MSSM Higgs search via bb decay at CMS experiment with Run 2016 Ultra Legacy proton proton collision data

Nazima Bi , Ye Chen

Institute of High Energy Physics, CAS Beijing

China LHC Physics Conference(CLHCP) , Qingdao

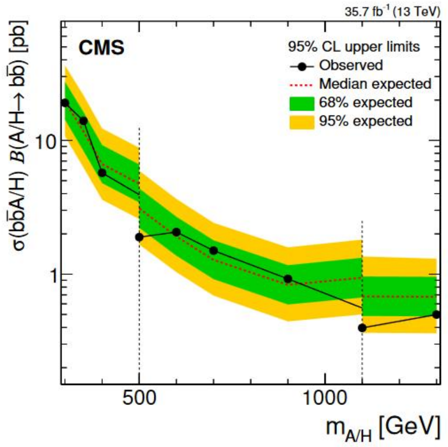
11/15/2024

Outline

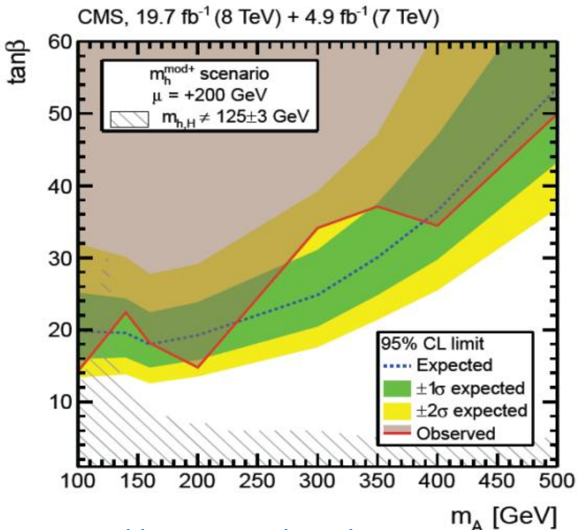
- Motivation
- Data samples
- Triggers
- Event selection
- Data / MC comparison
- Signal model
- Background model
- Systematic uncertainties
- Results & Interpretation
- Summary

Motivation

- Heavy neutral Higgs bosons predicted in many BSM extensions e.g. 2HDM and MSSM
  - enhanced b-couplings in various scenarios of these models
- Search for degenerate H and A in higher mass region:
  - dominant decay mode $H/A \rightarrow b\bar{b}$
- b-associated production:
  - cross section enhanced by ~2tan2β in MSSM and up to 2tan2β in 2HDM
  - better background control when require at least 3 b jets
- Main challenge: huge background rate from QCD multi jet production
- Run1+Run2 analyses achieved the best sensitivity in this channel to date :
  - aim to do parallel analysis with 13 TeV data



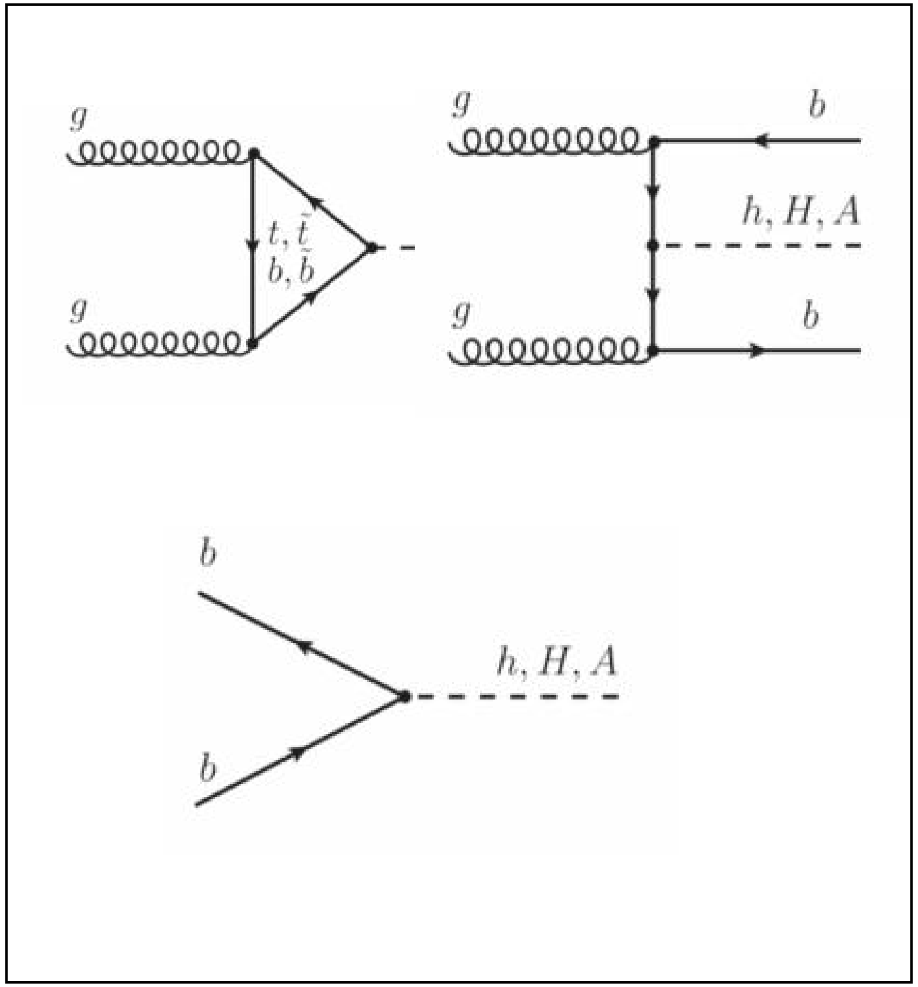
https://link.springer.com/article/10.1007/JHEP08(2018)113



https://arxiv.org/pdf/1506.08329

Analysis Strategy

- Focus on neutral Higgs bosons A/H
- Dedicated triggers requiring two online b-tagged jets
  - largely reduce rates of QCD events
- Event offline selection asks at least 3 jets to be b-tagged
- Search for a peak in dijet invariant mass (M12)
  - two leading pT b jets likely originate from Higgs
- Signal shape parametrized from MC
- Unified treatments to signal and background MC samples
  - correction and reshaping/calibration/scaling
- Full validation of background description
  - control trigger with little probability of signal production.
- Separate estimation of online selection impacts:
  - unified usage of bit selection.
  - online efficiency correction validation with defined control region.



Data & MC samples

Data Samples

❑ Data Samples

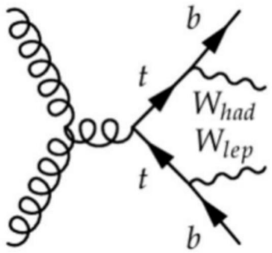
- ❑ Analysis performed on 35.7/fb at 13 TeV
  - golden JSON data collected in 2016
- ❑ Three datasets used (UL)
  - BTagCSV
  - JetHT

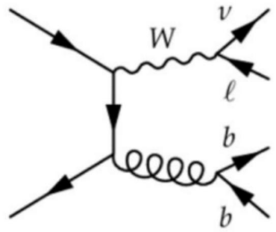
❑ Monte Carlo samples

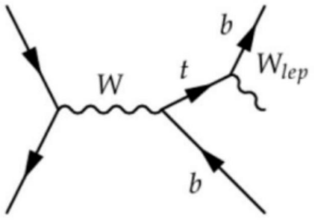
- ❑ Signal events simulated with Powheg at 23 different higgs mass points.

❑ Background

- QCD
- Single Top
- Top pair production
- WJets
- ZJets
- DYJets







Triggers

Triggers

- Physics triggers for analysis (unprescaled)

  HLT DoubleJetsC100 DoubleBTagCSV p026 DoublePFJetsC160

  HLT DoubleJetsC100 DoubleBTagCSV p014 DoublePFJetsC100MaxDeta1p6

- Control triggers (prescaled)

  HLT PFJet140 v*

- Two type of a control triggers (Background modeling):
  - HLT PFJet200 v* for mass range (900,1400)GeV
  - HLT PFJet260 v* for mass range (1400,2000) GeV

Event Selection

Event Selection

Signal Region

- ❑ Physics trigger bit
- ❑ 3 leading pT "Jets" loose ID
- ❑ 1st and 2nd jets
  - pT ≥ 170 GeV
  - |η| ≤ 2.1
  - Medium b tag WP
- ❑ 3rd jet
  - pT > 170 GeV
  - |η| ≤ 2.1
  - Medium b-tag WP

Control Region

- ❑ Physics trigger bit
- ❑ 3 leading pT "Jets" loose ID
- ❑ 1st and 2nd jets
  - pT > 170 GeV
  - |η| < 2.1
  - Medium b tag WP
- ❑ 3rd jet
  - pT > 170 GeV
  - |η| < 2.1
  - Medium b-tag WP

Signal Model

Signal Shape

- Signal shape (M12) from MC for mass of 400-2000 GeV
- Fitting Model: Novosibirsk function describes Higgs peaks

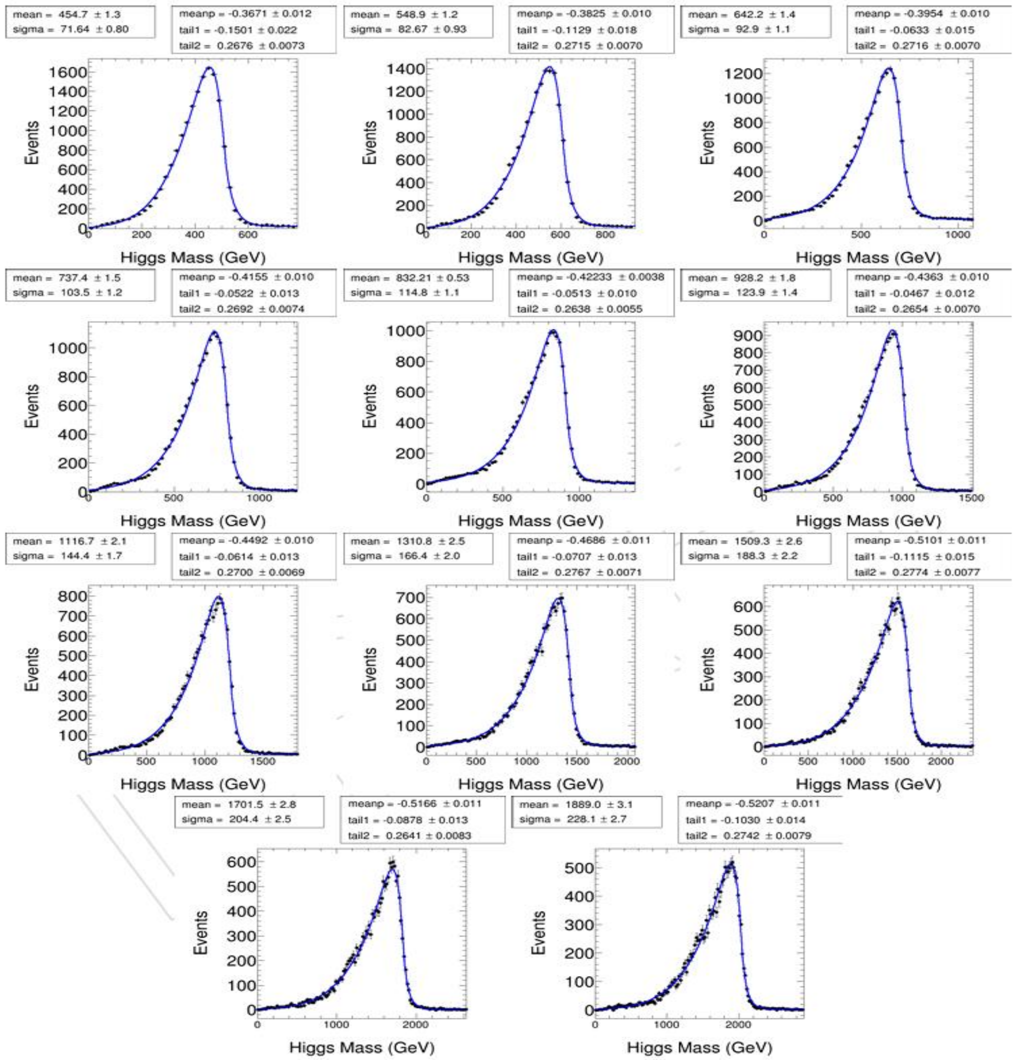
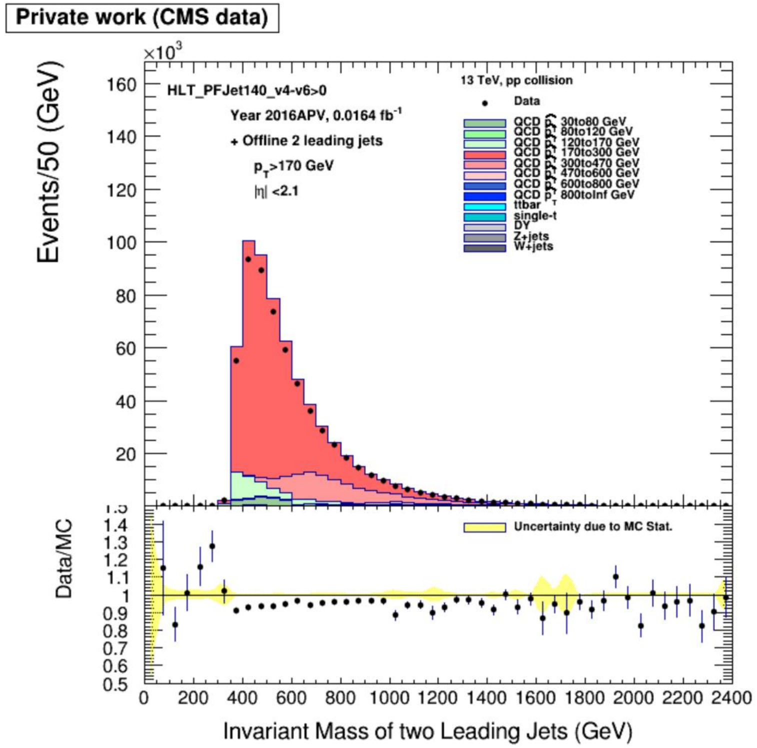


Figure: Fitted Invariant Mass spectrum with Novosibirsk function

Data & MC Comparison

Data MC Comparison: 2 Jets

- Data MC comparison of the vector sum of two leading jet invariant mass spectrum using control trigger and 2 offline leading jet criteria



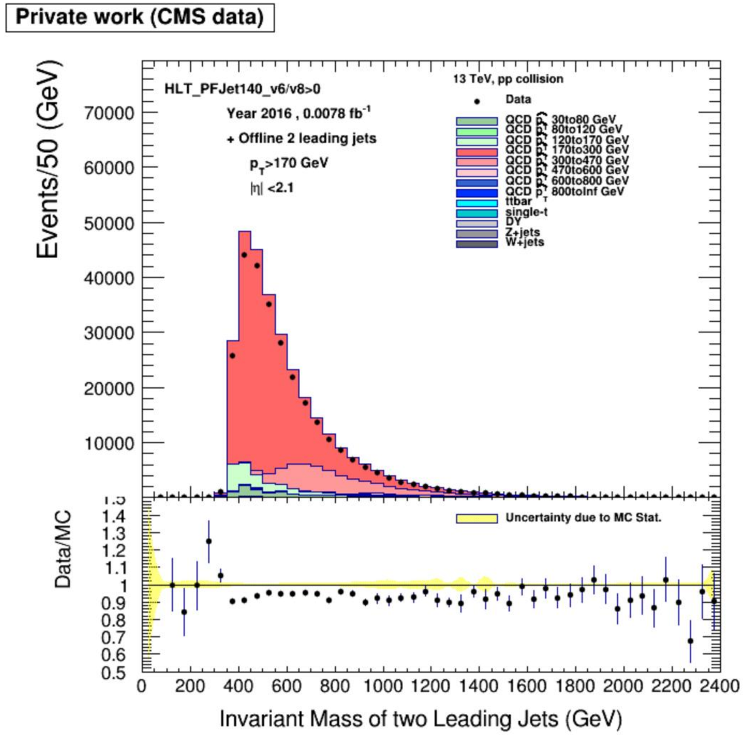
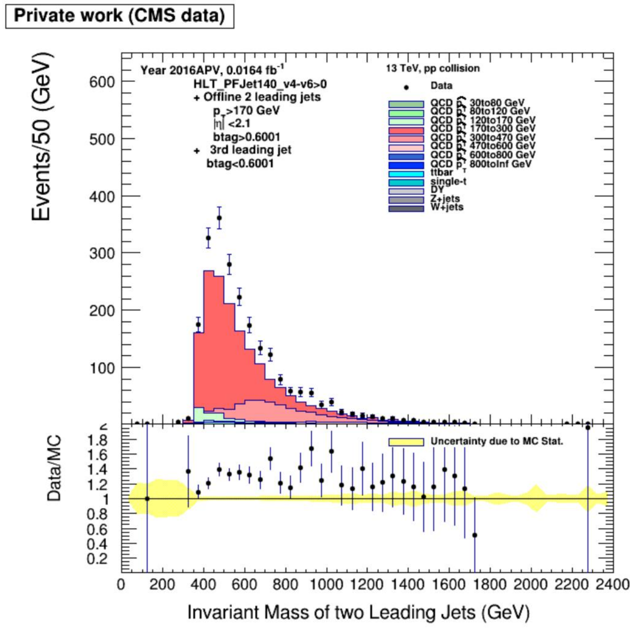


Figure: Data MC comparison plots for 2016APV (left) and 2016 (right)

Control Region with Pre-scaled

- Data MC comparison of the vector sum of two leading jet invariant mass spectrum using control trigger
- B-tagging scale factors are applied to correct for differences between the simulation and real data



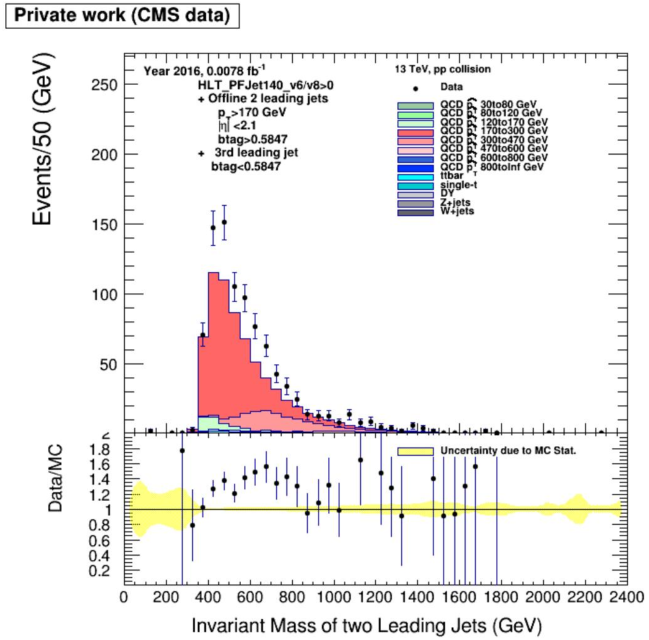


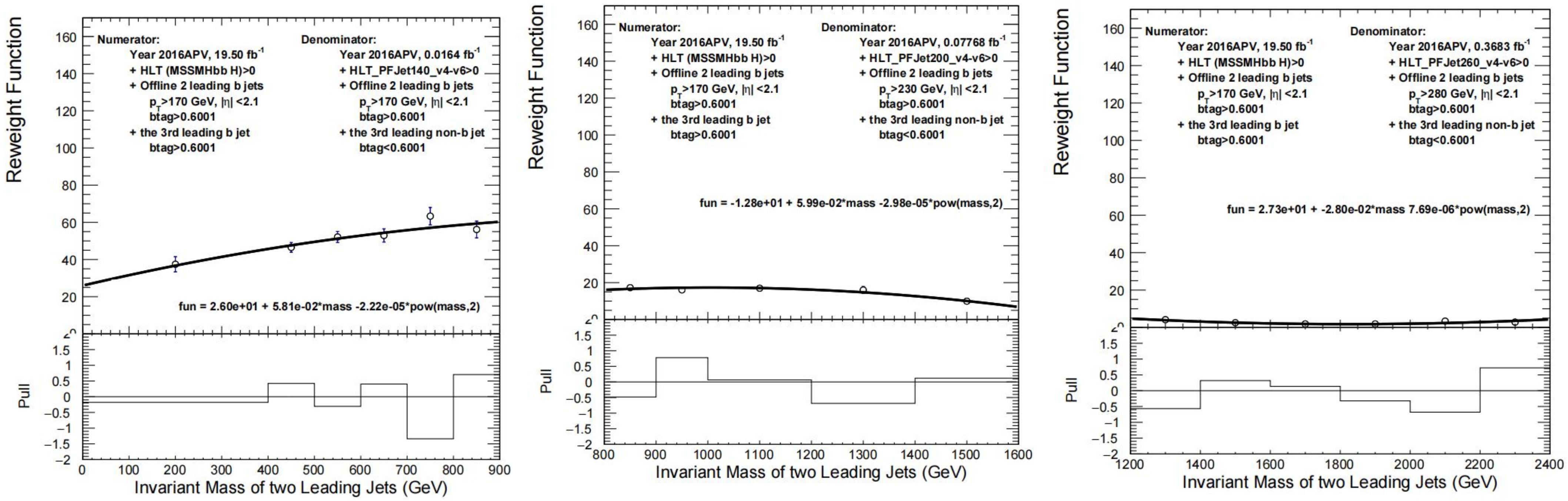
Figure: Data MC comparison plots for 2016APV (left) and 2016 (right)

- **conclusion:**
  - Improvements in Data/MC agreement in 2bNo3b control region
  - However background re-modeling still needed.

Background Re-modeling

Background Re-modeling

QCD Re-modeling: Reweight function



- QCD reweighted from data in control region taken by control trigger

$$ReweightFunction = \frac{N_{SignalRegionQCDMC(MSSMtriggeredwithallcorrections)}}{N_{ControlRegionMC(PFJettriggeredwithallcorrections)}}$$

$$N_{QCD(SignalRegion)} = RF * N_{datafromControlRegion(PFJettriggered)}$$

Background Re-modeling

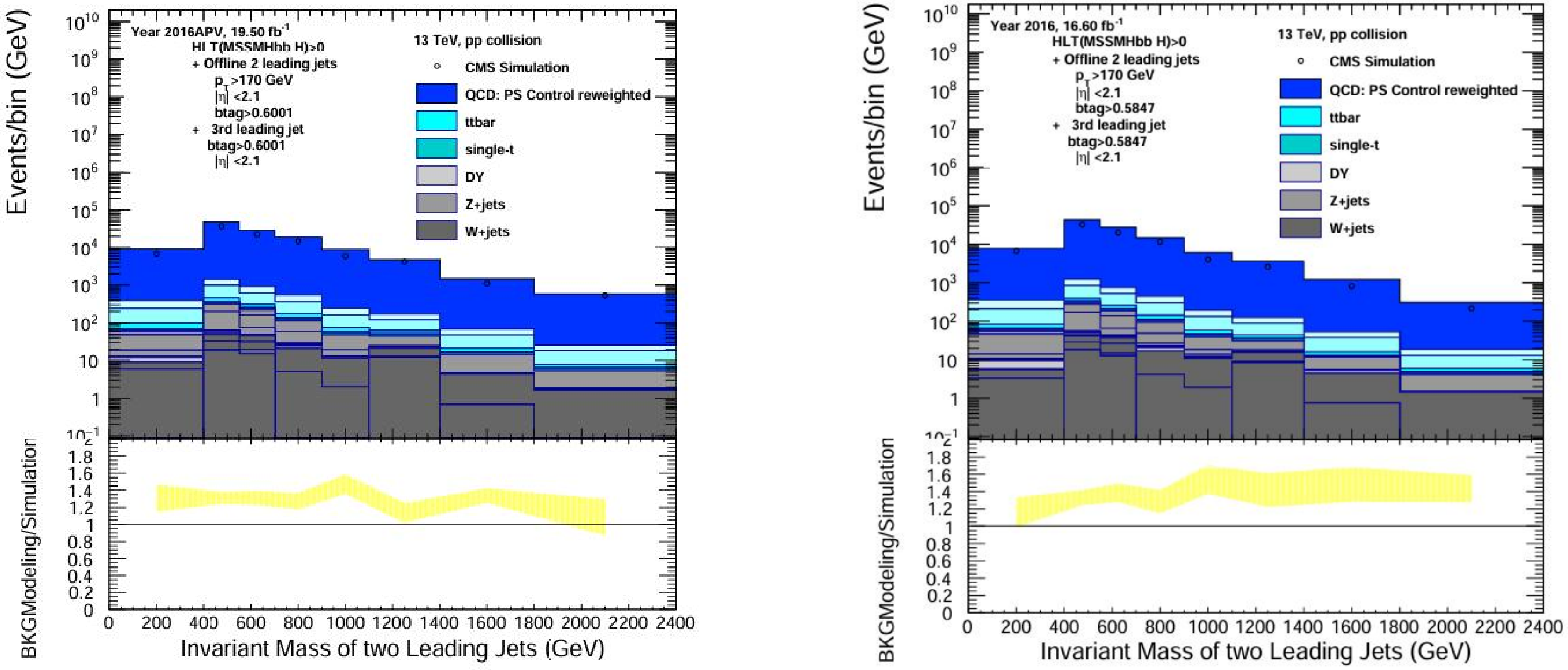


Figure:Comparison of background remodeling to simulation

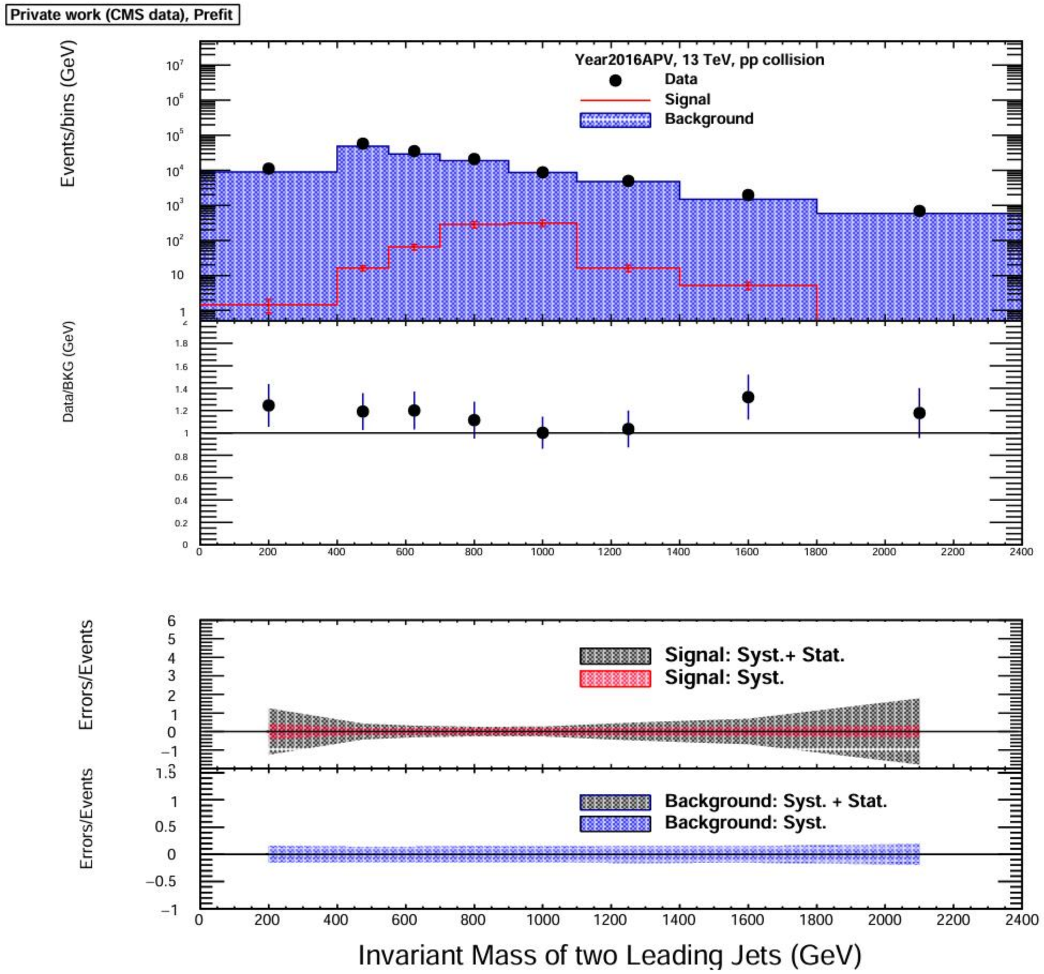
- ❑ Background re-modeling with QCD from prescaled data.
- ❑ Event by event unbinning reweight has been performed for QCD components

Systematic Uncertainty

Major Systematic Uncertainty

- There are many source of systematics:
- Experimental Uncertainties:
  - Luminosity: Relates to the precision of luminosity measurement in data.
  - Kinematical Trigger: Arises from the efficiency of triggers based on particle kinematics.
  - Pileup: Involves uncertainties due to overlapping events in the detector, especially at high collision rates.
  - Jet: Involves uncertainties in jet energy scale, resolution, and reconstruction in the detector.
  - BtagSF (B-tagging Scale Factor): Refers to the calibration of the b-tagging algorithm and how well it performs in data vs. simulation.
- Simulated Sample Uncertainty:
  - Background Modeling: Refers to the accuracy of background predictions, often derived from simulated processes or data-driven methods.
- These uncertainties are incorporated into the final fitting model to ensure accurate measurements and interpretations.

Prefit Distributions



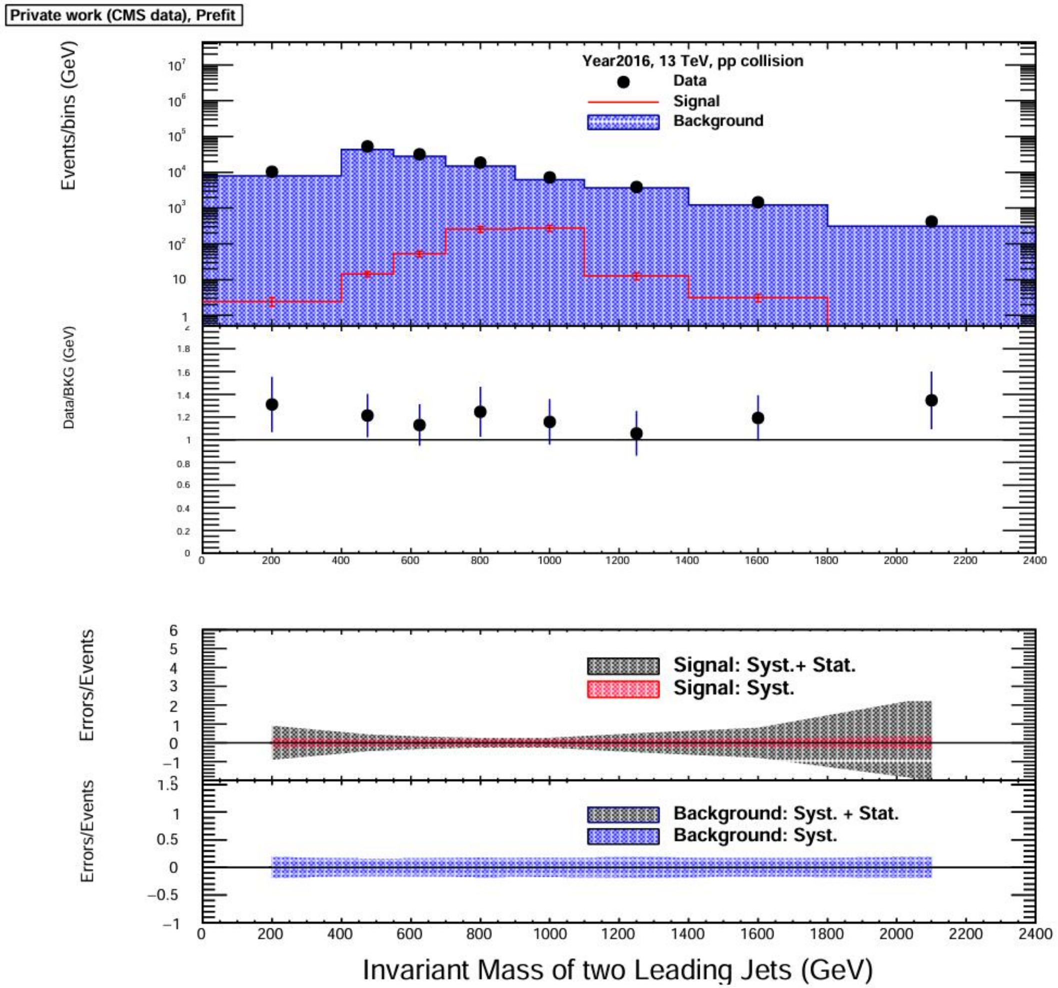
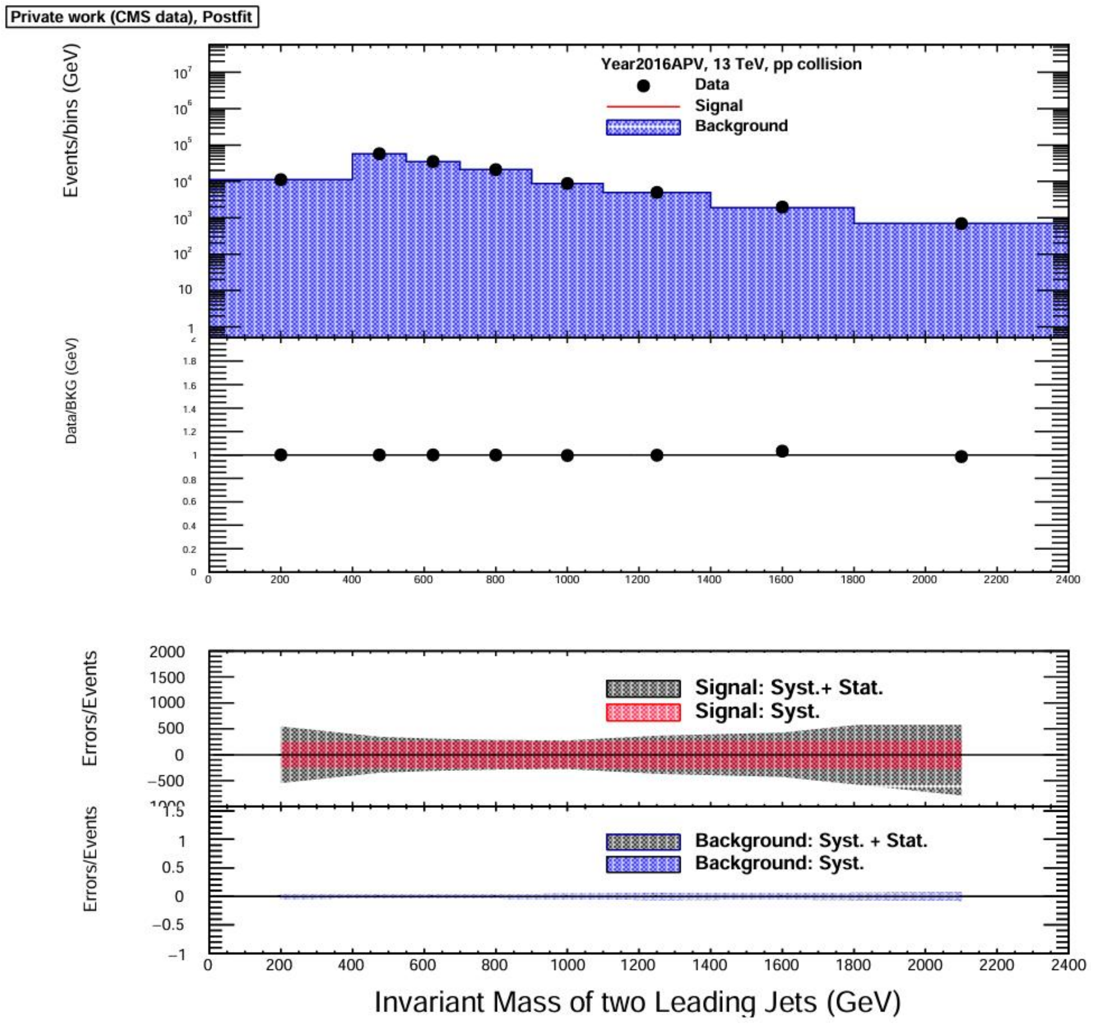


Figure:Prefit Distributions together with Data/BKG comparison. 2016APV(Left);2016(Right)

Postfit Distributions



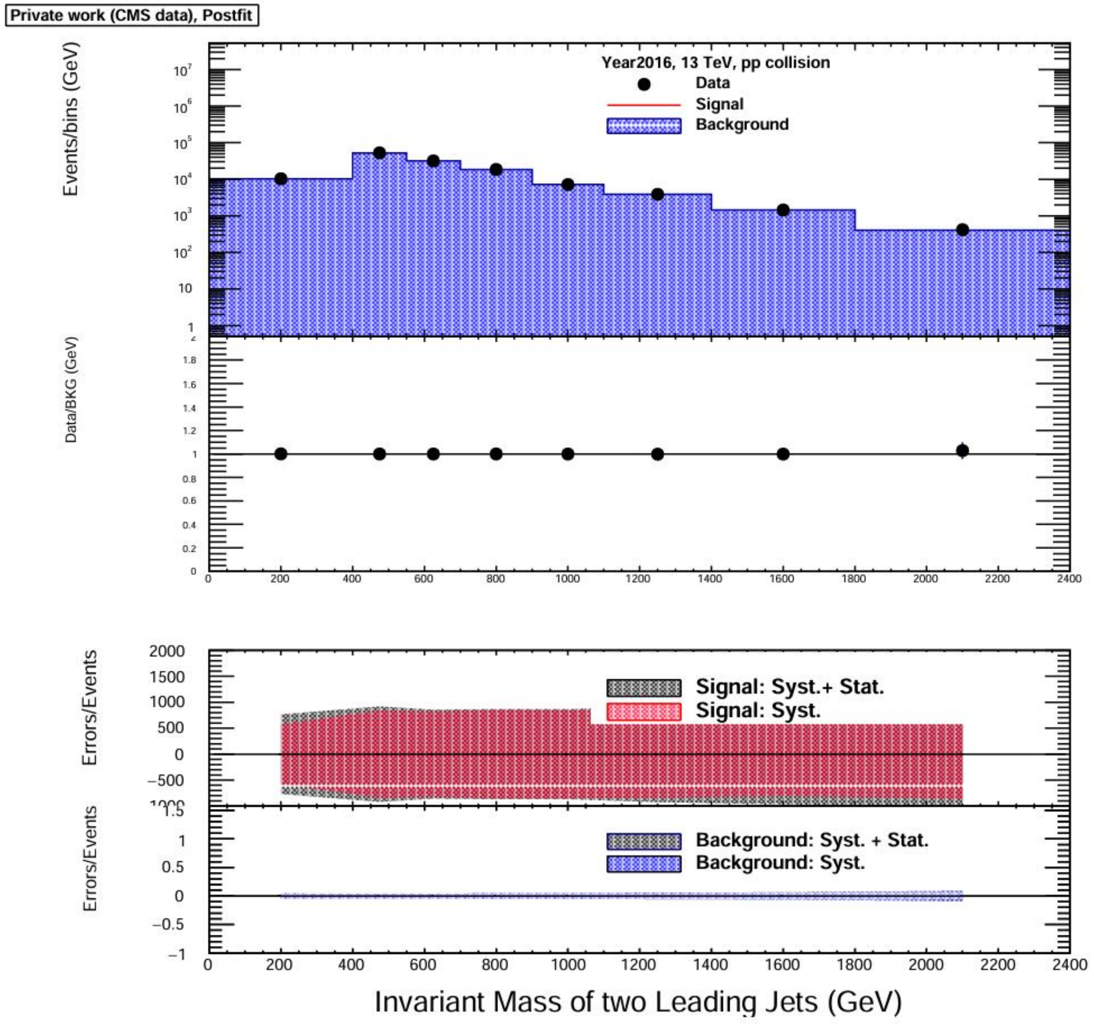
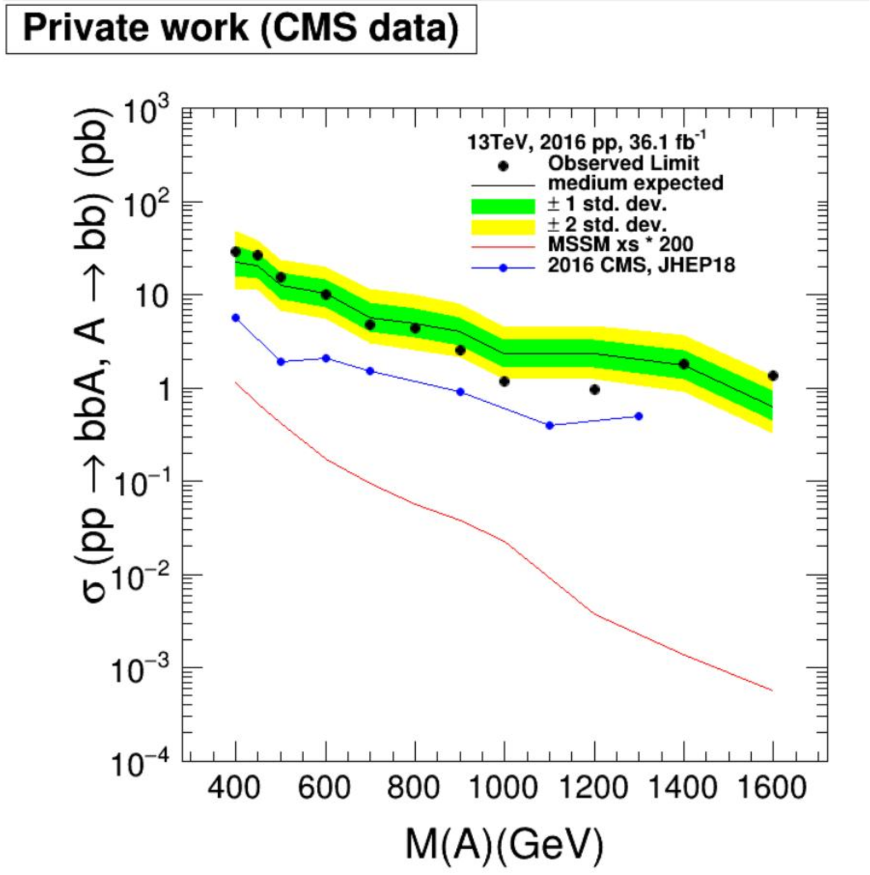


Figure:Postfit Distributions together with Data/BKG comparison. 2016APV(Left);2016(Right)

Year2016 Combined

- Upper limits of production cross section of MSSM Higgs.



- **Conculsion**
  - No signal is observed

Summary

- presented a search for high mass Higgs bosons in association with b quarks, specifically decaying into a b-quark pair, using the 2016 Ultra-legacy dataset collected at 13 TeV
- This analysis was conducted using a very conservative method, which has impacted the overall sensitivity and results.
- study contributes to the ongoing effort to explore BSM extensions and enhances our understanding of high mass Higgs bosons

Thank You

BACKUP

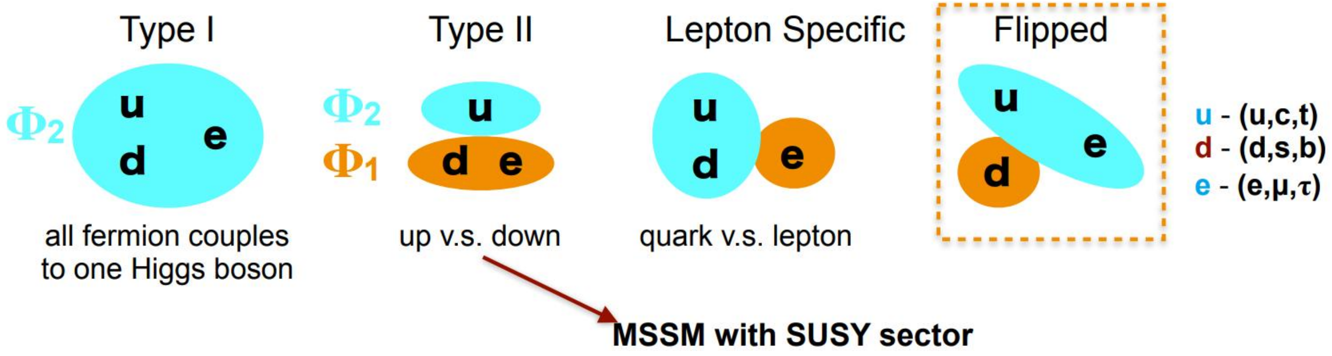
Two Higgs Doublet Model

- Two Higgs Doublet Models (2HDM) extend Higgs sector beyond the SM by including

two complex Higgs doublets, leads to 5 physical Higgs bosons

H+, H-, A (CP-odd), H and h (CP-even)

- There are 4 types of 2HDM which lead to natural flavor conservation, they differ in the way how 2 Higgs doublet fields couple to SM particles



- MSSM has the same Higgs sector structure as 2HDM type-II

▪ described with two free parameters m_A and tanβ $(\frac{\nu_u}{\nu_d})$ at tree leve

Physics objects

- Standard primary vertex (PV) identification:
  - offline slimmed primary vertex
  - $|z| < 24$ cm, $|dxy| < 2$cm (w.r.t. beam spot), n.d.o.f. in vertex fit > 4
  - not fake vertex
- Standard particle-flow (PF) anti-kT (R=0.4) jets with pile-up correction
  - slimmedJets: charged-hadron subtraction + per-event pile-up energy density and jet area ('L1FastJet correction') corrections
  - particle-jet based L2L3 corrections and (data only) L2L3Residual corrections
  - pT > 10 GeV
  - looseID
- b-tagging: combined secondary vertex (CSVv2) algorithm with medium working point

Signal Shape

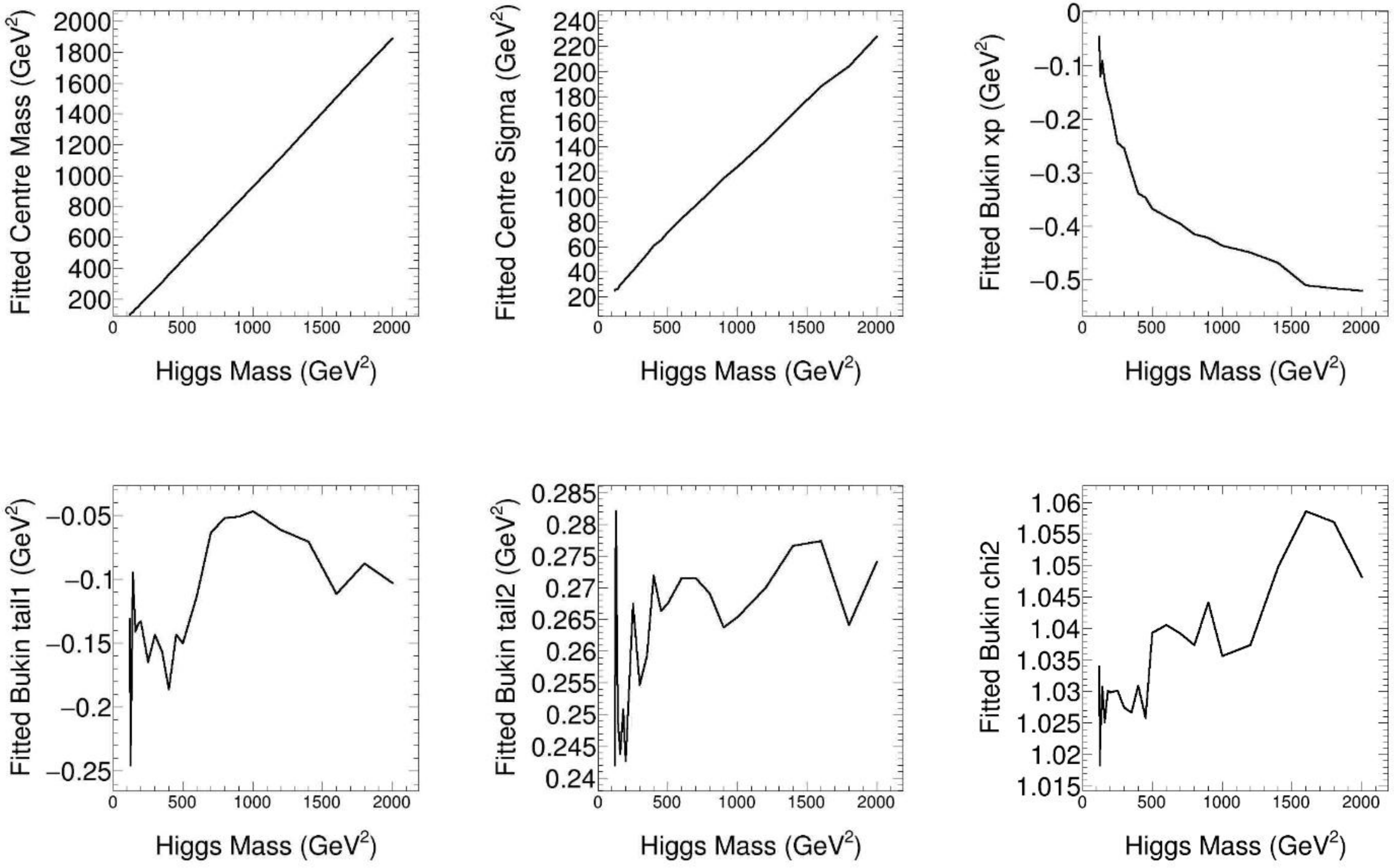


Figure: Fitted Parameters of the Novosibirsk functions for all Higgs mass points

Selection Optimization

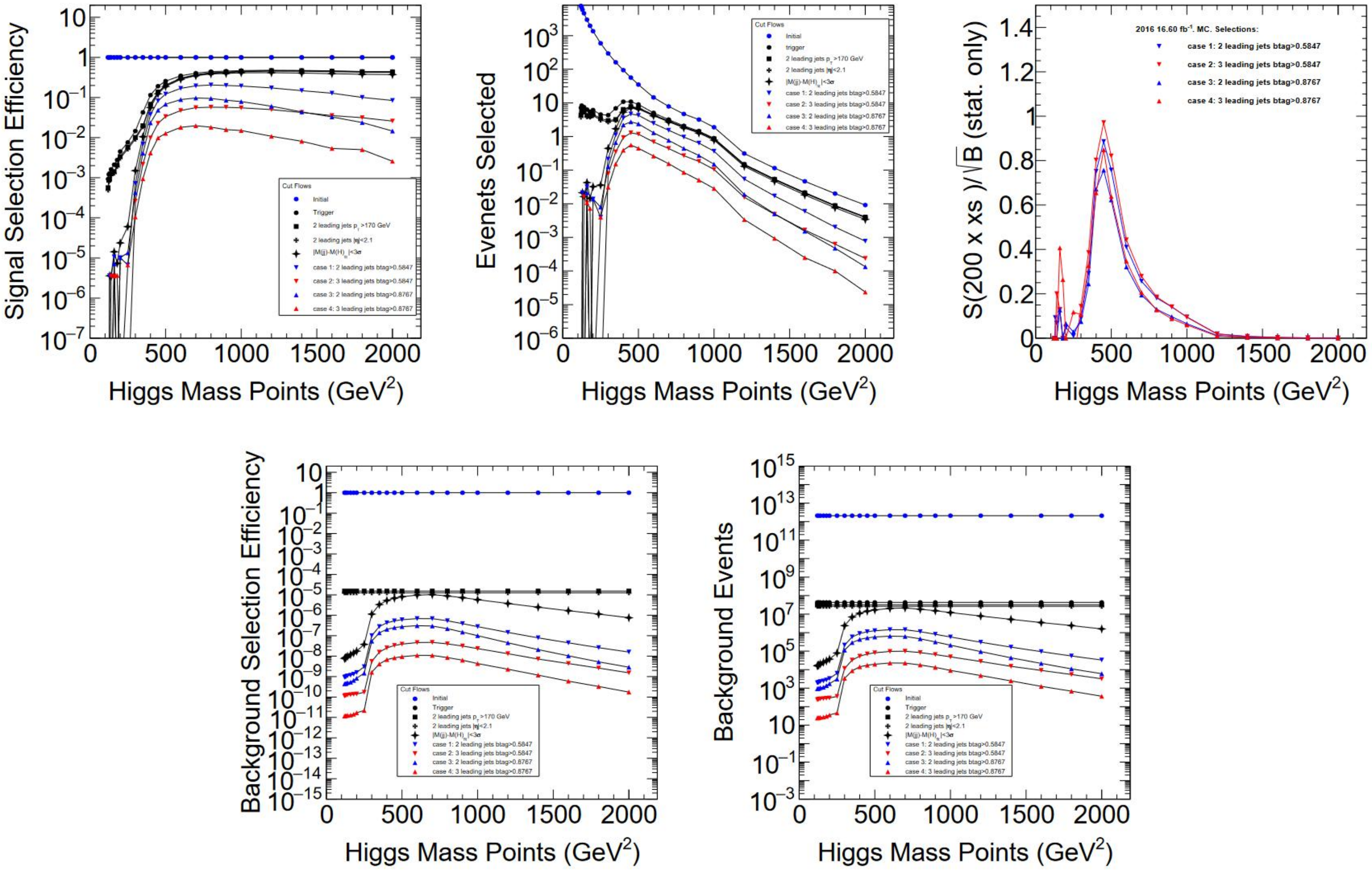
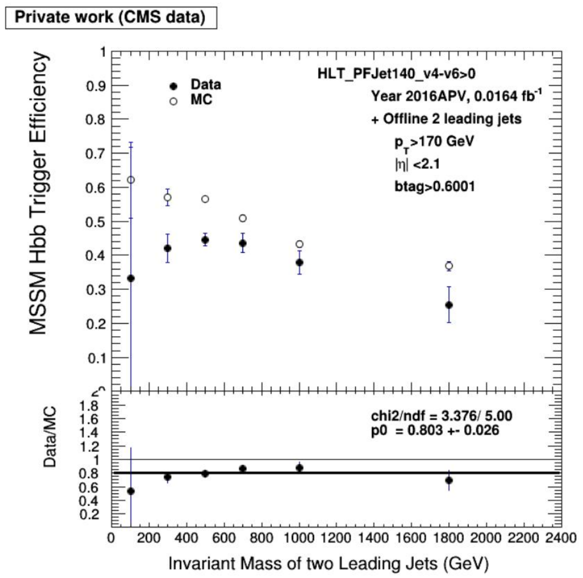


Figure: Offline selection cut flows with QCD and other background under each of MSSM Higgs samples

Trigger Efficiency

- Trigger efficiency is measured with control triggered datasets
- trigger efficiency has been defined

$$Eff = \frac{Npass(MSSMH_{bb}andTriggerPFJet_{140}TriggerandOfflineSelections)}{NPass(PFJet_{140}TriggerandOfflineSelections)}$$



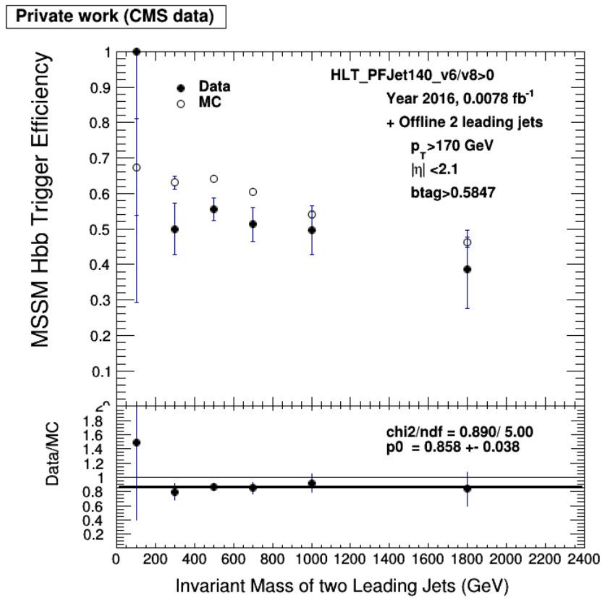
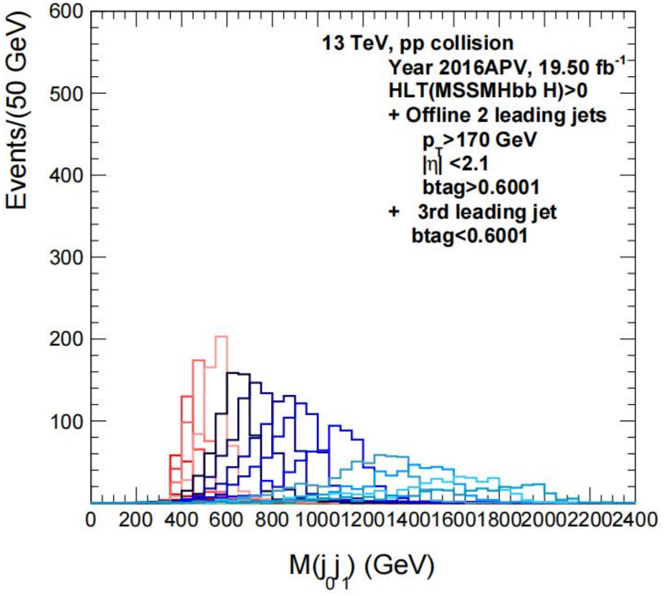


Figure: Rapid Trigger Efficiency Estimation or Year2016APV (left) and Year2016 (right)

MSSMHbb Triggered: Signal Region

- Comparison of the vector sum of two leading jet invariant mass spectrum
- MSSM Hbb trigger "HLT DoubleJetsC100 DoubleBTagCSV p026 DoublePFJetsC160 v*" and 3 offline leading b-tagging jet criteria
- B-tagging scale factors are applied to correct for differences between the simulation and real data



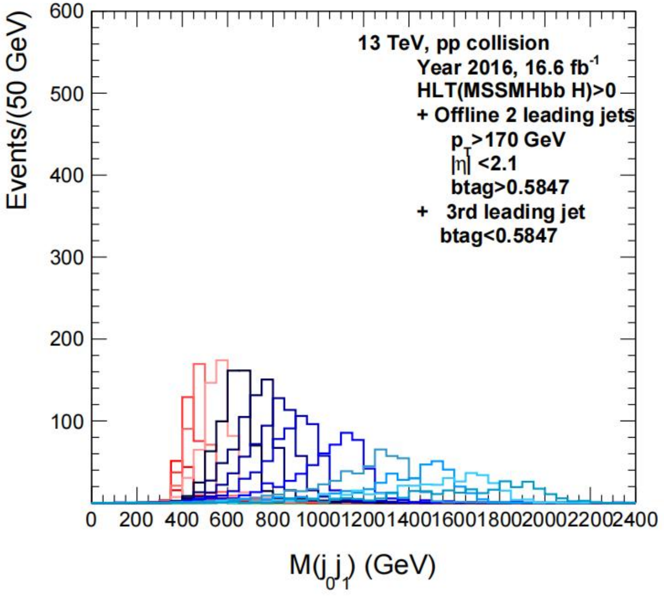


Figure: Comparison plots for 2016APV (left) and 2016 (right)

Control Region with MSSMHbb Trigger

- MSSM Hbb trigger "HLT DoubleJetsC100 DoubleBTagCSV p026 DoublePFJetsC160 v*" and 2 offline leading b-tagging jets plus 3 nonb-tagging jet criteria
- B-tagging scale factors are applied to correct for differences between the simulation and real data

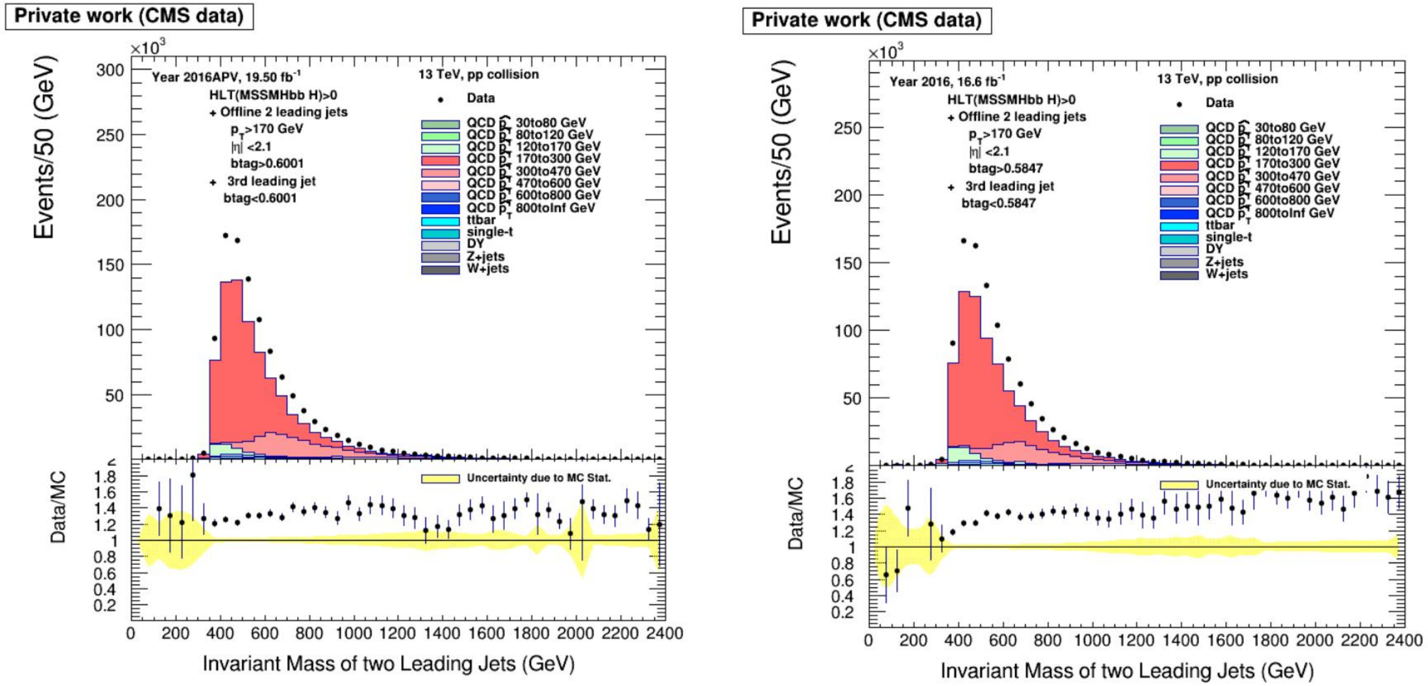
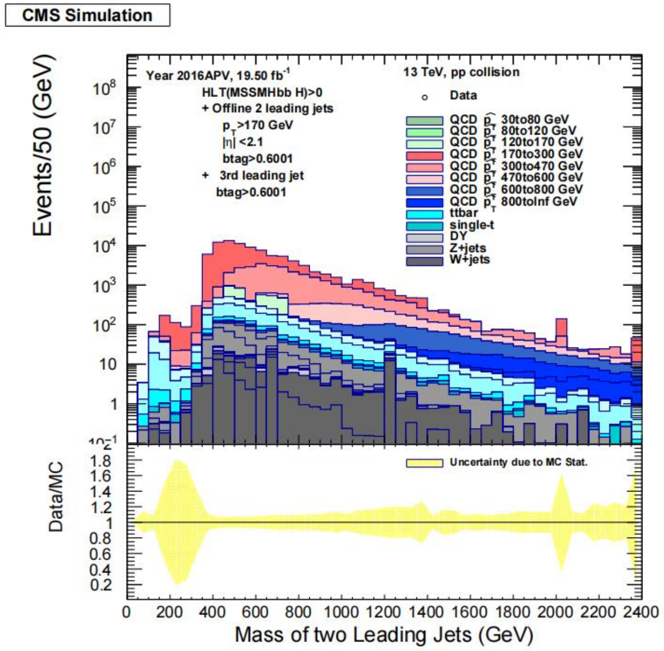


Figure: Data MC comparison plots for 2016APV (left) and 2016 (right)

- Comparison of the vector sum of two leading jet invariant mass spectrum
- MSSM Hbb trigger "HLT DoubleJetsC100 DoubleBTagCSV p026 DoublePFJetsC160 v*" and 3 offline leading b-tagging jet criteria
- B-tagging scale factors are applied to correct for differences between the simulation and real data



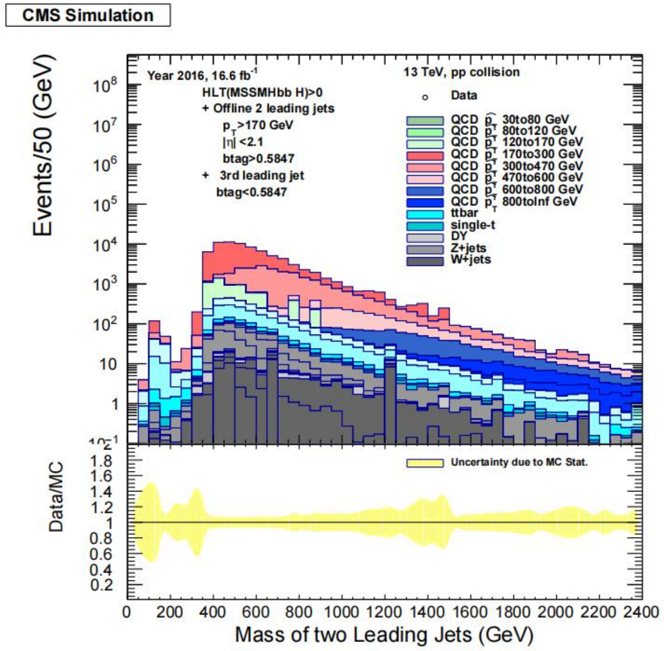
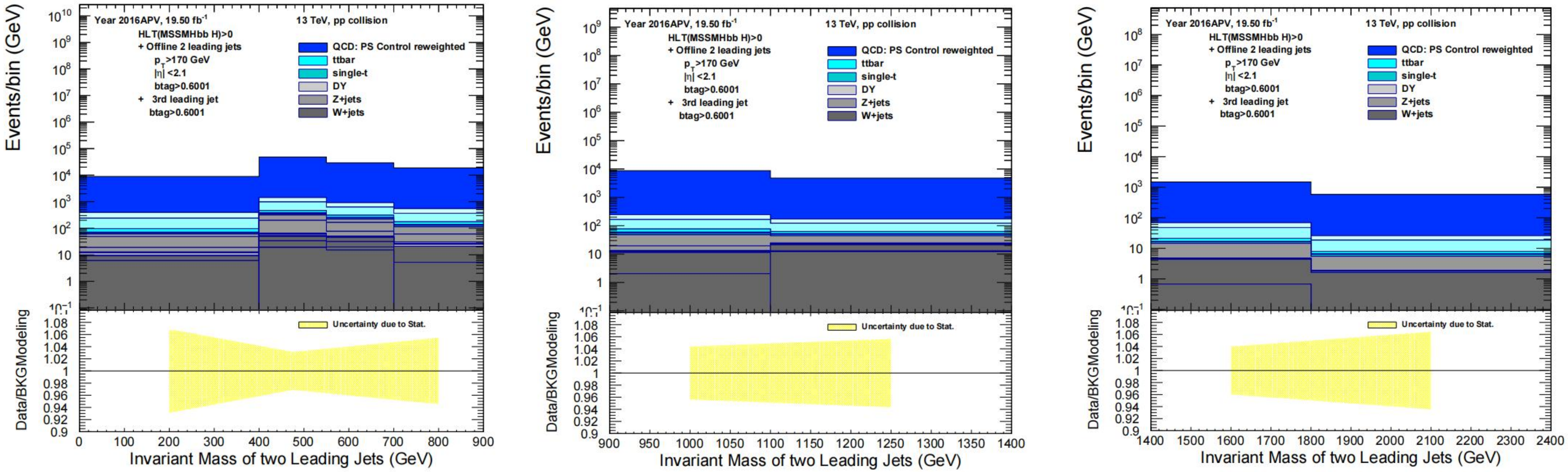


Figure: Comparison plots for 2016APV (left) and 2016 (right)

Background Re-modeling

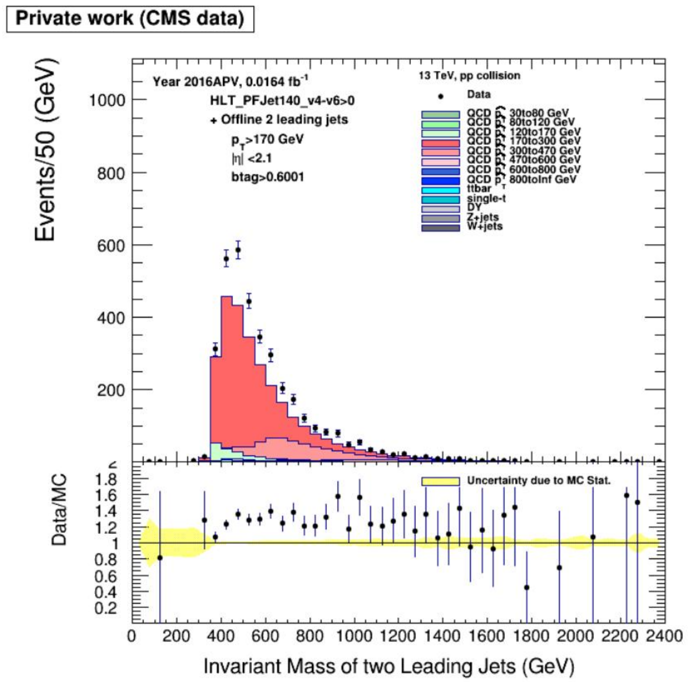
- MSSMHbb Triggered: remodeled Year2016APV signal region.



- Background re-modeling with QCD from prescaled data.
- Event by event unbinning reweight has been performed for QCD components.

Data MC Comparison: 2b Jets

- Data MC comparison of the vector sum of two leading b-tagging jet invariant mass spectrum
- control trigger HLT PFJet140 v*and 2 offline leading b-tagging jet criteria
- B-tagging scale factors are applied to correct for differences between the simulation and real data



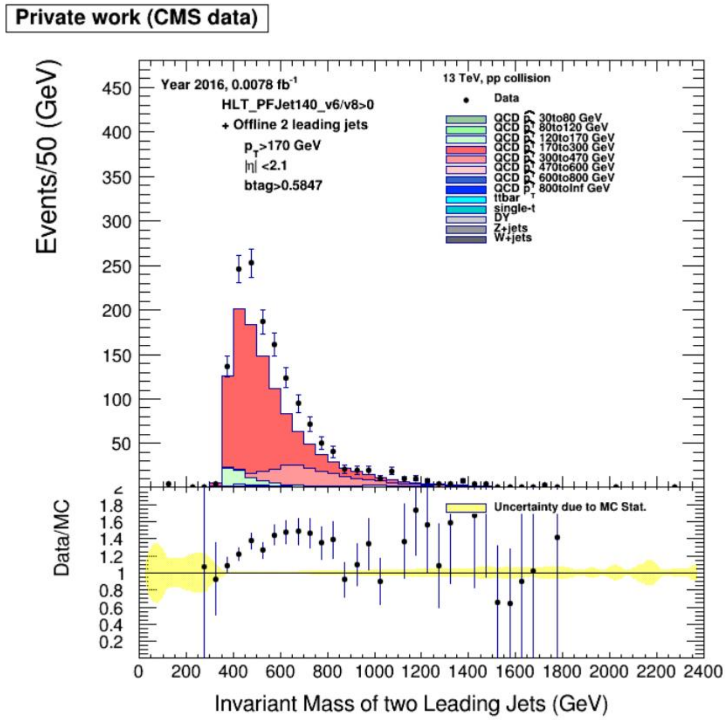
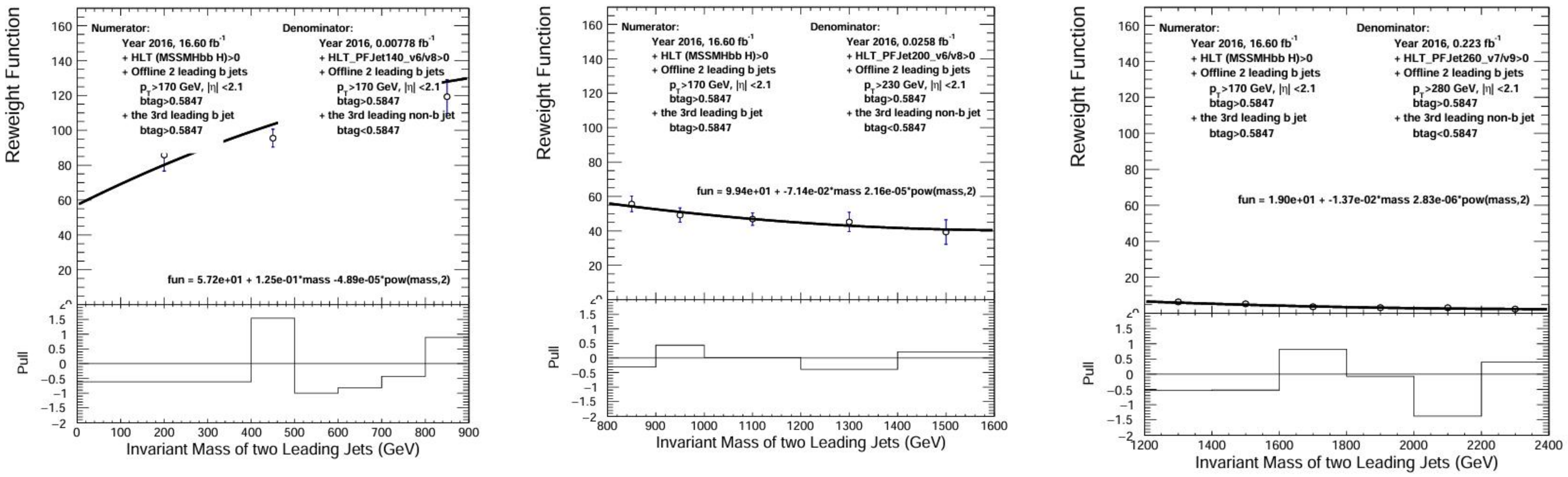


Figure: Data MC comparison plots for 2016APV (left) and 2016 (right)

Background Re-modeling

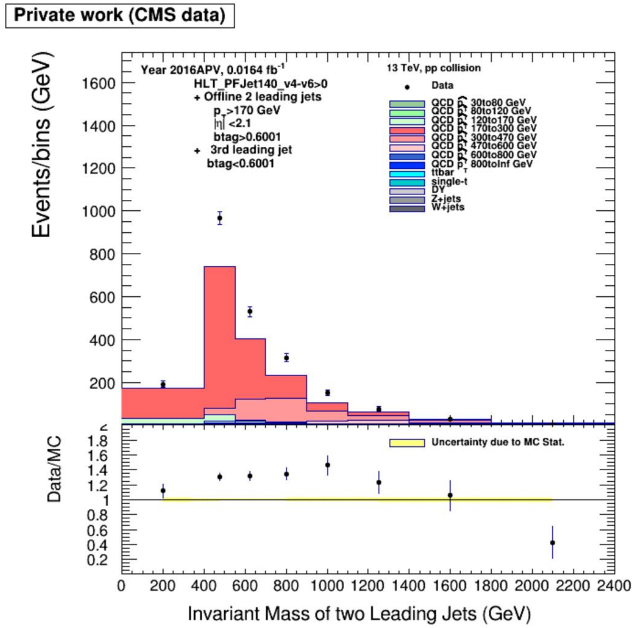
QCD Re-modeling: Reweight function (Year2016)

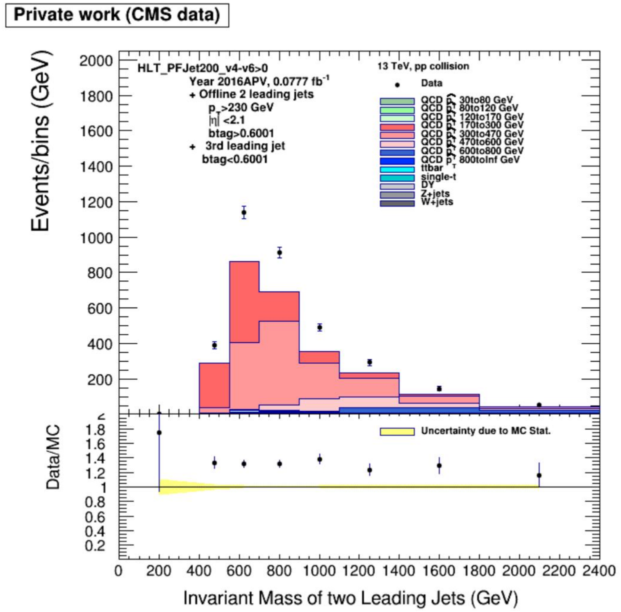


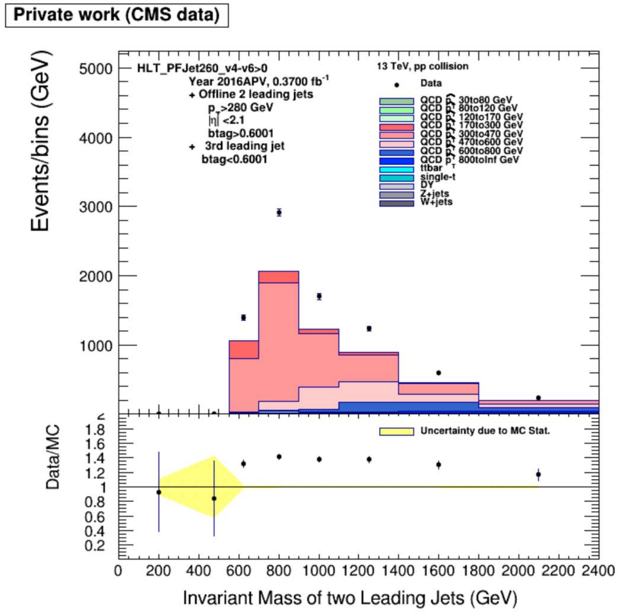
mass region	[0,900] GeV	[900,1400]GeV	[1400,2400] GeV
control trigger	HLT_PFJet140_v*	HLT_PFJet200_v*	HLT_PFJet260_v*
additional requirement	-	leading jet pt>230GeV 2nd jet pt>230GeV	leading jet pt>280 GeV 2nd jet pt>280 GeV

Background Re-modeling

- Control region by control trigger (Year2016APV)



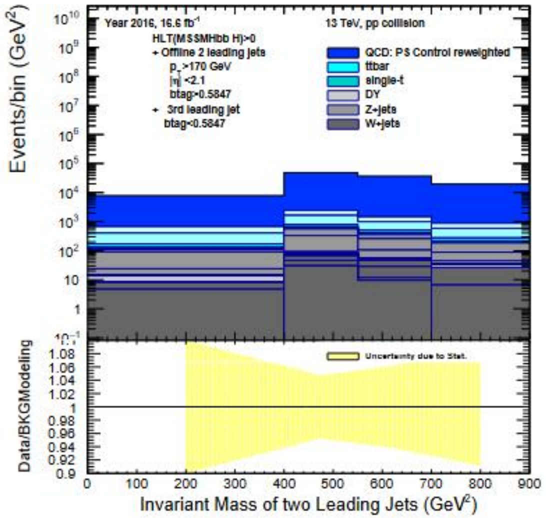


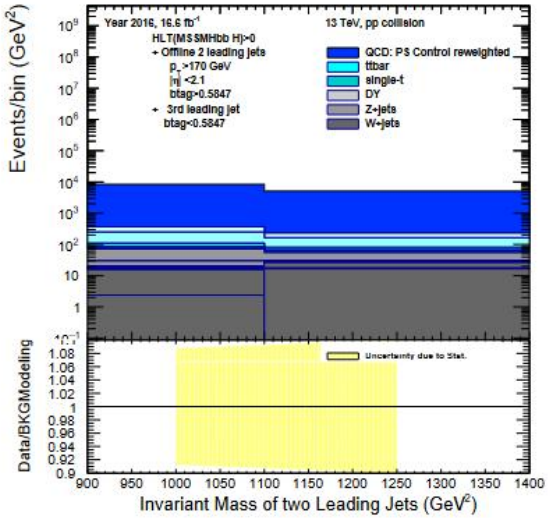


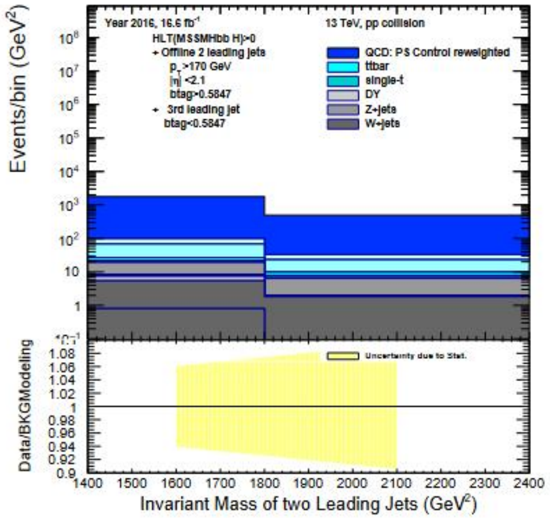
mass region	[0,900] GeV	[900,1400]GeV	[1400,2400] GeV
control trigger	HLT_PFJet140_v*	HLT_PFJet200_v*	HLT_PFJet260_v*
additional requirement	-	leading jet pt>230GeV 2nd jet pt>230GeV	leading jet pt>280 GeV 2nd jet pt>280 GeV

Background Re-modeling

□ **MSSMHbb Triggered: remodeled Year2016 signal region.**



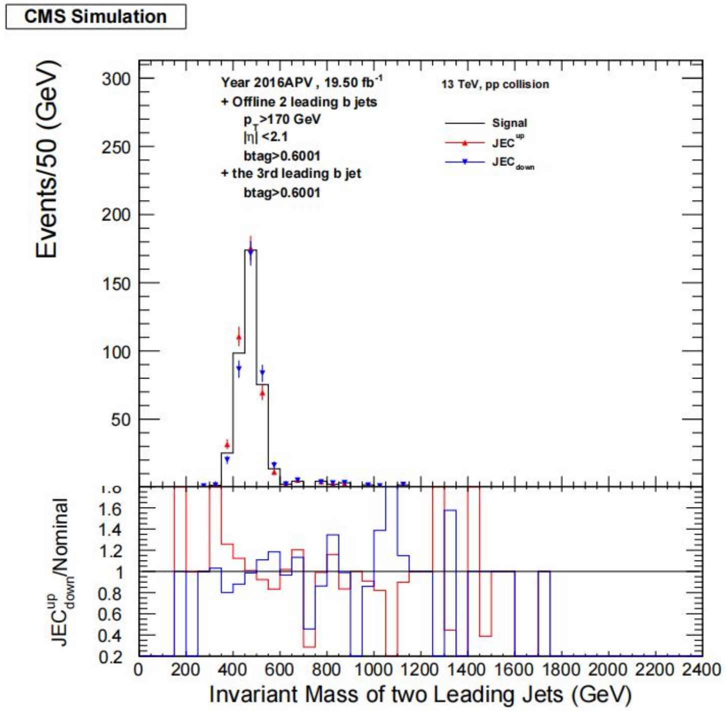




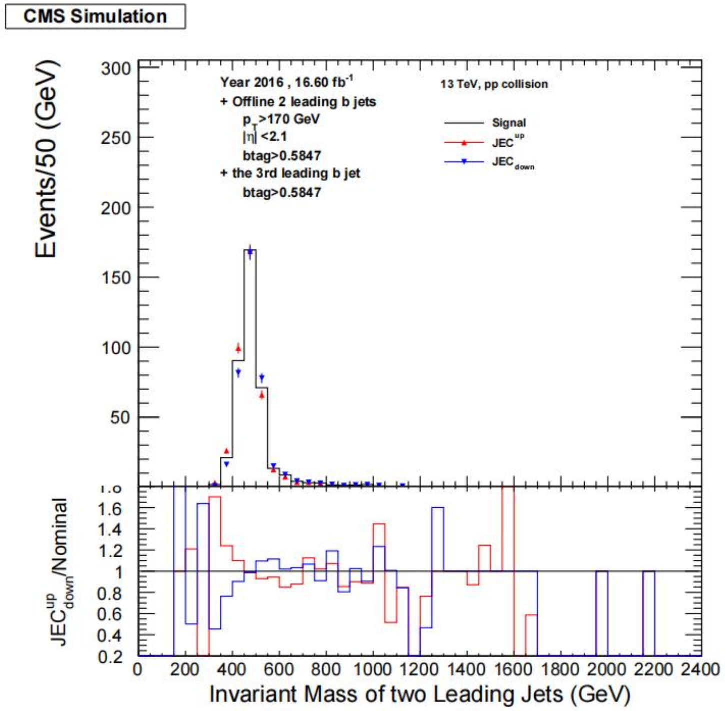
- ❑ Background re-modeling with QCD from prescaled data.
- ❑ Event by event unbinning reweight has been performed for QCD components

Nuisance Parameters: JEC

Year2016APV



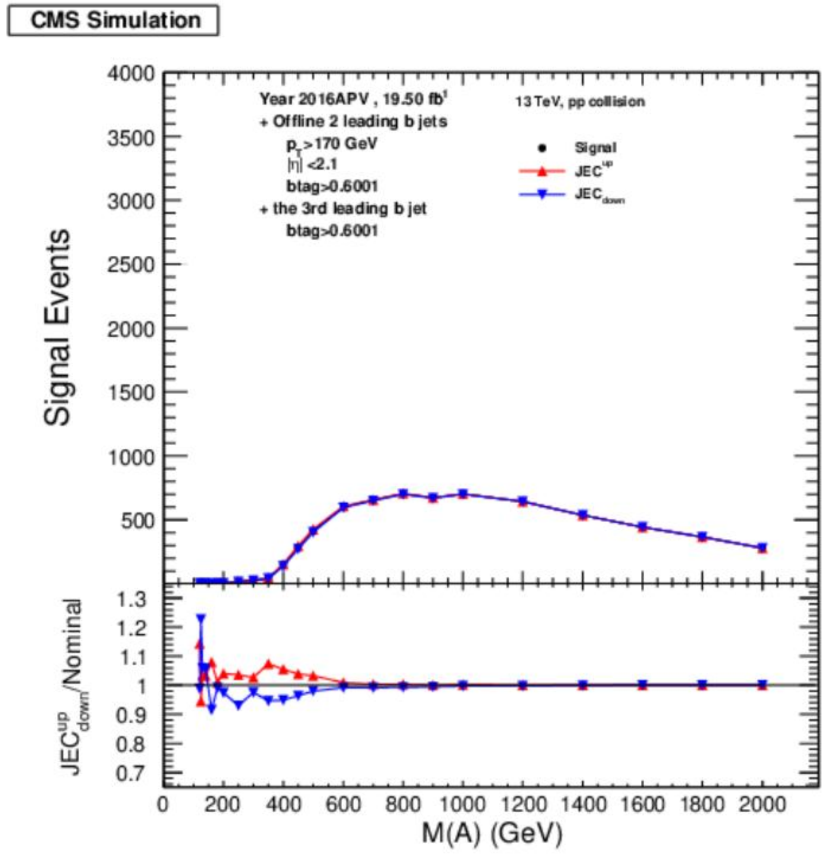
Year2016



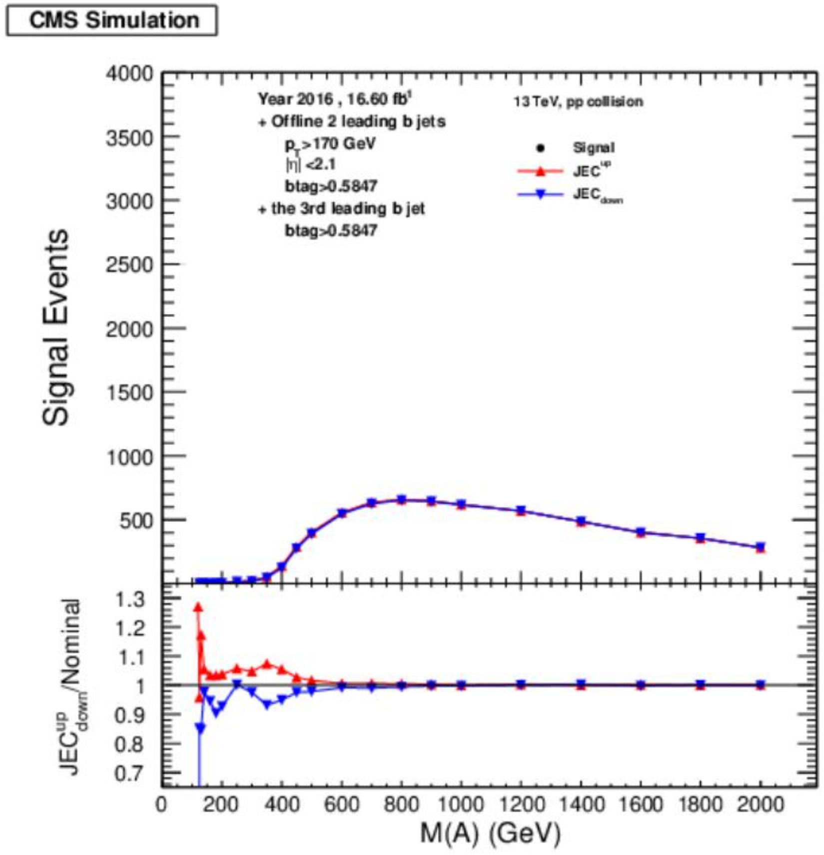
- Following JETMET POG official recommendation:
  - https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookJetEnergyCorrections#JetCorUncertainties

Nuisance Parameters: JEC

Year2016APV

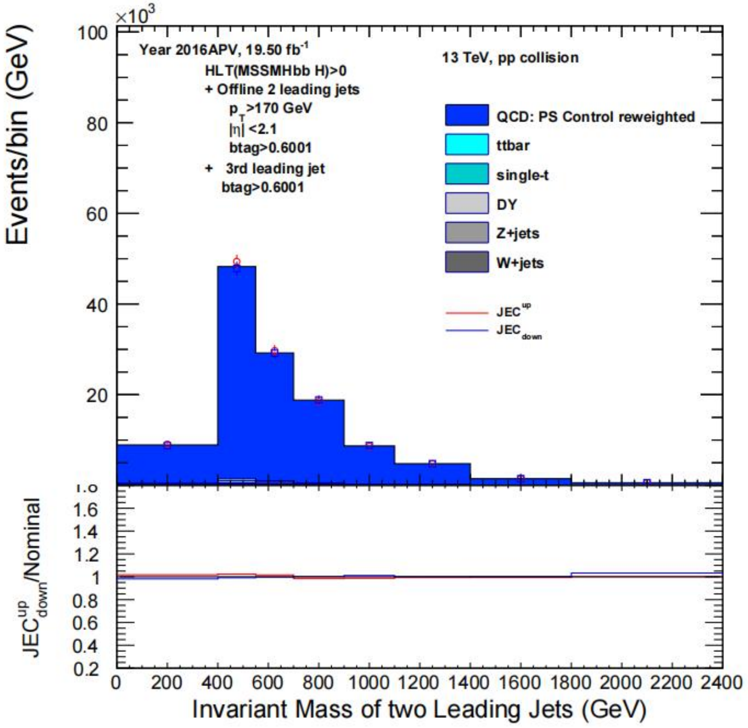


Year2016

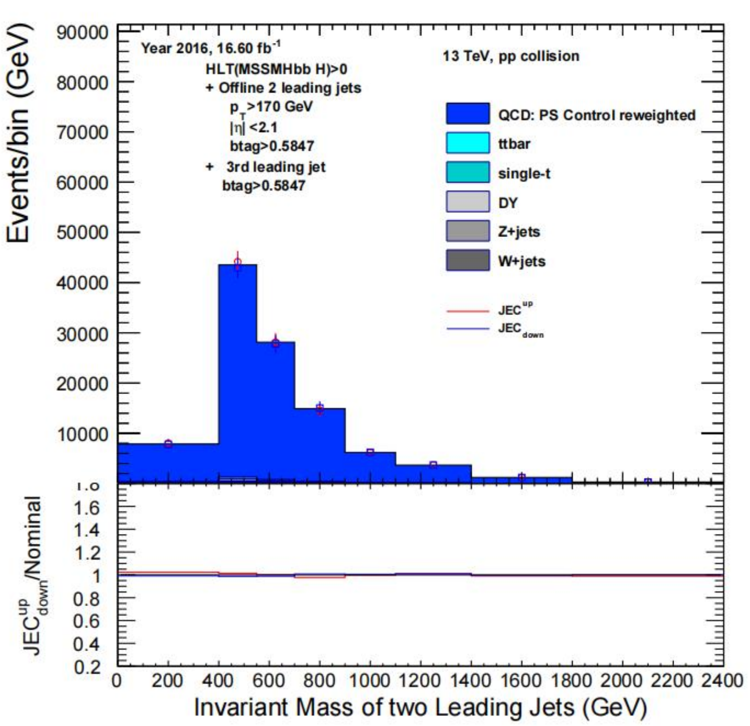


Nuisance Parameters: JEC

Year2016APV



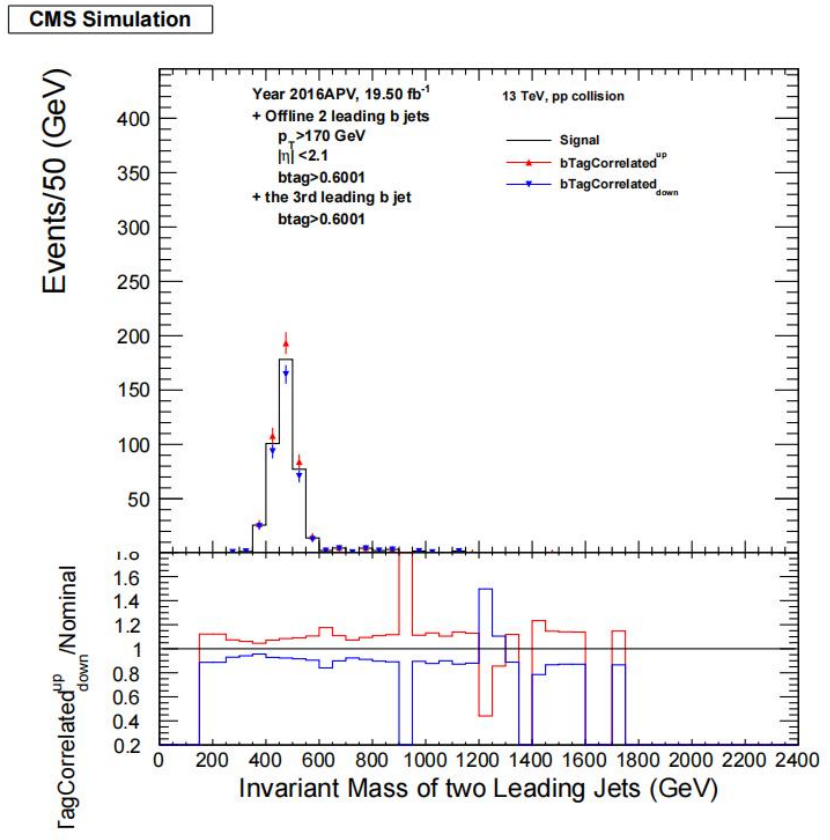
Year2016



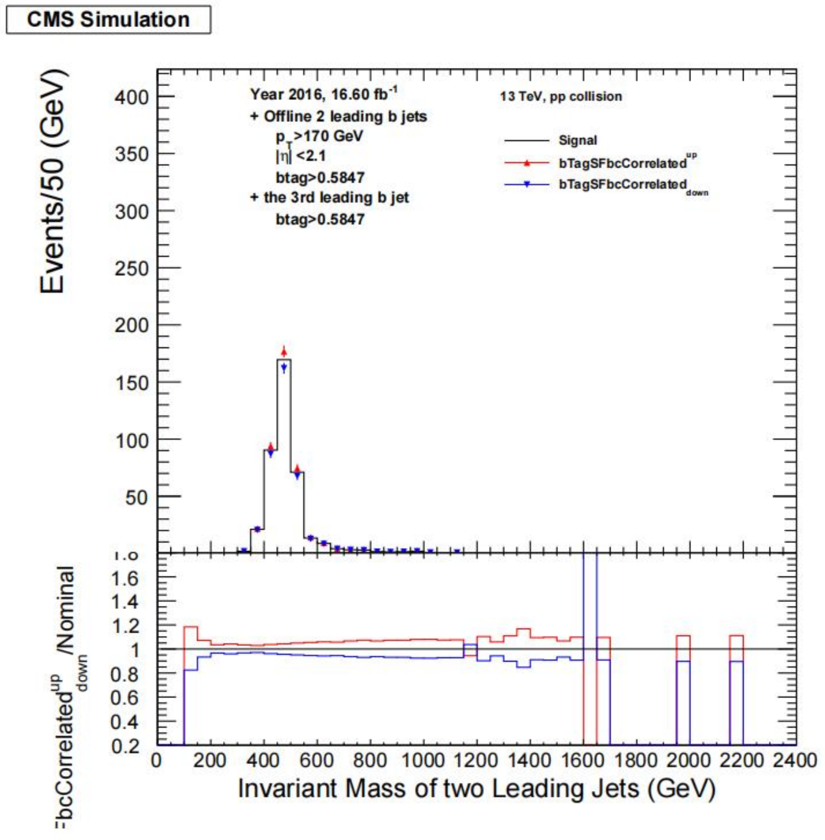
- QCD Reweighting function derived with JEC variance.
- top-related,W+jets, DY and Z+jets from simulation.

Nuisance Parameters: btag SFbc Correlated

Year2016APV

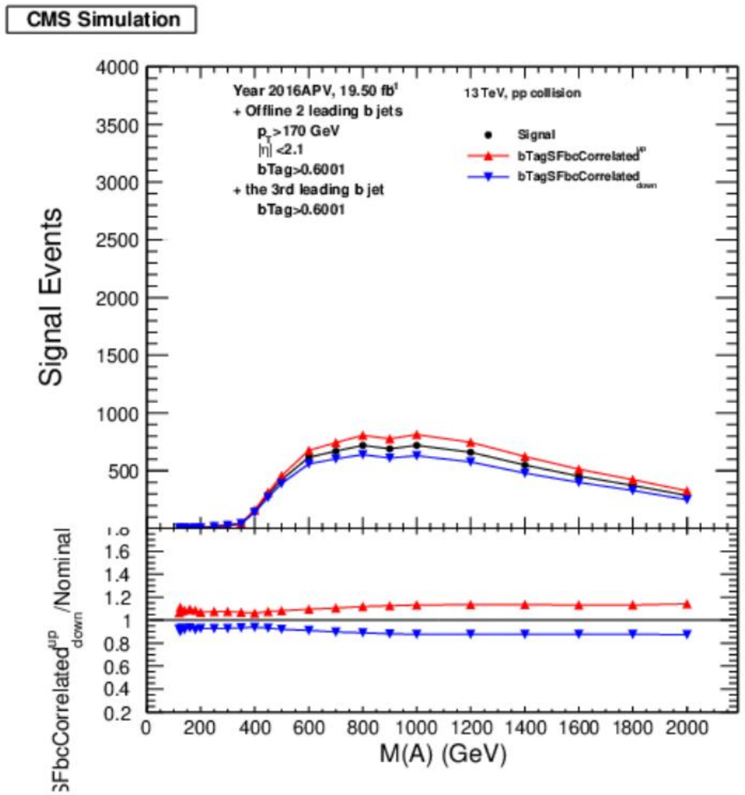


Year2016

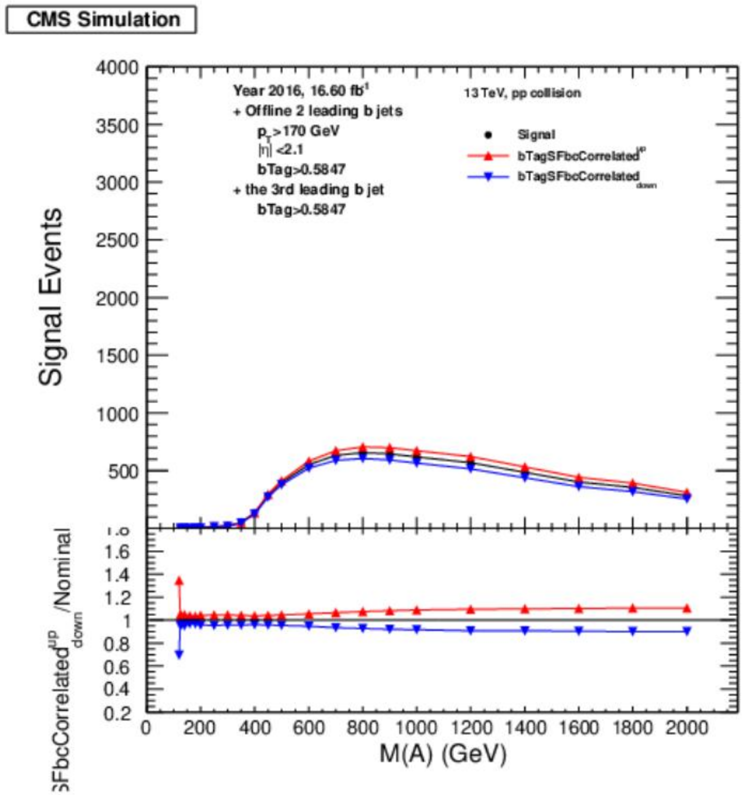


Nuisance Parameters: btag SFbc Correlated

Year2016APV

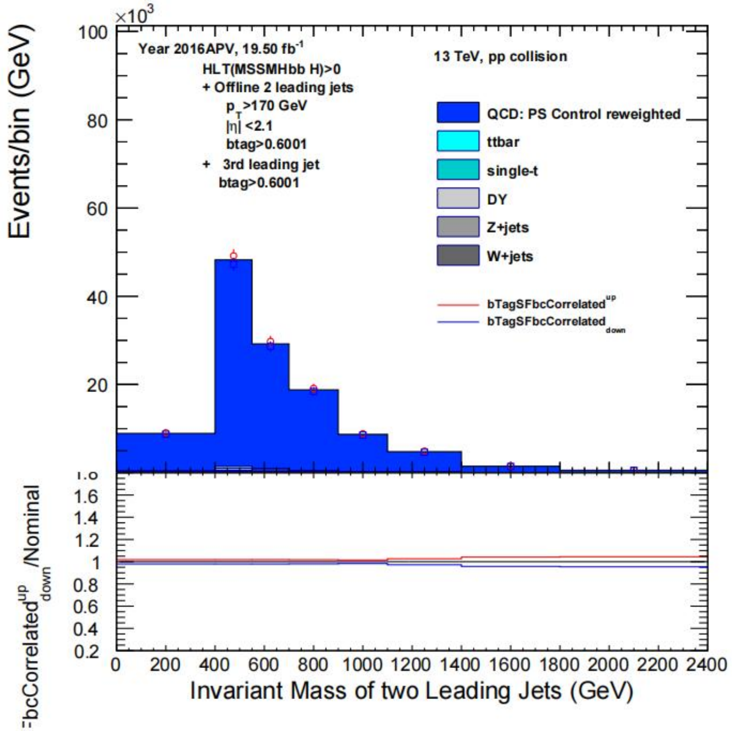


Year2016

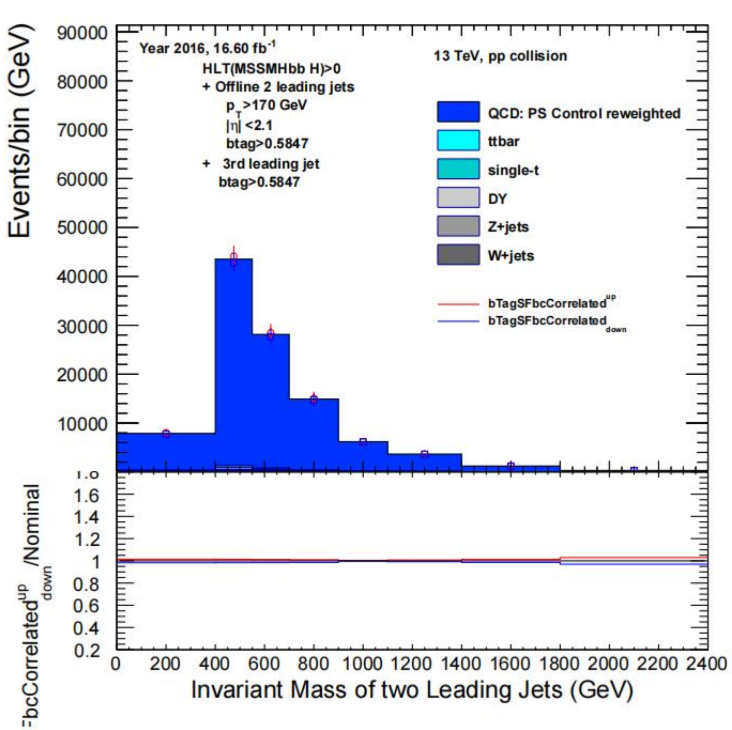


Nuisance Parameters: btag SFbc Correlated

Year2016APV



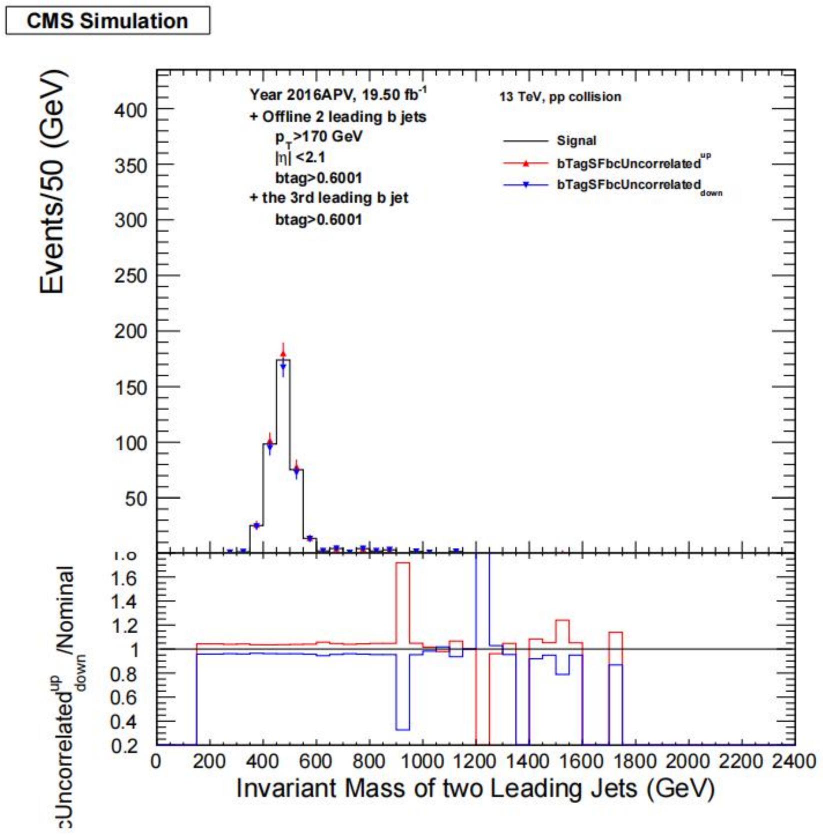
Year2016



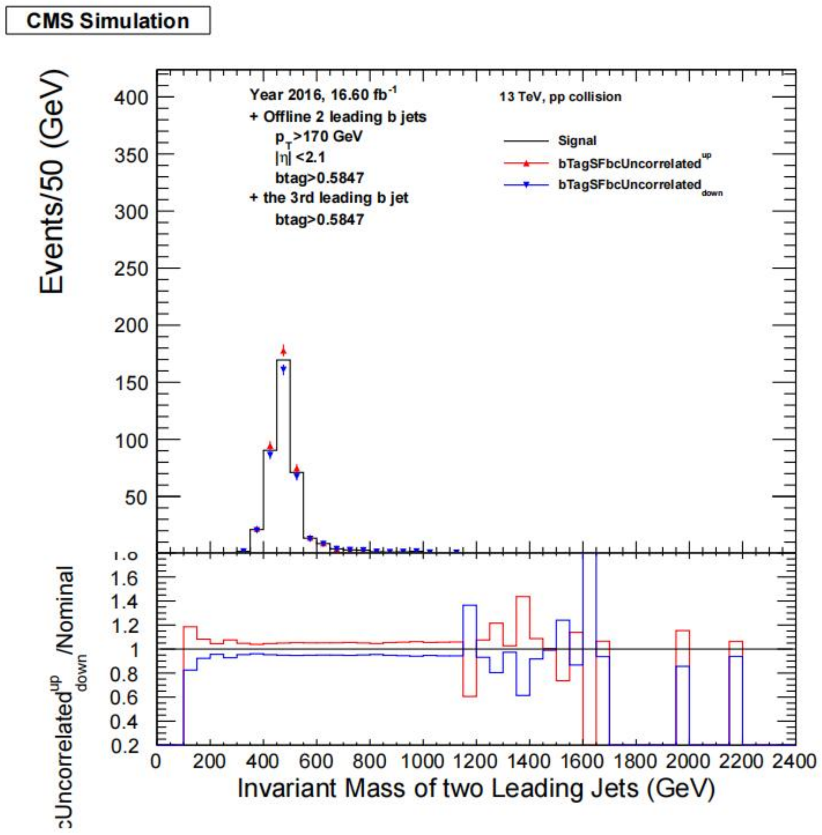
- QCD Reweighting function derived with bTagHf variance.
- top-related,W+jets, DY and Z+jets from simulation.

Nuisance Parameters: btag SFbc Uncorrelated

Year2016APV

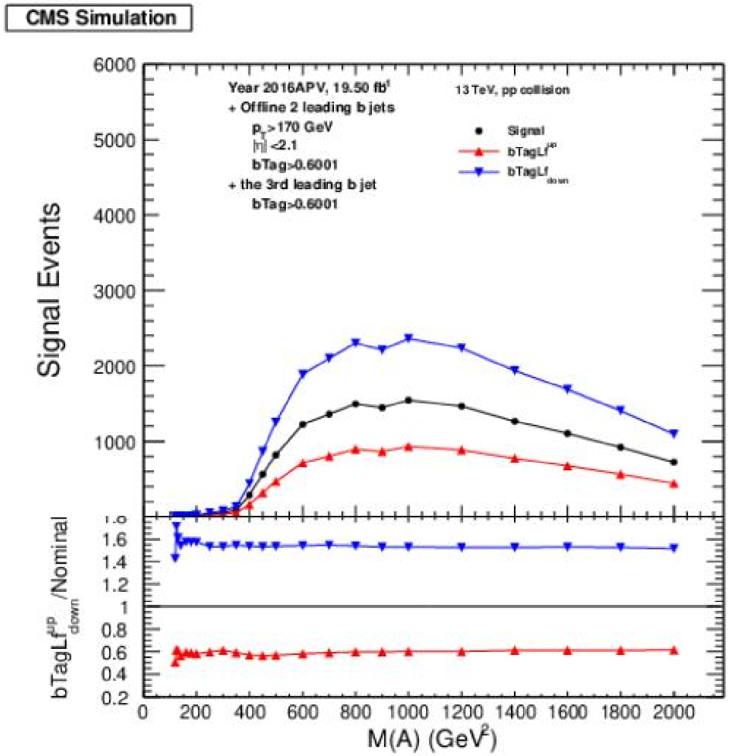


Year2016

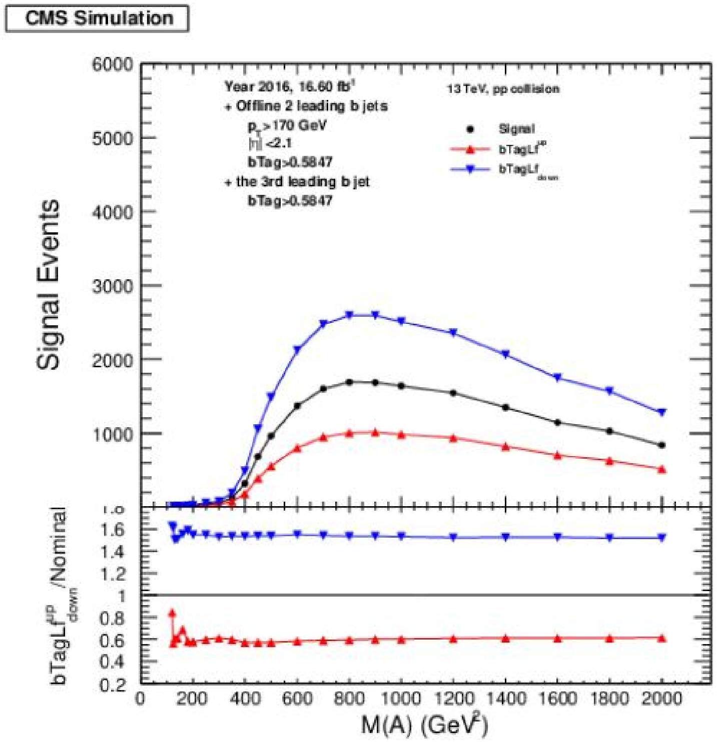


Nuisance Parameters: btag SFbc Uncorrelated

Year2016APV

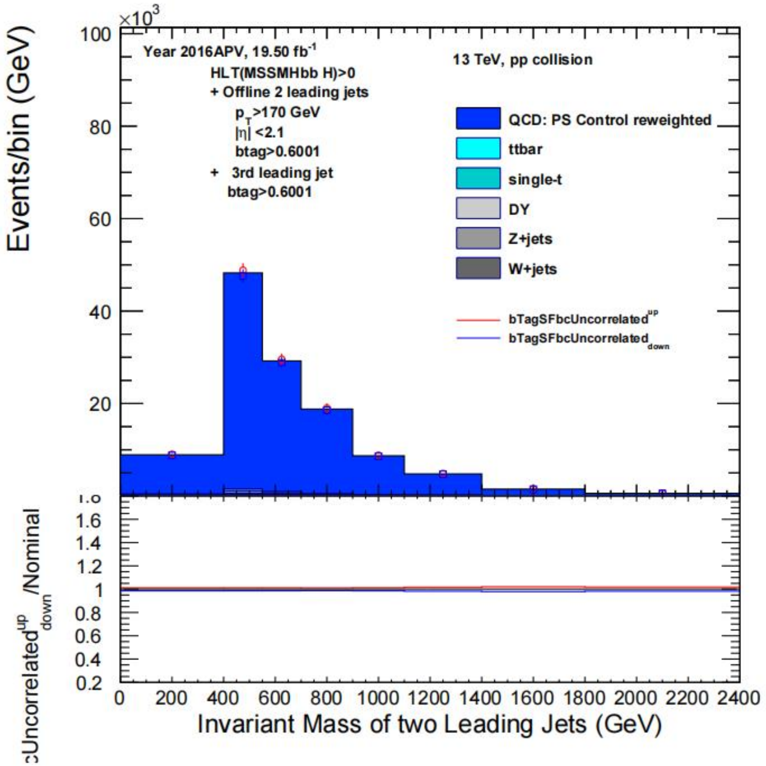


Year2016

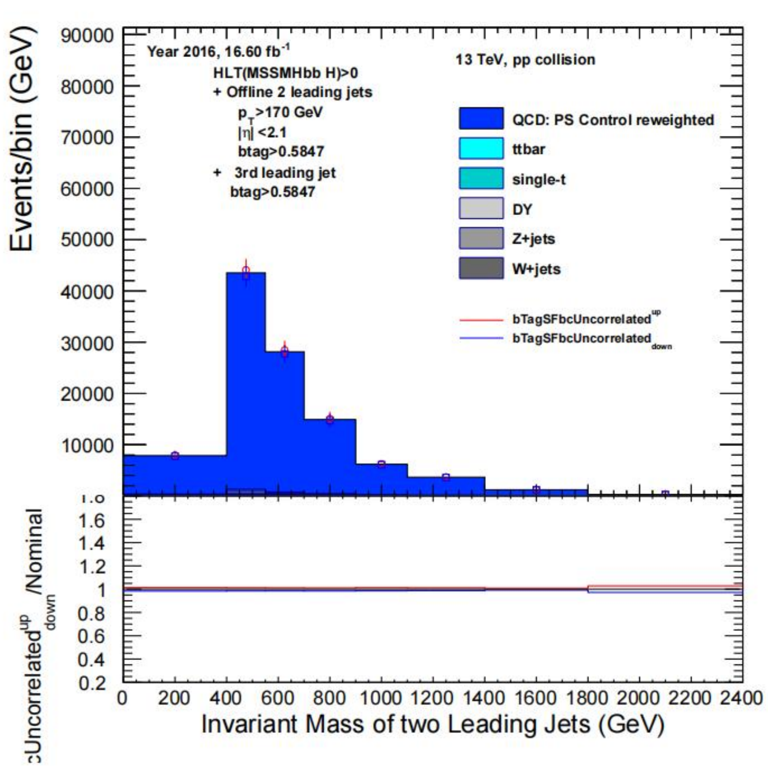


Nuisance Parameters: btag SFbc Uncorrelated

Year2016APV



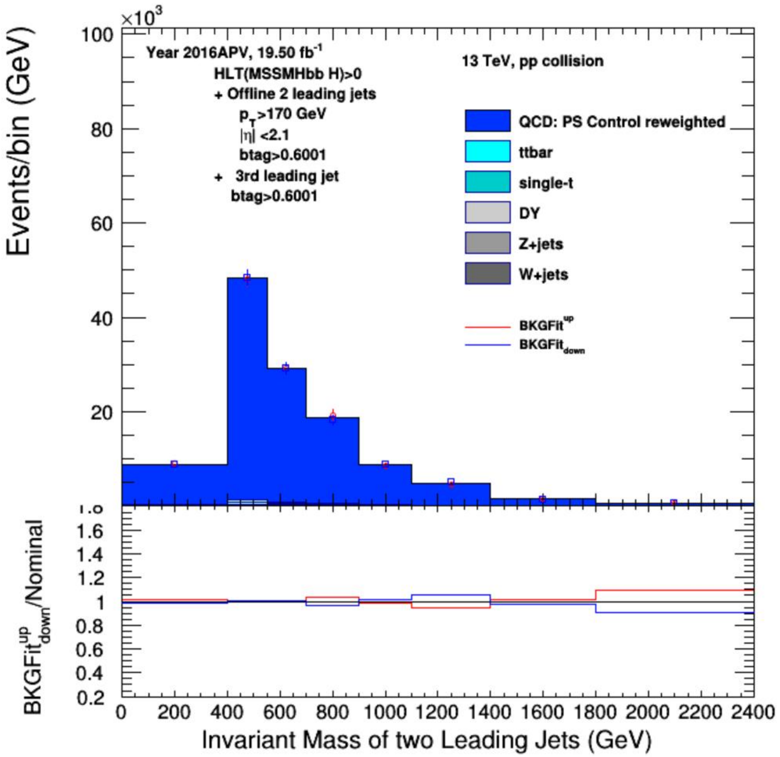
Year2016



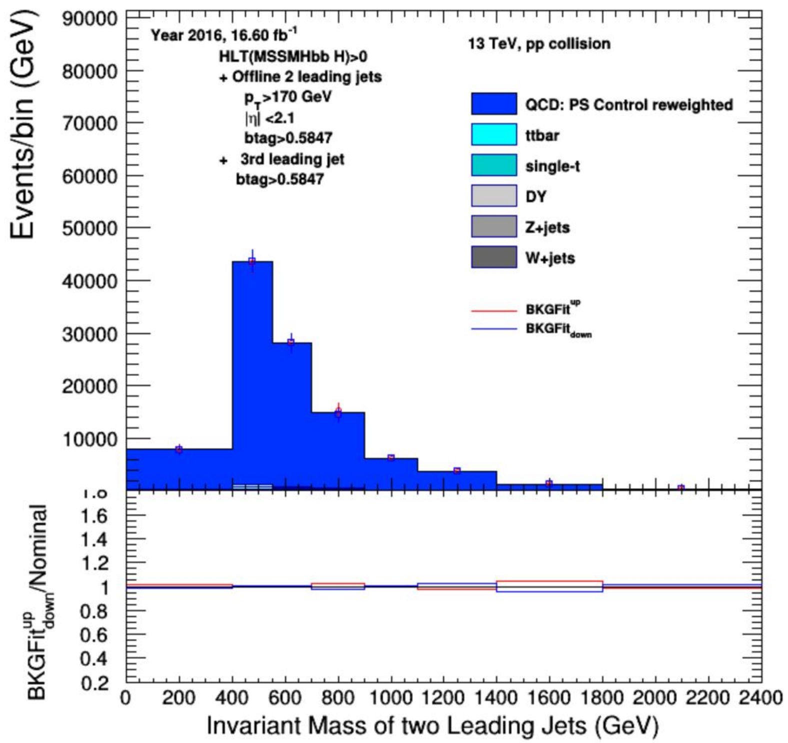
- ❑ QCD Reweighting function derived with bTagLf variance.
- ❑ top-related,W+jets, DY and Z+jets from simulation.

Nuisance Parameters: BKGFit

Year2016APV

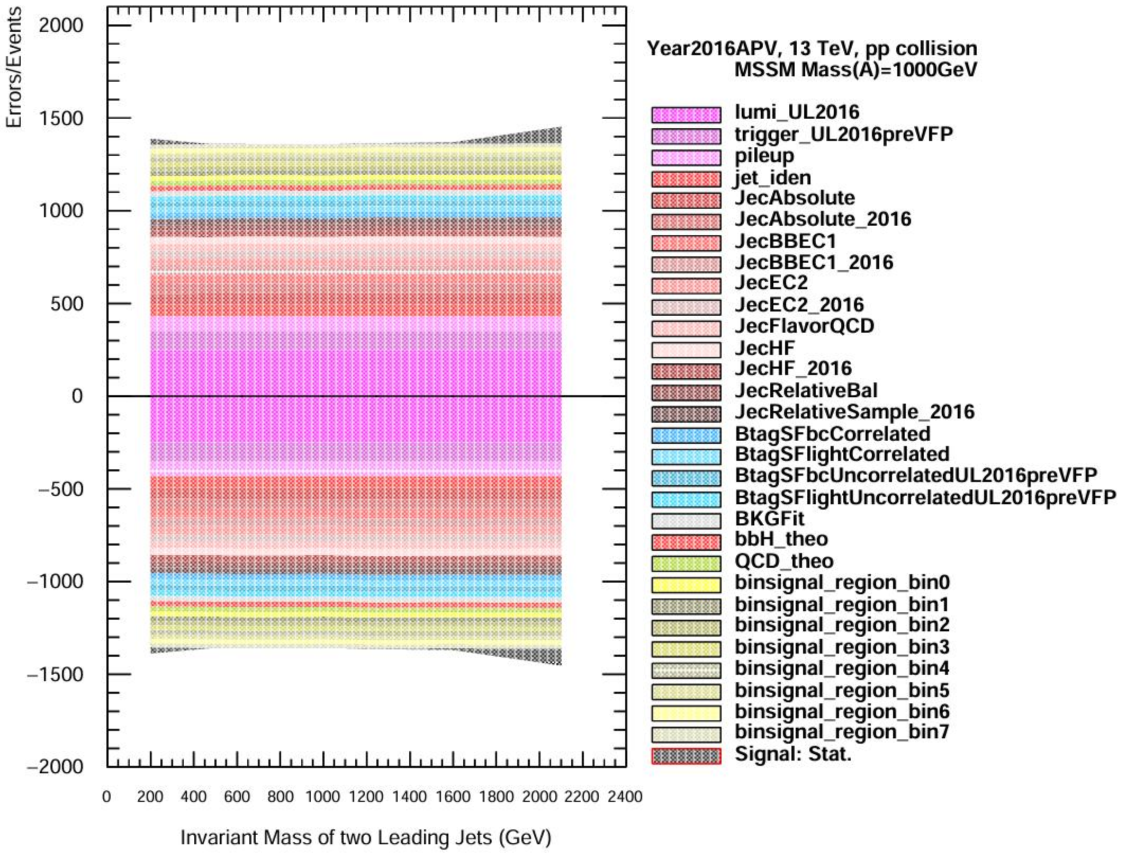


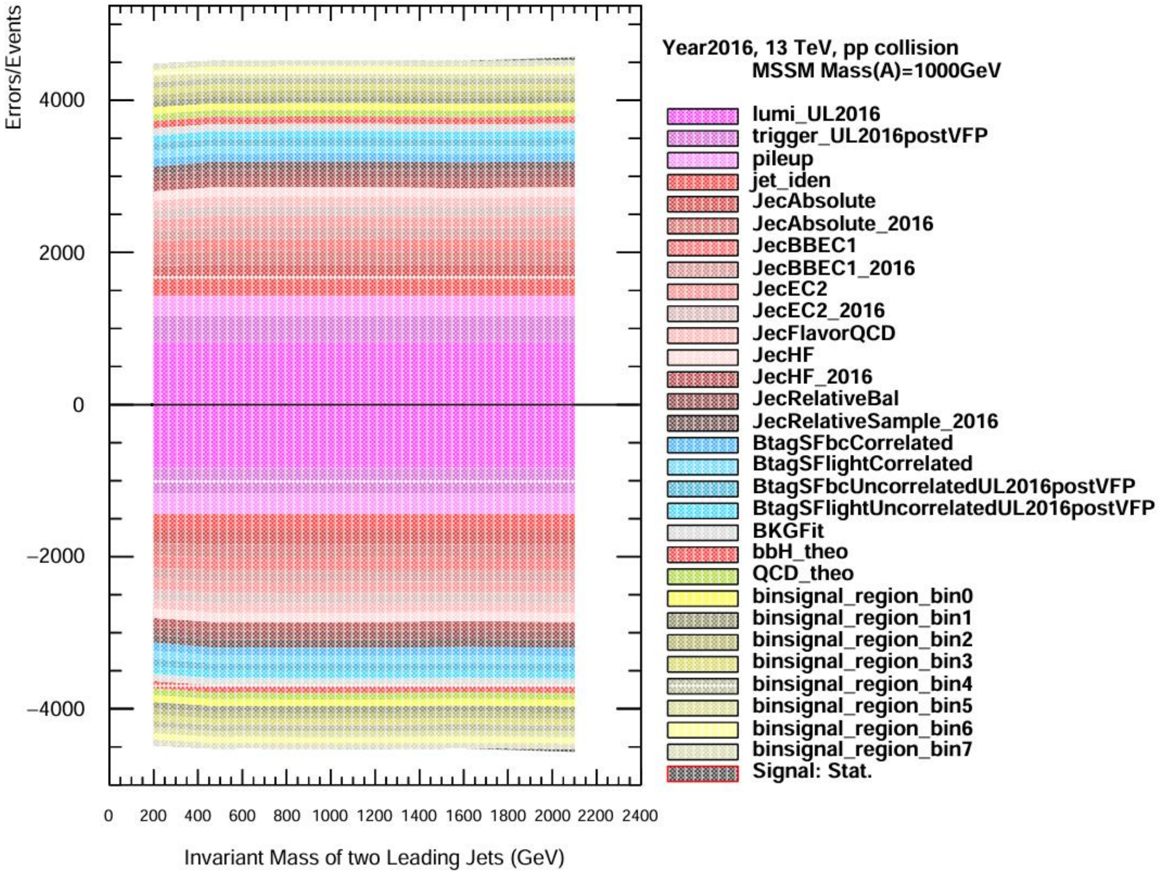
Year2016



- QCD Reweighting functions derived with retriving fitting errors.
- top-related, W+jets, DYand Z+jets from simulation.

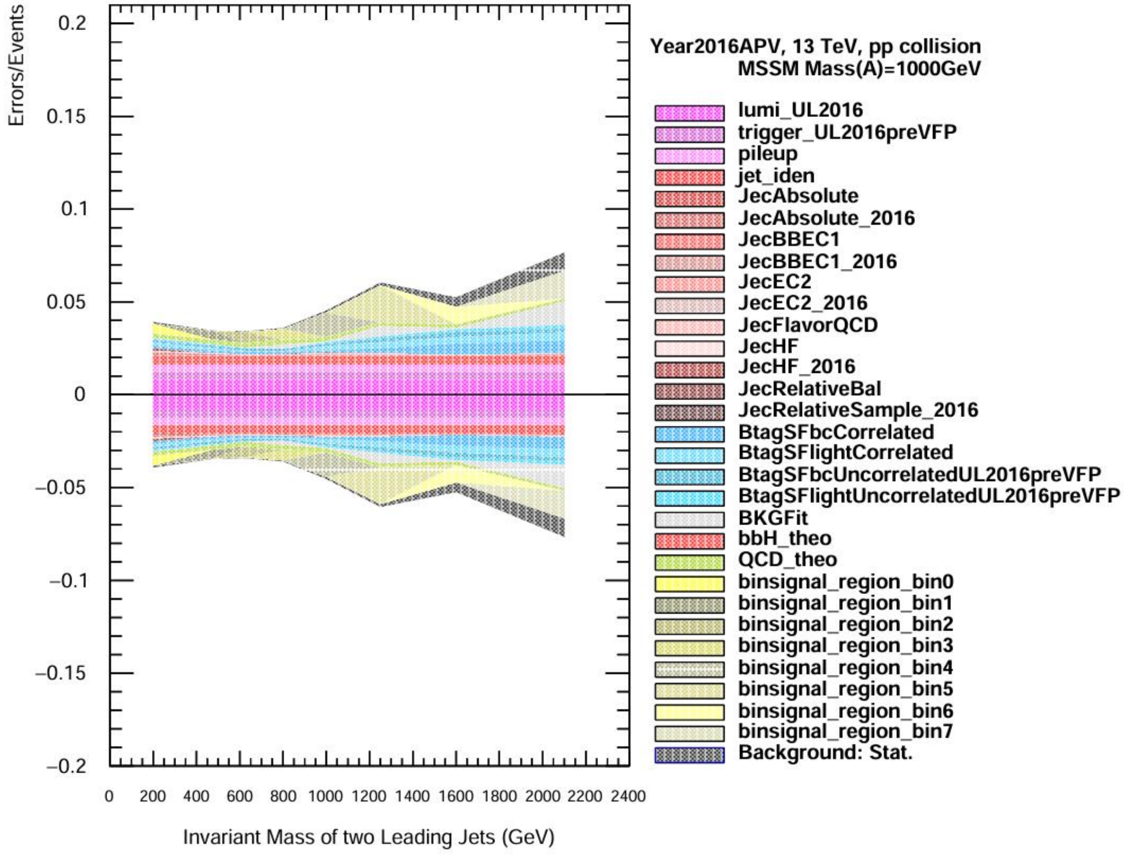
Postfit Distributions with Signal injected Asimov: Signal Uncertainties

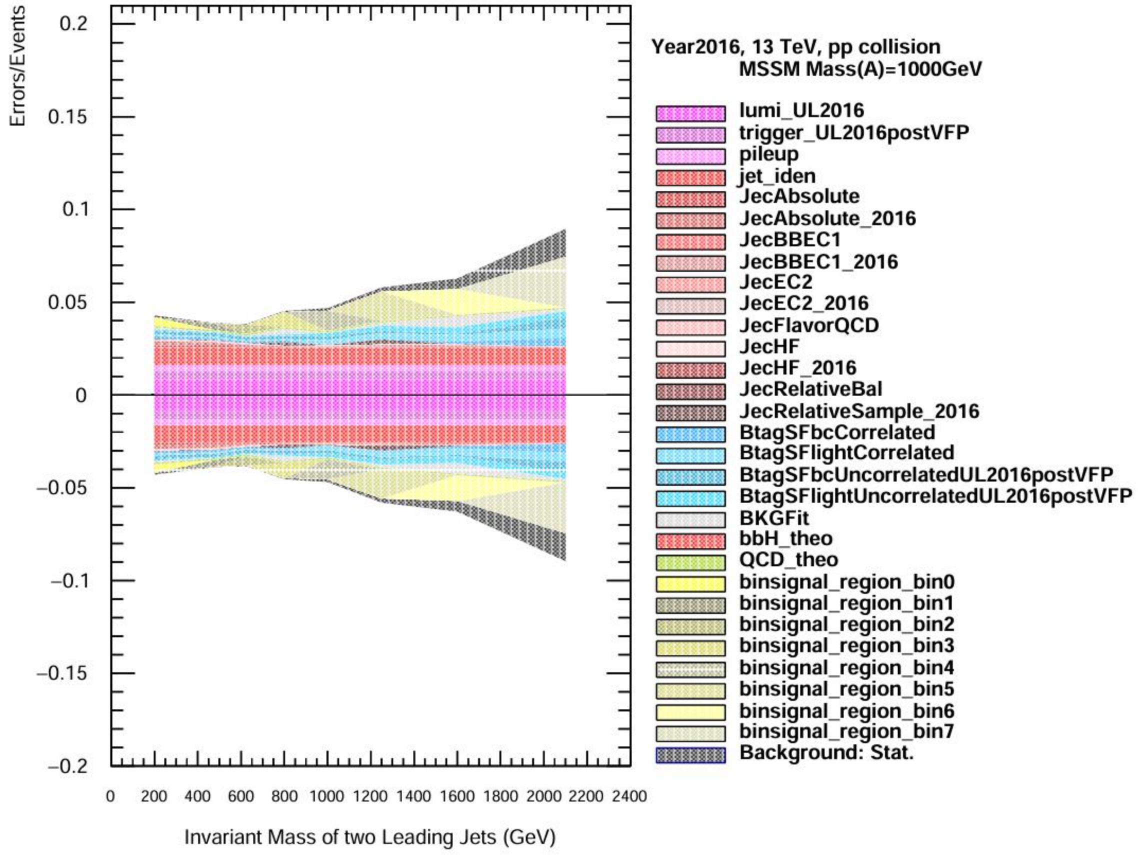




PostFitShapesFromWorkspace, withoption '–freeze'

Postfit Distributions with Signal injected Asimov: Background Uncertainties

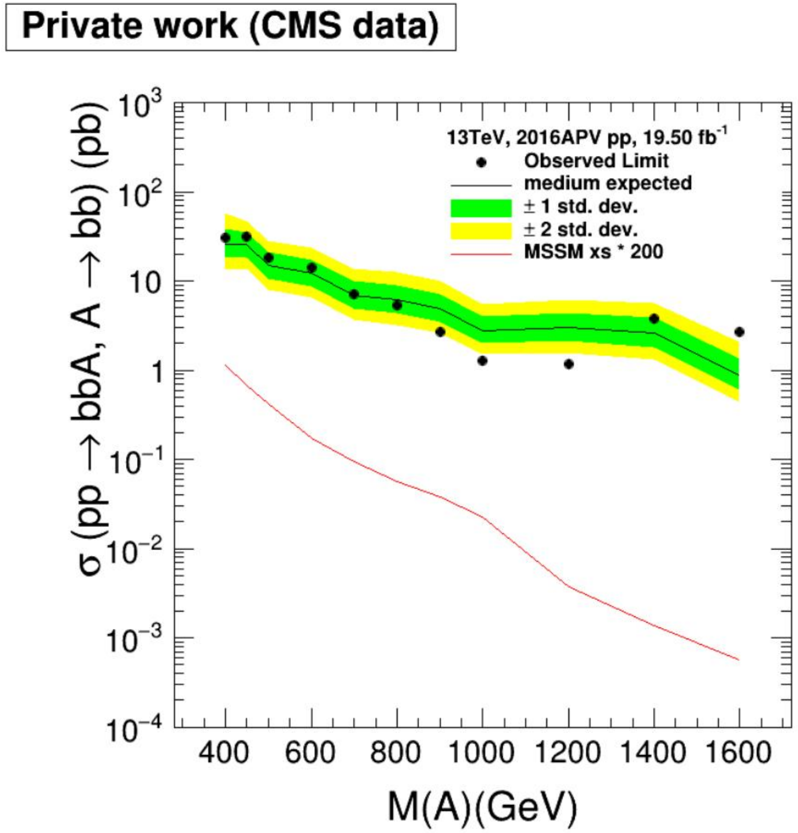




PostFitShapesFromWorkspace, withoption '–freeze'

Upper Limits

Year2016APV



Year2016

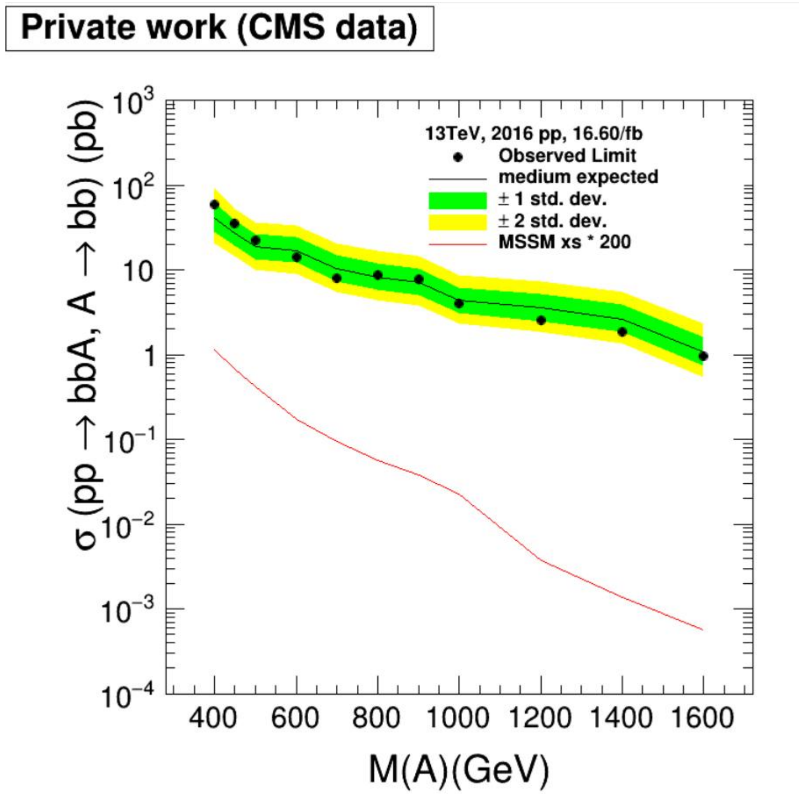
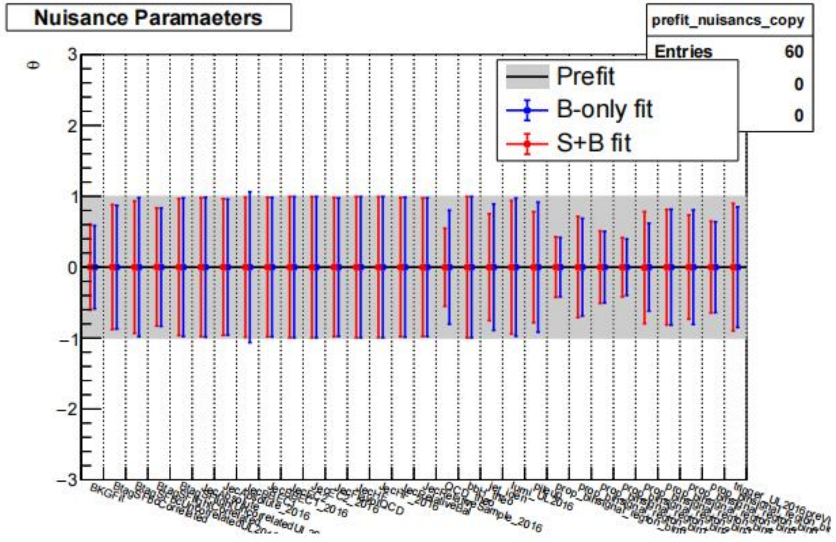


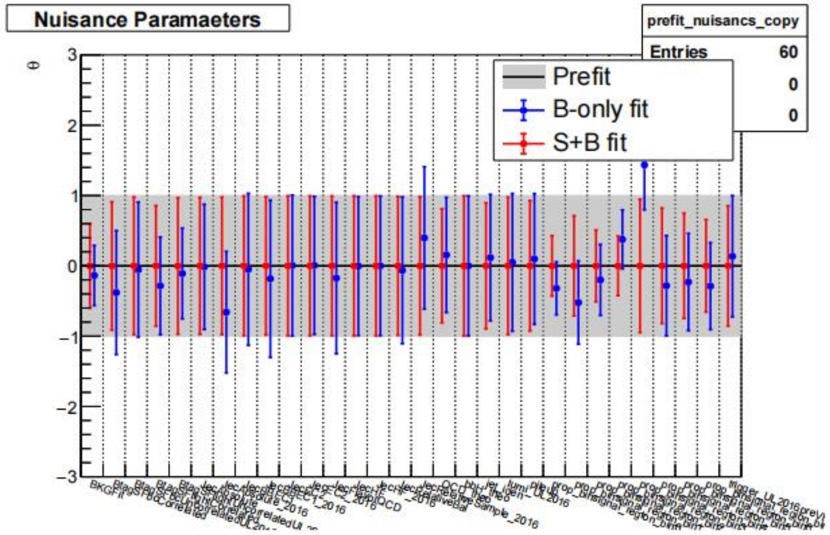
Figure:Year2016APV and Year2016 upper limits of production cross section of MSSM Higgs.

An example of Closure Test(Year2016APV)

Background Only Asimov Toy



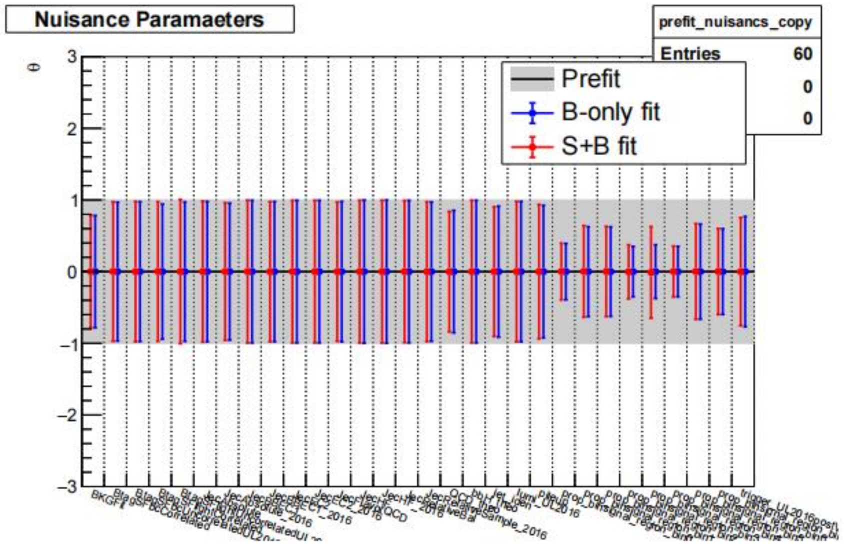
Signal+Background Asimov Toy



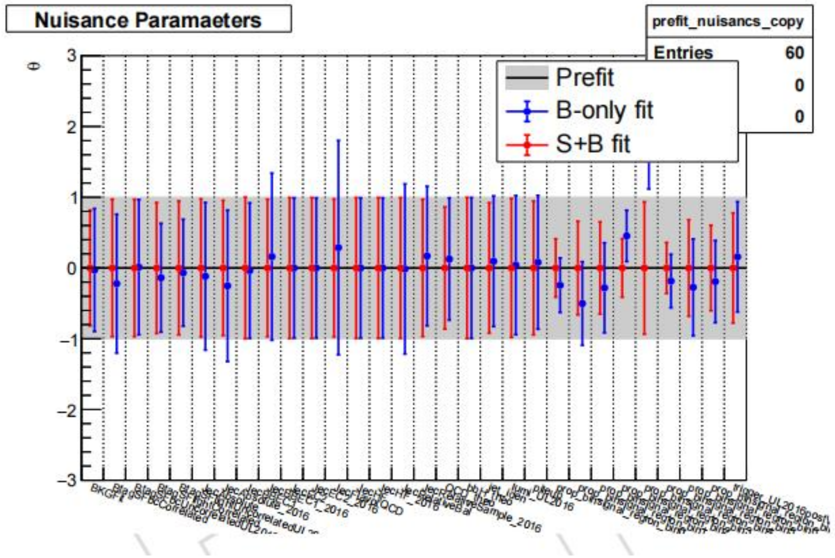
- best fit of background only Asimov toy agrees with inputs, r(cross-section): 0.0190051 -0.0190051/+1.14692
- best fit of signal+background Asimov toy agrees with inputs, r(cross-section): 2.80881 -2.80881/+1.22086

An example of Closure Test(Year2016)

Background Only Asimov Toy



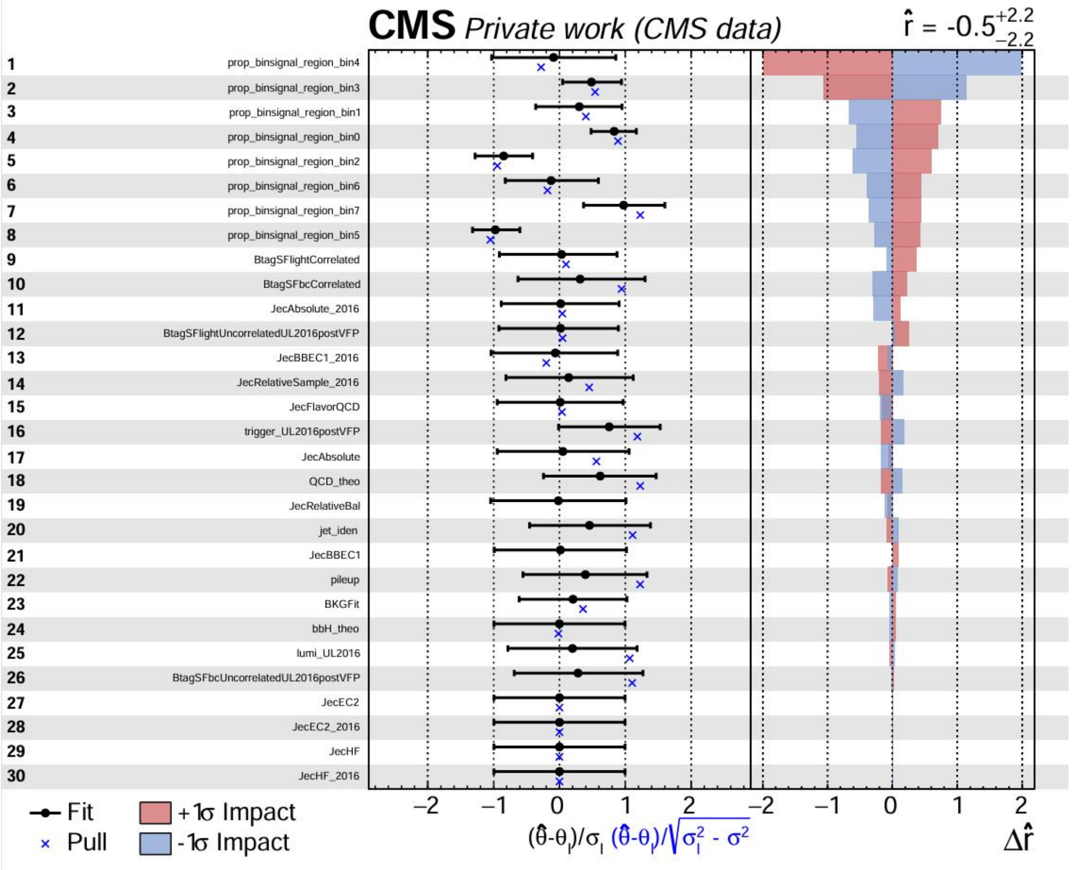
Signal+Background Asimov Toy



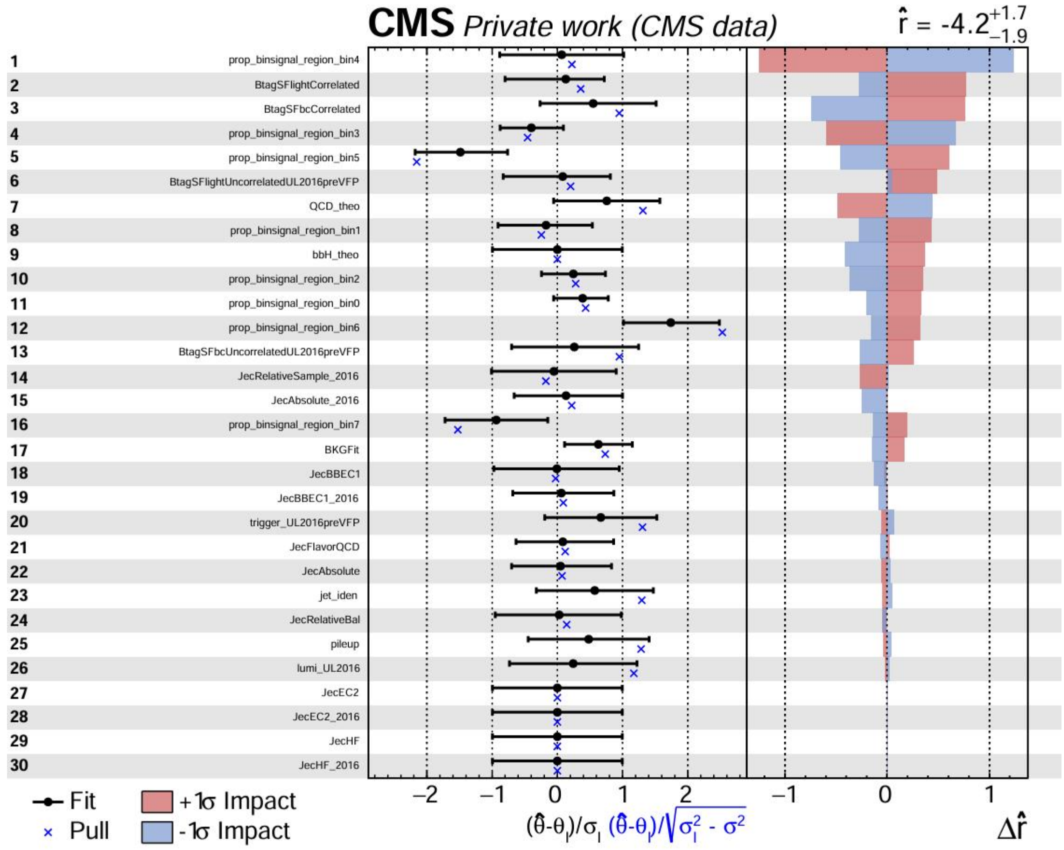
- ❑ best fit of background only Asimov toy agrees with inputs, r(cross-section): 0.0190051 -0.0190051/+1.14692
- ❑ best fit of signal+background Asimov toy agrees with inputs, r(cross-section): 2.80881 -2.80881/+1.22086

Impacts of Nuisance Parameters

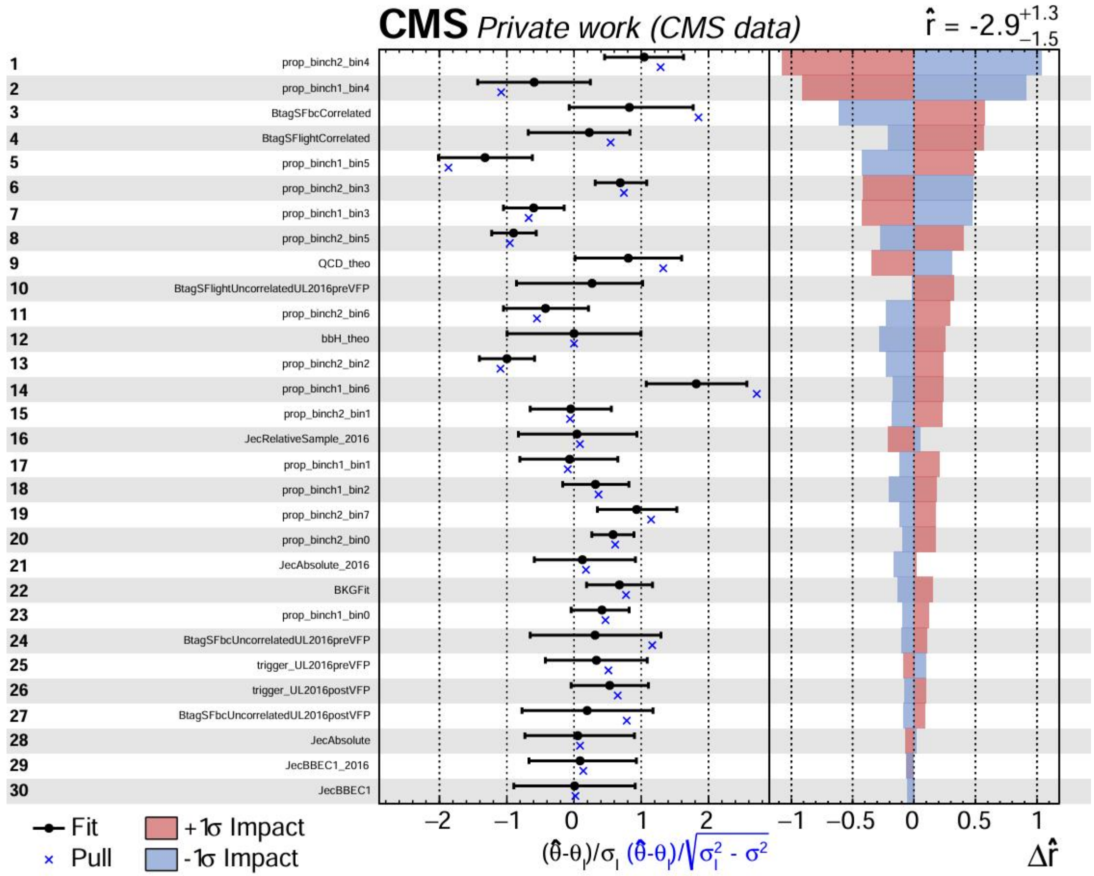
Year2016

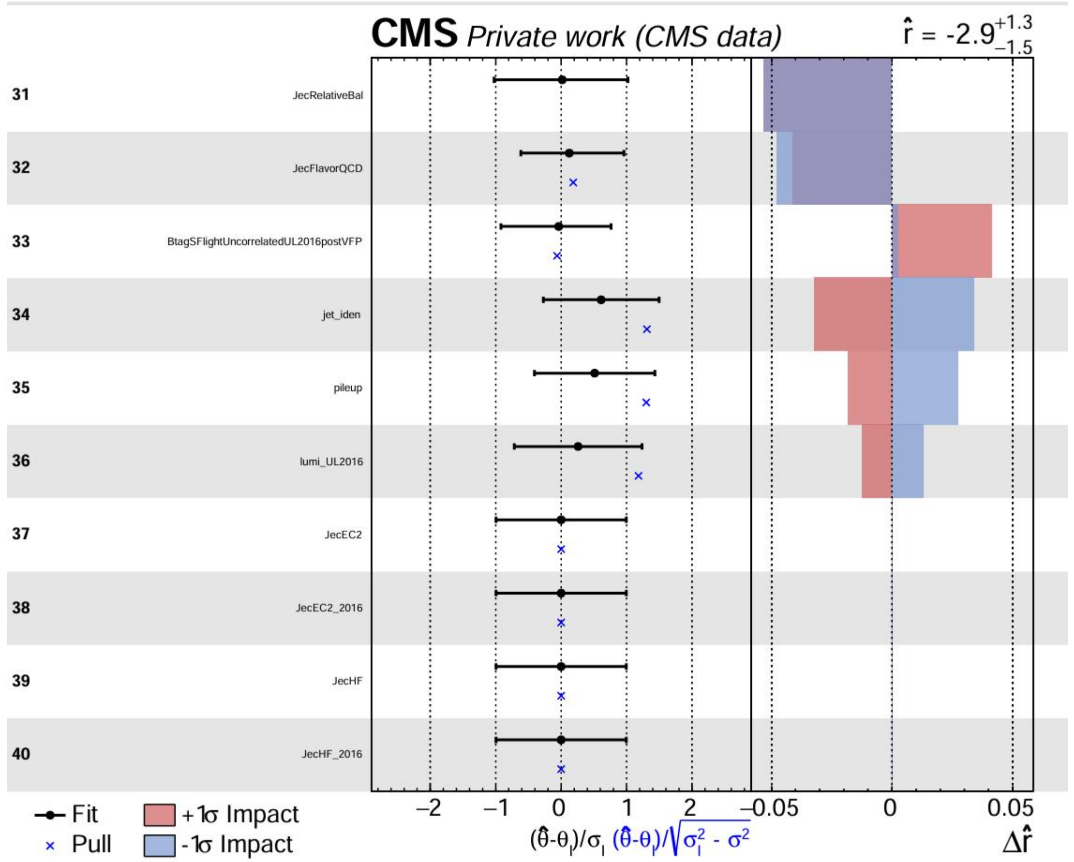


Year2016APV



Year2016 Combined





Impacts produced by signal injected Asimov datasets

2016 MSSM Hbb Ultra-Legacy signal MC samples as inputs