DNN Regression: Neural Network Application in Higgs Boson CP State Measurement

Prepared by: Qintong Li
Supervisor: David Colling
Assessor: Daniel Winterbottom

Introduction

- Discovery of Higgs boson, 2012 [1].
- Measurement of Higgs boson CP states:
  - As was stated in 2009 by Berge, one way to achieve that is through its decay to τ leptons [2].
  - Main limitation in previous measurements:
    - Detector precision
    - Detection of neutrinos
    - Limited application for sub-leading τ decay channels

Introduction

- History application of ML techniques
  - When the analysis of decay modes extended from 1-prong to 3-prong in 2016, ML techniques was implemented to improve the the available fraction of $H \rightarrow \tau\tau$ decay rate from 6.5 % to 11.9 % [3].
  - As it was supported by Barberio in 2017 [4], ML techniques can also be applied to differentiate the pseudo-scalar Higgs boson state from the scalar one.
  - The first successful measurement of Higgs boson CP structure was in 2021. Both multi-class neural networks and multi-class BDT algorithm combined with the XGBoost package were used to optimize the performance of categorization [5]

Methods

- Separation of Higgs signal from Drell-Yan background signal.
- Methods of φ_{CP} reconstruction
- Important variables
- Multi-output DNN Regressor

Separation of Higgs boson signal from Drell-Yan background signal

- The correlation between the τ-polarisation can be a distinctive way to separate the H → τ τ signal from the Drell-Yan background [6].
- The τ-polarisation can be presented as energy fractions R_i for different decay mode:

$$R_{i,gen} = \frac{E_\pi}{E_\tau} = \frac{E_\pi}{E_\pi + E_\nu} \qquad (1)$$

$$R_{i,gen/reco} = \frac{E_\rho}{E_{\tau_{vis}}} = \frac{E_{\pi^\pm}}{E_{\pi^\pm} + E_{\pi^0}} \qquad (2)$$

Methods of φ_{CP} reconstruction

- Impact parameter method
  The transverse impact parameter here is defined as the minimal distance between the primary vertex and the point on the track [7].
- Neutral-pion method
  For the ρ meson decays, the reconstructed four-momentum vector in laboratory frame is replaced by the four-momentum vector of the π^0 [8].
- Polarimetric vector method

Important variables

- Polarimetric vector
- Acoplanarity angle φ_{CP}

Polarimetric vector

- This method can be applied to all decay channels. Only two decay modes of the τ leptons are considered in this case: $\tau^\pm \rightarrow \pi^\pm \nu$ and $\tau^\pm \rightarrow \rho^\pm \nu$, followed by $\rho^\pm \rightarrow \pi^\pm\pi^0$.

- For π decay mode, it is defined as the four-momenta vector of the single charged pion boosted into the τ frames as a three-vector and then boosted back to the laboratory frame.

- Polarimetric vector for decay mode ρ [9]:

$$h^i = \mathcal{N}(2(q \cdot N)q^i - q^2 N^i) \qquad (3)$$

$$q \cdot N = (E_{\pi^\pm} - E_{\pi^0})m_\tau \qquad (4)$$

, where *N* is a normalization function, q is the difference of the $\pi^\pm$ and π^0 four-momenta and N is the four-momentum of the τ neutrino (τ frame)

Acoplanarity angle φ_{CP}

- Vector $h_{1,2}$ is the polarimetric vector (it can also be the vector computed by the other methods)[5]

$$\vec{k}_{1,2} = \frac{\vec{h}_{1,2} \times \vec{h}_{1,2}}{|\vec{h}_{1,2} \times \vec{h}_{1,2}|} \quad (5)$$

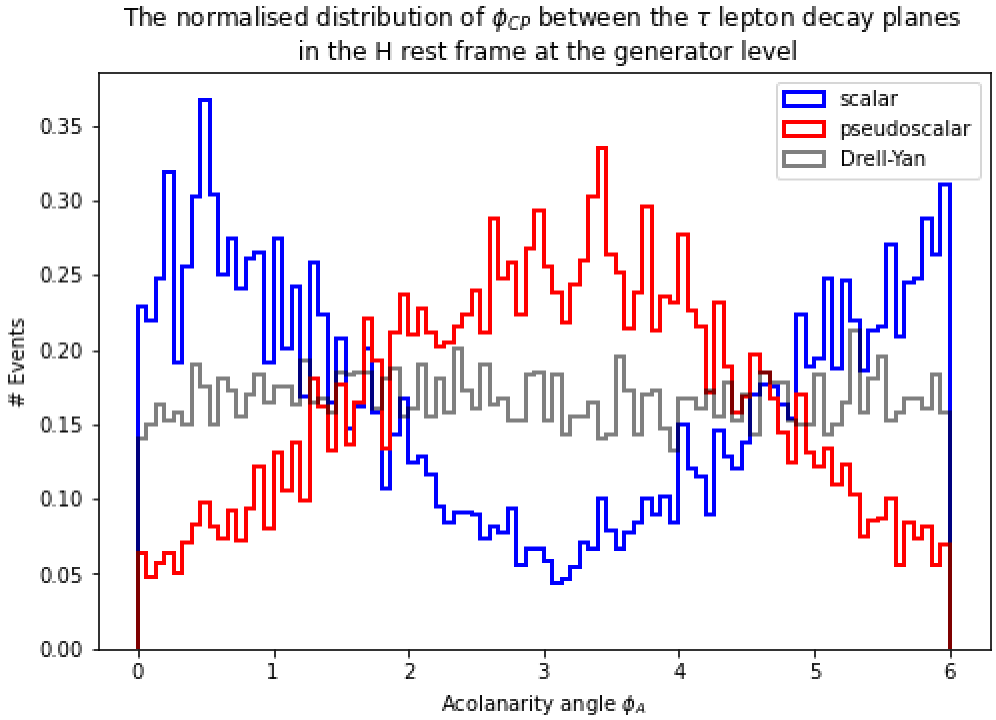
$$\phi^* = arccos(\vec{k}_1 \cdot \vec{k}_2) \quad (6)$$

$$O^* = -(\vec{h}_1 \times \vec{h}_2) \cdot \vec{h}_1 \quad (7)$$

$$\phi_{CP} = \begin{cases} \phi^* & ifO^* \geq 0 \\ 2\pi - \phi^* & ifO^* \leq 0 \end{cases} \quad (8)$$

Reconstruction of φ_{CP}

- Φ_{CP} can be used to differentiate the pseudo-scalar and scalar Higgs boson CP state.



DNN Regressor

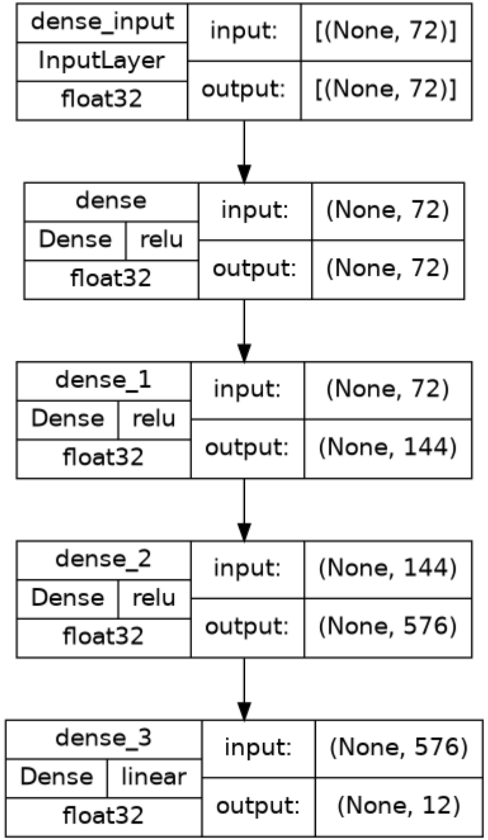
- Pre-proccessing of data
- Inputs and targets
- Loss function
- Built-in structure

Pre-proccessing of data

- Pre-selection:
  - Hadronic decays
  - The variables ‘dm_1’ and ‘dm_2’ used to pre-select the wanted events, are representing the decay mode for leading τ and sub-leading one.

dm_i number	Decay mode
0	π
1	ρ
10	a1

The structure of multi-output DNN regressor



Inputs and targets

Inputs (72 in total):

1. Single charged π and neutral π (four-momenta) kinematics decays from each τ .

2. Missing energy met in x and y direction(Cartesian coordinate).

3. Impact parameters for both taus in Cartesian coordinate.

4. Secondary vertex information.

5. Reconstructed tau kinematics.

Inputs and targets

The 6 output targets are also listed here for each τ, 12 in total:

1. **Neutrino four-momenta** in Cartesian coordinate(3 variables for each tau). Energy is the magnitude of the 3-vector.

2. **$\Delta R = \Delta\varphi^2 + \Delta\eta^2$**, where φ is the spherical polar angle (from 0 to π, inclusive) and η is the rapidity which is also a kind of spherical polar angle [34]. This quantity is defined to specify the difference between the neutrino vector and the visible decay product vectors.

3. **Scalar product** between the visible decay product and neutrino for each τ, which can also be treated as the projection of neutrino vector on visible decay product direction.

4. **Cross product magnitude** between the visible decay product and neutrino for each tau can also be treated as the distance between the neutrino and visible product vectors.

After the φ_{CP} is computed, the weight for the scalar and pseudo-scalar event included in the

same data set and will be applied for separation.

Loss function

Normally, the loss function is defined as:

$$MAE = \frac{1}{N}\sum_{j=1}^{n}|y_{pred} - y_{true}| \tag{10}$$

Due to the nature of each predicted physics quantity, the impact of each quantity on the final reconstruction is different. Therefore, to penalize each quantity equally or add bias to the preferred target, applying individual weight to each target value is done as followed:

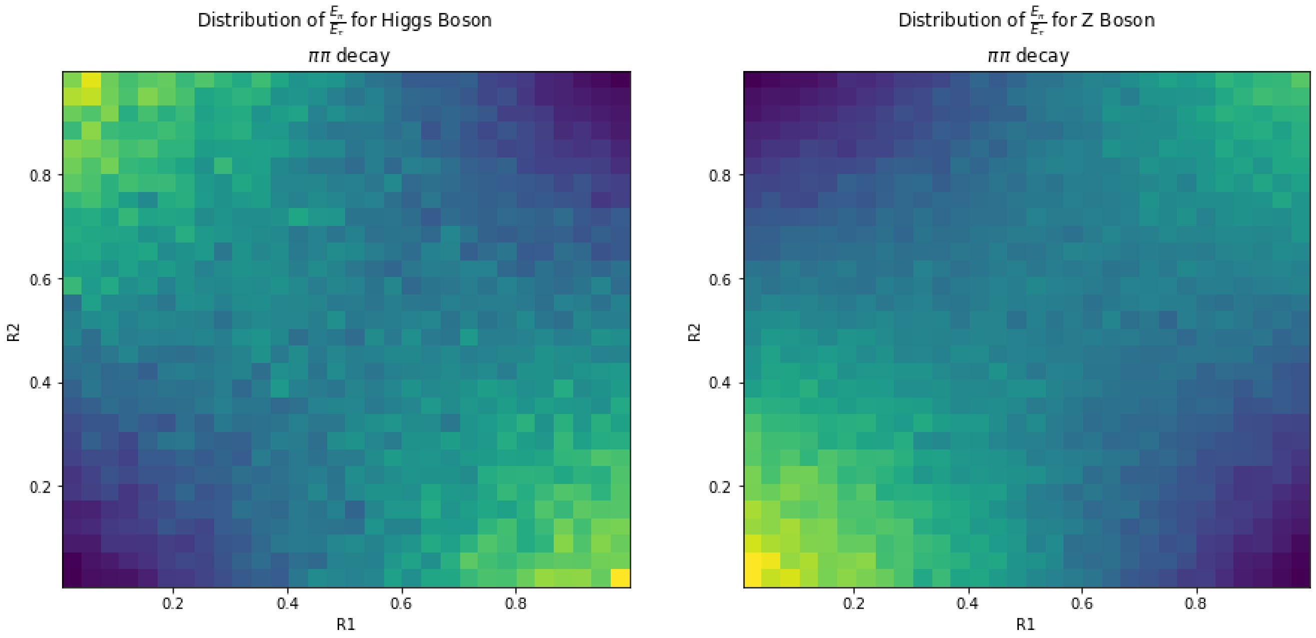
$$MAE_{multi} = \frac{1}{N}\sum_{j=1,i=1}^{n} w_i \cdot |y_{i,pred} - y_{i,true}| \tag{11}$$

Results

- Separation of Higgs signal from Drell-Yan background signal
- Regression of neutrino momentum components
- Evaluation of predictions
- Reconstruction of Acoplanarity angle φ_{CP}
- Comparison between methods

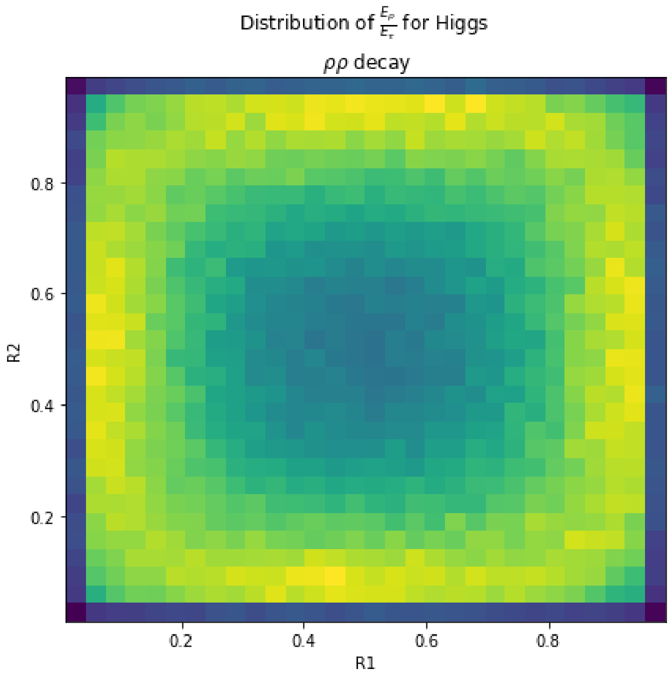
Separation of Higgs signal from Drell-Yan background signal

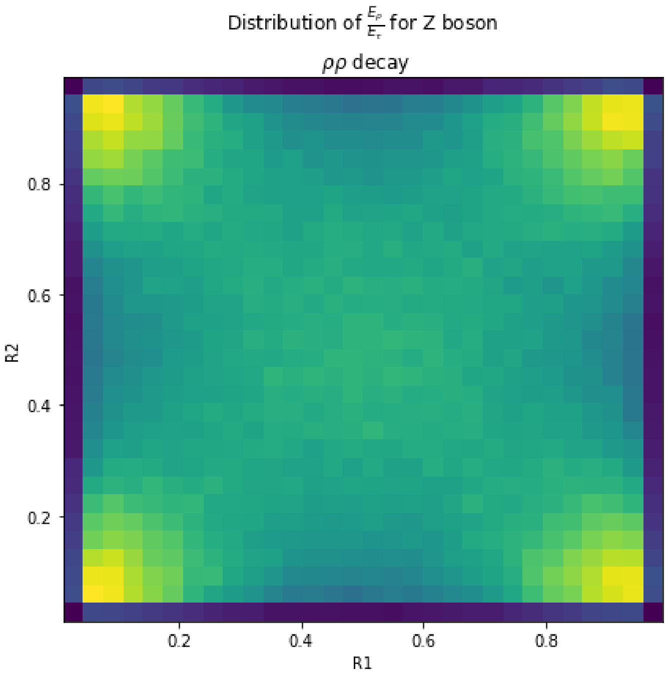
- Energy fraction distribution of π decay for both τ leptons



Separation of Higgs signal from Drell-Yan background signal.

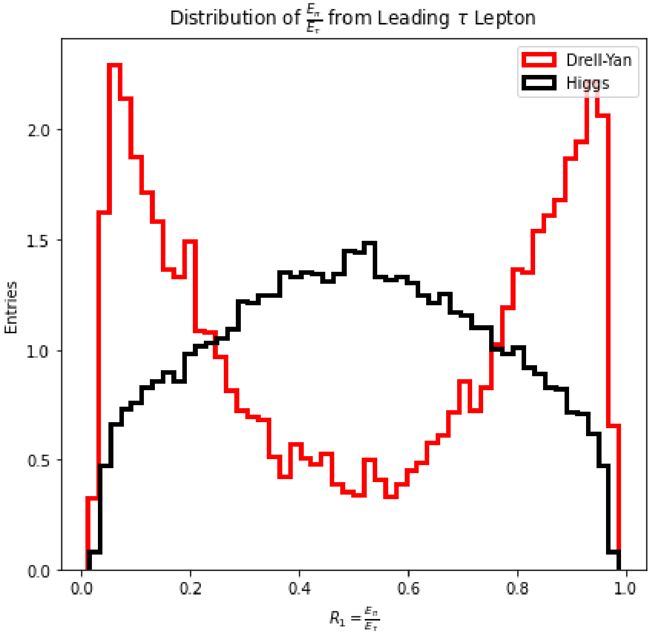
- Energy fraction distribution of ρ decay for both τ leptons

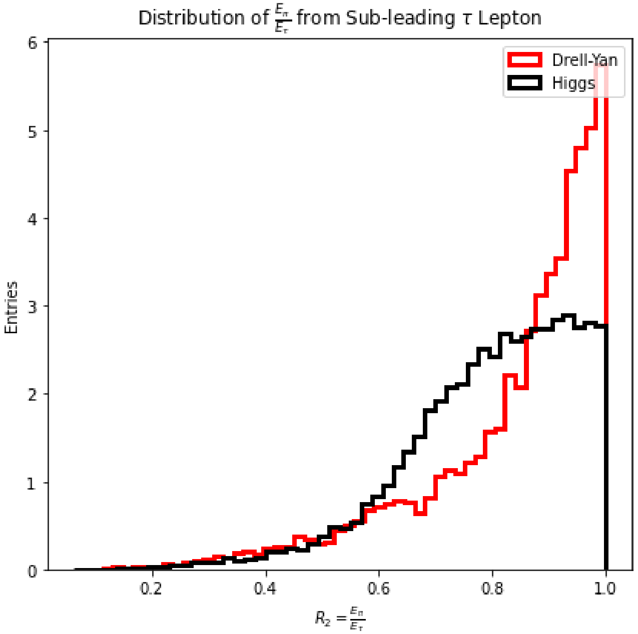




Separation of Higgs signal from Drell-Yan background signal.

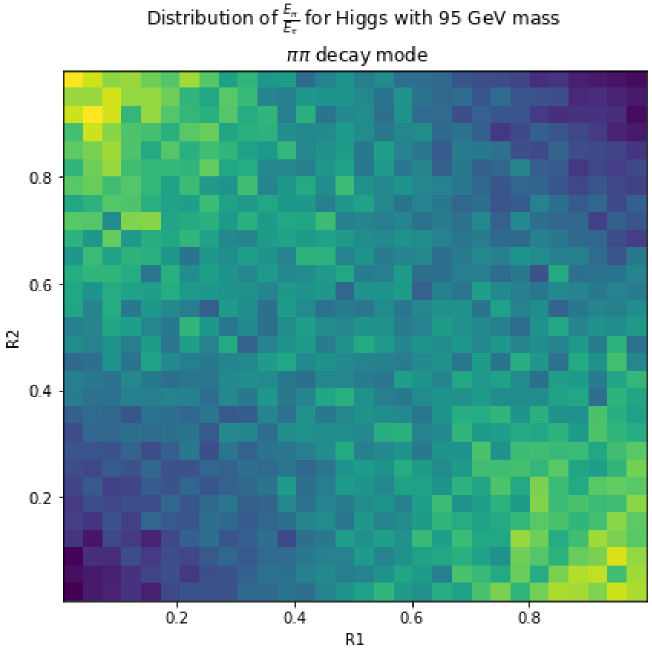
- Energy fraction distribution of ρ decay for leading τ lepton π decay for sub-leading τ

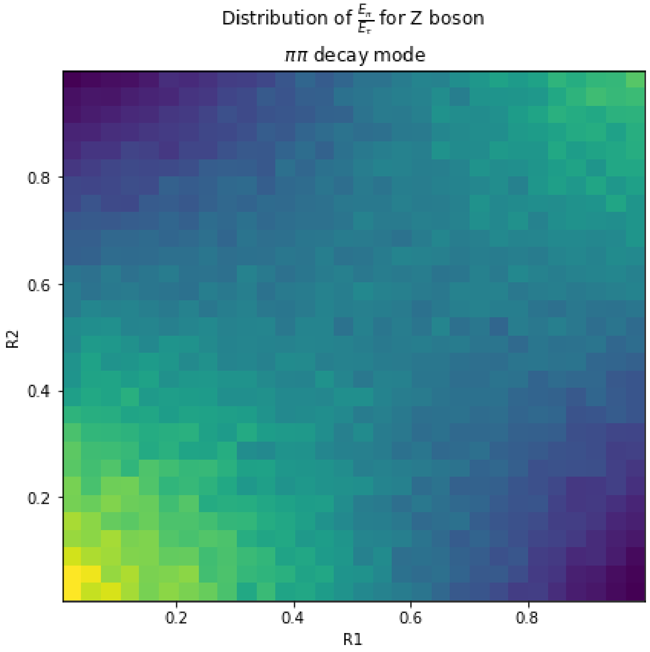




Separation of Higgs (95 GeV) signal from Drell-Yan background signal.

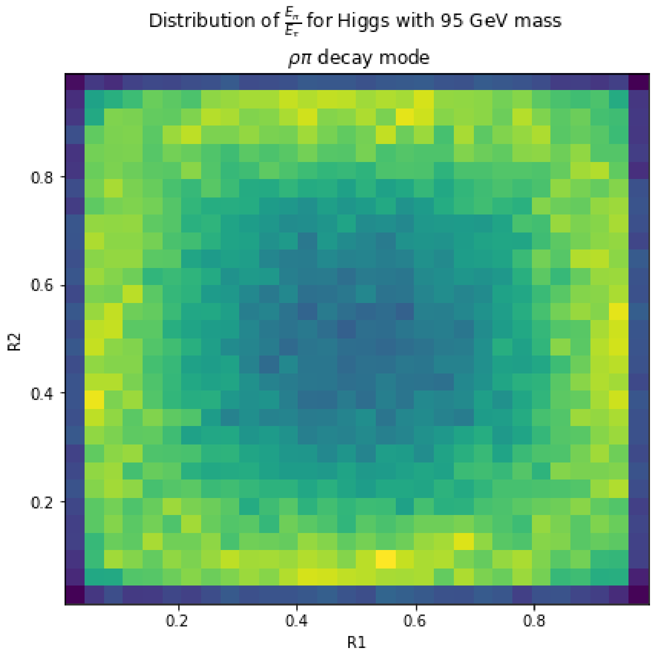
- Energy fraction distribution of π decay for both τ leptons

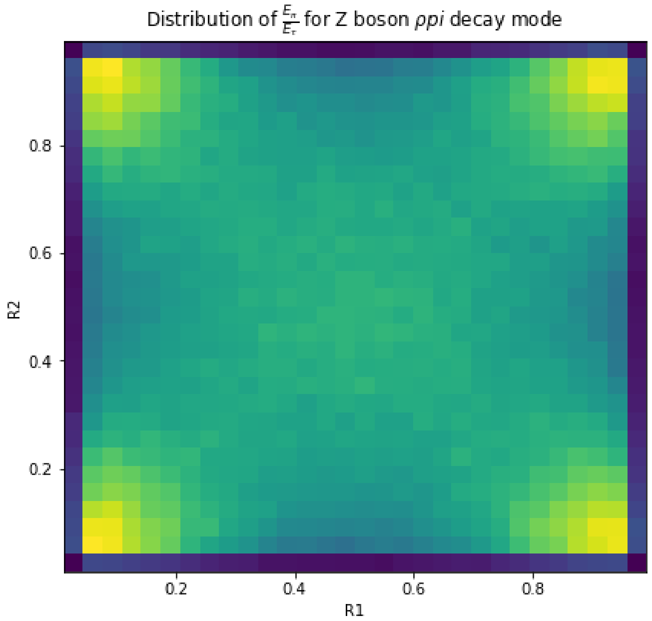




Separation of Higgs signal(95 GeV) from Drell-Yan background signal.

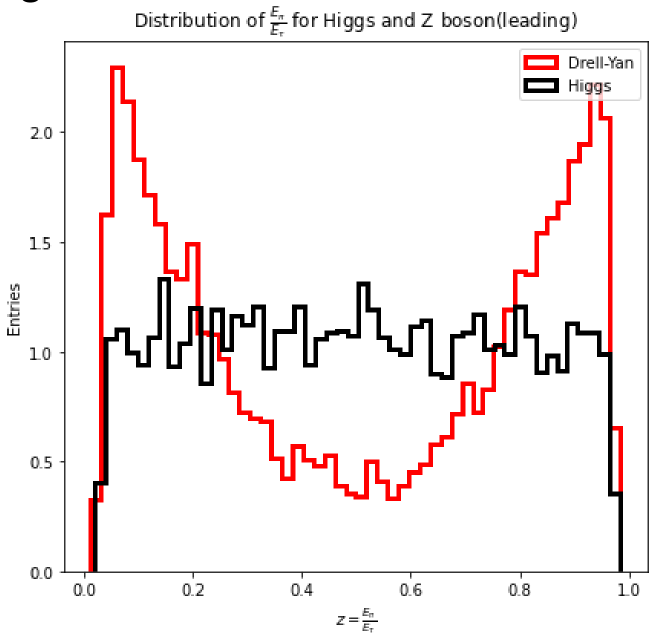
- Energy fraction distribution of ρ decay for both τ leptons

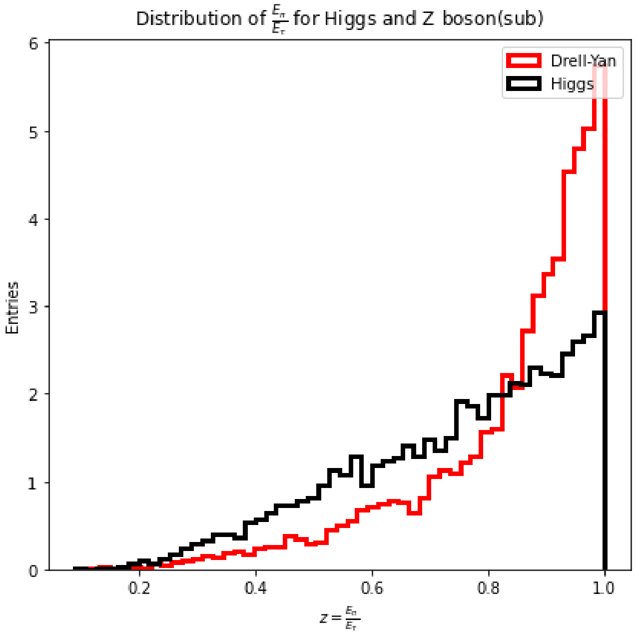




Separation of Higgs(95 GeV) signal from Drell-Yan background signal.

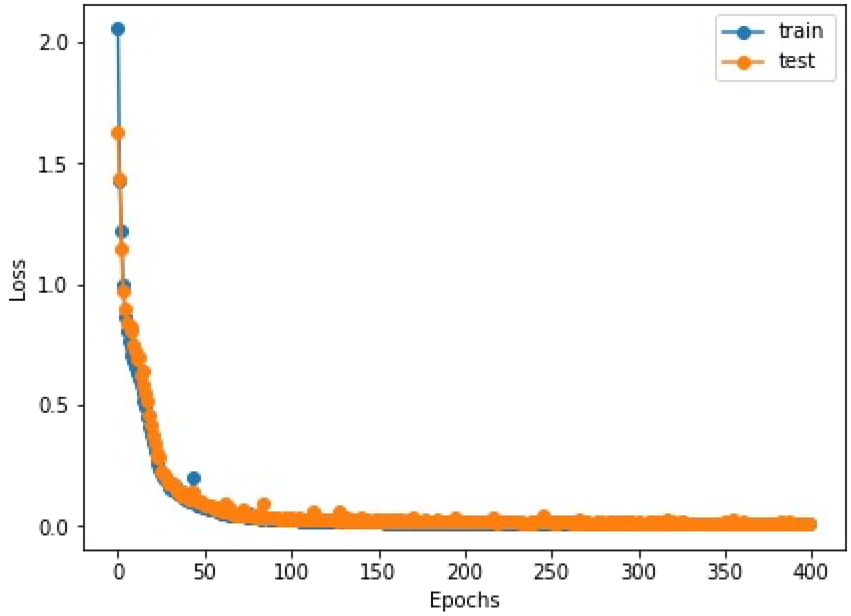
- Energy fraction distribution of ρ decay for leading τ lepton π decay for sub-leading τ

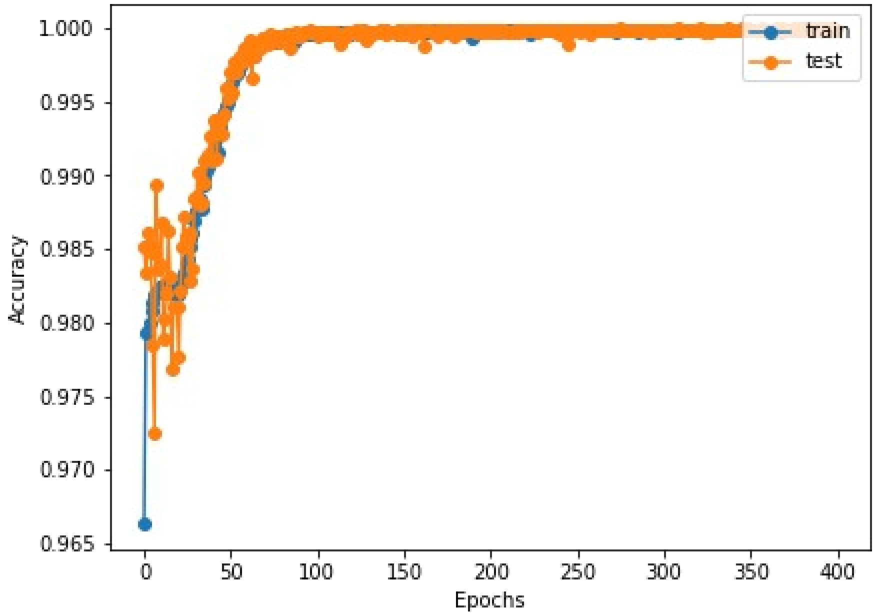




Linear regression for neutrino momentum components

- Learning curves and accuracy curves for training data set and testing data set



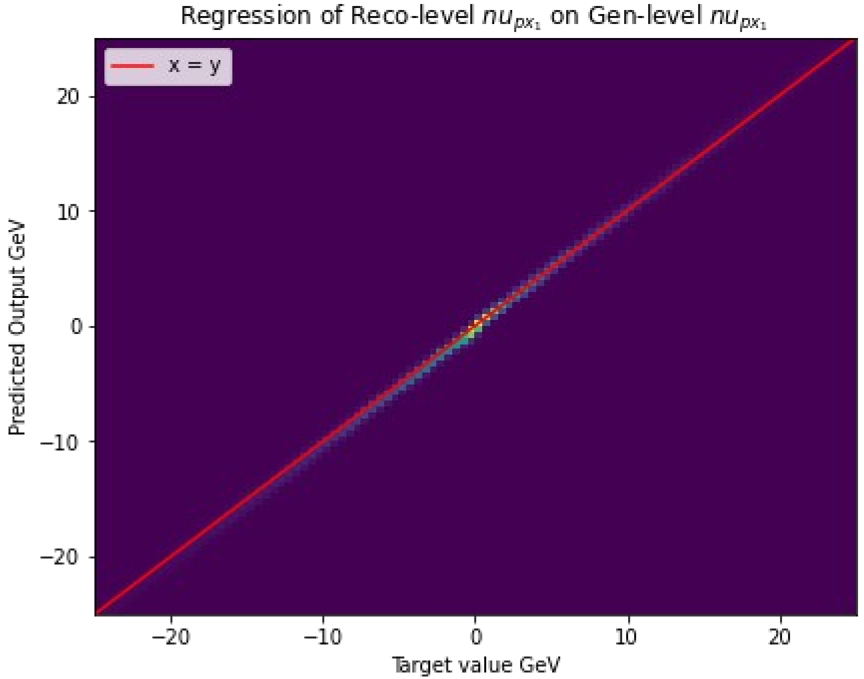


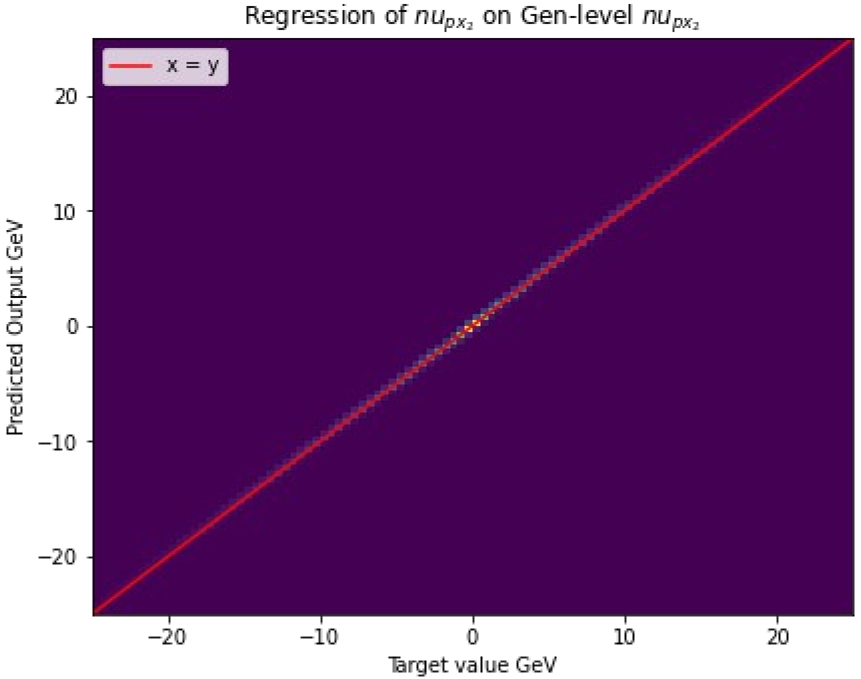
Evaluation of predictions

- 2D distribution of predictions v.s. true valus
- Difference between Montecarlo truth and prediction compared with the case of SV-fit.
- Comparison of other regressors by plotting τ masses using the predictions they made respectively.

Linear regression for neutrino momentum components

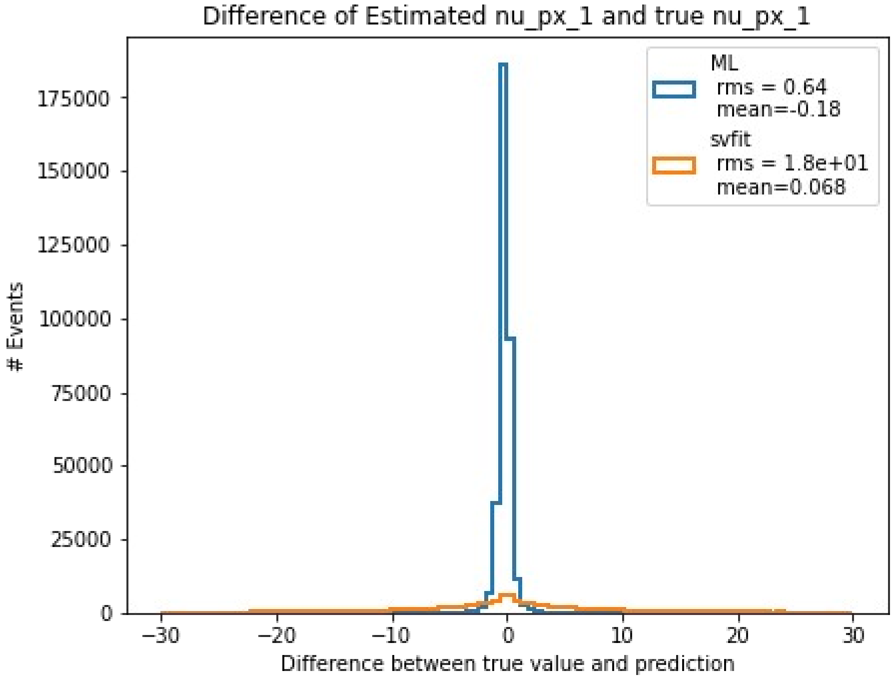
- Regression distribution (only x component)

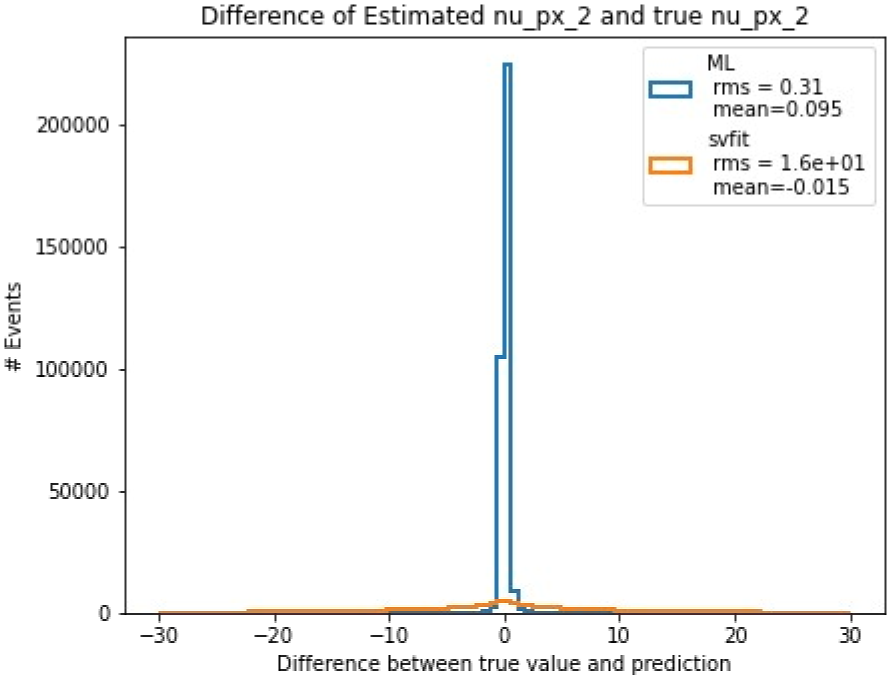




Evaluation of predictions

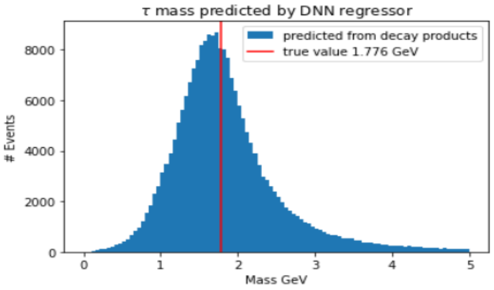
- SV-fit(only for x component)



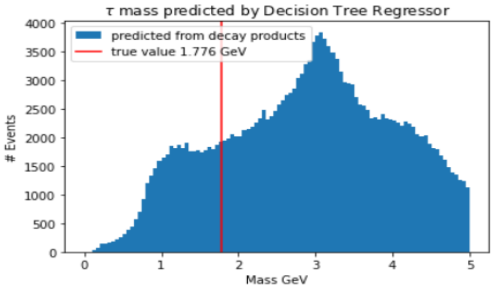


Evaluation of predictions

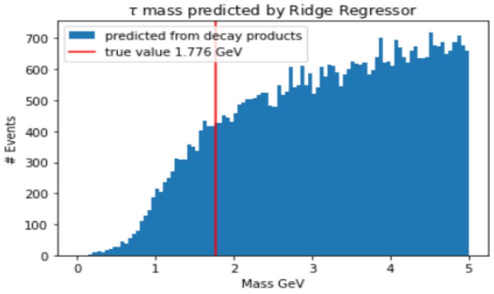
- Comparison between regressors



(a) A plot of τ mass computed from the predictions made by multi-output DNN regressor



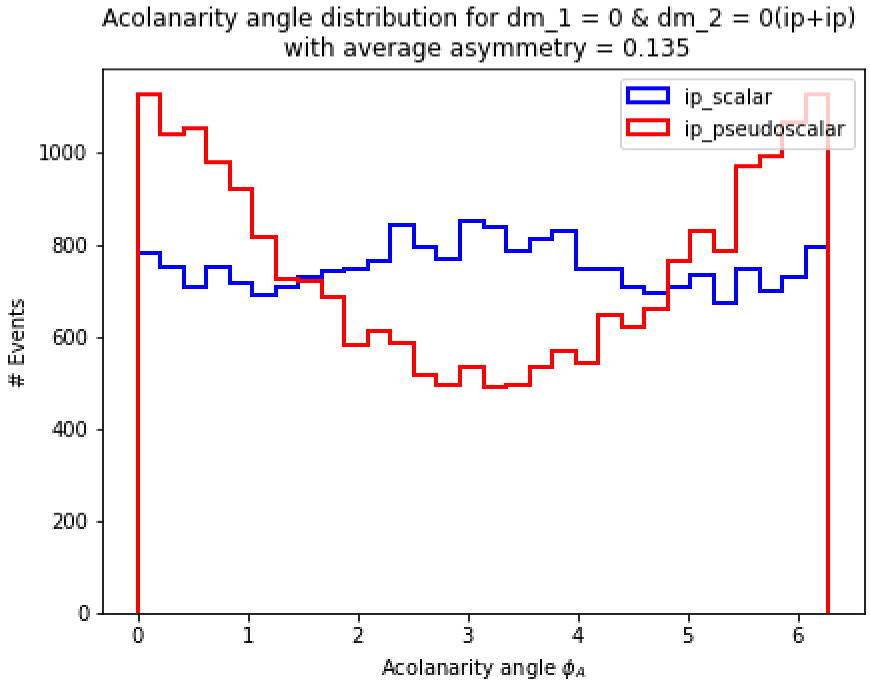
(b) A plot of τ mass computed from the predictions made by Decision Tree regressor

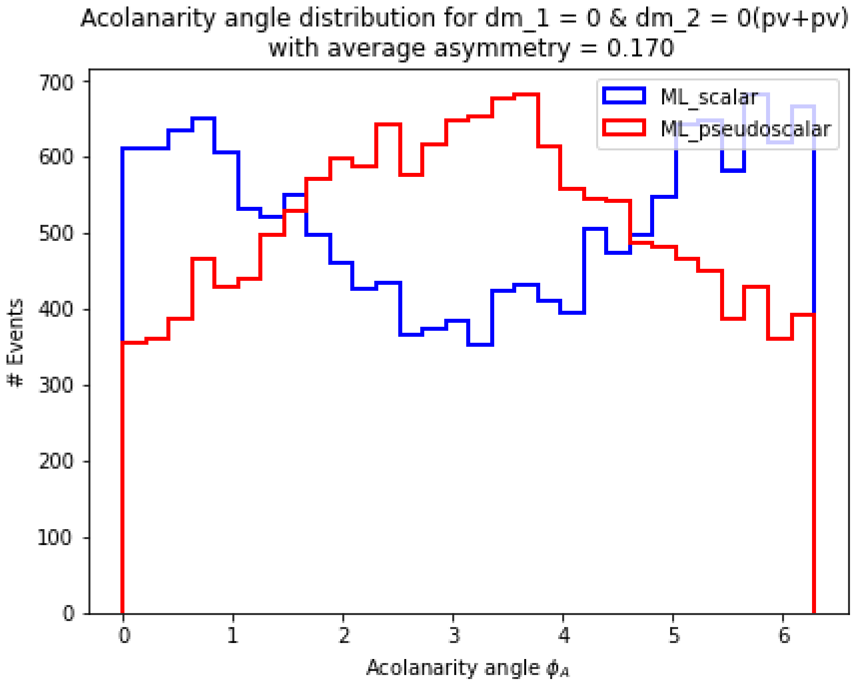


(c) A plot of τ mass computed from the predictions made by Ridge regressor

Reconstruction of φ_{CP}

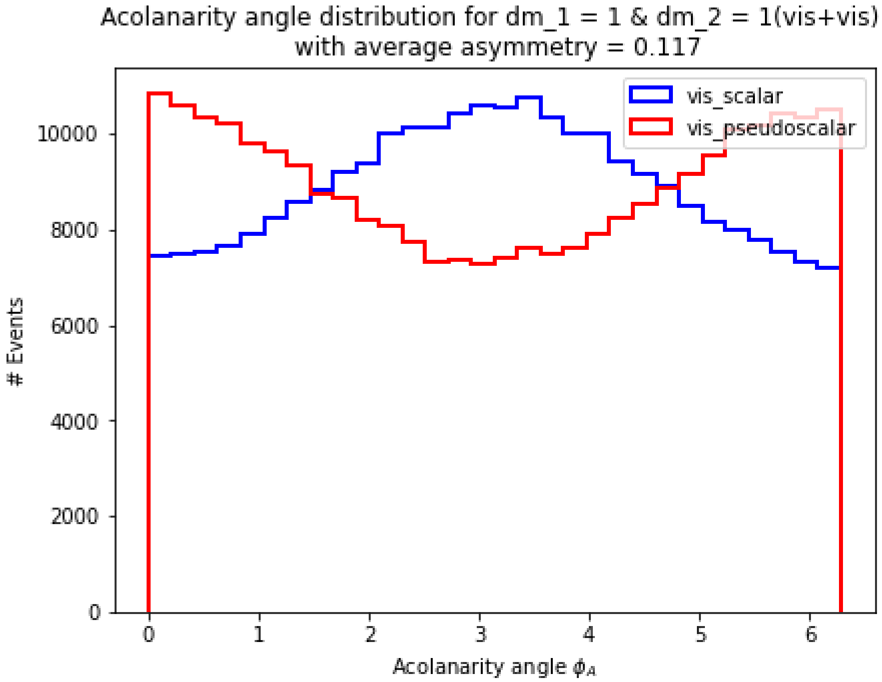
- φ_{CP} distribution when both τ leptons undergo π decays

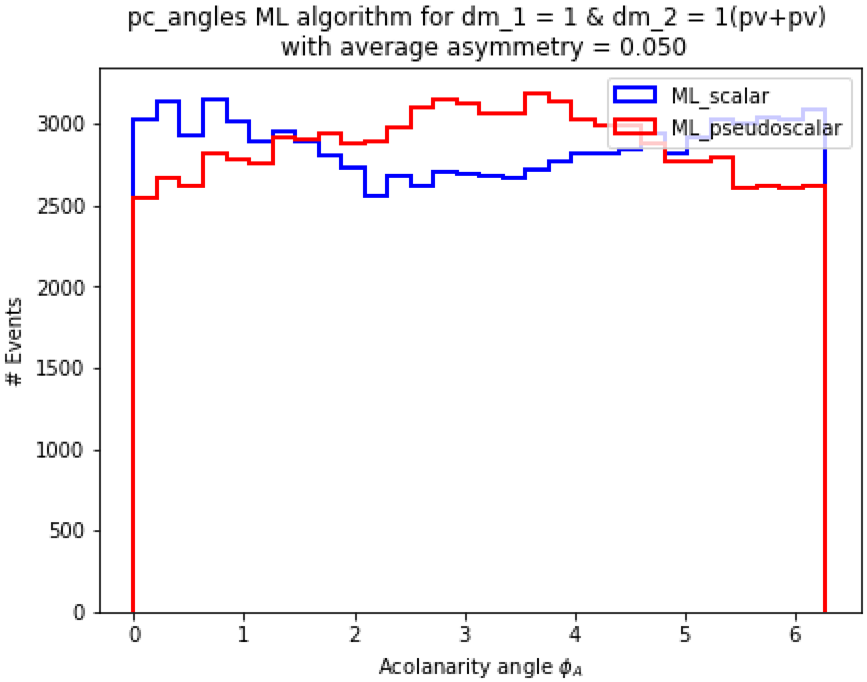




Reconstruction of φ$_{CP}$

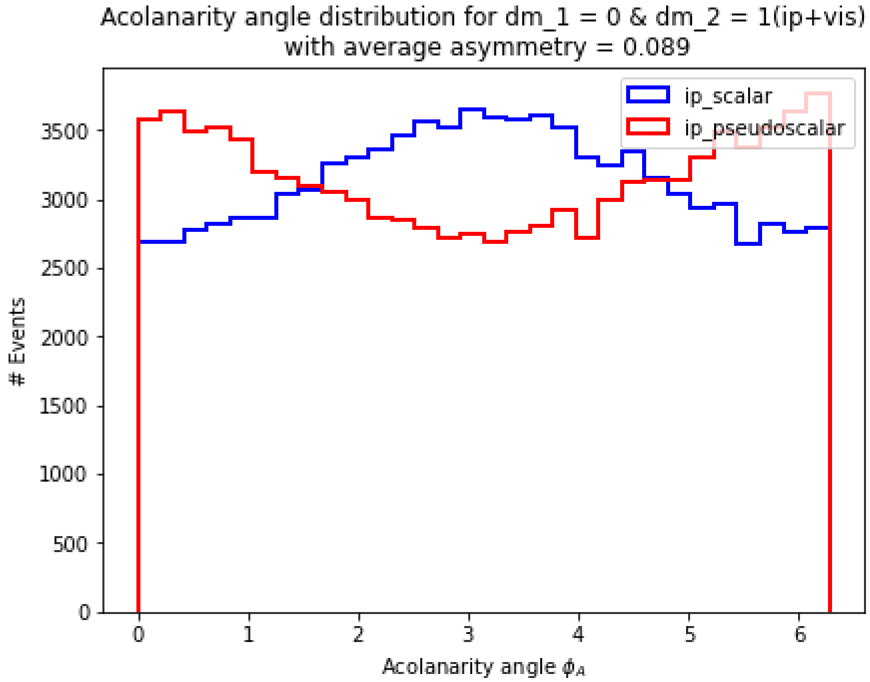
- φ$_{CP}$ distribution when both τ leptons undergo ρ decays

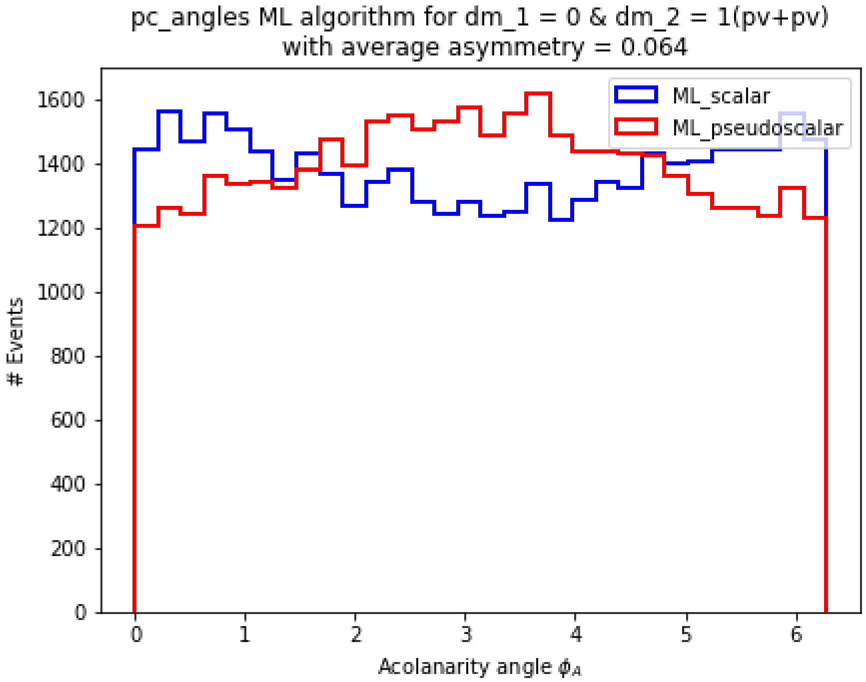




Reconstruction of φ_{CP}

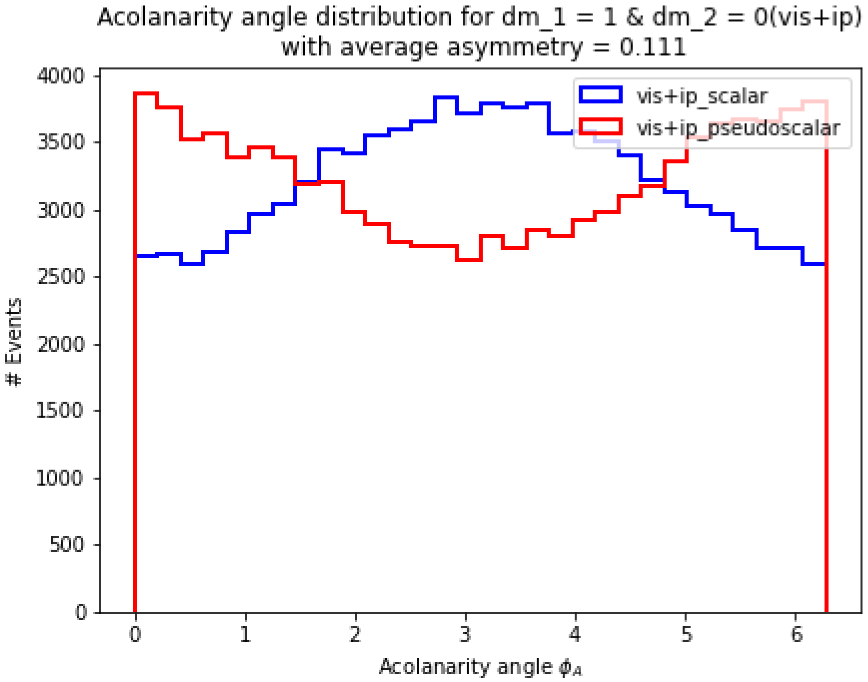
- φ_{CP} distribution when leading τ leptons undergo π decays sub-leading τ leptons undergo ρ decays

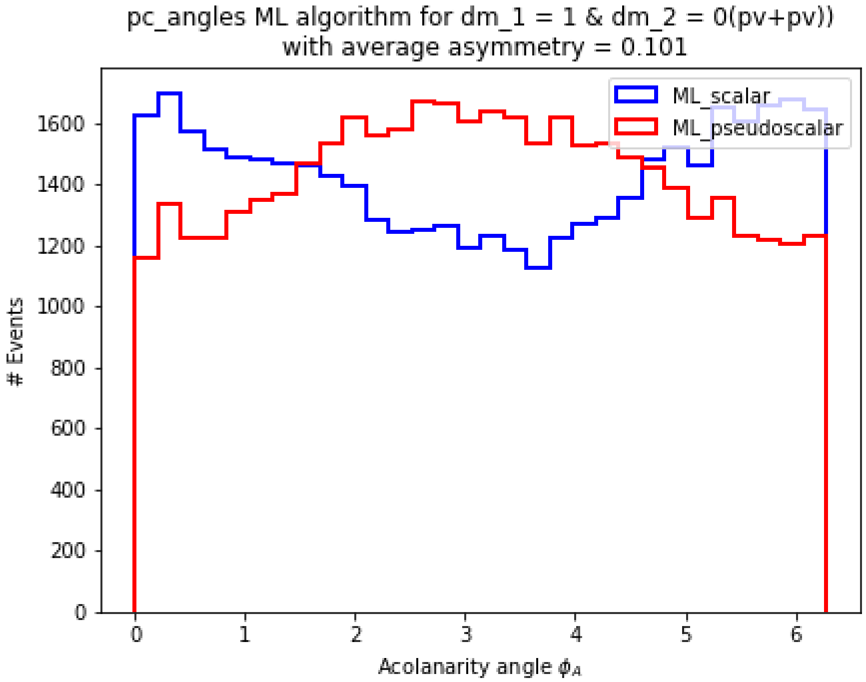




Reconstruction of φ_{CP}

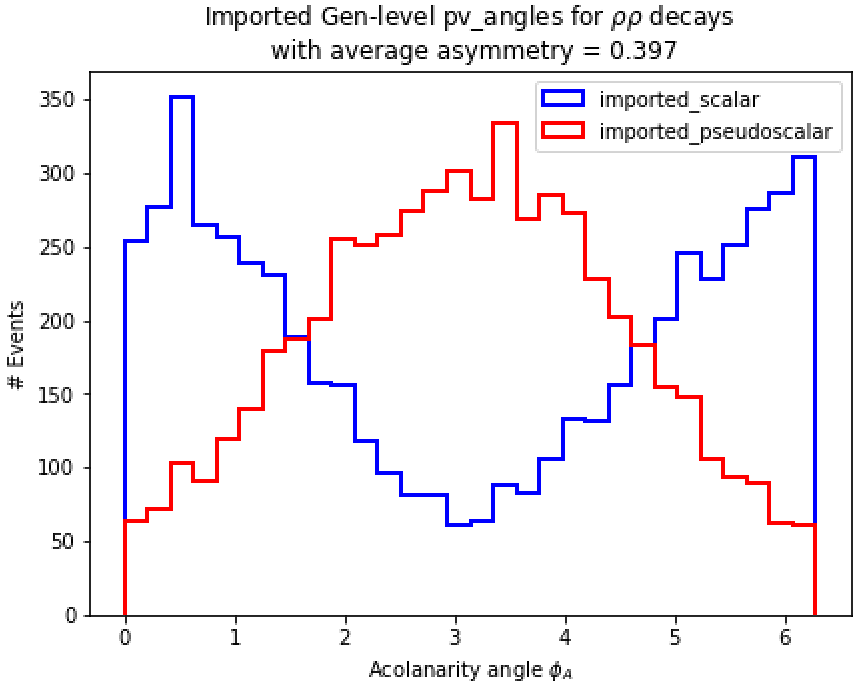
- φ_{CP} distribution when leading τ leptons undergo ρ decays sub-leading τ leptons undergo π decays

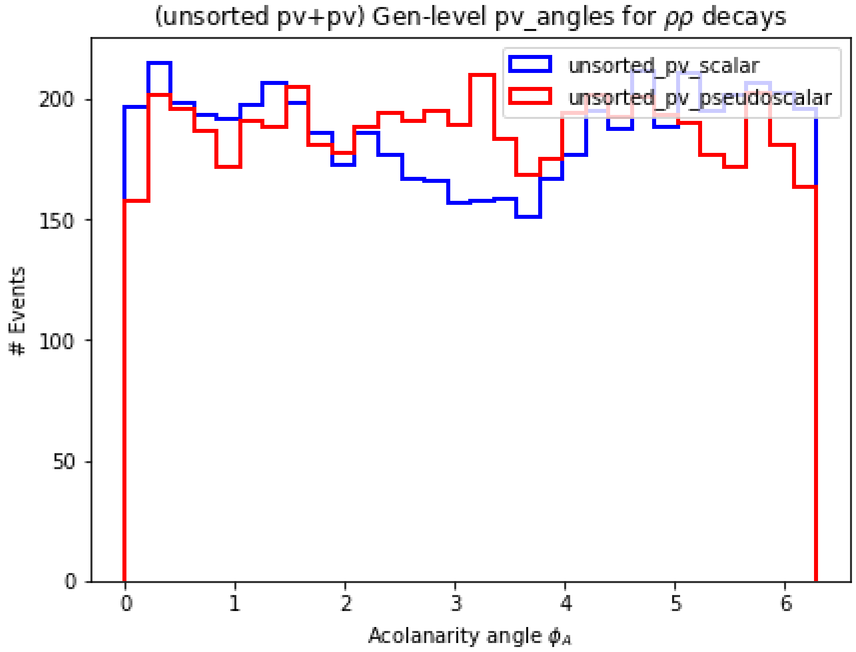




Reconstruction of φ_{CP}

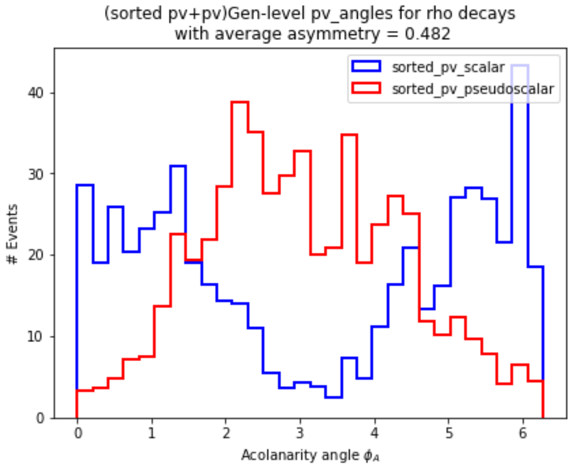
- φ_{CP} distribution when both τ leptons undergo ρ decays at Gen-level

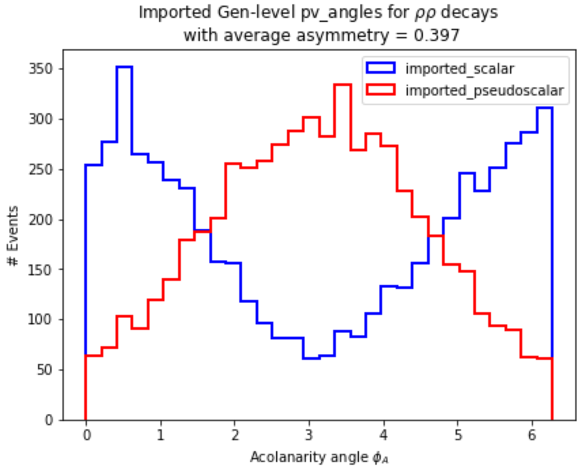


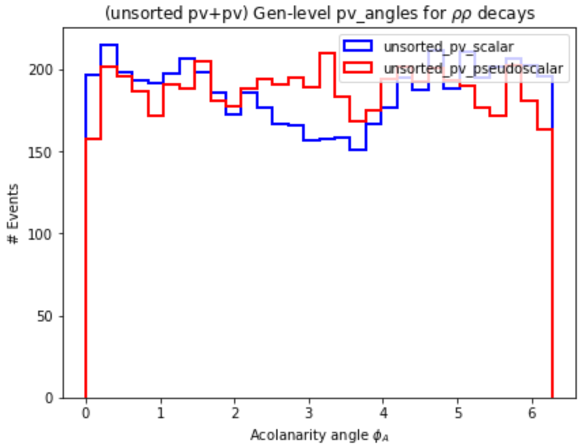


Comparison of Average asymmetry

- ϕ_{CP} distribution when both τ leptons undergo ρ decays at Gen-level







Conclusions

- Separating Higgs signals from Drell-Yan background signals with energy fractions or the $C_{1,2}$ is possible at the generator level.
- By customizing the loss function, the performance of the regressor was improved so that the average asymmetry of the CP_{even}(scalar) and CP_{odd}(pseudo-scalar) distributions for $\pi\pi$ decays improved from the 0.135 using the impact parameter method to 0.170.
- After fixing the problem with the order of events, this method should also apply to the other decay modes

Timeline

Jan
Literature review for key literature

Feb-Mar
- Set up database
- reconstruction with generator-level information

Apr
Separate Higgs signal from background using *energy fraction*

May
2d distribution of energy fraction for different decay Channels

Jun
Linear regression (Higgs)

Jul
DNN regression for neutrinos

Aug
- Evaluating and improving DNN regressor
- Report outline

Sept
Finalizing Report

References

[1]. S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, and etc. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. Physics Letters B, 716(1):30–61,2012.

[2]. Stefan Berge and Werner Bernreuther. Determining the cp parity of higgs bosons at the hc in the τ to 1-prong decay channels. Physics Letters B, 671(4):470–476, 2009.

[3]. R. J´ozefowicz, E. Richter-Was, and Z. Was. Potential for optimizing the Higgs boson CP measurement in H →τ τ decays at the LHC including machine learning techniques. ,94(9):093001, November 2016.

[4]. E. Barberio, B. Le, E. Richter-Was, Z. Was, and D. Zanzi. Optimizing higgs boson cp measurement in h → τ τ decay with ml techniques. Acta Physica Polonica B, 48(6):1049, 2017.

[5]. A. Tumasyan, W. Adam, J. W. Andrejkovic, and etc. Analysis of the CP structure of the yukawa coupling between the higgs boson and τ leptons in proton-proton collisions at √s = 13 TeV. Journal of High Energy Physics, 2022(6), 6 2022.

[6]. S Moretti and D.P Roy. The τ -polarisation test for the h → τ τ signal from the drell-yan background signal at the lhc. Physics Letters B, 545(3-4):329–336, 2002.

[7]. V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, and etc. Constraints on the spin-parity and anomalous hvv couplings of the higgs boson in proton collisions at 7 and 8 tev. Phys. Rev. D, 92:012004, 7 2015.

[8]. CMS Collaboration. Track impact parameter resolution for the full pseudo rapidity coverage in the 2017 dataset with the CMS Phase-1 Pixel detector. CERN, 2020.

[9]. Malgorzata Worek. Higgs cp from h/a -gt; tau tau decay. arXiv, 2003.

[10]. Ph.D. Ayyüce Kızrak. Comparison of activation functions for deep neural networks, Jan 2020.

[11]. James Dellinger. Weight initialization in neural networks: A journey from the basics to kaiming, Apr 2019.

Thank you for listening!
Thank you for the time you spent with me!