Note

Feng-Yang Hsieh

1 Signal and background

Signal events are 3 Higgs events in BSM. I use MadGraph to generate by following commands:

```
import model cxSM_VLF_EFT
generate g g > h h h
```

then pass the events through the Pythia and Delphes. In Pythia, I set the branching ratio of Higgs so that it only decays to b quarks.

Background events are 6b events in SM. I use MadGraph to generate by following commands:

```
generate p p > b b b b~ b~ b~
```

then pass the events through the Pythia and Delphes.

2 Transverse momentum and pseudorapidity distribution

In each event, there are 6 b quarks. I order these b quarks by transverse momentum p_{T}, then plot p_{T} and pseudorapidity η distributions.

Figure 1 is the p_{T} and η distributions of b quarks for background.

Figure 2 is the p_{T} and η distributions of b quarks for signal.

3 Delta R and transverse momentum distribution

For signal, I plot the ΔR(b,b) distribution of two b decays from the same Higgs. Then I plot ΔR(b,b) against p_{T}(b,b). The result is in Figure 3. As can be seen from the plot, ΔR get smaller at higher p_{T}(b,b).

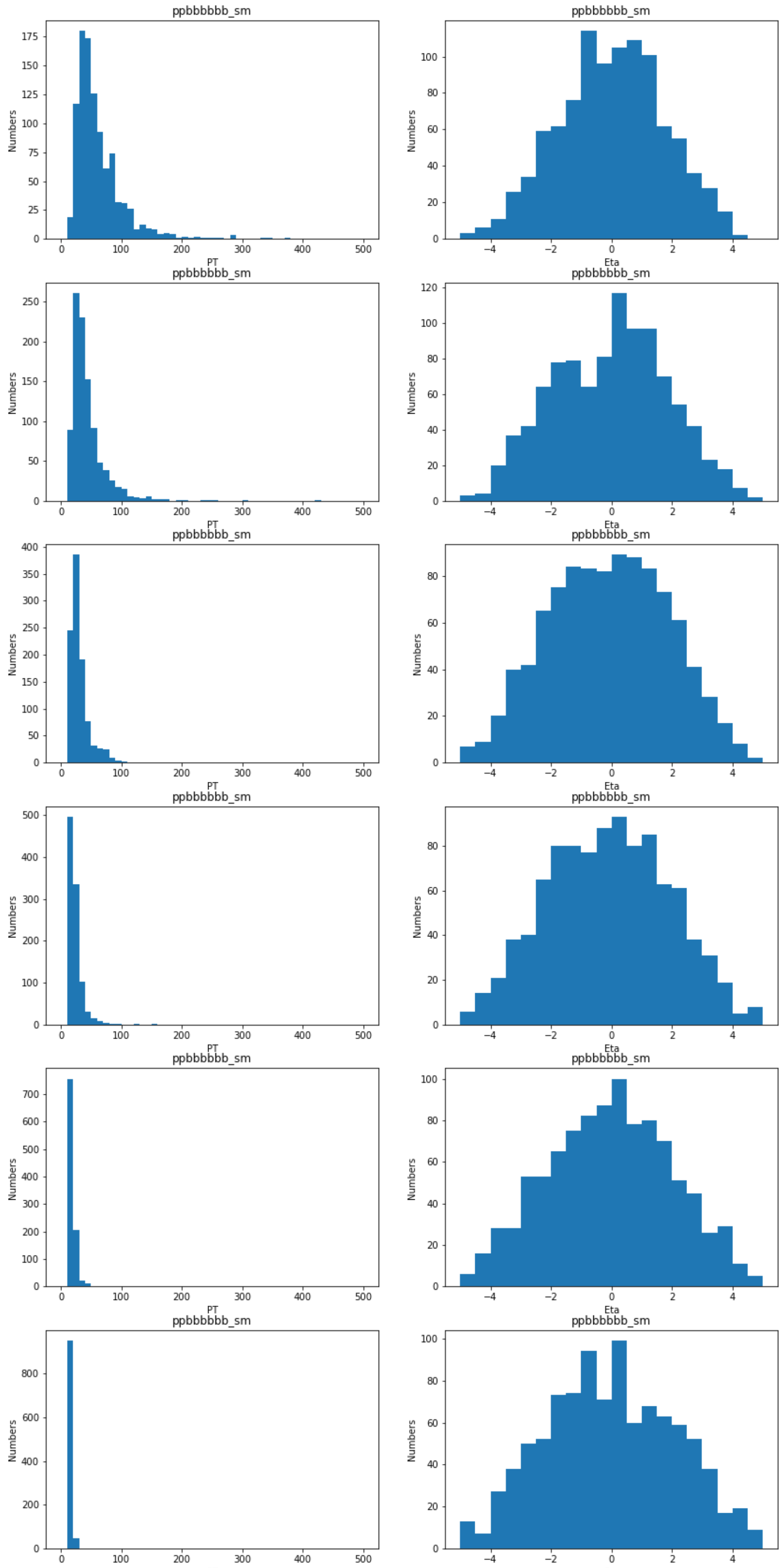


Figure 1: p_{T} and η distributions of b partons for background events. They are ordered by p_{T}.

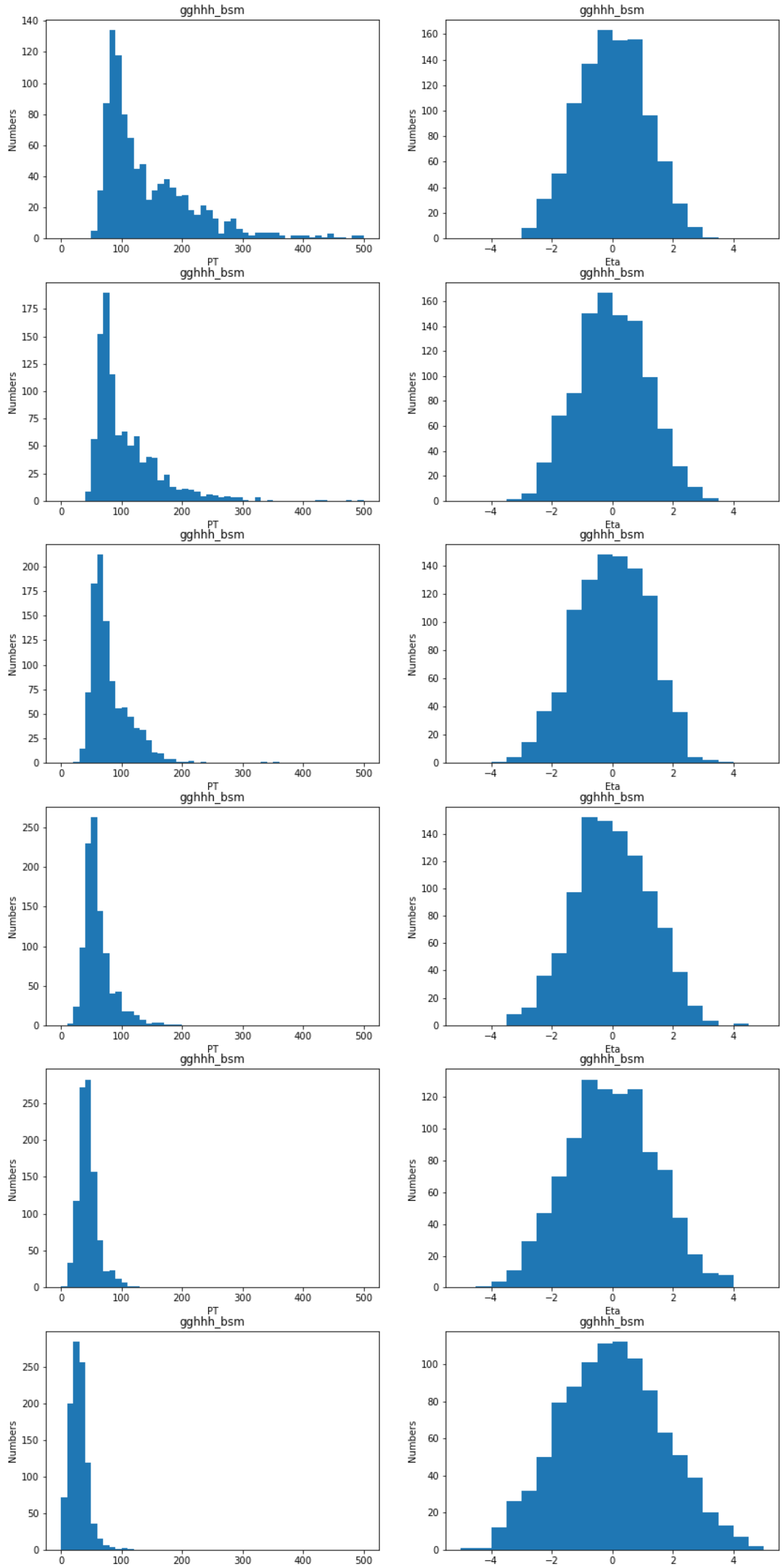


Figure 2: p_{T} and η distributions of b partons for signal events. They are ordered by p_{T}.

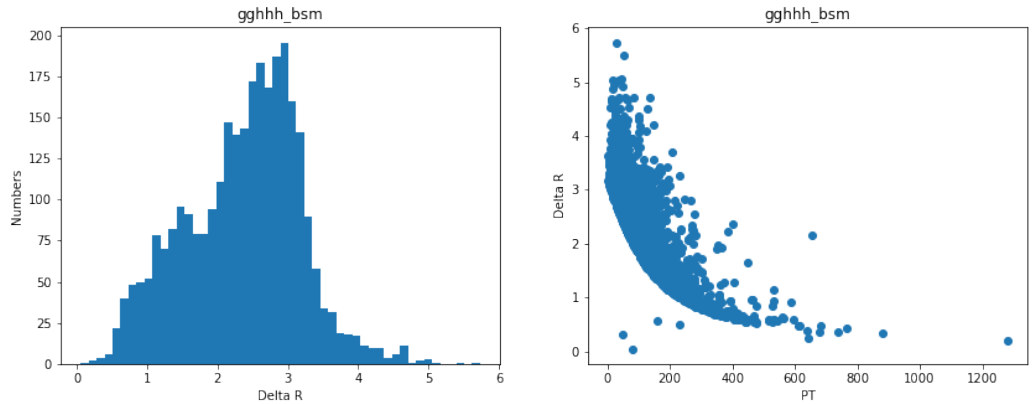


Figure 3: ΔR(b,b) distribution and ΔR(b,b) against p_{T}(b,b) plot

4 Cutflow table

I apply the η and p_{T} cuts sequentially and count how many events can pass the cuts. The results are in Table 1 and Table 2.

Table 1: 1000 background events. Those background events is generated in $\sqrt{s}$ = 13 TeV with the cut b $p_{\mathrm{T}} > 10$ GeV in MadGraph.

Cut	number of events pass
6 b-partons $\lvert\eta\rvert < 2.5$	330
6 b-partons $p_{\mathrm{T}} < 25$ GeV	11
4 b-partons $p_{\mathrm{T}} < 40$ GeV	8

Table 2: 1000 signal events.

Cut	number of events pass
6 b-partons $\lvert\eta\rvert < 2.5$	732
6 b-partons $p_{\mathrm{T}} < 25$ GeV	534
4 b-partons $p_{\mathrm{T}} < 40$ GeV	498

5 Construction of bb-pairs

To construct the bb-pairs from the same Higgs, the typical strategy is to try all combinatoric and find the minimal mass difference between all pairs, i.e. minimize this

$$\chi^2 = [M(\mathrm{b}_1\mathrm{b}_2) - M(\mathrm{b}_3\mathrm{b}_4)]^2 + [M(\mathrm{b}_1\mathrm{b}_2) - M(\mathrm{b}_5\mathrm{b}_6)]^2 + [M(\mathrm{b}_3\mathrm{b}_4) - M(\mathrm{b}_5\mathrm{b}_6)]^2$$

where $M(\mathrm{b}_i\mathrm{b}_j)$ means the invariant mass of b-parton i and b-parton j.

The other strategy is to minimize mass difference to the SM Higgs mass

$$\chi^2 = [M(\mathrm{b}_1\mathrm{b}_2) - M_\mathrm{H}]^2 + [M(\mathrm{b}_3\mathrm{b}_4) - M_\mathrm{H}]^2 + [M(\mathrm{b}_5\mathrm{b}_6) - M_\mathrm{H}]^2$$

I implement these two methods to construct bb-pairs and test on the 10,000 signal events at the parton level. The result is as follows:

- Method 1 (mass difference between all pairs): accuracy = 0.863
- Method 2 (mass difference to the SM Higgs mass): accuracy = 0.875

Method 2 is better for identifying the true Higgs pair.

6 Absolute cross section

The absolute cross section is defined as follows

$$\sigma_\mathrm{abs} = \sigma \times \frac{\text{number of events pass the cut}}{\text{number of events}}$$

I have generated background events with different cuts in MadGraph, the detailed information of cuts is in Table 3. I calculate their absolute cross section and check if it is the same or not. The result is in Table 4.

From Table 4, if the number of events is large enough the absolute cross section will be similar.

7 The problem for generating pp $\to$ 6b events

Failed to generate the requested number of events

Table 3: The cuts applied on the different runs. Where p_T means the minimum transverse momentum of b and $|\eta|$ means the range of b (-1 means no restriction).

No.	p_T (GeV)	$\|\eta\|$
1	0	-1
2	10	5
3	10	3
4	20	3
5	0	-1
6	10	-1

Table 4: Absolute cross section. Where the selection cut is $p_T > 20$ GeV, $|\eta| < 3$.

No.	total event	cross section	events pass selection	absolute cross section
1	339	2731.9362	0	0.0
2	4761	71.316651	190	2.846
3	9054	35.225852	694	2.700
4	4207	2.7698473	4183	2.754
5	1000	2531.9304	0	0.0
6	1000	70.158431	27	1.894

8 pp $\to$ 4b

I use MadGraph to generate 4b events by following commands:

```
generate p p > b b b~ b~
```

In MadGraph, I apply the cuts: $p_{\mathrm{T}} > 25$ GeV, $|\eta| < 2.5$ for b. Pass those events through Pythia, then check how many events there are 6 b-hadrons in the final state.

In 100,000 events, 6,916 events are having greater than or equal to 6 b-hadrons. The distribution of number of b-hadrons is in Figure 4.

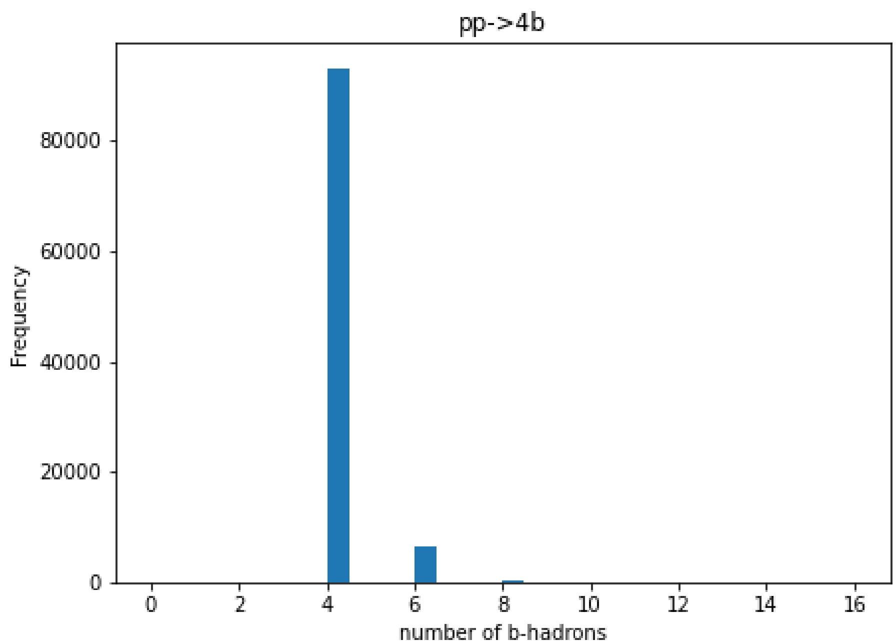


Figure 4: Number of b-hadrons in Pythia final state.

Pass events through Delphes. In Delphes the b-tagging efficiency is set to 1.

Figure 5 is the number of jets and number of b-jets distributions. In 100,000 events, 173 events are having greater than or equal to 6 b-jets.

9 Comparison for pp $\to$ 4b and pp $\to$ 6b event

To compare pp->4b and pp->6b events, I plot the p_{T}, η and total invariant mass of 6 b-jets for both.

The number of events of 6 b-jets for pp->4b is 2051 and for pp->6b is 1408. I scaled the number of events for pp->4b to be the same as pp->6b.

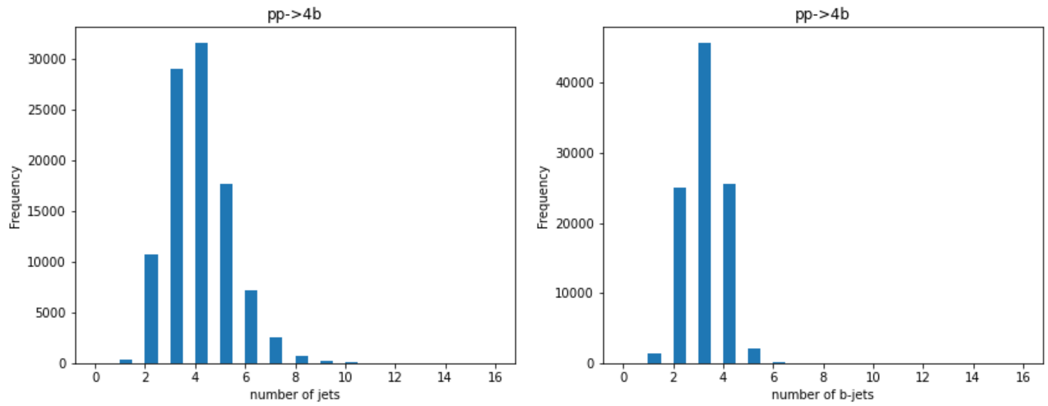


Figure 5: Number of jets and number of b-jets.

Figure 6 is p_{T} and η distributions. Figure 7 is total invariant mass of 6 b-jets distributions. Their distributions look similar.

10 Cutflow table for b-jets

I apply the following cut sequentially and count how many events can pass these cuts. Table 5 is the result for pp->4b events. Table 6 is the result for pp->6b events. Table 7 is the result for signal events.

- Cut 1: There are greater than or equal to 6 b-jets.
- Cut 2: There are greater than or equal to 6 b-jets satisfy $|\eta| < 2.5$.
- Cut 3: There are greater than or equal to 6 b-jets satisfy $p_{\mathrm{T}} > 25$ GeV.
- Cut 4: There are greater than or equal to 4 b-jets satisfy $p_{\mathrm{T}} > 40$ GeV.

Table 5: 1,000,000 pp->4b events. Those events are generated in $\sqrt{s} = 14$ TeV with the cuts: $p_{\mathrm{T}} > 25$ GeV, $|\eta| < 2.5$ for b in MadGraph.

Cut	number of event pass
1	1783
2	1565
3	1059
4	735

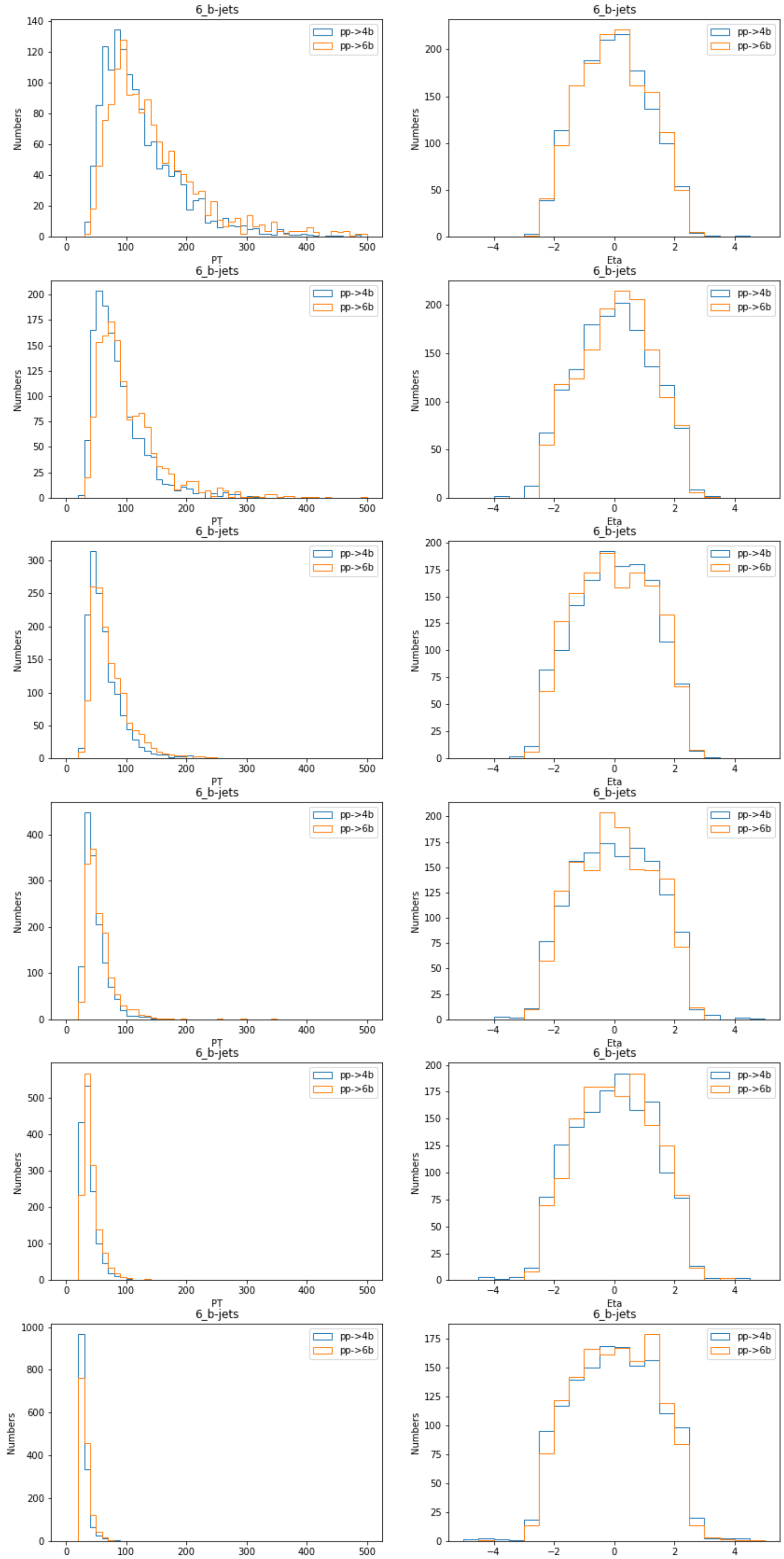


Figure 6: p_{T} and η distributions of b-jets for pp->4b and pp->6b events. They are ordered by p_{T}.

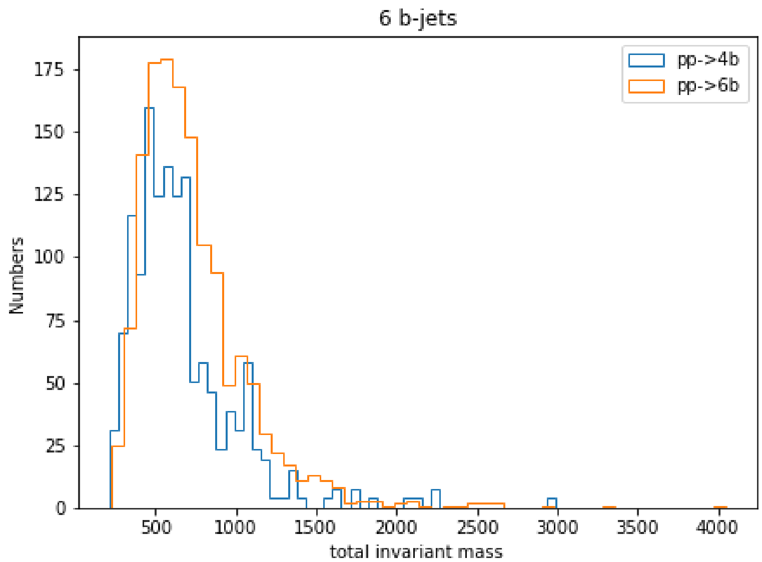


Figure 7: The distribution of total invariant mass of 6 b-jets for pp->4b and pp->6b events.

Table 6: 10,000 pp->6b events. Those events are generated in $\sqrt{s} = 14$ TeV with the cuts: $p_{\mathrm{T}} > 25$ GeV, $|\eta| < 2.5$ for b in MadGraph.

Cut	number of event pass
1	1408
2	1322
3	1052
4	822

Table 7: 100,000 signal events. Those events are generated in $\sqrt{s} = 14$ TeV.

Cut	number of event pass
1	21,814
2	21,254
3	17,130
4	14,142

11 Signal

Generate resonant channel gg->h3, h3->h2h, h2->hh in MadGraph by following commands:

```
import model cxSM_VLF_EFT
generate g g > h3, (h3 > h2 h, h2 > h h)
```

then check whether this channel is dominated in signal (gg->3h) or not.

I generate this channel and signal in $\sqrt{s} = 14$ TeV, the cross sections are 2.094 pb and 4.067 pb, respectively.

Figure 8 is the total invariant mass of 6 b-jets for signal events. Figure 9 is the total invariant mass of 6 b-jets for this resonant channel. From the results, there is a peak around 400 GeV, because the mass of h_3 is $m_3 = 420$ GeV.

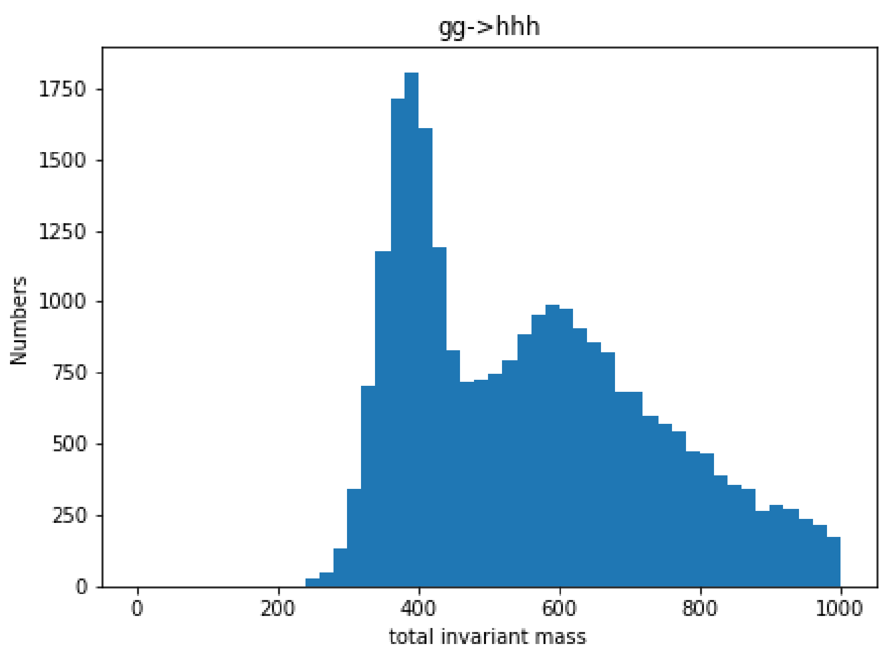


Figure 8: Total invariant mass of 6 b-jets for signal events.

A bump around 600 GeV in Figure 8? In the run card, some parameters are not set correctly. After setting all parameters correctly and regenerating these event, the results are in Figure 10. There is no bump around 600 GeV.

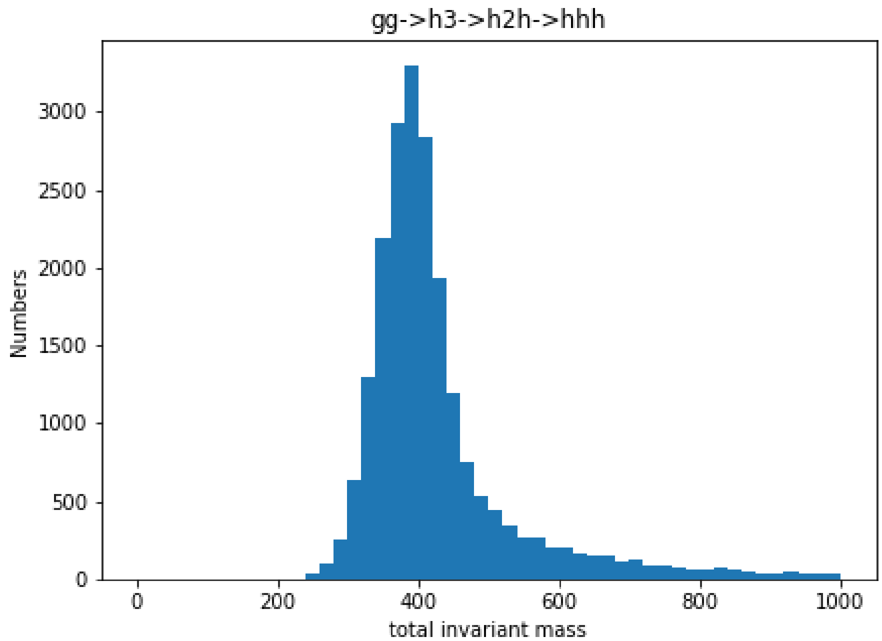


Figure 9: Total invariant mass of 6 b-jets for the resonant channel.

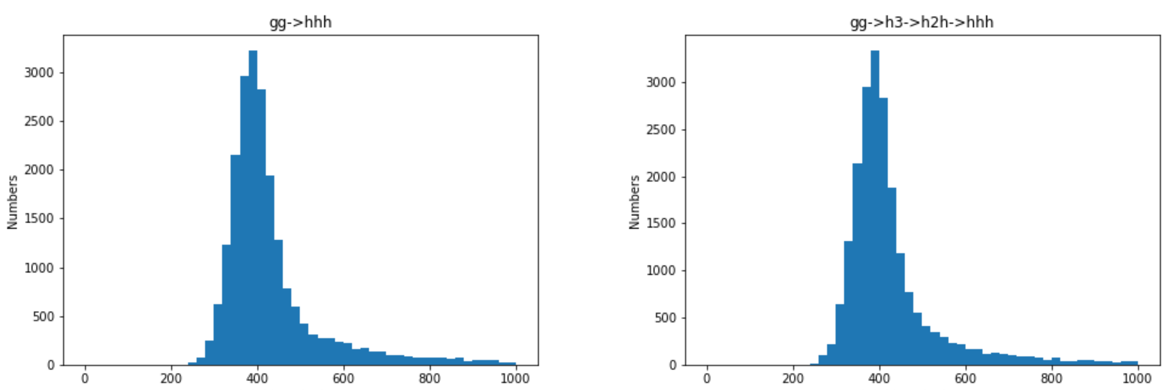


Figure 10: Total invariant mass of 6 b-jets for signal events and resonant channel (correct parameters).

12 The cross section in MadGraph

Table 8 is the cross sections calculated by MadGraph and in paper. They are different because in MadGraph we only consider LO. But in the paper, the numbers are quoted from the LHC Higgs Cross Section Working Group, and they have considered up to NNLO. The "k factor" is around 2.5, which accounts for the values in gg>h3.

Table 8: Cross section from MadGraph and paper

Process	σ MadGraph (fb)	σ in paper (fb)
g g > h3	21	55
gg > h3 > h h h	4.0	38.2

The problem of cross section: If we use the default run card to generate the decay process, the value of the cross section will be problematic.

Solution:

In madevent, use command `compute_widths` to compute decay widths of h_2, h_3, then replace the `run_card.dat` by `run_card_default.dat`.

Regenerate the signal and resonant event, the cross sections are 11.1 fb and 10.39 fb, respectively. This channel is indeed dominated in the signal.

13 Comparision for pp $\rightarrow$ hhh and gg $\rightarrow$ hhh

Generate pp->hhh in MadGraph by following commands:

```
import model cxSM_VLF_EFT
define p = p b b~
generate p p > h h h QCD<=8
```

Generate gg->hhh in MadGraph by following commands:

```
import model cxSM_VLF_EFT
generate g g > h h h
```

These events are generated in 14 TeV and the sample size is 100k. The cross section for pp->hhh and gg->hhh are 11.22 fb and 11.11 fb, respectively. They only differ by 1%.

14 Comparision for gg → hhh and resonant channel

To compare gg → hhh and resonant channel, I plot the p_{T}, η, and total invariant mass of 6 b-jets for both.

These events are generated in 14 TeV and the sample size is 100k. The cross section for gg->hhh and resonant channel are 11.11 fb and 10.38 fb, respectively.

The number of events of 6 b-jets for gg->hhh is 21,814 and for the resonant channel is 21,475. The numbers are very close. I scaled the number of events for gg->3h to be the same as the resonant channel.

Figure 11 is p_{T} and η distributions. Figure 12 is total invariant mass of 6 b-jets distributions. Their distributions look similar.

I apply the cuts same in Sec.10 and count how many events can pass these cuts. Table 9 is the result. The results for both events are similar.

Table 9: Number of events pass the selection cuts. The total number of events for gg->hhh and the resonant channel is 100,000.

Cut	gg->3h	resonant channel
1	21,814	21,475
2	21,254	20,898
3	17,130	16,828
4	14,142	13,730

15 Branching ratios and decay widths

The branching ratios and decay widths are calculated by MadGraph and Mathematica notebook. Mathematica notebook does not include uu dd ss ee decay modes. Since the contribution of these modes is very small.

Table 10 is calculated by MadGraph. Table 11 is calculated by Mathematica notebook.

The exact values from both are slightly different because MG can not correctly expand the parameter ϵ in the model.

Set the ϵ to 0.001 and calculate again. The results are in Table 12 and Table 13.

This time, the decay widths and branching ratios calculated by MG and Mathematica notebook become close.

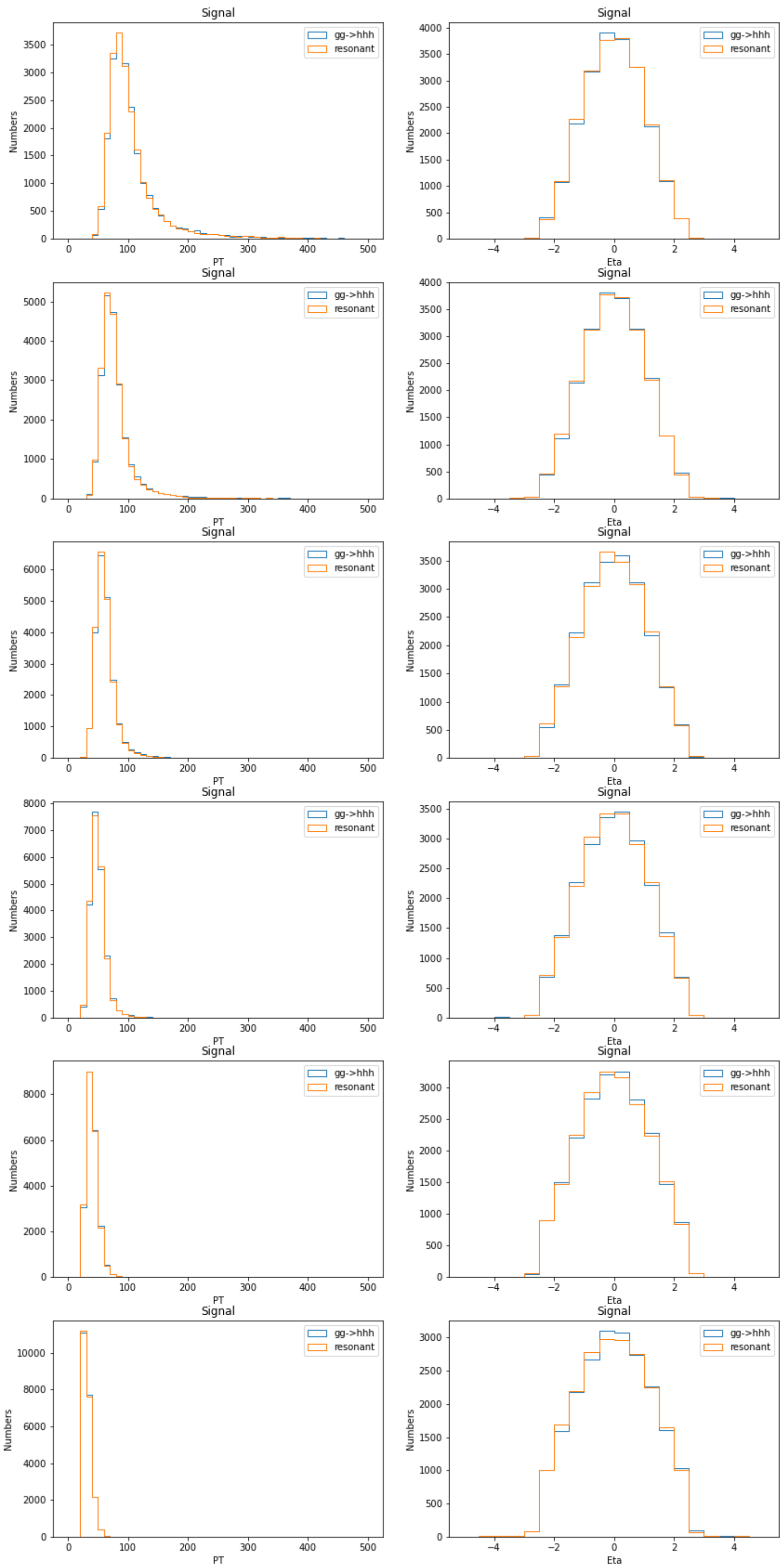


Figure 11: p_{T} and η distributions of b-jets for gg->3h and resonant channel. They are ordered by p_{T}.

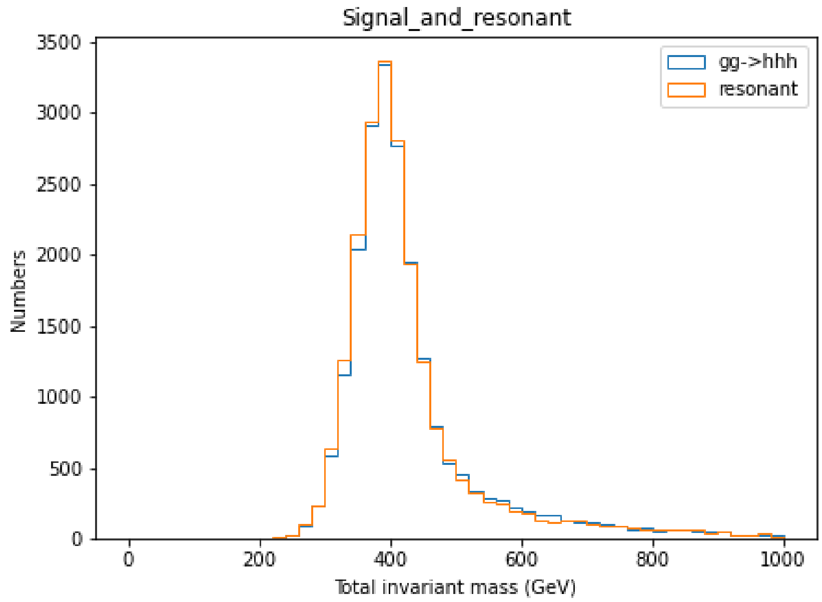


Figure 12: The distribution of total invariant mass of 6 b-jets for gg->3h and resonant channel.

Table 10: Decay widths and branching ratios calculated by MadGraph at BP1.

	BR	Width (GeV)		BR	Width (GeV)
$h_2 \to h_1 h_1$	0.7506	0.085616	$h_3 \to h_1 h_2$	0.6795	0.8414
$h_2 \to WW$	0.1734	0.019782	$h_3 \to h_1 h_1$	0.1711	0.21183
$h_2 \to ZZ$	0.07548	0.0086097	$h_3 \to WW$	0.08756	0.10842
$h_2 \to bb$	0.0004724	5.3887e-05	$h_3 \to ZZ$	0.04102	0.050792
$h_2 \to cc$	3.455e-05	3.9407e-06	$h_3 \to tt$	0.02065	0.025569
$h_2 \to \tau\tau$	2.254e-05	2.5714e-06	$h_3 \to gg$	8.662e-05	0.00010725
$h_2 \to ss$	2.185e-07	2.4927e-08	$h_3 \to bb$	8.158e-05	0.00010102
$h_2 \to gg$	1.561e-07	1.7811e-08	$h_3 \to cc$	5.961e-06	7.3813e-06
$h_2 \to \mu\mu$	7.972e-08	9.0933e-09	$h_3 \to \tau\tau$	3.89e-06	4.8167e-06
$h_2 \to \gamma\gamma$	6.495e-09	7.409e-10	$h_3 \to \gamma\gamma$	3.852e-07	4.7697e-07
$h_2 \to \gamma Z$	9.447e-10	1.0776e-10	$h_3 \to \gamma Z$	4.937e-08	6.1131e-08
$h_2 \to dd$	5.441e-10	6.207e-11	$h_3 \to ss$	3.77e-08	4.6686e-08
$h_2 \to uu$	1.393e-10	1.5889e-11	$h_3 \to \mu\mu$	1.375e-08	1.7031e-08
$h_2 \to ee$	1.865e-12	2.1269e-13	$h_3 \to dd$	9.389e-11	1.1625e-10
			$h_3 \to uu$	2.403e-11	2.976e-11
			$h_3 \to ee$	3.217e-13	3.9835e-13

Table 11: Decay widths and branching ratios calculated by Mathematica at BP1.

	BR	Width (GeV)		BR	Width (GeV)
$h_2 \to h_1h_1$	0.8321	0.1407	$h_3 \to h_1h_2$	0.8219	2.158
$h_2 \to WW$	0.1166	0.019721	$h_3 \to h_1h_1$	0.1049	0.27543
$h_2 \to ZZ$	0.05108	0.0086386	$h_3 \to WW$	0.04131	0.10846
$h_2 \to gg$	0.0001204	2.0359e-05	$h_3 \to ZZ$	0.01941	0.050962
$h_2 \to bb$	9.91e-05	1.6758e-05	$h_3 \to tt$	0.01242	0.032613
$h_2 \to \tau\tau$	1.526e-05	2.5802e-06	$h_3 \to gg$	6.364e-05	0.00016711
$h_2 \to cc$	4.59e-06	7.7613e-07	$h_3 \to bb$	1.195e-05	3.1386e-05
$h_2 \to \gamma\gamma$	3.161e-06	5.346e-07	$h_3 \to \tau\tau$	1.841e-06	4.8326e-06
$h_2 \to \gamma Z$	4.822e-07	8.154e-08	$h_3 \to cc$	5.536e-07	1.4536e-06
$h_2 \to \mu\mu$	5.43e-08	9.1826e-09	$h_3 \to \gamma\gamma$	1.823e-07	4.7857e-07
$h_2 \to tt$	4.355e-19	7.3649e-20	$h_3 \to \gamma Z$	2.334e-08	6.1277e-08
			$h_3 \to \mu\mu$	6.55e-09	1.7198e-08

Table 12: Decay widths and branching ratios calculated by MadGraph at BP1 with $\epsilon = 0.001$.

	BR	Width (GeV)		BR	Width (GeV)
$h_2 \to h_1h_1$	0.8311	1.4041e-05	$h_3 \to h_1h_2$	1.0	2.1493
$h_2 \to WW$	0.1174	1.9835e-06	$h_3 \to h_1h_1$	1.279e-05	2.748e-05
$h_2 \to ZZ$	0.0511	8.6329e-07	$h_3 \to WW$	5.058e-06	1.0871e-05
$h_2 \to bb$	0.0003198	5.4032e-09	$h_3 \to ZZ$	2.369e-06	5.0929e-06
$h_2 \to cc$	2.339e-05	3.9513e-10	$h_3 \to tt$	1.193e-06	2.5638e-06
$h_2 \to \tau\tau$	1.526e-05	2.5783e-10	$h_3 \to gg$	5.007e-09	1.0761e-08
$h_2 \to ss$	1.479e-07	2.4994e-12	$h_3 \to bb$	4.713e-09	1.0129e-08
$h_2 \to \mu\mu$	5.397e-08	9.1178e-13	$h_3 \to cc$	3.443e-10	7.4011e-10
$h_2 \to dd$	3.684e-10	6.2237e-15	$h_3 \to \tau\tau$	2.247e-10	4.8297e-10
$h_2 \to uu$	9.431e-11	1.5932e-15	$h_3 \to \gamma\gamma$	2.227e-11	4.7857e-11
$h_2 \to gg$	1.056e-11	1.7841e-16	$h_3 \to \gamma Z$	2.854e-12	6.1335e-12
$h_2 \to ee$	1.262e-12	2.1326e-17	$h_3 \to ss$	2.178e-12	4.6812e-12
$h_2 \to \gamma\gamma$	4.393e-13	7.4213e-18	$h_3 \to \mu\mu$	7.945e-13	1.7077e-12
$h_2 \to \gamma Z$	6.389e-14	1.0794e-18	$h_3 \to dd$	5.423e-15	1.1657e-14
			$h_3 \to uu$	1.388e-15	2.984e-15
			$h_3 \to ee$	1.858e-17	3.9942e-17

Table 13: Decay widths and branching ratios calculated by Mathematica at BP1 with $\epsilon = 0.001$.

	BR	Width (GeV)		BR	Width (GeV)
$h_2 \to h_1h_1$	0.8321	1.407e-05	$h_3 \to h_1h_2$	1.0	2.158
$h_2 \to WW$	0.1166	1.9721e-06	$h_3 \to h_1h_1$	1.276e-05	2.7543e-05
$h_2 \to ZZ$	0.05108	8.6386e-07	$h_3 \to WW$	5.026e-06	1.0846e-05
$h_2 \to gg$	0.0001204	2.0359e-09	$h_3 \to ZZ$	2.361e-06	5.0962e-06
$h_2 \to bb$	9.91e-05	1.6758e-09	$h_3 \to tt$	1.511e-06	3.2613e-06
$h_2 \to \tau\tau$	1.526e-05	2.5802e-10	$h_3 \to gg$	7.743e-09	1.6711e-08
$h_2 \to cc$	4.59e-06	7.7613e-11	$h_3 \to bb$	1.454e-09	3.1386e-09
$h_2 \to \gamma\gamma$	3.161e-06	5.346e-11	$h_3 \to \tau\tau$	2.239e-10	4.8326e-10
$h_2 \to \gamma Z$	4.822e-07	8.154e-12	$h_3 \to cc$	6.736e-11	1.4536e-10
$h_2 \to \mu\mu$	5.43e-08	9.1826e-13	$h_3 \to \gamma\gamma$	2.218e-11	4.7857e-11
$h_2 \to tt$	4.355e-15	7.3649e-20	$h_3 \to \gamma Z$	2.839e-12	6.1277e-12
			$h_3 \to \mu\mu$	7.969e-13	1.7198e-12

For color channels, still don't look the same. The reason is that in MadGraph only calculates to LO, but the numbers in the Mathematica notebook are calculated to NLO.

16 SPANet

Code: Symmetry Preserving Attention Networks

Symmetry Preserving Attention NETworks (Spa-Net) is used to do the jet assignment task. The jet assignment task is the identification of the original particle which leads to a reconstructed jet.

16.1 Prepare training data

1. Defining the event topology in `.ini` file. The structure of the `.ini` file follows this format:

```
[SOURCE]
FEATURE_1 = FEATURE_OPTION
FEATURE_2 = FEATURE_OPTION
FEATURE_3 = FEATURE_OPTION
...
```

```
[EVENT]
particles = (PARTICLE_1, PARTICLE_2, ...)
permutations = EVENT_SYMMETRY_GROUP

[PARTICLE_1]
jets = (JET_1, JET_2, ...)
permutations = JET_SYMMETRY_GROUP

[PARTICLE_2]
jets = (JET_1, JET_2, ...)
permutations = JET_SYMMETRY_GROUP

...
```

2. Create training dataset in HDF5 format.

3. Write option-file in JSON format.

16.2 Training

Training:

```
python train.py -of <OPTIONS_FILE> --log_dir <LOG_DIR>  --name <NAME> --gpus 1
```

<OPTIONS_FILE>: JSON file with option overloads. <LOG_DIR>: output directory. <NAME>: subdirectory name

Evaluation:

```
python test.py <log_directory> --gpu
```

<log_directory>: directory containing the checkpoint and options file.

```
python test.py <log_directory> -tf <TEST_FILE> --gpu
```

<TEST_FILE> will replace the test file in the option file.

Prediction:

```
python predict.py <log_directory> <output name> -tf <TEST_FILE> --gpu
```

17 SPANet for ttbar event

The training and testing datasets from here: Link
For full ttbar event:

- Training sample:
  - Total sample size: 10,009,520
  - 2t sample size: 2,967,955
- Testing sample size:
  - Total sample size: 358,946
  - 2t sample size: 116,342

Figure 13 is the training results for full ttbar events. The results are the same as the numbers given in the SPANet paper.

18 SPANet for di-Higgs event

Generate di-Higgs events in MadGraph by following commands:

```
import model 2HDMtII_NLO
generate p p > h2 [QCD] QED<=99 QCD<=99
```

then use the MadSpin let h2 decay to h1h1 and h1 decay to $b\bar{b}$. The mass of h2 is 1000 GeV.

The `.ini` file for di-Higgs event (h > b $\bar{\text{b}}$)

```
[SOURCE]
mass = log_normalize
pt = log_normalize
eta = normalize
phi = normalize

[EVENT]
particles = (h1, h2)
permutations = [(h1, h2)]

[h1]
```

Event Type: *t

Jet Limit	Event Proportion	Jet Proportion	Event Purity	T Purity
== 6	0.812	0.308	0.640	0.694
== 7	0.851	0.341	0.597	0.665
>= 8	0.881	0.351	0.523	0.610
Full	0.850	1.000	0.583	0.652

Event Type: 0t

Jet Limit	Event Proportion	Jet Proportion	Event Purity	T Purity
== 6	0.188	0.308	1.000	N/A
== 7	0.149	0.341	1.000	N/A
>= 8	0.119	0.351	1.000	N/A
Full	0.150	1.000	1.000	N/A

Event Type: 1t

Jet Limit	Event Proportion	Jet Proportion	Event Purity	T Purity
== 6	0.566	0.308	0.567	0.567
== 7	0.531	0.341	0.556	0.556
>= 8	0.484	0.351	0.523	0.523
Full	0.525	1.000	0.549	0.549

Event Type: 2t

Jet Limit	Event Proportion	Jet Proportion	Event Purity	T Purity
== 6	0.246	0.308	0.807	0.841
== 7	0.320	0.341	0.666	0.755
>= 8	0.397	0.351	0.523	0.663
Full	0.324	1.000	0.637	0.735

Figure 13: The training result for full ttbar events.

```
jets = (b1, b2)
permutations = [(b1, b2)]

[h2]
jets = (b1, b2)
permutations = [(b1, b2)]
```

There are three types of events, 0h, 1h, and 2h. The number means how many identifiable Higgs is in an event.

Definition of some parameters:

- Event proportion:

$$\text{Event proportion} \equiv \frac{\text{number of some type events with } i \text{ jets}}{\text{number of events with } i \text{ jets}} \tag{1}$$

- Jet proportion:

$$\text{Jet proportion} \equiv \frac{\text{number of events with } i \text{ jets}}{\text{total number of events}} \tag{2}$$

- Event purity:

$$\epsilon^{\text{event}} \equiv \frac{\text{number of some type events with and all Higgs are correctly identified}}{\text{number of some type events}} \tag{3}$$

- H purity:

$$\epsilon^{\text{h}} \equiv \frac{\text{number of correctly identified Higgs in some type events}}{\text{number of identifiable Higgs in some type events}} \tag{4}$$

18.1 Training sample for di-Higgs events

Di-Higgs events pp->h2, h2->hh, h->bb was generated at $\sqrt{s} = 14$ TeV using MadGraph. Then pass these events to Pythia8 for showering and hadronization. Then pass to Delphes for detector simulation.

Jets are reconstructed by the anti-kT algorithm with radius $R = 0.5$ and are required to have $p_{\text{T}} \geq 20$ GeV.

Preselection requirements: ≥ 4 jets in $|\eta| < 2.5$.

Define the correct jet assignments by matching them to the simulated truth quarks within an angular distance of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$.

18.2 Training result for di-Higgs event

For 100k di-Higgs event with $m_{\mathrm{h}_2} = 1000$ GeV:

- Training sample:
  - Total sample size: 78,785
  - 2h sample size: 36,765
  - 5% used on validation
- Testing sample:
  - Total sample size: 8,753
  - 2h sample size: 4,130

Event Type: *h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 4	0.887	0.291	0.945	0.960
== 5	0.925	0.331	0.908	0.940
>= 6	0.947	0.378	0.792	0.865
Full	0.922	1.000	0.874	0.914

Event Type: 0h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 4	0.113	0.291	1.000	N/A
== 5	0.075	0.331	1.000	N/A
>= 6	0.053	0.378	1.000	N/A
Full	0.078	1.000	1.000	N/A

Event Type: 1h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 4	0.545	0.291	0.911	0.911
== 5	0.442	0.331	0.888	0.888
>= 6	0.385	0.378	0.827	0.827
Full	0.450	1.000	0.876	0.876

Event Type: 2h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 4	0.342	0.291	1.000	1.000
== 5	0.483	0.331	0.926	0.963
>= 6	0.562	0.378	0.769	0.878
Full	0.472	1.000	0.871	0.933

Figure 14: The training result for 100k di-Higgs events.

Figure 14 is the training results for 100k di-Higgs events. For 2h events, $\epsilon^{\mathrm{event}} = 0.871, \epsilon^{\mathrm{h}} = 0.933$.

For 1M di-Higgs event with $m_{\mathrm{h}_2} = 1000$ GeV:

- Training sample:
  - Total sample size: 788,160
  - 2h sample size: 364,773
  - 5% used on validation
- Testing sample:
  - Total sample size: 87,702
  - 2h sample size: 40,695

Figure 15 is the training results for 1M di-Higgs events. For 2h events, $\epsilon^{\mathrm{event}} = 0.914, \epsilon^{\mathrm{h}} = 0.955$.

For di-Higgs event with $m_{\mathrm{h}_2} = 500$ GeV:

- Training sample:
  - Total sample size: 764,676
  - 2h sample size: 396,588
  - 5% used on validation
- Testing sample:
  - Total sample size: 84,687
  - 2h sample size: 43,818

Figure 16 is the training results for 500 GeV di-Higgs events. For 2h events, $\epsilon^{\mathrm{event}} = 0.779, \epsilon^{\mathrm{h}} = 0.871$.

18.3 Apply the pre-trained model on different $m_{\mathbf{h}_2}$ sample

The SPANet models in Sec.18.2 are trained on $m_{\mathrm{h}_2} = 500$ GeV and $m_{\mathrm{h}_2} = 1000$ GeV. Test these models on samples with $m_{\mathrm{h}_2} = 500$ GeV, 750 GeV, 1000 GeV.

The results are summarized in Table 14.

Event Type: *h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 4	0.877	0.301	0.950	0.964
== 5	0.920	0.326	0.919	0.946
>= 6	0.944	0.372	0.831	0.891
Full	0.916	1.000	0.894	0.928

Event Type: 0h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 4	0.123	0.301	1.000	N/A
== 5	0.080	0.326	1.000	N/A
>= 6	0.056	0.372	1.000	N/A
Full	0.084	1.000	1.000	N/A

Event Type: 1h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 4	0.533	0.301	0.917	0.917
== 5	0.445	0.326	0.885	0.885
>= 6	0.394	0.372	0.813	0.813
Full	0.452	1.000	0.873	0.873

Event Type: 2h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 4	0.345	0.301	1.000	1.000
== 5	0.475	0.326	0.950	0.975
>= 6	0.551	0.372	0.844	0.918
Full	0.464	1.000	0.914	0.955

Figure 15: The training result for 1M di-Higgs events.

Table 14: SPANet models test on different m_{h_2} samples.

Training m_{h_2} (GeV)	Testing m_{h_2} (GeV)	Event purity	H purity
500	500	0.779	0.871
500	750	0.408	0.672
500	1000	0.390	0.665
1000	500	0.383	0.638
1000	750	0.731	0.860
1000	1000	0.914	0.955

Event Type: *h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 4	0.872	0.319	0.879	0.914
== 5	0.923	0.325	0.802	0.868
>= 6	0.944	0.356	0.633	0.753
Full	0.914	1.000	0.763	0.836

Event Type: 0h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 4	0.128	0.319	1.000	N/A
== 5	0.077	0.325	1.000	N/A
>= 6	0.056	0.356	1.000	N/A
Full	0.086	1.000	1.000	N/A

Event Type: 1h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 4	0.497	0.319	0.793	0.793
== 5	0.391	0.325	0.748	0.748
>= 6	0.312	0.356	0.666	0.666
Full	0.397	1.000	0.743	0.743

Event Type: 2h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 4	0.375	0.319	0.994	0.994
== 5	0.532	0.325	0.841	0.912
>= 6	0.632	0.356	0.617	0.775
Full	0.517	1.000	0.779	0.871

Figure 16: The training result for di-Higgs events with $m_{\mathrm{h}_2} = 500$ GeV.

19 SPANet for tri-Higgs event

The tri-Higgs events are signal events (gg->hhh) generated in the previous.

19.1 Training sample for tri-Higgs event

Tri-Higgs events gg->hhh was generated at $\sqrt{s} = 14$ TeV using MadGraph. Then pass these events to Pythia8 for showering and hadronization. In Pythia8, the Higgs is forced to decay to $\mathrm{b\overline{b}}$. Then pass to Delphes for detector simulation.

Jets are reconstructed by the anti-kT algorithm with radius $R = 0.5$ and are required to have $p_{\mathrm{T}} \geq 20$ GeV.

Preselection requirements: ≥ 6 jets in $|\eta| < 2.5$.

Define the correct jet assignments by matching them to the simulated truth quarks within an angular distance of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$.

19.2 Training result for tri-Higgs event

For 1M tri-Higgs event:

- Training sample:
  - Total sample size: 542,681
  - 3h sample size: 104,467
  - 5% used on validation
- Testing sample:
  - Total sample size: 60,173
  - 3h sample size: 11,598

Figure 17 is the training results for 1M tri-Higgs events. For 3h events, $\epsilon^{\mathrm{event}} = 0.344, \epsilon^{\mathrm{h}} = 0.538$.

Event Type: 0h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 6	0.186	0.439	1.000	N/A
== 7	0.147	0.313	1.000	N/A
>= 8	0.119	0.248	1.000	N/A
Full	0.157	1.000	1.000	N/A

Event Type: 1h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 6	0.386	0.439	0.452	0.452
== 7	0.337	0.313	0.405	0.405
>= 8	0.312	0.248	0.334	0.334
Full	0.353	1.000	0.412	0.412

Event Type: 2h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 6	0.303	0.439	0.293	0.489
== 7	0.298	0.313	0.228	0.429
>= 8	0.285	0.248	0.152	0.343
Full	0.297	1.000	0.239	0.435

Event Type: 3h

Jet Limit	Event Proportion	Jet Proportion	Event Purity	H Purity
== 6	0.124	0.439	0.603	0.705
== 7	0.218	0.313	0.336	0.546
>= 8	0.283	0.248	0.152	0.400
Full	0.193	1.000	0.344	0.538

Figure 17: The training result for 1M tri-Higgs events.

20 χ^2 method

20.1 Di-Higgs

Try all possible combinations of final jets and find the minimal mass difference between SM Higgs mass, i.e. minimize this

$$\chi^2 = [M(j_1 j_2) - M_{\mathrm{H}}]^2 + [M(j_3 j_4) - M_{\mathrm{H}}]^2 \tag{5}$$

where $M(j_i j_j)$ means the invariant mass of jet i and jet j.

The results are presented in Table 15.

Table 15: Performance comparison for di-Higgs events.

	Event	SPANet Efficiency		χ^2 Efficiency	
N_{Jet}	Fraction	Event	Higgs	Event	Higgs
$= 4$	0.224	1.000	1.000	1.000	1.000
$= 5$	0.334	0.950	0.975	0.675	0.737
≥ 6	0.442	0.844	0.918	0.313	0.441
Total	1.000	0.914	0.955	0.582	0.665

20.2 Di-Higgs with b-tag

Similarly, try to minimize (5), but only try all possible permutations of b-tagged jets. Here, we only consider the events with ≥ 4 b-jets.

The results are presented in Table 16 and Table 17.

Table 16: Performance comparison for di-Higgs events with ≥ 4 b-jets (b-tagging efficiency $= 100\%$).

	Event	SPANet Efficiency		χ^2 Efficiency	
N_{Jet}	Fraction	Event	Higgs	Event	Higgs
$= 4$	0.224	1.000	1.000	1.000	1.000
$= 5$	0.334	0.980	0.990	0.974	0.979
≥ 6	0.442	0.946	0.972	0.916	0.933
Total	1.000	0.969	0.984	0.954	0.963

Table 17: Performance comparison for di-Higgs events with ≥ 4 b-jets (b-tagging efficiency = default).

	Event	SPANet Efficiency		χ^2 Efficiency	
N_{Jet}	Fraction	Event	Higgs	Event	Higgs
$= 4$	0.200	1.000	1.000	1.000	1.000
$= 5$	0.318	0.950	0.975	0.892	0.914
≥ 6	0.482	0.865	0.930	0.725	0.776
Total	1.000	0.919	0.958	0.833	0.865

20.3 Tri-Higgs

Try all possible combinations of final jets and find the minimal mass difference between SM Higgs mass, i.e. minimize this

$$\chi^2 = [M(j_1 j_2) - M_{\text{H}}]^2 + [M(j_3 j_4) - M_{\text{H}}]^2 + [M(j_5 j_6) - M_{\text{H}}]^2 \tag{6}$$

where $M(j_i j_j)$ means the invariant mass of jet i and jet j.

Testing samples are required with ≥ 0 b-jets and b-tagging efficiency: default. The results are presented in Table 18.

Table 18: Performance comparison for tri-Higgs events. The results are for complete events (3 Higgs).

	Event	SPANet Efficiency		χ^2 Efficiency	
N_{Jet}	Fraction	Event	Higgs	Event	Higgs
$= 6$	0.062	0.580	0.689	0.515	0.559
$= 7$	0.076	0.382	0.596	0.178	0.228
≥ 8	0.081	0.238	0.465	0.043	0.133
Total	0.291	0.384	0.564	0.223	0.303

Below, only consider the permutation of b-tagged jets. Testing events are required with ≥ 6 b-jets. The results are presented in Table 19 and Table 20.

21 SPANet for di-Higgs events with b-tag

In this section, SPANet will take the b-tagging information.

The `.ini` file for the di-Higgs event

Table 19: Performance comparison for tri-Higgs events with ≥ 6 b-jets (b-tagging efficiency = 100%). The results are for complete events (3 Higgs).

	Event	SPANet Efficiency		χ^2 Efficiency	
N_{Jet}	Fraction	Event	Higgs	Event	Higgs
$= 6$	0.150	0.633	0.729	0.517	0.560
$= 7$	0.183	0.550	0.675	0.459	0.512
≥ 8	0.196	0.443	0.597	0.363	0.425
Total	0.529	0.534	0.662	0.440	0.494

Table 20: Performance comparison for tri-Higgs events with ≥ 6 b-jets (b-tagging efficiency = default). The results are for complete events (3 Higgs).

	Event	SPANet Efficiency		χ^2 Efficiency	
N_{Jet}	Fraction	Event	Higgs	Event	Higgs
$= 6$	0.100	0.488	0.613	0.484	0.535
$= 7$	0.175	0.329	0.502	0.347	0.413
≥ 8	0.257	0.185	0.394	0.173	0.254
Total	0.532	0.290	0.417	0.289	0.359

```
[SOURCE]
mass = log_normalize
pt = log_normalize
eta = normalize
phi = normalize
btag = none

[EVENT]
particles = (h1, h2)
permutations = [(h1, h2)]

[h1]
jets = (b1, b2)
permutations = [(b1, b2)]

[h2]
jets = (b1, b2)
```

```
permutations = [(b1, b2)]
```

For training data preparation is almost the same in Sec.18.1. 2 b-tagged jets or 4 b-tagged jets will be required in the following training data. The default b-tagging efficiency in Delphes is

$$\text{eff.} = 0.85 \tanh(0.0025 p_{\mathrm{T}}) \frac{25.0}{1 + 0.063 p_{\mathrm{T}}} \tag{7}$$

21.1 Training result for di-Higgs event

Case 1: ≥ 0 b-tagged jets, b-tagging efficiency: 100%

- Training sample:
  - Total sample size: 788,108
  - 1h sample size: 357,082
  - 2h sample size: 364,717
  - 5% used on validation
- Testing sample:
  - Total sample size: 87,748
  - 1h sample size: 39,816
  - 2h sample size: 40,635

The training results for Case 1 is presented in Table 21.

Table 21: SPA-NET results on di-Higgs with at least 0 b-tagged jets (b-tagging efficiency: 100%).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 4$	0.224	1.000	1.000
$= 5$	0.333	0.981	0.990
≥ 6	0.443	0.948	0.973
Total	1.000	0.970	0.985

Case 2: ≥ 2 b-tagged jets, b-tagging efficiency: 100%

- Training sample:
  - Total sample size: 787,097
  - 1h sample size: 357,078
  - 2h sample size: 364,717
  - 5% used on validation
- Testing sample:
  - Total sample size: 87,641
  - 1h sample size: 39,816
  - 2h sample size: 40,635

The training results for Case 2 is presented in Table 22.

Table 22: SPA-NET results on di-Higgs with at least 2 b-tagged jets (b-tagging efficiency: 100%).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 4$	0.224	1.000	1.000
$= 5$	0.333	0.984	0.992
≥ 6	0.443	0.953	0.976
Total	1.000	0.974	0.987

Case 3: ≥ 4 b-tagged jets, b-tagging efficiency: 100%

- Training sample:
  - Total sample size: 454,768
  - 1h sample size: 85,074
  - 2h sample size: 363,904
  - 5% used on validation
- Testing sample:
  - Total sample size: 50,735

- 1h sample size: 9,547
- 2h sample size: 40,541

The training results for Case 3 is presented in Table 23.

Table 23: SPA-NET results on di-Higgs with at least 4 b-tagged jets (b-tagging efficiency: 100%).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 4$	0.224	1.000	1.000
$= 5$	0.333	0.980	0.990
≥ 6	0.443	0.946	0.972
Total	1.000	0.969	0.984

Case 4: ≥ 0 b-tagged jets, b-tagging efficiency: default

- Training sample:
  - Total sample size: 788,160
  - 1h sample size: 357,137
  - 2h sample size: 364,773
  - 5% used on validation
- Testing sample:
  - Total sample size: 87,702
  - 1h sample size: 39,659
  - 2h sample size: 40,695

The training results for Case 4 is presented in Table 24.

Case 5: ≥ 2 b-tagged jets, b-tagging efficiency: default

- Training sample:
  - Total sample size: 633,726
  - 1h sample size: 270,325
  - 2h sample size: 328,427

Table 24: SPA-NET results on di-Higgs with at least 0 b-tagged jets (b-tagging efficiency: default)

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 4$	0.224	1.000	1.000
$= 5$	0.334	0.961	0.981
≥ 6	0.442	0.887	0.941
Total	1.000	0.937	0.968

  - 5% used on validation

- Testing sample:

  - Total sample size: 70,436
  - 1h sample size: 29,847
  - 2h sample size: 36,742

The training results for Case 5 is presented in Table 25.

Table 25: SPA-NET results on di-Higgs with at least 2 b-tagged jets (b-tagging efficiency: default)

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 4$	0.223	1.000	1.000
$= 5$	0.332	0.964	0.982
≥ 6	0.445	0.891	0.943
Total	1.000	0.939	0.969

Case 6: ≥ 4 b-tagged jets, b-tagging efficiency: default

- Training sample:

  - Total sample size: 108,266
  - 1h sample size: 19,808
  - 2h sample size: 87,196
  - 5% used on validation

- Testing sample:

– Total sample size: 11,992

– 1h sample size: 2,205

– 2h sample size: 9,643

The training results for Case 6 is presented in Table 26.

Table 26: SPA-NET results on di-Higgs with at least 4 b-tagged jets (b-tagging efficiency: default)

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$=4$	0.200	1.000	1.000
$=5$	0.318	0.950	0.975
≥ 6	0.482	0.865	0.930
Total	1.000	0.919	0.958

Table 27 is the summary for b-tag training results.

Table 27: Summary for b-tag training results. 2b means at least 2 b-tagged jets. 100% and default are the b-tagging efficiencies in Delphes.

Case	Event Efficiency	Higgs Efficiency	Testing sample
no b-tag	0.914 ± 0.0014	0.955 ± 0.0007	40,695
≥ 0 b 100%	0.970 ± 0.0008	0.985 ± 0.0004	40,635
≥ 2 b 100%	0.974 ± 0.0008	0.987 ± 0.0004	40,635
≥ 4 b 100%	0.969 ± 0.0009	0.984 ± 0.0004	40,541
≥ 0 b default	0.937 ± 0.0012	0.968 ± 0.0006	40,695
≥ 2 b default	0.939 ± 0.0012	0.969 ± 0.0006	36,742
≥ 4 b default	0.919 ± 0.0028	0.958 ± 0.0014	9,643

22 SPANet for tri-Higgs events with b-tag

The `.ini` file for the tri-Higgs event

```
[SOURCE]
mass = log_normalize
pt = log_normalize
eta = normalize
```

```
phi = normalize
btag = none

[EVENT]
particles = (h1, h2, h3)
permutations = [(h1, h2, h3)]

[h1]
jets = (b1, b2)
permutations = [(b1, b2)]

[h2]
jets = (b1, b2)
permutations = [(b1, b2)]

[h3]
jets = (b1, b2)
permutations = [(b1, b2)]
```

22.1 Training sample preparation for tri-Higgs events

Tri-Higgs events gg->hhh was generated at $\sqrt{s} = 14$ TeV using MadGraph. Then pass these events to Pythia8 for showering and hadronization. In Pythia8, the Higgs is forced to decay to $\mathrm{b}\overline{\mathrm{b}}$. Then pass to Delphes for detector simulation.

Jets are reconstructed by the anti-kT algorithm with radius $R = 0.4$ and are required to have $p_{\mathrm{T}} \geq 20$ GeV.

Preselection requirements: ≥ 6 jets in $|\eta| < 2.5$.

Define the correct jet assignments by matching them to the simulated truth quarks within an angular distance of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$.

22.2 Trainging result for tri-Higgs events

In all cases, partial event training is set to true.

Case 1: ≥ 0 b-tagged jets, b-tagging efficiency: 100%

- Training sample:

- Total sample size: 519,372
- 1h sample size: 168,385
- 2h sample size: 175,616
- 3h sample size: 112,770
- 5% used on validation

- Testing sample:
  - Total sample size: 57,652
  - 1h sample size: 18,639
  - 2h sample size: 19,456
  - 3h sample size: 12,679

The training results for Case 1 is presented in Table 28.

Table 28: SPA-NET results on tri-Higgs with at least 0 b-tagged jets (b-tagging efficiency: 100%).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.062	0.608	0.711
$= 7$	0.076	0.519	0.652
≥ 8	0.081	0.426	0.584
Total	0.220	0.510	0.644

Case 2: ≥ 2 b-tagged jets, b-tagging efficiency: 100%

- Training sample:
  - Total sample size: 519,318
  - 1h sample size: 168,385
  - 2h sample size: 175,616
  - 3h sample size: 112,770
  - 5% used on validation
- Testing sample:

- Total sample size: 57,646
- 1h sample size: 18,639
- 2h sample size: 19,456
- 3h sample size: 12,679

The training results for Case 2 is presented in Table 29.

Table 29: SPA-NET results on tri-Higgs with at least 2 b-tagged jets (b-tagging efficiency: 100%).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.063	0.637	0.732
$= 7$	0.076	0.556	0.679
≥ 8	0.081	0.442	0.601
Total	0.220	0.537	0.665

Case 3: ≥ 4 b-tagged jets, b-tagging efficiency: 100%

- Training sample:
  - Total sample size: 503,404
  - 1h sample size: 164,104
  - 2h sample size: 175,572
  - 3h sample size: 112,770
  - 5% used on validation
- Testing sample:
  - Total sample size: 55,989
  - 1h sample size: 18,197
  - 2h sample size: 19,454
  - 3h sample size: 12,679

The training results for Case 3 is presented in Table 30.

Table 30: SPA-NET results on tri-Higgs with at least 4 b-tagged jets (b-tagging efficiency: 100%).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.064	0.635	0.732
$= 7$	0.079	0.563	0.684
≥ 8	0.084	0.443	0.602
Total	0.226	0.539	0.668

Case 4: ≥ 6 b-tagged jets, b-tagging efficiency: 100%

- Training sample:
  - Total sample size: 212,885
  - 1h sample size: 5,349
  - 2h sample size: 33,675
  - 3h sample size: 112,330
  - 5% used on validation
- Testing sample:
  - Total sample size: 23,886
  - 1h sample size: 3,853
  - 2h sample size: 6,804
  - 3h sample size: 12,639

The training results for Case 4 is presented in Table 31.

Table 31: SPA-NET results on tri-Higgs with at least 6 b-tagged jets (b-tagging efficiency: 100%).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.150	0.633	0.729
$= 7$	0.183	0.550	0.675
≥ 8	0.196	0.443	0.597
Total	0.529	0.534	0.662

Case 5: ≥ 0 b-tagged jets, b-tagging efficiency: default

- Training sample:
  - Total sample size: 519,404
  - 1h sample size: 168,781
  - 2h sample size: 175,235
  - 3h sample size: 112,740
  - 5% used on validation
- Testing sample:
  - Total sample size: 57,744
  - 1h sample size: 18,447
  - 2h sample size: 19,572
  - 3h sample size: 12,655

The training results for Case 5 is presented in Table 32.

Table 32: SPA-NET results on tri-Higgs with at least 0 b-tagged jets (b-tagging efficiency: default).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.062	0.580	0.689
$= 7$	0.076	0.382	0.569
≥ 8	0.081	0.238	0.465
Total	0.219	0.384	0.564

Case 6: ≥ 2 b-tagged jets, b-tagging efficiency: default

- Training sample:
  - Total sample size: 486,815
  - 1h sample size: 155,932
  - 2h sample size: 166,234
  - 3h sample size: 109,947

– 5% used on validation

- Testing sample:
  - Total sample size: 54,132
  - 1h sample size: 17,102
  - 2h sample size: 18,522
  - 3h sample size: 12,321

The training results for Case 6 is presented in Table 33.

Table 33: SPA-NET results on tri-Higgs with at least 2 b-tagged jets (b-tagging efficiency: default).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.064	0.575	0.683
$= 7$	0.079	0.386	0.573
≥ 8	0.085	0.236	0.460
Total	0.228	0.383	0.562

Case 7: ≥ 4 b-tagged jets, b-tagging efficiency: default

- Training sample:
  - Total sample size: 238,155
  - 1h sample size: 16,272
  - 2h sample size: 65,770
  - 3h sample size: 72,441
  - 5% used on validation
- Testing sample:
  - Total sample size: 26,561
  - 1h sample size: 7,261
  - 2h sample size: 9,300
  - 3h sample size: 8,124

The training results for Case 7 is presented in Table 34.

Table 34: SPA-NET results on tri-Higgs with at least 4 b-tagged jets (b-tagging efficiency: default).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.081	0.553	0.668
$= 7$	0.105	0.373	0.559
≥ 8	0.120	0.236	0.460
Total	0.306	0.367	0.549

Case 8: ≥ 6 b-tagged jets, b-tagging efficiency: default

- Training sample:
  - Total sample size: 20,335
  - 1h sample size: 3,300
  - 2h sample size: 5,639
  - 3h sample size: 10,840
  - 5% used on validation
- Testing sample:
  - Total sample size: 2,166
  - 1h sample size: 357
  - 2h sample size: 592
  - 3h sample size: 1,153

The training results for Case 8 is presented in Table 35.

Table 35: SPA-NET results on tri-Higgs with at least 6 b-tagged jets (b-tagging efficiency: default).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.100	0.488	0.613
$= 7$	0.175	0.329	0.502
≥ 8	0.257	0.185	0.394
Total	0.532	0.290	0.417

Case 9: Without b-tagging information.

- Training sample:
  - Total sample size: 519,404
  - 1h sample size: 168,781
  - 2h sample size: 175,235
  - 3h sample size: 112,740
  - 5% used on validation
- Testing sample:
  - Total sample size: 57,744
  - 1h sample size: 18,447
  - 2h sample size: 19,572
  - 3h sample size: 12,655

The training results for Case 9 is presented in Table 36.

Table 36: SPA-NET results on tri-Higgs without b-tagging information.

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.062	0.600	0.703
$= 7$	0.076	0.343	0.557
≥ 8	0.081	0.172	0.415
Total	0.219	0.351	0.545

Case 10: ≥ 6 b-tagged jets, b-tagging efficiency: default

- Training sample:
  - Total sample size: 200,020
  - 1h sample size: 32,269
  - 2h sample size: 55,500
  - 3h sample size: 106,820
  - 5% used on validation

- Testing sample:
  - Total sample size: 22,351
  - 1h sample size: 3,639
  - 2h sample size: 6,161
  - 3h sample size: 11,908

The training results for Case 10 is presented in Table 37.

Table 37: SPA-NET results on tri-Higgs with at least 6 b-tagged jets (b-tagging efficiency: default).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.099	0.631	0.729
$= 7$	0.164	0.435	0.604
≥ 8	0.269	0.272	0.481
Total	0.533	0.389	0.565

Case 11: ≥ 0 b-tagged jets, b-tagging efficiency: default

- Training sample:
  - Total sample size: 2,597,986
  - 1h sample size: 842,297
  - 2h sample size: 877,590
  - 3h sample size: 564,989
  - 5% used on validation
- Testing sample:
  - Total sample size: 289,131
  - 1h sample size: 93,303
  - 2h sample size: 98,004
  - 3h sample size: 62,944

The training results for Case 11 is presented in Table 38.

Table 38: SPA-NET results on tri-Higgs with at least 0 b-tagged jets (b-tagging efficiency: default).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.060	0.642	0.735
$= 7$	0.075	0.445	0.619
≥ 8	0.081	0.293	0.511
Total	0.218	0.443	0.611

Case 12: ≥ 0 b-tagged jets, b-tagging efficiency: default

- Training sample:
  - Total sample size: 5,194,850
  - 1h sample size: 1,681,832
  - 2h sample size: 1,754,176
  - 3h sample size: 1,132,044
  - 5% used on validation
- Testing sample:
  - Total sample size: 578,248
  - 1h sample size: 186,731
  - 2h sample size: 195,461
  - 3h sample size: 126,154

The training results for Case 12 is presented in Table 39.

Table 39: SPA-NET results on tri-Higgs with at least 0 b-tagged jets (b-tagging efficiency: default).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.061	0.669	0.755
$= 7$	0.075	0.478	0.643
≥ 8	0.082	0.322	0.531
Total	0.218	0.473	0.632

Case 13: ≥ 0 b-tagged jets, b-tagging efficiency: default

- Training sample:
  - Total sample size: 1,558,455
  - 1h sample size: 505,734
  - 2h sample size: 525,989
  - 3h sample size: 339,012
  - 5% used on validation
- Testing sample:
  - Total sample size: 173,474
  - 1h sample size: 55,738
  - 2h sample size: 58,879
  - 3h sample size: 37,864

The training results for Case 13 is presented in Table 40.

Table 40: SPA-NET results on tri-Higgs with at least 0 b-tagged jets (b-tagging efficiency: default).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.061	0.660	0.750
$= 7$	0.075	0.447	0.620
≥ 8	0.082	0.281	0.500
Total	0.218	0.444	0.611

Case 14: ≥ 0 b-tagged jets, b-tagging efficiency: default

- Training sample:
  - Total sample size: 3,636,395
  - 1h sample size: 1,178,817
  - 2h sample size: 1,227,899
  - 3h sample size: 791,457

    - 5% used on validation

- Testing sample:
    - Total sample size: 404,773
    - 1h sample size: 130,462
    - 2h sample size: 136,786
    - 3h sample size: 88,466

The training results for Case 14 is presented in Table 41.

Table 41: SPA-NET results on tri-Higgs with at least 0 b-tagged jets (b-tagging efficiency: default).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.061	0.685	0.767
$= 7$	0.075	0.462	0.634
≥ 8	0.082	0.301	0.516
Total	0.218	0.464	0.627

Case 15: ≥ 0 b-tagged jets, b-tagging efficiency: default

- Training sample:
    - Total sample size: 7,791,375
    - 1h sample size: 2,523,958
    - 2h sample size: 2,629,911
    - 3h sample size: 1,697,675
    - 5% used on validation

- Testing sample:
    - Total sample size: 866,656
    - 1h sample size: 280,093
    - 2h sample size: 293,008
    - 3h sample size: 188,956

The training results for Case 15 is presented in Table 42.

Table 42: SPA-NET results on tri-Higgs with at least 0 b-tagged jets (b-tagging efficiency: default).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.061	0.711	0.785
$= 7$	0.075	0.501	0.662
≥ 8	0.082	0.337	0.545
Total	0.218	0.497	0.652

Case 16: ≥ 0 b-tagged jets, b-tagging efficiency: default

- Training sample:
  - Total sample size: 10,388,500
  - 1h sample size: 3,364,562
  - 2h sample size: 3,507,744
  - 3h sample size: 2,263,437
  - 5% used on validation
- Testing sample:
  - Total sample size: 1,155,542
  - 1h sample size: 374,000
  - 2h sample size: 390,507
  - 3h sample size: 251,870

The training results for Case 16 is presented in Table 43.

Table 43: SPA-NET results on tri-Higgs with at least 0 b-tagged jets (b-tagging efficiency: default).

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.061	0.718	0.790
$= 7$	0.075	0.510	0.666
≥ 8	0.082	0.344	0.550
Total	0.218	0.505	0.656

Table 44 is the summary for tri-Higgs training results.

Figure 18 is the accuracy against 3h training sample size for 0b default case.

Table 44: Summary for tri-Higgs training results. Tagging eff. is the b-tagging efficiency in Delphes.

Case	$N_{\rm bjet}$	Tagging eff.	Event Efficiency	Higgs Efficiency	Testing sample
1	≥ 0	100%	0.510 ± 0.0044	0.644 ± 0.0025	12,679
2	≥ 2	100%	0.537 ± 0.0044	0.665 ± 0.0024	12,679
3	≥ 4	100%	0.539 ± 0.0044	0.668 ± 0.0024	12,679
4	≥ 6	100%	0.534 ± 0.0044	0.662 ± 0.0024	12,639
5	≥ 0	default	0.384 ± 0.0043	0.564 ± 0.0025	12,655
6	≥ 2	default	0.383 ± 0.0044	0.562 ± 0.0026	12,321
7	≥ 4	default	0.367 ± 0.0053	0.549 ± 0.0032	8,124
8	≥ 6	default	0.290 ± 0.0134	0.417 ± 0.0084	1,153
9	No b-tag		0.351 ± 0.0042	0.545 ± 0.0026	12,655
10	≥ 6	default	0.389 ± 0.0045	0.565 ± 0.0026	11,908
11	≥ 0	default	0.443 ± 0.0020	0.611 ± 0.0011	62,944
12	≥ 0	default	0.473 ± 0.0014	0.632 ± 0.0008	126,154
13	≥ 0	default	0.444 ± 0.0026	0.611 ± 0.0014	37,864
14	≥ 0	default	0.464 ± 0.0017	0.627 ± 0.0009	88,466
15	≥ 0	default	0.497 ± 0.0012	0.652 ± 0.0006	188,956
16	≥ 0	default	0.505 ± 0.0010	0.656 ± 0.0005	251,870
17	≥ 0	default	0.508 ± 0.0008	0.660 ± 0.0004	377,960

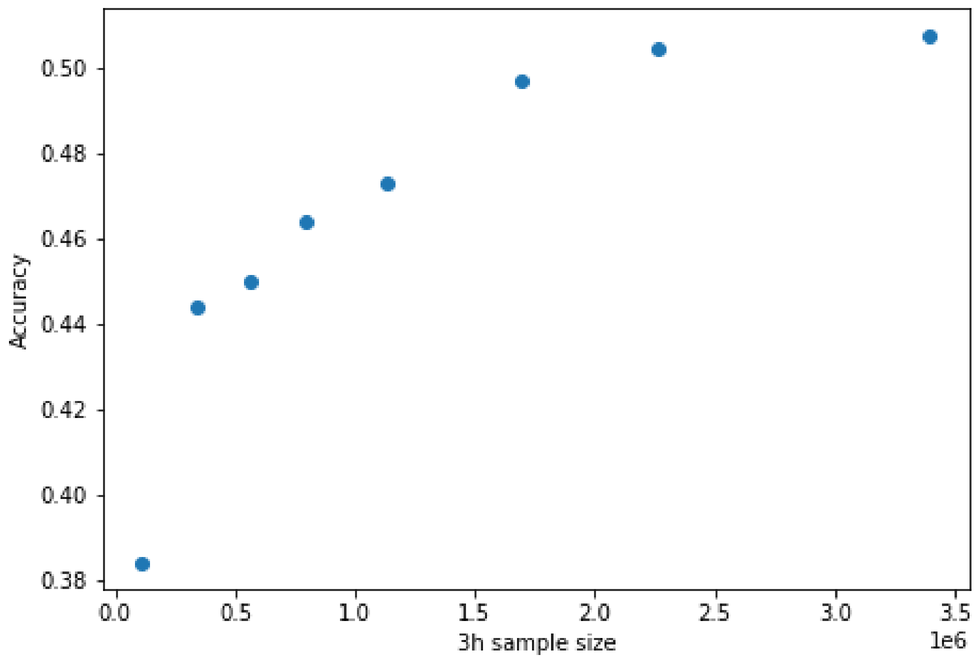


Figure 18: The accuracy against 3h training sample size.

22.3 Training curve

Figure 19 and 20 are the training curves for Case 9. After 50 epochs, the validation accuracy has not improved much.

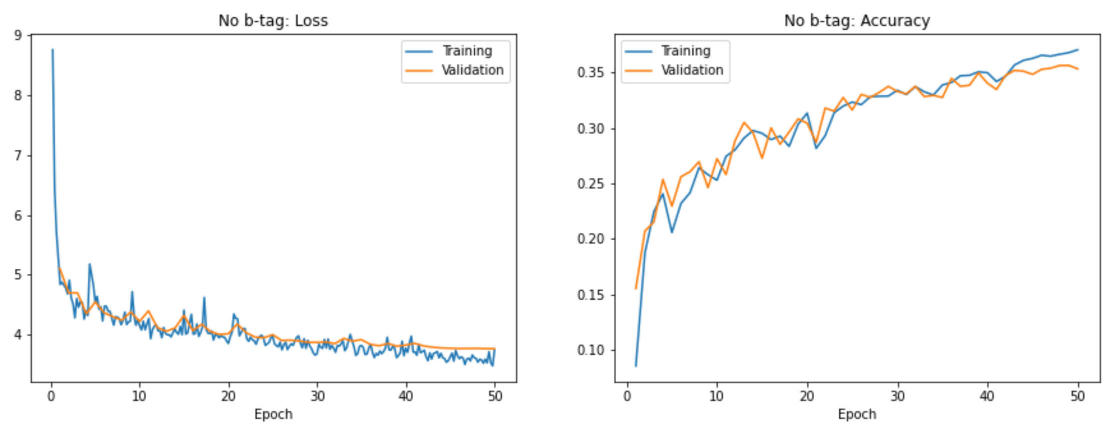


Figure 19: The 50 epoch training curves for Case 9.

Figure 21 is the training curves for Case 10.

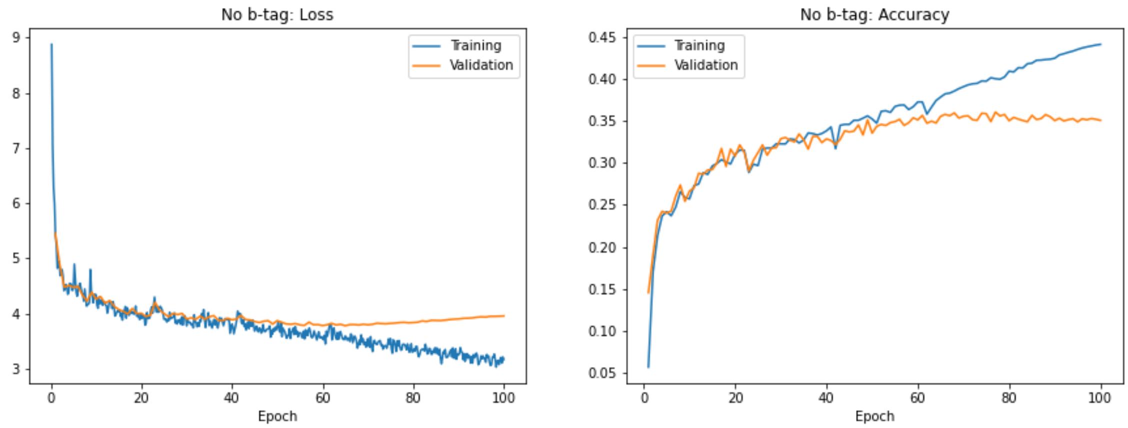


Figure 20: The 100 epoch training curves for Case 9.

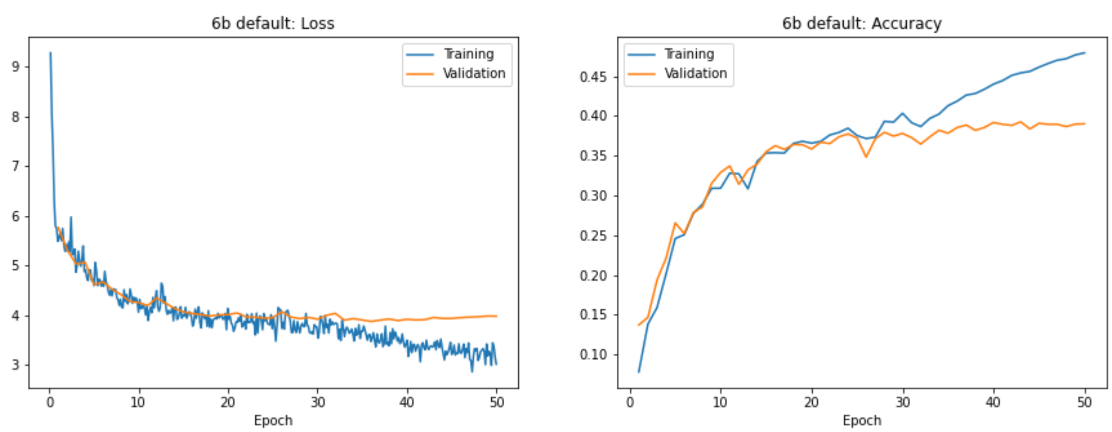


Figure 21: The training curves for Case 10 (6b default).

22.4 Test on different dataset

For example, use Case 5 samples for training and test on Case 5, 6, 7, and 8 samples. The results are presented in Table 45.

Table 45: 0b SPANet test on different samples.

Training samples	Testing samples	Event efficiency	Higgs efficiency
0b default	0b default	0.386	0.567
0b default	2b default	0.387	0.568
0b default	4b default	0.391	0.571
0b default	6b default	0.366	0.550

Use Case 6 samples for training and test on Case 5, 6, 7, and 8 samples. The results are presented in Table 46.

Table 46: 2b SPANet test on different samples.

Training samples	Testing samples	Event efficiency	Higgs efficiency
2b default	0b default	0.383	0.562
2b default	2b default	0.383	0.562
2b default	4b default	0.388	0.565
2b default	6b default	0.363	0.548

Use Case 7 samples for training and test on Case 5, 6, 7, and 8 samples. The results are presented in Table 47.

Table 47: 4b SPANet test on different samples.

Training samples	Testing samples	Event efficiency	Higgs efficiency
4b default	0b default	0.356	0.540
4b default	2b default	0.358	0.542
4b default	4b default	0.367	0.549
4b default	6b default	0.355	0.536

Use Case 10 samples for training and test on Case 5, 6, 7, and 8 samples. The results are presented in Table 48.

Use Case 12 samples for training and test on Case 6, 7, 10, and 12 samples. The results are presented in Table 49.

Table 48: 6b SPANet test on different samples.

Training samples	Testing samples	Event efficiency	Higgs efficiency
6b default	0b default	0.230	0.412
6b default	2b default	0.234	0.417
6b default	4b default	0.290	0.476
6b default	6b default	0.385	0.559

Table 49: 0b SPANet test on different samples.

Training samples	Testing samples	Event efficiency	Higgs efficiency
0b default	0b default	0.473	0.632
0b default	2b default	0.475	0.634
0b default	4b default	0.479	0.636
0b default	6b default	0.457	0.623

Table 50: SPANet trained on 0b test on 6b.

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.099	0.713	0.788
$= 7$	0.164	0.508	0.663
≥ 8	0.269	0.331	0.538
Total	0.533	0.457	0.623

23 Hyperparameters

Some hyperparameters of SPANet:

- Learning rate: 0.0004
- Central encoder count: 6
- Branch encoder count: 3
- Dropout percentage: 0.0
- L2 weight normalization: 0.0002
- L2 gradient clipping: 0.0

the above values are used in Sec.22.

23.1 Scan hyperparameters

In below, use the 0b default samples (Sec.22.2, Case 5) for hyperparameter scan.

Learning rate Range: $[0.0003, 0.0030]$. Step size: 0.0003. Figure 22 is the scanning results. The model cannot be trained when the learning rate is greater than 0.0006.

Central encoder count Range: $[2, 6]$. Step size: 1. Figure 23 is the scanning results. When central encoder layers are equal to 2 or 4, the model will get the best performance at the first 10 epochs.

Branch encoder count Range: $[2, 7]$. Step size: 1. Figure 24 is the scanning results. When branch encoder layers equal 2 or 4, the model will get the best performance at the first 10 epochs.

23.2 Training results for different hyperparamers

In this section, use the 0b default samples (Sec.22.2, Case 5) for training. Train 50 epochs. Table 51 is the training results for different hyperparameters.

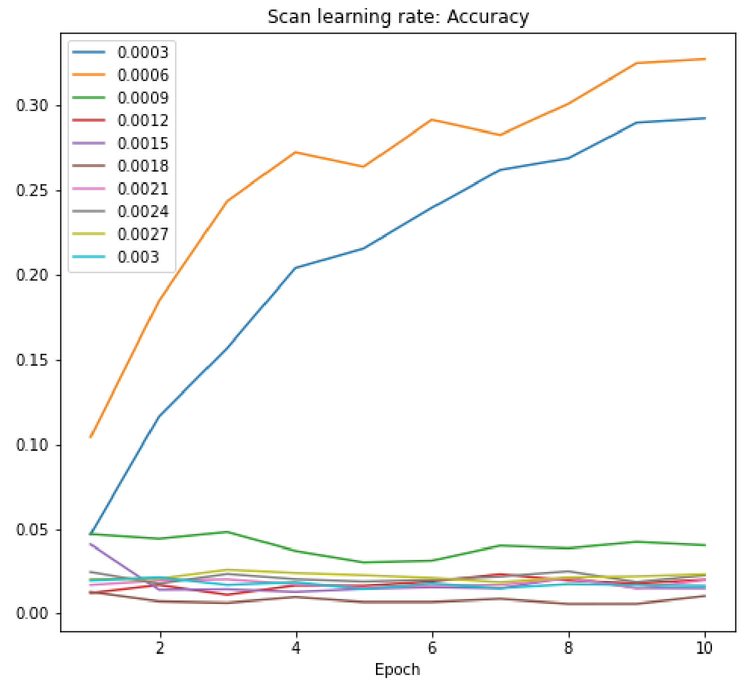


Figure 22: SPANet training results for different learning rates.

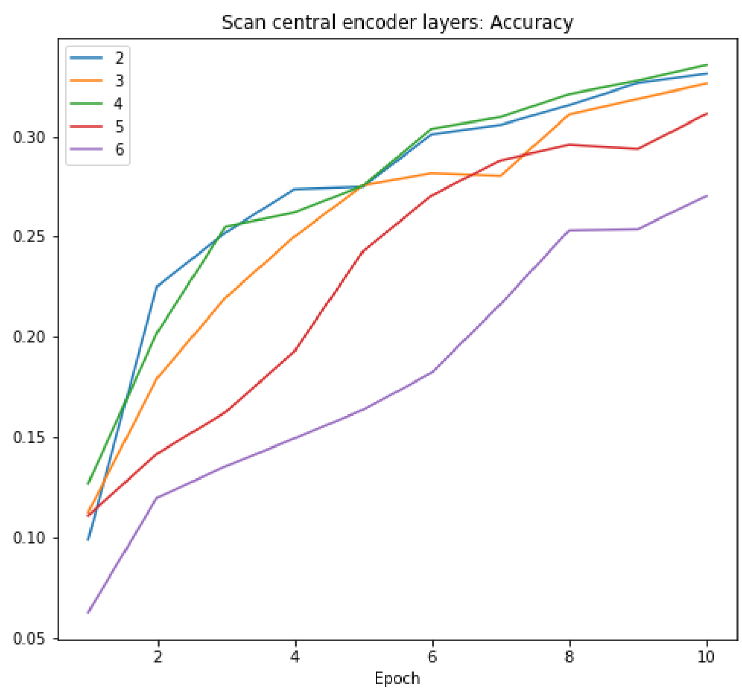


Figure 23: SPANet training results for different central encoder layers.

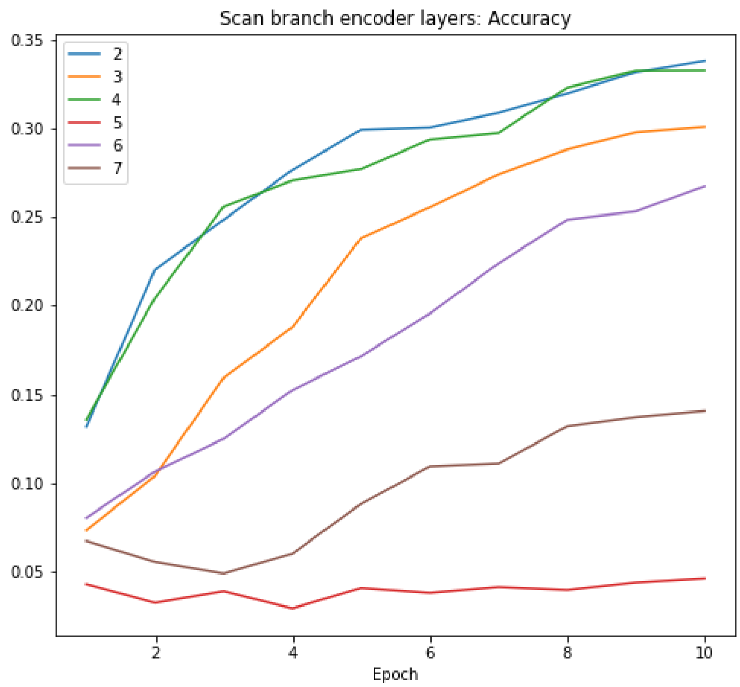


Figure 24: SPANet training results for different branch encoder layers.

Table 51: SPANet training results for different hyperparameters.

Central encoder	Branch encoder	learning rate	Event efficiency	Higgs efficiency
6	3	0.0004	0.386	0.567
2	2	0.0006	0.395	0.571
4	4	0.0006	0.362	0.547
4	2	0.0006	0.386	0.567

24 Signal and background

To separate the signal and background events, defined the cuts as follows:

- Eta PT cut: In $|\eta| < 2.5$ there are ≥ 6 jets. The p_{T} of jets are required over 25 GeV.
- bTag cut: ≥ 6 b-jets.
- PT cut: Require 4 b-jets over 40 GeV.
- inv mass cut: The total invariant mass of 6 b-jets ≤ 500 GeV.

The result is in Table 52.

Table 52: The cross sections for the signal and background processes at different cuts. $\mathcal{L} = 139\ \mathrm{fb}^{-1}$ is provided.

	Cross section (pb)			
	Signal	Background	S/B	$S/\sqrt{B}$
No cut	0.0118	7.18e+02	1.64e-05	0.164
Eta PT cut	0.00619	29.6	0.000209	0.424
bTag cut	0.00199	0.711	0.0028	0.881
PT cut	0.00164	0.503	0.00326	0.86
inv mass cut	0.0013	0.129	0.01	1.34

24.1 With SPANet

Define the cuts as follows:

- Eta PT cut: In $|\eta| < 2.5$ there are ≥ 6 jets. The p_{T} of jets are required over 25 GeV.
- PT cut: Require 4 jets over 40 GeV.
- Higgs mass cut: Use the SPANet to pair the jets and calculate the invariant mass of Higgs. Require the mass difference of SM Higgs χ^2 is smaller than 4000 GeV^2. The mass difference of SM Higgs χ^2 is calculated as follows

$$\chi^2 = [M(\mathrm{b}_1\mathrm{b}_2) - M_{\mathrm{H}}]^2 + [M(\mathrm{b}_3\mathrm{b}_4) - M_{\mathrm{H}}]^2 + [M(\mathrm{b}_5\mathrm{b}_6) - M_{\mathrm{H}}]^2$$

The result is in Table 53.

Table 53: The cross sections for the signal and background processes at different cuts.

	Cross section (fb)			$\mathcal{L} = 139$ fb^{-1}	$\mathcal{L} = 300$ fb^{-1}	$\mathcal{L} = 3000$ fb^{-1}
	Signal	Background	S/B	$S/\sqrt{B}$	$S/\sqrt{B}$	$S/\sqrt{B}$
No cut	11.8	7.18e+05	1.64e-05	0.164	0.241	0.761
Eta PT cut	6.19	2.96e+04	0.000209	0.424	0.623	1.97
PT cut	5.43	1.99e+04	0.000273	0.454	0.667	2.11
Higgs mass cut	3.06	5.59e+03	0.000548	0.483	0.709	2.24

24.1.1 Training SPANet

The training samples are required to pass the “Eta PT cut” and “PT cut”, ≥ 0 b-tagged jets, b-tagging efficiency is the default.

- Training sample:
  - Total sample size: 2,351,246
  - 1h sample size: 782,136
  - 2h sample size: 783,976
  - 3h sample size: 477,070
  - 5% used on validation
- Testing sample:
  - Total sample size: 261,717
  - 1h sample size: 86,638
  - 2h sample size: 87,394
  - 3h sample size: 53,226

The training results is presented in Table 54.

25 Di-Higgs signal background analysis

25.1 Signal and background event

Signal: same as the di-Higgs events in Sec.18.

Background: Generate background events in MadGraph by following commands:

Table 54: SPA-NET training results on the “PT cut” samples.

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 6$	0.059	0.668	0.752
$= 7$	0.072	0.463	0.630
≥ 8	0.072	0.297	0.509
Total	0.203	0.464	0.623

```
define p = p b b~
define j = j b b~
generate p p > j j j j
```

25.2 Cut

To separate the signal and background events, defined the cuts as follows:

- Eta PT cut: In $|\eta| < 2.5$ there are ≥ 4 jets. The p_{T} of jets are required over 25 GeV.
- bTag cut: ≥ 2 b-jets.
- Higgs cut: Pair the jets and calculate the invariant mass of Higgs. Require the mass difference of SM Higgs χ^2 is smaller than 1500 GeV^2. The mass difference of SM Higgs χ^2 is calculated as follows

$$\chi^2 = [M(\text{b}_1\text{b}_2) - M_{\text{H}}]^2 + [M(\text{b}_3\text{b}_4) - M_{\text{H}}]^2$$

  The pairing methods are described below.

Pairing method 1: Minimize the mass difference of 2 Higgs candidates.
Pairing method 2: Use the SPANet trained in Sec.25.3.
The result is in Table 55.

25.3 SPANet for di-Higgs jet pairing

The training samples are required to pass the “Eta PT cut”, ≥ 2 b-tagged jets. The b-tagging efficiency is the default.

- Training sample:
  - Total sample size: 617,340
  - 1h sample size: 266,790

Table 55: The cross sections for the di-Higgs signal and background processes at different cuts. Higgs cut 1 is for method 1. Higgs cut 2 is for method 2.

	Cross section (fb)			$\mathcal{L} = 139$ fb^{-1}	$\mathcal{L} = 300$ fb^{-1}	$\mathcal{L} = 3000$ fb^{-1}
	Signal	Background	S/B	$S/\sqrt{B}$	$S/\sqrt{B}$	$S/\sqrt{B}$
No cut	0.836	1.67e+07	5.02e-08	0.00242	0.00355	0.0112
Eta PT cut	0.714	1.39e+07	5.15e-08	0.00226	0.00332	0.0105
bTag cut	0.575	8.78e+05	6.55e-07	0.00723	0.01060	0.0336
Higgs cut 1	0.246	1.66e+05	1.48e-06	0.00711	0.01040	0.0330
Higgs cut 2	0.302	1.09e+05	2.77e-06	0.01080	0.01580	0.0501

  - 2h sample size: 315,219
  - 5% used on validation

- Testing sample:
  - Total sample size: 68,626
  - 1h sample size: 29,506
  - 2h sample size: 35,265

The training results is presented in Table 56.

Table 56: SPA-NET training results on the di-Higgs samples with "Eta PT cut" and "bTag cut".

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 4$	0.129	1.000	1.000
$= 5$	0.178	0.968	0.984
≥ 6	0.207	0.895	0.946
Total	0.514	0.947	0.973

25.4 ATLAS

The selection steps:

- Four tag: The event contains at least 4 b-tagged anti-kt small-R ($R = 0.4$) jets with $p_{\text{T}} > 40$ GeV and $|\eta| < 2.5$. The four jets with the highest b-tagging score are paired to construct two Higgs boson candidates.

- The four jets with the highest p_{T} are paired to construct two Higgs boson candidates in my samples.

- Delta R: Pairing jets to Higgs boson candidate need to satisfy the following requirements:

$$\left.\begin{aligned} \frac{360\ \text{GeV}}{m_{4\mathrm{j}}} - 0.5 < \Delta R_{jj,\text{lead}} < \frac{653\ \text{GeV}}{m_{4\mathrm{j}}} + 0.475 \\ \frac{235\ \text{GeV}}{m_{4\mathrm{j}}} < \Delta R_{jj,\text{subl}} < \frac{875\ \text{GeV}}{m_{4\mathrm{j}}} + 0.35 \end{aligned}\right\} \text{if } m_{4\mathrm{j}} < 1250\ \text{GeV}$$

$$\left.\begin{aligned} 0 < \Delta R_{jj,\text{lead}} < 1 \\ 0 < \Delta R_{jj,\text{subl}} < 1 \end{aligned}\right\} \text{if } m_{4\mathrm{j}} > 1250\ \text{GeV}$$

- If there are more than 2 pairings satisfy the Delta R requirement. Calculate D_{HH}

$$D_{HH} = \frac{\left|m_{2\mathrm{j}}^{\text{lead}} - \frac{120}{110} m_{2\mathrm{j}}^{\text{subl}}\right|}{\sqrt{1 + \left(\frac{120}{110}\right)^2}}$$

the pairing with the smallest value of D_{HH} is chosen.

- Higgs PT:

$$p_{\mathrm{T}}^{\text{lead}} > m_{4\mathrm{j}} \times 0.5 - 103\ \text{GeV}$$
$$p_{\mathrm{T}}^{\text{subl}} > m_{4\mathrm{j}} \times 0.33 - 73\ \text{GeV}$$

- Higgs Eta:

$$|\Delta\eta_{HH}| < 1.5$$

- Signal region:

$$X_{HH} = \sqrt{\left(\frac{m_{2\mathrm{j}}^{\text{lead}} - 120\ \text{GeV}}{0.1 m_{2\mathrm{j}}^{\text{lead}}}\right)^2 + \left(\frac{m_{2\mathrm{j}}^{\text{subl}} - 110\ \text{GeV}}{0.1 m_{2\mathrm{j}}^{\text{subl}}}\right)^2} < 1.6$$

- Top veto: Every possible pair of jets with $p_{\mathrm{T}} > 40$ GeV and $|\eta| < 2.5$, including those that were not selected for the H candidates, to form "W candidates". "Top quark candidates" are built by pairing W candidates with each remaining jet that was selected for H candidates

$$X_{Wt} = \sqrt{\left(\frac{m_W - 80\ \text{GeV}}{0.1 m_W}\right)^2 + \left(\frac{m_t - 173\ \text{GeV}}{0.1 m_t}\right)^2}$$

Events with the smallest $X_{Wt} < 1.5$ are vetoed.

Test the above selection steps on the 13 TeV samples with $m_{H_2} = 1000$ GeV and compare the results with ATLAS paper (Figure 25). The result is in Table 57. The biggest difference is at Delta R cut.

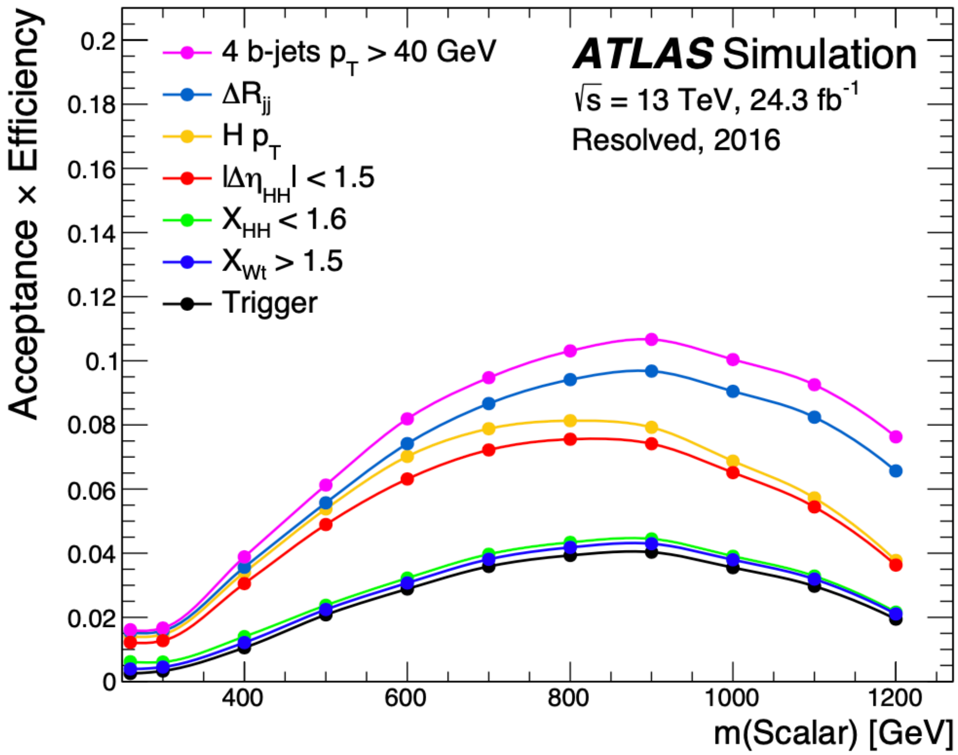


Figure 25: The selection acceptance times efficiency at each stage for the scalar.

25.4.1 Delta R cut

Change the b-tagging efficiency in Delphes such that the "Four tag" efficiency is the same as the ATLAS results. The b-tagging efficiency after the change is

$$\text{eff.} = 0.78 \tanh(0.0025 p_{\mathrm{T}}) \frac{25.0}{1 + 0.063 p_{\mathrm{T}}} \tag{8}$$

Generate events and apply the selection again. The results are in Table 59.

Define the correct jet assignments by matching them to the simulated truth quarks within an angular distance of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$.

The four b-jets with the highest p_{T} are paired to construct two Higgs boson candidates in my samples. When I checked the truth matching, I found the four b-jets may not all decay from Higgs. The event with a b-jet does not decay from Higgs is hard to pass the "Delta R cut". The number is in Table 60.

Table 57: The selection passing rate and efficiency at each stage. The resonant mass is 1000 GeV.

	ATLAS 1000 GeV		My sample 1000 GeV	
Cut	pass rate	efficiency	pass rate	efficiency
Four tag	0.1000	0.1000	0.1386	0.1386
Delta R	0.0900	0.9003	0.1089	0.7856
Higgs PT	0.0683	0.7584	0.0871	0.7996
Higgs Eta	0.0647	0.9469	0.0803	0.9225
Signal region	0.0390	0.6028	0.0422	0.5251
Top veto	0.0375	0.9612	0.0418	0.9912

Table 58: The selection passing rate and efficiency at each stage. The resonant mass is 600 GeV.

	ATLAS 600 GeV		My sample 600 GeV	
Cut	pass rate	efficiency	pass rate	efficiency
Four tag	0.0821	0.0821	0.1120	0.1120
Delta R	0.0739	0.9004	0.0856	0.7645
Higgs PT	0.0703	0.9508	0.0820	0.9571
Higgs Eta	0.0630	0.8966	0.0734	0.8951
Signal region	0.0324	0.5144	0.0287	0.3915
Top veto	0.0306	0.9439	0.0284	0.9903

Table 59: The selection passing rate and efficiency at each stage. The b-tagging efficiency is (8). The resonant mass is 1000 GeV.

	ATLAS 1000 GeV		My sample 1000 GeV	
Cut	pass rate	efficiency	pass rate	efficiency
Four tag	0.1000	0.1000	0.1006	0.1006
Delta R	0.0900	0.9003	0.0781	0.7758
Higgs PT	0.0683	0.7584	0.0627	0.8036
Higgs Eta	0.0647	0.9469	0.0575	0.9173
Signal region	0.0390	0.6028	0.0297	0.5163
Top veto	0.0375	0.9612	0.0294	0.9906

Table 60: Unmatched are the events that cannot define the correct jet assignments. Use the correct jet assignments to determine if four b-jets are all from Higgs. The total number of events is 100k. The resonant mass is 1000 GeV.

	Total	All from Higgs	Not all from Higgs	Unmatched
Four tag	13863	9348	2536	1979
Delta R	10891	8961	857	1073

25.4.2 Change b-tag efficiency

Change the b-tagging part in the Delphes card such that same as the MV2c10 b-tagger at 70% WP.

The b-jet efficiency is set to 0.70. The c-jet missing rate is set to 0.083. The light jet missing rate is set to 0.0026.

Regenerate the events and apply the cut. The result is in Table 61. The Delta R cut efficiency is 0.86, has a big difference compared to the before result of 0.78.

Table 61: The selection passing rate and efficiency at each stage. The b-tagging part is the same as the MV2c10 70% WP. The resonant mass is 1000 GeV.

	ATLAS 1000 GeV		My sample 1000 GeV	
Cut	pass rate	efficiency	pass rate	efficiency
Four tag	0.1000	0.1000	0.1257	0.1257
Delta R	0.0900	0.9003	0.1087	0.8643
Higgs PT	0.0683	0.7584	0.0896	0.8247
Higgs Eta	0.0647	0.9469	0.0826	0.9216
Signal region	0.0390	0.6028	0.0445	0.5391
Top veto	0.0375	0.9612	0.0441	0.9899

Table 62: Unmatched are the events that cannot define the correct jet assignments. Use the correct jet assignments to determine if four b-jets are all from Higgs. The total number of events is 100k. The b-tagging part is the same as the MV2c10 70% WP. The resonant mass is 1000 GeV.

	Total	All from Higgs	Not all from Higgs	Unmatched
Four tag	12573	10041	1181	1351
Delta R	10868	9635	337	856

25.4.3 Training SPANet for jet pairing

The training samples are required to pass the "Four tag cut", i.e., there are at least four b-tagged jets with $p_{\mathrm{T}} > 40$ GeV and $|\eta| < 2.5$. The b-tagging efficiency is the same as the MV2c10 b-tagger at 70% WP.

- Training sample:
  - Total sample size: 109,252
  - 1h sample size: 11,746
  - 2h sample size: 97,053
  - 5% used on validation
- Testing sample:
  - Total sample size: 12,140
  - 1h sample size: 1,199
  - 2h sample size: 10,883

The training results is presented in Table 63.

Table 63: SPA-NET training results on the di-Higgs samples with "Four tag cut".

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 4$	0.307	1.000	1.000
$= 5$	0.365	0.968	0.984
≥ 6	0.264	0.906	0.951
Total	0.896	0.961	0.980

25.4.4 Use SPANet for event selection

The "$\Delta R + \text{min-}D_{HH}$" pairing is replaced by SPANet pairing. Other cuts remained unchanged. The cutflow tables for signal and background are in Table 64, 65.

To separate the signal and background, apply the same selection steps to both samples. The background process are pp->4b and pp->tt->4j2b. Add up the cross sections of two background processes. The "$\Delta R + \text{min-}D_{HH}$" result is in Table 66. The SPANet result is in Table 67.

After the "Signal region cut", the SPANet pairing method has an improvement on $S/\sqrt{B}$.

Table 64: The selection passing rate and efficiency at each stage for "$\Delta R + \text{min-}D_{HH}$" and SPA-NET pairing. The resonant mass is 1000 GeV.

	$\Delta R + \text{min-}D_{HH}$				SPA-NET	
	ATLAS		My sample		My sample	
Cut	pass rate	efficiency	pass rate	efficiency	pass rate	efficiency
Four tag	0.1000	0.1000	0.1257	0.1257	0.1257	0.1257
Delta R	0.0900	0.9003	0.1087	0.8643	N.A.	N.A.
Higgs PT	0.0683	0.7584	0.0896	0.8247	0.0976	0.7759
Higgs Eta	0.0647	0.9469	0.0826	0.9216	0.0898	0.9202
Signal region	0.0390	0.6028	0.0445	0.5391	0.0468	0.5213
Top veto	0.0375	0.9612	0.0426	0.9555	0.0446	0.9534

Table 65: The selection passing rate and efficiency at each stage for "$\Delta R + \text{min-}D_{HH}$" and SPA-NET pairing.

	pp4b			
	$\Delta R + \text{min-}D_{HH}$		SPA-NET	
Cut	pass rate	efficiency	pass rate	efficiency
Four tag	0.0096	0.0096	0.0096	0.0096
Delta R	0.0050	0.5234	N.A.	N.A.
Higgs PT	0.0043	0.8462	0.0065	0.6714
Higgs Eta	0.0030	0.7127	0.0048	0.7380
Signal region	0.0005	0.1551	0.0001	0.0174
Top veto	0.0003	0.5975	0.0001	0.8675
	pptt			
	$\Delta R + \text{min-}D_{HH}$		SPA-NET	
Cut	pass rate	efficiency	pass rate	efficiency
Four tag	0.000747	0.000747	0.000747	0.000747
Delta R	0.000418	0.559572	N.A.	N.A.
Higgs PT	0.000370	0.885167	0.000548	0.733601
Higgs Eta	0.000310	0.837838	0.000444	0.810219
Signal region	0.000067	0.216129	0.000030	0.067568
Top veto	0.000021	0.313433	0.000010	0.333333

Table 66: The cross sections for the di-Higgs signal and background processes at different cuts. The pairing method is "$\Delta R + \text{min-}D_{HH}$".

	Cross section (fb)			$\mathcal{L} = 139$ fb^{-1}	$\mathcal{L} = 300$ fb^{-1}	$\mathcal{L} = 3000$ fb^{-1}
	Signal	Background	S/B	$S/\sqrt{B}$	$S/\sqrt{B}$	$S/\sqrt{B}$
No cut	0.644	8.92e+05	7.22e-07	0.00804	0.0118	0.0374
Four tag	0.081	6.26e+03	1.29e-05	0.0121	0.0177	0.0561
Delta R	0.070	3.28e+03	2.13e-05	0.0144	0.0212	0.0669
Higgs PT	0.058	2.78e+03	2.08e-05	0.0129	0.0190	0.0600
Higgs Eta	0.053	2.00e+03	2.67e-05	0.0140	0.0206	0.0652
Signal region	0.029	3.14e+02	9.13e-05	0.0191	0.0280	0.0886
Top veto	0.027	1.83e+02	0.00015	0.0239	0.0351	0.1110

Table 67: The cross sections for the di-Higgs signal and background processes at different cuts. The pairing method is SPA-NET.

	Cross section (fb)			$\mathcal{L} = 139$ fb^{-1}	$\mathcal{L} = 300$ fb^{-1}	$\mathcal{L} = 3000$ fb^{-1}
	Signal	Background	S/B	$S/\sqrt{B}$	$S/\sqrt{B}$	$S/\sqrt{B}$
No cut	0.644	8.92e+05	7.22e-07	0.00804	0.0118	0.0374
Four tag	0.081	6.26e+03	1.29e-05	0.0121	0.0177	0.0561
Higgs PT	0.063	4.21e+03	1.49e-05	0.0114	0.0168	0.0530
Higgs Eta	0.058	3.12e+03	1.85e-05	0.0122	0.0179	0.0567
Signal region	0.030	60.1	0.000502	0.0459	0.0674	0.2130
Top veto	0.029	47.9	0.000600	0.0490	0.0719	0.2270

25.4.5 Plots

The ΔR v.s. $m_{4\mathrm{j}}$ plots are in Figure 26.

The histograms for Higgs candidate mass are in Figure 27, 28.

26 Tri-Higgs signal background analysis

26.1 Tri-Higgs selection step

The selection steps:

- Six tag: The event contains at least 6 b-tagged anti-kt small-R ($R = 0.4$) jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. The six jets with the highest transverse momentum are paired to construct three Higgs boson candidates.
- Pick the pairing which minimizes the mass difference of 3 Higgs candidates, i.e., minimize

$$\chi^2 = [M(\mathrm{b_1b_2}) - M(\mathrm{b_3b_4})]^2 + [M(\mathrm{b_1b_2}) - M(\mathrm{b_5b_6})]^2 + [M(\mathrm{b_3b_4}) - M(\mathrm{b_5b_6})]^2$$

- Signal region:

$$X_{HHH} = \sqrt{[M(H_1) - M_{\mathrm{H}}]^2 + [M(H_2) - M_{\mathrm{H}}]^2 + [M(H_3) - M_{\mathrm{H}}]^2} < 40 \text{ GeV}$$

27 SPANET for di-Higgs mixing samples

27.1 Training sample preparation

The di-Higgs process pp->h2, h2->hh, h->bb was generated at $\sqrt{s} = 13$ TeV using MadGraph. Then pass these events to Pythia8 for showering and hadronization. Then pass to Delphes for detector simulation.

The b-tagging part in the Delphes card is set such that the efficiency is the same as the MV2c10 b-tagger at 70% WP. The b-jet efficiency is set to 0.70. The c-jet missing rate is set to 0.083. The light jet missing rate is set to 0.0026.

Jets are reconstructed by the anti-k_t algorithm with radius $R = 0.4$ and are required to have $p_{\mathrm{T}} \geq 20$ GeV.

Generate the samples with different h_2 mass. Then mix these samples.

Preselection cut: required at least 4 b-tagged jet with $p_{\mathrm{T}} > 40$ GeV in range $|\eta| < 2.5$.

Define the correct jet assignments by matching them to the simulated truth quarks within an angular distance of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$.

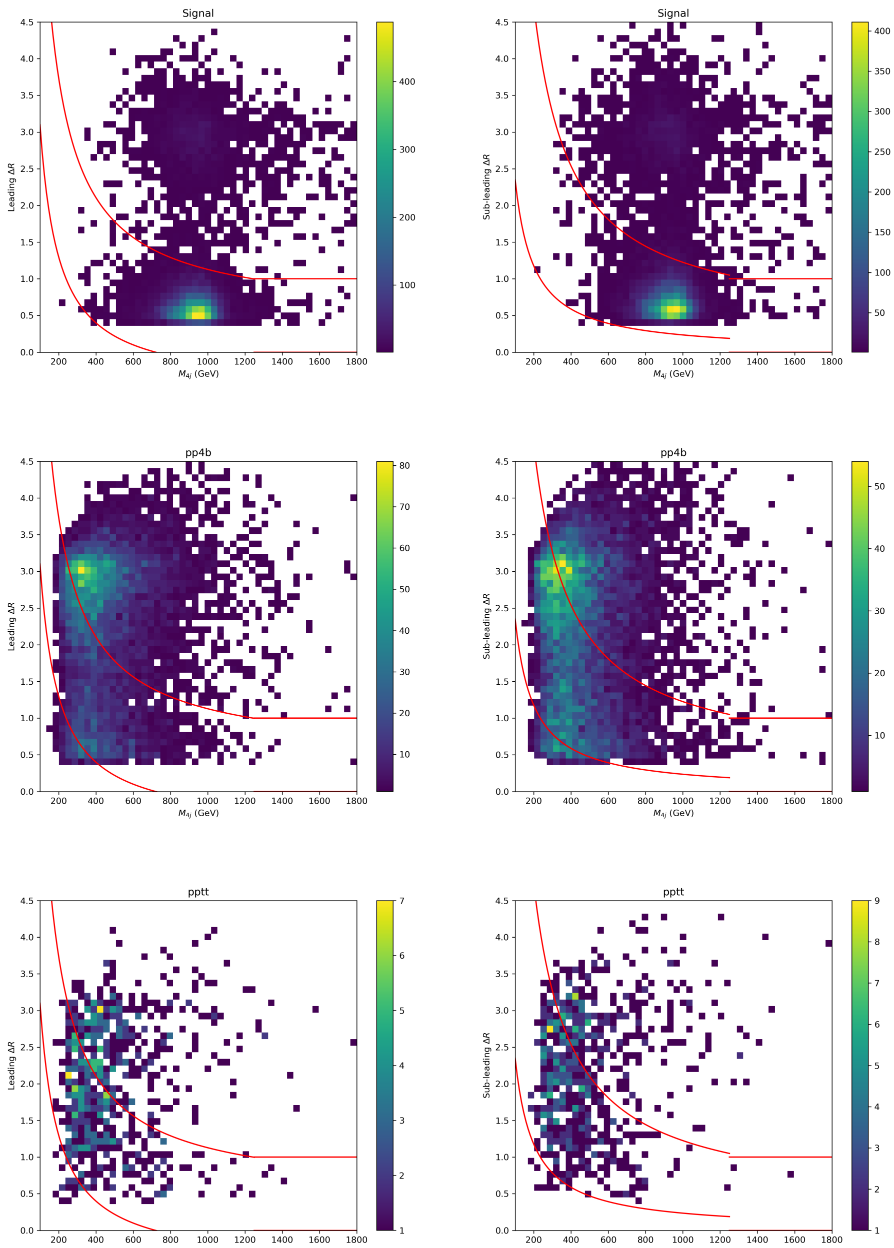


Figure 26: The ΔR_{jj} for Higgs candidate.

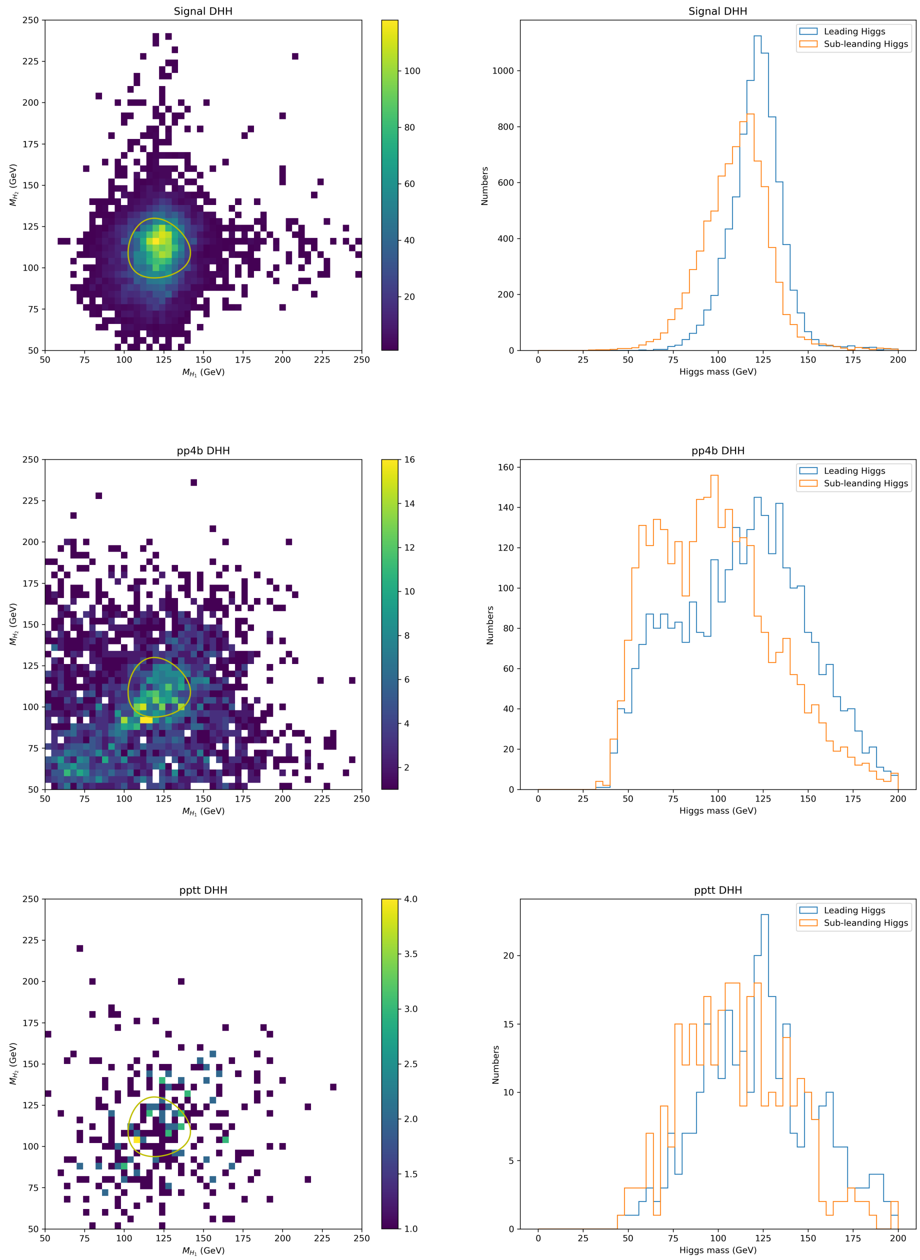


Figure 27: The mass plane and distribution for Higgs candidate. The pairing method is "$\Delta R + \text{min-}D_{HH}$".

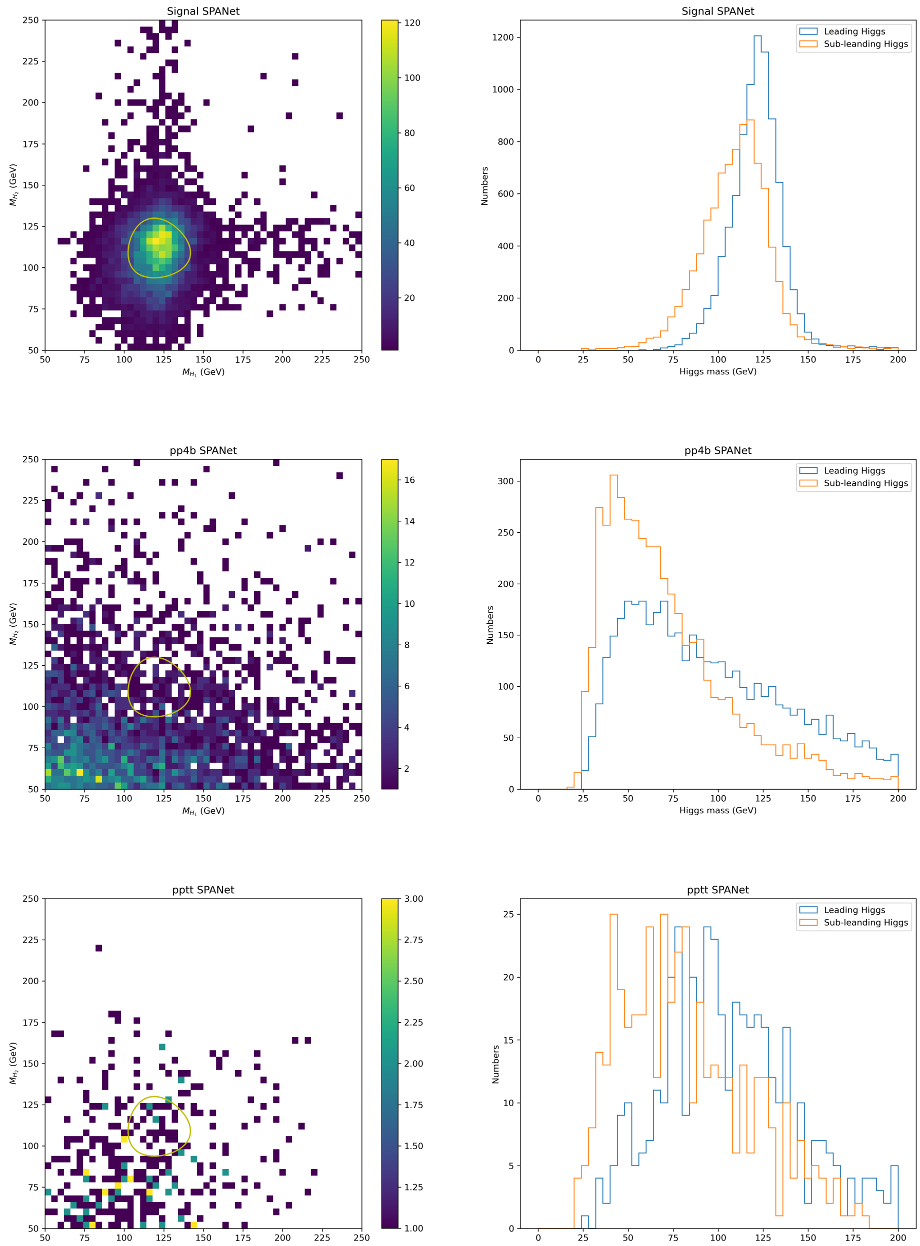
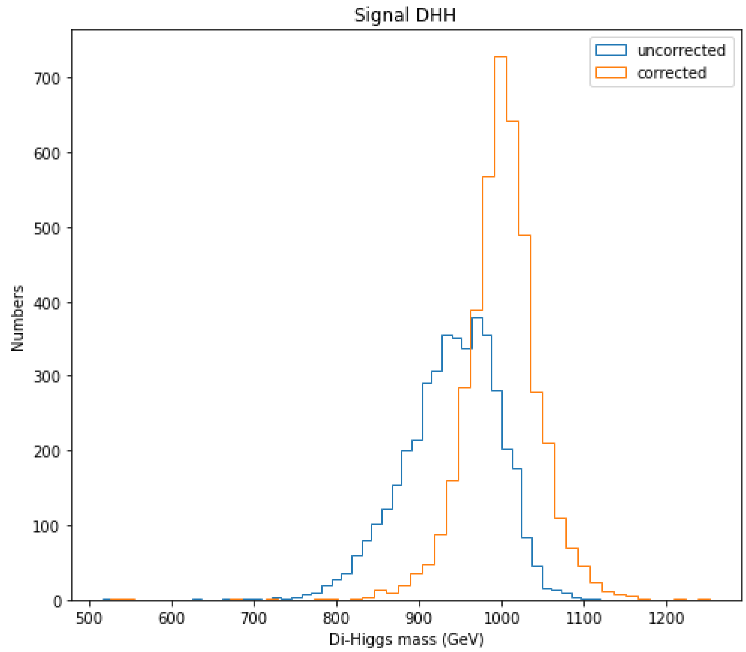


Figure 28: The mass plane and distribution for Higgs candidate. The pairing method is SPA-NET.



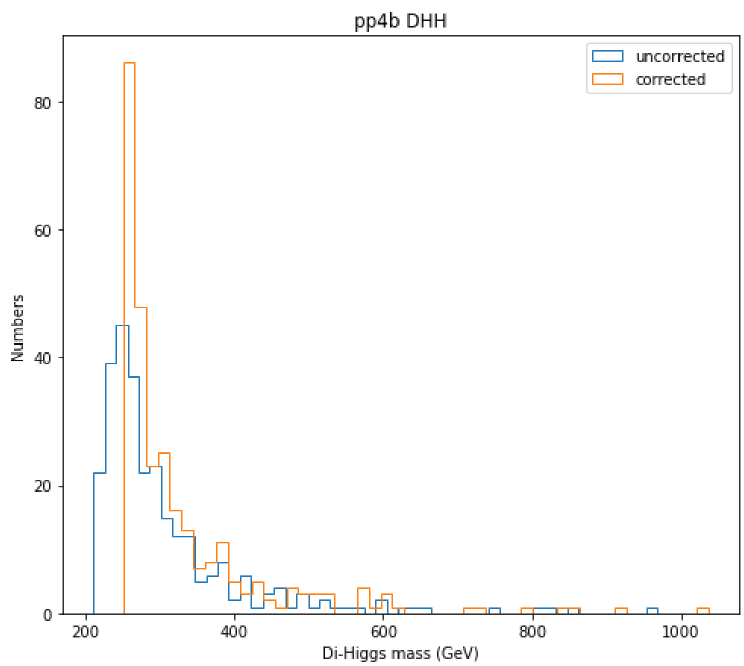


Figure 29: The di-Higgs mass distribution. The pairing method is "$\Delta R + \text{min-}D_{HH}$".

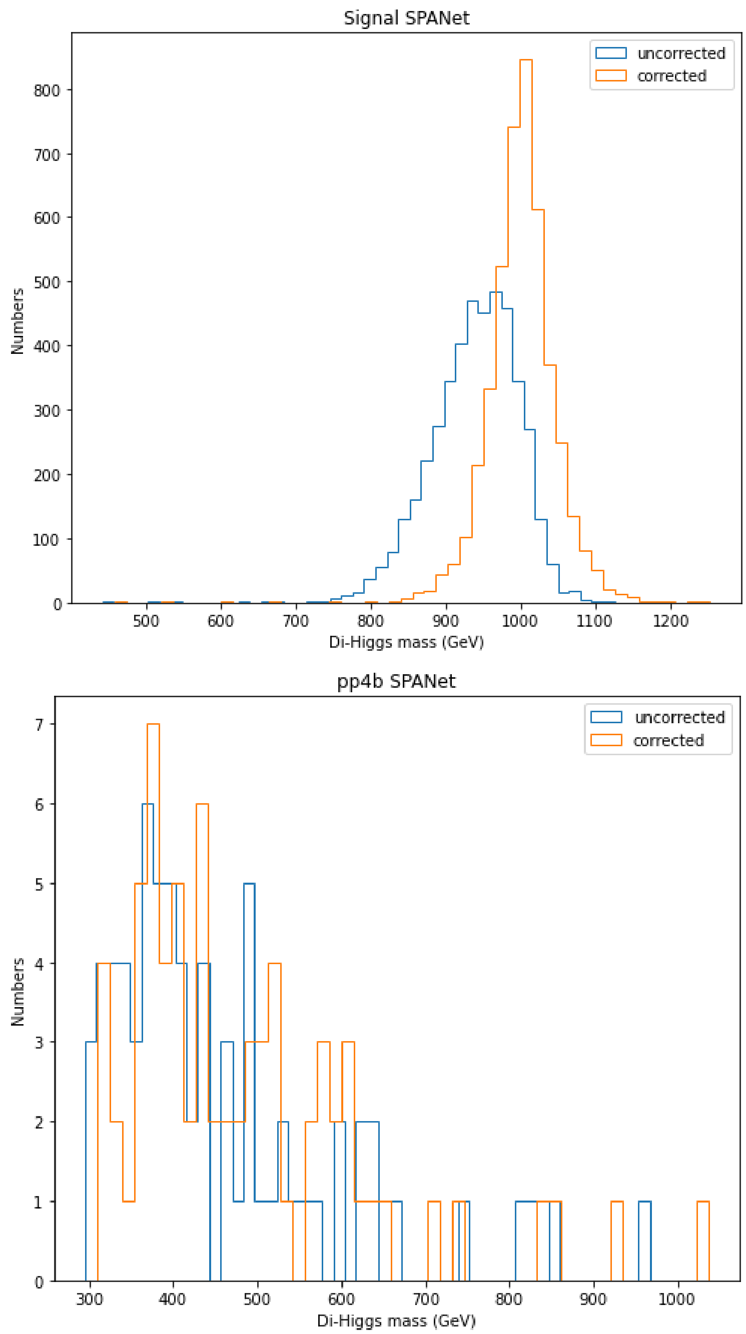


Figure 30: The di-Higgs mass distribution. The pairing method is SPA-NET.

27.2 Training results for mixing samples

Case 1: Mix m_H = 300 GeV, 600 GeV, 1000 GeV samples.

- Training sample:
  - Total sample size: 20,000 + 20,000 + 20,000 = 60,000
  - 1h sample size: 3,770 + 3,112 + 2,130 = 8,982
  - 2h sample size: 15,431 + 16,715 + 17,780 = 49,926
  - 5% used on validation
- Testing sample:
  - Total sample size: 2,000 + 2,000 + 2,000 = 6,000
  - 1h sample size: 368 + 309 + 232 = 909
  - 2h sample size: 1,547 + 1,675 + 1,759 = 4,981

The training results is presented in Table 68.

Table 68: SPA-NET training results on the mixing di-Higgs samples.

	Efficiency			
Case	300 GeV	600 GeV	1000 GeV	Total
1	0.765	0.907	0.956	0.880

27.3 Performance comparison

Calculate the pairing efficiency for D_{HH} method and the SPA-NET trained in Section 27.2.

These pairing methods are tested on the 2h sample in which the 4 quarks from Higgs can be matched to separate 4 jets.

The pairing efficiency is defined as

$$\text{efficiency} = \frac{\text{Correct pairing sample size}}{\text{Total 2h sample size}}$$

The results is in Figure 31.

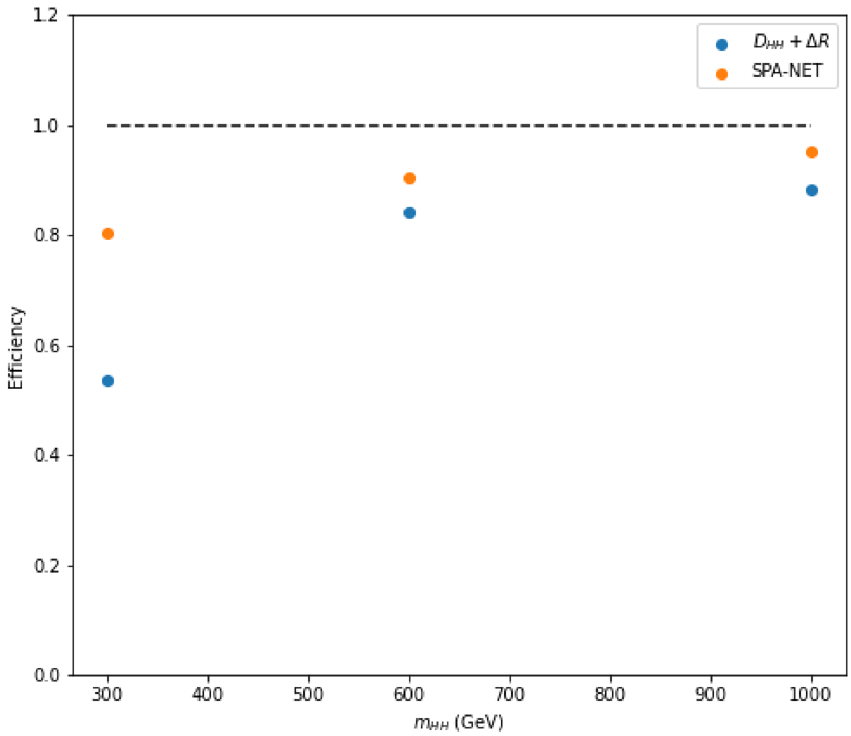


Figure 31: Performance of SPA-NET method compared to the $D_{HH} + \Delta R$ method.

28 Decay width

Check the heavy Higgs decay width in 2HDM. Check the MadGraph banner. For $m_H = 1000$ GeV, decay width is 35.5 GeV. For $m_H = 300$ GeV, decay width is 71.8 GeV. For $m_H = 600$ GeV, decay width is 56.9 GeV.

The above decay widths are copied from the 2HDMC output.

29 Generate di-Higgs samples in SM

Generate the double Higgs events in the standard model by MadGraph with `loop_sm` model. Following are the MadGraph scripts for generating di-Higgs samples:

```
import model loop_sm
generate p p > h h [QCD] QED^2<=99 QCD^2<=99
output /home/r10222035/CPVDM/Di-Higgs-SM/di-Higgs-sm

launch /home/r10222035/CPVDM/Di-Higgs-SM/di-Higgs-sm

shower=OFF
detector=OFF
analysis=OFF
done

set run_card nevents 10000
set run_card ebeam1 6500.0
set run_card ebeam2 6500.0

done
```

29.1 Variation with κ_λ

Reference: How to change the trilinear Higgs coupling in Madgraph?

The definition of κ_λ

$$\kappa_\lambda \equiv \frac{\lambda_{HHH}}{\lambda_{HHH}^{\mathrm{SM}}} \tag{9}$$

Following the below steps, we can add a parameter κ_λ in the model

1. Go to MadGraph model file directory. Copy `loop_sm` to `my_loop_sm`.

2. Go to `my_loop_sm` directory.

3. In `parameters.py`, add a new parameter for κ_λ by

   ```
   khhh = Parameter(name = 'khhh',
     nature = 'external',
     type = 'real',
     value = 1,
     texname = '\\text{khhh}',
     lhablock = 'SMINPUTS',
     lhacode = [ 10 ])
   ```

4. In `vertices.py`, we can find the coupling for three Higgs vertex in the form `GC_XX`.

5. In `couplings.py`, multiply the value for `GC_XX` found in step 4 by `khhh`.

6. In `restrict_default.dat`, add

   ```
   10 2.000000e+00 # khhh
   ```

   in Block SMINPUTS.

Finish the above setting we can use the following scripts to generate di-Higgs samples:

```
import model my_loop_sm
generate p p > h h [QCD] QED^2<=99 QCD^2<=99
output /home/r10222035/CPVDM/Di-Higgs-SM/di-Higgs-sm-kappa

launch /home/r10222035/CPVDM/Di-Higgs-SM/di-Higgs-sm-kappa

shower=OFF
detector=OFF
analysis=OFF
done

set param_card khhh 1

set run_card nevents 10000
set run_card ebeam1 6500.0
```

```
set run_card ebeam2 6500.0

done
```

29.2 Results

The cross section of various κ_λ is showed in Table 69.

Table 69: The cross section of various κ_λ. My data is the results from MadGraph. The reference data is from here.

	Cross section (fb)		
κ_λ	Ref.	My data	Ratio
-1	116.71	74.62	1.564
0	62.51	41.96	1.490
1	27.84	20.27	1.373
2	12.42	9.56	1.299
2.4	11.65	8.33	1.399
3	16.28	9.81	1.660
5	81.74	43.55	1.877

The m_{HH} distribution with various κ_λ is presented in Figure 32. In the left plot, the data is the parton level data from MadGraph. The right plot come from ATLAS reference.

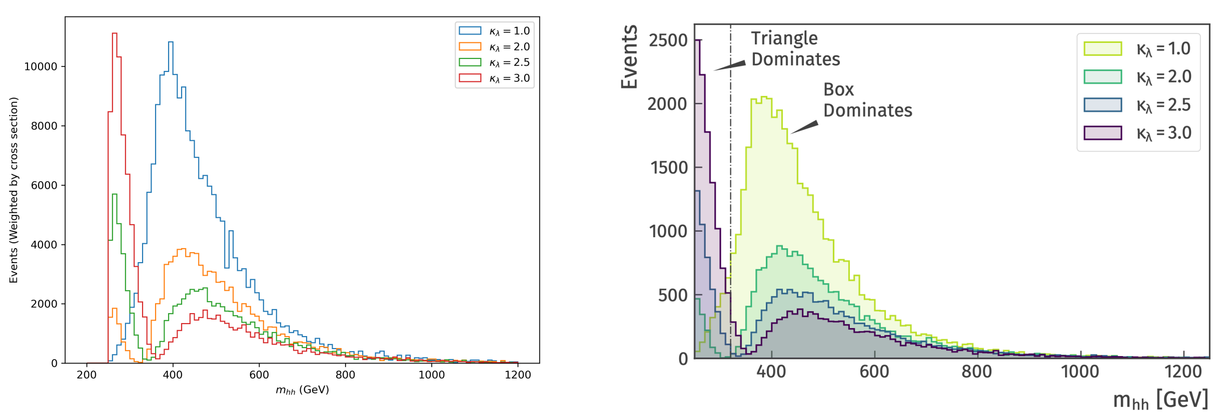


Figure 32: The m_{hh} distribution with various κ_λ. The bin height is weighted by the cross section.

Figure 33 and 34 are generated at $\sqrt{s} = 14$ TeV.

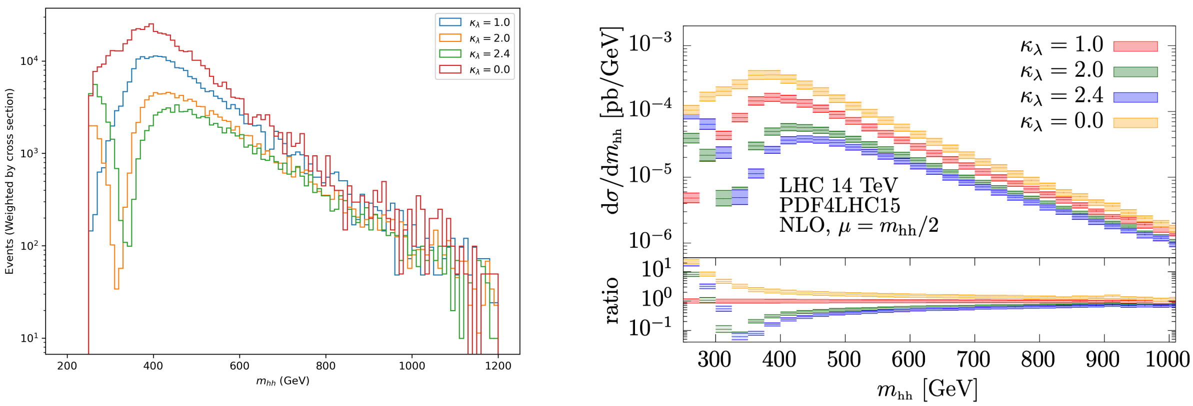


Figure 33: The m_{hh} distribution with various κ_λ. The bin height is weighted by the cross section.

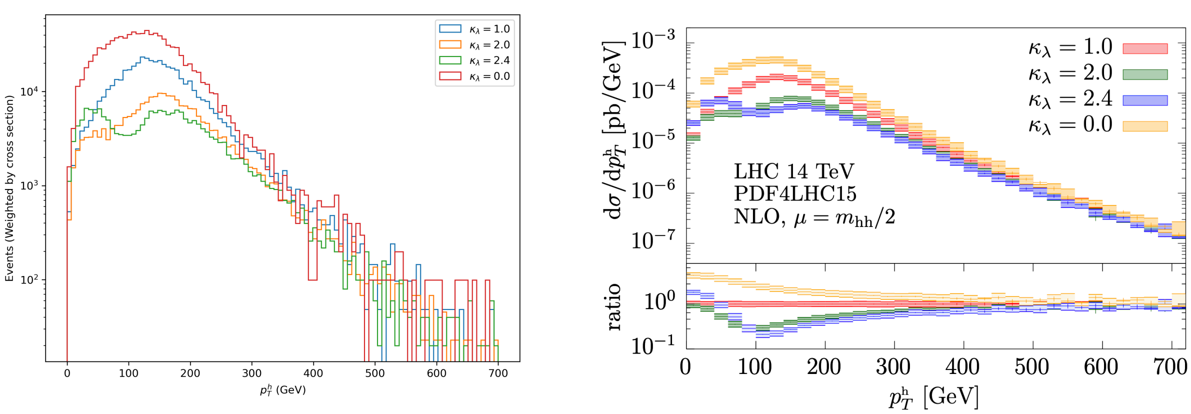


Figure 34: The p_{T}^{h} distribution with various κ_λ. The bin height is weighted by the cross section.