Search for Standard Model Higgs Decaying to $b\bar{b}$

Pierluigi Bortignon
on behalf of the CMS Collaboration
ETH Zurich

Higgs Hunting – Orsay – July 2012
Motivations

Higgs → b¯b

• A new resonance with mass near 125 GeV/c² has been observed in Higgs searches
• Properties of this resonance still unclear
• Combination of all CMS analyses shows fermion couplings (C_F) slightly suppressed
• H→bb key channel in the next future to characterize the new resonance
Introduction

Higgs -> b\bar{b} Status

- @ 125 GeV/c^2 b\bar{b} highest branching ratio
- At hadron colliders
  - ✗ Gluon fusion (huge QCD background)
  - ✔ Associated production with vector boson
- Analysis performed by 4 experiments
  - Tevatron (D0, CDF)
  - LHC (ATLAS, CMS)
  - None yet sensitive to SM cross section for m_H = 125 GeV/c^2
Analysis Strategy

Channels

• **Five channels** with different vector boson
  - \( W (l\nu_l) H(b\bar{b}) \), \( l = \) Electron, Muon
  - \( Z (ll) H(b\bar{b}) \), \( l = \) Electron, Muon, Neutrino

• **Large boost** on the vector boson important to suppress background

Vector Boson Selection

• Isolated lepton(s)
• Transverse Missing Energy

Higgs Candidate

• Two jets identified as B-Jets using dedicated likelihood discriminant
**Event display**

$M_{bb} = 128 \text{ GeV/c}^2$

$M_{LL} = 91 \text{ GeV/c}^2$

$p_{T LL} = 185 \text{ GeV/c}$
# Analysis Improvements

<table>
<thead>
<tr>
<th></th>
<th>7 TeV</th>
<th>Combined 7 – 8 TeV</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background Normalisation</td>
<td>Scale factors form solution of analytic system</td>
<td>Simultaneous fit using full shapes</td>
<td>Better confidence of background normalisation</td>
</tr>
<tr>
<td>Di-Jet Mass Resolution</td>
<td>Standard CMS Jet Energy Correction</td>
<td>Multivariate Jet Energy Regression in addition</td>
<td>~ 20% better mass resolution</td>
</tr>
<tr>
<td>Phase-Space Analyzed</td>
<td>Single bin in vector $p_T$</td>
<td>Add a lower vector $p_T$ bin to recover signal efficiency</td>
<td>~ 10% Higher sensitivity</td>
</tr>
<tr>
<td>Signal Extraction</td>
<td>Cut and Count on BDT discriminator output</td>
<td>Use of full shape information of BDT discriminator</td>
<td>~ 20% Higher sensitivity</td>
</tr>
</tbody>
</table>
Results (I)

Multivariate Analysis

- Combination of all 5 channels
- 7 TeV and 8 TeV data. 5 fb\(^{-1}\) each
- Two vector p\(_T\) categories
- MVA techniques (BDT) to classify signal against background
- BDT validated in control regions
- Fit on the shape of the BDT output distribution

Example of BDT discriminant output

- Combination of all 5 channels
- 7 TeV and 8 TeV data. 5 fb\(^{-1}\) each
- Two vector p\(_T\) categories
- MVA techniques (BDT) to classify signal against background
- BDT validated in control regions
- Fit on the shape of the BDT output distribution

Background-like

Signal-like
Results (II)

- ~ 50% improvements on the analysis since 2011 published results
- CMS is the most sensitive experiment for this channel by now
- About a factor 2.5 in integrated luminosity needed to be sensitive at $m_H = 125$ GeV/$c^2$

Higgs mass resolution ~ 8%

=> Excess compatible with $\gamma\gamma$ and $ZZ$

Higgs observation at 126 GeV/$c^2$
• Higgs to $b\bar{b}$ results are in the **official**

**Higgs combination** of CMS

• Essential channel to **characterize the**

**observation** in $\gamma\gamma$ and ZZ channels

• Important to **measure** Higgs couplings

• LHC is expected to provide about

20 fb$^{-1}$ before the next long shut-down

• **Improved analysis** will be sensitive

to SM Higgs with full LHC dataset
\[ M_{bb} = 128 \text{ GeV/c}^2 \]
\[ M_{LL} = 91 \text{ GeV/c}^2 \]
\[ p_{TLL} = 185 \text{ GeV/c} \]
References


Higgs -> b\bar{b} Status

- For low SM Higgs mass hypothesis bb decay channel has the highest branching ratio
- At hadron colliders very difficult to suppress QCD background if looking at gluon fusion production
  - Way out: require associated production with a vector boson
- Phenomenology paper shows the importance of vectors boson $p_T$ at LHC for this channel
- Tevatron not sensitive to SM Higgs at 125 GeV/c$^2$
- ATLAS has published 2011 results only
Higgs Branching Ratios

Branching ratios

\( \frac{\text{M}_{H}}{[\text{GeV}]} \)

- \( bb \)
- \( WW \)
- \( ZZ \)
- \( \gamma \gamma \)
- \( Z\gamma \)
- \( gg \)
- \( cc \)

-1, 10, -2, 10, -1, 10, 1

LHC Higgs XS WG 2010

Introduction
Gluon fusion production combined with $b\bar{b}$ decay is extremely challenging at hadron colliders

- B hadron production cross section is huge compared to the Higgs one
- QCD B hadron production is a mostly irreducible background
- B Jet energy resolution not optimal to look for a peak on the $m_{BB}$ distribution
- Difficult to design an efficient trigger for this topology
Associated production combined with $b\bar{b}$ decay is a better analysis strategy

- Main background $V + 2 \, b$
- $S/B$ increases by $\sim 4$ orders of magnitude
- Partially reducible background using vector boson momentum
- Easier to trigger using isolated leptons or high transverse missing energy
Low Mass Higgs at LHC

Higgs Production Cross Sections

\[ \sqrt{s} = 7 \text{ TeV} \]

- \( pp \rightarrow H \) (NNLO+NNLL QCD + NLO EW)
- \( pp \rightarrow qqH \) (NNLO QCD + NLO EW)
- \( pp \rightarrow WH \) (NNLO QCD + NLO EW)
- \( pp \rightarrow ZH \) (NNLO QCD +NLO EW)
- \( pp \rightarrow t\bar{t}H \) (NLO QCD)

\( ggg \) fusion:
- \( g \rightarrow t \rightarrow H^0 \)
- \( g \rightarrow W, Z \) fusion:
- \( q \rightarrow W, Z \)
- \( W, Z \) bremsstrahlung
• Five different channels
  – $W$ (EleNu) $H(bb)$
  – $W$ (MuNu) $H(bb)$
  – $Z$(EleEle) $H(bb)$
  – $Z$(MuMu) $H(bb)$
  – $Z$(NuNu) $H(bb)$

• Different channels have different background composition
• Five different channels
  – W (EleNu) H(bb)
  – W (MuNu) H(bb)
  – Z(EleEle) H(bb)
  – Z(MuMu) H(bb)
  – Z(NuNu) H(bb)

• Different channels have different background composition
Five different channels

- $W$ (EleNu) $H(bb)$
- $W$ (MuNu) $H(bb)$
- $Z$(EleEle) $H(bb)$
- $Z$(MuMu) $H(bb)$
- $Z$(NuNu) $H(bb)$

Different channels have different background composition
**Most Discriminating Variables**

B-Tag working point efficiency \(\sim 50\%\)

Invariant mass resolution \(\sim 10\%\)
### Analysis strategy

**Boosted Decision Tree Variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>W(ℓν)H</th>
<th>Z(ℓℓ)H</th>
<th>Z(νν)H</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_{ℓℓ} )</td>
<td></td>
<td>( 75 &lt; m_{ℓℓ} &lt; 105 )</td>
<td>–</td>
</tr>
<tr>
<td>( p_T(ℓ_1) )</td>
<td>( &gt; 30 )</td>
<td>( &gt; 20 )</td>
<td>( &gt; 80 )</td>
</tr>
<tr>
<td>( p_T(ℓ_2) )</td>
<td>( &gt; 30 )</td>
<td>( &gt; 20 )</td>
<td>( &gt; 20 )</td>
</tr>
<tr>
<td>( p_T(jj) )</td>
<td>( &gt; 120 )</td>
<td>–</td>
<td>( 120 – 160 )</td>
</tr>
<tr>
<td>( m(jj) )</td>
<td>( &lt; 250 )</td>
<td>( 80 &lt; m(jj) &lt; 150 )</td>
<td>( &lt; 250 )</td>
</tr>
<tr>
<td>( p_T(V) )</td>
<td>( 120 – 170 )</td>
<td>( 50 – 100 )</td>
<td>( &gt; 100 )</td>
</tr>
<tr>
<td>( CSV_{\text{max}} )</td>
<td>( &gt; 0.40 )</td>
<td>( 0.50 )</td>
<td>( &gt; 0.50 )</td>
</tr>
<tr>
<td>( CSV_{\text{min}} )</td>
<td>( &gt; 0.40 )</td>
<td>( 0.244 )</td>
<td>( &gt; 0.50 )</td>
</tr>
<tr>
<td>( N_{\text{al}} )</td>
<td>–</td>
<td>–</td>
<td>( = 0 )</td>
</tr>
<tr>
<td>( Δφ(E_T^{\text{miss}}, \text{jet}) )</td>
<td>–</td>
<td>–</td>
<td>( &gt; 0.5 )</td>
</tr>
<tr>
<td>( E_T^{\text{miss}} )</td>
<td>( &gt; 35 )</td>
<td>–</td>
<td>( 120 – 160 )</td>
</tr>
<tr>
<td>BDT</td>
<td>full distribution</td>
<td>full distribution</td>
<td>full distribution</td>
</tr>
</tbody>
</table>

**Preselection**
Background Normalisation

Main backgrounds are taken from the data. Simulation only used to extrapolate to the signal region.

The control regions are defined as close as possible to the signal region to minimize the extrapolation.

### Good agreement with theory predictions

**TTbar**

**W + B Jets**

**Z + B Jets**

### Data/MC Scale Factors

<table>
<thead>
<tr>
<th>Process</th>
<th>WH</th>
<th>$Z(\ell\ell)H$</th>
<th>$Z(\nu\nu)H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low $p_T$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W + udscg</td>
<td>0.88 ± 0.01 ± 0.03</td>
<td>–</td>
<td>0.89 ± 0.03 ± 0.03</td>
</tr>
<tr>
<td>$Wb\bar{b}$</td>
<td>1.91 ± 0.14 ± 0.31</td>
<td>1.36 ± 0.10 ± 0.15</td>
<td></td>
</tr>
<tr>
<td>$Z + udscg$</td>
<td>1.11 ± 0.03 ± 0.11</td>
<td>0.87 ± 0.01 ± 0.03</td>
<td></td>
</tr>
<tr>
<td>$Zb\bar{b}$</td>
<td>–</td>
<td>0.98 ± 0.05 ± 0.12</td>
<td>0.96 ± 0.02 ± 0.03</td>
</tr>
<tr>
<td>$\tau\tau$</td>
<td>0.93 ± 0.02 ± 0.05</td>
<td>1.03 ± 0.04 ± 0.11</td>
<td>0.97 ± 0.02 ± 0.04</td>
</tr>
</tbody>
</table>

| High $p_T$|            |               |               |
| W + udscg| 0.79 ± 0.01 ± 0.02 | –             | 0.78 ± 0.02 ± 0.03 |
| $Wb\bar{b}$| 1.49 ± 0.14 ± 0.19 | 1.48 ± 0.15 ± 0.20 |
| $Z + udscg$| 1.11 ± 0.03 ± 0.11 | 0.97 ± 0.02 ± 0.04 |
| $Zb\bar{b}$| –          | 0.98 ± 0.05 ± 0.12 | 1.08 ± 0.09 ± 0.06 |
| $\tau\tau$| 0.84 ± 0.02 ± 0.03 | 1.03 ± 0.04 ± 0.11 | 0.97 ± 0.02 ± 0.04 |

### Z + Light Jets

### Z + B Jets

### Z + B Jets

### Z + B Jets

### Z + B Jets

---

Image credit: CMS Preliminary, Higgs Hunting - 18th July 2012 - Orsay
Jet Energy Regression

Di-Jet Mass Resolution improvement

- Specialized b jet energy regression similar to CDF
  - => improve dijet invariant mass (and MET)
- Use MVA regression trained with b quark jets discriminant, properties of a secondary vertex, track information, charged constituents, variables related to the energy reconstruction of the jet etc.
- Attempts to recover the true b-jet energy.
  - Validated in MC and in data using $p_T$ balance in $Z + 2$ Jets events
- Upshot
  - 15-20% improved mass resolution
  - $m_{bb}$ distribution becomes more consistent with true generated mass spectrum
Results

Shape Fit of the BDT Output
## Results

<table>
<thead>
<tr>
<th>$m_H$ (GeV)</th>
<th>110</th>
<th>115</th>
<th>120</th>
<th>125</th>
<th>130</th>
<th>135</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp.</td>
<td>1.16</td>
<td>1.26</td>
<td>1.35</td>
<td>1.64</td>
<td>2.12</td>
<td>2.81</td>
</tr>
<tr>
<td>Obs.</td>
<td>1.39</td>
<td>1.82</td>
<td>2.24</td>
<td>2.11</td>
<td>4.20</td>
<td>3.39</td>
</tr>
</tbody>
</table>

**CMS Preliminary:**
- $\sqrt{s} = 7$ TeV, $L = 5.0$ fb$^{-1}$
- $\sqrt{s} = 8$ TeV, $L = 5.1$ fb$^{-1}$

VH(bb) Compatibility