event generation for flavor-tagging in ATLAS

Chris Pollard

on behalf of the ATLAS collaboration

2019 05 02

goals

- monte-carlo event generators are clearly a mainstay of data analysis at the LHC (and elsewhere)
- I'll give a very broad overview of MC production at ATLAS,
- focus a bit on how MC modeling affects flavor-tagging
- include some information about how we correct our simulation,
- give some details for which MCs we use for training,
- and point out some limitations from the currently available MC setups.
- lots of public documentation is available: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MCPublicResults

which generators?

many, many MC event generators (+ versions + tunes) are in use in ATLAS for various purposes

very loose list of general-purpose generators setups and their uses in ATLAS

- **pythia8 (Py8)**: "nominal" **parton shower (PS) + hadronization**; also primary **dijet sample; ATLAS tunes are used + EvtGen**
- **herwig7 (H7)**: "alternative" PS+ hadronization and dijet; **no specific ATLAS tunes (that I'm aware of) + EvtGen**
- **sherpa**: "nominal" *V*+jets, multiboson; "alternative" *tt*, dijet, etc. **no EvtGen**
- **PowHeg**: "nominal" *tt* and some "alternative" samples. **interfaced to Py8 and H7**
- **aMC@NLO**: "alternative" *tt*, *V*+jets, multiboson, etc. **usually interfaced to Py8**
- **MadGraph**: used for some leading-order setups and many BSM signal samples, **usually interfaced to Py8**

you can fact check all of this: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MCPublicResults

common underlying issues related to flavor-tagging

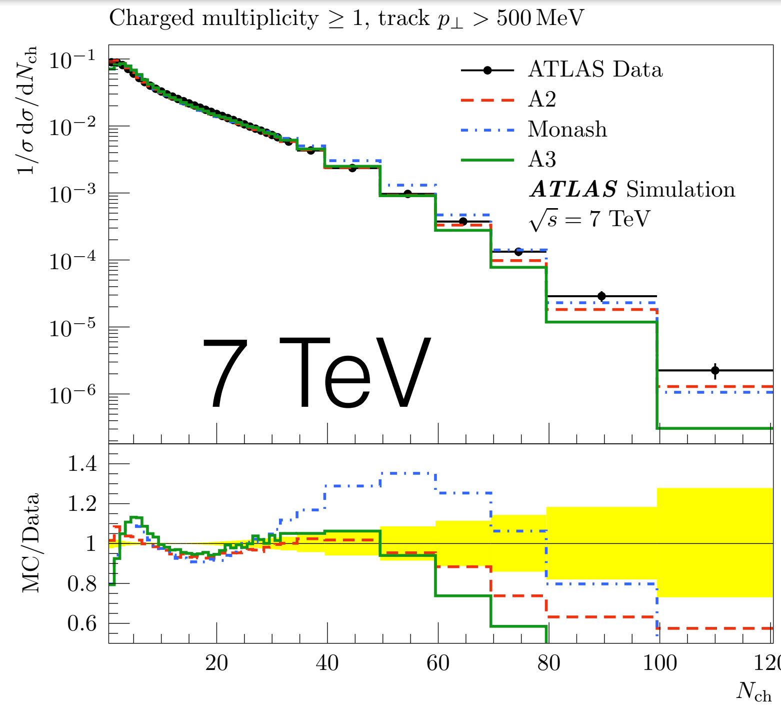
there are a number of (often vague or ill-defined) aspects of collision events that we need to model reasonably well to have flavor-tagging that is useful for physics analysis.

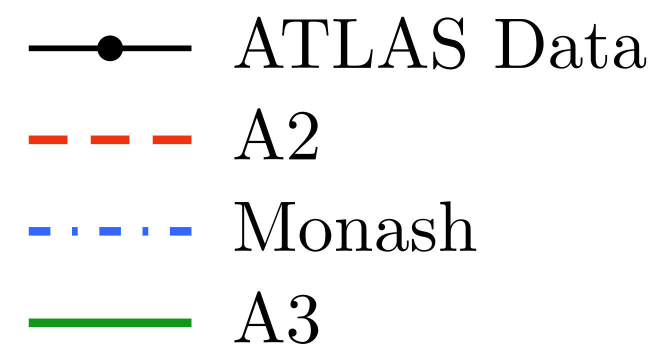
- pileup+underlying event
- "matrix-element kinematics"
- fragmentation and "jet shapes"
- hadronization and species production rates
- hadron decays
- list goes on...

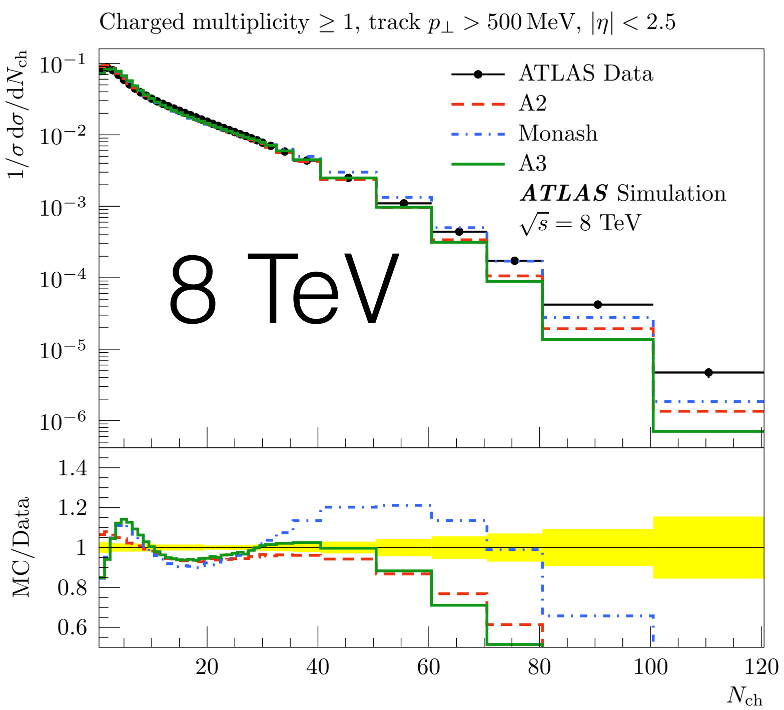
can we model each of these well enough for the event topologies that are interesting when it comes to flavor-tagging?

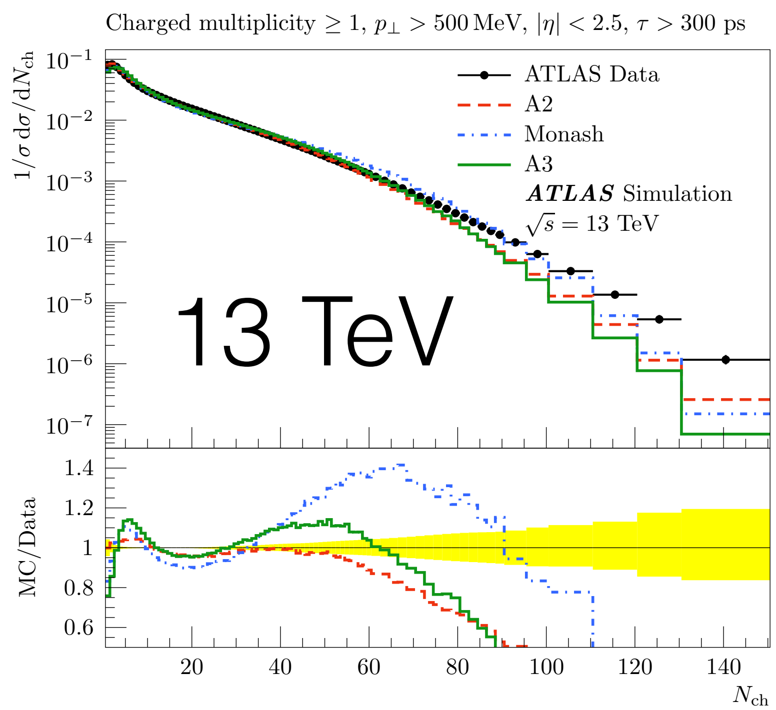
modeling pileup

- we use the **ATLAS A3 min-bias tune** for pythia8
- based on **charged particle distributions** and **inelastic cross section measurements** from Run I and early Run II
- in the end our **high-level taggers don't seem particularly sensitive to pileup**, but **low-level taggers can be**.





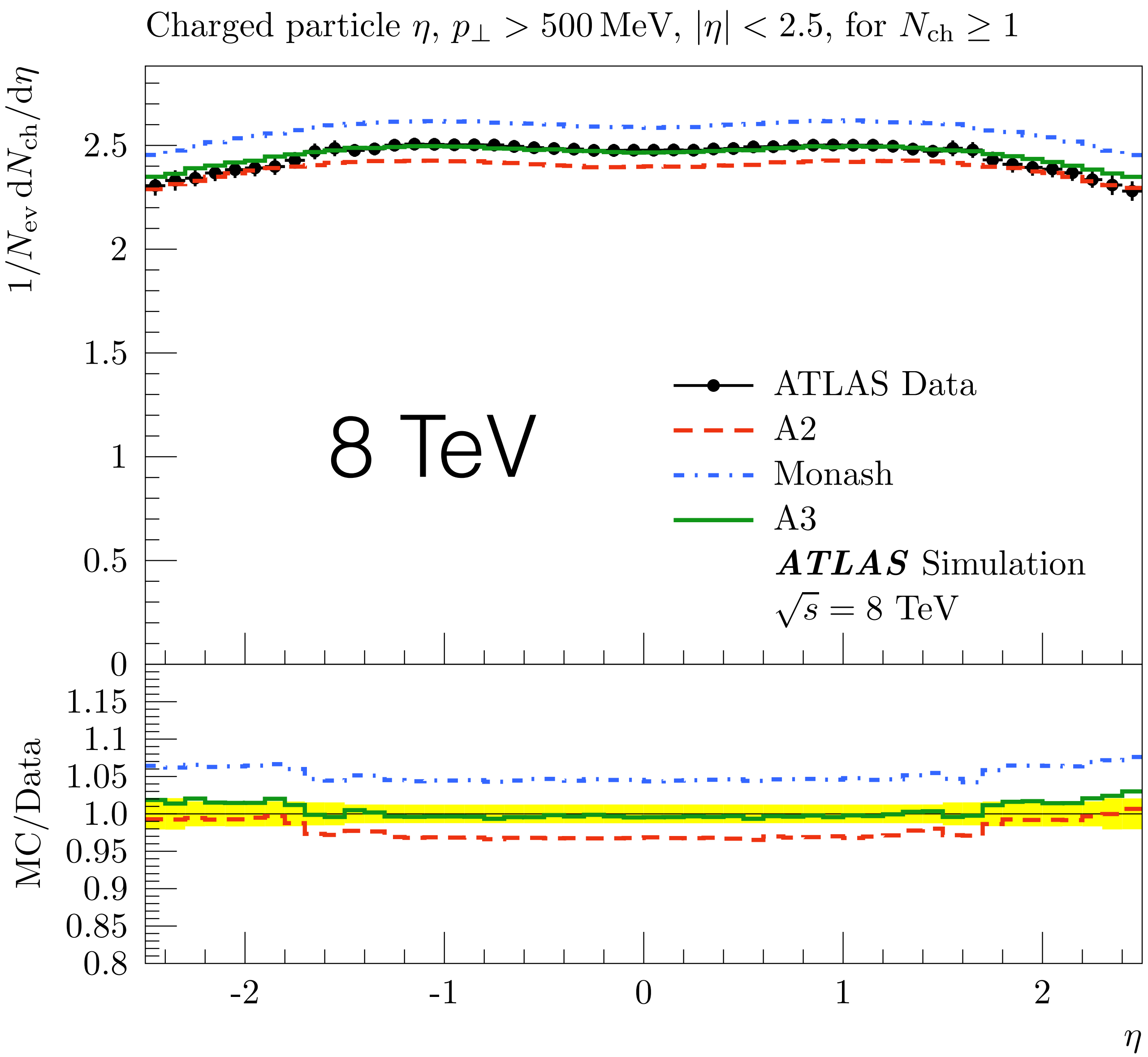




charged particle multiplicities with COM energy

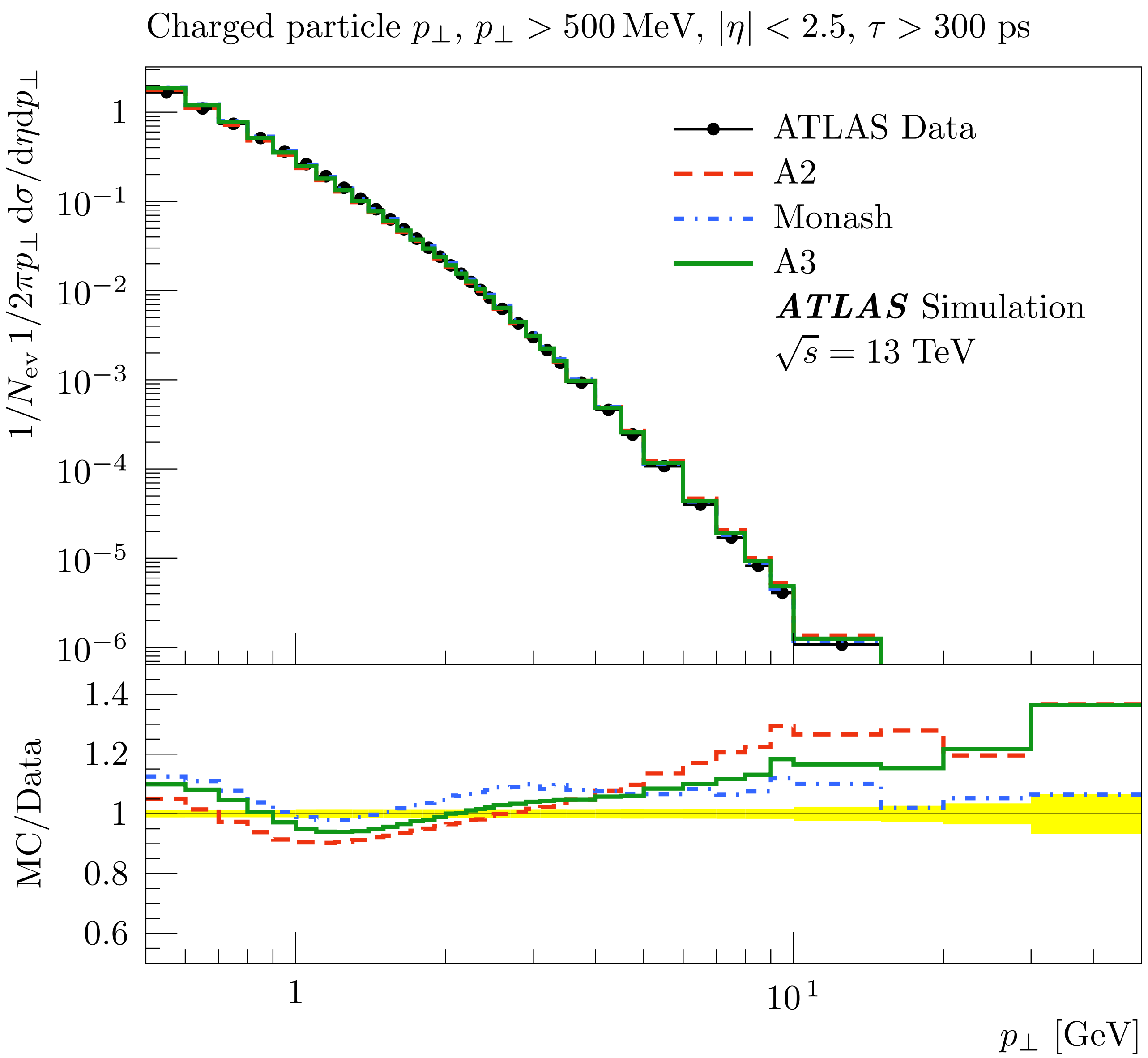
ATL-PHYS-PUB-2016-017

modeling pileup



ATL-PHYS-PUB-2016-017

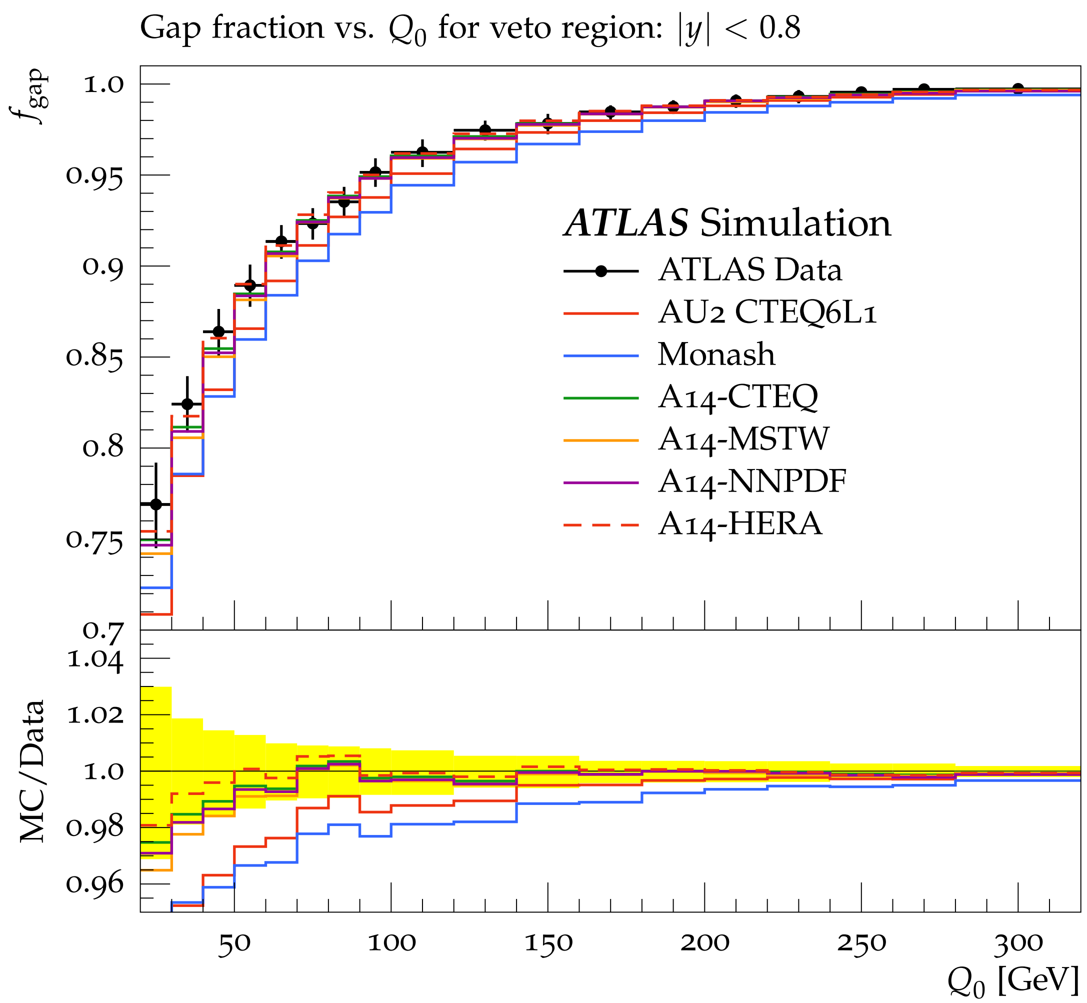
things are not perfect but improving!

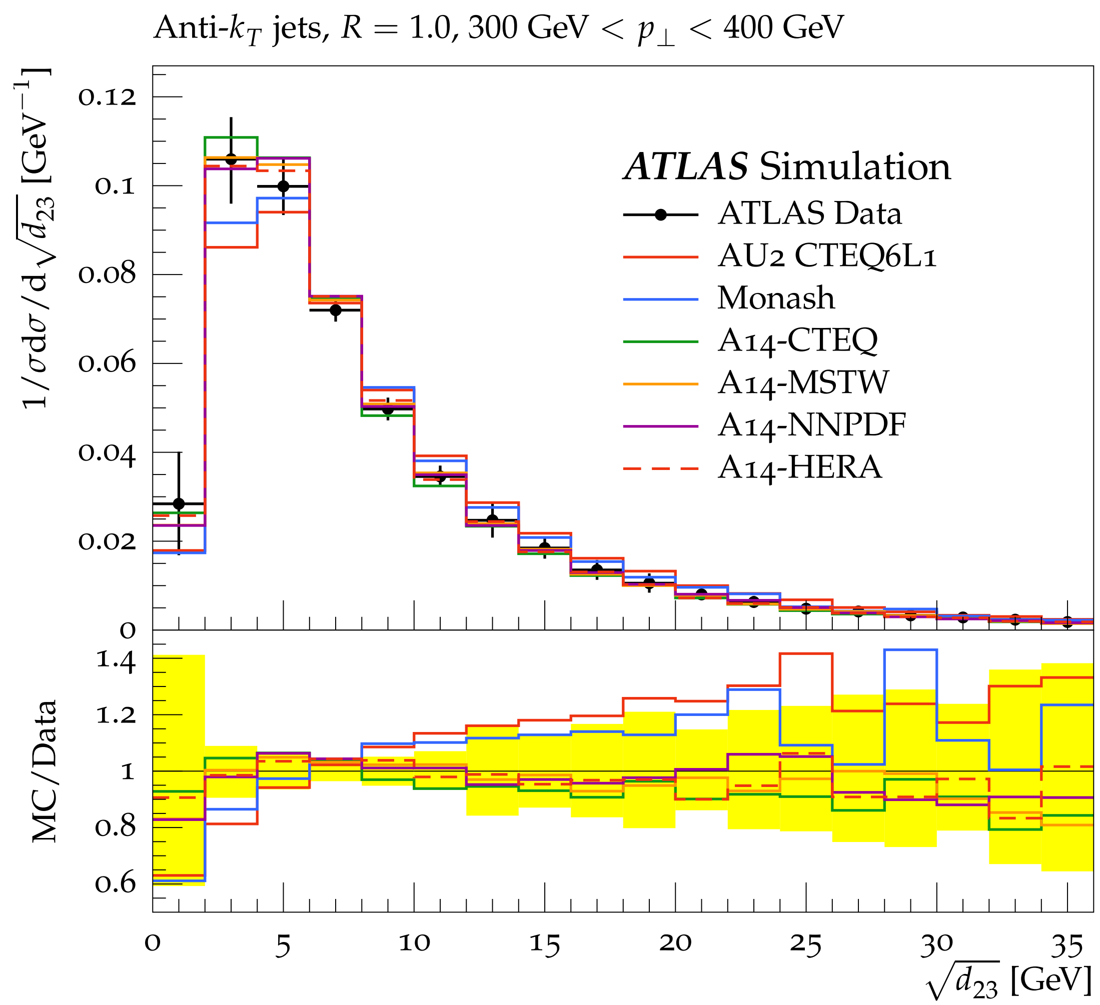


A14 pythia tune

ATL-PHYS-PUB-2014-021

- nominal **pythia8** samples in ATLAS use the **A14 tune**.
- this tune is based on ATLAS data measurements sensitive to the modeling of **underlying event**, **jet substructure**, and **jet activity above the lowest-order process**.
- can have a **strong impact on flavor-tagging modeling**...





heavy hadron production in MC

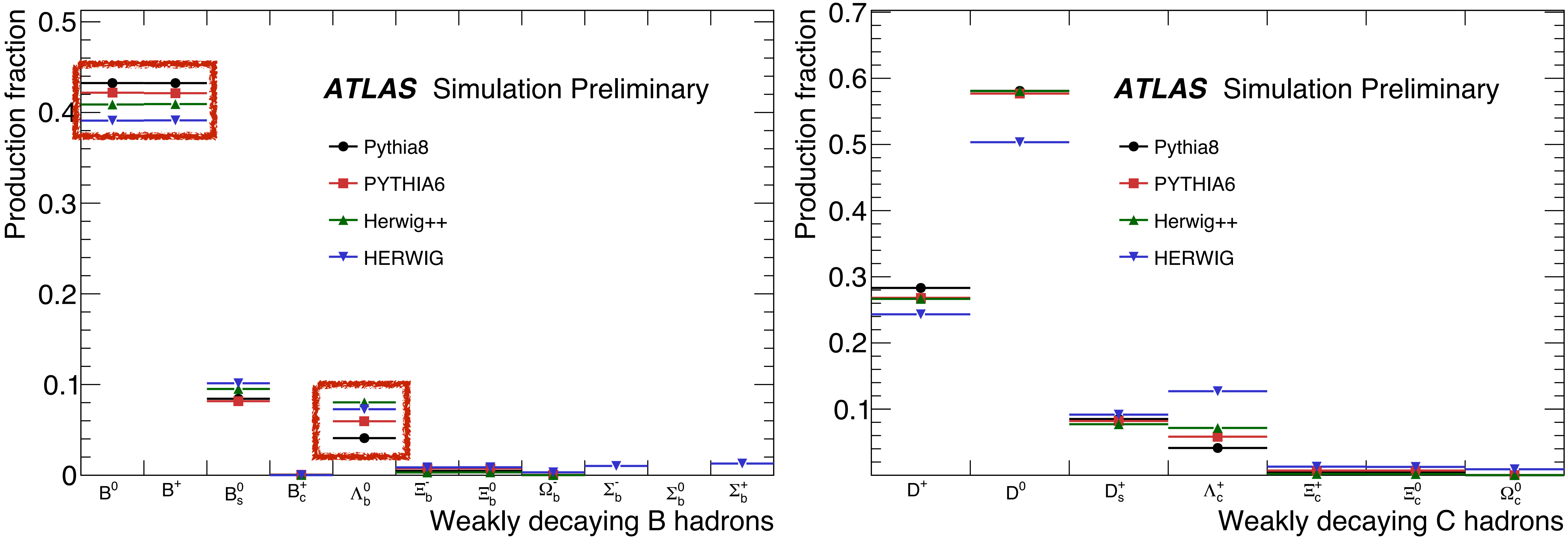
ATL-PHYS-PUB-2014-008

Species	Pythia8	Pythia6	Herwig++	Herwig	World Average [9]
B^+	43.2	42.1	40.9	39.1	40.2 ± 0.7
B^0	43.2	42.2	40.9	39.1	40.2 ± 0.7
B_s^0	8.4	8.2	9.5	10.1	10.4 ± 0.6
Baryons	5.1	7.3	8.6	11.7	9.3 ± 1.5

- the **ATLAS EvtGen note** has extensive studies of heavy hadron production and decay fractions observed in various **ATLAS parton shower+hadronization generator setups**
- there is reasonable **but far from perfect** agreement between these setups on heavy hadron production fractions
- side Q: **could we better constrain these** with LHC data (LHCb?)

heavy hadron production in MC

ATL-PHYS-PUB-2014-008

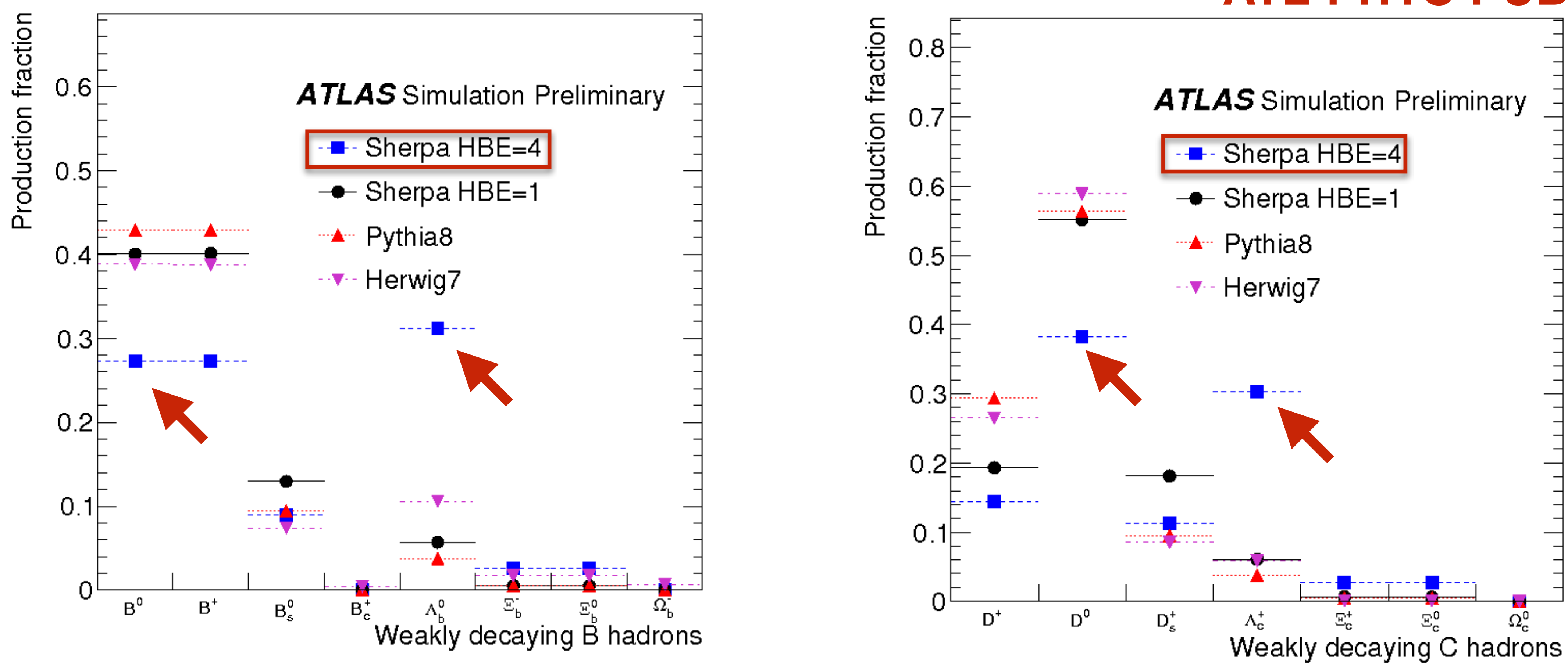


similar story, but in graphical form

fHERWIG was very far off from the rest (?)

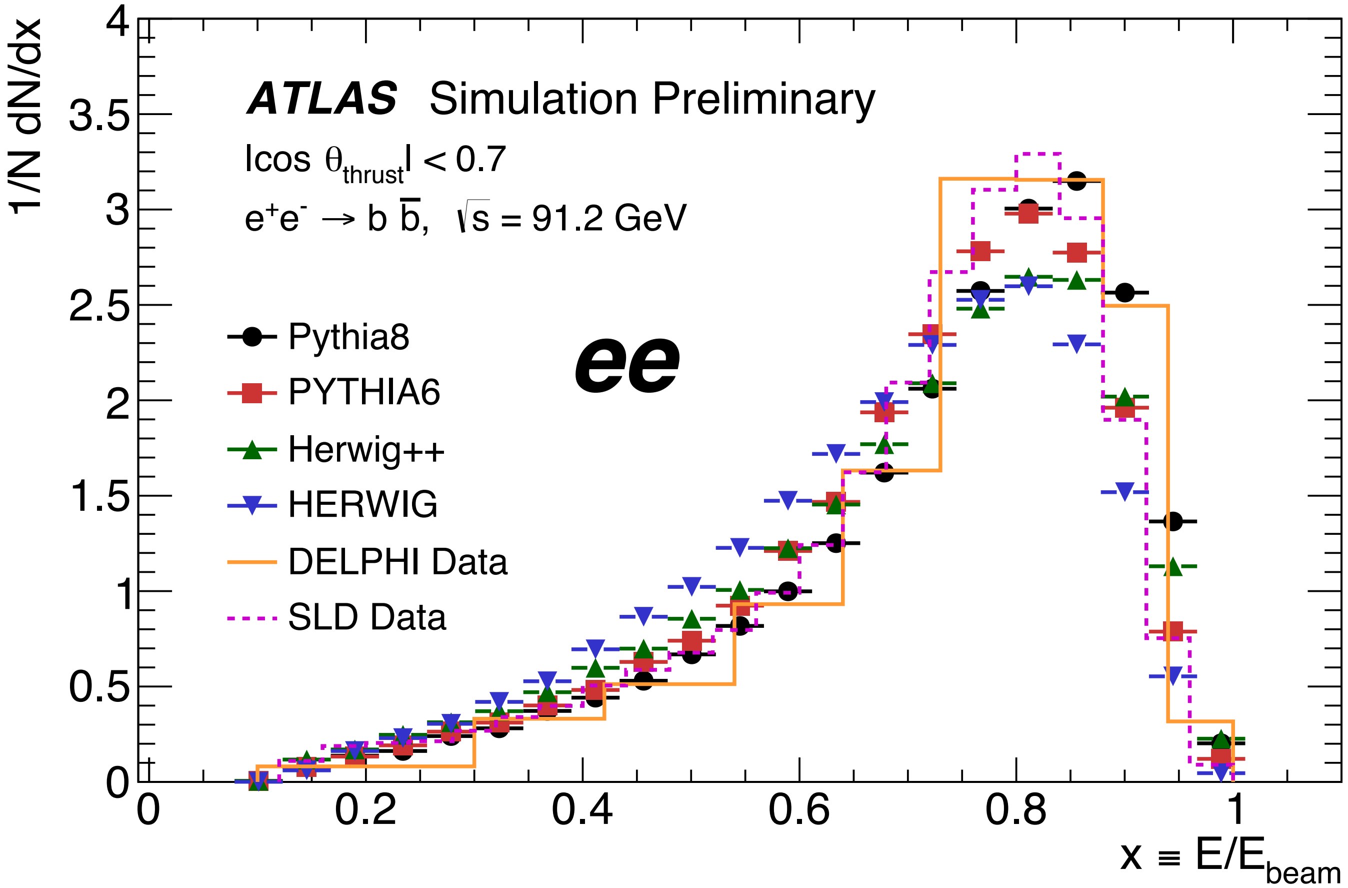
heavy hadron production in MC

ATL-PHYS-PUB-2017-007

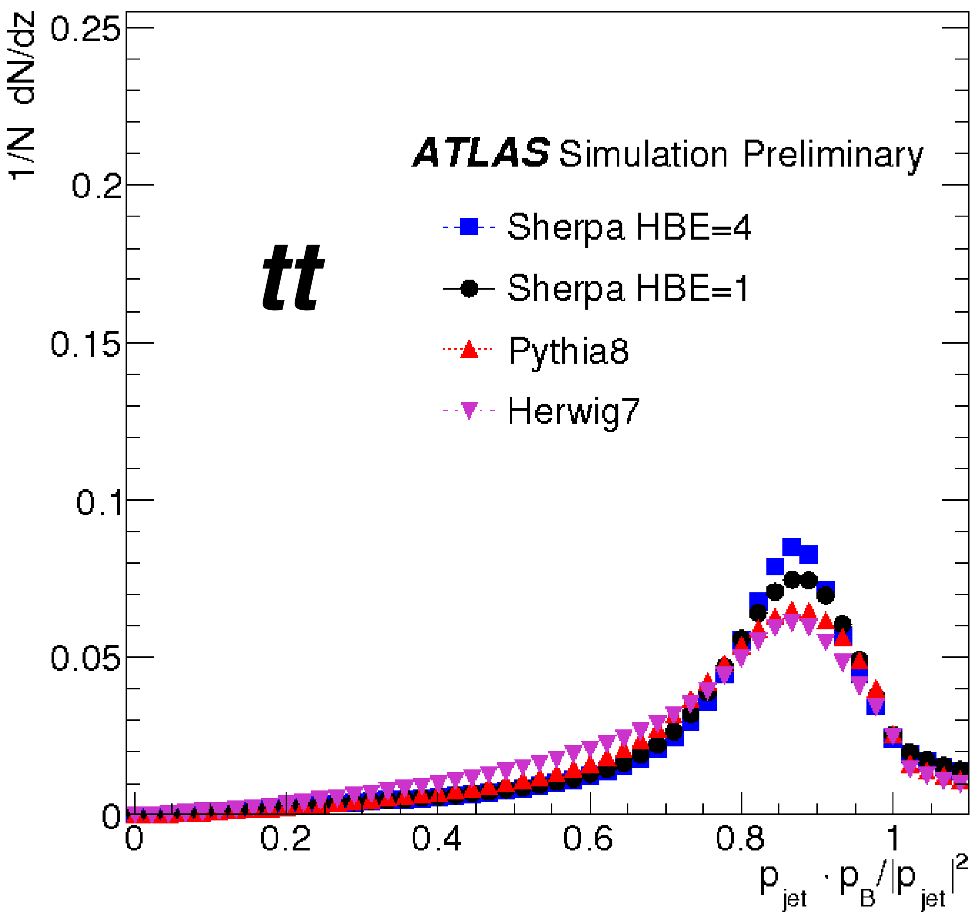


- there have, however, been some **bumps** in the road...
- for instance, we found some of the **Sherpa parameters** related to heavy hadron production were not ideal in some of our samples, and the *b*-tagging efficiency and mistag rates **needed to be corrected**. (more later)
- even with "better" choice of parameters, Sherpa still disagrees with other hadronization generators on charm production fractions.

heavy hadron production in MC



ATL-PHYS-PUB-2014-008



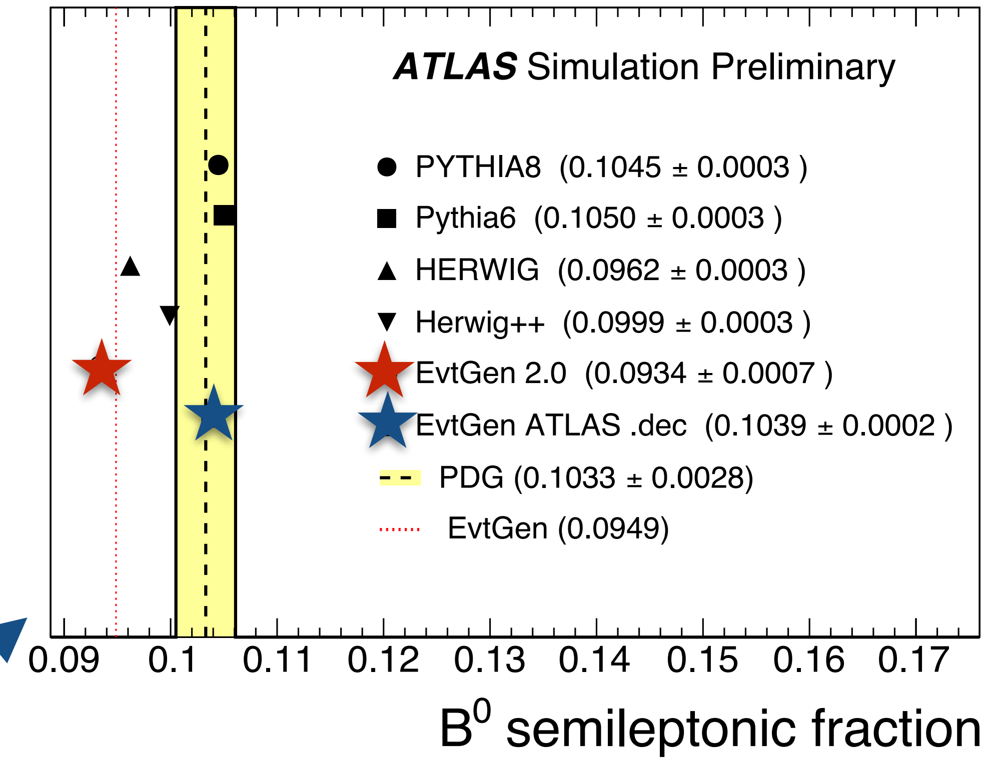
ATL-PHYS-PUB-2017-007

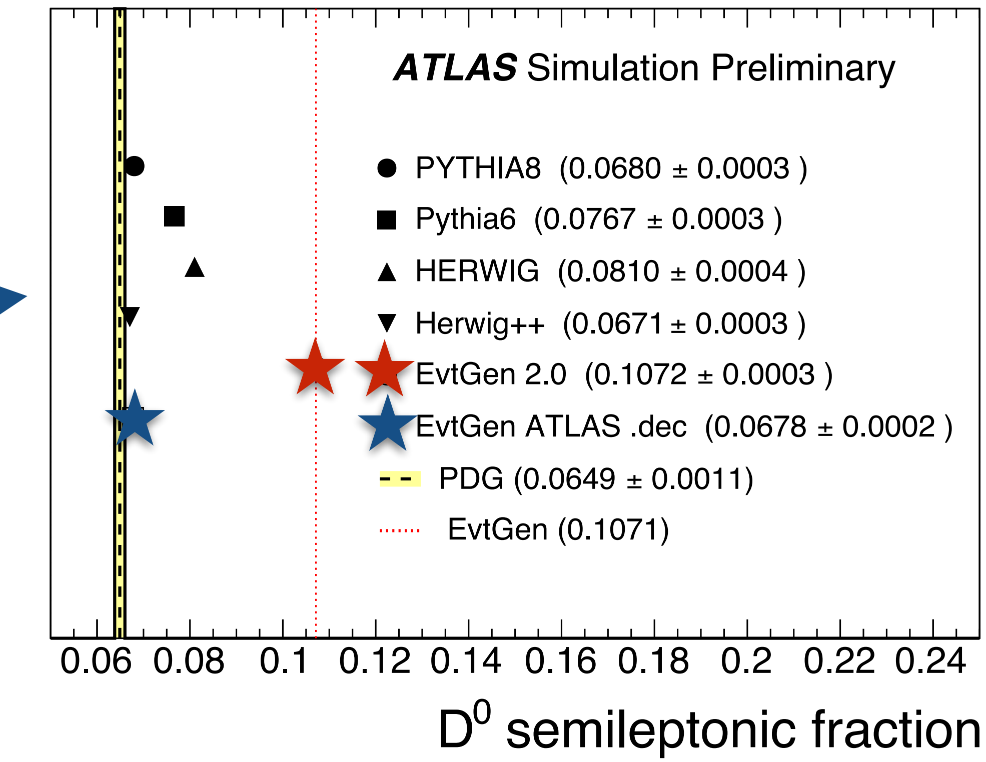
- **heavy quark fragmentation** is also clearly important to model well for the best flavor-tagging performance, jet response in the detector, etc...
- ... and it's not certain that we do this properly. (**more measurements sensitive to this**?)

EvtGen

ATL-PHYS-PUB-2014-008

- **EvtGen** is used extensively in ATLAS to yield **uniform heavy hadron decay fractions** and **lifetimes** in pythia and herwig.
- ***production fractions are not made uniform!***
- NB: ATLAS **Sherpa samples are not decayed with EvtGen**, so some differences are expected due to decay table discrepancies.
- NB II: some changes were made w.r.t. the default EvtGen tables.





correcting the MC predictions

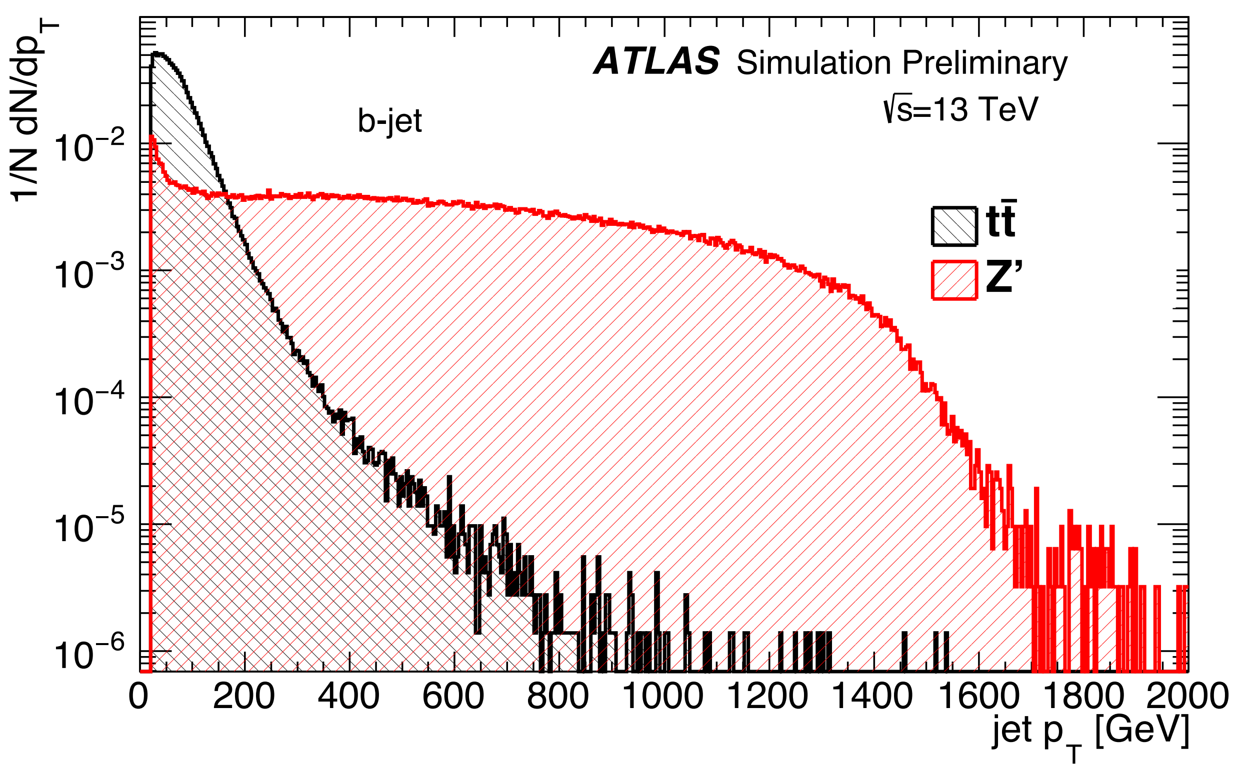
- in order to account for **differences in fragmentation, hadron production fractions, decay fractions**, etc...
- ... **each parton shower+hadronization generator has its own effective efficiency scale factor for each** jet flavor.
- this can even be separated by minor generator version if necessary.
  - e.g. Sherpa may change it **radiation patterns in top-quark decays** or **heavy hadron production fractions**, so a **different b-jet tagging efficiency** is observed.
- or, for instance, the way we currently **match aMC@NLO** and **PowHeg** to **pythia8** is not consistent, so **these setups predict different tagging efficiencies**.
- for *b*-jets the MCs agree on the efficiency to within 1-5%, depending on the kinematic region
- for *c*- and light-flavor jets, these **corrections can be quite large (>50% for light).**

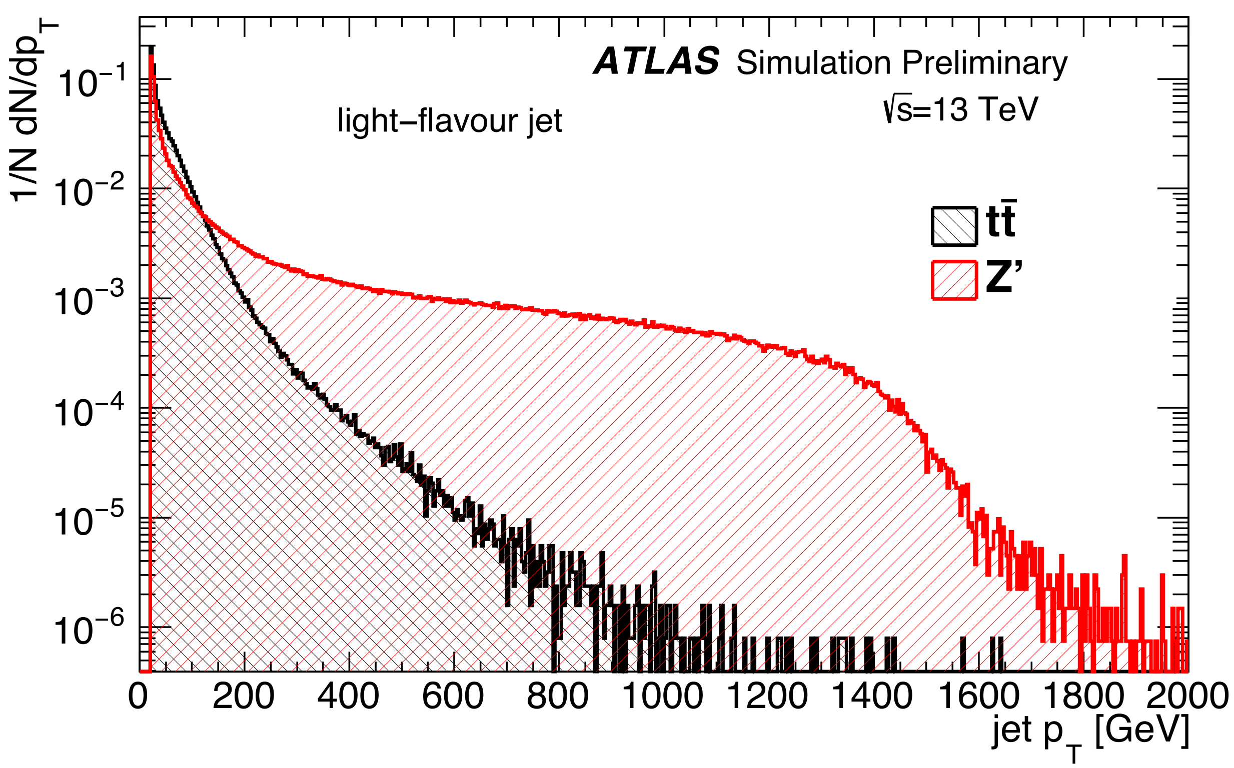
MCs for algorithm training

for training the *b*-tagging algorithms we use a combination of jets from *tt* and a *Z'* -> hadrons

- PowHeg+Pythia8 *tt* @NLO for jets with pT < 250 GeV
- pythia *Z'* flattened in pt for jets with pT > 250 GeV
- (loosely speaking)

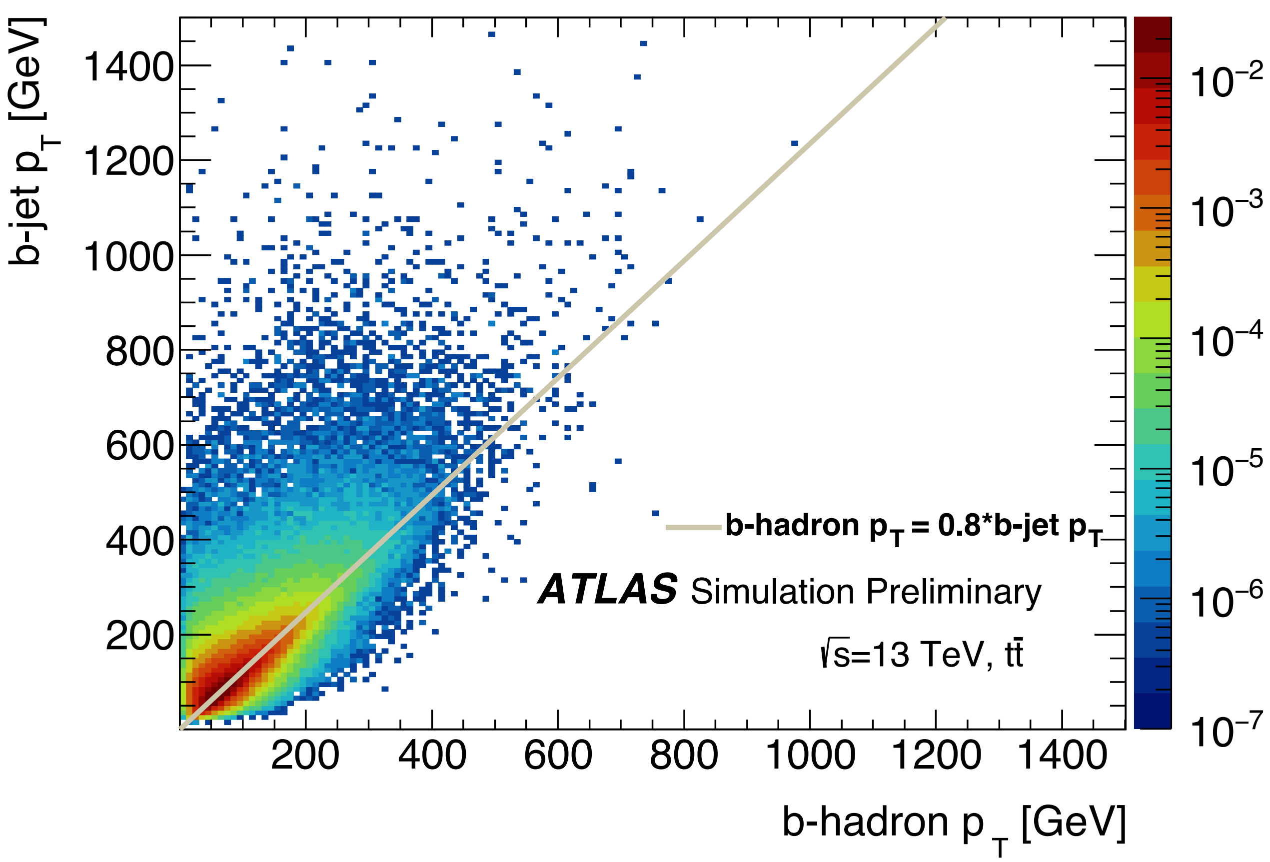
these samples are reweighted to have the same pT and eta distributions for signal and background jets

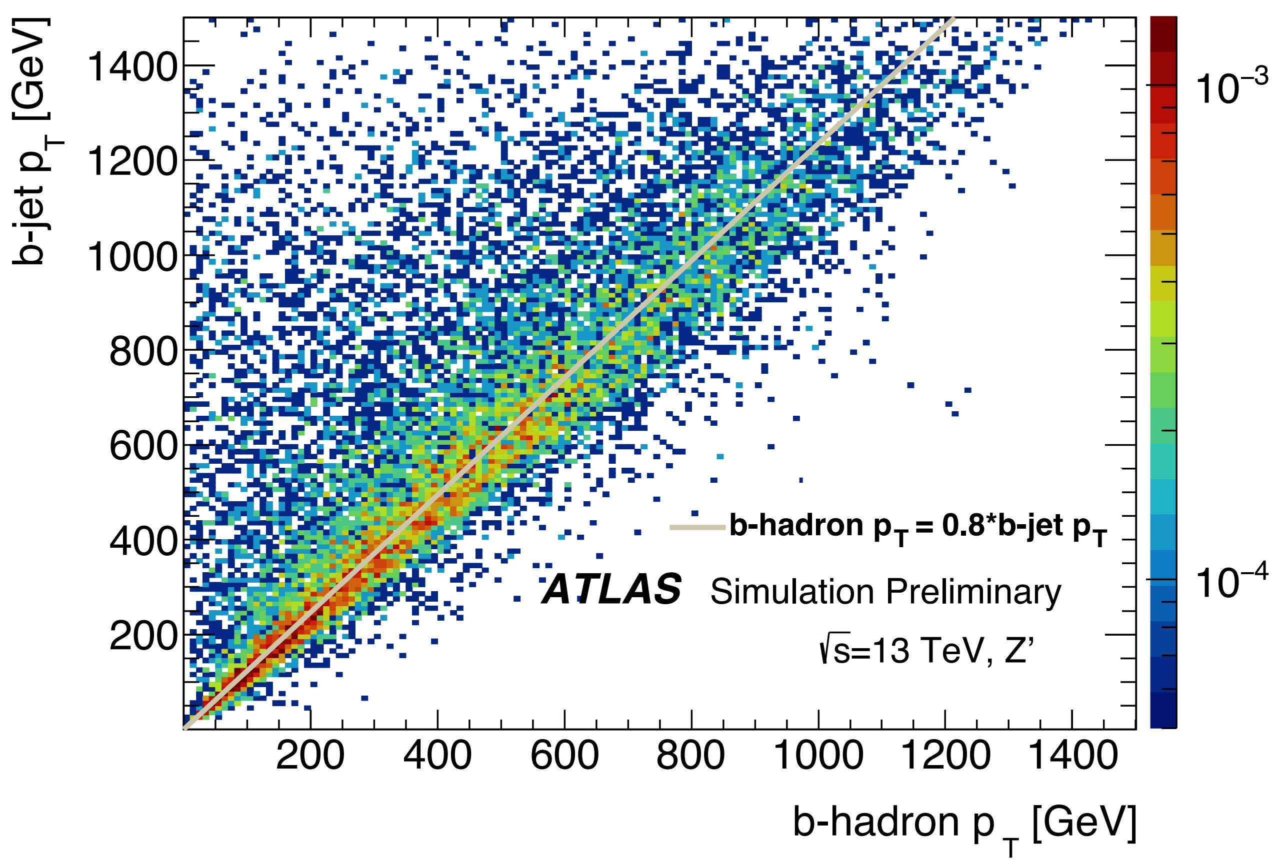




MCs for algorithm training

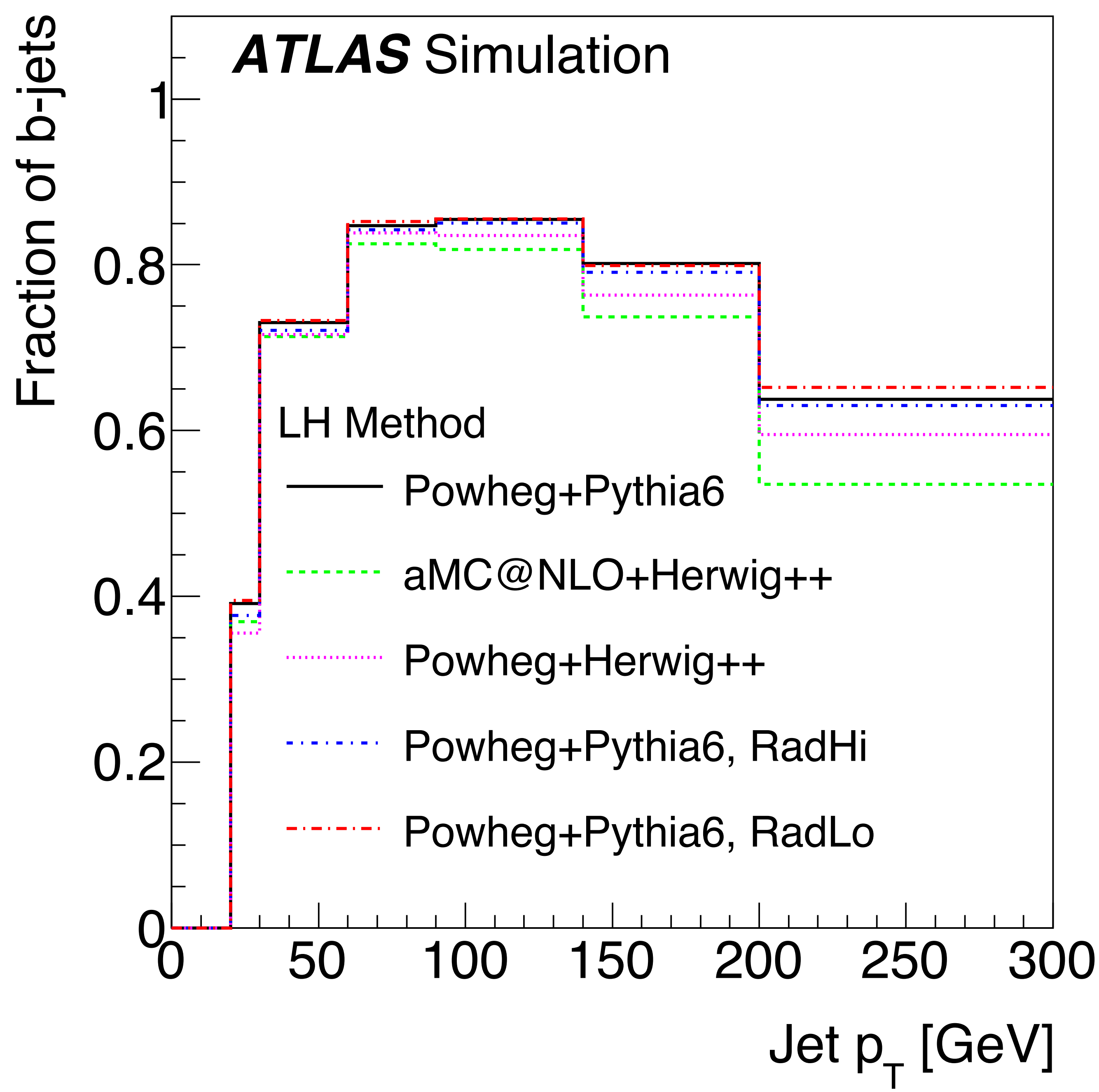
in fact the fragmentation functions can be quite different as a function of the *b*-jet pT, but we believe high-pT jets from the *Z'* sample are more indicative of what physics analyses are targeting.





MC for *tt*-based calibrations

- some of the dominant uncertainties in the *tt*-based calibrations for *b*- and *c*-flavor jets come from our **inability to model *tt* events precisely**
- for the *b*-jet calibrations, there is significant **disagreement between generators** on the **fraction of light-flavor background jets** in the data...
- ... leading to **substantial systematic uncertainties**.
- we have developed **data-driven corrections** that reduce these uncertainties significantly.
- the charm calibration also has leading uncertainties from *tt* modeling.

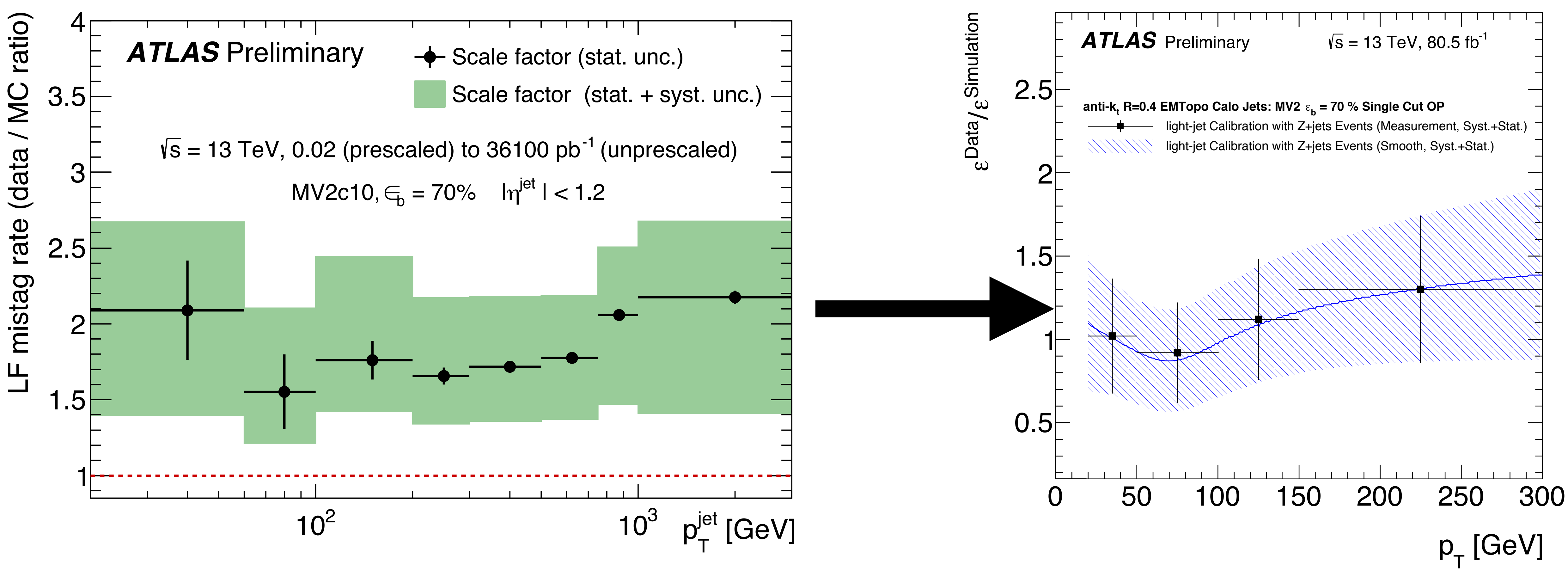


JHEP 08 (2018) 89

previously observed mismodeling in light-flavor jets

light flavor mistag rate SFs were ~2 in the past

- complicated to disentangle effects of generator and detector mismodeling
- better material description of the inner tracker has recently brought these SFs significantly closer to unity.
- need to beat down uncertainties before claiming victory here!



conclusions

- monte-carlo event generators are crucial for many physics analyses, and they are by-and-large doing a great job!
- MC events are clearly also crucial for training optimal *b*-taggers.
- despite significant tuning to data from LEP/*b*-factories, etc, there is still disagreement between generators on features relevant for flavor-tagging.
- we manage to correct somewhat these disagreements through generator-specific efficiency SFs.
- we are furthermore bringing in more data-driven corrections in calibration analyses to make ourselves insensitive to mismodeling of ISR, FSR, underlying event, etc.