VH(bb) and the search for boosted ggH(bb) in CMS

Martin Kwok (Brown University)

on behalf of CMS collaboration

QCD@LHC 2019, Buffalo NY
18 July, 2019

PhysRevLett.121.121801    Phys. Rev. Lett. 120.071802
Why Higgs to bb?

- Higgs gives mass to elementary particles via two different mechanisms.
- Higgs to bb is a **direct** test of mass generation to the charged fermion sector.

**EWSB**

\[
\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi)
\]

\[
V(\Phi) = m^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2
\]

**Yukawa coupling**

\[
\mathcal{L}_{\text{Yukawa}} = - \hat{h}_{d_{ij}} \bar{q}_{L_i} \Phi d_{R_j} - \hat{h}_{u_{ij}} \bar{q}_{L_i} \Phi u_{R_j} - \hat{h}_{l_{ij}} \bar{l}_{L_i} \Phi e_{R_j}
\]

---

**CMS**

\[
\mathcal{g}_{Hff} = \frac{m_f}{\nu}
\]

Experimental challenges

Overwhelming background at LHC: QCD dijet events

QCD dijet $\sim O(\text{mb})$

Gluon-gluon fusion

$\sim 48 \text{ pb}$

VH production

$\sim 2.2 \text{ pb}$
Analysis strategy

- At the cost of lower cross section, an associated vector boson gives many handles to reject backgrounds.
- 5 different final states depending on decay modes.
- To maximise the sensitivity, machine learning techniques helped from various aspects.

0-lepton

$Z(\nu\nu)$

1-lepton

$W(\mu\nu)$

$W(e\nu)$

2-lepton

$Z(ee)$

$Z(\mu\mu)$
Analysis strategy

W/Z selection  Background rejection, Triggering

Higgs candidate reconstruction

Signal extraction with full event information

0-lepton  1-lepton  2-lepton

\(Z(\nu\nu)\)  \(W(\mu\nu)\)  \(W(e\nu)\)

\(Z(ee)\)  \(Z(\mu\mu)\)

\(p_T^{miss}\)  \(p_T^{miss}\)  \(p_T^{miss}\)

\(\mu\)  \(e\)  \(e^+e^-\)

\(\mu^+\mu^-\)
Analysis strategy

W/Z selection
Background rejection, Triggering

Higgs candidate reconstruction

Signal extraction with full event information

0-lepton

\[ Z(\nu\nu) \]

1-lepton

\[ W(\mu\nu) \]

\[ W(e\nu) \]

2-lepton

\[ Z(ee) \]

\[ Z(\mu\mu) \]
Analysis strategy

- **W/Z selection**
- **Background rejection, Triggering**
- **Higgs candidate reconstruction**
- **Signal extraction with full event information**

**0-lepton**

\[ Z(\nu\nu) \]

- **1-lepton**
  - **W(\mu\nu)**
  - **W(e\nu)**

- **2-lepton**
  - **Z(ee)**
  - **Z(\mu\mu)**

**Background rejection, Triggering**

- **W/Z selection**
- **Higgs candidate reconstruction**
- **Signal extraction with full event information**
**W/Z selection**

- **0-lepton**
  - $p_T^{\text{miss}}$

- **1-lepton**
  - $p_T^{\text{miss}}$, $e$

- **2-lepton**
  - $e^+e^-$, $\mu^+\mu^-$

**Equations**

- $p_T(V) = p_T^{\text{miss}}$
- $p_T(V) = p_T^{\text{miss}} + p_T(l)$
- $p_T(V) = p_T(ll)$, $m_{ll} \in [75 - 105]$ GeV

**Requirements**

- $p_T(V) > 170$ GeV
- $p_T(V) > 150$ GeV
- $p_T(V) \in (50,150] \cup (150,\infty)$ GeV

**Additional**

- Remaining background are estimated with dedicated control regions (CR)
**W/Z selection**

0-lepton

1-lepton

2-lepton

- **W/Z momentum** can be reconstructed in the final states

\[ p_T(V) = p_T^{miss} \]

\[ p_T(V) = p_T^{miss} + p_T(l) \]

\[ p_T(V) = p_T(ll), m_{ll} \in [75 - 105] \text{ GeV} \]

- All backgrounds are substantially reduced by requiring large \( p_T(V) \)

\[ p_T(V) > 170 \text{ GeV} \]

\[ p_T(V) > 150 \text{ GeV} \]

\[ p_T(V) \in (50,150] \cup (150,\infty) \text{ GeV} \]

- Remaining background are estimated with dedicated control regions (CR)

![Graphs and plots](Mention the specific images or graphs if available.)
Higgs reconstruction
b-tagging
Higgs mass resolution
Improved b-tagging with DNN

- Higgs candidate is reconstructed with the 2 jets with highest b-tagging score
- 10-20% better performance w.r.t Run1 due to
  - Extra pixel layer installed for 2017 data-taking period
  - Improved machine learning architecture: Deep Neural Network (DNN)
  - More input variables: track variables up to 6 tracks
DNN based b-jet regression

- b-jets resolution is generally worse than light (quark/gluon) jets
- Dedicated DNN to correct for mis-measured energy due to escaping neutrino

Jet kinematics
- Jet energy fractions binned in radius
- Jet constituent multiplicity
- SV, Jet pT relative to soft-lepton

Per-jet resolution estimator
- ~double improvement on $m_{jj}$

3 Outputs
- **energy correction**
- **$q_{25\%}$ resolution estimator**
- **$q_{75\%}$ resolution estimator**
DNN based b-jet regression

• b-jets resolution is generally worse than light (quark/gluon) jets
• Dedicated DNN to correct for mis-measured energy due to escaping neutrino

![Diagram showing b-jet regression](image)

Significant improvement on jet mass resolution

<table>
<thead>
<tr>
<th>Before regression</th>
<th>After regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma/\mu$ (%)</td>
<td>~15%</td>
</tr>
<tr>
<td></td>
<td>~12%</td>
</tr>
</tbody>
</table>

CMS Simulation Supplementary

- PV (Primary Vertex)
- Secondary Vertex

- Identified b-jets → b-tagging
- Correct mismeasured energy due to escaping neutrino → b-jet energy regression

2017 (13 TeV)
Kinematic fit

2-lepton channel: Further improve jet mass resolution with kinematic fit

- leptons have better momentum resolution than jets
- No intrinsic missing momentum
- Fit lepton and jet momentum within uncertainty

Significant improvement on jet mass resolution

Before regression | After regression | kin. fit

\( \sigma/\mu(\%) \)

\( \sim 15\% \) | \( \sim 12\% \) | \( \sim 10\% \)
Signal extraction with DNN

- Invariant mass is the most intuitive variable
- Other variables may have some sensitivity as well:
  \[ p_T(V), \Delta R(j_1,j_2), N_j, \text{DeepCSV} \]
- Highly correlated => Maximise sensitivity with DNN

\[ l^+l^- \]

\[ m_{jj} \in [60,160] \text{ GeV} \]

\[ m_{jj} \in [90,150] \text{ GeV} \]

DNN score

\[ \text{CMS} \]

\[ \text{Supplementary} \]

\[ 2\text{H}, \text{Low } p_T, \text{Low } m_{jj} \]

\[ 41.3 \text{ fb}^{-1} (13 \text{ TeV}) \]
Signal extraction with DNN

- Invariant mass is the most intuitive variable
- Other variables may have some sensitivity as well: $p_T(V)$, $\Delta R(j1,j2)$, $N_j$, DeepCSV
- Highly correlated => Maximise sensitivity with DNN
- Separate DNNs in each channel.

DNN score

PhysRevLett.121.121801

S.R.

0-lepton: $m_{jj} \in [60,160]$ GeV
1,2-lepton: $m_{jj} \in [90,150]$ GeV
Signal extraction with DNN

- Fit simultaneously all control regions and signal regions.
- Allow normalisations of $\bar{t}t$, $W+b(b)$, $Z+b(b)$ backgrounds to float independently in each channel.

Control Regions

Signal Regions - DNN
**VZ(bb) Validation**

- Same final state and similar kinematic as VH
  - Standard Candle!
- Minimal changes to analysis strategy

<table>
<thead>
<tr>
<th></th>
<th>VH(bb)</th>
<th>VZ(bb)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>mass</strong></td>
<td>$m_{jj} \in [90,150]$ GeV</td>
<td>$m_{jj} \in [60,160]$ GeV</td>
</tr>
<tr>
<td><strong>Signal</strong></td>
<td>VH(bb)</td>
<td>VZ(bb)</td>
</tr>
<tr>
<td><strong>Background</strong></td>
<td>VZ(bb)</td>
<td>VH(bb)</td>
</tr>
</tbody>
</table>

Observed (Expected) = 5.2 (5.0) sigma

$$\mu_{Zbb} = 1.05 \pm 0.22$$
Results

Fit with DNN distribution
More sensitive, but no mass peak

Fit with di-jet mass distribution
Less sensitive, but visible mass peak
Reaching 5 sigma

- 2017 analysis has the best sensitivity 3.3(3.1) sigma obs.(exp.)
- Run 1 + Run 2 VH(bb) = 4.9 (4.8) sigma
- Combine with ttH and other analysis

Observed (Expected) = 5.6 (5.5) sigma

\[ \mu_{Hbb} = 1.04 \pm 0.20 \]

<table>
<thead>
<tr>
<th>Data set</th>
<th>Significance ((\sigma))</th>
<th>Expected</th>
<th>Observed</th>
<th>Signal strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-lepton</td>
<td>1.9</td>
<td>1.3</td>
<td>0.73 ± 0.65</td>
<td></td>
</tr>
<tr>
<td>1-lepton</td>
<td>1.8</td>
<td>2.6</td>
<td>1.32 ± 0.55</td>
<td></td>
</tr>
<tr>
<td>2-lepton</td>
<td>1.9</td>
<td>1.9</td>
<td>1.05 ± 0.59</td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td>3.1</td>
<td>3.3</td>
<td>1.08 ± 0.34</td>
<td></td>
</tr>
<tr>
<td>Run 2</td>
<td>4.2</td>
<td>4.4</td>
<td>1.06 ± 0.26</td>
<td></td>
</tr>
<tr>
<td>Run 1 + Run 2</td>
<td>4.9</td>
<td>4.8</td>
<td>1.01 ± 0.23</td>
<td></td>
</tr>
</tbody>
</table>

\[ \leq 5.1 \text{ fb}^{-1} (7 \text{ TeV}) + \leq 19.8 \text{ fb}^{-1} (8 \text{ TeV}) + \leq 77.2 \text{ fb}^{-1} (13 \text{ TeV}) \]
Why Boosted Higgs?

- Above ~400 GeV, two b-jets from Higgs starts to merge
- Exploit jet substructure to reject background
- Only possible with substantial understanding/development with jet substructures
- CMS pioneered a search into this new phase space

\[
\Delta R(b\bar{b}) \sim 2m_H/p_T
\]

**two separated b-jets (R=0.4)**

**one merged double b-jet (R=0.8)**
Why Boosted Higgs?

- New physics may show up as anomalous coupling.
- Inclusive production cross section alone cannot distinguish modification of $\kappa_T$ and $\kappa_g$

$$\sigma_{incl.}(\kappa_T, \kappa_g) \sim (\kappa_T + \kappa_g)^2 \times \sigma_{SM}$$
- Contribution could be sizeable at high pT!
Analysis strategy

- Two main handles to reject QCD background:
  - Jet substructure : $N_2$
  - Dedicated double-b tagging
- Look for excess on a smoothly falling QCD mass spectrum:
- Data-driven QCD background estimation
- Double-b tagging de-correlated with jet mass

Control Region
- double-b < 0.9
- $pT=[600,675]$

Signal Region
- double-b > 0.9

For each $(pT, \text{mass})$ bin

Boosted H to $b\bar{b}$
Jet substructure

- Use generalised Energy Correlation Function (ECF) to identify 2-prong substructure
- $N_2$: Ratio of 3-point over 2-point correlation functions (smaller $N_2$ for more 2-prong)

$$N_2^\beta = 1 = \frac{2e_3^\beta}{(1e_2^\beta)^2}$$

- Theoretically motivated
- IRC safe
- No subjet axes

soft-drop jet grooming algorithm

$z_i$: energy fraction

$\Delta R$: angular separation
Jet substructure decorrelation

- Substructure variables are **highly correlated** with jet mass.
- Take QCD MC, define

\[
N_{2}^{1,\text{DDT}}(\rho, p_T) = N_{2}(\rho, p_T) - N_{2}^{(\varepsilon=X\%)}(\rho, p_T)
\]

- By construction, \( N_{2}^{1,\text{DDT}} < 0 \) has constant efficiency of QCD and does not sculpt mass spectrum.

See Cristina’s talk for more

[HIG-PAS-17-010]
Data driven background

- Residue shape difference due to data/MC is fitted with a smooth transfer function

\[ N(\rho, p_T | \text{pass}) = N(\rho, p_T | \text{fail}) \times F(\rho, p_T) \]

- Simultaneously fit $Z(b\bar{b})$, $H(b\bar{b})$ and transfer function
**Results**

- With 2016 dataset, (<1/3 of full Run 2 data)
- Observed Z(bb) with 5.1 sigma (5.8 sigma expected)
- Observed H(bb) with 1.5 sigma (0.7 sigma expected)
Summary

- H(bb) is observed with associated production modes VH, ttH
- Boosted topology opens up possibility for probing ggF production
  - Updating with better double-b tagger + full Run 2 data
- Next goal: VH(cc) and ggH(cc)
Summary

- H(bb) is observed with associated production modes VH, ttH
- Boosted topology opens up possibility for probing ggF production
  - Updating with better double-b tagger + full Run 2 data
- Next goal: VH(cc) and ggH(cc)

Stay tuned for updates!

CMS-PAS-HIG-18-031
Additional materials
**VH(bb) control regions**

- Control regions are mutually exclusive with each other
- Defined using properties of on Higgs side (b-tagging) or vetos of extra jets

---

---

---

[*] W+L.F and W+H.F taken from 1-lepton analysis

[*] W+H.F splitted in high and low mass
Double-b tagger

- Combines tracking and vertexing information in a boosted decision tree (BDT) with 27 input observables
- Input variables are chosen to carefully to have no strong dependence on mSD or pT