Jets reconstruction and b-tagging: lessons learned and new strategies

Lorenzo Sestini
INFN-Padova

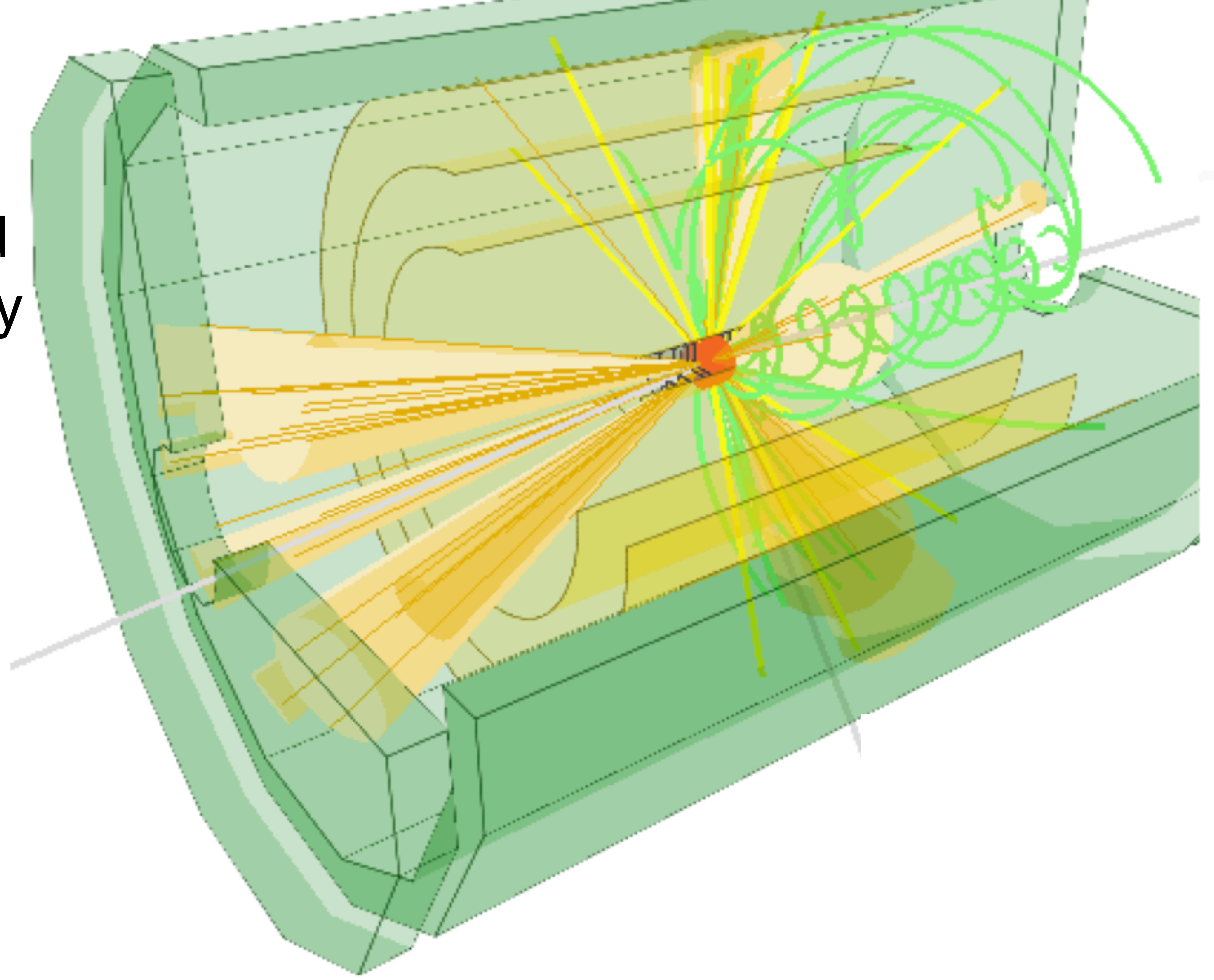
On behalf of the Muon Collider Detector and Physics group

Muon Collider Collaboration - Annual meeting - CERN- 12/10/2022

Jet reconstruction at the Muon Collider

- **Jet reconstruction is fundamental at the muon collider experiment, few examples:**
  - $H \rightarrow b\bar{b}$ main Higgs decay chanel
  - Semileptonic (leptons+jets) final states of ZZ and WW are often the best channels in terms of purity and significance
- **The environment for the jet reconstruction is quite unique compared to hadron colliders and e^+e^- colliders:**
  - The presence of the Beam-Induced Background poses several challenges

$\mu^+\mu^- \rightarrow H(\rightarrow b\bar{b})+X$ reconstructed event without cleaning and analysis cuts

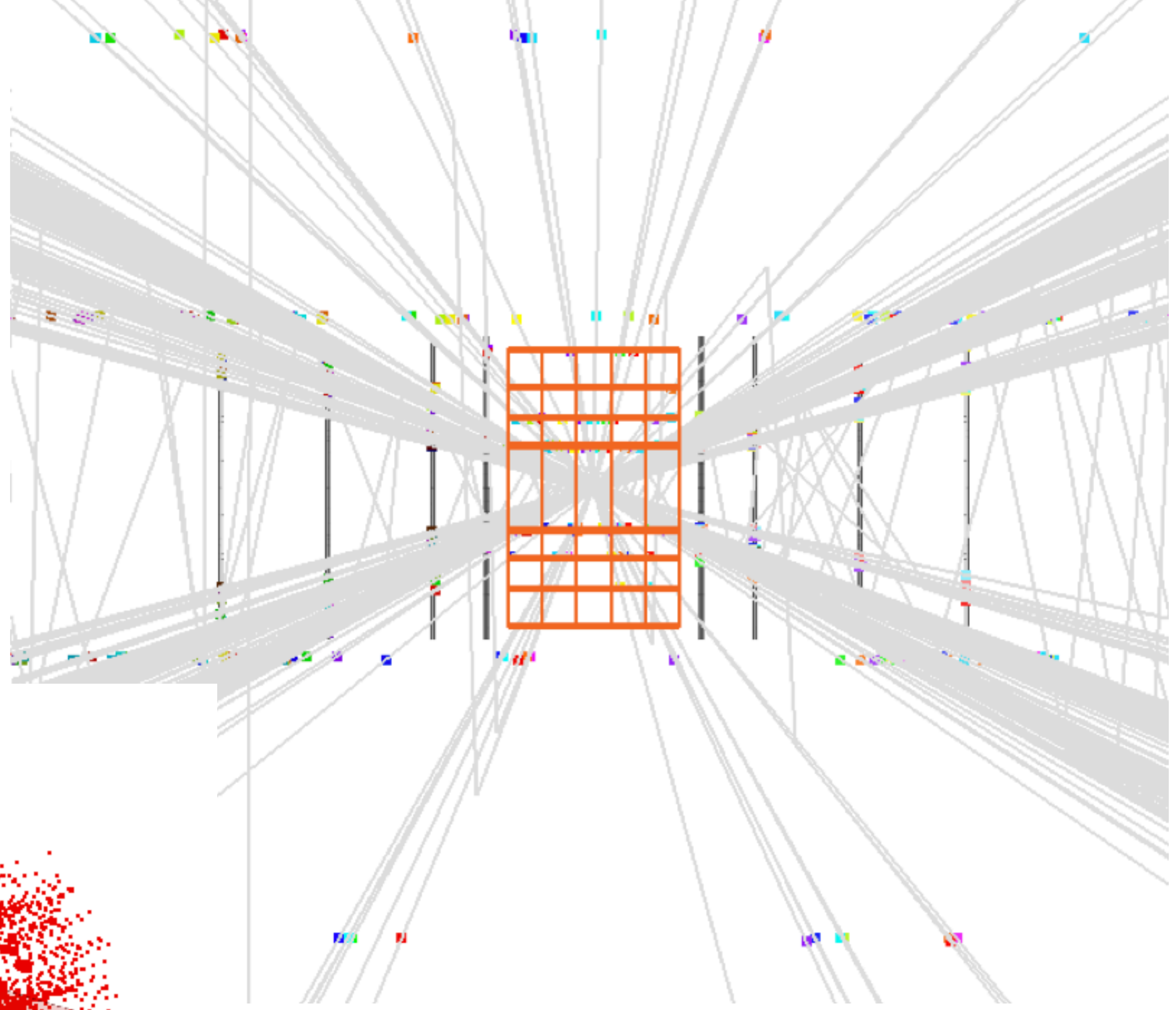


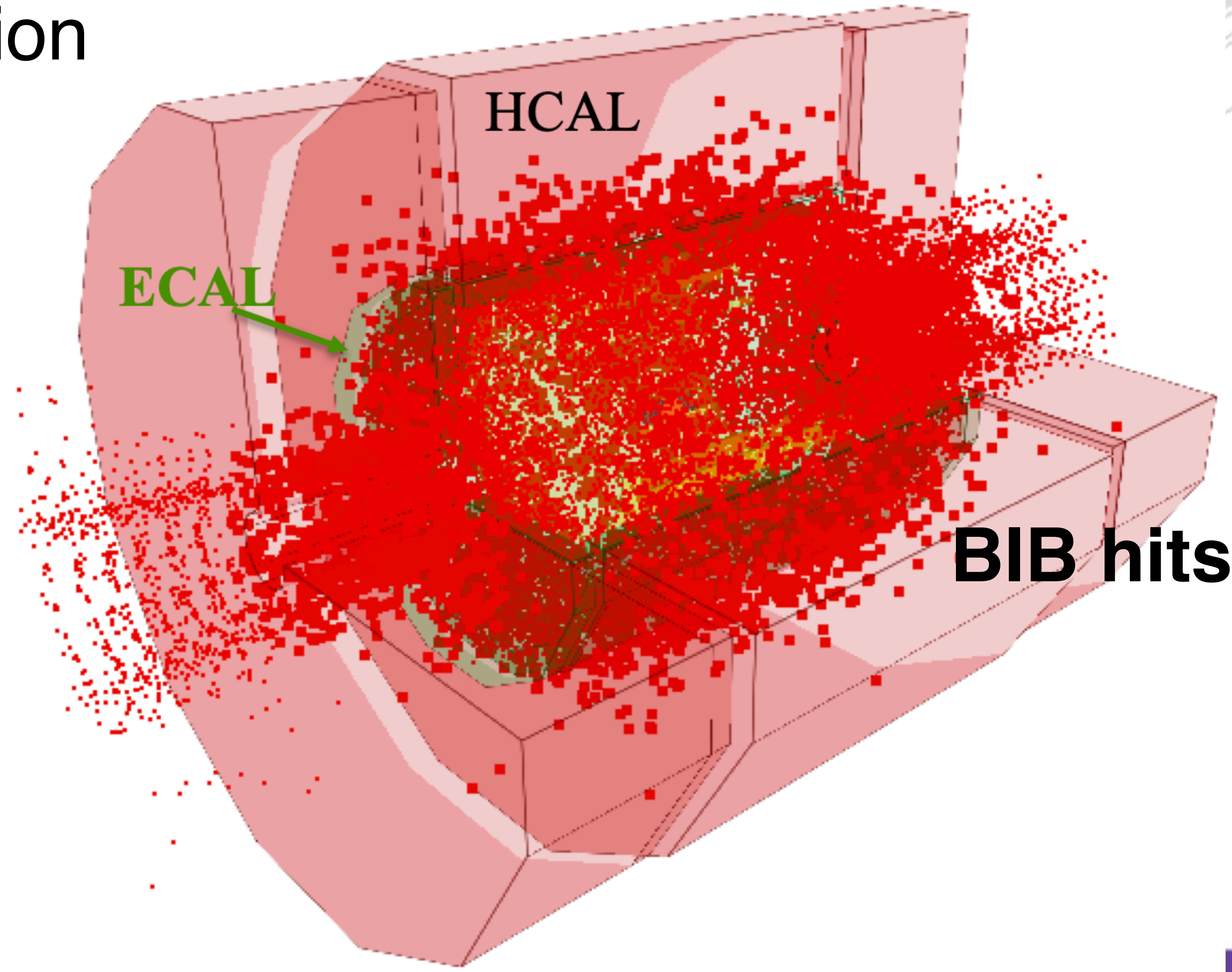
Challenges in jet reconstruction

- **Reconstruction** in this environment is not trivial:
  - the **high hit multiplicity from the BIB in the vertex detector/tracking modules** produces a significant combinatorial problem
  - A diffused BIB background is present in the calorimeters
  - The nozzles, that are fundamental for BIB mitigation, reduce the acceptance in the forward region

Most of the BIB particles are asynchronous with respect to the bunch crossing: timing is crucial for reconstruction

Combinatorial tracks from BIB

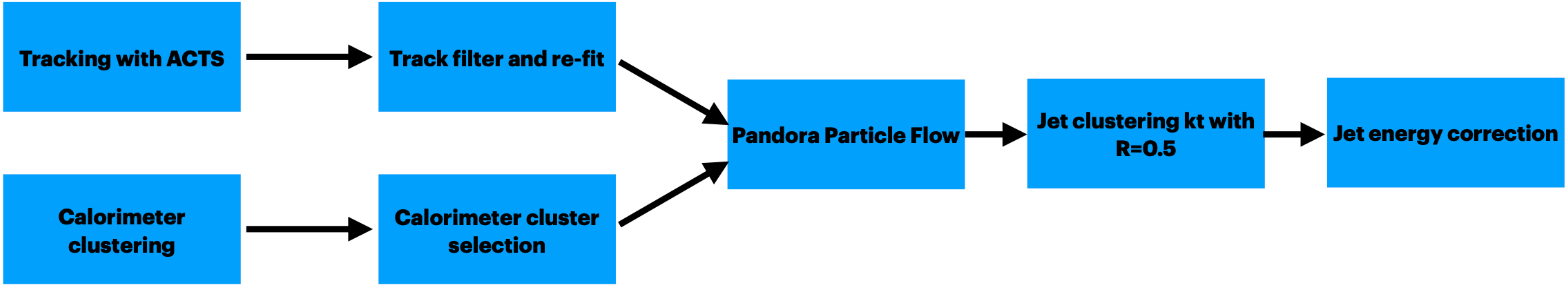




BIB hits in calorimeters

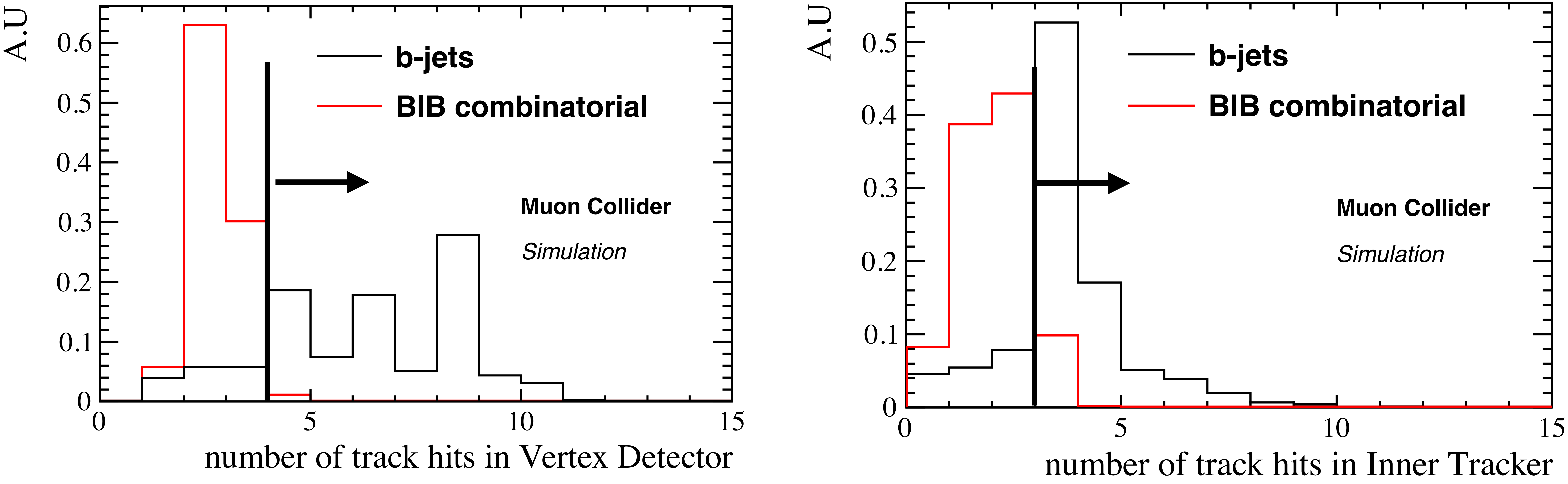
Jet algorithm

- Jet reconstruction algorithm is a “classical” **Particle Flow** algorithm
- **However several strategies should be applied in order to mitigate the effect of the BIB**



BIB mitigation: tracking

- Tracking strategy already described in the previous talk. **ACTS tracking is used for jet reconstruction** and **requirements on the hit time are applied as already discussed**
- In this way **a combinatorial of 100k-200k tracks** per bunch crossing (!) is obtained
- In the context of jet reconstruction, extra requirements on the number of hits in the Vertex Detector and Inner Tracker are applied: combinatorial **is reduced to O(100) tracks**

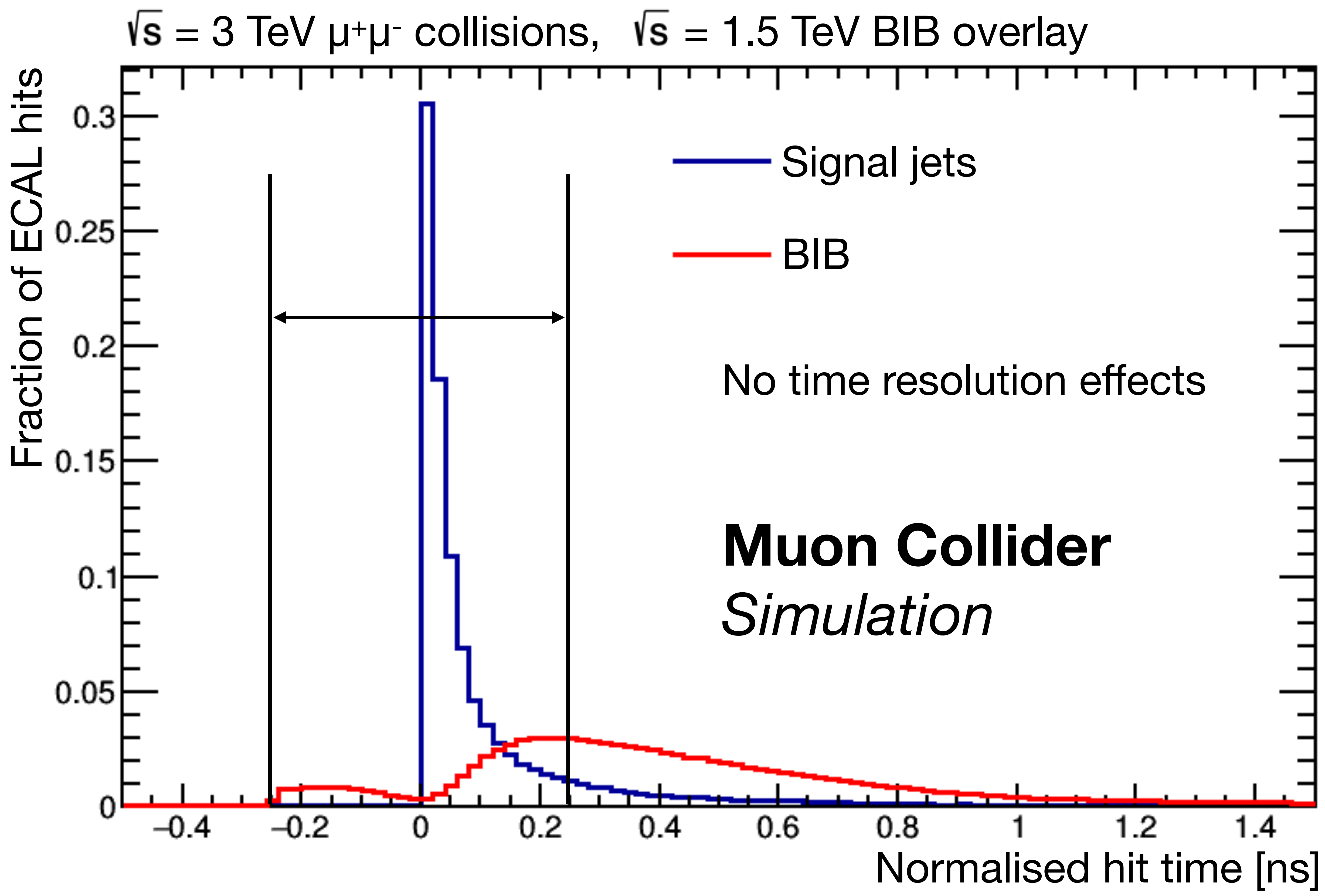


BIB mitigation: calorimeters

- Time requirements are also applied to calorimeter hits

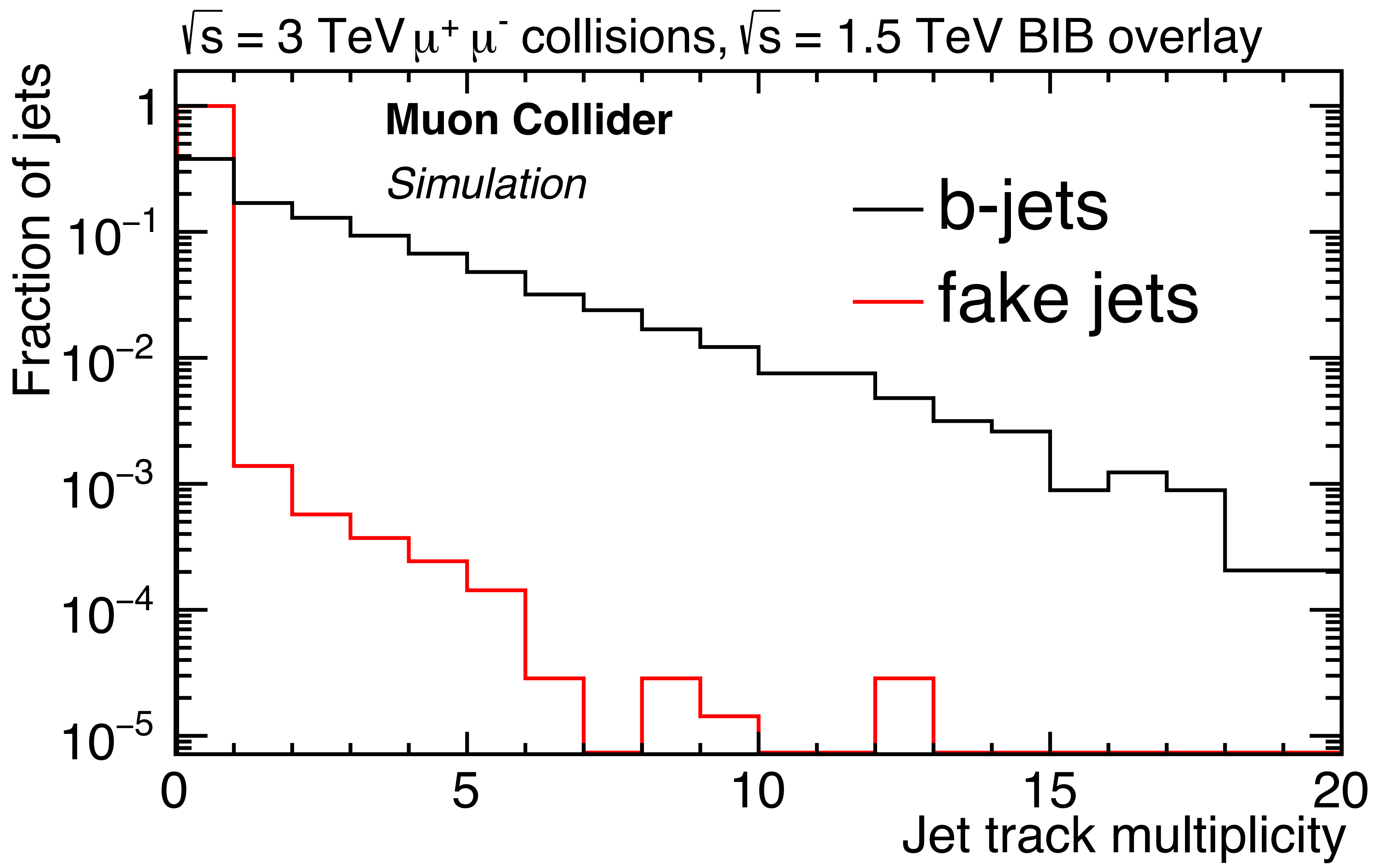
- **A (pretty high) threshold of 2 MeV is applied to the calorimeter hits**: in this way most of the BIB hits are removed.

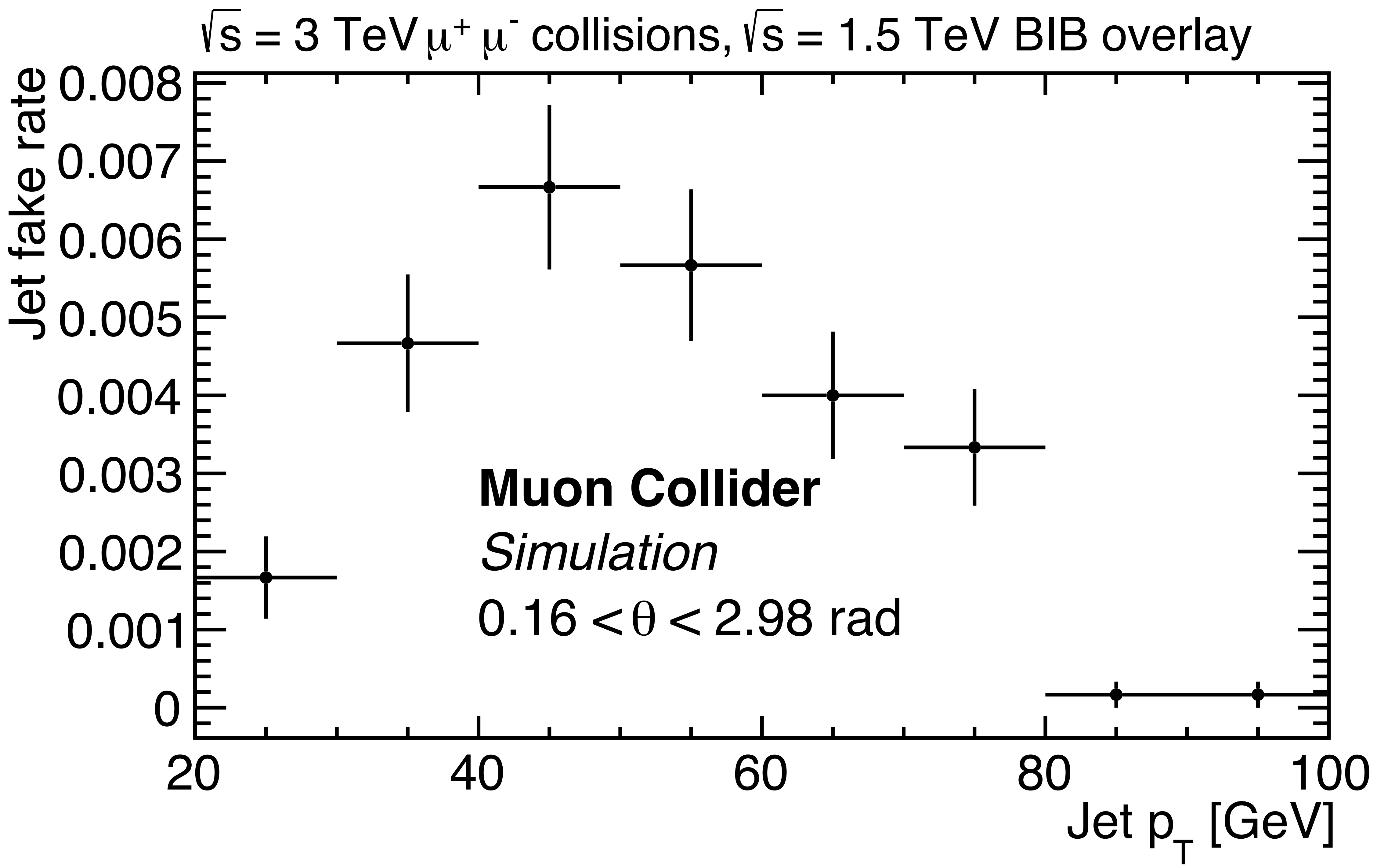
- A discussion on the optimization of this threshold will be done later



BIB mitigation: fake jet removal

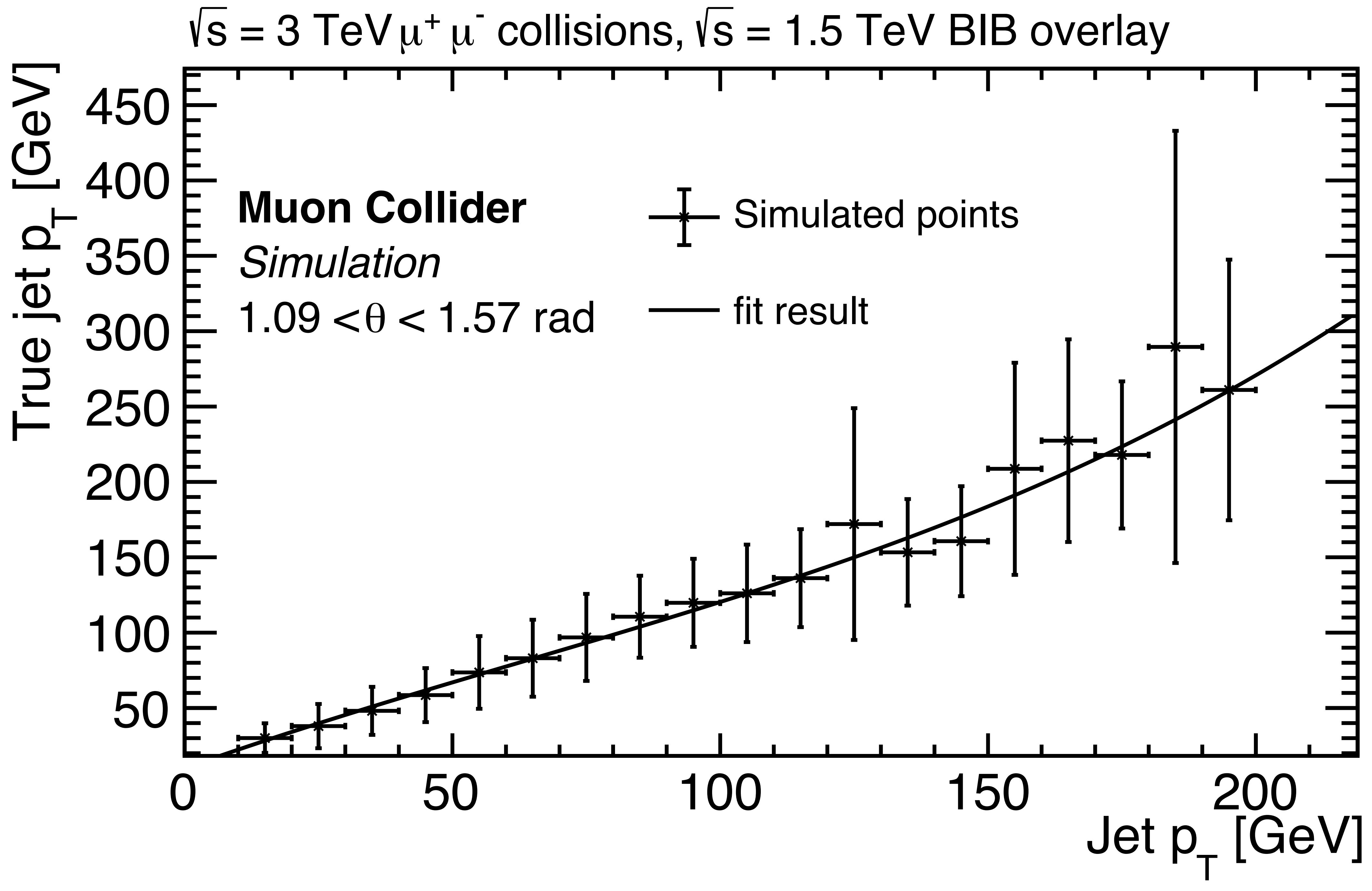
- Of about O(10) fake jets per bunch crossing are found
- Most of them are produced by the BIB calorimeter hits combinatorial
- A **requirement on the number of tracks** (>0) in the jet is applied to reduce them to a negligible level

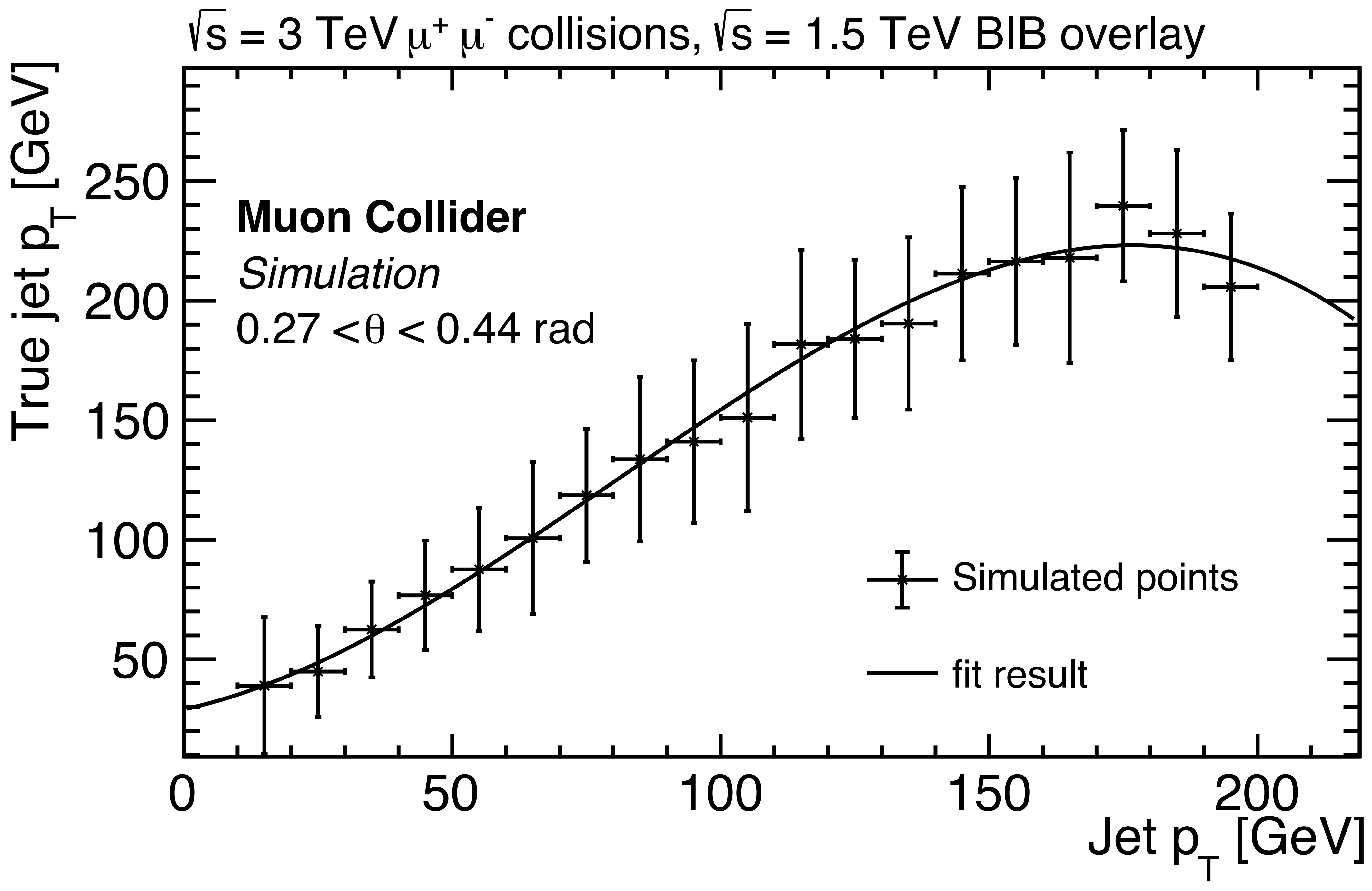




Jet energy correction

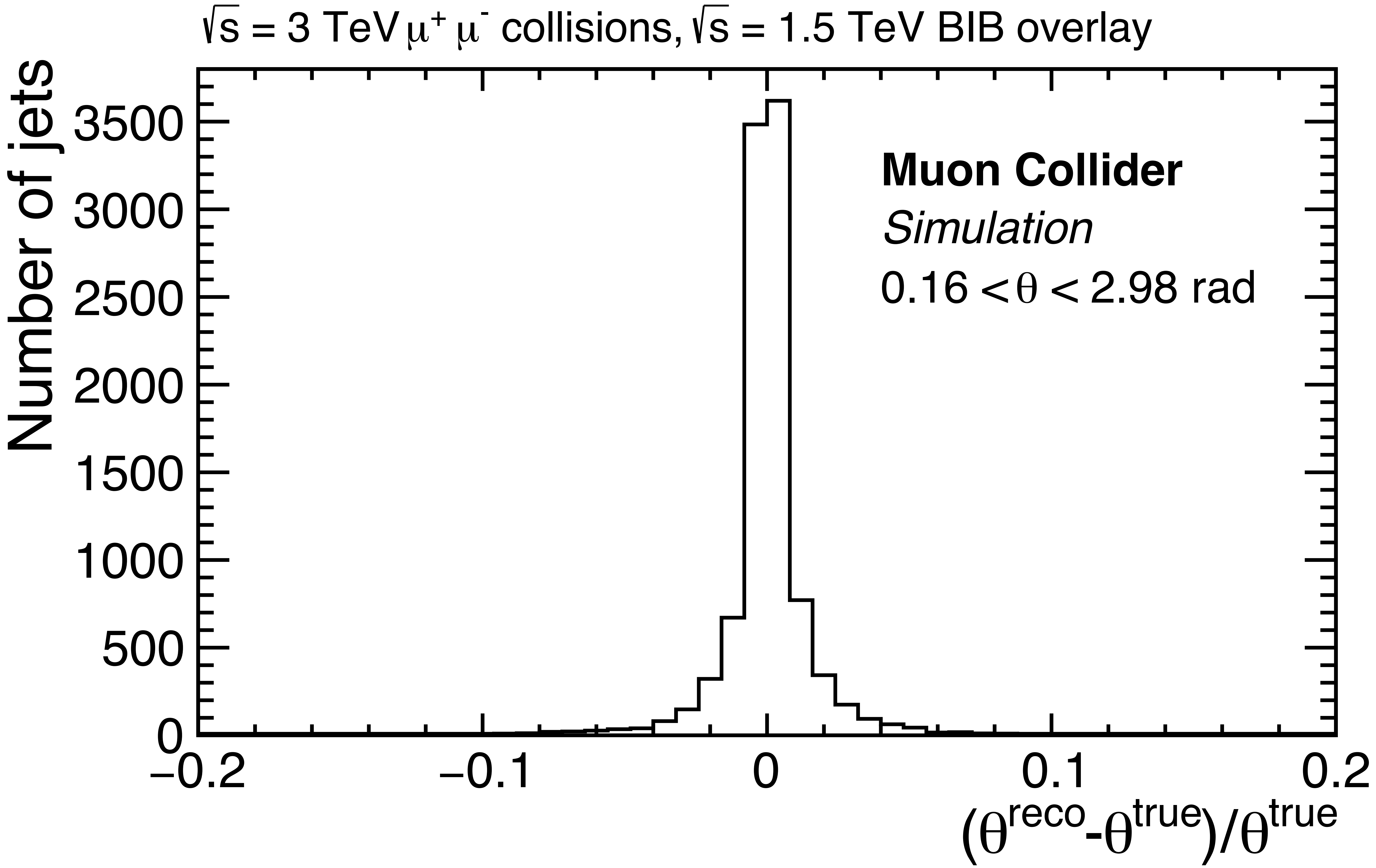
- A jet energy correction is obtained by comparing the reconstructed jets with the truth-level jets
- The jet energy correction **depends on the jet transverse momentum and on the jet direction**





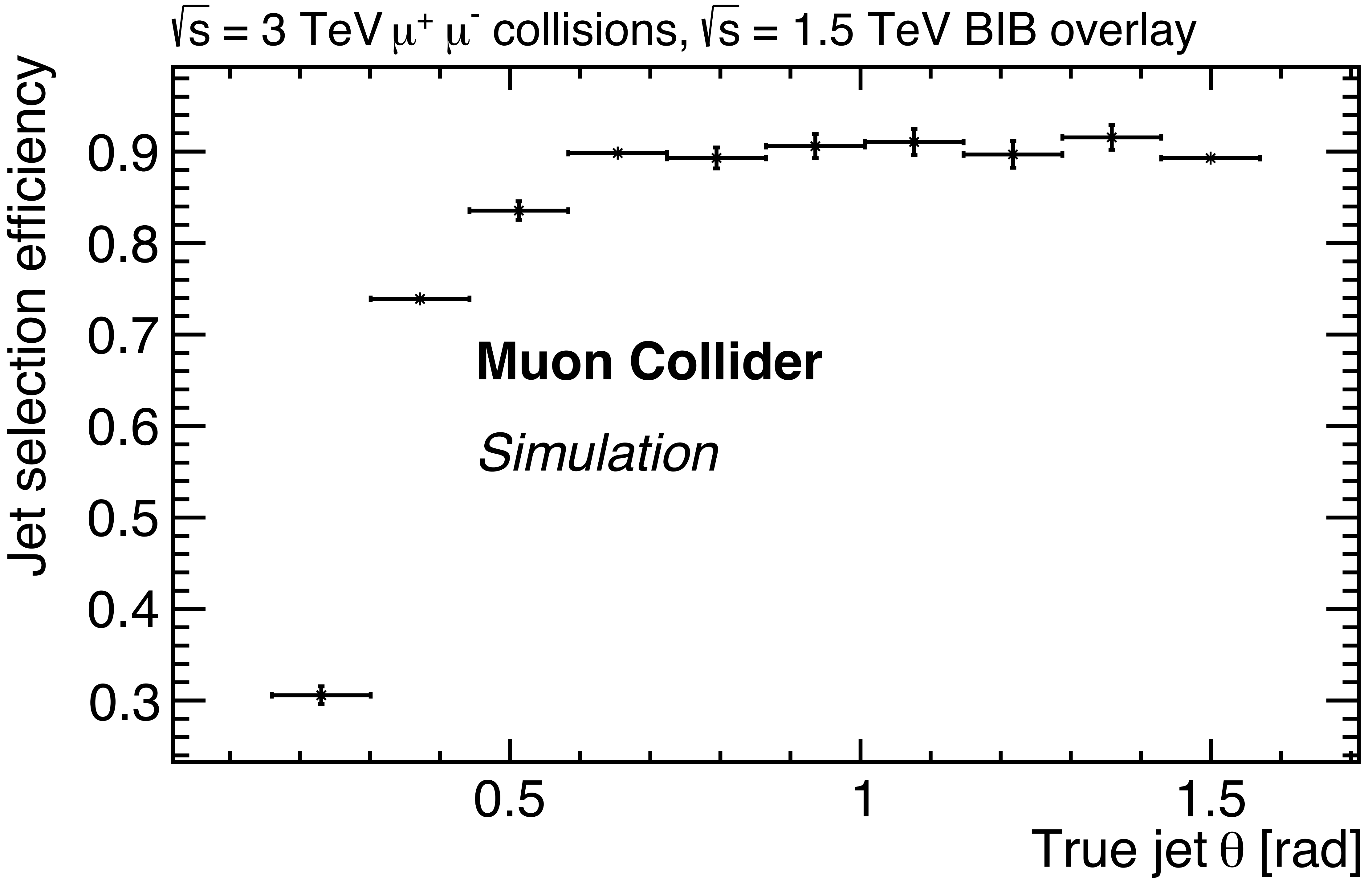
Jet performance: direction

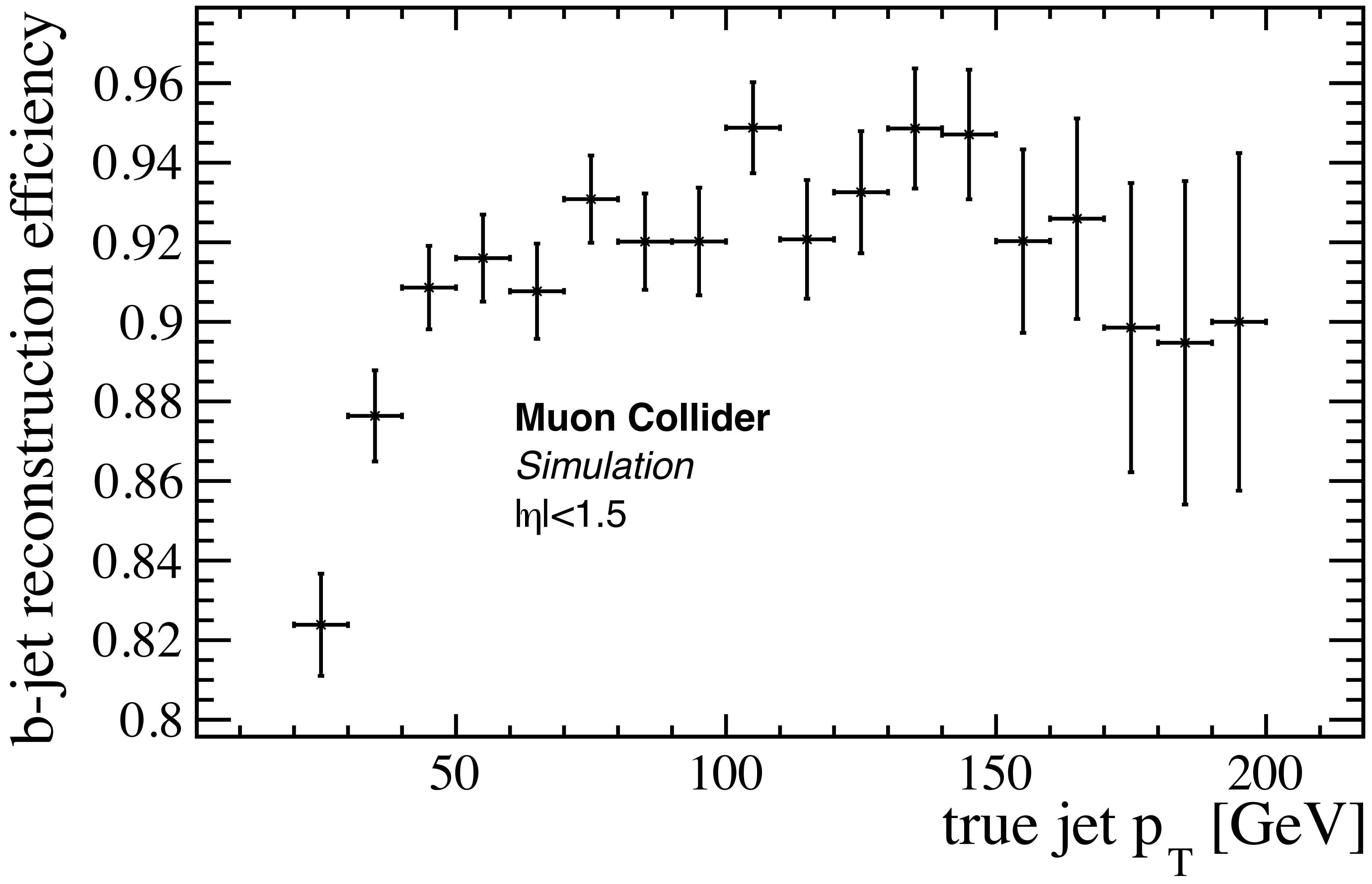
- Jet axis direction (= jet momentum direction) is reconstructed with a precision of 1.7% on θ



Jet performance: efficiency

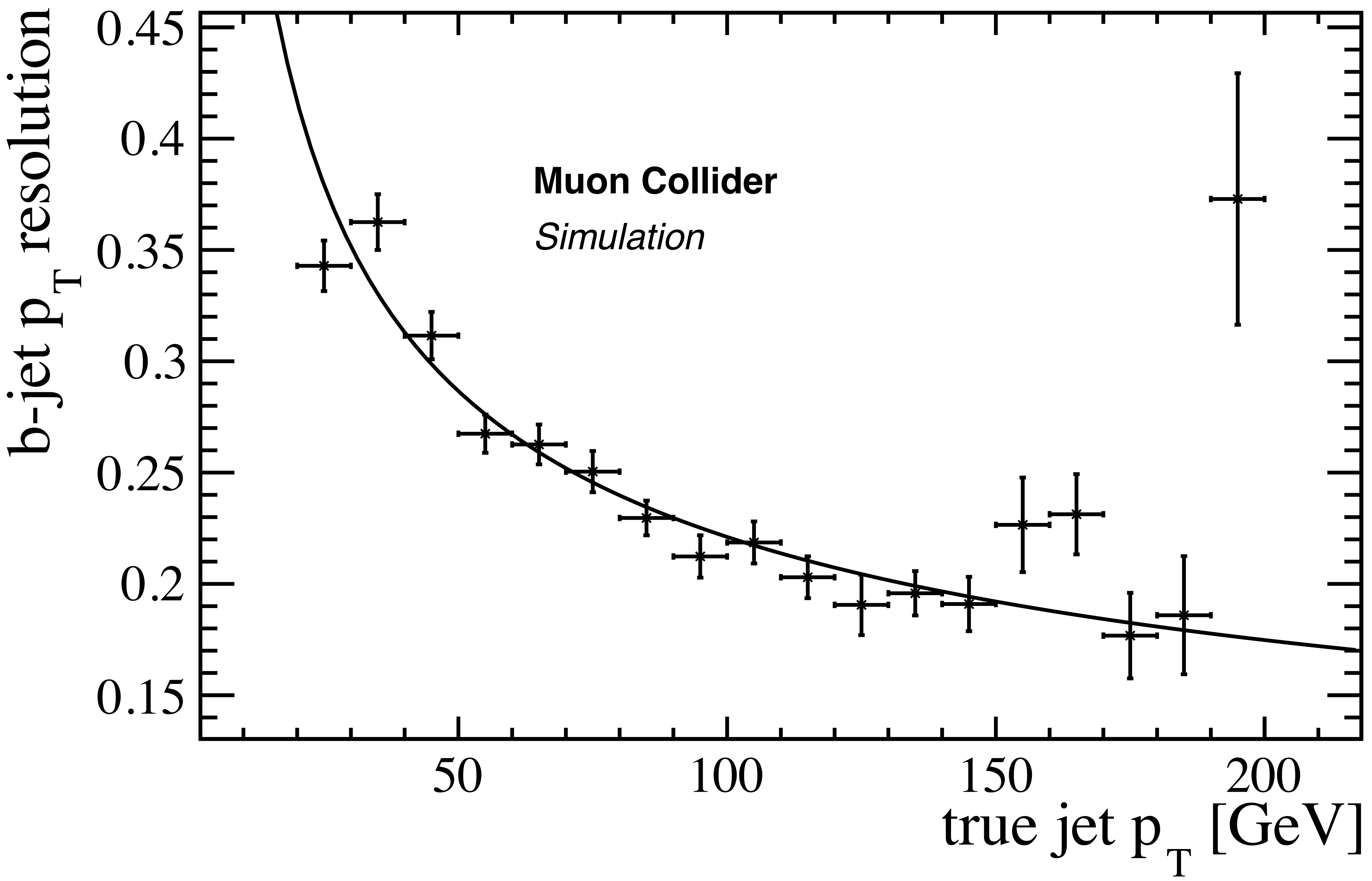
- **Jet reconstruction efficiency is low (30%) in the forward region and high (90%) in the central region**
- **This is mainly due to the fake jet removal:** in the forward region tracks are inefficiently reconstructed or removed by the BIB mitigation, resulting in many signal jets with just calorimeter clusters

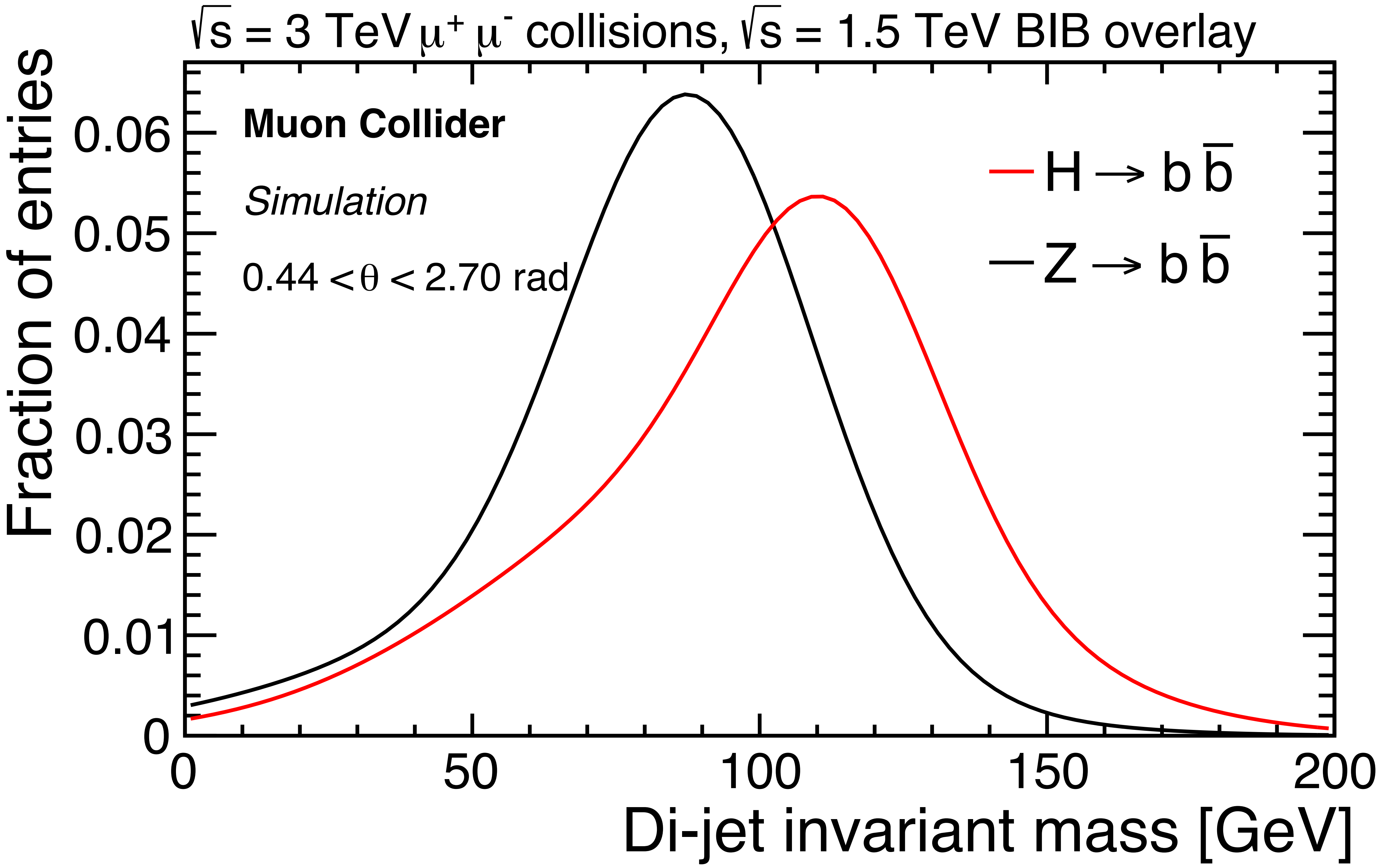




Jet performance: energy resolution

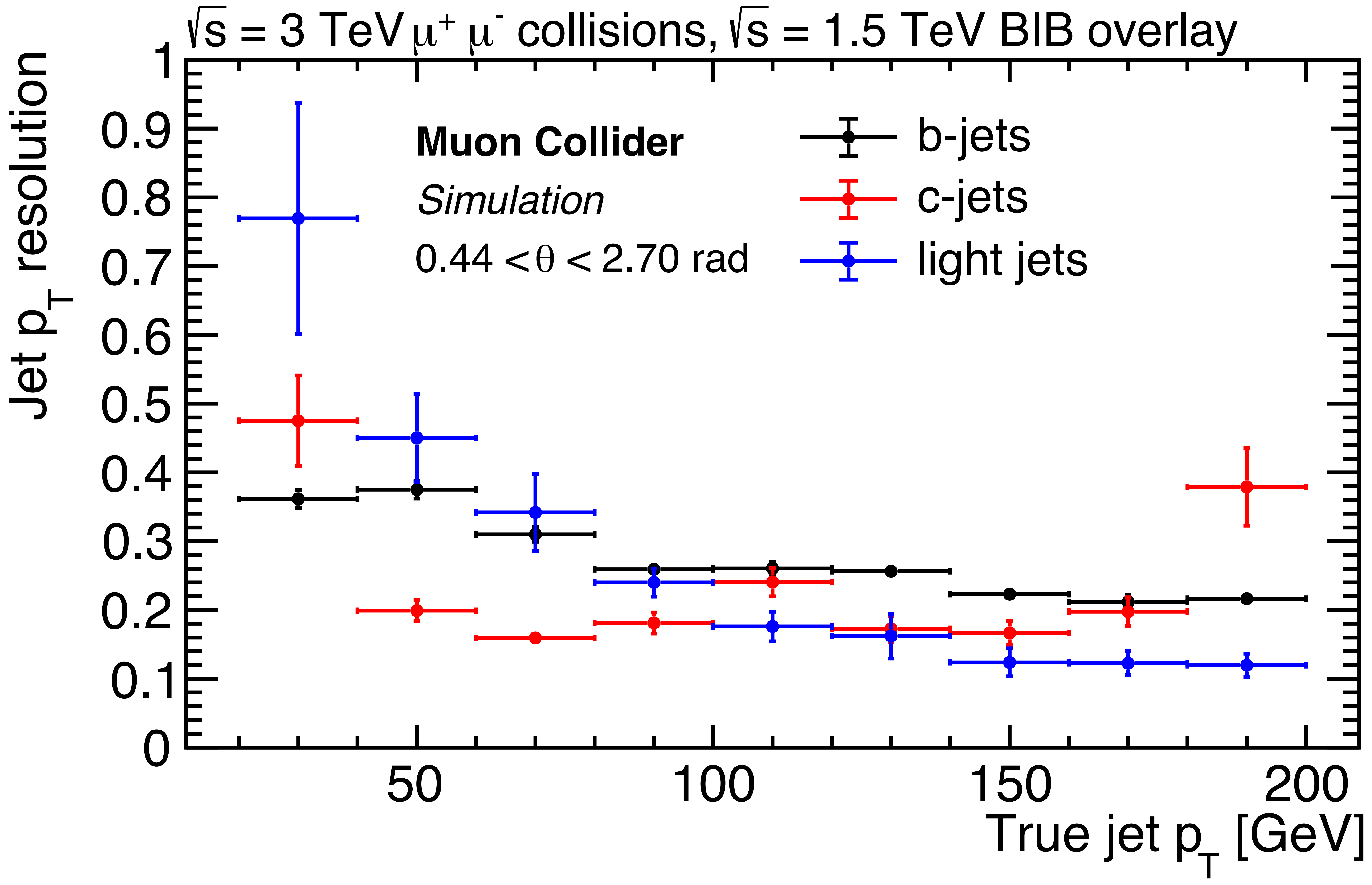
- Jet p_T resolution for b-jets goes from 35% (low p_T) to 20% (high p_T)
- At the moment it is more similar to the jet energy resolution we have at hadron colliders, while at e^+e^- colliders it is better
- Nevertheless **it is sufficient to separate $H\to b\bar{b}$ from $Z\to b\bar{b}$ with statistical techniques** (discussed in next talk)





Jet performance: flavour

- The jet reconstruction has been tested with physics samples of **different jet flavours**
- The jet energy resolution is of the same order but with relevant differences in many bins of p_T
- Since these are physics samples, they have different production processes and **the angular distribution is very different between them**



b-jet identification: Secondary Vertex

- **The current algorithm used for the b-jet identification is the Secondary Vertex (SV) tagging**

Primary Vertex Finding

Tracks with:
> 3 hits in the Vertex detector
$|D_0| \leq 0.1\,\mathrm{mm}$
$|Z_0| \leq 0.1\,\mathrm{mm}$

→

Tracks selection

- >3 total hits
- Max. values for D_0, Z_0 and errors
- p_T>0.8 GeV

→

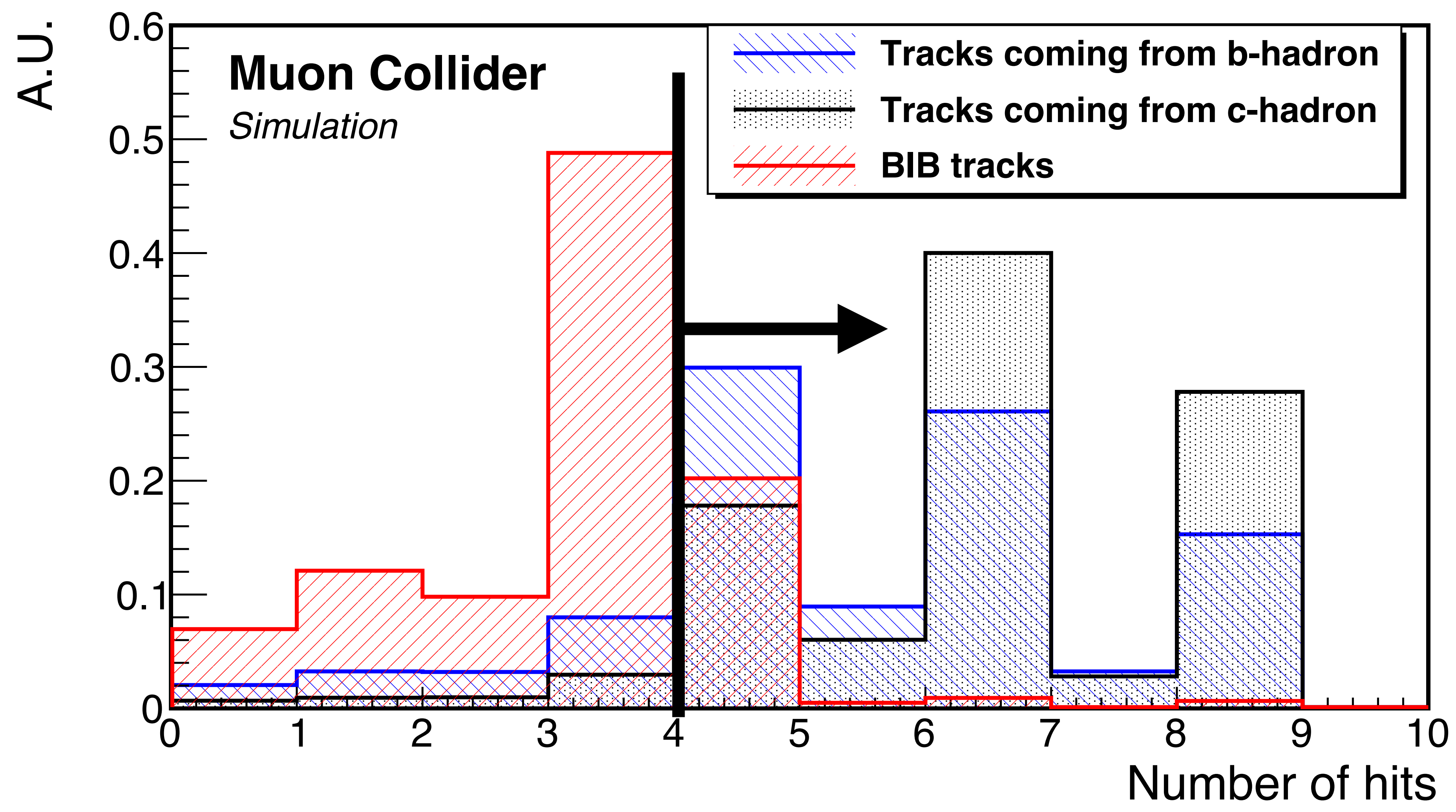
Secondary Vertex Finding

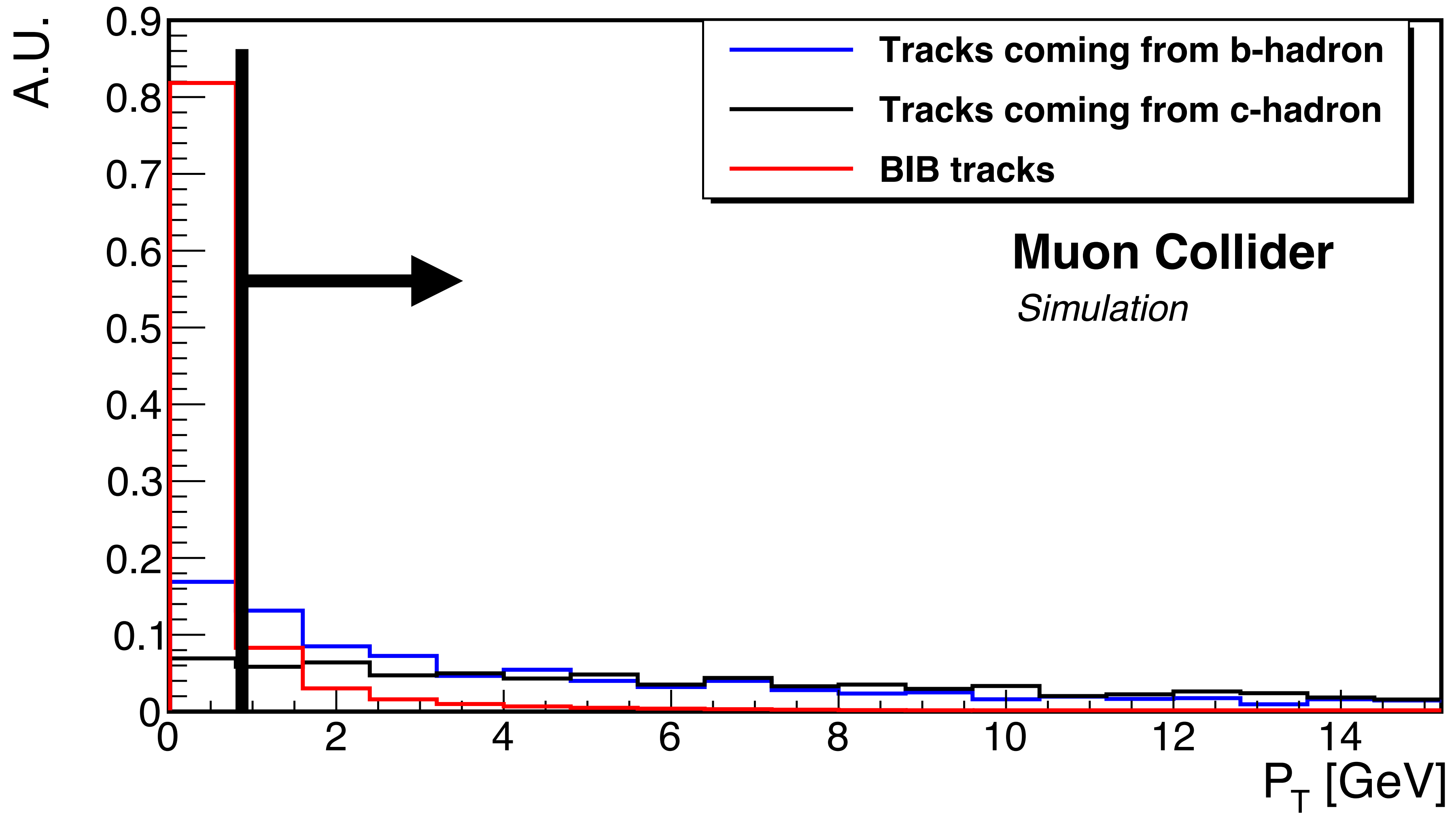
- Two-tracks vertices with invariant mass < 10 GeV
- Iteratively add tracks, on the basis of track-vertex χ2 requirements

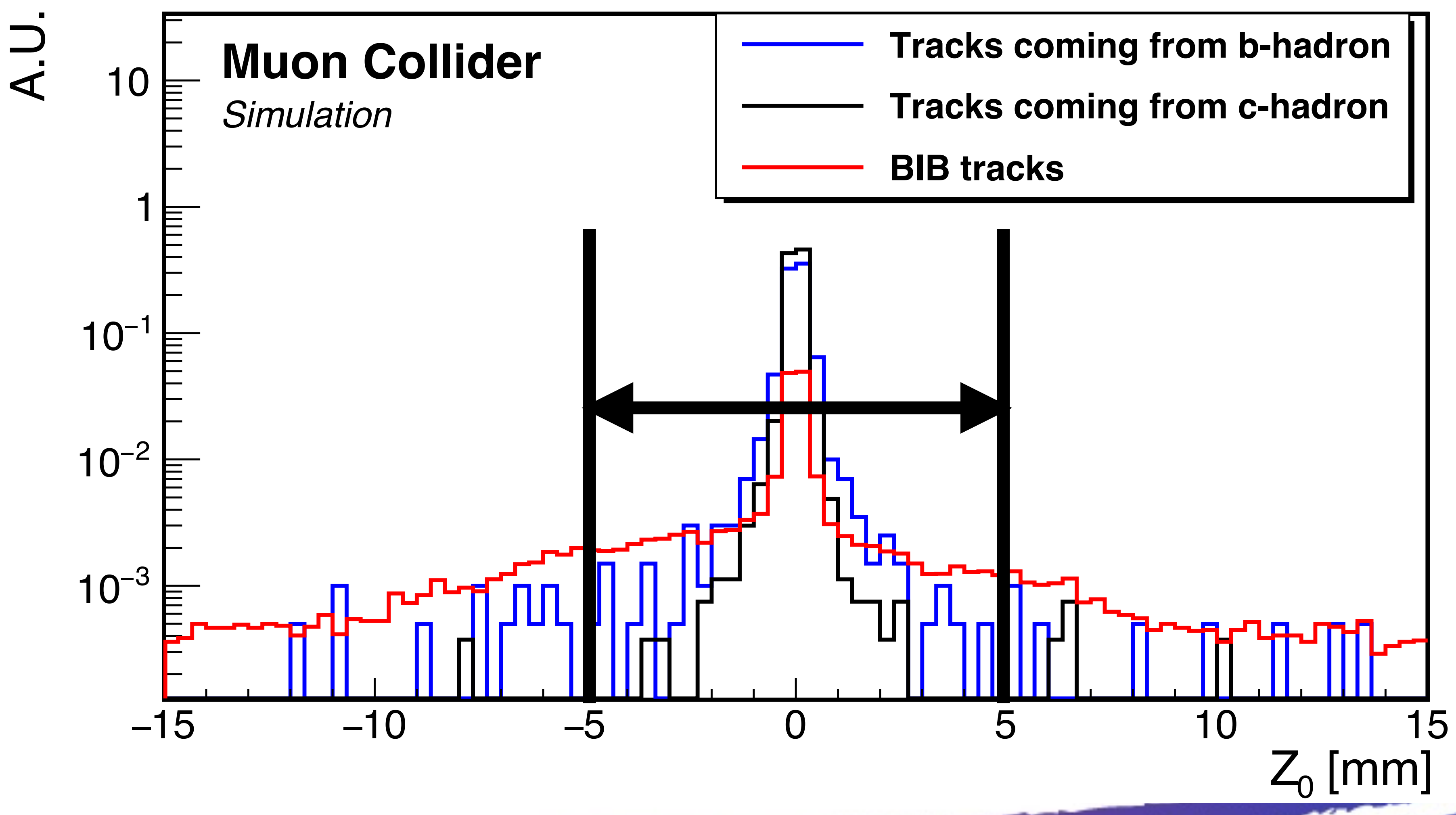
D_0 = longitudinal impact parameter
Z_0 = transverse impact parameter

Secondary Vertex: track selection

- For the SV finding, tracks are reconstructed with the **Conformal Tracking algorithm**
- **Track requirements are tuned in order to remove the combinatorial from BIB**
- By requiring p_T>0.8 GeV 80% of BIB tracks are rejected, while retaining 85-90% of b- and c-hadron tracks

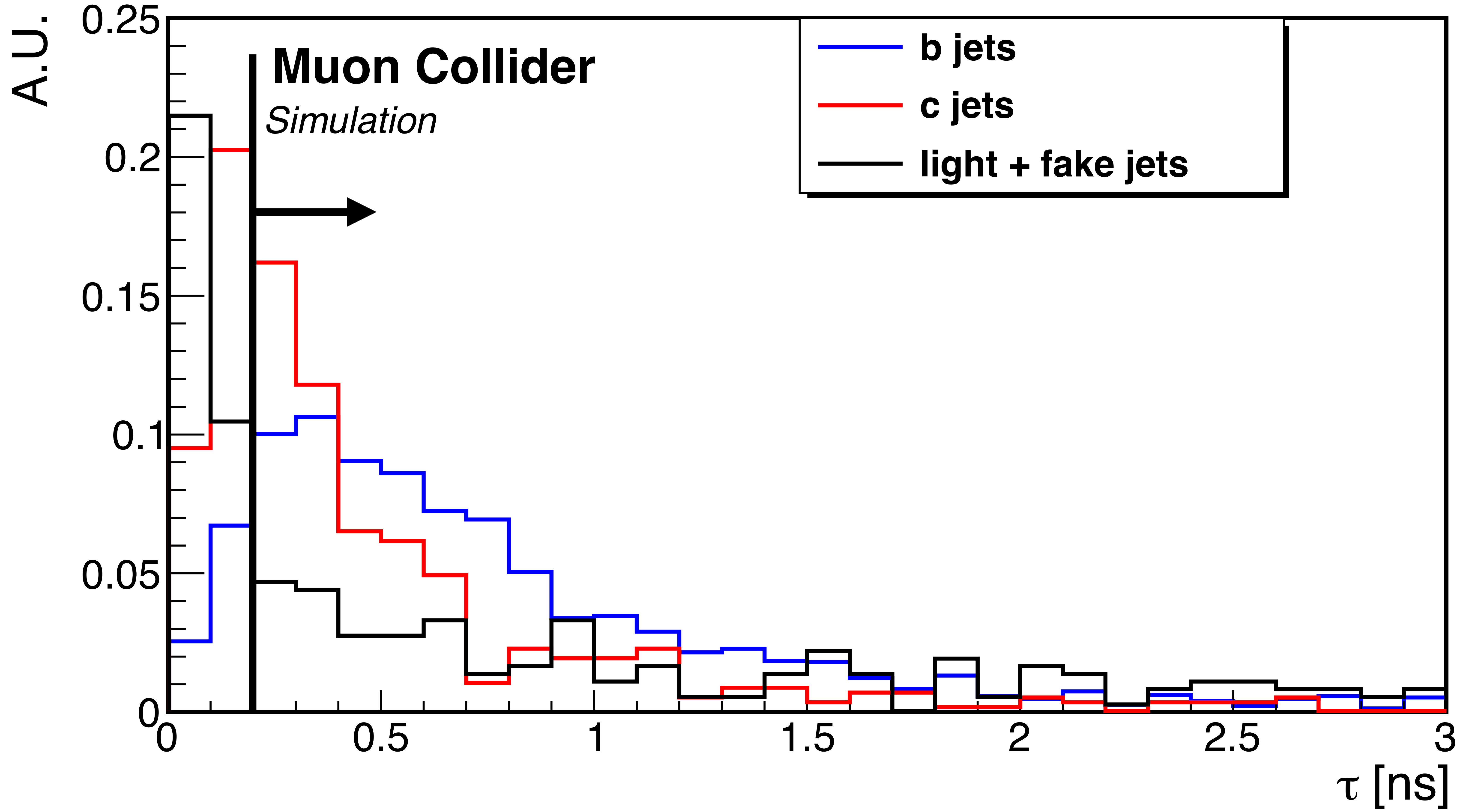






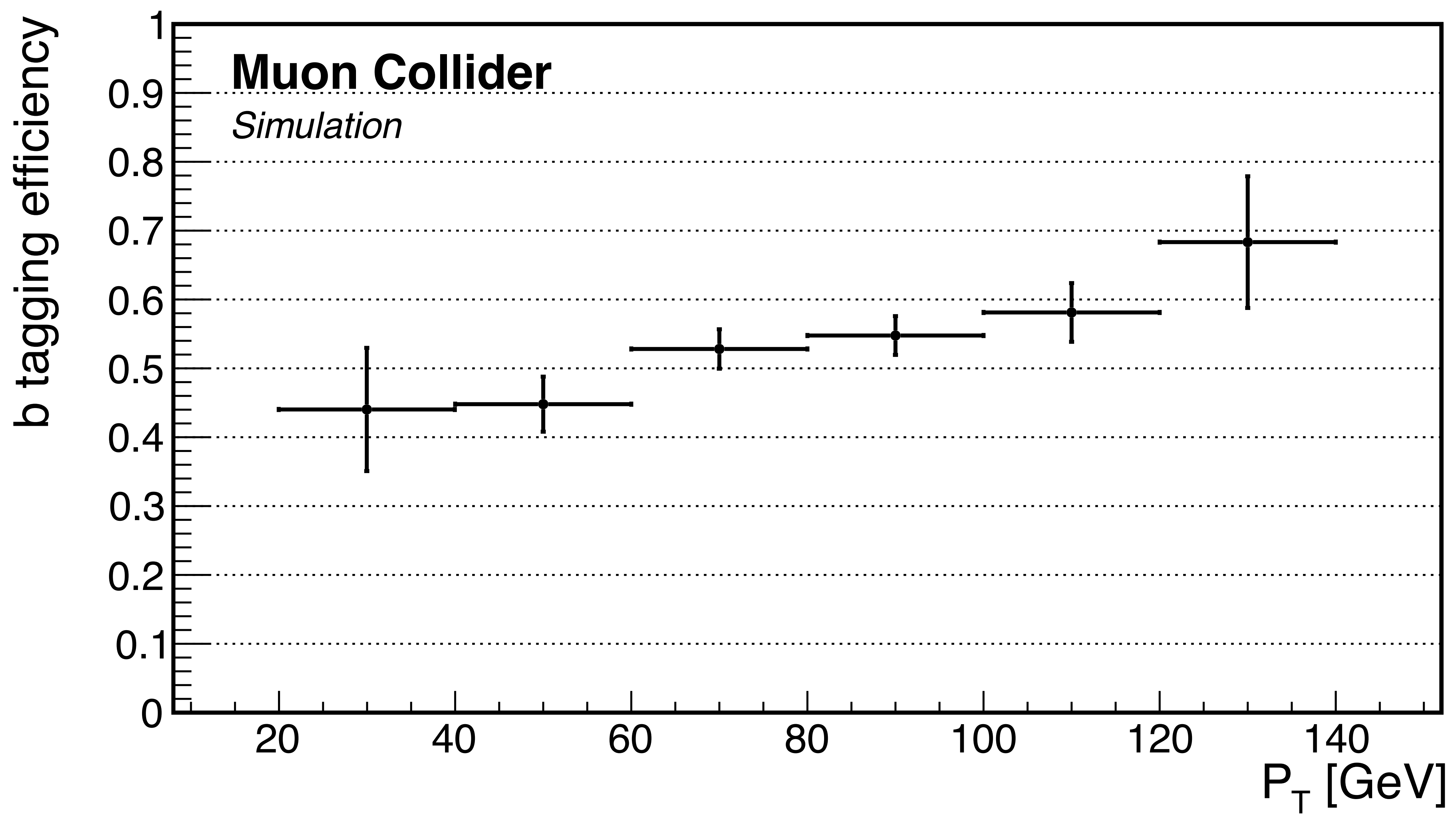
Separating b-jets from c-jets

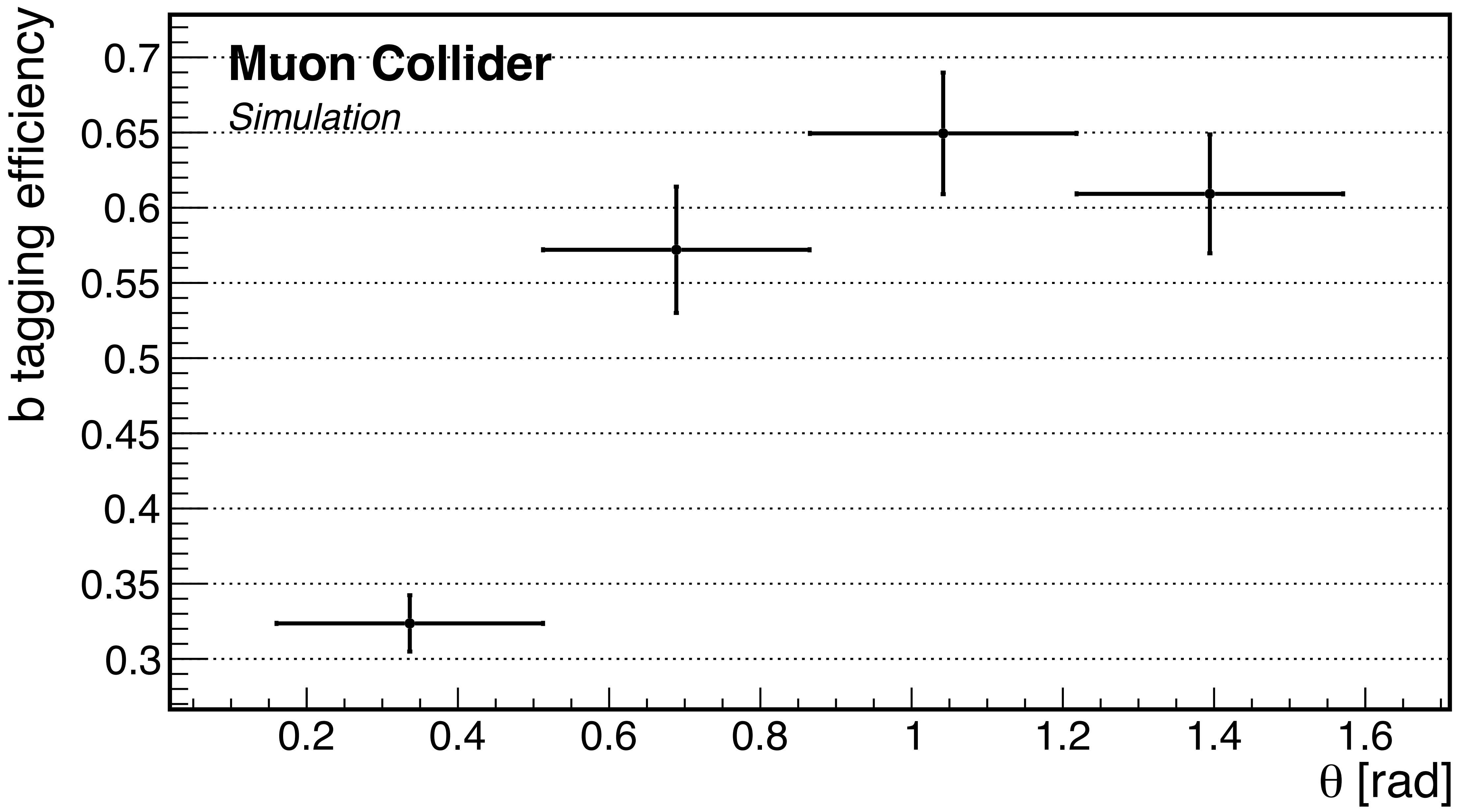
- In measurements with b-jets, it is useful to reduce the c-jets background
- SV features can be used to achieve this task
- **The SV proper lifetime (τ) is one of the most discriminant features**
- A cut on τ > 0.2 ns rejects 30% of c-jets while keeping 90% of b-jets



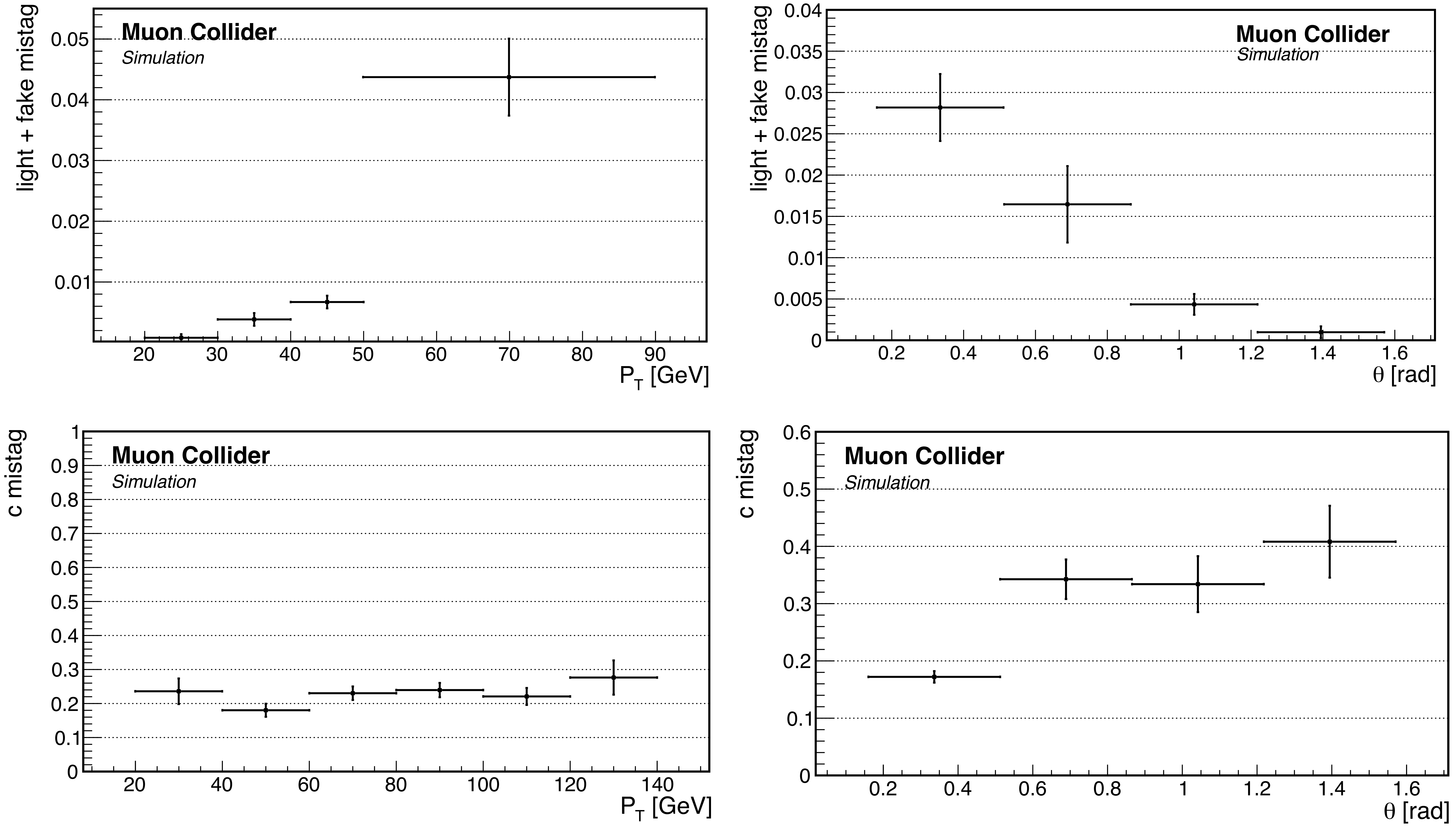
b-jet identification: efficiency

- b-tagging efficiency goes from 30% (forward region) to 65-70% (central region)
- As for the jet reconstruction, **the low efficiency in the forward region is due to tracking inefficiencies and tight requirements to reduce the BIB combinatorial**





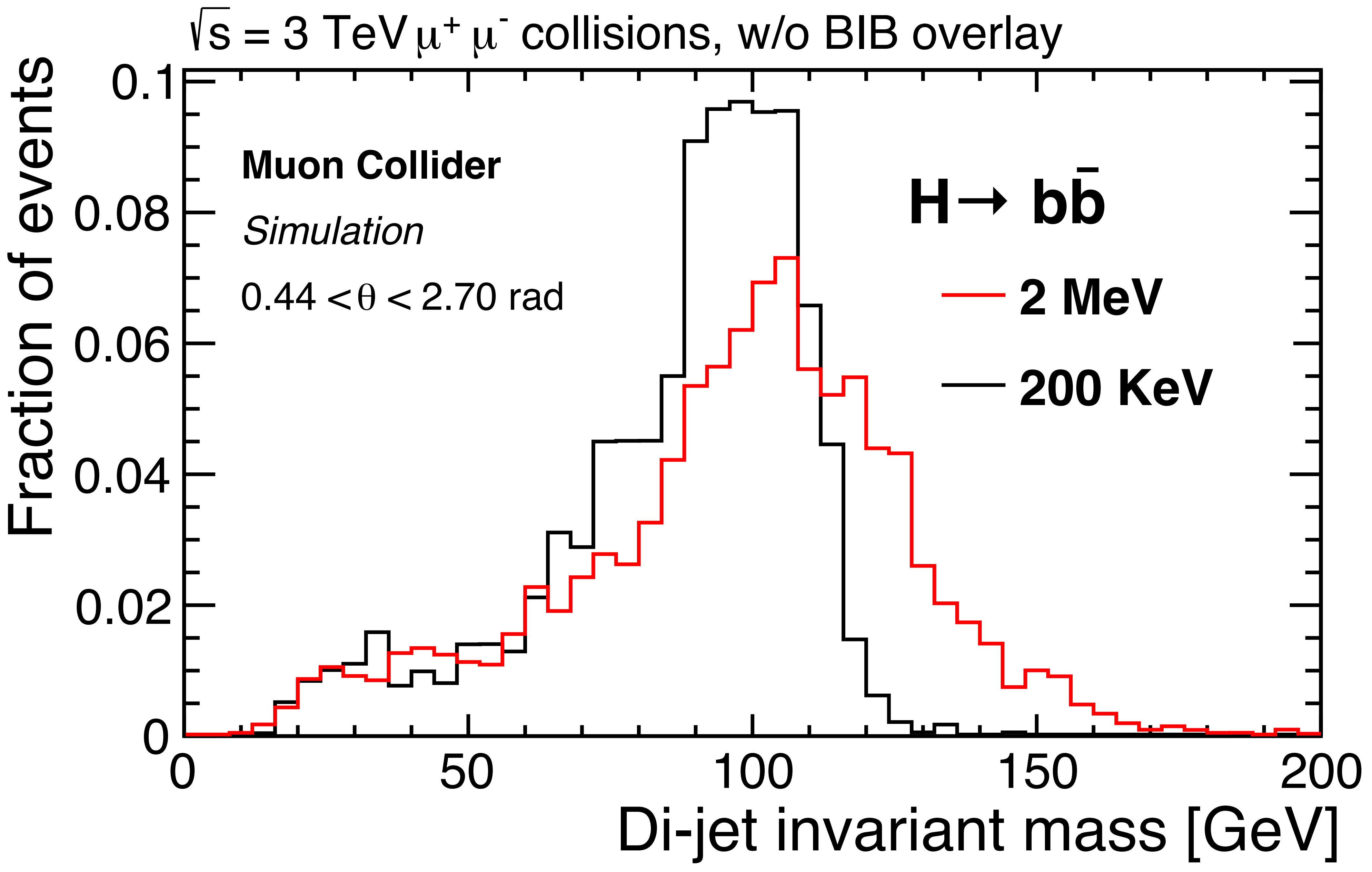
b-jet identification: mistag



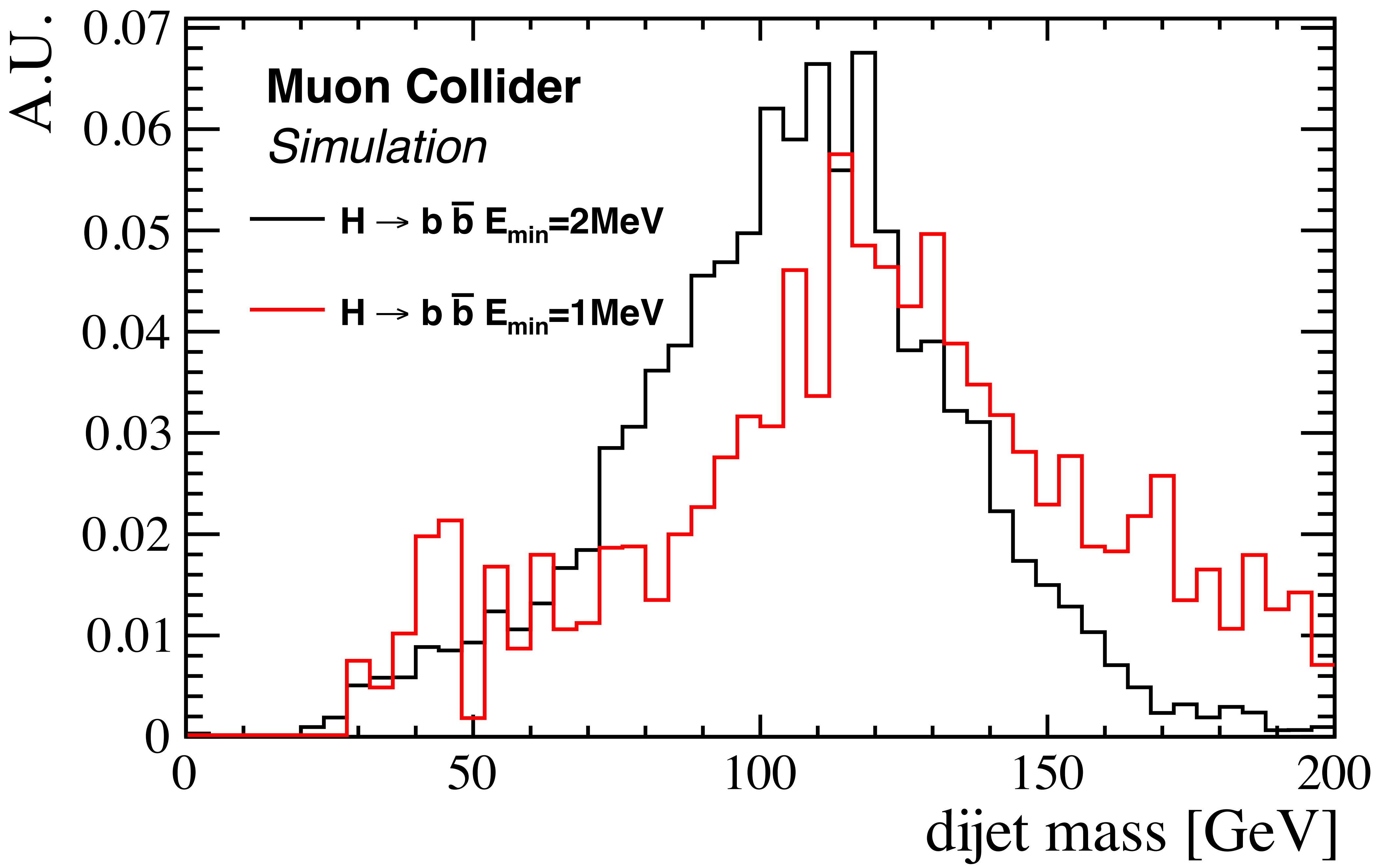
- Light/fake jet mistag rate is under control, higher in the forward region
- The requirement on τ is not sufficient to reduce the c-jet mistag to a negligible level

Prospects: calorimeter thresholds

- **The high threshold at 2 MeV limits the performance on the jet energy resolution**
- It can be demonstrated by reconstructing H->bb without the BIB, with thresholds of 2 MeV or 200 KeV

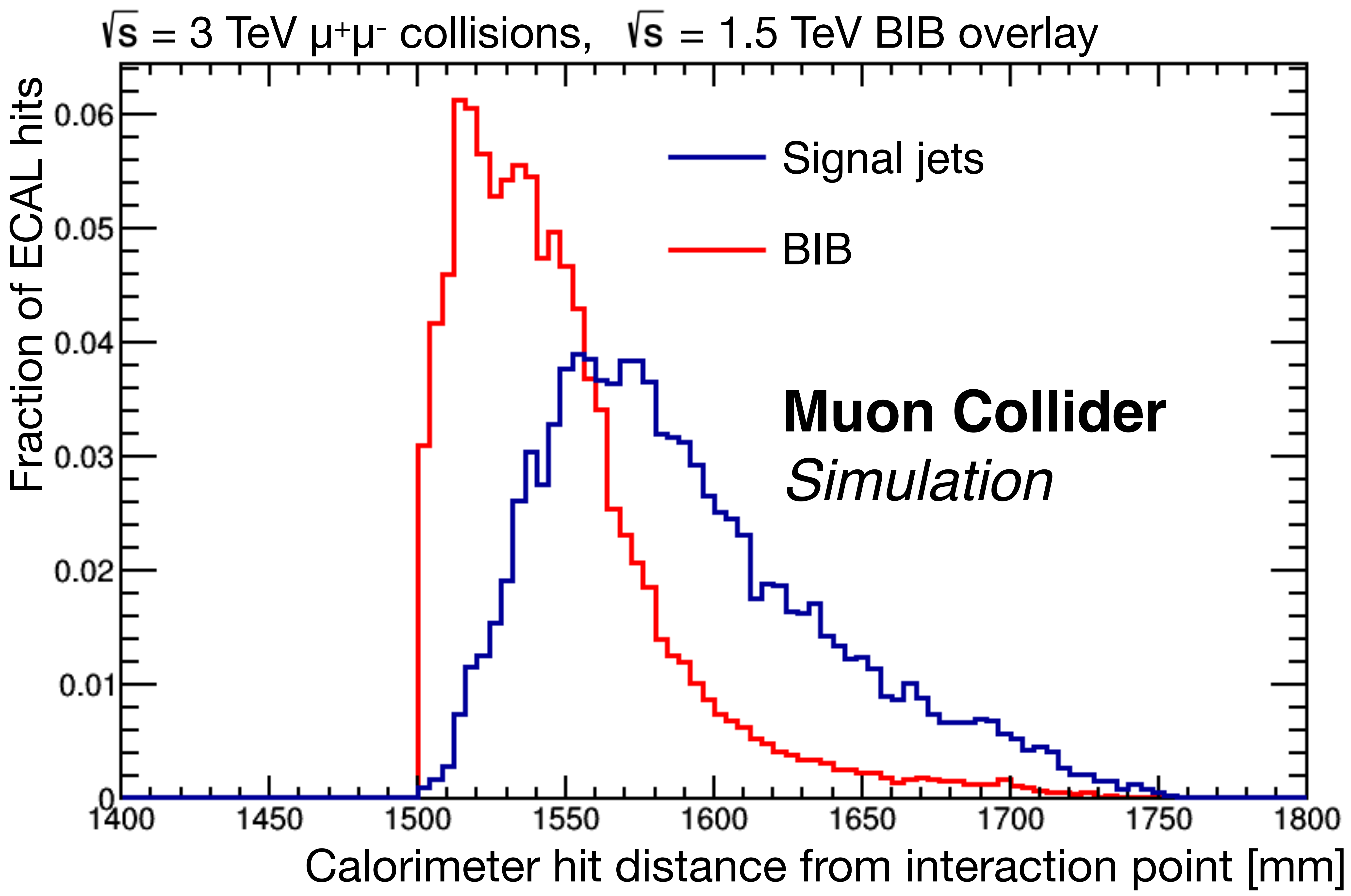


- **However, with the BIB, when the threshold is reduced from 2 MeV to 1 MeV, the jet energy resolution worsen**
- This is due to the fact that the accepted BIB energy is clustered in the jet, degrading the performance
- Tests with thresholds lower than 1 MeV are not easy, for the large computing time with many BIB calorimeter hits



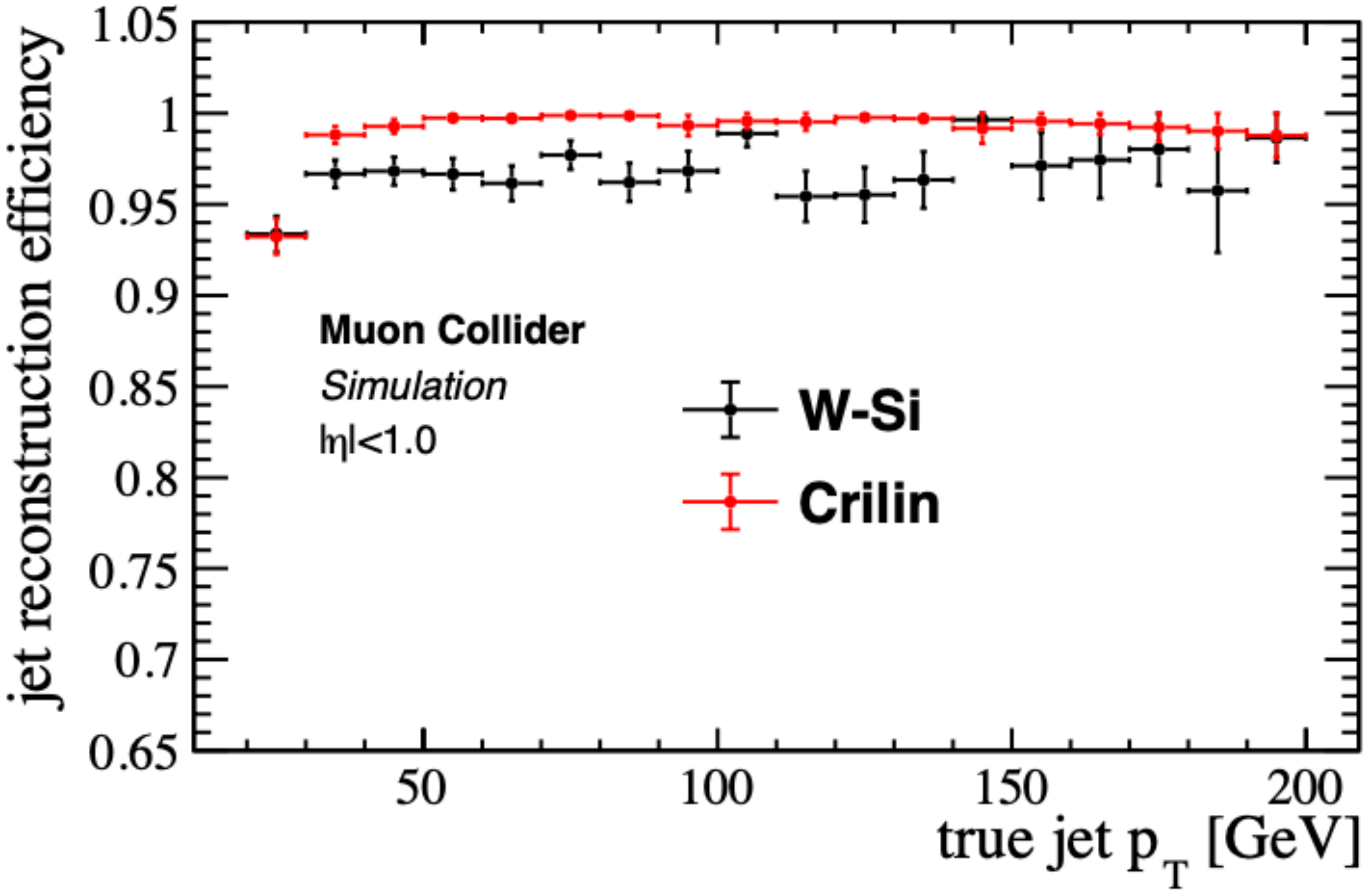
Prospects: calorimeter thresholds

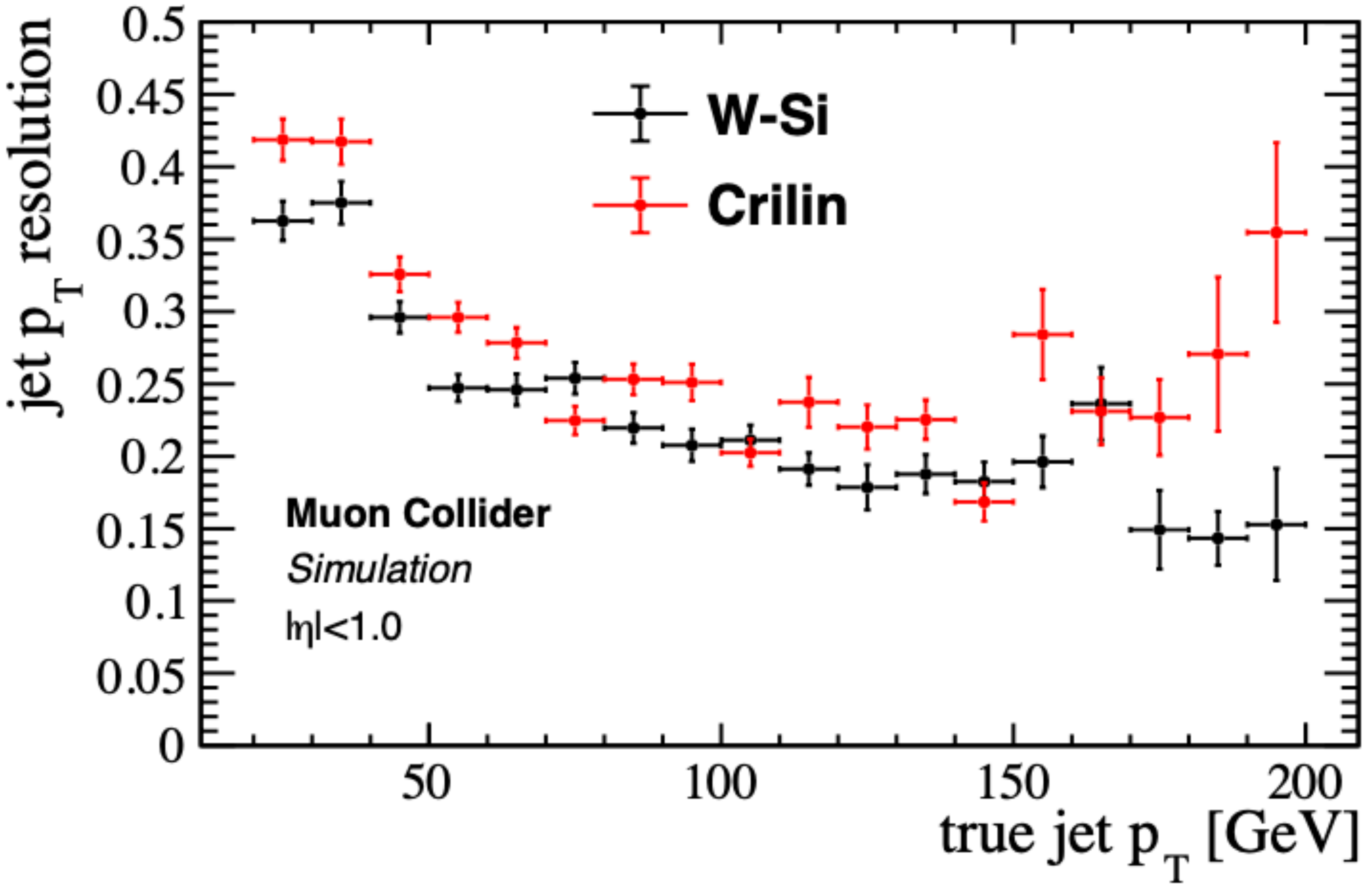
- **If we lower the threshold we would need other strategies to get rid of the BIB**
- One idea could be exploit the longitudinal segmentation of the calorimeter
- The longitudinal distribution of hits/energy released in the calorimeters is pretty different between signal jets and BIB
- In particular most of BIB energy is released in the first layers



Prospects: new calorimeters

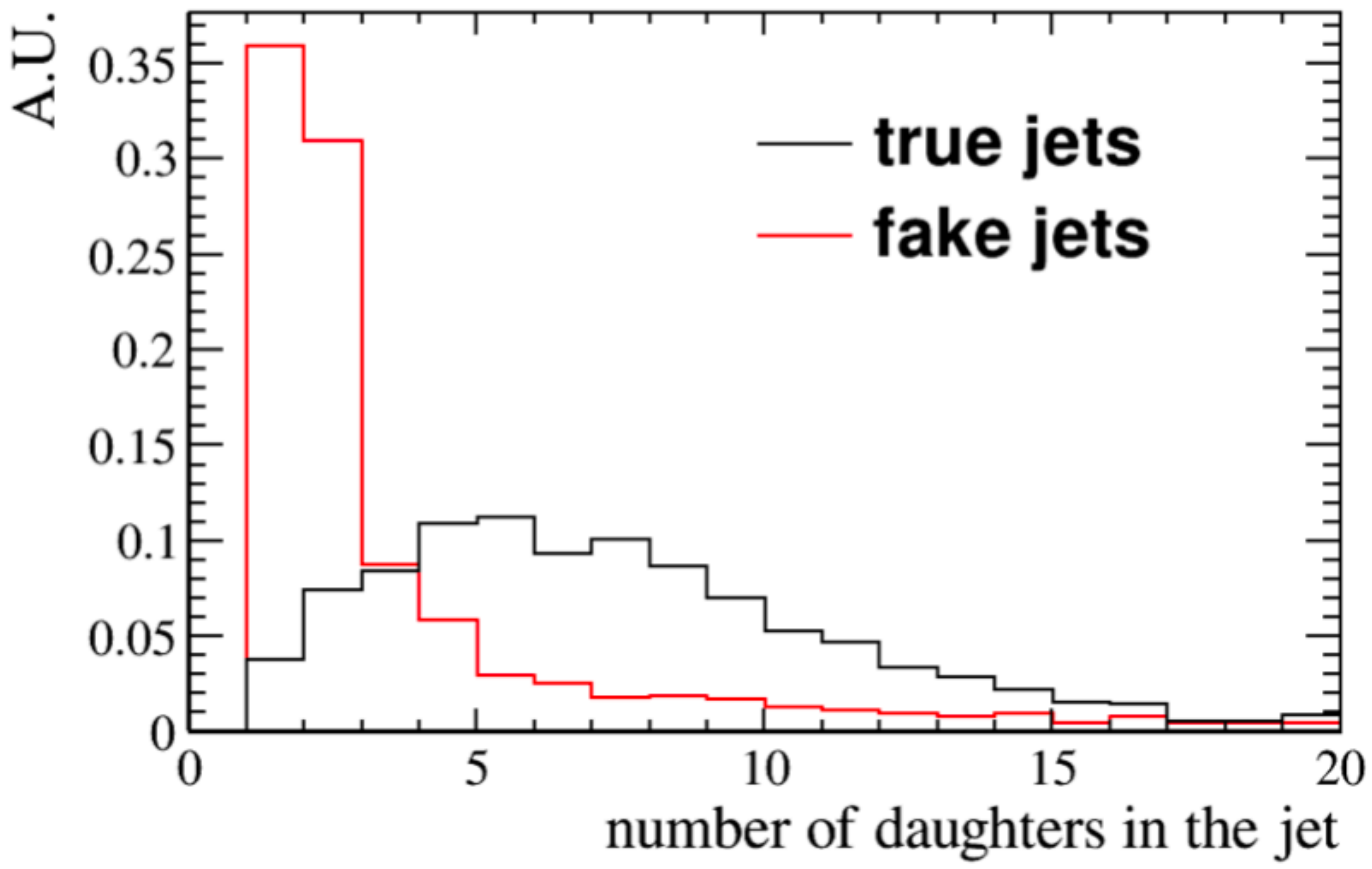
- Different calorimeters technology should be tested (baseline ECAL W-Si, baseline HCAL Steel-plastic scintillator)
- **As an example the Crilin ECAL has been used** (semi-homogeneous with five layers of PbF2, see talk tomorrow)
- Performance in the same order of W-Si (with similar fake rate) but different reconstruction techniques could be employed

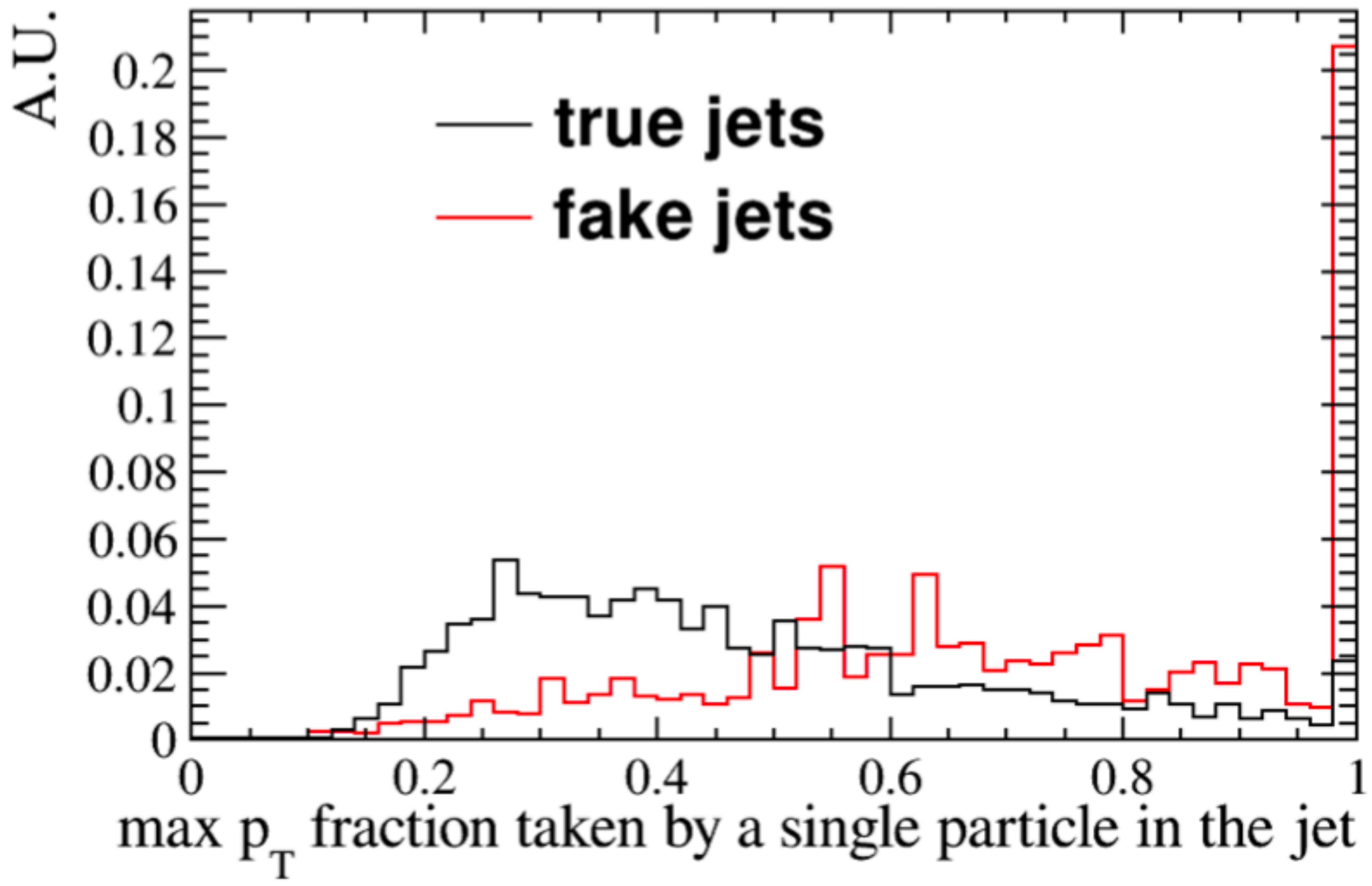




Prospects: efficiency and fake removal

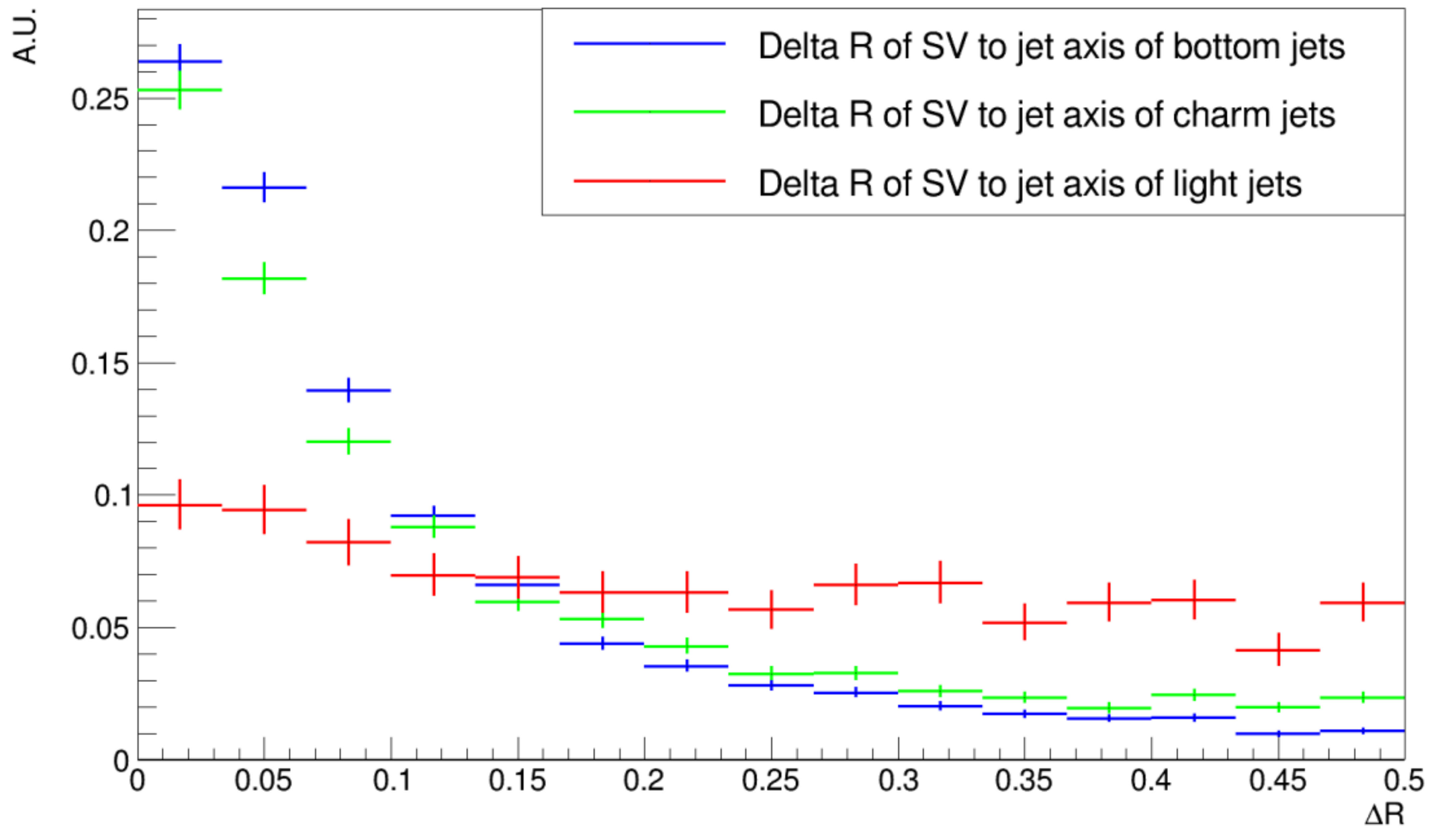
- In my view there are two way to improve the jet reconstruction efficiency in the forward region:
  - **Improve track reconstruction in the forward region**: many signal jets in the forward region does not have reconstructed tracks
  - **Exploit the jet substructure to remove fake jets**
  - **Longitudinal shape of calorimeter clusters?**

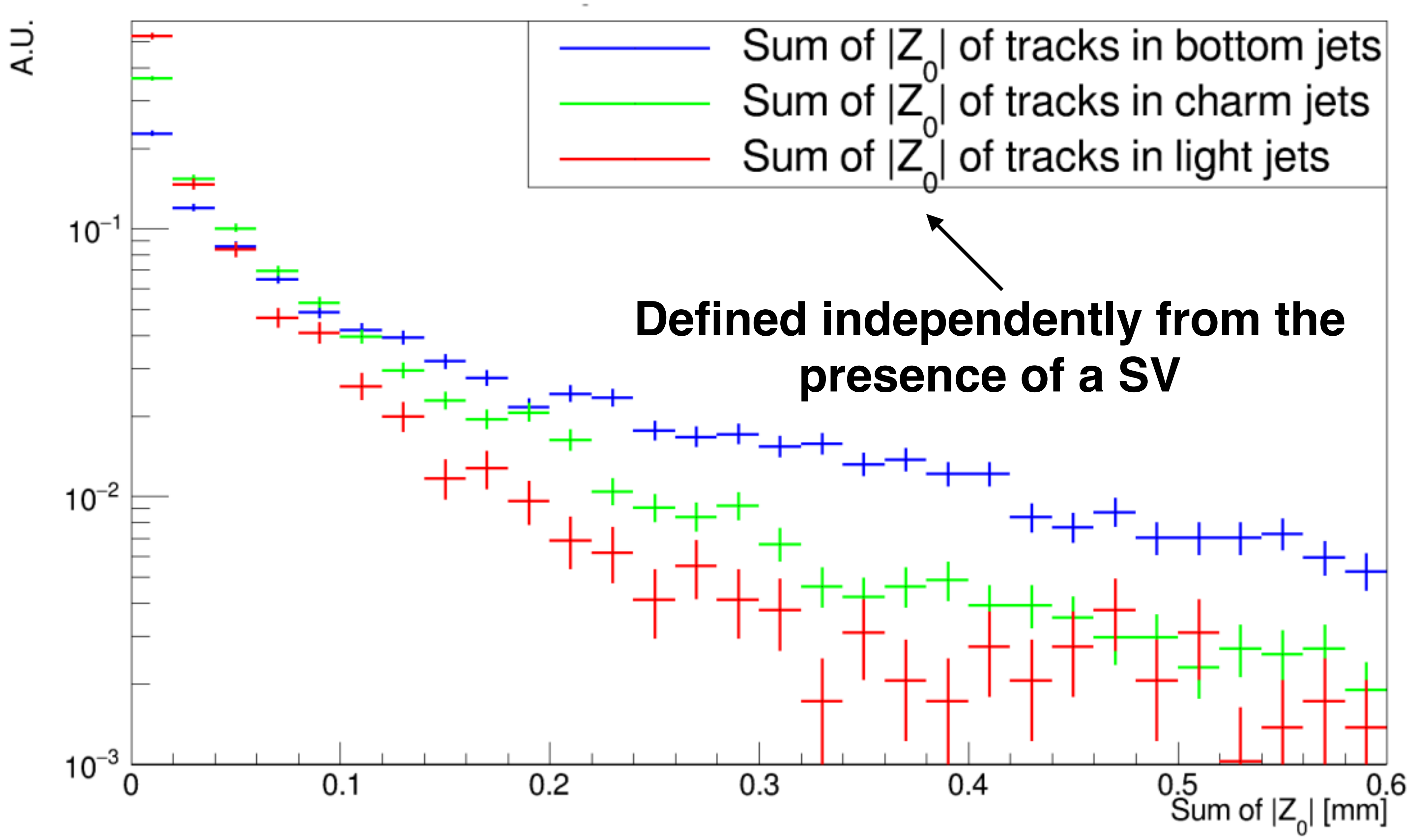




Prospects on jet identification

- Secondary Vertex reconstruction would benefit from a tracking strategy optimized for displaced tracks
- **The features of the SV should be studied in detail for improving the flavour separation**
- More advanced techniques that exploit not only SV information but also the jet substructure could be used
- Deep learning techniques could play a big role in this context





Conclusions

- **The jet reconstruction at the Muon Collider poses challenges that are different from other well-known environments**: interesting and important R&D!

- With the current algorithm/detector, **the jet performance in the central region is at the level of hadron collider experiments**, we need improvements in order to get closer to electron-positron colliders

- In the forward region we have low efficiencies, **an intensive R&D for this region is necessary**

- **However the jet reconstruction algorithm available now can already be used to perform good measurements** (check next talk)

- **There are many ideas and room for improvements**!

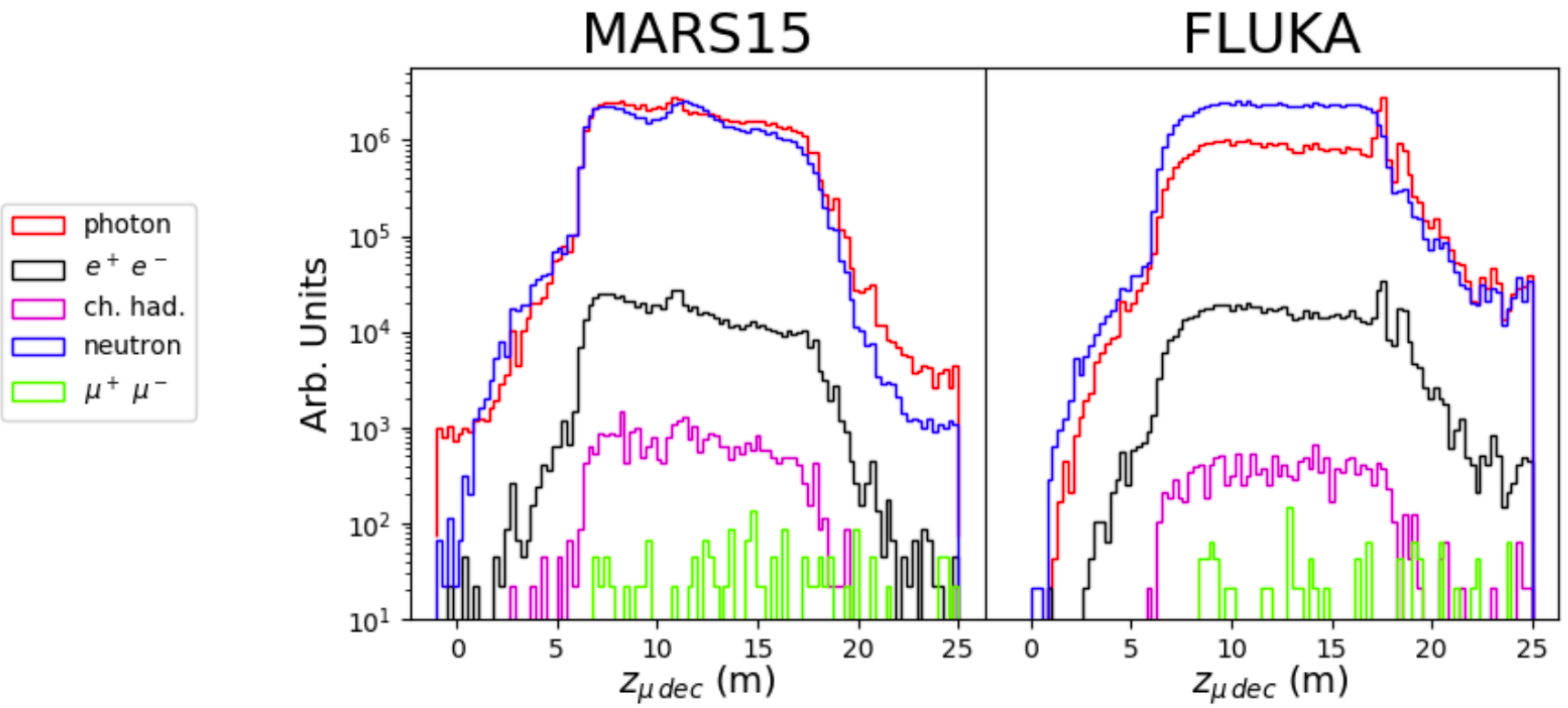
Thanks for your attention!

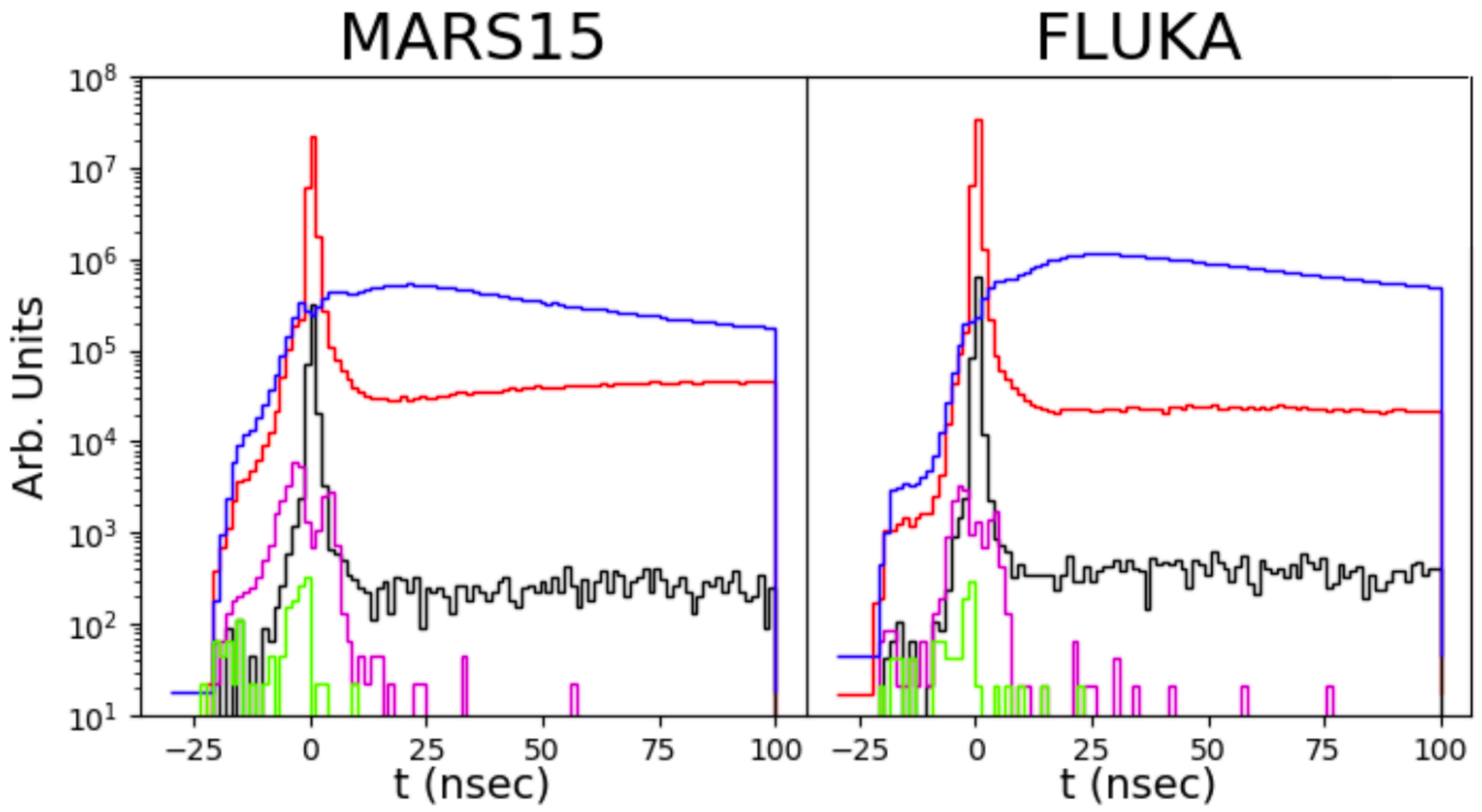
Backup

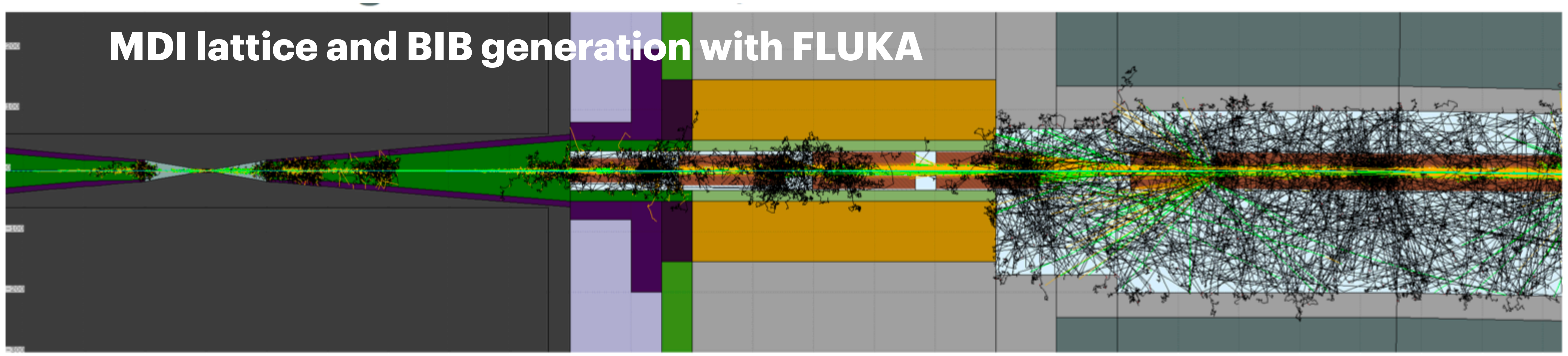
Beam-induced background (BIB)

- It is produced by the decay in flight of muons, and subsequent interactions
- The BIB is mitigated by the Machine Detector Interface (*e.g.* two tungsten nozzles are inserted)
- **In order to assess the physics reach of a Muon Collider it is fundamental to study the impact of the BIB on the detector**
- The BIB is simulated with MARS15 or FLUKA, by considering the machine and the Machine Detector Interface lattice

https://arxiv.org/abs/2105.09116







Detector

<u>Talk on Muon Collider calorimeter</u> by Eleonora Diociaiuti

hadronic calorimeter

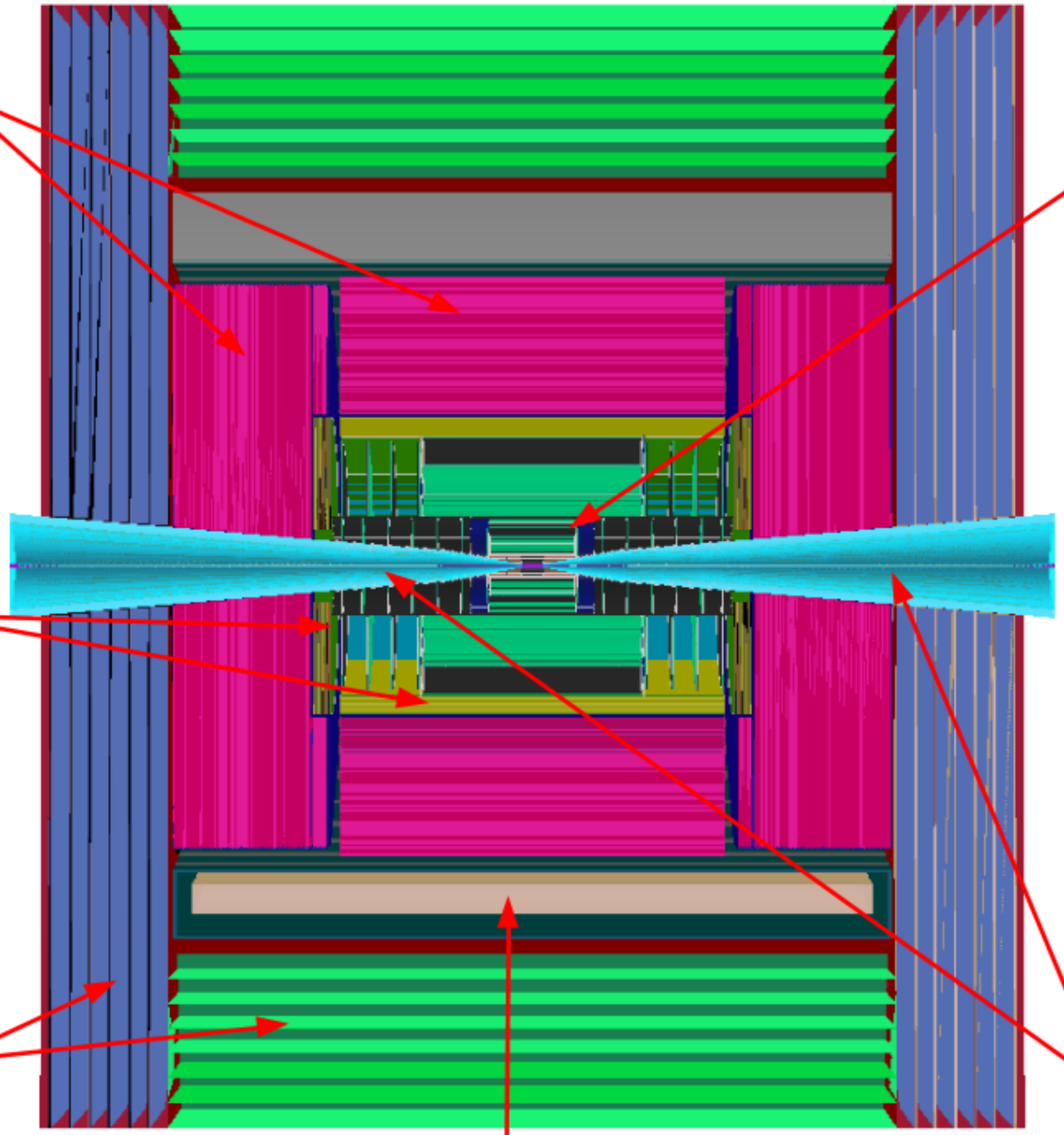
- 60 layers of 19-mm steel absorber + plastic scintillating tiles;
- 30x30 mm^2 cell size;
- 7.5 λ_I.

electromagnetic calorimeter

- 40 layers of 1.9-mm W absorber + silicon pad sensors;
- 5x5 mm^2 cell granularity;
- 22 X_0 + 1 λ_I.

muon detectors

- 7-barrel, 6-endcap RPC layers interleaved in the magnet's iron yoke;
- 30x30 mm^2 cell size.



superconducting solenoid (3.57T)

tracking system

- Vertex Detector:
  - double-sensor layers (4 barrel cylinders and 4+4 endcap disks);
  - 25x25 μm^2 pixel Si sensors.
- Inner Tracker:
  - 3 barrel layers and 7+7 endcap disks;
  - 50 μm x 1 mm macro-pixel Si sensors.
- Outer Tracker:
  - 3 barrel layers and 4+4 endcap disks;
  - 50 μm x 10 mm micro-strip Si sensors.

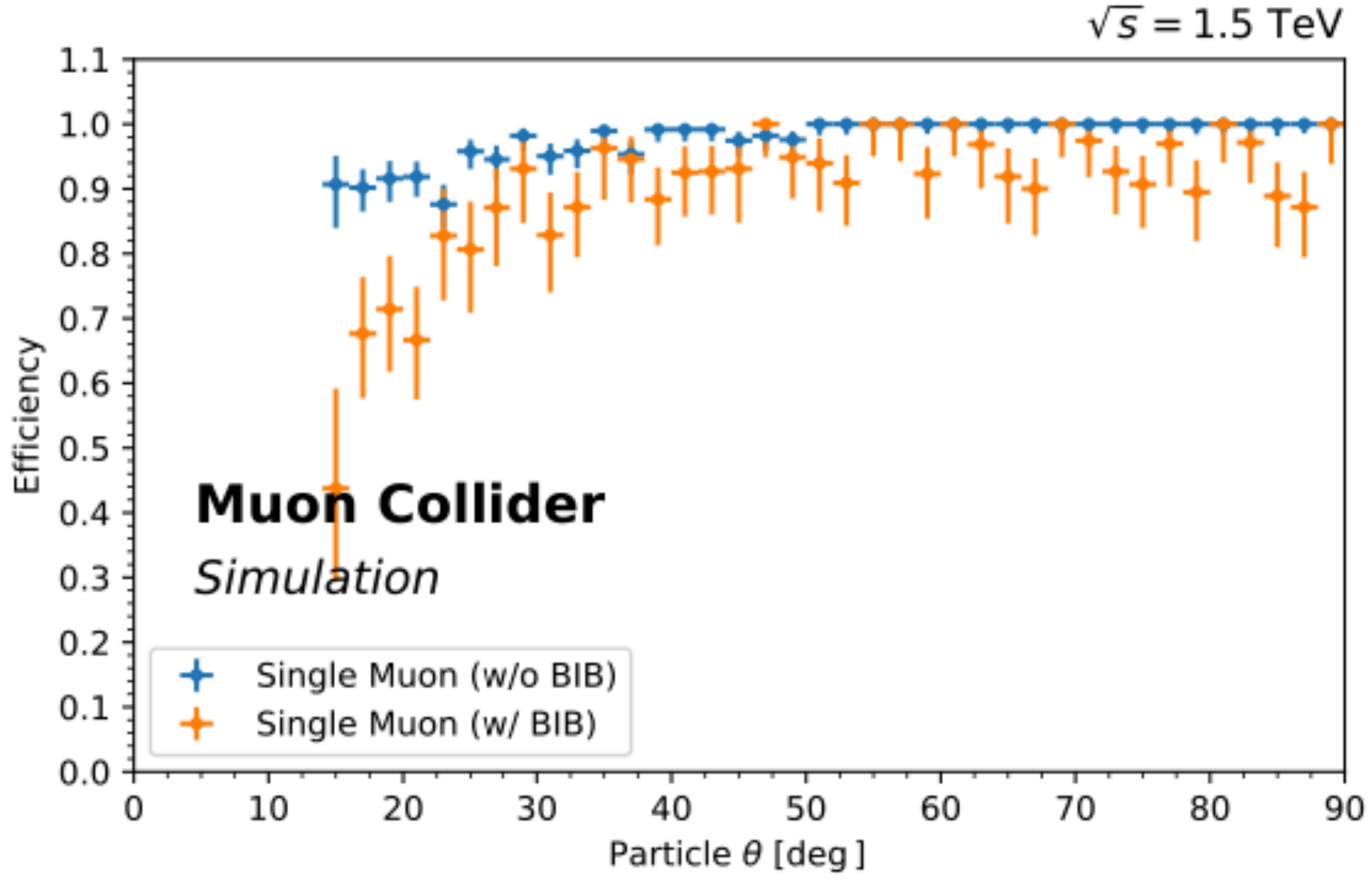
shielding nozzles

- Tungsten cones + borated polyethylene cladding.

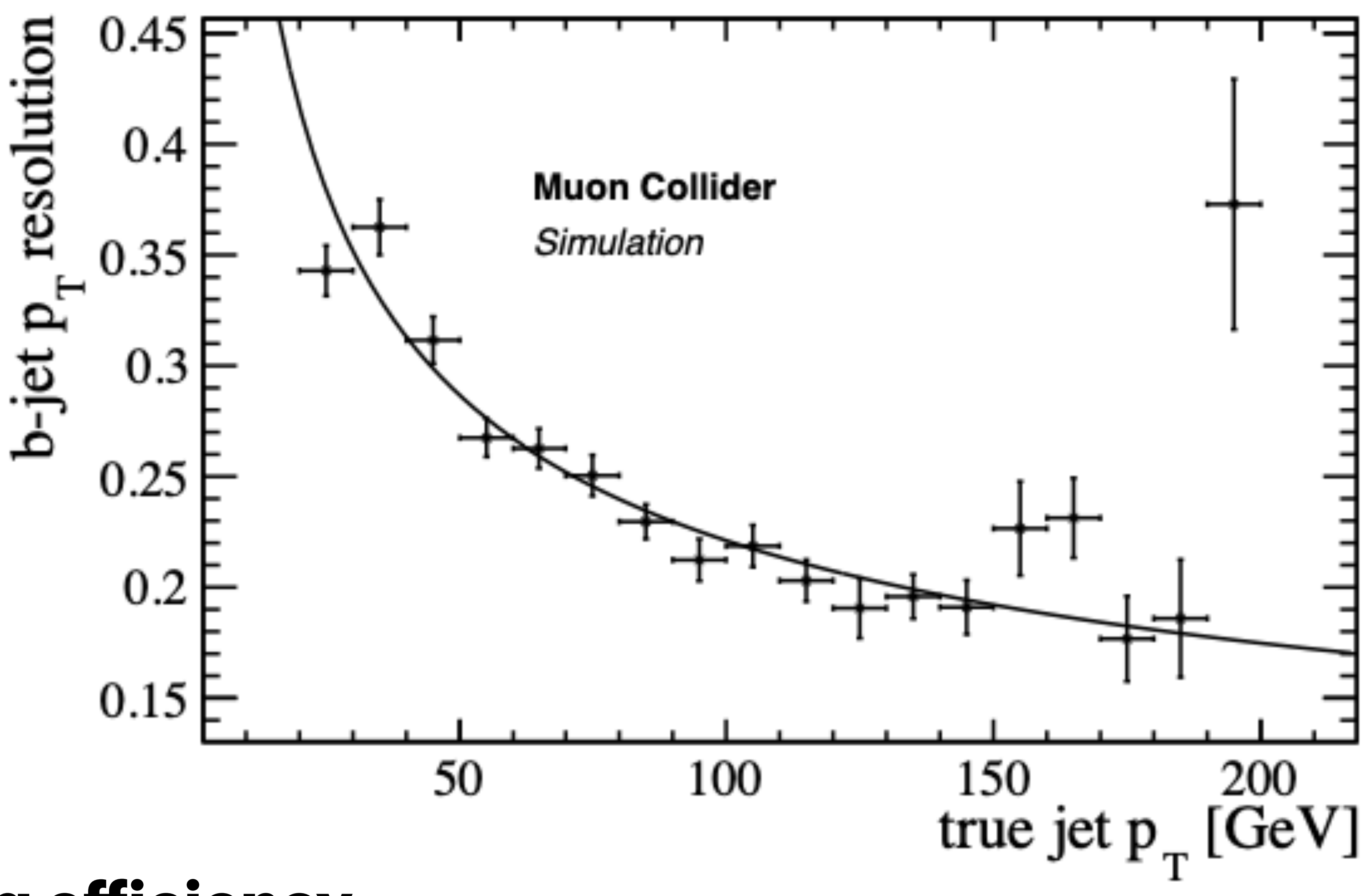
- **The interaction of BIB/signal with the detector is simulated with Geant4**

Physics objects reconstruction

Track reconstruction efficiency



Jet momentum resolution



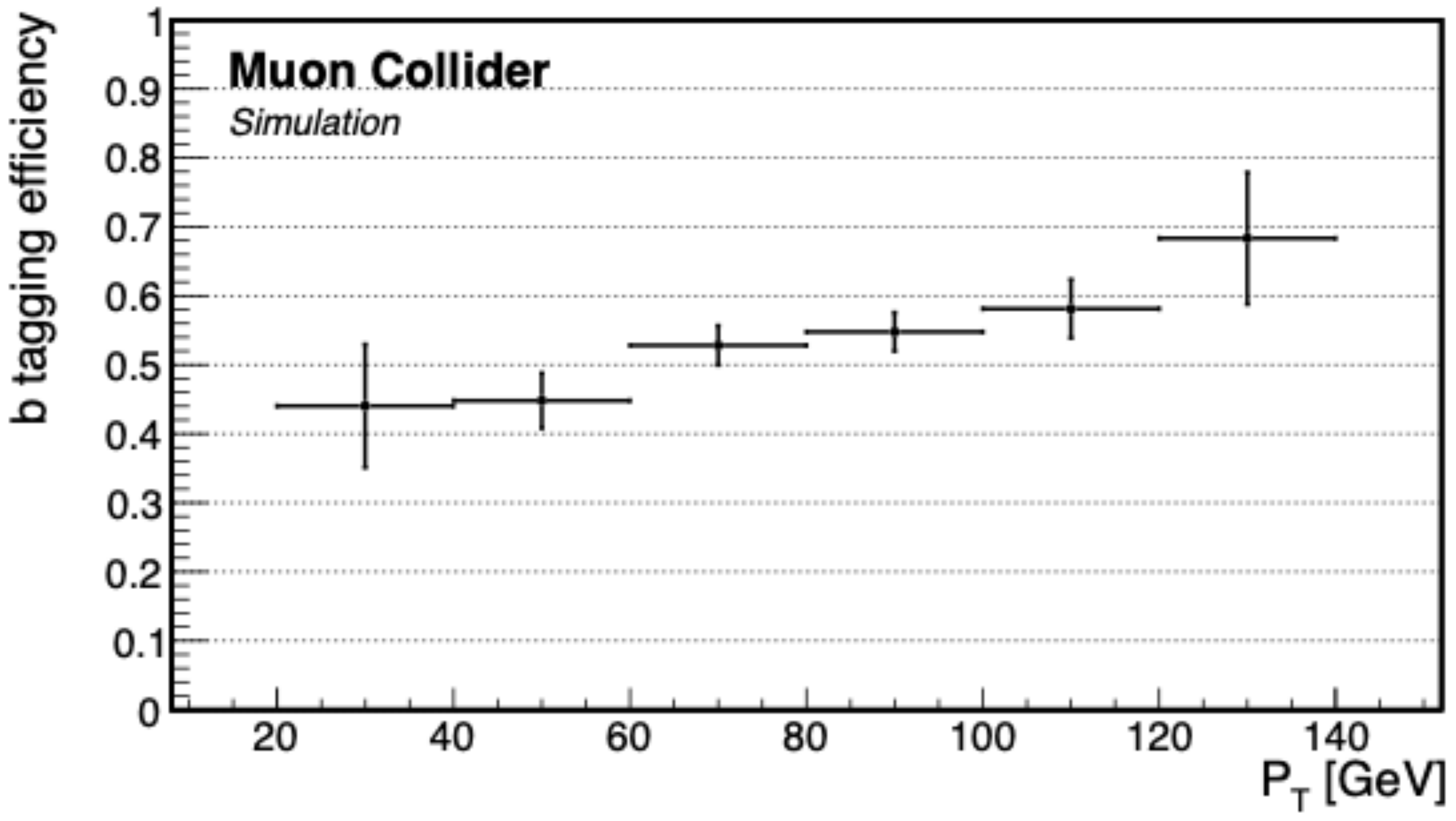
Reconstruction algorithms are not yet fully optimized

Muon fiducial region

- $10° < \theta < 170°$
- $p_T > 20$ GeV

https://arxiv.org/abs/2203.07964

Secondary Vertex tagging efficiency



Jet fiducial region

- $|\eta| < 2.5$
- $p_T > 20$ GeV