Status Report
Mass Fit and bTagging Efficiency
(2016 and 2017)

George Bakas, Ioannis Papakrivopoulos

Status Report

- Working with 2017 data and MC
  - Found a problem in the 2017 Data files, fix and new production
  - Missing subdominant W+Jets bkg (only HT 400-600, 600-800, 800-Inf whereas in 2016 we are using the HT180)
  - Simultaneous fit In 3 regions for 2017 also
  - Btagging efficiency and acceptance

- 2018 Analysis:
  - Production with 2018 data
  - MC's for signal and bkg are already being used
  - Subdominant bkg's: haven't found any MC's → request samples

- Preparation for a presentation in the following ttX meeting

- High Pt samples will not be used in our analysis
  - Maybe we can use them for the angular distributions for BSM analysis
  - No XSEC in the XSDB

- MATRIX: A tool for single & double differential x-section calculations at QCD NNLO precision
  - https://www.physik.uzh.ch/~grazzini/codes/MATRIX_ttbar.tar.gz
  - https://indico.cern.ch/event/830184/contributions/3476710/attachments/1869172/3075258/CMStop.pdf
  - Matrix allows the user to evaluate fully differential cross sections for a wide class of processes at hadron colliders in NNLO QCD

Simultaneous Fit in 3 regions

- As decided the previous week → Simultaneous fit in 3 regions (2btag, 1btag and 0btag) (now for 2017!)

$$D(x)^{(0)} = N_{tt}^{(0)} T^{(0)}(x, kMassScale, kMassResolution) + N_{bkg}^{(0)} B(x, \vec{p}) + N_{sub}^{(0)} O^{(0)}(x)$$
$$D(x)^{(2)} = N_{tt}^{(2)} T^{(1)}(x, kMassScale, kMassResolution) + N_{bkg}^{(2)} B(x, \vec{p})(1 + k_1 x) + N_{sub}^{(2)} O^{(1)}(x)$$
$$D(x)^{(1)} = N_{tt}^{(1)} T^{(2)}(x, kMassScale, kMassResolution) + N_{bkg}^{(1)} B(x, \vec{p})(1 + k_2 x) + N_{sub}^{(1)} O^{(2)}(x)$$

  - $N_{sub}^{(0)}$ is limited in $0.9 N_{sub,MC}^{(0)}$ *up to* $1.1 N_{sub,MC}^{(0)}$
  - We assume that $N_{tt}^{(0)} = (1 - \mathrm{e_b})^2 \mathrm{N_{tt}}, \mathrm{N_{tt}^{(2)}} = \mathrm{e_b^2 N_{tt}}$ and $\mathrm{N_{tt}^{(1)}} = 2(1 - e_b) e_b N_{tt}$ where e_b is the b tagging efficiency and N_{tt} is the total ttbar yield.

  We can either have e_b and N_{tt} as free parameters in the fit or $N_{tt}^{(0)}, N_{tt}^{(1)}, N_{tt}^{(2)}$

- We found out the the btagging efficiency and the Ntt yield are highly correlated.
  - We decided to try and fix the btagging parameter by measuring it ourselves
  - For the btagging efficiency calculation:

$e_b = \dfrac{\#subjets\ with\ flavour\ id\ requirement + deepCSV\ btagged}{\#subjets\ with\ flavour\ id\ requirement\ (b)}$, where all selected events pass baseline + parton selection

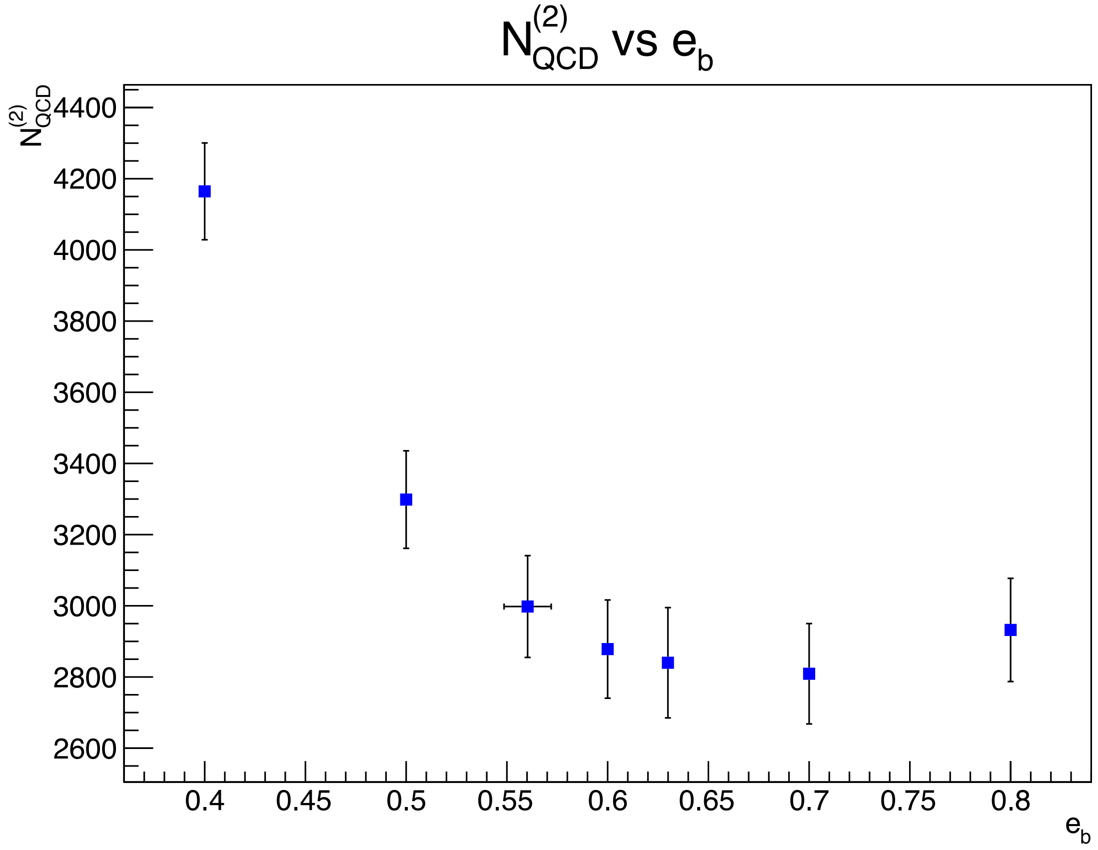
- Comparison for simultaneous fit for the 3 years
  - r for 2016: r = 1.02045 (when using the calculated btag eff constant r ≈ 0.85)
  - r for 2017: r = 0.867353 (when using the calculated btag eff constant r ≈ 0.61)

Overview of SR_A region

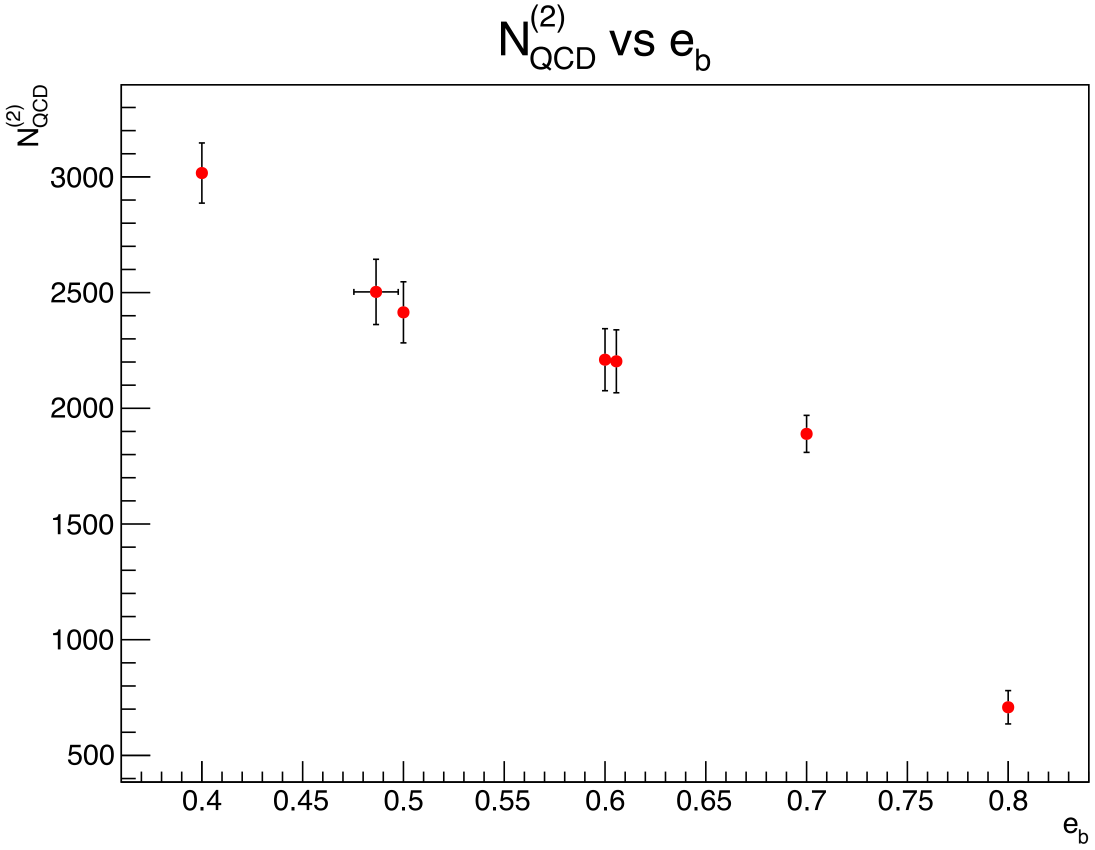
- Extension of Signal Region → SR_A = SR – Mass Selection cuts
- Selection:
  - Jet Matching
  - Parton cuts:
    - partonPt[0],[1] > 400
    - |partonEta[0],[1]| < 2.4
    - mTTbarParton > 1000
  - Reco cuts:
    - nJets > 1
    - nLeptons = 0
    - mJJ > 1000
    - jetPt[0],[1] > 400
    - |jetEta[0],[1]| < 2.4
    - bTagging cut (mediugm WP **deepCSV**) (2016: 0.6321, 2017: 0.4941, 2018: 0.4184)
    - Tagger cut (**top Tagger**) (2016: 0.2, 2017:0.0, 2018: 0.1)
    - TriggerBit

- We are checking for different values of e_b , the output of the $N^{(2)}_{QCD}$ for 2016 and 2017
- Points of interest are from 0.4 - 0.8 but especially 0.5-0.7 for 2016 and 0.45-0.6 5 for 2017
    - Calculated btagging efficiency for both years
    - btagging efficiency when the parameter is set as a free nuisance in the simultaneous fit
    - 2016: eb (fit) ≈ 0.56 and eb (calculated) ≈0.63
    - 2017: eb(fit) ≈ 0.49 and eb (calculated) ≈ 0.61

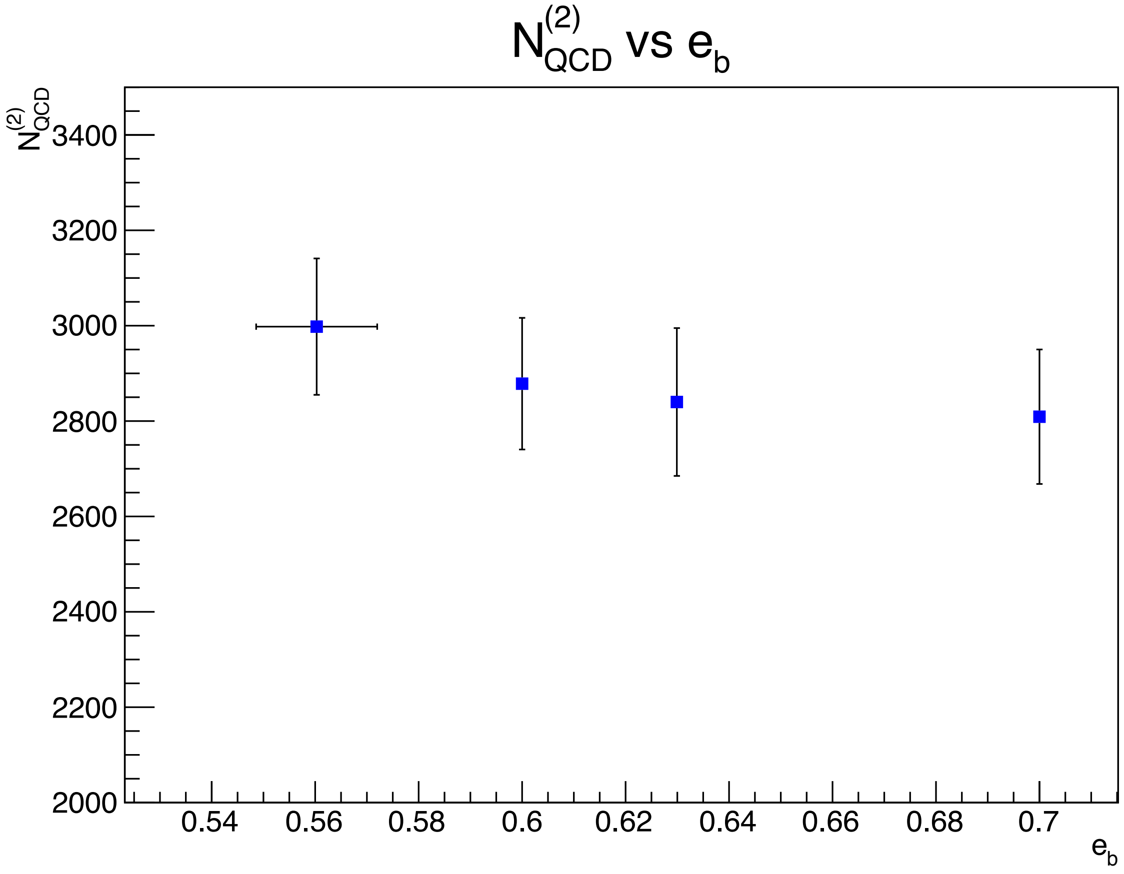
2016



2017

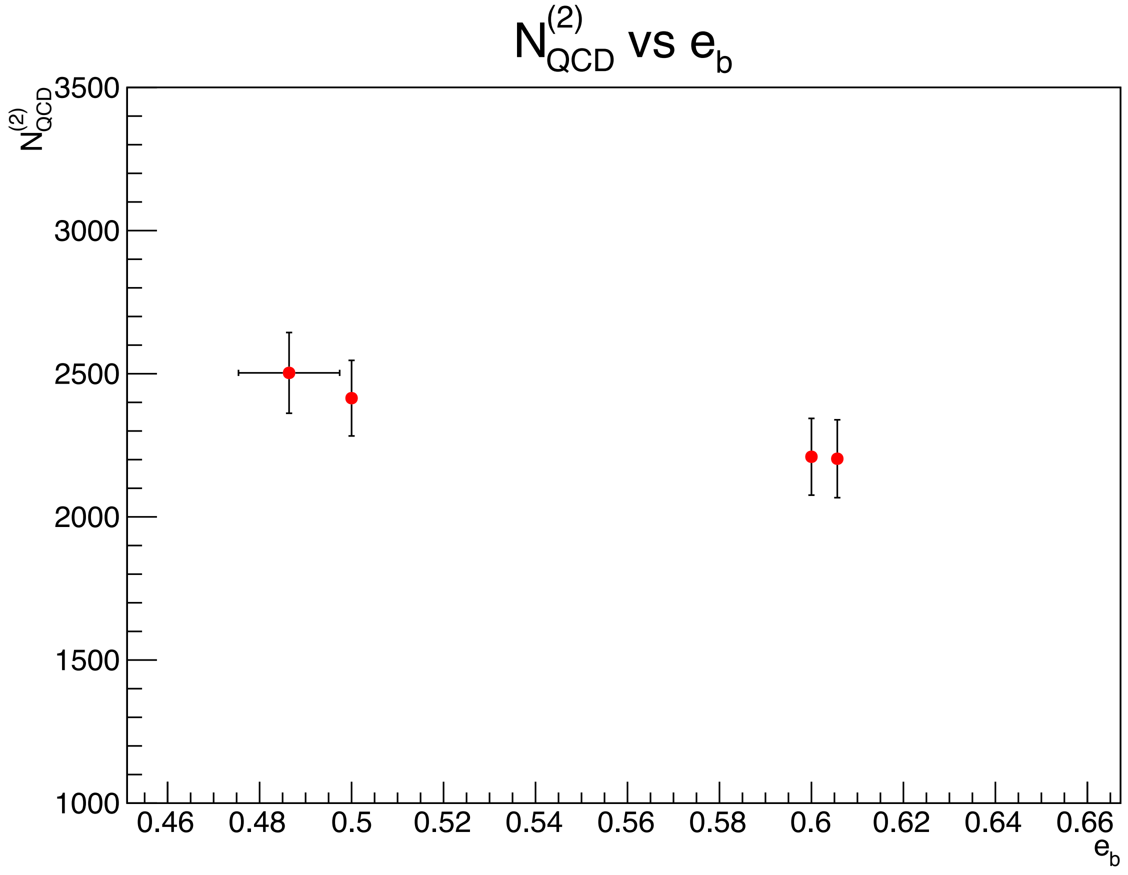


2016



$N_{QCD}^{(2)}$ vs e_b
$N_{QCD}^{(2)}$
3400
3200
3000
2800
2600
2400
2200
2000
0.54
0.56
0.58
0.6
0.62
0.64
0.66
0.68
0.7
e_b

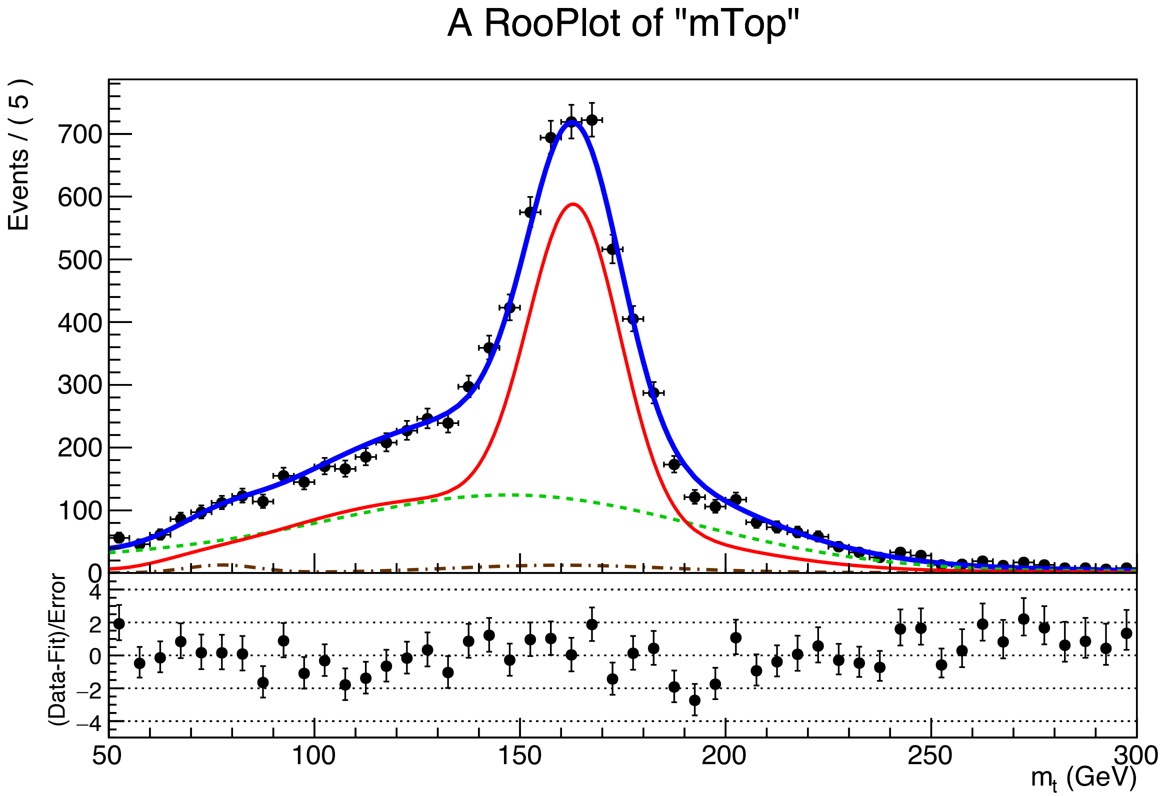
2017



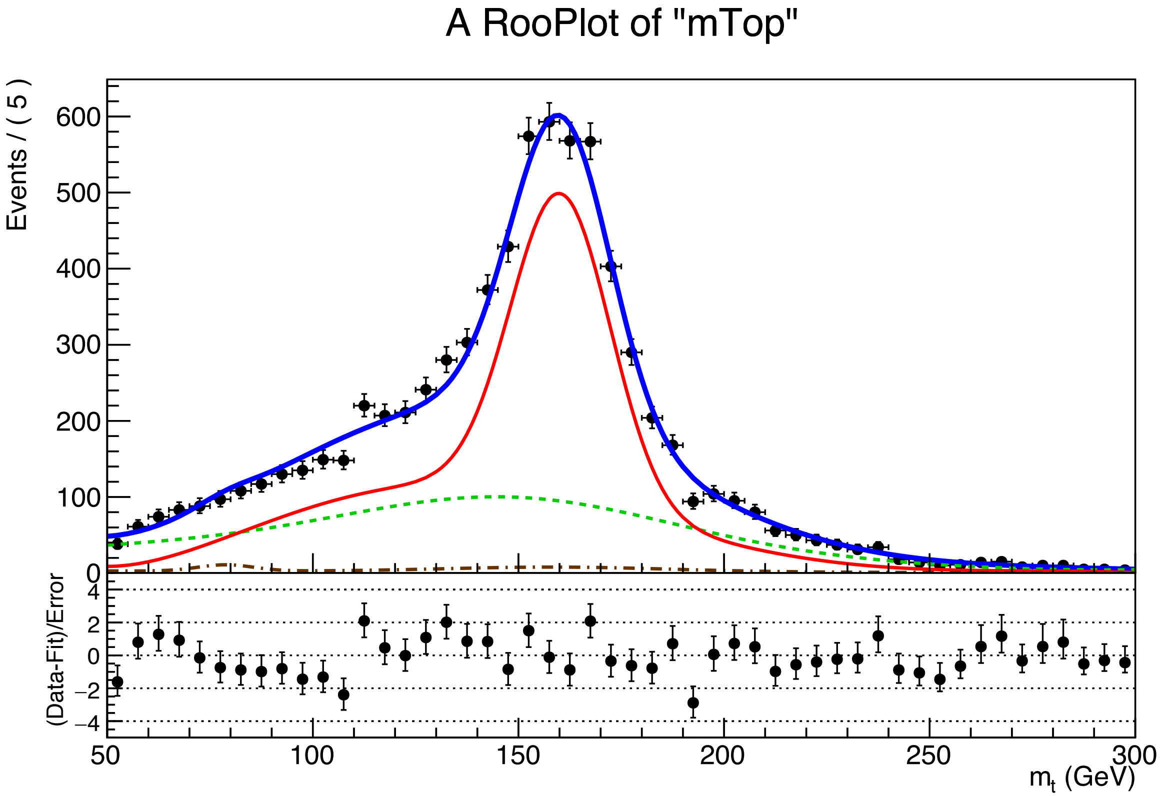
$N_{QCD}^{(2)}$ vs e_b
$N_{QCD}^{(2)}$
3500
3000
2500
2000
1500
1000
0.46
0.48
0.5
0.52
0.54
0.56
0.58
0.6
0.62
0.64
0.66
e_b

Simultaneous Fit in 3 regions for 2016 and 2017 when eb is free

Signal Region (2btag) (2016)



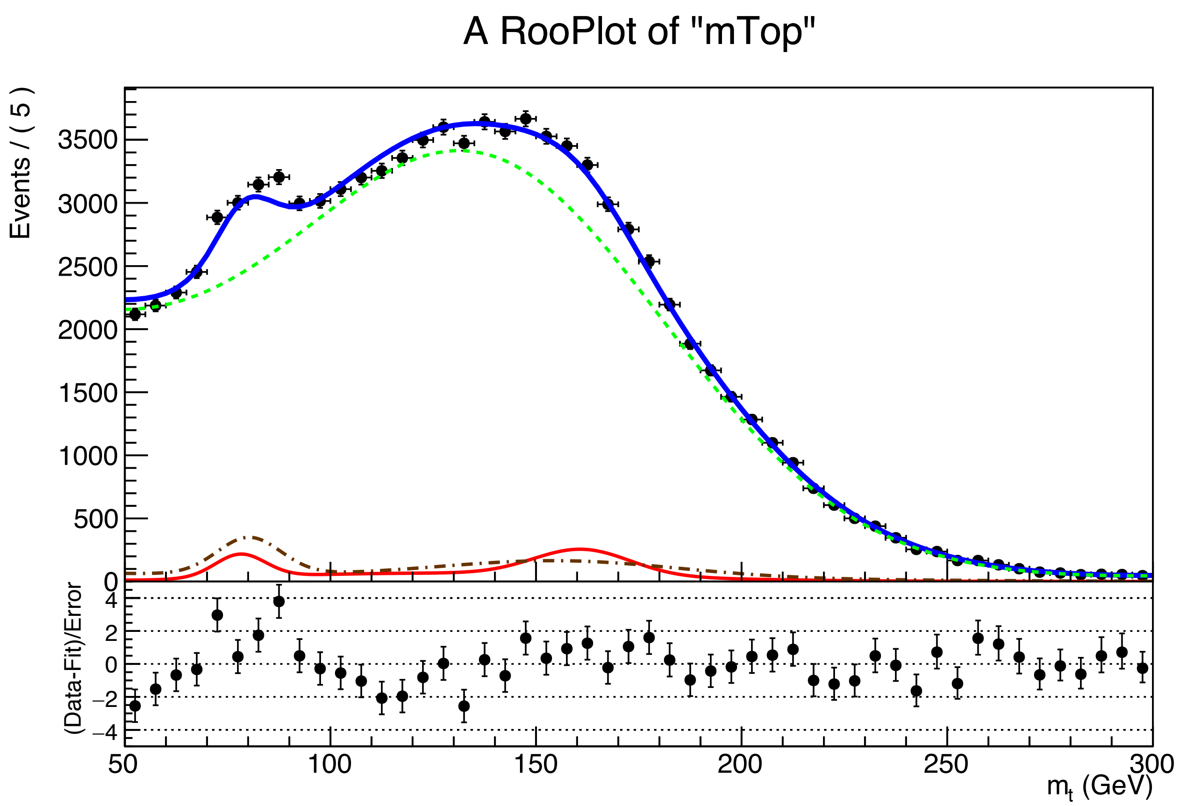
Signal Region (2btag) (2017)



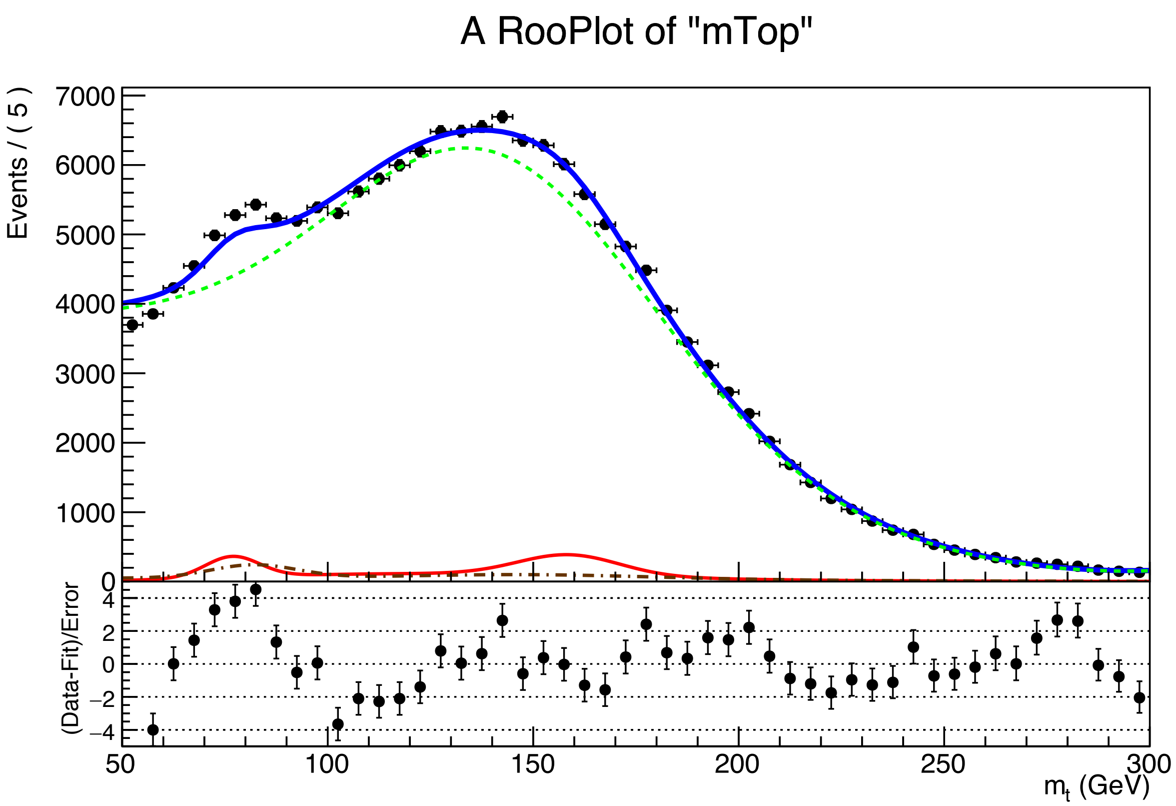
Result of the template fit on data in SR. The red line shows the ttbar contribution, the green line shows the QCD, and the brown line shows the subdominant backgrounds

Simultaneous Fit in 3 regions for 2016 and 2017 when eb is free

Control Region (0btag) 2016



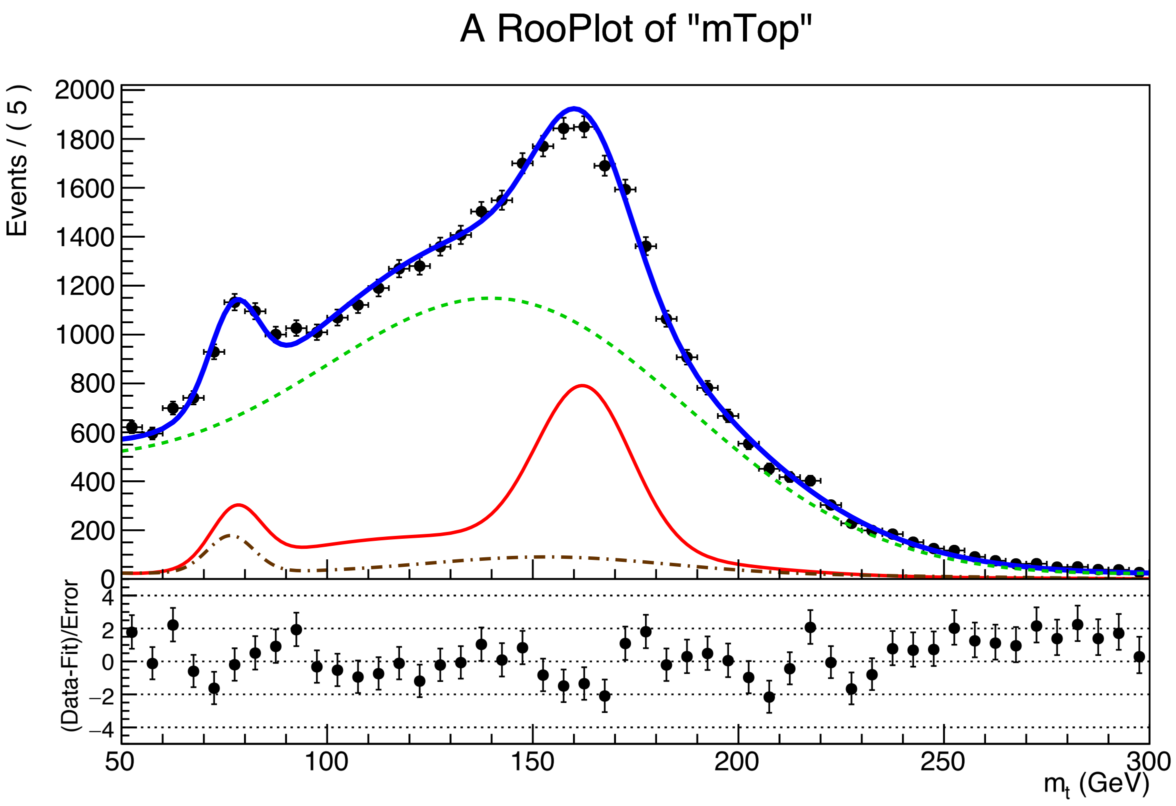
Control Region (0btag) (2017)



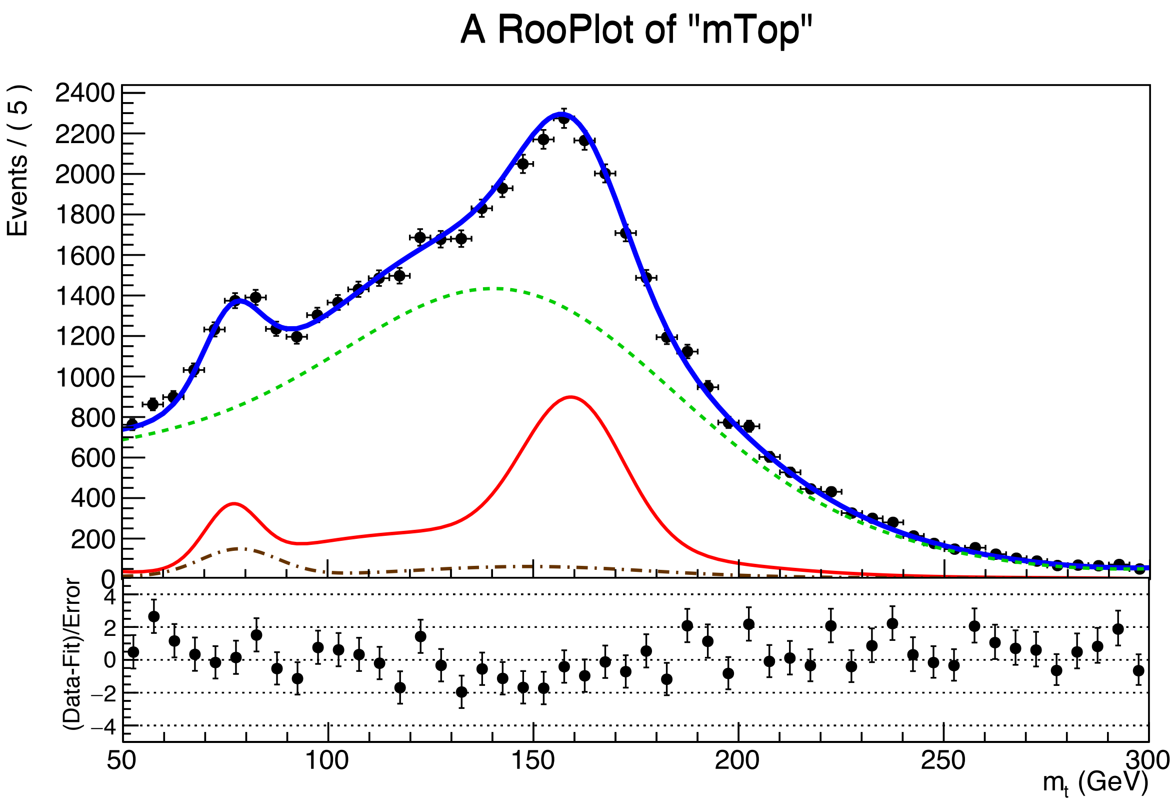
Result of the template fit on data in CR. The red line shows the tt contribution, the green line shows the QCD, and the brown line shows the subdominant backgrounds

Simultaneous Fit in 3 regions for 2016 and 2017 (1btag Region)

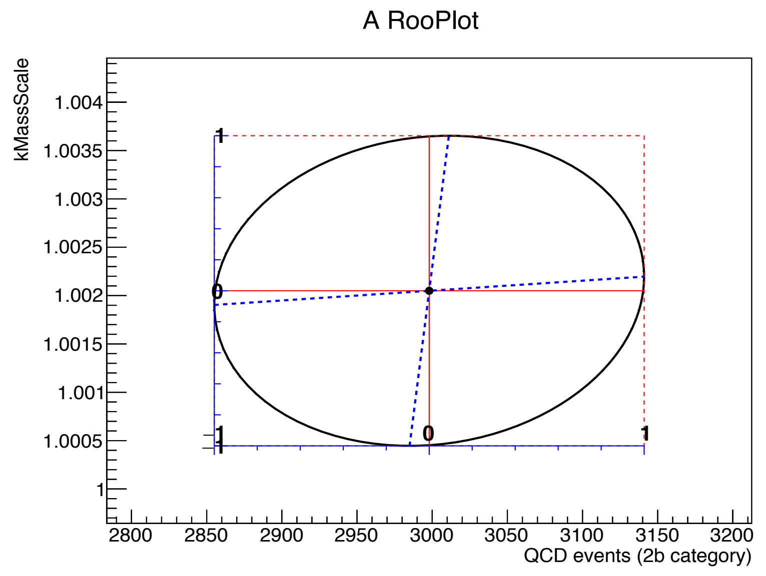
2016

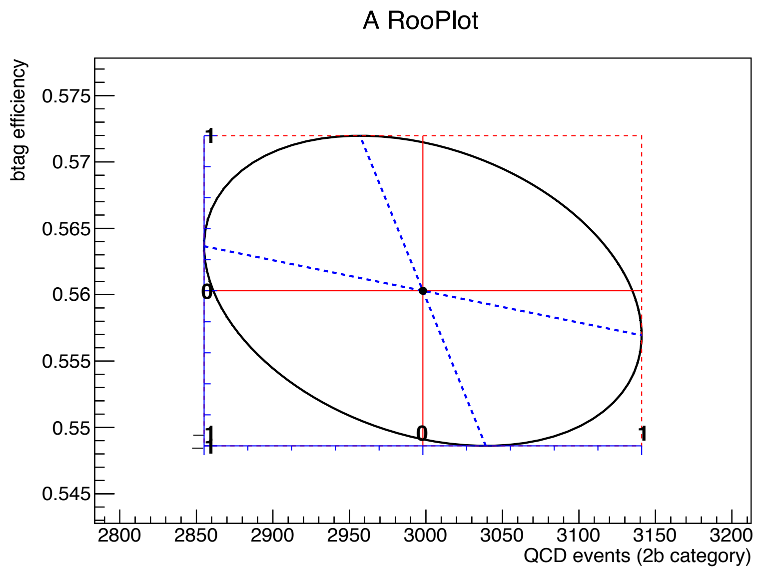


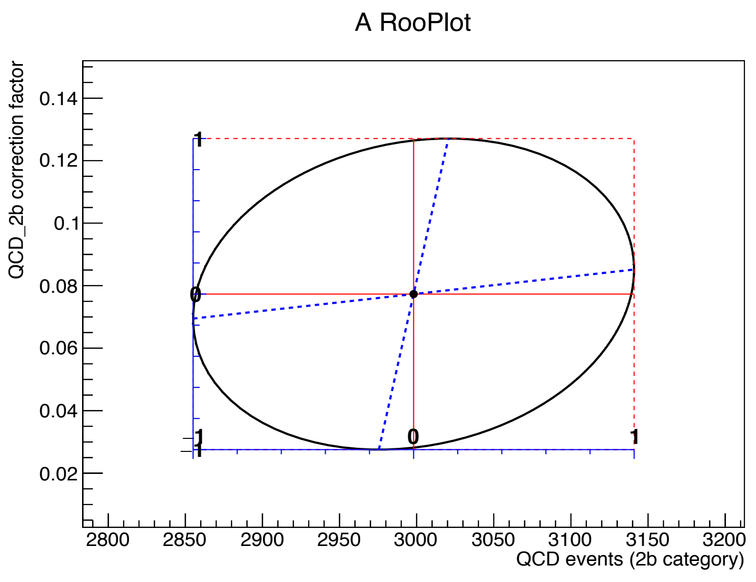
2017

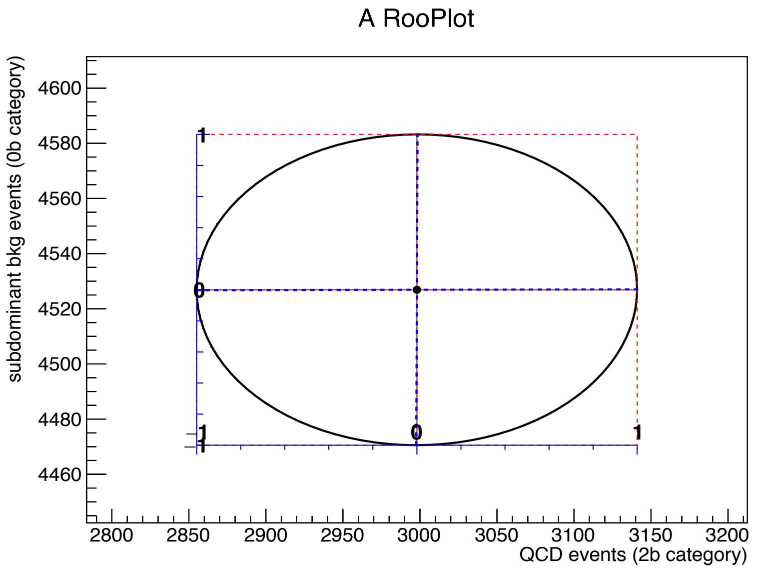


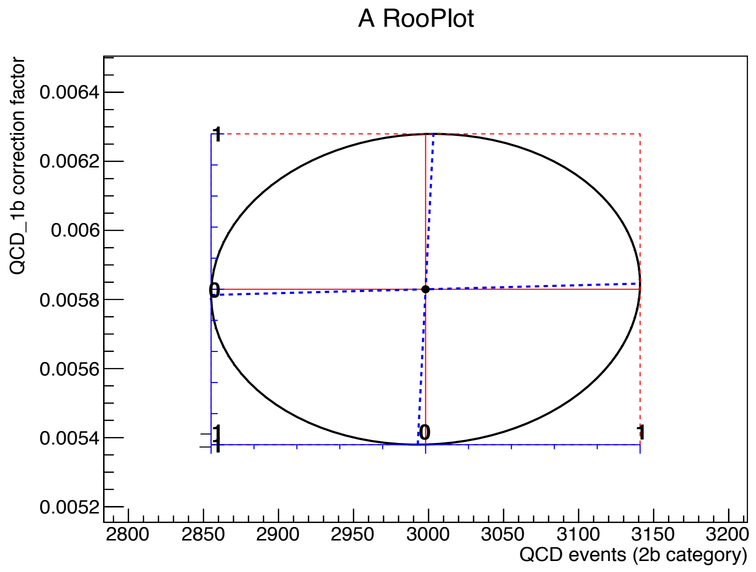
Correlation plots $N_{QCD(2)}$ vs all nuisances from fit when eb runs free

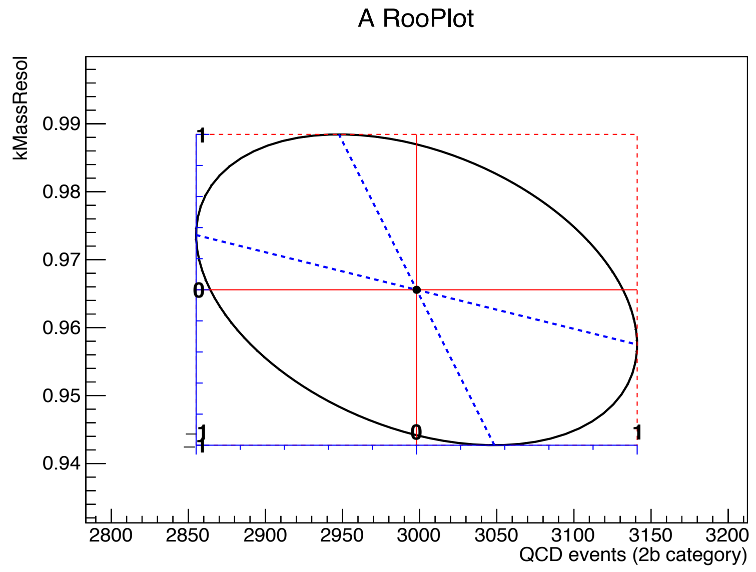




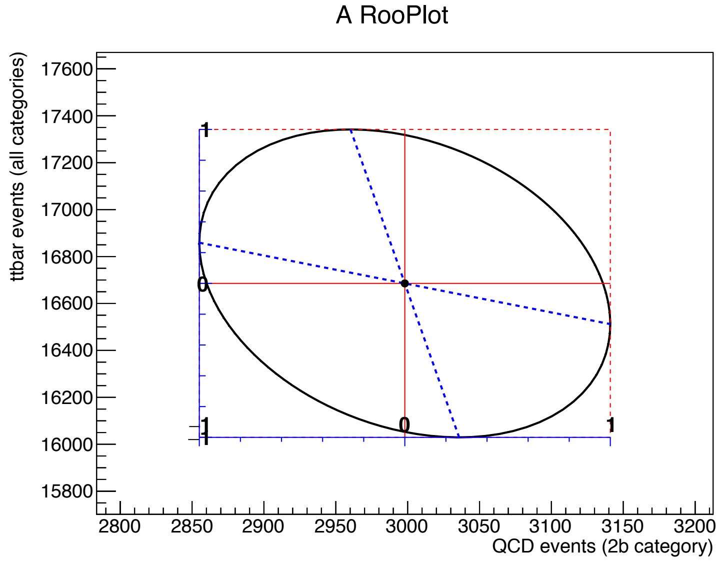


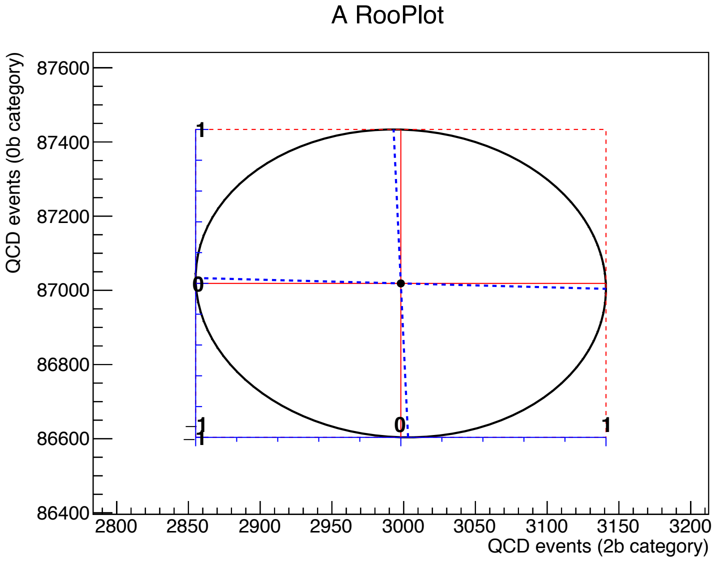


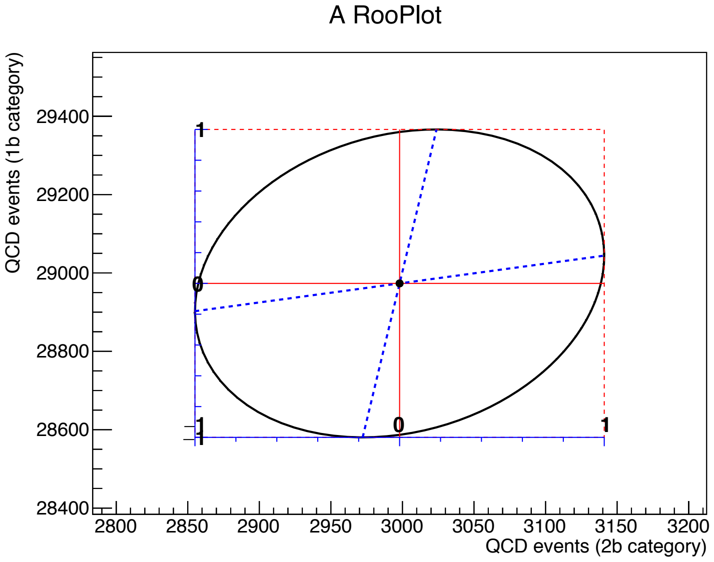


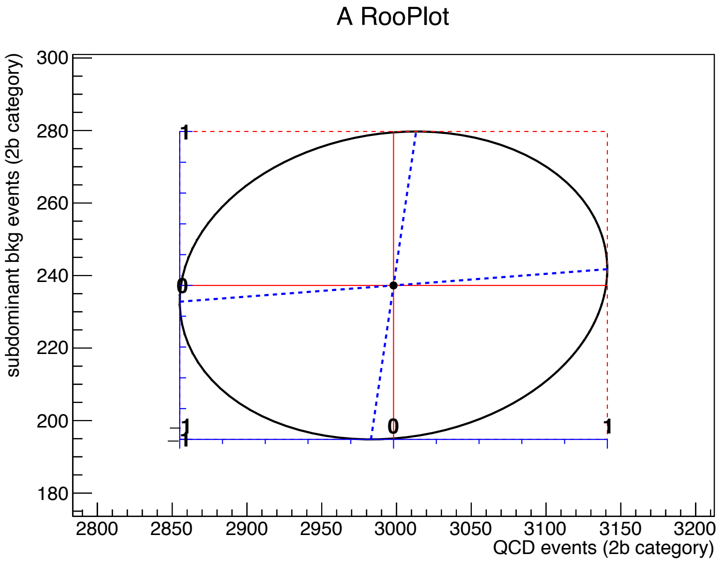


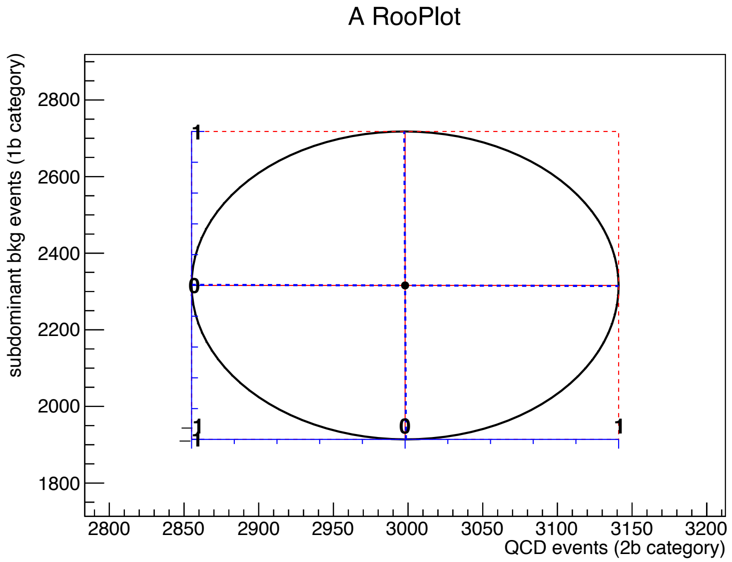
Correlation plots $N_{QCD(2)}$ vs all nuisances from fit when eb runs free





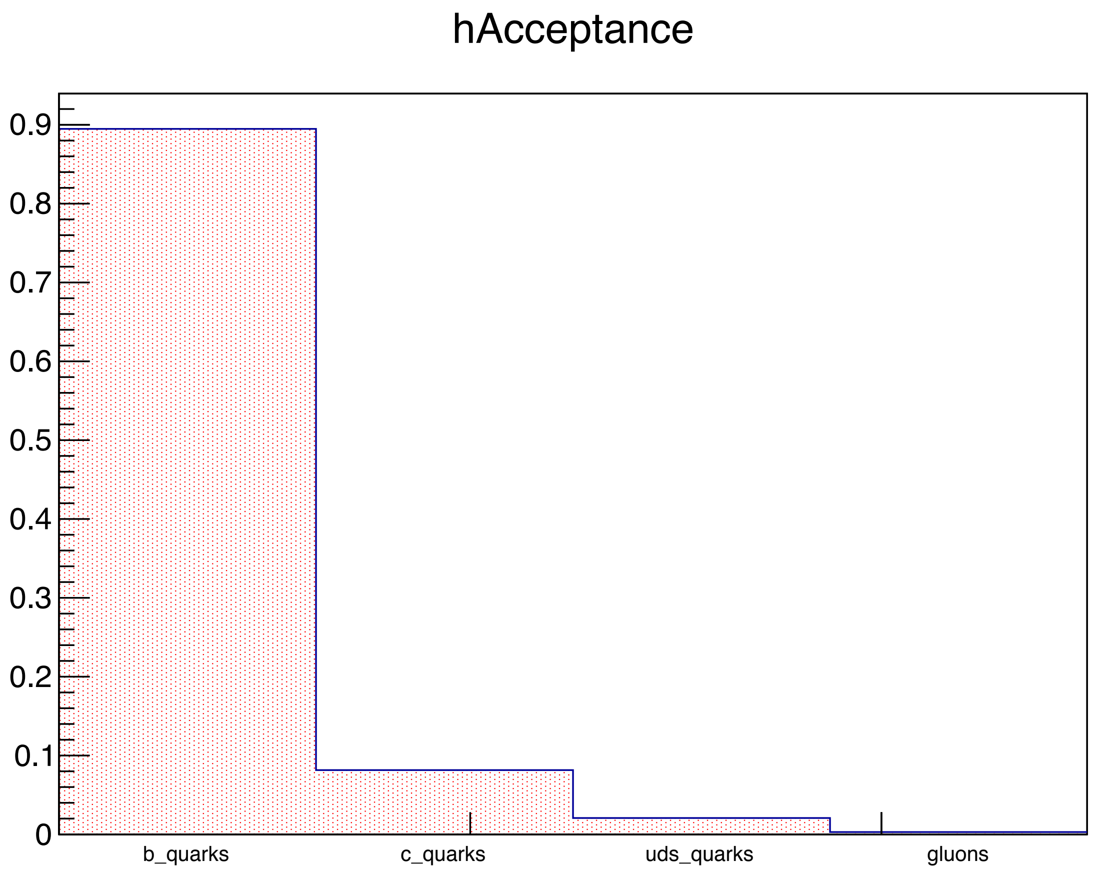




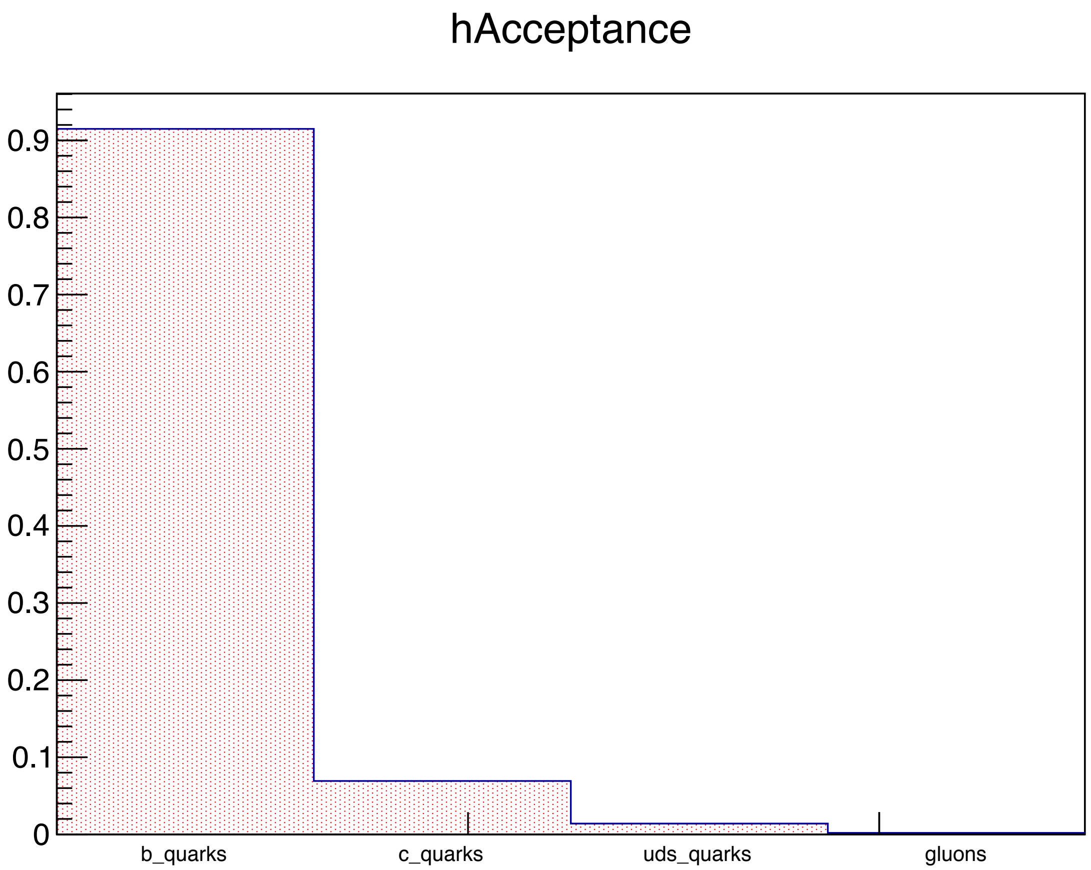


Btagging acceptance

2016

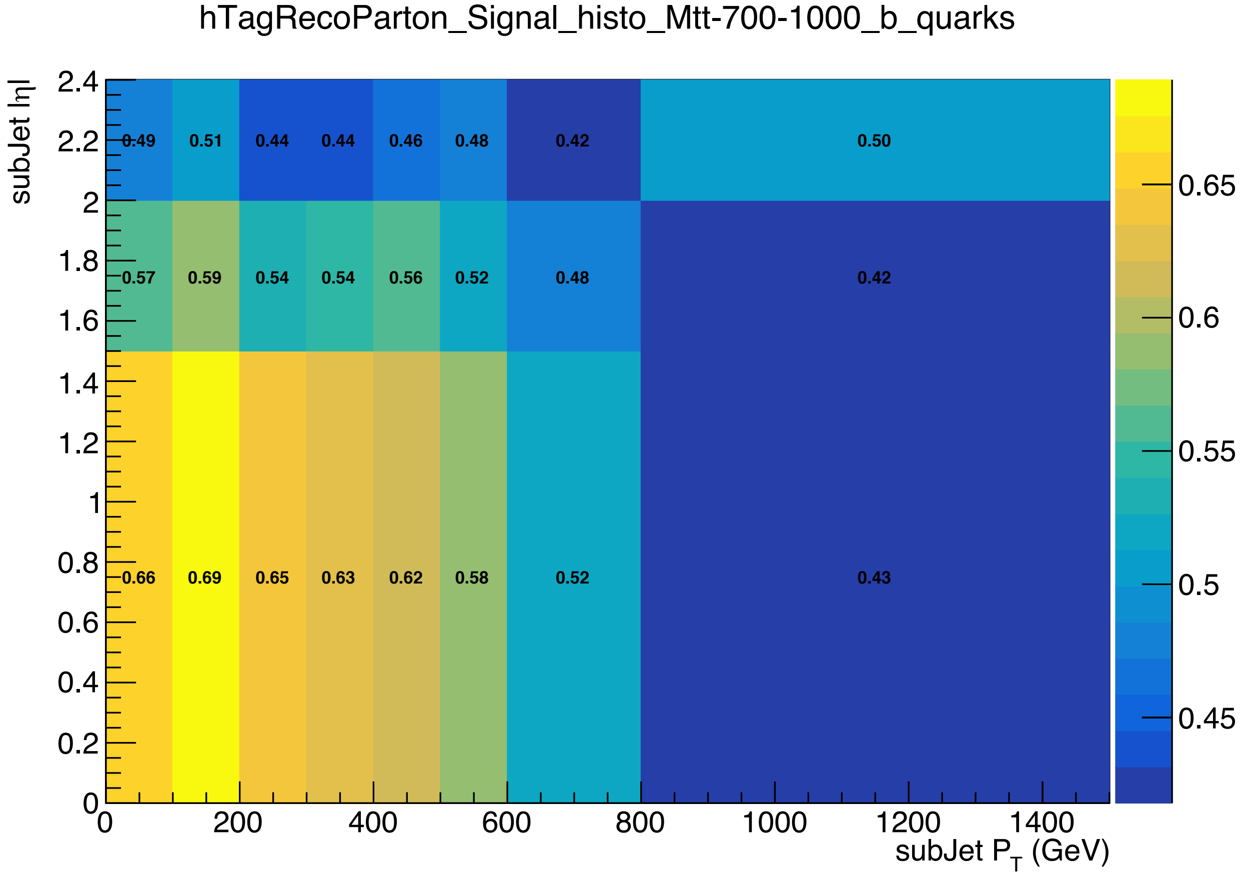


2017



Btagging efficiency in eta, pT_{subJet} phase space

2016



2017