Tagging Jets in Invisible Higgs Searches

[1712.03973]

Anke Biekötter

Heidelberg University

with Fabian Keilbach, Rhea Moutafis, Tilman Plehn

and Jennifer Thompson

VBScan Krakow, October 23, 2018

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Outline

- Introduction: Signatures of invisible Higgs decays
- Weak boson fusion and its backgrounds

- Comparison to associated Zh production

- Conclusion and discussion

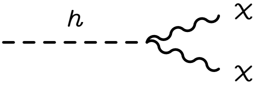
Outline

- Introduction: Signatures of invisible Higgs decays
- Weak boson fusion and its backgrounds
- **Quark gluon discrimination**
- **BDT analysis**
- Comparison to associated Zh production
- **low-level quark gluon discrimination using deep learning**
- Conclusion and **discussion**

Introduction

Motivation

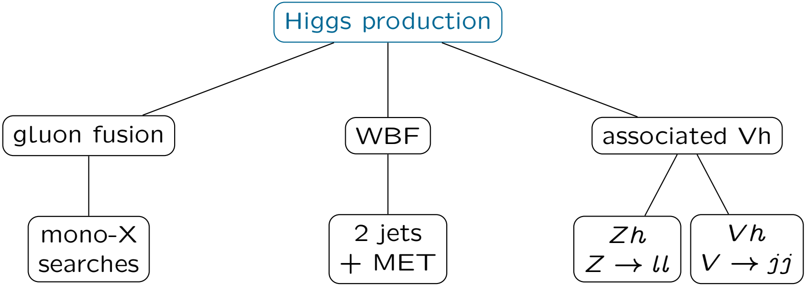
- Higgs decays to invisible particles
  - [Shrock, Suzuki, 1982]
- Higgs portal models
  - [Silveira, Zee, 1985]
  - [Burgess, Pospelov, Veldhuis, 2001]
  - [Patt, Wilczek, 2006]
  - [Englert, Plehn, Zerwas, Zerwas, 2011]
- Dark matter candidates
  - Scalar (minimal/extended Higgs sector)
  - Fermion (NMSSM) [Butter, Murgia, Plehn, Tait, 2016]
  - ...



Invisible Higgs decays

- Higgs production
  - gluon fusion
  - WBF
  - associated Vh

Invisible Higgs decays



Higgs production

gluon fusion

WBF

associated Vh

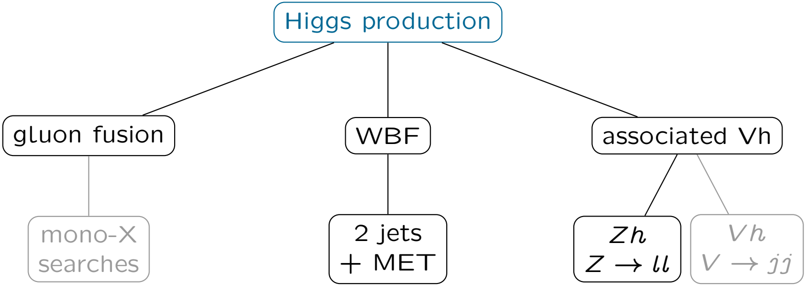
mono-X searches

2 jets + MET

Zh $Z \to ll$

Vh $V \to jj$

Invisible Higgs decays



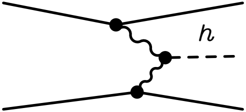
strongest channels

Weak boson fusion

WBF signature

EW process: Jets + missing energy

- 2 jets with large η separation
- opposite hemispheres $\eta_1 \cdot \eta_2 < 0$
- large MET
- no central jet activity



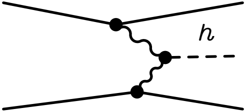
Trigger

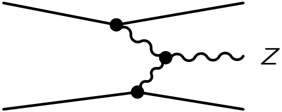
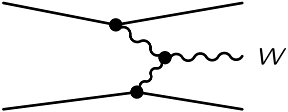
- CMS-HIG-16-016:
    - $p_{T,j} > 40$ GeV
    - $m_{jj} > 600$ GeV
    - $E_T^{\text{miss}} > 140$ GeV
    - $\Delta\eta_{jj} > 3.5$
    - $\eta_{j1} * \eta_{j2} < 0$
- outlook for HL-LHC
    - $E_T^{\text{miss}} > 200$ GeV?
    - . . . ?
    - How dangerous is this?

[Eboli, Zeppenfeld, 2000]

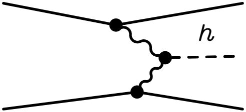
[Bernaciak, Plehn, Schichtel, Tattersall, 2014]

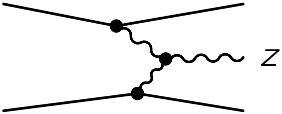
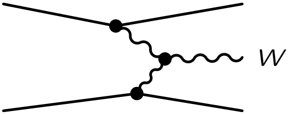
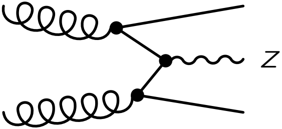
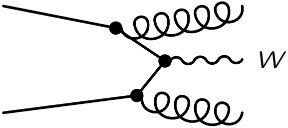
WBF backgrounds



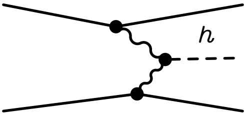
$Z \to \nu\nu$	$W \to (l)\nu$
 Z EW	 W EW

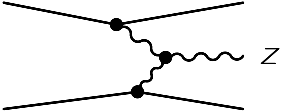
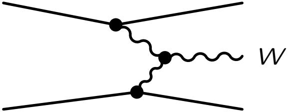
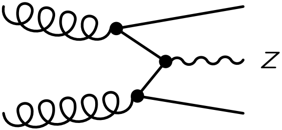
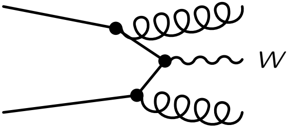
WBF backgrounds



$Z \to \nu\nu$	$W \to (l)\nu$
 Z EW	 W EW
 Z QCD	 W QCD

WBF backgrounds



$Z \to \nu\nu$	$W \to (\ell)\nu$
 Z EW	 W EW losing a lepton
 Z QCD	 W QCD losing a lepton

Tagging jet content

How to suppress QCD backgrounds?

- QCD dominates over EW processes (L**H**C)
- central jet veto

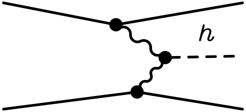
  [Eboli, Zeppenfeld, 2000]

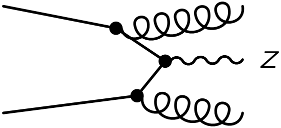
  ➜ analyze third jet, veto additional jets

  [Bernaciak, Plehn, Schichtel, Tattersall, 2014]
- recall: for QCD background tagging jets can be gluons

  ➜ can we use this to suppress QCD backgrounds?

  [AB, Keilbach, Moutafis, Plehn, Thompson, 2017]





Quark gluon discrimination

QCD backgrounds more likely to have hard gluon jets

- wider angle soft emissions
- more splittings in parton evolution

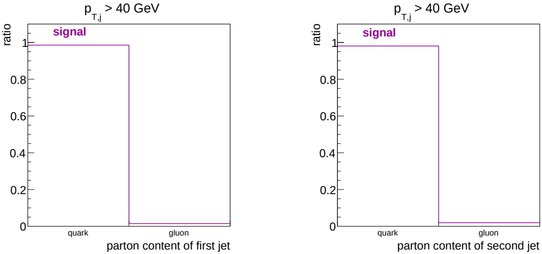
Variables for quark gluon discrimination

- n_{PF}: number of particle flow (PF) objects (Delphes)
- $$w_{\text{PF}} = \frac{\sum_{\text{PF}\in\text{jet}} p_{T,\text{PF}} \Delta R_{\text{PF,jet}}}{\sum_{\text{PF}\in\text{jet}} p_{T,\text{PF}}}$$
- $$C = \frac{\sum_{i_{\text{PF}}, j_{\text{PF}}} \; p_{T,i} p_{T,j} \; (\Delta R_{ij})^{0.2}}{\left(\sum_{i_{\text{PF}}} p_{T,i}\right)^2}$$
- $$p_T D = \frac{\sqrt{\sum_{\text{PF}\in\text{jet}} \; p_{T,\text{PF}}^2}}{\sum_{\text{PF}\in\text{jet}} \; p_{T,\text{PF}}}$$

[ATLAS-CONF-2016-034, CMS-PAS-JME-13-002]

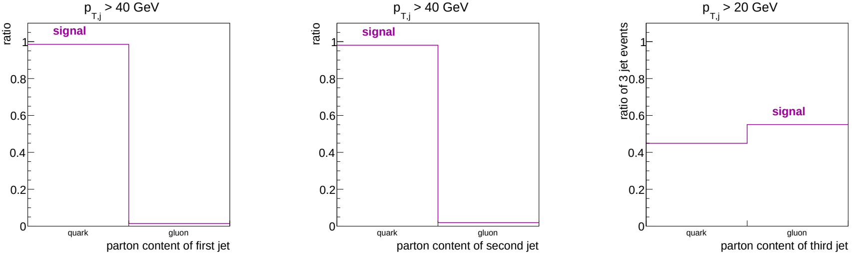
Theoretically not well defined

Parton content in WBF



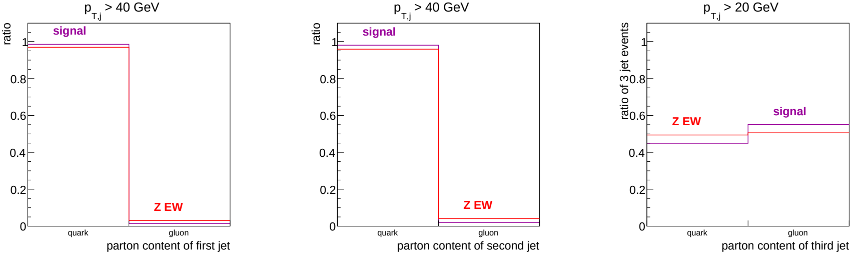
preselection: $p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\text{Lep}} = 0$, $p_T(V) > 80$ GeV

Parton content in WBF



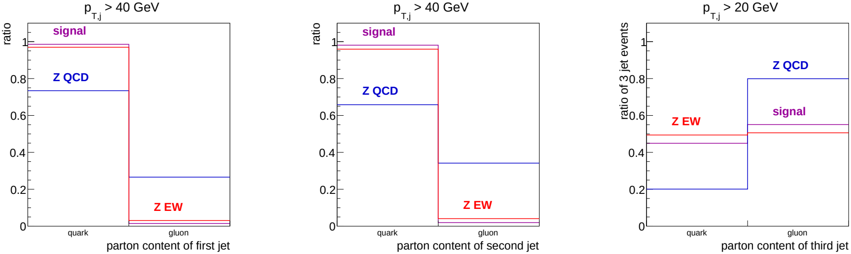
preselection: $p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\text{Lep}} = 0$, $p_T(V) > 80$ GeV

Parton content in WBF



preselection: $p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\text{Lep}} = 0$, $p_T(V) > 80$ GeV

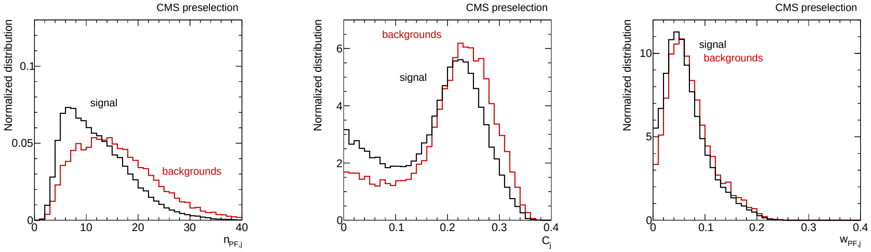
Parton content in WBF



Expect best discrimination power for second jet.

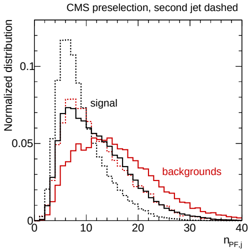
preselection: $p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\text{Lep}} = 0$, $p_T(V) > 80$ GeV

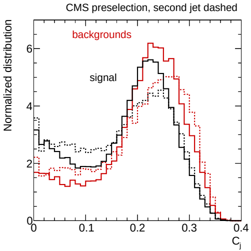
Quark gluon discrimination - distributions

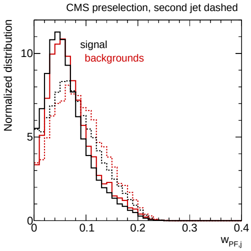


preselection: $p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $E_T^{\text{miss}} > 140$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\text{Lep}} = 0$

Quark gluon discrimination - distributions



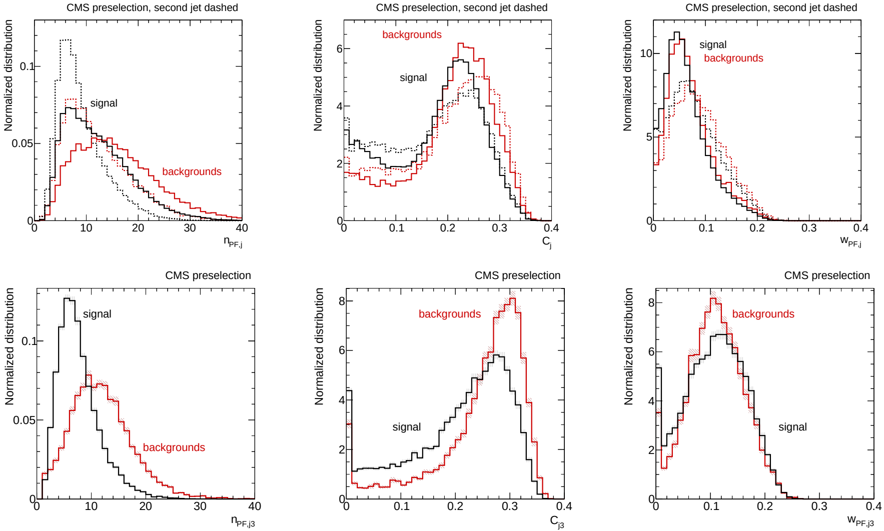




Quark gluon discrimination variables are p_T dependent

preselection: $p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $E_T^{\text{miss}} > 140$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\text{Lep}} = 0$

Quark gluon discrimination - distributions

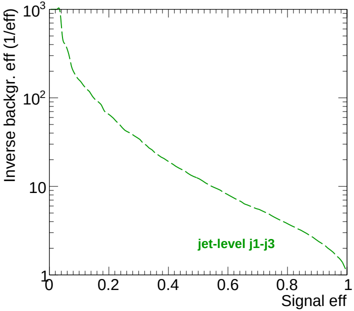


Third jet gives best separation (here: $p_T > 20$ GeV)

preselection: $p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $E_T^{\text{miss}} > 140$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\text{Lep}} = 0$

BDT analysis

Boosted decision tree analysis - TMVA

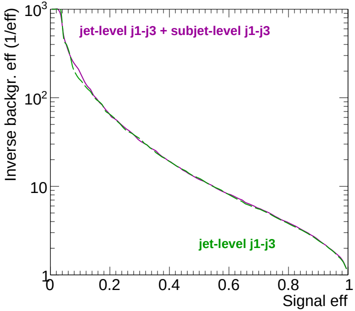


Set	Variables
jet-level j_1, j_2	p_{T,j_1} p_{T,j_2} $\Delta\eta_{jj}$ $\Delta\phi_{jj}$ m_{jj} $\not{E}_T$ $\Delta\phi_{j_1,\not{E}_T}$ $\Delta\phi_{j_2,\not{E}_T}$
subjet-level j_1, j_2	n_{PF,j_1} n_{PF,j_2} C_{j_1} C_{j_2} $p_T D_{j_1}$ $p_T D_{j_2}$
j_3 angular information	$\Delta\eta_{j_1,j_3}$ $\Delta\eta_{j_2,j_3}$ $\Delta\phi_{j_1,j_3}$ $\Delta\phi_{j_2,j_3}$
jet-level j_1-j_3	jet-level j_1, j_2 $+$ j_3 angular information $+$ p_{T,j_3}
subjet-level j_1-j_3	subjet-level j_1, j_2 $+$ n_{PF,j_3} C_{j_3} $p_T D_{j_3}$

preselection:

$p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $E_T^{\text{miss}} > 140$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\text{Lep}} = 0$, $p_{T,j3} > 20$ GeV

Boosted decision tree analysis - TMVA

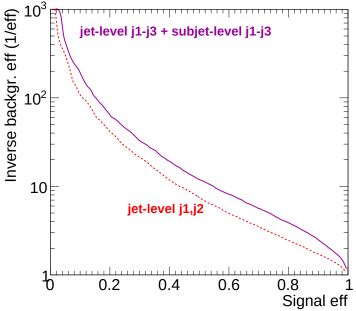


Set	Variables
jet-level j_1, j_2	p_{T,j_1} p_{T,j_2} $\Delta\eta_{jj}$ $\Delta\phi_{jj}$ m_{jj} $\not{E}_T$ $\Delta\phi_{j_1,\not{E}_T}$ $\Delta\phi_{j_2,\not{E}_T}$
subjet-level j_1, j_2	n_{PF,j_1} n_{PF,j_2} C_{j_1} C_{j_2} $p_TD_{j_1}$ $p_TD_{j_2}$
j_3 angular information	$\Delta\eta_{j_1,j_3}$ $\Delta\eta_{j_2,j_3}$ $\Delta\phi_{j_1,j_3}$ $\Delta\phi_{j_2,j_3}$
jet-level j_1-j_3	jet-level j_1, j_2 $+$ j_3 angular information $+$ p_{T,j_3}
subjet-level j_1-j_3	subjet-level j_1, j_2 $+$ n_{PF,j_3} C_{j_3} $p_TD_{j_3}$

preselection:

$p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $E_T^{\mathrm{miss}} > 140$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\mathrm{Lep}} = 0$, $p_{T,j3} > 20$ GeV

Boosted decision tree analysis - TMVA

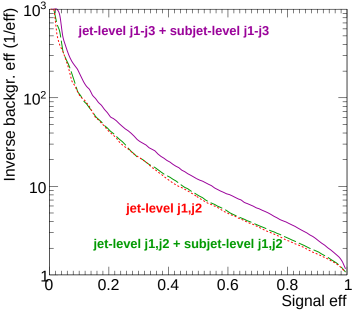


Set	Variables
jet-level j_1, j_2	p_{T,j_1} p_{T,j_2} $\Delta\eta_{jj}$ $\Delta\phi_{jj}$ m_{jj} $\not{E}_T$ $\Delta\phi_{j_1,\not{E}_T}$ $\Delta\phi_{j_2,\not{E}_T}$
subjet-level j_1, j_2	n_{PF,j_1} n_{PF,j_2} C_{j_1} C_{j_2} $p_TD_{j_1}$ $p_TD_{j_2}$
j_3 angular information	$\Delta\eta_{j_1,j_3}$ $\Delta\eta_{j_2,j_3}$ $\Delta\phi_{j_1,j_3}$ $\Delta\phi_{j_2,j_3}$
jet-level j_1-j_3	jet-level j_1, j_2 + j_3 angular information + p_{T,j_3}
subjet-level j_1-j_3	subjet-level j_1, j_2 + n_{PF,j_3} C_{j_3} $p_TD_{j_3}$

preselection:

$p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $E_T^{\mathrm{miss}} > 140$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\mathrm{Lep}} = 0$, $p_{T,j3} > 20$ GeV

Boosted decision tree analysis - TMVA

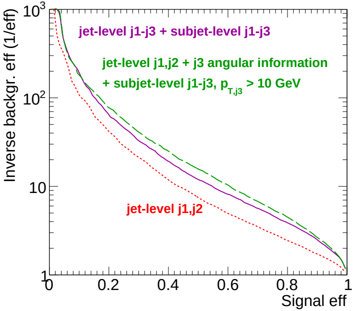


Set	Variables
jet-level j_1, j_2	p_{T,j_1} p_{T,j_2} $\Delta\eta_{jj}$ $\Delta\phi_{jj}$ m_{jj} $\not{E}_T$ $\Delta\phi_{j_1,\not{E}_T}$ $\Delta\phi_{j_2,\not{E}_T}$
subjet-level j_1, j_2	n_{PF,j_1} n_{PF,j_2} C_{j_1} C_{j_2} $p_T D_{j_1}$ $p_T D_{j_2}$
j_3 angular information	$\Delta\eta_{j_1,j_3}$ $\Delta\eta_{j_2,j_3}$ $\Delta\phi_{j_1,j_3}$ $\Delta\phi_{j_2,j_3}$
jet-level j_1-j_3	jet-level j_1, j_2 + j_3 angular information + p_{T,j_3}
subjet-level j_1-j_3	subjet-level j_1, j_2 + n_{PF,j_3} C_{j_3} $p_T D_{j_3}$

preselection:

$p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $E_T^{\mathrm{miss}} > 140$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\mathrm{Lep}} = 0$, $p_{T,j3} > 20$ GeV

Boosted decision tree analysis - TMVA



Set	Variables
jet-level j_1, j_2	$p_{T,j_1}\ p_{T,j_2}\ \Delta\eta_{jj}\ \Delta\phi_{jj}\ m_{jj}\ \not{E}_T\ \Delta\phi_{j_1,\not{E}_T}\ \Delta\phi_{j_2,\not{E}_T}$
subjet-level j_1, j_2	$n_{\text{PF},j_1}\ n_{\text{PF},j_2}\ C_{j_1}\ C_{j_2}\ p_T D_{j_1}\ p_T D_{j_2}$
j_3 angular information	$\Delta\eta_{j_1,j_3}\ \Delta\eta_{j_2,j_3}\ \Delta\phi_{j_1,j_3}\ \Delta\phi_{j_2,j_3}$, $p_{T,j_3} > 10$ GeV
jet-level j_1-j_3	jet-level j_1, j_2 + j_3 angular information + p_{T,j_3}
subjet-level j_1-j_3	subjet-level j_1, j_2 + $n_{\text{PF},j_3}\ C_{j_3}\ p_T D_{j_3}$

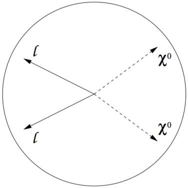
preselection:

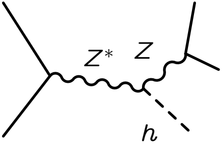
$p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $E_T^{\text{miss}} > 140$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\text{Lep}} = 0$, $p_{T,j3} > 20$ GeV

Associated Zh production

Zh production - signature

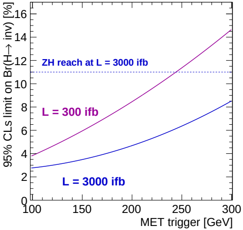
- boosted SFOS leptons $m_{ll} \sim m_Z$ + missing energy
- event generation with Sherpa + Delphes ➜ BDT
- Z+ jets not taken into account (irrelevant at high MET)





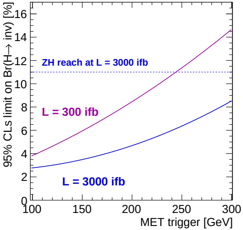
Goal: Compare Zh reach with WBF reach for invisible Higgs searches for different WBF triggers

WBF and Zh reach - triggering



default: $p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $E_T^{\text{miss}} > 140$ GeV

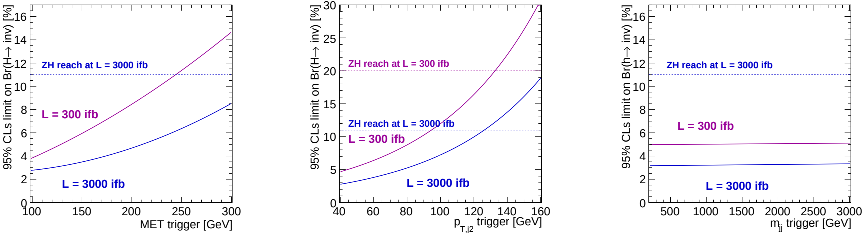
WBF and Zh reach - triggering



WBF constraints stronger for MET trigger $\lesssim$ 350 GeV

default: $p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $E_T^{\text{miss}} > 140$ GeV

WBF and Zh reach - triggering



WBF constraints stronger for MET trigger $\lesssim$ 350 GeV

default: $p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $E_T^{\text{miss}} > 140$ GeV

Deep learning for quark gluon discrimination

Nicholas Kiefer, Jennie Thompson, Tilman Plehn

Quark gluon discrimination - deep learning (1)

Motivation

- discrimination variables like n_{PF} theoretically not well defined (infrared safety)

  ➜ use 4-momenta of jet constituents
- use LoLa architecture developed for `DeepTopLoLa` tagger

  [Butter, Kasieczka, Plehn, Russell]
- tested on pure quarks/gluon samples
- application of mono-jet searches planned

Quark gluon discrimination - deep learning (2)

Combination layer (CoLa) acts on 4-momenta $k_{\mu,i}$ of the jet constituents

$$k_{\mu,i} \xrightarrow{\text{CoLa}} \widetilde{k}_{\mu,j} = k_{\mu,i}\, C_{ij} \qquad \text{with} \quad C = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 & C_{1,N+2} & \cdots & C_{1,M} \\ 1 & 0 & 1 & & \vdots & C_{2,N+2} & \cdots & C_{2,M} \\ \vdots & \vdots & \vdots & \ddots & 0 & \vdots & & \vdots \\ 1 & 0 & 0 & \cdots & 1 & C_{N,N+2} & \cdots & C_{N,M} \end{pmatrix} .$$

Pass CoLa output to a Lorentz layer (LoLa)

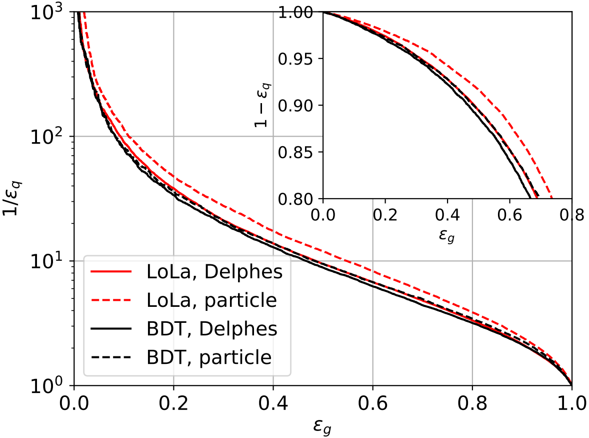
$$\tilde{k}_j \xrightarrow{\text{LoLa}} \hat{k}_j = \begin{pmatrix} m^2(\tilde{k}_j) \\ p_T(\tilde{k}_j) \\ p_T(\tilde{k}_j)\Delta R_{j,\text{jet}} \\ w_{jm}^{(E)}\, E(\tilde{k}_m) \\ w_{jm}^{(d)}\, d_{jm}^2 \\ E_T(\tilde{k}_j) E_T(\tilde{k}_m)(\Delta R_{jm})^{0.2} \end{pmatrix}$$

added for quark gluon discrimination

train deep neural network on GPUs

Quark gluon discrimination - deep learning (3)

Network trained and tested on **pure** quark/gluon samples



- LoLa performs better than standard BDT analysis
- loss of performance when Delphes is included

Conclusions

Conclusions

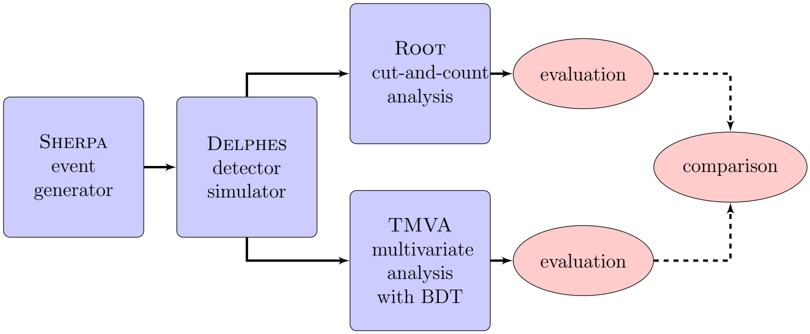
WBF

- Backgrounds: different behavior for N_{jets}
- Useful **quark gluon discrimination** variables: n_{PF}, C
- **Third jet** best for quark gluon discrimination $\mathbf{p_T > 10\ GeV}$
- However, no large improvement by QG variables when full information of additional jets is present (overconstrained)
- **WBF** will still provide strongest constraints after trigger update
- future: LoLa approach has improved sensitivity

Thank you for your attention!

Backup

Tool chain



SHERPA event generator

DELPHES detector simulator

ROOT cut-and-count analysis

TMVA multivariate analysis with BDT

evaluation

evaluation

comparison

BDT settings

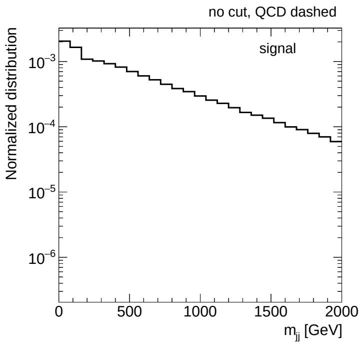
Use TMVA with

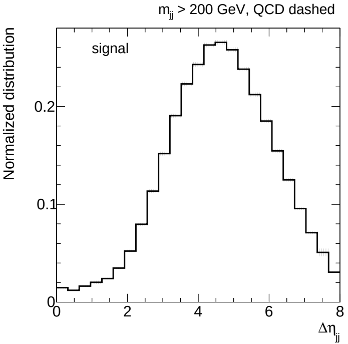
- 70 trees
- 3 layers
- nCuts = 20
- minimum node size 5 %
- preselection

WBF distributions

Sherpa + Delphes,

merged sample (2 + 3) jets

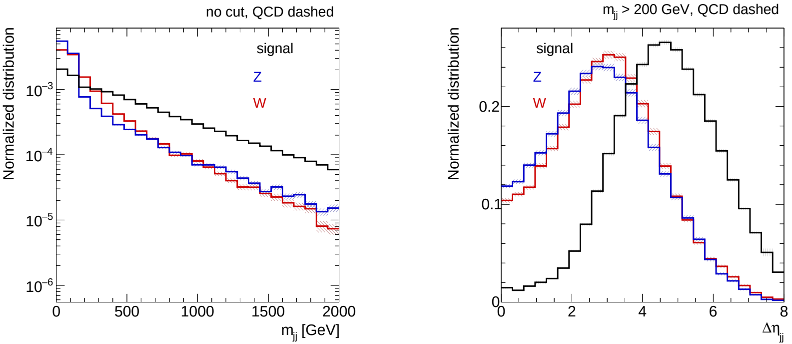




WBF distributions

Sherpa + Delphes,

merged sample (2 + 3) jets

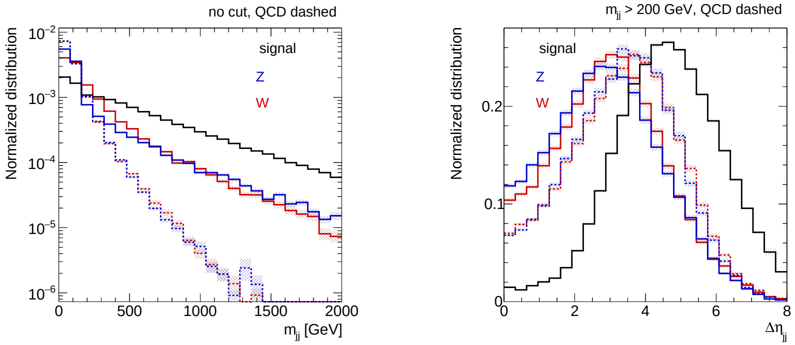


W and *Z* backgrounds similar in signal region

WBF distributions

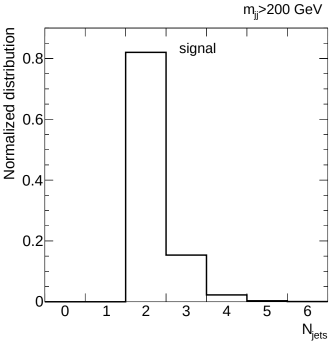
Sherpa + Delphes,

merged sample (2 + 3) jets



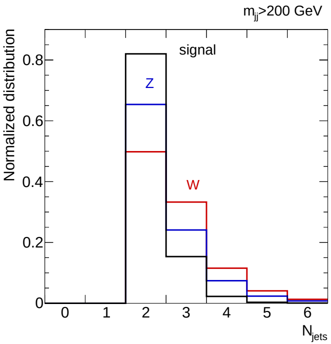
WBF distributions - N_{jets}

merged sample (2 + 3) jets



WBF distributions - N_{jets}

merged sample (2 + 3) jets

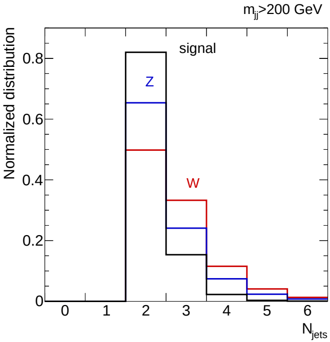


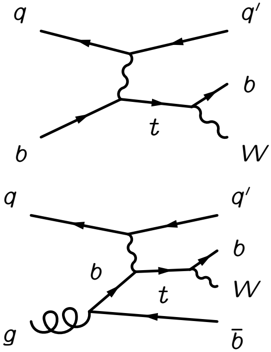
W and *Z* backgrounds different for N_{jets} distribution

- *W* background has more 3-jet events

WBF distributions - N_{jets}

merged sample (2 + 3) jets





W and Z backgrounds different for N_{jets} distribution

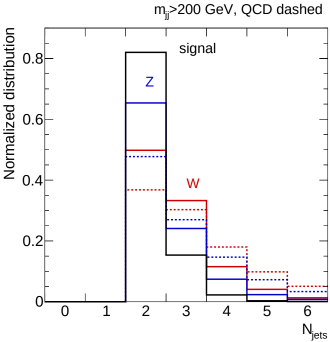
- W background has more 3-jet events
- W background contains **single-top** events

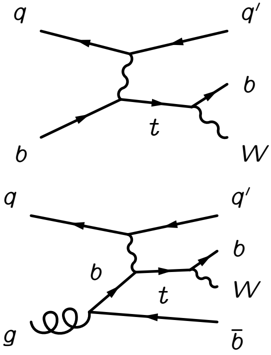
($m_{jj} > 200$ GeV: 30% 2jet, 50% 3jet; preselection: 5%, 12%)

preselection: $p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $\Delta\eta_{jj} > 3.5$, $p_T(V) > 80$ GeV

WBF distributions - N_{jets}

merged sample (2 + 3) jets





W and Z backgrounds different for N_{jets} distribution

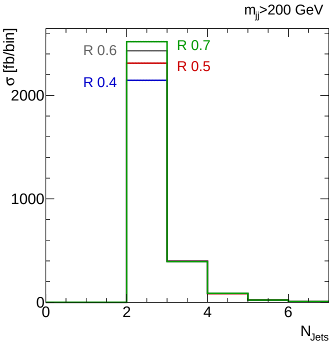
- W background has more 3-jet events
- W background contains **single-top** events

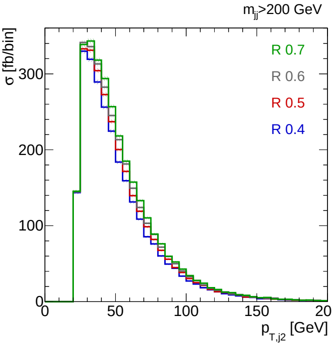
($m_{jj} > 200$ GeV: 30% 2jet, 50% 3jet; preselection: 5%, 12%)

WBF - dependence on jet cone size

Simulated process: $h + 2/3$ jets merged (Sherpa, parton shower)

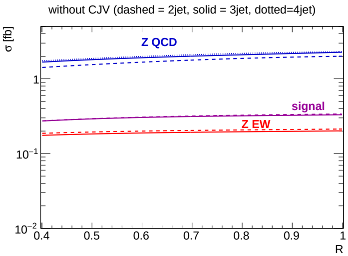
variation of jet cone size in Delphes

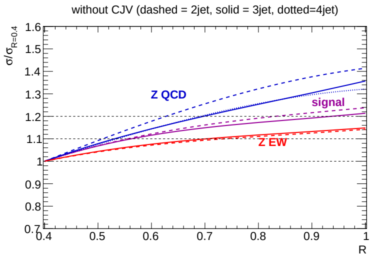




kinematics unchanged

WBF - dependence on jet cone size (2)


without CJV (dashed = 2jet, solid = 3jet, dotted=4jet)

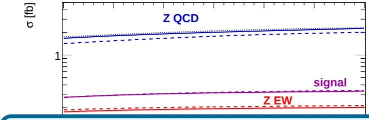

without CJV (dashed = 2jet, solid = 3jet, dotted=4jet)

Signal grows stronger with R than EW background

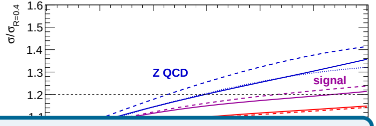
preselection: $p_{T,j} > 40$ GeV, $m_{jj} > 600$ GeV, $\Delta\eta_{jj} > 3.5$, $N_{\text{Lep}} = 0$, $p_T(V) > 80$ GeV

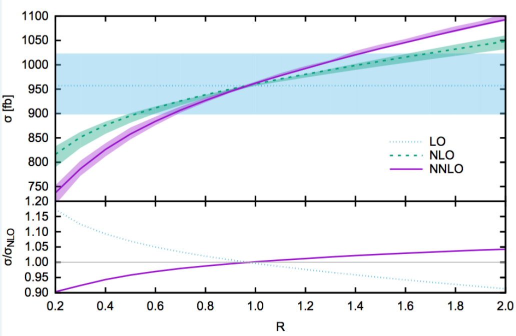
WBF - dependence on jet cone size (2)

without CJV (dashed = 2jet, solid = 3jet, dotted=4jet)



without CJV (dashed = 2jet, solid = 3jet, dotted=4jet)





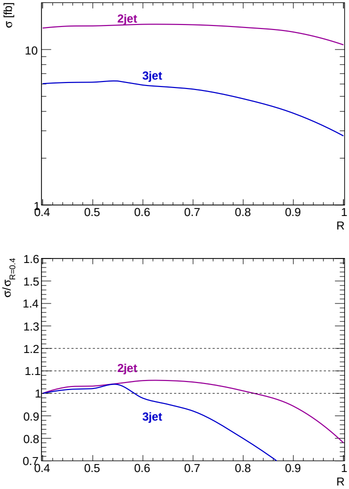
similar results in fixed-order calculation [Rauch, Zeppenfeld, 2017]

Dependence on jet cone size - hZ, Z $\rightarrow$ j j

same final state,

different topology

variable	cut
MET	$120 - 160$ GeV
N_{jets}	$2 - 3$
ΔR_{jj}	$0.7 - 2.0$
m_{jj}(2jets)	$70 - 100$
m_{jj}(3jets)	$50 - 100$



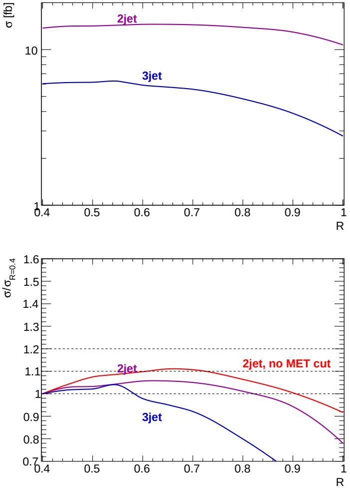
No strong dependence on R visible

Dependence on jet cone size - hZ, Z → j j

same final state,

different topology

variable	cut
MET	$120 - 160$ GeV
N_{jets}	$2 - 3$
ΔR_{jj}	$0.7 - 2.0$
m_{jj}(2jets)	$70 - 100$
m_{jj}(3jets)	$50 - 100$



No strong dependence on R visible

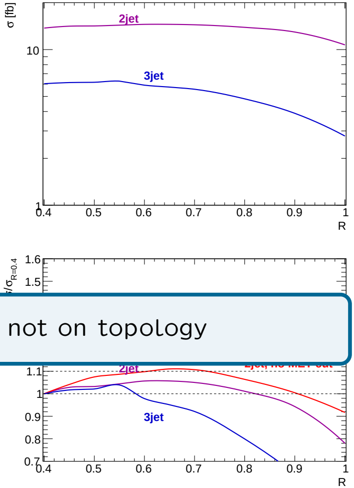
Dependence on jet cone size - hZ, Z → j j

same final state,

different topology

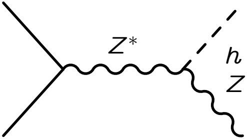
variable	cut
MET	$120 - 160$ GeV
N_{jets}	$2 - 3$
[illegible]	[illegible]
m_{jj}(3jets)	$50 - 100$

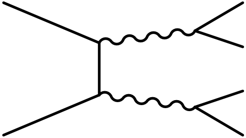
depends on **phase space**, not on topology



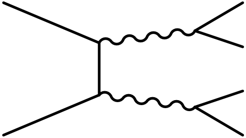
No strong dependence on *R* visible

Zh production - backgrounds

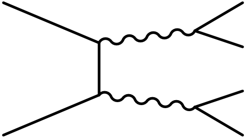




ZZ

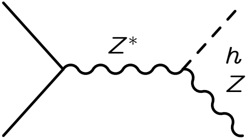


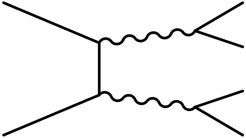
WZ



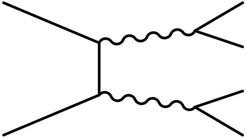
WW

Zh production - backgrounds

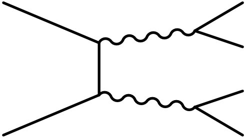




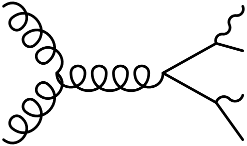
ZZ



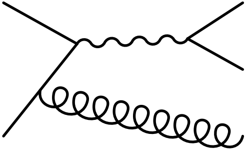
WZ



WW



$t\bar{t}$



Z+jets