

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 $\Delta R(b,b)$ distribution of two b decays from the same Higgs. Then I plot $\Delta R(b,b)$ 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: $\Delta R(b,b)$ distribution and $\Delta R(b,b)$ against $p_T(b,b)$ plot

4 Cutflow table

I apply the η and p_T cuts sequentially and count how many event 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 $p_T > 10$ GeV in MadGraph.

Cut	number of events pass
6 b-partons $ \eta < 2.5$	330
6 b-partons $p_T < 25$ GeV	11
4 b-partons $p_T < 40$ GeV	8

Table 2: 1000 signal events.

Cut	number of events pass
6 b-partons $ \eta < 2.5$	732
6 b-partons $p_T < 25$ GeV	534
4 b-partons $p_T < 40$ 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

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

where $M(b_i b_j)$ 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

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

I implement these two methods to construct bb-pairs and test on the 10,000 signal events in 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 to identify the true Higgs pair.

6 Absolute cross section

Absolute cross section is defined as follow

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

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 same or not. The result is in Table 4.

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

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

Failed to generate the requested number of events

Table 3: The cuts applied on the different run. 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 $pp \rightarrow 4b$

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

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

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

In 100,000 events, there are 6,916 events 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, there are 173 events having greater than or equal to 6 b-jets.

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

To compare $pp \rightarrow 4b$ and $pp \rightarrow 6b$ 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 $pp \rightarrow 4b$ is 2051 and for $pp \rightarrow 6b$ is 1408. I scaled the number of events for $pp \rightarrow 4b$ to be the same as $pp \rightarrow 6b$.



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 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 $p_T > 25$ GeV.
- Cut 4: There are greater than or equal to 4 b-jets satisfy $p_T > 40$ GeV.

Table 5: 1,000,000 pp->4b events. Those events are generated in $\sqrt{s} = 14$ TeV with the cuts: $p_T > 25$ 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 $pp \rightarrow 4b$ and $pp \rightarrow 6b$ 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: $p_T > 25$ 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 $gg \rightarrow h_3$, $h_3 \rightarrow h_2 h$, $h_2 \rightarrow hh$ in MadGraph by following commands:

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

then check this channel is dominated in signal ($gg \rightarrow 3h$) 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.

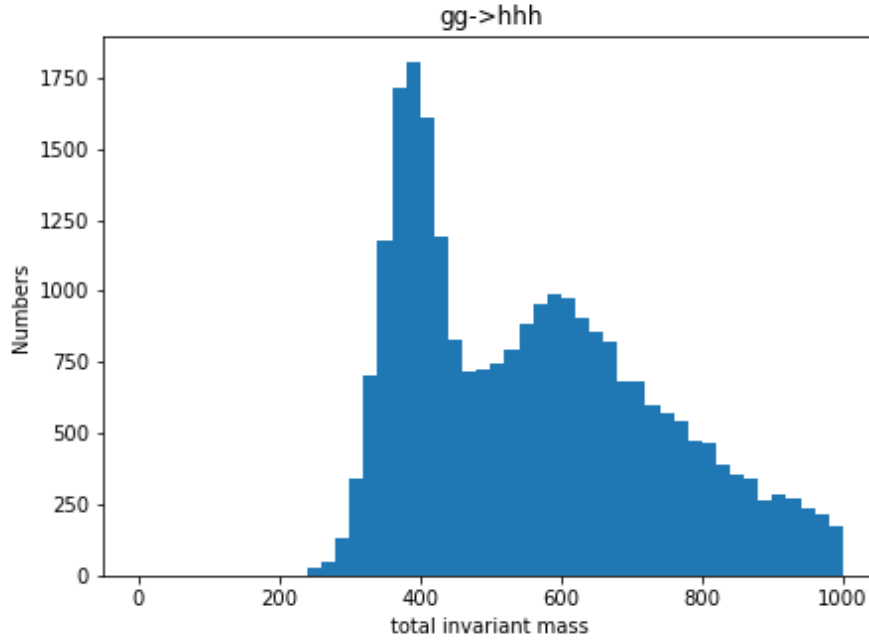


Figure 8: Total invariant mass of 6 b-jets for signal events.

A bump around 600 GeV in Figure 8? In 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 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 $gg \rightarrow h3$.

Table 8: Cross section from MadGraph and paper

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

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

13.1 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
```

```
...
```

Example: for diHiggs event ($h > b \bar{b}$)

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

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

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

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

2. Create training dataset in HDF5 format.

13.2 Training

13.3 Training result

Take diHiggs event as example:

Event purity:

$$\epsilon^{\text{event}} \equiv \frac{\text{both Higgs have all jets correctly assigned}}{\text{events with two identifiable Higgs}} \quad (1)$$

h purity:

$$\epsilon^h \equiv \frac{\text{Higgs have all jets correctly assigned}}{\text{number of identifiable Higgs}} \quad (2)$$

For 100k diHiggs event, $\epsilon^{\text{event}} = 0.851$, $\epsilon^h = 0.922$

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

Generate $pp \rightarrow hhh$ in MadGraph by following commands:

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

Generate $gg \rightarrow hhh$ in MadGraph by following commands:

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

These events are generated in 14 TeV. The cross section for $pp \rightarrow hhh$ and $gg \rightarrow hhh$ are 11.22 fb and 11.08 fb, respectively. They only differ by 1%.

15 Comparision for $gg \rightarrow hhh$ and resonant channel

To compare $gg \rightarrow 3h$ and resonant channel, 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 $gg \rightarrow 3h$ is 21,814 and for resonant channel is 21,475. The numbers are very close. I scaled the number of events for $gg \rightarrow 3h$ to be the same as 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. Those results for both events are similar.

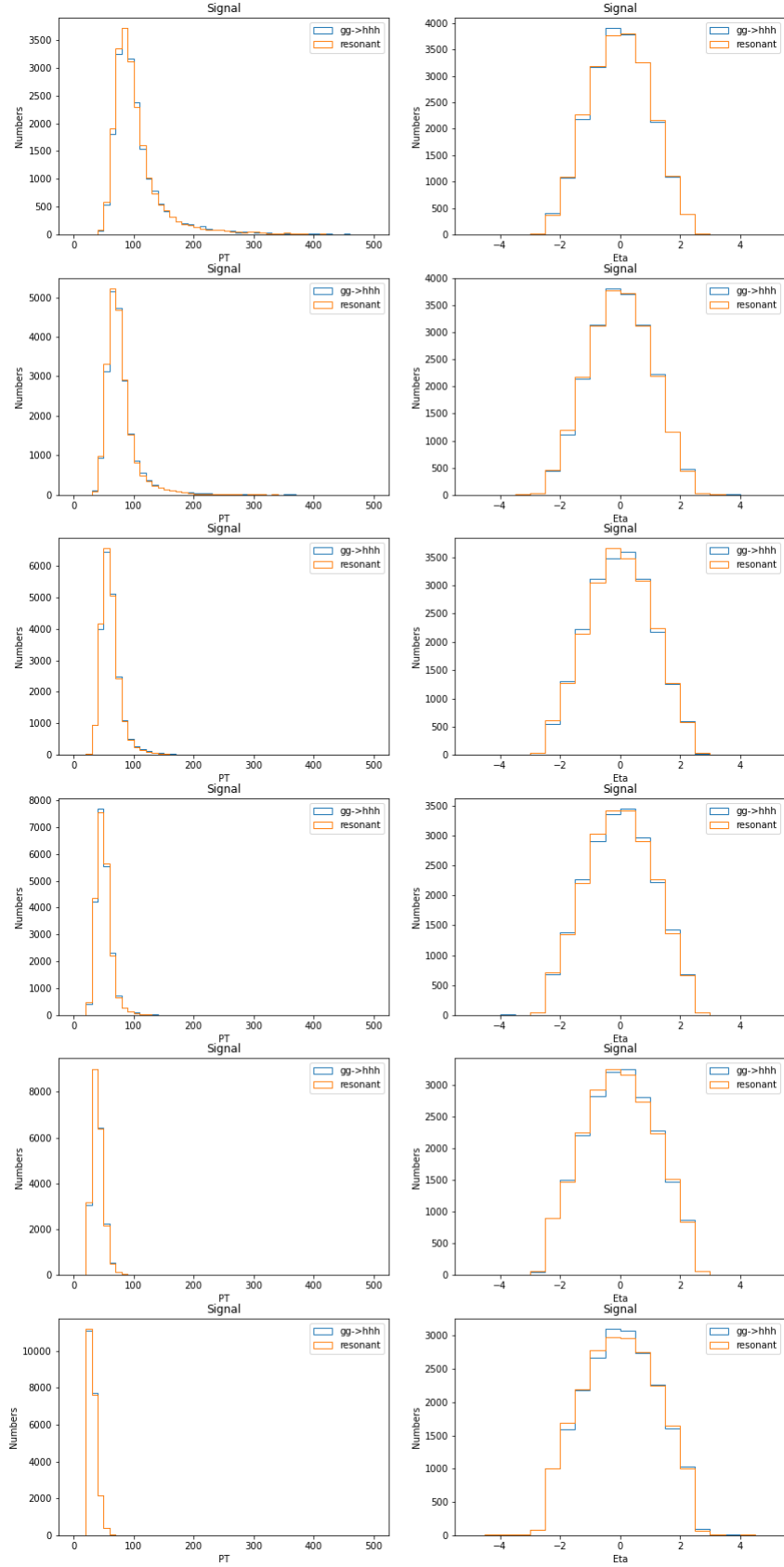


Figure 11: p_T and η distributions of b-jets for $gg \rightarrow 3h$ and resonant channel. They are ordered by p_T .

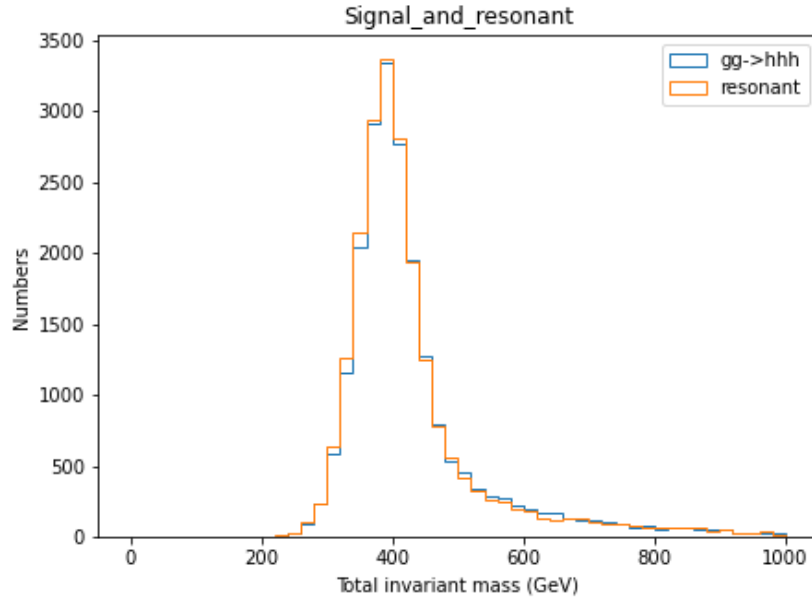


Figure 12: The distribution of total invariant mass of 6 b-jets for $gg \rightarrow 3h$ and resonant channel.

Table 9: Number of events pass the selection cuts. Total number of events for $gg \rightarrow hhh$ and resonant channel both are 100,000.

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