  $\vec{p}_T^{\text{miss}}$  trigger paths used in this analysis, calculated on SingleMuon dataset, as a function of the minimum of the variables  $E_T^{\text{miss}}$  (no  $\mu$ ) (eq. 4.1) and  $H_T^{\text{miss}}$  (eq. 4.2).

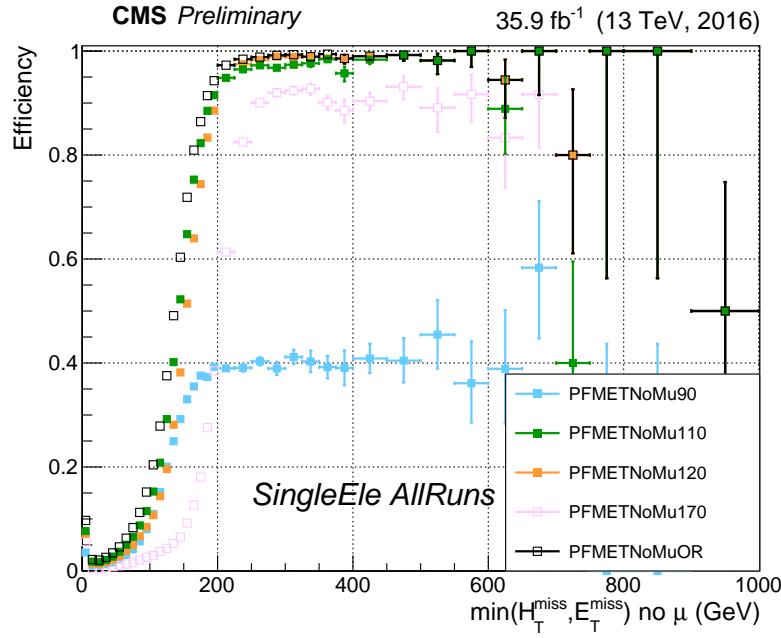


Figure 4.12:  $\vec{p}_T^{\text{miss}}$  trigger efficiency for the  $\vec{p}_T^{\text{miss}}$  trigger paths used in this analysis, calculated on SingleElectron dataset, as a function of the minimum of the variables  $E_T^{\text{miss}}$  (no  $\mu$ ) (eq. 4.1) and  $H_T^{\text{miss}}$  (eq. 4.2).

## 1309 4.3 Event selection

1310 In this section, the selections applied to the physics objects used in the analysis are presented and  
1311 motivated by performance and validation plots. Background events are represented as coloured  
1312 histograms:  $Z + \text{jets}$  events in light blue,  $W + \text{jets}$  events in violet,  $t\bar{t}$  events in yellow, single-top  
1313 events in orange, diboson (or  $VV$ ) events in blue, multi-jet (QCD) events in gray. Background un-  
1314 certainties are displayed as black shaded areas. Signal samples are represented as coloured shaded  
1315 histograms: the kind of signal (graviton or  $W'$ ), the mass and cross-section of the considered reso-  
1316 nance are reported in the legend. Data are represented with black markers, with their corresponding  
1317 Poissonian uncertainty bars. If data are displayed, the data-MC ratio is reported per each bin in the  
1318 bottom panel, along with the overall data-MC ratio calculated in the whole spectrum and the scores  
1319 of  $\chi^2$  and Kolmogorov-Smirnov goodness-of-fit tests.

### 1320 4.3.1 Vertex and Pile-up

1321 Due to the pile-up effect, several vertices are typically reconstructed in one event. The primary ver-  
1322 tex of the event is defined as the one with the highest sum of transverse momenta  $\sum p_T^2$  of clustered  
1323 physics objects associated to it, which passes the following selections:

- 1324 • number of degrees of freedom  $N_{DoF} > 4$
- 1325 • vertex position along the beampipe  $|z_{vtx}| < 24\text{cm}$
- 1326 • vertex distance with respect the beam pipe  $d_0 < 2\text{cm}$

1327 where  $z_{vtx}$  and  $d_0$  are the distance along and perpendicular to the beam line of the vertex with  
1328 respect the nominal interaction point  $(0, 0, 0)$ .

1329 The Monte Carlo samples listed in sec. 4.2 are generated simulating the pile-up conditions, as ex-  
1330 pected in the 25 ns bunch crossing pile-up scenario. Nevertheless, the MC pile-up description does  
1331 not match exactly the conditions in data, and there is therefore the need to reweight the simulated  
1332 events in order to improve the agreement with the data.

1333 The MC samples are reweighted assuming a total inelastic cross section of  $\sigma_{in} = 69\,200\mu b$ . The  
1334 comparison between the distributions of primary vertices in data and MC after the pile-up reweight-  
1335 ing is applied is shown in fig. 4.13 for an event selection (called inclusive selection, described in  
1336 sec. 4.4) requiring large amount of  $\vec{p}_T^{\text{miss}}$  recoiling against an AK8 fat jet (tab. 4.12).

### 1337 4.3.2 Electrons

1338 Electrons considered in this analysis, reconstructed from energy deposits in the ECAL matched to  
1339 tracks reconstructed in the silicon tracker, are required to pass the Particle-Flow criteria, and to fall  
1340 in the ECAL pseudorapidity fiducial range ( $|\eta| < 2.5$ ). The electron identification is defined with a  
1341 “cut-based” approach. In the isolation definition, the effect of pile-up is considered by taking into  
1342 account the energy deposits in the calorimeter, estimated through the so-called  $\rho$ -area method, by  
1343 subtracting the median energy density in the event  $\rho$  multiplied by the electron energy deposits  
1344 effective area. The isolation value is computed in a  $\Delta R$  cone of 0.3 centered along the lepton direc-  
1345 tion.

1346 Since in this analysis aims at a final state without any lepton, every electron identified with the looser  
1347 cut-based criteria (*veto Id*) and transverse momentum  $p_T > 10\text{ GeV}$  is vetoed. The detailed set of cuts  
1348 to define a *veto Id* cut-based electron are reported in tab. 4.5; this set of selections allow to identify  
1349 an electron with an efficiency of  $\sim 95\%$ . The supercluster width is indicated as  $\sigma_{in\eta}$ ;  $\Delta\eta_{in}^{seed}$  and  
1350  $\Delta\varphi_{in}$  are the difference in  $\eta$  and  $\varphi$  between the track position as it is measured in the inner layer,

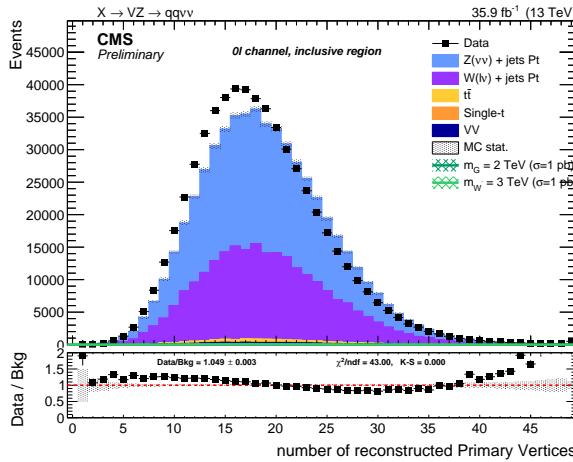


Figure 4.13: Primary vertices distributions in data and MC samples, after reweighting.

and then extrapolated to the interaction vertex and to the calorimeter, and the  $\eta$  of the seed cluster or the  $\varphi$  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;  $rellIso$  indicates the relative isolation calculated 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
$\sigma_{i\eta i\eta}$	< 0.0115	0.037
$\Delta\eta_{in}^{\text{seed}}$	< 0.00749	0.00895
$\Delta\varphi_{in}$	< 0.228	0.213
$H/E$	< 0.356	0.211
$rellIso$ (Effective Area)	< 0.175	0.159
$ 1/E - 1/p $	< 0.299	0.15
$ d_0 $	< 0.05	0.10
$ d_z $	< 0.10	0.20
missing hits	$\leq$ 2	3
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  $\rho$ -area method for the mitigation of the pile-up) and conversion-safe veto are applied. The isolation value is computed in a  $\Delta R$  cone of 0.3 and it is corrected for pile-up by subtracting the event-by-event energy density ( $\rho$ ) 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.( $\rho$ -corr)	< 1.295	1.011
PF neu.had.iso.( $\rho$ -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.( $\rho$ -corr)	< $3.630 + 0.0047 p_T$	$6.641 + 0.0034 p_T$
conversion veto	yes	yes

1365 **4.3.4 Muons**

1366 The minimal criteria to define a muon is that it must be identified by the Particle-Flow algorithm,  
 1367 and should be reconstructed either as a global muon or as a tracker muon (sec. 3.2.9). The muon  
 1368 isolation is defined in a cone with a radius of  $\Delta R = 0.4$  centered along the lepton direction. In the  
 1369 analysis event selection, all muons identified with the loosest criteria previously described,  $p_T$  over  
 1370 10 GeV, PF isolation below 0.25,  $\eta < |2.4|$  are vetoed.

1371 **4.3.5 Taus**

1372 The presence of hadronically decaying taus acts as a veto for the events both in the signal and in  
 1373 the control regions, in order to suppress electroweak backgrounds. The selection criteria for taus  
 1374 are  $p_T > 18$  GeV and  $|\eta| < 2.3$ . Loose identification criteria of the hadronic tau reconstruction algo-  
 1375 rithms are required and applied in order to identify possible tau candidates.

1376 **4.3.6 Jets**

1377 In this analysis, jets are considered if the corrected  $p_T$  is larger than 30 GeV for AK4 jets, and larger  
 1378 than 200 GeV for AK8 jets, and lie in the tracker acceptance ( $|\eta| < 2.4$ ). The requirement on AK8  
 1379 jets transverse momentum is motivated by the fact that  $p_T = 200$  GeV is the minimum kinematical  
 1380 threshold ensuring to enclose the lighter hadronically decaying vector boson (namely, the  $W$  bo-  
 1381 son) in the jet cone. Additionally, AK4 jets are required to pass *loose* jet identification requirements,  
 1382 AK8 are required to pass *tight* jet identification requirements defined in tab. 4.7. AK8 jets are used  
 1383 to reconstruct the hadronically decaying electroweak boson candidate, whilst AK4 jets are used to  
 1384 suppress the contribution of top and QCD background events. Jet energy corrections are applied  
 1385 to AK4 and AK8 CHS jets. Fig. 4.14- 4.16 show the data/simulation comparison after the analysis  
 1386 selections (tab. 4.12 without Top cleaning and Event cleaning).  
 1387 Since it has been measured that the jet energy resolution (JER) is not the same in data and MC,  
 1388 an additional smearing is applied in simulation, in order to get a better agreement. There are two  
 1389 independent ways to get the smearing. The scaling method rescales the corrected four-momentum  
 1390 of a reconstructed jet by a factor

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

1391 where  $p_T$  is the transverse momentum of the jet,  $p_T^{\text{gen}}$  is the transverse momentum of the genera-  
 1392 tor level particle corresponding to the reconstructed jet, and  $s_{\text{JER}}$  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

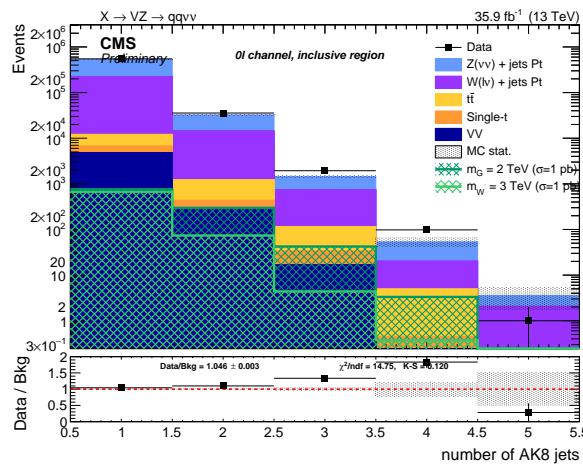
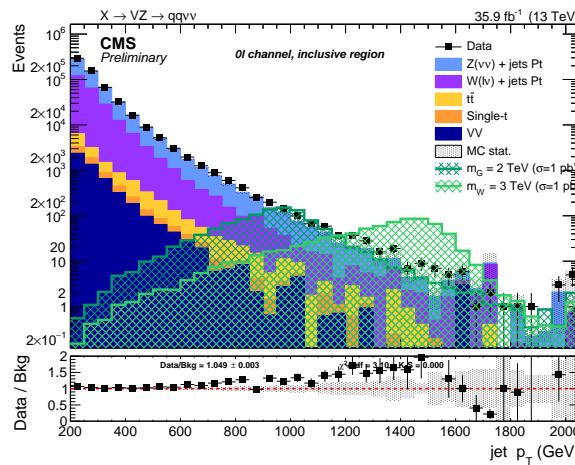


Figure 4.14: Number of reconstructed AK8 jets after inclusive selections.


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

1393 scale factor. The factor  $c_{\text{JER}}$  is positively defined, hence, when negative, it is set equal to zero. The

### 4.3 Event selection

---

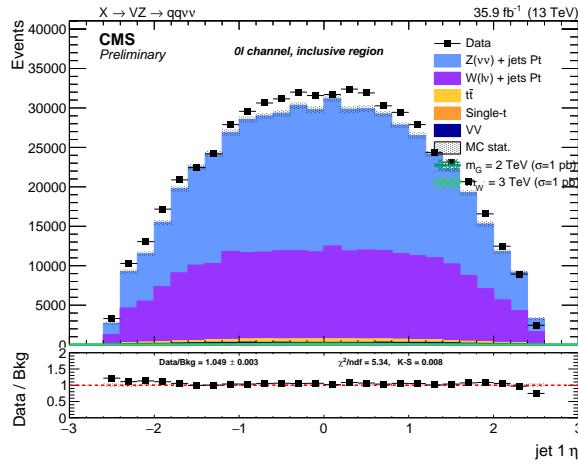


Figure 4.16: Leading AK8 jet  $\eta$  spectra after inclusive selections.

1394 generator level particle and a reconstructed jet are defined as matched if:

$$\Delta R < R_0/2, \quad |p_T - p_T^{\text{gen}}| < 3 \times \sigma_{\text{JER}} \times p_T, \quad (4.4)$$

1395 where  $R_0$  is the jet clustering parameter and  $\sigma_{\text{JER}}$  is the relative  $p_T$  resolution measured in simulation.  
1396

1397 The alternative approach is the stochastic smearing, and it does not require the matching with the  
1398 generator level particle. The jet four-momentum is rescaled by a factor

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

1399 where  $\sigma_{\text{JER}}$  is the relative  $p_T$  resolution in simulation,  $s_{\text{JER}}$  is the data-simulation scale factor, and  
1400  $\mathcal{N}(0, \sigma)$  is a random number extracted from a gaussian normal distribution, whose mean is zero  
1401 and variance  $\sigma^2$ . Scaling factor  $c_{\text{JER}}$  is positively defined.

1402 The smearing procedure adopted in this analysis is the hybrid method: when a matching jet at  
1403 generator level is found, the scaling method is adopted, else the stochastic smearing is chosen. The  
1404 smearing coefficients (scale factors, SF) as a function of the jet  $\eta$  and their uncertainties are reported  
1405 in tab. 4.8 for 2016 data [62].

#### 1406 4.3.7 Jet mass

1407 The jet mass is the main observable in distinguishing a  $V$  jet from a jet produced by colour interaction (QCD jets). Jet grooming procedure consists in the suppression of uncorrelated underlying  
1408 event, pile-up and soft radiation from the jet: it improves the signal and background discrimination,  
1409 by pushing the jet mass for QCD jets towards lower values of the spectrum, while maintaining  
1410 the jet mass for  $V$ -jets around the electroweak boson mass window.  
1411  
1412 The grooming technique of the analysis relies on the “soft drop declustering” algorithm, a jet sub-  
1413 structure technique that recursively removes soft wide-angle radiation from a jet [82], in order to  
1414 mitigate the contaminations from initial state radiation, along with pile-up and multiple scatterings.  
1415  
1416 The soft drop algorithm starts with a jet clustered with the anti- $k_T$  algorithm with a parameter  $R_0$ ;  
1417 the jet is then reclustered with the Cambridge-Aachen method [83], whose definition is included in

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

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

1418 eq. 3.10, once the exponent  $a$  is set  $a = 0$ . The soft drop algorithm is ruled by two parameters, a soft  
 1419 threshold  $z_{\text{cut}}$ , that cuts on the energy fraction of soft radiation, and an angular exponent  $\beta$ . The  
 1420 procedure is the following:

- 1421 • the jet is declustered into two subjets,  $j_1$  and  $j_2$ , by reverting the final step of Cambridge-  
 1422 Aachen algorithm;
- 1423 • if  $j_1$  and  $j_2$  respect the soft drop condition (eq. 4.6),  $j$  is defined as the groomed jet;
- 1424 • if they don't pass the condition, the leading subjet in  $p_T$  is redefined as the new  $j$ ;
- 1425 • if  $j$  can't be declustered anymore, it is defined as the groomed jet.

1426 The parameters  $z_{\text{cut}} = 0.1$  and  $\beta = 0$  are set in the soft drop condition:

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

1427 where  $p_T^1$  and  $p_T^2$  are the momenta of the constituents,  $\Delta R_{12}$  is their angular distance.  $z_{\text{cut}}$  and  $\beta$   
 1428 parameters affect the degree of jet grooming: if  $\beta \rightarrow \infty$  the jet remains ungroomed, while the more  
 1429  $\beta$  approaches zero, the more soft collinear radiation is removed.

1430 The net effect of the soft drop algorithm is studied in Monte Carlo simulations of a  $W$  hadronic  
 1431 decay process (signal), in association with jets, and of a multi-jet QCD process (background). Jets  
 1432 are clustered with the anti- $k_T$  algorithm with a parameter  $R_0 = 1$  and asked to have  $p_T > 500$  GeV  
 1433 and  $|\mathcal{Y}| < 4$ . The parameter  $z_{\text{cut}}$  is chosen such in a way that the number of events falling in the  $W$   
 1434 mass window ( $[70, 90]$  GeV) is the 35% of the total number of events. The results before (black curve)  
 1435 and after the application of soft drop algorithm (coloured curves, depending on the value of  $\beta$ ) are  
 1436 presented in fig. 4.17 [82]. In particular, by comparing the ungroomed jet mass (in black) with the  
 1437 mass groomed with a parameter  $\beta = 0$  (adopted in this analysis and displayed with a green curve),  
 1438 the soft drop mass of the leading jet is a very narrow distribution peaking around the nominal  $W$   
 1439 window in the signal sample, whilst it is pushed at lower values in the background sample.

1440 The soft drop algorithm is used in association with the Pile Up Per Particle Identification algorithm  
 1441 (PUPPI) [84], designed to combine detector informations in order to compute a local metric  $\alpha$ , that

### 4.3 Event selection

---

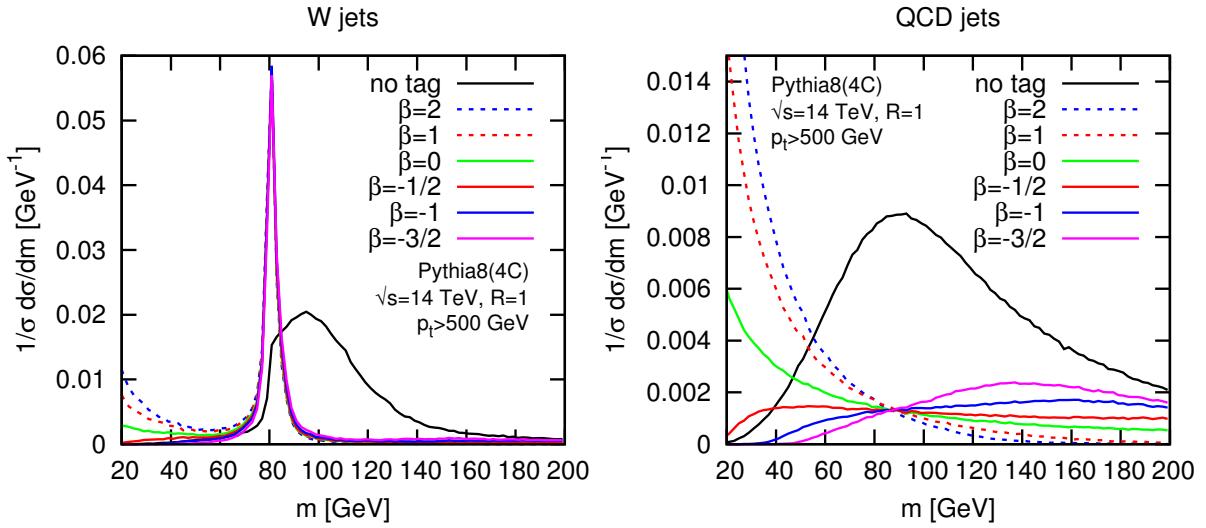


Figure 4.17: Distributions of the jet mass in  $W +$  jeta signal simulations (left) and multi-jet QCD background (right), before (in black) and after applying soft drop algorithm. Each curve corresponds to a different value of the parameter  $\beta$ . [82]

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  $\alpha$  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  $\Theta$  is the Heaviside step function,  $\Delta R_{ij}$  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,  $\alpha_i$  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  $\alpha_i$  fluctuates from the median of the pile-up local shape  $\bar{\alpha}_{PU}$  (that has a variance  $\sigma_{PU}^2$ ), and it is distributed like a  $\chi^2$  with 1 degree of freedom. The PUPPI weight is defined as the cumulative  $\chi^2$  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:

- 1461 1. a  $p_T$ -dependent correction to account for a small shift in the generated vector boson mass,  
 1462 applied only on simulated samples,
- 1463 2. a  $p_T$ - $\eta$ -dependent correction to the reconstructed jet mass, applied separately for jets in the  
 1464 barrel and endcaps regions.

1465 Fig. 4.18- 4.19 show the jet mass for hadronically decaying  $W$  or  $Z$  bosons in bulk graviton and  
 1466  $W'$  signal samples, before and after the correction, without applying any selections. In fig. 4.20,  
 1467 the distribution of soft drop PUPPI jet mass is shown for the expected backgrounds of the analysis  
 1468 (coloured histograms) and data (black markers), before and after the corrections.

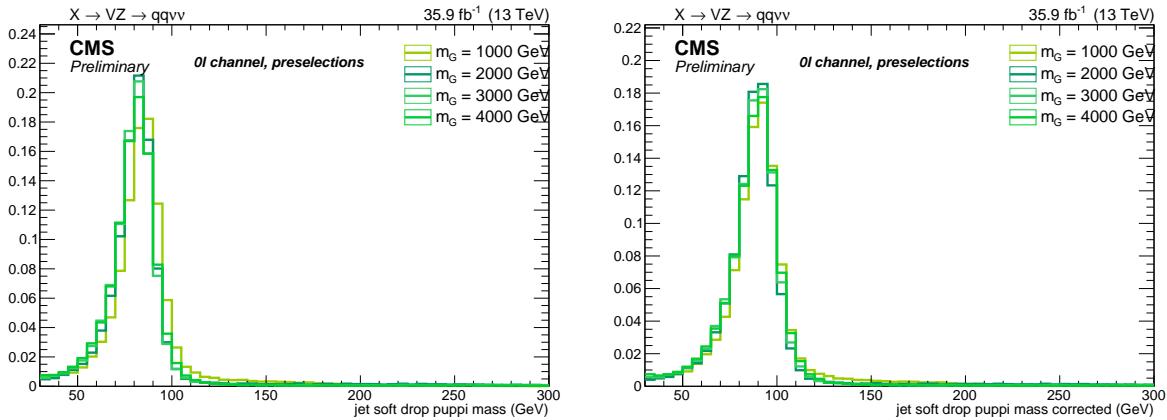


Figure 4.18: Soft drop PUPPI mass of AK8 jet reconstructed for different bulk graviton signal samples, before corrections (left) and after corrections (right).

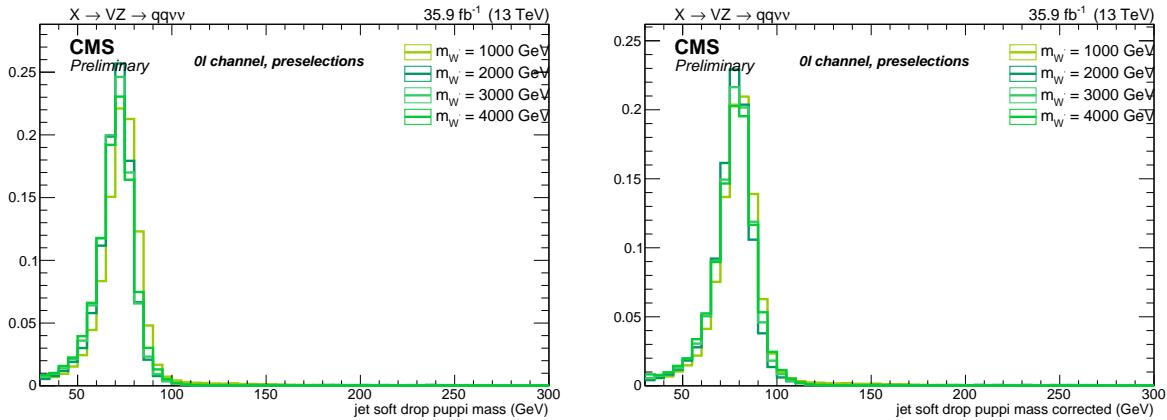


Figure 4.19: Soft drop PUPPI mass of AK8 jet reconstructed for different  $W'$  signal samples, before corrections (left) and after corrections (right).

1469 In order to obtain a better data-Monte Carlo agreement, a smearing procedure has been applied to  
 1470 the soft drop PUPPI jet mass, by using the stochastic method, with a constant smearing coefficient  
 1471 ( $1.00 \pm 0.20$ ), that does not depend on jet pseudorapidity, if it is restricted to  $|\eta| < 2.5$ .

1472 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

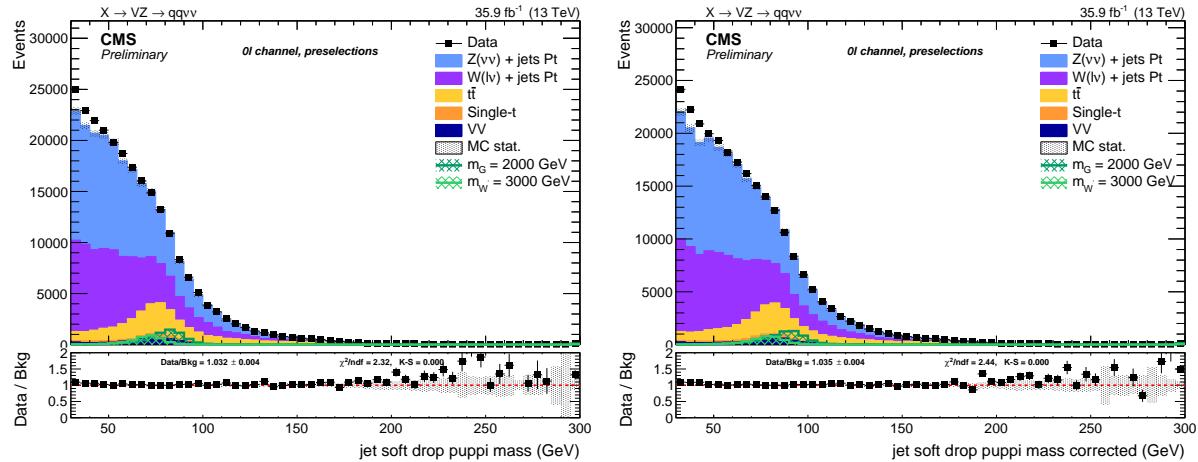


Figure 4.20: Soft drop PUPPI mass of AK8 jet; left: before corrections. right: after corrections.

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 it has to provide a sufficient data and simulation statistics for the control regions (the regions outside the mass cut). The soft drop PUPPI mass variable is used to define the following regions:

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

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 expected. Events with a jet mass value lower than 30 GeV are discarded, because of the high background contamination. The jet mass distribution of the V candidate, in the sidebands and in the signal region, is shown in fig. 4.21. If the soft drop, PUPPI corrected mass of a large-con jet falls into the V-region, the jet is defined as V-tagged.

#### 4.3.8 Jet substructure

In order to further discriminate signal from background, the inner structure of the jet is investigated. Studying the distribution of the jet constituents with respect to the jet axis allows to test the hypothesis of the existence of multiple substructures, that could be an evidence of jets originated by more than one parton. The constituents of the considered jet are clustered again with the  $k_T$  algorithm, and it is forced to return  $n$  subjets. The  $n$ -subjettiness [85],  $\tau_n$ , is defined as

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

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

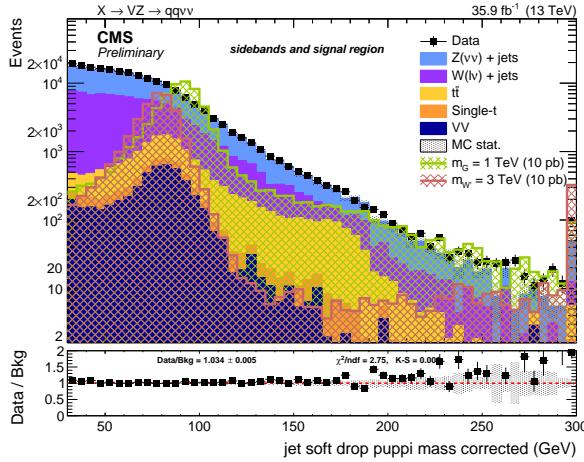


Figure 4.21: Distribution of the soft drop PUPPI corrected mass of the leading AK8 jet, selected as the hadronically decaying  $V$  candidate, in the sidebands and control region of the analysis, for expected SM background, bulk graviton signal,  $W'$  signal, and data.

1491 The parameter  $d_0$  is a normalization factor:

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

1492 where  $R_0$  is the clustering parameter of the considered jet. The  $\tau_n$  variable describes to what degree  
1493 a jet can be considered as composed by  $n$  substructures; smaller values of  $\tau_n$  correspond to higher  
1494 compatibility to the  $n$ -prong hypothesis. A large-cone jet generated by the hadronic decay of an  
1495 electroweak boson is expected to be a 2-prong object, whilst light flavour and gluon jets generated  
1496 by colour interaction have a 1-prong monolithic structure. The  $\tau_2$  or the  $\tau_1$  alone, by the way, do  
1497 not provide an optimal signal and background discrimination, as shown in fig. 4.22 (left and center);  
1498 by looking at fig. 4.22 (right), it is clear that the most powerful discriminating variable is their ratio  
1499  $\tau_{21} = \tau_2 / \tau_1$ :

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

1500 In fig. 4.23, the distributions of the  $\tau_{21}$  variable are displayed for background and data, after applying  
1501 the PUPPI algorithm (left), and for different bulk graviton mass hypotheses (right). The signal  
1502 distribution is expected to peak at low values of the  $\tau_{21}$  subjettness variable.

1503 The  $\tau_{21}$  variable is used to classify the events into two exclusive categories, in order to improve  
1504 the signal discovery reach. Events are included in either the high-purity ( $\tau_{21} < 0.35$ ) or low-purity  
1505 ( $0.35 < \tau_{21} < 0.75$ ) category.

1506 The choice of the  $\tau_{21}$  categorization listed above is based on a study of the analysis sensitivity. An-  
1507 other  $\tau_{21}$  categorization is probed, according to which events are grouped into different high-purity  
1508 ( $\tau_{21} < 0.40$ ) and low-purity ( $0.40 < \tau_{21} < 0.75$ ) categories. This different set of  $\tau_{21}$  cuts has been  
1509 tested, along with that chosen for this analysis. Two figures of merit are considered: the discovery  
1510 reach, namely the bulk graviton signal significance (displayed in fig. 4.24), and the expected exclu-  
1511 sion limit on cross-section times branching fraction at 95% CL (displayed in fig. 4.25), as a function  
1512 of the reconstructed transverse mass of the resonance. To this purpose, the entire analysis workflow

### 4.3 Event selection

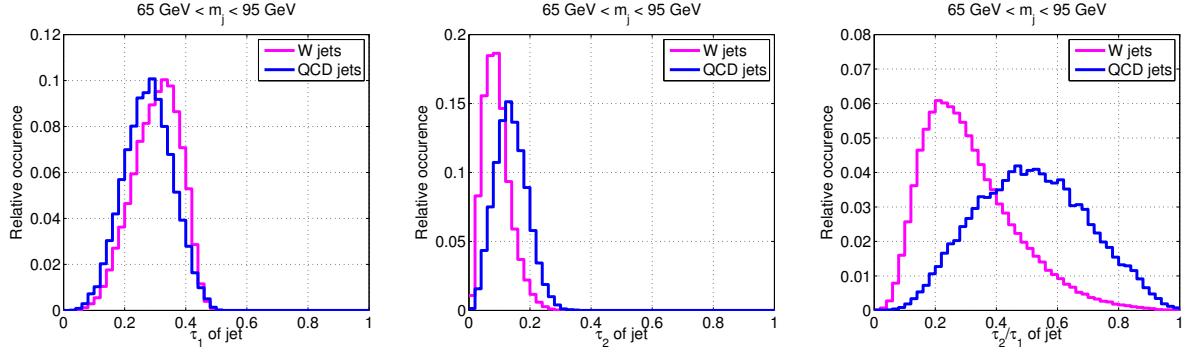


Figure 4.22: Distribution of  $\tau_1$  (left),  $\tau_2$  (center), and  $\tau_{21}$  (right) variables, in simulations of a  $W$  plus jets process (in pink) and for a multi-jet QCD originated process (in blue). A selection on the leading jet mass is applied:  $65 < m_j < 95$  GeV; jets are clustered with a parameter  $R_0 = 0.6$ ,  $p_T > 300$  GeV,  $|\eta| < 1.3$  [85].

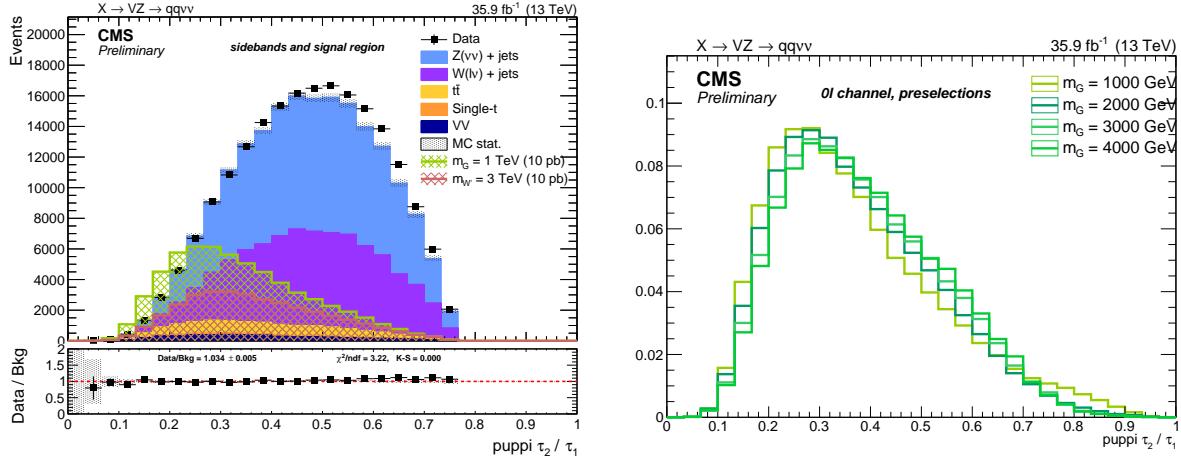


Figure 4.23: Distribution of the  $\tau_{21}$  subjetteness of the leading AK8 jet, selected as the hadronically decaying  $V$  candidate, for expected SM background and data (left), and for bulk graviton signal (right).

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

When doing the  $\tau_{21}$  categorization,  $V$ -tagging scale factors have been taken into account to correct data and simulation discrepancies introduced by the  $n$ -subjettiness. They are described in sec. 4.3.8.1.

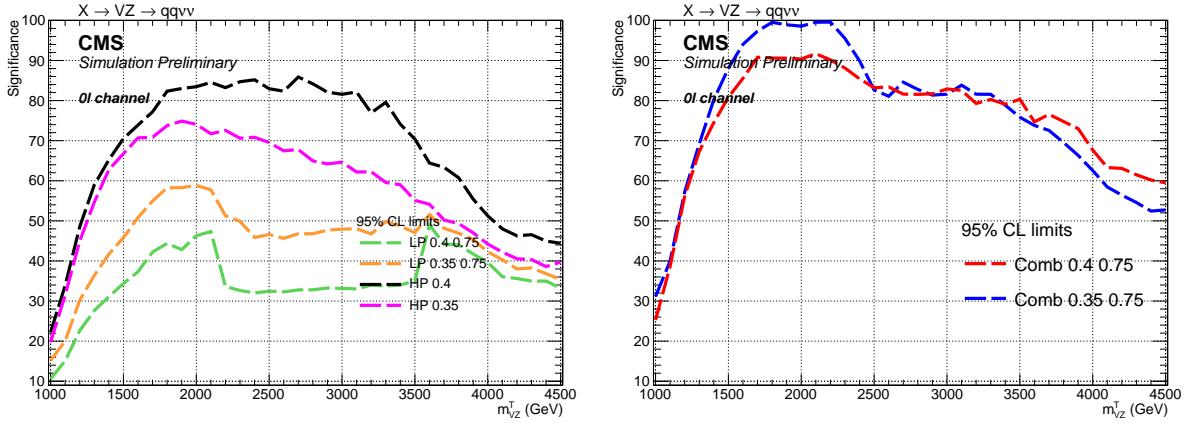


Figure 4.24: Analysis sensitivity to bulk graviton signals, computed by applying different  $\tau_{21}$  categorizations, considering the categories separately (left) and combining them together (right), as a function of the resonance mass.

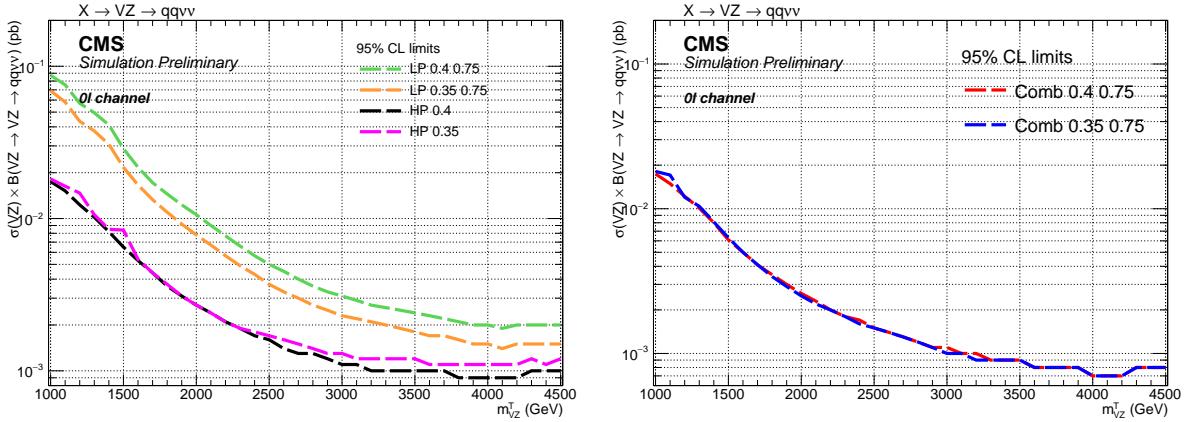


Figure 4.25: Exclusion limit on cross-section time branching fraction at 95% CL of bulk graviton signals, computed by applying different  $\tau_{21}$  categorizations, considering the categories separately (left) and combining them together (right), as a function of the resonance mass.

### 1524 4.3.8.1 Corrections induced by jet substructure variables

1525 By applying a selection on the jet  $\tau_{21}$ , the jet mass spectrum is sculpted, hence the effects of the  
 1526  $V$ -tagging procedure shall take into account both the selections on mass and on substructure si-  
 1527 multaneously. The distributions of the groomed jet mass and  $\tau_{21}$  subjetiness have been compared  
 1528 in data and simulations, by selecting samples of di-jet,  $t\bar{t}$  and  $W +$  jets events, and a significative  
 1529 discrepancy has been observed (10%) [76]. Scale factors are extracted by selecting a  $t\bar{t}$  sample in  
 1530 data, because an high  $p_T$   $W$  boson is produced by the top quark decay. The hadronically decaying  
 1531  $W$  boson is tagged by choosing events where the soft drop mass of a large-cone jet lies in a win-  
 1532 dow centered around the nominal  $W$  mass. The jet mass distributions of events passing and failing  
 1533 the selection on the  $\tau_{21}$  variable ( $\tau_{21} < 0.35$  and  $0.35 < \tau_{21} < 0.75$ , considered separately) are fit-  
 1534 ted simultaneously, both in data and in simulations. The  $V$ -tagging scale factors are defined as the  
 1535 ratio of the  $\tau_{21}$  categorization efficiencies in data and MC, and they are summarized in tab. 4.10.

### 4.3 Event selection

---

1536 The systematic uncertainties depend on the simulation of the  $t\bar{t}$  process, they cover the discrepancies  
1537 observed while using different Monte Carlo simulations, and due to the choice of the fitting  
1538 function.

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

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

1539 **4.3.9 b-tagging**

1540 The presence of a b-tagged quark can be an hint to identify the top quark decays, representing a  
1541 potential background to the search. The CSV b-tagging algorithm [65] is applied to the AK4 jets.  
1542 The jet is considered as tagged if the CSV discriminator value is above a threshold value; the b-tag  
1543 efficiency is defined as the number of jets fulfilling this requirement, divided by the total number of  
1544 jets. Since the purpose of the b-tagging is to reject the top quark events, the working point with the  
1545 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%

1546 Events where an AK4 jet, not laying in the AK8 jet cone, is b-tagged with the *loose* working point  
1547 threshold, are rejected. This veto allows to suppress the single-top events and  $t\bar{t}$  events by one half.  
1548 The b-tagging efficiency is not the same in data and MC. In order to take into account this difference,  
1549 b-tagging scale factors for b-jets and mis-tagged light jets, measured for different physics processes,  
1550 are calculated. A weight is extracted on a per-event basis, as a function of the b-tagging status of the  
1551 jets and their kinematic variables [86].

1552 **4.3.10 Missing Energy**

1553 As pointed out in sec. 3.2.9.8, Type-I corrected  $E_T^{\text{miss}}$  is used in the analysis, along with dedicated  
1554 filters to remove detector noise and events with bad reconstruction. In order to lie in the plateau  
1555 of the trigger efficiency,  $E_T^{\text{miss}} > 200\text{GeV}$ . Fig. 4.26 shows the  $E_T^{\text{miss}}$  distribution for data and Monte  
1556 Carlo after the corrections and filters.

1557 **4.3.11 Diboson candidate reconstruction**

1558 **4.3.11.1  $V \rightarrow q\bar{q}$  reconstruction**

1559 The identification of jets produced by the hadronic decays of one vector boson is based on the two  
1560 concepts:

- 1561 • *Jet mass*: jets produced by the decay of a massive particle should have an invariant mass  
1562 around the nominal mass of the original particle. Oppositely, jets originated by QCD radia-  
1563 tion are produced by the emission of quarks or gluons and typically have smaller invariant  
1564 mass. This effect is further enhanced by the grooming techniques (sec. 4.3.7).

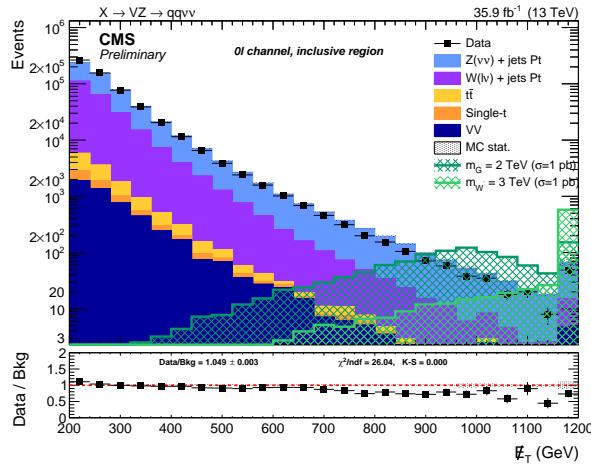


Figure 4.26: Type-1 corrected  $E_T^{\text{miss}}$  distribution after inclusive selections.

- *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.

The leading AK8 jet respecting the jet mass and jet substructure selections is tagged as the  $V$  candidate.

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

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

#### 4.3.11.3 Composite $VZ$ candidate reconstruction

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

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

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

#### 4.3.12 Final analysis selections

Events considered in this analysis have to pass a certain number of selections before being considered as suitable signal candidates, both in data and in simulations. The selections are reported below and in tab. 4.12. The selections applied to group the events in purity category, defined on the PUPPI corrected  $\tau_{21}$  subjetiness variable (sec. ??), and into signal or control region, defined on the PUPPI corrected soft drop mass (sec. 4.3.7) are reported in tab. 4.13. The final signal efficiency is 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  $p_T > 200 \text{ GeV}$ ;
- $\eta$ :  $|\eta| < 2.4$ ;
- *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  $\tau_{21}$  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:*
  - $p_T$ :  $p_T > 10 \text{ GeV}$ ;
  - $\eta$ :  $|\eta| < 2.5$ ;
  - *Id: veto* cut-based working point;
- *Veto on muons:*
  - $p_T$ :  $p_T > 10 \text{ GeV}$ ;
  - $\eta$ :  $|\eta| < 2.4$ ;
  - *Id: loose Id*;
  - *Isolation:* Particle-Flow Isolation  $< 0.25$ ;
- *Veto on hadronic taus:*
  - $p_T$ :  $p_T > 18 \text{ GeV}$ ;
  - $\eta$ :  $|\eta| < 2.4$ ;
  - *Id: loose Id*;
- *Veto on photons:*
  - $p_T$ :  $p_T > 15 \text{ GeV}$ ;

- 1618           –  $\eta: |\eta| < 2.5$ ;
- 1619           – *Id: loose cut-based working point.*
- 1620   Further selections are applied to suppress spurious events.
- 1621       • *Event cleaning:* events where the  $V$  and the  $Z$  candidates are collinear are rejected:  
 1622        $\Delta\varphi(V, \vec{p}_T^{\text{miss}}) > 2$ .
- 1623       • *Top rejection:* as discussed in sec. 4.3.9, a b-tag veto is imposed on AK4 jets lying outside the  
 1624       AK8 cone; this reduces the top quark background contamination by 50%.
- 1625       • *QCD rejection:* a minimum angular separation  $\Delta\varphi > 0.5$  is imposed in the transverse plane be-  
 1626       tween the  $\vec{p}_T^{\text{miss}}$  vector and the momenta of all the AK4 jets in the event, lying outside the AK8  
 1627       cone and not tagged as b-quark initiated jets. The effect of this cut is to suppress the multi-jet  
 1628       QCD background: it has been studied by considering additional QCD simulated samples to  
 1629       the analysis backgrounds. As it can be inferred by looking at the distribution of the minimum  
 1630       azimuthal separation between  $\vec{p}_T^{\text{miss}}$  and the AK4 jets, shown in fig. 4.27 (where looser selec-  
 1631       tions are applied w.r.t. the nominal selections of the analysis, *i.e.*, no QCD event cleaning is  
 1632       performed), if a minimum  $\Delta\varphi = 0.5$  threshold is imposed, the QCD contribution is reduced  
 1633       from 32% to 5%. In the final signal region, the QCD event yield amounts to 2%, and hence it  
 1634       is negligible (3% in low-purity, less than 1% in high-purity).

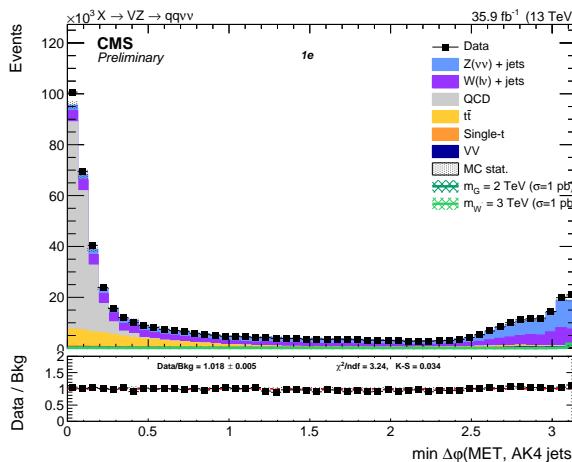


Figure 4.27: Distribution of the minimum azimuthal separation between  $\vec{p}_T^{\text{miss}}$  and the momenta of all the AK4 jets present in each event. By imposing  $\min\Delta\varphi > 0.5$ , the QCD background (in gray) is suppressed.

- 1635   The final selections of the analysis are summarized in tab. 4.12-4.13. The detection efficiencies due  
 1636   to each cut sequentially applied to bulk graviton signal samples (fig. 4.3.12.3, left) and  $W'$  signal  
 1637   samples (fig. 4.3.12.3, right) are shown. The signal efficiency for bulk graviton ranges from  $\sim 30\%$   
 1638   at 1 TeV, up to 20% at 4.5 TeV for low-purity category, whilst it's around 20% for the high-purity  
 1639   category in the whole mass range. The signal efficiency for  $W'$  ranges from  $\sim 40\%$  at 1 TeV, up to 25%  
 1640   at 4.5 TeV for low-purity category, whilst it's around 25% for the high-purity category in the whole  
 1641   mass range. The different detection efficiencies for the two signals are related to their production  
 1642   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
$E_T^{\text{miss}}$	Type-I corrected $> 200 \text{ GeV}$
Veto	$e, \mu, \tau, \gamma$
$V$	$p_T > 200 \text{ GeV, tight Id}$ $\text{nhf} < 0.8; \text{chf} > 0.2$
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.

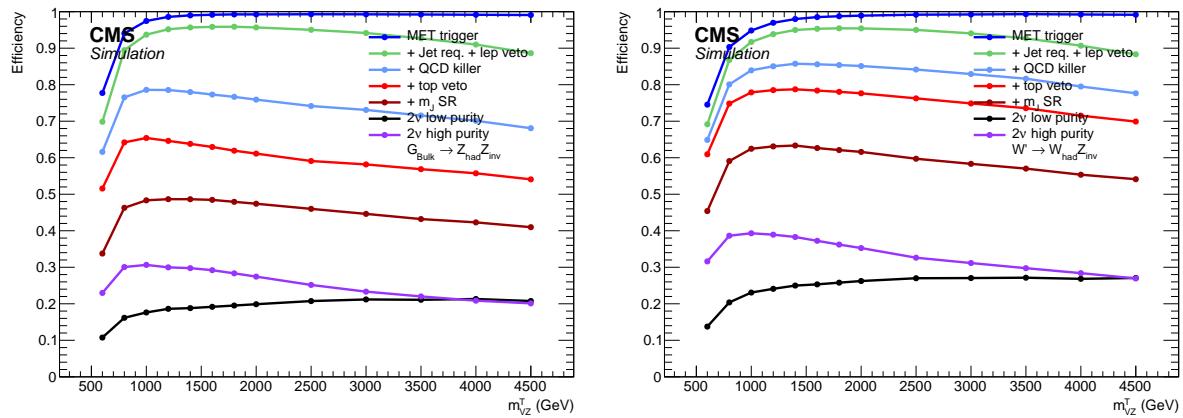


Figure 4.28: Signal efficiency for a spin-2 bulk graviton decaying into a pair of  $Z$  bosons (left), and for a spin-1  $W'$  decaying into a  $W$  and a  $Z$  bosons (right), as a function of the mass of the heavy particle. The efficiencies are separated by purity category after the signal region selections.

## 1645 4.4 Data and simulations comparison

1646 In this section, a comparison between data and simulation is reported for various kinematic ob-  
 1647 servables. It can be seen that the dominant background contribution comes from the  $Z + \text{jets}$  and  
 1648  $W + \text{jets}$  production, while sub-leading contributions from top ( $t\bar{t}$  and single-top) production and  
 1649 dibosons can be minor yet non-negligible.

1650 In the following plots (fig. 4.29-4.38), the comparison is performed in three different regions. On  
 1651 top of the selections defined above, additional criteria are defined:

- 1652 • *Inclusive*: no selection is performed on top of the selections, except for a veto on the jet mass  
 1653  $65 < m_V < 135 \text{ GeV}$  to avoid potential signal contamination from  $VZ$  signals;
- 1654 • *Sidebands (SB)*: only events in the sidebands, defined the interval between  $30 < m_V < 65 \text{ GeV}$   
 1655 and  $m_V > 135 \text{ GeV}$  are considered. This region can be considered as signal-depleted. The  
 1656 main difference with the previous regions is that the bulk of the jet mass distribution, peaking  
 1657 at  $m_V \sim 20 \text{ GeV}$ , is not included. The region selected is thus much closer kinematically to the  
 1658 signal region.
- 1659 • *Signal region (SR)*: it represents the phase space where is signal is expected.

1660 A summary of the number of expected events from Monte Carlo simulations, per each sample, along  
 1661 with the number of events observed in data in each category is reported in tab. 4.14. No significant  
 1662 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

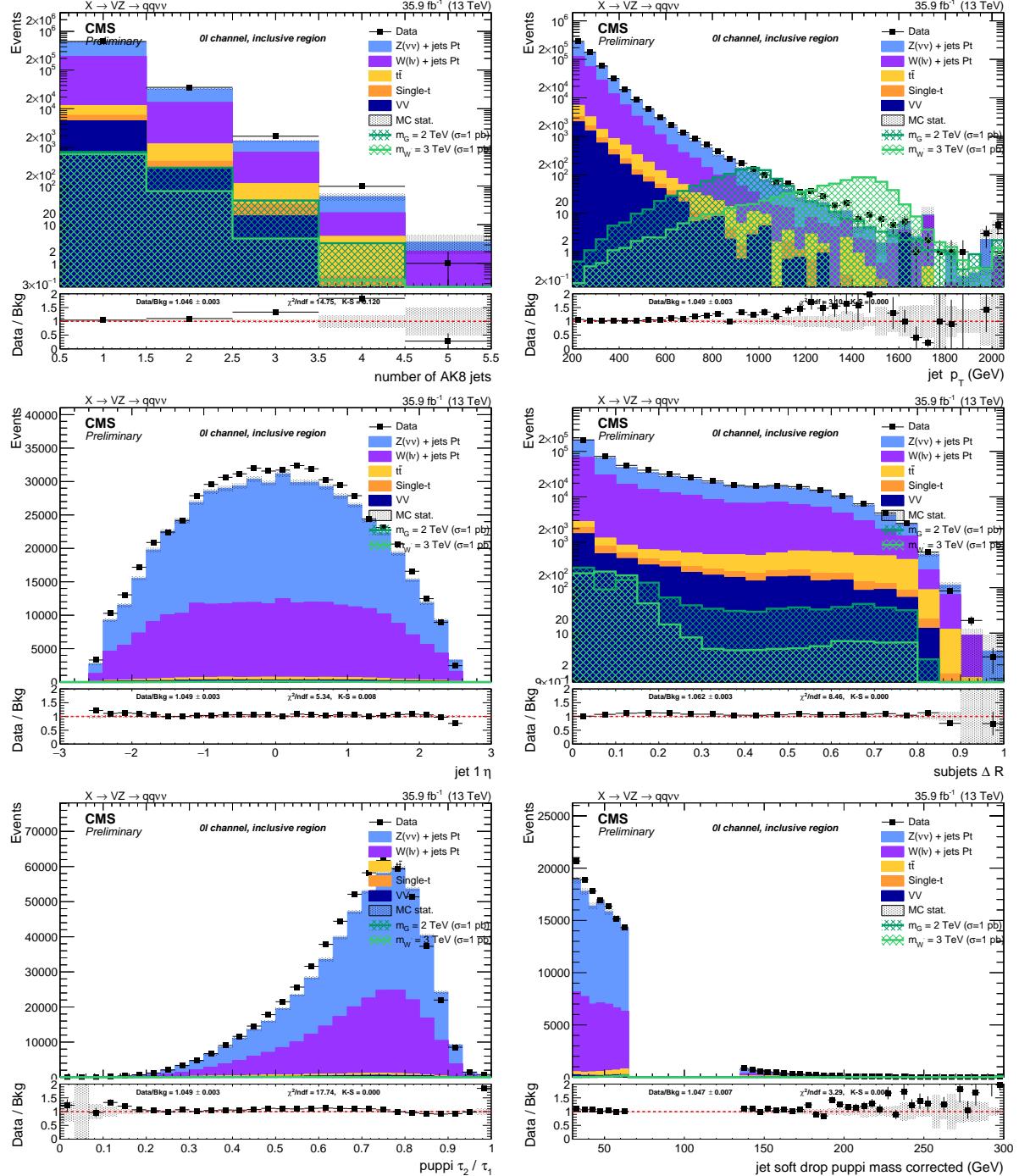


Figure 4.29: Top: number of AK8 jets in the event (left) and  $V$  jet candidate  $p_T$  (right). Center:  $V$  jet candidate  $\eta$  (left) and angular separation  $\Delta R$  between the constituents leading subjets (right). Bottom:  $V$  jet candidate  $\tau_{21}$  subjettness after PUPPI correction (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

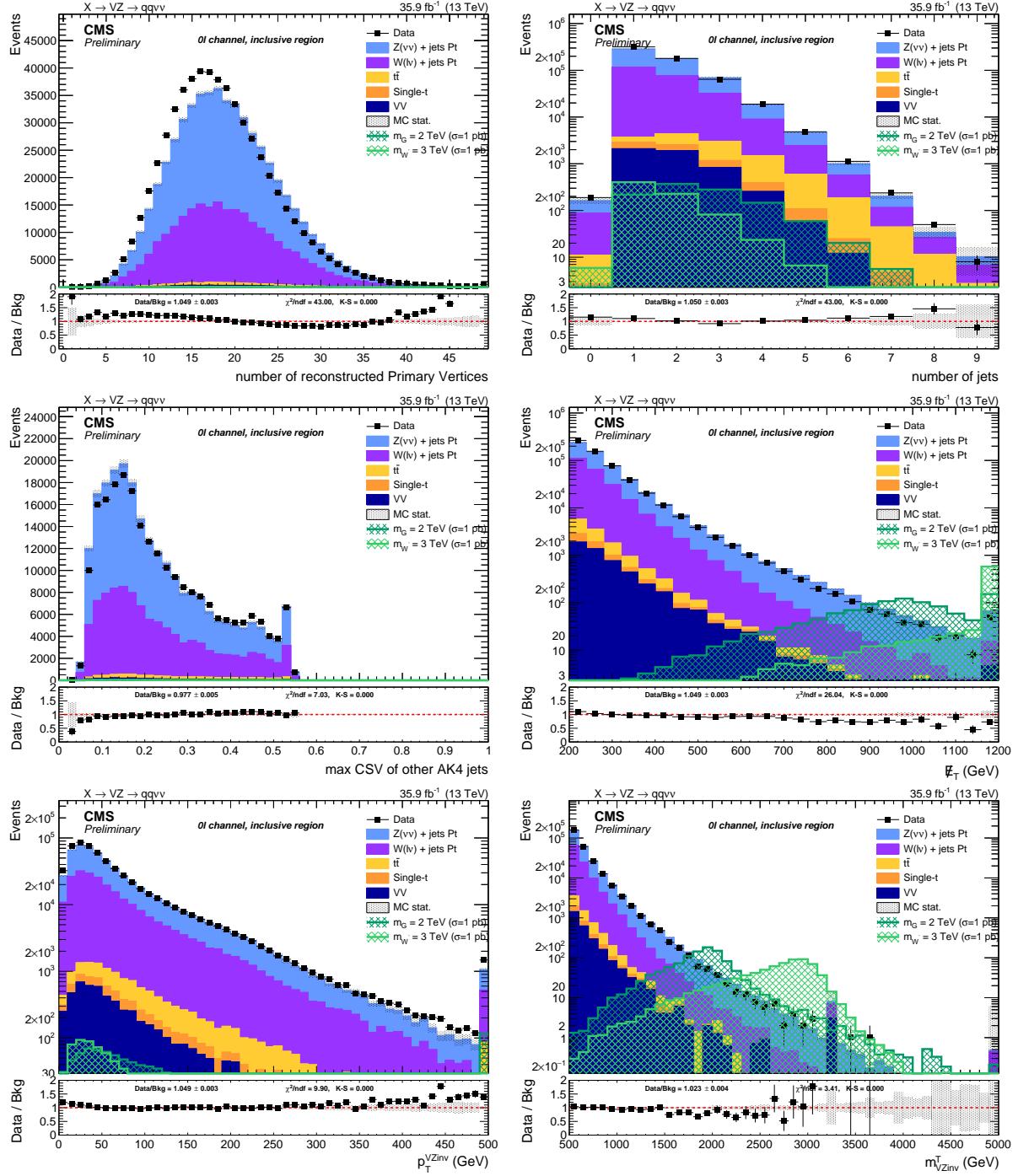


Figure 4.30: Top: number of reconstructed primary vertices (left) and number of AK4 jets in the event (right). Center: distribution of the b-tagging multivariate discriminant for the AK4 jets not included in the  $V$  jet cone (left) and  $E_T^{\text{miss}}$  distribution (right). Bottom:  $p_T$  of 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

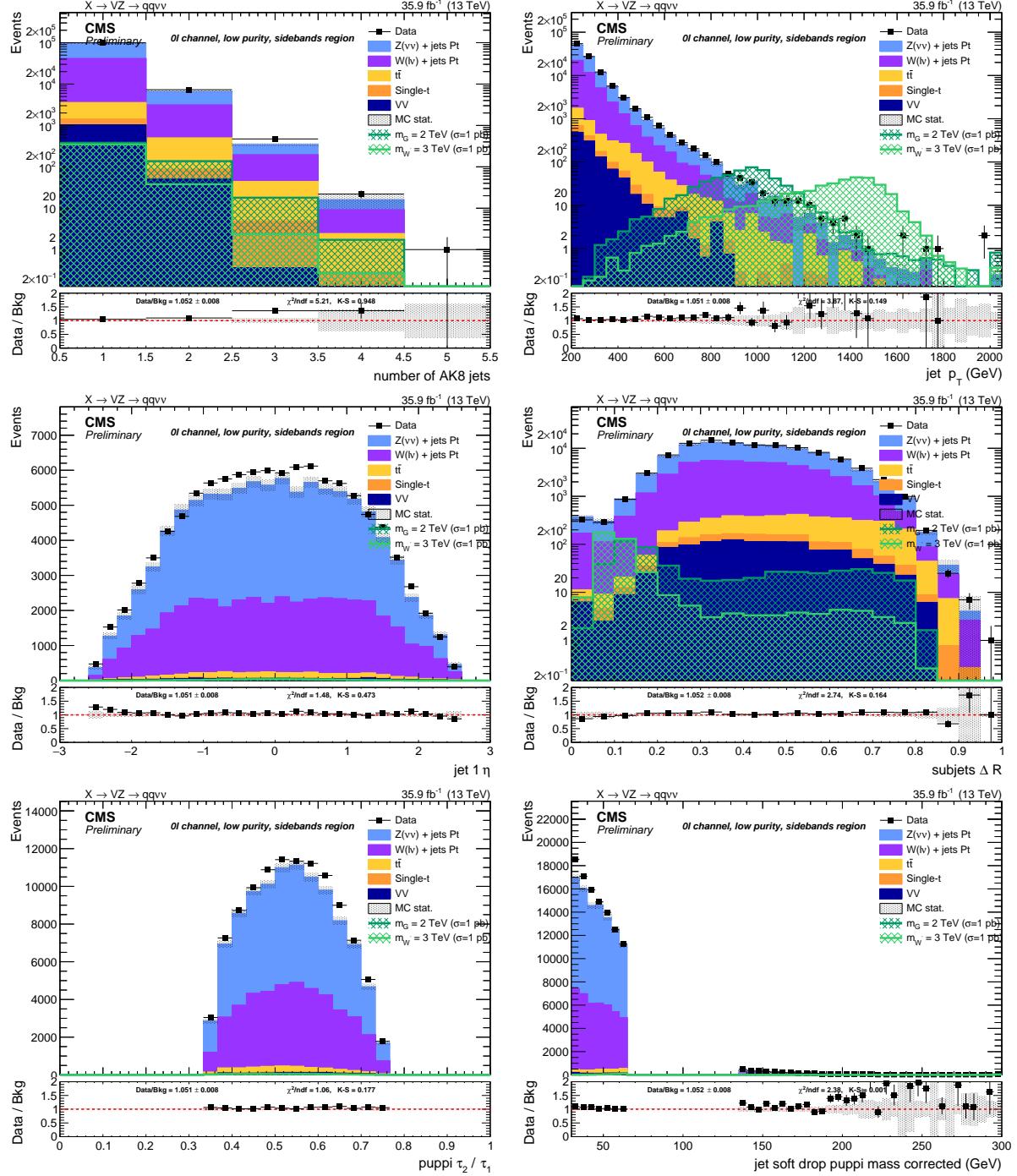


Figure 4.31: Top: number of AK8 jets in the event (left) and  $V$  jet candidate  $p_T$  (right). Center:  $V$  jet candidate  $\eta$  (left) and angular separation  $\Delta R$  between the constituents leading subjets (right). Bottom:  $V$  jet candidate  $\tau_{21}$  subjetiness after PUPPI correction (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.

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

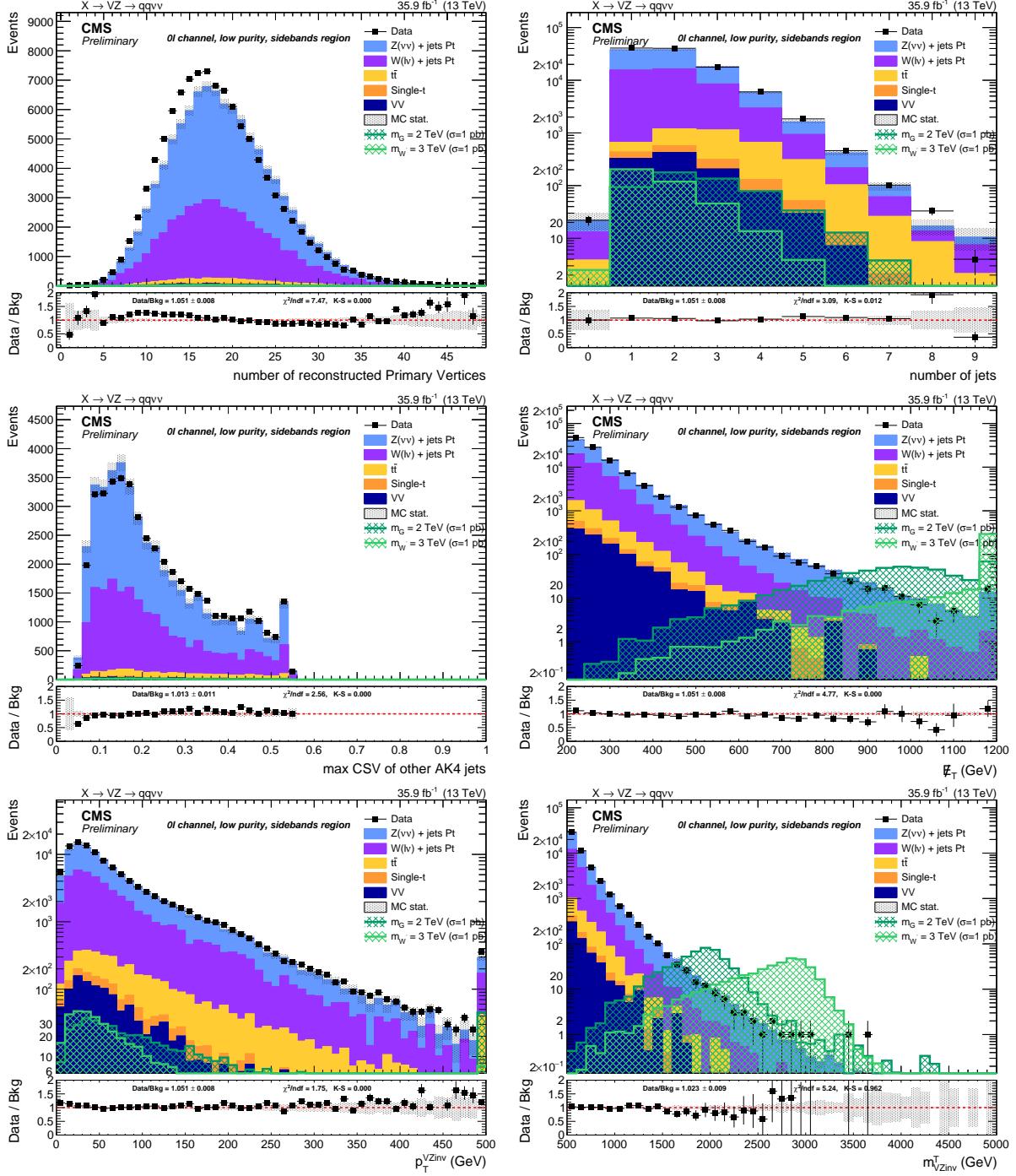


Figure 4.32: Top: number of reconstructed primary vertices (left) and number of AK4 jets in the event (right). Center: distribution of the b-tagging multivariate discriminant for the AK4 jets not included in the V jet cone (left) and  $E_T^{\text{miss}}$  distribution (right). Bottom:  $p_T$  of the 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

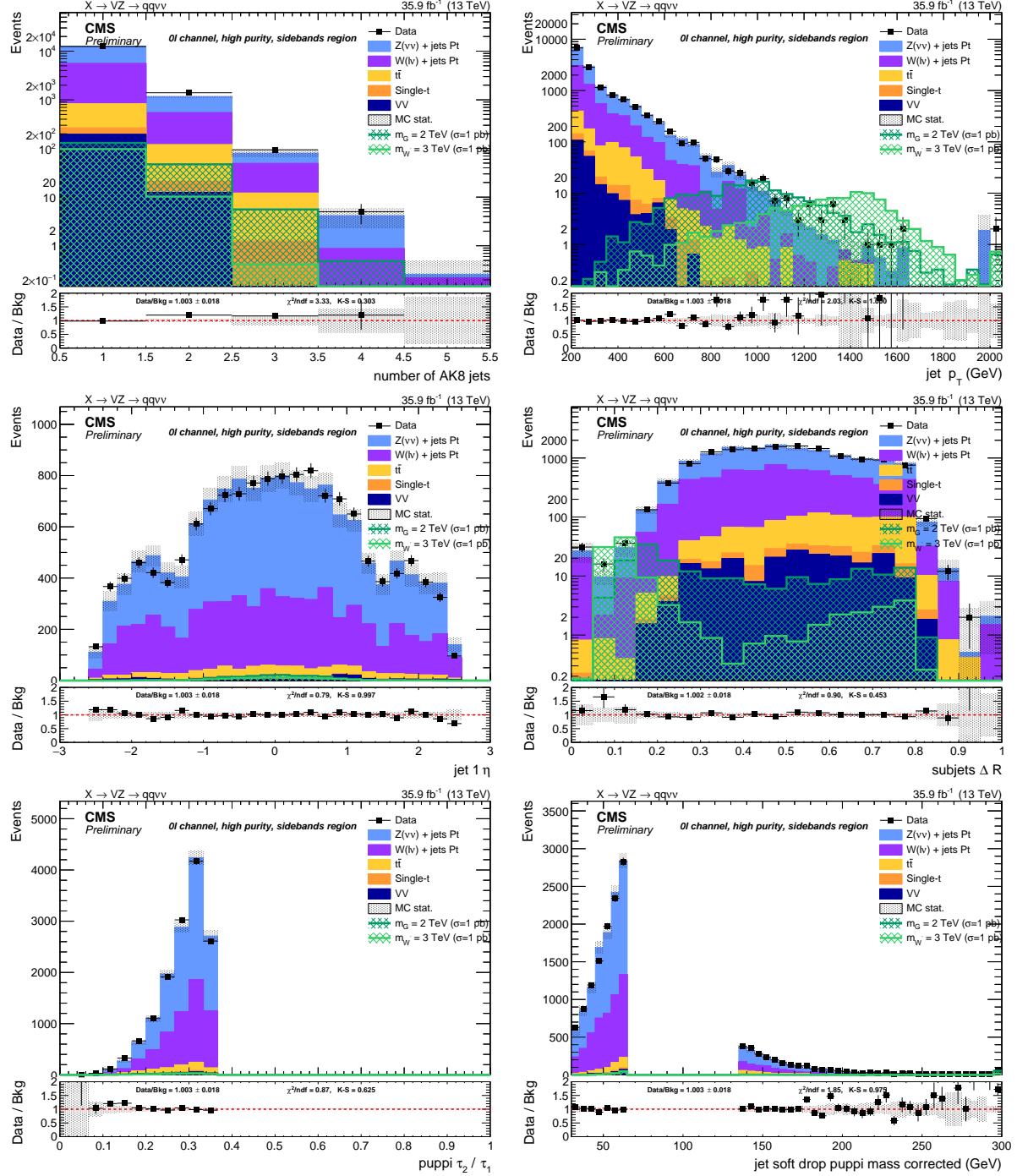


Figure 4.33: Top: number of AK8 jets in the event (left) and  $V$  jet candidate  $p_T$  (right). Center:  $V$  jet candidate  $\eta$  (left) and angular separation  $\Delta R$  between the constituents leading subjets (right). Bottom:  $V$  jet candidate  $\tau_{21}$  subjetiness after PUPPI correction (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

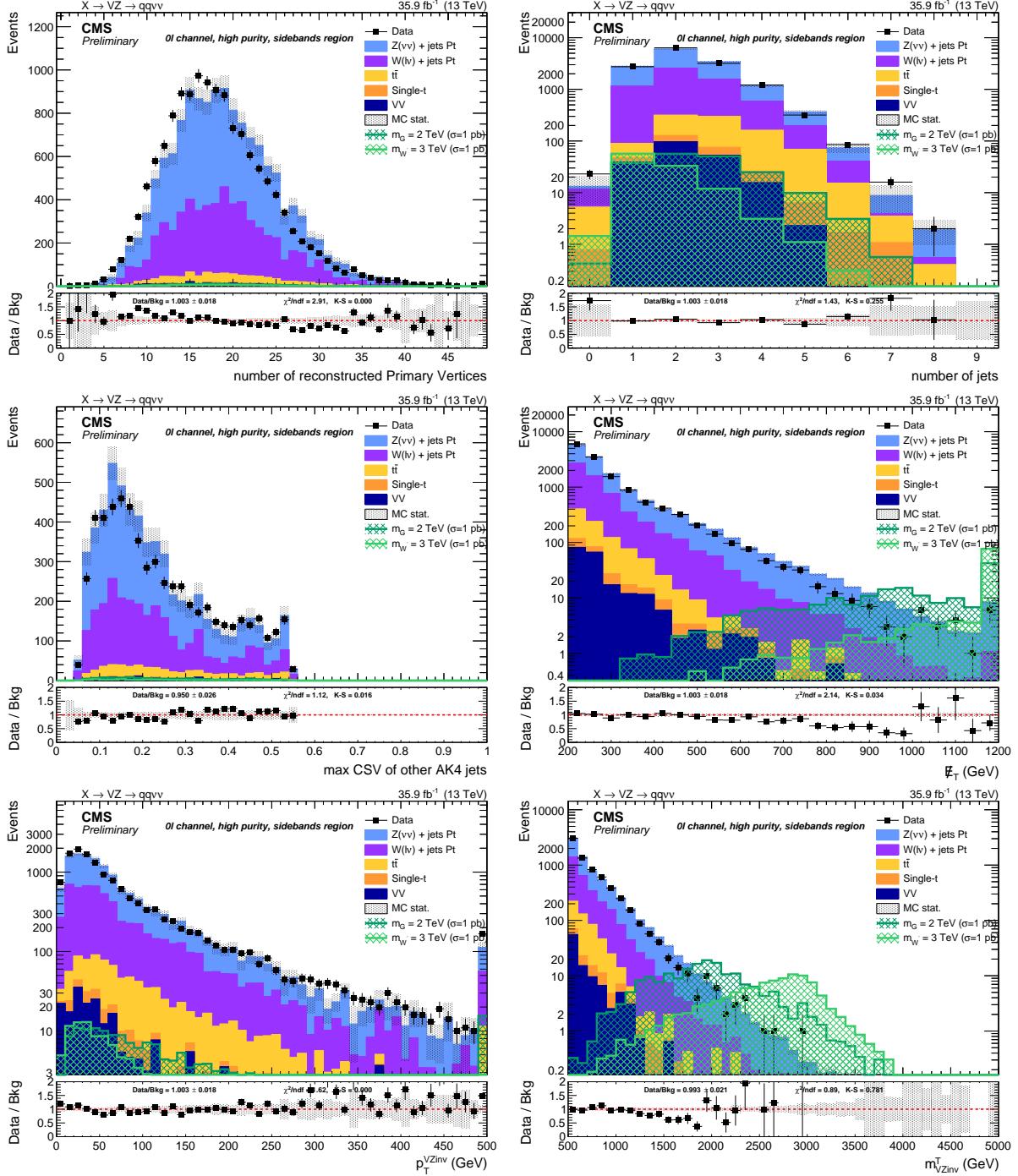


Figure 4.34: Top: number of reconstructed primary vertices (left) and number of AK4 jets in the event (right). Center: distribution of the b-tagging multivariate discriminant for the AK4 jets not included in the  $V$  jet cone (left) and  $E_T^{\text{miss}}$  distribution (right). Bottom:  $p_T$  of 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

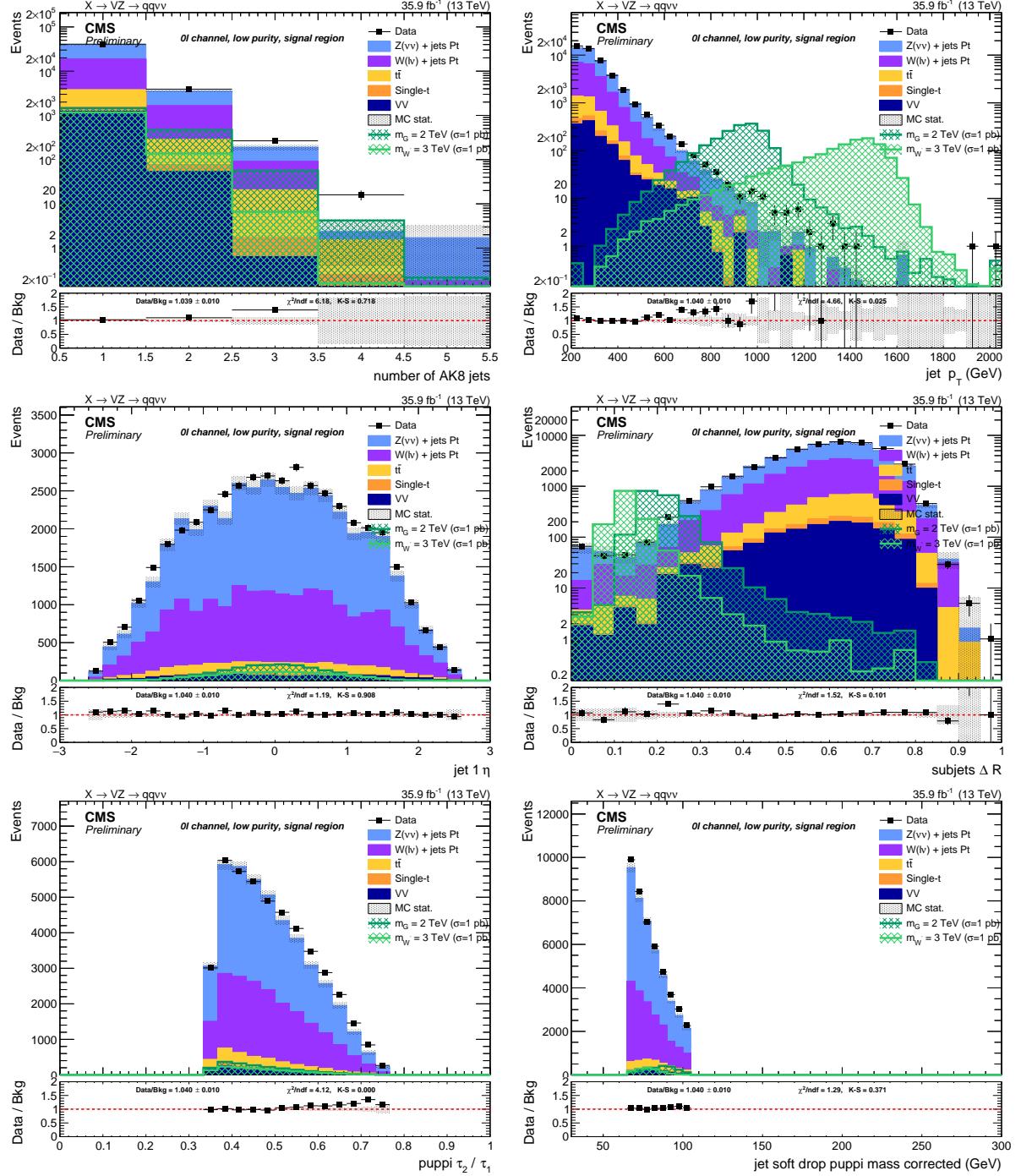


Figure 4.35: Top: number of AK8 jets in the event (left) and  $V$  jet candidate  $p_T$  (right). Center:  $V$  jet candidate  $\eta$  (left) and angular separation  $\Delta R$  between the constituents leading subjets (right). Bottom:  $V$  jet candidate  $\tau_{21}$  subjettness after PUPPI correction (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.

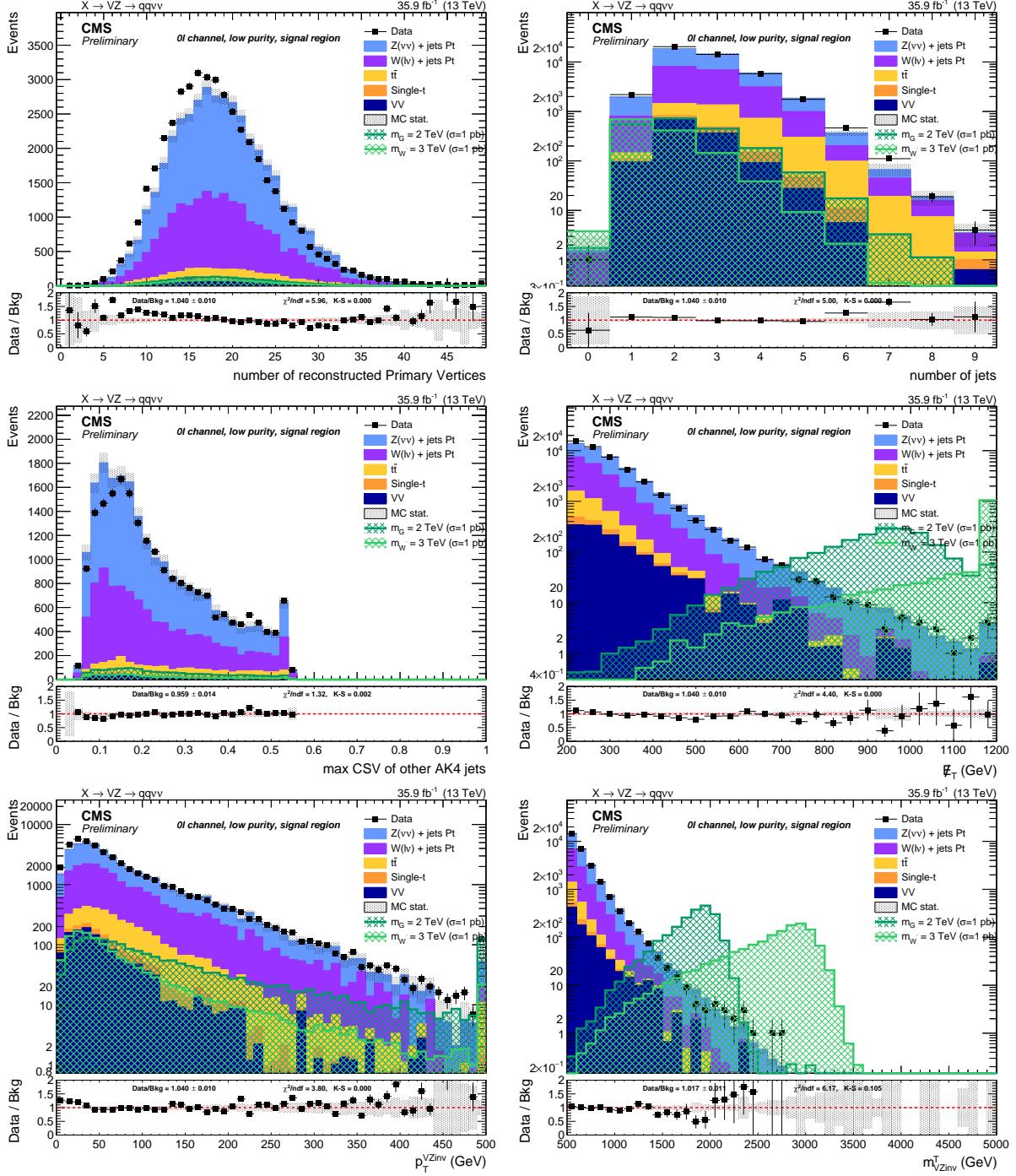


Figure 4.36: Top: number of reconstructed primary vertices (left) and number of AK4 jets in the event (right). Center: distribution of the b-tagging multivariate discriminant for the AK4 jets not included in the  $V$  jet cone (left) and  $E_T^{\text{miss}}$  distribution (right). Bottom:  $p_T$  of 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

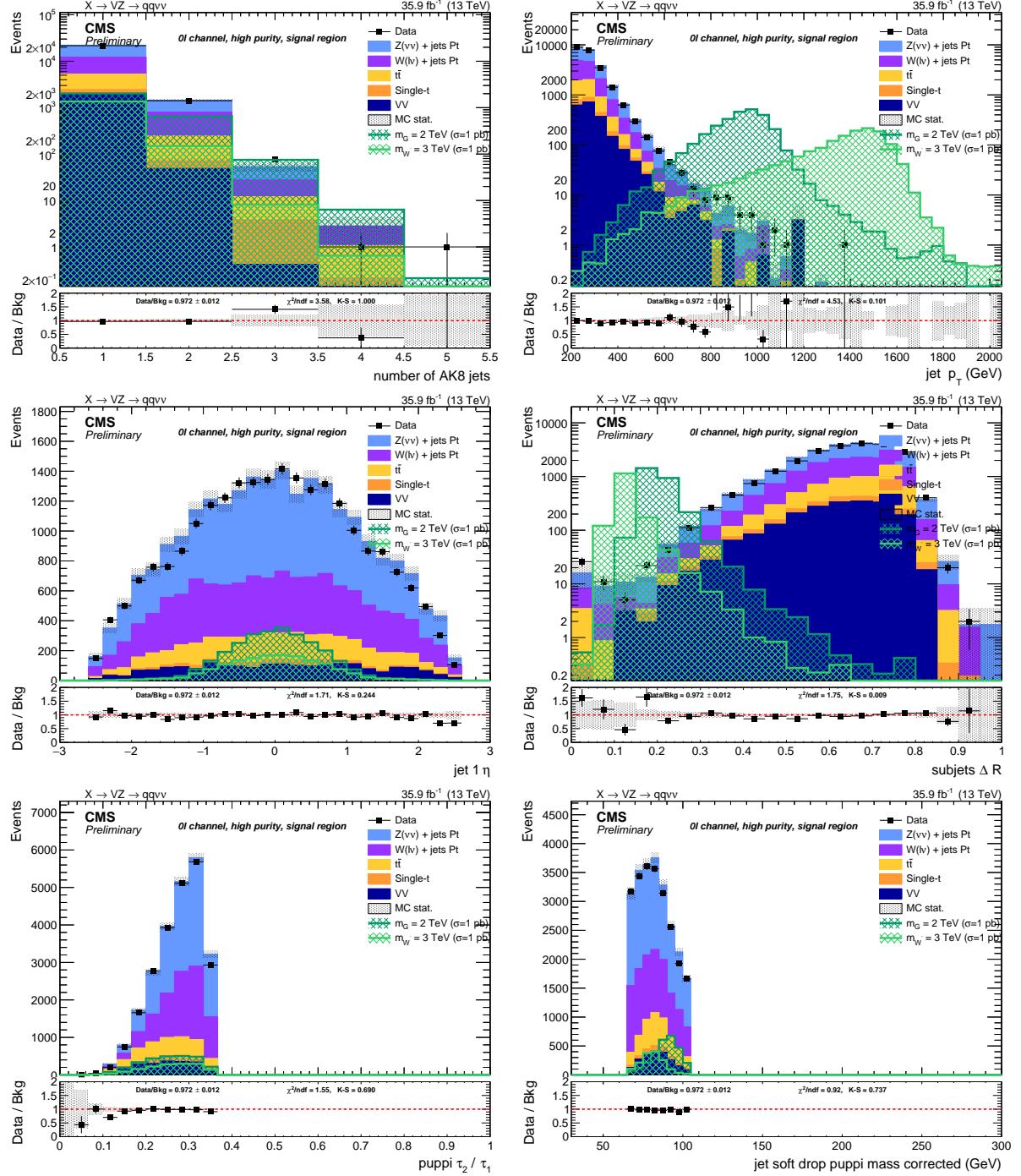


Figure 4.37: Top: number of AK8 jets in the event (left) and  $V$  jet candidate  $p_T$  (right). Center:  $V$  jet candidate  $\eta$  (left) and angular separation  $\Delta R$  between the constituents leading subjets (right). Bottom:  $V$  jet candidate  $\tau_{21}$  subjetiness after PUPPI correction (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

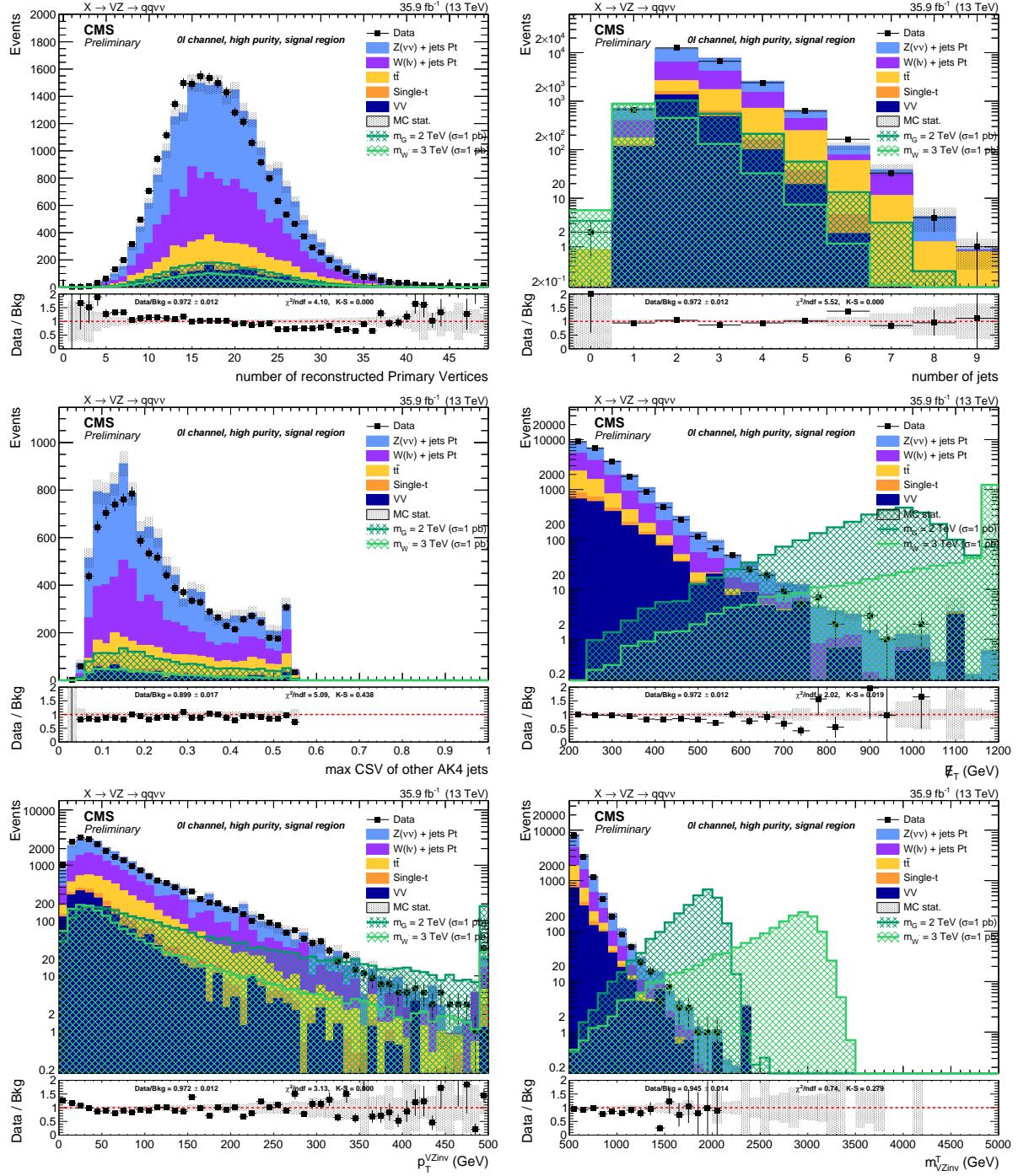


Figure 4.38: Top: number of reconstructed primary vertices (left) and number of AK4 jets in the event (right). Center: distribution of the b-tagging multivariate discriminant for the AK4 jets not included in the  $V$  jet cone (left) and  $E_T^{\text{miss}}$  distribution (right). Bottom:  $p_T$  of 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  $\alpha$  method is used in searches for heavy resonances since Run 1 [87], 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  $\alpha$  method extracts a predicted shape from the data in the SB to the SR using a transfer function (the  $\alpha$  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  $\alpha$ -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  $\alpha$  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  $\alpha$  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  $\alpha$  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

1703 to the data in the jet mass sidebands. During the fit to data SB, the parameters of the  $V +$  jets back-  
 1704 ground are left free to float and adapt to the data distribution. The integral of the final sum of the  
 1705 fitted functions over the SR jet mass range represents the background yield prediction in the SR.  
 1706 The empirical functional forms for each background are chosen to reflect the physics properties of  
 1707 the samples. In the low-purity category, the  $V +$  jets background is a falling background with no  
 1708 peaks, modelled as a power law, while in the high-purity category the  $V +$  jets background compo-  
 1709 nent is characterized by a broad distribution roughly centered at  $m_V$ , modelled as a gaussian, with  
 1710 an exponential tail at high mass values. The exponential falling  $VV m_V$  spectrum shows a peak,  
 1711 corresponding to the reconstruction of a vector boson hadronic decay. The hadronic decays of  $W$   
 1712 and  $Z$  bosons cannot be distinguished, hence they are modelled together as a gaussian. For the  
 1713 jet mass spectrum of the Top backgrounds, two peaks corresponding to the  $W$  and top quark mass  
 1714 can be observed; they are modelled as gaussian functions, superimposed to a falling exponential  
 1715 background.  
 1716 An extended likelihood fit is performed, hence the functional forms chosen to build the jet mass  
 1717 templates are normalized to unity (becoming probability density functions) through normalization  
 1718 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  $w$ ):

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

1719 The choice of the functions is category-dependent, and it is summarized in tab. 4.15. In order to  
 1720 make the background evaluation less dependent as possible from the choice of the function de-  
 1721 scribing the jet mass of the main  $V +$  jets background, an alternative function has been used to  
 1722 fit the  $V +$  jets mass spectrum. The absolute difference bewteen the number of expected events  
 1723 calculated with the main  $V +$  jets function and the alternative is taken as systematic uncertainty.  
 1724 The following plots (fig. 4.39-4.40) show the fits to the jet mass distributions in Monte Carlo samples,  
 1725 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	$\pm 52.5$	$\pm 16.0$	$\pm 1.1$	2314
SR	low-purity	1093.2	$\pm 48.1$	$\pm 16.4$	$\pm 49.1$	1153
SB	high-purity	779.8	$\pm 29.1$	$\pm 13.1$	$\pm 0.3$	774
SR	high-purity	254.4	$\pm 15.3$	$\pm 17.9$	$\pm 7.8$	271

1726 dotted lines. The background estimation, after the fit to data SB, is shown in fig. 4.41. The bottom  
 1727 panels of each plot display the fit pulls (per-bin), namely, the number of events observed in data  
 1728 (or in Monte Carlo simulations) minus the number of events predicted by the fit, divided by the  
 1729 uncertainty in the data (or simulations). Table 4.16 summarizes the expected background yield in  
 1730 the signal region, that is in agreement with observation in both the purity categories. The quoted  
 1731 uncertainties are calculated as follows:

- 1732 • the statistic uncertainty is the uncertainty of the fit to the  $V + \text{jet}$  background performed on  
 1733 data SB;  
 1734 • the systematic uncertainty is the propagation of the uncertainties of the fits to the  $VV$  and  
 1735 Top backgrounds performed on simulations, to the fit performed on data SB to extract the  $V$   
 1736 + jets functional parameters;  
 1737 • the alternative function uncertainty describes the discrepancy in the background yield in SR  
 1738 depending on the choice of the function to describe the  $V + \text{jets}$  background.

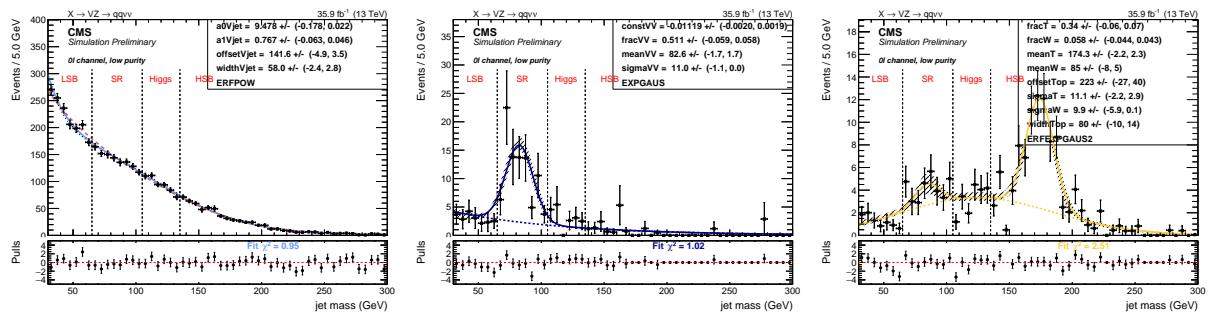


Figure 4.39: Fit to the simulated  $m_V$  in the low-purity category for the three backgrounds:  $V + \text{jets}$  (left),  $VV$  (center), Top (right). For the main background prediction, the alternative function is displayed with a dotted red line, superimposed to the main choice (continuous light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

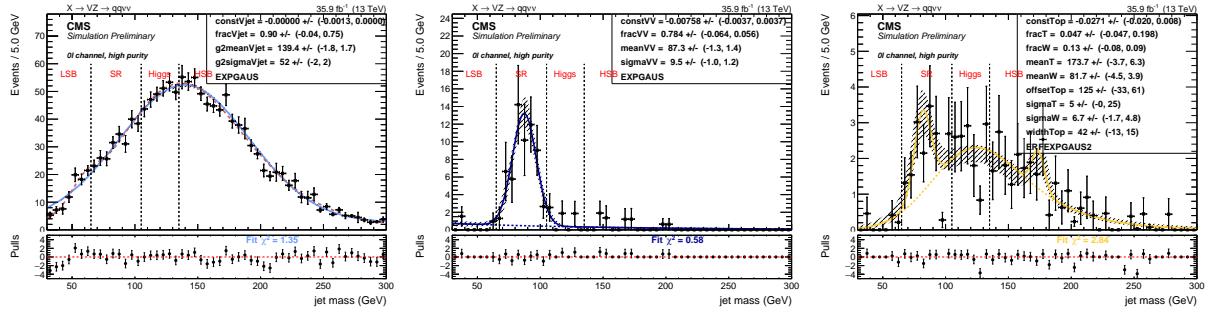


Figure 4.40: Fit to the simulated  $m_V$  in the high-purity category for the three backgrounds:  $V +$  jets (left),  $VV$  (center), Top (right). For the main background prediction, the alternative function is displayed with a dotted red line, superimposed to the main choice (continuous light blue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

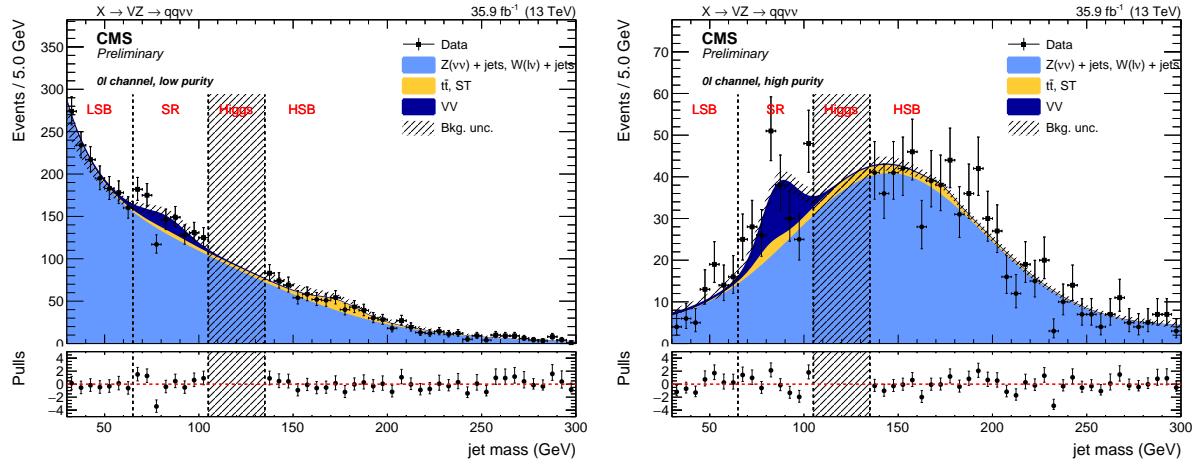


Figure 4.41: Background yield prediction in the signal region, after the fit to data sidebands, in the low- (left) and high-purity category (right). Data and predictions are in agreement.

## 1739 4.5.2 Background shape

1740 The second part of the background prediction consists in estimating the background shape of the  
 1741 transverse mass of the diboson candidate,  $m_{VZ}^T$ . Each transverse mass spectrum is parametrized  
 1742 separately for the  $V +$  jets background ( $f_{SR}^{MC, V + jets}(m_{VZ}^T)$ ,  $f_{SB}^{MC, V + jets}(m_{VZ}^T)$ ), Top production ( $f_{SR}^{MC, Top}(m_{VZ}^T)$ ,  
 1743  $f_{SB}^{MC, Top}(m_{VZ}^T)$ ), and diboson background ( $f_{SR}^{MC, VV}(m_{VZ}^T)$ ,  $f_{SB}^{MC, VV}(m_{VZ}^T)$ ). The parameters of these  
 1744 functions are extracted by fitting the simulated  $m_{VZ}^T$  spectra in SR and SB, respectively. The top and  
 1745 the diboson spectra are normalized to luminosity; the  $V +$  jets spectrum is normalized according  
 1746 to the data-driven prediction obtained in sec. 4.5.1. The functions describing the  $V +$  jets back-  
 1747 ground, calculated from simulations, are used to define the  $\alpha$  function, that has the purpose to take  
 1748 into account the kinematical differences of the SR compared to SB:

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

1749 The parameters describing the main background are then left free to float and extracted through a  
 1750 fit to data in the SB, after subtracting the corresponding Top and  $VV$  contribution from data. The  
 1751 resulting shape is then multiplied by the  $\alpha$  function in order to get the main background expectation

## 4.5 Background estimation technique

---

1752 in the SR. Finally, the Top and diboson contributions in the SR are added to the main background  
 1753 estimation.

1754 In formulas, the procedure used to extract the total background prediction is the following:

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

1755 where the expression in brackets represents the main background evaluation in data SB; the  $\alpha$ -ratio  
 1756 is the expression enclosed in square brackets.

1757 The functions probed to parametrize the  $m_{VZ}^T$  distributions are smoothly falling exponential func-  
 1758 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-  
 1758 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

1759 The functions chosen to parametrize the backgrounds and extract the  $\alpha$  function are reported in  
 1760 tab. 4.17, for each category. As a cross-check for the main  $\alpha$  function used in the background esti-  
 1761 mation, an additional  $\alpha$  function is extracted with alternative function choices for the  $V + \text{jets}$  back-  
 1762 ground. Table 4.17 reports both the main function and the alternative function. In fig. 4.42 (4.44),  
 1763 the fits to each simulated background are reported for sidebands and signal region respectively, for  
 1764 low- (high-) purity categories. In fig. 4.43 (4.45), the results of the fit to data SB are presented for  
 1765 the low- (high-) purity categories: the expected background distribution in SB, where parameters  
 1766 describing the  $V + \text{jets}$  background are extracted according to data distribution (left); the  $\alpha$ -ratio  
 1767 function, calculated with the main function to describe the  $V + \text{jets}$  background (black solid line)  
 1768 and the alternative function (gray dotted line) (center); the full background estimation performed  
 1769 with the main and alternative functions for describing the  $V + \text{jets}$  background: the background  
 1770 shape in SB (blue solid curve for the main function, light blue dotted curve for the alternative) and  
 1771 the final background shape in SR (red solid line for the main function, green dotted line for the al-  
 1772 ternative) (right). A proof to the compatibility of the two predictions in SR is presented in sec. 4.5.3.  
 1773 The bottom panels in the plots display the fit pulls (per-bin), namely, the number of events observed  
 1774 in data (or in Monte Carlo simulations) minus the number of events predicted by the fit, divided by  
 1775 the uncertainty in the data (or simulations).

1776 Fig. 4.46 summarizes the final background predictions as a function of the search variable, the trans-  
 1777 verse mass. Data and predictions are in agreement in both the categories.

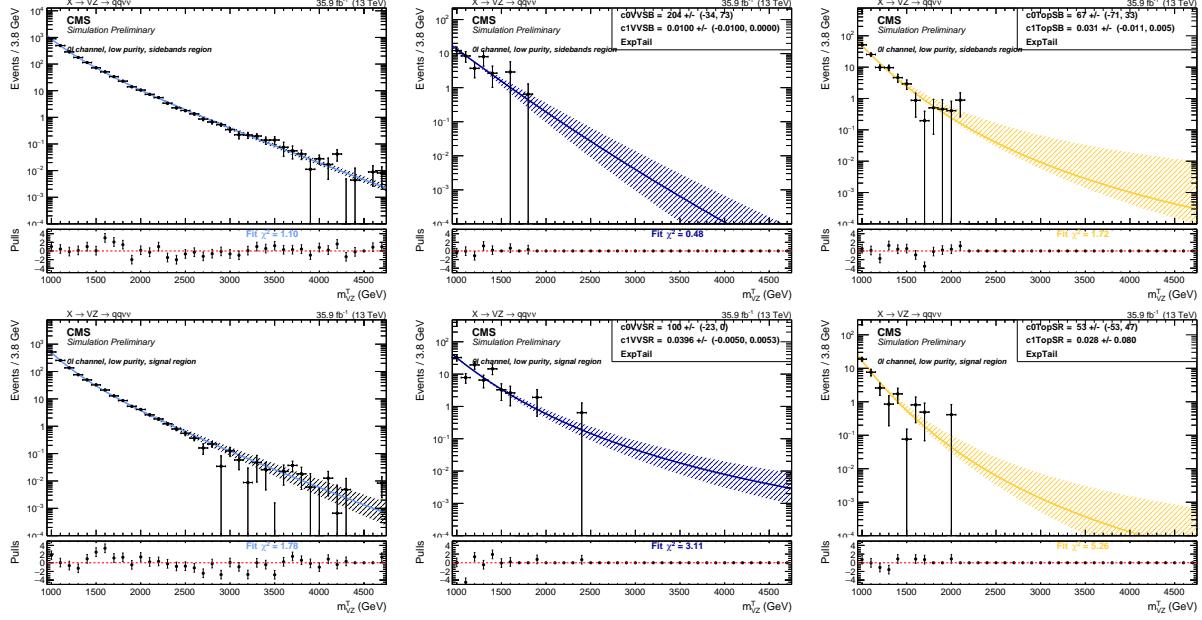


Figure 4.42: Low-purity category. Top: fits to the simulated background components  $V + \text{jets}$  (left),  $\text{VV}$  (center),  $\text{Top}$  (right) in the sidebands (SB). Bottom: fits to the simulated background components  $V + \text{jets}$  (left),  $\text{VV}$  (center),  $\text{Top}$  (right)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

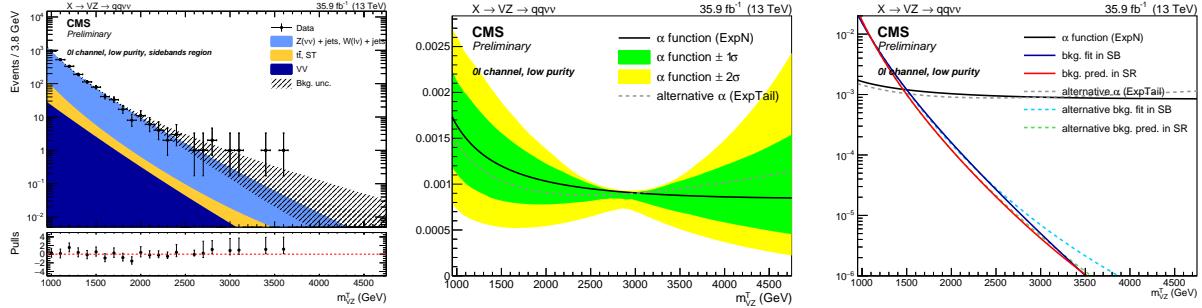


Figure 4.43: Low-purity category. Result of the fit to data in the SB (left),  $\alpha$ -ratio function (center), and  $\alpha$  function compared to the background shape in both SB and SR (right). The black line, with the corresponding  $1\sigma$  (green) and  $2\sigma$  (yellow) uncertainty bands, represents the  $\alpha$  function. The gray line is the alternative  $\alpha$  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.

### 1778 4.5.3 Alpha method validations

1779 The first required validation is performed in order to legitimate the choice of putting the  $Z + \text{jets}$   
 1780 and the  $W + \text{jets}$  backgrounds together while performing the background estimation. The full pro-  
 1781 cedure has been repeated, by keeping the two background contributions separated. Fit results per-  
 1782 formed in SB (top plots) and SR (bottom plots) in MC samples are displayed in fig. 4.47 ( 4.48) for  
 1783 low- (high-) purity category, for  $Z + \text{jets}$  and  $W + \text{jets}$  background, separately, and for the combina-  
 1784 tion of the two. In fig. 4.49, the  $\alpha$  functions calculated for  $Z + \text{jets}$  background (red dotted line) and  
 1785 for  $W + \text{jets}$  background (blue dotted line) are in agreement with the  $\alpha$  function used in the analysis  
 1786 (black solid line), calculated by merging together the two backgrounds, both in low- (left plot) and

## 4.5 Background estimation technique

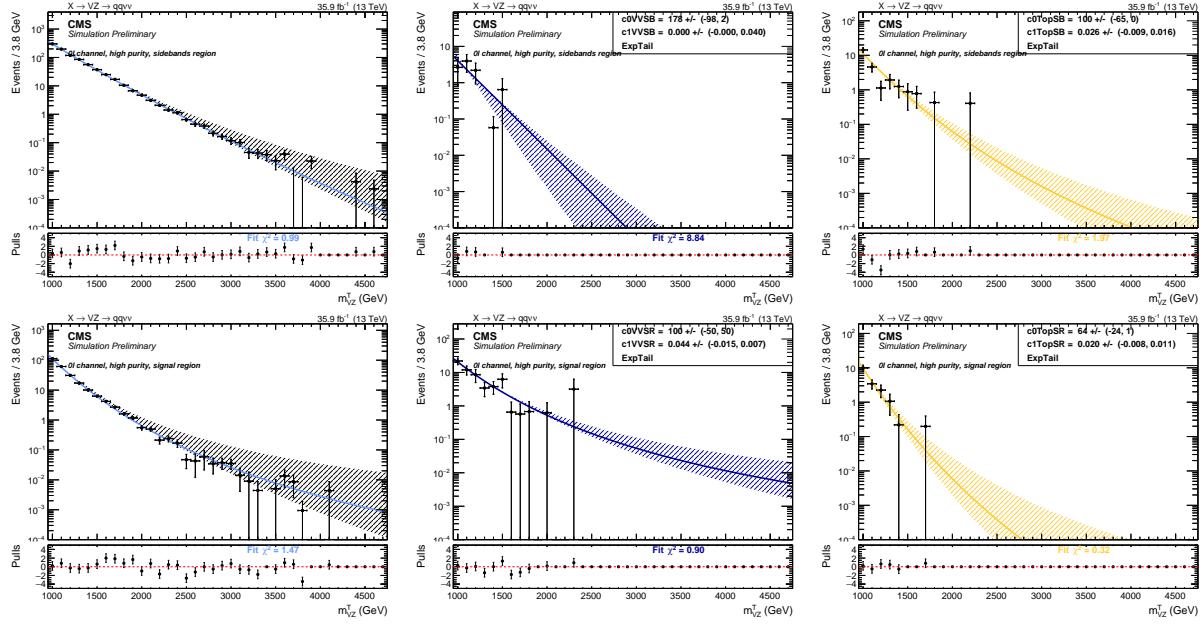


Figure 4.44: High-purity category. Top: fits to the simulated background components  $V + \text{jets}$  (left),  $VV$  (center), Top (right) in the sidebands (SB). Bottom: fits to the simulated background components  $V + \text{jets}$  (left),  $VV$  (center), Top (right) in the signal region (SR).

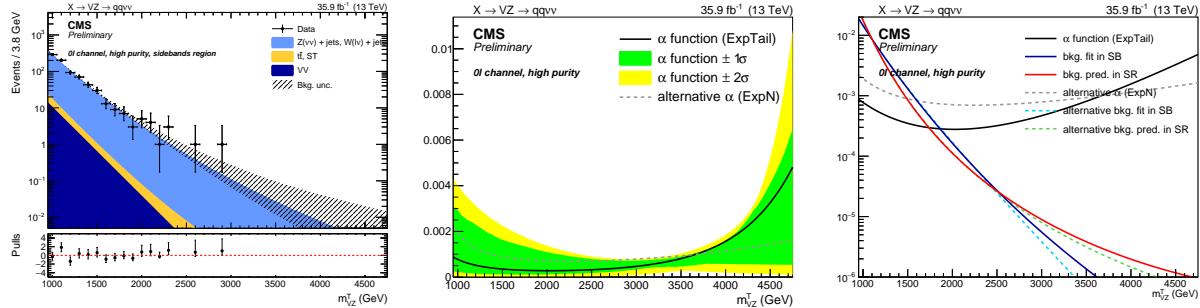


Figure 4.45: High-purity category. Result of the fit to data in the SB (left),  $\alpha$ -ratio function (center), and  $\alpha$  function compared to the background shape in both SB and SR (right). The black line, with the corresponding  $1\sigma$  (green) and  $2\sigma$  (yellow) uncertainty bands, represents the  $\alpha$  function. The gray line is the alternative  $\alpha$  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.

1787 high-purity category (right plot).

1788 As a robustness check of the  $\alpha$ -ratio method, a closure test is performed on data. Instead of pre-  
1789 dicting the background in the real SR from both the lower and the upper jet mass sidebands, the SB  
1790 and SR are redefined for the purposes of this test. The low sideband is splitted into two sub-regions:  
1791 30 – 50 GeV (LSB) and 50 – 65 GeV (SR). The former is considered as the new low sideband, while  
1792 the latter is exploited as a pseudo-signal region. The high sideband is instead effectively used in the  
1793 fit without any modifications with respect to the standard  $\alpha$ -ratio method. With this configuration,  
1794 the prediction of the background in the SR region is estimated from the fit to the LSB region and

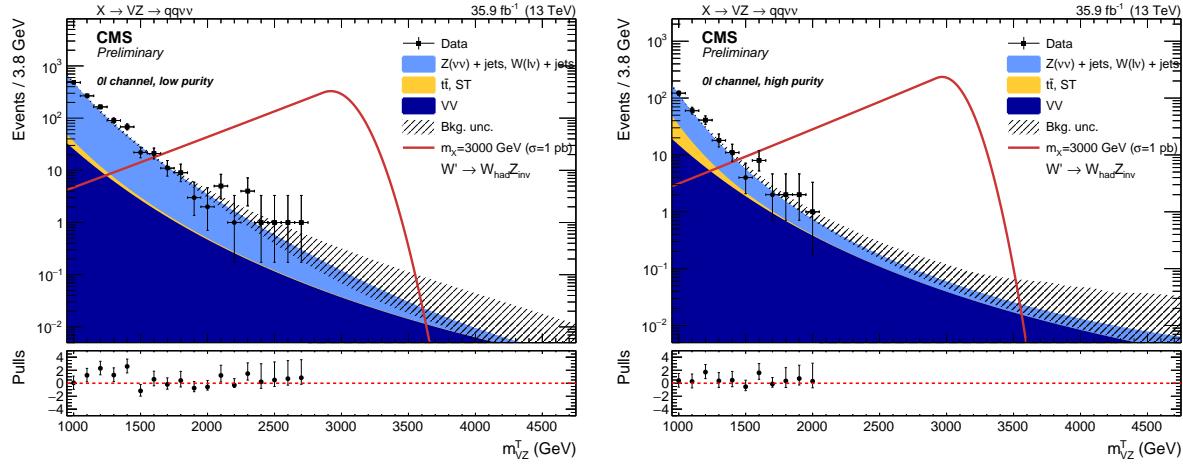


Figure 4.46: Expected background predicted with the  $\alpha$  method in the low- (left) and high-purity category (right), compared to observations (black markers) and a signal hypothesis of a spin-1  $W'$  of mass 3 TeV.

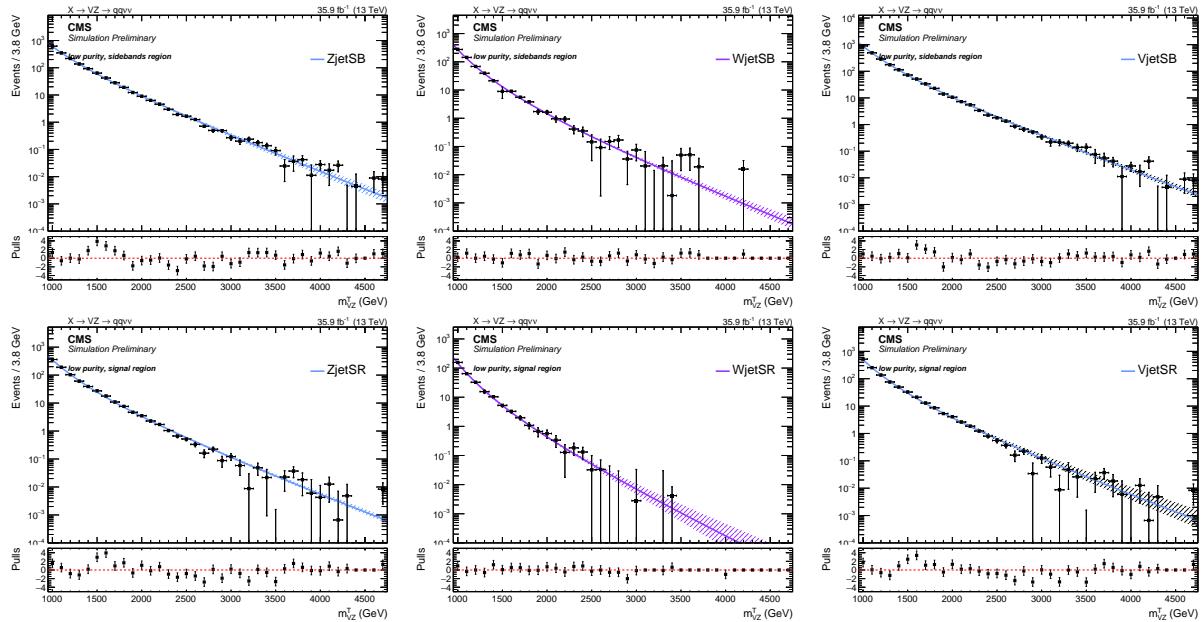


Figure 4.47: Validation of the  $\alpha$  method, low-purity category. Top: fits to the simulated background components  $Z + \text{jets}$  (left),  $W + \text{jets}$  (center), and their combination  $V + \text{jets}$  (right), in the sidebands (SB). Bottom: fits to the simulated background components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

1795 the high-sidebands, and checked with data for both shape and normalization. This test has been  
1796 performed before the unblinding of the signal region of the analysis.

1797 In fig. 4.50 and tab. 4.18, the predicted shapes and normalizations are compared to the observed  
1798 ones in data. A good overall agreement both in normalization and shape is obtained. There is a  
1799 bit of tension in normalization for high-purity category, due to an upper fluctuation in data around  
1800 60 GeV. This cross check confirms that the method to extract the  $V + \text{jets}$  background is reliable  
1801 and can be used to model the background in the search for potential excesses in the signal region

## 4.5 Background estimation technique

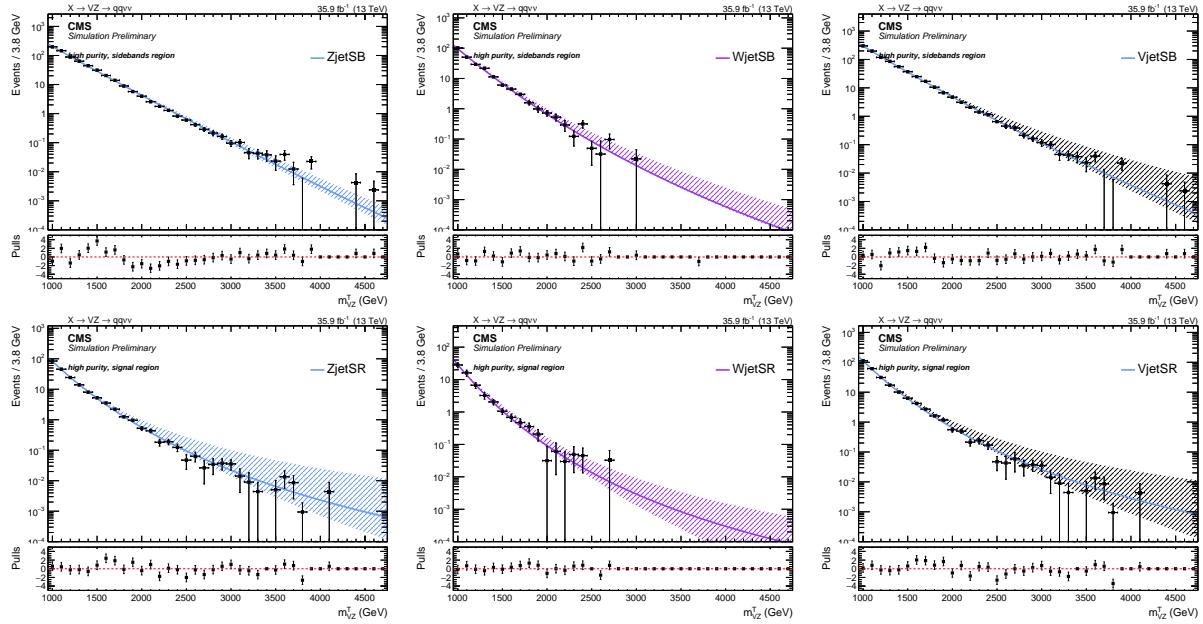


Figure 4.48: Validation of the  $\alpha$  method, high-purity category. Top: fits to the simulated background components  $Z + \text{jets}$  (left),  $W + \text{jets}$  (center), and their combination  $V + \text{jets}$  (right), in the sidebands (SB). Bottom: fits to the simulated background components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

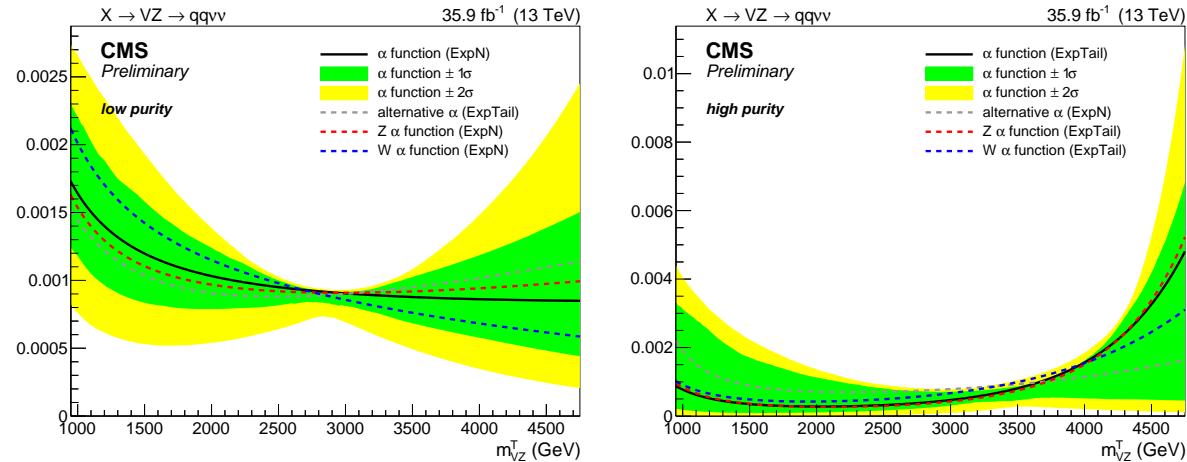


Figure 4.49: Validation of the  $\alpha$  method:  $\alpha$  functions calculated for  $Z + \text{jets}$  background (red dotted line) and for  $W + \text{jets}$  background (blue dotted line) separately, and  $\alpha$  function for the total  $V + \text{jets}$  background (black solid line). Left: low-purity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

1802 defined in the analysis.

1803 The last check performed is a study of the impact of the choice of the function to describe the  $V$   
 1804 + jets background on the very last result of the analysis, namely the exclusion limit on the signal  
 1805 cross-section times branching fraction. The procedure of the limit extraction is discussed in detail  
 1806 in sec. 4.7. The main and alternative functions chosen to parametrize the dominant background  
 1807 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	$1841.3 \pm 45.7$	1793
pseudo-SR	low-purity	$529.9 \pm 37.8$	521
SB	high-purity	$728.5 \pm 29.9$	725
pseudo-SR	high-purity	$39.3 \pm 5.2$	49

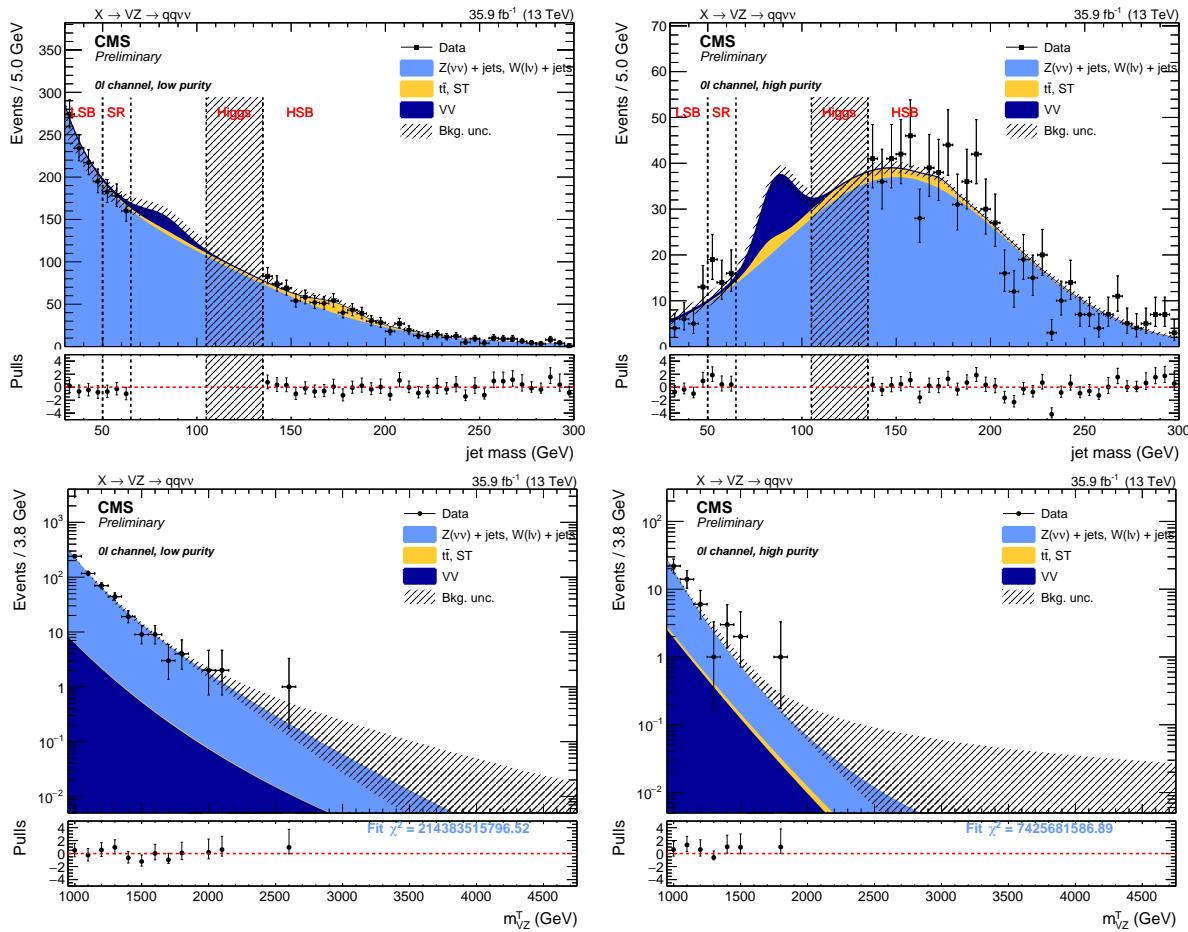


Figure 4.50: Top: results of the fit to the  $m_V$  spectrum in data, in the sidebands defined for the  $\alpha$  method validation: low-sideband ( $30 < m_V < 50$  GeV) and high-sideband ( $m_V > 135$  GeV) (left: low-purity category, right: high-purity category). Bottom: results of the fits to the  $m_{VZ}^T$  spectrum, in the pseudo-signal region ( $50 < m_V < 65$  GeV) defined for the  $\alpha$  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 ( $<< 1\%$ ).

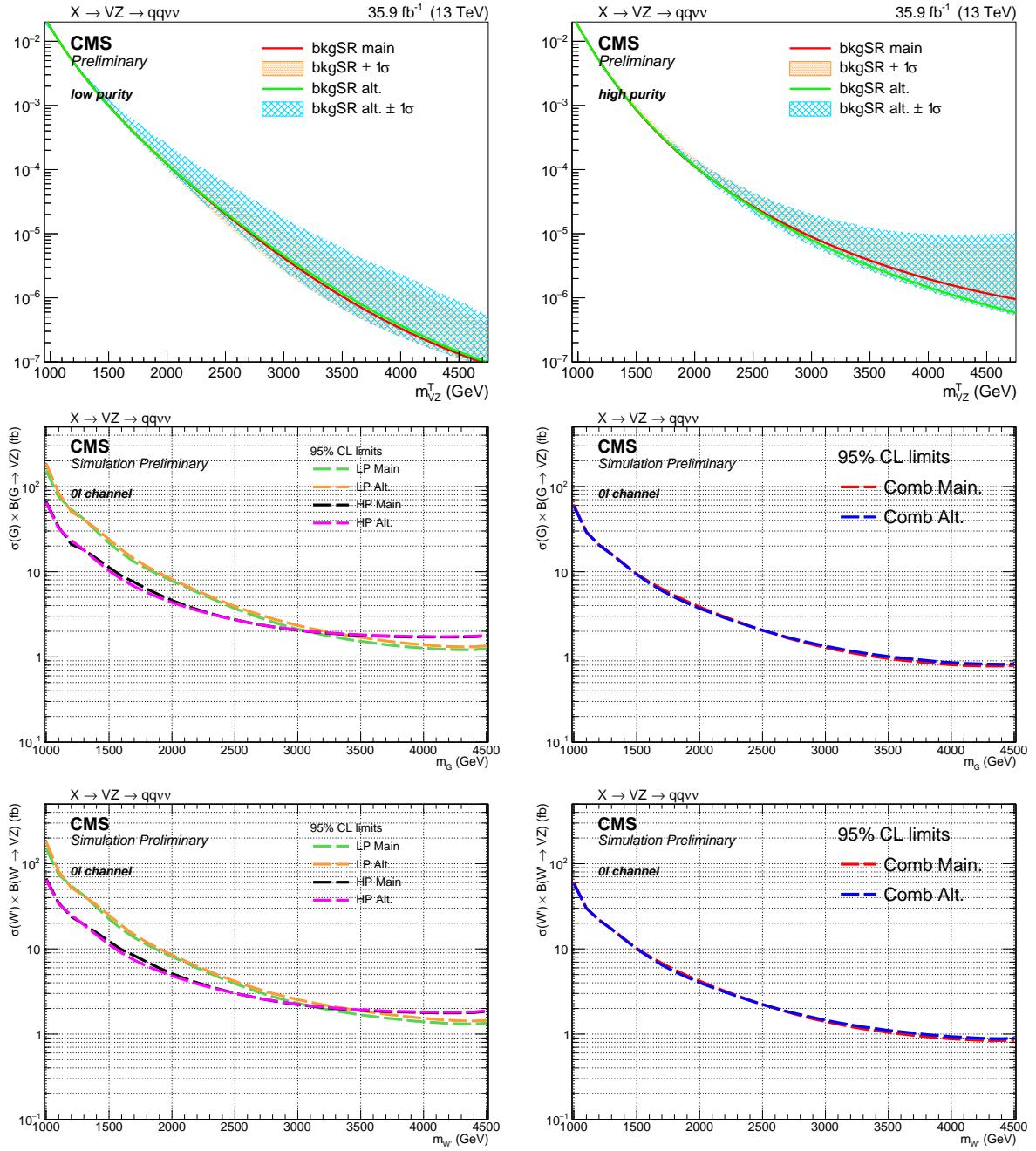


Figure 4.51: Validation of the  $\alpha$  method: impact of the choice of the function to describe the dominant  $V + \text{jets}$  background. Top: fit results of the background shape prediction in the SR obtained with the main function (red curve) and the alternative function (green curve), for low- (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

---

### 1820 4.5.4 Signal modeling

1821 The simulated signal samples, with different resonance mass hypotheses, are fitted in the SR with an  
 1822 empirical function in order to be able to perform an unbinned likelihood fit for the signal extrac-  
 1823 tion. The function chosen to model the signal samples is a *Crystal Ball* function [88, 89], which is  
 1824 composed by a gaussian-like core convoluted to two power-law tails. Both spin-2 (fig. ??) and spin-1  
 1825 (fig. 4.53) signal samples are fitted.

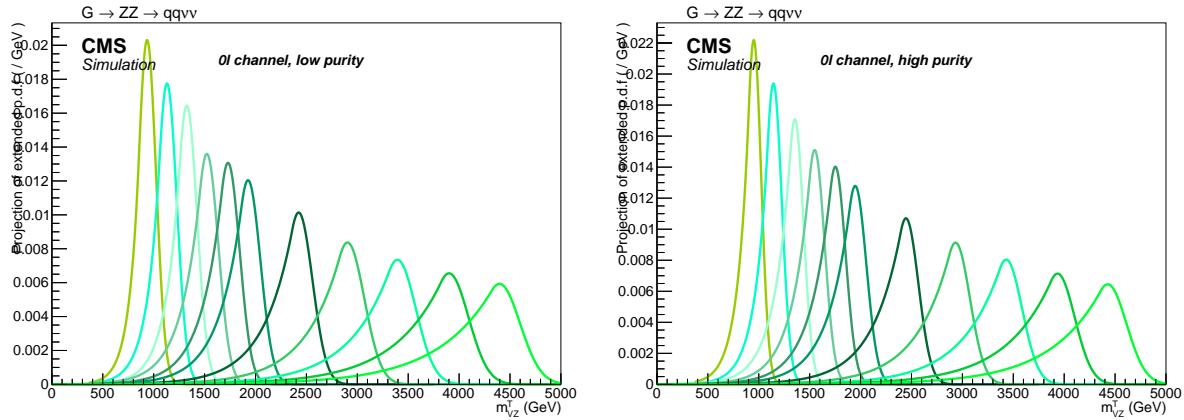


Figure 4.52: Interpolation of the signal as a function of the resonance transverse mass  $m_{VZ}^T$ , for a spin-2 (bulk graviton) signal hypothesis with an arbitrary cross section of 1 pb in the low- (left) and high-purity category (right).

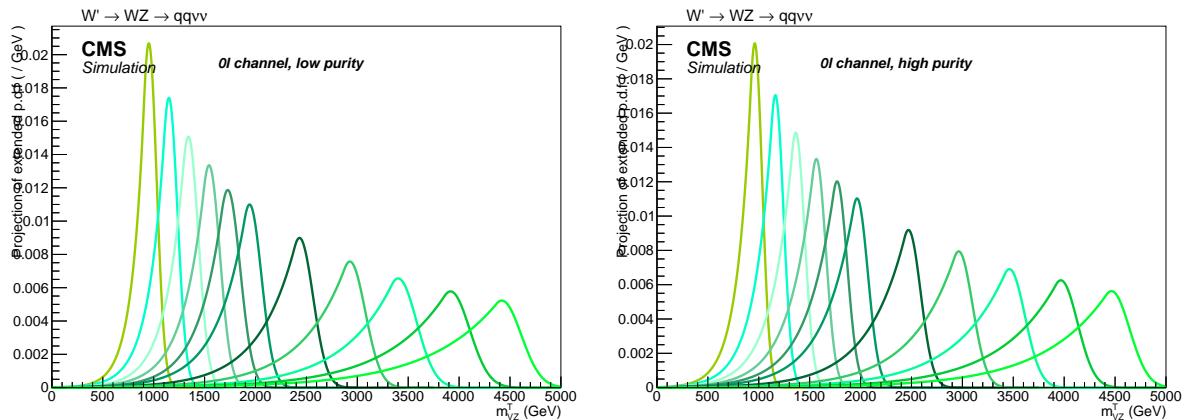


Figure 4.53: Interpolation of the signal as a function of the resonance transverse mass  $m_{VZ}^T$ , for a spin-1 ( $W'$ ) signal hypothesis with an arbitrary cross section of 1 pb in the low- (left) and high-purity category (right).

### 1826 4.5.4.1 Signal parametrization

1827 The signal is parametrized by interpolating the fitted parameters separately for each category in or-  
 1828 der to have a continuous variation of the signal shape for every possible  $m_{VZ}^T$  value within the range.  
 1829 A linear fit is performed on the mean and the width of the gaussian core of the Crystal Ball functions.  
 1830 The interpolations are shown in fig. 4.54- 4.55 for the spin-2 signal model, and in fig. 4.56- 4.57 for

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

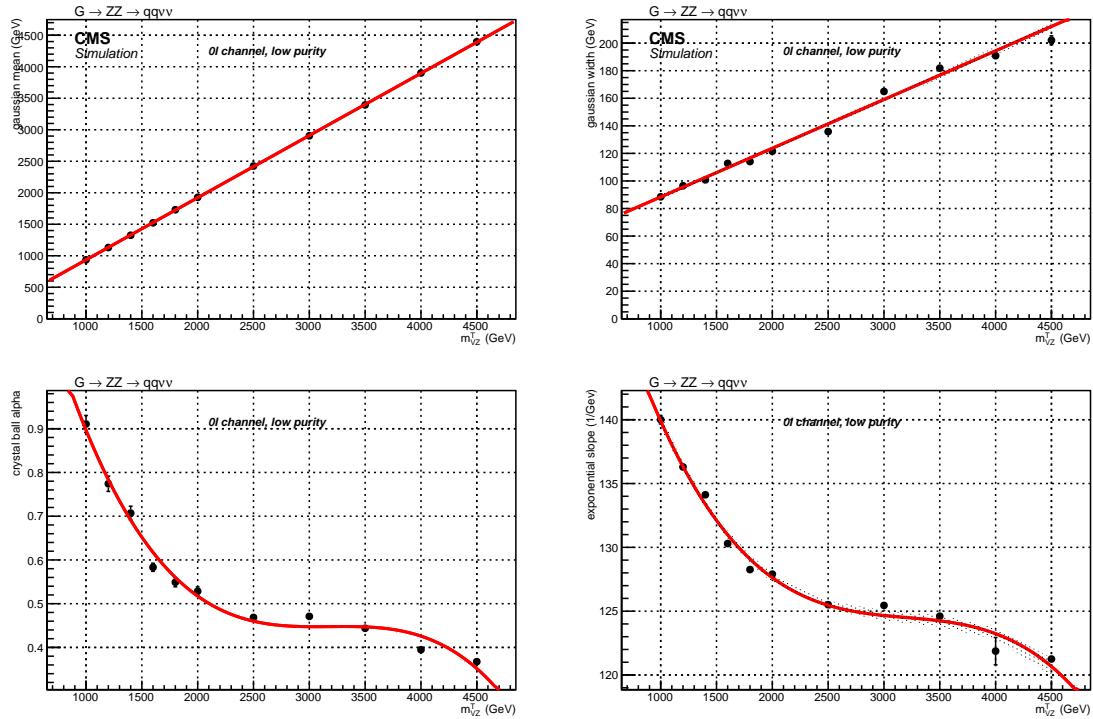


Figure 4.54: Interpolation of the fitted parameters as a function of the resonance transverse mass  $m_{VZ}^T$ , for a spin-2 (bulk graviton) signal hypothesis, low-purity category.

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

## 4.5 Background estimation technique

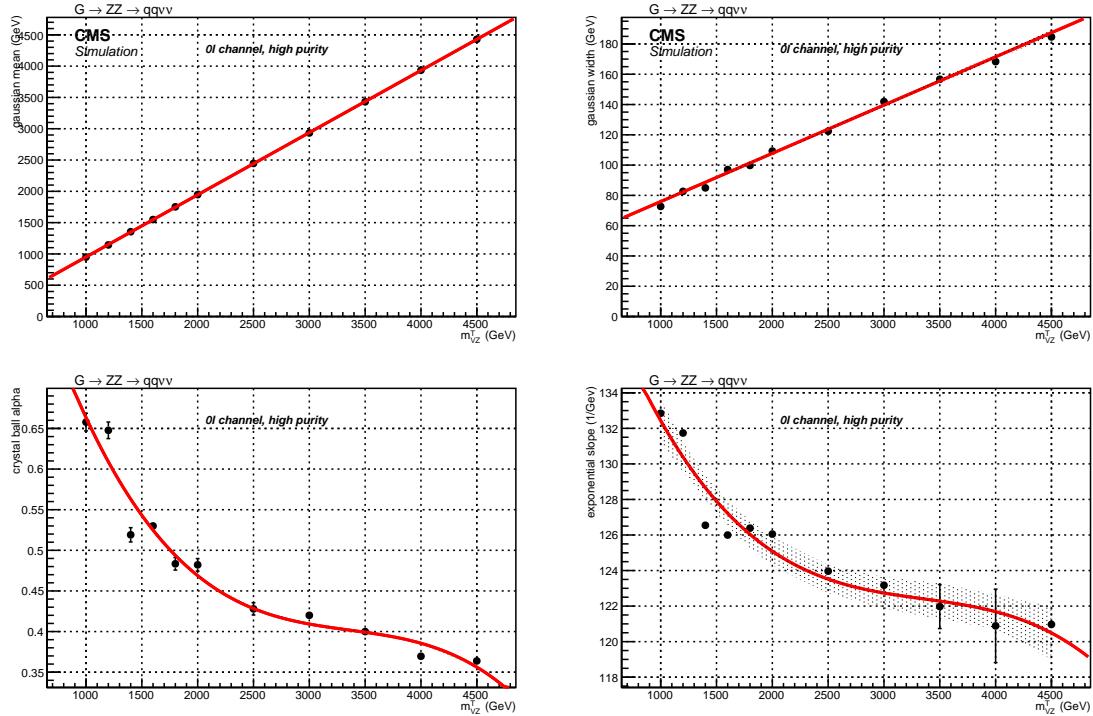


Figure 4.55: Interpolation of the fitted parameters as a function of the resonance transverse mass  $m_{VZ}^T$ , for a spin-2 (bulk graviton) signal hypothesis, high-purity category.

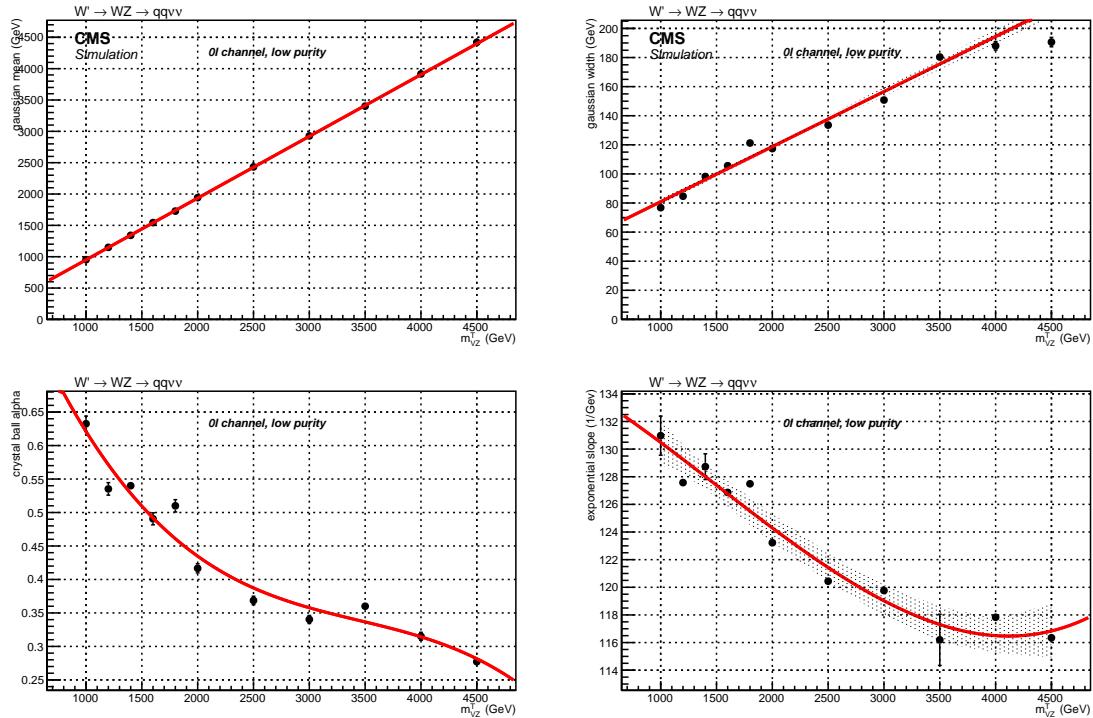


Figure 4.56: Interpolation of the fitted parameters as a function of the resonance transverse mass  $m_{VZ}^T$ , for a spin-1 ( $W'$ ) signal hypothesis, low-purity category.

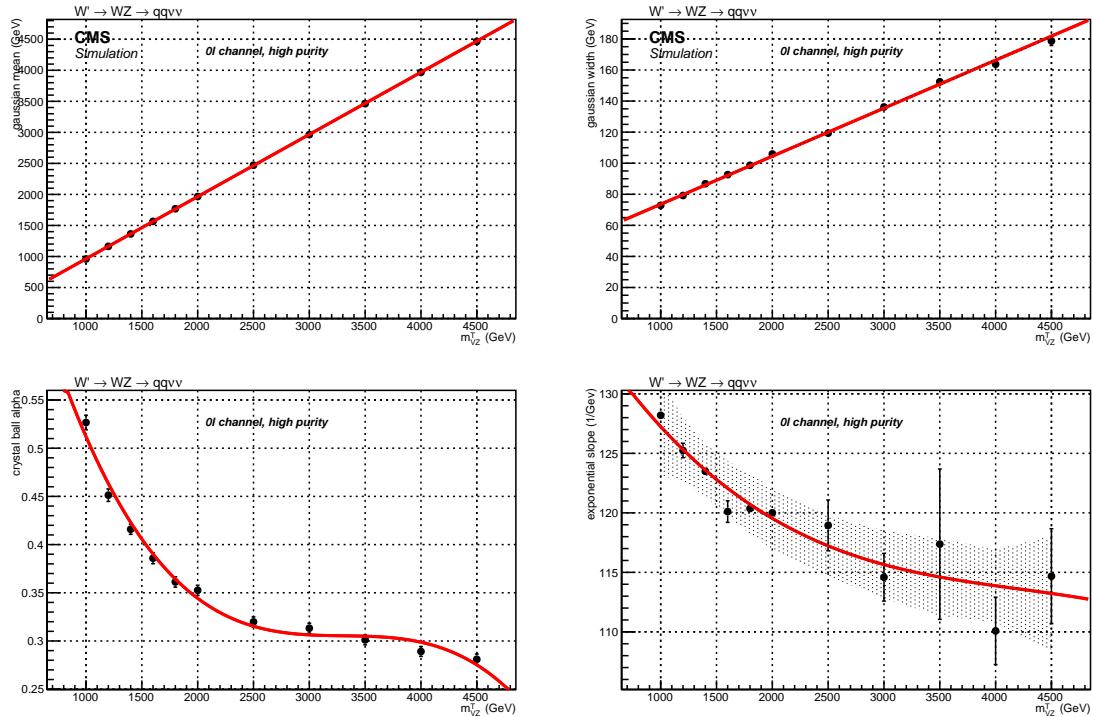


Figure 4.57: Interpolation of the fitted parameters as a function of the resonance transverse mass  $m_{VZ}^T$ , for a spin-1 ( $W'$ ) signal hypothesis, high-purity category.

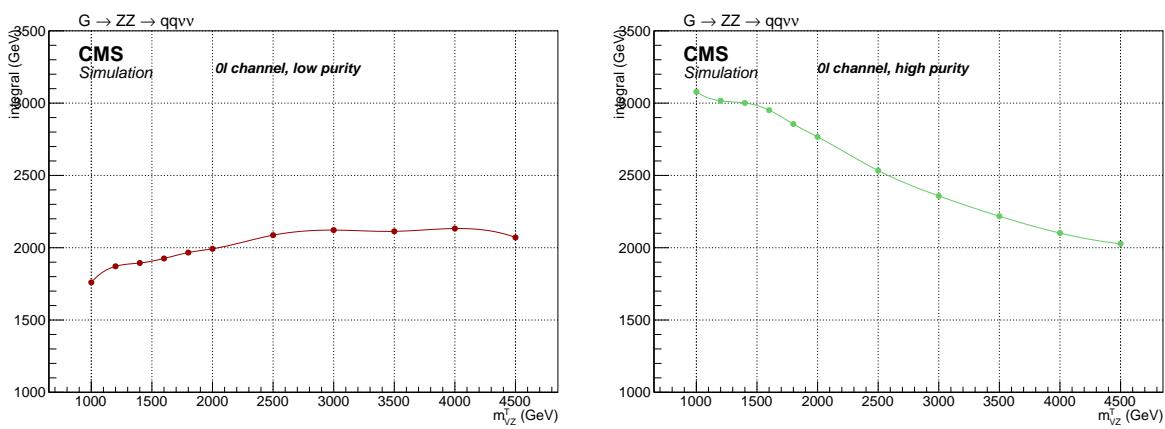


Figure 4.58: Interpolation of the signal normalization as a function of the resonance transverse mass  $m_{VZ}^T$ , for a spin-2 (bulk graviton) signal hypothesis. From left to right: low-purity, high-purity.

## 4.5 Background estimation technique

---

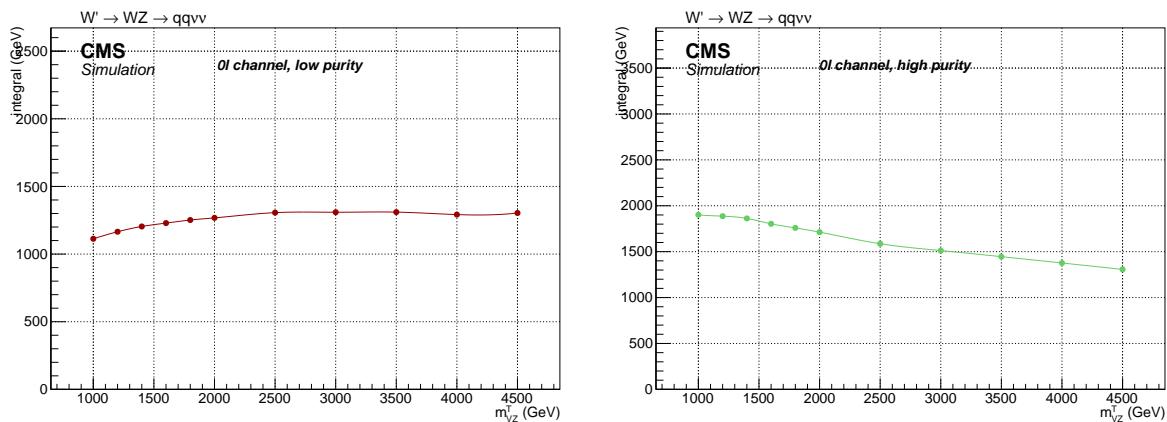


Figure 4.59: Interpolation of the signal normalization as a function of the resonance transverse mass  $m_{VZ}^T$ , for a spin-1 ( $W'$ ) signal hypothesis. From left to right: low-purity, high-purity.

---

## 1839 4.6 Systematic uncertainties

1840 The background and signal predictions are affected by systematic uncertainties that have to be es-  
 1841 timated and taken into account in the signal extraction procedure. This section includes a list of the  
 1842 relevant systematic uncertainties for this analysis and how they are estimated.

### 1843 4.6.1 Uncertainties affecting the data-driven main background estimation

#### 1844 4.6.1.1 Normalization

1845 The prediction of the normalization and shape of main background,  $V + \text{jets}$ , are both taken from  
 1846 data. The normalization is extracted from fits to the jet mass sidebands with arbitrary functions  
 1847 tested on simulation. The effects related to the contribution of the sub-dominant backgrounds are  
 1848 also taken into account, for both the normalization and the shape.

1849 The uncertainties on the sub-dominant backgrounds normalization, namely the uncertainties on  
 1850 the parameters describing the jet mass spectra obtained with the fits performed on simulations,  
 1851 are propagated to the main background yield prediction. An additional uncertainty on the main  
 1852 background yield comes from the fit with the alternative function. In this case, the difference in  
 1853 the predicted number of events due to the function choice is taken as a systematic uncertainty.  
 1854 The limited number of events in data in the sidebands is treated separately as a source of statistical  
 1855 uncertainty. Numerical values are reported in tab. 4.16.

#### 1856 4.6.1.2 Shape

1857 The shape uncertainties on the main background are determined with the  $\alpha$  method, discussed in  
 1858 sec. 4.5.2. The uncertainties on the parameters of the main background prediction in the signal  
 1859 regions are affected by the parameter uncertainties of the fit to  $m_{VZ}^T$  in data in the jet mass side-  
 1860 bands, and by the parameter uncertainties of the two components of the  $\alpha$  function (numerator  
 1861 and denominator), that are the  $m_{VZ}^T$  fits to the simulated  $V + \text{jets}$  distributions in SR and SB. These  
 1862 uncertainties are propagated to the shape of the main background in the signal region. Before being  
 1863 provided to the likelihood fit, these parameters are decorrelated through a linear transformation.

### 1864 4.6.2 Uncertainties affecting the signal and the sub-dominant backgrounds

#### 1865 4.6.2.1 Trigger uncertainty

1866 Trigger uncertainty is evaluated shifting by one standard deviation (*i.e.* 1%, as discussed in sec. 4.2.6)  
 1867 the  $E_T^{\text{miss}}$  trigger efficiency calculated on data, that is applied as per-event weight to MC samples.  
 1868 The impact has been studied in signal and secondary background samples: it amounts to 0.7–0.5%  
 1869 for signal samples, depending on the mass hypothesis, whilst it affects by 1% the top and diboson  
 1870 normalization. No effect can be appreciated in signal and background shapes.

#### 1871 4.6.2.2 Jet momentum uncertainties

1872 Jet uncertainties are evaluated in the signal regions by moving up and down by one standard devia-  
 1873 tion the source of the uncertainty. The two sources are the uncertainty on the jet energy correction,  
 1874 also identified as jet energy scale (JES) [61], [62], and the uncertainty due to the different jet mo-  
 1875 mentum resolution (JER) [62].

1876 Considering the jet energy scale, the transverse momenta of the jets are shifted by the uncertainty  
 1877 value of the corresponding jet energy correction. The impact on the normalization due to the jet

## 4.6 Systematic uncertainties

---

- 1878 energy correction is evaluated in the signal region, by taking into account its effect on jets and on  
 1879  $E_T^{\text{miss}}$  simultaneously, in a correlated fashion.  
 1880 The JER effect is evaluated (together with its impact on  $E_T^{\text{miss}}$ ) by smearing the jet  $p_T$  by the  $\eta$ -  
 1881 dependent coefficients listed in tab. 4.8, up and down by one standard deviation, using the hybrid-  
 1882 method (sec. 4.3.6).  
 1883 The impact of JEC uncertainties is evaluated also on the signal and background shapes. The result-  
 1884 ing normalization and shape uncertainties are reported in sec. 4.6.2.6.

1885 **4.6.2.3 Jet mass uncertainties**

- 1886 The soft drop PUPPI corrected jet mass is affected by two different uncertainties sources.  
 1887 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  
 1888 of the JES, described above, and  $\text{JMS}_{\text{unc.}} = 0.0094$  is a constant coefficient (4.3.7, [61], [62]). The  
 1889 impact is calculated on signal and secondary backgrounds, both in normalization and shape.  
 1890 As regarding the smearing, the soft drop PUPPI corrected jet mass of the signal samples and sub-  
 1891 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  $\Delta$  indicates the variation for each variable, due to the considered systematic uncertainty shift.

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

1892

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

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

- 1893 Results are presented in detail in tab. 4.19-4.20, for JMS and JMR uncertainties. Shape uncertainties  
 1894 on signal are evaluated as the variation in the mean and variance of the transverse mass distribution.  
 1895 Shape uncertainties on top and diboson backgrounds are quoted as the relative variation in the  
 1896 slope of the exponential falling distribution of  $m_{VZ}^T$ , and their effects are shown in fig. 4.60-4.61.

1897 **4.6.2.4 V-tagging uncertainties**

- 1898 Data-Monte Carlo V-tagging scale factors are applied to the signal and secondary background yields  
 1899 (sec. 4.3.8.1), and their uncertainty is taken as systematic. The contribution of the uncertainty is 11%

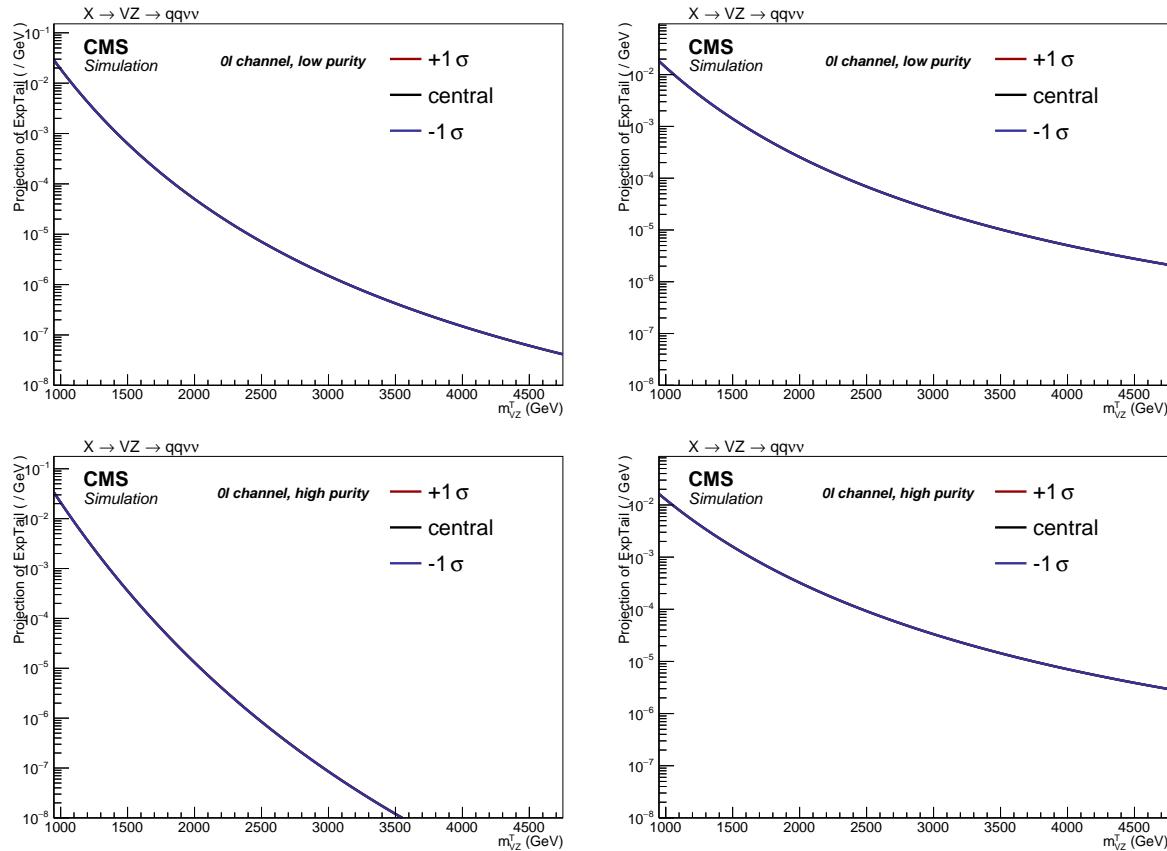


Figure 4.60: Shape variations due to jet mass calibration corrections obtained in the Top (left) and diboson (right) backgrounds, in the low-purity (top) and high-purity (bottom) category.

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 [90]). 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 [86]. 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

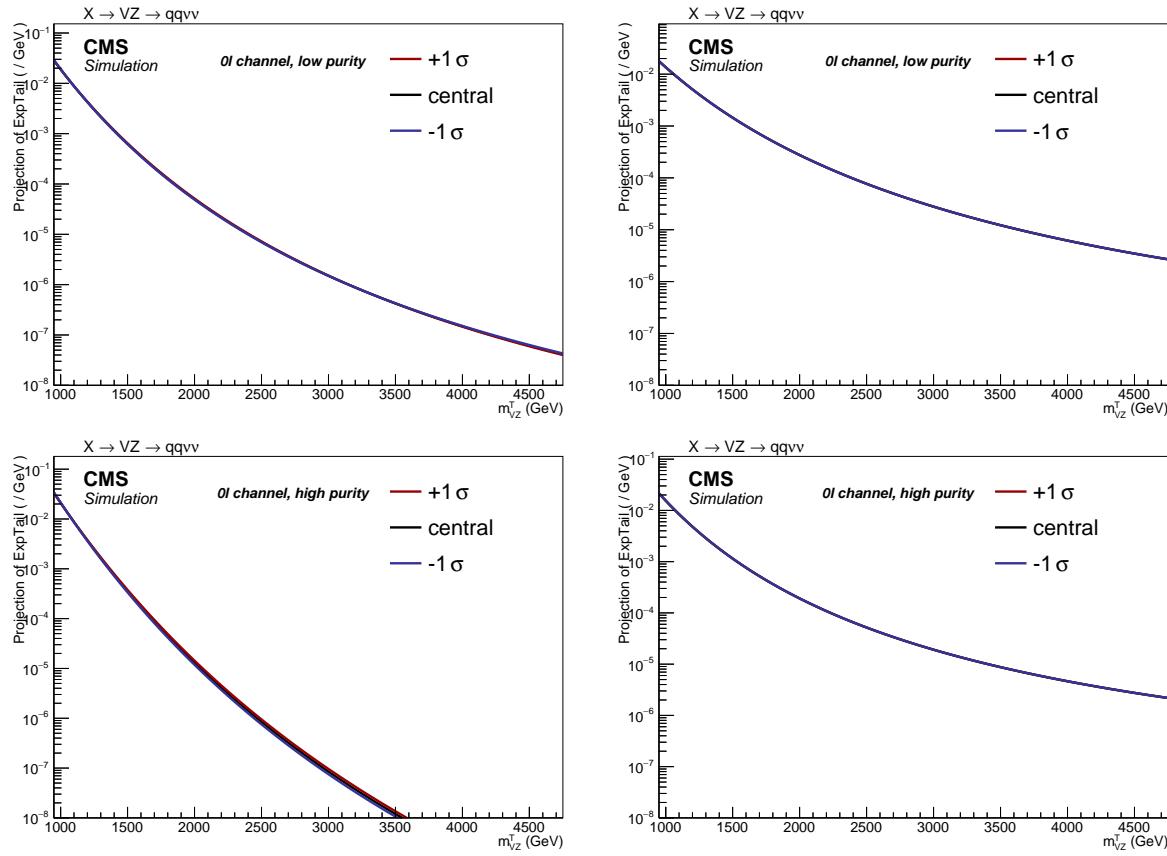


Figure 4.61: Shape variations due to jet mass resolution obtained in the Top (left) and diboson (right) backgrounds, in the low-purity (top) and high-purity (bottom) category.

### 1918 4.6.2.6 Missing Energy uncertainties

- 1919 As described in sec. 3.2.9.8, the  $E_T^{\text{miss}}$  evaluation depends on all the reconstructed particles in the  
 1920 event, and on their uncertainties. Missing energy uncertainties are calculated by factorizing  $\vec{p}_T^{\text{miss}}$   
 1921 in components: electrons, photons, muons, taus, jets and unclustered energy. Dedicated uncer-  
 1922 tainties are derived by propagating the original object scales and resolutions to the  $E_T^{\text{miss}}$  itself.  
 1923 In this analysis, a leptonic veto is applied, hence the  $E_T^{\text{miss}}$  uncertainties are due to jets and unclus-  
 1924 tered energy. The effect of JES is evaluated on  $E_T^{\text{miss}}$  in a correlated way with jets, by scaling up or  
 1925 down the central value of JES by one sigma, both on  $E_T^{\text{miss}}$  and on jets  $p_T$ . The result is a negli-  
 1926 gible uncertainty on signal normalization, 0.2% and less than 0.1% uncertainty on top and diboson  
 1927 normalizations, negligible impact on signal, top and diboson shapes.  
 1928 The same procedure applies for the uncertainties related to jet JER, that are varied up and down by  
 1929 one sigma in both jets and  $\vec{p}_T^{\text{miss}}$  at the same time. The result is a negligible uncertainty on signal and  
 1930 diboson normalizations, 0.3% uncertainty on top normalization, and negligible effects on signal  
 1931 and background shapes.  
 1932 The last contribution in  $E_T^{\text{miss}}$  uncertainty is related to unclustered energy, whose impact is eval-  
 1933 uated scaling up or down the central value by its own resolution, depending on the particle type. The  
 1934 uncertainty is negligible on signal and background normalizations and shape.

---

 1935 **4.6.2.7 Pile-up uncertainty**

1936 An additional source of systematic uncertainty is the limited knowledge of the total proton-proton  
 1937 inelastic cross-section at 13 TeV, used to get the expected number of vertices distribution for the  
 1938 pile-up reweighting procedure. A 4.6% uncertainty is assumed for the default value of 69 200 mb,  
 1939 and the vertices distributions are varied accordingly (fig. 4.62). Changing the pile-up weight varies  
 1940 also the MC normalizations in the signal region, and the relative difference is estimated to be 0.2%  
 1941 for the diboson background, 0.3% for top processes, and 0.4-0.7% for signal samples. Pile-up im-  
 1942 pacts on signal shapes are negligible, and it affects by 0.8% and 0.4% the diboson and top shapes  
 1943 (fig. 4.63).

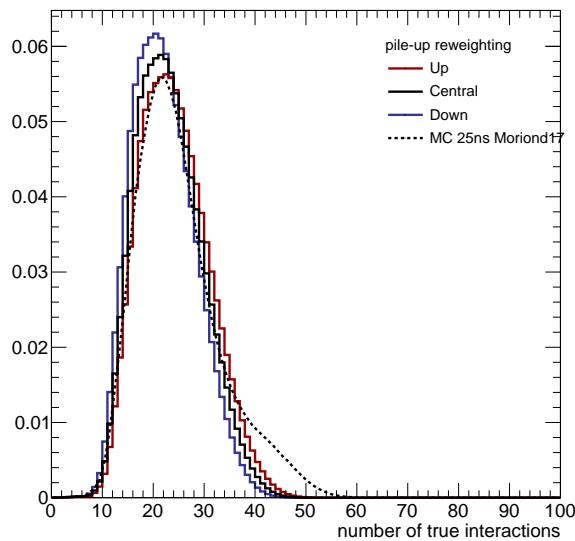


Figure 4.62: Pile-up scenario in 2016 data (black curve), and scenarios obtained by shifting up (red curve) or down (blue curve) the central value of the total inelastic cross-section (69 200 mb), compared to pile-up distribution simulated in Monte Carlo samples (dotted curve).

---

 1944 **4.6.2.8 QCD renormalization and factorization scale uncertainties**

1945 Divergencies appearing in perturbative QCD calculation, used to predict the cross-sections and the  
 1946 spectra of the observables in Monte Carlo simulations, are absorbed in the renormalization and  
 1947 factorization scales,  $\mu_R$  and  $\mu_F$ . Per-event weights are calculated for a variation of these scales by  
 1948 a factor 2. The two scales can be varied separately and independently, or together assuming 100%  
 1949 correlation; the first approach is adopted. The weight is propagated up to the final distributions,  
 1950 accounting for normalization and shape uncertainties.  
 1951 The QCD variations have negligible effect on signal acceptance and on the mean and sigma of the  
 1952 gaussian core of the Crystal Ball functions. The QCD factorization has an impact on top background  
 1953 shape (1.1%) and normalization (3.1%), and on diboson normalization (0.9%). The QCD renorm-  
 1954 alization affects the top normalization (7.3%) and diboson normalization (1.3%).

## 4.6 Systematic uncertainties

---

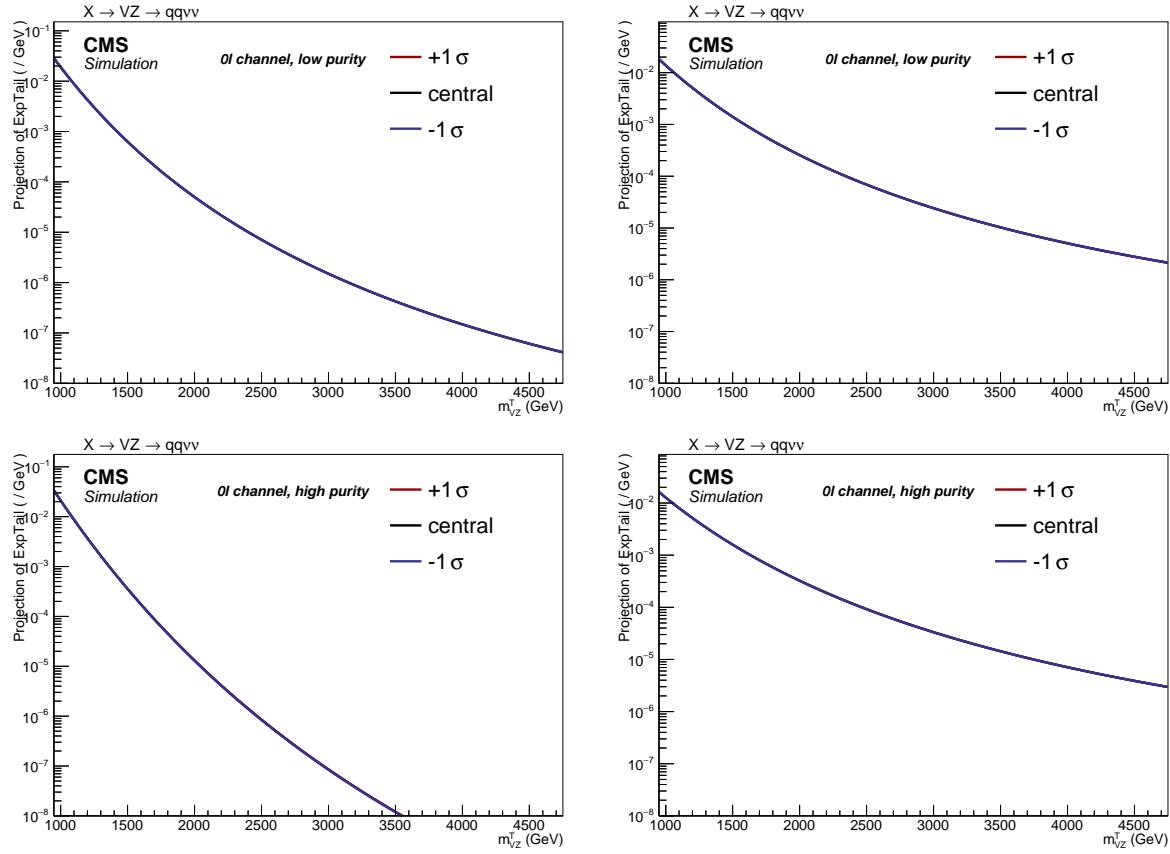


Figure 4.63: Shape variations due to pile-up uncertainty obtained in the Top (left) and diboson (right) backgrounds, in the low-purity (top) and high-purity (bottom) category.

1955 **4.6.2.9 PDF**

1956 Systematic uncertainties related to the PDFs parameters are estimated according to the PDF4LHC  
 1957 prescriptions [91], and using the NNPDF3.1 [92] set. Each parameter describing the PDFs is varied  
 1958 within its uncertainty, resulting in a set of per-event weights. The 100 shifted weights have been  
 1959 considered together, by calculating the effect of their envelope, compared to their central values,  
 1960 on the expected event yield and on the  $m_{VZ}^T$  distributions, and propagated as a normalization or  
 1961 shape uncertainty. The effect of the PDF uncertainty on the signal acceptance is found to be negli-  
 1962 gible, and it amounts to 10.3% for top background normalization and 2.1% for diboson background  
 1963 normalization. PDF uncertainties affect top background shape by 1.2%.

1964 **4.6.3 Summary**

1965 A summary of all the systematic uncertainties is listed in tab. 4.21. In addition to those described in  
 1966 the previous sections, an uncertainty of 10% on top background normalization is assumed, that is  
 1967 the uncertainty on the top production cross-sections obtained from CMS measurements (sec. 4.2.3),  
 1968 and an uncertainty of 15% is assigned to the diboson background normalization, due to the uncer-  
 1969 tainty on the cross-section measurements performed by CMS. An additional 3% covers the uncer-  
 1970 tainty related to the tau veto, and an uncertainty of 2.5% is assigned to the data integrated luminos-  
 1971 ity [81].

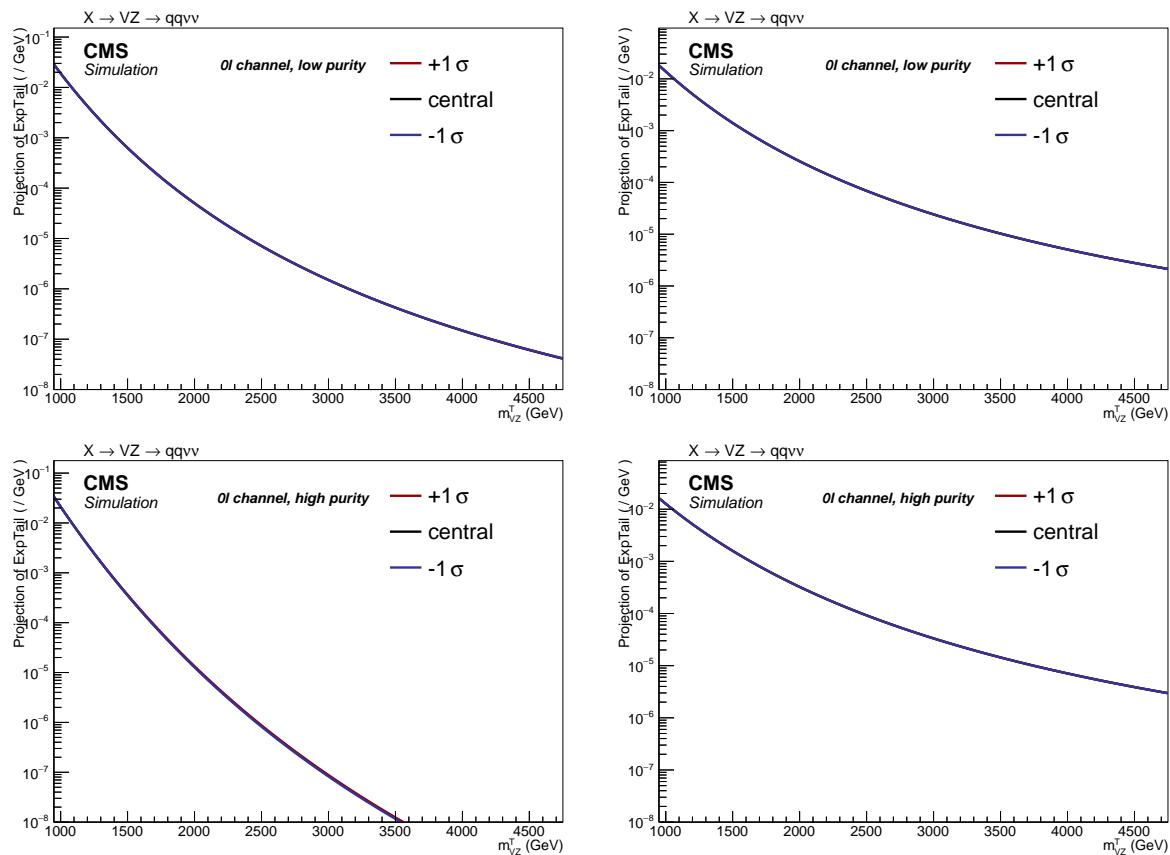


Figure 4.64: Shape variations due to QCD factorization in the Top (left) and diboson (right) backgrounds, in the low-purity (top) and high-purity (bottom) category.

## 4.6 Systematic uncertainties

---

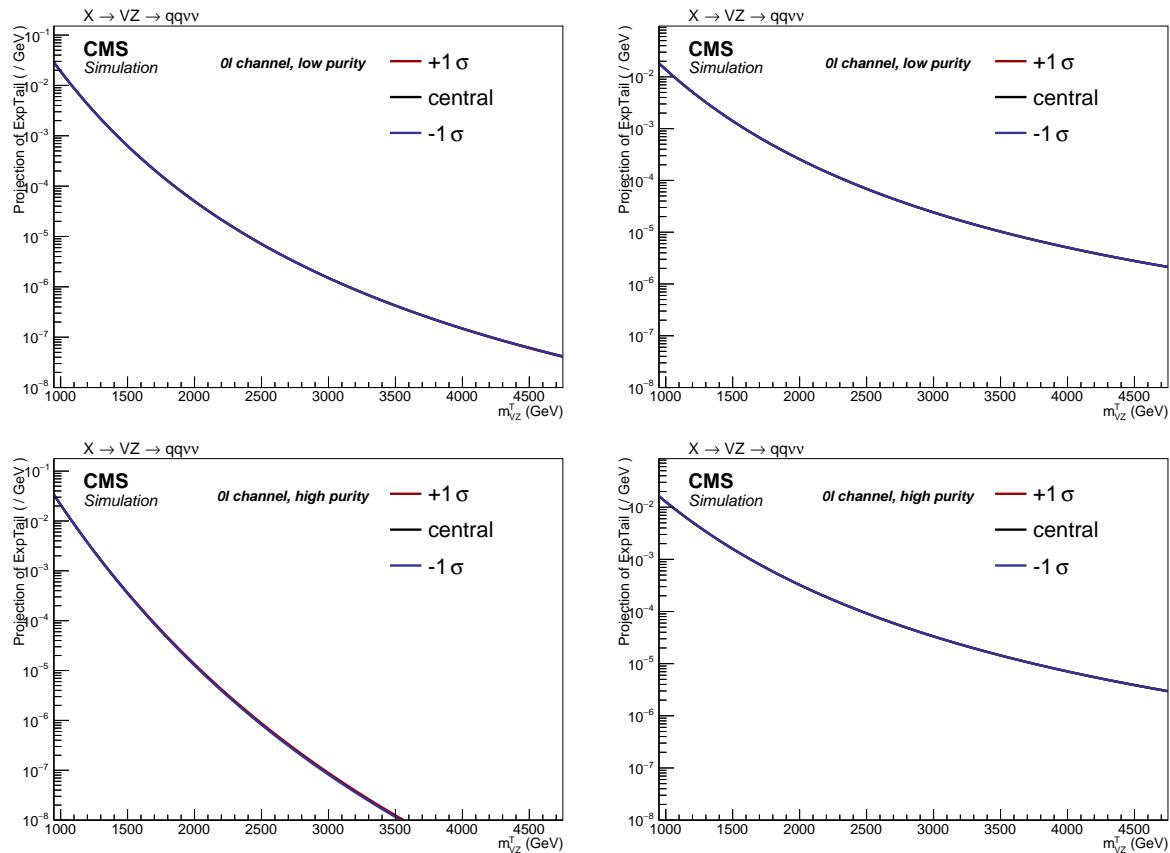


Figure 4.65: Shape variations due to PDF scale in the Top (left) and diboson (right) backgrounds, in the low-purity (top) and high-purity (bottom) category.

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
$\alpha$ -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 ( $\tau_{21}$ )	-	-		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%

### 1972 4.7 Results and interpretation

#### 1973 4.7.1 Statistical approach

##### 1974 4.7.1.1 The modified frequentist approach: asymptotic formulae to extract an upper limit

1975 The modified frequentist approach, also known as  $CL_s$  criterion [93–95], is used to determine the  
1976 95% confidence level upper limit on the signal contribution in the data.

1977 The main parameters describing the method are the background event yield,  $b$ , the signal event  
1978 yield, predicted by the theoretical model,  $s$ , the signal strength modifier  $\mu$ , parametrizing how much  
1979 the signal yield deviates from the model expectation  $s$ , and the nuisance parameters  $\theta$ , namely, the  
1980 uncertainties affecting the signal and background yields, that can be seen as functions  $b(\theta)$ ,  $s(\theta)$ .

1981 The uncertainties are considered as fully correlated (100%) or uncorrelated.

1982 The likelihood function is built starting from a Poissonian probability density function:

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

1983 where data can either be real or generated pseudo-data. For an unbinned likelihood, where  $k$  events  
1984 have been observed,

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

1985 where  $f_s$  and  $f_b$  are the probability density functions for signal and background for an observable  
1986  $x$ , and  $S$  and  $B$  are the total expected signal and background event yields.

1987 The measurement of the compatibility of data with the signal plus background or the background  
1988 only hypotheses is performed by defining a likelihood ratio test statistics  $\tilde{q}_\mu$  [96],

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

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

1989 The quantities  $\hat{\mu}$  and  $\hat{\theta}$  are global maximums of the likelihood, while  $\hat{\theta}_\mu$  is the conditional maximum,  
1990 given  $\mu$ . The signal strength  $\hat{\mu}$  is defined positive, the upper boundary  $\hat{\mu} \leq \mu$  is set in order  
1991 to avoid to consider upward fluctuations in data (namely, when the global maximum is larger than  
1992 the hypothesis  $\mu$ ) as an incompatibility with the signal hypothesis ( $\mu$ ).

1993 Given the  $\mu$  hypothesis, the test statistic value is measured on data, and labelled as  $\tilde{q}_\mu^{\text{obs}}$ . Parameters  
1994  $\hat{\theta}_0^{\text{obs}}$  and  $\hat{\theta}_\mu^{\text{obs}}$  are calculated by maximizing the likelihood function 4.18. Toy Monte Carlo  
1995 pseudo-data are then generated to build the probability density functions  $f(\tilde{q}_\mu | \mu, \hat{\theta}_\mu^{\text{obs}})$  (signal with  
1996  $\mu$  strength hypothesis) and  $f(\tilde{q}_\mu | 0, \hat{\theta}_0^{\text{obs}})$  (background-only hypothesis). Nuisance parameters are fixed  
1997 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 evaluate  $\tilde{q}_\mu$ .

1999 The p-values associated to signal plus background and background-only hypotheses are defined as:

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

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

2000 The  $CL_s$  is defined as the ratio of the p-values:

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

2001 Given the a-priori confidence level  $\alpha$ , if  $CL_s \leq \alpha$ , a model with signal strength  $\mu = 1$  is excluded at  
2002  $(1 - \alpha)$  confidence level (CL). The 95% CL *observed* upper limit on the theoretical model is set by  
2003 extracting  $\mu$  from the equation  $CL_s = 0.05$ .

2004 Similarly to the observed limit, an upper *expected* limit, along with the  $\pm 1\sigma$  and  $\pm 2\sigma$  uncertainty  
2005 bands, can be extracted by generating pseudo-data under the background-only hypothesis, and by  
2006 calculating the  $CL_s$  and 95% upper limit for each one. A cumulative distribution is then constructed:  
2007 the 50% quantile corresponds to the median expected, the 2.5%, 16%, 84%, 95.5% quantiles corre-  
2008 spond respectively to  $-2\sigma, -1\sigma, +1\sigma, +2\sigma$  uncertainty bands.

2009 Generating a large number of pseudo-data, however, can be a very expensive computational effort.  
2010 This problem is overcome by profiting of asymptotic formulae [96], derived through Wilk's [97] and  
2011 Wald's [98] theorems. The set of pseudo-data is replaced by only one dataset, the Asimov dataset:  
2012 it corresponds to a dataset where the statistical fluctuations are suppressed, and hence every pa-  
2013 rameter is set to its expectation value. These values are then equivalent to the outcomes of a large  
2014 sample of Monte Carlo simulations. The *expected* limit can therefore be calculated from the Asimov  
2015 dataset.

2016 By using the asymptotic formulae, the distribution of the test statistic  $\tilde{q}_\mu$  is given by:

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

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

2017 where the test statistic  $\tilde{q}_{\mu,A}$  is evaluated in the Asimov dataset. Once defined the inverse of the cu-  
2018 mulative Gaussian distribution  $\Phi$ , the asymptotic expression of the  $CL_s$  simplifies into:

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

2019 The expected upper limits and its  $N$  uncertainty bands are given by:

$$\begin{aligned} \mu_{up} &= \sigma \cdot \Phi^{-1}(1 - 0.5\alpha), \\ \mu_{up+N} &= \sigma \cdot [\Phi^{-1}(1 - \alpha\Phi(N)) + N]. \end{aligned} \quad (4.25)$$

2020 **4.7.1.2 Treatment of the systematic uncertainties**

2021 The nuisance parameters  $\theta$ , introduced to describe the systematic uncertainties, are expected to  
2022 have their own probability density function,  $\rho(\theta)$ , called *prior*, that is inferred by an additional set  
2023 of measurements  $\tilde{\theta}$ , used to define the mean, the shape and the width of each uncertainty. The dis-  
2024 tribution of the priors depend on the type of uncertainty considered. Flat priors (namely, a constant  
2025 value) are assigned to nuisances a-priori unconstrained; gaussian priors are assigned to nuisances  
2026 allowed to assume both negative and positive values; log-normal priors are used to positively de-  
2027 fined nuisances (such as cross-sections, efficiencies, luminosity, scale factors). For the purpose of  
2028 this search, log-normal priors are being adopted. Partially correlated uncertainties, *i.e.* those asso-  
2029 ciated to the  $\alpha$  method parameters, are decorrelated through linear transformations.

## 4.7 Results and interpretation

---

### 2030 4.7.1.3 Computation of local p-values

2031 The discovery of a signal can be inferred on data by an observation of a p-value that is incompatible  
2032 with the background-only hypothesis. The discovery test statistics is defined as:

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

2033 The boundary  $\hat{\mu} \geq 0$  is motivated by the fact that an underfluctuation of the background is not  
2034 considered as an evidence against the background-only hypothesis. The distribution  $f(q_0 | 0, \hat{\theta}_0^{obs})$   
2035 is again built with pseudo-data, generated under the background-only hypothesis with nuisances  
2036  $\hat{\theta}_0^{obs}$ . The exact p-value is therefore:

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

2037 that can be reverted into a significance  $Z$ , once the convention of the one-sided Gaussian tail is  
2038 adopted:

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

2039 By taking advantage of the Wilk's theorem, the p-value can be approximated as:

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

2040 Since the p-value depends on the phase-space considered (specifically, on the resonance mass hy-  
2041 potesis), eq. 4.27 is known as the *local p-value*. A scan of the local p-values indicates a local de-  
2042 parture from the background-only hypothesis. In case of a local excess, the global significance is  
2043 computed by correcting the local significance with trial factors, that takes into account the so-called  
2044 *look-elsewhere* effect [99], namely, the probability to observe the same excess anywhere in the whole  
2045 mass range.

### 2046 4.7.2 Signal extraction strategy for the analysis

2047 The background prediction, estimated with the  $\alpha$  method (sec. 4.5), the signal parametrization  
2048 (sec. 4.5.4.1), and the observed data are used as inputs for the signal extraction procedure. An un-  
2049 binned maximum likelihood fit is performed on each purity category, and on the combination of  
2050 the categories, in order to present a global limit on the production cross-section times branching  
2051 ratio, that is the parameter describing the signal yield and defining the signal strength  $r$  (equivalent  
2052 to the signal strength  $\mu$  described in the previous section).

#### 2053 4.7.2.1 Fit diagnostics: nuisances pulls and impacts

2054 The systematic uncertainties, treated as log-normal nuisance parameters, are allowed to vary around  
2055 their nominal values and are profiled during the maximum likelihood estimation of the signal strength.  
2056 As a diagnostic, the profiled values (post-fit) of the nuisance parameters  $\hat{\theta}$  are compared to their  
2057 a-priori expectations (pre-fit)  $\theta_0$ , in unities of the width of the gaussian core of the nuisance param-  
2058 eter  $\Delta\theta$ . The quantities  $(\hat{\theta} - \theta_0)/\Delta\theta$  are called nuisance pulls, and they have been computed both  
2059 in the background-only hypothesis (blue bars) and in the signal plus background hypothesis (green  
2060 bars), for the low- (fig. 4.66) and high-purity (fig. 4.67) categories. In fig. 4.66-4.67, the signal of a  
2061 spin-2 bulk graviton with a mass of 3 TeV is considered. The distribution of pulls does not show any

2062 anomaly, since pulls are centered around zero (no discrepancies with the a-priori expectations) and  
2063 their widths are around one (no strong deviations from the original assumption on the width of the  
2064 nuisance distributions), for both the background and signal hypothesis. The only pulls with a mean  
2065 value a bit shifted from zero or with a width smaller than one are related to  $\alpha$ -method parameters,  
2066 that are under controls.

2067 Impacts are defined as the shifts induced in the signal strength ( $r$ , the cross-section times branching  
2068 fraction in this case) as the nuisance parameter  $\theta$  is fixed and brought to its  $+1\sigma$  or  $1\sigma$  post-fit val-  
2069 ues, while all the other nuisance parameters are simultaneously profiled as log-normal. In fig. 4.68,  
2070 impacts are calculated by combining the two purity categories, assuming a signal hypothesis of a  
2071 spin-2 bulk graviton of mass 2.5 TeV. As expected a-priori (sec. 4.6), the most relevant systematic im-  
2072 pacting the determination of the signal strength is represented by the uncertainty on the  $V$ -tagging  
2073 procedure. No pathological behaviour can be observed.

## 4.7 Results and interpretation

---

Figure 4.66: Nuisance pulls for the low-purity category, in 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.

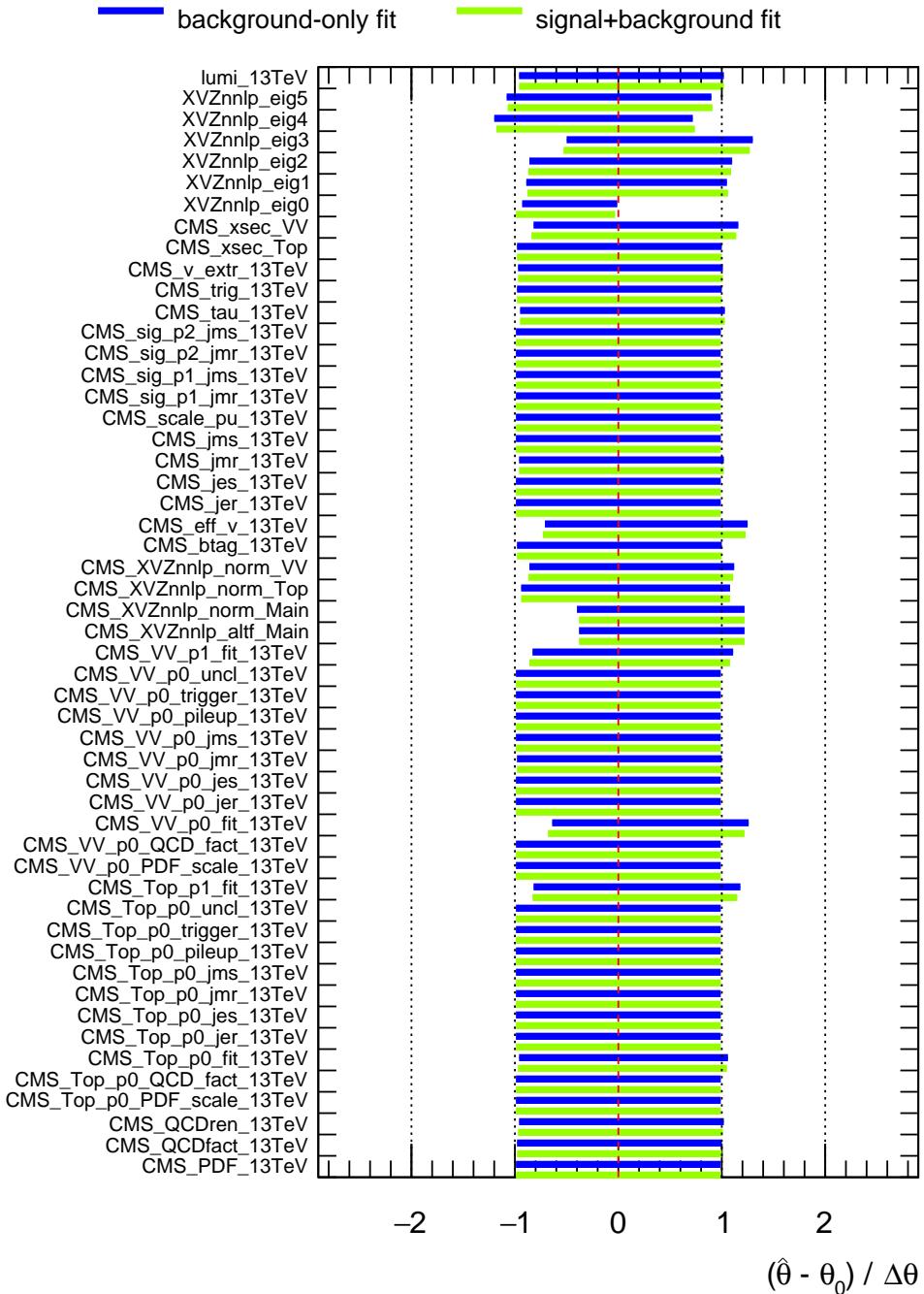
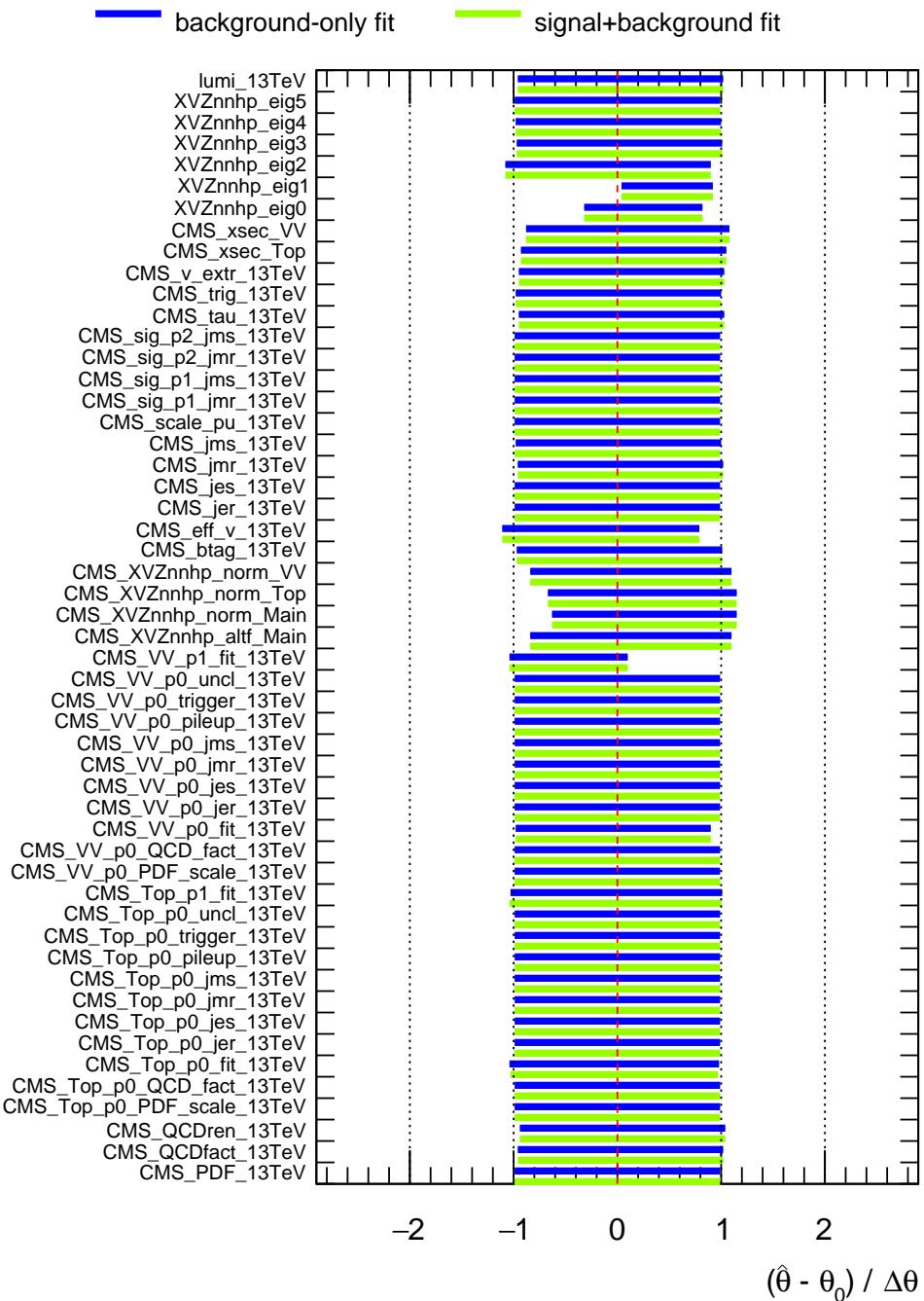


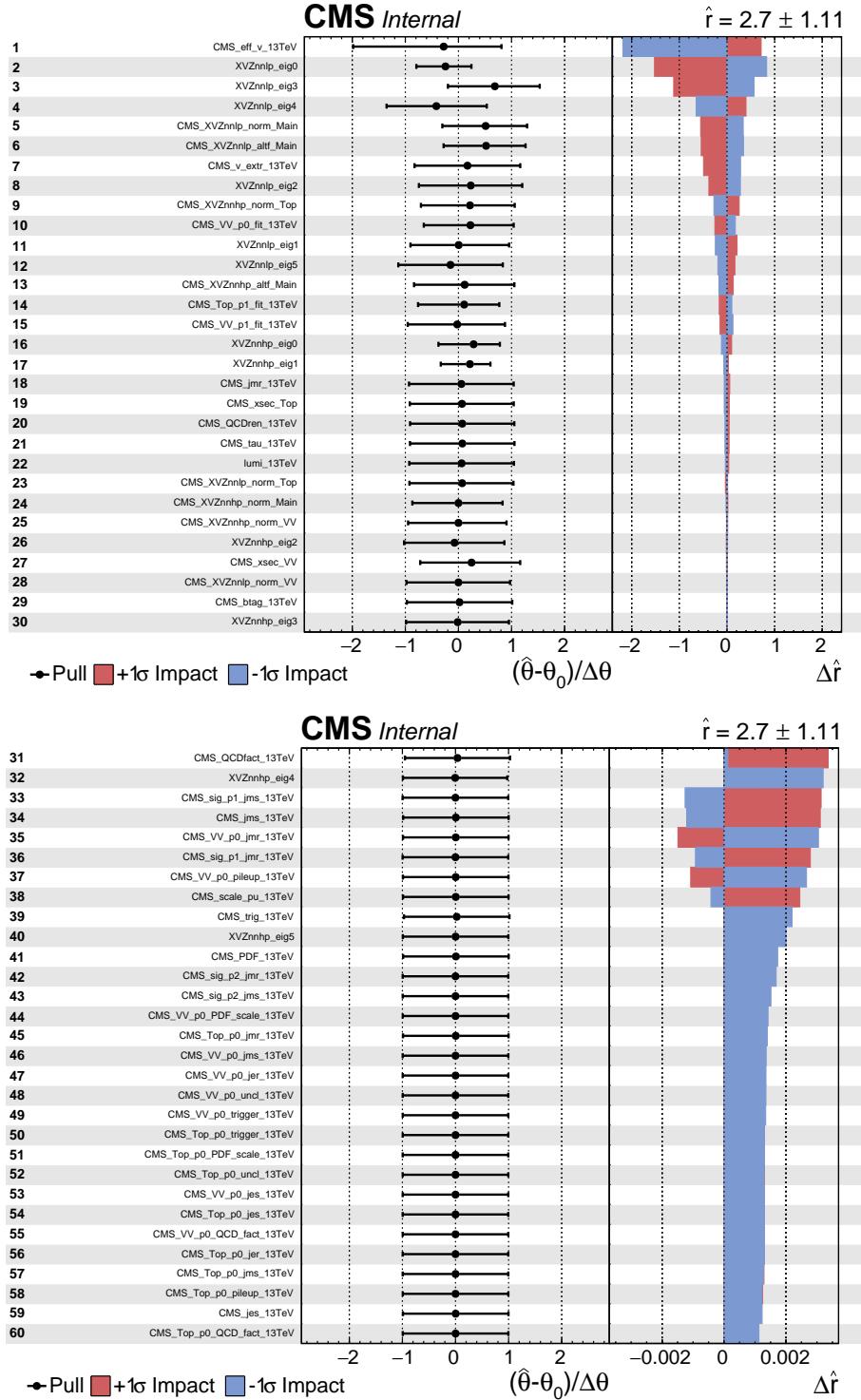
Figure 4.67: Nuisance pulls for the high-purity category, in both the background-only (blue bars) and signal plus background hypotheses (green bars). A signal hypothesis of a spin-2 bulk graviton of mass 3 TeV is considered.



## 4.7 Results and interpretation

---

Figure 4.68: Impacts of the nuisance parameters on the  $\sigma \times \mathcal{B}$  estimation, for the combination of the purity categories. A signal hypothesis of a spin-2 bulk graviton of mass 3 TeV is considered.  $\theta_0$  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  $\theta$  parameter is fixed and brought to its  $+1\sigma$  or  $-1\sigma$  post-fit values, with all other parameters profiled as log-normal.



## 2074 4.7.2.2 Results: expected and observed limits

- 2075 The observed upper limit on the resonance cross-section times branching fraction  $\sigma \mathcal{B}(X \rightarrow V_{\text{had}} Z_{\text{inv}})$ ,  
 2076 as well as the expected limit and its relative 68% and 95% uncertainty bands, are reported as a func-  
 2077 tion of the resonance mass. The limits are obtained by considering separately a spin-2 bulk gravi-  
 2078 ton and a spin-1 ( $W'$ ) heavy resonances in the narrow-width approximation. For the spin-2 case  
 2079 (fig. 4.69), data are compared to theoretical predictions on  $\sigma \mathcal{B}(G \rightarrow Z_{\text{had}} Z_{\text{inv}})$  obtained by impos-  
 2080 ing a curvature parameter of the fifth extra-dimension  $\tilde{k} = 0.5$  (red curve) and  $\tilde{k} = 1.0$  (blue curve).  
 2081 In case of spin-1 hypothesis (fig. -4.70), HVT model A (red curve) and model B (blue curve) theoretical  
 2082 predictions on  $\sigma \mathcal{B}(W' \rightarrow W_{\text{had}} Z_{\text{inv}})$  are reported.  
 2083 No significant excess is observed in data with respect to the background-only hypothesis. A spin-1  
 2084  $W'$  in the model A scenario ( $g_V = 1$ ) is excluded up to a mass of 3.11 TeV. A spin-1  $W'$  in the model  
 2085 B scenario ( $g_V = 3$ ) is excluded up to a mass of 3.41 TeV. A spin-2 bulk-graviton, once assumed a  
 2086 curvature parameter  $\tilde{k} = 0.1$ , is excluded up to 1.14 TeV.

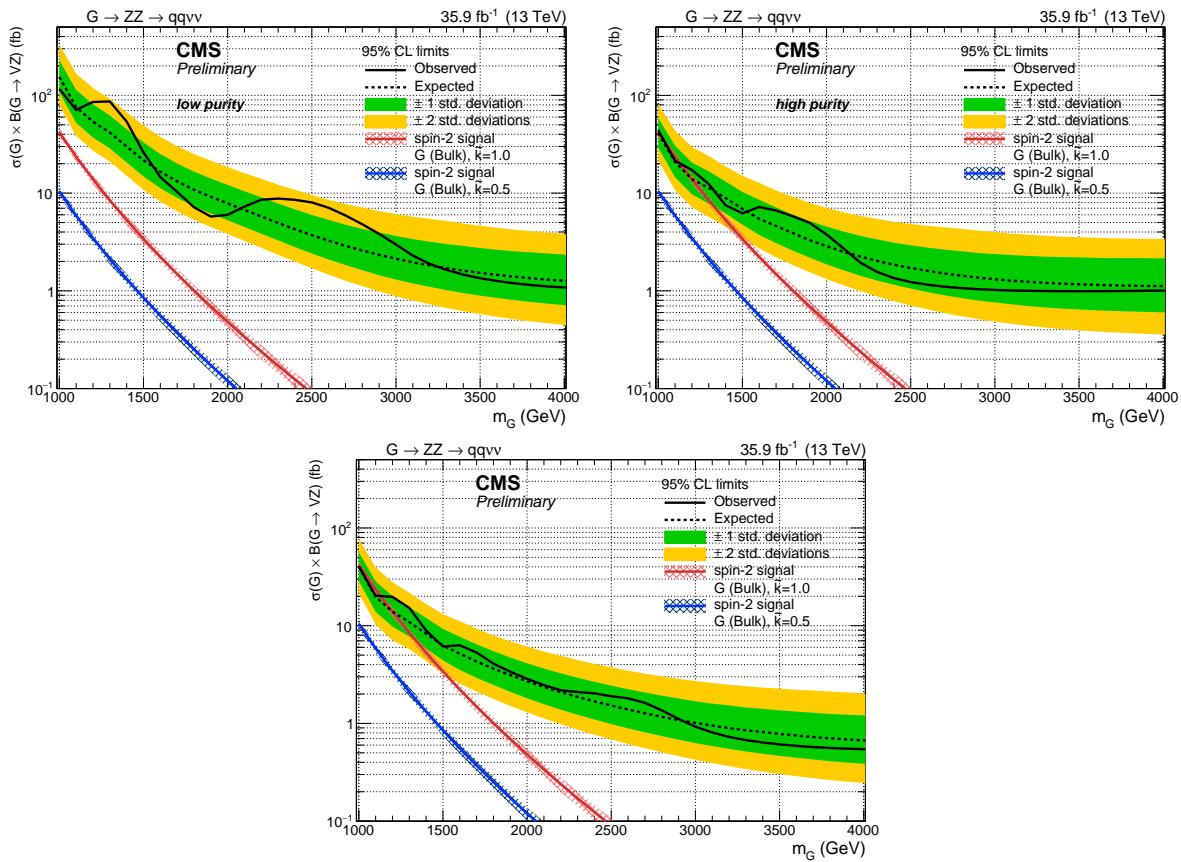


Figure 4.69: Top: observed and expected (with  $\pm 1(2)\sigma$  band) 95% C.L. upper limit on  $\sigma \mathcal{B}(G \rightarrow Z_{\text{had}} Z_{\text{inv}})$  for a spin-2 (Bulk Graviton) signal, low purity (left) and high purity (right) categories, including all statistical and systematics uncertainties.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

## 4.7 Results and interpretation

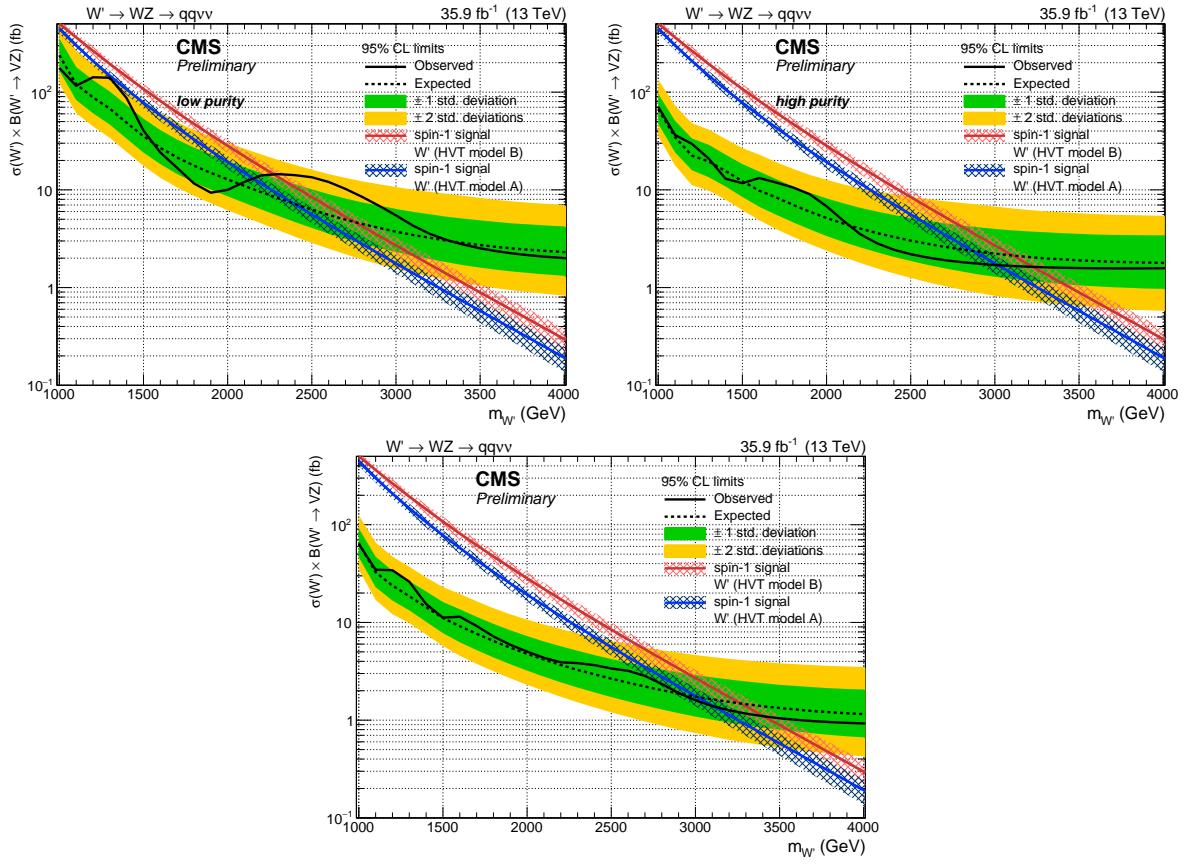


Figure 4.70: Top: observed and expected (with  $\pm 1(2)\sigma$  band) 95% C.L. upper limit on  $\sigma \mathcal{B}(W' \rightarrow W_{\text{had}} Z_{\text{inv}})$  for a spin-1 (HVT) signal, low purity (left) and high purity (right) categories, including all statistical and systematics uncertainties.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

2087 **4.7.2.3 Results: local p-value scan**

2088 Scans of the local significance (left plots) and of the local p-values (right plots), as a function of  
 2089 the resonance mass, are presented in fig. 4.71 (spin-1 signal) and in fig. 4.72 (spin-2 signal). No  
 2090 significant deviation is observed with regards to the background only hypothesis.

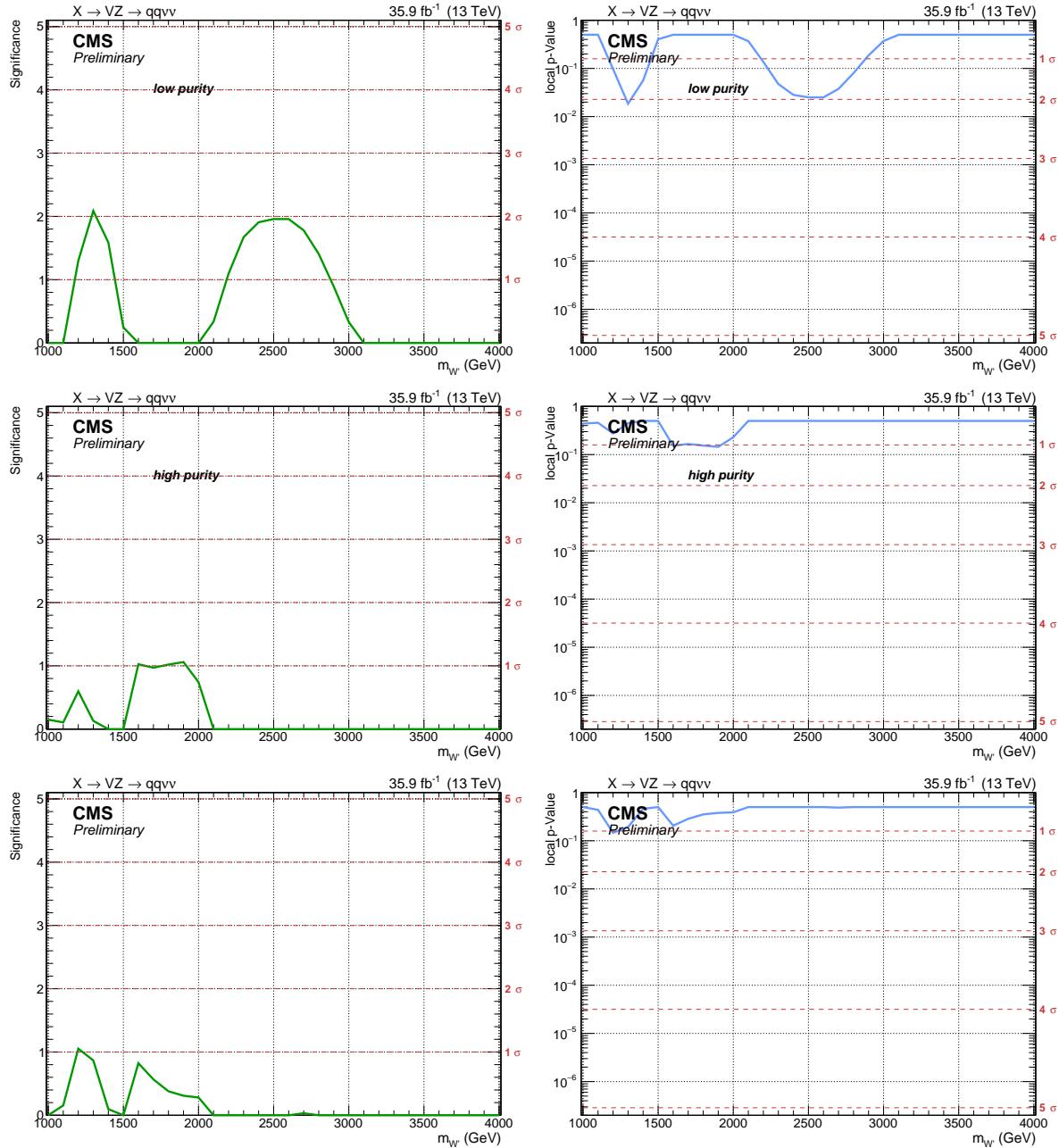


Figure 4.71: Local significances (left plots) and local p-values (right plots) as a function of the resonance mass, for a spin-1  $W'$  hypothesis, in the low- (top), high-purity categories (center), and in the combination of the categories (bottom).

## 4.7 Results and interpretation

---

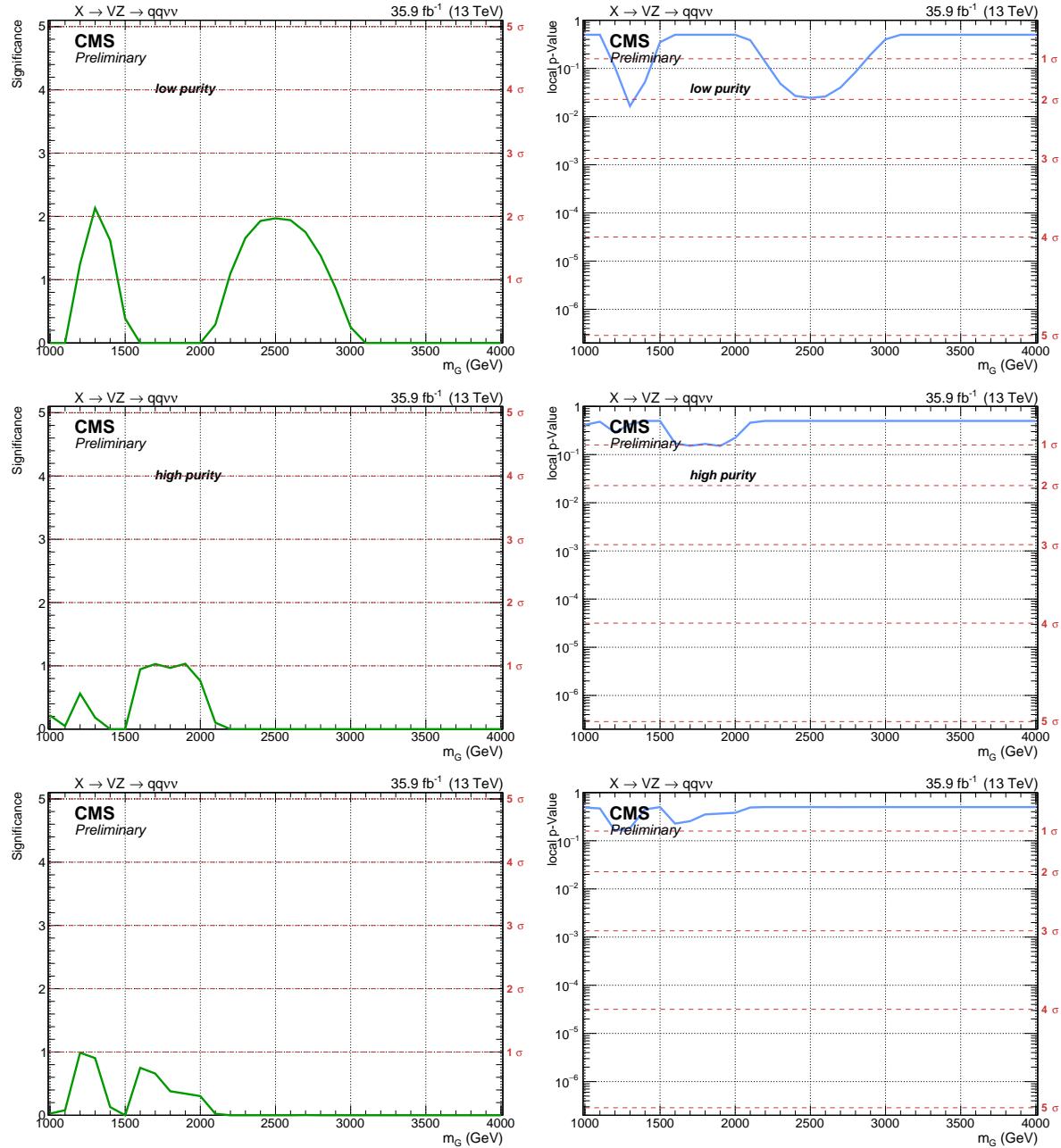


Figure 4.72: Local significances (left plots) and local p-values (right plots) as a function of the resonance mass, for a spin-2 bulk graviton hypothesis, in the low- (top), high-purity categories (center), and in the combination of the categories (bottom).

### 4.7.3 Interpretation of the results in the HVT model

For the HVT 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 triplets 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  $(g_V = 3, c_H = 0.976, c_F = 1.024)$  [9]. This search is sensitive to the charged components of the vector triplet, namely to the  $(W^{+'}, W^{-'})$  doublet. 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 parameters excluded by the observations on 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).

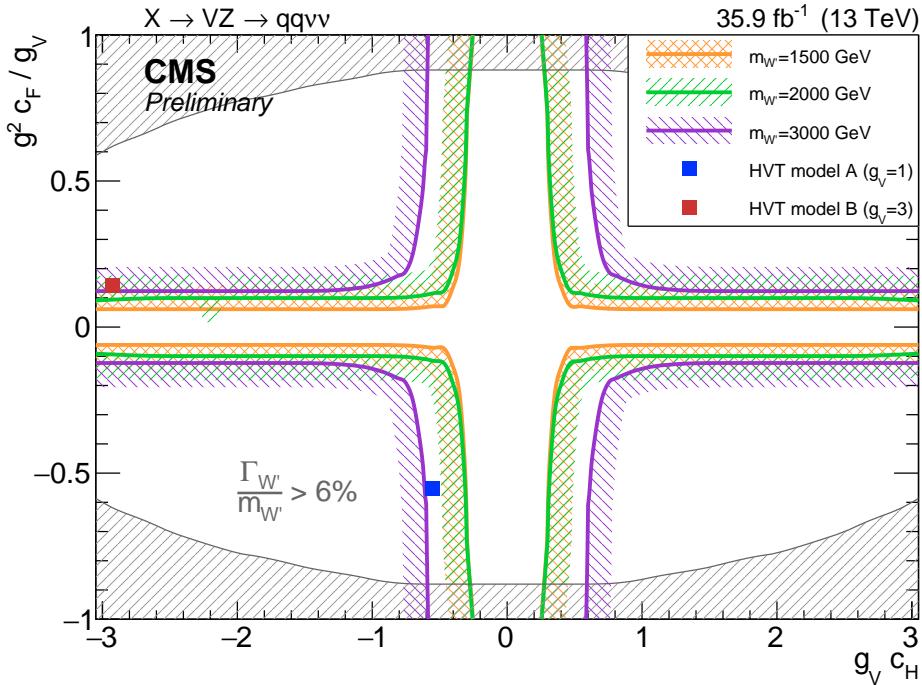


Figure 4.73: Exclusion in the parameter space of the HVT model. Coloured curves represents the contours of exclusion for a spin-1  $W'$  resonance of mass 1.5 TeV (in orange), 2 TeV (in green), 3 TeV (in violet). Benchmark model A and B are represented as blue and red markers. The shaded gray area indicates the parameter space where the narrow width approximation fails.

# 2108 Conclusions

2109

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

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

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 analysis is inserted in a set of searches for heavy resonances decaying into dibosons. The future perspectives of the analysis consist both on the combination of this final state with other diboson searches sharing the same treatment of the hadronic part of the  $V$  decay (namely, the same definition of sidebands and signal regions), and on the combination of the 2016 data with the newly collected 2017 data. The luminosity planned to be delivered by the LHC collider in 2017 is comparable to what was collected in 2016 ( $\sim 40 \text{ fb}^{-1}$ ). By doubling the statistics, marginal improvements are foreseen; hence, a larger improvement can be achieved by decreasing the impacts of systematic uncertainties. New jet substructure techniques are currently being tested, in order to improve the jet mass resolution (recursive soft drop), suppress the pile-up contribution (PUPPI associated to SoftKiller algorithm [100]), exploit the jet substructure and tag the nature of a large-cone jet (originating from  $W, Z$ , Higgs boson or top quark) with machine learning techniques. Preliminary results on these new methods seem to be promising.



2138

## Bibliography

2139

2140

- 2141 [1] ATLAS Collaboration, ATLAS Collaboration, *Observation of a new particle in the search for*  
2142 *the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012)  
2143 1, [1207.7214].
- 2144 [2] CMS Collaboration, CMS Collaboration, *Observation of a new boson at a mass of 125 GeV*  
2145 *with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30, [1207.7235].
- 2146 [3] CMS Collaboration, CMS Collaboration, *Observation of a new boson with mass near 125 GeV*  
2147 *in pp collisions at  $\sqrt{s} = 7$  and 8 TeV*, *JHEP* **06** (2013) 081, [1303.4571].
- 2148 [4] ATLAS Collaboration, ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson*  
2149 *using ATLAS data*, *Phys. Lett. B* **726** (2013) 120, [1307.1432].
- 2150 [5] CMS Collaboration, CMS Collaboration, *Precise determination of the mass of the Higgs boson*  
2151 *and tests of compatibility of its couplings with the standard model predictions using proton*  
2152 *collisions at 7 and 8 TeV*, *Eur. Phys. J. C* **75** (2015) 212, [1412.8662].
- 2153 [6] ATLAS Collaboration, ATLAS Collaboration, *Measurement of the Higgs boson mass from the*  
2154  *$H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^* \rightarrow 4\ell$  channels in pp collisions at center-of-mass energies of 7 and*  
2155 *8 TeV with the ATLAS detector*, *Phys. Rev. D* **90** (2014) 052004, [1406.3827].
- 2156 [7] CMS and ATLAS Collaborations, “Combined Measurement of the Higgs Boson Mass in pp  
2157 Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments.” 2015.
- 2158 [8] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, G. Isidori et al., *Higgs mass*  
2159 *and vacuum stability in the Standard Model at NNLO*, *JHEP* **08** (2012) 098, [1205.6497].
- 2160 [9] D. Pappadopulo, A. Thamm, R. Torre and A. Wulzer, *Heavy vector triplets: bridging theory*  
2161 *and data*, *Journal of High Energy Physics* **2014** (2014) 1–50.
- 2162 [10] V. D. Barger, W.-Y. Keung and E. Ma, *A Gauge Model With Light W and Z Bosons*, *Phys. Rev.*  
2163 **D22** (1980) 727.
- 2164 [11] C. Grojean, E. Salvioni and R. Torre, *A weakly constrained W' at the early LHC*, *JHEP* **07** (2011)  
2165 002, [1103.2761].

- 2166 [12] R. Contino, D. Pappadopulo, D. Marzocca and R. Rattazzi, *On the effect of resonances in*  
2167 *composite higgs phenomenology*, *Journal of High Energy Physics* **2011** (2011) 1–50.
- 2168 [13] B. Bellazzini, C. Csáki and J. Serra, *Composite Higgses*, *Eur. Phys. J.* **C74** (2014) 2766,  
2169 [1401.2457].
- 2170 [14] G. F. Giudice, C. Grojean, A. Pomarol and R. Rattazzi, *The Strongly-Interacting Light Higgs*,  
2171 *JHEP* **06** (2007) 045, [[hep-ph/0703164](#)].
- 2172 [15] A. Wulzer, *An Equivalent Gauge and the Equivalence Theorem*, *Nucl. Phys.* **B885** (2014)  
2173 97–126, [[1309.6055](#)].
- 2174 [16] CMS Collaboration, A. M. Sirunyan et al., *Search for massive resonances decaying into WW,*  
2175 *WZ, ZZ, qW, and qZ with dijet final states at  $\sqrt{s} = 13$  TeV*, [1708.05379](#).
- 2176 [17] CMS Collaboration, CMS Collaboration, *Search for massive resonances decaying into WW,*  
2177 *WZ, ZZ, qW and qZ in the dijet final state at  $\sqrt{s} = 13$  TeV*, CMS Physics Analysis Summary  
2178 CMS-PAS-B2G-17-001, CERN, Geneva, 2017. <http://cds.cern.ch/record/2256663>.
- 2179 [18] CMS Collaboration, CMS Collaboration, *Search for heavy resonances decaying into a vector*  
2180 *boson and a Higgs boson in hadronic final states with 2016 data*, .
- 2181 [19] CMS Collaboration, CMS Collaboration, *Search for heavy resonances decaying into a vector*  
2182 *boson and a Higgs boson in hadronic final states with 2016 data*, CMS Physics Analysis  
2183 Summary CMS-PAS-B2G-17-002, CERN, Geneva, 2017. <http://cds.cern.ch/record/2256742>.
- 2184 [20] CMS Collaboration, CMS Collaboration, *Search for heavy resonances decaying into a Z boson*  
2185 *and a W boson in the  $\ell^+\ell^-q\bar{q}$  final state*, CMS Physics Analysis Summary  
2186 CMS-PAS-B2G-16-022, CERN, Geneva, 2017. <http://cds.cern.ch/record/2242955>.
- 2187 [21] CMS Collaboration, V. Khachatryan et al., *Search for heavy resonances decaying into a vector*  
2188 *boson and a Higgs boson in final states with charged leptons, neutrinos, and b quarks*, *Phys.*  
2189 *Lett.* **B768** (2017) 137–162, [[1610.08066](#)].
- 2190 [22] CMS Collaboration, CMS Collaboration, *Search for new resonances decaying to*  
2191 *WW/WZ  $\rightarrow \ell\nu qq$* , CMS Physics Analysis Summary CMS-PAS-B2G-16-020, CERN, Geneva,  
2192 2016. <http://cds.cern.ch/record/2205880>.
- 2193 [23] CMS Collaboration, CMS Collaboration, *Search for heavy resonances decaying into a Z boson*  
2194 *and a vector boson in the  $\nu\nu q\bar{q}$  final state*, CMS Physics Analysis Summary  
2195 CMS-PAS-B2G-17-005, CERN, Geneva, 2017. <http://cds.cern.ch/record/2273910>.
- 2196 [24] ATLAS Collaboration, M. Aaboud et al., *Search for diboson resonances with boson-tagged jets*  
2197 *in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, [1708.04445](#).
- 2198 [25] ATLAS Collaboration, ATLAS Collaboration, *Search for WW/WZ resonance production in*  
2199  *$\ell\nu qq$  final states in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, ATLAS Conference  
2200 Note ATLAS-CONF-2017-051, CERN, Geneva, Jul, 2017. <http://cds.cern.ch/record/2273867>.
- 2201 [26] ATLAS Collaboration, ATLAS Collaboration, *Searches for heavy ZZ and ZW resonances in the*  
2202  *$llqq$  and  $\nu\nu qq$  final states in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, ATLAS  
2203 Conference Note ATLAS-CONF-2016-082, CERN, Geneva, Aug, 2016.  
2204 <https://cds.cern.ch/record/2206275>.

## BIBLIOGRAPHY

---

- [2205] [27] ATLAS Collaboration, M. Aaboud et al., *Search for heavy resonances decaying to a W or Z boson and a Higgs boson in the  $q\bar{q}^{(\prime)} b\bar{b}$  final state in  $p\bar{p}$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, 1707.06958.
- [2208] [28] ATLAS Collaboration, 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  $36.1 \text{ fb}^{-1}$  of  $p\bar{p}$  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>.
- [2212] [29] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, *Phys. Rev. Lett.* **83** (1999) 3370–3373, [[hep-ph/9905221](#)].
- [2214] [30] L. Randall and R. Sundrum, *An Alternative to compactification*, *Phys. Rev. Lett.* **83** (1999) 4690–4693, [[hep-th/9906064](#)].
- [2216] [31] K. Agashe, H. Davoudiasl, G. Perez and A. Soni, *Warped Gravitons at the LHC and Beyond*, *Phys. Rev.* **D76** (2007) 036006, [[hep-ph/0701186](#)].
- [2218] [32] 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](#)].
- [2220] [33] 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](#)].
- [2222] [34] T. Gherghetta and A. Pomarol, *Bulk fields and supersymmetry in a slice of AdS*, *Nucl. Phys.* **B586** (2000) 141–162, [[hep-ph/0003129](#)].
- [2224] [35] 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](#)].
- [2226] [36] A. Oliveira, *Gravity particles from Warped Extra Dimensions, predictions for LHC*, 1404.0102.
- [2228] [37] CMS Collaboration, CMS Collaboration, *Search for diboson resonances in the  $2l2\nu$  final state*, CMS Physics Analysis Summary CMS-PAS-B2G-16-023, CERN, Geneva, 2017. <http://cds.cern.ch/record/2264700>.
- [2231] [38] CMS Collaboration, CMS Collaboration, *Search for heavy resonances decaying to a pair of Higgs bosons in the four b quark final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, CMS Physics Analysis Summary CMS-PAS-B2G-16-026, CERN, Geneva, 2017. <http://cds.cern.ch/record/2264684>.
- [2235] [39] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.
- [2236] [40] CMS Collaboration, CMS Collaboration, *The CMS experiment at the CERN LHC*, *JINST* **3** (2008) S08004.
- [2238] [41] CMS Collaboration, CMS Collaboration, *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997, <https://cds.cern.ch/record/368412>.
- [2241] [42] CMS Collaboration, CMS Collaboration, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, *JINST* **9** (2014) P10009, [[1405.6569](#)].

- 2243 [43] CMS Collaboration, CMS Collaboration, *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997, <https://cds.cern.ch/record/349375>.
- 2244
- 2245
- 2246 [44] CMS Collaboration, CMS Collaboration, *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997, <https://cds.cern.ch/record/357153>.
- 2247
- 2248
- 2249 [45] CMS Collaboration, CMS Collaboration, *The CMS muon project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997, <https://cds.cern.ch/record/343814>.
- 2250
- 2251 [46] CMS Collaboration, CMS Collaboration, *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical Design Report CMS. CERN, Geneva, 2002, <https://cds.cern.ch/record/578006>.
- 2252
- 2253
- 2254 [47] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth.* **A389** (1997) 81–86.
- 2255
- 2256 [48] CMS Collaboration, A. M. Sirunyan et al., *Particle-flow reconstruction and global event description with the cms detector*, *JINST* **12** (2017) P10003, [[1706.04965](#)].
- 2257
- 2258 [49] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, *Nucl. Instrum. Meth.* **A262** (1987) 444–450.
- 2259
- 2260 [50] K. Rose, *Deterministic annealing for clustering, compression, classification, regression, and related optimization problems*, in *Proceedings of the IEEE*, vol. 86, pp. 2210 – 2239, 12, 1998.
- 2261
- 2262 [51] R. Frühwirth, W. Waltenberger and P. Vanlaer, *Adaptive vertex fitting*, *J. Phys.* **G34** (2007) N343.
- 2263
- 2264 [52] CMS Collaboration, CMS Collaboration, *Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at  $\sqrt{s} = 7$  TeV*, *JINST* **8** (2013) P09009, [[1306.2016](#)].
- 2265
- 2266
- 2267 [53] W. Adam, R. Frühwirth, 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](#)].
- 2268
- 2269
- 2270
- 2271 [54] CMS Collaboration, CMS Collaboration, *Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at  $s = 8$  TeV*, *JINST* **10** (2015) P06005, [[1502.02701](#)].
- 2272
- 2273
- 2274 [55] CMS Collaboration, CMS Collaboration, *Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at  $\text{sqrt}(s) = 8$  TeV*, *JINST* **10** (2015) P08010, [[1502.02702](#)].
- 2275
- 2276
- 2277 [56] CMS Collaboration, CMS Collaboration, *Performance of CMS muon reconstruction in pp collision events at  $\sqrt{s} = 7$  TeV*, *JINST* **7** (2012) P10002, [[1206.4071](#)].
- 2278
- 2279 [57] CMS Collaboration, 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>.
- 2280
- 2281

## BIBLIOGRAPHY

---

- [58] CMS Collaboration, CMS Collaboration, *Pileup Removal Algorithms*, Tech. Rep. CMS-PAS-JME-14-001, CERN, Geneva, 2014. <https://cds.cern.ch/record/1751454>.
- [59] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896, [1111.6097].
- [60] M. Cacciari, G. P. Salam and G. Soyez, *The anti- $k_t$  jet clustering algorithm*, *JHEP* **04** (2008) 063, [0802.1189].
- [61] CMS Collaboration, CMS Collaboration, *Determination of jet energy calibration and transverse momentum resolution in cms*, *Journal of Instrumentation* **6** (2011) P11002.
- [62] CMS Collaboration, 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>.
- [63] CMS Collaboration, CMS Collaboration, *Performance of tau-lepton reconstruction and identification in CMS*, *JINST* **7** (2012) P01001, [1109.6034].
- [64] CMS Collaboration, 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>.
- [65] CMS Collaboration, CMS Collaboration, *Identification of b-quark jets with the CMS experiment*, *JINST* **8** (2013) P04013, [1211.4462].
- [66] CMS Collaboration, CMS Collaboration, *Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector*, .
- [67] CMS Collaboration, CMS Collaboration, *Missing transverse energy performance of the CMS detector*, *JINST* **6** (2011) P09001, [1106.5048].
- [68] CMS Collaboration, CMS Collaboration, *Performance of Missing Transverse Momentum Reconstruction Algorithms in Proton-Proton Collisions at  $p s = 8$  TeV with the CMS Detector*, .
- [69] ATLAS Collaboration, ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [70] ALICE Collaboration, ALICE Collaboration, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.
- [71] LHCb Collaboration, LHCb Collaboration, *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005.
- [72] 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].
- [73] Torbjorn Sjostrand and Stephen Mrenna and Peter Skands, *A brief introduction to pythia 8.1*, *Comput.Phys.Commun.* **178** (2008) 852–867, [0710.3820].
- [74] CMS Collaboration, CMS Collaboration, *Event generator tunes obtained from underlying event and multiparton scattering measurements*, *Eur. Phys. J. C* (2016) 155, [1512.00815].

- 2319 [75] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce et al., *Geant4a simulation*  
2320 *toolkit, Nuclear Instruments and Methods in Physics Research Section A: Accelerators,*  
2321 *Spectrometers, Detectors and Associated Equipment* **506** (2003) 250 – 303.
- 2322 [76] CMS Collaboration, V. Khachatryan et al., *Identification techniques for highly boosted W*  
2323 *bosons that decay into hadrons, JHEP* **12** (2014) 017, [1410.4227].
- 2324 [77] Rikkert Frederix and Stefano Frixione, *Merging meets matching in mc@nlo, JHEP* **12** (2012)  
2325 061, [1209.6215].
- 2326 [78] Emanuele Re, *Single-top wt-channel production matched with parton showers using the*  
2327 *powheg method, Eur.Phys.J.* **C71** (2011) 1547, [1009.2450].
- 2328 [79] John M. Campbell, R. Keith Ellis, Paolo Nason, Emanuele Re, *Top-pair production and decay*  
2329 *at nlo matched with parton showers, JHEP* **04** (2015) 114, [1412.1828].
- 2330 [80] S. Kallweit, J. M. Lindert, S. Pozzorini, M. Schönher and P. Maierhöfer, *NLO QCD+EW*  
2331 *automation and precise predictions for V+multijet production, in Proceedings, 50th Rencontres*  
2332 *de Moriond, QCD and high energy interactions*, pp. 121–124, 2015, 1505.05704,  
2333 <http://inspirehep.net/record/1372103/files/arXiv:1505.05704.pdf>.
- 2334 [81] CMS COLLABORATION Collaboration, *CMS Luminosity Measurements for the 2016 Data*  
2335 *Taking Period, CMS Physics Analysis Summary CMS-PAS-LUM-17-001*, CERN, Geneva, 2017.  
2336 <https://cds.cern.ch/record/2257069>.
- 2337 [82] A. J. Larkoski, S. Marzani, G. Soyez and J. Thaler, *Soft Drop, JHEP* **05** (2014) 146, [1402.2657].
- 2338 [83] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better jet clustering algorithms,*  
2339 *JHEP* **08** (1997) 001, [hep-ph/9707323].
- 2340 [84] D. Bertolini, P. Harris, M. Low and N. Tran, *Pileup per particle identification, Journal of High*  
2341 *Energy Physics* **2014** (2014) 59.
- 2342 [85] J. Thaler and K. Van Tilburg, *Identifying boosted objects with n-subjettiness, Journal of High*  
2343 *Energy Physics* **2011** (2011) 15.
- 2344 [86] CMS Collaboration, *Performance of b-Tagging Algorithms in Proton Collisions at 13 TeV using*  
2345 *the 2016 Data, Physics Analysis Summary CMS-DP-2016-042*, CERN, Geneva, Jul, 2016.  
2346 <https://cds.cern.ch/record/2202967>.
- 2347 [87] CMS Collaboration, V. Khachatryan et al., *Search for a higgs boson in the mass range from 145*  
2348 *to 1000 GeV decaying to a pair of W or Z bosons, JHEP* **10** (2015) 144, [1504.00936].
- 2349 [88] M. Oreglia, *A Study of the Reactions  $\psi' \rightarrow \gamma\gamma\psi$* , Ph.D. thesis, SLAC, 1980.
- 2350 [89] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and*  
2351 *Upsilon resonances*, Ph.D. thesis, Cracow, INP, 1986.
- 2352 [90] M. Bahr, S. Gieseke, M. A. Gigg, D. Grellscheid, K. Hamilton, O. Latunde-Dada, S. Platzer, P.  
2353 Richardson, M. H. Seymour, A. Sherstnev, J. Tully, B. R. Webber, *Herwig++ physics and*  
2354 *manual, Eur.Phys.J.* **C58** (2008) 639–707, [0803.0883].
- 2355 [91] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II, J. Phys.* **G43** (2016) 023001,  
2356 [1510.03865].

## BIBLIOGRAPHY

---

---

- 2357 [92] NNPDF Collaboration, R. D. Ball et al., *Parton distributions from high-precision collider data*,  
2358 *Eur. Phys. J.* **C77** (2017) 663, [1706.00428].
- 2359 [93] A. L. Read, *Presentation of search results: the  $CL_s$  technique*, *J. Phys.* **G28** (2002) 2693.
- 2360 [94] T. Junk, *Confidence level computation for combining searches with small statistics*, *Nucl.*  
2361 *Instrum. Meth.* **A434** (1999) 435, [9902006].
- 2362 [95] The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group,  
2363 *Procedure for the LHC Higgs boson search combination in Summer 2011*, Tech. Rep.  
2364 CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug, 2011.  
2365 <https://cds.cern.ch/record/1379837>.
- 2366 [96] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests*  
2367 *of new physics*, *The European Physical Journal C* **71** (2011) .
- 2368 [97] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite*  
2369 *hypotheses*, *The Annals of Mathematical Statistics* **9** (03, 1938) 60–62.
- 2370 [98] A. Wald, *Tests of statistical hypotheses concerning several parameters when the number of*  
2371 *observations is large*, *Transactions of the American Mathematical Society* **54** (1943) 426–482.
- 2372 [99] E. Gross and O. Vitells, *Trial factors for the look elsewhere effect in high energy physics*, *The*  
2373 *European Physical Journal C* **70** (Nov, 2010) 525–530.
- 2374 [100] M. Cacciari, G. P. Salam and G. Soyez, *SoftKiller, a particle-level pileup removal method*, *Eur.*  
2375 *Phys. J.* **C75** (2015) 59, [1407.0408].