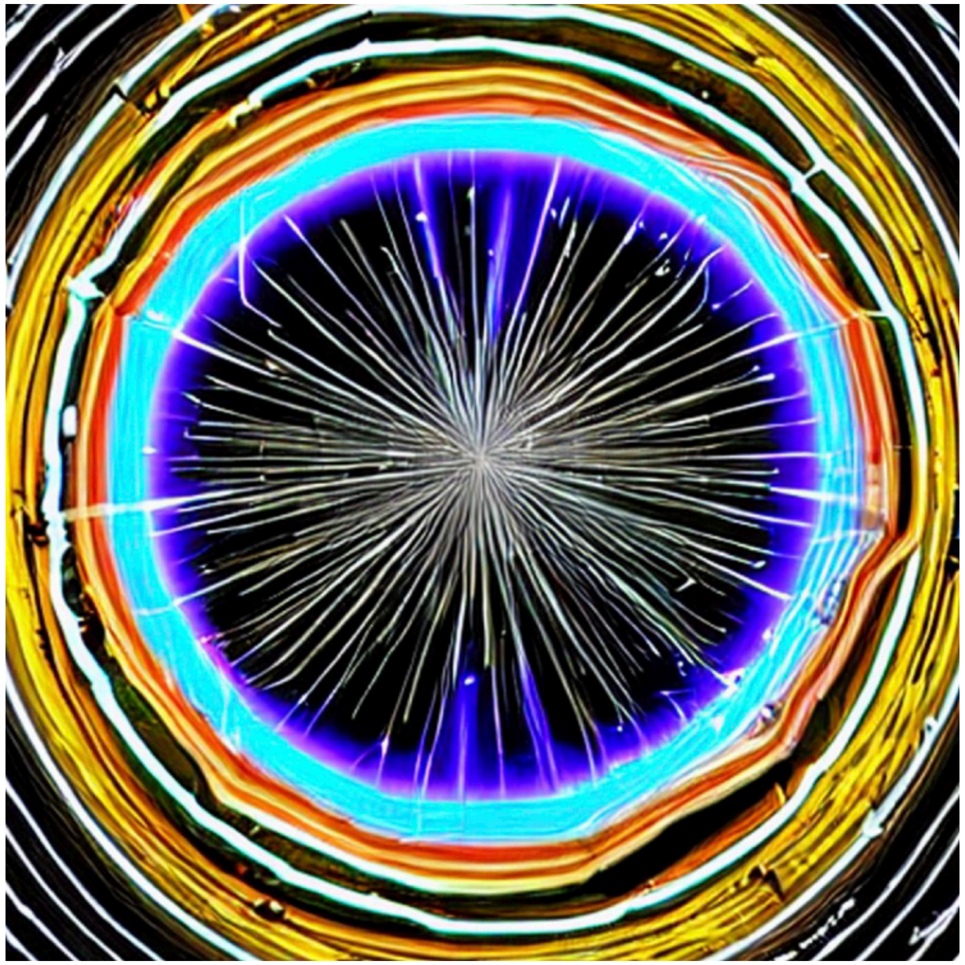
New Ideas in Jet Clustering

Roman Kogler
DESY

JetMET Workshop Brussels
May 16, 2023

Overview

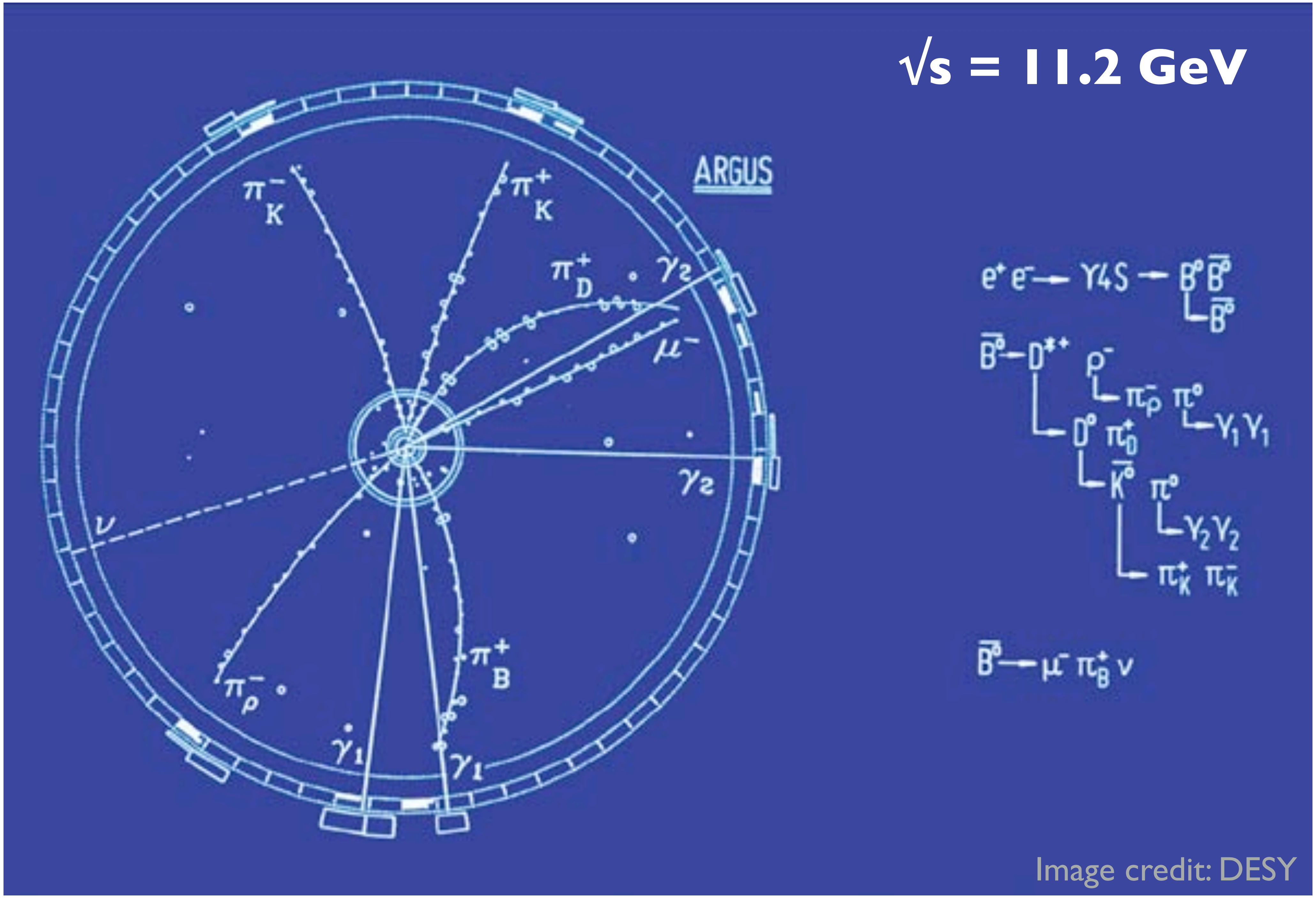
- Introduction
- Quark and Gluon Jets
- Exclusive Clustering
- Variable R Jets
- Scale Invariant Jet Clustering



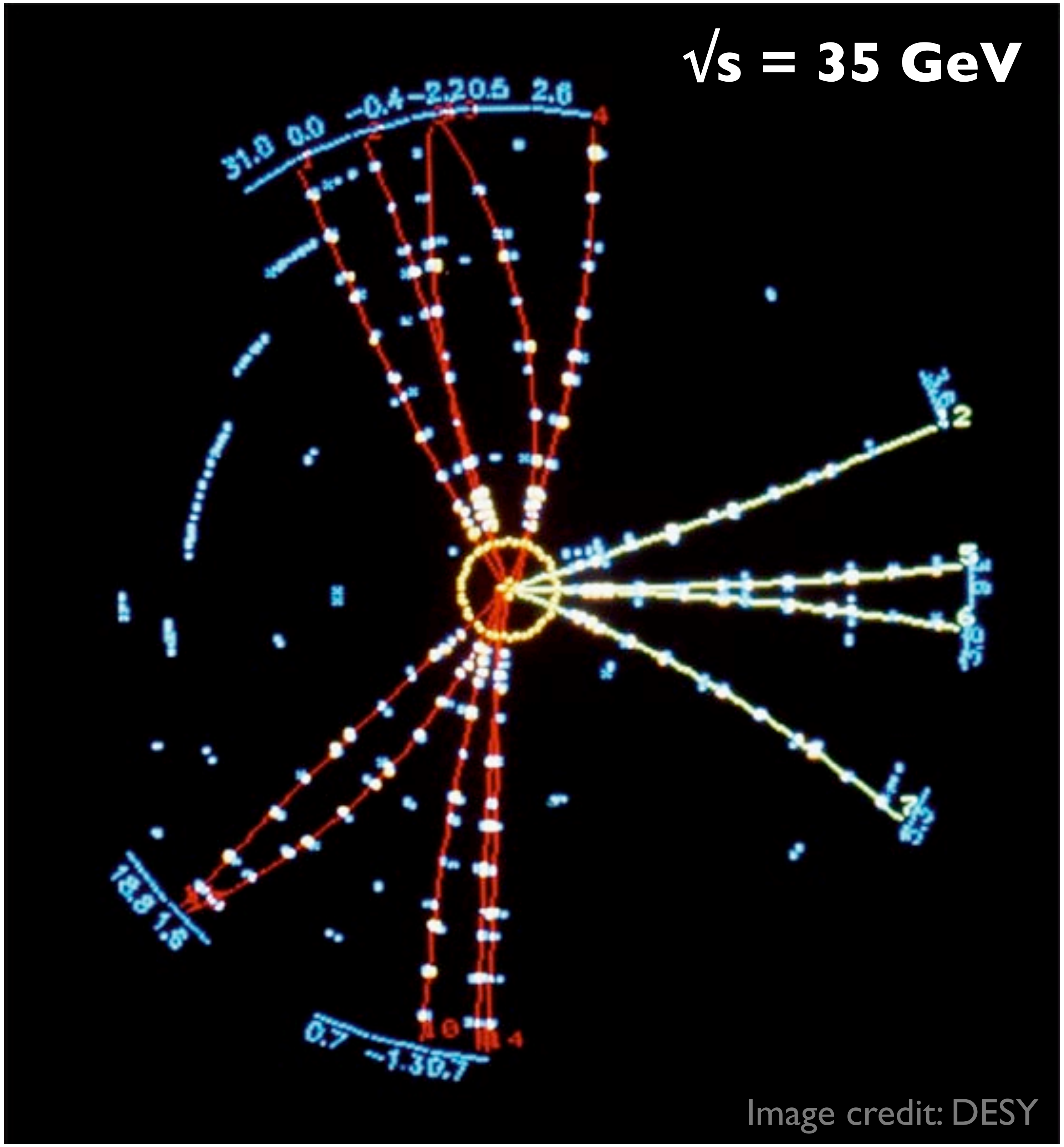
picsart.com

I will not cover machine learning techniques.

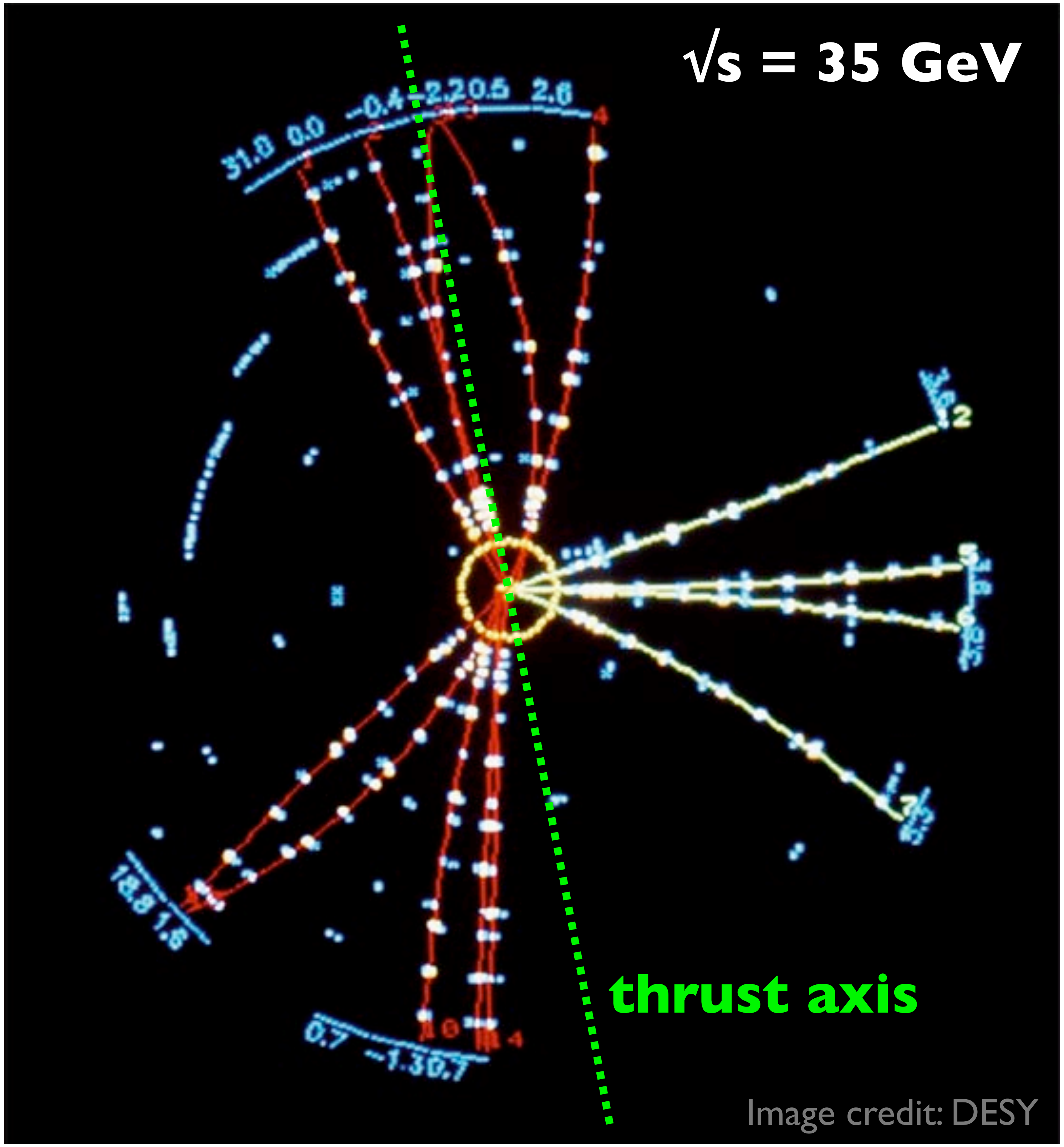
ARGUS at DORIS, 1987



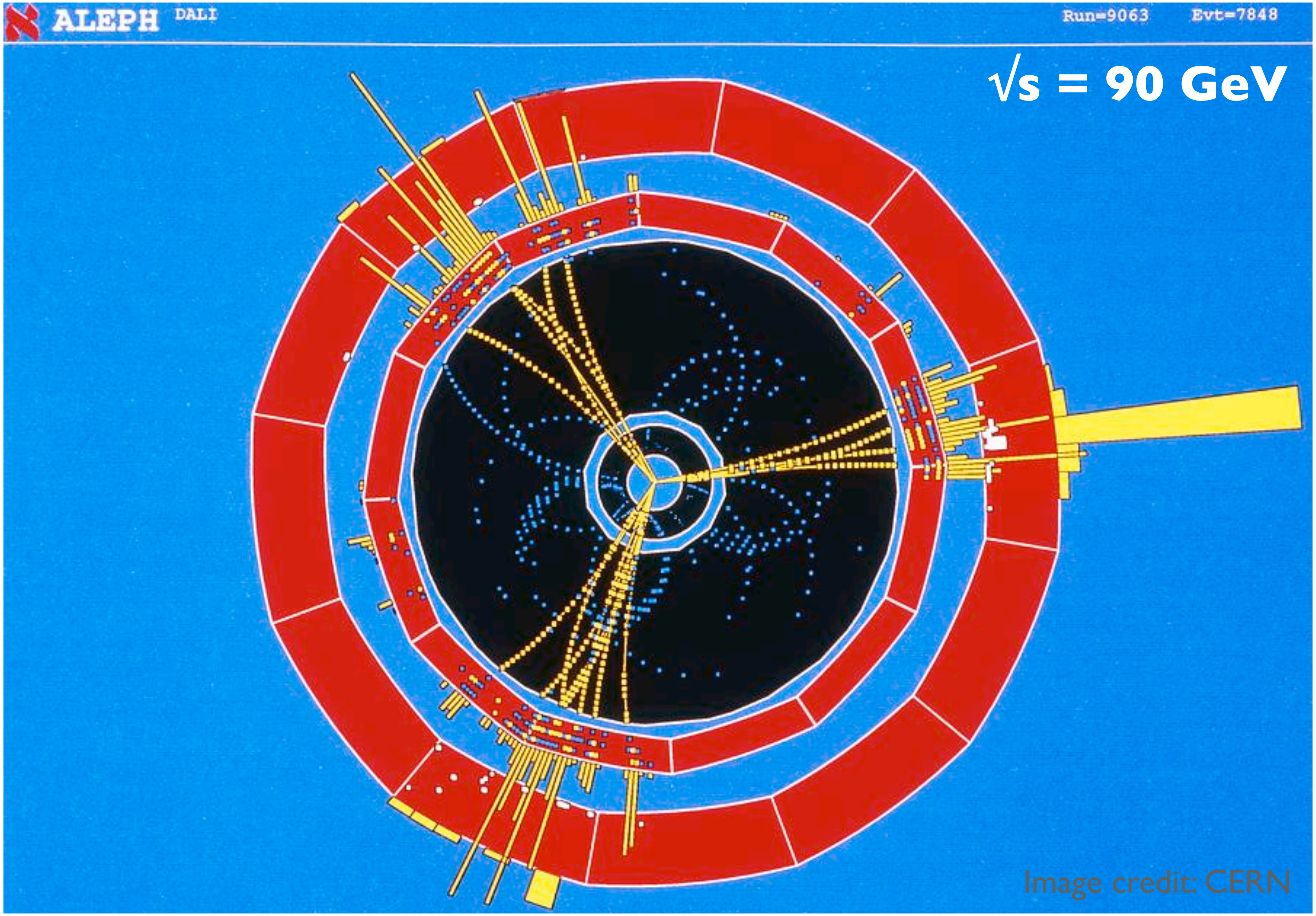
TASSO at PETRA, 1979



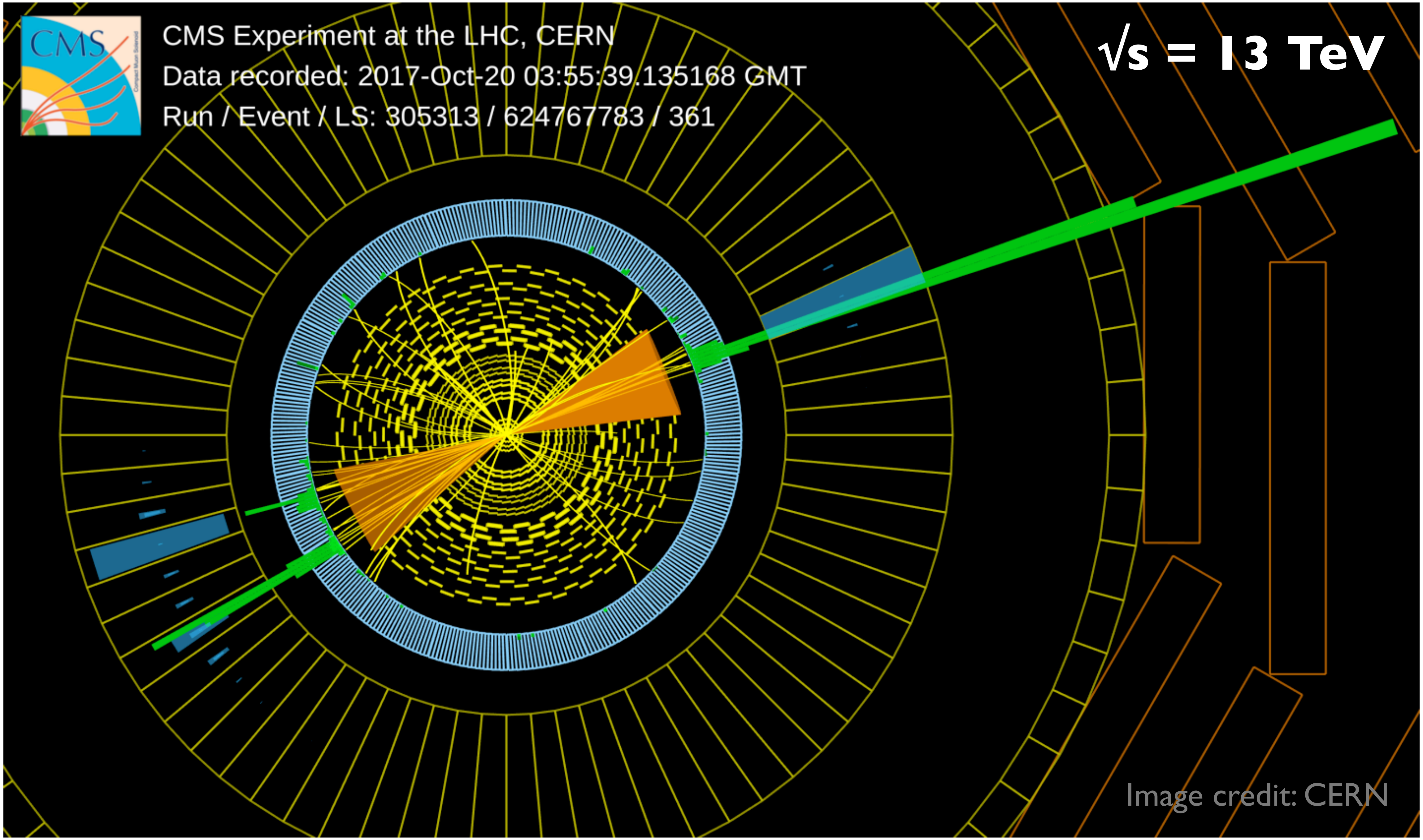
TASSO at PETRA, 1979



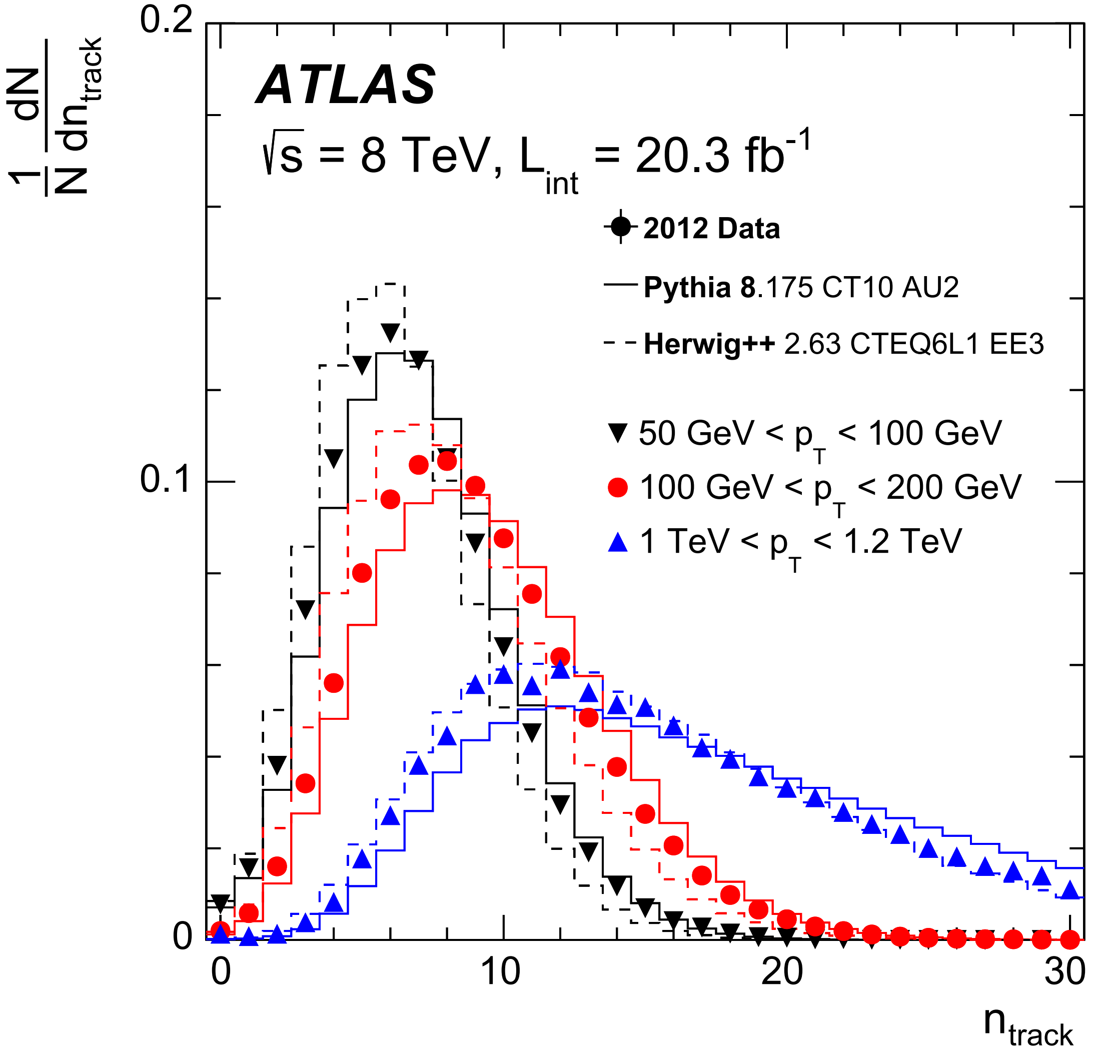
ALEPH at LEP, 1992



CMS at LHC, 2017



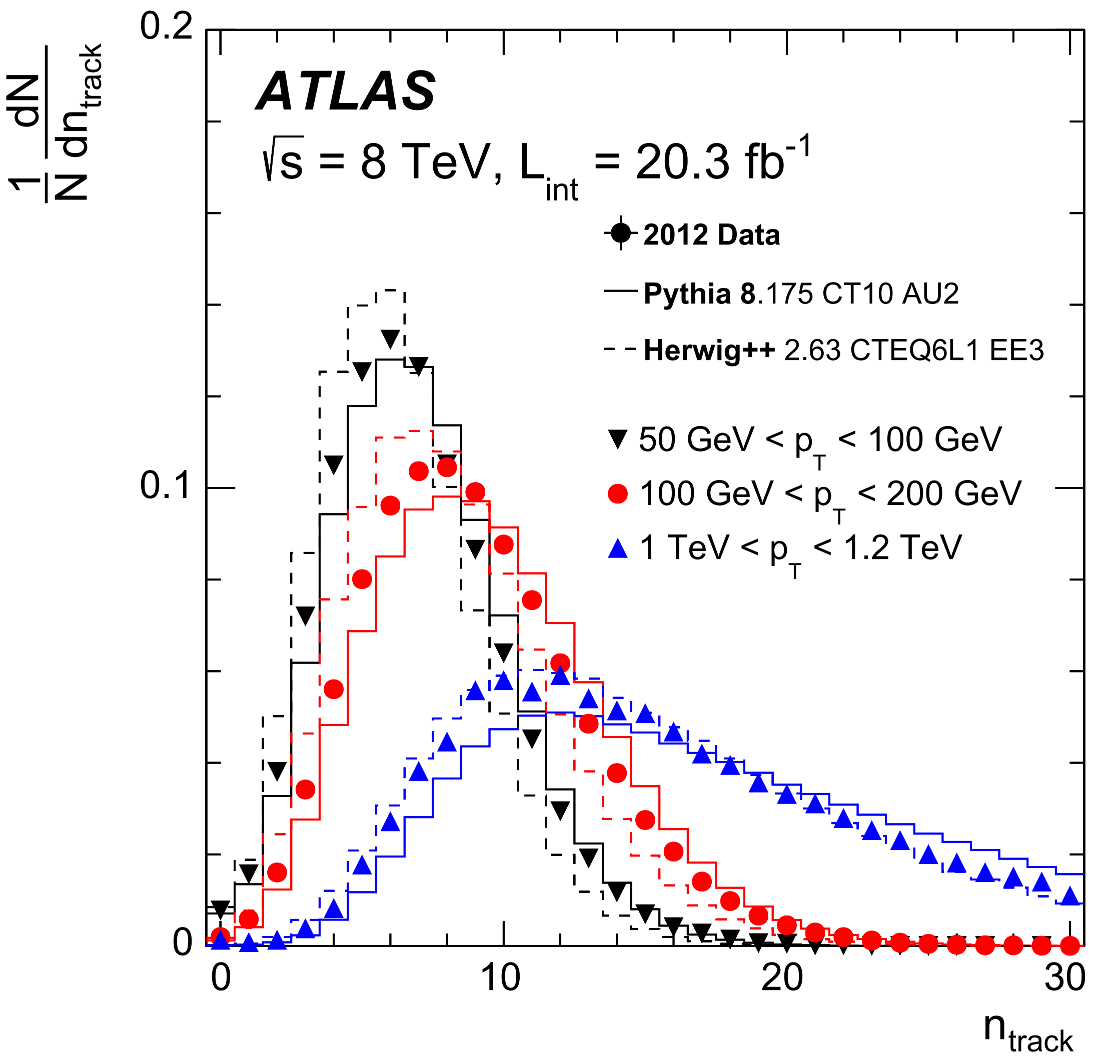
Charged Particles in Jets

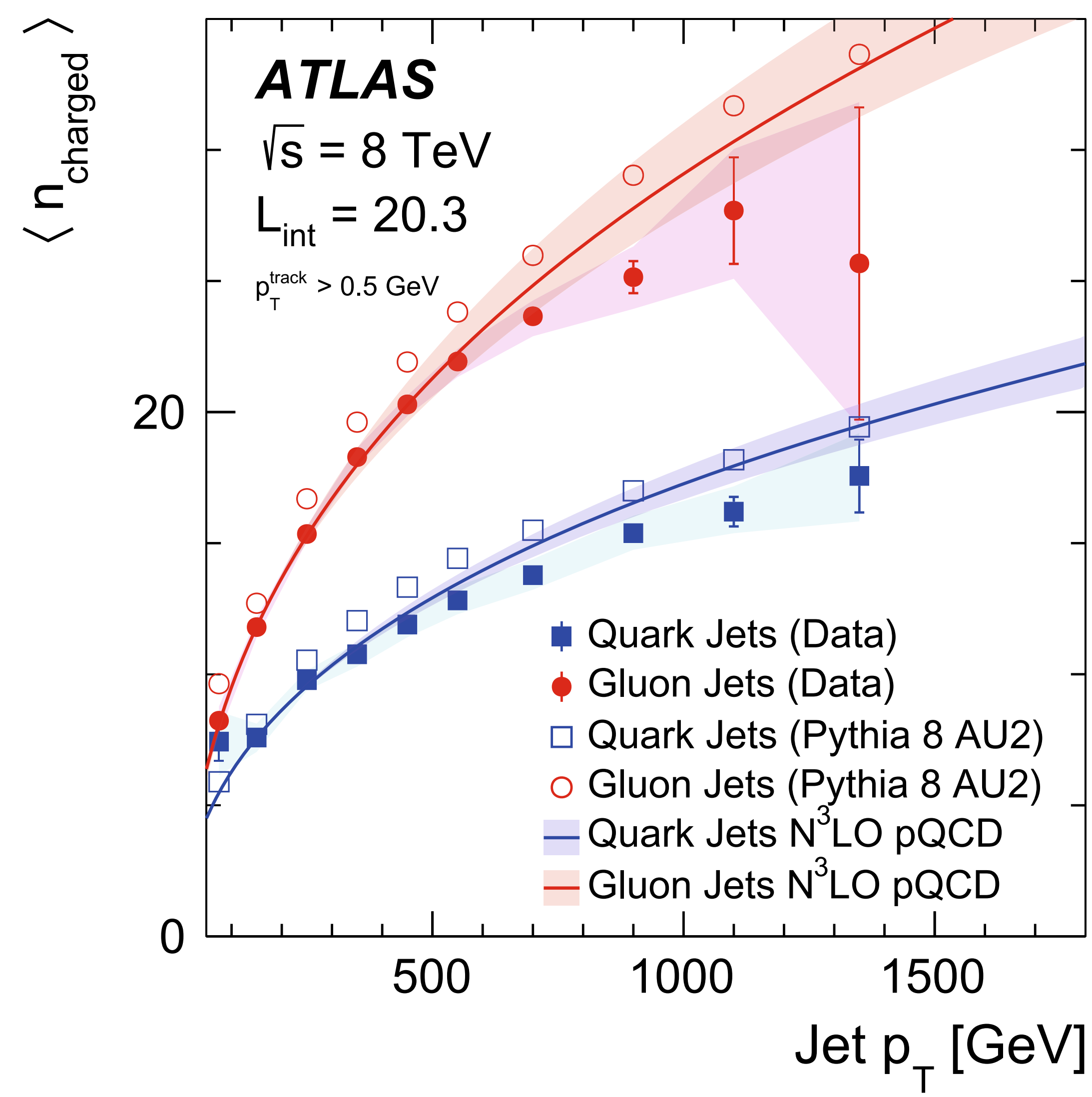


- Approximate particle content in a jet: $\pi^+ : \pi^- : \pi^0 = 1 : 1 : 1$ (+10% Kaons, Protons…)

[ATLAS, EPJC 76, 322 (2016)]

Charged Particles in Jets





- Approximate particle content in a jet: $\pi^+ : \pi^- : \pi^0 = 1 : 1 : 1$ (+10% Kaons, Protons…)
- Gluon jets have higher multiplicity (colour factor C_A compared to C_F)

[ATLAS, EPJC 76, 322 (2016)]

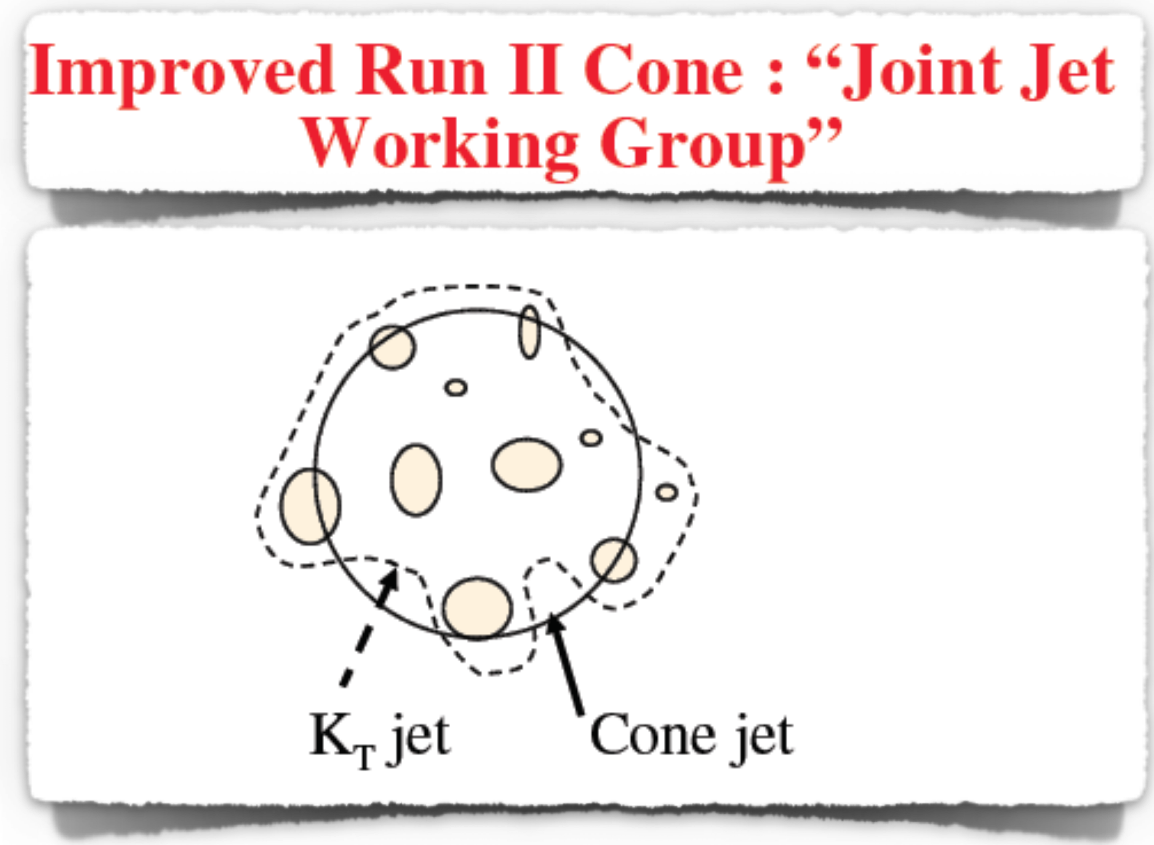
Jet Algorithms

The whole world is using anti-k_T jets, right?

Jet Algorithms

The whole world is using anti-k_T jets, right?

- Before 2007
  - Durham/Jade (LEP)
  - Cambridge/Aachen (LEP)
  - k_T (HERA)
  - midpoint-cone (Tevatron)
- In 2008, the LHC was ready to ramp up to 14 TeV
  - ATLAS and CMS were eager to start with data taking using cone-type jet algorithms



[Mikołaj Ćwiok, Moriond QCD, 2007]

Jet Algorithms

- Catastrophic incident in Sep 2008
  - Magnet quench resulted in explosive Helium release
  - Repairs delayed the start by 14 months

[CDS, CERN 2008, 1185822]

Jet Algorithms

- Catastrophic incident in Sep 2008
  - Magnet quench resulted in explosive Helium release
  - Repairs delayed the start by 14 months
- At the same time, important ideas / breakthroughs

[CDS, CERN 2008, 1185822]

Dispelling the N^3 myth for the k_t jet-finder

Matteo Cacciari, Gavin P. Salam *

LPTHE, Universities of Paris VI and VII and CNRS, Paris, France

Received 10 July 2006; received in revised form 11 August 2006; accepted 11 August 2006

Available online 28 August 2006

Editor: N. Glover

The anti-k_t jet clustering algorithm

Matteo Cacciari and Gavin P. Salam

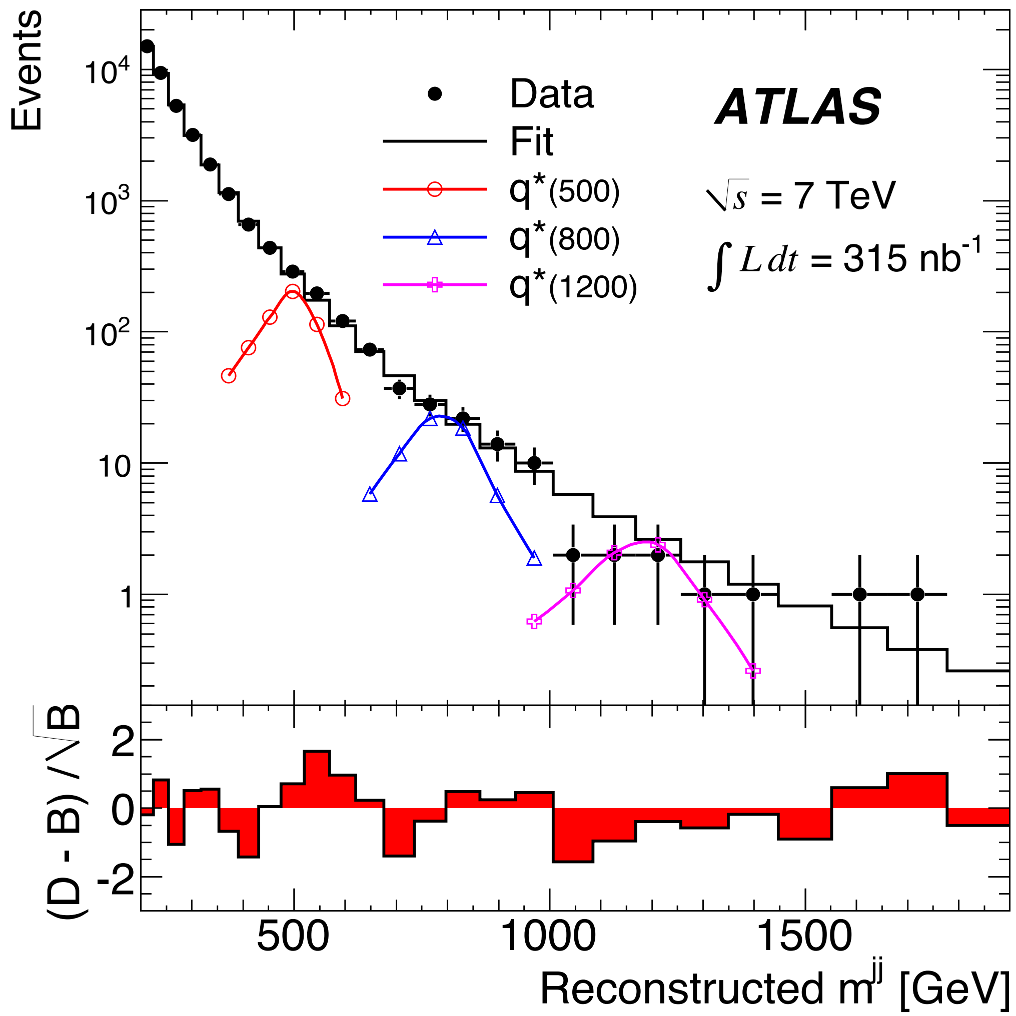
LPTHE, UPMC Université Paris 6,
Université Paris Diderot — Paris 7,
CNRS UMR 7589, Paris, France
E-mail: cacciari@lpthe.jussieu.fr, salam@lpthe.jussieu.fr

Gregory Soyez

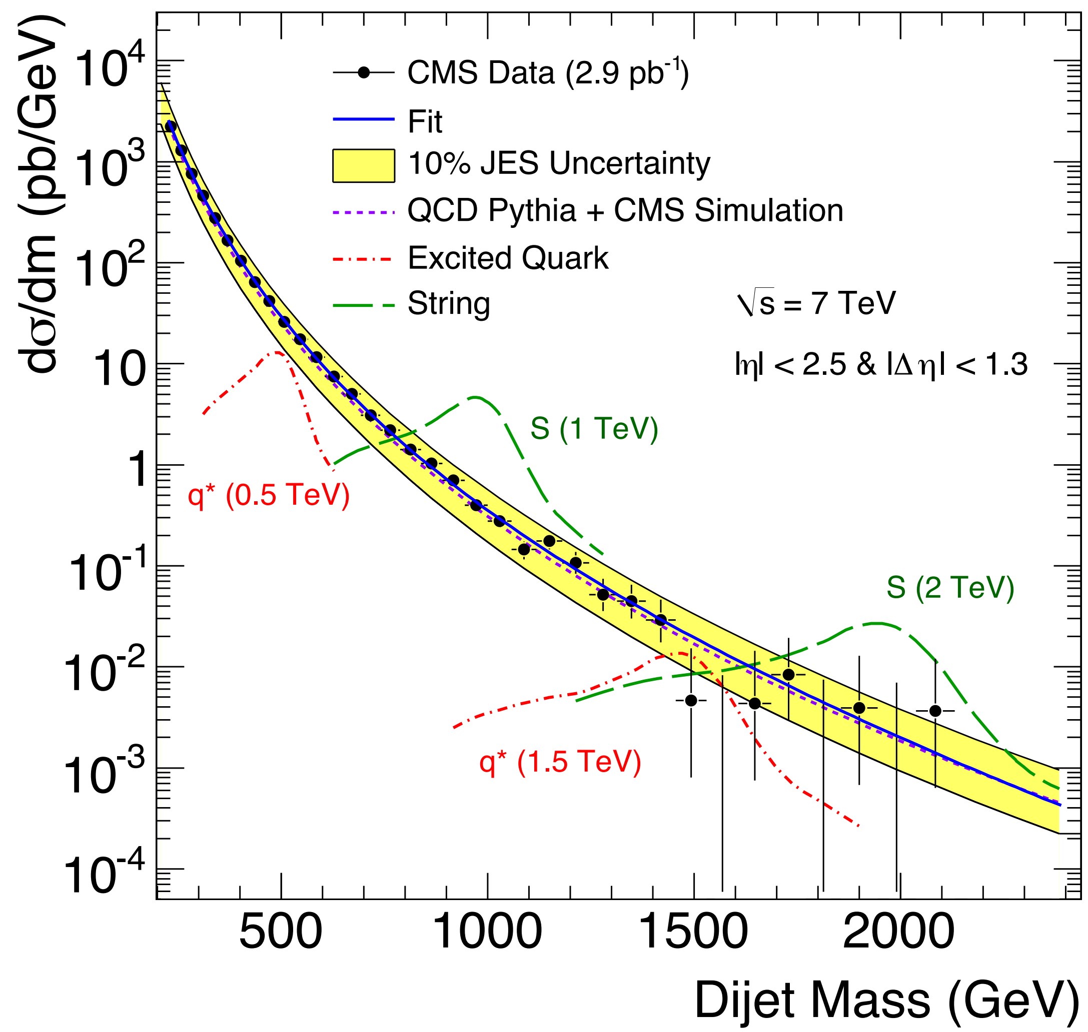
Brookhaven National Laboratory,
Upton, NY, U.S.A.
E-mail: gsoyez@quark.phy.bnl.gov

The Rest is History…

[ATLAS PRL 105, 161801]



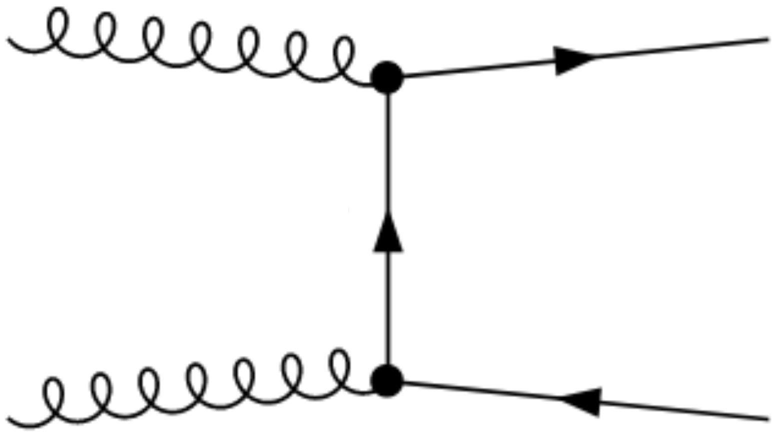
[CMS PRL 105, 211801]

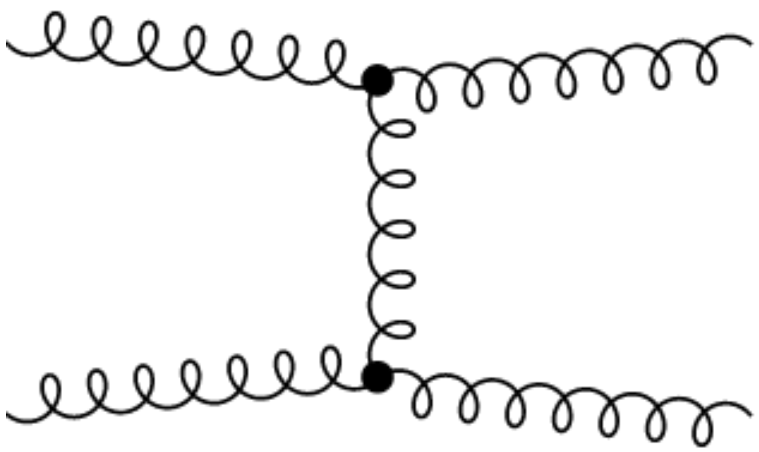


ATLAS dijet search using AK6 jets Sep. 2010

CMS dijet search using AK7 jets Oct. 2010

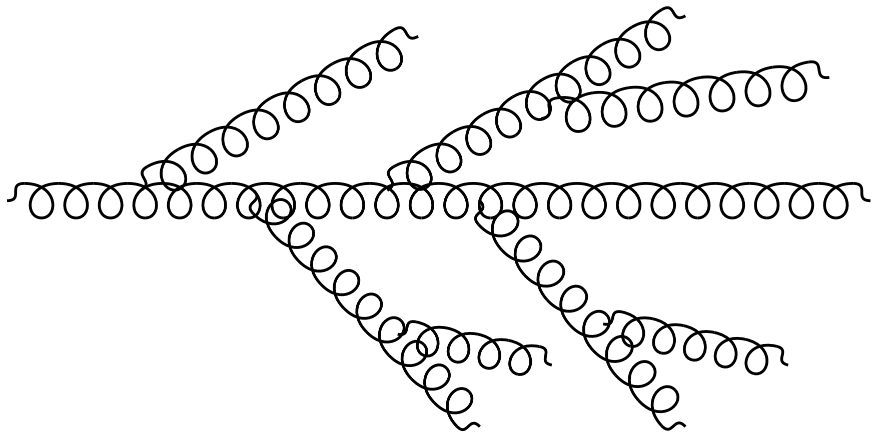
Quark and Gluon Jets





Corrections to Jet Observables

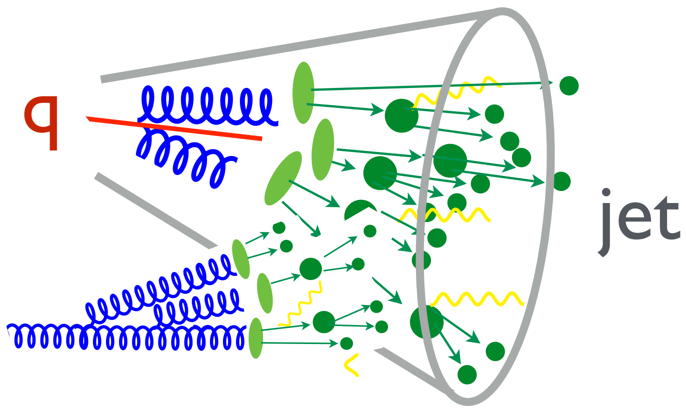
Perturbative effects



Lose E and p because of splittings

$$\frac{\langle \delta p_t \rangle_{\text{pert}}}{p_t} = \frac{\alpha_s}{\pi} L_i \ln R + \mathcal{O}(\alpha_s)$$

Non-perturbative effects

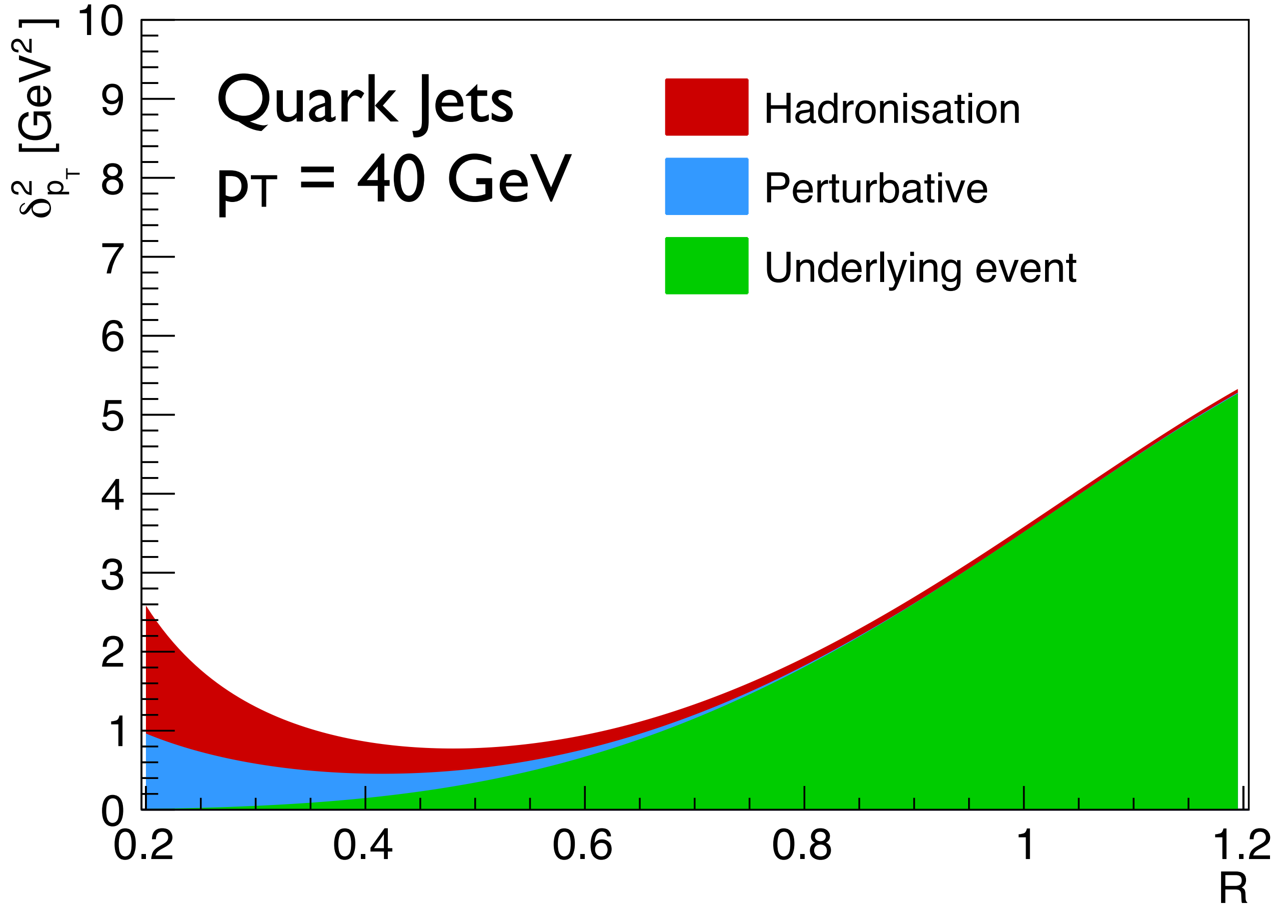


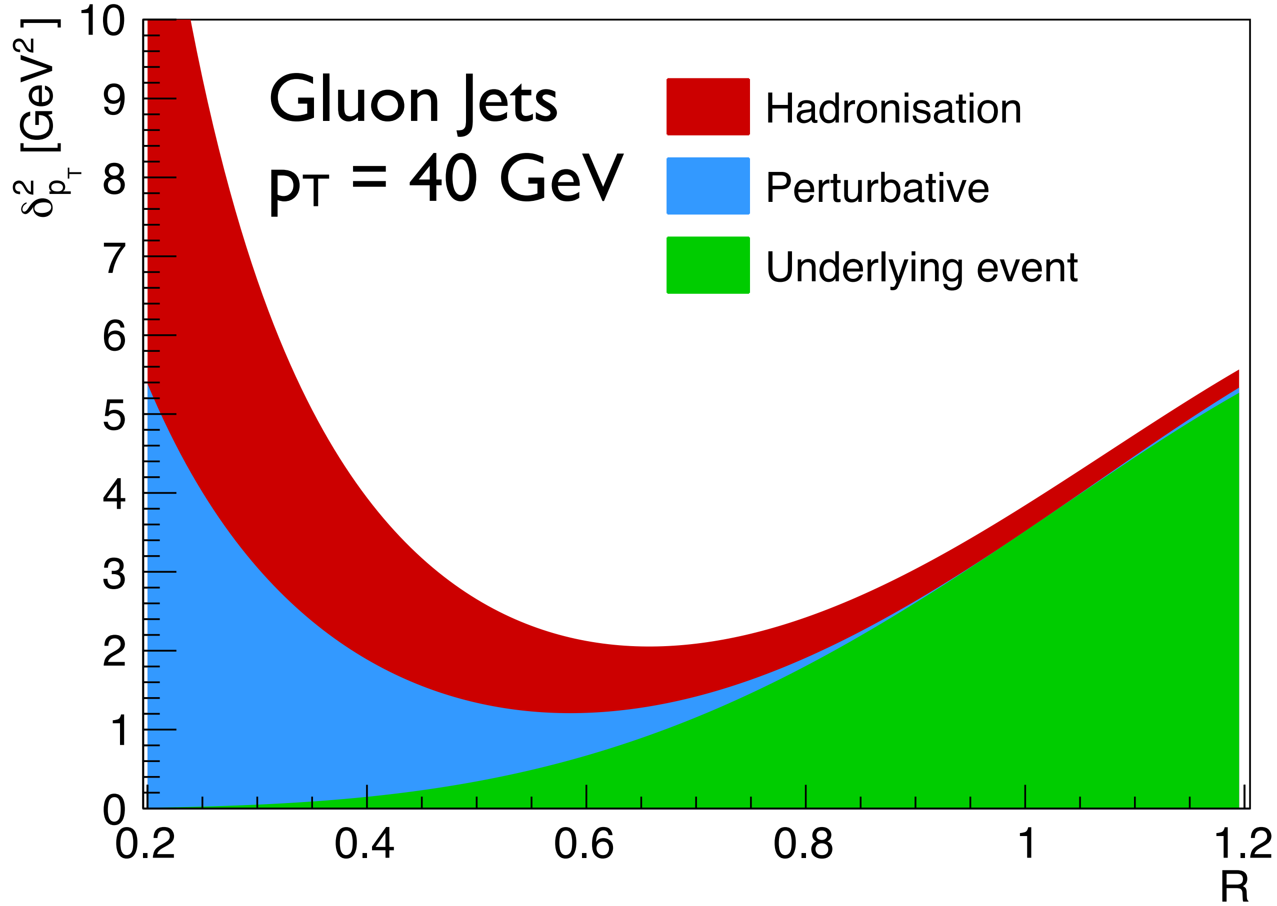
$$\langle \delta p_t \rangle_{\text{NP}} \sim -\frac{2C_F \Lambda}{\pi R}$$

$$\langle \delta p_t \rangle_{\text{UE}} \simeq \Lambda_{\text{UE}} R J_1(R) = \Lambda_{\text{UE}} \left(\frac{R^2}{2} - \frac{R^4}{8} + \dots \right)$$

[Salam, EPJC 67, 637 (2010)]

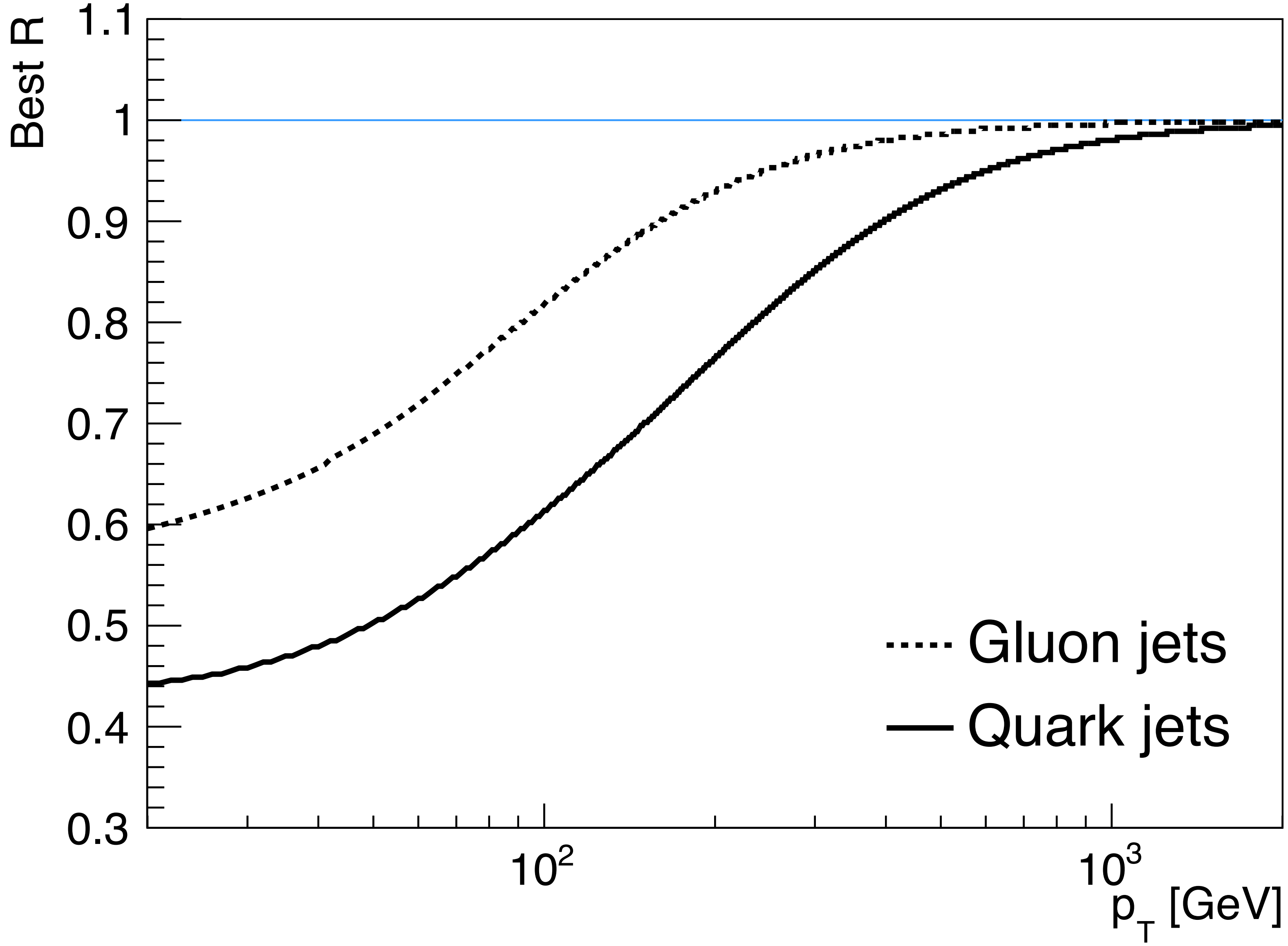
Uncertainties from Corrections



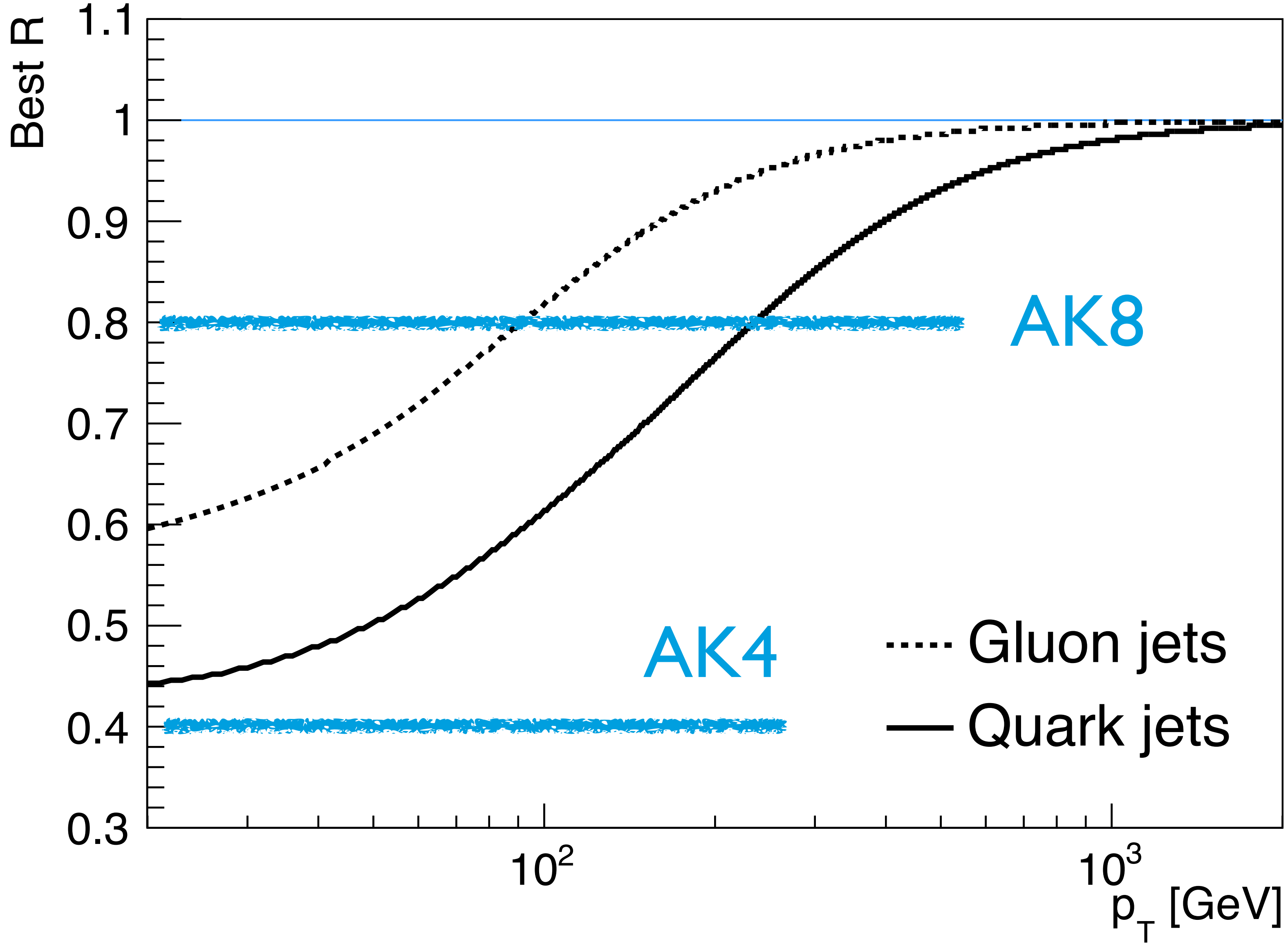


Smaller corrections for quark jets
Minimum at around R = 0.5 - 0.6

Is there an optimal R?



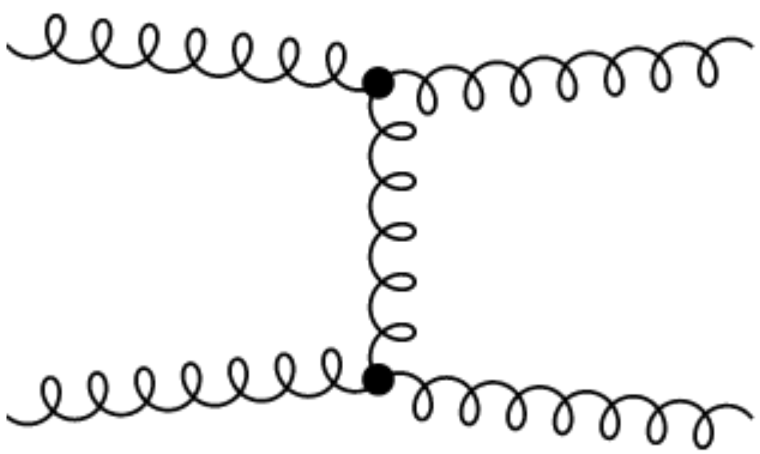
Is there an optimal R?



It depends…

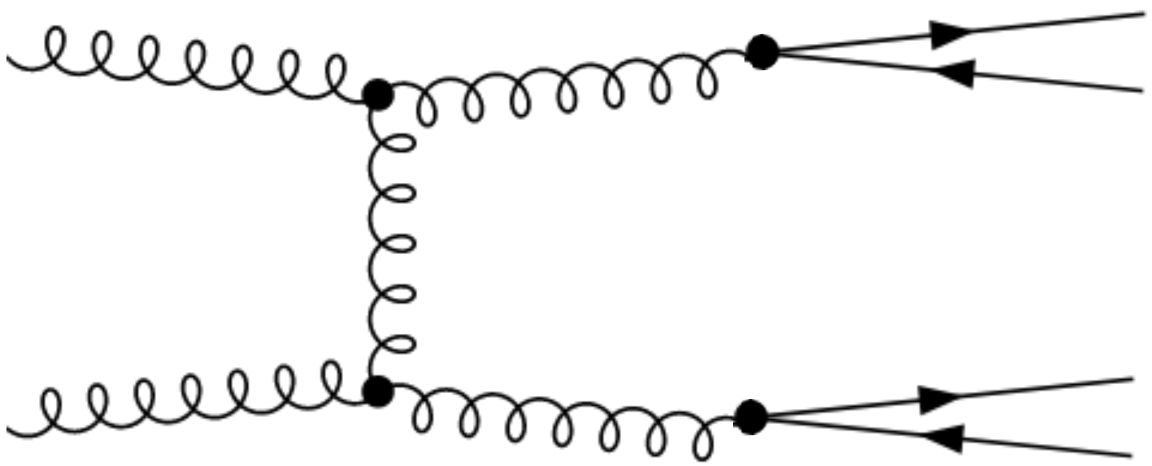
… on p_T and flavour

Defining Quark and Gluon Jets



Obviously two gluon jets…

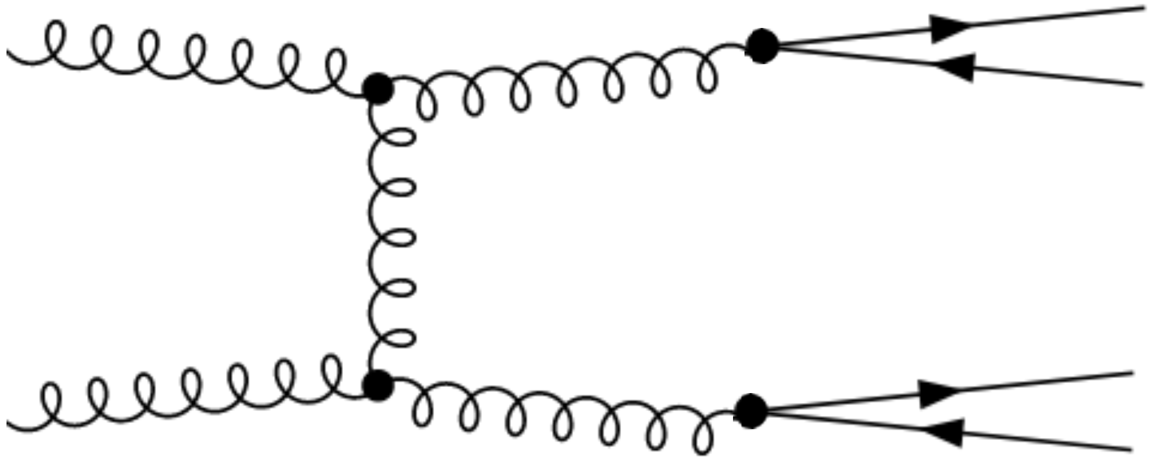
Defining Quark and Gluon Jets



Obviously two gluon jets…

… or not?!?

Defining Quark and Gluon Jets



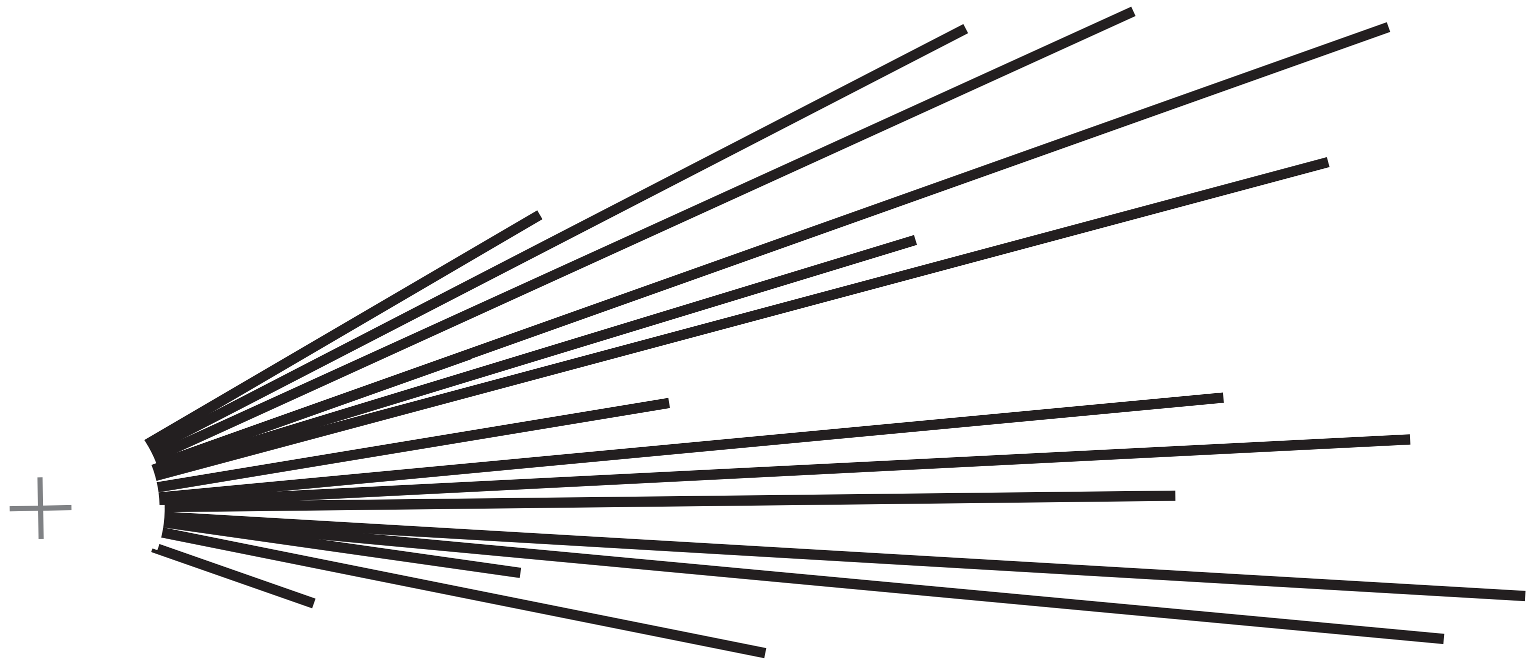
Obviously two gluon jets…

… or not?!?

- Parton flavour (from hard matrix element) is intrinsically flawed
- Physically meaningful definitions (not exhaustive)
  - N-Subjettiness [Larkoski, Metodiev, EPJC 10, 014 (2019)]
    - Possibility to unambiguously define quark jets ($\tau_N \rightarrow 0$)
    - Gluon jets always contaminated by quark jets, $(C_F/C_A)^{N_{emissions}}$
  - Flavour-k_T [Banfi, Salam, Zanderighi, EPJC 47, 113 (2006)]
  - Jet topics [Komiske, Metodiev, Thaler, JHEP 11 059 (2018)]
  - Fragmentation approach (WTA axis) [Caletti et al., JHEP 10 158 (2022)]

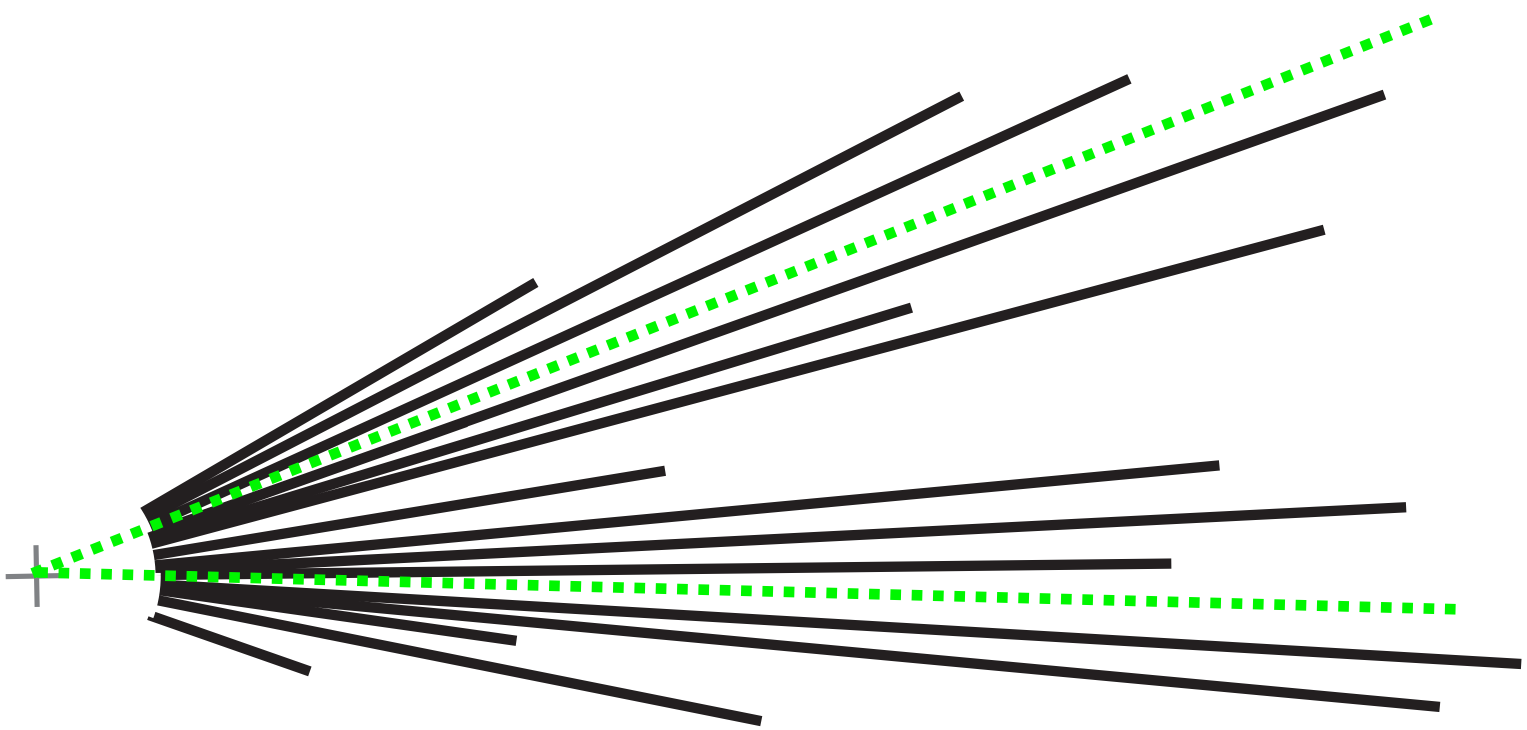
Jet substructure

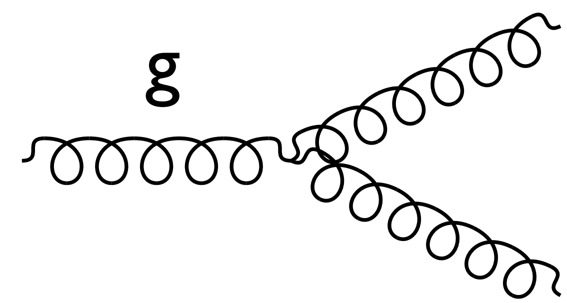
- Remove unwanted / soft radiation from jets
- Aid the jet reconstruction and calibration
- Distinguish quark/gluon jets
- Tagging of fully merged W, Z, H and top jets

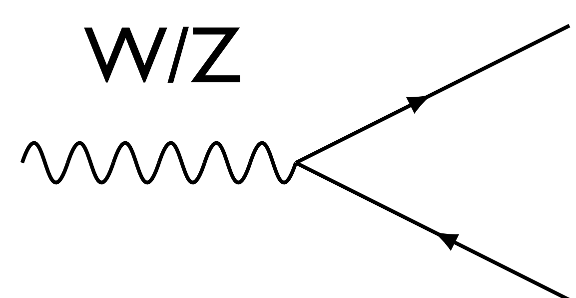


Jet substructure

- Remove unwanted / soft radiation from jets
- Aid the jet reconstruction and calibration
- Distinguish quark/gluon jets
- Tagging of fully merged W, Z, H and top jets



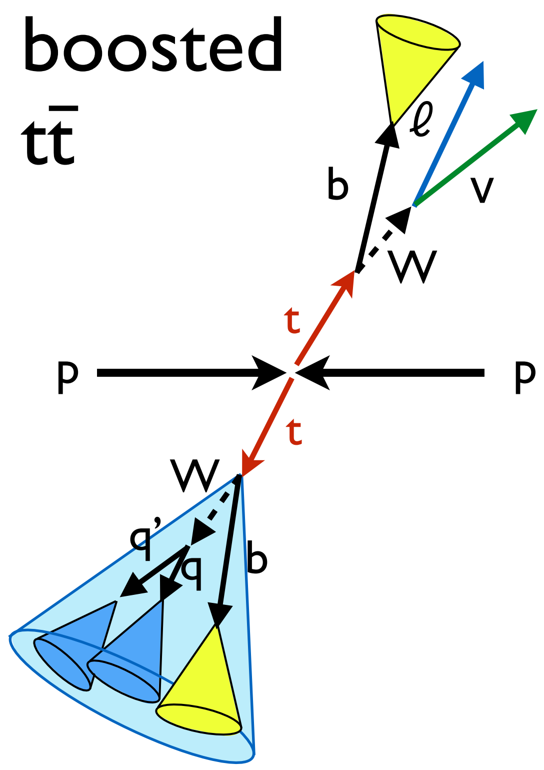


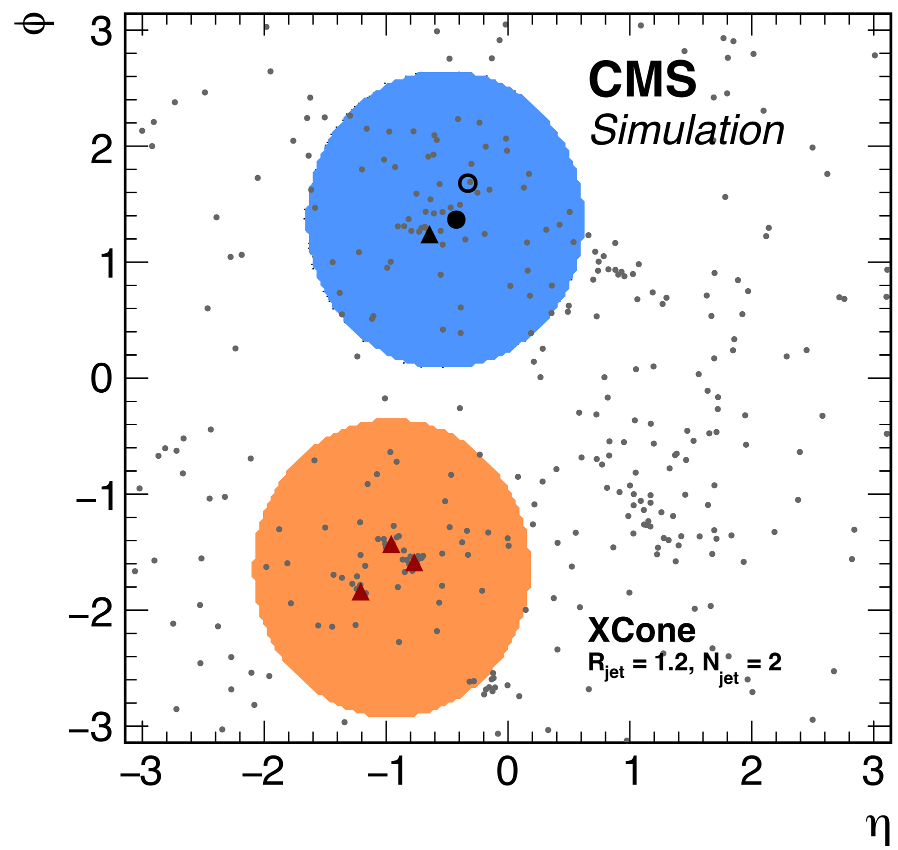


Exclusive Clustering

XCone

Use number of expected jets when event topology is known



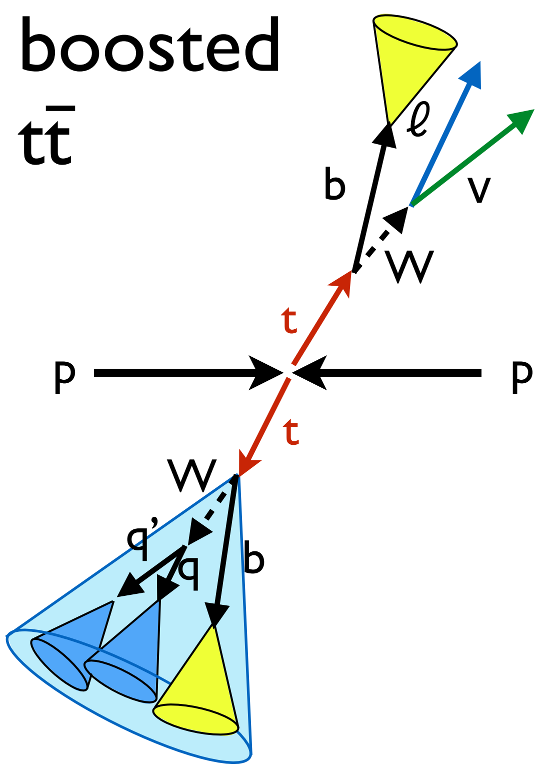


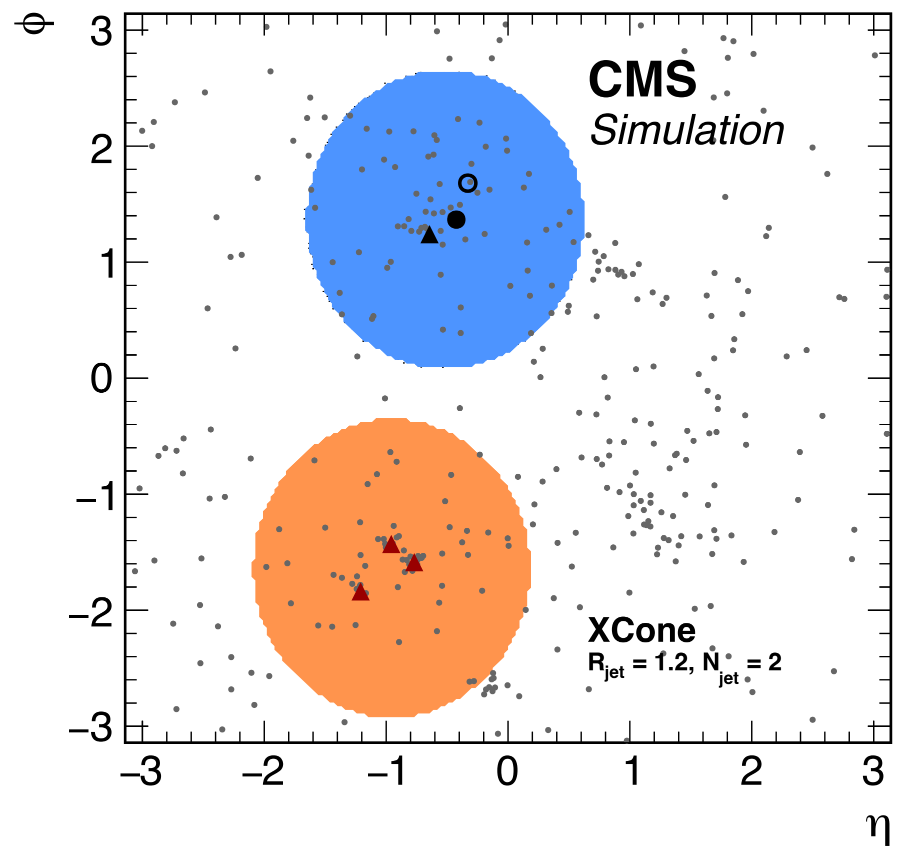
- XCone assigns particles based on N-jettiness axes
- Natural transition resolved ↔ boosted

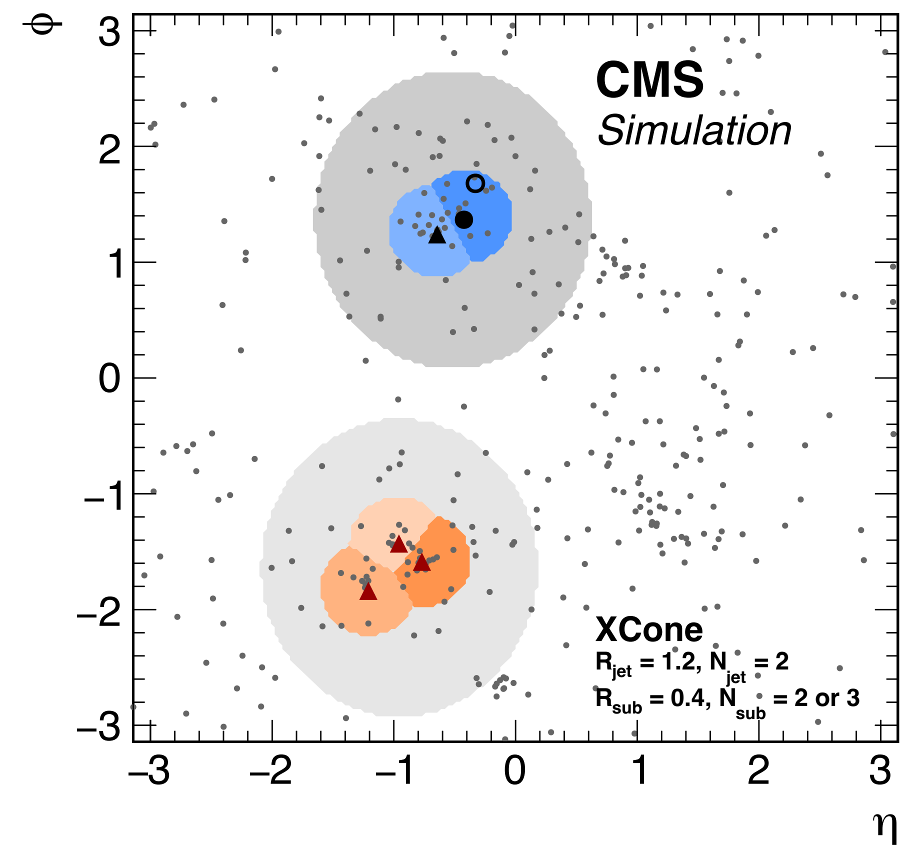
[Stewart et al., JHEP 11, 072 (2015)]

XCone

Use number of expected jets when event topology is known



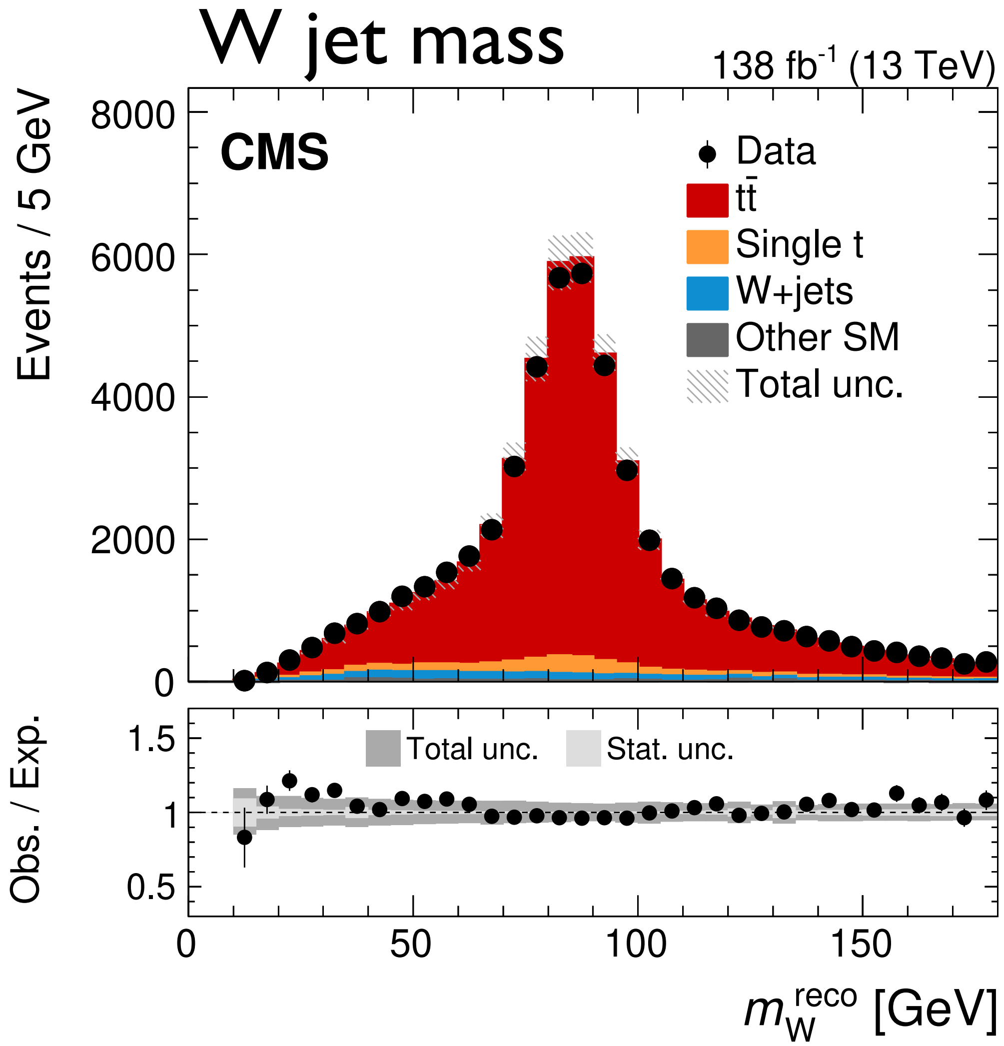


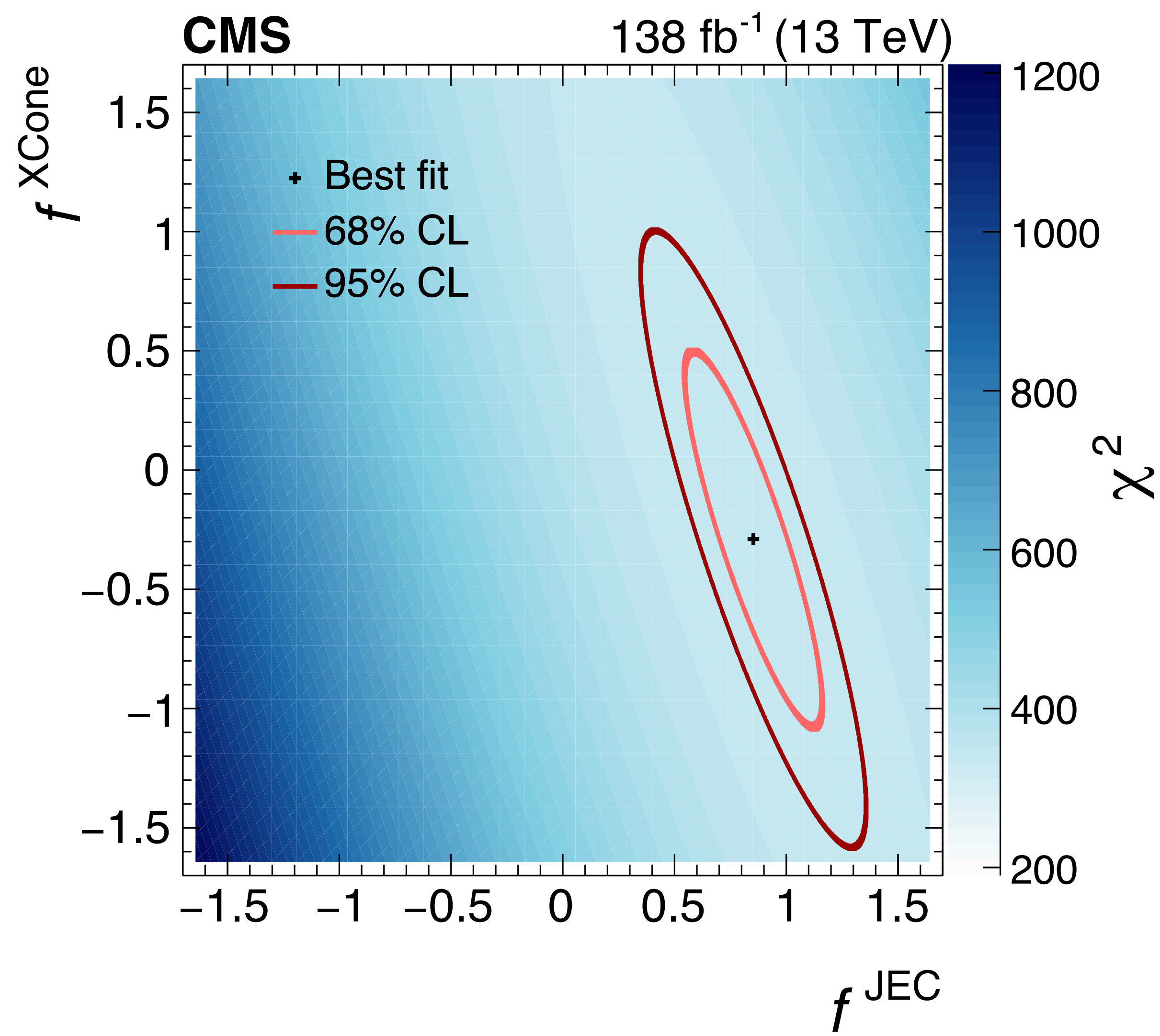


- XCone assigns particles based on N-jettiness axes
- Natural transition resolved ↔ boosted

[Stewart et al., JHEP 11, 072 (2015)]

Measuring with XCone

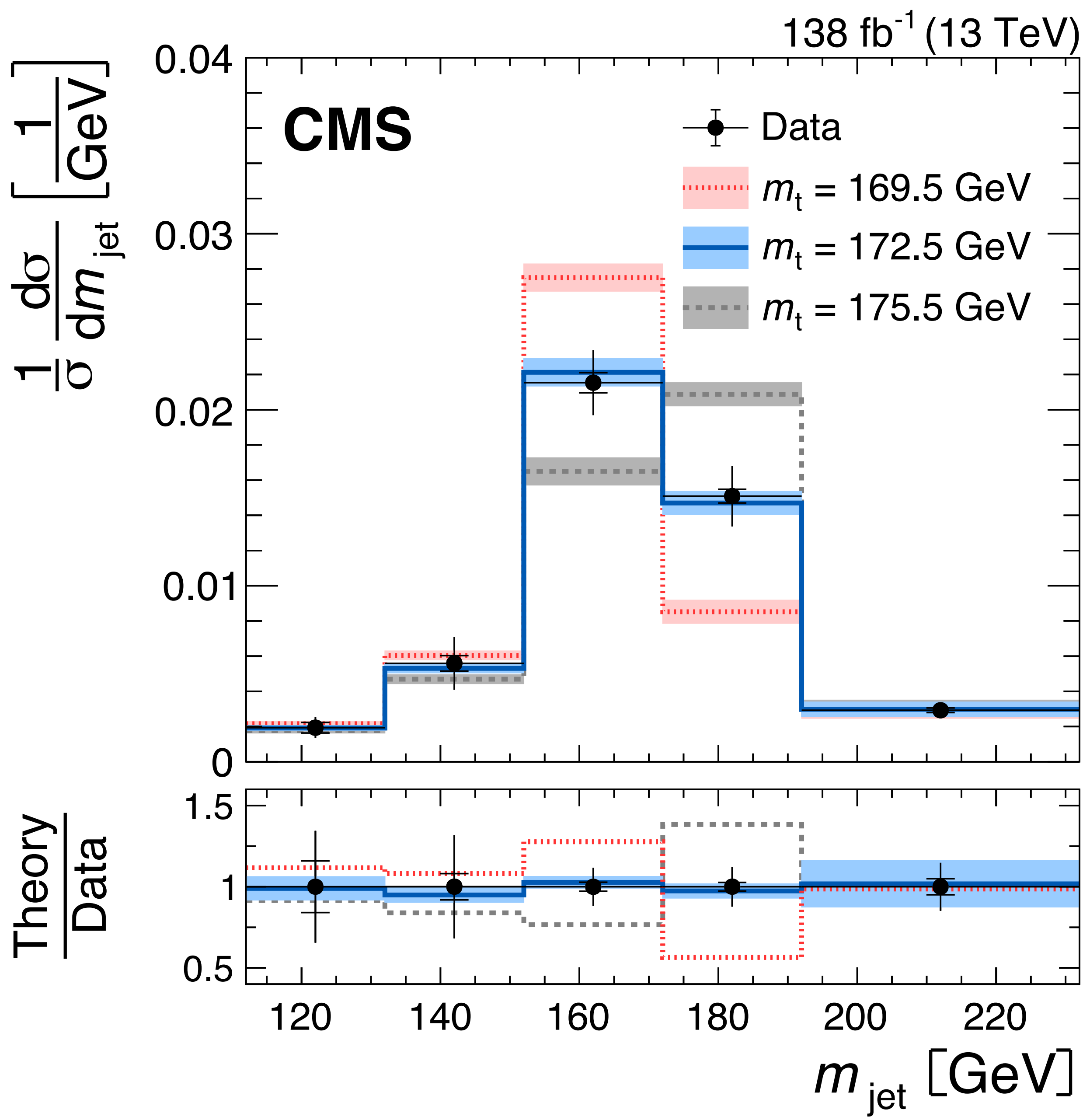




- Calibrate jet mass using "standard candle" M_W
- Excellent jet mass resolution of 6-8%

[CMS, arXiv:2211.01456]

Top Quark Mass with XCone



- Top quark mass from unfolded cross section
  - Uncertainty of ~ 0.8 GeV

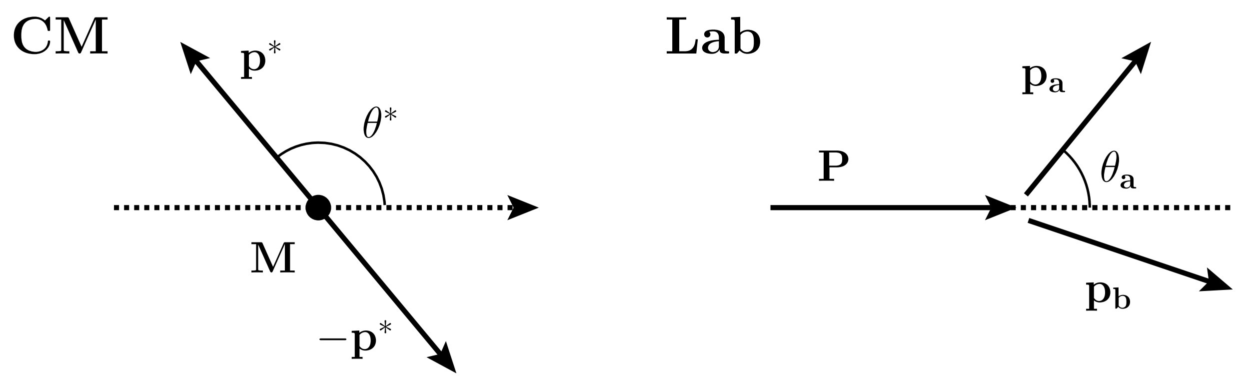
[CMS, arXiv:2211.01456]

Variable R Jets

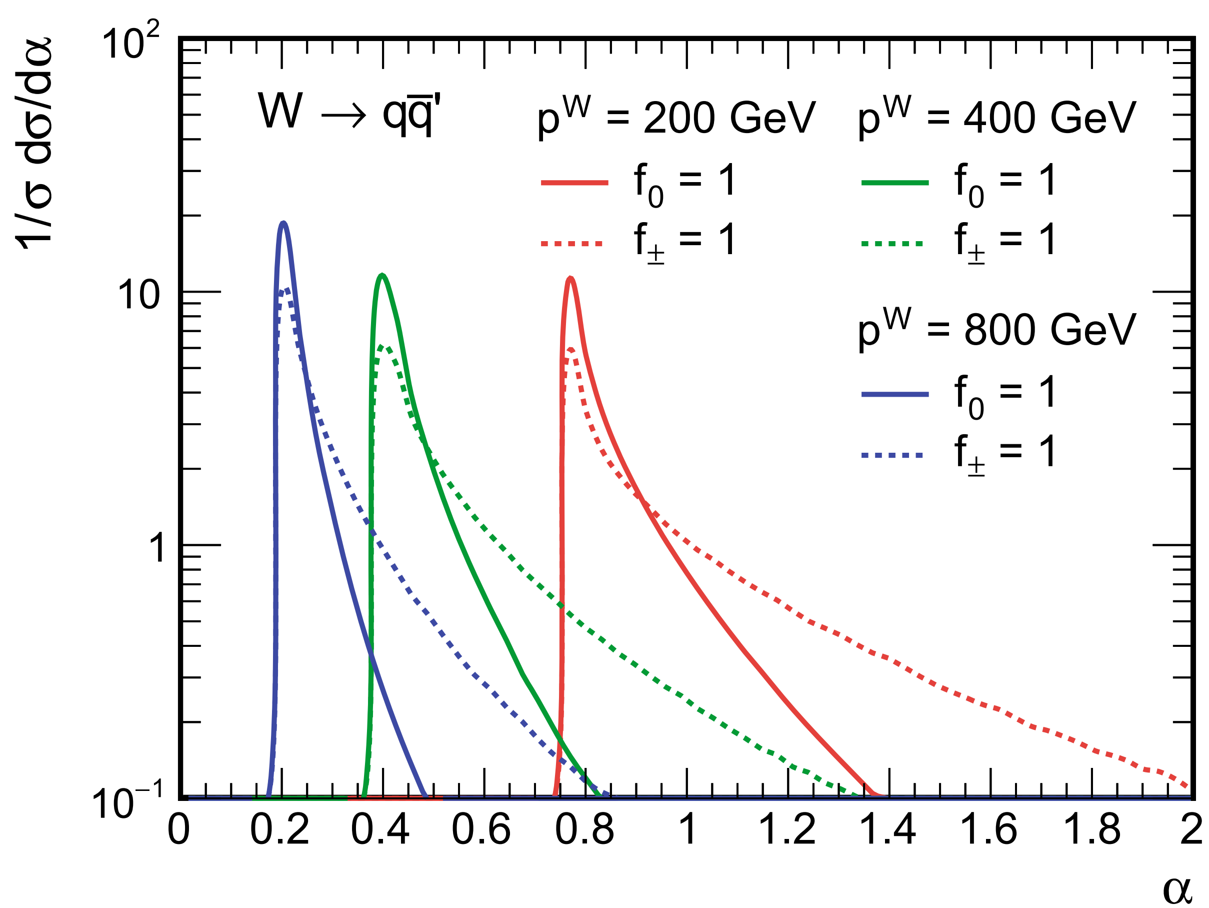
Particle Decays

W and Z bosons

$B_{W\to had} = 67.5\%$
$B_{Z\to had} = 69.2\%$



$$\frac{1}{\sigma}\frac{d\sigma}{d|\cos\theta^*|} = f_\pm \frac{3}{4}\left(1+|\cos\theta^*|^2\right) + f_0\frac{3}{2}|\sin\theta^*|^2$$



$$\alpha_{\min} \approx \frac{2M}{P}$$

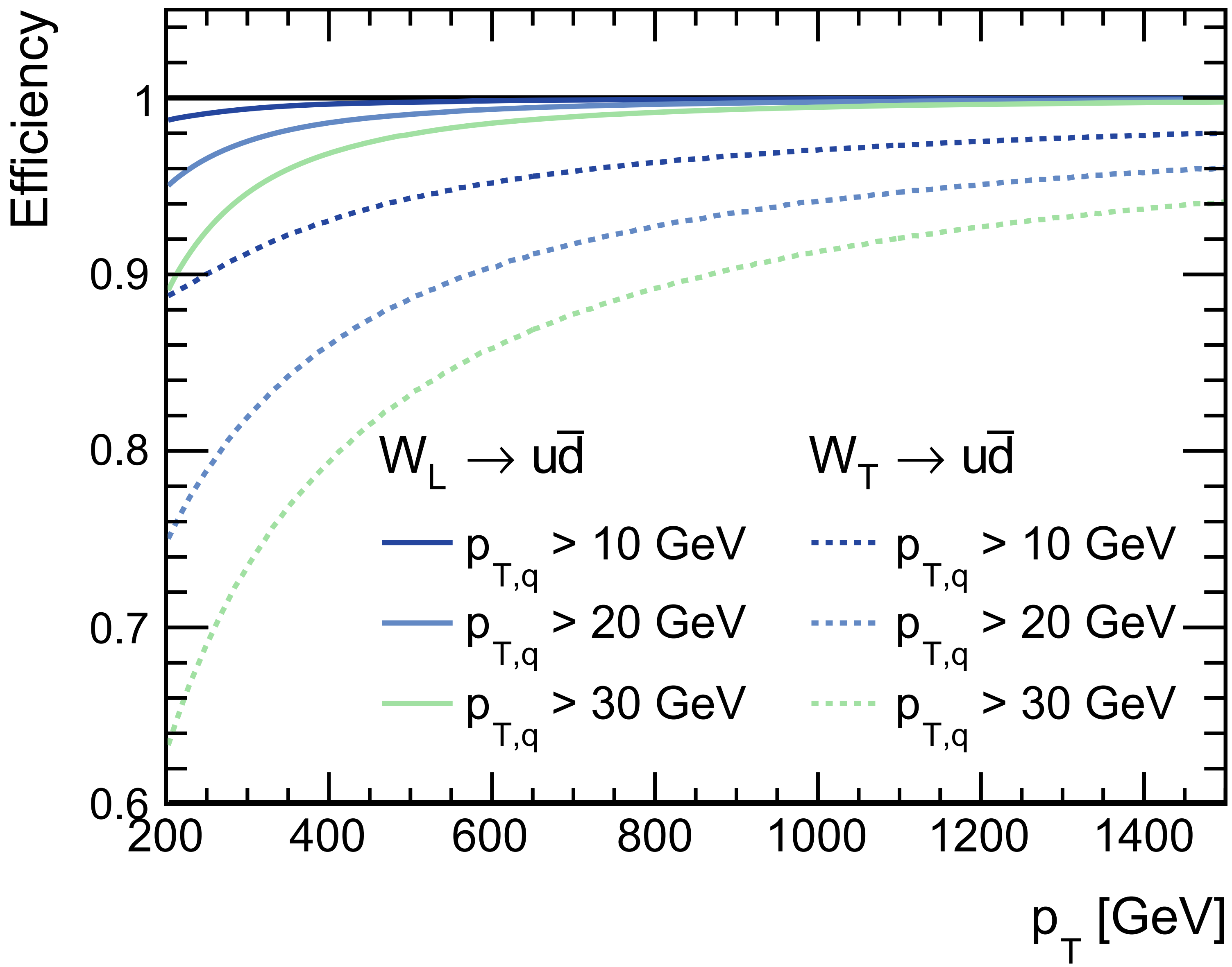
and consequently

$$\Delta R \approx \frac{2M}{P_T}$$

(holds for $P_T \gg M$)

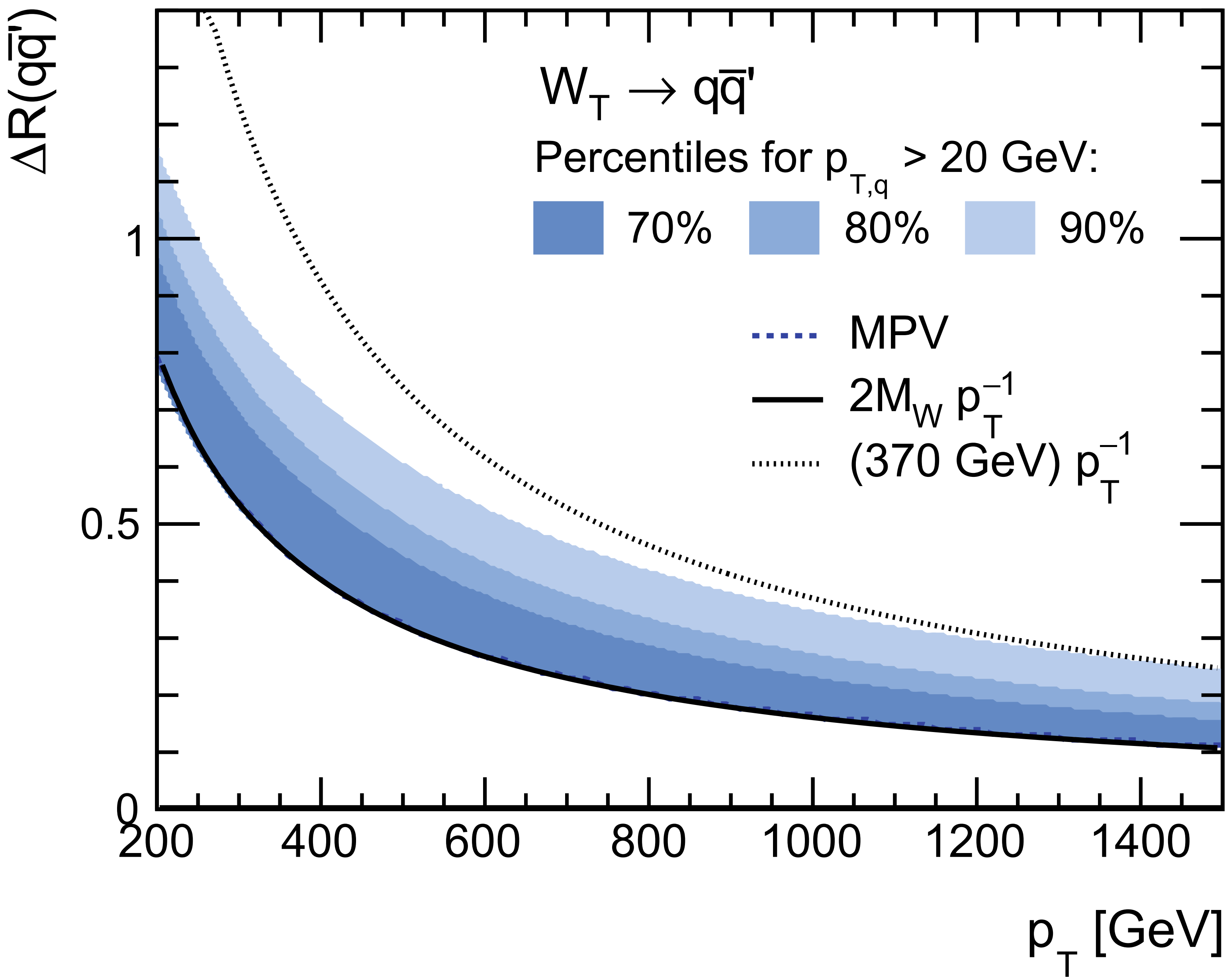
[RK, STMP 284 (2021)]

Quark (subjet) pT thresholds



[RK, STMP 284 (2021)]

Decay distance



- Similar picture for top quarks

[RK, STMP 284 (2021)]

Heavy Object Tagger with Variable R

[Lapsien, Haller, RK, EPJC 76, 600 (2016)]

One-pass clustering with integrated subjet finding

- Jet distance measures (with variable R)

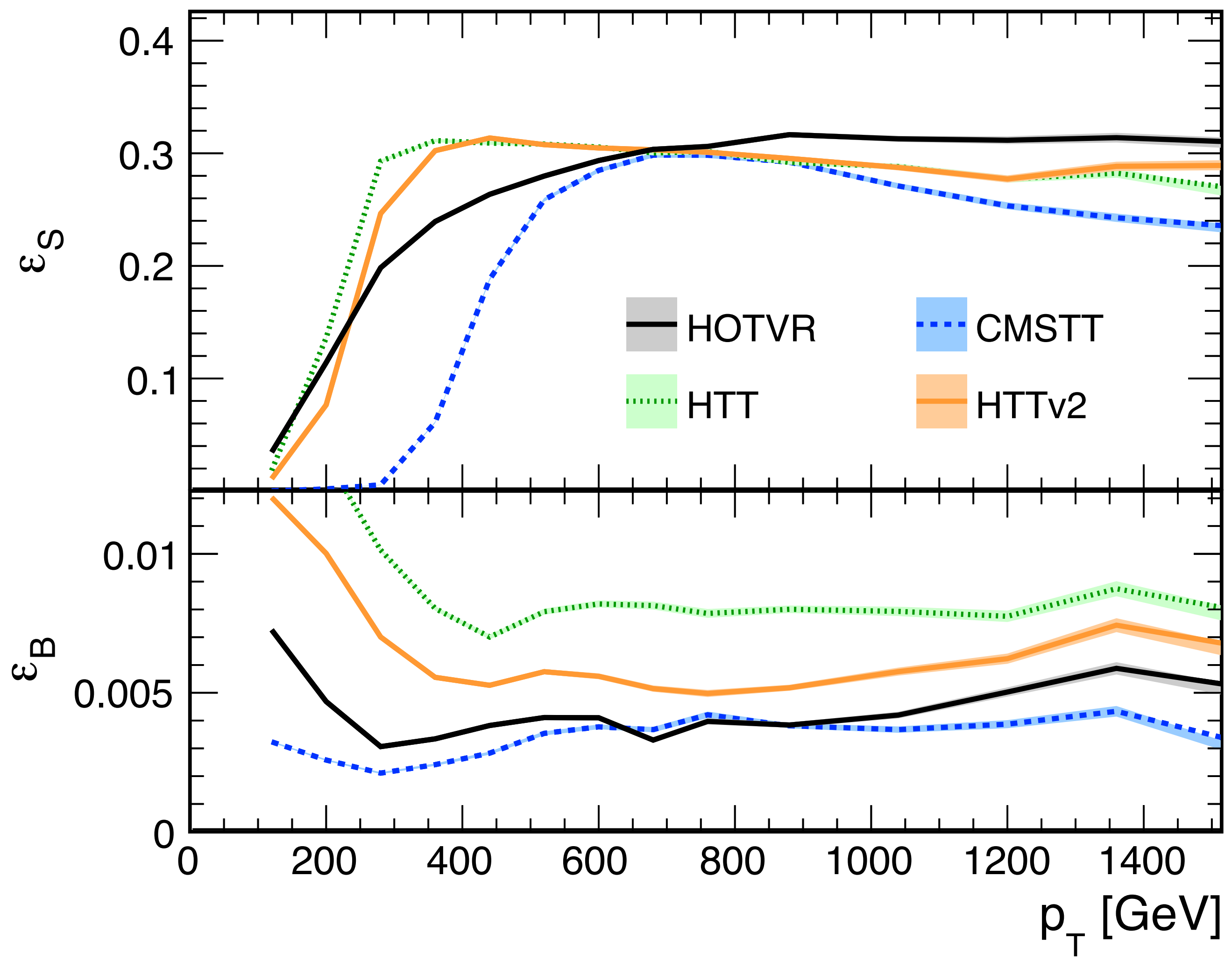
$$d_{ij} = \min[p_{T,i}^{2n}, p_{T,j}^{2n}]\Delta R_{ij}^2$$

$$d_{iB} = p_{T,i}^{2n} R_{\text{eff}}^2 \qquad R_{\text{eff}} = \frac{\rho}{p_T}$$

- Clustering veto at each step
  - mass jump veto
- Store objects i and j as subjets
- Used in tW resonance search

[CMS, JHEP 04, 048 (2022)]

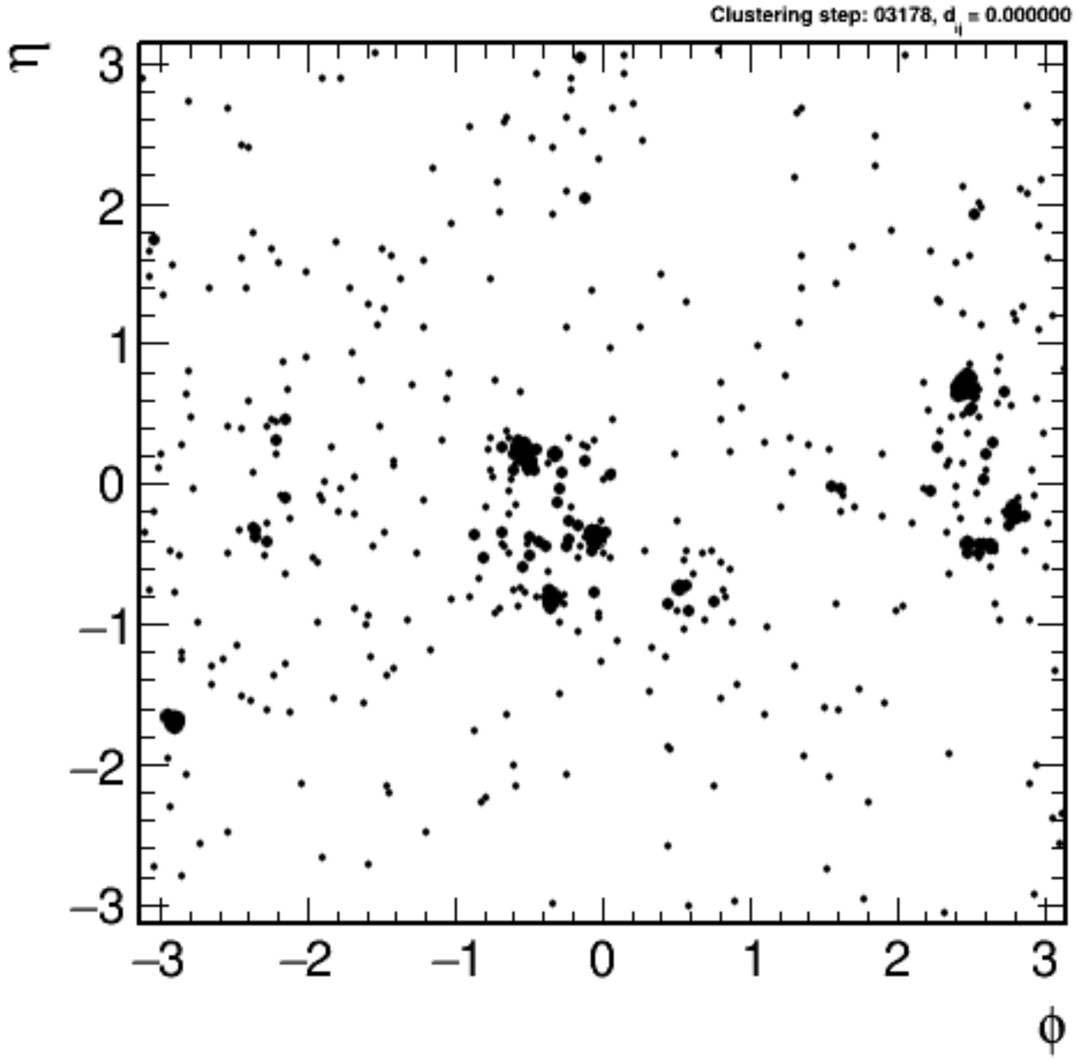
- Works beautifully, but can be improved



[RK, STMP 284 (2021)]

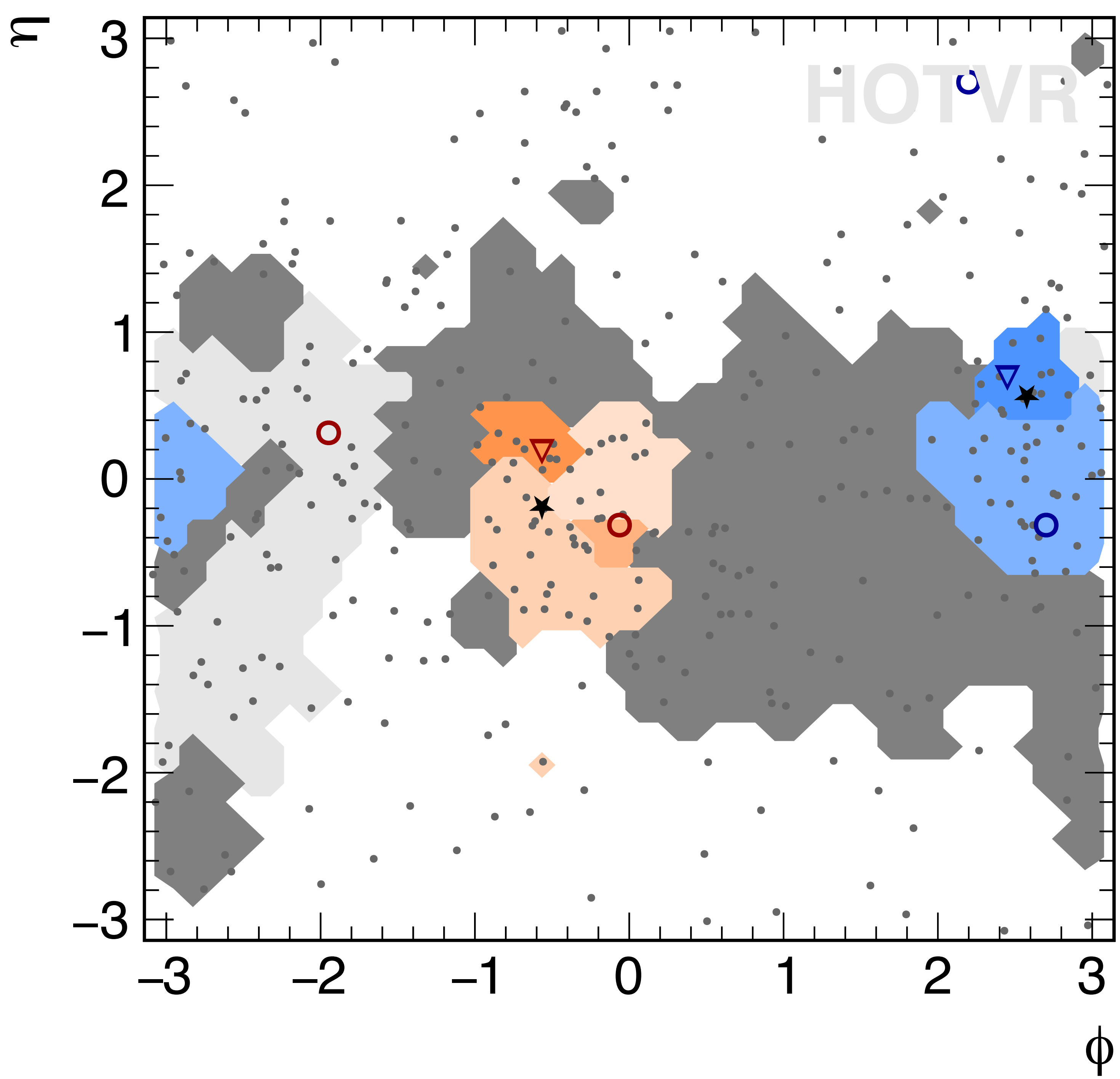
HOTVR with soft drop

- Use soft drop veto instead of mass jump
  - At each clustering step, test $\frac{\min(p_{T,i}, p_{T,j})}{p_{T,i} + p_{T,j}} > z_{\text{cut}} \left(\frac{\Delta R_{ij}}{R_{\text{eff}}} \right)^{\beta}$
  - Remove softer subjet if not fulfilled
  - Active area exactly 0, because ghosts get groomed
- Expand with mass-dependent R (work in progress)



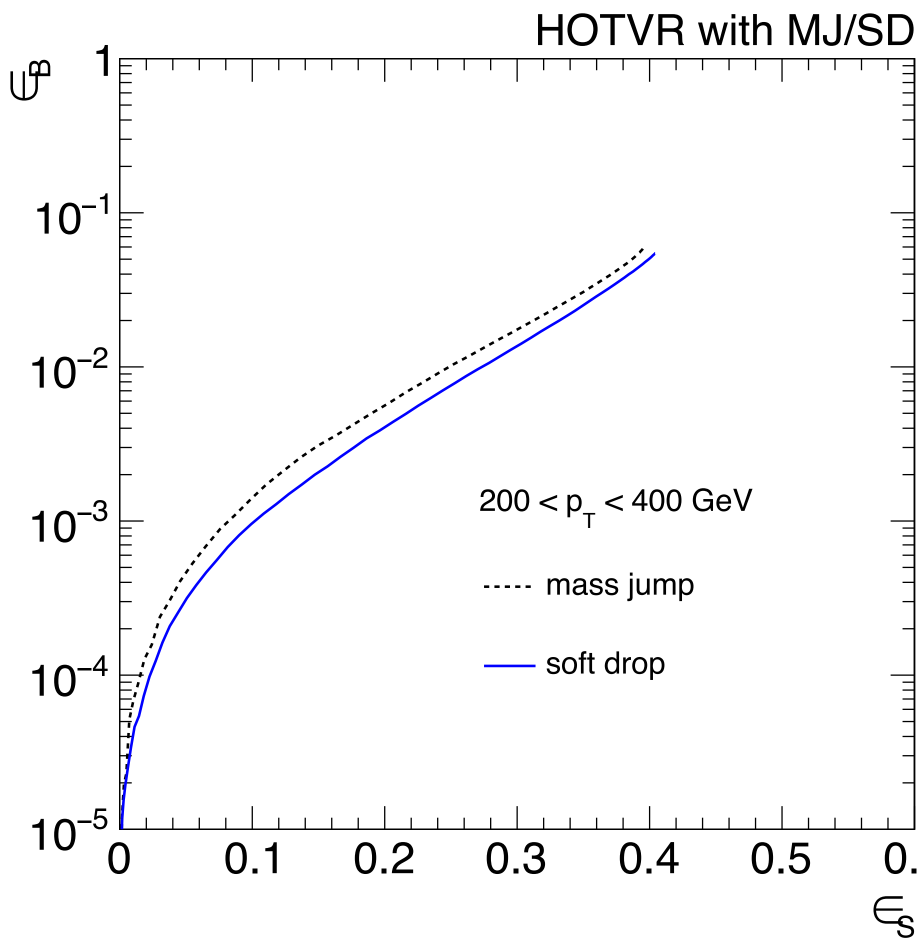
[Albrecht, Benecke, RK, work in progress]

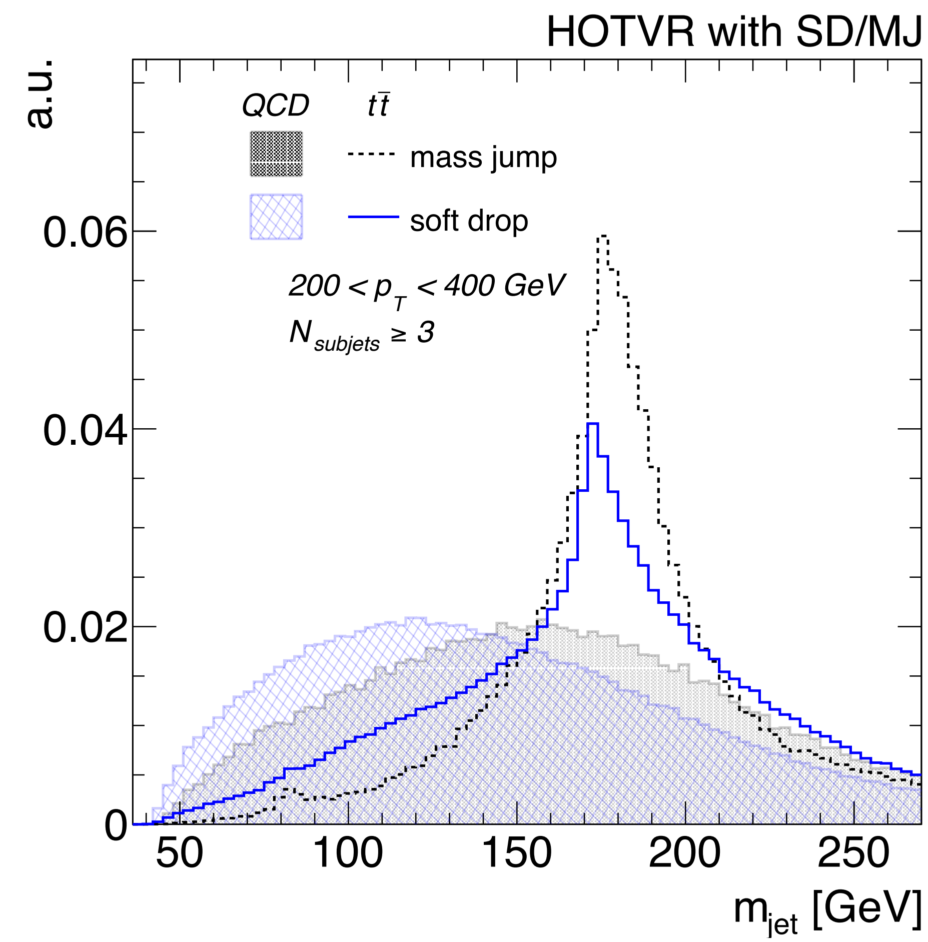
HOTVR with soft drop



[Albrecht, Benecke, RK, work in progress]

HOTVR with soft drop





- Stronger grooming with soft drop
- No essential tagging information is lost with HOTVR-SD jets compared to plain Variable R jets
  - Better starting point for (ML) taggers

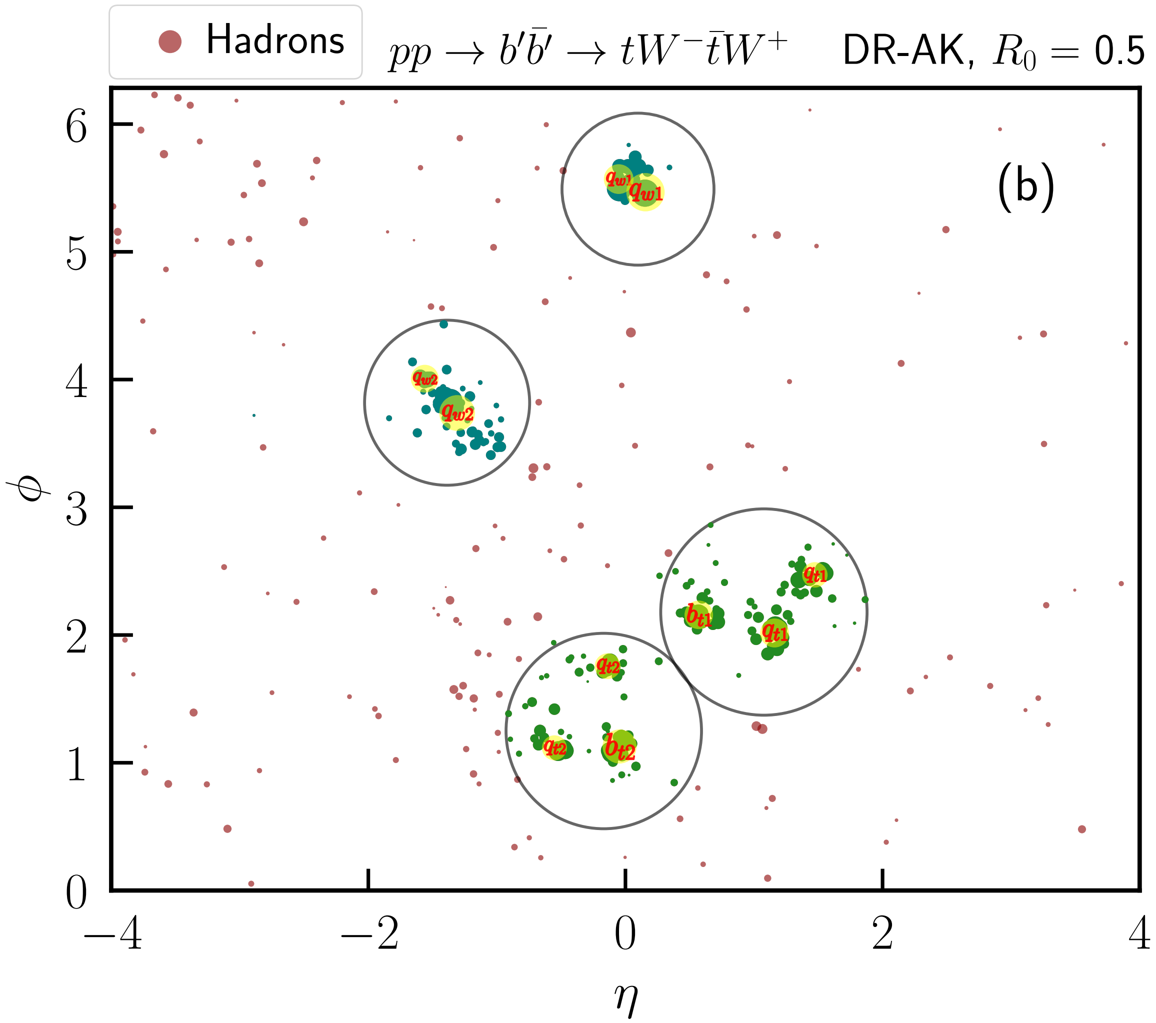
[Albrecht, Benecke, RK, work in progress]

Scale Invariant Jets

Other Variable-Size Jets

Local, dynamical R [Mukhopadhyaya, Samui, Singh, JHEP 2023, 19 (2023)]

$$R_{d_i} = R_0 + \sigma_i . \qquad \sigma_i^2 = \frac{\sum_{a<b} p_{T_a} p_{T_b} \Delta R_{ab}^2}{\sum_{a<b} p_{T_a} p_{T_b}} - \left(\frac{\sum_{a<b} p_{T_a} p_{T_b} \Delta R_{ab}}{\sum_{a<b} p_{T_a} p_{T_b}} \right)^2 ,$$



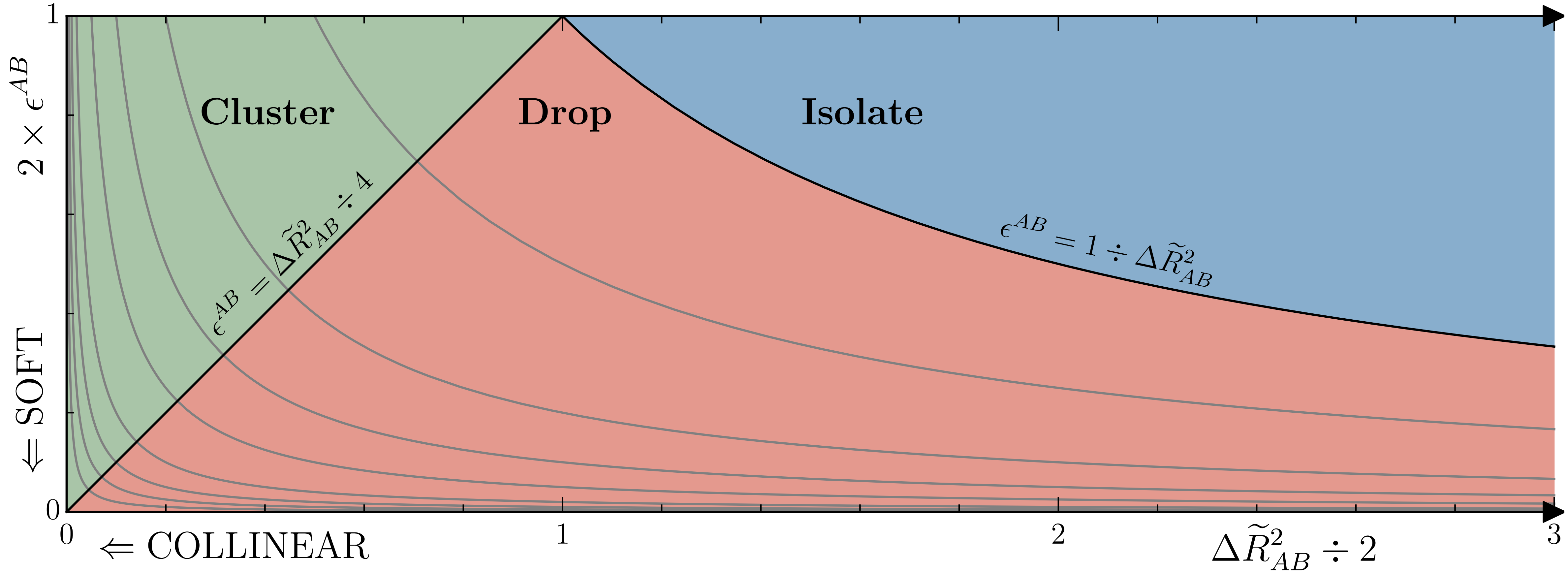
- Useful in searches with high p_T and multi-prong resonances
- Minimum R_0 needed, can not have jets smaller than that
- Adjustment of R_0 to analysis needs

Scale Invariant Jets

- Optimal distance parameter R depends on energy scale of event
- Idea: a scale-invariant algorithm, independent of R

$$\delta_{AB} = \epsilon^{AB} \times \Delta\widetilde{R}^2_{AB} = \frac{\cosh \Delta y_{AB} - \xi^A \xi^B \cos \Delta\phi_{AB}}{\cosh \Delta u_{AB}}$$

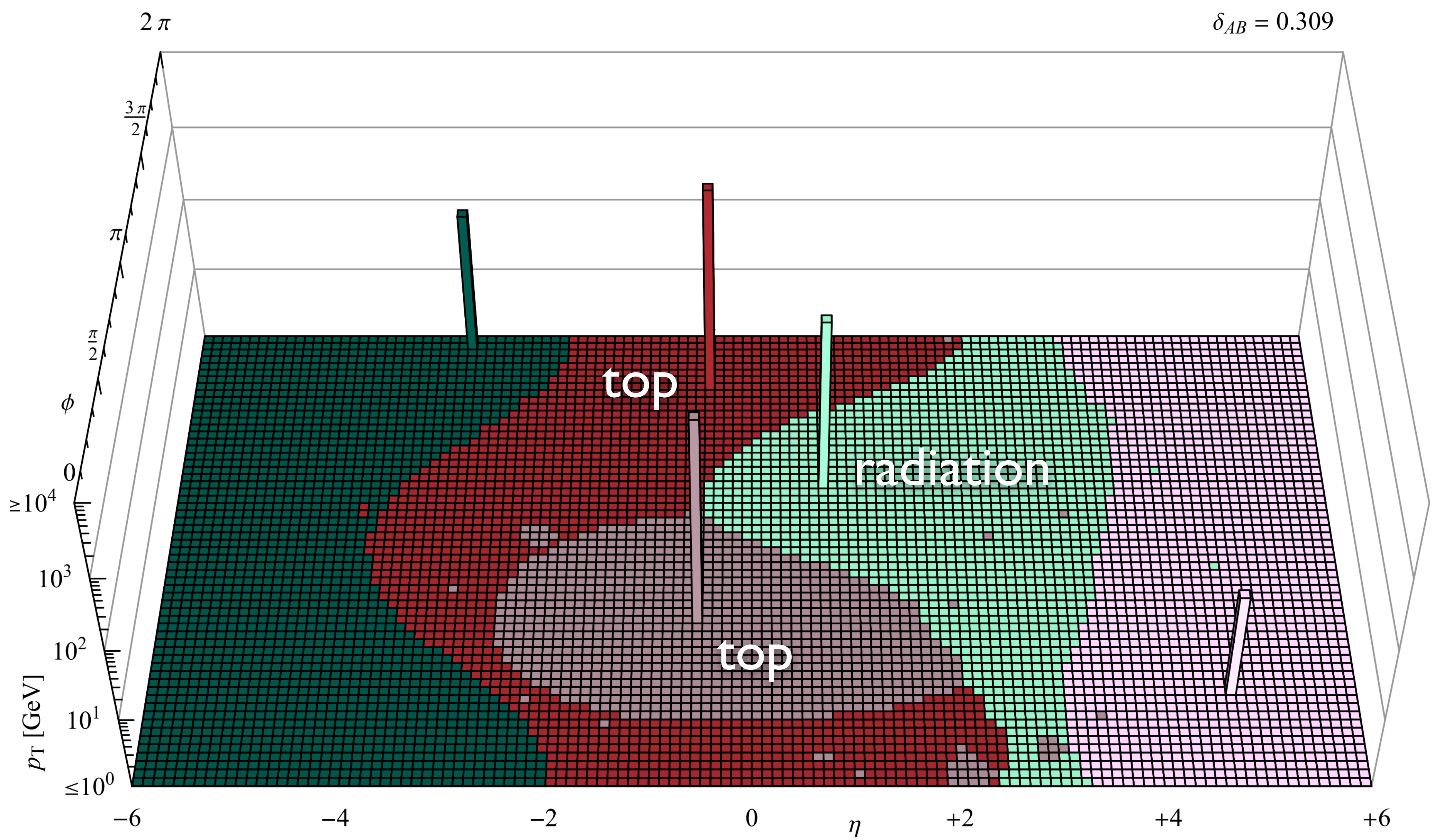
- Inherent soft-drop-like grooming in "Drop" region



[Larkoski, Rathjens, Veatch, Walker, arXiv:2302.08609]

Scale Invariant Jets

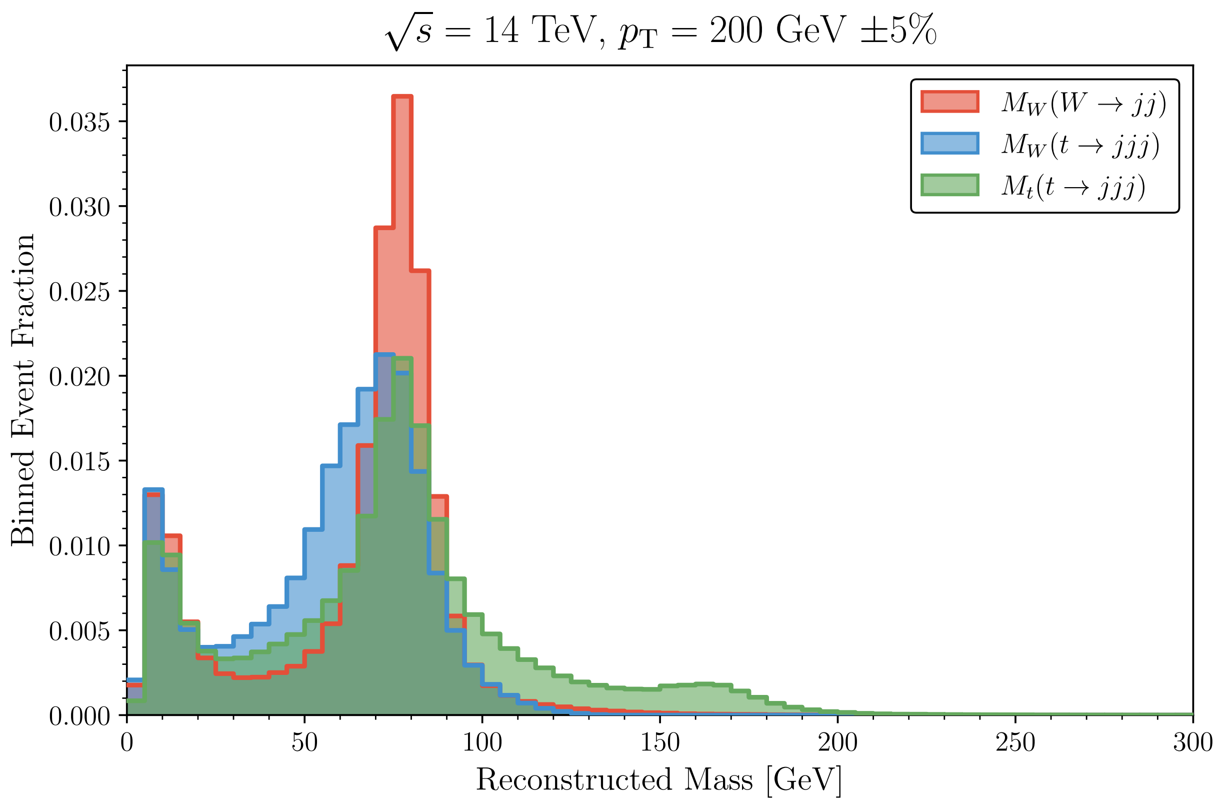
- If left running, the whole event will merge into one large jet
  - Large discontinuity in distance measure δ_{AB} in the last steps

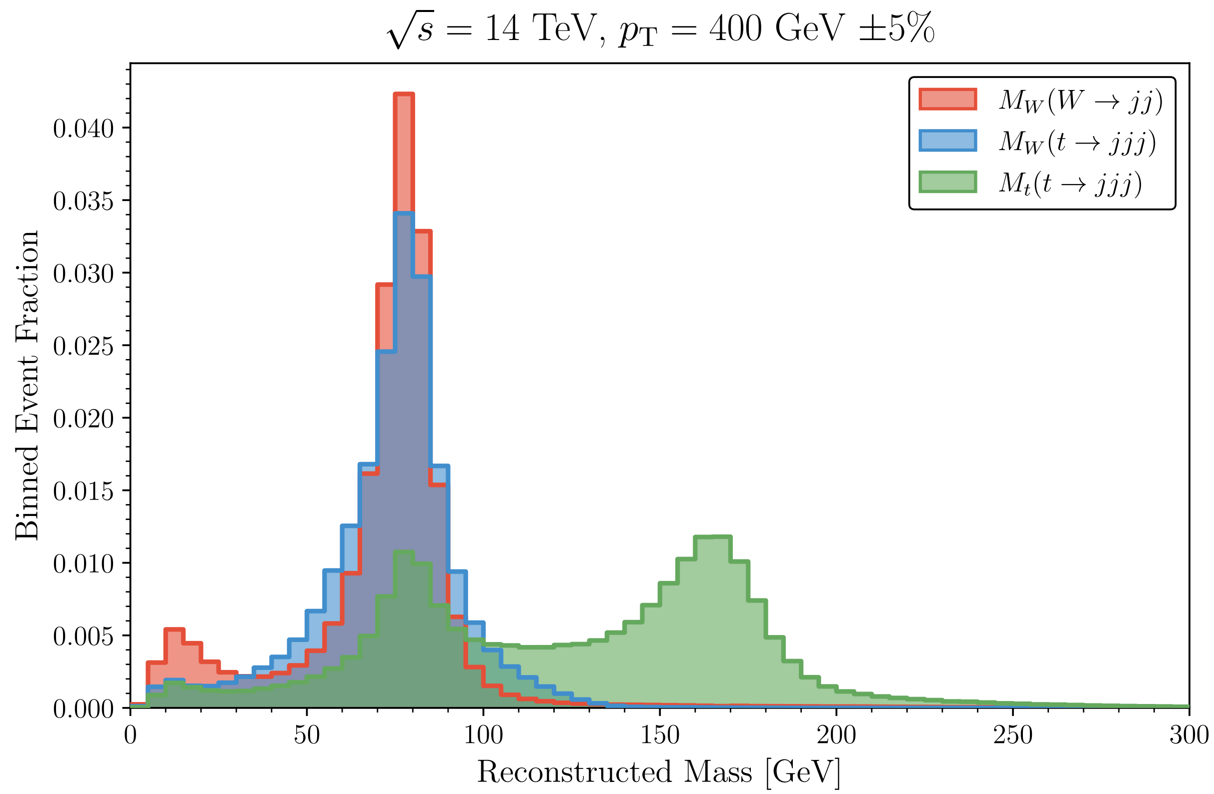


[Larkoski, Rathjens, Veatch, Walker, arXiv:2302.08609]

Scale Invariant Filtered Tree (SIFT)

Clustering history (N-subjet tree): Exclusive (sub)jet counts





[Larkoski, Rathjens, Veatch, Walker, arXiv:2302.08609]

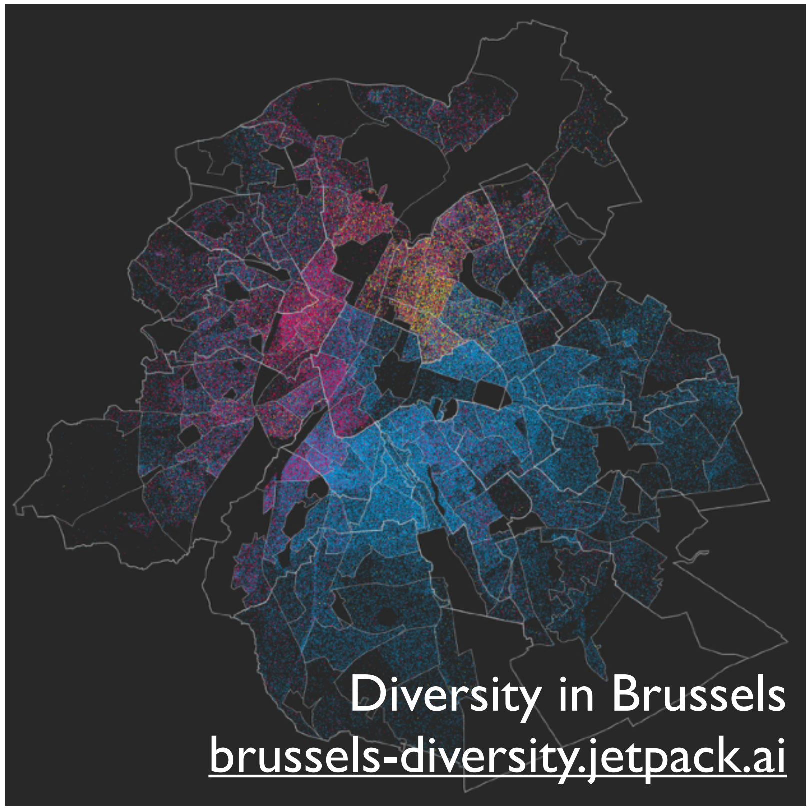
- Promising results over a large range of p_T
- Tagging results (obtained with BDT) better than for fixed-R jets

Comprehensive comparison of all algorithms needed

Summary

Last 10 years huge progress in jets and jet substructure

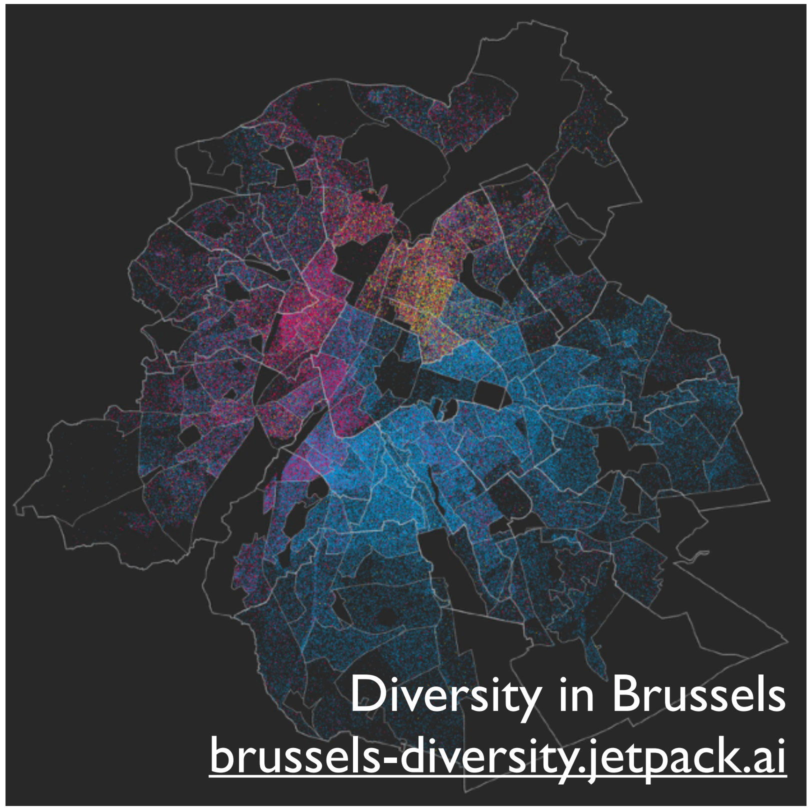
- Why are we still using AK4 for measurements and AK8 for tagging?
- Should be using:
  - Large jets (or R ~ p_T) for measurements
  - Decreasing jets R ~ $1/p_T$ for tagging
  - Unambiguous (IRC safe) definition of q/g jets



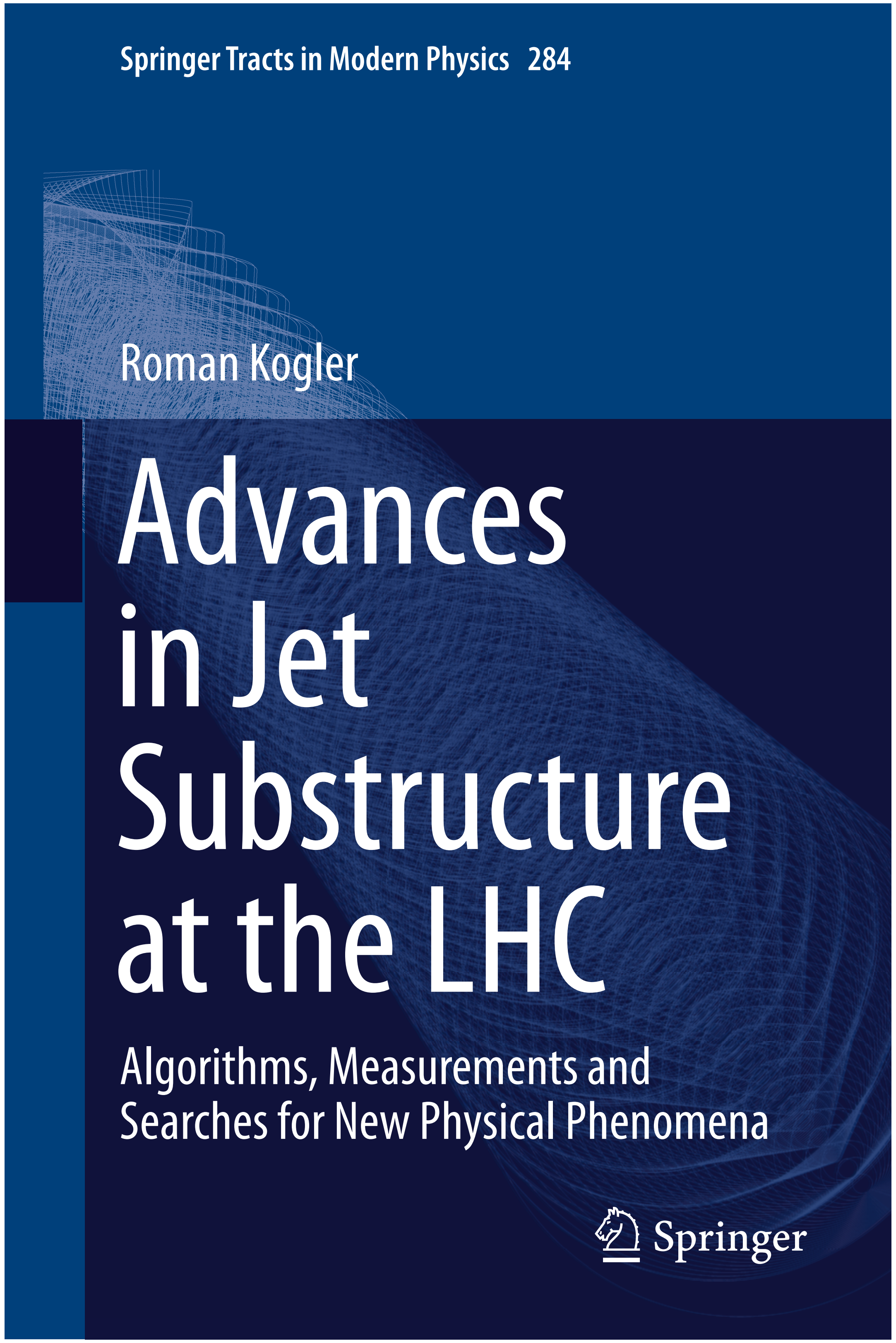
Summary

Last 10 years huge progress in jets and jet substructure

- Why are we still using AK4 for measurements and AK8 for tagging?
- Should be using:
  - Large jets (or R ~ p_T) for measurements
  - Decreasing jets R ~ $1/p_T$ for tagging
  - Unambiguous (IRC safe) definition of q/g jets



- We are all busy with Run 3 (and 2)
- Hopefully, no catastrophic incident is needed for the next consolidation of our jet usage
  - Preparation of HL-LHC: chance for new ideas
  - Be open for new techniques and strategies
  - Start with data formats, analyses will follow



Thanks

CMS group of the University of Hamburg:

Johannes Haller, Andreas Hinzmann, Alexander Schmidt, Thomas Peiffer, Valentina Sola, Jochen Ott, Anastasia Karavdina, Robin Aggleton, Kristin Göbel, Tobias Lapsien, Mareike Meyer, Dominik Nowatschin, Daniel Gonzalez, Marc Stöver, Vilius Kripas, Torben Dreyer, Arne Reimers, Mehdi Mamoumi, Andreas Kell, Arne Reimers, Eugen Trapp, Jens Multhaup, Anna Benecke, Melanie Eich, Dennis Schwarz, Alexander Fröhlich, Andrea Malara, Tim Christensen, Alexander Paasch, Mathis Frahm, Jan Skottke, Steffen Albrecht, Serge Rosen, Henrik Jabusch, Nino Ehlers, Christopher Matthies, Ksenia de Leo, Irene Zoi, Anna Albrecht, Finley Quinton, Finn Labe, Tom Sokolinski

CMS Physics Group “Beyond 2 Generations”

CMS Physics Group “Top quarks”

CMS Object Group “Jets and Missing Energy”

All colleagues from ATLAS, CMS and Theory

My family and friends

[RK, STMP 284 (2021)]