HEP NTUA
Weekly Report

30/9/2020

G. Bakas

Summary

- Analysis:
  - Switched to Tune CP5 nominal MC's for 2016
  - Start investigating ttbar Systematic Uncertainties

  - Investigate on how to combine the measurements between the three years
    - Combine them in the fiducial level and extract the cross section?
    - Extract the cross sections individually and combine the measurements in the unfolded level?

  - Top Angular Distributions: chi, |cosθ*| leading and subleading
    - Responses, Signal Extraction →Unfolding
    - Results!

  - AN 2020/156

Signal Selection

Variables	Selected Cut
pT (both leading jets)	> 400 GeV
Njets	> 1
N leptons	= 0
|eta| (both leading jets)	< 2.4
mJJ	> 1000 GeV
jetMassSoftDrop (only for fit)	(50,300) GeV
Top Tagger	> 0.2, 0, 0.1
B tagging (2 btagged jets)	> Medium WP
Signal Trigger	

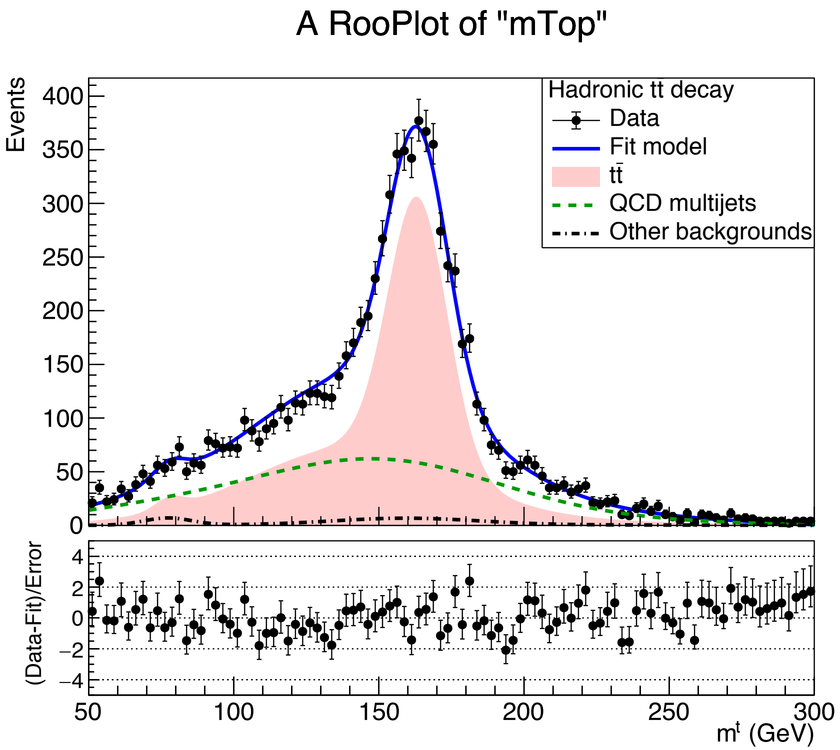
Control Region Selection

Variables	Selected Cut
pT (both leading jets)	> 400 GeV
Njets	> 1
N leptons	= 0
|eta| (both leading jets)	< 2.4
mJJ	> 1000 GeV
jetMassSoftDrop (only for fit)	(50,300) GeV
Top Tagger	> 0.2, 0, 0.1
B tagging (0 btagged jets)	< Medium WP
Control Trigger	

Mass Fit in Extended SR (SR_A)

- Both SR and Control Region use the Medium btag WP.
- Intuition is to remove the ttbar and subdominant bkg contribution from the data Control Region

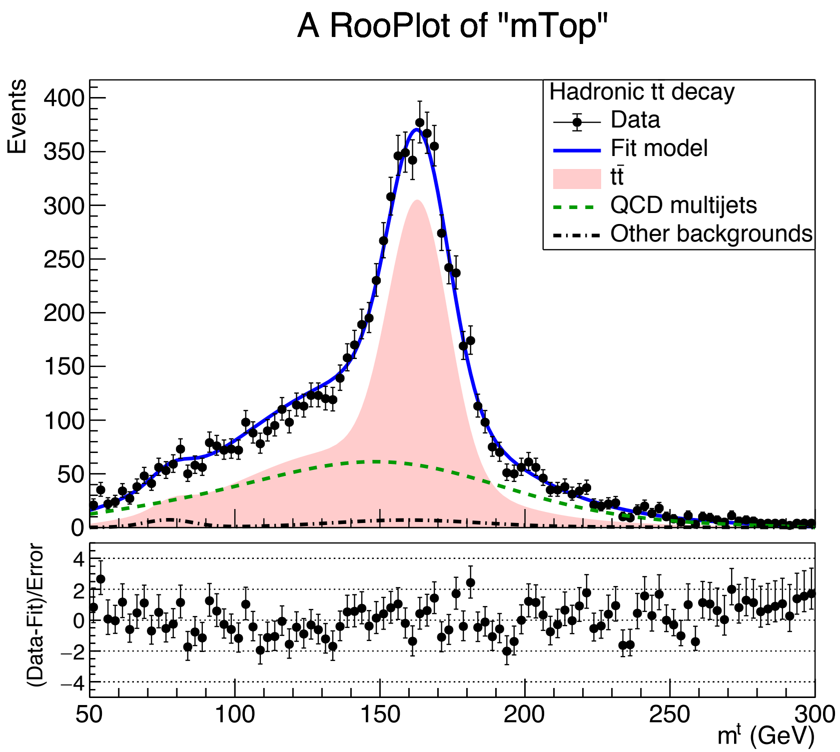
$$QCD_0(m^t) = D_0(m^t) - T_0(m^t) - Sub_0(m^t)$$



(2016)
Tune CUE

Floating Parameter	FinalValue +/-	Error
kMassResol	9.2251e-01 +/-	2.73e-02
kMassScale	9.9891e-01 +/-	2.01e-03
kQCD_2b	6.9753e-02 +/-	5.26e-02
nFitBkg_2b	2.4472e+02 +/-	1.47e+02
nFitQCD_2b	2.9890e+03 +/-	1.74e+02
nFitSig2b	5.2763e+03 +/-	1.67e+02

$t\bar{t}$ Signal strength: r = 0.686668 ± 0.0263103

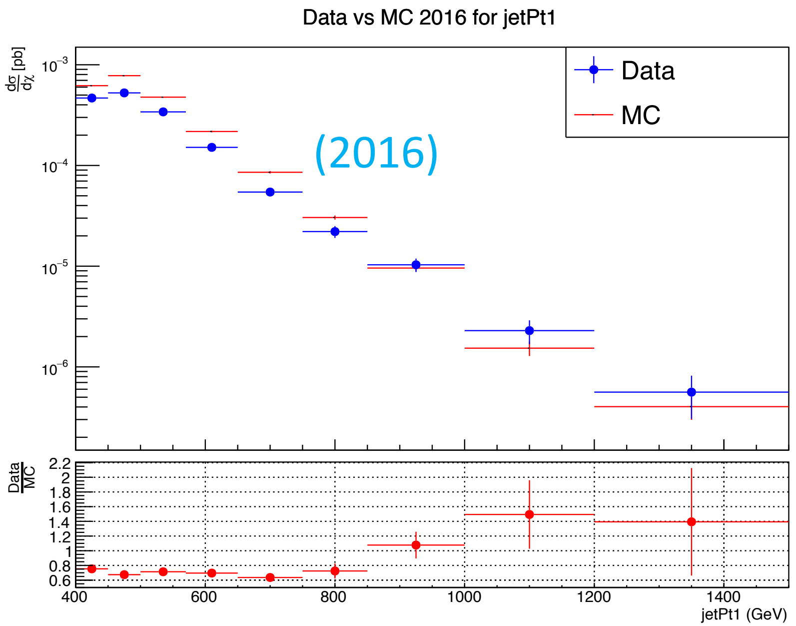


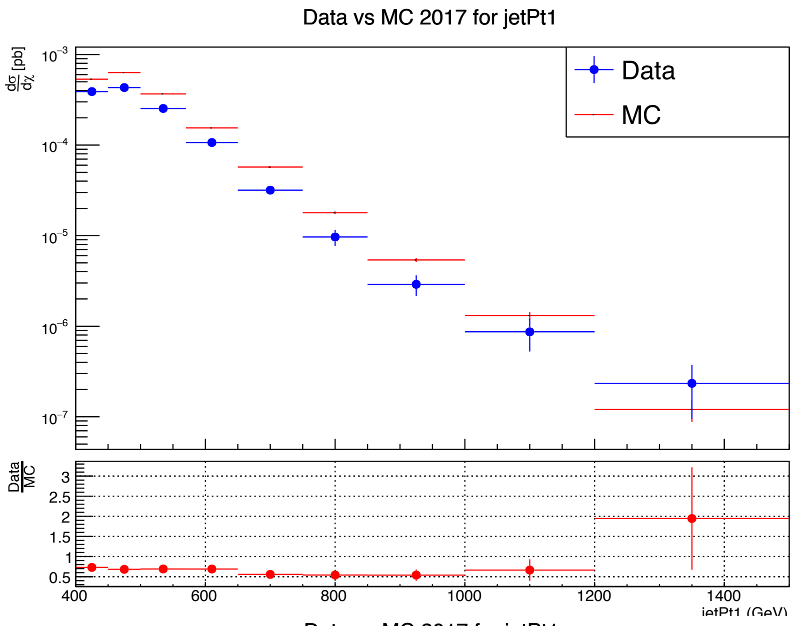
(2016)
Tune CP5

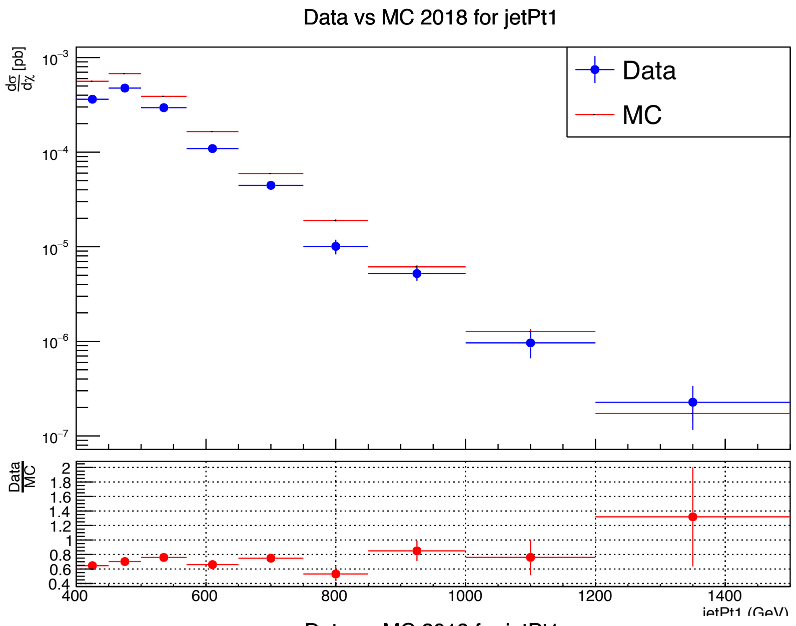
Floating Parameter	FinalValue +/-	Error
kMassResol	9.3798e-01 +/-	2.75e-02
kMassScale	9.9751e-01 +/-	1.99e-03
kQCD_2b	2.3668e-01 +/-	5.05e-01
nFitBkg_2b	2.5758e+02 +/-	1.47e+02
nFitQCD_2b	2.9242e+03 +/-	1.76e+02
nFitSig2b	5.3282e+03 +/-	1.67e+02

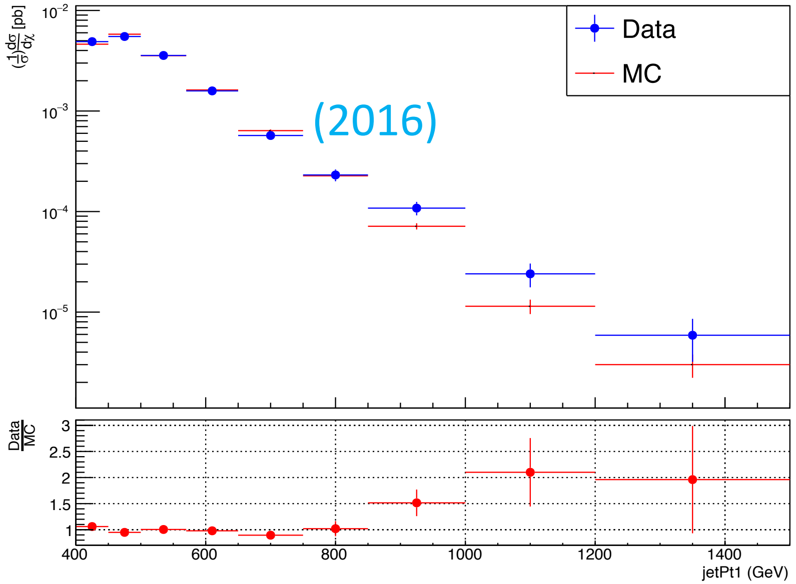
$t\bar{t}$ Signal strength: r = 0.641241 ± 0.0238714 (new)

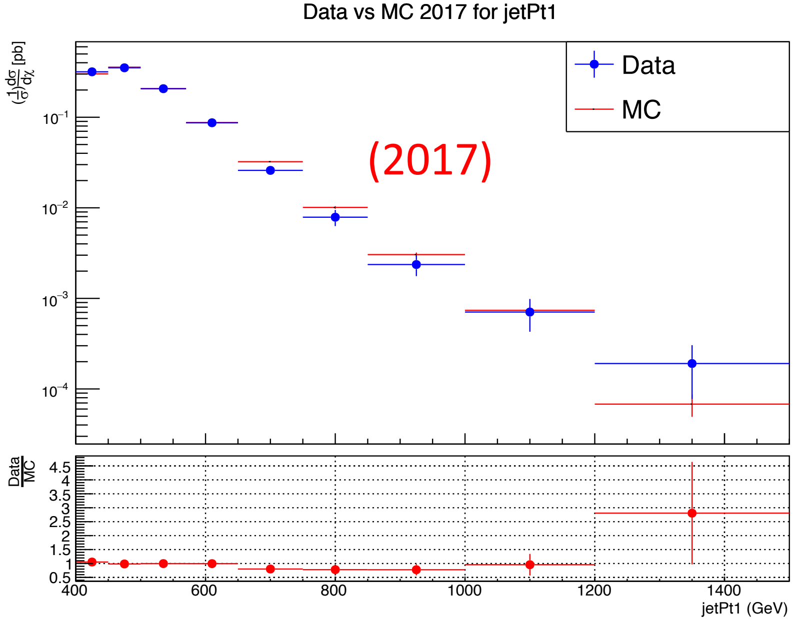
Fiducial Differential Cross Section

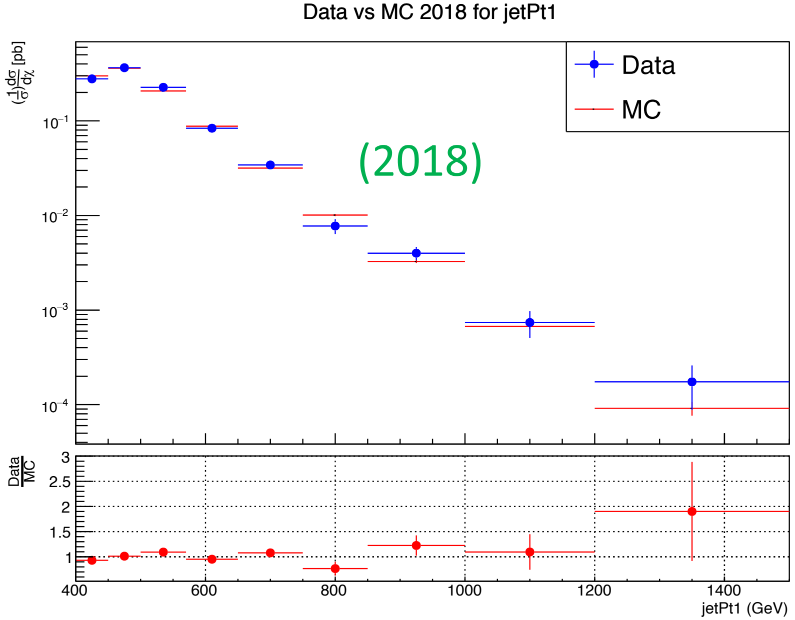




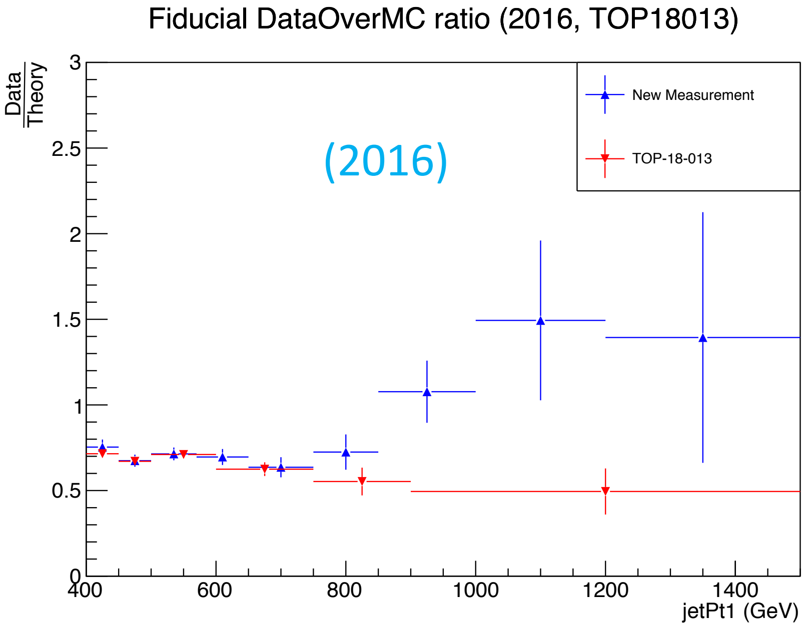


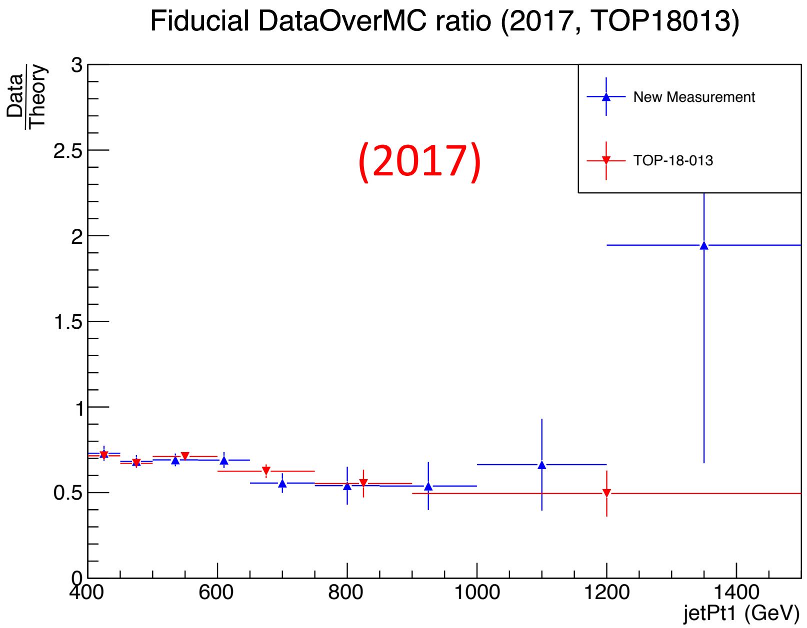


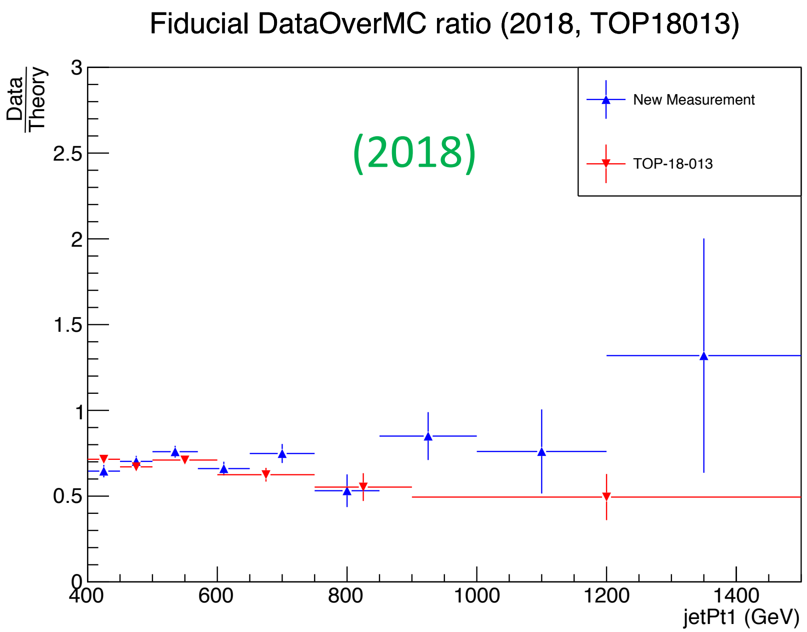


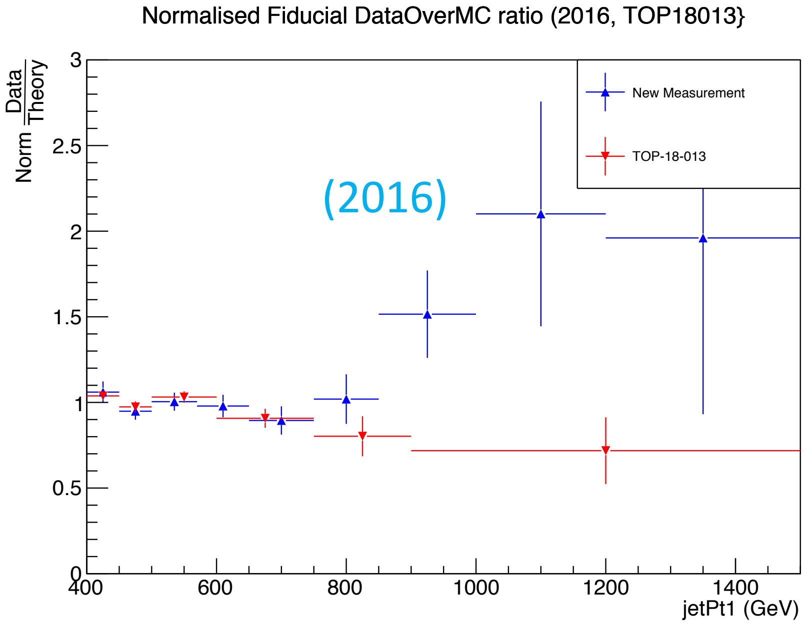


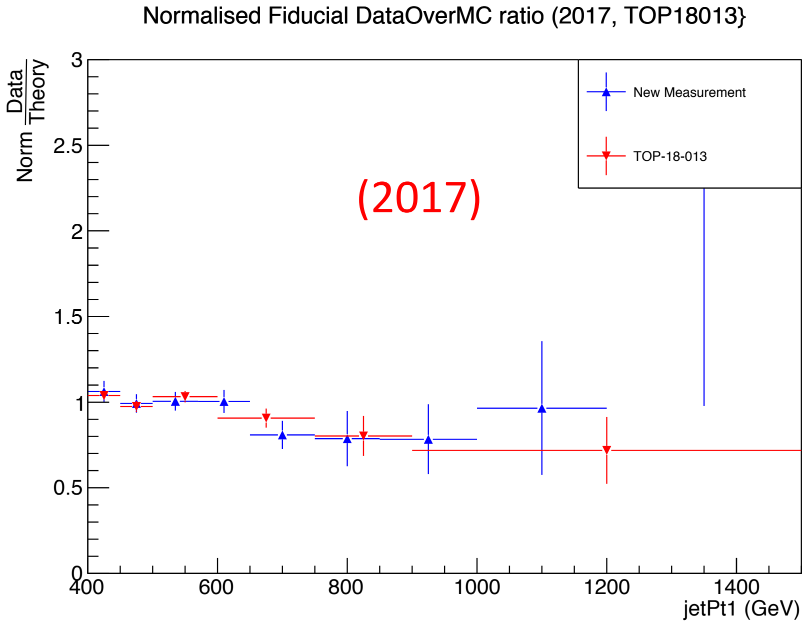
Fiducial Differential Cross Section Comparison

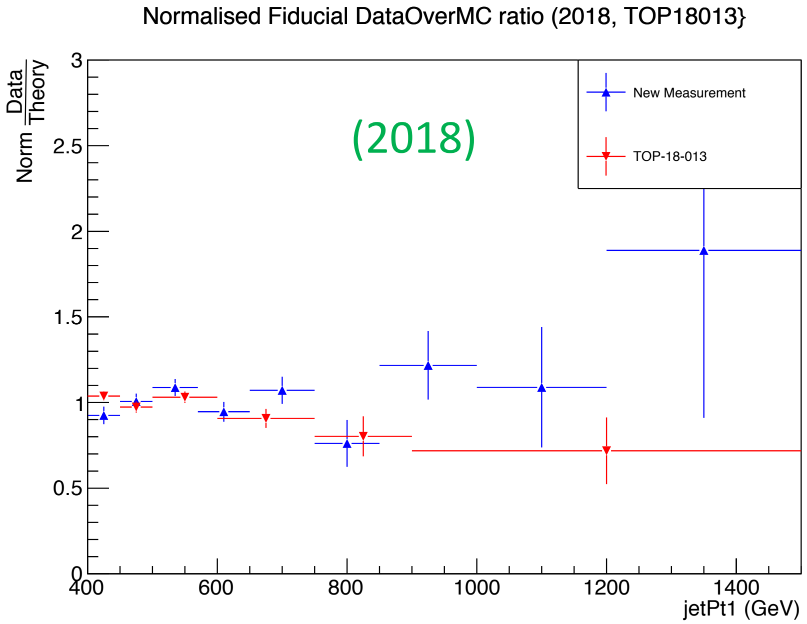




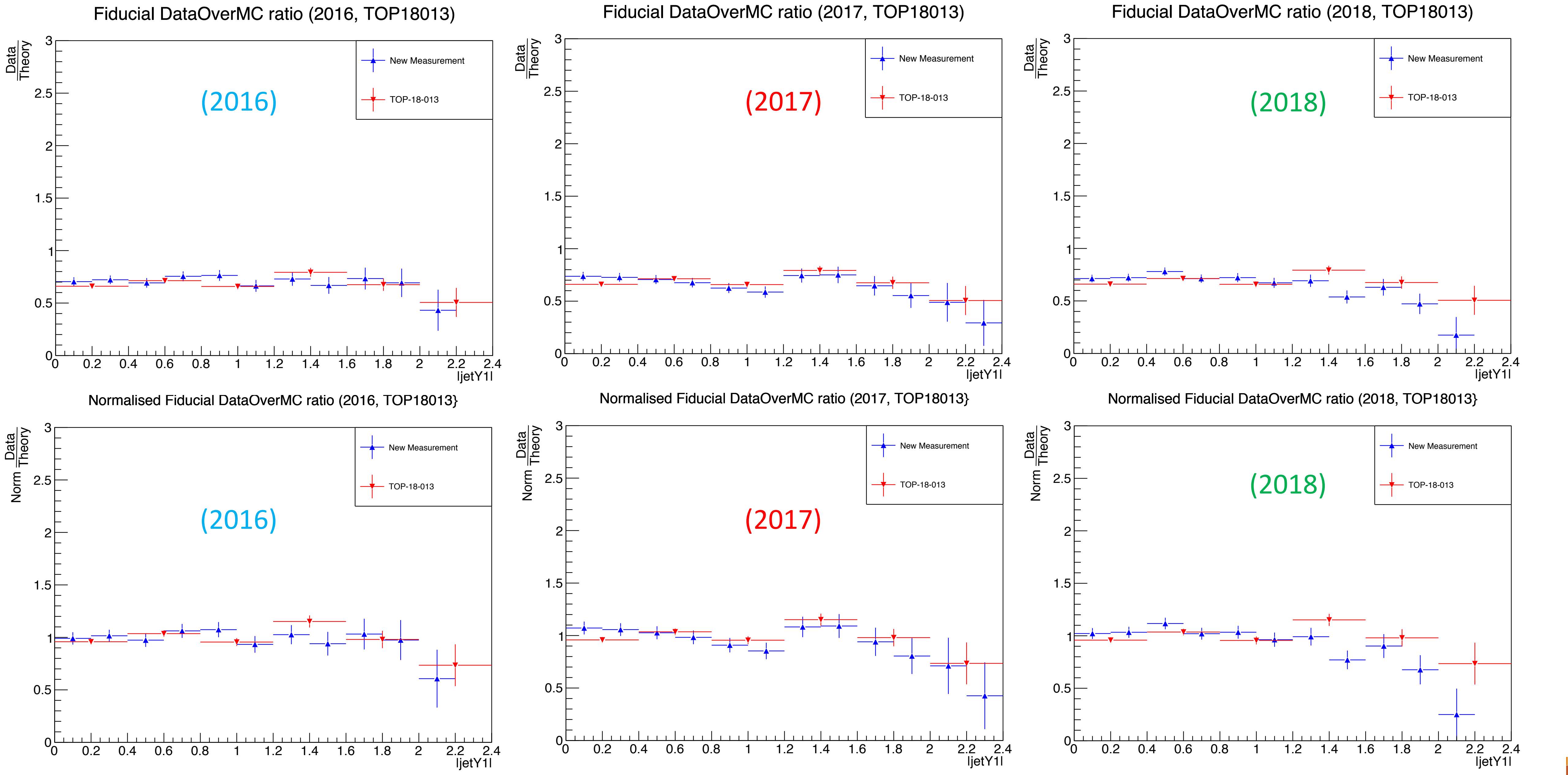




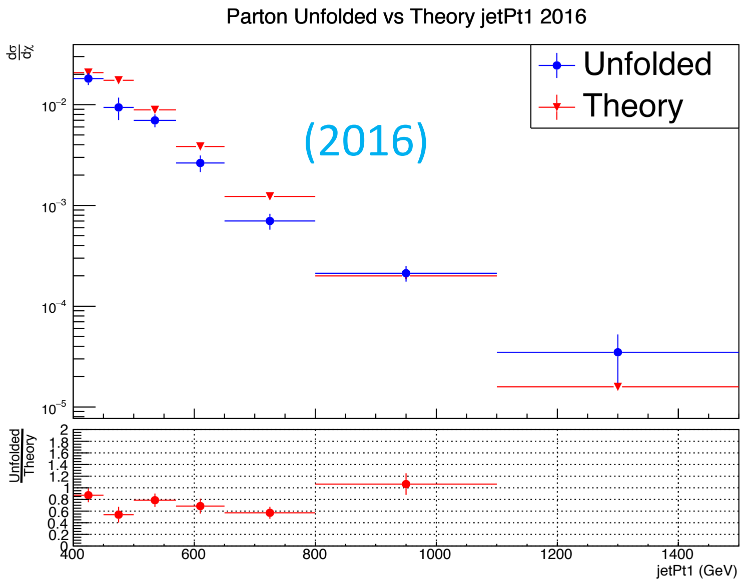


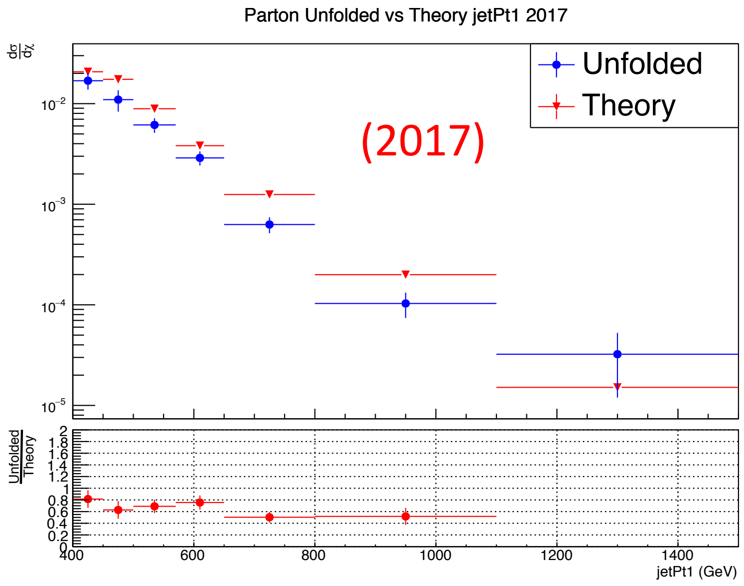


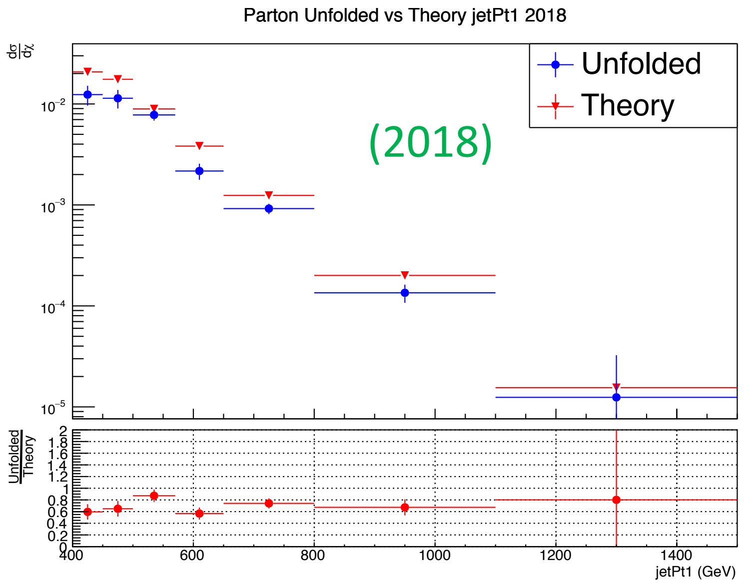
Fiducial Differential Cross Section Comparison

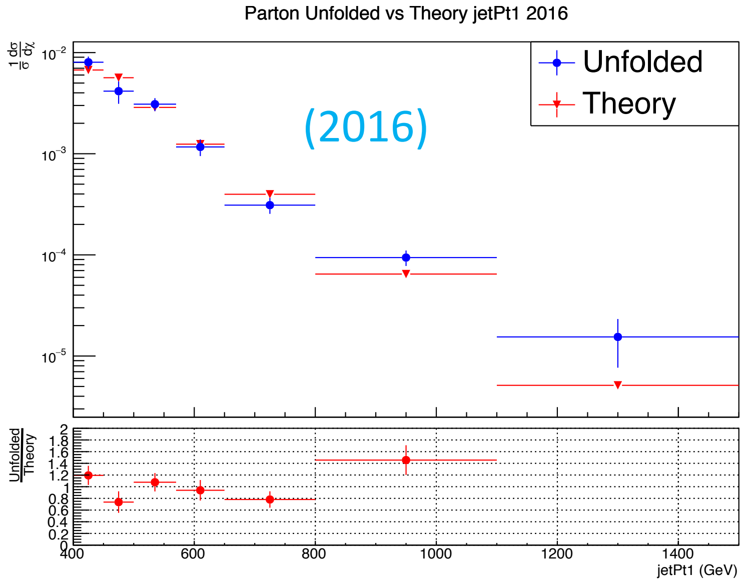


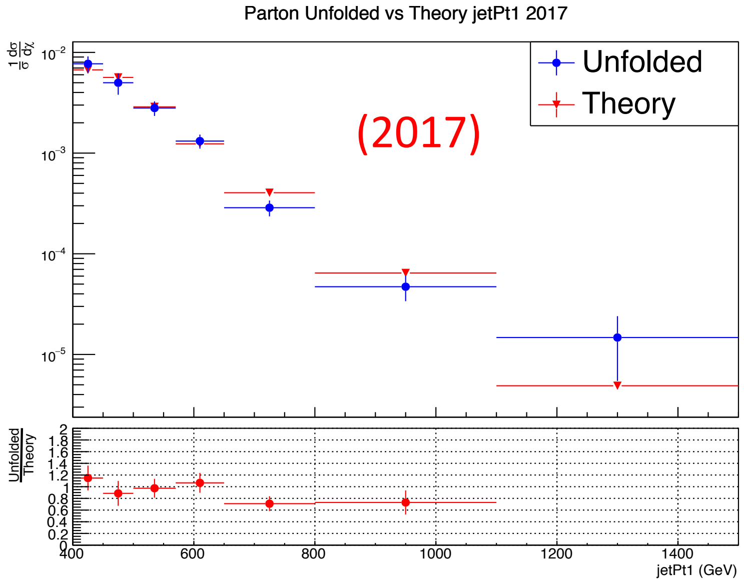
Parton Differential Cross Section

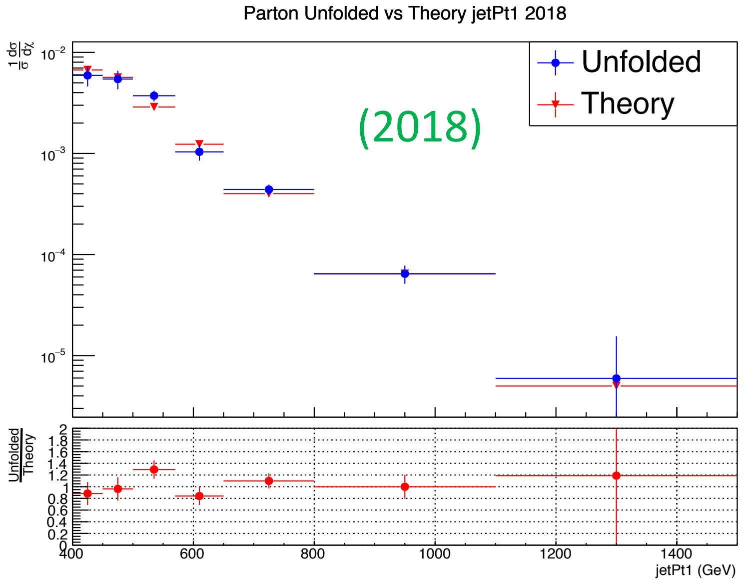






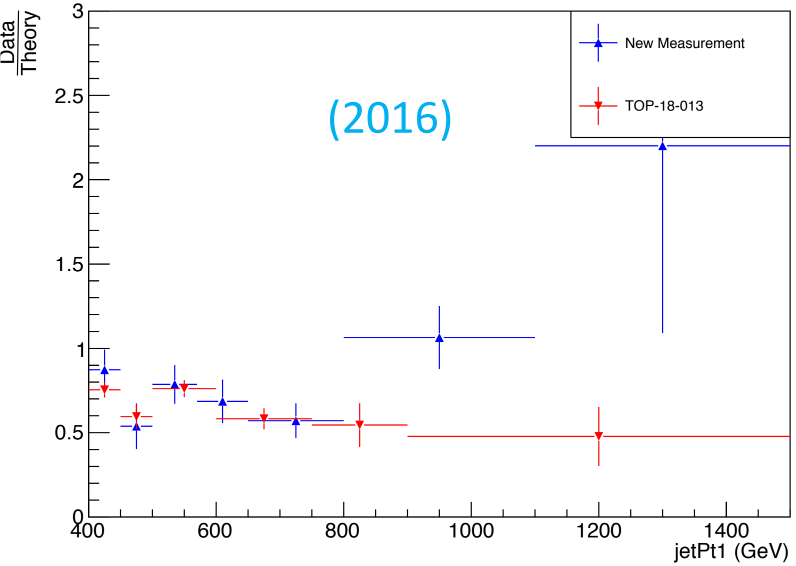




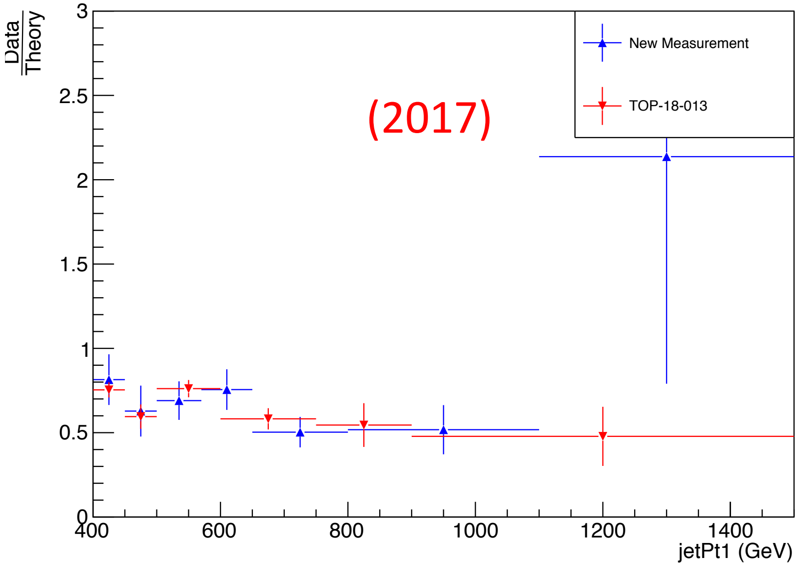


Parton Differential Cross Section Comparison

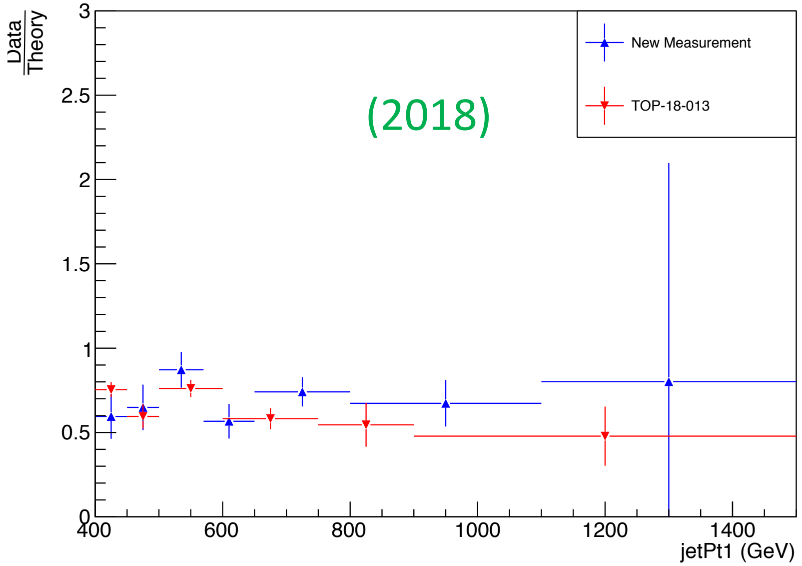
Parton DataOverMC ratio (2016, TOP18013)



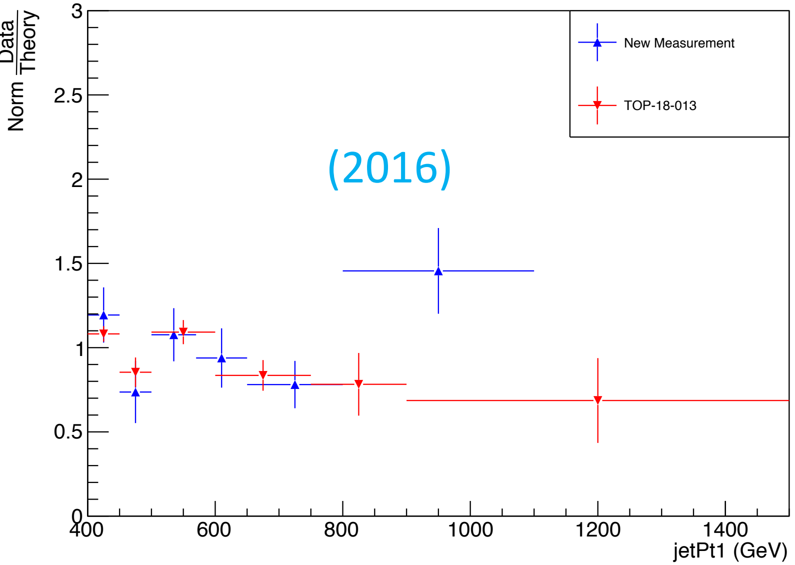
Parton DataOverMC ratio (2017, TOP18013)



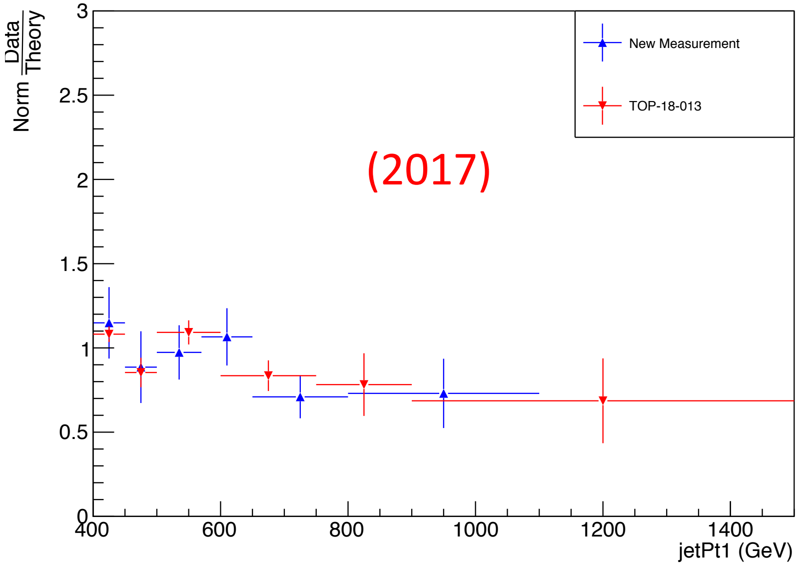
Parton DataOverMC ratio (2018, TOP18013)



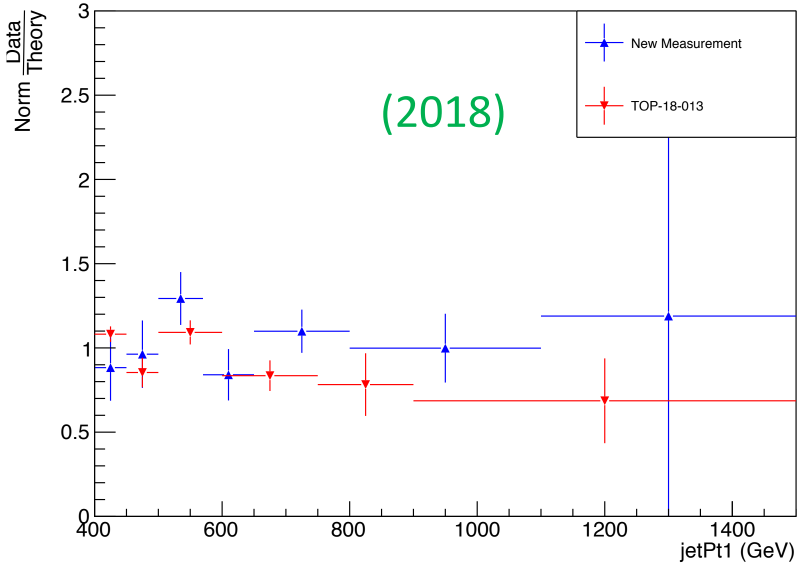
Normalised Parton DataOverMC ratio (2016, TOP18013}



Normalised Parton DataOverMC ratio (2017, TOP18013}

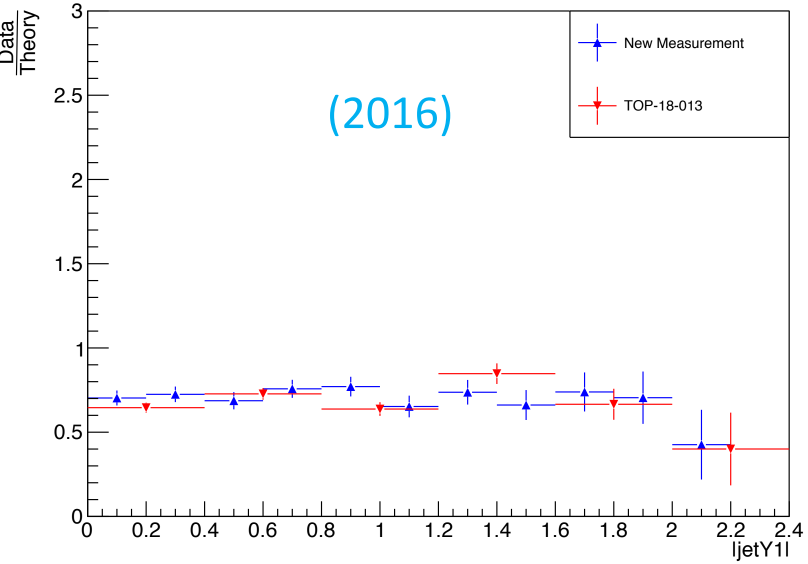


Normalised Parton DataOverMC ratio (2018, TOP18013}

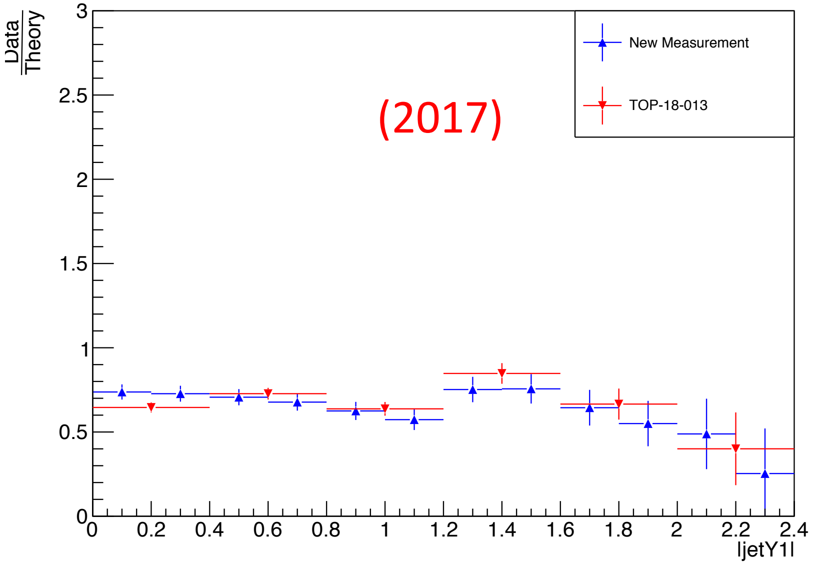


Parton Differential Cross Section Comparison

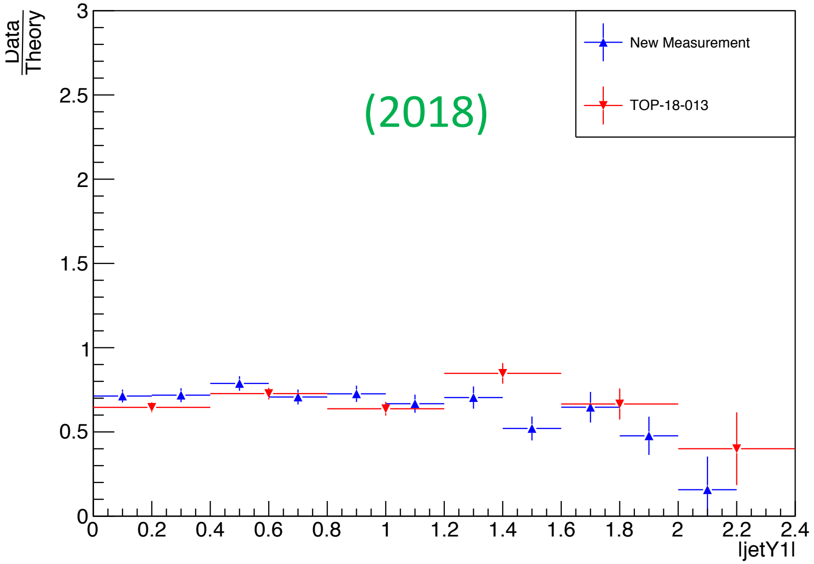
Parton DataOverMC ratio (2016, TOP18013)



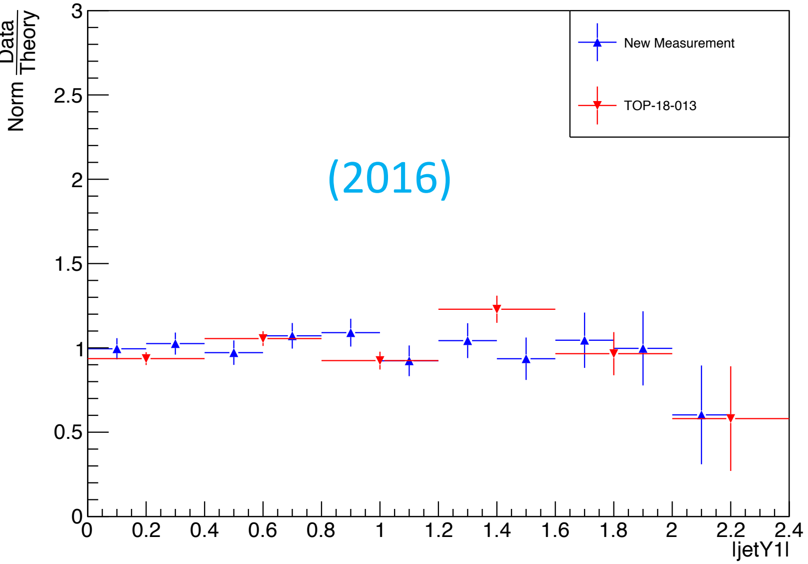
Parton DataOverMC ratio (2017, TOP18013)



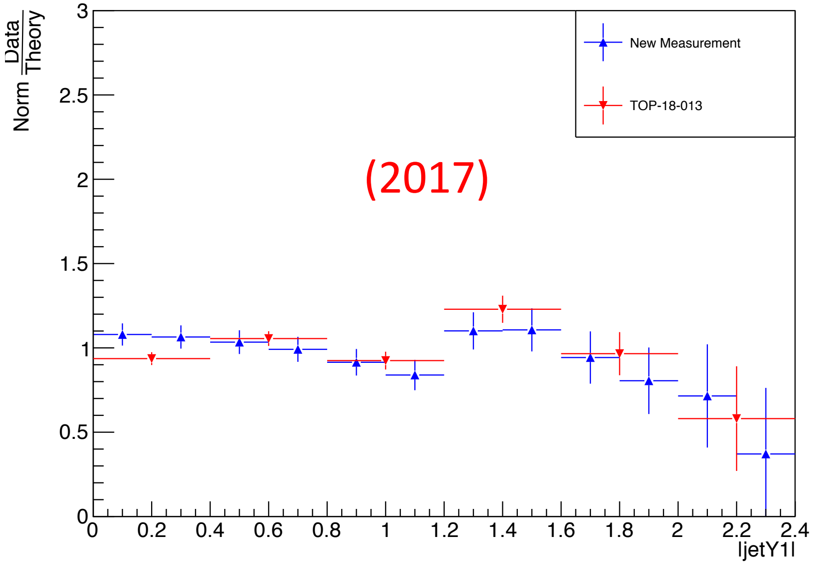
Parton DataOverMC ratio (2018, TOP18013)



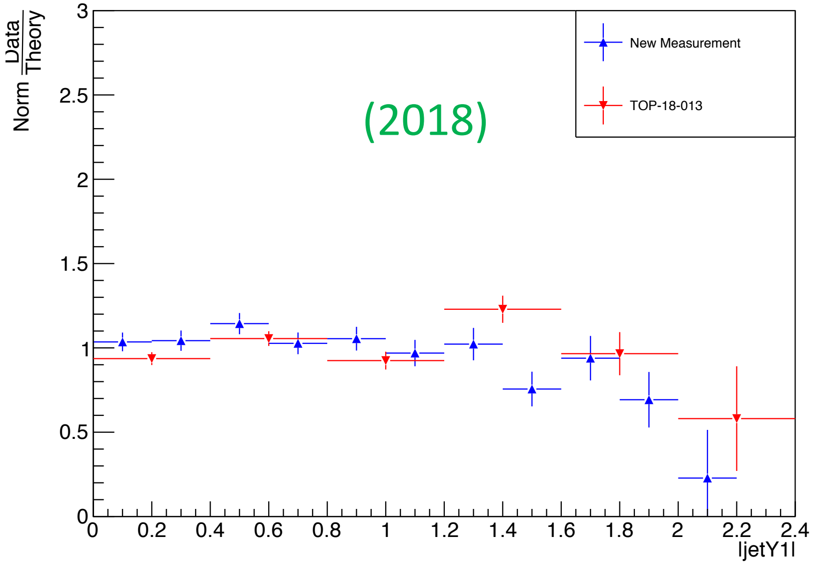
Normalised Parton DataOverMC ratio (2016, TOP18013}



Normalised Parton DataOverMC ratio (2017, TOP18013}



Normalised Parton DataOverMC ratio (2018, TOP18013}



Tag And Probe Calculations

(2016) Tune CUE

eff data: 0.781 ± 0.038
eff ttbar: 0.772 ± 0.014

Efficiency per Pt region
eff data pT[400-600]: 0.761 ± 0.042
eff ttbar pT[400-600]: 0.778 ± 0.016
----
eff data pT[600-800]: 0.851 ± 0.100
eff ttbar pT[600-800]: 0.748 ± 0.031
----
eff data pT[800-Inf]: 0.886 ± 0.160
eff ttbar pT[800-Inf]: 0.775 ± 0.063

(2016) Tune CP5

eff data: 0.781 ± 0.038
eff ttbar: 0.813 ± 0.01

Efficiency per Pt region
eff data pT[400-600]: 0.761 ± 0.042
eff ttbar pT[400-600]: 0.812 ± 0.011
----
eff data pT[600-800]: 0.851 ± 0.100
eff ttbar pT[600-800]: 0.820 ± 0.023
----
eff data pT[800-Inf]: 0.886 ± 0.160
eff ttbar pT[800-Inf]: 0.806 ± 0.053

Top Angular Distributions

- We employ the dijet angular variable χ from the rapidities of the two leading jets
- Why χ?
  - The distributions associated with the final states produced via QCD interactions are relatively flat in comparison with the distributions of the BSM models or new particles, which typically peak at low values of x

- We can measure the variable χ in two ways

1. By measuring the difference of the rapidities of the two leading jets such as the corresponding rapidity in the ZMF is:

$$y^* = \frac{1}{2}(y_1 - y_2)$$

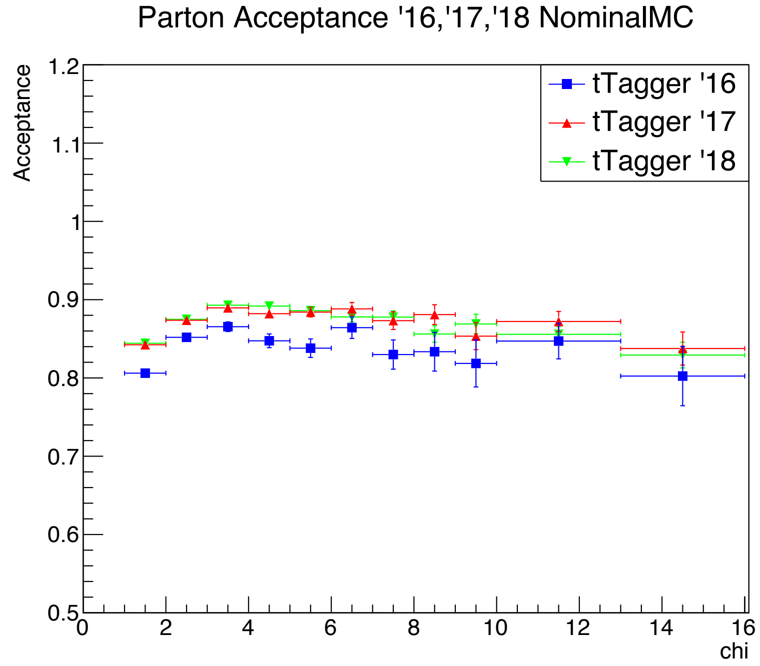
X is defined as $\chi = e^{|y^*|} = e^{|y_1 - y_2|}$ (1) and can be measured by creating the TLorentzVector, boost it to the ZMF and find the rapidity difference of the two leading jets

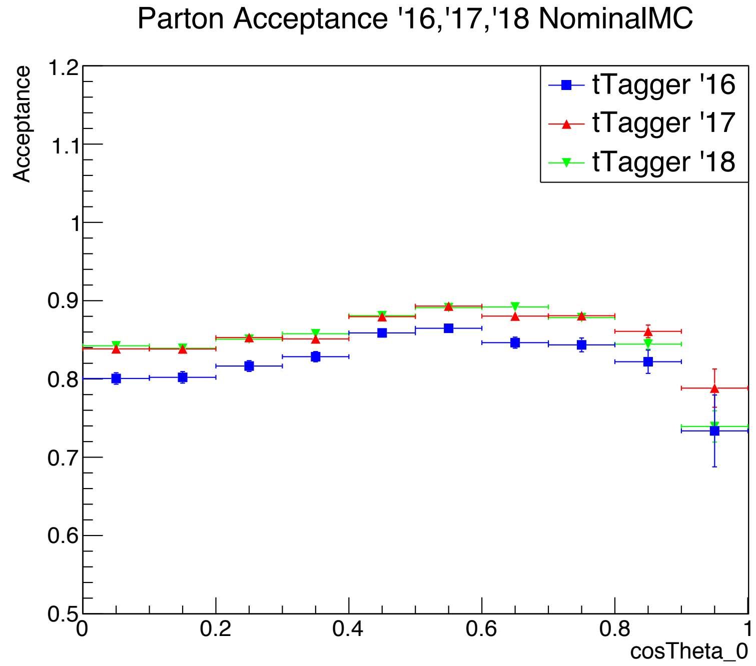
2. By measuring the scattering angle θ* (angle between top quark and z-axis in the Zero Momentum Frame)

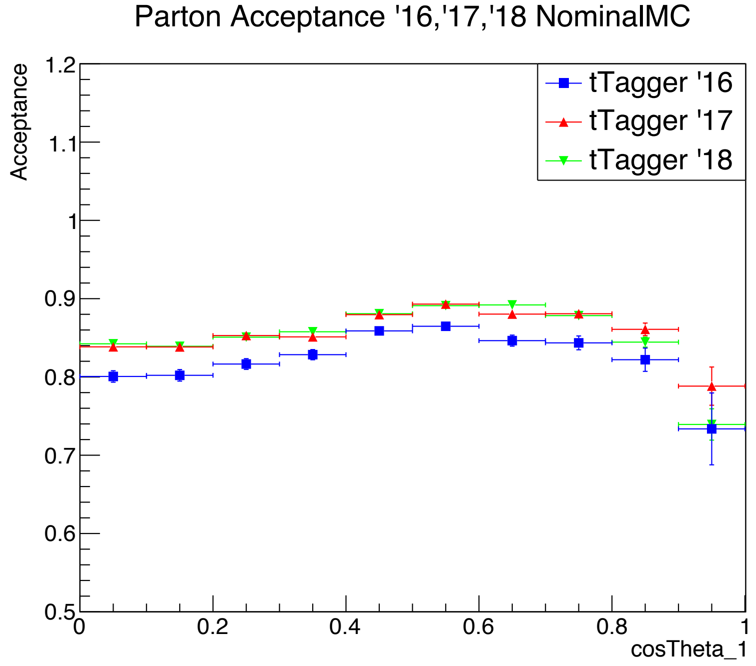
We define as $y^* = \frac{1}{2}\ln(\frac{1+|cos\theta^*|}{1-|cos\theta^*|})$ and from (1) we can find that:

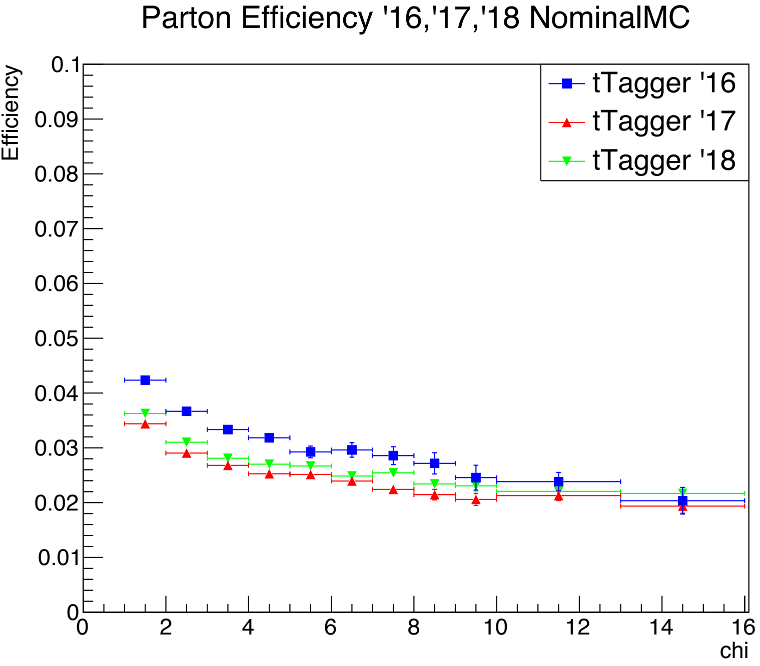
$$\chi = \frac{1 + |cos\theta^*|}{1 - |cos\theta^*|}$$

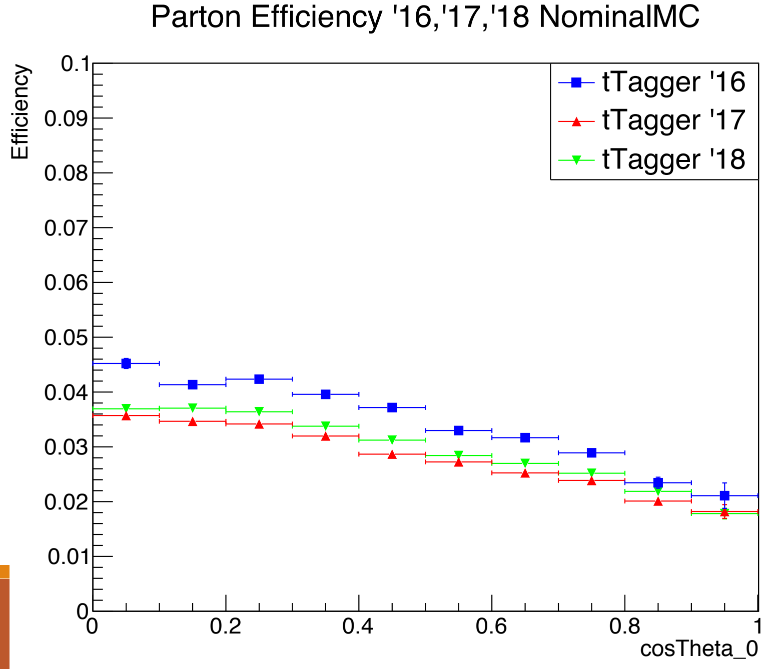
Efficiency, Acceptance

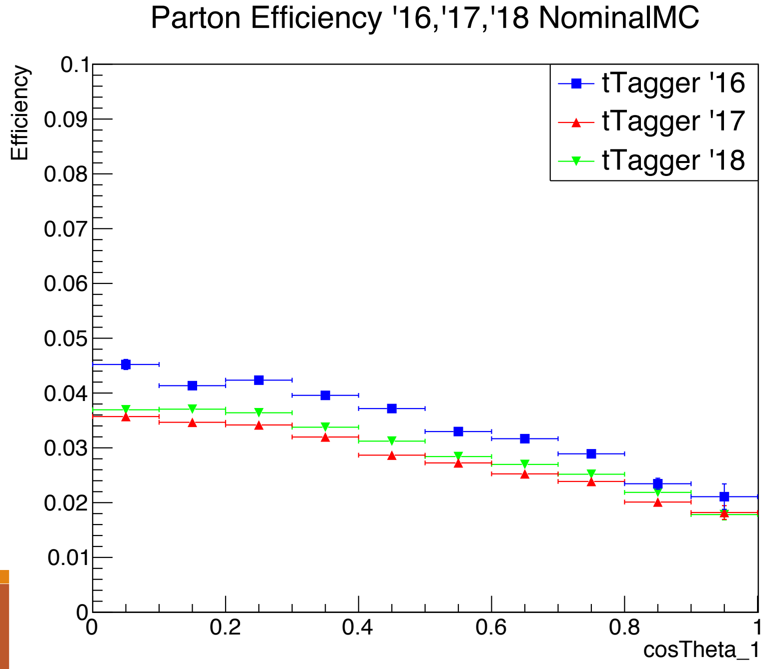




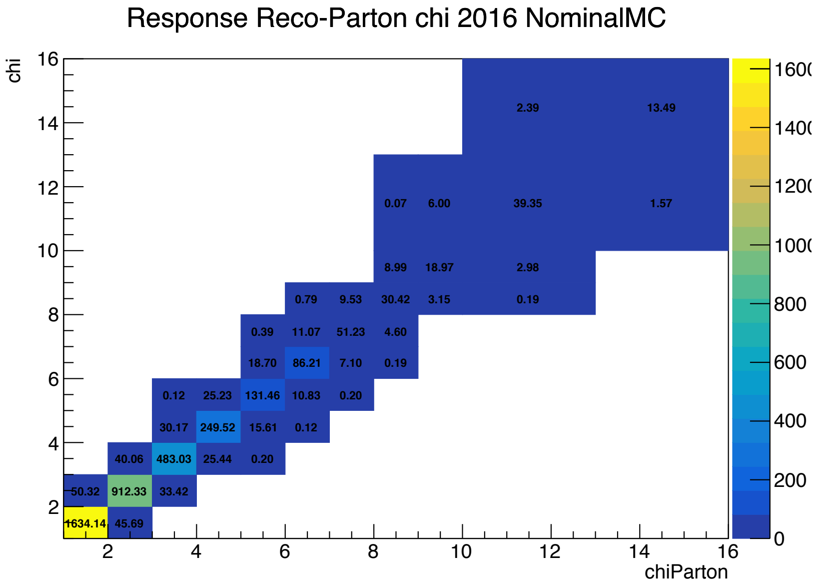


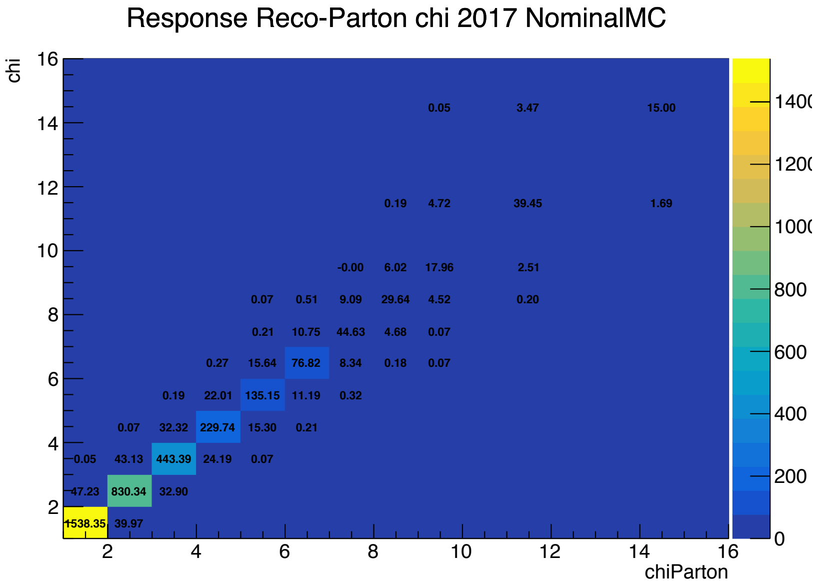


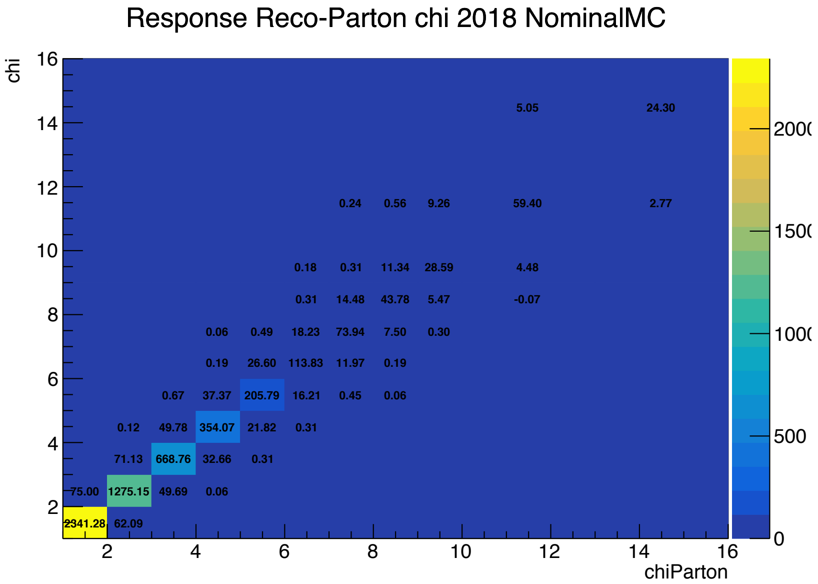




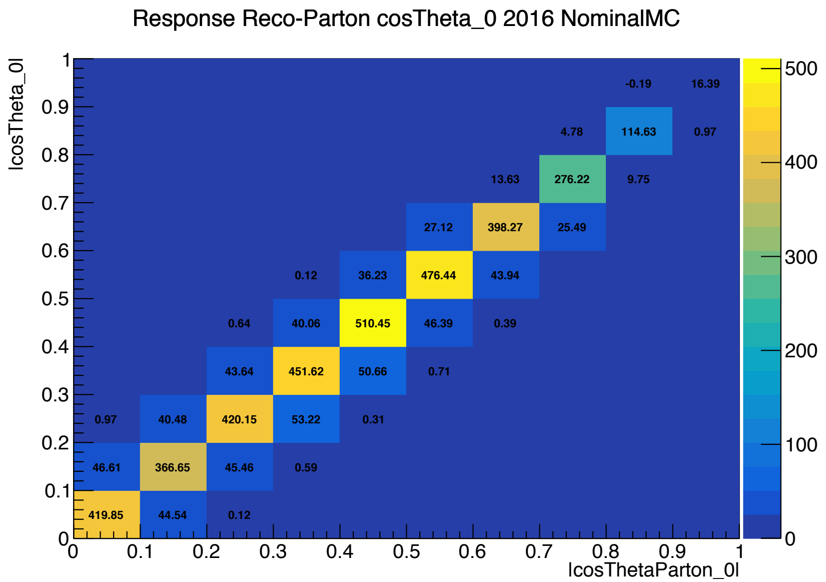
Response Matrices

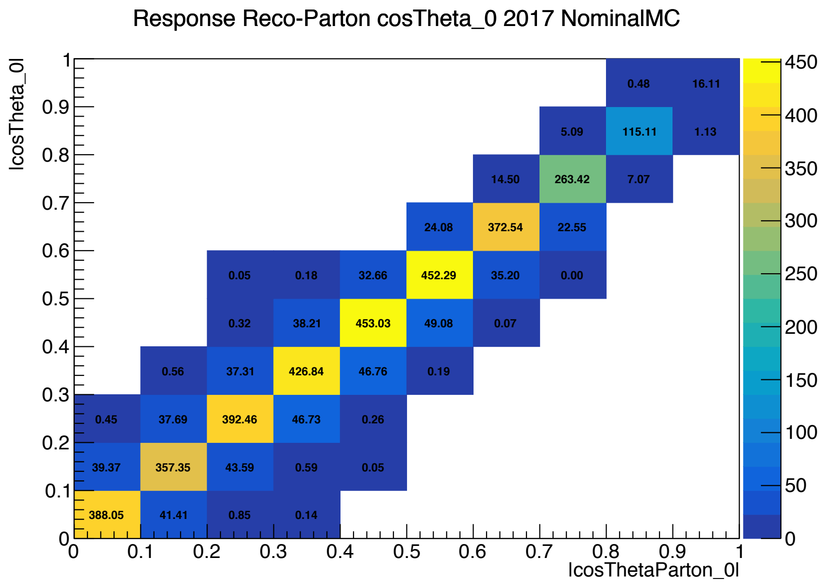


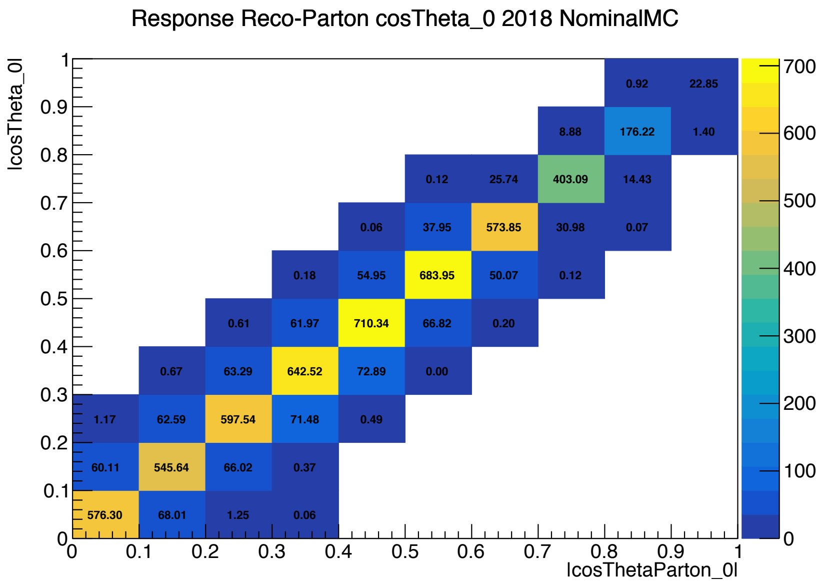




Response Matrices







Signal Extraction

Measured dist from data

Fitted number of QCD events in SR_A

QCD shape taken from Data (CR)

$$S(x_{reco}) = D(x_{reco}) - C_{bkg}^{yield} N_{QCD}^{fit} C_{QCD}^{shape}(x_{reco}) Q(x_{reco}) - B(x_{reco})$$

Fiducial Yield

Transfer factor from SR_A to SR

QCD shape correction factor

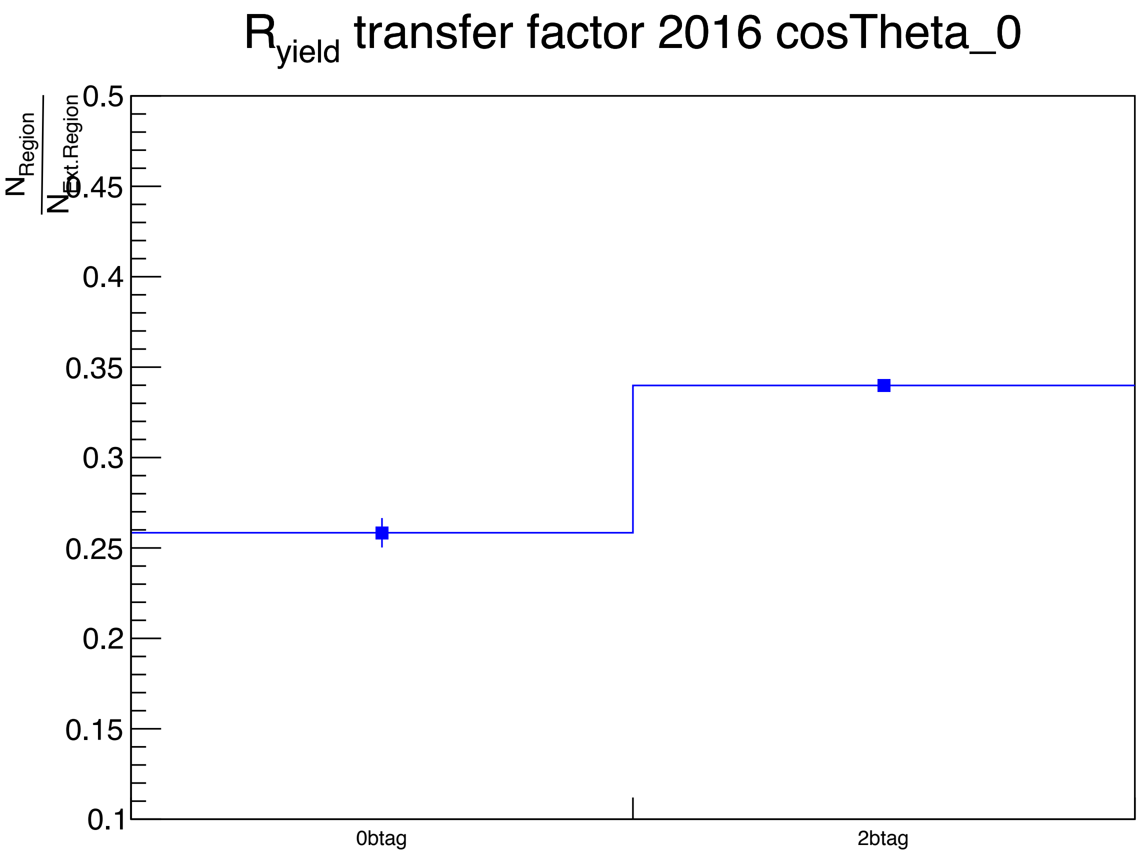
Subdominant bkg shape and contribution (MC)

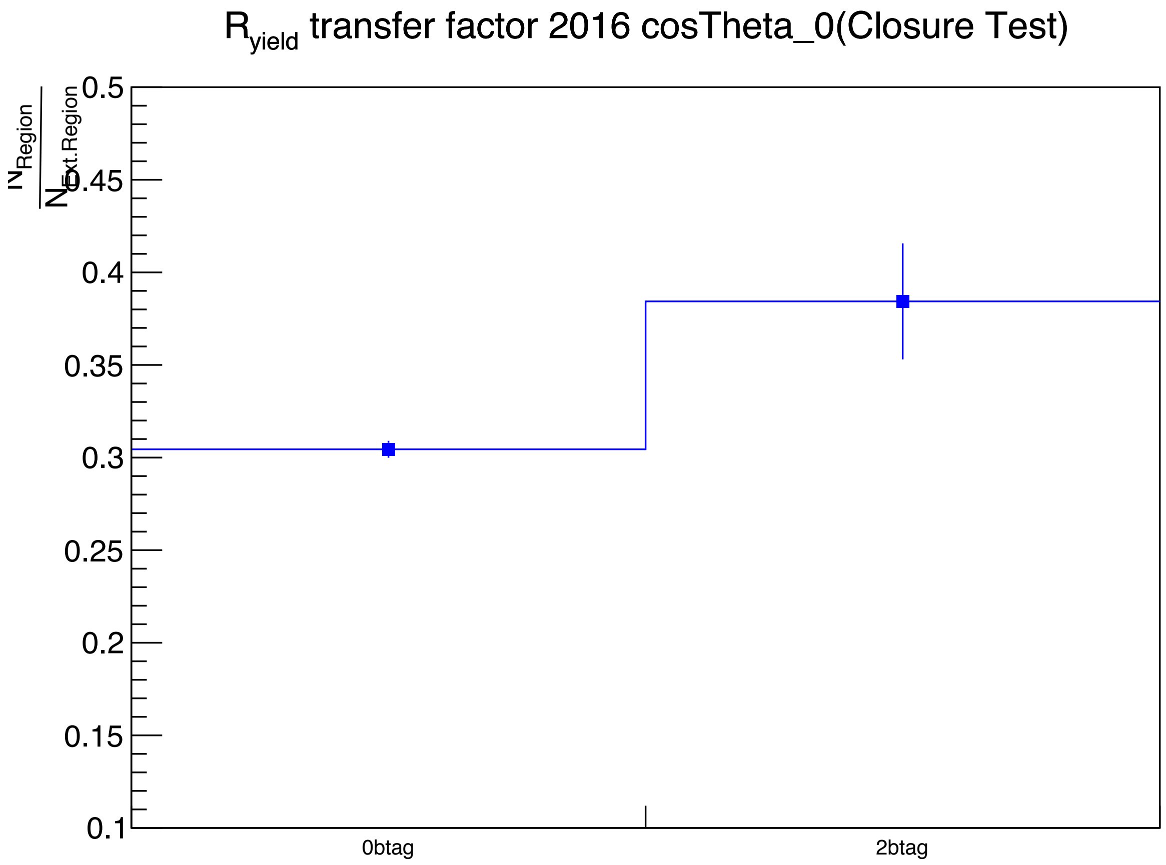
- Where x_{reco} is the respected variable of interest (ttbar mass, pt, rapidity, leading and subleading jetPt and |jetY|)
- We deploy a fit in the Signal Region (2btag) to extract the N_{QCD}^{fit}

$$D(m^t)^{(i)} = N_{tt}^{(i)} T^{(i)}(m^t, k_{MassScale}, k_{MassResolution}) + N_{bkg}^{(i)} B(m^t)(1 + k_1 x) + N_{sub}^{(i)} O^{(i)}(m^t)$$

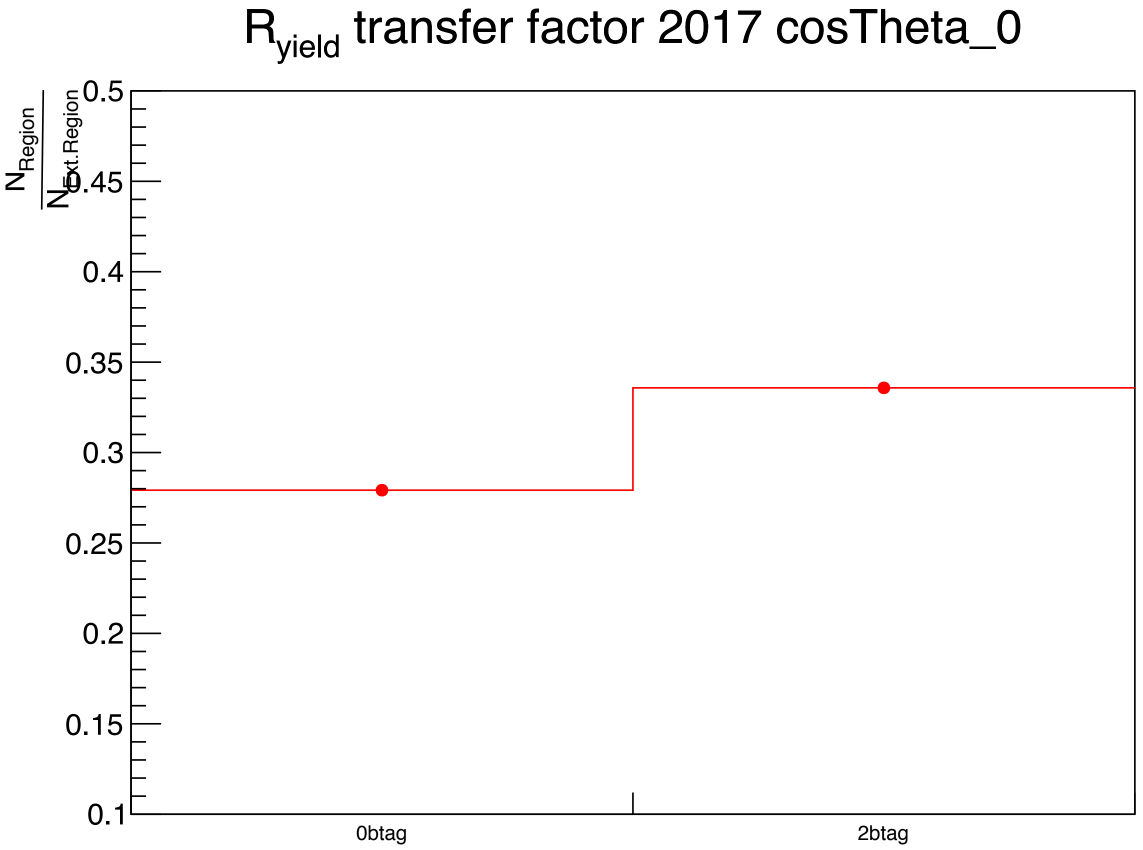
- Our data CR is contaminated from ttbar and subdominant bkgs which has to be dealt with.

R yield

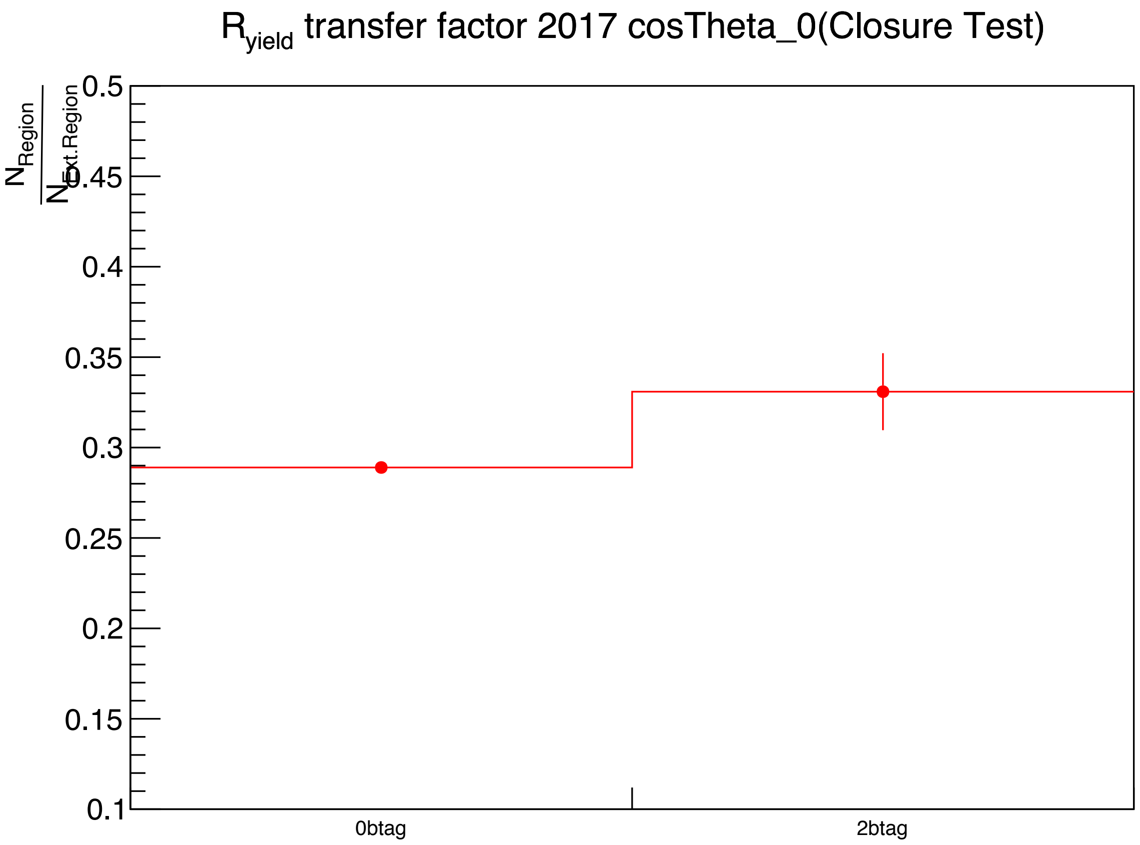

R_{yield} transfer factor 2016 cosTheta_0
$\frac{N_{Region}}{N_{Ext.Region}}$
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0btag
2btag


R_{yield} transfer factor 2016 cosTheta_0(Closure Test)
$\frac{N_{Region}}{N_{Ext.Region}}$
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0btag
2btag

R yield

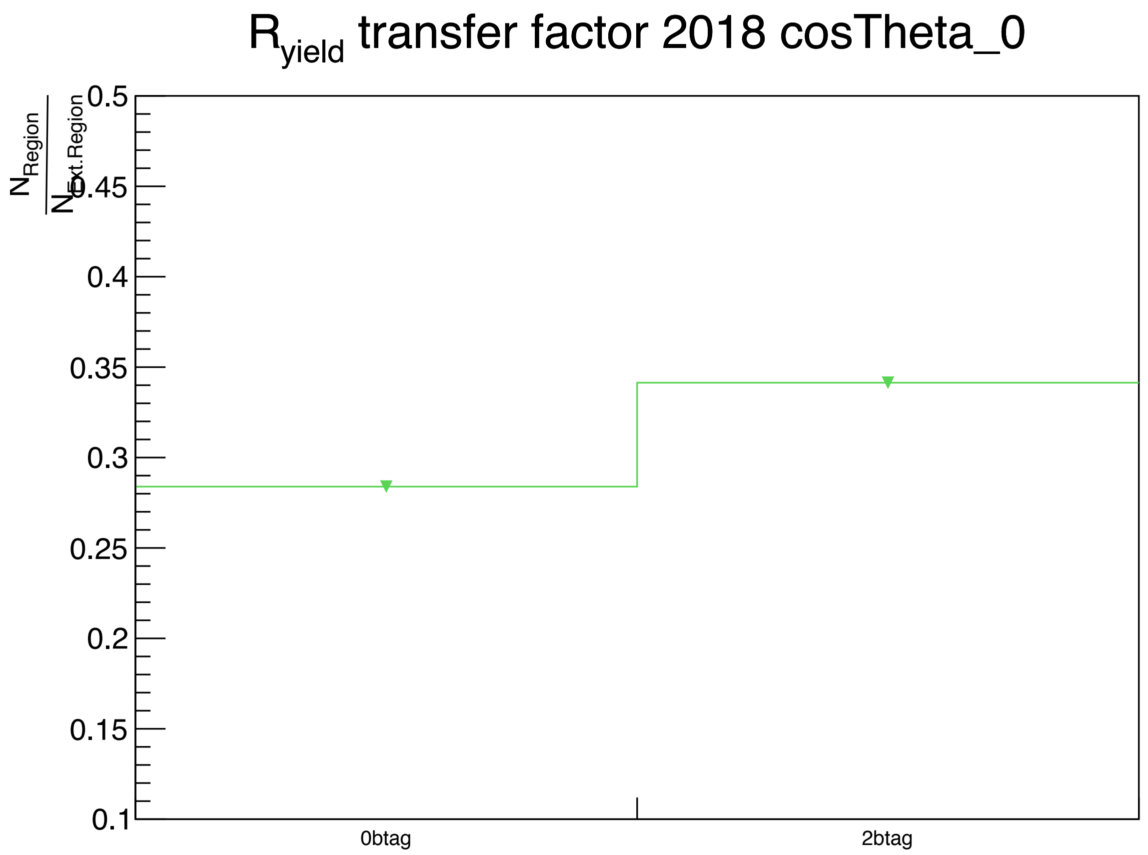


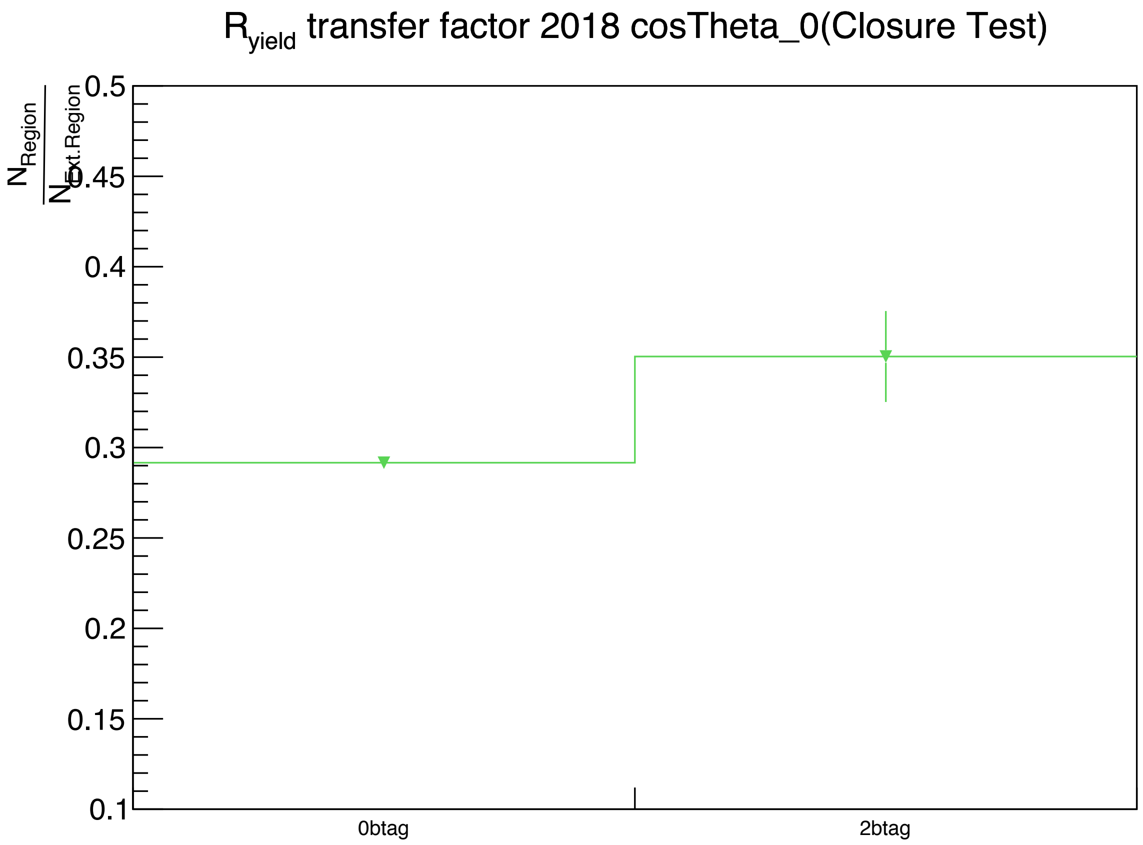
R_{yield} transfer factor 2017 cosTheta_0
$\frac{N_{Region}}{N_{Ext.Region}}$
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0btag
2btag



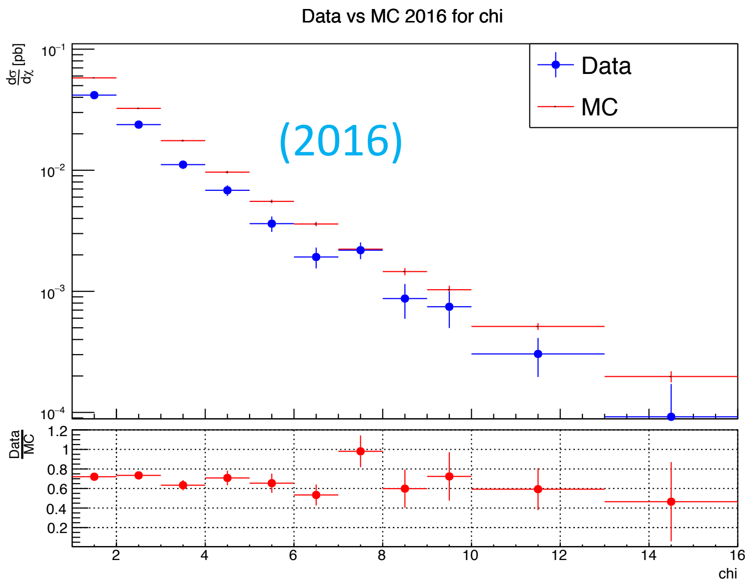
R_{yield} transfer factor 2017 cosTheta_0(Closure Test)
$\frac{N_{Region}}{N_{Ext.Region}}$
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0btag
2btag

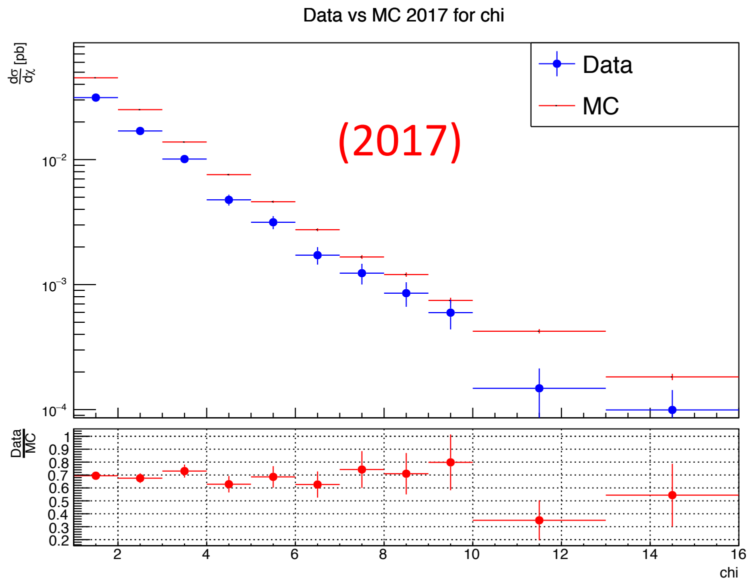
R yield

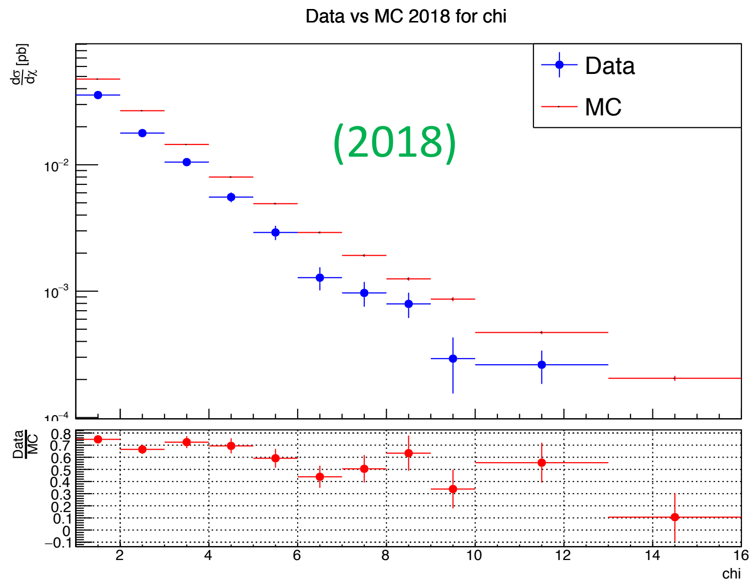

R_{yield} transfer factor 2018 cosTheta_0
$\frac{N_{Region}}{N_{Ext.Region}}$
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0btag
2btag

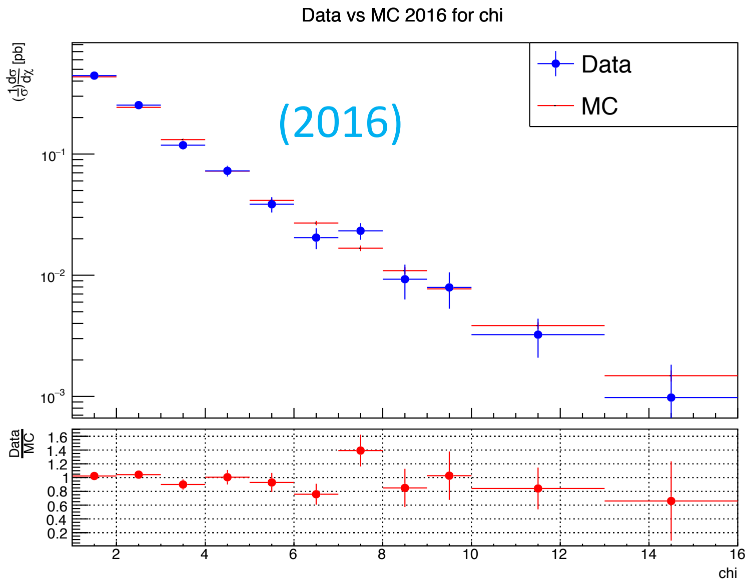

R_{yield} transfer factor 2018 cosTheta_0(Closure Test)
$\frac{N_{Region}}{N_{Ext.Region}}$
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0btag
2btag

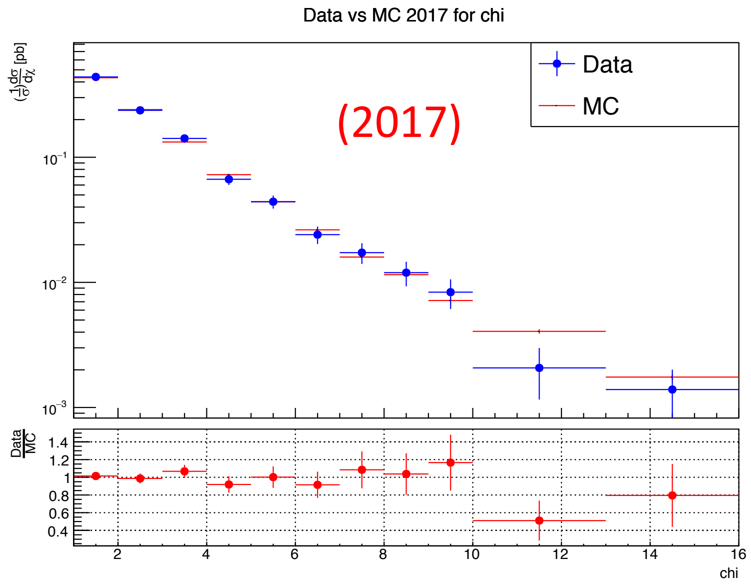
Fiducial Measurement

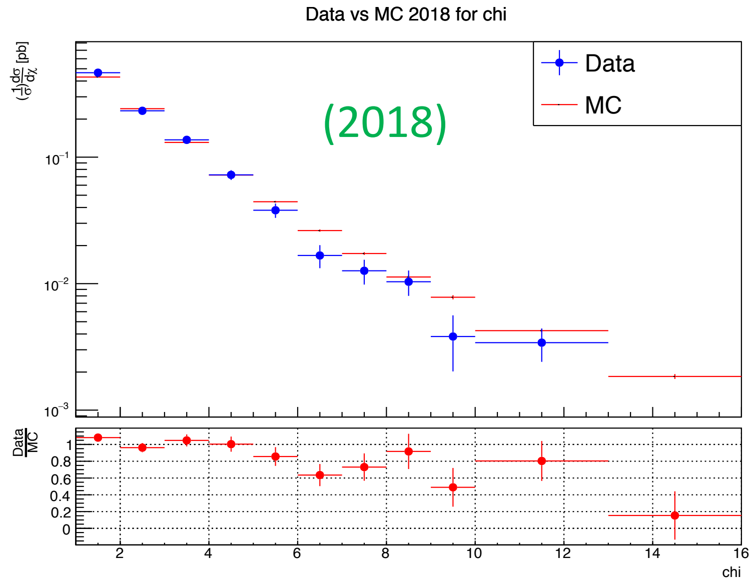




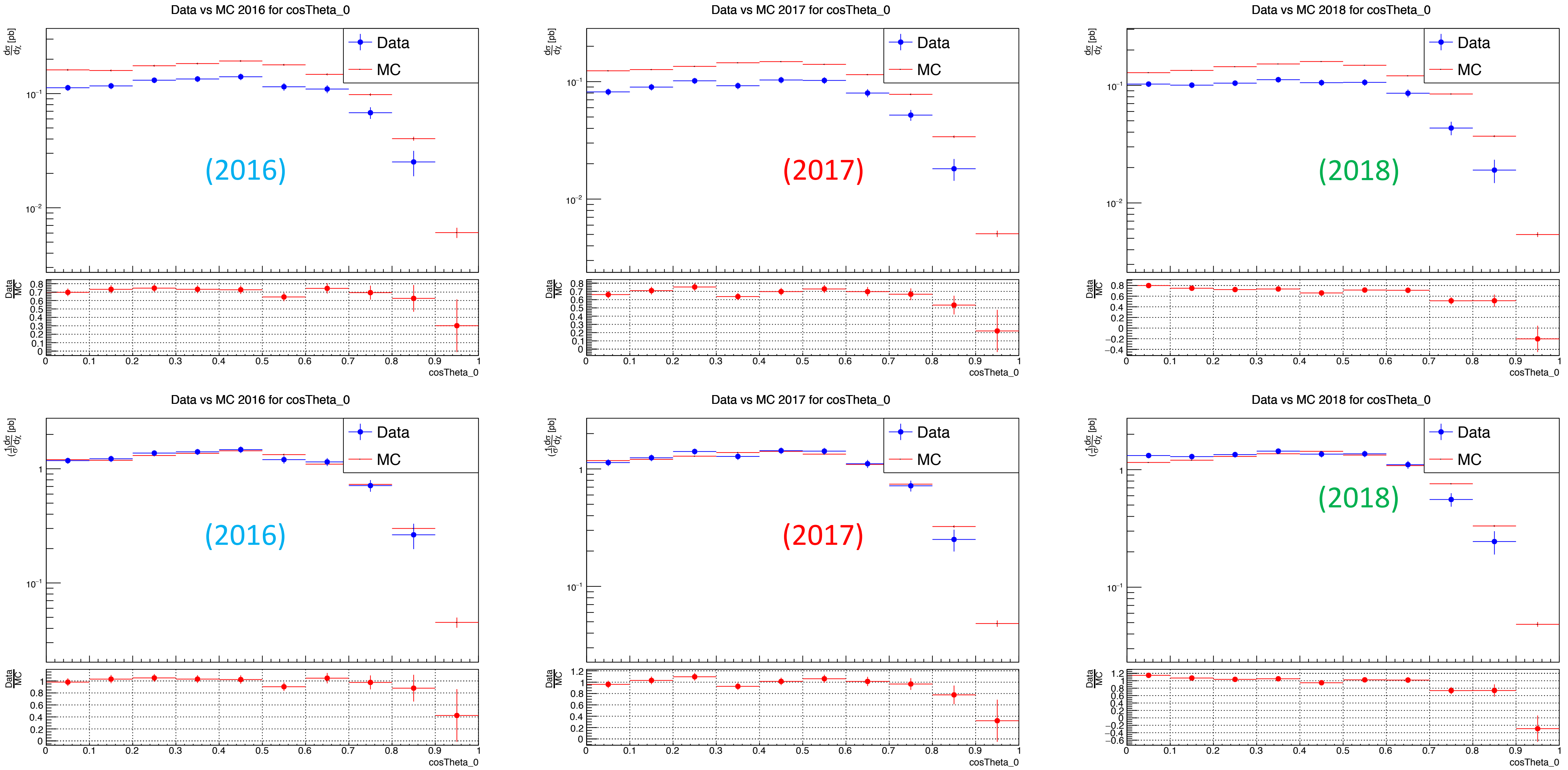




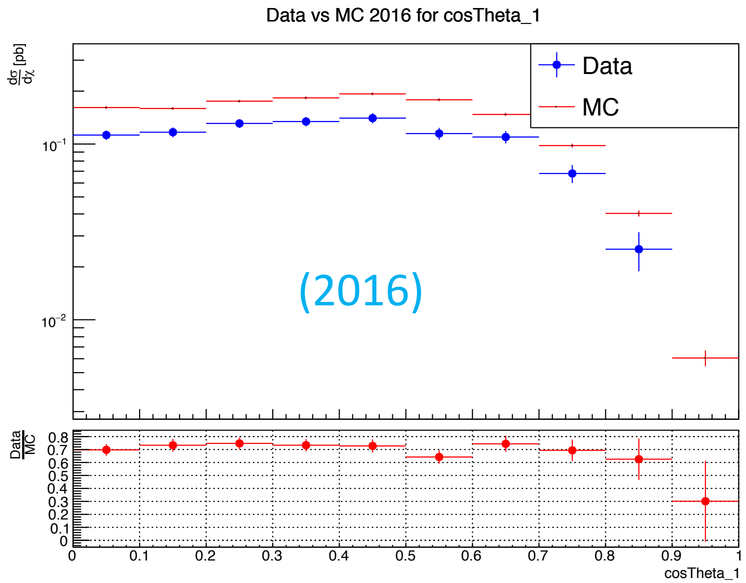


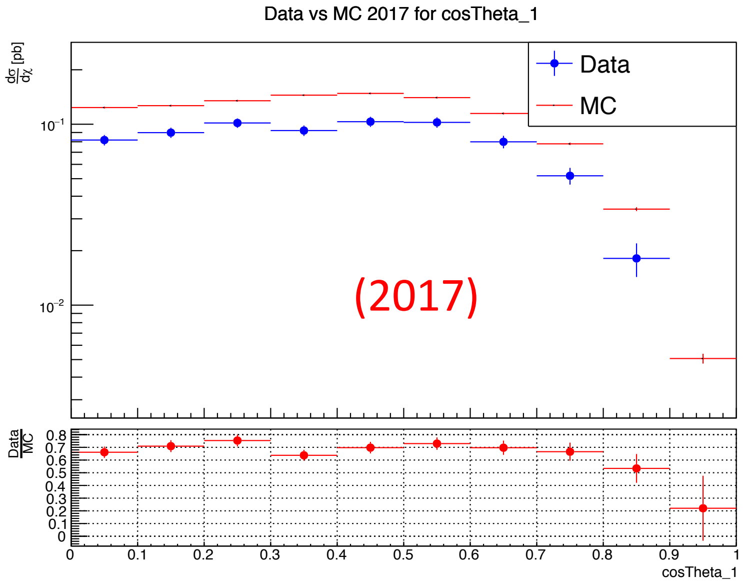


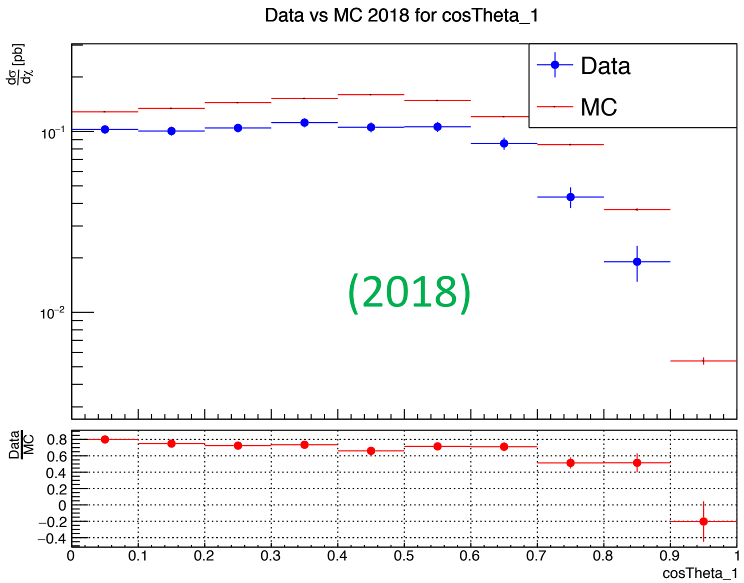
Fiducial Measurement

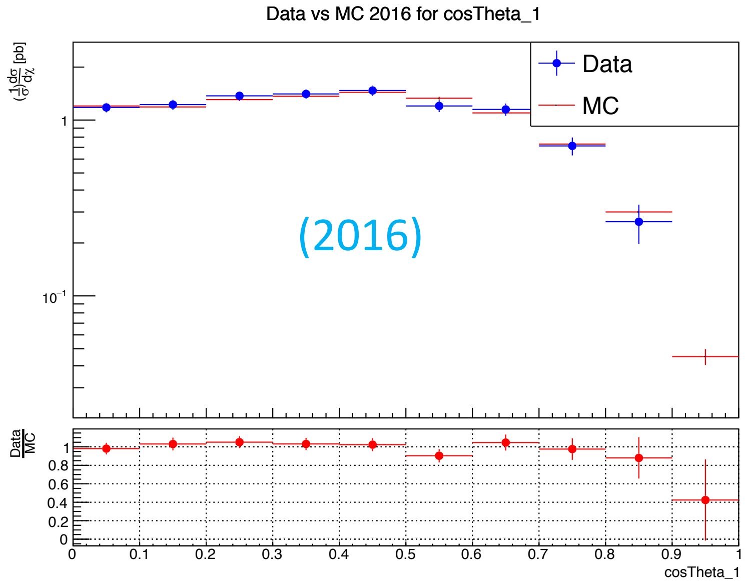


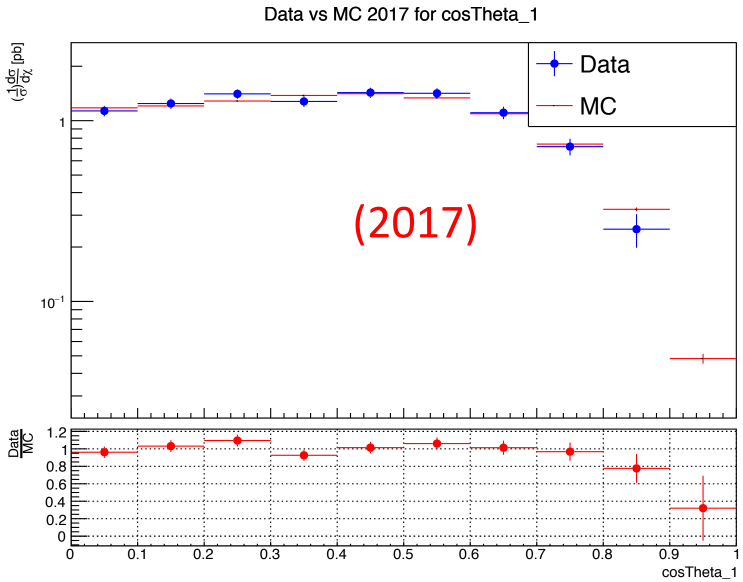
Fiducial Measurement

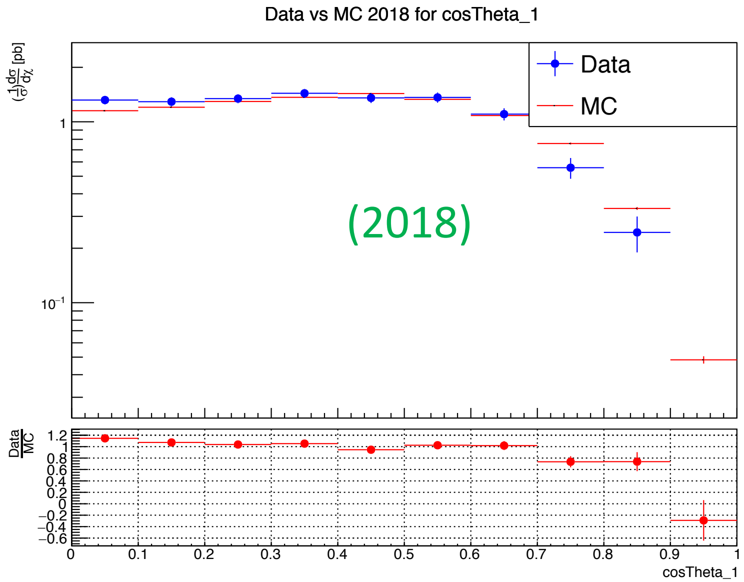




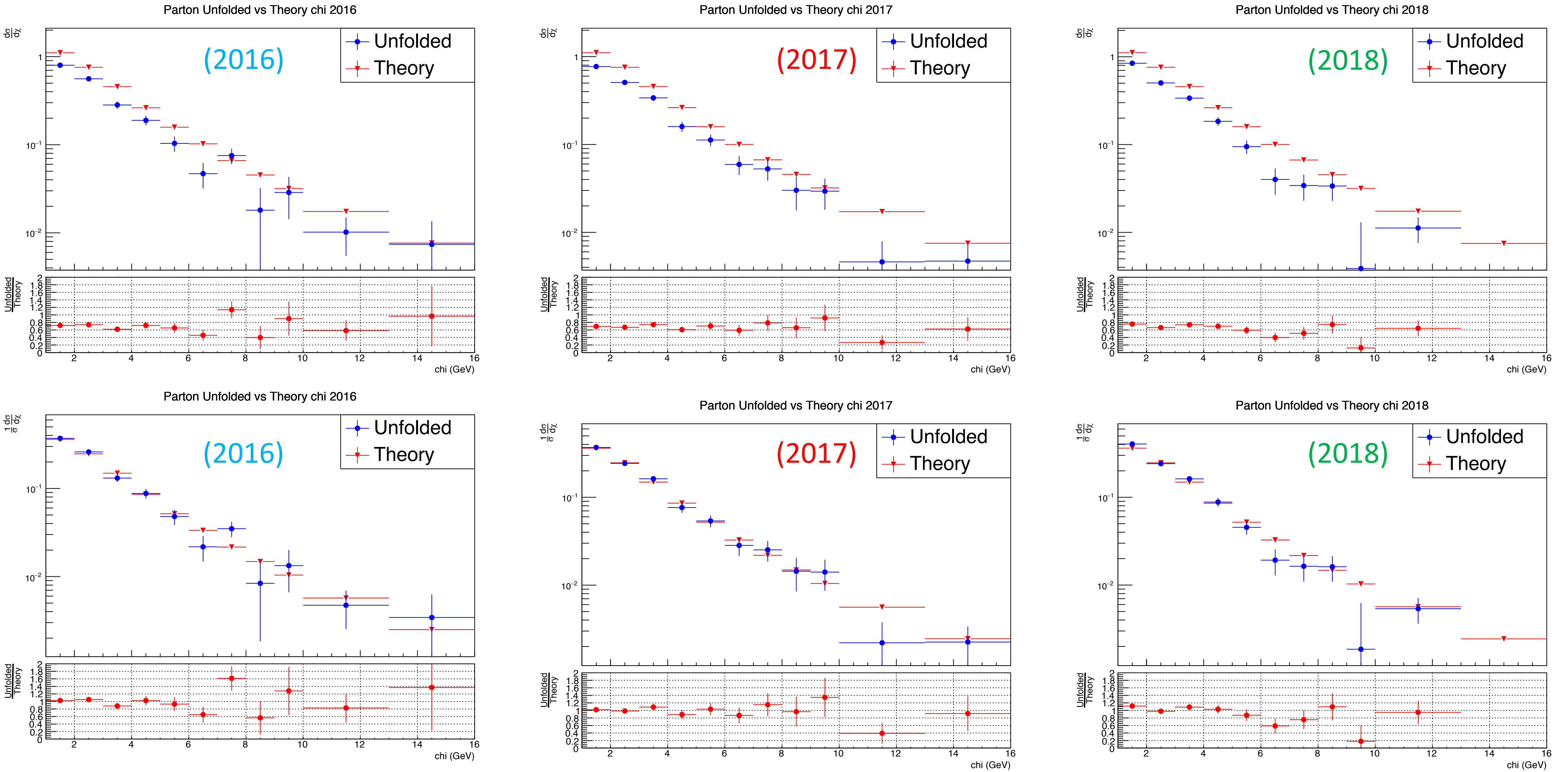




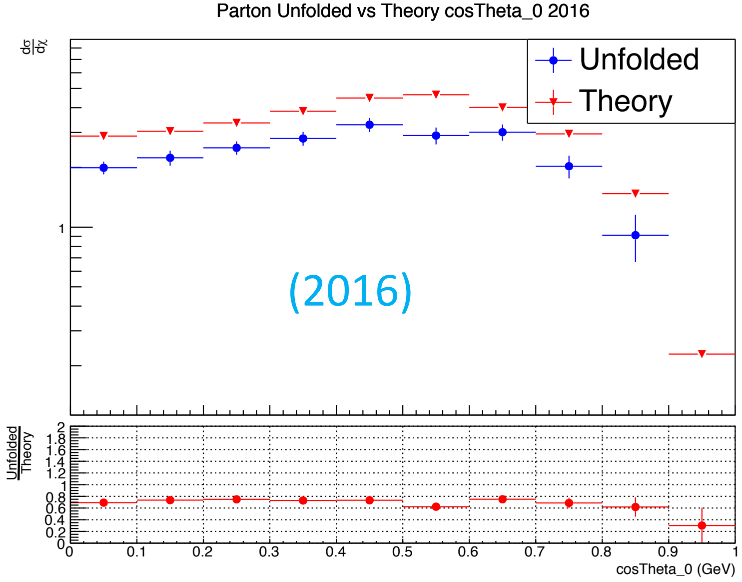


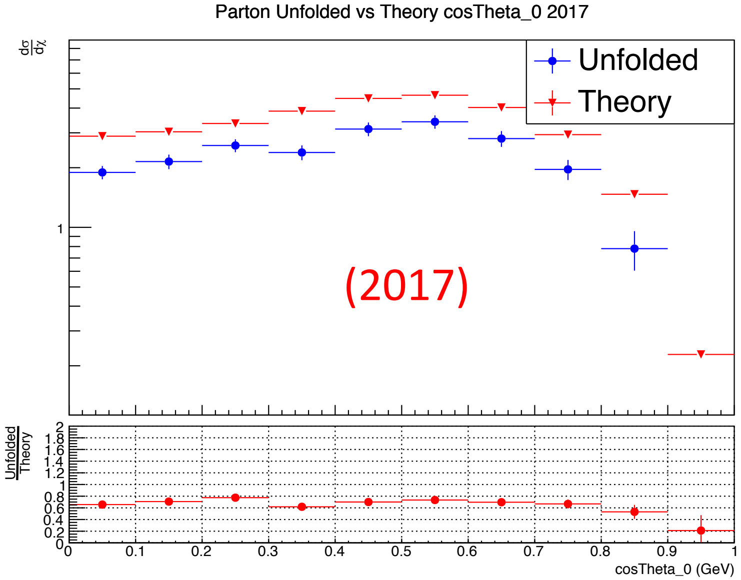


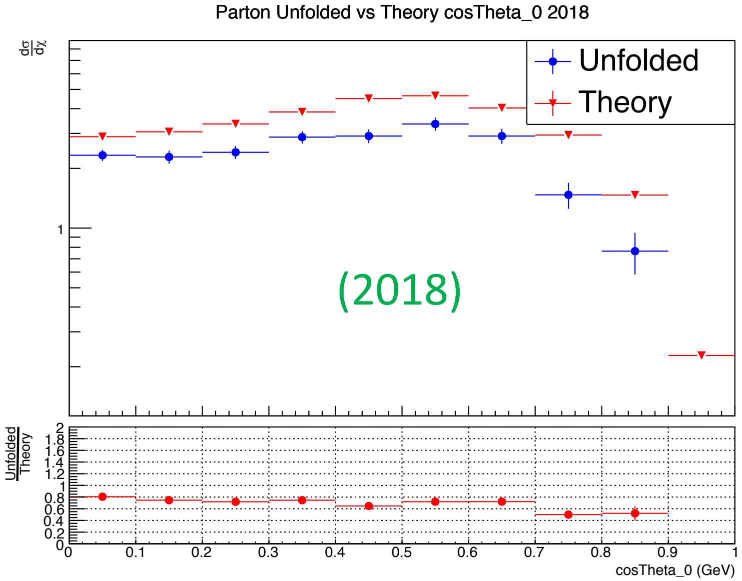
Parton Measurement

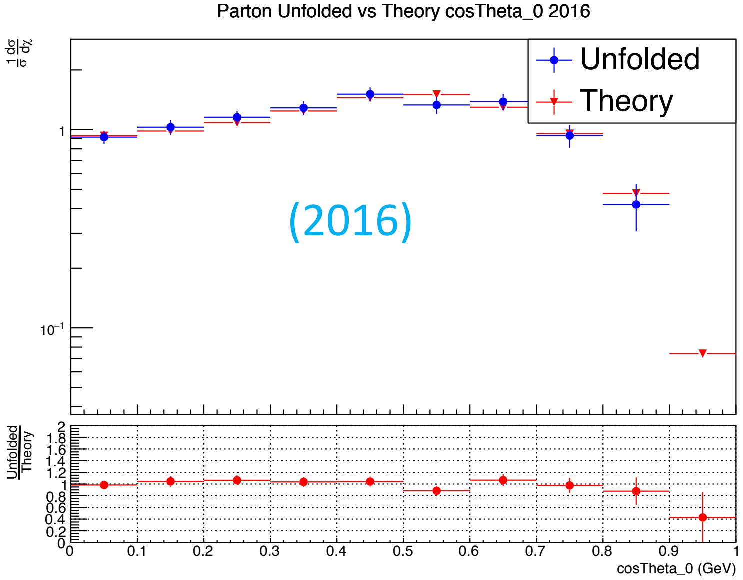


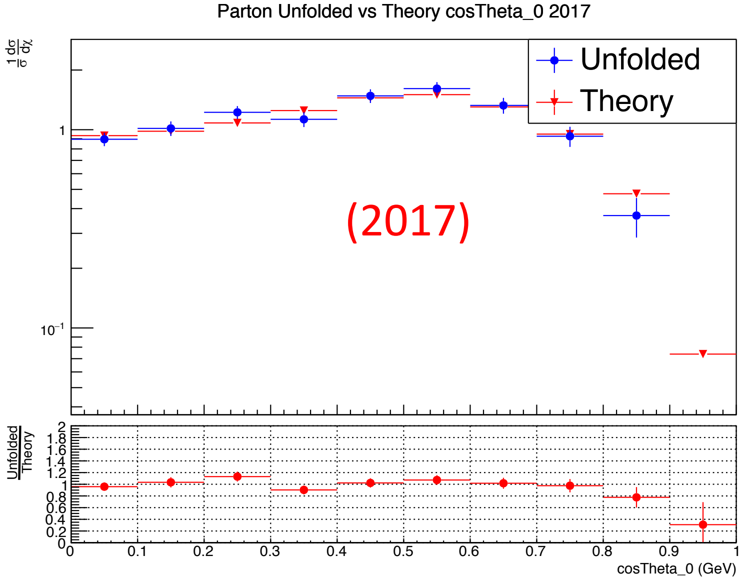
Parton Measurement

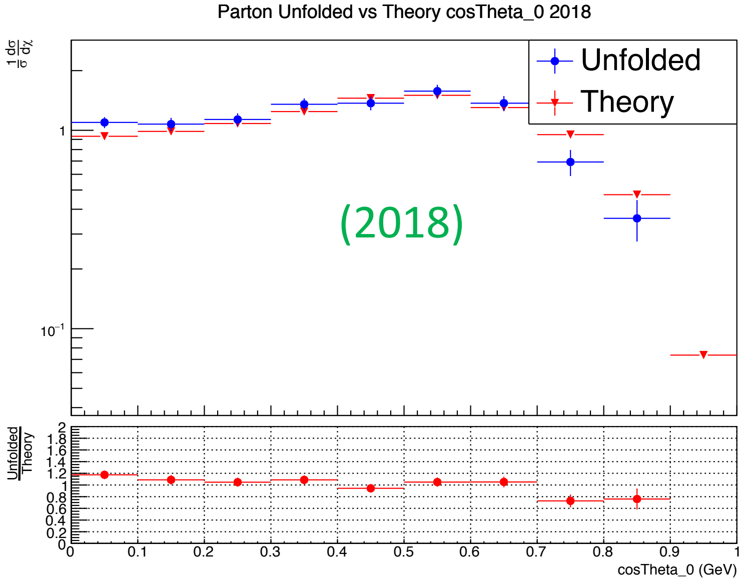




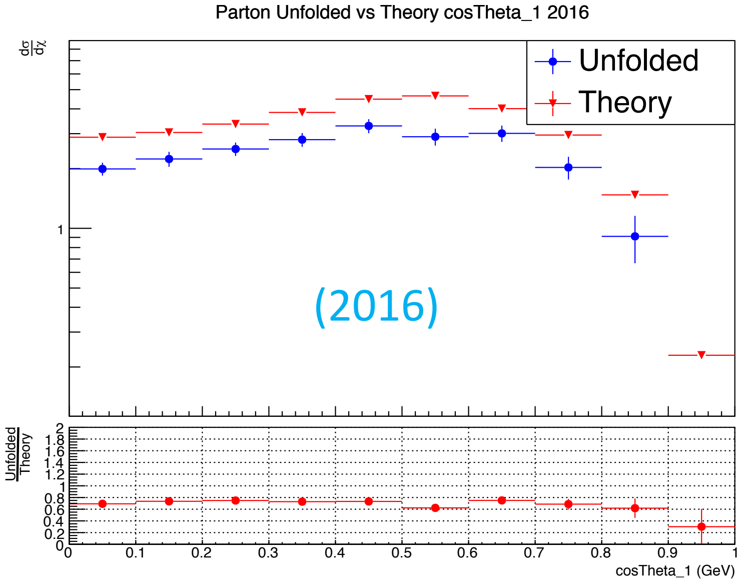


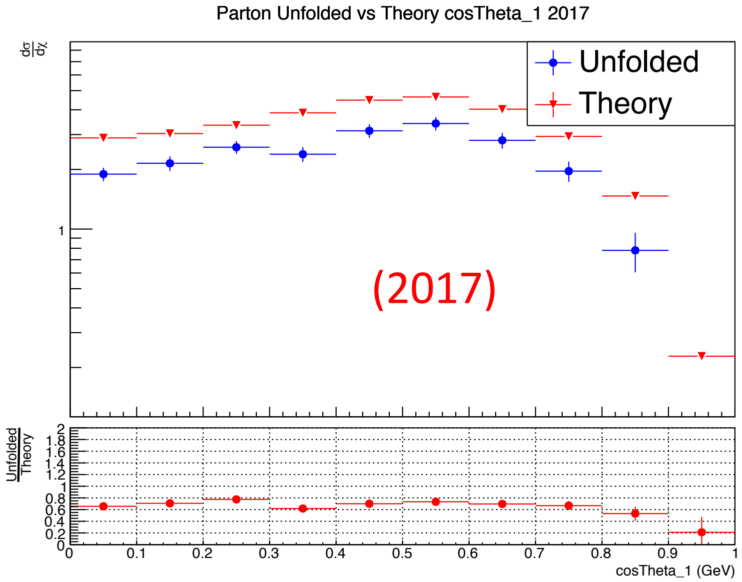


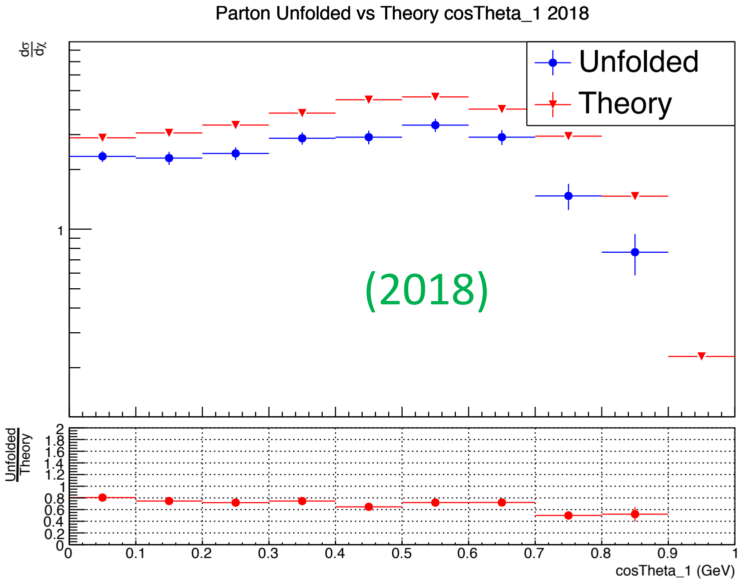


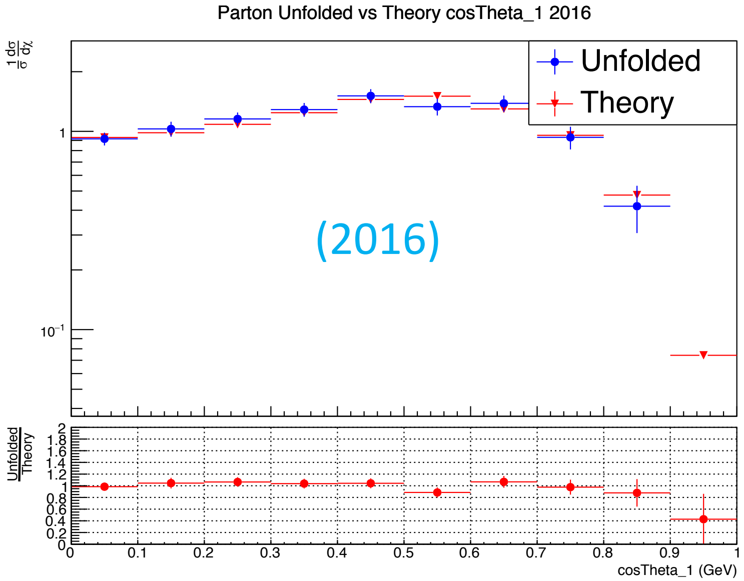


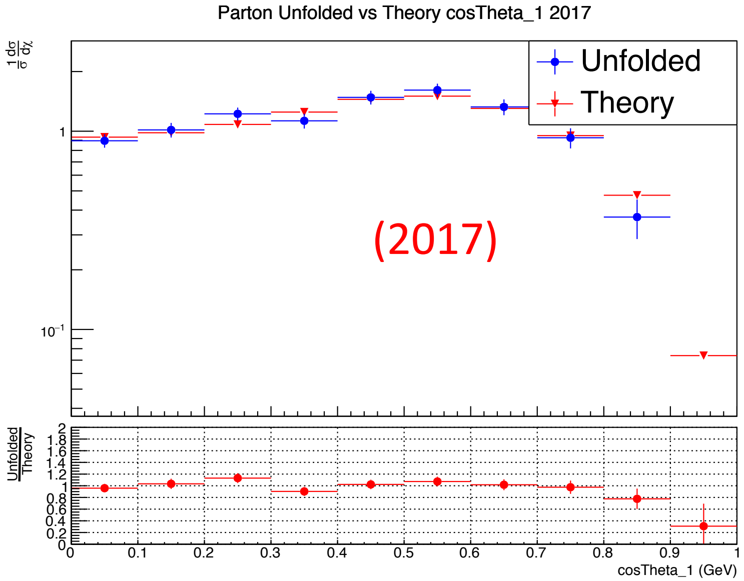
Parton Measurement

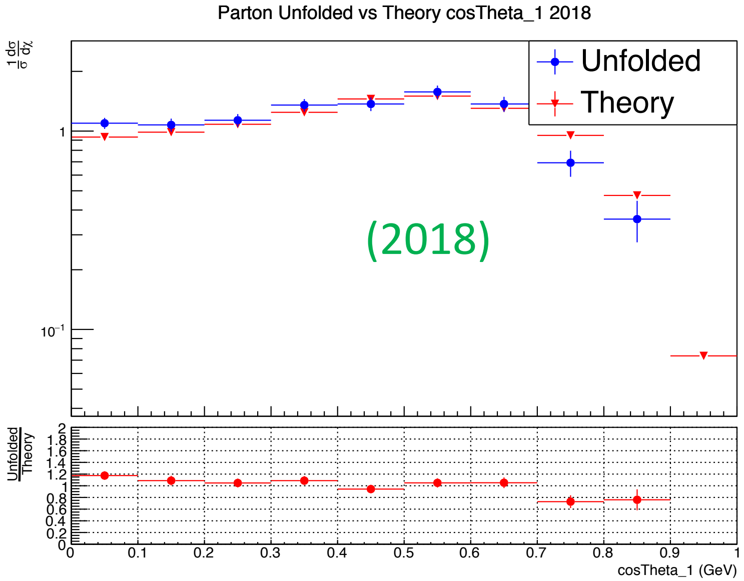












BACKUP

Signal Extraction

Measured dist from data

Fitted number of QCD events in SR_A

QCD shape taken from Data (CR)

$$S(x_{reco}) = D(x_{reco}) - C_{bkg}^{yield} N_{QCD}^{fit} C_{QCD}^{shape}(x_{reco}) Q(x_{reco}) - B(x_{reco})$$

Fiducial Yield

Transfer factor from SR_A to SR

QCD shape correction factor

Subdominant bkg shape and contribution (MC)

- Where x_{reco} is the respected variable of interest (ttbar mass, pt, rapidity, leading and subleading jetPt and |jetY|)

- We deploy a fit in the Signal Region (2btag) to extract the N_{QCD}^{fit}

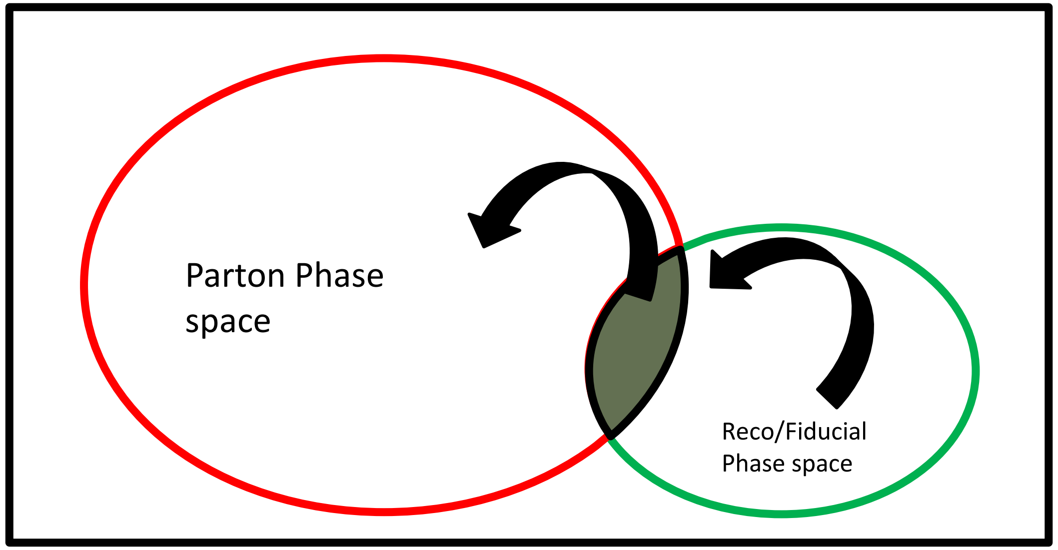
$$D(m^t)^{(i)} = N_{tt}^{(i)} T^{(i)}(m^t, k_{MassScale}, k_{MassResolution}) + N_{bkg}^{(i)} B(m^t)(1 + k_1 x) + N_{sub}^{(i)} O^{(i)}(m^t)$$

- Our data CR is contaminated from ttbar and subdominant bkgs which has to be dealt with.

Parton & Particle levels

Parton

Observable	Requirement
$p_T^{t,\bar{t}}$	> 400 GeV
$\lvert\eta^{t,\bar{t}}\rvert$	< 2.4
$m_{t\bar{t}}$	> 1000 GeV



Particle level Top Candidates

Observable	Requirement
N_{jets}	>1
$p_T^{jet1,2}$	> 400 GeV
$\lvert\eta^{jet1,2}\rvert$	< 2.4
$m_{SD}^{jet1,2}$	(120, 220) GeV
m_{jj}	> 1000 GeV

$$\frac{d\sigma_i^{\mathrm{unf}}}{dx} = \frac{1}{\mathcal{L}\cdot\Delta x_i}\cdot\frac{1}{f_{2,i}}\cdot\sum_j\left(R_{ij}^{-1}\cdot f_{1,j}\cdot S_j\right)$$

efficiency of the reco+true selection

migration matrix

reco efficiency of the reco+true selection

Unfolding: simple response matrix inversion w/o regularisation

Tag And Probe

- Top Tagger Scale Factors
  - Validation method to ensure that no SF’s are needed
  - From data we subtract QCD and Subdominant bkgs (MC) so that the data sample is pure

$$efficiency = \frac{\#\,(1\ jet\ pass\ baseline + Tight\ TopTagger\ Cut\ AND\ 1\ jet\ pass\ SR)}{\#\,(1\ jet\ pass\ baseline + Tight\ TopTagger\ Cut\ AND\ 1\ jet\ pass\ only\ baseline)}$$

  - Randomization: Randomly select leading/subleading jet to use as tag or probe to avoid pT bias
  - Divide the phase space into pT regions based on the topTagger categories: [400-600] GeV, [600-800] GeV, [800-Inf] GeV

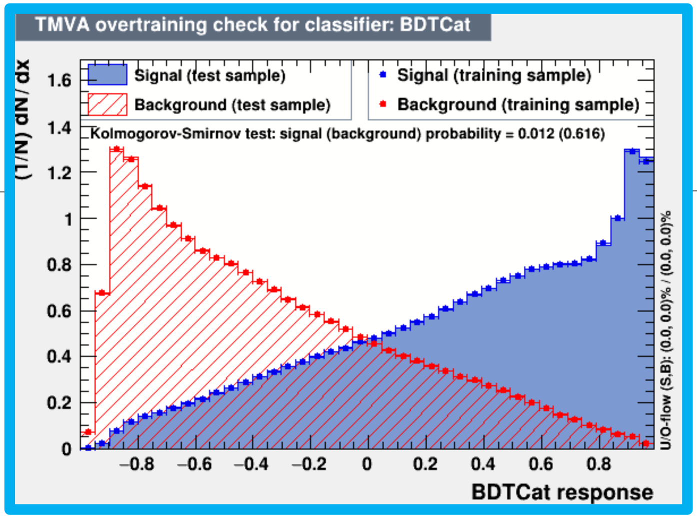
Year	Type of File	DAS
2016	TT Mtt 700-1000	/TT_Mtt-700to1000_TuneCUETP8M2T4_13TeV-powheg-pythia8/RunIISummer16MiniAODv3-PUMoriond17_94X_mcRun2_asymptotic_v3-v2/MINIAODSIM
	TT Mtt 1000-Inf	/TT_Mtt-1000toInf_TuneCUETP8M2T4_13TeV-powheg-pythia8/RunIISummer16MiniAODv3-PUMoriond17_94X_mcRun2_asymptotic_v3-v2/MINIAODSIM
	TT Nominal	/TT_TuneCUETP8M2T4_13TeV-powheg-pythia8/RunIISummer16MiniAODv3-PUMoriond17_94X_mcRun2_asymptotic_v3-v1/MINIAODSIM
2017	TT Mtt 700-1000	--
	TT Mtt 1000-Inf	--
	TT Nominal Hadronic	/TTToHadronic_TuneCP5_13TeV-powheg-pythia8/RunIIFall17MiniAODv2-PU2017_12Apr2018_94X_mc2017_realistic_v14-v1/MINIAODSIM
	TT Nominal Semilepton	/TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8/RunIIFall17MiniAODv2-PU2017_12Apr2018_94X_mc2017_realistic_v14-v2/MINIAODSIM
	TT Nominal Dilepton	TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIIFall17MiniAODv2-PU2017_12Apr2018_94X_mc2017_realistic_v14-v2/MINIAODSIM
2018	TT Mtt 700-1000	--
	TT Mtt 1000-Inf	--
	TT Nominal Hadronic	/TTToHadronic_TuneCP5_13TeV-powheg-pythia8/RunIIAutumn18MiniAOD-102X_upgrade2018_realistic_v15-v1/MINIAODSIM
	TT Nominal Semilepton	/TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8/RunIIAutumn18MiniAOD-102X_upgrade2018_realistic_v15-v1/MINIAODSIM
	TT Nominal Dilepton	/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8/RunIIAutumn18MiniAOD-102X_upgrade2018_realistic_v15-v1/MINIAODSIM

Overview: Discriminator, Efficiency and Acceptance

The discriminator is a BDT trained individually for 2016, 2017 and 2018

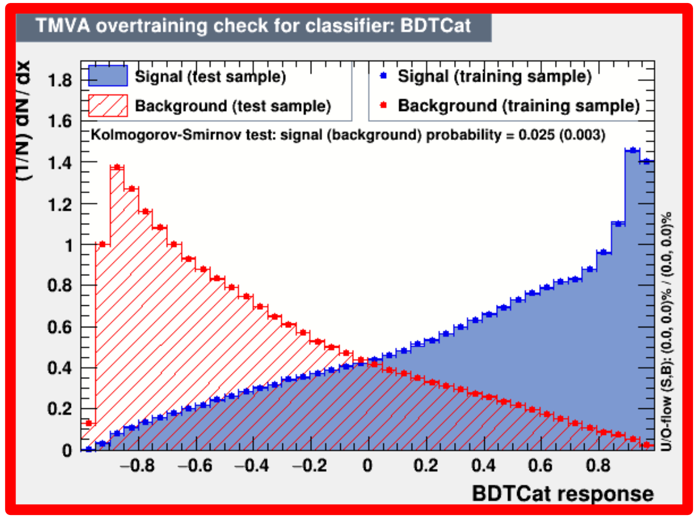
Category training: split the sample in categories based on Pt

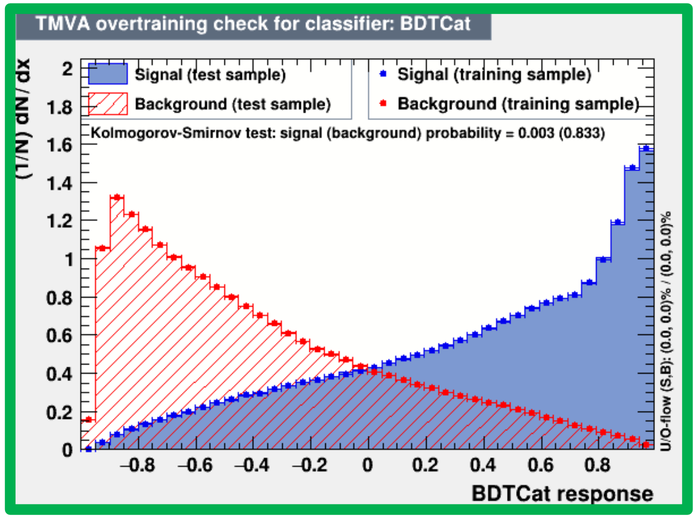
- Bins:
    - [400, 600] GeV
    - [600, 800] GeV
    - [800, 1200] GeV
    - [1200, inf) GeV
- BDT, used variables:
    - Leading and Sub-leading subjet mass
    - N-Subjetiness variables (tau1, tau2, tau3)
    - fraction of the jetPt over the total pt sum of the event.
    - Energy correlation functions (ecfB1N2,ecfB1N3, ecfB2N2, ecfB2N3)
- BDT Output consistency for the 3 years
- Calculation of Efficiency and acceptance for each year
    - We choose the WP's for each year so that the leading jet p_T efficiency is similar for all years



2016

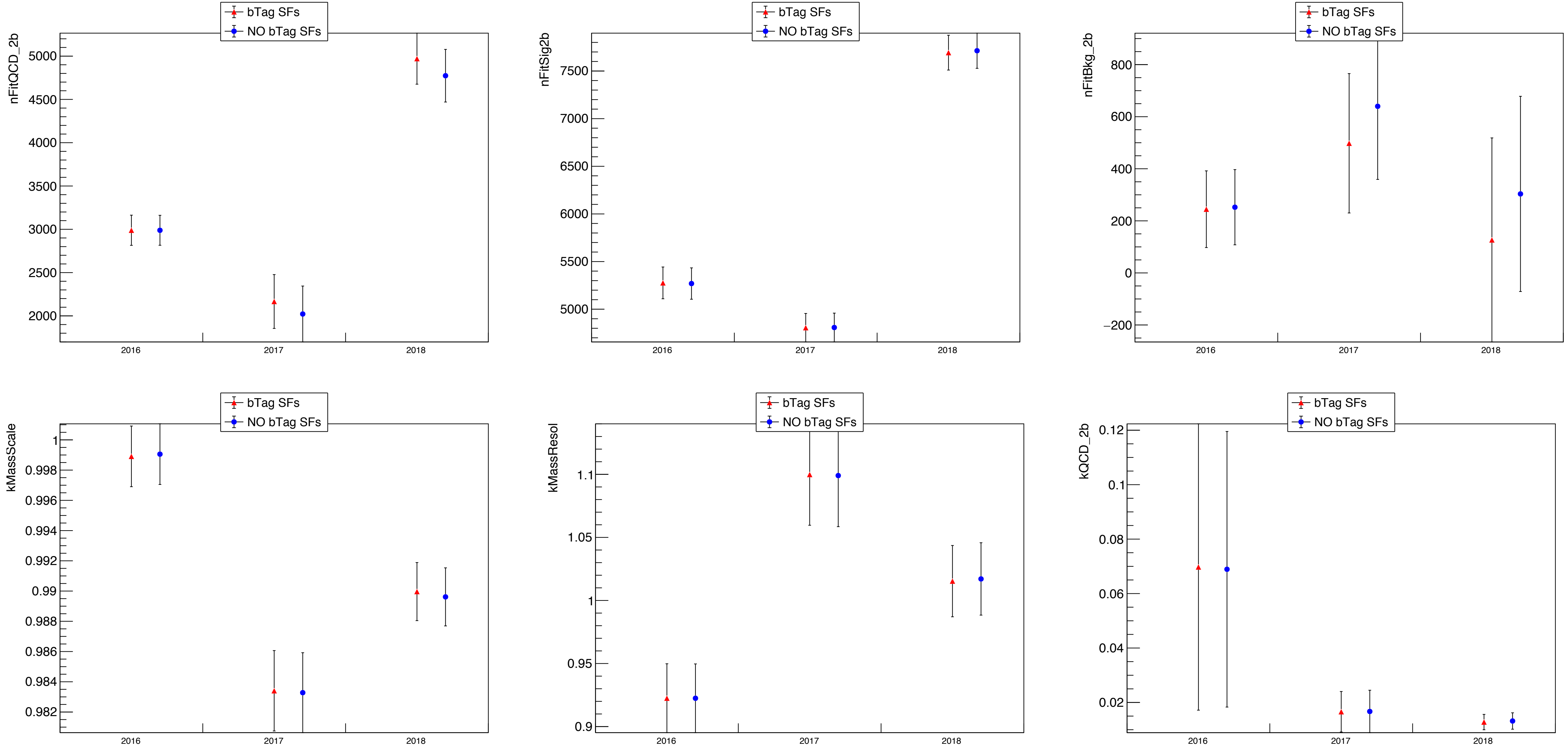
2017





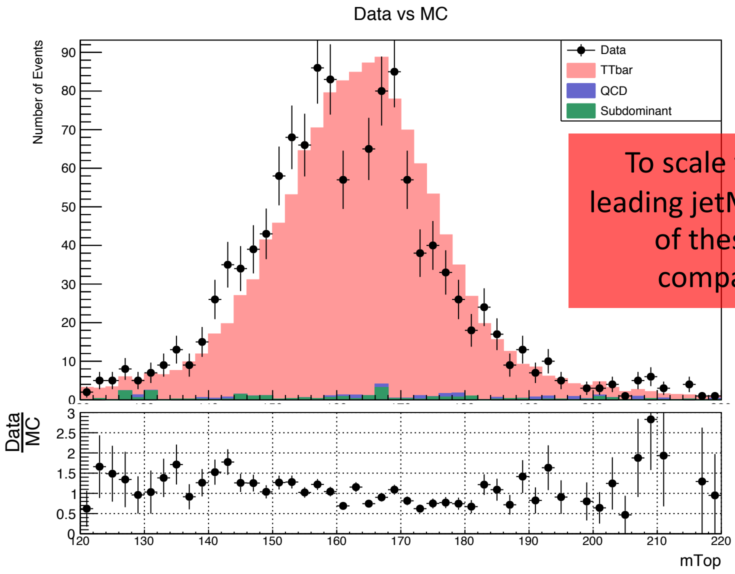
2018

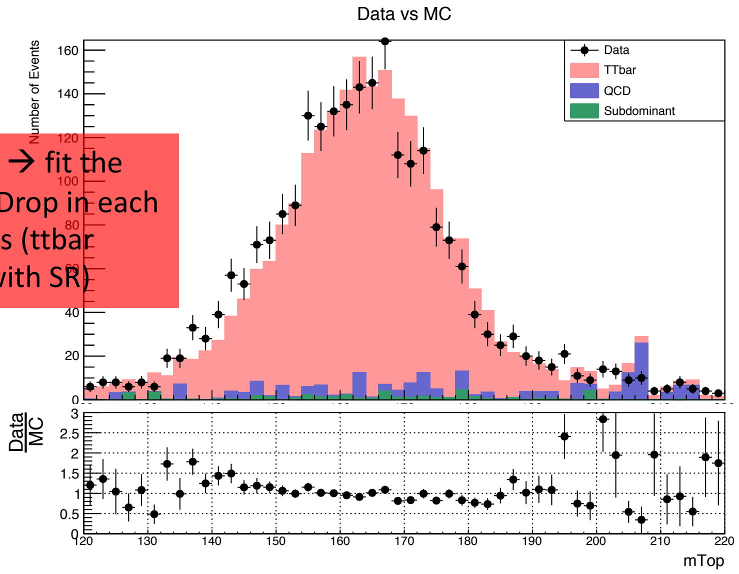
Fit Params Results Comparison



TagAndProbe Efficiency Plots

Tight TopTagger + SR TopTagger





To scale the ttbar → fit the leading jetMassSoftDrop in each of these regions (ttbar compatible ~ with SR)

Tight TopTagger + Probe