Data-MC discrepancy study

Yang Tianyi

Samples

- Data
  - Run 2022C-G Muon
  - Golden json luminosity 5.0104, 2.9700, 5.8070, 17.7819, 3.0828 fb^{-1}
  - Run 2023C-D Muon0 & Muon1
  - Golden json luminosity: 17.794, 9.451 fb^{-1}
- Trigger menu
  - From Run 3 TWZ analysis note AN -2024/226
  - Single muon trigger `HLT_IsoMu{20|24}{_eta2p1} HLT_IsoMu27 HLT_Mu50`
  - Dimuon trigger `HLT_Mu{17|19}_TrkIsoVVL_Mu{8|9}_TrkIsoVVL_DZ_Mass{3p8|8}`
  - Trimuon triggers `HLT_TripleMu_5_3_3_Mass3p8_DZ HLT_TripleMu_10_5_5_DZ HLT_TripleMu_12_10_5`
- MET flag:
  - `Flag_goodVertices, Flag_globalSuperTightHalo2016Filter, Flag_EcalDeadCellTriggerPrimitiveFilter, Flag_BadPFMuonFilter, Flag_eeBadScFilter, Flag_BadPFMuonDzFilter, Flag_hfNoisyHitsFilter`

MC samples

- Dominant Drell-Yan global background:
  - DYto2L-4Jets_MLL-50_TuneCP5_13p6TeV_madgraphMLM-pythia8/NANOAODSIM/
- Cross-section (NNLO value): 6345.99 pb
- Other backgrounds:
  - $t\bar{t}$ DL/SL
  - Diboson, triboson
  - $t\bar{t}t\bar{t}$
  - $t\bar{t}VV$
  - Higgs: $t\bar{t}H$, $t\bar{t}VH$, tH

Object selections

- Muon p_T need to take Rochester corrections
- Muon selection
  - GoodMuon25: $|\eta| < 2.4$, tight ID, tight Iso, $p_T > 25\ \mathrm{GeV}$
  - GoodMuon15: $|\eta| < 2.4$, tight ID, tight Iso, $p_T > 15\ \mathrm{GeV}$
  - Tight isolation: Muon_pfRelIso04_all<0.15
- Dimuon preselection
  - Leading Muon fulfill $p_T > 25\ \mathrm{GeV}$ condition
  - Other Muons fulfill $p_T > 15\ \mathrm{GeV}$ condition
  - nGoodMuon_15 > 1 && nGoodMuon_25 > 0

MC corrections

- Pile-up correction
  - From https://gitlab.cern.ch/cms-nanoAOD/jsonpog-integration.git
  - Applying based on truth interactions
  - Adding additional weights
- Muon scale corrections
  - From https://gitlab.cern.ch/cms-nanoAOD/jsonpog-integration.git
  - Applying for the leading and subleading Muon
  - Adding additional scaling factors
- Drell-Yan Z p_T correction
  - From https://gitlab.cern.ch/cms-higgs-leprare/hleprare.git
  - Based on the truth Z boson p_T to fix the Muon momentum
  - Adding additional weights

Data-MC yield comparison

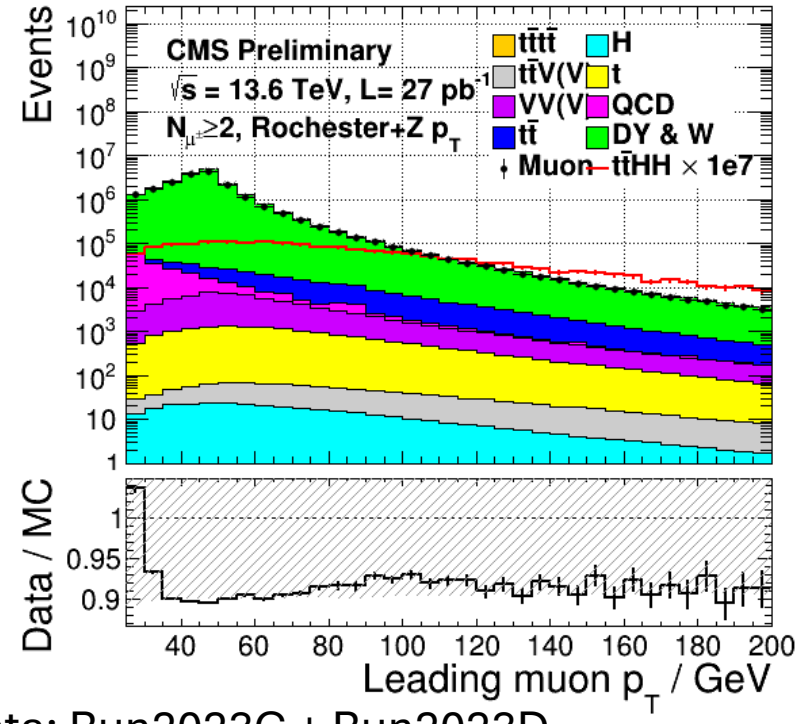
- The data-MC yield comparison of different eras.

MC campaign	22		22EE			23		23BPix	
Drell-Yan yield	5516392.2		17604112			14454755.88		7238221.833	
Era	Run2022C	Run2022D	Run2022E	Run2022F	Run2022G	Run2023C		Run2023D	
Primary dataset	Muon	Muon	Muon	Muon	Muon	Muon0	Muon1	Muon0	Muon1
Yield	3334682	2203560	4412469	12549031	2189719	7022361	7021051	3821148	3818724
Data/MC	1.003960886		1.087883274			0.971542662		1.055490171	

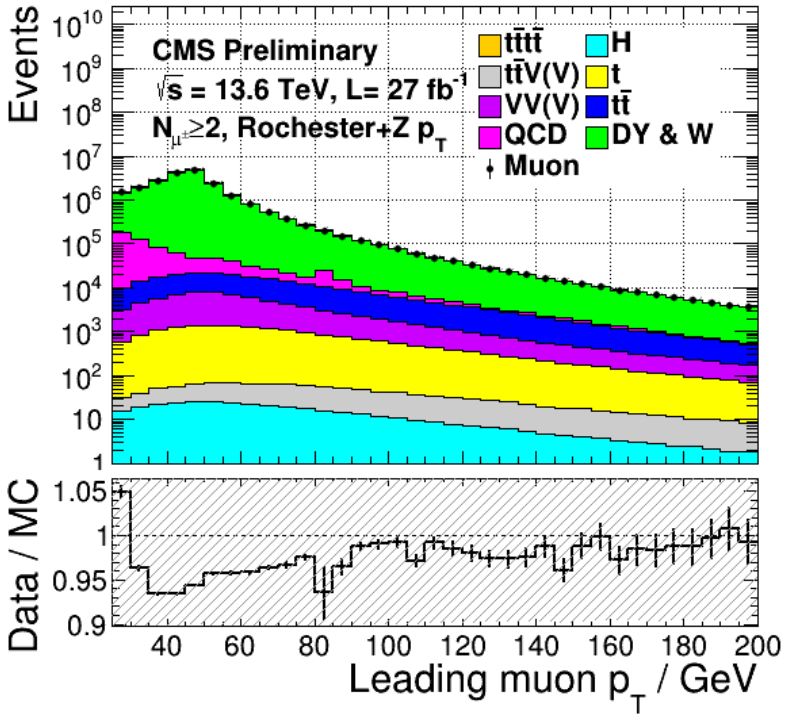
- Data/MC ratio variation is seen.
- Only around 3% overshooting seen in Run2023C.
- 22EE and 23BPix has data MC gap instead.

Trigger for data-MC unmatching in Run2023

Only di-muon trigger



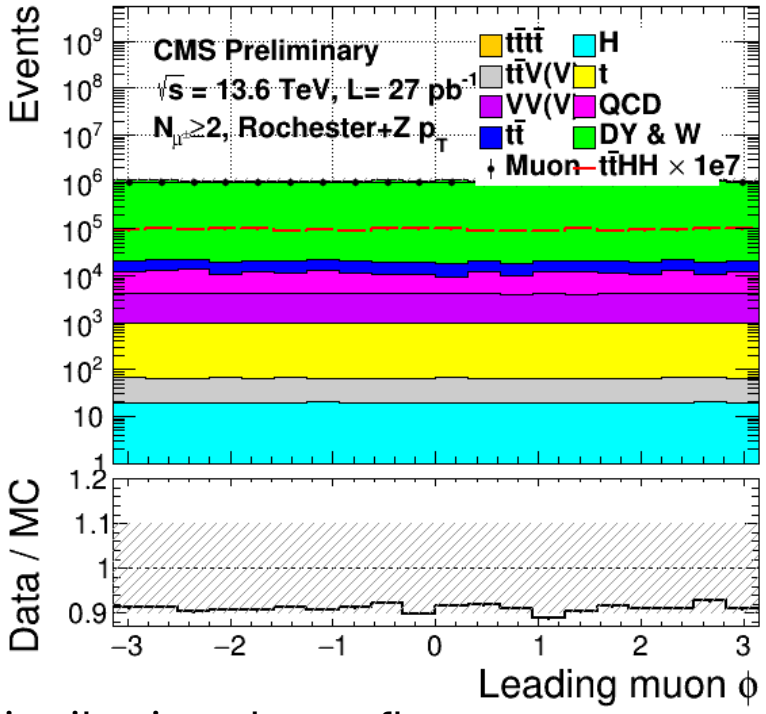
Current trigger scheme



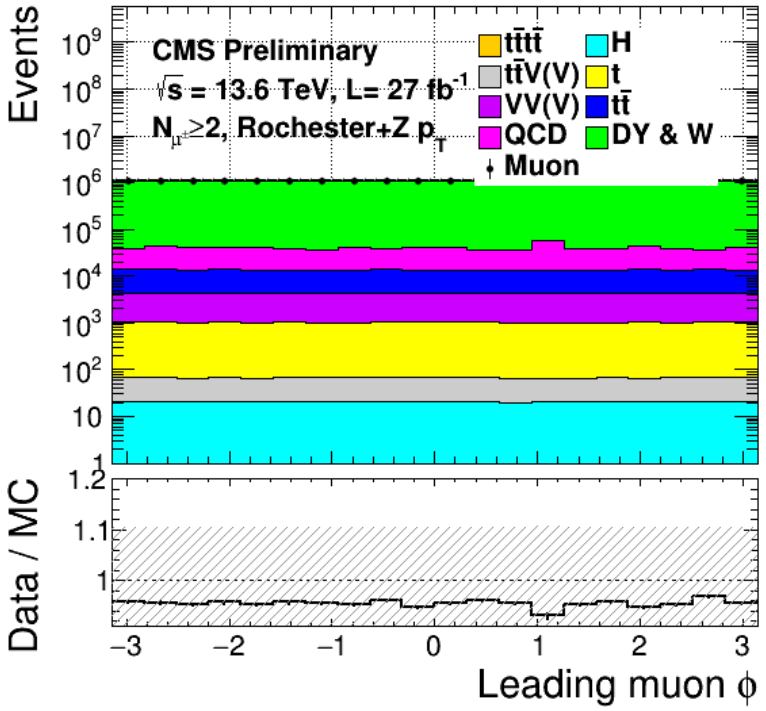
- Data: Run2023C + Run2023D
- MC: Run3Summer23NanoAODv12 + Run3Summer23BPixNanoAODv12.
- Including a larger trigger menu solve the overshooting.

Trigger for data-MC unmatching in Run2023

Only di-muon trigger



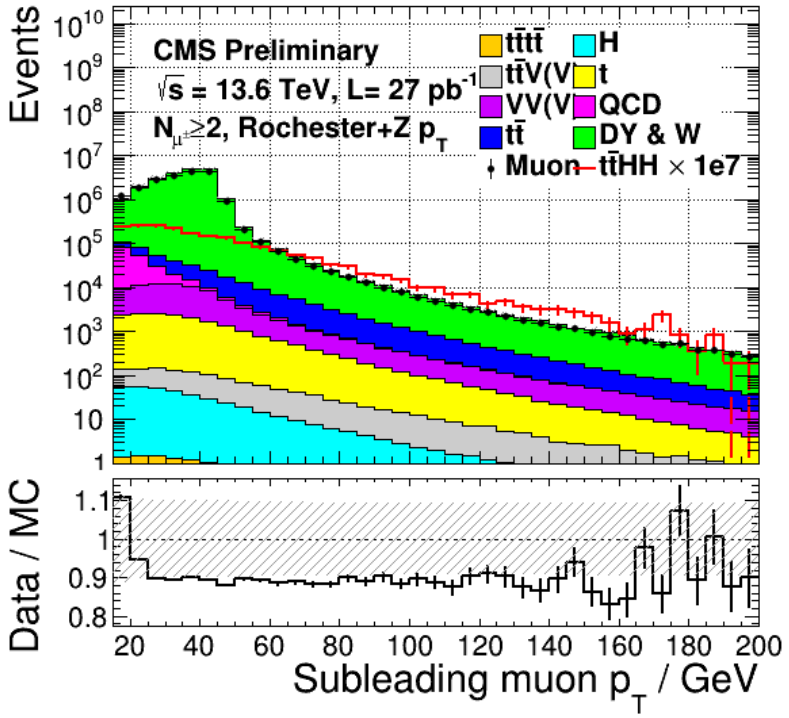
Current trigger scheme



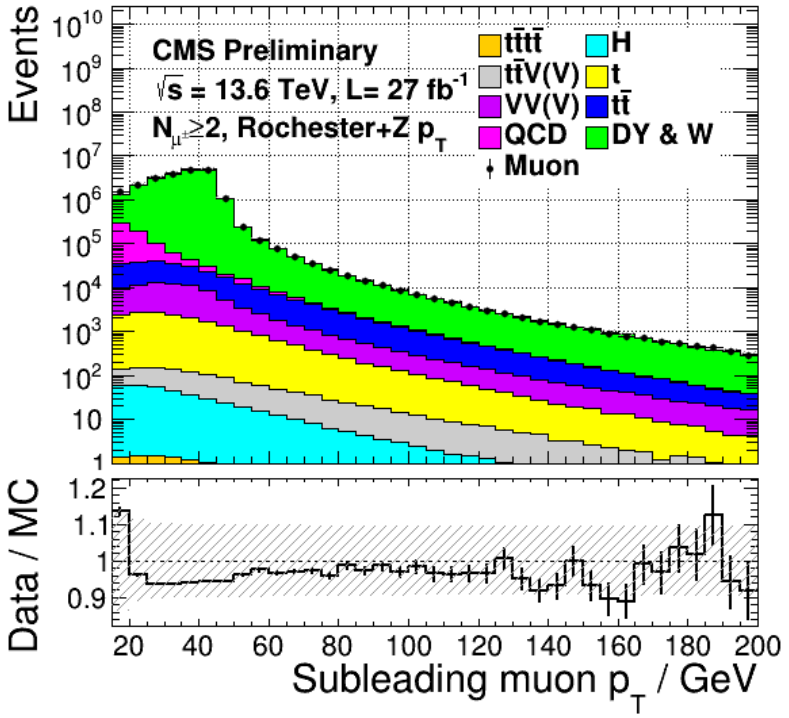
- ϕ distribution almost flat.
- View the normalization difference more clearly.

Trigger for data-MC unmatching in Run2023

Only di-muon trigger



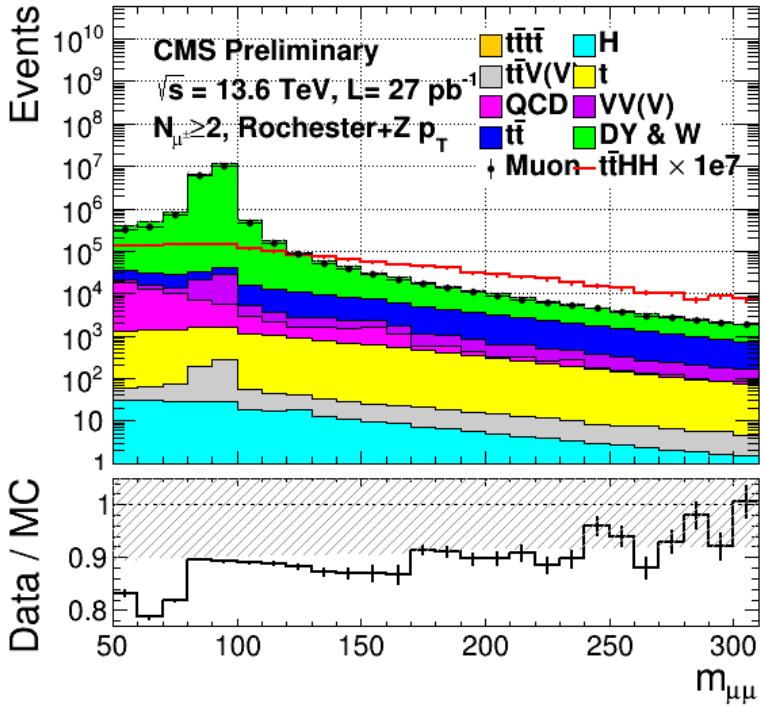
All trigger path



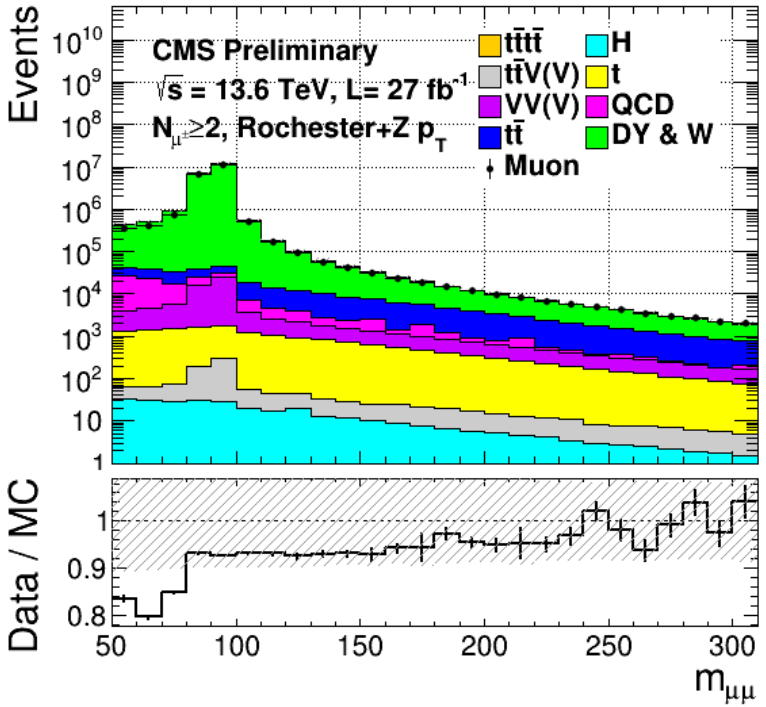
- Subleading Muon p_T distribution similar to the leading Muon.

Run2023 di-Muon

Only di-muon trigger



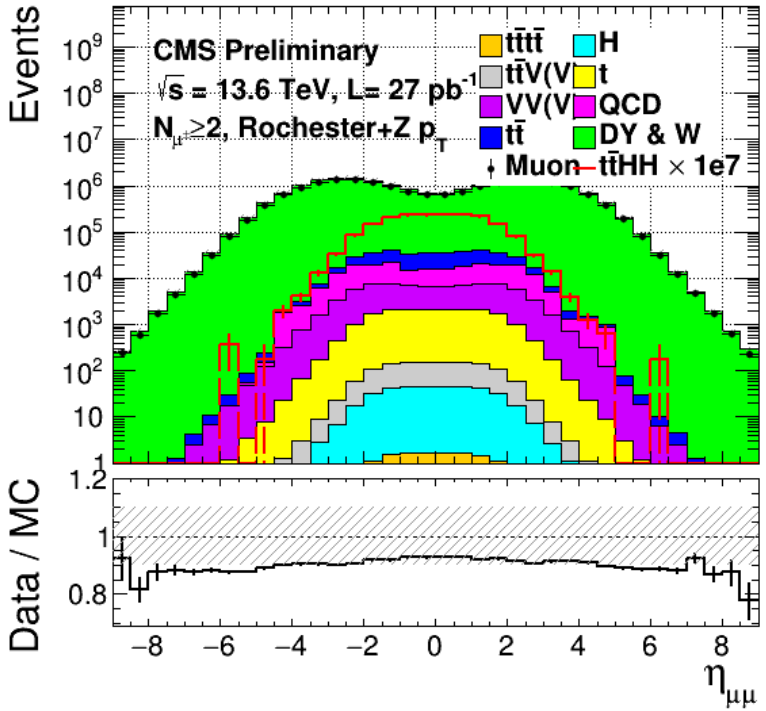
All trigger path



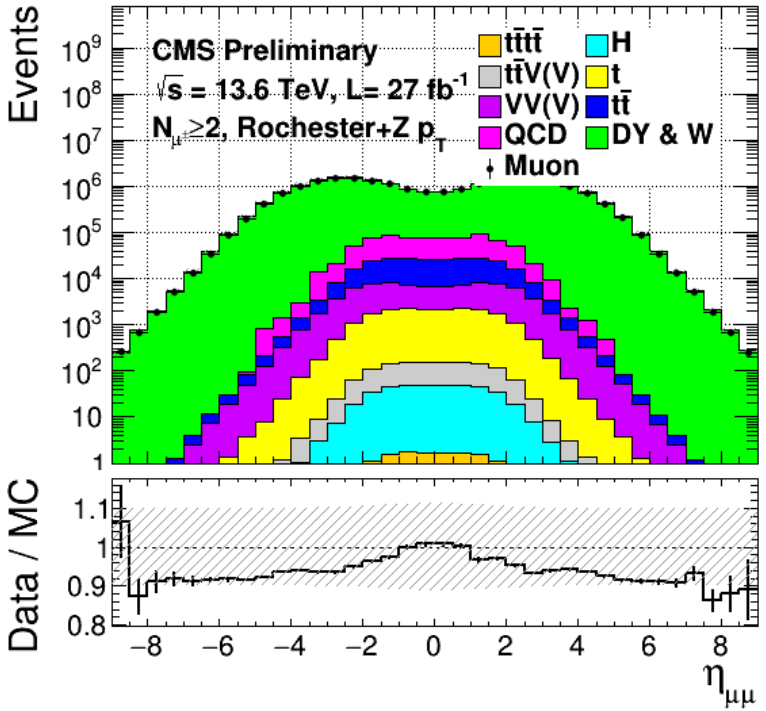
- $m_{\mu\mu}$ also have better yield agreement.
- Lower region has a gap.

Run2023 di-Muon variables

Only di-muon trigger



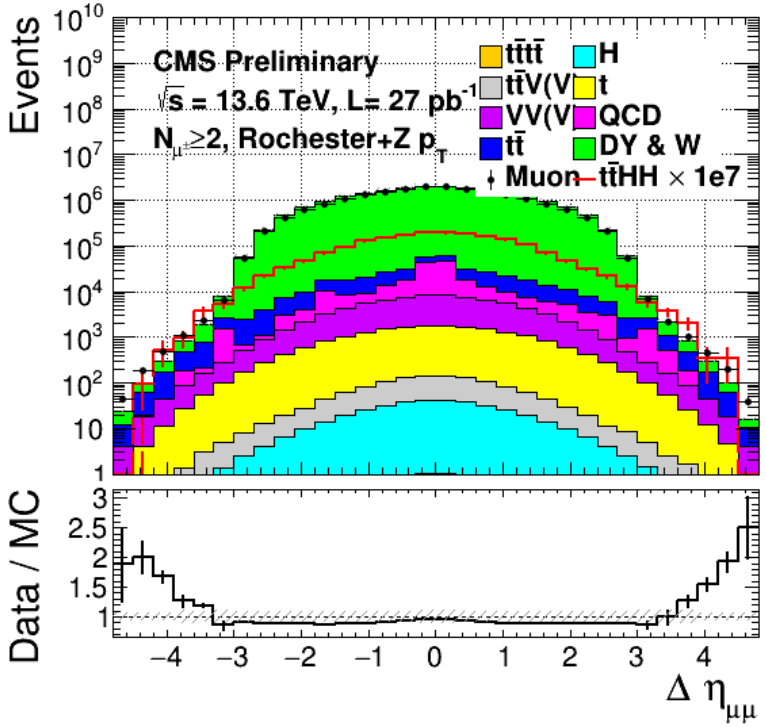
All trigger path



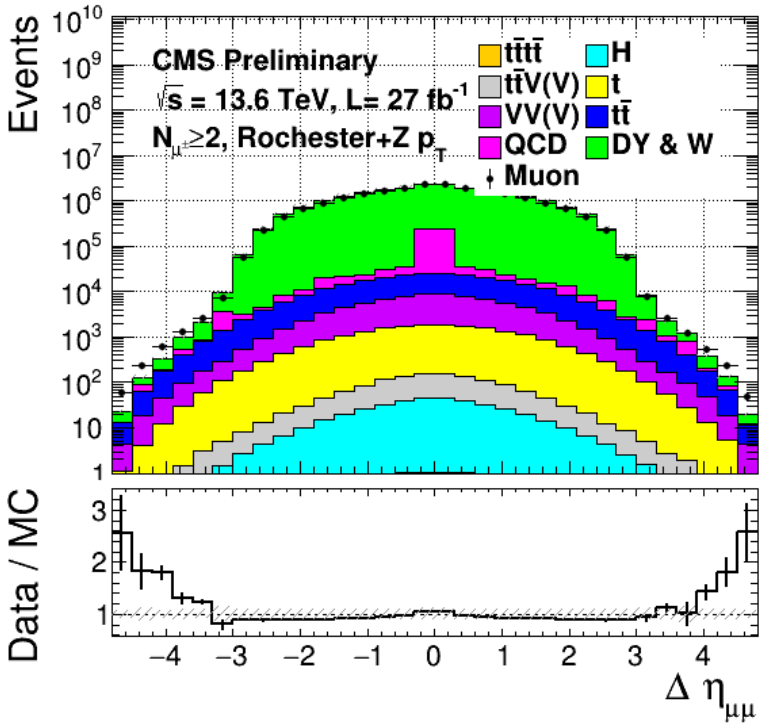
- Dimuon η matching better for the normalization.
- The central part shows up.

Run2023 di-Muon variables

Only di-muon trigger



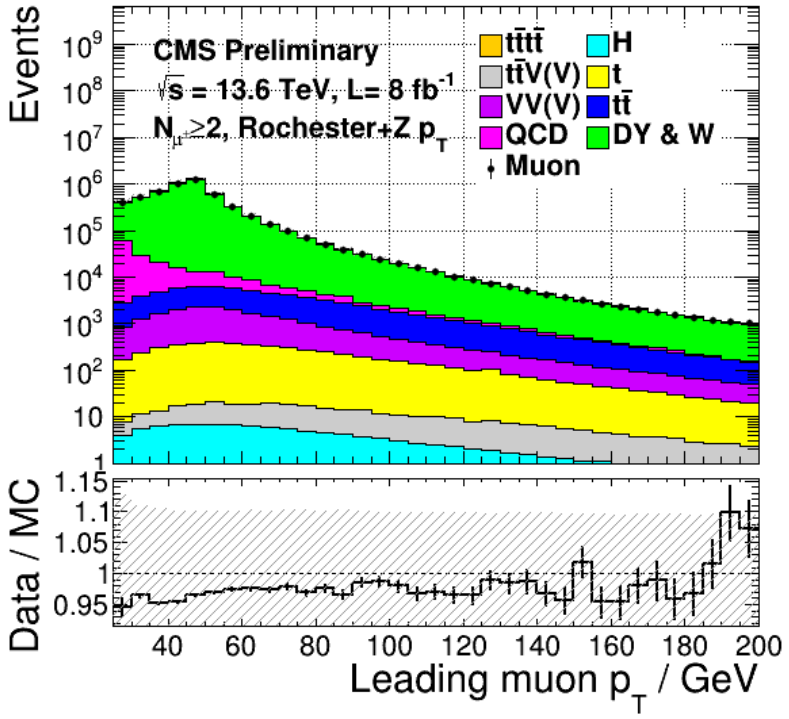
All trigger path



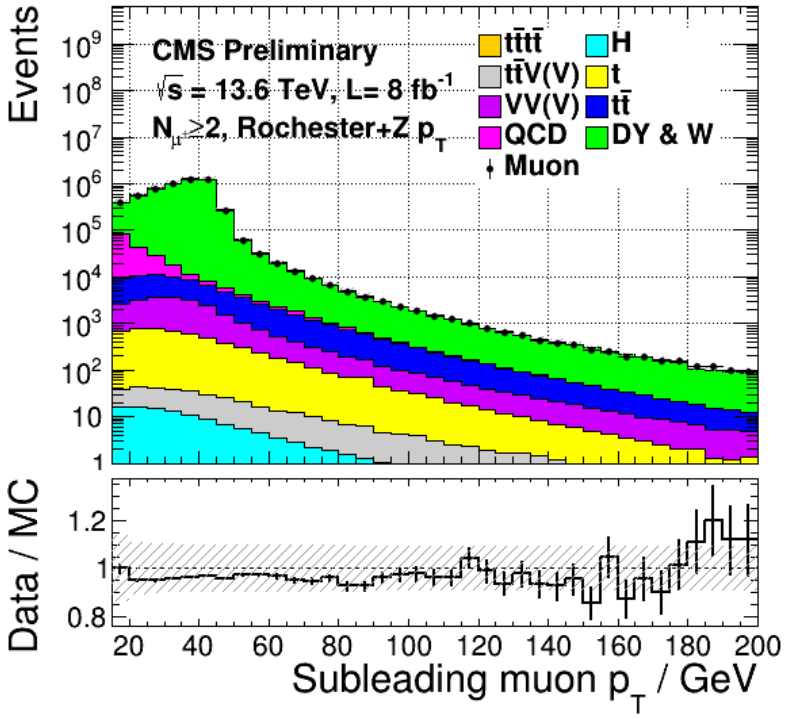
- Cannot solve the $\Delta\eta_{\mu\mu}$ gap in the large value region.

Run2023C adding 10-50 DY sample

Leading Muon

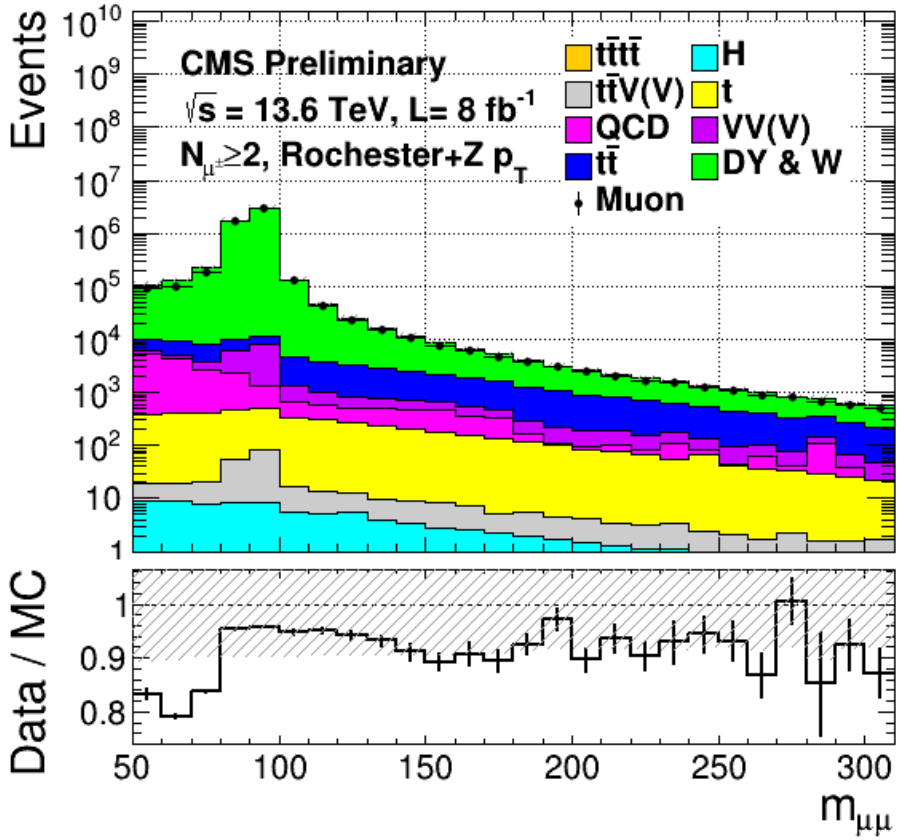


Subleading Muon



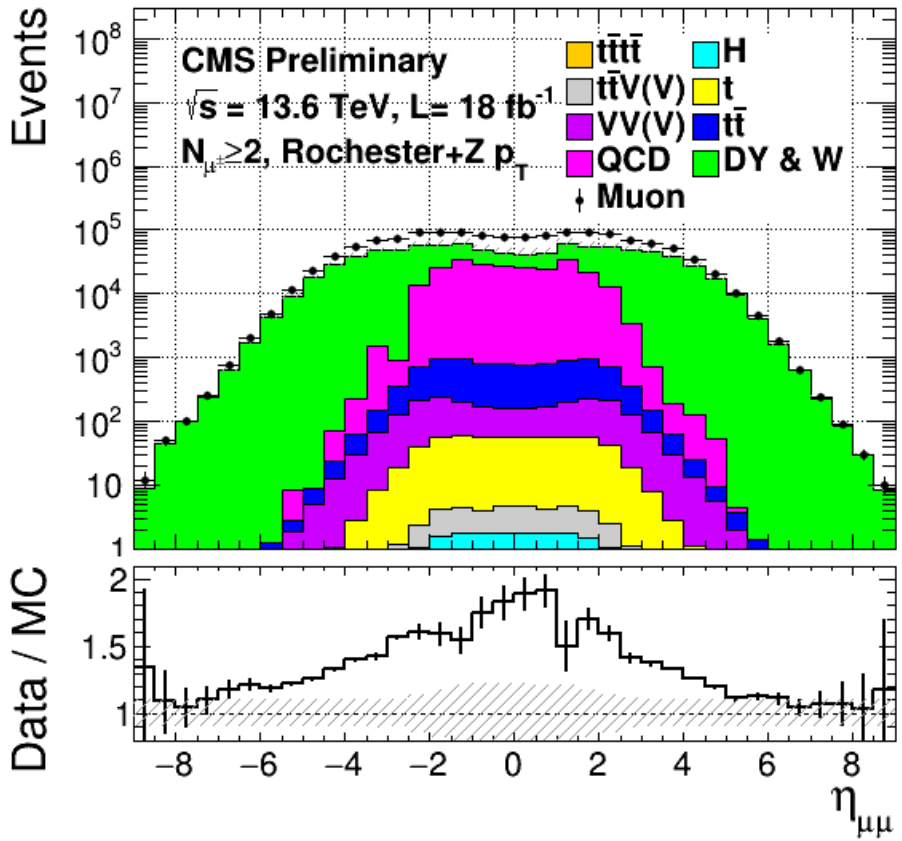
- The lower p_T gap is fixed by adding the DY low mass sample.

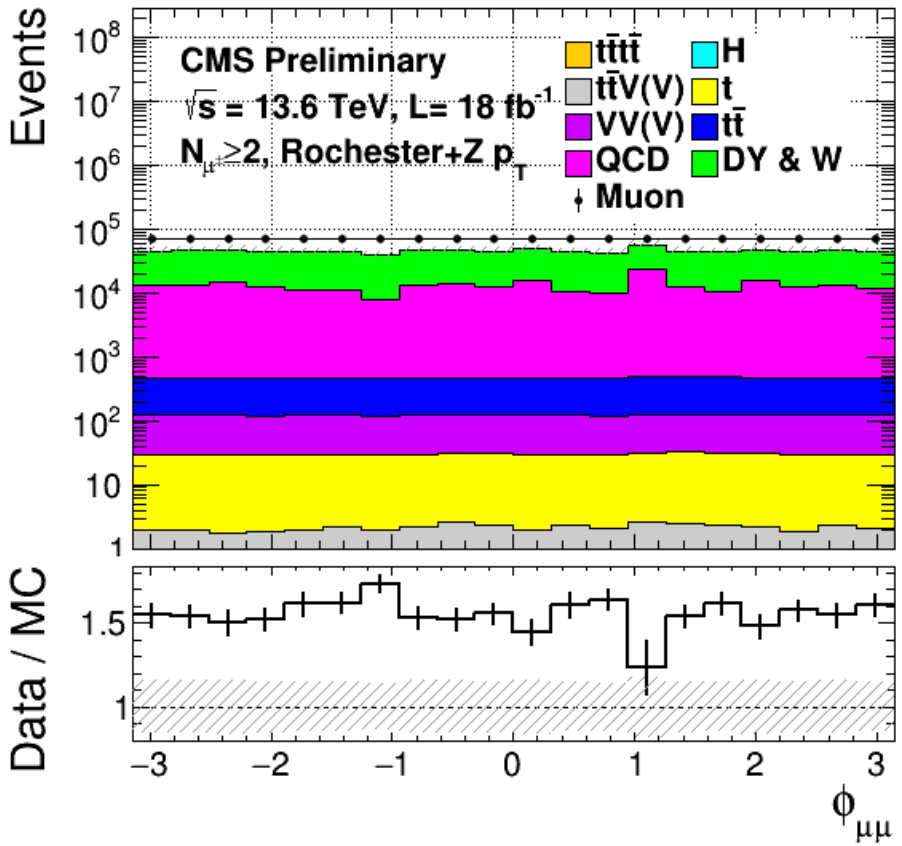
Run2023C adding 10-50 DY sample



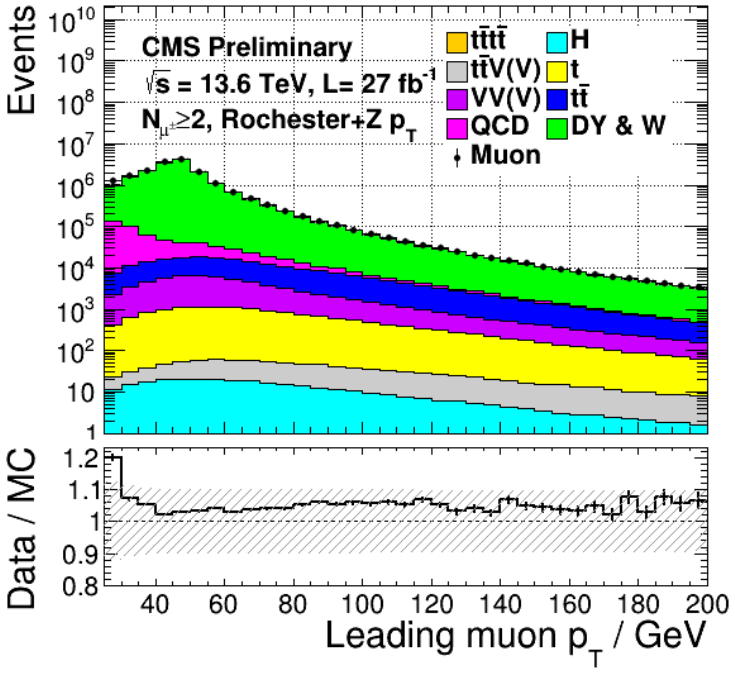
- The lower p_T gap is fixed by adding the DY low mass sample.

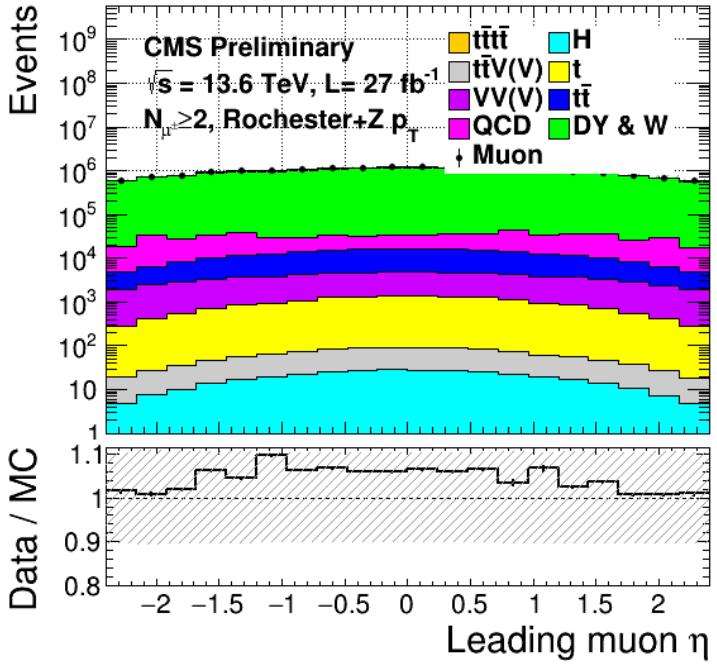
Single Muon

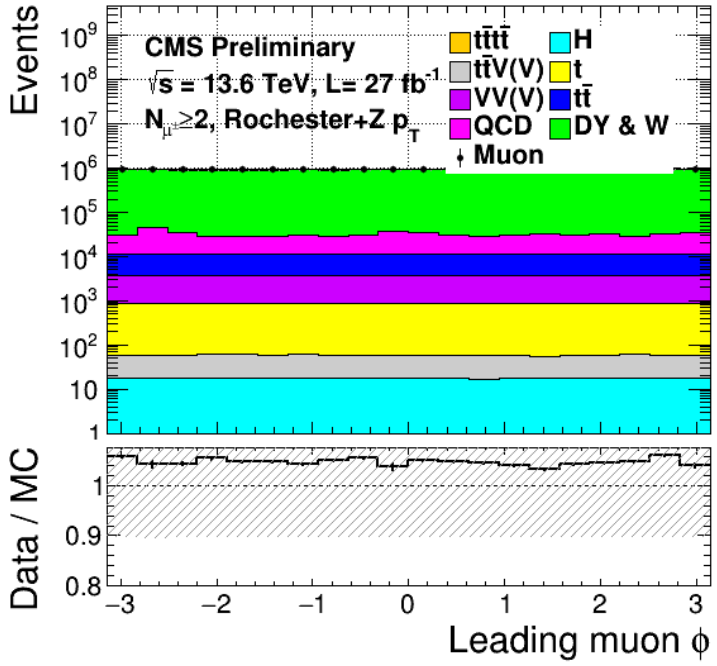

CMS Preliminary
√s = 13.6 TeV, L= 18 fb-1
Nμ≥2, Rochester+Z pT
tttt
tt̄V(V)
VV(V)
QCD
H
t
tt̄
DY & W
Muon
Events
Data / MC
ημμ


CMS Preliminary
√s = 13.6 TeV, L= 18 fb-1
Nμ≥2, Rochester+Z pT
tttt
tt̄V(V)
VV(V)
QCD
H
t
tt̄
DY & W
Muon
Events
Data / MC
ϕμμ

Run2022 Leading Muon variables



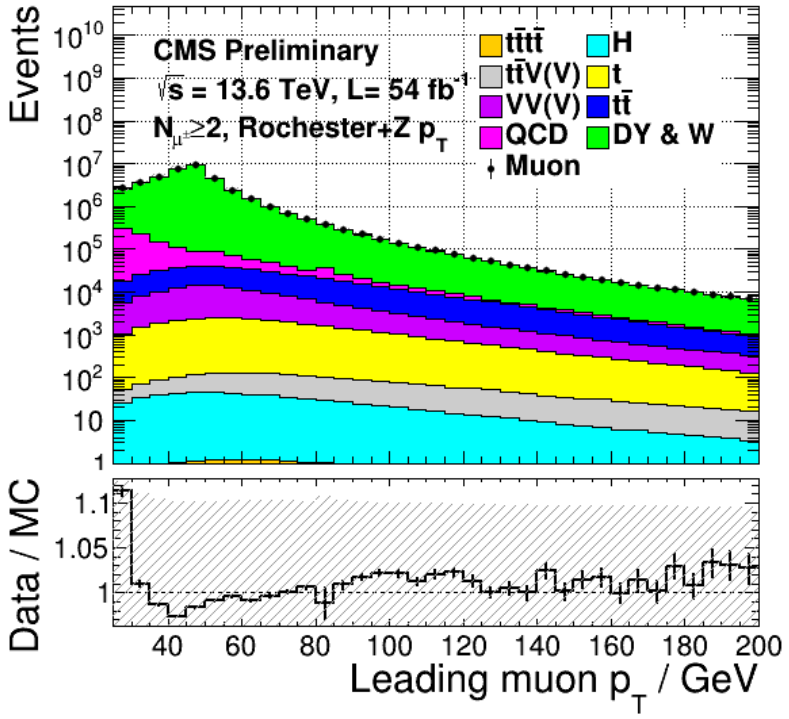




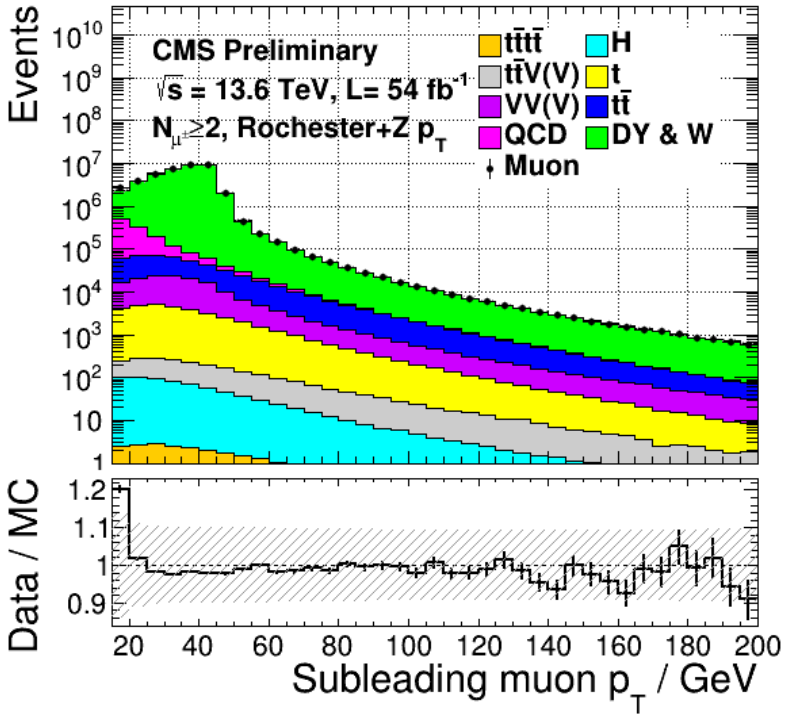
- Data: Run2022E/F/G
- MC: Run3Summer22EENanoAODv12.
- The 22EE comparison has more data than MC as seen in normalization.

Run2022EE + Run2023

Leading muon



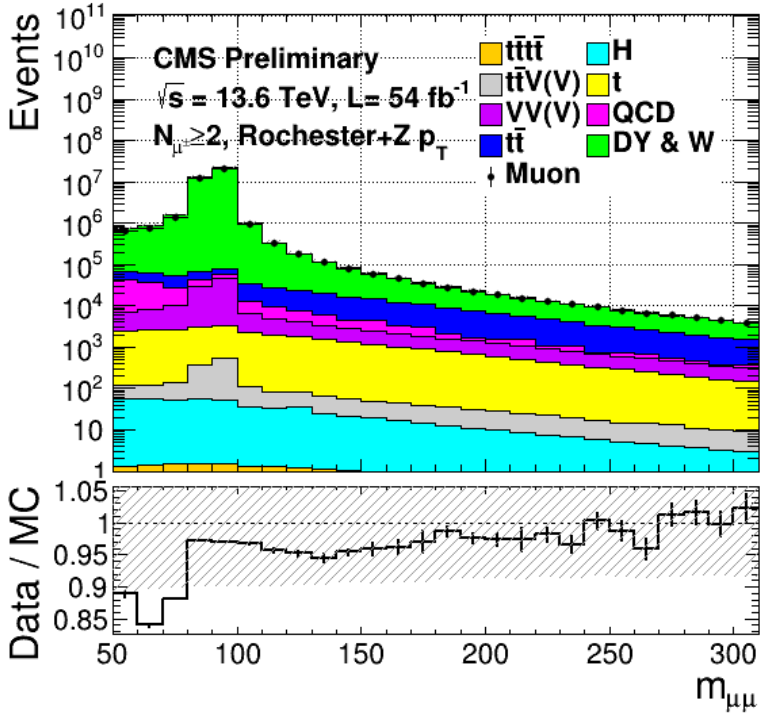
Subleading muon



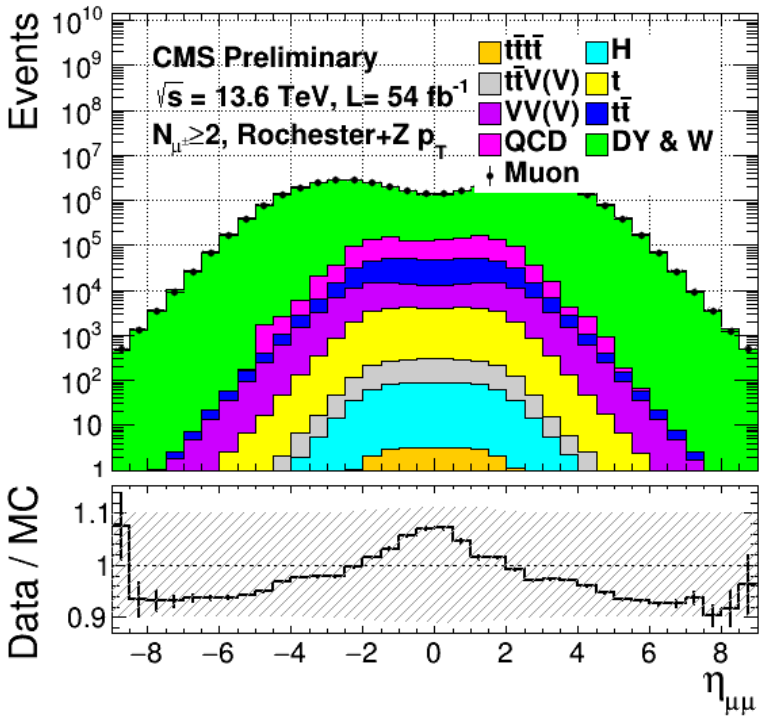
- Add 2022EE and Run2023 together.
- Cancellation in different periods.

Run2022EE + Run2023

Dimuon mass



Dimuon eta



- Add 2022EE and Run2023 together.
- Cancellation in different periods.