Progress Report

21/12/2020

Georgios Bakas

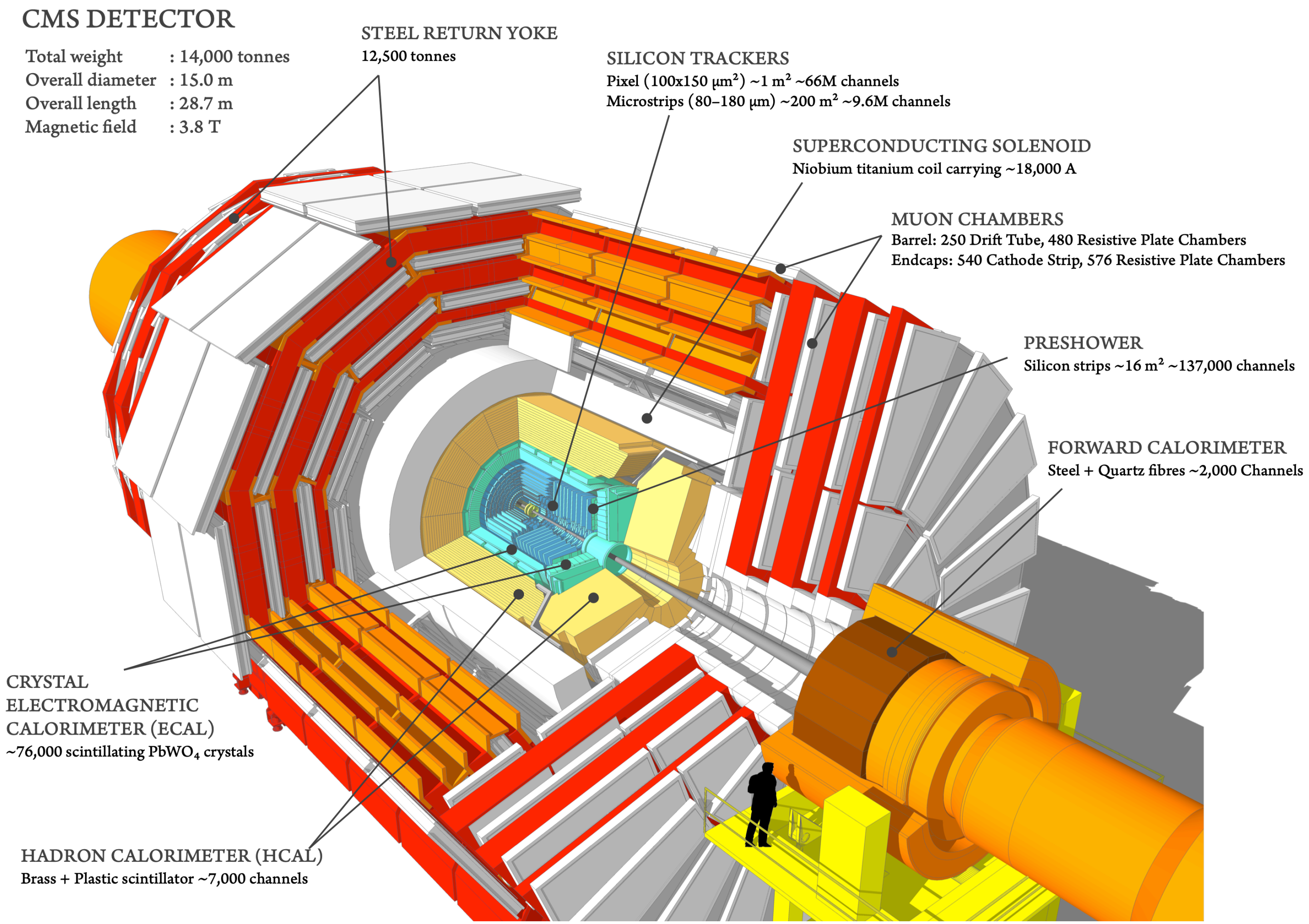
Outline

- CERN and the CMS Experiment
- Detector Control System
- Data Analysis
- Detector Upgrades
- RD51
- Conclusions

CERN – LHC

- CERN, European laboratory for particle and nuclear physics.
- Various experiments.
- Large Hadron Collider (LHC)
  - 27Km circumference
  - Colliding Protons
  - Design energy up to 14 TeV.
- Beams collide at 4 predefined points.
- 4 detectors/experiments.
- ALICE, ATLAS, CMS, LHCb

CMS Experiment

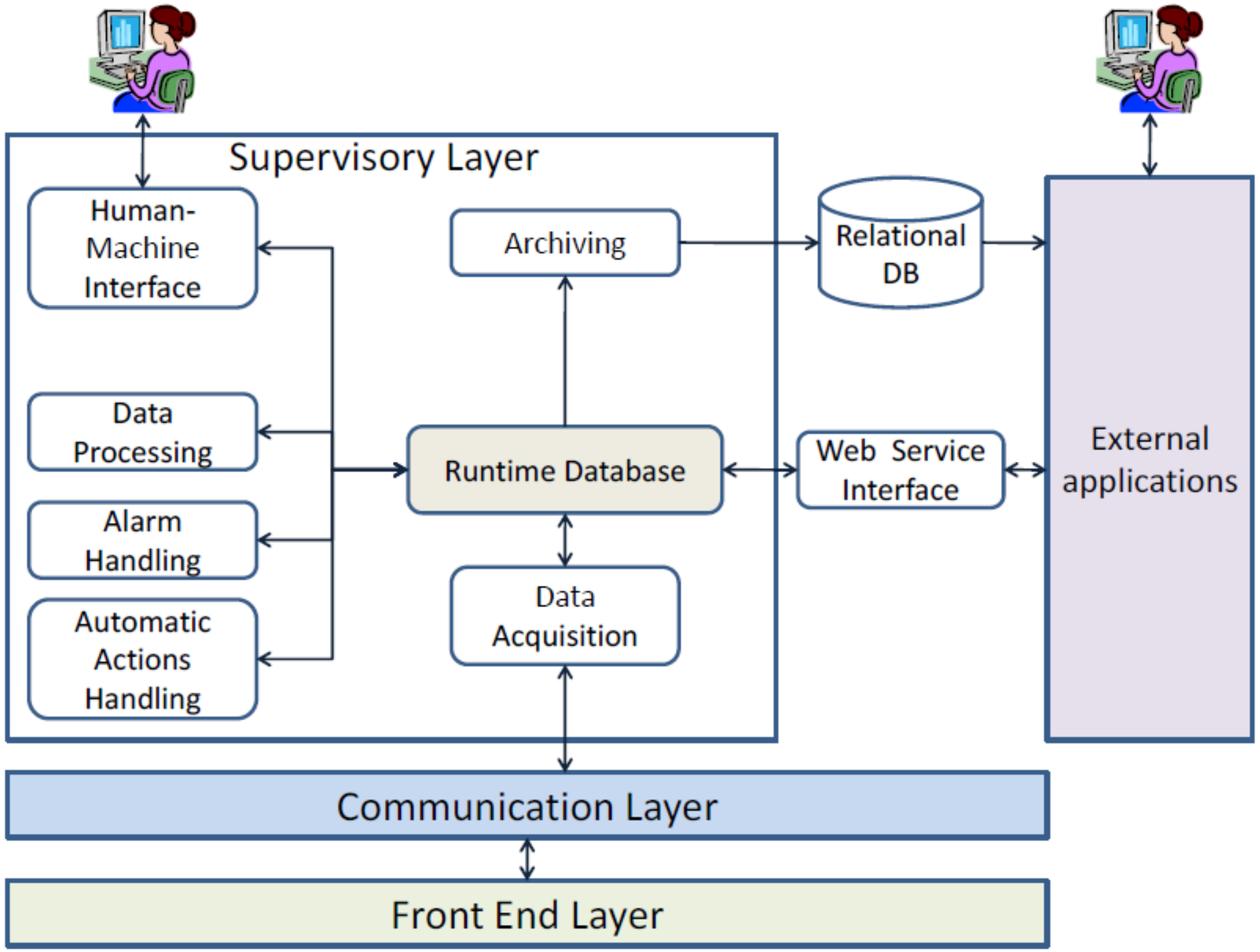


Detector Control System

- Control the detector
- Monitor the conditions under which it operates
- Sophisticated and sensitive equipment (more than 6 million parameters)
- Take necessary actions
- Use the collected data as a first indication for determining the quality of the physics data collected.

Control Systems

- SCADA system (Supervisory Control And Data Acquisition)
- Monitor and Control a remote process
- Factories, airports, physics experiments
- Toolkits
- Tools for developing the supervisory layer
- Ability to connect to hardware (communication layer, drivers)



Control Systems at CERN

Experiment's Framework

JCOP FW

CERN

ALICE

ATLAS EXPERIMENT

CMS

LHCb

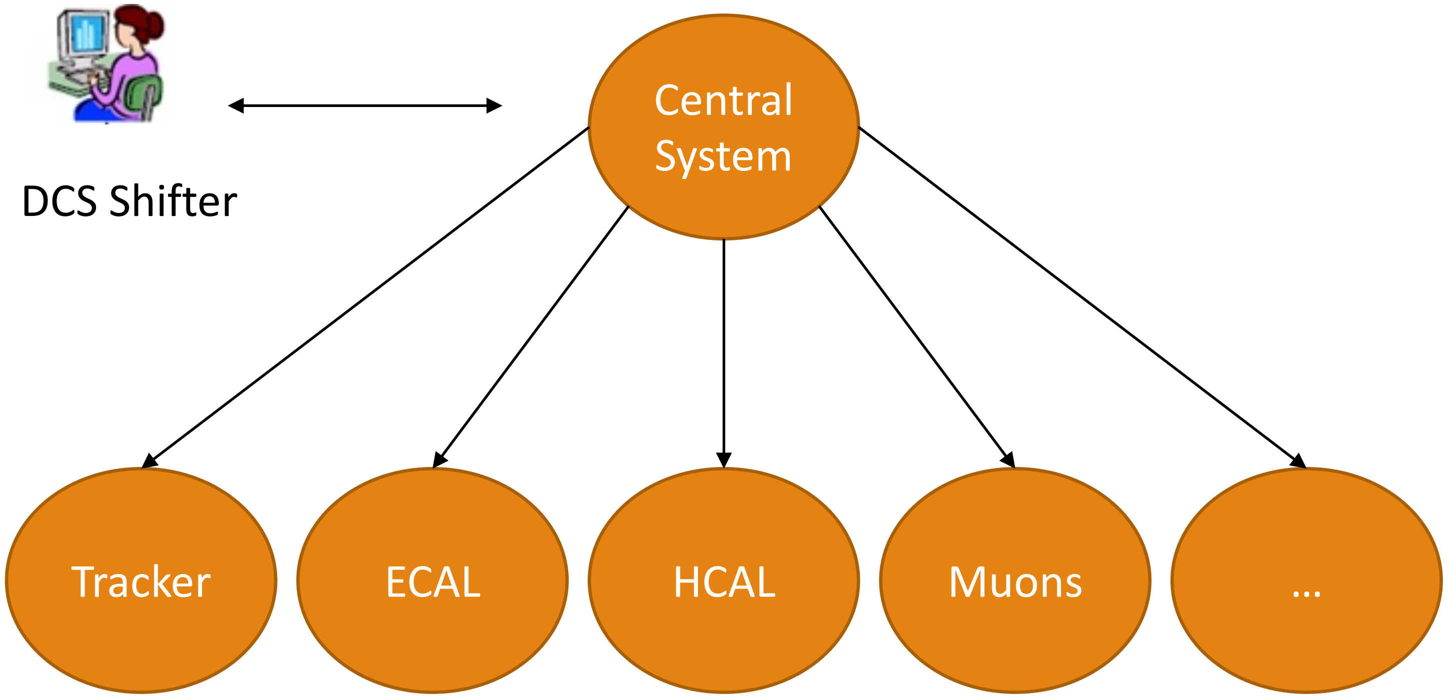
- Distributed Systems
- Runtime Database
- OS Independent
- Native language, GEDI
- Embedded drivers

WinCC_OA

SIEMENS

The CMS Detector Control System

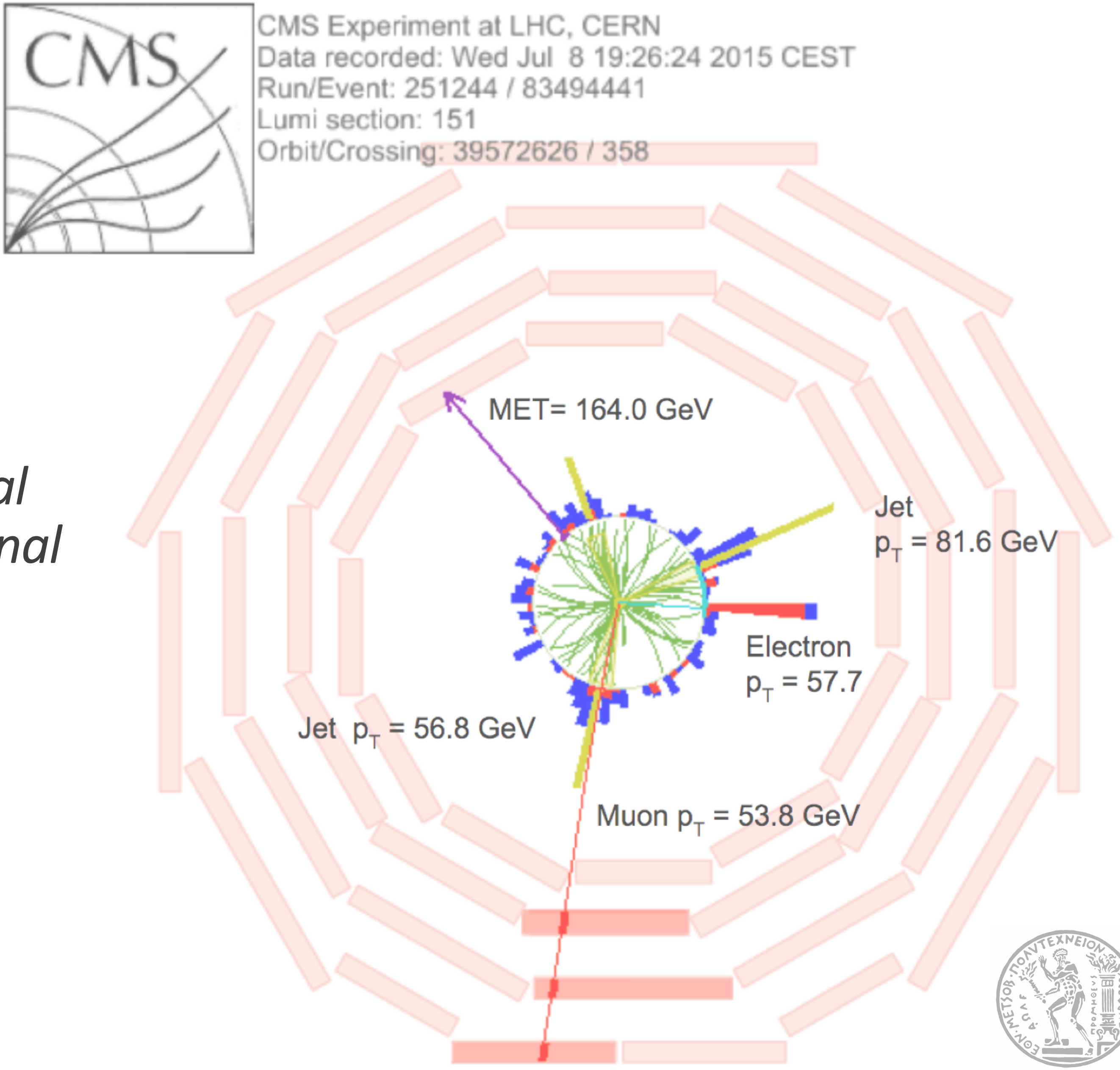
- More than 35 individual systems
- Minimal systems modified by a bunch of JCOP-like CMS components
- Centralized
- Central DCS responsible for the maintenance of all the projects



Role in the DCS

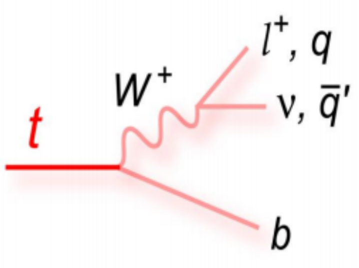
- Members of the CMS central DCS team
- System administration
- Maintenance and improvement of the existing system
- Developed/improved various applications used either in the control room by the operators or by subsystem experts
- Support in the CMS experiment in control system manners. Problem fixing or suggestions for future developments

Measurement of the top-anti-top differential production cross section in the all-hadronic final state using the full Run II proton-proton collision data at $\sqrt{s}$ = 13 TeV

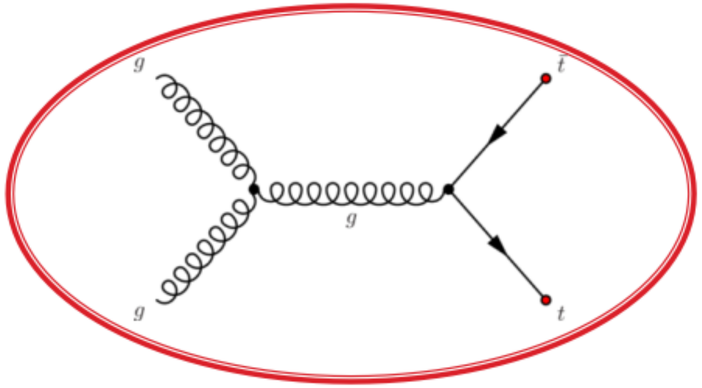


Top Quark -- Top AntiTop system decay

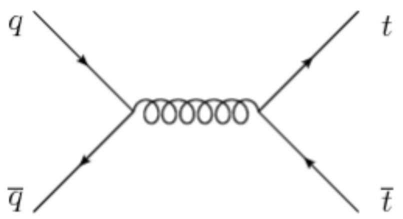
- Mass: $172.44 \pm 0.13\ \frac{GeV}{c^2}$
- Top Quark decay:
  - $t \rightarrow W^+ + b\ (\bar{t} \rightarrow \mathrm{W}^- + \bar{b})$



- Top quark pair production
  - $q + \bar{q} \rightarrow t + \bar{t}$
  - $g + g \rightarrow t + \bar{t}$

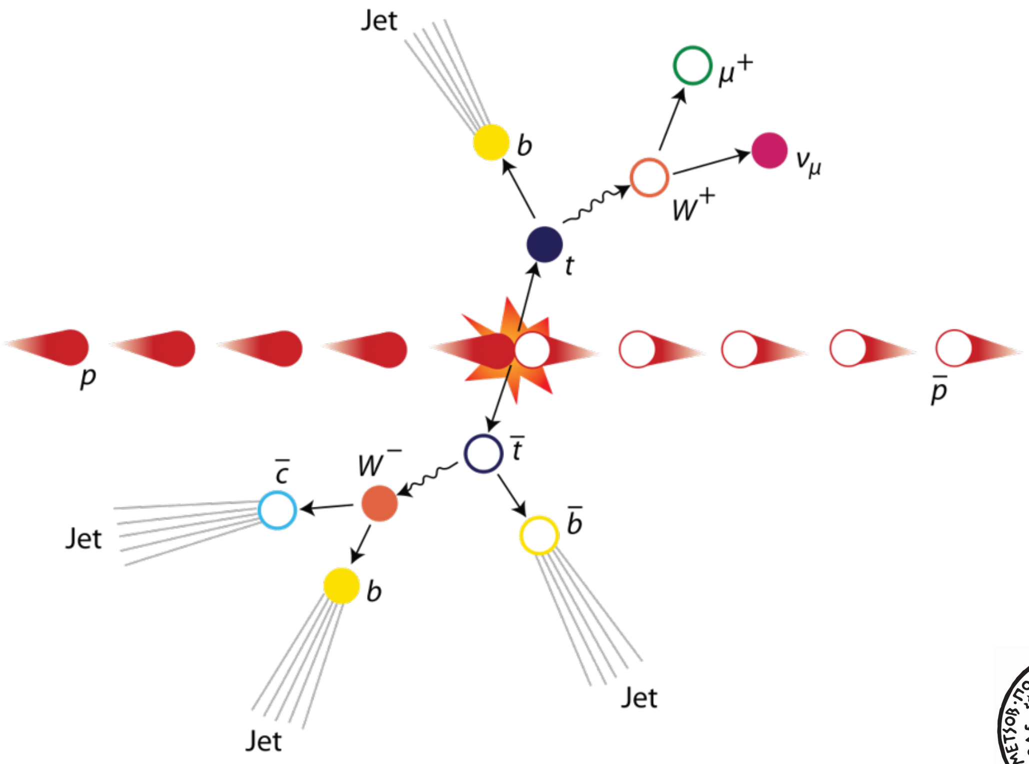


Gluon Fusion is dominant at LHC



(b)

1. $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}b\, q''\bar{b}\bar{q}''$ (45.7 %) → hadronic
2. $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'b\, l^-\bar{\nu}_l\bar{b} + l^+\nu_l b q''\bar{q}'''\bar{b}$ (43.8 %) → semileptonic
3. $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow l^+\nu_l b l'\bar{\nu}'\bar{b}$ (10.5 %) → dileptonic



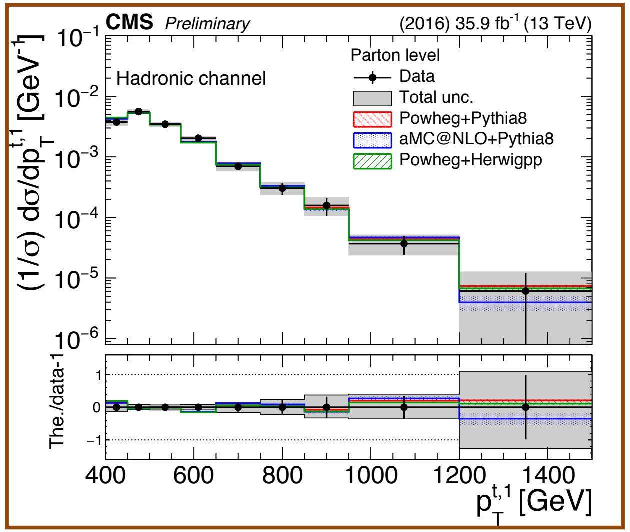
Boosted Jets

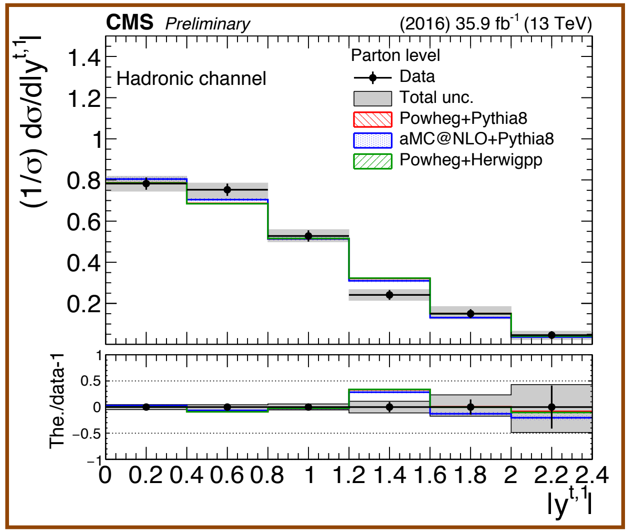
- Boosted Jets are jets with high p_T (> 400 GeV)
- Aim is the reconstruction of two big jets that contain the decay products of the top-antitop quark pair decay

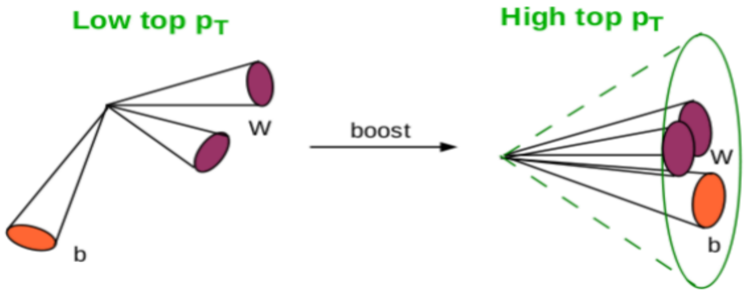
- Motivation
  - With resolved hypothesis we measure the top pair cross section up to ~500GeV
  - There is an interesting discrepancy with theory (p_T slope)
  - In order to see what happens in bigger p_T 's → boosted

- Why Boosted jets?
  - Single "fat" jet: No combinatorial background
  - At high top p_T the hadronic decay is easier to reconstruct than the leptonic

- In order to identify boosted jets
  - Use of sophisticated reconstruction techniques to identify the substructure within the jet
  - SoftDrop technique to eliminate soft contributions





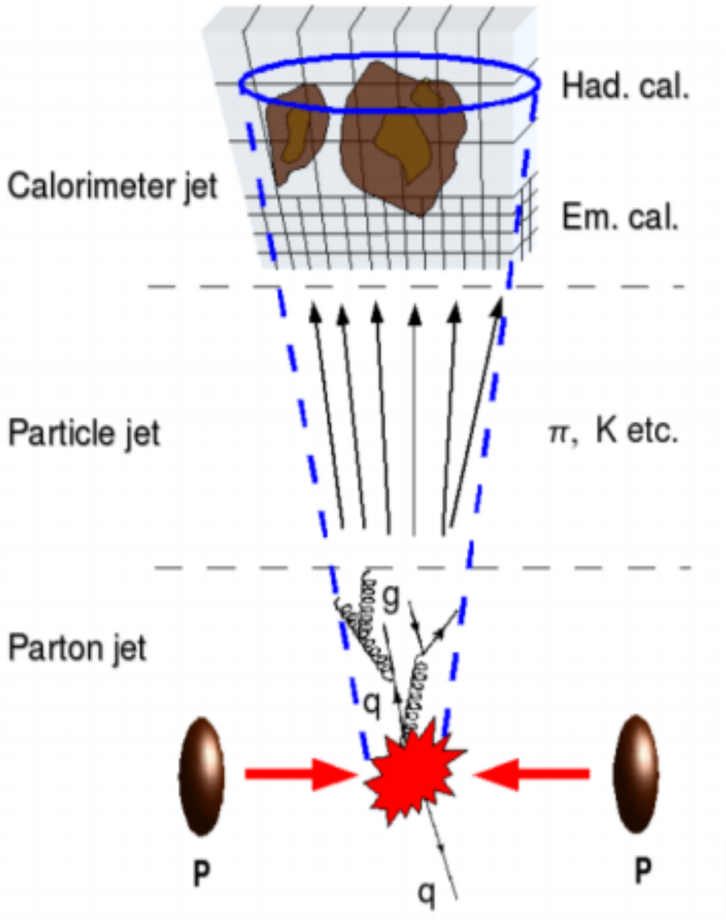


Analysis Overview

Levels:
- Parton level: collision products
- Particle level: Hadronization
- Reconstructed level: collected data.

You use the unfolding technique to go from the reconstructed level to the particle and parton levels.

$$y = Ax \Rightarrow x = A^{-1}y$$



Variables of interest:
- leading and subleading top p_T
- leading and subleading top $|y|$
- mass of the ttbar system
- p_T of the ttbar system
- rapidity of the ttbar system
- leading and subleading $\cos(\theta^*)$
- $\chi = e^{|y^*|} = e^{|y_1 - y_2|}$

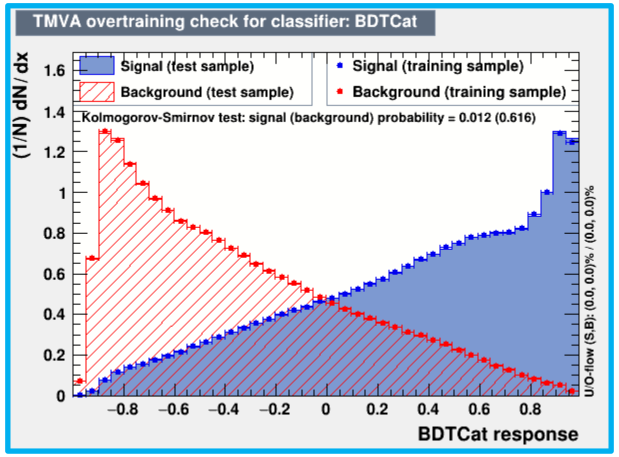
Overview

- Baseline Reconstructed level cuts:
  - nJets > 1, nLeptons = 0, Dijet mass (mJJ) > 1000 GeV
  - Leading and Subleading jet p_T > 400 GeV
  - Leading and Subleading absolute jet eta $|\eta|<2.4$
- Btagging selection:
  - How many subjets are b-tagged
- In house developed top Tagger:
  - Boosted Decision Tree
  - **Working points** (2016: 0.2, 2017: 0.0, 2018: 0.1)

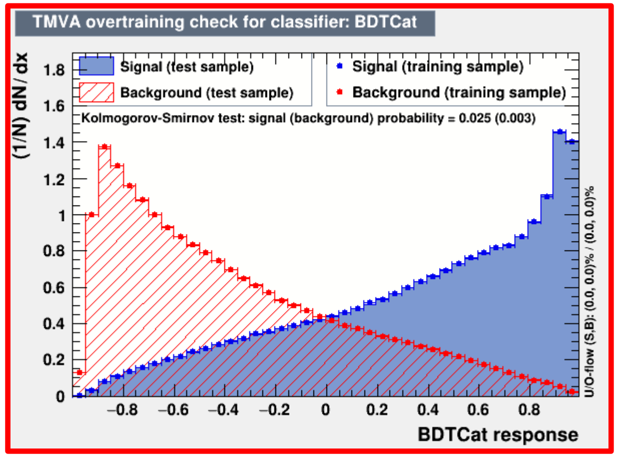
Goal is to Unfold to the Parton And Particle Levels

- Data driven method for QCD subtraction

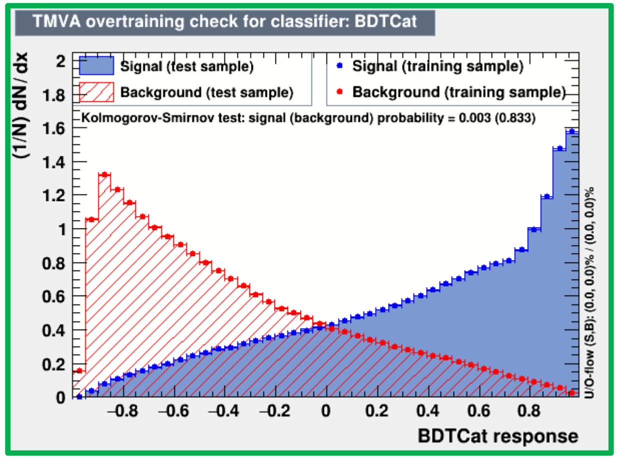
2016



2017

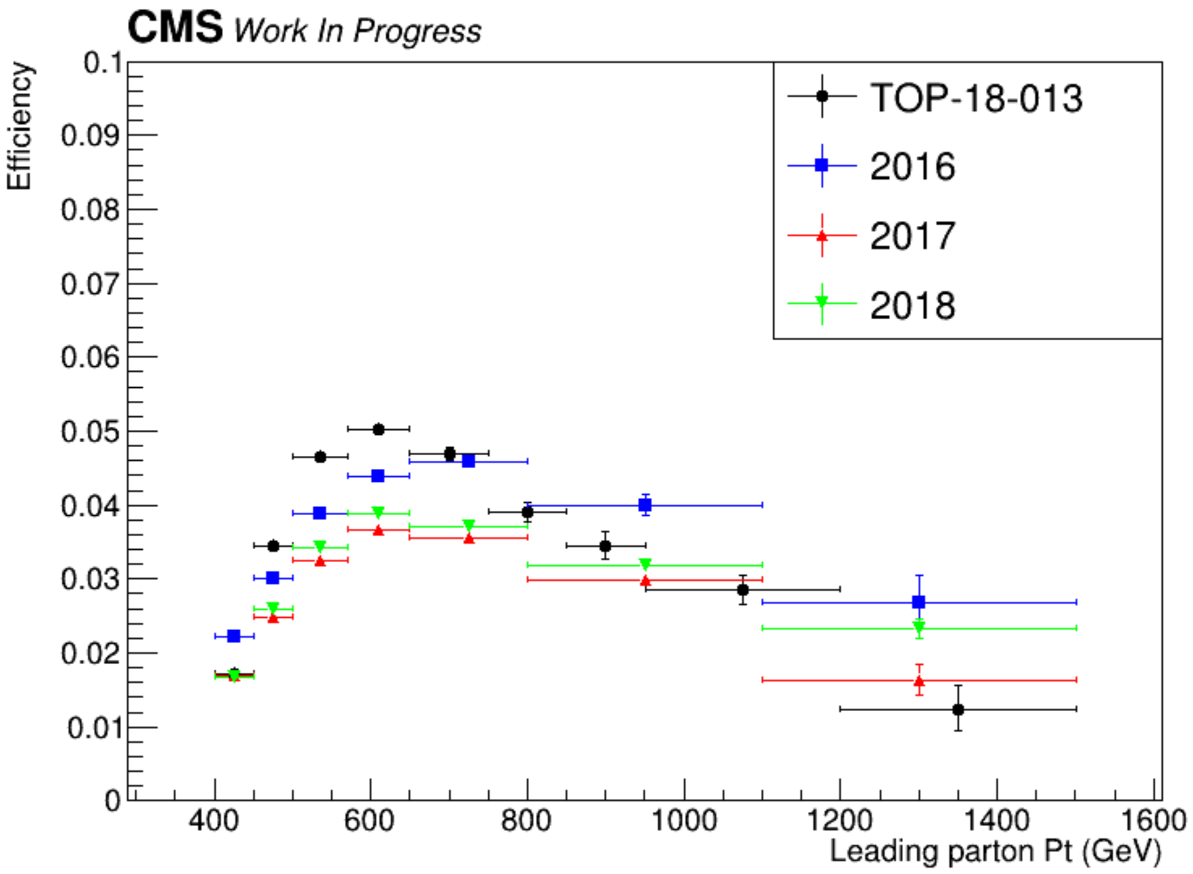


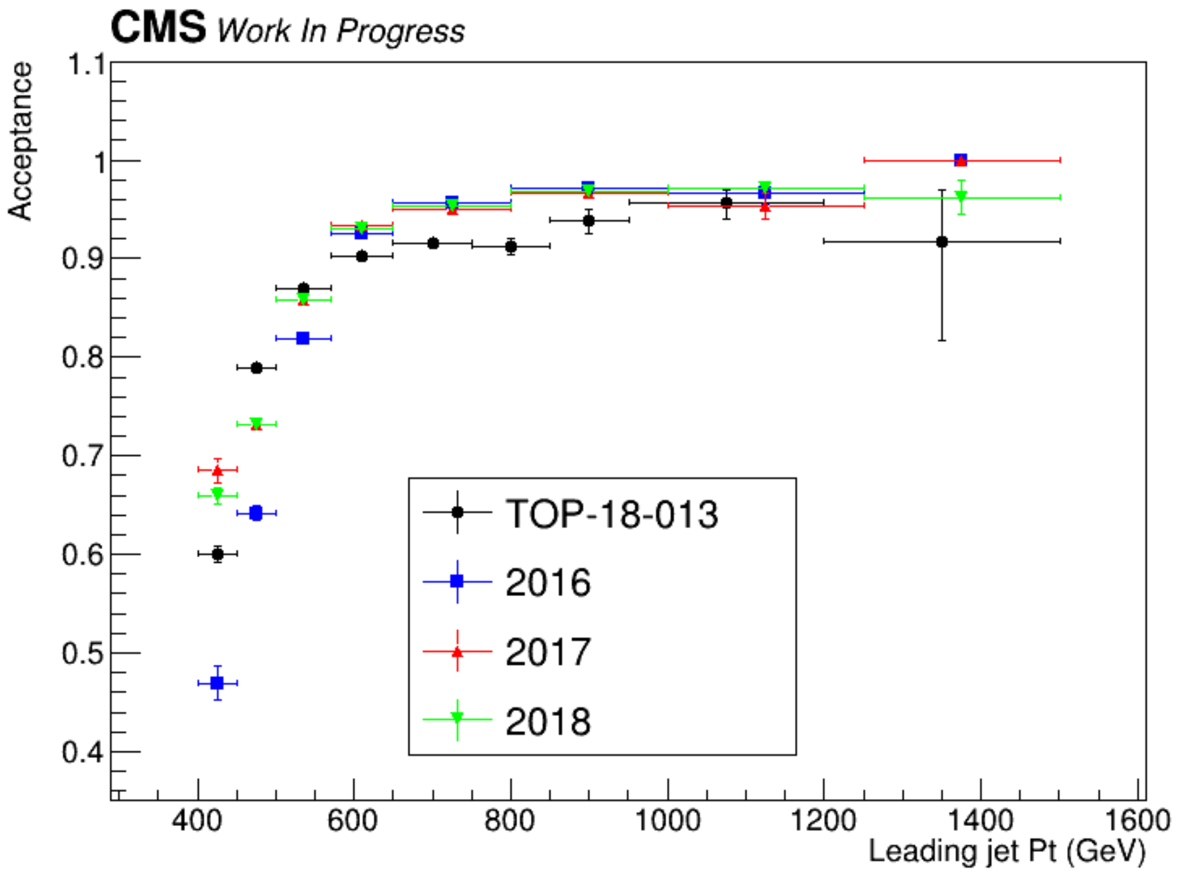
2018



Region	**Requirements**
Signal Region (SR)	Baseline + topTagger + $m_{SD}^{jet1,2} \in (120,220) GeV + 2btags$
Control Region (CR)	Baseline + topTagger + $m_{SD}^{jet1,2} \in (120,220) GeV + 0btags$
Extended SR (SR_A) (QCD fit region)	Baseline + topTagger + $m_{SD}^{jet1,2} \in (50,300) GeV + 2btags$

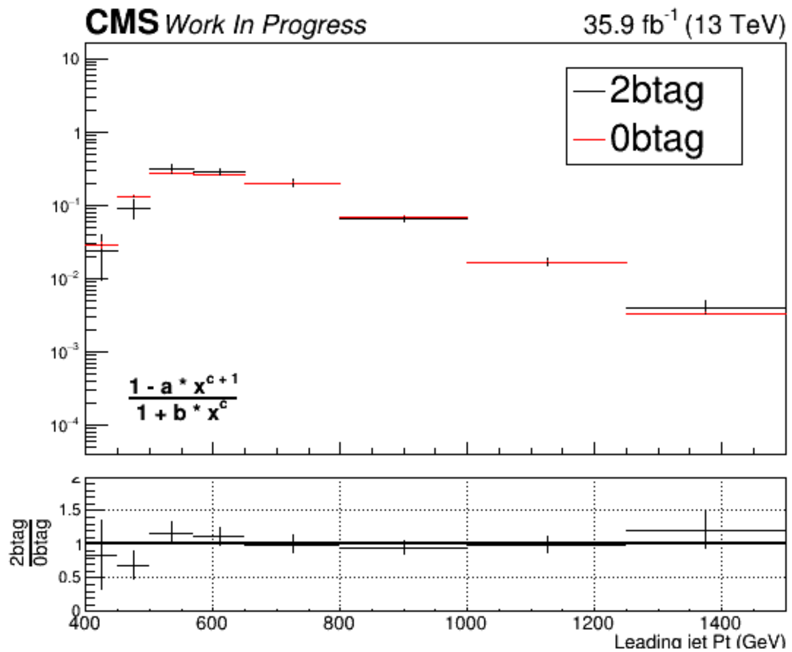
Efficiency and Acceptance for 2016, 2017 and 2018 and previous 2016 analysis



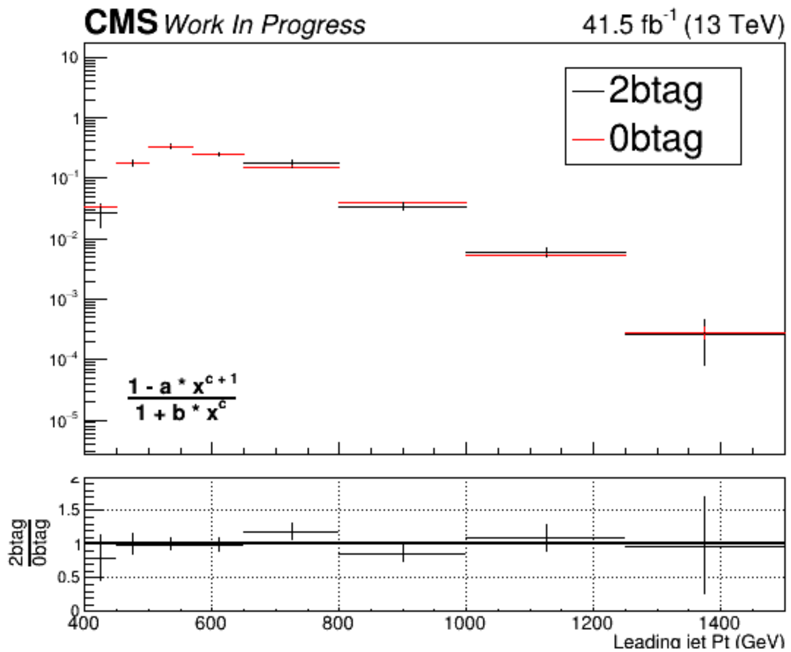


QCD Closure Tests ’16, ‘17, ’18 (corrected) for leading JetPt

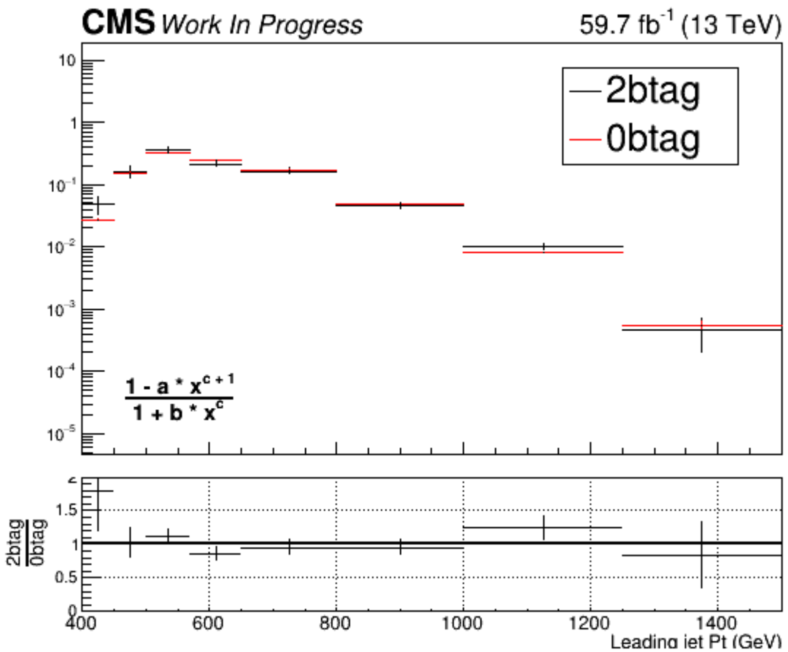
2016



2017

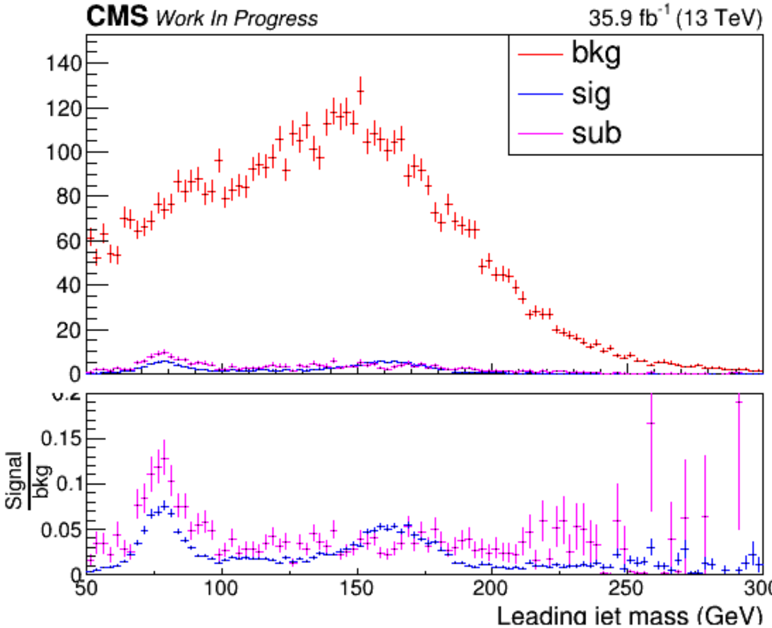


2018

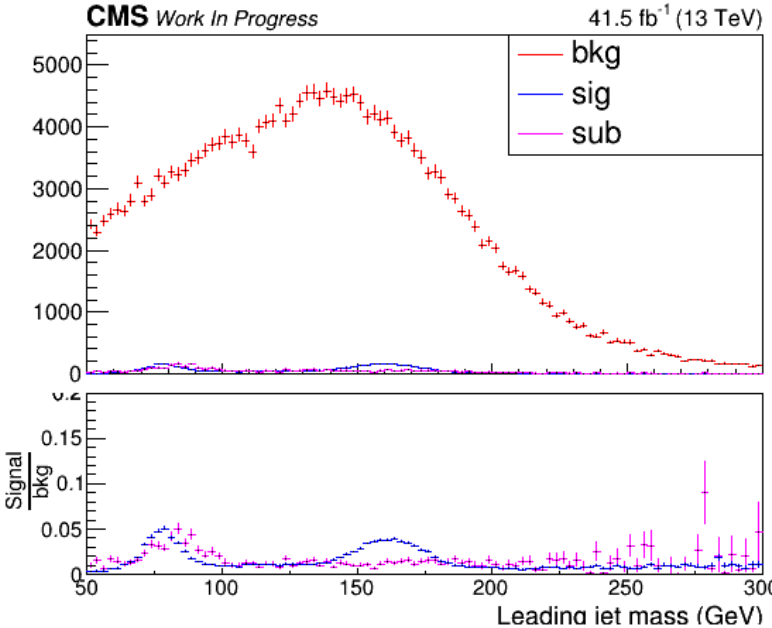


CR Contamination ’16,’17,’18 for the leading Jet Mass (softDrop)

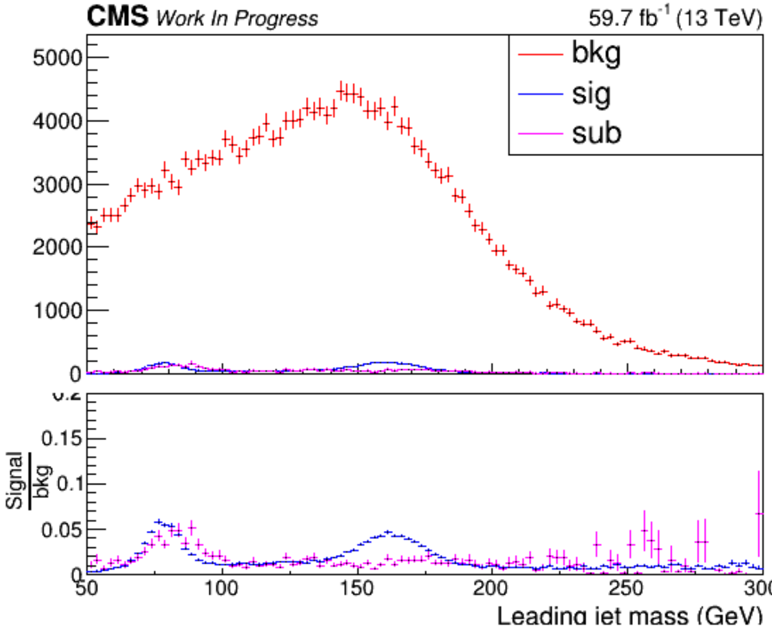
2016



2017

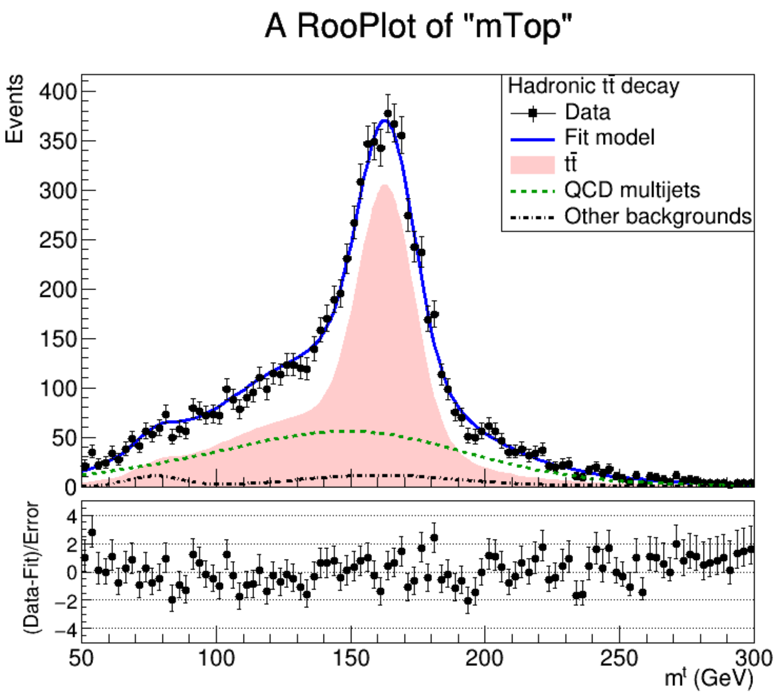


2018



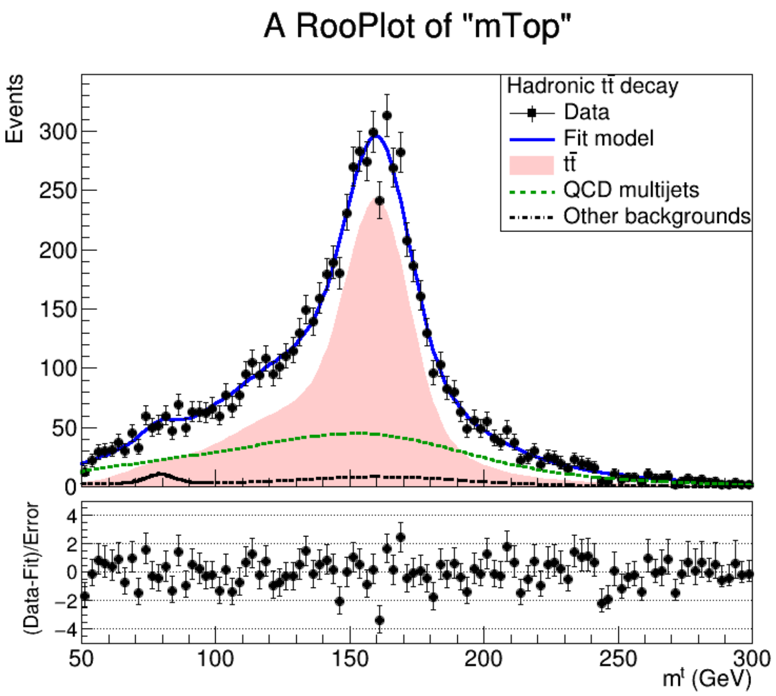
Mass Fit in Signal Region for 2016, 2017, 2018

Signal Region (2btag) (2016)



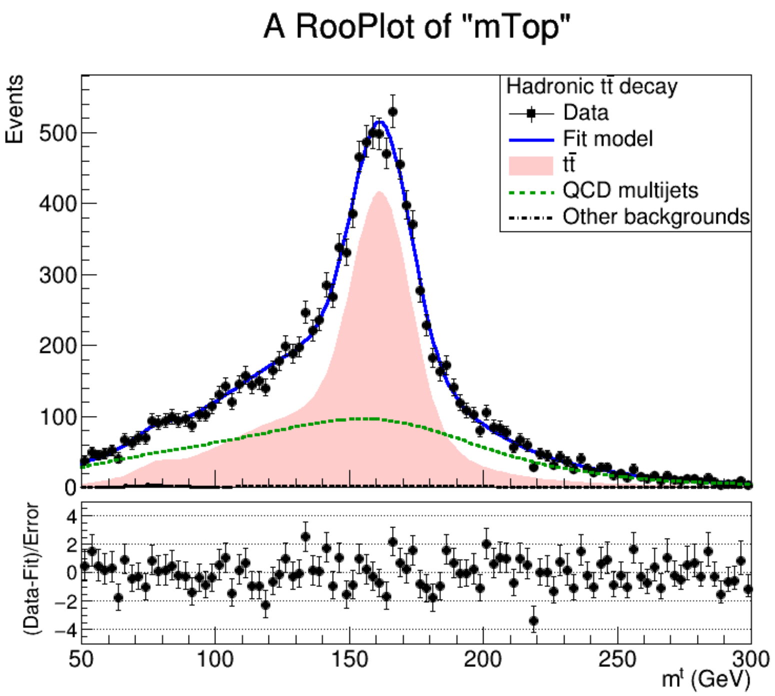
Signal Strength: r = 0.69 ± 0.026

Signal Region (2btag) (2017)



Signal Strength: r = 0.64 ± 0.024

Signal Region (2btag) (2018)



Signal Strength: r = 0.68 ± 0.02

Result of the fit on data in SR. The red area shows the ttbar contribution, the green line shows the QCD, and the black line shows the subdominant backgrounds

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)

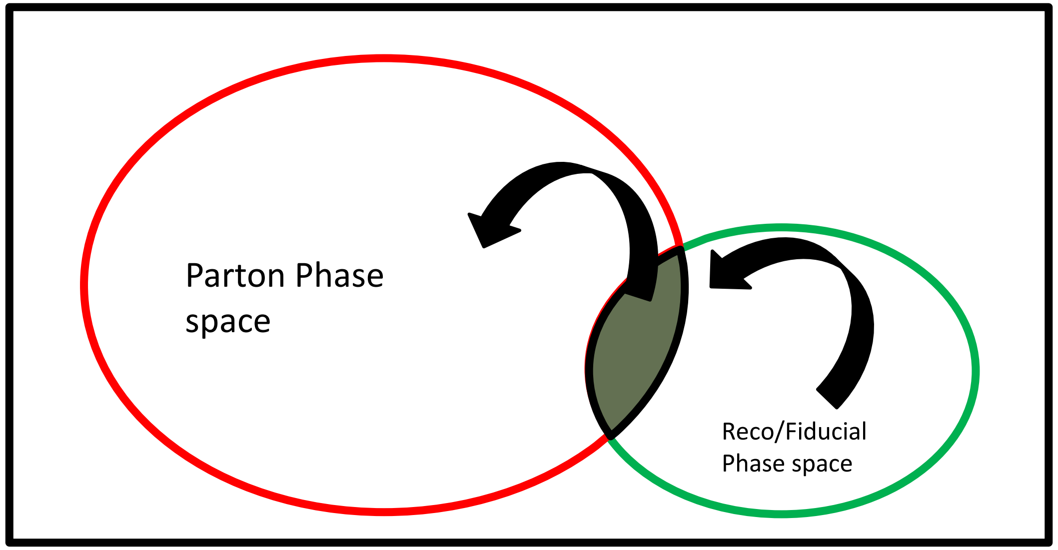
- Where x_{reco} is the respected variable of interest (ttbar mass, pt, rapidity, leading and subleading jetPt and |jetY|)
- We deploy a fit in the Signal Region (2btag) to extract the N_{QCD}^{fit}

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

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

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

efficiency of the reco+true selection

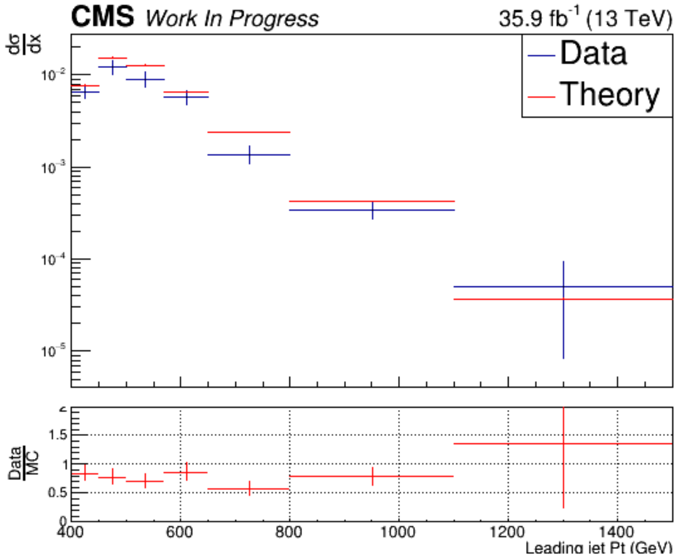
migration matrix

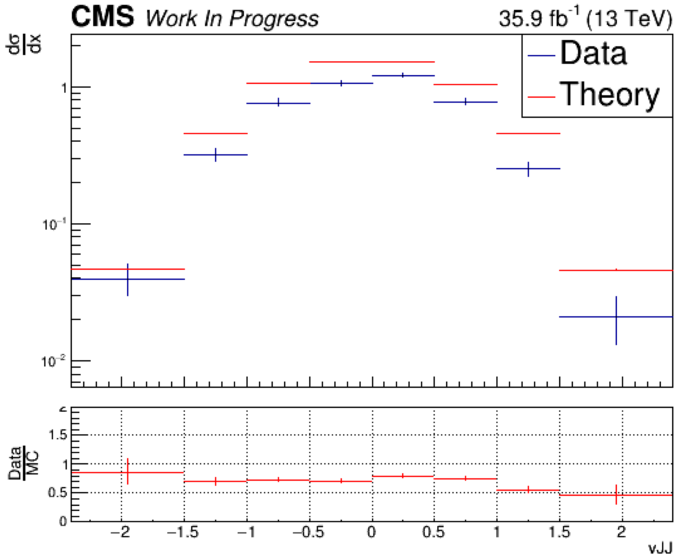
reco efficiency of the reco+true selection

Unfolding: simple response matrix inversion w/o regularisation

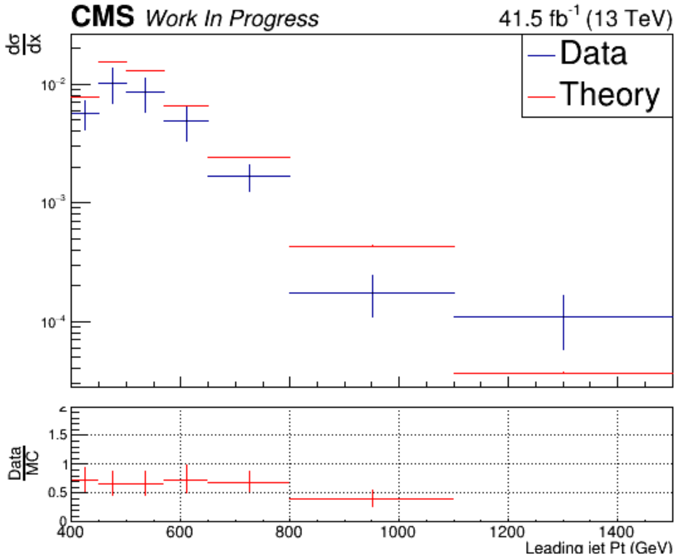
Unfolding - Extrapolation to Parton Level, Data vs MC

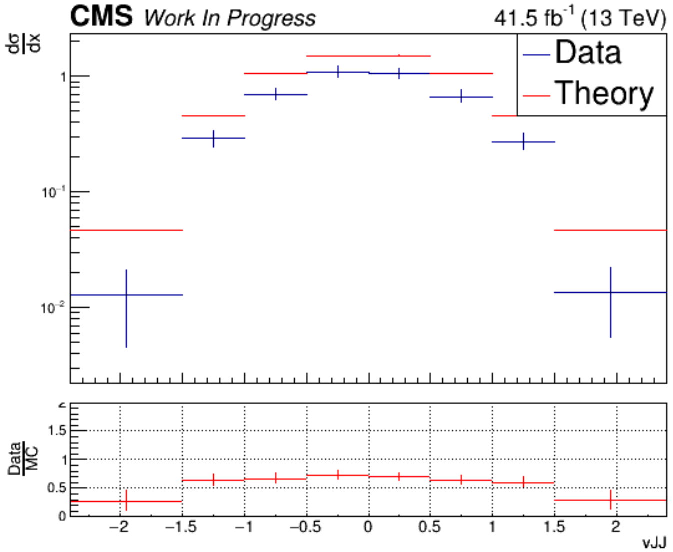
(2016)



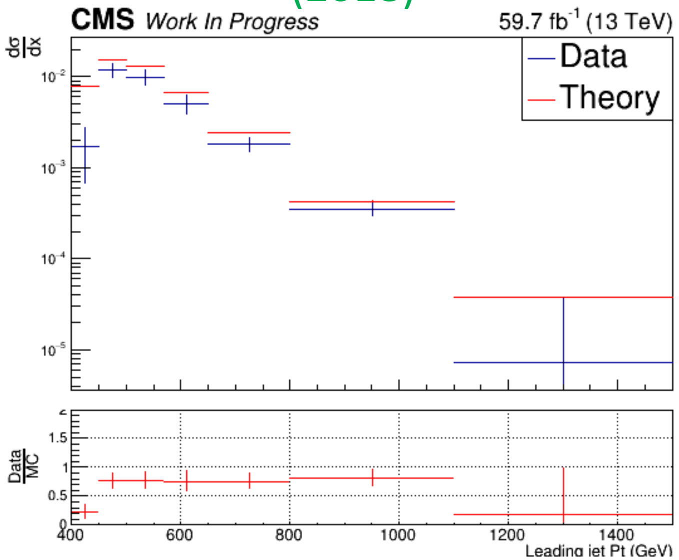


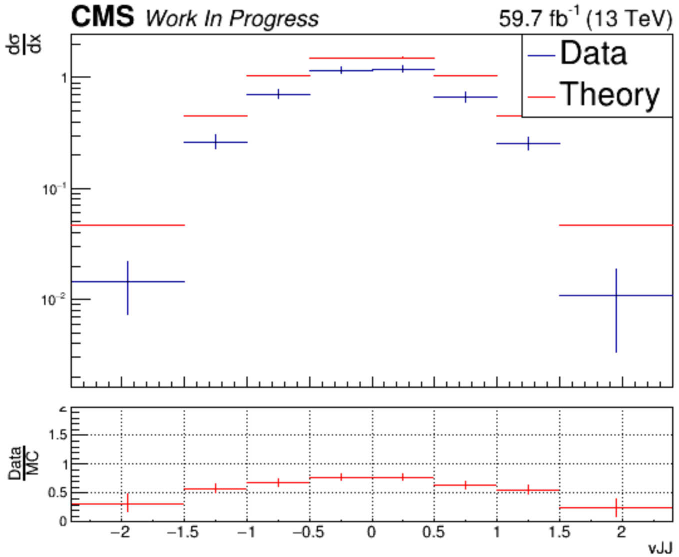
(2017)



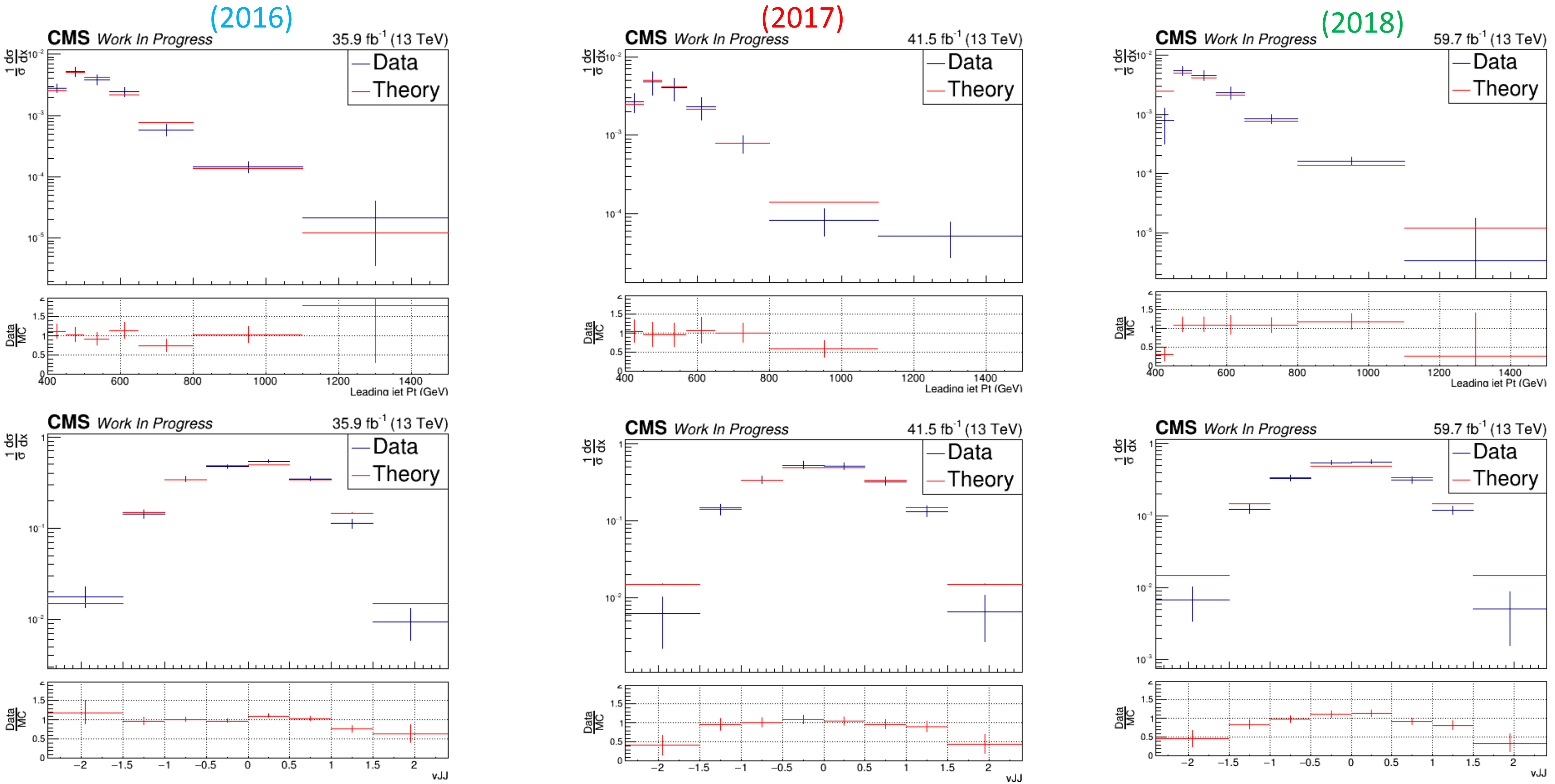


(2018)





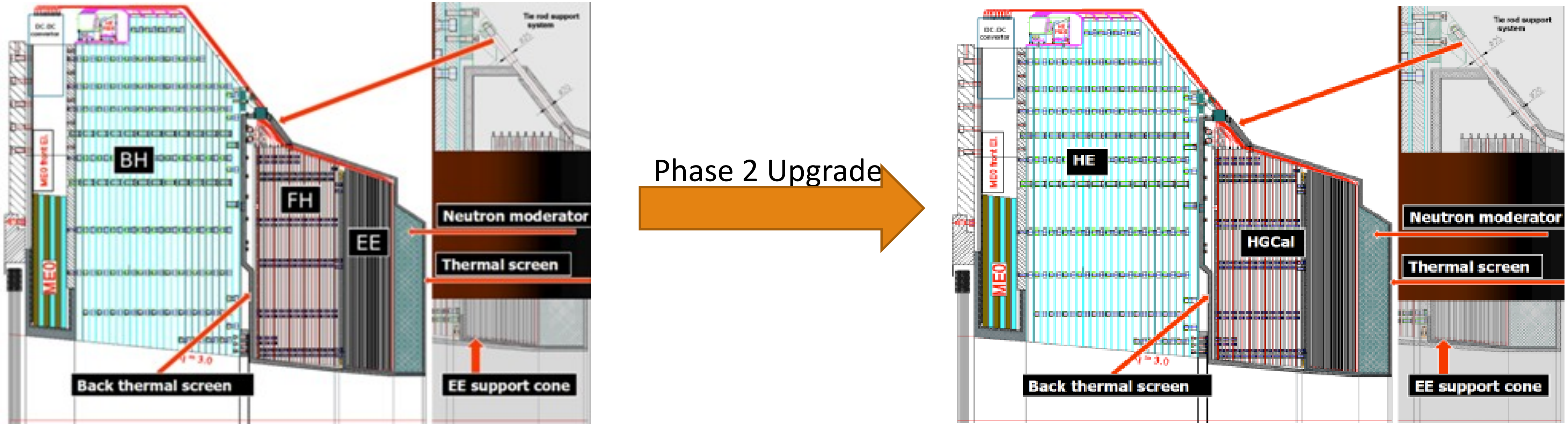
Unfolding - Extrapolation to Parton Level, Data vs MC (Normalised)



CERN Phase2 upgrade

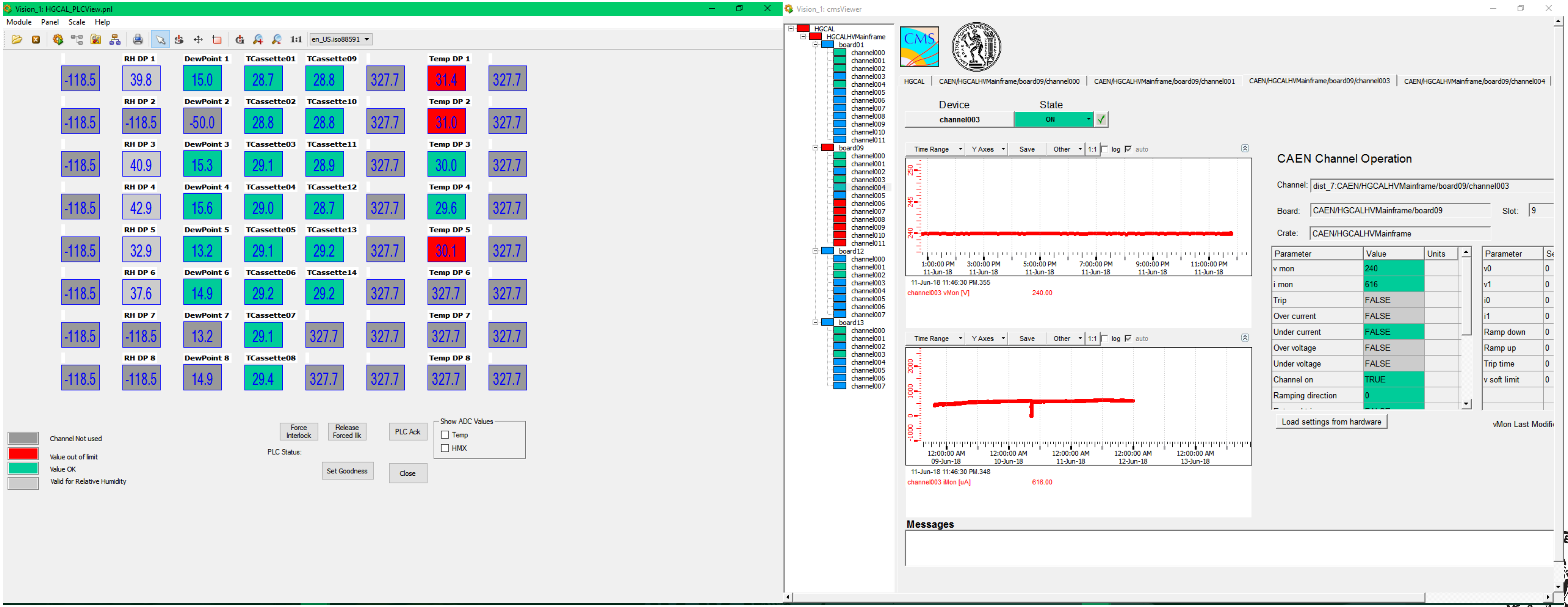
- LHC will upgrade to High Luminosity LHC (HL - LHC) between 2023 and 2025
- Higher luminosity
- Higher radiation doses than the ones the detectors were designed for
- Higher pileup
- CMS will replace most of its subdetectors to mitigate with the LHC upgrades
- High Granularity CALorimeter (HGCAL) will replace the endcap calorimeters.

High Granularity Calorimeter (HGCAL)



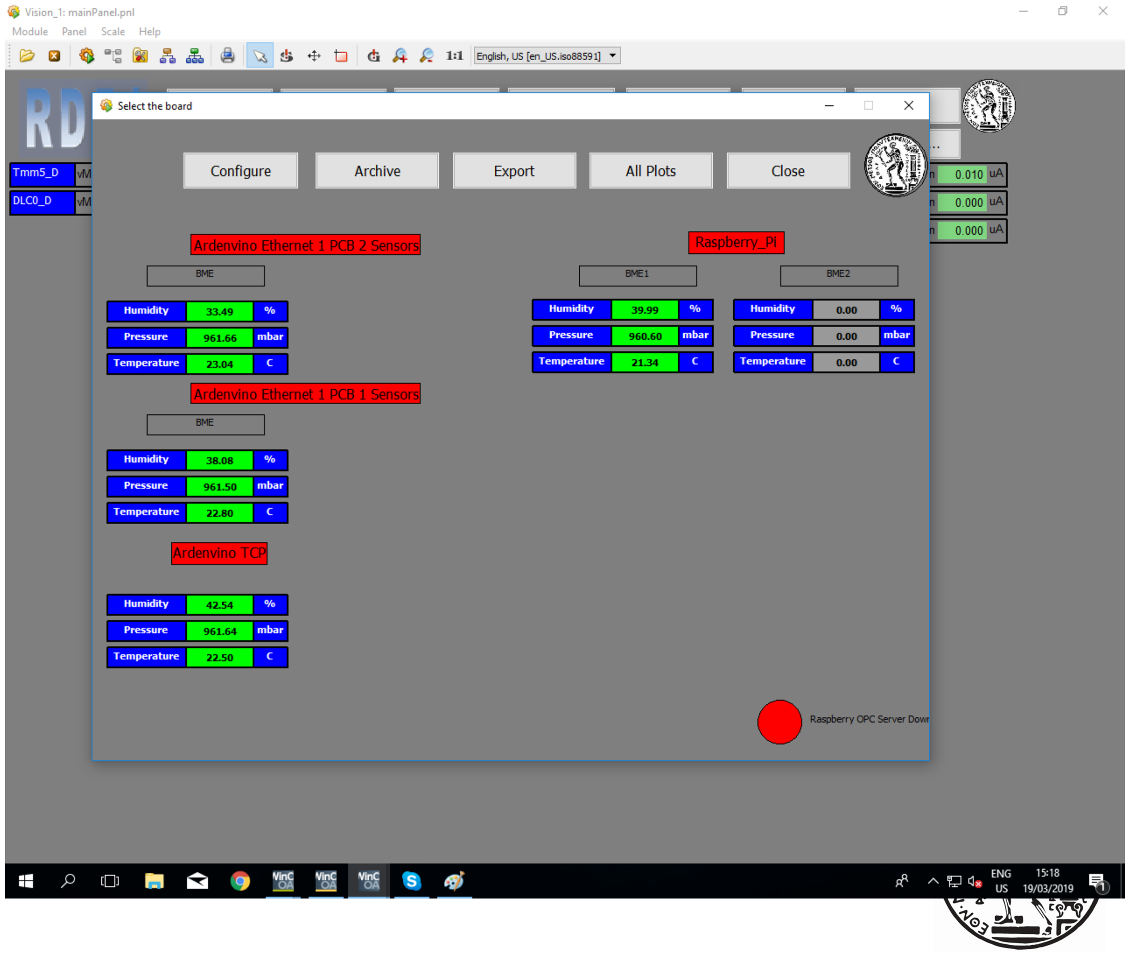
1. ECAL and HCAL won’t exist in near future on endcaps
2. HGCAL in endcaps. Why?
   - Radiation hardness because of bigger lumi
   - Endcaps have maximum radiation
   - Bigger Lumi → higher granularity → more modules/ind. Components
   - Silicon detector

High Granularity Calorimeter (HGCAL)



Environmental Monitoring for the RD51 Test Beams

- Development of the slow control system
- Developed environmental monitoring system
- Support during test beams



Conclusions

- Contribution to the CMS DCS group
    - Maintenance and development of new tools

- Measurement of Top – antitop quark production cross section for Full Run II data collected with CMS detector
    - Combination of three years
    - Systematic variations evaluation

- HGCAL
    - DCS
    - PLC control and monitoring

- RD51
    - Environmental monitoring parameters
    - Slow Control System for High Voltage monitoring

BACKUP SLIDES

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	

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:

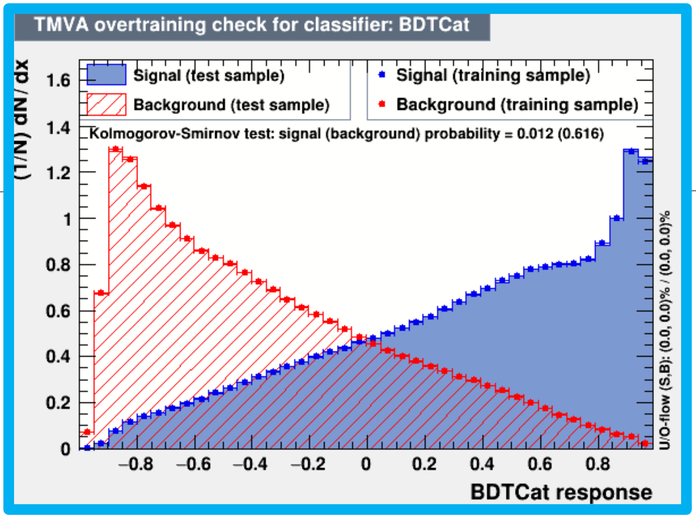
$$\chi = \frac{1 + |cos\theta^*|}{1 - |cos\theta^*|}$$

Overview: Discriminator, Efficiency and Acceptance

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

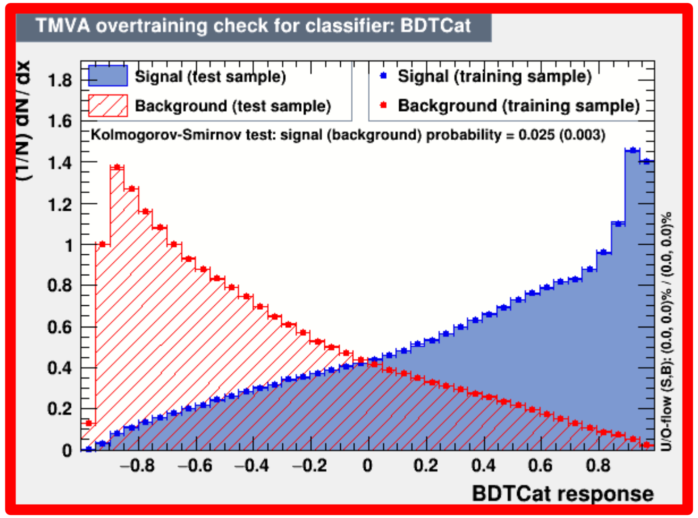
Category training: split the sample in categories based on Pt

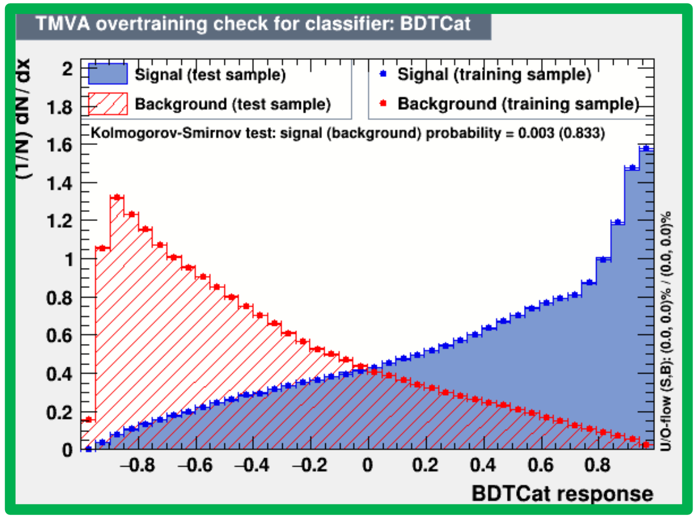
- Bins:
  - [400, 600] GeV
  - [600, 800] GeV
  - [800, 1200] GeV
  - [1200, inf) GeV
- BDT, used variables:
  - Leading and Sub-leading subjet mass
  - N-Subjetiness variables (tau1, tau2, tau3)
  - fraction of the jetPt over the total pt sum of the event.
  - Energy correlation functions (ecfB1N2,ecfB1N3, 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

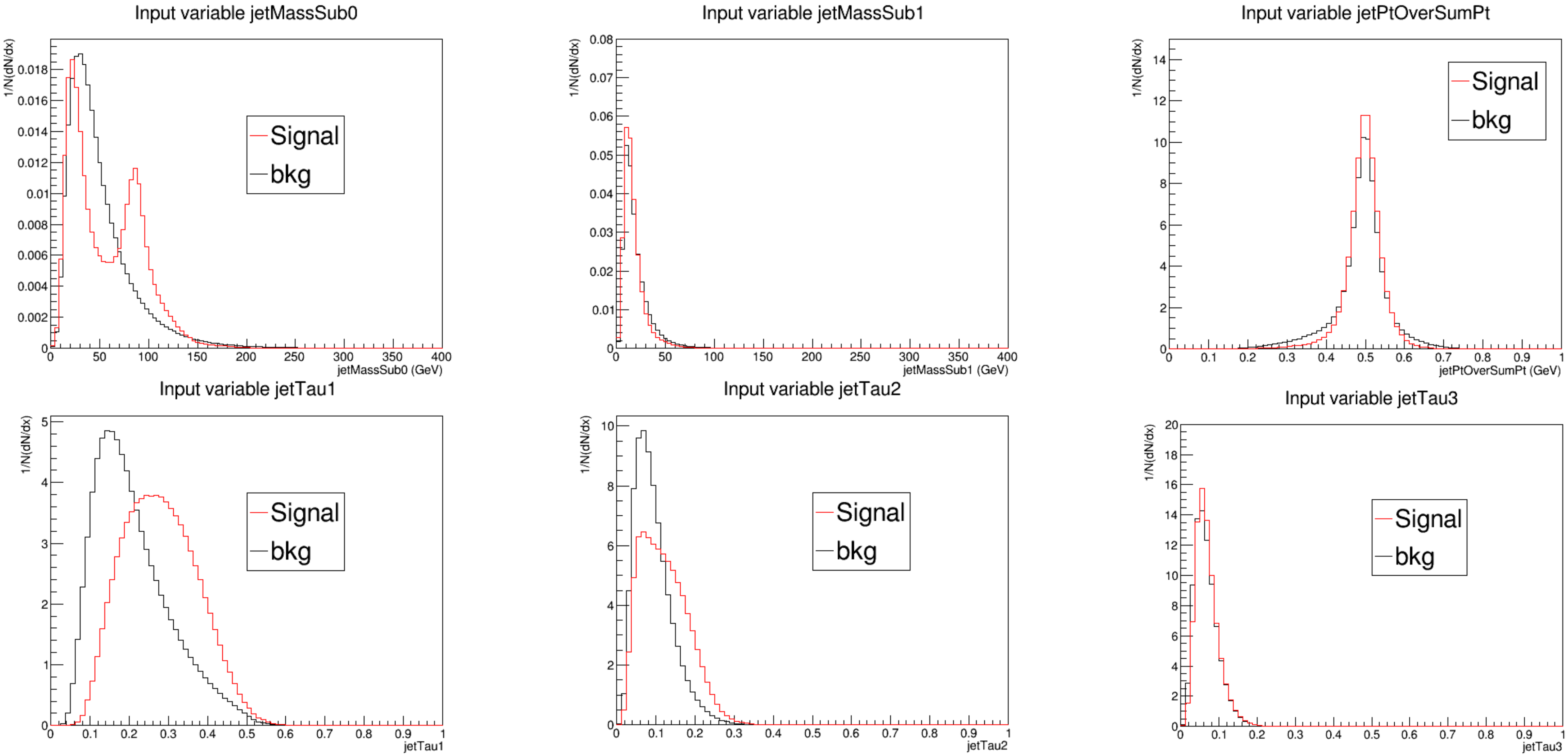
2017



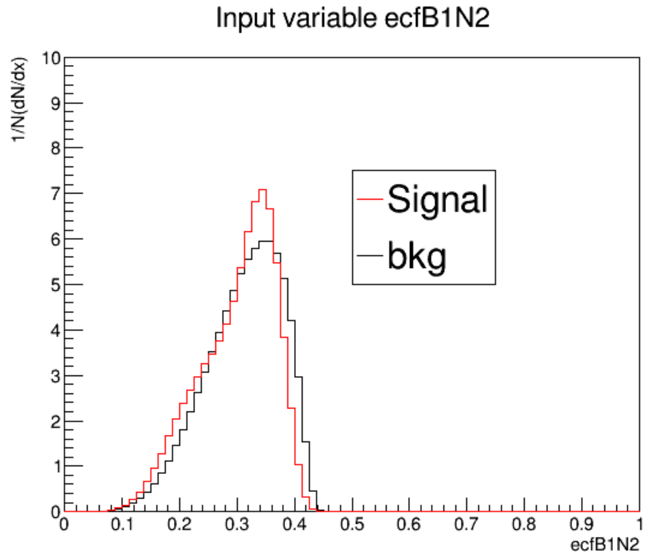


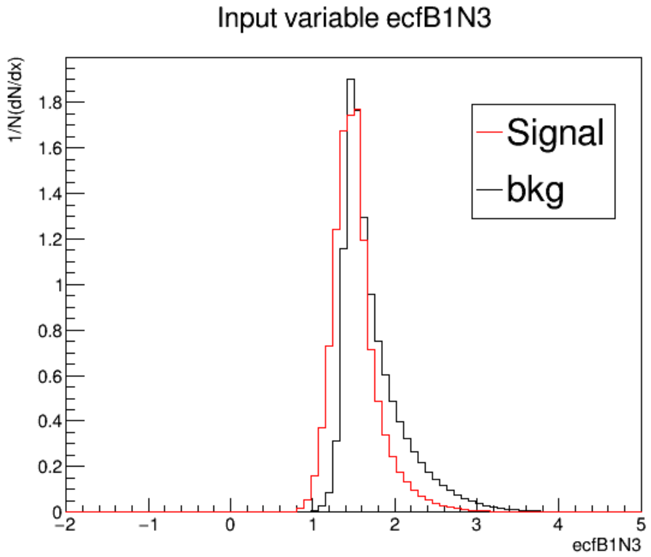
2018

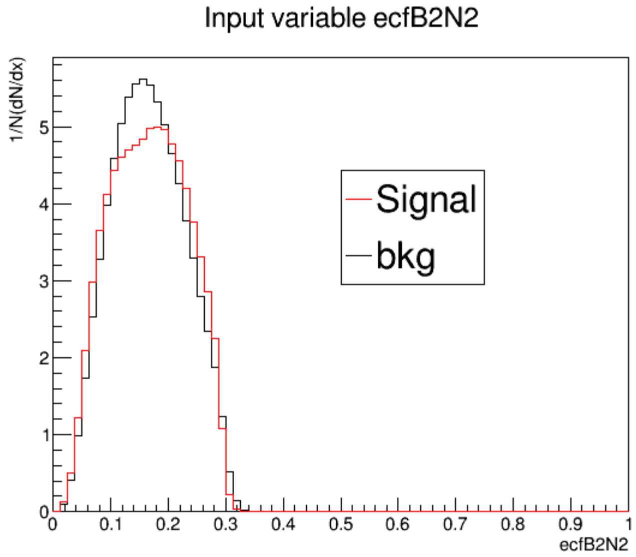
Training variables 2017

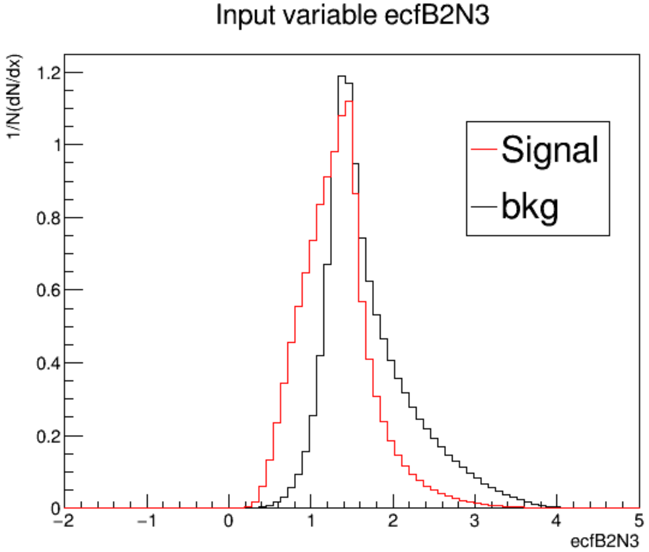


Training variables 2017









Signal Extraction

$$S_{1.5TeV}(x_{reco}) = D_{1.5TeV}(x_{reco}) - QCD_{1.5TeV}(x_{reco}) - Sub_{1.5TeV}(x_{reco}) \rightarrow$$

Where $QCD_{1.5TeV}(x_{reco}) = D^{0-btag}_{1.5TeV,shape}(x_{reco})\boldsymbol{x}N_{SR(1.5TeV)}\boldsymbol{x}C^{shape\ SF}_{closure}$

and $N_{SR(1.5TeV)} = R^{1TeV\rightarrow 1.5TeV}_{yield}\boldsymbol{x}N^{QCD}_{SR(1TeV)} = R^{1TeV\rightarrow 1.5TeV}_{yield}\boldsymbol{x}R^{SR_A\rightarrow SR}_{yield}\boldsymbol{x}N^{QCD}_{SR_A}$

- The variable of interest here: x_{reco} → χ
- 1.5 TeV refers to the mJJ cut

- We deploy a fit in the Signal Region (2btag) to extract the N^{fit}_{QCD} in SRA (mJJ > 1TeV)

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