

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
- MC sample
 - Drell-Yan samples: DYto2L-4Jets_MLL-50_TuneCP5_13p6TeV_madgraphMLM-pythia8/NANOAODSIM/
 - Cross-section (NNLO value): 6345.99 pb

Event selections

- Dimuon preselection
 - Taking Rochester correction: Muon_pt_Rcorr
 - GoodMuon25: (abs(Muon_eta)<2.4) && Muon_tightId && (Muon_pfRelIso04_all<0.15) && (Muon_pt_Rcorr>25.0)
 - GoodMuon15: (abs(Muon_eta)<2.4) && Muon_tightId && (Muon_pfRelIso04_all<0.15) && (Muon_pt_Rcorr>15.0)
 - nGoodMuon_15 > 1 && nGoodMuon_25 > 0
- Trigger Menu: (HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8 || HLT_IsoMu20 || HLT_IsoMu24 || HLT_IsoMu24_eta2p1 || HLT_IsoMu27 || HLT_Mu50 || HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8 || HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8 || HLT_Mu19_TrkIsoVVL_Mu9_TrkIsoVVL_DZ_Mass3p8 || HLT_Mu19_TrkIsoVVL_Mu9_TrkIsoVVL_DZ_Mass8 || HLT_TripleMu_5_3_3_Mass3p8_DZ || HLT_TripleMu_10_5_5_DZ || HLT_TripleMu_12_10_5)
- MET filter: (Flag_goodVertices && Flag_globalSuperTightHalo2016Filter && Flag_EcalDeadCellTriggerPrimitiveFilter && Flag_BadPFMuonFilter && Flag_eeBadScFilter && Flag_BadPFMuonDzFilter && Flag_hfNoisyHitsFilter)

MC additional corrections

- Pile-up correction
 - From <https://gitlab.cern.ch/cms-nanoAOD/jsonpog-integration.git>, for MC only.
 - This correction will enlarge the total yield of Drell-Yan MC sample, thus making the discrepancy slightly larger.
- Muon scale corrections
 - From <https://gitlab.cern.ch/cms-nanoAOD/jsonpog-integration.git>, for MC only.
 - This correction suppress the DY yield $\sim 1\%$.
- Drell-Yan Z p_T correction
 - From <https://gitlab.cern.ch/cms-higgs-leprare/hleprare.git>, for MC only.

Sample cutflow

- Run2023C raw data (Golden json 17.794 fb-1)

	Golden json	Dimuon pre	+ Trigger	+MET filter	+PU correction	+Muon scale
Muon0	221,904,154	7,124,317	7,034,031	7,022,361		
Muon1	221,871,332	7,122,949	7,032,677	7,021,051		
Drell-Yan	120,765,800.3	14,624,971.1	14,539,953.7	14,538,304.1	14,596,438.5	14,454,755.9
Data / MC	367.47%	97.42%	96.75%	96.60%	96.21%	97.15%

- The corrections do not change yield much.
- The previous discrepancy is mainly caused by only using one dimuon trigger. With more trigger paths included, the discrepancy between data and MC are well-released.

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	

- After pileup and muon scale corrections.
- Data/MC ratio variation is seen. Only around 3% overshooting seen in Run2023C.