Updates on top mass & width measurement

Jiwon Park
DESY top meeting

21 January, 2025

UNTERSTÜTZT VON / SUPPORTED BY
Alexander von HUMBOLDT STIFTUNG

Current progress in a nutshell

Technical developments

- Histograms from data / nominal samples are processed for full run 2 datasets
- Most of systematic uncertainties are implemented
- Combine fit for unfolding and xFitter for parameter extraction has been implemented and tested

Missing items

- bb4l width variations: bb4l v2 gridpack setting is not producing the variation, v1 is being tested
- Several ttbar only uncertainties (pythia8 tune, color reconnection -> to be taken from hvq)

Challenges

- Year incompatibility: As usual, 2016APV (pre) is over 10% lower than the other years in data/mc
- Combine fit with many years: With over 10 POIs, the fit is not working very well
  - Currently working with 2017+2018 as a minimal example

Documentations

- AN is being prepared, in parallel with Valentina's thesis

Samples and object selections

Data samples

- DoubleEG, MuonEG, DoubleMuon, SingleElectron, SingleMuon, and EGamma (2018) MiniAODv2

Simulated samples

- bb4lv2: ttbar+tW+interferences in dilepton channel
  - Normalized to tt + tW dilepton ~ 96.45 pb
- ttbar hvq (l+jets, all jets), single top (s, t-channel) DY (NLO), W+jets (NLO), TTW, TTZ, ttH, VV
- Sample list in the backup

Triggers

- Dilepton + single lepton triggers (backup)

Object selections

- Good PV > 0
- MET filters applied
- Muon: p_T > 25 (20) GeV, $|\eta|$ < 2.4, Iso_{rel} < 0.15, Cut-based ID tight WP
- Electron: p_T > 25 (20) GeV, $|\eta|$ < 2.4, Ecal gap veto, MVA ID wp80
- Dilepton pair: OS leptons, m_{ll} > 20 GeV
  - ee, $\mu\mu$ (SF) channel: veto 76 < m_{ll} < 106 GeV
- Jet: AK4Puppi, p_T > 30 GeV, $|\eta|$ < 2.4
  - b-tagging: DeepJet tight WP
  - Jets are cleaned with respect to the selected leptons
- MET: p_T^{miss} > 40 GeV for ee, $\mu\mu$ pair

Signal definition and event selection

Particle level definition

- Particle level objects constructed by Rivet plugin (twiki, config, NOTE, tZq differential)
- Leptons: Dressed leptons, clustering stable photons around prompt leptons with dR = 0.1,
  - p_T > 25 (20) GeV, $|\eta|$ < 2.4, m_{ll} > 20 GeV, veto [76, 106] GeV for SF channel
- Jet: Clustering stable particles with anti-kt, dR=0.4, excluding prompt leptons and neutrinos. b-jets are identified by the ghost particles
  - p_T > 30 GeV, $|\eta|$ < 2.4
- Signal is defined by the events with exactly two leptons and two b-jets
- $m_{lb}^{\mathrm{minimax}}$ ∈ [0, 40, 60, 80, 100, 120, 140, 160, 180, 220, 270, inf]: Reduced to 11 bins

Event selections: Dilepton selections + Two or more jets, including exactly two b-jets

Event categorization: bb4l, other ttbar (l+jets and all jets), DY, and others (including single top)

Data / MC per year

Too many plots…

Data / MC for Run 2 and each year

- Plots: Link
- Yield table: Link -> With hvq tt + tW cross section, MC is over estimated

Tried ttbar dilepton cross section instead (88.96 pb)

- Plots: Link
- Yield table: Link

Discussion…

- Data/MC ratio is significantly low in 2016APV and a bit high in 2017 (known issues)
- Directly affects simultaneous fit

Systematic uncertainties

Experimental uncertainties

- Pileup, luminosity
- Lepton efficiencies, scale and smearing
- Unclustered energy
- JES (23 sources + HEM)
- JER (η dependent)
- b-tagging
- L1 Prefiring

Background normalization

- ttbar: +4.2 -4.84% (ref: TOP-21-008)
- DY: 30%
- Other: 30%
- Limited statistics: Barlow-Beeston-lite approach

Modeling uncertainties

- Missing higher order (μ_F and μ_R)
- PS scale (ISR and FSR)
- ME-PS matching (h_{damp}): tt only, DCTR method
  - For bb4l, applied to 'events with two tops' only
- NNPDF31 PDF α_S + 100 eigenvectors
- B fragmentation (Bowler-Lund model)
- top quark mass (3 GeV variation scaled by 1/3)

Currently missing

- Pythia 8 tune
- Color reconnection
- Top p_T reweighting to NNLO
- tt hvq + tW template will be used for fit

Systematic uncertainties

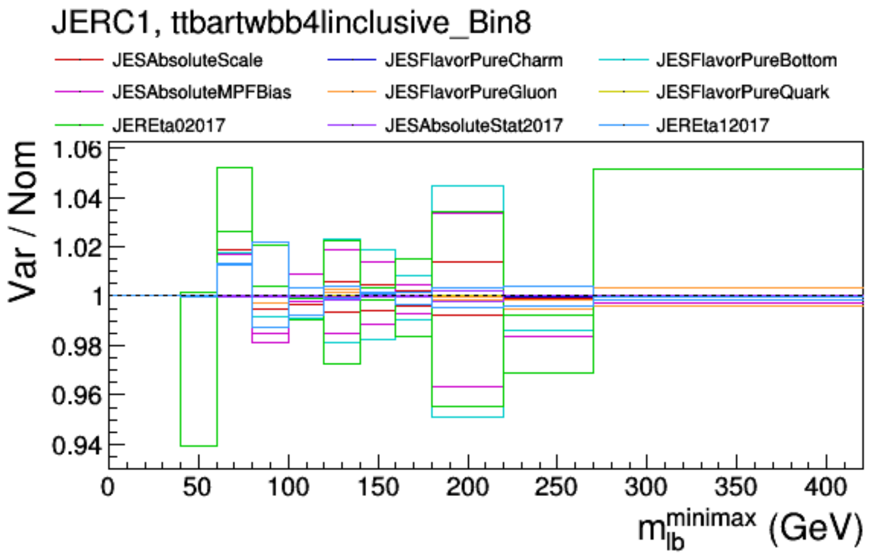
Some templates are further processed

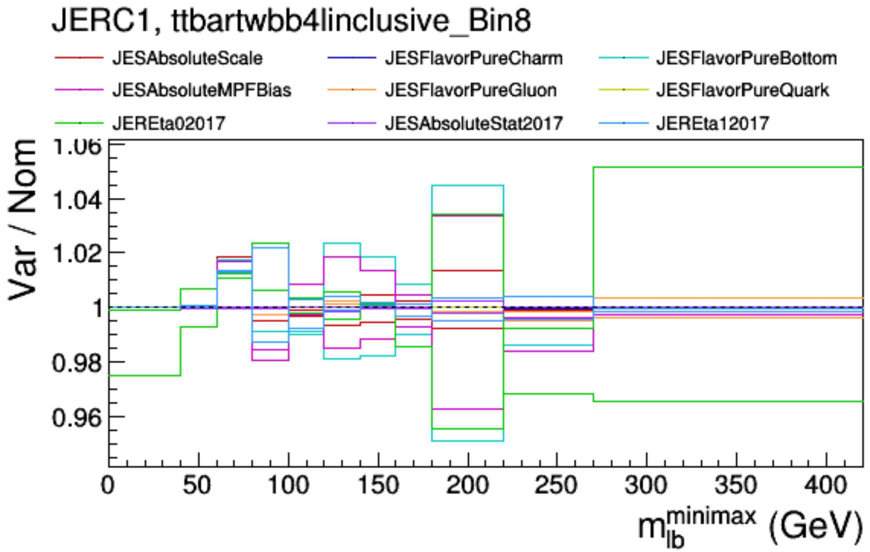
Smoothing

- bb4l Bin 1 (except for mass)
- Electron and muon scale
- FSR
- JER (2017, central eta region)

Smoothing is applied based on flucuation and impacts

- Few more sources need to be treated
- Further rebin?





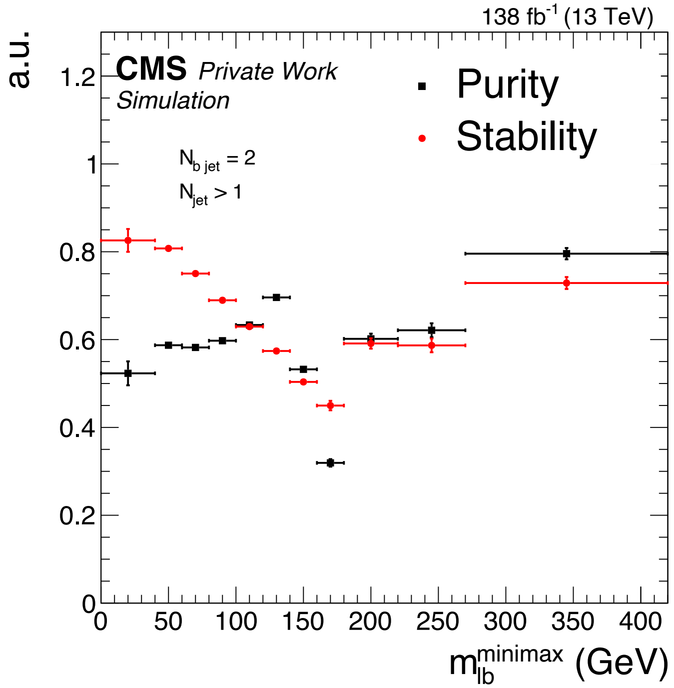
Maximum likelihood unfolding

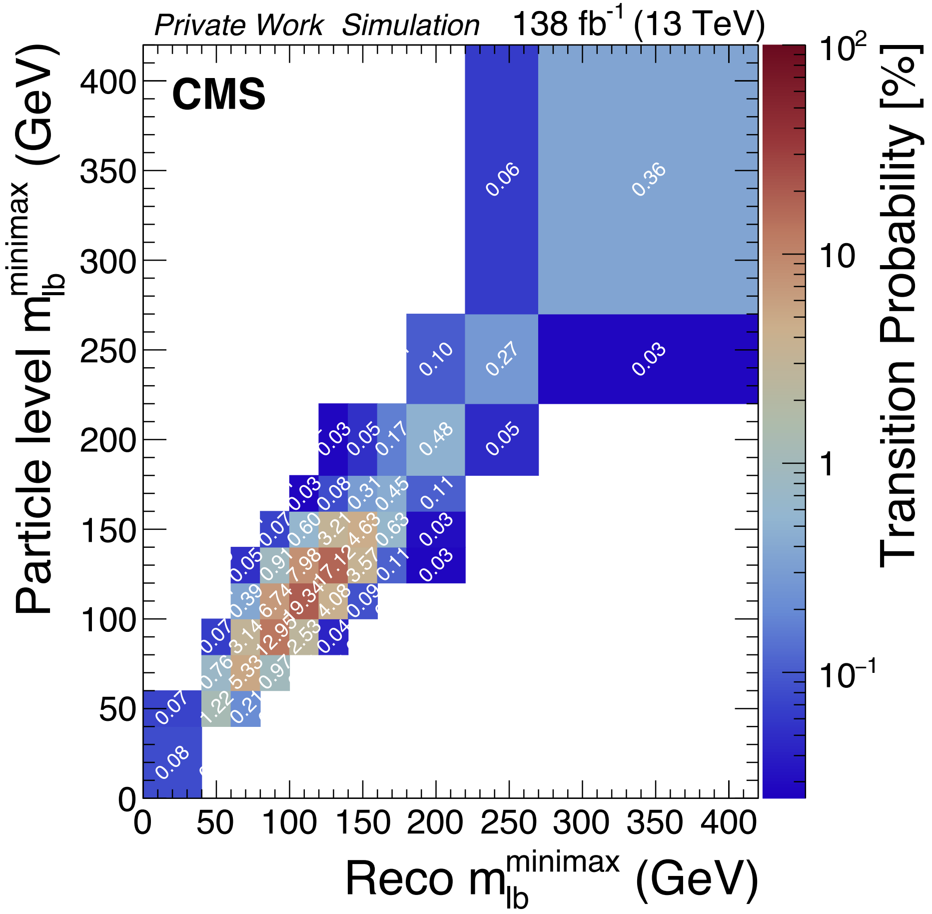
Normalization of each particle-level bin is considered by assigning a signal strength (POI)

- Maximum likelihood fitting calculates bin migration by changing (relative) normalization of particle level bins
- Total 11 POIs for the bb4l particle level bins
- Bin0 normalization: free floating nuisance parameter

Purity & stability

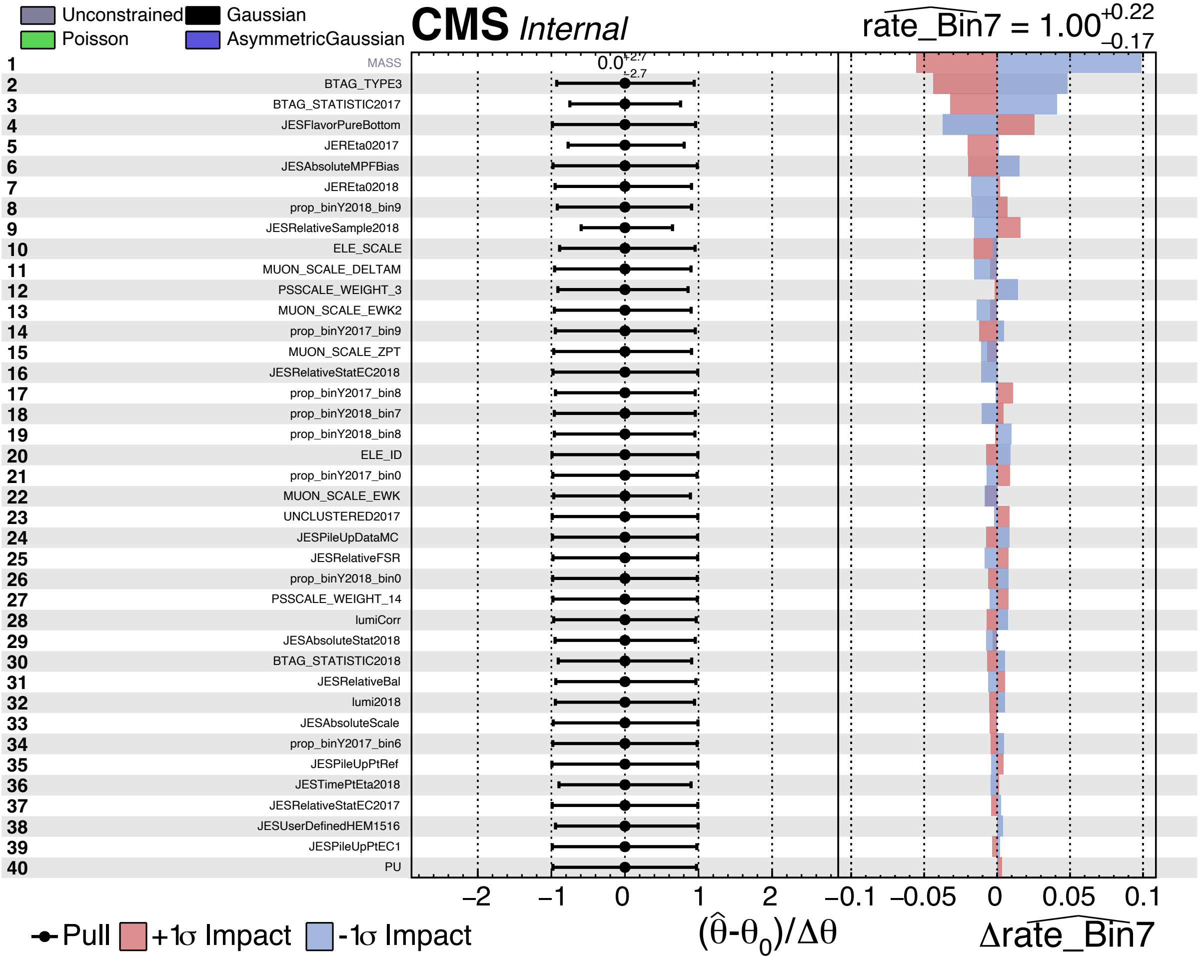
- Purity = N(reco&gen) / N(gen)
- Stability = N(reco&gen) / N(reco)
- Condition number = 7.24

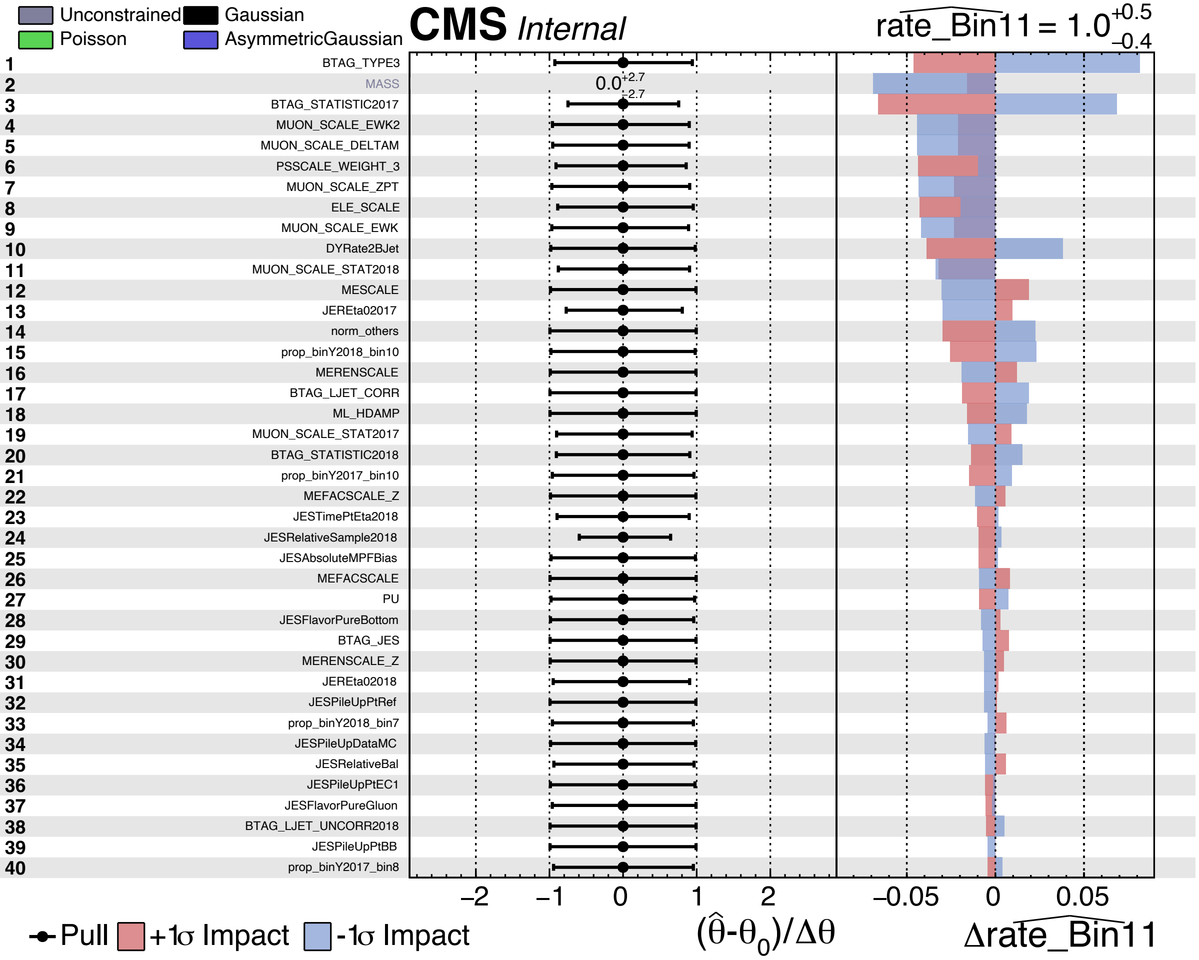




Nuisance parameter impacts and pulls (Asimov)

2017+2018, Bin7 and 11

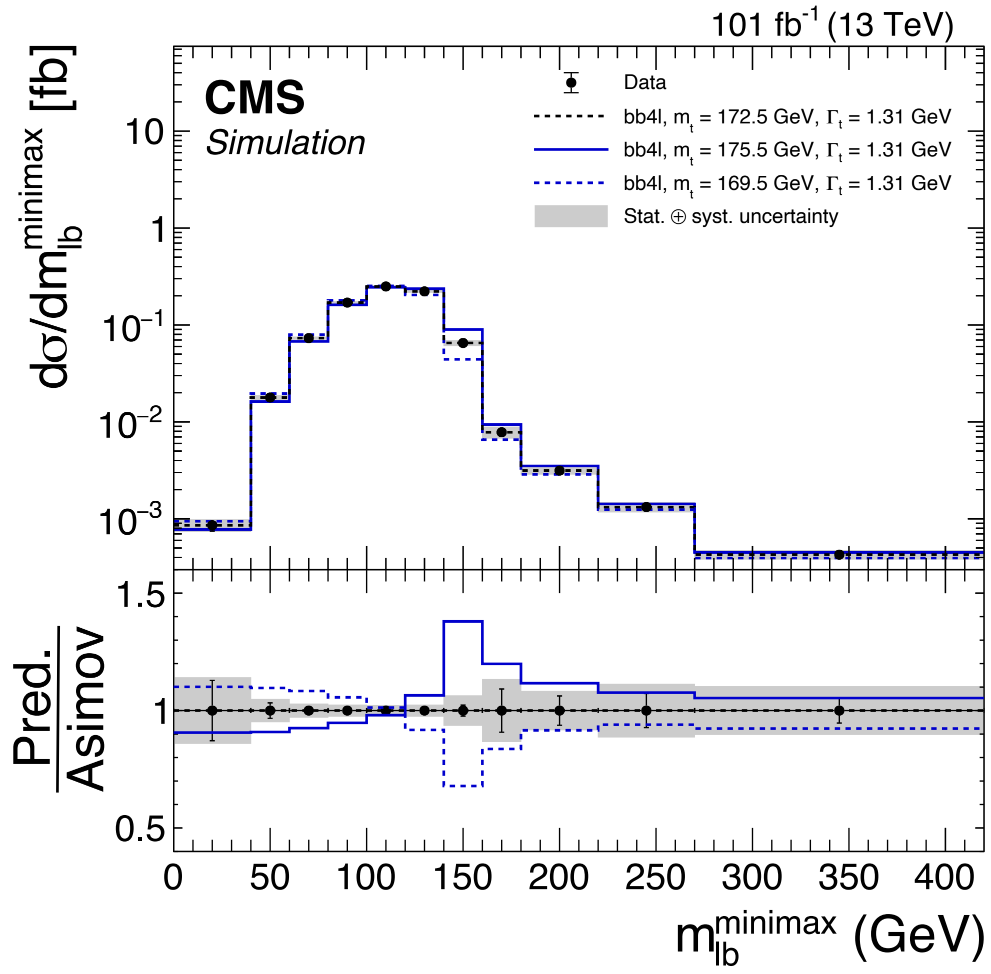


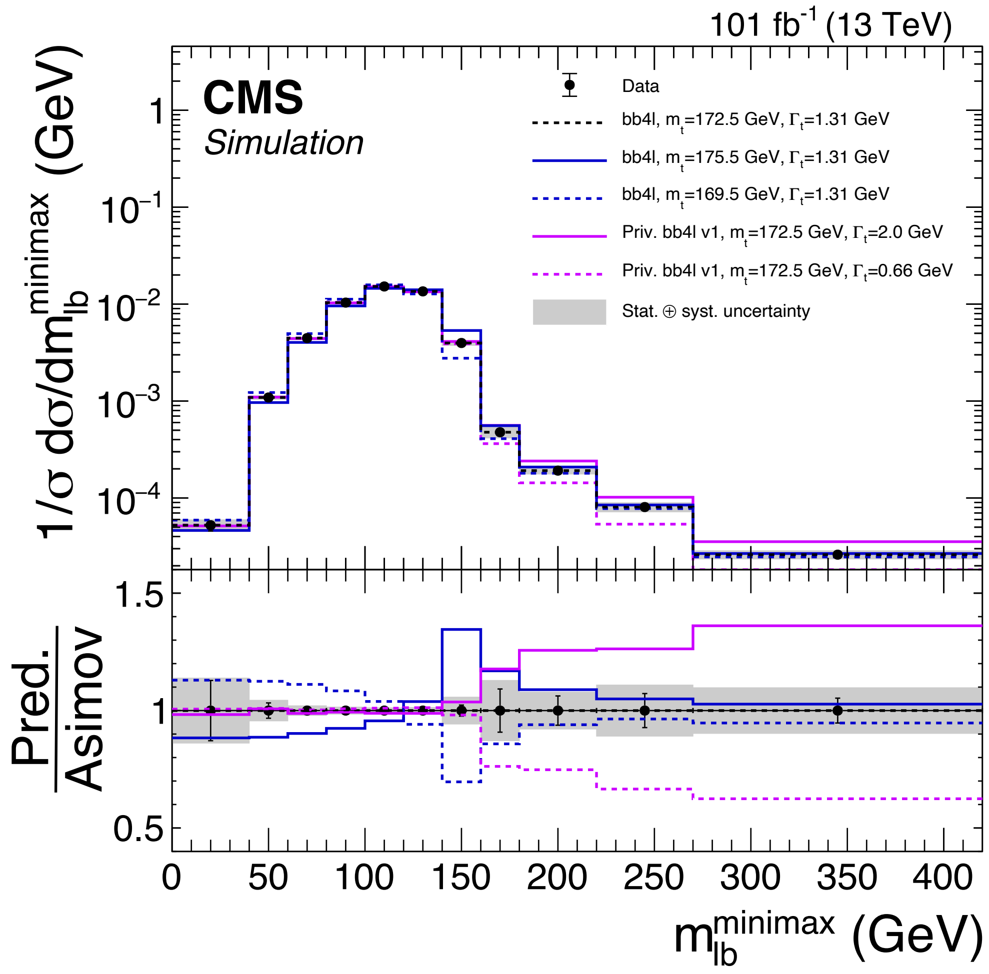


Differential cross section

Absolute and normalized differential cross section, 2017+2018, blinded

- Width variation is available only in the normalized distributions

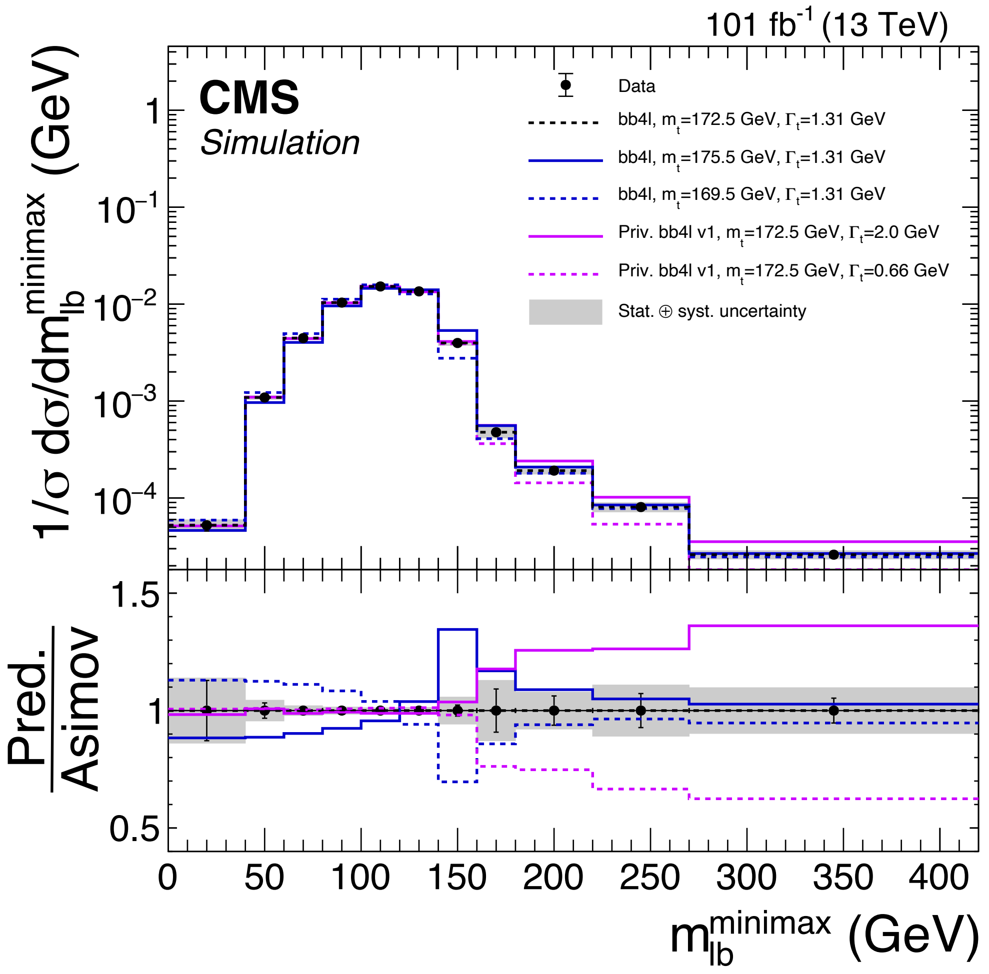


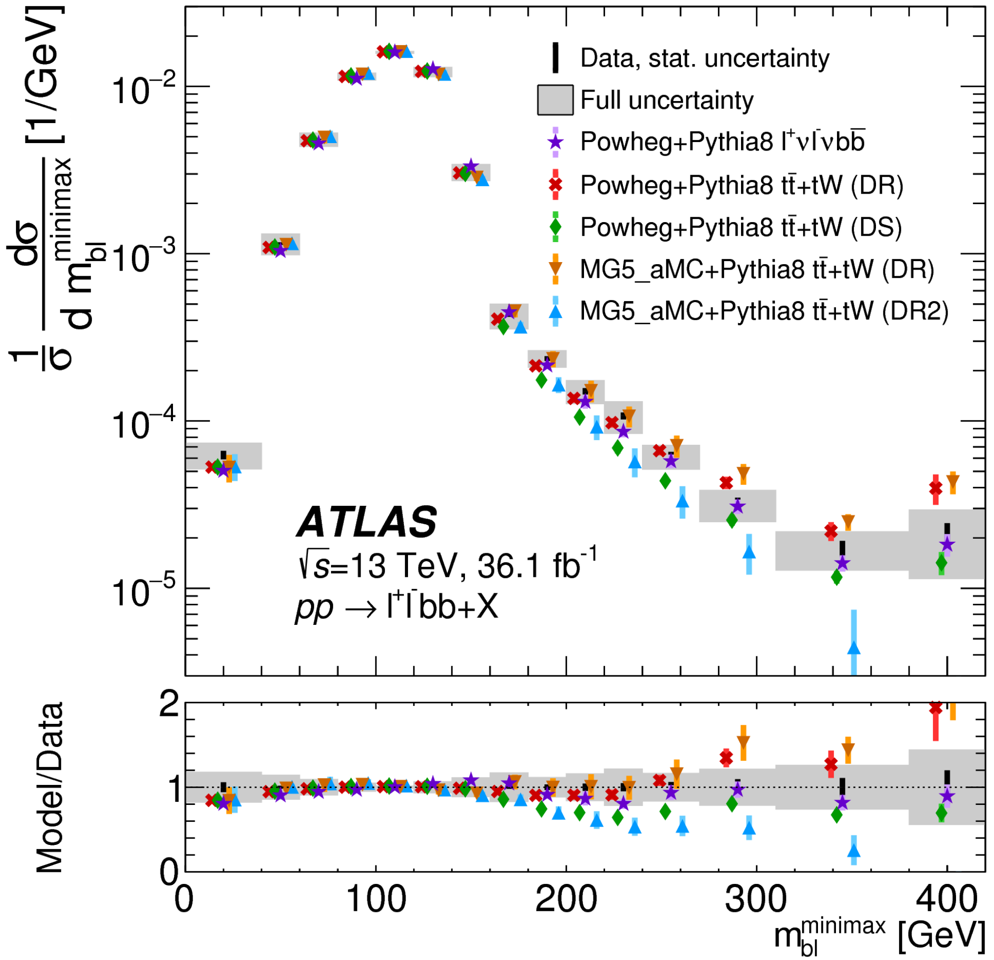


Differential cross section

Normalized differential cross section compared to ATLAS

- Uncertainties in the tail improved significantly

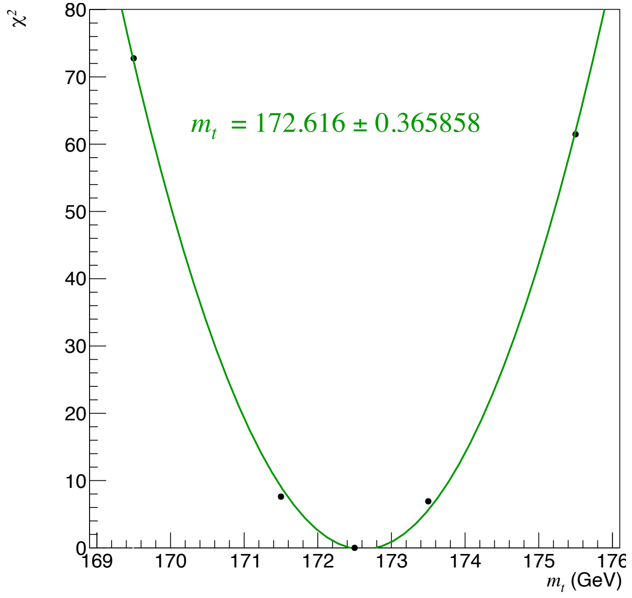


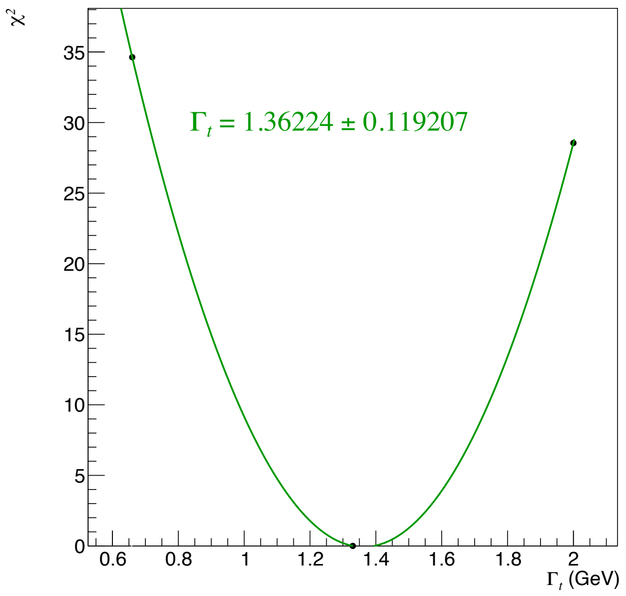


Extraction of mass and width

Utilize xFitter to calculate χ^2 for each mass and width point

- Normalized differential cross section, total correlation matrix, and particle level bb4l predictions are used
- 5 points for mass and width: m_t = [169.5, 171.5 172.5, 173.5, 175.5] GeV and Γ_t = [0.66, 1.32, 2.00] GeV
  - Caveat: Width template is constructed from bb4l v1 private sample, by taking relative variation and applying to bb4l v2
  - Modeling uncertainties on particle level bb4l prediction is not included (scales and PDF)





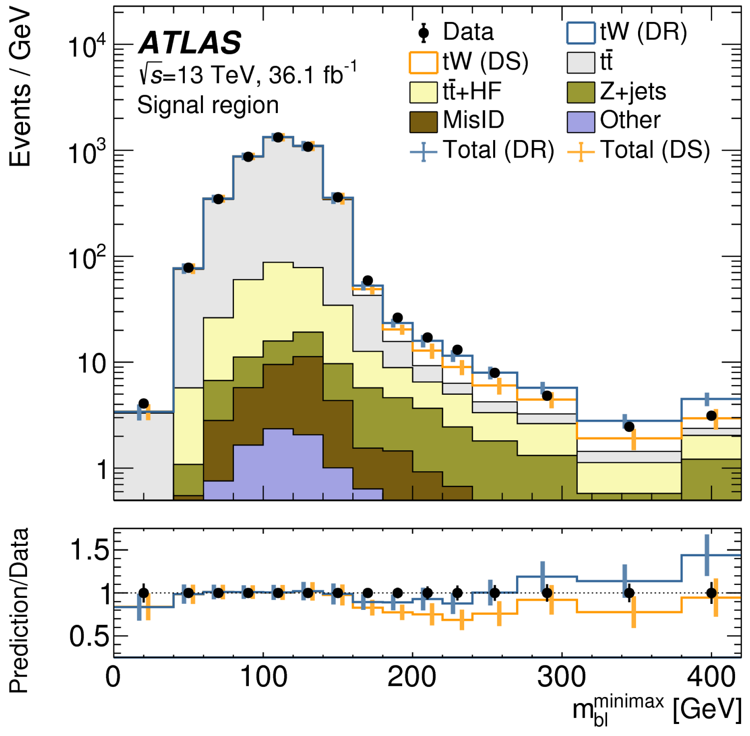
< 10% uncertainty

$$m_{\text{top}} = 172.62 \pm 0.37\,\text{GeV}\ \ (\text{blinded}) \qquad \Gamma_{\text{top}} = 1.36 \pm 0.12\,\text{GeV}\ \ (\text{blinded})$$

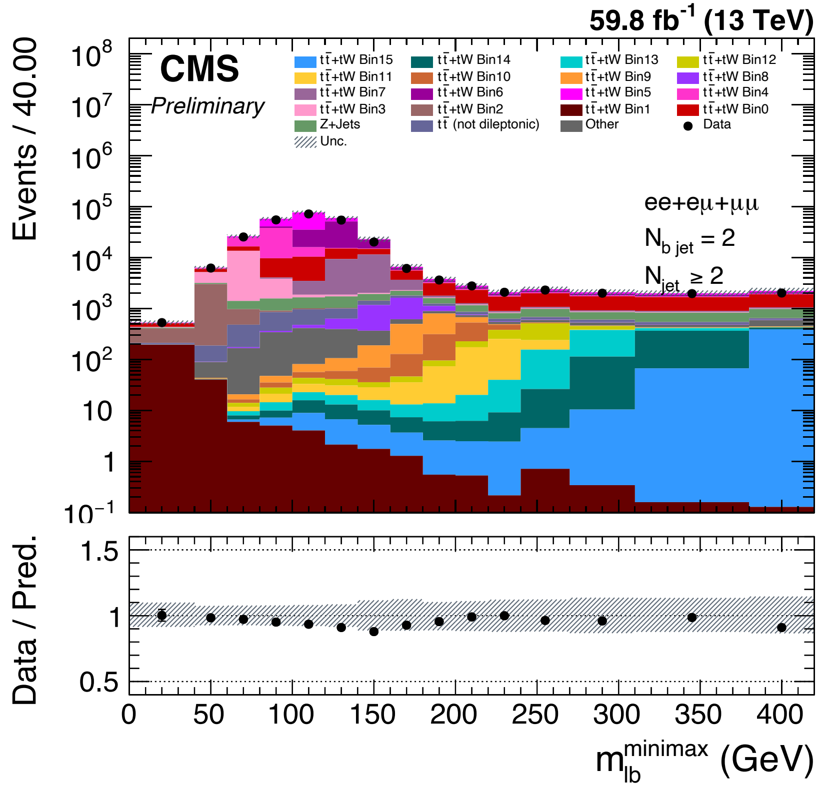
Discussion

If data prefers DR scheme, large deviation is expected

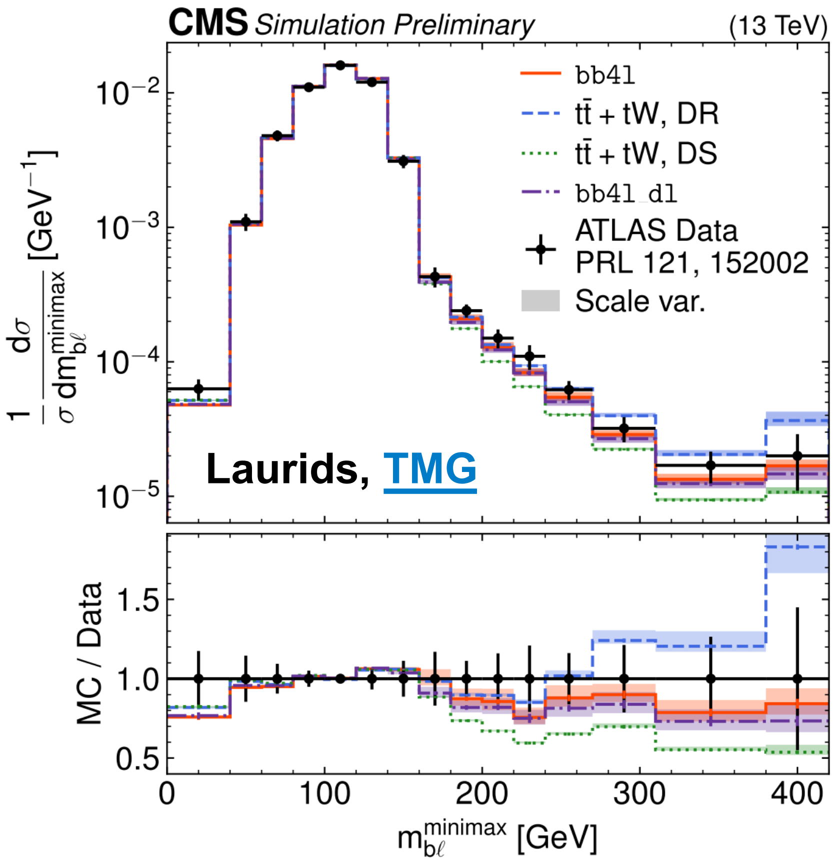
- 20-50% less MC in the tail at particle level



[Phys. Rev. Lett. 121 (2018) 152002](https://doi.org/10.1103/PhysRevLett.121.152002)



tt hvq + tW but loose btag (old plot)
Low purity, not good for unfolding



Discussion

What is missing in the combine fit?

- DR-DS difference was considered in ATLAS unfolding – this can alter tail shape without largely changing POI
  - Need width variation in the fit for shape and residual normalization, as done for mass
- Modeling uncertainty at particle level (around 0.2-0.3 GeV, according to Simone's work)

Year incompatibility

- Oviedo provided Data/MC yield plot, giving around 0.88 (2016APV) and 0.92 (2016) in emu 2b2j
  - We have 0.81 and 0.90 in emu channel (with tt+tW cross section for bb4l)
  - In our framework, MC yield ratio in 2016APV/2016 is similar to lumi ratio, while data ratio is different
  - Differences (Oviedo vs DESY): Nano vs MiniAOD, medium vs tight b-tag, CHS+ JetPUId vs Puppi, n jets
  - We don't include small MCs (VVV or VV NLO) but this should be negligible for data/MC

To do

Short term

- Add uncertainty mimicking width variation in combine fit and check behaviour
- Try medium b-tagging / n jet == 2 selections
- Refine systematic uncertainties

Long term

- To better identify the problem, would it be worth to run a small synchronization test? (maybe using pepper?)
- bb4l width variation is ongoing, aiming for private samples in the end

Thank you

Contact

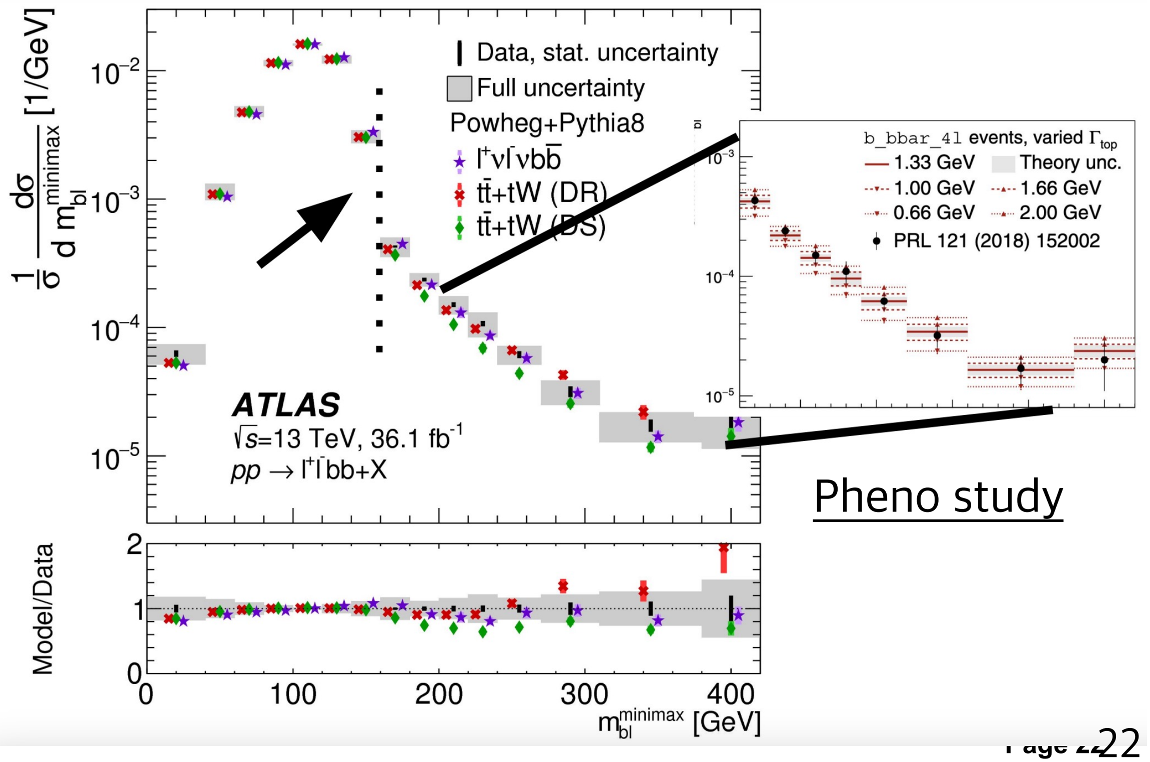
Deutsches Elektronen-
Synchrotron DESY

www.desy.de

Jiwon Park
CMS
jiwon.park@cern.ch

Overview

- Simultaneous extraction of direct top mass and width with Run 2 dataset
  - Past presentations: round table, TMP
- Observable $m_{\ell b}^{\text{minimax}} = \min\{\max(m_{\ell_1 b_1}, m_{\ell_2 b_2}), \max(m_{\ell_2 b_1}, m_{\ell_1 b_2})\}$
- bb4l simulation provides necessary information on the observable
  - m_{lb}^{minimax} tail is sensitive to the top width and is enhanced by tWb diagrams
  - bb4l provides WbWb process including off-shell and tt / tW interference effects
- Analysis strategy
  - Measure differential cross section of m_{lb}^{minimax} with maximum likelihood unfolding
  - Interpretation with bb4l prediction (mass / width var.)
- Personpower (DESY)
  - Jiwon Park, Valentina Guglielmi, Katerina Lipka (DESY)
  - Mass & width extraction is performed by Valentina



Pheno study

Trigger

Year	Dataset	Run range	Trigger selection
2016	$\mu\mu$	273158-284044	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v HLT_IsoMu24_v HLT_IsoTkMu24_v
	ee	273158-284044	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_DoubleEle33_CaloIdL_MW_v
		273158-278822	HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_v
		273158-284044	HLT_Ele27_WPTight_Gsf_v HLT_Ele25_eta2p1_WPTight_Gsf_v HLT_Photon175_v
	$e\mu$	273158-280385	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v
		278273-284044	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v
		273158-284044	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v HLT_IsoMu24_v HLT_IsoTkMu24_v HLT_Ele27_WPTight_Gsf_v HLT_Ele25_eta2p1_WPTight_Gsf_v HLT_Photon175_v
2017	$\mu\mu$	299368-306460	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8_v
		297050-306460	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8_v HLT_IsoMu27_v
	ee	297050-306460	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v HLT_DoubleEle33_CaloIdL_MW_v HLT_Ele32_WPTight_Gsf_L1DoubleEG_v HLT_Ele35_WPTight_Gsf_v
	$e\mu$	299368-306460	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v
		2297050-306460	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_IsoMu27_v HLT_Ele32_WPTight_Gsf_L1DoubleEG_v HLT_Ele35_WPTight_Gsf_v
2018	$\mu\mu$	315257-325172	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8_v HLT_IsoMu24_v
	ee	315257-325172	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v HLT_DoubleEle25_CaloIdL_MW_v HLT_Ele32_WPTight_Gsf_v
	$e\mu$	315257-325172	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_IsoMu24_v HLT_Ele32_WPTight_Gsf_v

MC samples

Simulated samples	σ [pb]
BBLLNuNu_TuneCP5_13TeV-powheg-pythia8	96.447
TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	366.5
TTToHadronic_TuneCP5_13TeV-powheg-pythia8	377.6
DYJetsToLL_M-10to50_TuneCP5_13TeV-amcatnloFXFX-pythia8	18610.1
DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8	6077.22
DYJetsToLL_0J_TuneCP5_13TeV-amcatnloFXFX-pythia8	4826
DYJetsToLL_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8	905.4
DYJetsToLL_2J_TuneCP5_13TeV-amcatnloFXFX-pythia8	347.4
ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8	10.32*0.3258
ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	134.02
ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	80.0
TTWJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
TTWJetsToQQ_TuneCP5_13TeV-amcatnloFXFX-madspin-pythia8	0.4062
TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8	0.2529
TTZToQQ_TuneCP5_13TeV-amcatnlo-pythia8	0.5297
ttHTobb_M125_TuneCP5_13TeV-powheg-pythia8	0.5269
ttHToNonbb_M125_TuneCP5_13TeV-powheg-pythia8	0.5638
WW_TuneCP5_13TeV-pythia8	118.7
WZ_TuneCP5_13TeV-pythia8	46.75
ZZ_TuneCP5_13TeV-pythia8	16.91
WJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-pythia8	61526
WJetsToLNu_0J_TuneCP5_13TeV-amcatnloFXFX-pythia8	49063
WJetsToLNu_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8	9141
WJetsToLNu_2J_TuneCP5_13TeV-amcatnloFXFX-pythia8	3505