CP measurement in H→$\tau\tau$ decay

Analysis Status and Future Plans

Vinay Krishnan, Diwakar Vats, Arun Nayak
Institute of Physics, Bhubaneswar
India

10th Oct. 2019
Higgs-CP Workshop
DESY - Hamburg

Overview

- Introduction
- Quick overview of some earlier studies about detector resolution effects
- Generator level studies on $\pi^+\pi^-$, $\rho\rho$ channels using IP method
- Studies on vertex resolution
- Reco level studies on acoplanarity angle
- Studies on 2017 data
  - Data/MC agreement
  - MVA/NN implementation
  - Implementation of FF method underway
- Datacard production, combine setup – Near future

Introduction

arXiv 1308.2674

Angle between tau decay planes (IP vectors in Higgs rest frame) is sensitive to Higgs CP

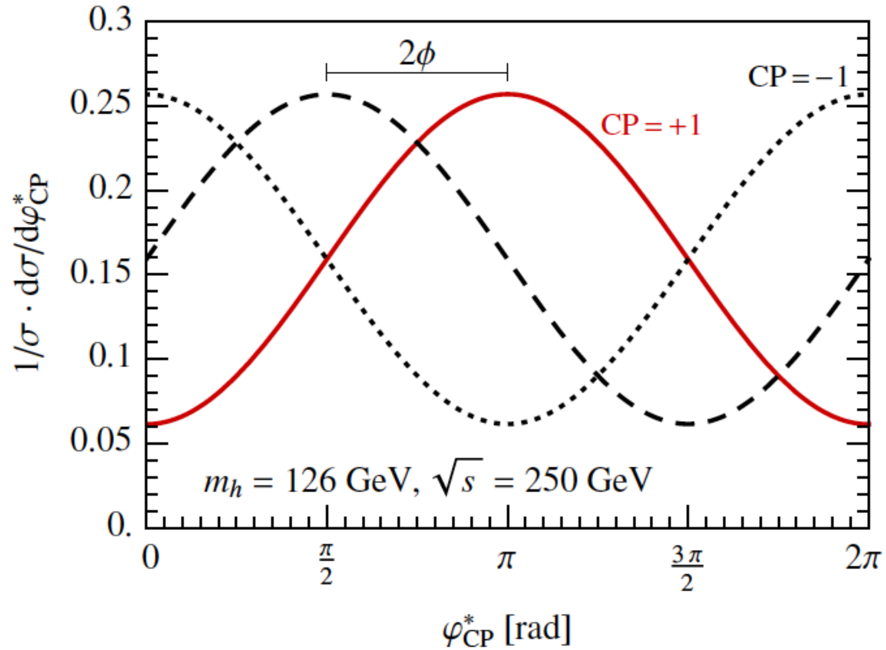
$$\varphi^* = \arccos(\hat{\mathbf{n}}_\perp^{*+} \cdot \hat{\mathbf{n}}_\perp^{*-}),$$

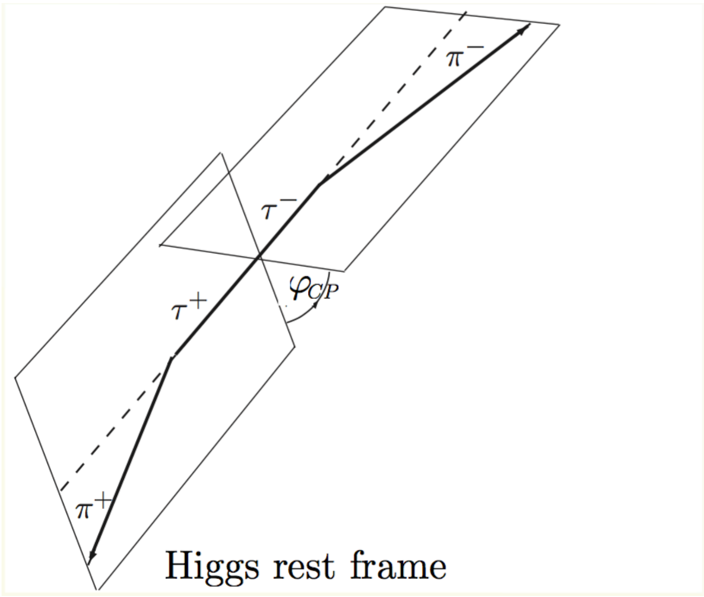
Discriminates CP-odd from CP-even

$$\mathcal{O}^*_{CP} = \hat{\mathbf{q}}^*_- \cdot (\hat{\mathbf{n}}_\perp^{*+} \times \hat{\mathbf{n}}_\perp^{*-}),$$

Discriminates CP-invariant from CP-mixture

$$\varphi^*_{CP} = \begin{cases} \varphi^* & if \;\; \mathcal{O}^*_{CP} \geq 0, \\ 2\pi - \varphi^* & if \;\; \mathcal{O}^*_{CP} < 0. \end{cases}$$





Decay plane of each tau is reconstructed by:

- Momentum of charged pion track
- Vector between primary vertex (PV) and point-of-closest approach (PCA) of pion track to PV (IP vector)

Earlier Studies

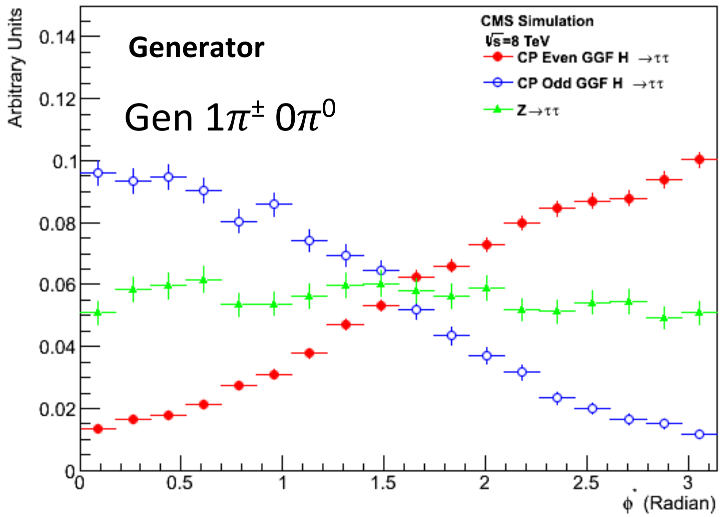
- We had performed some studies earlier, using 8 TeV MC samples, to understand the impact of detector resolutions on CP observable
  - Documented in a small AN,

  (AN-2015/169 -- AN, Alexei Raspereza, Christian Veelken)
- We had considered mostly the $1\pi^{\pm}+0\pi^{0}$ decay modes, and partially $1\pi^{\pm}+1\pi^{0}$ modes
- Studies were performed using AOD samples

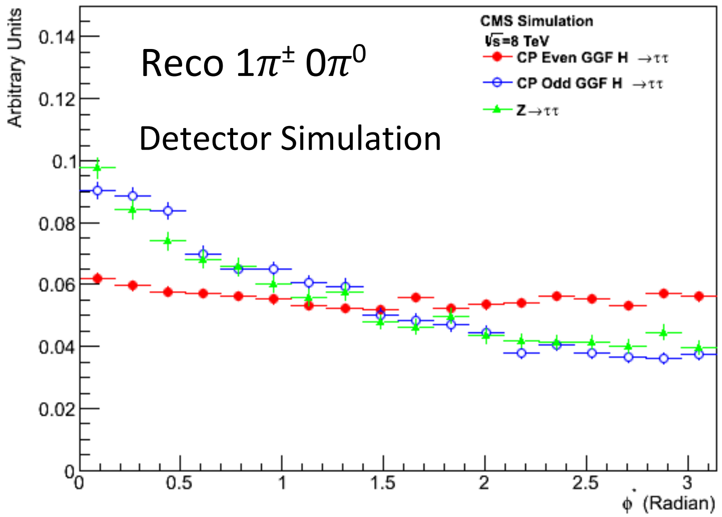
Sensitivity with Simulation

Separation between Higgs CP=+1 and CP=-1 as expected, if we take momentum of charged Pion, PV and PCA position from generator level

With detector simulation and reconstruction,the sensitivity is much reduced and the distribution is shifted towards lower angle



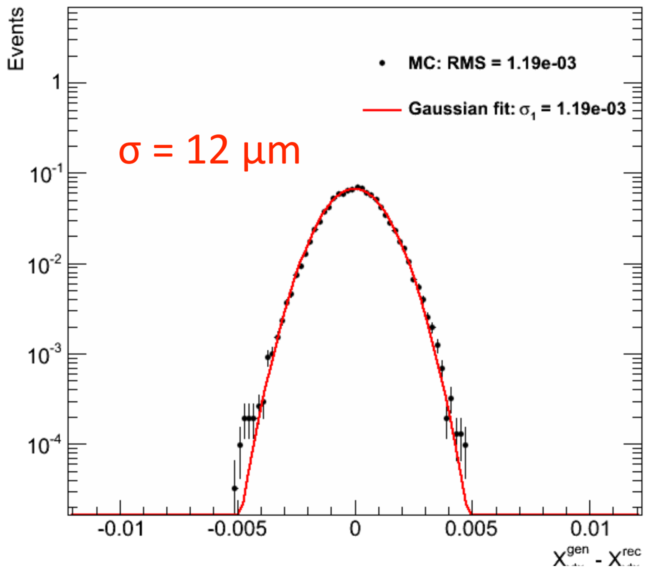
$P_T(\tau) > 20$ GeV
$|\eta(\tau)| < 2.1$



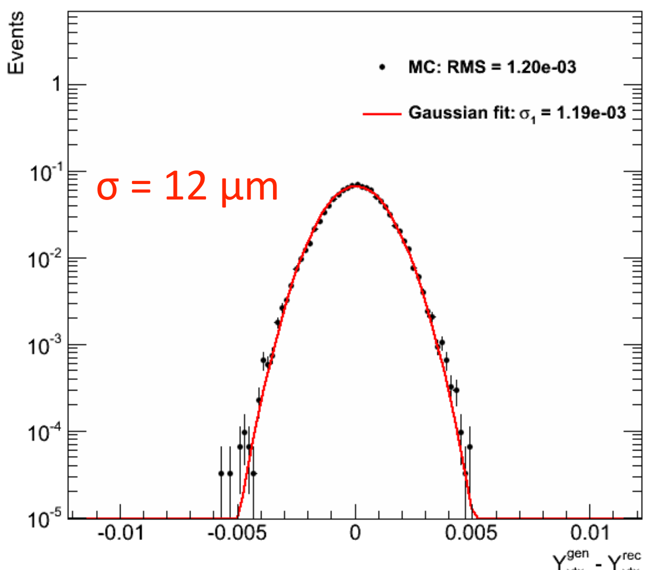
Two effects observed:

1. Reduction in separation between CP Odd and Even.
2. The whole distribution is shifted towards CP-odd direction

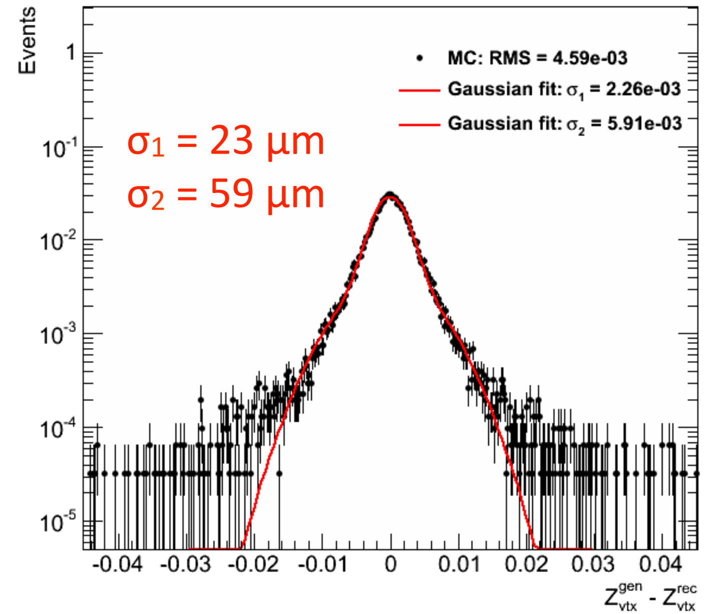
Primary Vertex Resolution



ΔX



ΔY



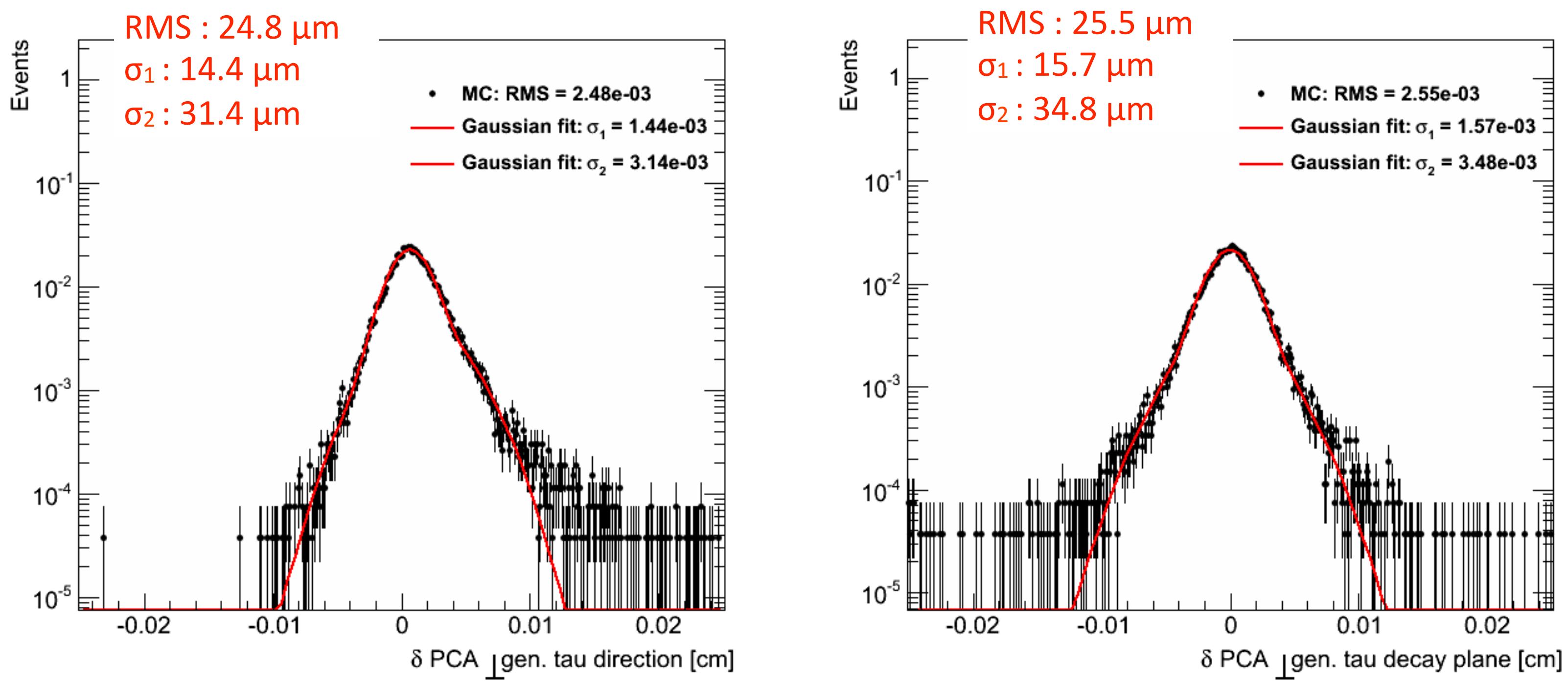
ΔZ

→ Primary vertex re-fitted after excluding tracks from tau candidates

→ Non-Gaussian tails negligible for PV resolution in x, y direction

→ Presence of non-Gaussian tails have a sizeable effect on the RMS in z direction

Resolution on PCA



➔ Non-Gaussian tails have a sizeable effect on the RMS

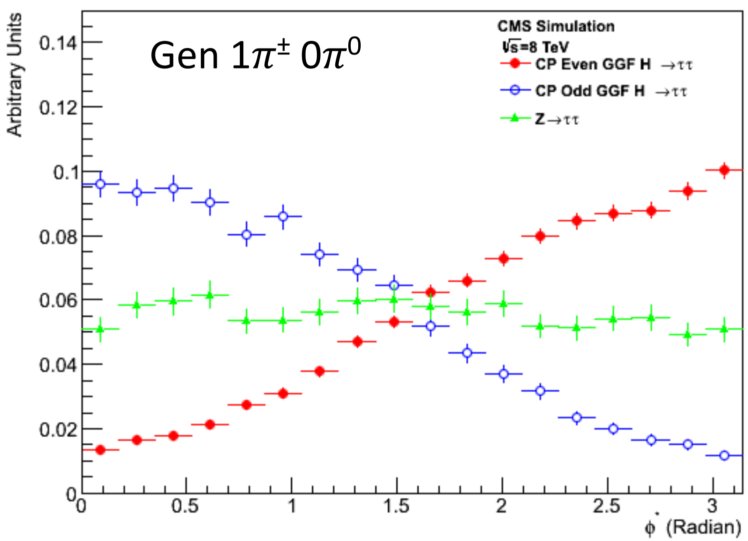
We tried to smear the vertex and PCA position in the generator level according to their detector level resolution and check if it reproduces the detector level behaviour.

Vertex & PCA Smearing

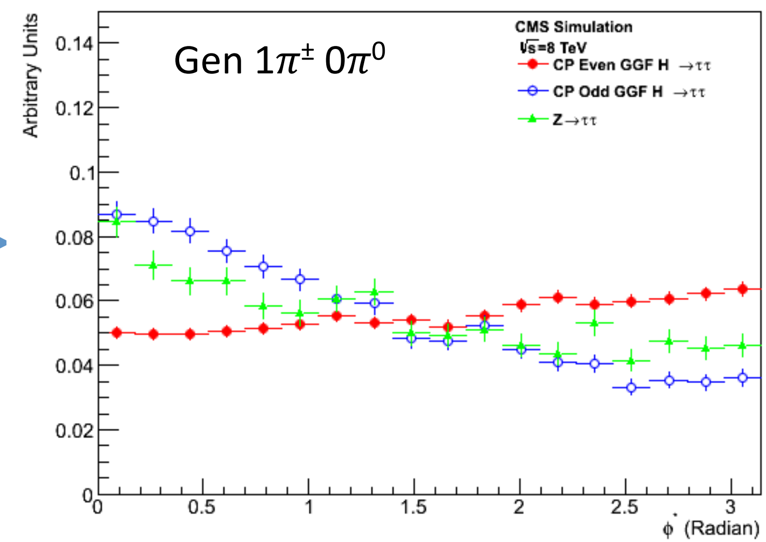
Vertex position smeared with Gaussian: σ_x = 10 µm, σ_y = 10 µm, σ_z = 30 µm

PCA smeared with Gaussian: σ_x = 20 µm, σ_y = 20 µm, σ_z = 20 µm

No Smearing



With both vertex and PCA Smearing

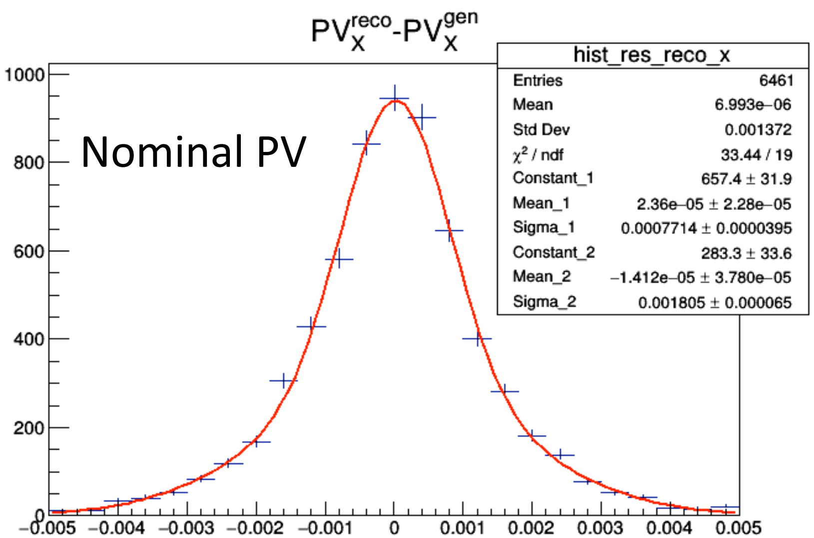


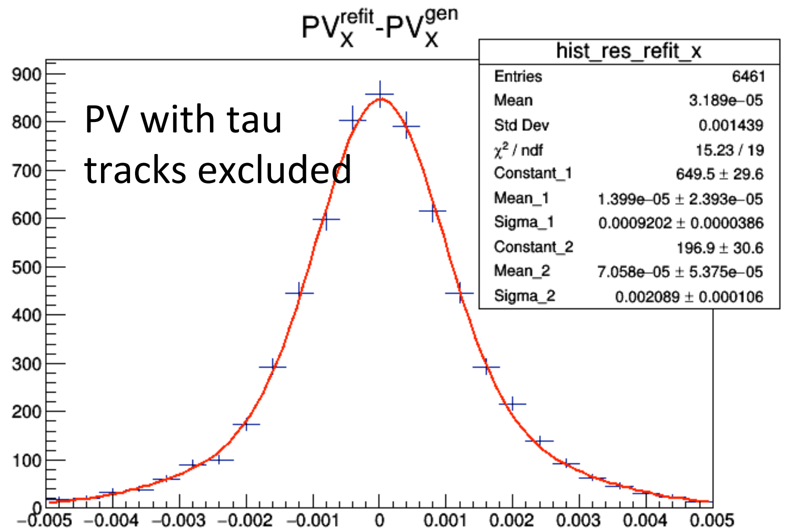
The smearing almost reproduces the effect of the detector resolution

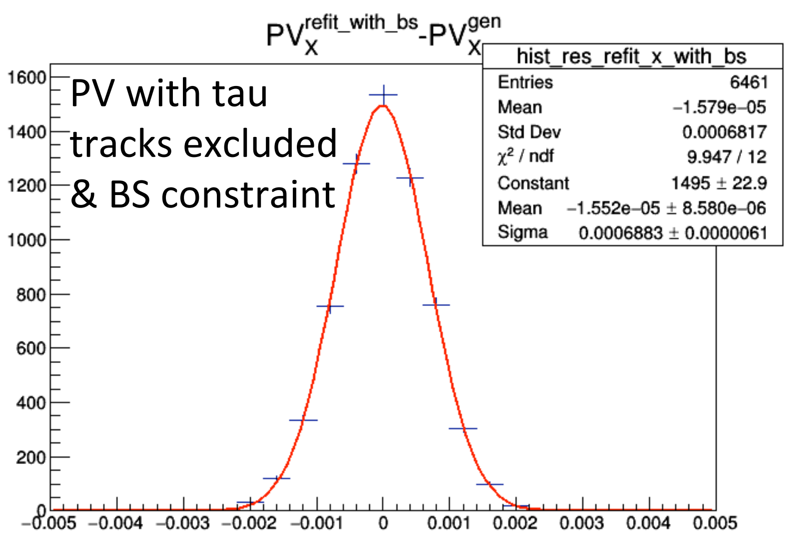
Generator level studies with 2017 MC

- Vertex studies with, removal of tracks belonging to tau, and with beamspot constraint
  - Use miniAOD level vertex refitting, developed by Aachen/M. Bluj
- Studies of acoplanarity angle for $\pi\pi$, $\pi\rho$, $\rho\rho$ channels using IP method.
  - Impact of vertex choice, cut on IP length, and tau pT cut

Reconstructed PV Resolution





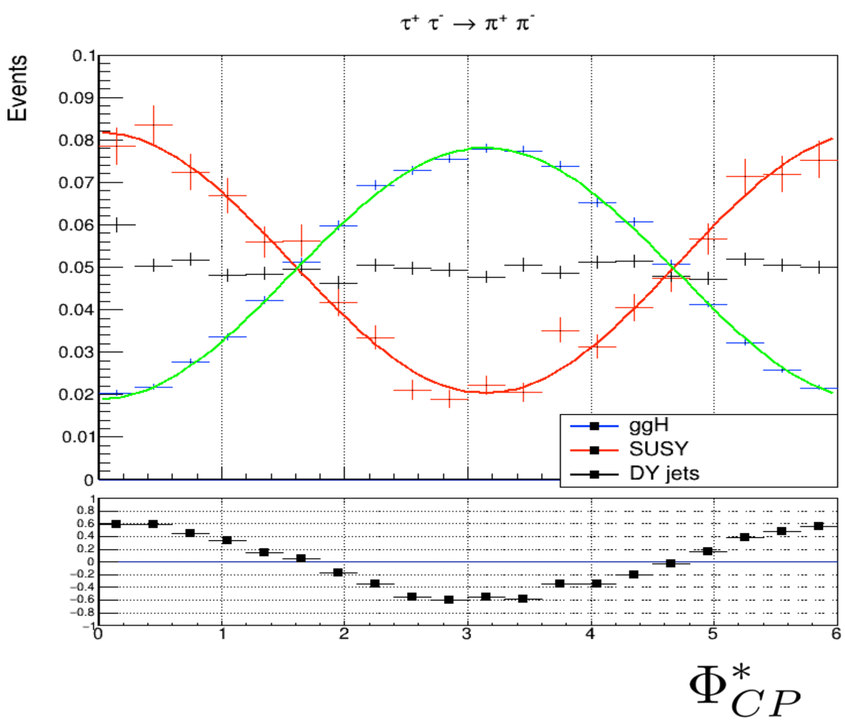


Vertex	σ1	σ2
Nominal PV	7.7×10^{-4}	1.8×10^{-3}
PV with tau tracks excluded	9.2×10^{-4}	2.1×10^{-3}
PV with tau tracks excluded & BS constraint	$\sigma = 6.8 \times 10^{-4}$	

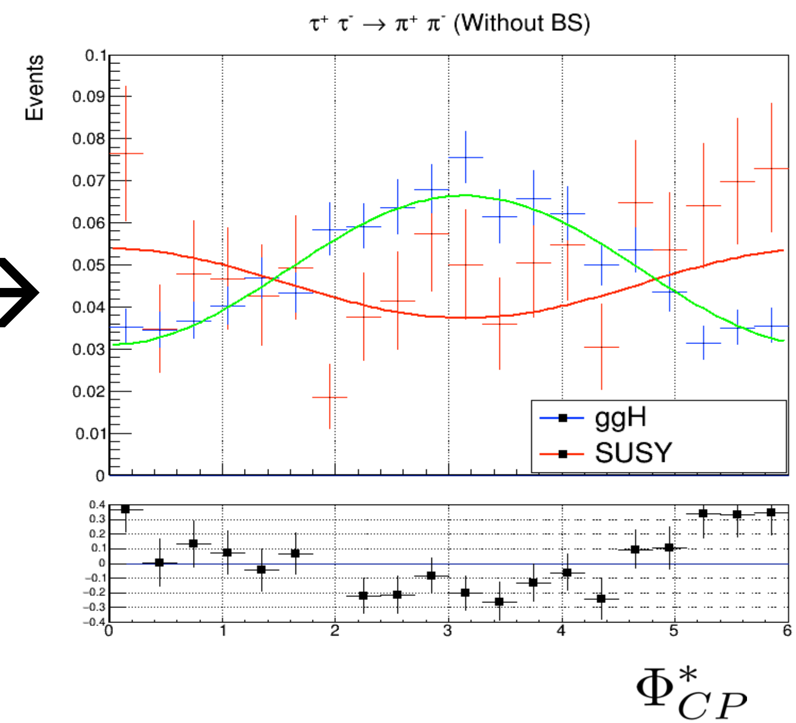
No improvement upon only tau tracks removal but good improvement upon BS constraint

Reconstruction of Acoplanarity angle in $\pi^+\pi^-$ channel

Significant degradation of discrimination due to detector resolution

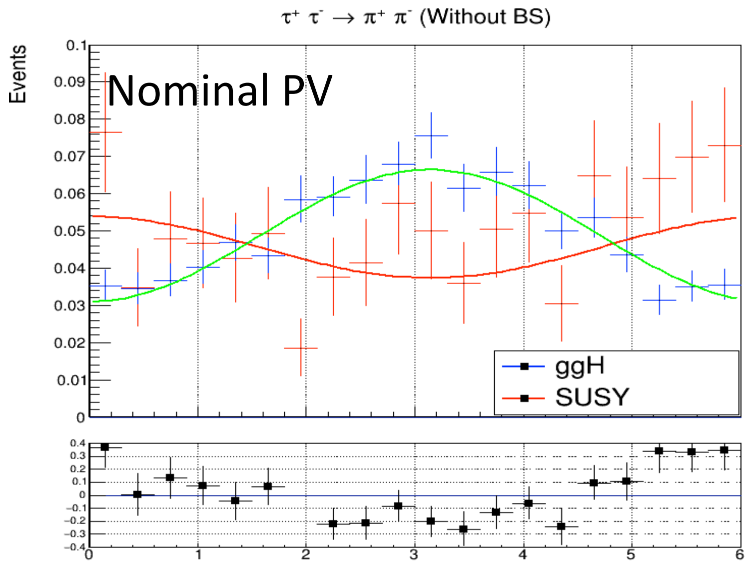


(Generator Level)

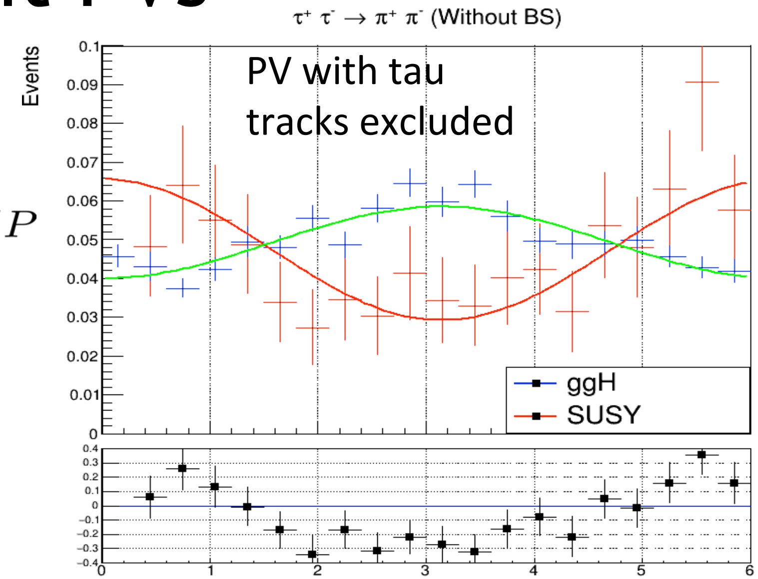


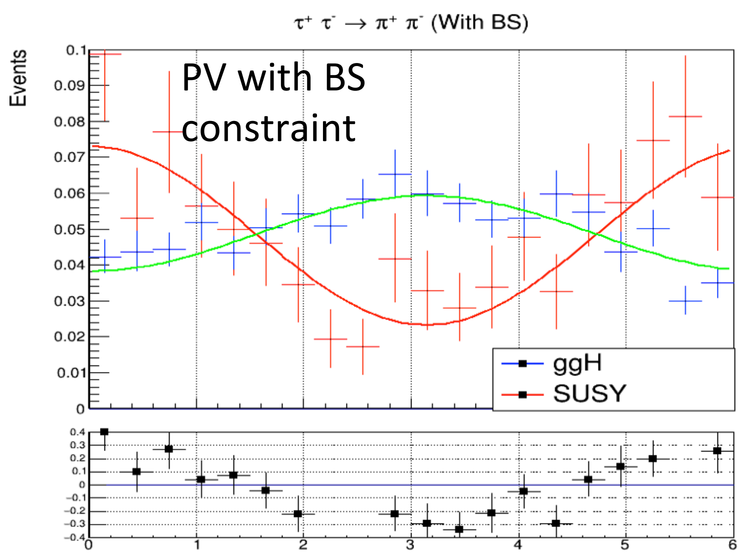
(Reconstruction Level)

IP method in $\pi^+\pi^-$ channel using different PVs



X axis : Φ^*_{CP}





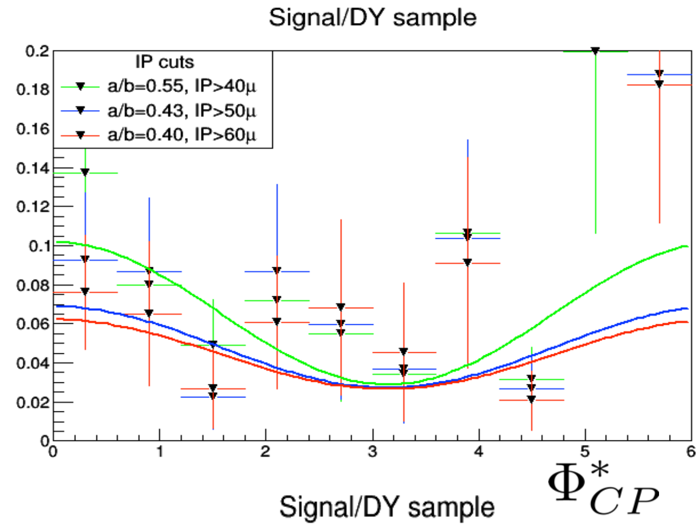
Fitted curve: b + a*cos Φ^*_{CP}

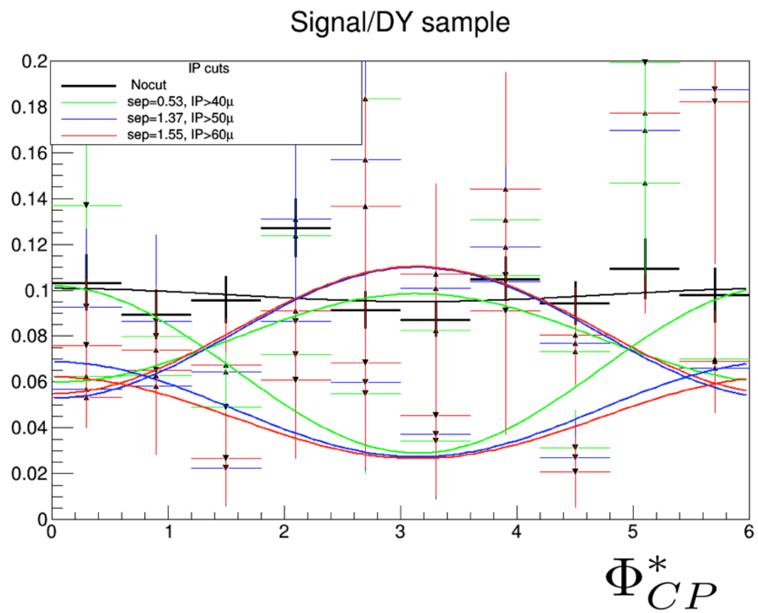
Vertex	a/b (CP even)	a/b (CP odd)
Nominal PV	-0.36	0.18
PV with tau tracks excluded	-0.22	0.38
PV with tau tracks excluded & BS constraint	-0.21	0.51

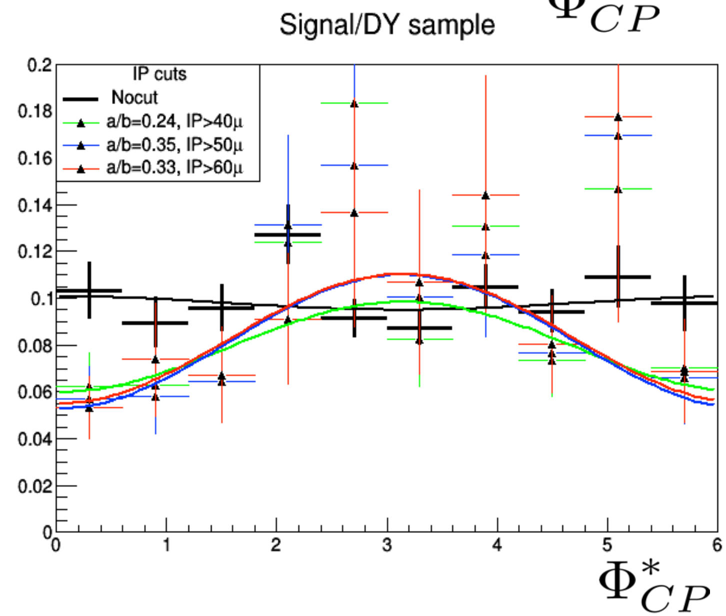
Impact of cut on IP length

Top right: IP cuts of 40,50,60μm on CP odd
Bottom right: IP cuts of 40,50,60 on CP even

IP cut: 40μm , Separation = 0.53
IP cut: 50μm , Separation = 1.37
IP cut: 60μm , Separation = 1.55







Low Statistics. More detailed studies by Aachen group.

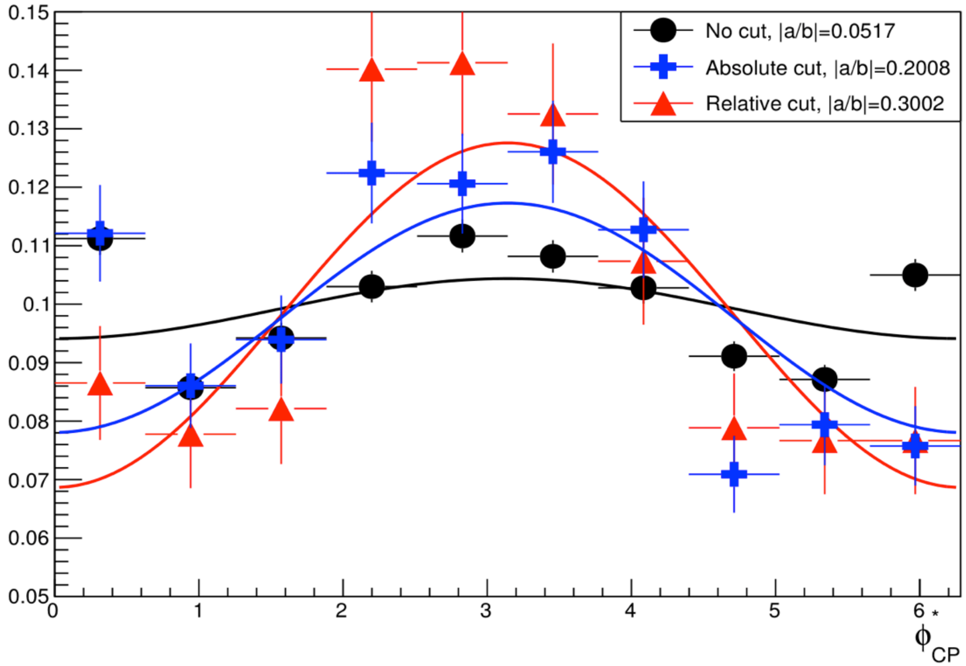
Impact of tau p_T cut in IP method ($\tau\tau \rightarrow \rho\rho$)

In tau+tau → rho+rho, putting cuts on charged pions pT is expected to improve separation between CP even and odd.

Case 1: no cuts on charged pions p_T

Case 2 (Absolute cut): One charged pion p_T>40GeV, second charged pion can have p_T>30 GeV

Case 3 (Relative cut): Both charged pions have p_T > 0.7 * (tau p_T)



For CP even case, there is good improvement in |a/b| ratio

Data/MC and NN Studies

A few words about our analysis framework:

- We share the analysis framework from DESY (helps in sync with the object ID etc.. already)
- However, DESY was not participating in $\tau_h\tau_h$ channel before
- Needed to develop the analysis macros for $\tau_h\tau_h$ channel, some macros are shared between all channels (computation of systematics etc..)
- Have already implemented most of the things: Data/MC studies, multi-class NN framework etc.., working on FF method and Datacard production
  - Good support from Andrea, Merijn and others…

Data & MC samples for 2017

DataSet Name	run range	Luminosity[/fb]
/Tau/Run2017B-31Mar2018-v1/MINIAOD	297046–299329	4.823
/Tau/Run2017C-31Mar2018-v1/MINIAOD	299368–302029	9.664
/Tau/Run2017D-31Mar2018-v1/MINIAOD	302030–303434	4.252
/Tau/Run2017E-31Mar2018-v1/MINIAOD	303824–304797	9.278
/Tau/Run2017F-31Mar2018-v1/MINIAOD	305040–306462	13.540

Dataset Description	Dataset Name(2017)	Cross-Section[pb]
$t\bar{t}$	/TT TuneCUETP8M1 13TeV-powheg-pythia8	831.76
$Z \rightarrow \tau\tau$	/DYJetsToLL M-50 TuneCUETP8M1 13TeV	5765.4
W+jets	/WJetsToLNu TuneCUETP8M1 13TeV-pythia8	61526.7

Stitched Njet sample

Trigger Info

HLT Path	L1 seed	Tau filter to match
HLT DoubleMediumIso PFTau35 Trk1	L1 DoubleIsoTau	hltDoublePFTau35TrackPt1MediumIsolationDz02
HLT DoubleTightIso PFTau40 Trk1	L1 DoubleIsoTau	hltDoublePFTau40TrackPt1TightChargedIsolationDz02Reg
HLTDoubleMedium CombinedIso PFTau35	L1 DoubleIsoTau	hltDoublePFTau35TrackPt1MediumCombinedDz02

Event selection for $H \rightarrow \tau_h \tau_h$

The method and strategy taken from 2017 Higgs to ditau Analysis (HIG-18-032).

Requirements on $\tau_h \tau_h$ pair:

- Opposite charges
- ΔR > 0.5
- Each tau passes:
  - Tight iso
  - pT > 40, |η|<2.1
  - byDecayModeFinding discriminator.
  - d_z < 0.2 cm
  - match to trigger objects within ΔR < 0.5

Lepton vetos:

Require exactly 0 muons with:

- pT > 10 , |η| < 2.4
- passing medium muon id
- iso < 0.3 times muon pT

Require exactly 0 electrons with:

- pT > 10 , |η| < 2.5
- passing the 90% efficiency working point of electron ID MVA
- iso < 0.3 times electron pT

Jet selection:

- AK4, PFJet
- pT > 30
- |η| < 4.7

MET:

PF MET with Type-1 and recoil corrections

- Performed b-tagging using DeepCSV algorithm using pfDeepCSVJetTags:probb and pfDeepCSVJetTags:probbb discriminator.
- Higgs pT > 50 to suppress $Z \rightarrow \tau_h \tau_h$ events

Background Estimation

- W+jets, ttbar, DY estimated from MC
- QCD estimated from data using ABCD method
  - Failed iso region is when at least one tau passes the VLoose isolation but not the tight isolation.

SS	B	C
OS	A	D
	pass tau-iso	fail tau-iso

$$\text{DATA(OS)}_A = \frac{D}{C} \times (\text{Data(SS)} - \text{MC(SS)})_B$$

Working on implementation of FF method, should be ready asap.

Weights / Scale Factors applied so far

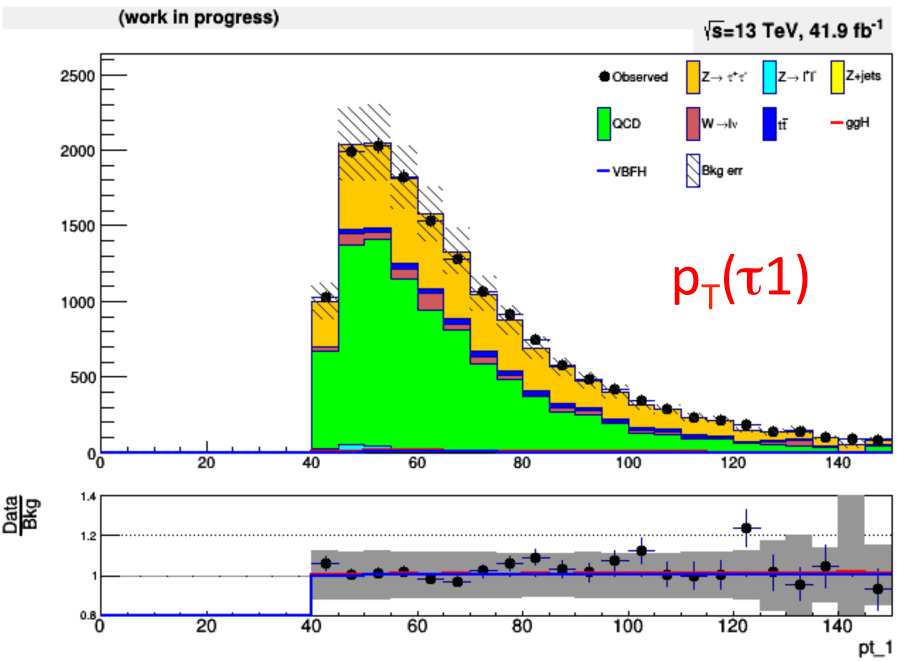
1. MC x-section weight
2. Pileup weight
3. Tau ID scale factor
4. Tau Trigger Scale Factor
5. B-Tagging Scale factor

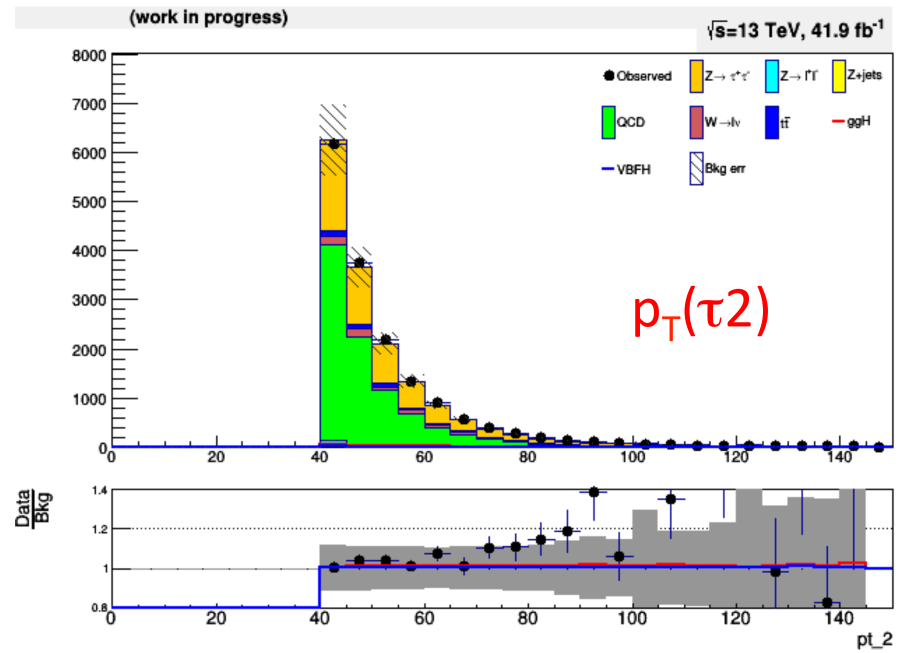
Uncertainties

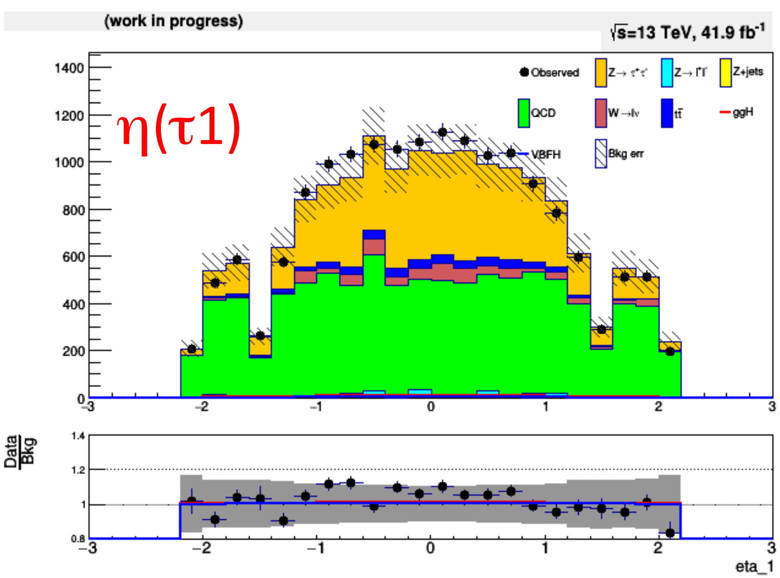
- Cross sectional uncertainities:
  - ZtoTauTau : 6%
  - W+jets : 10%
  - TTbar : 7%
  - QCD : 15%
- Tau ID scale factor : 10%
- Luminosity : 3%
- Statistical uncertainties

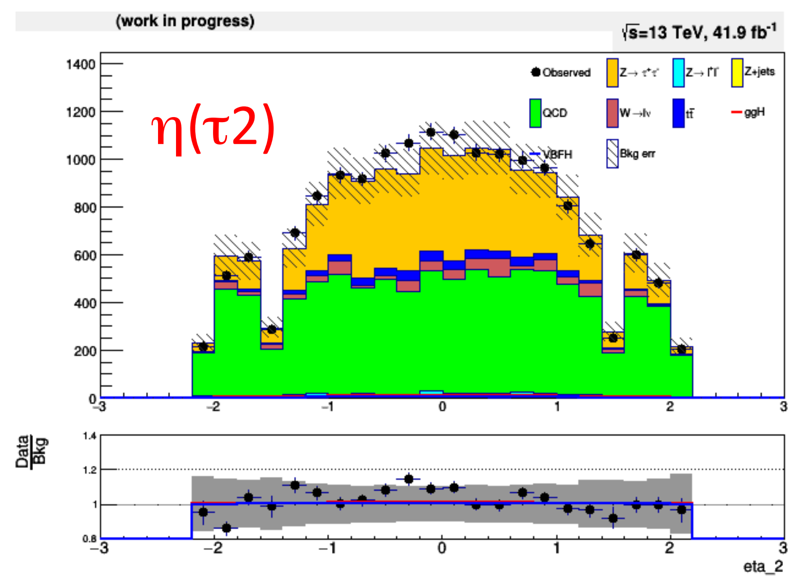
JES, b-Tag scale, tau E-Scale uncertainties are not yet included

Tau Kinematics

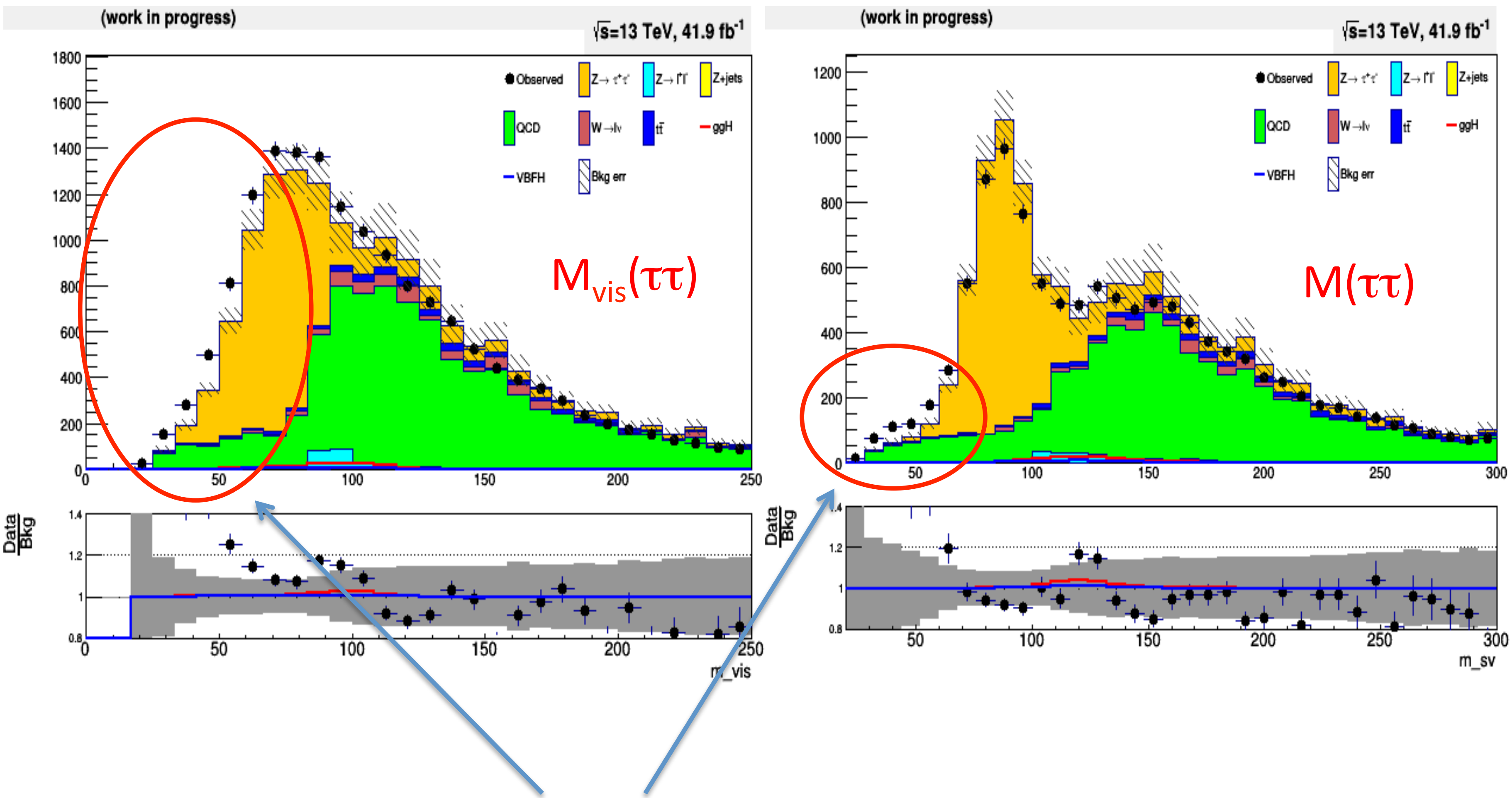






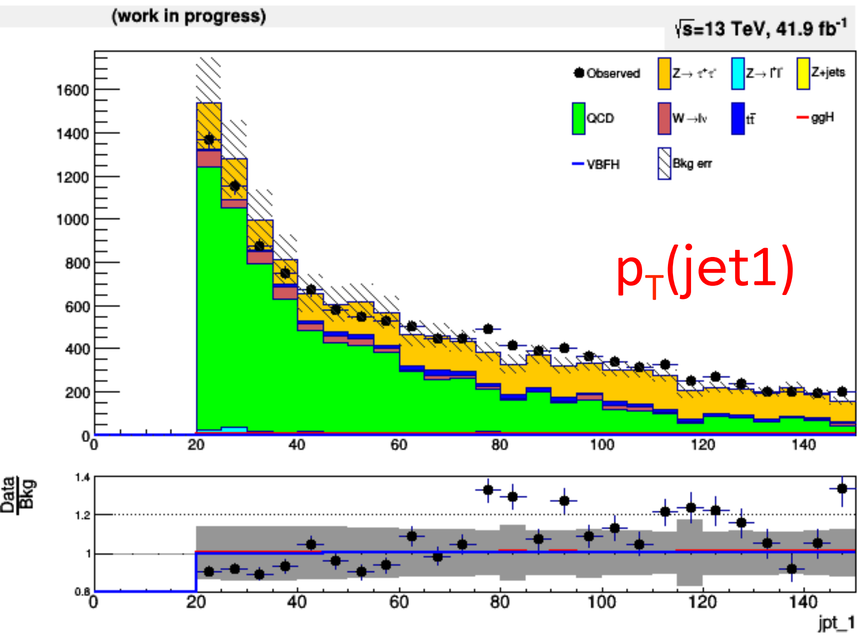


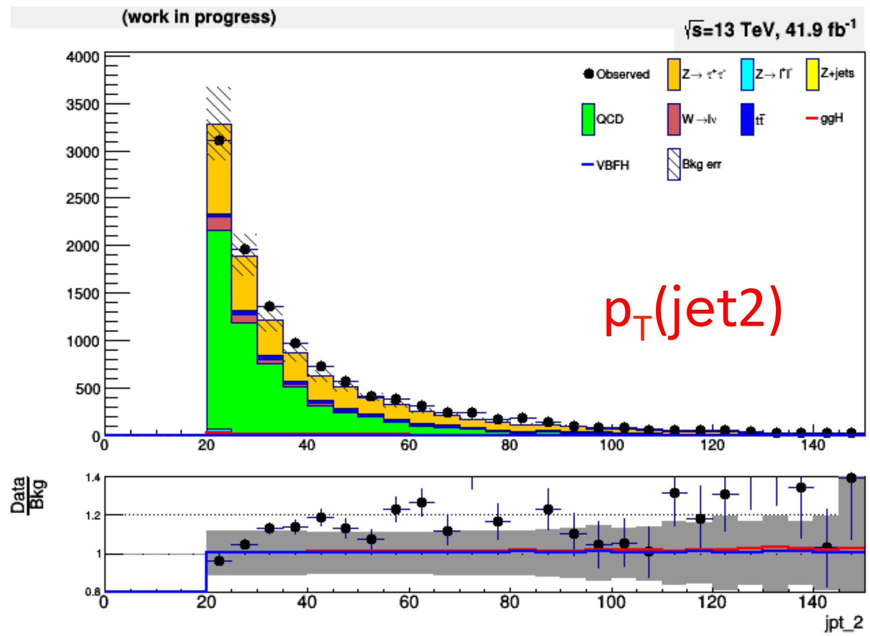
di-Tau Mass

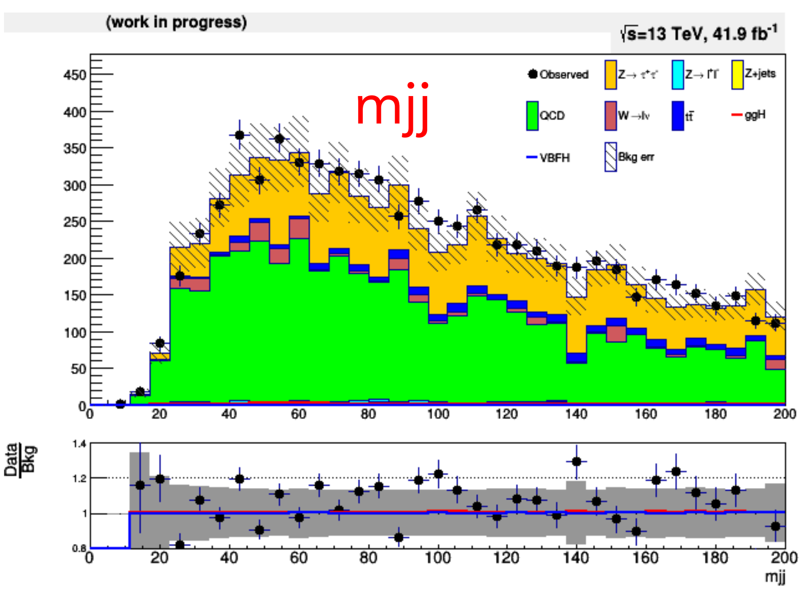


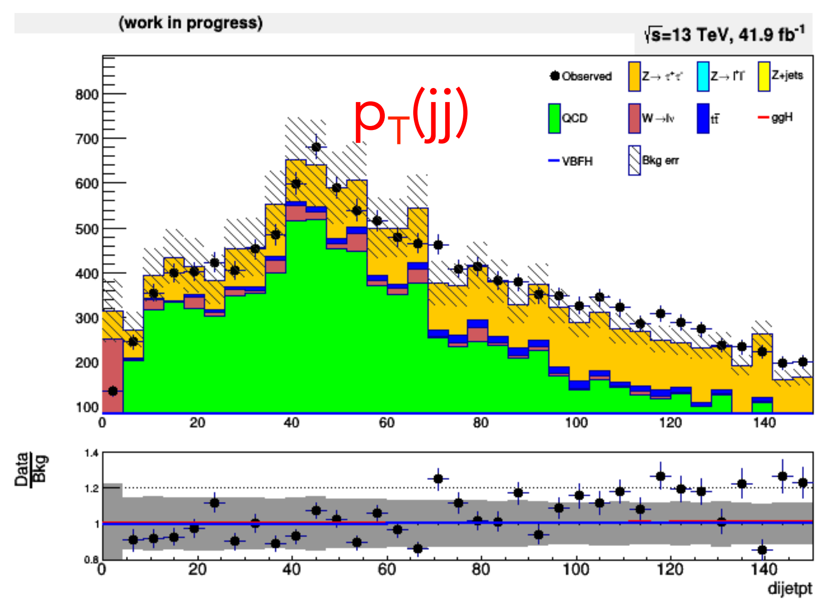
Could be due to the absence of low mass DY

Jet Distributions



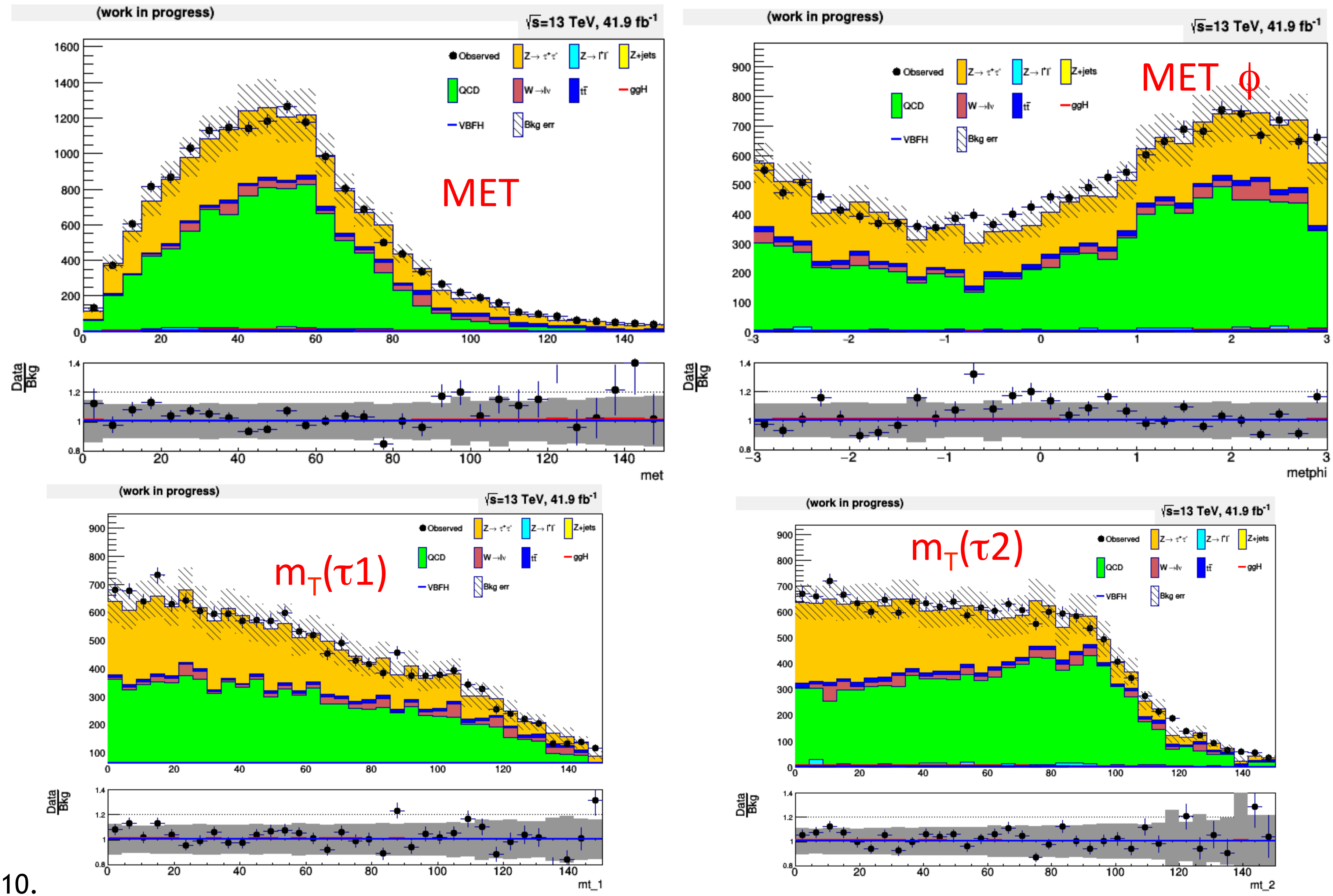




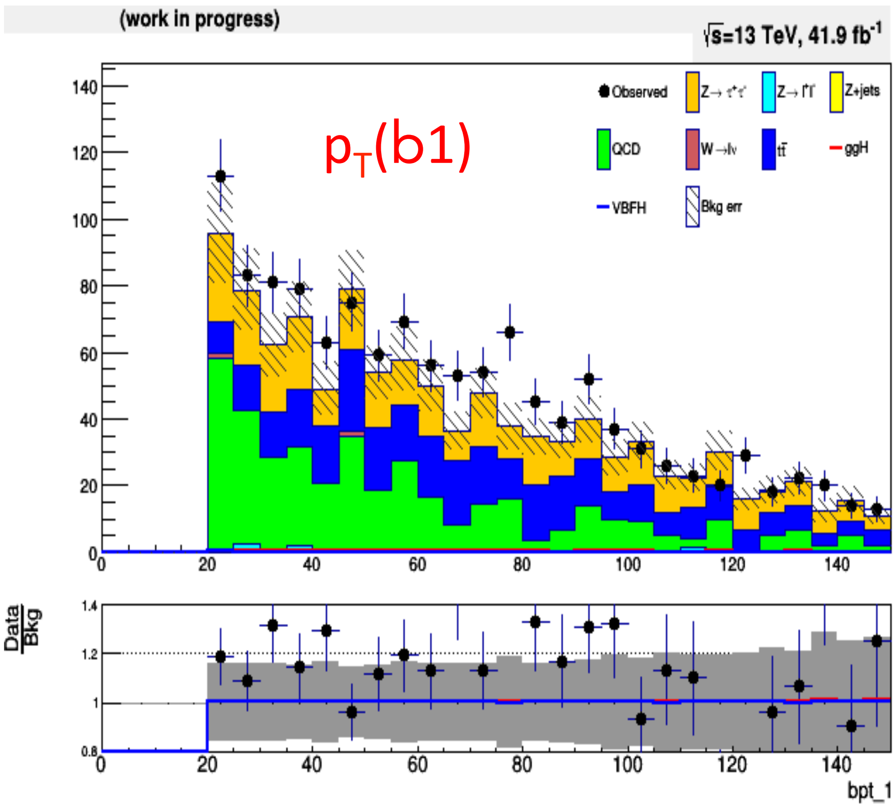


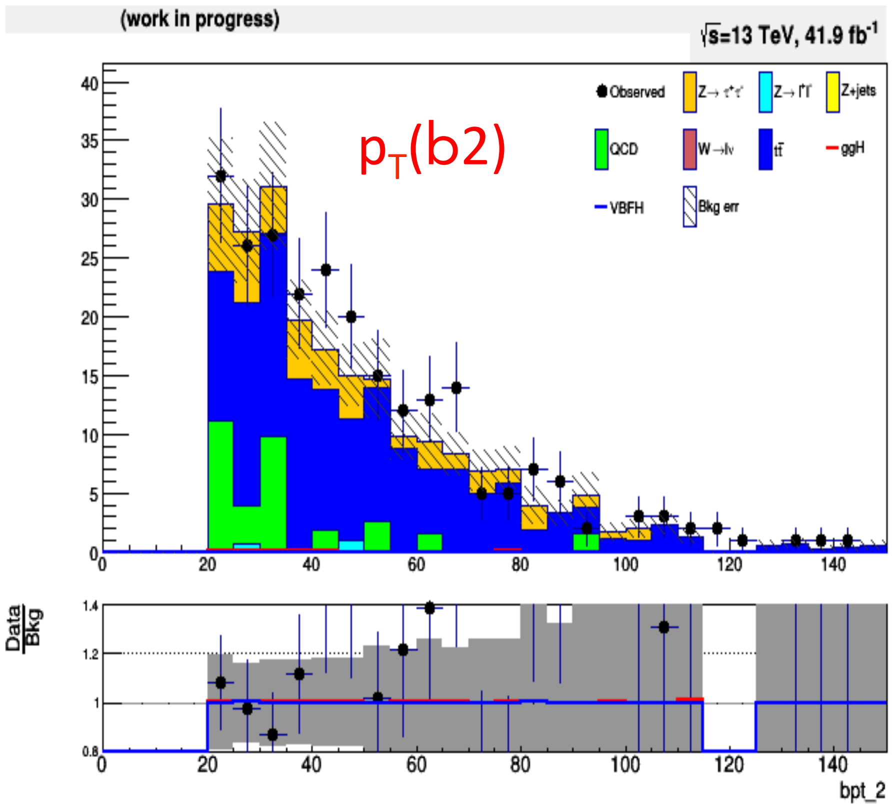
10.

MET distributions



b-Tagged Jets





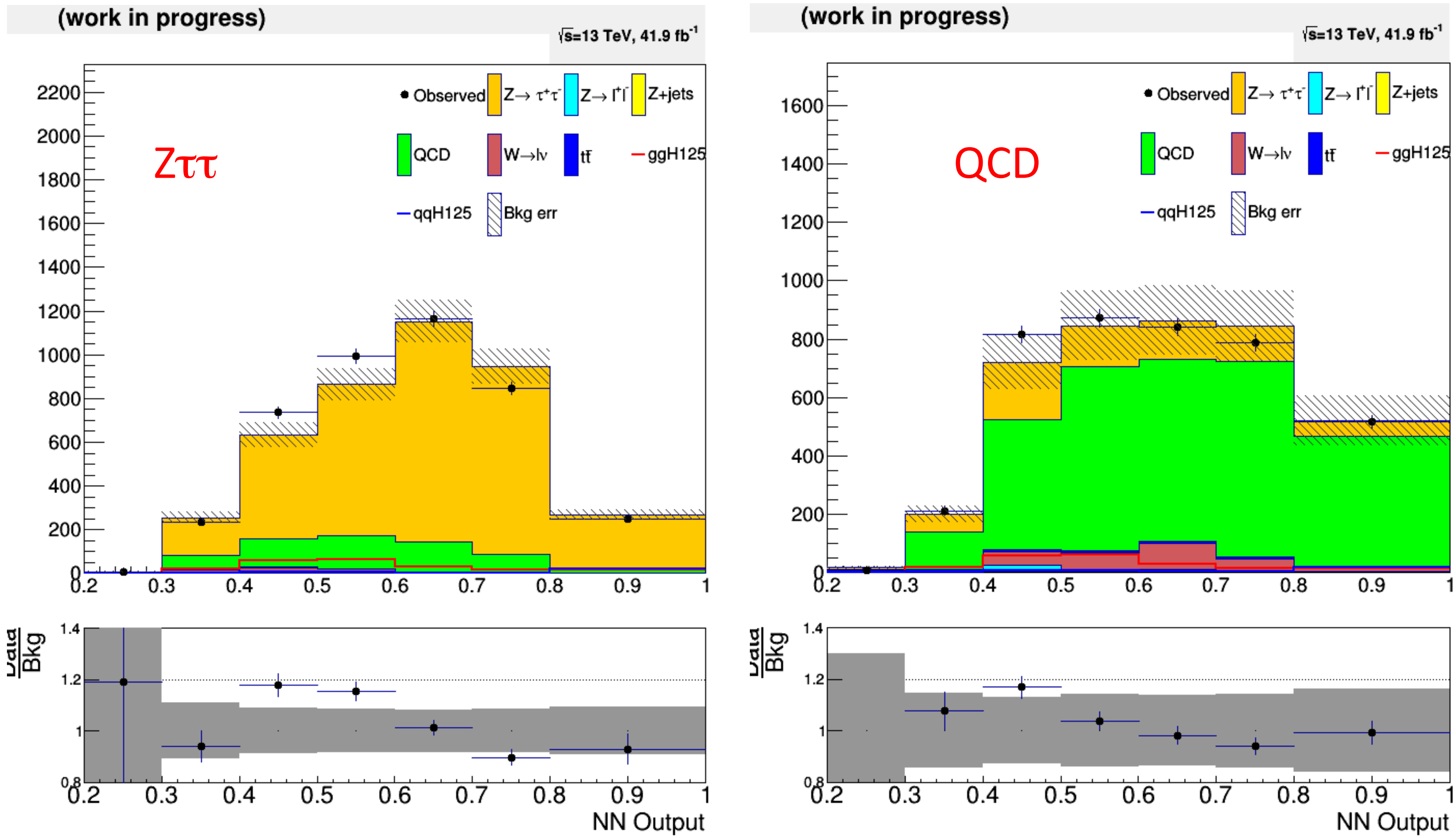
NN Studies

Implemented the multi-class Neural Network, as developed by SM $H \rightarrow \tau\tau$ analysis [HIG-18-032].
The following variables are used for $\tau_h\tau_h$ category, for 2017 data:

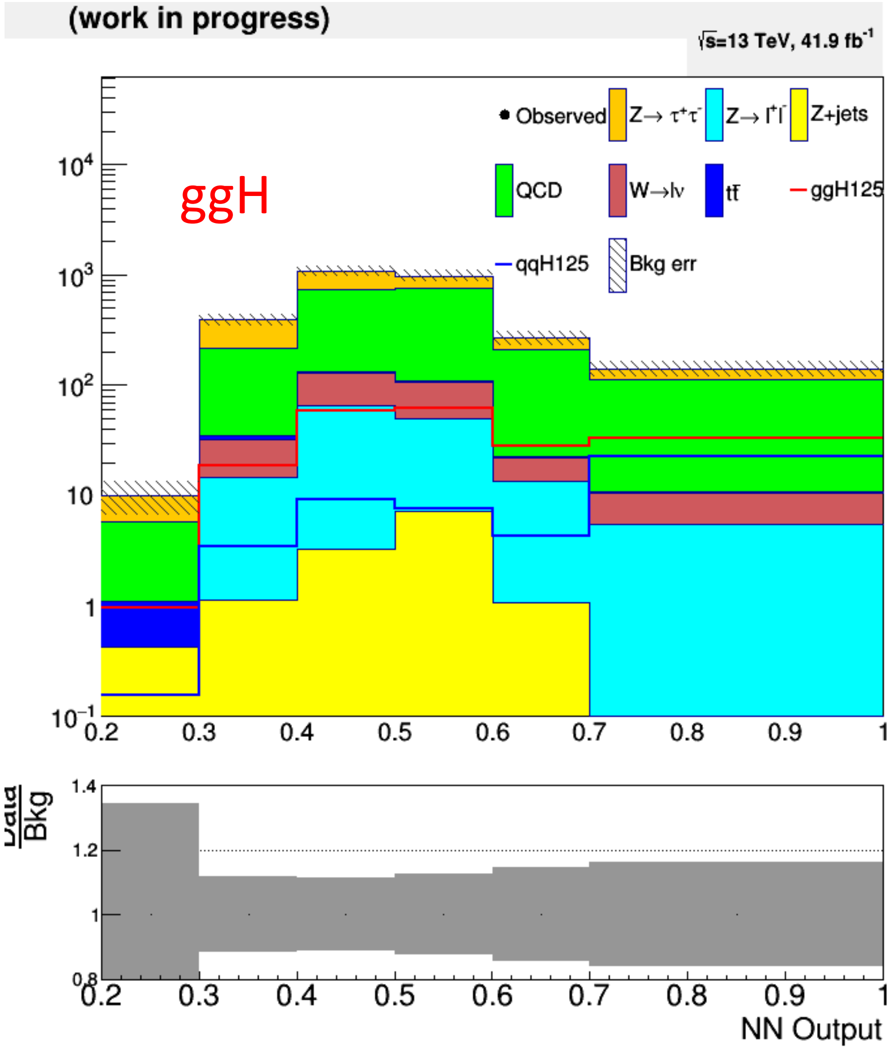
$p_T(\tau 1)$, $m_T(\tau 1)$, $m_T(\tau 2)$,
$p_T(jet2)$, $p_T(b1)$, $p_T(b2)$
N-jets, N-b-jets, $m_{\tau\tau}^{vis}$,
$m_{\tau\tau}^{svFit}$, $p_T^{svFit}(\tau\tau)$, m_{jj},
$\Delta\eta(jj)$, $p_T(jj)$

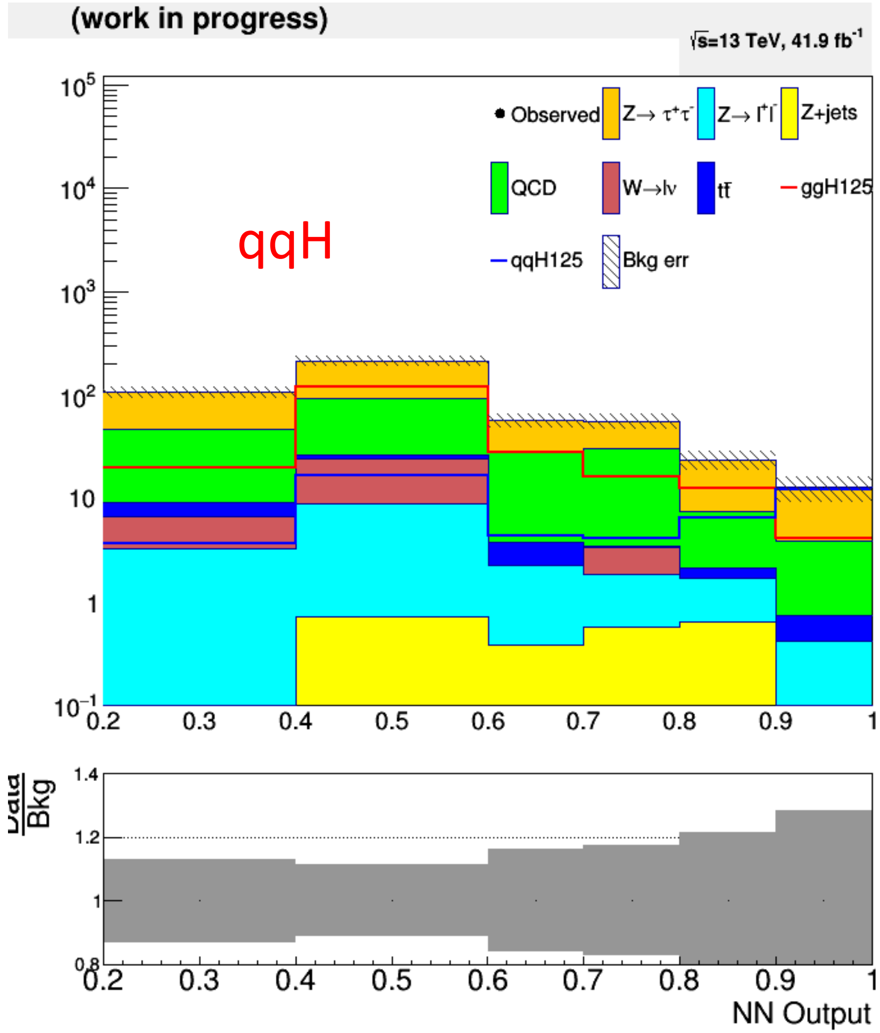
Output Categories: ggH, qqH (VBF), $Z\tau\tau$, QCD, misc.

NN Background categories



NN Signal categories

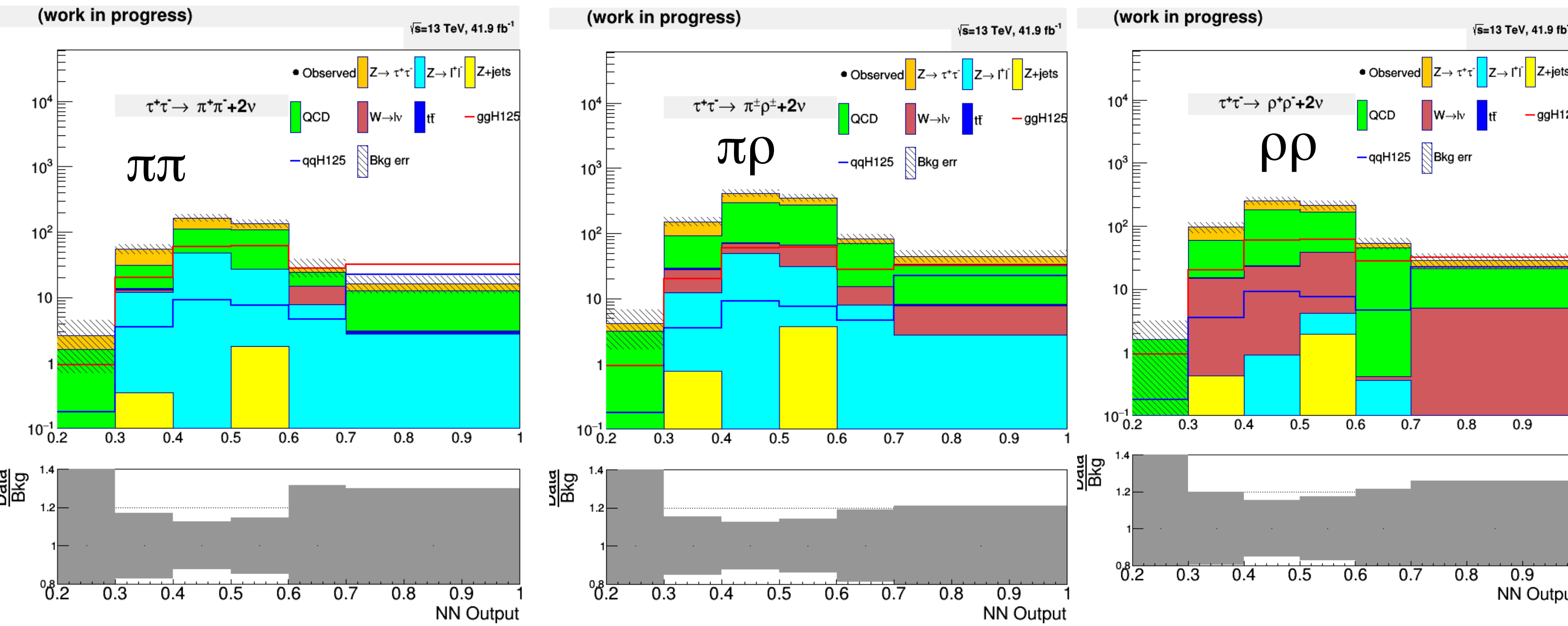




Plots are blinded in these signal categories

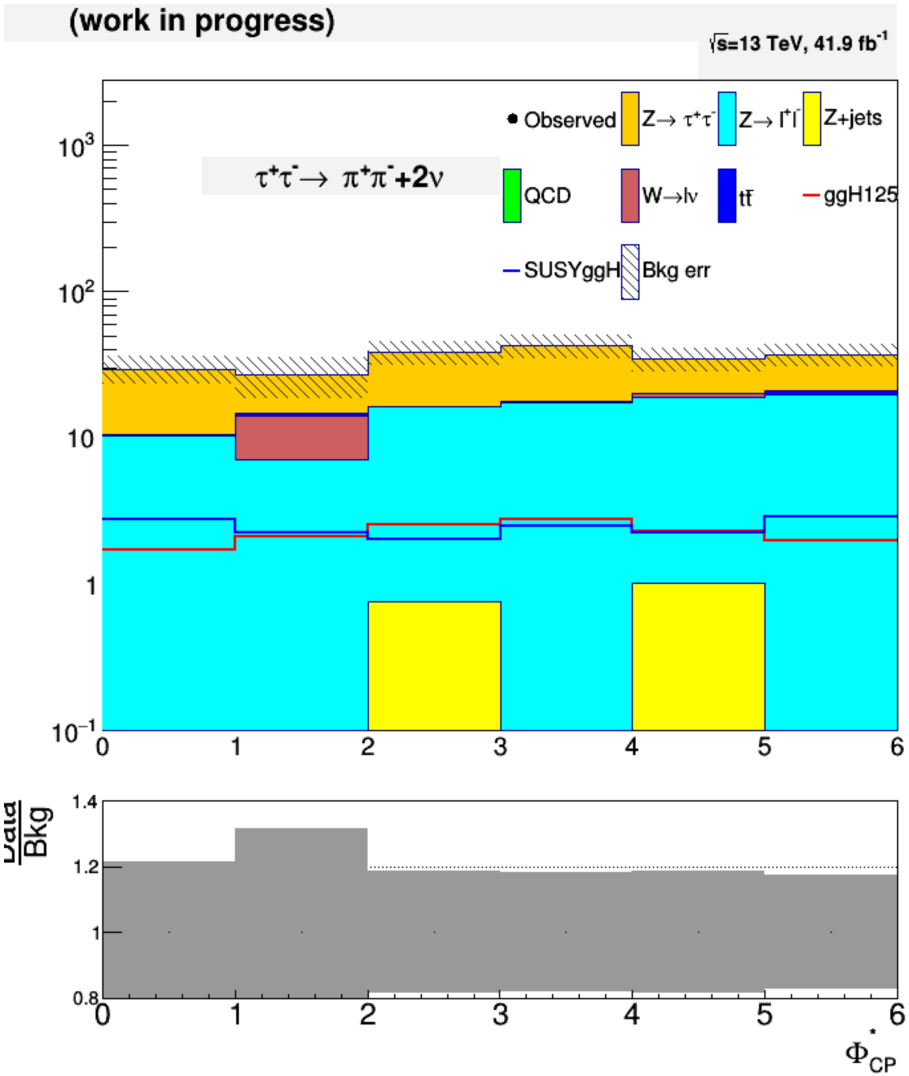
10. 10. 2019

Signal Categories in separate tau decay modes



The shape in all three modes look similar, though the bkg content are different

Acoplanarity angle for $\pi^+\pi^-$ channel, in ggH category



No cut on NN output value, because of low stat.

Just show the ϕ^* distribution in ggH category, for $\pi^+\pi^-$ channel.

Summary & Plans

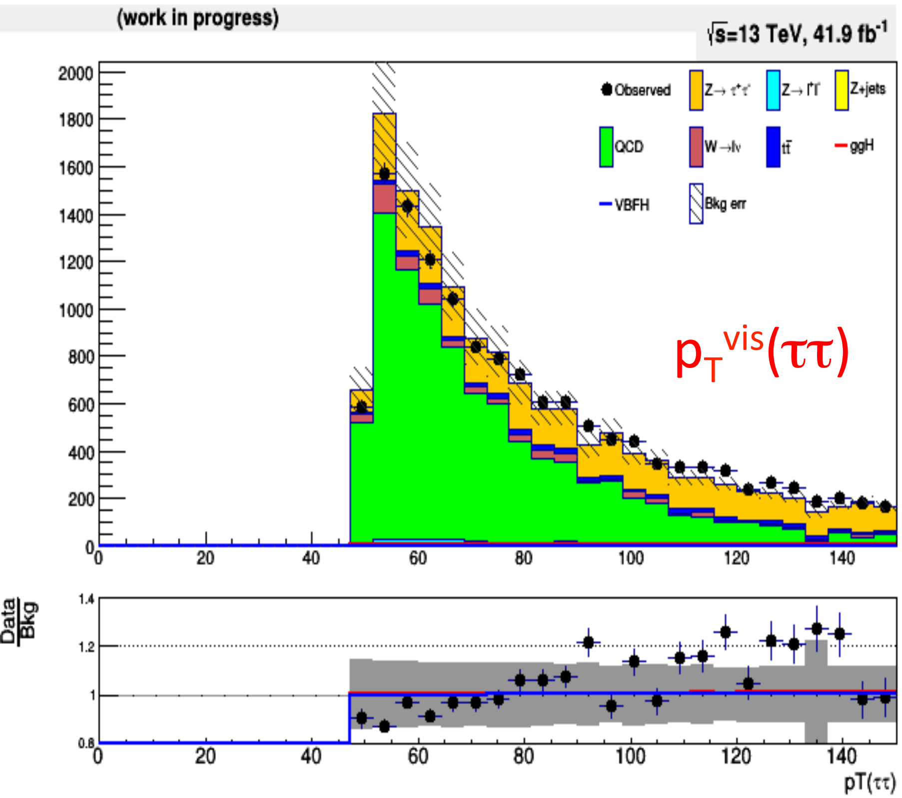
- Performed studies on vertex resolution and reconstruction of acoplanarity angle using IP method.
  - Some indications about possible improvement
- Performed Data/MC studies with 2017 data, and implemented multi-class NN framework developed in HTT analysis.
  - Data/MC agreement looks good, still working on FF method.
  - Framework should be ready asap to produce datacards
  - Will look at 2016/2018 data, ntuples are being produced by DESY group.
  - Also include MVA decay mode, IP using helix propagation (Ingredients are already in the ntuple)

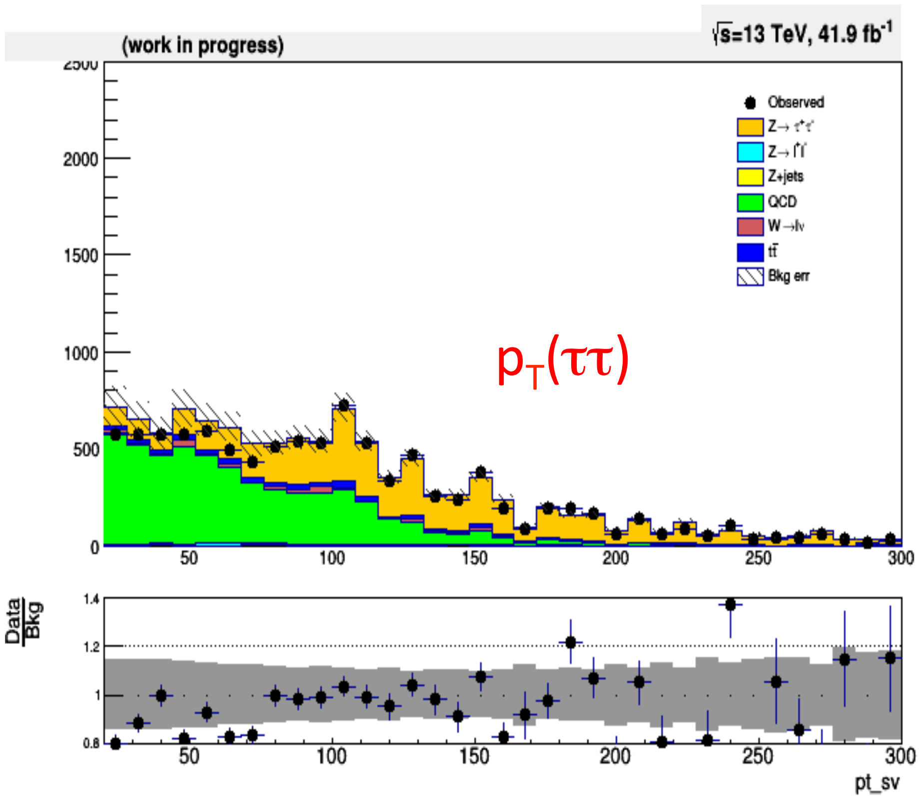
Other Plans

- Plan to perform some further studies on improving discriminating power in IP method
  - Optimizing cuts on IP/IP-significance
  - Optimizing cuts on charged pion (for $\tau \rightarrow \rho\nu$ decay mode)
  - …..
- Not sure whether we can have manpower to participate in MVA development
  - We can contribute to validation
  - We can try to study MVA for $\pi\pi$, $\pi\rho$ categories (if we plan to study MVA separately for different DM categories)

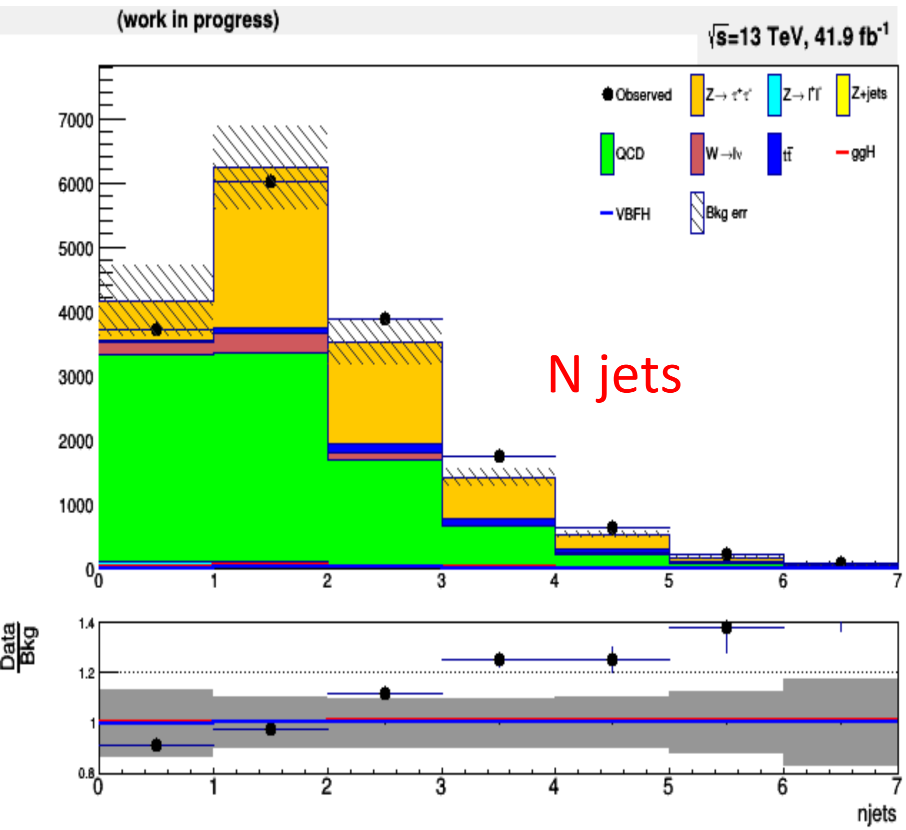
Thank You

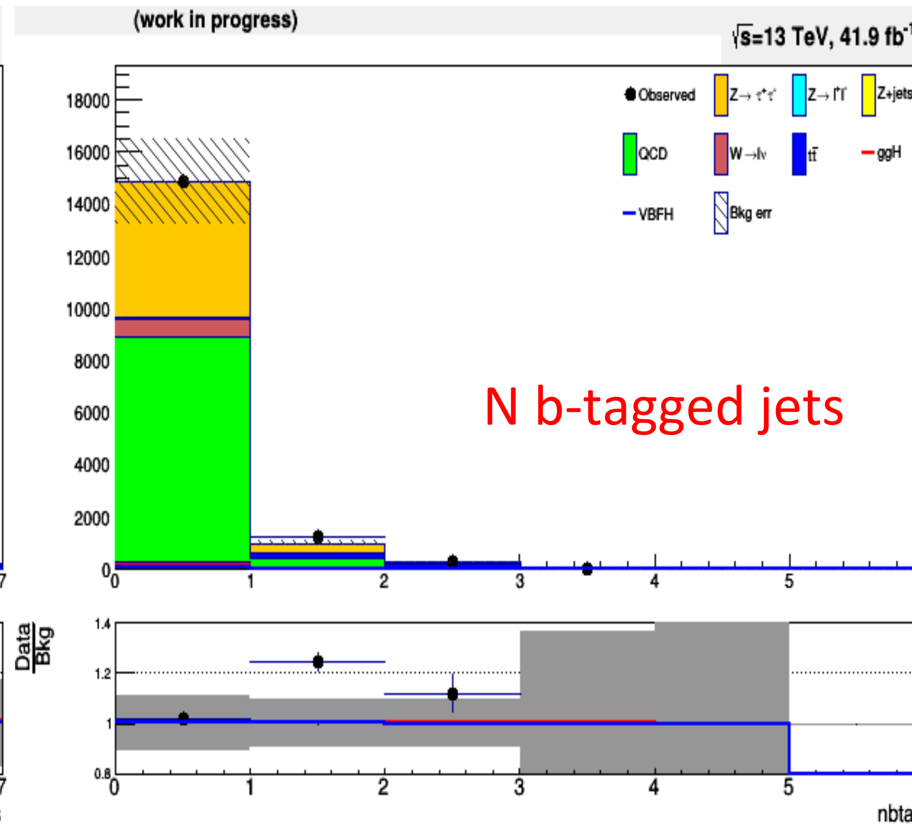
Di-Tau Distributions



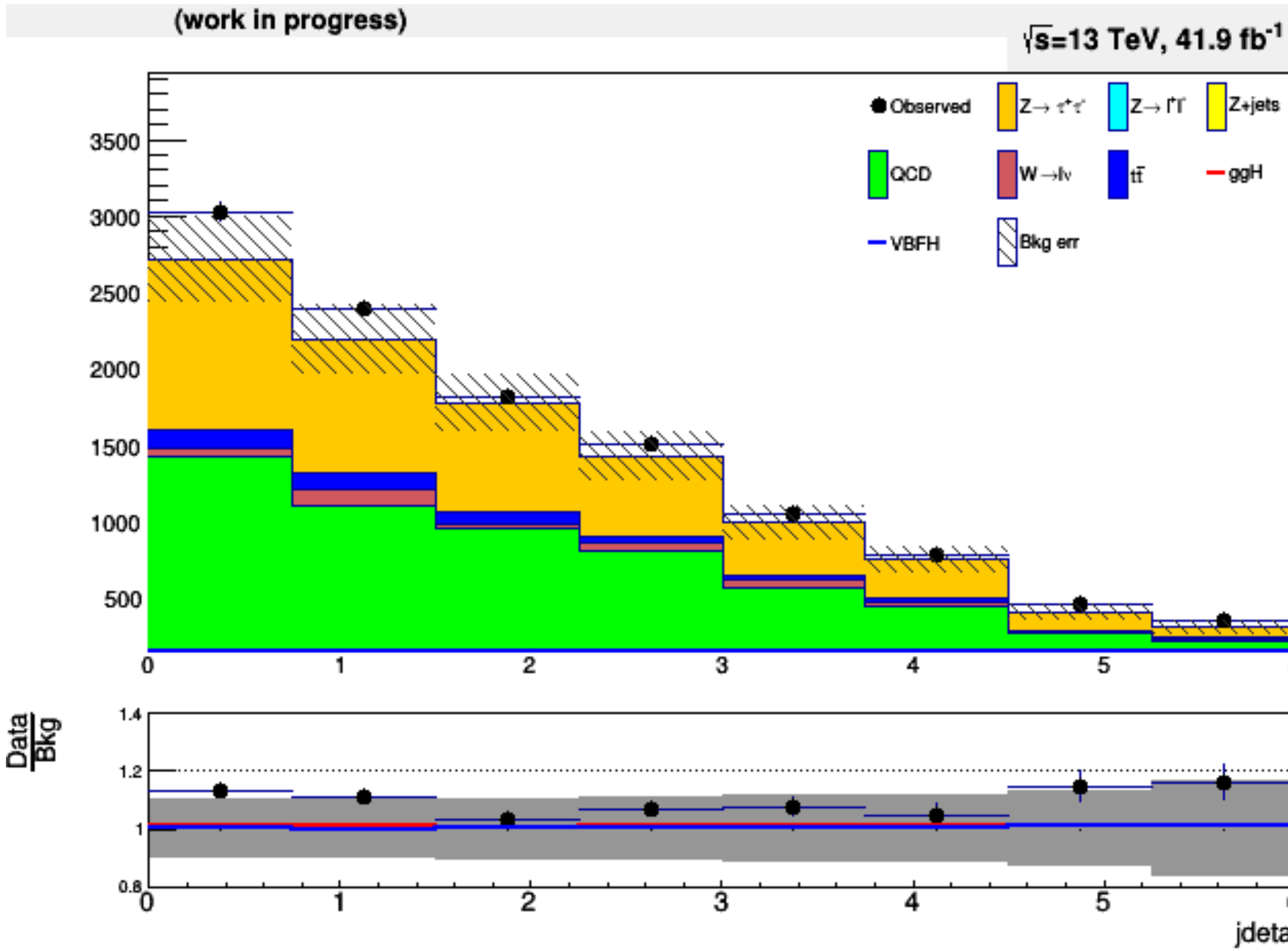


Jet Multiplicity

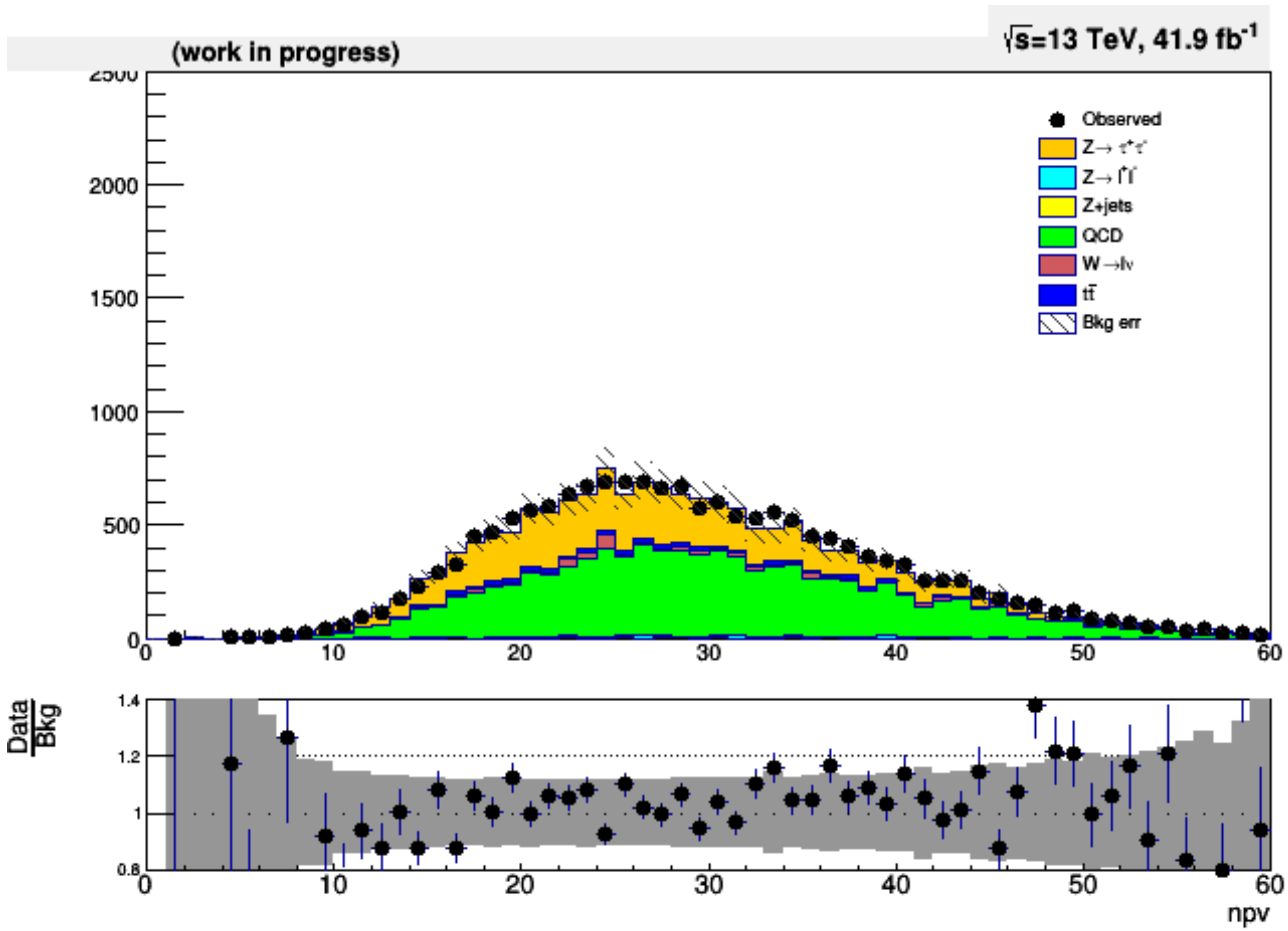




Δη(jet1, jet2)


(work in progress)
√s=13 TeV, 41.9 fb-1
Observed
Z→ τ+τ-
Z→ l+l-
Z+jets
QCD
W→lν
tt̄
ggH
VBFH
Bkg err
3500
3000
2500
2000
1500
1000
500
0
1
2
3
4
5
6
Data/Bkg
1.4
1.2
1
0.8
jdeta

Pileup Modeling



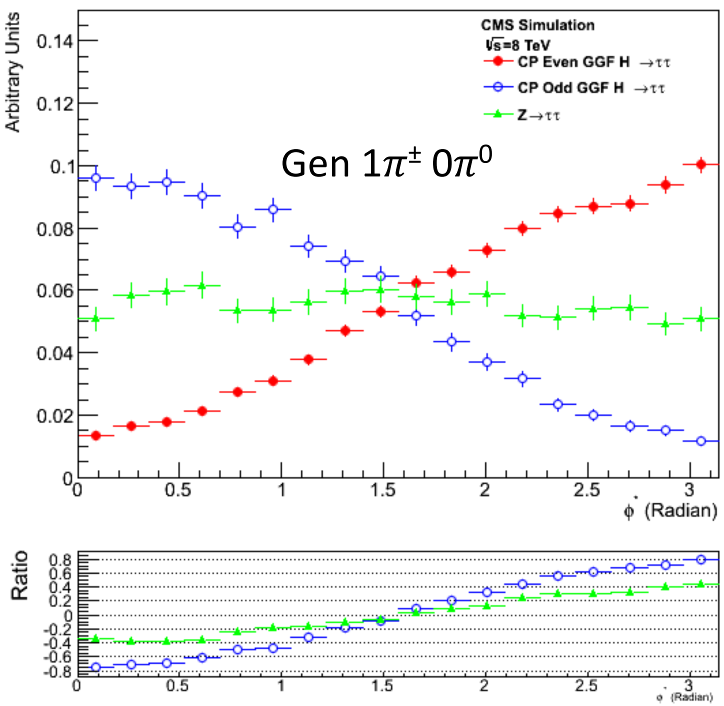
NN Architecture for tt-chanel

Number of hidden layers	2
Number of nodes per hidden layer	200
Activation functions of the hidden layers	Hyperbolic tangent
Regularization	Dropout (probability = 30%), L2 (10^{-5})
Optimizer algorithm	Adam (learning rate = 10^{-4})
Batch size	100
Early stopping	50 epochs
Validation split	25%
Weight initialization	Glorot (uniform)

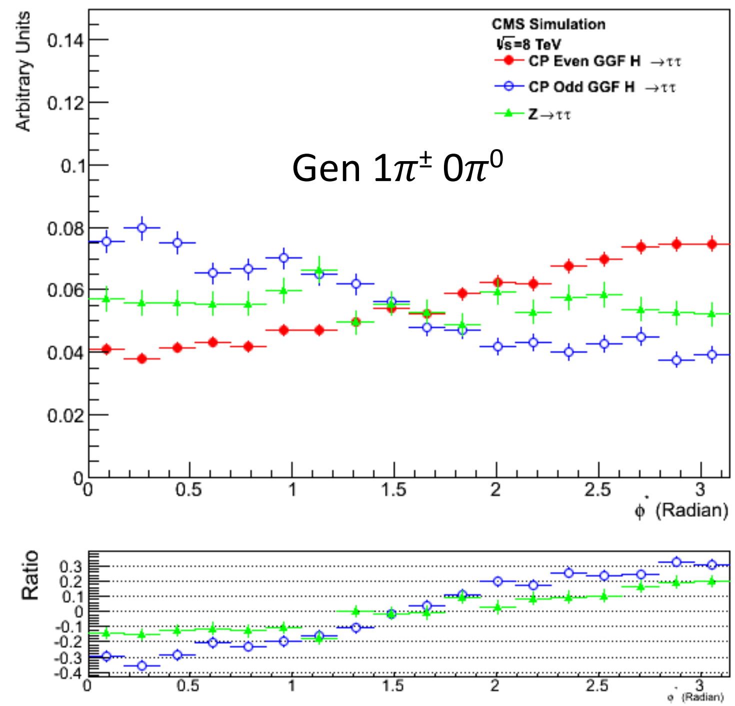
PCA Smearing

Smear PCA with Gaussian
σ_x = 20 µm, σ_y = 20 µm, σ_z = 20 µm

No Smearing



With PCA Smearing



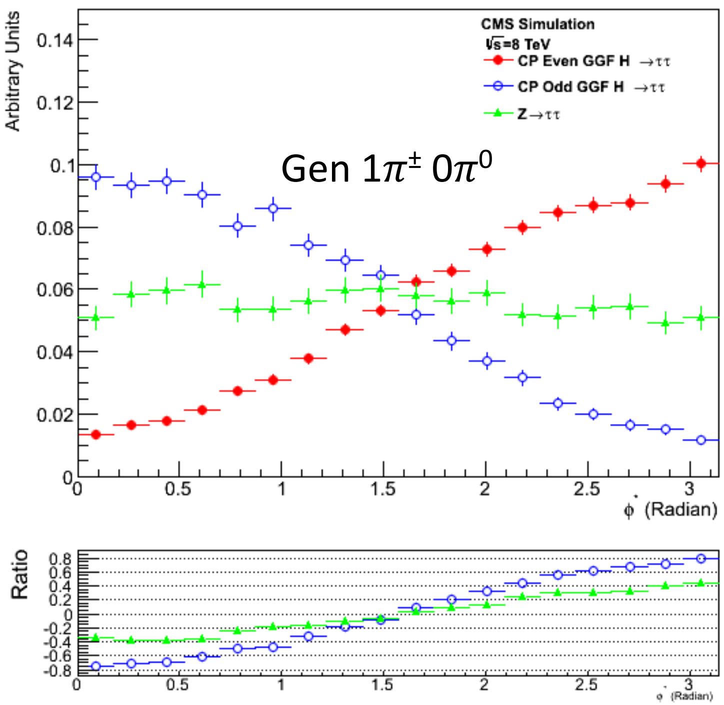
PCA smearing reduces the discriminating power

Vertex Smearing

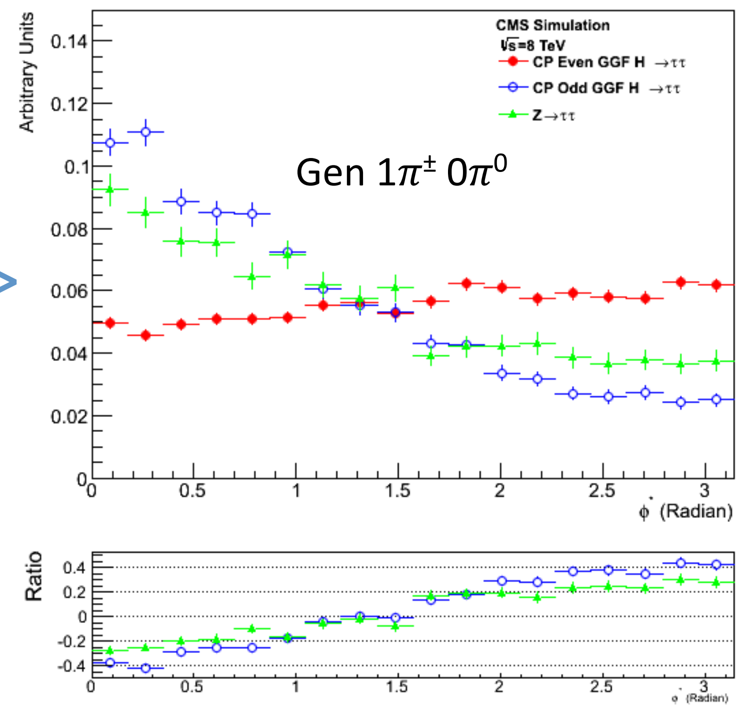
Smear Vertex position with Gaussian

σ_x = 10 μm, σ_y = 10 μm, σ_z = 30 μm

No Smearing



With Vertex Smearing



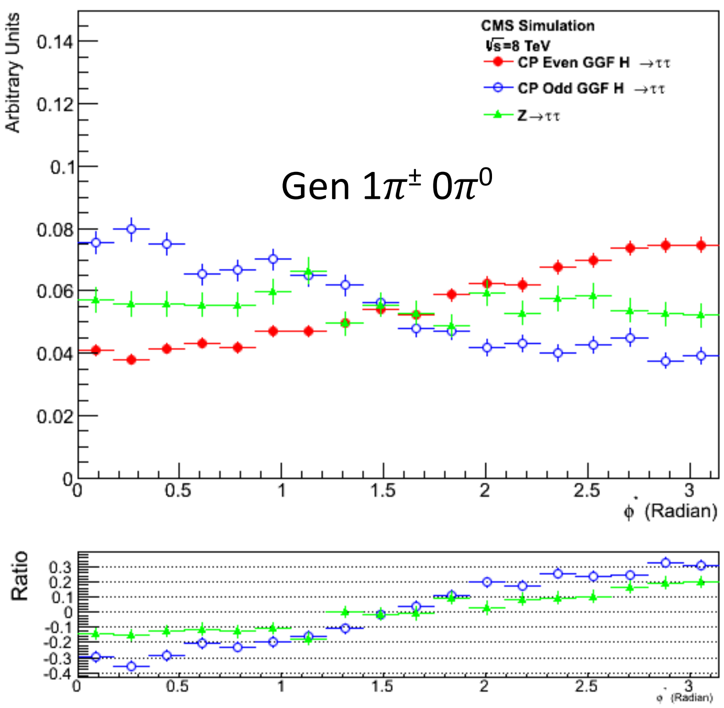
➡ Vertex smearing reduces the discriminating power and the distribution of Z➔$\tau\tau$ becomes more closer to that of CP-odd boson

Vertex & PCA Smearing

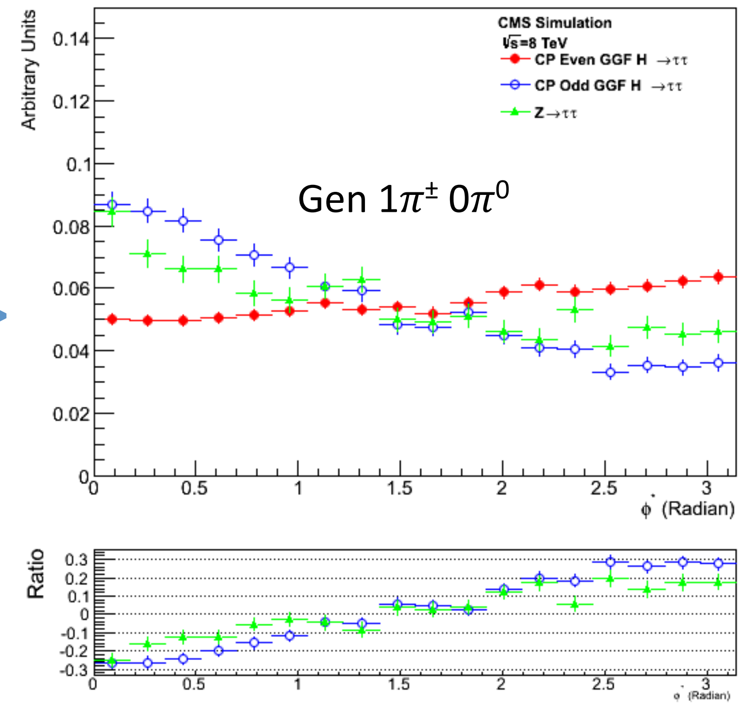
Vertex position smeared with Gaussian
σ_x = 10 µm, σ_y = 10 µm, σ_z = 30 µm

PCA smeared with Gaussian
σ_x = 20 µm, σ_y = 20 µm, σ_z = 20 µm

With only PCA Smearing



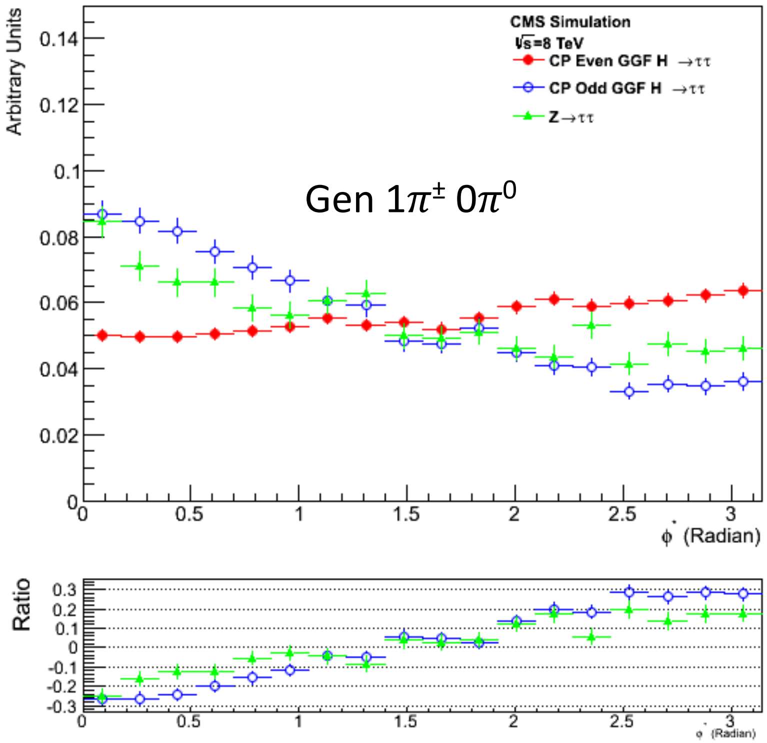
With both vertex and PCA Smearing



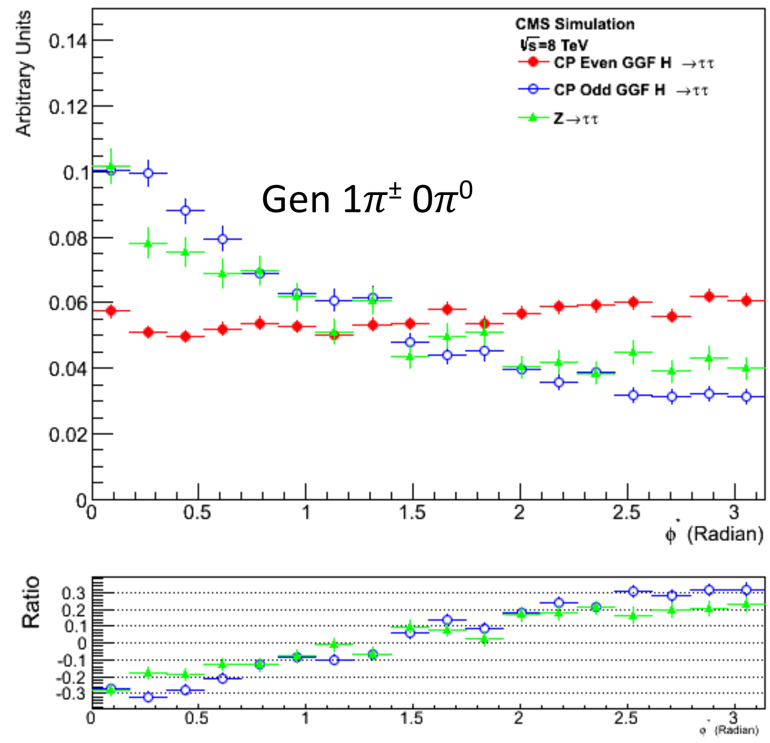
- The PCA resolution is the leading source of degradation
- The Vertex resolution does not cause further loss of discrimination power between CP-odd and CP-even, but changes the shape of all the distribution and makes Z$\rightarrow\tau\tau$ and CP-odd similar

Gen. Smearing Vs Detector Resolution

Generated distribution
With both vertex and PCA Smearing



Reconstructed distribution excluding effect of mis-identified taus

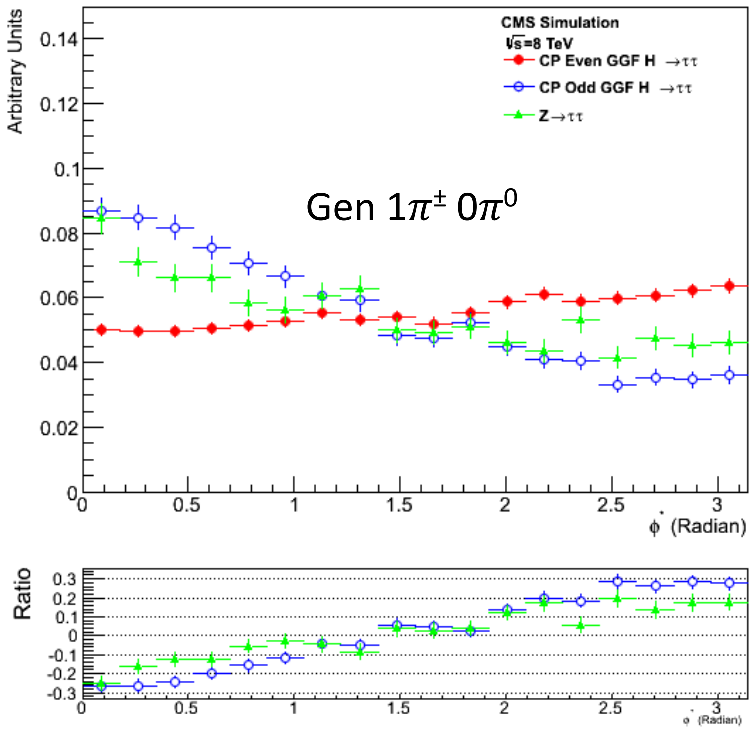


✓ Smearing of vertex and PCA together re-produces the effect of detector resolution

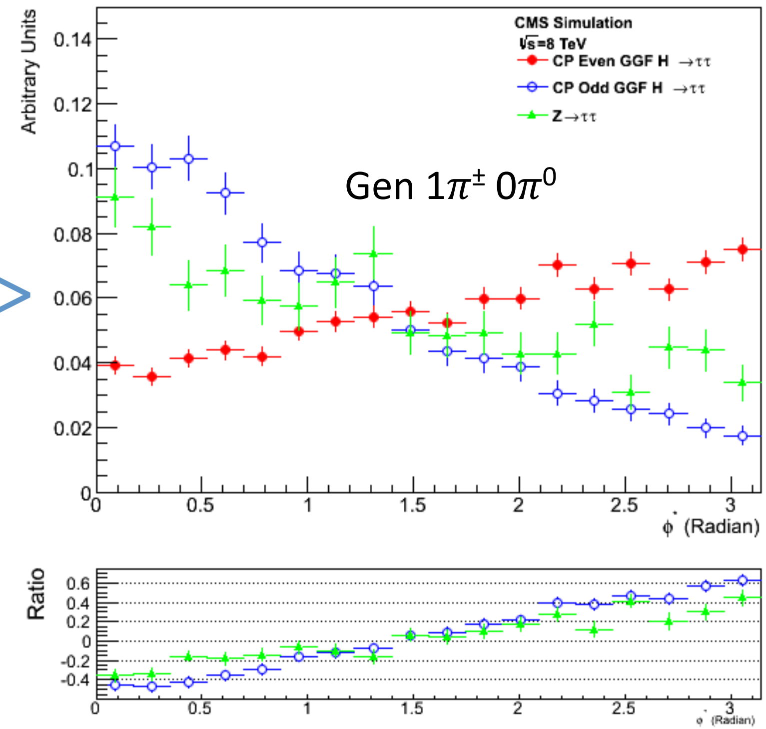
Impact of cut on IP length

Generated distribution with both vertex and PCA Smearing

No cut on IP length



IP length > 50 μm



Cut on IP length seems to reduce some effect of PCA and vertex smearing

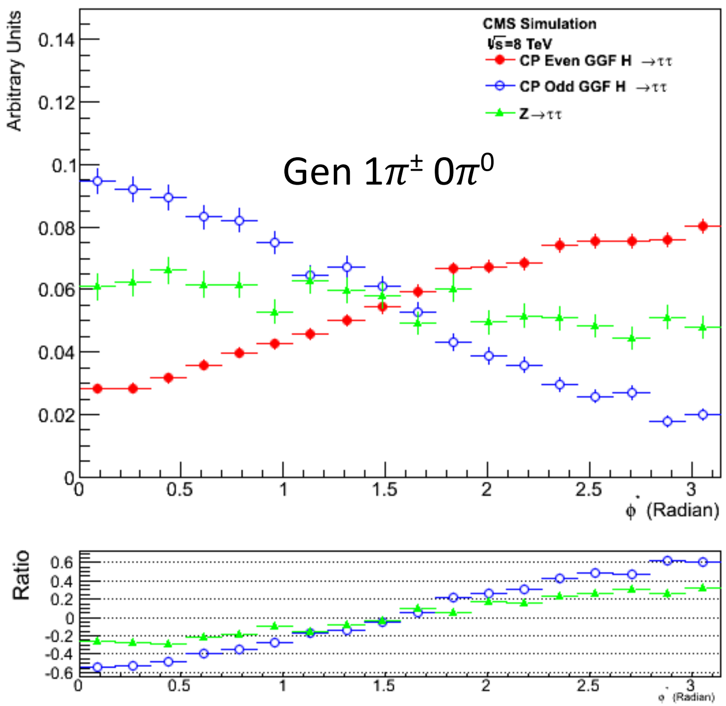
Vertex Smearing

Smear Vertex position with Gaussian

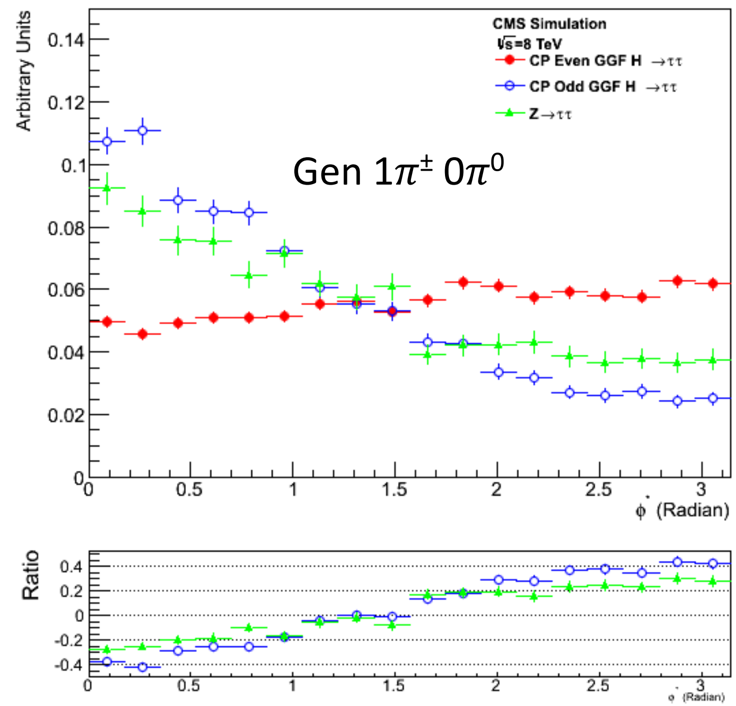
σ_x = 10 μm, σ_y = 10 μm, σ_z = 30 μm

Effect of σ_z (Vertex)

Smearing of X and Y only



Smearing of X, Y and Z



The rotation of the distribution is mainly due to the smearing of vertex Z-position