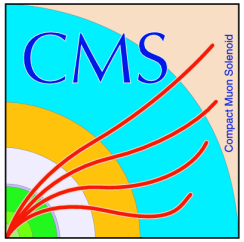


Search for Higgs boson decays to Zγ with the CMS detector in proton-proton collisions at √s=13TeV with 137 fb^{-1}

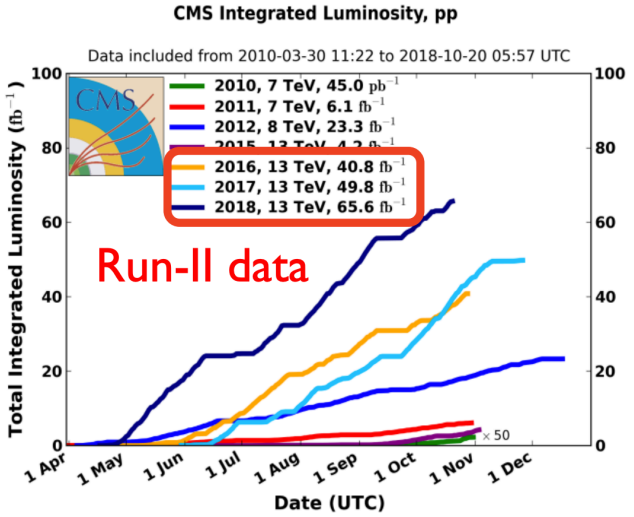
Presenter: Ming-Yan, Lee
Supervisor: Chia-Ming, Kuo
National Central University
2021/01/22, TIDC meeting

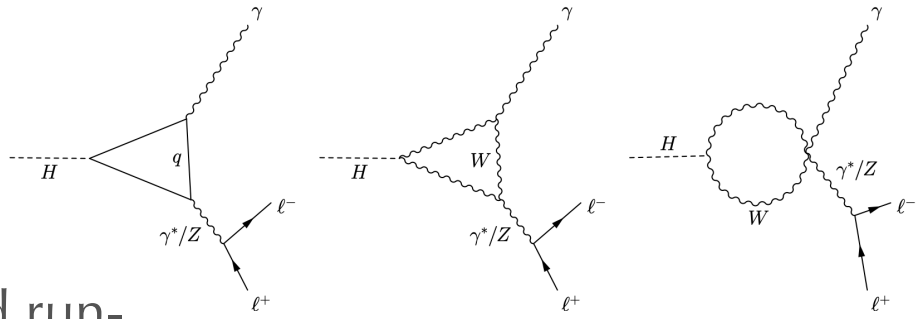
Introduction

- A rare decay(BR(H→Zγ→llγ)=1.5x10^{-4}) is induced by loop diagram, becoming more important from the upcoming periods
- New Physics
  - Sensitive on the new physics inside loop
- SM physics measurements
  - NLO contribution of H→ℓℓγ
- Both CMS and ATLAS performed the analysis of H→Zγ search in run-I and run-II(2016), ATLAS publish full run2 recently

→No significant excess over background, but starts to have sensitivity to probe this channel

→Analysis methods are crucial!





Limit	Observed	Expected
CMS(2016)	8	5.1
Atlas (2016)	6.6	5.2
Atlas run2	3.6	1.7

FSR γ recovery & kinematic refit

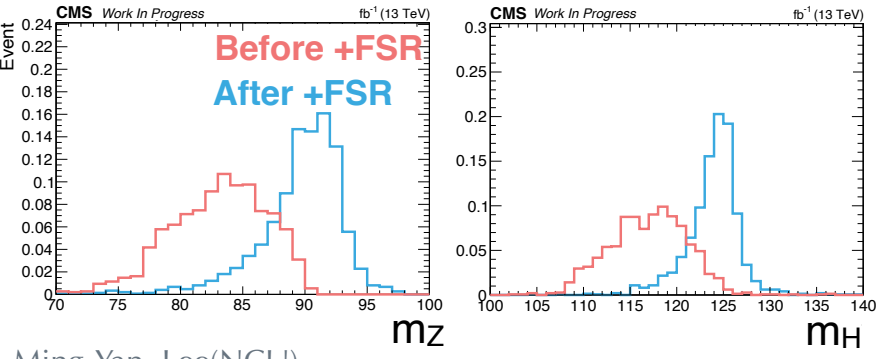
Improving mass scale and resolution, 12% improvement on expected limit

FSR photon recovery

- Recover FSR γ for the muons to improve mass scale and resolution

➔Improves 2.5% on the Higgs mass resolution, 1% events have an FSR γ selected

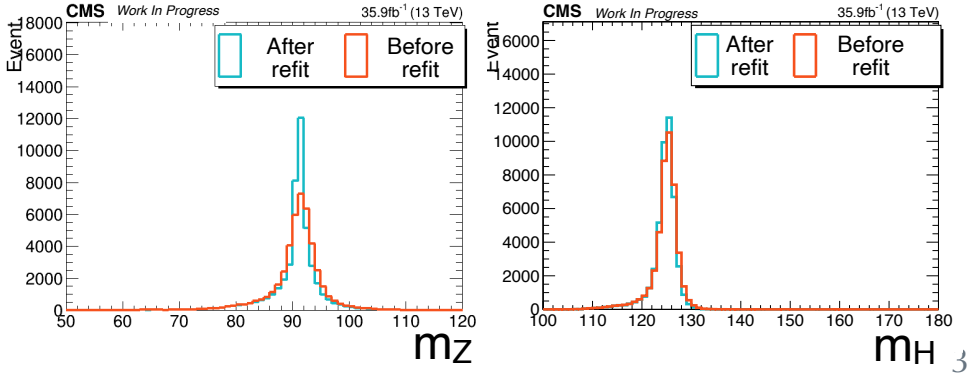
Events with FSR



Kinematic refit

- Perform a constraint ML fit with known quantities(Z mass) on the measured quantities(pT of leptons) make it close to the true value

➔Improves the Z mass resolutions up to 18%

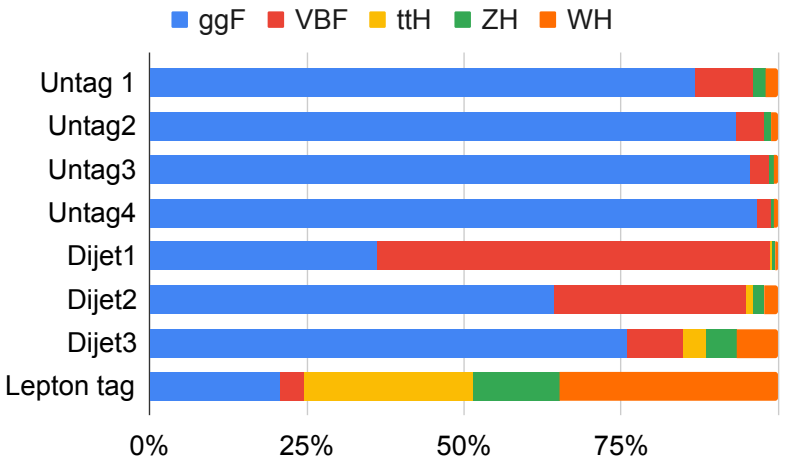


Categorization with MVA

Improve signal and background separation, improves another 12% on expected limit

The Multivariate analysis(MVA) is to build up the multidimensional cut to suppress background and categorizes events

The fraction of each Higgs production mode per category



Lepton tag
(VH+ttH)

Select events with additional lepton(e/μ) with >7/5GeV
It can be : ee+μ&μμ+e& ee+e&μμ+μ **1% contribution in signal**

VBF tag*3
Dijet MVA

Event with dijet, the boundary is defined by dijet MVA.

19% contribution in signal

Untagged*4
kinematic MVA

Use kinematic MVA to determine the boundary of untagged category

80% contribution in signal

Dijet MVA & Dijet categories

- Use dijet information to tag VBF events
- Preselection:Only one jet pair candidate for each event, highest pTs are picked

  (the purity of VBF signal is 65%)
- 3 categories for the dijet categories

Dijet angles

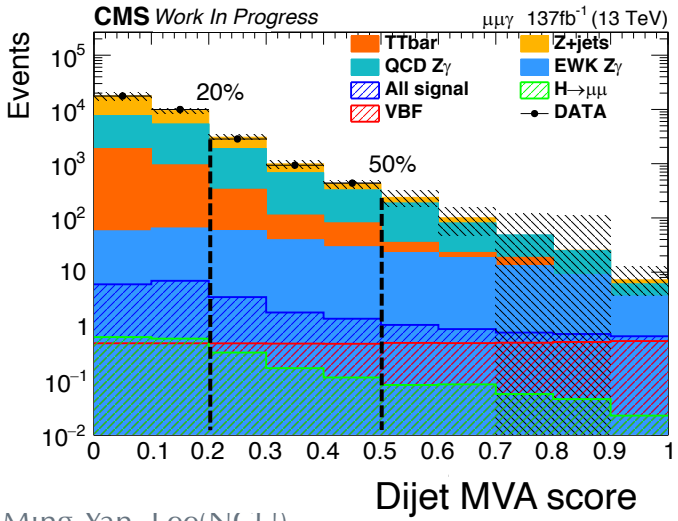
$\lvert\Delta\varphi_{Z\gamma jj}\rvert$	$\Delta R(\gamma,j)$
$\lvert\Delta\eta_{jj}\rvert$	$\lvert\Delta\varphi_{jj}\rvert$
$\lvert\eta_{\gamma}-(\eta_{j1}+\eta_{j1})/2\rvert$(Zeppenfeld)	

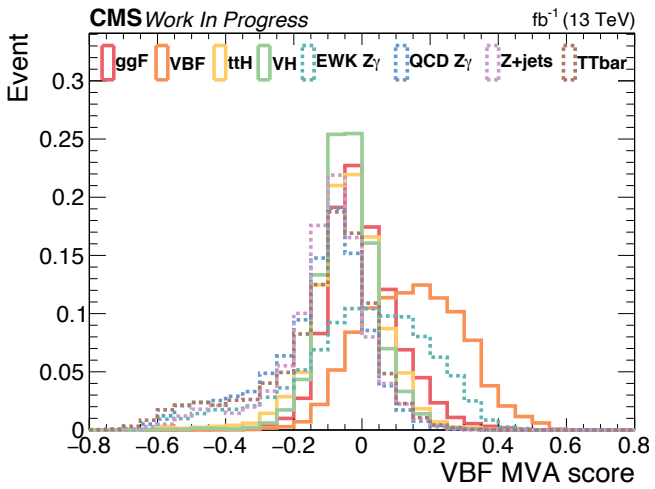
Dijet&llγ pT

Jet1/2 p_T
$\lvert\sum_{Z,\gamma,j1,j2}\vec{p}^{\,i}_{transverse}\rvert / \sum_{Z,\gamma,j1,j2} p_T$ system balance
$PTt_{Z\gamma} = p_T^H \times \frac{\vec{p_T}^{\gamma}-\vec{p_T}^{Z}}{\lvert p_T^{\gamma}-p_T^{Z}\rvert}$

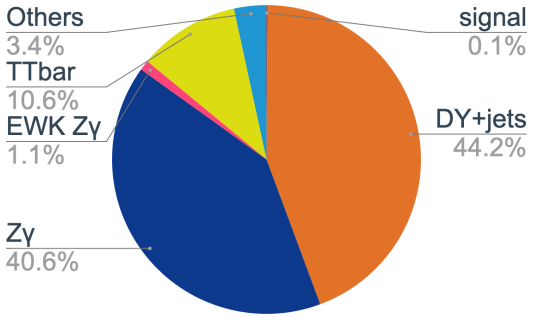
llγ info

Kinematic MVA



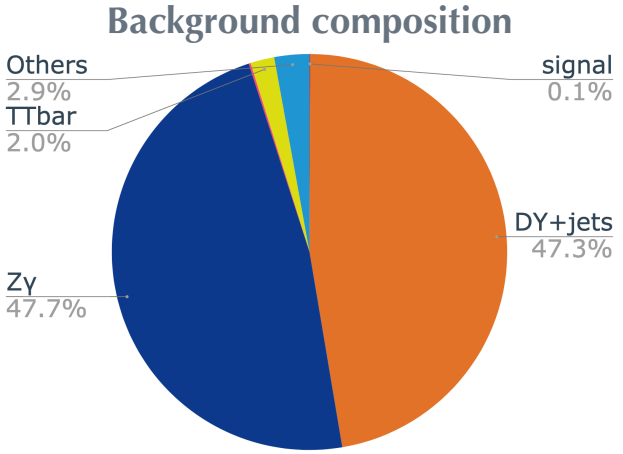


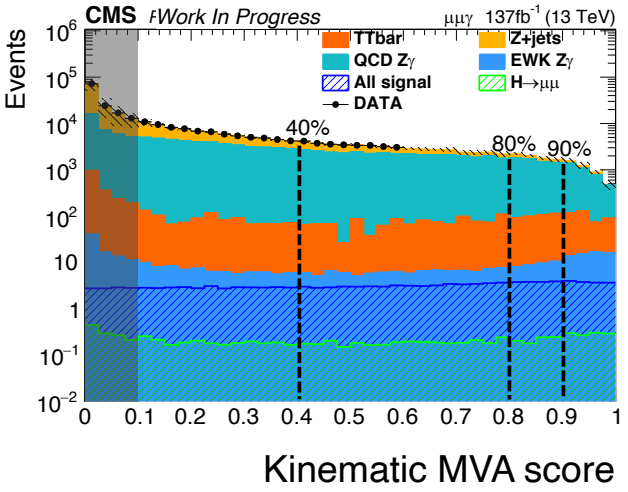
Background composition

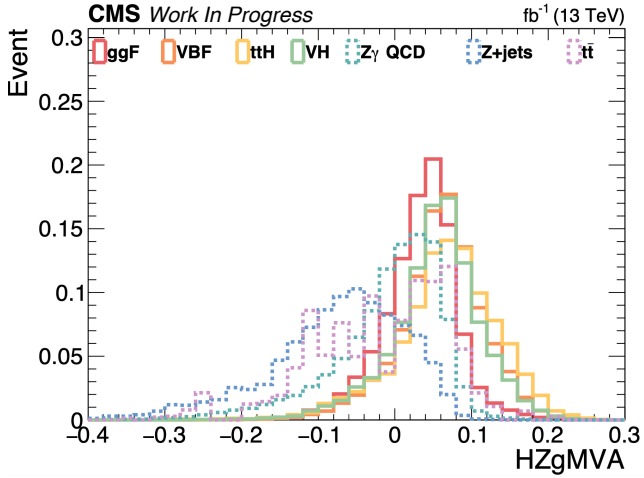


Kinematic MVA& untagged categories

- The signal and background kinematic differences are exploited to build the kinematic MVA
- Events passing selections for all 3 years are combined during training and weighted by luminosity
- We have 4 untagged categories by the kinematic MVA







Angles between final state particles

$\Delta R(l,\gamma)$

$\cos\theta$

Φ

$\cos\Theta$

Kinematics of final state particles

η_γ

η_{l1}

η_{l2}

$IDMVA_\gamma$

σ_γ/E_γ

$pT_{ll\gamma}/m_{ll\gamma}$

Higgs boost **Photon variables**

Summary

- The $H \rightarrow Z\gamma \rightarrow ll\gamma$ is an important rare decays for new physics and understanding on SM
- This analysis approaches to Standard Model sensitivity, improving sensitivity becomes crucial
- This analysis is based on the full run2 data($137fb^{-1}$) taken by CMS
- New analysis techniques such as FSR γ recovery, kinematic refit and MVA methods are implemented for the full run2 analysis
- Categorized by MVAs, 7 categories(3 Dijet+4 Untag)and 1 lepton tag category
- The analysis techniques bring **~24%** improvement w.r.t. the cut-based method

Comparison of analysis techniques

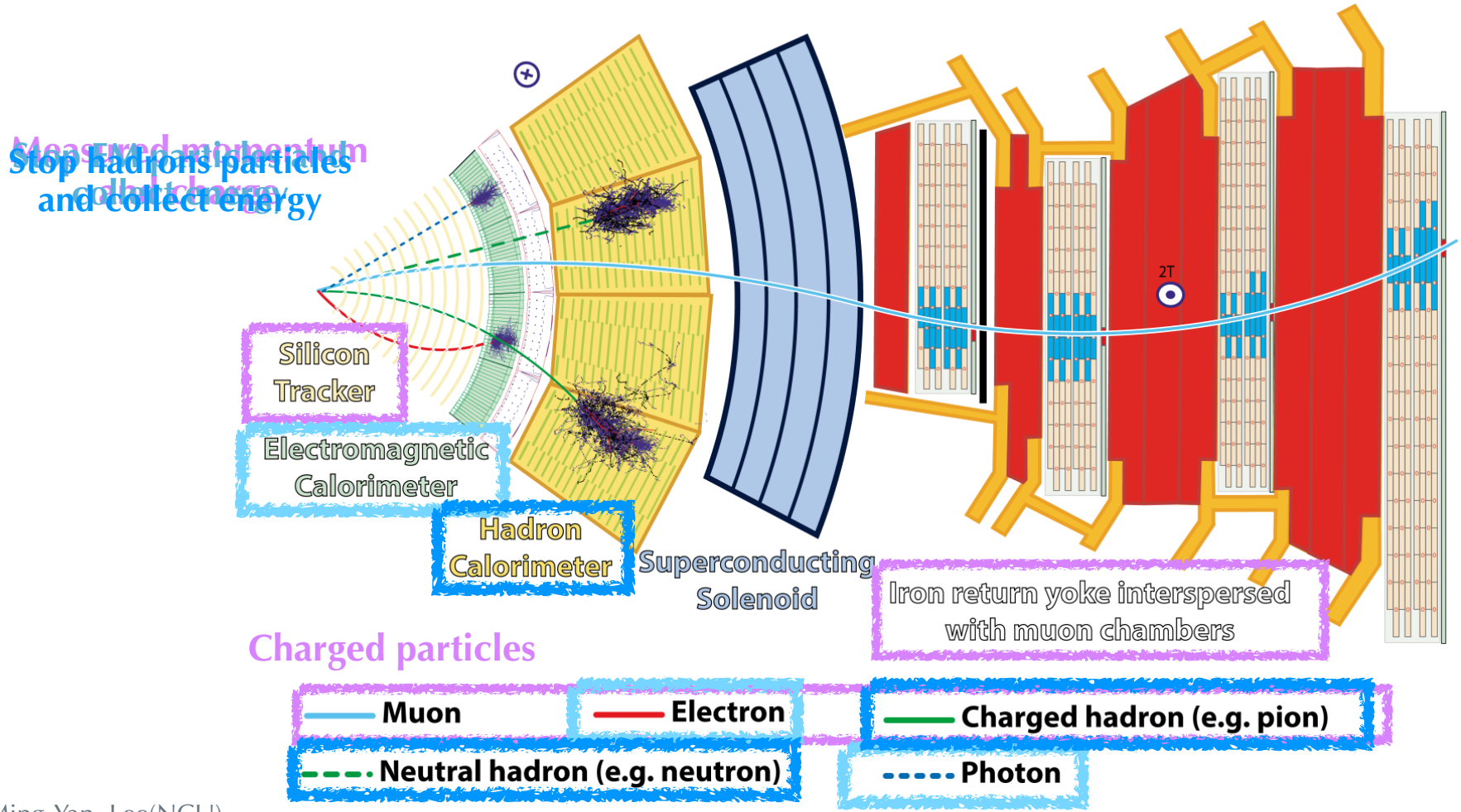
$$Z\gamma\ P_{T_t} = |\vec{P}_T^{Higgs} \times \frac{\vec{P}_T^{\gamma} - \vec{P}_T^{Z}}{\left|\vec{P}_T^{\gamma} - \vec{P}_T^{Z}\right|}|$$

- Compare the analysis techniques between published results with 2016 data and this work
- We changed cutbased analysis to the MVA analysis

		Atlas	HIG-17-007	HIG-19-014
VBFMVA		Yes	No	NEW!
Kinematic MVA		No	No	NEW!
FSR recovery		Yes	No	NEW!
Kinematic refit		Yes	No	NEW!
categorization	Lepton tag	No	Yes	Yes
	VBF tag	Use MVA	Cutbased	Use MVA NEW!
	Boosted tag	Use PTt	Use p_T^H	**Not need!**
	Untagged	Use PTt	Use $R9_\gamma$, η_γ	Use MVA NEW!

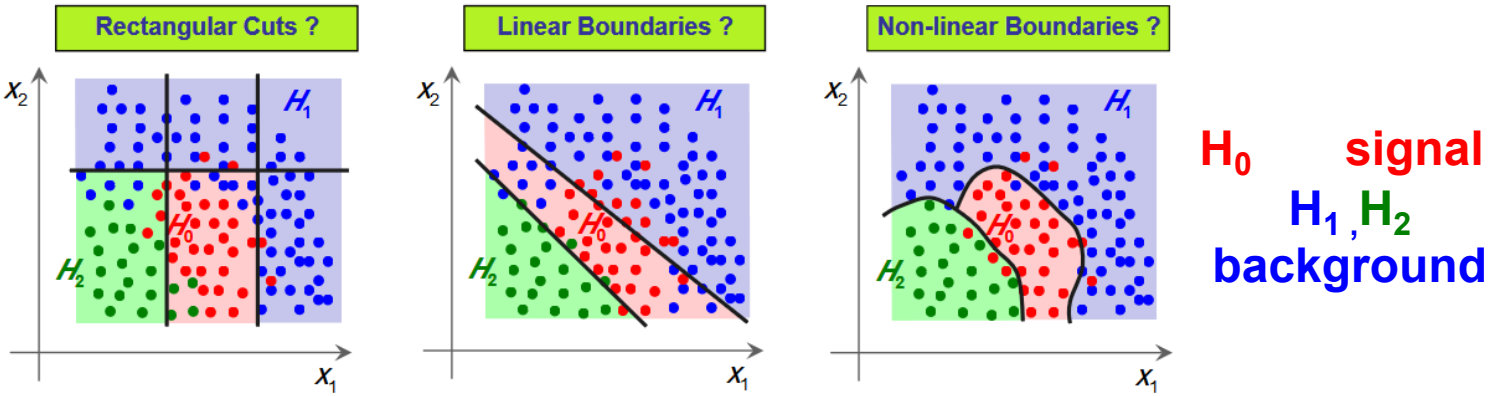
CMS detector

- General purpose detector, include p-p and heavy-ion physics



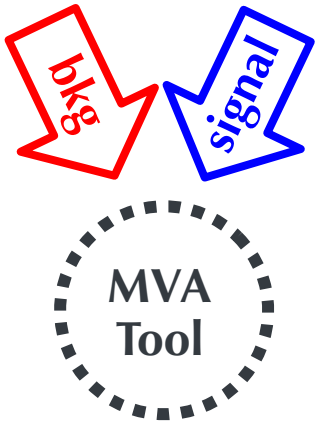
Info for MVA

- What is MVA?(Multivariate analysis)
  - Analyze multiple variables(measurements) on the objects studied, we use Machine Learning method to do it
- Why we need this?
  - Find a multidimensional cut to separate signal and background in our analysis → **better classification**
- Tools : Use TMVA Toolkit→ contains many algorithms



How to use MVA?

Use MC to put into training → Avoid overtraining → Check the performance by ROC curve → Validation on data → Use the MVA score to do the optimization

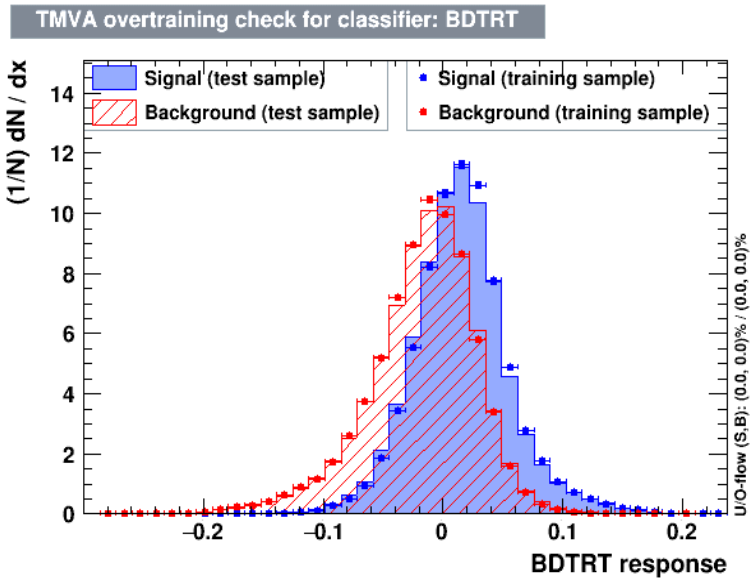


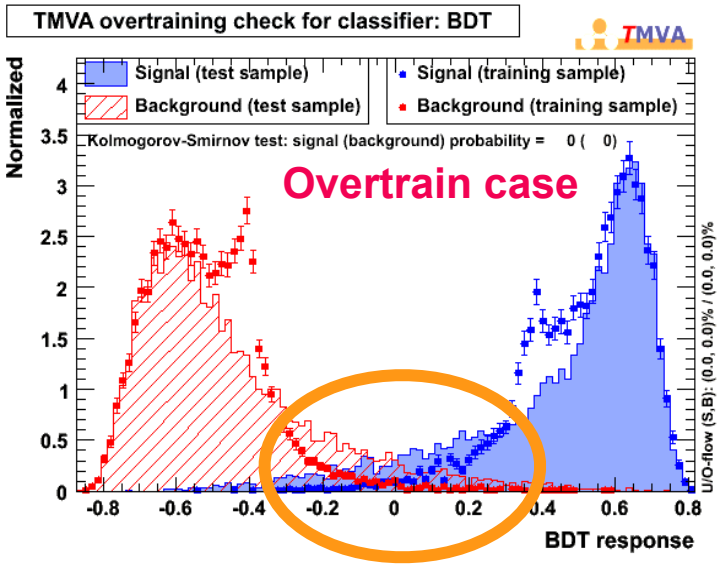
Train sample– build the model

Test sample – test the prediction of the model

How to use MVA?

Use MC to put into training → Avoid overtraining → Check the performance by ROC curve → Validation on data → Use the MVA score to do the optimization

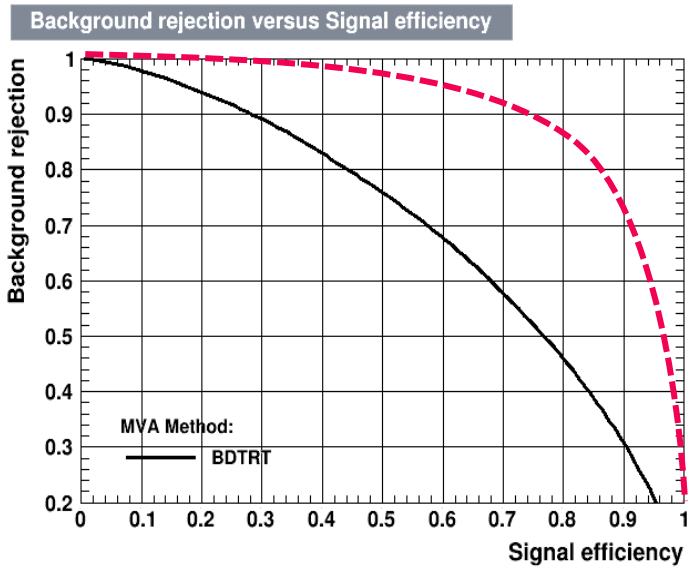




How to use MVA?

Use MC to put into training → Avoid overtraining → Check the performance by ROC curve → Validation on data → Use the MVA score to do the optimization

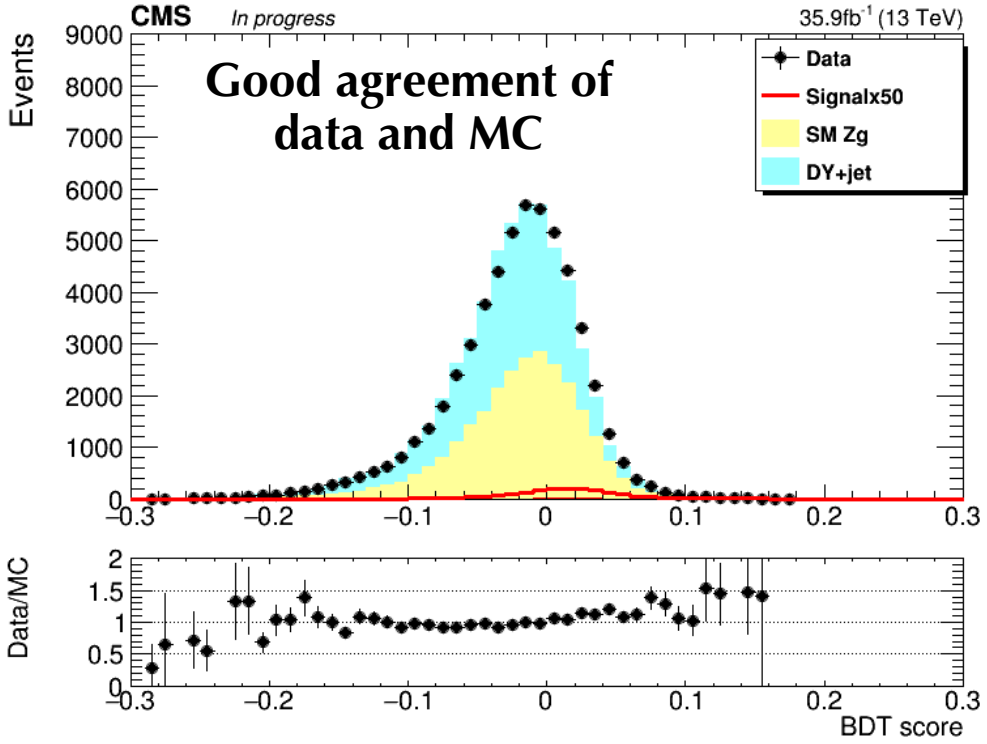
Bkg rejection power = 1- bkg eff



Higher signal efficiency & higher background rejection power

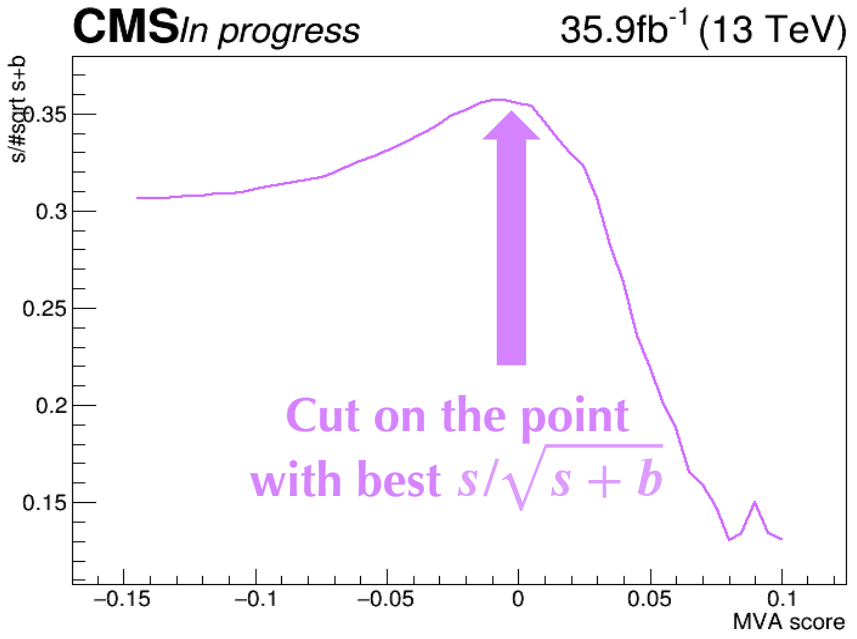
How to use MVA?

Use MC to put into training → Avoid overtraining → Check the performance by ROC curve → Validation on data → Use the MVA score to do the optimization



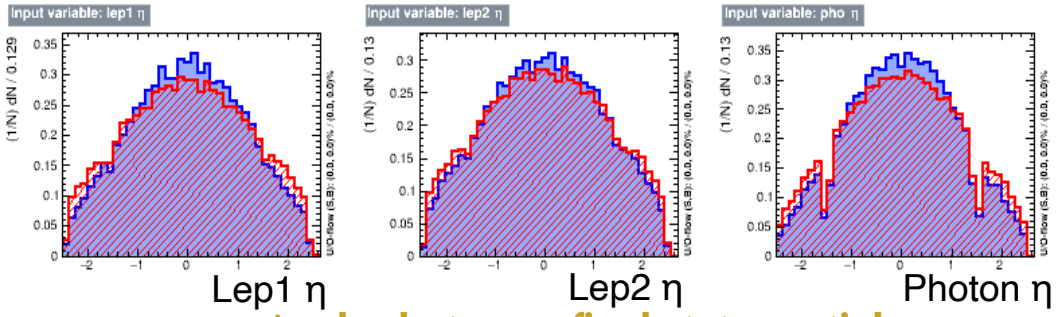
How to use MVA?

Use MC to put into training → Avoid overtraining → Check the performance by ROC curve → Validation on data → Use the MVA score to do the optimization



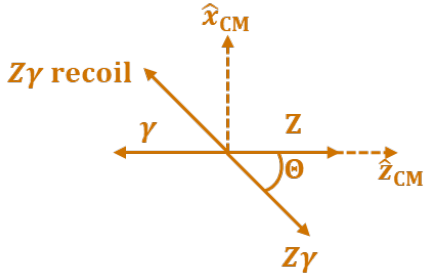
Kinematic MVA
Training variables

The lepton/photon η correlate to m_{lly} resolution

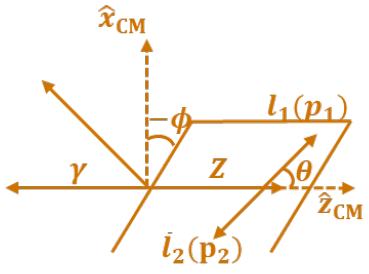


Angles between final state particles

CM frame

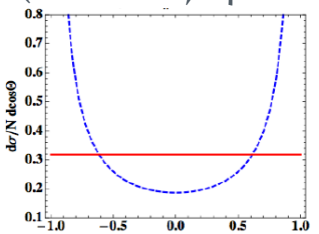


Z rest frame

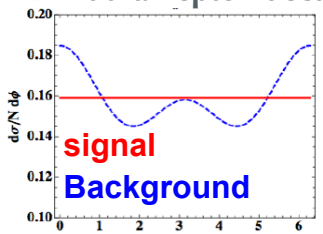


arXiv:1112.1405
arXiv.1108.2274

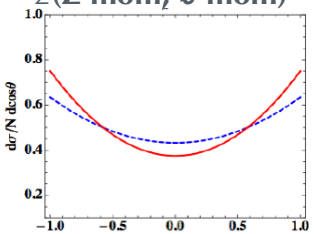
CM frame

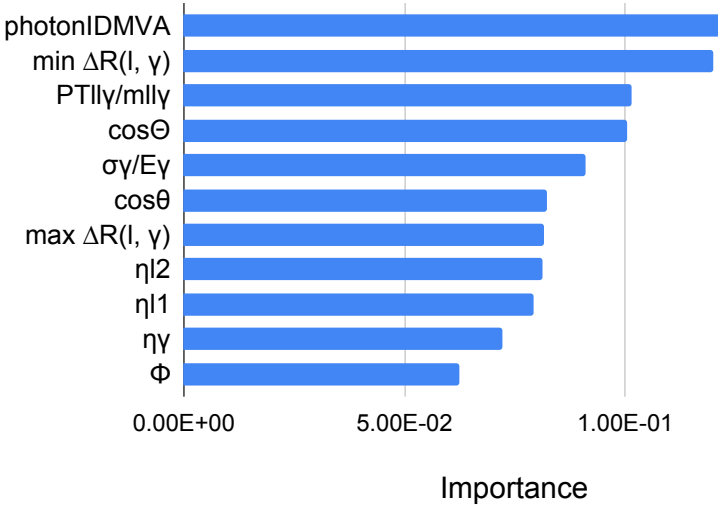


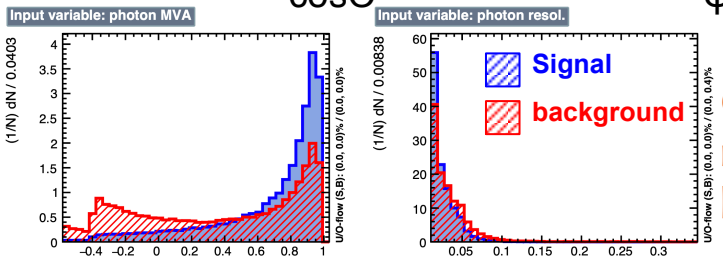
Z rest frame



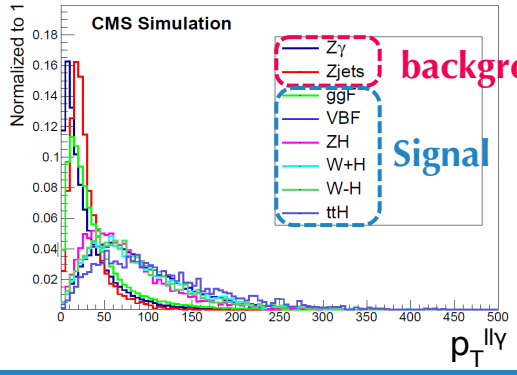
Z CM frame





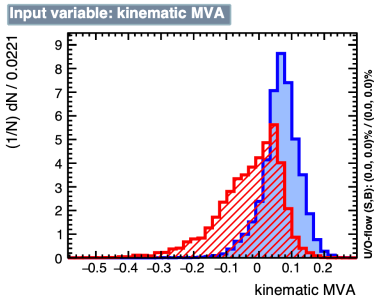


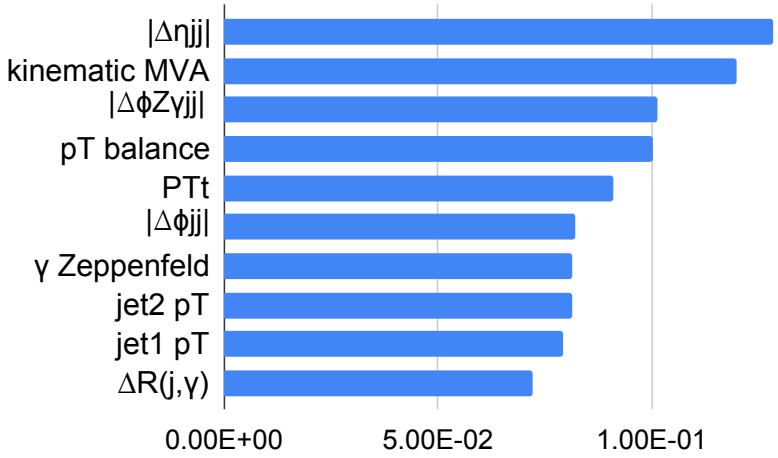
Good photon measurements give better m_{lly} resolution



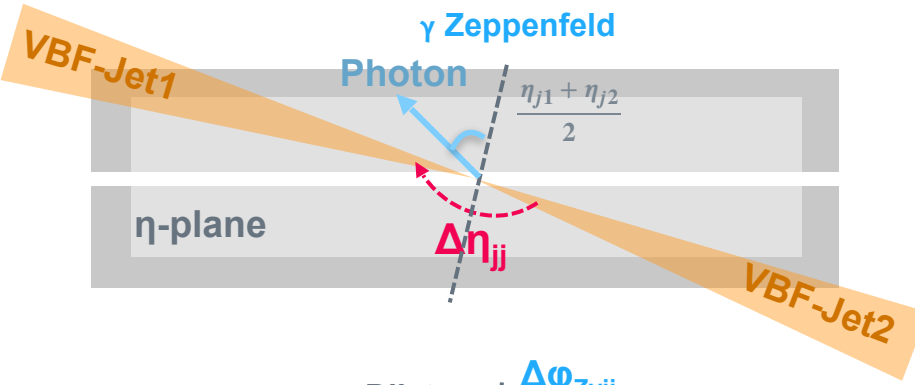
Relative lower bkg for Higgs with high P_T

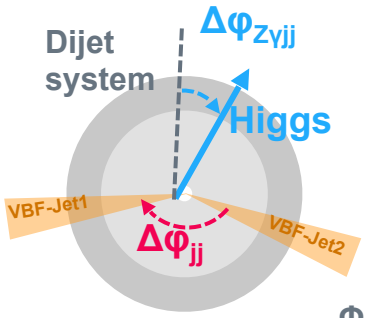
Dijet MVA
Training variables





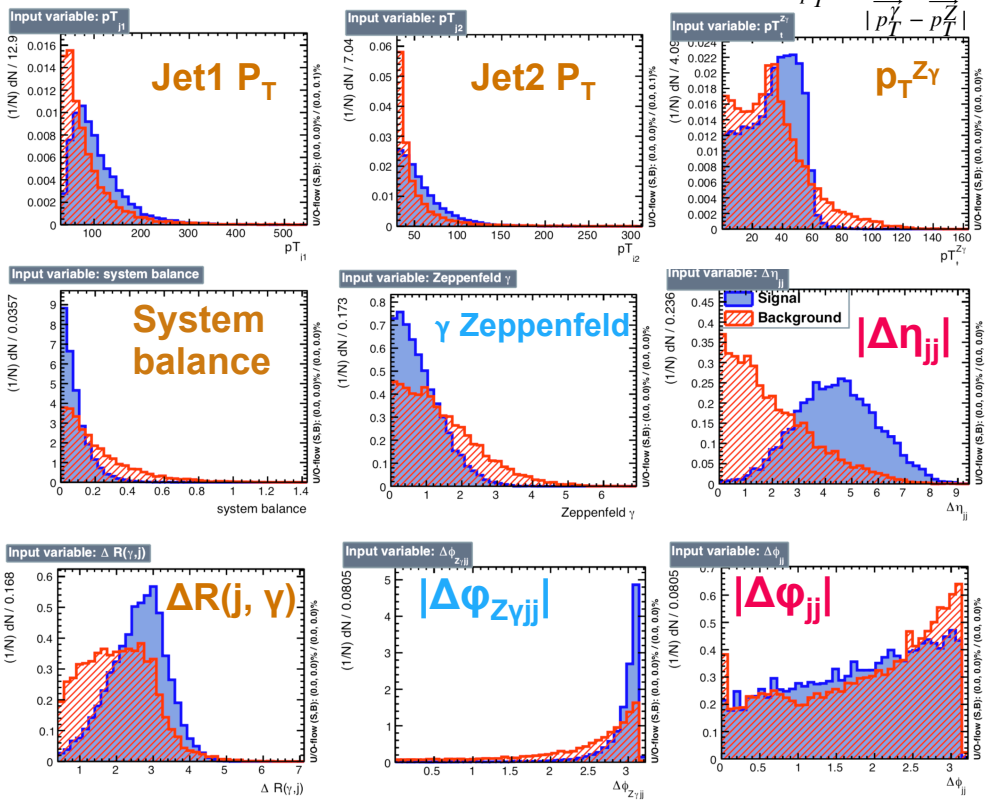
- Comparing to the QCD jets, the VBF events have clear signature
  - larger separation between 2 jets
  - Two jets with high P_T





Importance

$p_T^H \times \frac{\vec{p_T^\gamma} - \vec{p_T^Z}}{|\vec{p_T^\gamma} - \vec{p_T^Z}|}$

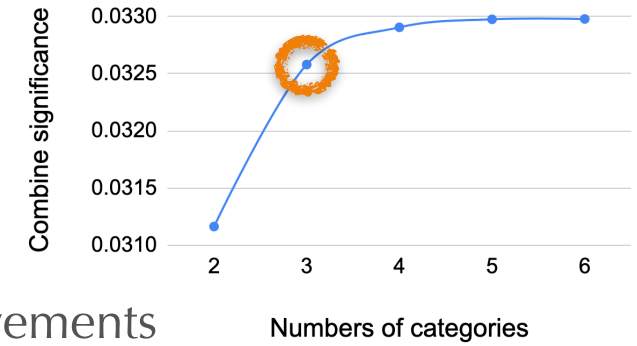


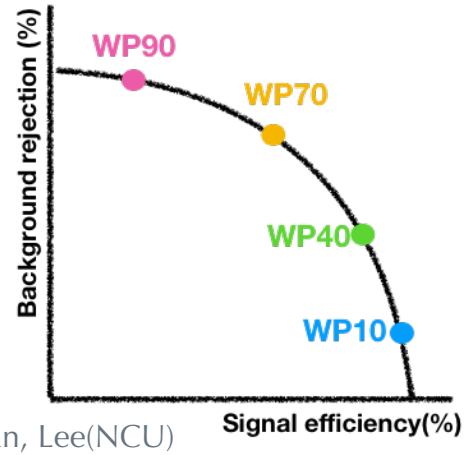
Methodology of categorization(MVA-based)

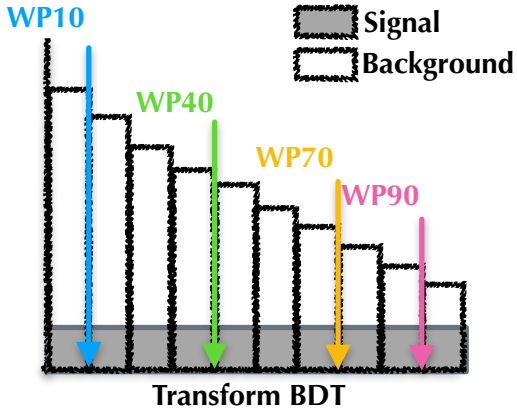
To prevent the bias on the MVA score, we use the **combine significance** for categorization

1. Define the boundary by the signal efficiency (Use transform BDT for intuitive expression)
2. Split to **n** categories and check all possible combination to compute the combined significance
3. Pick up the point starting to reach the plateau
4. Remove the worst s^2/b category to check the improvements

The combination of n categories that yields the combined significance is on the plot







$$\text{Combine significance} \equiv \sum_i^n \frac{S_i^2}{B^i}$$

S_i - signal yield(MC)

B_i- background yield(MC)

Calculate event yield (from MC) within $m_{ll\gamma}$ 120-130GeV