Jets and their structure

Simone Marzani
Università di Genova &
INFN Sezione di Genova

Sangam@HRI 2017

my apologies

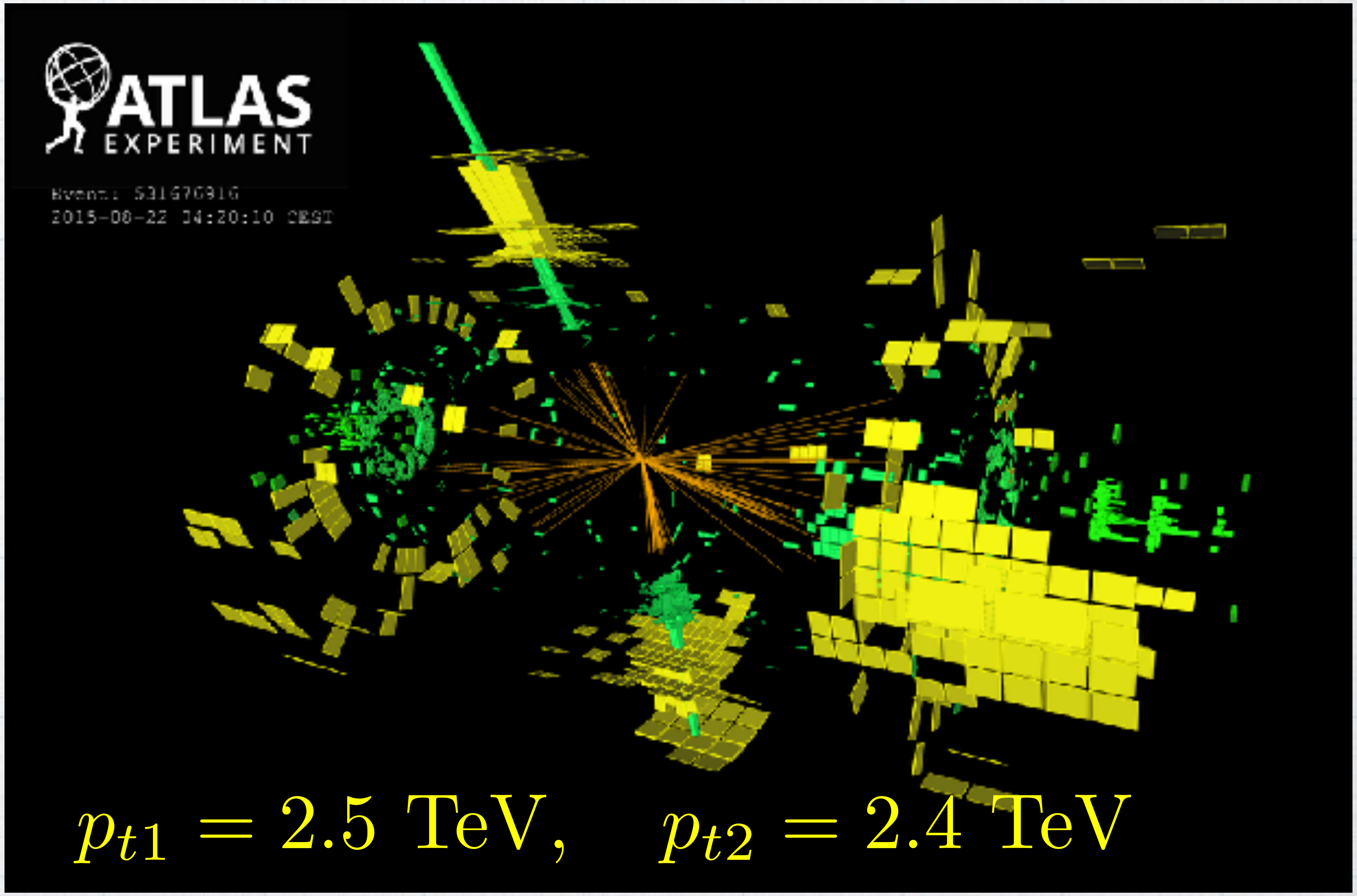
* I'm very sorry for not being able to be with you
* I was looking forward to visiting HRI and interacting with you
* lectures via skype not ideal, but they hopefully work
* feel free to e-mail me questions/doubts at simone.marzani@ge.infn.it

Lecture plan

* lecture 1: jets and jet algorithms
* lecture 2: calculating jet properties
* lecture 3: jet substructure

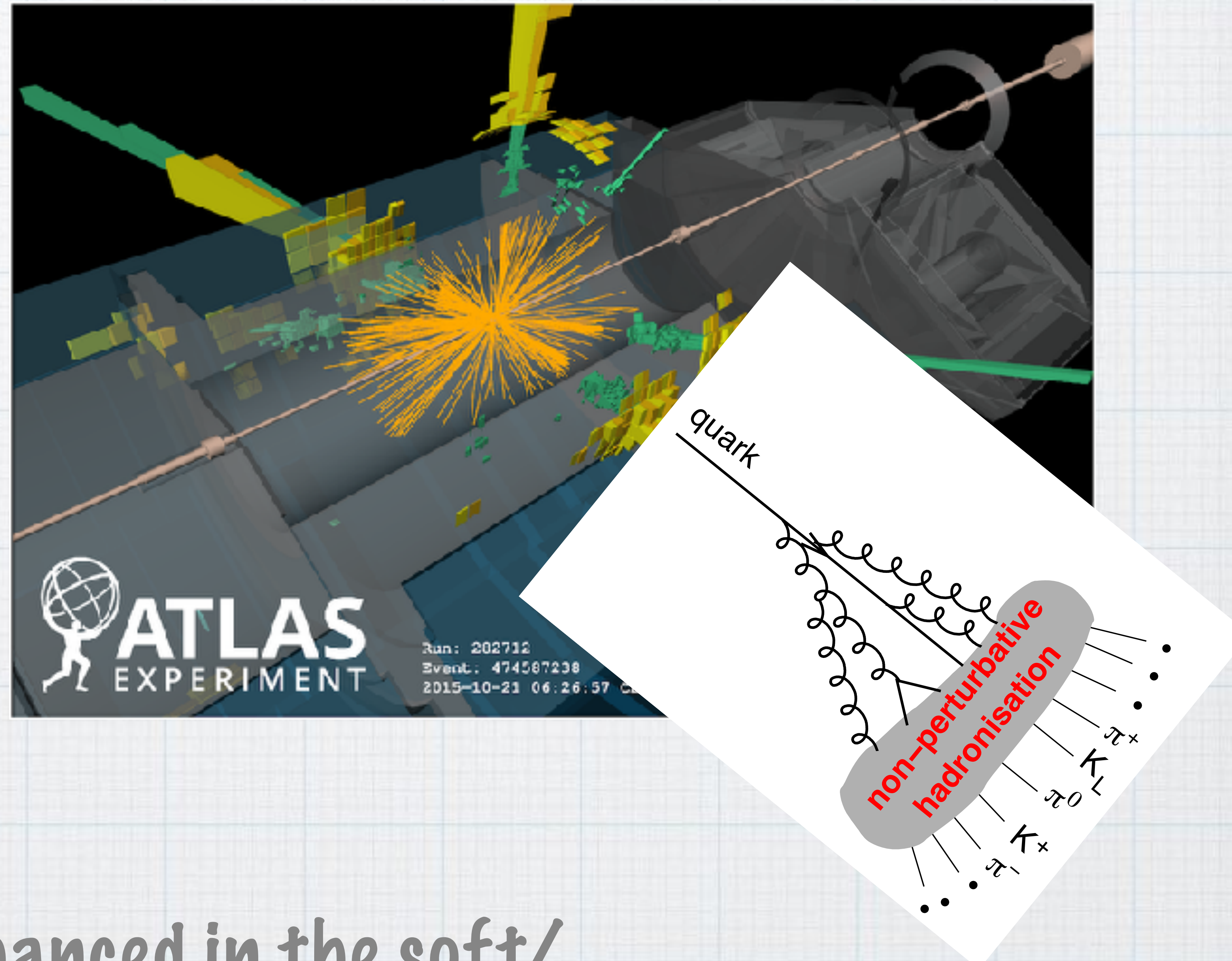
Lecture 1: Jets and jet algorithms

* jet definition(s)
* IRC safety
* sequential recombination
* pile-up



jets for experimentalists

- high-energy collisions offer results into collimated sprays of particles



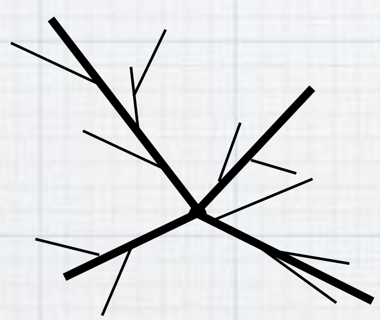
- why?

gluon emission enhanced in the soft/collinear limit

$$\int \frac{dE}{E}\frac{d\theta}{\theta}\alpha_s \gg 1$$

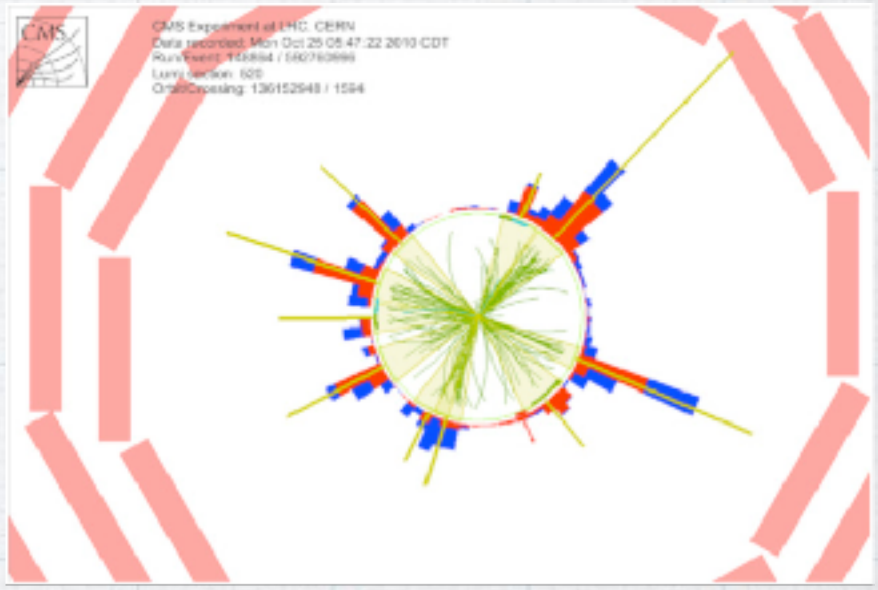
jets for theorists

* jets are extremely useful for theorists
* powerful way of turning calculations into predictions



theory-land:
(quarks & gluons)

??



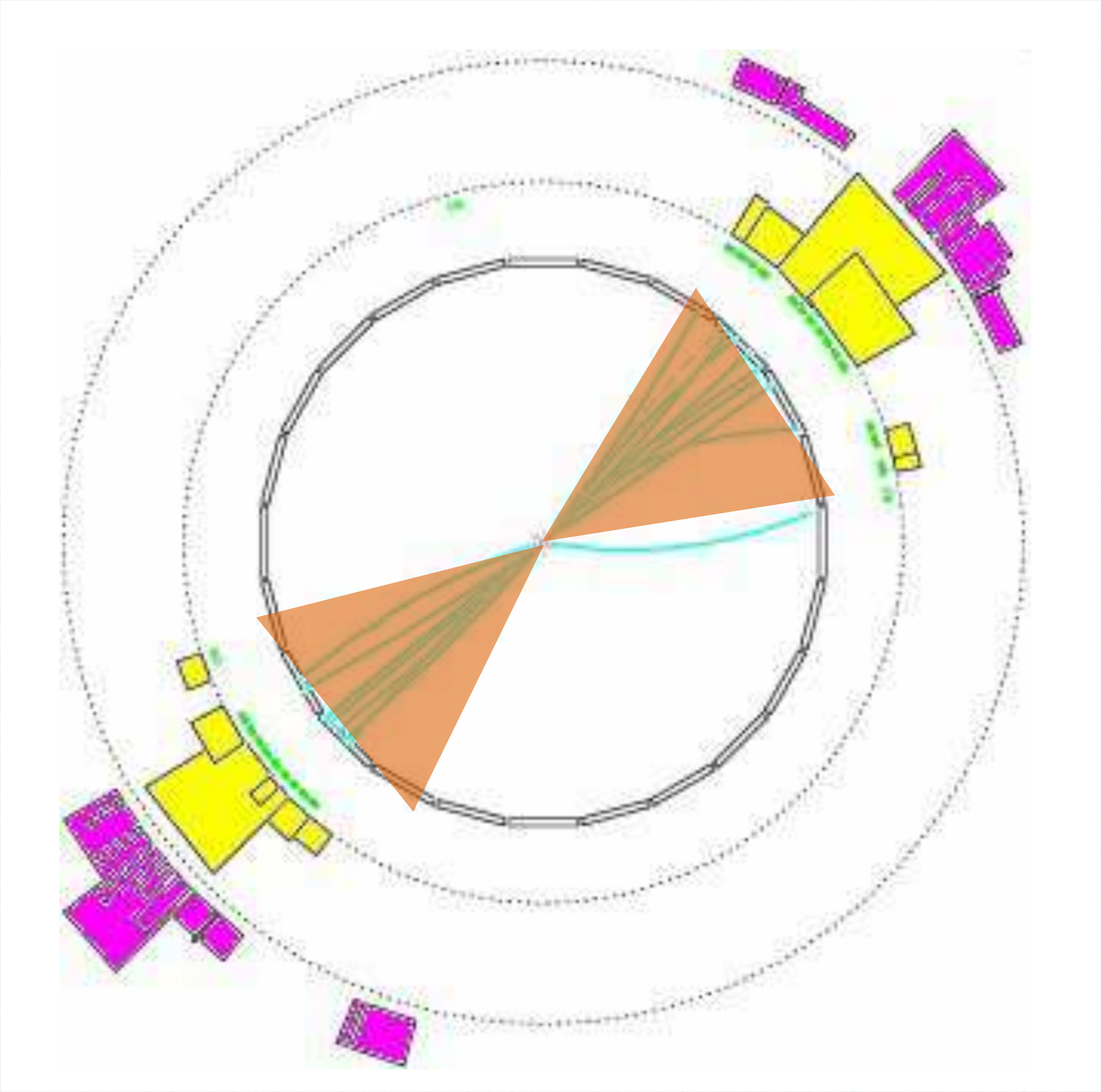
real data
(π,K,p,e,μ)

jets

thanks to jets we can
reduce the complexity of the final state, simplifying many hadrons to simpler objects that one can hope to calculate

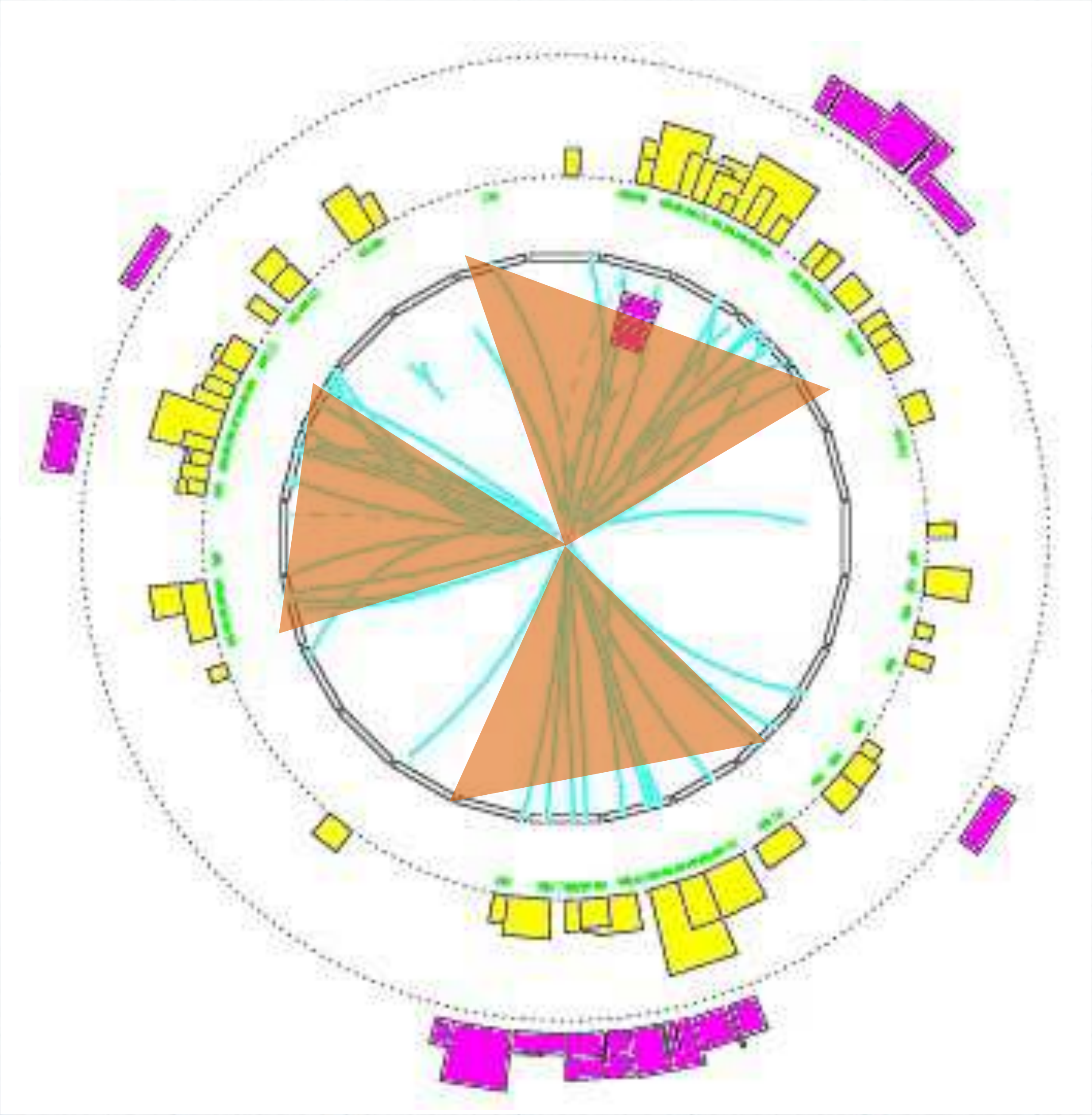
what is a jet?

* how many jets do you see?
* two is probably a good guess
* eyeballing not good enough!



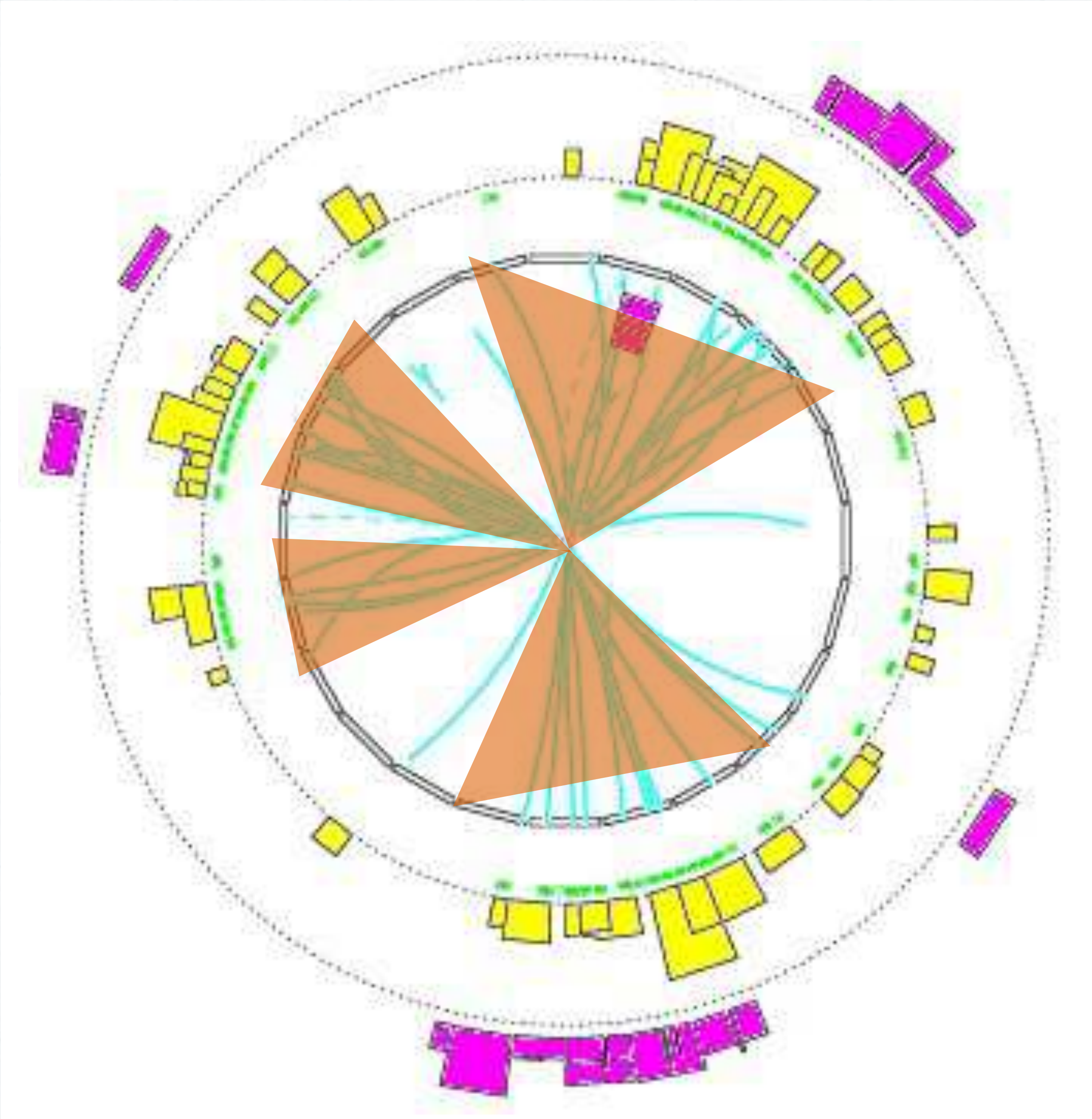
what is a jet?

* what about now ?
* messy events are more ambiguous
* 3 jet event?



what is a jet?

* what about now ?
* messy events are more ambiguous
* or 4 jet event?



* we need a way to define jets in a given event

jet definition

a jet algorithm
+
its parameters (e.g. R)
+
a recombination scheme
=
a jet definition

* examples of recombination schemes:
    * E-scheme: sum all the four momenta
    * winner-take-all

jet clustering algorithm

* an algorithm that maps the momenta of the final state particles into the momenta of a certain number of jets

$$\{p_i\} \xrightarrow{\text{jet algorithm}} \{j_k\}$$

particles,
4-momenta,
calorimeter towers, ...

jet algorithm

often comes with resolution parameters e.g. R

jets

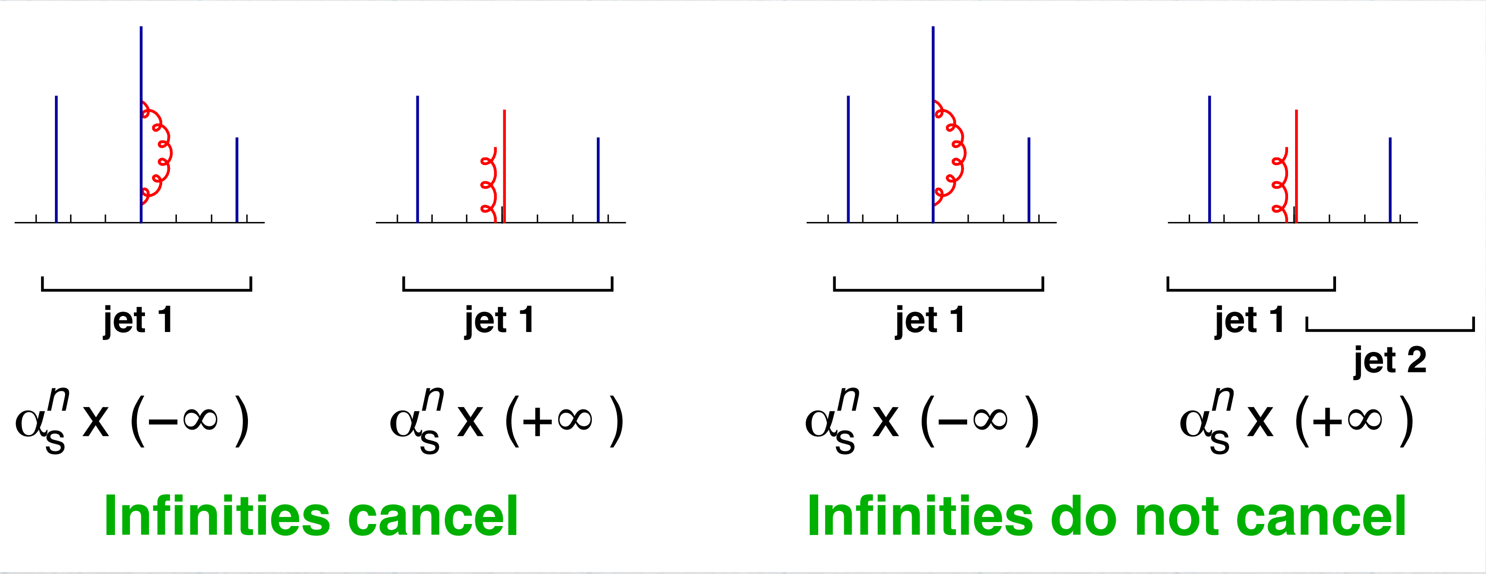
* jet definitions must make sense for both theorists and experimentalists!

what do theorists want?

- Infra-Red and Collinear Safety!
- An observable is **IRC safe** if, in the limit of a collinear splitting, or the emission of an infinitely soft particle, the observable remains **unchanged:**

$$O(X; p_1, \ldots, p_n, p_{n+1} \to 0) \to O(X; p_1, \ldots, p_n)$$

$$O(X; p_1, \ldots, p_n \parallel p_{n+1}) \to O(X; p_1, \ldots, p_n + p_{n+1})$$



we need IRC safety if we want to compute things beyond LO!

homework 1

* which of the following observables are IRC safe?
  * the p_T distribution of the highest p_T jet
  * the p_T distribution of the highest p_T jet in Z+jet events
  * the jet invariant mass
  * the invariant mass of tracks in a jet

what do experimentalists want?

* jet algorithms must be usable on real events
* fast and easy to calibrate
* a thousand particles in each event
* CMS high-level trigger output rate 50kHz

types of algorithms

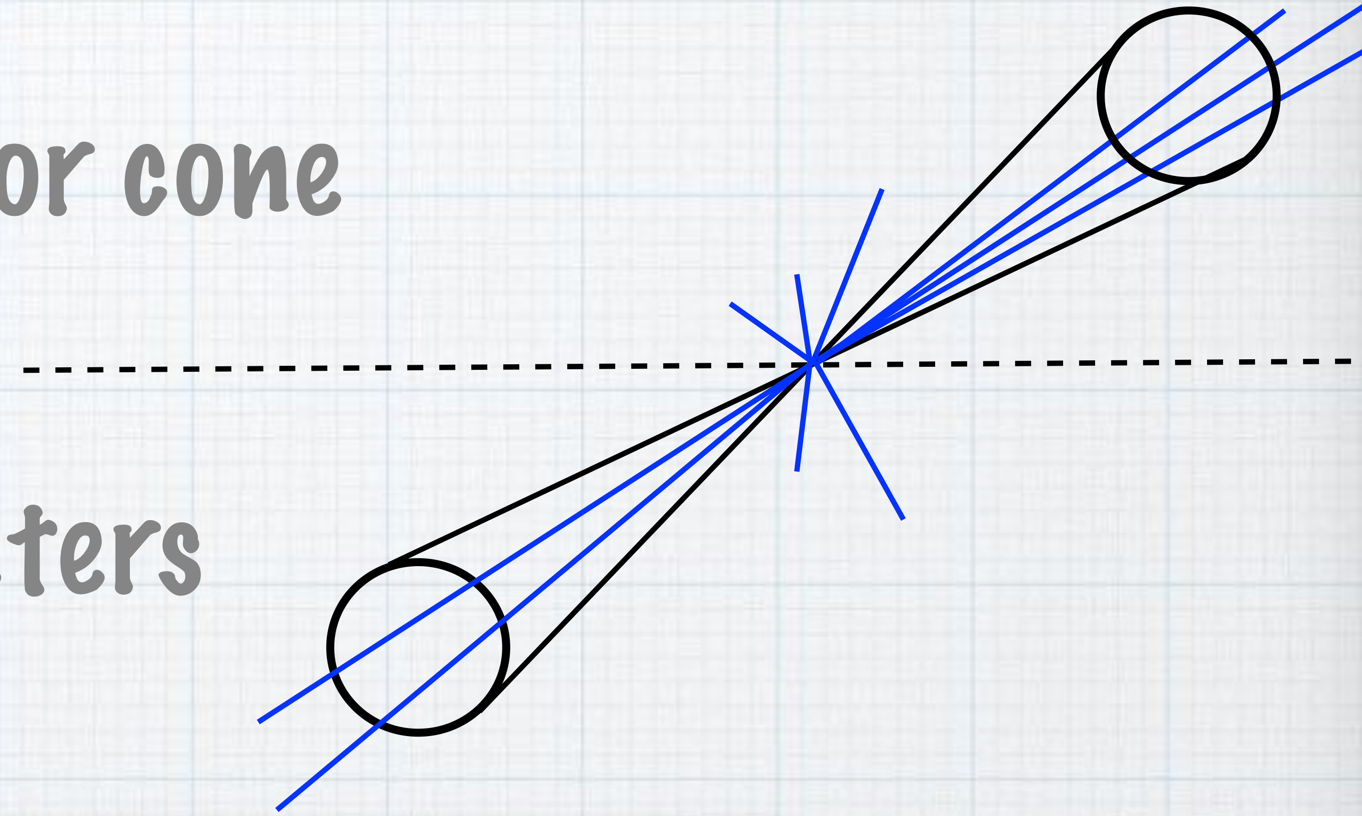
* sequential recombination algorithms
  * bottom-up approach: combine particles starting from closest ones
  * how? Choose a distance measure, iterate recombination until few objects left, call them jets
  * usually trivially made IRC safe, but their algorithmically complex (unless you're clever)
  * Examples: Jade, k_t, Cambridge/Aachen, anti-k_t ...

* cone algorithms
  * top-down approach: find coarse regions of energy flow.
  * How? Find stable cones (i.e. their axis coincides with sum of momenta of particles in it)
  * Can be programmed to be fairly fast, at the price of being complex and IRC unsafe
  * Examples: JetClu, MidPoint, ATLAS cone, CMS cone, SISCone ...

for a complete review see G. Salam, Towards jetography (2009)

a bit of history

* first calculation done for cone algorithm
* two resolution parameters

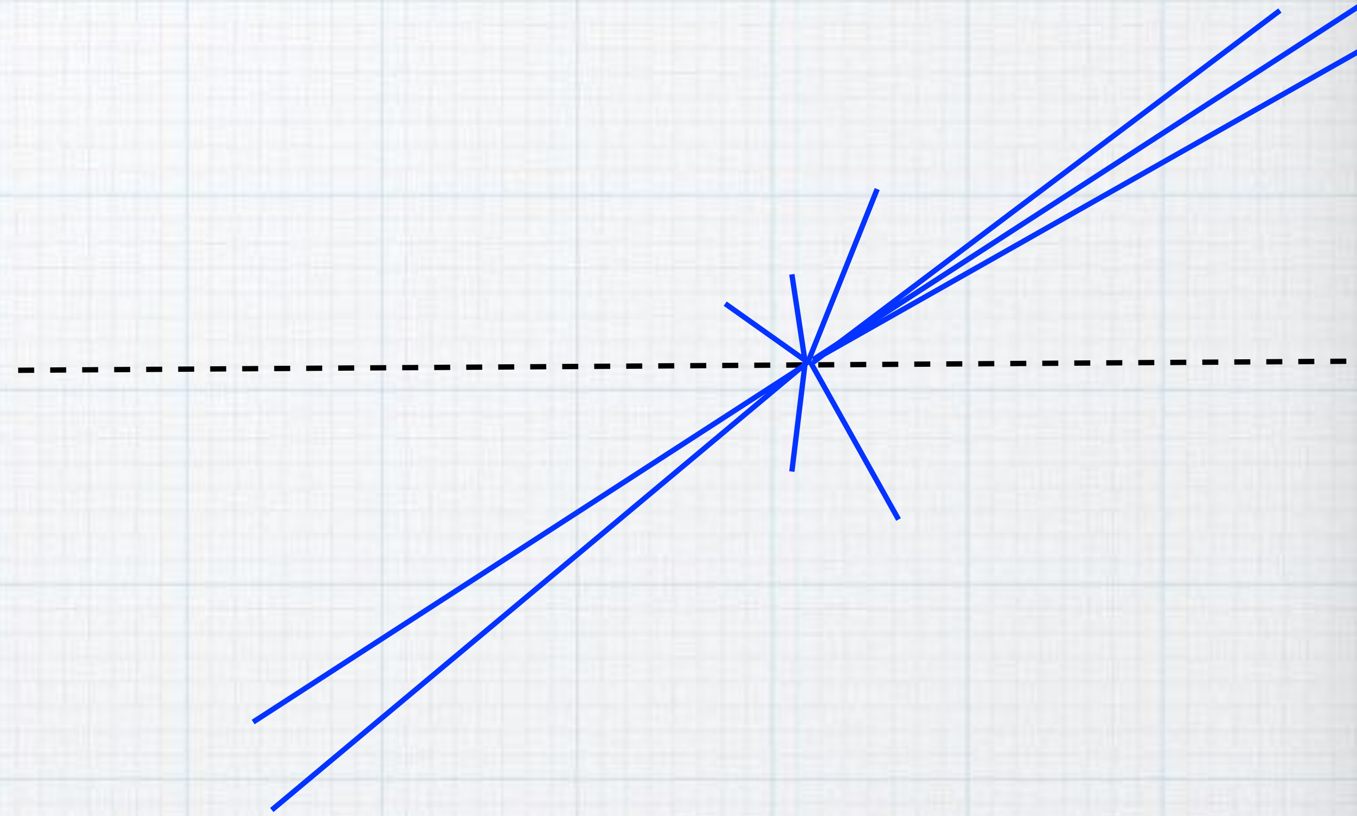


To study jets, we consider the partial cross section $\sigma(E,\theta,\Omega,\varepsilon,\delta)$ for e^+e^- hadron production events, in which all but a fraction $\varepsilon \ll 1$ of the total e^+e^- energy E is emitted within some pair of oppositely directed cones of half-angle $\delta \ll 1$, lying within two fixed cones of solid angle Ω (with $\pi\delta^2 \ll \Omega \ll 1$) at an angle θ to the e^+e^- beam line. We expect this to be measur-

Sterman and Weinberg,
Phys. Rev. Lett. 39, 1436 (1977):

sequential recombination

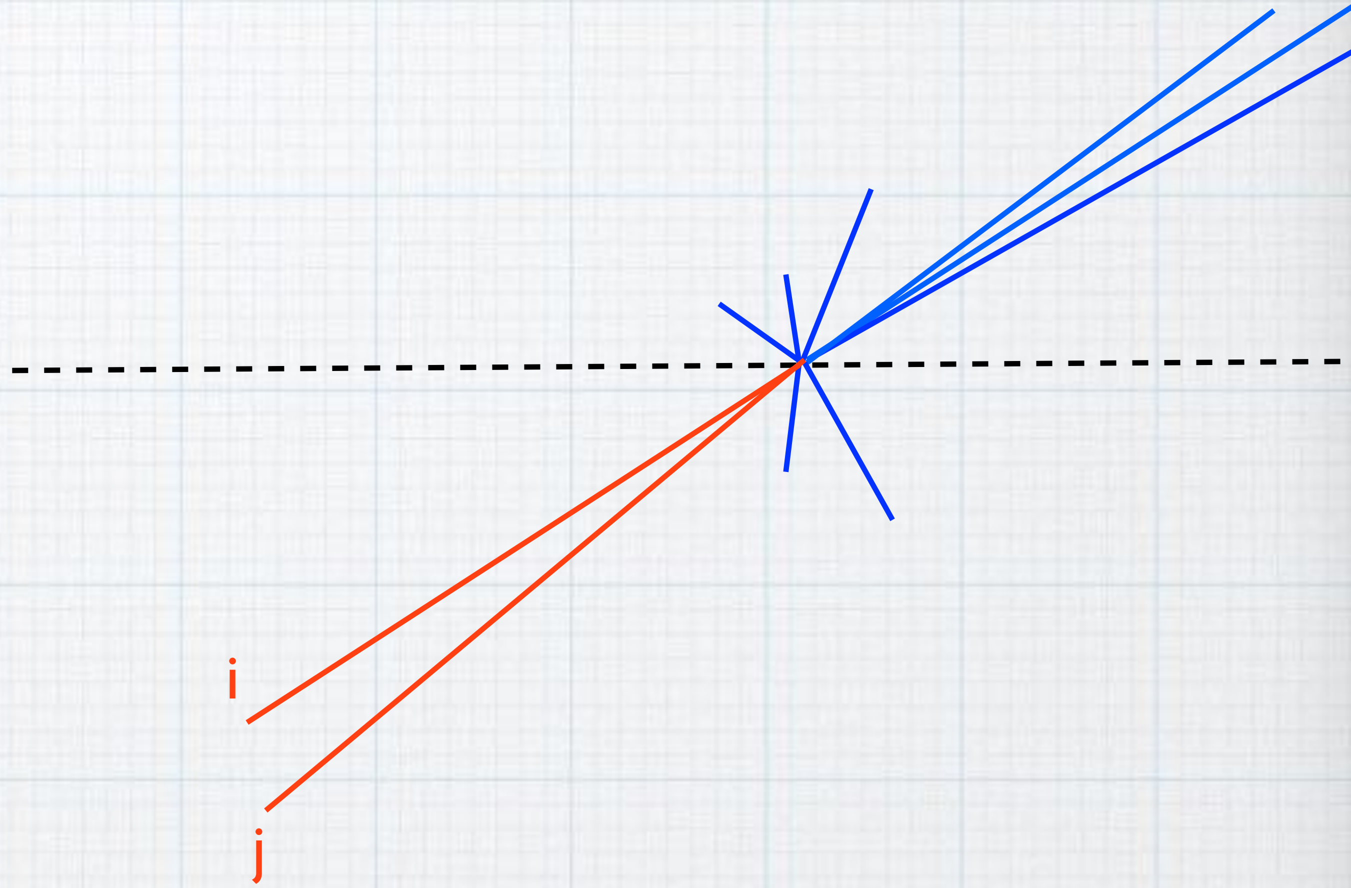
* start with a list of particles,
* compute all distances d_{ij} and d_{iB}
* find the minimum of all d_{ij} and d_{iB}



d_{ij} (weighted) distance between i j
d_{iB} external parameter or distance from the beam ...

sequential recombination

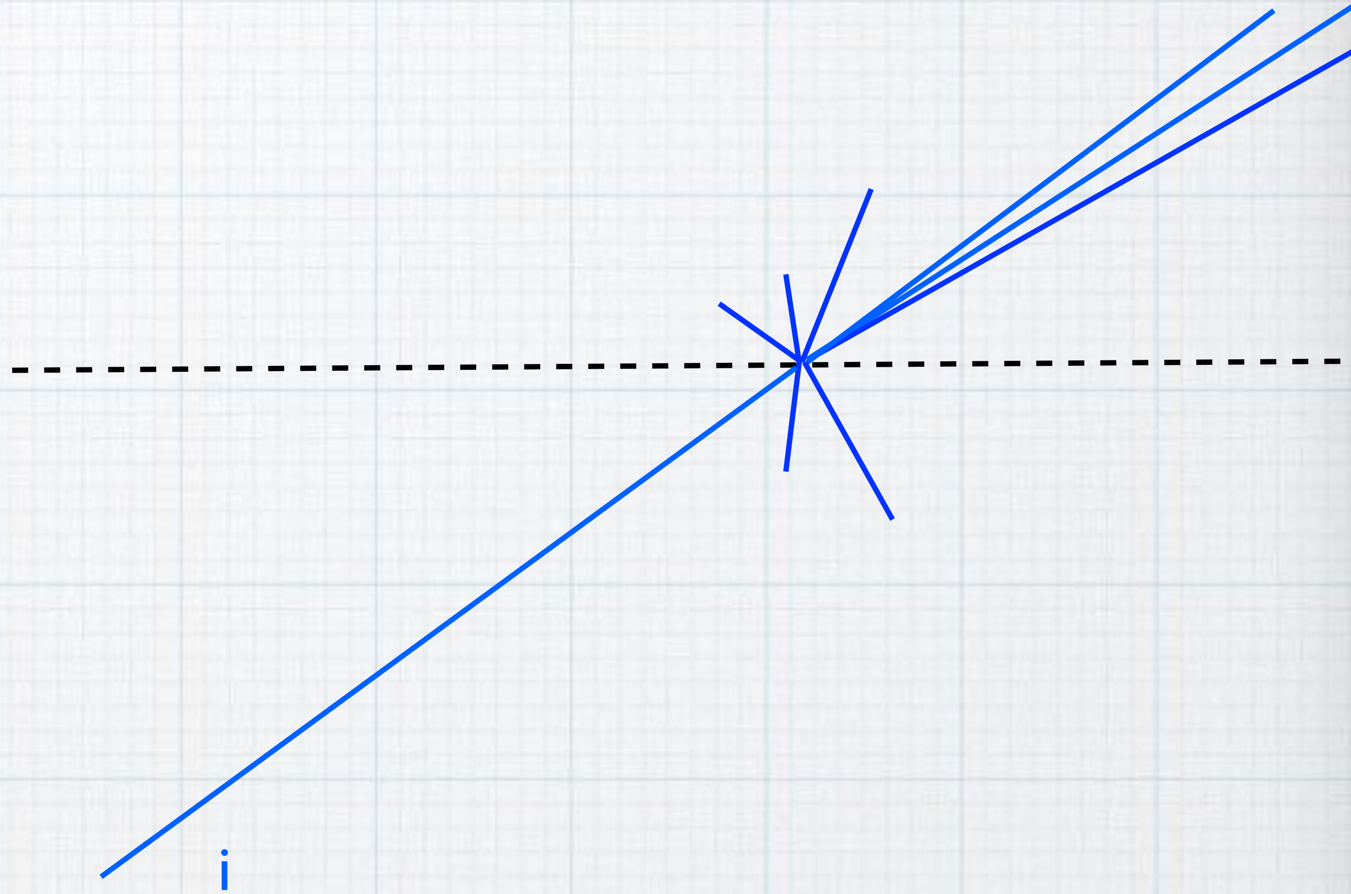
* start with a list of particles,
* compute all distances d_{ij} and d_{iB}
* find the minimum of all d_{ij} and d_{iB}
* if the minimum is a d_{ij}, recombine i and j and iterate



d_{ij} (weighted) distance between i j
d_{iB} external parameter or distance from the beam ...

sequential recombination

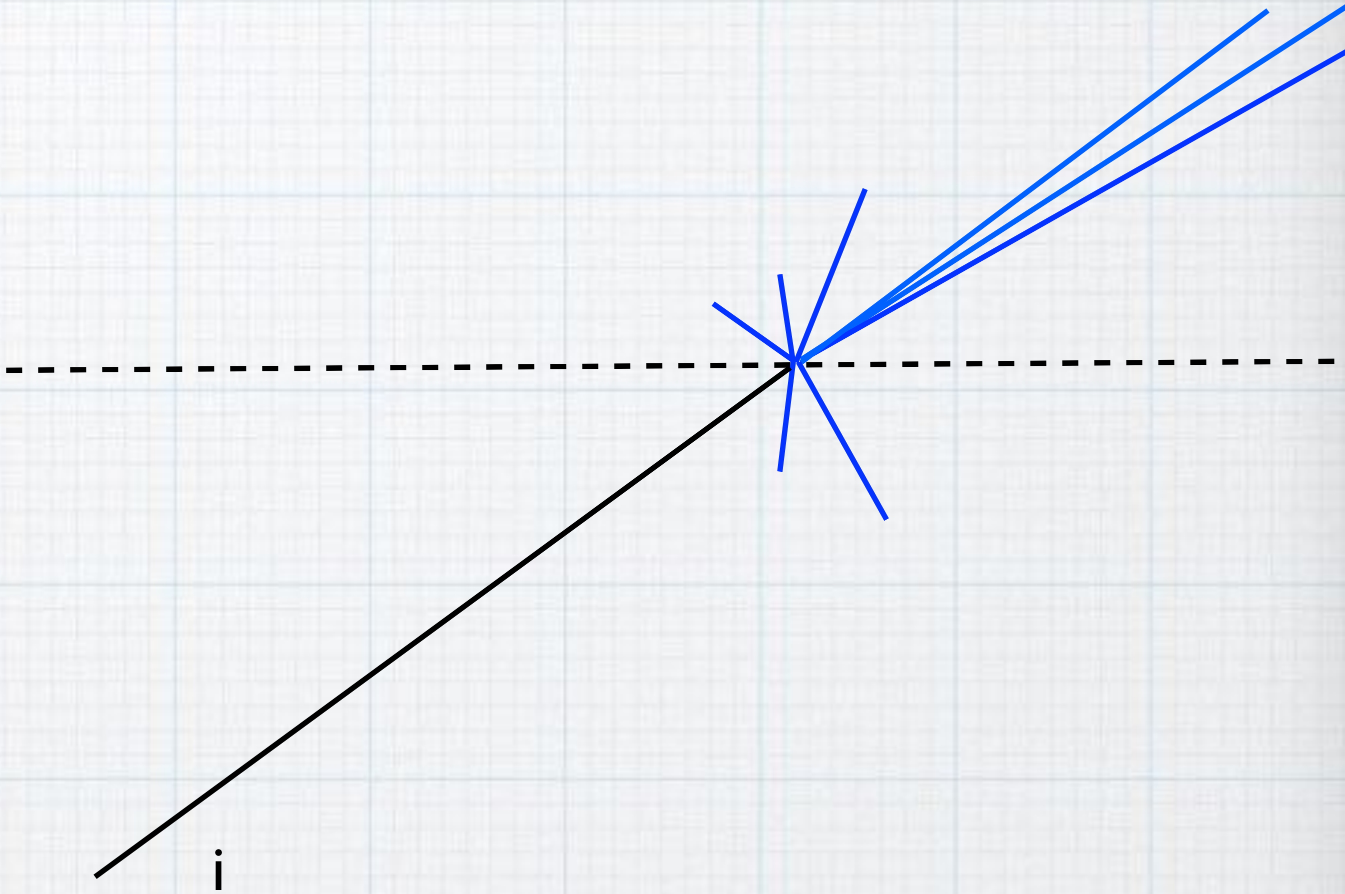
* start with a list of particles,
* compute all distances d_{ij} and d_{iB}
* find the minimum of all d_{ij} and d_{iB}
* if the minimum is a d_{ij}, recombine i and j and iterate



d_{ij} (weighted) distance between i j
d_{iB} external parameter or distance from the beam ...

sequential recombination

* start with a list of particles,
* compute all distances d_{ij} and d_{iB}
* find the minimum of all d_{ij} and d_{iB}
* if the minimum is a d_{ij}, recombine i and j and iterate

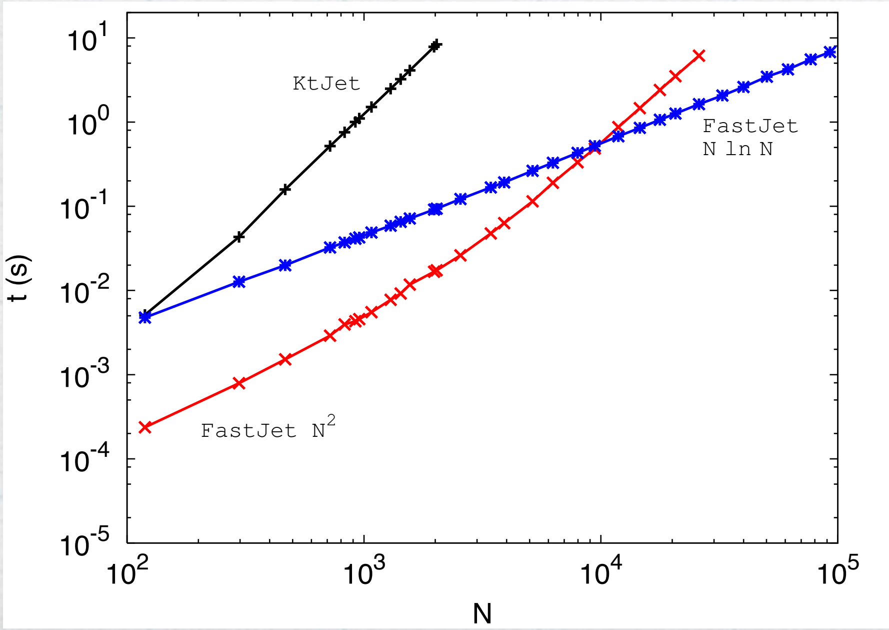


d_{ij} (weighted) distance between i j
d_{iB} external parameter or distance from the beam ...

* otherwise call i a final-state jet, remove it from the list and iterate

speeding-up the algorithms

* from combinatorics sequential recombination should scale like N^3
* an approach based on geometry (Voronoi diagrams) leads to notable improvements
* Sequential recombination algorithms could be implemented with $O(N^2)$ or even $O(N\ln N)$ complexity rather than $O(N^3)$
  Cacciari, Salam, 2006
* Cone algorithms could be implemented exactly (and therefore made IRC safe) with $O(N^2 \ln N)$ rather than $O(N 2^N)$ complexity
  Salam, Soyez, 2007



method implemented in FastJet

the generalised k_t family

* actual choice of d_{ij} determines the algorithm

$$d_{ij} = \min(p_{ti}^{2p}, p_{tj}^{2p}) \frac{\Delta y^2 + \Delta\phi^2}{R^2}$$

IRC behaviour

p = 1 **k_t algorithm**

S. Catani, Y. Dokshitzer, M. Seymour and B. Webber, Nucl. Phys. B406 (1993) 187
S.D. Ellis and D.E. Soper, Phys. Rev. D48 (1993) 3160

new soft particle ($p_t \to 0$) means that $d \to 0$ $\Rightarrow$ clustered first, no effect on jets

new collinear particle ($\Delta y^2 + \Delta\Phi^2 \to 0$) means that $d \to 0$ $\Rightarrow$ clustered first, no effect on jets

p = 0 **Cambridge/Aachen algorithm**

Y. Dokshitzer, G. Leder, S.Moretti and B. Webber, JHEP 08 (1997) 001
M.Wobisch and T.Wengler, hep-ph/9907280

new soft particle ($p_t \to 0$) can be new jet of zero momentum $\Rightarrow$ no effect on hard jets

new collinear particle ($\Delta y^2 + \Delta\Phi^2 \to 0$) means that $d \to 0$ $\Rightarrow$ clustered first, no effect on jets

p = -1 **anti-k_t algorithm**

M. Cacciari, G. Salam and G. Soyez, arXiv:0802.1189

new soft particle ($p_t \to 0$) means $d \to \infty$ $\Rightarrow$ clustered last or new zero-jet, no effect on hard jets

new collinear particle ($\Delta y^2 + \Delta\Phi^2 \to 0$) means that $d \to 0$ $\Rightarrow$ clustered first, no effect on jets

the k_t algorithm

* the k_t distance is the inverse of the QCD splitting probability

$$\frac{dP_{k\to ij}}{dE_i d\theta_{ij}} \sim \frac{\alpha_s}{\min(E_i, E_j)\theta_{ij}}$$

* the algorithm roughly inverts the QCD shower, bringing us back to the hard scattering
* the clustering history has physical meaning
* jets grow around soft particles, which is a problem in a noisy environment as the LHC

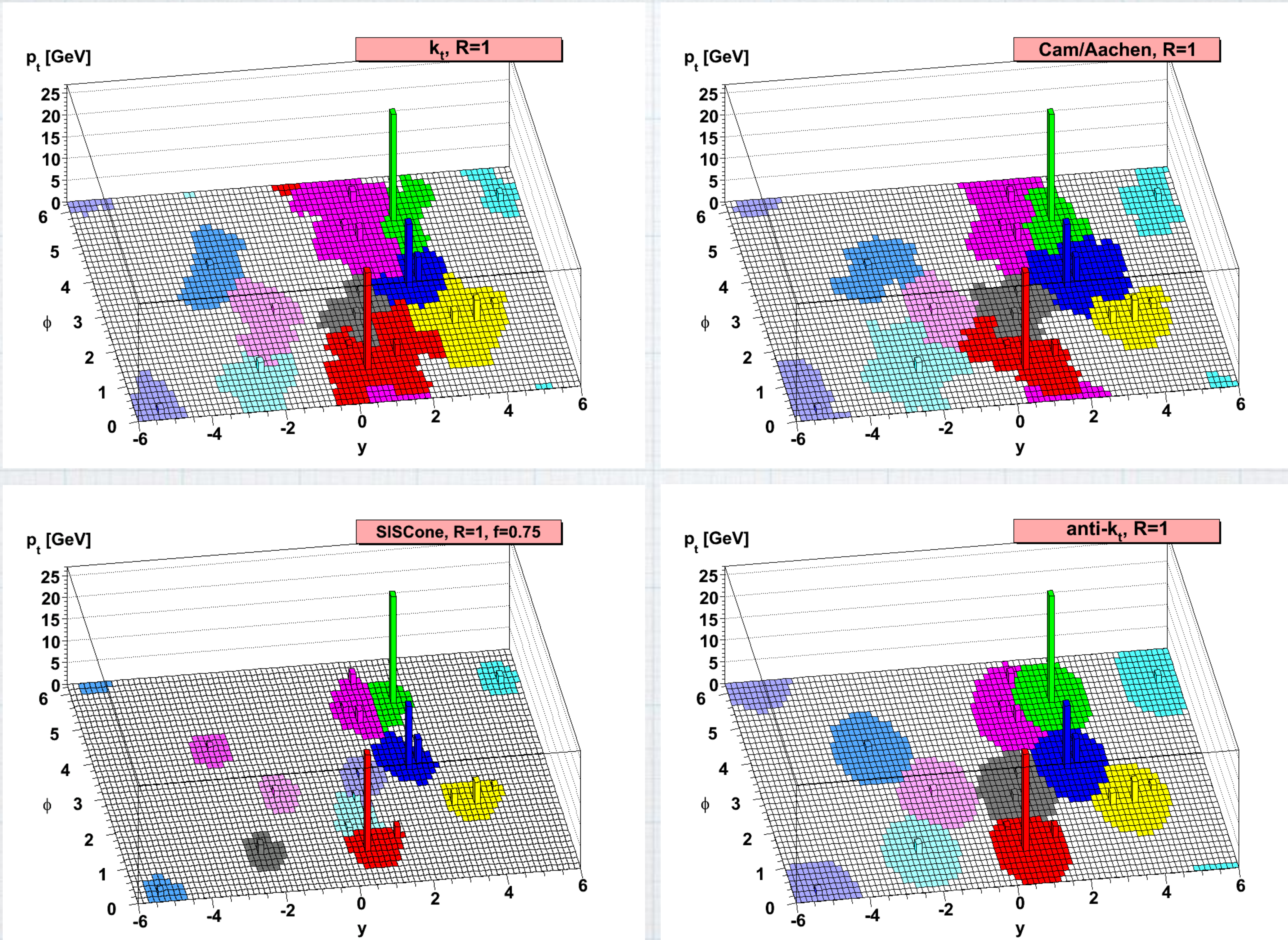
the anti-k_t algorithm

* with this measure soft particles are always far away
* jets grow around hard cores
* if no other hard particles are around the algorithm provides (ironically) perfect cones
* however, the clustering history carries little physics (re-clustering)

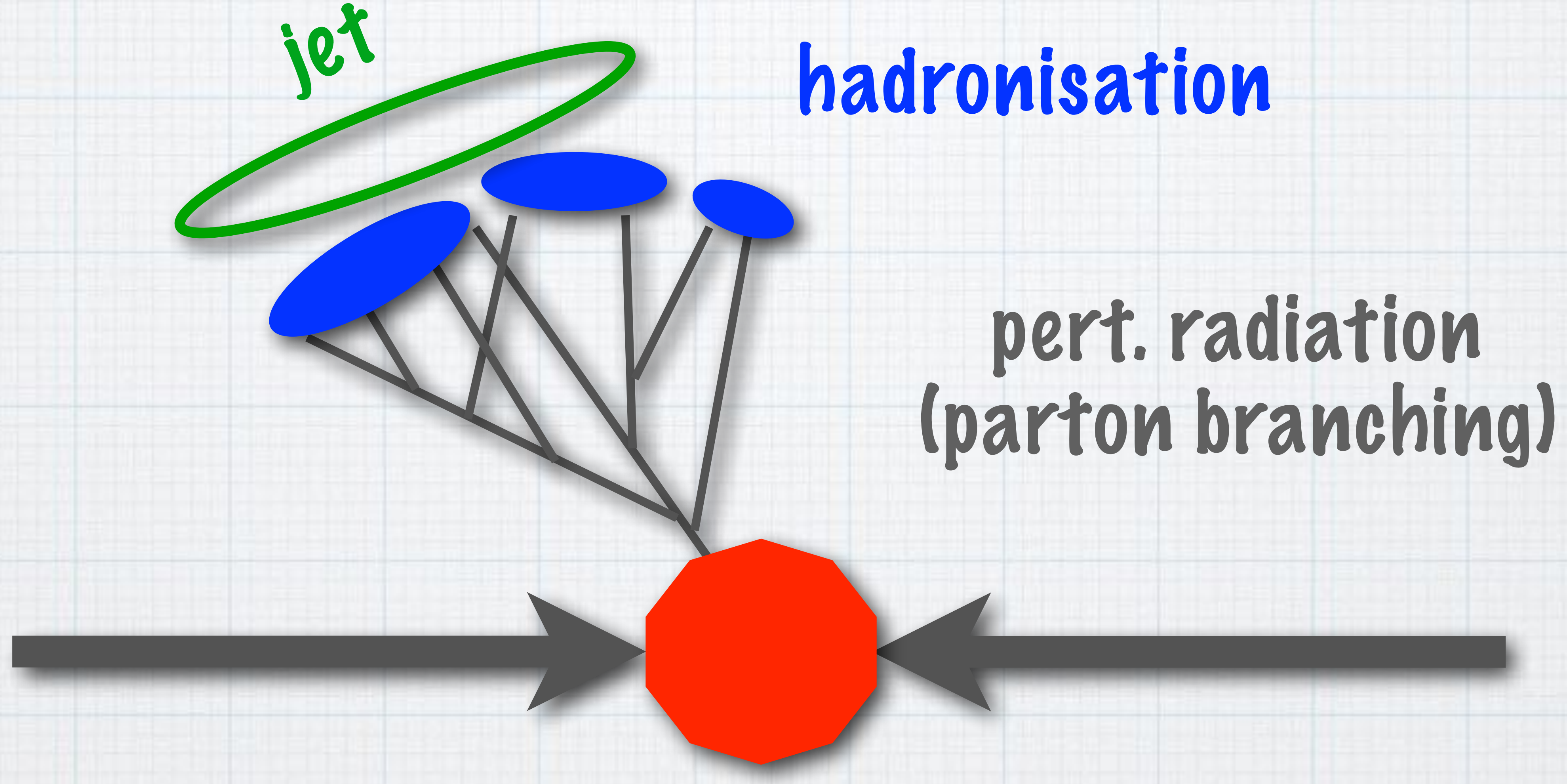
homework 2

* show that for an event made up of two particles all gen. k_t algorithms recombine them is their azimuth-rapidity distance is less than R
* things dramatically changes with many particles!

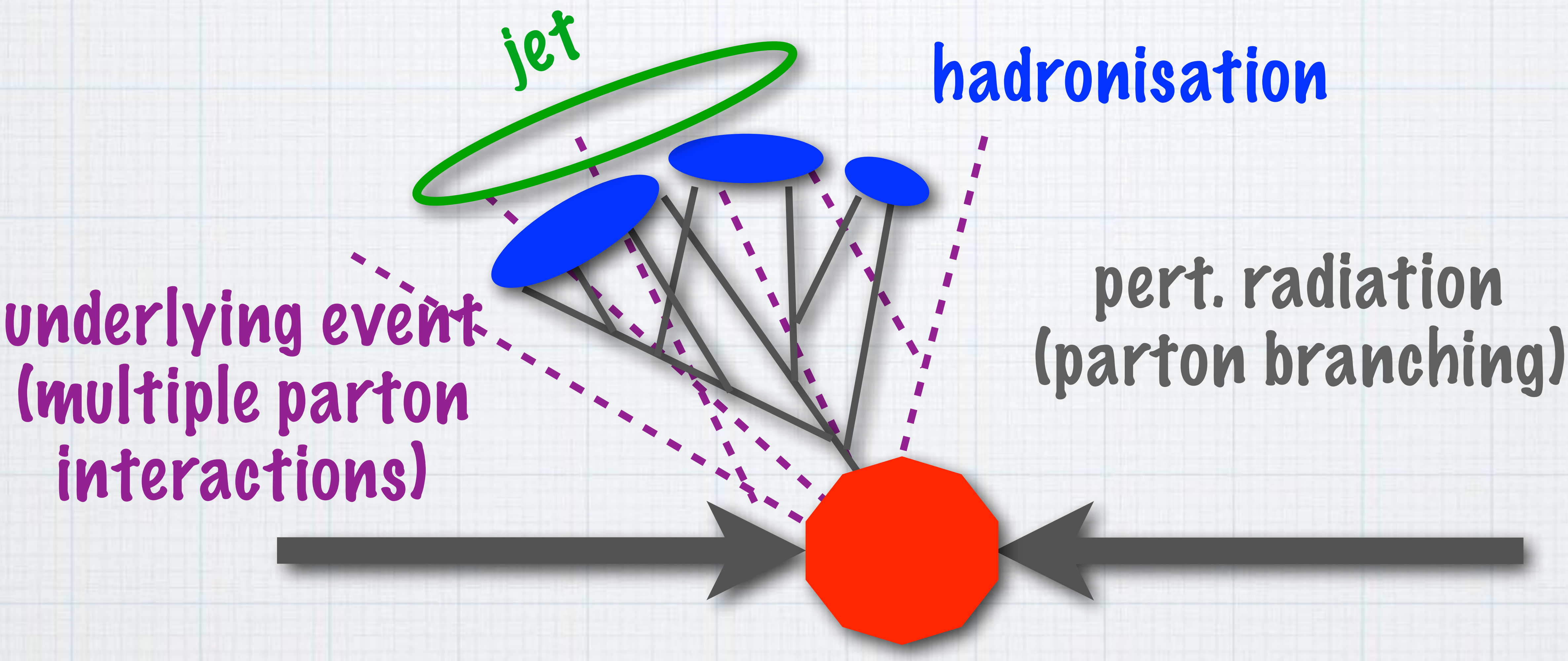
comparing them all



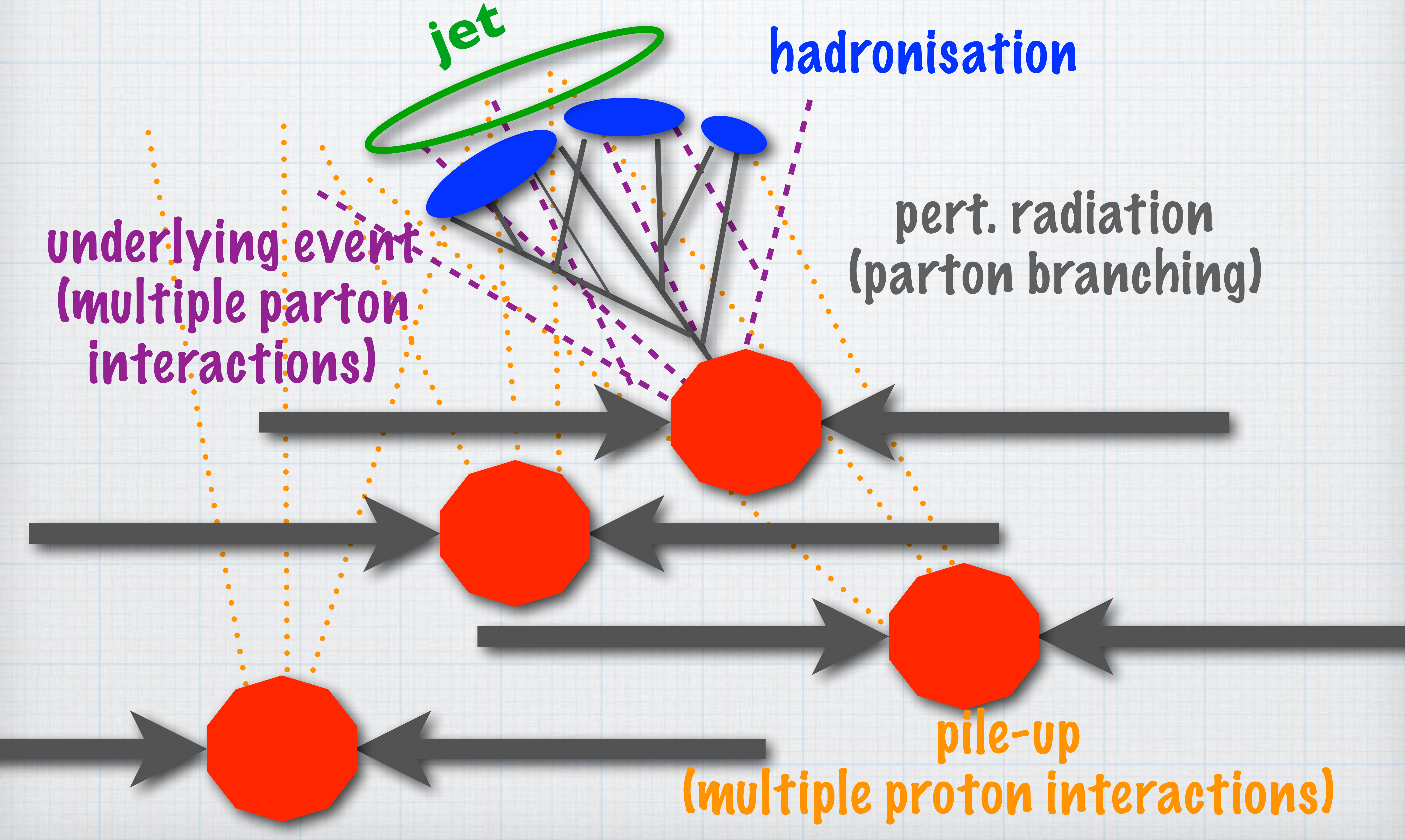
a useful cartoon



a useful cartoon



a useful cartoon



estimating p_t shifts

* we can use soft emission kinematics to estimate the changes in p_t from the hard parton to the measured quantities

* assume a finite coupling in the IR

PT radiation:

$$q: \quad \langle \Delta p_t \rangle \simeq \frac{\alpha_s C_F}{\pi} p_t \ln R$$

Hadronisation:

$$q: \quad \langle \Delta p_t \rangle \simeq -\frac{C_F}{R} \cdot 0.4 \text{ GeV}$$

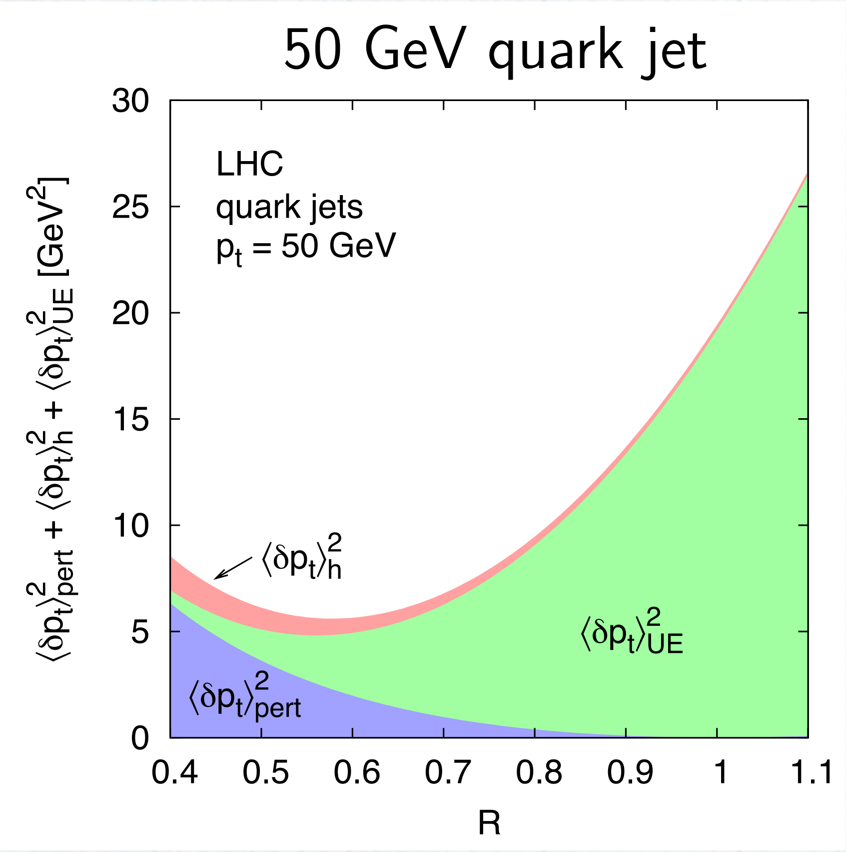
Underlying event:

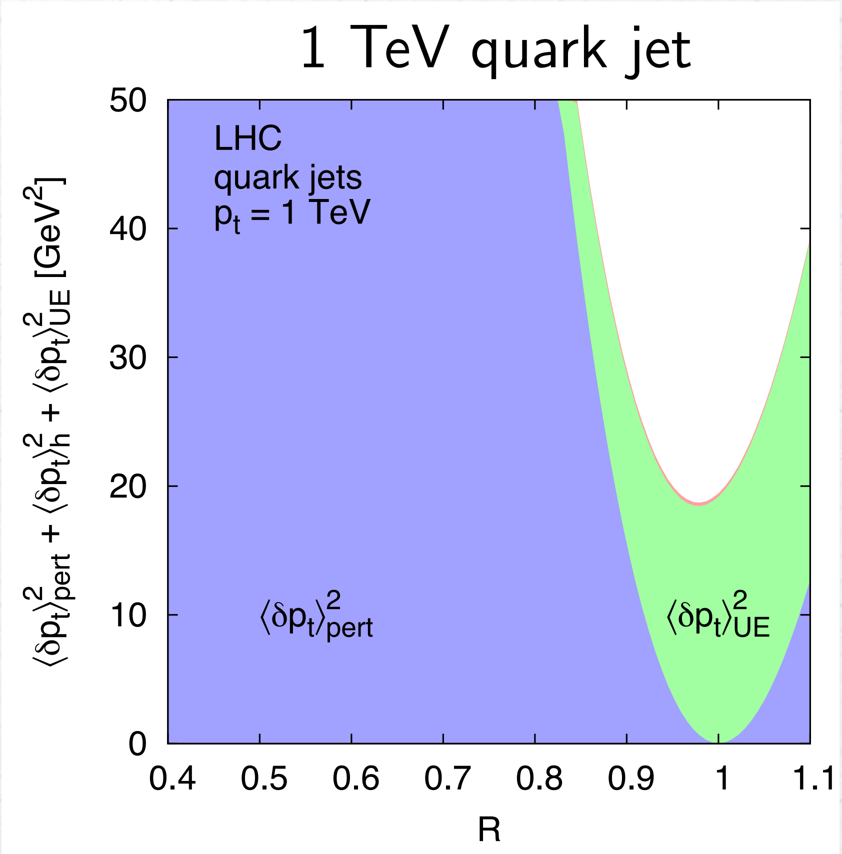
$$q, g: \quad \langle \Delta p_t \rangle \simeq \frac{R^2}{2} \cdot 2.5\text{–}15 \text{ GeV}$$

Dasgupta, Magnea, Salam (2007)

calculation 1

what is the optimal R?

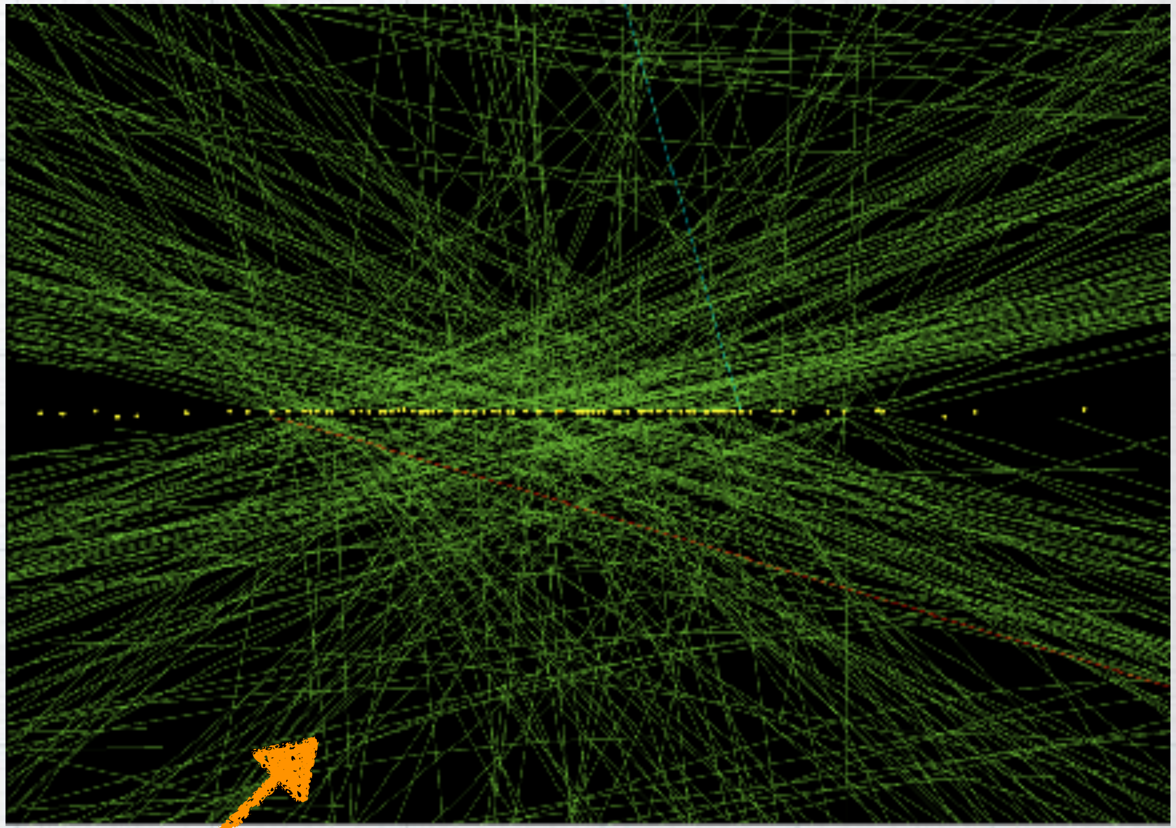




* at low p_t small R (0.4-0.6) reduces the impact of UE
* at high p_t perturbative effects dominate (see lecture 2)
* at high p_t R=1 seems excellent (good also for boosted-object, see lecture 3)

pile-up

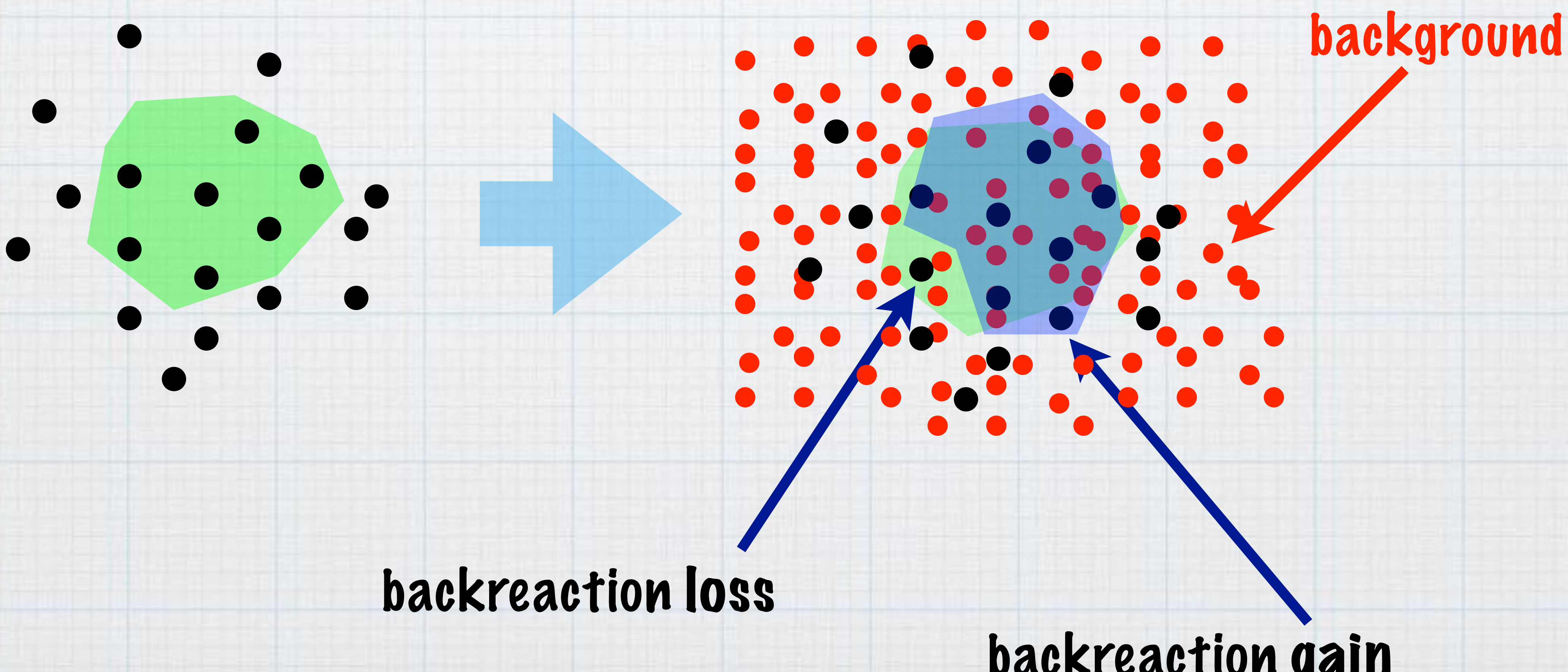
* pile-up can deposit several tens of GeV (or even hundreds, in a heavy ion collision) into a medium-sized jet
* it's a direct consequence of the desired high luminosity
* it hampers how ability of extracting useful information about the hard scatters



a 78-vertices event from CMS

hard jets and pile-up

* **susceptibility** measures how much background is picked up (jet area)
* **resiliency** measures how much the original jet is modified (backreaction)



hard jets and pile-up

* susceptibility measures how much background is picked up (jet area)
* resiliency measures how much the original jet is modified (backreaction)

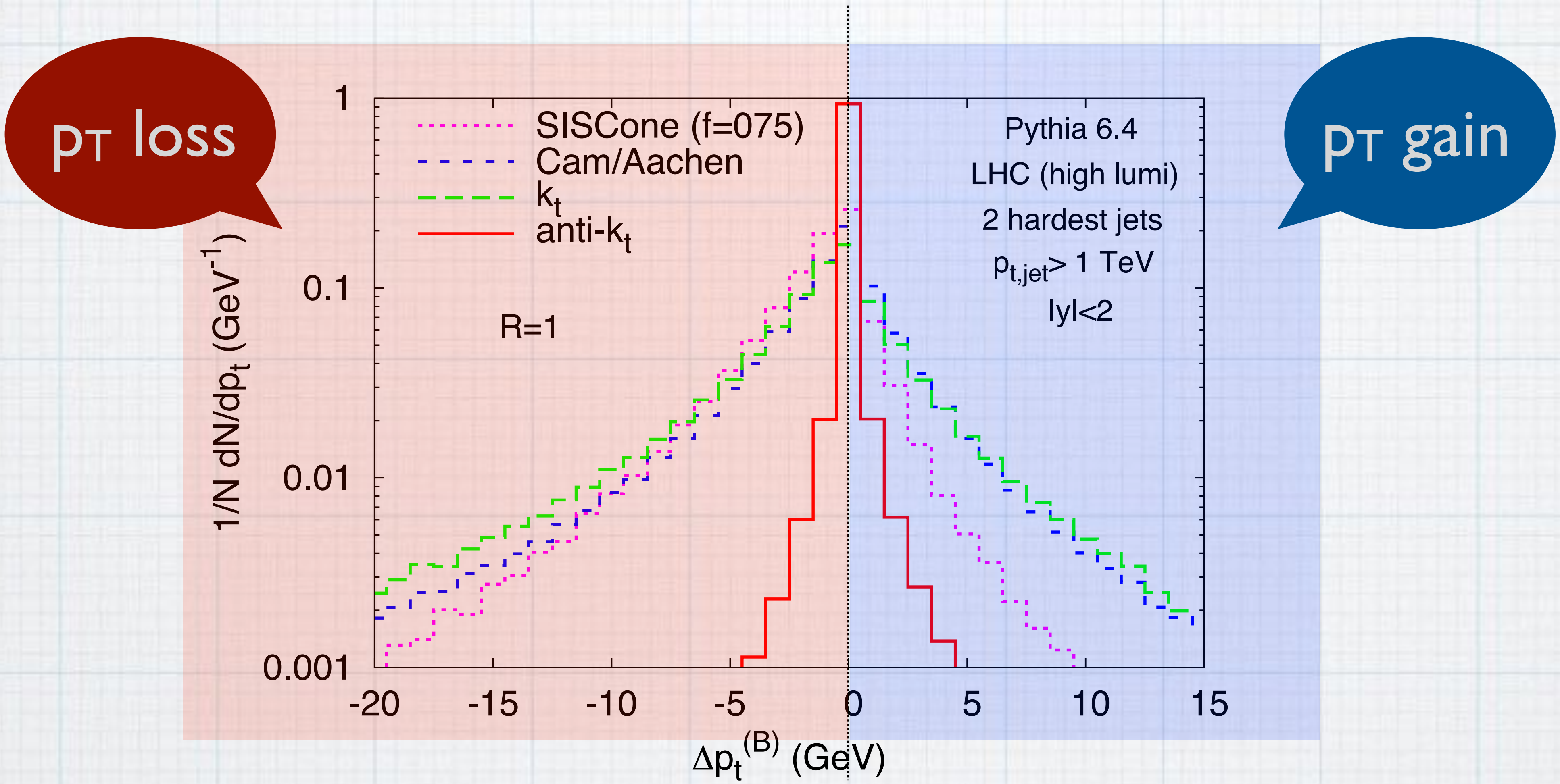
$$\Delta p_t = \rho A \pm (\sigma\sqrt{A} + \sigma_\rho A + \rho\sqrt{\langle A^2\rangle - \langle A\rangle^2}) + \Delta p_t^{BR}$$

background momentum density (per unit area)

background 'susceptibility'

backreaction 'resiliency'

resiliency



* anti-k_t jets are much more resilient to changes from background immersion
* their regular shape makes them easier to correct for detector effects
* default choice for LHC collaborations

mitigating pile-up

* Jet-based
  * Cluster the full event, determine the event-specific (ρ) and jet-specific (A) quantities, and subtract the relevant contamination from a given observable
  * Pros: largely unbiased subtraction
  * Cons: slow, potentially large(er) residual uncertainty
  * Examples: `jet area/median' in FastJet, GenericSubtractor for jet shapes, JetFFMoments for fragmentation functions, ….
* Particle-based
  * Produce a reduced event, by dropping some of the particles. Cluster this reduced event, and calculate from it the observables
  * Pros: fast, often small(er) residual uncertainty
  * Cons: not natively unbiased, can depend on choice of parameters
  * Examples: ConstituentSubtractor, SoftKiller, PUPPI, ….

for a complete review see G. Soyez, "Pile-up mitigation at the LHC: a theorist's view (2018)

summary of lecture 1

* jet definitions and jet algorithm
* the generalised k_t family
* the issue of pile-up