On the switch to UL campaign and nanoAOD data tier

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

December 1, 2021

Introduction

- It is now **time to switch to UL** campaign (**mandatory** after Moriond 2021)
- To save disk space, we are **considering** switching to **nanoAOD** data tier **too**
  - When using nanoAOD, we could benefit from Duncan's tWIHEPFramework
- Huiling, Duncan and me had a group chat to discuss technical details
  - Only major bottleneck seems to be how to integrate CMSSW-based code into tWIHEPFramework (e.g., resolved HOT)
  - But...
- **Question**: first of all, do we have all the information we need in nanoAOD data tier?
- I've been **investigating this**

Information in nanoAOD data tier

- At this link (thanks Duncan!), we can see all the branches saved in nanoAOD for 2016, together with a brief description
- Went through our object definitions to see if that's enough to implement them

Electron ID

Cuts/Equations for MVA (2016 - MVANoIso94XV2)

1

Different from 2017/2018

Tight

Region	MVA value, 10<ePt<40	MVA value ePt>=40
$\|\eta\| < 0.8$	> 3.447 + 0.063(pt - 25)	> 4.392
$0.8 \leq \|\eta\| < 1.479$	> 2.522 + 0.058(pt - 25)	> 3.392
$1.479 \leq \|\eta\| < 2.5$	> 1.555 + 0.075(pt - 25)	> 2.680

VLoose

Region	ePt: 5-10	10<ePt<25	ePt>=25
$\|\eta\| < 0.8$	> 1.309	> 0.887 + 0.088(pt - 25)	> 0.887
$0.8 \leq \|\eta\| < 1.479$	> 0.373	> 0.112 + 0.099(pt - 25)	> 0.112
$1.479 \leq \|\eta\| < 2.5$	> 0.071	> -0.017 + 0.137(pt - 25)	> -0.017

VLooseFO

Region	ePt: 5-10	10<ePt<25	ePt>=25
$\|\eta\| < 0.8$	> -0.259	> -0.388 + 0.109(pt - 25)	> -0.388
$0.8 \leq \|\eta\| < 1.479$	> -0.256	> -0.696 + 0.106(pt - 25)	> -0.696
$1.479 \leq \|\eta\| < 2.5$	> -1.630	> -1.219 + 0.148(pt - 25)	> -1.219

cms.InputTag("electronMVAValueMapProducer:ElectronMVAEstimatorRun2Fall17NoIsoV2RawValues")

- nanoAOD provides the "Electron_mvaFall17V2noIso" branch; should be OK

Electron ISO

- Relative miniIsolation is OK

$$\text{ISO}^{\text{rel}}_{\text{mini}} = \frac{\text{ISO}^{\text{rel}}_{\text{mini}}}{p_{\text{T}}^{e}}$$

- jet p_{T} ratio **?!?**

$$p_{\text{T}}^{\text{ratio}} = \frac{p_{\text{T}}^{e}}{p_{\text{T}}^{\text{ejet}}}$$

where *ejet* is the jet containing the electron (or the closest to the electron)

- $p_{\text{T}}^{\text{rel}}$ **?!?**

$$p_{\text{T}}^{\text{rel}} = \frac{\left|(\vec{p_{\text{T}}}^{\,\text{ejet}} - \vec{p_{\text{T}}}^{\,e}) \times \vec{p_{\text{T}}}^{\,e}\right|}{\left|\vec{p_{\text{T}}}^{\,\text{ejet}} - \vec{p_{\text{T}}}^{\,e}\right|}$$

Electron ISO

```
for(const pat::Jet &j : *jets){
  pat::Jet jet = j;
  double dr = deltaR(ele->p4(),jet.p4());// #include "DataFormats/Math/interface/deltaR.h"
  for(unsigned int i1 = 0 ; i1 < ele->numberOfSourceCandidatePtrs();i1++){
      const reco::CandidatePtr  &c1s=ele->sourceCandidatePtr(i1);
      for(unsigned int i2 = 0 ; i2 < jet.numberOfSourceCandidatePtrs();i2++) {
          const reco::CandidatePtr  &c2s=jet.sourceCandidatePtr(i2);
          if(c2s== c1s){
              elejet = jet;
              elejet_mindr = dr;
              lepjetidx = currjetpos;
              break;  // take leading jet with shared source candidates
          }
      }
      if(lepjetidx >=0)break;// take leading jet with shared source candidates
  }
  currjetpos++;
  if(lepjetidx >=0)break;// take leading jet with shared source candidates
}
```

- Don't think we can repeat the exercise with nanoAOD (or do we?)

Electron IP

- **dXY**: we have it
- **dZ**: in BSM we store

  ```
  patElectron_gsfTrack_dz_pv.push_back(el->gsfTrack()->dz(firstGoodVertex.position()));
  ```

  while in nanoAOD we have "dz (with sign) wrt first PV, in cm". They could be the same thing
- **3D IP significance**: we have it
- **Missing inner hits**: in BSM we store

  ```
  expectedMissingInnerHits.push_back(el->gsfTrack()-
  >hitPattern().numberOfLostHits(reco::HitPattern::MISSING_INNER_HITS));
  ```

  while in nanoAOD we have "number of missing inner hits", maybe it's the same?
- **Conversion veto**: we have it in nanoAOD

Muons

- **ID**: we have mediumID flag in nanoAOD
- **ISO**: same problems of electrons
- **dZ**: in BSM we store

  ```
  Muon_dz_bt.push_back(mu->muonBestTrack()->dz(firstGoodVertex.position()));
  ```

  while in nanoAOD we have "dz (with sign) wrt first PV, in cm"
- **dXY**: same thing
- **3D IP significance**: we have it

Hadronic taus

- **dZ**: we have it in nanoAOD
- **decayModeFindingNewDMs**: we have it in nanoAOD
- **decayMode**: we have it in nanoAOD
- **vsEle, vsMu & vsJet** discriminators: we have all WPs in nanoAOD

Jets

- **Neutral had & EM energy fractions**: we have them in nanoAOD
- **Charged had & EM energy fractions**: we have them in nanoAOD
- **Muon energy fraction**: we have it in nanoAOD
- **Number of constituents**: we have it in nanoAOD
- **Charged multiplicity ?!?**: can't find it
- **DeepJet score**: we have it in nanoAOD
- **Remark**: in miniAOD, need to rerun on top to get DeepJet score. Do we need all these cuts when using nanoAOD?

DeepTau scale factors

- As an open item, we have the implementation of DeepTau SFs
- **Needed ingredients** were to be added in BSM
- Do we have them **in nanoAOD?**

The SFs are meant for the following campaigns:

Year label	MC campaign	Data campaign
2016Legacy	RunIISummer16MiniAODv3	17Jul2018
2017ReReco	RunIIFall17MiniAODv2	31Mar2018
2018ReReco	RunIIAutumn18MiniAOD	17Sep2018 / 22Jan2019
UL2016_preVFP	RunIISummer20UL16*APV	UL2016_MiniAODv1
UL2016_postVFP	RunIISummer20UL16	(HIPM_)UL2016_MiniAODv1
UL2017	RunIISummer20UL17	UL2017_MiniAODv1
UL2018	RunIISummer20UL18	UL2018_MiniAODv1

- Documentation looks like only for miniAOD **?!?**

Extras: multiISO with LepAware JEC

- As ZhangYu pointed out, we are not doing things correctly with lepton ISO inside BSM
- **Not applying JECs correctly** when computing variables used for ISO
- This has **to be fixed** whatever choice we made on data tier & framework!

Conclusions

- **Not all the ingredients** we need appear to be **ready** in nanoAOD
- We could lack info for electrons, muons, jets, DeepTau SFs, [...]
- How to deal with this?
  - **Is it really convenient** to switch to nanoAOD?
  - Will it require too much time?
  - Will we have all the ingredients we need?
- Also, let's **remember to fix details on lepton multiISO** (I added this item to the Google Doc I recently created)