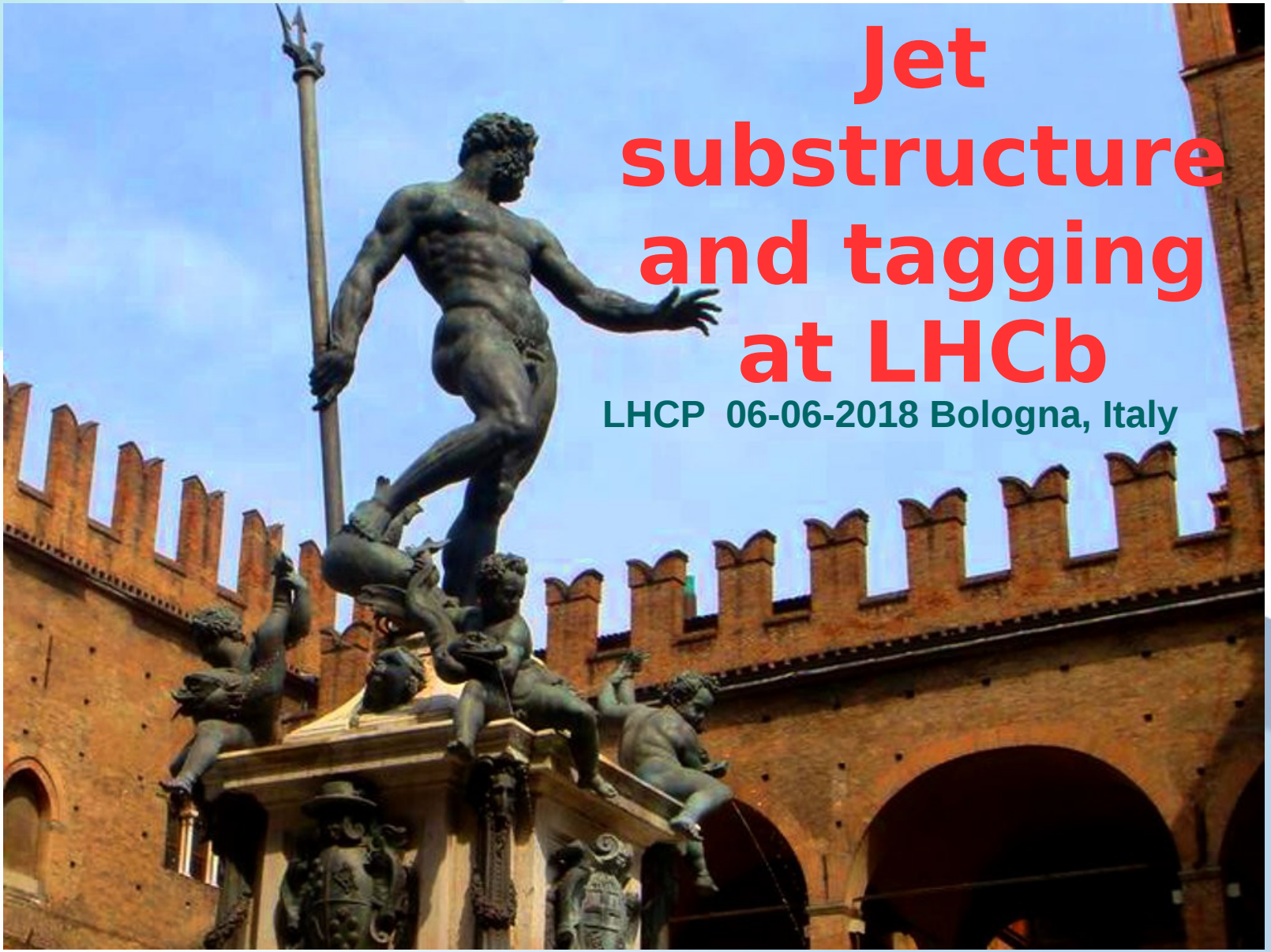
Jet substructure and tagging at LHCb

LHCP 06-06-2018 Bologna, Italy

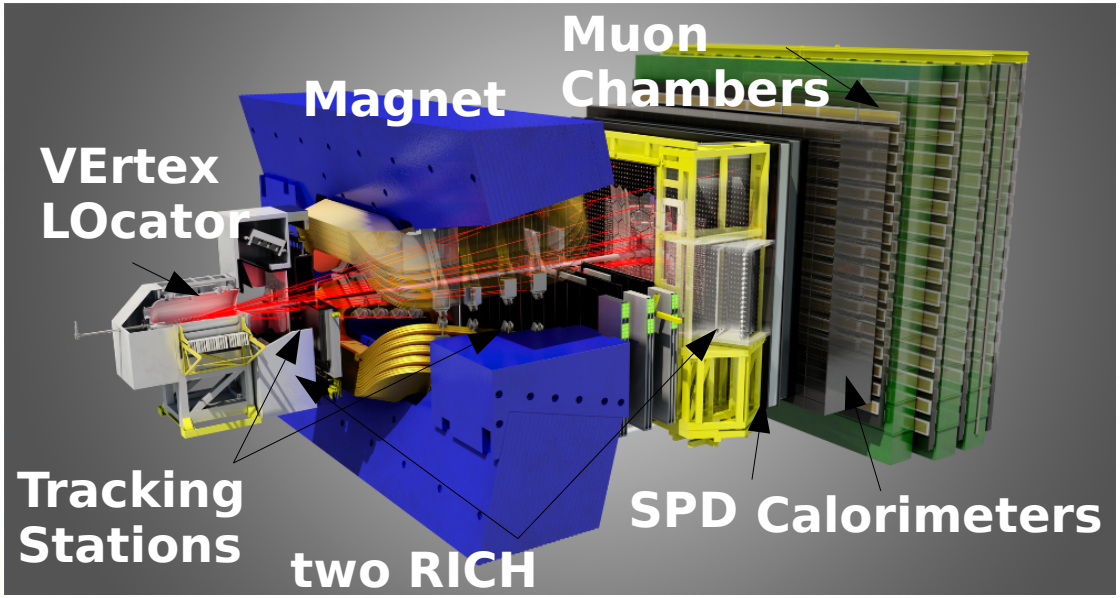


Lorenzo Sestini
Università di Padova
and INFN

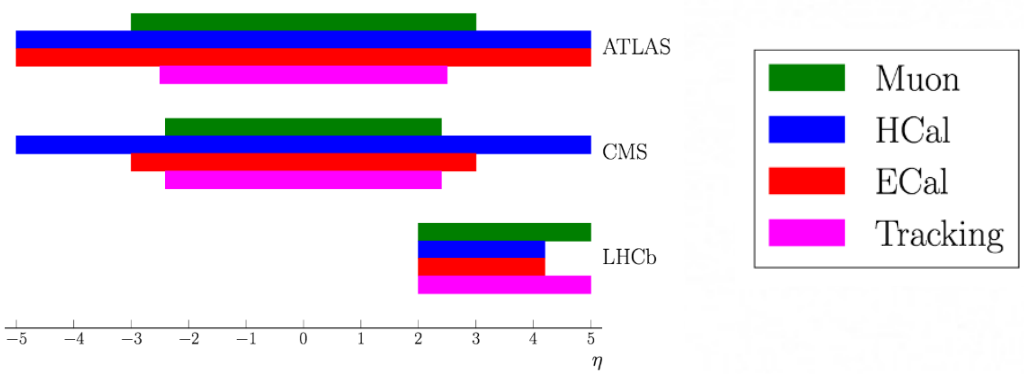
On behalf of the LHCb collaboration

LHCb: a general purpose forward detector

JINST 3 (2008) S08005



- Forward region: **unique acceptance** within the LHC experiments ($2 < \eta < 5$).
- **Momentum resolution**: 0.4% at 5 GeV and 0.6% at 100 GeV.
- **Muon ID efficiency**: 97% with 1-3% $\mu \rightarrow \pi$ mis-identification.
- **Electron reconstruction:** bremsstrahlung recovery and well-measured direction.
- **Excellent vertex reconstruction:** tagging of b and c jets.



LHCb demonstrated its capability in EW and jet physics → **General purpose forward detector**

Jet reconstruction algorithm

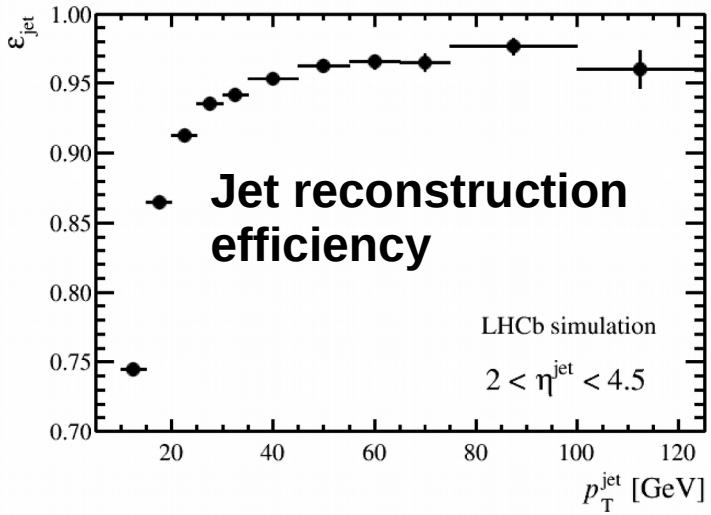
JHEP 01 (2014) 033

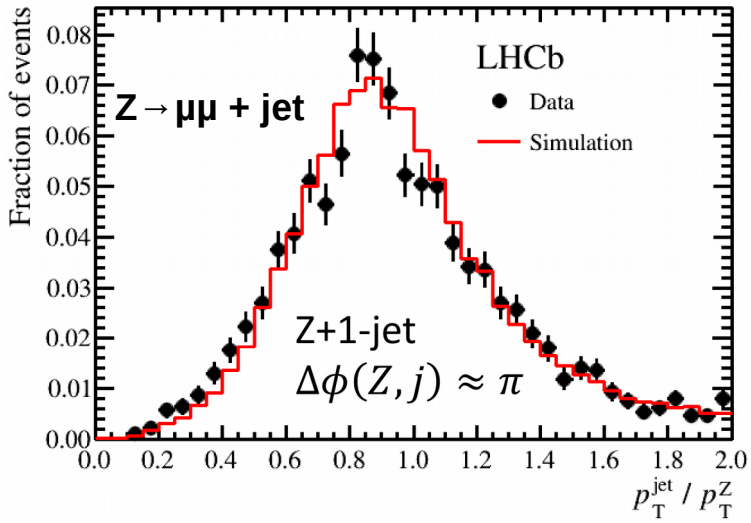
Particle Flow → **Tracks and calorimeter clusters are selected as inputs**

Anti-kt clustering → **Inputs are used to form jets of a given cone radius R=0.5**

Jet four-momentum → $E_{jet} = \sum_i E_i \quad \vec{p}_{jet} = \sum_i \vec{p}_i$

Jet energy correction → **The jet energy is corrected for detector inefficiencies (calibrated on simulation)**





Energy resolution of final jets δE/E ≈ 10-16% in 20-100 GeV p_T range.

Jet substructure

Study of jet substructure at LHCb

Secondary Vertices for b-jets and c-jets identification

J/ψ in jets for QCD studies

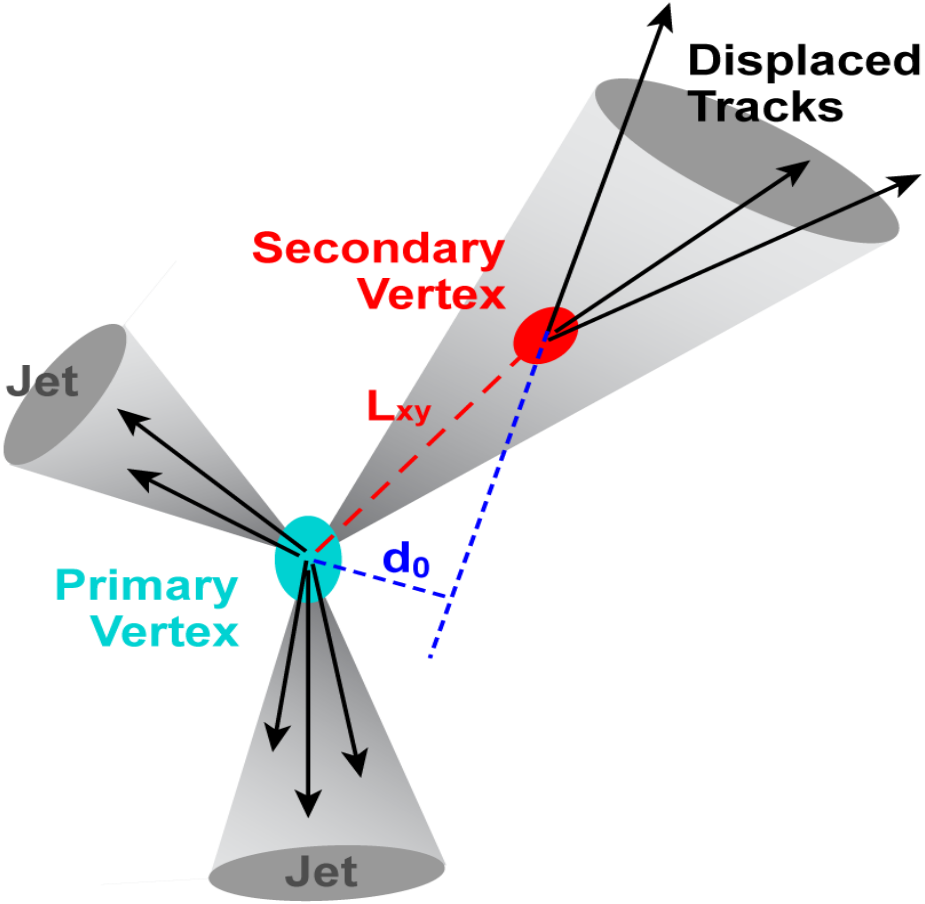
Jet tagging at LHCb

JINST 10 (2015) P06013

- **The jet tagging system takes advantage of LHCb features → precise vertex reconstruction!**

- As first step **Secondary Vertices** are reconstructed using tracks.

- A jet is identified to be generated from a **b** or **c** quark (**b-jet** or **c-jet**) if a **Secondary Vertex** is reconstructed within the jet cone (**ΔR <0.5**).



Jet tagging at LHCb

JHEP 01 (2014) 033

- Two **Boosted Decision Trees** (BDT) are used to identify b and c jets.

BDT(bc|udsg)
To separate **heavy flavour** jets from **light** jets

BDT(b|c)
To separate **b-jets** from **c-jets**

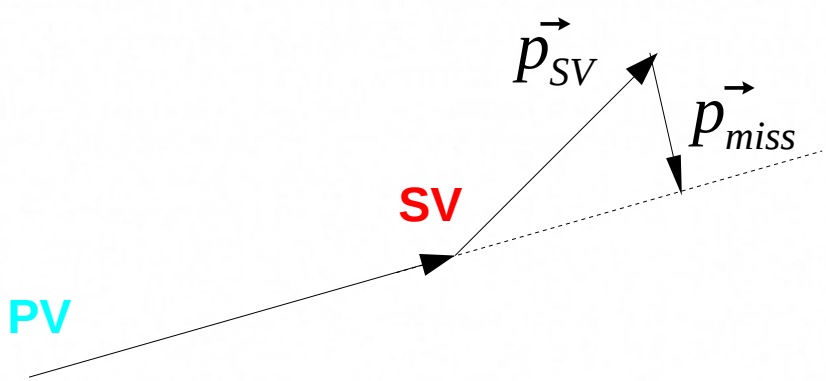
- **Some observables in input to the BDTs:**

SV mass

Fraction of jet p_T taken by the SV

Flight distance χ^2

SV corrected mass

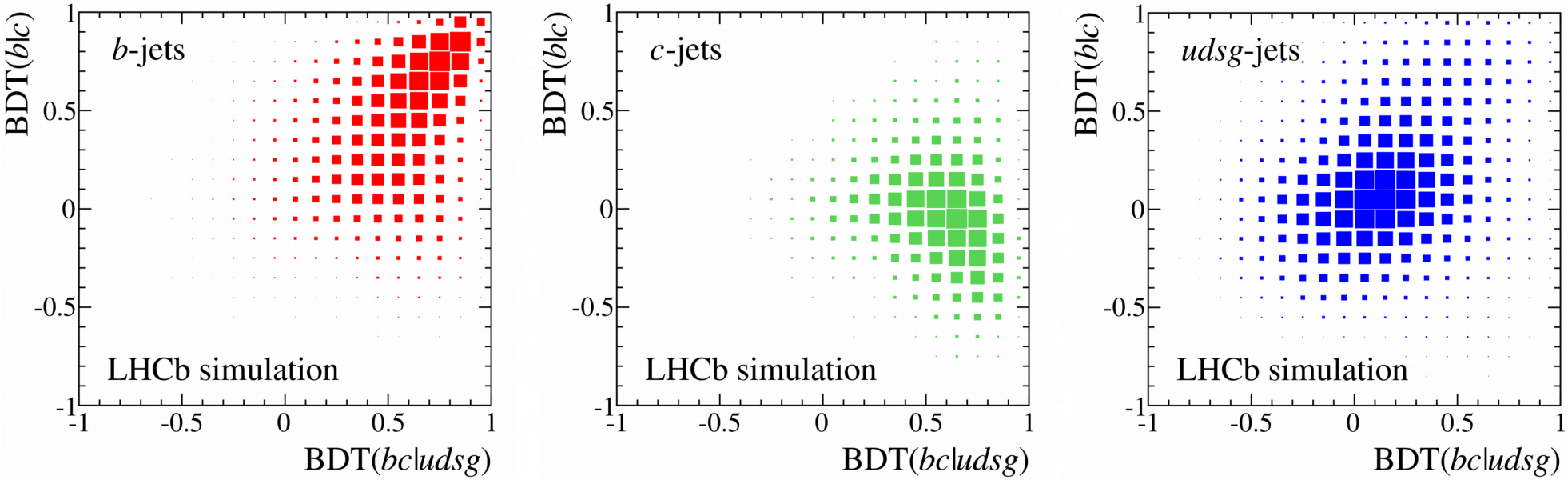


$$M_{corr} = \sqrt{M_{SV}^2 + p_{miss}^2} + p_{miss}$$

Jet tagging at LHCb

JHEP 01 (2014) 033

- Training samples of **b-jets**, **c-jets** and **light jets** are obtained from the Monte Carlo simulation.
- **A good discrimination power is achieved!**



Jet tagging at LHCb

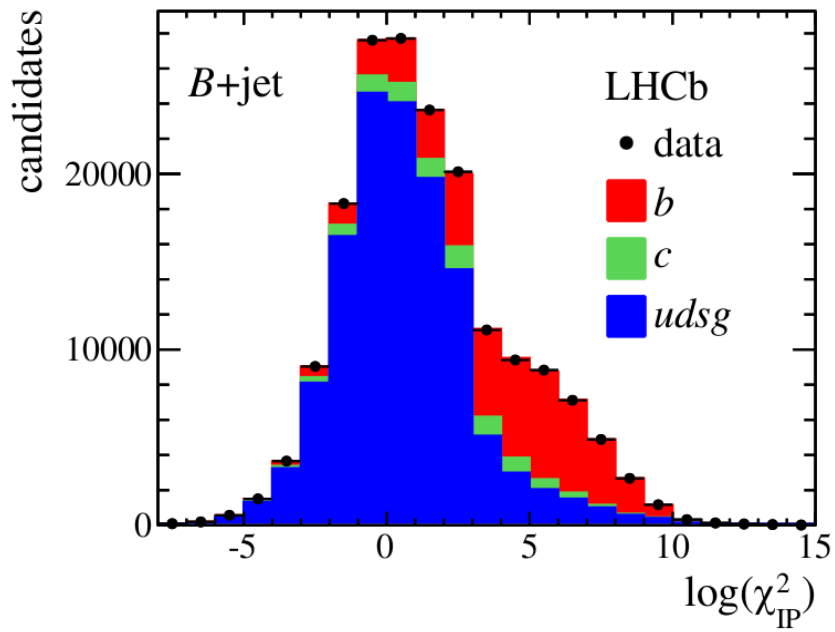
JHEP 01 (2014) 033

- **Tagging efficiencies are measured in data (Run I).**
- Events with a **jet** and a **probe** back-to-back to the jet in the azimuthal plane are selected.
- The **probe** identifies the jet flavour
  - **B+jet** (**b** enriched)
  - **D+jet** (**b** and **c** enriched)
  - **μ(displaced)+jet** (**b** and **c** enriched)
  - **μ(isolated and high P_T)+jet** (**light jets** enriched)

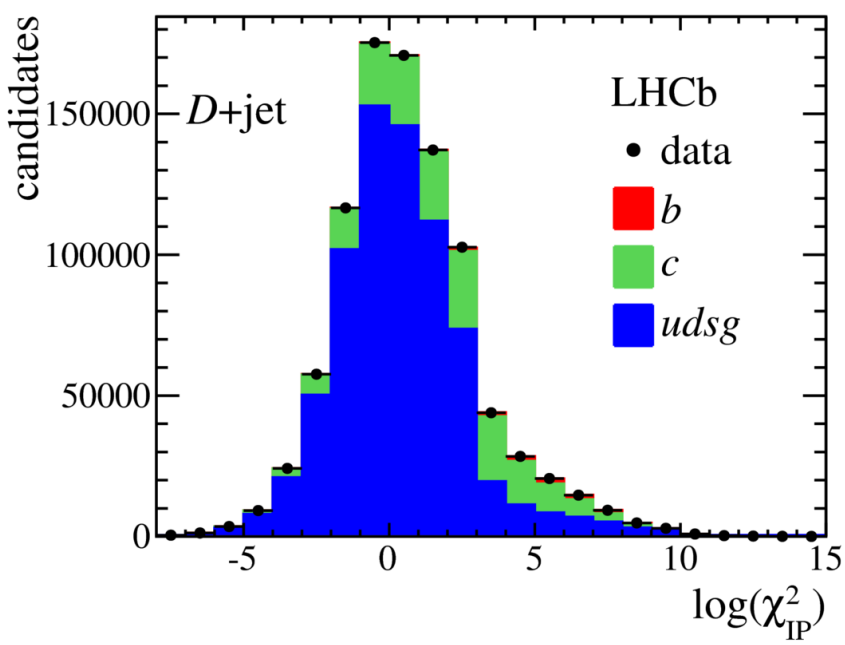
Jet tagging at LHCb

JHEP 01 (2014) 033

- Yields of **b**, **c** and **light** jets **prior to apply the SV-tag** are measured by fitting the distribution of the χ^2_{IP} associated to the highest p_T track in the jet.



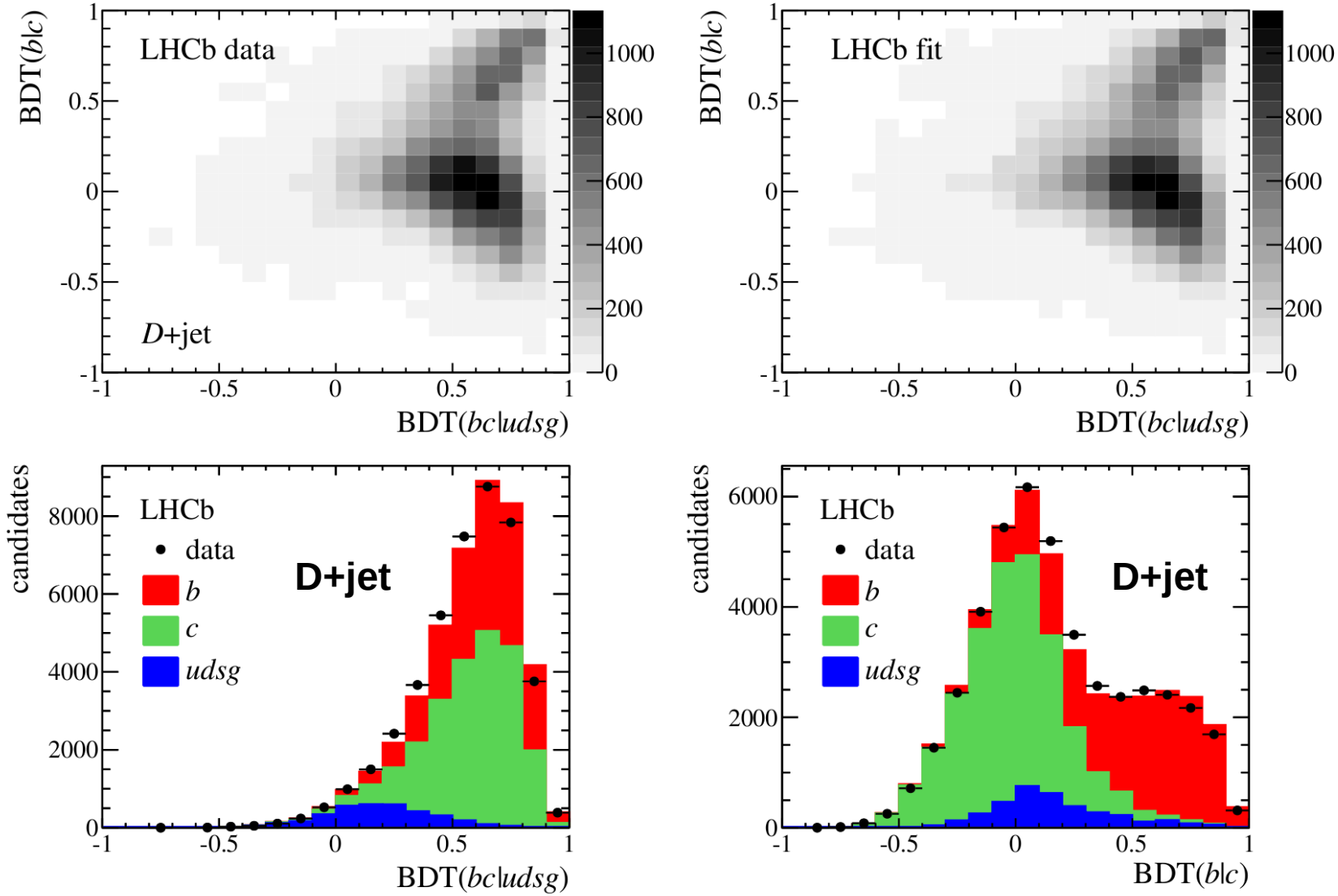
- Yields of **b**, **c** and **light** jets **after applying the SV-tag** are measured by fitting the 2-dimensional distribution of the BDTs (**next slide**).



- **Two-dimensional templates are obtained from simulation.**

Jet tagging at LHCb

JHEP 01 (2014) 033



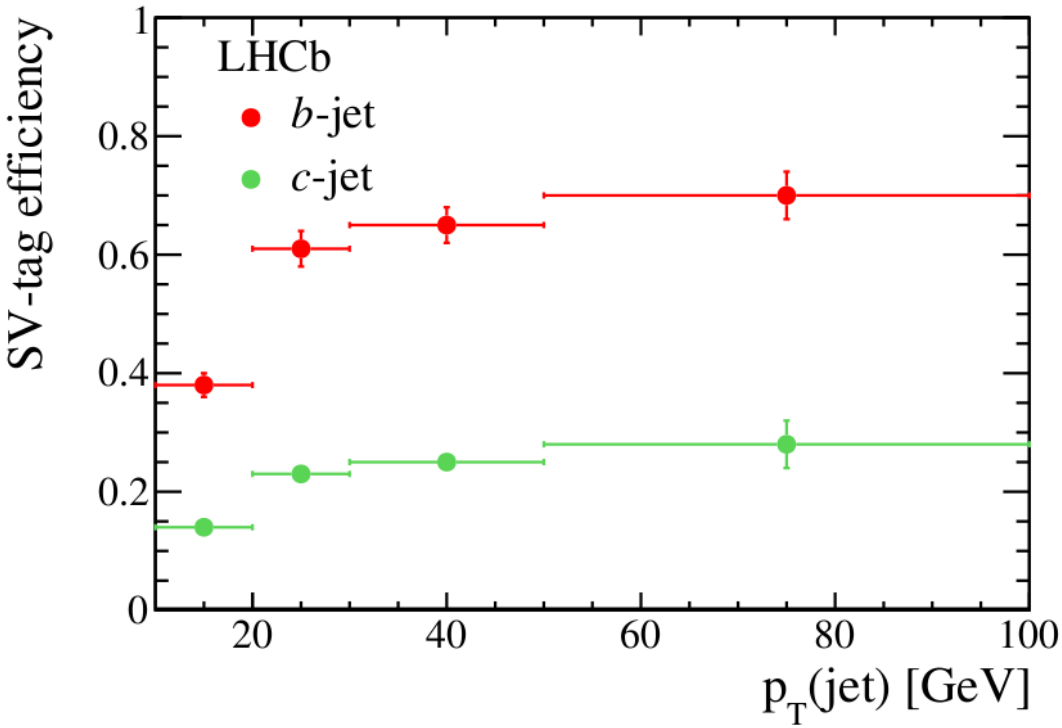
Jet tagging at LHCb

JHEP 01 (2014) 033

- Efficiencies obtained with: $\epsilon = \frac{N(pass)}{N(tot)}$

from BDTs fit

from χ^2_{IP} fit



Probability for a **b-jet** to be selected ~ **65%**

Probability for a **c-jet** to be selected ~ **25%**

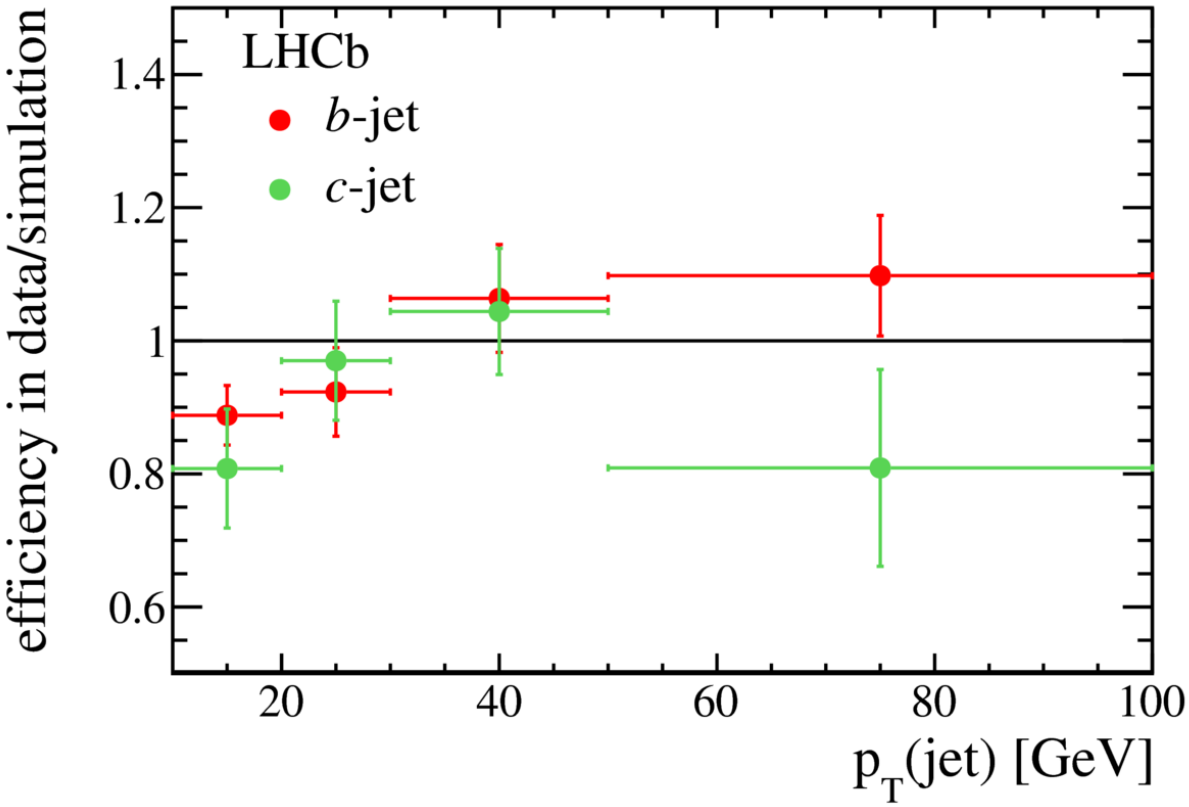
Probability to wrongly select a **light jet (g,u,d,s)** **~ 0.3%**

- Uncertainty due to the limited statistics of the data samples and to the modeling of the templates

Jet tagging at LHCb

JHEP 01 (2014) 033

- **Data/simulation efficiencies scale factors:**

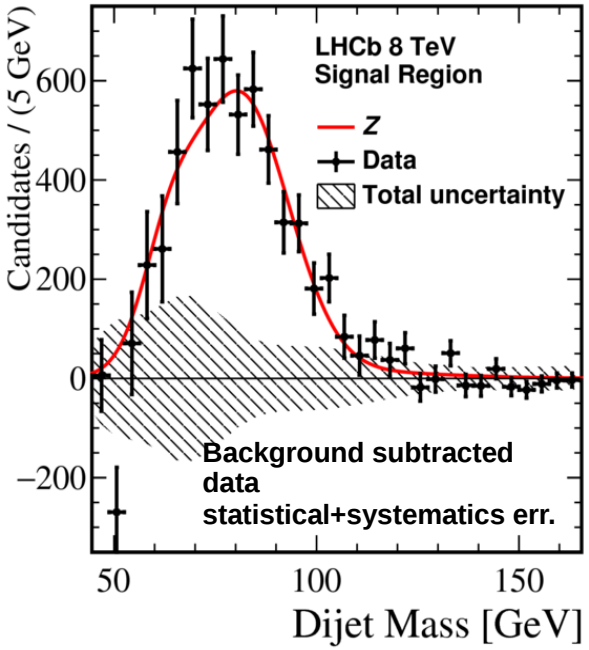


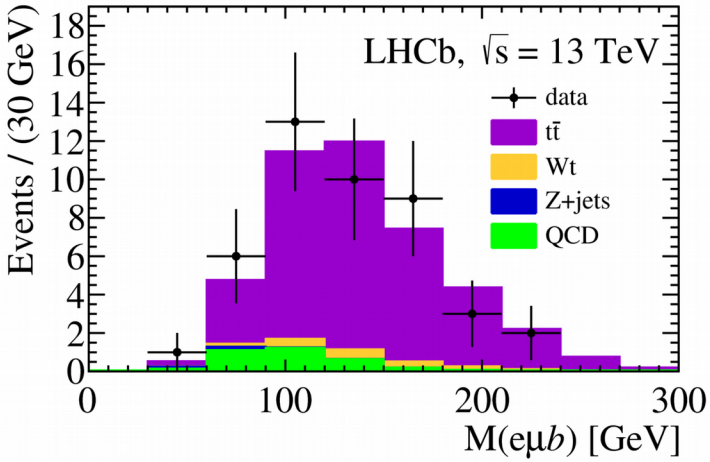
- Uncertainty due to the limited statistics of the data samples and to the modeling of the templates, in the order of ~10%

Jet tagging at LHCb

- **Many measurements have been performed thanks to the jet tagging!**
- The most recent ones:
  - $Z \to b\bar{b}$ cross section in the forward region (**dedicated talk in Electroweak session on Friday**)
  - $t\bar{t}$ production in the $t\bar{t} \to \mu e b$ final state (**dedicated talk in Top session tomorrow by D. Lucchesi**)

Phys. Lett. B776 (2017) 430-439



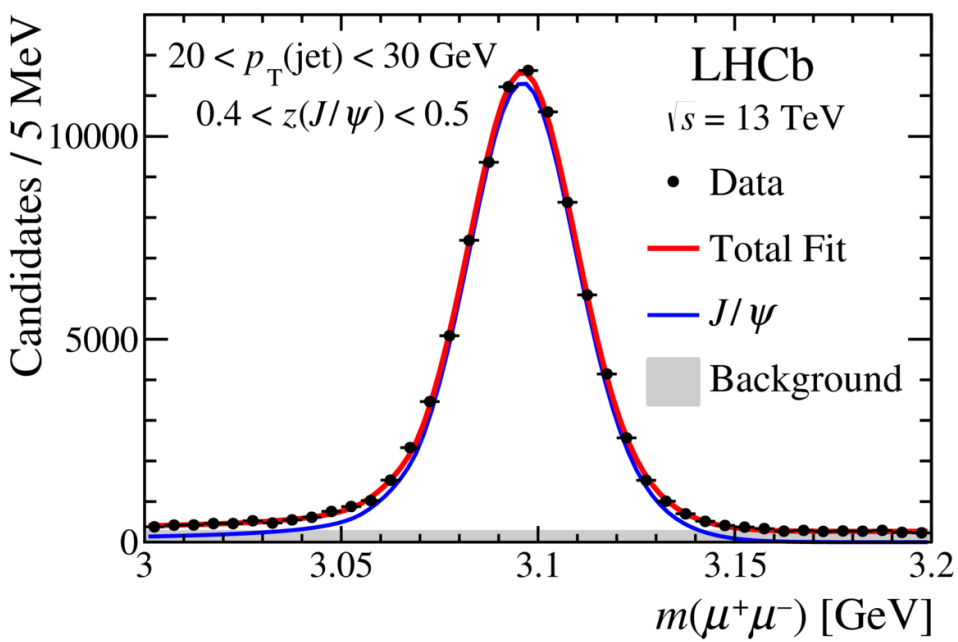


LHCb-PAPER-2017-050

Study of J/ψ in jets

Phys. Rev. Lett. 118, 192001 (2017)

- **Measurement of J/ψ production in jets can be used as probe of QCD.**
- LHCb measured **$z(J/\psi) = p_T(J/\psi)/p_T(jet)$** for prompt J/ψ and for those produced in b-hadron decays, using the 13 TeV dataset (1.4 fb^{-1}).
- Jets must have p_T>20 GeV and 2.5 < η < 4.0 → **Reconstructed at trigger level!**
- **$J/\psi \rightarrow \mu^+\mu^-$** in the jet must have $p_T(\mu)$ > 0.5 GeV and p(μ)>5 GeV

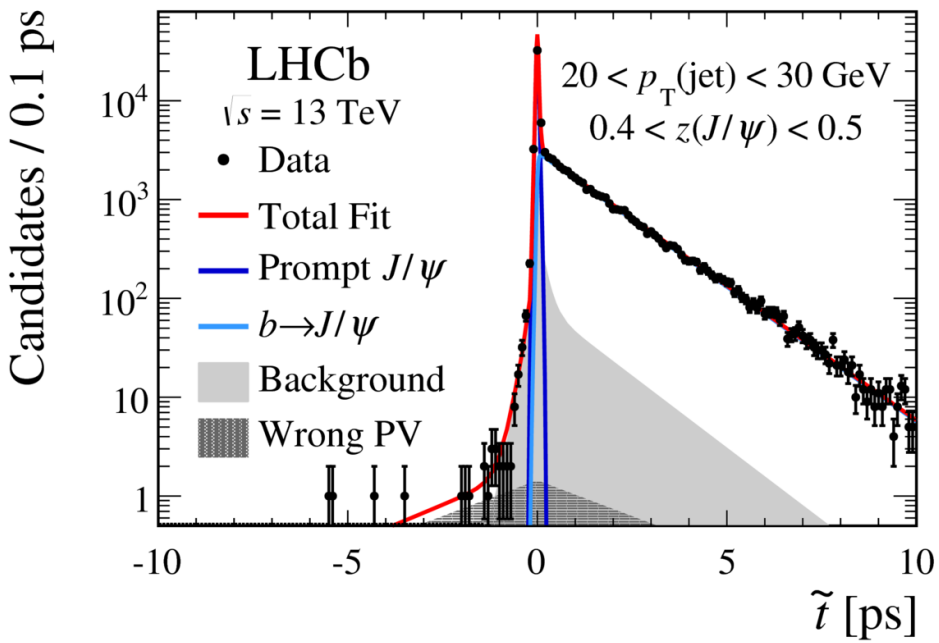


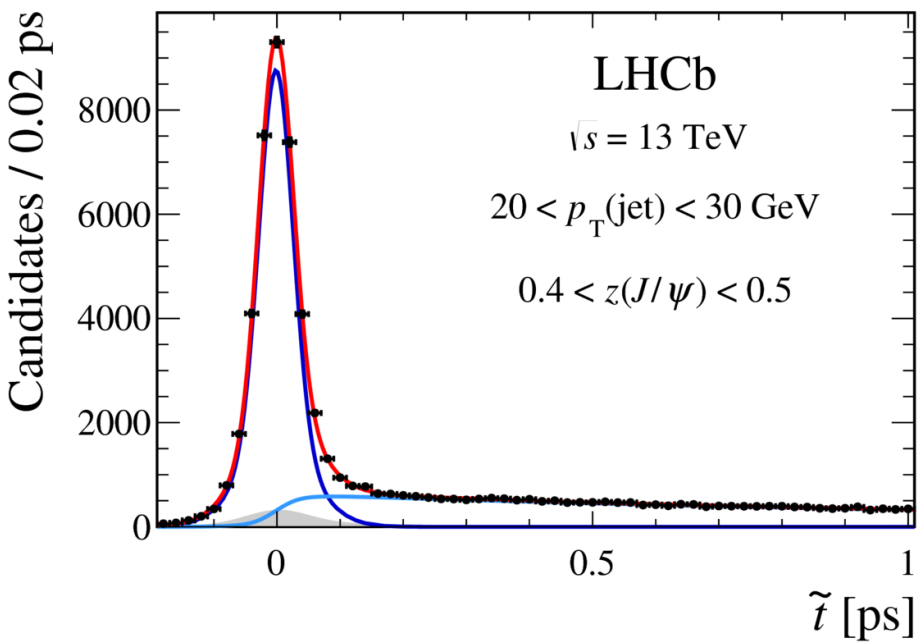
Di-muon mass fit in one particular p_T(jet) and z(J/ψ) bin, used to determine the J/ψ yield

Study of J/ψ in jets

Phys. Rev. Lett. 118, 192001 (2017)

- A fit to the pseudo decay time $\tilde{t} = \lambda\, m(J/\psi)/p_L(J/\psi)$ is performed in $[p_T(\text{jet}), z(J/\psi)]$ bins to determine the prompt and displaced J/ψ yields.



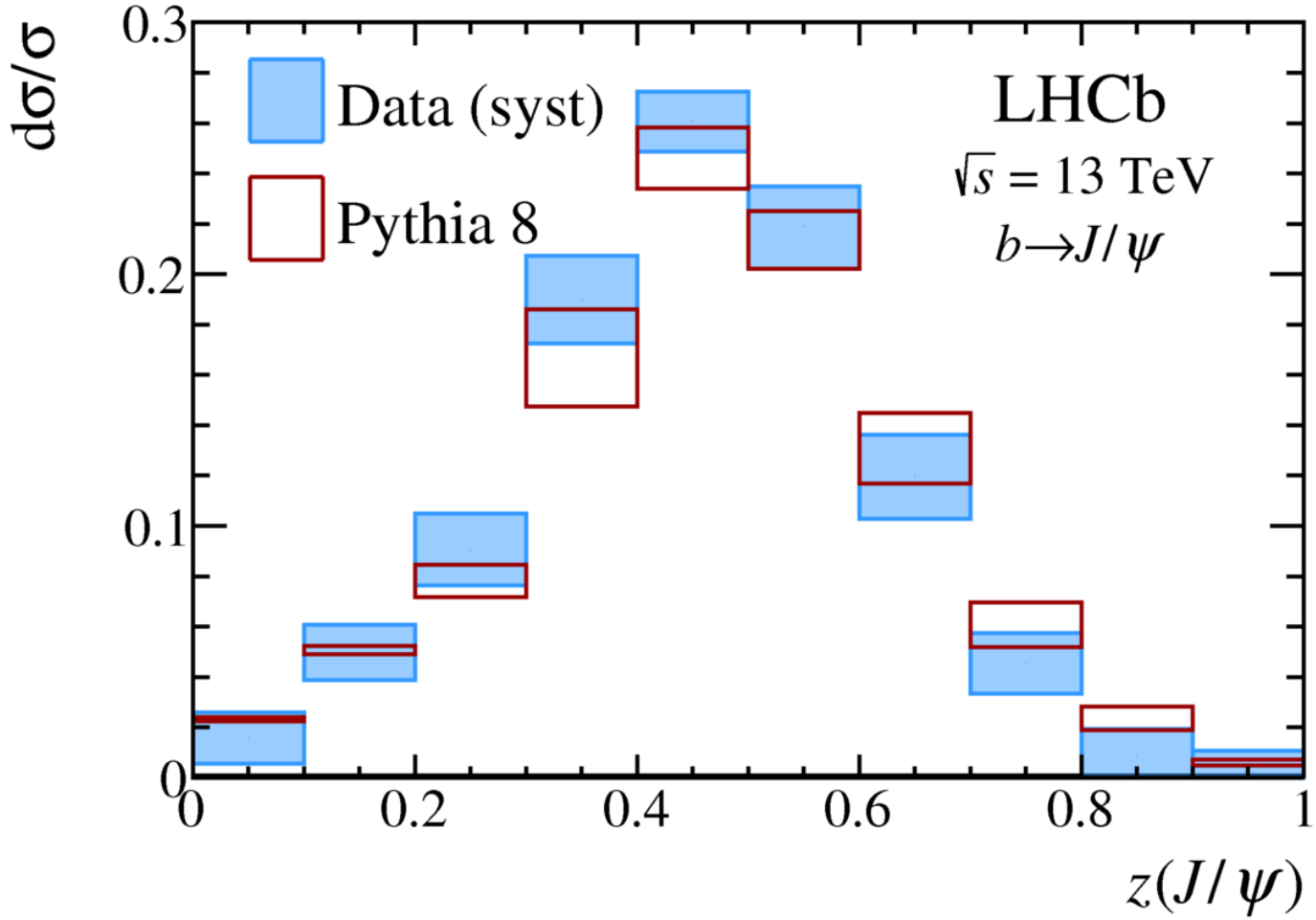


Study of J/ψ in jets

Phys. Rev. Lett. 118, 192001 (2017)

- The measured z(J/ψ) distributions are corrected for the selection efficiencies and unfolded for the detector response.

- Measurements are compared with theorical predictions:

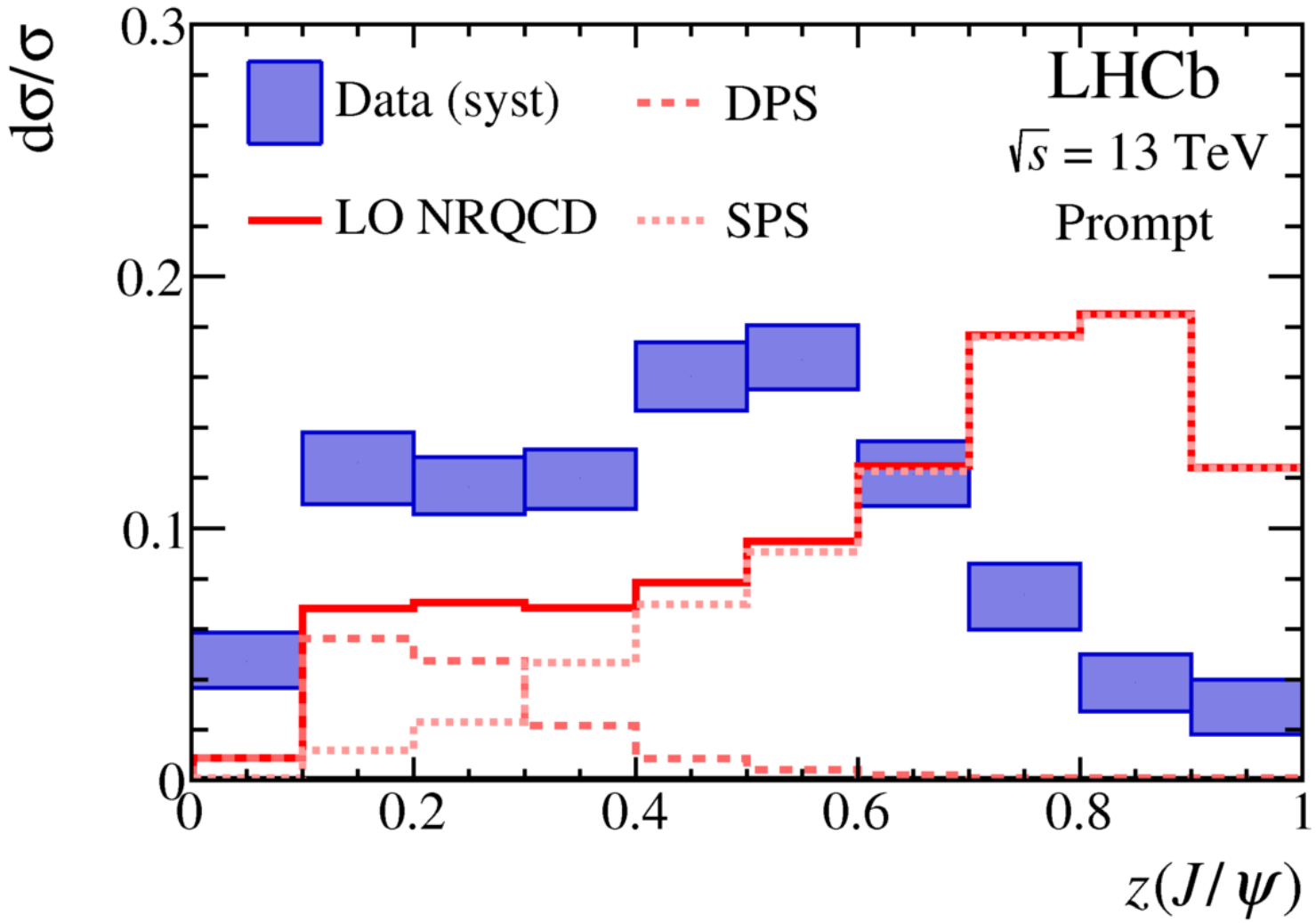


Displaced case

Study of J/ψ in jets

Phys. Rev. Lett. 118, 192001 (2017)

- Measurements are compared with theorical predictions: **discrepancies in the prompt case trigger a discussion among theorists**!



Prompt case

Future prospects and conclusions

- At **LHCb** the jet substructure has been studied for the identification of b- and c- jets (secondary vertices) and for QCD studies (J/ψ in jets).

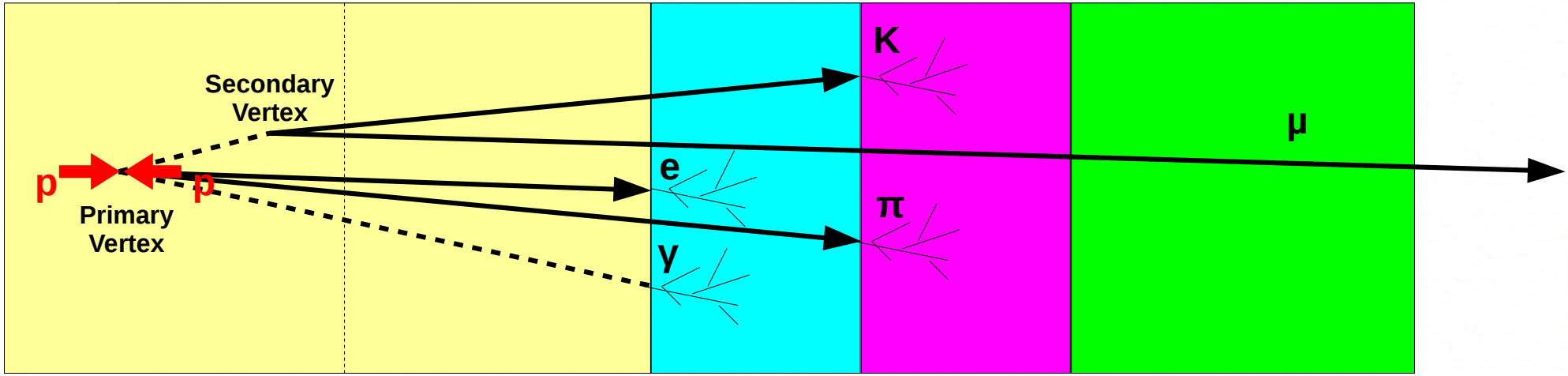
- Thanks to LHCb unique features an **excellent heavy flavour tagging system** has been developed.

- **Work is in progress to improve jet tagging performances:** we are considering to use all the jet constituents, not only the secondary vertices.

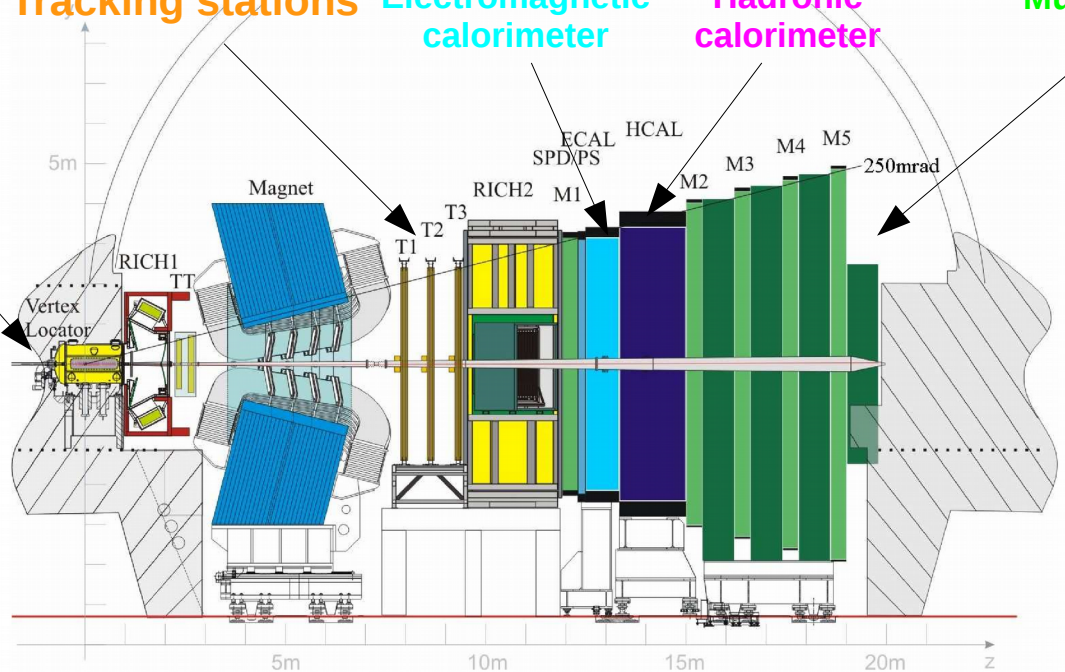
- **We are studying the possibility to tag merged bb and cc fat jets at LHCb:**
  - ➔ it will be an important tool for low mass resonances searches
  - ➔ it can be used to study the b-pairs production mechanism (i.e. gluon splitting)

Backup slides

Jet detection at LHCb

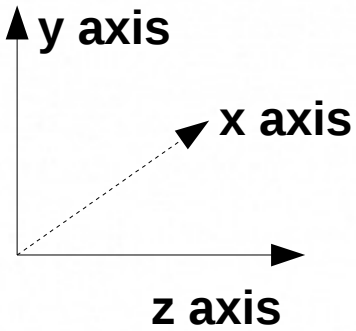


Vertex Locator | Tracking stations | Electromagnetic calorimeter | Hadronic calorimeter | Muon System



Jet reconstruction inputs:

- Tracks
- Calorimeter clusters
- Metastable particles (like K_S)



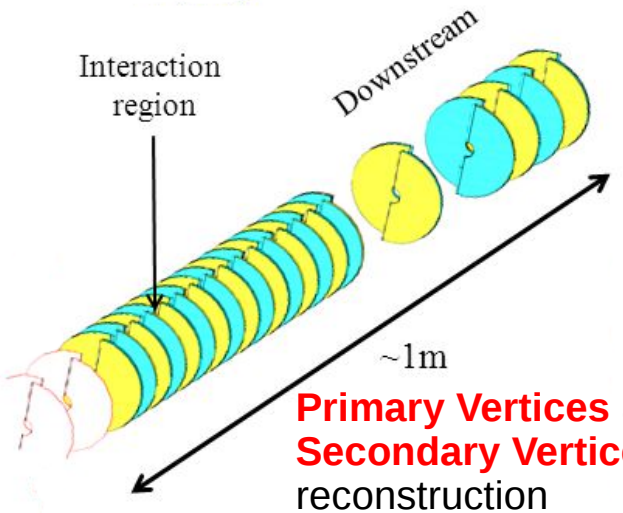
Tracking system

Tracking at LHCb: **silicon microstrip** (VELO, Inner Tracker), **drift tubes** (Outer tracker)

VErtex LOcator (VELO)

↓

21 stations → (r,φ) coordinates

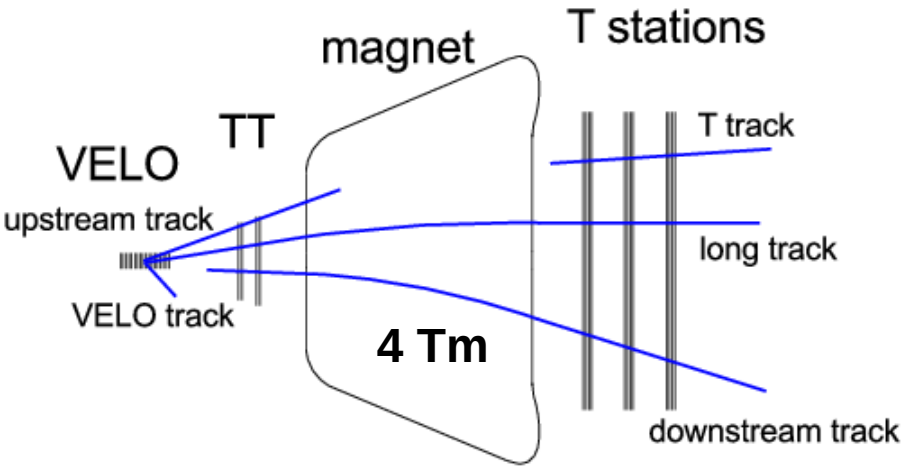


Primary Vertices and **Secondary Vertices** reconstruction

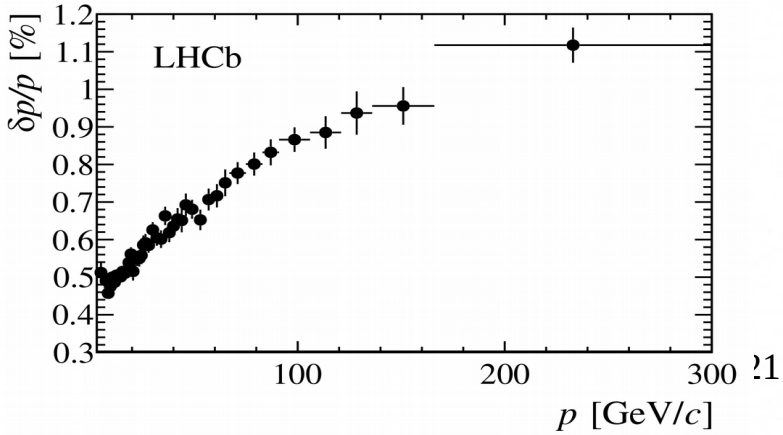
Tracking stations

↓

4 stations → (x,y) coordinates



The **momentum** of **charged particles** is determined by measuring **the curvature of the trajectory in the magnetic field**

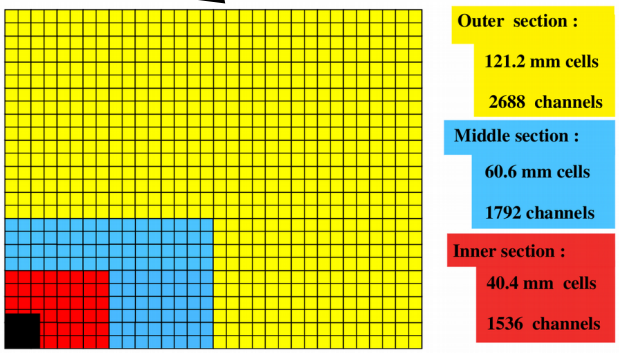


Calorimeters

Electromagnetic calorimeter

e, γ, π^0 produce electromagnetic showers in lead layers

showers are detected by layers of scintillating fibers

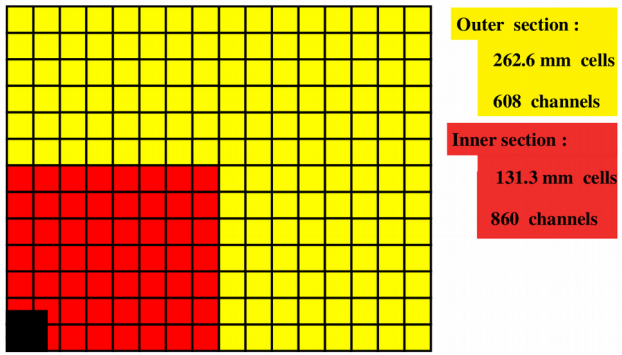


$$\frac{\sigma_E}{E} = \frac{10\,\%}{\sqrt{E}} \oplus 1\,\%$$

Hadronic calorimeter

K, π and other hadrons produce hadronic showers in iron layers

showers are detected by layers of scintillating tiles



$$\frac{\sigma_E}{E} = \frac{69\,\%}{\sqrt{E}} \oplus 10\,\%$$

Not optimal for jets physics!

Limitations due to saturation

Energy resolution

Inputs for jets reconstruction

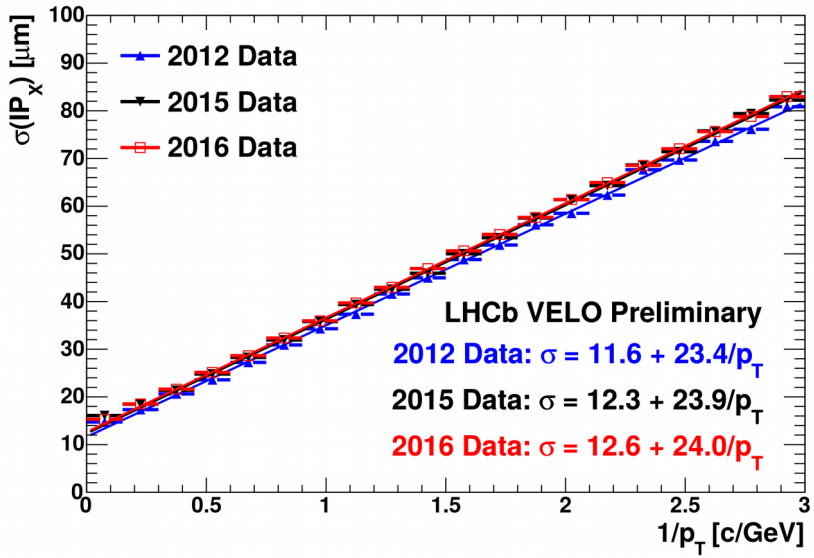
- Clusters isolated from tracks (neutral particles)
- Excesses of energy nearby tracks (neutral recovery)

LHCb performance

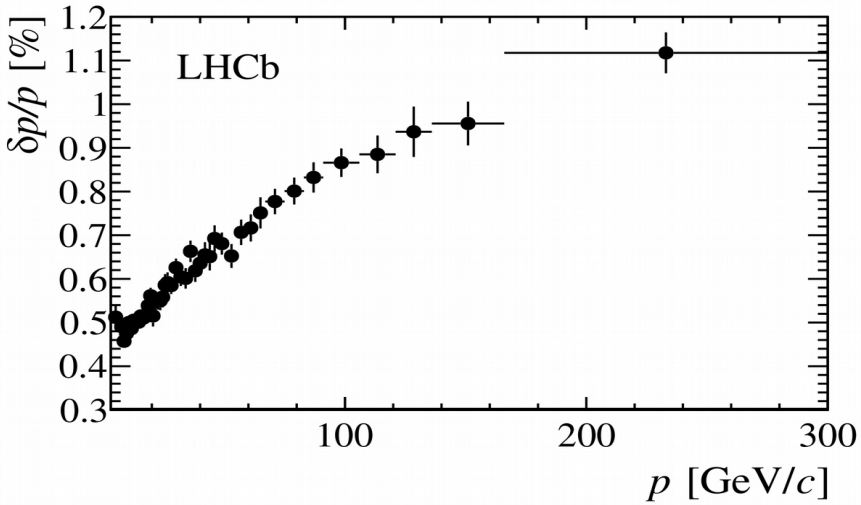
Int. J. Mod. Phys. A 30, 1530022 (2015)

Excellent IP resolution

→ very important for SV reconstruction and tagging



Tracks momentum resolution



Electromagnetic calorimeter

$$\frac{\sigma_E}{E} = \frac{10\,\%}{\sqrt{E}} \oplus 1\,\%$$

Limitations due to saturation

Energy resolution

Hadronic calorimeter

$$\frac{\sigma_E}{E} = \frac{69\,\%}{\sqrt{E}} \oplus 10\,\%$$

Calorimeter clusters in input to jets reconstruction

→ **Clusters isolated from tracks (neutral particles)**

→ **Excesses of energy nearby tracks (neutral recovery)**