Search for heavy resonances in final states with 4ℓ and missing transverse energy or jets

Abdualazem Fadol

AIs from the HDBS approval meeting

中国科学院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences

June 19, 2023

ATLAS EXPERIMENT

List of AIs received during the HDBS approval

(1) Need to check overlap removal for 6ℓ between two sets of $A \to ZH$ samples.

(2) For the $A \to ZH$, what happens in events with 6 leptons then? How are the 4 leptons chosen?

(3) Are signal regions orthogonal? Yes, but the diagram and text need to be improved.

(4) Is the number of b-jets a subset of the number of central jets?

(5) Confirm that the interpolated/generated behaviour is similar for many mass points (difference between interpolated and generated is smaller than the signal modelling variations) and if so, just use the interpolated/generated difference as the uncertainty.

(6) Make sure that there's nothing wrong/missing. Try to artificially inflate a leading uncertainty and see if it induces a constraint.

(7) Add category names to each label (rather than numbers) and add a few more mass points so that we can see how consistent the normalization factors are.

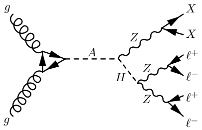
(8) Do a comparison for a low mass (high stats) and a high mass (low stats) point.

AI number (1)

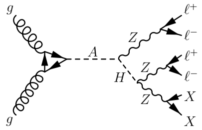
- ☐ **Need to check overlap removal for 6ℓ between two sets of $A \to ZH$ samples.**
  Answer:

- ☐ Forcing the total number of lepton to be exactly 4ℓ in the generation of the $A \to Z(\to 2\ell)H(\to 2\ell + X)$ signal, so X is everything except 2ℓ in this case.

```
#----------------------------------------#
#- The A->Z(->X)H(->4l) signal generation -#
#----------------------------------------#
process="""
import model 2HDM_GF
define p = g u c d s u~ c~ d~ s~ b b~
define j = g u c d s u~ c~ d~ s~ b b~ a
define vl = ve vm vt
define vl~ = ve~ vm~ vt~
define l+ = e+ mu+ ta+
define l- = e- mu- ta-
define inc = u c d s u~ c~ d~ s~ b b~ e+ mu+ ta+ e- mu- ta- ve vm vt ve~ vm~ vt~
# Define multiparticle labels
# Specify process(es) to run
"""
if mA > mH:
      process += """
generate     g g > h3 > z h2 , z > inc inc, h2 > z z
output -f
"""
#------------------------------------------#
#- The A->Z(->2l)H(->2l+X) signal generation -#
#------------------------------------------#
process="""
import model 2HDM_GF
define p = g u c d s u~ c~ d~ s~ b b~
define j = g u c d s u~ c~ d~ s~ b b~ a
define vl = ve vm vt
define vl~ = ve~ vm~ vt~
define l+ = e+ mu+
define l- = e- mu-
define inc = u c d s u~ c~ d~ s~ b b~ e+ mu+ ta+ e- mu- ta- ve vm vt ve~ vm~ vt~
# Define multiparticle labels
# Specify process(es) to run
"""
if mA > mH:
      process += """
generate     g g > h3 > z h2 , z > l+ l- , h2 > z z
output -f
"""
MultiLeptonWithParentFilter.NLeptonsMin = 4
MultiLeptonWithParentFilter.NLeptonsMax = 4
```



$X \to \ell/q/\nu$



$X \to q/\nu$

There's **no overlap** between the two signals. However, we need to clarify that in the text, such as changing the symbol for one of the signals.

AI number (2)

- ☐ For the $A \rightarrow ZH$, what happens in events with 6 leptons then? How are the 4 leptons chosen?

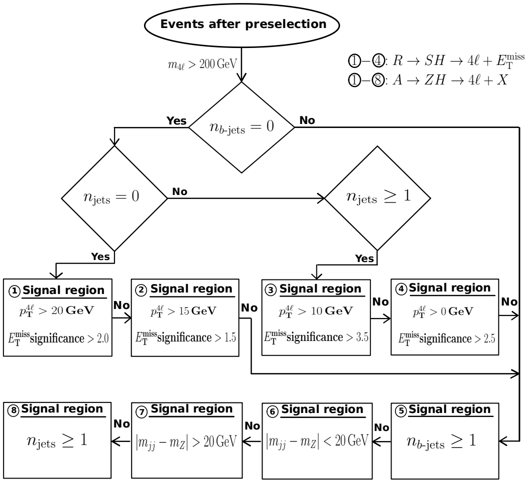
  Answer: the following answer is provided by RD (HZZ framework expert and developer)

- ☐ When we find a quadruplet and there is an extra lepton above 12 GeV, passing isolation and d0 too, then we use a matrix element calculation based on the decay of the 4 leptons for all of the quads. Only in this case do we allow to 'change our minds' and pick the quad with the highest value for the ME. This is described (a bit) in the HZZ common support note.

AI number (3)

- ☐ Are signal regions orthogonal? Yes, but the diagram and text need to be improved.

  Answer: We improved the flowchart and the text for the signal regions.



AI number (4)

- **Is the number of *b*-jets a subset of the number of central jets?**

Answer:

	$q\bar{q} \to ZZ$	$gg \to ZZ$	$q\bar{q} \to ZZ(EW)$	$t\bar{t}V$	VVV	$Z + jets$	WZ	$t\bar{t}$	Expected
Preselection	2518.94±0.00	349.23±0.00	32.77±0.00	47.18±0.00	19.05±0.00	3.85±0.00	4.89±0.00	2.80±0.00	2978.72±0.00
$n_{\text{jets}} = 0$	1622.87±3.64	213.09±0.53	2.97±0.11	2.44±0.09	9.38±0.07	2.52±0.89	2.64±0.23	0.82±0.09	1856.73±5.65
$n_{\text{jets}} = 0$ & $n_{b\text{-jets}} = 0$	1622.87±3.64	213.09±0.53	2.97±0.11	2.44±0.09	9.38±0.07	2.52±0.89	2.64±0.23	0.82±0.09	1856.73±5.65
$n_{\text{jets}} \geq 1$	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
$n_{\text{jets}} \geq 1$ & $n_{b\text{-jets}} = 0$	896.07±2.17	136.15±0.42	29.80±0.24	44.74±0.43	9.66±0.08	1.33±0.73	2.25±0.21	1.99±0.17	1121.98±4.46
$n_{\text{jets}} \geq 1$ & $n_{b\text{-jets}} \geq 1$	835.39±2.11	129.36±0.41	28.00±0.24	9.80±0.21	9.16±0.08	1.21±0.73	2.12±0.21	1.04±0.11	1016.07±4.09

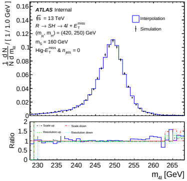
- Yes, the number of *b*-jets is a subset of the number of enteral jets. So here will be no b-jets if there're no central jets.
- However, the signal regions were developed by divided events into 0- and 1-jets for events with no *b*-jets. Then events fail this criteria move to the second selection.
- This is shown in the digram on the previous slide.

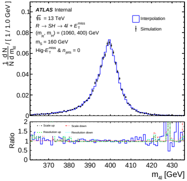
AI number (5)

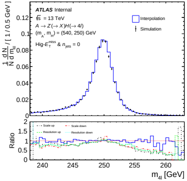
- ☐ Confirm that the interpolated/generated behaviour is similar for many mass points (difference between interpolated and generated is smaller than the signal modelling variations) and if so, just use the interpolated/generated difference as the uncertainty.

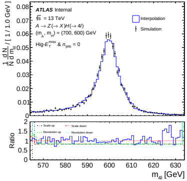
Answer:

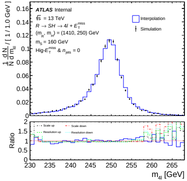
- ☐ We check different samples with low, medium and high energy gap ($m_{A/R} - m_H$)

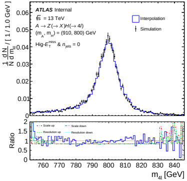


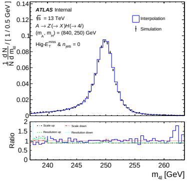








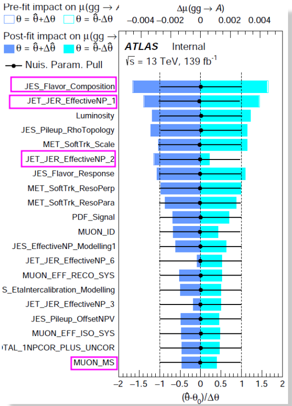




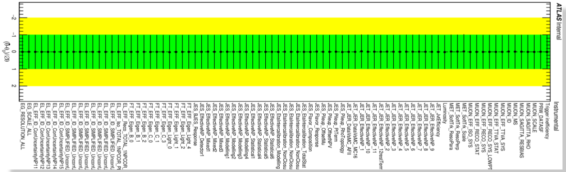
- ☐ Because the systematic uncertainties are overestimated on the edges we decided to use the difference between the MC and e interpolated histograms.

AI number (6)

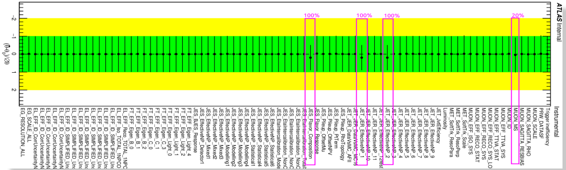
- Make sure that there's nothing wrong/missing. Try to artificially inflate a leading uncertainty and see if it induces a constraint.



$(m_A, m_H) = (510, 380)$ GeV



Before inflating NPs

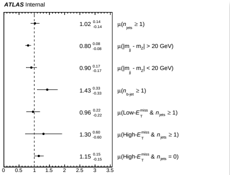


After inflating NPs

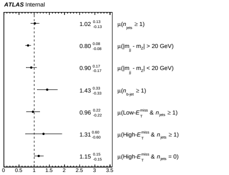
- **Inflating some nuisance parameters produces pulls and constrains on the parameters.**

AI number (7)

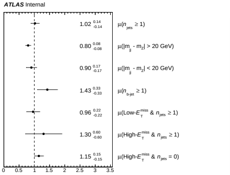
- Add category names to each label (rather than numbers) and add a few more mass points so that we can see how consistent the normalization factors are.



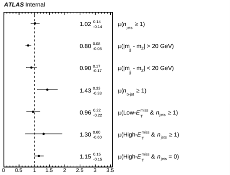
$(m_A, m_H) = (420, 250)$ GeV



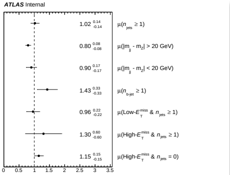
$(m_A, m_H) = (420, 240)$ GeV



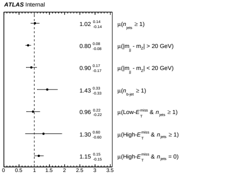
$(m_A, m_H) = (500, 250)$ GeV



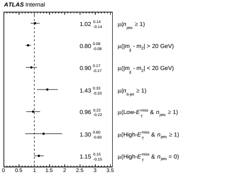
$(m_A, m_H) = (800, 250)$ GeV



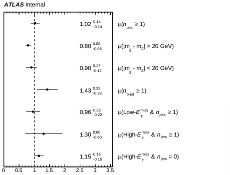
$(m_A, m_H) = (510, 380)$ GeV



$(m_A, m_H) = (470, 300)$ GeV



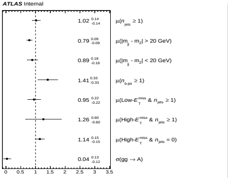
$(m_A, m_H) = (800, 300)$ GeV



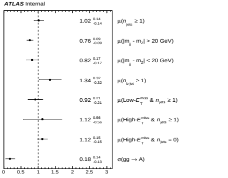
$(m_A, m_H) = (820, 620)$ GeV

- The background normalisation under the background-only hypothesis for different mass points is shown.

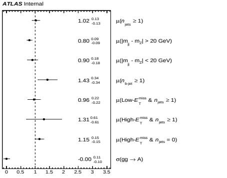
- Add category names to each label (rather than numbers) and add a few more mass points so that we can see how consistent the normalization factors are.



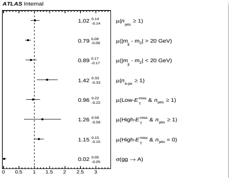
$(m_A, m_H) = (420, 250)$ GeV



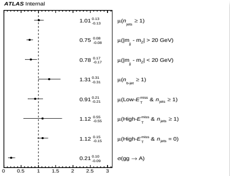
$(m_A, m_H) = (420, 240)$ GeV



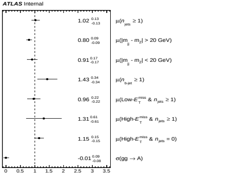
$(m_A, m_H) = (500, 250)$ GeV



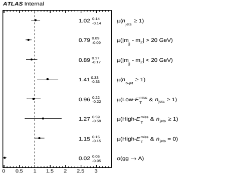
$(m_A, m_H) = (800, 250)$ GeV



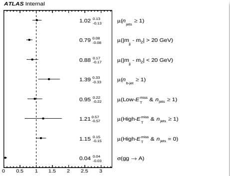
$(m_A, m_H) = (510, 380)$ GeV



$(m_A, m_H) = (470, 300)$ GeV



$(m_A, m_H) = (800, 300)$ GeV



$(m_A, m_H) = (820, 620)$ GeV

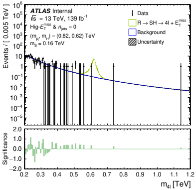
- The background normalisation under the signal-plus background hypothesis for different mass points is shown.

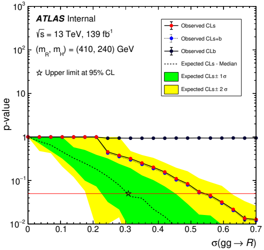
AI number (8)

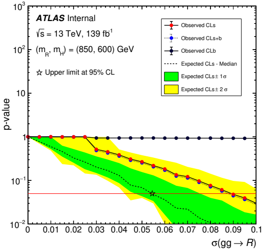
☐ Do a comparison for a low mass (high stats) and a high mass (low stats) point.

Answer:

- Comparing toy MC and asymptotic upper limits for the $R \to SH \to 4\ell + E_{\mathrm{T}}^{\mathrm{miss}}$ signal.







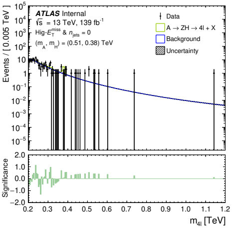
- Two mass points peak around high ($(m_R, m_H) = (410, 240)$ GeV) and low ($(m_R, m_H) = (850, 600)$GeV) statistics were chosen for the study. The number of toy MC generated was 5900 toys.
- The comparison is shown on the table below:

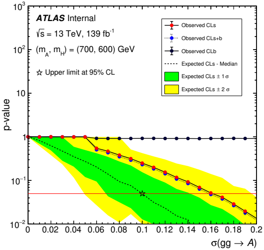
Upper limit	$(m_R, m_H) = (410, 240)$ GeV		$(m_R, m_H) = (850, 600)$ GeV	
	Expected	Observed	Expected	Observed
Toy MC	0.308	0.544	0.054	0.089
Asymptotic	0.303	0.532	0.053	0.085
Toys/Asymptotic	1.02	1.02	1.02	1.05

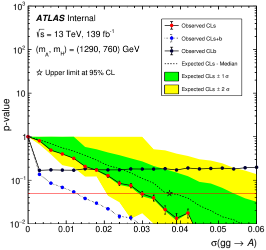
- The signal strength scan vs p-values for each mass point is shown on the right.

AI number (8)

- Comparing toy MC and asymptotic upper limits for the $A \to ZH \to 4\ell + X$ signal.







- Two mass points peak around high $((m_A, m_H) = (510, 380)$ GeV) and low $((m_A, m_H) = (700, 600)$GeV) statistics were chosen for the study. The number of toy MC generated was N toys.
- The comparison is shown on the table below:

Upper limit	$(m_A, m_H) = (510, 380)$ GeV $N = 5800$		$(m_A, m_H) = (700, 600)$ GeV $N = 5900$		$(m_A, m_H) = (1290, 760)$ GeV $N = 5600$	
	Expected	Observed	Expected	Observed	Expected	Observed
Toy MC	0.181	0.395	0.100	0.162	0.038	0.031
Asymptotic	0.166	0.360	0.097	0.162	0.038	0.026
Toys/Asymptotic	1.10	1.10	1.03	1.0	1.0	1.19

- The signal strength scan vs p-values for each mass point is shown on the right.

Summary

- We addressed all the action items (AIs) brought up during the HDBS approval meeting.
- The INT note was modified according to the suggestions and comments in the CDS.
- And a new version (1.4) of the INT note was uploaded to the CDS.