Trigger studies with UL2016 pre/postVFP & ttbarRes datacard

Huiling Hua [1] **Fabio Iemmi** [1] Duncan Leggat [2]
Hongbo Liao [1] Hideki Okawa [2] Yu Zhang [2]

[1]Institute of High Energy Physics (IHEP), Beijing

[2]Fudan University, Shanghai

March 9, 2022

4tops

Introduction

- Goal is to **repeat previous trigger studies**
- Hopefully no major changes when switching to UL
- **UL2016 pre/postVFP** are downloaded and ready $\Longrightarrow$ start from them
- Today's **plots implement missing items from last time**
- **Apply** needed **SFs**:
  - Muon ID
  - **No ele ID**: select events with $== 1$ leptons, $== 1$ muon
  - Tau ID
  - b tagging (**FixedWP scale factors**)
  - Evt weight, PU weight, prefiring weight
  - **ttbb correction**

Trigger efficiency studies

- **Compare** results for **data and MC**, **extract** trigger **SF** if needed
- Trigger efficiency definition:

$$\varepsilon(v) = \frac{\mathrm{N}_{\mathrm{trig+presel}}}{\mathrm{N}_{\mathrm{presel}}}(v)$$

- **N.B.**: **in data, we never have all the events that pass the offline preselection**
- In data, events are **always** collected with a trigger
  - In other words, denominator meaningless for data

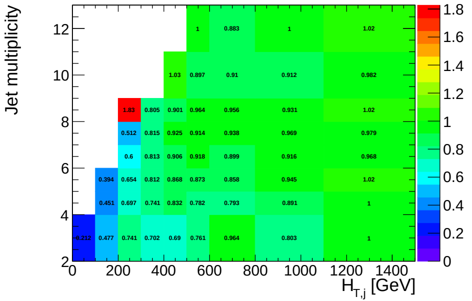
Trigger efficiency studies

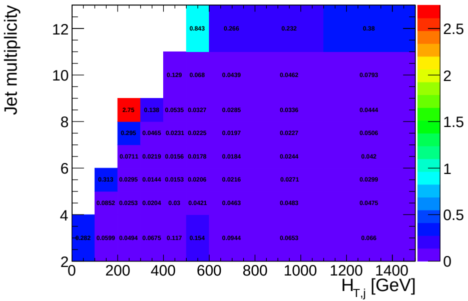
- We need an **unbiased sample of events**
- This should be **collected with a reference trigger** with looser and (if possible) orthogonal criteria
- Then the efficiency definition becomes

$$\varepsilon(v) = \frac{\mathrm{N}_{\mathrm{trig+presel+reference}}}{\mathrm{N}_{\mathrm{presel+referece}}}(v)$$

  which makes sense for data as well
- Obviously the **reference should be unbiased**, i.e., should not change MC efficiency distribution

Pre-UL results

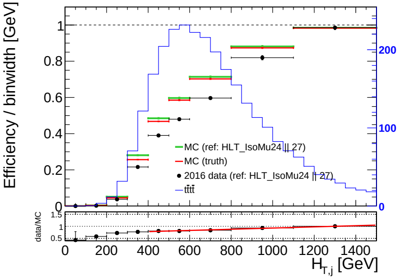


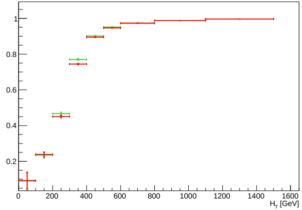


- nJets vs H_T trigger efficiency
- Left: data/MC efficiency ratio; right: corresponding errors
- **Add $H_T > 400$ GeV cut to analysis selection** to make trigger efficient
- Use these histograms as **trigger efficiency scale factors and uncertainties**

UL2016 preVFP

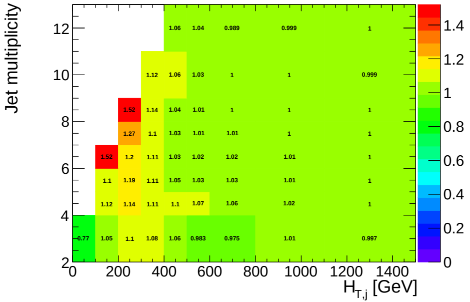
1D trigger efficiency plots

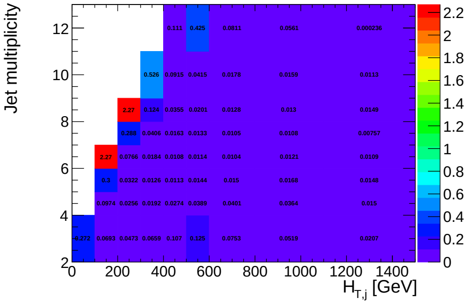




- Trigger choice: unchanged wrt preUL
- OR of:
  - `HLT_PFHT450_SixJet40_BTagCSV_p056`
  - `HLT_PFHT400_SixJet30_DoubleBTagCSV_p056`
  - `HLT_PFJet450`
- Reference triggers: OR of
  - `HLT_IsoMu24`
  - `HLT_IsoMu27`
- Selection:
  - Preselection
  - $== 1\ \ell$, $== 1\ \mu$
  - Designed to have reference firing

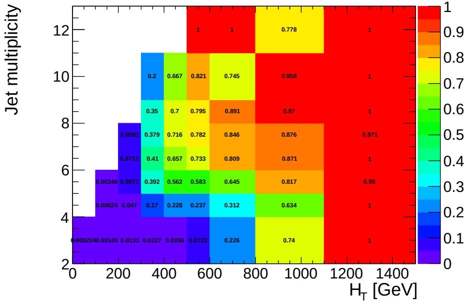
2D plots: checking the reference trigger

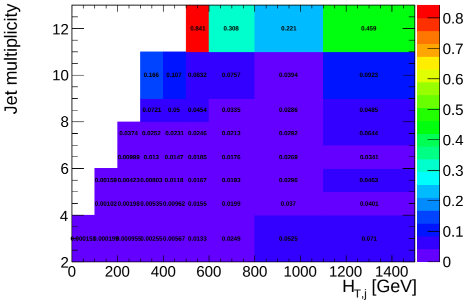




- nJets vs H_T trigger efficiency
- Left: MC/MCtruth efficiency ratio; right: corresponding errors
- The reference is unbiased in 2D too above 400/500 GeV

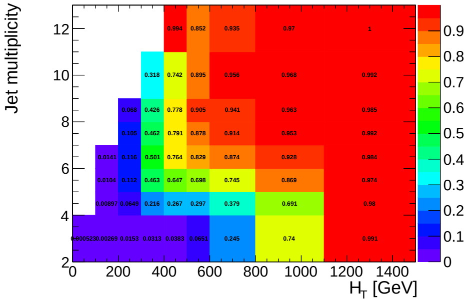
2D plots: trigger efficiency in data

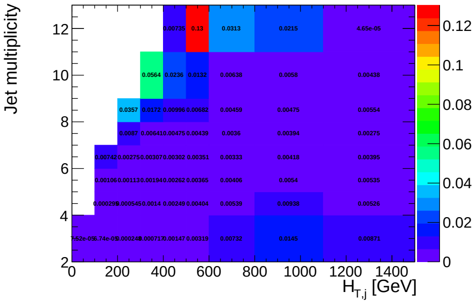




- nJets vs H_T trigger efficiency in data
- Left: trigger efficiency; right: corresponding errors
- Decent increase in efficiency in some bins

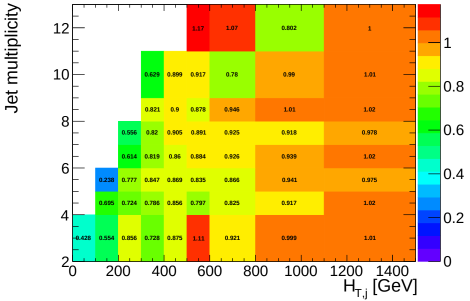
2D plots: trigger efficiency in MC

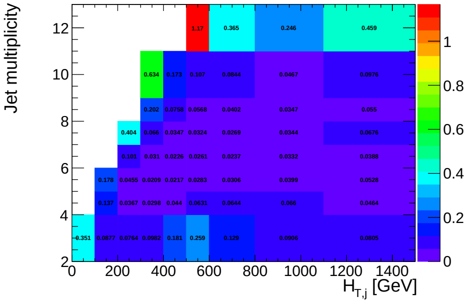




- nJets vs H_T trigger efficiency in data
- Left: trigger efficiency; right: corresponding errors
- Efficiency higher in MC

2D plots: data/MC efficiency ratio

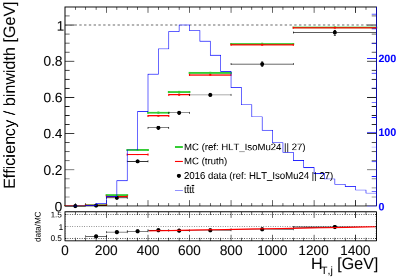


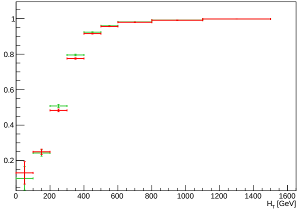


- nJets vs H_T data/MC efficiency ratio
- Left: trigger efficiency; right: corresponding errors
- Not very different from previous results
- Somehow bigger errors and a bit more spiky (lower data stats?)

UL2016 postVFP

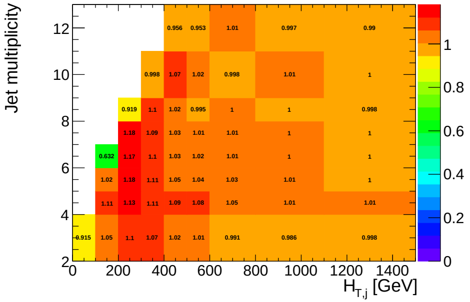
1D trigger efficiency plots

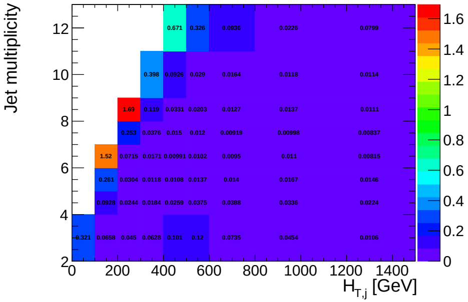




- Trigger choice: unchanged wrt postUL
- OR of:
  - `HLT_PFHT450_SixJet40_BTagCSV_p056`
  - `HLT_PFHT400_SixJet30_DoubleBTagCSV_p056`
  - `HLT_PFJet450`
- Reference triggers: OR of
  - `HLT_IsoMu24`
  - `HLT_IsoMu27`
- Selection:
  - postselection
  - $== 1\ \ell$, $== 1\ \mu$
  - Designed to have reference firing

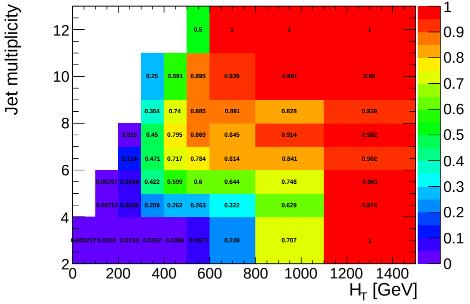
2D plots: checking the reference trigger

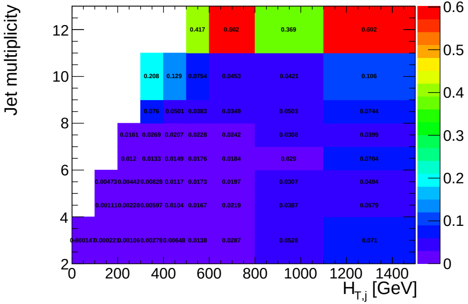




- nJets vs H_T trigger efficiency
- Left: MC/MCtruth efficiency ratio; right: corresponding errors
- The reference is unbiased in 2D too above 400/500 GeV

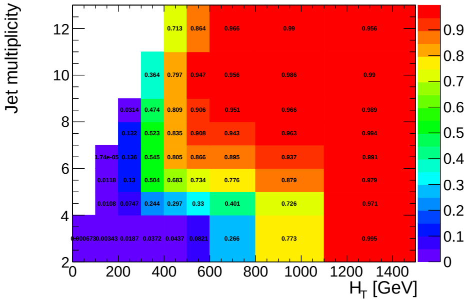
2D plots: trigger efficiency in data

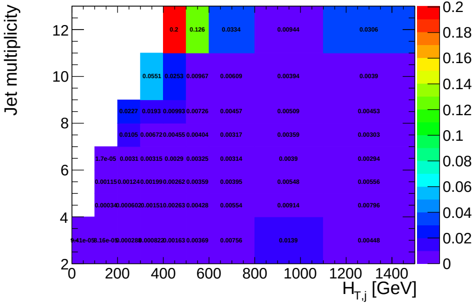




- nJets vs H_T trigger efficiency in data
- Left: trigger efficiency; right: corresponding errors
- Decent increase in efficiency in some bins

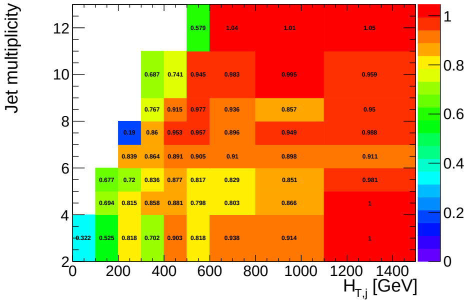
2D plots: trigger efficiency in MC

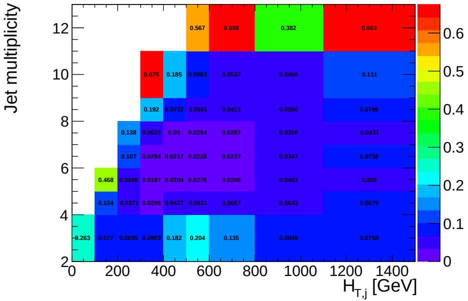




- nJets vs H_T trigger efficiency in data
- Left: trigger efficiency; right: corresponding errors
- Efficiency higher in MC

2D plots: data/MC efficiency ratio





- nJets vs H_T data/MC efficiency ratio
- Left: trigger efficiency; right: corresponding errors
- Not very different from postvious results
- Somehow bigger errors and a bit more spiky (lower data stats?)

Trigger studies in 2018

- 2018 simulation has the following multijet triggers available:
  - HLT_PFHT400_SixPFJet32_DoublePFBTagDeepCSV_2p94
  - HLT_PFHT450_SixPFJet36_PFBTagDeepCSV_1p59
- In 2018 data, two different sets of triggers run, depending on era:
  - 2018 A:
    - HLT_PFHT380_SixPFJet32_DoublePFBTagDeepCSV_2p2
    - HLT_PFHT430_SixPFJet40_PFBTagDeepCSV_1p5
  - 2018 B, C and D:
    - HLT_PFHT400_SixPFJet32_DoublePFBTagDeepCSV_2p94
    - HLT_PFHT450_SixPFJet36_PFBTagDeepCSV_1p59
- **Problem: trigger that run in 2018 A are not emulated in MC**

Trigger studies in 2018

- Not 100% about how to proceed
- An idea (maybe?):
  - When running on MC, ask for the triggers available
  - When running on data, ask for triggers available in different eras
  - When adding together TEfficiency objects for data, **reweight each era for it's integrated luminosity**
  - Extract trigger SF
  - This should take into account the differences in trigger in 2018 A

Summary and next step

- **2016 results** look **similar to** what we saw in **pre-UL**
- Plots **a bit more spiky** and having bigger uncertainties
  - Most likely due to splitting of data samples
- **Plans**:
  - **2017** is ready: submit jobs and **repeat study** there
  - Is 2017 downloaded? Plan to make studies there asap

ttbarRes

Datacard

- Meg provided me with the first nominal templates for ttbar resonance analysis
- I wrote the very first, very preliminary datacard

```
File Edit Options Buffers Tools Help
imax 6 number of bins
jmax * number of processes minus 1
kmax * number of nuisance parameters

shapes * * TemplateShapes.root  $PROCESS$CHANNEL $PROCESS$CHANNEL$SYSTEMATIC

-----------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------

bin             C1      C2      C3      C4      C5      C6
observation     -1      -1      -1      -1      -1      -1

-----------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------

bin             C1      C1      C1      C2      C2      C2      C3      C3      C3      C4      C4      C4      C5      C5      C5      C6      C6      C6
process         RSG1000 QCD     TTbar   RSG1000 QCD     TTbar   RSG1000 QCD     TTbar   RSG1000 QCD     TTbar   RSG1000 QCD     TTbar   RSG1000 QCD     TTbar
process         0       1       2       0       1       2       0       1       2       0       1       2       0       1       2       0       1       2
rate            -1      -1      -1      -1      -1      -1      -1      -1      -1      -1      -1      -1      -1      -1      -1      -1      -1      -1

-----------------------------------------------------------------------------------------------
-----------------------------------------------------------------------------------------------

* autoMCStats 10 1
```

Some checks

1. `combine -M FitDiagnostics datacard.txt -t -1 -expectSignal 1`
   - Fitted value is r = 1, as expected
2. `combine -M FitDiagnostics datacard.txt -t -1 -expectSignal 0`
   - Fitted value is r = 0, as expected
3. Putting sum of MC bkg in data_obs and extracting upper limit with `combine -M AsymptoticLimits datacard.txt` gives same expected and observed limit
   - Good, as the expected limit is computed under the bkg-only hypothesis
   - Expected, stat only limit: $r < 0.0610$
4. Adding MC stat uncertainties: $r < 0.2363$