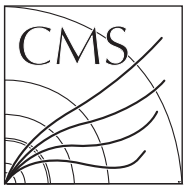
Revision : 261942

Available on the CMS information server **CMS AN-14-027**



The Compact Muon Solenoid Experiment

CMS Draft Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

2014/10/31
Head Id:
Archive Id: 261309:266008
Archive Date: 2014/09/25
Archive Tag: trunk

Search for massive WH and ZH resonances in dijet systems containing jets tagged as W or Z or H boson decays in pp collisions at $\sqrt{s} = 8\,\text{TeV}$

Yongjie Xin[3], Andreas Hinzmann[2], Petar Maksimović[3], Maxime Gouzevich[4], Alexandra Carvalho[4], Maurizio Pierini[1], and Thiago Fernandez[1]

[1] CERN
[2] University of Zurich
[3] Johns Hopkins University
[4] Universite Claude Bernard-Lyon I

Abstract

A search for massive resonances decaying into a Higgs boson (125 GeV) and a W or a Z boson is reported. The analysis is performed on an inclusive sample of multijet events corresponding to an integrated luminosity of $19.7\,\text{fb}^{-1}$, collected in proton-proton collisions at a centre-of-mass energy of 8 TeV with the CMS detector at the LHC. Jets from Higgs decaying into b-quark pairs are identified by subjets/jets b-tagging. The search also uses jet-substructure identification techniques that provide sensitivity to the presence of highly boosted Higgs bosons decaying into hadronic W pairs and vector bosons decaying into a pair of quarks. Exclusion limits are set at a confidence level of 95% on the production of Z' decaying into HZ for masses in [1.0, 1.1] TeV and [1.3, 1.5] TeV, and W' decaying into HW for masses in [1.0, 1.6] TeV.

Contents

1 Introduction

Several physics models beyond the standard model (SM) predicts the existence of vectorial resonances with masses above 1 TeV that decay into a W or Z vector boson, denoted as V boson, and a Higgs boson [1, 2] . In proton-proton (pp) collisions at the energies reached at the Large Hadron Collider (LHC), bosons emerging from such decays usually would have sufficiently large momenta so that the hadronization products of their decays would merge into a single massive jet [3]. We present a search for events containing jets of this kind in pp collisions at a centre-of-mass energy of $\sqrt{s} = 8$ TeV. The data sample, corresponding to an integrated luminosity of 19.7 fb^{-1}, was collected with the CMS detector at the LHC.

The signal of interest is a heavy vector resonance X, produced via the process , $\text{pp} \to \text{X} \to \text{HV}$, as in Figure 1, where the V boson decay hadronically and the Higgs boson decay either to a b-quark pair or to a pair of fully hadronic W bosons[1].

The resuts are interpreted in the following models of $\text{W}' \to \text{HW}$ and $\text{Z}' \to \text{HZ}$ resonances. The Composite-Higgs [4–6] and Little Higgs models [7] provide a direct solution to the hierarchy problem and predict many new particles, including addtional gauge bosons, e.g. heavy W' or Z' bosons. Models of such type are generalized in the Heavy Vector Triplet (HVT) model [8]. Of particular interest for this search is the so called HVT scenario B model. The W' and Z' decay to respective WH and ZH become dominant and almost degenarate with $\text{W}' \to \text{WZ}$, such that these parameter points are not constrained so far from experiments (see Fig 3.2 in Ref. [8]). For W' with the SM coupling, the most stringent limits are reported in searches with leptonic final states [9, 10], and the current lower limit on the W' mass is 2.9 TeV. The limit varies by 0.1 TeV, depending on the chirality of the W' couplings. Specific searches in the WZ final state have also been reported [11–13] setting a lower limit of 1.1 TeV. For Z' with the SM coupling, the most stringent limits are reported in searches with leptonic final states [14, 15], and the current lower limit on the Z' mass is 2.8 TeV.

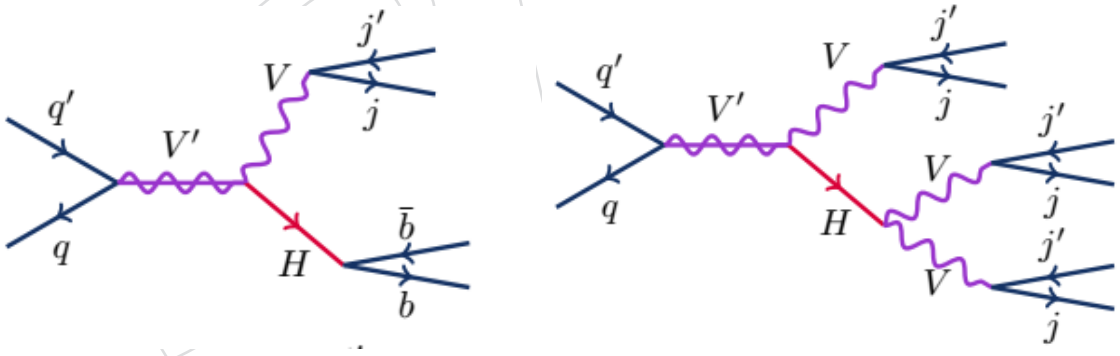


Figure 1: X decays.

The signal is characterized by a peak in the dijet invariant mass distribution m_{jj} over a continuous background from SM processes, comprised mainly of multijet events from quantum chromodynamic (QCD) processes. The sensitivity to jets from Higgs decays to b-quark pair is enhanced through subjets b-tagging [16]. While jets from Higgs decays to hadronic W-pairs, and also jets from W or Z bosons are identified with jet-substructure techniques [17, 18]. Results for W′ and Z′, which decays to Higgs boson and one vector boson V, are presented in this analysis.

[1] As the Higgs boson branching ratio to Z bosons is about one-tenth of its branching ratio to W bosons, we only consider the $H \to WW$ decay.

This is the first search for the heavy resonances decaying to all hadronic HV final states, as well as the first application of the highly boosted $H \to WW^* \to 4q$ decays.

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 [19, 20].
2. We identify events with W or Z: in each jet which is a candidate to originate from merging of V daughter jets:
   - we require a pruned jet mass cut, and
   - an N-subjettiness cut preferring two-prong decays

   (This is identical to EXO-12-024 [21].)
3. We identify events with a highly boosted Higgs boson:
   - we require a pruned jet mass cut, and
   - two b tagged subjets, or
   - (when there are no two b tagged subjets) a N-subjettiness cut preferring four subjets

   (The $H \to b\bar{b}$ tagging is synchronized with our sister analysis, the Radion search to the HH final state [22].)
4. 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.
5. We model the background distribution with a smoothly falling analytical function. (The functional form is identical to the one used in EXO-12-024.)
6. We form the joint likelihood of several dijet distributions of V tagged and H tagged jets. We include both two types of Higgs tags, and also low-purity Higgs and V taggers. The background estimate procedure is the same in all channels – analytical prameterization – but is performed separately for each channel.
7. Finally, we set the limits on the various simplified models for resonances decaying to HV final states.

2 Data and Monte Carlo samples

The data sample of proton-proton collisions at $\sqrt{s} = 8$ TeV was collected in 2012 and corresponds to an integrated luminosity of 19.7 fb^{-1}. The datasets and also the certifications used are summarized in Table 1. The dijet sample is dominated by light flavored and gluon jets, which we denote as the "QCD background".

We list part of our monte carlo simulated signal(from 1 TeV to 2.6 TeV) in Table 2. Signal samples are generated exclusively of the specific Higgs decay mode and W/Z decay mode. Model parameters and detailed cross sections are summarized in Appendix B. The matrix element is calculated with Madgraph 5.1.5.11 [23]. The signals of interest, are showered and hadronized with PYTHIA 6.426 [24], and HERWIG++ 2.5.0 [25], using simulation of the CMS detector, based on GEANT4 [26]. Tune Z2* [27] is used in PYTHIA, while the version 23 tune [25] is used in HERWIG++. The CTEQ61L [28] parton distribution functions (PDF) are used for PYTHIA and the MRST2001 [29] leading-order (LO) PDF for HERWIG++ W′ and Z′ are generated with resonance widths at ≈4% of the resonance mass, slightly smaller than the experimental resolution in m_{jj} for resonance masses considered in the analysis. Samples showered from PYTHIA are used in the analysis. While, samples from ++ are used to evaluate the systematic uncertainty by comparing the difference of hadronization from PYTHIA.

Dataset
/Jet/Run2012A-22Jan2013-v1/AOD
/JetHT/Run2012B-22Jan2013-v1/AOD
/JetHT/Run2012C-22Jan2013-v1/AOD
/JetHT/Run2012D-22Jan2013-v1/AOD

Table 1: Summary of 8 TeV collision data used in this analysis. The certification file used for these data is `Cert_190456-208686_8TeV_22Jan2013ReReco_Collisions12_JSON.txt` .

Process	mass (GeV)	Events	X-sec[pb]
Z′ → HZ	1000	20000	8.56E-02
	1500	20000	1.19E-02
	2000	20000	1.93E-03
	2500	20000	3.39E-04
W′ → HW	1000	20000	1.71E-01
	1500	20000	2.55E-02
	2000	20000	4.25E-03
	2500	20000	7.31E-04

Table 2: Examples of the simulated Monte Carlo samples used in this analysis for process V′ → VH. Cross sections are calculated from production cross sections of V′ times its BR(W′ →HW or Z′ → HZ). These samples are generated using Madgrap5 and hadronized with Pythia6.

3 Trigger

Events are selected if one of the following triggers has fired: HLT_HT750, HLT_PFHT650, HLT_PFNoPUHT650 HLT_FatDiPFJetMass750_DR1p1_Deta1p5. All versions of each of these triggers is used. None of these triggers are prescaled druing the 2012 data taking period. HLT_PFNoPUHT650 trigger is used for the data set after the RunC(including RunC), while HLT_PFHT650 trigger is only used for RunA and RunB data sets.

Figs 3, 4 and 5 show the trigger efficiency. The trigger efficiency has been measured with respect to a lower-threshold, but prescaled, HLT_HT550 trigger. The trigger is 99% effiecient above 890GeV for the untagged, HbbVqq-tagged , and HwwVqq-tagged data.

Fig 2 shows the turn-on curve of the reference trigger on the signal MC. 1.0 TeV signals are used here in Fig 2,and the plot shows HwwVqq and HbbVqq signals are, fully efficient for HLT_HT550 trigger, which is not prescaled in MC. So other signals, having resonance mass bigger than 1.0 TeV, will surely be fully efficient for the triggers.

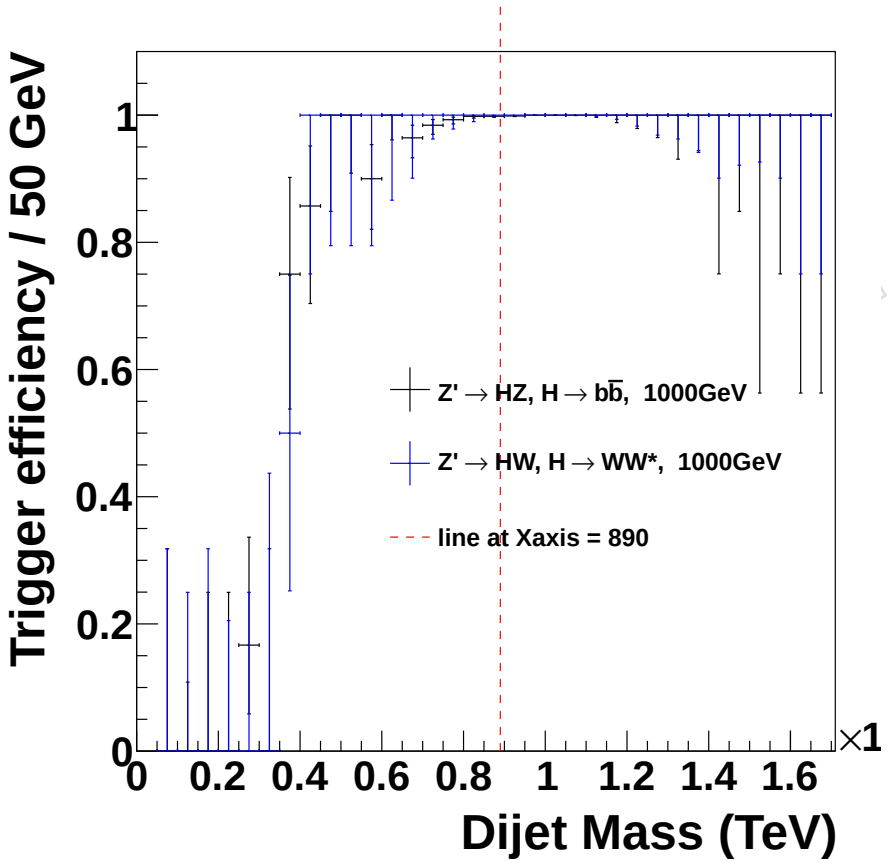


Figure 2: Reference trigger efficiency of signal MC.

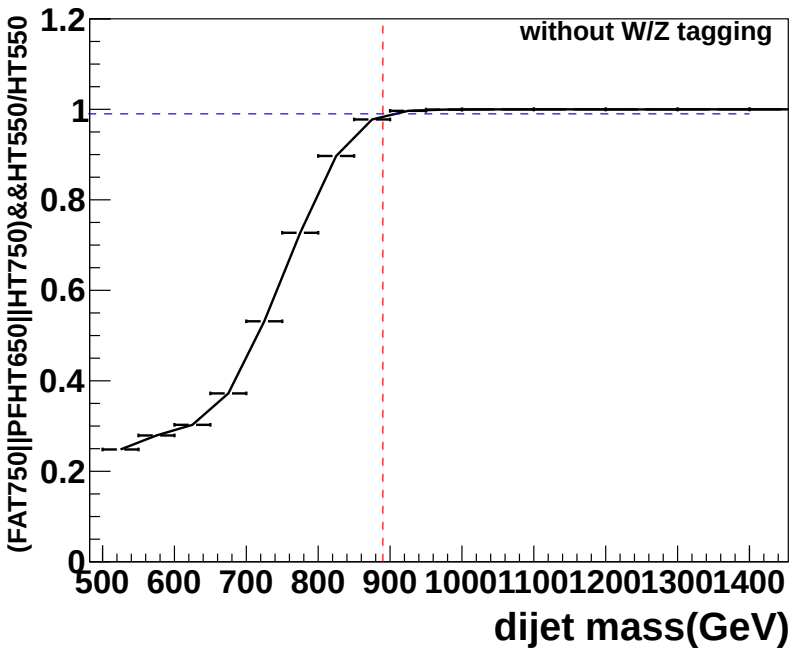


Figure 3: trigger efficiency for untagged data of fat_750||hlt_pf(nopu)ht650||hlt_ht750 measured using data collected by lower threshold h_t550 trigger. the dashed red line is drawn at m_{jj} equal 890 GeV, the blue line is at efficiency at 99%.

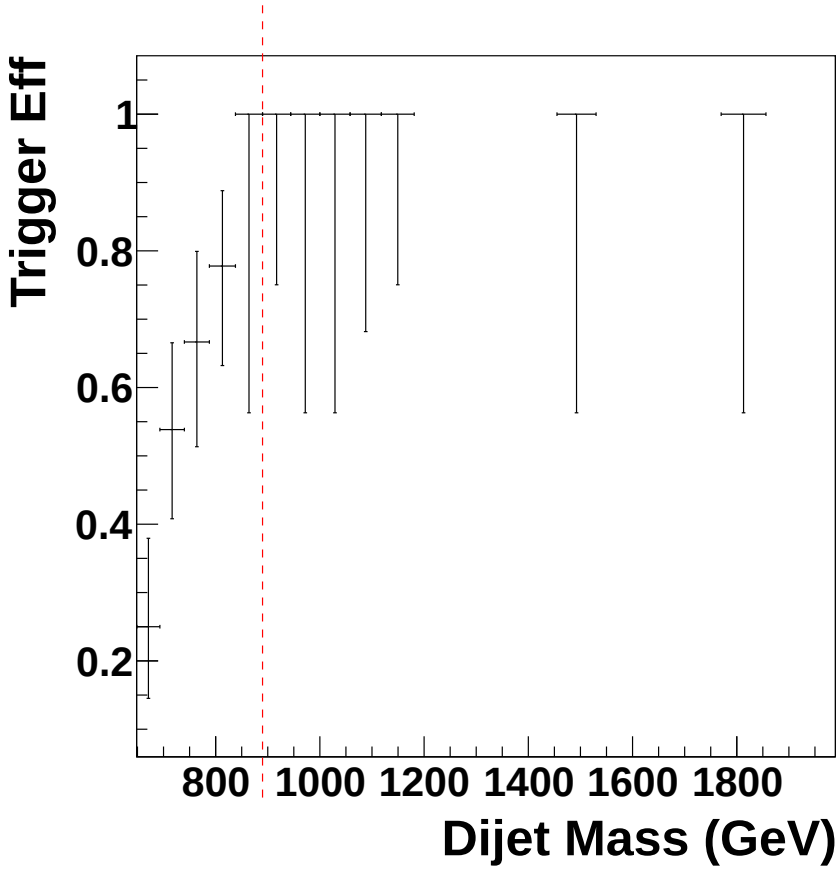


Figure 4: Trigger efficiency for HbbVqq tagged data of FAT_750||HLT_PF(NoPU)HT650||HLT_HT750 measured using data collected by lower threshold H_T550 trigger. The dashed red line is drawn at m_{jj} equal 890 GeV.

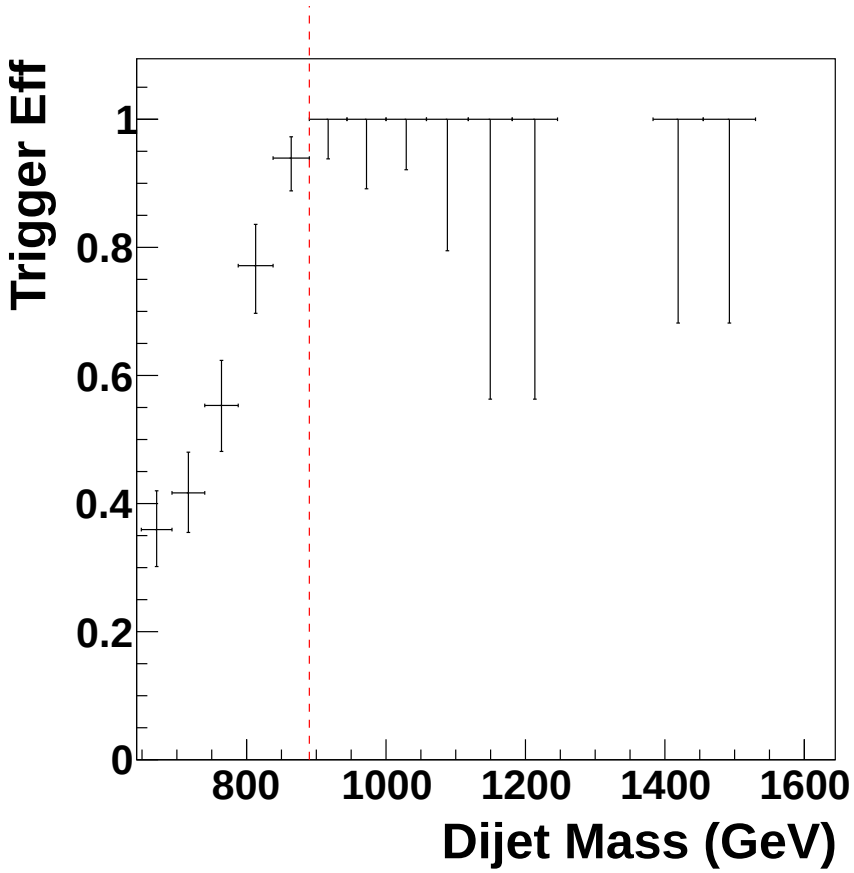


Figure 5: Trigger efficiency for HqqqqVqq tagged data of FAT_750||HLT_PF(NoPU)HT650||HLT_HT750 measured using data collected by lower threshold H_T550 trigger. The dashed red line is drawn at m_{jj} equal 890 GeV.

4 Event preselection

4.1 Jet reconstruction

Based on CMSSW 5.3.x software package, events are reconstructed using the particle-flow reconstruction algorithm [30] that attempts to reconstruct 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 Cambridge-Aachen (CA) [31, 32] jet clustering algorithm, as implemented in FastJet version 3.0.1 [33, 34], 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 $R = 0.8$ is used for the CA algorithm. These jets are hereafter referred to as CA8 jets.

Charged hadrons identified as pileup are removed from the list of PF candidates input to the jet clustering algorithms. The remaining neutral component of pileup is removed by applying a residual area-based correction as described in Ref. [35, 36]. The mean 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 [37, 38].

The jet energy corrections for the CA8 jets are derived from studies using the anti-k_T $R = 0.7$ jet algorithm. Simulation studies confirm that these anti-k_T-derived jet corrections are adequate for the CA8 jet algorithm for the jet momenta considered here [17].

4.2 Event cleaning

Events are selected using the following cuts:

- The event must have a well reconstructed primary vertex as computed by a deterministic annealing filter (DAF) ($|z_{\text{Primary Vertex}}| < 24$ cm, $N_{\text{DOF}} > 4$).
- The following recommended noise event filters are used:
  - CSC tight beam halo filter
  - HBHE noise filter with isolated noise rejection
  - HCAL laser event filter (HBHE) and HCAL laser event filter 2012
  - ECAL dead cell trigger primitive (TP) filter
  - Bad EE supercrystal filter
  - The tracking failure filter
  - Good primary vertex filter
  - Tracking coherent noise filter
  - Tracking TOBTEC fakes filter
- The events are required to have at least two ungroomed CA8 jets with
  - $p_T > 30\,\text{GeV}$, $|\eta| < 2.5$
  - to have muon energy fraction < 0.8
  - pass tight particle flow jet ID. The tight PF jet ID is listed below:
    - Neutral Hadron (EM) Fraction $< 0.90 (< 0.90)$, for all jet η
    - Number of Constituents > 1, for all jet η

- Charged Hadron (EM) Fraction $> 0 (< 0.99)$, for jet $|\eta| < 2.4$
- Charged Multiplicity> 0, for jet $|\eta| < 2.4$

- Beam background events are removed using the following requirements:
  - In events with at least 10 tracks, a minimum of 25% of these tracks must be high purity tracks.
- The events must pass $|\Delta\eta| < 1.3$, $m_{jj} > 890\,\text{GeV}$

This sample of dijet events is then searched for W/Z and Higgs bosons. One jet must be identified as a hadronically decaying W or Z boson, whereas the other jet must be identified as one of the two types of hadronic Higgs decays.

5 The H tagging and W/Z tagging algorithms

The products of hadronic decays of Higgs, W, and Z bosons can fall within a single jet if these particles are highly boosted. In this analysis, we aim to cover as much of the Higgs branching ratio as possible. The Standard Model Higgs with a mass of 125 GeV decays to $b\bar{b}$ with a branching fraction of 57.7%, and to WW^* with a braching fraction of 21.4%. [39]. Using these two decay modes in a VH search, where WW^* specifically decays to four quarks, is the main topic of this note. (The semileptonic decay mode $H \to WW \to 2q\ell\nu$ is viable, but its reconstruction is more involved and will be covered in a subsequent analysis.)

The algorithms to identify W/Z, $H \to b\bar{b}$ and $H \to WW^*$ jets are necessarily different, but they use similar jet-level variables: N-subjettiness (described in Section 5.1) and jet pruning (Section 5.2). The W/Z-tagger is described in Section 5.3, and the two H-taggers in Sections 5.4 and 5.5.

5.1 N-subjettiness

N-subjettiness [40–42] exploits the fact that the pattern of the hadronic decay of a heavy object is reflected through the presence of distinctive energy lobes corresponding to the decay products, as opposed to QCD jets which present a more uniformly spread energy configuration. The inclusive jet shape N-subjettiness is defined, in its generalized version [40], as

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k}\, min((\Delta R_{1,k})^\beta, (\Delta R_{2,k})^\beta ... (\Delta R_{N,k})^\beta) \tag{1}$$

where the index k runs over the jet constituents and the distances $\Delta R_{n,k}$ are calculated with respect to the axis of the n^{th} subjet. The normalization factor d_0 is calculated as $d_0 = \sum_k p_{T,k} R_0^\beta$, setting R_0 to the jet radius of the original jet. In the analysis, we use onepass_kt_axes definition of subjet axes and the N-subjettiness is calculated from the unpruned jets with the parameter $\beta = 1$. It has been shown in the literature [40] that the best way to separate (N+1)-prong from the N-prong decays merged into a single jet is to select jets with low value of the ratio τ_{N+1}/τ_N. In this vein, the variable able to best discriminate between W/Z jets and QCD jets is $\tau_{21} = \tau_2/\tau_1$. The distribution of τ_{21} for the VV signal and QCD background is given in Fig. 6.

5.2 Jet Pruning

Jet pruning consists of the removal of the softest components of the jets [17, 43]. The jet pruning algorithm uses the CA $R = 0.8$ jets as inputs. In the process, soft and wide-angle particles (relative to the parent in the clustering) are ignored and are not clustered. The same parameters are chosen for the jet pruning algorithm as in the original paper [44, 45].

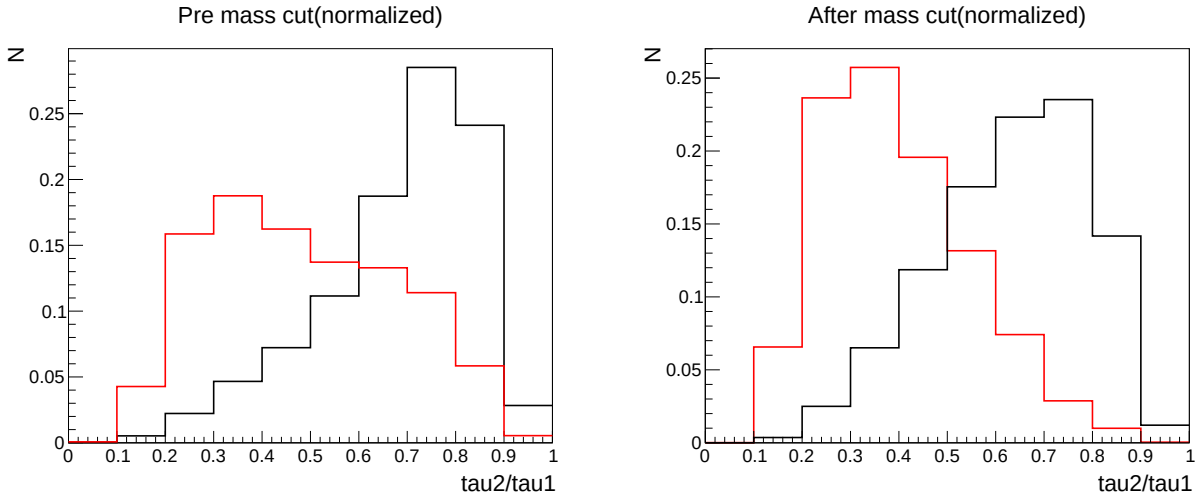


Figure 6: Comparison for τ_2/τ_1 distribution between signal (red) and background (black) before the jet mass cut (left) and after the jet mass cut of [70, 100] GeV applied (right). The signal MC used here is Herwig WW 1.5 TeV, and background is Herwig QCD.

The result of jet pruning on the CA8 jets is two fold, i.e., the invariant jet mass reconstruction and subjet indentification. In all cases, we use the jet invariant mass computed from the whole (or "fat") pruned jet. This quantity is referred below to as the pruned jet mass. For W/Z tagging, we use pruned jet mass between 70 and 100 GeV. For the identification of Higgs jets, we require the pruned jet mass to lie between 110 and 135 GeV. The distribution of the pruned jet mass of the Higgs candidate jet is shown on Fig. 7

The main role of jet pruning is to allow better delineation of subjets within the jet. In $H \to b\bar{b}$ tagger, the axes of the pruned subjets are used as the basis for b tagging.

5.3 W/Z tagging

For the identification of W/Z jets, we employ the same tagging algorithm previously used in published searches [21]. W/Z jets are selected using the following requirements:

- **Pruned jet mass m_{jet}** - Require the total pruned jet mass to satisfy $70\,\text{GeV} < m_{\text{jet}} < 100\,\text{GeV}$.
- **N-subjettiness** - We split the events into two categories, "high purity" W/Z jets by requiring $\tau_{21} \leq 0.5$, while $0.5 < \tau_{21} < 0.75$ defines the "low purity" W/Z jets. The thresholds are taken from the published VV search.

The performance of the W/Z tagger has been documented in detail in Ref. [18].

5.4 $H \to b\bar{b}$

To identify Higgs jets arising from the shower and hadronization of two collimated b quarks, we apply b tagging either on the two subjets or the fat jet , based on the angular separation of the two subjets, which is recommended by BTV-13-001 [16].

we use the following selection, syncrhonized with the radion search to $HH \to 4b$ [22] and the search for HW resonances in the semileptonic channel [46]:

- **Pruned jet mass m_{jet}** - Require the total pruned jet mass to satisfy $110\,\text{GeV} < m_{\text{jet}} < 135\,\text{GeV}$.

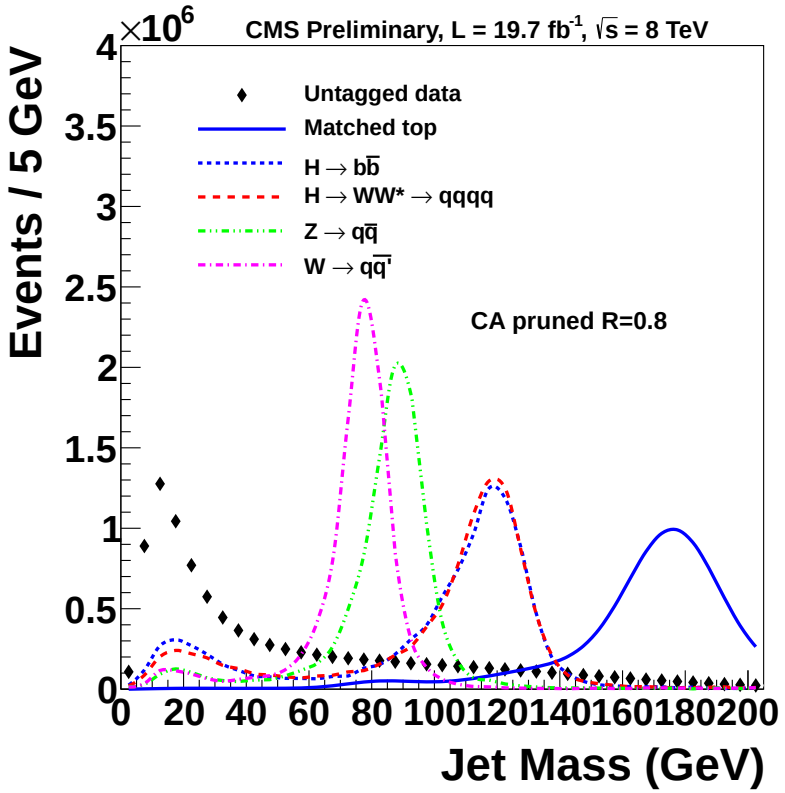


Figure 7: Pruned jet mass in signal MC, data and $t\bar{t}$ MC. MC samples are normalized to data. MC distributions are plotted as smooth curves connecting the histogram entries; the MC histograms have the same binning as the data. Higgs, W/Z and top jets are matched to their generator level particles, respectively.

- **Subjet b-tagging**
  - if ΔR between the CA8 subjets is bigger than 0.3: *both* subjets must pass the CSV Loose working point.
  - if ΔR between the CA8 subjets is smaller than 0.3: require the *fat* CA8 jet to pass the CSV Loose working point.

5.5 $H \rightarrow WW^* \rightarrow 4q$

In this channel, Higgs decays to two W bosons, one real and one virtual (denoted with an asterisk). Given that this is effectively a three-body decay $H \rightarrow Wqq$, the jets from the four quarks are not on an even footing – the subjets from the real W are harder, and they also form a W mass. The subjets from the softer two quarks are less well defined.

A naive $H \rightarrow 4q$ tagger would require a fat jet with four subjets. However, a study done using the subjets as defined by the CMS Top Tagging algorithm (which reruns the CA8 jet clustering with additional weak pruning [47]) removes $\approx$ 90% of the signal. Compounded with a decreasing angular separation between Higgs decay products, as a function of the Higgs p_T, at higher resonance masses, *e.g.* at 2 TeV, only 1% of signal passes this selection. The distribution of the number of subjets of the reconstructed Higgs jets in signal MC (obtained from the CMS Top Tagging) is shown in Fig. 8

As an alternative, we explore the N-subjettiness, in particular the variables involving τ_4. The

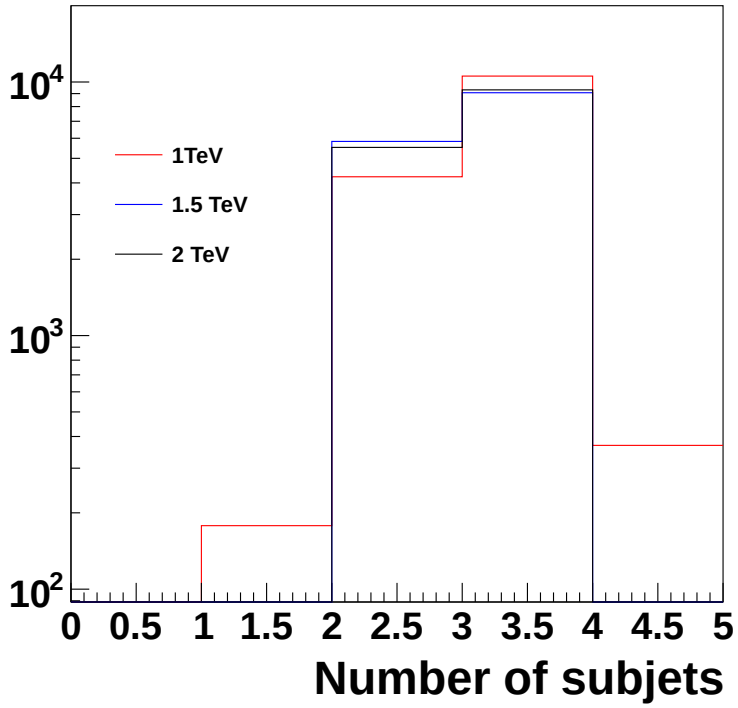


Figure 8: Number of subjets of the Higgs jets, in $H \to WW^* \to 4q, W/Z \to \bar{q}q'$signal MC. Subjets are abtained from the CMSTopTag jet collection.

ratio $\tau_{42} \equiv \tau_4/\tau_2$ has the best separation between the $H \to 4q$ signal and not only QCD background, but also Z and top jets. Figs 9 and 10 show the discriminating power of τ_{42} against $t\bar{t}$ and QCD, for 1 TeV and 2 TeV resonance masses respectively.

We also explore other combinations of $\tau_{NM} \equiv \tau_N/\tau_M$, which are listed in Appendix. E. The ROC (receiver operating characteristic) curve of for several τ_{NM} cuts (but the same pruned jet mass cut) is shown in Fig. 11. The signal efficiency is evaluated using Higgs jets in 2 TeV signal MC, and the false positive rate (*i.e.*, mistag rate) is derived from QCDPT300to470 MC sample. From the figure, it is clear that τ_{42} outperforms any other single τ_{NM} variable.

After optimizing the cut on τ_{42} (documented in Sec. 5.5.1 below), the full selection of the $H \to WW^* \to 4q$ tagger is:

- **Pruned jet mass** m_{jet} - We require the total pruned jet mass to satisfy $110\,\text{GeV} < m_{\text{jet}} < 135\,\text{GeV}$.
- **N-subjettiness** - We split the events into two categories, "high purity" Higgs jets by requiring $\tau_{42} \leq 0.55$, while $0.55 < \tau_{42} < 0.65$ defines the "low purity" Higgs jets.

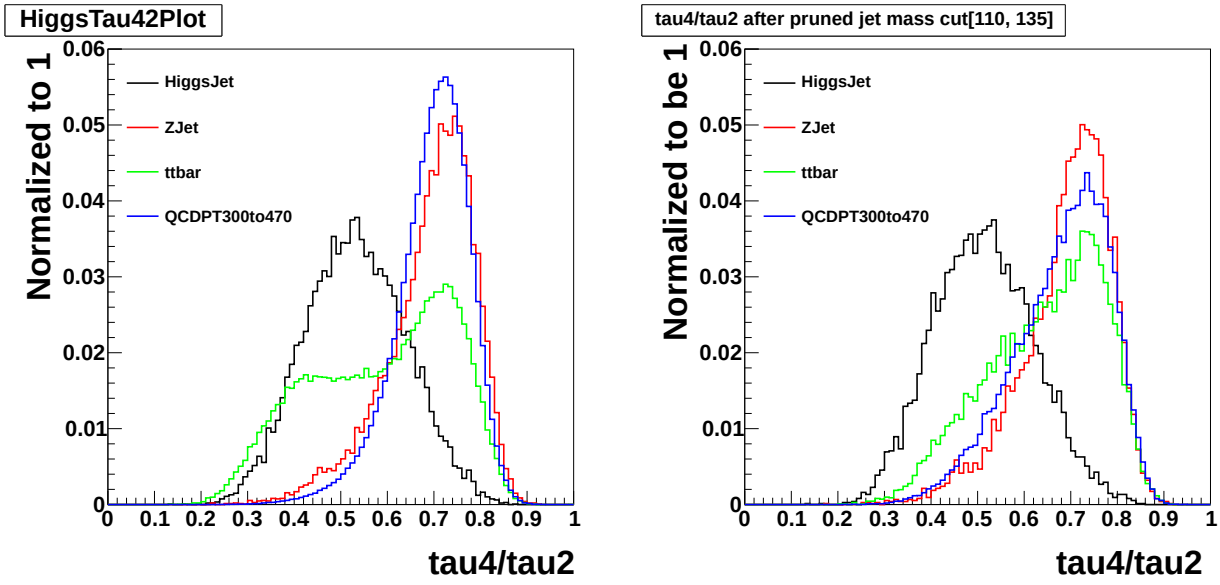


Figure 9: Distribution for τ_4/τ_2 in data and in simulations of signal (1.0 TeV) and background events. All simulated distributions are scaled to match the number of events in data, W/Z, matched top and Higgs jets are required to match their generator level particles, respectively.

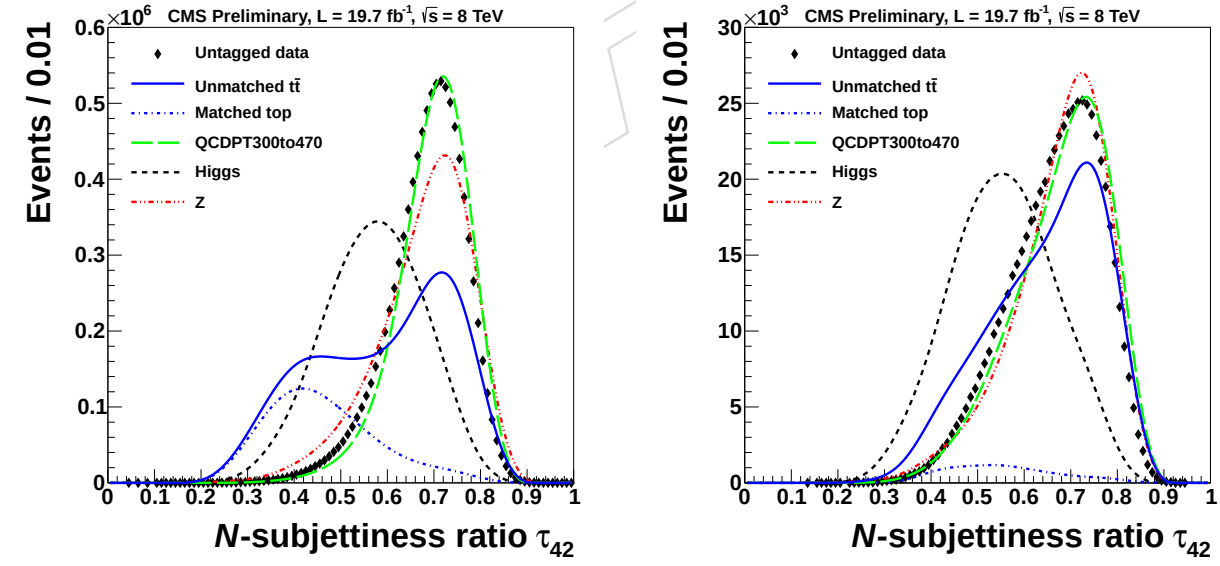


Figure 10: Distribution for τ_4/τ_2 in data and in simulations of signal (2.0 TeV) and background events. All simulated distributions are scaled to match the number of events in data, except that matched top is scaled to its fraction of unmatched $t\bar{t}$ times the number of data events. W/Z, matched top and Higgs jets are required to match their generator level particles, respectively.

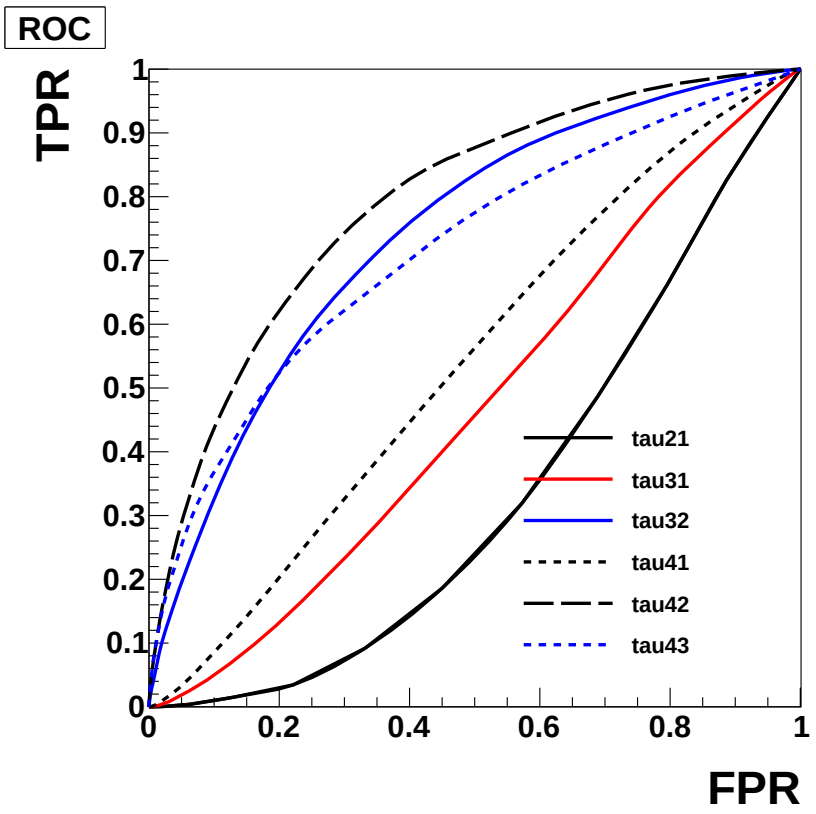


Figure 11: ROC curves for different τ_{NM} after the cut on the pruned jet mass. The false positive rate (FPR) is obtained from QCDPT300to470 and the true positive rate (TPR) from Higgs jets in 2 TeV signal MC sample. Using τ_{42} to select Higgs jets outperforms all other τ_{NM} variables.

5.5.1 Optimization of the τ_4/τ_2 threshold

Having selected τ_{42} as the discriminating variable, we next optimize its upper value. In this study, the jet mass is confined within $[110, 135]$ GeV. We use the limit setting method (described in Sec. 9) and evaluate the expected limits of several signal resonance masses at different τ_{42} working points. These expected limits are presented in Table. 3. Given our focus on the resonance masses above 1500 GeV, we choose to cut on $\tau_{42} < 0.55$. In the following analysis, to compensate the signal efficiency loss at higher resonance mass, we introduce an additional categories for $H \to WW^* \to 4q$ tagger as $0.55 < \tau_{42} < 0.65$. This is chosen from back-of-envelope calculation based on Figs 9 and 10, since this category provides very limited sensitivity.

Table 3: Upper limits (in units of 0.01 pb) for high purity HW and HZ signals at different resonance masses and also different τ_{42} working points.

HW / τ_{42}	0.45	0.5	0.55	0.6
1000	4.14	4.09	4.46	4.91
1500	0.97	0.88	0.86	0.91
2000	0.89	0.64	0.51	0.47
2500	1.36	0.82	0.53	0.40
HZ / τ_{42}				
1000	4.31	4.36	4.63	5.05
1500	0.98	0.89	0.86	0.90
2000	0.70	0.55	0.42	0.39
2500	0.96	0.61	0.41	0.32

6 Signal efficiency

We search for several models of heavy resonances decaying to a W or Z boson on one side, and a Higgs on the other, where both bosons decay to quarks producing merged jet. This analysis is focused on two channels:

- $H \to b\bar{b}, W/Z \to \bar{q}q'$, and
- $H \to WW^* \to 4q, W/Z \to \bar{q}q'$

As previously discussed, we use one V-tagging and two Higgs tagging algorithms to identify such events. After subdividing the events according to high purity and low purity tags, we end up with five distinct categories, as shown in Table 5.

In this section, we discuss various issues related to the evaluation of the signal efficiency.

6.1 Cross-talk between the Higgs decay channels

In order to combine events from all categories into a single joint likelihood, the categories must be mutually exclusive. However, a cross-talk between the Higgs channels is neverthelss possible: for example, $H \to b\bar{b}$ tagger can identify other two-prong Higgs decay modes like $H \to gg$, $H \to \tau\tau$, or $H \to c\bar{c}$, although this kind of 'false positive' tag happens only rarely (the efficiency is $\lesssim 6$ %). Similarly, events from two-prong Higgs decay channels can also pass the τ_{42} cut in the $H \to WW^* \to 4q$ selection. In this case, the channel $H \to b\bar{b}$, because of its large branching ratio, contributes a non-negligible number of events to the sample of 4q tags. This effect is illustrated by Fig. 12, where it can be seen that most of the low-τ_{42} tail of the $H \to b\bar{b}$ curve will be below the cut value of 0.55.

Table 4 provides an overview of the cross-talk between the various channels. The Higgs branching ratios correspond to the Higgs mass of 125 GeV. For $H \to WW^* \to 4q$, the branching ratio of the hadronic decay of (real) W boson is already included, so that the final state is four quarks. The table is normalized to 100,000 standard model Higgs bosons, and the numbers in the table show the number of Higgs decays that pass the tagger for each channel, with the branching ratio taken into account. For example, let us consider $H \to c\bar{c}$ channel. At the Z' resonance mass of 1 TeV, out of 100,000 Higgs decays, 104 events are tagged by the $H \to b\bar{b}$ tagger and 82 pass $H \to WW^* \to 4q$ tagger but fail $H \to b\bar{b}$ tagger. For $H \to ZZ$ decays, we take its tagging efficiency the same as $H \to WW^* \to 4q$ signals. So the number of $H \to ZZ$ to pass $H \to b\bar{b}$ and $H \to WW^* \to 4q$ tagger is estimated by efficiency of $H \to WW^* \to 4q$ signal times BR(Hzz)*BR(Zqq)*BR(Zqq) divided by BR(Hww)*BR(Wqq)*BR(Wqq).

From Table 4, it can be seen that for various signal resonance masses, the contribution of other decay channels to the sample of $H \to b\bar{b}$ tags never exceeds 6%. We will assign additional systematics for the cross-talk.

Since the $H \to b\bar{b}$ tagger has significantly lower background than $H \to WW^* \to 4q$, it takes precedence in selecting events: we first identify the events that pass the $H \to b\bar{b}$ tagger, and only if they fail, we test them for the presence of the $H \to WW^* \to 4q$ tag.

The effect of the $H \to b\bar{b}$ tagger veto on the $H \to WW^* \to 4q$ tagged dijet mass distribution background (data) is shown in Appendix F.

Table 4: Number of Higgs jets falls into two exclusive categories, assuming we have 100,000 SM Higgs (125 GeV) decays to all channels. $H \to ZZ^* \to 4q$ signals are estimated by its branching ratio times the efficiency of $H \to WW^* \to 4q$ signals divied by the branching ratio of $H \to WW^* \to 4q$ channel.

	Branching ratio (%)	Pass $H \to b\bar{b}$	Fail $H \to b\bar{b}$, pass $H \to WW^* \to 4q$
1.0 TeV			
$H \to b\bar{b}$	57.70	11871	804
$H \to WW^* \to 4q$	9.94	86	2360
$H \to ZZ^* \to 4q$	1.30	11	309
$H \to c\bar{c}$	3.00	104	82
$H \to \tau\tau$	6.30	17	37
$H \to gg$	10.00	14	314
1.5 TeV			
$H \to b\bar{b}$	57.70	11444	755
$H \to WW^* \to 4q$	9.94	228	1916
$H \to ZZ^* \to 4q$	1.30	29	250
$H \to c\bar{c}$	3.00	121	88
$H \to \tau\tau$	6.30	12	57
$H \to gg$	10.00	69	174
2.0 TeV			
$H \to b\bar{b}$	57.70	13816	551
$H \to WW^* \to 4q$	9.94	449	1435
$H \to ZZ^* \to 4q$	1.30	58	187
$H \to c\bar{c}$	3.00	228	99
$H \to \tau\tau$	6.30	42	74
$H \to gg$	10.00	157	262

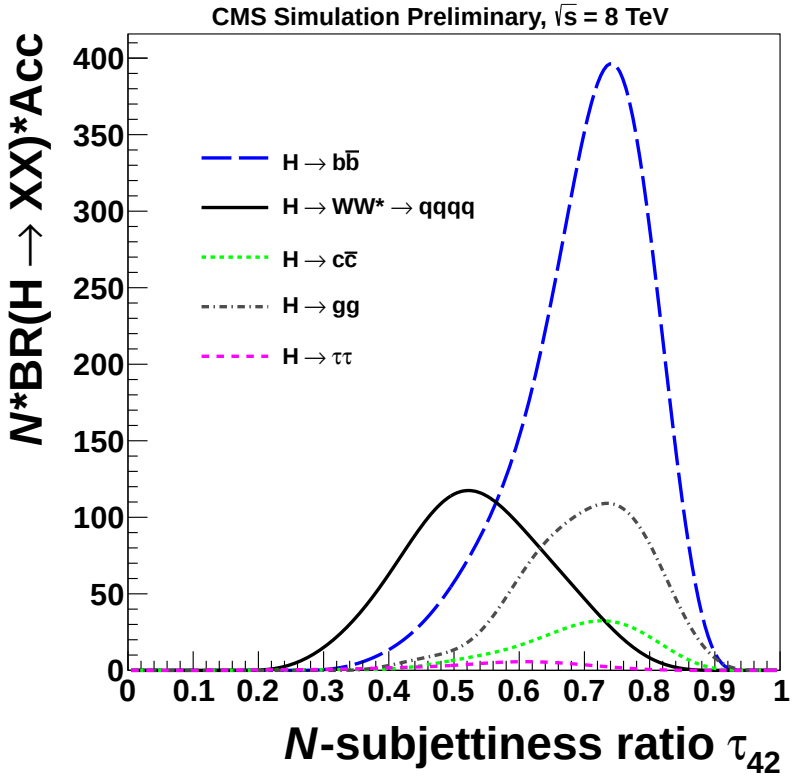


Figure 12: Comparison of τ_{42} distribution for H(ww$\rightarrow$ $qqqq$), H(gg), H(bb), H(cc), H($\tau\tau$) channels. All curves are drawn for Higgs jets after pruned jet mass cut, and also normalized with respective branching ratios.

6.2 Summary of Higgs and W/Z tagging categories

The W or Z jets from the signal are selected by the V-tagger, and the Higgs candiadates are selected by an OR of the two Higgs taggers, $H \to b\bar{b}$ and $H \to WW^* \to 4q$. Both V-tagger and $H \to WW^* \to 4q$ taggers have high-purity and low-purity categories. The latter are added to increase the sensitivity of the analysis at high resonance masses, where the QCD background is low, and a higher signal efficiency is at the premium.

We first identify the events that pass the $H \to b\bar{b}$ tagger, and only if they fail, we test them for the presence of the $H \to WW^* \to 4q$ tag. Thus we arrive at the final division of events into mutually exclusive categories listed in Table 5. For the $H \to WW^* \to 4q, W/Z \to \bar{q}q'$ channel, we drop the low-purity Higgs and low-purity V-tagging category, because it adds only a negligible sensitivity.

Table 5: The five exclusive event categories used in this analysis. We also assign specific names (in parenthesis) for each category, which will be used in following sections.

$H \to b\bar{b}, W/Z \to \bar{q}q'$	$H \to WW^* \to 4q, W/Z \to \bar{q}q'$
high-purity V-tag (Hbb1)	high-purity H-tag, high-purity V-tag (Hww1)
low-purity V-tag (Hbb2)	high-purity H-tag, low purity V-tag (Hww2)
	high-purity V-tag, low purity H-tag (Hww3)

The events from the $H \to b\bar{b}, W/Z \to \bar{q}q'$decay could contribute to all the five categories, due to its large branching ratio. The $H \to WW^* \to 4q, W/Z \to \bar{q}q'$signal events contribute only in events that fail $H \to b\bar{b}$ but pass $H \to WW^* \to 4q$ tagger; their contribution to $H \to b\bar{b}$ tagged sample is negligible. The contributions from other Higgs decay modes to all these five categories is tiny compared to $H \to b\bar{b}, W/Z \to \bar{q}q'$and $H \to WW^* \to 4q, W/Z \to \bar{q}q'$yields. We will not specificly study them, but include them as systematic uncertainties.

6.3 Tagging efficiency for $H \to b\bar{b}$ jets

We study the Higgs tagging efficiency in MC by matching the jet to the Higgs generator-level particle. This jet is referred to as the Higgs jet. The Higgs tagging efficiency is obtained from the MC simulation as the fraction of the Higgs jets that passes the given H-tagging selection. It is given in Fig. 13. The same Figure shows the W/Z tagging efficiency for the other jet in the event. The total event efficiency is a product of these two efficiencies.

In the $H \to b\bar{b}$ channel, the $H \to b\bar{b}$ tagging efficienc start rising after ~ 1.6 TeV. The reason is explained as follows. For the resonance masses above 1.6 TeV, the Higgs jets are sufficiently boosted that the ΔR between the two b subjets is ≤ 0.3. When $\Delta R \leq 0.3$, we are switching from 2 subjets b tagging to CSV loose fat jet b tagging. This causes the rising tagging efficiency.

Since the CSV tagging uses the cone of 0.3 to associate the candidate tracks to the jet, when the two subjets are at angular distance of 0.6, they begin sharing tracks. This effect becomes important for $\Delta R \leq 0.3 \sim 0.4$. For this reason, when the subjets are closer than 0.3, following the BTV POG recommendation, we switch to using the CSV b tagging decision for the fat jet. (We use CSVL, the loose operating point.)

Note that if for the fat jet b tagging CSVM operating point is used, the $H \to b\bar{b}$ tagging efficiency is smooth, as shown in Fig. 14. For the $H \to b\bar{b}, Z \to q\bar{q}$ analysis, we have compared the limits of these two different fat jet b tagging methods (CSVL *vs.* CSVM), Unsurprisingly, we have found that using the more-efficient CSVL b tagging results in better expected limits in

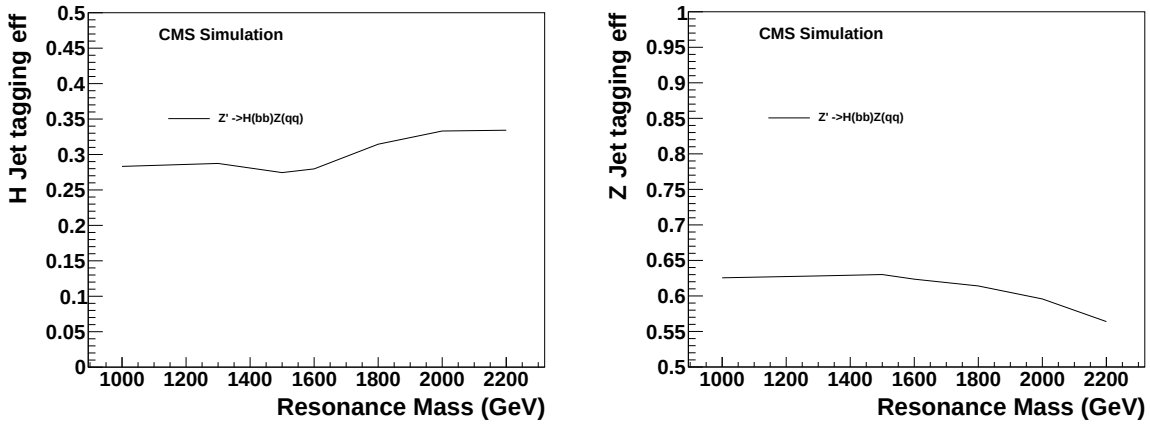


Figure 13: Higgs jets and Z jets tagging efficiencies in signal MC simulation. Left: $H \to b\bar{b}$. Right: $Z \to q\bar{q}$. The total event efficiency is a product of the two, resulting in a relatively flat efficiency for reconstructing $X \to HV$.

this background-poor region than using the CSVM b tagging operating point. (The details of this study are given in the Appendix.)

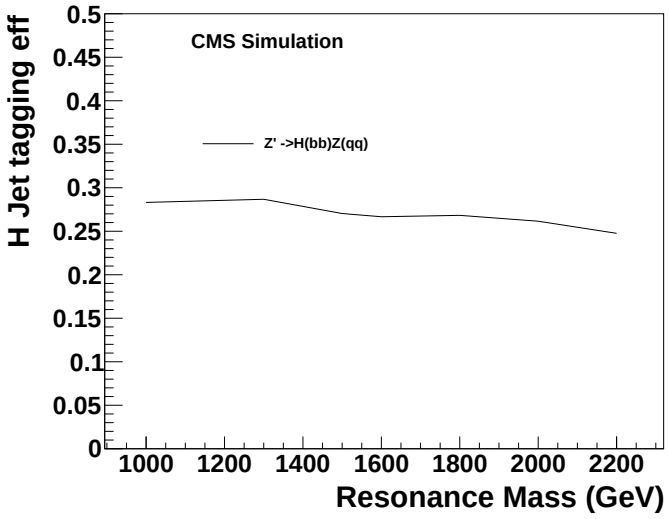


Figure 14: Higgs tagging efficiency in signal MC, for $H \to b\bar{b}$ channel. Changing fat jet b tagging to CSVM instead of CSVL.

6.4 Signal acceptance and total efficiency for $H \to b\bar{b}, Z \to q\bar{q}$ channel

Signal acceptance is defined as the number of signal events pass all the kinematic event selection (that is, without the two jet-tagging algorithms) divided by the number of generated events. The signal acceptance for $H \to b\bar{b}, Z \to q\bar{q}$ channel is shown on Fig. 15.

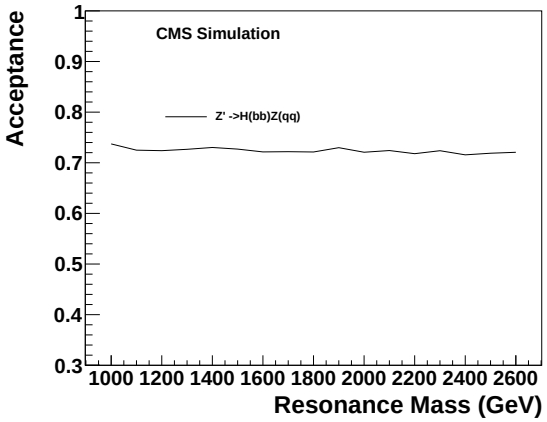


Figure 15: Acceptance in $H \to b\bar{b}, Z \to q\bar{q}$ signal.

The combined tagging rate of H and Z tagging, is defined as the number of events pass the HZ-tagging dvided by the number of events after events selection, which is shown in Fig. 16 for $H \to b\bar{b}, Z \to q\bar{q}$ tagging.

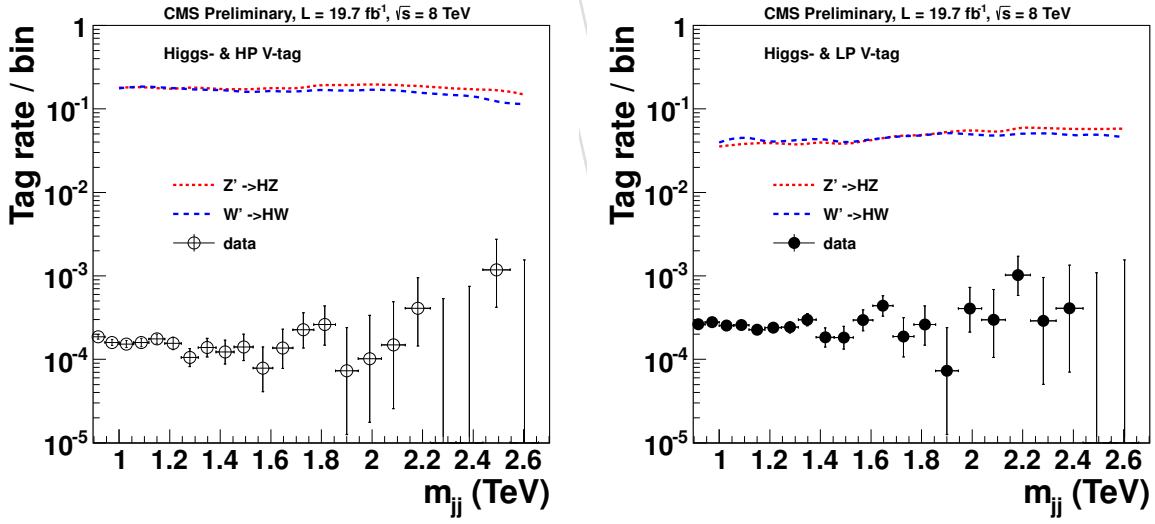


Figure 16: Tagging rates in $H \to b\bar{b}, Z \to q\bar{q}$ and signal channels and data. Horizontal bars in data indicates variable binning size.

6.5 Tagging efficiency for $H \to WW^* \to 4q$ jets

The tagging efficiency for $H \to WW^* \to 4q$ jets, as a fraction of getJets that passes the selection, is shown in Fig. 17.

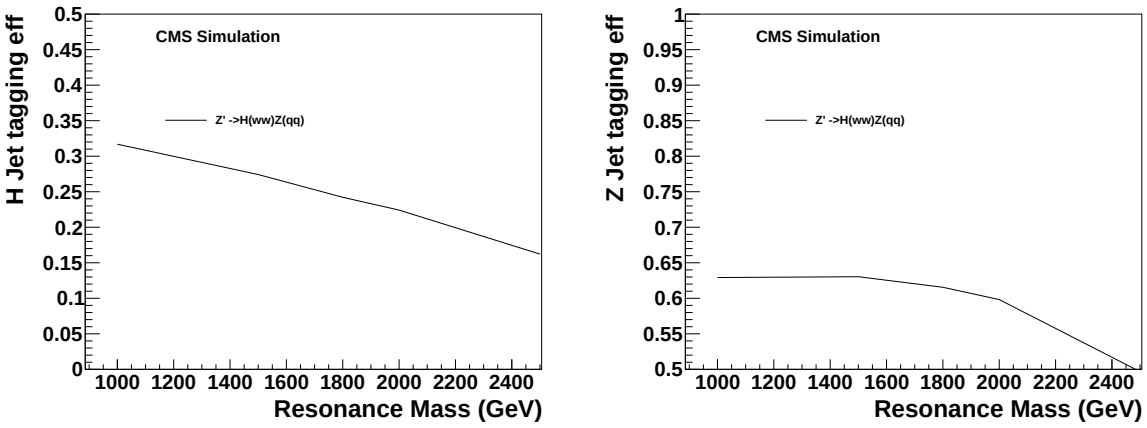


Figure 17: Higgs jets and Z jets tagging efficiencies in signal MC simulation. Left: $H \to WW^* \to 4q$. Right: $Z \to q\bar{q}$. The total event efficiency is a product of the two, resulting in a following spectrum efficiency for reconstructing $X \to HV$.

In the $H \to WW^* \to 4q$ all-hadronic channel, to compensate the efficiency loss in the high resonance mass, we also add two low purity categories, low purity H-tagging and low purity V-tagging. The tagging efficiency of low purity H/V-tagging on the H/Z jets is shown on Fig. 18. And low purity Higgs is defined as $0.55 < \tau_{42} < 0.65$, pruned jet mass in $[110, 135]$ GeV.

For low purity W/Z tagging, τ_{21} must be in $[0.5, 0.75]$, and the pruned jet mass in the window $[70, 100]$ GeV.

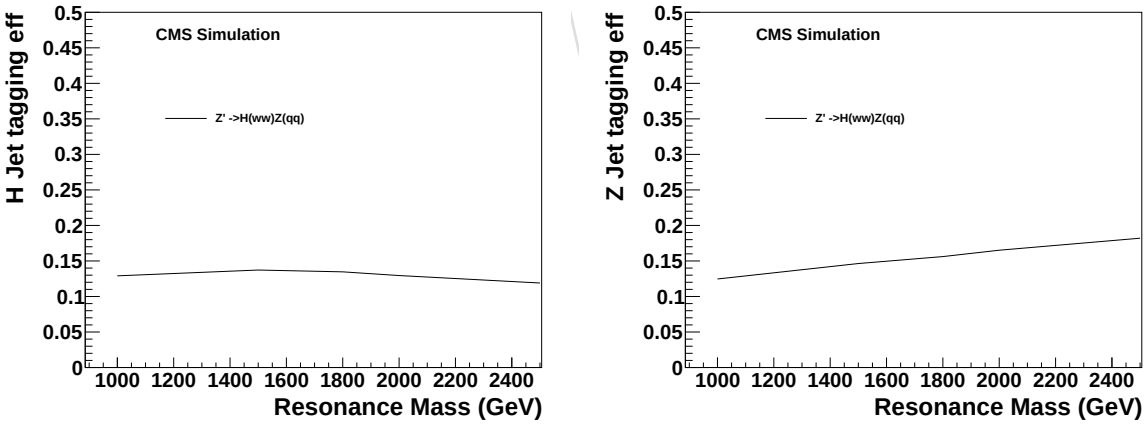


Figure 18: Lwo purity Higgs jets and Z jets tagging efficiencies in signal MC simulation. Left: $H \to WW^* \to 4q$. Right: $Z \to q\bar{q}$.

6.6 Signal acceptance and total efficiency for $H \to WW^* \to 4q, Z \to q\bar{q}$ channel

The acceptance of the $H \to WW^* \to 4q, Z \to q\bar{q}$ channel is shown on Fig. 19

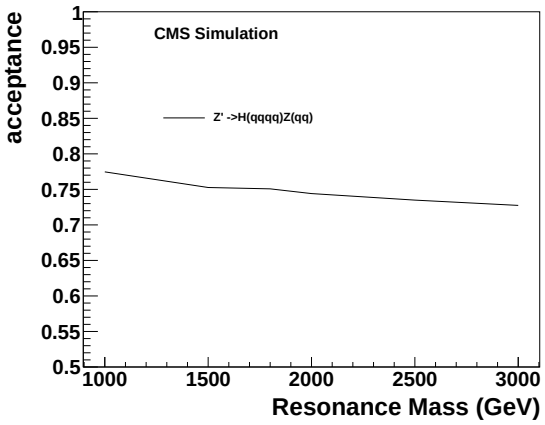


Figure 19: Acceptance in signal.

The combined tagging rates of H and Z tagging, for $H \to WW^* \to 4q, Z \to q\bar{q}$ channel in signal and data are shown on Fig. 20 .

The signal shapes are shown in Fig. 21

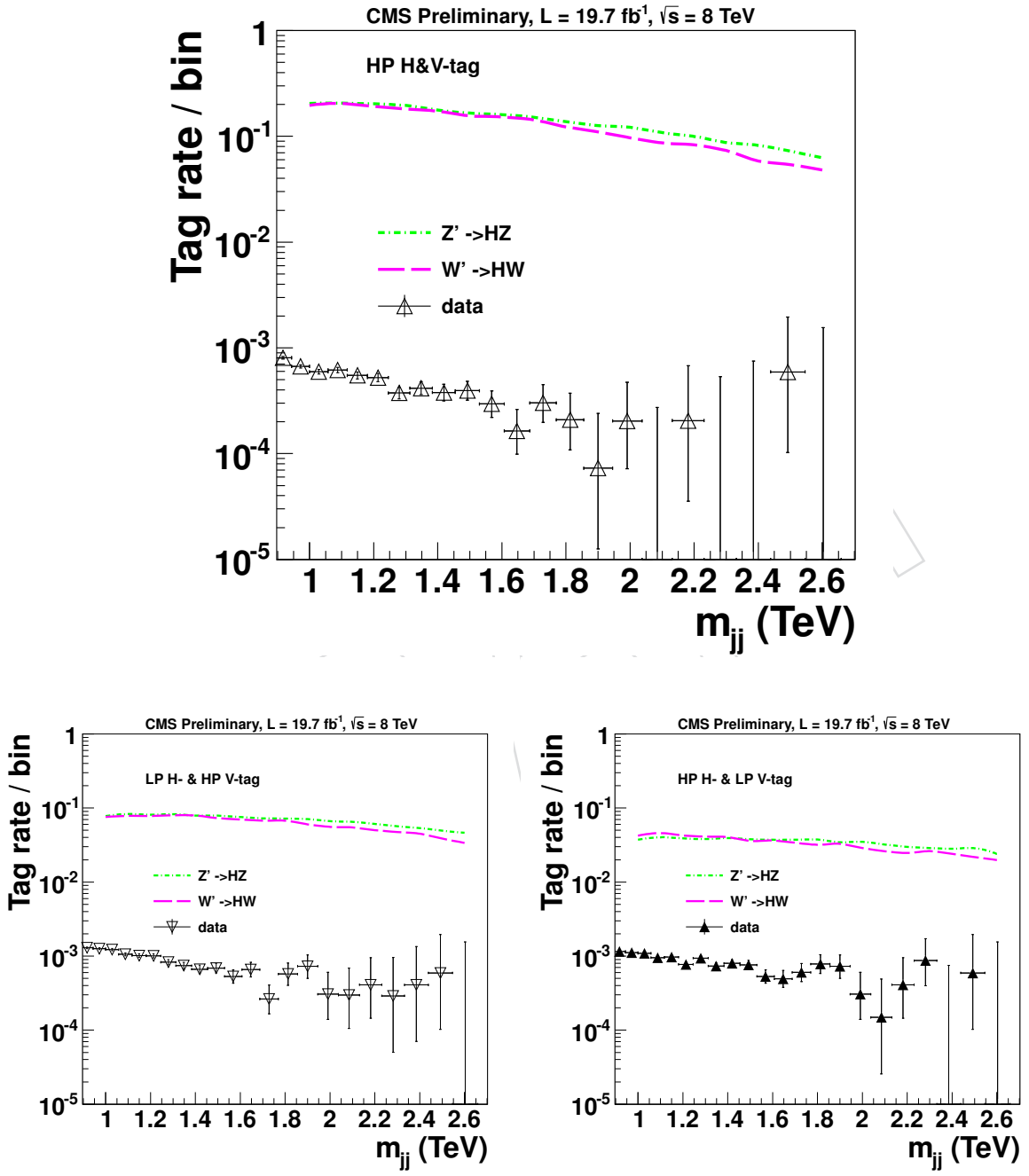


Figure 20: Tagging rates in $H \to WW^* \to 4q, Z \to q\bar{q}$ and signal channels and data. Horizontal bars in data indicates variable bins.

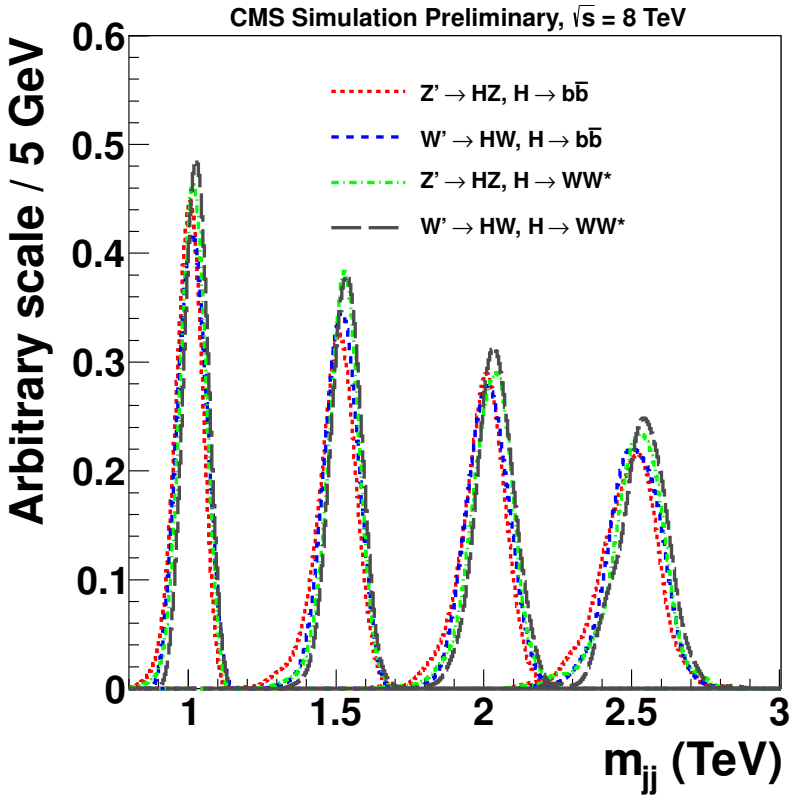


Figure 21: Signal shapes for Z′ and W′ signals at 1.0, 1.5, 2.0 and 2.5 TeV resonances.

7 Background shape parametrization

Background from multijet events is modelled by a smoothly falling distribution for each event category, given by the empirical probability density function

$$P_D(m_{\mathrm{jj}}) = \frac{P_0(1 - m_{\mathrm{jj}}/\sqrt{s})^{P_1}}{(m_{\mathrm{jj}}/\sqrt{s})^{P_2}} \,. \tag{2}$$

For each category, the normalization factor P_0 and the two parameters P_1 and P_2 are treated as uncorrelated. This parameterization was deployed successfully in searches in dijet mass spectra [19]. A Fisher F-test [48] is used to check that no additional parameters are needed to model the individual background distribution, for each of the four cases considered.

The background-only fit plot is only to demonstrate that the fit function works nice if no signal is assumed. However we don't use that function. We search for a peak on top of the falling background spectrum by means of a maximum likelihood fit to the data.

Figure 22 and Figure 23 show the dijet mass spectra from H(bb)V-tagged and H(ww)V-tagged data fitted to Eq. (2), respectively. The bottom panes show corresponding pull distributions, demonstrating the agreement between the background-only probability density function and the data.

The largest resonance mass in each channel is described in Table 6

Table 6: Highest resonance mass in each channel.

Category	Resoanance(GeV)
Hbb1	2481
Hbb2	2354
Hww1	2449
Hww2	2448
Hww3	2471

No sizeable deviation from the background-only hypothesis is seen, exclusion limits are set on the product of cross section, acceptance, and branching fraction for V' to HV search.

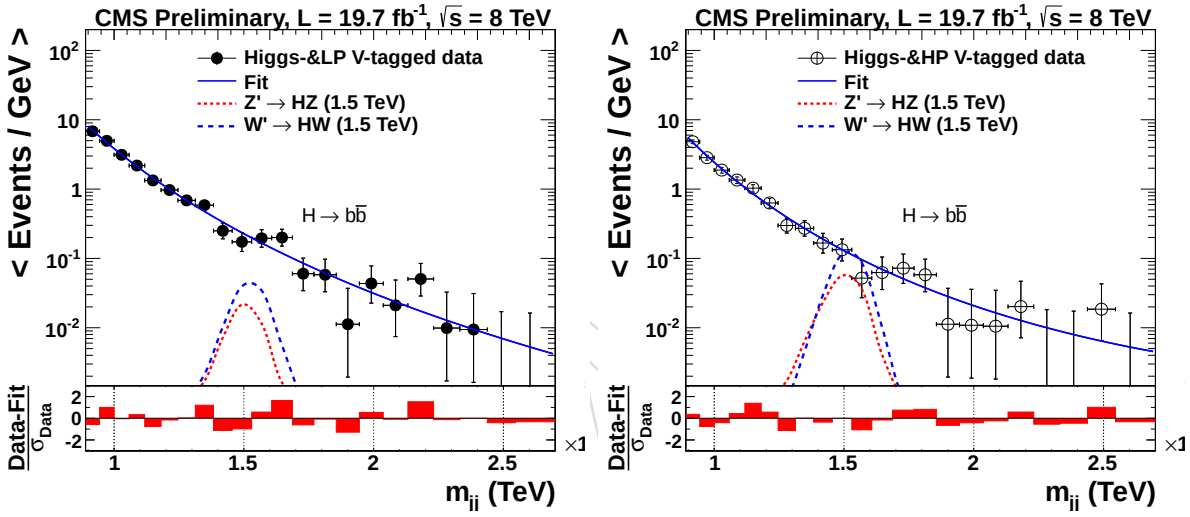


Figure 22: Distributions in m_{jj}, respectively, for LP V-tag(left), HP V-tag(right). The solid curves represent the results of fitting Eq. (2) to the data. The distributions for $H \to b\bar{b}, Z \to q\bar{q}$ and contributions, scaled to their corresponding cross sections, are given by the dash-dotted curves. Horizontal bars in data indicates variable binning size. The corresponding pull distributions ($\frac{\text{Data}-\text{Fit}}{\sigma_{\text{Data}}}$, where σ_{Data} represents the statistical uncertainty in the data in a bin in m_{jj}) are shown below each m_{jj} plot.

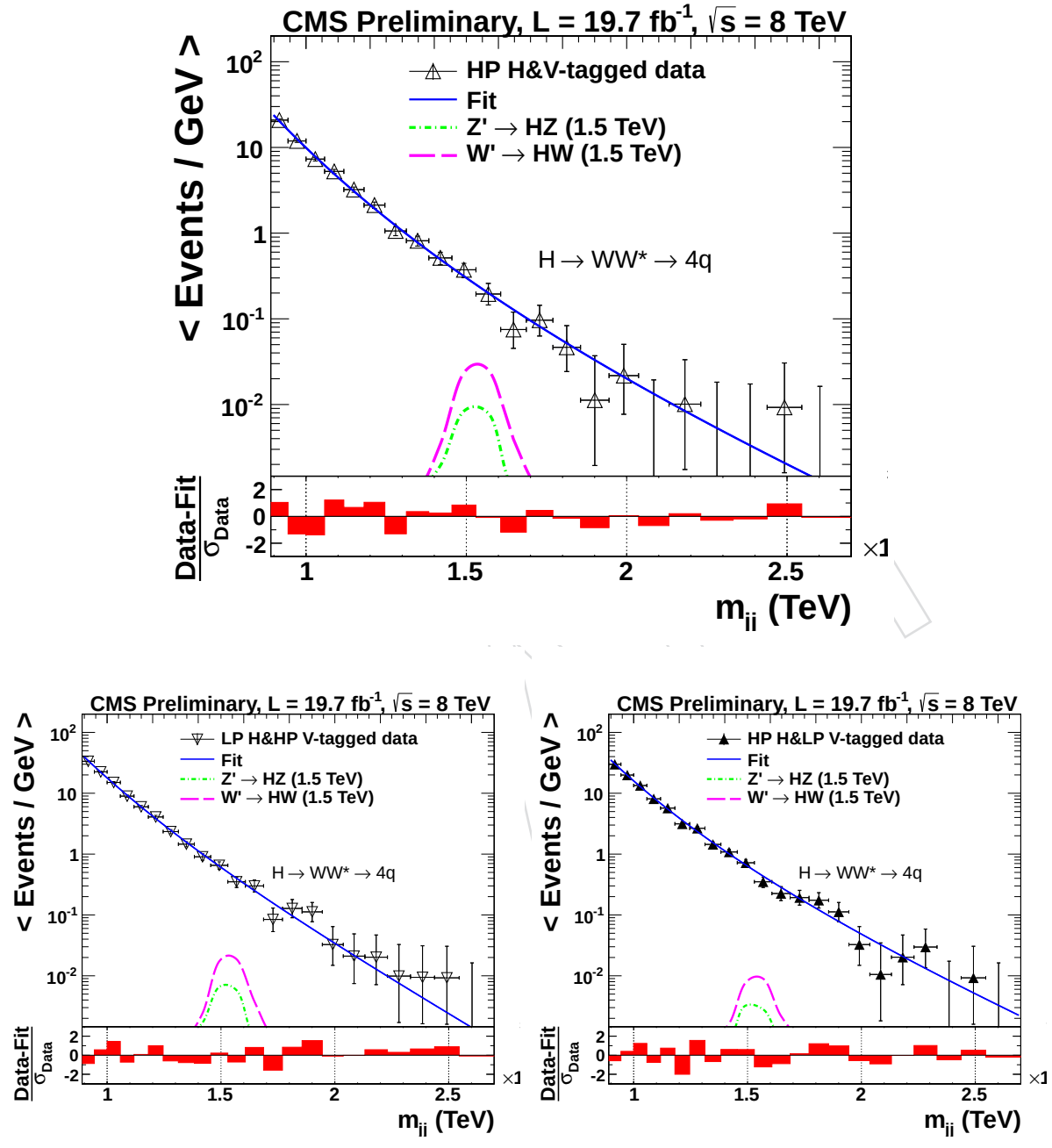


Figure 23: Distributions in m_{jj}, respectively, for HP(top), LP H-tag(bottom left) and LP V-tag(bottom right). The solid curves represent the results of fitting Eq. (2) to the data. The distributions for and $H \to WW^* \to 4q, Z \to q\bar{q}$ contributions, scaled to their corresponding cross sections, are given by the dash-dotted curves. Horizontal bars in data indicates variable binning size. The corresponding pull distributions ($\frac{\text{Data}-\text{Fit}}{\sigma_{\text{Data}}}$, where σ_{Data} represents the statistical uncertainty in the data in a bin in m_{jj}) are shown below each m_{jj} plot.

hyperref

8 Systematic uncertainties

The sources of systematic uncertainties are summarized here: :H-tagging efficiency, b-tagging scale factor, PDF uncertainties, W/Z-tagging efficiency, Jet Energy Scale(JES), Jet Energy Resolution(JER), luminosity, cross-talk of various signals.

8.1 b-tagging scale factor

We use method 1c recommended by the BTV group to apply b-tagging scale factor and uncertainties, which is `https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods#1c_Event_reweighting_using_scale`.

if ΔR of the two subjets is bigger than 0.3, we apply the b-tagging scale factor on the two subjets. While ΔR is between 0.3 to 0.4, we double the b-tagging scale factor uncertainties. When ΔR is smaller than 0.3, we apply the b-tagging scale factor on the fat jet. The uncertainty from b-taggging scale factor is evaluated by taking the largest difference in tagging efficiency by shifting b-tagging scale factor 1 σ, which results in 15% uncertainty.

8.2 W/Z-tagging efficiency

W/Z-tagging efficiency scale factor is studied in hadronic VV search [18, 21], details following. The W/Z-tagging efficiency is determined from the Monte Carlo simulation. We cross-check the MC modelling of the signal efficiency by measuring the W/Z-tagging efficiency in semileptonic $t\bar{t}$ data, and compare it with the same efficiency obtained using identical procedure from $t\bar{t}$ Monte Carlo sample generated with MadGraph [23] and showered with Pythia6 Tune Z2*. The ratio of the two efficiencies defines a scale factor, which is then applied to the efficiencies for signals in the dijet data.

Combining the efficiencies of the τ_{21} and jet mass cuts, a data-MC scale factor of 0.86 ± 0.07 (1.39 ± 0.75) for the high (low) purity selection for the W-tagging efficiency is determined. We assume that the same scale factor applies to Z-tagging as well. The errors on the scale factor are propagated into the systematic uncertainties on the overal signal efficiency.

The efficiency error on a single W/Z-tagging is estimated with a control sample of semileptonic $t\bar{t}$ events as described above. The uncertainties of 7.5% (54%) on the scale factors for high (low) purity tagging include sources from control sample statistics, pruned jet mass scale and pruned jet mass resolution. Since we estimate the scale factor only in the kinematic regime of the $t\bar{t}$ sample where the W decay products merge, but the b-quarks are still reconstructed as separate jets, we need to rely on the simulation to extrapolate to higher jet p_{T}.

Therefore, we estimate how the W/Z jets tagging efficiency varies as a function of p_{T} for two different showering and hadronization models using PYTHIA 6 and HERWIG++. We find that the differences are within 4% (12%) for the high (low) purity tagging [21], significantly smaller than the statistical uncertainties in the scale factors.

8.3 H-tagging efficiency for $\mathrm{H} \to \mathrm{WW}^* \to 4\mathrm{q}$ tagger

We extrapolate the H-tagging efficiency scale factor from the W/Z-tagging efficiency scale factor, more details are in Appendix C. In $\mathrm{H} \to \mathrm{WW}^* \to 4\mathrm{q}$, Higgs will decay to one on-shell W and one off-shell W. The on-shell W decays similarly as a single W/Z jet. while the off-shell W, is soft. So the overall Higgs jet could be viewed as a real W/Z jet plus soft part. To the first

order approximation, we apply the same scale factor to H-tagging as W/Z-tagging, but with an additional uncertainty.

For W/Z tagging, which uses τ_{21}, we know exactly the scale factor at low pt and need no additional uncertainty from Pythia/Herwig. So the Herwig efficiency as a function of pT can be normalized in such a way that Pythia and Herwig agree at low pT, but are different at high pT.

For τ_{42} we don't know if it is correct at low pT and high pT. Therefore, we should be more conservative and take the Pythia-Herwig difference without normalizing them at low pT.

We estimate how the Higgs jet tagging efficiency varies as a function of p_T for two different showering and hadronization models using PYTHIA 6 and HERWIG++. We find that the differences are within 7% (7%) for the high (low) purity tagging, the results are summarized in Table 7.

Table 7: The difference of $H \to WW^* \to 4q$ jet tagging efficiency in signal MC by showering and hadronization with Pythia and Herwig. Higgs jets are the jets matched to Higgs generator particles.

Resonance	1000GeV	1500GeV	2000GeV	2500GeV
High Purity Higgs jets	4.75%	2.28%	1.52%	6.31%
Low Purity Higgs jets	5.91%	6.33%	0.79%	4.39%

8.4 Other uncertainties

Because of the rejection of charged particles not originating from the primary vertex and also the application of pruning, the pileup dependence on the H/W/Z-tagging efficiency is weak, and the uncertainty of the modeling of the pileup distribution is less than 3%. Modeling of the underlying event, estimated by switching it off in PYTHIA 6, and also by comparing different tunes of PYTHIA 6, impacts the tagging efficiency by less than 1%.

In the jet p_T and η regions considered in this analysis, the Jet Energy Scale is known to a precision of 1-2% [37, 49]. For JES, p_T and η dependent uncertainty is propagated to the reconstructed dijet invariant mass, and taken into account by shifting the resonance dijet mass in the statistical analysis.

The Jet Energy Resolution(JER) is known to a precision of 10% and its tails are in agreement between data and MC [37]. The JER is taken into account in the statistical analysis by a variation of the resonance width by 10%.

The JES and JER of b jets are studied by JETMet, showing a smaller uncertainty than the standard QCD mixture of light quarks and gluon jets, whose JES and JER are applied for all flavor jets in this analysis. So we don't apply additional specific JES and JER uncertainty for b jets. The uncertainty on the pruned jet mass cut of $H \to WW^* \to 4q$ is included in the scale factor extrapolated from W-tagging. For $H \to b\bar{b}$ tagger, pruned jet mass uncertainty is evaluated as 2.6%, synchronized with EXO-12-053.

The uncertainty related to the PDF used to model the signal acceptance is estimated from the eigenvectors of the CT10, MRST2008 and NNPDF sets of PDF. The envelope of the upward and downward variations of the estimated acceptance of the three sets is assigned as uncertainty and found to be 5%-15% in the resonance mass range of interest.

The luminosity has been measured with an uncertainty of 2.6% [50], and is also taken into

account in the statistical analysis.

For the cross-talk of Higgs signals, we assign them as systematic uncertainty. In this analysis, signals are generated in exclusive decay channels, for example, $H \to b\bar{b}, W/Z \to \bar{q}q'$signals only have $H \to b\bar{b}$ decays, no other Higgs decays. So in category Hbb1 and Hbb2, the only signal is $H \to b\bar{b}, W/Z \to \bar{q}q'$, other Higgs decays passing the $H \to b\bar{b}$ tagger will be assigned as systematic uncertainties, which is evaluated according to Equation 3.

$$Uncertainty_{Cross-talk} = \frac{Numberof ExpectedNuisanceSignals + NumberofSignalOfInterest}{NumberofSignalOfInterest} - 1 \tag{3}$$

For $H \to b\bar{b}$ tagger, signal of interest is $H \to b\bar{b}$ signals. All other Higgs decays are taken as nuisance signals. For $H \to WW^* \to 4q$ tagger, signal of interest would be $H \to b\bar{b}$ plus $H \to WW^* \to 4q$ and all other Higgs decays are taken as nuisance signals. This number is evaluated across various resonance signal masses, resulting in 7% for category Hbb1 and Hbb2. This uncertainty is 31% for Hww1, Hww2 and Hww3, including 9.4% uncertainty from $H \to ZZ^* \to 4q$ decays, and 21% from Hcc,Hgg, etc. $H \to ZZ$ is taken as having the same efficiency as $H \to WW^* \to 4q$, as discussed in Table 4.

Table 8 shows a summarization of all the systematics applied.

Table 8: Summarization of systematics.Numbers in parenthesis are for low purity categories.

Systematics/Signals	Relevant quantity	$H \to b\bar{b}, W/Z \to \bar{q}q'$signals		$H \to WW^* \to 4q, W/Z \to \bar{q}q'$signals
		Hbb1, Hbb2	Hww1, Hww2, Hww3	Hww1,Hww2,Hww3
Background fit	Resonance shape	shape	shape	shape
Integrated Luminosity	Yield (per event)	2.6%	2.6%	2.6%
cross-talk	Yield (per event)	7.0%	31%	31%
Dijet mass shift due to JES	Resonance shape	1.0%	1.0%	1.0%
Dijet mass shift due to JER	Resoancne shape	10.0%	10.0%	10.0%
PileUp	Efficiency (per jet)	1.5%	1.5%	1.5%
PDF	Yield (per event)	5-15.0%	5-15.0%	5-15.0%
B-tagging SF	Yield (per event)	15.0%	1.0%	15.0%
Higgs mass	Efficiency (per jet)	2.6%	2.6%	-
W-tagging tau21	Efficiency (per jet)	7.5%(54%)	7.5%(54%)	7.5%(54%)
W-tagging tau21 shower/hadronization	Efficiency (per jet)	4%(12%)	4%(12%)	4%(12%)
H-tagging tau42	Efficiency (per jet)	-	7.5%(54%)	7.5%(54%)
H-tagging tau42 shower/hadronization	Efficiency (per jet)	-	7%(7%)	7%(7%)

9 Limit setting procedure

We search for a peak on top of the falling background spectrum by means of a maximum likelihood fit to the data. The likelihood $\mathcal{L}$, computed using events binned as a function of m_{jj}, is written as

$$\mathcal{L} = \prod_i \frac{\lambda_i^{n_i} \mathrm{e}^{-\lambda_i}}{n_i!}, \tag{4}$$

where $\lambda_i = \mu N_i(S) + N_i(B)$, μ is a scale factor for the signal, $N_i(S)$ is the number expected from the signal, and $N_i(B)$ is the number expected from multijet background. The parameter n_i quantifies the number of events in the i^{th} m_{jj} mass bin. The background $N_i(B)$ is described by the functional form of Eq. (2). While maximizing the likelihood as a function of the resonance mass, μ as well as the parameters of the background function are left floating.

The asymptotic approximation [51] of the LHC $\mathrm{CL_s}$ method [52, 53] is used to set upper limits on the cross sections for resonance production. The dominant sources of systematic uncertainties are treated as nuisance parameters associated with log-normal priors in those variables. For a given value of the signal cross section, the nuisance parameters are fixed to the values that maximize the likelihood, a method referred to as profiling. The dependence of the likelihood on parameters used to describe the background in Eq. (2) is removed in the same manner, and no additional systematic uncertainty is therefore assigned to the parameterization of the background.

10 All combined limits

Figure 24 shows the limits for combining $H \to b\bar{b}$ and $H \to WW^* \to 4q$ decaying channels together in the five categories. The Higgs and V bosons branching ratios are already taken into account. In HVT B model, W′ and Z′ are degenerate, having about the same mass. So we also show the combined limit of W′ and Z′ here.

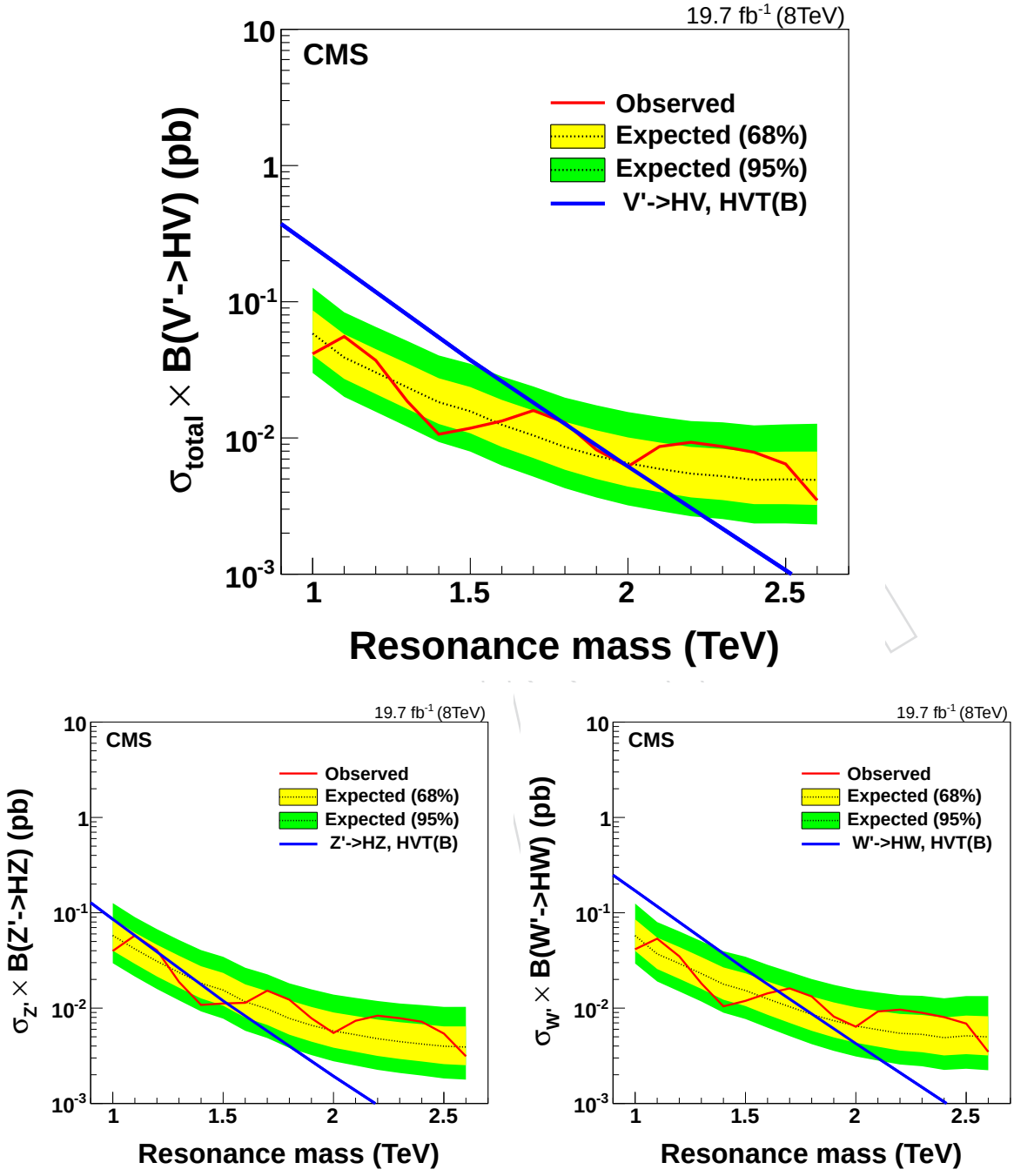


Figure 24: W′ and Z′ combined limit is the top plot, by considering them having the same mass, including all the categories in this analysis. Combined expected and observed limits for $Z' \to HZ$ (bottom left) and $W' \to WH$ (bottom right), including $H \to b\bar{b}$ and $H \to WW^* \to 4q$ channels. Branching ratios of Higgs and V decays are already taken into account. Theory model used here is HVT scenario B, arXiv:1402.4431.

11 $H \to b\bar{b}$ tagger limits (categories Hbb1, Hbb2)

Figure 25 shows the limits for $H \to b\bar{b}, W/Z \to \bar{q}q'$signals passing the $H \to b\bar{b}$ tagger. Limits of combining catogories Hbb1 and Hbb2 are presented. The $H \to b\bar{b}$ and V bosons braching ratios are already taken into account.

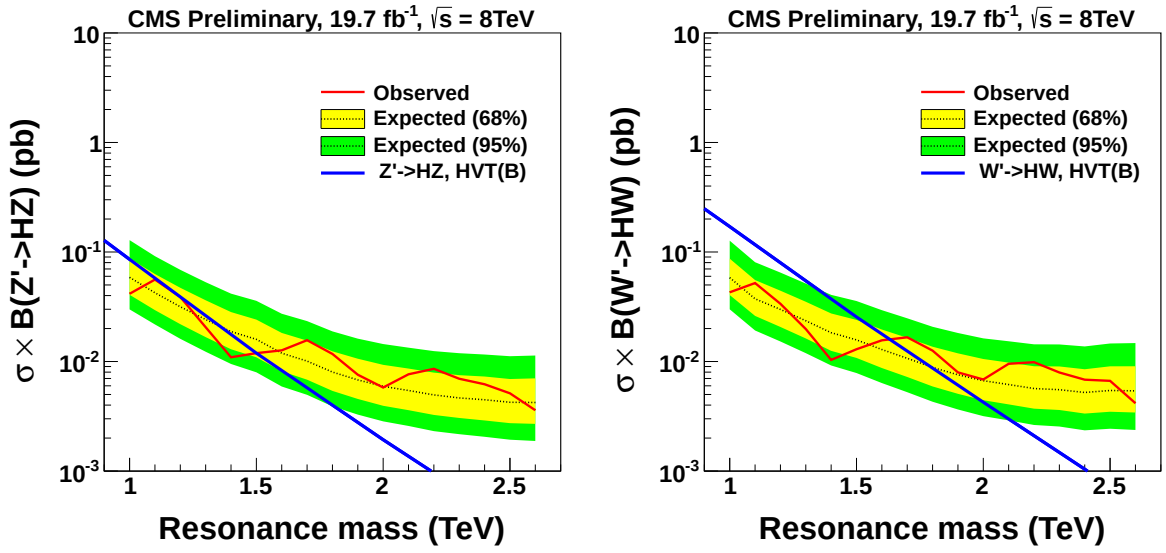


Figure 25: Expected and observed limits for $Z' \to HZ$(left) and $W' \to WZ$(right) search, in $H \to b\bar{b}$ decay mode. Branching ratios of $H \to b\bar{b}$ and V decays are taken into account.Theory model used here is HVT scenario B, arXiv:1402.4431.

12 $H \to WW^* \to 4q$ tagger limits (categories Hww1, Hww2, Hww3)

Figure 26 shows the cobmined limits for $H \to WW^* \to 4q, W/Z \to \bar{q}q'$and $H \to b\bar{b}, W/Z \to \bar{q}q'$signals failing the $H \to b\bar{b}$ tagger but passing the $H \to WW^* \to 4q$ tagger. Limits of combining category Hww1, Hww2 and Hww3 are presented. The $H \to WW^* \to 4q$, $H \to b\bar{b}$ and V bosons braching ratios are already taken into account.

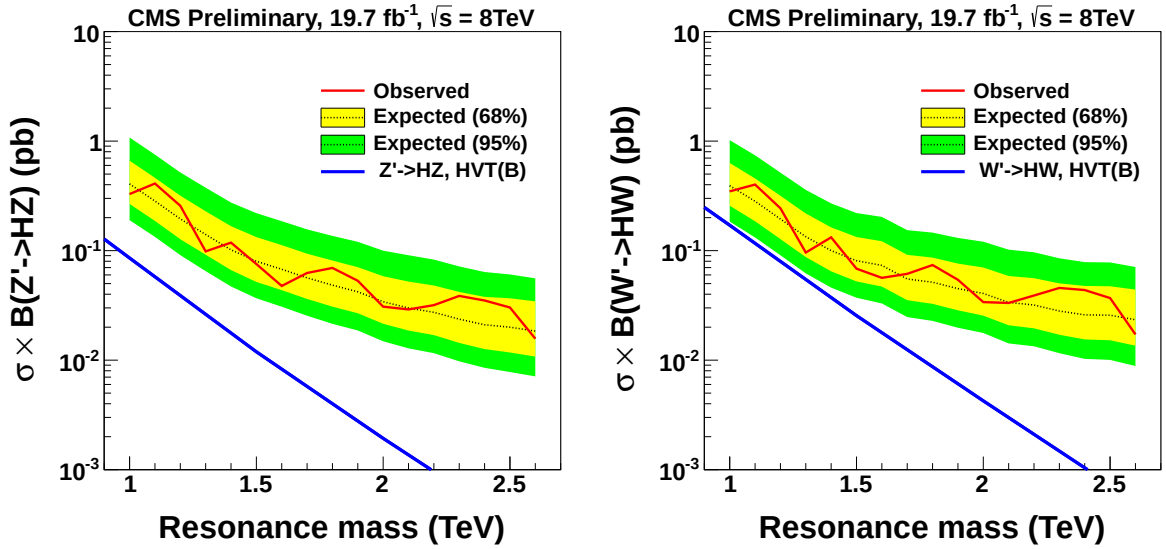


Figure 26: Expected and observed limits for $Z' \to HZ$(left) and $W' \to WZ$(right) search for categories Hww1, Hww2 and Hww3. Branching ratios of $H \to WW^* \to 4q$, $H \to b\bar{b}$ and V decays are taken into account. Theory model used here is HVT scenario B, arXiv:1402.4431.

13 Results and conclusions

Our limits results are summarized in Table 9.

A data sample corresponding to an integrated luminosity of 19.7 fb^{-1}collected in pp collisions at $\sqrt{s} = 8$ TeV with the CMS detector has been used to measure the W/Z− and H-tagged dijet mass spectrum using the two leading jets within the pseudorapidity range $|\eta| < 2.5$ and with pseudorapidity separation $|\Delta\eta| < 1.3$. The QCD background is suppressed using jet substructure tagging techniques and/or b-tagging, which identify boosted bosons decaying into hadrons. In particular, we use the invariant mass of pruned jets and the N-subjettiness ratios τ_{21} and τ_{42}, as well as b tagging applied to the subjets of the Higgs jet, to discriminate against the initially overwhelming QCD background. The remaining QCD background is estimated from a fit to the parameterized shape. We have searched for the signal as a peak on top of the smoothly falling QCD background. Z′ is excluded in resonance mass regions, [1.0, 1.1] and [1.3, 1.5] TeV. While W′ is excluded in resonance mass regions [1.0, 1.6] TeV. In HVT B model, W′ and Z′ are degenerate, so we combine them together. The exclusion limit is set on [1.0, 2.1] TeV.

Table 9: Summary of observed limits on resonance masses at 95% CL and their expected values, assuming a null hypothesis. The analysis is sensitive to resonances heavier than 1 TeV.

Process	Observed excluded mass limit (TeV)	Expected excluded mass limit (TeV)
$V' \to VH$	[1.0, 1.7], [1.9, 2.0]	2.0
$Z' \to HZ(qq)$	[1.0, 1.1], [1.3, 1.5]	1.3
$W' \to HW(qq)$	[1.0, 1.6]	1.7

A Appendices

A.1 Event displays

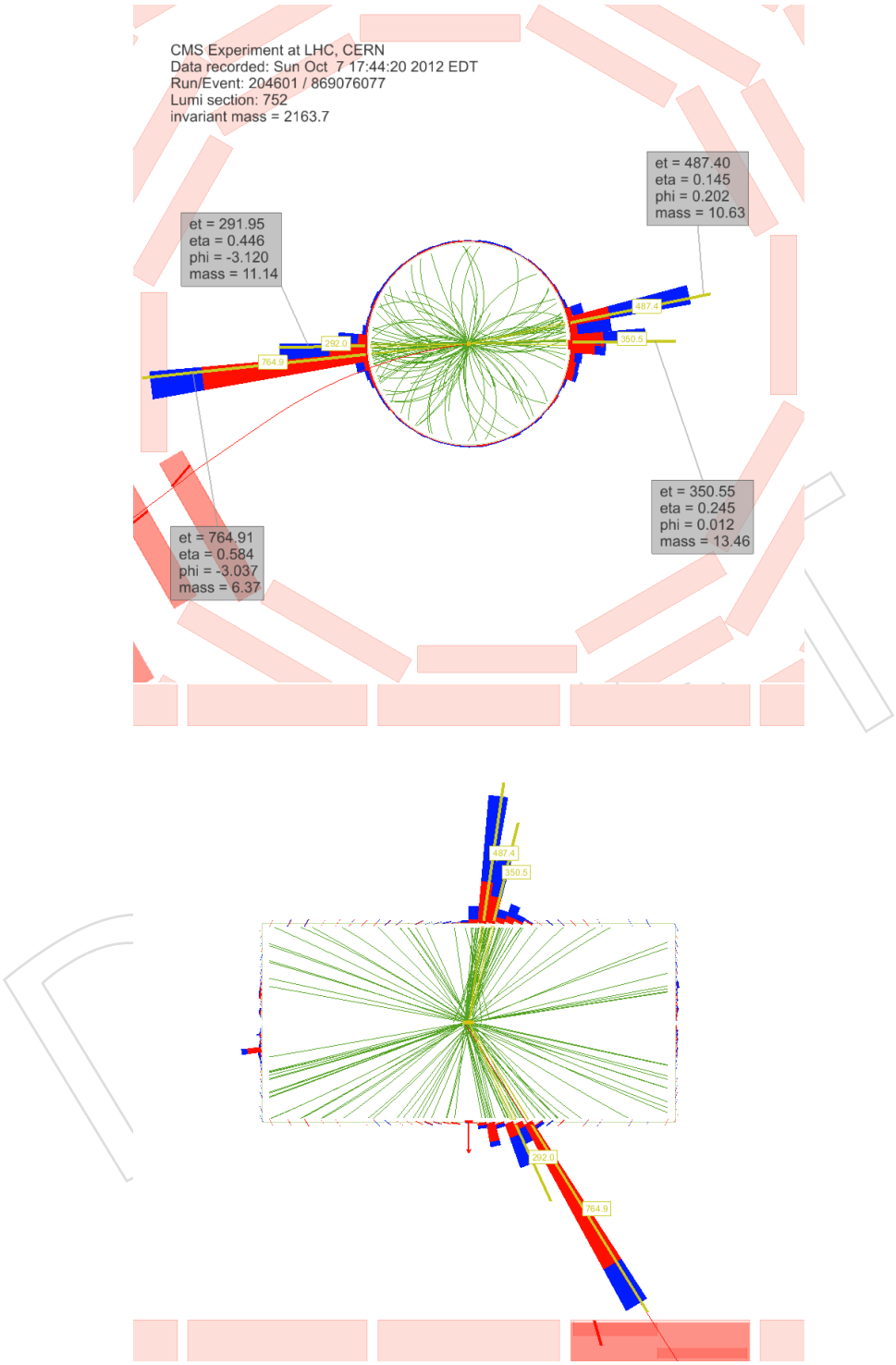


Figure 27: Event display of double W/Z-tagged event with the highest dijet invariant mass of 2.16 TeV. The transverse momenta of the two leading jets are 1.1 TeV and 0.92 TeV. The invariant mass of the two leading pruned CA8 jets is 97.82 GeV and 85.08 GeV.

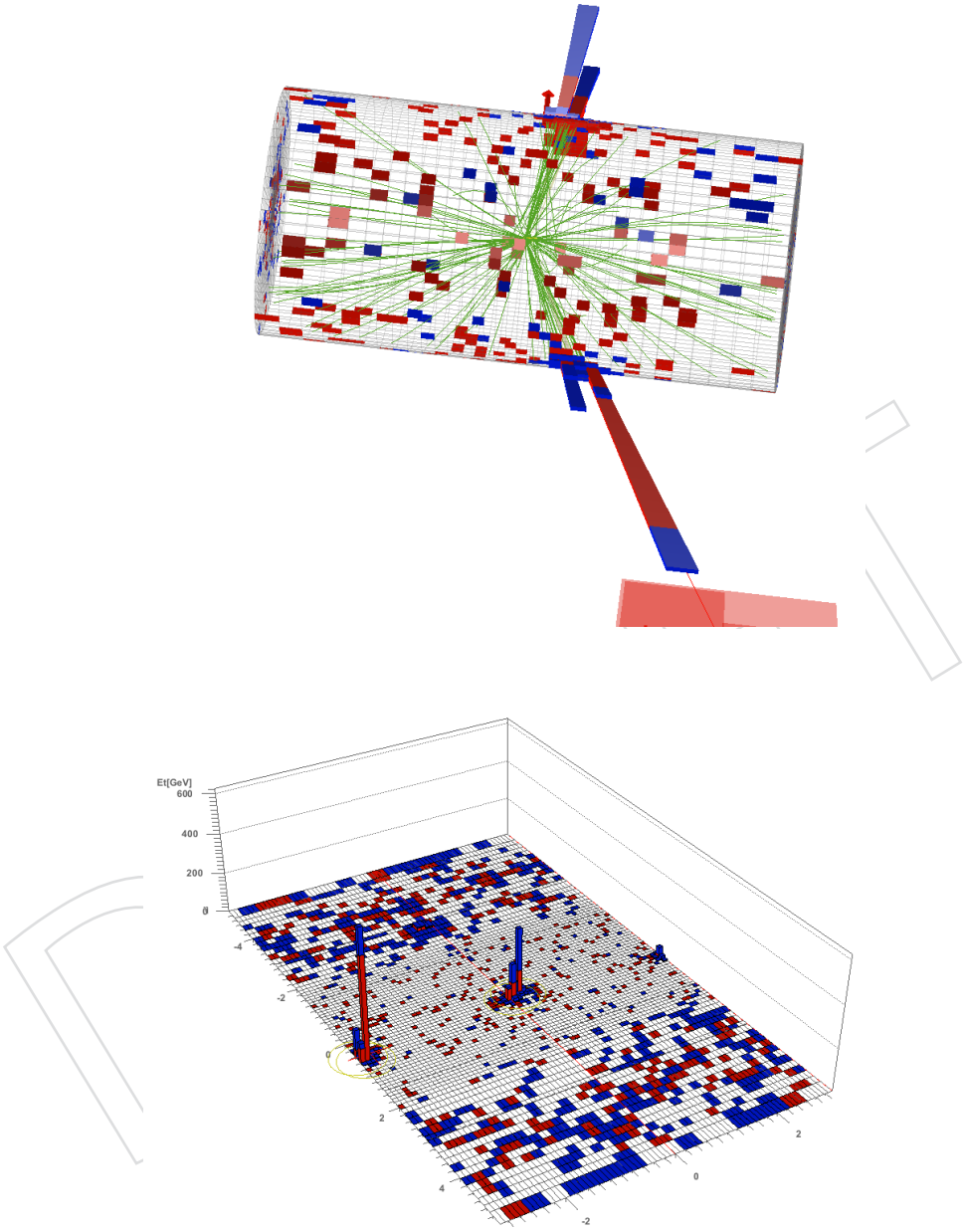


Figure 28: Event display of double W/Z-tagged event with the highest dijet invariant mass of 2.16 TeV. The transverse momenta of the two leading jets are 1.1 TeV and 0.92 TeV. The invariant mass of the two leading pruned CA8 jets is 97.82 GeV and 85.08 GeV.

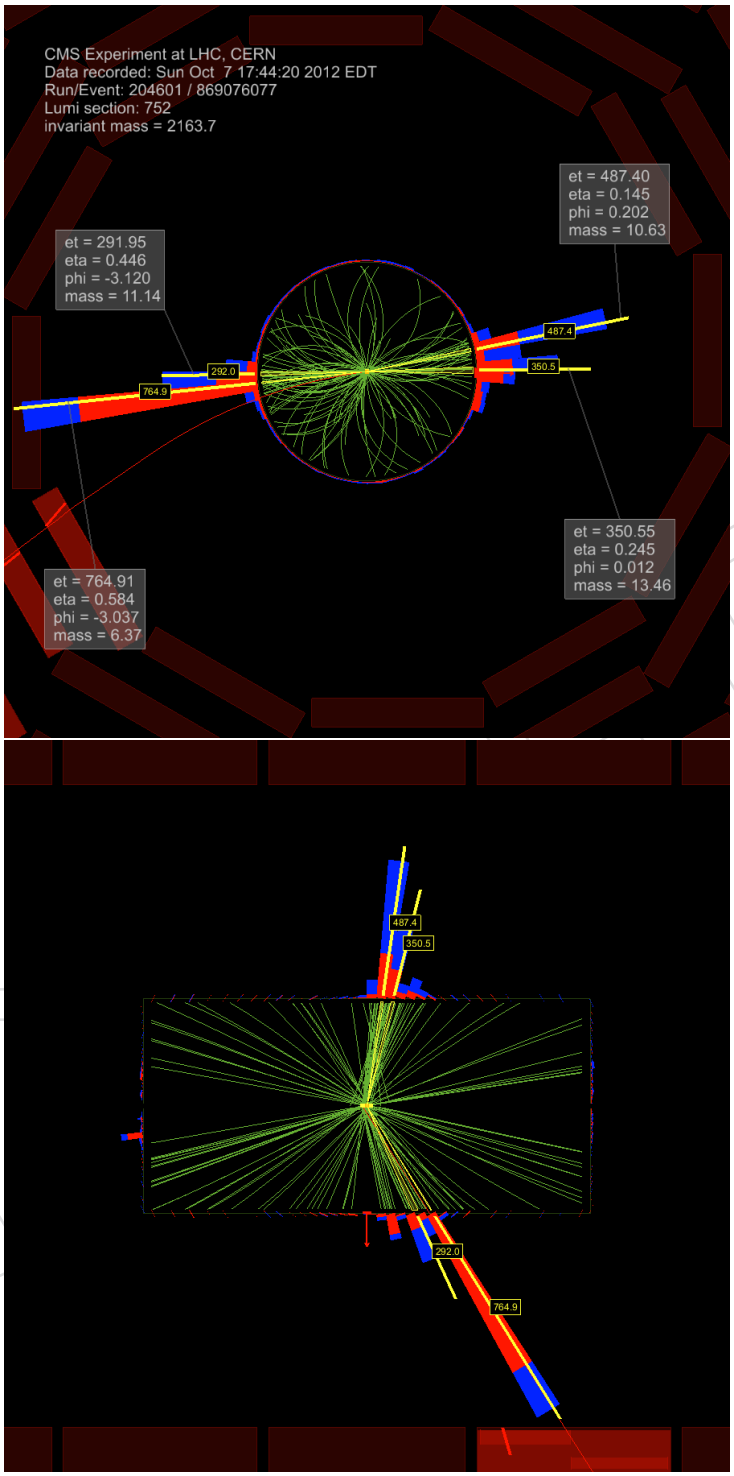


Figure 29: Event display of double W/Z-tagged event with the highest dijet invariant mass of 2.16 TeV. The transverse momenta of the two leading jets are 1.1 TeV and 0.92 TeV. The invariant mass of the two leading pruned CA8 jets is 97.82 GeV and 85.08 GeV.

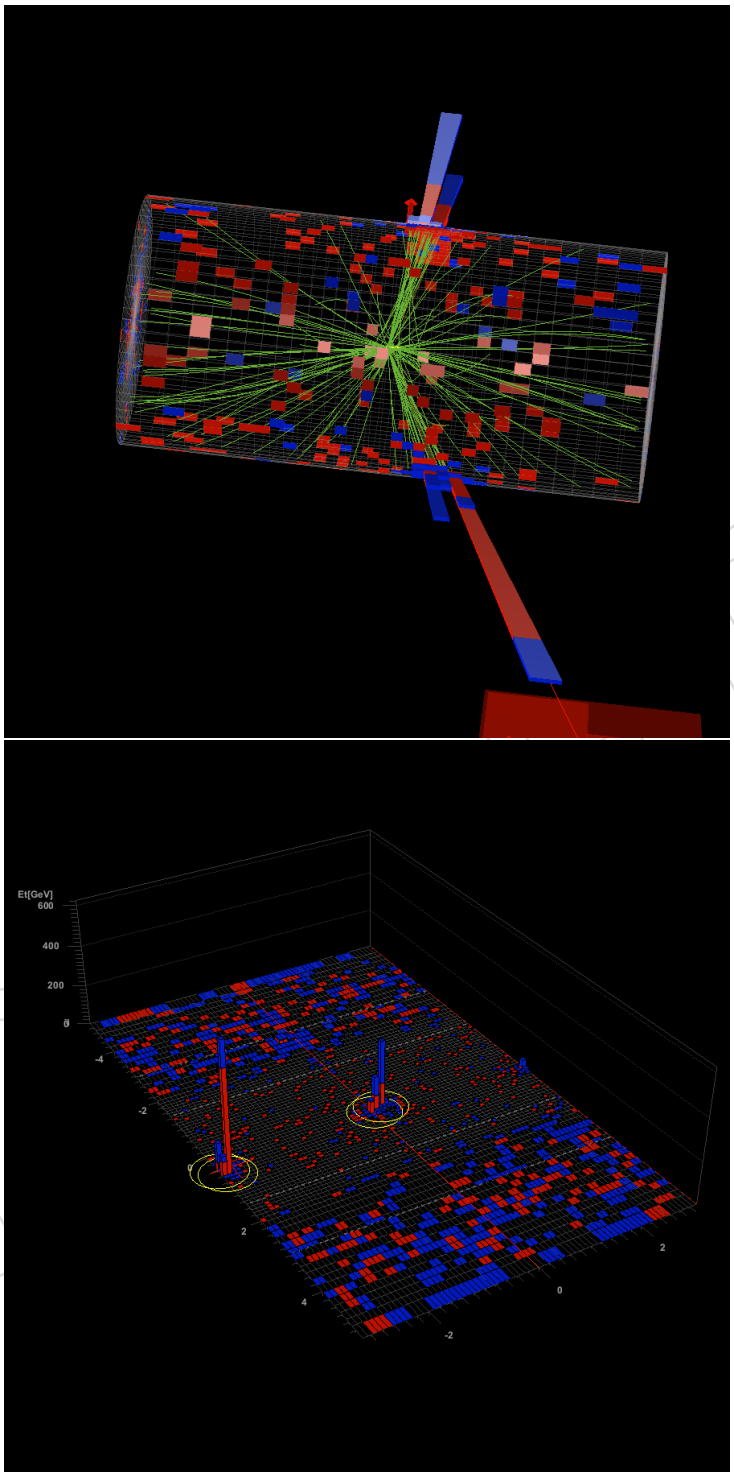


Figure 30: Event display of double W/Z-tagged event with the highest dijet invariant mass of 2.16 TeV. The transverse momenta of the two leading jets are 1.1 TeV and 0.92 TeV. The invariant mass of the two leading pruned CA8 jets is 97.82 GeV and 85.08 GeV.

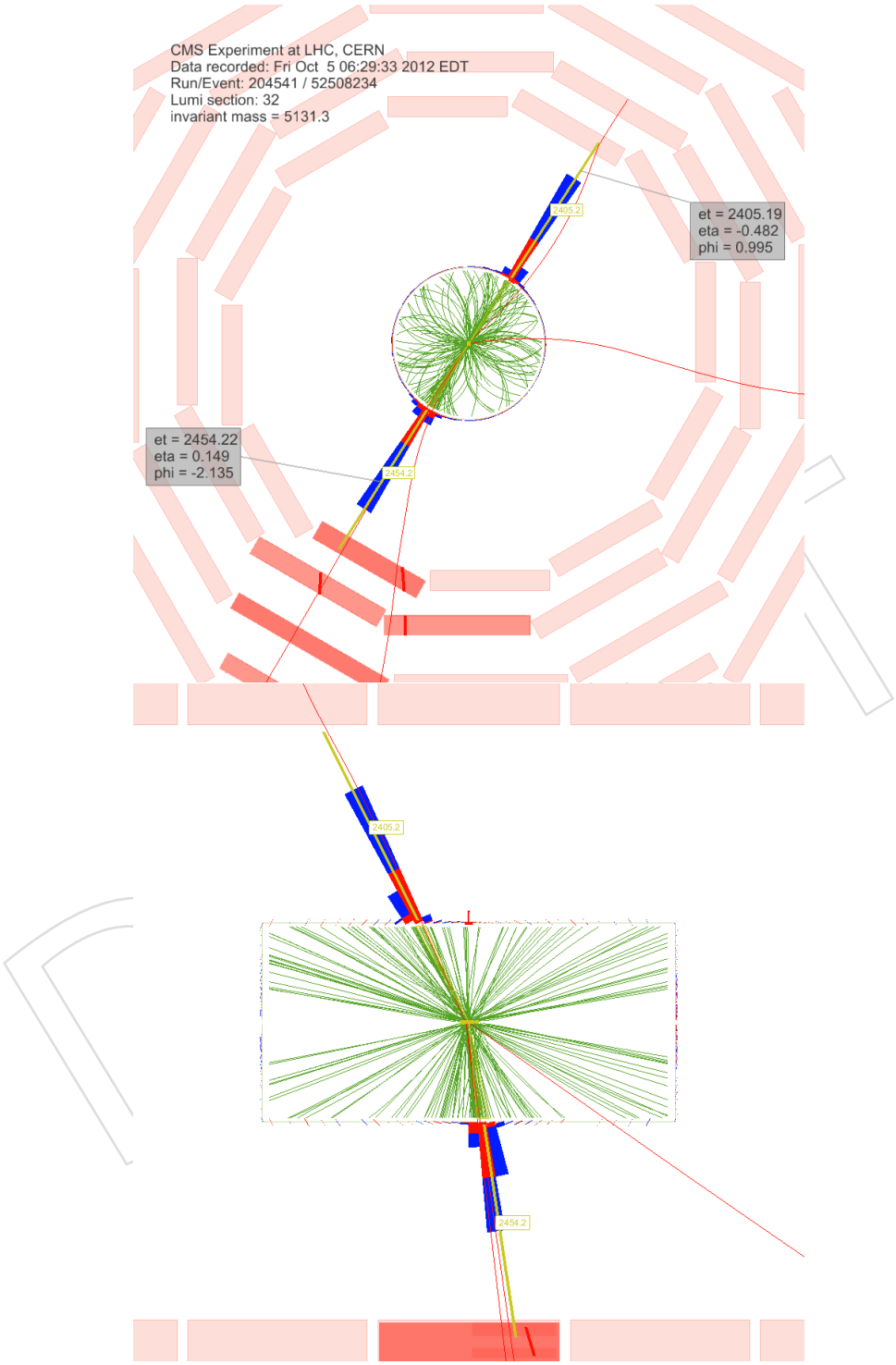


Figure 31: Event display of event with the highest dijet invariant mass of 5.13 TeV. The transverse momenta of the two leading AK5 jets are 2.45 TeV and 2.40 TeV.

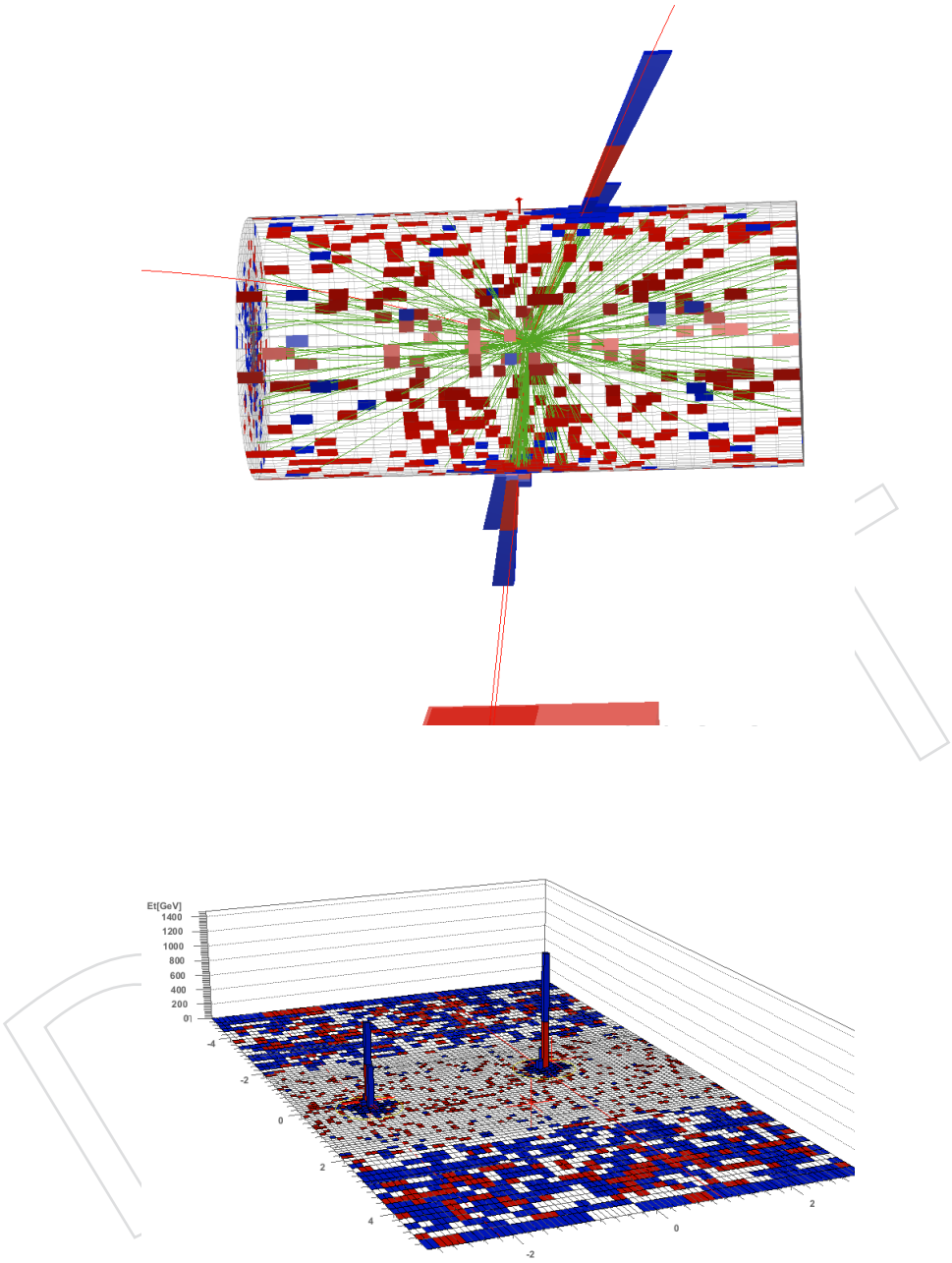


Figure 32: Event display of event with the highest dijet invariant mass of 5.13 TeV. The transverse momenta of the two leading AK5 jets are 2.45 TeV and 2.40 TeV.

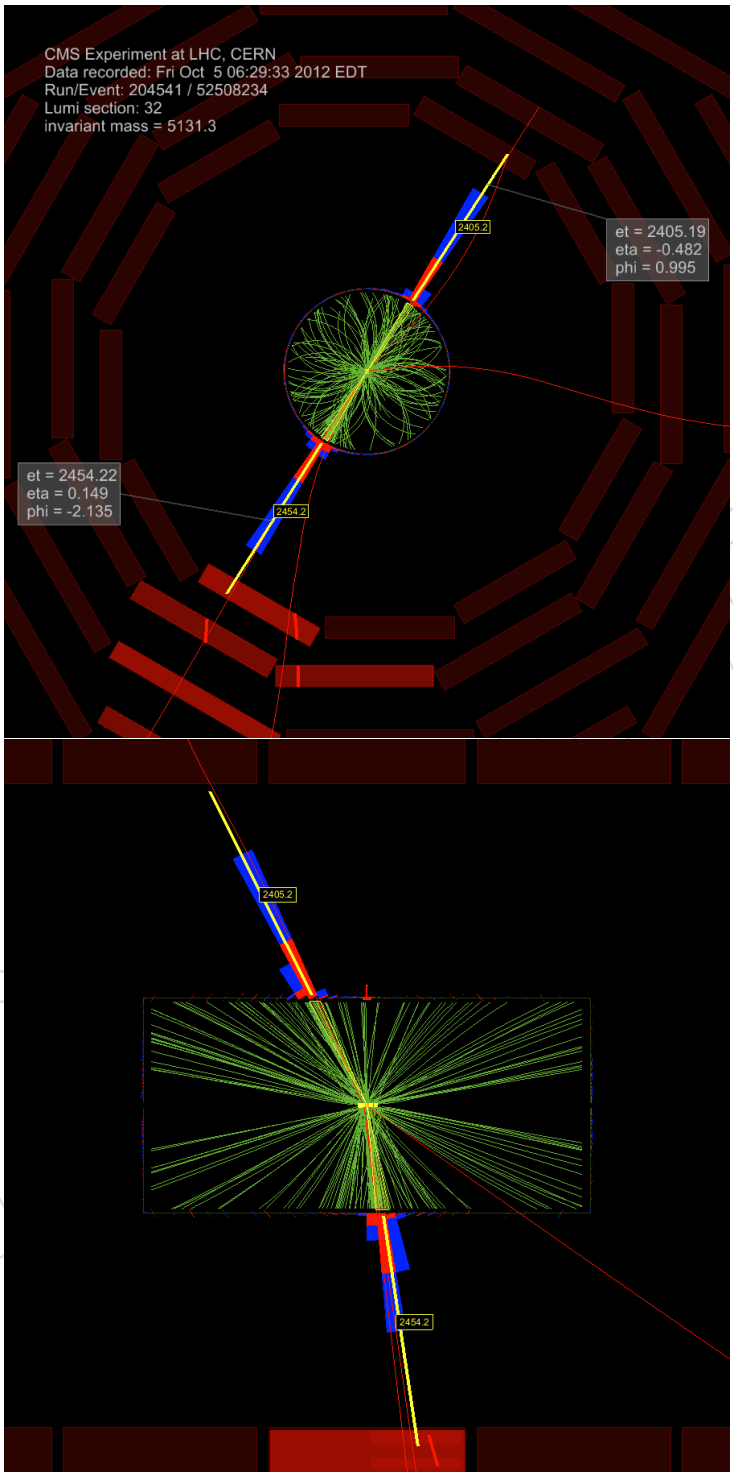


Figure 33: Event display of event with the highest dijet invariant mass of 5.13 TeV. The transverse momenta of the two leading AK5 jets are 2.45 TeV and 2.40 TeV.

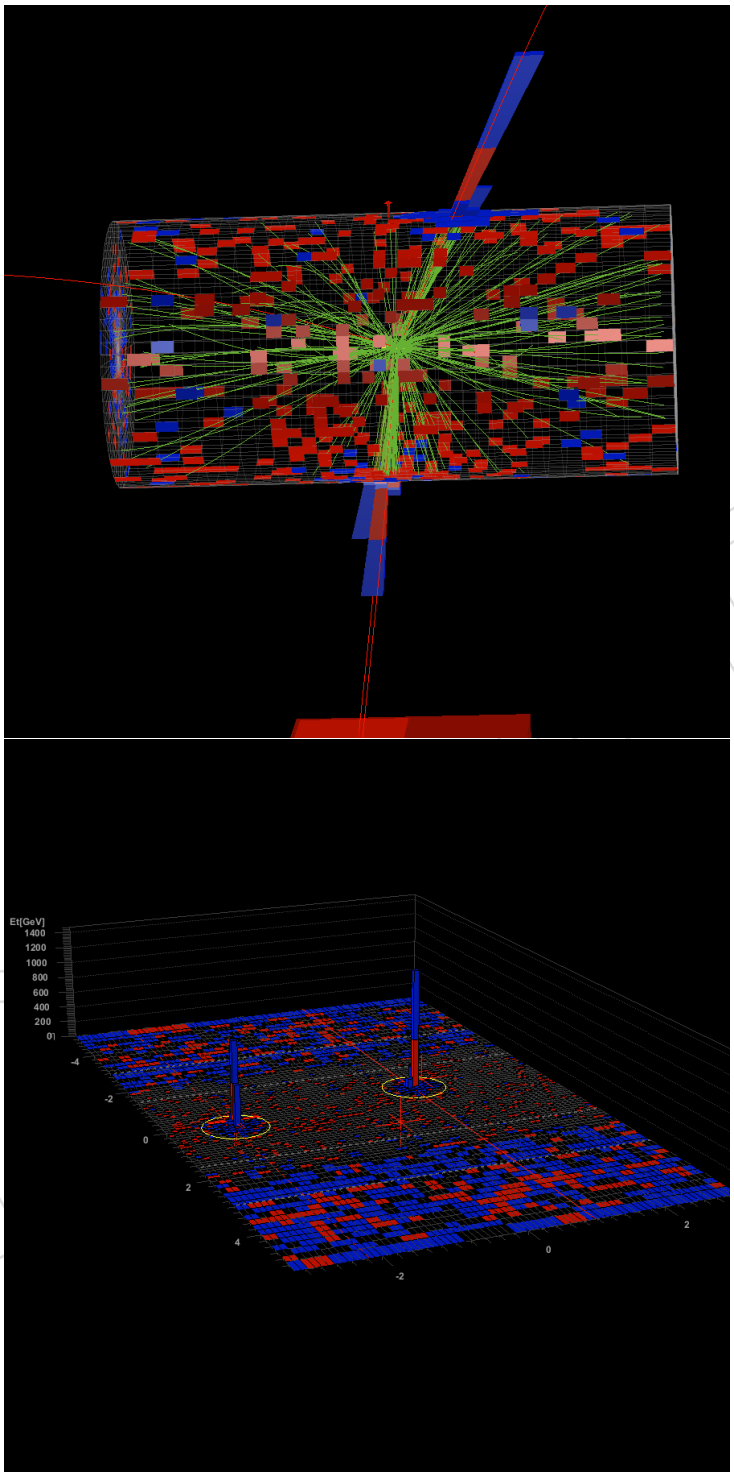


Figure 34: Event display of event with the highest dijet invariant mass of 5.13 TeV. The transverse momenta of the two leading AK5 jets are 2.45 TeV and 2.40 TeV.

B Model parameters and cross sections

We use scenario B of HVT model, in which we set $c_q = c_l = c_3 = 1.024$, $c_H = -0.976$, $g_V = 3$, $c_{VVV} = 0.928$, $c_{VW} = 1$, $c_{VVHH} = -0.024$. Other parameters are listed in Table 10. In this table, we show the width of the generated W′ and Z′, also their production cross sections in different resonance masses.

Table 10: Table of model parameters. CX+ is the cross section for W'+. CX- is the cross section for W'-. CX0 is the cross section for Z'.

M0	g	Width_Z'(GeV)	BRhZ	Width_W'(GeV)	BRWh	CX+(pb)	CX0(pb)	CX-(pb)
890	0.641	32.911	0.536	32.806	0.524	3.58E-01	2.47E-01	1.33E-01
900	0.641	33.069	0.534	32.969	0.522	3.47E-01	2.39E-01	1.28E-01
950	0.642	33.975	0.525	33.895	0.514	2.95E-01	2.00E-01	1.06E-01
1000	0.642	35.017	0.518	34.953	0.508	2.49E-01	1.65E-01	8.70E-02
1050	0.643	36.149	0.512	36.095	0.503	2.08E-01	1.37E-01	7.10E-02
1100	0.643	37.343	0.508	37.298	0.499	1.74E-01	1.12E-01	5.79E-02
1150	0.644	38.584	0.504	38.546	0.496	1.45E-01	9.26E-02	4.72E-02
1200	0.644	39.861	0.501	39.828	0.494	1.21E-01	7.63E-02	3.84E-02
1250	0.645	41.165	0.498	41.136	0.492	1.01E-01	6.29E-02	3.14E-02
1300	0.645	42.492	0.496	42.467	0.490	8.41E-02	5.19E-02	2.56E-02
1350	0.645	43.838	0.494	43.816	0.488	7.02E-02	4.29E-02	2.10E-02
1400	0.645	45.200	0.492	45.181	0.487	5.86E-02	3.55E-02	1.72E-02
1450	0.646	46.575	0.491	46.558	0.486	4.90E-02	2.94E-02	1.41E-02
1500	0.646	47.961	0.489	47.946	0.485	4.10E-02	2.44E-02	1.16E-02
1550	0.646	49.357	0.488	49.343	0.484	3.43E-02	2.02E-02	9.56E-03
1600	0.646	50.762	0.487	50.749	0.483	2.88E-02	1.68E-02	7.89E-03
1650	0.646	52.174	0.486	52.162	0.482	2.41E-02	1.40E-02	6.52E-03
1700	0.646	53.592	0.485	53.582	0.482	2.02E-02	1.17E-02	5.40E-03
1750	0.646	55.017	0.485	55.008	0.481	1.70E-02	9.77E-03	4.47E-03
1800	0.647	56.447	0.484	56.438	0.481	1.42E-02	8.16E-03	3.71E-03
1850	0.647	57.882	0.483	57.874	0.480	1.20E-02	6.82E-03	3.08E-03
1900	0.647	59.320	0.483	59.313	0.480	1.00E-02	5.71E-03	2.56E-03
1950	0.647	60.763	0.482	60.756	0.479	8.44E-03	4.79E-03	2.13E-03
2000	0.647	62.209	0.482	62.203	0.479	7.10E-03	4.01E-03	1.78E-03
2050	0.647	63.659	0.481	63.653	0.479	5.96E-03	3.37E-03	1.48E-03
2100	0.647	65.111	0.481	65.106	0.478	5.01E-03	2.83E-03	1.24E-03
2150	0.647	66.567	0.480	66.562	0.478	4.21E-03	2.37E-03	1.03E-03
2200	0.647	68.024	0.480	68.020	0.478	3.54E-03	2.00E-03	8.62E-04
2250	0.647	69.484	0.480	69.480	0.478	2.97E-03	1.68E-03	7.21E-04
2300	0.647	70.946	0.479	70.942	0.477	2.50E-03	1.41E-03	6.02E-04
2350	0.647	72.410	0.479	72.407	0.477	2.10E-03	1.19E-03	5.03E-04
2400	0.647	73.876	0.479	73.873	0.477	1.76E-03	9.99E-04	4.21E-04
2450	0.647	75.344	0.479	75.340	0.477	1.48E-03	8.41E-04	3.52E-04
2500	0.647	76.813	0.478	76.810	0.477	1.24E-03	7.08E-04	2.95E-04
2550	0.647	78.284	0.478	78.281	0.476	1.04E-03	5.96E-04	2.47E-04
2600	0.647	79.756	0.478	79.753	0.476	8.70E-04	5.01E-04	2.07E-04
2650	0.648	81.229	0.478	81.226	0.476	7.28E-04	4.22E-04	1.73E-04
2700	0.648	82.703	0.477	82.701	0.476	6.09E-04	3.55E-04	1.45E-04
2750	0.648	84.179	0.477	84.177	0.476	5.09E-04	2.99E-04	1.21E-04
2800	0.648	85.656	0.477	85.654	0.476	4.25E-04	2.52E-04	1.01E-04
2850	0.648	87.134	0.477	87.132	0.476	3.55E-04	2.12E-04	8.45E-05
2900	0.648	88.613	0.477	88.611	0.476	2.96E-04	1.78E-04	7.05E-05
2950	0.648	90.092	0.477	90.090	0.476	2.47E-04	1.50E-04	5.89E-05
3000	0.648	91.573	0.477	91.571	0.475	2.05E-04	1.26E-04	4.91E-05

C tau42 scale factor extrapolation

We extrapolate our H $\rightarrow$ WW* $\rightarrow$ 4q tagging scale factor, as in Equation 5, from hadronic W tagging scale factor, as in Equation 6, plus an additional systematic uncertainty ϵ, as in Equation 7.

$$SF_H = \frac{\tau_{42}^{Data}}{\tau_{42}^{MC}} \tag{5}$$

$$SF_W = \frac{\tau_{21}^{Data}}{\tau_{21}^{MC}} \tag{6}$$

$$SF_H = SF_W + \epsilon \tag{7}$$

Derived from equation 7, we have equation 8.

$$SF_H = SF_W \Longleftrightarrow \frac{\tau_{42}^{Data}}{\tau_{42}^{MC}} = \frac{\tau_{21}^{Data}}{\tau_{21}^{MC}} \Longleftrightarrow \frac{\tau_{42}^{Data}}{\tau_{21}^{Data}} = \frac{\tau_{42}^{MC}}{\tau_{21}^{MC}} \tag{8}$$

We validate equation 8 by comparing the $\frac{\tau_{42}}{\tau_{21}}$ in data and PYTHIA and HERWIG QCD MC, as shown in Figure 35, 36 and 37. In this plot MC shows resonably good agreement with data.

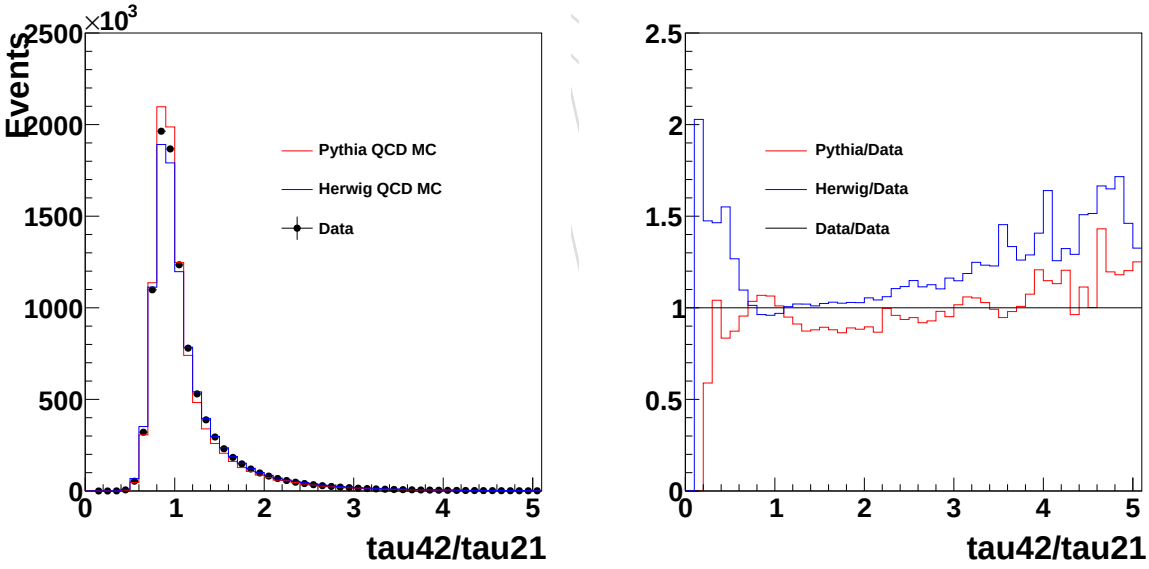


Figure 35: $\frac{\tau_{42}}{\tau_{21}}$ in data (black) compared to PYTHIA QCD MC (red), and HERWIG QCD MC (blue). Left hand plot is logY scale. Plot on the right hand is corresponding ratio plot of left hand.

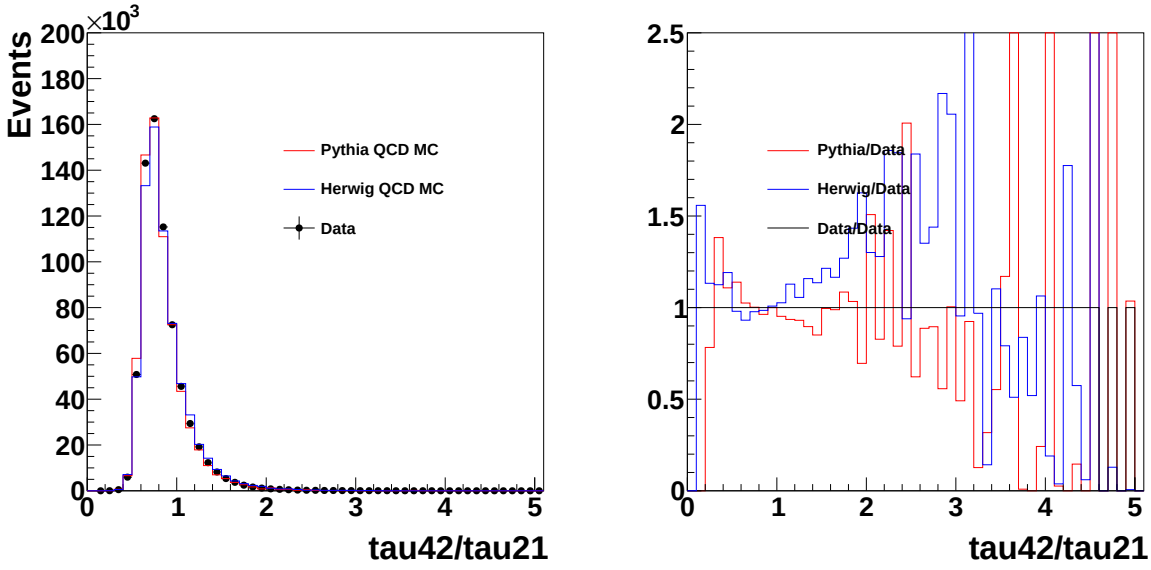


Figure 36: $\frac{\tau_{42}}{\tau_{21}}$ in data (black) compared to PYTHIA QCD MC (red), and HERWIG QCD MC (blue). Region of $\tau_{42} < 0.55$ is shown. Left hand plot is logY scale. Plot on the right hand is corresponding ratio plot of left hand.

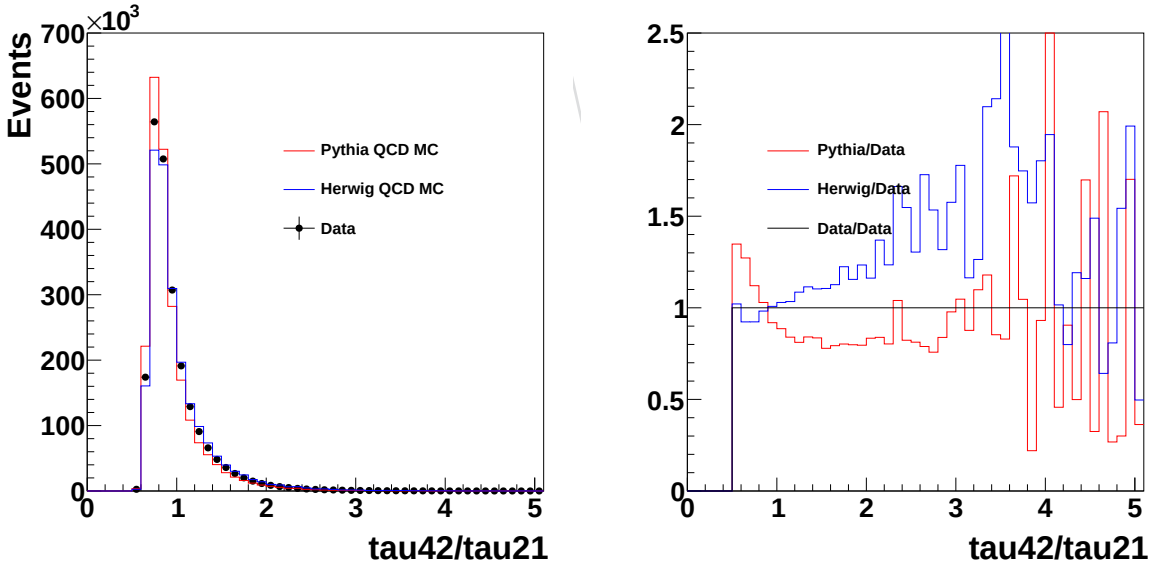


Figure 37: $\frac{\tau_{42}}{\tau_{21}}$ in data (black) compared to PYTHIA QCD MC (red), and HERWIG QCD MC (blue). Region of $0.55 < \tau_{42} < 0.65$ is shown. Left hand plot is logY scale. Plot on the right hand is corresponding ratio plot of left hand.

Since Nsubjettiness τ_N is directly correlated with jet p_T, we further study the $\frac{\tau_{42}}{\tau_{21}}$ with respect to the jet p_T, which are shown in Figure 38.

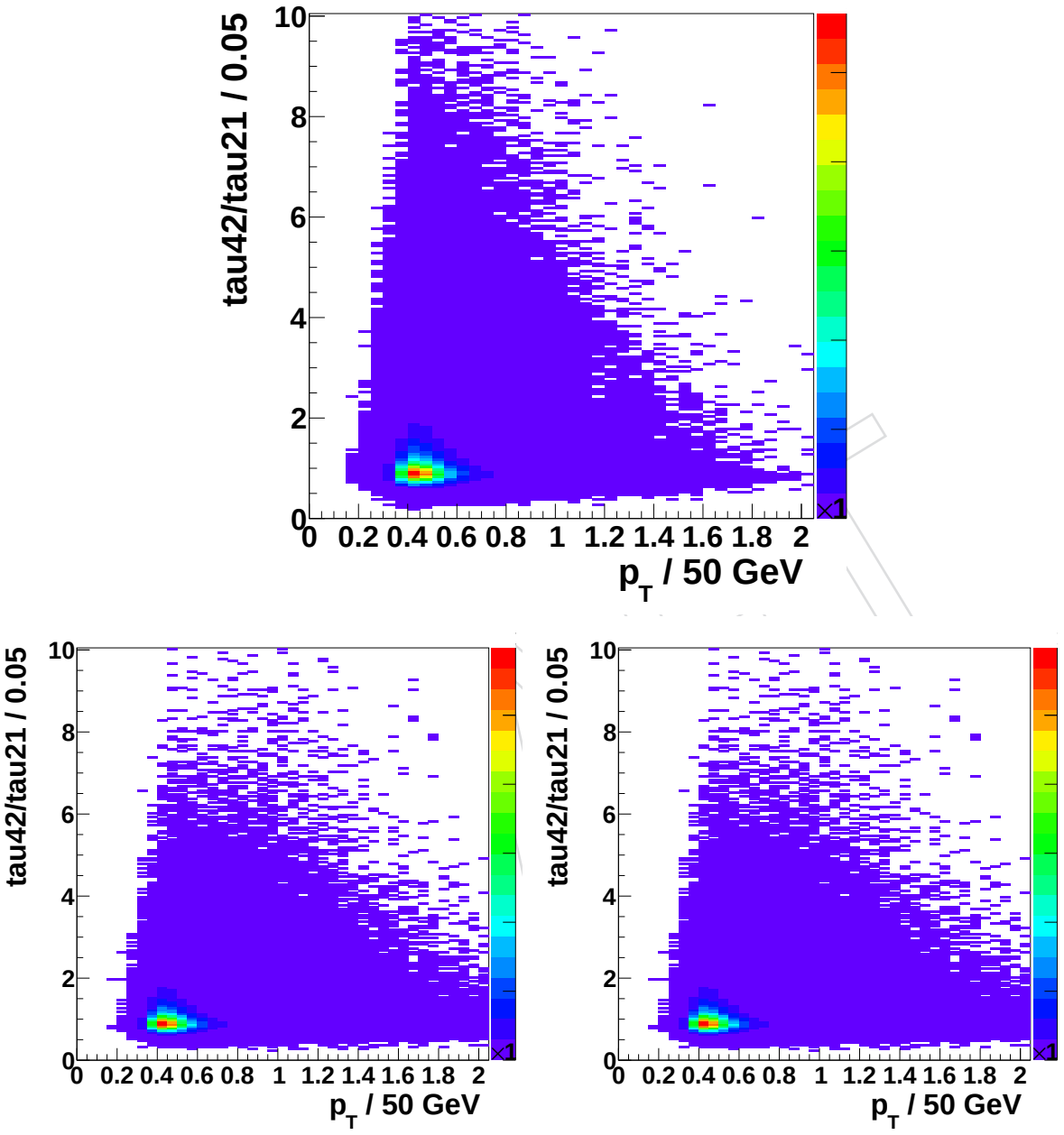


Figure 38: 2D Plot of $\frac{\tau_{42}}{\tau_{21}}$ in Y axis, jet p_T in X axis, in data (top) compared to PYTHIA QCD MC (bottom left), and HERWIG QCD MC (bottom right).

The corresponding profile plot of Figure 38 is shown in Figure 39. From this plot, QCD MC agrees well with data. However, they still have a small discrepancy respect to data, and also between they seleves, especially at high p_T. To compensate this difference, we add an additional uncertainty ϵ, as mentioned in the beginning of this section. ϵ represents the shower and hadronization difference of MC tools, i.e., PYTHIA, HERWIG. We estimate ϵ by taking the largest difference in Higgs-tagging efficiency across various signal resonance masses, which results in 7%.

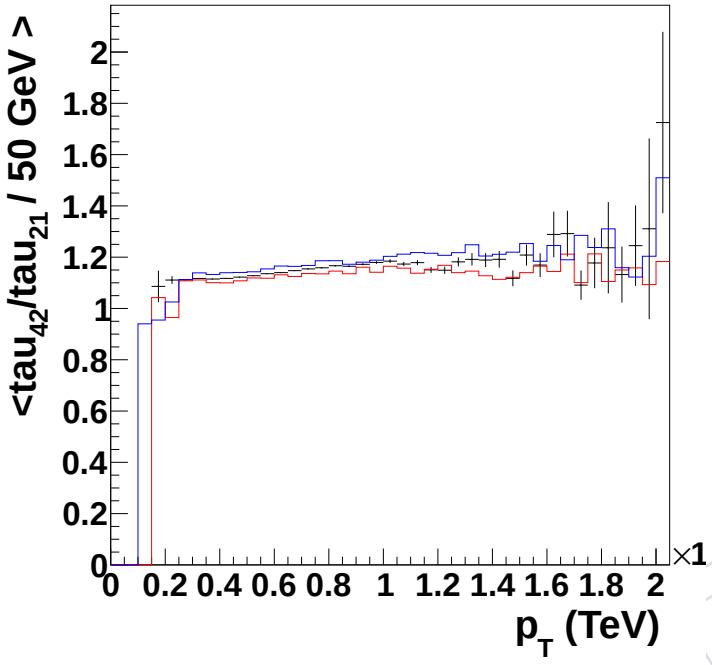


Figure 39: Profile plot, mean of $\frac{\tau_{42}}{\tau_{21}}$ in Y axis, jet p_{T} in X axis, in data (black) compared to PYTHIA QCD MC (red), and HERWIG QCD MC (blue).

So the extrapolated scale factor for H-tagging is $0.86 \pm 7.6\% \pm 7.0\%$, for $\tau_{42} \leq 0.55$, while $1.39 \pm 54\% \pm 7.0\%$ for $0.55 < \tau_{42} < 0.65$.

D CSVL Vs CSVM fat jet b tagging

In H(bb)Z analysis, We compare the csvl fat jet b tagging vs csvm on the limits. Fig.40 is showing the dijet mass spectrum with CSVL fat jet b tagging vs CSVM. And Fig.41 is showing the limits comparison between CSVL fat jet b tagging Vs CSVM. On the table, the limits from CSVL and CSVM are very close, so we show the limits in Table.11.

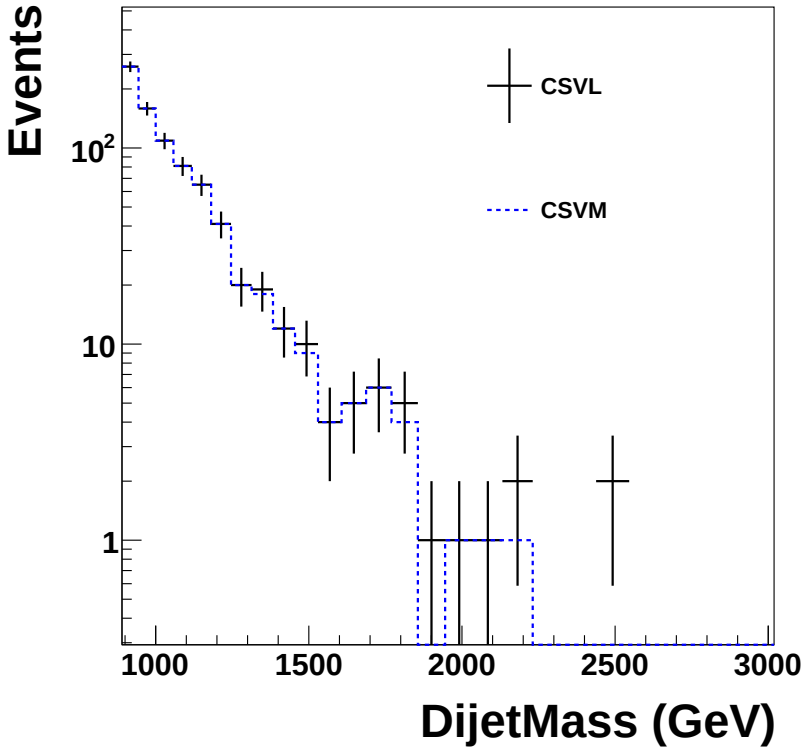


Figure 40: DijetMass distribution for using CSVL Vs. CSVM.

Resonance(TeV)	CSVM(fb)	CSVL(fb)
1	29.9	29.4
1.5	6.86	6.98
2	2.98	2.67

Table 11: Limits on different resonance mass for CSVL VS CSVM.

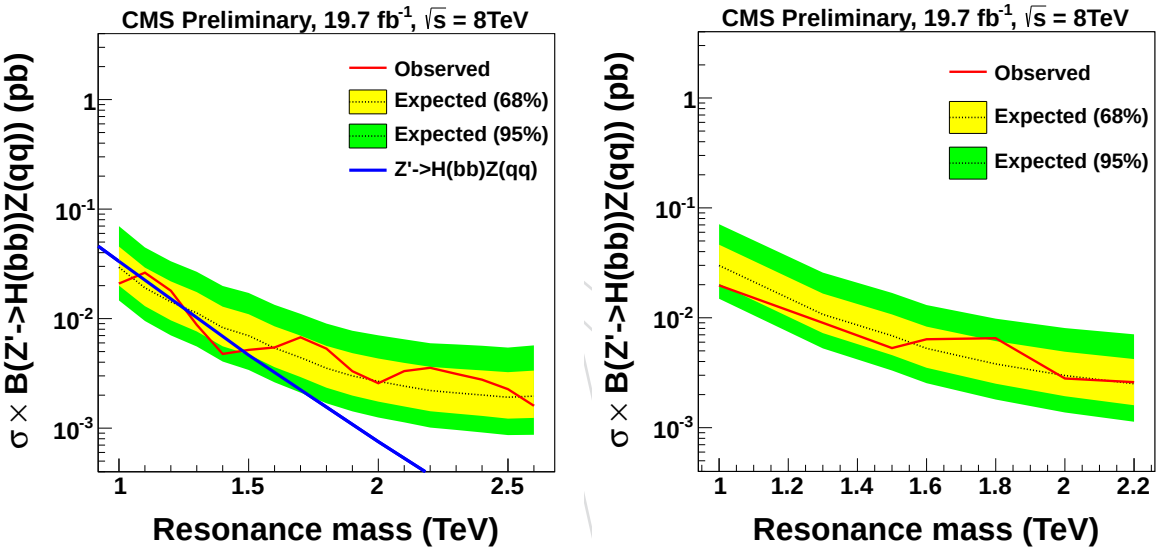


Figure 41: Comparison for limits using CSVL fat jet b tagging(left), and CSVM fat jet b tagging(right).

E tauNM distribution

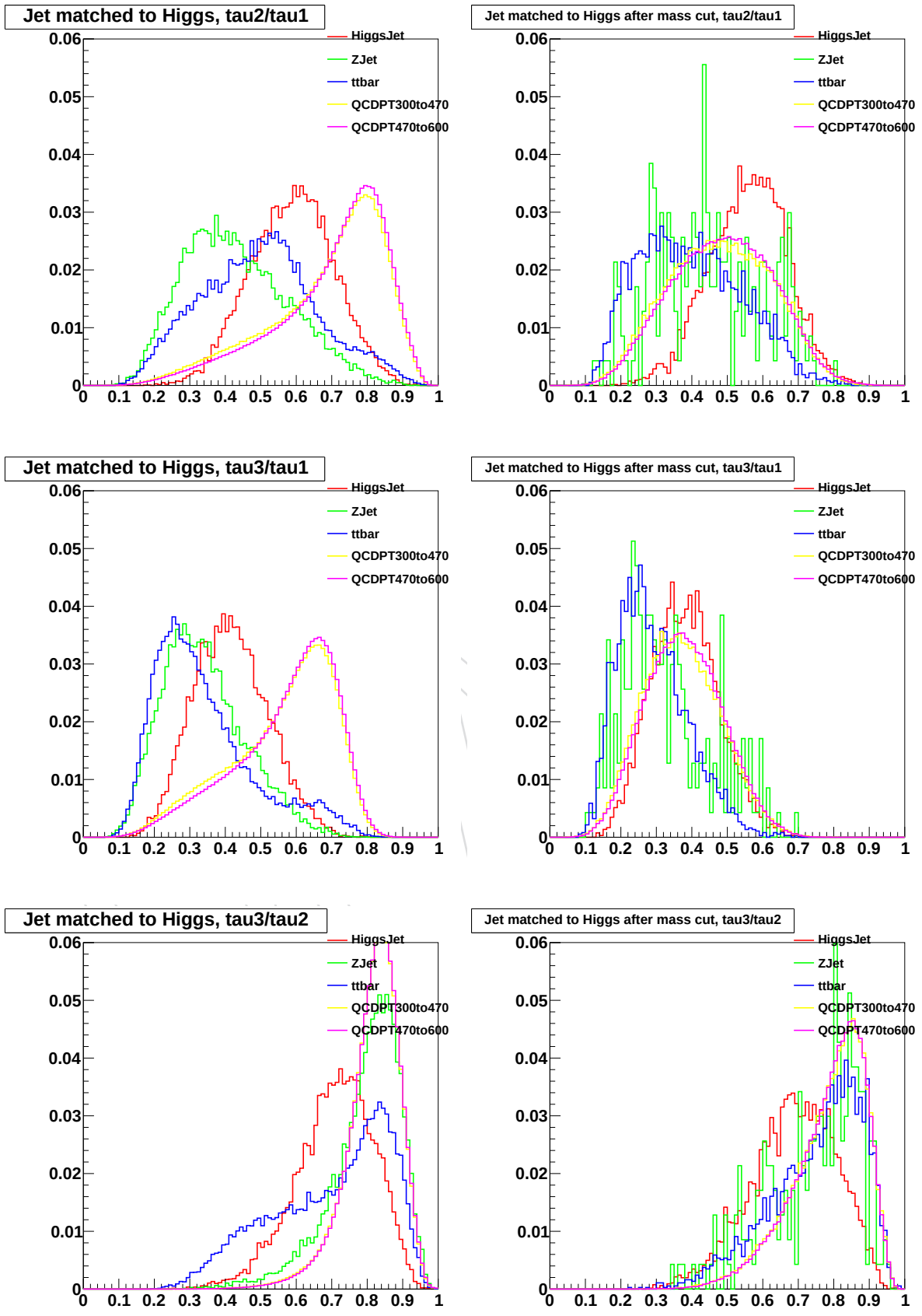


Figure 42: list of tauNM plots between Higgs genJet and Z genJet, hadronic top and QCD. Signal used is 2 TeV Z′.

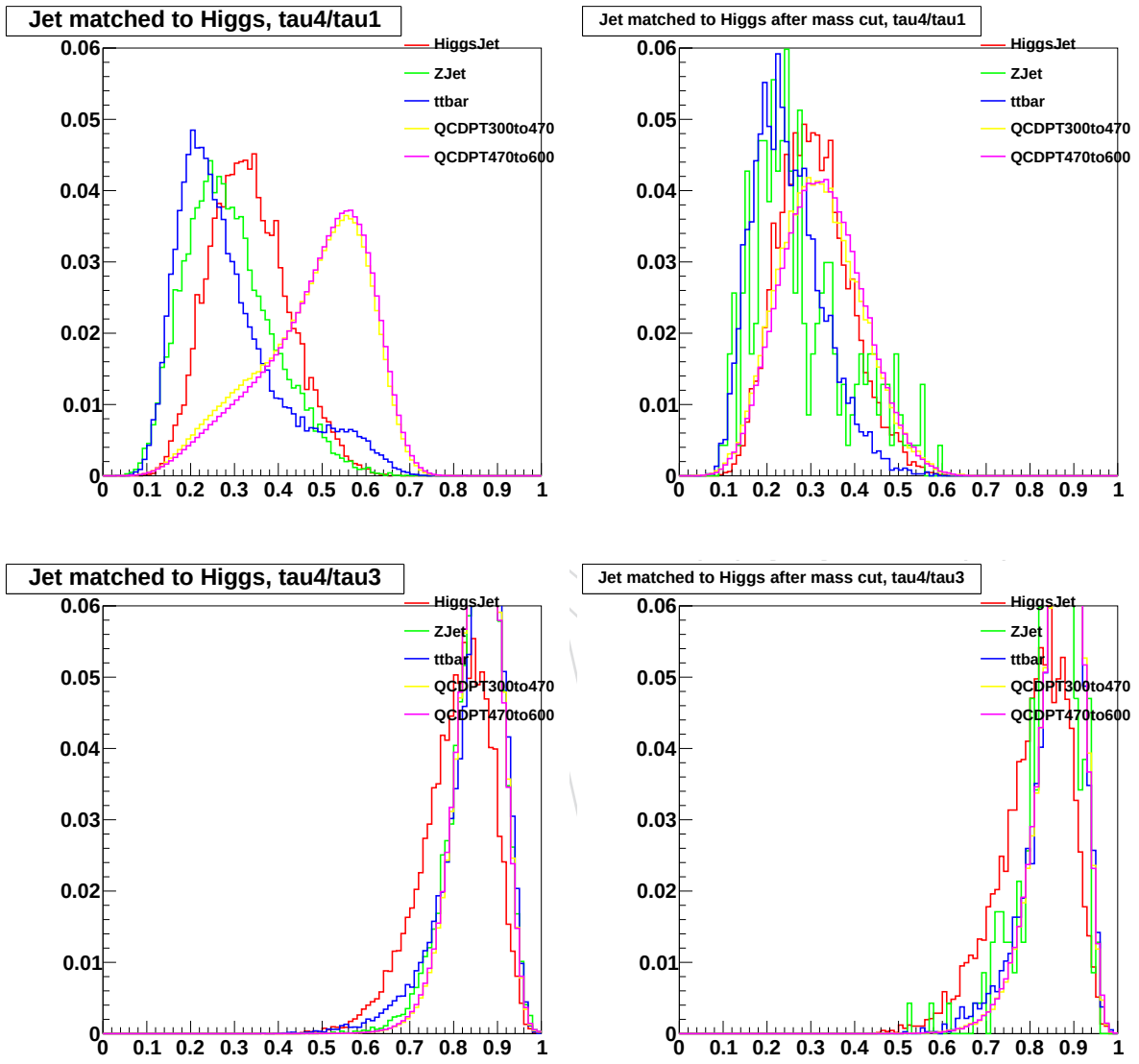


Figure 43: list of tauNM plots between Higgs genJet and Z genJet, hadronic top and QCD. Signal used is 2 TeV Z′.

F Cross-talk in data

The effect of the $H \to b\bar{b}$ tagger veto on the $H \to WW^* \to 4q$ tagged dijet mass distribution background (data) is shown in Figures 44.

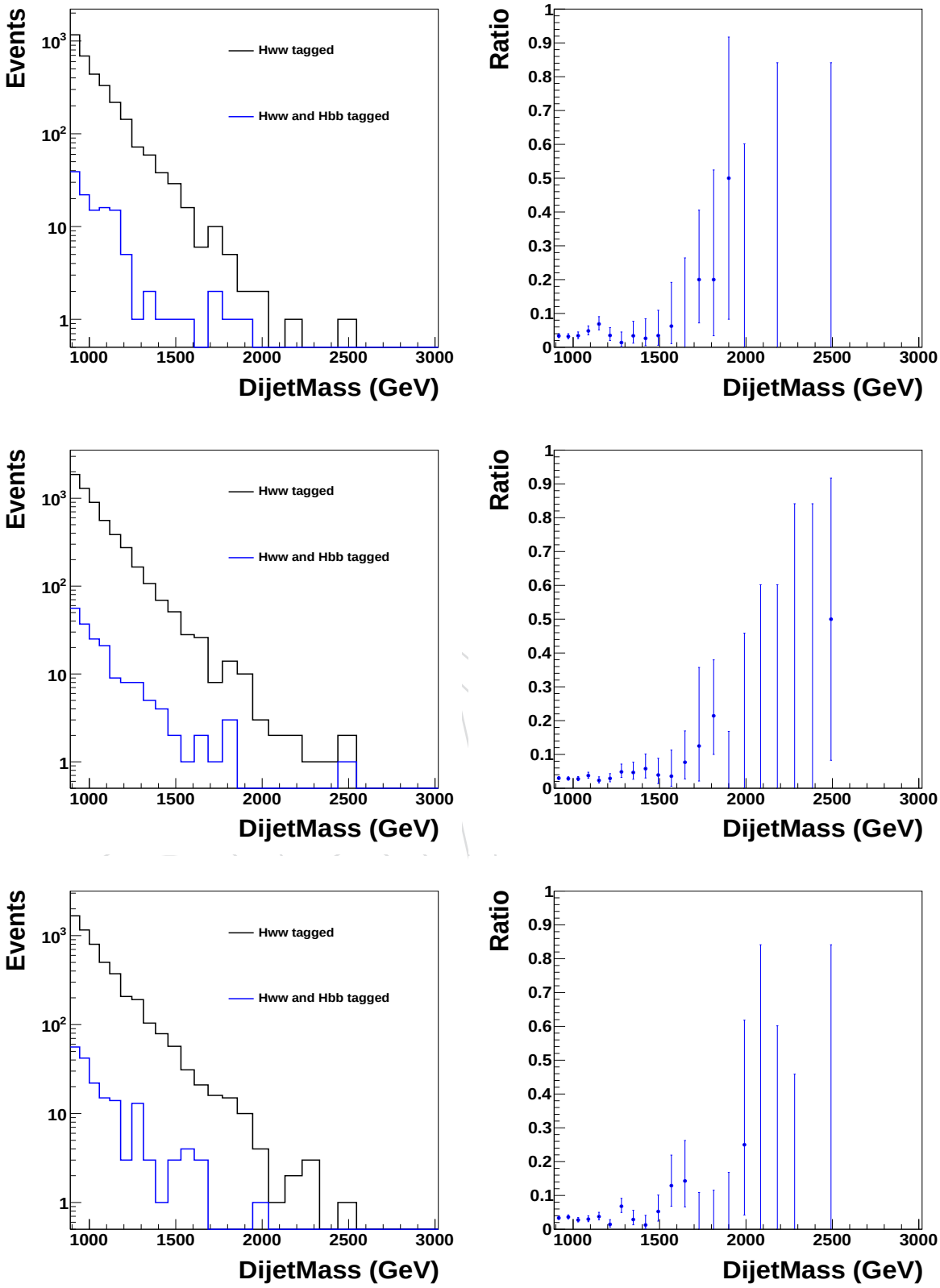


Figure 44: Left column: dijet mass distribution in data, for events passing the $H \to WW^* \to 4q$ tagger (black), and for a subset of these events passing also the $H \to b\bar{b}$ tagger (blue). Right column: the fraction of $H \to WW^* \to 4q$ tagged events also tagged by $H \to b\bar{b}$. Top row: the high purity $H \to WW^* \to 4q$ tagger and high purity V-tagger. Middle row : the low purity $H \to WW^* \to 4q$ tagger, high purity V tagger. bottom row : the high purity $H \to WW^* \to 4q$ tagger, low purity V tagger.

References

[1] CMS Collaboration Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys.Lett.* **B716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[2] ATLAS Collaboration Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys.Lett.* **B716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[3] M. Gouzevitch et al., "Scale-invariant resonance tagging in multijet events and new physics in Higgs pair production", arXiv:1303.6636.

[4] B. Bellazzini, C. Csaki, and J. Serra, "Composite Higgses", (2014). arXiv:1401.2457.

[5] R. Contino, D. Marzocca, D. Pappadopulo, and R. Rattazzi, "On the effect of resonances in composite Higgs phenomenology", *JHEP* **1110** (2011) 081, doi:10.1007/JHEP10(2011)081, arXiv:1109.1570.

[6] D. Marzocca, M. Serone, and J. Shu, "General Composite Higgs Models", *JHEP* **1208** (2012) 013, doi:10.1007/JHEP08(2012)013, arXiv:1205.0770.

[7] T. Han, H. E. Logan, B. McElrath, and L.-T. Wang, "Phenomenology of the little Higgs model", *Phys. Rev. D* **67** (2003) 095004, doi:10.1103/PhysRevD.67.095004, arXiv:0301040.

[8] D. Pappadopulo, A. Thamm, R. Torre, and A. Wulzer, "Heavy Vector Triplets: Bridging Theory and Data", arXiv:1402.4431.

[9] CMS Collaboration, "Search for new physics in final states with a lepton and missing transverse energy in pp collisions at the LHC", *Phys. Rev. D* **87** (2013) 072005, doi:10.1103/PhysRevD.87.072005.

[10] ATLAS Collaboration, "Search for a heavy gauge boson decaying to a charged lepton and a neutrino in 1 fb^{-1} of pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector", *Phys. Lett. B* **705** (2011) 28, doi:10.1016/j.physletb.2011.09.093, arXiv:1108.1316.

[11] CMS Collaboration, "Search for a $\text{W}^{'}$ or ρ_{TC} decaying to WZ in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **109** (2012) 141801, doi:10.1103/PhysRevLett.109.141801.

[12] ATLAS Collaboration, "Search for resonant diboson production in the WW/WZ $\to lvjj$ decay channels with the ATLAS detector at $\sqrt{s}$ = 7 TeV", *Phys. Rev. D* **87** (2013) 112006, doi:10.1103/PhysRevD.87.112006, arXiv:1305.0125.

[13] ATLAS Collaboration, "Search for resonant WZ production in the WZ to l nu l' l' channel in $\sqrt{s}$ = 7 TeV pp collisions with the ATLAS detector", *Phys. Rev. D* **85** (2012) 112012, doi:10.1103/PhysRevD.85.112012, arXiv:1204.1648.

[14] ATLAS Collaboration, "Search for high-mass dilepton resonances in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector", *submitted to Phys. Rev D* (2014) 28, arXiv:1405.4123.

[15] CMS Collaboration, "Search for heavy narrow dilepton resonances in pp collisions at $\sqrt{s}$ = 7 TeV and $\sqrt{s}$ = 8 TeV", *Phys. Lett. B* **720** (2013) 63, doi:10.1016/j.physletb.2013.02.003, arXiv:1212.6175.

[16] CMS Collaboration, "Performance of b tagging at $\sqrt{s}$ = 8 TeV in multijet, $t\bar{t}$ and boosted topology events", *CMS PAS BTV-13-001* (2013).

[17] CMS Collaboration, "Study of Jet Substructure in pp Collisions at 7 TeV in CMS", CMS Physics Analysis Summary CMS-PAS-JME-10-013, (2010).

[18] CMS Collaboration, "Identifying Hadronically Decaying Vector Bosons Merged into a Single Jet", CMS Physics Analysis Summary CMS-PAS-JME-13-006, (2013).

[19] CMS Collaboration, "Search for resonances in the dijet mass spectrum from 7 TeV pp collisions at CMS", *Phys. Lett. B* **704** (2011) 123, doi:10.1016/j.physletb.2011.09.015, arXiv:1107.4771.

[20] CMS Collaboration, "Search for Narrow Resonances using the Dijet Mass Spectrum with 19.6fb-1 of pp Collisions at sqrts=8 TeV", *CMS Physics Analysis Summary* **CMS-PAS-EXO-12-059** (2013).

[21] CMS Collaboration Collaboration, "Search for heavy resonances in the W/Z-tagged dijet mass spectrum in pp collisions at 8 TeV",.

[22] CMS Collaboration, "Search for heavy resonances decaying into two Higgs bosons in hadronic final states", *CMS Physics Analysis Summary* **CMS-PAS-EXO-12-053** (2013).

[23] J. Alwall et al., "MadGraph 5: Going Beyond", *JHEP* (2011) arXiv:1106.0522.

[24] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[25] S. Gieseke et al., "Herwig++ 2.5 Release Note", arXiv:1102.1672.

[26] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[27] R. Field, "Early LHC Underlying Event Data – Findings and Surprises", in *Proceeedings of the Hadron Collider Physics Symposium 2010*. 2010. arXiv:1010.3558.

[28] J. Pumplin et al., "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP* **07** (2002) 012, doi:10.1088/1126-6708/2002/07/012, arXiv:hep-ph/0201195.

[29] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "MRST2001: partons and alpha_S from precise deep inelastic scattering and Tevatron jet data", *Eur. Phys. J. C* **23** (2002) 73, doi:10.1007/s100520100842, arXiv:hep-ph/0110215.

[30] CMS Collaboration, "Particle–Flow Event Reconstruction in CMS and Performance for Jets, Taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, (2009).

[31] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross sections in deep-inelastic scattering", (1998). arXiv:hep-ph/9907280.

[32] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better jet clustering algorithms", *JHEP* **08** (1997) 001, doi:10.1088/1126-6708/1997/08/001, arXiv:hep-ph/9707323.

[33] M. Cacciari and G. P. Salam, "Dispelling the N**3 myth for the k(t) jet-finder", *Phys. Lett. B* **641** (2006) 57, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

[34] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[35] M. Cacciari, G. P. Salam, and G. Soyez, "The Catchment Area of Jets", *JHEP* **04** (2008) 005, doi:10.1088/1126-6708/2008/04/005, arXiv:0802.1188.

[36] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett. B* **659** (2008) 119, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378.

[37] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *J. Instrum.* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[38] CMS Collaboration, "Jet Energy Corrections and Uncertainties Detector Performance Plots for 2012", *CMS Detector Performance Plots* **DP-2012/012** (2012).

[39] M. Carena, C. Grojean, M. Kado, and V. Sharma, "Status of Higgs Boson Physics", *PDG* (2013).

[40] J. Thaler and K. Van Tilburg, "Identifying Boosted Objects with N-subjettiness", *JHEP* **1103** (2011) 015, doi:10.1007/JHEP03(2011)015, arXiv:1011.2268.

[41] J. Thaler and K. Van Tilburg, "Maximizing Boosted Top Identification by Minimizing N-subjettiness", *JHEP* **1202** (2012) 093, doi:10.1007/JHEP02(2012)093, arXiv:1108.2701.

[42] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, "N-Jettiness: An Inclusive Event Shape to Veto Jets", *Phys.Rev.Lett.* **105** (2010) 092002, doi:10.1103/PhysRevLett.105.092002, arXiv:1004.2489.

[43] CMS Collaboration, "A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging", CMS Physics Analysis Summary CMS-PAS-JME-09-001, (2009).

[44] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Techniques for improved heavy particle searches with jet substructure", *Phys. Rev. D* **80** (2009) 051501, doi:10.1103/PhysRevD.80.051501, arXiv:0903.5081.

[45] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches", *Phys. Rev. D* **81** (2010) 094023, doi:10.1103/PhysRevD.81.094023, arXiv:0912.0033.

[46] CMS Collaboration, "Search for a BSM resonance decaying into W and Higgs bosons in lvbb final state", *CMS Physics Analysis Summary* **CMS-PAS-EXO-14-010** (2014).

[47] M. D.E.Kaplan, K.Rehermann and B.Tweedie, "Top tagging: A method for identifying boosted hadronically decaying top quarks", *Phys. Rev. Lett* **101** (2008) 142001.

[48] R. G. Lomax and D. L. Hahs-Vaughn, "Statistical Concepts: A Second Course". p.10, Routledge Academic, 2007.

[49] CMS Collaboration, "Status of the 8 TeV Jet Energy Corrections and Uncertainties based on 11/fb of data in CMS", *CMS Detector Performance Plots* **DP-2013/011** (2013).

[50] CMS Collaboration, "CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update", CMS Physics Analysis Summary CMS-PAS-LUM-13-001, (2013).

[51] G. Cowan, K. Cranmer, E. Gross, O. Vitells, "Asymptotic formulae for likelihood based tests of new physics", *Eur. Phys. J.* **C71** (2011) 1554, arXiv:physics/1007.1727v2.

[52] A. L. Read, "Presentation of search results: The CL(s) technique", *J. Phys.* **G28** (2002) 2693–2704, doi:10.1088/0954-3899/28/10/313.

[53] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2.