Using machine learning to constrain the Higgs total width

1910.02082

Phil Harris, Dylan Rankin (MIT)
Cristina Mantilla Suarez (Johns Hopkins)

ML4Jets 2020

January 17th, 2020
Introduction

• Higgs properties are crucial test of SM

• Inclusive Higgs cross section and Higgs width are poorly constrained

• Difficult measurement at hadron collider

• New proposed method along with machine learning can make competitive measurement possible
Constraints - LHC

- Peak width
  - Using \( h \to \gamma\gamma/ZZ^* \)
  - Limited by detector resolution
- \( \Gamma_H < 1.1 \text{ GeV} (270 \times \Gamma_{SM}) \)
- \( HL-LHC: \delta \Gamma_H/\Gamma_{SM} \approx 80 \)
Constraints - LHC

- Off-shell measurement

\[
\mu = \frac{\sigma_{\text{on-shell}}^{gg\to H\to ZZ^*}}{\sigma_{\text{off-shell}}^{gg\to H^*\to ZZ}} \sim \frac{g_{ggH}^2 g_{HZZ}^2}{m_H \Gamma_H} \]

- Assumption: \(g_{ggH}^2 g_{HZZ}^2\) is the same on and off-shell

- \(\Gamma_H < 9.1\text{ MeV } (2.2 \times \Gamma_{\text{SM}})\)

- **HL-LHC**: \(\delta \Gamma_H/\Gamma_{\text{SM}} \approx 0.25-0.5\)
**Constraints - Future**

- Muon collider
  - $\delta \Gamma_H/\Gamma_{SM} \approx 0.05$

- Electron collider: recoil method
  - $\delta \Gamma_H/\Gamma_{SM} \approx 0.1$ [e.g. ILC]
  
  1) **Measure** $\sigma(e^+e^- \to Zh) \propto g_{hZZ}^2$

  2) **Measure** $h \to ZZ$ decay

\[
\sigma_{Zh\to XX} \propto \frac{g_{hZZ}^2 g_{hXX}^2}{\Gamma_h} \quad \Rightarrow \quad \Gamma_h \propto \frac{[\sigma(e^+e^- \to Zh)]^2}{\sigma(e^+e^- \to Zh) / \sigma(e^+e^- \to Zh \to ZZ)}
\]
\( \Gamma_h \propto LHC \)

- Similar method to recoil at electron collider
- Use Higgs + 1 jet instead of Zh
  - 1) Measure inclusive cross section from reconstructed \( m_h \)
  - 2) Complement with other measurements
    - Boosted \( h \rightarrow bb \)
    - \( W + h \rightarrow bb \)
    - \( W + h \rightarrow WW \)

\[
\Gamma_h \propto \frac{1}{\sigma(W + h \rightarrow WW)} \times \left( \frac{\sigma(gg \rightarrow h)}{\sigma(gg \rightarrow h \rightarrow bb)} \times \frac{\sigma(W + h \rightarrow bb)}{\sigma(gg \rightarrow h \rightarrow bb)} \right)^2
\]
\[ \Gamma_h \propto \frac{1}{\sigma(W + h \rightarrow WW)} \times \left( \frac{\sigma(gg \rightarrow h) \times \sigma(W + h \rightarrow bb)}{\sigma(gg \rightarrow h \rightarrow bb)} \right)^2 \]

\[ \frac{\delta \sigma}{\sigma_{SM}} = 0.05 \]

\[ \frac{\delta \sigma}{\sigma_{SM}} = 0.09 \]

\[ \frac{\delta \sigma}{\sigma_{SM}} = 0.25 \times \delta \sigma_{gg \rightarrow h} \]
Γ_\text{h} \propto \frac{1}{\sigma(W + h \rightarrow WW)} \times \left( \frac{\sigma(gg \rightarrow h) \times \frac{\sigma(W + h \rightarrow bb)}{\sigma(gg \rightarrow h \rightarrow bb)}}{2} \right)

Can we measure inclusive Higgs cross section? 
\( \delta\sigma/\sigma_{\text{SM}} = ? \)

This is the challenge we need to solve
Measurement Strategy

- Require high $p_T$ jet
- Tag/select Higgs jet
- Fit mass
Measurement Strategy

- Require high $p_T$ jet
  - Decay products contained in fat-jet
- Tag/select Higgs jet
  - Reduce background without bias
- Fit mass
  - Extract total cross section
Measurement Strategy

- Require high $p_T$ jet
  - Decay products contained in fat-jet
- Tag/select Higgs jet
  - Reduce background without bias
- Fit mass
  - Extract total cross section
Higgs Tagging

- Dominant background with minimal jet selection is QCD

- Can existing tools tag all Higgs decays simultaneously?

<table>
<thead>
<tr>
<th>Higgs Decay</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>H(bb)</td>
<td>60%</td>
</tr>
<tr>
<td>H(cc)</td>
<td>2.8%</td>
</tr>
<tr>
<td>H(WW)</td>
<td>24%</td>
</tr>
<tr>
<td>H(ZZ)</td>
<td>2.6%</td>
</tr>
<tr>
<td>H(\tau\tau)</td>
<td>6%</td>
</tr>
<tr>
<td>H(gg)</td>
<td>8%</td>
</tr>
</tbody>
</table>

QCD
Higgs Tagging

- $\tau_{21}$ not optimal for $WW/ZZ/gg$
- Useful as a benchmark

<table>
<thead>
<tr>
<th>Higgs</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>H(bb)</td>
<td>60%</td>
</tr>
<tr>
<td>H(cc)</td>
<td>2.8%</td>
</tr>
<tr>
<td>H(WW)</td>
<td>24%</td>
</tr>
<tr>
<td>H(ZZ)</td>
<td>2.6%</td>
</tr>
<tr>
<td>H(\tau\tau)</td>
<td>6%</td>
</tr>
<tr>
<td>H(gg)</td>
<td>8%</td>
</tr>
</tbody>
</table>

QCD
Benchmark Comparison

- Jet $p_T > 500$ GeV
- Using $\tau_{21}^{DDT}$ for Higgs tagging
  - $\Gamma_h < 7.8 \times \Gamma_{SM} = 32$ MeV
  - equivalently $\delta\sigma = 3.6 \times \sigma_{SM}$
- If lowered to jet $p_T > 400$ GeV
  - $\Gamma_h < 19$ MeV ($\delta\sigma = 2.2 \times \sigma_{SM}$)

$\Gamma_h \lesssim 10$ MeV from off-shell method
Can We Improve?

- Jet selection/mass reconstruction
- Improved tagger
Jet Selection/Mass (MET)

- Missing energy spreads/shifts jet mass peak
  - Simple MLP for regression of true MET (input: jet variables)
- Highest $p_T$ jet not always Higgs jet
  - Especially WW, $\tau\tau$
- Select jet using $p_T$ of jet+MET$_{\text{reg}}$ system
ML is a great tool for tagging Higgs jets (ex. $h \to bb$)
Higgs Tagging

- Design network to simultaneously tag all Higgs decays
- GRU with 20 particle inputs ($p_T$, $\eta$, $\phi$, ID)
Higgs Tagging

- Design network to simultaneously tag all Higgs decays
- GRU with 20 particle inputs ($p_T$, $\eta$, $\phi$, ID)
Higgs Tagging

- Ideal case is uniform performance across decays
  - Avoid vertexing
- Similar performance excluding $h \rightarrow \tau\tau$ and $h \rightarrow gg$
Higgs Tagging - gg

- Particle content of $h \rightarrow gg$ jet similar to QCD jet
- Dedicated GRU struggles to differentiate
- MLP with expert features - jet mass ratios
- Similar concept to **collinear drop** [1907.11107], isolates color singlet
Higgs Tagging - gg

- Particle content of $h \rightarrow gg$ jet similar to QCD jet
  - Dedicated GRU struggles to differentiate
- MLP with expert features - jet mass ratios
  - Similar concept to collinear drop [1907.11107], isolates color singlet

![Signal efficiency vs. background efficiency graph](image-url)
Sensitivity

**GRU\textsuperscript{DDT}**

- **GRU\textsuperscript{DDT}**
Results

\( \delta \sigma_{\text{inc.}} / \text{BR at 3 ab}^{-1} \)

- **dim4 poly fit**
- **template fit**
- **no syst**

- CMS Offshell 3 ab\(^{-1}\)
- ATLAS Offshell 3 ab\(^{-1}\)
- \( e^+e^- \) proj.
Results
Results

\[ \delta \sigma_{\text{inc.}} / \text{BR at 3 ab}^{-1} \]

- CMS Offshell 3 ab\(^{-1}\)
- ATLAS Offshell 3 ab\(^{-1}\)
- e\(^+\)e\(^-\) proj.

Classes:
- Inclusive
- \(\tau_2/\tau_1^{\text{DDT}}\) (6%)
- Mass DNN\(^{\text{DDT}}\) (10%)
- GRU\(^{\text{DDT}}\) + Mass DNN\(^{\text{DDT}}\) (10%)
- GRU\(^{\text{DDT}}\) (1%)
- GRU (1%)

ML4Jets 2020

12/17/2020
Results - GRU+DNN

- Tight GRU + DNN (1%) is good overall: benchmark
Results - GRU+DNN

- Tight GRU + DNN (1%) is good overall: benchmark

\[ \delta \sigma_{\text{inc}} / \sigma_{\text{SM}} \approx 0.14 \]
\[ \delta \Gamma_h / \Gamma_{\text{SM}} \approx 0.33 \]
\[ \delta \Gamma_h \approx 1.4 \text{ MeV} \]
Biases/Assumptions

• Experimental challenges:
  • Jet $p_T > 400$ GeV
  • Fit convergence, necessary polynomial order
  • Signal efficiency (uncertainty and uniformity)

• Method assumptions:
  • Higgs decays can be reconstructed in single jet (semi-vis/invis?)
Semi-Visible/Invisible

- SM semi-visible decays included in study (WW, $\tau\tau$)
- MET regression, MET + jet selection could be improved
- Further categorization possible using MET for fully invisible

Projections for 3 ab$^{-1}$: ($h\rightarrow\tau\tau$)

- $\frac{\delta\sigma_{\text{inc}}}{\sigma_{\text{SM}}} \sim 0.25$
- $\frac{\delta\Gamma_h}{\Gamma_{\text{SM}}} \sim 0.55$
- $\delta\Gamma_h \sim 2.3$ MeV

$\sigma_{\text{inv}} < 5\% \sigma_{\text{SM}}$

$\sigma < 14\% \sigma_{\text{SM}}$

Extended categorization
Other Signatures

- $h \rightarrow gg$ difficult to handle
  - Some dedicated techniques show promise (mass DNN)
- Exotic decays
  - Prompt caught by tagger
  - LLP possible with further categories?

### Projections for 3 ab$^{-1}$: ($h \rightarrow gg$)

- $\delta \sigma_{\text{inc}}/\sigma_{\text{SM}} \sim 0.41$
- $\delta \Gamma_h/\Gamma_{\text{SM}} \sim 0.83$
- $\delta \Gamma_h \sim 3.5$ MeV

---

**Features (e.x. Long-Lived)**

- $h \rightarrow \tau\tau$
- $h \rightarrow WW^*$
- $h \rightarrow \text{prompt}$
  - $h \rightarrow c\bar{c}$
  - $h \rightarrow b\bar{b}$
- $h \rightarrow \text{LLPs}$

**This paper**

- Improved results through categorization
- Missing energy
Improvements

• Improved tagger?
• Better performance
  • Alternate architectures (VAE? Normalizing flow?)
• Stability across decays
  • Investigated use of adversarial network (large reduction in performance)
• New ideas (from LHC olympics?) can have significant impact here
Conclusions

• Proposed method to measure Higgs width in high $p_T$ events at LHC

• Constraints comparable to off-shell method possible
  • $\delta \Gamma_h \sim 1.4 - 3.5$ MeV
  • Alternative measurements important

• Method has much room to grow, improvements possible

• Welcome suggestions/thoughts
Benchmark Comparison

- Selection using hadronic triggers
  - Jet $p_T > 500$ GeV, Higgs candidate is highest $p_T$ jet
  - Using $\tau_{21}^{DDT}$ for Higgs tagging
  - Fit jet mass using same procedure as boosted $h \rightarrow bb$ (Bernstein polynomial, predict multijet from failing data)

\[ \Gamma_h < 7.8 \times \Gamma_{SM} = 32 \text{ MeV} \]

equivalently $\delta \sigma = 3.6 \times \sigma_{SM}$

- If lowered to jet $p_T > 400$ GeV

\[ \Gamma_h < 19 \text{ MeV} \ (\delta \sigma = 2.2 \times \sigma_{SM}) \]

$\Gamma_h \lesssim 10 \text{ MeV}$ from off-shell method
Benchmark @ 3000 fb\(^{-1}\)

- This method: \(\delta \sigma = 0.55 \times \sigma_{SM}\)
  - Without systematics: \(\delta \sigma = 0.24 \times \sigma_{SM}\)

- Off-shell projections:
  - CMS: \(\delta \sigma = 0.09 \times \sigma_{SM}\)
  - ATLAS: \(\delta \sigma = 0.17 \times \sigma_{SM}\)

- Electron collider projections
  - \(\delta \sigma = 0.01 \times \sigma_{SM}\)
Sensitivity

$\tau_{21}^{DDT}$
Sensitivity

$\text{GRU}^{\text{DDT}} + \text{Mass DNN}^{\text{DDT}}$

- $\text{GRU}^{\text{DDT}} + \text{Mass DNN}^{\text{DDT}}$ (1%)
Fit Uncertainty

\[ \sigma_{\text{xs}/\text{BR}} \]

- **organic ddt+massratios (1%)**
- **Uncertainty Band (pol4)**

No Systematics
Template
0-pars
1-pars
2-pars
3-pars
4-pars
5-pars
6-pars
• Inclusive fits do not converge
Results - GRU

- GRU has significant mass sculpting, fit is problematic
• Loose GRU + DNN (10%) able to select $h \rightarrow gg$ decays
• Tight GRU + DNN (1%) is good overall: benchmark