Note

Feng-Yang Hsieh

1 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
```

1.1 Variation with κ_λ

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

The definition of κ_λ

$$\kappa_\lambda \equiv \frac{\lambda_{HHH}}{\lambda_{HHH}^{\text{SM}}} \tag{1}$$

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

1. Go to the 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
```

1.2 Results

The cross sections of various κ_λ are shown in Table 1.

Table 1: The cross sections of various κ_λ. My data is the results from MadGraph. The reference data is from here.

	13 TeV			14 TeV				
	Cross section (fb)			Cross section (fb)				
κ_λ	Ref.	My data	Ref./My	Ref.	My data	Ref./My	Ref. K	Ref.K/My K
-1	116.71	74.62	1.564	136.91	87.93	1.56	1.86	1.19
0	62.51	41.96	1.490	73.64	49.45	1.49	1.79	1.20
1	27.84	20.27	1.373	32.88	24.05	1.37	1.66	1.21
2	12.42	9.56	1.299	14.75	11.34	1.30	1.56	1.20
2.4	11.65	8.33	1.399	13.79	9.90	1.39	1.65	1.18
3	16.28	9.81	1.660	19.07	11.55	1.65	1.90	1.15
5	81.74	43.55	1.877	95.22	50.68	1.88	2.14	1.14

The m_{HH} distribution with various κ_λ is presented in Figure 1. In the left plot, the data is the parton level data from MadGraph. The right plot comes from the ATLAS reference. Here, the $\sqrt{s} = 13$ TeV

Figure 2 and 3 are generated at $\sqrt{s} = 14$ TeV.

2 Non-resonant di-Higgs event selection

2.1 Sample

Non-resonant Higgs pair process is generated by MadGraph. Then pass to Pythia for showering and hadronization. Then pass to Delphes for detector simulation.

Jets are reconstructed using the anti-k_t algorithm with radius parameter $R = 0.4$.

The b-tagging part in the Delphes card is changed such that same as the DL1r b-tagger at 77% WP. The b-jet efficiency is set to 0.77. The c-jet missing rate is set to 0.204. The light jet missing rate is set to 0.0077.

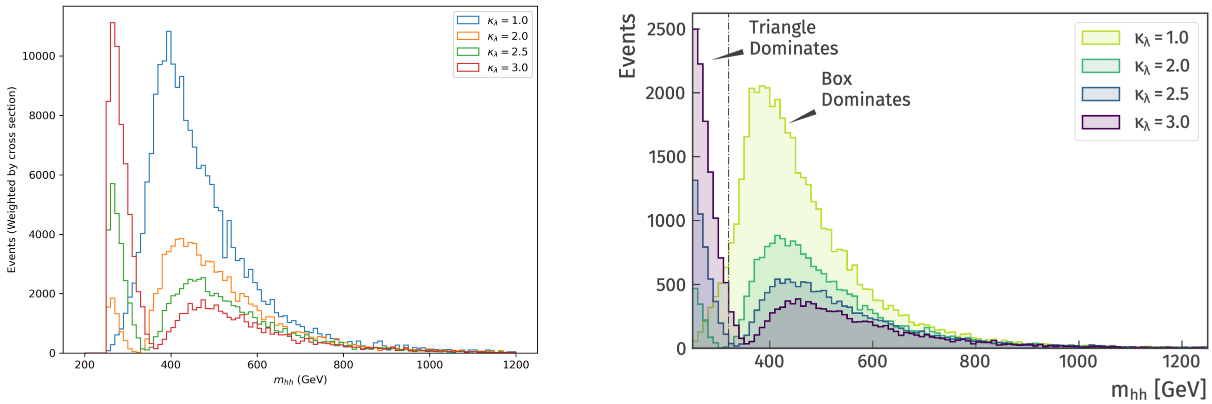

Figure 1: The m_{hh} distribution with various κ_λ. The bin height is weighted by the cross section.

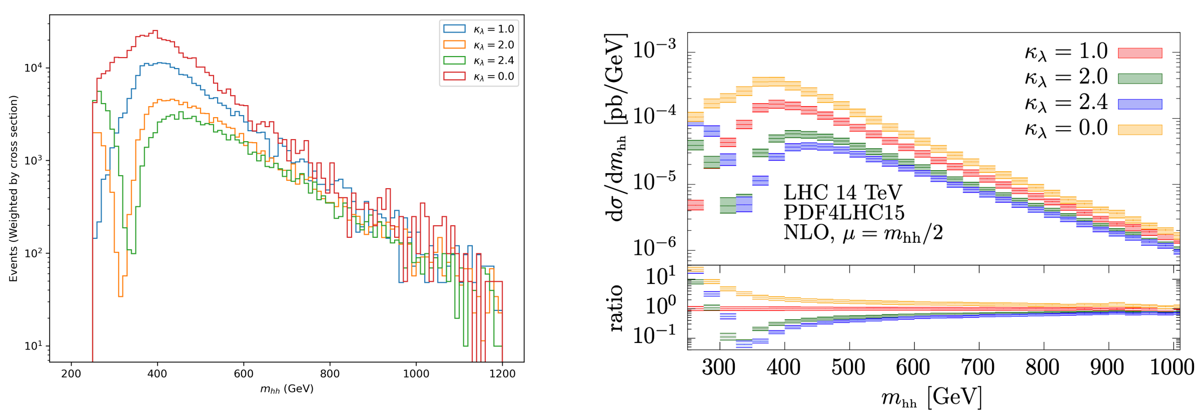

Figure 2: The m_{hh} distribution with various κ_λ. The bin height is weighted by the cross section.

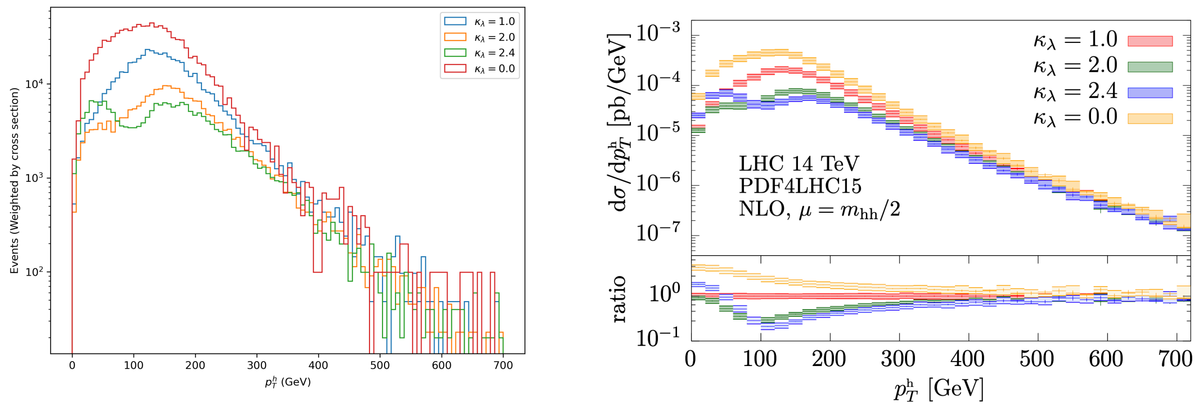

Figure 3: The p_{T}^h distribution with various κ_λ. The bin height is weighted by the cross section.

2.2 Event selection

Reference: ATLAS CONF Note CONF-HDBS-2022-35

The selection steps:

- Four tag: The event contains at least 4 b-tagged anti-k_t $R = 0.4$ jets with $p_{\mathrm{T}} > 40$ GeV and $|\eta| < 2.5$.
- The four jets with the highest p_{T} are paired to construct two Higgs boson candidates.
- min-ΔR pairing method: Choose the pairing in which the higher-p_{T} jet pair has the smallest ΔR separation.
- Higgs Eta:
$$|\Delta\eta_{HH}| < 1.5$$
- 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. The quantity X_{Wt} is defined as
$$X_{Wt} = \sqrt{\left(\frac{m_W - 80.4\ \text{GeV}}{0.1 m_W}\right)^2 + \left(\frac{m_t - 172.5\ \text{GeV}}{0.1 m_t}\right)^2}$$
Events with the smallest $X_{Wt} < 1.5$ are vetoed.

- Signal region:

$$X_{HH} = \sqrt{\left(\frac{m_{H_1} - 124\text{ GeV}}{0.1 m_{H_1}}\right)^2 + \left(\frac{m_{H_2} - 117\text{ GeV}}{0.1 m_{H_2}}\right)^2} < 1.6$$

Table 2: The selection passing rate and efficiency at each stage. The b-tagging part is the same as the DL1r 77% WP.

	ATLAS		My sample	
Cut	pass rate	efficiency	pass rate	efficiency
Four tag	0.0649	0.0649	0.0852	0.0852
Higgs Eta	0.0543	0.8360	0.0688	0.8074
Top veto	0.0456	0.8401	0.0553	0.8044
Signal region	0.0220	0.4818	0.0181	0.3283

Correct selection: Consider the events in which four jets can be matched one-to-one (within $\Delta R < 0.3$) to the four b-quarks decayed from the Higgs bosons. For the highest p_{T} there are 89% of simulated signal events reaching this selection.

Correct pairing: Consider the correct selection events, for min-ΔR pairing method there 85% of events are correctly paired.

Figure 4 shows the Higgs mass distribution. There is a deviation between the mass distribution peak and the signal region's center.

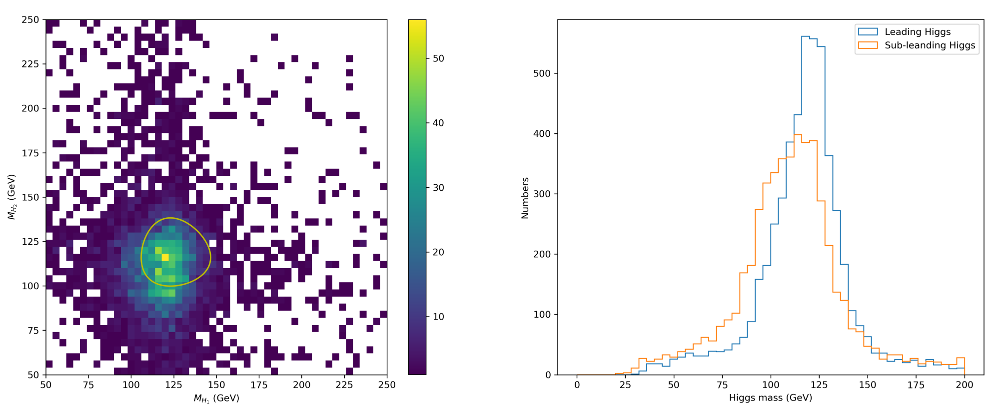

Figure 4: The mass plane and distribution for Higgs candidate.

2.2.1 Old method

Reference: Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

The selection steps:

- Four tag: The event contains at least 4 b-tagged anti-kt small-R ($R = 0.4$) jets with $p_{\mathrm{T}} > 40$ 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:

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

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

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

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

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

- Higgs PT:

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

- Higgs Eta:

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

- Signal region:

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

- 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.

The results are in Table 3.

Table 3: The selection passing rate and efficiency at each stage.

	ATLAS		My sample	
Cut	pass rate	efficiency	pass rate	efficiency
Four tag	0.0490	0.0490	0.0563	0.0563
Delta R	0.0448	0.9143	0.0471	0.8370
Higgs PT	0.0422	0.9420	0.0446	0.9480
Higgs Eta	0.0380	0.9005	0.0398	0.8911
Signal region	0.0193	0.5079	0.0170	0.4280
Top veto	0.0179	0.9275	0.0145	0.8537

2.3 Background event selection

Apply the same selection step to the background samples. The cutflow table is in Table 4 and the mass distribution is in Figure 5.

Table 4: The selection passing rate and efficiency at each stage for "min-ΔR" pairing method.

	pp4b min-ΔR	
Cut	pass rate	efficiency
Four tag	0.0096	0.0096
Higgs Eta	0.0056	0.5835
Top veto	0.0042	0.7423
Signal region	0.0001	0.0232

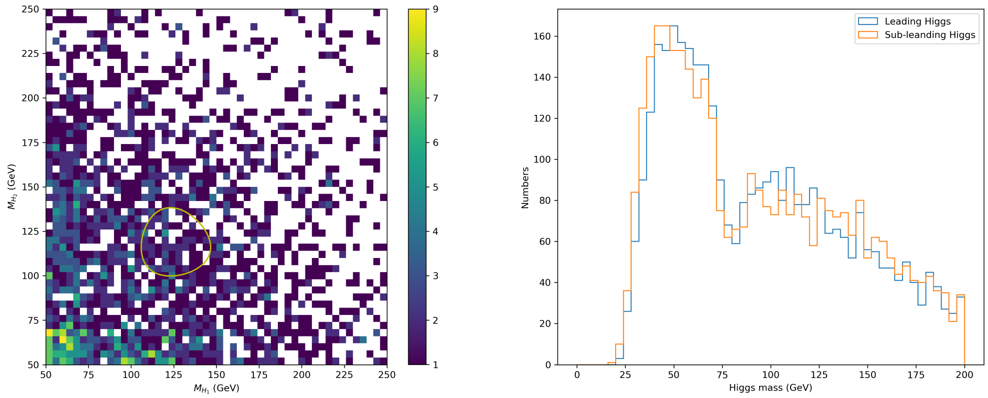

Figure 5: The mass plane and distribution for background events Higgs candidate.

3 Apply SPANET on non-resonant di-Higgs event

3.1 Training SPANET

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 DL1r b-tagger at 77% WP.

- Training sample:
  - Total sample size: 76,131
  - 1h sample size: 14,527
  - 2h sample size: 60,122
  - 5% used on validation
- Testing sample:
  - Total sample size: 8,460
  - 1h sample size: 1,577
  - 2h sample size: 6,744

The training results are presented in Table 5.

Table 5: SPA-NET training results on the di-Higgs samples with "Four tag cut".

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 4$	0.280	0.961	0.961
$= 5$	0.287	0.878	0.913
≥ 6	0.229	0.740	0.819
Total	0.797	0.868	0.903

Table 6: The selection passing rate and efficiency at each stage for "min-ΔR" and SPA-NET pairing.

	min-ΔR				SPA-NET	
	ATLAS		My sample		My sample	
Cut	pass rate	efficiency	pass rate	efficiency	pass rate	efficiency
Four tag	0.0649	0.0649	0.0852	0.0852	0.0852	0.0852
Higgs Eta	0.0543	0.8360	0.0688	0.8074	0.0635	0.7454
Top veto	0.0456	0.8401	0.0553	0.8044	0.0508	0.8006
Signal region	0.0220	0.4818	0.0181	0.3283	0.0027	0.0541

Table 7: The selection passing rate and efficiency at each stage for "min-ΔR" and SPA-NET pairing.

	pp4b			
	min-ΔR		SPA-NET	
Cut	pass rate	efficiency	pass rate	efficiency
Four tag	0.0096	0.0096	0.0096	0.0096
Higgs Eta	0.0056	0.5835	0.0055	0.5733
Top veto	0.0042	0.7423	0.0042	0.7607
Signal region	0.0001	0.0232	0.0001	0.0181

3.2 Use SPANET for event selection

The "min-ΔR" pairing is replaced by SPA-NET pairing. Other cuts remained unchanged. The cutflow tables for signal and background are in Table 6, 7.

Figure 6 shows the Higgs mass distribution for SPA-NET pairing.

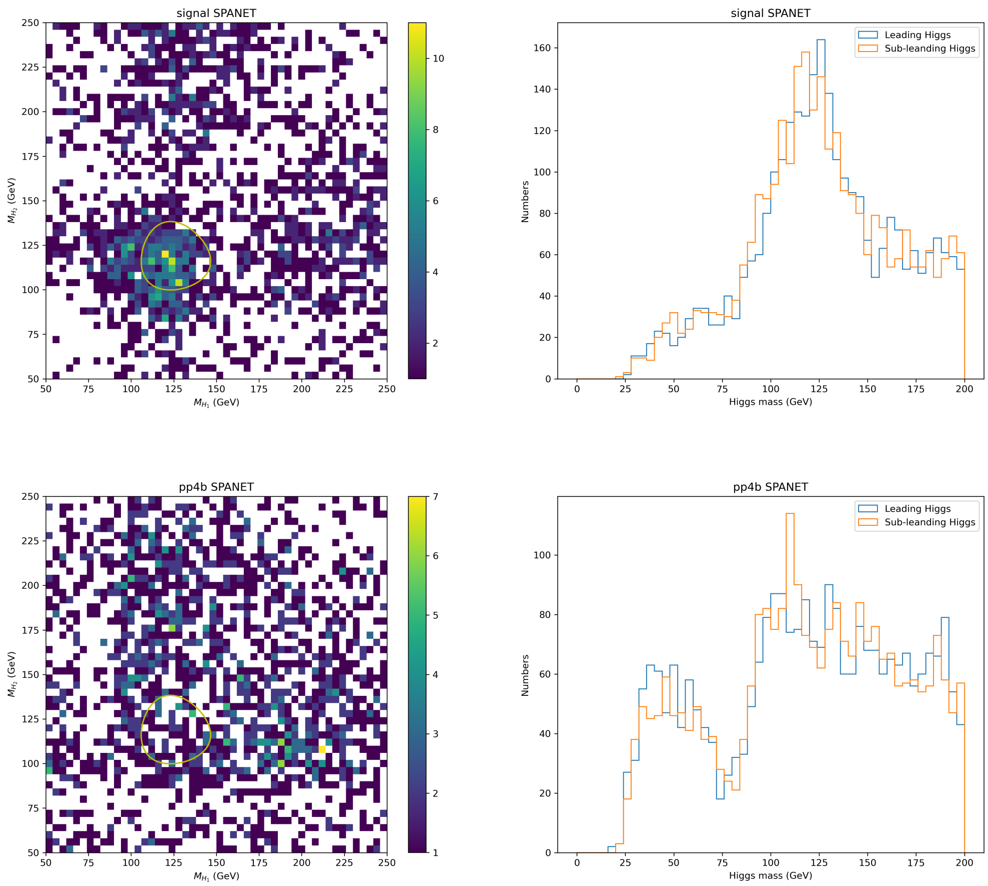

Figure 6: The mass plane and distribution for Higgs candidate for SPA-NET pairing method. The above figure is for the signal sample and the below one is for the background sample.

4 Pairing performace

The truth pairing is defined by matching the jets to the simulated quarks within the angular distance of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$.

The "min-ΔR" pairing method contain two-step: first select four b-jets with the highest p_{T}, and second, choose min-ΔR configuration. For the first step, the correct selection rate is 89%. For the second step, the correct pairing rate is 91%. The pairing efficiency is 81%.

The “SPA-NET” pairing method: The pairing efficiency is 86.8%.

The mass distribution for truth pairing is in Figure 7. The results are similar to the “min-ΔR” ones, but much different to the SPA-NET’s. There are indeed some bugs in the SPA-NET pairing code.

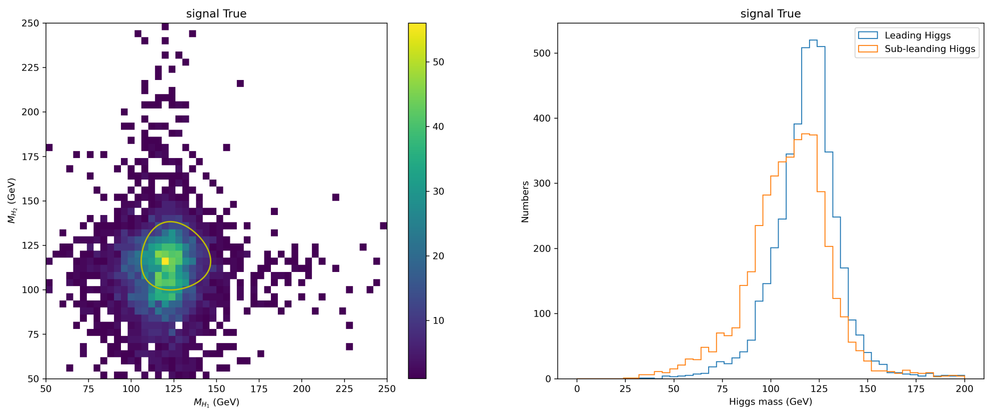

Figure 7: The mass plane and distribution for truth pairing.

5 Apply SPANET to non-resonant di-Higgs event again

The bugs for Sec.3: Some variables are misdefined twice, so they will get an incorrect number when calculating some physical quantities.

Fix this bug, then the cutflow tables for signal and background are in Table 8, 9.

Table 8: The selection passing rate and efficiency at each stage for “min-ΔR” and SPA-NET pairing.

	min-ΔR				SPA-NET	
	ATLAS		My sample		My sample	
Cut	pass rate	efficiency	pass rate	efficiency	pass rate	efficiency
Four tag	0.0649	0.0649	0.0852	0.0852	0.0852	0.0852
Higgs Eta	0.0543	0.8360	0.0688	0.8074	0.0676	0.7942
Top veto	0.0456	0.8401	0.0553	0.8044	0.0544	0.8051
Signal region	0.0220	0.4818	0.0181	0.3283	0.0194	0.3567

Figure 8 shows the Higgs mass distribution for SPA-NET pairing.

The “min-ΔR” result is in Table 10. The SPANet result is in Table 11.

Table 9: The selection passing rate and efficiency at each stage for "min-ΔR" and SPA-NET pairing.

	pp4b			
	min-ΔR		SPA-NET	
Cut	pass rate	efficiency	pass rate	efficiency
Four tag	0.0096	0.0096	0.0096	0.0096
Higgs Eta	0.0056	0.5835	0.0050	0.5202
Top veto	0.0042	0.7423	0.0037	0.7341
Signal region	0.0001	0.0232	0.0002	0.0443

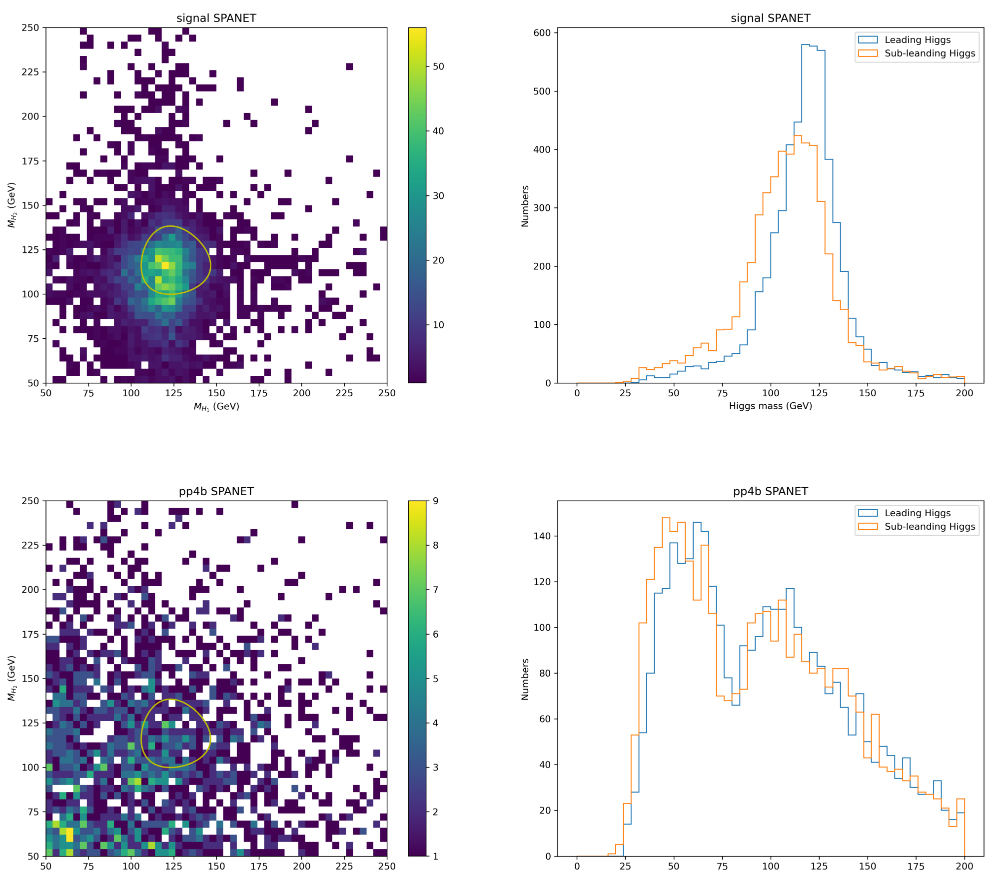

Figure 8: The mass plane and distribution for Higgs candidate for SPA-NET pairing method. The above figure is for the signal sample and the below one is for the background sample.

The SPA-NET method will let more signal events pass the Higgs signal cut, but also true for the background process. Therefore, the SPA-NET pairing method can not improve the $S/\sqrt{B}$.

Table 10: The cross sections for the di-Higgs signal and background processes at different cuts. The pairing method is "min-ΔR".

	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	6.774	6.29e+05	1.08e-05	0.101	0.148	0.468
Four tag	0.577	6.06e+03	9.52e-05	0.0874	0.128	0.406
Higgs Eta	0.466	3.54e+03	0.000132	0.0923	0.136	0.429
Top veto	0.375	2.62e+03	0.000143	0.0862	0.127	0.401
Higgs signal	0.123	61.0	0.00202	0.186	0.273	0.862

Table 11: 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	6.774	6.29e+05	1.08e-05	0.101	0.148	0.468
Four tag	0.577	6.06e+03	9.52e-05	0.0874	0.128	0.406
Higgs Eta	0.458	3.15e+03	0.000145	0.0962	0.141	0.447
Top veto	0.369	2.31e+03	0.000159	0.0904	0.133	0.420
Higgs signal	0.132	1.02e+02	0.00128	0.153	0.225	0.712

6 Train on resonant test on SM

In Figure 1, there is a peak around 450 GeV in the SM ($\kappa_\lambda = 1$) m_{hh} distribution. This section trains a SPA-NET on the resonant samples with resonant mass 450 GeV, then test its performance on SM samples.

6.1 SPA-NET training on 450 GeV

Set $m_H = 450$ GeV. 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 $|\eta| < 2.5$.

- Training sample:
  - Total sample size: 51,145
  - 1h sample size: 9,320
  - 2h sample size: 40,991
  - 5% used on validation
- Testing sample:
  - Total sample size: 5,683
  - 1h sample size: 1,011
  - 2h sample size: 4,582

The training results are presented in Table 12.

Table 12: SPA-NET training results on the resonant di-Higgs samples with "Four tag cut".

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 4$	0.316	0.992	0.992
$= 5$	0.282	0.903	0.936
≥ 6	0.208	0.768	0.838
Total	0.806	0.903	0.933

6.2 Test SPA-NET on SM event

In the below, the SPA-NET trained in Sec.6.1 is called "resonant SPA-NET". Test the resonant SPA-NET on the non-resonant events. The testing results are in Table 13. The event efficiency is 0.805. This result is worse than "train on resonant test on resonant 0.903" and "train on SM test on SM 0.868".

6.3 Apply the resonant SPA-NET on SM pairing

The "min-ΔR" pairing is replaced by SPA-NET pairing. Other cuts remained unchanged. The cutflow tables for signal and background are in Table 14, 15.

Figure 9 shows the Higgs mass distribution for resonant SPA-NET pairing. The $\frac{S}{\sqrt{B}}$ of resonant SPA-NET results are in Table 16.

Compared to "min-ΔR" method (Table 10), the resonant SPA-NET will let roughly the same number of signal events pass the Higgs signal cut, and let the smaller number of the

Table 13: SPA-NET testing results on the non-resonant di-Higgs samples. Where the SPA-NET is trained on the resonant samples.

N_{Jet}	Event Fraction	Event Efficiency	Higgs Efficiency
$= 4$	0.280	0.940	0.940
$= 5$	0.287	0.790	0.861
≥ 6	0.229	0.649	0.756
Total	0.797	0.805	0.861

Table 14: The selection passing rate and efficiency at each stage for "min-ΔR" and SPA-NET pairing. The SPA-NET is trained in Sec.6.1

	min-ΔR				SPA-NET	
	ATLAS		My sample		My sample	
Cut	pass rate	efficiency	pass rate	efficiency	pass rate	efficiency
Four tag	0.0649	0.0649	0.0852	0.0852	0.0852	0.0852
Higgs Eta	0.0543	0.8360	0.0688	0.8074	0.0685	0.8022
Top veto	0.0456	0.8401	0.0553	0.8044	0.0550	0.8048
Signal region	0.0220	0.4818	0.0181	0.3283	0.0181	0.3296

Table 15: The selection passing rate and efficiency at each stage for "min-ΔR" and SPA-NET pairing. The SPA-NET is trained in Sec.6.1

	pp4b			
	min-ΔR		SPA-NET	
Cut	pass rate	efficiency	pass rate	efficiency
Four tag	0.0096	0.0096	0.0096	0.0096
Higgs Eta	0.0056	0.5835	0.0054	0.5589
Top veto	0.0042	0.7423	0.0039	0.7326
Signal region	0.0001	0.0232	0.0001	0.0200

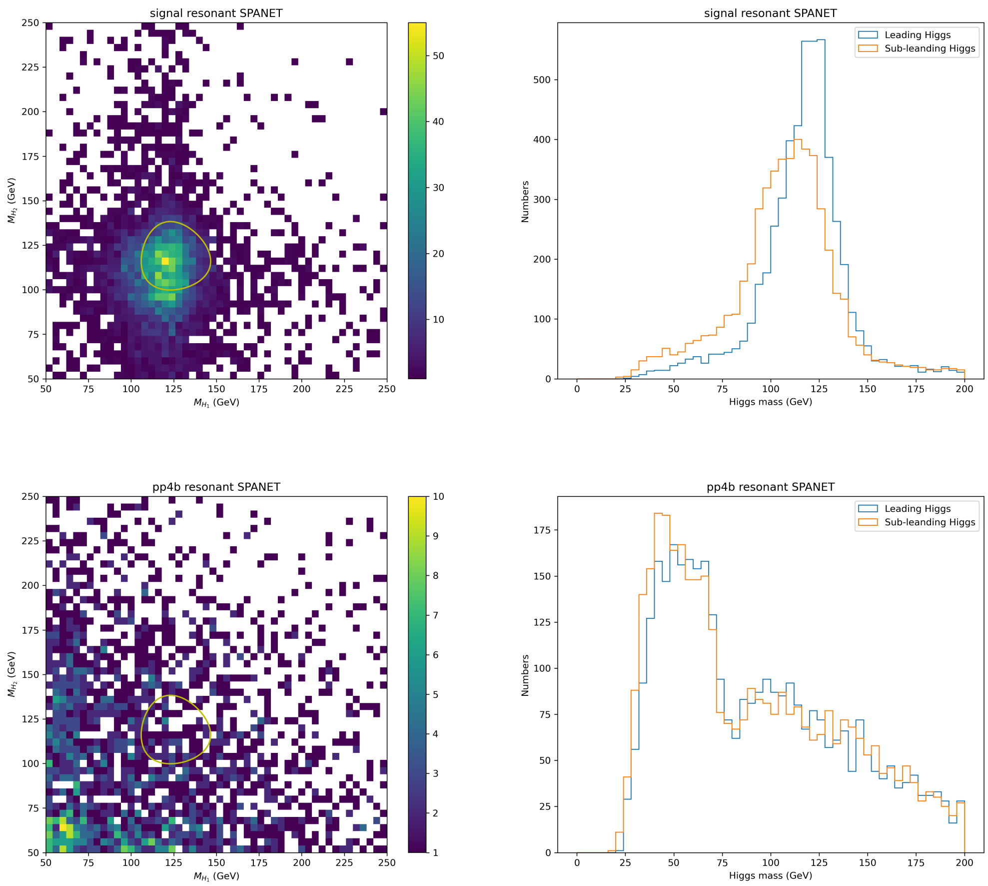

Figure 9: The mass plane and distribution for Higgs candidate for resonant SPA-NET pairing. The above figure is for the signal sample and the below one is for the background sample.

background events pass, but not so much difference. Therefore, the resonant SPA-NET can a little improve the $S/\sqrt{B}$.

Table 16: The cross sections for the di-Higgs signal and background processes at different cuts. The pairing method is resonant SPA-NET. Compared to Table 10, there is a little improvement.

	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	6.774	6.29e+05	1.08e-05	0.101	0.148	0.468
Four tag	0.577	6.06e+03	9.52e-05	0.0874	0.128	0.406
Higgs Eta	0.464	3.39e+03	0.000137	0.0941	0.138	0.437
Top veto	0.372	2.48e+03	0.00015	0.0881	0.129	0.410
Higgs signal	0.123	49.7	0.00247	0.205	0.302	0.954

7 Apply the min-ΔR method on the resonant sample

In the previous, the pairing method for resonant samples is ΔR+min-D_{HH} method. This section will test the min-ΔR pairing method for resonant samples. The "$\Delta R +$ min-D_{HH}" pairing is replaced by min-ΔR pairing. Other cuts remained unchanged. The cutflow tables for signal and background are in Table 17, 18.

Table 17: The selection passing rate and efficiency at each stage for ΔR+min-D_{HH}, min-ΔR and SPA-NET pairing. The resonant mass is 1000 GeV.

	$\Delta R +$ min-D_{HH}		min-ΔR		SPA-NET	
Cut	pass rate	efficiency	pass rate	efficiency	pass rate	efficiency
Four tag	0.1257	0.1257	0.1257	0.1257	0.1257	0.1257
Delta R	0.1087	0.8644	N.A.	N.A.	N.A.	N.A.
Higgs PT	0.0896	0.8247	0.0936	0.7444	0.0978	0.7775
Higgs Eta	0.0826	0.9216	0.0863	0.9226	0.0900	0.9203
Higgs signal	0.0445	0.5391	0.0445	0.5155	0.0468	0.5197
Top veto	0.0425	0.9555	0.0425	0.9553	0.0446	0.9536

Figure 10 shows the Higgs mass distribution for resonant signal events with different pairing methods. Figure 11 is for background events with different pairing methods. The

Table 18: The selection passing rate and efficiency at each stage for ΔR+min-D_{HH}, min-ΔR and SPA-NET pairing.

	pp4b					
	$\Delta R+$ min-D_{HH}		min-ΔR		SPA-NET	
Cut	pass rate	efficiency	pass rate	efficiency	pass rate	efficiency
Four tag	0.0096	0.0096	9.6420e-03	0.0096	9.6420e-03	0.0096
Delta R	0.0050	0.5234	9.6420e-03	1.0000	9.6420e-03	1.0000
Higgs PT	0.0043	0.8462	6.6460e-03	0.6893	6.7200e-03	0.6970
Higgs Eta	0.0030	0.7127	5.0040e-03	0.7529	4.8580e-03	0.7229
Higgs signal	0.0005	0.1551	1.1600e-04	0.0232	7.5000e-05	0.0154
Top veto	0.0003	0.5975	9.4000e-05	0.8103	6.6000e-05	0.8800

mass planes for the signal process all look similar. For background, $\Delta R+$ min-D_{HH} pairing is much different from the other two methods.

The results of $S/\sqrt{B}$ are presented in Table 19, 20, 21. Compared to the ΔR+min-D_{HH} pairing method, the min-ΔR and SPA-NET pairing method can reduce more background events, so they can improve the $S/\sqrt{B}$. The SPA-NET method has the best performance.

Table 19: The cross sections for the di-Higgs signal and background processes at different cuts. The pairing method is $\Delta R+$ 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	6.29e+05	1.02e-06	0.0096	0.0141	0.0445
Four tag	0.081	6.06e+03	1.34e-05	0.0123	0.0180	0.0570
Delta R	0.070	3.17e+03	2.21e-05	0.0147	0.0215	0.0681
Higgs PT	0.058	2.68e+03	2.15e-05	0.0131	0.0193	0.0610
Higgs Eta	0.053	1.91e+03	2.78e-05	0.0143	0.0211	0.0666
Higgs signal	0.029	2.97e+02	9.67e-05	0.0196	0.0288	0.0912
Top veto	0.027	1.77e+02	0.000155	0.0243	0.0357	0.1128

8 Hyperparameter tuning

Some hyperparameters of SPA-NET:

- Learning rate: 0.0007

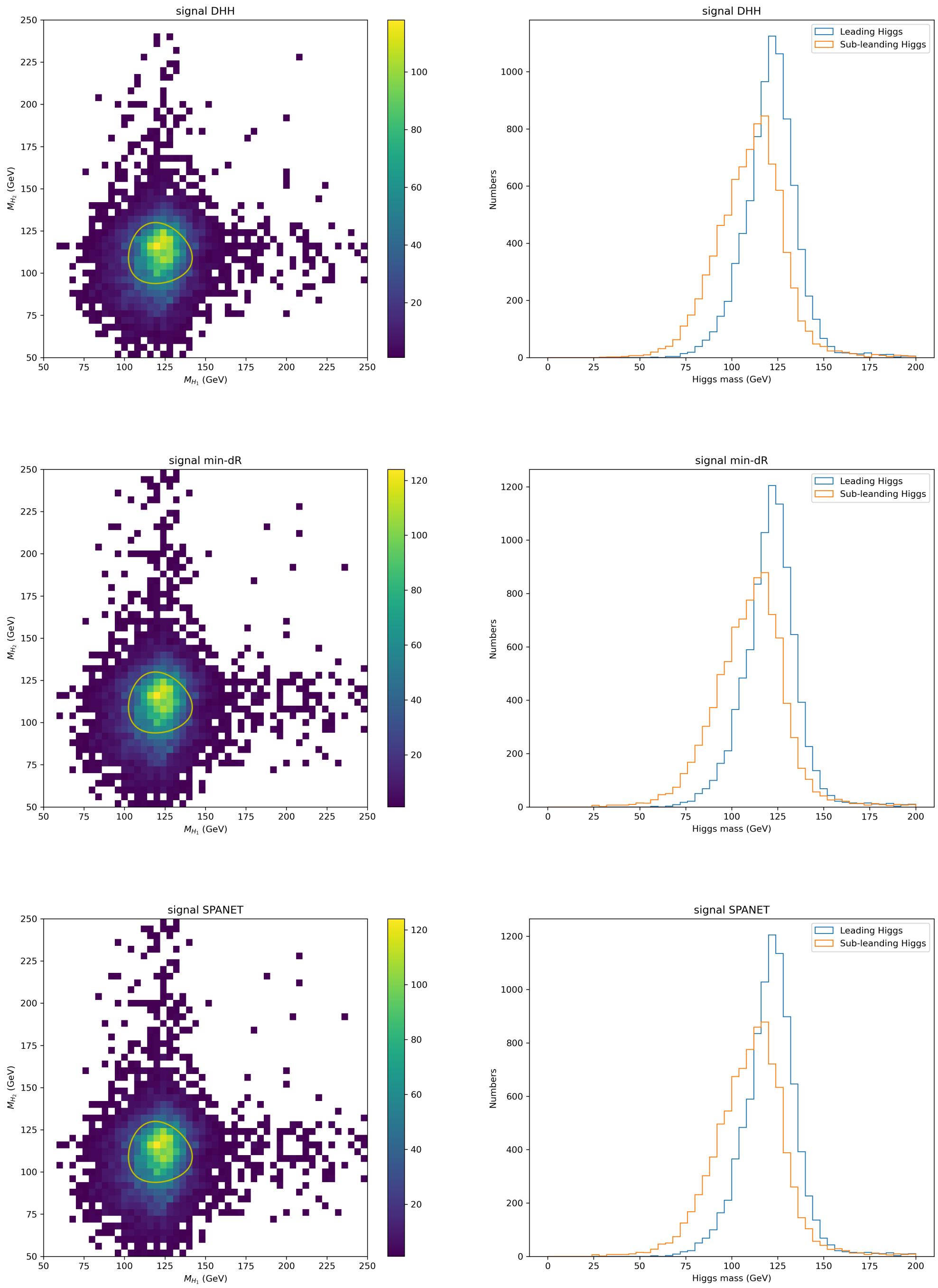

Figure 10: The mass plane and distribution of Higgs candidate for resonant signal events with different pairing methods. The top one is $\Delta R + \text{min-}D_{HH}$ pairing, the middle one is min-ΔR pairing, bottom one is SPA-NET pairing.

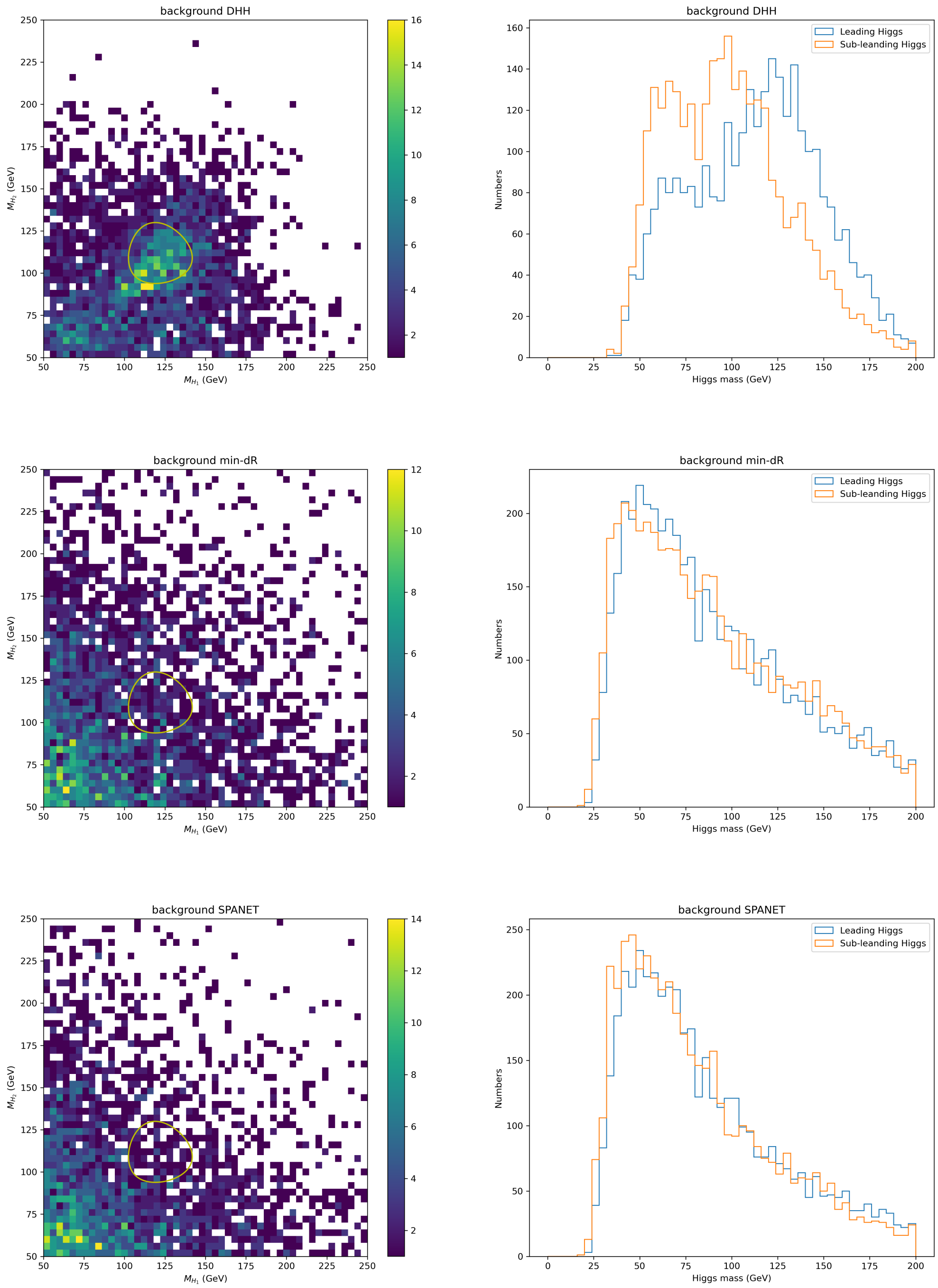

Figure 11: The mass plane and distribution of Higgs candidate for background events with different pairing methods. The top one is ΔR+min-D_{HH} pairing, the middle one is min-ΔR pairing, bottom one is SPA-NET pairing.

Table 20: The cross sections for the di-Higgs signal and background processes at different cuts. The pairing method is min-ΔR.

	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	6.29e+05	1.02e-06	0.0096	0.0141	0.0445
Four tag	0.081	6.06e+03	1.34e-05	0.0123	0.0180	0.0570
Higgs PT	0.060	4.18e+03	1.44e-05	0.0110	0.0162	0.0511
Higgs Eta	0.056	3.15e+03	1.77e-05	0.0117	0.0172	0.0543
Higgs signal	0.029	72.9	0.000393	0.0396	0.0582	0.1839
Top veto	0.027	59.1	0.000464	0.0420	0.0617	0.1952

Table 21: 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	6.29e+05	1.02e-06	0.0096	0.0141	0.0445
Four tag	0.081	6.06e+03	1.34e-05	0.0123	0.0180	0.0570
Higgs PT	0.063	4.22e+03	1.49e-05	0.0114	0.0168	0.0531
Higgs Eta	0.058	3.05e+03	1.9e-05	0.0124	0.0182	0.0574
Higgs signal	0.030	47.1	0.000639	0.0517	0.0760	0.2403
Top veto	0.029	41.5	0.000692	0.0526	0.0772	0.2443

- 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.6.

8.1 Scan hyperparameters

In the below, use the same samples in Sec.6.1 for hyperparameter scan.

Learning rate Range: $[0.0003, 0.0030]$. Step size: 0.0003. Figure 12 is the scanning results. When the learning rate is equal to 0.0018, the model will get the best performance at the first 10 epochs, but the difference between different learning rates is very small. The model cannot be trained when the learning rate is greater than 0.0030.

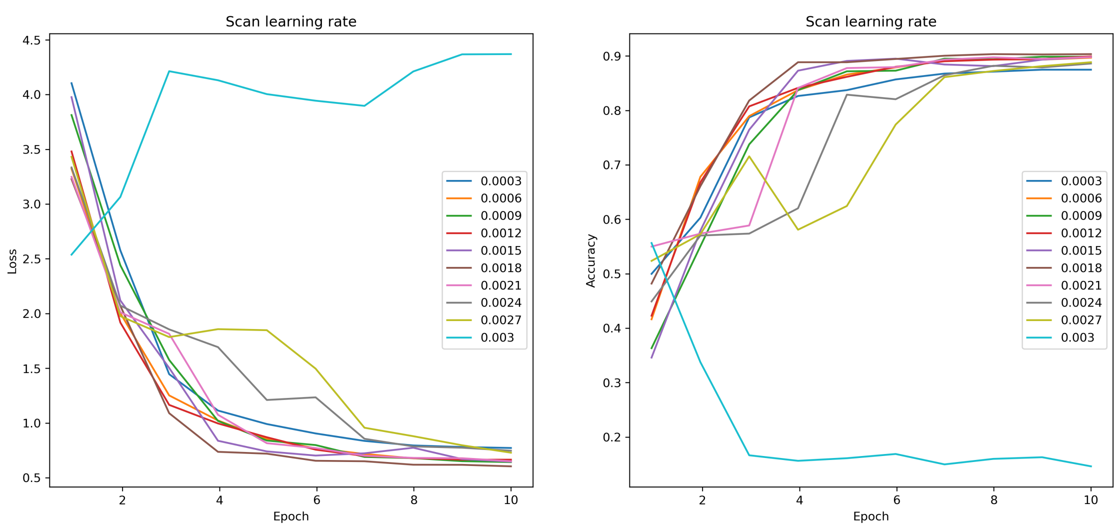

Figure 12: SPA-NET training results for different learning rates.

Central encoder count Range: $[2, 6]$. Step size: 1. Figure 13 is the scanning results. When the number of central encoder layers is equal to 3, the model will get the best performance at the first 10 epochs.

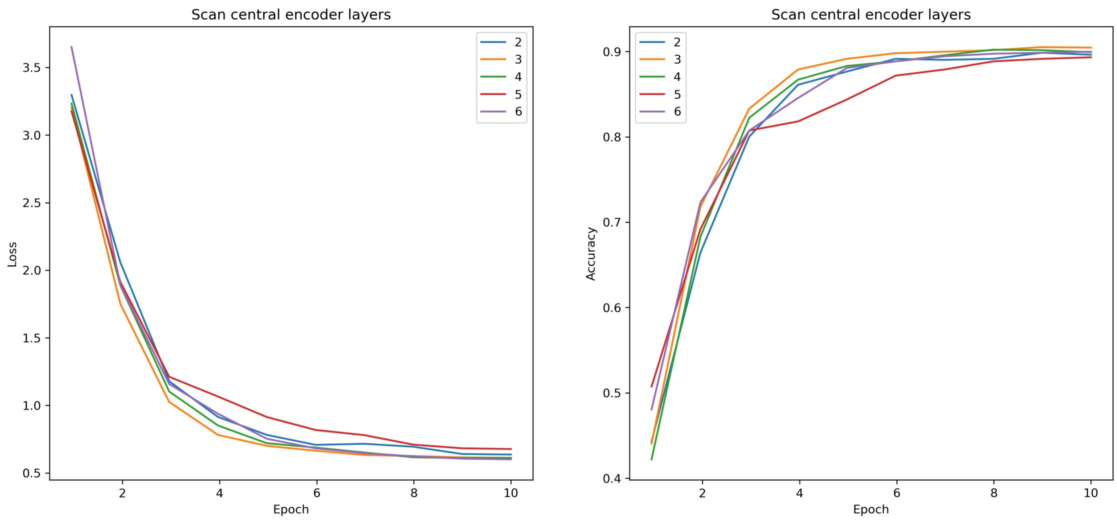

Figure 13: SPA-NET training results for different central encoder layers.

Branch encoder count Range: $[2, 7]$. Step size: 1. Figure 14 is the scanning results. When the number of branch encoder layers is equal to 6, the model will get the best performance at the first 10 epochs.

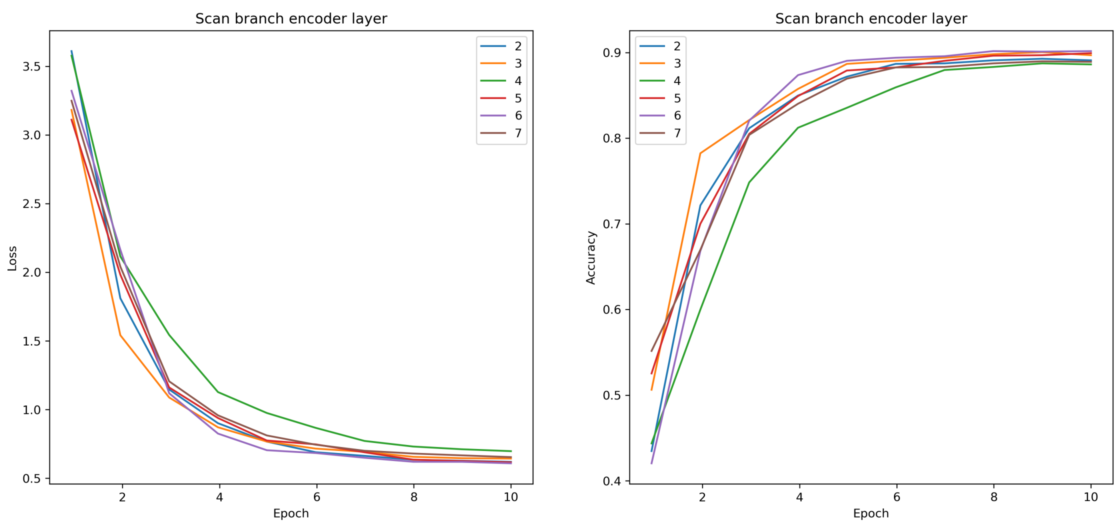

Figure 14: SPA-NET training results for different branch encoder layers.

8.2 Training results for different hyperparamers

This section uses the same samples in Sec.6.1 for training. Train 50 epochs. Table 22 is the training results for different hyperparameters. Table 23 is the testing results on SM samples.

Table 22: SPA-NET training results for different hyperparameters.

Case	Central encoder	Branch encoder	learning rate	Event efficiency	Higgs efficiency
1	6	3	0.0007	0.902	0.932
2	6	3	0.0018	0.870	0.898
3	3	3	0.0007	0.908	0.936
4	6	6	0.0007	0.911	0.941
5	3	6	0.0018	0.883	0.911
6	3	6	0.0007	0.910	0.940

Table 23: SPA-NET testing results on SM samples for different hyperparameters.

Case	Central encoder	Branch encoder	learning rate	Event efficiency	Higgs efficiency
1	6	3	0.0007	0.810	0.863
2	6	3	0.0018	0.824	0.858
3	3	3	0.0007	0.831	0.876
4	6	6	0.0007	0.812	0.865
5	3	6	0.0018	0.829	0.863
6	3	6	0.0007	0.793	0.854

8.3 Use the best SPA-NET on selection

The section uses the best SPA-NET trained in Sec.8.2 for pairing (Case 3 and 5). Other cuts remained unchanged. The cutflow tables for signal and background are presented in Table 24, 25.

Compared to the SPA-NET trained in Sec.6.1 (Table 16), the results are similar. Consider $S/\sqrt{B}$, for case 3: 0.2123, for case 5: 0.2060, for case 1: 0.2050 ($\mathcal{L} = 139$ fb^{-1}).

9 Constraints on κ_λ

In this section, the m_{HH} distribution is used to set the constraints of κ_λ.

The binned m_{HH} distribution is used. The likelihood function L consisting of a product of Poisson distributions

$$L(\kappa_\lambda \mid \text{data}) = \prod_{i=1}^{B} \text{Pois}(n_i \mid n_{i,\text{exp}}) \tag{2}$$

where B is the number of bins, n_i is the number of events in bin i from data, $n_{i,\text{exp}}$ is the expected number of events in bin i. The expected number of events is the sum of the signal

Table 24: The cross sections for the di-Higgs signal and background processes at different cuts. The pairing method is the SPA-NET trained in Sec.8.2 case 3. The results are similar compared to Table 16.

	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	6.774	6.29e+05	1.08e-05	0.1007	0.1480	0.4680
Four tag	0.577	6.06e+03	9.52e-05	0.0874	0.1284	0.4059
Higgs Eta	0.464	3.34e+03	0.000139	0.0948	0.1392	0.4403
Top veto	0.373	2.44e+03	0.000153	0.0891	0.1309	0.4139
Higgs signal	0.125	48.4	0.00259	0.2123	0.3118	0.9862

Table 25: The cross sections for the di-Higgs signal and background processes at different cuts. The pairing method is the SPA-NET trained in Sec.8.2 case 5. The results are similar compared to Table 16.

	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	6.774	6.29e+05	1.08e-05	0.1007	0.1480	0.4680
Four tag	0.577	6.06e+03	9.52e-05	0.0874	0.1284	0.4059
Higgs Eta	0.459	3.3e+03	0.000139	0.0941	0.1383	0.4374
Top veto	0.369	2.43e+03	0.000152	0.0882	0.1296	0.4100
Higgs signal	0.125	50.9	0.00245	0.2060	0.3026	0.9568

and background events, so $n_{i,\mathrm{exp}}$ depend on κ_λ.

From Poisson distribution

$$\mathrm{Pois}(n \mid \lambda) = \frac{\mathrm{e}^{-\lambda}\lambda^n}{n!} \tag{3}$$

then

$$\begin{aligned} \ln\left(\prod_{i=1}^{B} \mathrm{Pois}(n_i \mid n_{i,\mathrm{exp}})\right) &= \sum_{i=1}^{B} \ln\left(\frac{\mathrm{e}^{-n_{i,\mathrm{exp}}} n_{i,\mathrm{exp}}^{n_i}}{n_i!}\right) \\ &= \sum_{i=1}^{B} \left[-n_{i,\mathrm{exp}} + n_i \ln(n_{i,\mathrm{exp}}) - \ln(n_i!)\right] \end{aligned} \tag{4}$$

$\sum_{i=1}^{B} \ln(n_i)$ is independent of κ.

Define the test statistic as follow

$$-2\Delta\ln(L) \equiv -2\ln\left(\frac{L(\kappa_\lambda)}{L(\hat{\kappa}_\lambda)}\right) \tag{5}$$

where $\hat{\kappa}_\lambda$ is the maximum likelihood estimate of κ_λ.

9.1 m_{HH} distribution

Figure 15 and 16 are the m_{HH} distribution of min-ΔR and SPA-NET pairing method, respectively. The events passing all selection cuts are used. For signal, both pairing methods give similar results. For background, the results look different.

Parameter setting:

- Number of bins: 10
- Range: $[200, 1000]$
- Luminosity: $\mathcal{L} = 139\ \mathrm{fb}^{-1}$

9.2 Expected limit of κ_λ

For the expected limit setting, n_i is the background data. The likelihood of different κ_λ is calculated. The range of κ_λ is $[-10, 15]$ by step 1.

Figure 17 and 18 are the profile likelihood ratio scans for κ_λ with min-ΔR and SPA-NET pairing method. The 2σ exclusion limit: The values of κ_λ beyond $[-5.01, 12.21]$ are excluded for min-ΔR method; The values of κ_λ beyond $[-5.30, 13.02]$ are excluded for SPA-NET method.

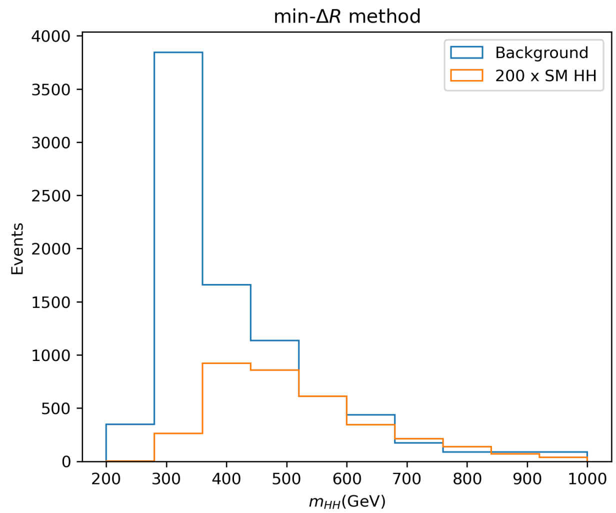

Figure 15: The m_{HH} distribution with min-ΔR pairing method.

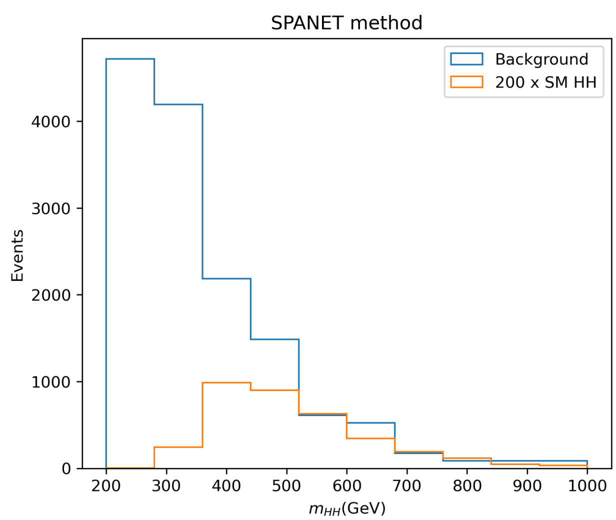

Figure 16: The m_{HH} distribution with SPA-NET pairing method.

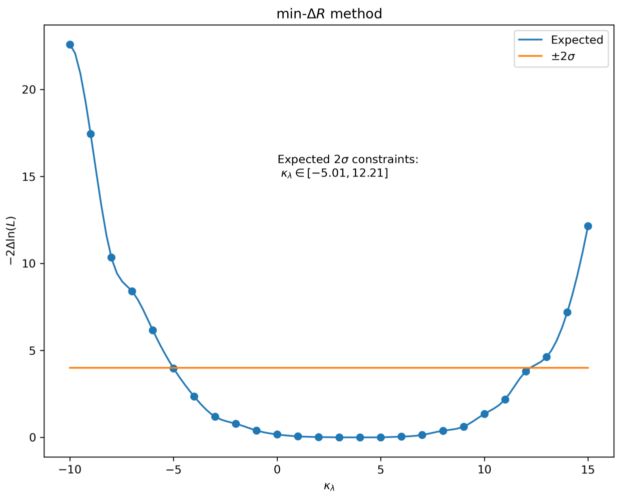

Figure 17: The profile likelihood ratio scans for κ_λ. The pairing method is the min-ΔR method. The orange line indicates the 2σ exclusion boundary.

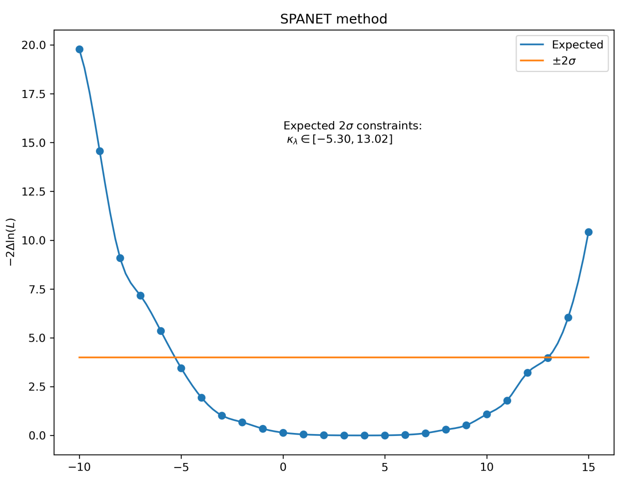

Figure 18: The profile likelihood ratio scans for κ_λ. The pairing method is the SPA-NET method. The orange line indicates the 2σ exclusion boundary.

9.3 CLs method

This section uses the CL_s method to set the constraints of κ_λ. The signal strength is chosen as the parameter of interest. The POI is excluded at the 95% confidence level when the CL_s is less than 0.05. Where the package `pyhf` is used to calculate the upper limit. When the upper limit of signal strength is got, then it can be converted to the upper limit of the cross-section.

Figure 19 and 20 are the upper limits scans for κ_λ with min-ΔR and SPA-NET pairing method. The values of κ_λ beyond $[-4.93, 12.03]$ are excluded for min-ΔR method; The values of κ_λ beyond $[-5.22, 12.86]$ are excluded for SPA-NET method.

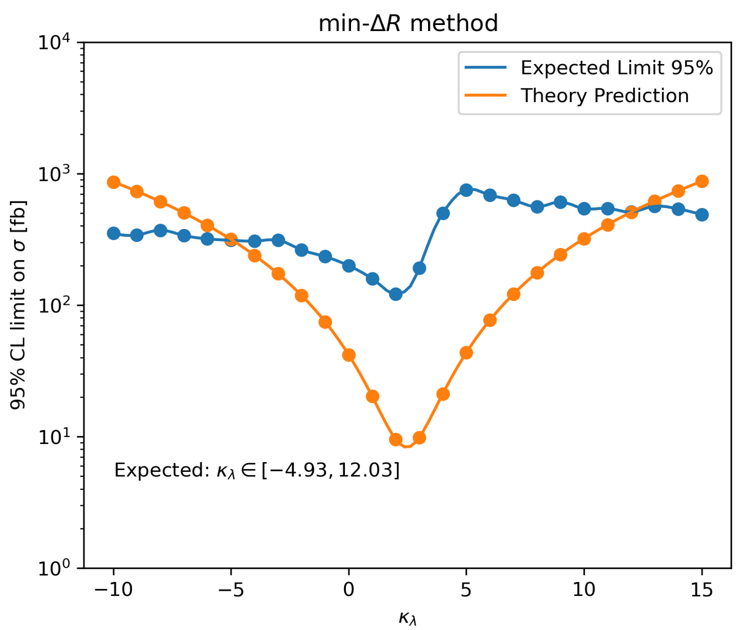

Figure 19: The upper limit of the cross-section with different κ_λ. The pairing method is the min-ΔR method.

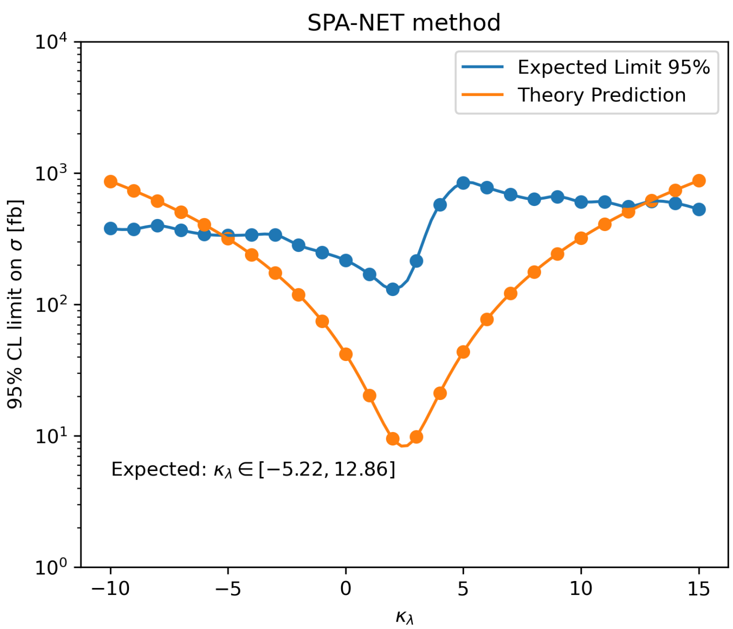

Figure 20: The upper limit of the cross-section with different κ_λ. The pairing method is the SPA-NET method.