Search for diboson resonances at CMS

identifying highly energetic boson decays and discriminating new physics signals from the standard model background

Clemens Lange (CERN)
DESY Physics Seminar
Hamburg/Zeuthen
31st May/1st June 2016
About diboson resonances

- heavy resonances?
- boson jets?
- background estimation?

wasn’t there this diphoton resonance?

[https://ideas.lego.com/projects/94885]
We have a problem

> how can we explain the big difference between the **weak force** and **gravity**?

\[ \mu^2 = \lambda v^2 = \frac{\lambda}{g^2} 4 M_W^2 \sim 10^4 \text{ GeV}^2 \ll M_{Pl} \sim 10^{38} \text{ GeV}^2 \]

> no **symmetry** in the standard model (SM) protects the Higgs mass

> \[ \mu |H^2| \] always a singlet under phase transformations

> „natural“ explanation would be that SM is replaced by another theory at the TeV scale: \[ \mu^2 \sim (\text{heavier scale})^2 \rightarrow \text{new particles} \]

> these theories could be:

> **SUSY**: protecting the Higgs mass by a symmetry

> **Composite Higgs**: the Higgs is not elementary

> **Large/warped extra dimensions**: gravity is strong at electroweak scale
the Higgs could be non-fundamental

instead: **bound state** of a new strong interaction

e.g. size of $10^{-18}$ m $\sim$ *Fermi scale* (100 GeV)

- light Higgs like a pion from a new sector

solves hierarchy problem, and brings along **new heavy particles**/states

heavy partners of SM particles decay to lighter ones ($W$, $Z$, $H$, top, …)
Large extra dimensions?

> another attempt to solve the hierarchy problem

> SM fields are confined to four-dimensional „membrane“, gravity propagates in additional dimensions

> effectively, change power law of gravity from $1/r^2$ to $1/r^2+N$, where $N =$ number of extra dimensions

> this only applies to particles with $r \ll N$ - smaller things have more possibilities to move

> „large“, because of size 1 mm to ~1/TeV

> proposed by Arkani-Hamed, Dimopoulos, and Dvali (ADD)
Warped extra dimensions?

> often referred to as **Randall-Sundrum** (RS) models

> warping causes energy scale at one end of the extra dimension to be much larger than at the other end

> SM models reside on **TeV-brane** (in RS1 models)

> **bulk graviton** models allow SM particles into 5D-bulk

> overlap of 5-D profiles at TeV-brane (and Higgs) determine particle masses

> additionally, if distance between two branes is not fixed, additional fluctuations can occur
How do we observe these models at the LHC?

- should be able to observe excitations/resonances/fluctuations

- **composite Higgs**: electroweak composite vector resonances
  - mostly spin-1 (W', Z')
  - decay to pairs of W, Z, H

- **Randall-Sundrum**: Kaluza-Klein excitations of gravitons + radion fluctuations

- **gravitons** (spin-2):
  - RS1: decay predominantly to leptons
  - bulk: decay to pairs of W, Z
  - ADD: broader excess from many narrow-spaced resonances

- **radions** (spin-0):
  - only used for signal modelling here

- focus here on **narrow resonances** (width < detector resolution), mass ≥ 600 GeV
What about the diphoton resonance?

> neutral resonance could be **graviton** or **radion** as in diboson searches

> resonance cannot directly couple to photons $\rightarrow$ loop of **charged particles** (e.g. W, top, ?) **in decay** (and production?)

> there must be **more than just a di-photon resonance**

> searches presented in this talk **constrain** what **physics models** this potential resonance could be
Reconstructing heavy resonances

- Bosons will be very energetic → **collimated decay products**
- Need to develop dedicated reconstruction methods
- Hadronic decays of bosons:
  - "*boson-tagging*"
  - Exploiting substructure of jets
- Leptonic decays:
  - Special **isolation** for dileptonic decays
  - Dedicated reconstruction algorithms for high-\(p_T\) leptons
  - New tau-identification algorithms

Focus here on Run-2 developments and analyses
Hadronic boson identification

- at CMS use anti-\( k_T \) jet algorithm with \( R = 0.4 \)
- already for resonances of 1 TeV a significant fraction of cases where the boson decay is contained in a single jet
- increase jet size to \( R = 0.8 \) to contain full decay within „fat“ jet

\[ \Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} \]

back of the envelope calculation:

for a resonance of mass 1 TeV the bosons from the decay will have \( p_T \approx 0.4 \) TeV \( \Rightarrow \Delta R \approx 0.4 \)
tracking detectors and calorimeters contained in magnetic field

particle flow algorithm makes use of sub-detectors with best resolution (both spatial and energy)

actual „particles“ enter jet clustering
Jet pruning

> we know the **masses** of $W$, $Z$ and Higgs very well ➔ can use them as **constraints**

> however, large number of particles in jet ➔ rather bad resolution

> jet pruning (generally grooming) removes **soft and large angle radiation**

> **strategy:**

  ▪ reclustering jet using Cambridge-Aachen (CA) jet algorithm

?
Reminder: jet clustering algorithms

- $k_T$-algorithms: **sequential clustering**
- examine four-vector inputs pairwise and construct jets hierarchically
- **anti-$k_T$**: preferentially merge constituents with high $p_T$ with respect to their nearest neighbours first
- **Cambridge-Aachen**: no $p_T$-weighting, merge based on **spatial separation** only $\Rightarrow$ undoing clustering yields **subjets**
Jet pruning

> we know the masses of W, Z and Higgs very well → can use them as constraints

> however, large number of particles in jet → rather bad resolution

> jet pruning (generally grooming) removes soft and large angle radiation

> strategy:

  ▪ recluster jet using Cambridge-Aachen (CA) jet algorithm
  ▪ „soft“: \( \min(p_T^i, p_T^j)/\tilde{p}_T < 1 \)
  ▪ „large angle“: \( \Delta R_{ij} > \frac{m_{\text{orig}}}{\tilde{p}_T} \), \( \text{orig} = \text{unpruned CA jet} \)

> cut on mass window (~±10 GeV)
N-subjettiness

- for boson-tagging: want to quantify how 2-subjetty a jet is
- to what extent is energy flow aligned along 2 momentum directions (N=2)?

\[
\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}, \Delta R_{2,k}, \ldots, \Delta R_{N,k})
\]

normalisation sum over particles minimise distance to candidate subjets

low values of \(\tau_N\) \(\rightarrow\) compatibility with the hypothesis of N axes

for each calorimeter cell is proportional to the logarithm of the particle energies in the cell. The open circles indicate the two subjet directions. The open square indicates the total jet direction and the dashed line gives the approximate boundary of the jet. The jets are clustered with the anti-k algorithms.

W-jet

QCD-jet

Less obvious is how best to use the variable for identifying boosted QCD jets with invariant mass near \(\tau_6\), as shown in Fig. 2(b). Similarly, though, as shown in Fig. 2(c). Di-jet QCD jets are shown in Fig. 2.
N-subjettiness ratio

> bare $\tau_N$ has very little discrimination power

> take ratio $\tau_2/\tau_1$ instead

> mind: rather complicated variable, difficult to model $\Rightarrow$ need to validate in data

> clean sample of W-jets: top-antitop quark pairs used for calibration

![Graph showing distribution of $\tau_2/\tau_1$ ratio for CMS data, MC, and various processes like $g_{\text{Bulk}}(2 \text{ TeV}) \rightarrow WW$, $W'$(2 TeV) $\rightarrow$ WZ, and QCD PYTHIA8. The graph displays events per bin with $65 \text{ GeV} < M_\ell < 105 \text{ GeV}$ and $2.6 \text{ fb}^{-1}(13 \text{ TeV})$. The data shows a peak around $\tau_2/\tau_1 = 0.5$ with a comparison to the MC predictions and CMS data points.](image)
Higgs\(\rightarrow bb\) tagging

- Higgs has higher mass than W/Z bosons \(\rightarrow \tau_2/\tau_1\) less important, exploit b-jet content instead
- two different strategies:
  - identify b-subjets
  - tag fat jet
- currently, both show comparable performance
- 50% lower mis-tagging rate than W-/Z-tagging
- dedicated Higgs-tagger available soon

\[
H\rightarrow WW\rightarrow qqqq\text{ tagging is done using }\tau_4/\tau_2\text{ ratio (cf. }\tau_2/\tau_1\text{ for W/Z)}
\]
Higgs→ττ tagging

> τ-lepton can decay hadronically and leptonically

> need to take into account potential overlap between the two τ-leptons
  ▪ remove tracks/particles entering other isolation cone

> discrimination against q-/g-jets: MVA-based isolation
  ▪ sum reconstructed particle energies in various cones around τ decay products

> neutrinos in decay cannot be reconstructed → missing energy
  ▪ ττ-reconstruction using templates from Monte Carlo simulation (SVfit)

Currently only 8 TeV results
Lepton (muon + electron) reconstruction

>two **isolation** issues:
  - radiation from highly energetic leptons spoils isolation
  - leptons spoil each other’s isolation

>employ **dedicated high-\(p_T\) algorithms** to preserve high efficiency

>loosen selection criteria one of the leptons in \(Z \rightarrow ll\) decays

>leptonic W-decay: need to recover z-component of neutrino
  - use W-mass constraint for reconstruction
### Strategy/event selection for VV analyses

<table>
<thead>
<tr>
<th>Trigger</th>
<th>Lepton(s)</th>
<th>V-Jet</th>
<th>V Boson Candidate(s)</th>
<th>X</th>
<th>Search for bump in (m_{VV}) distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H_T) trigger (800 GeV) or jet+groomed mass</td>
<td>—</td>
<td>anti-(k_T) (R=0.8), (p_T &gt; 200) GeV, exploit substructure (\tau_2/\tau_1), use groomed mass</td>
<td>(\Delta \eta &lt; 1.3)</td>
<td>additional cuts on separation of bosons in (\Delta \phi) and (\Delta R)</td>
<td>reconstruct (X) using both reconstructed vector bosons</td>
</tr>
</tbody>
</table>

**Models and final states**

- Looking for \(m > \sim 600\) GeV \(VV/VH/HH\) resonances
- Models of extra dimensions
  - KK-Graviton \(G\) ! \(WW/ZZ/HH\)
  - Radion \(R\) ! \(WW/ZZ/HH\)
- Bulk graviton model prefers di-boson decays, while classical Randall-Sundrum model prefers di-fermion decays
- Models of compositeness, new forces
  - \(W'\) ! \(WZ/WH\), \(Z'\) ! \(WW/ZH\)
  - Technicolor \(\rho T C\) ! \(WZ\)
- Composite Higgs, Little Higgs model prefer di-boson decays, while sequential standard model prefers di-fermion decays
- Generalized models of heavy vector triplets (HVT) consisting of \(X^+\) and \(X^0\)
  - Bosons get high (\(p_T > \sim 200\) GeV) boost
  - special reconstruction techniques for jets and taus
  - special isolation techniques for leptons

**Diboson-like topology:**

- \(\Delta R(\ell, W_{had}) > \pi/2\)
- \(\Delta \phi(W_{had}, E_{miss}) > 2\)
- \(\Delta \phi(W_{had}, W_{lep}) > 2\)
Dijet (VV) analysis

> **trigger** at \( \sim 100\% \) efficiency at \( m_{JJ} > 1 \text{ TeV} \)
  
  - apply cut on reconstructed dijet system

> **define different** \( \tau_2/\tau_1 \) **regions:**
  
  - high purity to suppress background
  - low purity to recover signal efficiency at high masses

> **split W and Z** samples based on pruned jet mass (65-85, 85-105 GeV)

> still dominated by **QCD multi-jet** events

> difficult to obtain sufficient MC simulation statistics

> need a data-driven approach

> exponentially falling spectrum: use **fit function**
VV background estimation

- naively, fitting the $m_{JJ}$ spectrum could swallow signal
- also, need to avoid claiming false discovery (in particular in tail)
- fit function:
  \[
  \frac{dN}{dm_{jj}} = \frac{P_0}{(m_{jj}/\sqrt{s})^{P_2}} \quad \text{or} \quad \frac{dN}{dm_{jj}} = \frac{P_0(1 - m_{jj}/\sqrt{s})^{P_1}}{(m_{jj}/\sqrt{s})^{P_2}}
  \]
  2 parameters 3 parameters
- number of free parameters determined by F-test:
  - check if quality of fit improves by > 10% confidence level
  - if not, stick with current fit function
- extensive bias tests conducted
- combined signal+background fit performed

very similar background estimation strategy as for diphoton resonance search
Intermezzo: Zγ search overview

- recently published Z→ll + γ search
- inspired by γγ „excess“
- same photon ID as γγ search, dilepton ID as in ZV search, fit background
- limited by statistics, no significant excess
VW/VZ analysis overview

> for 2015 data, two separate analyses performed:
  - „low mass“: 600-1000 GeV
  - „high mass“: 1-4 TeV
  - VZ analysis not yet public

> difference low vs. high mass:
  - lower boost ➔ can use isolated lepton triggers with 27 GeV thresholds

> requiring an isolated lepton suppresses QCD multi-jet background significantly

> dominant backgrounds:
  - Drell-Yan/W+Jets
  - top-antitop quark production

> can estimate individual background components from sidebands
VW analysis background estimation

- statistics in MC simulated samples still limited
- furthermore, analysis performed in extreme phase space
- use pruned mass sidebands (40-65 GeV, 135-160 GeV) to exploit correlation between pruned jet mass and resonance mass
  - Higgs mass region kept blind
- determine ratio of simulated to data distributions in sideband
- extrapolate to signal region using transfer function (based on simulation)
- method accounts for data-MC differences in shape and normalisation
Intermezzo: \( \gamma \gamma \) vs. \( Z \gamma \) vs. \( WW \)

> limits for narrow resonances

- caveat: slightly different models used

> minimal upward fluctuations?

![Graph showing limits for narrow resonances](image1)

![Graph showing minimal upward fluctuations](image2)
VH analysis overview

- 2015 data: $H\rightarrow bb$, leptonic $W/Z$ decays
  - $W\rightarrow l\nu$
  - $Z\rightarrow ll$
  - $Z\rightarrow \nu\nu$

- Categorise in **single** and **double subjet b-tag** categories

- Same background estimation method as for VW search
Signal modelling and uncertainties

- depending on spin of new particles, **polarisation of bosons** different

- Bulk graviton (spin-2) and $W'/Z'$ (spin-1) models primarily couple to **longitudinal components of $W/Z$**

- Analytical description of **signal shapes** based on fully simulated benchmark mass points
  - double-sided Crystal-Ball function
  - linearly interpolation between benchmark points

- Signal efficiency up to 15% depending on analysis category

- Largest uncertainties:
  - background estimation
  - jet energy and mass scale
  - boson-tagging
Model interpretation

- several different analyses performed
- advantageous to use common benchmark models for easier interpretation
- need to know individual couplings to bosons
  - individual analysis: e.g. $\sigma(gg \rightarrow G \rightarrow WW)$
  - combination: e.g. $\sigma(gg \rightarrow G)$
  - mind also production mechanism

![Feynman diagrams](image-url)
Model tuning

> models described before can be tuned by a **handful of parameters**

> bulk graviton:

  - mass of graviton
  - coupling constant determining production cross section and width

> heavy vector triplet (HVT):

  - phenomenological Lagrangian
  - describes production and decay of heavy spin-1 resonances
  - 4 parameters for resonance mass, interaction strength, couplings to bosons and fermions
  - focus here on „Model B“ with enhanced couplings to bosons

---

![Feynman diagrams](https://example.com/feynman_diagram.png)

**Figure 1.1:** Heavy X particle production and decay.

---

![Graph](https://example.com/hvt_graph.png)

**Figure 2.1:**

Upper panel: Branching Ratios for the two body decays of the neutral vector \( V^0 \) for the benchmarks A \( g_V = 1 \) (left) and B \( g_V = 3 \) (right). Lower panel: Total widths corresponding to different values of the coupling \( g_V \) in the models A (left) and B (right).

Therefore, for model B \( g_V = 3 \) the dominant BRs are into di-bosons and the fermionic decays are extremely suppressed, of the order of one percent to one per mil. Moreover, the total width increases with increasing \( g_V \) since it is dominated by the di-boson width which grows with \( g_V \) as expected from Eq. (2.33). Finally, in model B we see that a very large coupling \( g_V \) (the case of \( g_V = 8 \) is shown in the Figure) leads to an extremely broad resonance, with \( \Gamma / M \ll 1 \), for which the experimental searches for a narrow resonance are no longer motivated. For this reason we expect, if no further suppression is present in the parameter \( c_H \), to be able to constrain heavy vector models from direct searches only up to \( g_V \) of the order 6–7. For larger couplings different searches become important, like for instance those for four fermion contact interactions (see for instance Refs. [83, 84]).
Putting it all together

- currently, larger number of channels has been covered with 8 TeV data
- differentiate between final states (number of leptons and jets)
- violet analyses interpreted in dark matter scenarios

<table>
<thead>
<tr>
<th></th>
<th>$W \rightarrow lv$</th>
<th>$Z \rightarrow ll$</th>
<th>$V \rightarrow qq$</th>
<th>$Z \rightarrow vv$</th>
<th>$H \rightarrow qqqq$</th>
<th>$H \rightarrow \tau\tau$</th>
<th>$H \rightarrow bb$</th>
<th>$H \rightarrow \gamma\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W \rightarrow lv$</td>
<td></td>
<td></td>
<td>13 TeV</td>
<td></td>
<td></td>
<td></td>
<td>13 TeV</td>
<td></td>
</tr>
<tr>
<td>$Z \rightarrow ll$</td>
<td></td>
<td></td>
<td>13 TeV</td>
<td></td>
<td></td>
<td></td>
<td>13 TeV</td>
<td></td>
</tr>
<tr>
<td>$V \rightarrow qq$</td>
<td>13 TeV</td>
<td>13 TeV</td>
<td>13 TeV</td>
<td>13 TeV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Z \rightarrow vv$</td>
<td>13 TeV</td>
<td></td>
<td></td>
<td>13 TeV</td>
<td></td>
<td></td>
<td>13 TeV</td>
<td></td>
</tr>
<tr>
<td>$H \rightarrow qqqq$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H \rightarrow \tau\tau$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H \rightarrow bb$</td>
<td>13 TeV</td>
<td>13 TeV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H \rightarrow \gamma\gamma$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- several analyses repeated and improved w.r.t. 8 TeV (considering $m_X > 1$ TeV)
- several more to come this year
While 13 TeV has opened up new energy regime, integrated luminosity recorded in 2015 significantly below the one of 2012

- 20 fb\(^{-1}\) vs. \(~3\) fb\(^{-1}\)

LHC is hadron collider - √s ≠ energy available in collision

need to consider parton luminosities

exceed 2012 reach with 2015 data already at 1-2 TeV resonance mass

nevertheless, worthwhile combining results
Combination of diboson analyses

> example here: $V'$ combination in HVT model B
> seven 8 TeV analyses, three at 13 TeV
> how to combine upper cross section limits from two different $\sqrt{s}$?

![Graph showing cross sections vs. $M_{V'}$](image)
Combination of diboson analyses

- convert cross section limits into **signal strength limits**
- 8+13 TeV limits comparable at lower masses, 13 TeV dominates high mass
- lower masses: leptonic analyses, higher masses: hadronic final states
Combination of diboson analyses

- can translate observed limits into **exclusion contours** in the HVT couplings space
- additional input to theorists for model building
Summary

- discovery of a diboson resonance might solve hierarchy problem
- however, currently no sign of new physics
- 13 TeV results already exceed 8 TeV ones
- expect another boost with 2016 data
W-tagging calibration

- cutting on $\tau_2/\tau_1$-ratio $\Rightarrow$ need to know efficiency of cut in data and simulation

- select at generator level clean W-events and those that do not match

- perform simultaneous fit
### Boson-Tagging Efficiencies

<table>
<thead>
<tr>
<th>Tagger</th>
<th>BR($W/Z/H \rightarrow xx$)</th>
<th>Efficiency</th>
<th>Mistag Rate (q-/g-jets)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W/Z\rightarrow qq$</td>
<td>70 %</td>
<td>35 %</td>
<td>1.2 %</td>
</tr>
<tr>
<td>$H\rightarrow bb$</td>
<td>57 %</td>
<td>35 %</td>
<td>0.5 %</td>
</tr>
<tr>
<td>$H\rightarrow WW\rightarrow qqqq$</td>
<td>10 %</td>
<td>35 %</td>
<td>1.5 %</td>
</tr>
<tr>
<td>$H\rightarrow tt$</td>
<td>6 %</td>
<td>35 %</td>
<td>0.03 %</td>
</tr>
</tbody>
</table>
Zγ search limits

> recently published $Z \rightarrow ll + \gamma$ search

![Graph showing Zγ search limits](image-url)

- **0.014% width**
  - Observed
  - Expected ± 1σ
  - Expected ± 2σ

- **5.4% width**
  - Observed
  - Expected ± 1σ
  - Expected ± 2σ

Resonance Mass [GeV]

CMS Preliminary

2.7 fb⁻¹ (13 TeV)

95% CL UL on $\sigma \times BR(A \rightarrow Z\gamma)$ [fb]