b tagging scale factors and their uncertainties

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

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

[2]Fudan University, Shanghai

November 10, 2021

Introduction

- Several methods are available to apply b tagging SFs
- We use the whole b tagging discriminant spectrum in our analysis (BDT variables)
  - We need the BTagShapeCalibration method
  - Aims to correct the whole b-tagging discriminant distribution in MC to match that in data
  - Method does not induce migration between b jet multiplicity bins
  - Per event weight is

$$\omega_{\text{event}} = \prod_{i}^{\text{njets}} SF(D_i, p_{\text{T}i}, \eta_i, F_i)$$

Phase space extrapolation factors

- Important **note on event yields**
- **Before** applying any **b tag** selection criteria, expected **event yields should be preserved**
  - Analysis phase space (PS) can be different from PS where SFs were extracted
  - This condition is not automatically fulfilled
- Analysts should measure the **sum of event weights before and after applying b-tag event weights, without requiring any b-tag selection**
- Ratio

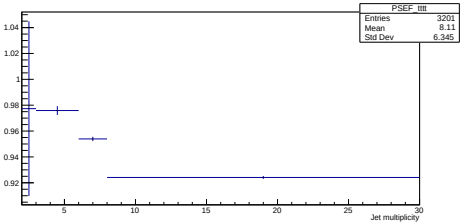
$$\sum \omega_{\text{before}} / \sum \omega_{\text{after}}$$

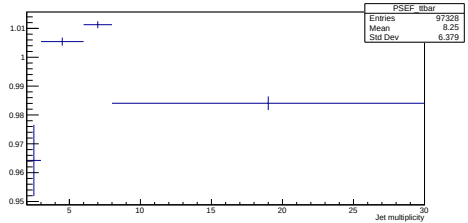
  is a phase space extrapolation factor (PSEF) to be multiplied by the SF
- In high jet multiplicity analysis, **should be done per jet multiplicity**

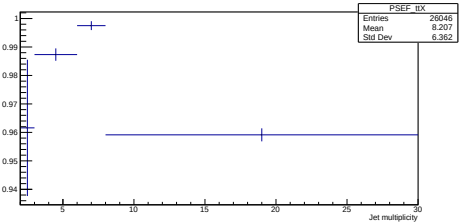
Phase space extrapolation factors

- We have **b tagging requirements in our preselection**
  - Need to produce **new ntuples with no requirement on b tags**
  - Launched jobs
- Produced one PSEF for each "kind" of process we consider
  - $t\bar{t}t\bar{t}$
  - $t\bar{t}$
  - QCD $t\bar{t}X$
  - Single top
  - Single Higgs
- **Split them in jet multiplicity bins**
- Bin width is 2 (following our category definitions): $\{2, 4, 6, 8, 30\}$

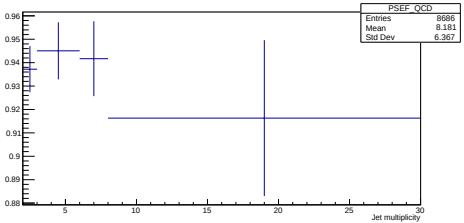
Phase space extrapolation factors

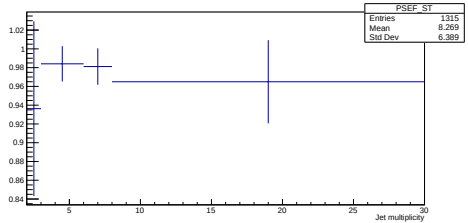


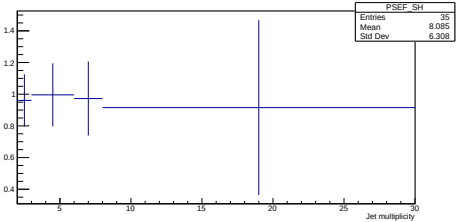




Phase space extrapolation factors







Systematic uncertainties

- BTagShapeCalibration is subject to many sources of uncertainties; among these there's JES
- Two .csv files are available, with full or reduced JES uncertainty sources
- I picked the second (still, 12 sources of JES uncertainty...)
- **I have a C++ method to deal with b tag uncertainties**
- Will soon be able to produce upwards-/downwards-shifted shapes for b tagging
  - Expect many histograms: 1 central value + $\times$ 20 sources of unc. $\times$ 2 = 41

Open items

- What about **categories where we don't have a BDT**? Should we treat b tagging SFs differently there?
- We don't use the full b tagging discriminant shape there
- If a different treatment is needed, I need to study (always used BTagShapeCalibration)
- I can "easily" **provide** upwards-/downwards-shifted shapes for **other uncertainties**
  - Muon SFs
  - Electron SFs
  - Trigger SFs
  - QCD shape estimation
- **Will do so** when done with b tagging uncertainties