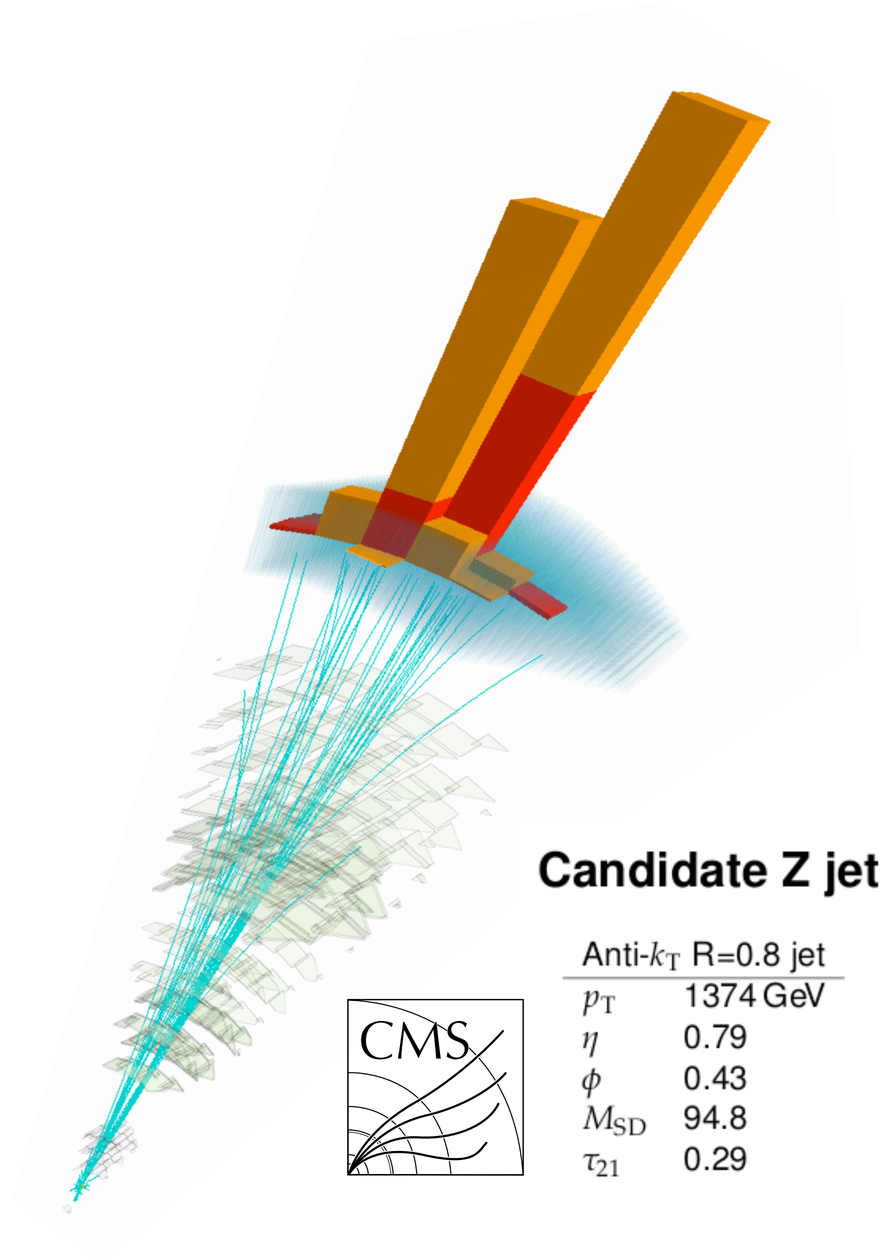
The Coming of Age of Jet Substructure at the LHC

Roman Kogler
DESY and Universität Hamburg

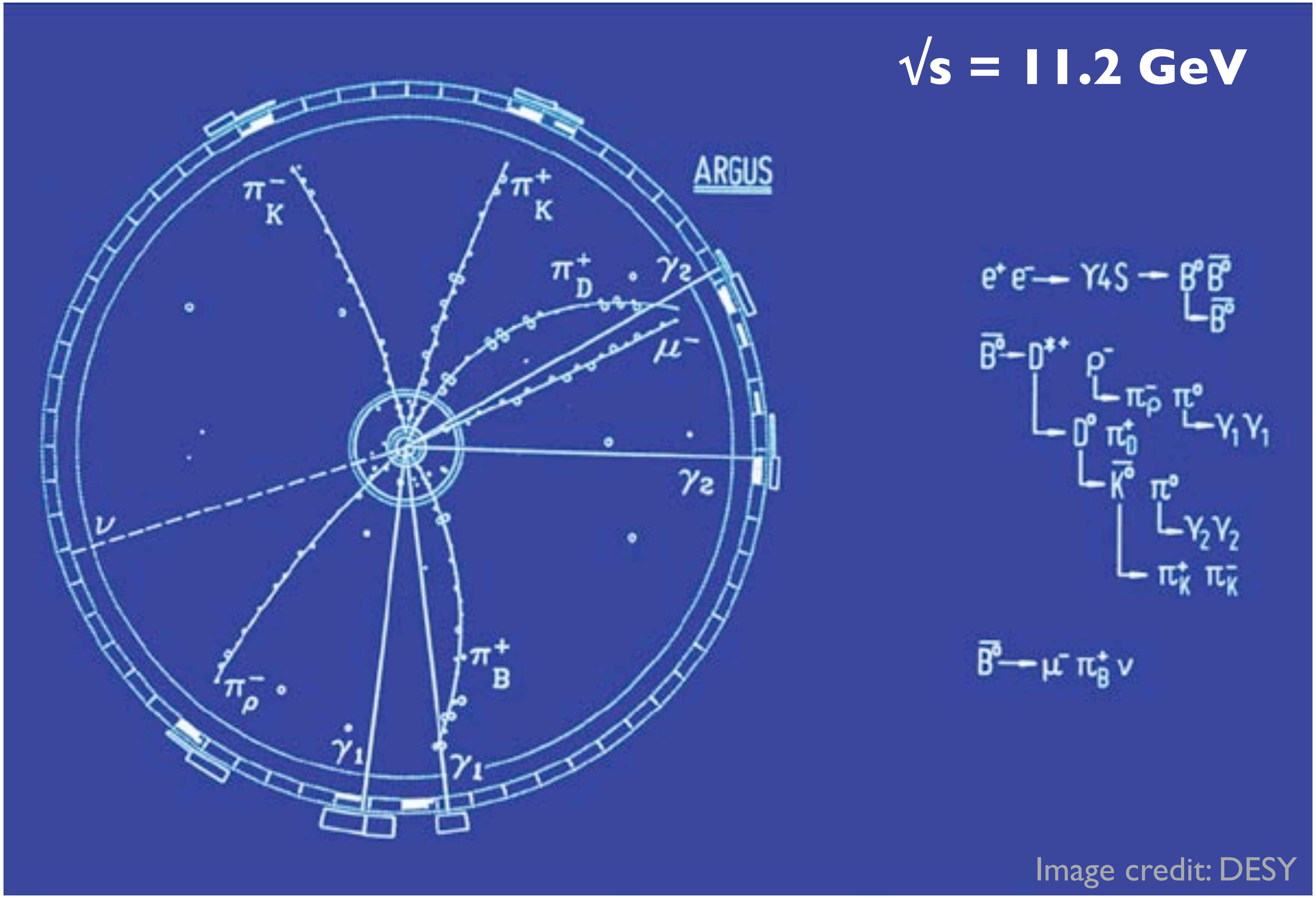
Universität Zürich
Sep 18, 2023

Overview

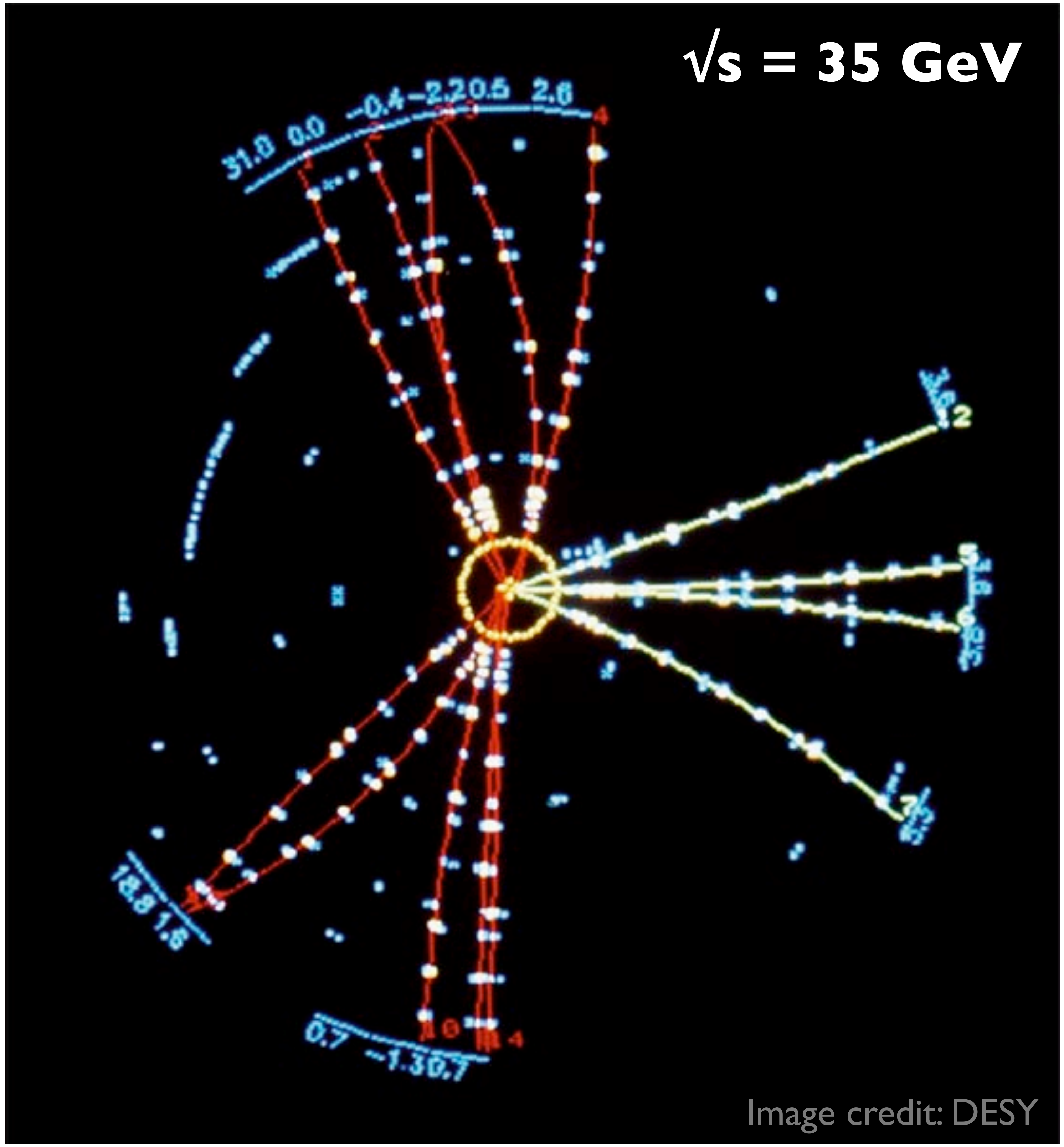
- Introduction
- Jet algorithms and substructure
- Jet tagging with substructure
- Calibration and commissioning
- Standard model measurements
- Searching for new phenomena
- Summary



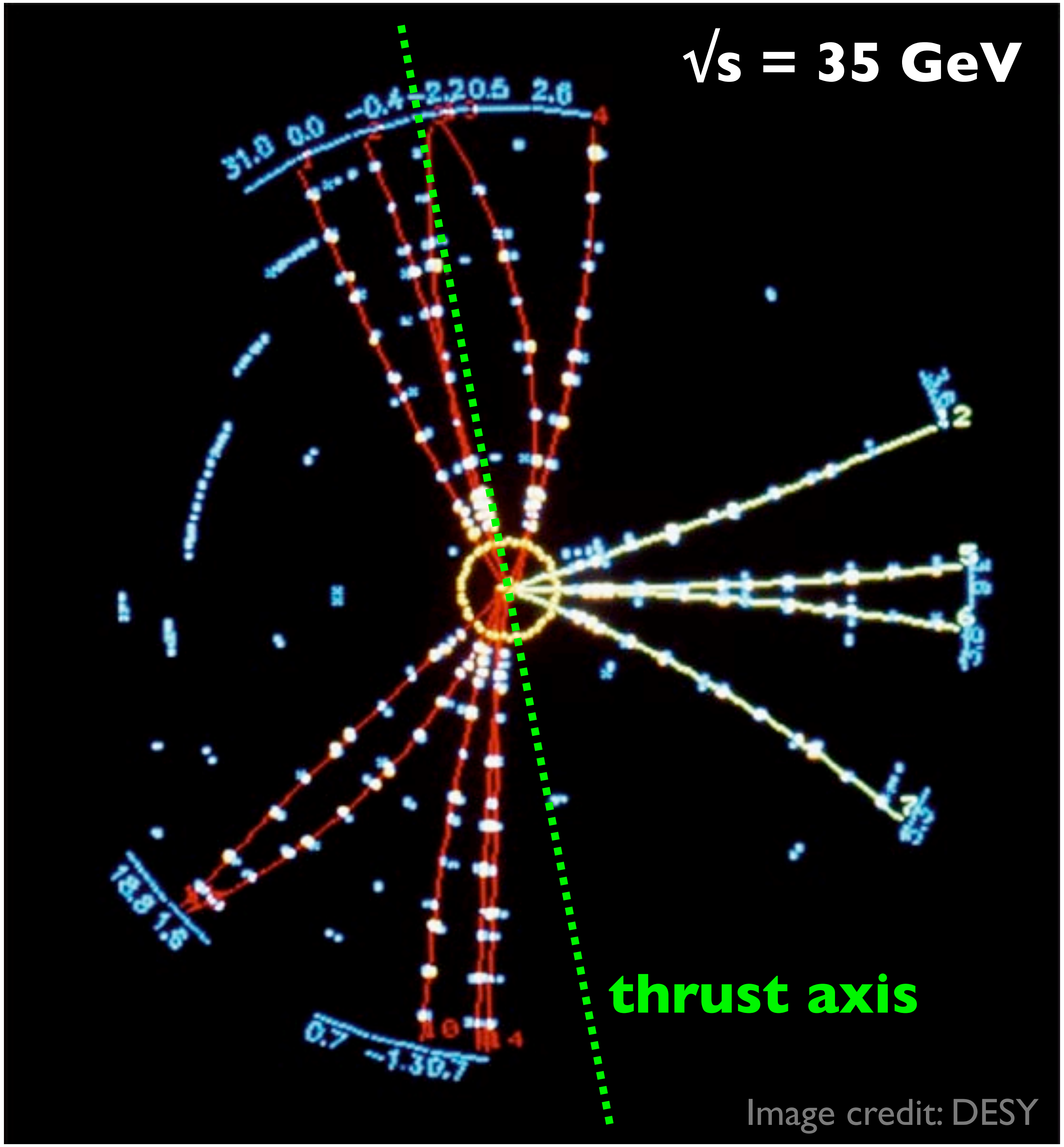
ARGUS at DORIS, 1987



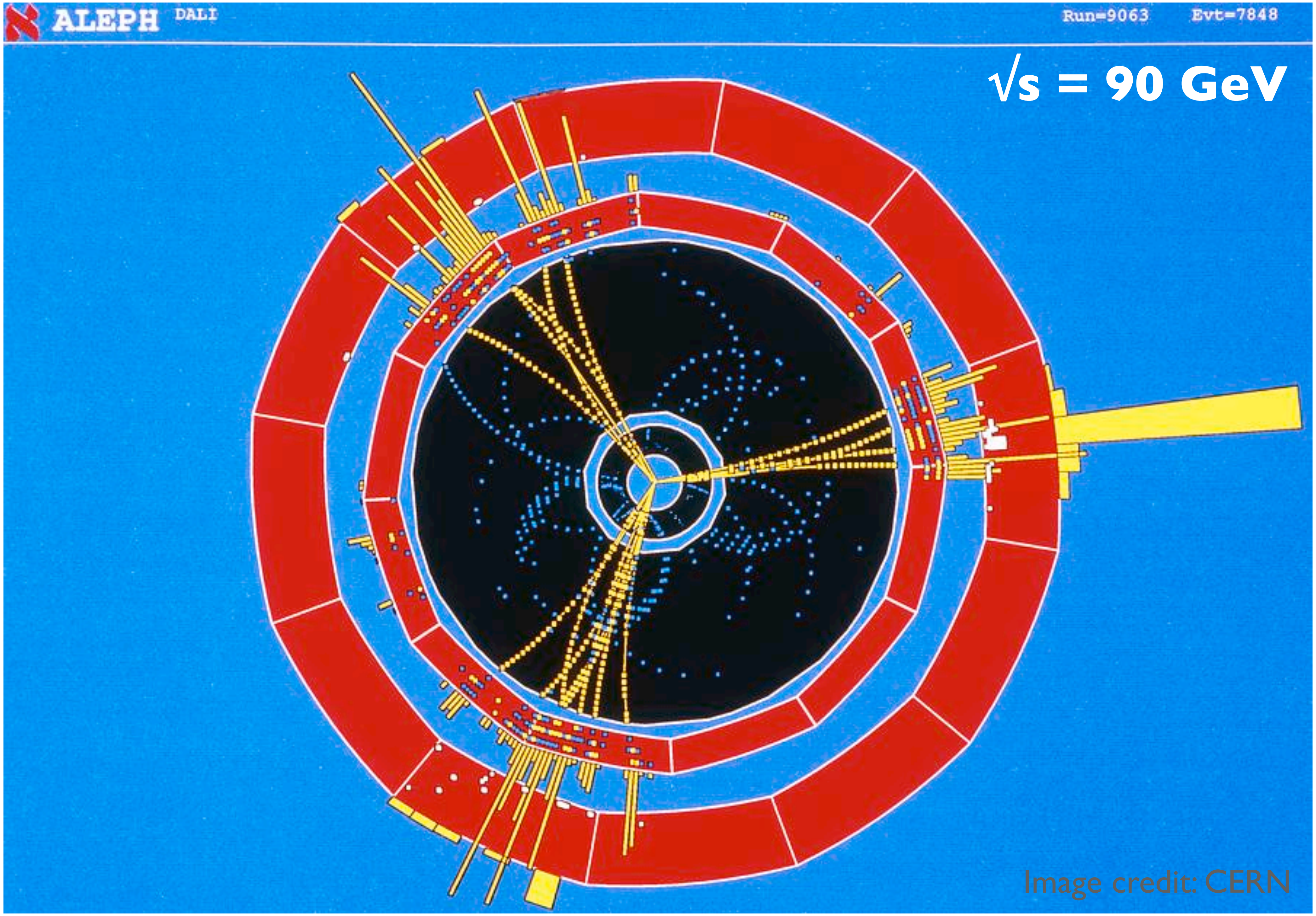
TASSO at PETRA, 1979



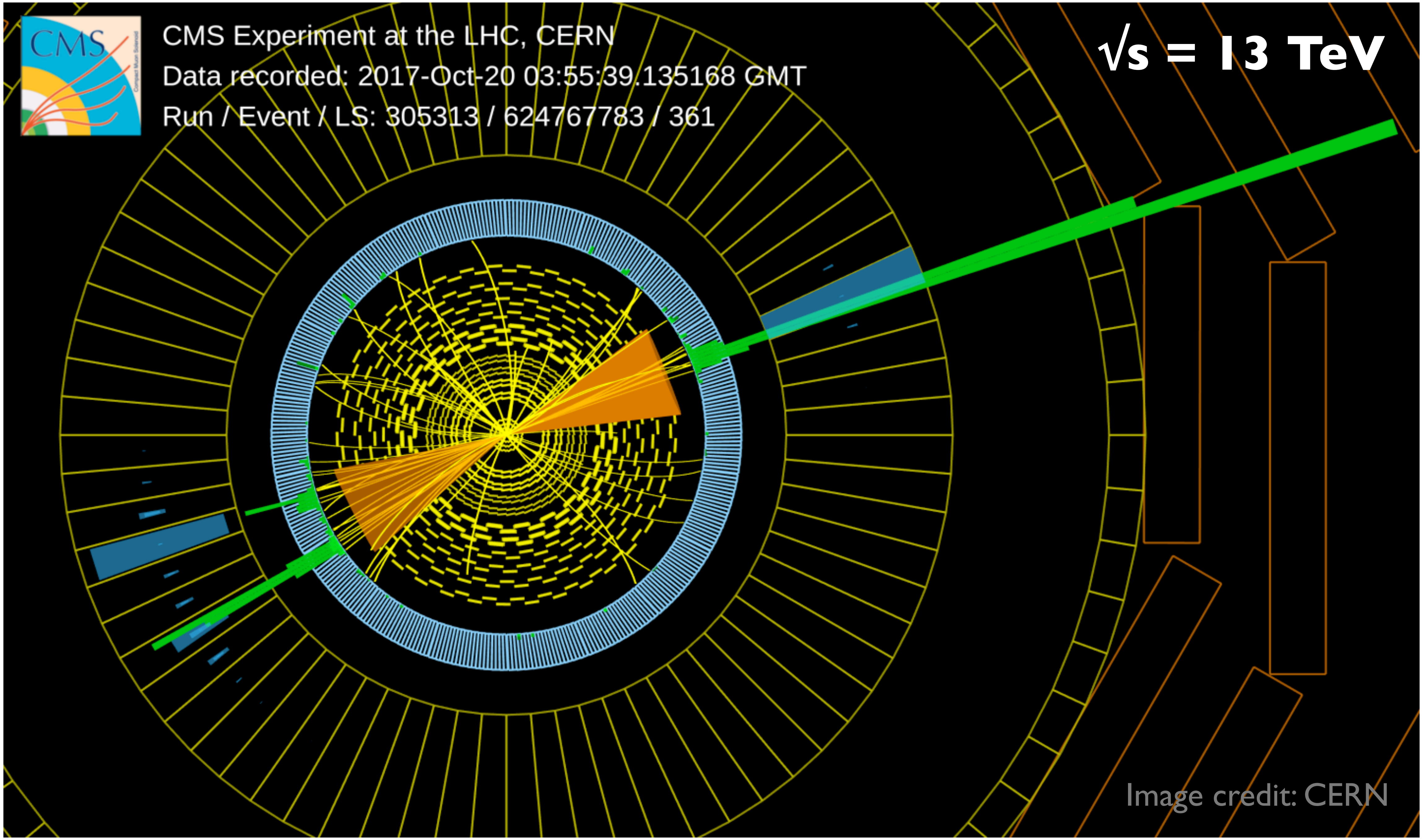
TASSO at PETRA, 1979



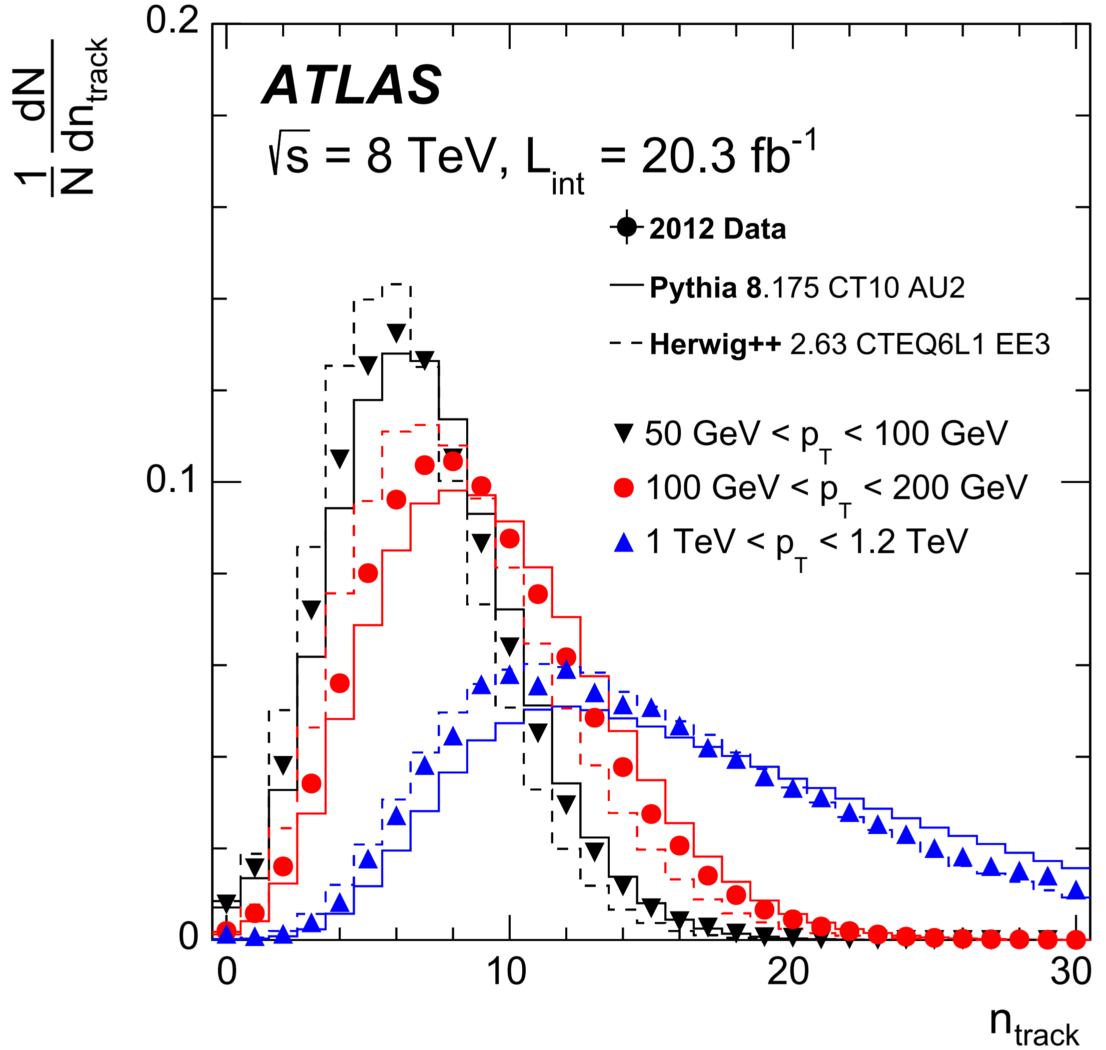
ALEPH at LEP, 1992



CMS at LHC, 2017



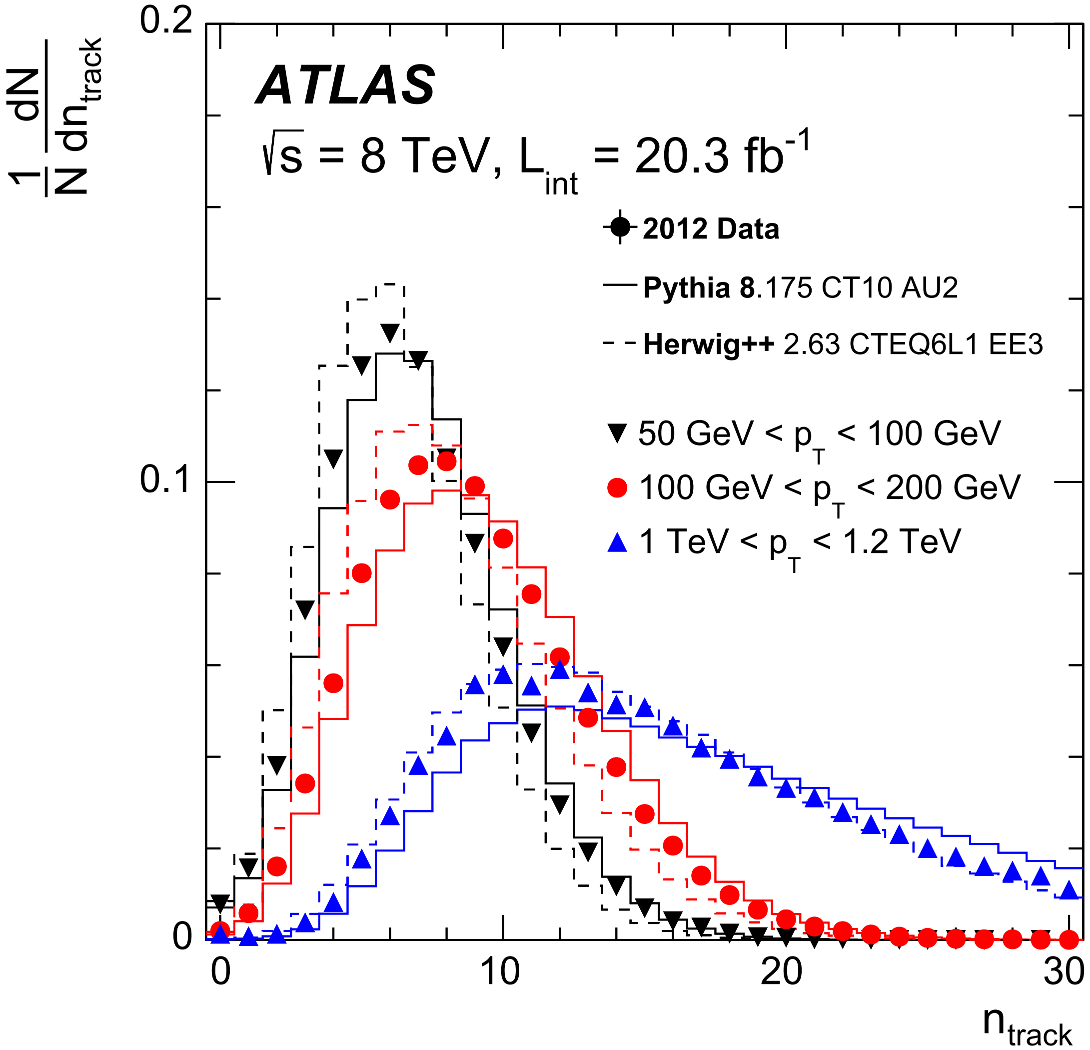
Charged particles in jets

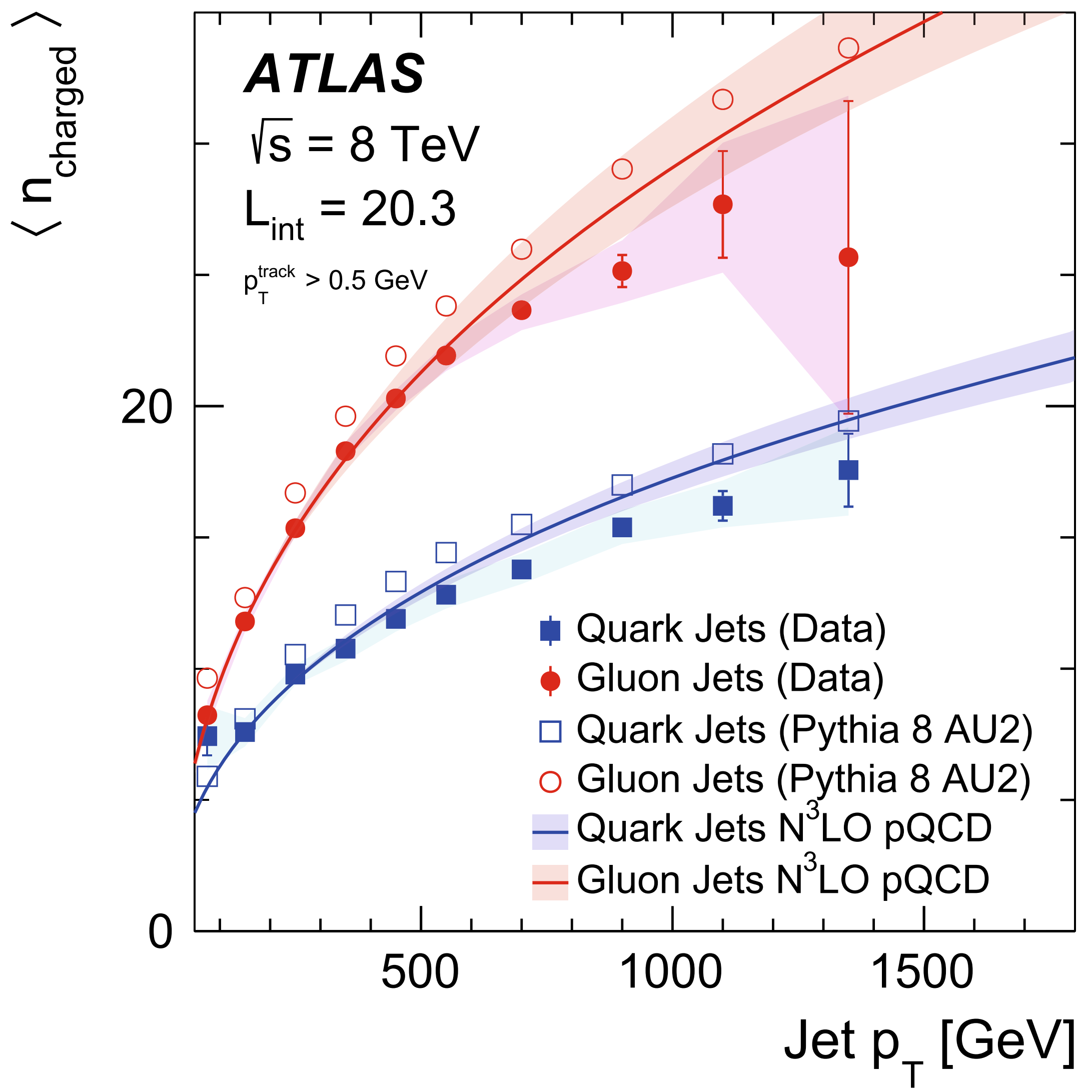


- Approximate particle content in a jet: $\pi^+ : \pi^- : \pi^0 = 1 : 1 : 1$ (+10% Kaons, Protons…)

[ATLAS, EPJC 76, 322 (2016)]

Charged particles in jets

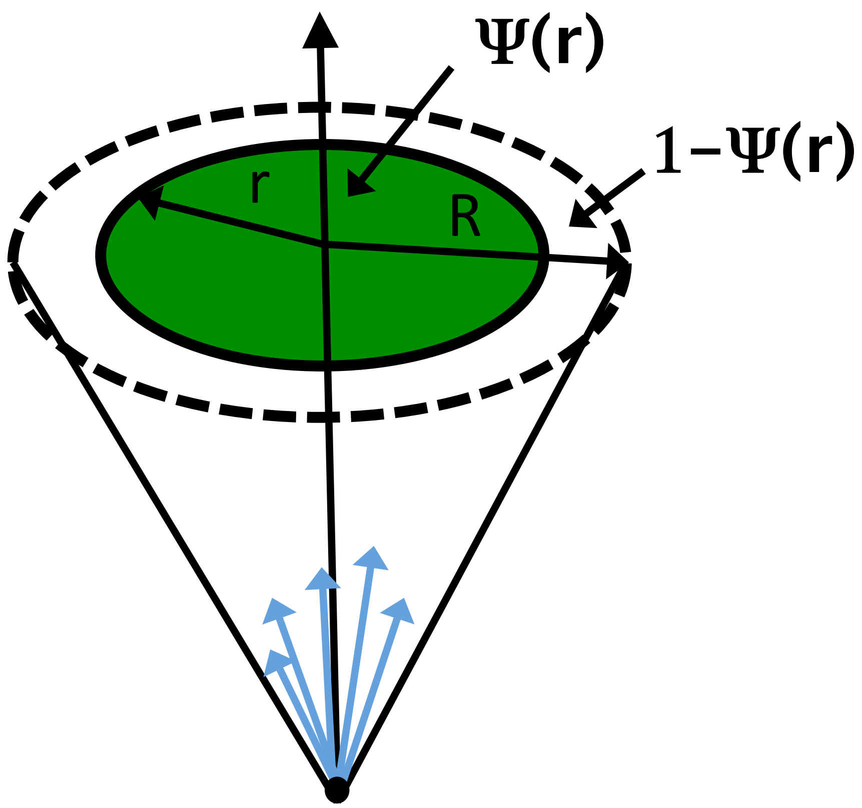




- Approximate particle content in a jet: $\pi^+ : \pi^- : \pi^0 = 1 : 1 : 1$ (+10% Kaons, Protons…)
- Gluon jets have higher multiplicity (colour factor C_A compared to C_F)

[ATLAS, EPJC 76, 322 (2016)]

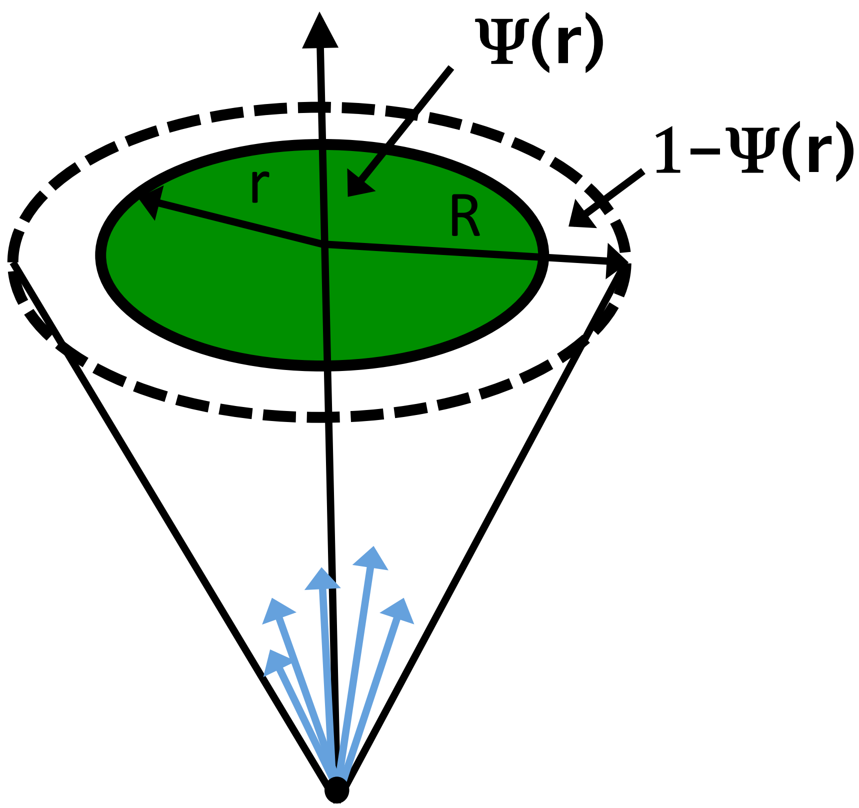
Jet substructure



Jet shape $\Psi(r) = \frac{\sum_{r_i<r} p_{\mathrm{T},i}}{\sum_{r_i<R} p_{\mathrm{T},i}}$

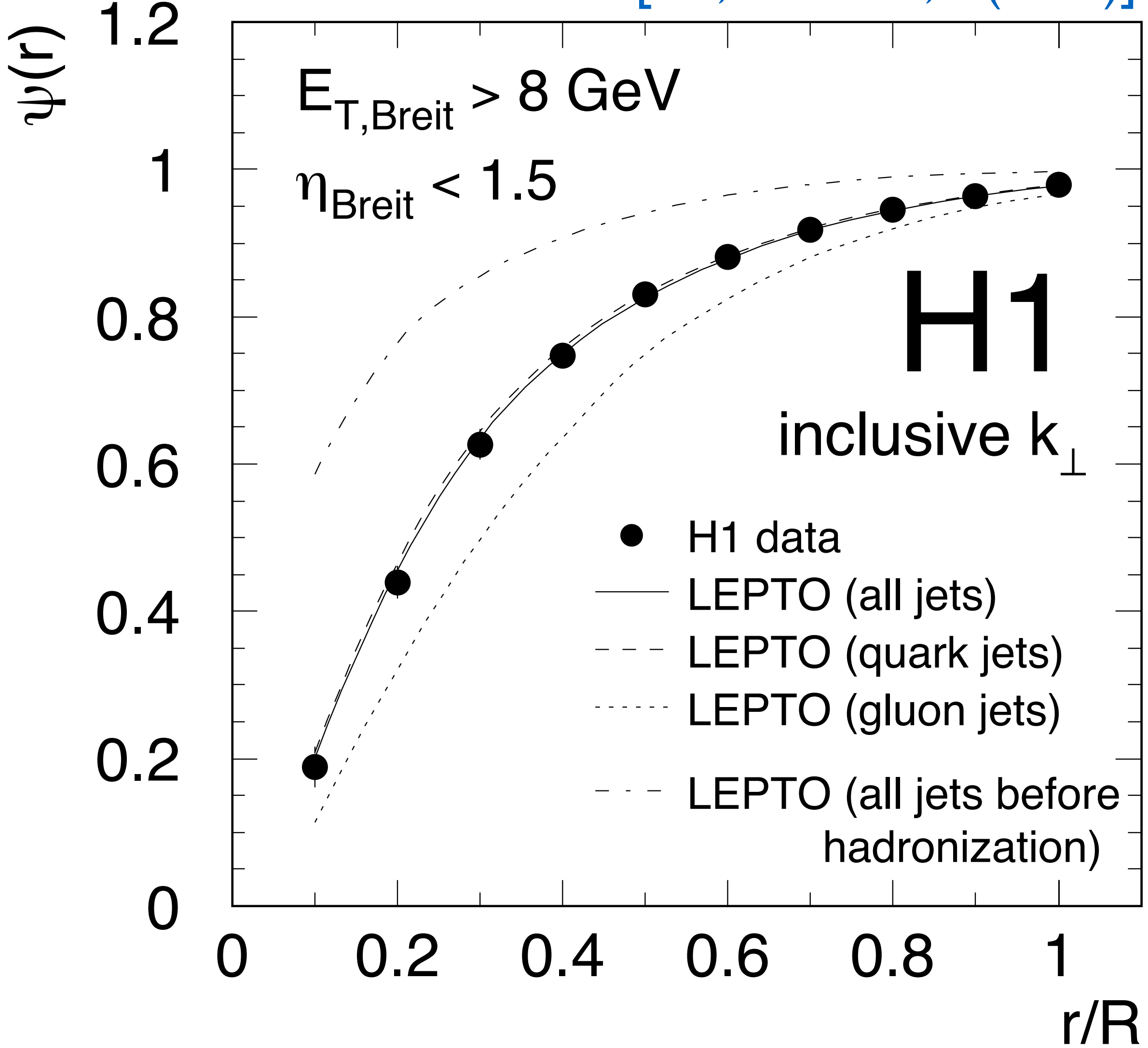
[CMS, JHEP 06, 160 (2012)]

Jet substructure



Jet shape $\Psi(r) = \frac{\sum_{r_i<r} p_{T,i}}{\sum_{r_i<R} p_{T,i}}$

[H1, NPB 545, 3 (1999)]

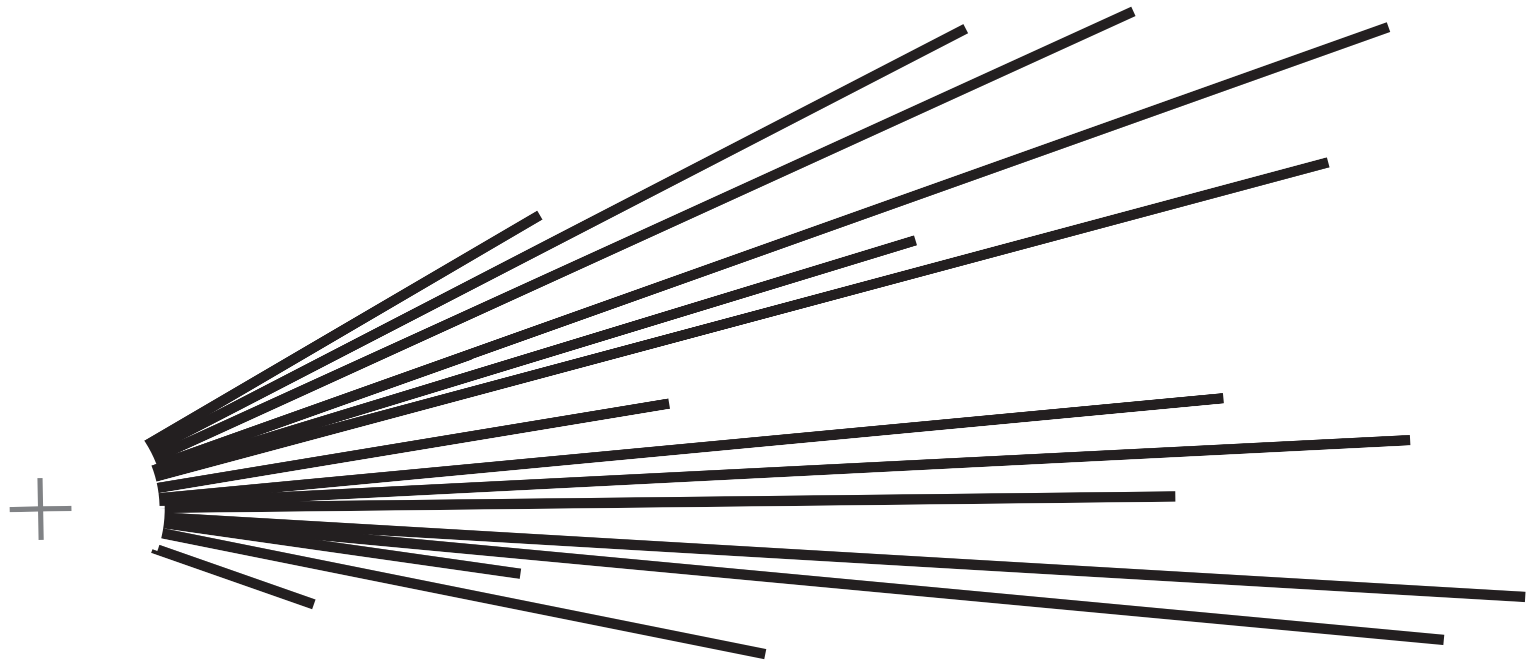


- Quark jets are narrower than gluon jets

[CMS, JHEP 06, 160 (2012)]

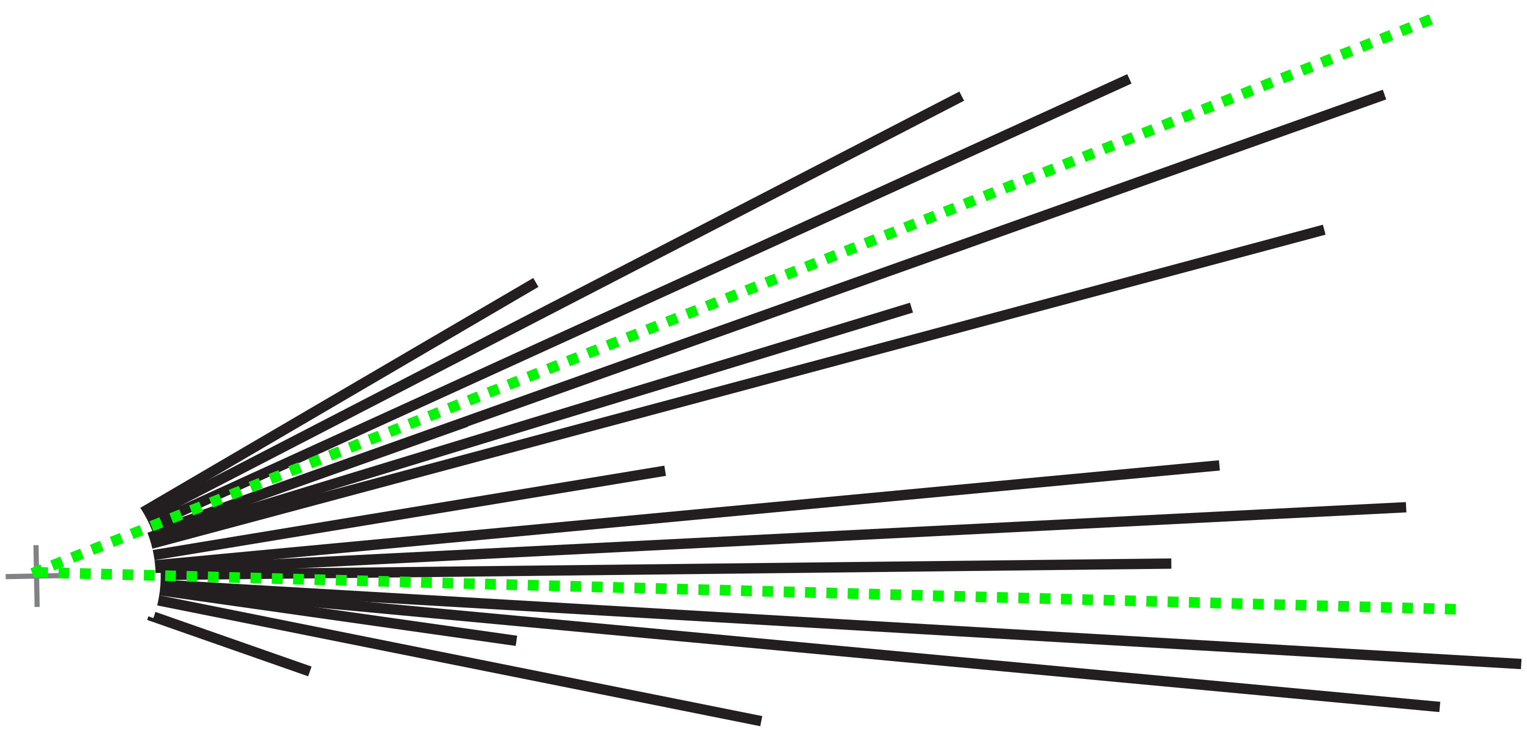
Jet substructure today

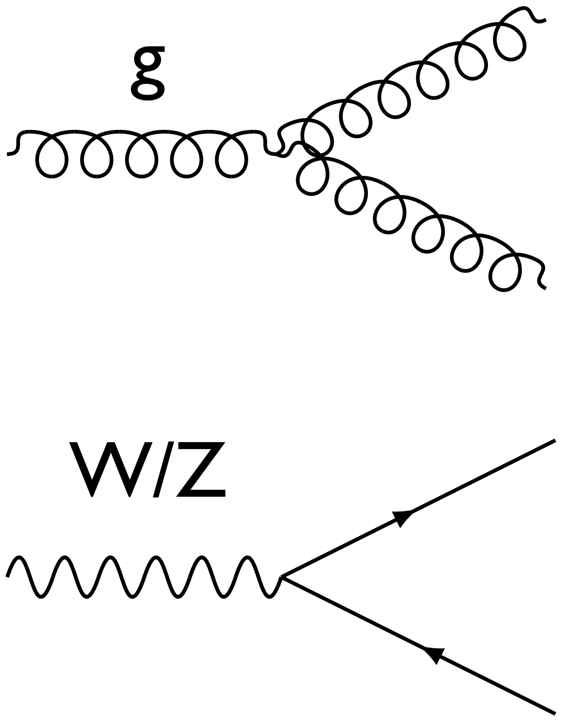
- Remove unwanted / soft radiation from jets
- Aid the jet reconstruction and calibration
- Distinguish quark/gluon jets
- Tagging of fully merged W, Z, H and top jets



Jet substructure today

- 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



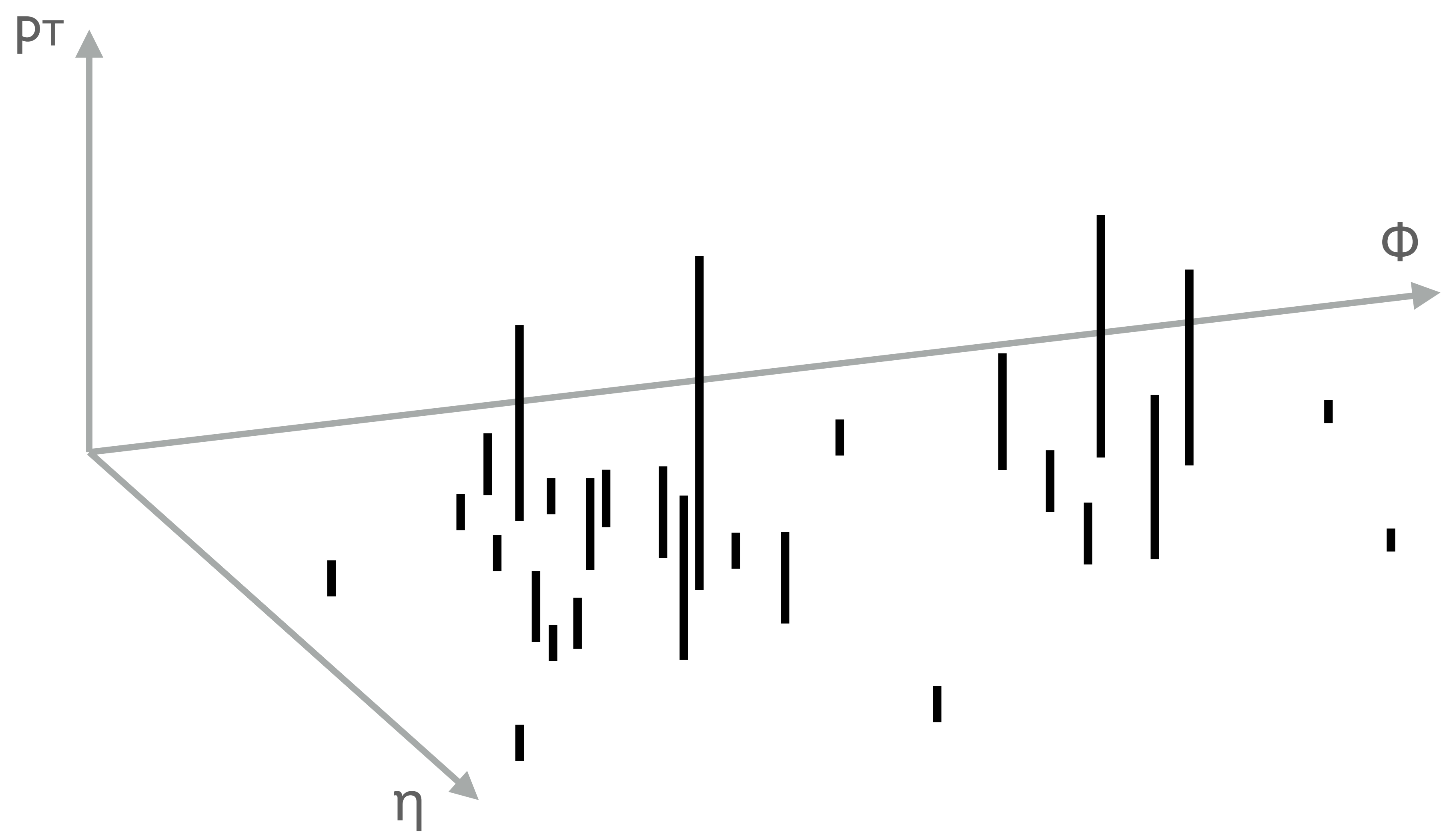


Jet algorithms and substructure

Jet clustering (anti-k_T)

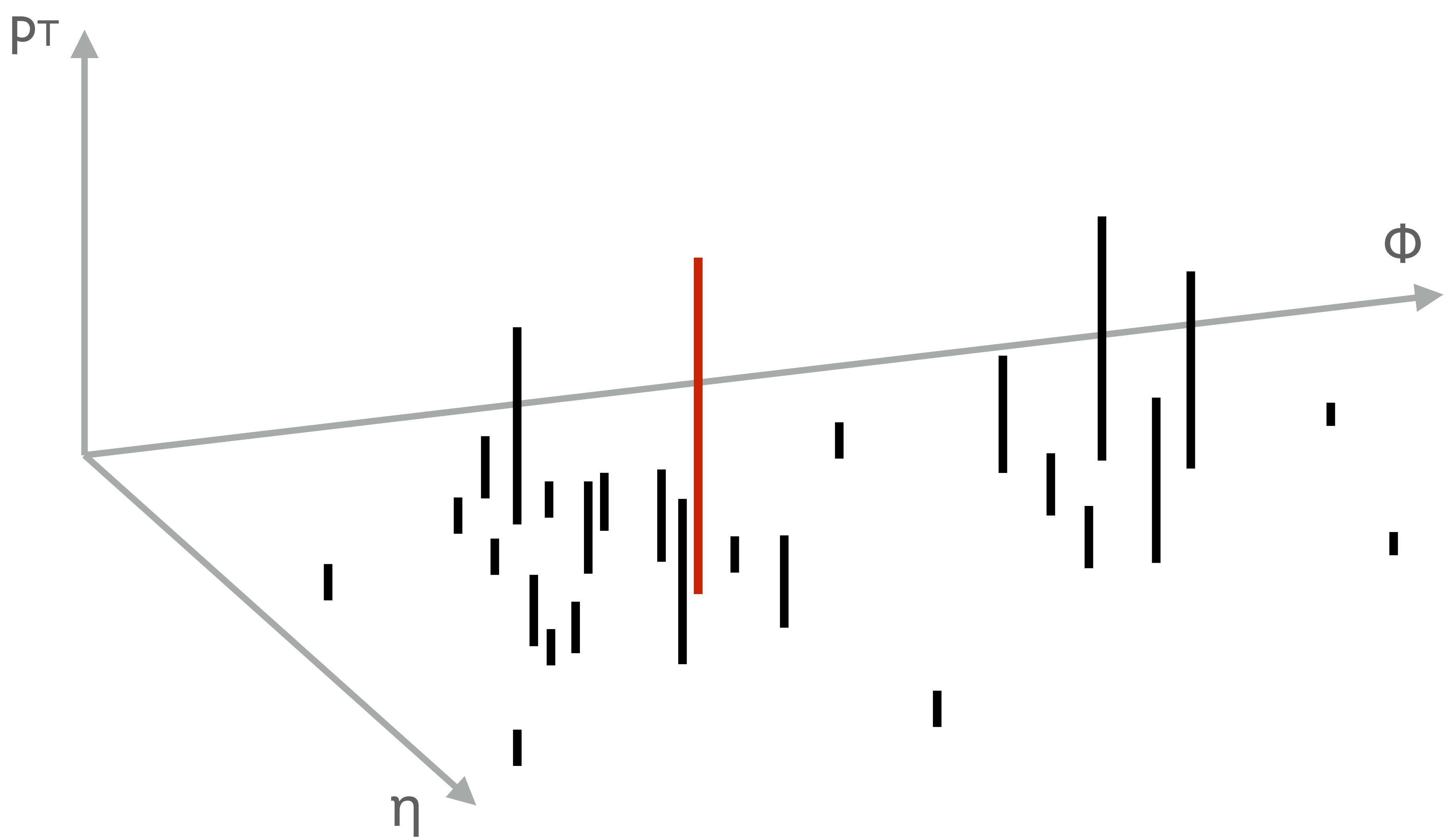
$$d_{ij} = \min(p_{t,i}^{2k}, p_{t,j}^{2k}) \frac{\Delta R^2}{R^2} \quad \text{with } k = -1, \;\; \Delta R^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

$$d_{iB} = p_{t,i}^{2k}$$



Jet clustering (anti-k_T)

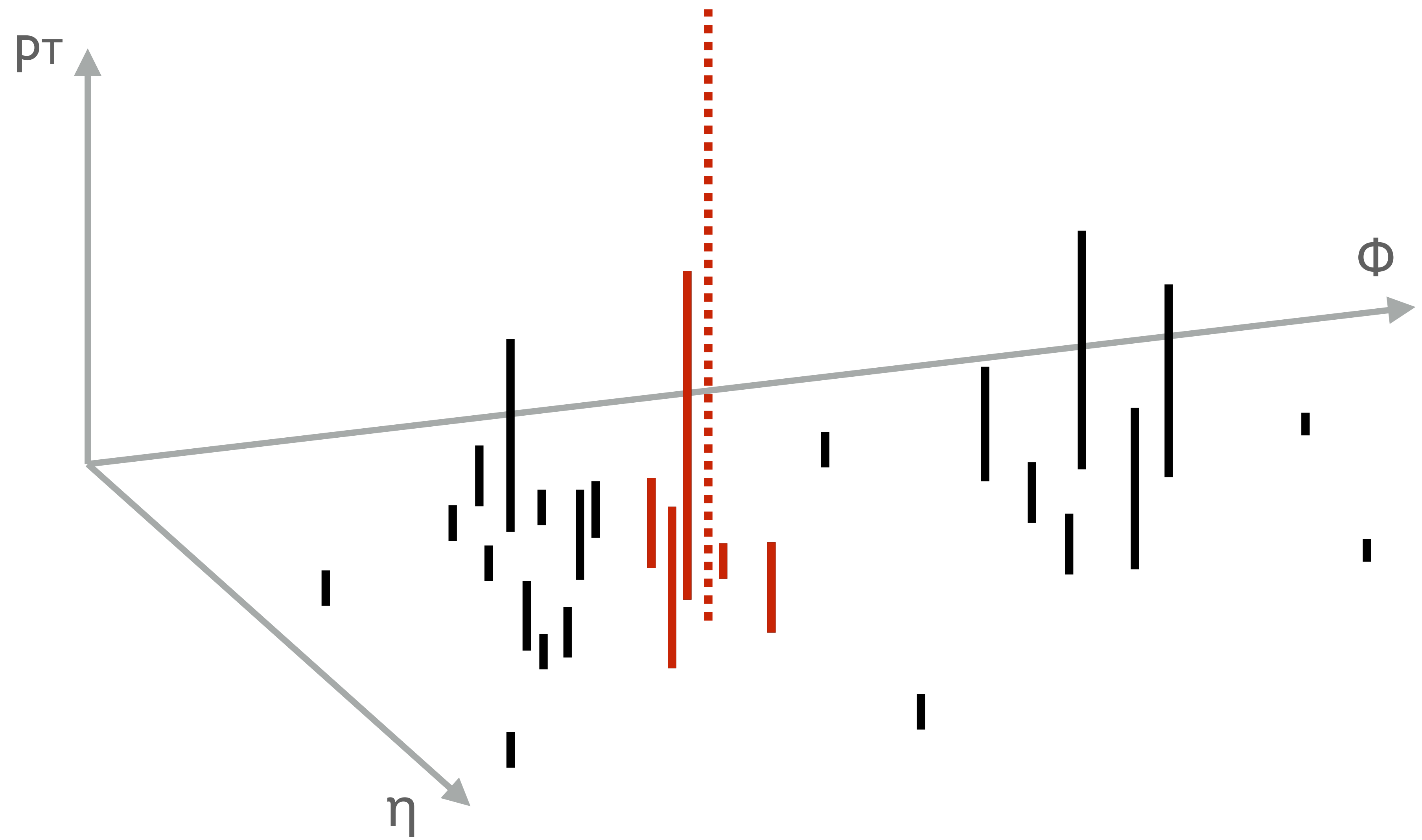
$$d_{ij} = \min(p_{t,i}^{2k}, p_{t,j}^{2k}) \frac{\Delta R^2}{R^2} \text{ with } k = -1, \ \Delta R^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$
$$d_{iB} = p_{t,i}^{2k}$$



Jet clustering (anti-k_T)

$$d_{ij} = \min(p_{t,i}^{2k}, p_{t,j}^{2k}) \frac{\Delta R^2}{R^2} \quad \text{with } k=-1,\ \ \Delta R^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

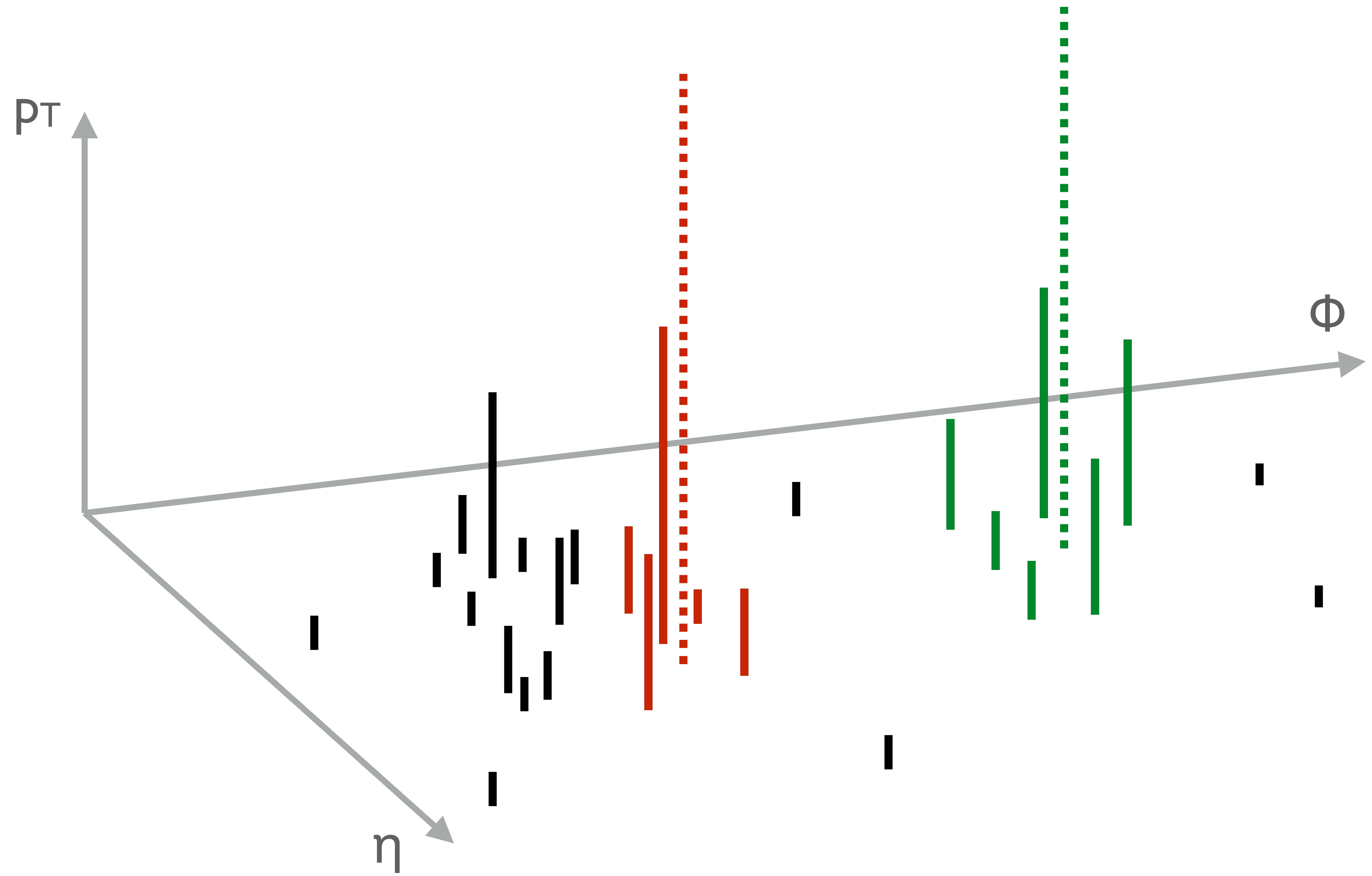
$$d_{iB} = p_{t,i}^{2k}$$



Jet clustering (anti-k_T)

$$d_{ij} = \min(p_{t,i}^{2k}, p_{t,j}^{2k}) \frac{\Delta R^2}{R^2} \quad \text{with} \quad k=-1, \quad \Delta R^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

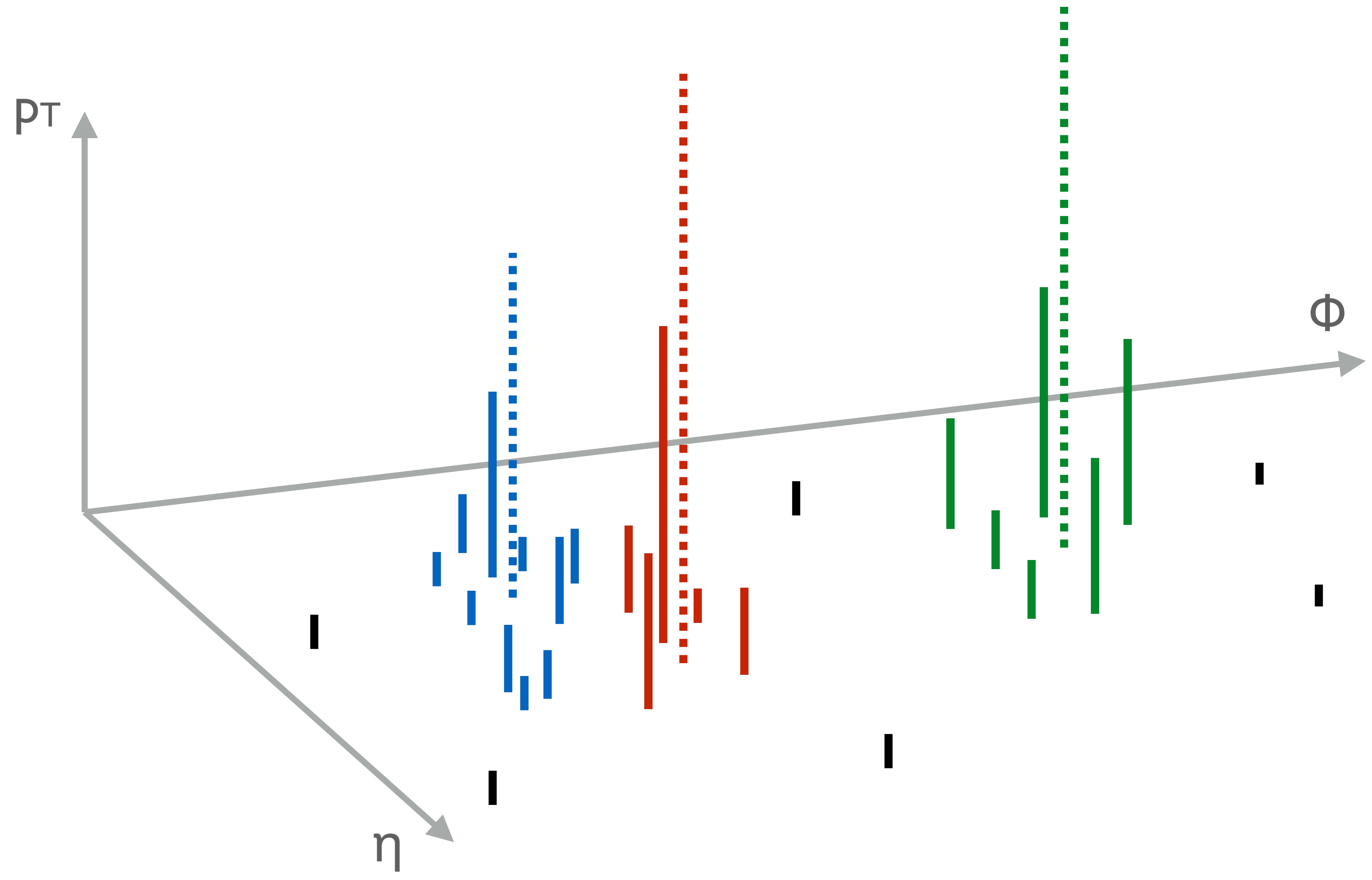
$$d_{iB} = p_{t,i}^{2k}$$



Jet clustering (anti-k_T)

$$d_{ij} = \min(p_{t,i}^{2k}, p_{t,j}^{2k}) \frac{\Delta R^2}{R^2} \quad \text{with } k = -1, \quad \Delta R^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

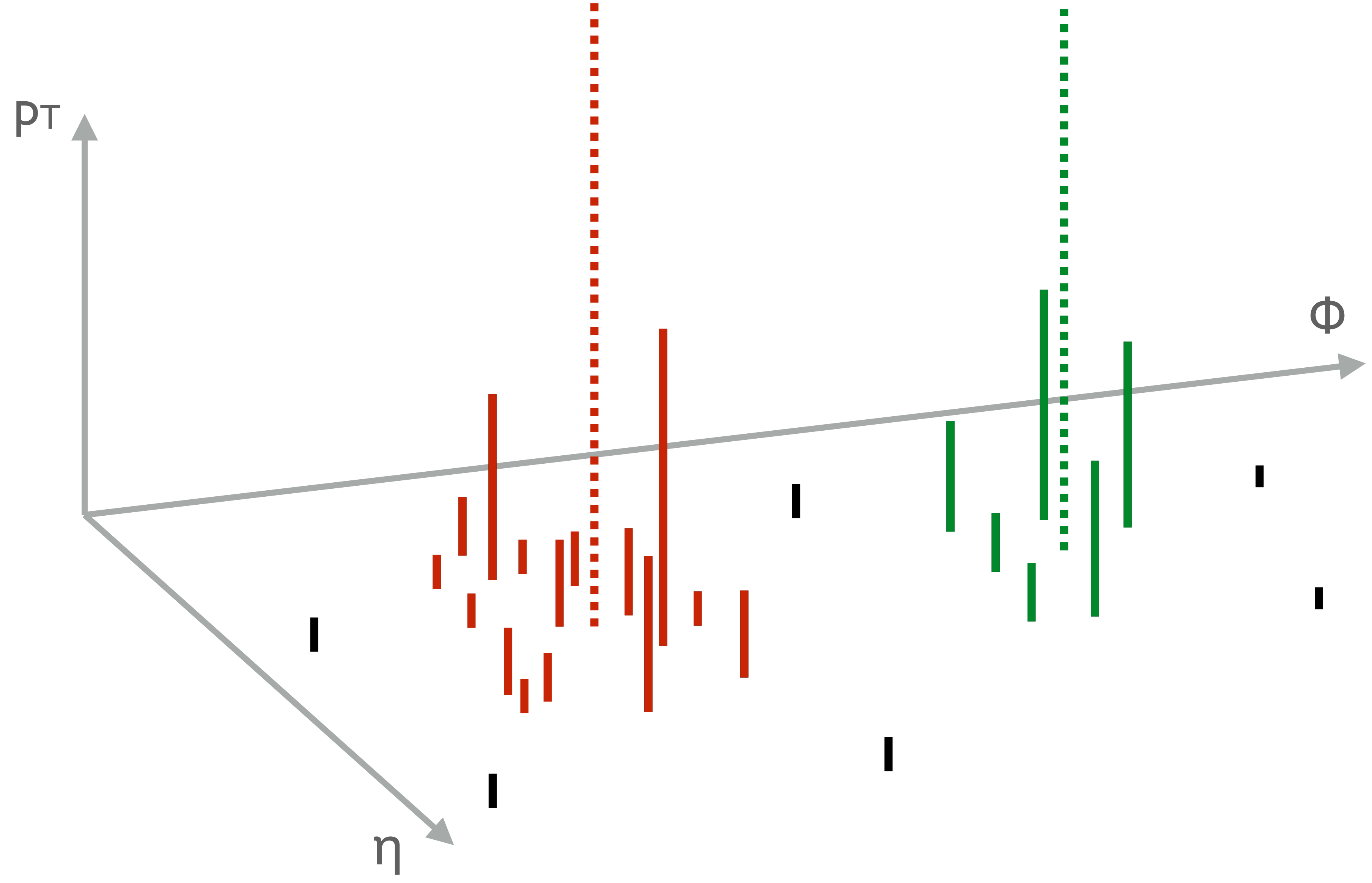
$$d_{iB} = p_{t,i}^{2k}$$



Jet clustering (anti-k_T)

$$d_{ij} = \min(p_{t,i}^{2k}, p_{t,j}^{2k}) \frac{\Delta R^2}{R^2}$$ with $$\Delta R^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

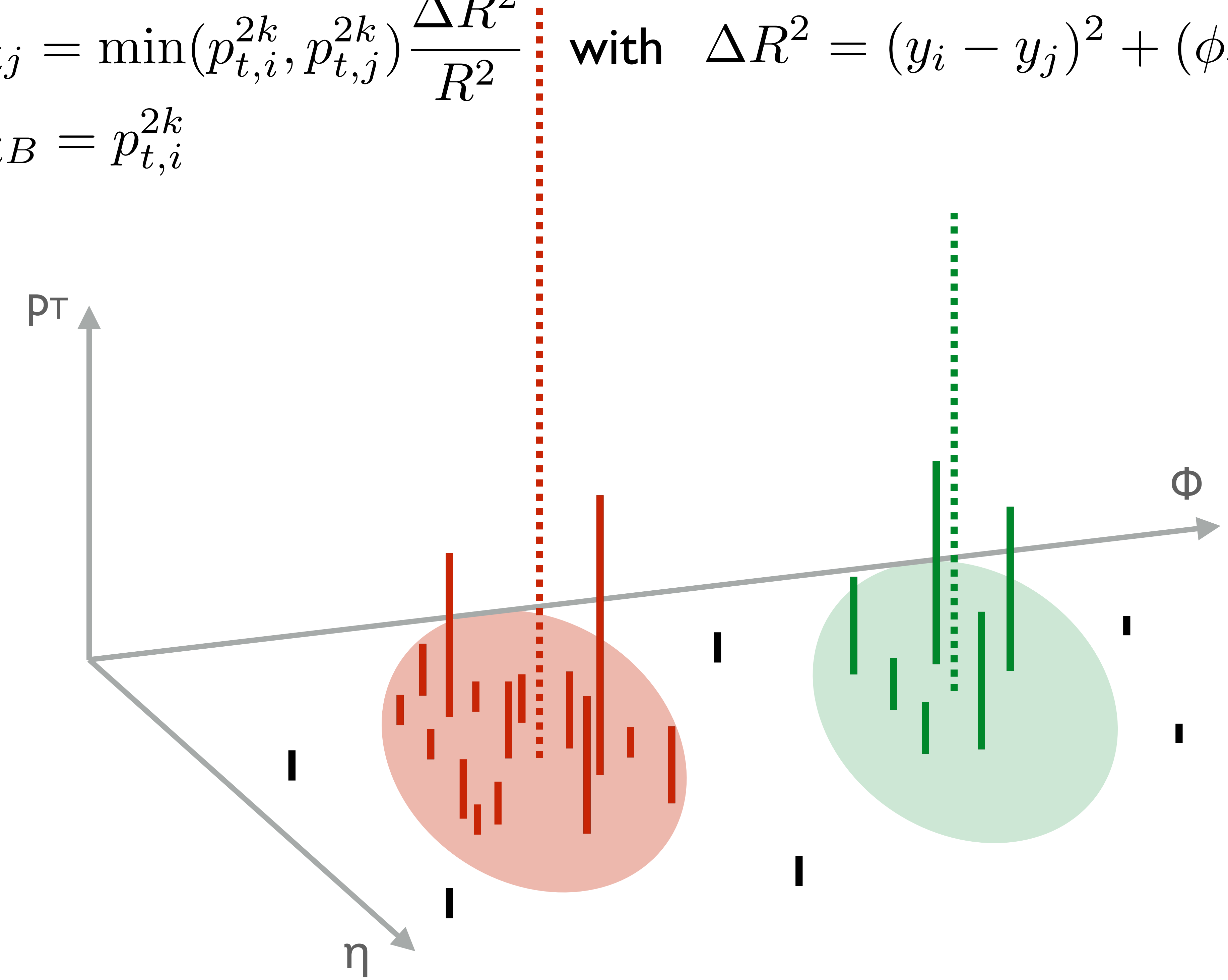
$$d_{iB} = p_{t,i}^{2k}$$



Jet clustering (anti-k_T)

$$d_{ij} = \min(p_{t,i}^{2k}, p_{t,j}^{2k}) \frac{\Delta R^2}{R^2} \quad \text{with} \quad \Delta R^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

$$d_{iB} = p_{t,i}^{2k}$$

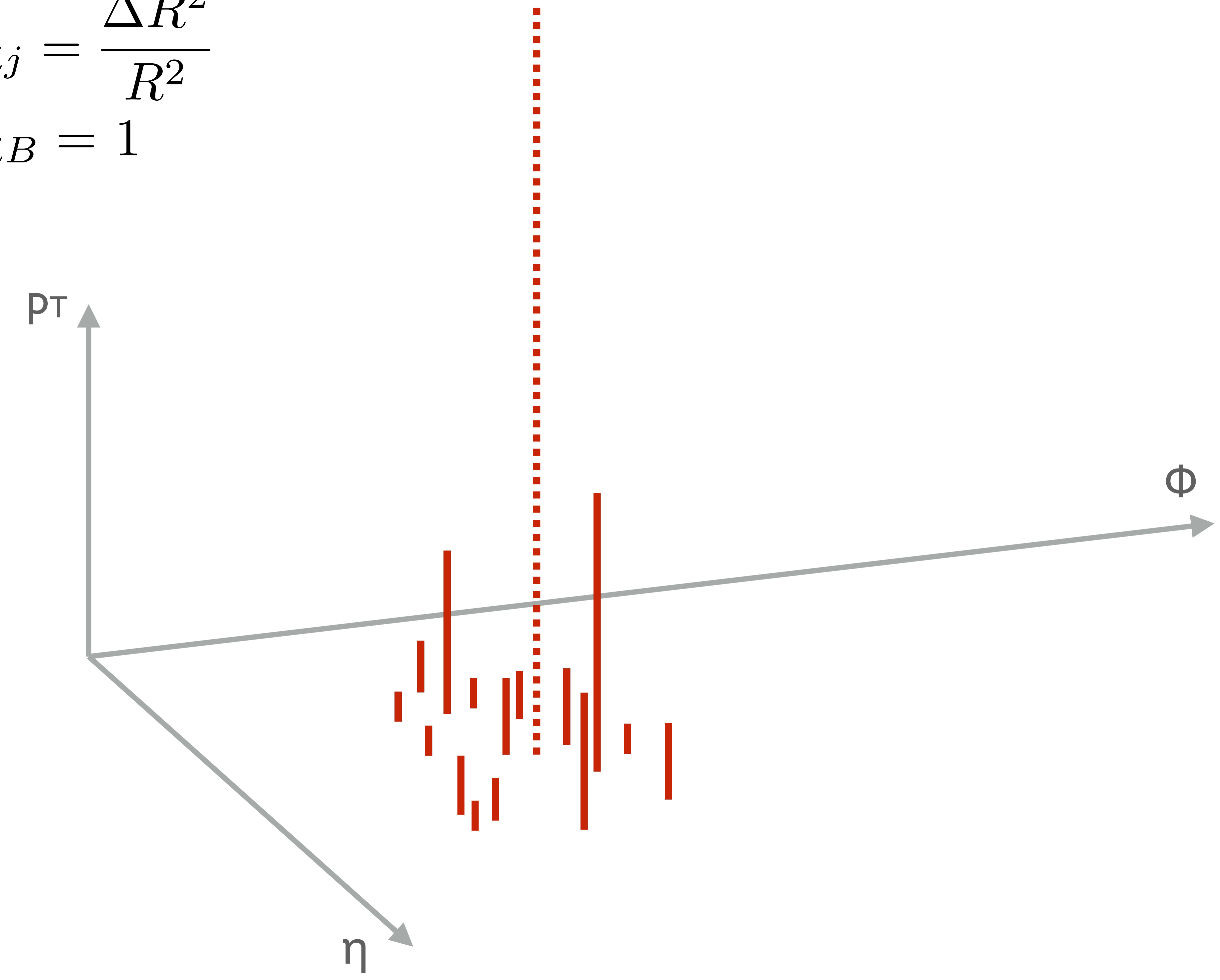


Jet clustering (C/A)

$$d_{ij} = \frac{\Delta R^2}{R^2}$$
$$d_{iB} = 1$$

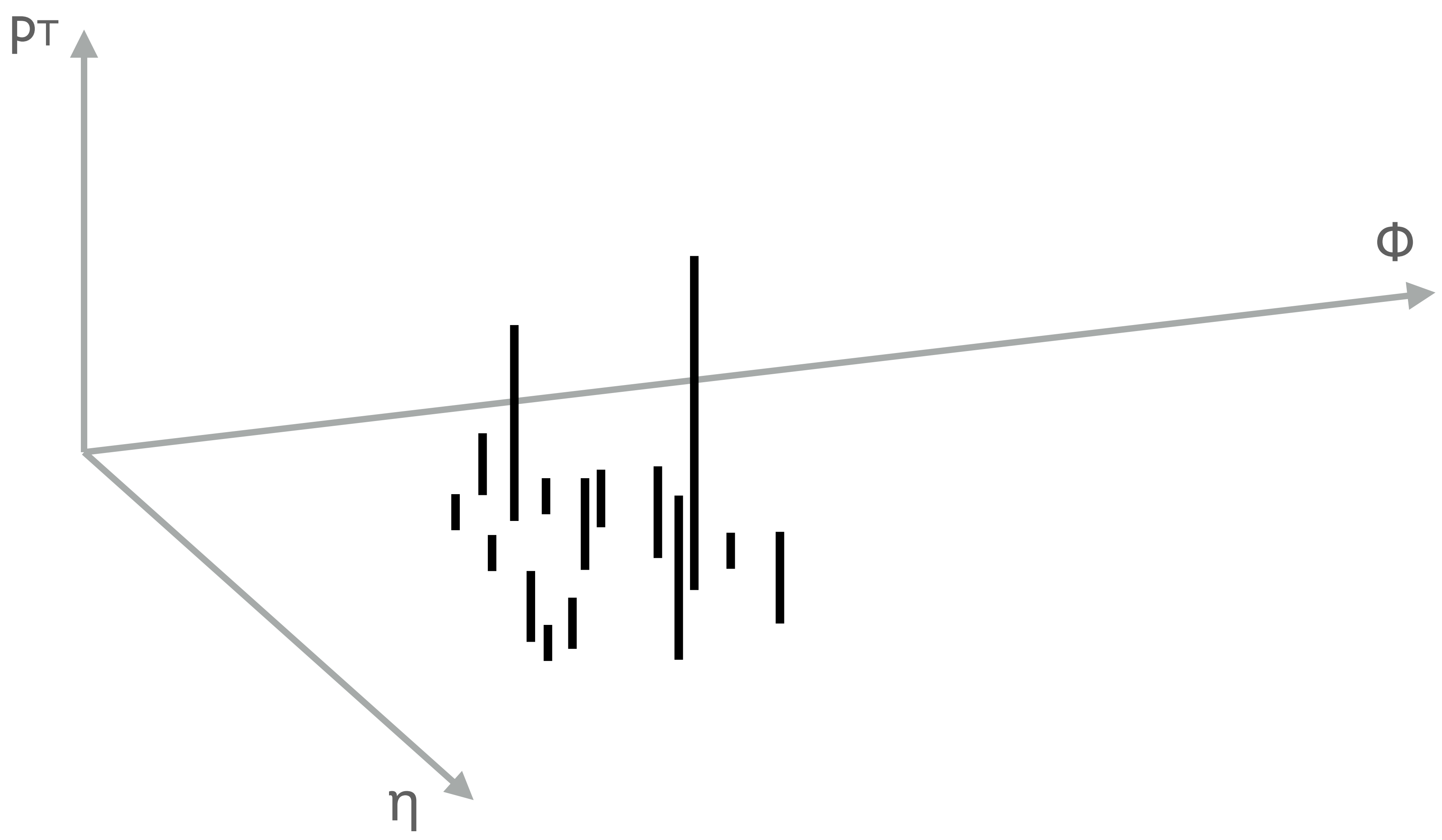
Jet clustering (C/A)

$$d_{ij} = \frac{\Delta R^2}{R^2}$$
$$d_{iB} = 1$$



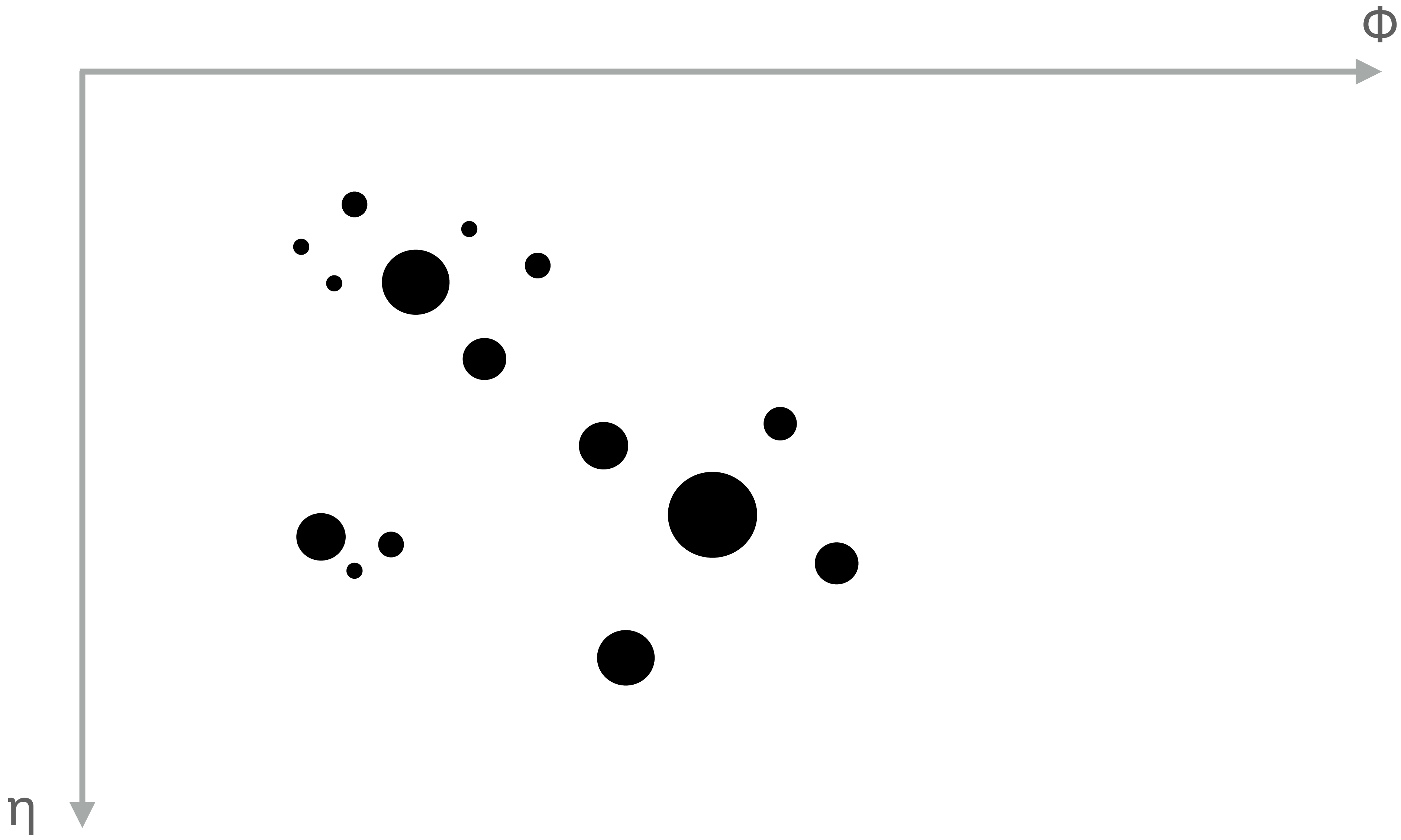
Jet clustering (C/A)

$$d_{ij} = \frac{\Delta R^2}{R^2}$$
$$d_{iB} = 1$$



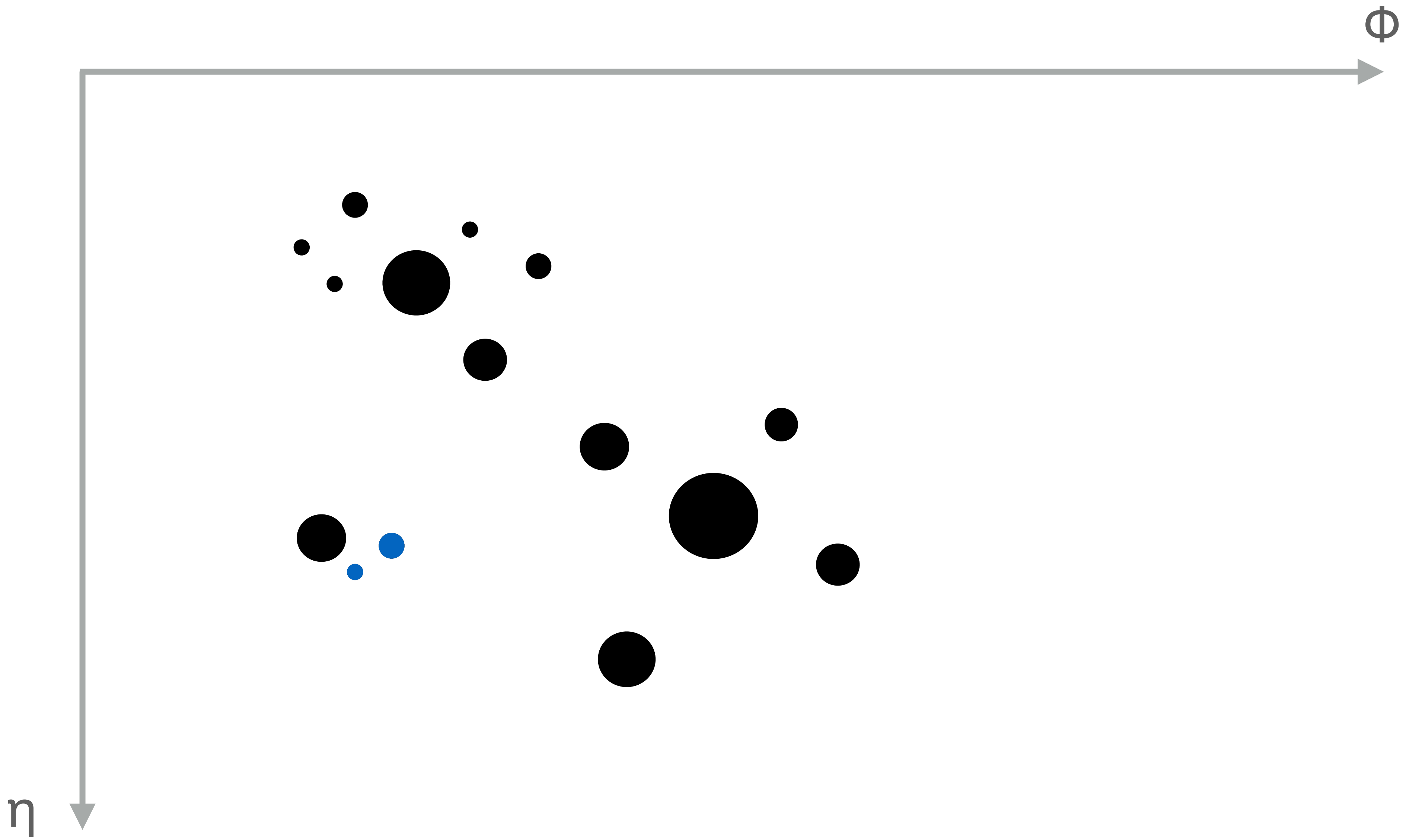
Jet clustering (C/A)

$$d_{ij} = \frac{\Delta R^2}{R^2}$$
$$d_{iB} = 1$$



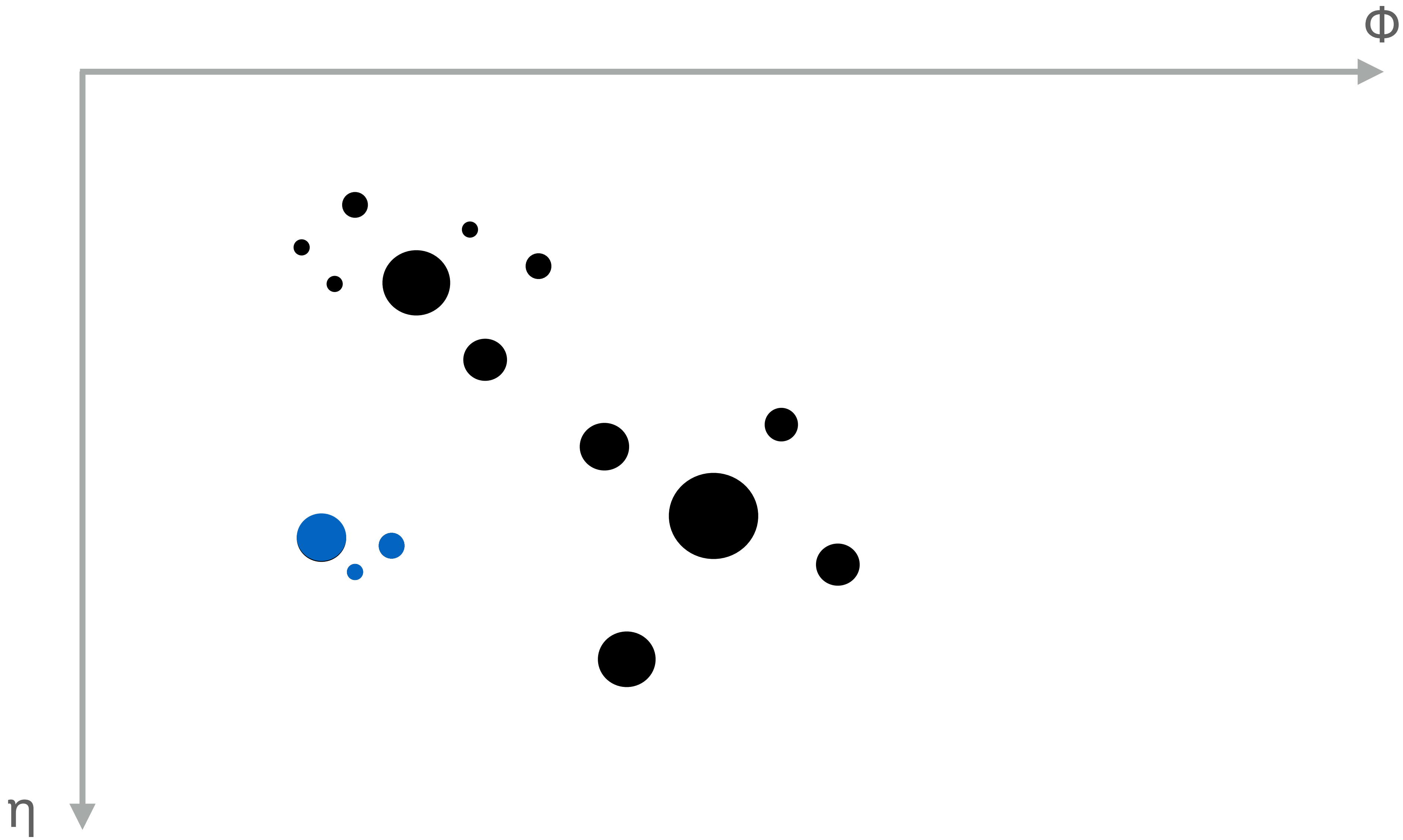
Jet clustering (C/A)

$$d_{ij} = \frac{\Delta R^2}{R^2}$$
$$d_{iB} = 1$$



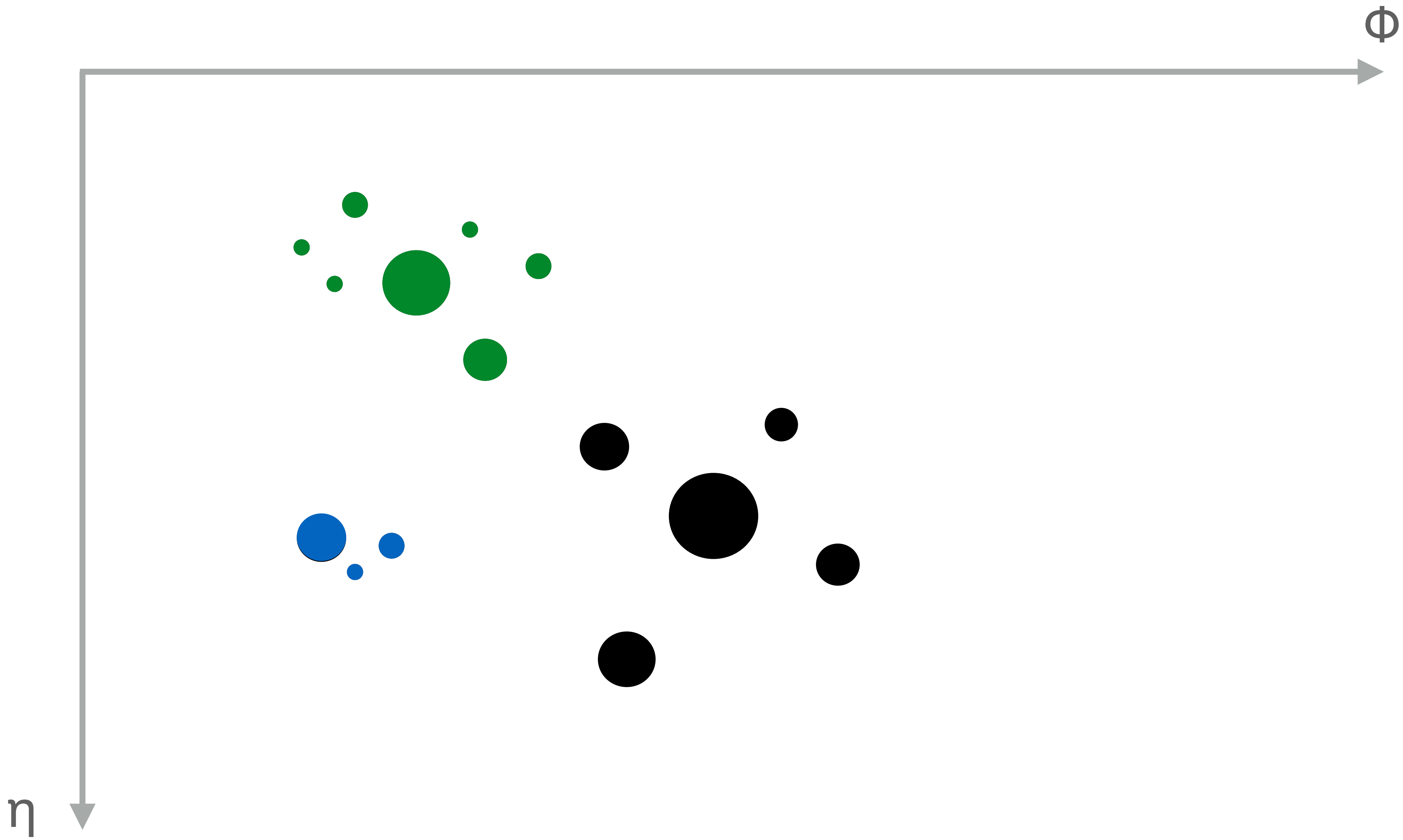
Jet clustering (C/A)

$$d_{ij} = \frac{\Delta R^2}{R^2}$$
$$d_{iB} = 1$$



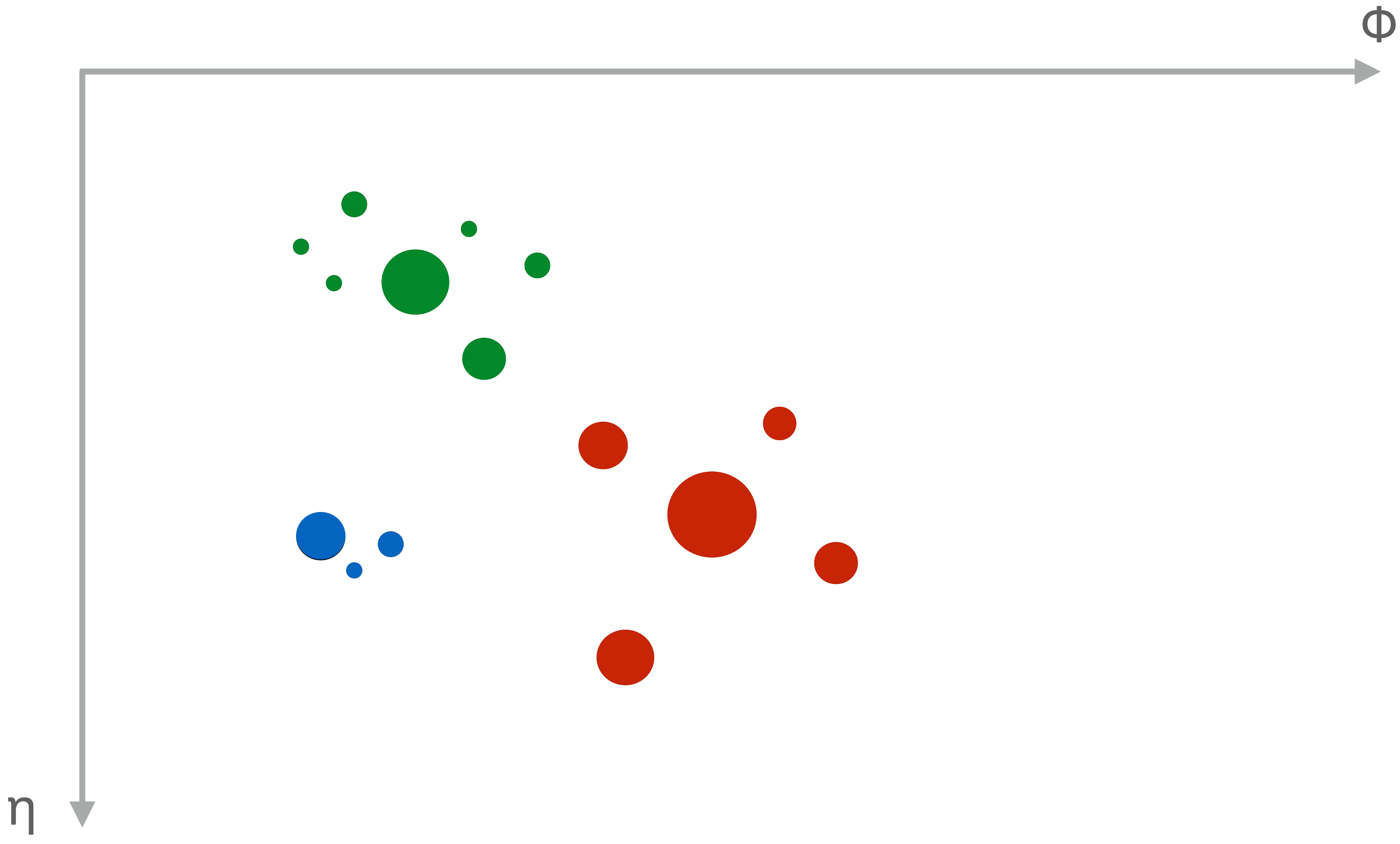
Jet clustering (C/A)

$$d_{ij} = \frac{\Delta R^2}{R^2}$$
$$d_{iB} = 1$$



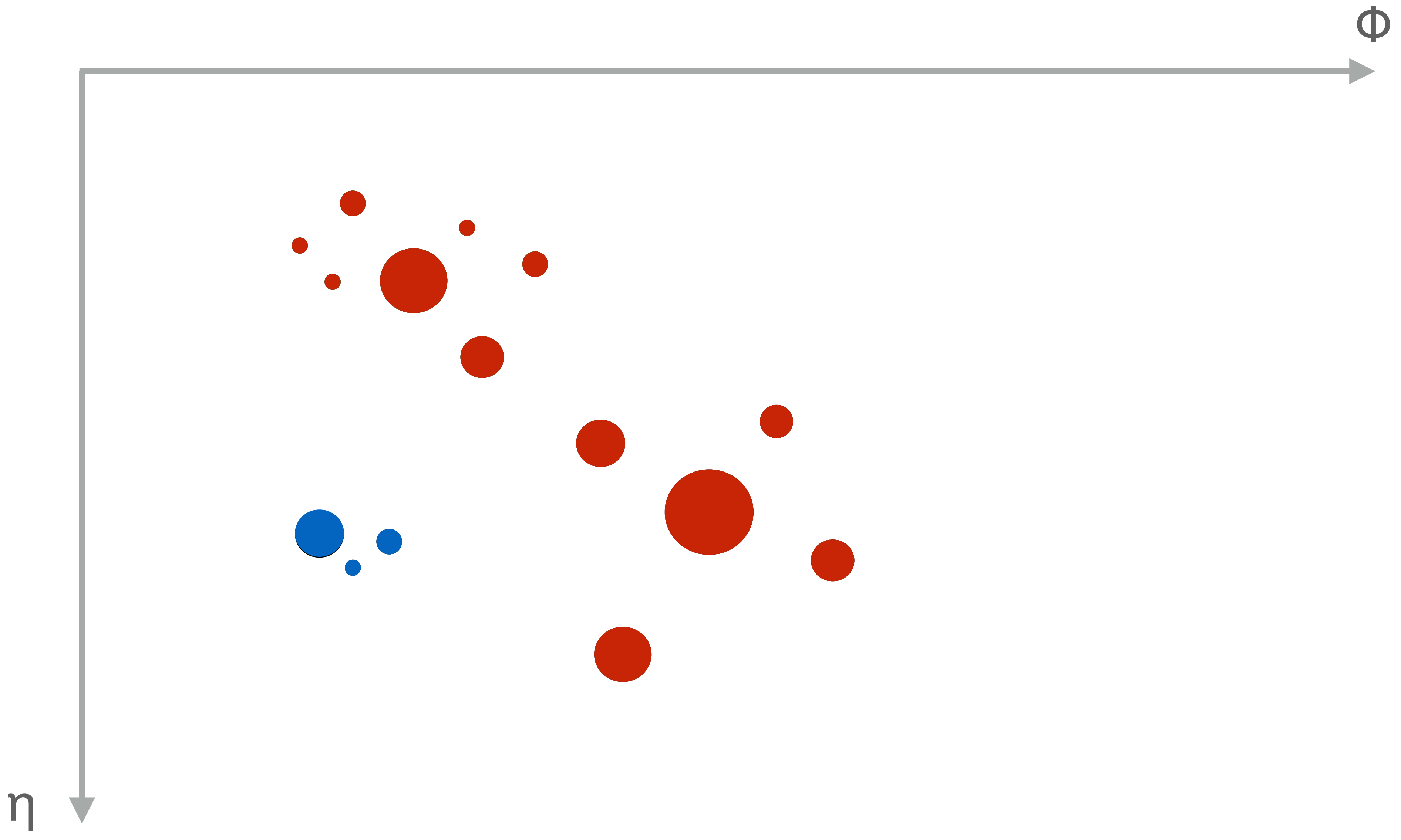
Jet clustering (C/A)

$$d_{ij} = \frac{\Delta R^2}{R^2}$$
$$d_{iB} = 1$$



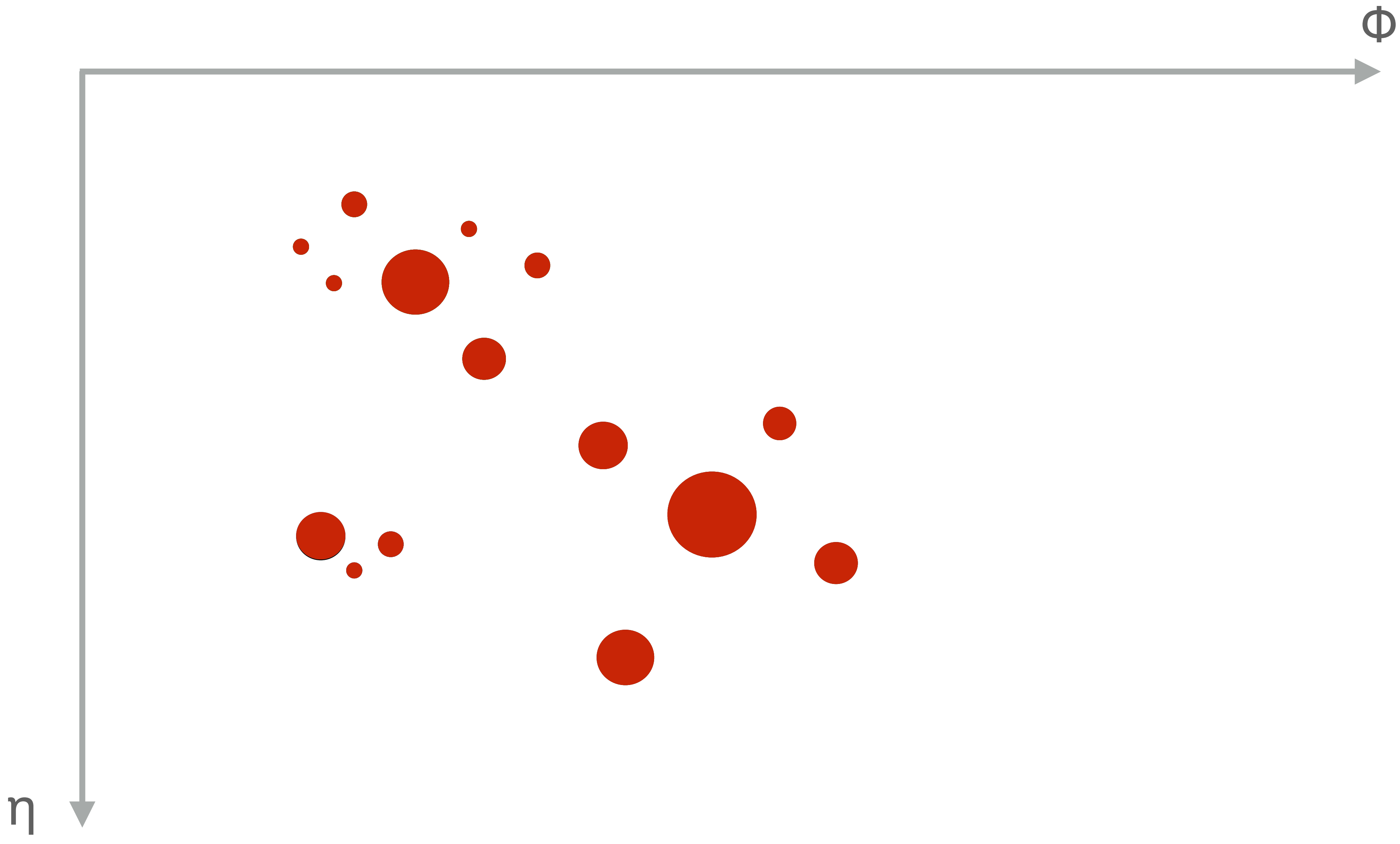
Jet clustering (C/A)

$$d_{ij} = \frac{\Delta R^2}{R^2}$$
$$d_{iB} = 1$$



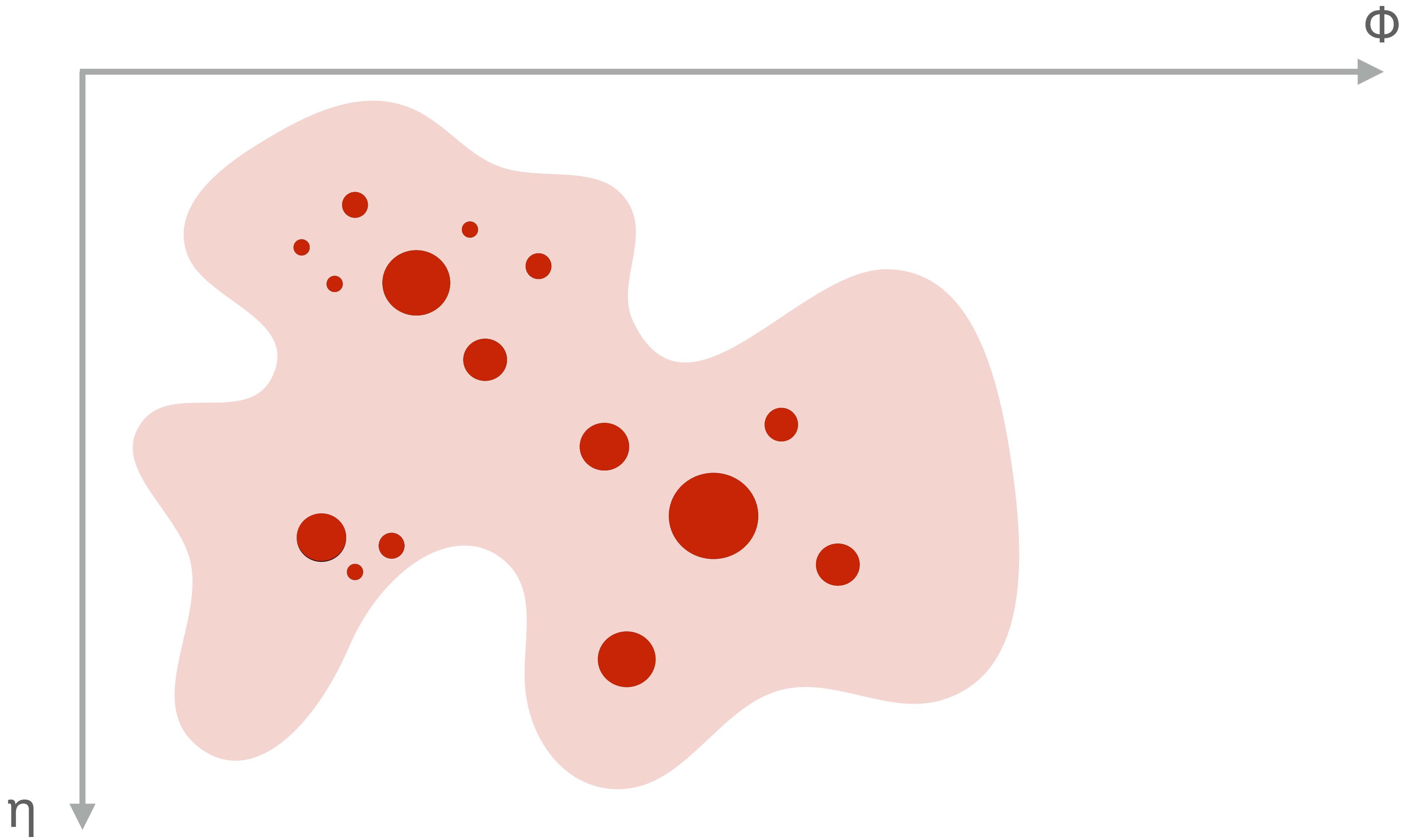
Jet clustering (C/A)

$$d_{ij} = \frac{\Delta R^2}{R^2}$$
$$d_{iB} = 1$$



Jet clustering (C/A)

$$d_{ij} = \frac{\Delta R^2}{R^2}$$
$$d_{iB} = 1$$



Jet grooming (soft drop)

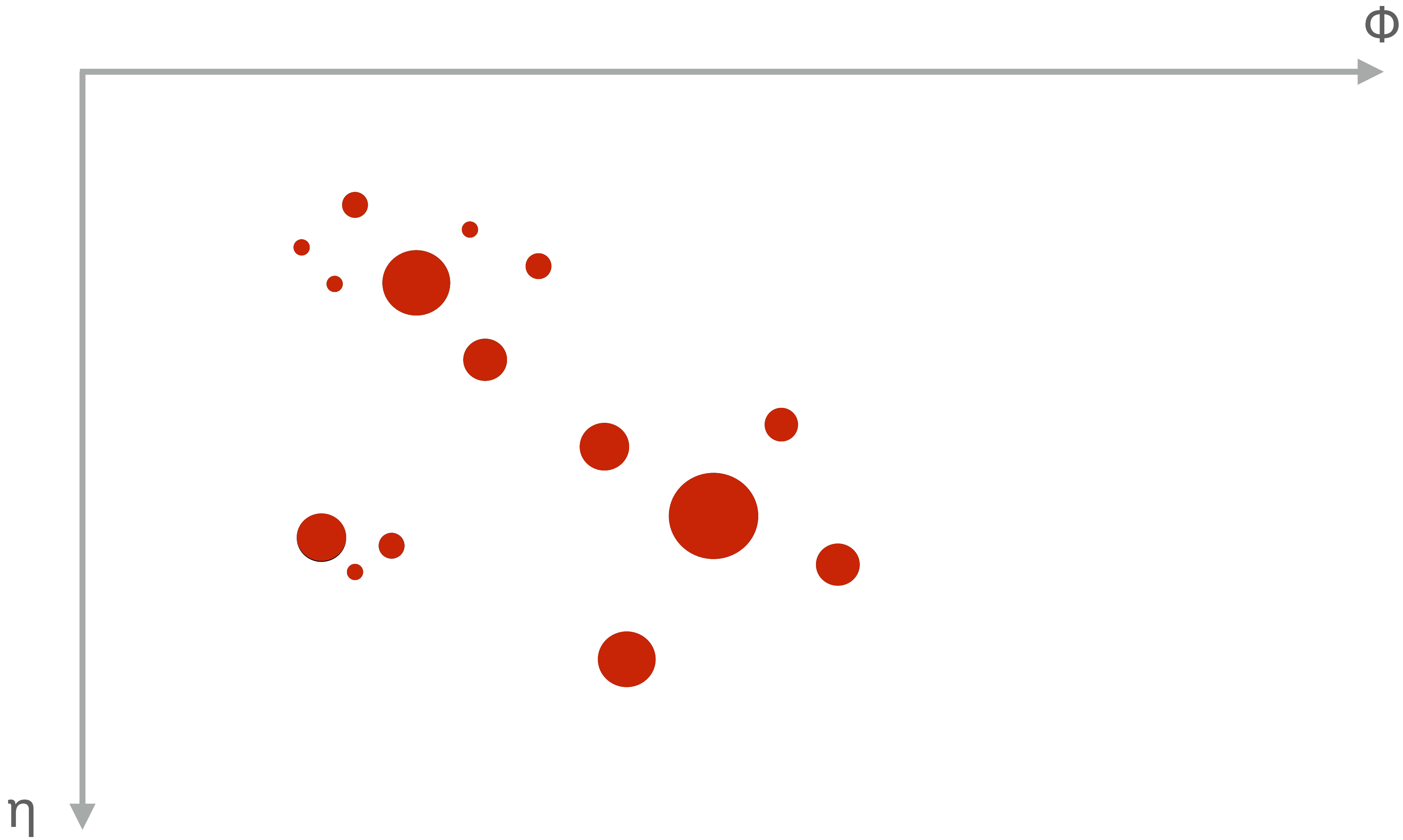
Invert clustering history with a veto condition:

$$\min(p_{\mathrm{T,i}}, p_{\mathrm{T,j}}) > z_{\mathrm{cut}} \left(p_{\mathrm{T,i}} + p_{\mathrm{T,j}}\right) \left(\frac{\Delta R_{ij}}{R}\right)^{\beta}$$

Jet grooming (soft drop)

Invert clustering history with a veto condition:

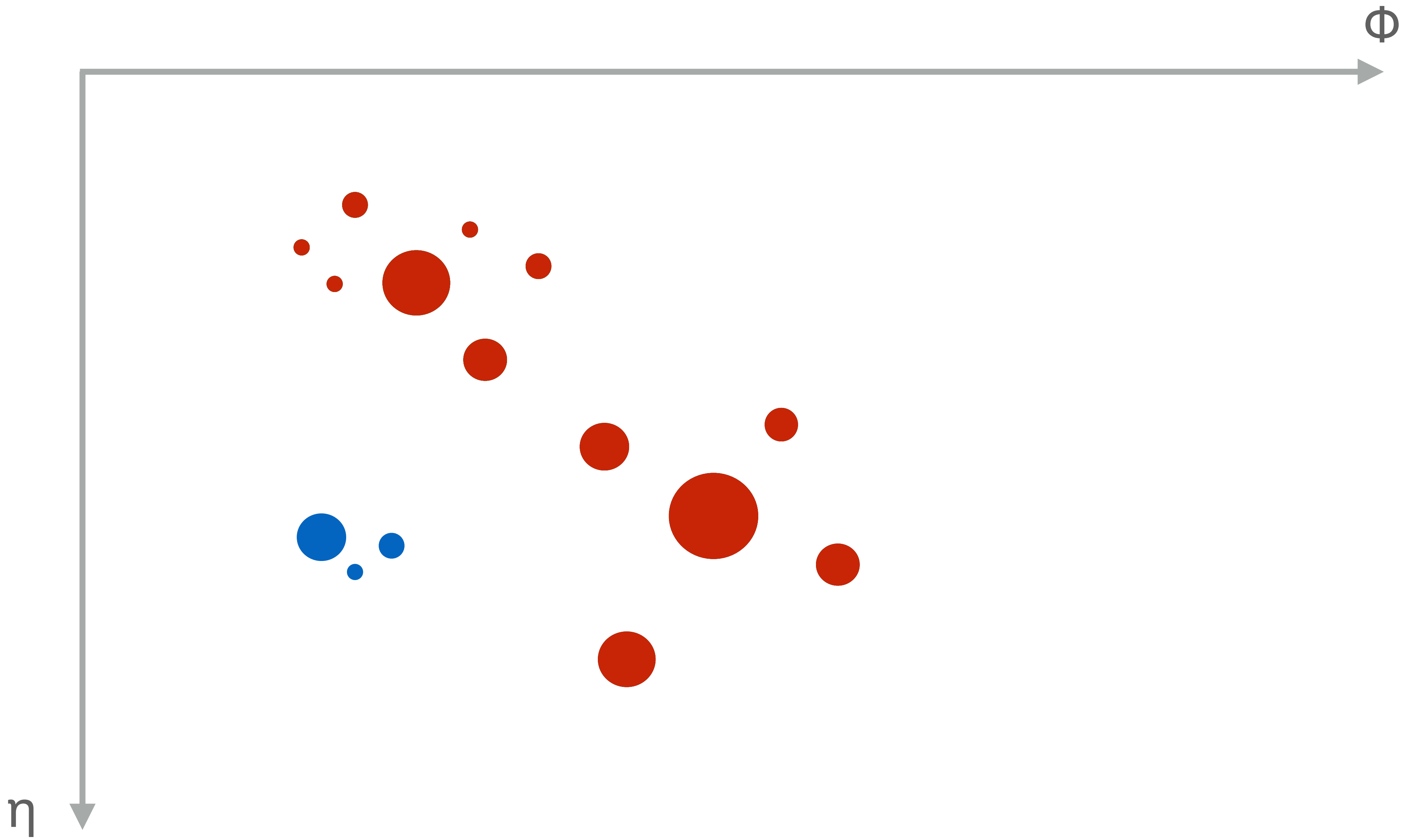
$$\min(p_{\mathrm{T,i}}, p_{\mathrm{T,j}}) > z_{\mathrm{cut}}\,(p_{\mathrm{T,i}} + p_{\mathrm{T,j}}) \left(\frac{\Delta R_{ij}}{R}\right)^{\beta}$$



Jet grooming (soft drop)

Invert clustering history with a veto condition:

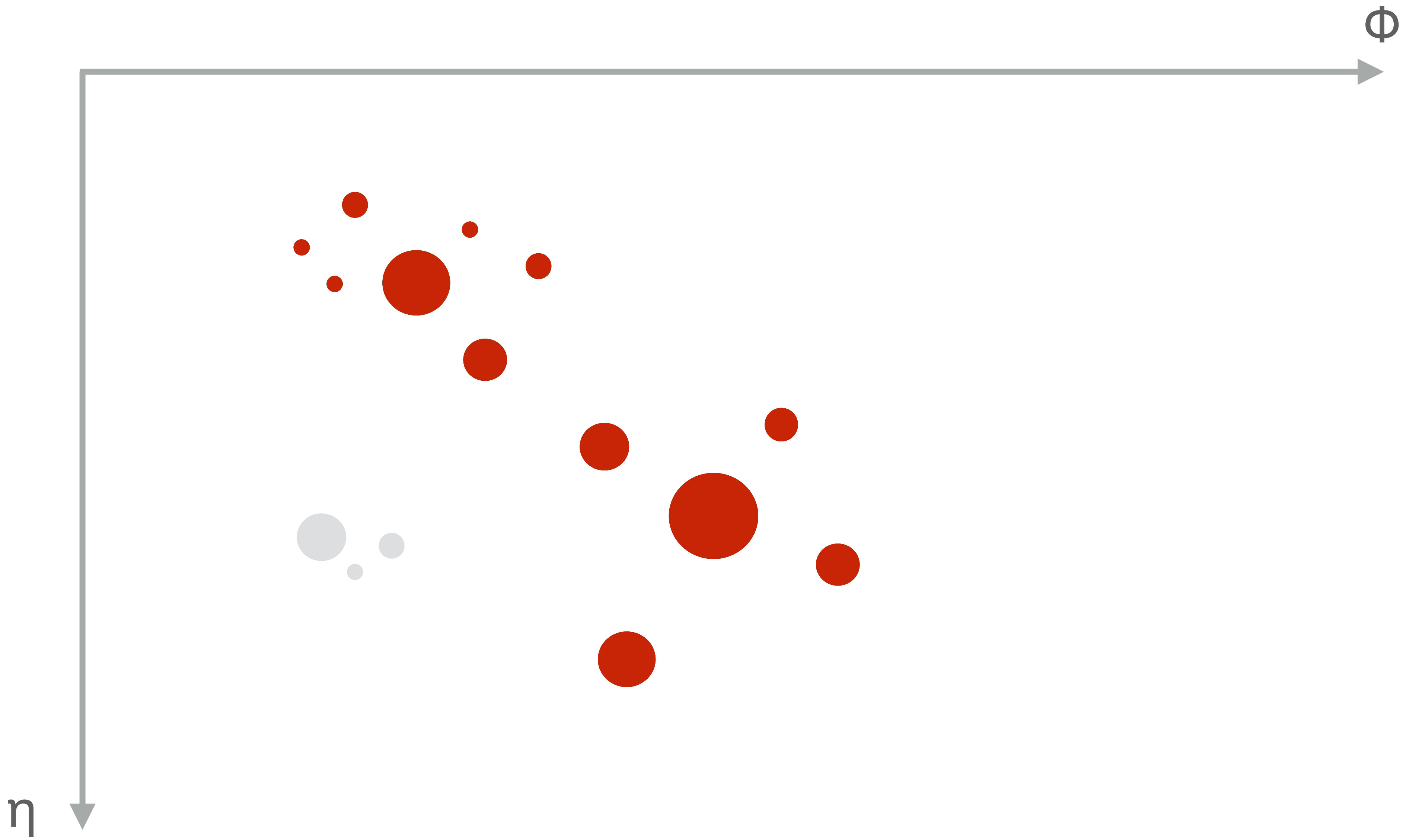
$$\min(p_{\mathrm{T,i}}, p_{\mathrm{T,j}}) > z_{\mathrm{cut}} \, (p_{\mathrm{T,i}} + p_{\mathrm{T,j}}) \left(\frac{\Delta R_{ij}}{R} \right)^{\beta}$$



Jet grooming (soft drop)

Invert clustering history with a veto condition:

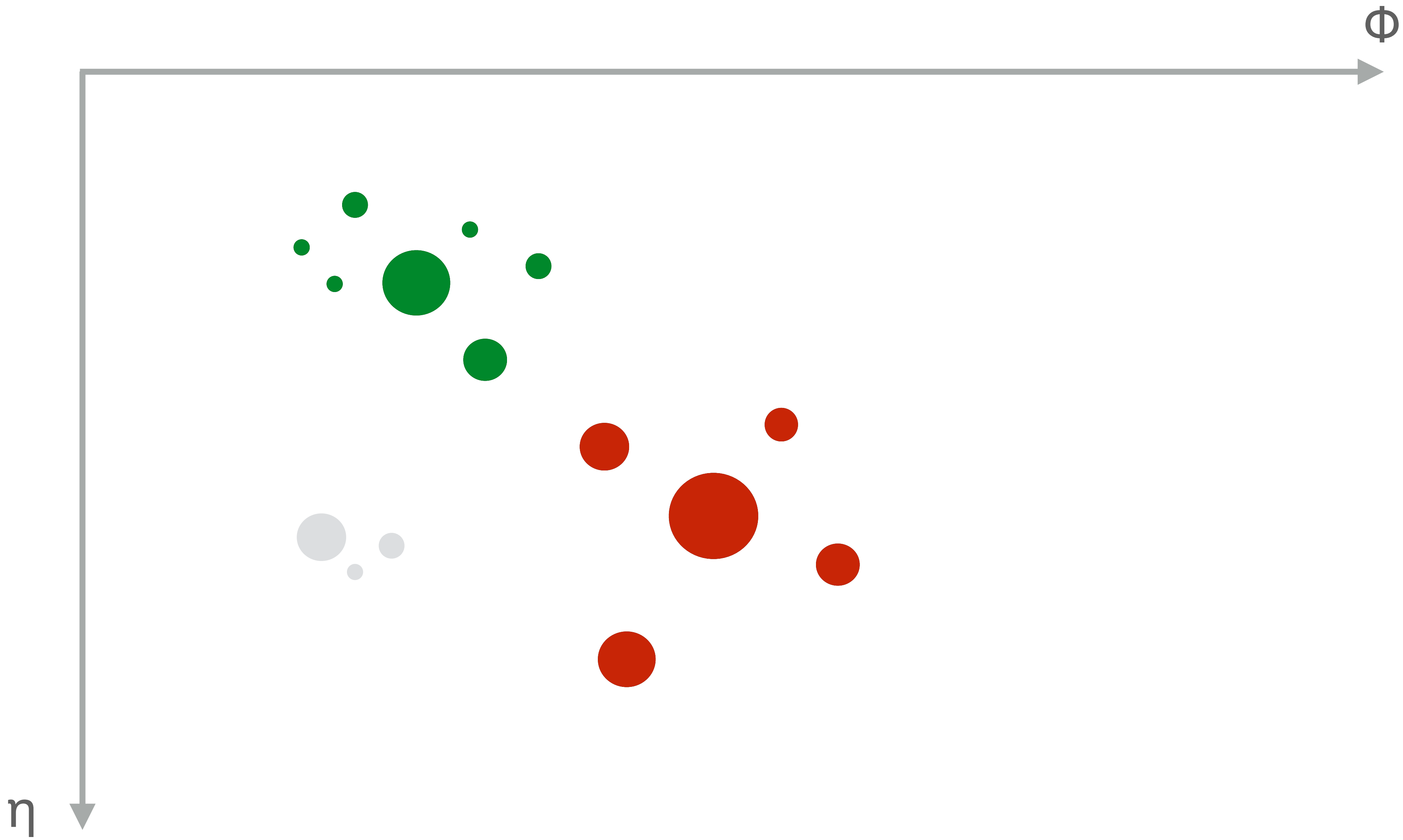
$$\min(p_{\mathrm{T,i}}, p_{\mathrm{T,j}}) > z_{\mathrm{cut}}\,(p_{\mathrm{T,i}} + p_{\mathrm{T,j}}) \left(\frac{\Delta R_{ij}}{R}\right)^{\beta}$$



Jet grooming (soft drop)

Invert clustering history with a veto condition:

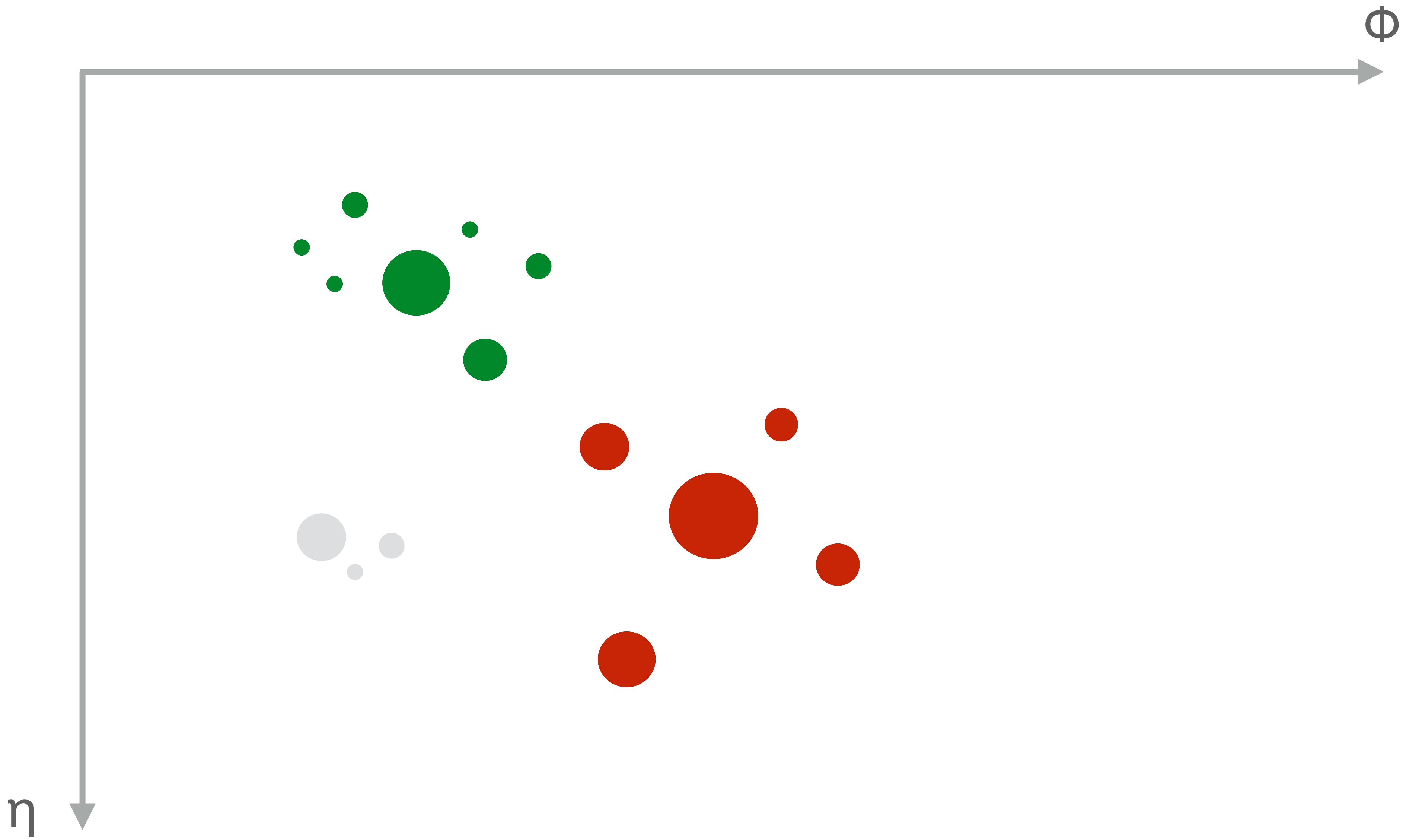
$$\min(p_{\mathrm{T,i}}, p_{\mathrm{T,j}}) > z_{\mathrm{cut}}\,(p_{\mathrm{T,i}} + p_{\mathrm{T,j}}) \left(\frac{\Delta R_{ij}}{R}\right)^{\beta}$$



Jet grooming (soft drop)

Invert clustering history with a veto condition:

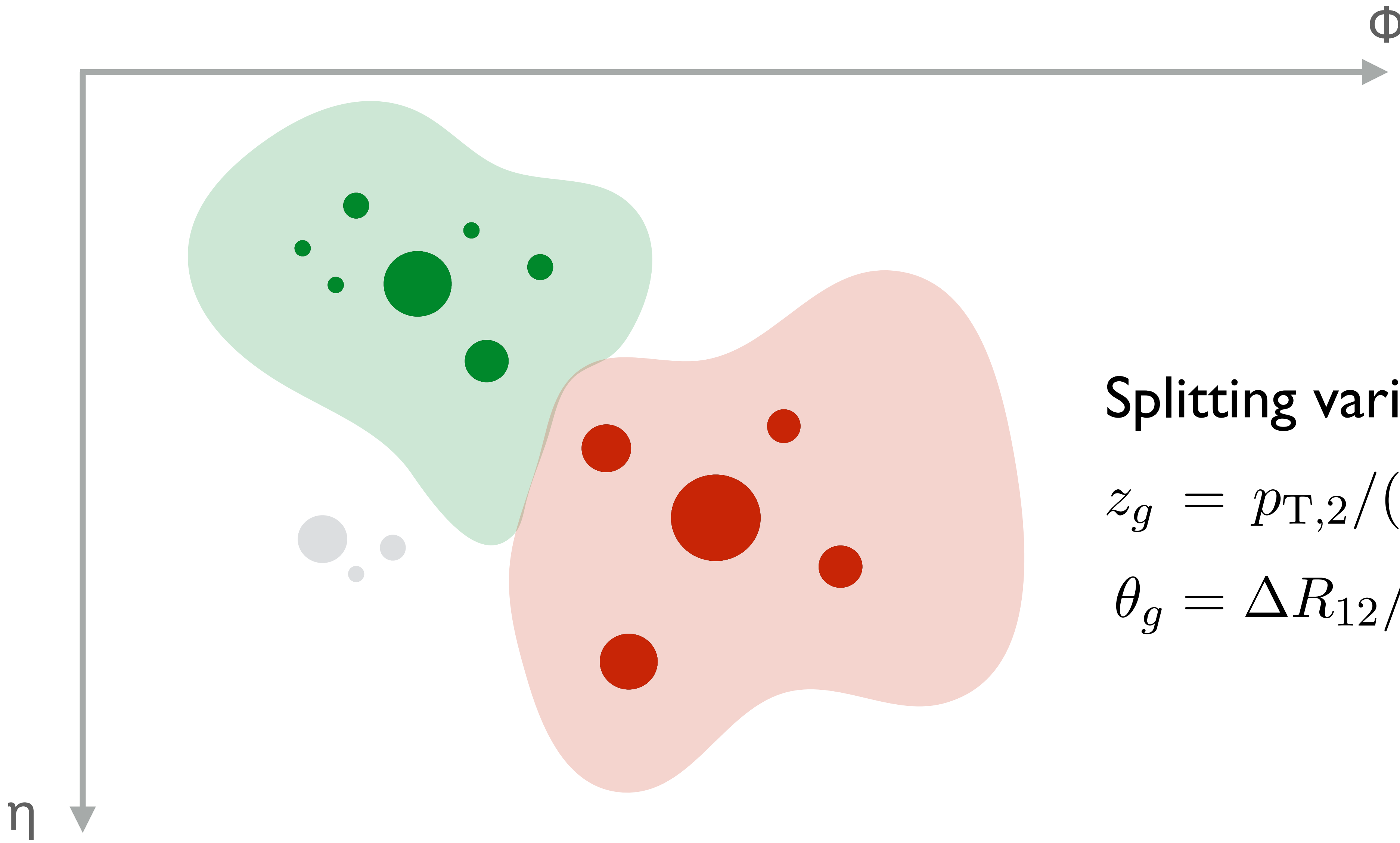
$$\min(p_{\mathrm{T,i}}, p_{\mathrm{T,j}}) > z_{\mathrm{cut}}\,(p_{\mathrm{T,i}} + p_{\mathrm{T,j}}) \left(\frac{\Delta R_{ij}}{R}\right)^{\beta} \quad ✓$$



Jet grooming (soft drop)

Invert clustering history with a veto condition:

$$\min(p_{\mathrm{T,i}}, p_{\mathrm{T,j}}) > z_{\mathrm{cut}}\,(p_{\mathrm{T,i}} + p_{\mathrm{T,j}}) \left(\frac{\Delta R_{ij}}{R}\right)^{\beta} \quad ✓$$



Splitting variables:

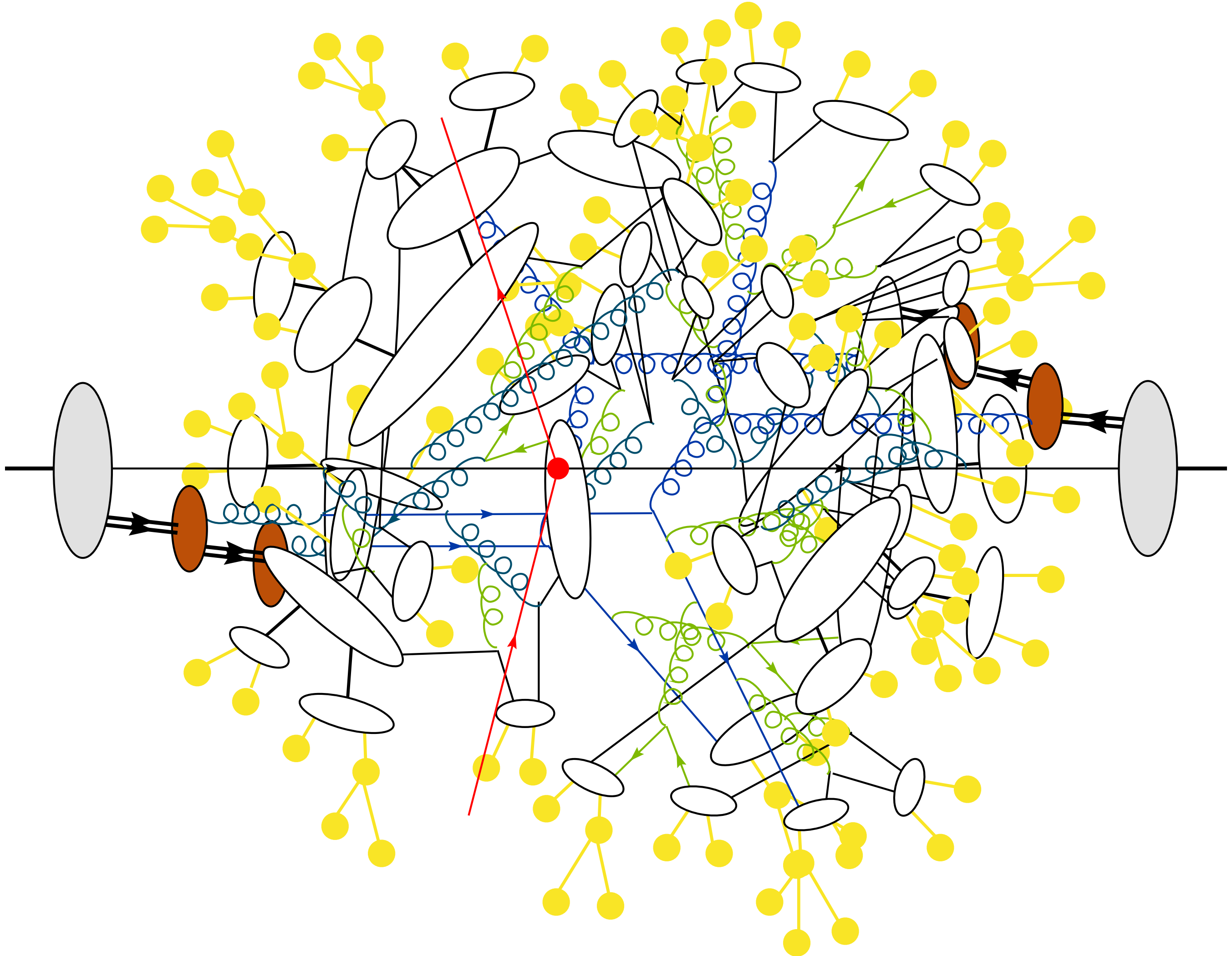
$$z_g = p_{\mathrm{T},2}/(p_{\mathrm{T},1} + p_{\mathrm{T},2})$$

$$\theta_g = \Delta R_{12}/R$$

Importance of grooming

Jet kinematics affected by

- Perturbative radiation
- Hadronisation
- Multiple pp interactions
- Pileup



[S. Gieseke, MCNet 2010]

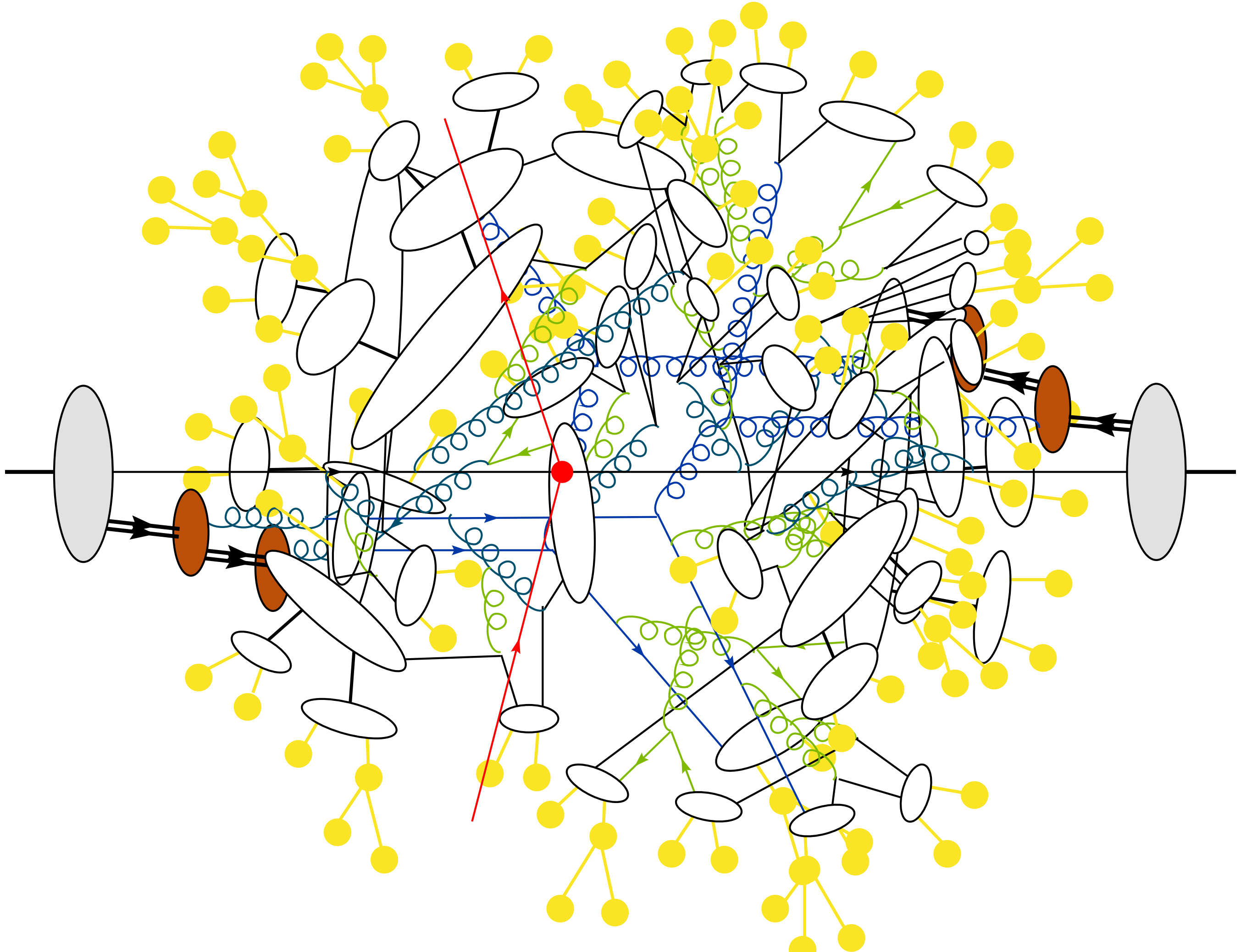
Importance of grooming

Jet kinematics affected by

- Perturbative radiation
- Hadronisation
- Multiple pp interactions
- Pileup

$$\langle \delta p_{\mathrm{T}} \rangle_h = 2 C_R A(\mu_I) M \left(-\frac{1}{R} + \mathcal{O}(R) \right)$$

$$\langle \delta p_{\mathrm{T}} \rangle_{\mathrm{UE}} = \frac{\Lambda_{\mathrm{UE}}}{2} \left(R^2 - R^4/8 + \mathcal{O}(R^6) \right)$$



[S. Gieseke, MCNet 2010]

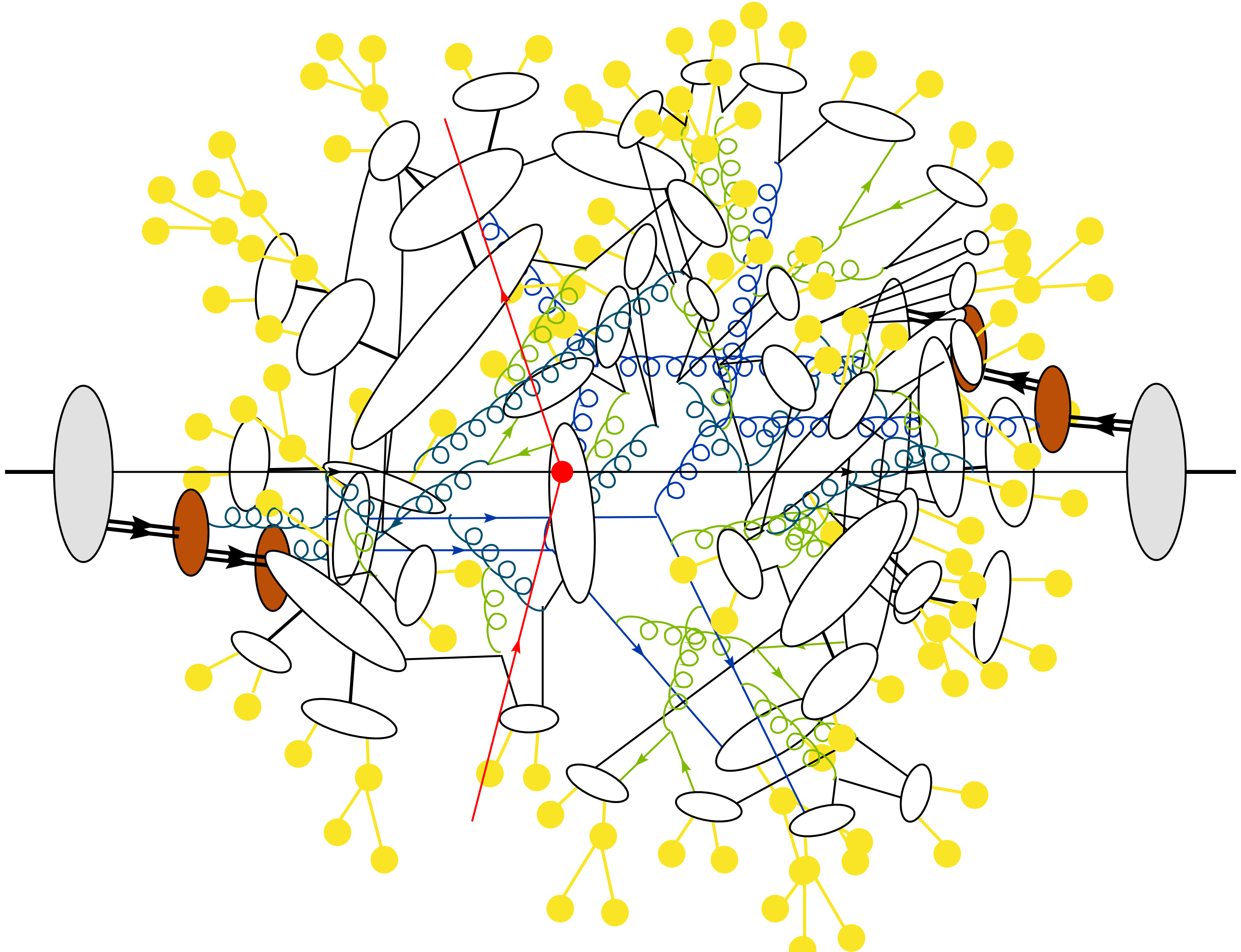
Importance of grooming

Jet kinematics affected by

- Perturbative radiation
- Hadronisation
- Multiple pp interactions
- Pileup

$$\langle \delta p_{\mathrm{T}} \rangle_h = 2 C_R A(\mu_I) M \left(-\frac{1}{R} + \mathcal{O}(R) \right)$$

$$\langle \delta p_{\mathrm{T}} \rangle_{\mathrm{UE}} = \frac{\Lambda_{\mathrm{UE}}}{2} \left(R^2 - R^4/8 + \mathcal{O}(R^6) \right)$$



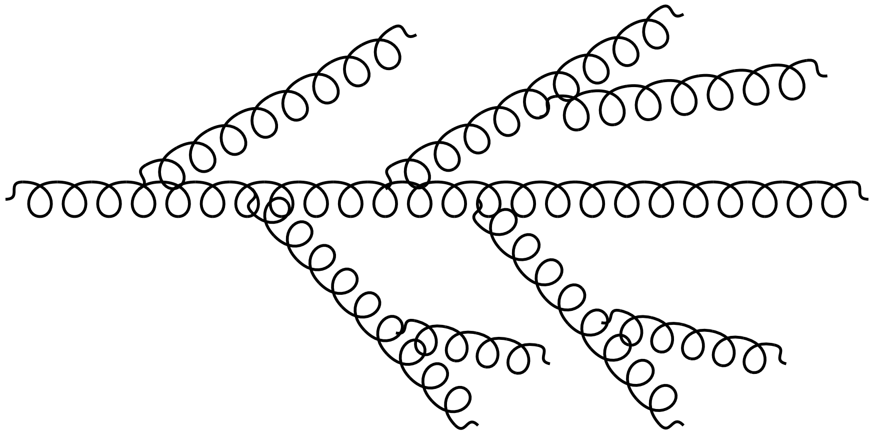
[S. Gieseke, MCNet 2010]

Grooming

- Reduces non-perturbative effects
- Gives access to internal jet structure

Jet mass

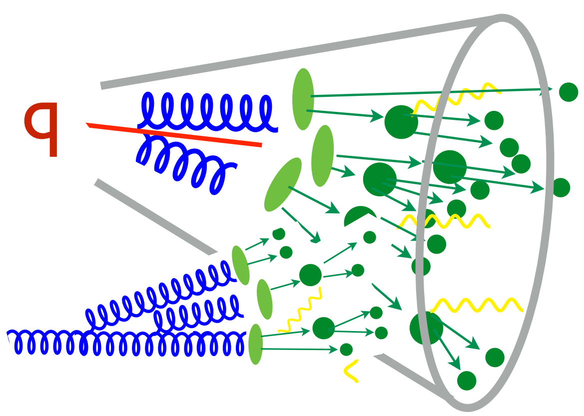
Perturbative effects



Mass generated at each splitting:

$$m^2 = (p_1 + p_2)^2 \approx p_{\mathrm{T},1} p_{\mathrm{T},2} R_{12}^2$$

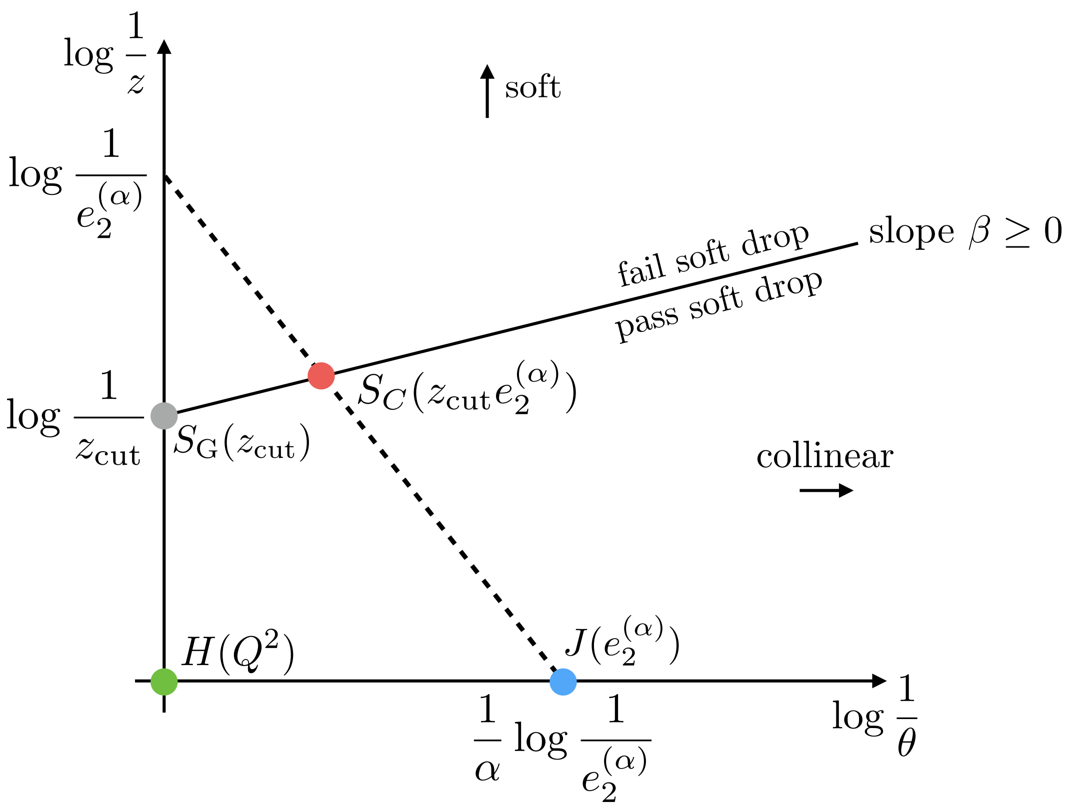
Non-perturbative effects



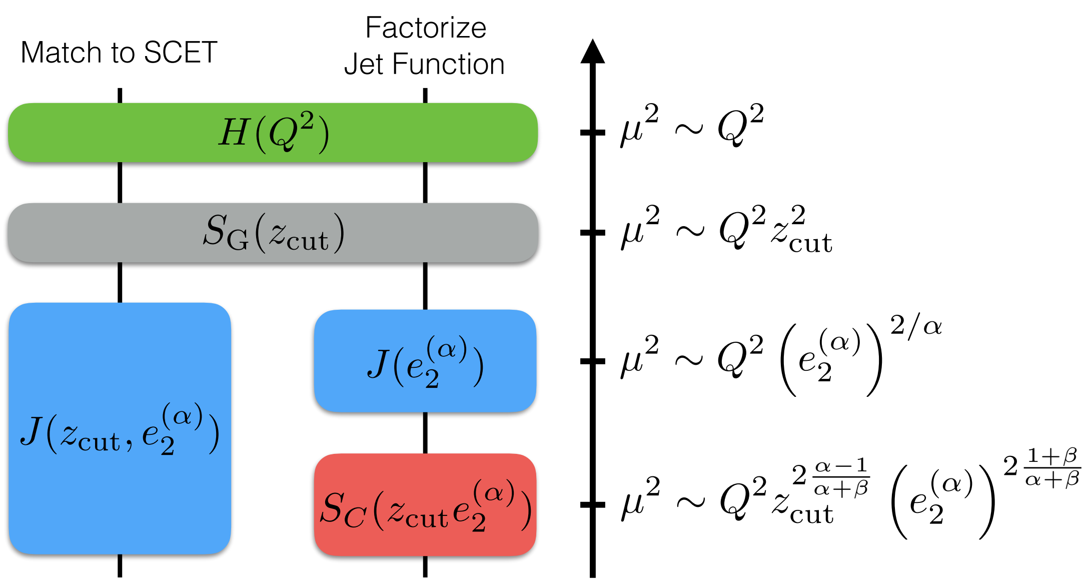
$$\langle \delta m^2 \rangle_h = 2 C_R A(\mu_I) M p_\mathrm{T} \left(R + \mathcal{O}(R^3) \right)$$

$$\langle \delta m^2 \rangle_\mathrm{UE} = C_R \frac{A(\mu_I)}{4} \left(p_\mathrm{T} R^4 + \mathcal{O}(R^6) \right)$$

Jet mass

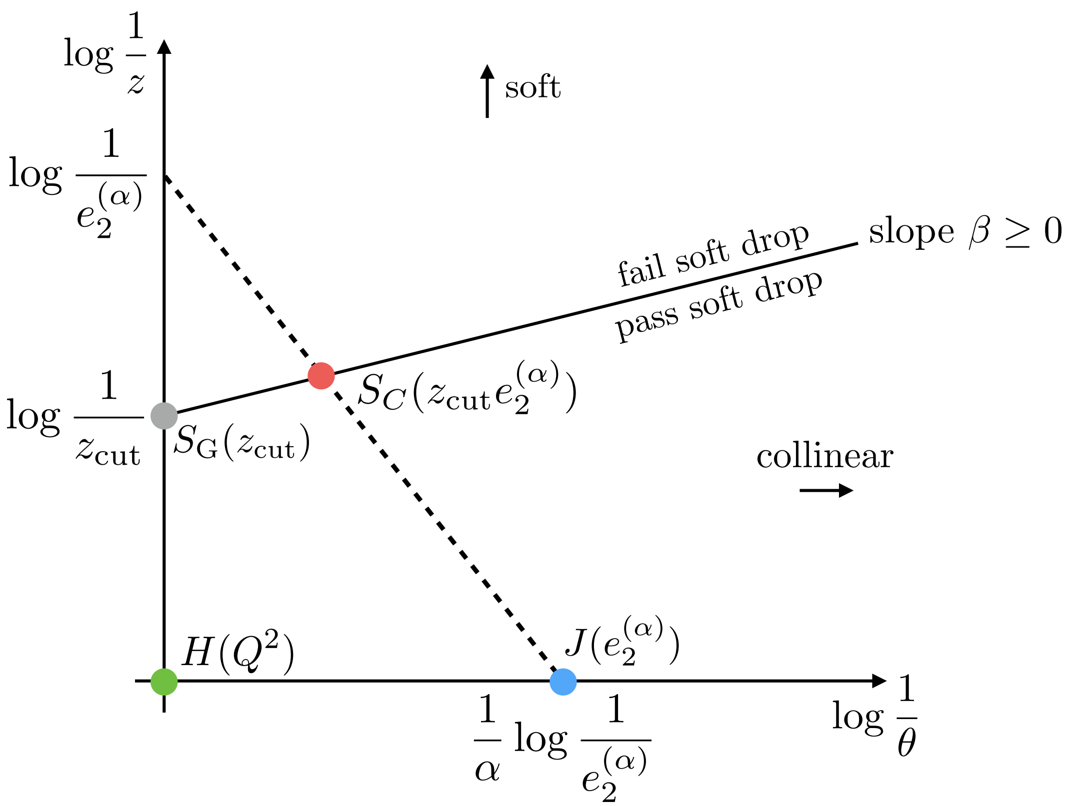


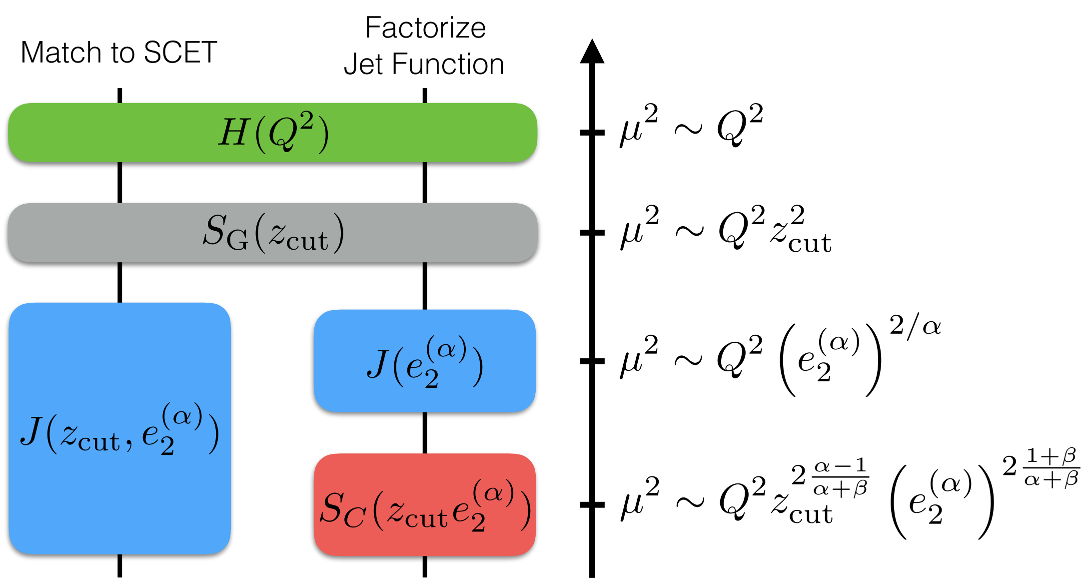
$\log \frac{1}{z}$
soft
$\log \frac{1}{e_2^{(\alpha)}}$
fail soft drop
pass soft drop
slope $\beta \geq 0$
$\log \frac{1}{z_{\text{cut}}}$
$S_C(z_{\text{cut}} e_2^{(\alpha)})$
$S_G(z_{\text{cut}})$
collinear
$H(Q^2)$
$J(e_2^{(\alpha)})$
$\frac{1}{\alpha} \log \frac{1}{e_2^{(\alpha)}}$
$\log \frac{1}{\theta}$



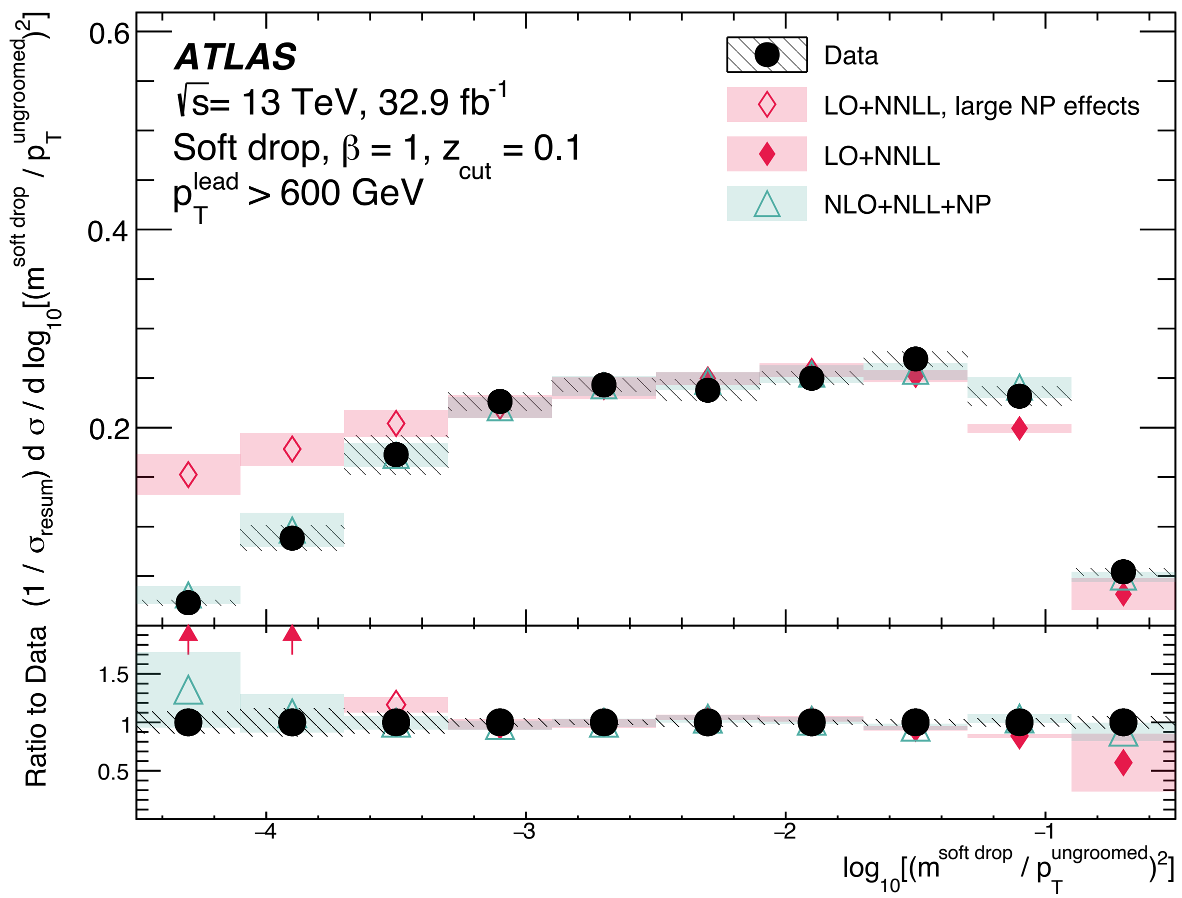
Match to SCET
Factorize Jet Function
$H(Q^2)$
$S_G(z_{\text{cut}})$
$J(z_{\text{cut}}, e_2^{(\alpha)})$
$J(e_2^{(\alpha)})$
$S_C(z_{\text{cut}} e_2^{(\alpha)})$
$\mu^2 \sim Q^2$
$\mu^2 \sim Q^2 z_{\text{cut}}^2$
$\mu^2 \sim Q^2 \left(e_2^{(\alpha)}\right)^{2/\alpha}$
$\mu^2 \sim Q^2 z_{\text{cut}}^{2\frac{\alpha-1}{\alpha+\beta}} \left(e_2^{(\alpha)}\right)^{2\frac{1+\beta}{\alpha+\beta}}$

Jet mass





[ATLAS, PRL 121, 092001 (2018)]



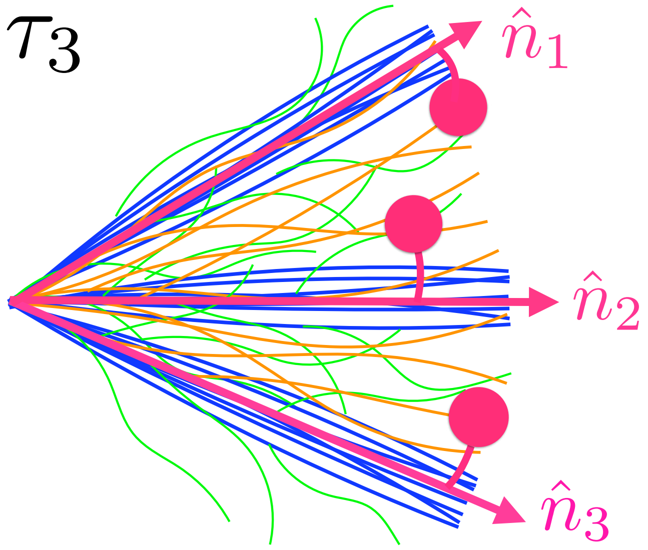
Soft drop

- Enables precise calculations
- Allows tests of QCD in different energy regimes

Jet substructure observables

1. Angularities and energy correlations

N-subjettiness τ_3

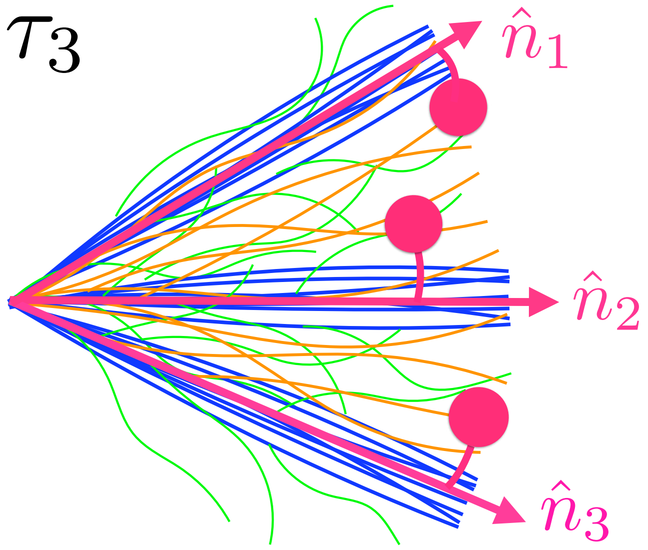


[I. Moult et al, JHEP 12, 153 (2016)]

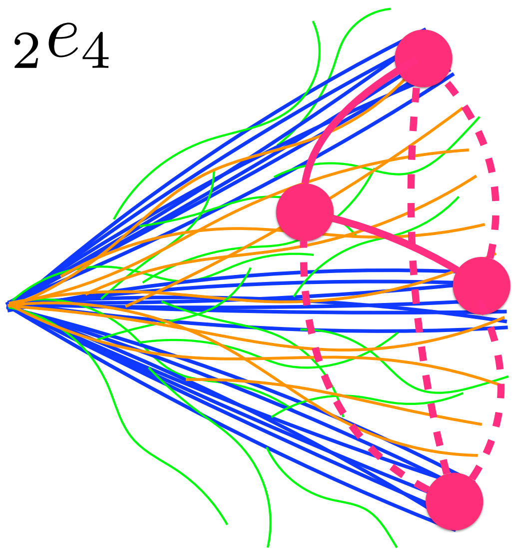
Jet substructure observables

1. Angularities and energy correlations

N-subjettiness τ_3



Normalized energy correlation functions $_2e_4$

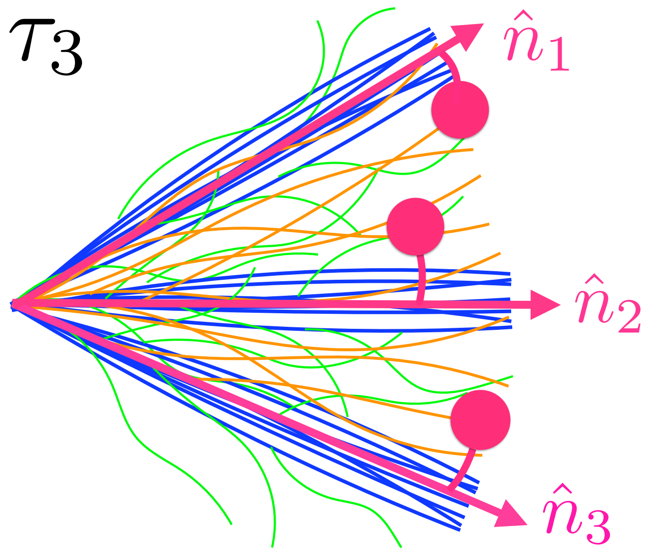


[I. Moult et al, JHEP 12, 153 (2016)]

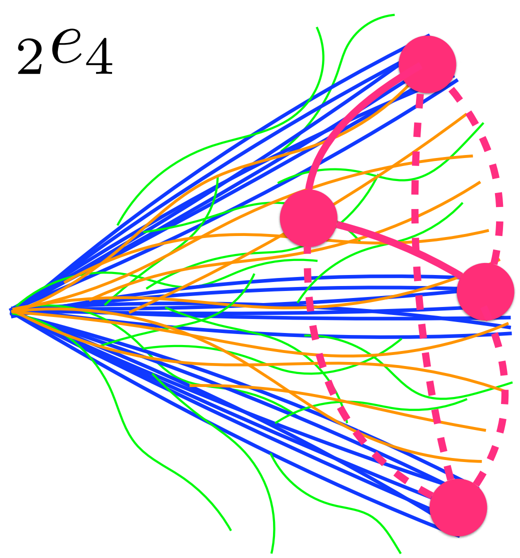
Jet substructure observables

1. Angularities and energy correlations

N-subjettiness τ_3



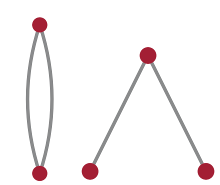
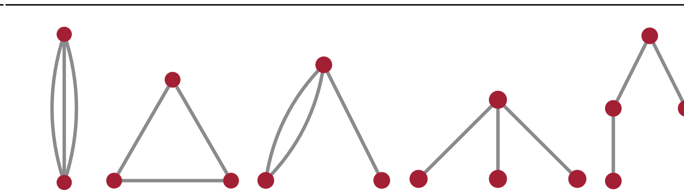
Normalized energy correlation functions $_2e_4$



[I. Moult et al, JHEP 12, 153 (2016)]

Energy Flow Polynomials

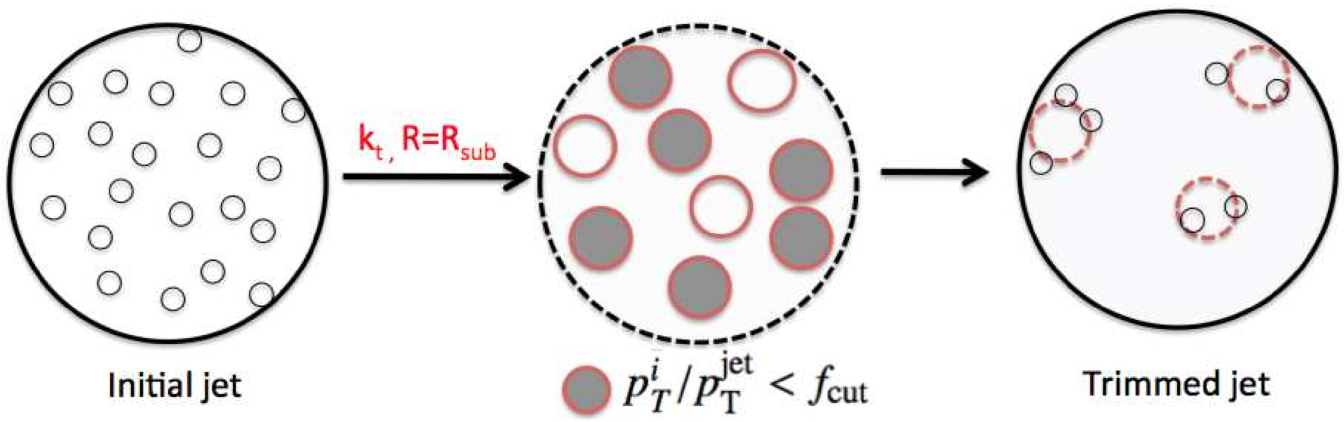
$$\text{EFP}_G = \sum_{i_1=1}^{M} \cdots \sum_{i_N=1}^{M} z_{i_1} \cdots z_{i_N} \prod_{(k,\ell)\in G} \theta_{i_k i_\ell}$$

Degree	Connected Multigraphs
$d=0$	
$d=1$	
$d=2$	
$d=3$	

[P.T. Komiske et al, JHEP 12, 153 (2016)]

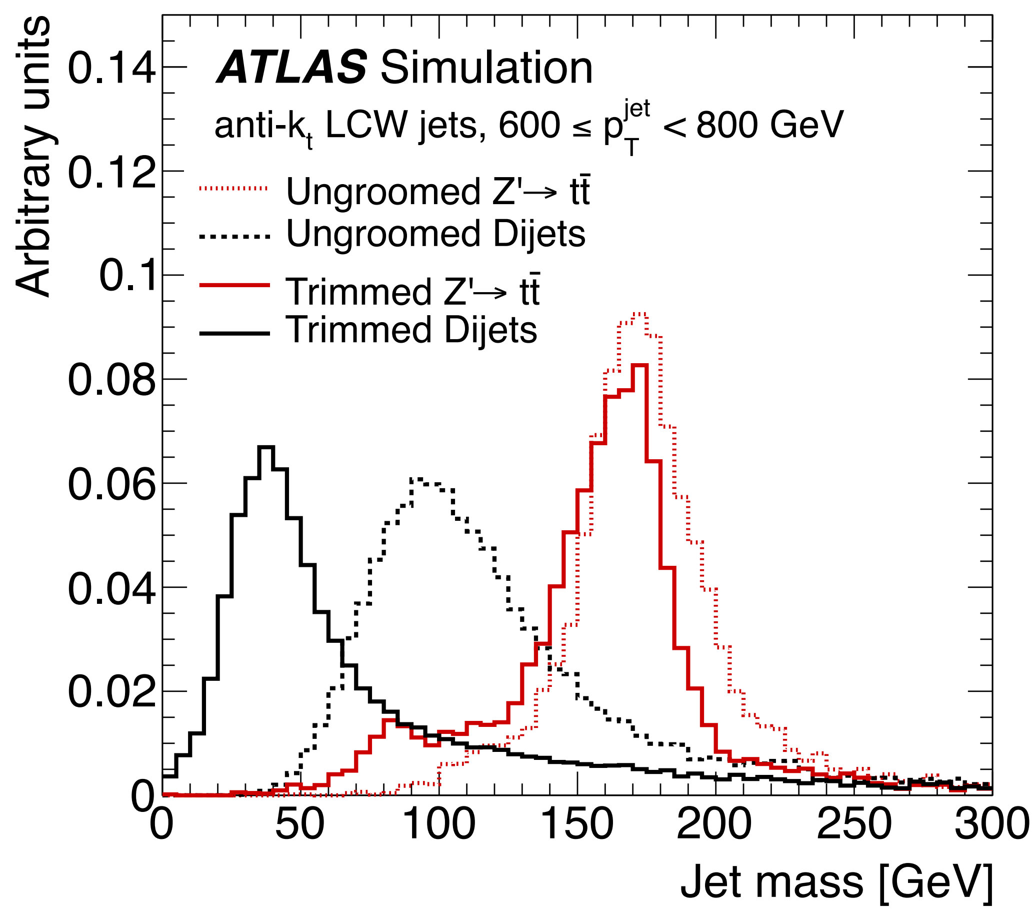
Jet substructure observables

2. Grooming and N-prong taggers



Trimming
(standard in ATLAS)

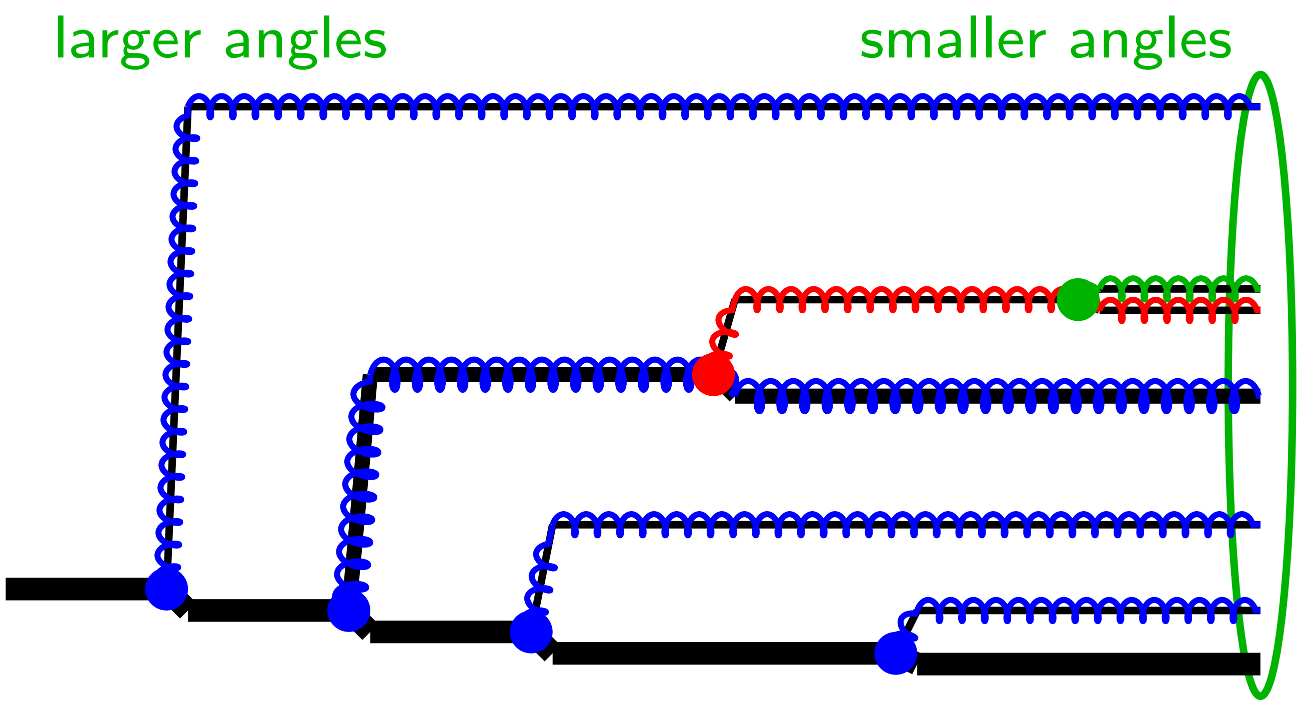
also studied:
Filtering, Pruning, soft drop…



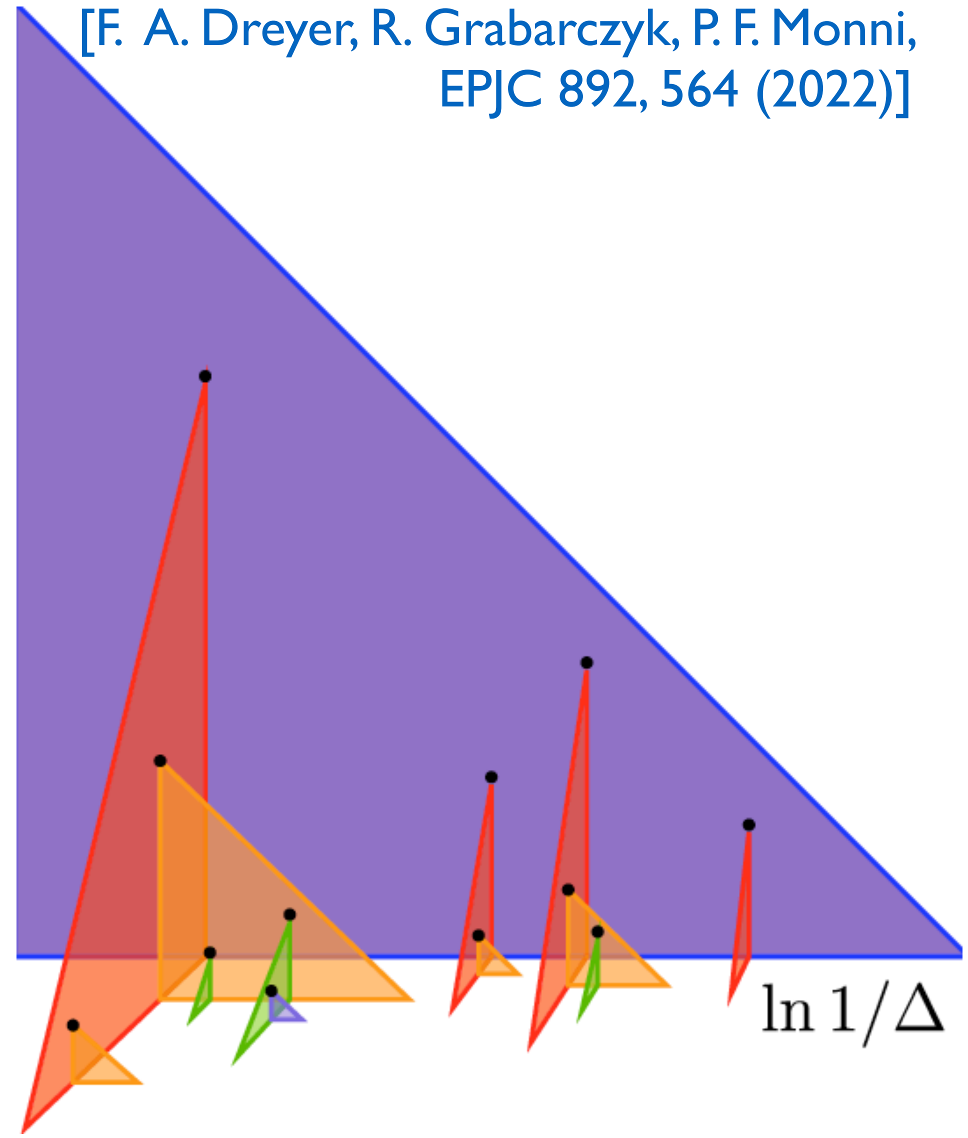
[ATLAS, JHEP 09, 076 (2013)]

Lund jet plane

[G. Soyez, 2022]



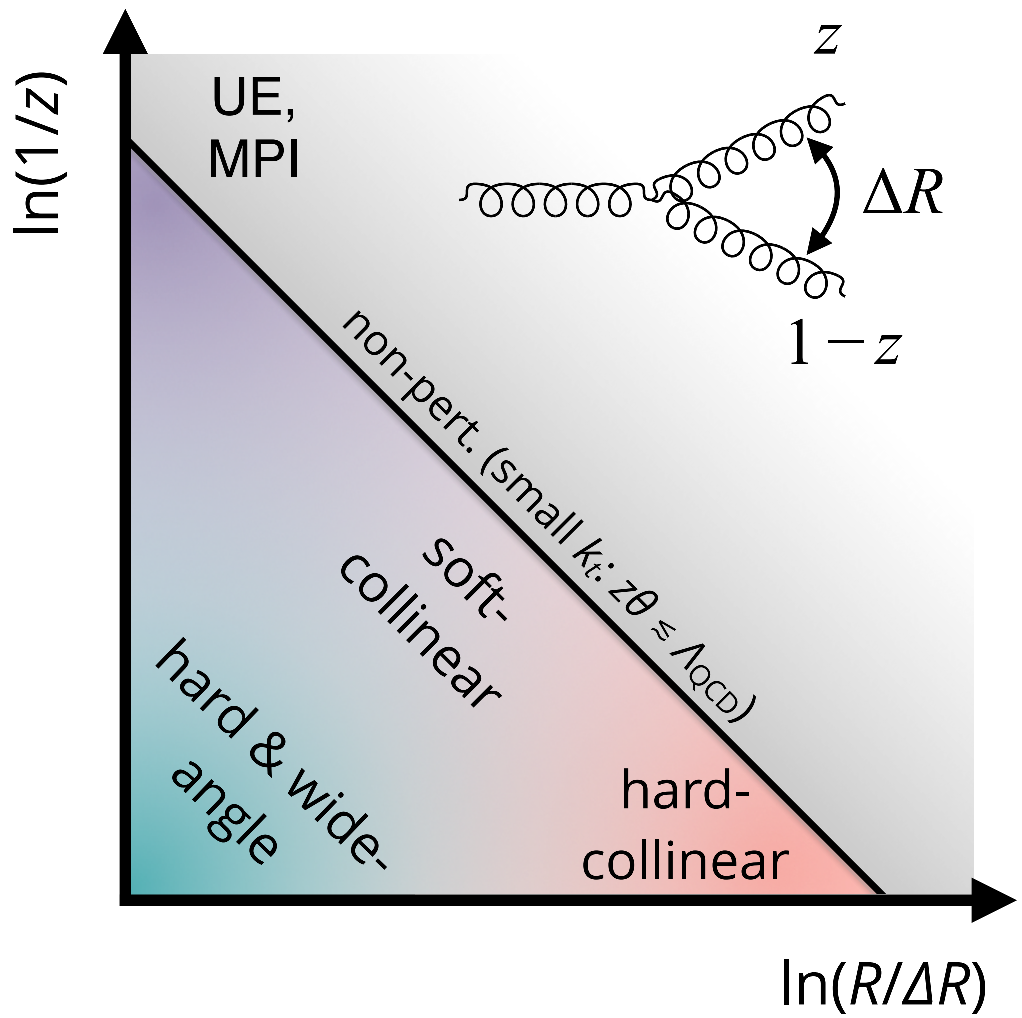
[F. A. Dreyer, R. Grabarczyk, P. F. Monni, EPJC 892, 564 (2022)]



- Decluster a jet to access shower history
- Separate emissions in angular scale and hardness

[F. A. Dreyer, G. P. Salam, G. Soyez, JHEP 12, 064 (2018)]

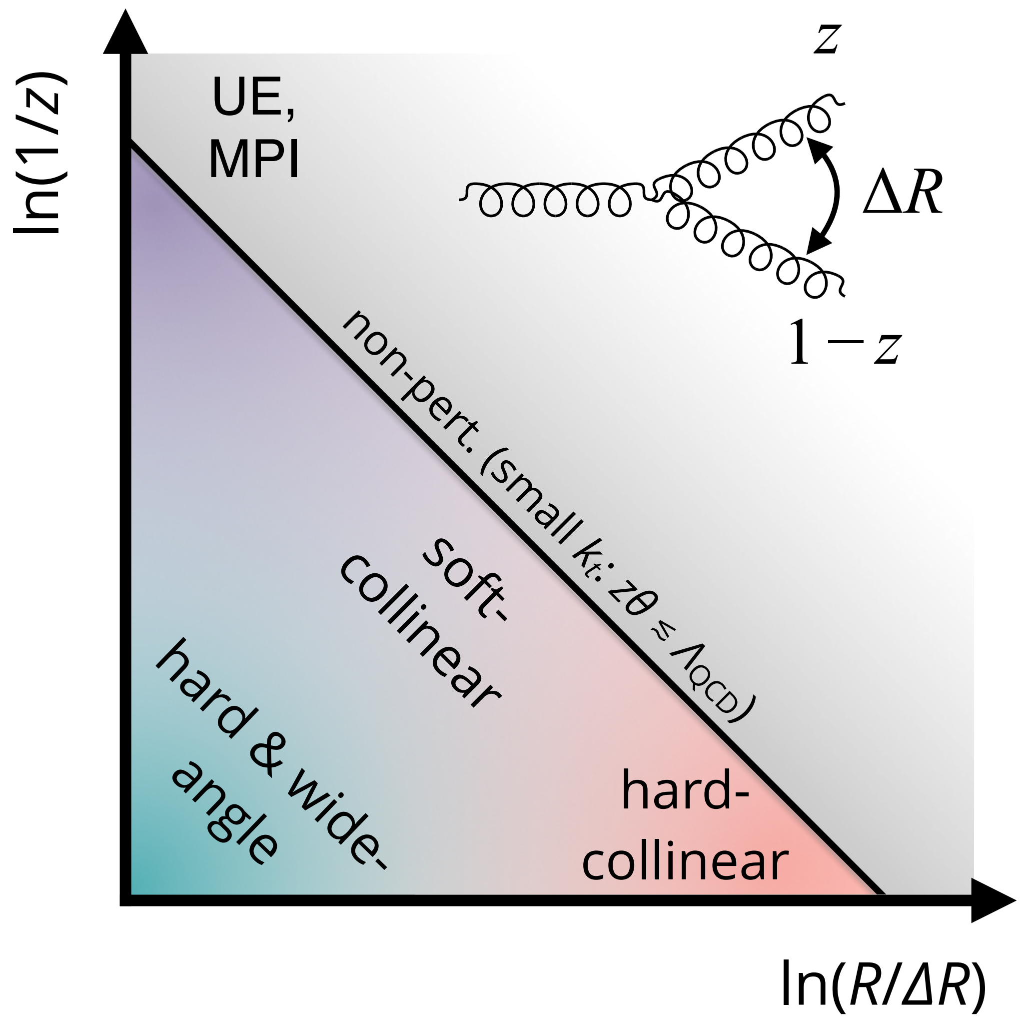
Lund jet plane

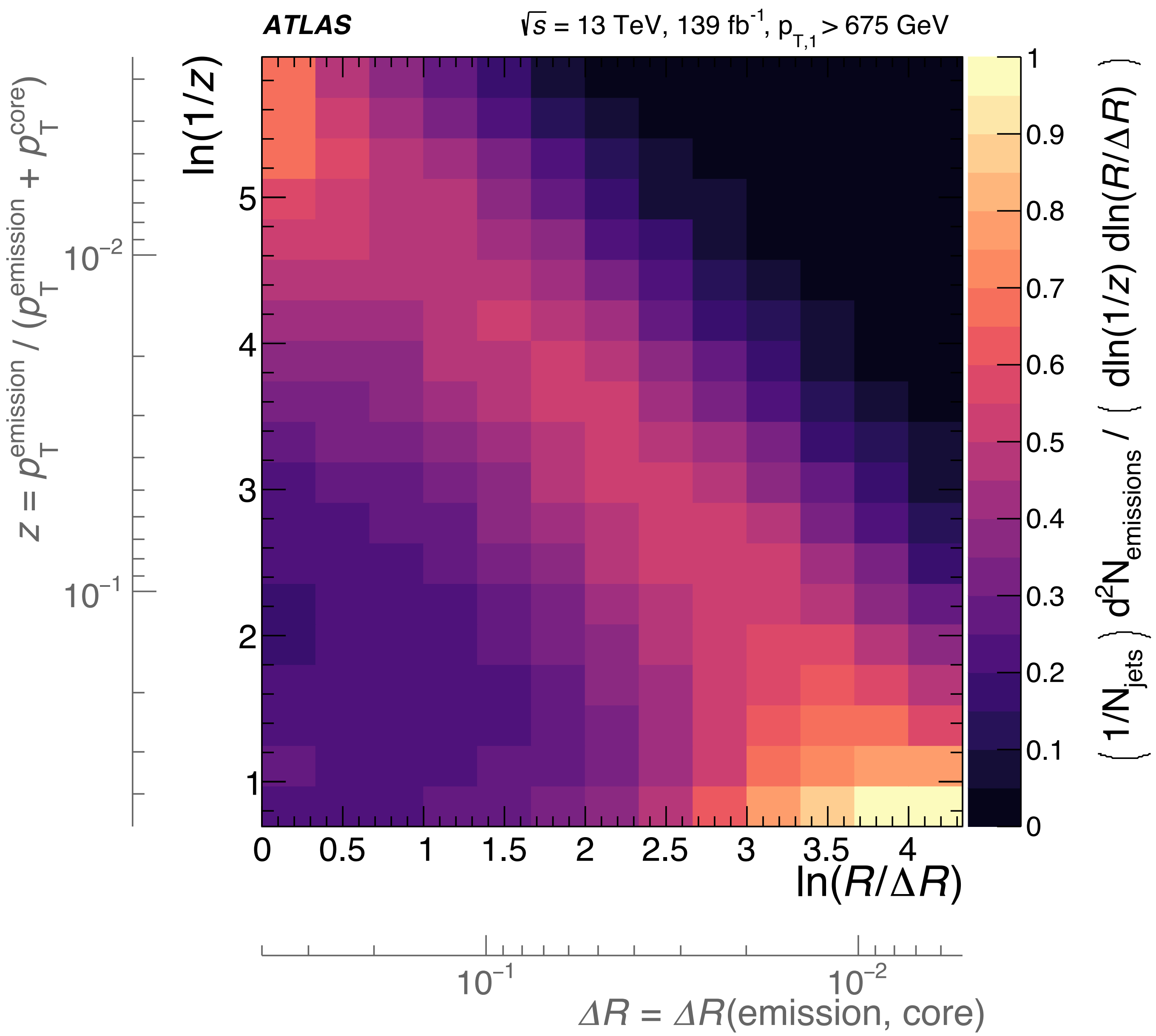


- Access splitting function, jet fragmentation and non-pert. radiation

[ATLAS, PRL 124, 22202 (2020)]

Lund jet plane





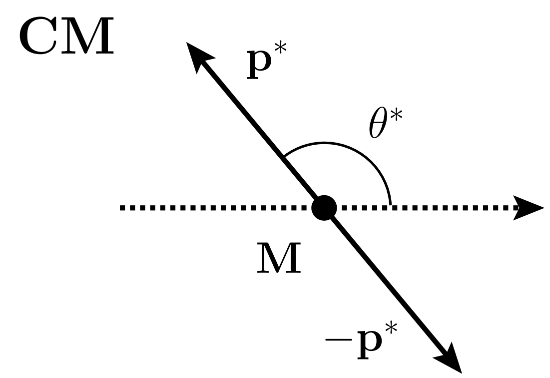
- Access splitting function, jet fragmentation and non-pert. radiation
- Soft-wide angle and hard-collinear radiation dominant

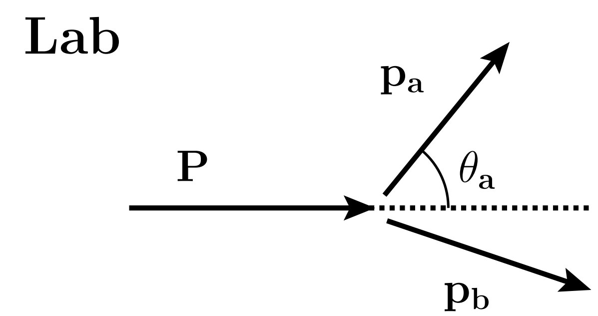
[ATLAS, PRL 124, 22202 (2020)]

Jet tagging with substructure

Particle Decays

W and Z bosons

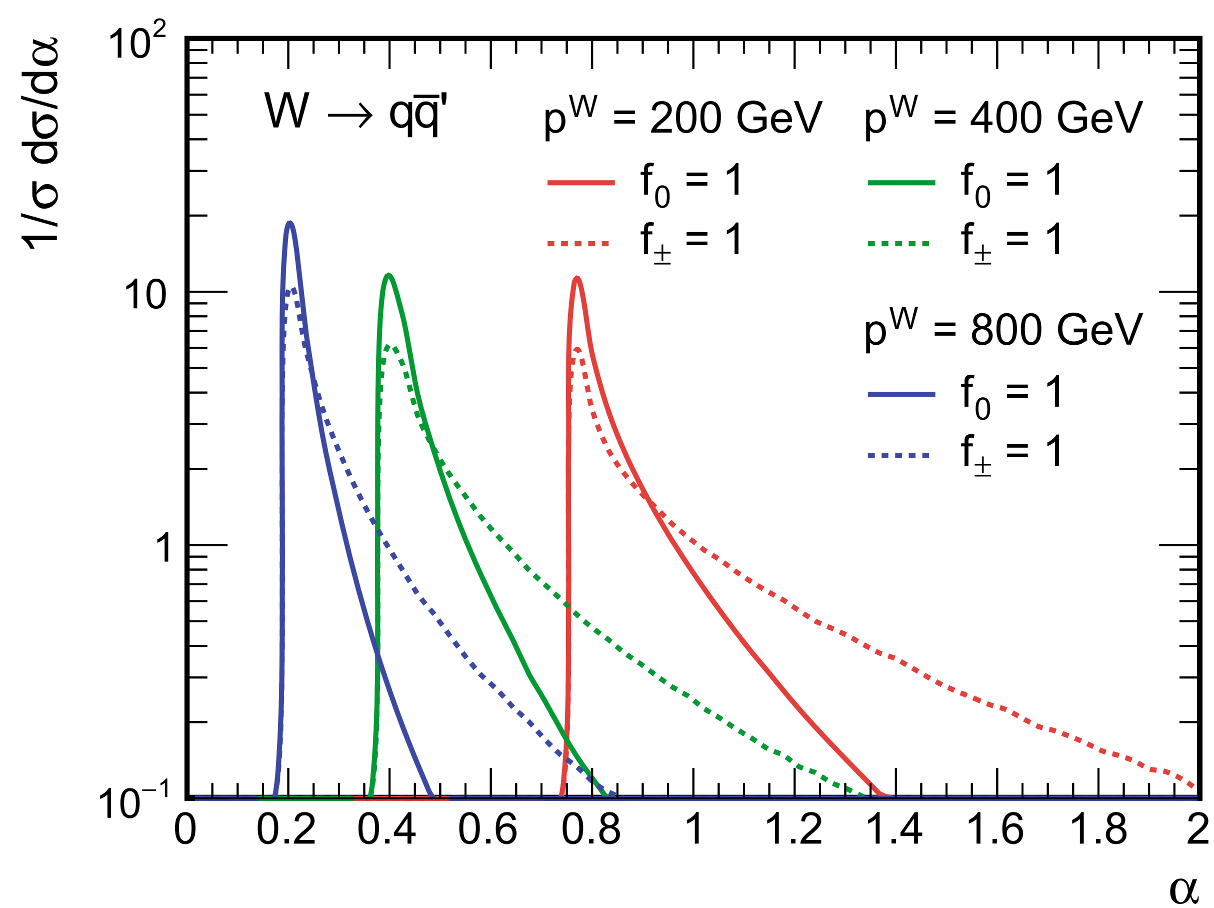




$B_{W\to had} = 67.5\%$

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

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



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

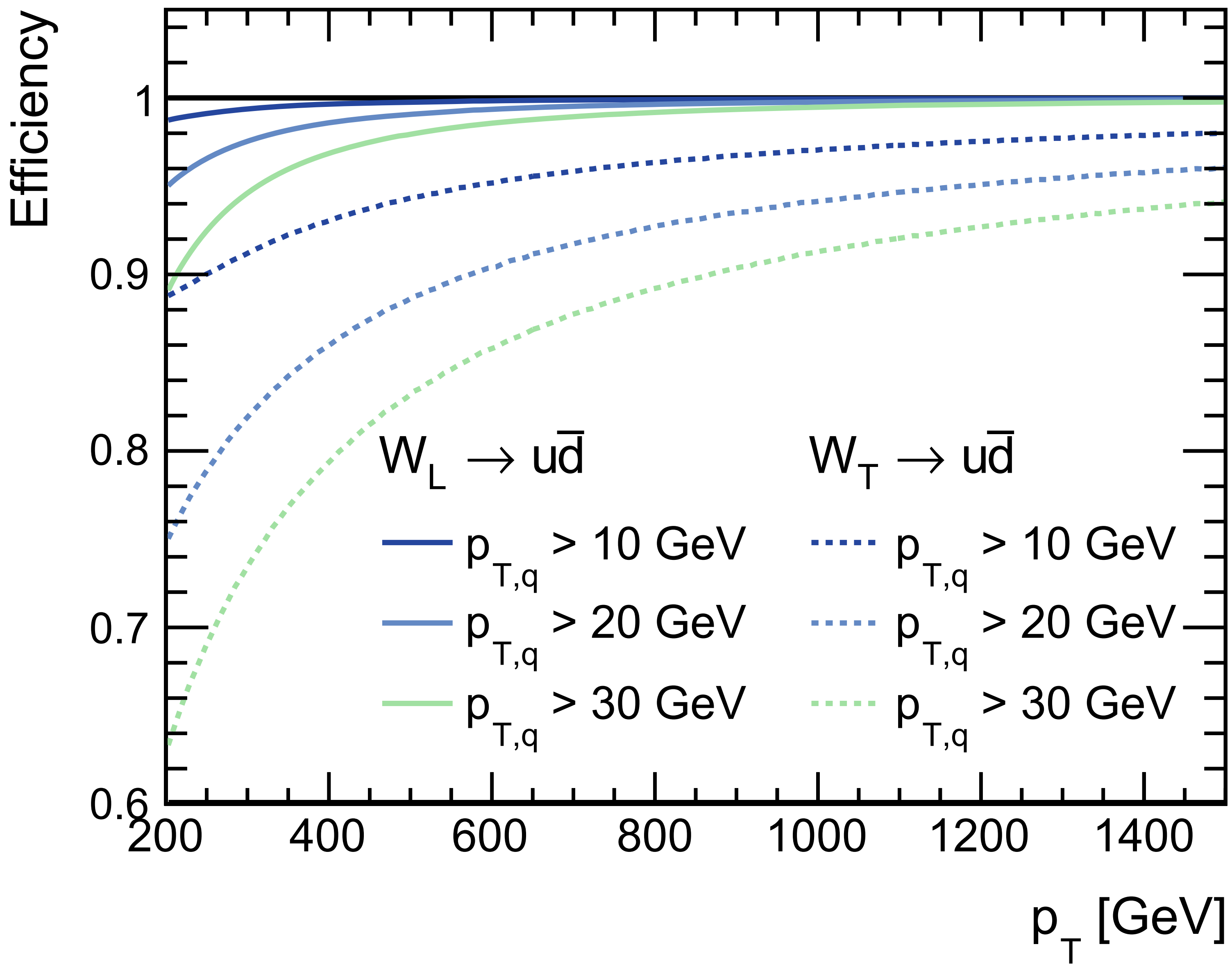
and consequently

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

(holds for $P_T \gg M$)

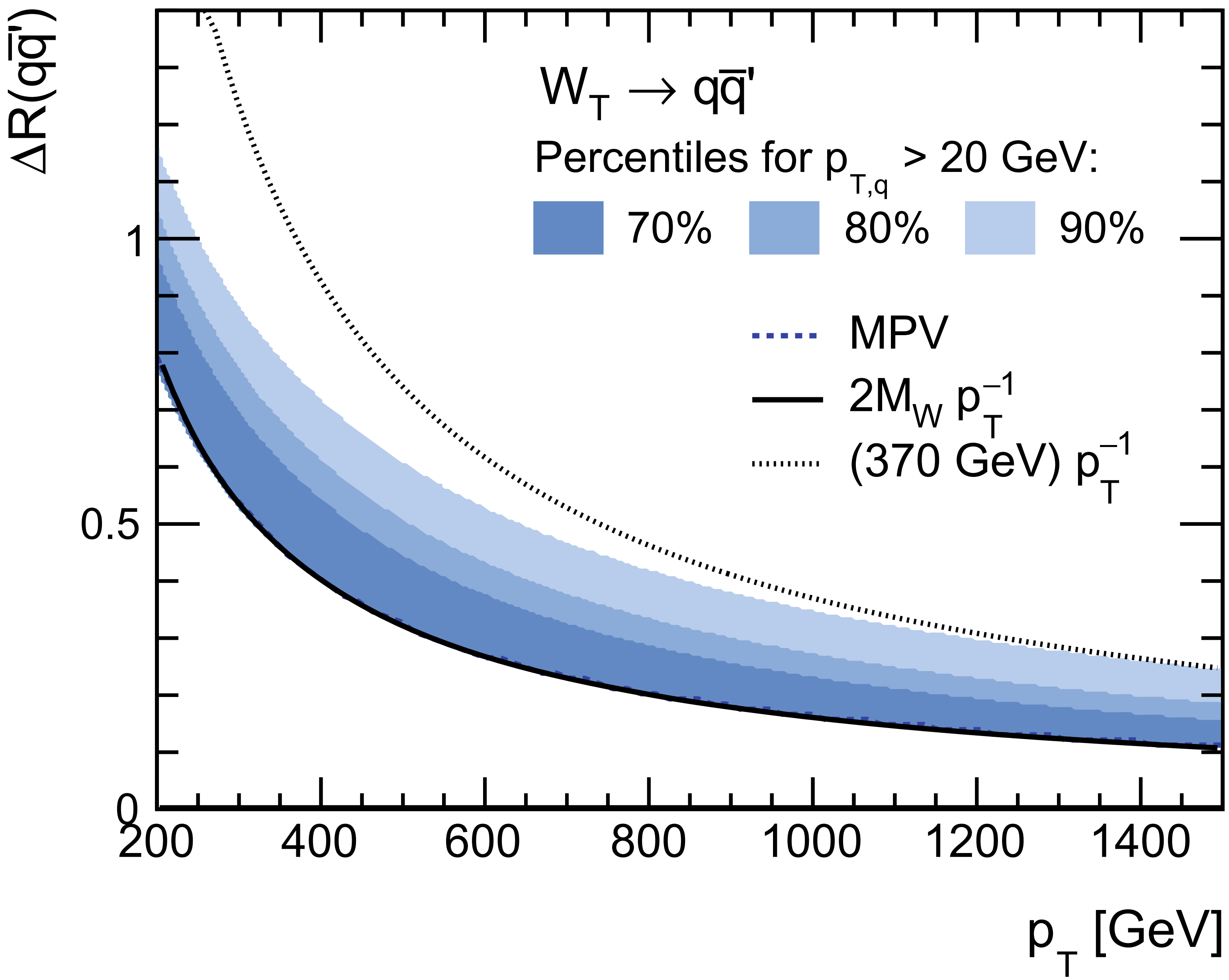
[RK, STMP 284, 2021]

Quark (subjet) p_T thresholds



[RK, STMP 284, 2021]

Decay distance



- Similar picture for top quarks

[RK, STMP 284, 2021]

Heavy Object Tagger with Variable R

One-pass clustering with integrated subjet finding

- jet distance measures (with variable R)

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

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

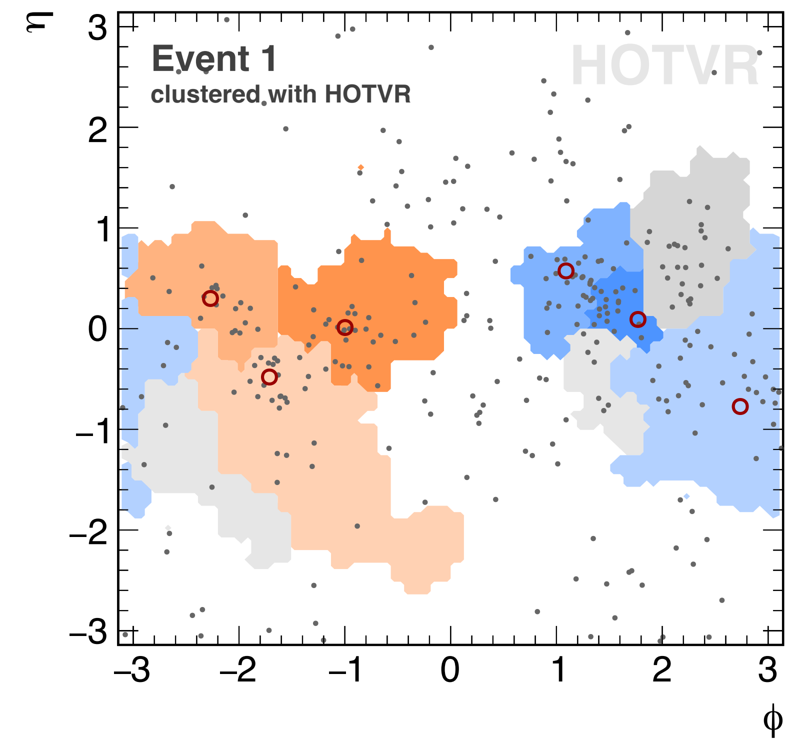
- clustering veto at each step

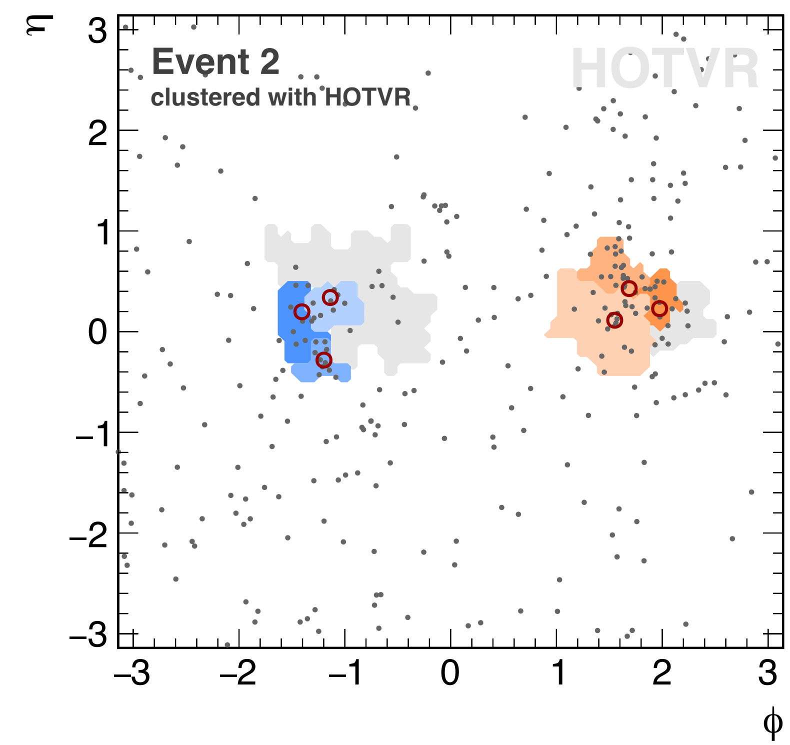
$$m_{ij} < \mu$$

$$\theta \cdot m_{ij} > \max[m_i, m_j]$$

- store objects i and j as subjets if

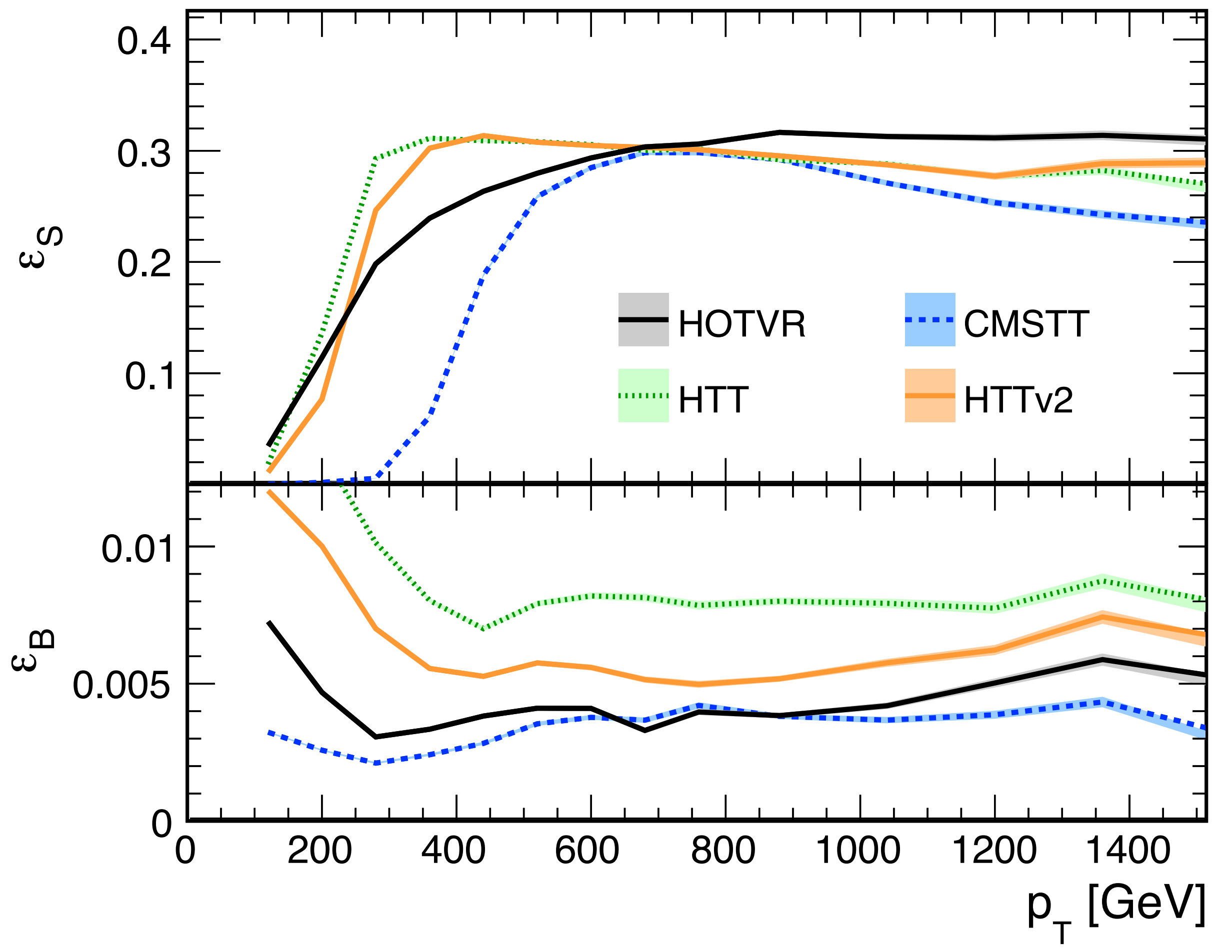
$$p_{T,i}, p_{T,j} > p_{T,\text{sub}}$$





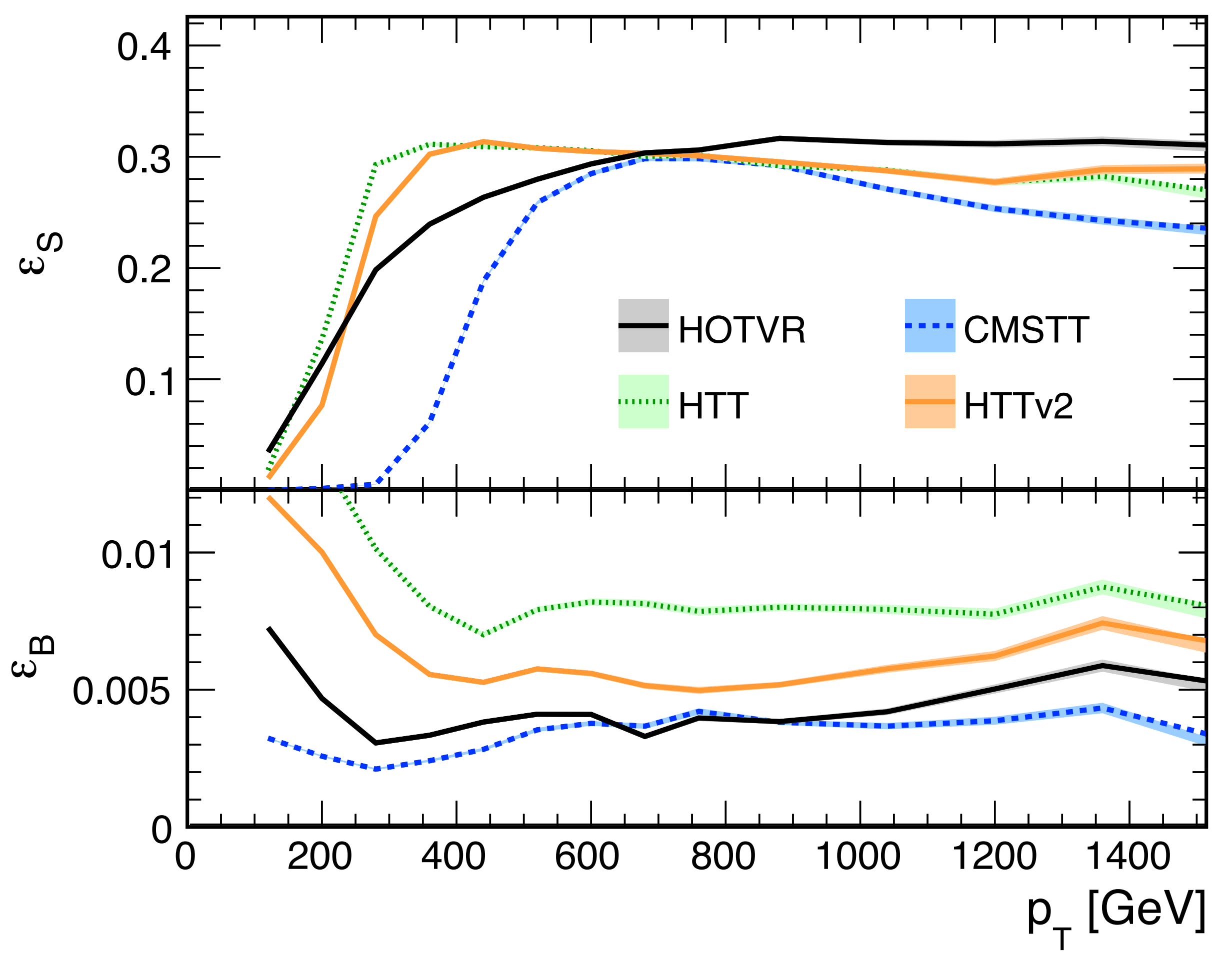
[T. Lapsien et al, EPJC 76, 600 (2016)]

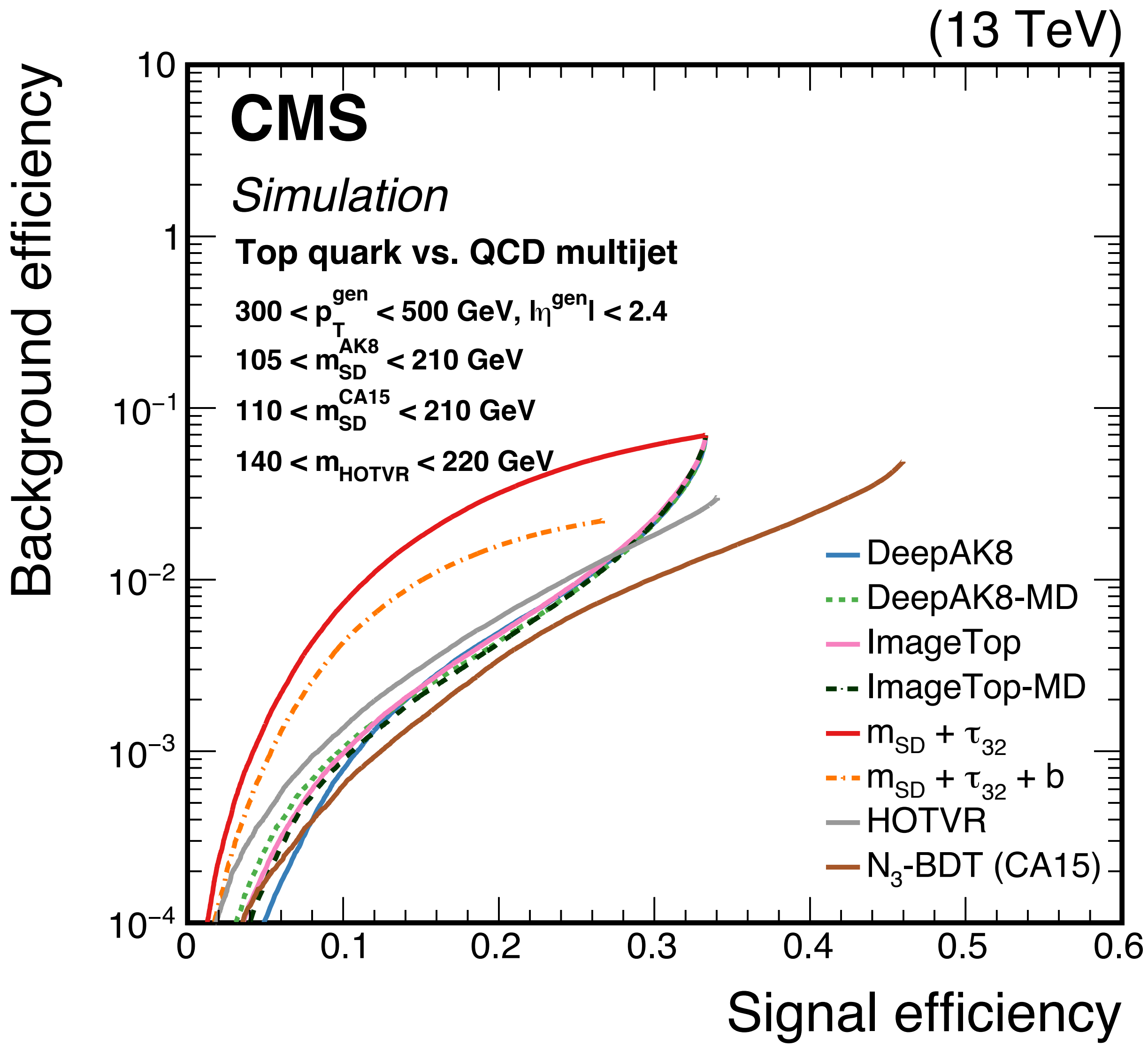
Top Quark Tagging



[RK, STMP 284, 2021]

Top Quark Tagging

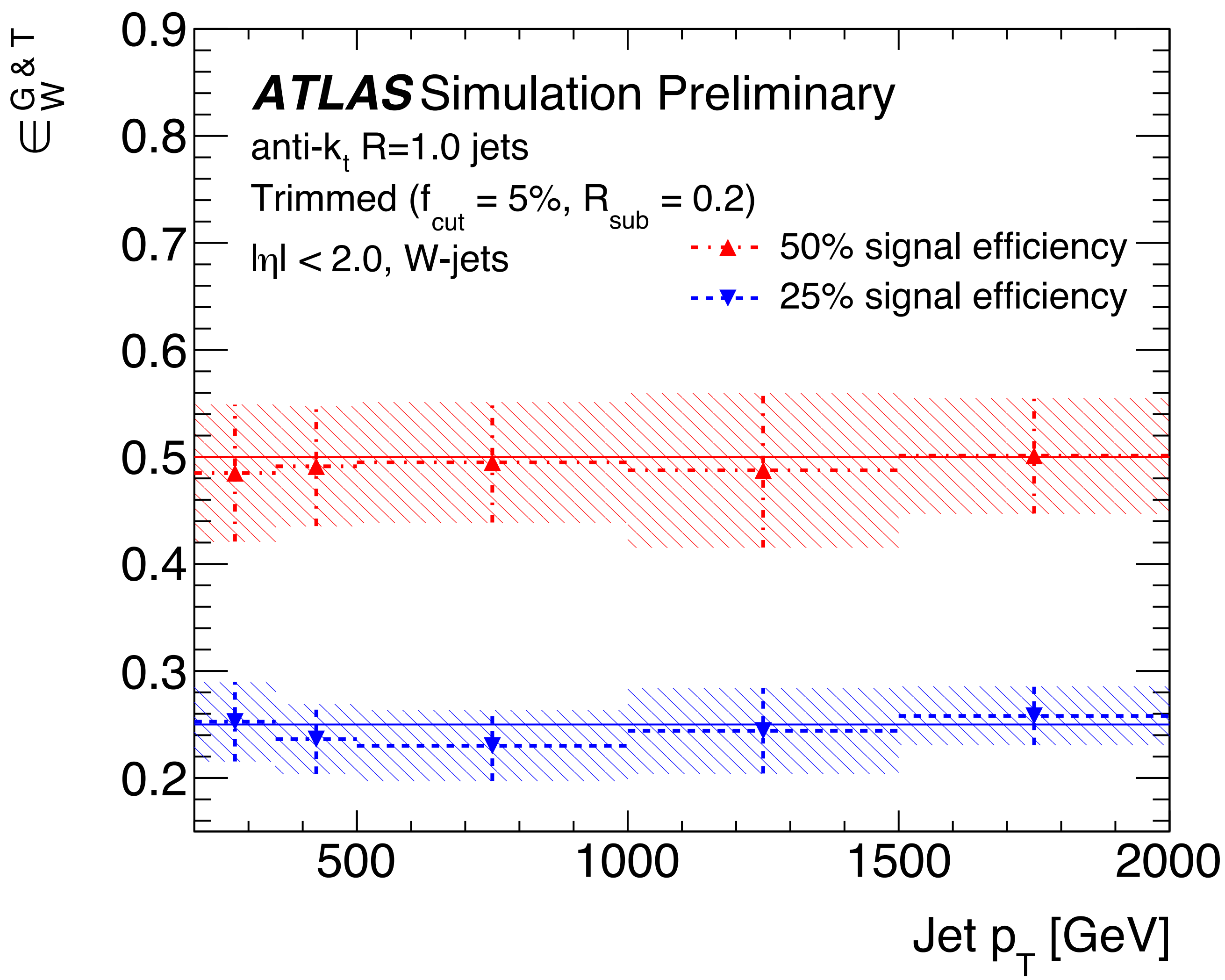


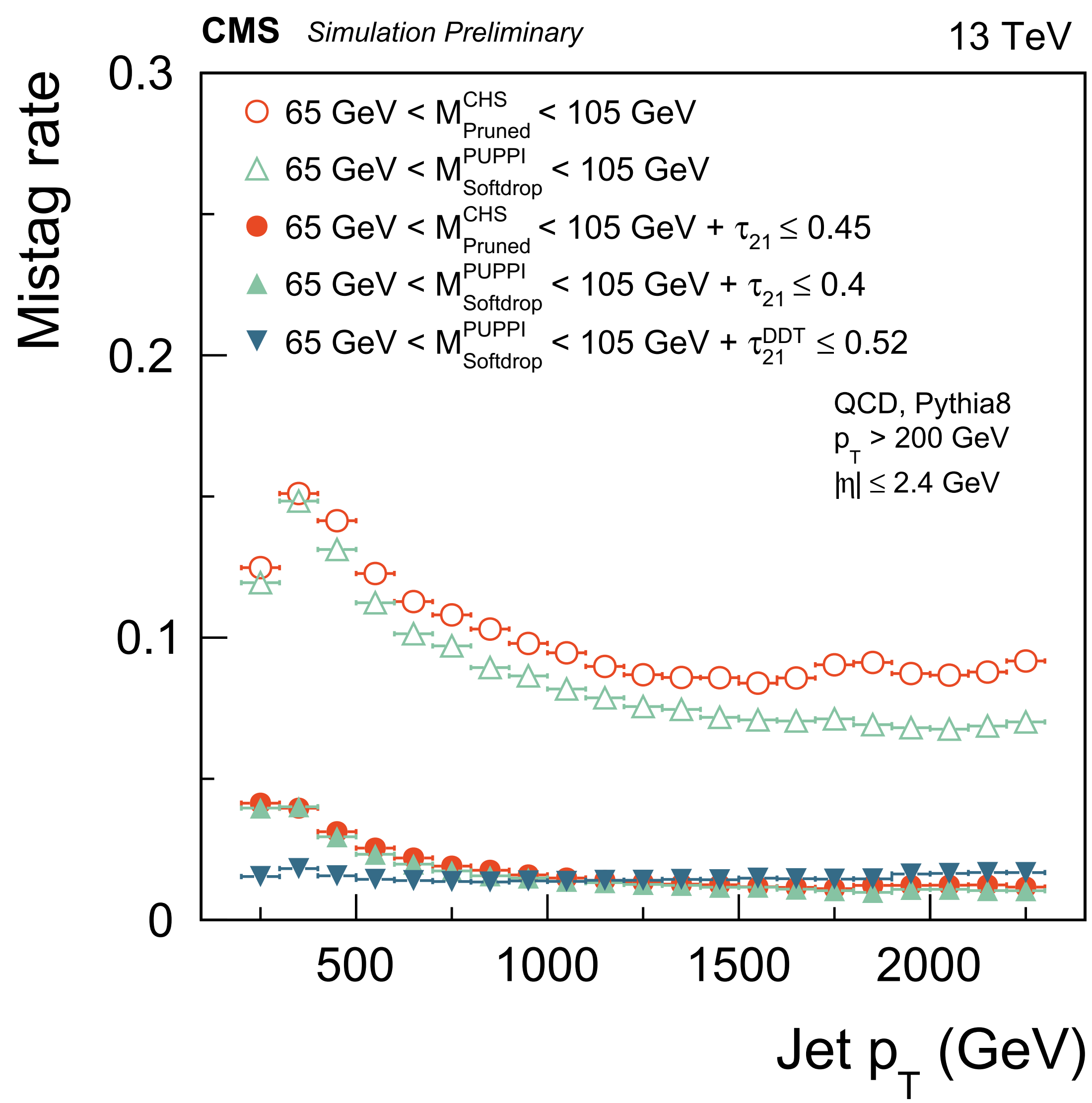


[RK, STMP 284, 2021]

[CMS, JINST 15, P06005 (2020)]

W/Z Tagging





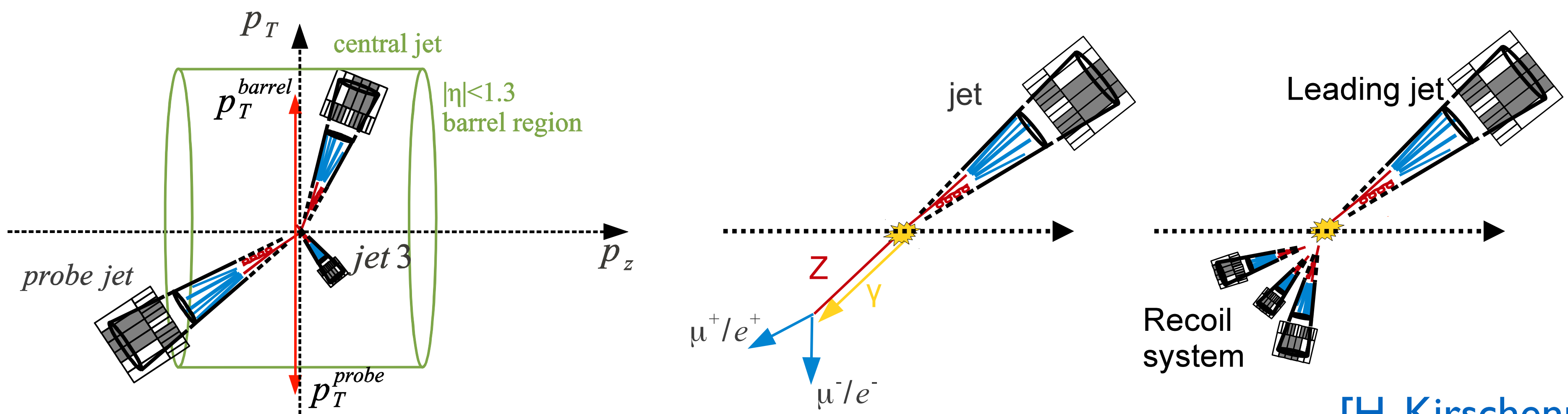
- Constant signal efficiency by p_T-dependent selection
- Constant background efficiency with DDT technique

[ATL-PHYS-PUB-2015-033] [CMS-PAS-JME-16-003]

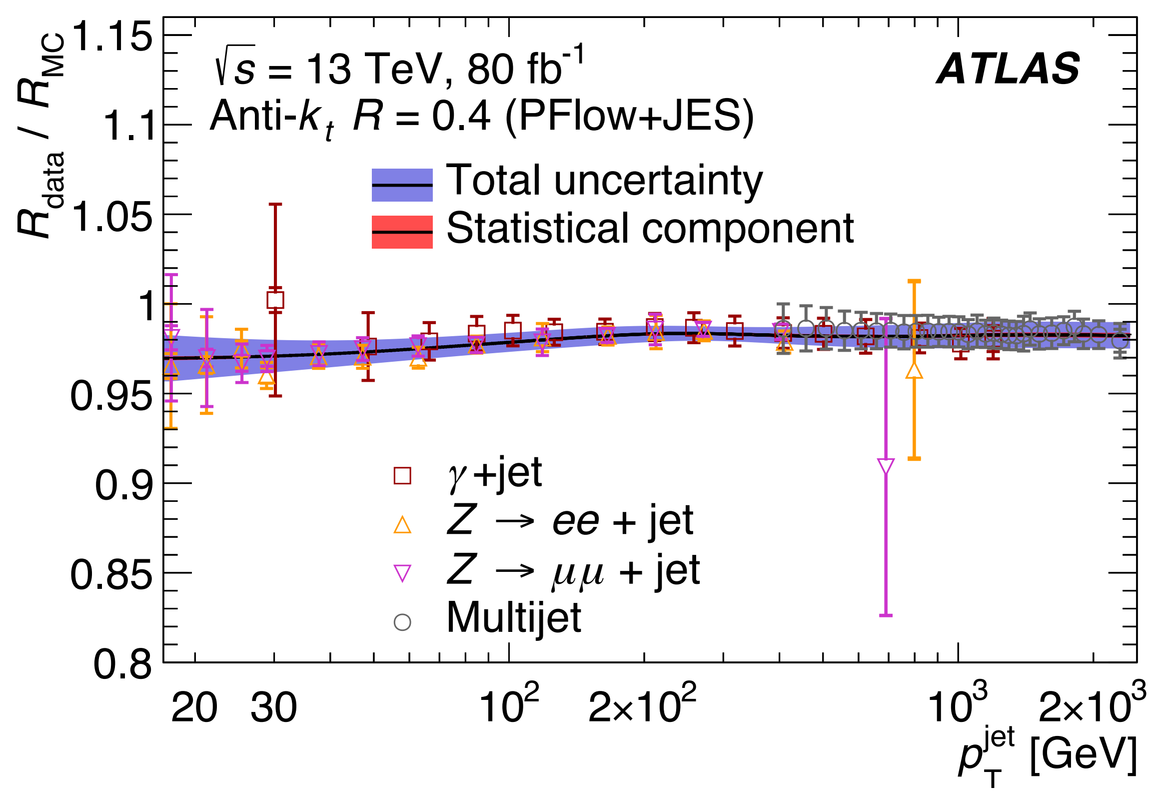
Calibration and Commissioning

Jet calibration

- Local calorimeter calibration based on test beam data and simulation
- Pileup corrections based on jet area and average energy density

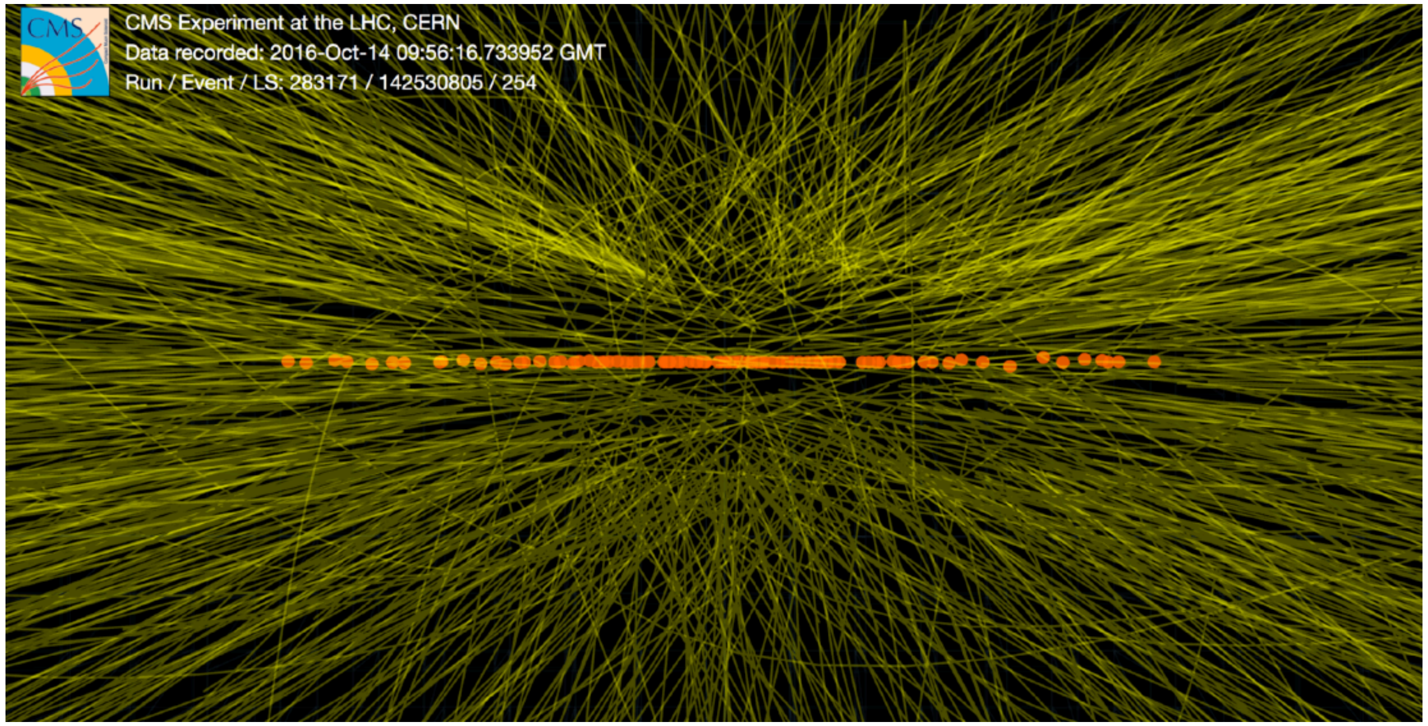


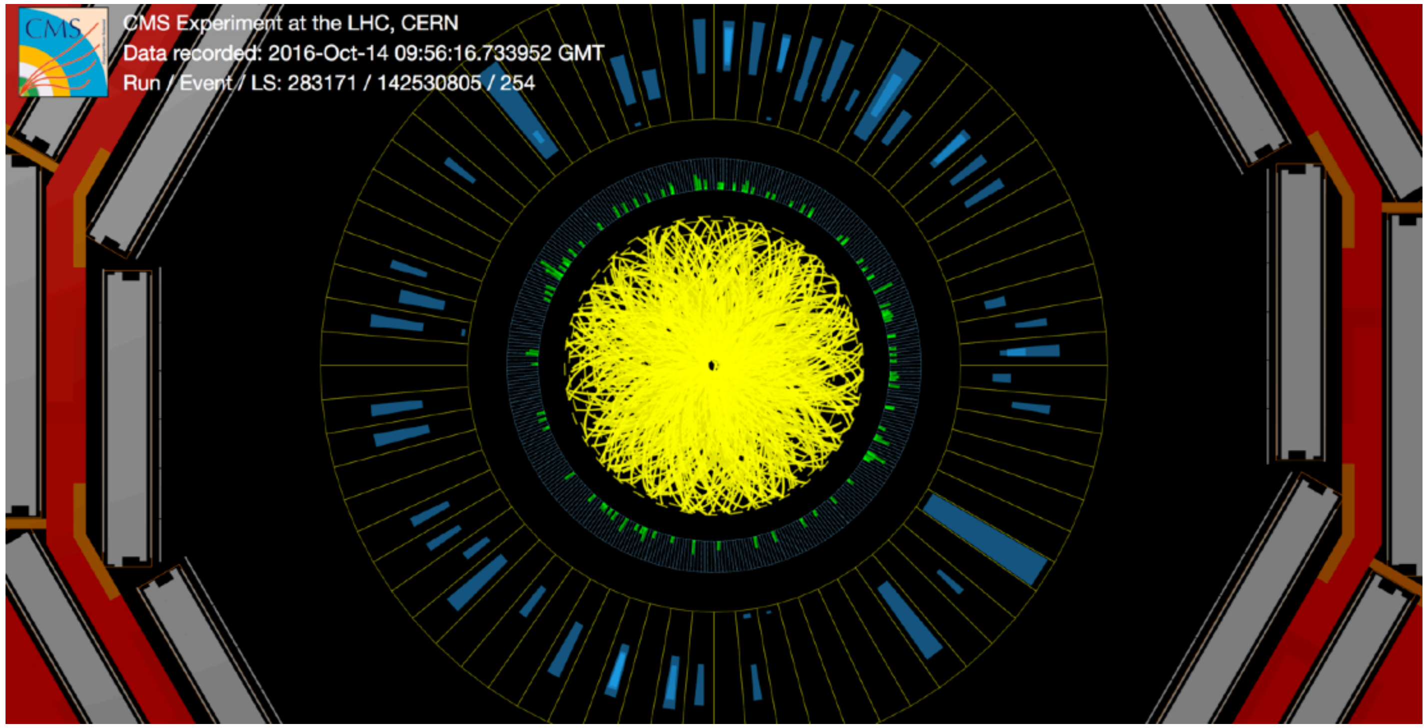
[H. Kirschenmann, M. Schröder]



[ATLAS, EPJC 81, 689 (2021)]

Pileup mitigation

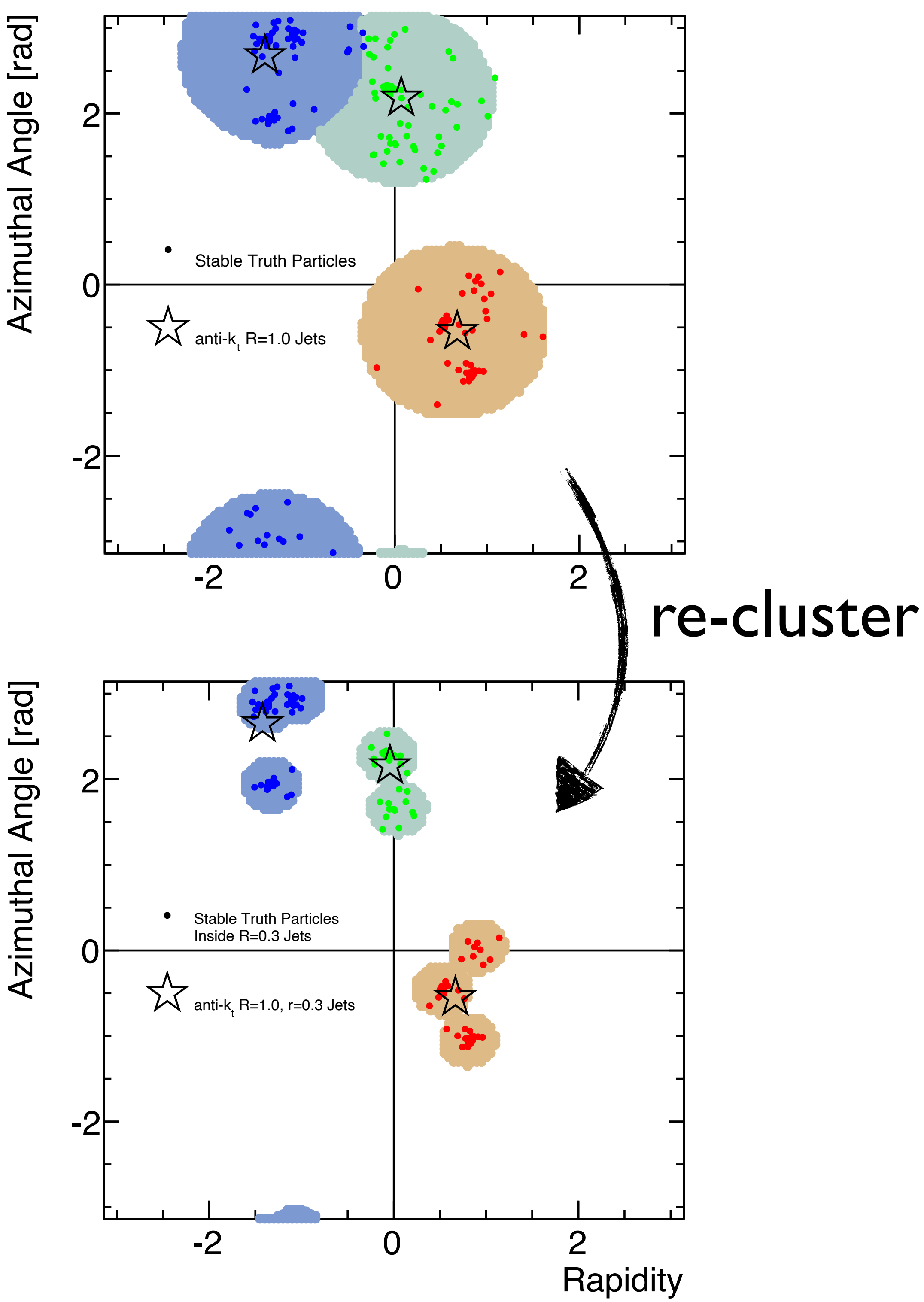




13 TeV, 2016
high-PU fill

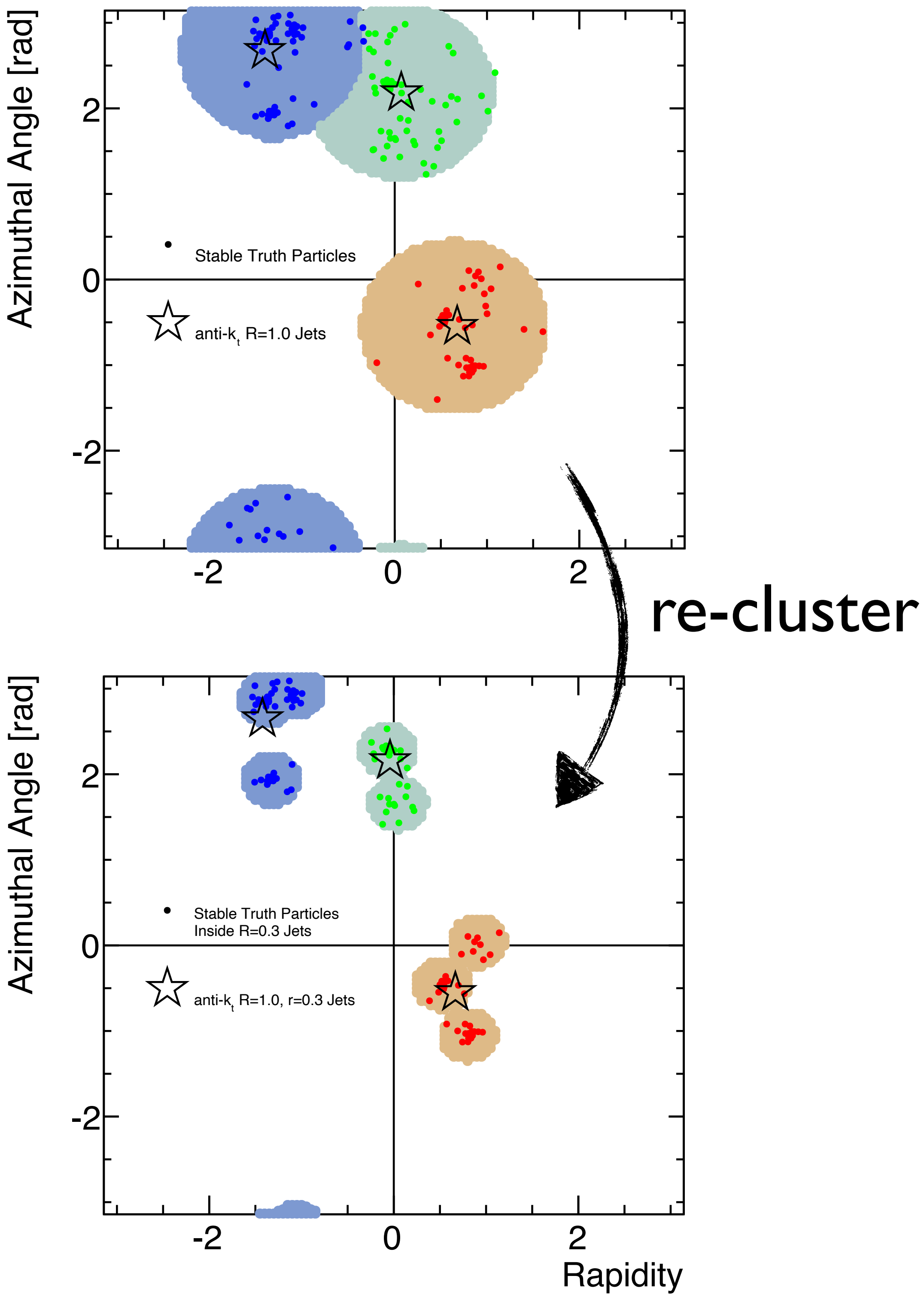
Image credit: CERN

Pileup mitigation

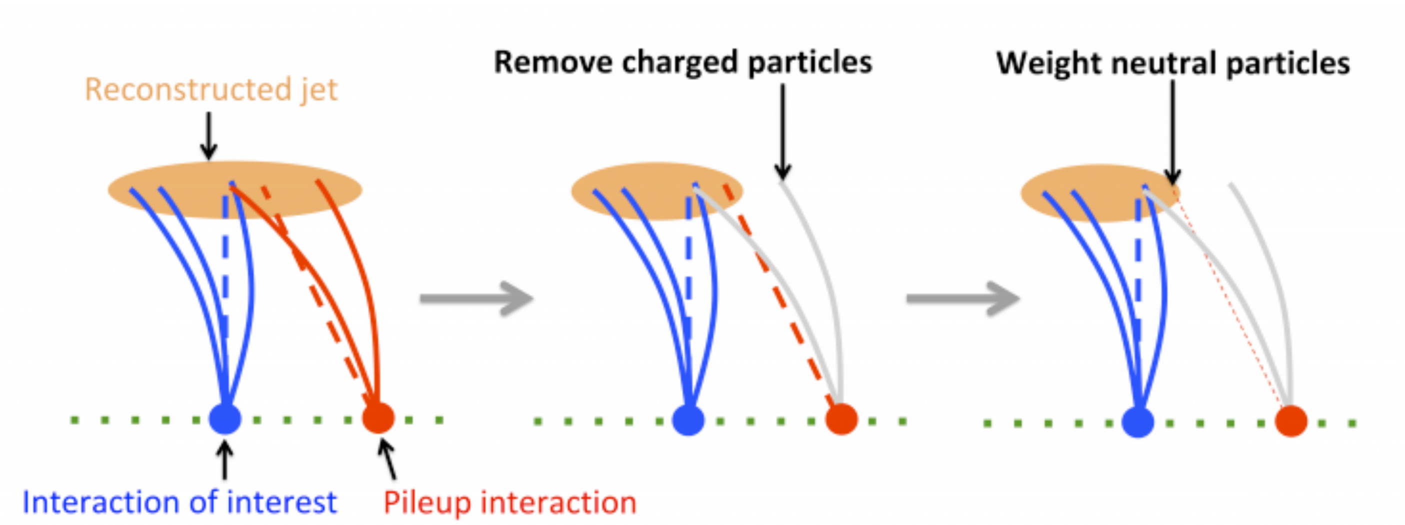


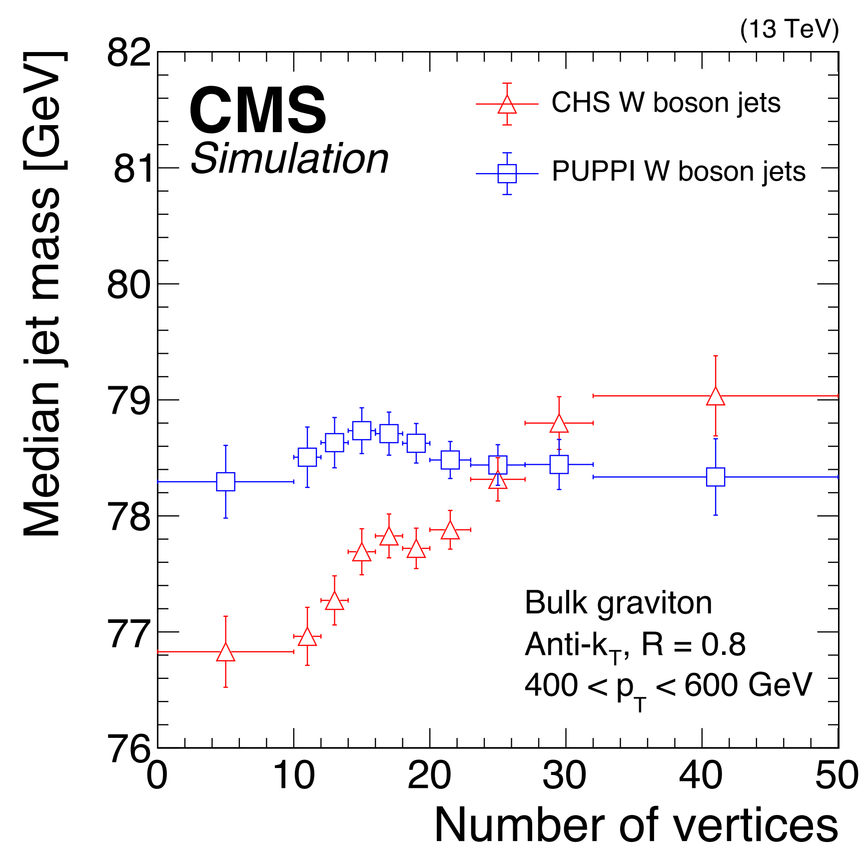
[B. Nachmann, JHEP 02, 075 (2015)]

Pileup mitigation



Pileup-per-particle identification (PUPPI)



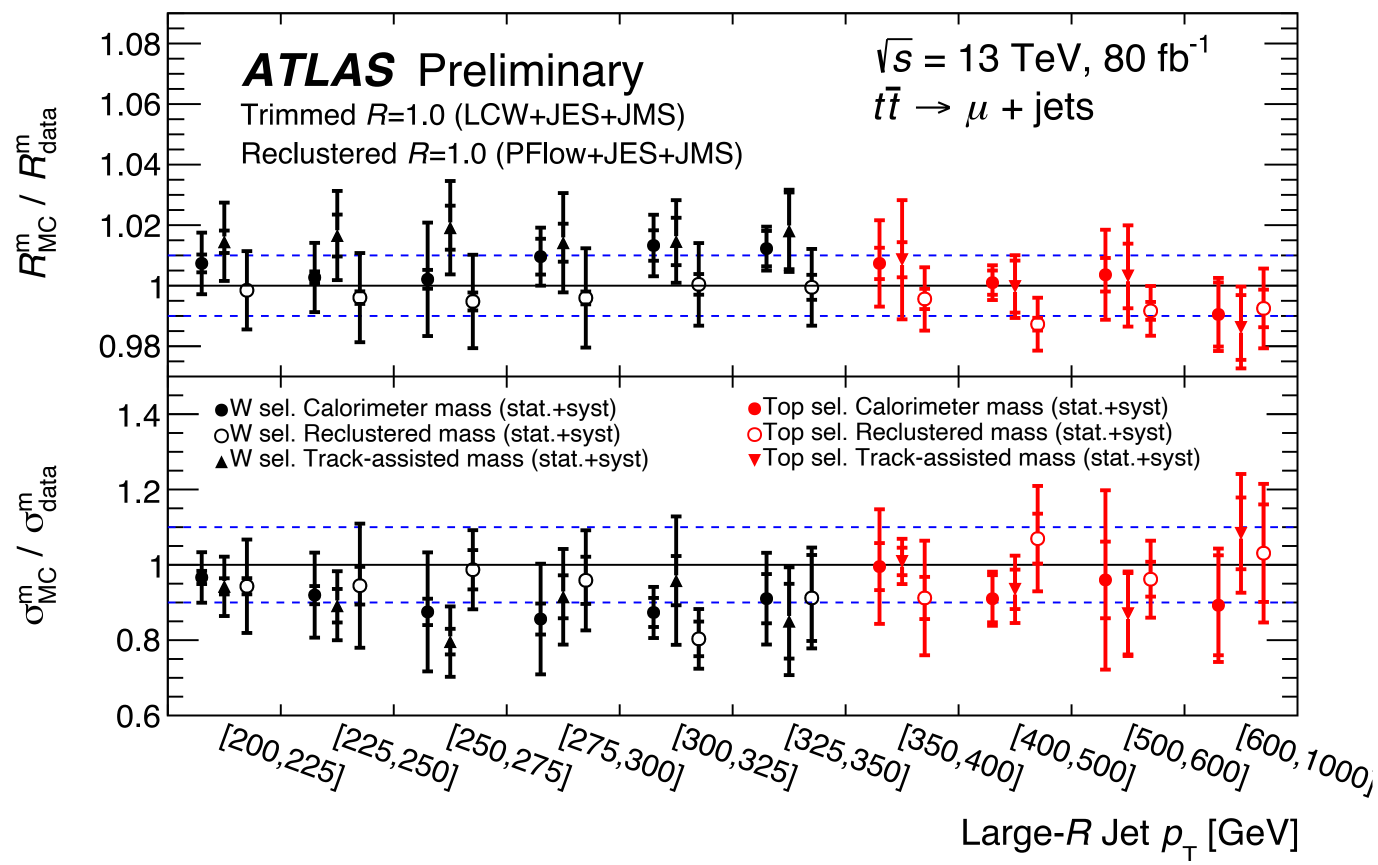


[B. Nachmann, JHEP 02, 075 (2015)]

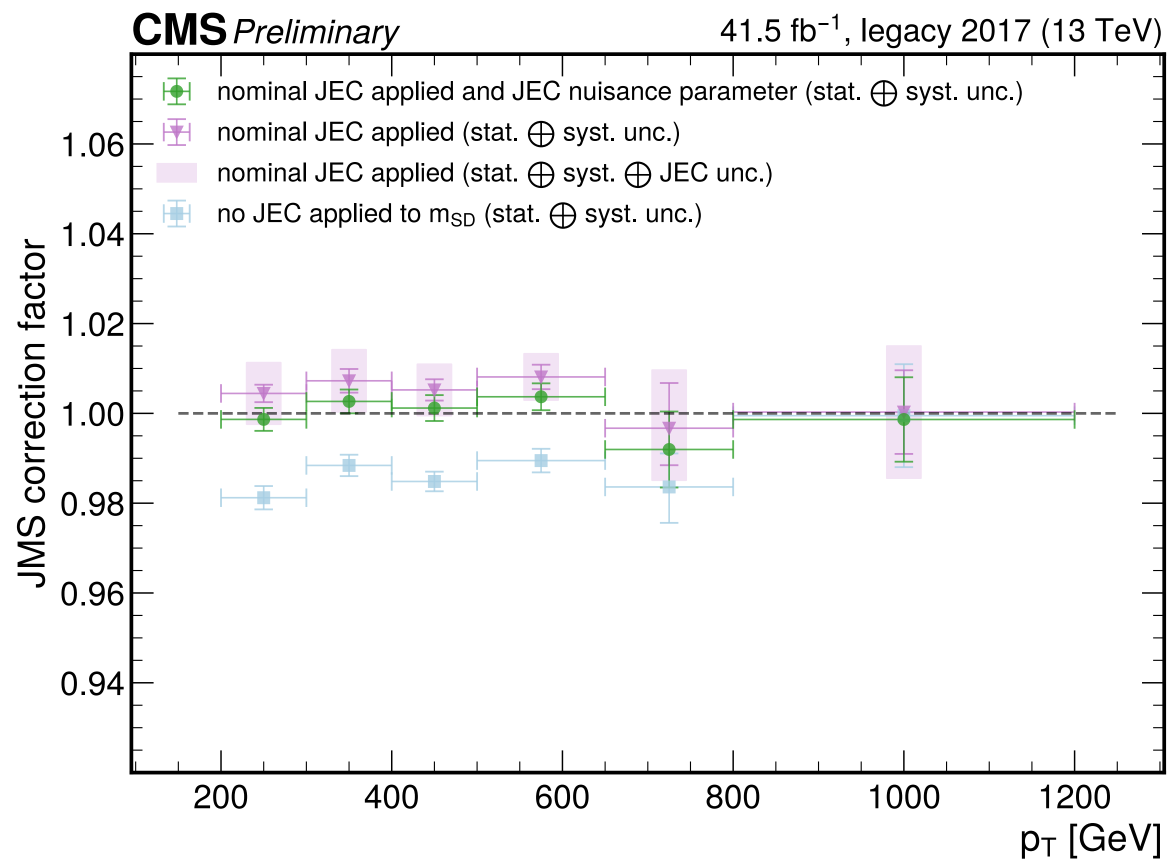
[CMS, JINST 15, P09018 (2020)]

Jet mass calibration

[ATLAS-CONF-2020-022]



[CMS, DP-2023/044]

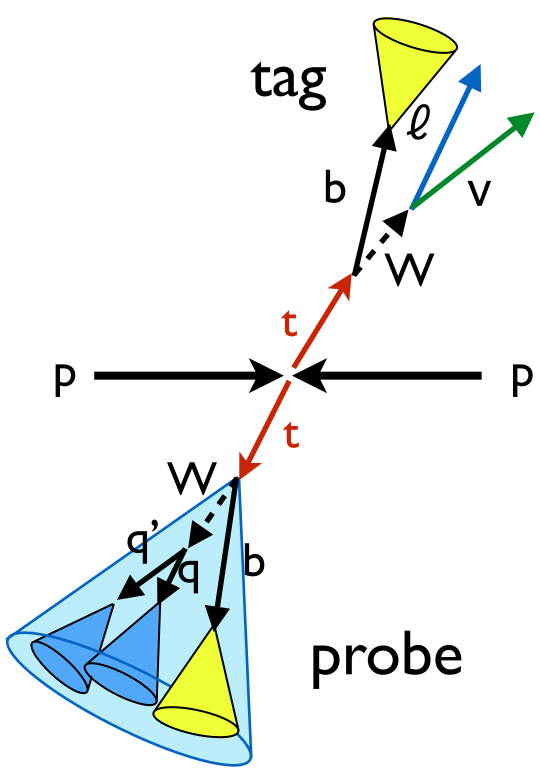


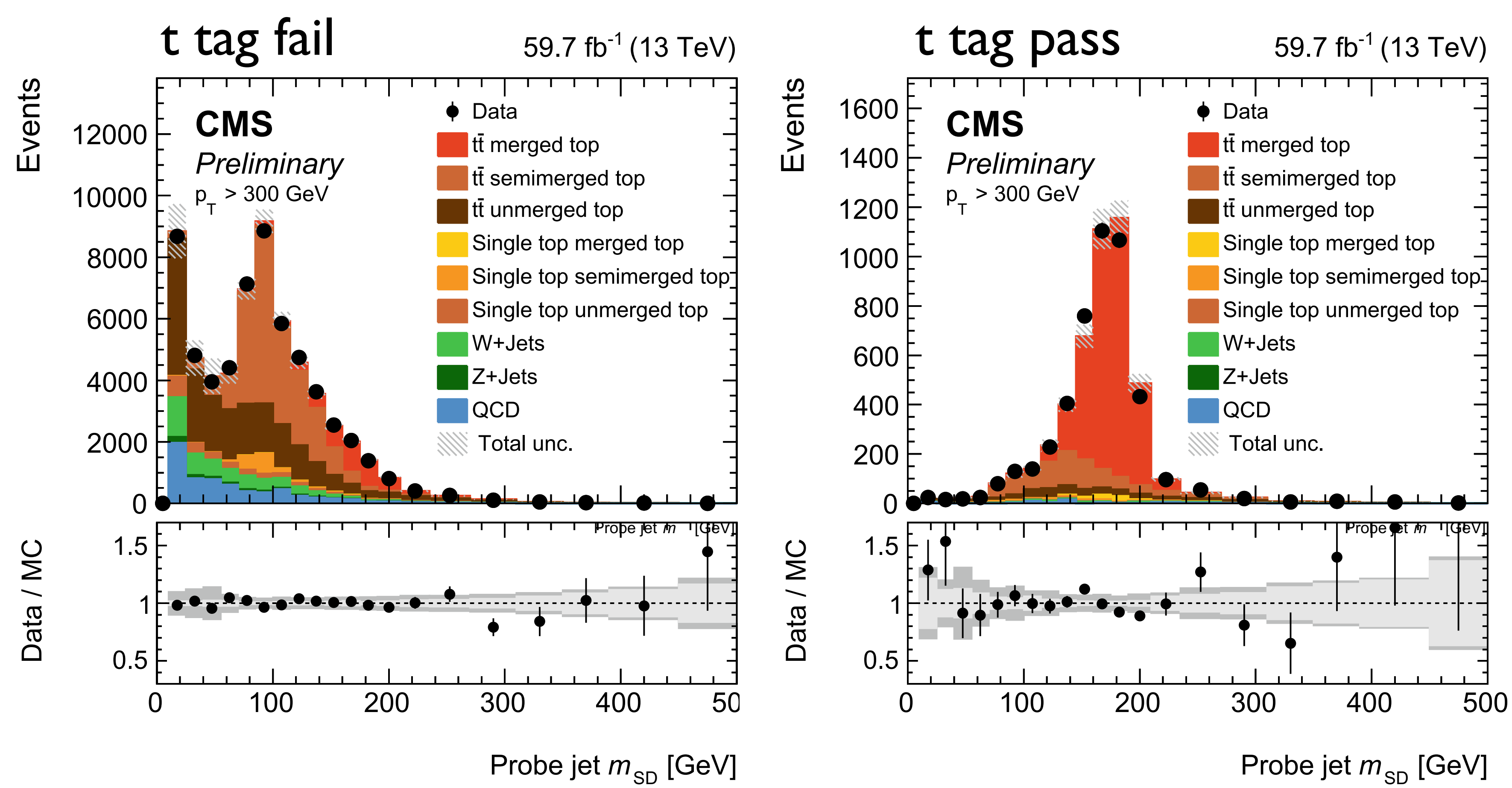
- ATLAS: Forward-folding approach for absolute scale and resolution
- CMS: Template fit to W peak in tt and W+jets
- Uncertainties of 1-2% (scale) and 10-20% (resolution) achieved

Tagging Efficiencies

Tag-and-probe measurements

- $t\bar{t}$ production for W and t tagging



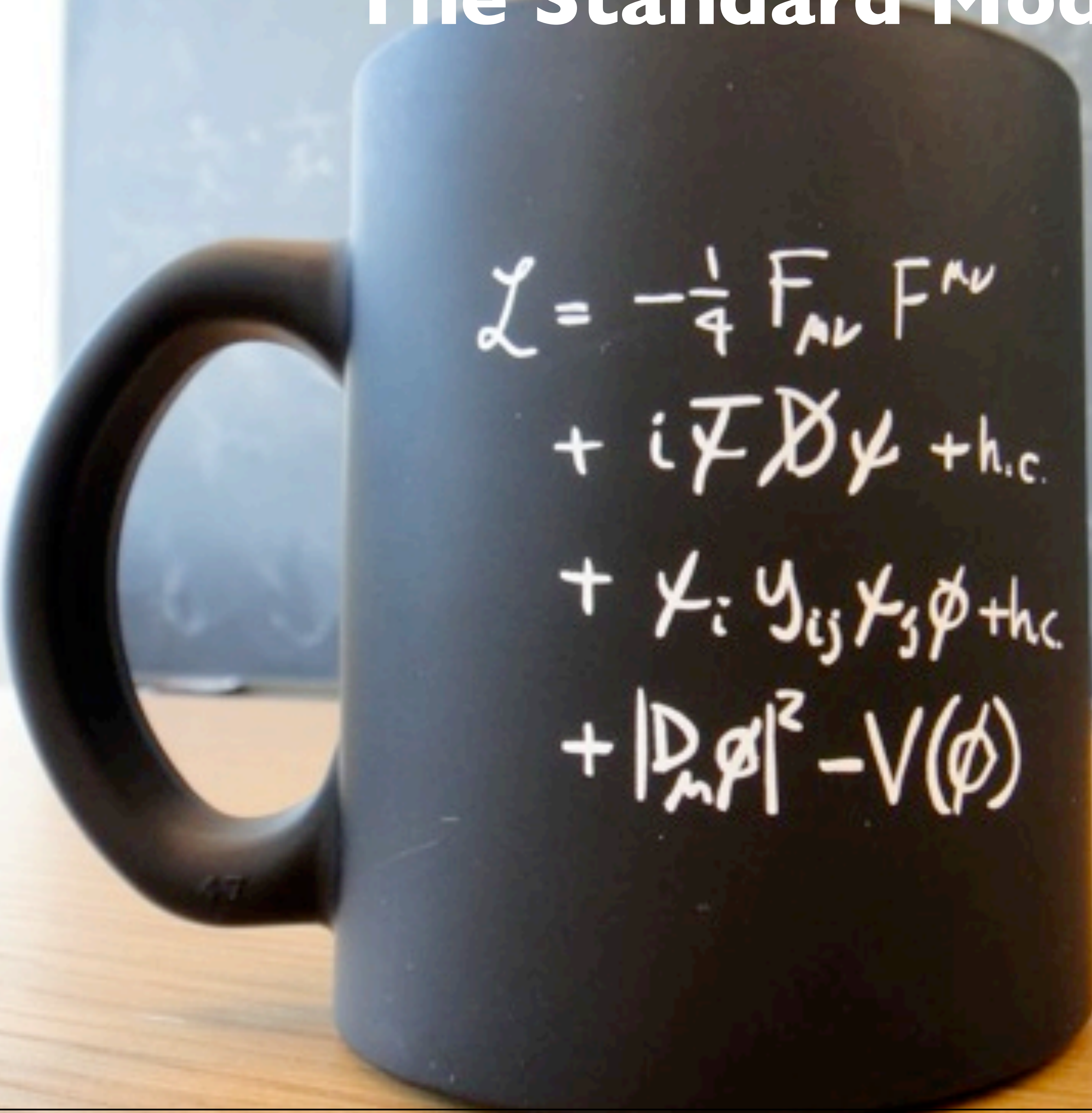


- Extrapolations to Z and H from simulation

[CMS, DP-2020/025]

Standard model measurements

The Standard Model



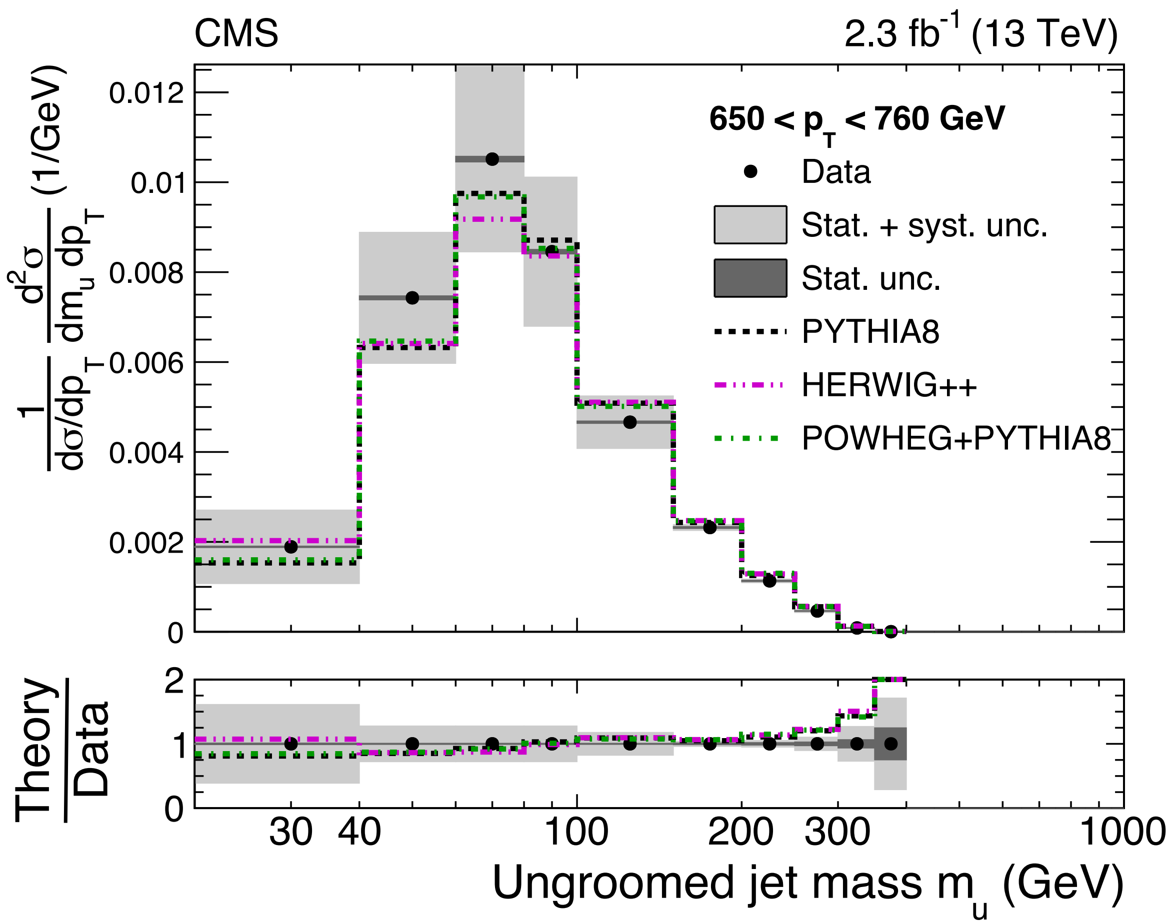
c (SLAC, Brookhaven ’74)
τ (SLAC ’75)
b (Fermilab ’77)
g (DESY, ’78-79)
W/Z (CERN ’83)
t (Fermilab ’95)

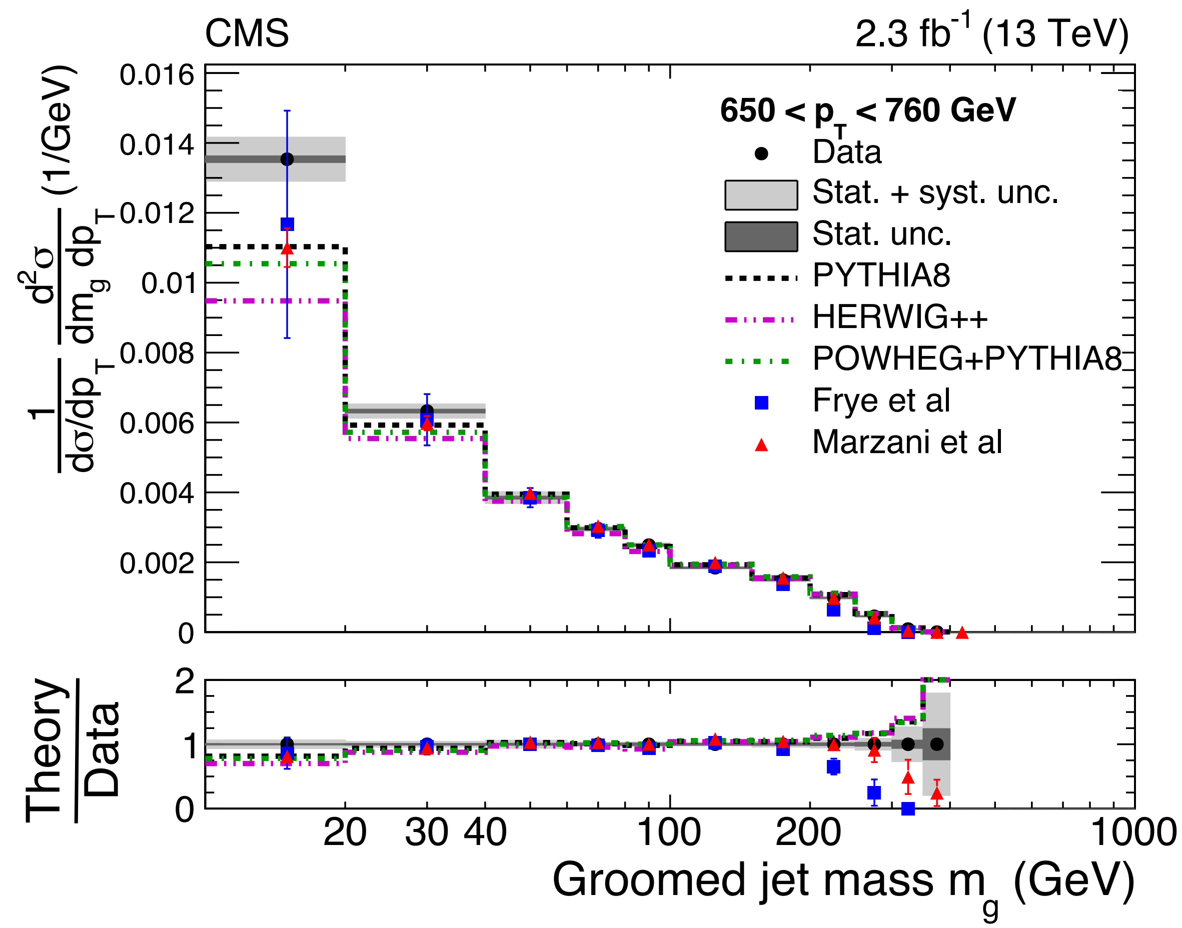
… did not mention the ν sector

Yukawa interactions (CERN ’16-18)

H (CERN ’12) and its gauge interactions

Jet mass of q/g jets

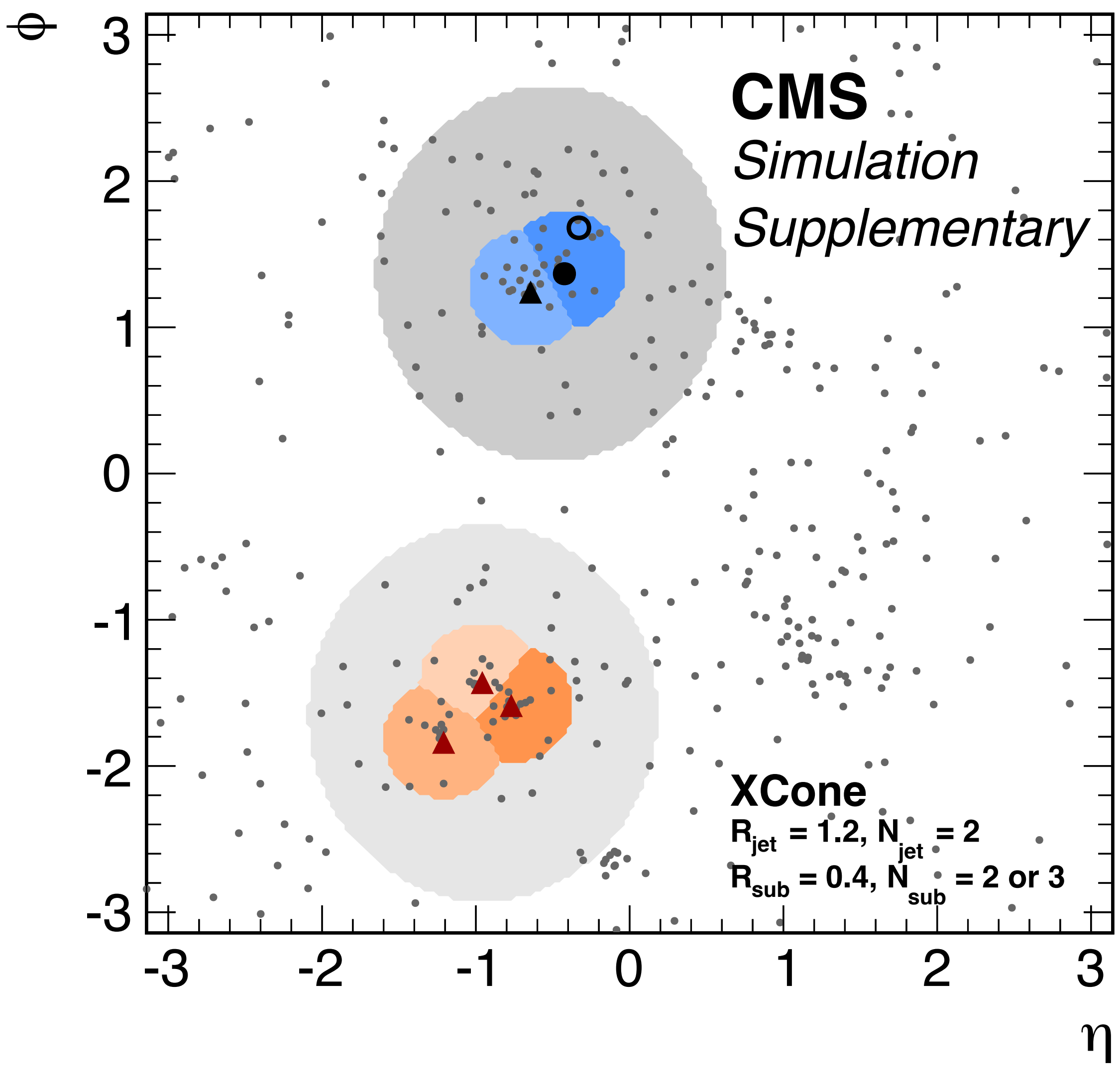


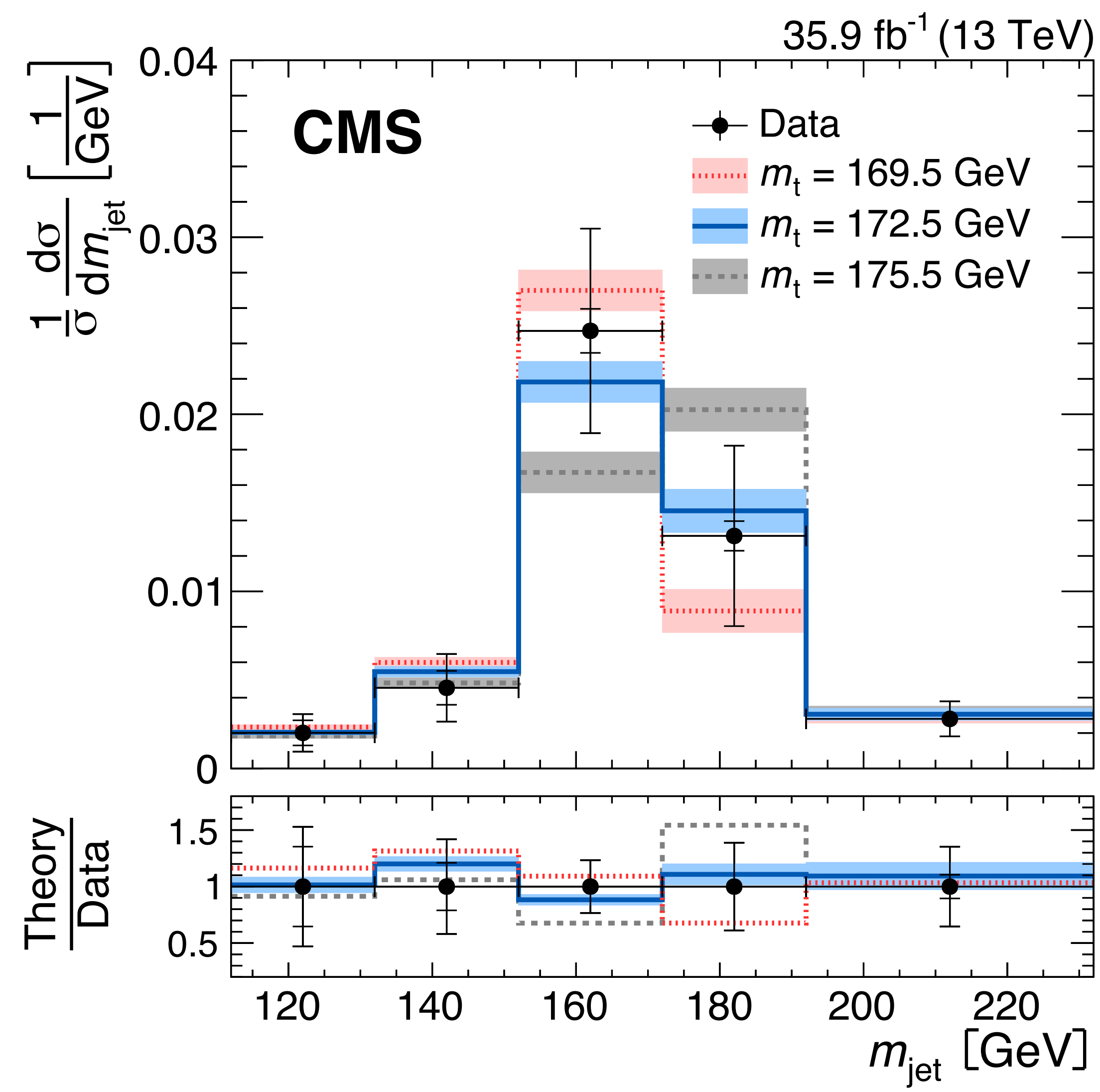


- Sudakov peak shifts to smaller values for soft drop grooming
- Soft drop grooming reduces modelling uncertainties

[CMS, JHEP 11, 113 (2018)]

Jet mass of top quark jets

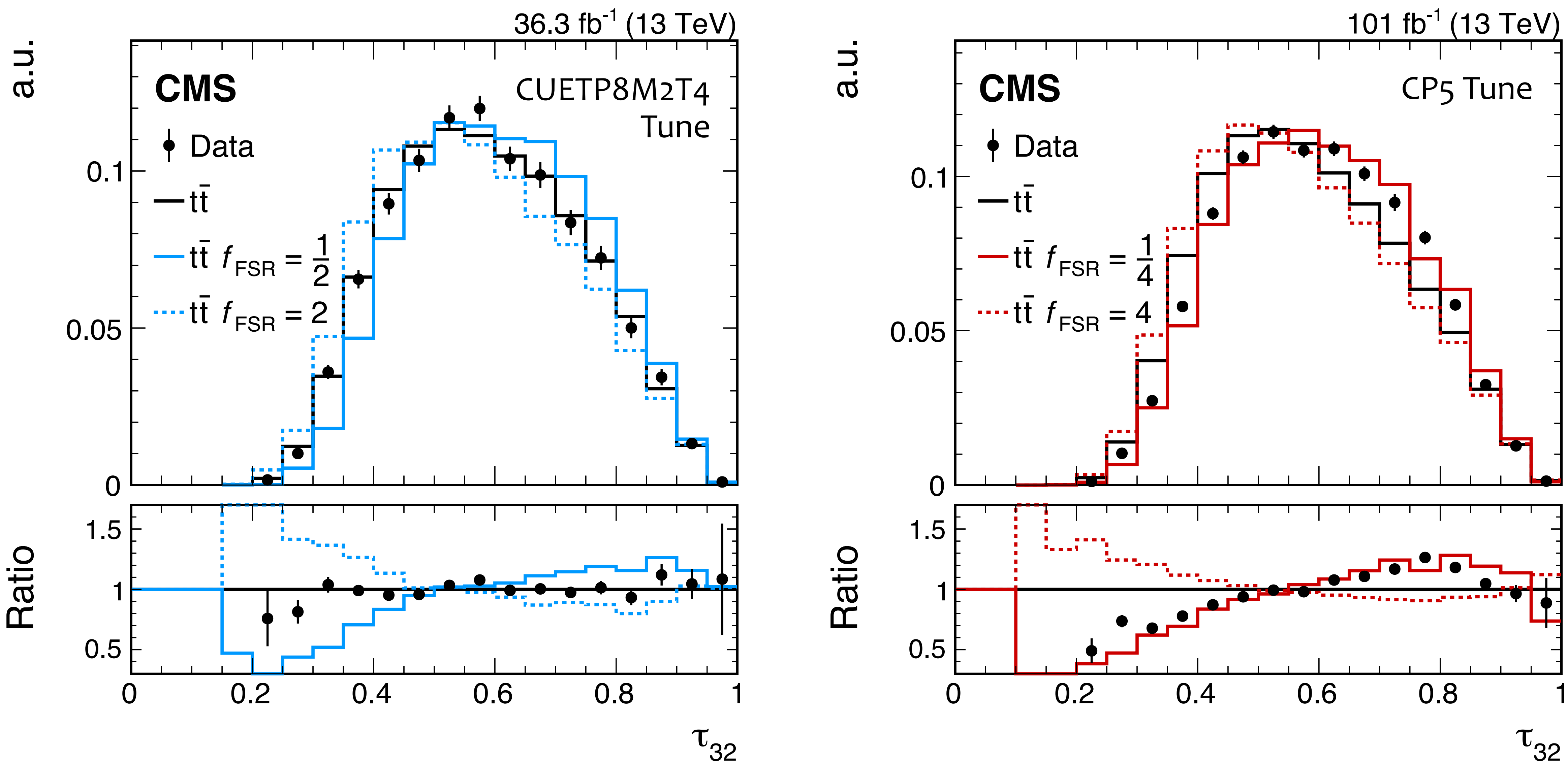




- XCone clustering improves mass resolution by a factor of two
- Measurement of top quark mass: m_t = 172.6 ± 2.5 GeV
  - Equal size of stat, modelling and JES uncertainties

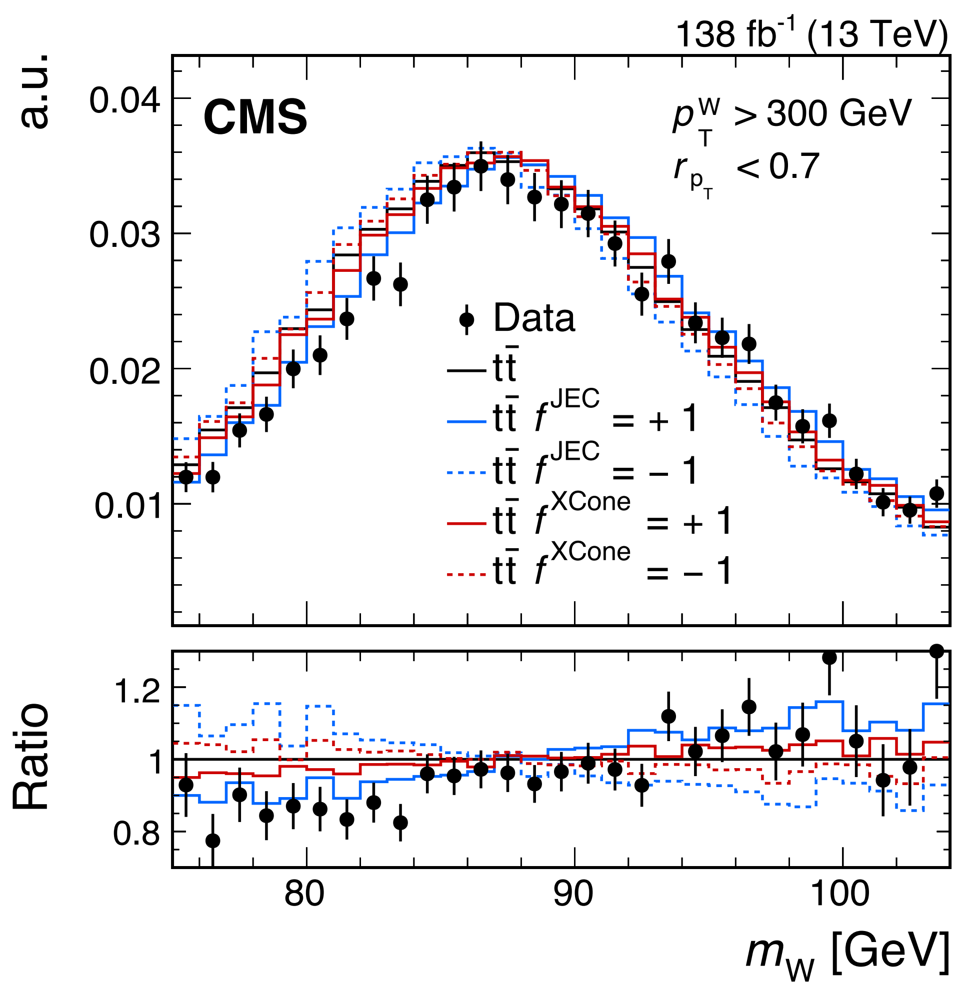
[CMS, PRL 124, 202001 (2020)]

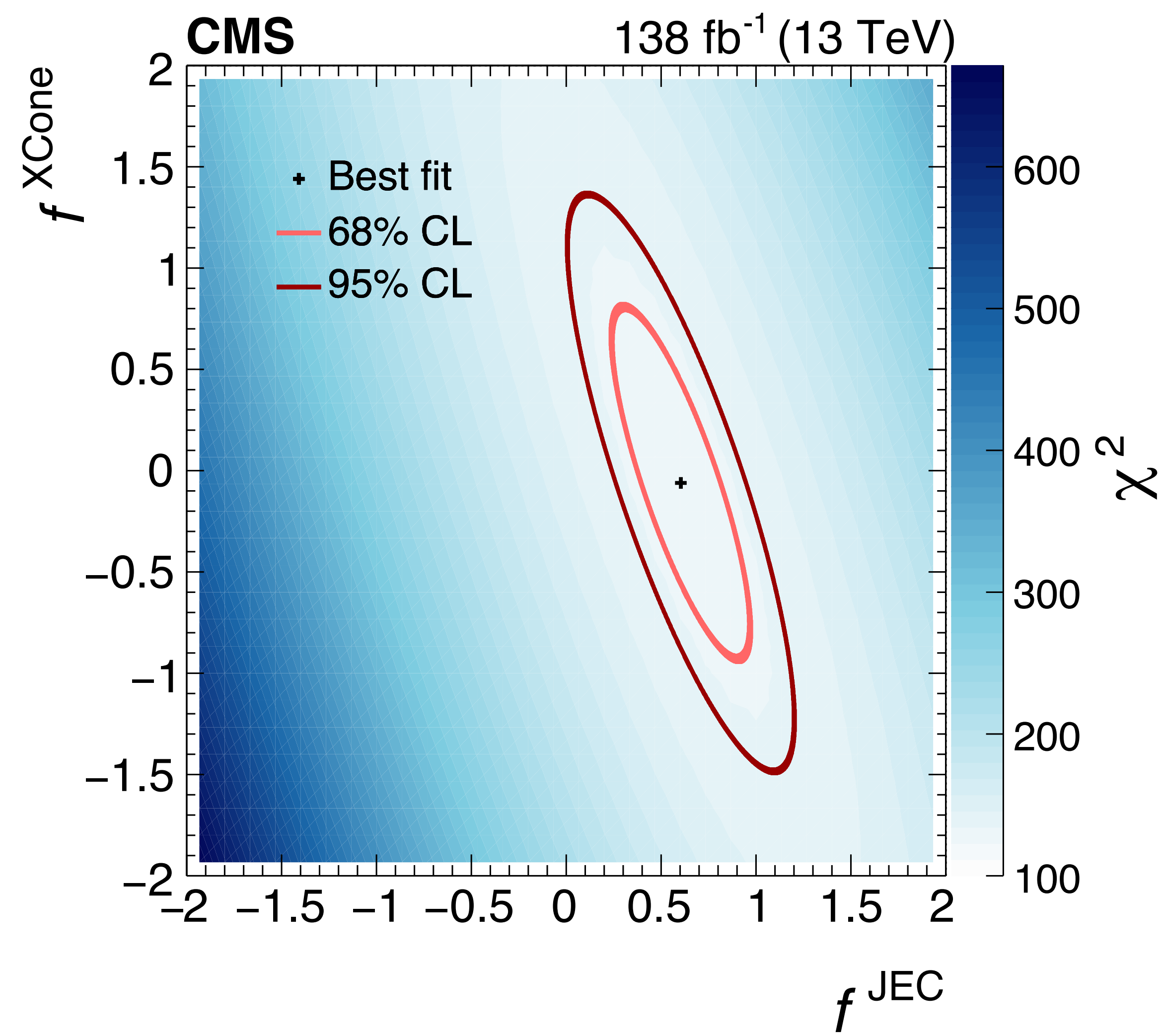
Constraining the Parton Shower



- Adjust f_{FSR} in $\alpha_S^{FSR}(f_{FSR}\,\mu_0)$, equivalent to choosing different $\alpha_S^{FSR}(M_Z)$

Jet mass calibration

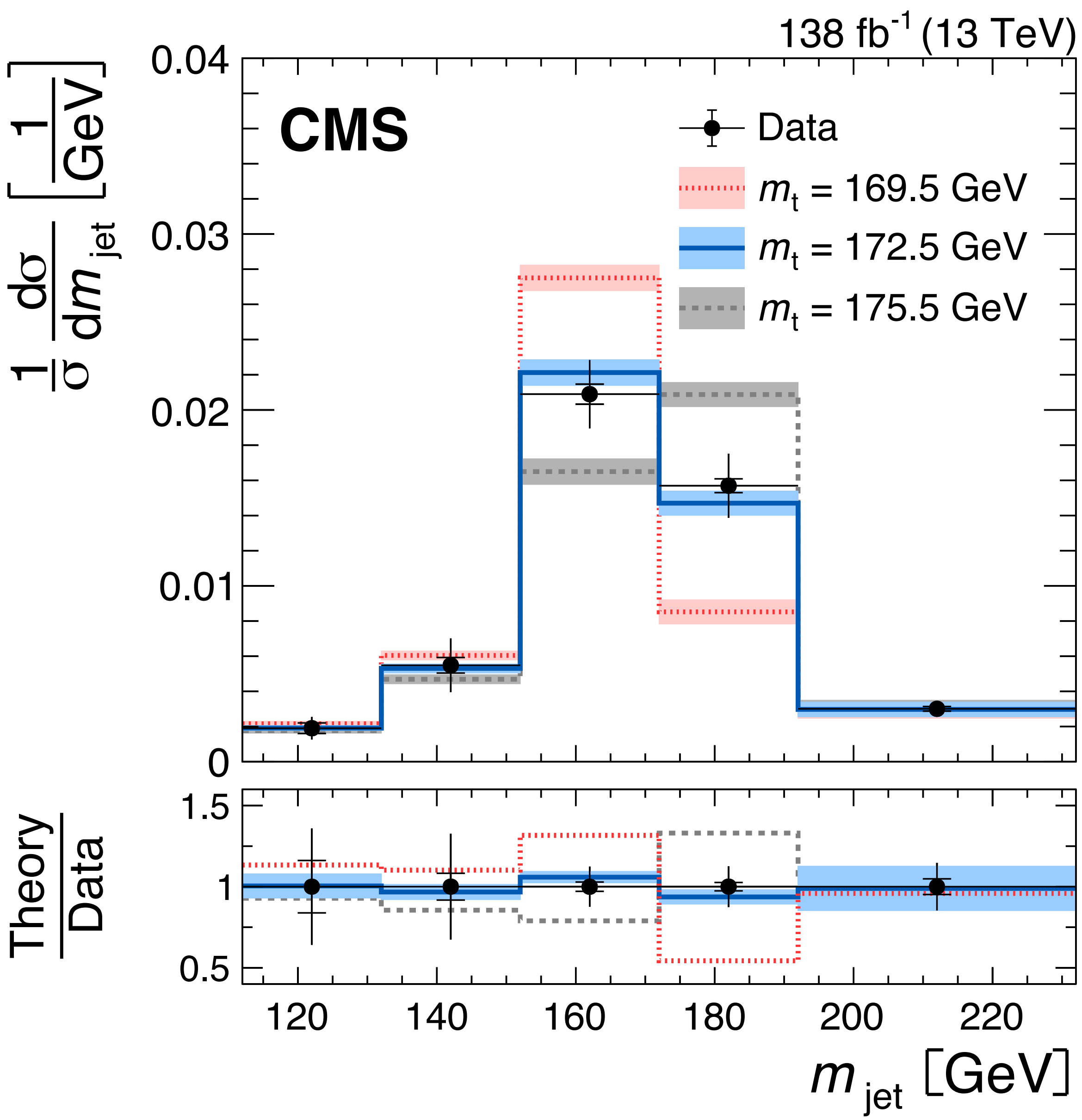




- Reconstruct m_W using the two light-flavoured subjets
- Measure m_W in four regions
- Measure jet mass scale (JMS) by adjusting XCone and JES corrections

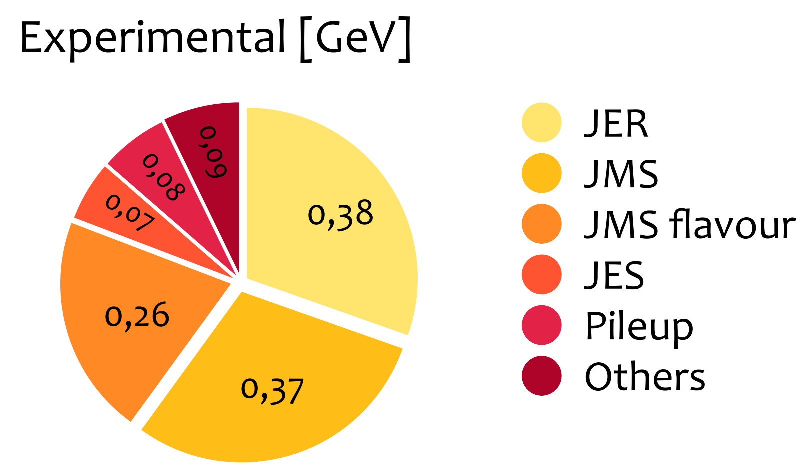
[CMS, EPJC 83, 560 (2023)]

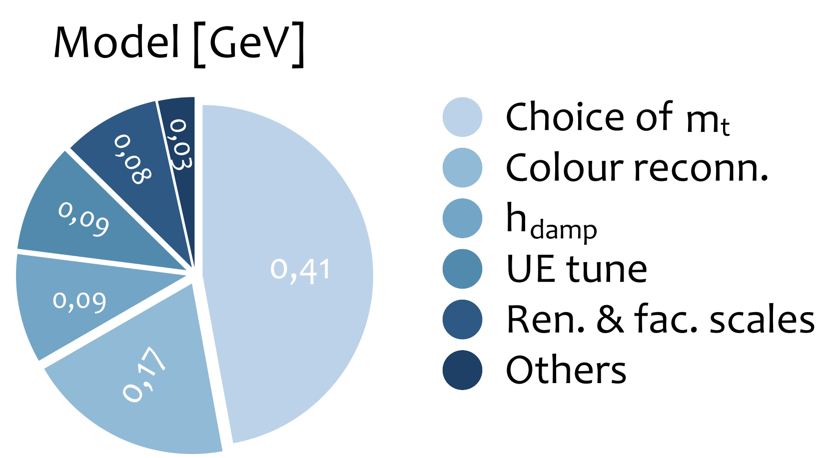
After many improvements…



$$m_t = 173.06 \pm 0.84\,\text{GeV}$$

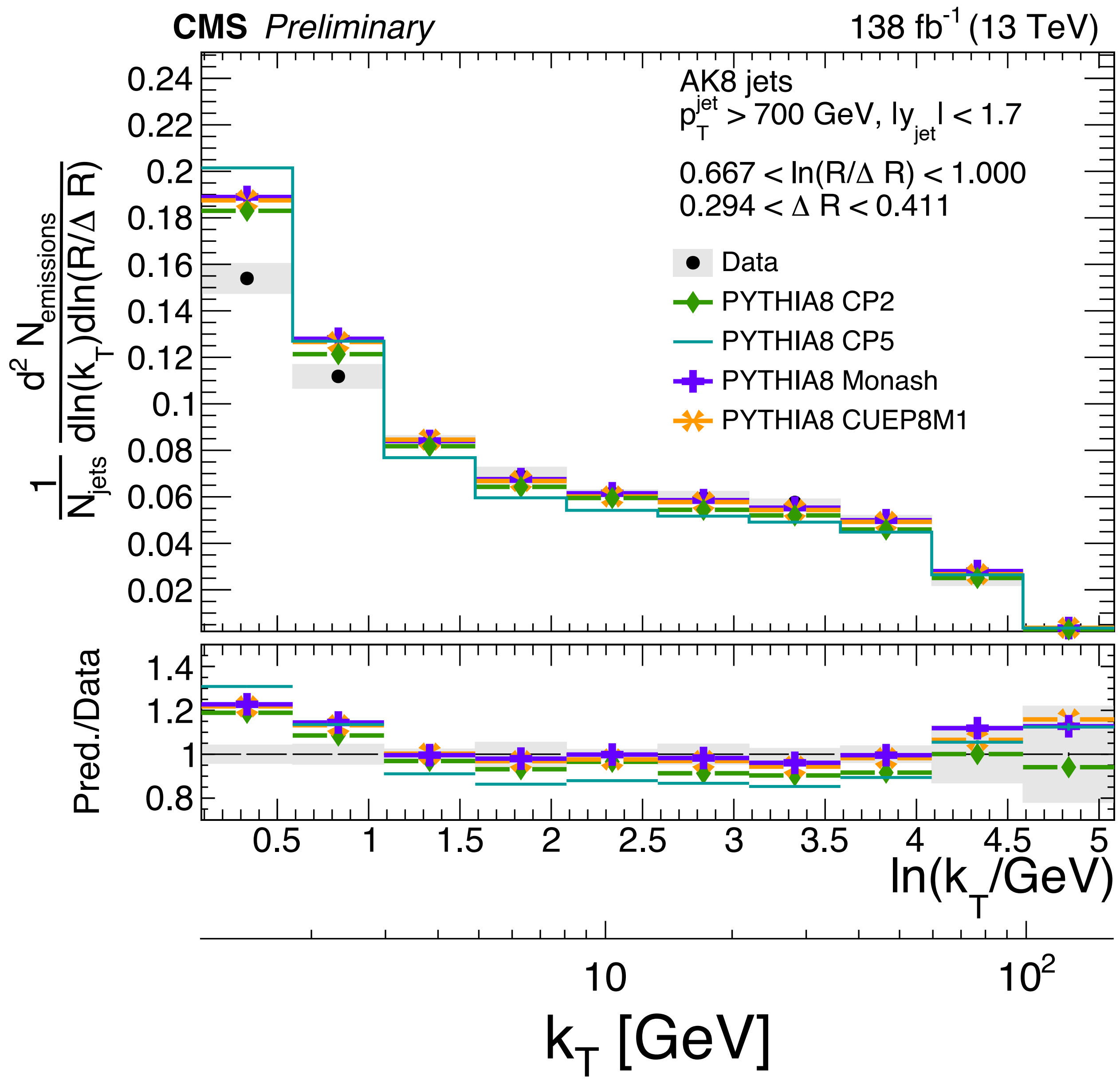
± 0.24 (stat) ± 0.61 (exp) ± 0.47 (model) ± 0.23 (theo) GeV

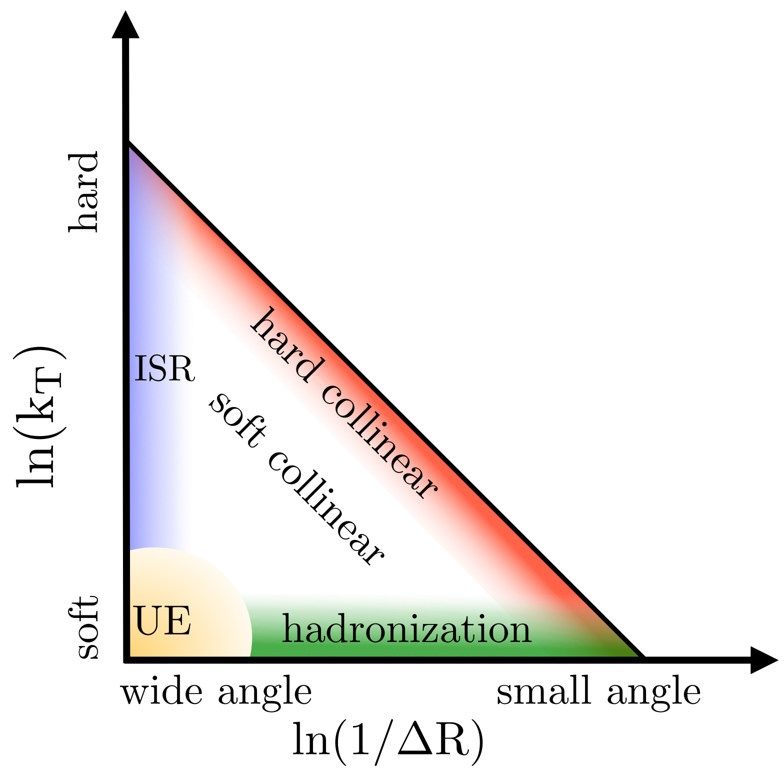




[CMS, EPJC 83, 560 (2023)]

Dedicated Measurements



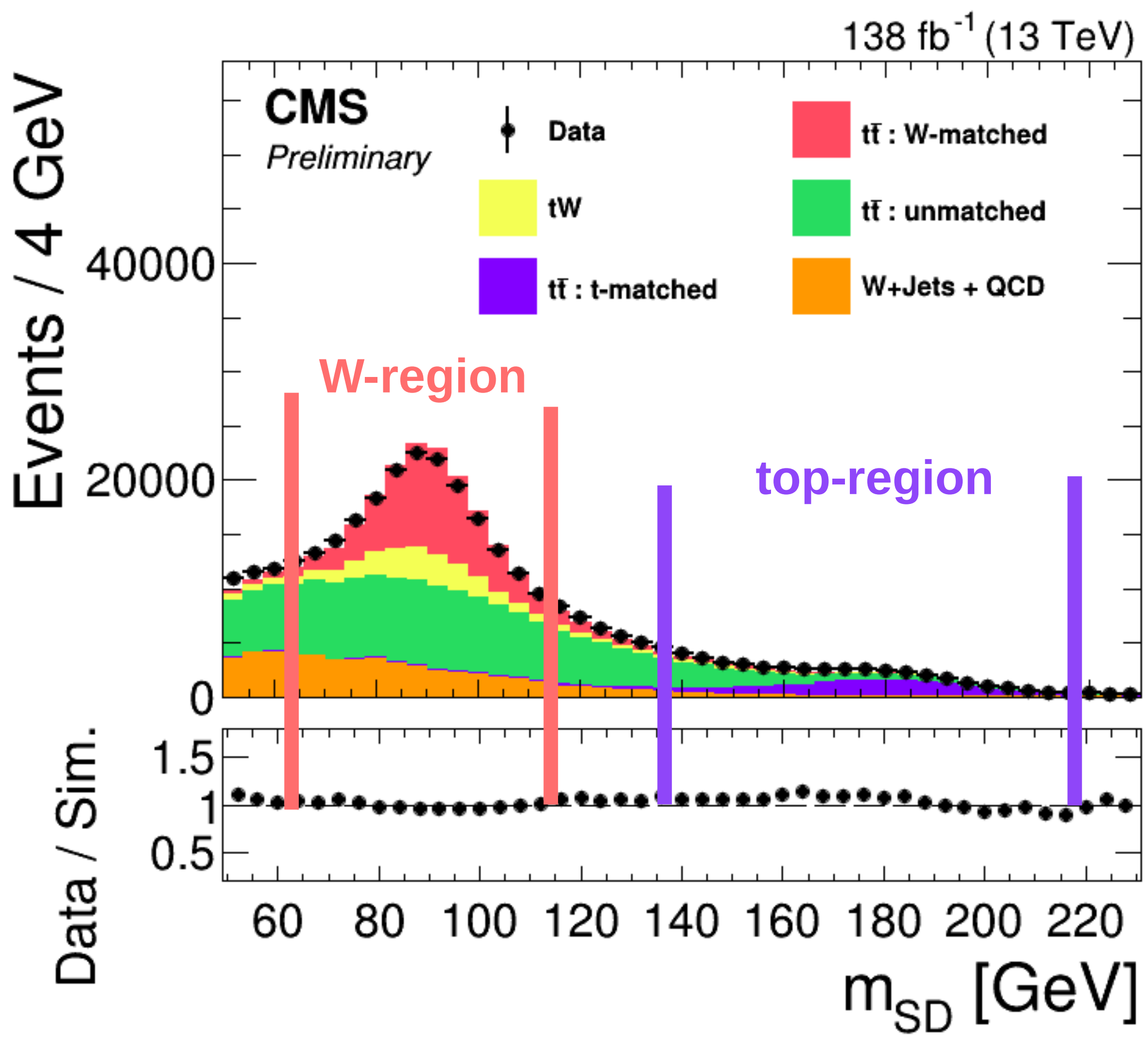


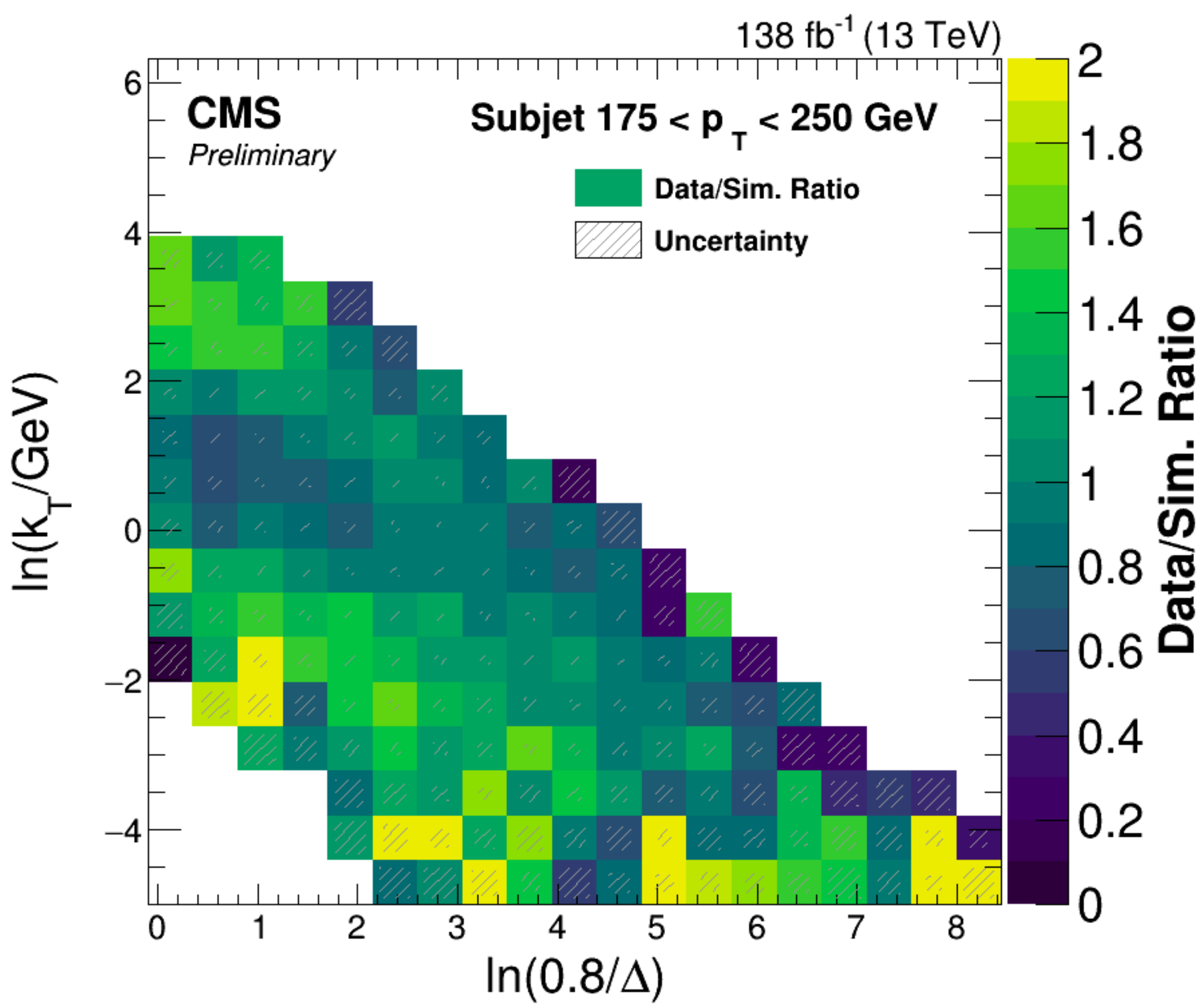
- Unfolded measurement of Lund jet plane
- Can be used to test new PS / hadronization models and tunes

[CMS, PAS-SMP-22-007 (2023)]

Ad-hoc Corrections with LJP

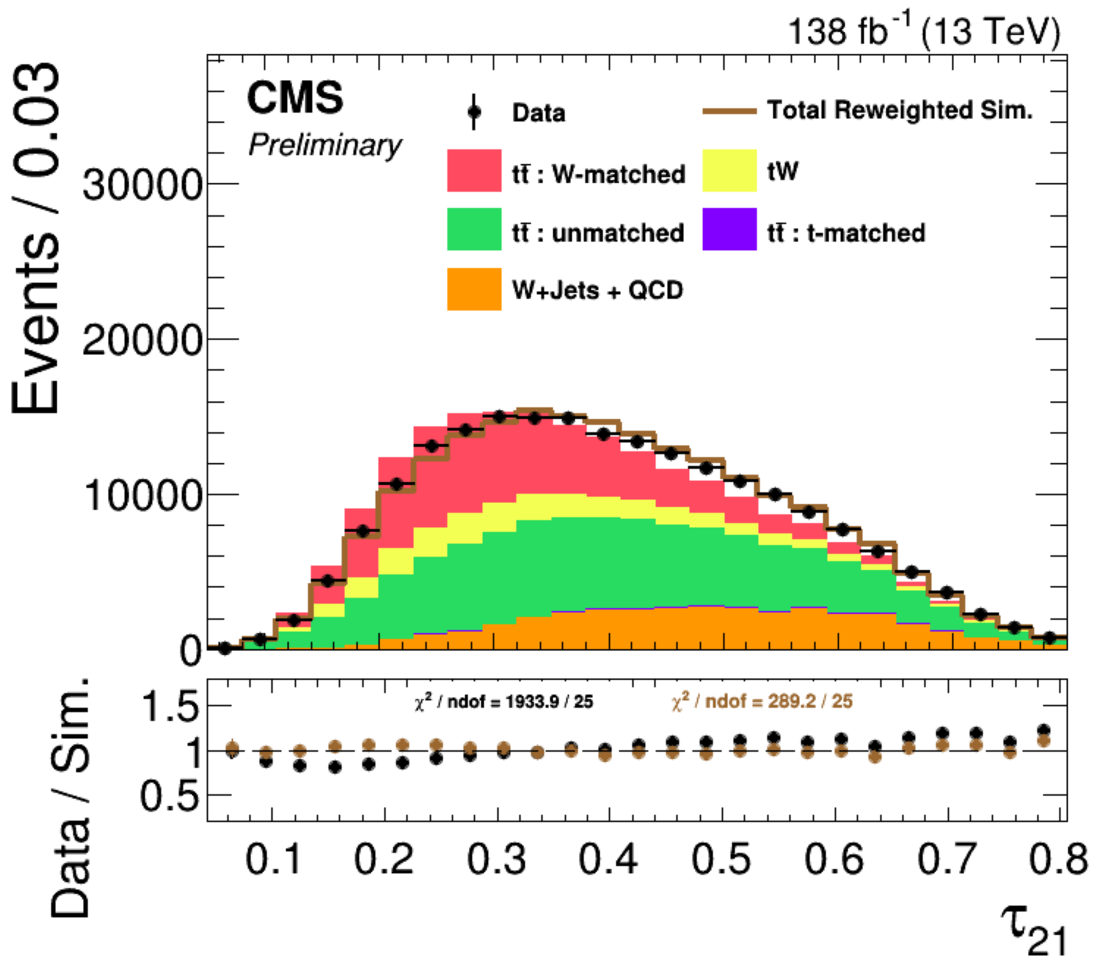
- Derive corrections 'per prong' to calibrate $N > 3$ prong jets
- Use merged W two-prongs to derive corrections, test on top quarks

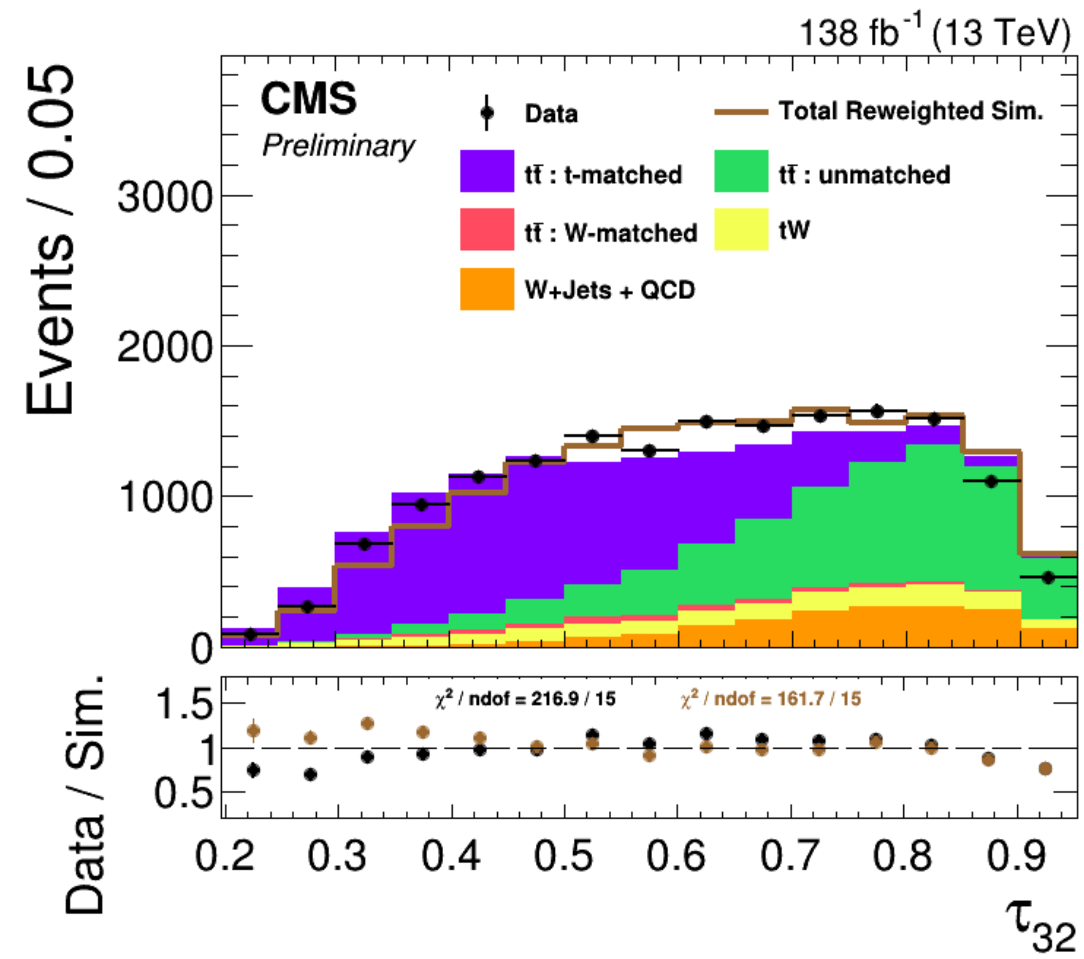




[CMS, DP-2023/046]

Ad-hoc Corrections with LJP



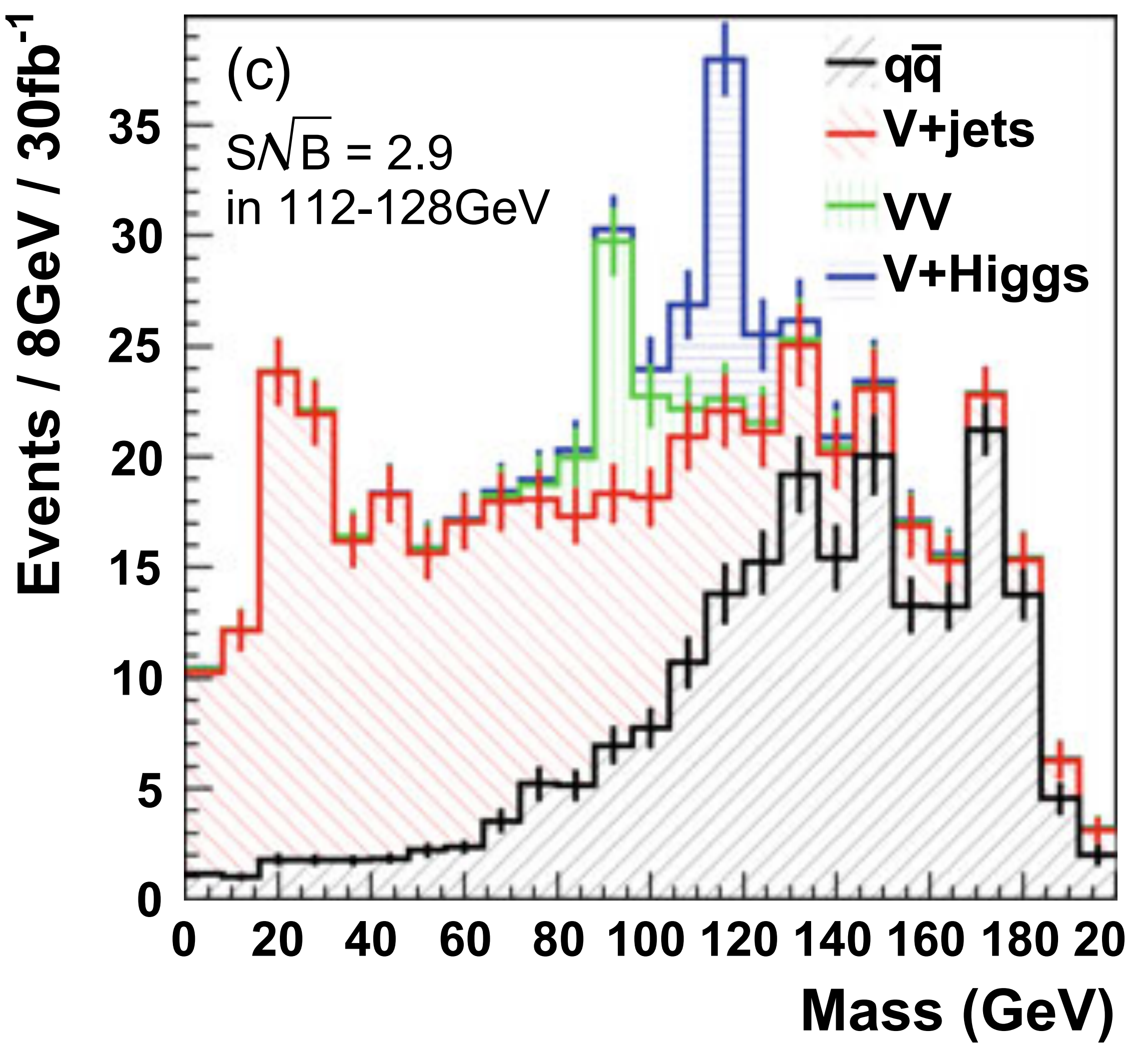


- Correction works perfectly for boosted Ws
- Improvement for boosted top quarks, but uncertainties of 15-20%
- Best way to calibrate high-prong decays (H→4q, BSM)

[CMS, DP-2023/046]

Higgs production in VH($b\bar{b}$)

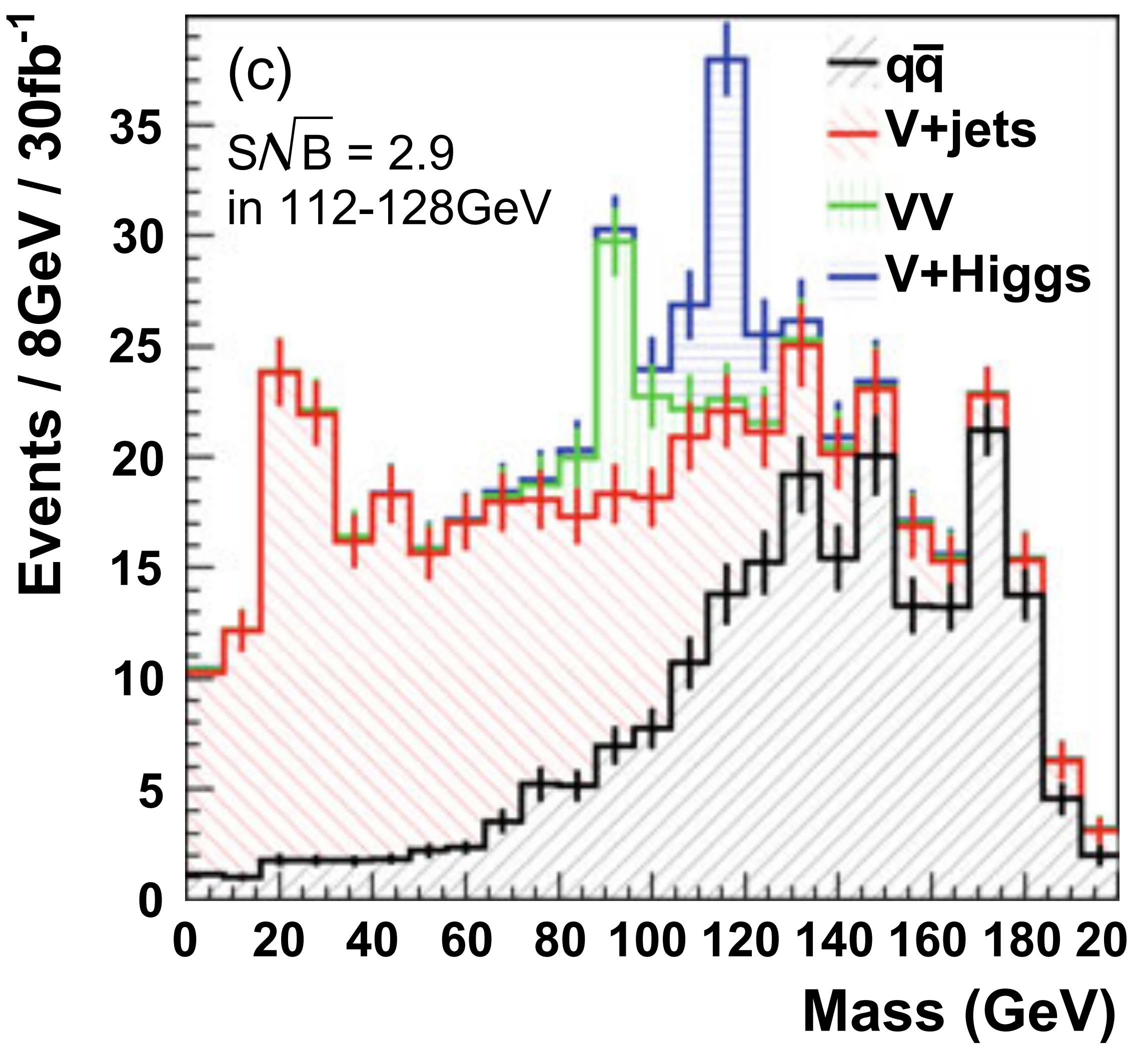
2008



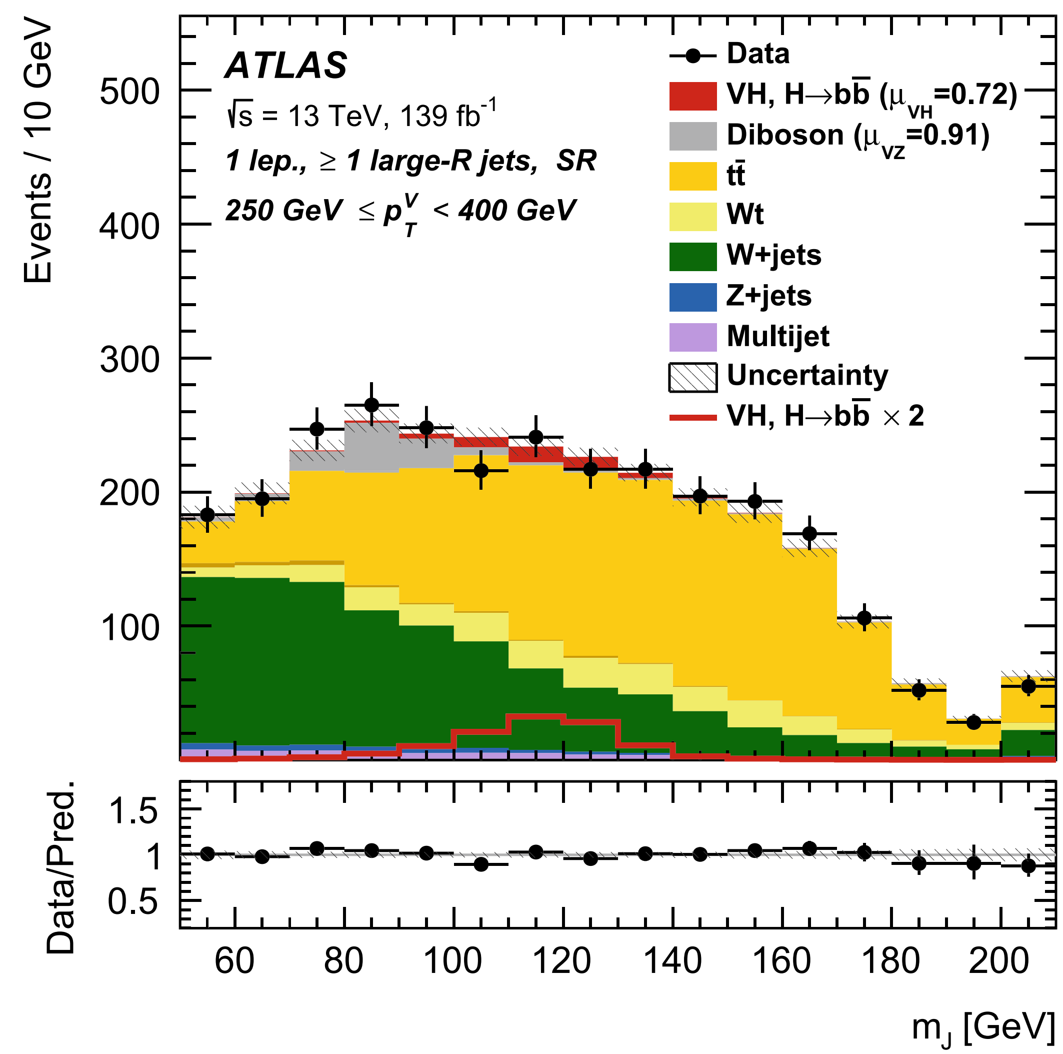
[J. M. Butterworth et al, PRL 100, 242001 (2008)]

Higgs production in VH($b\bar{b}$)

2008



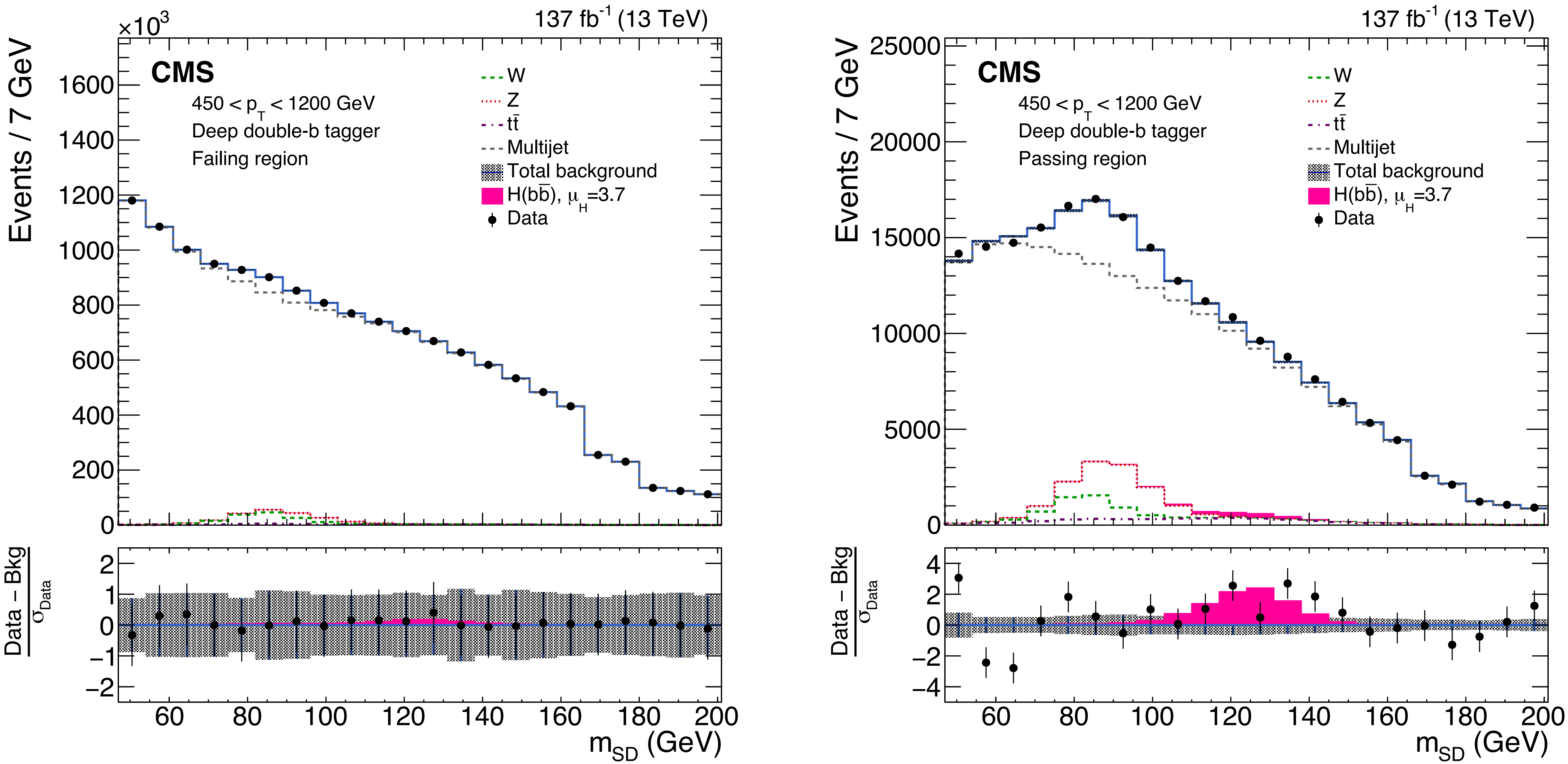
2020



[J. M. Butterworth et al, PRL 100, 242001 (2008)]

[ATLAS, PLB 816, 136204 (2021)]

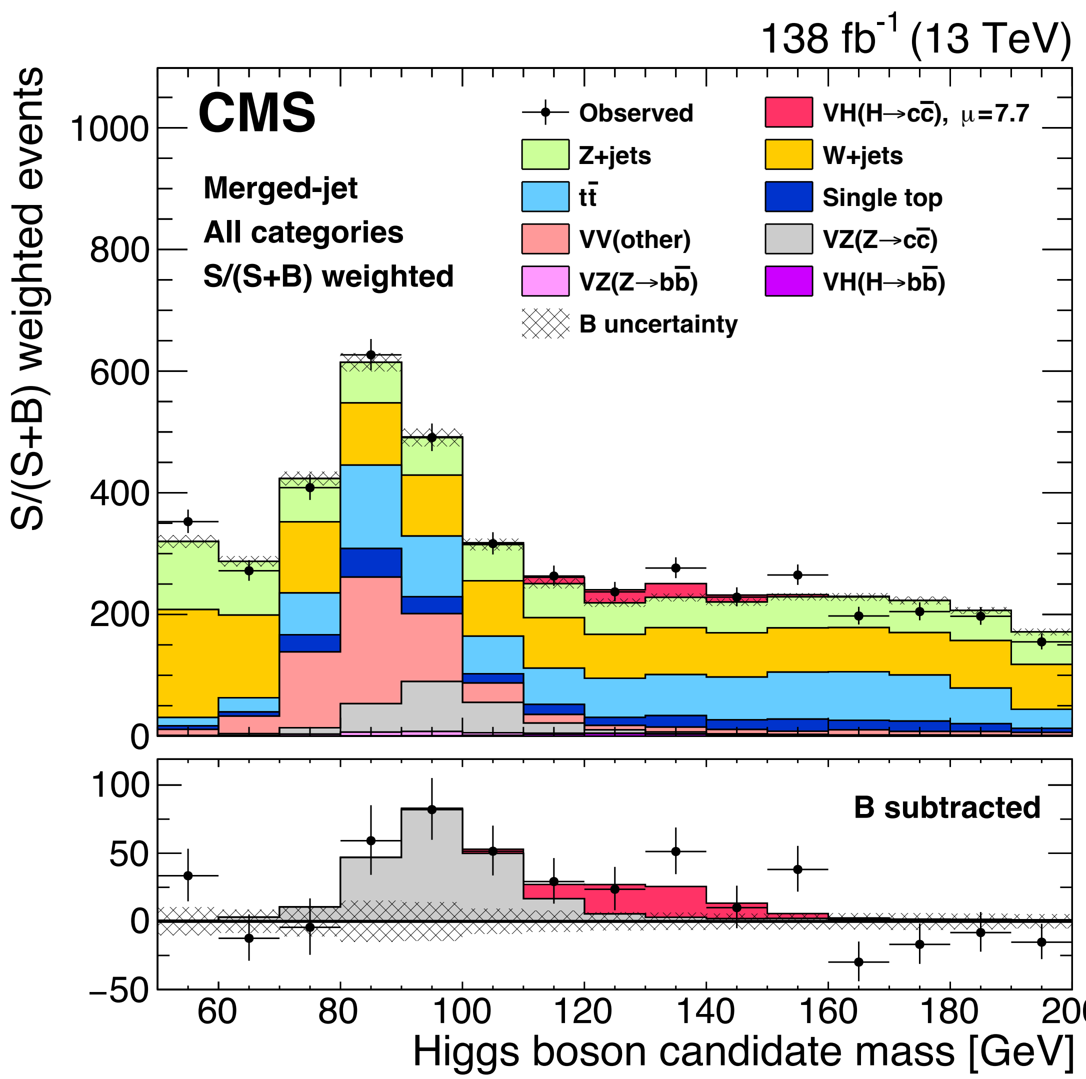
$H(b\bar{b})$ in ggF

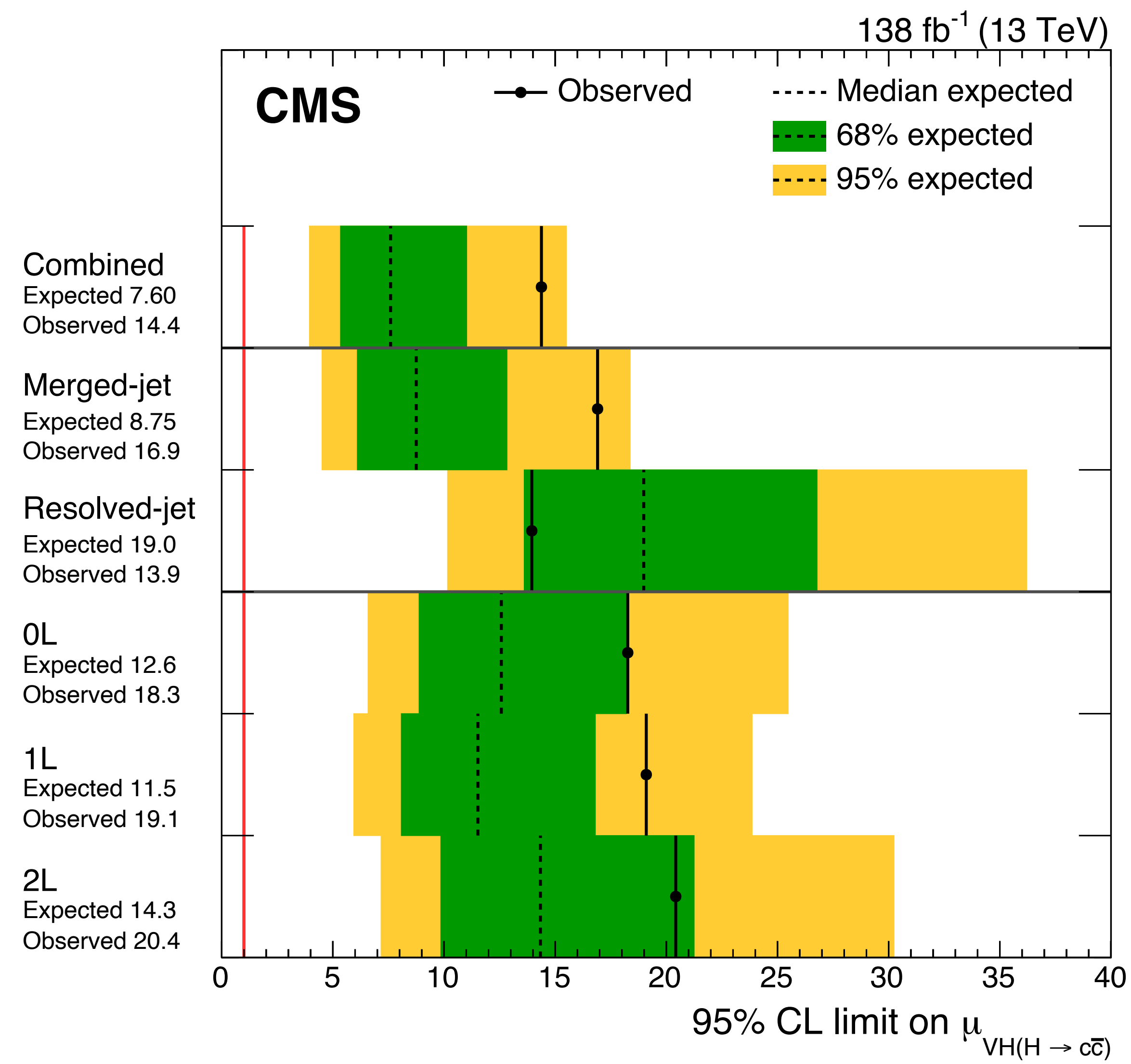


- Multijet background prediction from pass-to-fail ratio
- Observed significance of 2.5σ (0.7 expected)

[CMS, JHEP 12, 085 (2020)]

Doing the impossible: H($c\bar{c}$)





- Made possible by ParticleNet tagger (graph-based NN)
- Observed significance 14 × SM (7.6 × SM expected)

[CMS, PRL 131, 061801 (2023)]

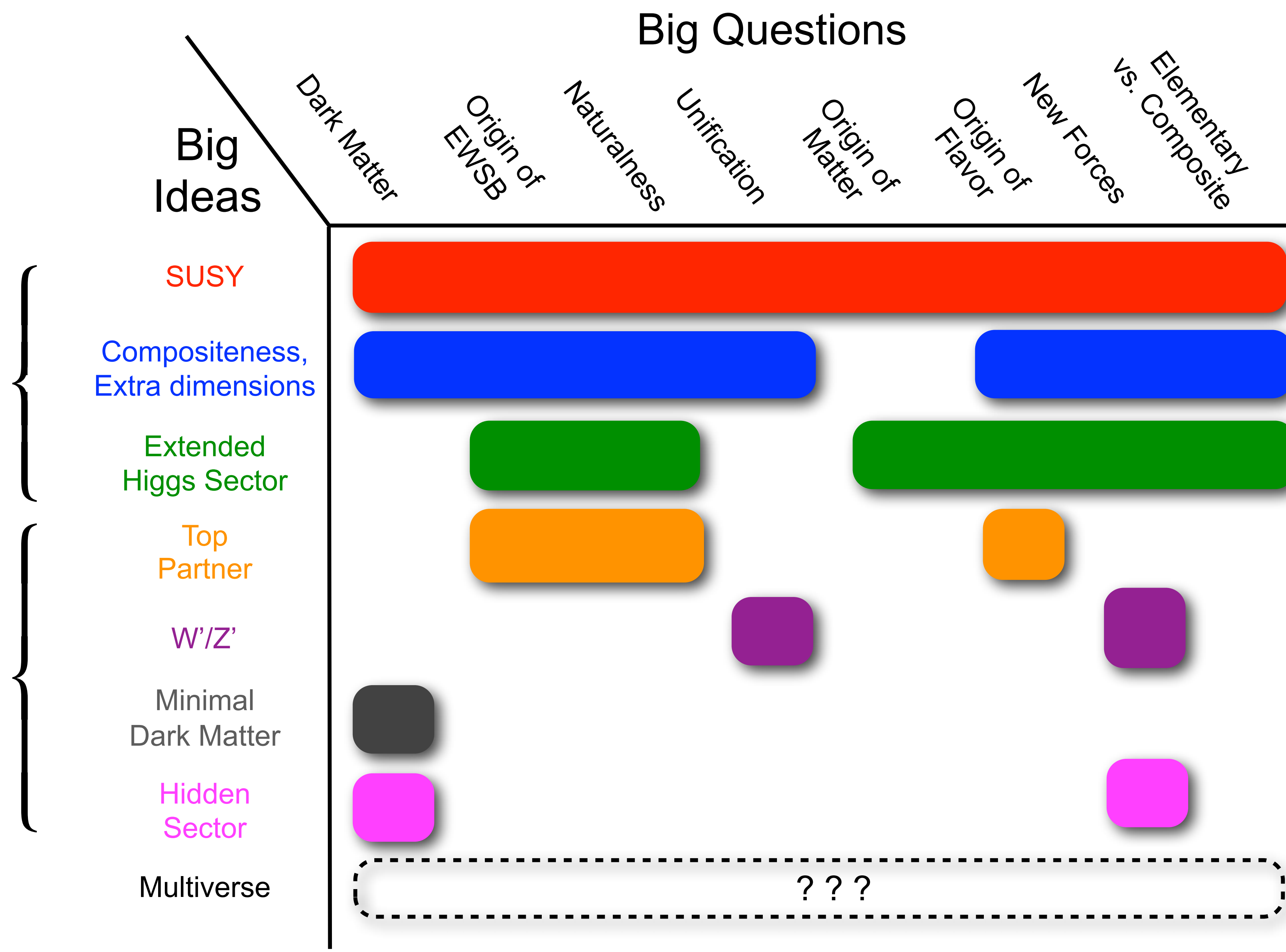
Searches for new phenomena

Beyond the Standard Model

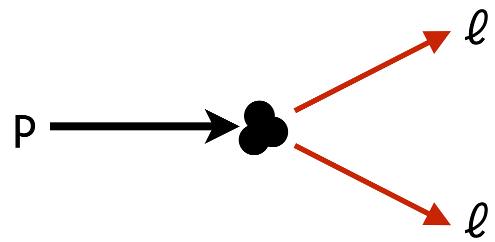
Complete models

New matter fields / interactions

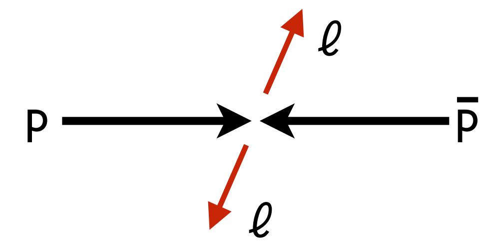
The unknown

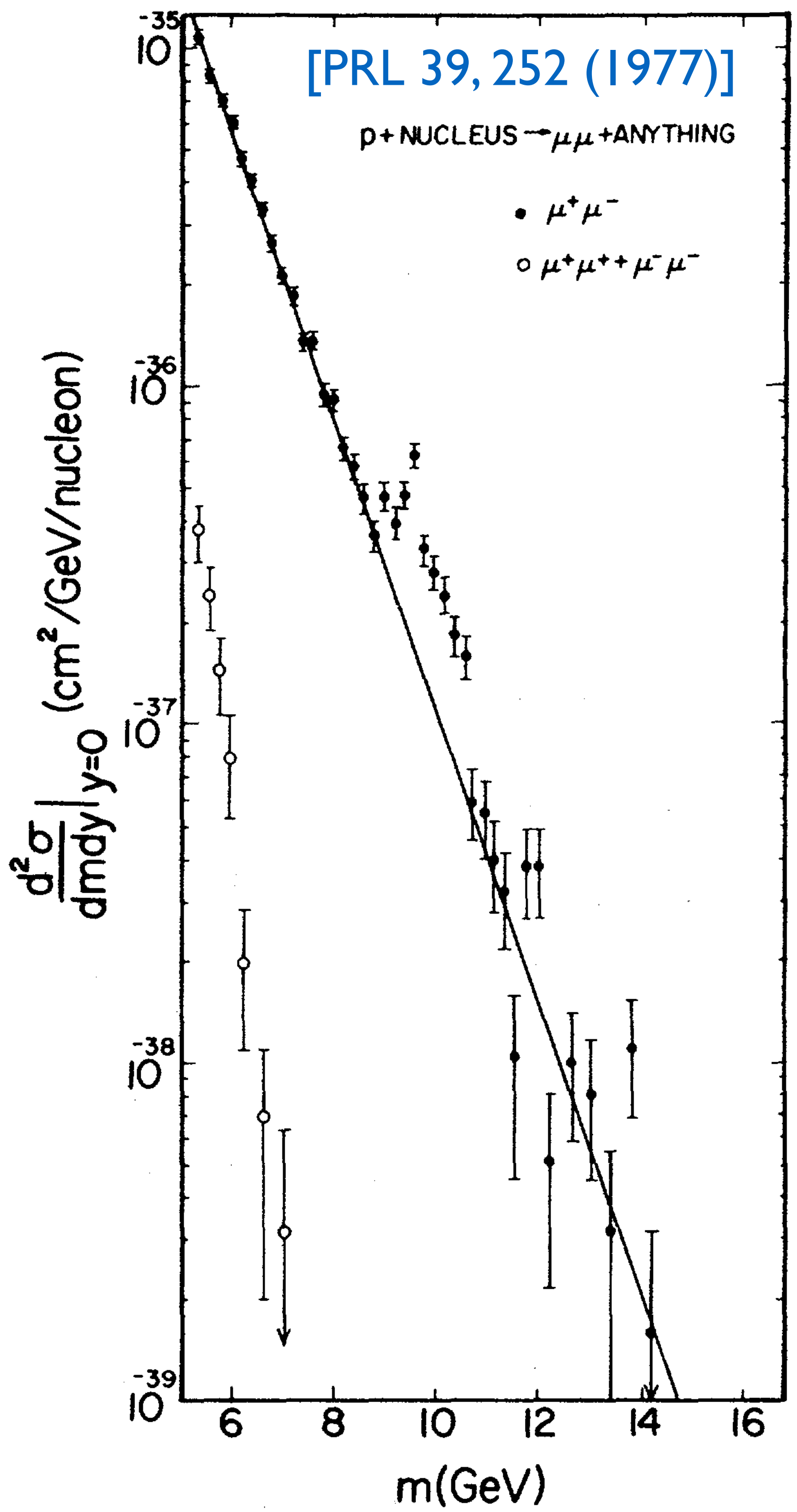


[Snowmass White Paper, arXiv:1311.0299]

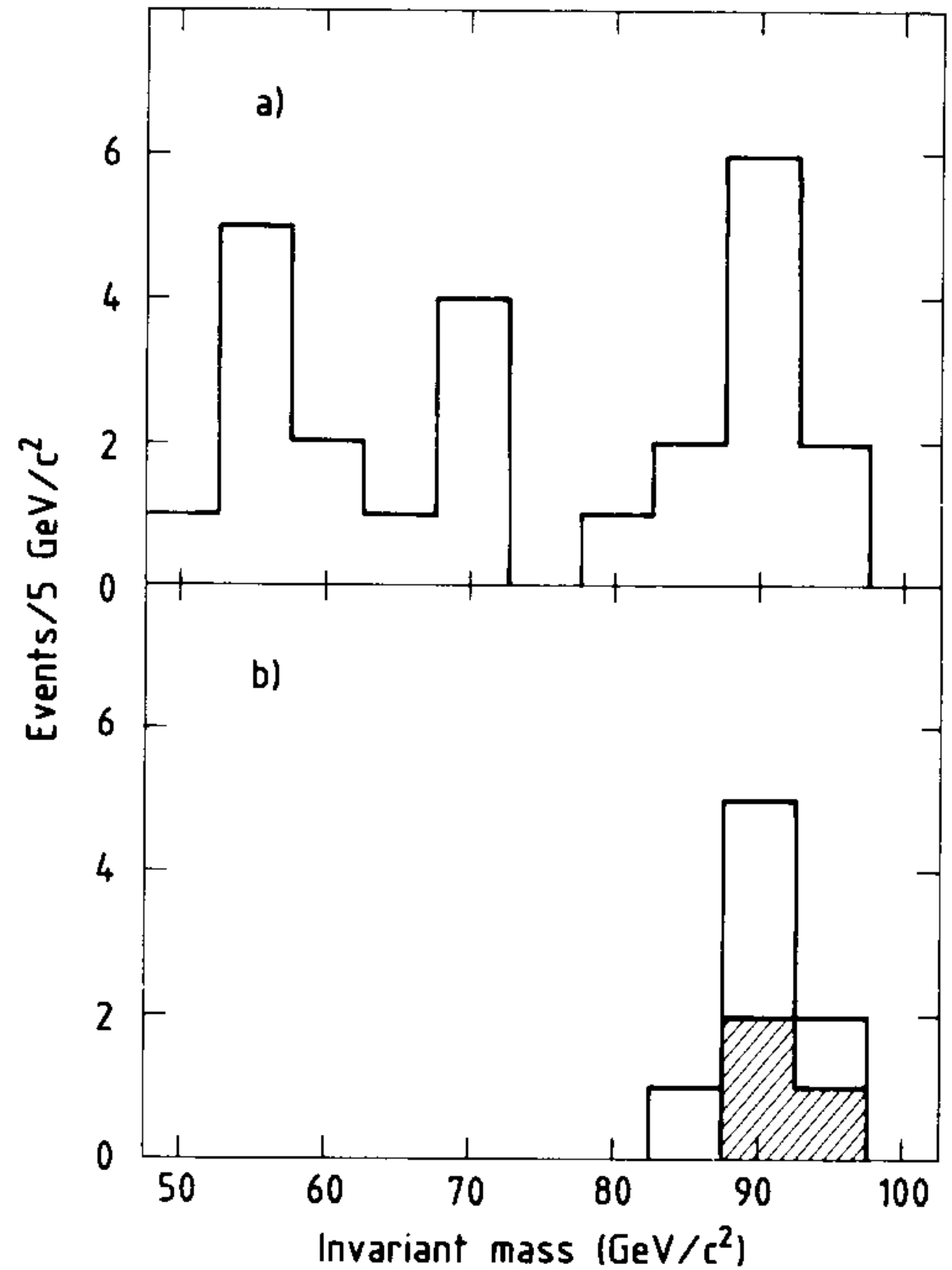


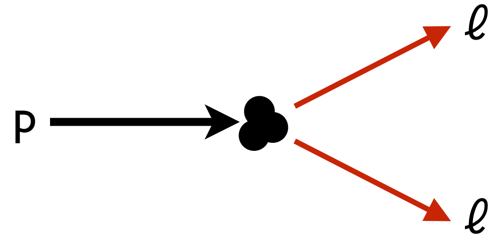
ℓℓ Resonances



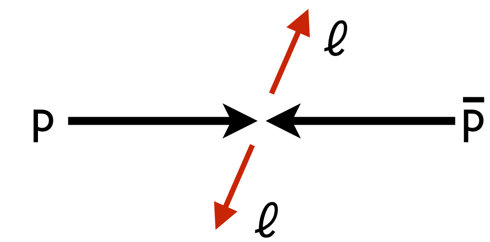


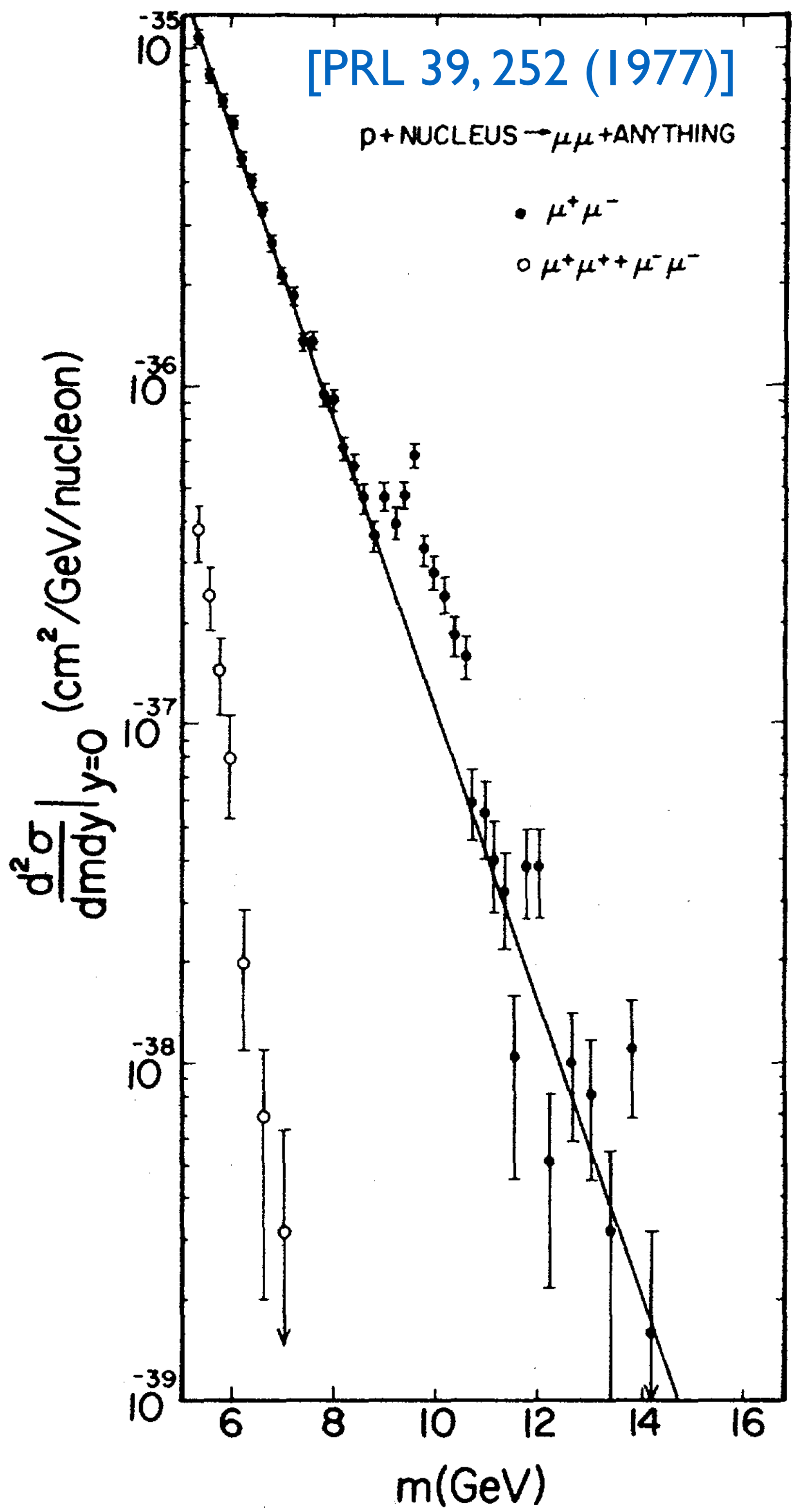
[UA1, PLB126, 398 (1983)
UA2, PLB129, 130 (1983)]



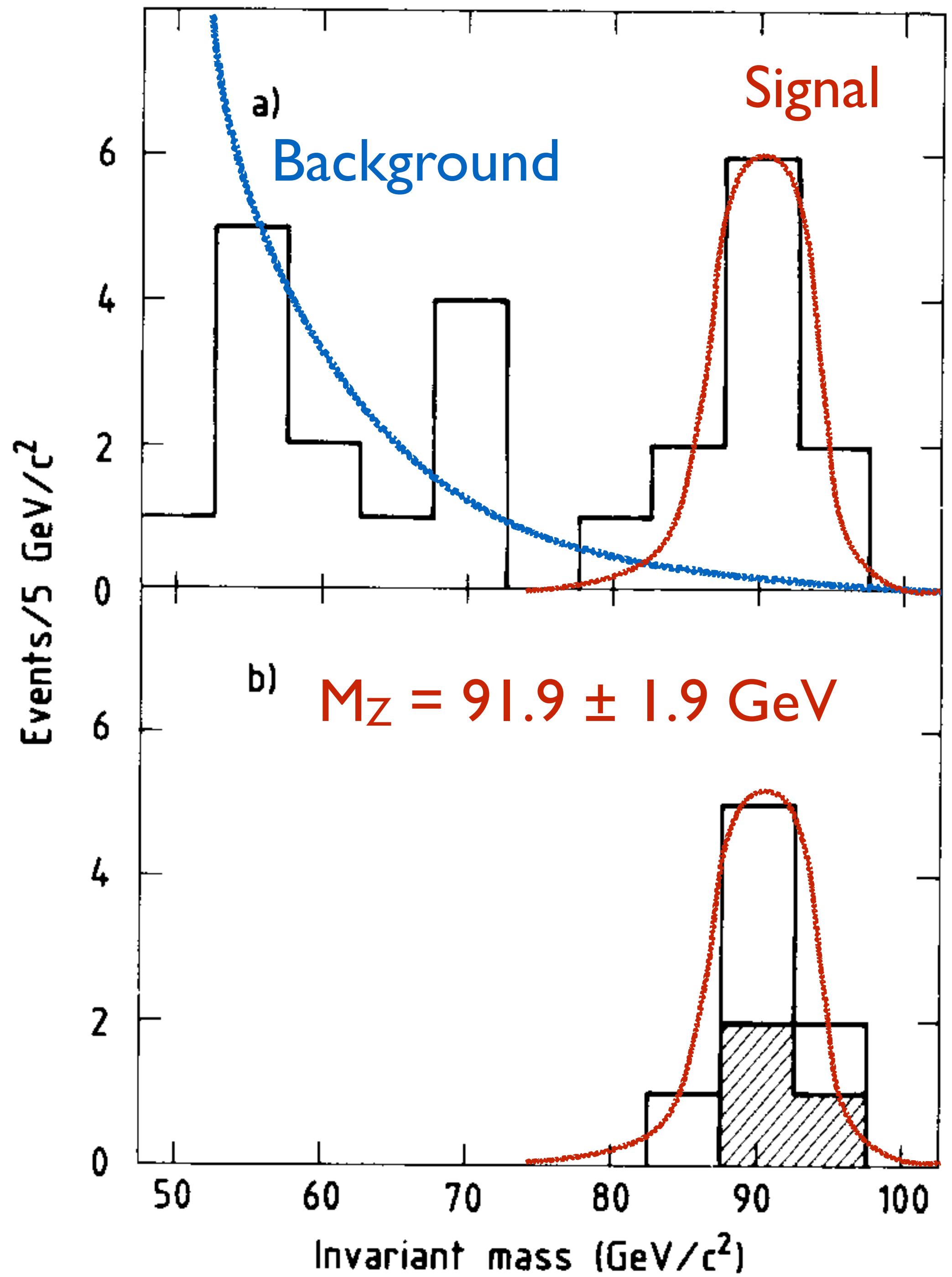


ℓℓ Resonances

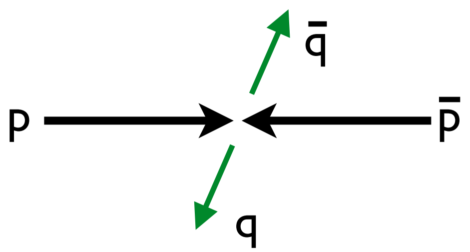




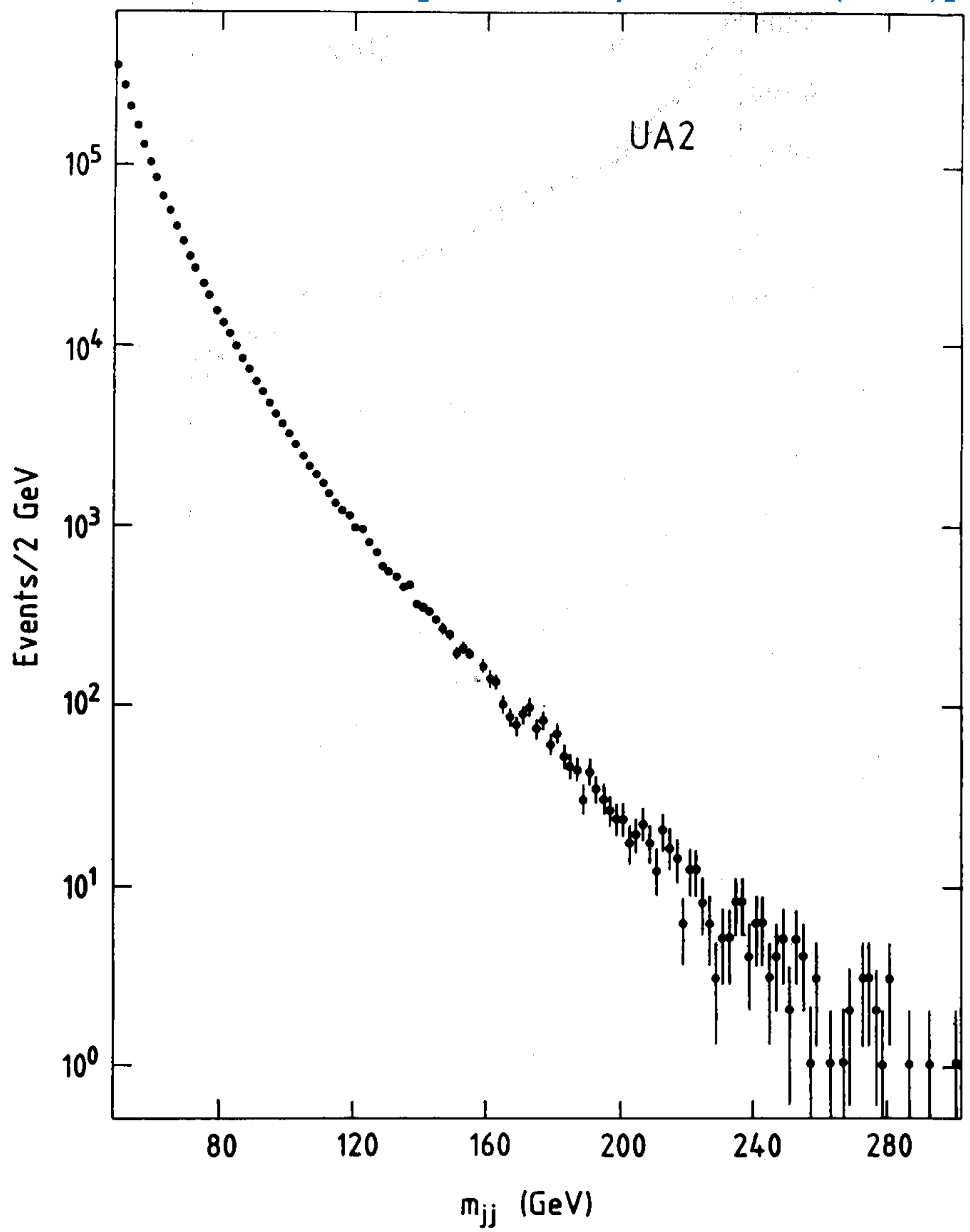
[UA1, PLB126, 398 (1983)
UA2, PLB129, 130 (1983)]



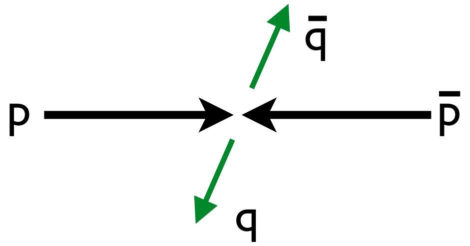
qq/gg Resonances



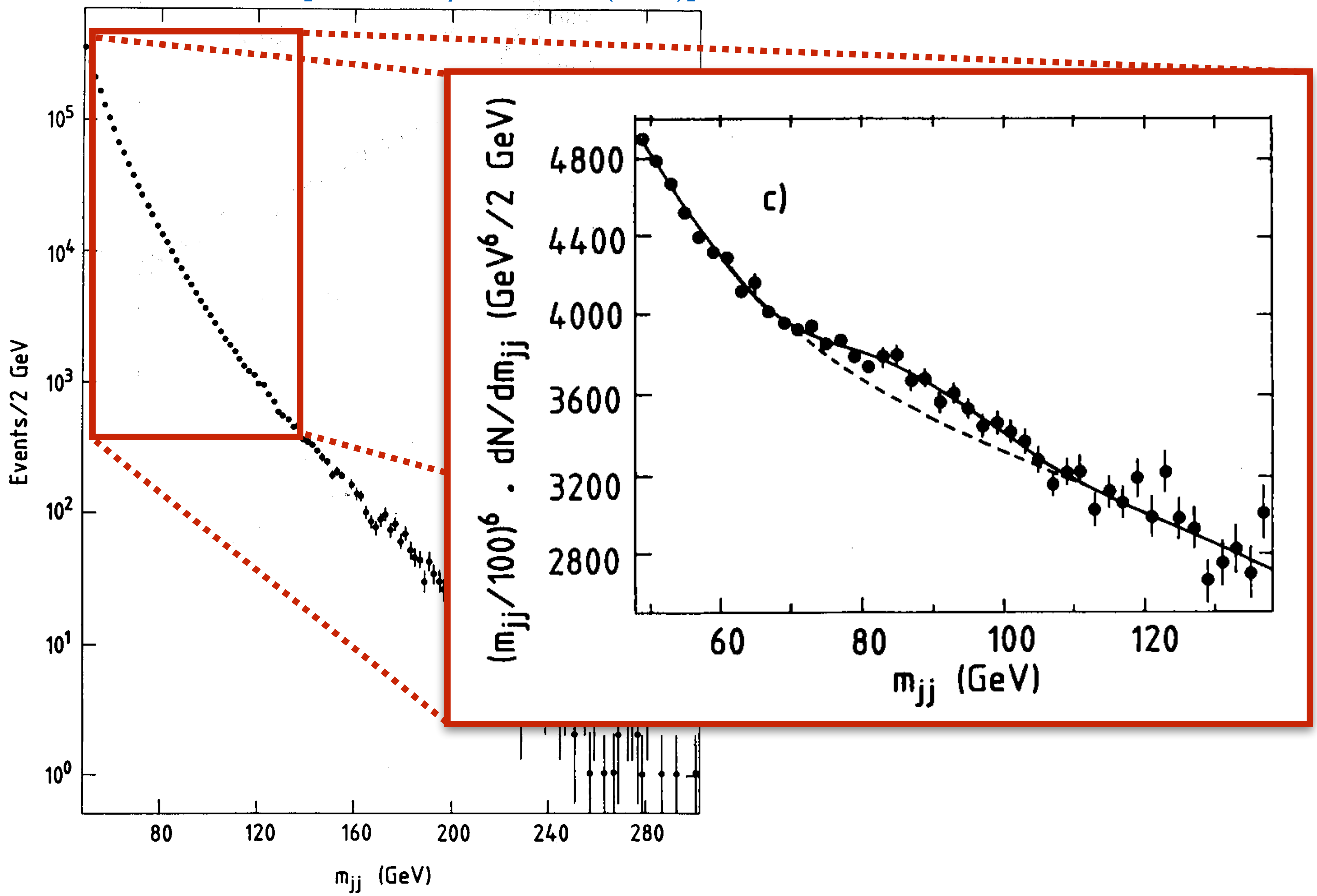
[UA2, Z. Phys. C 49, 17 (1991)]



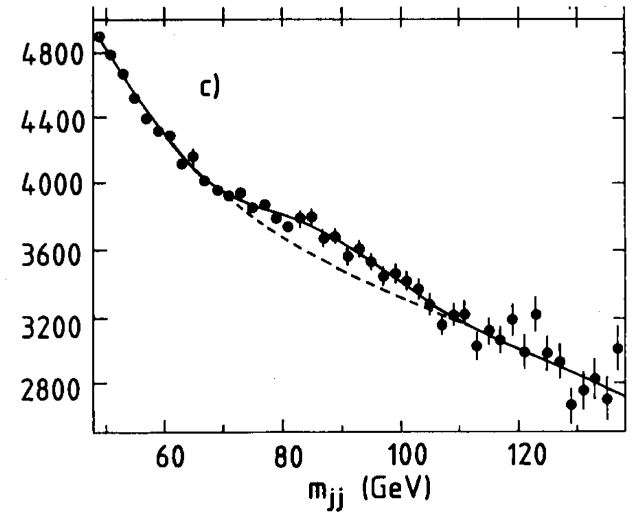
qq/gg Resonances

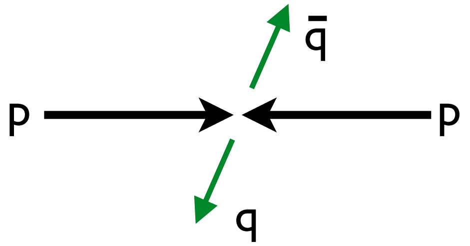


[UA2, Z. Phys. C 49, 17 (1991)]



qq/gg Resonances

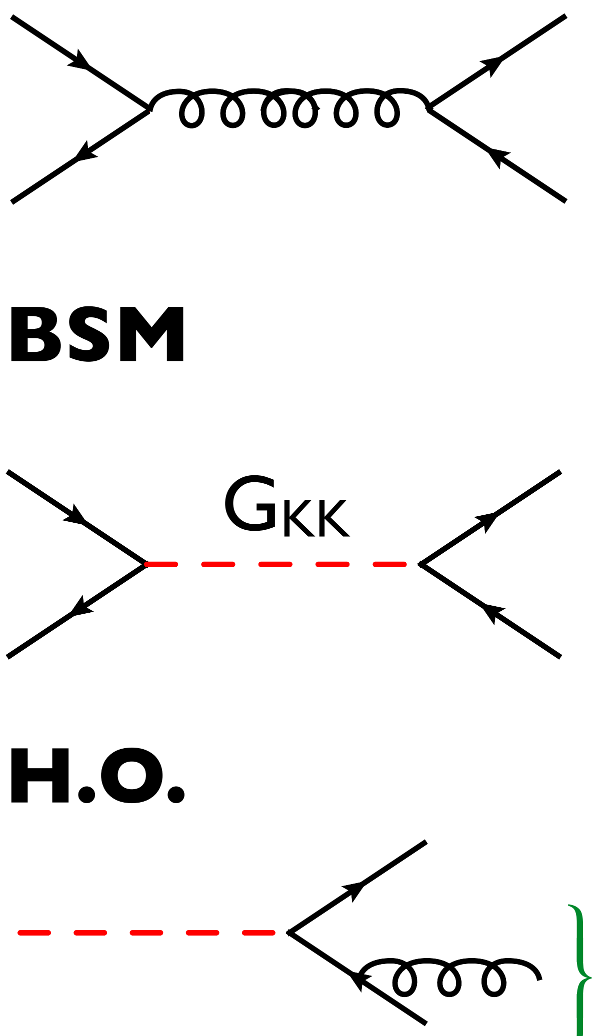




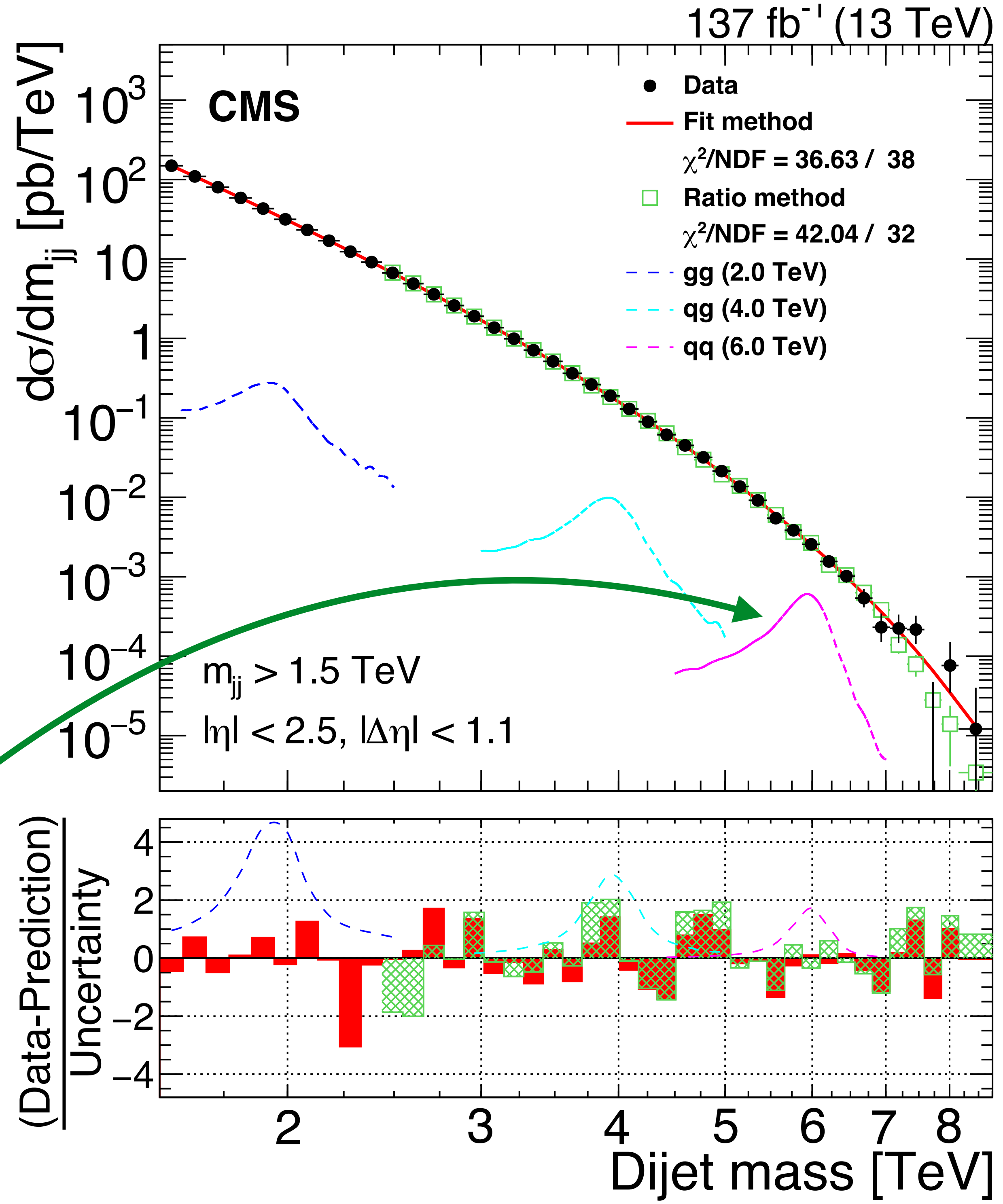
SM

BSM

H.O.



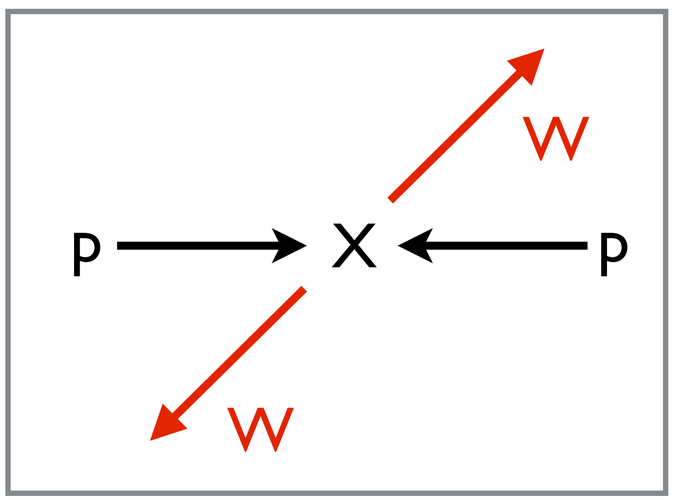
combine



[CMS, JHEP 05, 033 (2020)]

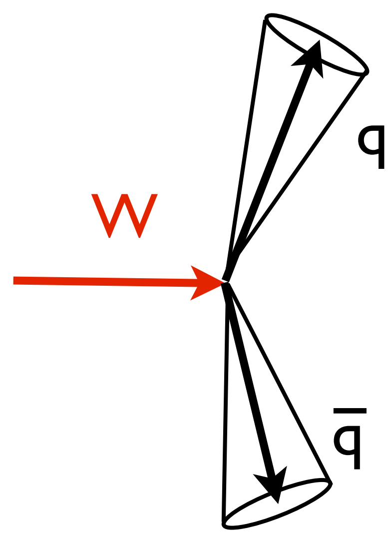
Boost!

pp→X→WW→Jets

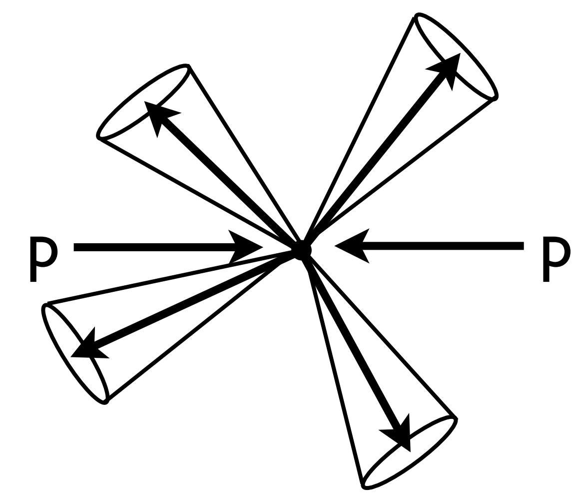


$M_X \sim 2\,M_W$

p_T^W small, $\gamma \lesssim 5$



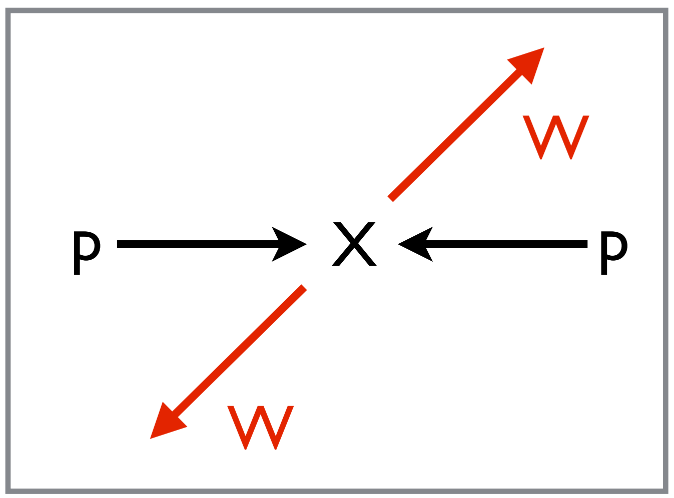
4 jet final state



combinatorics, background!

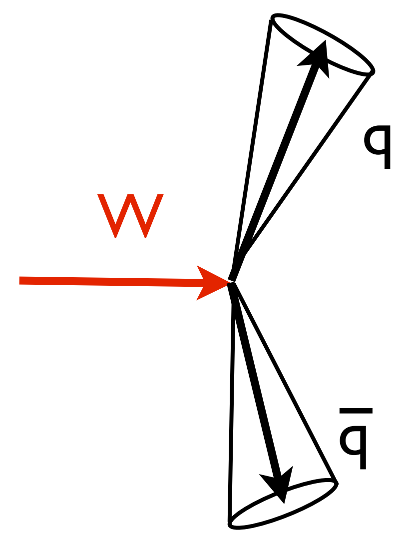
Boost!

pp→X→WW→Jets

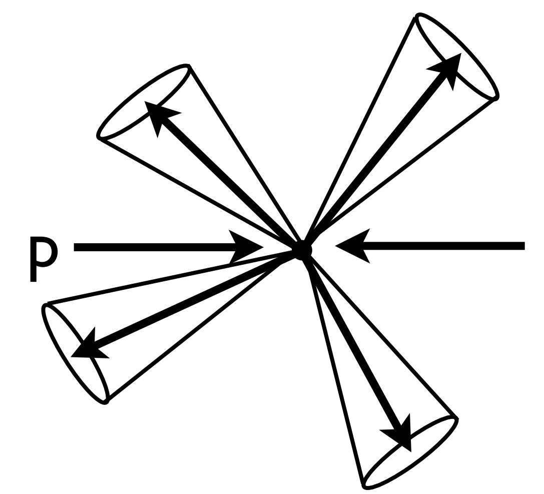


$M_X \sim 2\ M_W$

p_T^W small, $\gamma \lesssim 5$



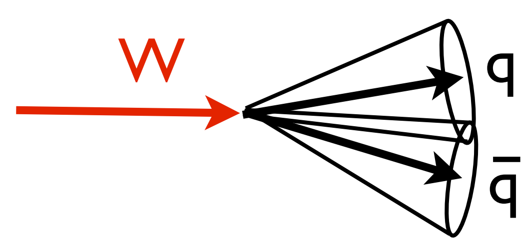
4 jet final state



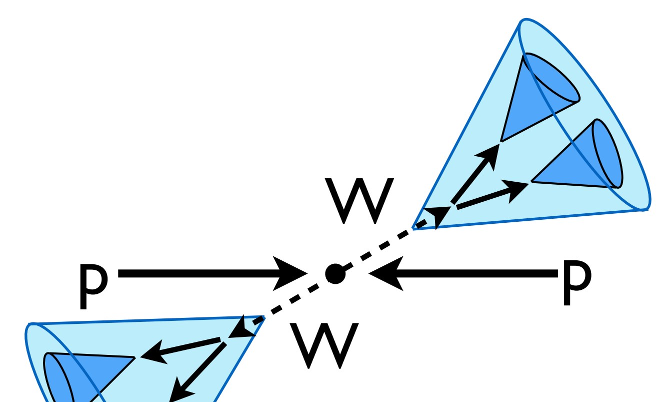
combinatorics, background!

$M_X \gg 2\ M_W$

p_T^W large, $\gamma \gg 5$



Dijet final state



Jet substructure

Diboson-tagged dijet event, M_{JJ} = 5.0 TeV

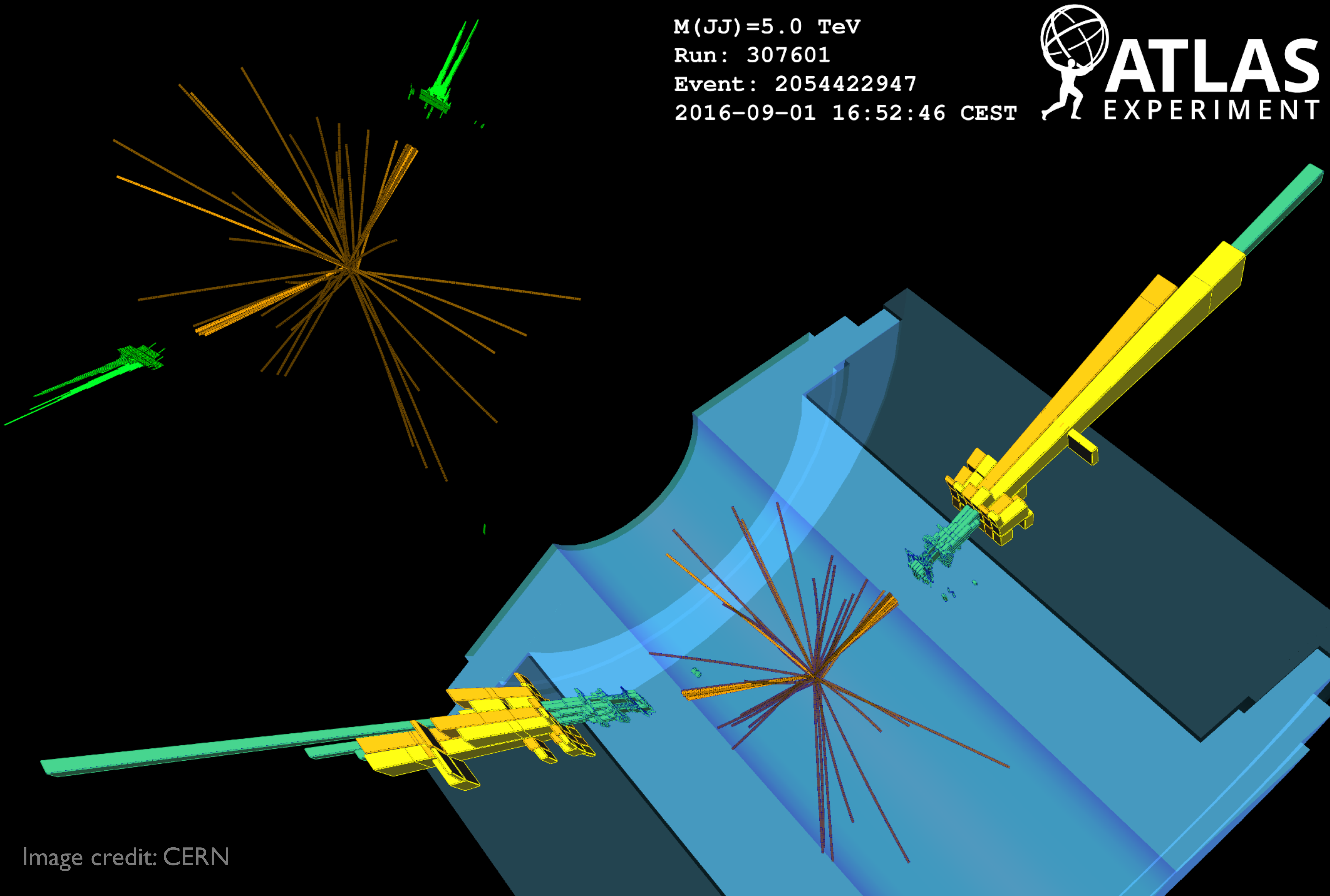
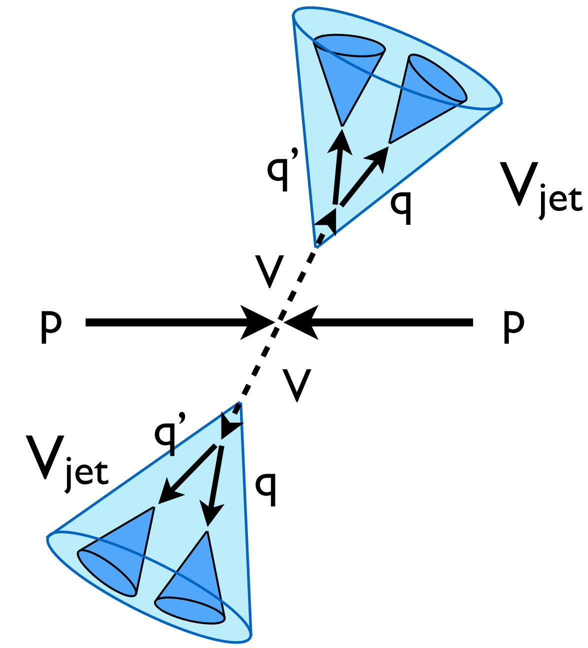


Image credit: CERN

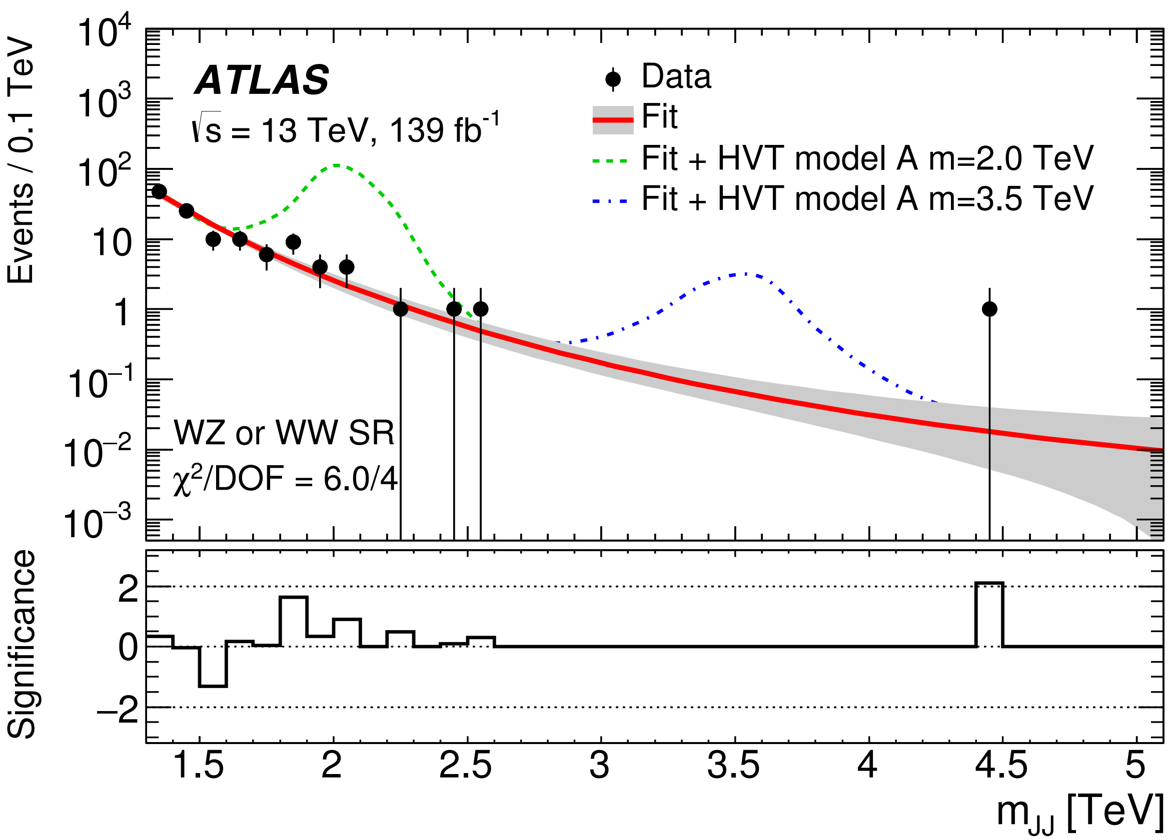
VV Resonances

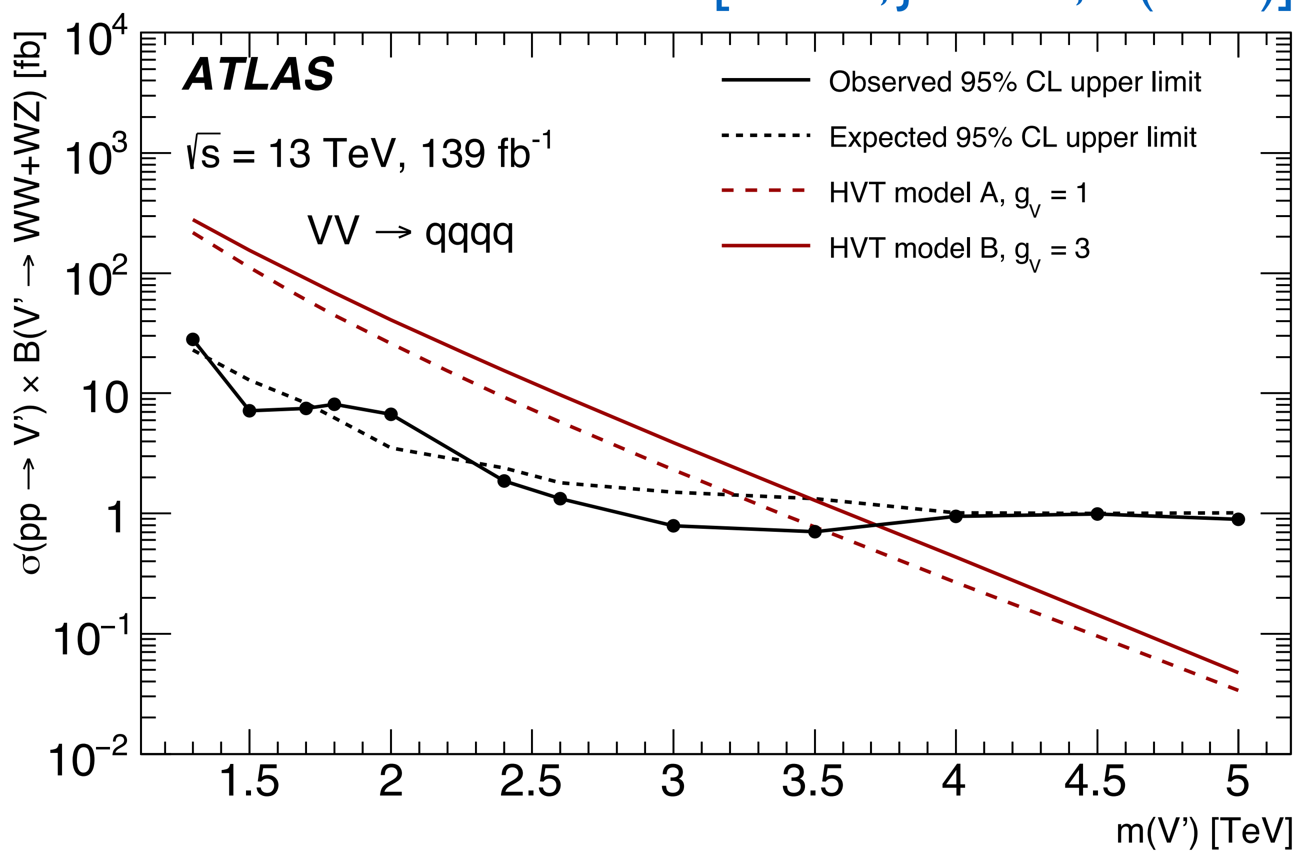


Improved jet substructure resolution with tracking information (TCCs): 50% improvement at high p_T

Optimal S/B with p_T dependent mass and D_2 selections

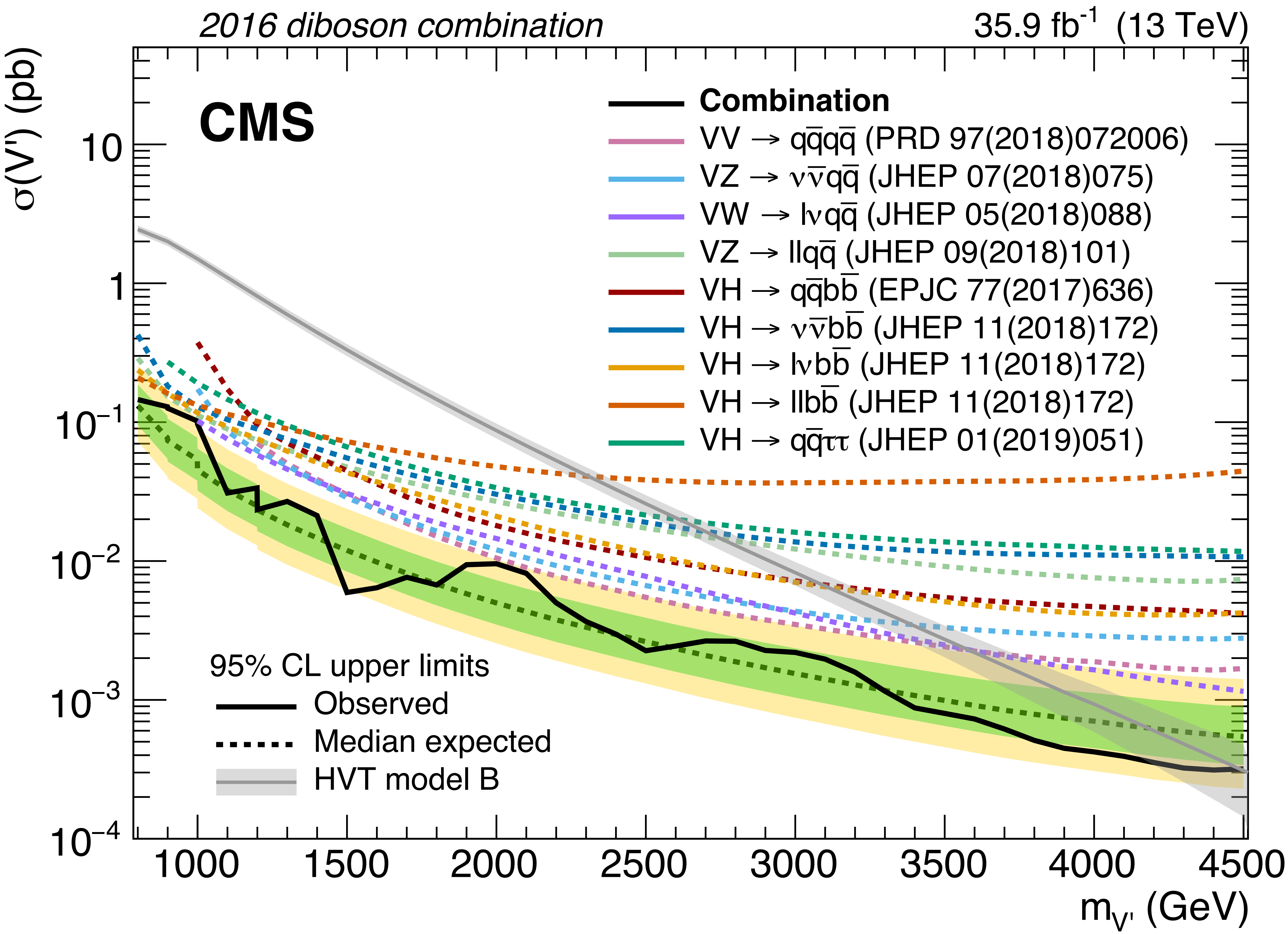
[ATLAS, JHEP 09, 1 (2019)]





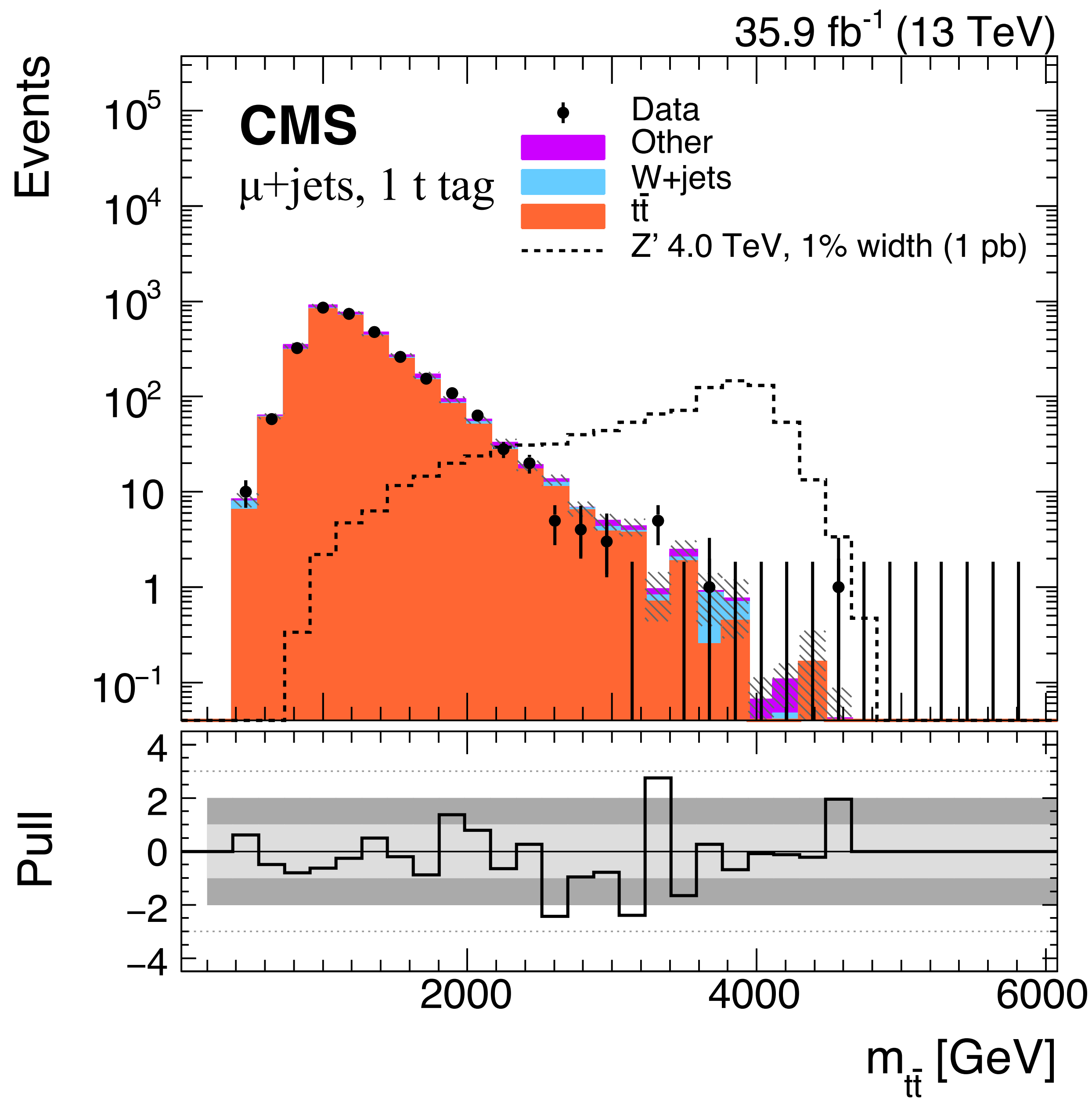
Extension to 4- and 5-prongs: [CMS, PRL 121, 141802 (2018)]

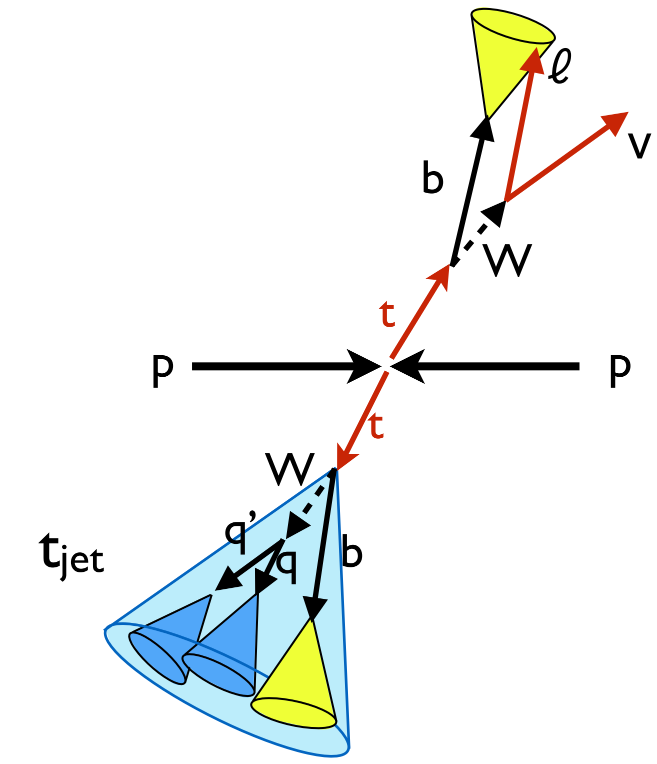
Combination



[CMS, PLB 798, 134952 (2019)]

$t\bar{t}$ Resonances





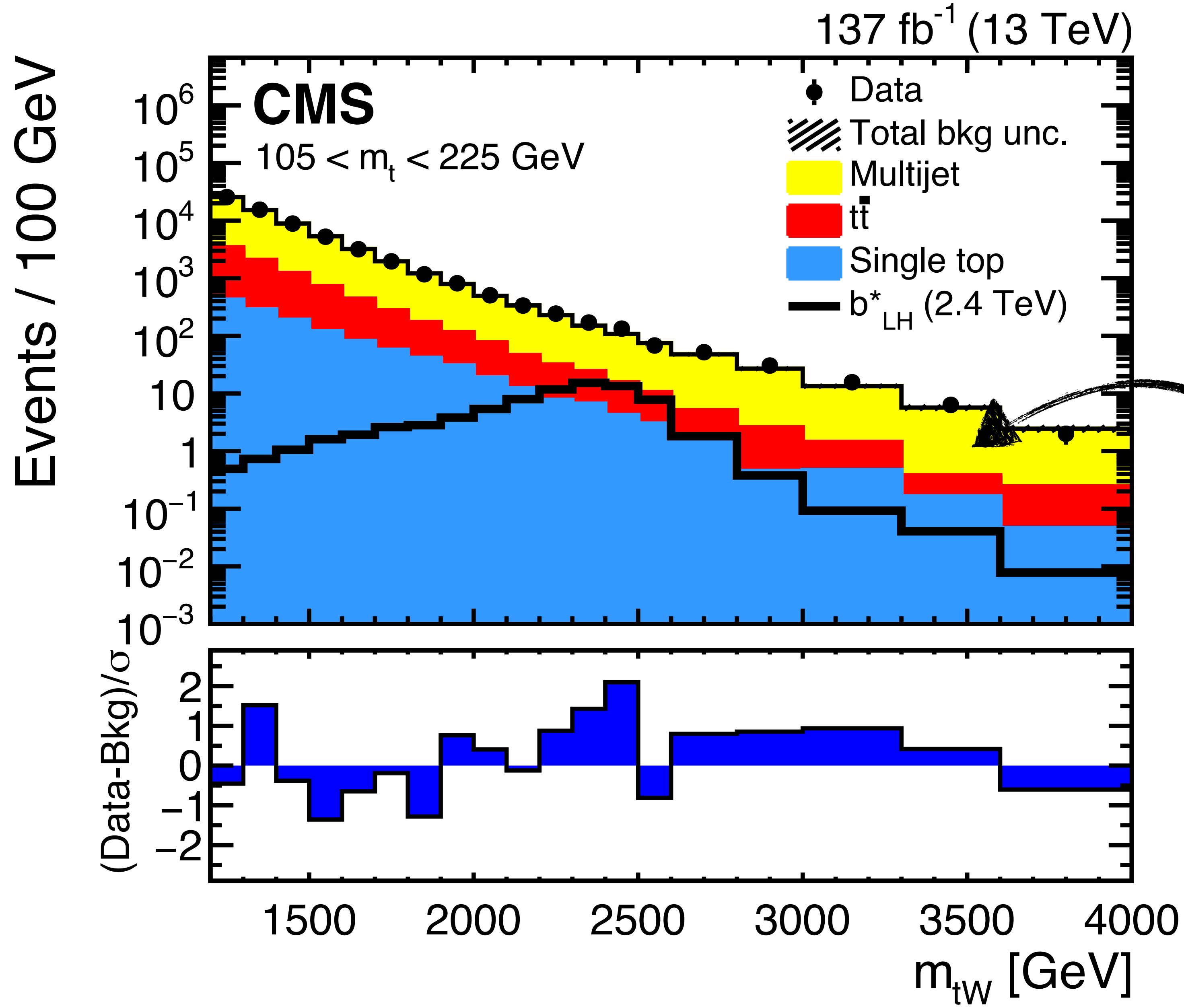
Combination of ℓℓ, ℓ+jets and all-hadronic channels:

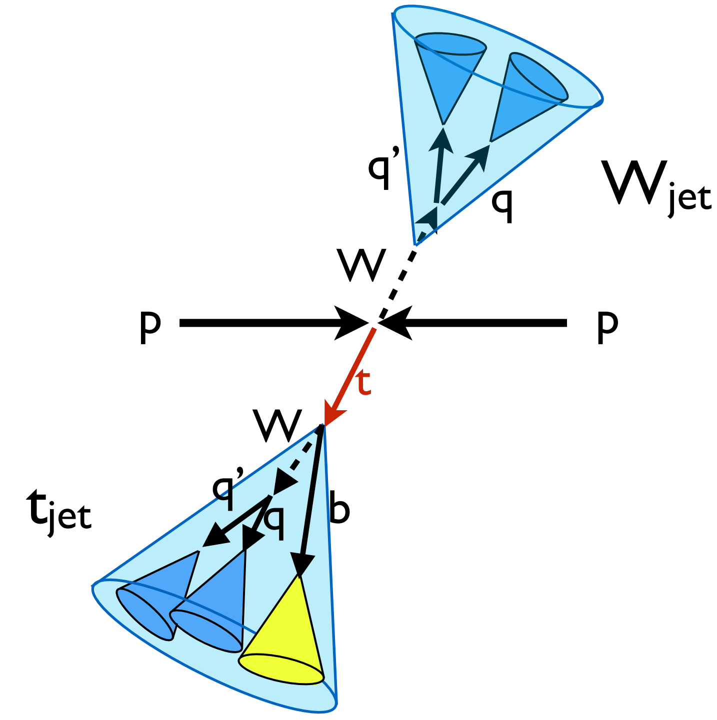
Kaluza-Klein gluons excluded below **4.6 TeV**

- BDT for W+jet suppression
- Sidebands to constrain backgrounds

[CMS, JHEP 04, 031 (2019)]

tW resonances



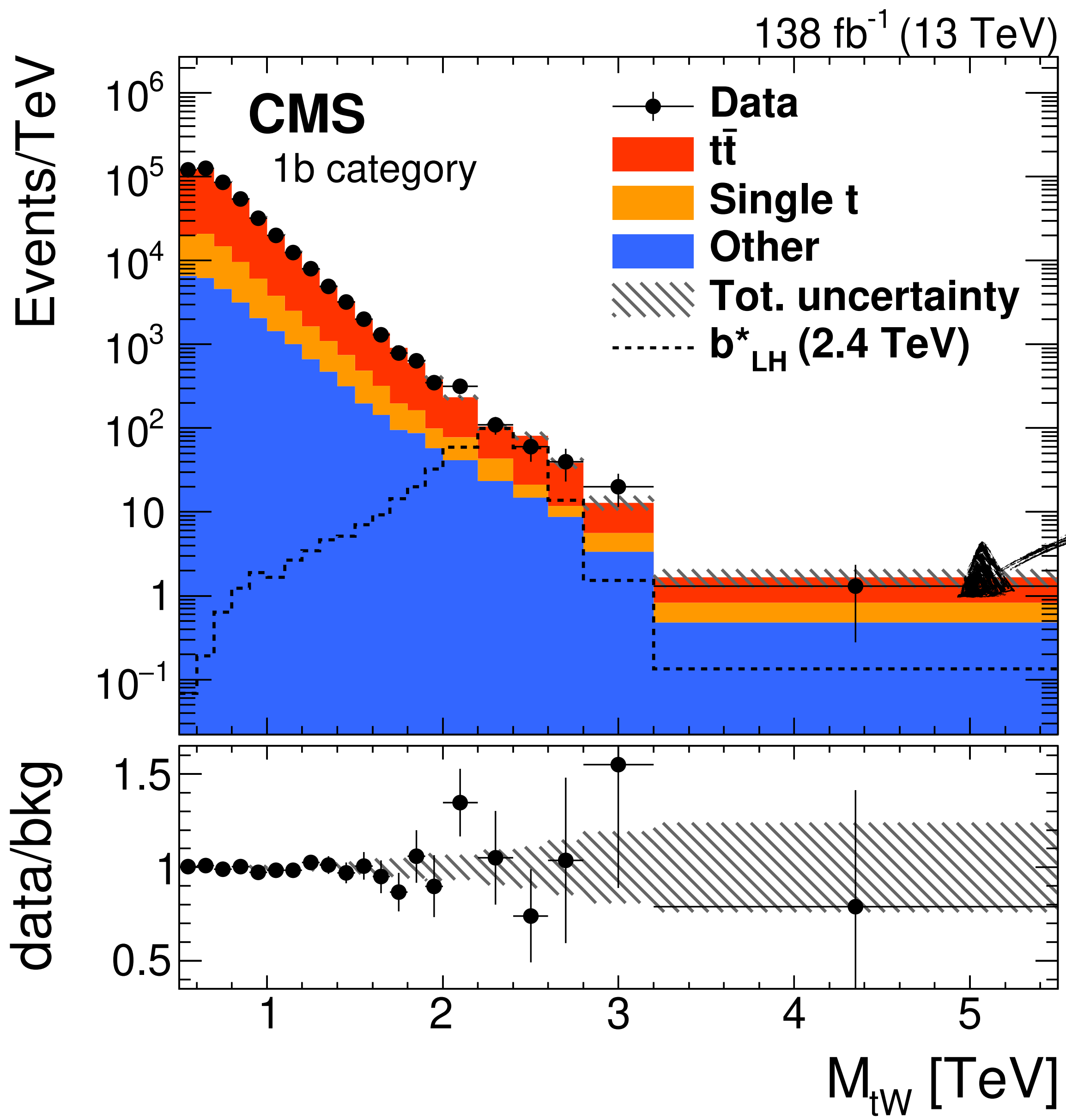


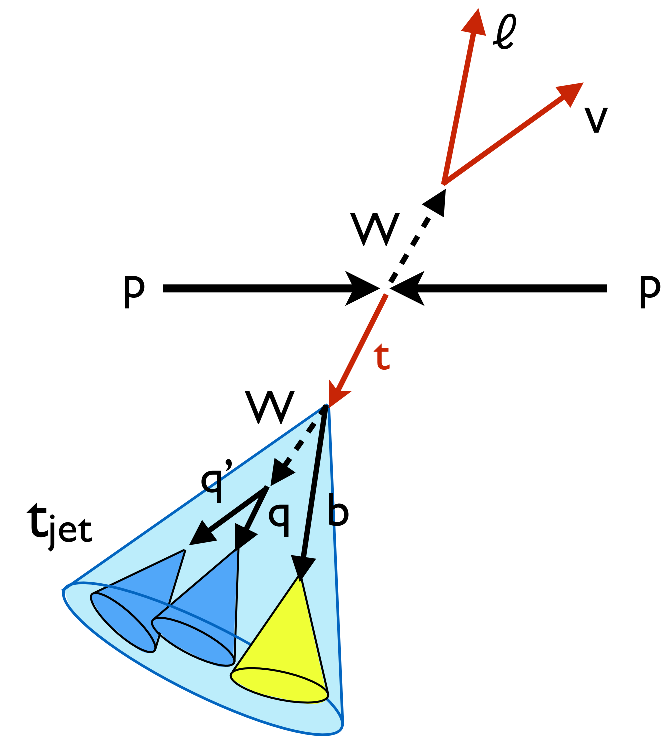
Multijet background estimated from a 2D pass-fail ration in (m_t, m_{tW}) plane

- Sensitivity to excited b* quarks and vector-like quarks

[CMS, JHEP 12, 106 (2021)]

tW resonances



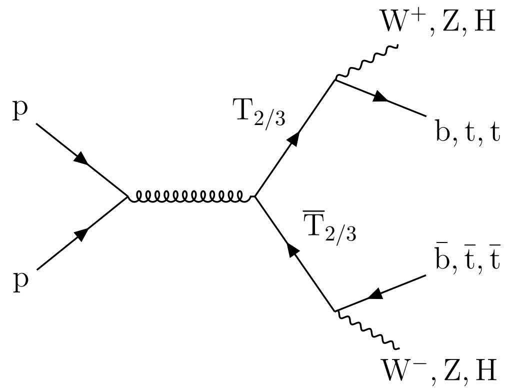


Genuine top quark background dominates (tt, tW)

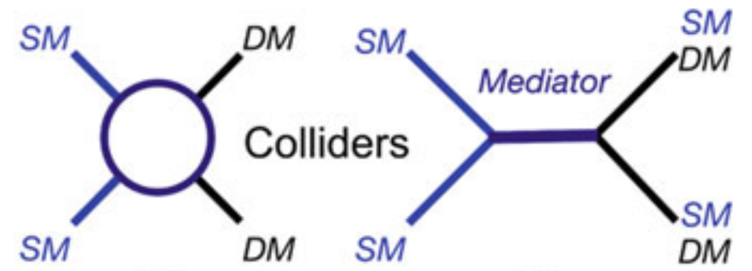
- Extend sensitivity down to 700 GeV using lepton triggers and HOTVR
- Background from misidentified t jets extrapolated from sideband

[CMS, JHEP 04, 048 (2022)]

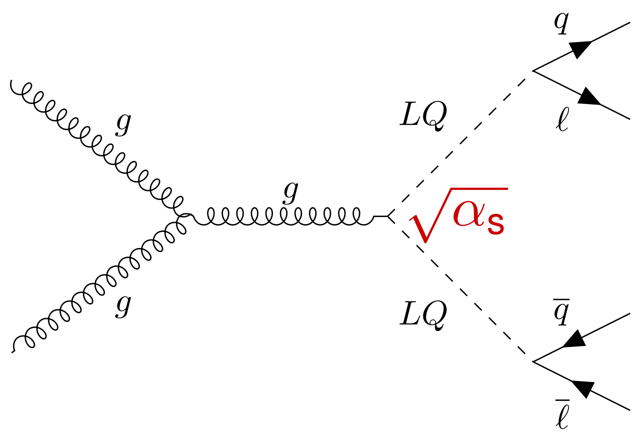
Searches, Searches, Searches



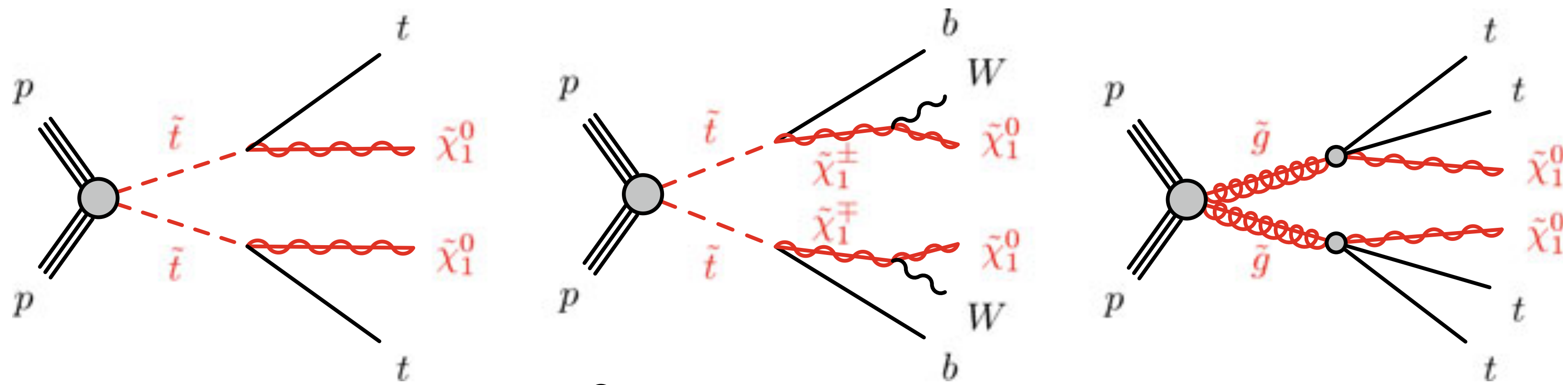
Vector-like quarks



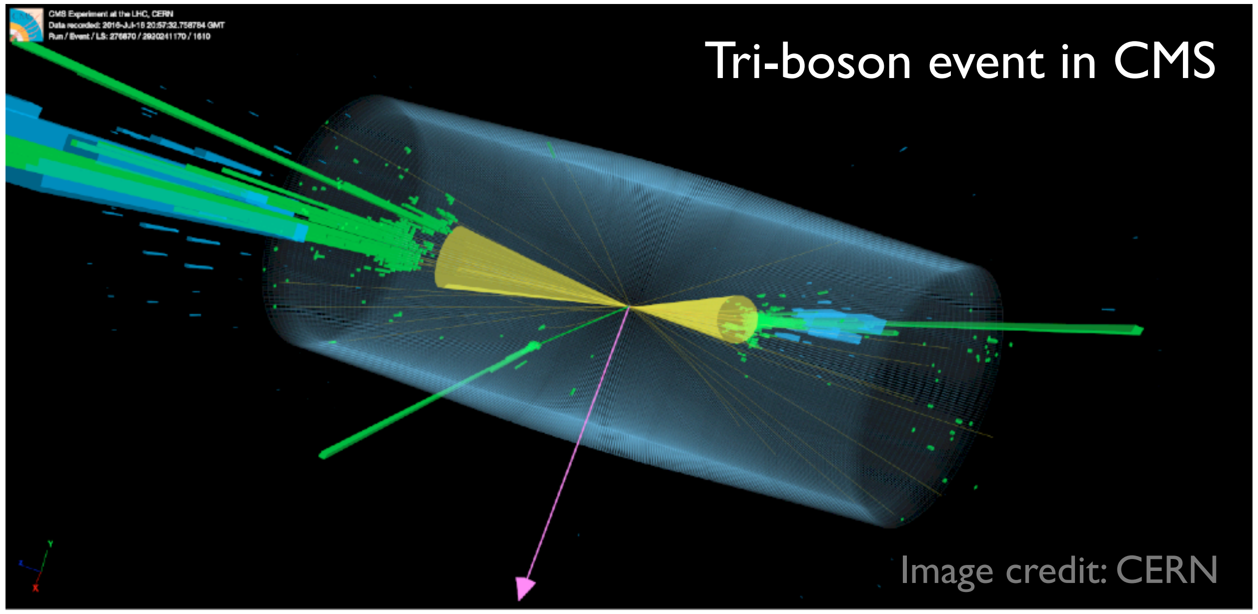
Dark matter



Leptoquarks



Supersymmetry



[CMS, JHEP 08, 177 (2018)]

[https://atlas.cern/updates/atlas-feature/dark-matter]

[ATLAS, JHEP 06, 108 (2018)]

[ATLAS, JHEP 10, 062 (2020)]

Summary

Jet substructure ubiquitous at the LHC

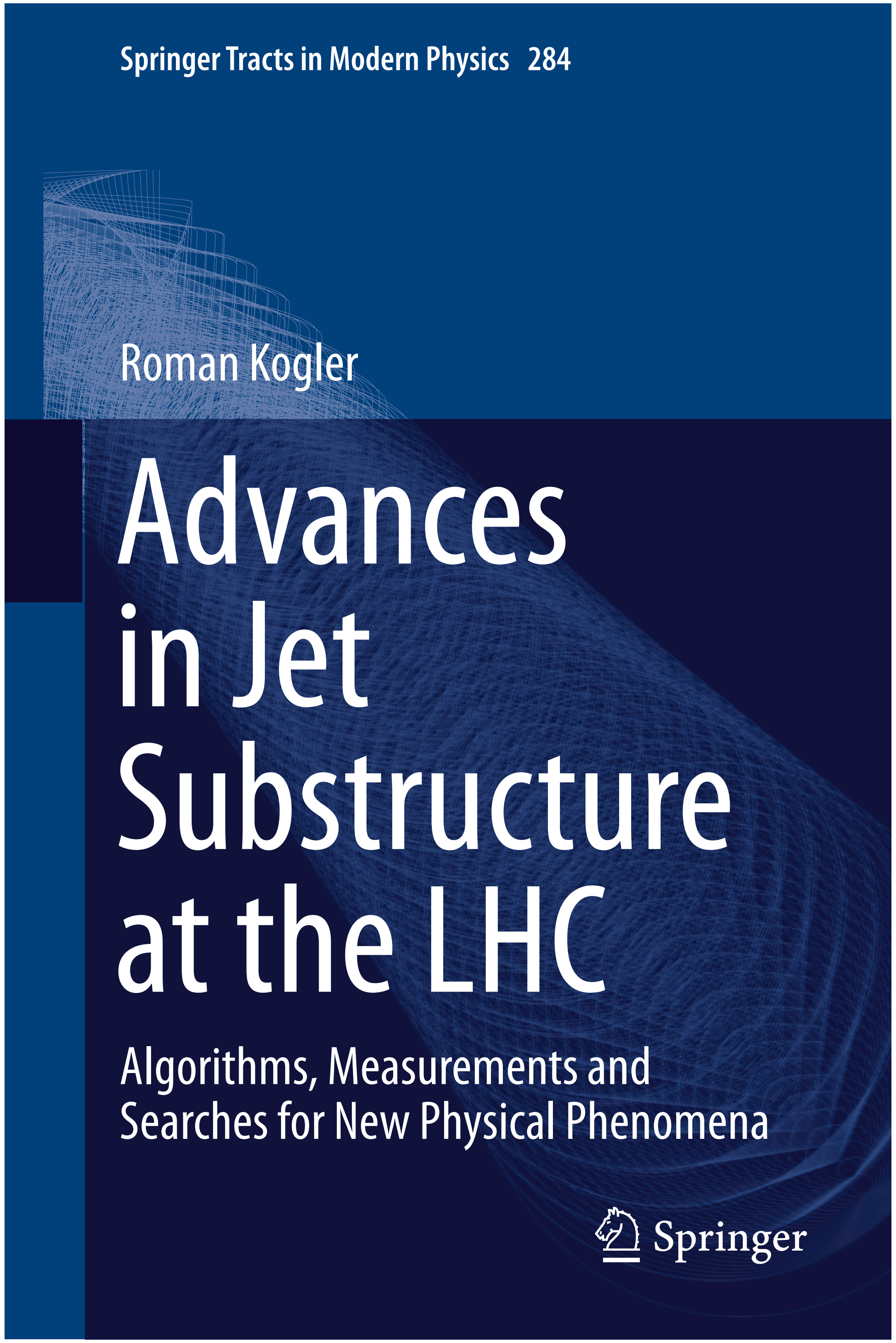
- Remarkable progress in the last 15 years
- Exciting interplay between:
  - theory
  - model building
  - tools development
  - commissioning
  - application

[M. Schwartz at BOOST 2012]

- Jet substructure is past its adolescence: Precision QCD!
- Coming years will bring many novel measurements, searches and algorithms with jet substructure


Image credit: Acton Memorial Library

[RK, STMP 284, 2021]



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