Study of the Vector Boson Fusion of the Z boson and MTD DAQ developments

20/09/2024

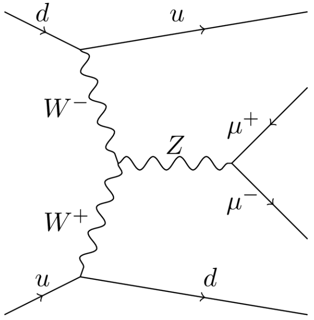
Giorgio Pizzati[a]

a: Università e INFN di Milano Bicocca

Second year of Ph.D. outline

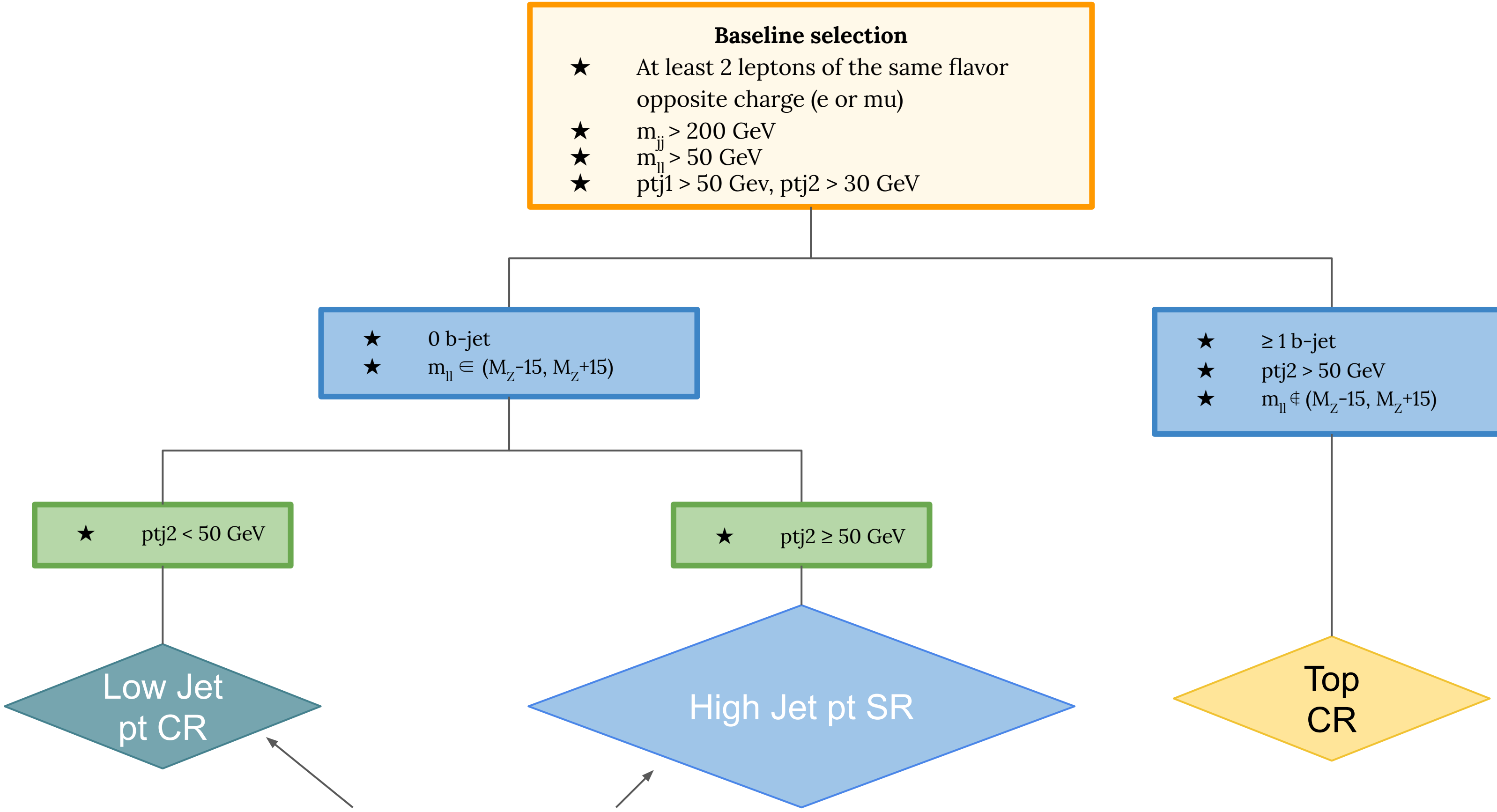
- Finishing the analysis for VBF Z + jets for SM with Full Run II Data
  - Gained a much more detailed understanding of the analysis
  - Left the old latinos post-processing → everything is run on the fly
  - Can easily add or substitute new samples
  - Had to check and fix all the processing chain
  - Full Run 2 is finally ready
- MTD DAQ developments in Bicocca and @ TIF
  - Started around February, still have to learn a lot
  - Many different areas of development are going on in parallel

Vector Boson Fusion



- VBF Topology can be identified by two jets with high invariant mass and high separation in pseudorapidity (opposite hemispheres)
- The two leptons originating from the Z are required to have an invariant mass under the Z mass peak
- Interested in Z decay into charged leptons → clean process (no MET)
- Inside the same sample (EW Zjj) one has VBF as other processes

Analysis event selection



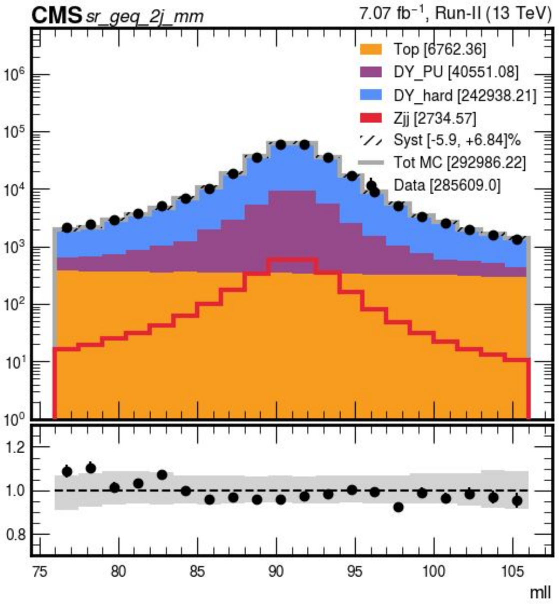
In the next slides DY-PU CR SR Inclusive

JER studies

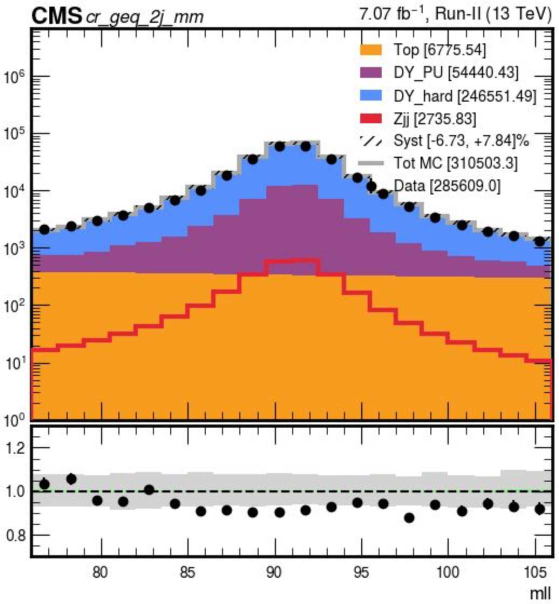
- We've performed some tests on Jet smearing to validate our analysis approach
- We started by not applying any smearing and then applying the official recipe
- Best agreement found by not applying the smearing
- JER uncertainty do not cover the case of “no smearing”
- The discrepancy that we used to observe for the DY normalization has its origin in the Jet smearing
- Latinos Working Group recipe is to not apply JER in the Horns (2.8 < |eta| < 3.0) for low pt jets (pt < 50.0, mostly from PU) -> no JER in Horn and is blessed by JERC
- Next plots are in DY inclusive region with at least 2 jets

JER comparison

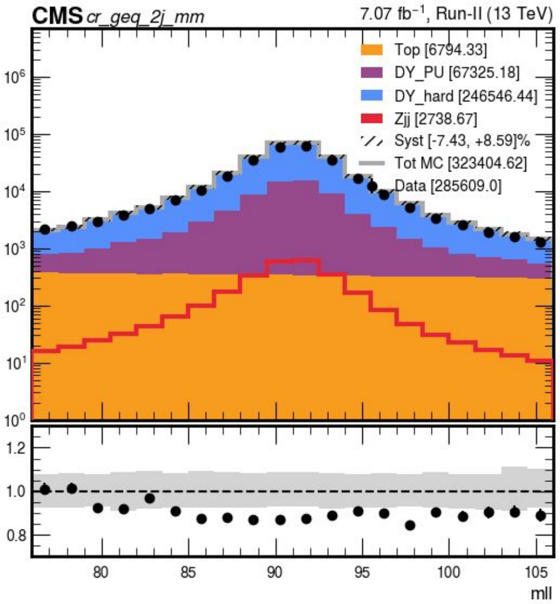
No JER



No JER In Horn

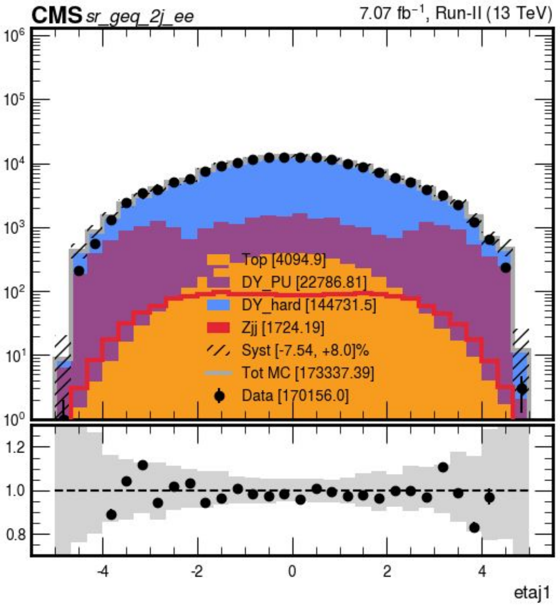


JER

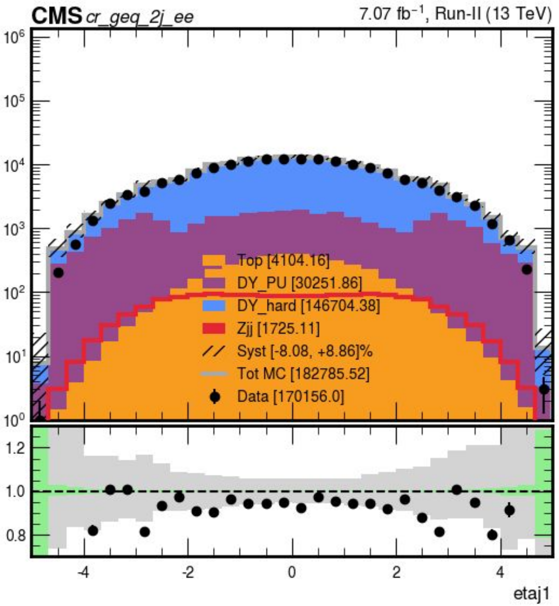


JER comparison

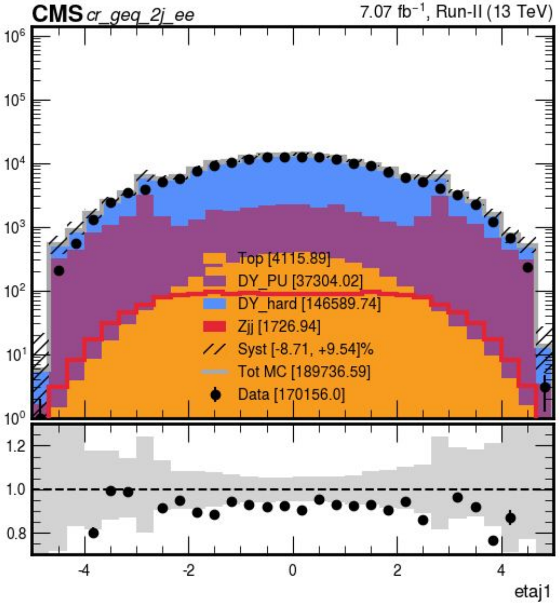
No JER



No JER In Horn



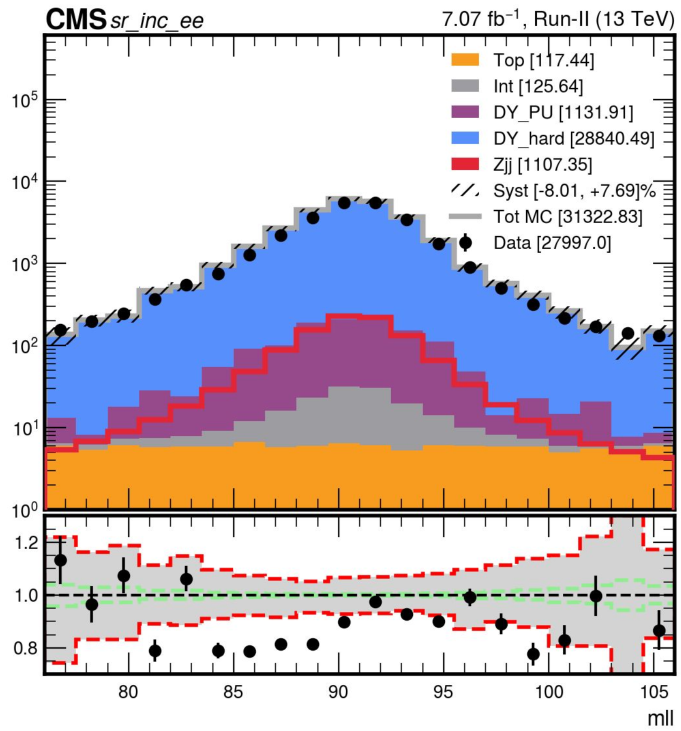
JER

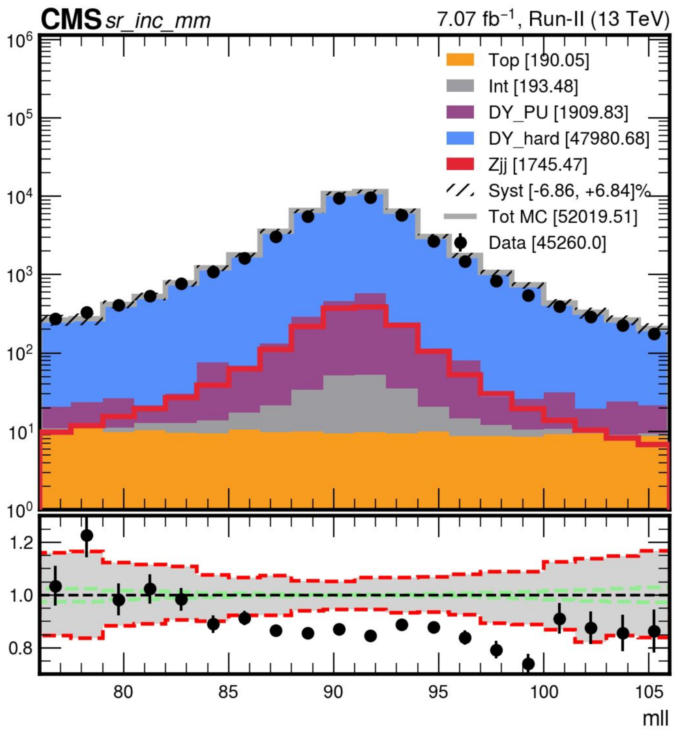


No JER in Horn in the Analysis Phase Space

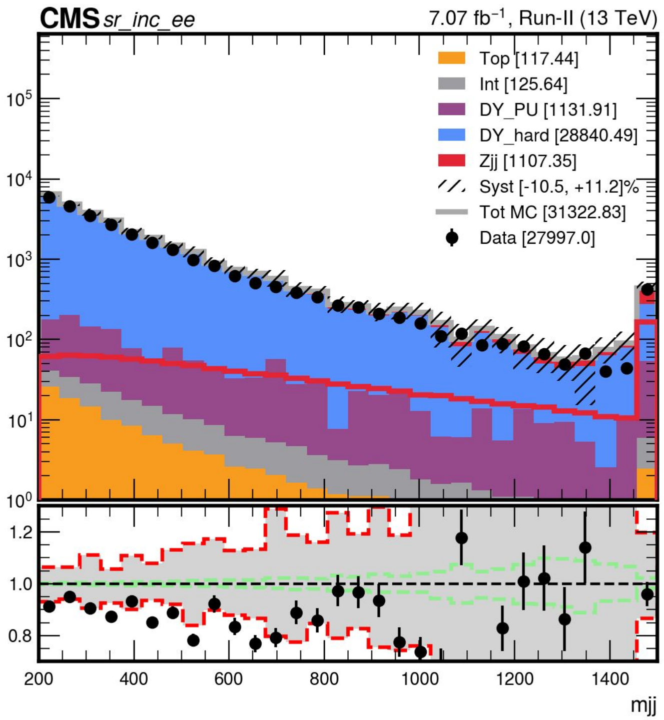
- The blessed approach of No JER in Horn seemed to have a reasonable agreement in an inclusive 2 jets DY phase space
- When going into the analysis phase space (mjj > 200, high pt jets) the impact of smearing becomes once again important leading to a ~ 20% disagreement
- We already know from previous studies that the disagreement can be taken care of with rate parameters on the DY (both hard and PU) but we finally know its origin

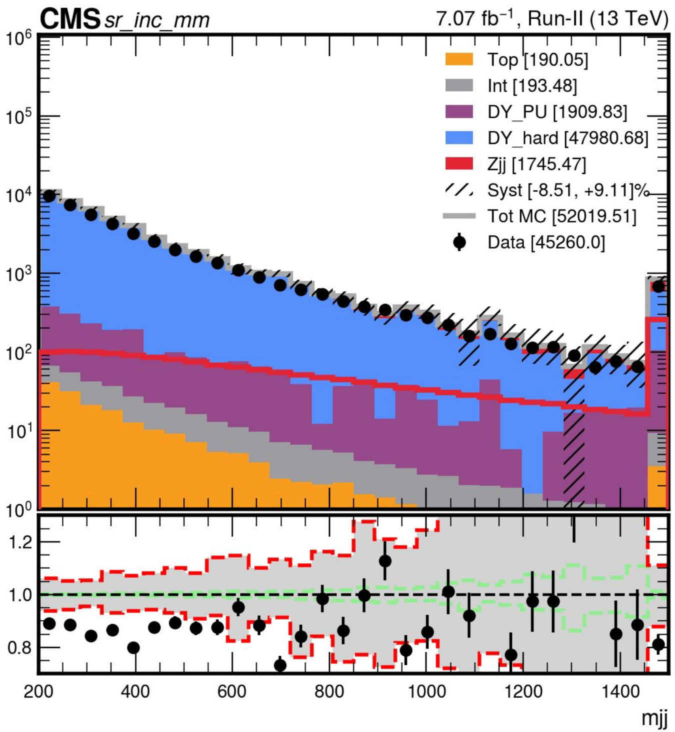
No JER in Horn in the Analysis Phase Space





No JER in Horn in the Analysis Phase Space

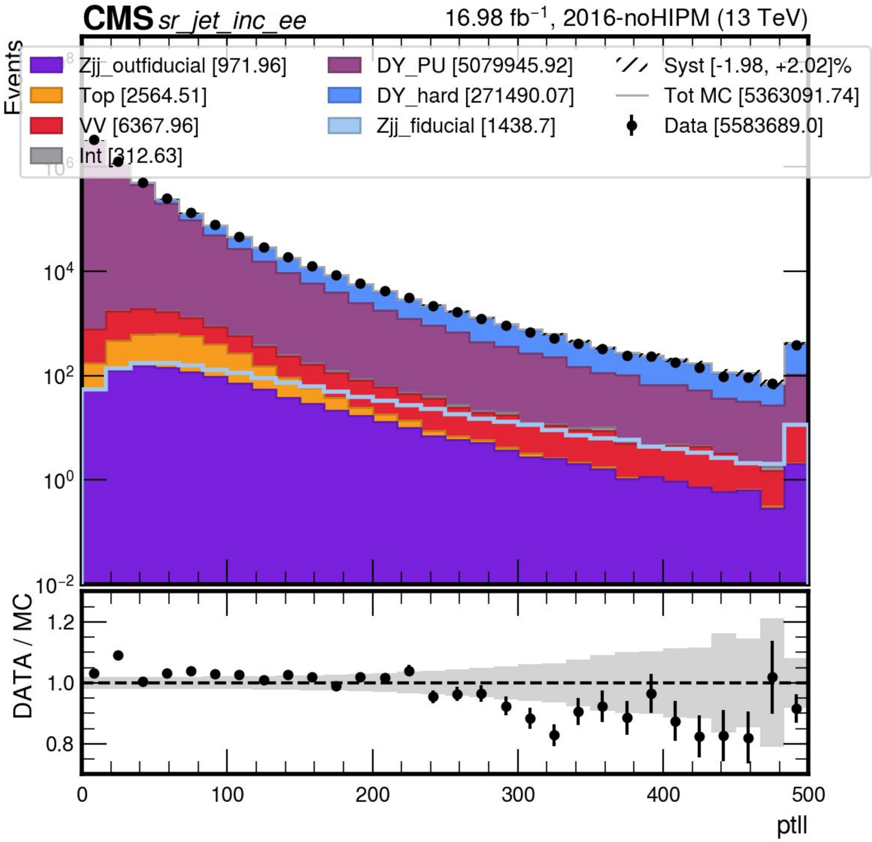


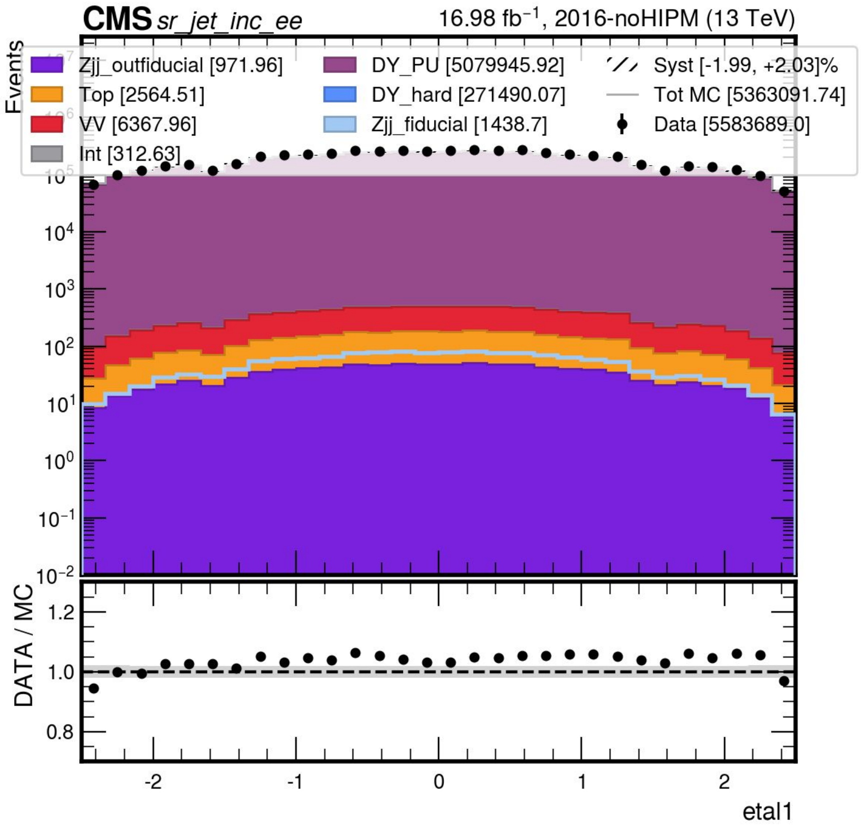


Final JER approach

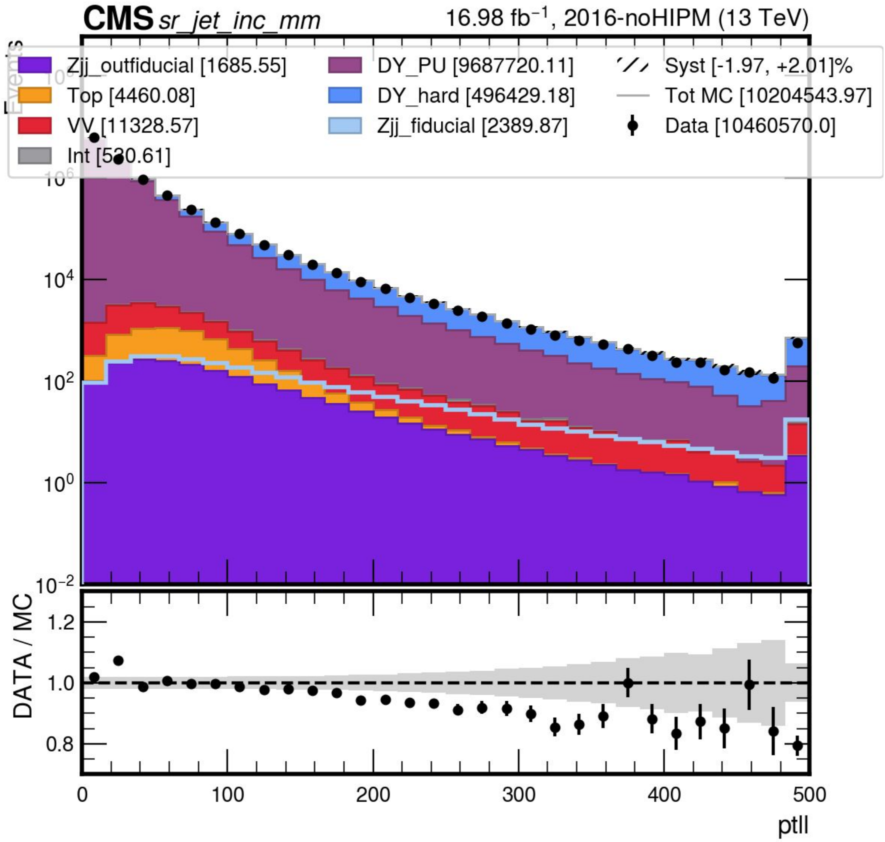
- Given the JER disagreement not only in the horns but also in the forward region the JER is not applied for jets with $\eta>2.5$

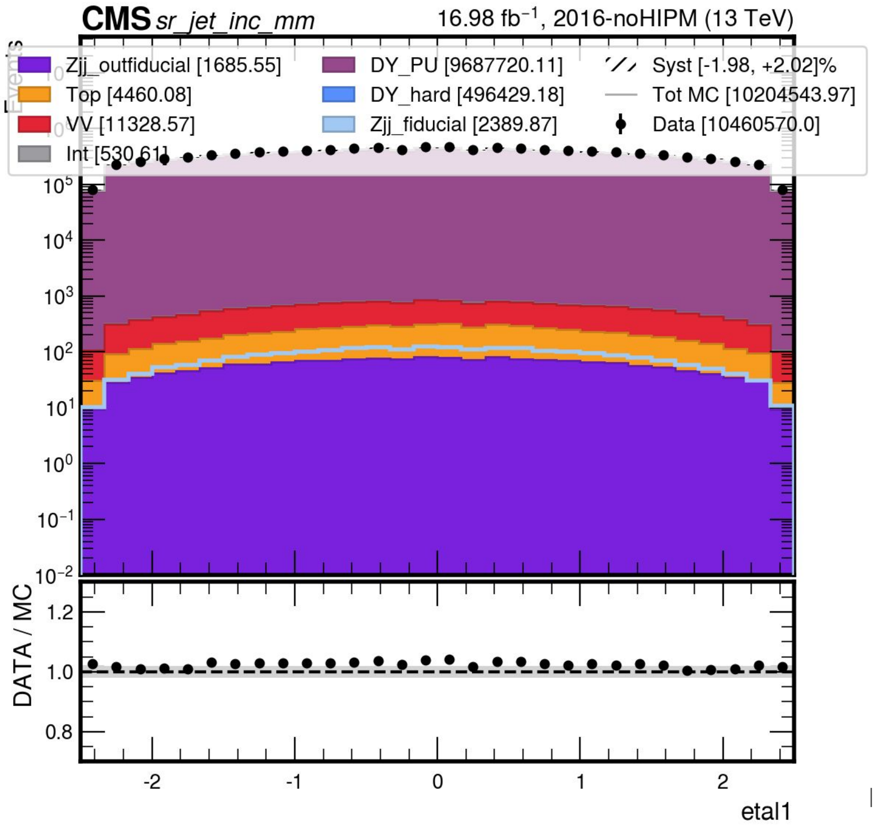
Z+Jets (ee) inclusive (no njet cut)



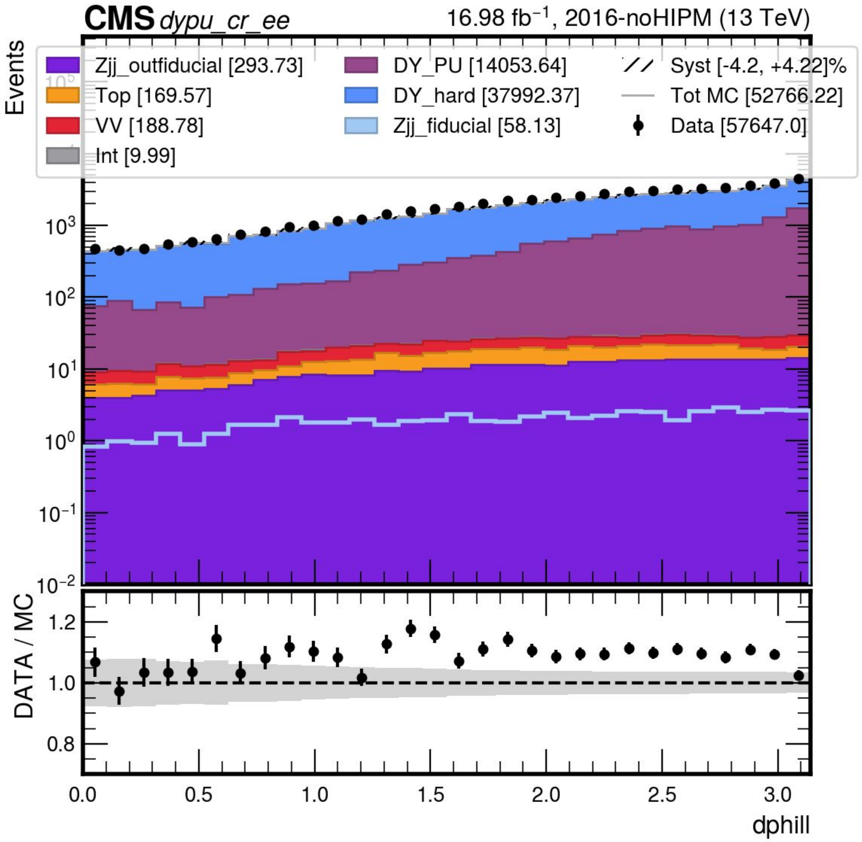


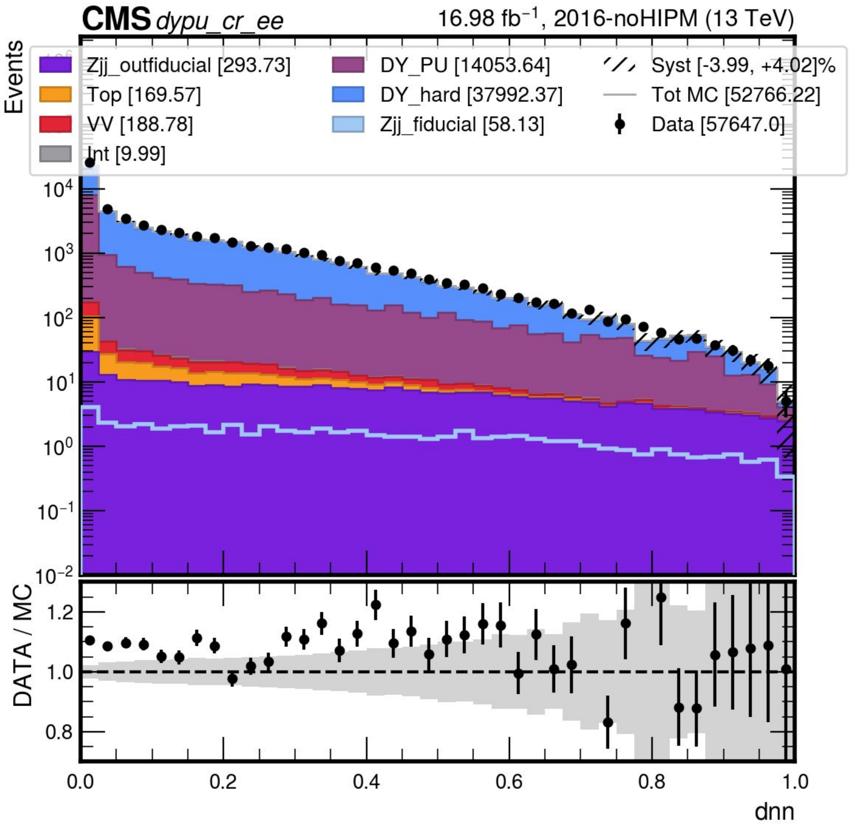
Z+Jets (mm) inclusive (no njet cut)



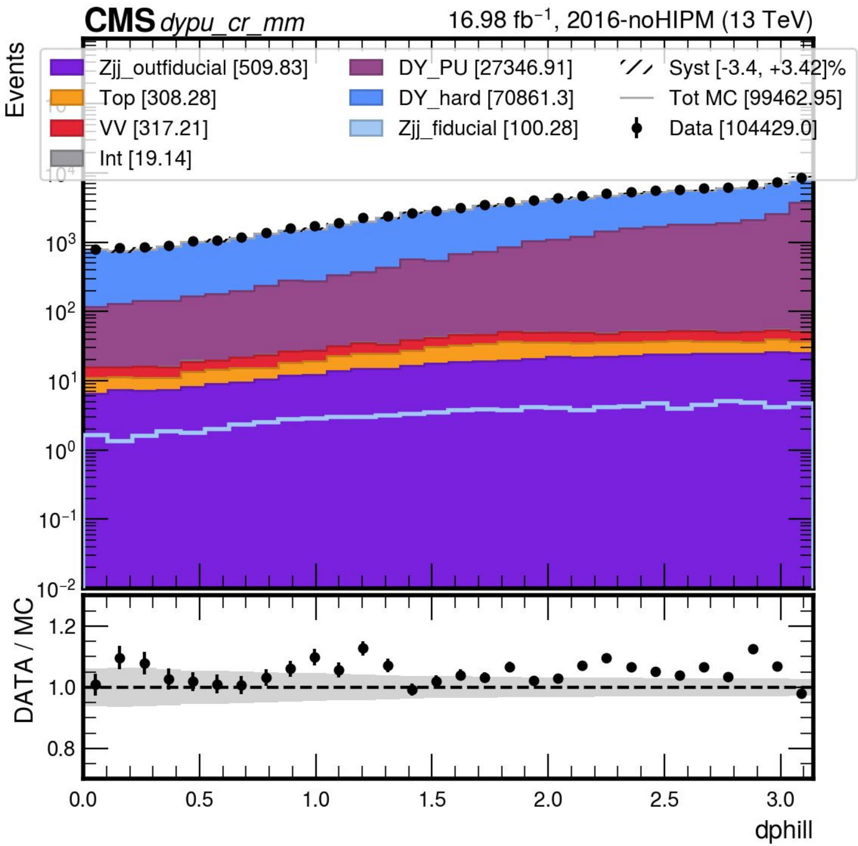


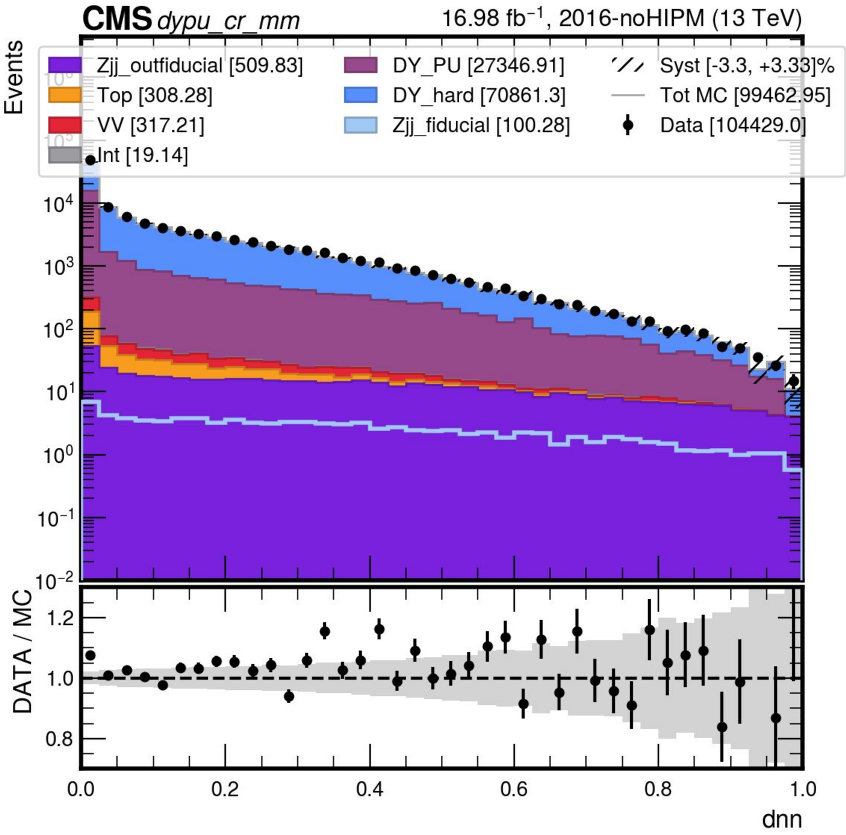
The DY PU CR


CMS dypu_cr_ee
16.98 fb^{-1}, 2016-noHIPM (13 TeV)
Zjj_outfiducial [293.73]
Top [169.57]
VV [188.78]
Int [9.99]
DY_PU [14053.64]
DY_hard [37992.37]
Zjj_fiducial [58.13]
Syst [-4.2, +4.22]%
Tot MC [52766.22]
Data [57647.0]
Events
DATA / MC
dphill

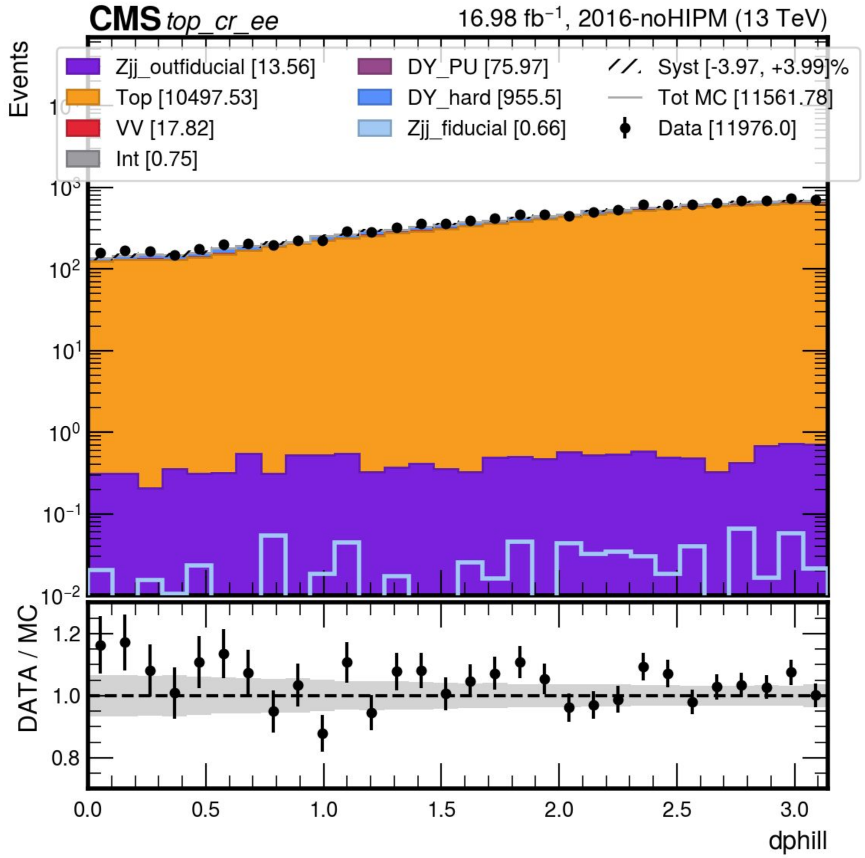

CMS dypu_cr_ee
16.98 fb^{-1}, 2016-noHIPM (13 TeV)
Zjj_outfiducial [293.73]
Top [169.57]
VV [188.78]
Int [9.99]
DY_PU [14053.64]
DY_hard [37992.37]
Zjj_fiducial [58.13]
Syst [-3.99, +4.02]%
Tot MC [52766.22]
Data [57647.0]
Events
DATA / MC
dnn

The DY PU CR

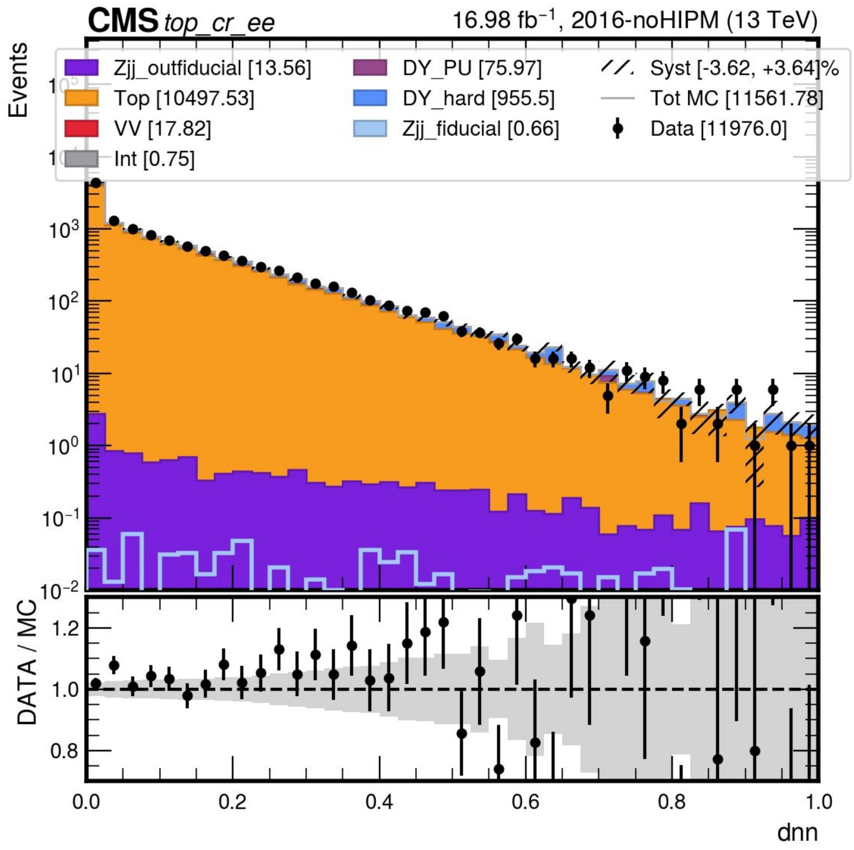

CMS dypu_cr_mm
16.98 fb^{-1}, 2016-noHIPM (13 TeV)
Zjj_outfiducial [509.83]
Top [308.28]
VV [317.21]
Int [19.14]
DY_PU [27346.91]
DY_hard [70861.3]
Zjj_fiducial [100.28]
Syst [-3.4, +3.42]%
Tot MC [99462.95]
Data [104429.0]
Events
DATA / MC
dphill


CMS dypu_cr_mm
16.98 fb^{-1}, 2016-noHIPM (13 TeV)
Zjj_outfiducial [509.83]
Top [308.28]
VV [317.21]
Int [19.14]
DY_PU [27346.91]
DY_hard [70861.3]
Zjj_fiducial [100.28]
Syst [-3.3, +3.33]%
Tot MC [99462.95]
Data [104429.0]
Events
DATA / MC
dnn

The Top CR

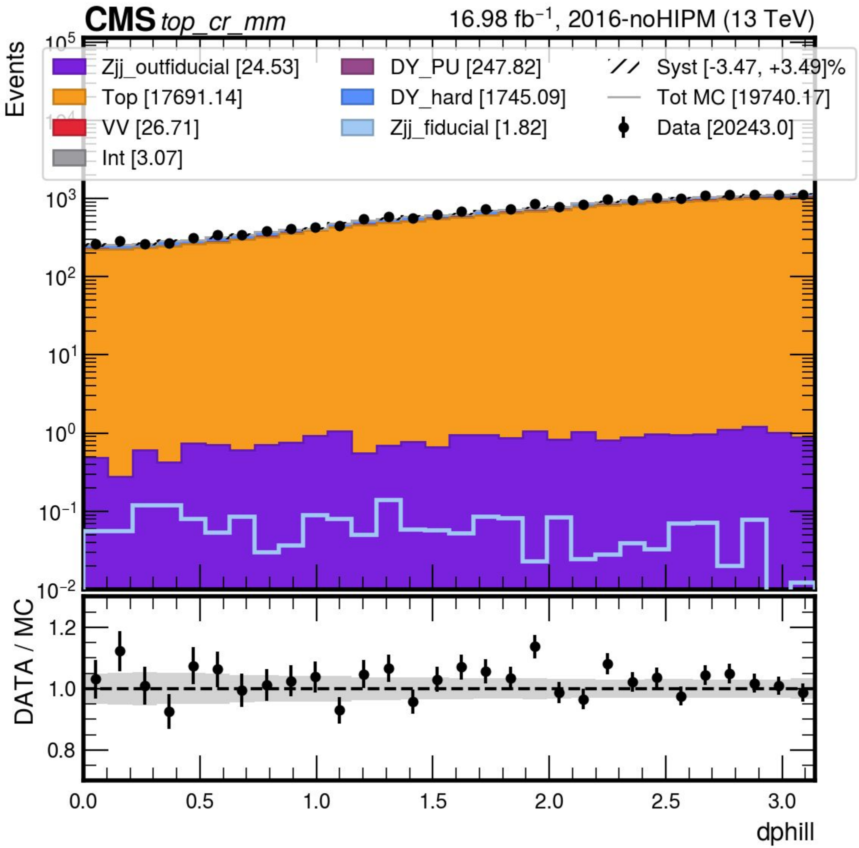


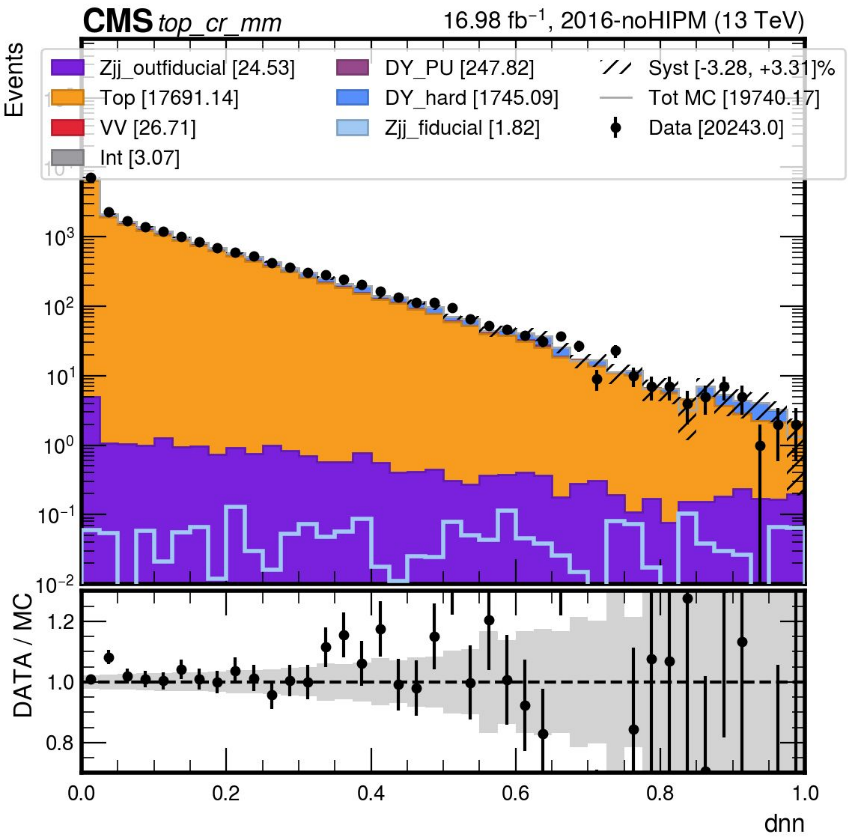
CMS top_cr_ee
16.98 fb^{-1}, 2016-noHIPM (13 TeV)
Zjj_outfiducial [13.56]
Top [10497.53]
VV [17.82]
Int [0.75]
DY_PU [75.97]
DY_hard [955.5]
Zjj_fiducial [0.66]
Syst [-3.97, +3.99]%
Tot MC [11561.78]
Data [11976.0]
Events
DATA / MC
dphill



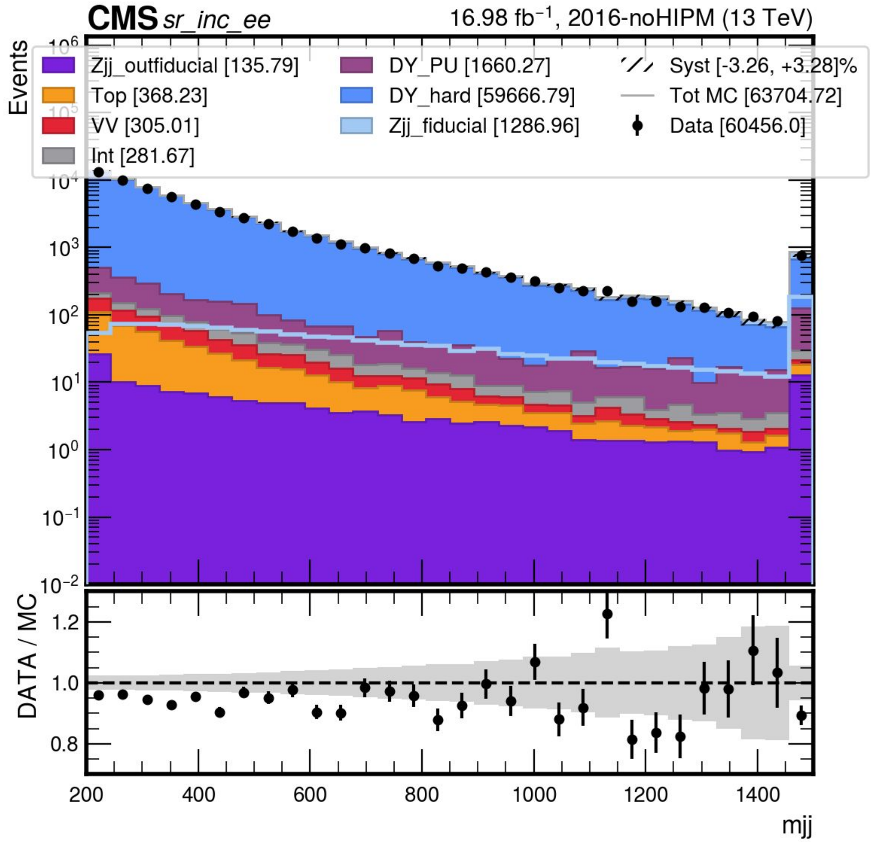
CMS top_cr_ee
16.98 fb^{-1}, 2016-noHIPM (13 TeV)
Zjj_outfiducial [13.56]
Top [10497.53]
VV [17.82]
Int [0.75]
DY_PU [75.97]
DY_hard [955.5]
Zjj_fiducial [0.66]
Syst [-3.62, +3.64]%
Tot MC [11561.78]
Data [11976.0]
Events
DATA / MC
dnn

The Top CR

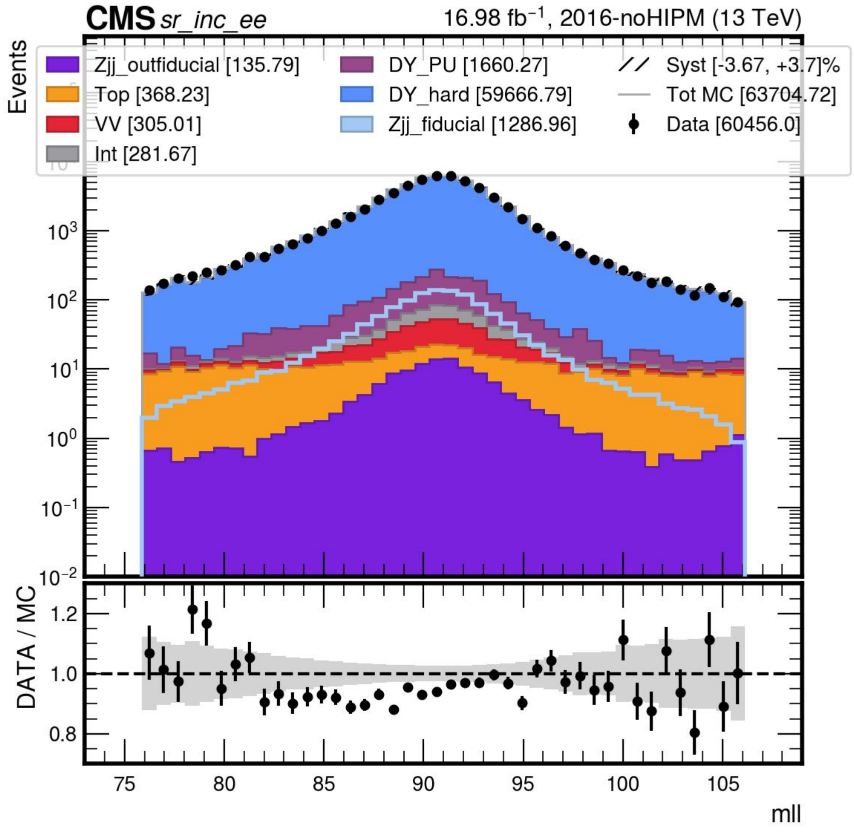

CMS top_cr_mm
16.98 fb^{-1}, 2016-noHIPM (13 TeV)
Events
Zjj_outfiducial [24.53]
Top [17691.14]
VV [26.71]
Int [3.07]
DY_PU [247.82]
DY_hard [1745.09]
Zjj_fiducial [1.82]
Syst [-3.47, +3.49]%
Tot MC [19740.17]
Data [20243.0]
DATA / MC
dphill


CMS top_cr_mm
16.98 fb^{-1}, 2016-noHIPM (13 TeV)
Events
Zjj_outfiducial [24.53]
Top [17691.14]
VV [26.71]
Int [3.07]
DY_PU [247.82]
DY_hard [1745.09]
Zjj_fiducial [1.82]
Syst [-3.28, +3.31]%
Tot MC [19740.17]
Data [20243.0]
DATA / MC
dnn

The “SR”

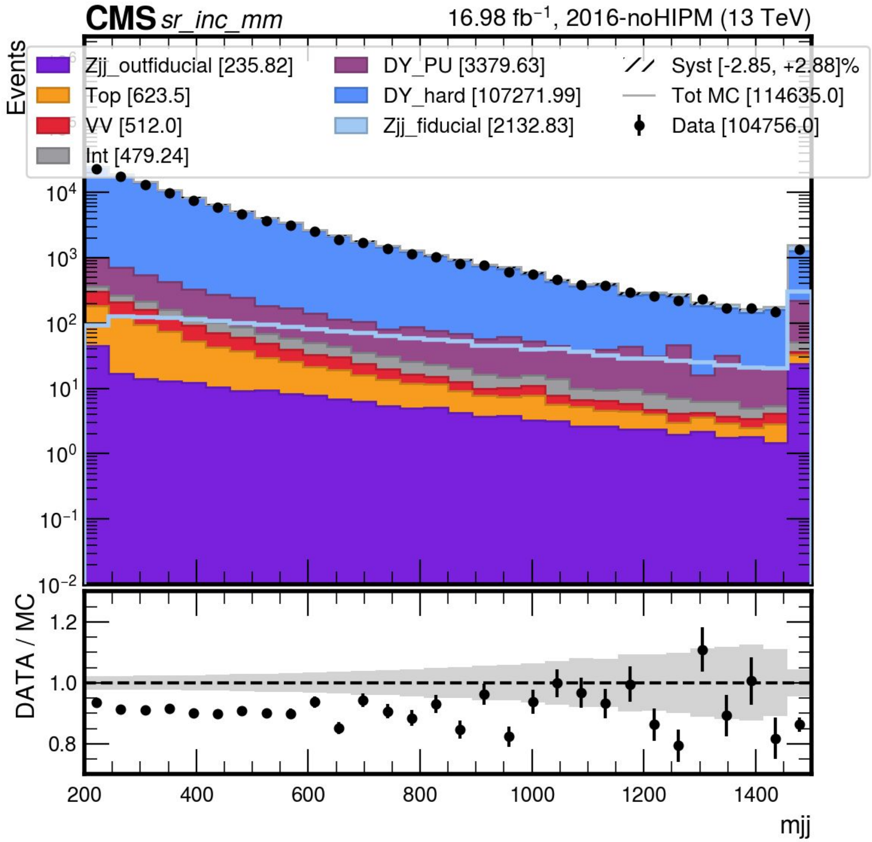


CMS sr_inc_ee
16.98 fb⁻¹, 2016-noHIPM (13 TeV)
Events
Zjj_outfiducial [135.79]
Top [368.23]
VV [305.01]
Int [281.67]
DY_PU [1660.27]
DY_hard [59666.79]
Zjj_fiducial [1286.96]
Syst [-3.26, +3.28]%
Tot MC [63704.72]
Data [60456.0]
DATA / MC
mjj

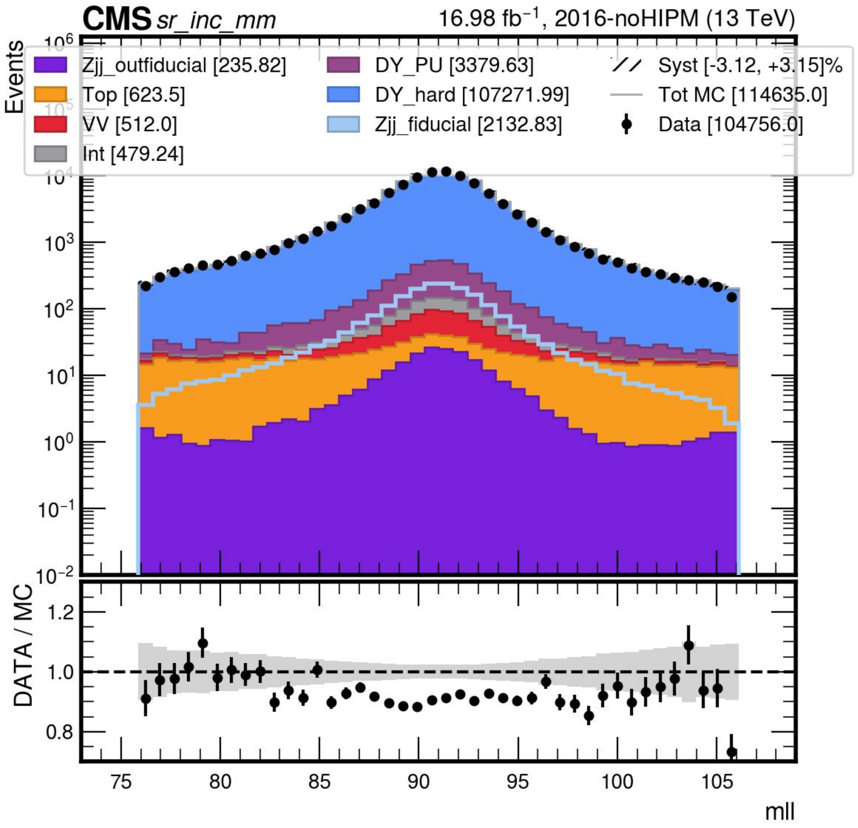


CMS sr_inc_ee
16.98 fb⁻¹, 2016-noHIPM (13 TeV)
Events
Zjj_outfiducial [135.79]
Top [368.23]
VV [305.01]
Int [281.67]
DY_PU [1660.27]
DY_hard [59666.79]
Zjj_fiducial [1286.96]
Syst [-3.67, +3.7]%
Tot MC [63704.72]
Data [60456.0]
DATA / MC
mll

The “SR”

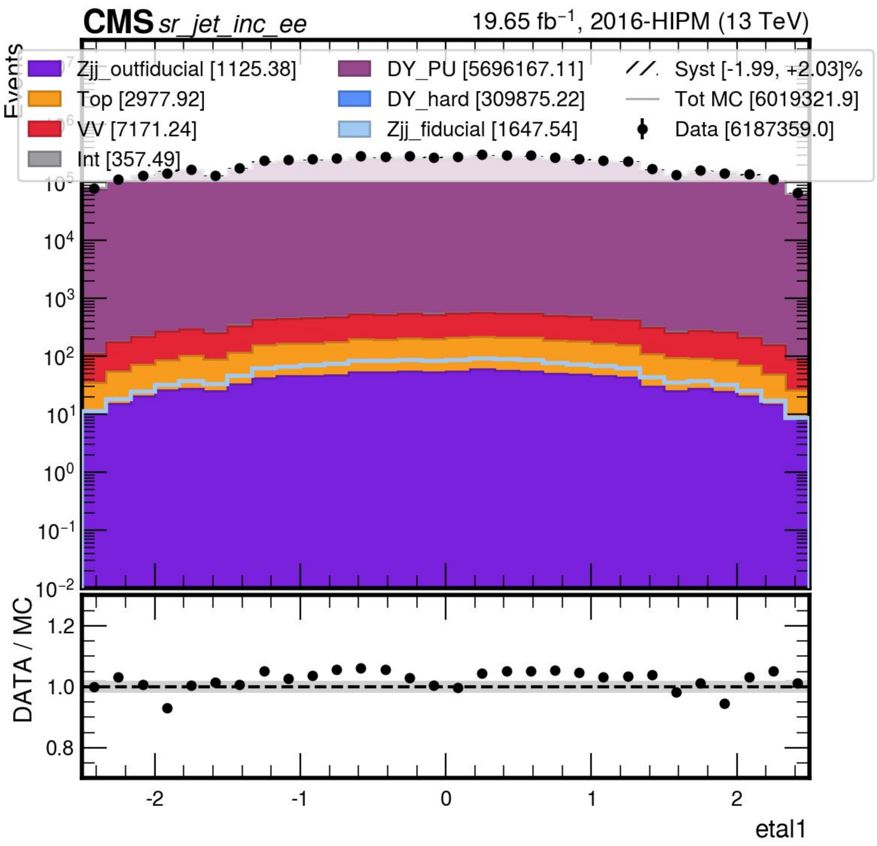


CMS sr_inc_mm
16.98 fb^{-1}, 2016-noHIPM (13 TeV)
Events
Zjj_outfiducial [235.82]
Top [623.5]
VV [512.0]
Int [479.24]
DY_PU [3379.63]
DY_hard [107271.99]
Zjj_fiducial [2132.83]
Syst [-2.85, +2.88]%
Tot MC [114635.0]
Data [104756.0]
DATA / MC
mjj

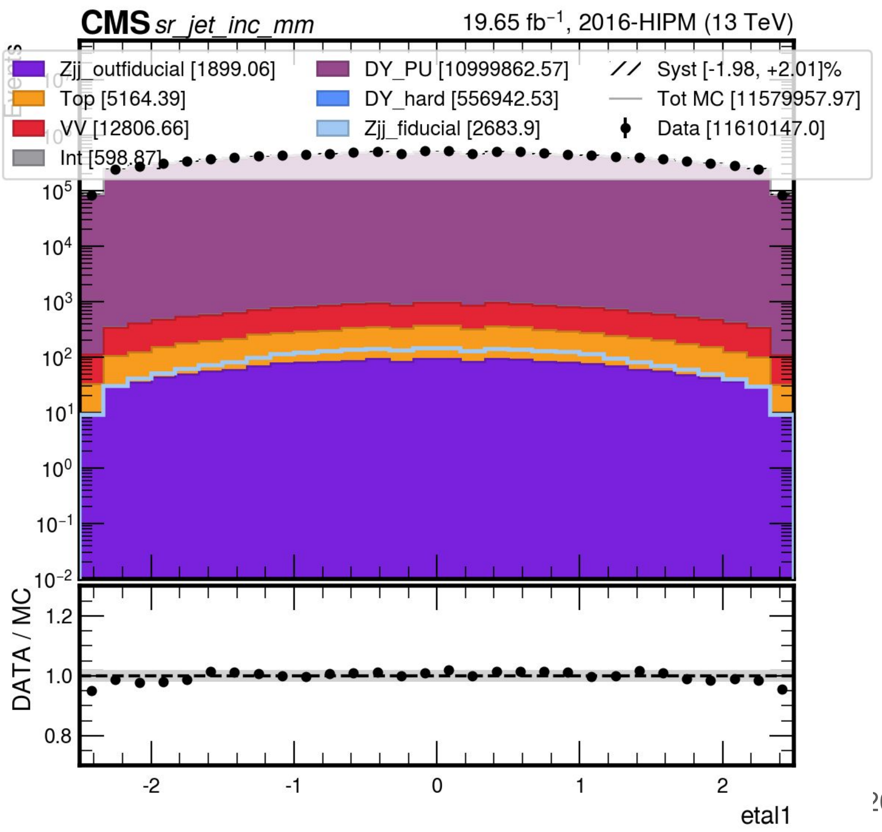


CMS sr_inc_mm
16.98 fb^{-1}, 2016-noHIPM (13 TeV)
Events
Zjj_outfiducial [235.82]
Top [623.5]
VV [512.0]
Int [479.24]
DY_PU [3379.63]
DY_hard [107271.99]
Zjj_fiducial [2132.83]
Syst [-3.12, +3.15]%
Tot MC [114635.0]
Data [104756.0]
DATA / MC
mll

16 HIPM

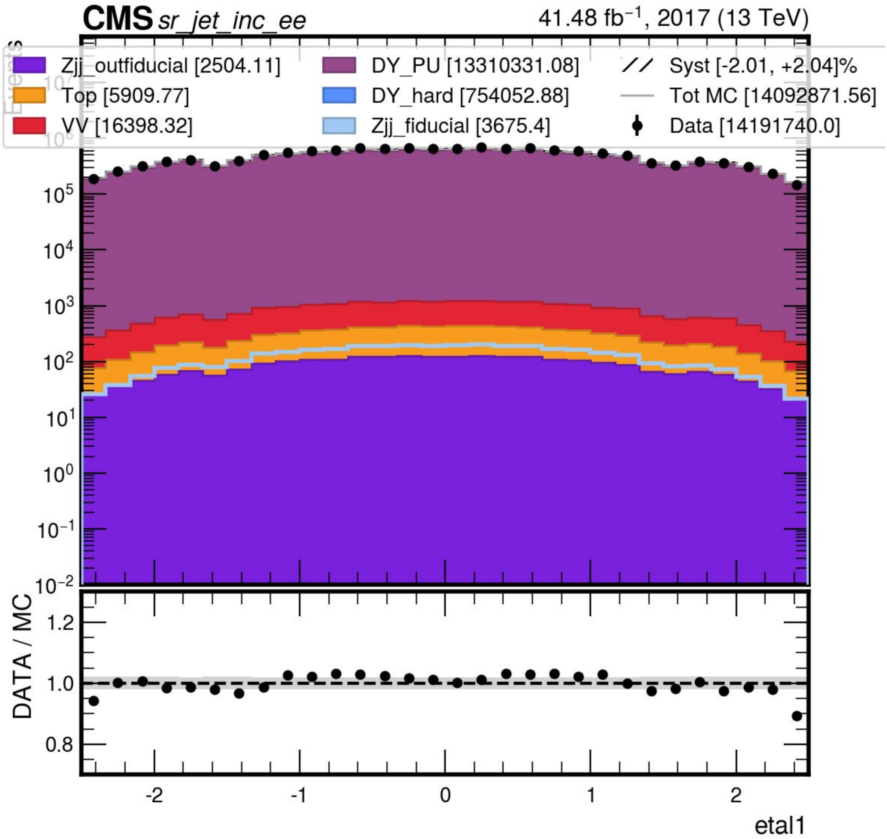


CMS sr_jet_inc_ee
19.65 fb^{-1}, 2016-HIPM (13 TeV)
Zjj_outfiducial [1125.38]
Top [2977.92]
VV [7171.24]
Int [357.49]
DY_PU [5696167.11]
DY_hard [309875.22]
Zjj_fiducial [1647.54]
Syst [-1.99, +2.03]%
Tot MC [6019321.9]
Data [6187359.0]
Events
DATA / MC
etal1

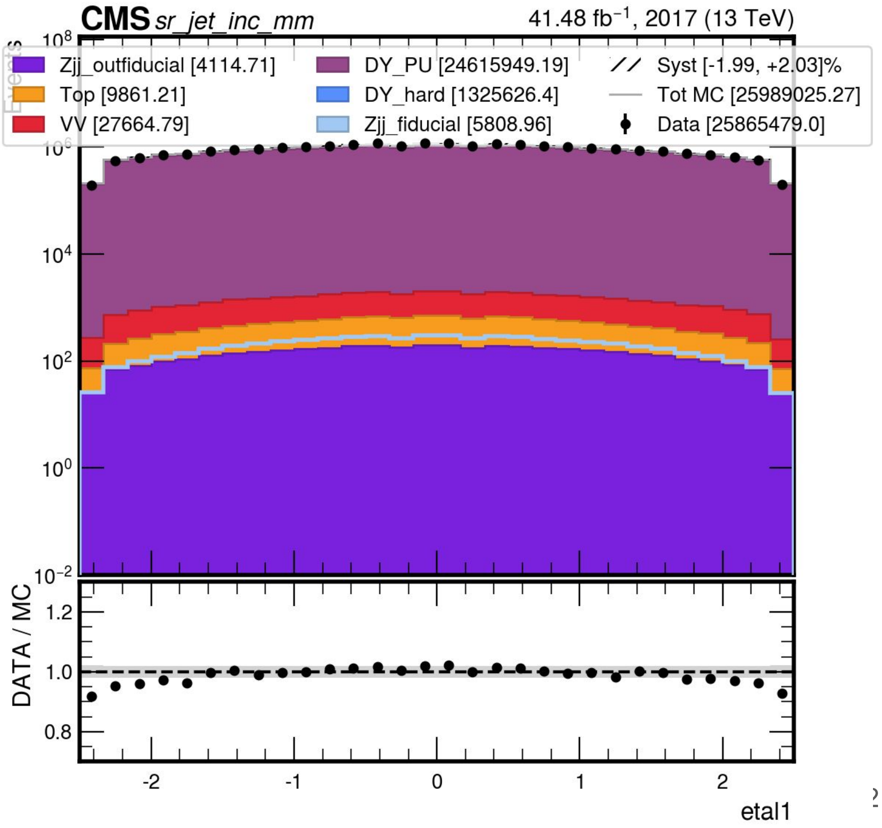


CMS sr_jet_inc_mm
19.65 fb^{-1}, 2016-HIPM (13 TeV)
Zjj_outfiducial [1899.06]
Top [5164.39]
VV [12806.66]
Int [598.87]
DY_PU [10999862.57]
DY_hard [556942.53]
Zjj_fiducial [2683.9]
Syst [-1.98, +2.01]%
Tot MC [11579957.97]
Data [11610147.0]
Events
DATA / MC
etal1

17

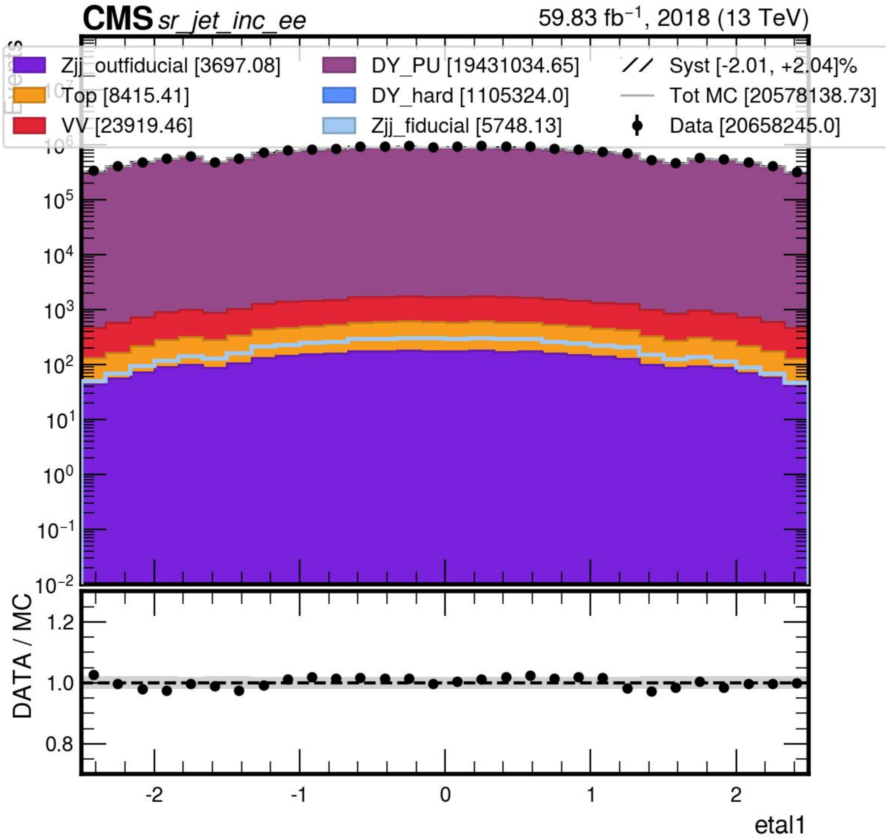


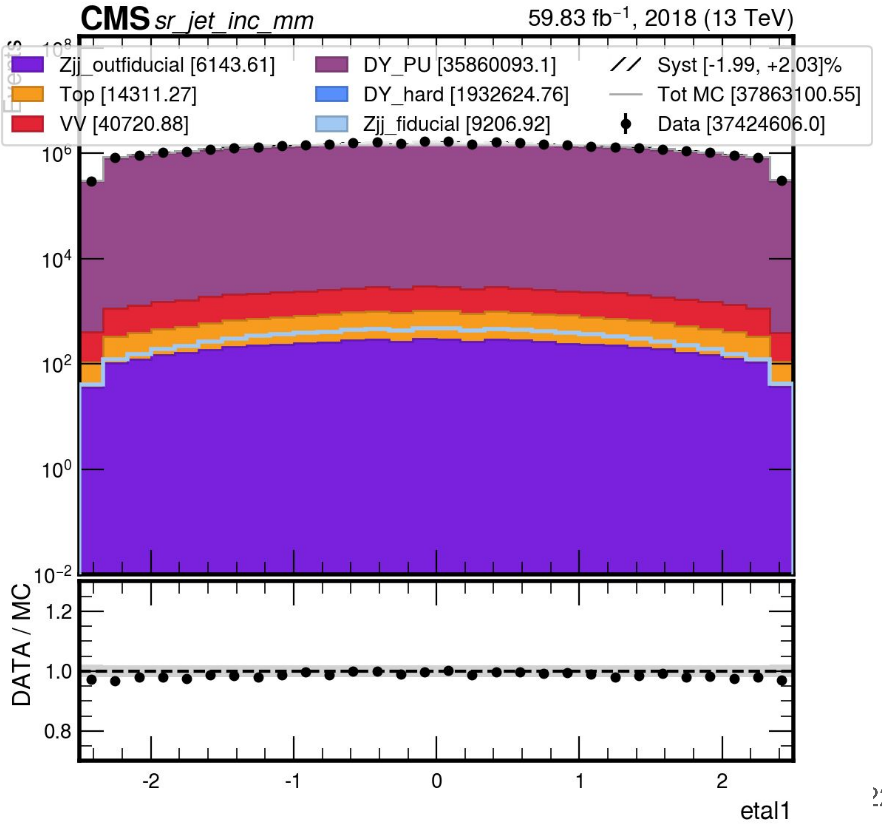
CMS sr_jet_inc_ee
41.48 fb^{-1}, 2017 (13 TeV)
Zjj_outfiducial [2504.11]
Top [5909.77]
VV [16398.32]
DY_PU [13310331.08]
DY_hard [754052.88]
Zjj_fiducial [3675.4]
Syst [-2.01, +2.04]%
Tot MC [14092871.56]
Data [14191740.0]
DATA / MC
etal1



CMS sr_jet_inc_mm
41.48 fb^{-1}, 2017 (13 TeV)
Zjj_outfiducial [4114.71]
Top [9861.21]
VV [27664.79]
DY_PU [24615949.19]
DY_hard [1325626.4]
Zjj_fiducial [5808.96]
Syst [-1.99, +2.03]%
Tot MC [25989025.27]
Data [25865479.0]
DATA / MC
etal1

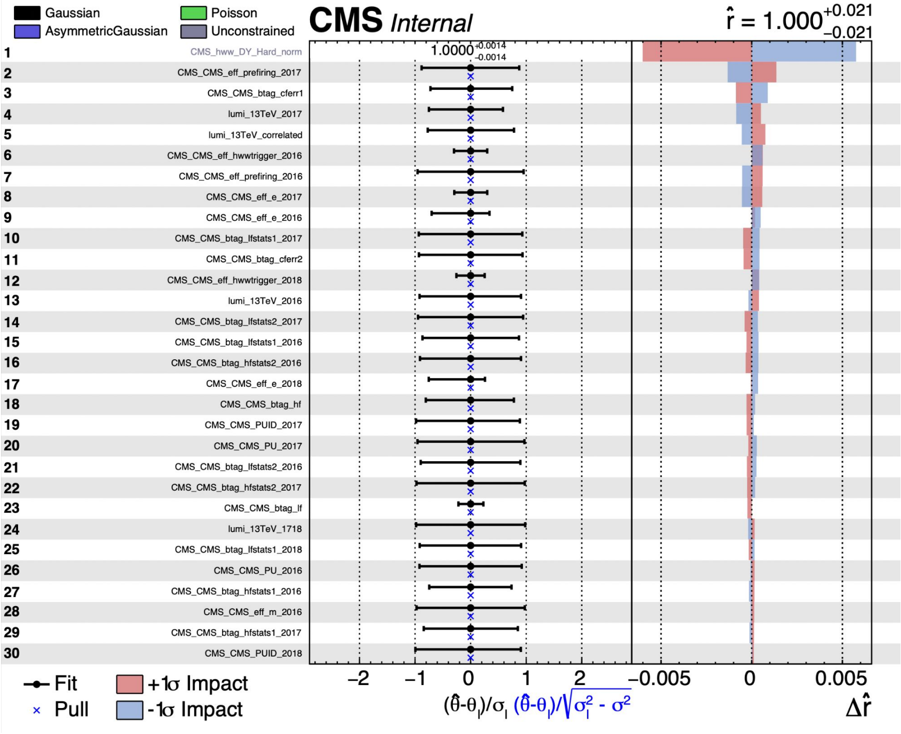
18


CMS sr_jet_inc_ee
59.83 fb^{-1}, 2018 (13 TeV)
Zjj_outfiducial [3697.08]
Top [8415.41]
VV [23919.46]
DY_PU [19431034.65]
DY_hard [1105324.0]
Zjj_fiducial [5748.13]
Syst [-2.01, +2.04]%
Tot MC [20578138.73]
Data [20658245.0]
DATA / MC
etal1

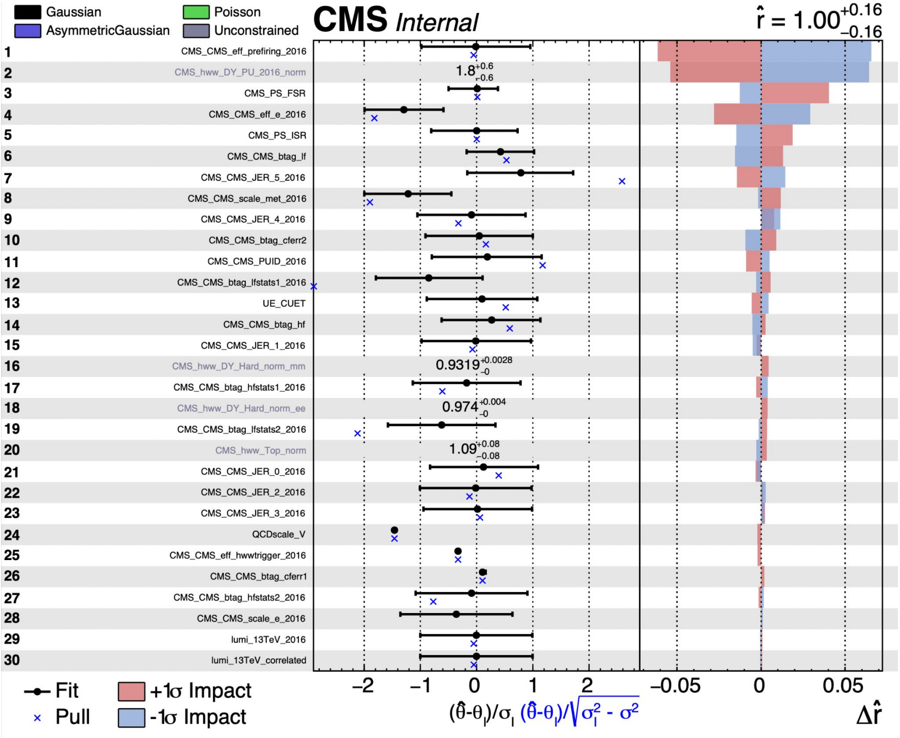

CMS sr_jet_inc_mm
59.83 fb^{-1}, 2018 (13 TeV)
Zjj_outfiducial [6143.61]
Top [14311.27]
VV [40720.88]
DY_PU [35860093.1]
DY_hard [1932624.76]
Zjj_fiducial [9206.92]
Syst [-1.99, +2.03]%
Tot MC [37863100.55]
Data [37424606.0]
DATA / MC
etal1

Impacts, should be updated with latest strategy

Full Run 2 with few nuisances



2016 HIPM with most nuisances



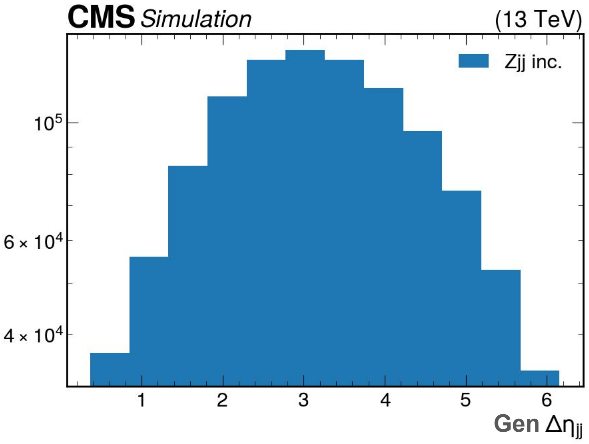
Differential cross section and unfolding

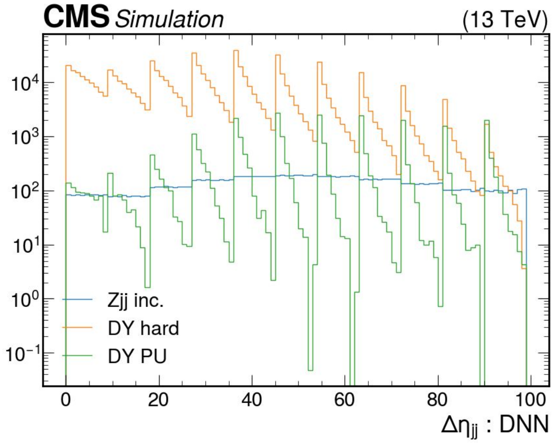
- Signal cross section is measured differentially as a function of relevant kinematic variables
- The unfolding of the detector response is performed via a Maximum-Likelihood-based method through Combine
- Each generator level bin is regarded as a different signal
- All gen level signals are fitted simultaneously
- Pros of this method (as reported in Combine documentation)
  - Background subtraction is accounted for directly in the likelihood
  - Systematic uncertainties are accounted for directly during the unfolding as nuisance parameters
  - We can profile the nuisance parameters during the unfolding to make the most of the data available
- No regularization procedure is applied

Unfolding strategy

- The current strategy is to perform a fit on a 2D variable: the reco variable to unfold and the DNN output (e.g. plot on the right)
- The DNN is used to separate the signal vs the background
- The reco variable is used to have sensitivity on the gen-level bins fit

Example of gen level distribution of $\Delta\eta_{jj}$





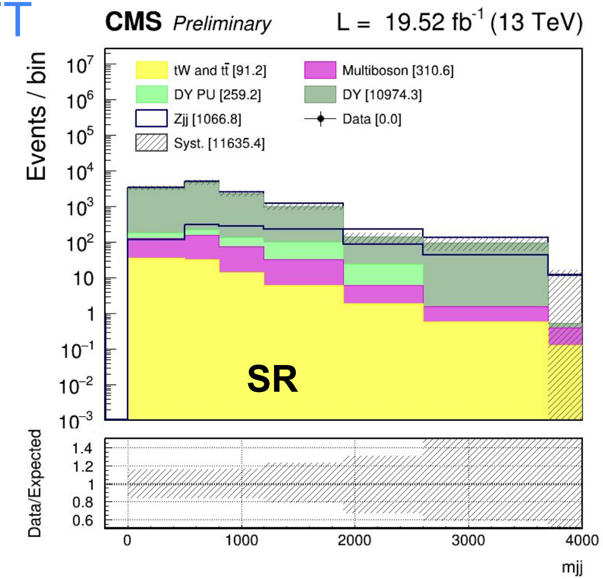
Variables to unfold

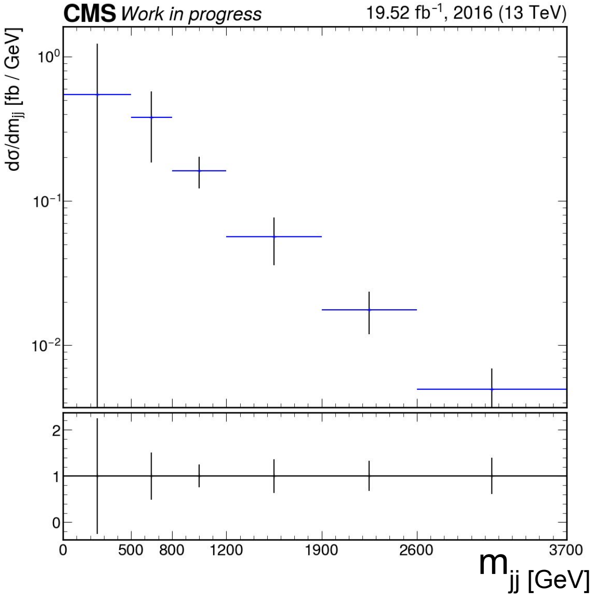
Useful for VBF and EFT

Variables to unfold:

- m_{jj}, p^T_{jj}, $\Delta\varphi_{jj}$, $\Delta\eta_{jj}$
- p^T_{ll}, $\Delta\varphi_{ll}$, $\Delta\eta_{ll}$
- p^T_{j1}, p^T_{j2}
- p^T_{l1}, p^T_{l2}
- p^T_{j3} (0 if no 3rd jet)
- HT (scalar sum of all non leading jets that have a pt > 20 GeV)

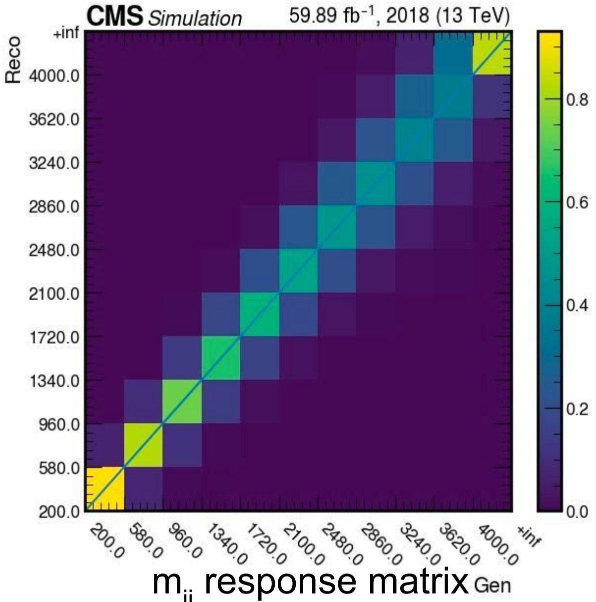
Useful for PS and NLO QCD studies





Response matrix example: dijet invariant mass

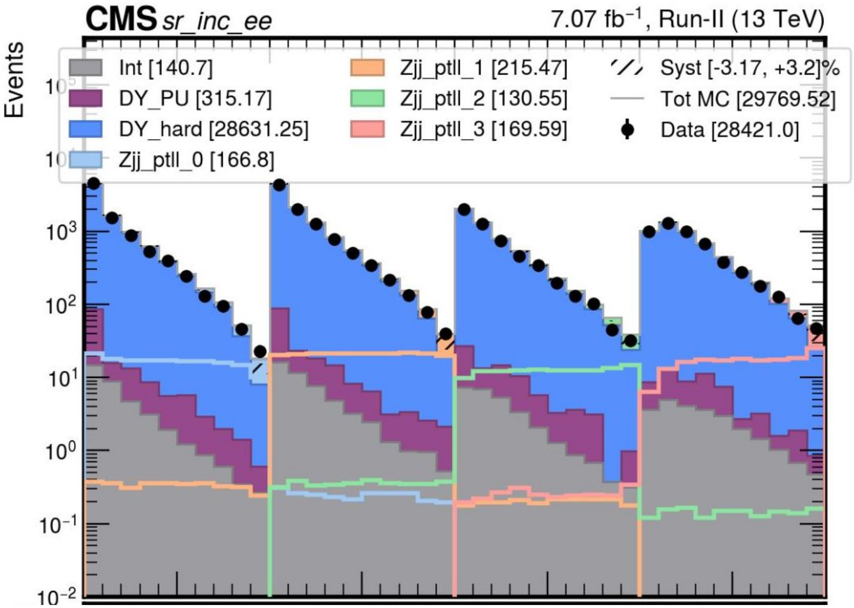
- The response matrix is built comparing each gen bin (that will be a single signal) to its reco level distribution



m_{jj} response matrix

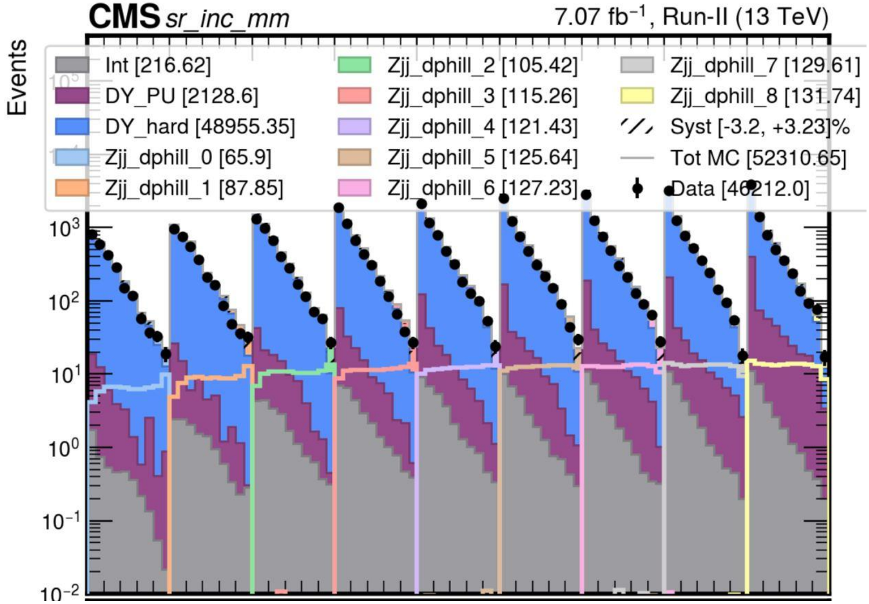
Ptll unfolding

Bins:
[0., 50., 100., 150., 200.]



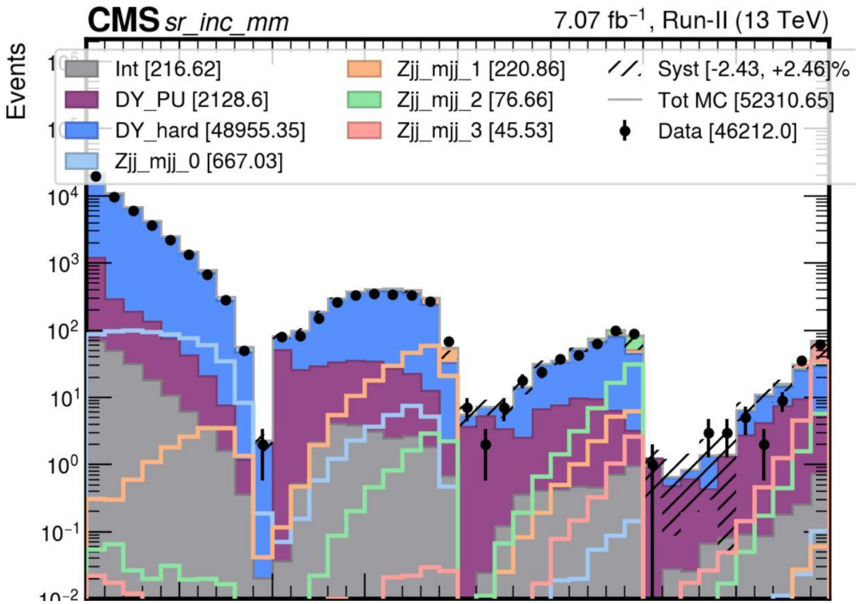
dphill unfolding

Bins:
[0.0, 0.35, 0.7, 1.05, 1.4, 1.74, 2.09, 2.44, 2.79, 3.14]



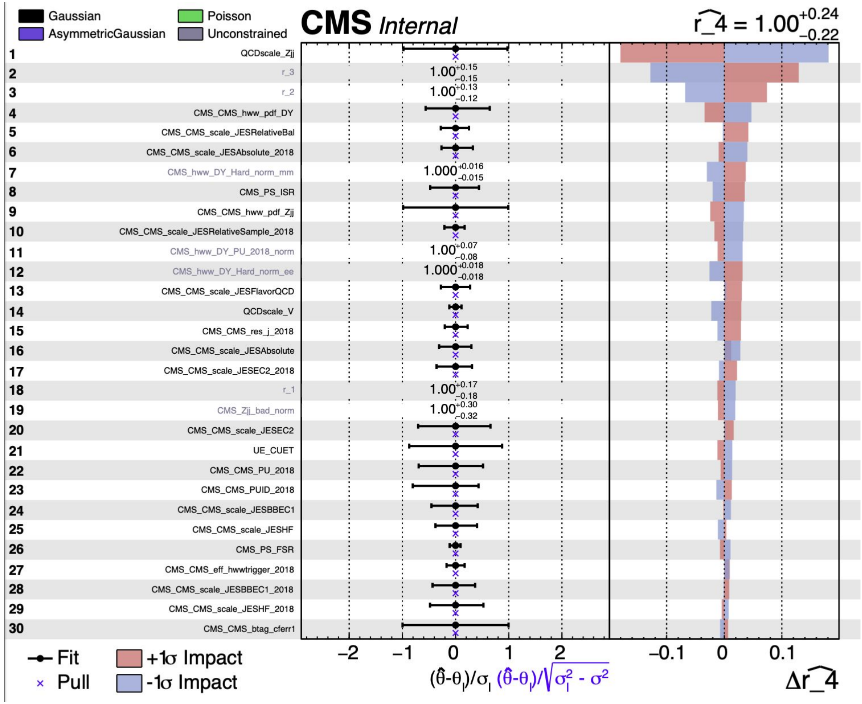
mjj unfolding

Bins:
[200.0, 900.0, 1600.0, 2300.0, 3000.0]



Expected impacts

- Example of impacts for r_4: signal strength modifier of last gen bin



Summary and status of the VBF-Z analysis

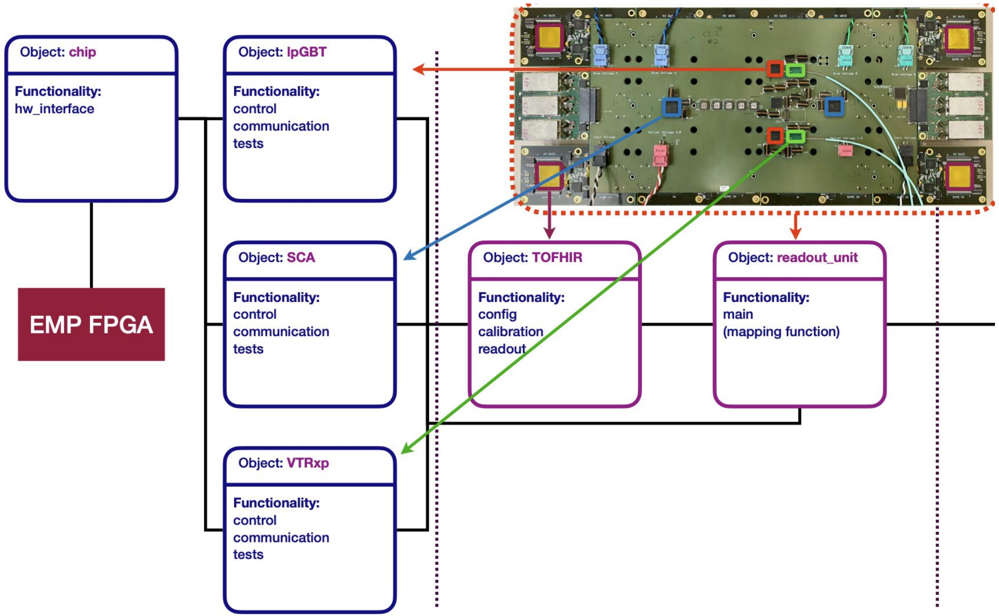
- All corrections, SF and regions are finally under control
- First Full Run 2 impacts with these new processing will be done next week
- The unfolding is already in place and is just missing the activation of systematics for impacts
- Once we have the impacts we're ready for pre approval!

MTD

- Starting from february I began working on the MTD DAQ
- Outline:
  - DAQ Software organization
  - Orchestration of the DAQ initialization and basic slow control readout
  - Tofhir raw output and reconstruction

DAQ Software organization

- Serenity relies on the EMP framework that creates the interface between the firmware and software
- Each physical chip is mapped to a DAQ sw object that inherits from the basic EMP chip



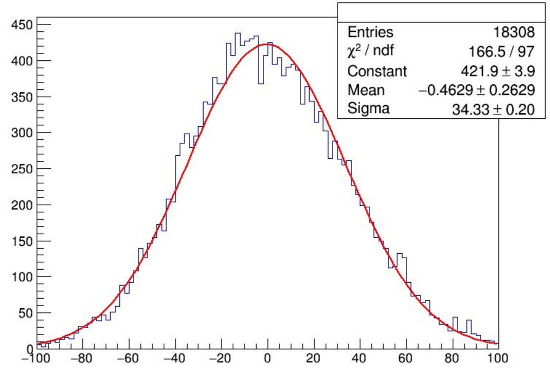
Orchestration of Serenity initialization and basic slow control readout

- Given the complex physical and software structure of the DAQ an orchestration of the configurations and initialization is needed even for basic operations as the temperature readings (slow control)
- Assembly Centers have to check the temperature readings and that the DAQ is able to communicate with freshly built RUs
  - We simplified and hid the complexity in automatized scripts in order for the average user to be able to perform basic DAQ tasks
- A construction DB is filled by the AC with the bar codes of each component
  - The DAQ is currently missing an interface with such a DB to retrieve configuration and calibration files → my current task

Tofhir raw output and reconstruction

- The processing of the raw output gives for a single hit:
    - Absolute time
    - ToT (time over threshold)
    - Energy
- A proper calibration of energies and times is done with calibration files
- Hits get grouped together within a time window to form an event
- Calibration routines have still to be ported to the EMP framework MTD DAQ

- Bits 0-4: channel identifier;
- Bits 5-4: identifier of the time-to-amplitude converter in multi-buffer TAC;
- Bits 15-6: charge measurement;
- Bits 25-16: fine counter of the 2nd time measurement;
- Bits 35-26: fine counter of the 1st time measurement;
- Bits 45-36: coarse counter of the previous event crossing the timing threshold;
- Bits 55-46: coarse counter of the end of charge integration;
- Bits 65-56: coarse counter of the 2nd time measurement;
- Bits 81-66: coarse counter of the 1st time measurement;

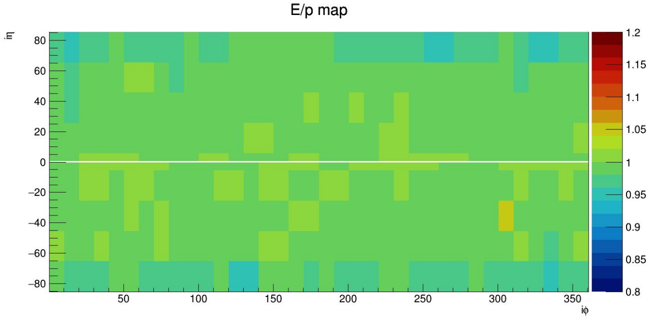


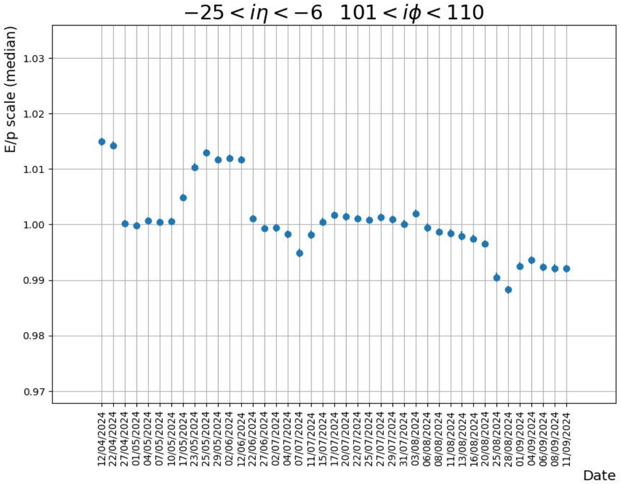
Calibrated Time Resolution [ps]

Technical tasks

ECAL Calibration and Monitoring Automation: E/p

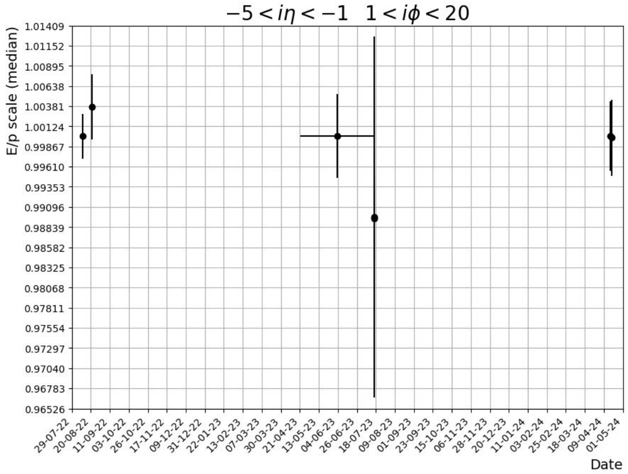
- ECAL calibration and monitoring was moved to an automated chain, where for each fill/run jobs are automatically run in sequence and resubmitted by the automation tool
- There were some requests from the DPG convener that were implemented but not yet integrated into production (will have to do that to validate the EPRs)





ECAL Calibration and Monitoring Automation: E/p

- Developments of 2024:
  - Integration into the automation production pipeline
  - Now renormalizing to the first IC
  - New plot for the cumulative view of ICs that takes into account the time duration of the IOV and a correct scaling for the x axis



Ph.D. Courses and Schools

- Physics courses:
  - Physics at Colliders 2 CFU (in three weeks)

- Physics Schools:
  - AIPHY 2 CFU (in two weeks)

With these courses I'll have 10 CFU from physics courses and 3 CFU from interdisciplinary.

Backup

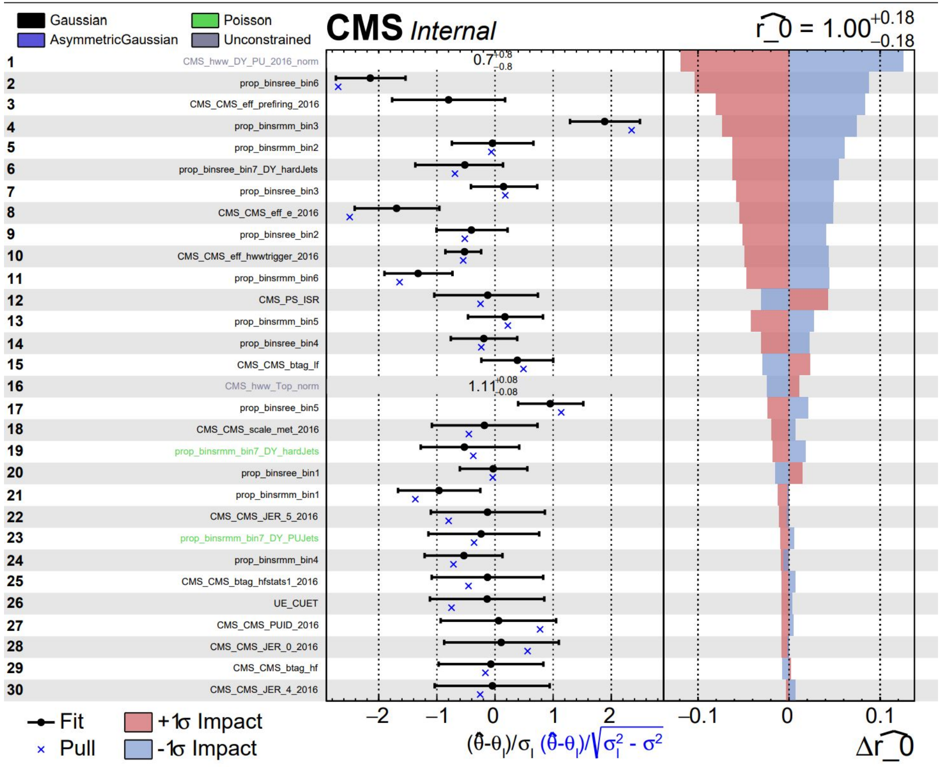
Object selections and corrections (detailed view)

- Trigger filters for each run era
- MET filters
- Jets selected from AK4CHS, tightID and PU ID for pt < 50 GeV
- Jets cleaned from loose Leptons
- Prompt gen matching is required for the two leading leptons
- Rochester corrections are applied to muons
- Trigger SF applied for the two leptons
- Leptons SF
- Jets are vetoed with the JME POG map, HEM issue of 2018 is solved by removing the jets in the affected zone
- L1L2L3Res JEC are applied
- JER applied only for eta < 2.5
- PU Weight SF and PUID SF applied
- L1PreFiring for 2016-2017 is taken into account
- B-tag SF applied

OUTDATED

Final fit (inclusive): 2016 HIPM (first era of 2016)

- Data-asimov fit
- We know that JER/JES together with QCD scales represents important contribution
- DY normalization factor for now is the leading systematic
- Counting the shifts of + sigma and - sigma of the systematics fit, the result is reasonable



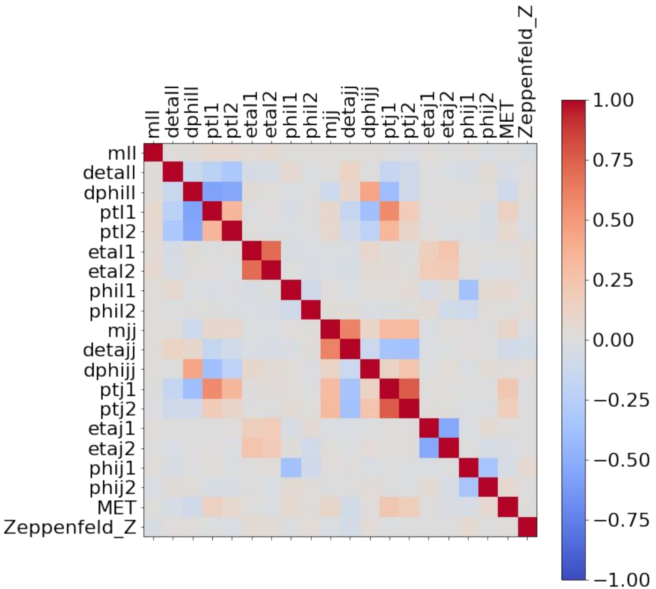
Chi2: 19.57
Chi2/ndof: 0.4164

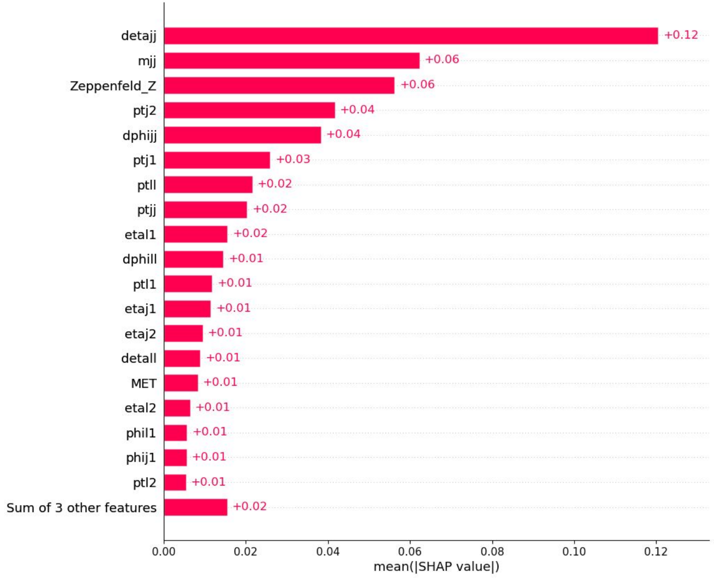
Deep Neural Network

- A Deep Neural Network has been trained to separate signal and background in a phase space with loose and generic cuts
- The chosen model was a DNN with 4 dense layers each with 128 neurons and the
- Training was performed with the Binomial Cross Entropy as loss function
- Since EW VBF-Z and DY appear to be the most difficult to separate, the model is trained only with these two samples
- The separation of the signal with the all the other backgrounds is reached even with the above training
- The DNN output, i.e. the score given by the DNN to each event during the evaluation step, is transformed with a simple function in order to have a flat signal and all the backgrounds peaking at 0
- The function chosen that satisfied this requirement is the cumulative of the DNN output evaluated on the signal

DNN input variables

- Do not use Parton Shower sensible variables





DNN evaluation and performances