Chapter 1

Introduction

1.0.1 Standard Model in a nutshell

Standard Model (SM) is a summarization of what is known in the subatomic world and also it predicts the existence of new particles.

Chapter 2

The CMS detector

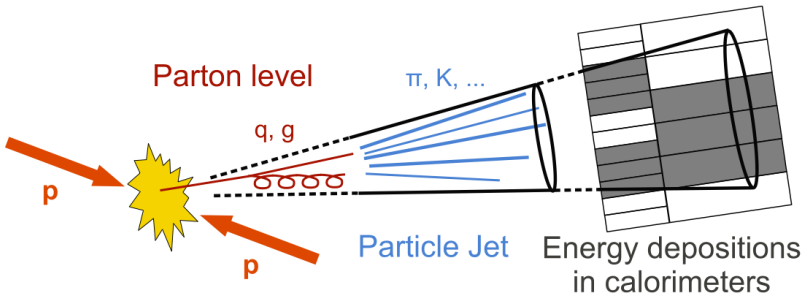

Figure 3.1: Sketch of pp-collision and resulting collimated spray of particles, a jet.

Chapter 3

Event reconstruction and selections

Jets are the experimental signatures of quarks and gluons produced in high-energy processes such as head-on proton-proton collisions at LHC. As quarks and gluons have a net colour charge and cannot exist freely due to colour-confinement, they are not directly observed in Nature. Instead, they come together to form colour-neutral hadrons, a process called hadronisation that leads to a collimated spray of hadrons called a jet.

As these jets of particles propagate through the CMS detector, they leave signals in components such as the tracker and the electromagnetic and hadronic calorimeters. These signals are combined using jet algorithms to form a reconstructed jet.

3.1 Event reconstruction

Here events are reconstructed using the particle-flow reconstruction algorithm [**?**], which attempts to reconstruct all stable particles in an event by combining information from all subdetectors. The algorithm categorizes all particles into five types: muons, electrons, photons, charged and neutral hadrons. The resulting particle flow candidates are passed to each jet clustering algorithm, in this case the Cambridge-Aachen (CA) [**?**, **?**] jet clustering algorithm, as implemented in FastJet version 3.0.1 [**?**, **?**], to create "particle flow jets". The CA clustering sequence is only determined by the distance between clusters and is not weighted by their momentum, as is done for the k_{T} and anti-k_{T} algorithms. A distance parameter of size $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.8$ is used for the CA algorithm.

Charged hadrons identified as pileup are removed from the inputs to the jet clustering algorithms. The remaining neutral component of pileup is removed by applying a residual area-based correction as described in Ref. [**?**, **?**]. The mean $\mathrm{p_T}$ per unit area is computed with the k_{T} algorithm with the "active area" method, with a distance parameter of 0.6, and the jet energy is corrected by the amount of pileup expected

in the jet area. The amount of energy expected from the underlying event is added back into the jet. The pileup-subtracted jet four momenta are finally corrected for nonlinearities in η and p_T with simulated data, with a residual η-dependent correction added to correct for the difference in simulated and true responses [**?**, **?**].

3.2 W, Z and H tagging

3.2.1 jet pruning

As the mass of the V or H boson is larger than the mass of a typical QCD jet, the jet mass is the primary observable that distinguishes such a jet from a QCD jet. The bulk of the V or H jet mass arises from the kinematics of the two or more jet cores that correspond to the decay quarks. In contrast, the QCD jet mass arises mostly from soft gluon radiation. For this reason, the use of jet pruning [**?**, **?**] improves discrimination by removing the softer radiation, as this shifts the jet mass of QCD jets to smaller values, while maintaining the jet mass for V and H jets close to the masses of Z, W or H bosons. Jet pruning is implemented as additional cuts in the process of CA jet clustering. This algorithm starts from a set of "protojets" given by the PF particles that form the original CA jet within a cone of $R = 0.8$. These protojets are iteratively combined with each other until a set of jets is found, however the large angle and low p_T protojets are removed in the process. The details of this procedure are given in [**?**]. The distributions of the pruned jet mass (m_j) for simulated

signal and background samples, are shown in Fig. **??**. Jets from boosted W and Z decays are expected to generate peaks at $m_j \approx 80$ and $m_j \approx 90$ GeV, respectively. Jets from boosted H decays are expected to peak at $m_j \approx 120$ GeV. Hadronic top-quark jets, where the b quark and the two different light quarks from the $t \to Wb \to q\bar{q}'b$ decay are required to be within a reconstructed jet of size R = 0.8, peak at $m_j \approx 175$ GeV. Jets from multijet events and not-fully-merged W, Z and H bosons give rise to a peak around 20GeV, whose size depends particularly on the spin and polarization of the boson. All peaks are slightly shifted to lower masses, due to removal of soft radiation by jet pruning. If the pruned jet has a mass (m_j) within $70 < m_j < 100 GeV$ ($110 < m_j < 135$ GeV), candidate.

3.2.2 b tagging

Jet pruning can also provide a good delineation of subjets within the CA8 jet. To tag jets from $H \to b\bar{b}$ decays, denoted as H_{bb} jets, the pruned subjets, given by reversing the last step of the CA8 pruning recombination algorithm, are used as the basis for b tagging. Jets arising from the hadronization of b quarks (b jets) are identified using the combined secondary vertex (CSV) b tagging algorithm [**?**], which uses information from tracks and secondary vertices associated with jets to build a likelihood-based discriminator to distinguish between jets from b quarks and those from charm or light quarks and gluons. The b tagging discriminator can take values between 0 and 1 with higher values indicating higher probability of the jet to originate

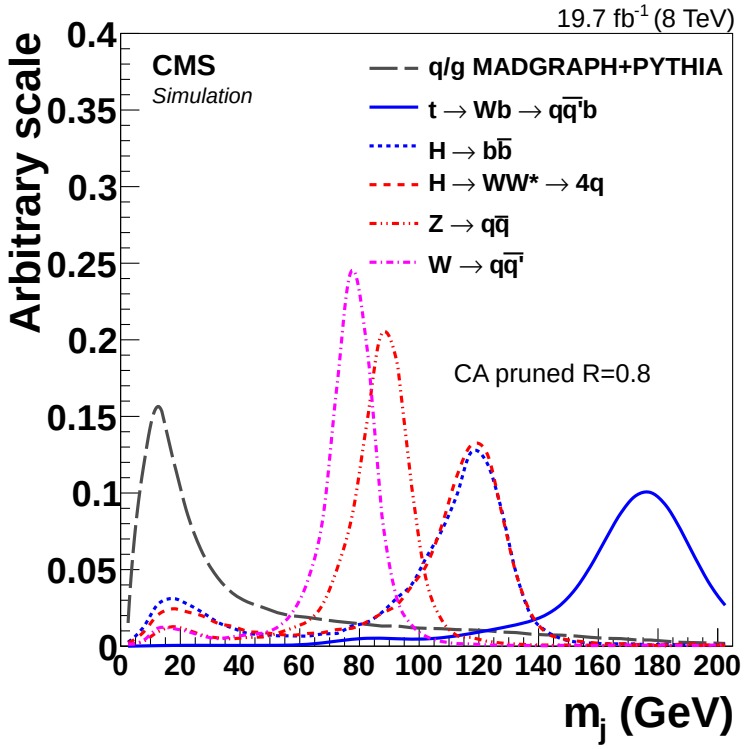

Figure 3.2: Distribution of pruned jet mass in simulation of signal and background processes. All simulated distributions are normalized to 1. The W/Z, H, and top-quark jets are required to match respective generator level particles in the event. The W/Z and H jets are from 1.5 TeV $W' \to WH$ and $Z' \to ZH$ signal samples.

from a b quark. The loose working point of the CSV algorithm [**?**] is chosen, which is found to be optimal for both subjets and jets b tagging. It gives a b-tagging efficiency of $\approx 85\%$, with mistagging probabilities of $\approx 40\%$ for c-quark jets and $\approx 10\%$ for light-quark and gluon jets with $p_T \approx 80 GeV$. The b-tagging efficiency ratio between data and the simulated signal events. To identify CA8 jets originating from H decays resulting in two collimated b jets, we apply b tagging either on the two subjets or the CA8 jet, based on the angular separation of the two subjets (Δ R) [**?**]. If ΔR

between the CA8 subjets is bigger (smaller) than 0.3, b tagging is applied on both of the two subjets (on the CA8 jet).

3.2.3 N subjettiness

We achieve additional discrimination against multijet events by considering the distribution of jet constituents relative to the jet axis. In particular, we quantify how well the constituents of a given jet can be arranged into N subjets. This is done by reconstructing the full set of jet constituents (before pruning) with the $k_T$algorithm [**?**] and halting the reclustering when N distinguishable protojets are formed. The directions of the N jets are used as the reference axes to compute the N-subjettiness [**?**,**?**,**?**] τ_N of the original jet, defined as

$$\tau_N = \frac{1}{d_0} \sum_k p_{\mathrm{T},k} \, \min(\Delta R_{1,k}, \Delta R_{2,k}, \ldots, \Delta R_{N,k}), \tag{3.1}$$

where $p_{\mathrm{T},k}$ is the $\mathrm{p_T}$ of the particle constituent k of the original jet, and $\Delta R_{n,k}$ is its angular distance from the axis of the nth subjet (with $n = 1, 2, \ldots, N$). The normalization factor d_0 for τ_N is $d_0 = \sum_k p_{\mathrm{T},k} R_0$, with R_0 set to the distance parameter R of the original CA jet. To improve the discriminating power, we perform a one-pass optimization of the directions of the subjet axes by minimizing τ_N [**?**, **?**]. By using the smallest $\Delta R_{n,k}$ to weight the value of $p_{\mathrm{T},k}$ in Eq. (**??**), τ_N yields small values when the jet originates from the hadronization of N quarks. The $\tau_{ij} = \tau_i/\tau_j$ ratios

τ_{21}, τ_{31}, τ_{32}, τ_{41}, τ_{42}, and τ_{43} have been studied to identify the best discriminators for jets from $H \to WW^* \to 4q$ and $W/Z \to qq'$ decays. We find that the ratio τ_{42} works best to discriminate the four-pronged $H \to WW^* \to 4q$ events against QCD jets, and τ_{21} to identify $W/Z \to qq'$ [**?**]. The discriminating power of τ_{21} and τ_{42} for different resonance models can be seen in Fig. **??** and Fig. **??**, respectively.

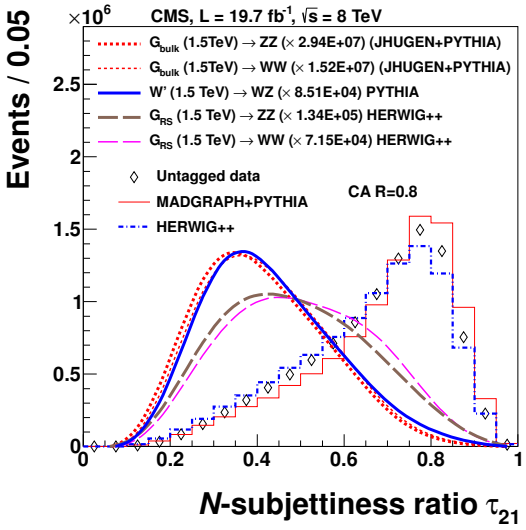

Figure 3.3: Distribution for jet N-subjettiness ratio τ_{21} in data, and in simulations of signal and background events. All simulated distributions are scaled to match the number of events in data. MADGRAPH/PYTHIAand HERWIG++ refer to QCD multijet event simulations.

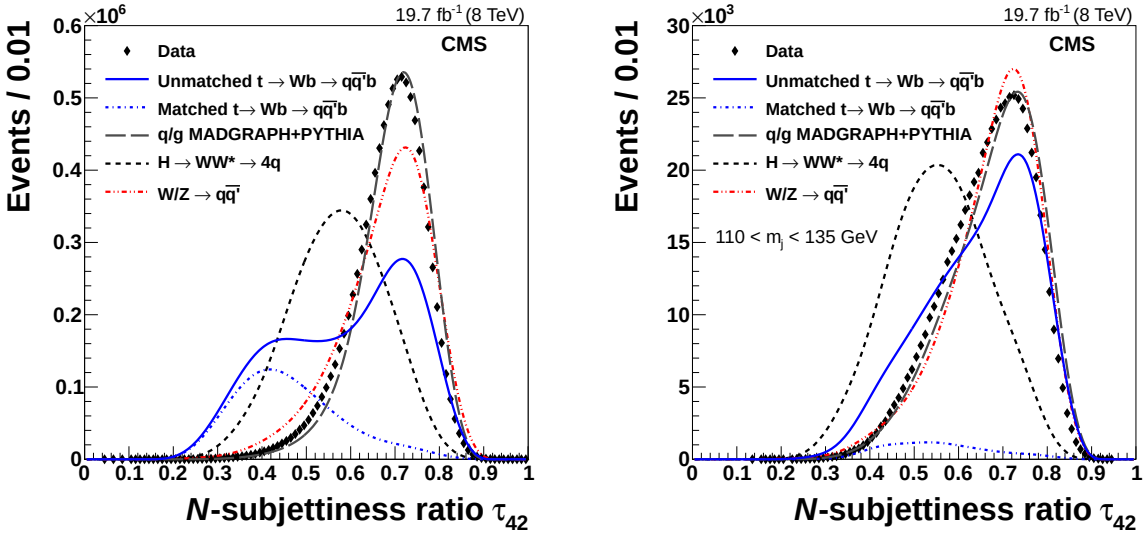

Figure 3.4: Distributions of τ_{42} in data and in simulations of signal (2 TeV) and background events, without applying the pruned jet mass requirement (left) and with the pruned jet mass requirement applied (right). W/Z, matched top-quark, and H_{WW} jets are required to match their generator level particles, respectively. All simulated distributions are scaled to match the number of events in data, except that matched top-quark background is scaled to the fraction of unmatched $t\bar{t}$ events times the number of data events.

Chapter 4

EXO-12-024

4.1 Introduction

Several scenarios of physics beyond the standard model predict the existence of resonances with masses above 1 TeVwhich decay into a quark jet and a W or Z boson, or into a pair of W or Z bosons. A vector boson emerging from such a decay is usually sufficiently boosted that the hadronization products from its daughter quarks merge into a single massive jet [**?**]. We present a search for these special dijet topologies, performed in the pp collision data collected by the CMS experiment in 2012 at a center-of-mass energy of 8 TeV.

The signal is characterized by a peak in the dijet invariant mass, emerging from the falling standard model background comprised mainly of QCD events with a dijet topology. The analysis presented here enhances the sensitivity to processes with

jets from W/Z bosons by the application of techniques that can identify W/Z-jets and suppress quark and gluon jets ("W/Z-tagging"). This search is an update of a previous study [?] performed on the 7 TeV CMS data. Besides a larger dataset and an increased signal cross section due to a higher center-of-mass energy, this update also benefits from an improved W/Z-tagger based on "N-subjettiness" variables [?].

We consider four benchmark scenarios that would produce singly- or doubly-tagged events: an excited quark q^* [?] decaying into a quark and a W or Z boson; a Randall–Sundrum (RS) graviton G_{RS}decaying to WW or ZZ [?]; a Bulk graviton G_{Bulk}decaying to WW or ZZ [?, ?, ?]; and a heavy partner of the SM W boson W' which decays to WZ [?]. The most stringent limits on the q^* model have been set in dijet resonance searches at the LHC by considering inclusive all-hadronic final states [?, ?, ?]. The most stringent lower limit (at 95% CL) on the q^* mass to date is 3.47TeV [?]. Specific searches for the qW and qZ final states have previously been reported at the Tevatron [?, ?], which exclude resonances decaying to qW or qZ with masses up to 0.54 TeV, and at the LHC [?, ?], which extends the mass exclusion of $qW(qZ)$ to 2.38TeV(2.15TeV). The WW and ZZ final states have also been explored experimentally [?, ?, ?, ?], setting lower limits on the $G_{\mathrm{RS}}$and G_{Bulk}mass as a function of the coupling parameter $k/\overline{M}_{\mathrm{Pl}}$, where k determines the curvature of the warped space and $\overline{M}_{\mathrm{Pl}}$ is the reduced Planck mass ($\overline{M}_{\mathrm{Pl}} \equiv M_{\mathrm{Pl}}/\sqrt{8\pi}$). The $G_{\mathrm{Bulk}}$and G_{RS}models differ in the fact that the G_{Bulk}favors the decay into vector bosons rather than photons or fermions and favors the production of longitudinal polarized W or Z

bosons. For the W', the most stringent limits are reported in searches with leptonic final states [?, ?], and the current lower limit on the W' mass is 2.9TeV. The limit varies by 0.1TeV, depending on the chirality of the W' couplings. Specific searches in the WZ final state have also been reported [?, ?, ?] setting a lower limit of 1.1TeV.

This analysis is the reload of the previous analysis EXO-11-095 [?]. This analysis proceeds via the following steps:

1. The search is performed in the dijet sample, using the same preselection as the standard search for resonances decaying to dijets [?].

2. We identify events with substructure: in each jet which is a candidate to originate from merging of W or Z daughter jets

   - we require a pruned jet mass cut, and
   - an n-subjettiness cut

3. After the full event selection, a potential signal would be characterized as a peak in the dijet invariant mass, on top of a falling background distribution.

4. We model the background distribution with a smoothly falling analytical function.

5. Finally, in absense of a signal, we set the limits on the production cross section of models with qW, qZ, WW, WZ and ZZ final states using the CL_s approach.

4.2 The signal: dijet resonance

We search for dijet resonances corresponding to several models. Using the W/Z-tagging algorithm, we examine both single W/Z-tag and double W/Z-tag events.

The pruned jet mass and jet τ_{21} distributions in signal MC, data and background MC are shown in Fig. **??**. Fully merged jets from Ws and Zs peak around 80-90 GeVin pruned jet mass while QCD jets and not fully merged Ws and Zs peak around 20 GeV. The discriminating power of the pruned jet mass and τ_{21} for the different signals is evident.

For both the pruned jet mass and τ_{21}, differences are observed between the HERWIG++ (G_{RS}) and PYTHIA6 (G_{Bulk}, q^*, W') distributions, which arise from differences in the polarization of the W/Z boson and the showering and hadronization models used by these generators. In particular this is the reason why the WZ prediction for τ_{21} is different from the WW, ZZ predictions. The differences due to showering and hadronization models are into account in the estimate of the systematic uncertainties on the tagging efficiency, as discussed below.

The full event selection efficiency is estimated using simulated signal samples. Less than 1% of the ZZ or WW events which pass the full selection are from $ZZ \to llqq$ or $WW \to l\nu qq$ decays, where l can be a muon or electron. While 3% of the selected ZZ events are from $ZZ \to \tau\tau qq$ decays, less than 1% of the selected WW events are from $WW \to \tau\nu qq$ decays. To within 10% accuracy the full selection efficiency can therefore be approximated by the product of the W/Z-tagging efficiency and an

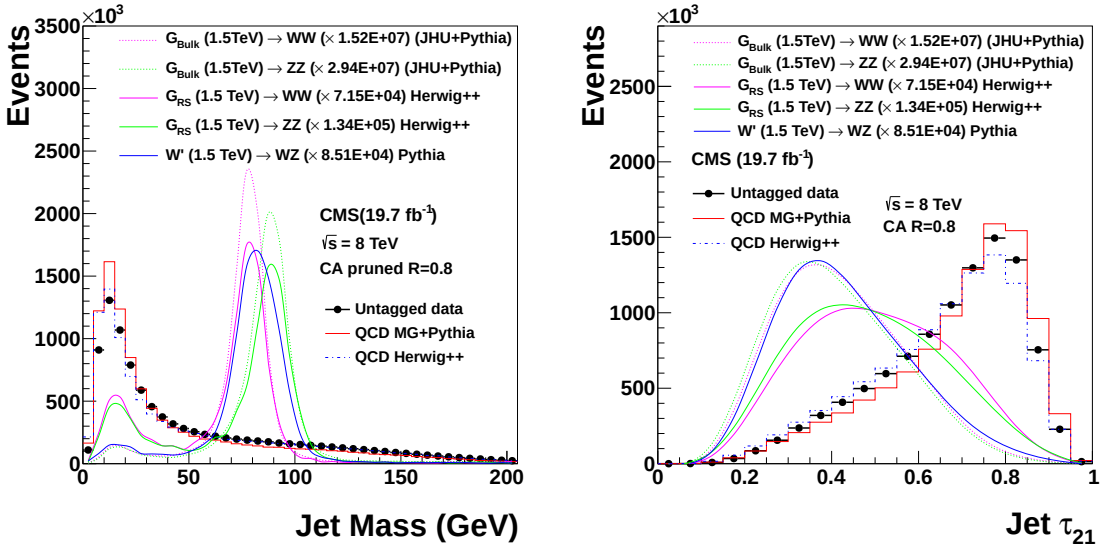

Figure 4.1: Pruned jet mass and τ_{21} in signal MC, data and background MC. All curves are plotted with the same binning. The signal MC distributions are plotted as smooth curves connecting the histogram entries. MC are normalized according to data.

approximate acceptance. This acceptance is shown in Fig. **??** and takes into account the angular acceptance ($|\eta| < 2.5$, $|\Delta\eta| < 1.3$), the branching into quark final states, BR($W/Z \to$ quarks) and a matching within $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.5$ between the generated W/Z bosons and the reconstructed jets.

The W/Z-tagging efficiency is shown in Fig. **??** and Fig. **??**.

The signal shapes for all five processes considered in this analysis are shown in Fig. **??** and Fig. **??**. For the qW and qZ final states the shape with a single W/Z-tag required is shown, while for the other signals two W/Z-tags are required.

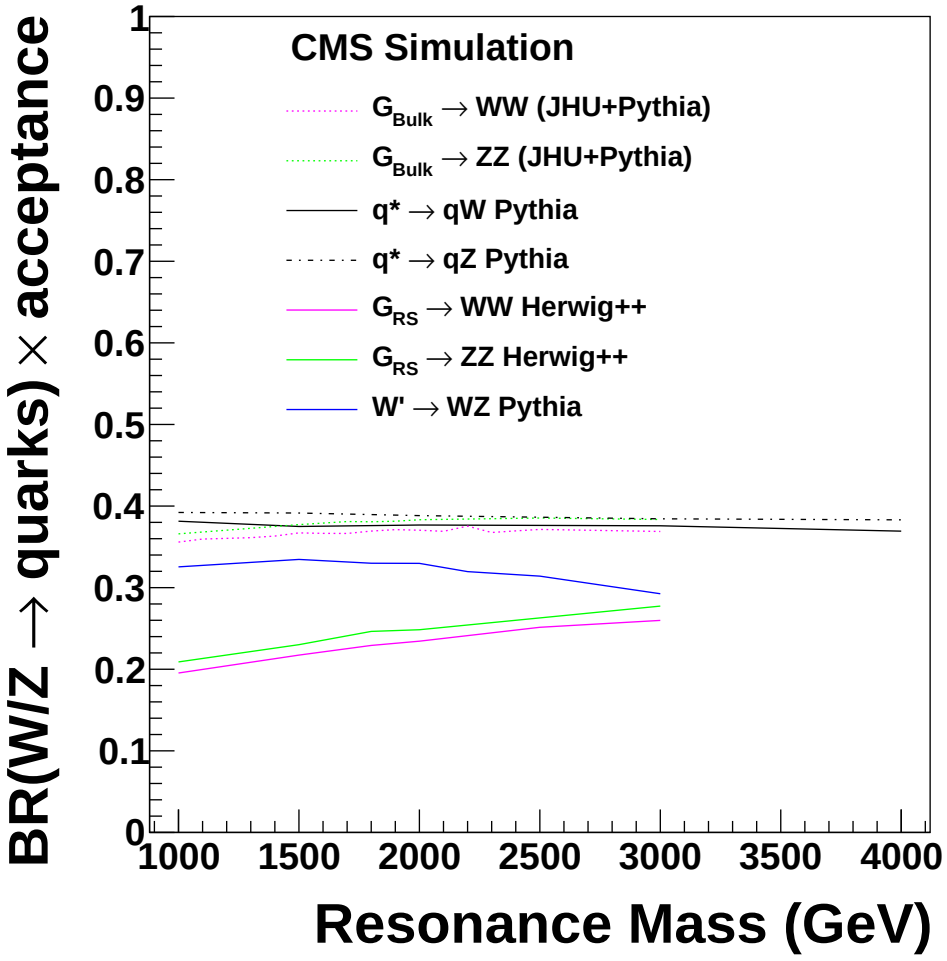

Figure 4.2: Fraction of events with branching into quark final states, BR($W/Z \rightarrow$ quarks), which are reconstructed as dijets (quarks $\rightarrow$ jets) and pass the angular acceptance ($|\eta| < 2.5$, $|\Delta\eta| < 1.3$.

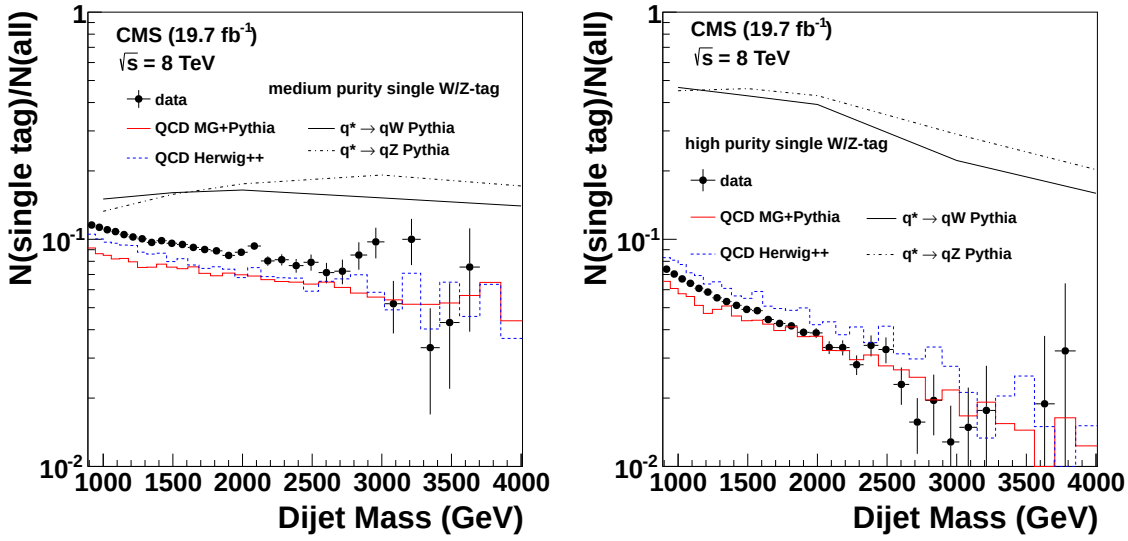

Figure 4.3: The fraction of singly-tagged events, requiring one medium purity (left) and high purity (right) W/Z-tag in data, signal and background simulations for events passing the angular acceptance requirement ($|\eta| < 2.5$, $|\Delta\eta| < 1.3$).

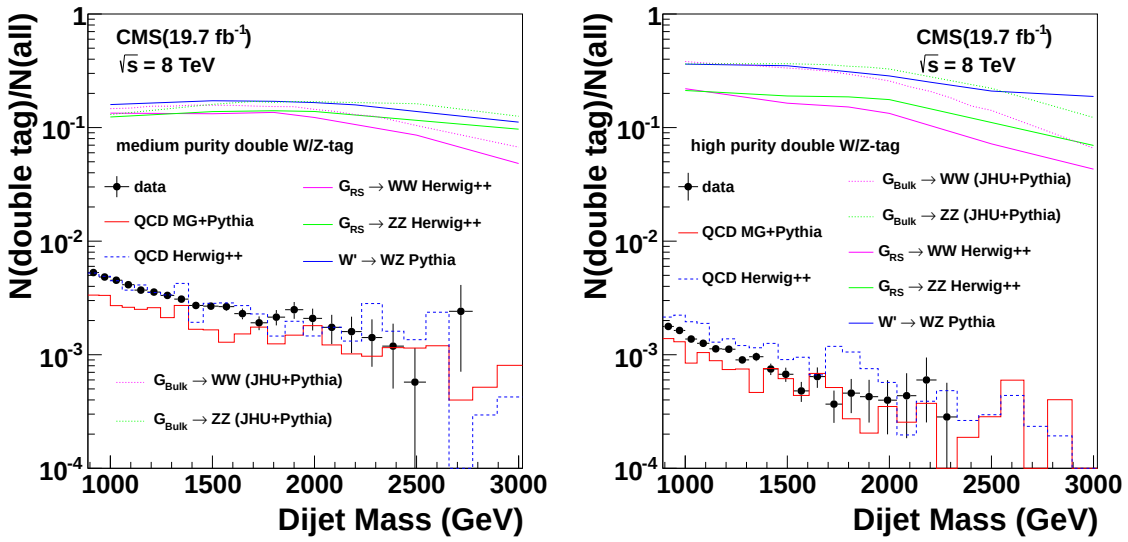

Figure 4.4: The fraction of doubly-tagged events, requiring two medium purity (left) and high purity (right) W/Z-tags in data, signal and background simulations for events passing the angular acceptance requirement ($|\eta| < 2.5$, $|\Delta\eta| < 1.3$).

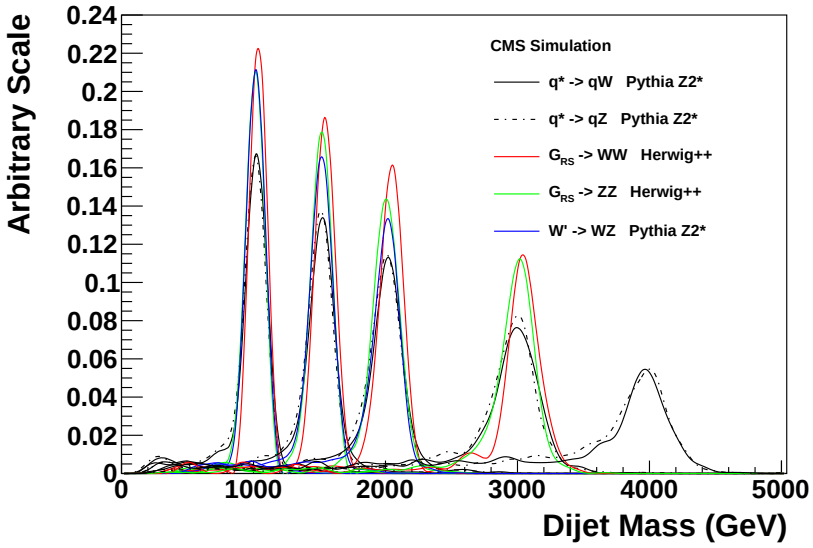

Figure 4.5: The normalized medium purity signal resonance distribution for $G_{RS} \to WW$, $G_{RS} \to ZZ$, $W' \to WZ$, $q* \to qW$, and $q* \to qZ$ resonances of dijet invariant mass 1.0TeV, 1.5TeV, 2.0TeV, 2.5 TeV, 3.0TeV, 4.0TeV.

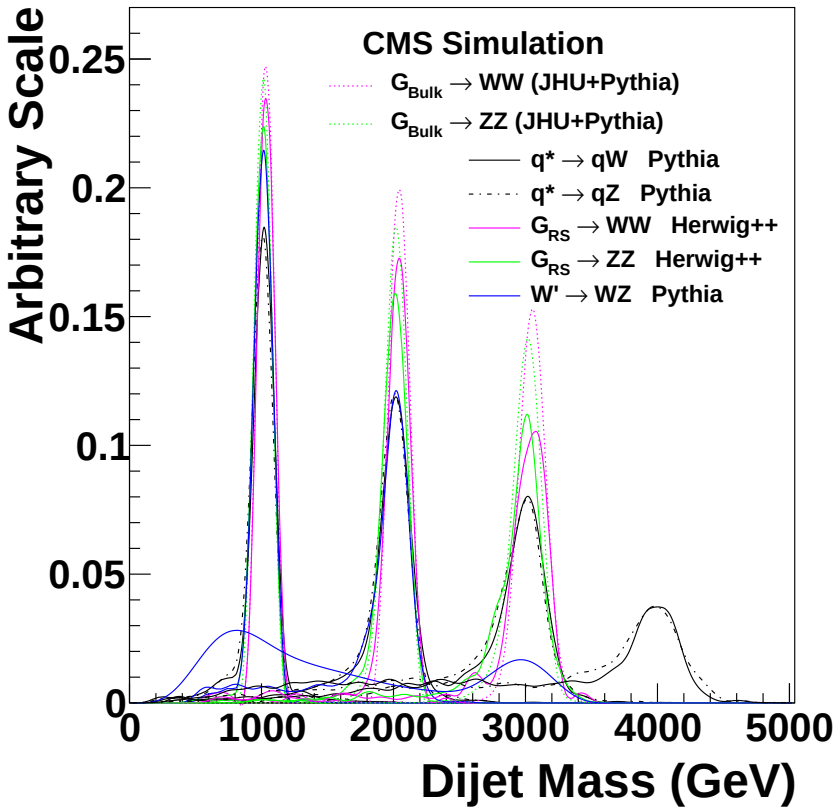

Figure 4.6: The normalized high purity signal resonance distribution for $G_{RS} \to WW$, $G_{RS} \to ZZ$, $W' \to WZ$, $q* \to qW$, and $q* \to qZ$ resonances of dijet invariant mass 1.0TeV, 1.5TeV, 2.0TeV, 3.0TeV, 4.0TeV.