QT Updates

Jan 20, 2020

Xiaoning Wang

University of Illinois at Urbana Champaign

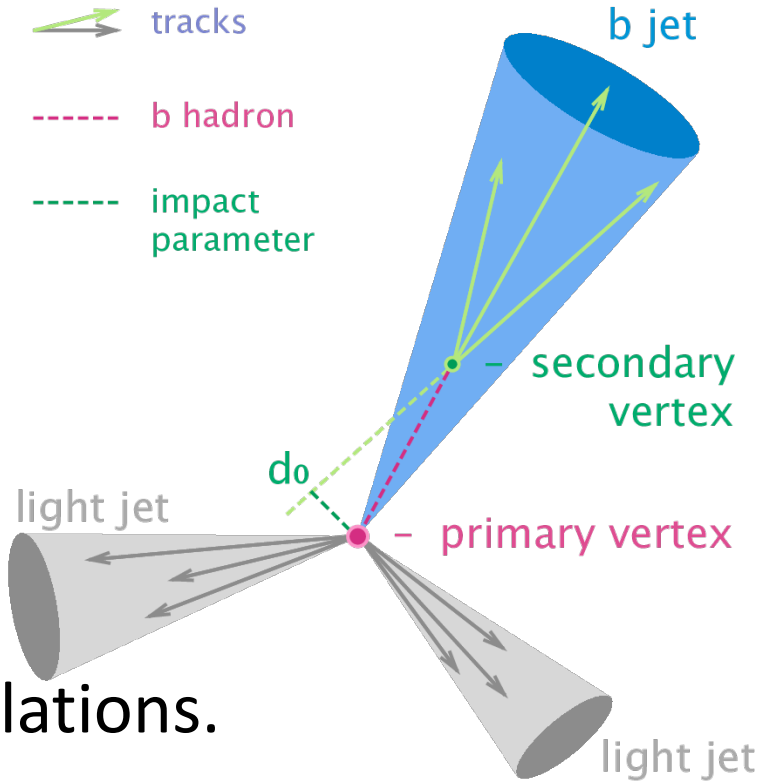
Summary of Plan for Workfest

- 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."

Description of Task (AFT-455)

- B-tagging: the identification of jets containing b-hadrons
  - Several dedicated algorithms exploiting specific proprieties like long lifetime, high mass and decay multiplicity of b-hadrons and the hard b-quark fragmentation.
  - Low Level Discriminants:
    - IP3: Impact Parameter (d0, z0) based
    - SV1: Secondary Vertex based
    - JetFitter: Secondary Vertex based
  - As input of High Level Discriminants:
    - MV2 and DL1



- High level discriminants are now trained on *pp* simulations.
- B-tagging algorithms can now run heavy ion data from AFT-233 (Rui).

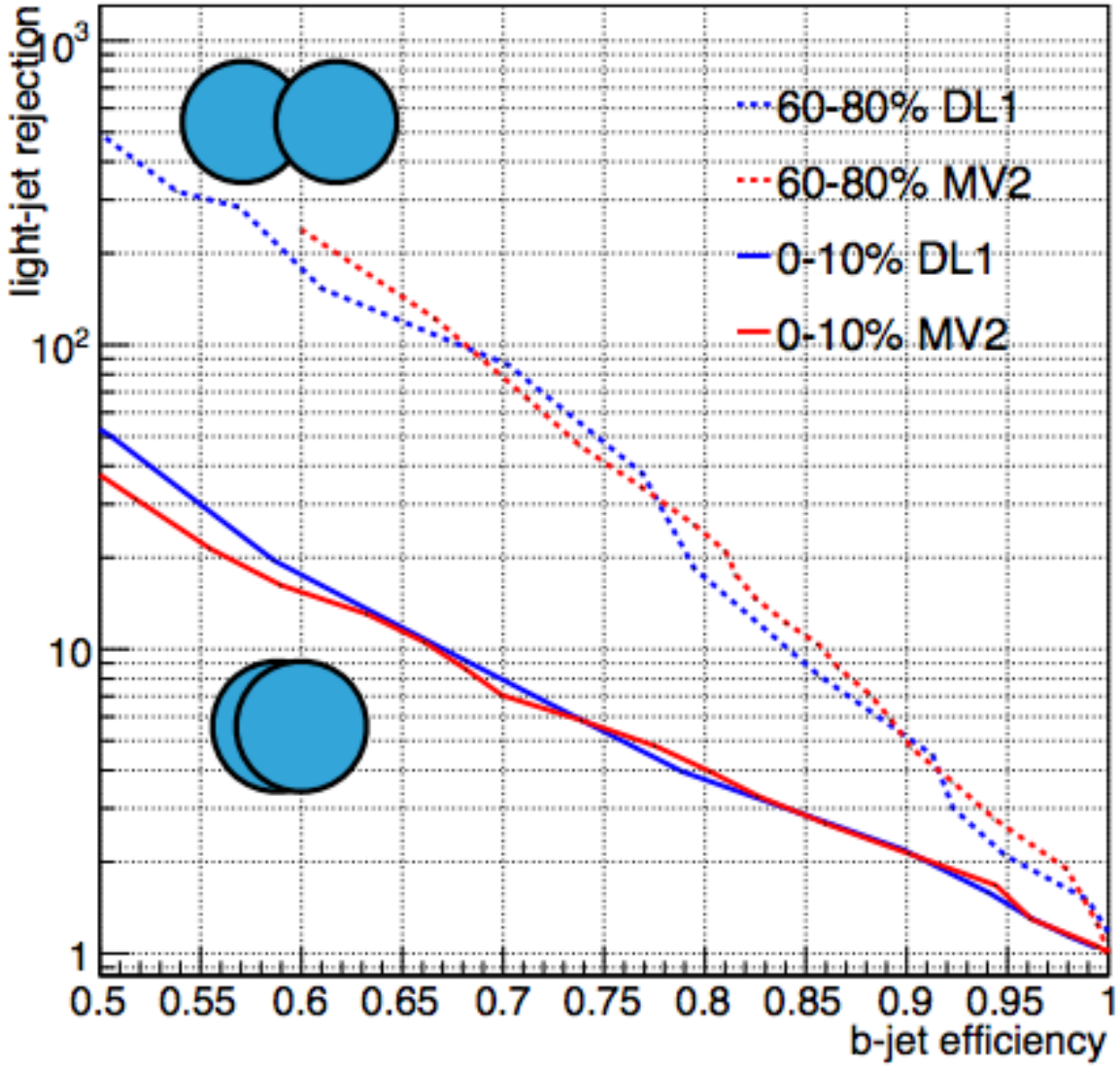
Scope of This Study

- PbPb 2018 5.02 TeV data with integrated luminosity 1.76 nb^{-1}.
- We estimate to have order of 100 b-jets at p_T 200-300 GeV, and order of 10 b-jets at p_T 300-400 GeV, and the latter will be the upper limit of this optimization.
  - The estimation was made for all centrality using the $L_{int} \times \sigma^{pp}_{b-jets} \times \langle N_{coll} \rangle \times \frac{\sigma^{tot}_{AA}}{\sigma^{pp}_{tot}}$, where $\langle N_{coll} \rangle \times \frac{\sigma^{tot}_{AA}}{\sigma^{pp}_{tot}}$ was estimated as 40,000. (N_{coll} is the expected number of binary nucleon-nucleon collisions)
  - The theoretical calculation for b-jets cross section is from this paper by Hai Tao Li and Ivan Vitev: https://arxiv.org/abs/1811.07905

- There are ongoing measurements within the HI group of b-jet cross-sections in PbPb and *pp* at 5.02 TeV using muon based tagging, which can provide cross checking reference.
  - The link to this internal note is here: https://cds.cern.ch/record/2683608?#

B-Tagging Performance in Overlay MC

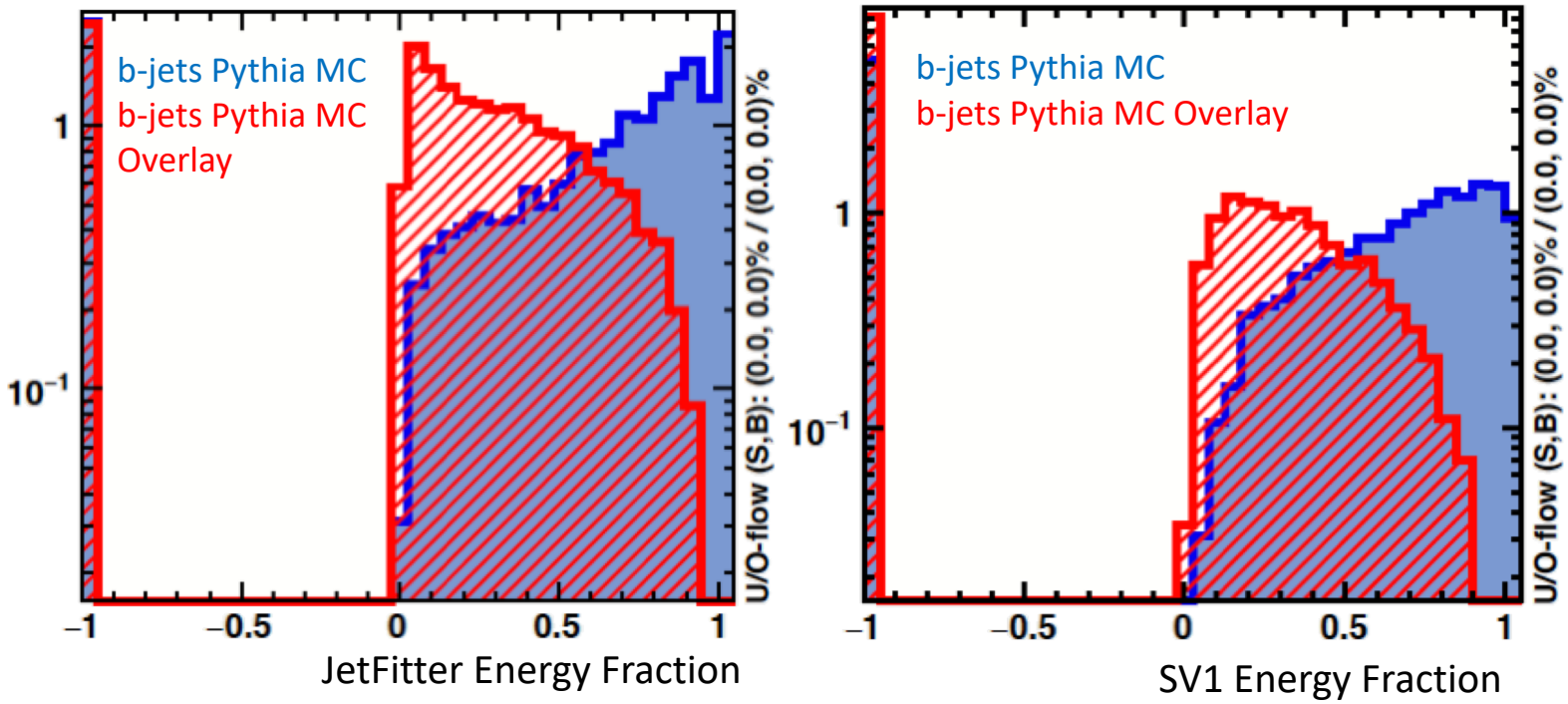
Inclusive Jet-MC

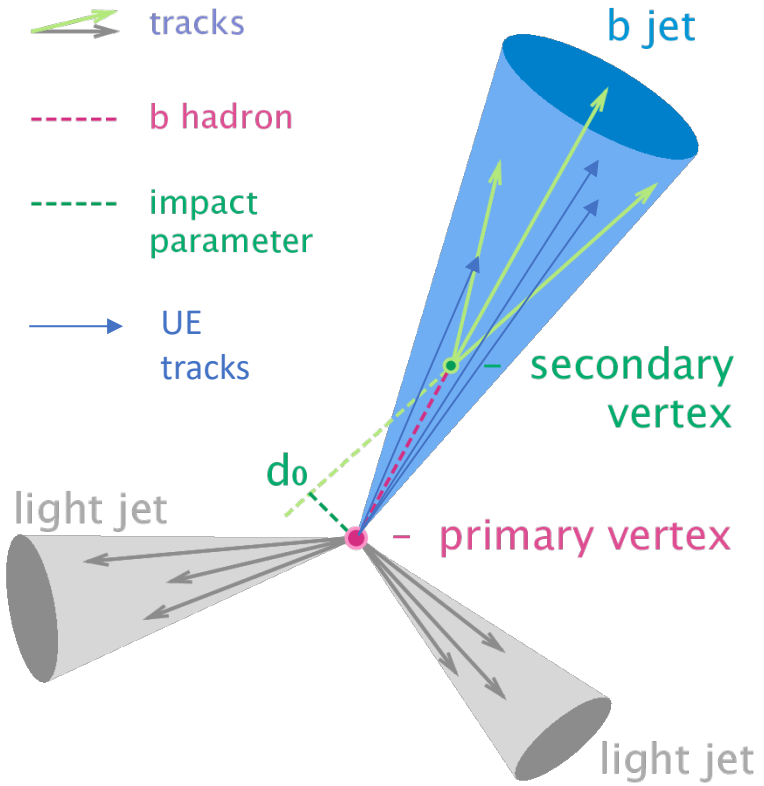


For the same efficiency in b-jet efficiency, central collisions have an order of magnitude lower of light-jet rejection than peripheral collisions.

Improve Second Vertex Reconstruction based inputs

- Tracks from Underlying Events (UE) are mistaken as part of jet tracks.
  - Wrong energy fraction (efc=$E_{SV.trks}/E_{All\ trks}$), UE are included in denominator.





Planned Effort

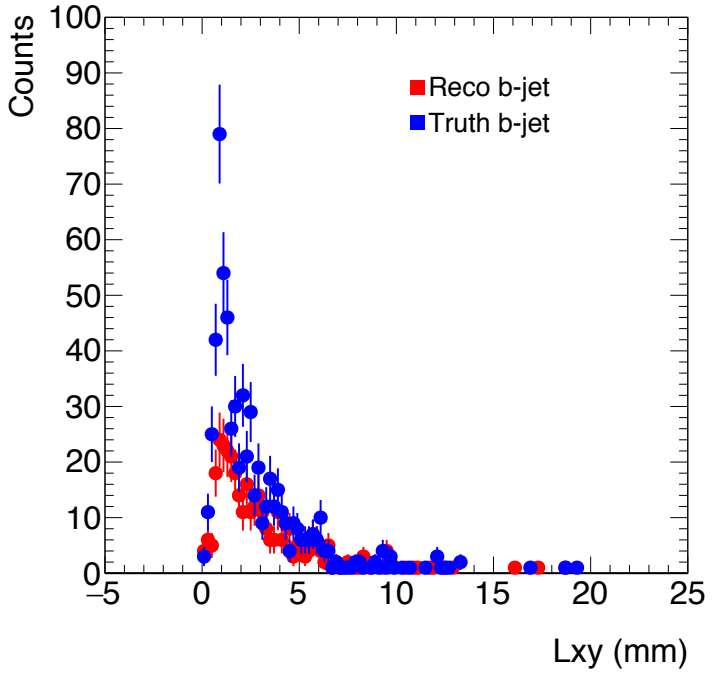
- Ongoing works:
  - Xiaoning will be looking at the efficiency, purity and position resolution of the secondary vertex reconstruction as a function of centrality using the JetFitter and SV1 algorithms.
    - Dataset running.
  - Will be working on the track selections for lower level taggers.
    - 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.

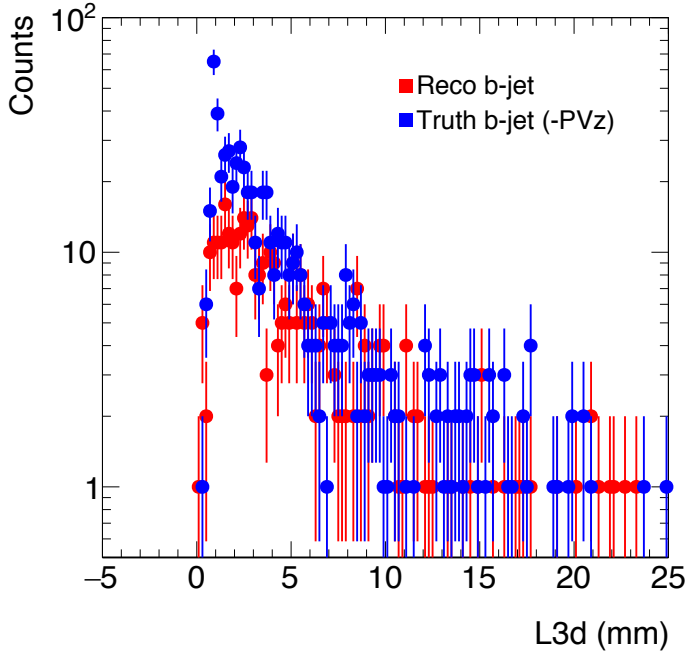
Back-Up

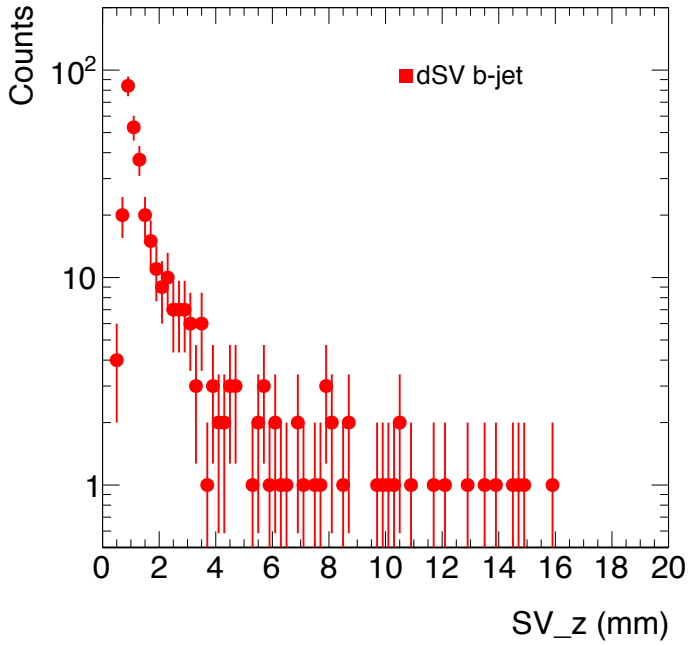
MC Samples used

- 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


Counts
Reco b-jet
Truth b-jet
Lxy (mm)


Counts
Reco b-jet
Truth b-jet (-PVz)
L3d (mm)


Counts
dSV b-jet
SV_z (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.