Searching for HH-3lepton @ ATLAS

Yulei_Zhang
Shanghai Jiao Tong University

Di-Higgs 2020 Workshop @ Shanghai
November 20th, 2020

Outline

- Introduction
- Light Lepton channel ($3\ell 0\tau_h$)
  - Selection
  - Fake Factor Method
  - Higgsness
  - BDT Study & Sensitivity
- Summary

0 τ_h channels	$2\ell SS$ $2\ell OS$ **3ℓ** 4ℓ
τ_h channels	$1\ell 1\tau_h$ $1\ell 2\tau_h$ $1\ell 3\tau_h$ $2\ell 1\tau_h$ $2\ell 2\tau_h$ $3\ell 1\tau_h$
$\gamma\gamma + ML(\tau_h)$	$\gamma\gamma + 2\ell(e,\mu,\tau)$ $\gamma\gamma + 1\ell(e,\mu,\tau)$
$HH \to bbZZ \to bb + 4\ell$ $X \to SH \to ML$	

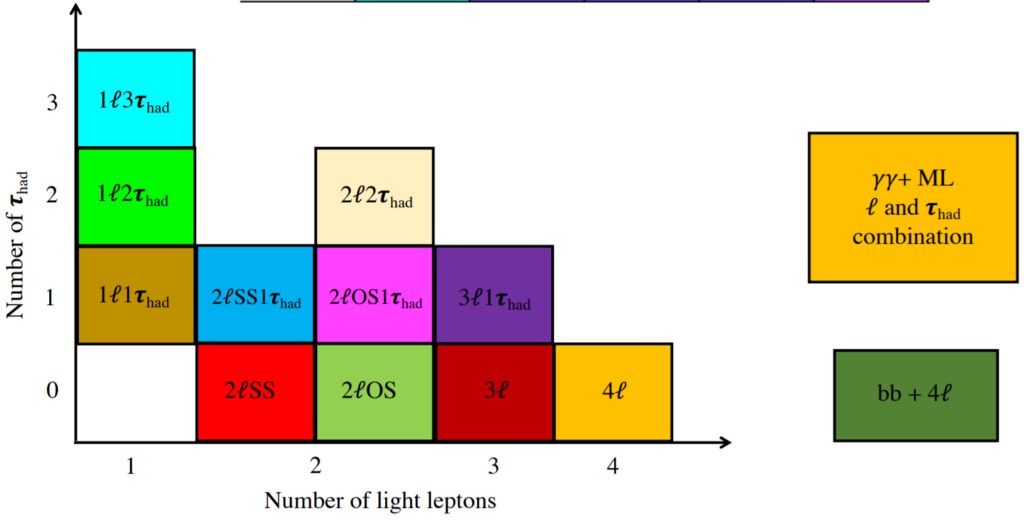
Red means mainly contributed channel

Introduction

- Search for Higgs pair production in the multilepton final state, focusing on non-resonant HH production and $X \rightarrow SH$ final states.
- Several HH decay modes with small BRs. Many of these not covered by dedicated analyses.
- Hard to fully reconstruct HH system
- ~ 5% of total branching ratio
- TWiki page

Katharine Leney & Marcus Morgenstern

	bb	WW	ττ	ZZ	γγ
bb	33%				
WW	25%	4.6%			
ττ	7.4%	2.5%	0.39%		
ZZ	3.1%	1.2%	0.34%	0.076%	
γγ	0.26%	0.10%	0.029%	0.013%	0.0005%



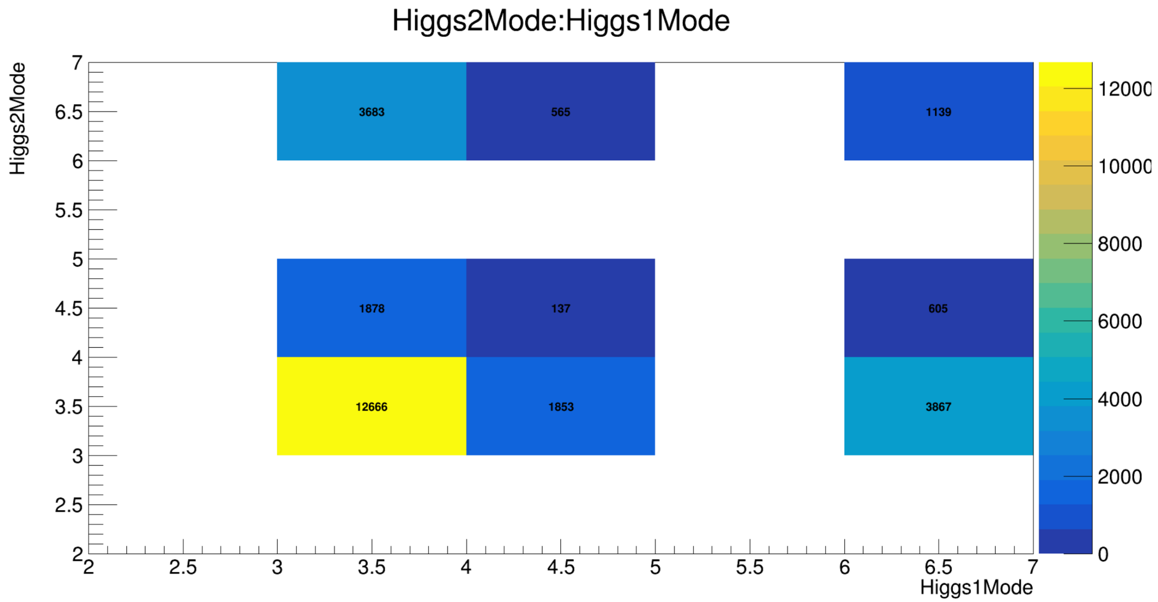
$HH \rightarrow 3\ell + \text{jets}$

Light Lepton channels ($3\ell 0\tau_h$)

Final State: $3\ell 0\tau_h$

- $3 = WW$
- $4 = ZZ$
- $6 = \tau\tau$

- non-resonant HH production (SM)
- $HH \rightarrow WW^*WW^* \rightarrow 3\ell + \text{jets}$ is the dominant signal process.
- Total cross section: 0.4180 fb @ 13 TeV
- Lepton label:
  - ℓ_1: the opposite sign lepton
  - ℓ_2: the lepton closer to ℓ_1 (the smaller ΔR)
  - ℓ_3: the remaining lepton
- Previous $HH \rightarrow 4W$ paper $(36\ \text{fb}^{-1})$



$$\frac{\sigma(HH\rightarrow 4W\rightarrow 3\ell)}{\sigma(HH\rightarrow (WW,\ ZZ,\tau\tau)\rightarrow 3\ell)} = 54.6\%$$

Selection

- Three lepton with total charge of ± 1
  - One opposite sign lepton(ℓ_1): $p_T > 10$ GeV, loose quality
  - Two same sign leptons(ℓ_2, ℓ_3): $p_T > 15$ GeV, tight quality
- Hadronic tau Veto, Z Mass Veto and Low Mass Veto
  - **Low Mass Veto:** invariant mass of any SFOS pair should be larger than 12 GeV
  - **Z Mass Veto:** invariant mass of any SFOS pair should exclude Z Mass region by at least 10 GeV
- b-jets veto and jet multiplicity
  - $N_{jets} == 1$: Control region to calculate fake lepton background in Signal Region
  - $N_{jets} \geq 2$: Signal Region for $3\ell 0\tau_h + jj$
- Select variables with good separation power to train boosted decision tree(BDT)
  - A novel variable called "Higgsness" is introduced.

Sample Consideration

- Main backgrounds are dominated by diboson and fake lepton backgrounds.
  - **Prompt backgrounds:** $VV, t\bar{t}H, t\bar{t}Z, t\bar{t}W, tZ, VVV, VH$
  - **Fake lepton backgrounds:** $V\gamma$, *Jet Fakes* $(t\bar{t}, Z + \text{jets})$
  - Jet Fake backgrounds are estimated by ***Fake Factor Method***.
- Data are included from year 15 to year 18 with luminosity of $139\ \text{fb}^{-1}$ @ $13\ \text{TeV}$.

Fake Factor Method

- A data-driven technique that can be used to estimate both the normalization and shape of the fakes
- Define fake factor regions dominated by fakes
- Numerator: events with *tight leptons*
- Denominator: define an *anti-tight* selection by inverting the identification and isolation of the lepton
  - Very little signal contamination
  - Enhanced in the fakes
- Apply the fake factors to a control region identical to the signal region, except one of the lepton is *anti-tight*

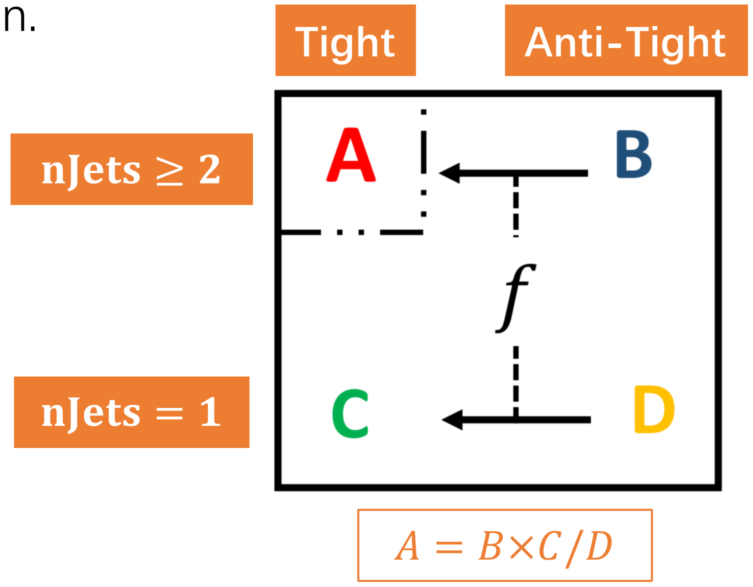
Fake Factor Method (Cont.)

- Prompt Backgrounds: $t\bar{t}H, t\bar{t}Z, t\bar{t}W, tZ, VV, VVV, VH$
- Jet Fakes MC: $t\bar{t}$, Z+jets
- **Anti-Tight Event**: SS pair with one *tight* lepton and one *anti-tight* lepton.
- **Tight Event**: SS pair with two *tight* leptons.

- **Region D** *(CR for calculating FF):* Anti-Tight Event, $N_{jet} = 1$
- **Region C** *(CR for calculating FF):* Tight Event, $N_{jet} = 1$
- **Region B** *(CR for calculating Fakes in SR):* Anti-Tight Event, $N_{jet} > 1$
- **Region A** *(Fakes in SR):* Tight Event, $N_{jet} > 1$
- $f_\ell(i) = \frac{N_\ell^C(i)}{N_\ell^D(i)}$ → $N_{\text{fake}}^{A\,(SR)}(i) = f_\ell(i) \times N_\ell^B(i)$

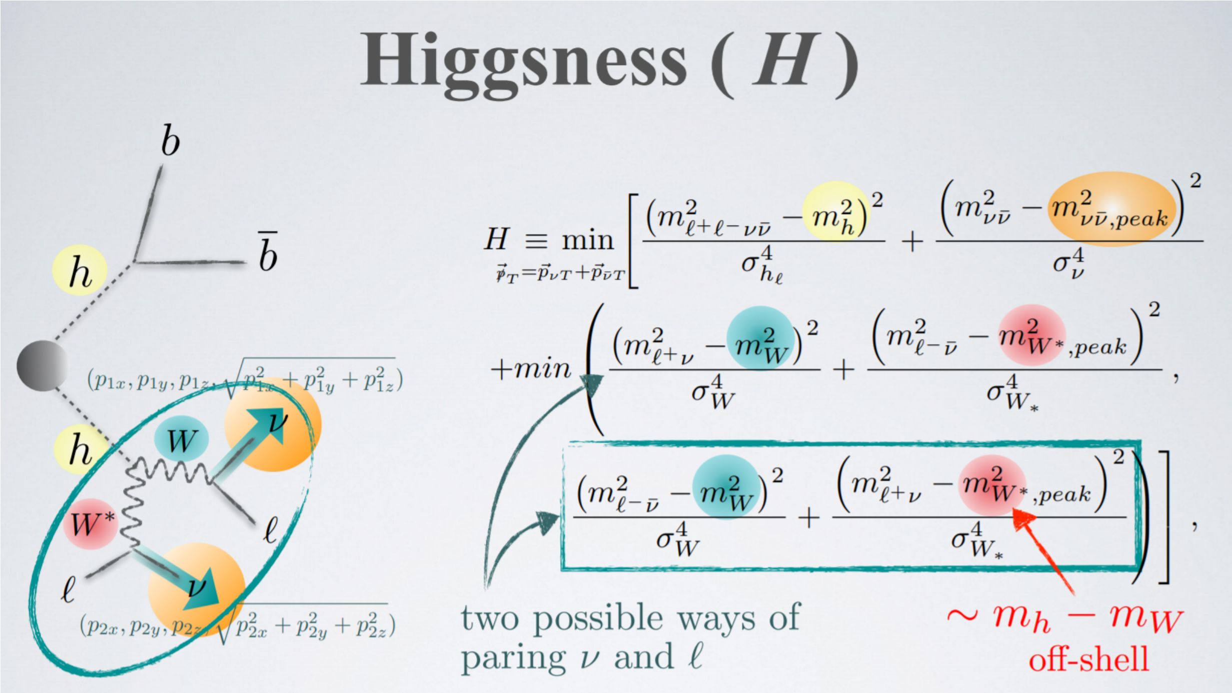
$\ell = e, \mu$ and i refers to p_T bin

p_T-parameterized Data-Driven Fakes



Higgsness

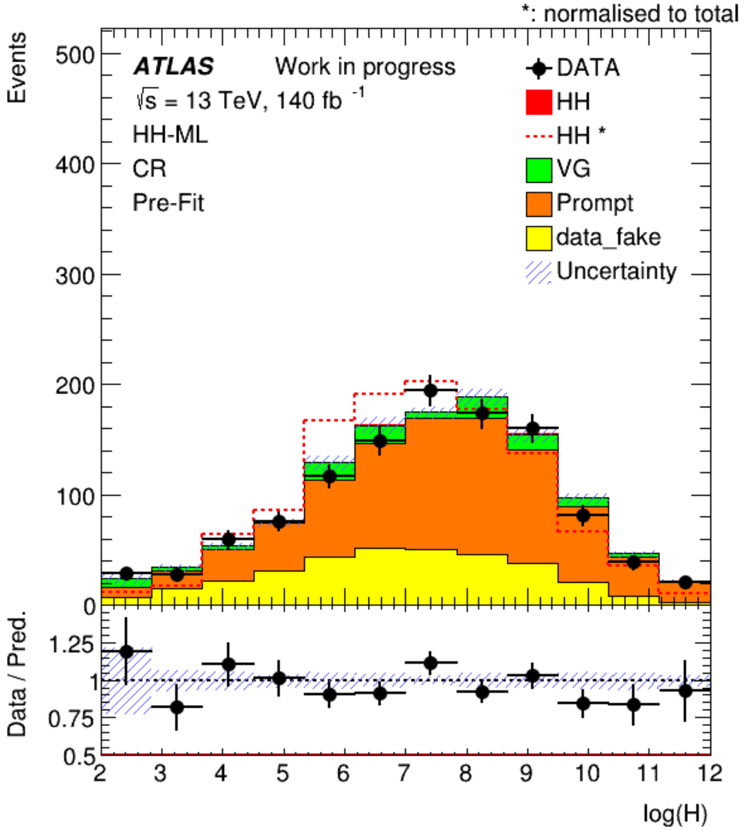
- A novel kinematic variable to estimate if the sample is signal-like.
- The basic idea is to find an optimal combination of the four momentum of ν_0, ν_1 and ν_2 to fit the signal features best.



PRL Paper

First try in 3ℓ channel

- Signal Process: $HH \to WW^*WW^* \to \ell_1\nu_1 + \ell_2\nu_2 + \ell_3\nu_3 + jj$
- 9 constraints: $m_{\text{total}}, m_{H(\ell\nu\ell\nu)}, m_{H(\ell\nu jj)}, m_{\nu\nu\nu}, m_W, m_{W^*}, p_Z^{\ell_3}$ and missing transverse momentum conservation
- $2\times 2 = 4$ combinations of diHiggs:
  - $H_1 \to \ell_1\nu_1 + \ell_2\nu_2\,,\ \ H_2 \to \ell_3\nu_3 + jj$ (either can decay from on-shell or off-shell W)
- $$\text{Higgsness} = \min\left[\frac{(m_{\ell\nu\ell\nu\ell\nu jj} - m_{\text{total}})^2}{\sigma_{\text{total}}^2} + \frac{(m_{\nu\nu\nu} - m_{\nu\nu\nu,peak})^2}{\sigma_{\nu\nu\nu}^2} + \frac{(p_Z^{\ell_3} - 0)^2}{\sigma_{p_Z}^2} + H_{123}\right]$$
- $$H_{123} = \frac{(m_{\ell\nu\ell\nu} - m_{H_{\ell\nu\ell\nu}})^2}{\sigma_{H_{\ell\nu\ell\nu}}^2} + \frac{(m_{\ell\nu jj} - m_{H_{\ell\nu jj}})^2}{\sigma_{H_{\ell\nu jj}}^2} + W_{12} + W_{3jj}$$
- $$W_{12} = \min\left[\frac{(m_{\ell^-\nu} - m_{W,peak})^2}{\sigma_W^2} + \frac{(m_{\ell^+\nu} - m_{W^*,peak})^2}{\sigma_{W^*}^2}, \frac{(m_{\ell^+\nu} - m_{W,peak})^2}{\sigma_W^2} + \frac{(m_{\ell^-\nu} - m_{W^*,peak})^2}{\sigma_{W^*}^2}\right]$$
- $$W_{3jj} = \min\left[\frac{(m_{\ell\nu} - m_{W,peak})^2}{\sigma_W^2} + \frac{(m_{jj} - m_{W^*,peak})^2}{\sigma_{W^*}^2}, \frac{(m_{jj} - m_{W,peak})^2}{\sigma_W^2} + \frac{(m_{\ell\nu} - m_{W^*,peak})^2}{\sigma_{W^*}^2}\right]$$



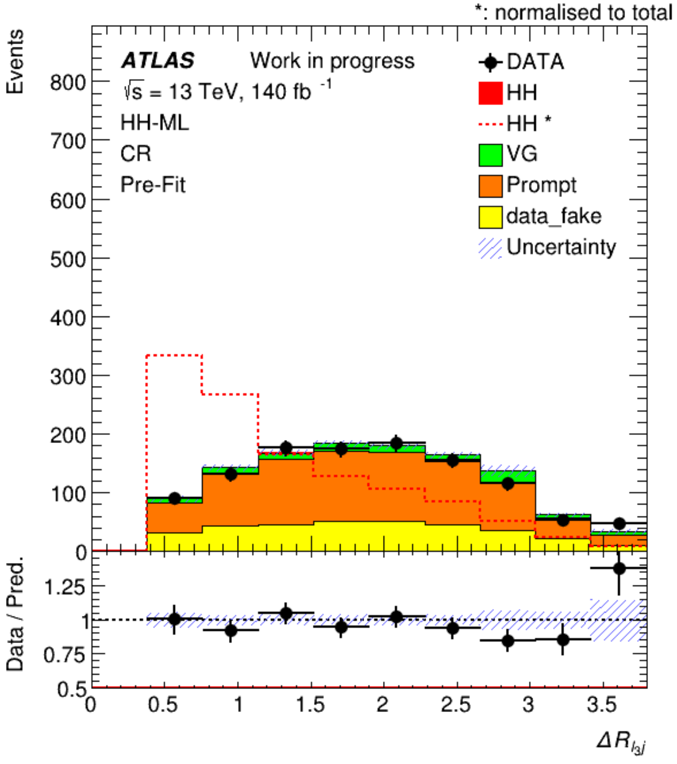
- Shapes are sensitive to kinematic constraints.
- Continuing optimization.

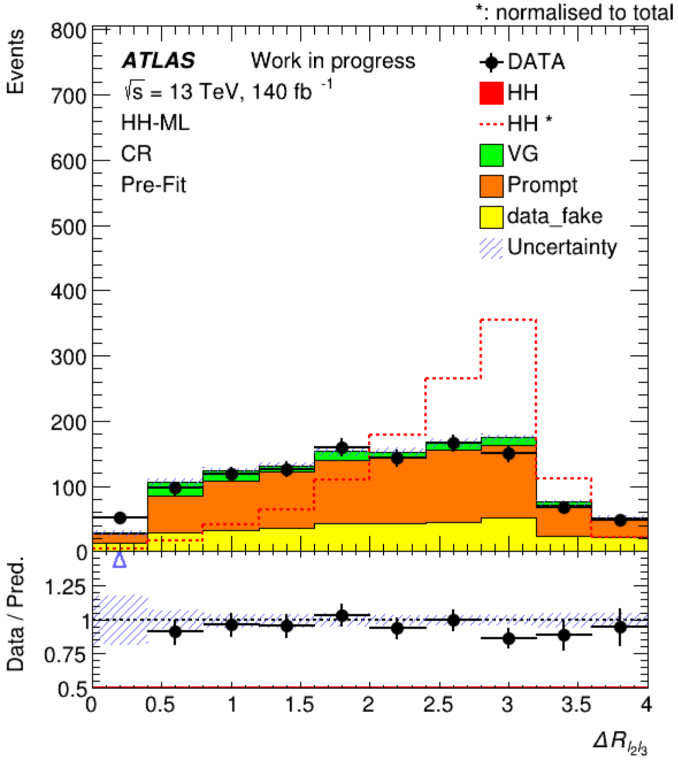
BDT Study

- First try of BDT with semi-Data-Driven Fakes in multilepton group
  - Directly train signal with full backgrounds.
  - Shape from MC, normalized by fake factor from orthogonal data
  - Fully Data-Driven fakes under study

- Preliminary expected limit **with statistics uncertainty only**:
  - Preliminary results with relatively low statistics for background and signal samples
  - VV and fakes from $t\bar{t}$, $Z + jets$ are major backgrounds
  - BDT optimization ongoing

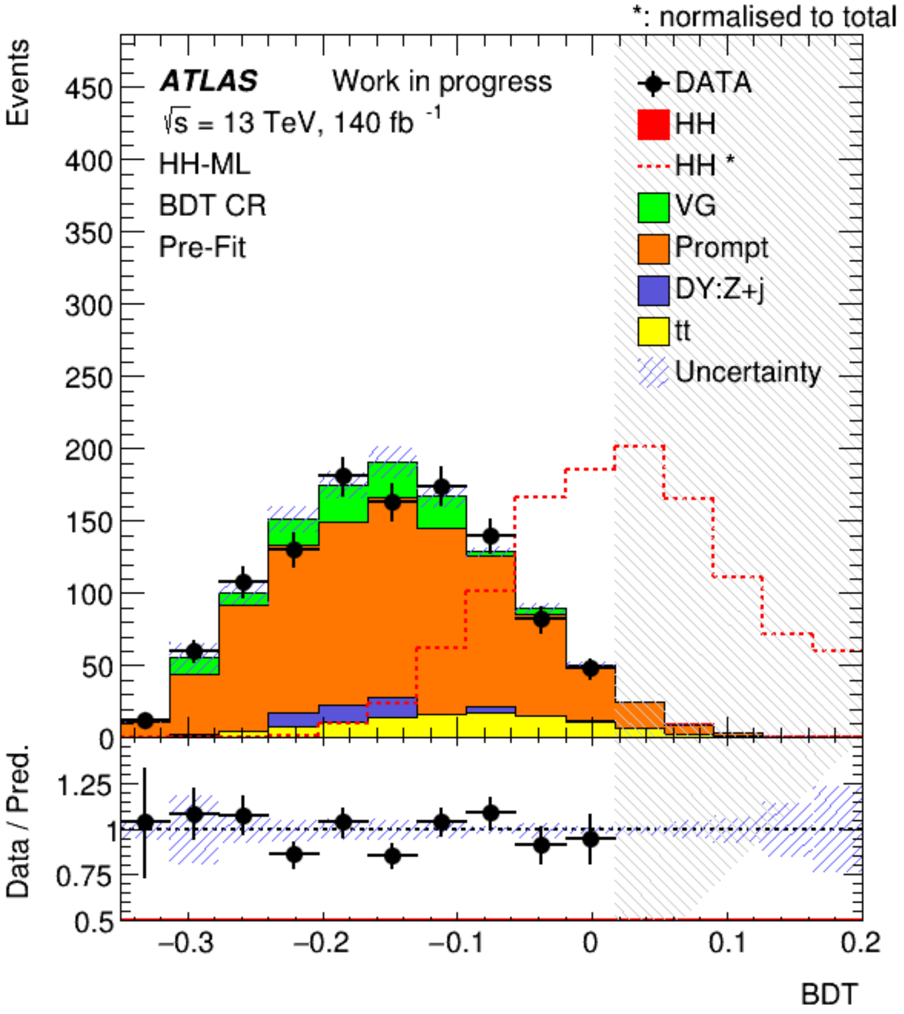
Discriminating variables





- Data/MC agreement reasonable.

Single BDT Result



$CR: BDT < 0.01$
$SR: BDT > 0.01$
Blind region: BDT>0.01

Single BDT without *Higgsness*	Significance	0.096
	Limits	28.09
Single BDT with *Higgsness*	Significance	0.102
	Limits	27.54

- Expected Limits and Significance with stats only.

Summary

- Non-resonant Higgs pair to 3-lepton channel has been studied with full Run 2 data.
- Fake factor method has been applied to the estimation of fake lepton backgrounds.
- BDT is introduced in multilepton group for the first time and shows ~50% improvement in sensitivity (**statistics only**).
- Higgsness variable has shown some improvements as well. More study and optimization are needed.

Thanks!