Jet Algorithms

Beatrice Liang-Gilman, 290E Jets Seminar, October 4 2023

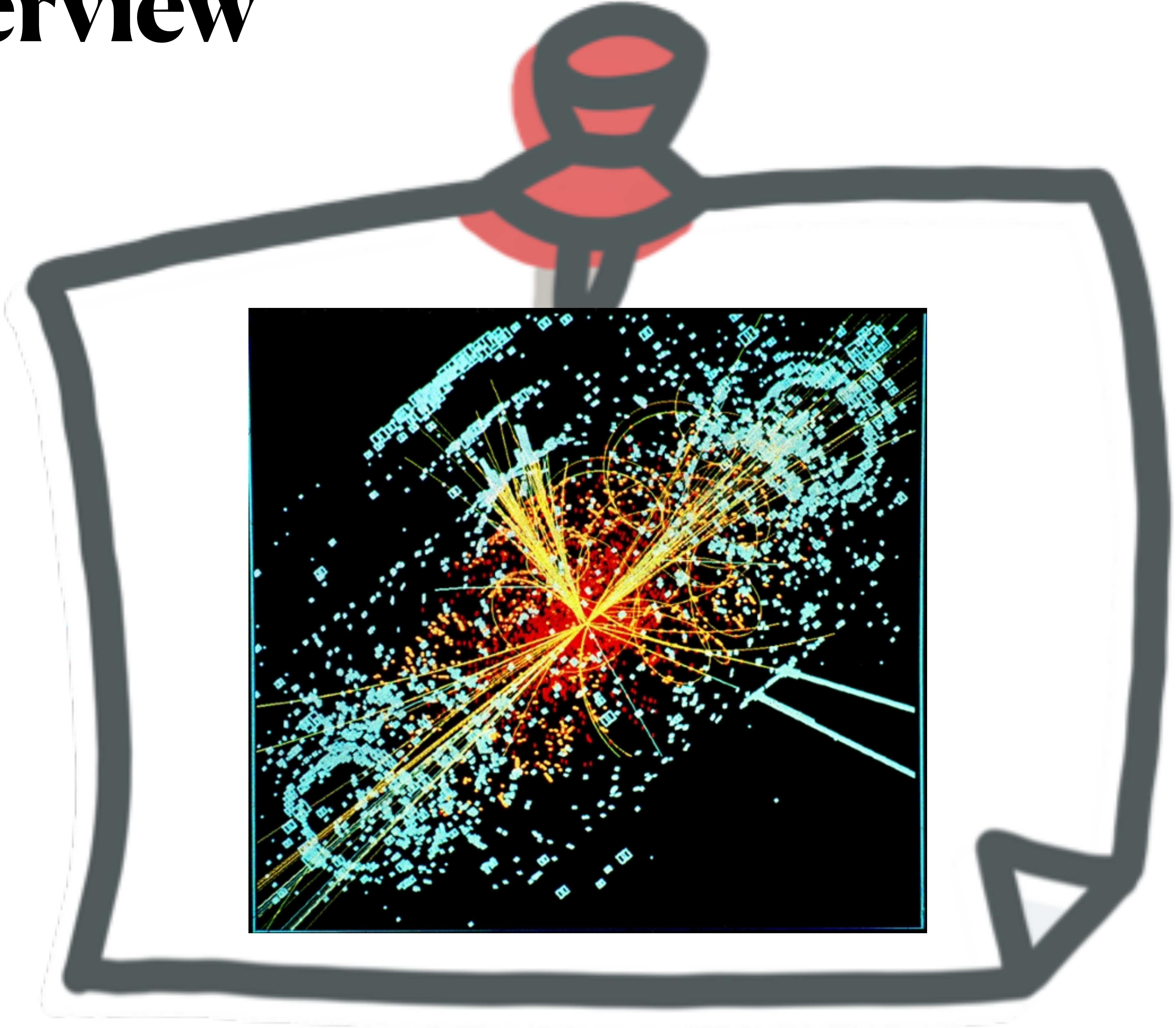
Overview

Jet Requirements

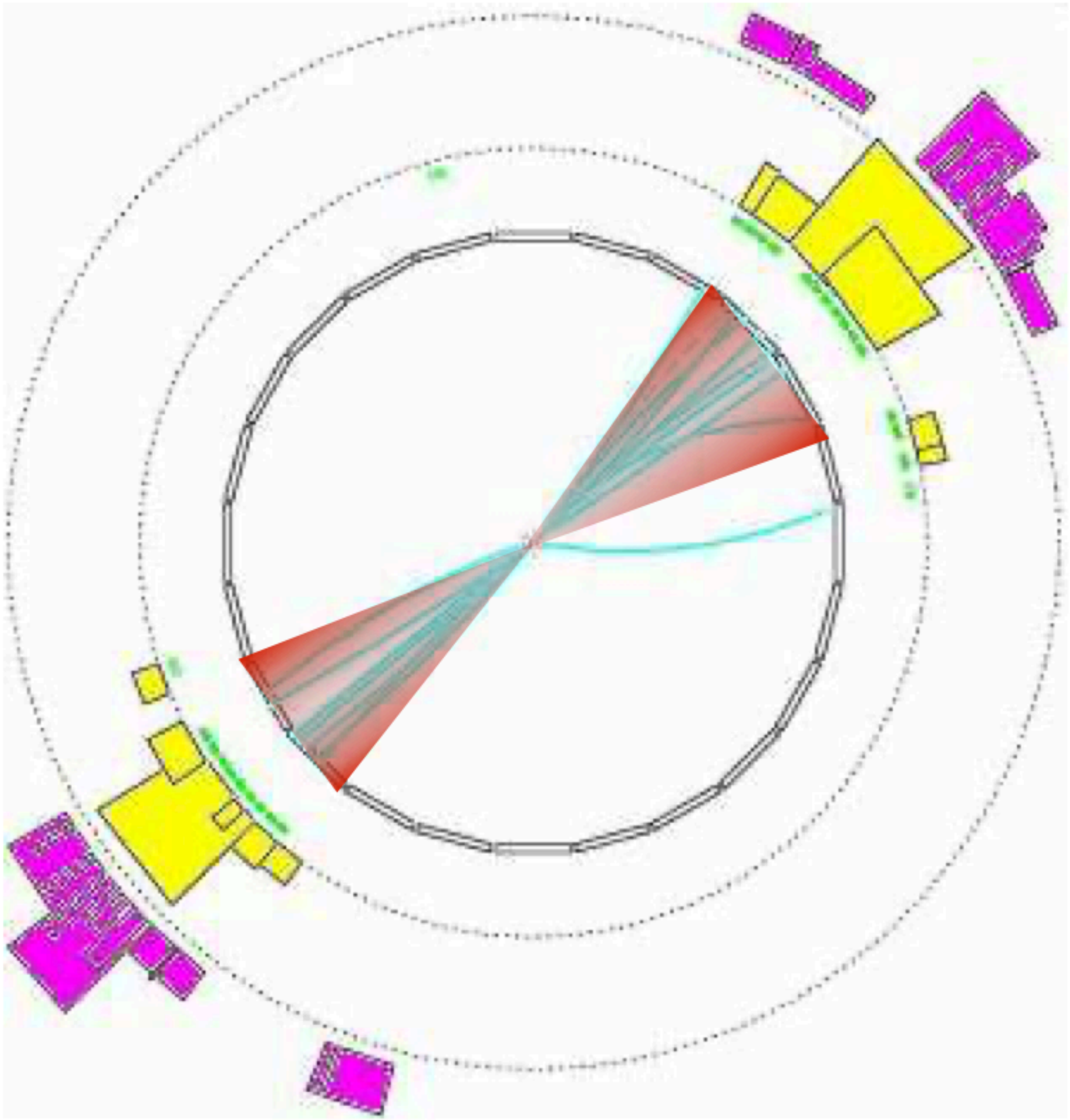
Jet Algorithms

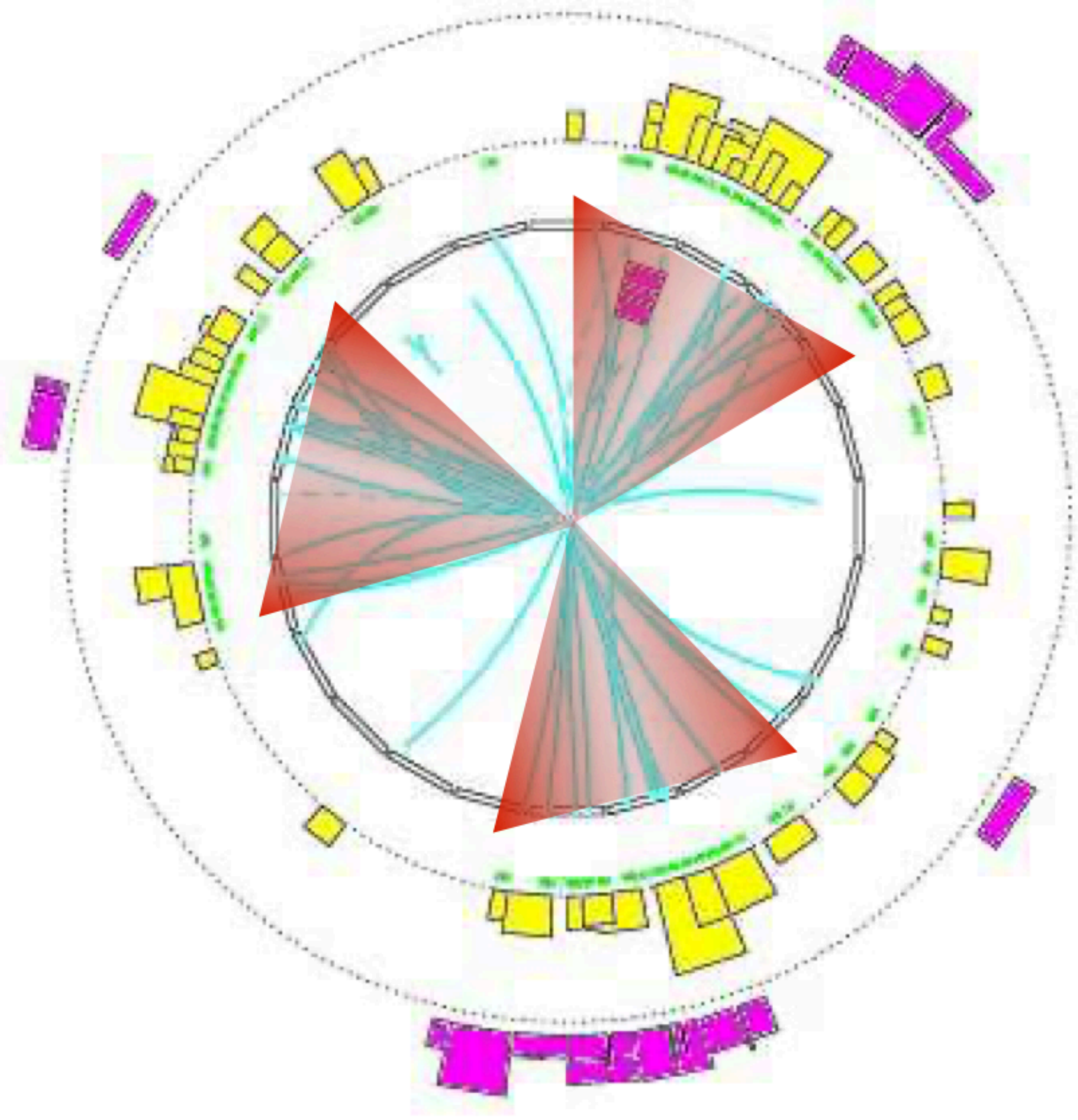
- Iterative Cone
- Sequential Recombination

How Jet Algorithms are Used

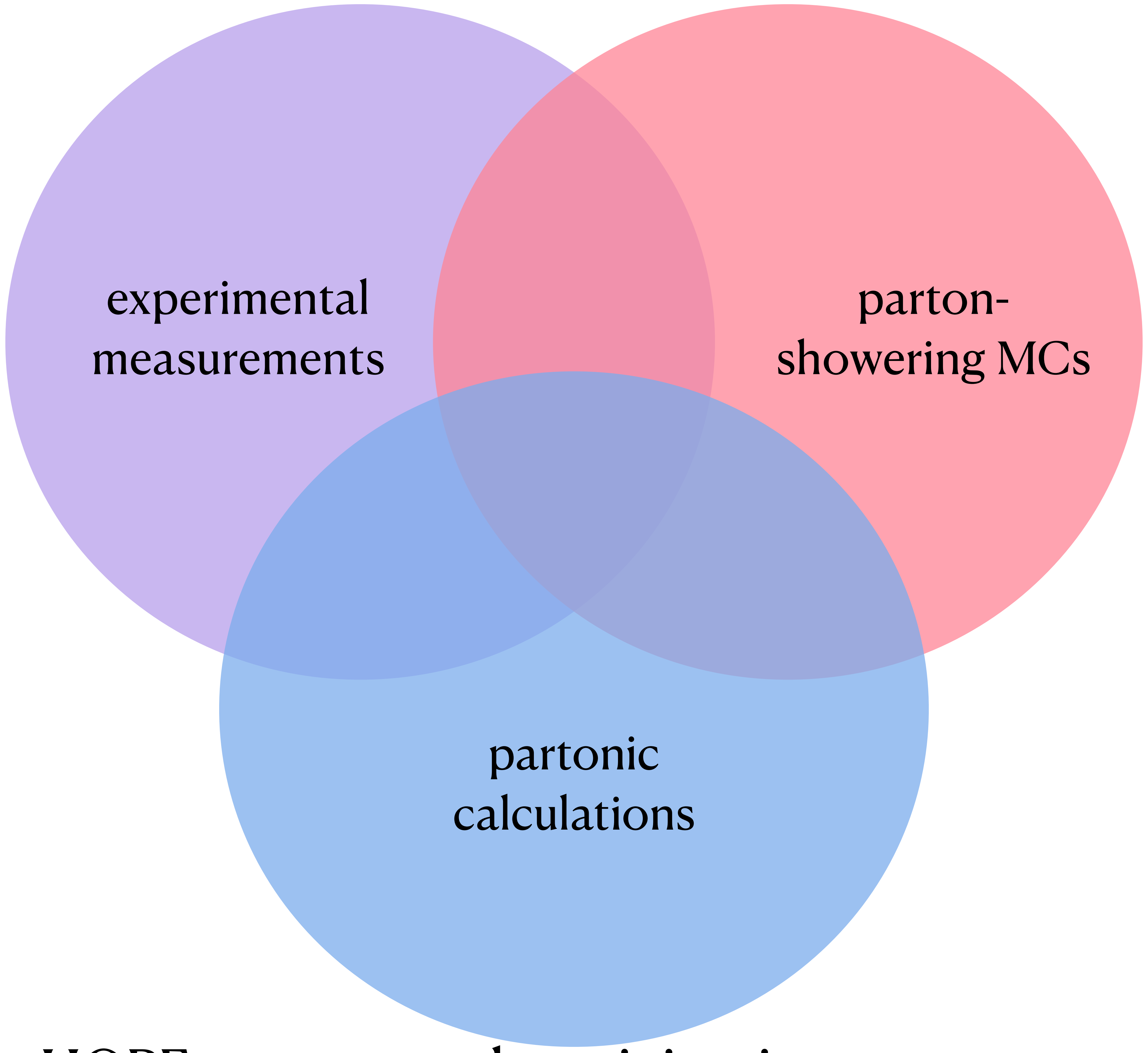


How many jets?





What makes a jet?



- How to group particles into a jet and assign the momentum?
- Define a radius (jet R)
- What is the momentum (p_T)

- Jets (in experiment) can have:
  - inherent angular scale (θ)
  - momentum fraction (z) due to vacuum splitting

- Jet = group of particles grouped together by some algorithm - HOPE to recover the originating parton BUT jet $\neq$ parton

Challenges in Defining Jets

- Not always obvious where to place jet cones (physically or computationally)
- How to define events with more than 2 cones?
- Overlapping cones?

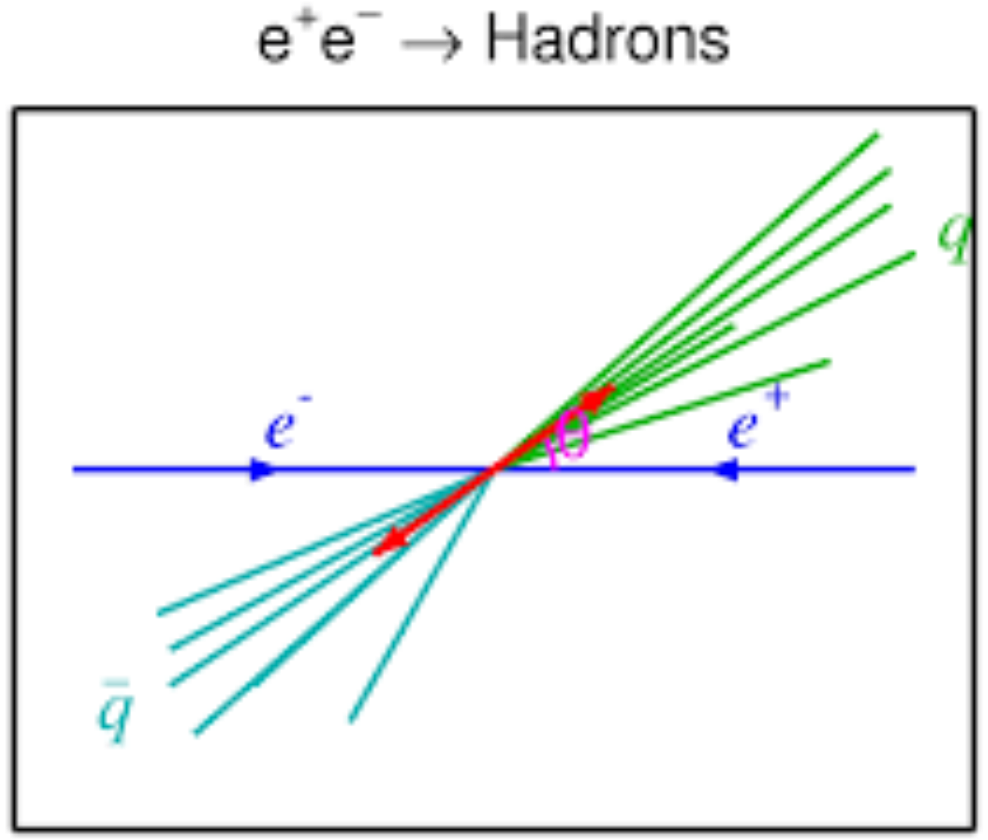
- Ambiguous association between jets and partons
  - ex. boosted $W(Z) \rightarrow XX$

- Computational power scales with multiplicity, N

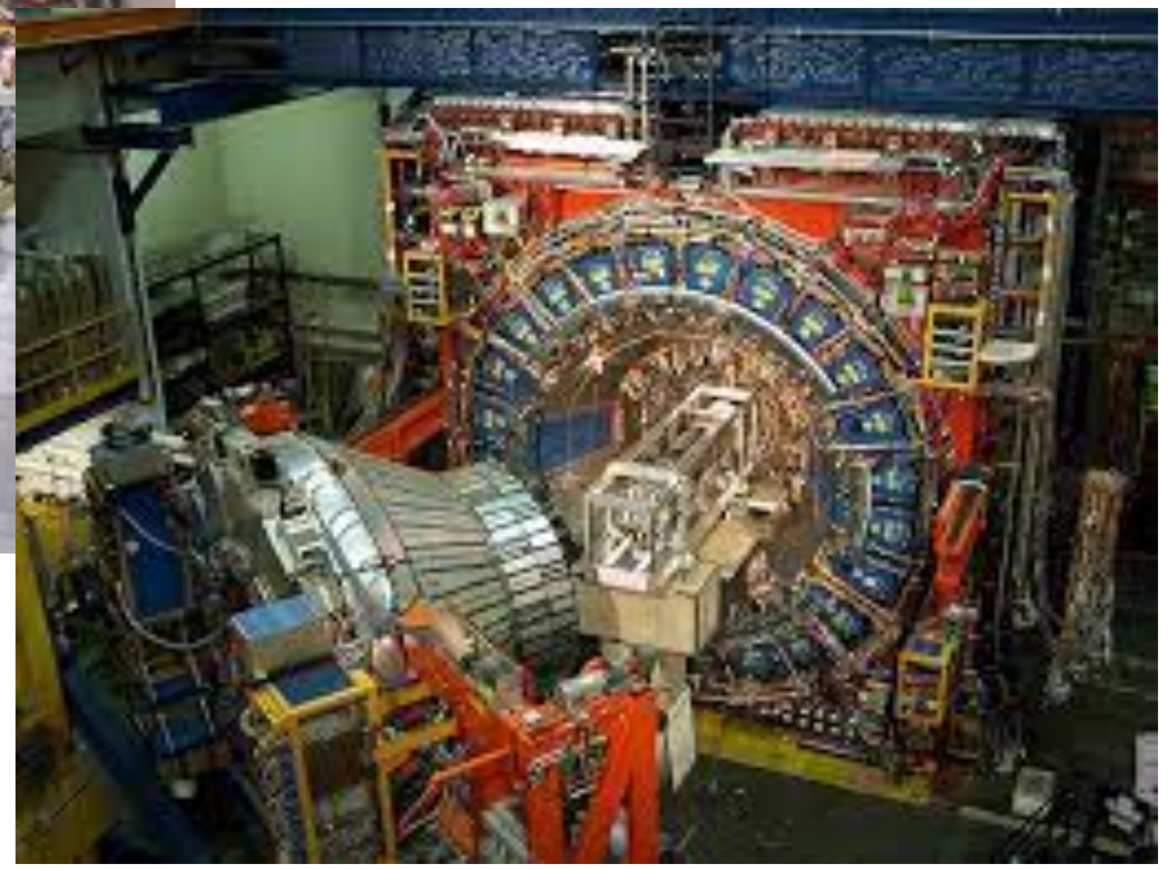
History of Jet Algorithms

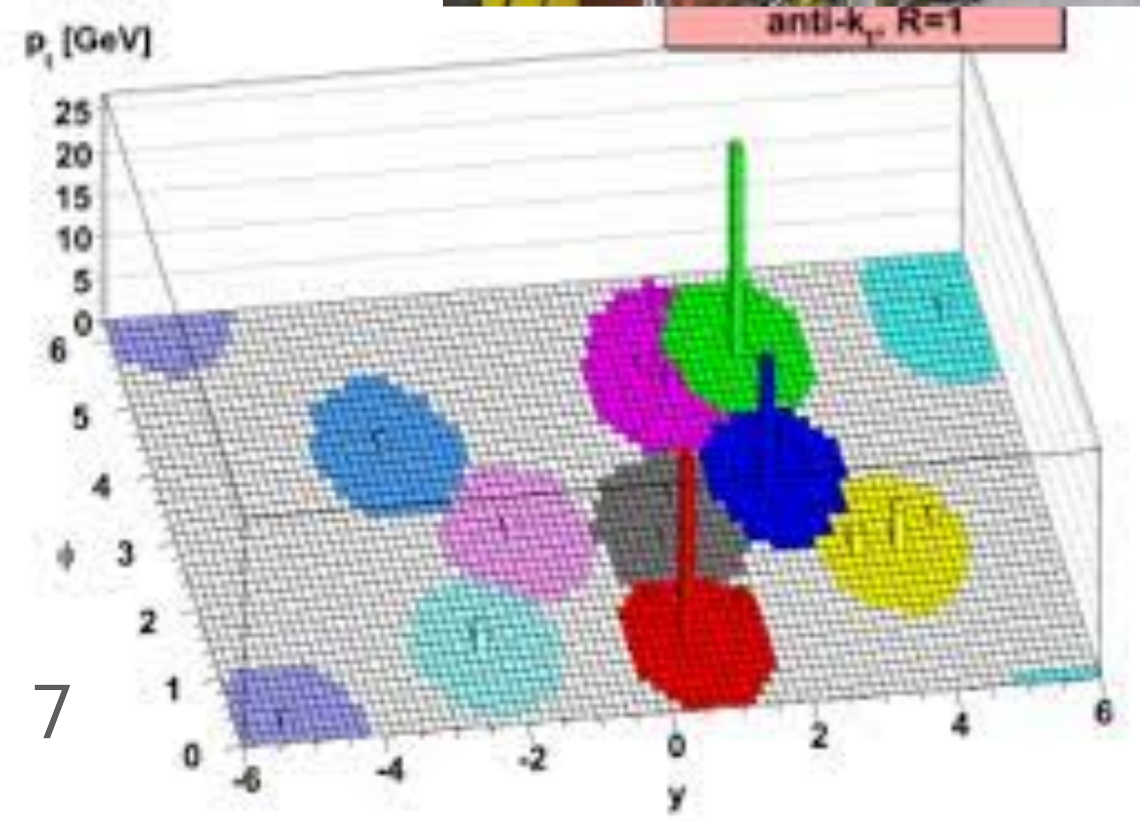
History of Jet Algorithms

Key Dates



- 1977: Sterman-Weinberg cone-type algorithms for QCD
- 1980s: Iterative cone algorithms for hadron colliders (SppS, Tevatron)
- 1990s: New algorithms, LEP
  - k_T: a theory-friendly approach
  - Start of sequential recombination algorithms
- 1999: C/A
- 2007: SISCone
- 2008: anti-k_T

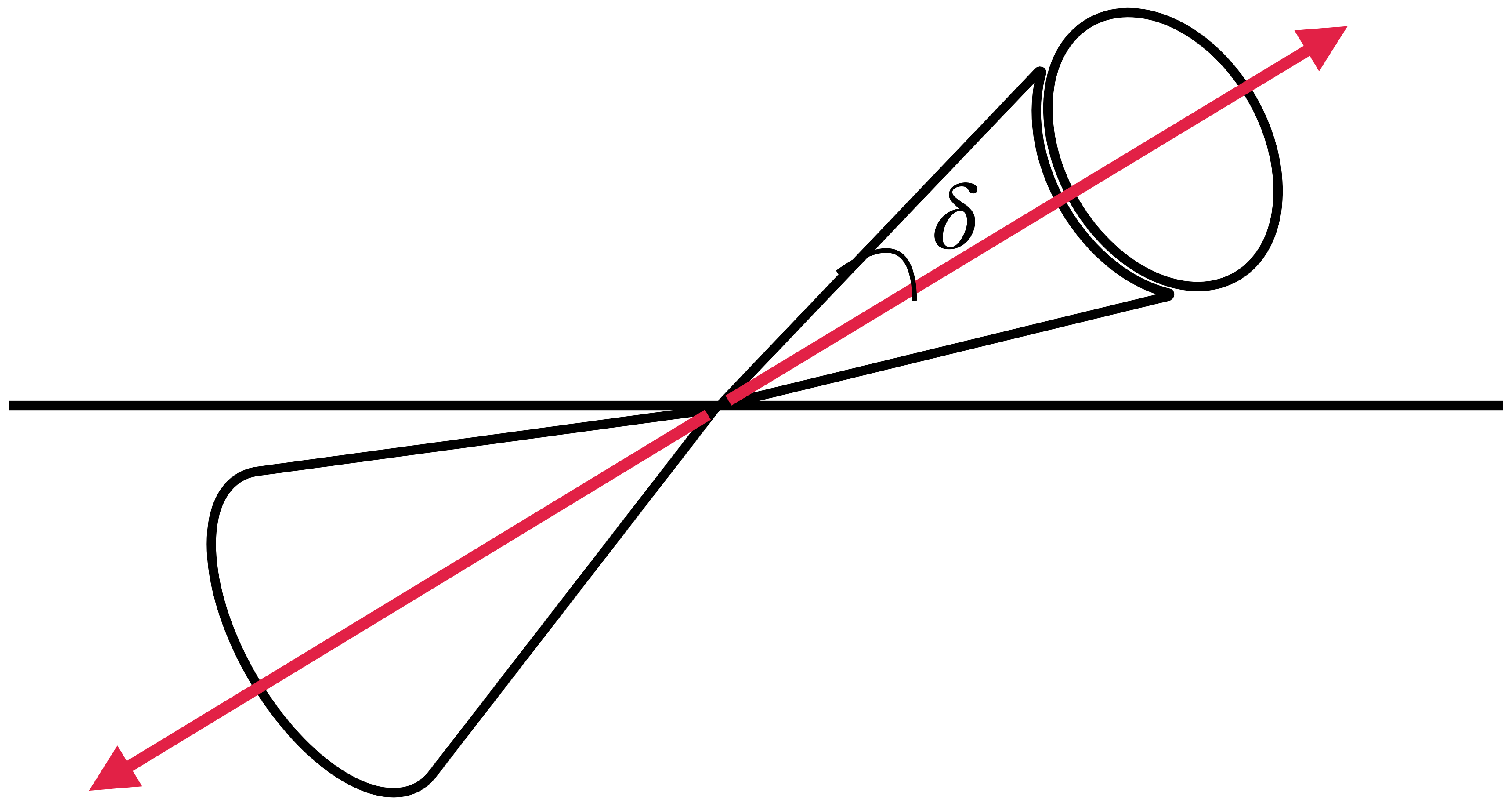




Sterman-Weinberg Jet Definition

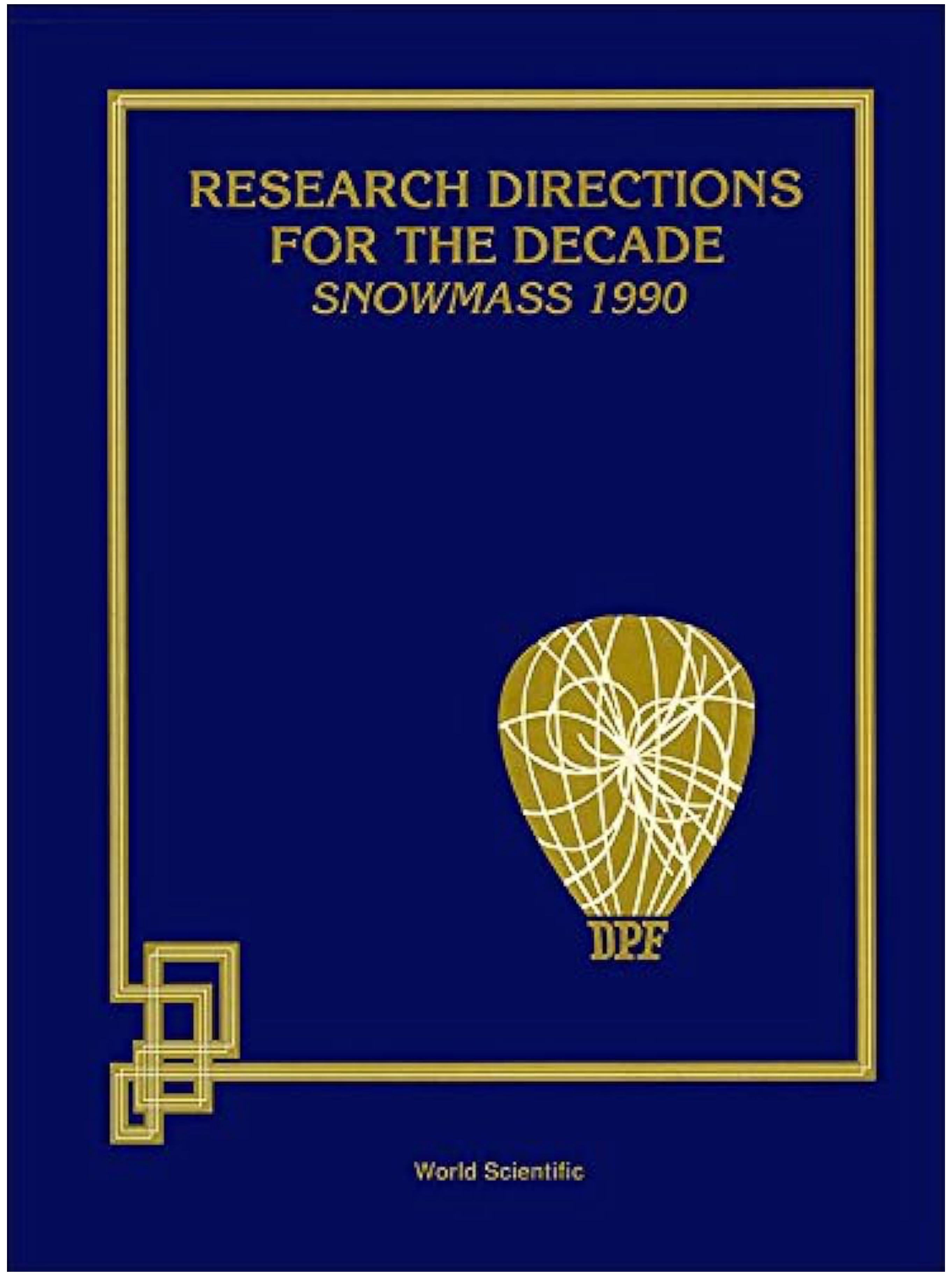
1977

- Intended for e+e- collisions
  - Cross section dominated by 2-jet events
- 2-jet event = at least a fraction $1 - \epsilon$ of the event's energy was contained in two cones of opening half-angle δ
  - Arbitrary parameters
- But…
  - What about > 2 jets? Where are the cones placed?
  - What about the energy in hadron colliders?



1990

“Snowmass accord”



Set of rules to define a jet

1. Simple to implement in an experimental analysis
2. Simple to implement in the theoretical calculation
3. Defined at any order of perturbation theory
4. Yields finite cross sections at any order of perturbation theory
5. Yields a cross section that is relatively insensitive to hadronisation

https://arxiv.org/pdf/0906.1833.pdf

IRC Safety

- Theoretical requirements to insure that jets remain unchanged due to hadronisation or orders of perturbation theory
- **Infrared safety**
  - the set of jets will not change if some soft emission is added to the event
- **Collinear safety**
  - the set of jets will not change if the event is modified by a collinear splitting

Perturbative calculations for jet observables are only possible if jets are IRC safe!

Recombination Schemes

- **(Massless) E_t weighted**
  - Recombine particles according to $E_{t,\text{jet}} = \sum_i E_{ti},\ \eta_{\text{jet}} = \frac{1}{E_{t,\text{jet}}} \sum_i E_{ti}\eta_i,\ \phi_{\text{jet}} = \frac{1}{E_{t,\text{jet}}} \sum_i E_{ti}\phi_i$
  - Not invariant under longitudinal boosts if particles are massive
- **E-scheme**
  - Sums the components of the four-vectors
  - Most commonly used
- **Winner-Take-All (WTA)**
  - Recombine the two particles with the y, ϕ, and m of the particle with the larger p_t, and a p_t equal to the sum of the two p_t's
  - Reduces effects related to the recoil of the jet axis when computing jet observables that share similarities with the event-shape broadening

Two Categories of Jet Algorithms

- **Iterative cone algorithms**: finding regions where lots of energy flows
  - Rigid, circular jet boundaries
  - Works because QCD branching and hadronisation only modifies energy flow at small scales
  - Ex. IC-PR, IC-SM, SISCone
- **Sequential recombination algorithms**: combine particles by closest distance
  - Repeatedly combine the closest pair of particles in momentum space
  - Closely connected with probabilistic pictures for parton branching
  - Ex. k_t, C/A, anti-k_t

Iterative Cone Algorithms

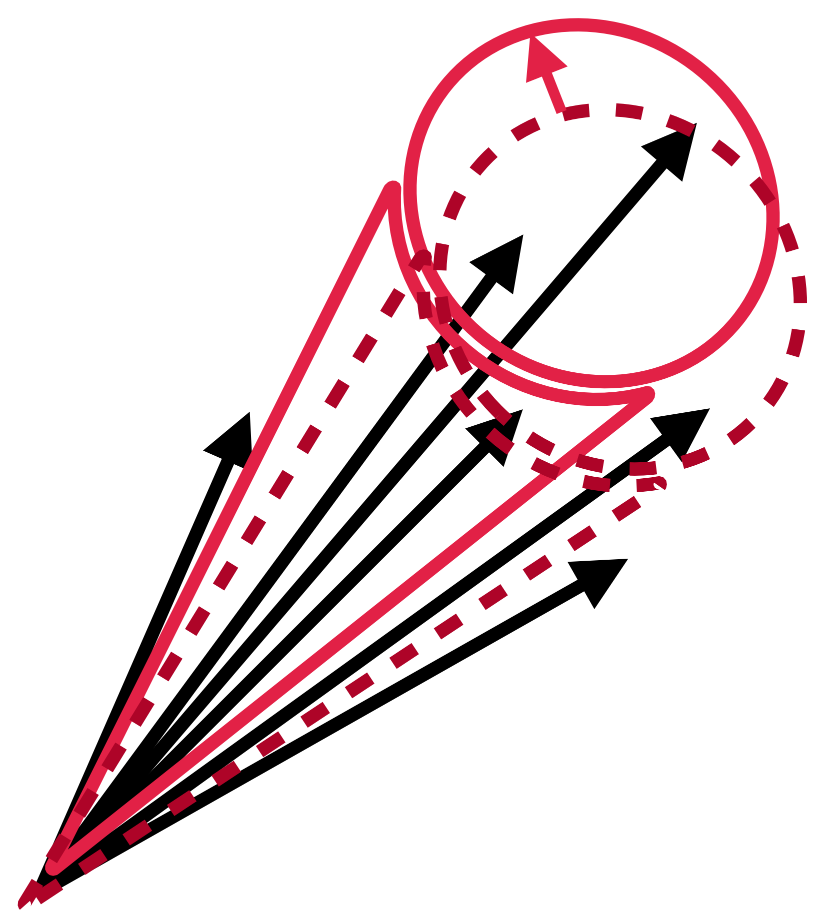
Iterative Cone

A key component to the iterative cone algorithms is how to find a stable cone:

- Seed particle i sets some initial direction
- Sum the momenta of all particles j within a "cone" of radius R around i in (ϕ and y)

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 < R^2$$

- Direction of the resulting sum is used as the new seed direction
- Repeat until direction of the resulting cone is stable

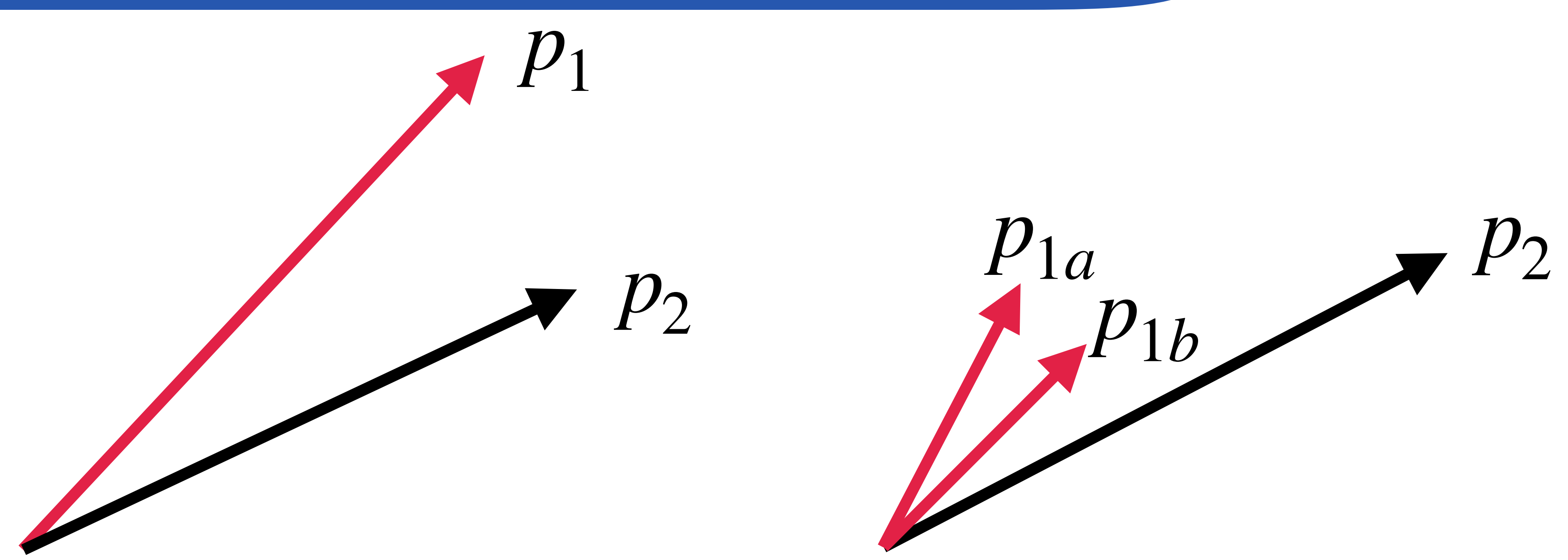


- No specific energy cut, but later only consider jets above a certain p_T threshold

Iterative Cone Approaches

Progressive Removal (IC-PR)

- Take hardest seed particle in the event, iterate to find a stable cone -> jet!
- Remove particles contained in that jet from the event
- Repeat procedure with remaining particles until no seeds left
- Collinear unsafe!



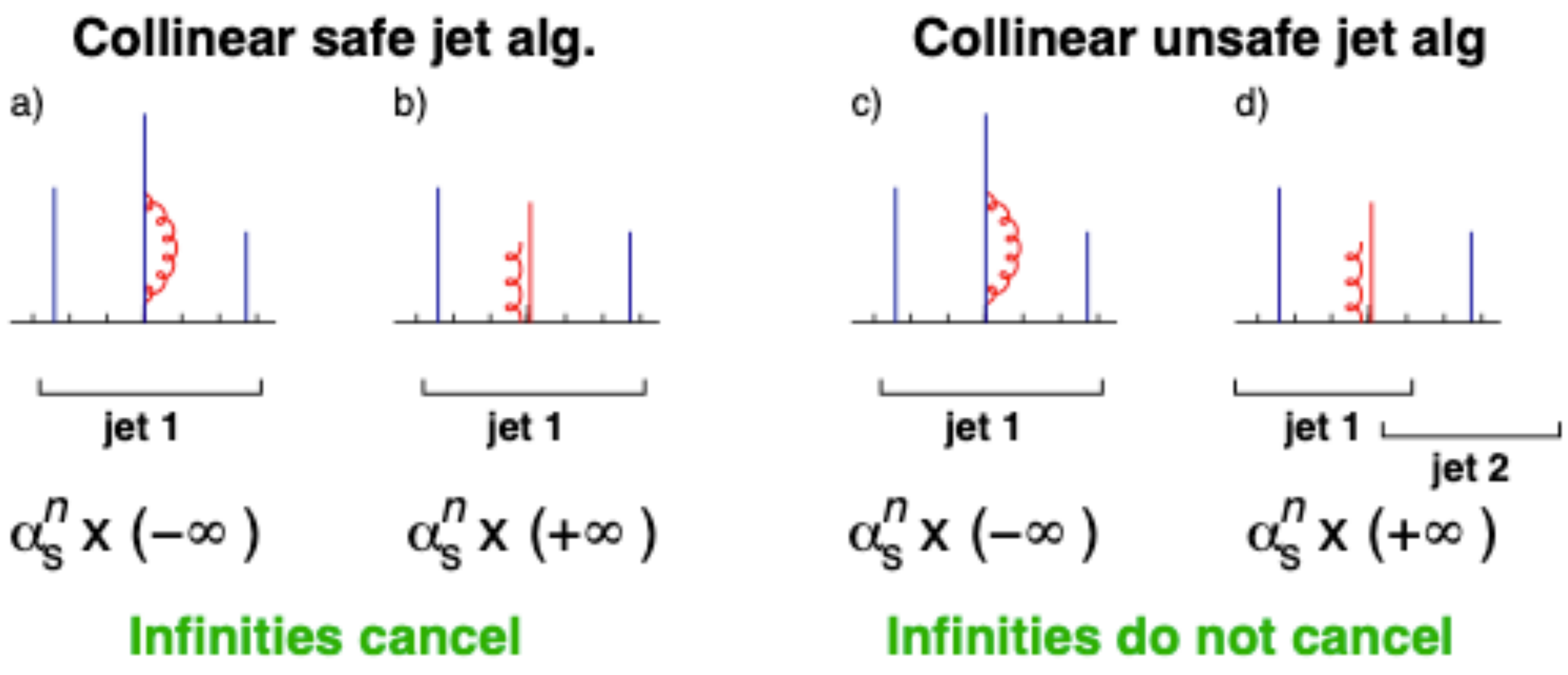
- **Fixed Cones (IC-FC)**:
  - Do not iterate the cone direction,

    but instead identify a fixed cone around the seed direction and call that a jet

Iterative Cones and IRC Safety

Progressive Removal (IC-PR)

- Vertical lines are partons, and their heights translate to transverse momentum. The horizontal axis is rapidity.
- (a) and (b) are some collinear safe jet algorithm
- (c) and (d) are IC-PR algorithm
- For (c), the middle jet provides a seed first
- In (d), leftmost particle provides the seed first, and misses adding rightmost particle into that jet



NOT a collinear safe algorithm!

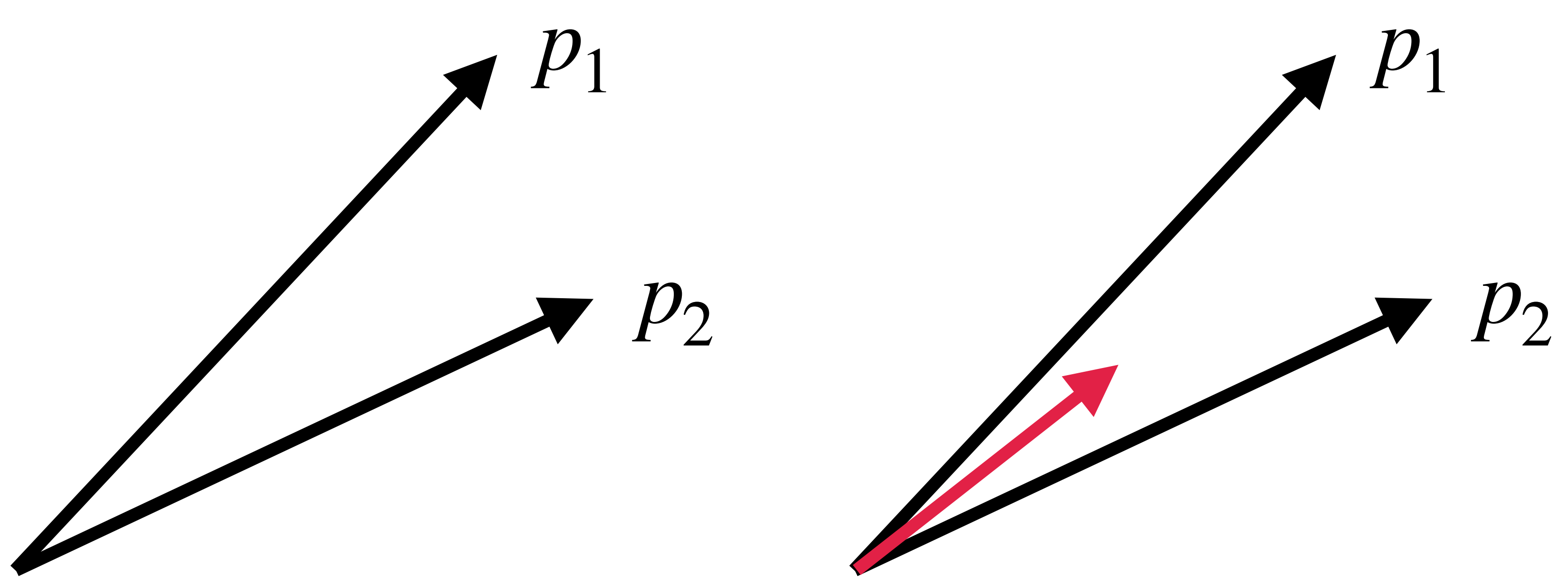
Iterative Cone Approaches

Split Merge (IC-SM)

Process:

- Iterate over all seeds and find all stable cones (of radius R) = "protojets"
  - Can be above some seed threshold
- Run split-merge procedure to deal with overlapping stable cones
  - Label largest p_T protojet *a*
  - Label the next largest p_T protojet *b*. If no such protojet, add *a* to the final jet list
  - If $p_{T,\text{shared}}/p_{t,b} > f$, merge protojets *a* and *b*
    - f = overlap threshold, a free parameter, generally chosen to be 0.5 or 0.75
  - Otherwise, assign shared particles to the closer cone
  - Then repeat

- Infrared unsafe!

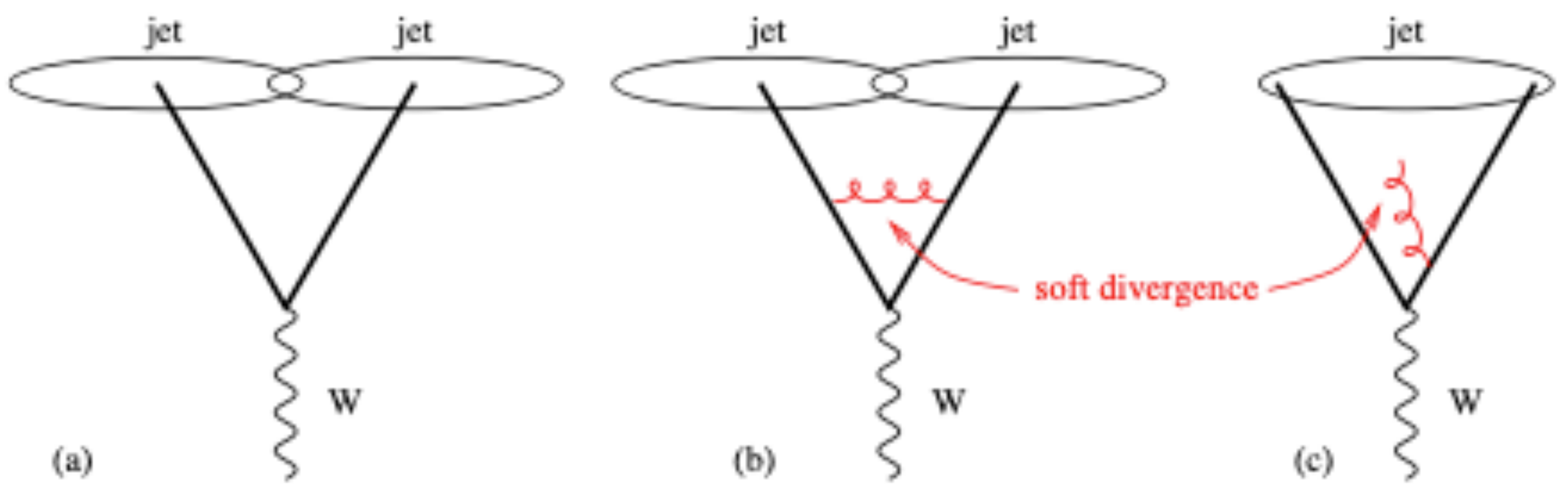


- **Split-Drop (IC-SD)**:
  - Drop the non-shared particles that are part of the softer of two overlapping cones
  - Ex. PxCone

Iterative Cones and IRC Safety

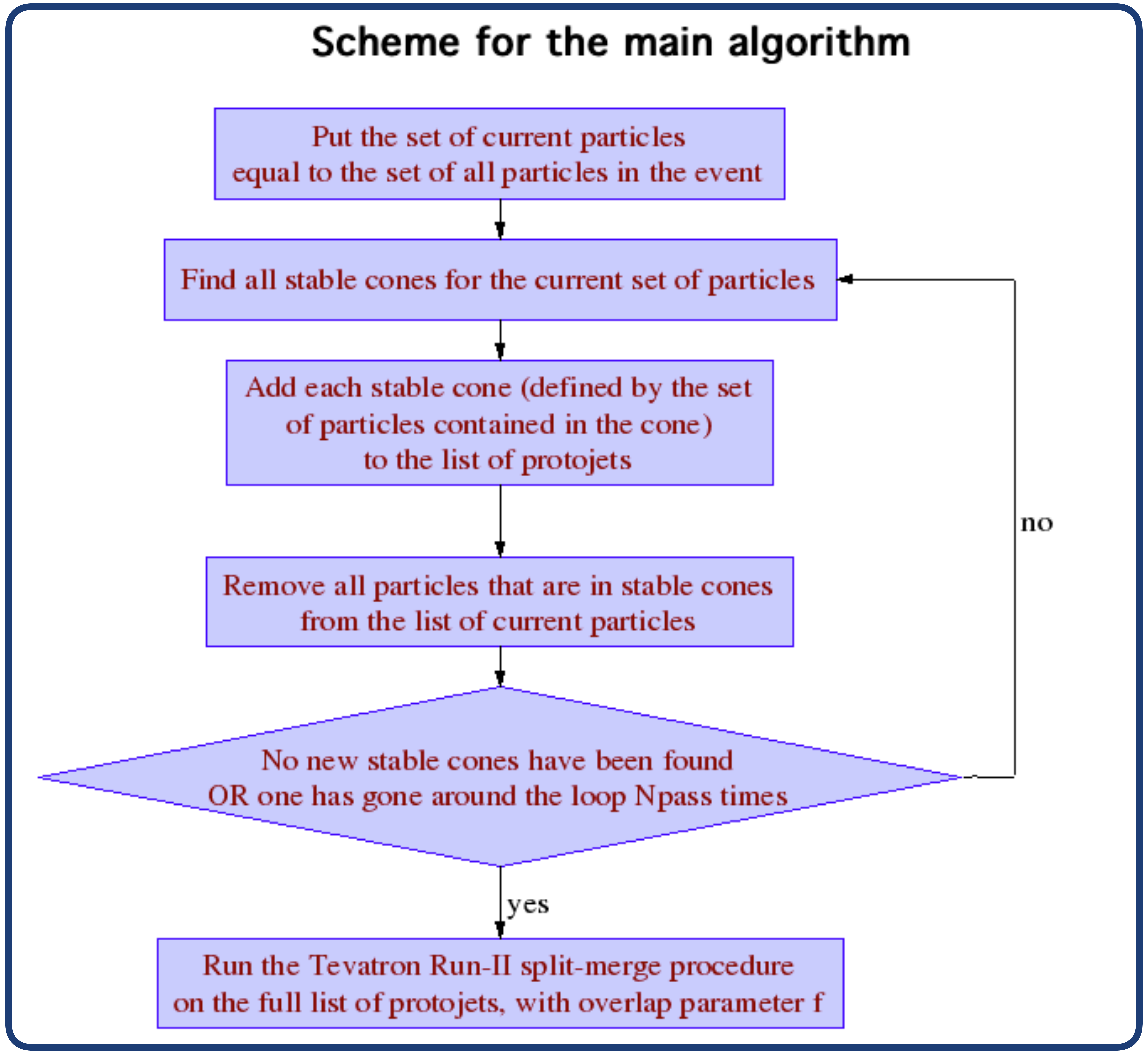
Progressive Removal (IC-SM)

- An extra soft particle can act as a new seed and find a new stable cone
- (a) and (b) show two stable cones and two jets
- (c) has an extra soft gluon that provides a new seed
  - If two hard partons have similar momenta and are separated by $< 2R$, new stable cone is found

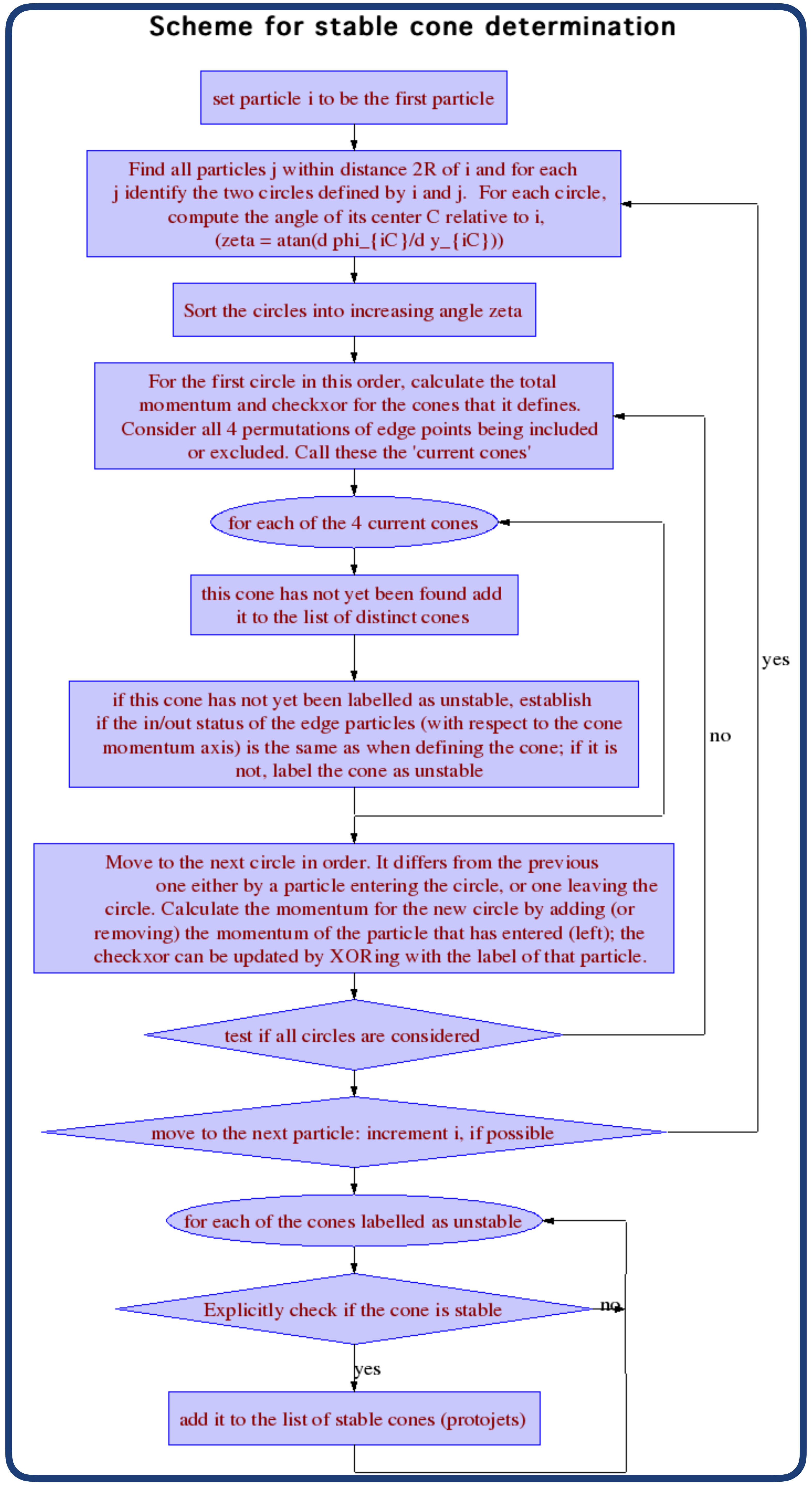


NOT an infrared safe algorithm!

SISCone



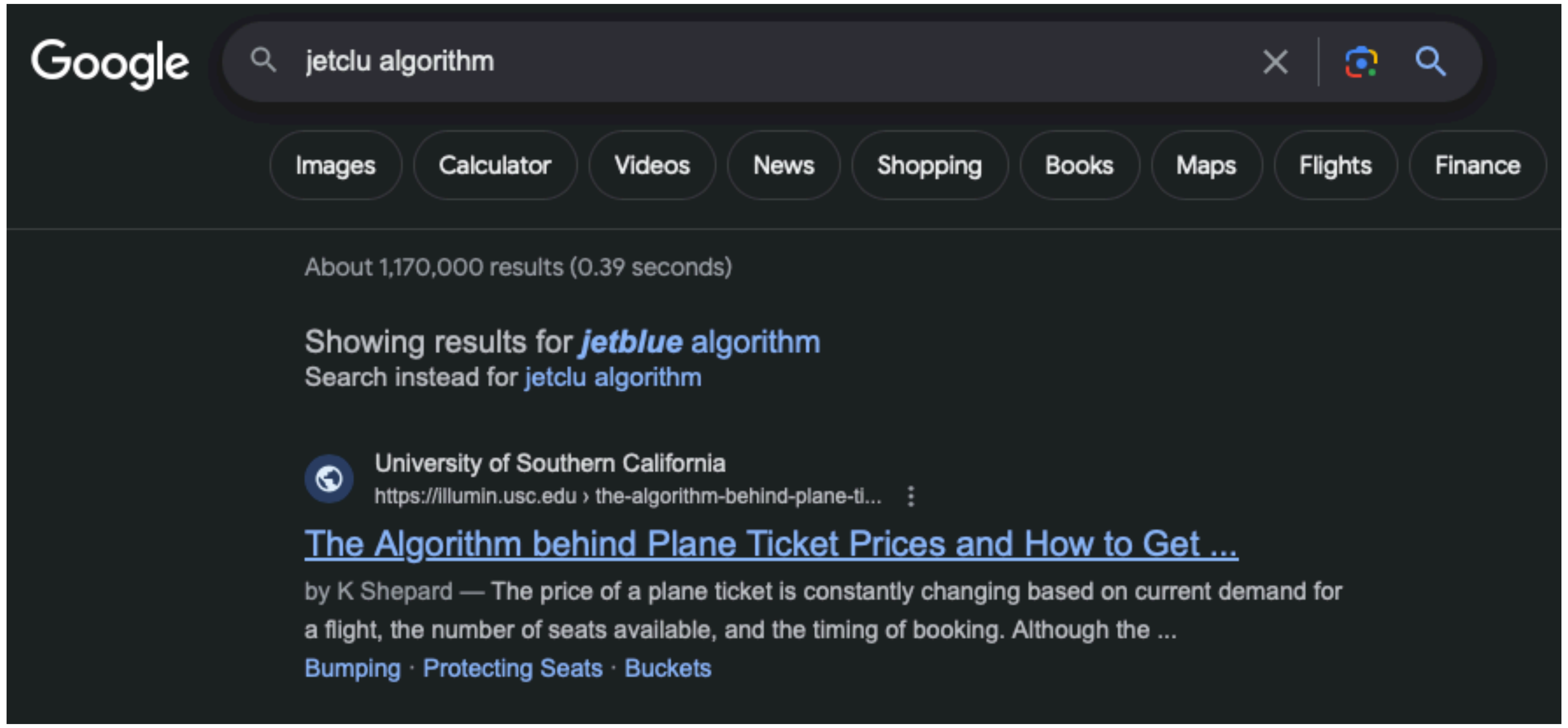
- Seedless Infrared-Safe Cone Algorithm
- Only IRC-safe cone algorithm
- Is soft-resilient, but would differ perturbatively from IC-PR at NLO



Specific Iterative Cone Algorithms

- IC-PR:
  - CMS iterative cone
  - Pythia CellJet
- IC-SM:
  - JetClu (CDF, Run II)
  - Midpoint cone (CDF, Run I & II)
  - DØ cone
  - ATLAS cone

*Note: many of these algorithms were implemented in FastJet, but some have since been deprecated or are harder to find



Sequential Recombination Algorithms

Sequential Recombination

General Procedure

- d_{ij} = distance between two entities
- d_{iB} = distance between an entity and the beam
- Find all d_{ij} and d_{iB}
- Find the smallest of the differences
  - If d_{ij}: recombine entities *i* and *j*
  - If d_{iB}: call *i* a jet and remove it from the list of entities
- Then recalculate distances and repeat process until no entities left

$$d_{ij} = \min(k_{T_i}^{2p}, k_{T_j}^{2p})\frac{\Delta_{ij}^2}{R^2}$$

$$d_{iB} = k_{T_i}^{2p}$$

$$\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

p = an arbitrary parameter

(inclusive) k_t: p = 1

CA: p = 0

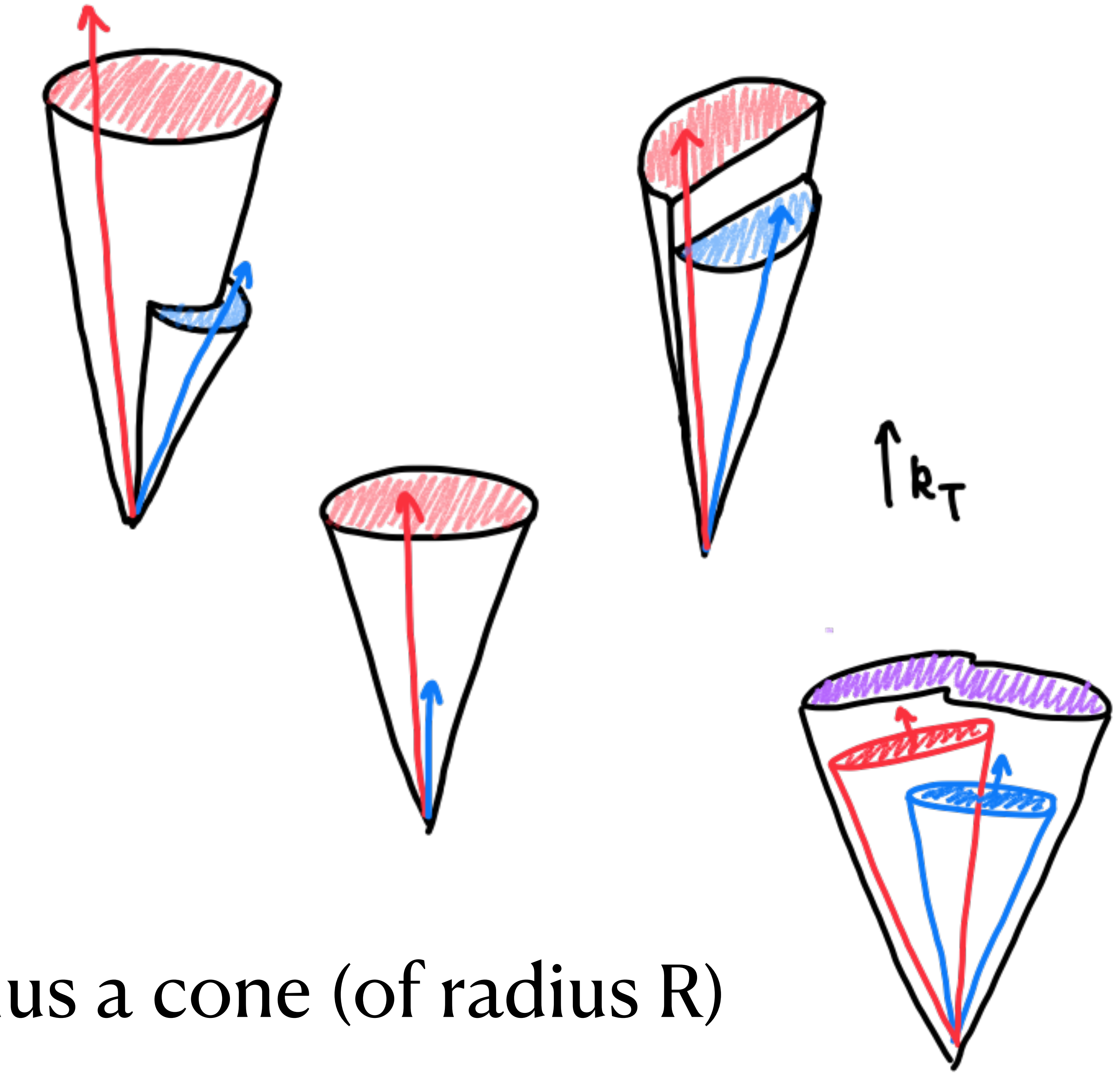
anti-k_t: p = -1

Anti-k_t

Behavior

$$d_{ij} = \min\left(\frac{1}{k_{T_i}^2}, \frac{1}{k_{T_j}^2}\right)\frac{\Delta_{ij}^2}{R^2}$$

- If a hard particle has no hard neighbors within a distance $2R$, it will accumulate all the soft particles within a circle of radius R —> perfectly conical jet
- If another hard particle 2 is present s.t. $R < \Delta_{12} < 2R$ —> 2 hard jets
  - If $k_{t1} \gg k_{t2}$ —> jet 1 will be conical, and jet 2 will be partly conical
  - If $k_{t1} \approx k_{t2}$ —> jets will be split by a straight line through the middle
- If $\Delta_{12} < R$ —> particles 1 and 2 will cluster to form a single jet
  - If $k_{t1} \gg k_{t2}$ —> conical jet centered about k_{t1}
  - If $k_{t1} \approx k_{t2}$ —> union of cones (radius < R) around each hard particle plus a cone (of radius R) centered on the final jet

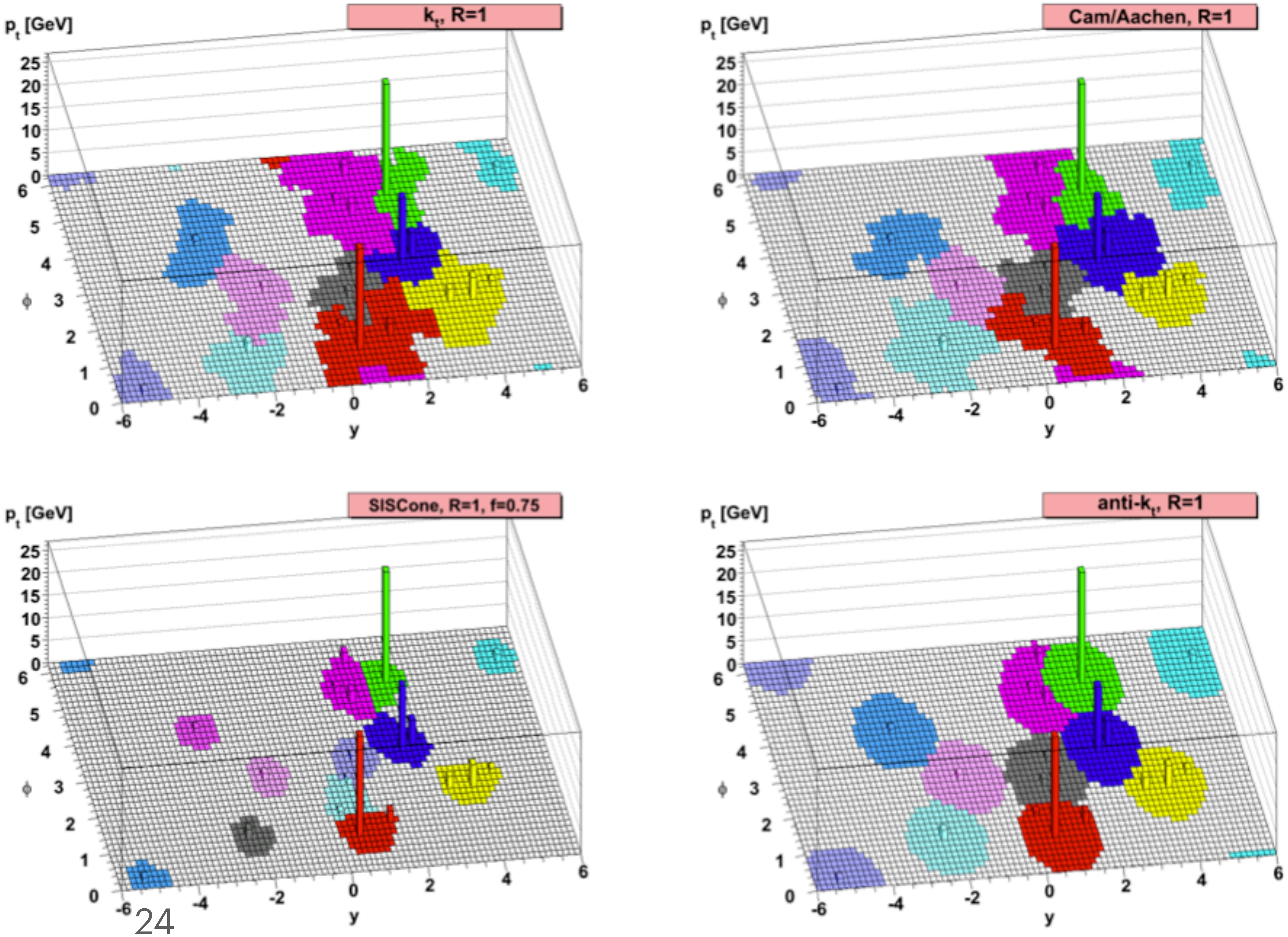


Cacciari, Salam, Soyez arXiv:0802.1189

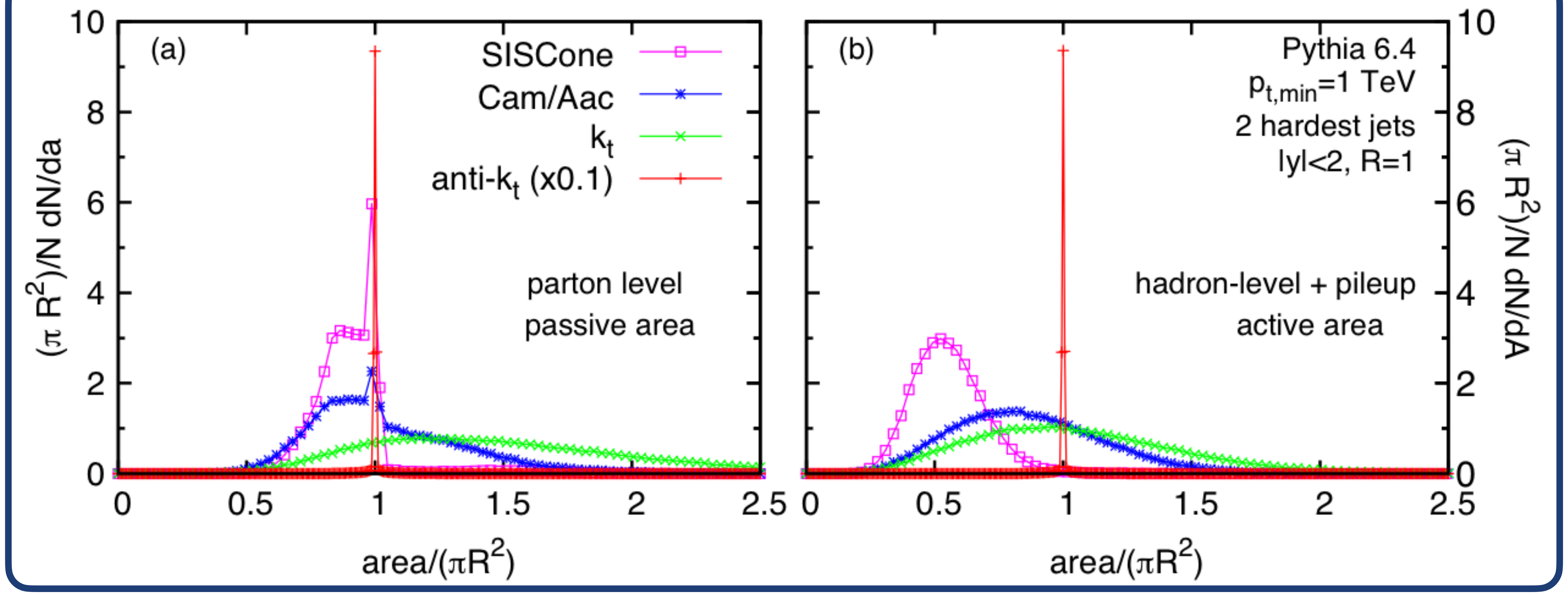
pictures thanks to Anjali Nambrath!

Jet Shapes

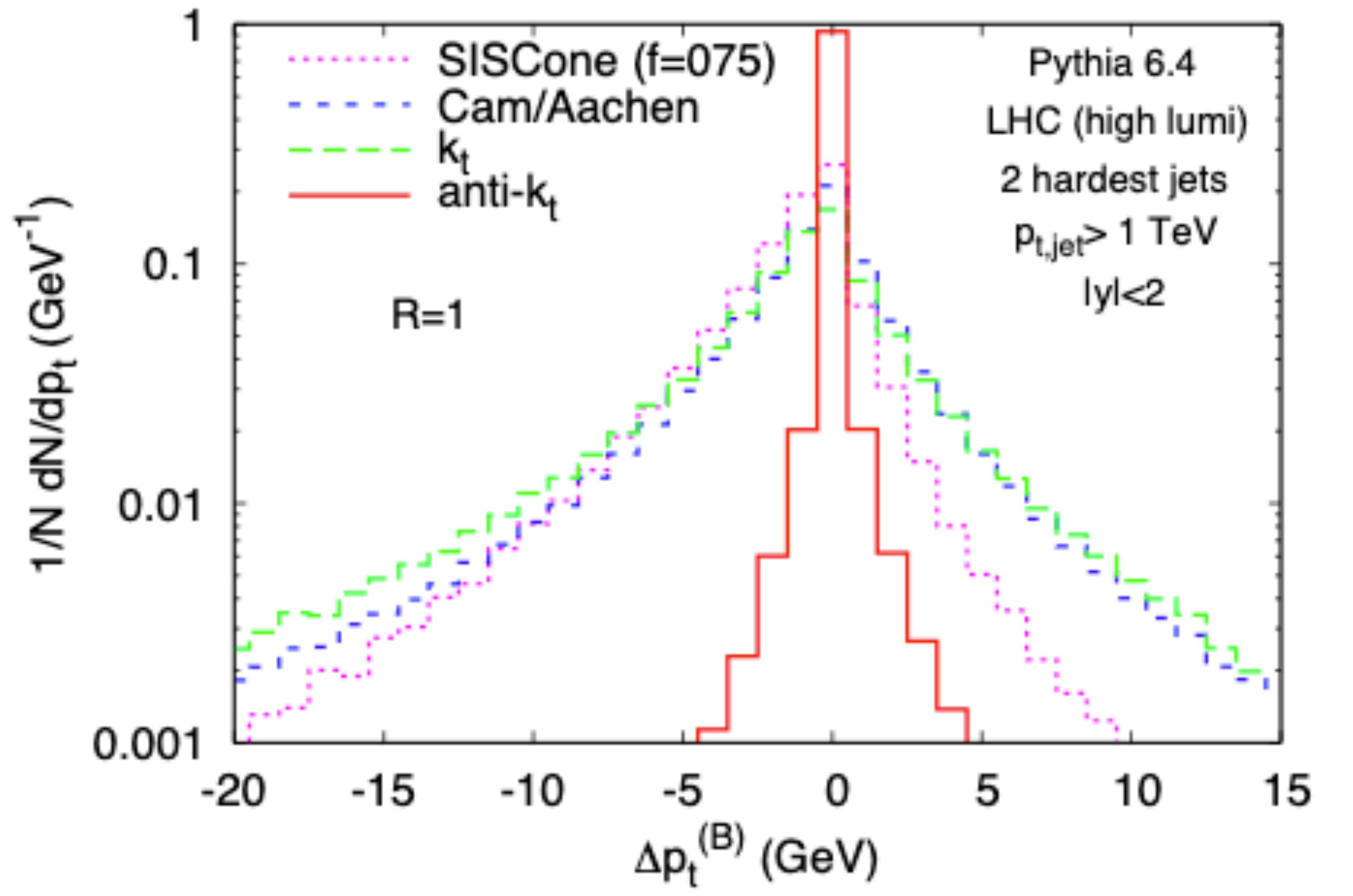
- Parton-level event together with ~10^4 random soft 'ghost' particles
- k_t and CA: region depends on set of (random) ghosts
- SISCone: single-particle jets are regular, composite jets have more varied shapes
- anti-k_t: circular jets



Jet Area-Related Properties



- Back reaction = having a different subset of jets due to additional UE/pileup
- Back reaction suppressed for anti-k_t



- Jet "area" = measure susceptibility to soft radiation
  - passive = point-like radiation
  - active = diffuse radiation
- anti-k_t has the most stable jet area

Using Jet Algorithms

Pros and Cons of Different Algorithms

	Speed	Regularity	UE contamination	Backreaction	Hierarchical substructure
k_t	☺ ☺ ☺	☂	☂☂	☁☁	☺ ☺
Cambridge /Aachen	☺ ☺ ☺	☂	☂	☁☁	☺ ☺ ☺
anti-k_t	☺ ☺ ☺	☺ ☺	☁/ ☺	☺ ☺	✘
SISCone	☺	☁	☺ ☺	☁	✘

Array of tools with different characteristics.
Pick the right one for the job

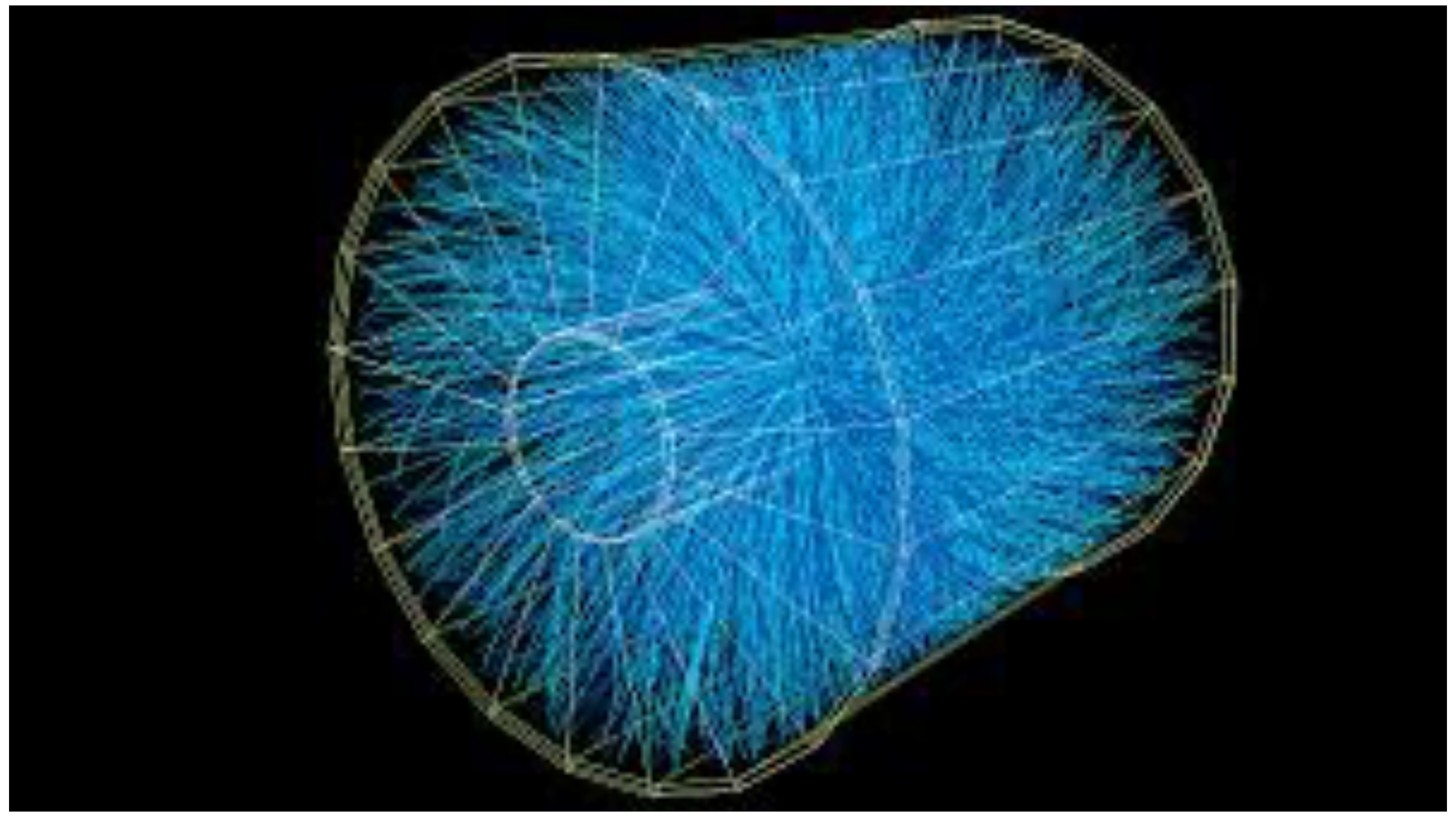
Pythia, Fastjet

- Many implementations for theoretical and phenomenological comparisons
  - Present-day LHC, HERA, LEP use(d) sequential algorithms
  - Tevatron & preparatory LHC work used cone algorithms
- Faster clustering algorithms

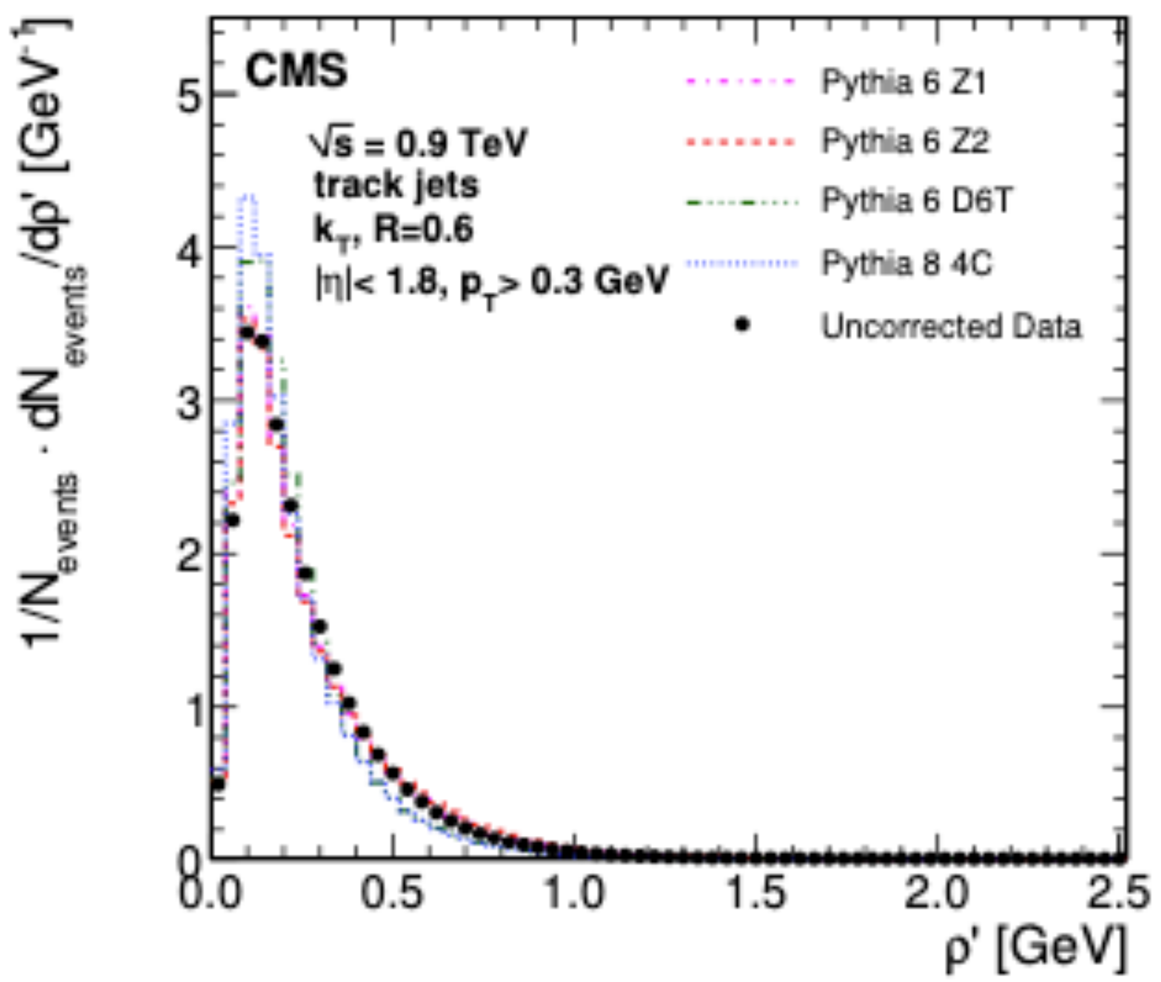
Cacciari, Salam, Soyez arXiv:1111.6097

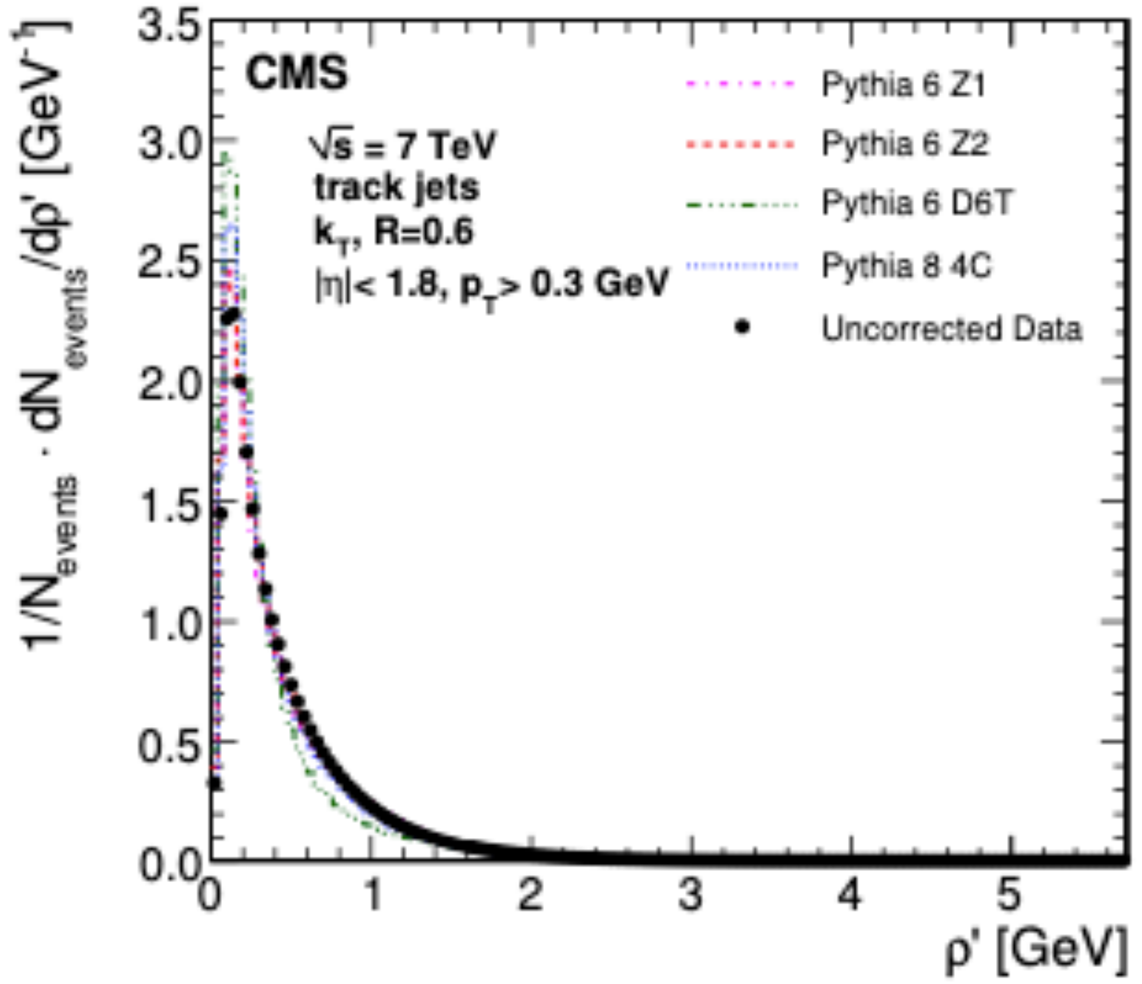
Jets can be used as tools or observables!

Jets in Heavy Ions



- Big UE
- Reclustering: provides hierarchical tree of jet evolution
  - Use C/A
    - Merges smallest angles together first
  - Substructure observables





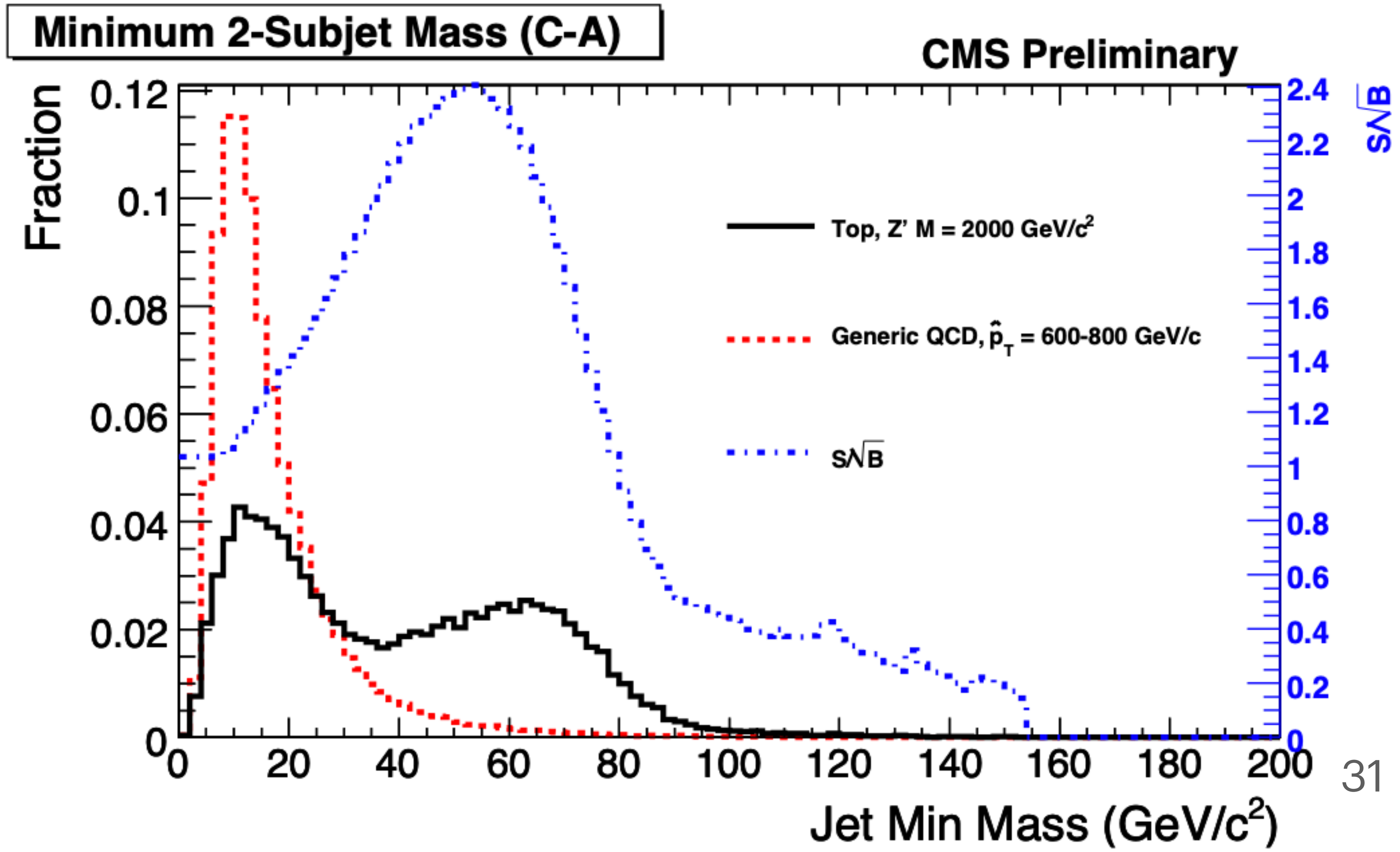
Data-driven background subtraction

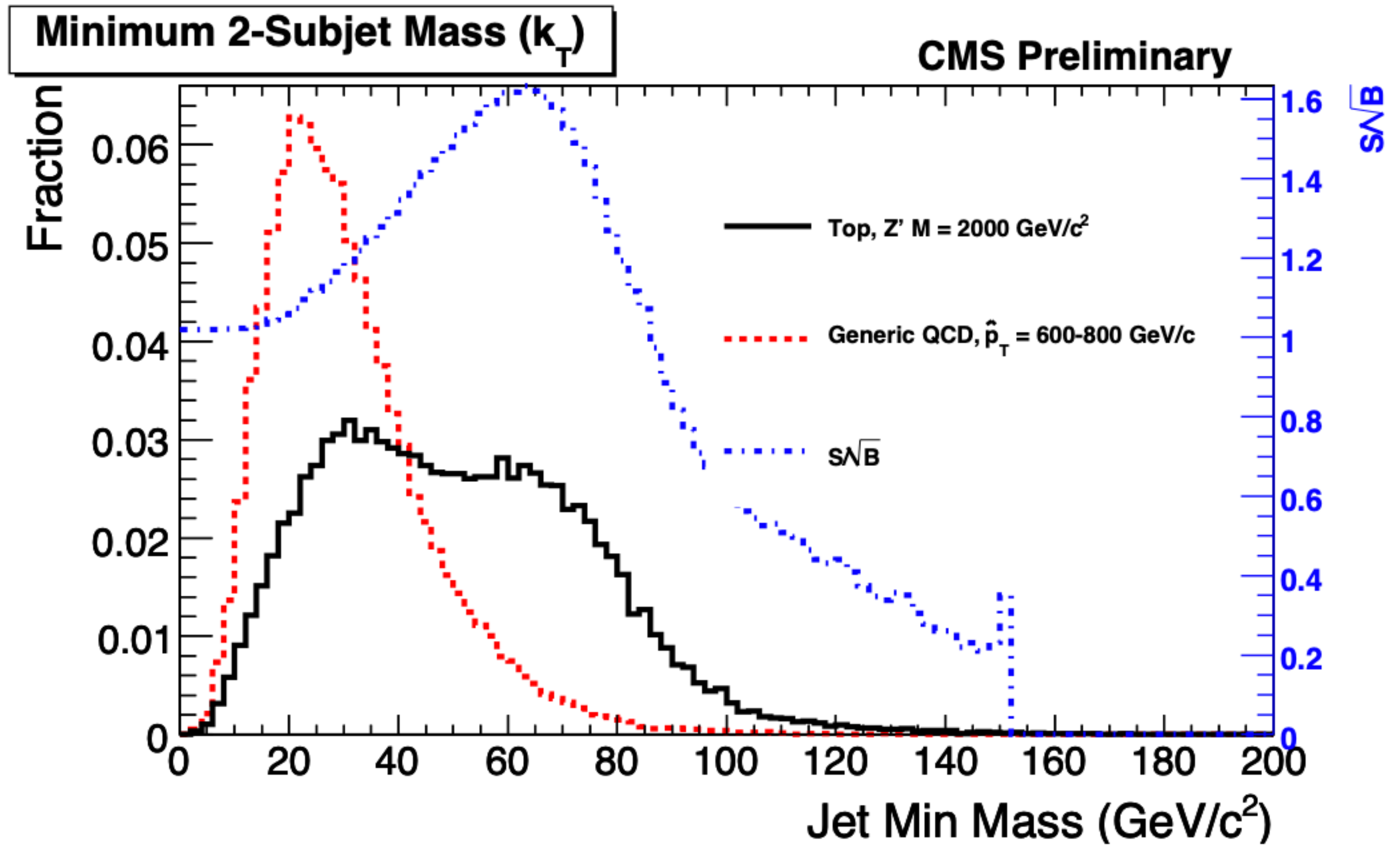
- anti-k_t clusters jets, but leaves some UE
  - UE inflates jet p_T
  - Also cluster with k_t to access soft particles
  - Find energy density (p_T/area) of median of k_t jets —> scale anti-k_t jet area and subtract off of anti-k_t jets

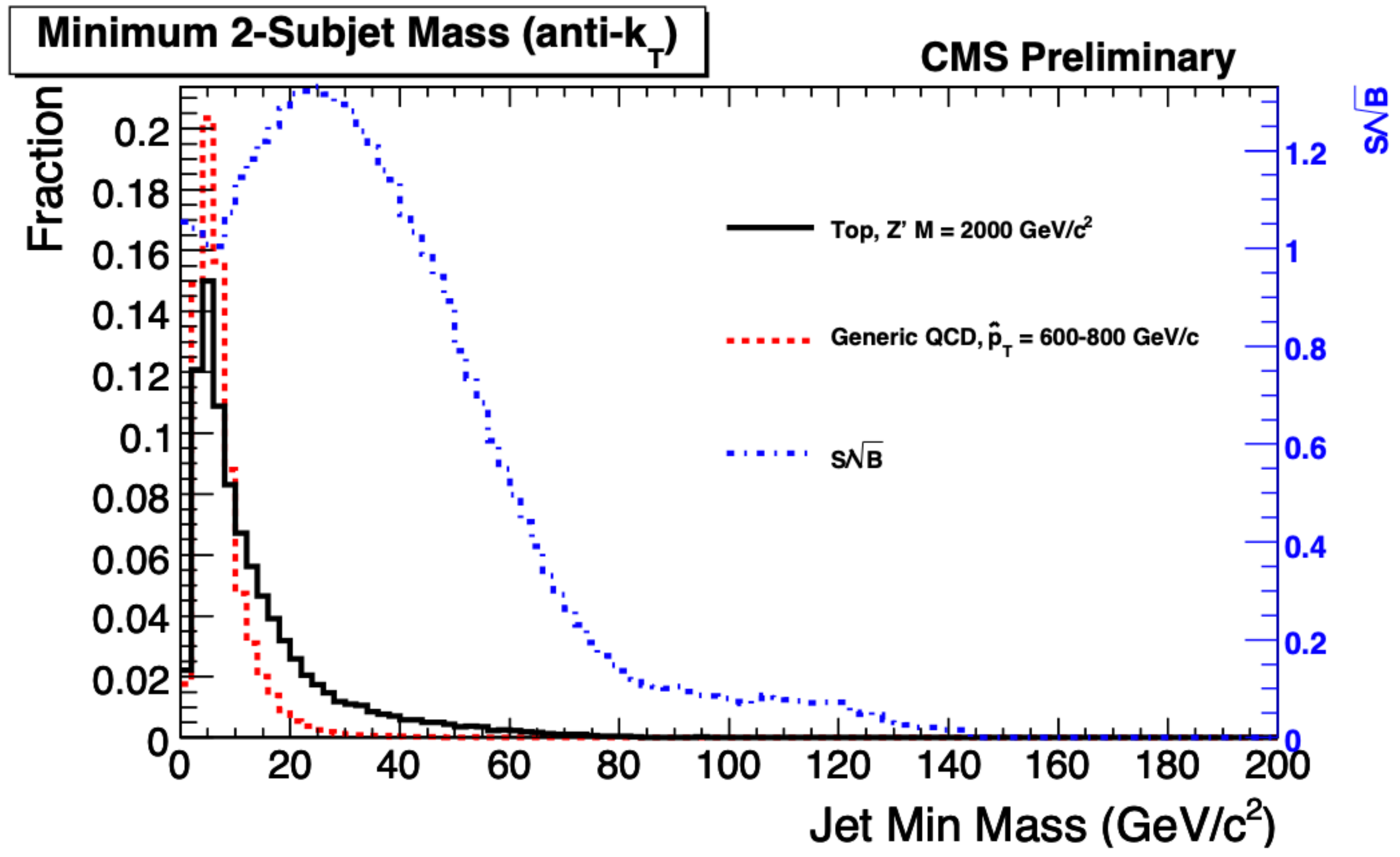
https://link.springer.com/content/pdf/10.1007/JHEP08(2012)130.pdf?pdf=button

Top Mass Discrimination

- Look at the angular ordering in jets
- Top mass discrimination
- $S/\sqrt{B}$ = significance of signal
- C/A selects subjets closest to the hard jet axis





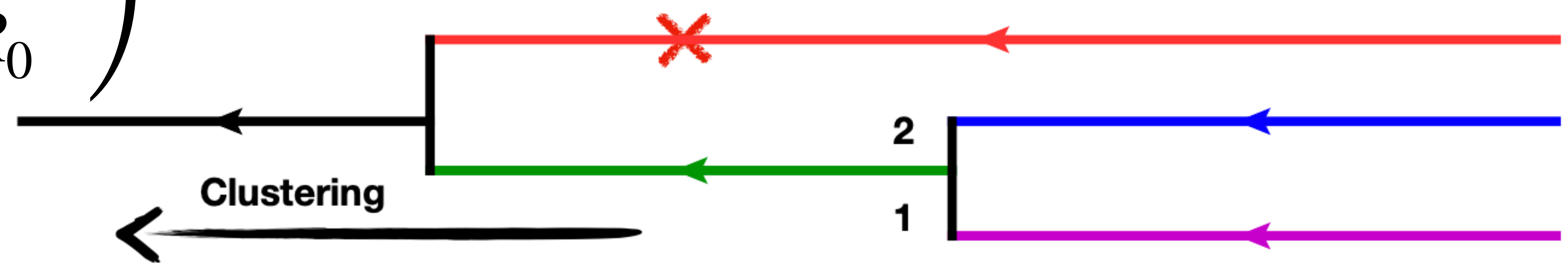


https://cds.cern.ch/record/1194489/files/JME-09-001-pas.pdf

Soft Drop

$$d_{ij} = \frac{\Delta_{ij}^2}{R^2}\,,\quad d_{iB} = k_{T_i}^{2p}$$

- Use anti-k_t to cluster, use C/A to de-cluster
  - Soft Drop Condition : $\dfrac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{\text{cut}} \left(\dfrac{\Delta R_{12}}{R_0}\right)^{\beta}$

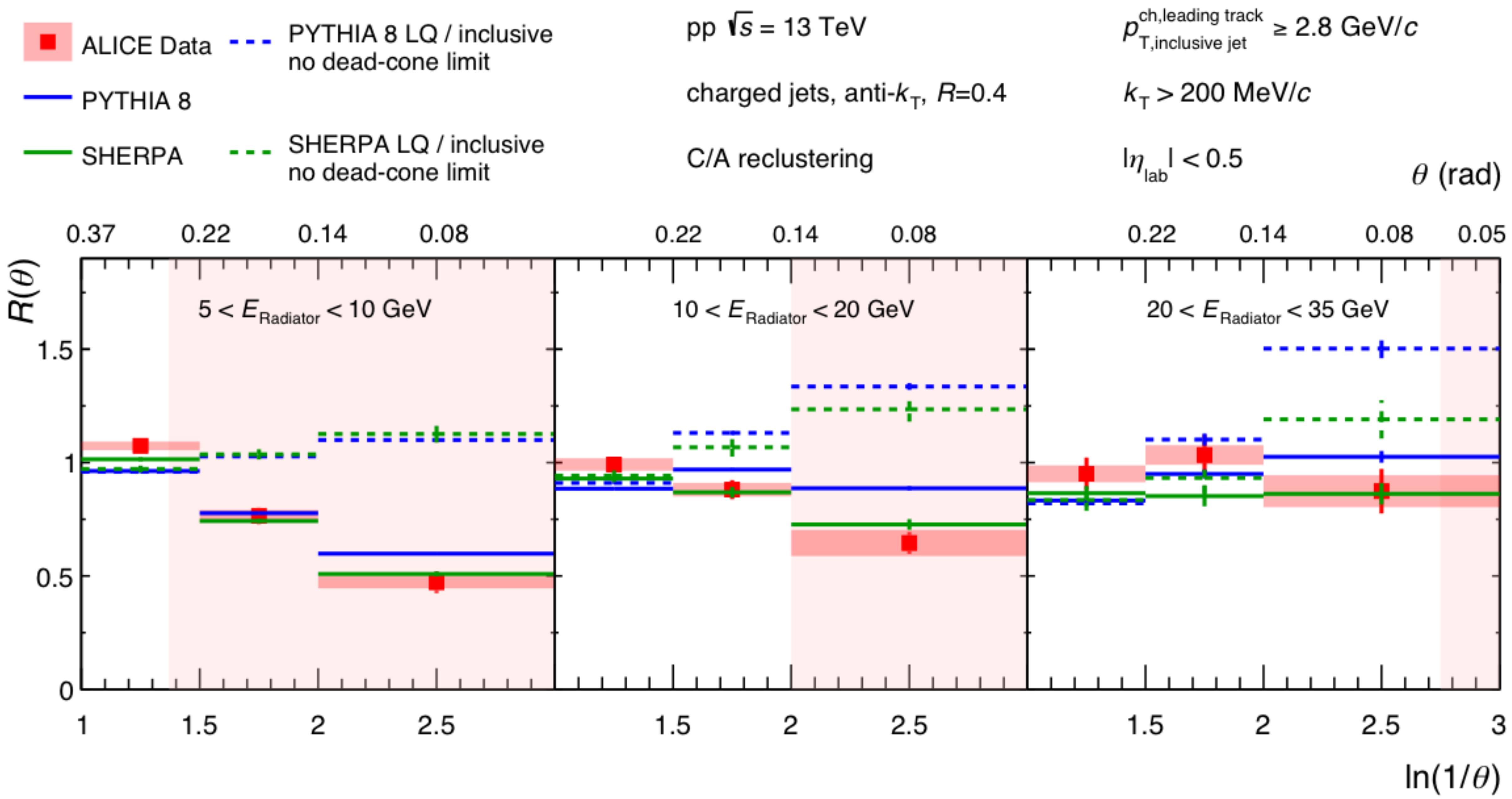


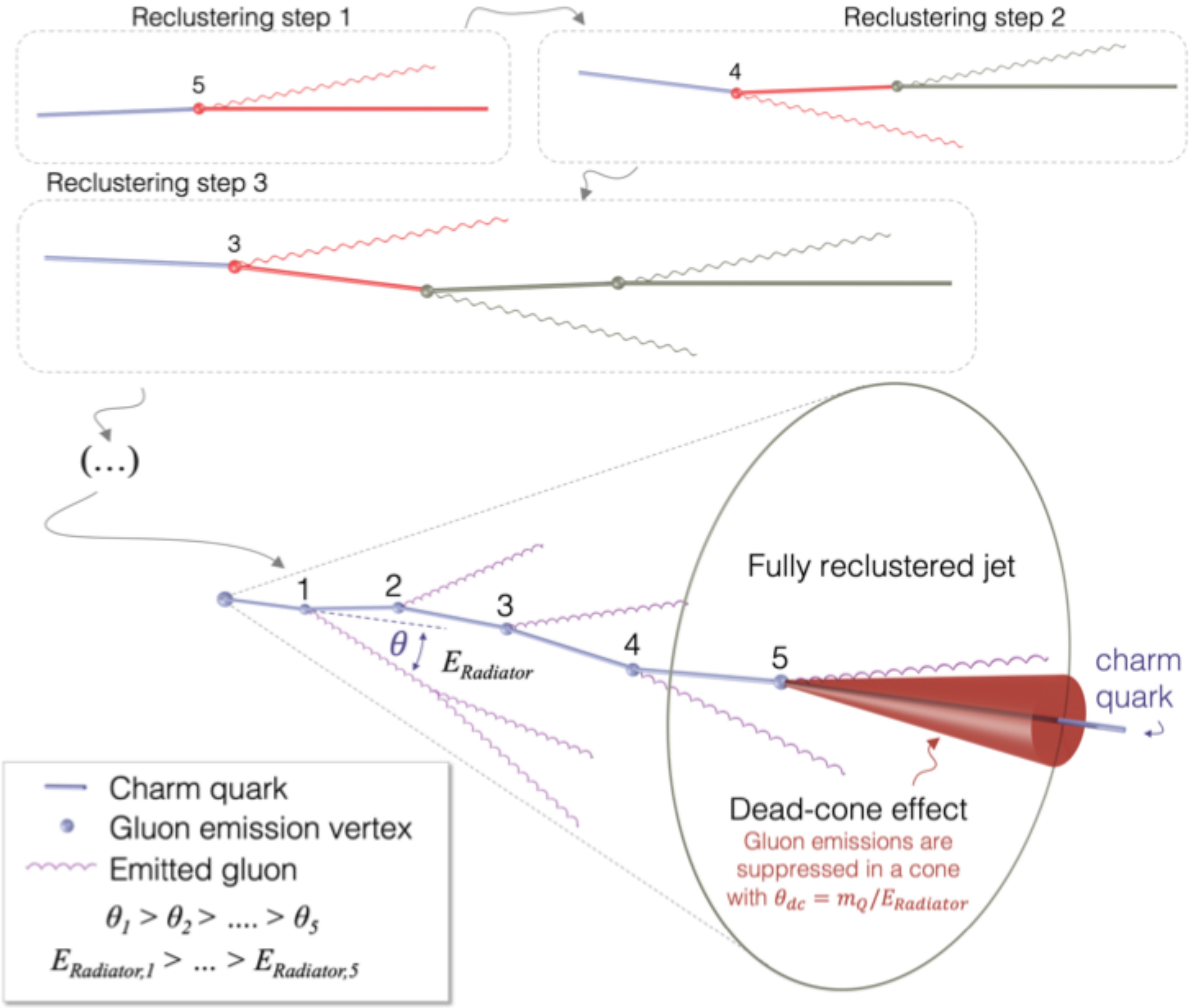
- Process:
  - Break jet j into two subjets by undoing last stage of C/A clustering -> j_1 and j_2
  - If subjets pass SD condition, deem j the final soft-drop jet
  - Otherwise, redefine j = subjet with the larger p_T and iterate the procedure
  - If j is a singleton, either remove j from consideration ("tagging mode") or leave j as final soft-drop jet ("grooming mode")

Larkoski, Marzani, Soyez, Thaler arXiv:1402.2657

Iterative Declustering to See the Dead Cone

- Use anti-k_t to get the jets and find D^0-tagged jets
- Reconstruct (recluster) internal cascade process using C/A
- Then iteratively decluster the reclustering history to access building blocks of jet shower





https://www.nature.com/articles/s41586-022-04572-w

More in Emma's presentation on Nov 1!

Summary

- There are many jet algorithms out there! They mainly fall into two categories - **iterative cone** and **sequential recombination**.
- **IRC Safety** is an important feature of jets to bridge the gap between theorists and experimentalists. **SISCone** and the **sequential recombination algorithms** meet this criteria.
- Different algorithms have different advantages - we like to use **anti-k_t to resolve jets**, and **Cambridge/Aachen to look at jet substructure observables**.

References

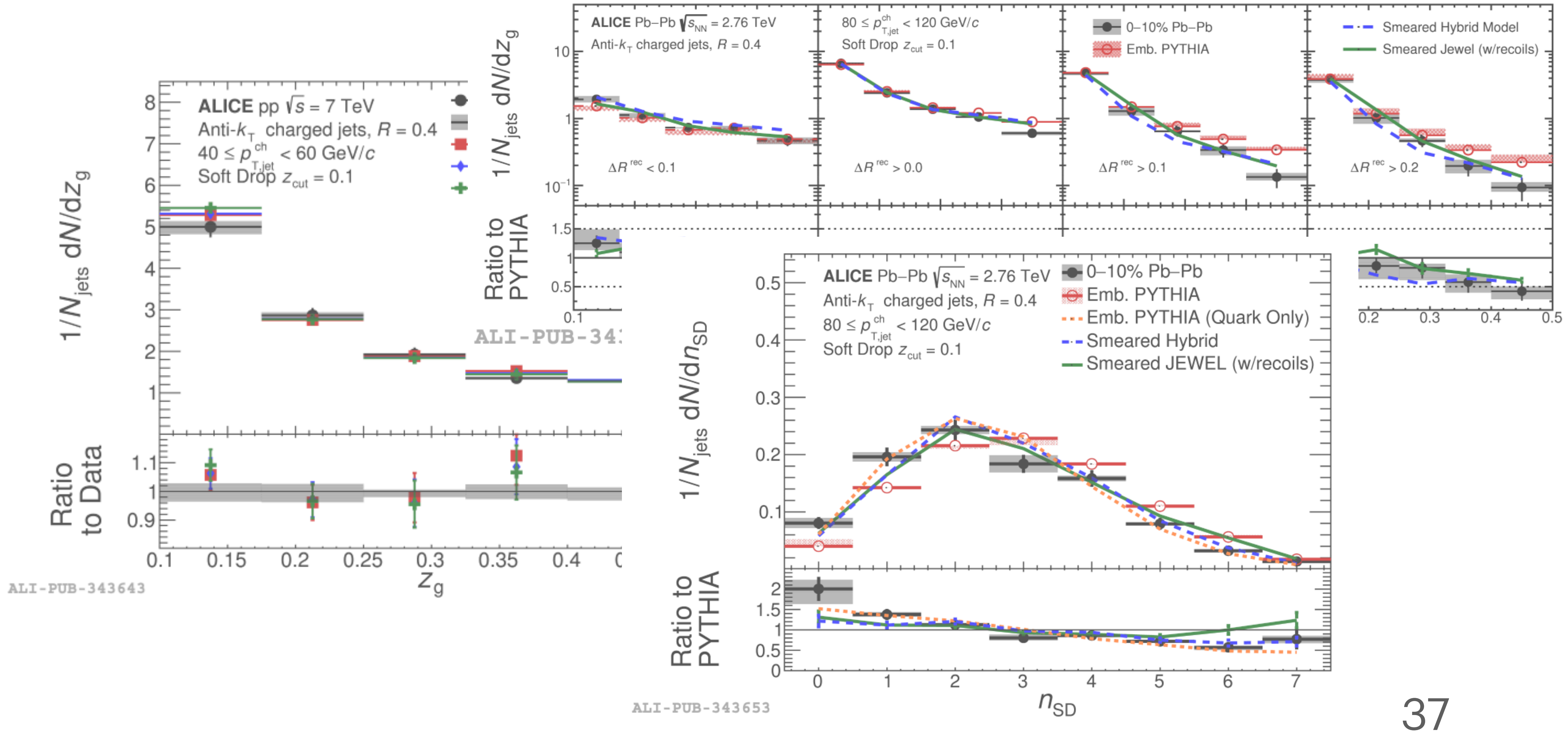
- Towards Jetography: https://arxiv.org/pdf/0906.1833.pdf
- Cacciari talk, HardProbes 2016 : https://indico.cern.ch/event/502239/contributions/2279188/attachments/1341921/2021242/jets.pdf
- Elayavalli talk, QM2023 : https://indico.cern.ch/event/1139644/contributions/5355031/attachments/2707618/4700891/student_day_intro_jet_raghav_3Sept2023.pdf
- Schieferdecker, 2009: https://twiki.cern.ch/twiki/bin/viewfile/Sandbox/Lecture?rev=1;filename=Philipp_Schieferdeckers_Lecture.pdf
- Cacciarai 2018: https://indico.ihep.ac.cn/event/7822/contributions/98092/attachments/52114/60078/lecture1and2.pdf
- Jets and Jet Algorithms: https://inspirehep.net/files/dee666238cd7b18138b1380a421a7474
- Looking Inside Jets: an introduction to jet substructure and boosted-object phenomenology: https://arxiv.org/pdf/1901.10342.pdf
- Review of jet reconstruction algorithms: https://iopscience.iop.org/article/10.1088/1742-6596/645/1/012008/pdf

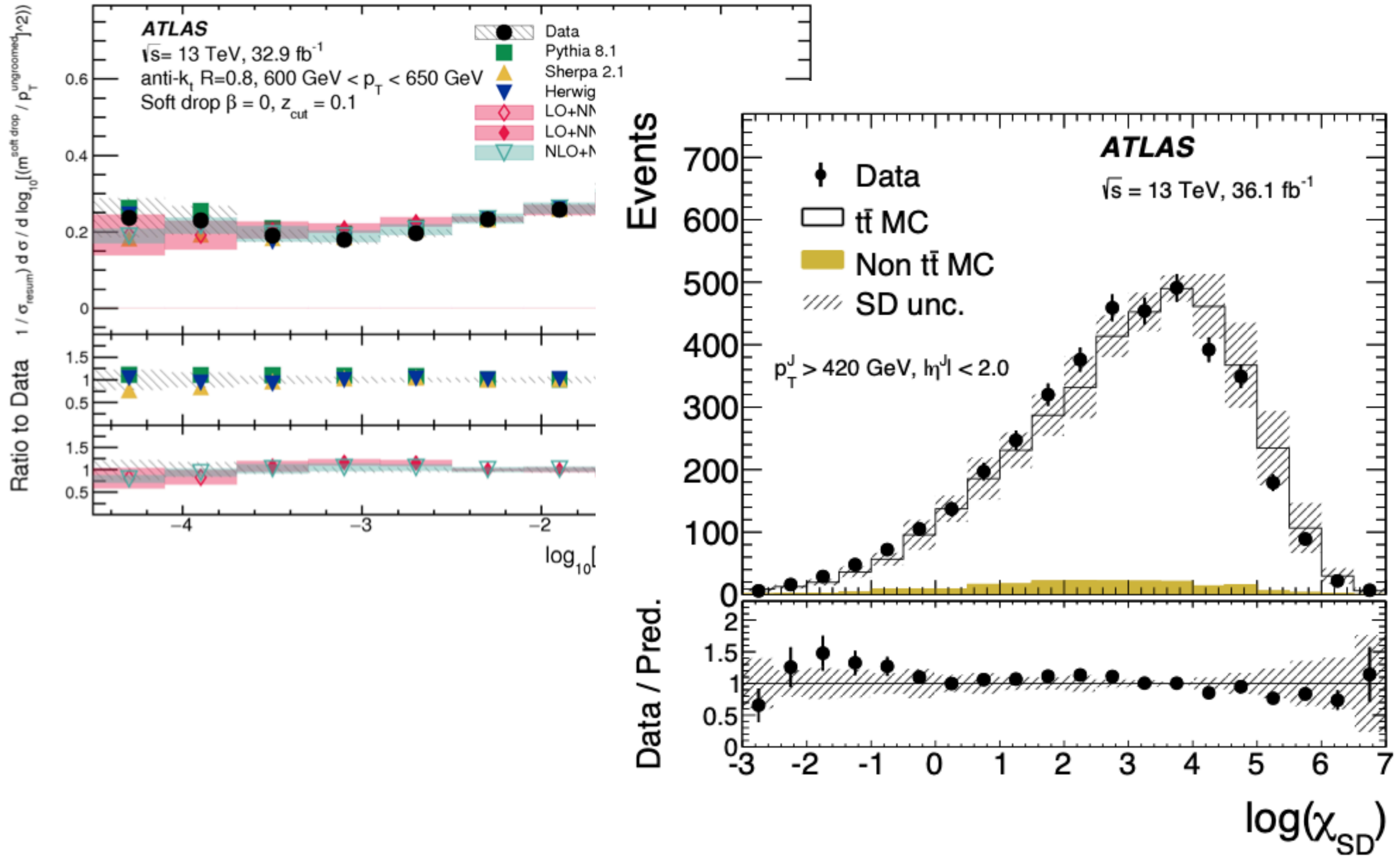
Backup

Anti-k_t

Examples

- Most analyses at LHC today use anti-k_t
- Recombine particles around the hardest branch instead of recombining individual subjets first -> ruins QCD branching history





https://cds.cern.ch/record/2799973/files/10.1051_epjconf_202023501002.pdf

https://arxiv.org/pdf/1901.10342.pdf

Choosing R

$$R = \sqrt{2}\left(\frac{2C_i \mathscr{M}\mathscr{A}(\mu_I)}{\pi \Lambda_{\mathrm{UE}}}\right)^{1/3}$$

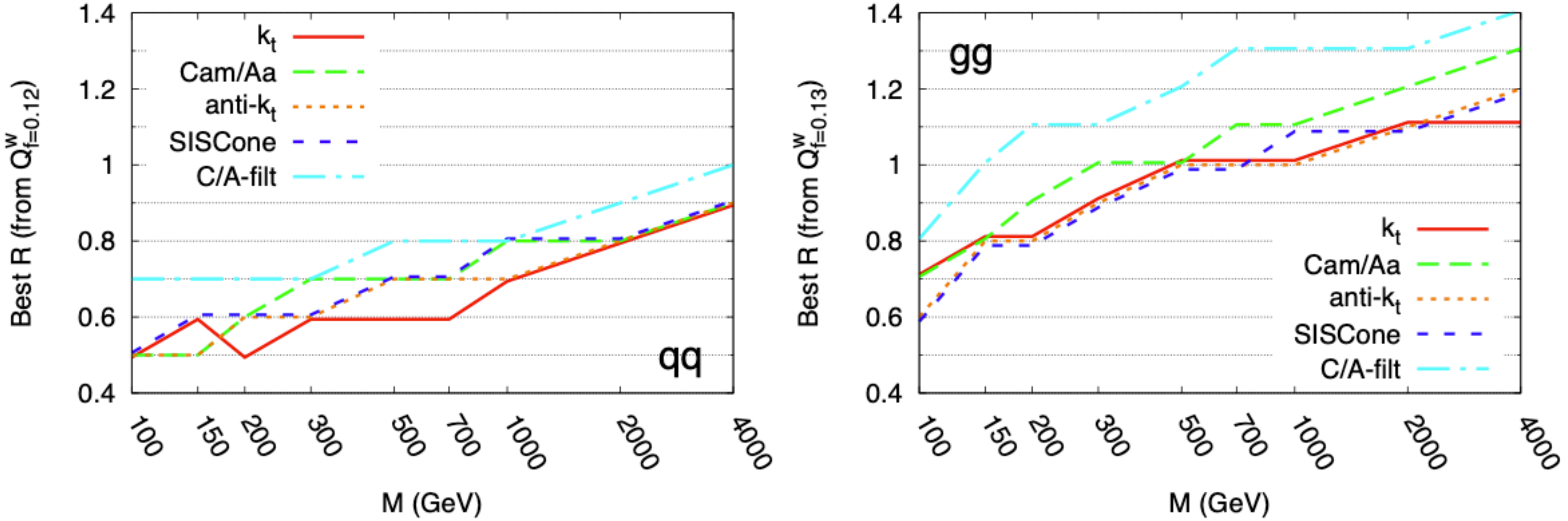


Figure 23: The optimal value for R as a function of the mass of the $q\bar{q}/gg$ system (left/right), as determined from the $Q^w_{f=z}$ quality measure for various jet algorithms. Note that the exact results for the optimal R depend a little on the choice of quality measure, however the observed trends do not.

Variable-R Jet Algorithm

- Krohn, Thaler, and Wang (KRT) analysis
- R scales as $1/p_{t,\text{jet}}$—> $R \sim \cosh \frac{\Delta y}{2}$

Algorithm	500 GeV	1 TeV	2 TeV	3 TeV
anti-k_t → anti-k_t VR	18% (0.9, 200)	14% (1.0, 450)	10% (1.2, 1000)	8% (1.3, 1500)
C/A → C/A VR	17% (0.9, 175)	14% (1.0, 400)	7% (1.2, 1000)	9% (1.3, 1500)

Table 7: Percentage improvement in the number of events from a resonance X that have been reconstructed in the mass window $m_X \pm 25$ GeV, comparing a fixed-R algorithm at its best R (first number in brackets) with the variable-R algorithm (the second number in brackets, ρ/GeV, sets the jet radius as $R(p_t) = \rho/p_t$). Results taken from [84].

Older Sequential Recombination Algorithms

- Jade algorithm (exclusive k_t)
- e+e- k_t (Durham) algorithm
- k_t algorithm with incoming hadrons
- Longitudinally invariant k_t (HERA)

Jets in Heavy Ions

- C/A shows expected features
  - Slow enhancement of radiation with increasing k_T
  - Features at large angles are from underlying event

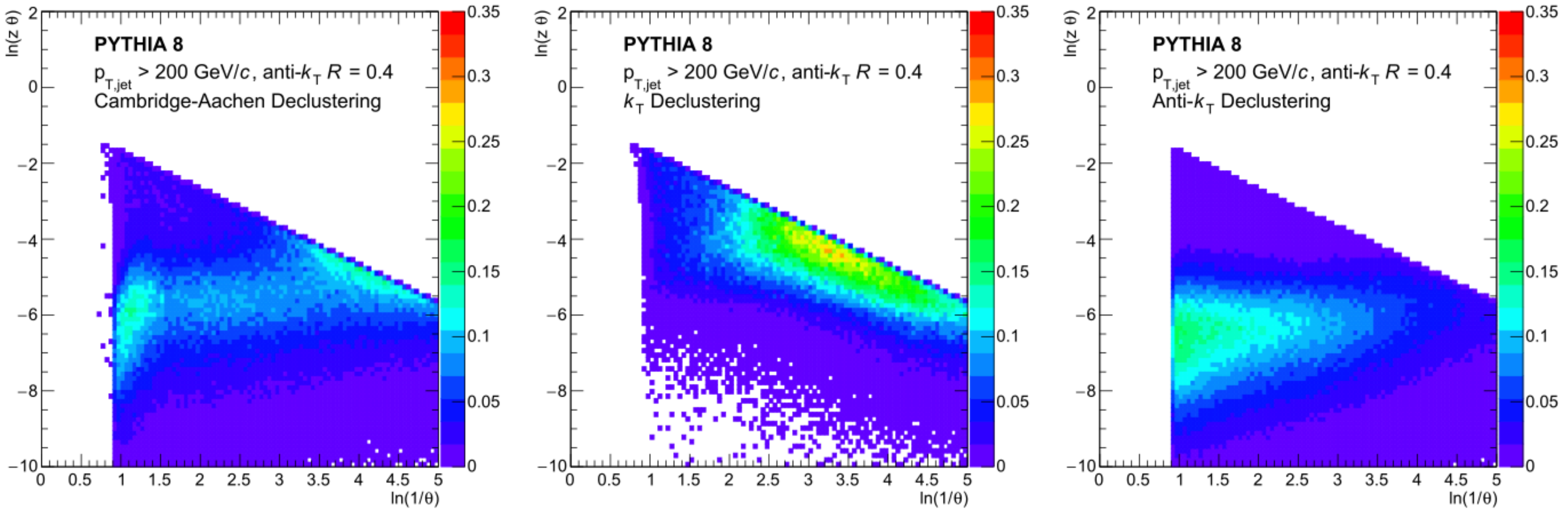
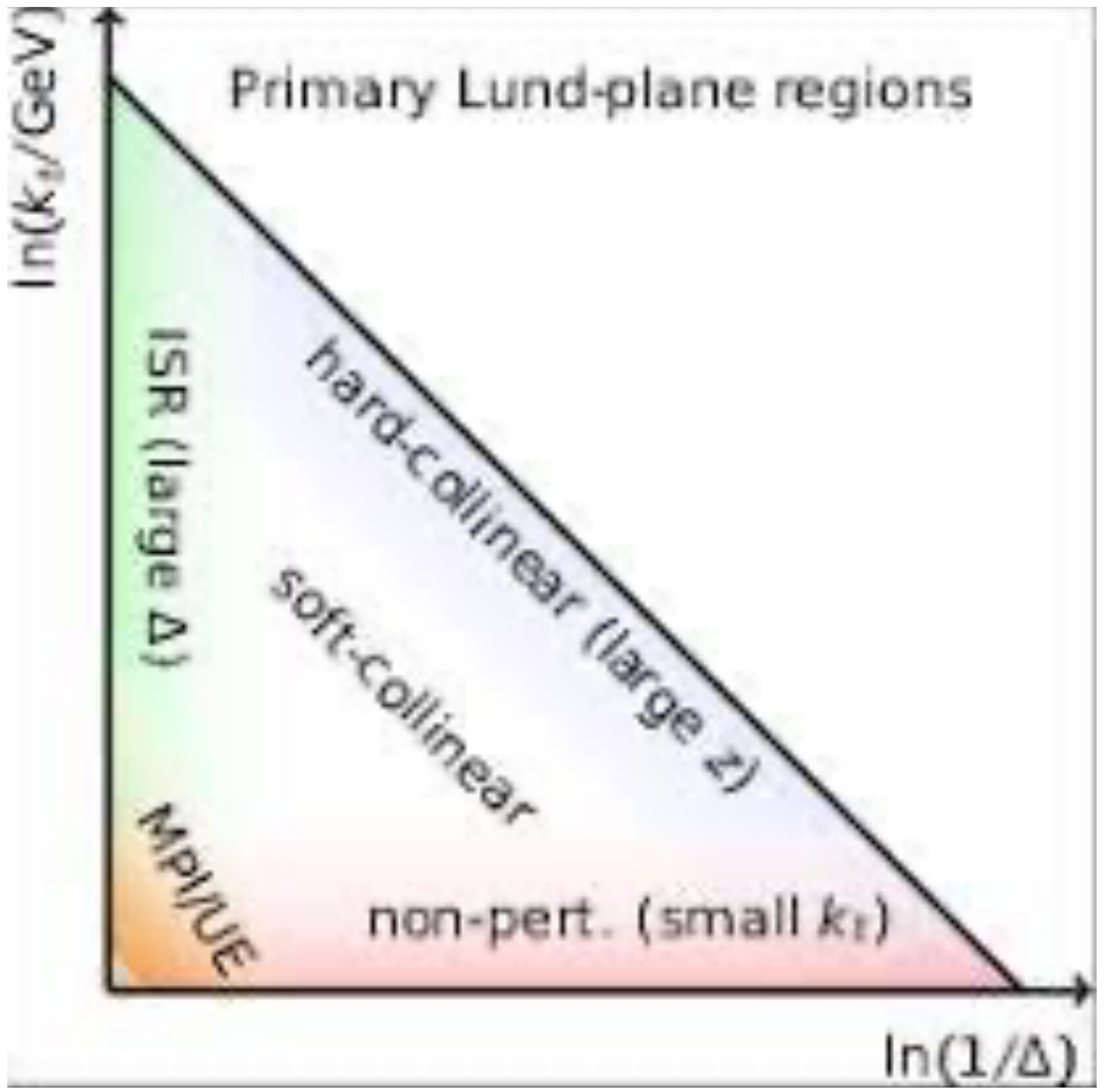


Figure 3. Lund diagrams reconstructed from a sample anti-k_T $R = 0.4$ jets generated by PYTHIA8. Three reclustering strategies were considered: C/A (left), k_T(middle), and anti-k_T (right).



https://iopscience.iop.org/article/10.1088/1361-6471/ab7cbc/pdf