Fully hadronic Higgs decay $H \rightarrow WW^* \rightarrow qqqq$

in Higgsstrahlung $HZ, Z \rightarrow qq$ at 250 GeV CepC

Mila Pandurović

Vinca Institute of Nuclear Sciences
University of Belgrade, Serbia
Introduction

- Analyzed HZ fully hadronic decay, signal: $Z \rightarrow qq$, $H \rightarrow WW^* \rightarrow qqqq$
- $\text{BF}_{H126 \rightarrow WW} \sim 23.0\%$, $\text{BF}_{WW \rightarrow qqqq} \sim 45.4\% \Rightarrow \text{signal} \sim 10\% \text{ of Higgs decays}$
- $\sigma_{HZ, Z \rightarrow qq} \sim 143.39\ fb \ (\text{unpolarized beams})$
- $\sigma_{(HZ, Z \rightarrow qq, H \rightarrow WW^* \rightarrow qqqq)} \sim 16.12\ fb$

- Measurement of the relative branching fraction

- Signal signature: 6 central jets in the final state
- Goal of the analysis:
  - Calculate the statistical potential for the determination of the specific Higgs couplings
  - Verify the analysis strategy
Two Analysis strategies

Fast Jet: Forcing events into 6 jets

<table>
<thead>
<tr>
<th>Preselection</th>
<th>MV selection</th>
</tr>
</thead>
</table>

\[ \frac{\Delta \sigma}{\sigma} = \frac{\sqrt{S + B}}{S} \]

Relative statistical precision

Lcfi Vertexing

Static cuts analysis
Signal reconstruction

- $k_T$ exclusive, particle flow with Arbor v3.1
- Jet formation: force events into 6 jets, do the jet pairing to form H ($W$ and $W^*$), Z

- Fit in boson the peak vicinity ($\pm 10$ GeV, $\pm 5$ GeV,) for the Higgs and the Z boson for several jet openings $R=0.8, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5$
- The best result is obtained for $R=1.5$

![Histograms](image.png)
Reconstruction of the Higgs, Z and W bosons

- In order to reconstruct the Higgs, Z, W boson reconstruction the event is forced into six jets
- Obtained jets are grouped into three pairs to form the W, W* and Z bosons
- From WW* pair - the Higgs boson
- The combination which minimizes the $\chi^2$ is chosen:

$$\chi^2 = \frac{(m_{ij} - m_w)^2}{\sigma^2_w} + \frac{(m_{kl} - m_z)^2}{\sigma^2_z} + \frac{(m_{ijmn} - m_H)^2}{\sigma^2_H}$$

- For the corresponding $\sigma$ are the WA width was taken $\sigma^2_{H,W,Z}$
Reconstructed boson invariant masses for signal
### Signal and background samples

<table>
<thead>
<tr>
<th>Sample</th>
<th>$\sigma [fb]$</th>
<th>#evts/5ab$^{-1}$</th>
<th>#evts used</th>
</tr>
</thead>
<tbody>
<tr>
<td>$qqh \rightarrow q\bar{q}WW^* \rightarrow q\bar{q}q\bar{q}q\bar{q}$</td>
<td>16.12</td>
<td>80600</td>
<td>74342</td>
</tr>
<tr>
<td>Other Higgs decays</td>
<td>127.27</td>
<td>636350</td>
<td>644354</td>
</tr>
<tr>
<td>Non $qqh \rightarrow q\bar{q}WW^* \rightarrow q\bar{q}q\bar{q}q\bar{q}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$2f$</td>
<td>49561.30</td>
<td>247806500</td>
<td>100000</td>
</tr>
<tr>
<td>$4f_{ww_cuxx}$</td>
<td>3395.48</td>
<td>16977400</td>
<td>1220200</td>
</tr>
<tr>
<td>$4f_{ww_ccbs}$</td>
<td>5.74</td>
<td>28700</td>
<td>99400</td>
</tr>
<tr>
<td>$4f_{ww_ccds}$</td>
<td>165.57</td>
<td>827850</td>
<td>1343474</td>
</tr>
<tr>
<td>$4f_{ww_uubd}$</td>
<td>0.05</td>
<td>250</td>
<td>99800</td>
</tr>
<tr>
<td>$4f_{ww_uusd}$</td>
<td>165.94</td>
<td>829700</td>
<td>691057</td>
</tr>
<tr>
<td>$4f_{Mix_udud}$</td>
<td>1570.40</td>
<td>7852000</td>
<td>2782962</td>
</tr>
<tr>
<td>$4f_{Mix_cscs}$</td>
<td>1568.94</td>
<td>7844700</td>
<td>2375076</td>
</tr>
<tr>
<td>$4f_{zz_utut}$</td>
<td>83.09</td>
<td>415450</td>
<td>400000</td>
</tr>
<tr>
<td>$4f_{zz_dtdt}$</td>
<td>226.20</td>
<td>1131000</td>
<td>332600</td>
</tr>
<tr>
<td>$4f_{zz_uu_notd}$</td>
<td>95.65</td>
<td>478250</td>
<td>477400</td>
</tr>
<tr>
<td>$4f_{zz_cc_notd}$</td>
<td>96.04</td>
<td>480200</td>
<td>337400</td>
</tr>
</tbody>
</table>
Investigated variables

- Invariant masses: $m_{\text{Higgs}}$, $m_Z$, $m_W$, $m_{w^*}$
- Number of particle flow objects NPFO
- Visible energy $E_{\text{vis}}$
- The highest transverse momentum of the jet in the event – $\text{highestPtJet}$
- Transverse momentum of the Higgs boson $\text{PtOfHiggsJets}$
- Event shape variables: thrust, oblatness, sphericity, aplanarity
- Jet transitions: $y_{12}$, $y_{23}$, $y_{34}$, $y_{45}$, $y_{56}$, $y_{67}$
- Force event into 2 jet: $\text{btag1}$, $\text{btag2}$, $\text{btag1*btag2}$
- $\text{ctag1}$, $\text{ctag2}$
- Force event into 6 jet: $\text{btag}_i$, $\text{ctag}_i$
- Angle between jets that comprise $W$ boson: $\text{ThetaWqq}$
- Angle between jets that comprise $Z$ boson: $\text{ThetaZqq}$
- Angle between $W$ and $W^*$ that comprise the Higgs boson: $\text{ThetaHiggsW1W2}$
- Arithmetic variable $\text{Energy*Theta}$ of the $W$, Higgs and $Z$ boson
Invariant masses
Number of particle flow objects

The graph shows the distribution of particle flow objects with counts on the y-axis and a range from 0 to 2500 on the x-axis. The categories include:

- Signal $Z \rightarrow \ell\ell$, $H \rightarrow WW \rightarrow \ell\ell\ell\ell$
- $HZ$, other $H$ decays
- 2f hadronic
- 4f $WW$ cuxx
- 4f ZZ dtdt
- 4f WW/ZZ udud

The NPFO value is indicated on the graph.
The event shape variables

![Histograms showing event shape variables](image-url)
Jet transitions

- The $k_t$ values at which the number of jet goes from $i \rightarrow i + 1$ number of jets
New variable construction based on signal Monte Carlo information

The distribution of the energy of the W real boson versus the angle between jets that comprise it.
New variable construction energy theta of the Z boson

The distribution of the energy of the W real boson versus the angle between jets that comprise it.
Arithmetic Variables Energy*Theta for W boson

Rec Signal

Rec nonWW Higgs decays

Rec 2f

Rec 4f_WW_cuxx
## Multivariate approach: preselection

- $8000 < \text{Energy} \times \Theta < 14000$. $10000 < \text{Energy} \times \Theta < 17000$. NPFO $> 80$.

<table>
<thead>
<tr>
<th>sample</th>
<th>$\sigma [fb]$</th>
<th>#evts/5ab$^{-1}$</th>
<th>$\epsilon_{\text{pres}} [%]$</th>
<th>evts after preselection</th>
</tr>
</thead>
<tbody>
<tr>
<td>$qqh \rightarrow q\bar{q}WW^* \rightarrow q\bar{q}q\bar{q}q\bar{q}$</td>
<td>16.12</td>
<td>80600</td>
<td>70.0</td>
<td>56380</td>
</tr>
<tr>
<td>other Higgs decays</td>
<td>127.27</td>
<td>636350</td>
<td>43.0</td>
<td>273975</td>
</tr>
<tr>
<td>non $qqh \rightarrow q\bar{q}WW^* \rightarrow q\bar{q}q\bar{q}q\bar{q}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2f</td>
<td>49561.30</td>
<td>247806500</td>
<td>0.8</td>
<td>1990414</td>
</tr>
<tr>
<td>4f $ww_{cuxx}$</td>
<td>3395.48</td>
<td>16977400</td>
<td>16.7</td>
<td>2838452</td>
</tr>
<tr>
<td>4f $ww_{ccbs}$</td>
<td>5.74</td>
<td>28700</td>
<td>22.5</td>
<td>6453</td>
</tr>
<tr>
<td>4f $ww_{ccds}$</td>
<td>165.57</td>
<td>827850</td>
<td>18.3</td>
<td>151787</td>
</tr>
<tr>
<td>4f $ww_{uubd}$</td>
<td>0.05</td>
<td>250</td>
<td>19.8</td>
<td>50</td>
</tr>
<tr>
<td>4f $ww_{uud}$</td>
<td>165.94</td>
<td>829700</td>
<td>15.3</td>
<td>127241</td>
</tr>
<tr>
<td>4f $ww_{ZZ_uud}$</td>
<td>1570.40</td>
<td>7852000</td>
<td>16.0</td>
<td>1255551</td>
</tr>
<tr>
<td>4f $ww_{ZZ_cscs}$</td>
<td>1568.94</td>
<td>7844700</td>
<td>17.9</td>
<td>1406147</td>
</tr>
<tr>
<td>4f $ZZ_{utut}$</td>
<td>83.09</td>
<td>415450</td>
<td>22.0</td>
<td>91366</td>
</tr>
<tr>
<td>4f $ZZ_{dtdt}$</td>
<td>226.20</td>
<td>1131000</td>
<td>27.5</td>
<td>311025</td>
</tr>
<tr>
<td>4f $ZZ_{uu_notd}$</td>
<td>95.65</td>
<td>478250</td>
<td>23.7</td>
<td>113345</td>
</tr>
<tr>
<td>4f $ZZ_{cc_nots}$</td>
<td>96.04</td>
<td>480200</td>
<td>27.8</td>
<td>133496</td>
</tr>
</tbody>
</table>
Multivariate analysis

- The training of BDTG was performed on ten background samples excluding:
  - 2f backgrounds, 4f_WW_ccbs and 4f_WW_uubd

- The variables set was optimized to a set with the minimal stable relative statistical error (41 variables investigated – 18 final variables)
  - Invariant masses: \( m_{\text{Higgs}} \), \( m_Z \), \( m_W \)
  - Number of particle flow objects NPFO
  - Highest PtJet, transverse momentum of jets that comprise Higgs boson - PtOfHiggsJets
  - Event shape variables: thrust, oblatness, aplanarity
  - Jet transitions: \( y_{12} \), \( y_{34} \), \( y_{45} \), \( y_{56} \), \( y_{67} \)
  - Force event into 2 jet: btag1, btag2
  - ctag1
  - Arithmetic variable Energy*Theta of the Z boson
After preselection and multivariate analysis ~99% of the background is reduced

<table>
<thead>
<tr>
<th>sample</th>
<th>$\sigma [fb]$</th>
<th>#evts /5ab$^{-1}$</th>
<th>evts after preselection</th>
<th>$\varepsilon_{\text{tmva}}$ [%]</th>
<th>$\varepsilon_{\text{total}}$ [%]</th>
<th>evts after final selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q\bar{q}h \to q\bar{q}WW^* \to q\bar{q}q\bar{q}q\bar{q}$</td>
<td>16.12</td>
<td>80600</td>
<td>56380</td>
<td>41.3</td>
<td>28.85</td>
<td>23257</td>
</tr>
<tr>
<td>other Higgs decays</td>
<td>127.27</td>
<td>636350</td>
<td>273975</td>
<td>14.1</td>
<td>6.1</td>
<td>38629</td>
</tr>
<tr>
<td>non $q\bar{q}h \to q\bar{q}WW^* \to q\bar{q}q\bar{q}q\bar{q}$</td>
<td>127.27</td>
<td>636350</td>
<td>273975</td>
<td>14.1</td>
<td>6.1</td>
<td>38629</td>
</tr>
<tr>
<td>2f</td>
<td>49561.30</td>
<td>247806500</td>
<td>1990414</td>
<td>0.25</td>
<td>0.002</td>
<td>4976</td>
</tr>
<tr>
<td>4f$_{ww}$cuxx</td>
<td>3395.48</td>
<td>16977400</td>
<td>2838452</td>
<td>1.45</td>
<td>0.24</td>
<td>41188</td>
</tr>
<tr>
<td>4f$_{ww}$ccbs</td>
<td>5.74</td>
<td>28700</td>
<td>6453</td>
<td>1.7</td>
<td>0.38</td>
<td>110</td>
</tr>
<tr>
<td>4f$_{ww}$ccds</td>
<td>165.57</td>
<td>827850</td>
<td>151787</td>
<td>1.5</td>
<td>0.28</td>
<td>2294</td>
</tr>
<tr>
<td>4f$_{ww}$uubd</td>
<td>0.05</td>
<td>250</td>
<td>50</td>
<td>2.0</td>
<td>0.4</td>
<td>1</td>
</tr>
<tr>
<td>4f$_{ww}$uusd</td>
<td>165.94</td>
<td>829700</td>
<td>127241</td>
<td>0.8</td>
<td>0.13</td>
<td>1073</td>
</tr>
<tr>
<td>4f$<em>{ww}$zz$</em>{udud}$</td>
<td>1570.40</td>
<td>7852000</td>
<td>1255551</td>
<td>1.5</td>
<td>0.24</td>
<td>19102</td>
</tr>
<tr>
<td>4f$<em>{ww}$zz$</em>{cscs}$</td>
<td>1568.94</td>
<td>7844700</td>
<td>1406147</td>
<td>1.6</td>
<td>0.29</td>
<td>22514</td>
</tr>
<tr>
<td>4f$_{zz}$utut</td>
<td>83.09</td>
<td>415450</td>
<td>91366</td>
<td>5.5</td>
<td>1.2</td>
<td>4997</td>
</tr>
<tr>
<td>4f$_{zz}$dtd</td>
<td>226.20</td>
<td>1131000</td>
<td>311025</td>
<td>6.4</td>
<td>1.8</td>
<td>19845</td>
</tr>
<tr>
<td>4f$<em>{zz}$uu$</em>{notd}$</td>
<td>95.65</td>
<td>478250</td>
<td>113345</td>
<td>5.9</td>
<td>1.4</td>
<td>6675</td>
</tr>
<tr>
<td>4f$<em>{zz}$cc$</em>{nots}$</td>
<td>96.04</td>
<td>480200</td>
<td>133496</td>
<td>6.0</td>
<td>1.7</td>
<td>7949</td>
</tr>
</tbody>
</table>
The relative statistical uncertainty: MVA method

- The dominant background after final selection are $ee \rightarrow qqqq$ backgrounds
- The high cross-section $2f \rightarrow q\bar{q}$ background show good response to the preselection and multivariate analysis. The obtained relative statistical precision is 1.9% with the corresponding signal efficiency of 29%

$$\frac{\Delta \sigma}{\sigma} = \frac{\sqrt{S + B}}{S} \approx 1.9\%$$
Static cut analysis

- The staticic cut variables used:
  - Invariant masses: $80 < m_Z < 100$ GeV
  - Invariant masses: $115 < m_H < 135$ GeV
  - Number of particle flow objects NPFO $> 90$ GeV
  - Highest PtJet $< 90$
  - transverse momentum of jets that comprise Higgs boson $< 80$ GeV
  - Jet transitions: $y_{23} < 2.4$
  - $y_{34} < 2.4$
  - $Y_{45} < 2.7$
  - $y_{56} < 3.2$
  - $y_{67} < 3.5$
  - Arithmetic variable Energy*Theta of the $Z$ boson $8000 < EnThW < 14000$
  - Arithmetic variable Energy*Theta of the $Z$ boson $10000 < EnThZ < 17000$
### Static cuts analysis results

<table>
<thead>
<tr>
<th>sample</th>
<th>$\sigma [fb]$</th>
<th>#evts/5ab$^{-1}$</th>
<th>$\varepsilon_{\text{tot mva}}$ [%]</th>
<th>$\varepsilon_{\text{static}}$ [%]</th>
<th>evts after final selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>$qqh \rightarrow q\bar{q}WW^* \rightarrow q\bar{q}q\bar{q}$</td>
<td>16.12</td>
<td>80600</td>
<td>28.85</td>
<td>28.9</td>
<td>23293</td>
</tr>
<tr>
<td>other Higgs decays</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non $qqh \rightarrow q\bar{q}WW^* \rightarrow q\bar{q}q\bar{q}$</td>
<td>127.27</td>
<td>636350</td>
<td>6.1</td>
<td>8.1</td>
<td>51544</td>
</tr>
<tr>
<td>$2f$</td>
<td>49561.30</td>
<td>247806500</td>
<td>0.002</td>
<td>0.02</td>
<td>49561</td>
</tr>
<tr>
<td>$4f_{ww_cuxx}$</td>
<td>3395.48</td>
<td>16977400</td>
<td>0.24</td>
<td>1.5</td>
<td>254661</td>
</tr>
<tr>
<td>$4f_{ww_ccbs}$</td>
<td>5.74</td>
<td>28700</td>
<td>0.38</td>
<td>1.9</td>
<td>545</td>
</tr>
<tr>
<td>$4f_{ww_ccds}$</td>
<td>165.57</td>
<td>827850</td>
<td>0.28</td>
<td>1.6</td>
<td>13246</td>
</tr>
<tr>
<td>$4f_{ww_uubd}$</td>
<td>0.05</td>
<td>250</td>
<td>0.4</td>
<td>1.8</td>
<td>5</td>
</tr>
<tr>
<td>$4f_{ww_uusd}$</td>
<td>165.94</td>
<td>829700</td>
<td>0.13</td>
<td>1.3</td>
<td>10786</td>
</tr>
<tr>
<td>$4f_{ww_zz_udud}$</td>
<td>1570.40</td>
<td>7852000</td>
<td>0.24</td>
<td>1.4</td>
<td>109928</td>
</tr>
<tr>
<td>$4f_{ww_zz_cscs}$</td>
<td>1568.94</td>
<td>7844700</td>
<td>0.29</td>
<td>1.6</td>
<td>125515</td>
</tr>
<tr>
<td>$4f_{zz_utut}$</td>
<td>83.09</td>
<td>415450</td>
<td>1.2</td>
<td>2.4</td>
<td>9971</td>
</tr>
<tr>
<td>$4f_{zz_dtdt}$</td>
<td>226.20</td>
<td>1131000</td>
<td>1.8</td>
<td>2.9</td>
<td>32799</td>
</tr>
<tr>
<td>$4f_{zz_uu_notd}$</td>
<td>95.65</td>
<td>478250</td>
<td>1.4</td>
<td>2.5</td>
<td>11956.</td>
</tr>
<tr>
<td>$4f_{zz_cc_nots}$</td>
<td>96.04</td>
<td>480200</td>
<td>1.7</td>
<td>2.9</td>
<td>13926</td>
</tr>
</tbody>
</table>

- After the static cut analysis ~98% of the background is reduced.
- The obtained relative statistical uncertainty 3.6 % with the corresponding signal efficiency of 29%
Summary

- The fully hadronic decay is most abundant channel in the $H \to WW^*$ decay
- In Higgsstrahlung, $Z \to qq$, this decay leads complex central six jet final state
- High cross section hadronic backgrounds
- The channel is analysed with two types of analysis flow:
  - multivariate analysis
  - static cut analysis
- The multivariate approach showed better reduction capabilities in comparison to the static cut analysis
- This is due to lack of distinct cut variables for hadronic final state
- The obtained relative statistical precision with the static cut analysis is 3.6% with the signal efficiency of 29%, while the result obtained with the multivariate analysis is 1.9% with the corresponding signal efficiency of 29% also
- The result is obtained for the integrated luminosity of 5 ab$^{-1}$
Fully hadronic Higgs decay $H \rightarrow ZZ^* \rightarrow qqqq$

in Higgsstrahlung $HZ, Z \rightarrow qq$ at 250 GeV CepC

Very preliminary test of the estimation of the stat. uncertainty
Introduction

- Analyzed HZ fully hadronic decay, signal: \( Z \rightarrow qq, H \rightarrow WW^* \rightarrow qqqq \)
- \( BF_{H126 \rightarrow WW} \approx 2.67\% \), \( BF_{WW \rightarrow qqqq} \approx 48.87\% \)
- \( \sigma_{HZ, Z \rightarrow qq} \approx 143.39 \text{ fb (unpolarized beams)} \)
- \( \sigma_{(HZ, Z \rightarrow qq, H \rightarrow WW^* \rightarrow qqqq)} \approx 1.87 \text{ fb} \)

- Measurement of the relative branching fraction

- Signal signature: 6 central jets in the final state
- Low cross section

- Possibility of the estimation of the rel. statistical uncertainty of the cross-section \( xBF \)
Reconstruction of the Higgs, Z and W bosons

- In order to reconstruct the Higgs, Z, Z,Z* boson reconstruction the event is forced into six jets
- Obtained jets are grouped into three pairs to form the Z, Z* from Higgs decay and Z boson from HZ
- The combination which minimizes the $\chi^2$ is chosen:

$$
\chi^2 = \frac{(m_{ij} - m_Z)^2}{\sigma^2_Z} + \frac{(m_{kl} - m_Z)^2}{\sigma^2_Z} + \frac{(m_{ijmn} - m_H)^2}{\sigma^2_H}
$$

- For the corresponding $\sigma$ are the WA width was taken $\sigma^2_{H,Z}$
- The used chi2 is favouring the reconstruction of the Z boson which is coming from the Higgs decay
Reconstructed invariant masses
Jet transitions
Discussion

- The chi2 which was used is favoring the reconstruction of the Z boson which is coming from the decay of the Higgs boson – refinement of the chi2 is needed
- The cross-section of the other Higgs decays is nearly two orders of magnitude higher
- Fast simulation does not contain btagging tools to reduce H-bb background
- Severe other backgrounds especially 4f_WW_cuxx
- Using the fast simulation using this channel would lead to severe underestimation of the rel. statistical uncertainty
- At this point suggest to use the extrapolation that is currently used for the white paper