QT Updates

Jan 30, 2020

Xiaoning Wang

University of Illinois at Urbana Champaign

Summary of Plan

- Goal: work on qualification task on optimizing the inputs of high-level discriminates (DL1 & MV2) for b-tagging in HI collisions.
- Problem: HI collisions have large number of Underlying Event (UE) tracks that modify some inputs.
- First step plans:
  - Look into current performance of SV reconstruction in HI collisions using MC overlay and compare with MC signal.
  - Experiment with selections on tracks in HI MC for all lower level taggers to see how it affects performance.

"The optimization of the inputs of high-level taggers(DL1 and MV2) for b-tagging in heavy ion collisions, following the work done in a previous QT described in AFT-233. It is known that some inputs for the taggers training are affected (like ipxd probabilities and jet fitter and sv1 energy fraction) by the large number of tracks coming from the HI collision underlying event (UE). This degrades the performance for central collisions and induces a strong centrality dependence. This effect can be reduced by implementing tighter tracking selections or an UE subtraction at the tracking level prior the calculation of the tagger inputs. If time permits, following the optimization, the calibration of the taggers will be done using HI data control samples that have a specific flavor composition e.g. jets with a muon from a heavy flavor semi-leptonic decay. This study will be documented in an internal note and the analysis recommendations will be described on a twiki."

Planned Effort

- Ongoing works:
  - Xiaoning will be looking at the efficiency, purity and position resolution of the secondary vertex reconstruction as a function of centrality using the JetFitter and SV1 algorithms.
  - Will be working on the track selections for lower level taggers.
    - Have compiled the machinery to do track selections for lower level taggers
  - Dominik is going to start comparing the existing data and MC we're using.

MC Samples used

- MC Overlay:
  - mc16_5TeV.420271.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ1_bbfilter.recon.AOD.e7383_d1521_r11472
  - mc16_5TeV.420272.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ2_bbfilter.recon.AOD.e7383_d1521_r11472
  - mc16_5TeV.420273.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ3_bbfilter.recon.AOD.e7383_d1521_r11472
  - mc16_5TeV.420274.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ4_bbfilter.recon.AOD.e7383_d1521_r11472

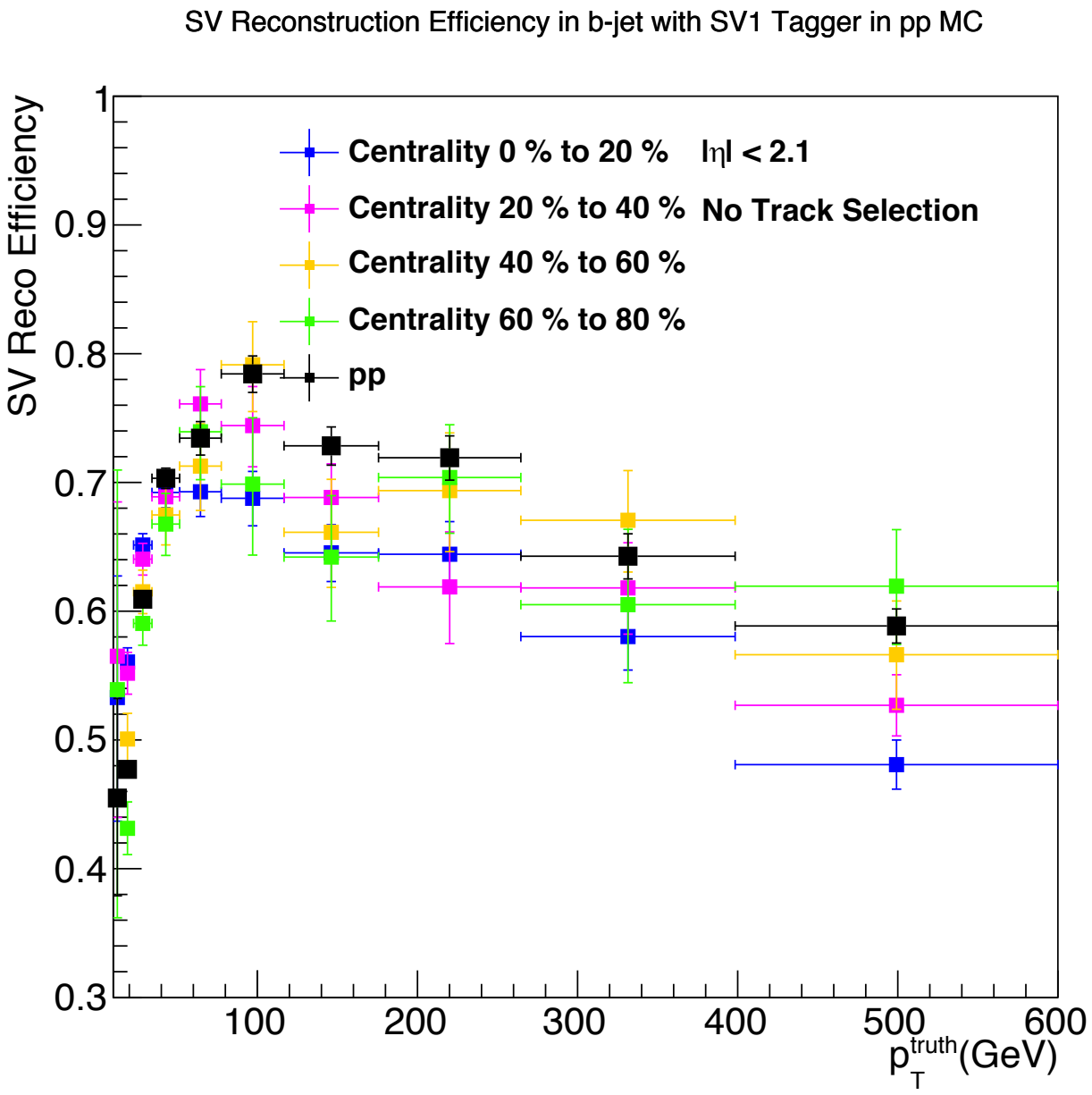
- MC Signal:
  - mc16_5TeV.420274.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ4_bbfilter.recon.AOD.e7383_s3428_r11320
  - mc16_5TeV.420273.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ3_bbfilter.recon.AOD.e7383_s3428_r11320
  - mc16_5TeV.420272.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ2_bbfilter.recon.AOD.e7383_s3428_r11320
  - mc16_5TeV.420271.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ1_bbfilter.recon.AOD.e7383_s3428_r11320

SV Reconstruction using SV1 Tagger

- Selection on Jets:
  - Reco jets must be truth matched using dR < 0.3
  - Truth jets matched must have $p_T > 10$ GeV
  - Pt of truth jets matched do not differ with reco pt by more than a factor of 2.
  - B-Jets are defined as jets with a truth B hadron associated with it.
    - B hadron has $p_T > 5$ GeV
    - B hadron and jet has dR < 0.3

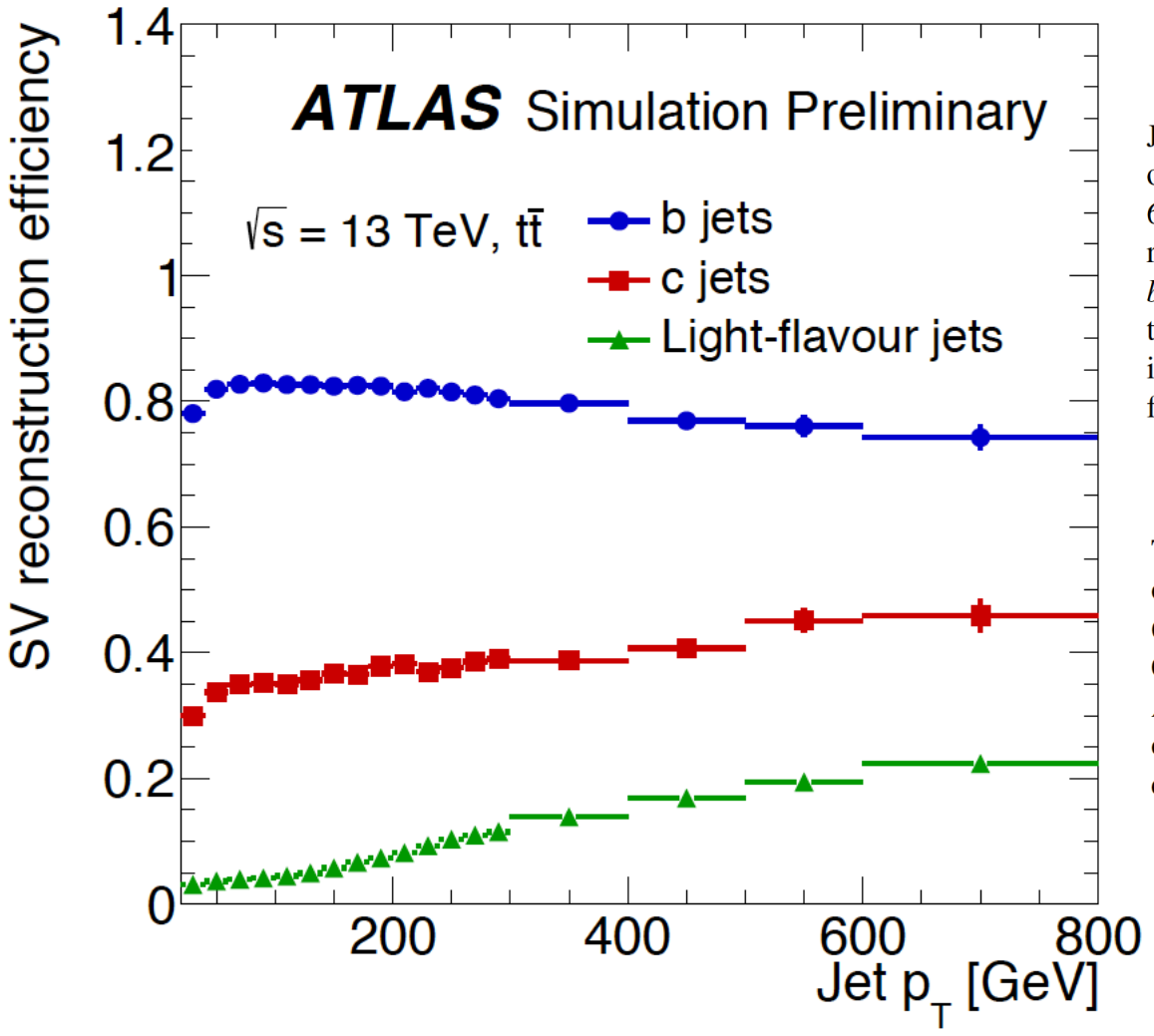
- Selection on Tracks:
  - No selection
  - HILoose minus d0 z0 selections.
    - Tracks have absolute eta < 2.5
    - Use min bias inner most layer cut (pass the track if expected hit on IBL is absent)
    - Minimum pixel of hits is 1
    - Pt Cutoffs for Sct Hits = [0.0, 300.0, 400.0]
    - Min N Sct Hits Above Pt = [2, 4, 6]

SV Reco Efficiency in SV1



- Eff = # of Jets with SV reconstructed/# of Jets with truth B hadron matched
- SV Reconstruction for PbPb MC for central collision is slightly (~5%) lower than pp.
- Minor dependence on centrality
- There's about 10% drop in efficiency from low pt to high pt jets.

SV Reco Efficiency Quoted in SV1 paper



- https://cds.cern.ch/record/2270366

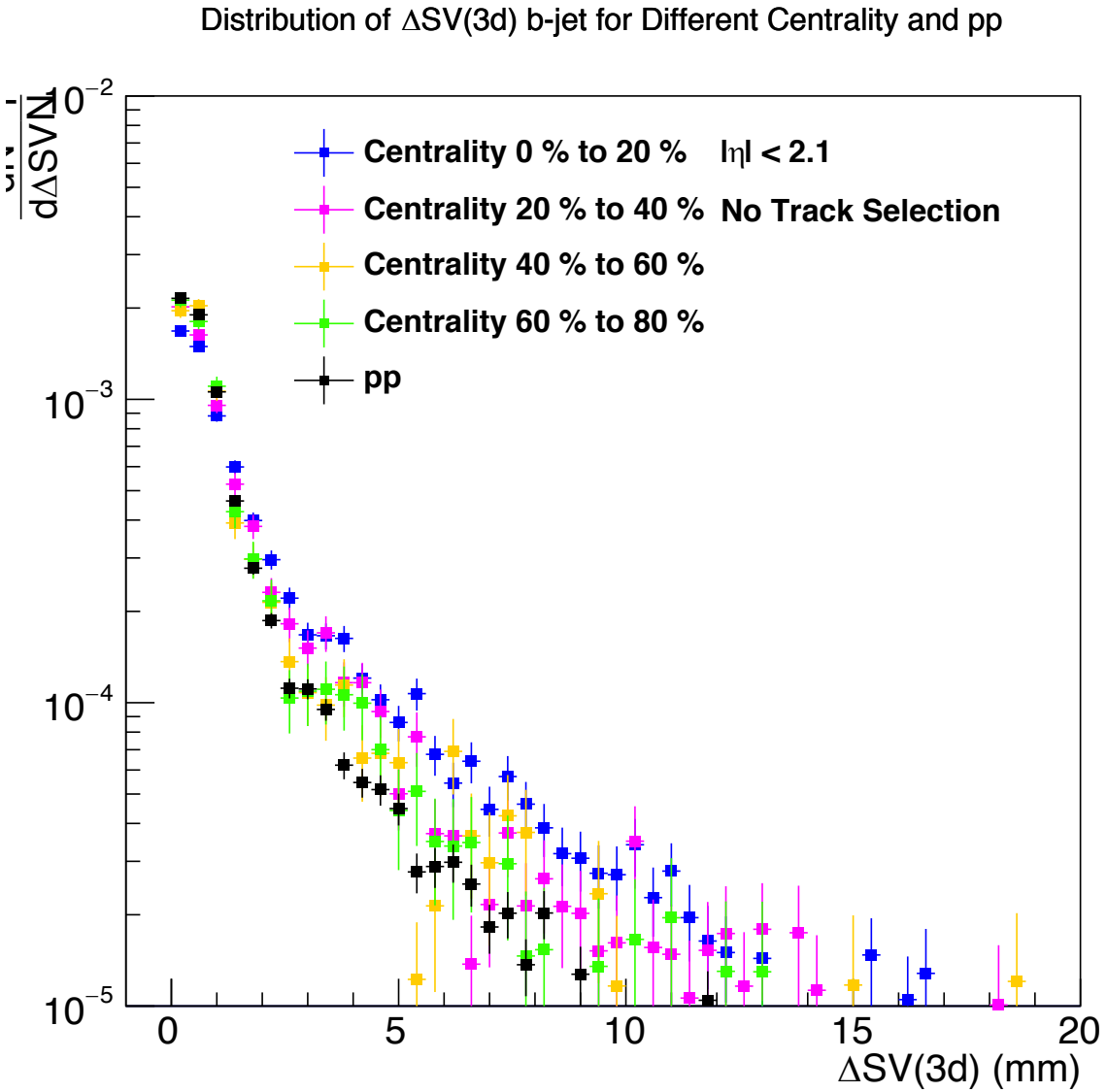
Jets are reconstructed using the anti-k_t algorithm [19]. Jets are required to have a transverse momentum of $p_T > 20$ GeV and a pseudorapidity of $|\eta| < 2.5$. In addition, jets with a transverse momentum of $p_T < 60$ GeV and pseudorapidity of $|\eta| < 2.4$ are required to pass the Jet Vertex Tagger requirement, aimed at reducing the impact of pile-up [20] . The jets are flavour labeled by matching them to weakly decaying b- and c-hadrons in the event generator record. If a b-hadron is found within a distance of $\Delta R < 0.3$ from the jet axis, then the jet is labeled as a b-jet. If no b-hadron is found, the search is repeated for c-hadrons, if a c-hadron is found and no τ leptons are found, the jet is then labeled as a c-jet. If no match is found for c, b, or τ, the jet is labeled as a light-flavour jet.

- ^same flavor labeling method as what we have.

The results in this note are obtained using $t\bar{t}$ simulated events corresponding to a centre-of-mass energy of $\sqrt{s} = 13$ TeV proton-proton collisions, generated using the Powheg generator [11] interfaced to Pythia 6 [12] with the Perugia 2012 tune. EvtGen [13] is used to model the decays of b- and c-hadrons. The CT10 [14] parton distribution function set is used in the simulation. More details are available in Ref. [15]. Additional minimum-bias pp interactions are simulated with the Pythia 8 generator [16] and are overlaid on the simulated events with an average of 20 pileup interactions per event to emulate the data-taking conditions of Run-2.

- Different jet/track reconstruction?
- Plan: look at a standard pp sample to see whether there's still difference.

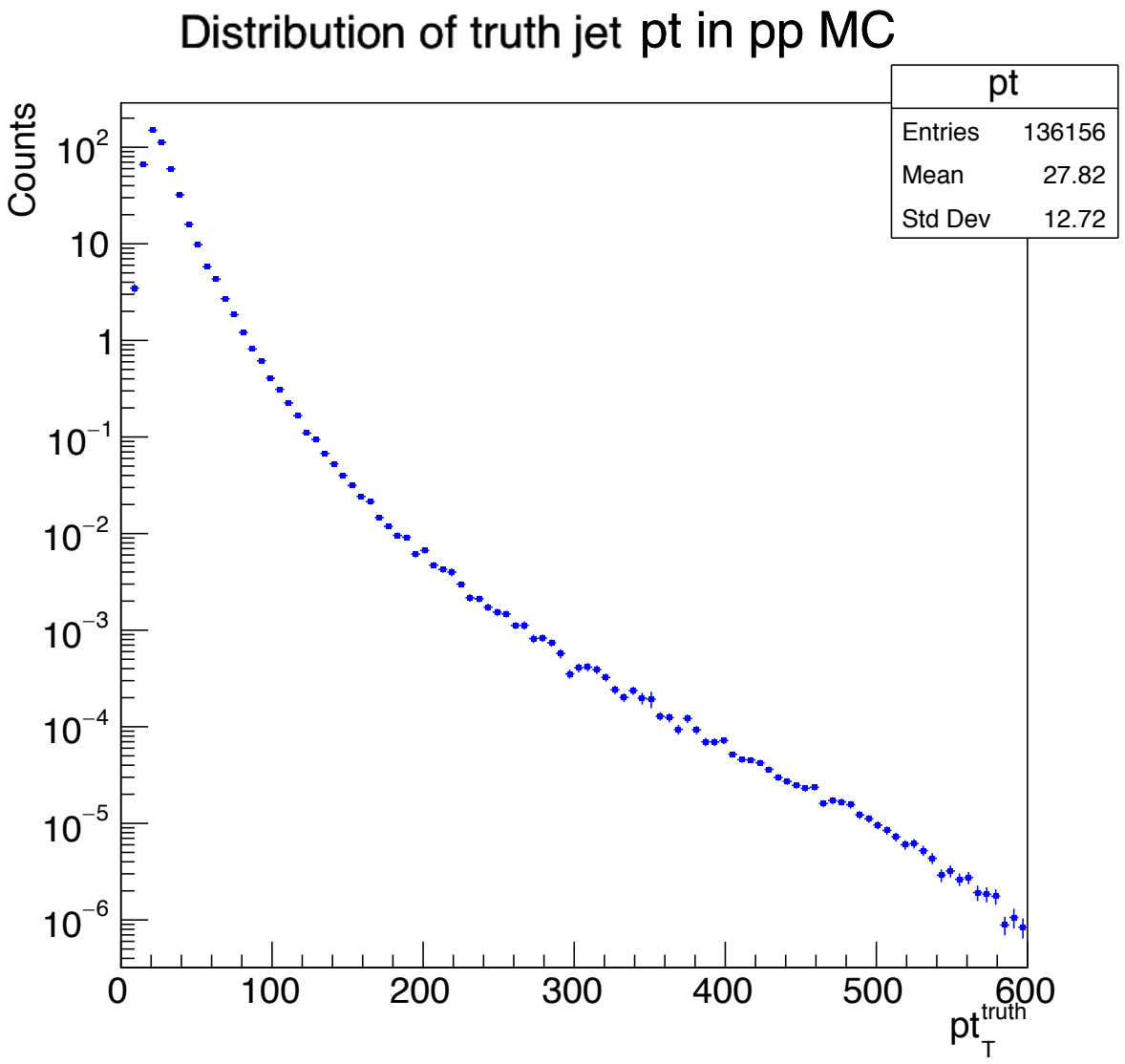
SV Reco Resolution in PbPb Overlay and pp MC

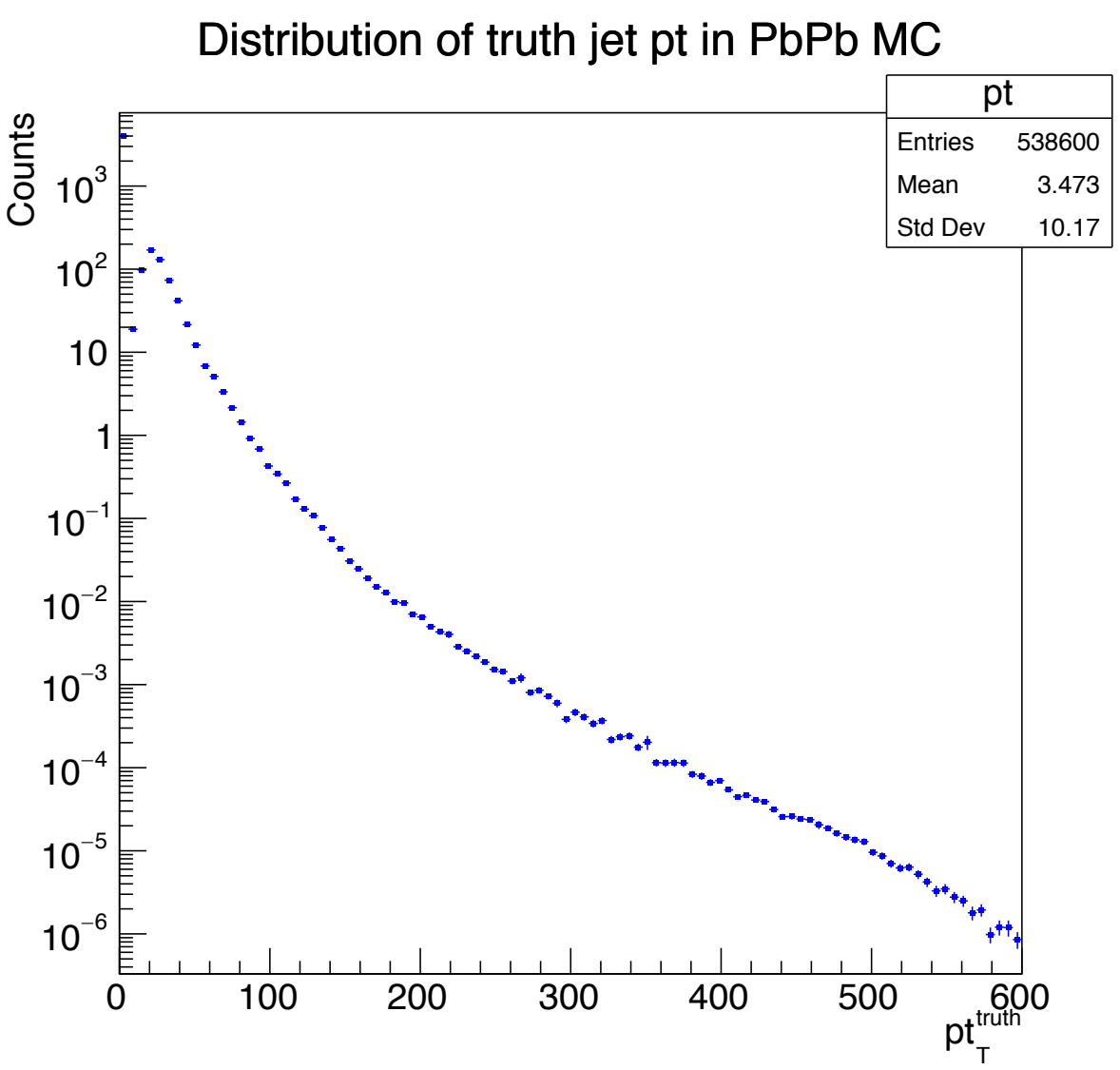


- SV reconstruction resolution here is defined as decay length difference.
  - Similar dependence on centrality can be seen in the distance between reco and truth SV, see back-up slide.
- More central events have more widely distributed d(L3d) (abs(reco-truth) decay length), as expected.

- The SV resolution of pp MC is similar to overlay peripheral, as expected

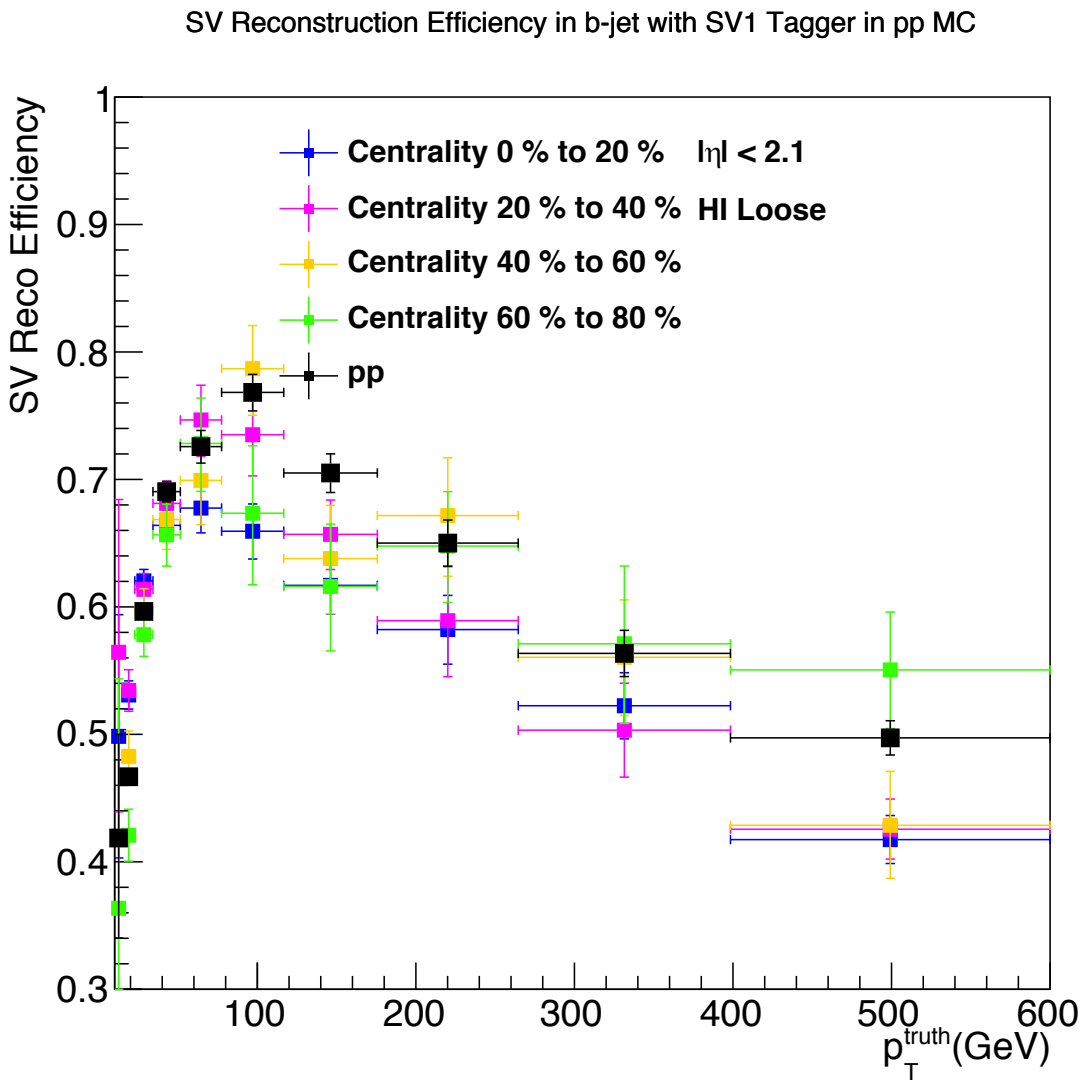
MC pt Distribution for Validation Purpose

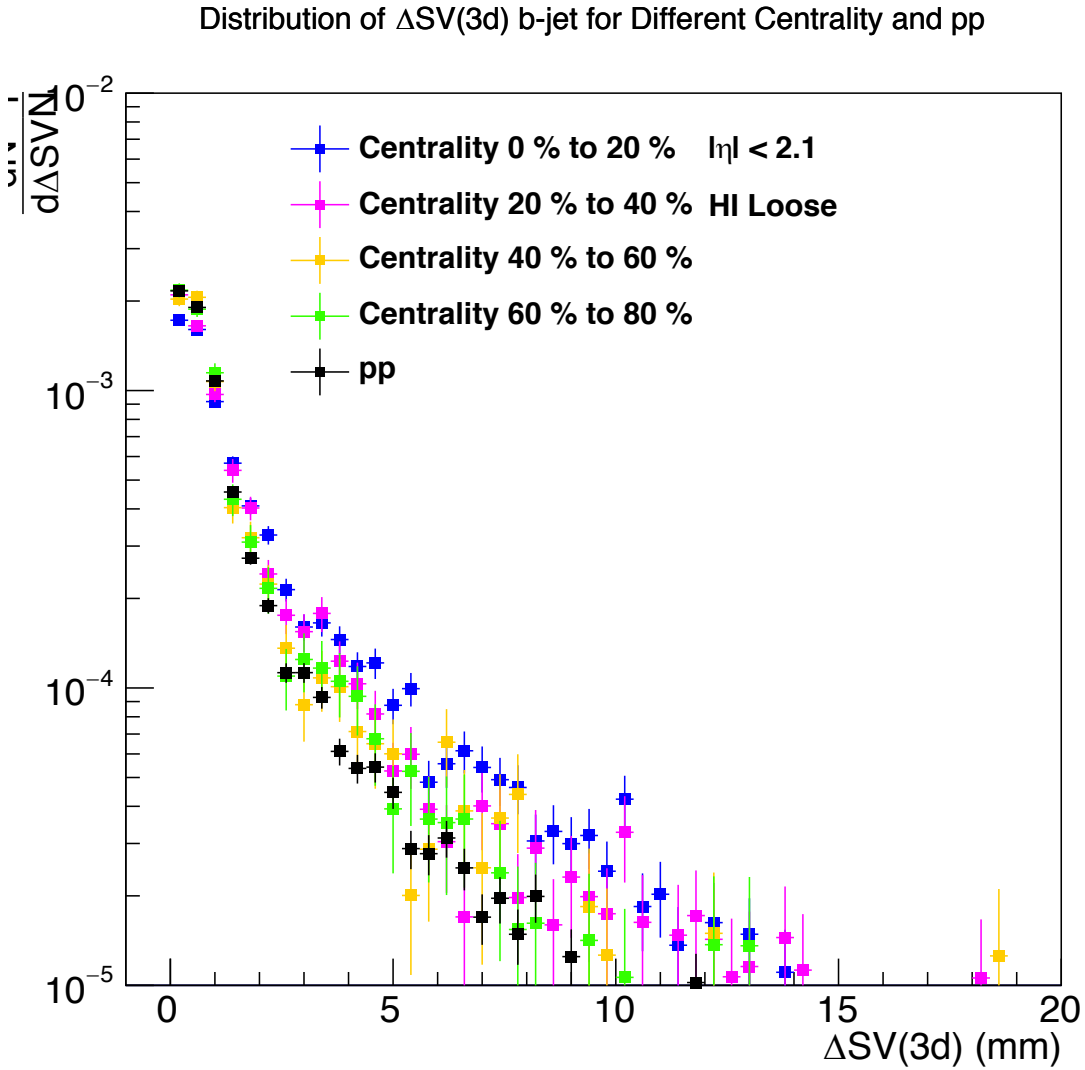


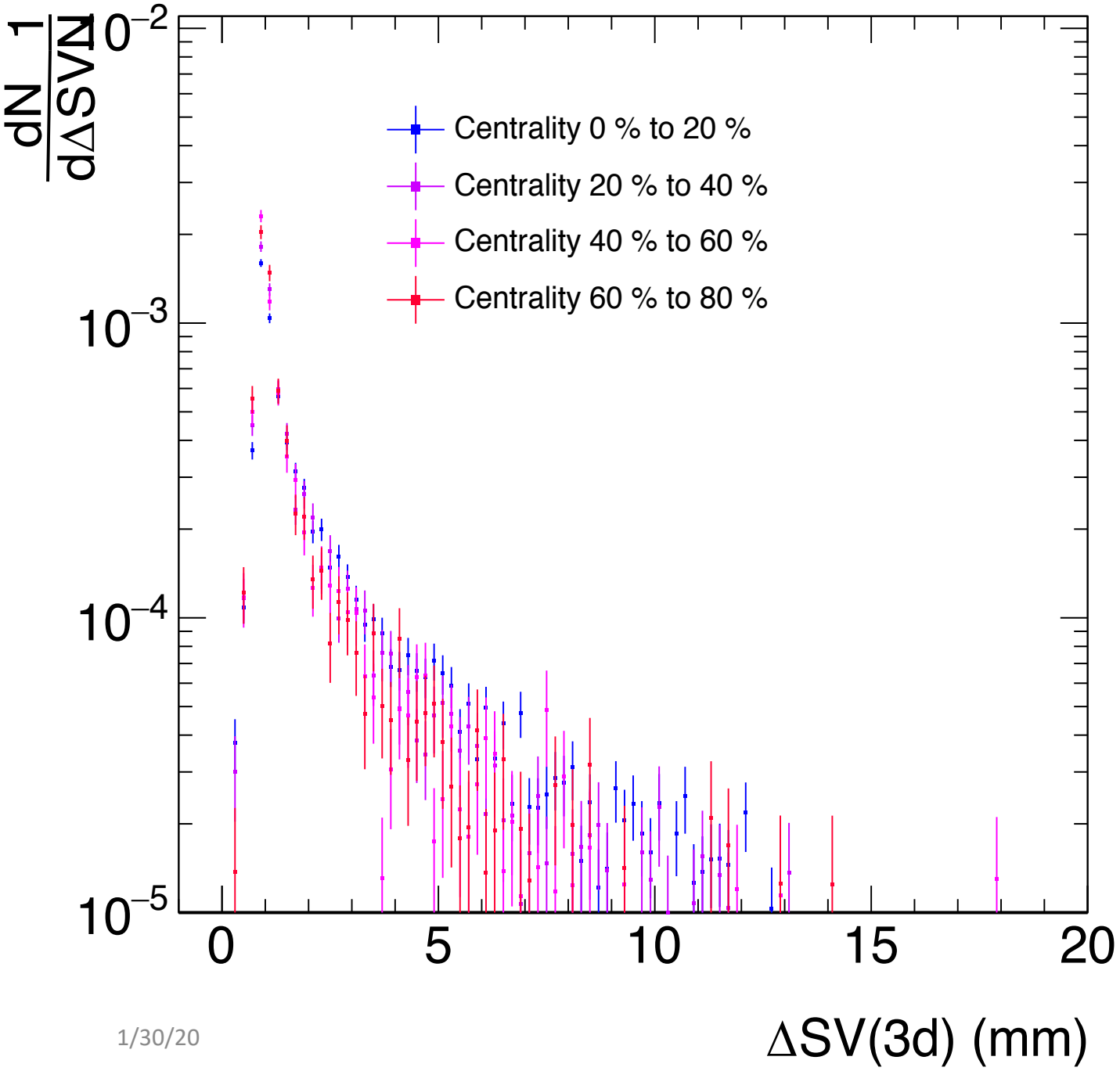


Back-Up

HI Loose Track Selection






$\frac{dN}{d\Delta SV} \frac{1}{N}$
10^{-2}
10^{-3}
10^{-4}
10^{-5}
Centrality 0 % to 20 %
Centrality 20 % to 40 %
Centrality 40 % to 60 %
Centrality 60 % to 80 %
0
5
10
15
20
ΔSV(3d) (mm)

Goal of the Week

- For b-jet and c-jet using 50k MC overlay and 50k MC signal, for the default options of SV and JF algorithms respectively,
  - Plot decay length (L3d) and x-y plane decay length (Lxy) resolution graphs and compare.
  - Plot fake rate (SV in light jet) and efficiency of SV reconstruction in b-jet and c-jet.
  - For MC overlay, add in centrality information

- Compile, test and try more options with track selection tools
  - With small local samples, compile and test the different pT, cutLevel, etc.
  - If lower tagger inputs are improved, try with large samples.