Detecting anomaly in vector boson scattering

Jinmian Li

四川大學
SICHUAN UNIVERSITY

Jan. 28th, 2021

VBS at Snowmass Workshop

based on arXiv: 2010.13281, JL, Shuo Yang, Rao Zhang

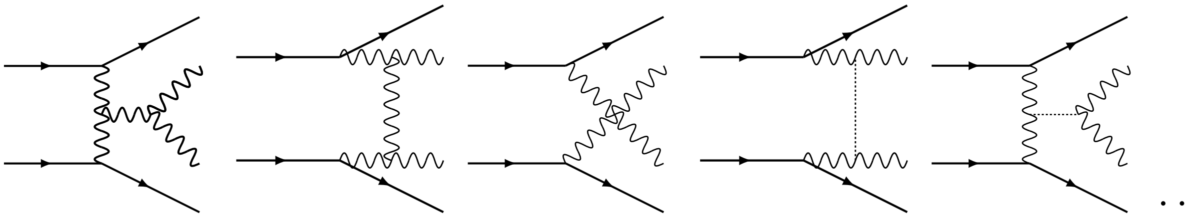
Outline

1. Vector boson scattering in the SM
   - The W^+W^-jj signal and its backgrounds
2. The neural network and shape analysis
3. Resolve the features of polarization for W^+W^-jj channel
4. Applications to new physics searches
   - The effective field theory
   - Two Higgs Doublet Model

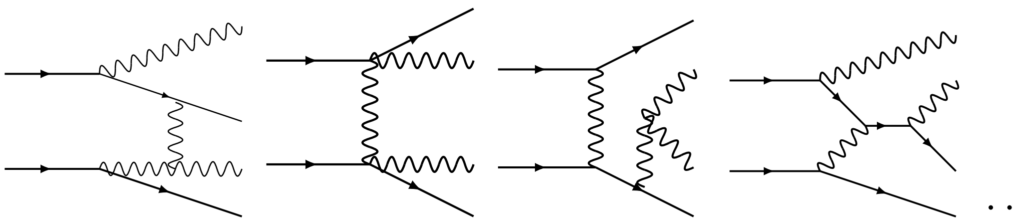
Outline

1. Vector boson scattering in the SM
   - The W^+W^-jj signal and its backgrounds
2. The neural network and shape analysis
3. Resolve the features of polarization for W^+W^-jj channel
4. Applications to new physics searches
   - The effective field theory
   - Two Higgs Doublet Model

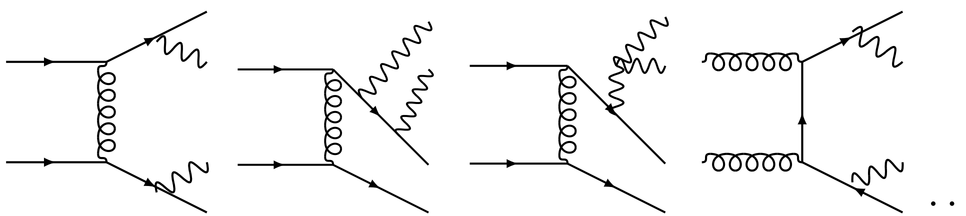
- The vector boson scattering processes ($\mathcal{M}^2 \propto \mathcal{O}(\alpha_{\text{EW}}^4)$)



- Irreducible electroweak backgrounds ($\mathcal{M}^2 \propto \mathcal{O}(\alpha_{\text{EW}}^4)$)

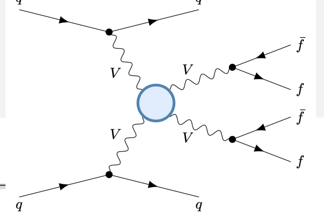


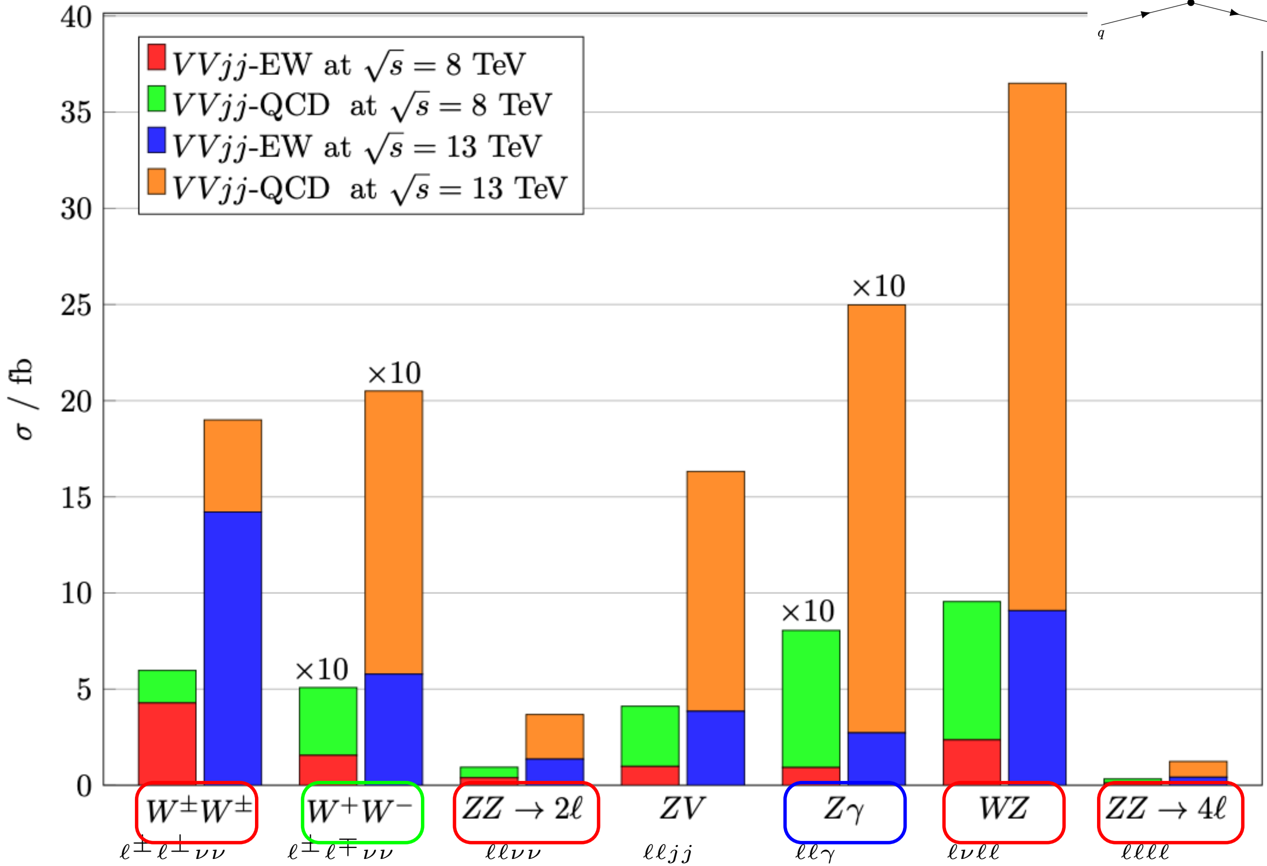
- QCD induced contributions ($\mathcal{M}^2 \propto \mathcal{O}(\alpha_{\text{EW}}^2 \alpha_S^2)$)



- Interference between the EW and QCD amplitudes ($\mathcal{M}^2 \propto \mathcal{O}(\alpha_{\text{EW}}^3 \alpha_S)$)
  - Typically smaller than $\mathcal{O}(3)\%$.

Cross section of VBS processes at the LHC





[CERN-THESIS-2014-105]

Why measure the polarization modes of W^+W^-jj?

- The amplitude for the longitudinal vector boson scattering:

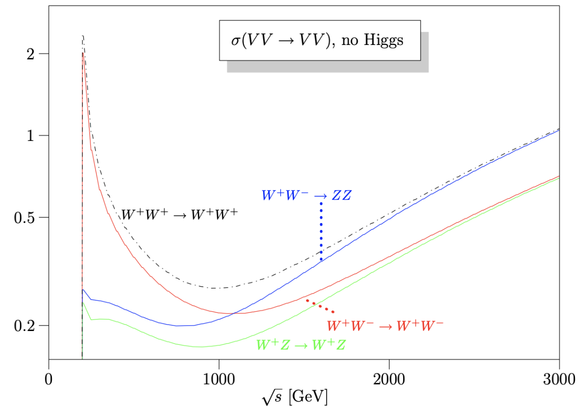
$$\mathcal{M}^{\text{no Higgs}} = -\frac{g_w^2}{4m_V^2}u + \mathcal{O}(s^0)$$

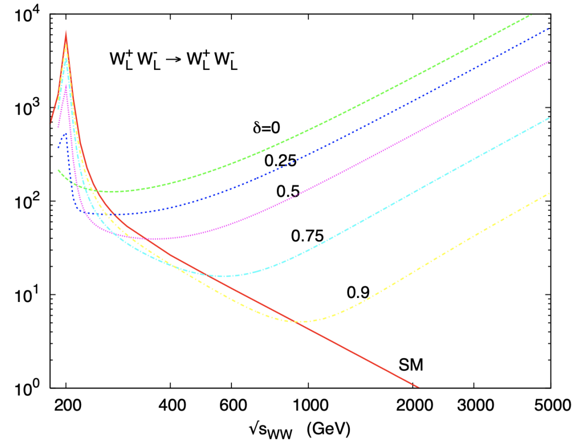
- it cancels with the amplitude from Higgs exchange

$$\mathcal{M}^{\text{Higgs}} = -\frac{g_w^2}{4m_W^2}\left[\frac{(s-m_W^2)^2}{s-m_H^2} + \frac{(t-m_W^2)^2}{t-m_H^2}\right]$$

$$\overset{s,t,u \gg m_W, m_H}{\sim} \frac{g_w^2}{4m_W^2}u$$

- The longitudinal mode $W_L W_L jj$ is the most sensitive to the new physics which affect the EWSB



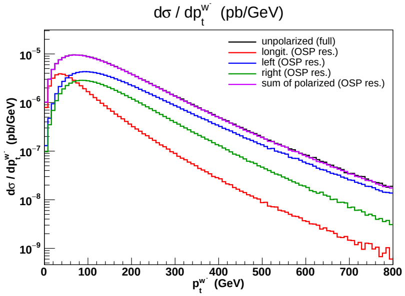


K. Cheung et.al.(2008), A. Alboteanu et. al.(2008)

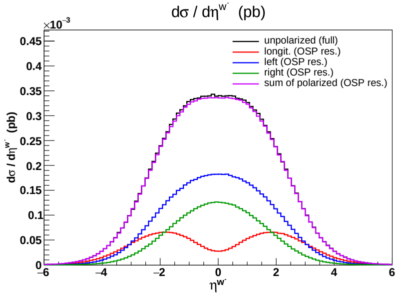
Variables sensitive to polarizations of W^+W^-jj

W boson is reconstructable in semi-leptonic channel:

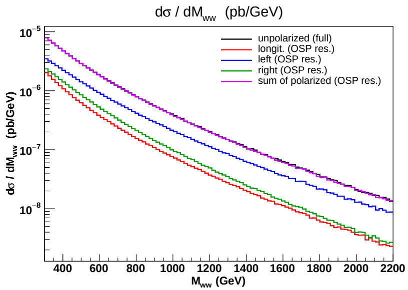
p_T^W



η^W

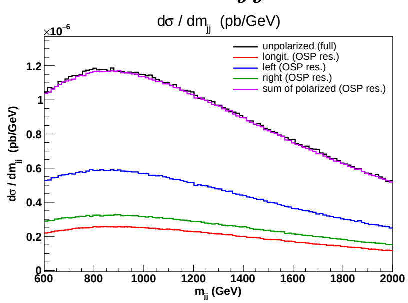


m^{WW}

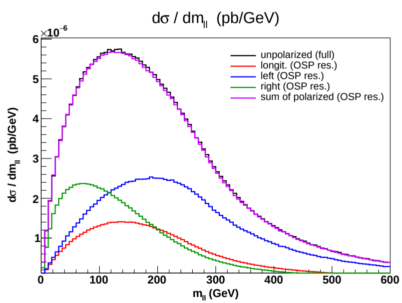


W boson is NOT reconstructable in dileptonic channel:

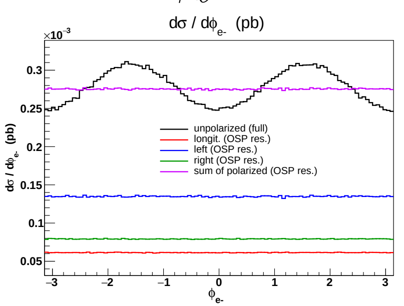
m_{jj}



$m_{\ell\ell}$



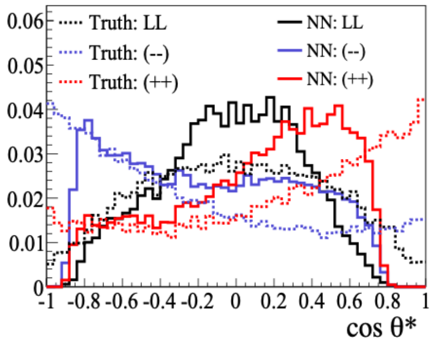
ϕ_{ℓ^-}

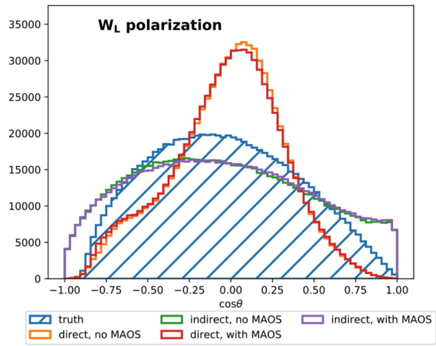


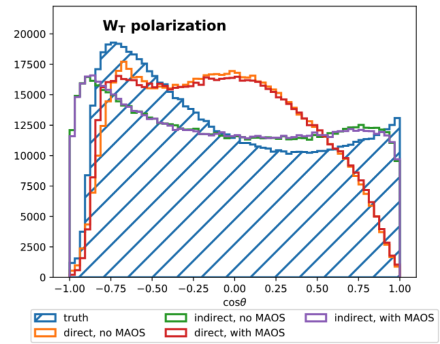
[JHEP03(2018)170]

Machine learning methods for polarizations discrimination

- Regression of the lepton angle in the gauge boson rest frame ($W_\ell^\pm W_\ell^\pm jj$)

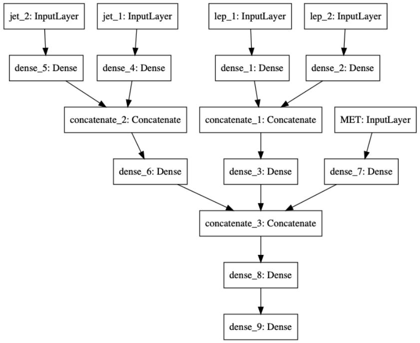


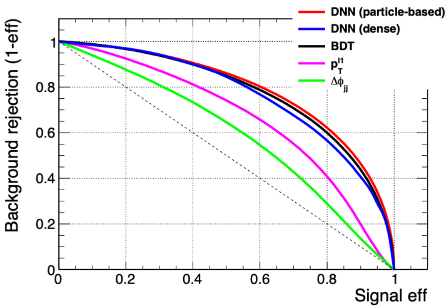


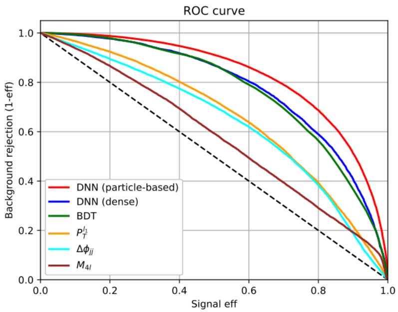


[J. Searcy *et.al* (2016), M. Grossi *et.al* (2020)]

- Classification of events from different polarizations for $W_\ell^\pm W_\ell^\pm jj$ and $ZZjj$





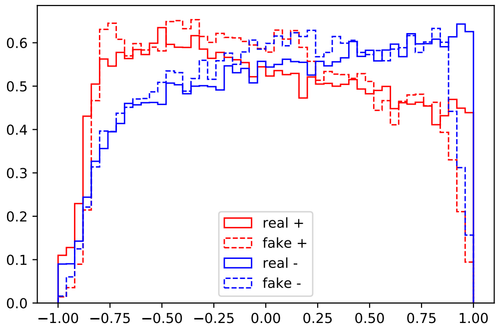


[J. Lee *et.al* (2019)] Signal: LL; Background: TT&TL

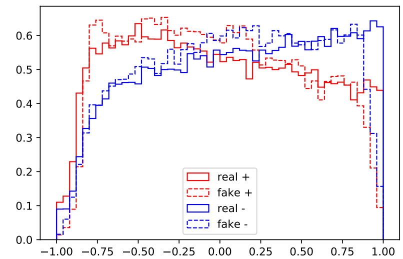
Resolve polarization of different processes? Study in progress · · ·

Train a network that is capable of regress the lepton angle in the W boson rest frame for the dileptonic SM $WWjj$ channel (template fitting).

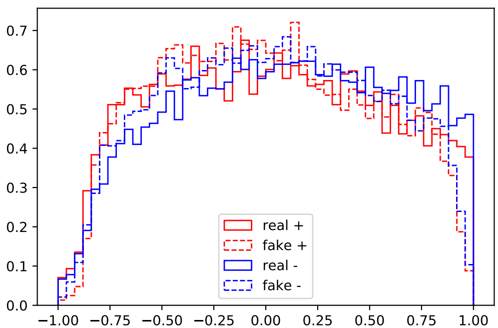
$\bar{c}_H = -1$	LL	LT	TL	TT
True	0.10964	0.20203	0.19765	0.49091
Pred.	0.12239	0.17786	0.13348	0.56637



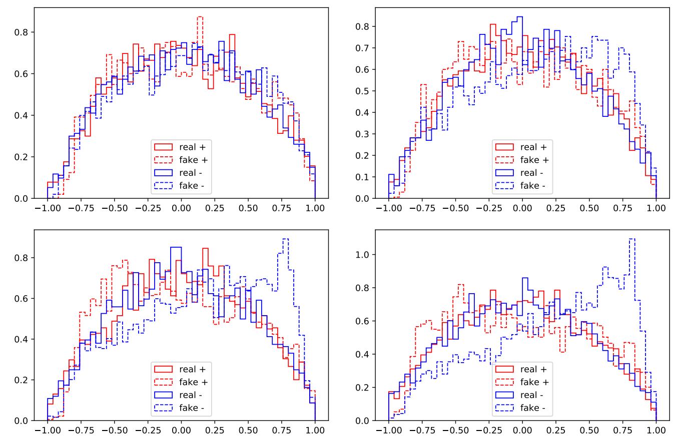
2HDM_700_095	LL	LT	TL	TT
True	0.09586	0.19379	0.18666	0.5244
Pred.	0.09335	0.22685	0.16057	0.51946



2HDM_300_07	LL	LT	TL	TT
True	0.25975	0.21256	0.22687	0.30299
Pred.	0.28877	0.18845	0.22318	0.30053



$pp \to H \to WWjj$, with m_H =400, 600, 800, 1200 GeV



The event generation setup

The aim of the talk

Focus on the $W^{\pm}W^{\mp}jj$ process, considering the polarization information; propose a method to search general new physics in this final state

- Use the new feature of MadGraph_aMC@NLO to simulate the polarized sample:

generate p p > w+{0/T } w-{0/T} j j QED<=4 QCD=0

- Consider the interferences of amplitudes at α_{EW}^4, ignore the interferences with amplitude at $\alpha_S^2\alpha_{EW}^2$.
- Madspin framework for vector boson decay.
- Backgrounds for semi-(di-)leptonic channels

	σ^{fid} [pb]	$\sigma^{\ell\ell}$ [fb]	$\sigma^{\ell j}$ [fb]
tt_ℓ	210.3	139.8	3007.6
$tW_\ell/t_\ell W$	15.9	11.6	224.6
$W_\ell Wjj^{\text{QCD}}$	4.68	14.7	340.5
$W_\ell Zjj^{\text{QCD}}$	2.20	4.49	165.7
$W_\ell Zjj^{\text{EW}}$	0.487	3.68	22.2
$W_\ell Wjj^{\text{EW}}$	0.738	4.36	37.3

superscript QCD (EW) corresponds to amplitudes at $\alpha_S^2\alpha_{EW}^2$ (α_{EW}^4).
$\sigma^{\ell\ell}(\sigma)^{\ell j}$ corresponds to cross section after preselection cuts

Preselection cuts

Di-leptonic cuts:

- exactly two opposite sign leptons with $p_T(\ell) > 20$ GeV, $|\eta(l)| < 2.5$
- at least two jets with $p_T(j) > 20$ GeV, $|\eta(j)| < 4.5$
- two jets with leading p_T should have $m_{jj} > 500$ GeV and $|\Delta\eta|_{jj} > 3.6$
- no b-tagged jet in the final state.

Input to the network:

- momenta and charge of two leptons;
- forward and backward jets;
- sum of all detected particles ($\vec{\not{p}}$);
- sum of jets that are not assigned as forward-backward jets

Semi-leptonic cuts:

- exactly one charged lepton with $p_T(\ell) > 20$ GeV, $|\eta(l)| < 2.5$
- at least four jets with (b-veto) $p_T(j) > 20$ GeV, $|\eta(j)| < 4.5$
- the pair of jets with the largest invariant mass ($m_{jj} > 500$ GeV) that also satisfies $|\Delta\eta|_{jj} > 3.6$ is taken as the forward-backward jet pair
- In the remaining jets, the jet pair with invariant mass closest to the W boson mass is regarded as the jet pair from W decay.

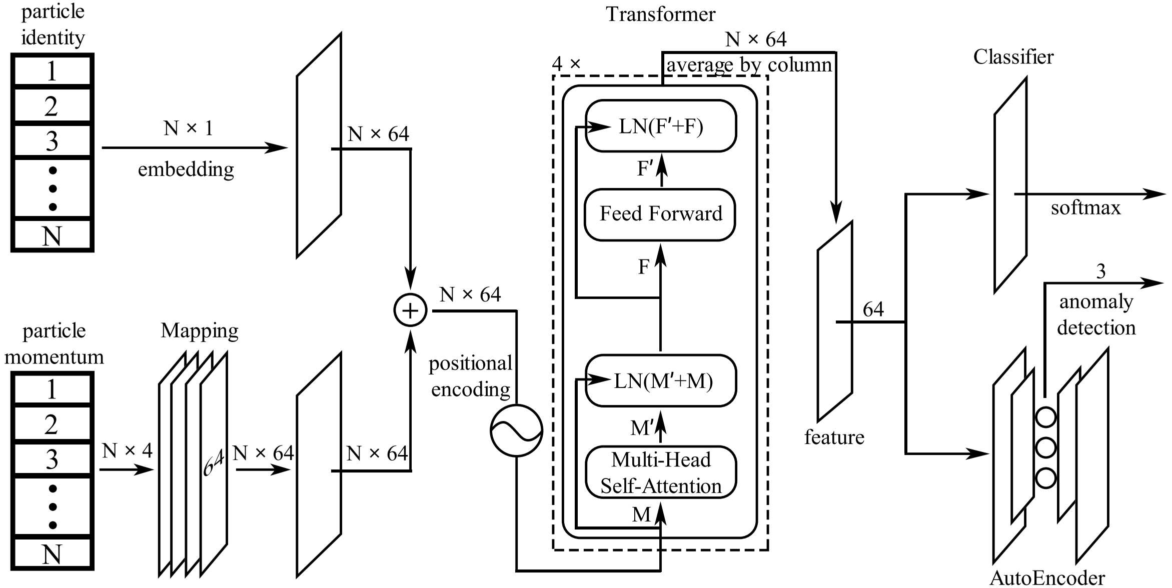
Inputs to the Network:

- momenta of the lepton;
- forward and backward jets;
- two jets from W decay;
- sum of all detected particles;
- sum of remaining jets;

Outline

1. Vector boson scattering in the SM
   - The W^+W^-jj signal and its backgrounds
2. The neural network and shape analysis
3. Resolve the features of polarization for W^+W^-jj channel
4. Applications to new physics searches
   - The effective field theory
   - Two Higgs Doublet Model

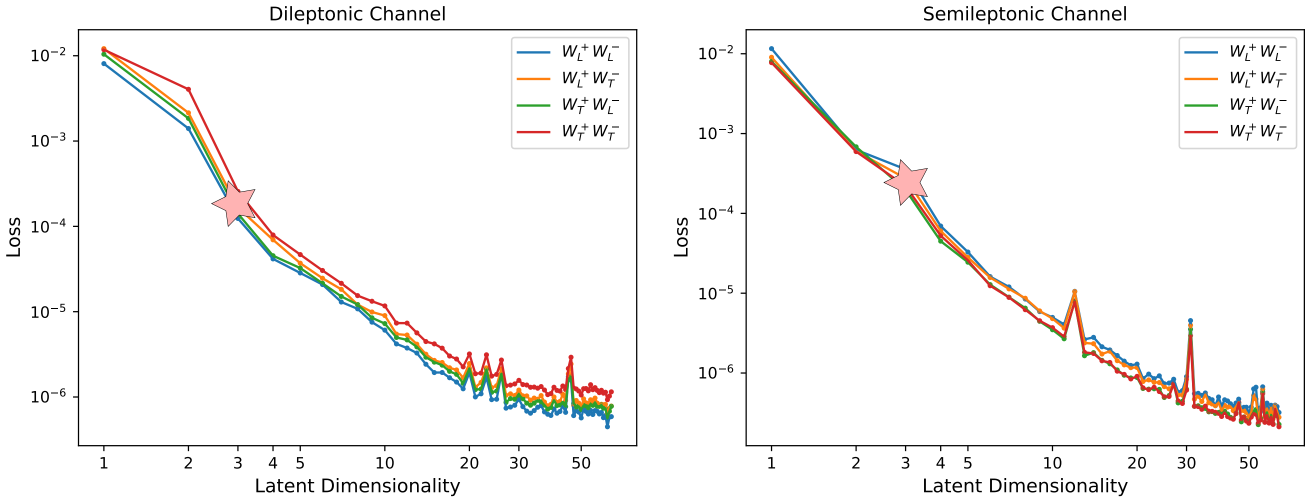
Architecture of neural network



- Normalization: $\hat{p}_i^{\mu} = \frac{p_i^{\mu} - \bar{p}^{\mu}}{\sigma_{p^{\mu}}}$
- Self-attention layer: $W^{Q,K,V} = M_{N\times 64} \cdot W'^{Q,K,V}_{64\times 16}$
- $W'^{Q,K,V,O}$ are trainable parameters

$$M'_{N\times 64} = [\text{Softmax}(\frac{W_1^Q (W_1^K)^T}{8}) W_1^V, \cdots, \text{Softmax}(\frac{W_4^Q (W_4^K)^T}{8}) W_4^V]_{N\times 64} \cdot W'^{O}_{64\times 64}$$

Training and Hyper parameters



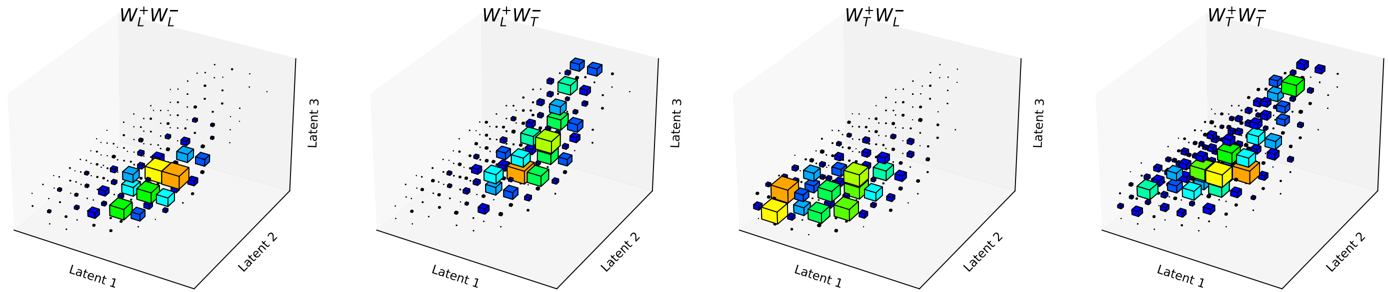
Competition of statistical uncertainty and information completeness.

Shape analysis (binned log-likelihood test)

- 3-dimensional latent space is divided into $8 \times 8 \times 8$ bins for dileptonic channel and $10 \times 10 \times 10$ bins for semi-leptonic channel,
- the bins which contain at least 1% of total signal events, ten with highest signal to background ratios are selected for the log-likelihood test.
  In realistic experiments, the number of signal in each bin can be obtained by subtracting the predicted background event number from the measured number.
- This procedure selects $\sim 30\%$ of signal events and $\sim 0.5\%$ of total background events in most of the cases ($N_B \gtrsim N_S$)
- The pseudo-data is obtained by generating a random number from Poissonian (statistical uncertainty) plus Gaussian distribution (systematical uncertainty)
- The p-value is calculated by assuming observation is given by null hypothesis.

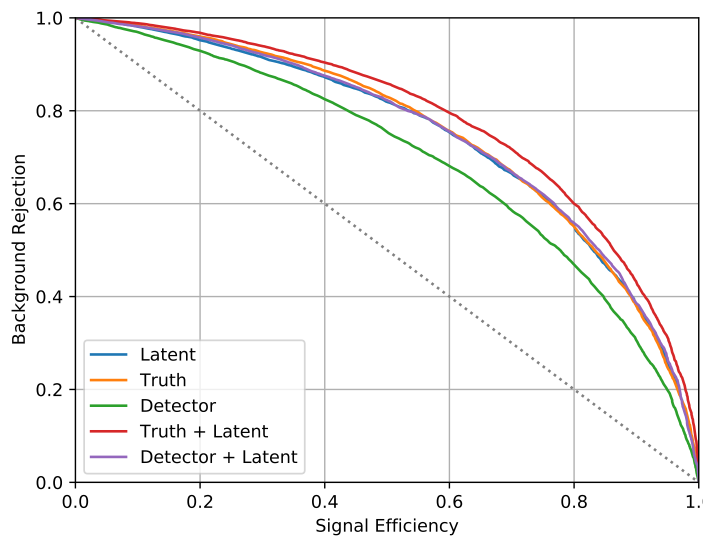
Outline

1. Vector boson scattering in the SM
   - The W^+W^-jj signal and its backgrounds
2. The neural network and shape analysis
3. Resolve the features of polarization for W^+W^-jj channel
4. Applications to new physics searches
   - The effective field theory
   - Two Higgs Doublet Model

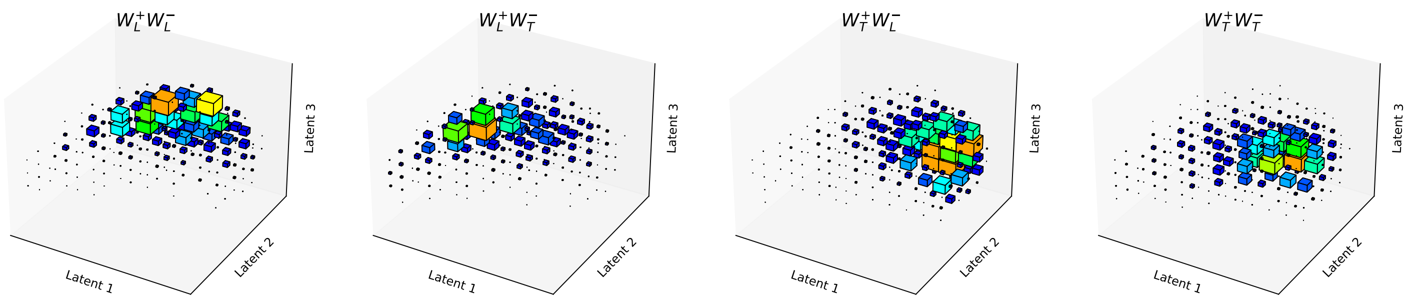


Comparison study (GBDT method)

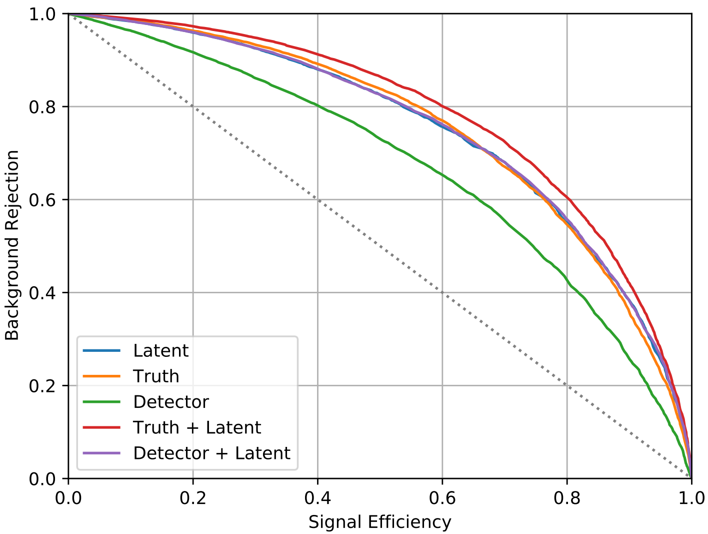
- **Detector level:** the transverse momenta of two leptons $p_T(\ell_{1,2})$ and the forward-backward jets $p_T(j_{1,2})$; the azimuthal angle difference between the forward and backward jets $\Delta\phi(j,j)$.
- **Truth level:** the transverse momenta of two W bosons $p_T(W^{\pm})$; the lepton angle in the W boson rest frame $\cos(\theta_{l^{\pm}})$.



Features of the W boson polarization (Semileptonic channel)



- **Detector level:** the lepton $p_T(\ell)$ and $\eta(\ell)$; azimuthal angle difference between forward backward jets $\Delta\phi(j,j)$; the transverse momentum of W boson pair $p_T(W,W)$
- **Truth level:** transverse momenta of two W bosons $p_T(W^\pm)$, the lepton angle in the W boson rest frame $\cos(\ell)$ and the invariant mass of the forward backward jets m_{jj}



Reconstructable W_ℓ,
but ambiguity in assigning jets

Outline

SM + Effective operator

$$\mathcal{O}_H = \frac{\bar{c}_H}{2v^2}\partial^\mu[\Phi^\dagger\Phi]\partial_\mu[\Phi^\dagger\Phi] \Rightarrow \frac{\bar{c}_H}{2}\partial^\mu h\partial_\mu h$$

Leads the changes to SM couplings:

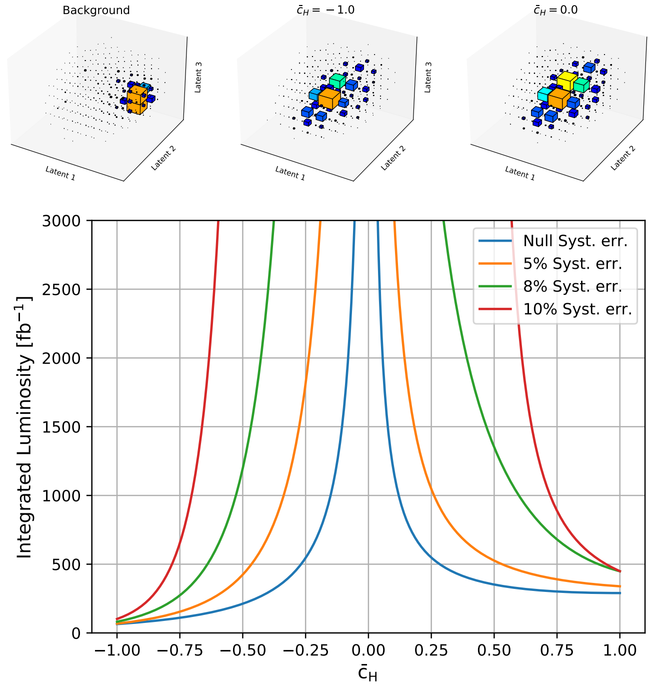
$$\mathcal{L}_H \supset \frac{g m_W}{c_W^2}[1-\frac{1}{2}\bar{c}_H]Z_\mu Z^\mu h + g m_W[1-\frac{1}{2}\bar{c}_H]W_\mu^\dagger W^\mu h + [\frac{y_f}{\sqrt{2}}[1-\frac{1}{2}\bar{c}_H]\bar{f}P_R f h + h.c.]$$

$\bar{c}_H$	-1.0	-0.5	0	0.5	1.0
$\sigma^0_{m_{jj}>500}$ [fb]	440.6	421.8	419.7	426.7	436.2
σ_{ll} [fb]	4.82	4.44	4.36	4.48	4.62
σ_{lj} [fb]	40.2	37.7	37.3	37.9	39.3
$\sigma^{LL}_{m_{jj}>500}$ [fb]	46.29	29.68	25.84	28.79	34.01
σ^{LL}_{ll} [fb]	0.754	0.397	0.314	0.356	0.462
σ^{LL}_{lj} [fb]	5.28	3.04	2.40	2.79	3.50

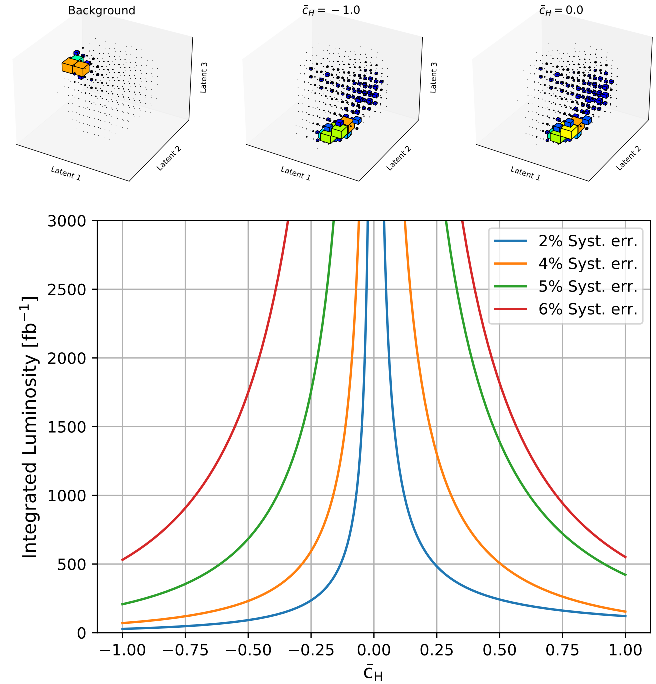
- Larger $|\bar{c}_H|$ lead to increased LL component
- Preselection increase the LL fraction
- Semileptonic channel has larger cross section

SM + Effective field theory operator

Dileptonic channel:



Semileptonic channel:



- weighted sum of backgrounds
- The required integrated luminosity to achieve 95% C.L. probing
- semi-leptonic channel outperforms the dileptonic channel if Syst. err. $\lesssim 5\%$

Two Higgs Doublet Model

Parameter choices:

$m_{H_1} = 125$ GeV, $\tan\beta = 5$, m_A and $m_{H^\pm}$ are not relevant;

varying m_{H_2} and $\sin(\alpha - \beta)$

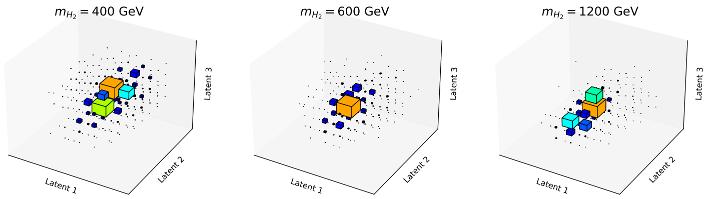
The H_2 decay width can be calculated:

$$\Gamma(H_2) = \Gamma(H_2 \to WW) + \Gamma(H_2 \to ZZ) + \Gamma(H_2 \to \bar{t}t) + \Gamma(H_2 \to \bar{b}b)$$

$(m_{h_2}, \sin(\beta-\alpha))$	(300,0.7)	(300,0.9)	(700, 0.7)	(700,0.9)
$\sigma^0_{m_{jj}>500}$ [fb]	636.2	492.5	461.9	428.5
σ_{ll} [fb]	8.362	5.853	5.527	4.842
σ_{lj} [fb]	64.07	46.52	43.70	39.33
$\sigma^{LL}_{m_{jj}>500}$ [fb]	170.75	79.81	71.58	42.65
σ^{LL}_{ll} [fb]	2.91	1.27	1.30	0.676
σ^{LL}_{lj} [fb]	20.78	9.35	9.50	5.06

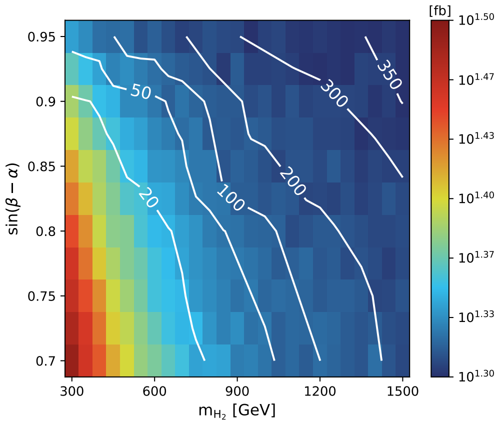
The cross section in 2HDM can be smaller than that in SM when the mass of the H_2 is heavy and decay width of the H_2 is large, because of the destructive interference between H_1 and H_2 in some phase space.

- The LL fraction is considerably larger than that of the SM
- The preselections can increase the fraction even further
- the resonance in the latent space

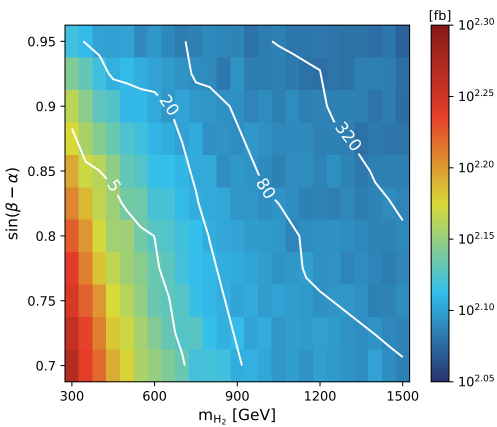


Two Higgs Doublet Model

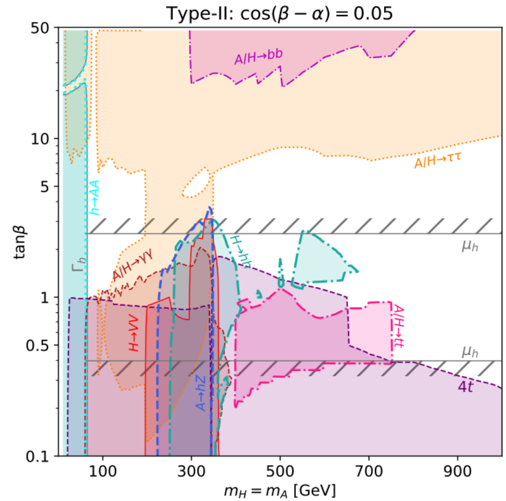
Dileptonic channel:



Semileptonic channel



- The color grades correspond to the fiducial cross sections ($m_{jj} > 500$ GeV)
- The systematic uncertainties are set to 5% for both channel
- the sensitivity of the method is roughly determined by the cross section, even though a slightly better sensitivity can be achieved in the small $\sin(\beta - \alpha)$ region
- model specific search may be more sensitive in some parameter space



F. Kling *et.al.* (2020)

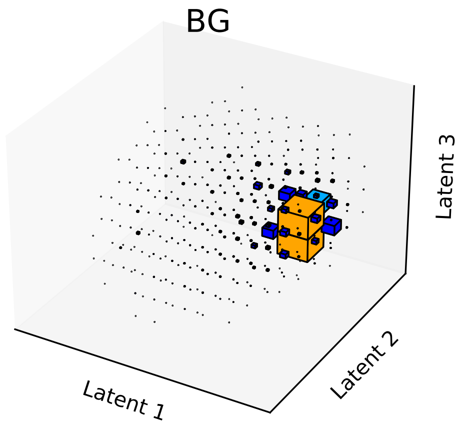
Conclusion

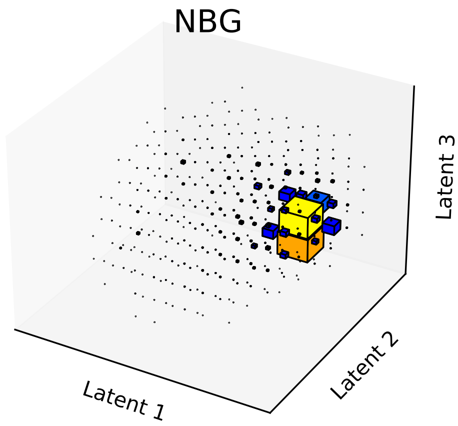
- The W^+W^-jj process and its polarization serve as a important window to test the accuracy of the SM
- A network based on transformer architecture is proposed to detect the features of this channel (discriminate their polarizations and discriminate the signal from backgrounds)
  - The network takes reconstructed four-momenta of final state leptons and jets as inputs.
  - classifying different polarization modes
  - detect the anomalous contributions from the EFT operator $\mathcal{O}_H = \partial^\mu[\Phi^\dagger\Phi]\partial_\mu[\Phi^\dagger\Phi]$ and 2HDM
- Model-specific approaches might have better sensitivity, the advantage of the method is its generality.

Future prospects:

- A process ignorsic method of polarization fraction extraction for the dileptonic channel?

Backup (Simulation dependence)





Two independent smearing (Delphes-ATLAS):

- **Electron:** $\sigma(\phi,\eta) = 0.002$, $\sigma(E) = \sqrt{0.03^2 + E^2 \times 0.0013^2}$
- **Muon:** $\sigma(\phi,\eta) = 0.001$, $\sigma(E) = \sqrt{0.01^2 + E^2 \times 0.0001^2}$
- **Tracker:** $\sigma(\phi,\eta) = 0.002$, $\sigma(E) = \sqrt{0.06^2 + E^2 \times 0.0013^2}$
- **ECal:** $\sigma(\phi,\eta) = 0.02$, $\sigma(E) = \sqrt{E \times 0.101^2 + E^2 \times 0.0017^2}$
- **HCal:** $\sigma(\phi,\eta) = 0.1$, $\sigma(E) = \sqrt{1.59^2 + E \times 0.5205^2 + E^2 \times 0.0302^2}$

Each of the maximal bins takes $\sim 10\%$ of total background events.

Backup (Discrimination power of the score(s))