ATLAS EXPERIMENT

Results and future outlook for the H->Zy analysis on the ATLAS experiment.

Xiang Li

Supervisor： Kun Liu

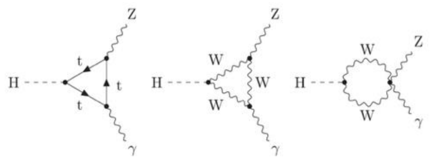
November 2025

李政道研究所
TSUNG-DAO LEE INSTITUTE

Overview

Analysis Objective: $\mathbf{H \to Z\gamma \to \ell\ell\gamma}$

- a potential channel to explore physics **Beyond the Standard Mode**
- rare decay only via loop diagrams in the SM



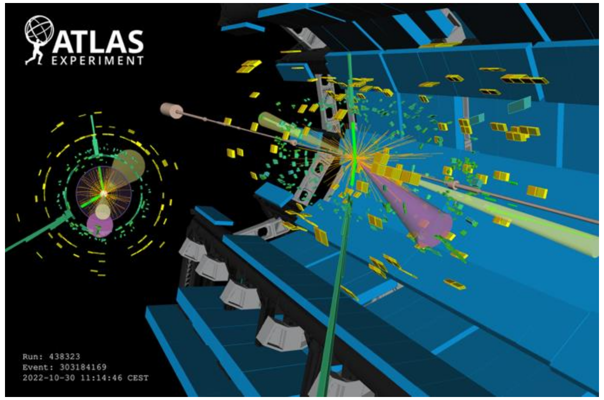
Partial Run 3 Data: **13.6 TeV** and integrated luminosity of **165 fb^{-1}**

Irreducible background

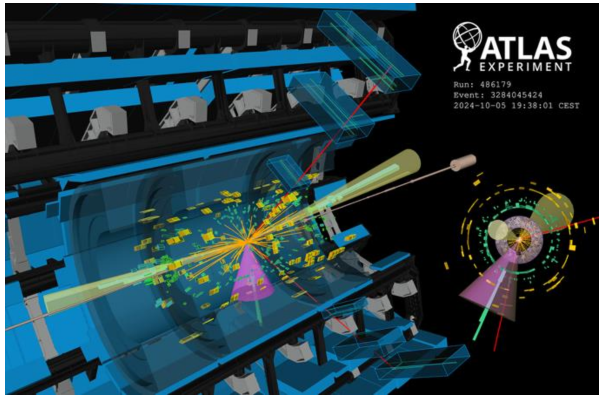
- the non-resonant production of Z bosons in association with a photon.

Reducible background

- inclusive Z boson production in which a hadronic jet is misidentified as a photon

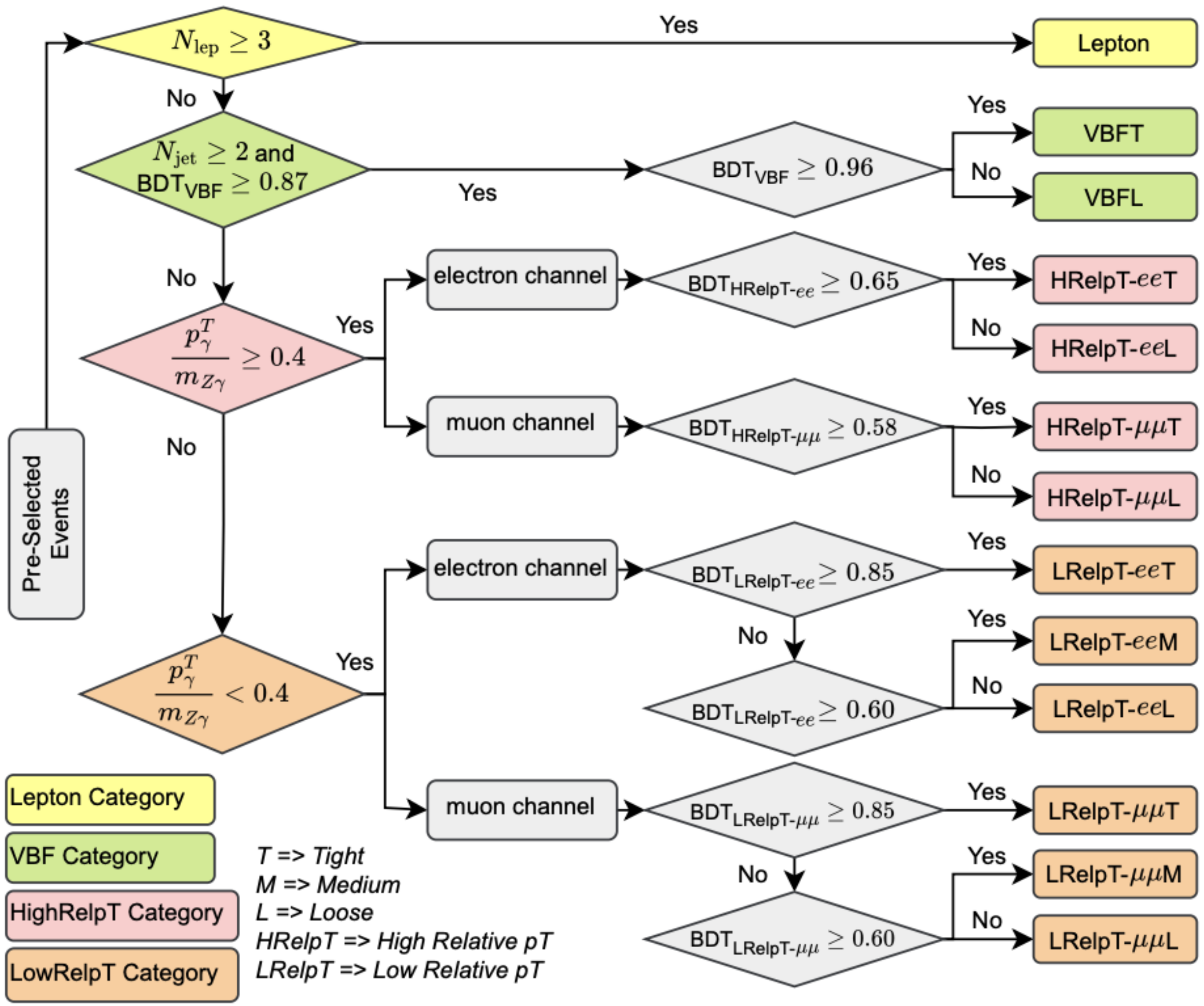


Event display of $H \to Z\gamma \to ee\gamma$



Event display of $H \to Z\gamma \to \mu\mu\gamma$

categories



Events are classified into **13 categories**.

Lepton category: Additional leptons
VBF category : at least 2 jets
High/Low Rel-*pT* category :
Separation between $ee\gamma/\mu\mu\gamma$ final states

BDT classifications for VBF, HRelpT and LRelpT categories.

Split events by their BDT score into **Tight, (Medium) Loose** regions.

Fit range optimization

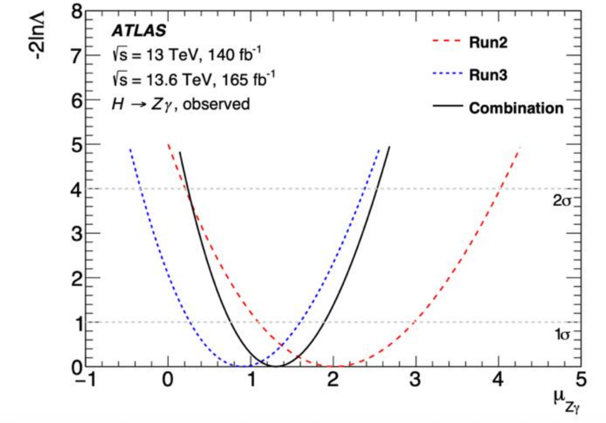
李政道研究所 TSUNG-DAO LEE INSTITUTE

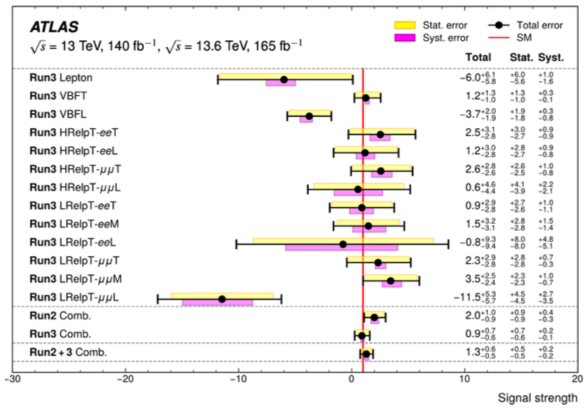
- In each category we build signal modeling and background modeling
- A larger fitting range will result in the best function tending to be of higher order and a larger spurious signal. Therefore, we adjust fit range depend on each category distribution.
- Calculated the results for multiple mass ranges to find the optimal range for the background model.

Mass range	Pvalue for same mu case	significance	DNLL
110-150	0.0326985	1.84253	3.39493
111-150	0.0339938	1.82509	3.33095
113-150	0.032083	1.85102	3.42629
114-150	0.0319835	1.85241	3.43142
115-140	0.0249291	1.96118	3.84622
115-150	0.0284554	1.904	3.6252

Statistic result

Finally statistical results from the un-binned likelihood fit:





- **Run 3 results**:

measured signal strength $\boldsymbol{\mu} = \ 0.9^{+0.7}_{-0.6}\ (\text{stat})\ \ ^{+0.2}_{-0.1}\ (\text{syst})$

Run 3 observed (expected) significance:1.4 $\boldsymbol{\sigma}$ (1.5 $\boldsymbol{\sigma}$).

- **Run 2 + Run 3 combination**:

signal strength $\boldsymbol{\mu}$ = 1.3 ± 0.5(stat) ± 0.2(syst)

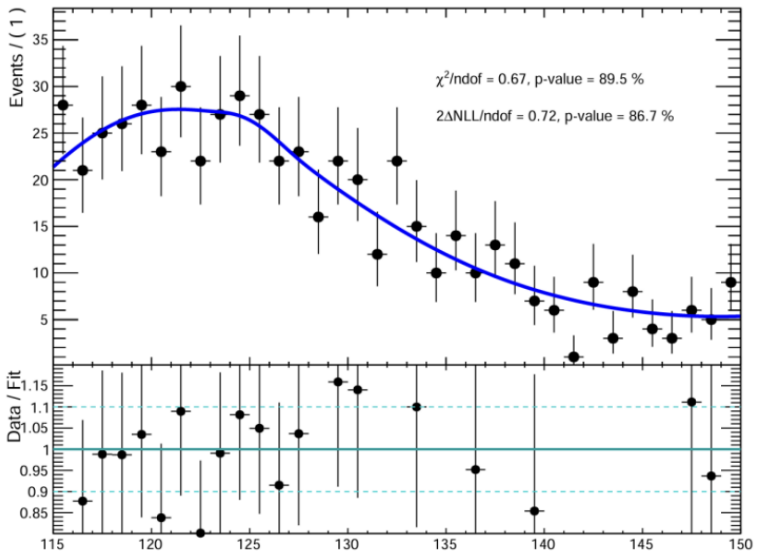
arXiv:2507.12598 (2025)

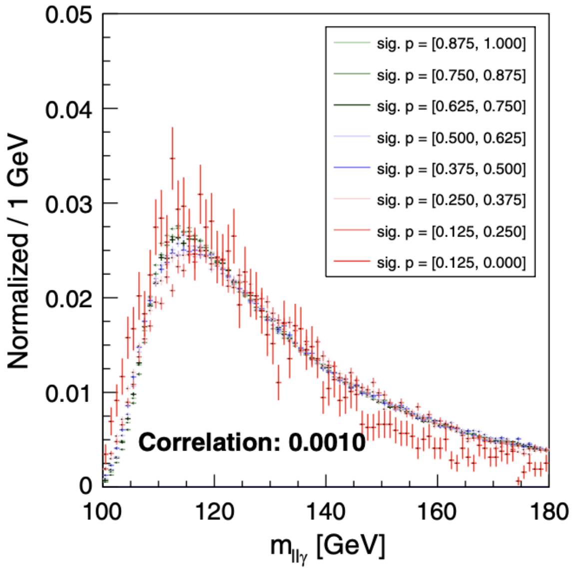
Run 2 + Run 3 combined observed (expected) significance: 2.5$\boldsymbol{\sigma}$ (1.9$\boldsymbol{\sigma}$).

Motivation of New Classifier

Tsung-Dao Lee Institute

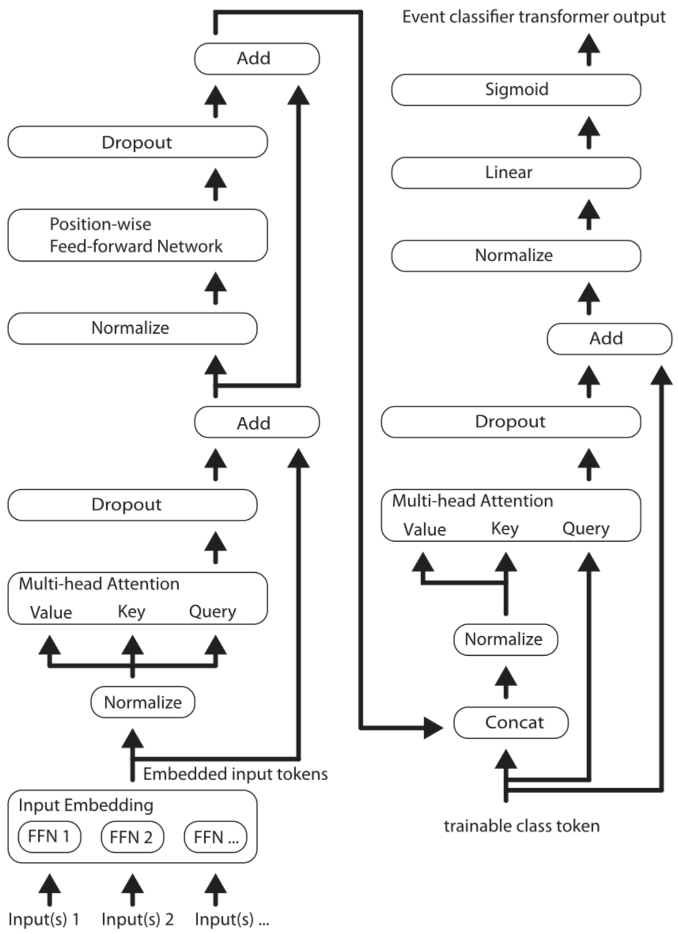
- In partial Run3 analysis, categorized by BDT shows high correlation with mass, which leads to difficulty in background modeling and spurious signal test.
- Recently study(kim2024PRD) shows that a new transformer neural network with special loss function based on CMS H->Zy MC dataset can enhance the significance and reduce the correlation between the network's output and the reconstructed mass.
- The purpose of this work is to test and evaluate the feasibility of the new event classifier on Atlas H→Zγ data, improve and use it in full Run3 H->Zy analysis.





Method

- **Event classifier transformer neural network**



Specialized loss function

Distance correlation (Disco) measures the dependence between output and mass

The DisCo term penalizes the neural network when y and mass are correlated.

$$\mathrm{dCov}^2(X,Y) = \int d^p s\, d^q t \,|f_{X,Y}(s,t) - f_X(s) f_Y(t)|^2 w(s,t),$$

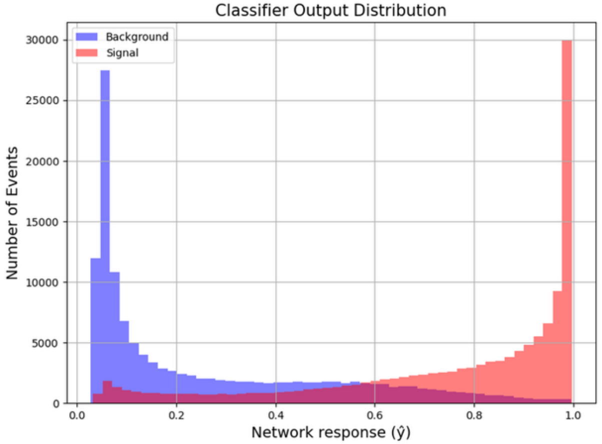
$$\mathrm{dCorr}^2(X,Y) = \frac{\mathrm{dCov}^2(X,Y)}{\mathrm{dCov}(X,X)\mathrm{dCov}(Y,Y)}$$

$$\mathrm{Loss} = \mathrm{Loss}_{\mathrm{classifier}}(\hat{y}, y) + \lambda \cdot \mathrm{DisCo}(\mathrm{mass}, \hat{y}).$$

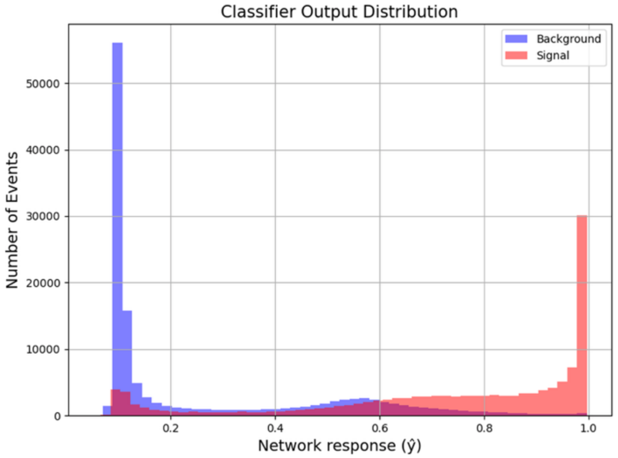
Result of VBF training set

李政道研究所 TSUNG-DAO LEE INSTITUTE

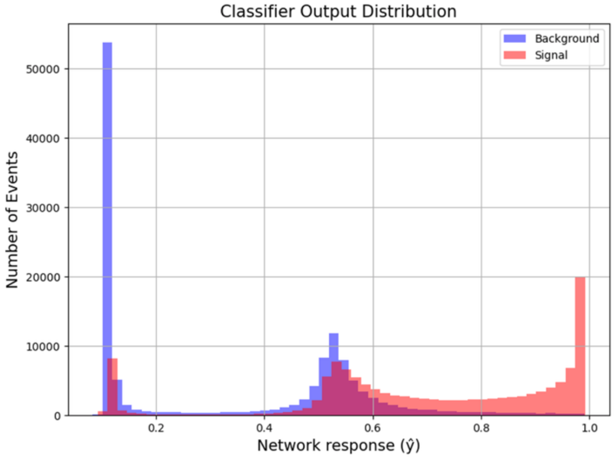
BCE + 10 DisCo



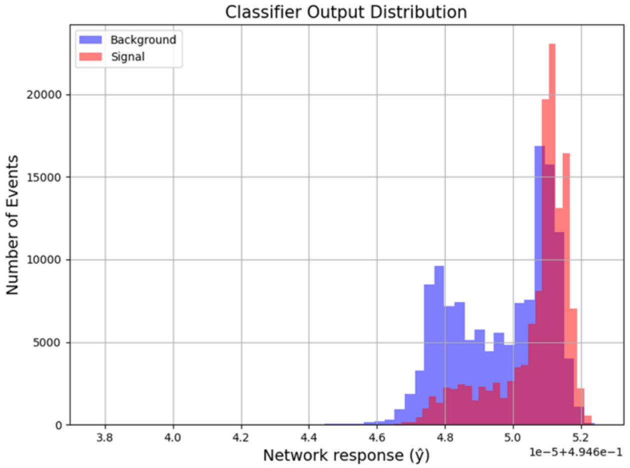
BCE + 15 DisCo



BCE + 20 DisCo



BCE + 30 DisCo

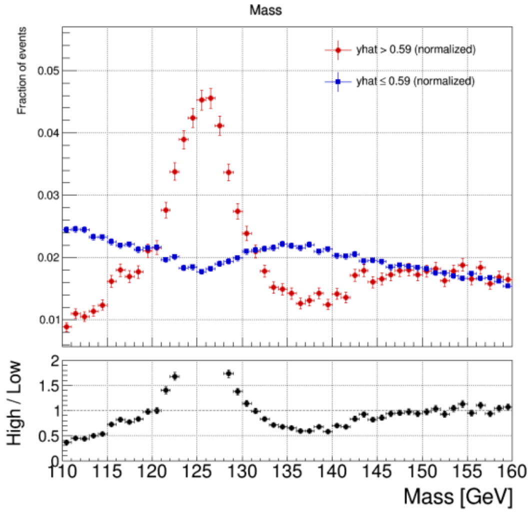


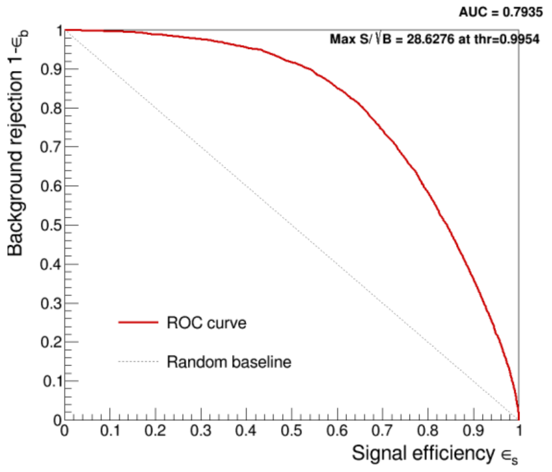
Factor of disco	10	15	20	30
avg loss	0.481855	0.512391	0.584306	0.856971
Significance	46.635 +- 8.847	42.973 +- 4.999	36.411 +- 3.738	23.782 +-5.549

Result of VBF training set

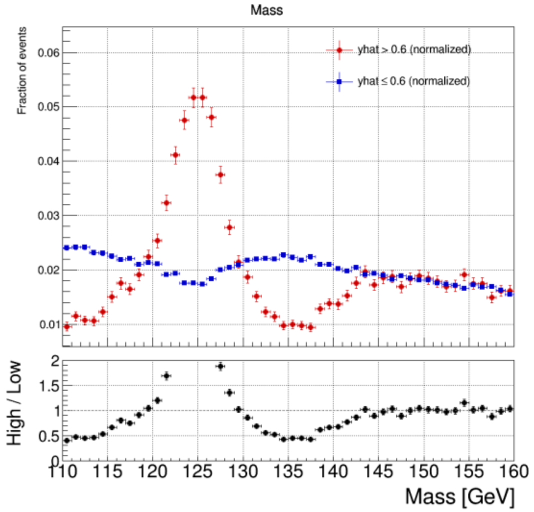
李政道研究所 TSUNG-DAO LEE INSTITUTE

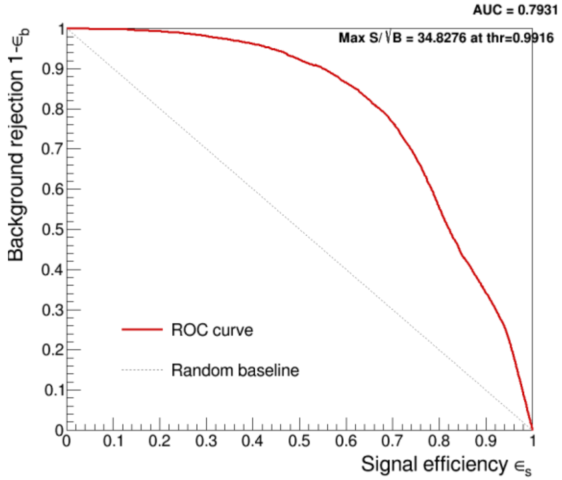
BCE + 10 DisCo



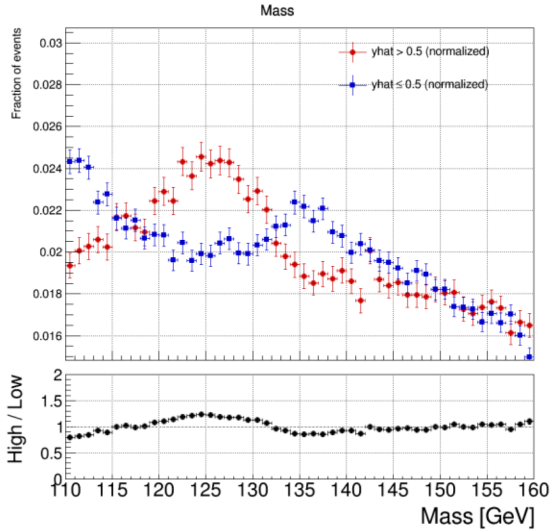


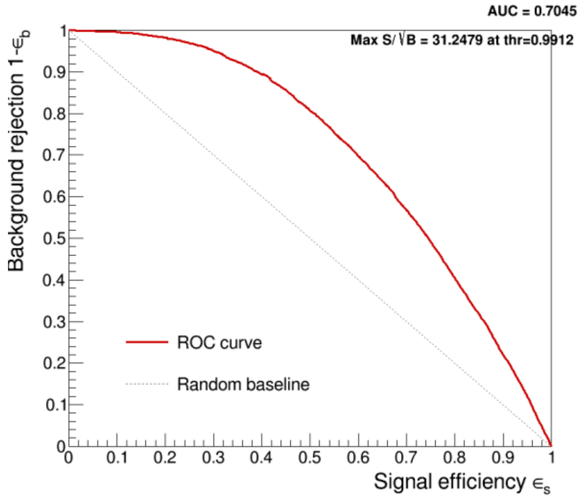
BCE + 15 DisCo



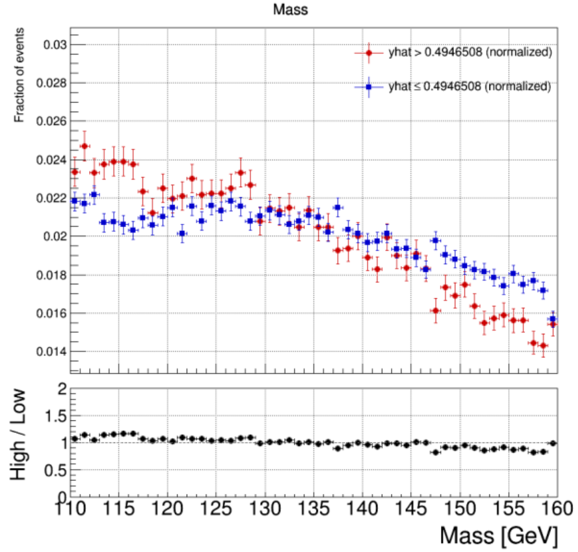


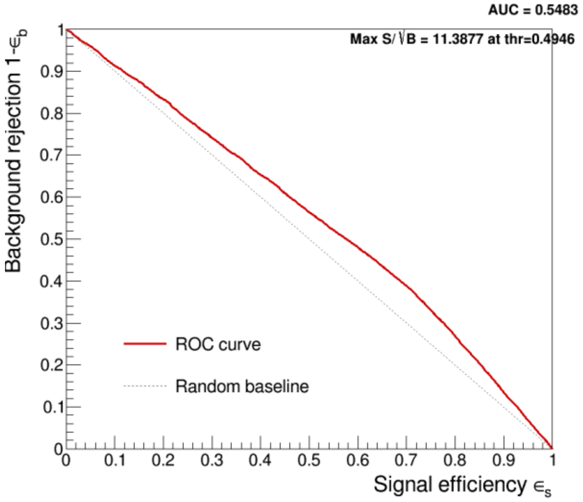
BCE + 20 DisCo





BCE + 30 DisCo





Summary

- The Disco loss helps reduce the strong correlation with mass, but it also weakens classifier performance.

- I have applied the classifier transformer method and obtained preliminary results, next step will be to optimize the network parameters further and better understand its behavior.

- partial Run 3 HZy analysis.

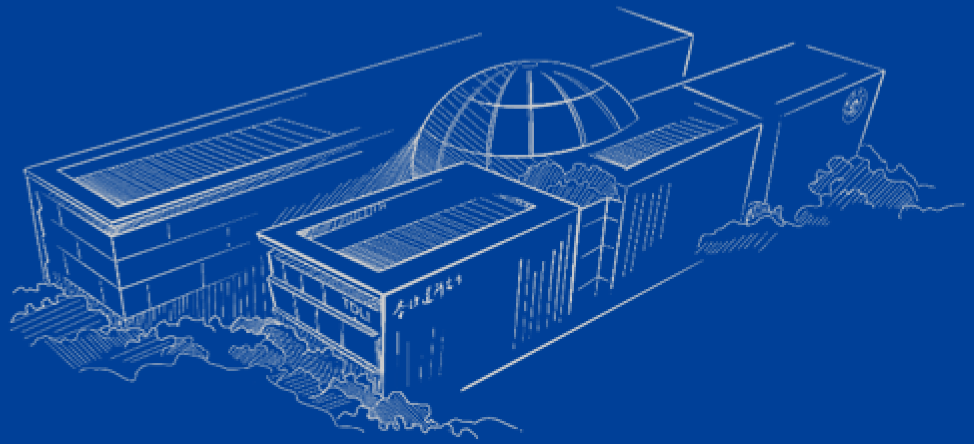
statistical work signal background model cross-check

- plan to the Full Run 3 analysis:

event classifier, the signal and background models, statistical analysis.

Thank you for your attention!

Back up



Setup

- **Input** : x:features (25) y:classID (signal1 background0)
- **Spectators**: llg_m_Zmassconstraint (for Disco)
- **Model**:event classifier transformer neural network
- **Loss function :**BCE ＋ λ DisCo

$$\mathrm{BCE}(\hat{y}, y) = -y \ln(\hat{y}) - (1-y)\ln(1-\hat{y}),$$

- **output**: probability of the event (sigmod)
- Batchsize 1000 epoch 1000

Key factor: loss, significance, AUC of ROC

Model selection: Significance-best

- Divide the dataset into 5 bins based on output. The bins are constructed to have an equal number of signal events. Calculate the significance of each bin and combine the significances of the bins.

$$\text{Significance} = \sqrt{2\left[(N_S + N_B)\ln\left(1 + \frac{N_S}{N_B}\right) - N_S\right]},$$

$$\text{Total significance} = \sqrt{\sum_i^n (\text{Significance}_i)^2}$$

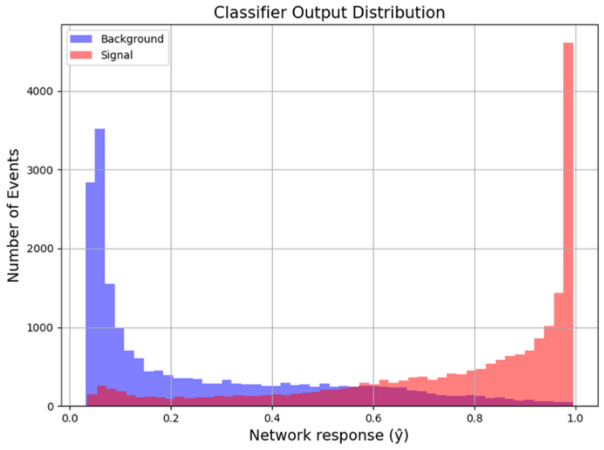
```
# VBF_DRmin_y_j
# VBF_Dphi_Zy_jj
# VBF_Dy_j_j
# VBF_N_j
# VBF_Zepp
# VBF_eta_j1
# VBF_eta_j2
# VBF_m_jj
# VBF_mass_j1
# VBF_mass_j2
# VBF_pT_j1
# VBF_pT_j2
# VBF_pT_jj
# VBF_pTt_Zy
# Zy_Dphi_j1
# ll_eta
# ll_pt
# llg_angles_costheta_ginH
# llg_angles_costheta_linZ
# llg_deta_Zy
# llg_dphi_Zy
# llg_eta
# llg_pt
# ph_eta
# ph_pt
```

Variables	Definition
M_{jj}	Di-jet mass
$\Delta\eta_{jj}$	Pseudo-rapidity separation of dijet
pT_{j1}	lead-jet pT
M_{j1}	lead-jet mass
$\eta^{Zeppenfeld}$	$\lvert\eta_{Z\gamma} - 0.5 * (\eta_{j1} + \eta_{j2})\rvert$
η_{j2}	sublead-jet eta
pT_{jj}	Di-jet pT
$\Delta\Phi_{Z\gamma,jj}$	Azimuthal angle between Zgamma and dijet system
M_{j2}	sublead-jet mass
$\Delta R^{min}_{\gamma or Z,j}$	Minimum ΔR between one object of the Zgamma and jets
pT_{j2}	sublead-jet pT
η_{j1}	lead-jet eta
$pT_{ll\gamma}$	$ll\gamma$ pT
$\Delta\Phi_{Z\gamma,j1}$	Azimuthal angle between Zgamma and lead-jet
pT_{γ}	photon pT
pT_{ll}	ll pT
N_j	Number of jets pT
$\Delta\Phi_{Z,\gamma}$	Azimuthal angle between di-lepton system and photon
$cos\theta(ll,\gamma)inZ$	cosθ in Z rest system
η_{ll}	eta of Z
$\eta_{ll\gamma}$	eta of $ll\gamma$
$cos\theta(ll,\gamma)inH$	cosθ in Higgs rest system
η_{γ}	eta of photon
p_{Tt}	Zgamma p_{T} projected perpendicular to the Zgamma thrust axis
$\Delta\eta_{Z,\gamma}$	Pseudo-rapidity separation of Z γ

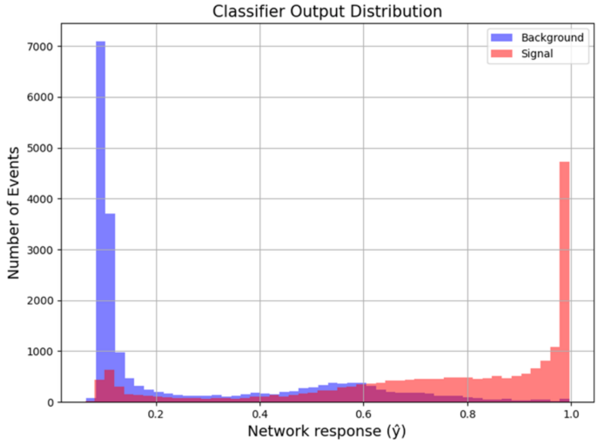
Result of validation set

李政道研究所 TSUNG-DAO LEE INSTITUTE

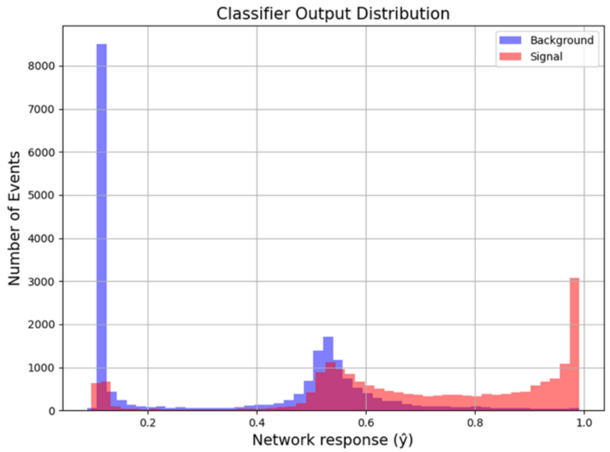
BCE + 10 DisCo



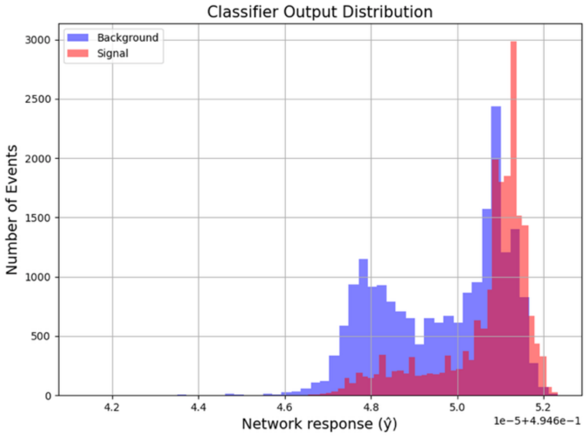
BCE + 15 DisCo



BCE + 20 DisCo



BCE + 30 DisCo

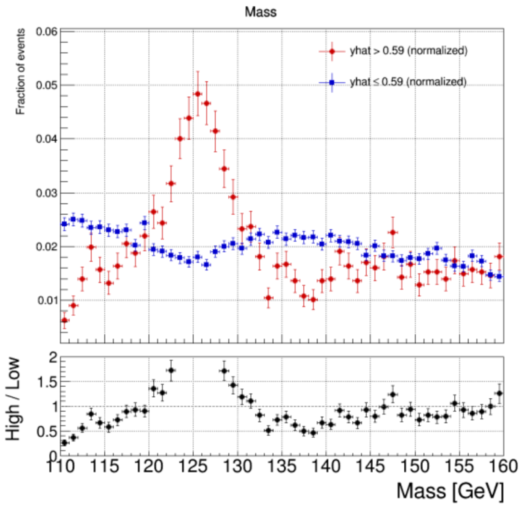


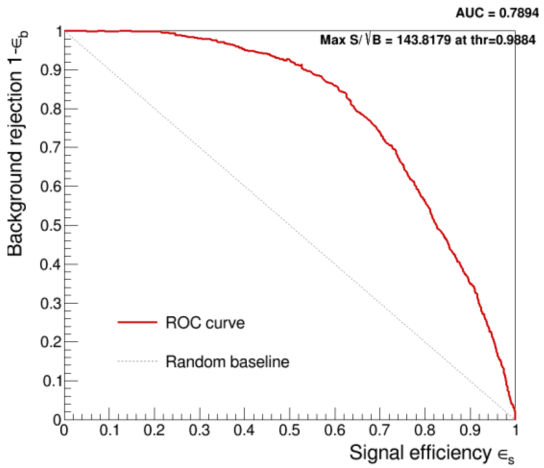
Factor of disco	10	15	20	30
avg loss	0.500804	0.516626	0.585283	0.868690
Significance	13.923 +- 1.817	13.648 +-1.817	14.397 +- 2.479	9.428 +- 0.616

Result of validation set

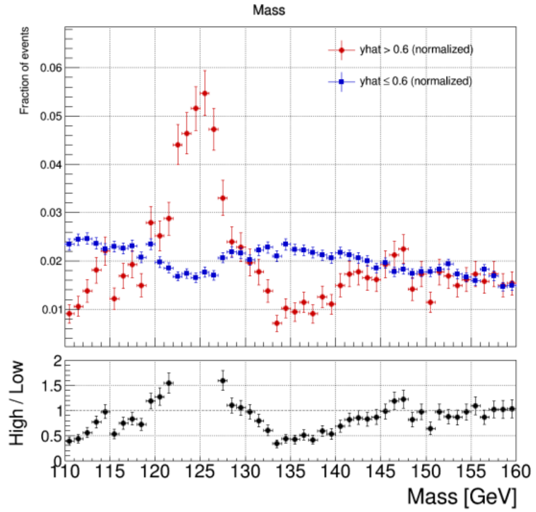
李政道研究所
TSUNG-DAO LEE INSTITUTE

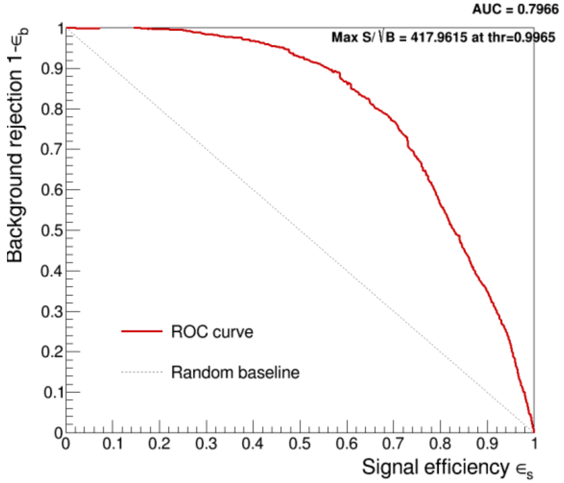
BCE + 10 DisCo



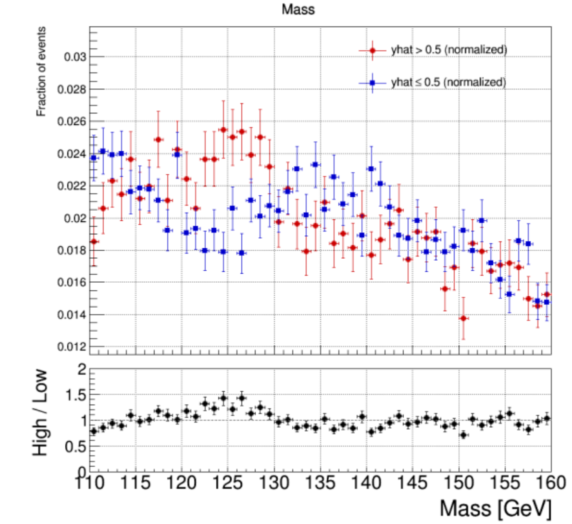


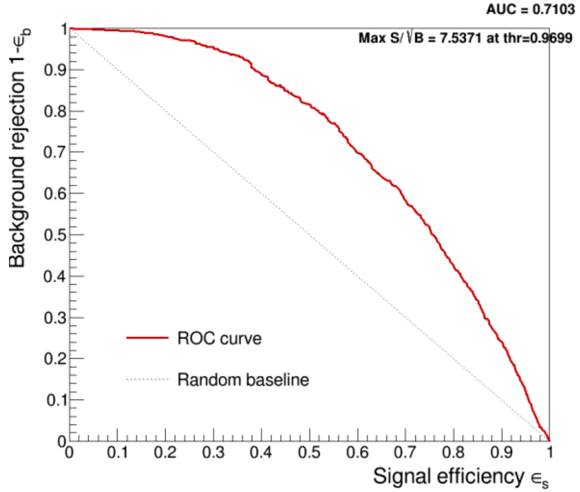
BCE + 15 DisCo





BCE + 20 DisCo





BCE + 30 DisCo