Transformer networks for constituent-based b-jet calibration with the ATLAS detector

Brendon Bullard

Postdoc @ SLAC

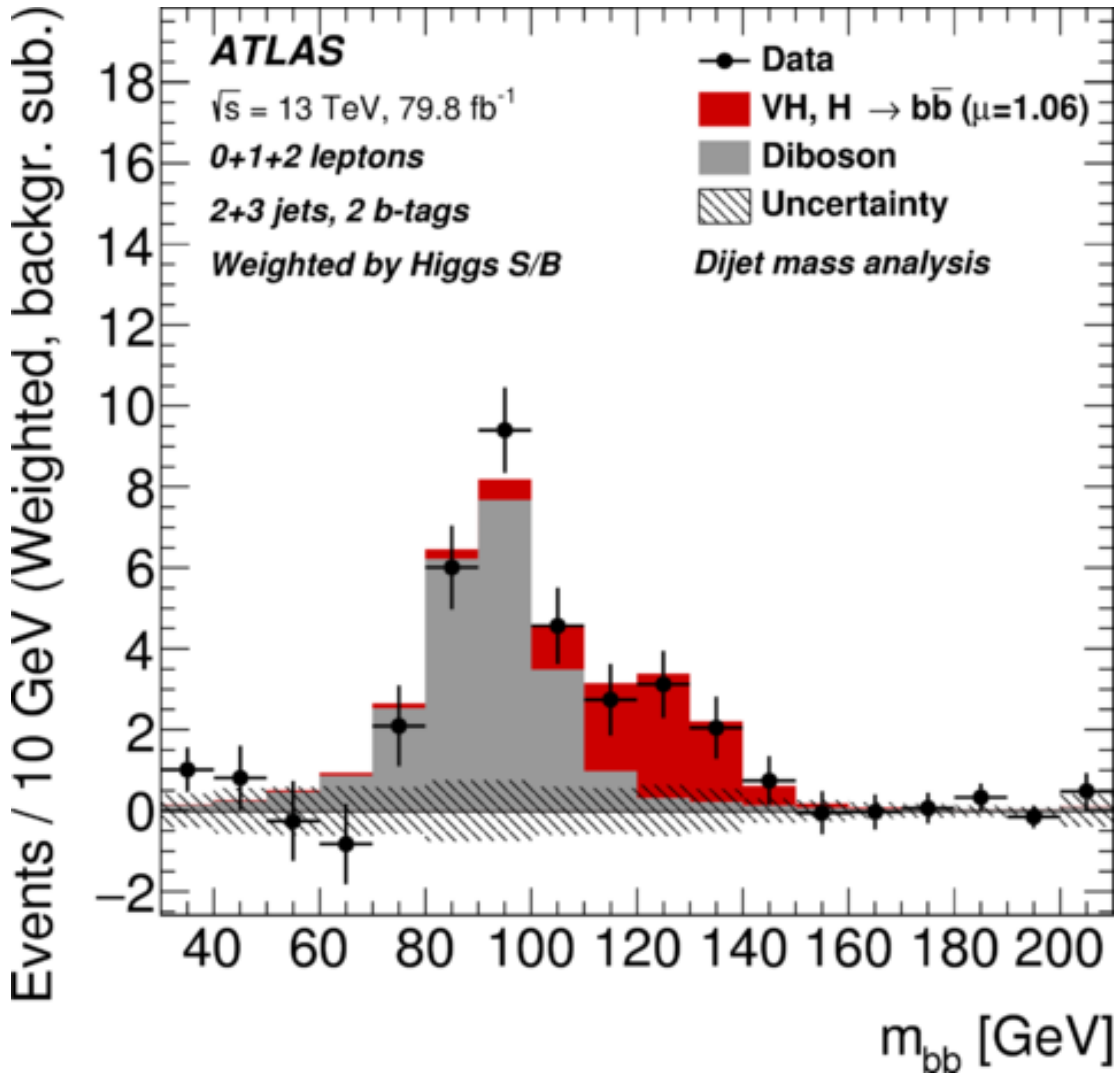
US-LUA lightning talk
December 18, 2024

ATLAS EXPERIMENT

SLAC NATIONAL ACCELERATOR LABORATORY

Physics motivation

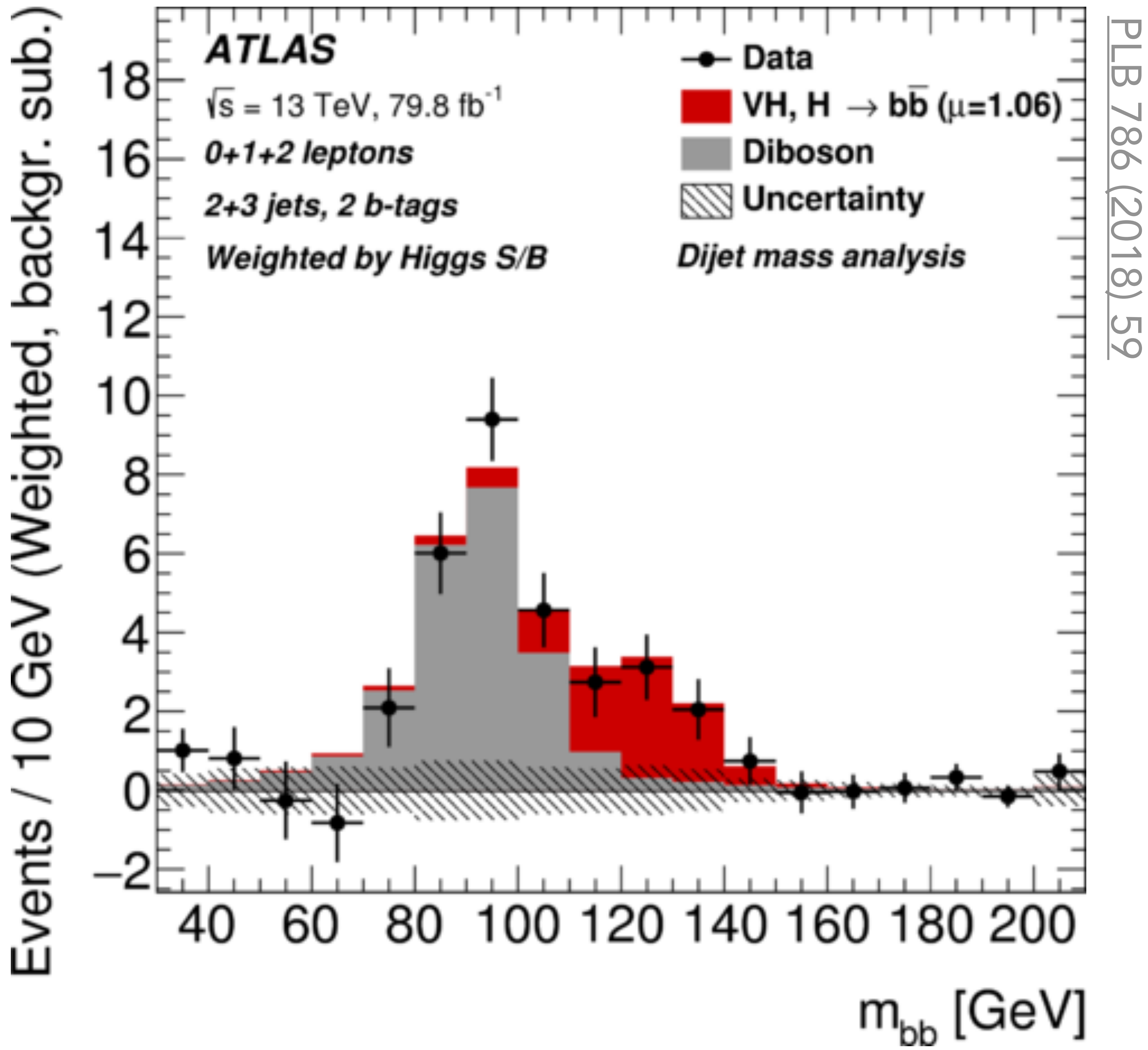
Measuring $H \to b\bar{b}$ constrains **bottom Yukawa**



PLB 786 (2018) 59

Physics motivation

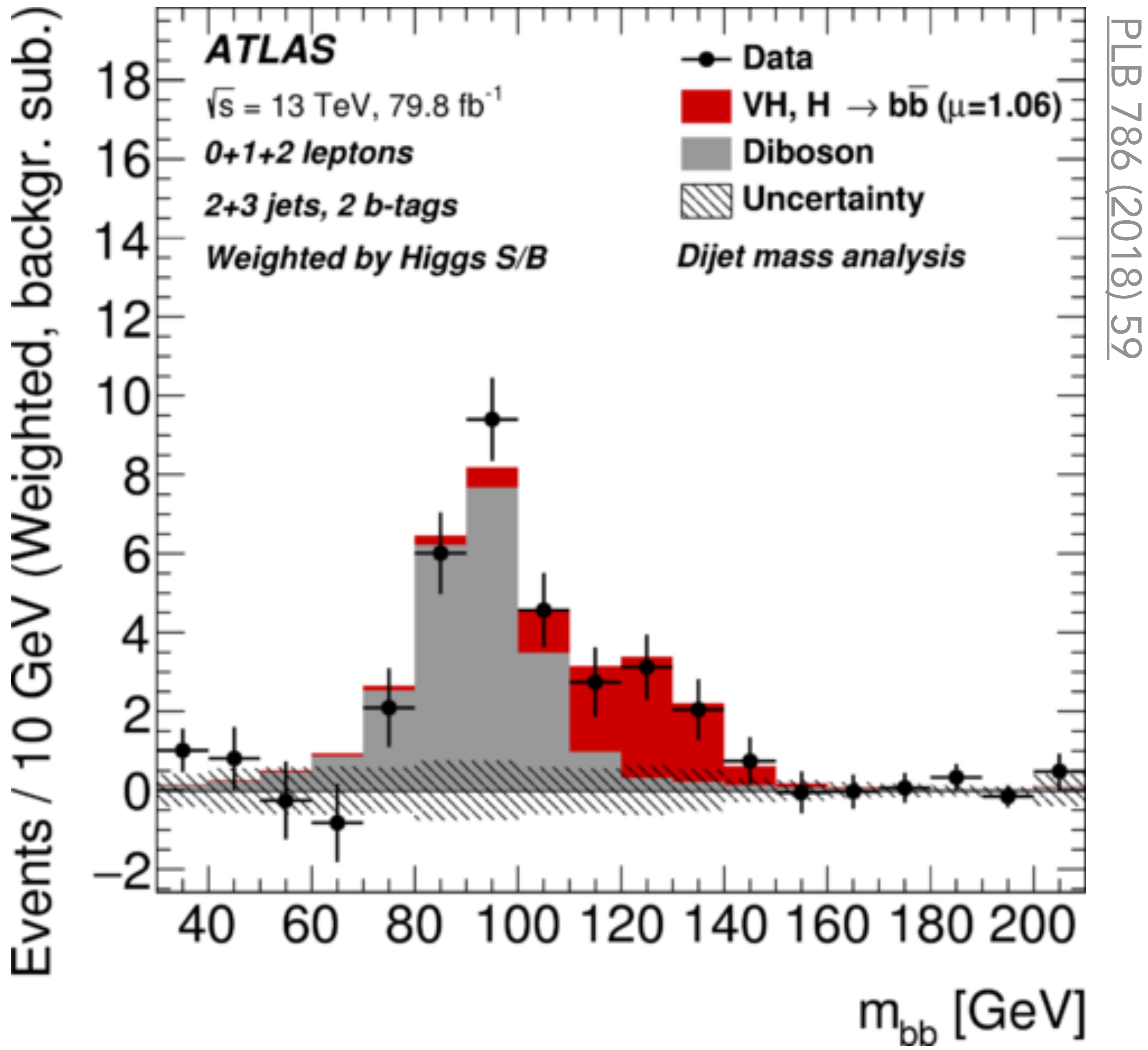
Measuring $H \to b\bar{b}$ constrains **bottom Yukawa**



Limited by **poor jet momentum resolution**

Physics motivation

Measuring $H \to b\bar{b}$ constrains **bottom Yukawa**



Limited by **poor jet momentum resolution**

Di-Higgs production is a critical target, gives handle on Higgs **self-coupling**

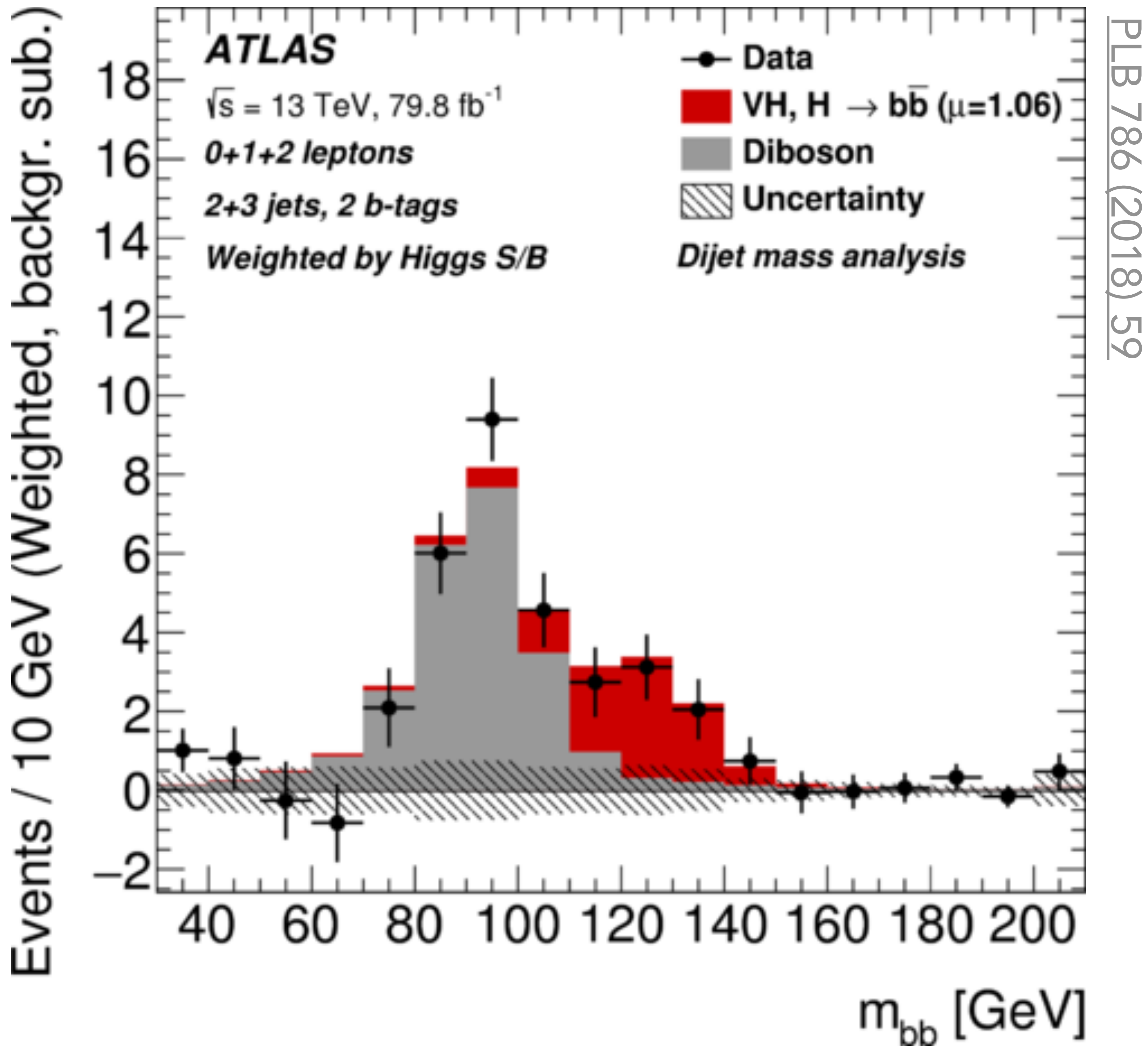
	bb	WW	ττ	ZZ	γγ
bb	34%				
WW	25%	4.6%			
ττ	7.3%	2.7%	0.39%		
ZZ	3.1%	1.1%	0.33%	0.069%	
γγ	0.26%	0.10%	0.028%	0.012%	0.0005%

70%!

Branching fractions of the the two Higgs

Physics motivation

Measuring $H \rightarrow b\bar{b}$ constrains **bottom Yukawa**



Limited by **poor jet momentum resolution**

Di-Higgs production is a critical target, gives handle on Higgs **self-coupling**

	bb	WW	ττ	ZZ	γγ
bb	34%				
WW	25%	4.6%			
ττ	7.3%	2.7%	0.39%		
ZZ	3.1%	1.1%	0.33%	0.069%	
γγ	0.26%	0.10%	0.028%	0.012%	0.0005%

70%!

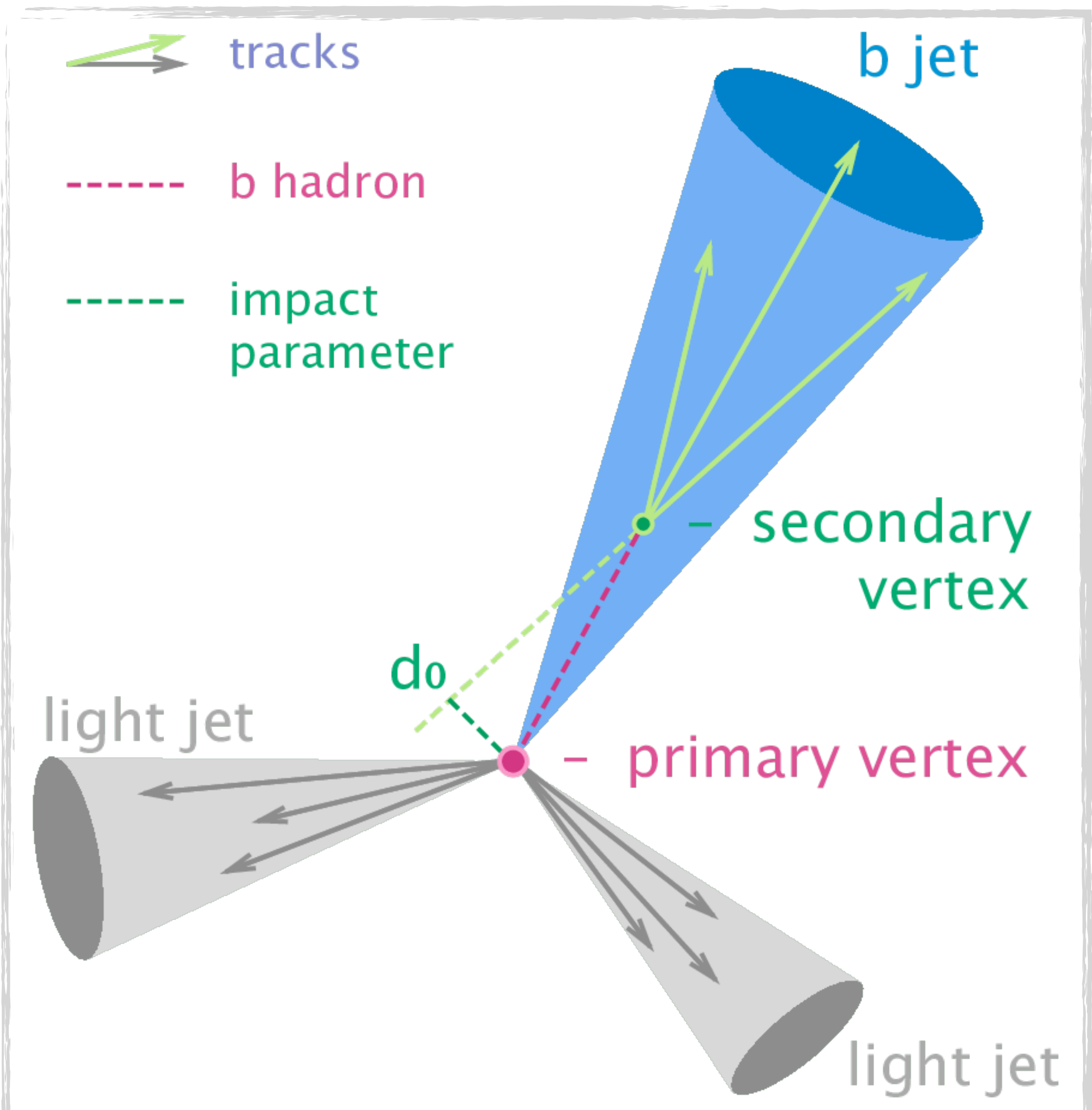
Branching fractions of the the two Higgs

Improving the **reconstruction of b-quark jets** has huge impact

Machine learning for b-jets

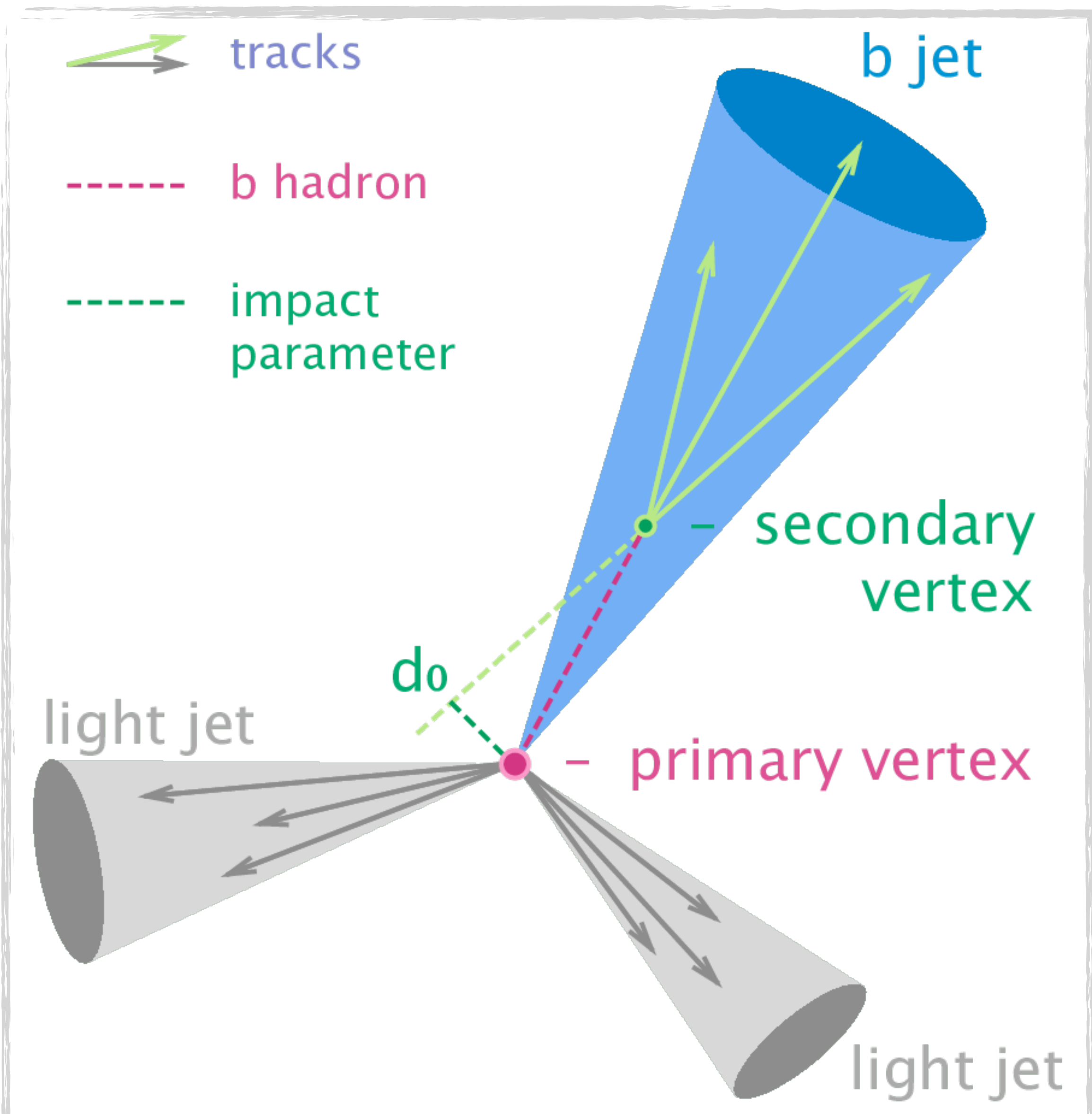
Machine learning for b-jets

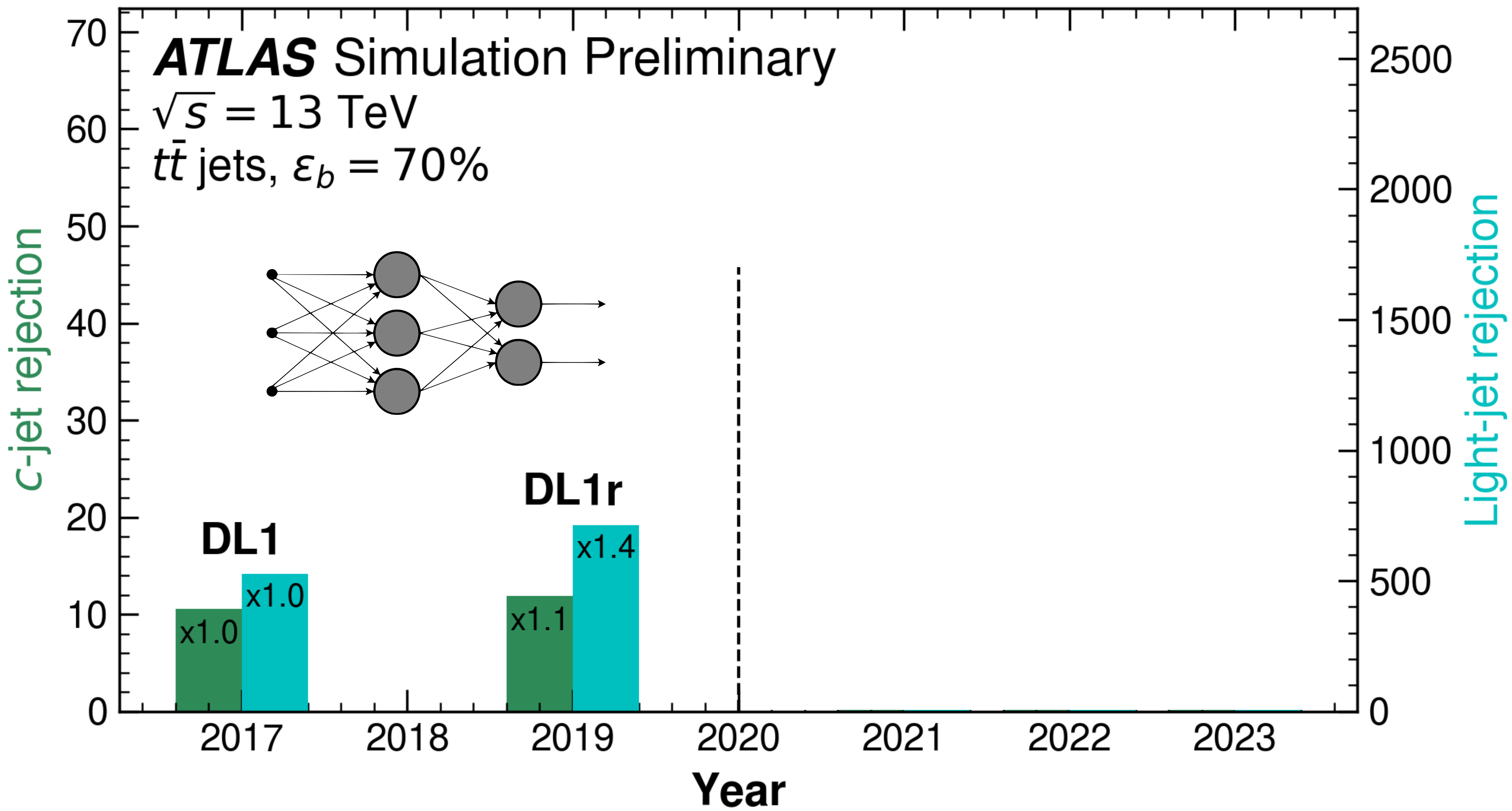
- Identification of b-jets has had a long history of using ML
  - Secondary vertices, many tracks and unique radiation pattern product rich substructure



Machine learning for b-jets

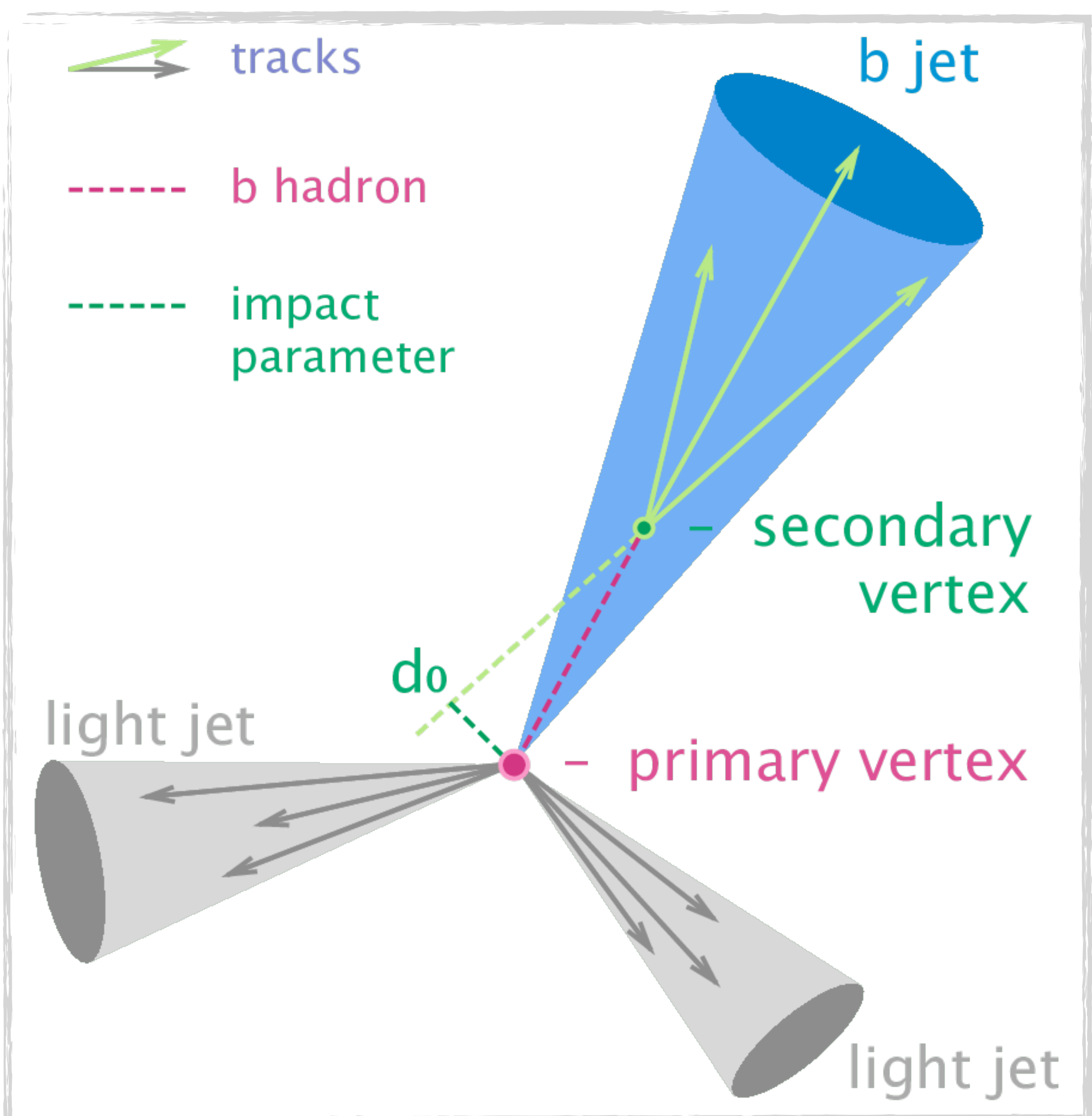
- Identification of b-jets has had a long history of using ML
  - Secondary vertices, many tracks and unique radiation pattern product rich substructure

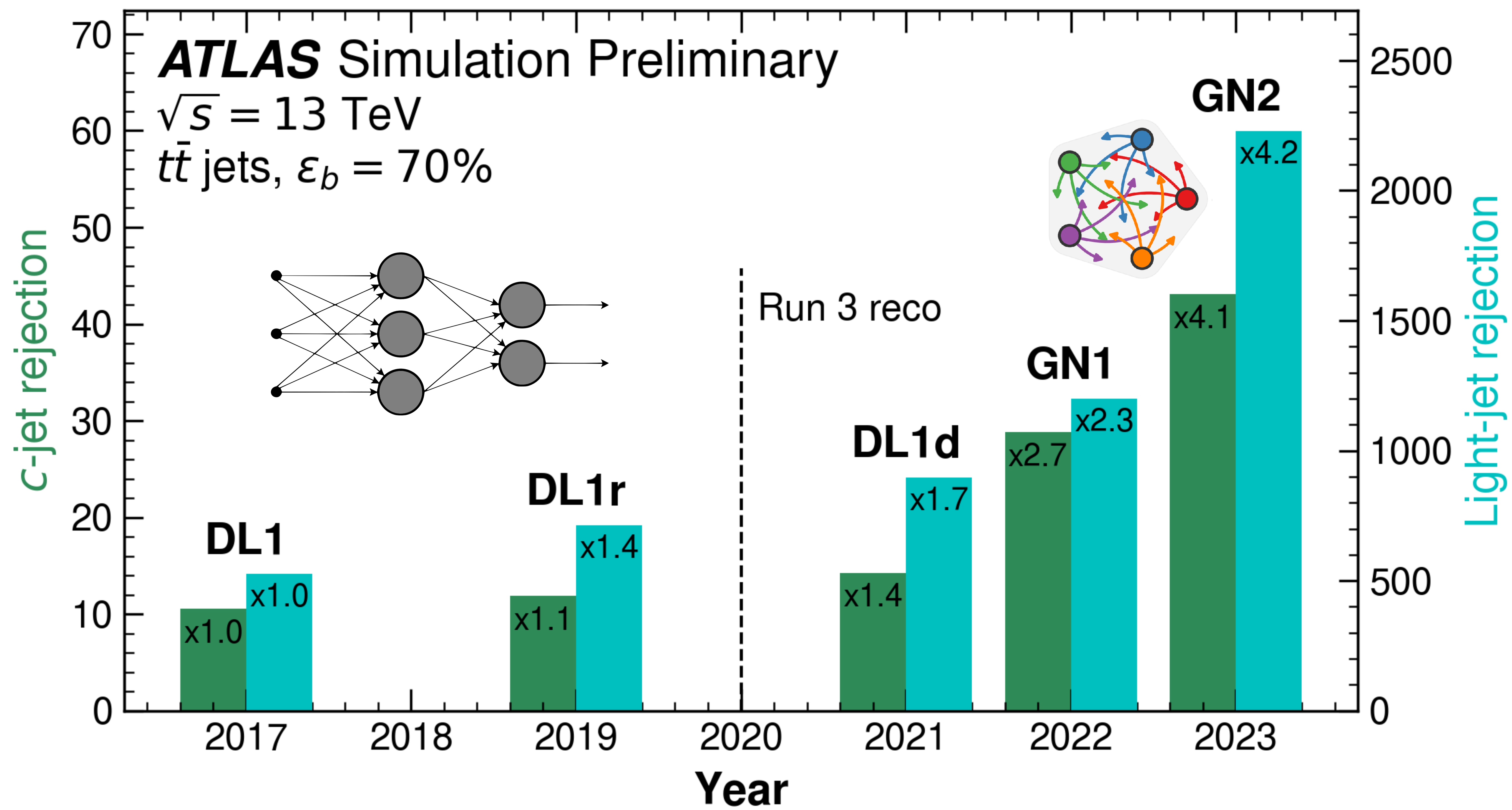




Machine learning for b-jets

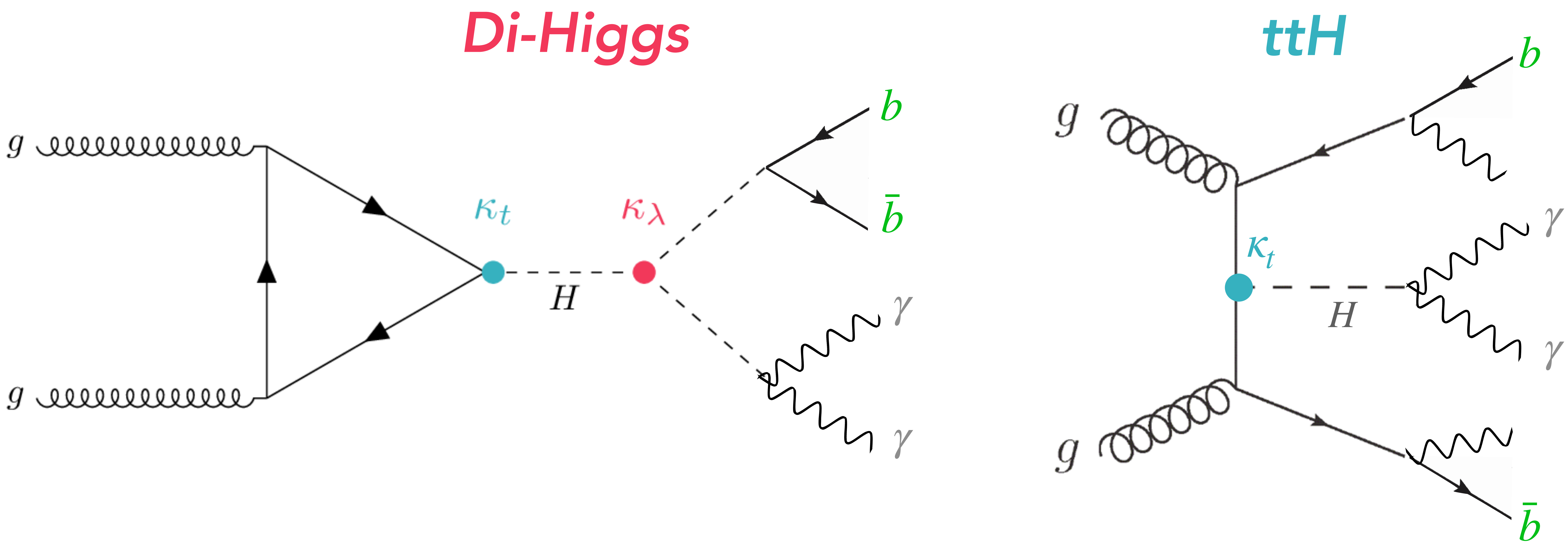
- Identification of b-jets has had a long history of using ML
  - Secondary vertices, many tracks and unique radiation pattern product rich substructure
- Flavor classification has dramatically improved with graph and transformer neural networks trained on low-level information → apply to p_T regression





The limits of flavor tagging

- In searches for Di-Higgs, single Higgs is now a background
  - Not improved with light-jet rejection - primary differentiator is $m_{b\bar{b}}$!



Need to exploit **well-resolved kinematics** to separate and exploit complementarity

ATLAS PUB Note

ATL-PHYS-PUB-2024-015

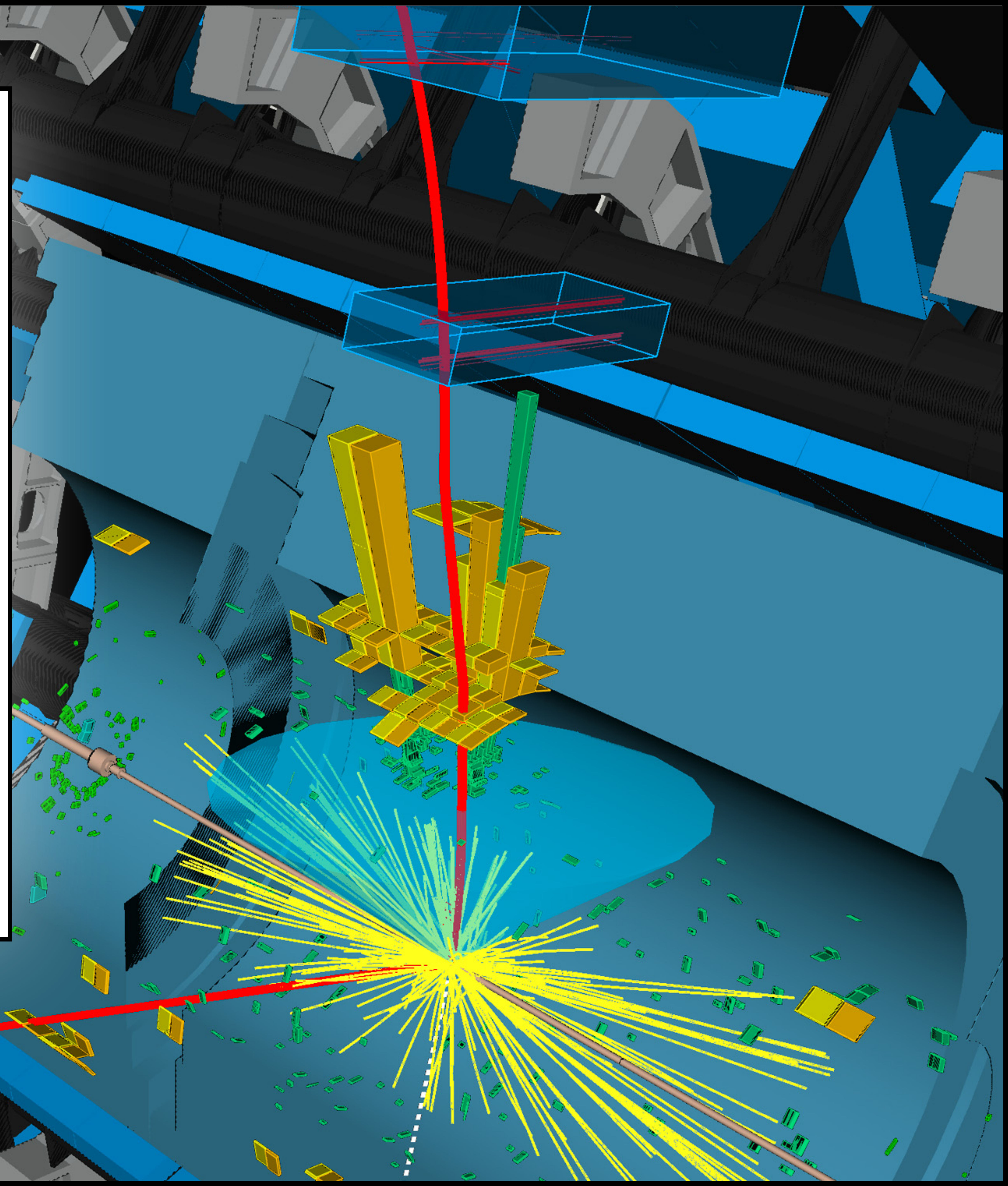
27th July 2024

Transformer networks for constituent-based b-jet calibration with the ATLAS detector

The ATLAS Collaboration

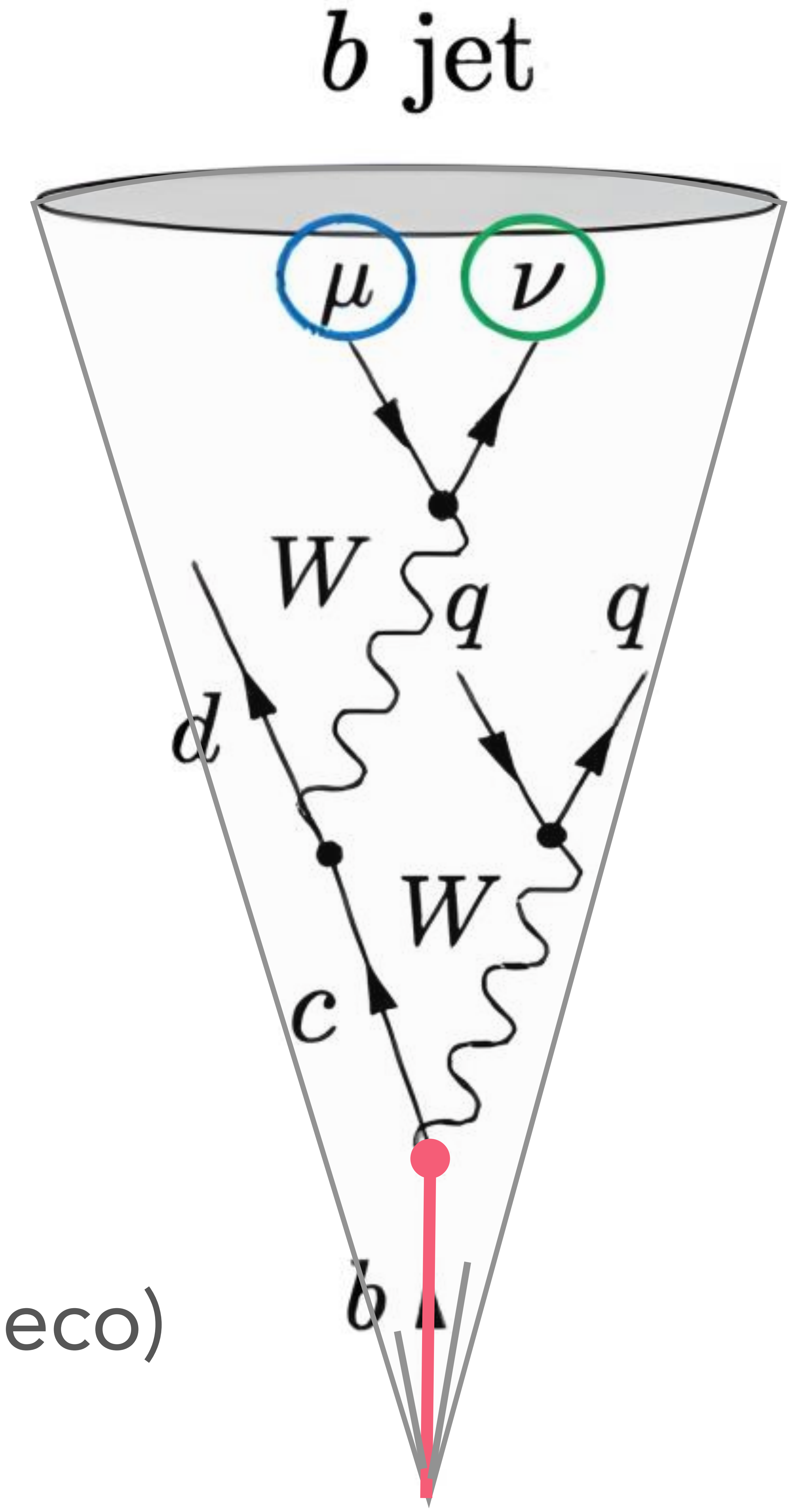
The precise measurement of a jet's kinematics is a critical component of the physics program based on proton–proton collision data recorded by the ATLAS detector at the Large Hadron Collider. The determination of the energy and mass of jets containing bottom quarks b-jets is particularly difficult as, for example, they have different radiation patterns compared to the average jet and can contain heavy-flavour decays into a charged lepton and an unobserved neutrino. This document reports on a novel calibration technique for jets focusing on b-jets using transformer-based neural networks trained on simulation samples to correct reconstructed jet properties to the true values. Separate simulation-based regression methods have been developed to estimate the transverse momentum of small-radius jets and the transverse momentum and mass of large-radius jets. In both cases, the regression methods move the median measurement closer to the true value. A relative resolution improvement with respect to the nominal calibration between 18% and 31%, depending on the transverse momentum, is demonstrated for small-radius jets. Both the large-radius jet transverse momentum and mass resolution are shown to improve by 25–35%.

ATL-PHYS-PUB-2024-015



Physics of b-jets

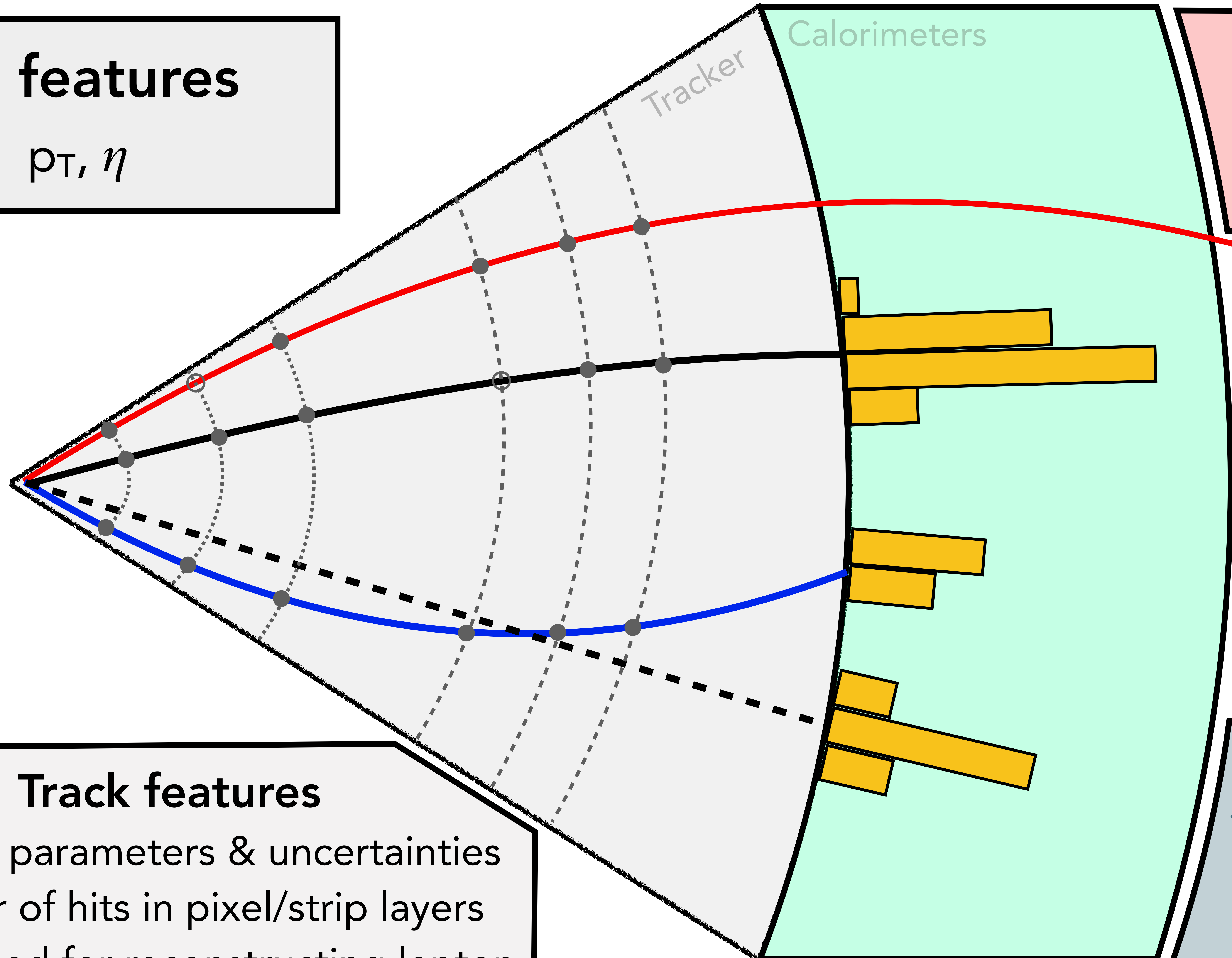
- Unique due to secondary vertex
  - B-hadron carries >80% of jet energy
  - Semi-leptonic decays (~15% BR), ν carries 20% of the jet energy
  - Charged lepton should correlate to ν
  - Leptons more important than in tagging
- Baseline strategy:
  - **µ-in-jet** addition to jet 4-mom
  - Coarse-grained correction binned in jet p_T, split into leptonic and hadronic decays (PtReco)



Neural network input features

Jet features

p_T, η



Soft muon features

Kinematics, perigee, p_T balance in ID/MS,

Soft electron features

Kinematics, perigee, E/p, E_{ECal}/E_{HCal}, layer cluster ratios, p_{HF}

Training dataset

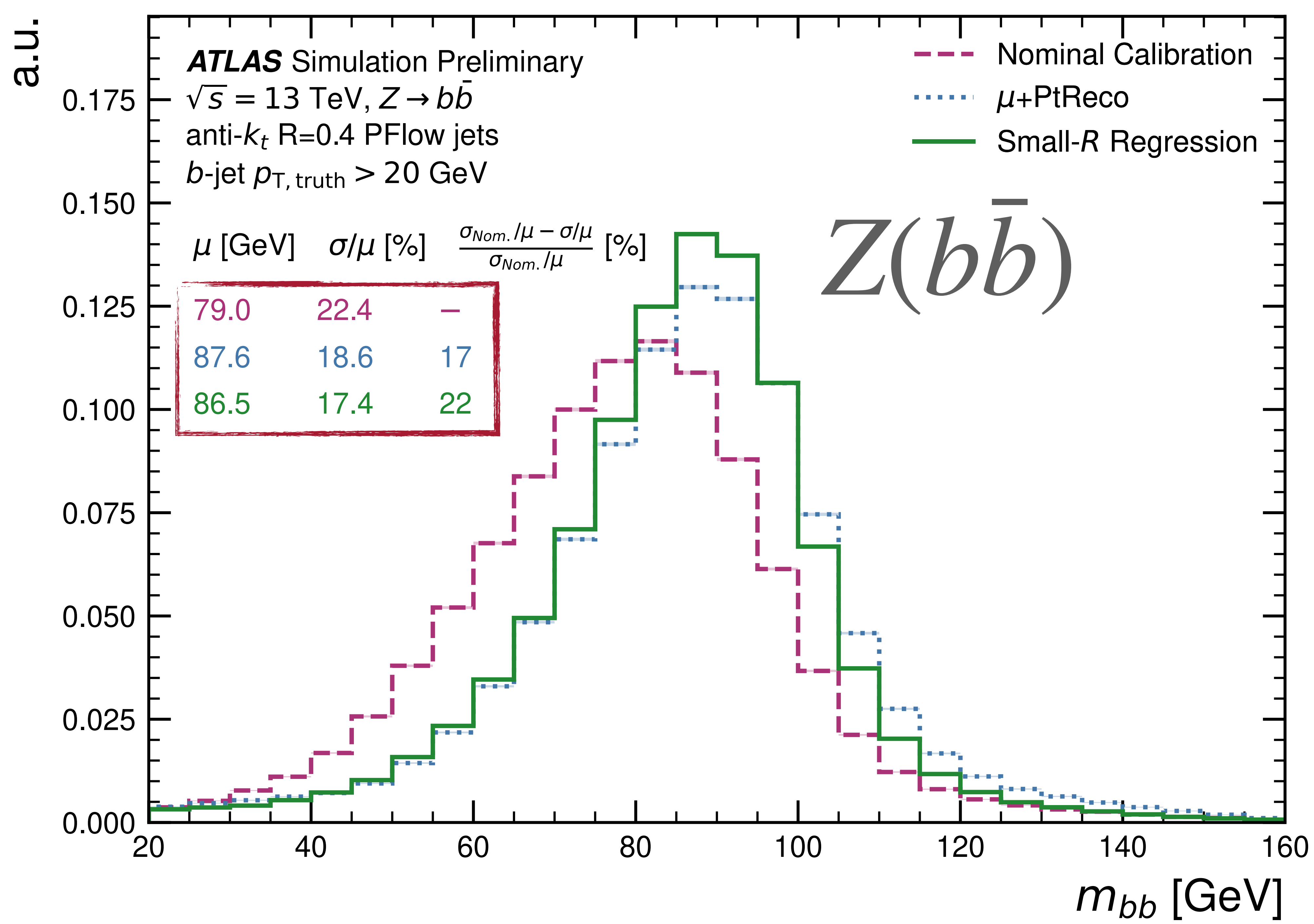
~260M jets from $t\bar{t}$, Z+jet

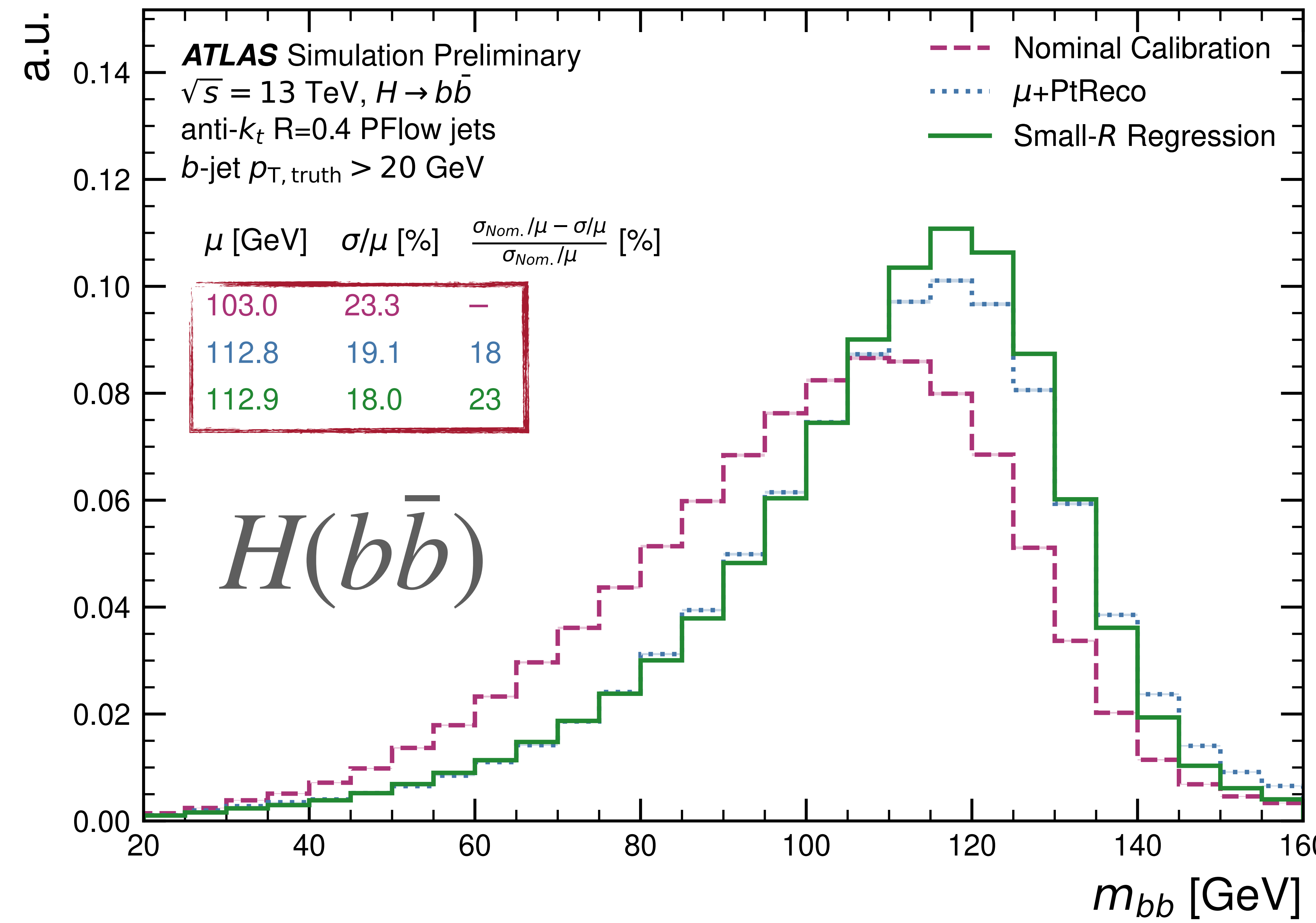
Mixture of light, charm, and b-jets (!)

Track features

- Perigee parameters & uncertainties
- Number of hits in pixel/strip layers
- Track used for reconstructing lepton

Resonance resolution

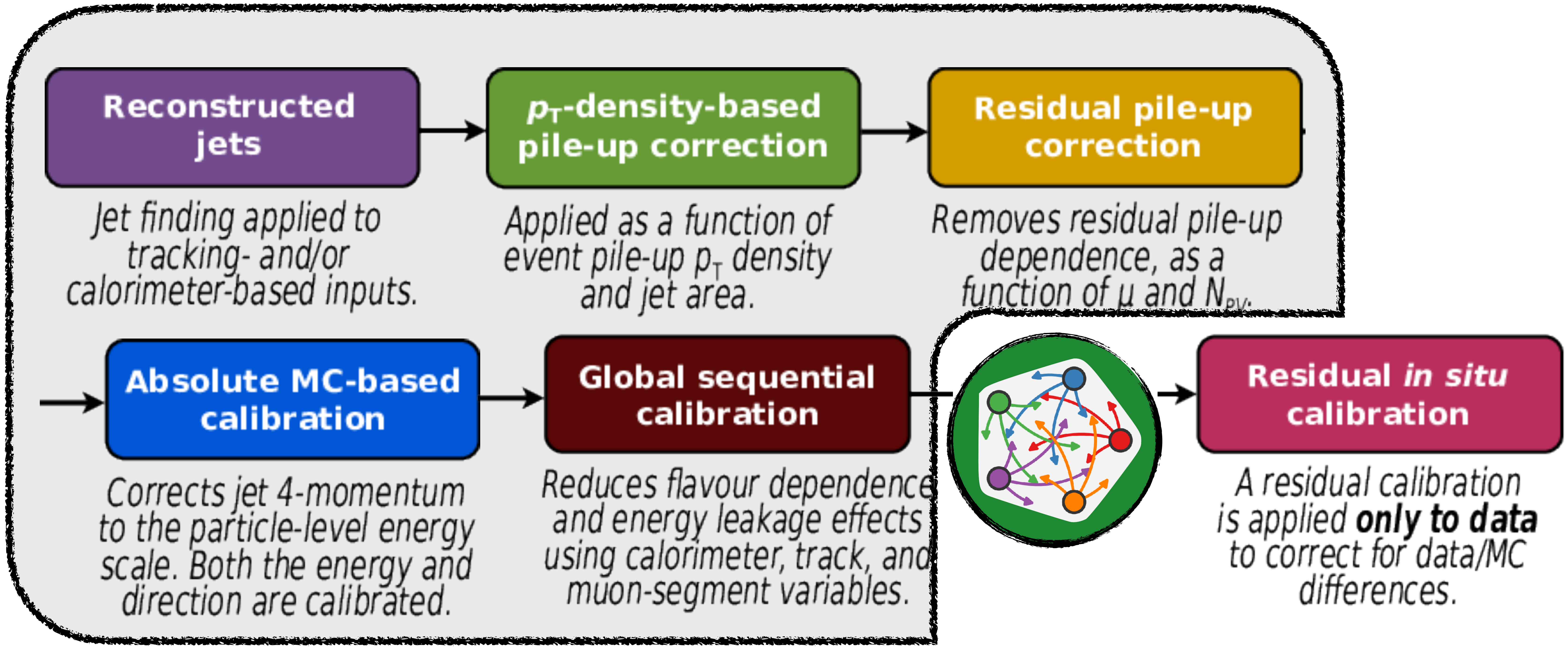


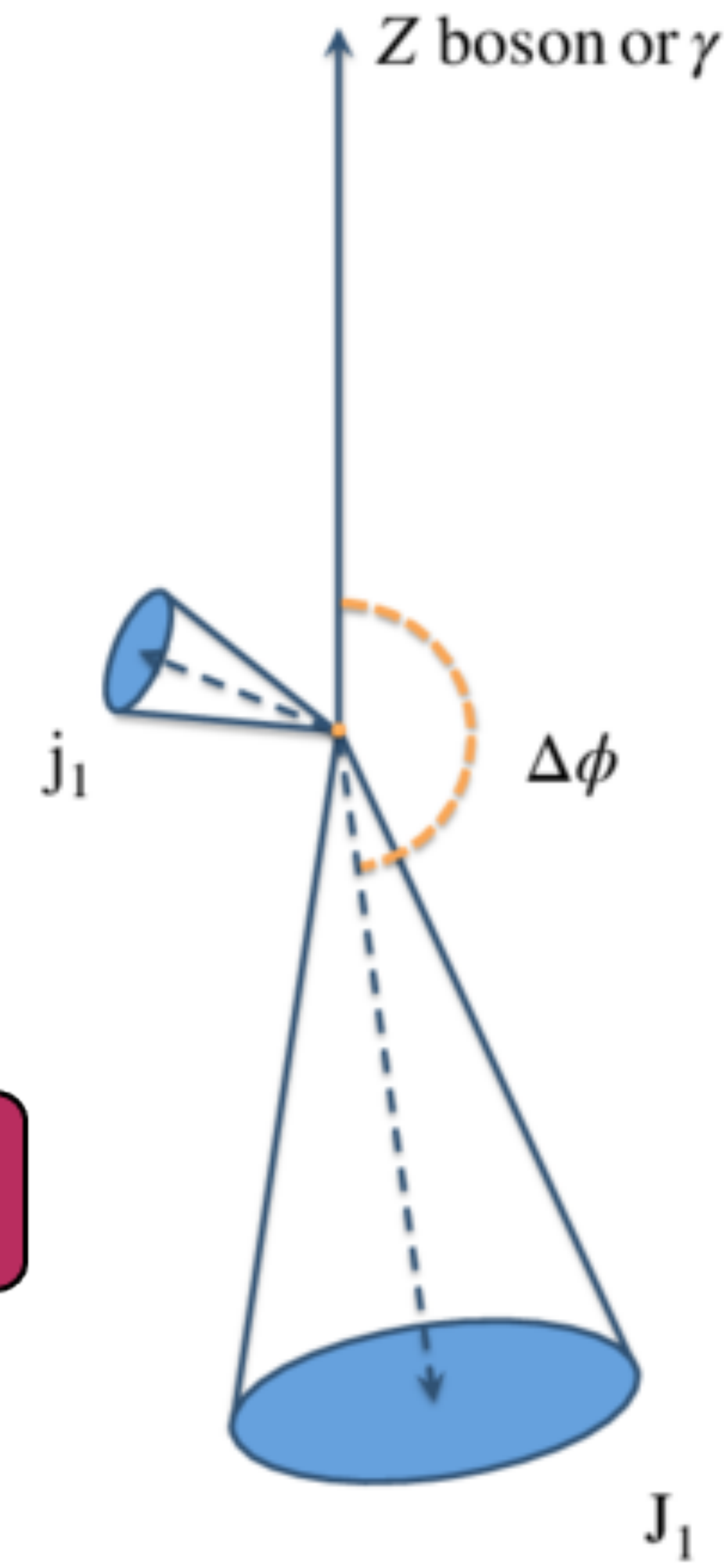


- 23% reduction in resolution on the $H(b\bar{b})$ peak!
  - More modest gain of 5% relative to μ+PtReco corrections
  - Can be improved via optimization (Z/H samples in training, calocells)

From AI/ML to physics

- So far the model only been tested on simulation
  - **Needs calibration** on data for ultimate use for physics!
  - Opportunity to integrate with in-situ, **training on data**
  - Exploring synergies with tagging - **all in one model?**



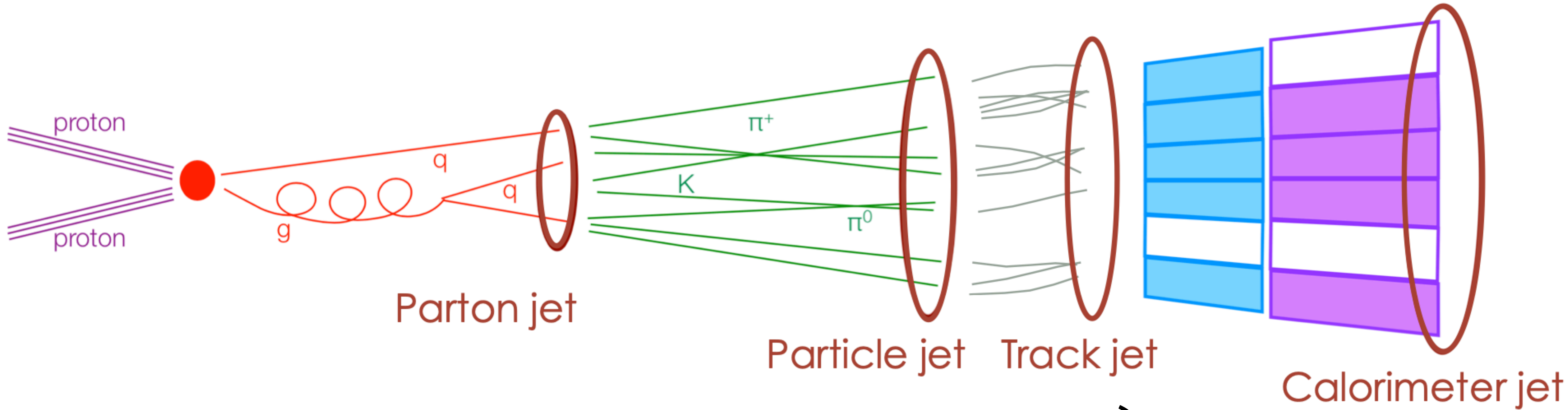


Thank you

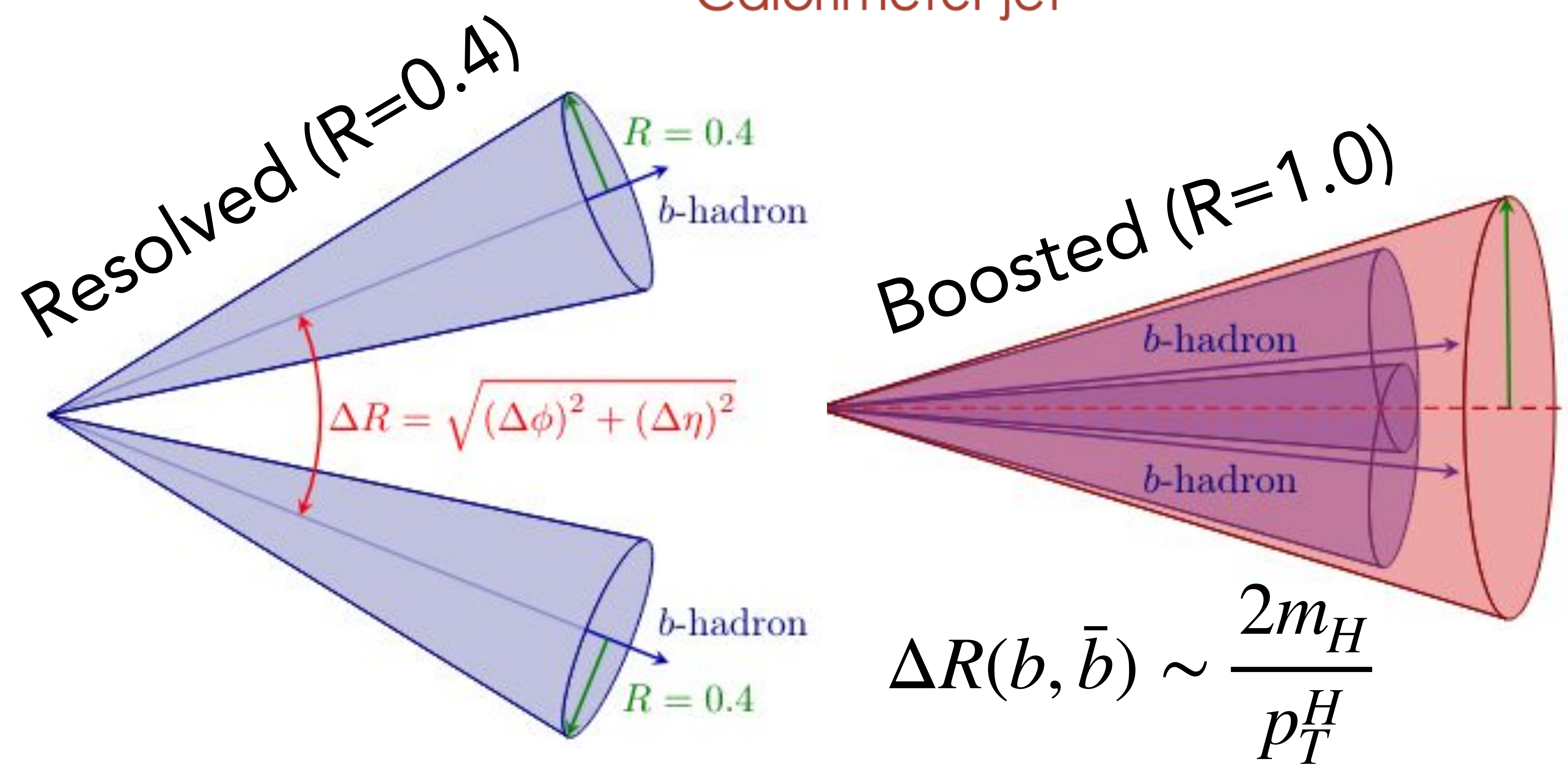
Backup

Jet reconstruction

- Jets are the most complex objects produced at colliders
  - Needs careful combination of signatures in tracker and calorimeters



- Cluster constituents using anti-k_T algorithm
  - Use different radius parameters depending on the target phase space (e.g. low/high-p_T Higgs)



$$\Delta R(b,\bar{b}) \sim \frac{2m_H}{p_T^H}$$

Jet constituents

Eur. Phys. J. C 77 (2017) 466

Eur. Phys. J. C 81 (2021) 689

ATL-PHYS-PUB-2017-15

Jet constituents

- Originally ATLAS only used calorimeter cells for jets

Eur. Phys. J. C 77 (2017) 466

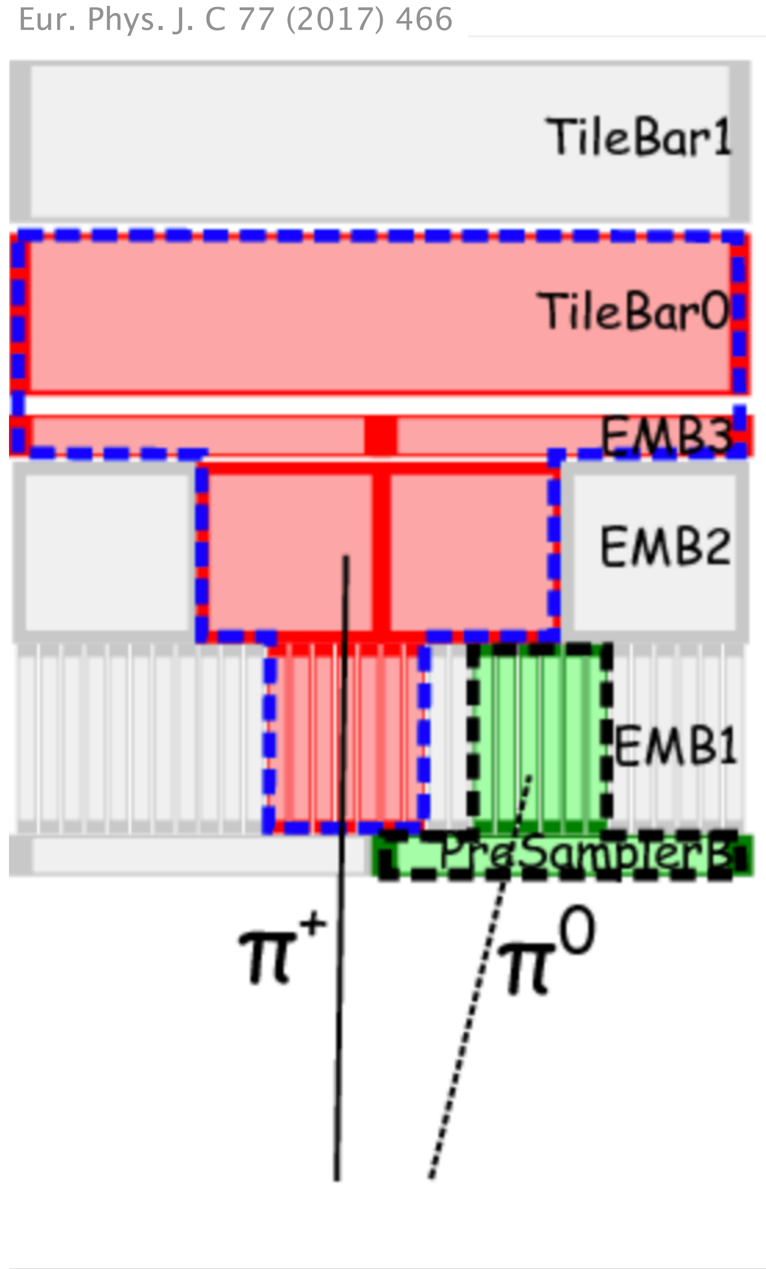
Eur. Phys. J. C 81 (2021) 689

ATL-PHYS-PUB-2017-15

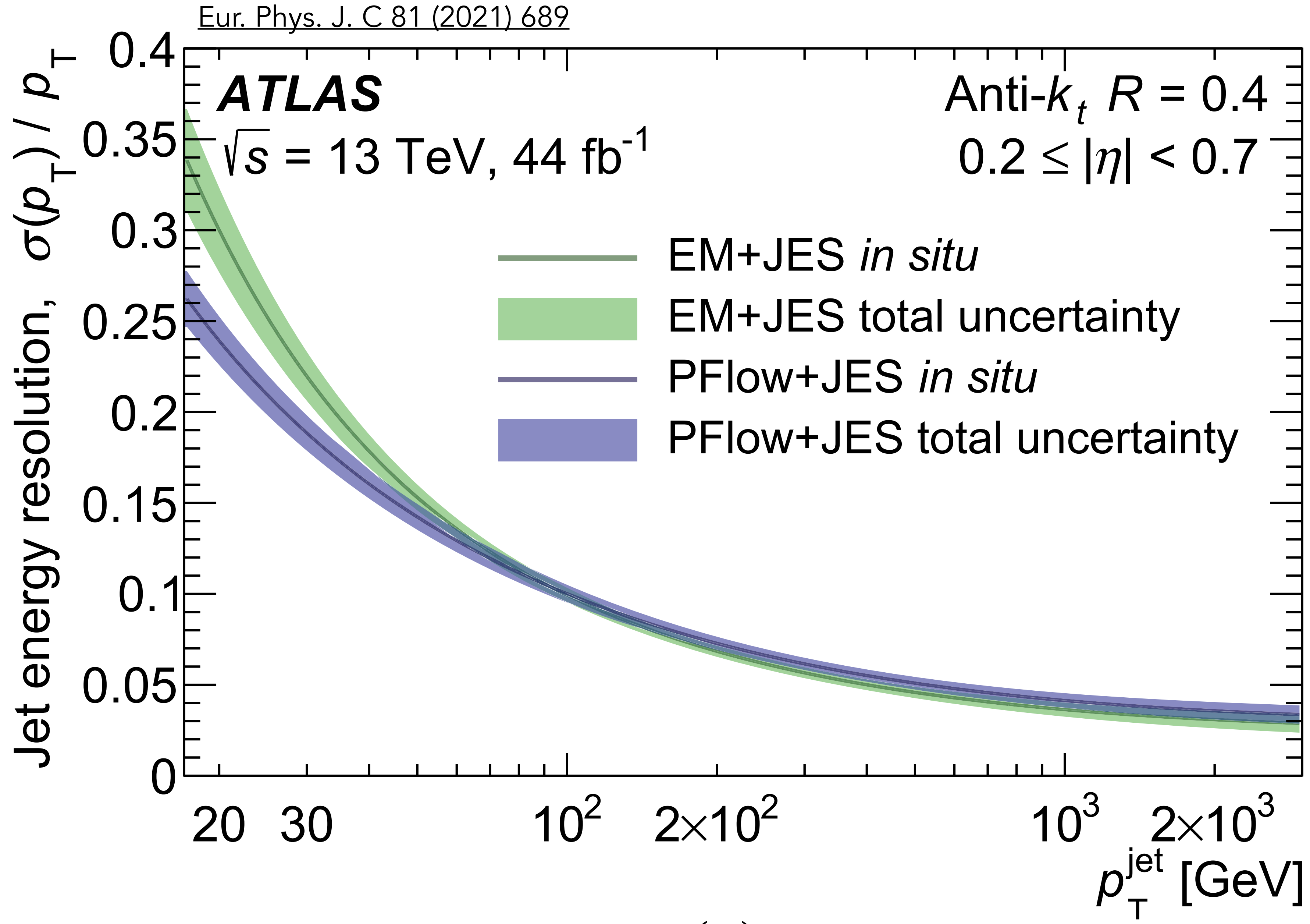
Jet constituents

- Originally ATLAS only used calorimeter cells for jets
- Now we are using **particle flow** objects: combine tracks and calo-clusters
  - Avoid double-counting energy/momentum, boosts performance at low-p_T (used for Small-R)

Eur. Phys. J. C 77 (2017) 466



Eur. Phys. J. C 81 (2021) 689

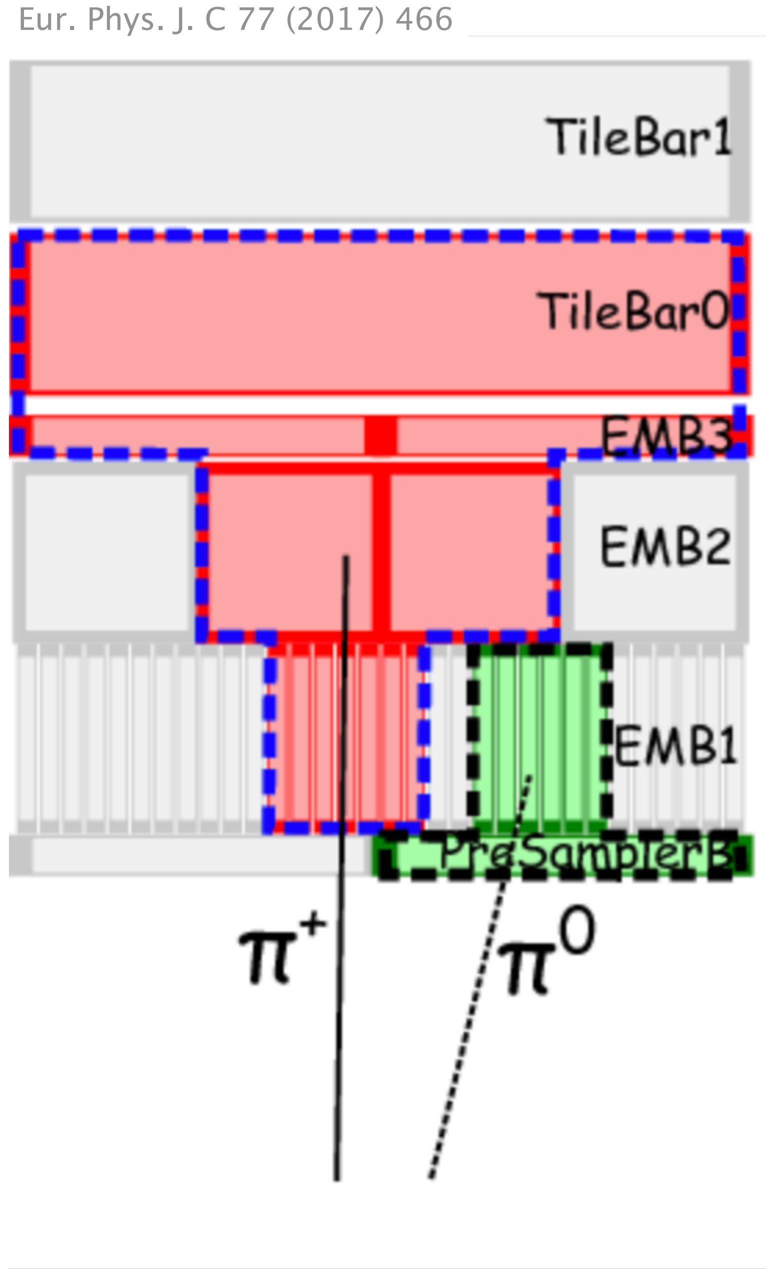


ATL-PHYS-PUB-2017-15

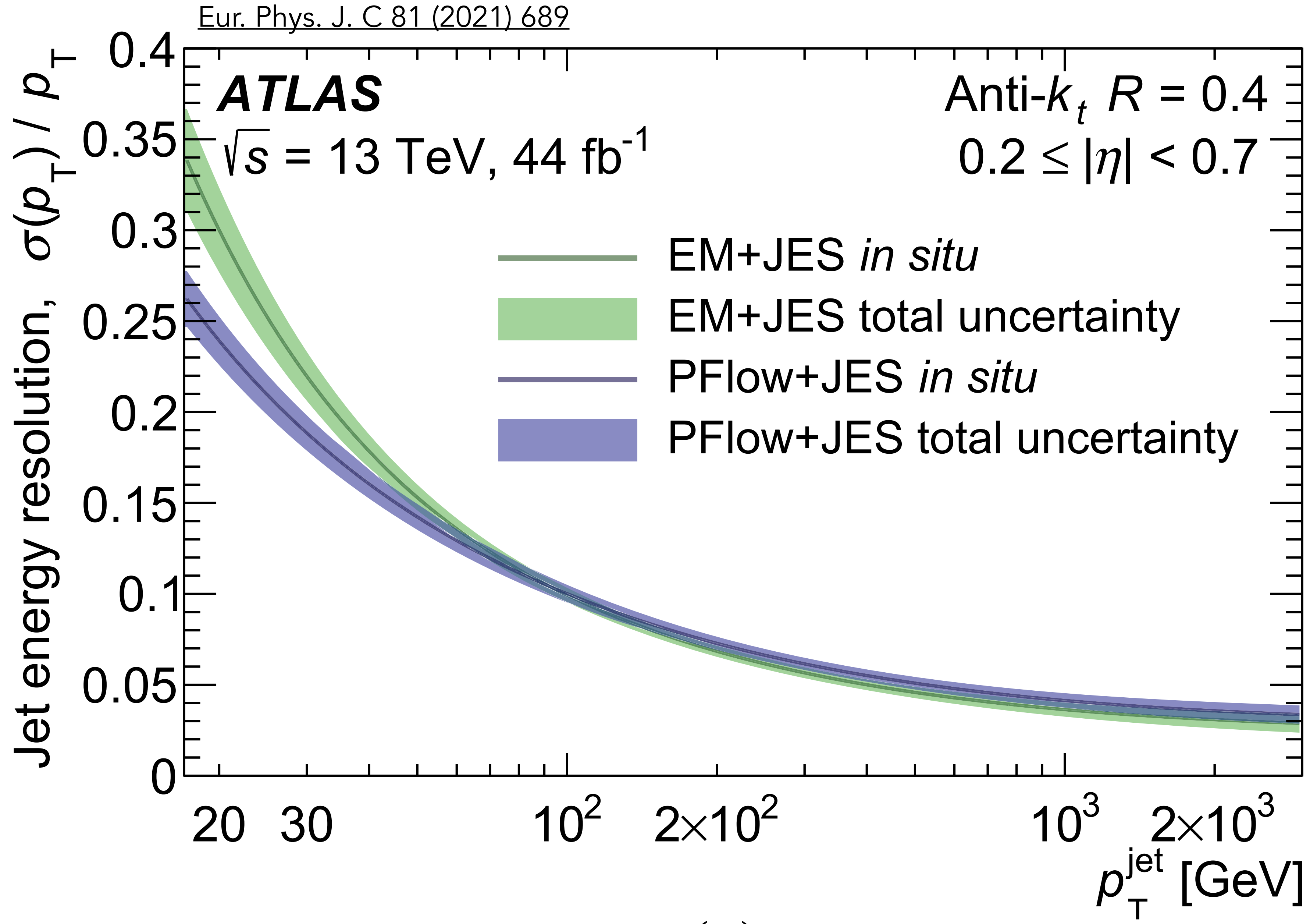
Jet constituents

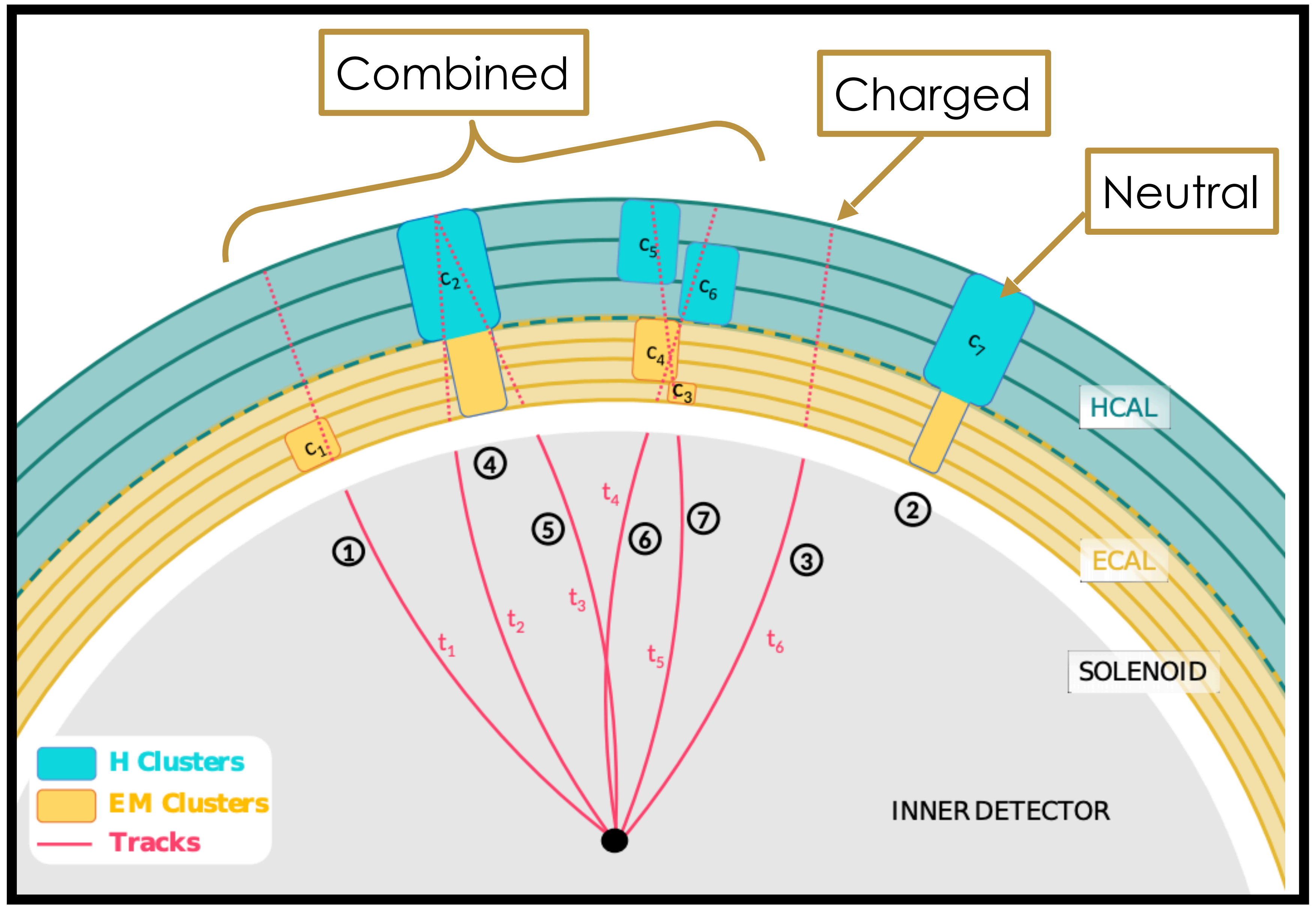
- Originally ATLAS only used calorimeter cells for jets
- Now we are using **particle flow** objects: combine tracks and calo-clusters
  - Avoid double-counting energy/momentum, boosts performance at low-p_T (used for Small-R)
- Recently developed **unified flow** objects, leverage angular resolution of tracker and energy resolution of calorimeter at high-p_T (used for Large-R jets)

Eur. Phys. J. C 77 (2017) 466



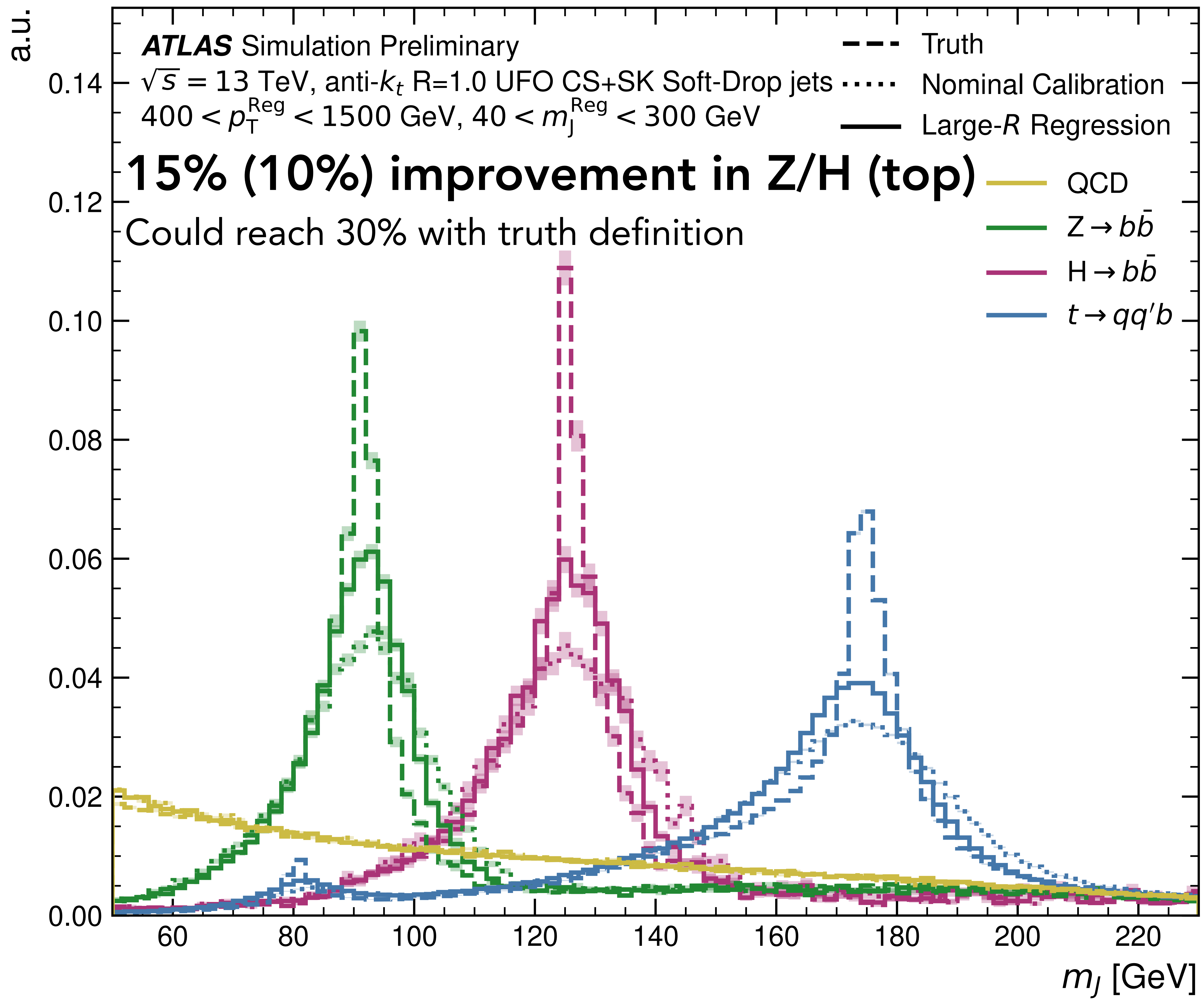
Eur. Phys. J. C 81 (2021) 689





ATL-PHYS-PUB-2017-15

Large-R resonance spectrum



- ✦ Evaluate on SM resonances (Z/H/top)
  - Significant sharpening of Z/H mass peaks
  - Still a long way to go to reach truth-level
- ✦ No mass sculpting in the QCD continuum
  - Use of flat-mass samples eliminates SM mass point bias

Neural network architecture

- Based on ATLAS flavor-tagging architecture
  - anti-k_T R=0.4 PFlow (small-R) jets use **track** constituents
  - anti-k_T R=1.0 UFO (large-R) jets use **track** and **flow object** constituents