Boosted $(VH)H \rightarrow b\bar{b}$?

A brief sensitivity study

Hannah Arnold, Valerio Dao, Tristan du Pree, Brian Moser

February 7, 2018

Nikhef (Dutch National Institute for Subatomic Physics)

Boosted regimes: Setting the stage

Using the Higgs sector as a portal:

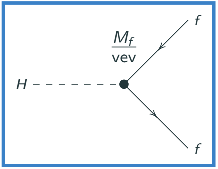
- Most of the (predicted) Higgs decay channels have been already observed or are at the edge ...
- As investigations on the Higgs sector improve it can also be used as a portal to look for physics beyond the standard model
- Signal strengths and couplings (so far) compatible with SM prediction $\rightarrow$ differential distributions?

Using the Golden Channels:

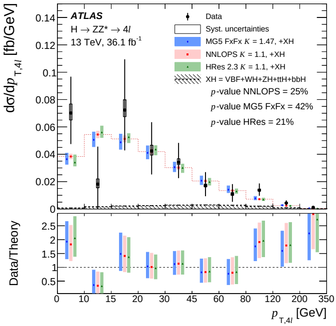
[arXiv:1708.02810]

- Very "clean" channels, signal statistically already observed
- **But:** They do not reach the ultra high-p_T tails with enough statistics

$\rightarrow$ **New approaches needed, e.g. looking at $H \rightarrow bb$**



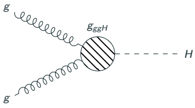
$$\mathrm{BR}(H \rightarrow bb) = 58.1\%$$

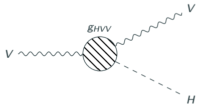


The (theoretical) motivation behind boosted regimes

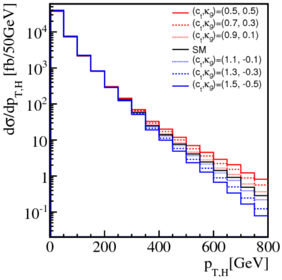
Effective Field Theory approach:

$$\mathcal{L} = \mathcal{L}_{\text{SM}} + \sum_{d>4} \sum_{i=1}^{n(d)} \frac{c_i^{(d)}}{\Lambda^{d-4}} \times \mathcal{O}_i^{(d)}$$

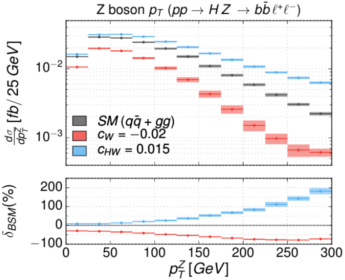




[arXiv:1405.4295]



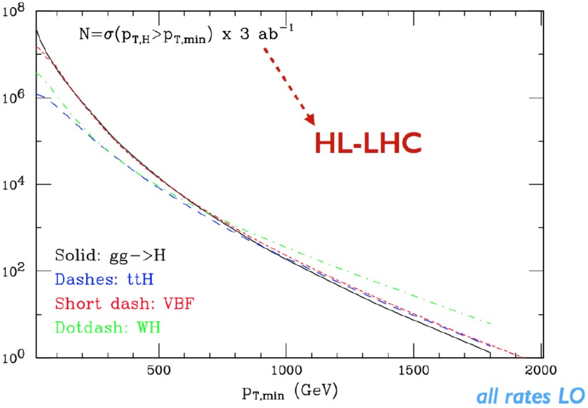
[arXiv:1512.02572]



An other motivation for high-p_T Higgsstrahlung

- At very high Higgs-p_T, the VH production cross-section can be even higher than the ggF one

[Talk at LHCHXSWG ggF meeting]

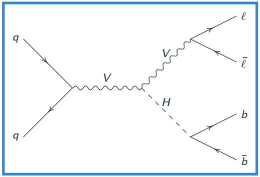


Caution: LO $\rightarrow$ Large corrections in both directions expected when considering higher preturbative orders

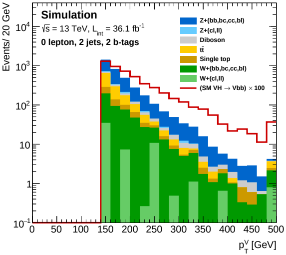
The current $(VH)H \to bb$ analysis in a nutshell

Analysis strategy in a nutshell:

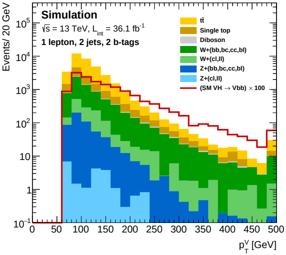
- Split the analysis into subchannels according to the charged lepton multiplicity
- Try to reconstruct the two b-jets using the anti-k_t-4 algotithm and use their invariant mass as a major discriminant (actually a BDT is trained adding additional variables)
- Focus at the phase space with $p_T^V > 150\,\text{GeV}$ in order to increase the signal-to-background ratio



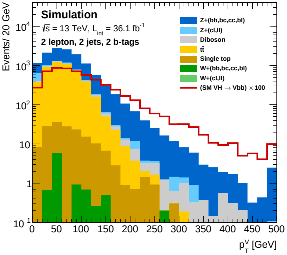
0 lepton:



1 lepton:



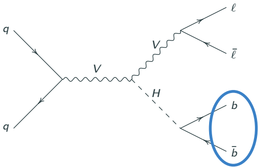
2 lepton:

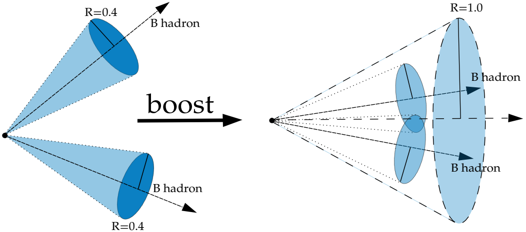


How to deal with very boosted Higgs bosons?

Operating the LHC at $\sqrt{s} = 13$ TeV

- The current $(VH)H \rightarrow bb$ search aims at individually reconstructing the jets that originated from the hard scattered b-quarks
- If these get too close in ΔR the usual (anti-k_t-4) jet cones merge and can no longer be resolved individually





Rule of thumb:

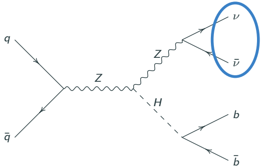
$$\Delta R(j_1, j_2) \simeq \frac{m_{\text{Higgs}}}{p_T^{\text{Higgs}}} \frac{1}{\sqrt{z(1-z)}} \overset{z=1/2}{=} \frac{2 m_{\text{Higgs}}}{p_T^{\text{Higgs}}}$$

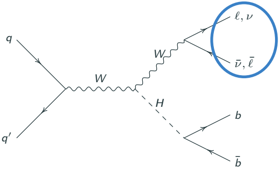
with z being the total momentum fraction of one jet

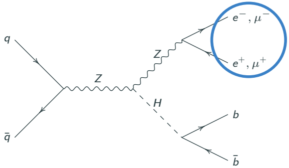
> Merging regime for anti-k_t-4 jets ($R = 0.4$):
> $p_T^H \approx 5 \times m_H = 625$ GeV (assuming $z = 1/2$)

A generator-level study

Setup of the study - What has been done?

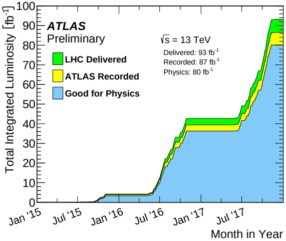






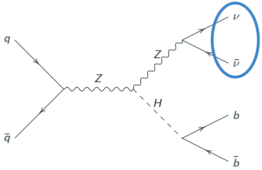
- Simulated separate signal events according to the number of charged final state leptons **without** the detector simulation:
  - 0 lepton: $\sigma \times \mathrm{BR} = 103.2\,\mathrm{fb}$ (sum over all 3 flav.)
  - 1 lepton: $\sigma \times \mathrm{BR} = 269.0\,\mathrm{fb}$ (sum over all 3 flav.)
  - 2 lepton: $\sigma \times \mathrm{BR} = 34.7\,\mathrm{fb}$ ($ee + \mu\mu$ only)

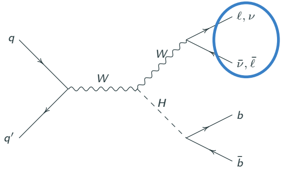
    Only electron and muon decays have been simulated to take care of the event selection acceptance[1].
- Yield scaled up to correspond to an integrated luminosity of $L_{\mathrm{int}} = 80\,\mathrm{fb}^{-1}$ (2015+2016+2017 data set).

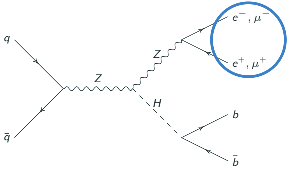


[1] Leptonic tauon decays would fail the selection because of the missing transverse energy due to the neutrinos. More details on the cross sections are in the backup.

Setup of the study - What has been done?







- Simulated separate signal events according to the number of charged final state leptons **without** the detector simulation:
  - 0 lepton: $\sigma \times \mathrm{BR} = 103.2\,\mathrm{fb}$ (sum over all 3 flav.)
  - 1 lepton: $\sigma \times \mathrm{BR} = 269.0\,\mathrm{fb}$ (sum over all 3 flav.)
  - 2 lepton: $\sigma \times \mathrm{BR} = 34.7\,\mathrm{fb}$ ($ee + \mu\mu$ only)
    Only electron and muon decays have been simulated to take care of the event selection acceptance[1].
- Yield scaled up to correspond to an integrated luminosity of $L_{\mathrm{int}} = 80\,\mathrm{fb}^{-1}$ (2015+2016+2017 data set).

How many signal events are expected in the high p_T^{Higgs} tails and how many of them can be reconstructed using either two small-R jets or one large-R jet?

[1] Leptonic tauon decays would fail the selection because of the missing transverse energy due to the neutrinos. More details on the cross sections are in the backup.

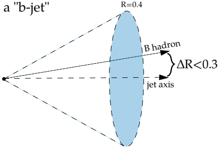
The event selection/categorization

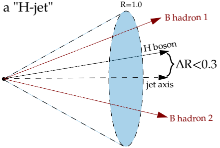
The need of an event selection:

- Mimic the existing VH analysis selection [JHEP12(2017)024] in order to get an acceptance estimate.

The introduced selection:

- Two regions have been defined with selection either on small-R jets (**resolved**) or large-R jets (**boosted**).
- The below selection has been commonly applied for all three subchannels $\rightarrow$ Only difference: $p_T^V > 75$ GeV in 2-lepton vs. $p_T^V > 150$ GeV in 0/1-lepton





Resolved:

Selection	Jet p_T	Leading b-jet p_T	Jet $\|\eta\|$	n_{jet}	$n_{\text{b-jets}}$
Value	> 20 GeV	> 45 GeV	< 2.5	2 or 3	2

Boosted:

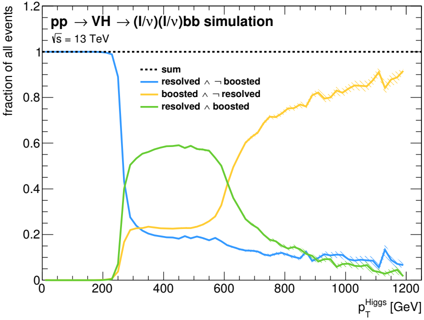
Selection	Jet p_T	Jet $\|\eta\|$	n_{jet}	$n_{\text{H-jets}}$
Value	> 250 GeV	< 2.0	1	1

Categorization and results

Intuitive approach:

- Categorize events in an exclusive, non-overlapping way into:
  1. boosted $\wedge$ ¬resolved
  2. resolved $\wedge$ ¬boosted
  3. resolved $\wedge$ boosted

The fraction of events that can be reconstructed using a resolved approach drops at $p_T^{\text{Higgs}} \sim 600$ GeV - complementary the fraction rises for the boosted approach, as expected

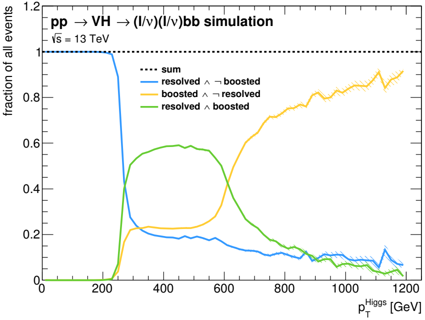


Categorization and results

Intuitive approach:

- Categorize events in an exclusive, non-overlapping way into:
  1. boosted $\wedge$ ¬resolved
  2. resolved $\wedge$ ¬boosted
  3. resolved $\wedge$ boosted

What if an event can be reconstructed in both approaches?



A "priority resolved" scheme:

- An event is first tried to be reconstructed in a resolved approach. Only if this fails a boosted reconstruction is considered. $\rightarrow$ No overlap.
- This is motivated by the desired orthogonality to the existing analysis and the fact that the m_{BB} resolution is better in a resolved reconstruction.

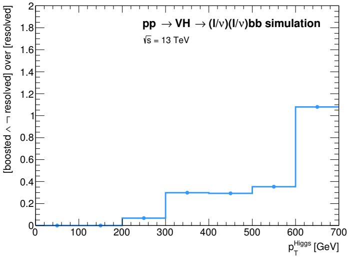
Counting events: Expected signal yields

Expected signal yields for $p_T^{\text{Higgs}} > 500$ GeV:

	0 lepton	1 lepton	2 lepton	combined
all	16.86 ± 0.11	48.8 ± 0.2	5.7 ± 0.1	71.4 ± 0.2
resolved	6.11 ± 0.05	17.01 ± 0.08	1.79 ± 0.04	24.91 ± 0.11
boosted	9.16 ± 0.06	24.02 ± 0.10	2.33 ± 0.05	35.51 ± 0.12
boosted $\wedge \neg$resolved	4.38 ± 0.04	11.79 ± 0.07	1.15 ± 0.04	17.33 ± 0.09
resolved $\vee$ boosted	10.49 ± 0.07	28.80 ± 0.11	2.95 ± 0.06	42.24 ± 0.14

Conclusions:

- Already in the medium p_T^H range events could be gained
- At very high p_T^H **boosted reconstruction** is the way to go
- The expected number of signal events is smallest in the 2-lepton channel
- Largest signal yield in 1-lepton channel, as expected



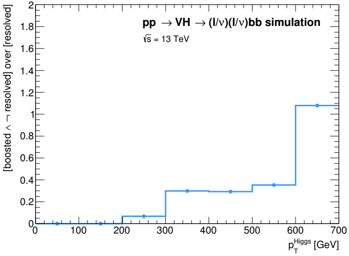
Counting events: Expected signal yields

Expected signal yields for $p_T^{\text{Higgs}} > 500$ GeV:

	0 lepton	1 lepton	2 lepton	combined
all	16.86 ± 0.11	48.8 ± 0.2	5.7 ± 0.1	71.4 ± 0.2
resolved	6.11 ± 0.05	17.01 ± 0.08	1.79 ± 0.04	24.91 ± 0.11
boosted	9.16 ± 0.06	24.02 ± 0.10	2.33 ± 0.05	35.51 ± 0.12
boosted $\wedge \neg$resolved	4.38 ± 0.04	11.79 ± 0.07	1.15 ± 0.04	17.33 ± 0.09
resolved $\vee$ boosted	10.49 ± 0.07	28.80 ± 0.11	2.95 ± 0.06	42.24 ± 0.14

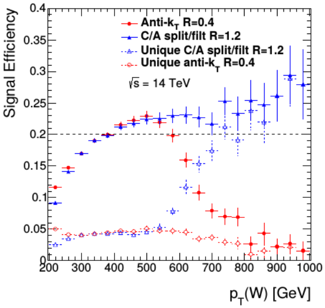
Conclusions:

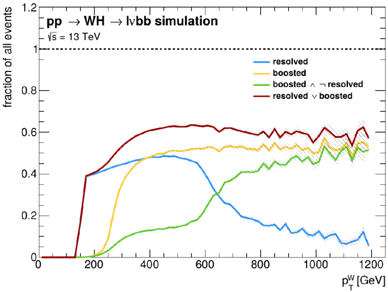
- Already in the medium p_T^H range events could be gained
- At very high p_T^H **boosted reconstruction** is the way to go
- The expected number of signal events is smallest in the 2-lepton channel
- Largest signal yield in 1-lepton channel, as expected



Crosschecking the study

- Compare the results qualitatively to the sensitivity study published by J. Butterworth et al [arXiv:1506.04973] $\rightarrow$ **Similar efficiency behaviour visible**





Why do the plots differ?

- Different jet algorithms used (Cambridge/Aachen vs. anti-k_t) and preselection cuts do not match completely.
- The events in the left graph are weighted to mimic a b-tagging efficiency of 75% $\rightarrow$ difference in signal efficiency of a factor 2

$\rightarrow$ **The two results are in very good agreement**

A full simulation study

Towards a complete study: Adding the detector simulation

What to gain from this step?

- The detector related efficiencies and acceptances obviously further penalize the expected signal yield.
- Having in addition a simulation of all the relevant backgrounds allows the estimation of an expected observation significance $Z_{\text{exp.}}$.

A few technicalities:

- Used the CxAODFramework's publication tag to process the v.28 CxAODs (AnalysisBase 20.7)
- Applied the **resolved** SM cutbased event selection described in [JHEP12(2017)024]
- All event yields have consequently been scaled up to $L_{\text{int}} = 80\,\text{fb}^{-1}$ as well.

About the boosted selection

- At least one large-R jet with exactly two b-tagged anti-k_t-2 sub-(track-)jets
- No additional b-tags outside the large-R jet
- MV2c10 @ 70% signal efficiency on the track jets
- Muon-in-jet correction for large-R jets
- Remaining selection is close to the resolved one

More information on the event selection can be found in the backup

Full simulation results - expected stat. significance (1)

The figure of merit - $Z_{exp.}$:

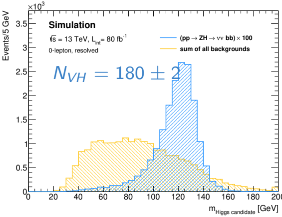
- To approximate the expected significance of the SM $(VH)H \to bb$ signal the usual

$$Z_{\mathrm{exp}} = \mathrm{med}[Z_0|\mu = 1] = \sqrt{2((s+b)\ln(1+s/b) - s)}$$

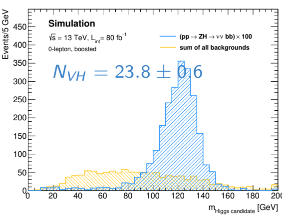
is calculated in a binned way on the invariant mass distribution of the Higgs candidate[2].

Examplary distributions for the 0 lepton channel ($p_T^V > 150$ GeV):

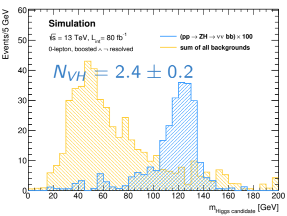
resolved



boosted



boosted ∧ ¬resolved

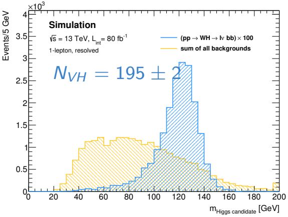


[2] In bins of 5 GeV - if $s_b/b > 0.2$, bins are added until this is no longer the case.

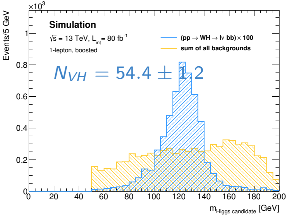
Full simulation results - expected stat. significance (2)

Examplary distributions for the 1 lepton channel ($p_T^V > 150$ GeV):

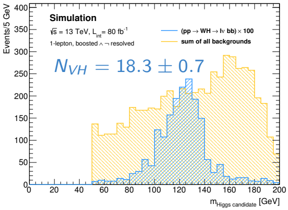
resolved



boosted

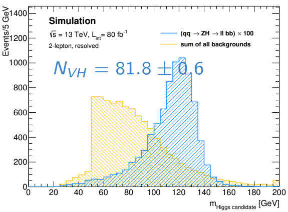


boosted ∧ ¬resolved

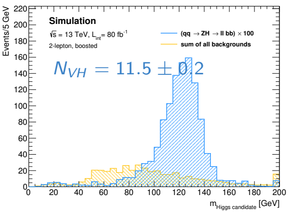


Examplary distributions for the 2 lepton channel ($p_T^V > 75$ GeV):

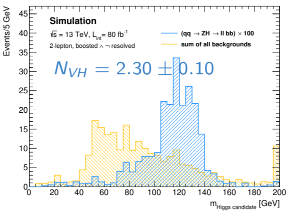
resolved



boosted



boosted ∧ ¬resolved



Full simulation results - expected stat. significance (3)

The figure of merit - $Z_{\text{exp.}}$:

- To approximate the expected significance of the SM $(VH)H \to bb$ signal the usual

$$Z_{\text{exp}} = \text{med}[Z_0|\mu = 1] = \sqrt{2((s+b)\ln(1+s/b) - s)}$$

is calculated in a binned way on the invariant mass distribution of the Higgs candidate.

channel	boosted	boosted $\wedge \neg$resolved	resolved	JHEP res. $\times\sqrt{2}$
0 lep	1.20 ± 0.13	0.27 ± 0.05	2.15 ± 0.08	2.4
1 lep	1.03 ± 0.08	0.37 ± 0.05	2.24 ± 0.10	2.5
2 lep	0.95 ± 0.09	0.25 ± 0.05	1.61 ± 0.06	2.7

Caution: 2 lep is $qq \to ZH \to \ell\ell bb$ only (16% $\sigma \times$ BR missing) and for $p_T^V \geq 75$ GeV

Inclusive significances are in the same range as the expected ones from the publication $\to$ nice crosscheck

The overall signal sensitivity of the $(VH)H \to bb$ search can only be slightly increased by this approach

$\to$ Look especially for high-p_T^H tails

Expected stat. significances at ultra high p_T^V

Requiring $p_T^V > 500$ GeV:

	boosted	boosted $\wedge \neg$resolved	resolved
0 lep	0.48 ± 0.06	0.26 ± 0.04	0.32 ± 0.06
1 lep	0.58 ± 0.08	0.35 ± 0.06	0.26 ± 0.04
2 lep*	0.54 ± 0.13	0.36 ± 0.12	0.29 ± 0.08
sum	0.98 ± 0.18	0.60 ± 0.16	0.53 ± 0.12

*2 lep is $qq \to ZH \to \ell\ell bb$ only (16% $\sigma \times$ BR missing)

Observations:

- The middle column represents the gain from a "priority resolved" scheme $\rightarrow\sim \sqrt{2}$ in sensititvity $\rightarrow$ Promising!
- Maybe consider a boosted only approach after a certain p_T^V threshold, e.g. 500 GeV

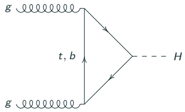
Conclusions

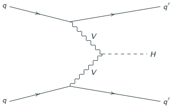
- High-p_T Higgs boson production is sensitive to possible BSM physics also in VH production
  - $H \to bb$ most promising due to large BR
  - Sensitivity studies for most recent data set (2015+2016+2017) have been started

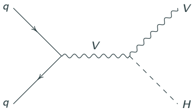
1. Generator-level signal studies:
   - Categorization done to mimic the existing $(VH)H \to bb$ analysis
   - Boosted reconstruction gets dominant for $p_T^H > 600\,\text{GeV}$
   - This is in agreement with the rough estimate of $\frac{p_T^H}{m_H} \sim 5$
2. Full simulation studies:
   - Boosted sensitivity gets compatitive for $p_T^H > 500\,\text{GeV}$ with $Z_{\text{exp}} \sim 0.5$
   - Combining the 0-/1-/and 2-lepton channels yields $Z_{\text{exp}}^{\text{tot}} \sim 1.0$ for boosted reconstruction and $Z_{\text{exp}}^{\text{tot}} \sim 0.6$ for boosted $\wedge \neg$resolved
   - Attention: expected significances are for cut-based approach and stat. only
   - **Next:** Investigation on poissible sensitivity improvements
     - Can we further suppress the major backgrounds?

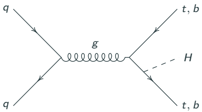
Appendix

Where to look for $H \to bb$?









Several production mechanisms targeted at the LHC:

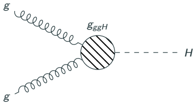
- Gluon-gluon fusion:
  - Huge QCD background, triggering possible at high Higgs-p_T but very low signal-to-background ratio
- Vector-boson fusion:
  - Large QCD background, additional photon helps for triggering and increases signal-to-background ratio
- Higgsstrahlung:
  - The **main** search channel, exploits leptons for triggering and to reduce the QCD background
- ttH/bbH production:
  - Difficult background modelling (*ttbb* or *bbbb*), in top-case leptonic signatures for triggering

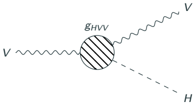
This talk focusses mainly on Higgsstrahlung but includes as well a comparison to already existing searches in gluon fusion.

The (theoretical) motivation behind boosted regimes

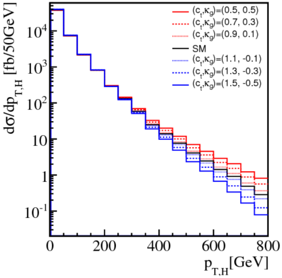
Effective Field Theory approach:

$$\mathcal{L} = \mathcal{L}_{\text{SM}} + \sum_{d>4} \sum_{i=1}^{n(d)} \frac{c_i^{(d)}}{\Lambda^{d-4}} \times \mathcal{O}_i^{(d)}$$

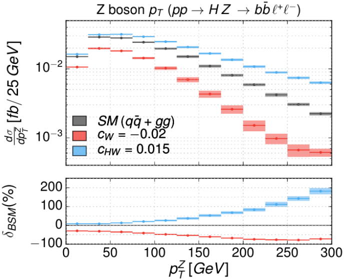




[arXiv:1405.4295]



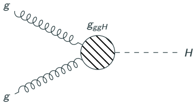
[arXiv:1512.02572]

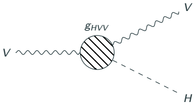


The (theoretical) motivation behind boosted regimes

Effective Field Theory approach:

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}} + \sum_{d>4} \sum_{i=1}^{n(d)} \frac{c_i^{(d)}}{\Lambda^{d-4}} \times \mathcal{O}_i^{(d)}$$

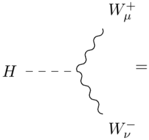




Caution - Operator Phenomenology:

[Johann Brehmer: Higgs EFT]

- EFT approach is limited at high p_T^H as can be shown by the following dim-6 example $\mathcal{O}_W = i\frac{g}{2}(D^\mu \phi)^\dagger \sigma^k (D^\nu \phi) W^k_{\mu\nu}$



$$= \mathrm{i} g m_W \left[g_{\mu\nu} + \frac{f_W}{2\Lambda^2} p_H^2 \, g_{\mu\nu} + \frac{f_W}{2\Lambda^2} \left(p^H_\mu p^+_\nu + p^-_\mu p^H_\nu \right) \right]$$

$$\rightarrow g_{HWW} \sim \mathrm{SM} + \frac{c_W (p_T^H)^2}{\Lambda^2}$$

is assumed to be **small** in this ansatz

Comparison: Status of boosted Higgs boson searches

ATLAS $(VH)H \to bb$ [arXiv:1708.03299]:

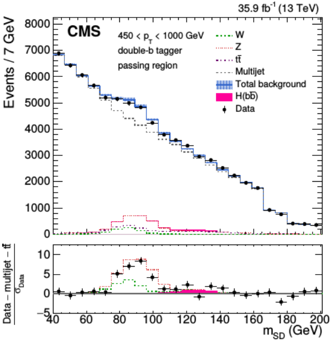
- Most of the sensitivity comes from the medium Higgs-p_T range $p_T \in [150\,\text{GeV}, 300\,\text{GeV}] \to$ Boosted region not (yet) targeted
- **Idea:** Use an orthogonal recycling approach to access this regime as a portal for possible BSM physics
- At the moment we are trying to estimate the expected sensitivity (prob. $\sim 1\sigma$ for $L_{\text{int}} = 80\,\text{fb}^{-1}$) but there are significant improvements possible both in the reconstruction and in the b-tagging

CMS $H \to bb$ + ISR [arXiv:1711.10508]:

- Looking for $gg \to H \to bb$ together with a very high-p_T ISR jet to trigger on
- Expected significance $\sim 1\sigma$ for $L_{\text{int}} = 80\,\text{fb}^{-1}$
- Similar study also started within the ATLAS collaboration

Table 1: Fitted signal strength, expected and observed significance of the Higgs and Z boson signal.

	H	H no p_T corr.	Z
Observed signal strength	$2.3^{+1.8}_{-1.6}$	$3.2^{+2.2}_{-2.0}$	$0.78^{+0.23}_{-0.19}$
Expected significance	0.7σ	0.5σ	5.8σ
Observed significance	1.5σ	1.6σ	5.1σ



Cross-sections for the simulated signal processes (1)

Process	$\sigma \times$ BR [fb]
$qq \to WH \to \ell\nu bb$	269.0
$qq/qg \to ZH(q) \to \ell\ell bb(q)$, $gg \to ZHqq \to qq\ell\ell bb$	29.9
$qq/qg \to ZH(qg) \to \nu\nu bb(q)$, $gg \to ZHqq \to qq\nu\nu bb$	89.0
$gg \to ZH \to \ell\ell bb$	4.8
$gg \to ZH \to \nu\nu bb$	14.2

Table 1: Cross section times branching ratio for the various signal processes, calculated from the values of the LHCHXSWG at $\sqrt{s} = 13$ TeV and $m_H = 125.09$ GeV. The cross sections taken have been calculated at NNLO in QCD and include as well NLO EW corrections, except for the $gg \to ZH \to \ldots$ processes that have been calculated at NLO+NLL in QCD. The branching ratios take into account the most recent partial decay-width calculations and assume the SM particle content. For the $Z \to \ell\ell$ decay, only electrons and muons are taken into account in the calculation whereas for $W \to \ell\nu$ all three lepton categories entered the calculation. This has been done to approximate the event selection acceptance for events involving subsequently decaying tauons.

Cross-sections for the simulated signal processes (2)

0 lepton:

- $qq \to ZH \to \nu\nu bb$
  $\sigma \times \mathrm{BR} = 761.2\,\mathrm{fb} \times 0.5824 \times 0.06729 = 88.66\,\mathrm{fb}$
- $gg \to ZH \to \nu\nu bb$
  $\sigma \times \mathrm{BR} = 122.7\,\mathrm{fb} \times 0.5824 \times 0.06729 = 14.29\,\mathrm{fb}$

1 lepton:

- $pp \to W^+H \to \ell^+\nu bb$
  $\sigma \times \mathrm{BR} = 94.26\,\mathrm{fb} \times 0.5824 \times 3 = 164.69\,\mathrm{fb}$
- $pp \to W^-H \to \ell^-\nu bb$
  $\sigma \times \mathrm{BR} = 59.83\,\mathrm{fb} \times 0.5824 \times 3 = 104.53\,\mathrm{fb}$

2 lepton:

- $qq \to ZH \to \ell\ell bb$
  $\sigma \times \mathrm{BR} = 761.2\,\mathrm{fb} \times 0.5824 \times 0.06729 = 29.83\,\mathrm{fb}$
- $gg \to ZH \to \ell\ell bb$
  $\sigma \times \mathrm{BR} = 122.7\,\mathrm{fb} \times 0.5824 \times 0.06729 = 4.81\,\mathrm{fb}$

Why do the cross-sections differ that much?

- In electroweak unification (GSW) one has four vector fields W^i_μ and B^μ that, after symmetry breaking, mix to the physical mass eigenstates $W^\pm$, Z^0 and A^μ.
- The W boson is a mixture of only the W^1_μ and W^2_μ fields whilst the Z boson is a mixture of both the neutral W^3_μ component and the B^μ field $\rightarrow$ different vertex factors

$\frac{-ig_W}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)$ ($W^\pm$ vertex factor) $\frac{-ig_Z}{2}\gamma^\mu(c^f_V - c^f_A\gamma^5)$ (Z^0 vertex factor)

- The W boson coupling is pure V-A, whilst the Z boson coupling (because of being a mixture of W^3_μ and B^μ) has the coefficients c^f_V and c^f_A that can be calculated if the weak mixing angle is known.
- This accounts for different branching ratios, the W boson decay is in LO completely flavour universal while the Z boson decay is not.

f	c_V	c_A
ν_e, ν_μ, ν_τ	$\frac{1}{2}$	$\frac{1}{2}$
e^-, μ^-, τ^-	$-\frac{1}{2} + 2\sin(\theta_W)$	$-\frac{1}{2}$
u, c, t	$\frac{1}{2} - \frac{4}{3}\sin^2(\theta_W)$	$\frac{1}{2}$
d, s, b	$-\frac{1}{2} - \frac{2}{3}\sin^2(\theta_W)$	$-\frac{1}{2}$

The Branching Ratios:

- $\mathrm{BR}(W \rightarrow l\nu) \approx 1/9$
- $\mathrm{BR}(Z \rightarrow ll) \approx 3.4\%$
- $\mathrm{BR}(Z \rightarrow \nu\nu) \approx 6.8\%$

(always calculated for one flavor)

The event selection in detail (1)

Selection	0 lepton channel	1 lepton channel		2 lepton channel
		e sub-channel	μ sub-channel	
Trigger	E_T^{miss}	Single lepton	E_T^{miss}	Single lepton
Leptons	0 loose leptons with $p_T > 7$ GeV	1 tight e $p_T > 27$ GeV	1 medium μ $p_T > 25$ GeV	2 loose leptons with $p_T > 7$ GeV ≥ 1 leptons with $p_T > 27$ GeV
E_T^{miss}	> 150 GeV	> 30 GeV	-	-
m_{ll}	-	-		81 GeV$< m_{ll} < 101$ GeV
Jets	Exactly 2 or 3 jets			Exactly 2 or ≥ 3 jets
Jet p_T	> 20 GeV			
b-jets	Exactly 2 b-tagged jets			
Leading b-tagged jet p_T	> 45 GeV			
H_T	> 120 GeV (2 jets), > 150 GeV (3 jets)	-		-
$\min[\Delta\phi(E_T^{\text{miss}}, \text{jets})]$	$> 20^\circ$ (2 jets), $> 30^\circ$ (3 jets)	-		-
$\Delta\phi(E_T^{\text{miss}}, jj)$	$> 120^\circ$	-		-
$\Delta\phi(j_1, j_2)$	$< 140^\circ$	-		-
$\Delta\phi(E_T^{\text{miss}}, E_{T,\text{trk}}^{\text{miss}})$	$< 90^\circ$	-		-
p_T^V regions	> 150 GeV			75 GeV$-$150 GeV, > 150 GeV
Signal regions	no add. selection	$m_{jj} \geq 75$ GeV or $m_{\text{top}} \leq 225$ GeV		Same-flavor leptons, opposite-sign charge in $\mu\mu$ sub-channel
Control regions	-	$m_{jj} < 75$ GeV and $m_{\text{top}} > 225$ GeV		Different-flavor leptons

The event selection in detail (2)

Table 4: Summary of the event selection criteria in the 0-, 1- and 2-lepton channels for the dijet-mass analysis, applied in addition to those described in Table 2 for the multivariate analysis.

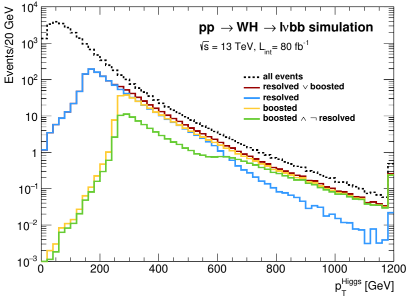
Channel			
Selection	0-lepton	1-lepton	2-lepton
m_{T}^{W}	-	< 120 GeV	-
$E_{\mathrm{T}}^{\mathrm{miss}} / \sqrt{S_{\mathrm{T}}}$	-	-	$< 3.5\ \sqrt{\mathrm{GeV}}$

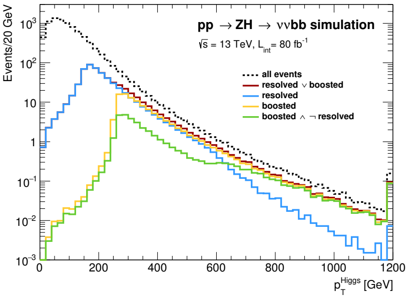
p_{T}^{V} regions			
p_{T}^{V}	(75, 150] GeV (2-lepton only)	(150, 200] GeV	(200, ∞) GeV
$\Delta R(\boldsymbol{b}_1, \boldsymbol{b}_2)$	<3.0	<1.8	<1.2

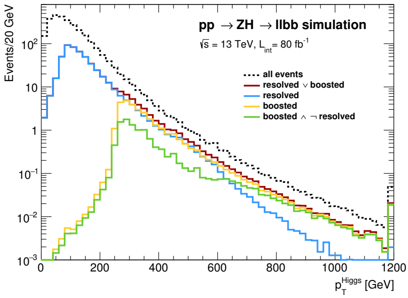
Closer Look: The Higgs boson p_T

Information:

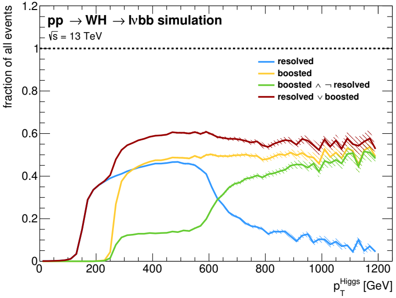
- The transverse momentum of the underlying Higgs boson is shown for all three charged lepton subchannels separately.
- The reference ("all events") has no preselection applied.





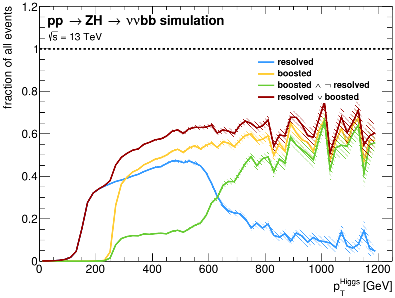


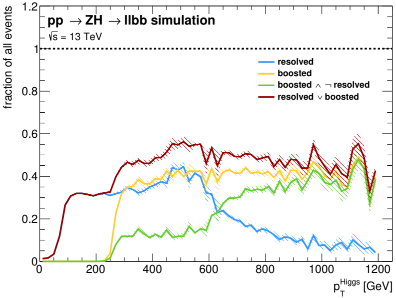
Closer Look: Relative fraction of each category



Information:

- The categories' relative fraction as a function of the transverse momentum of the underlying Higgs boson is shown for all three charged lepton subchannels separately.
- The reference ("all events") has no preselection applied.





Generator study: Expected inclusive signal yields

Expected inclusive signal yields:

	0 lepton	1 lepton	2 lepton	combined
all	8226 ± 5	$21\,552 \pm 9$	2774 ± 3	$32\,550 \pm 10$
resolved	524.4 ± 0.7	1133.6 ± 0.8	489 ± 2	2147 ± 2
boosted	103.8 ± 0.3	248.8 ± 0.3	29.3 ± 0.3	381.9 ± 0.5
boosted $\wedge \neg$resolved	32.1 ± 0.2	77.8 ± 0.2	10.7 ± 0.3	120.6 ± 0.4
resolved $\vee$ boosted	556.5 ± 0.7	1211.4 ± 0.8	500 ± 2	2267 ± 2

Counting events: Expected signal yields

Expected signal yields for $p_T^{\text{Higgs}} > 500$ GeV:

	0 lepton	1 lepton	2 lepton	combined
all	16.86 ± 0.11	48.8 ± 0.2	5.7 ± 0.1	71.4 ± 0.2
resolved	6.11 ± 0.05	17.01 ± 0.08	1.79 ± 0.04	24.91 ± 0.11
boosted	9.16 ± 0.06	24.02 ± 0.10	2.33 ± 0.05	35.51 ± 0.12
boosted $\wedge \neg$resolved	4.38 ± 0.04	11.79 ± 0.07	1.15 ± 0.04	17.33 ± 0.09
resolved $\vee$ boosted	10.49 ± 0.07	28.80 ± 0.11	2.95 ± 0.06	42.24 ± 0.14

Conclusions:

- At very high p_T^H **boosted reconstruction** is the way to go ...
- The expected number of signal events is smallest in the 2-lepton channel $\rightarrow$ How much does this further decrease after the full detector simulation?
- The 1-lepton subchannel might be of high interest since the BR penalty is not that big ...

Yields at ultra high p_T^V

2l: Event yields and expected observation significance

Event yields and expected stat. significance:

	boosted	boosted $\wedge \neg$resolved	resolved	resolved ($p_T^V > 250$ GeV)
n_{sig}	11.5 ± 0.2	2.30 ± 0.10	81.8 ± 0.6	12.9 ± 0.2
n_{bkg}	413 ± 10	224 ± 9	8964 ± 79	397 ± 8
Z_{exp}	0.95 ± 0.09	0.25 ± 0.05	1.61 ± 0.06	1.30 ± 0.10
$Z_{\text{exp}}^{\text{contr.}}$	0.82 ± 0.02	0.226 ± 0.012	1.353 ± 0.011	1.04 ± 0.02

2 lep: Yields and expected stat. significances at ultra high p_T^V

Requiring $p_T^V > 500$ GeV:

	boosted	boosted $\wedge \neg$resolved	resolved
n_{sig}	1.12 ± 0.07	0.54 ± 0.04	0.55 ± 0.05
n_{bkg}	18.9 ± 1.3	10.7 ± 1.0	19.9 ± 1.3
Z_{exp}	0.54 ± 0.13	0.36 ± 0.12	0.29 ± 0.08

Requiring $p_T^V > 600$ GeV:

	boosted	boosted $\wedge \neg$resolved	resolved
n_{sig}	0.47 ± 0.04	0.31 ± 0.03	0.14 ± 0.02
n_{bkg}	7.3 ± 0.8	4.3 ± 0.7	8.6 ± 0.8
Z_{exp}	0.39 ± 0.13	0.38 ± 0.15	0.12 ± 0.06

Only $qq \rightarrow ZH \rightarrow \ell\ell bb$ for 80 fb^{-1}!

1 lep: Yields and expected stat. significances

	boosted	boosted $\wedge \neg$resolved	resolved
n_{sig}	54.4 ± 1.2	18.3 ± 0.7	195 ± 2
n_{bkg}	6920 ± 90	5650 ± 80	23100 ± 300
Z_{exp}	1.03 ± 0.08	0.37 ± 0.05	2.24 ± 0.10

Requiring $p_T^V > 250$ GeV:

	boosted	boosted $\wedge \neg$resolved	resolved
n_{sig}	49.1 ± 1.1	15.3 ± 0.6	52.8 ± 1.2
n_{bkg}	4670 ± 70	3620 ± 60	1920 ± 240
Z_{exp}	1.15 ± 0.10	0.39 ± 0.06	2.1 ± 0.2

1 lep: Yields and expected stat. significances at ultra high p_T^V

Requiring $p_T^V > 500$ GeV:

	boosted	boosted $\wedge \neg$resolved	resolved
n_{sig}	6.1 ± 0.4	3.3 ± 0.3	2.0 ± 0.2
n_{bkg}	362 ± 18	275 ± 17	78 ± 13
Z_{exp}	0.58 ± 0.08	0.35 ± 0.06	0.26 ± 0.04

Requiring $p_T^V > 600$ GeV:

	boosted	boosted $\wedge \neg$resolved	resolved
n_{sig}	2.4 ± 0.2	1.6 ± 0.2	0.39 ± 0.08
n_{bkg}	142 ± 12	106 ± 11	29 ± 5
Z_{exp}	0.31 ± 0.05	0.25 ± 0.04	0.08 ± 0.02

0 lep: Yields and expected stat. significances

	boosted	boosted $\wedge \neg$resolved	resolved
n_{sig}	23.8 ± 0.6	2.4 ± 0.2	180 ± 2
n_{bkg}	1180 ± 30	470 ± 20	20400 ± 200
Z_{exp}	1.20 ± 0.13	0.27 ± 0.05	2.15 ± 0.08

Requiring $p_T^V > 250$ GeV:

	boosted	boosted $\wedge \neg$resolved	resolved
n_{sig}	21.3 ± 0.6	2.0 ± 0.2	38.1 ± 0.7
n_{bkg}	850 ± 20	322 ± 14	1050 ± 30
Z_{exp}	1.29 ± 0.13	0.26 ± 0.04	2.2 ± 0.2

0 lep: Yields and expected stat. significances at ultra high p_T^V

Requiring $p_T^V > 500$ GeV:

	boosted	boosted $\wedge \neg$resolved	resolved
n_{sig}	2.1 ± 0.2	0.79 ± 0.09	1.40 ± 0.12
n_{bkg}	38 ± 4	19 ± 2	31 ± 4
Z_{exp}	0.48 ± 0.06	0.26 ± 0.04	0.32 ± 0.06

Requiring $p_T^V > 600$ GeV:

	boosted	boosted $\wedge \neg$resolved	resolved
n_{sig}	0.87 ± 0.08	0.48 ± 0.06	0.33 ± 0.05
n_{bkg}	17 ± 2	10 ± 2	11 ± 3
Z_{exp}	0.27 ± 0.04	0.24 ± 0.04	0.24 ± 0.04