

ADVANCED!

Reconstruction

Nhan Tran, Fermilab
Aug 30/31, 2018

Hadron Collider Physics Summer School 2018

TABLE OF CONTENTS

Part 1: Building blocks

- a. Charged particle tracking, vertexing
- b. Precision Timing
- c. Calorimetry

Part 2: Particle reconstruction

- a. Muons
- b. Photons/Electrons
- c. Taus, Hadrons
- d. special topic: LHCb RICH detector
- e. Particle Flow

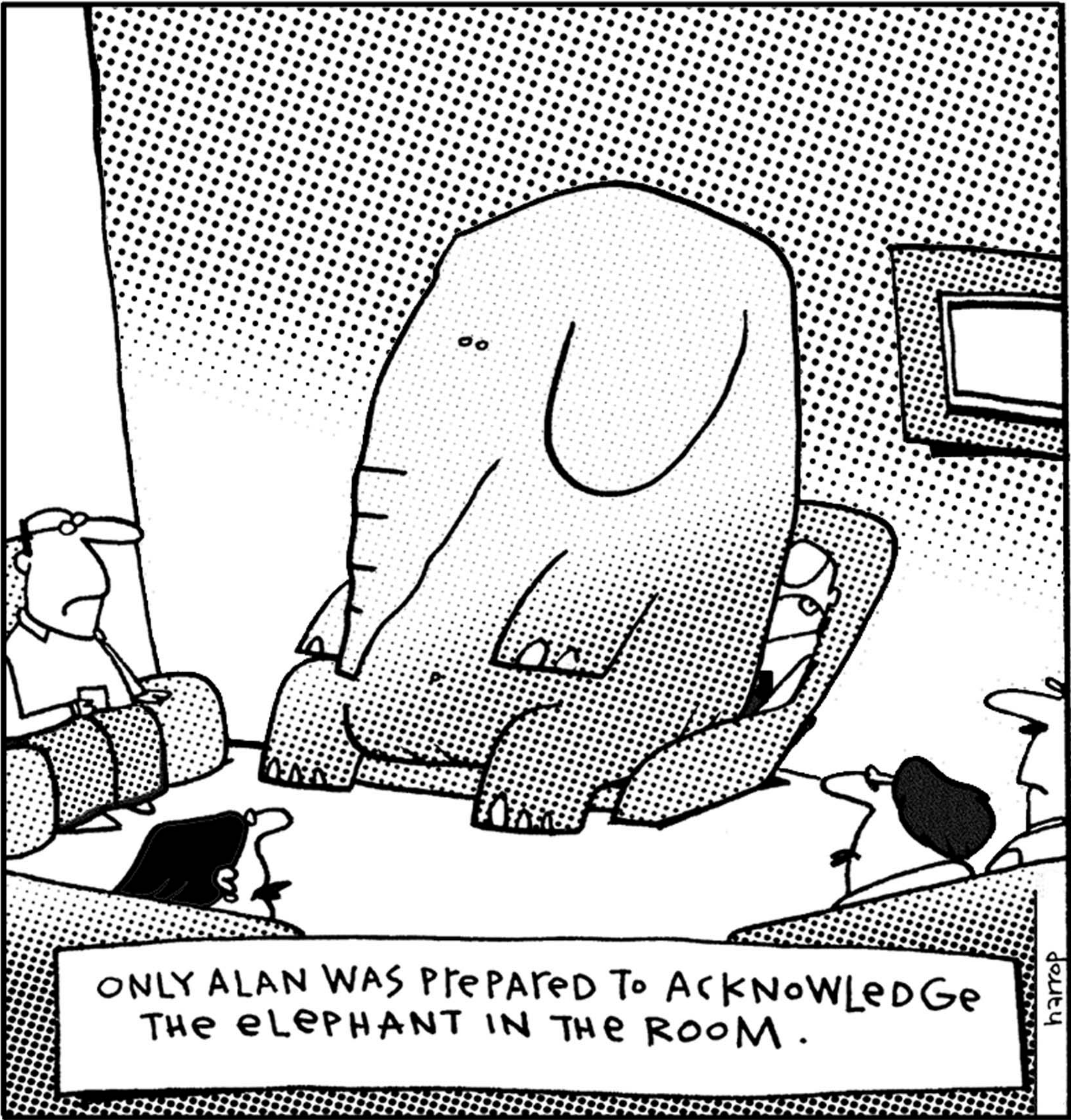
← YOU ARE HERE

A LOT OF GROUND TO COVER!
MY STRATEGY: GIVE YOU AN IDEA
OF MANY THINGS RATHER THAN
FOCUS ON A FEW

Part 3: Composite objects and beyond

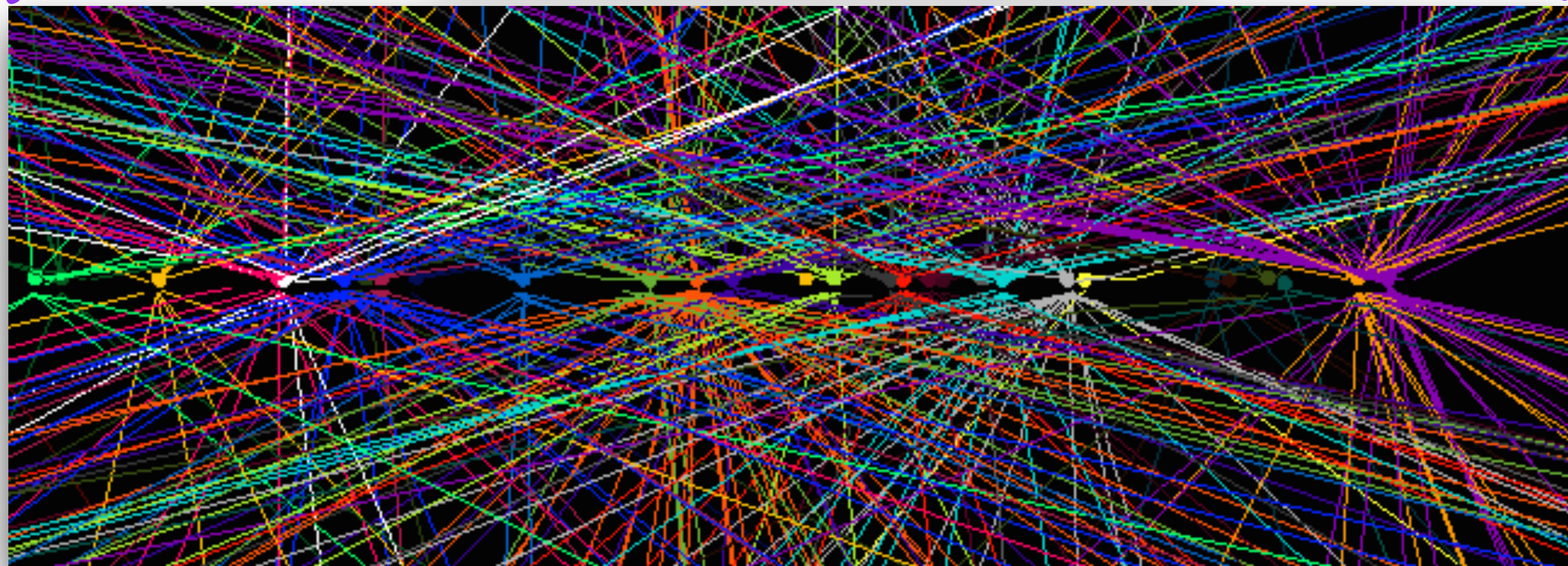
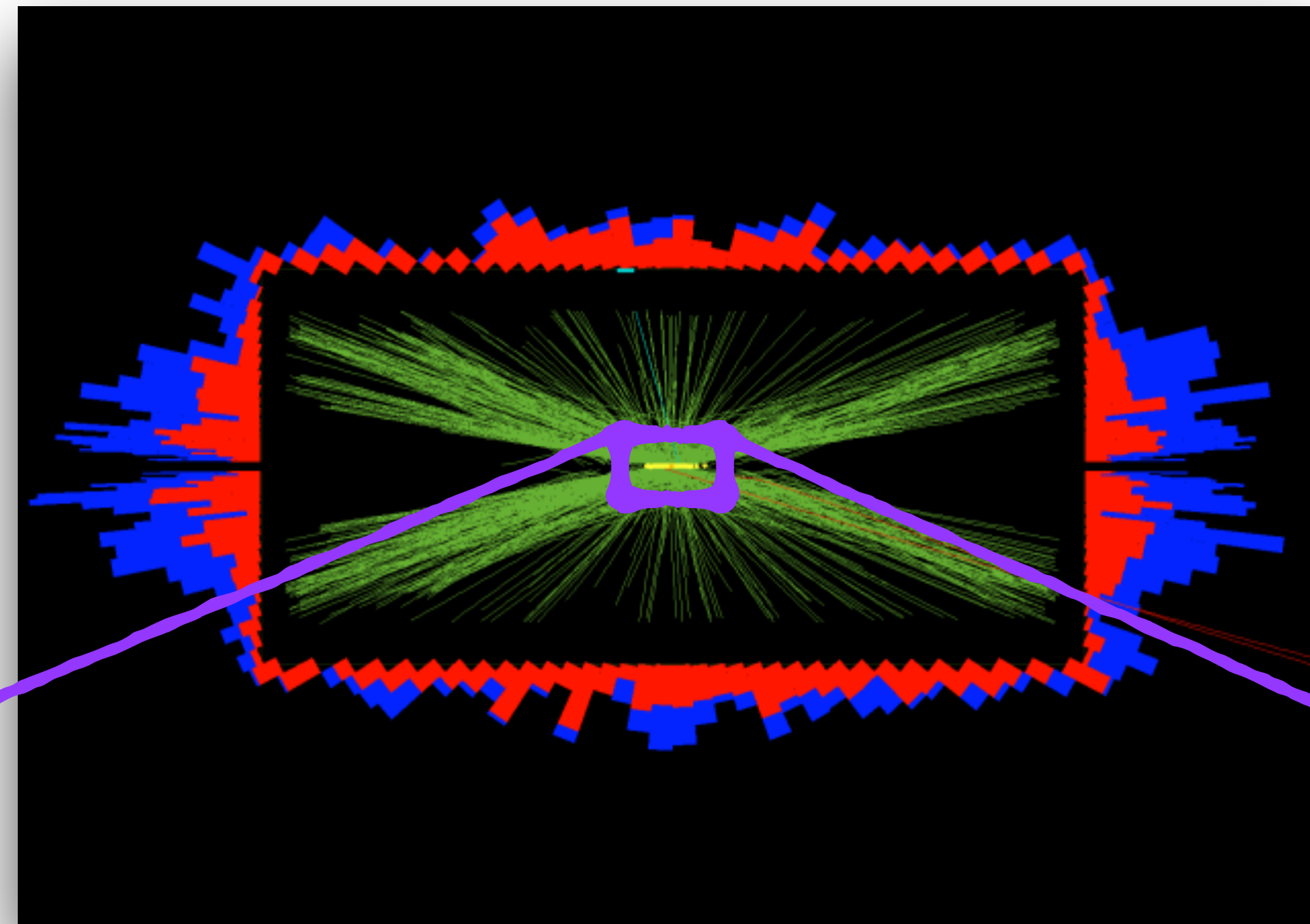
- a. Jets, MET
- b. Jet substructure
- c. Pileup Mitigation
 - c.ii. special topic: Underlying event in heavy ions
- d. Displaced/Exotic objects

PILEUP: THE ELEPHANT IN THE ROOM



WHAT IS PILEUP?

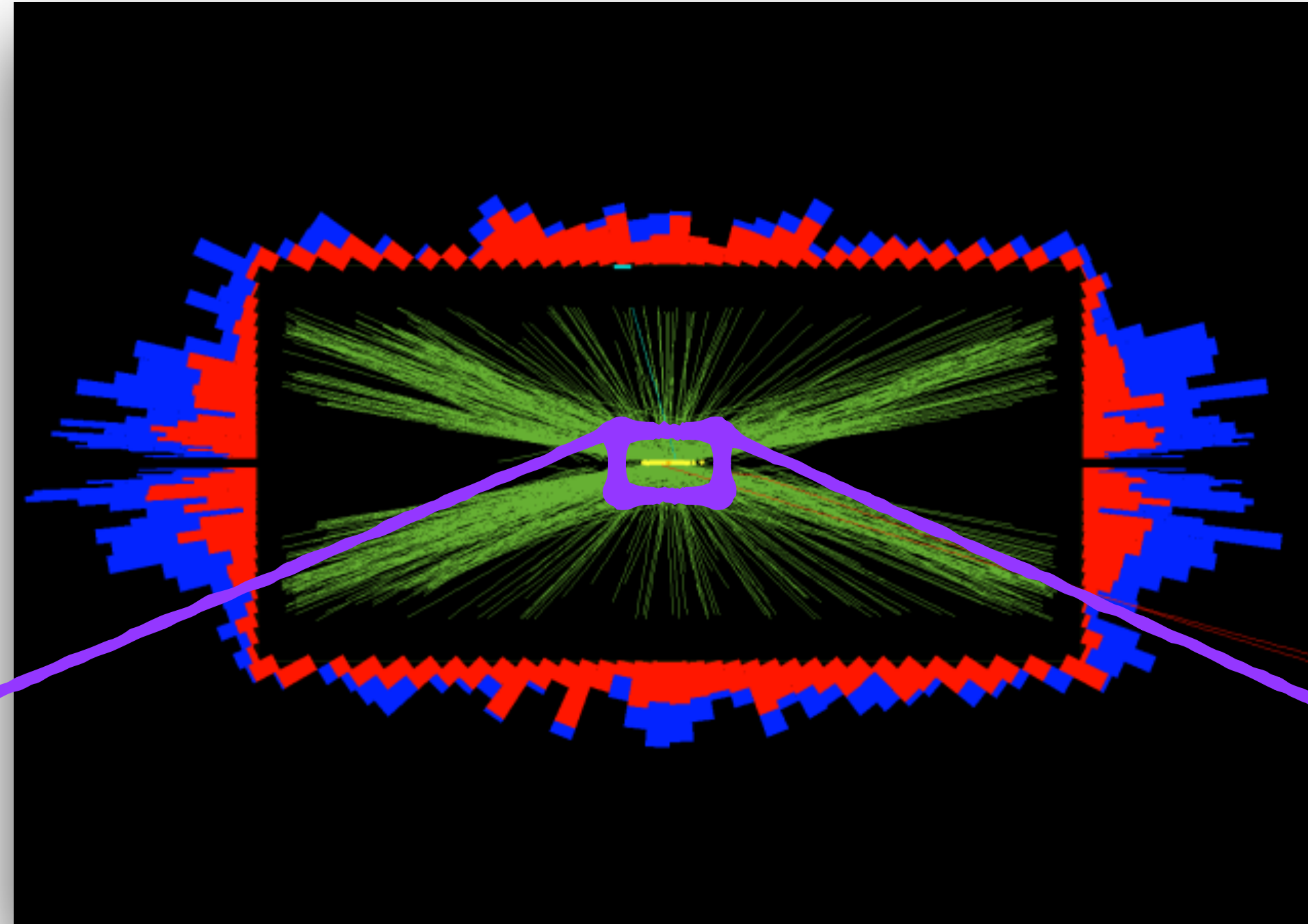
Multiple pp collisions in the same beam crossing (mostly minimum bias events)



← ~10 cm →

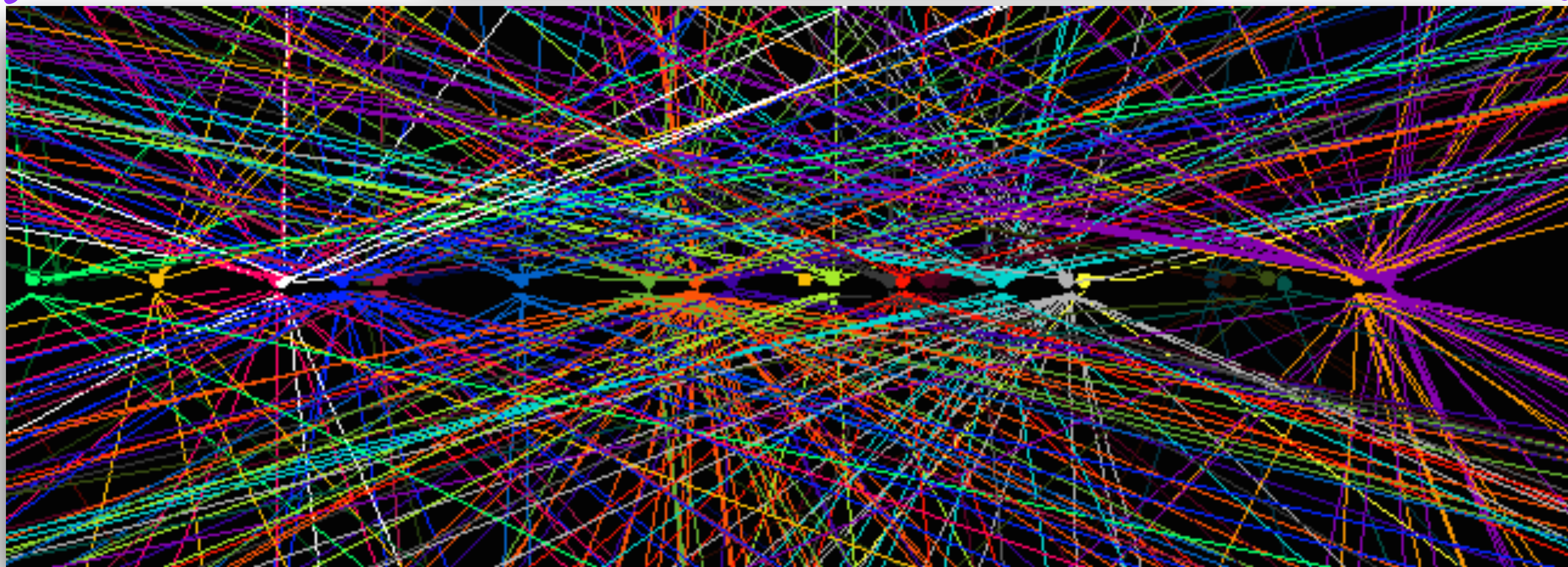
WHAT IS PILEUP?

Multiple pp collisions in the same beam crossing (mostly minimum bias events)



2012: $\langle \text{PU} \rangle \sim 20$
2016: $\langle \text{PU} \rangle \sim 20-40$
2017: $\langle \text{PU} \rangle \sim 50$
Run 3: > 50
HL-LHC: 140-200

to give a sense of scale:
1 PU vertex ~ 0.7 GeV of energy per unit area



\longleftrightarrow
 ~ 10 cm

Also was sometimes referred to as “global event description”

Combine the sub-detector information in a complementary way in a single algorithm

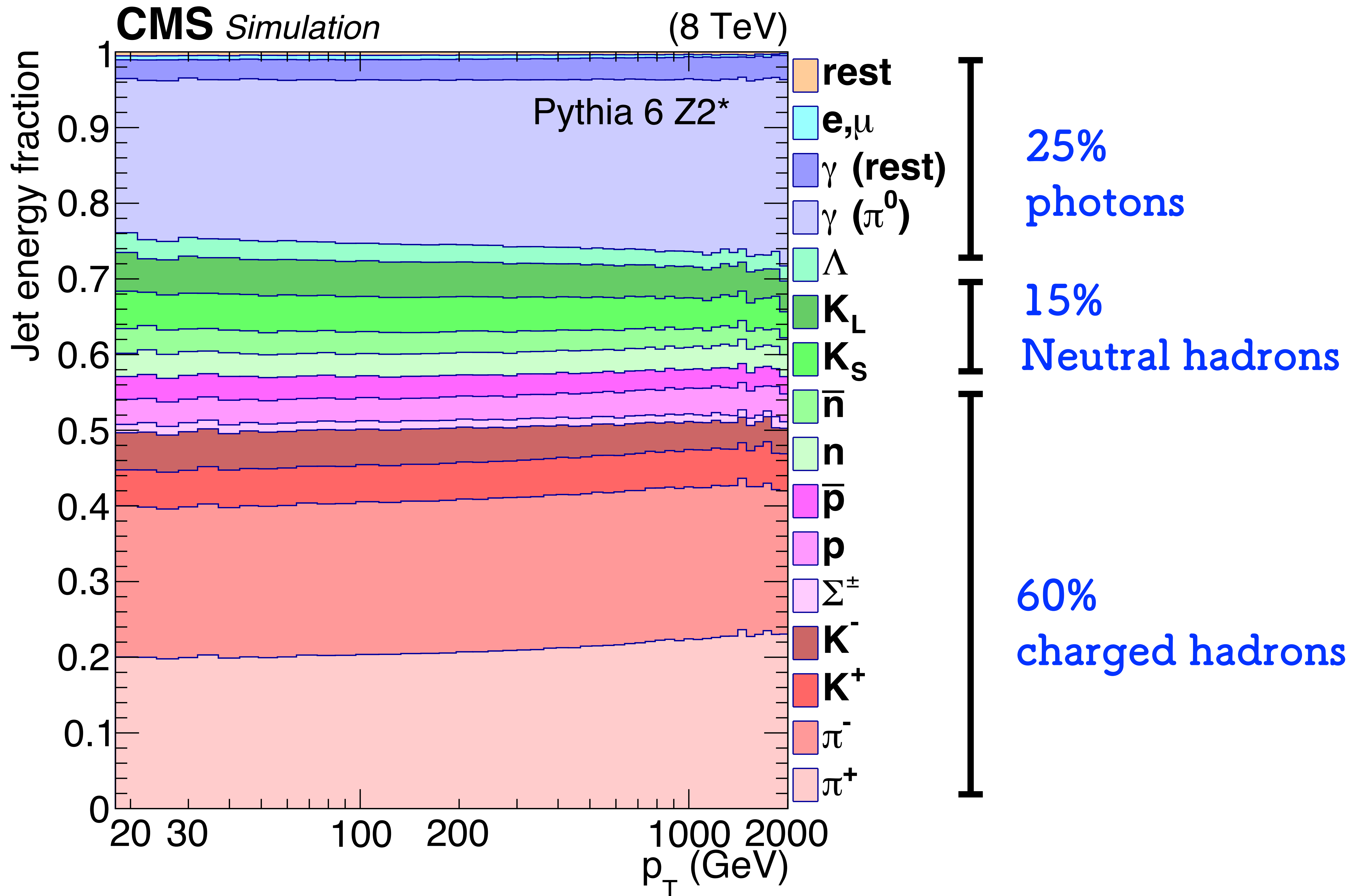
Outputs a list of particles:

muons, electrons, photons, neutral hadrons, charged hadrons

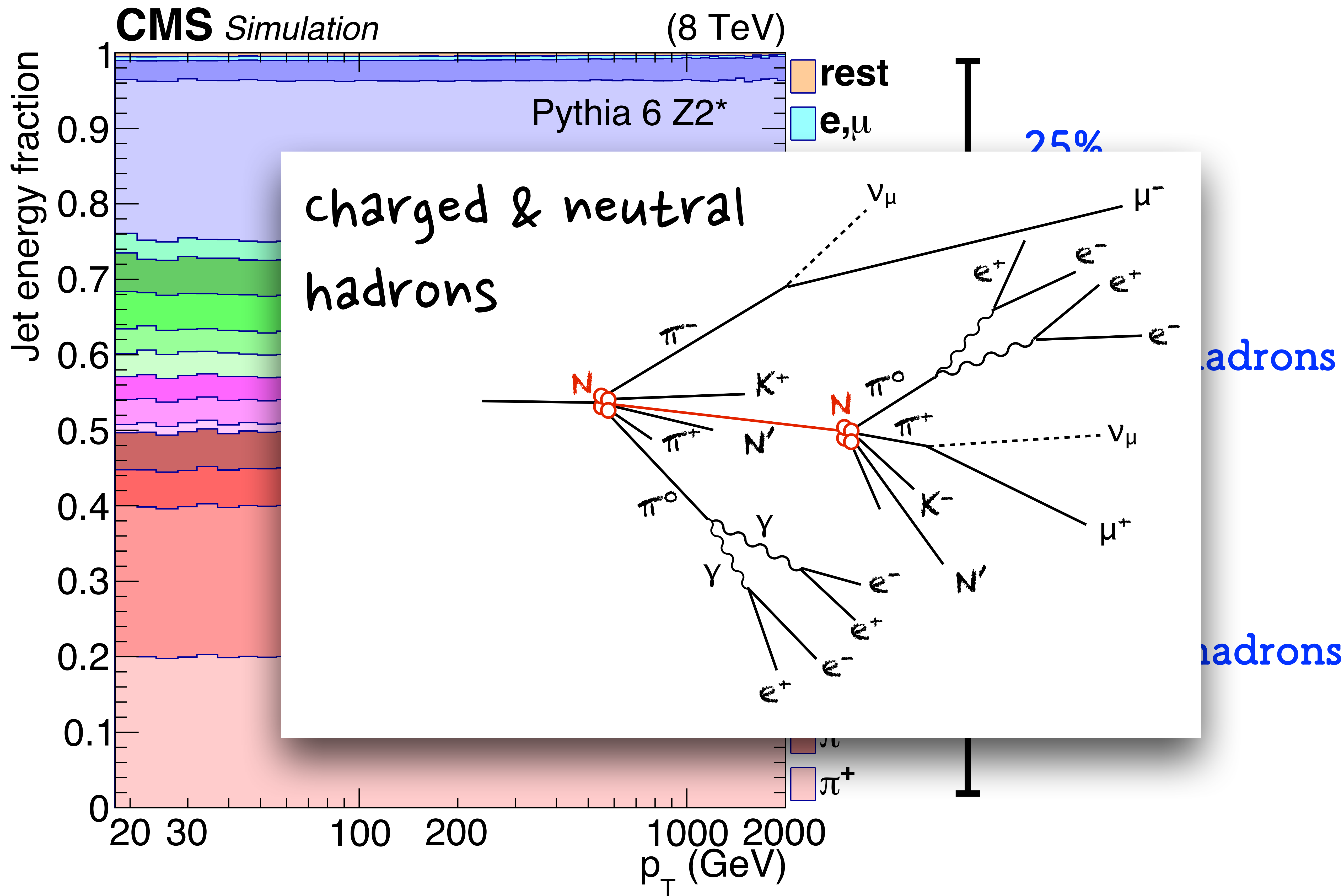
Avoids double-counting of the energy to create a self-consistent view of the event

Breaking down the event at the particle level can aid in things like jet substructure and pileup mitigation (more later)

SETTING THE STAGE, JET COMPOSITION



SETTING THE STAGE, JET COMPOSITION

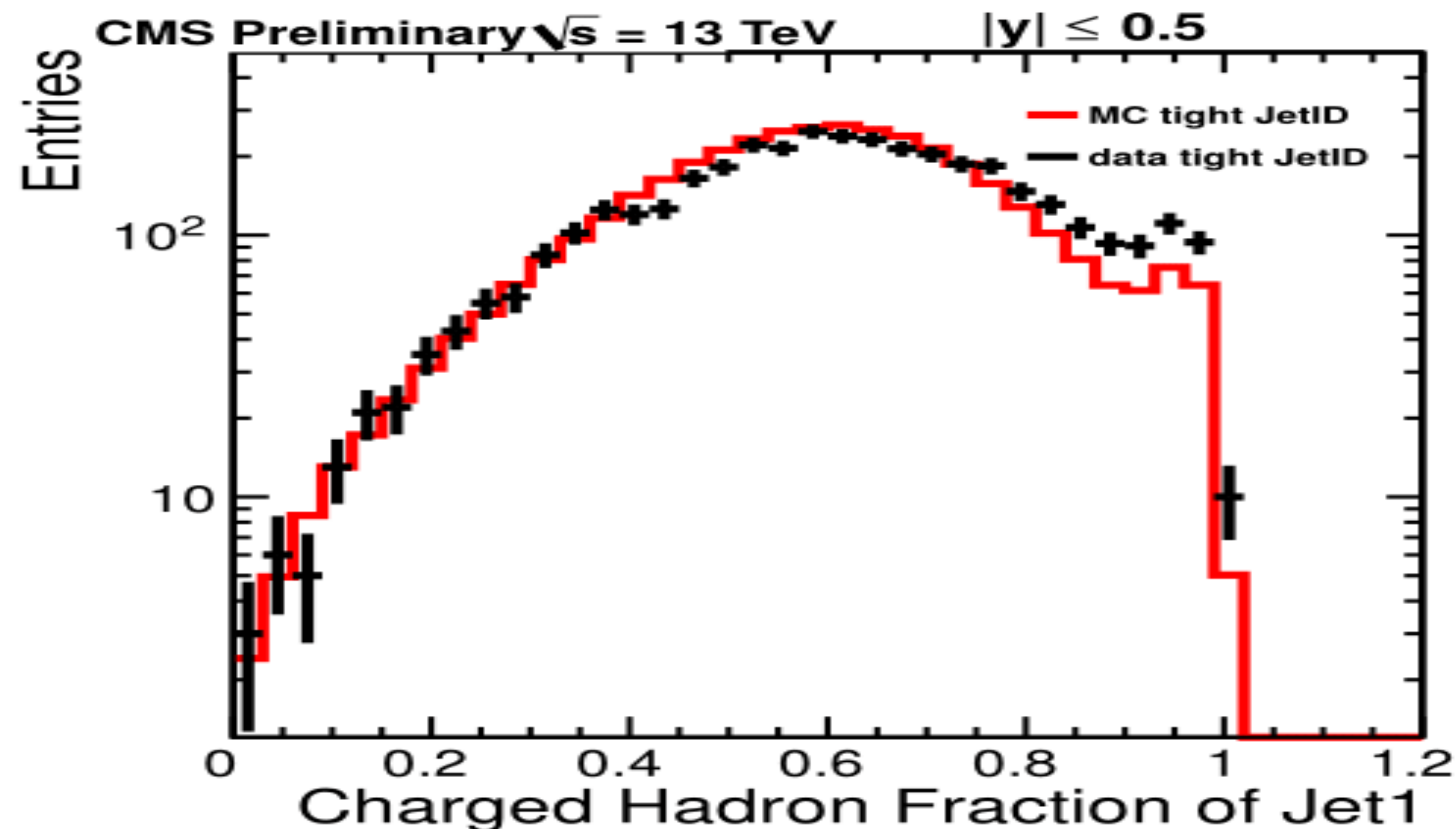


BUT...FLUCTUATIONS

The fraction of the jet energy that is charged/neutral hadron and photon fluctuates quite a bit

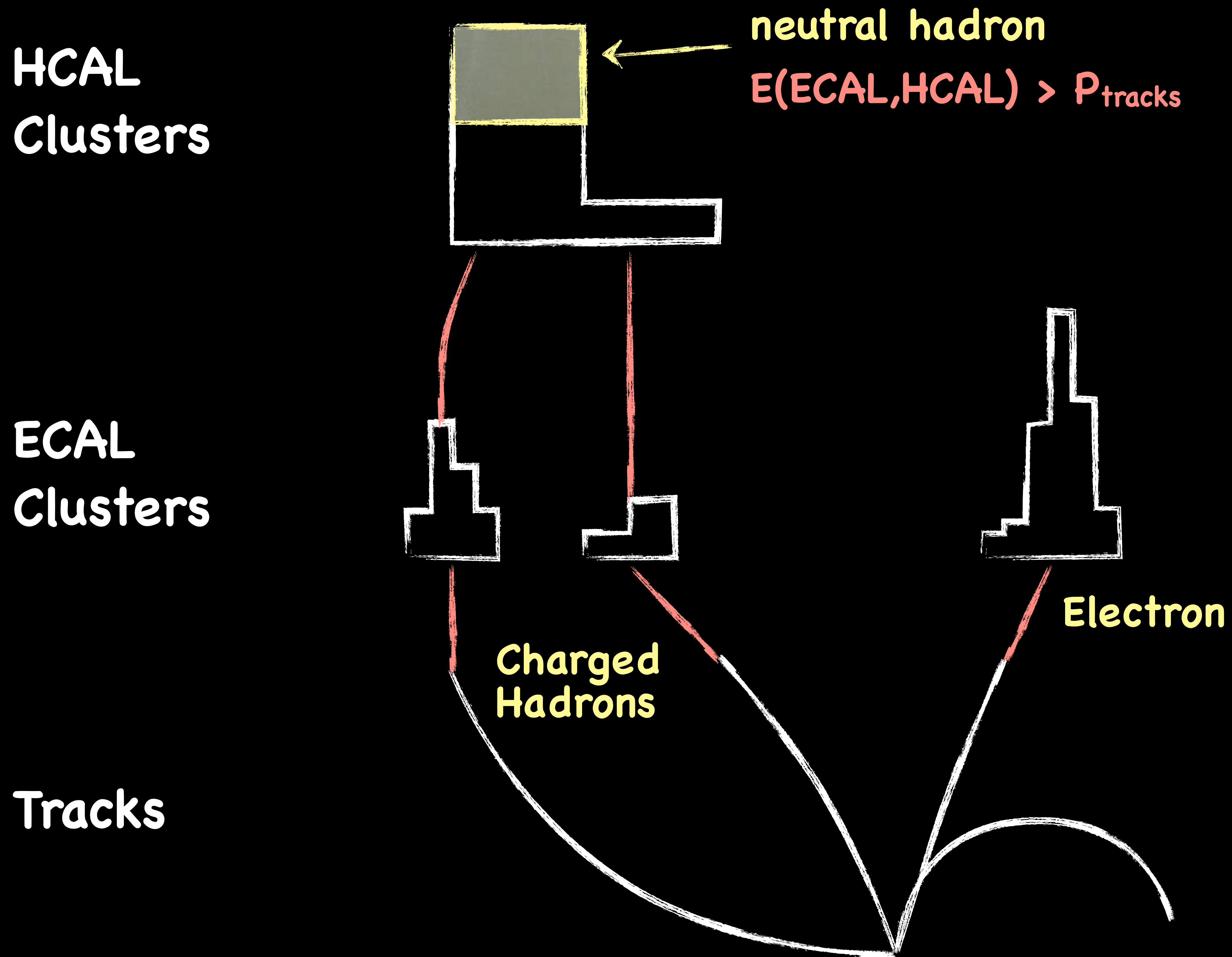
Fluctuations on the order of 20-30% of the jet energy

Therefore, you still have to measure all the energy in the event!



How to reconstruct individual particles?

First Associate Hits within Each Detector



Very basic view of the Particle Flow Algorithm

Clean the event during reconstruction

Find and "remove" muons (σ_{track})

Find and "remove" electrons ($\min[\sigma_{\text{track}}, \sigma_{\text{ECAL}}]$)

Find and "remove" converted photons ($\min[\sigma_{\text{track}}, \sigma_{\text{ECAL}}]$)

Find and "remove" charged hadrons (σ_{track})

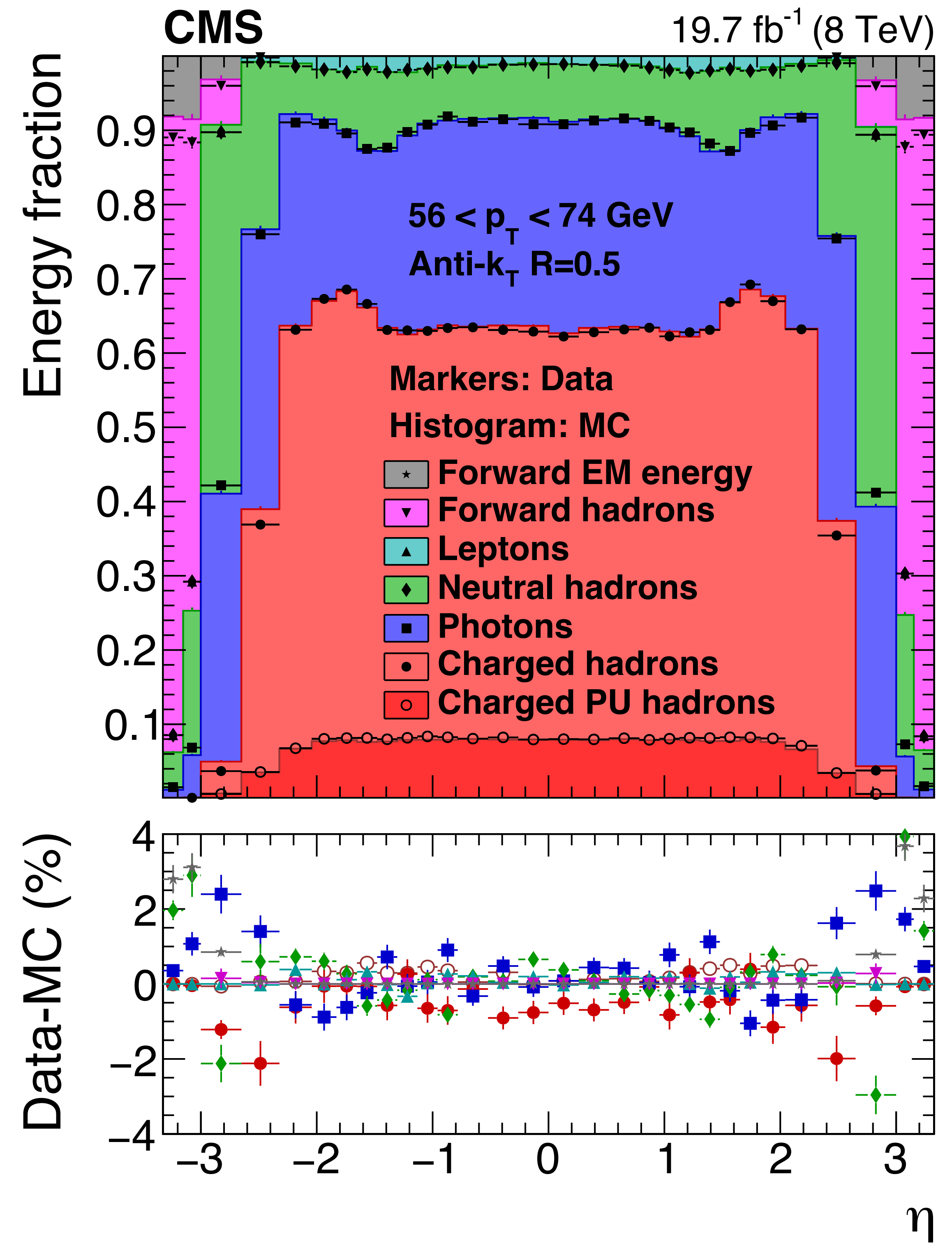
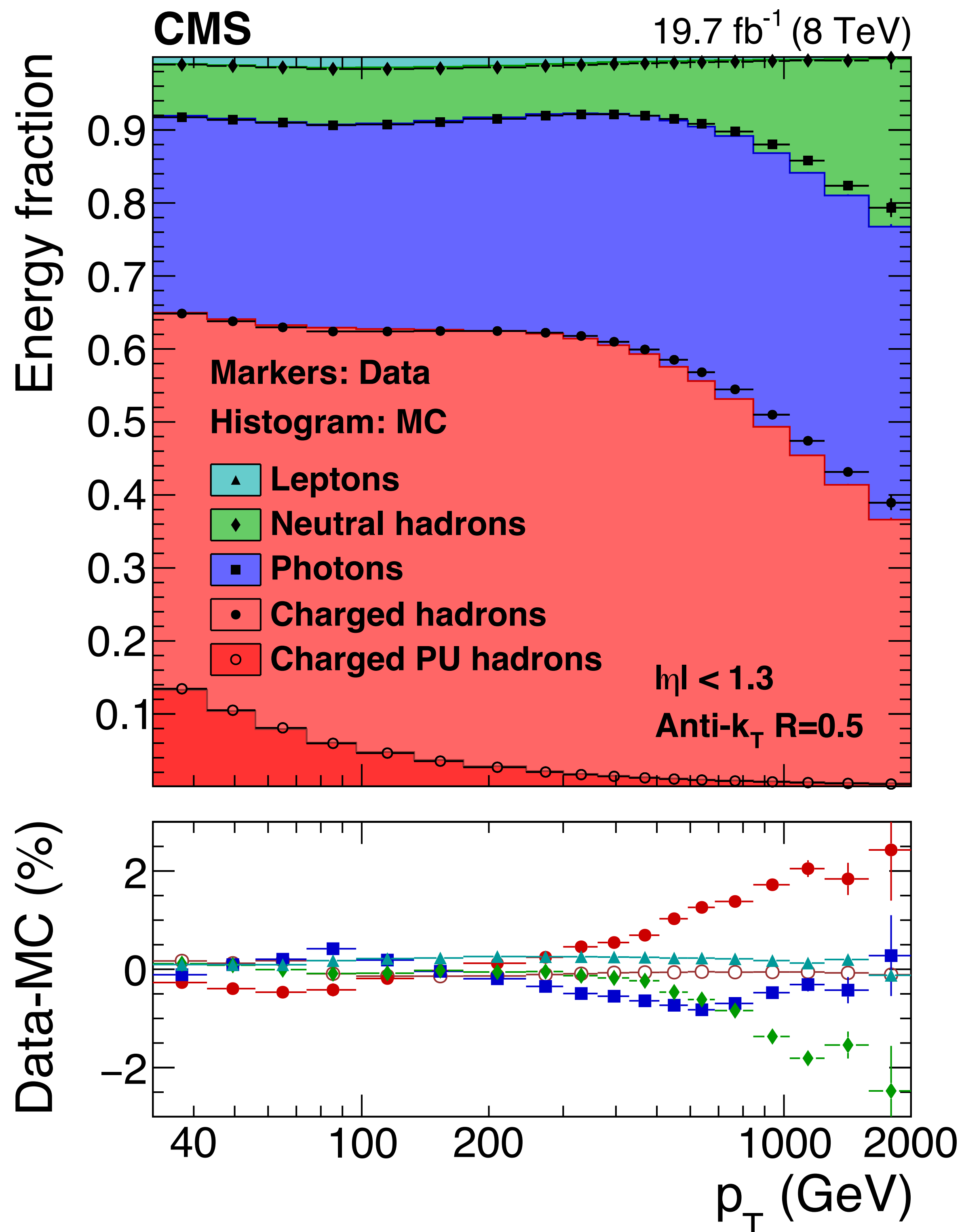
Find and "remove" V0's (σ_{track})

Find and "remove" photons (σ_{ECAL})

Left with neutral hadrons (10%) ($\sigma_{\text{HCAL}} + \text{fake}$)

Use above list of Reconstructed Particles to describe the entire event!

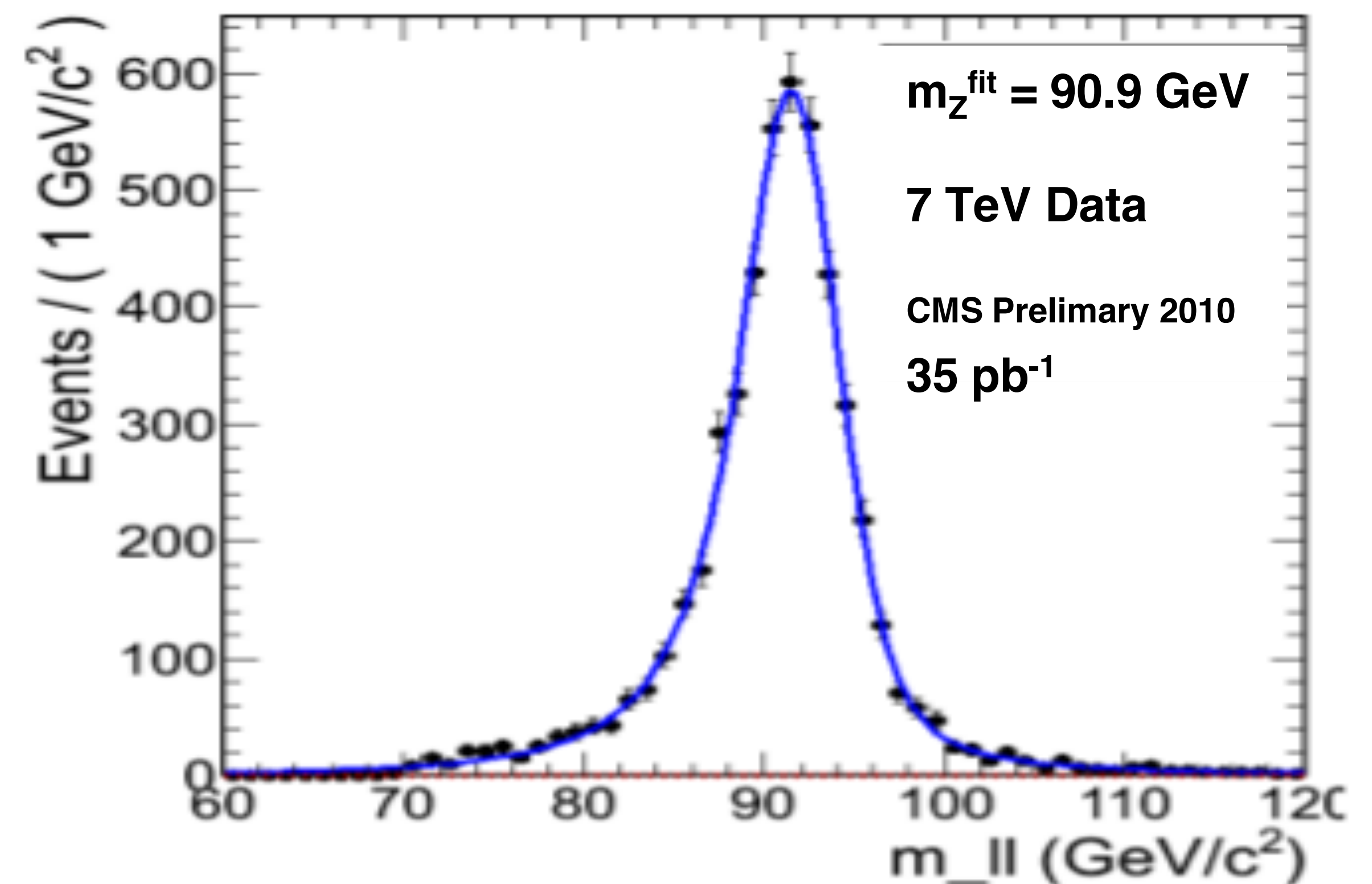
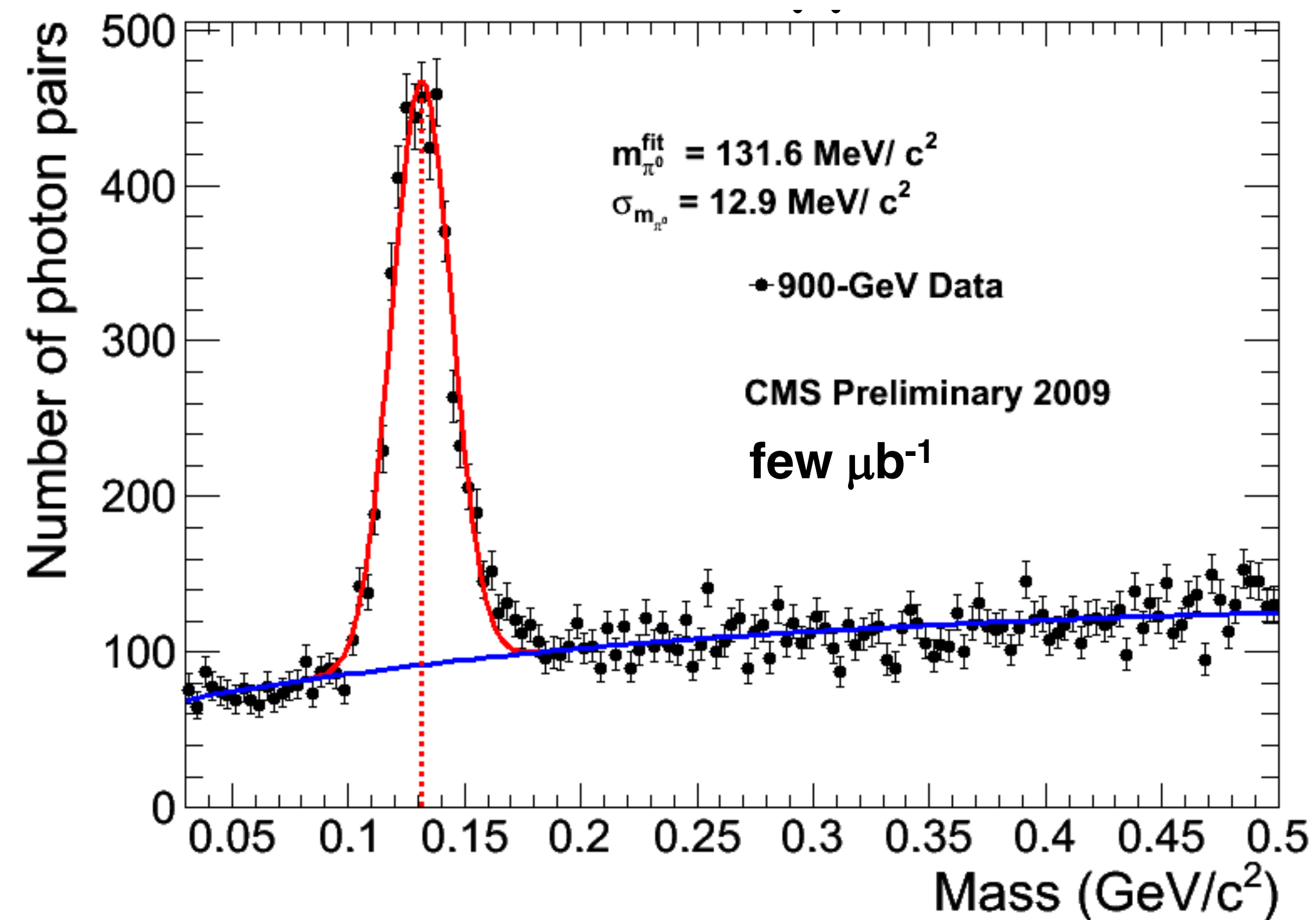
SETTING THE STAGE, JET COMPOSITION



WHY THE CHANGE AT HIGH PT?

In-situ calibration of particle flow candidates:

Electrons/photons/muons use the Z and π^0

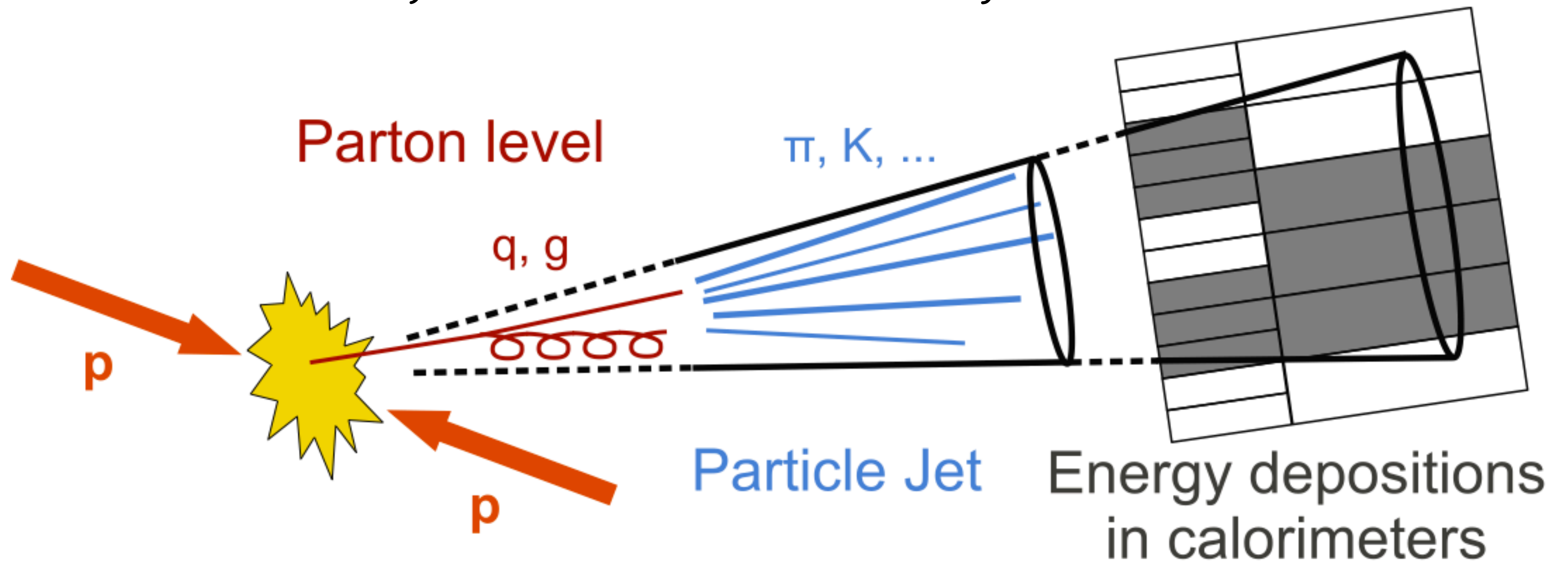


Calibration of the charged/neutral pions use isolated tracking to fit for the energy of charged hadrons calorimeter energy

$$E = a + b(p, \eta) E_{\text{ecal}} + c(p, \eta) E_{\text{hcal}}$$

3. COMPOSITE OBJECTS AND BEYOND

Now that we have multiple particles,
let's talk about jets a little more formally now

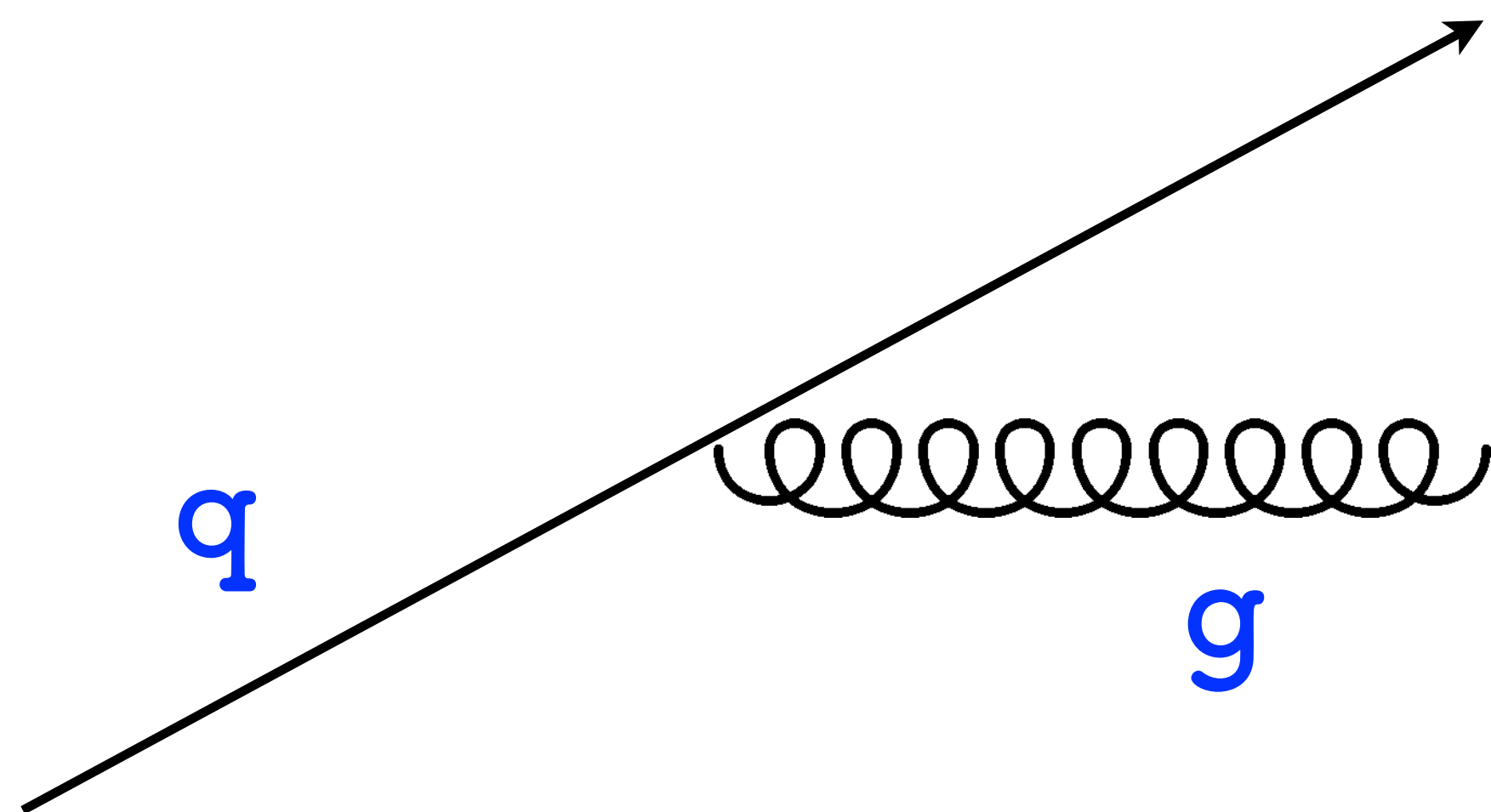


Jet = a spray of stuff (typically from q/g)
reconstructed as a single object

How to group particles/deposits/etc. together to make a jet?

Jet clustering algorithms have a looong history, but to keep it short — for precise predictions, it is important to have a formal connection between theory and experiment

Often referred to as “IRC safe”



The result of the jet algorithm stable against infinitely soft and collinear emissions

Infrared, IR: As $E \rightarrow 0$

Collinear, C: As $\Delta R \rightarrow 0$

Hierarchical jet clustering algorithms

Compute a “distance” between each particle

Recombine particles pairwise based on smallest “distance” until some condition is met

Distance measure: $d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta R_{ij}}{R^2}$ **Condition:** $d_{ij} < d_{iB} = k_{ti}^{2p}$

Jet distance parameter, R

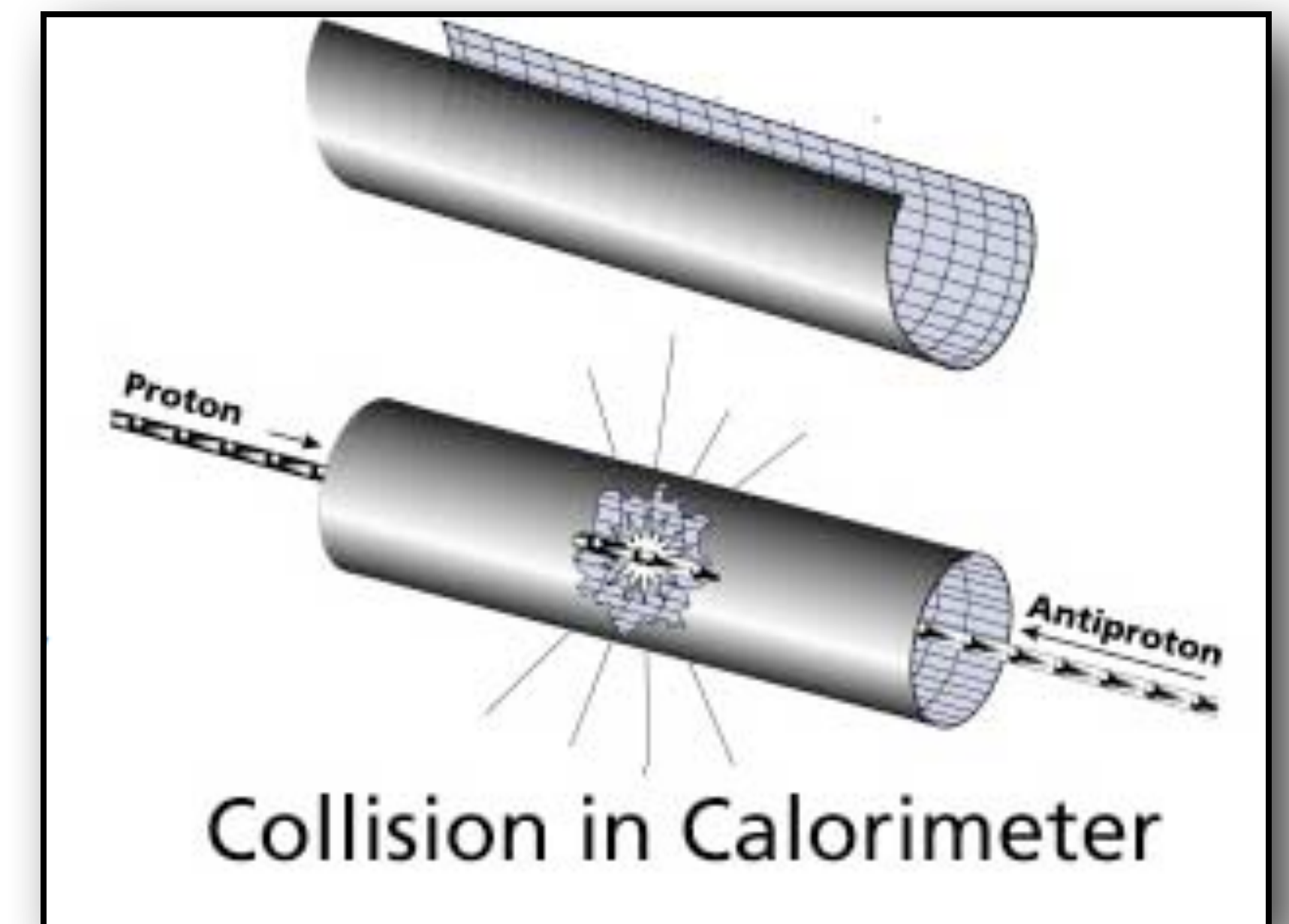
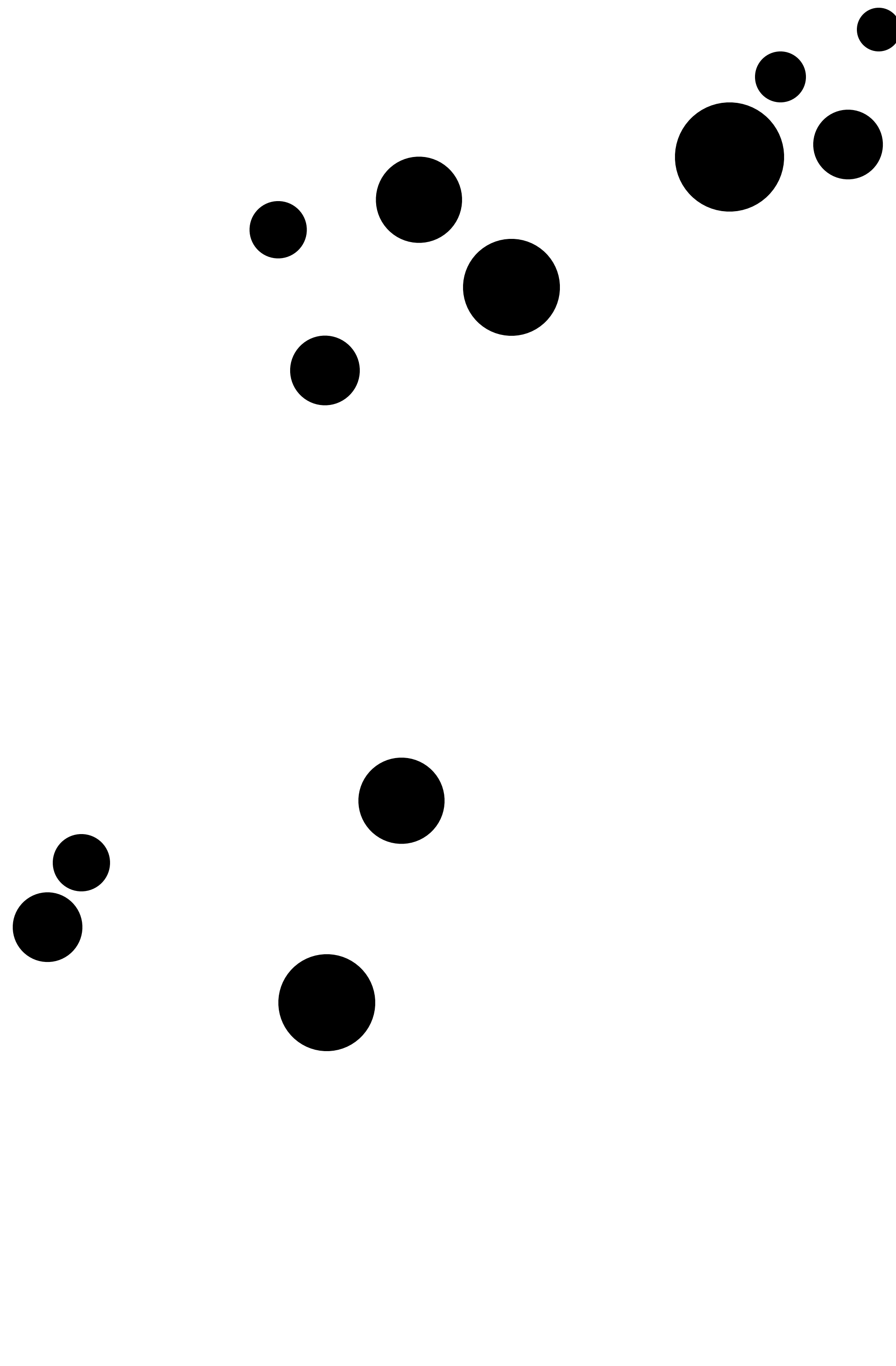
When:

p = 1, kT algorithm - start with softest particles

p = 0, CA algorithm - start with closest particles

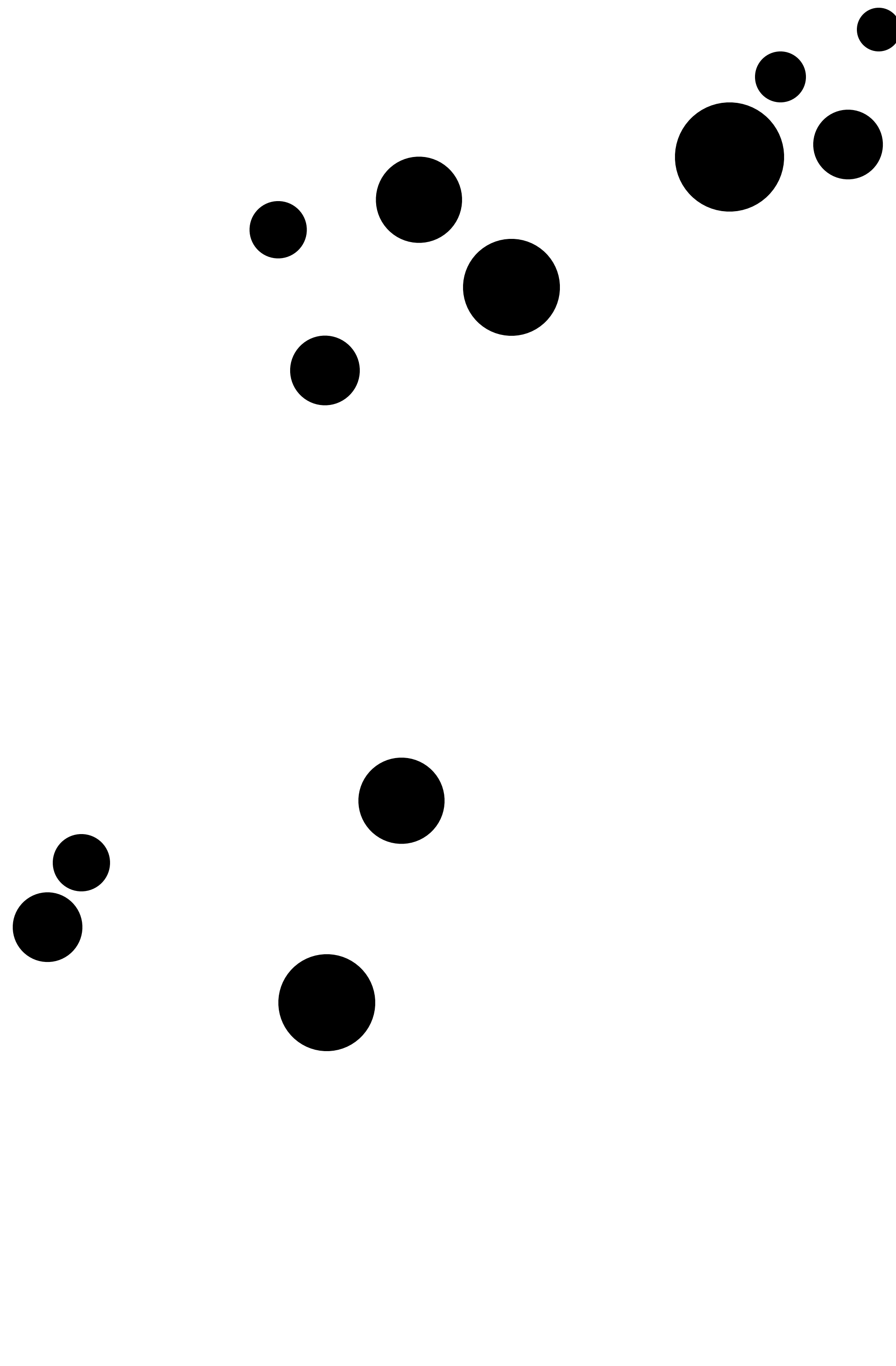
p = -1, anti-kT algorithm - start with hardest particles

Cartoon event display - PF particles



Circle = position of particle within the detector
Area \sim energy of particle

Example: Cambridge Aachen Jet Clustering



Merging conditions in CA:

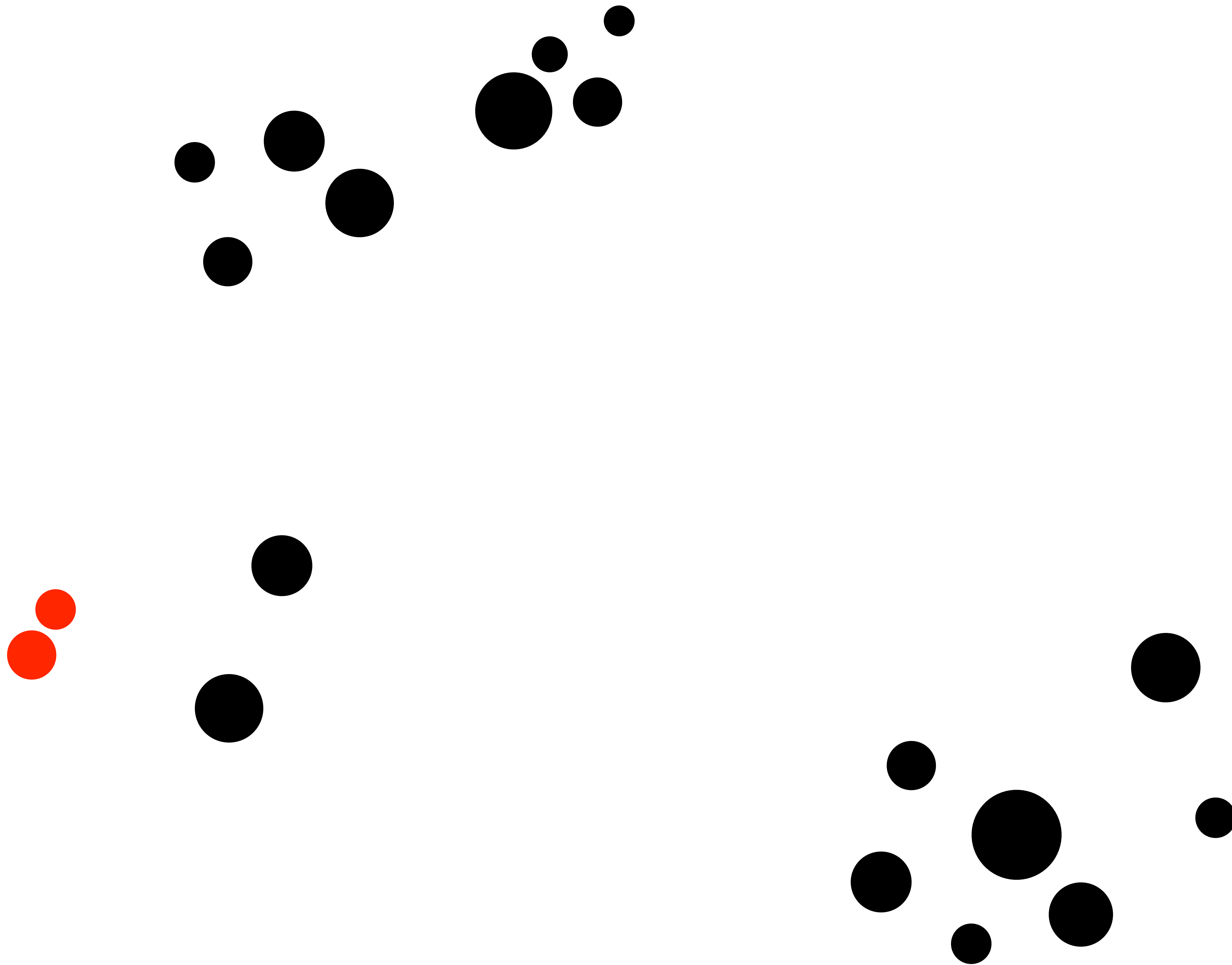
$$d_{ij} = \Delta R/R < d_{iB}$$

$$d_{iB} = 1 \text{ for CA}$$

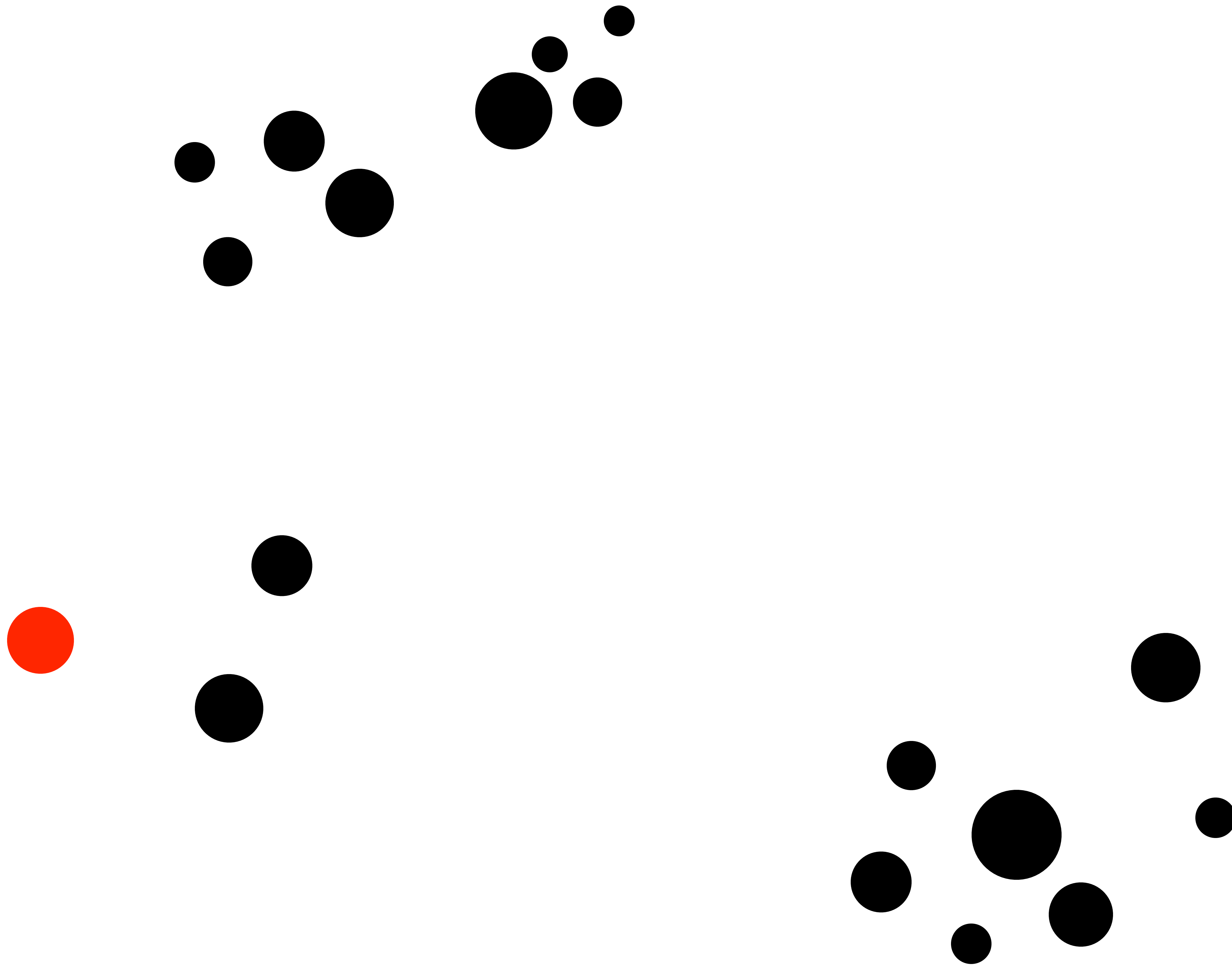
$$\Delta R/R < 1$$

$$\Delta R < R (!)$$

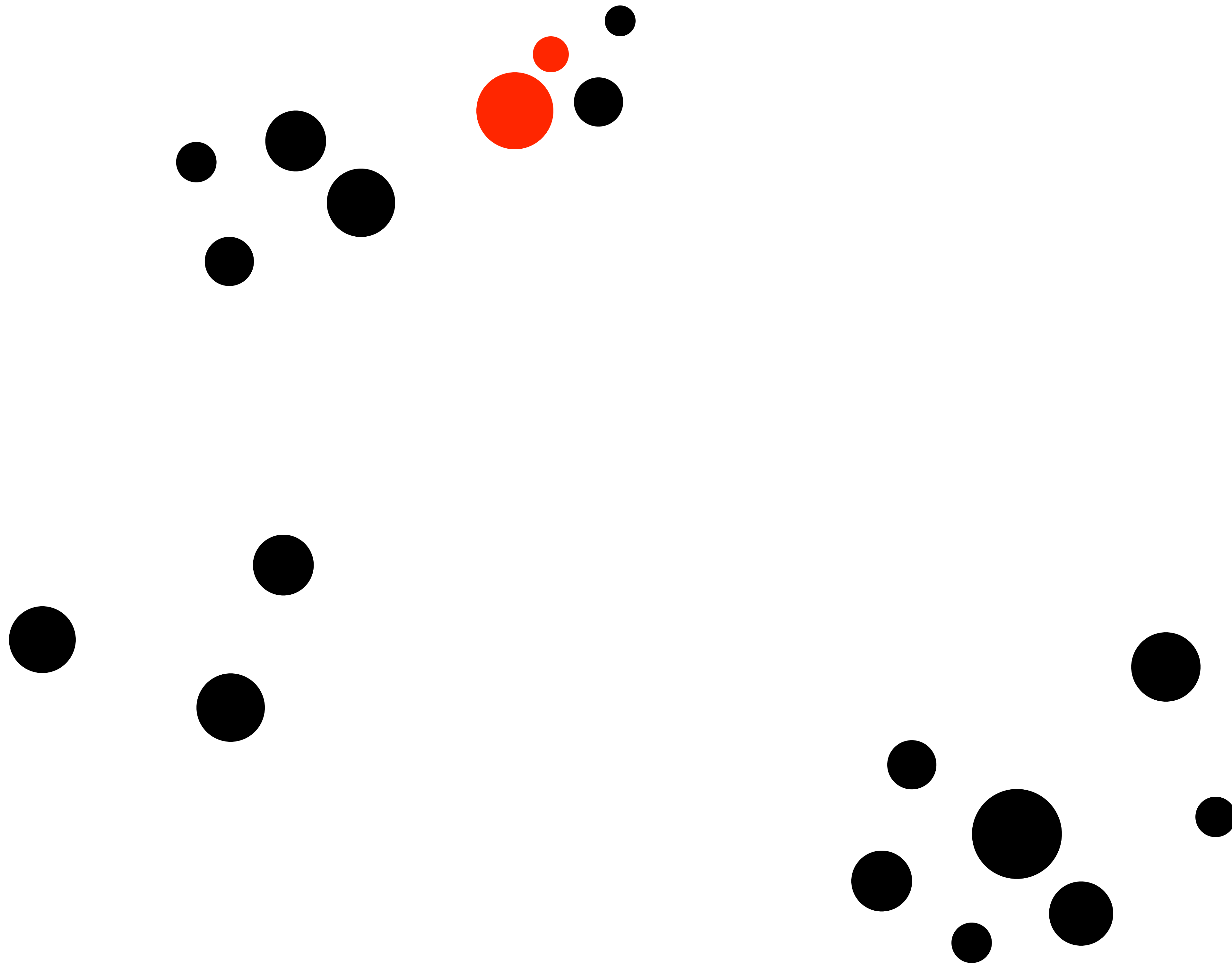
Find the closest pair

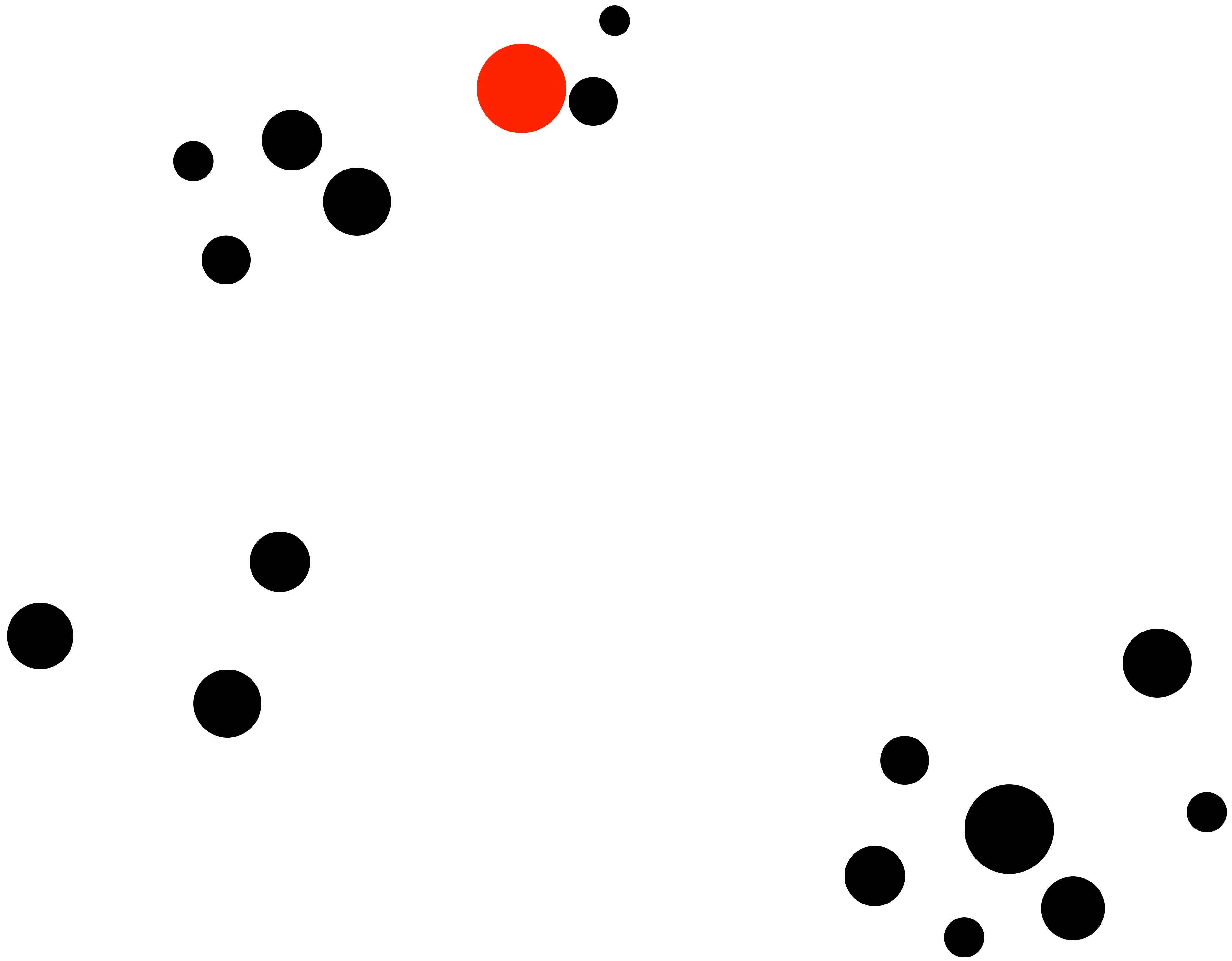


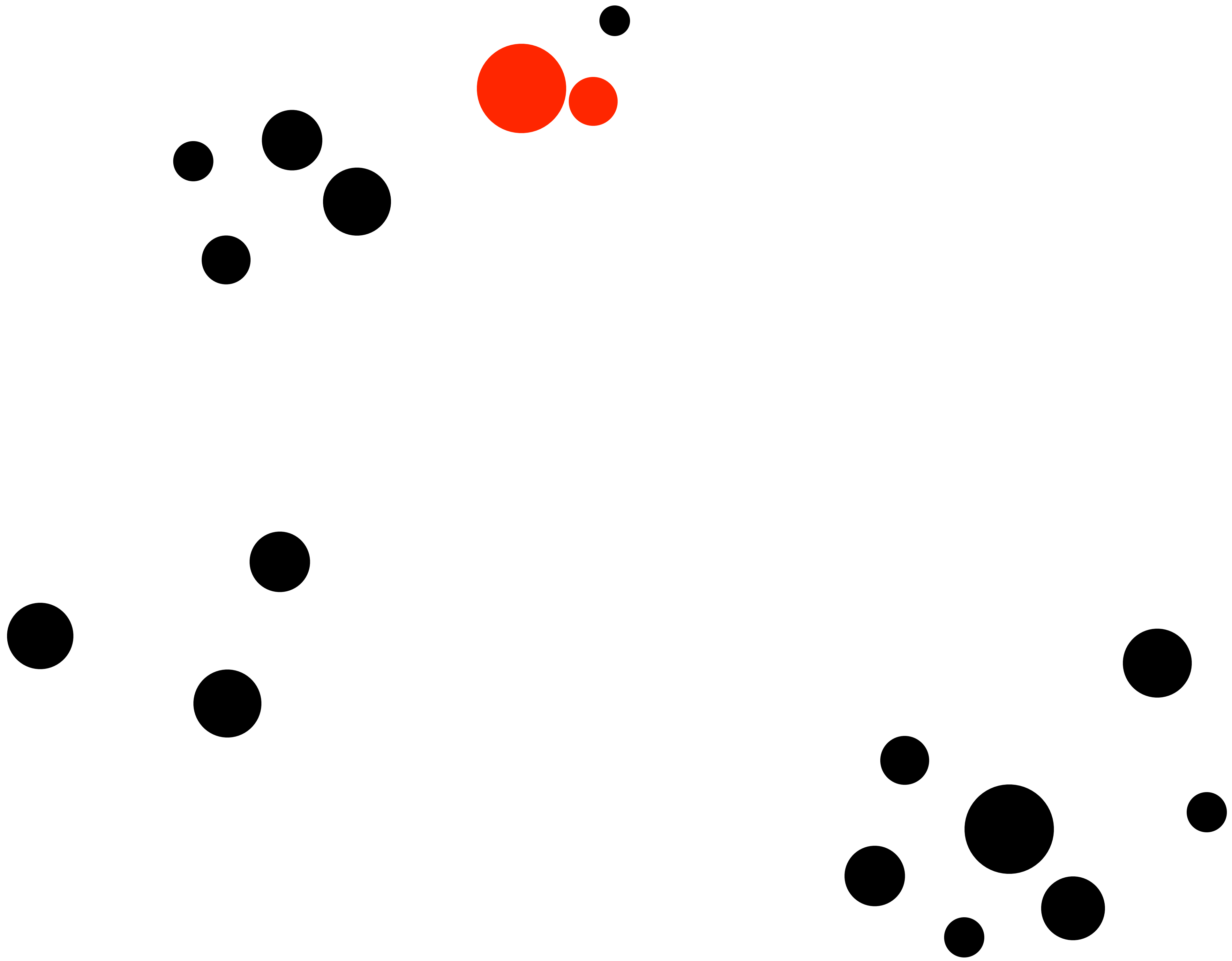
If they are closer than d_{ij} , combine their 4 vectors

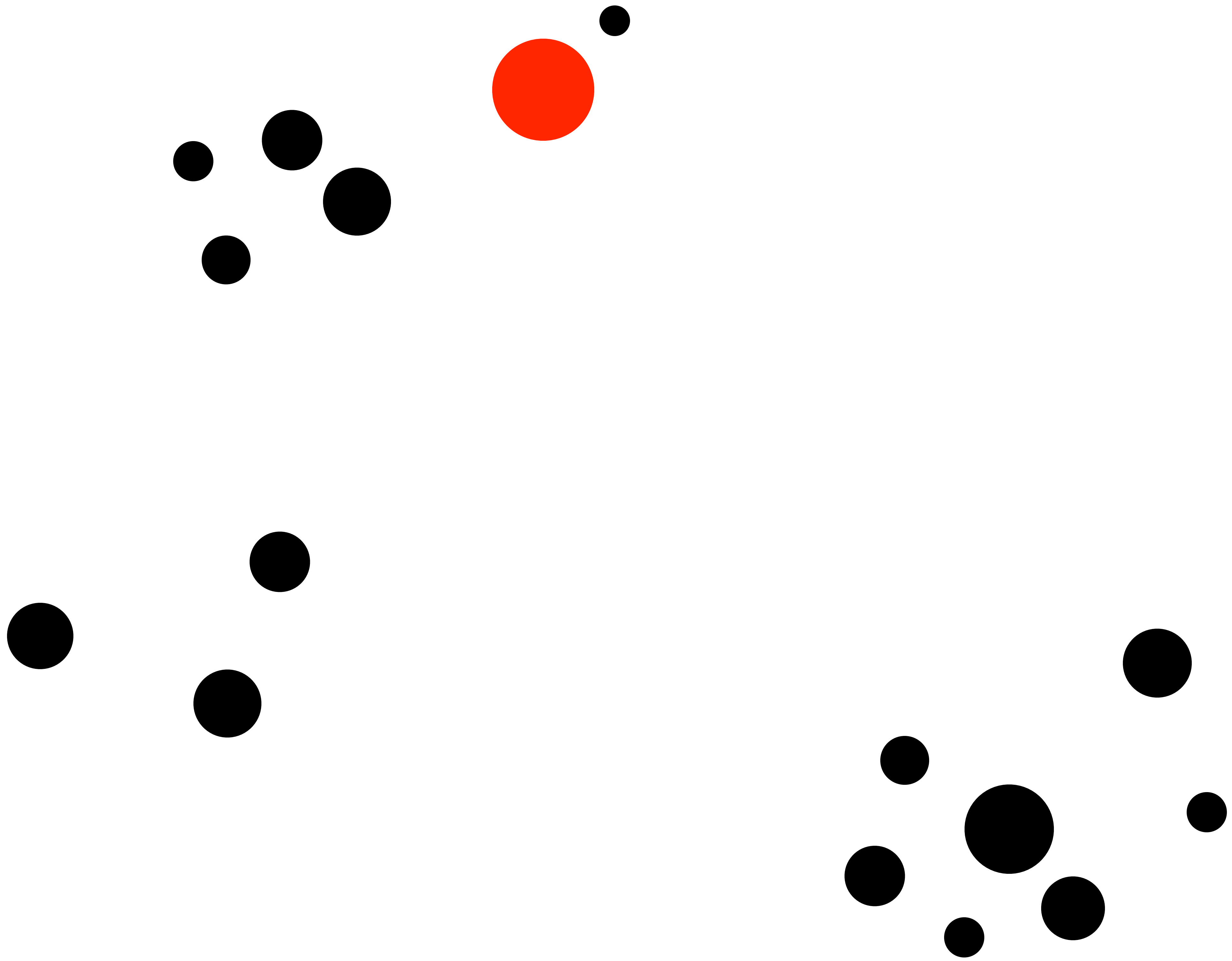


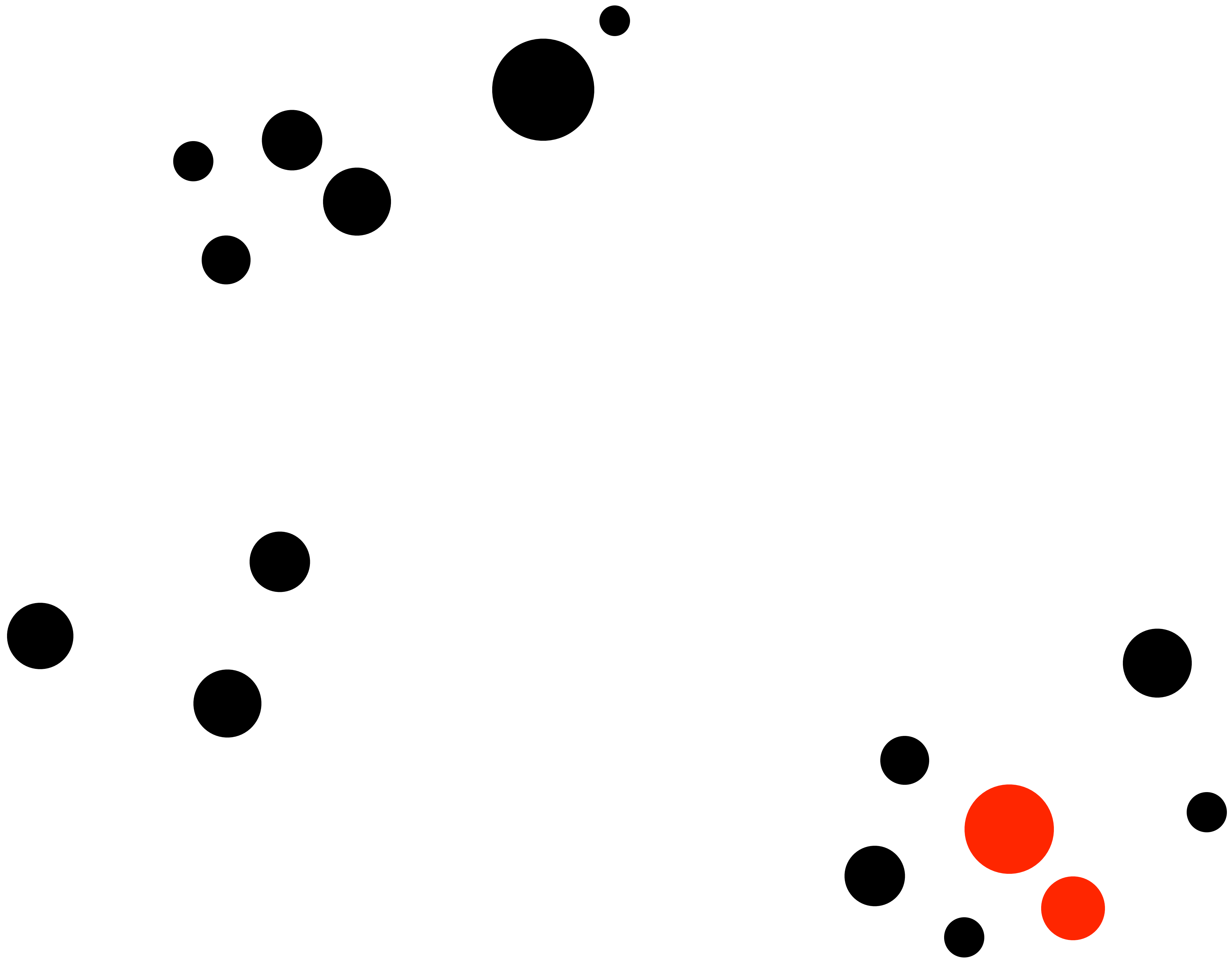
Repeat on the new closest pair

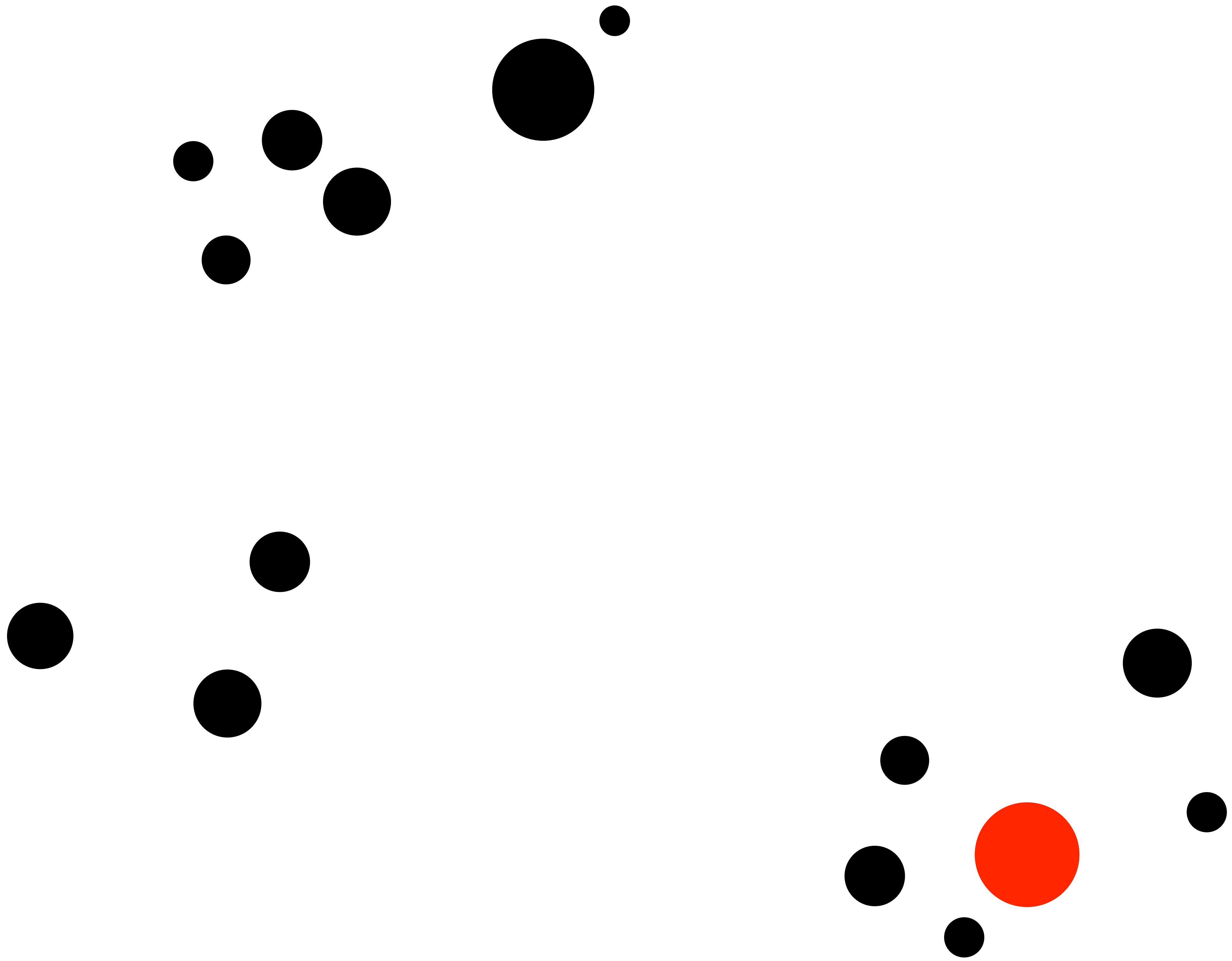


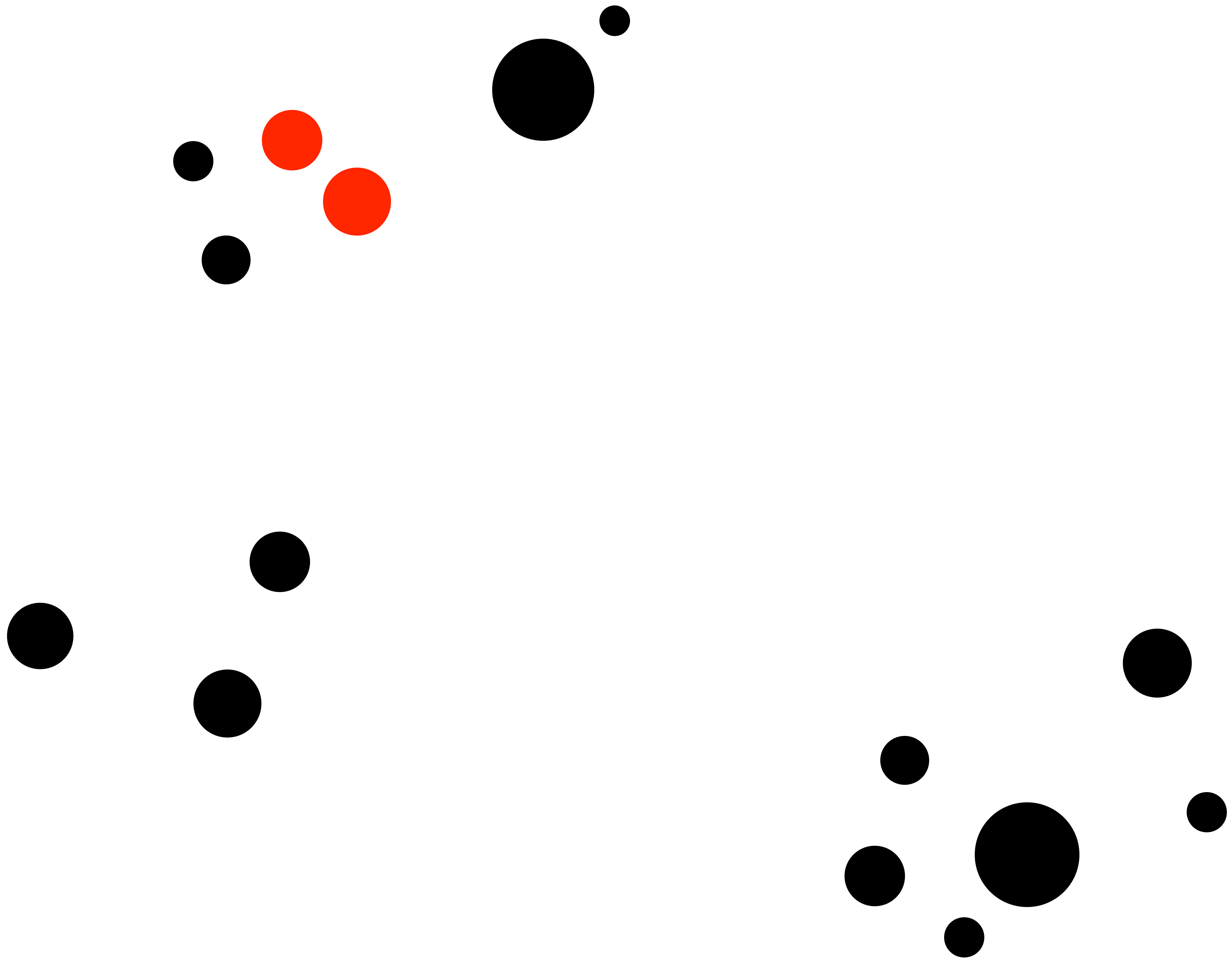


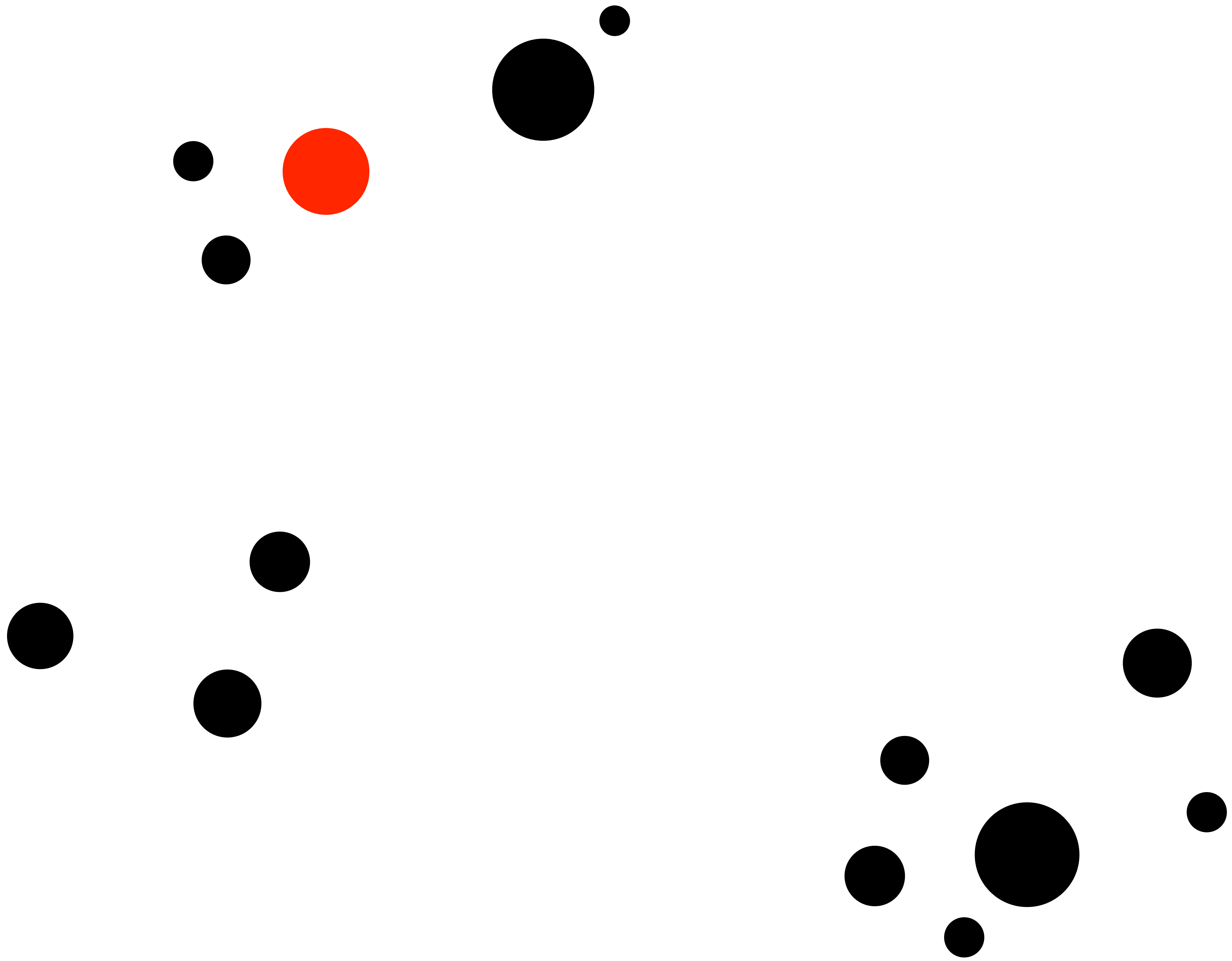


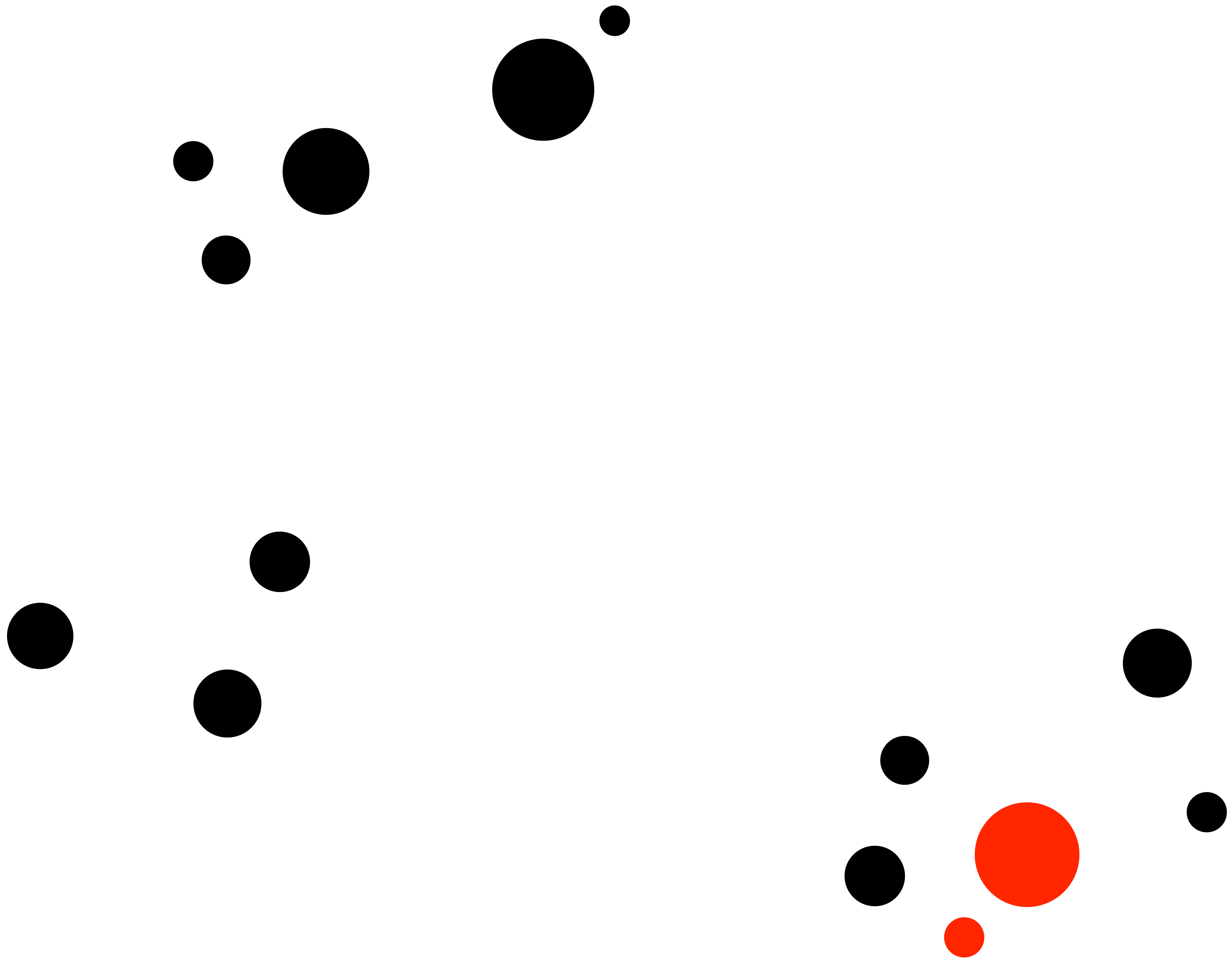


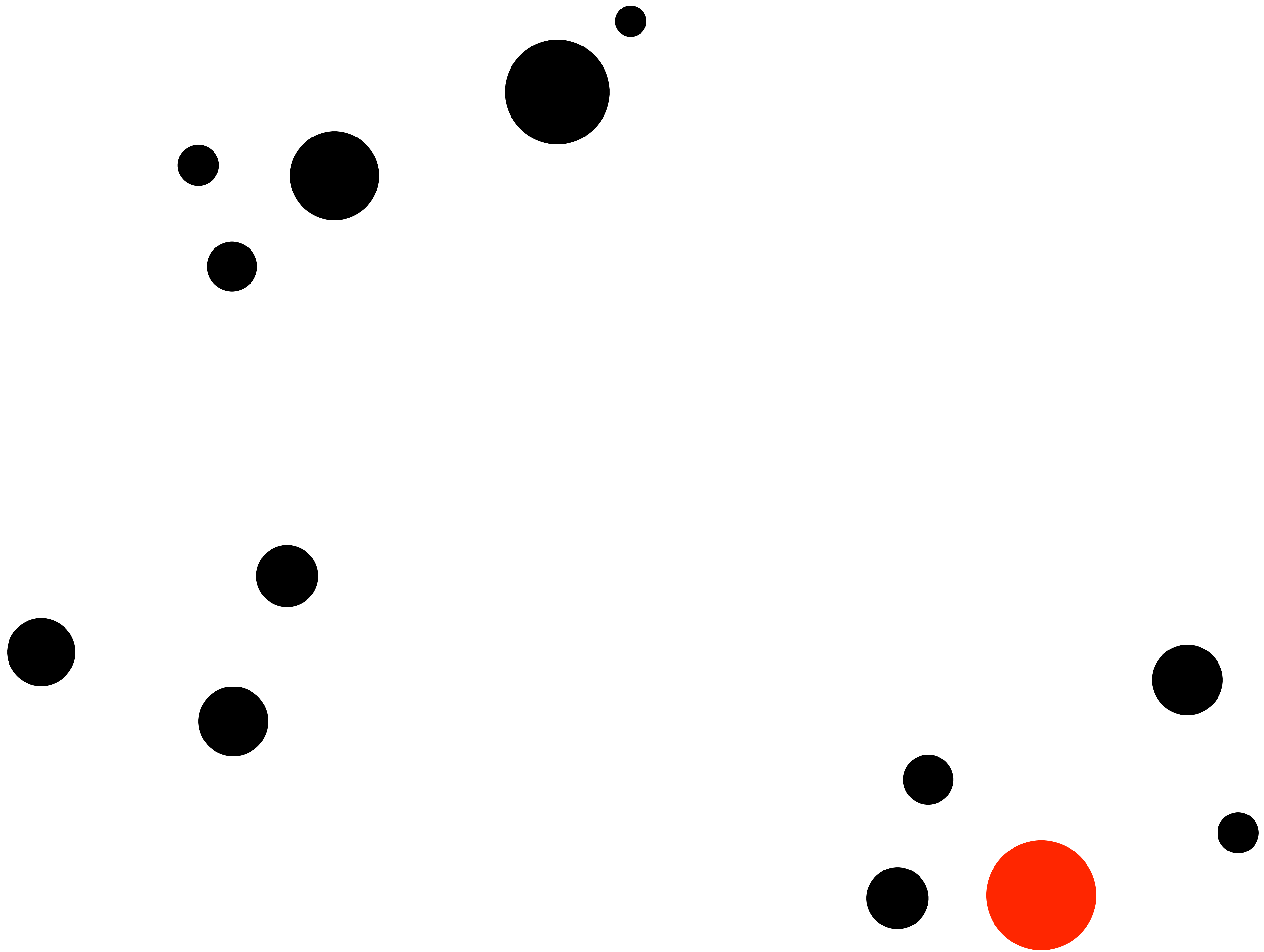


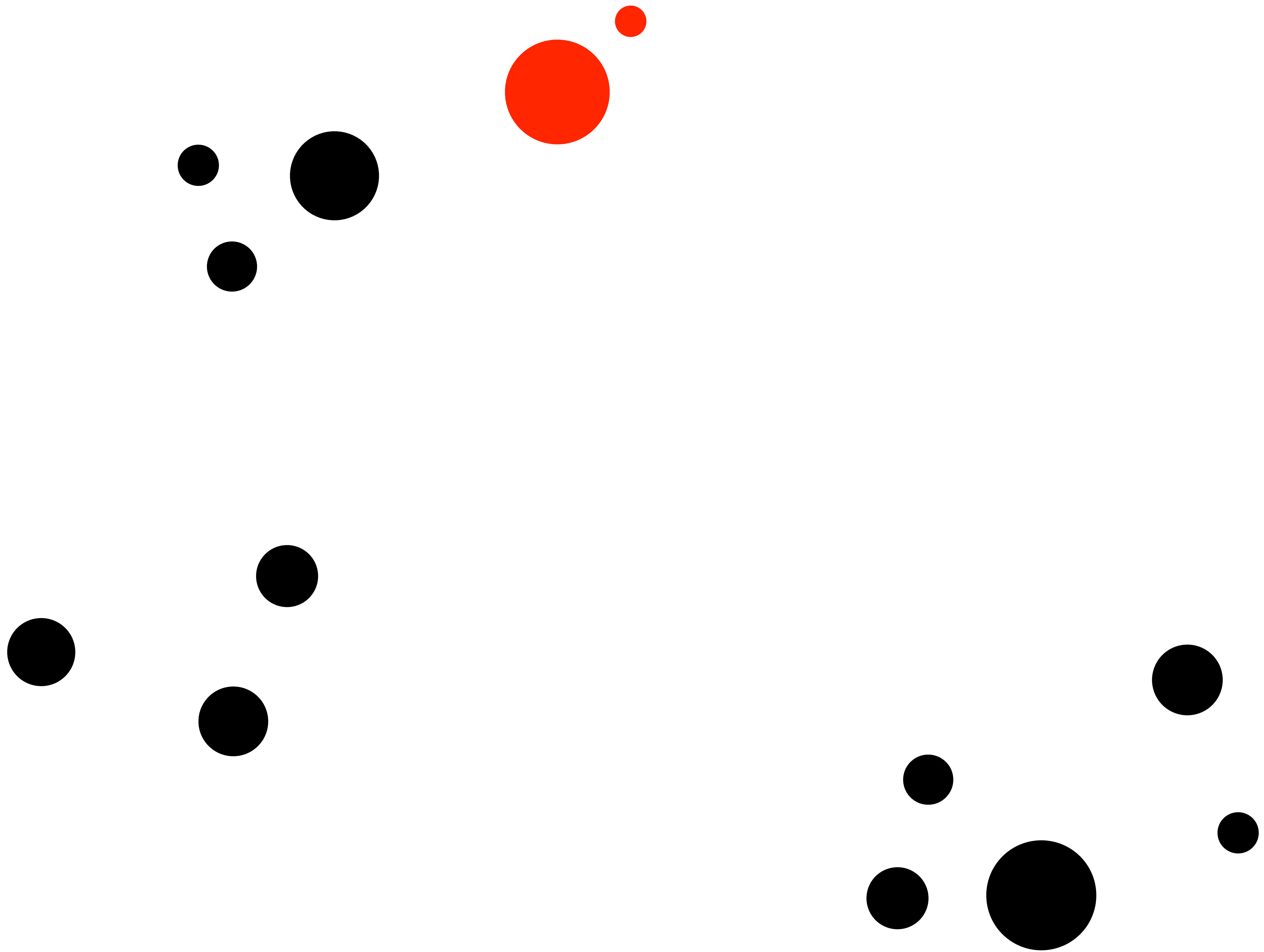


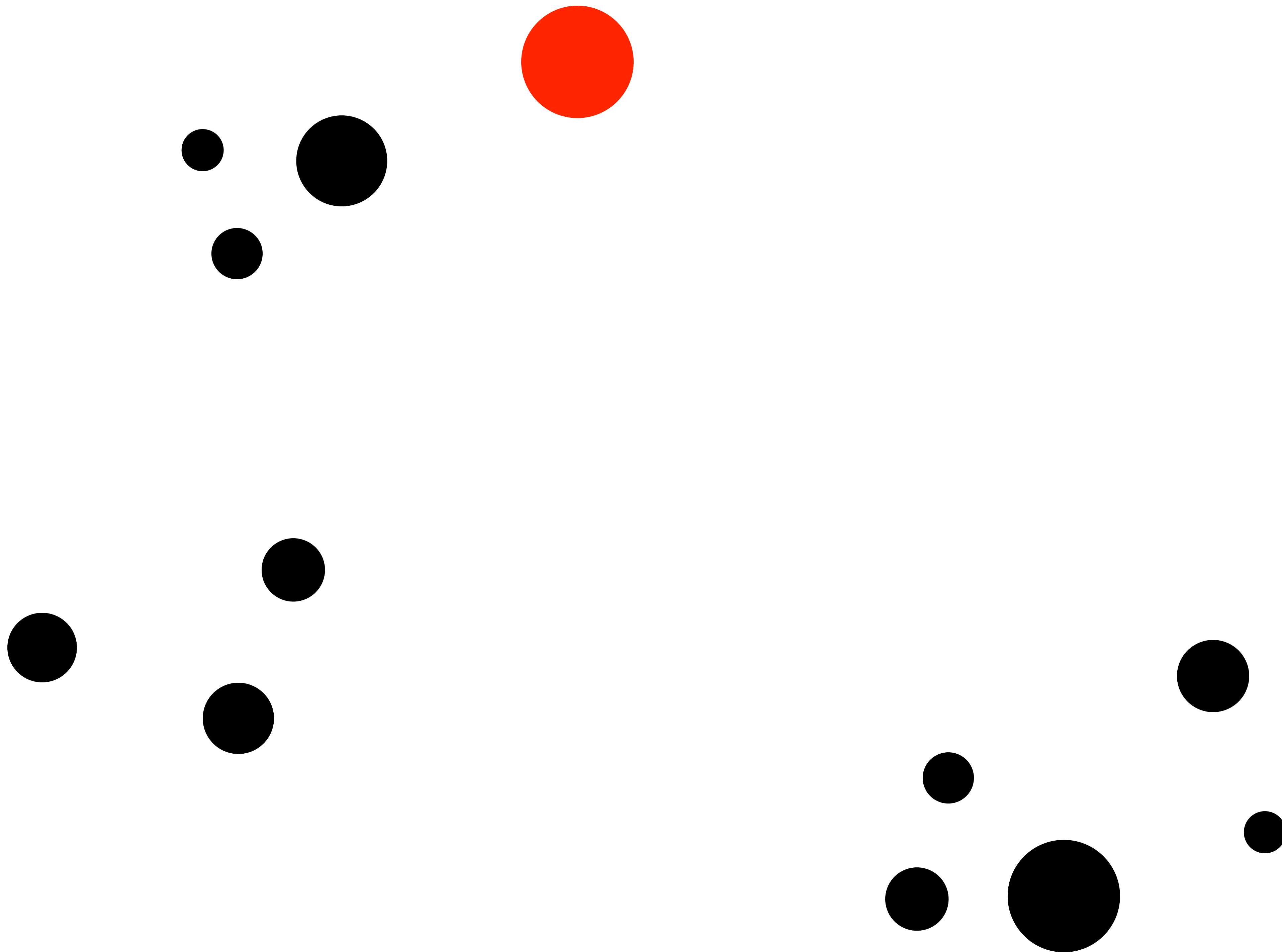


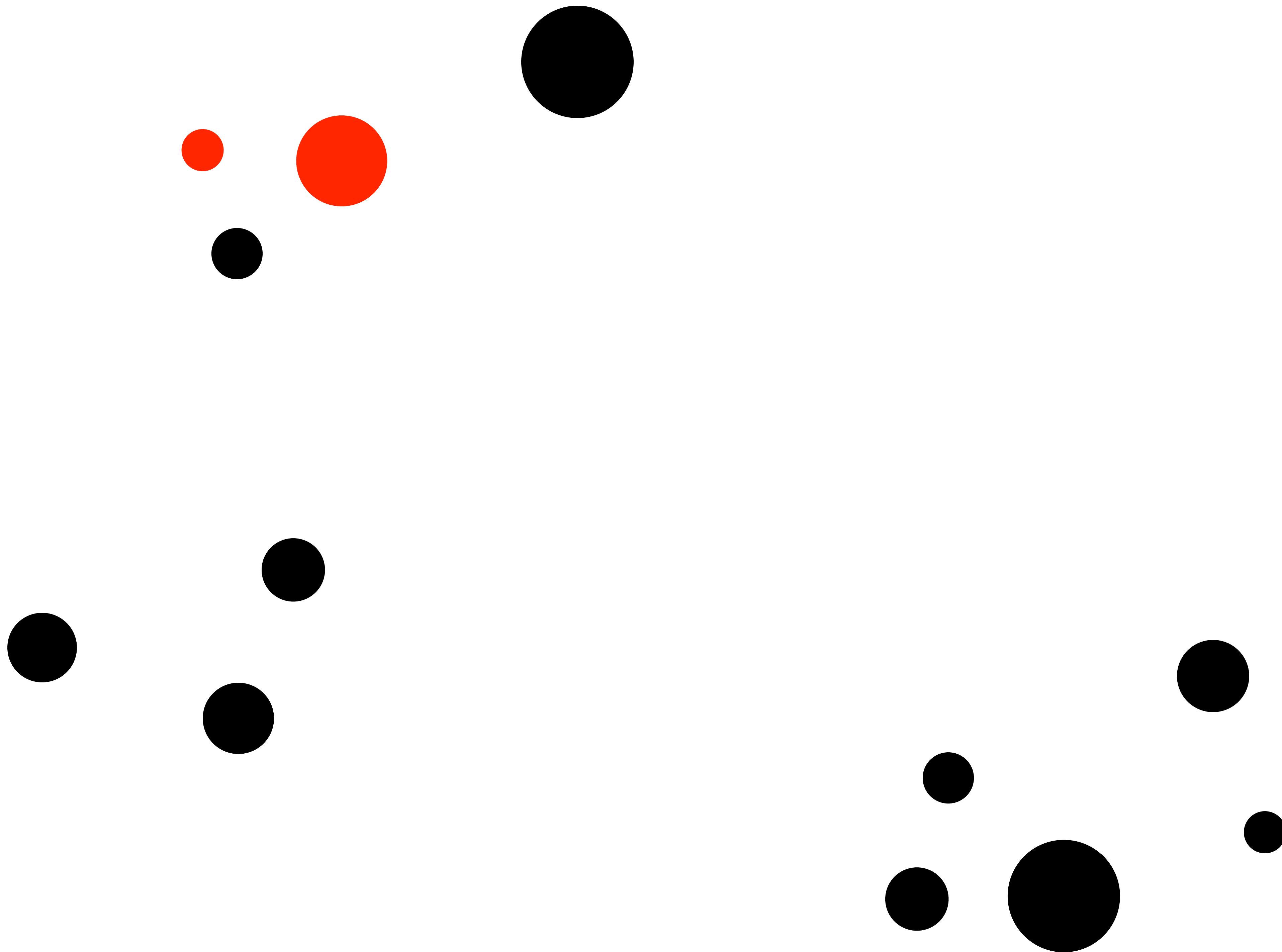


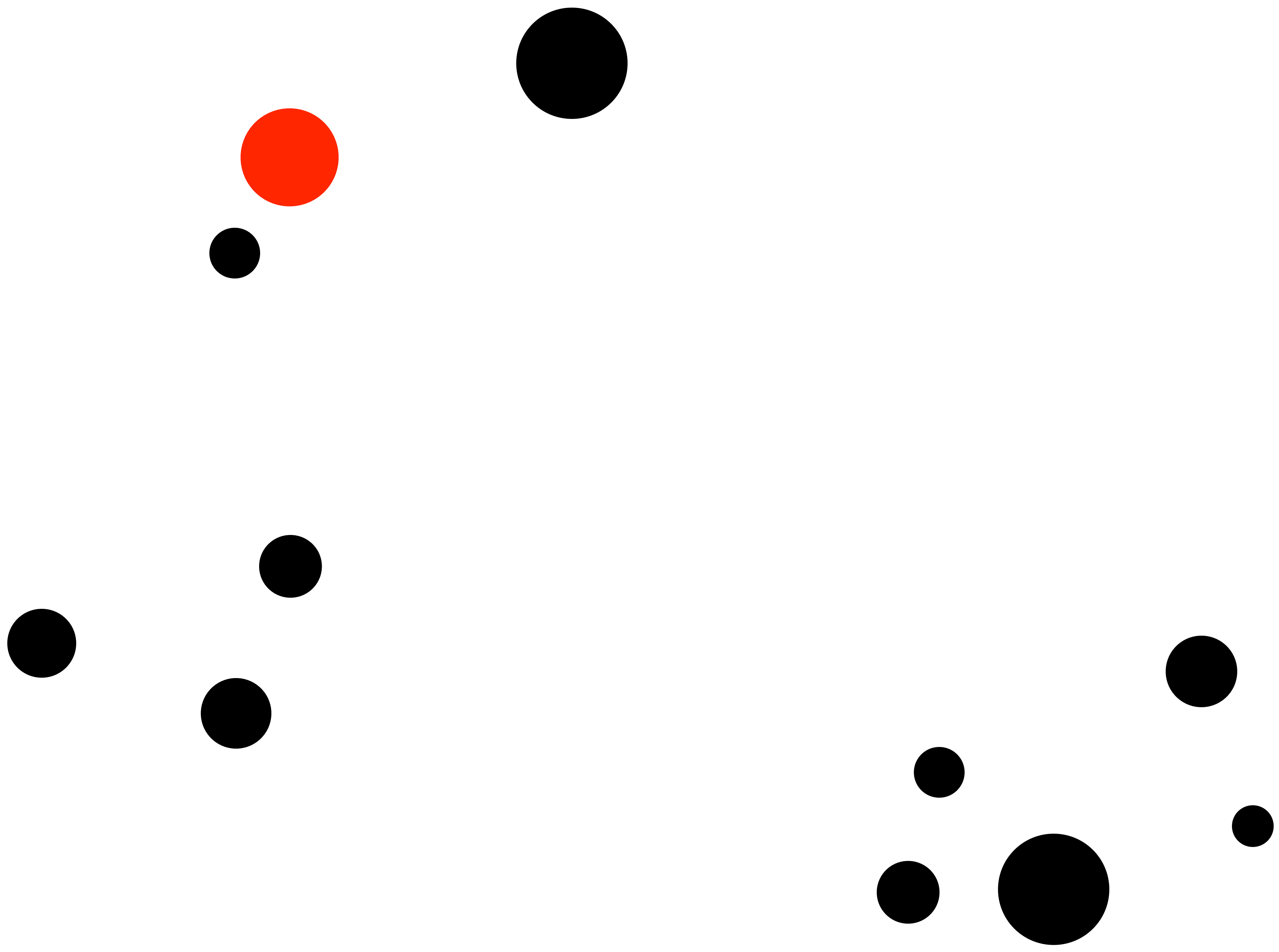


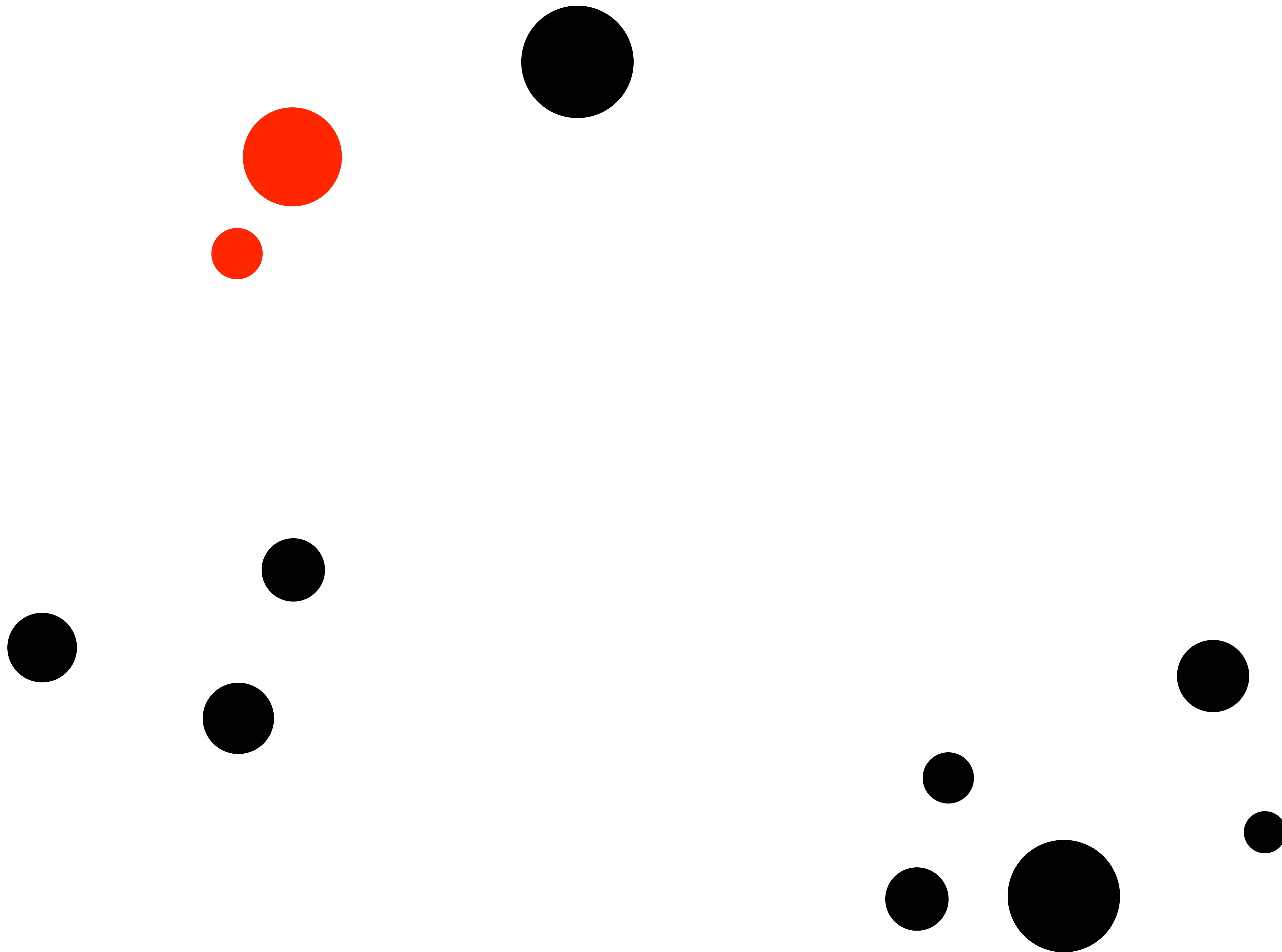


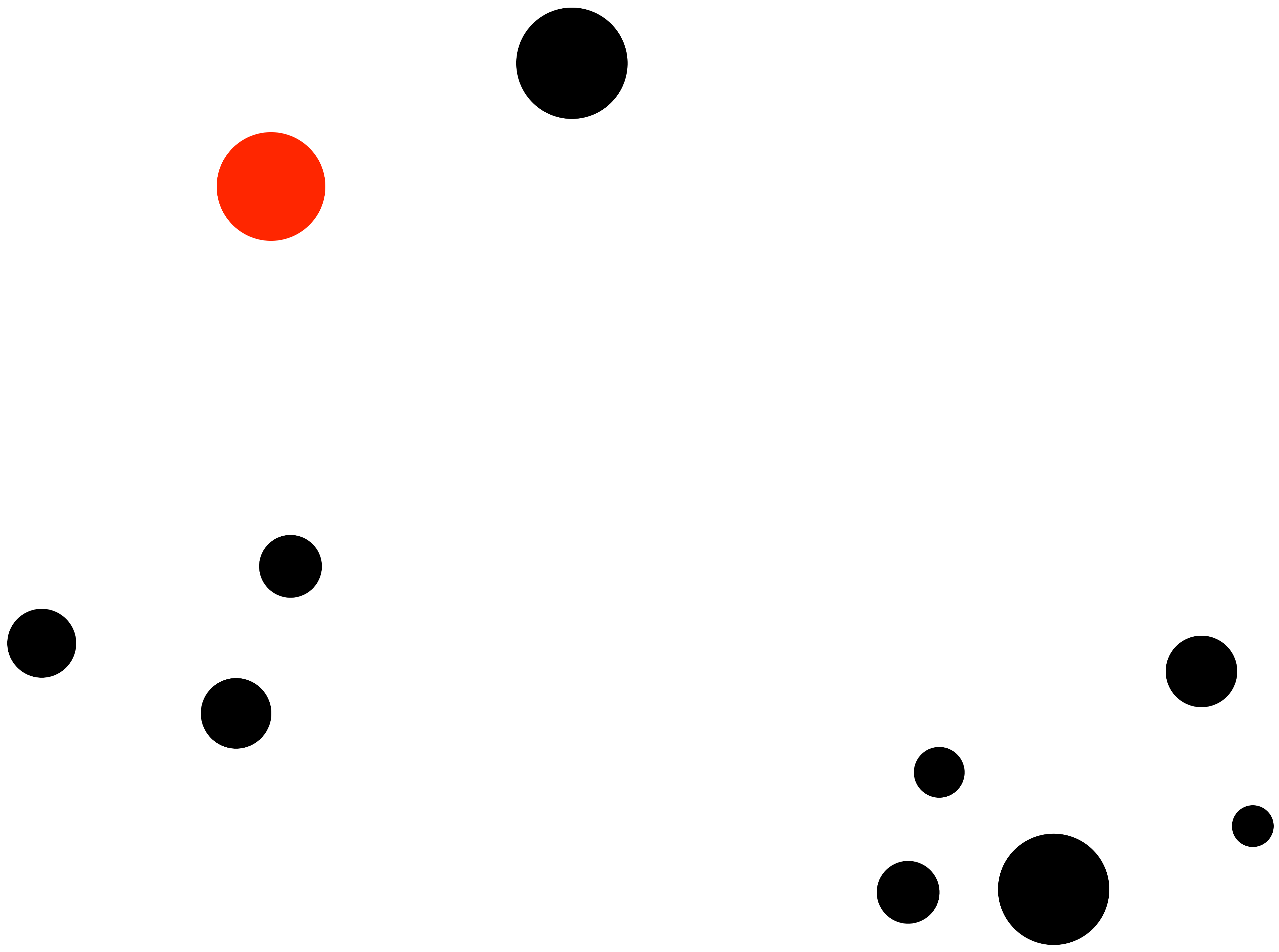


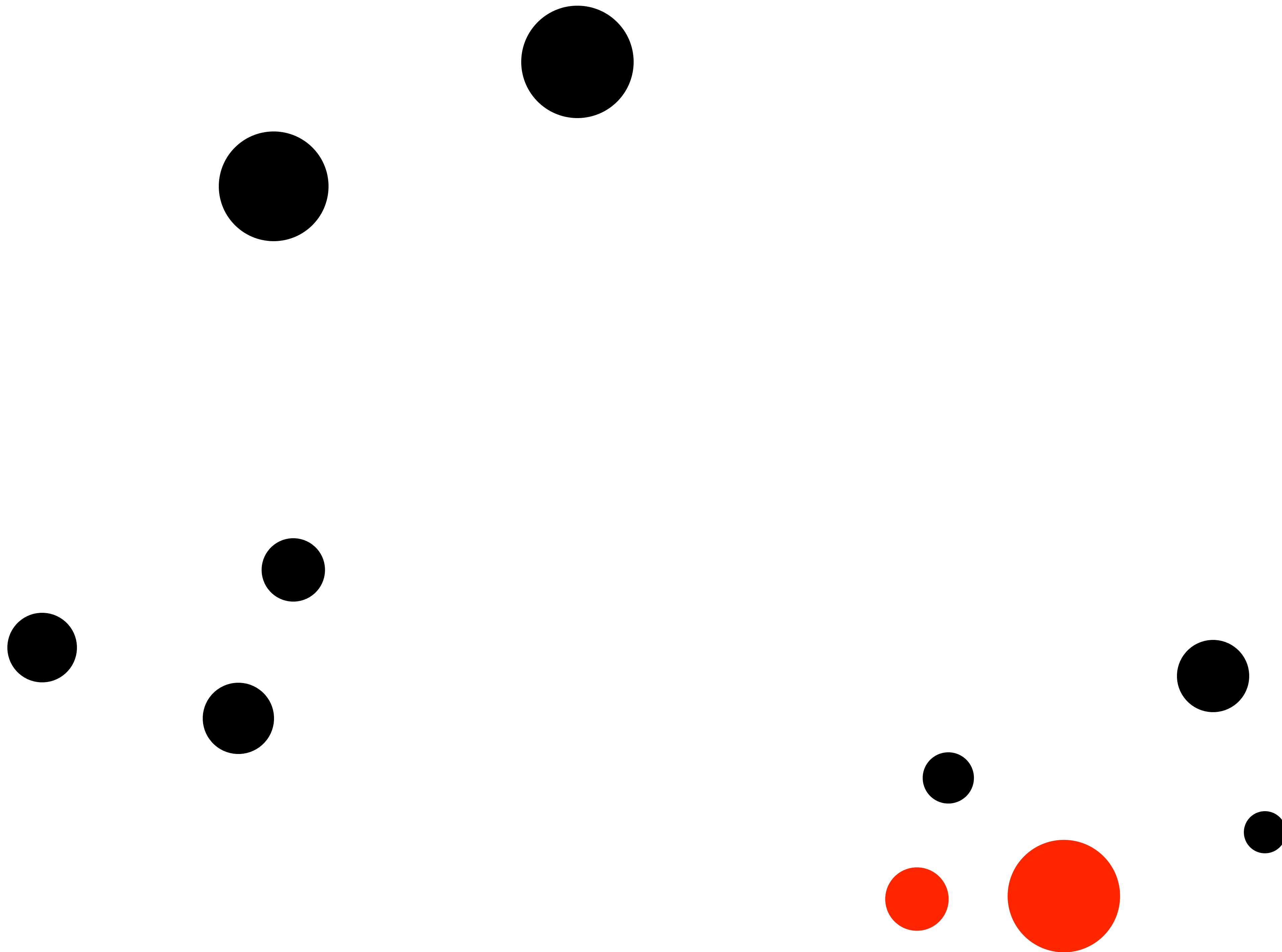


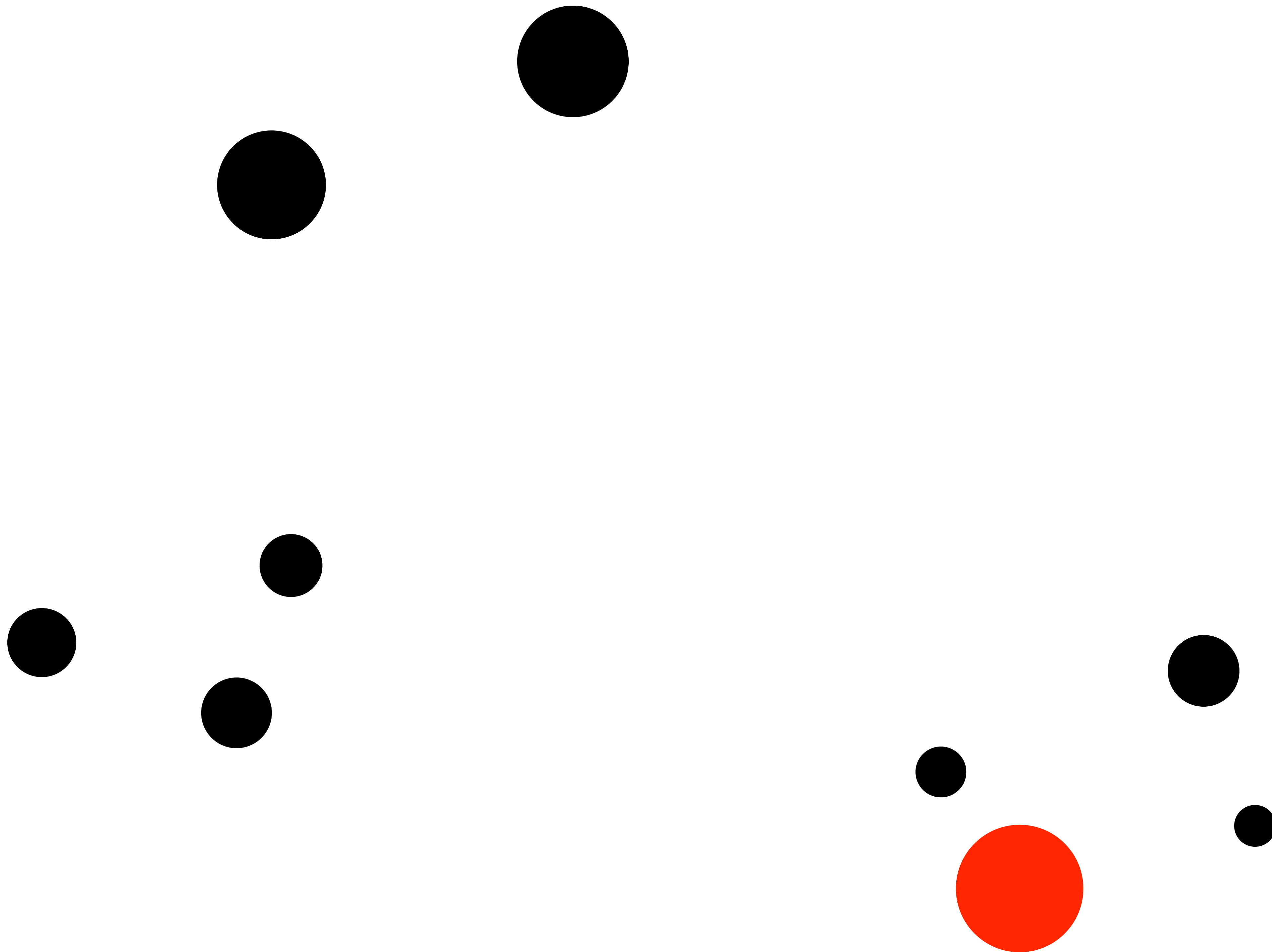


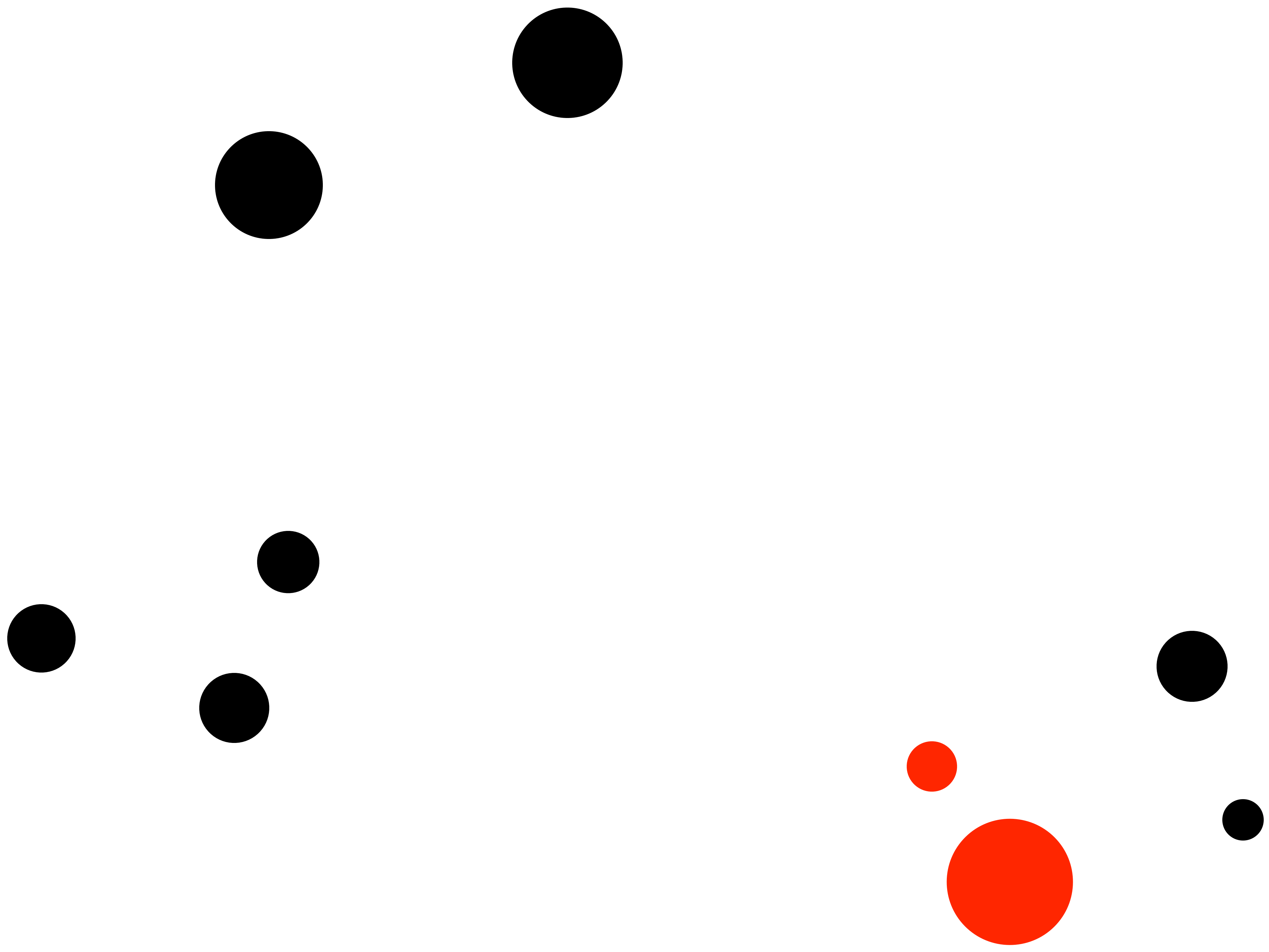


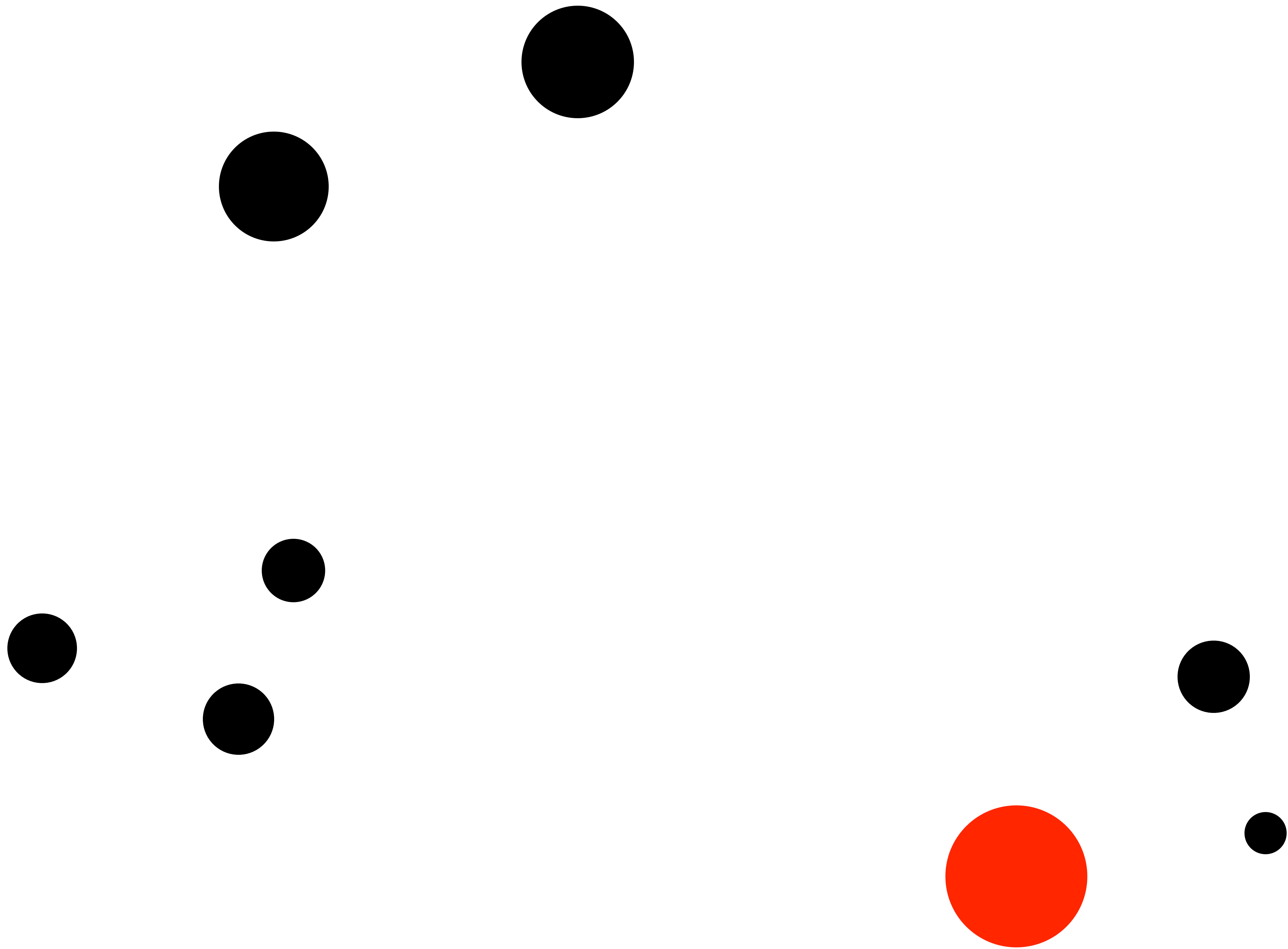


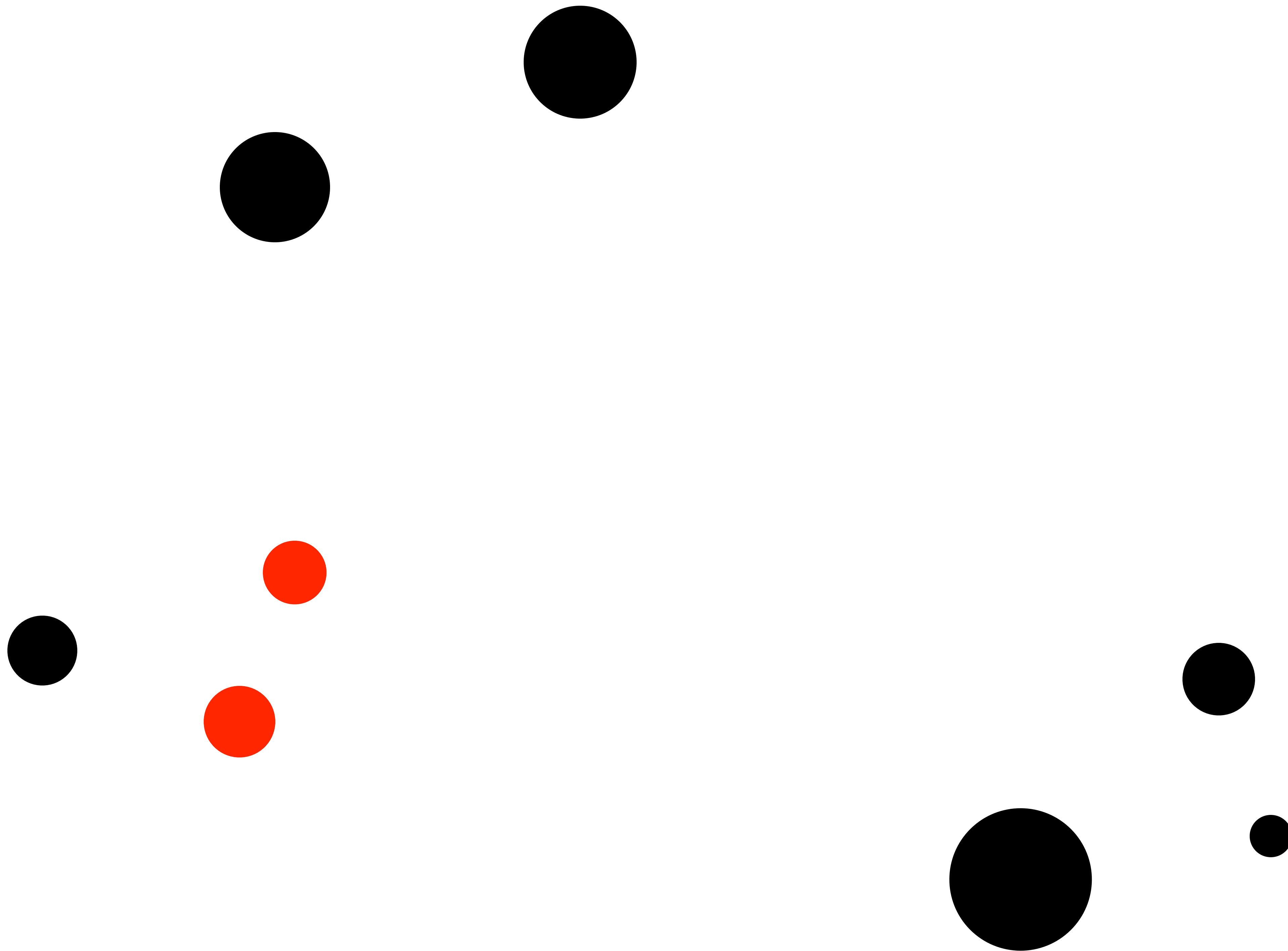


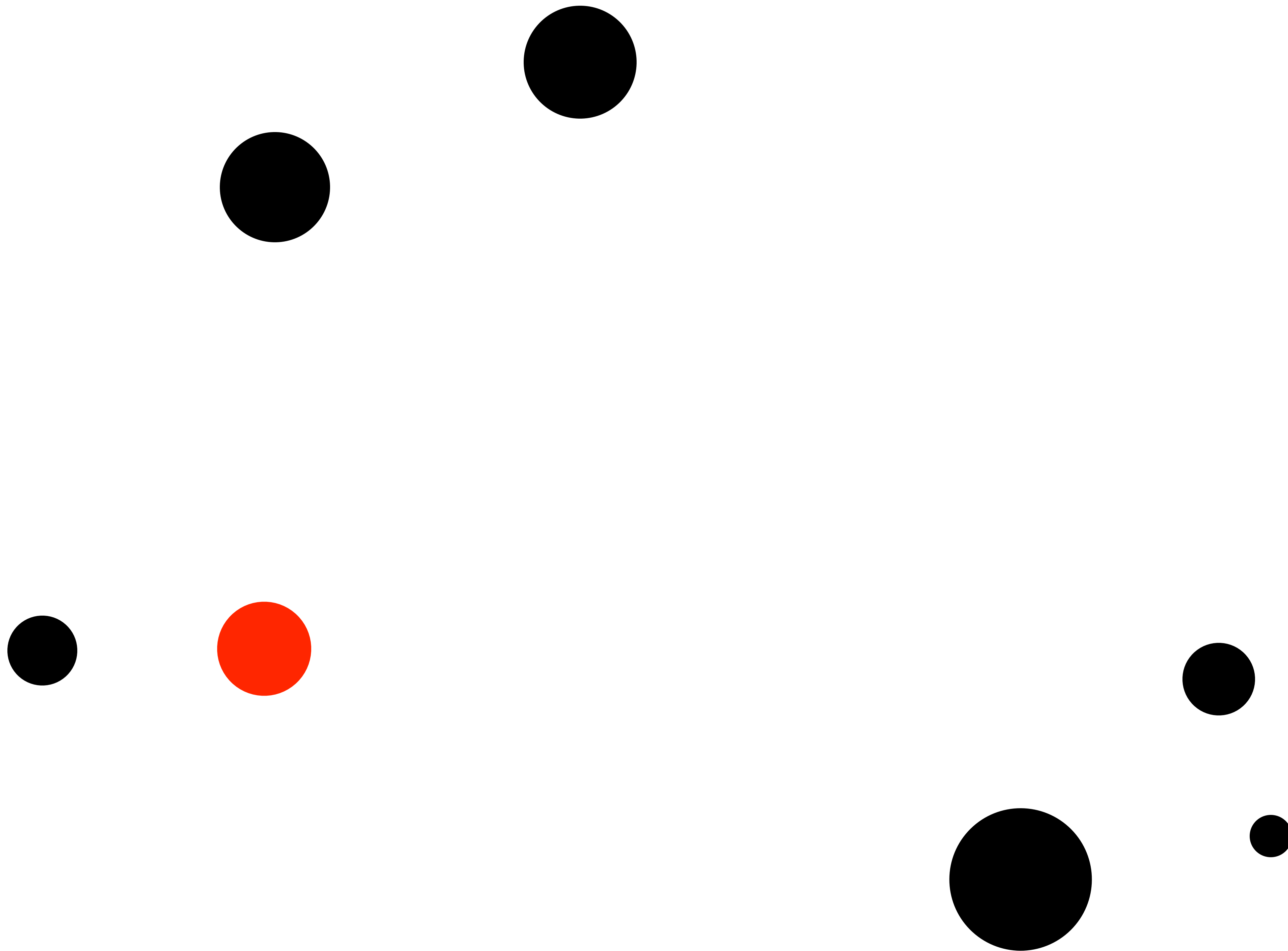


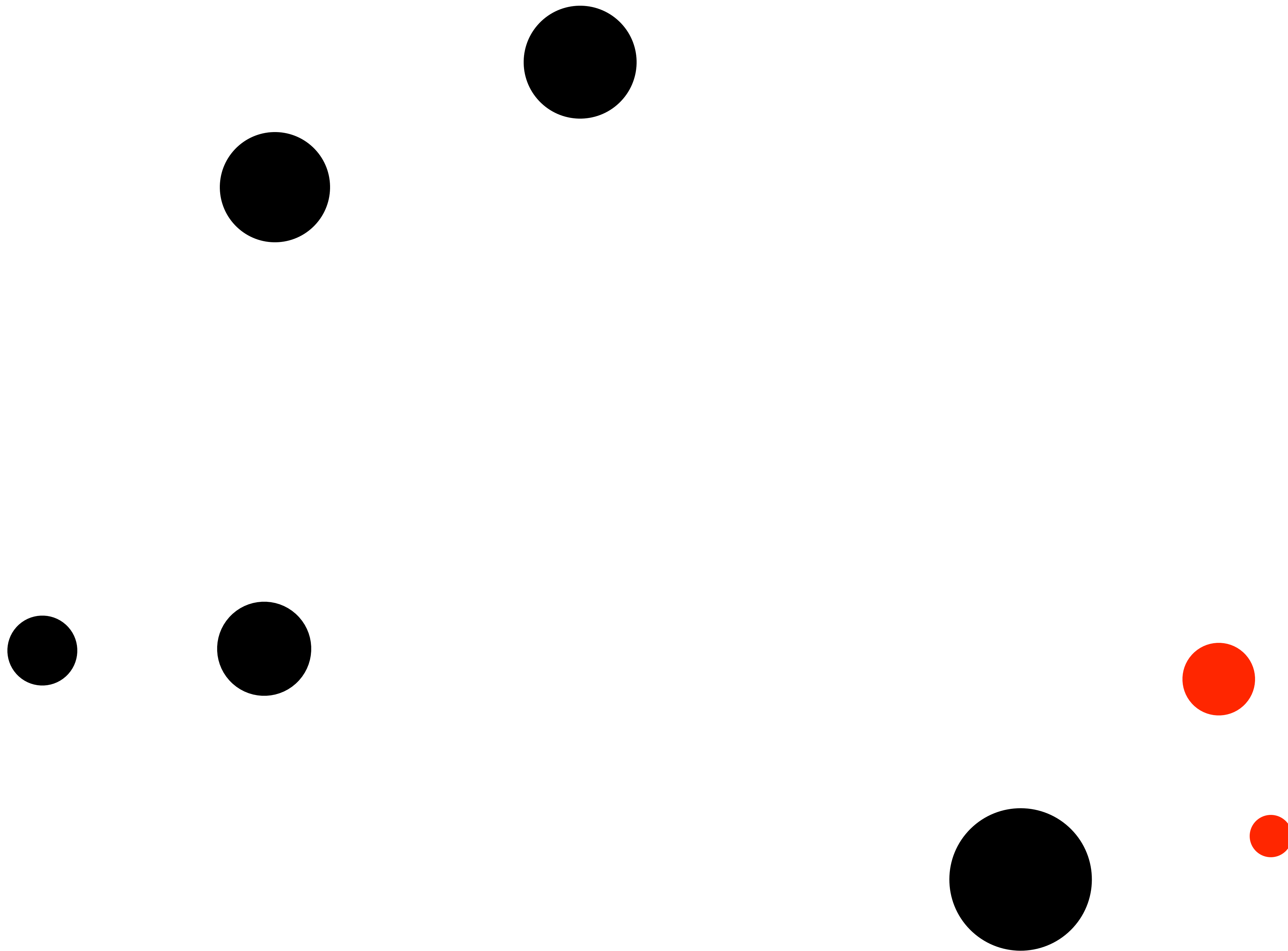


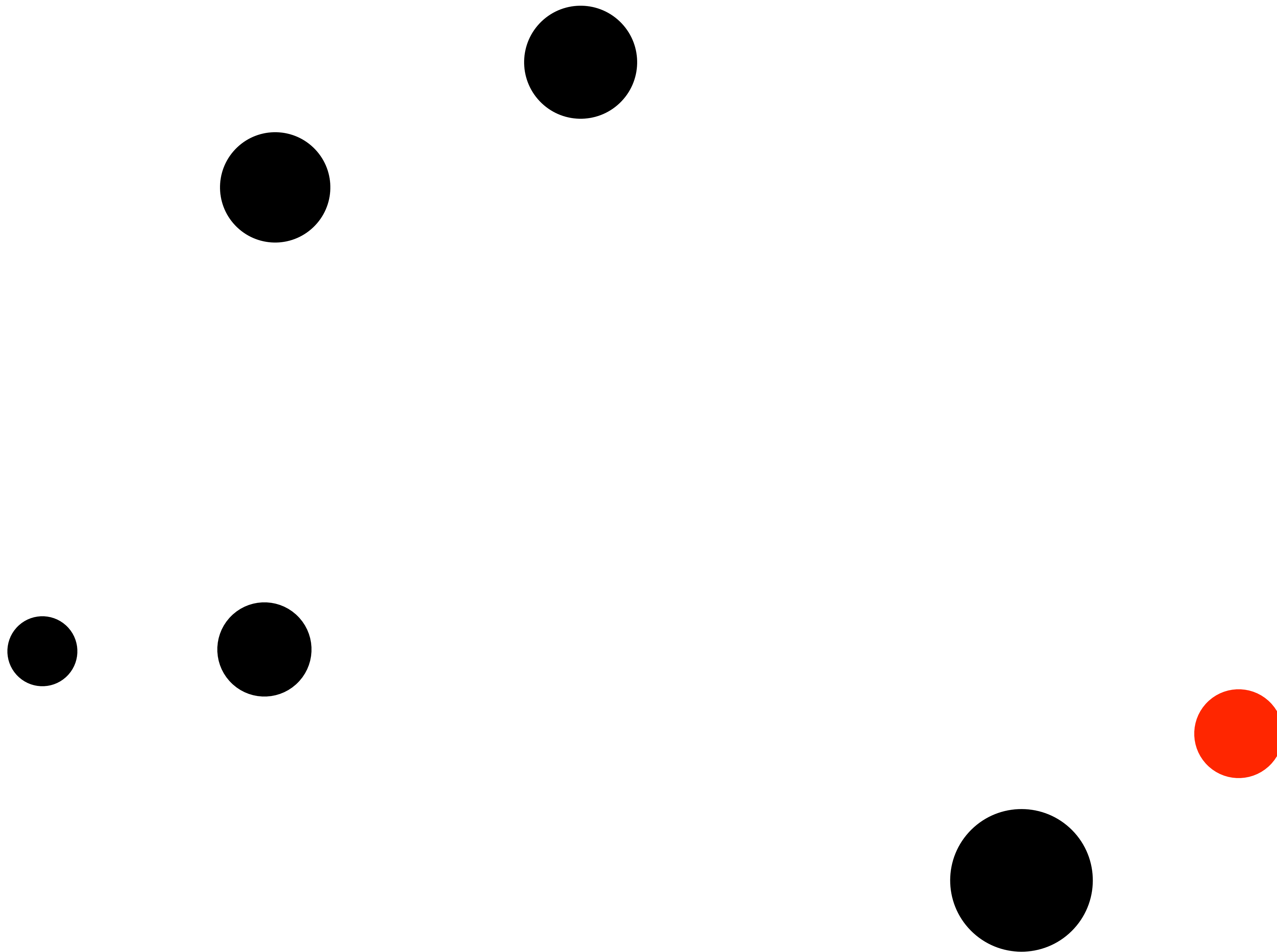


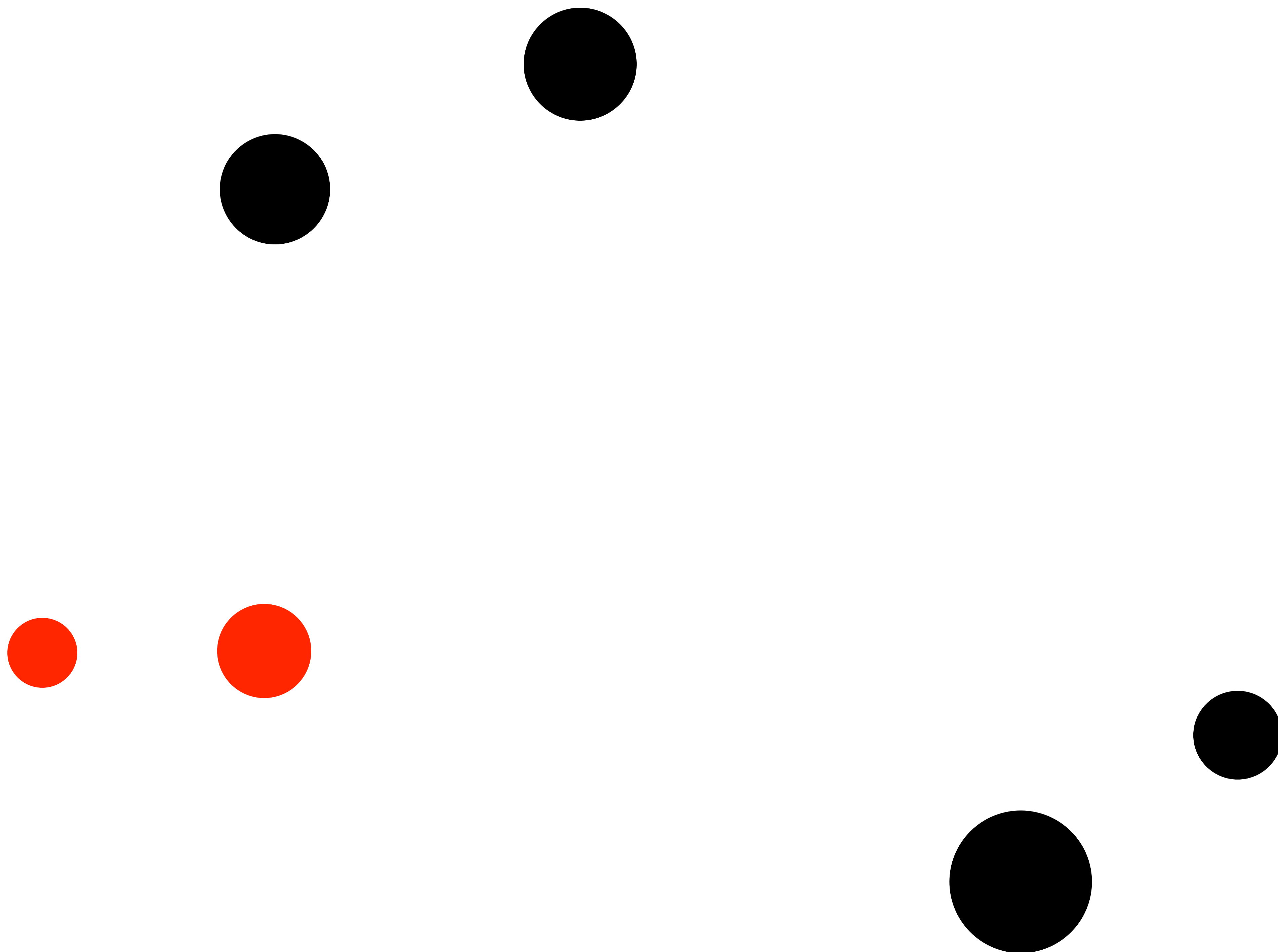


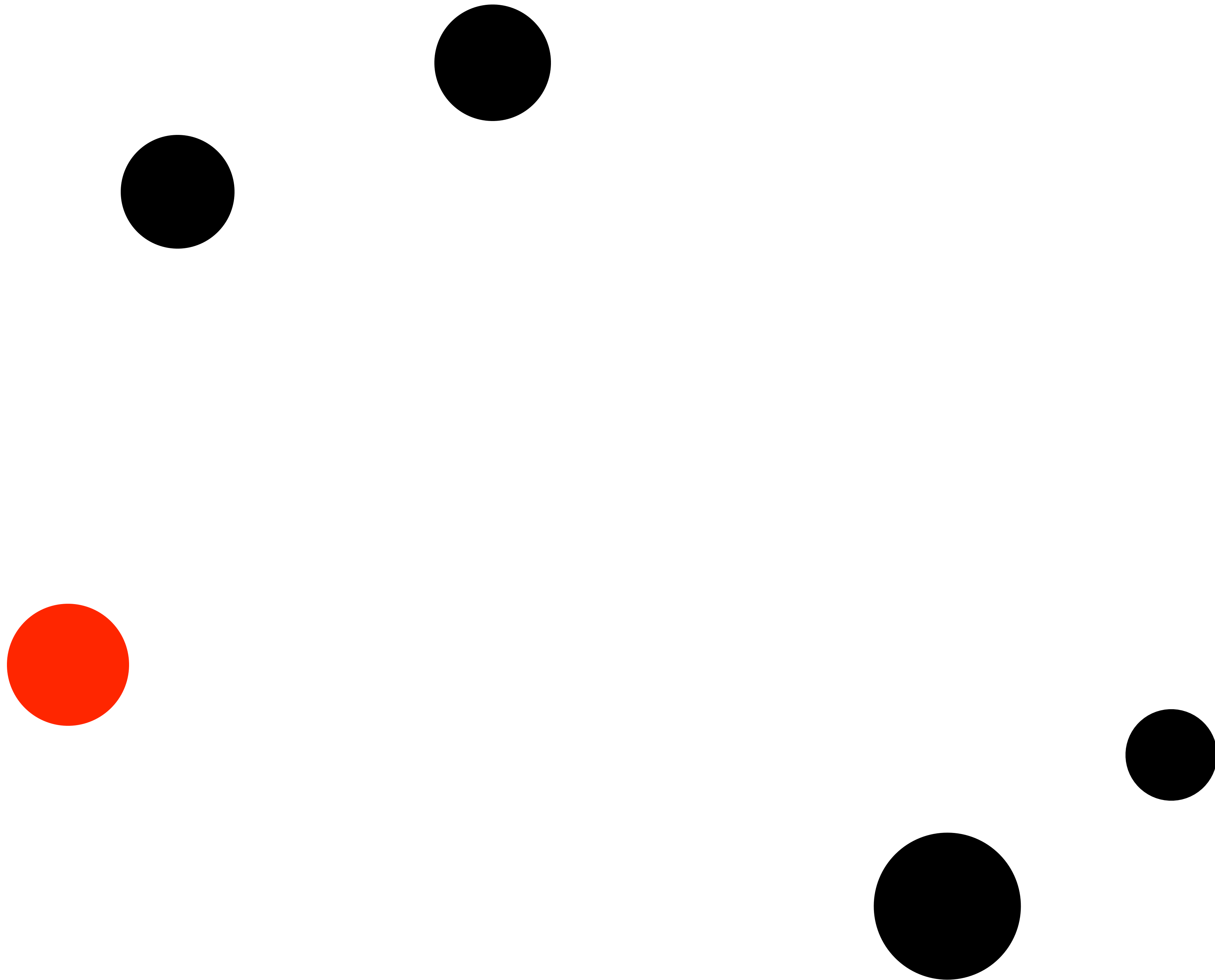


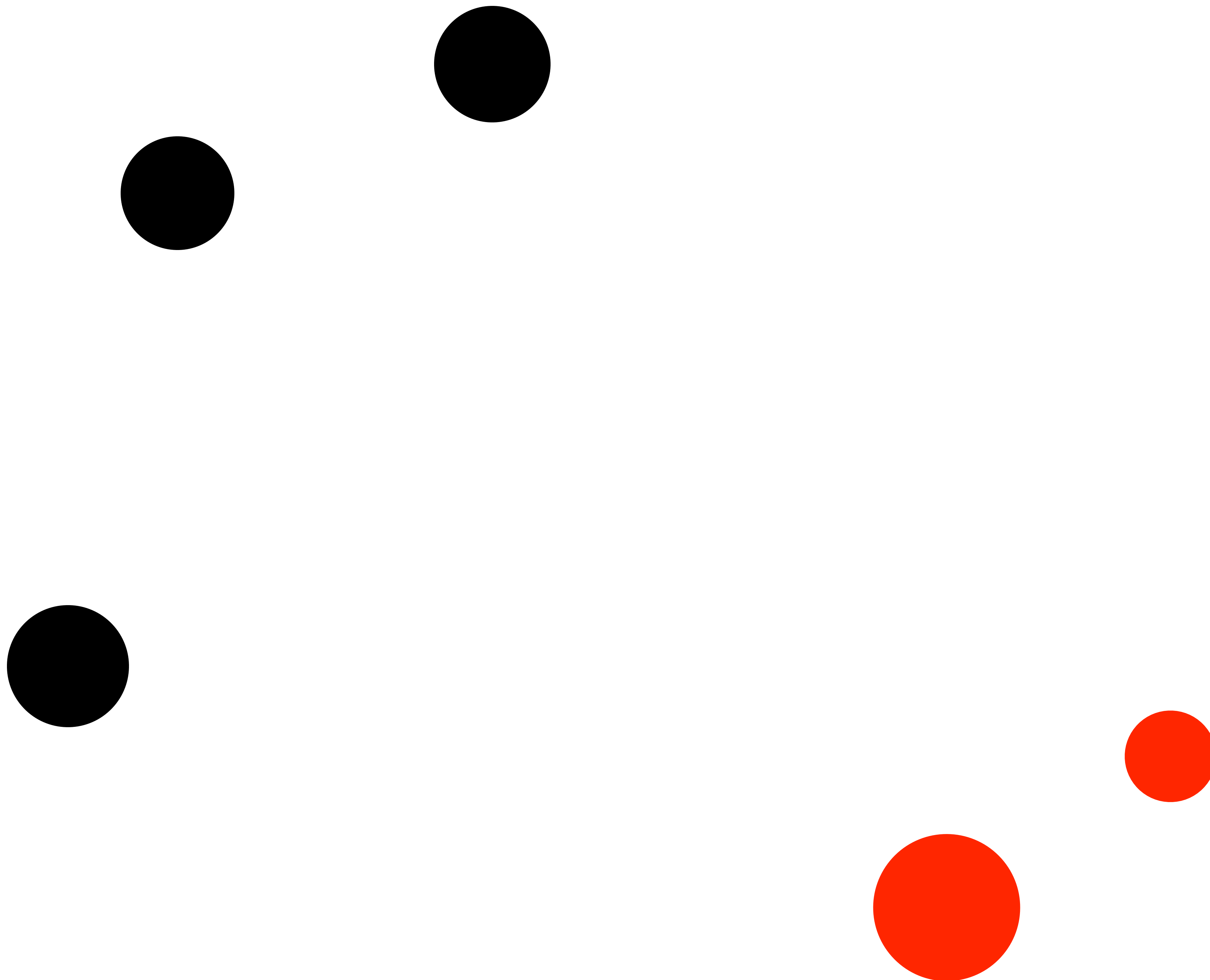


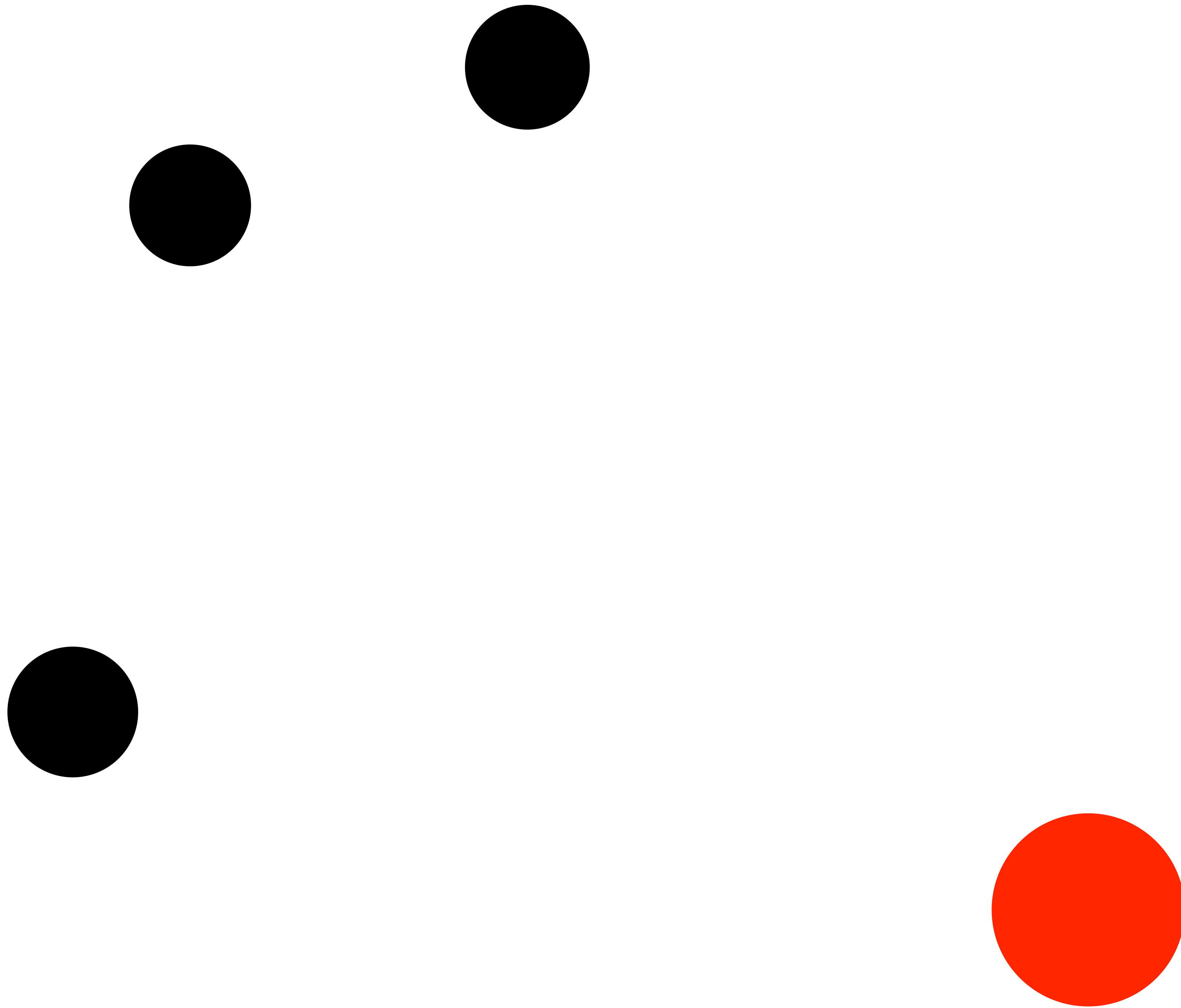


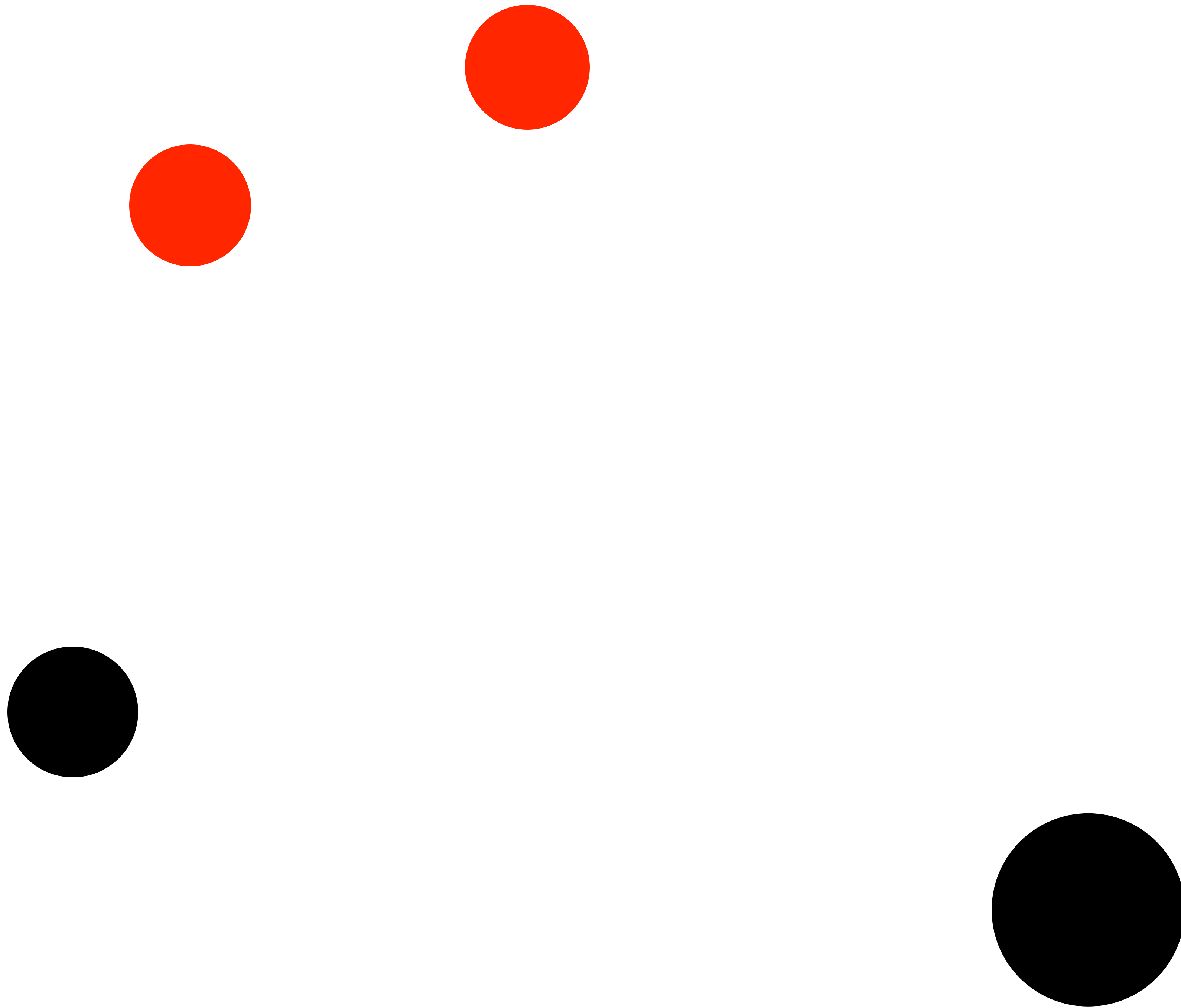


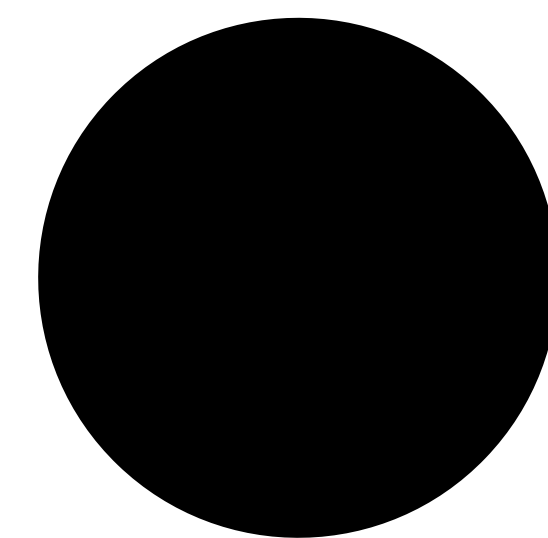
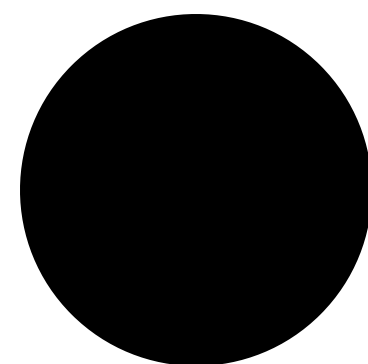
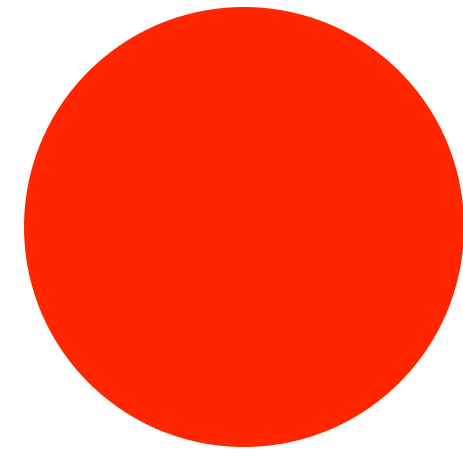






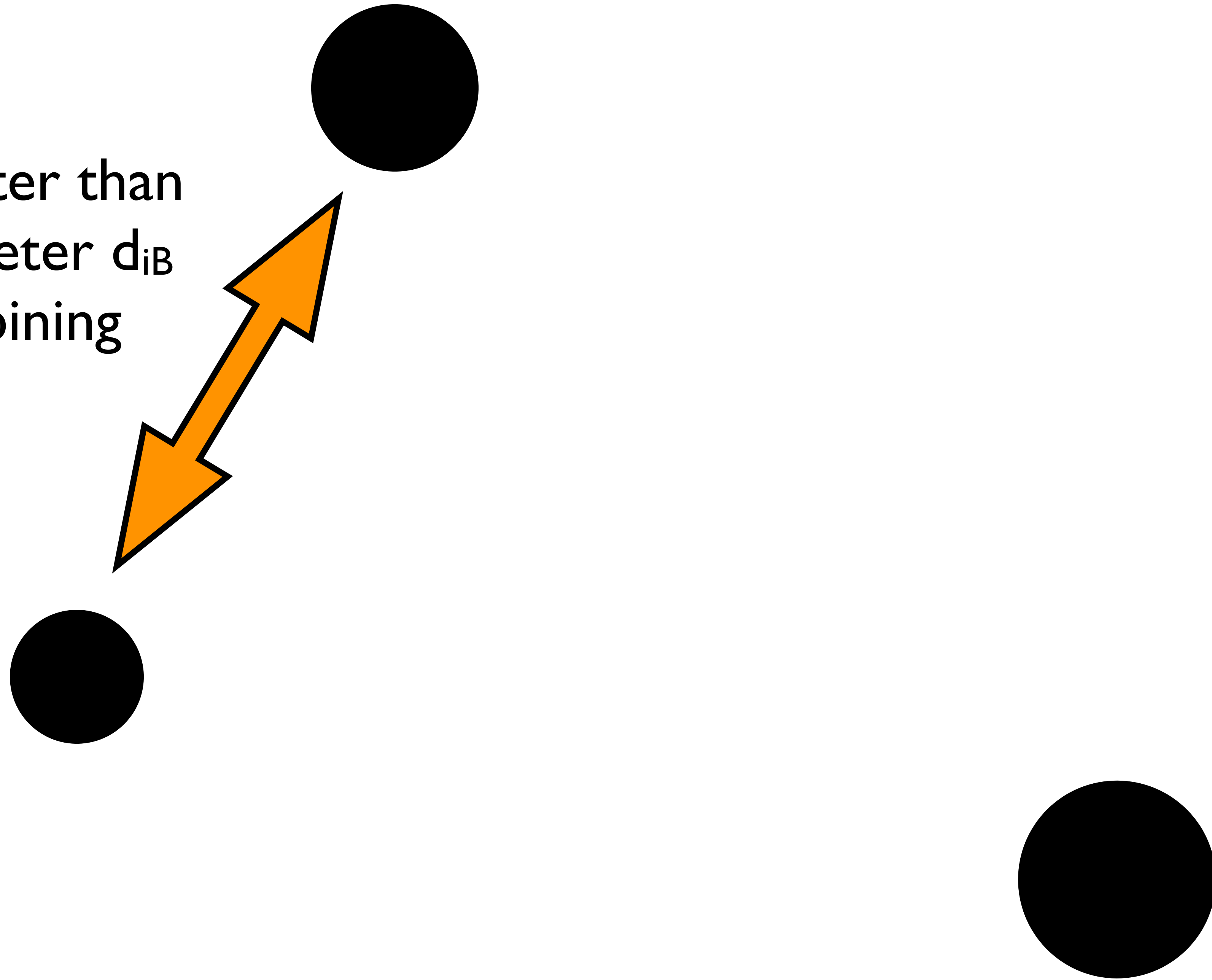




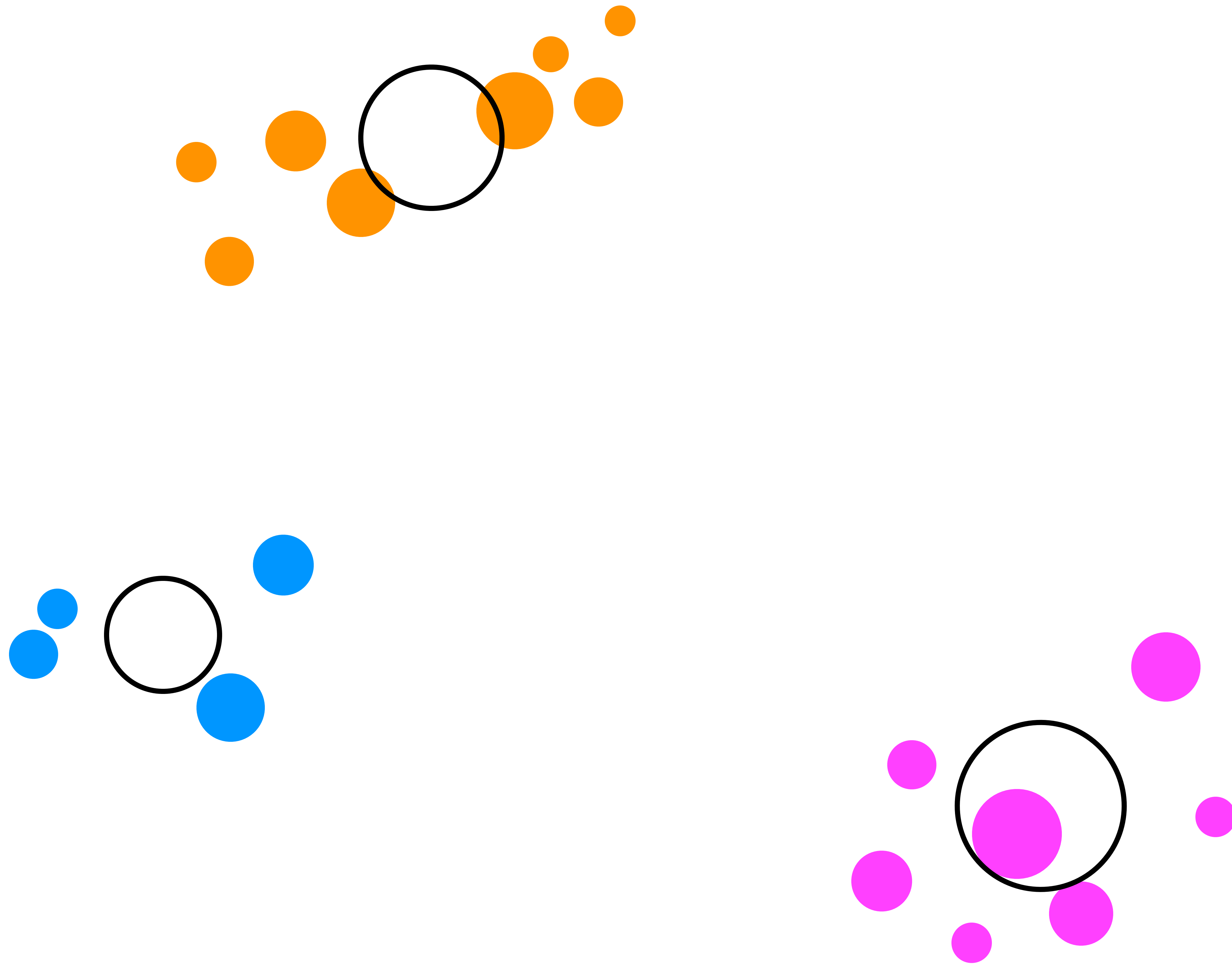


Stop when the closest pair is separated by $\Delta R > R$

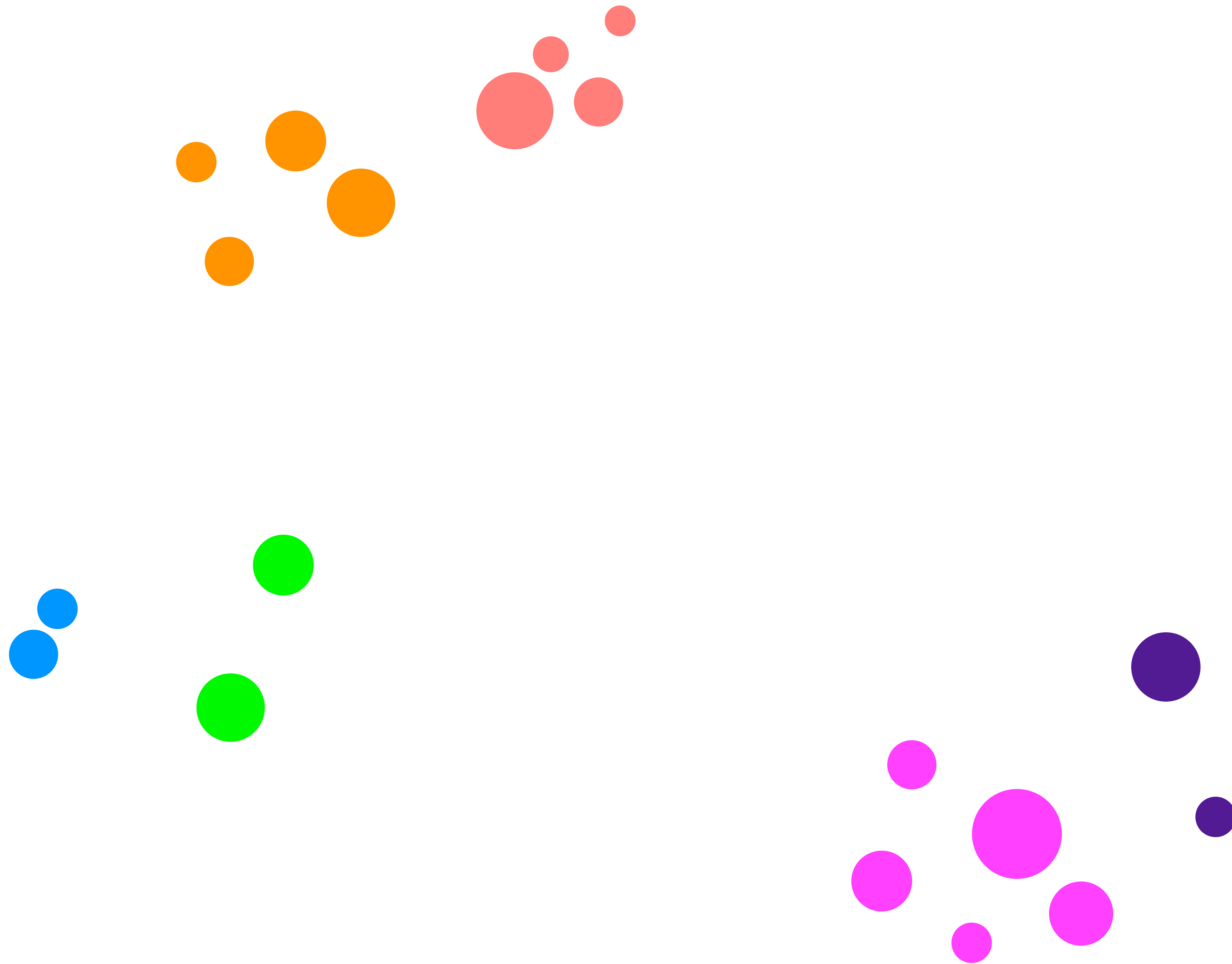
separation greater than
distance parameter d_{iB}
 \Rightarrow stop combining

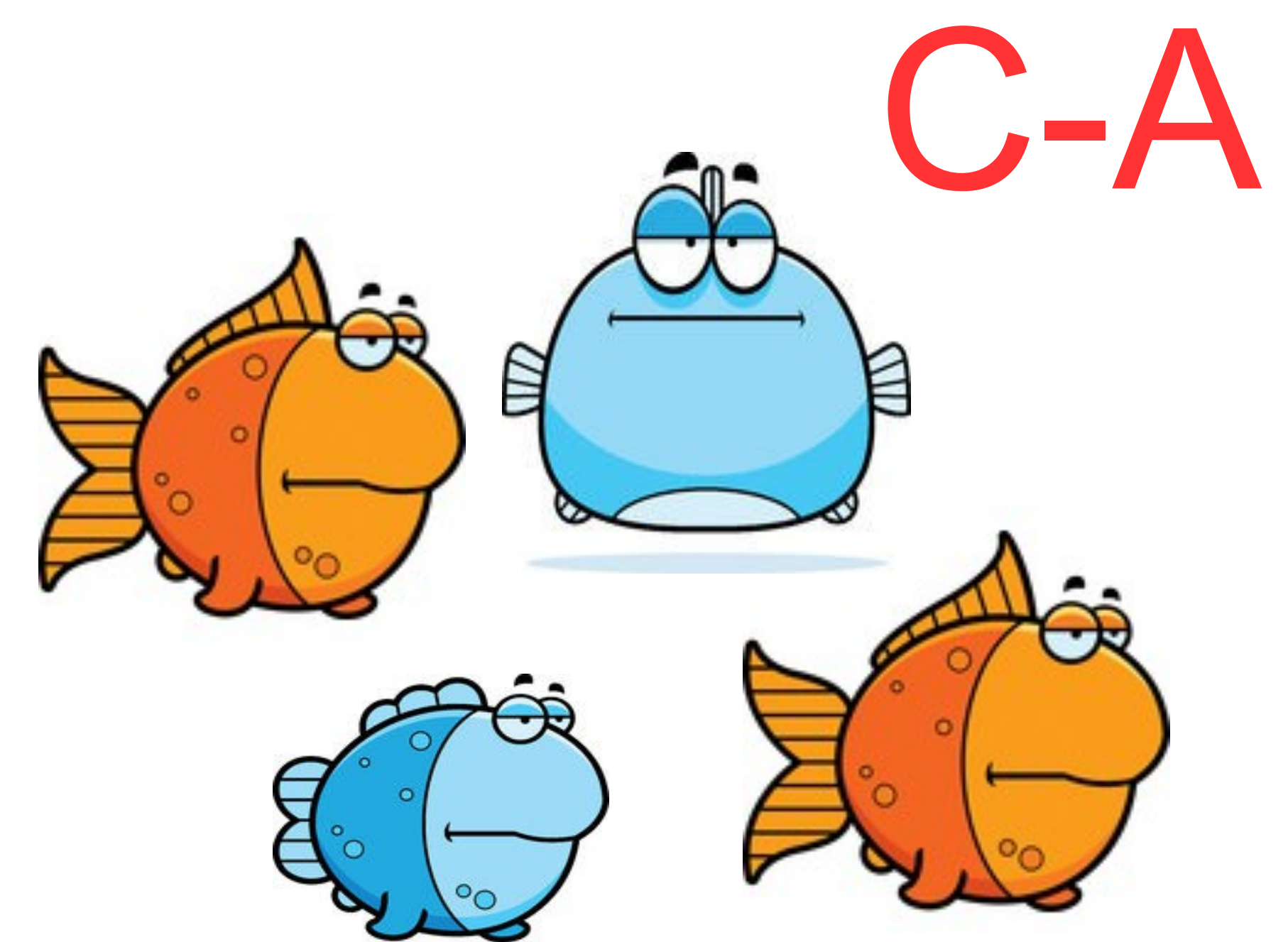
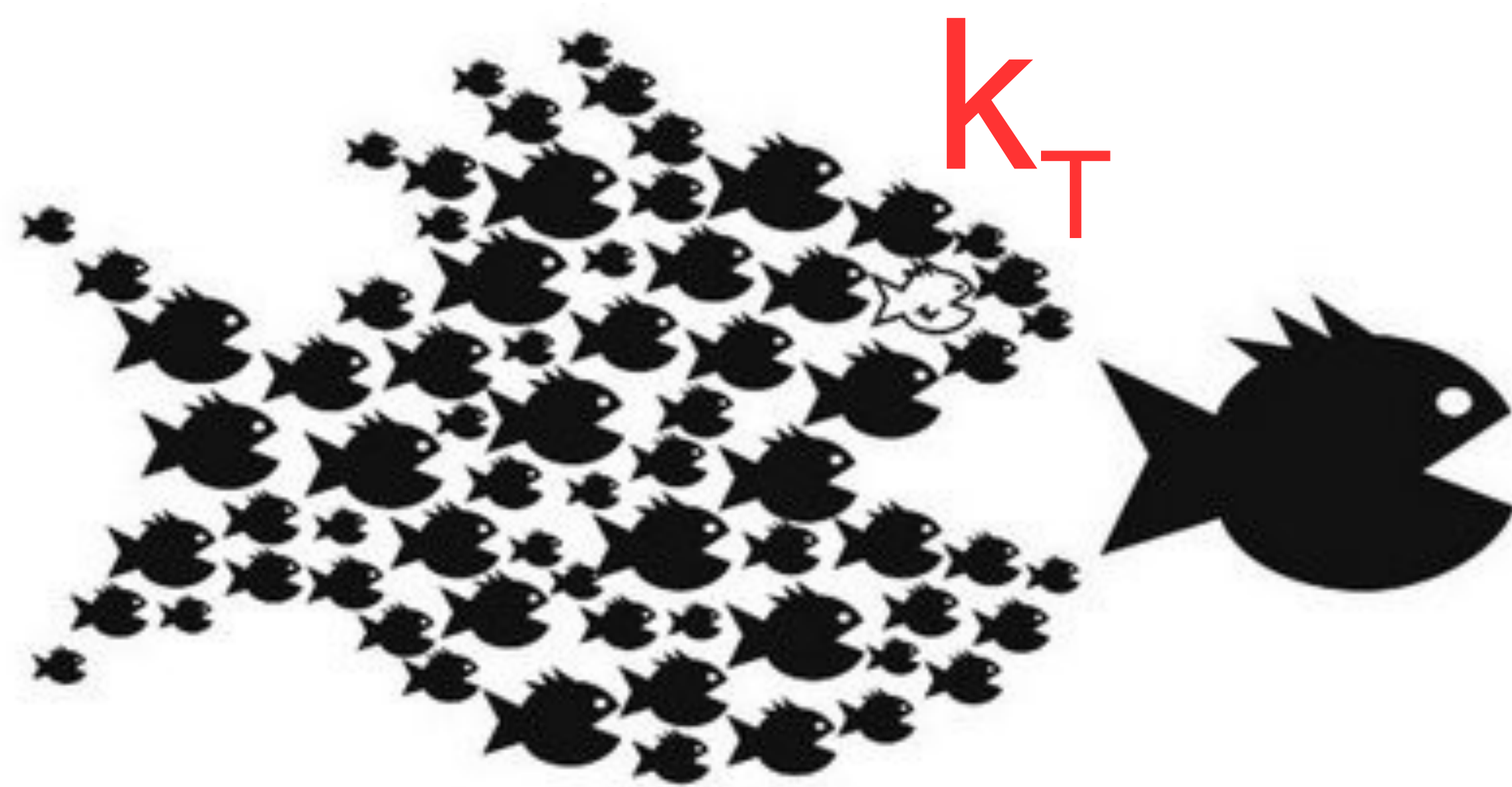
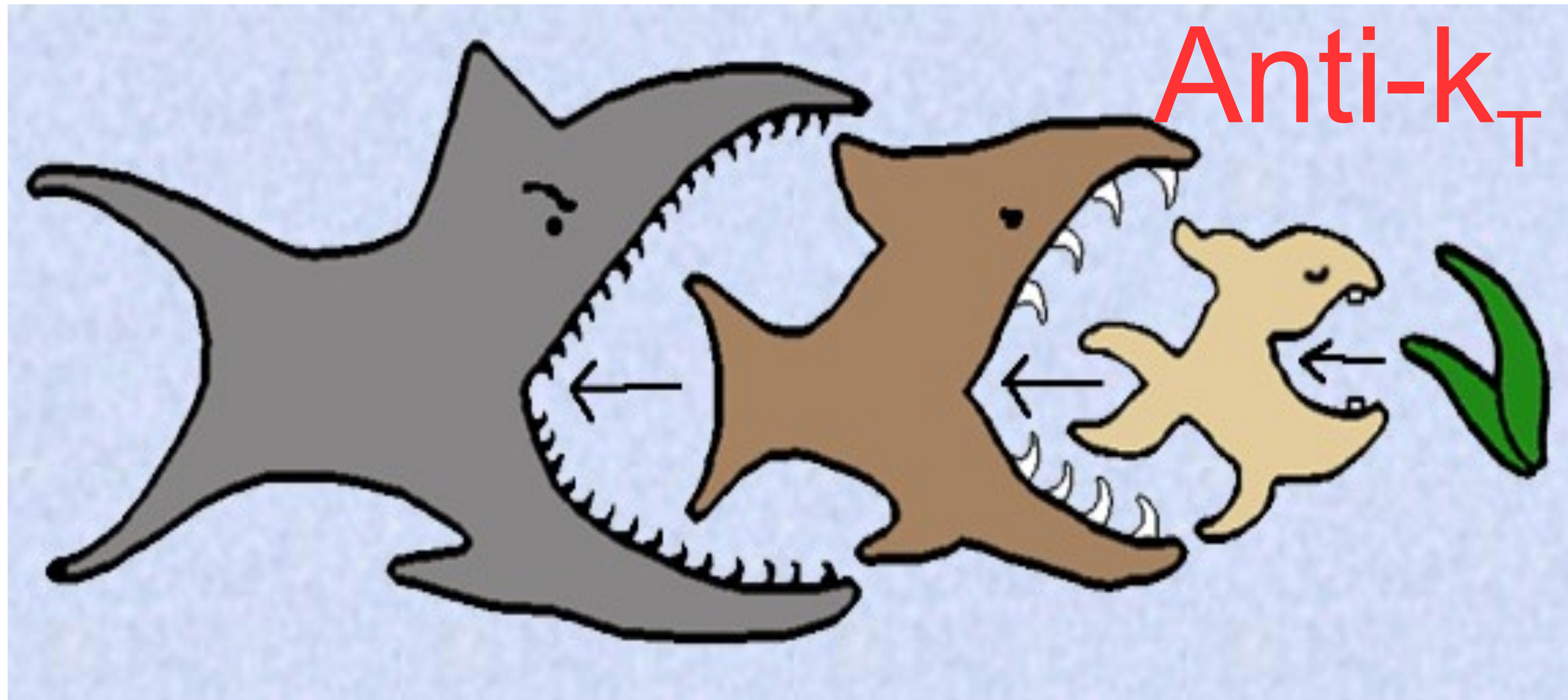


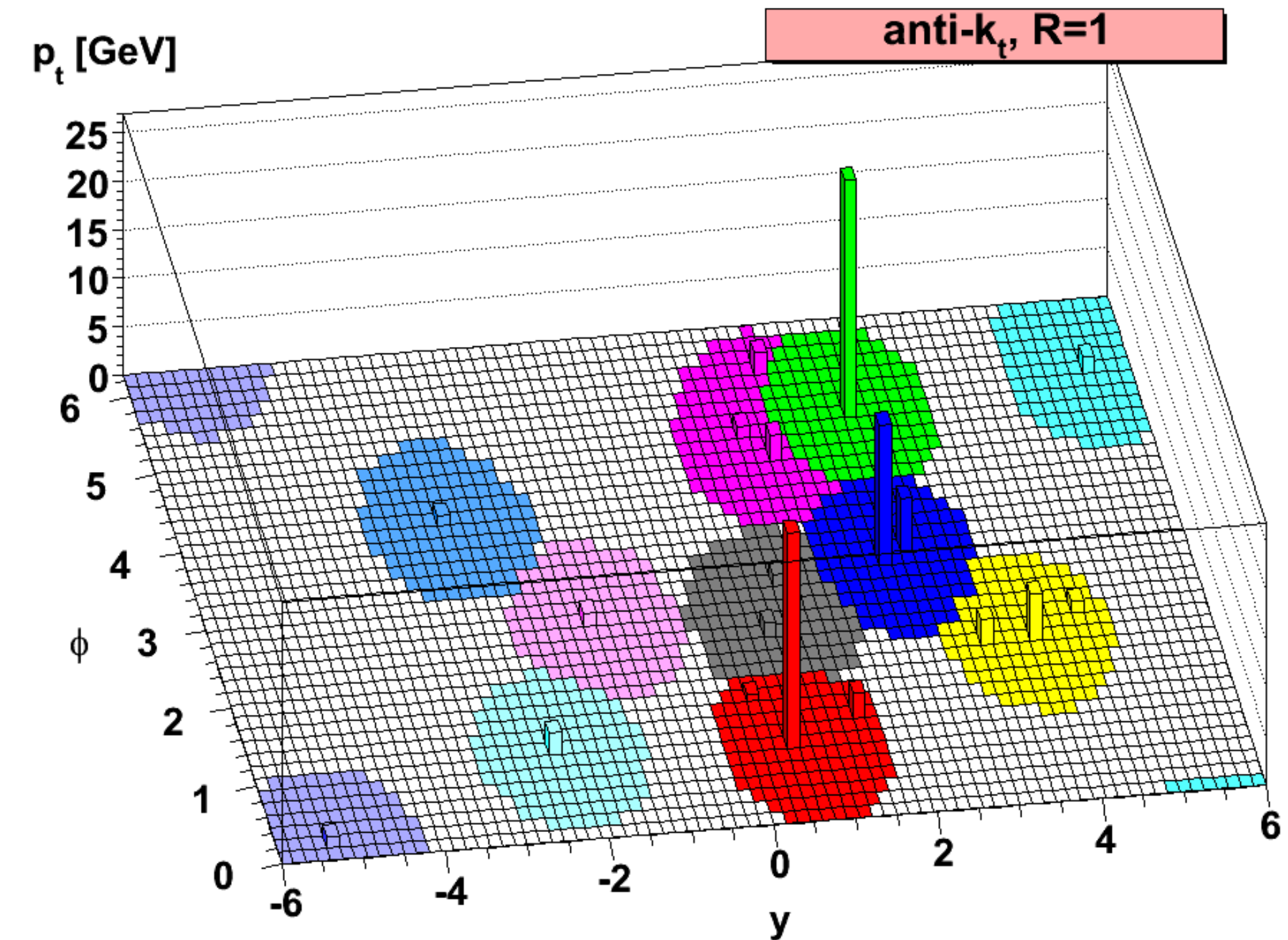
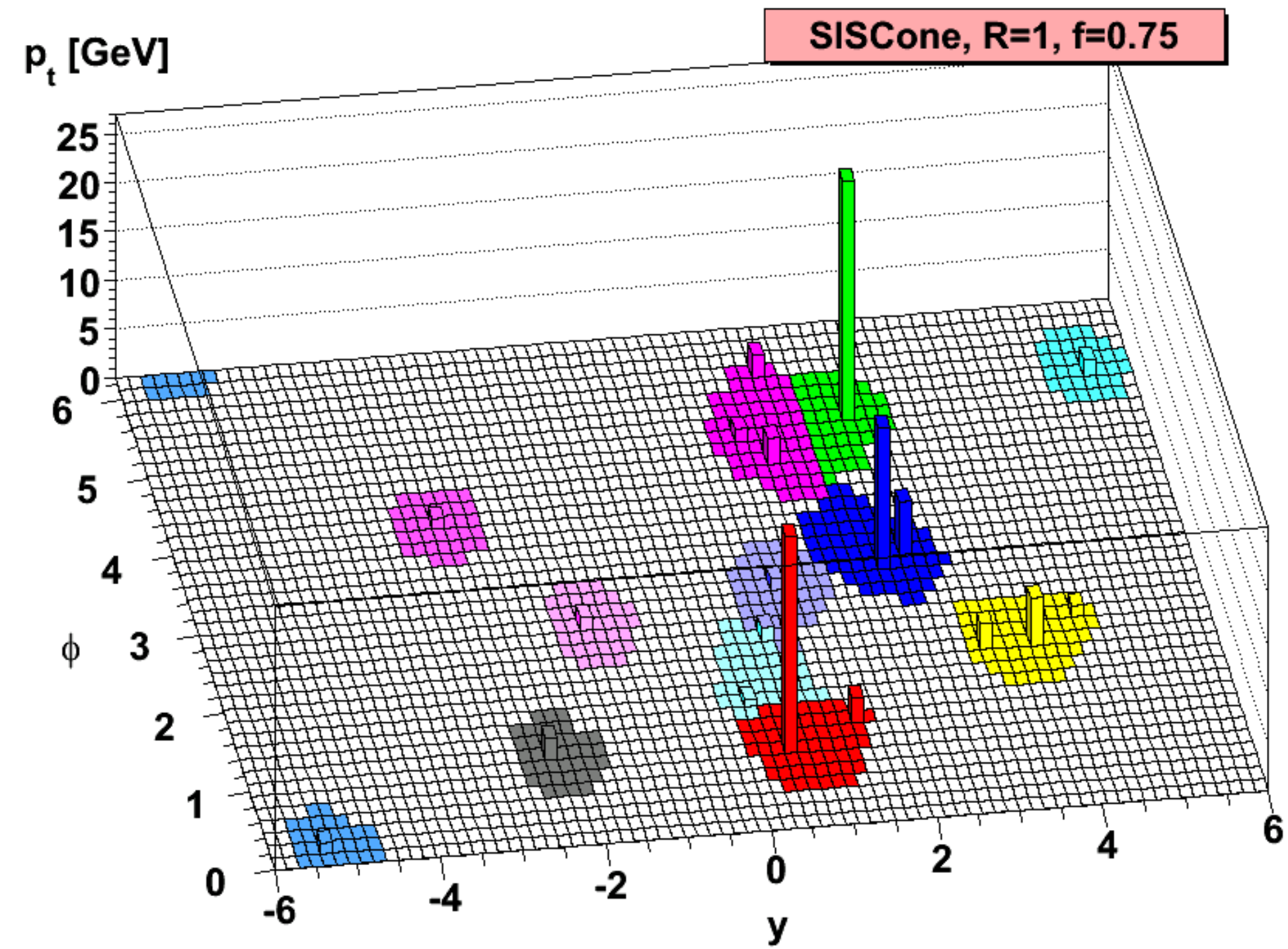
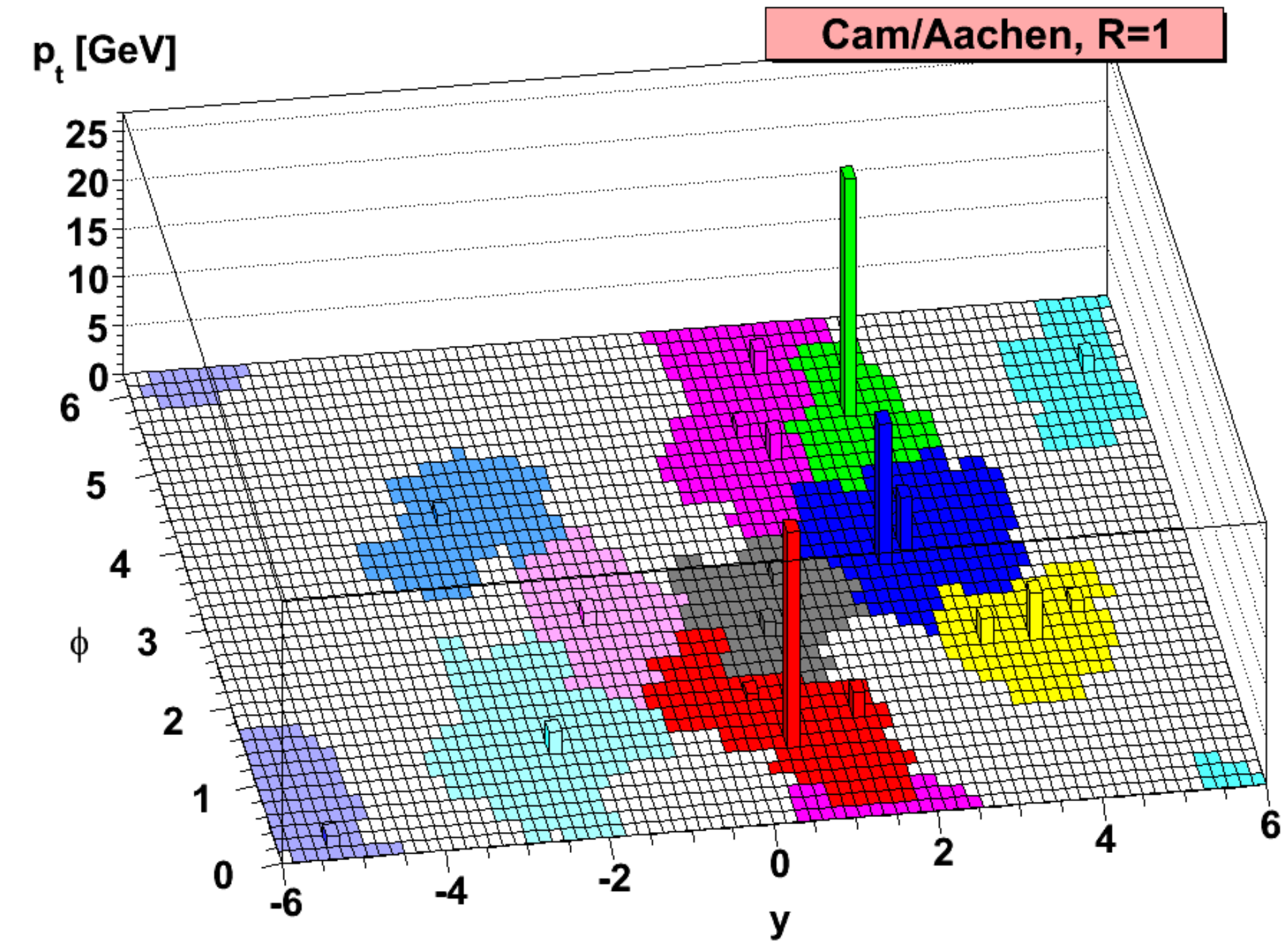
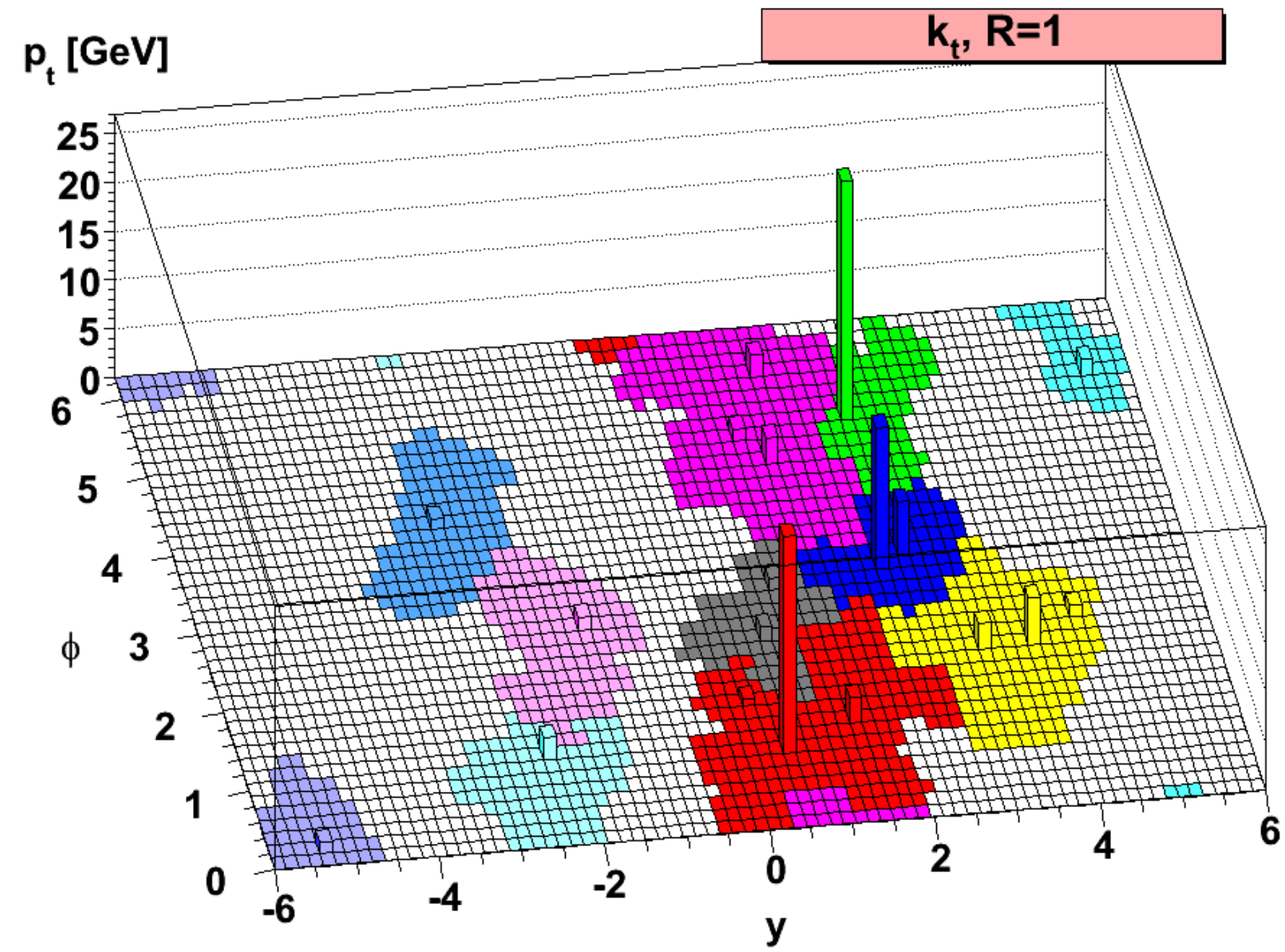
The algorithm found 3 jets, each with 4-vector equal to the sum of its components



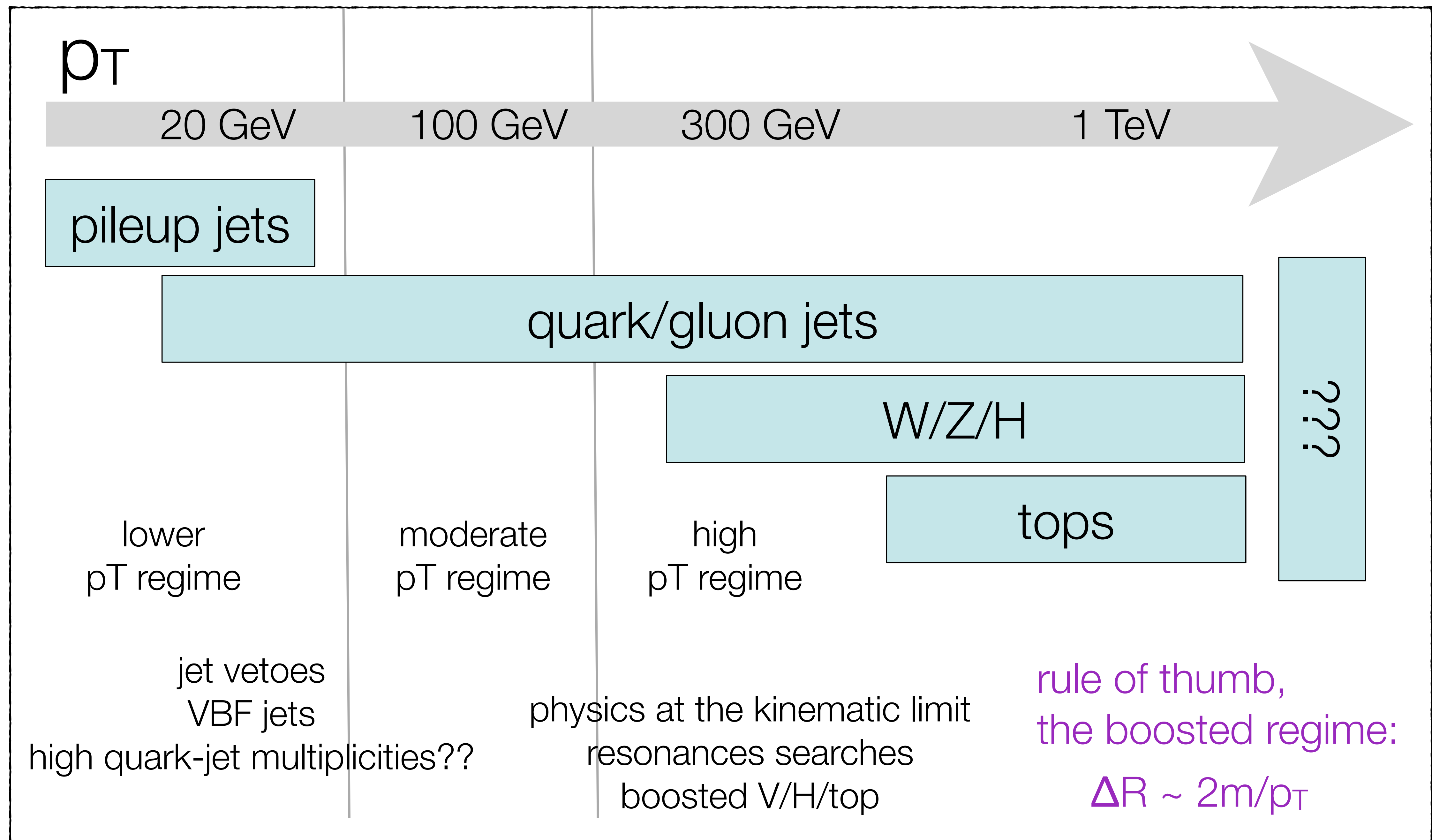
If we had used a different distance parameter, the answer would have been much different (6 jets instead of 3)







WHICH R?...WHICH JET?

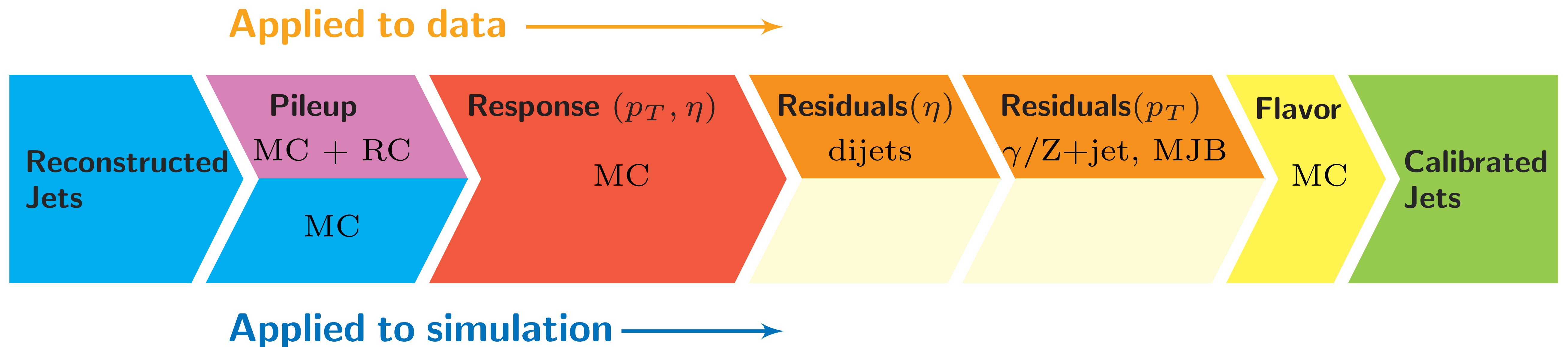


In the end, you pick the R that is appropriate for your analysis.

Discuss this more when talking about jet substructure

Most popular jet algorithm is AK4

A good choice for q/g jets with $p_T > 25$ GeV



This is an example of the CMS chain of jet energy corrections

Basic chain:

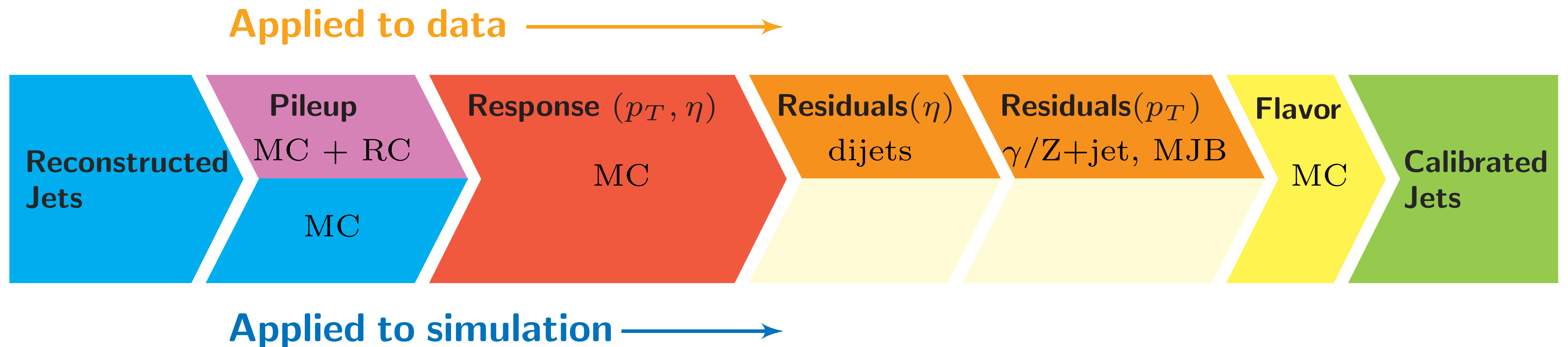
- Correct for pileup (on average)

- Correct for detector effects

 - Can be many things depending on detector: out-of-cone effects, detector response, material loss, etc.

- Correct for data/MC

- Correct for flavor of jet (q,g,b,etc.)



This is an example of the CMS chain of corrections.

Basic chain:

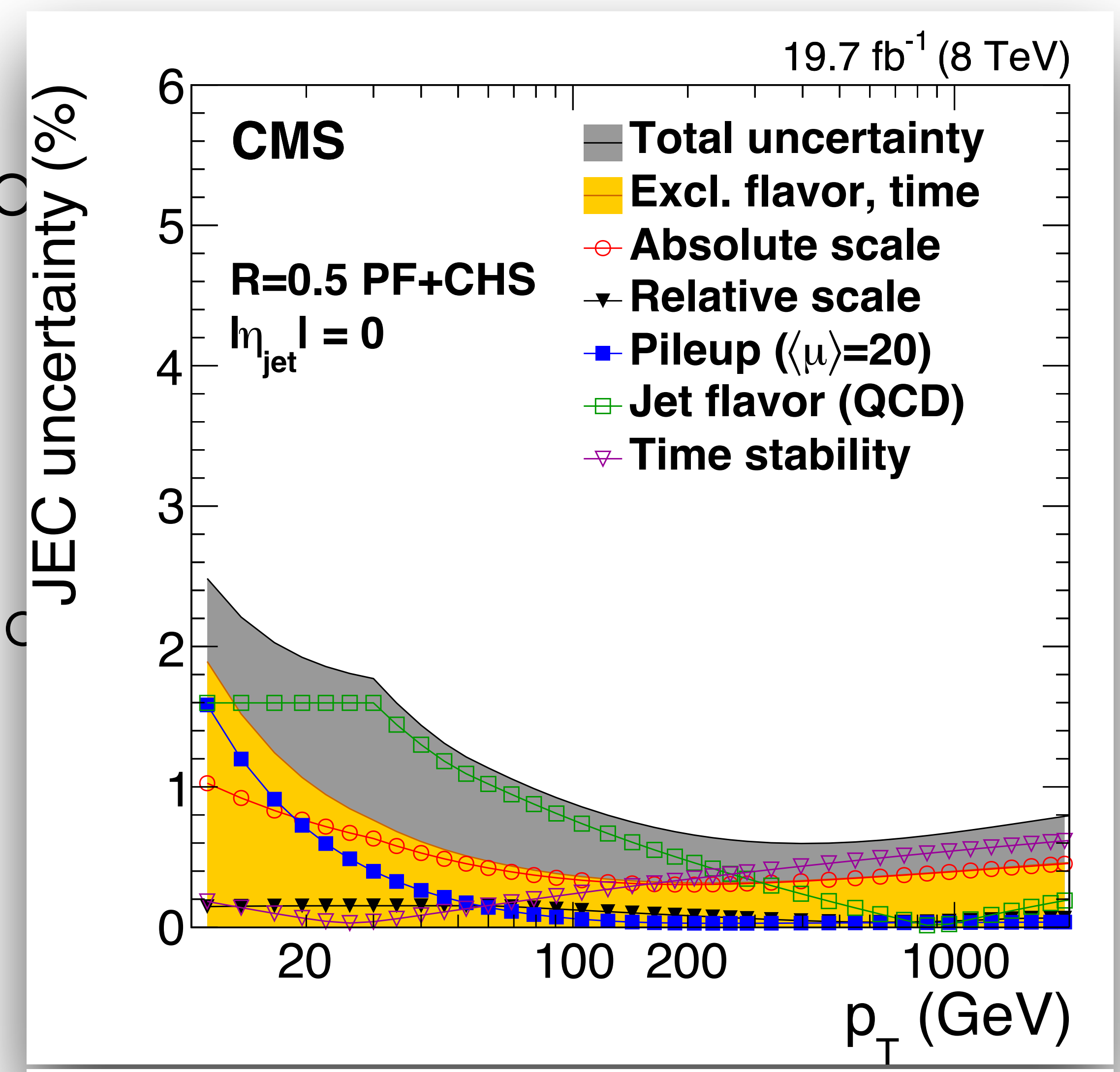
Correct for pileup (on average)

Correct for detector effects

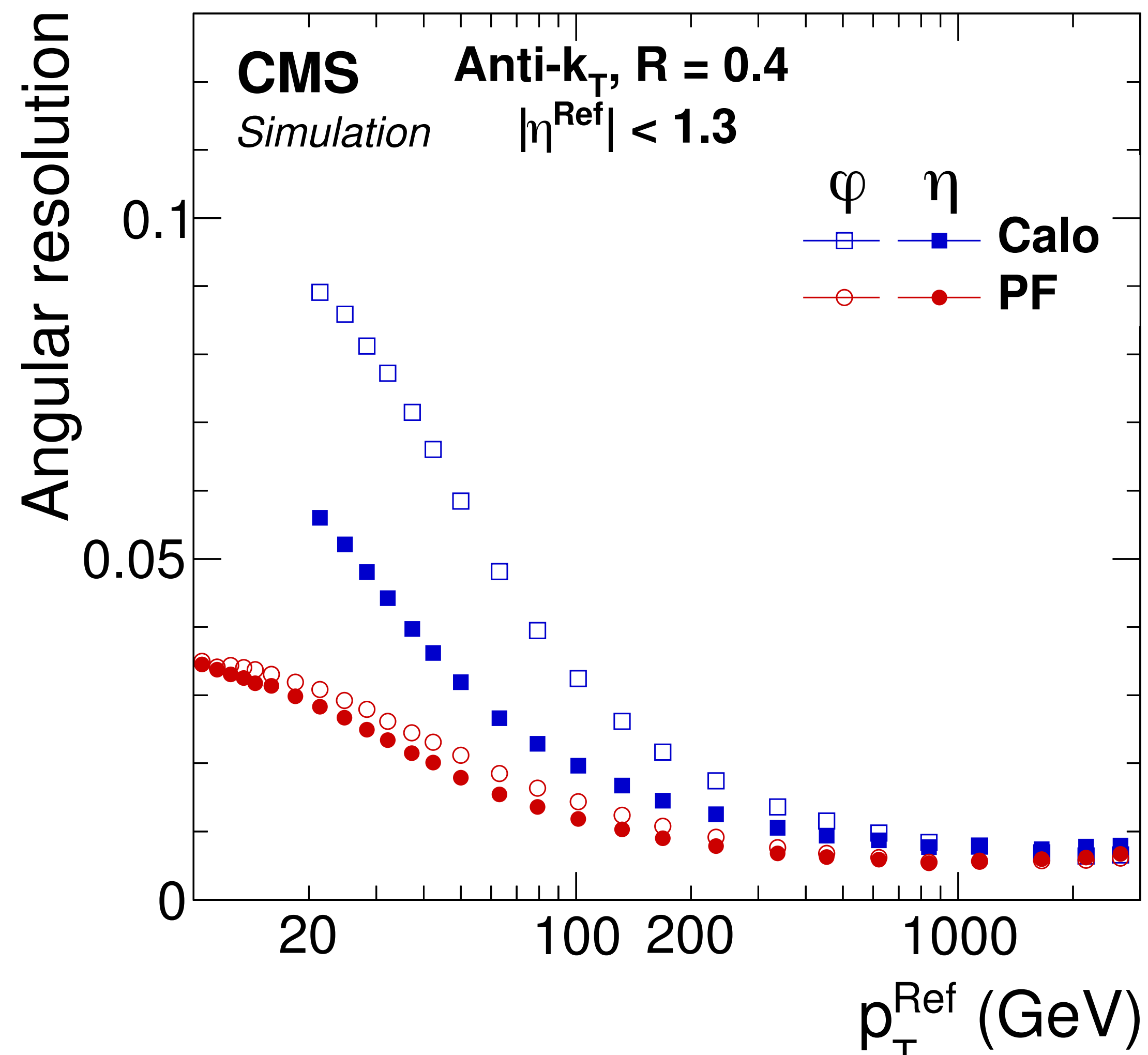
Can be many things depending on detector
response, material loss, etc.

Correct for data/MC

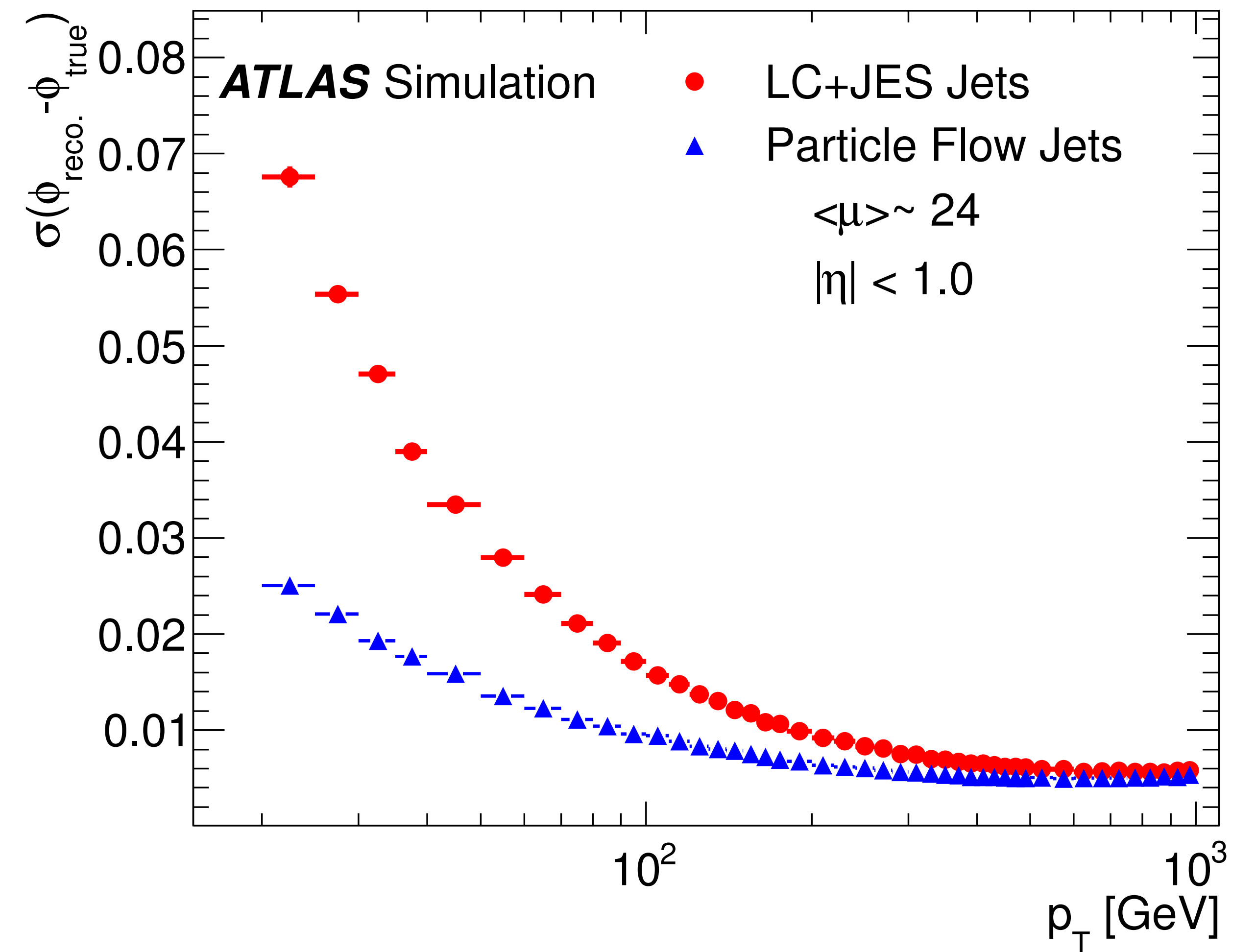
Correct for flavor of jet (q,g,b,etc.)



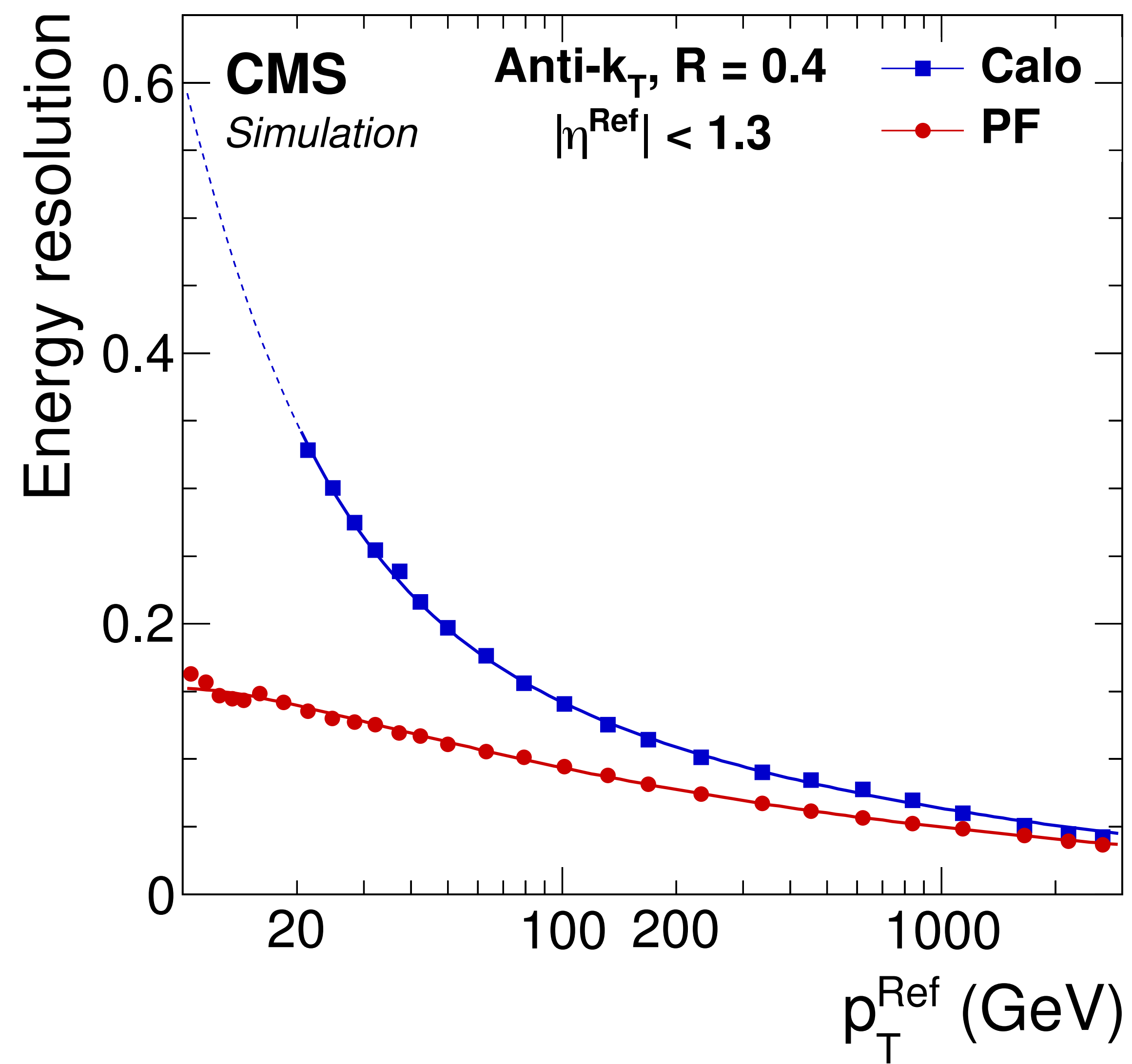
CMS



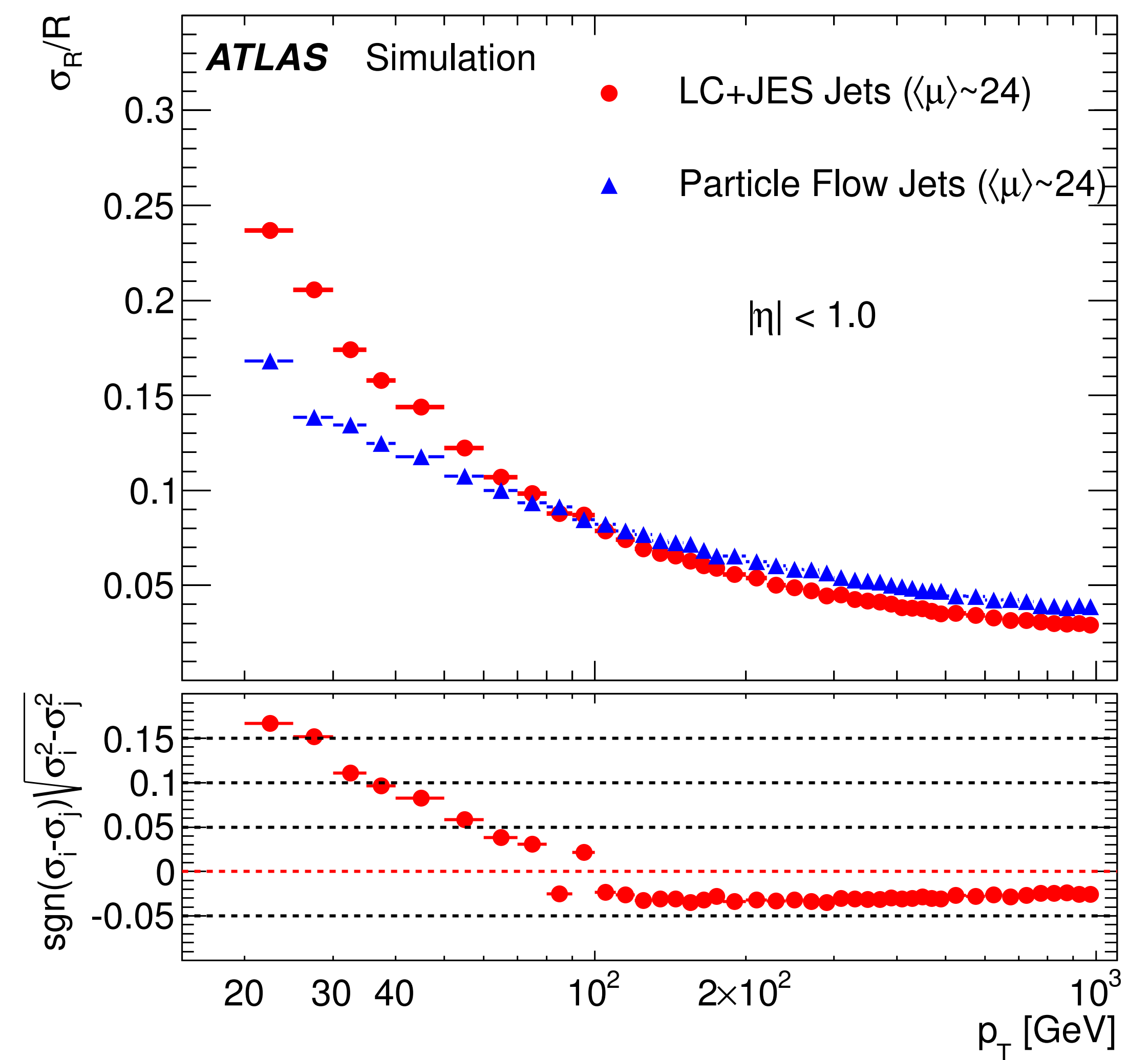
ATLAS



CMS



ATLAS



Comparing ATLAS & CMS

Ecal+Hcal pion resolution	$\frac{\sigma}{E_T} \approx \frac{40\%}{\sqrt{E_T}}$	$\frac{\sigma}{E_T} \approx \frac{100\%}{\sqrt{E_T}} \oplus 7\%$
Missing momentum resolution (TDR)	$\frac{\sigma(E_T)}{\Sigma E_T} \approx \frac{50\%}{\sqrt{\Sigma E_T}}$	$\frac{\sigma(E_T)}{\Sigma E_T} \approx \frac{120\%}{\sqrt{\Sigma E_T}} \oplus 2\%$
Inner tracker resolution (TDR)	$\frac{\sigma(p_T)}{p_T} \approx 1.8\% \oplus 60\% p_T$ (p_T in TeV)	$\frac{\sigma(p_T)}{p_T} \approx 0.5\% \oplus 15\% p_T$ (p_T in TeV)
B field inner region	2 Tesla : p_T swept < 350 MeV	4 Tesla : p_T swept < 700 MeV

ATLAS has better calorimetry; CMS has better tracking

Good jet & MET resolution important!

Improve CMS Jet & MET resolution using full detector

MET: the garbage collector

You need to understand EVERYTHING in your detector before you can understand missing energy!

MET is the absence of energy in your detector

Important for signals with neutrinos, e.g. τ , W , Z , t

Important for beyond the SM signals like dark matter!

Important:

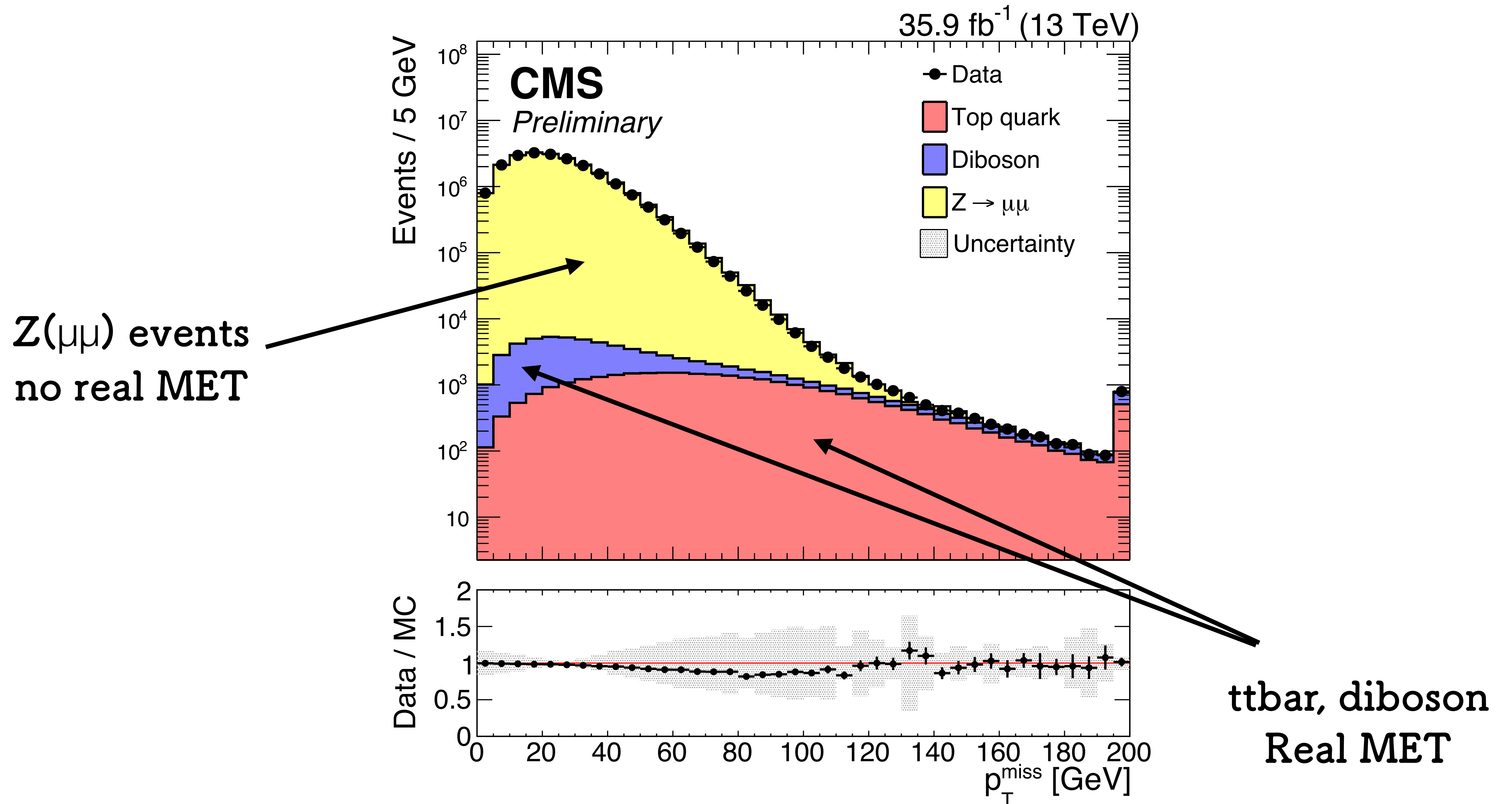
MET resolution - how well can you measure the energy of everything else without creating imbalances?

Physics: missing energy coming from resonances like $t\bar{t}b\bar{a}$

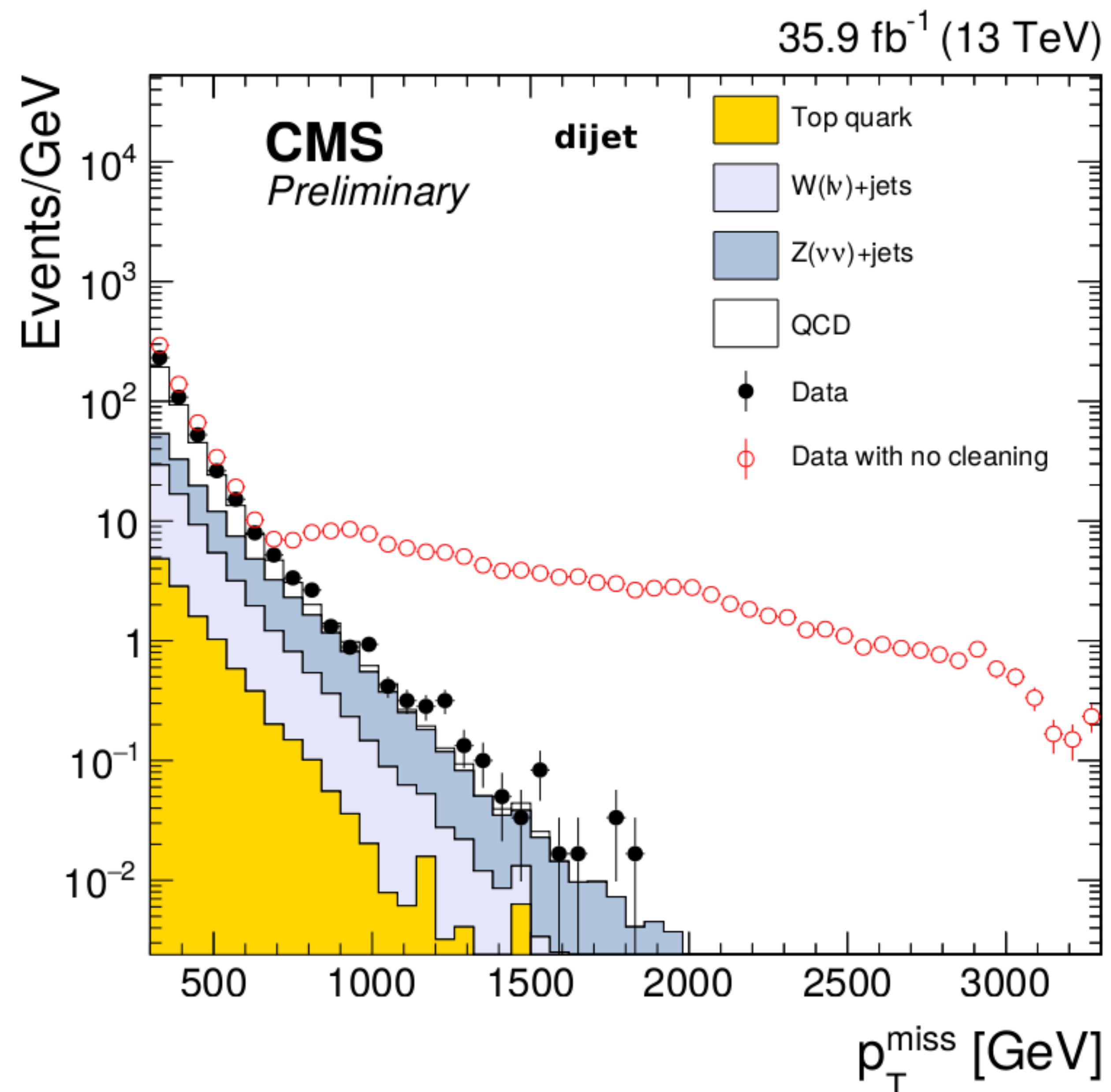
MET tails - how well can you understand the rare/pathological things in your reconstruction

Physics: non-resonant, high invisible energy like mono-jet

Core MET resolution vs. real MET tails



The better the MET resolution, the better you can identify real MET
Driven by jet resolution and how you handle soft unclustered deposits



Noise cleaning and filtering

cleaning - remove anomalous spikes before doing reconstruction

filtering - remove anomalous events from the dataset

Sources:

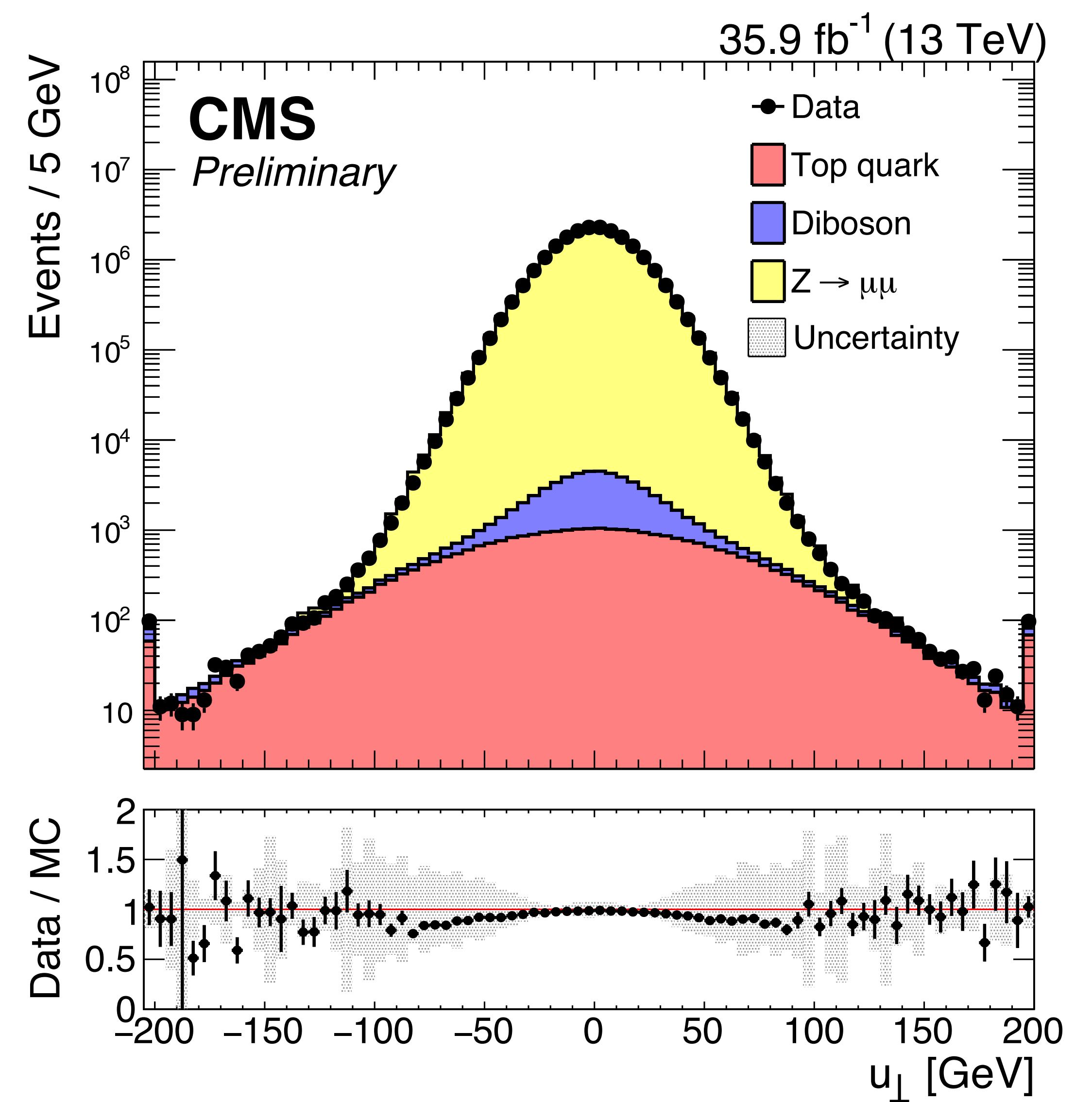
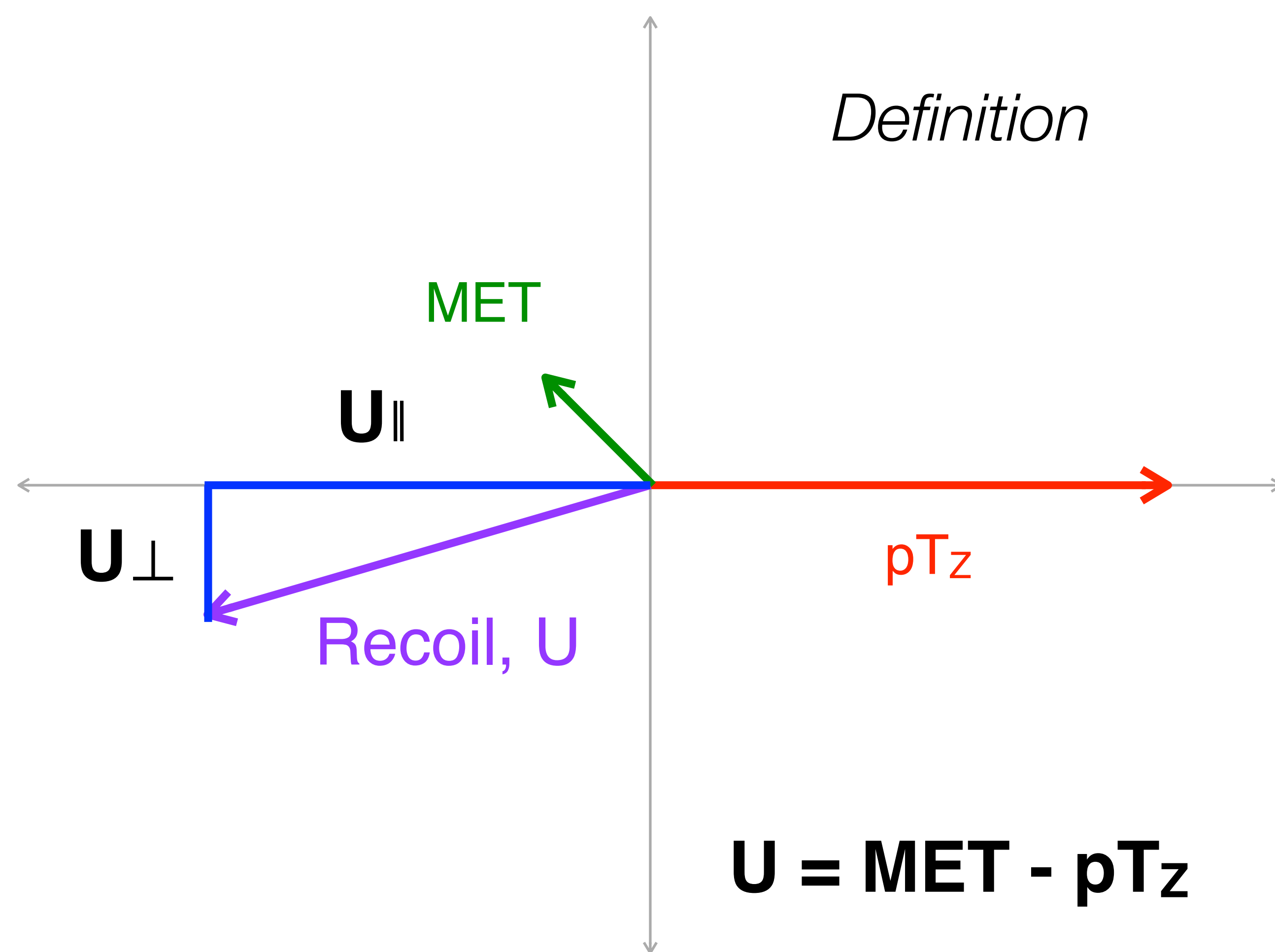
Electronics/detector noise, e.g. spurious interactions with photodectors

Physics signals like beam halo muons

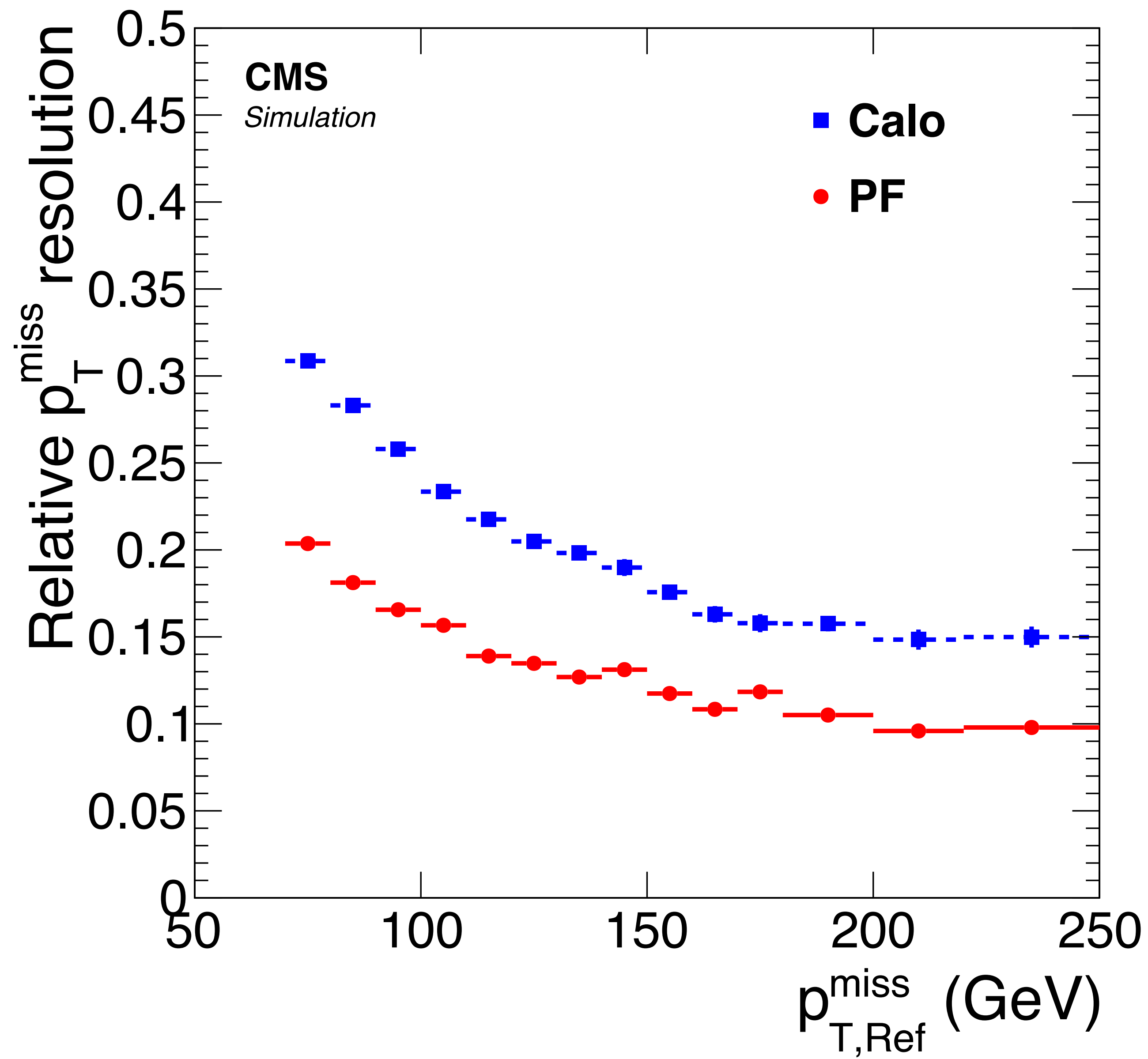
Reconstruction effects, poorly id'ed low p_T muons

MET VALIDATION IN DATA

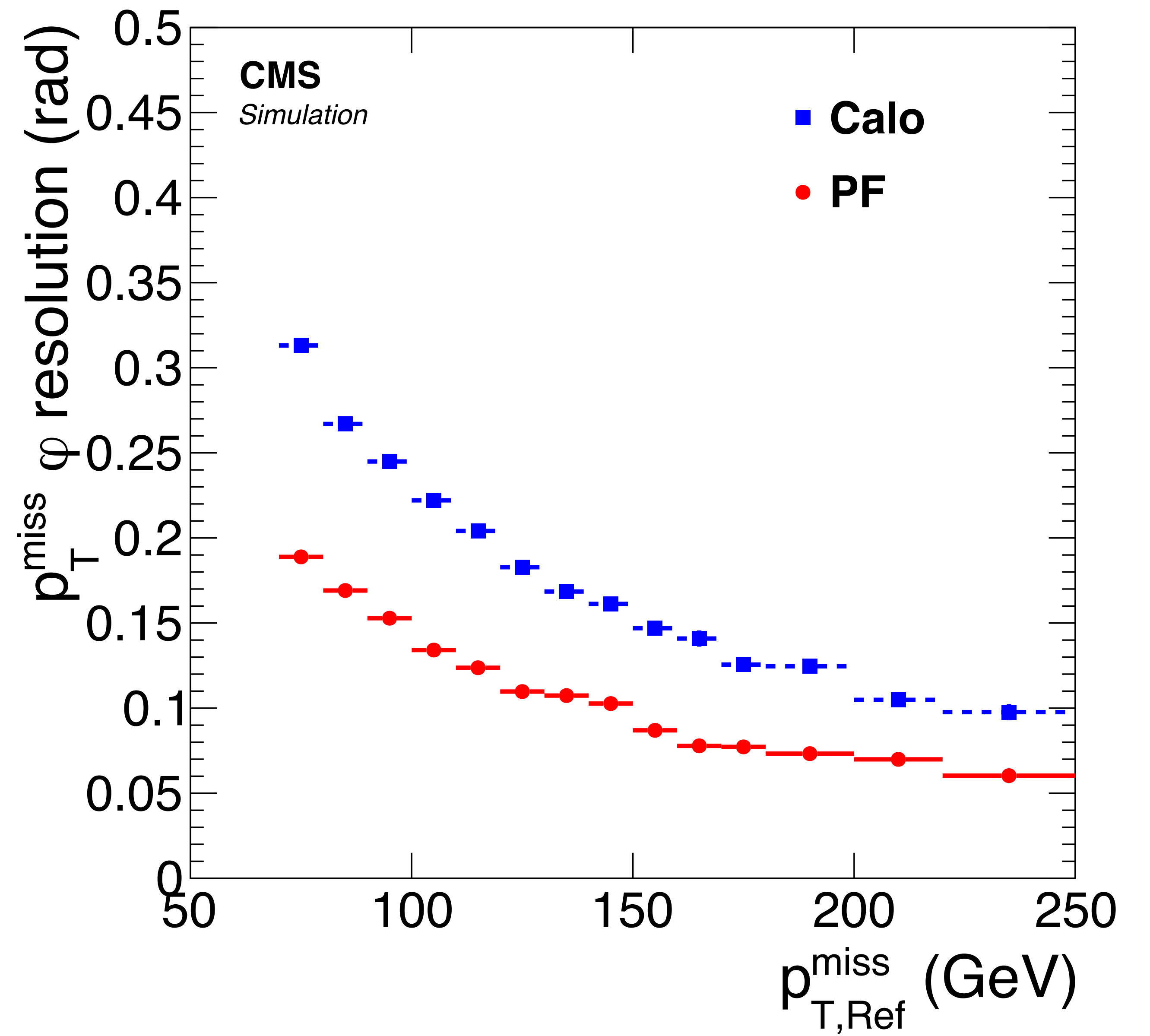
Use Drell-Yan events where a well-measured Z boson can be treated as MET to understand the recoil



scale resolution



angular resolution

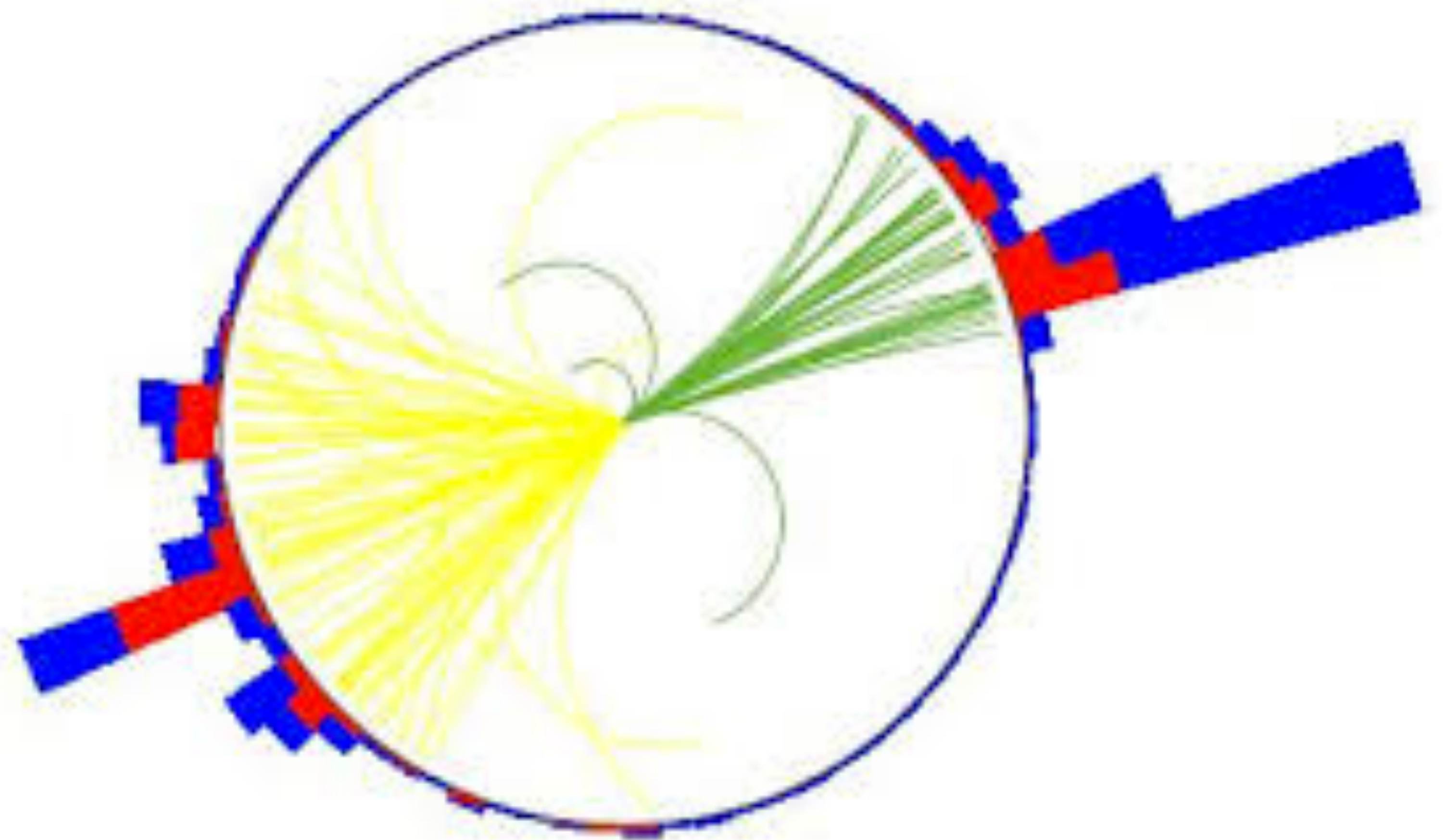
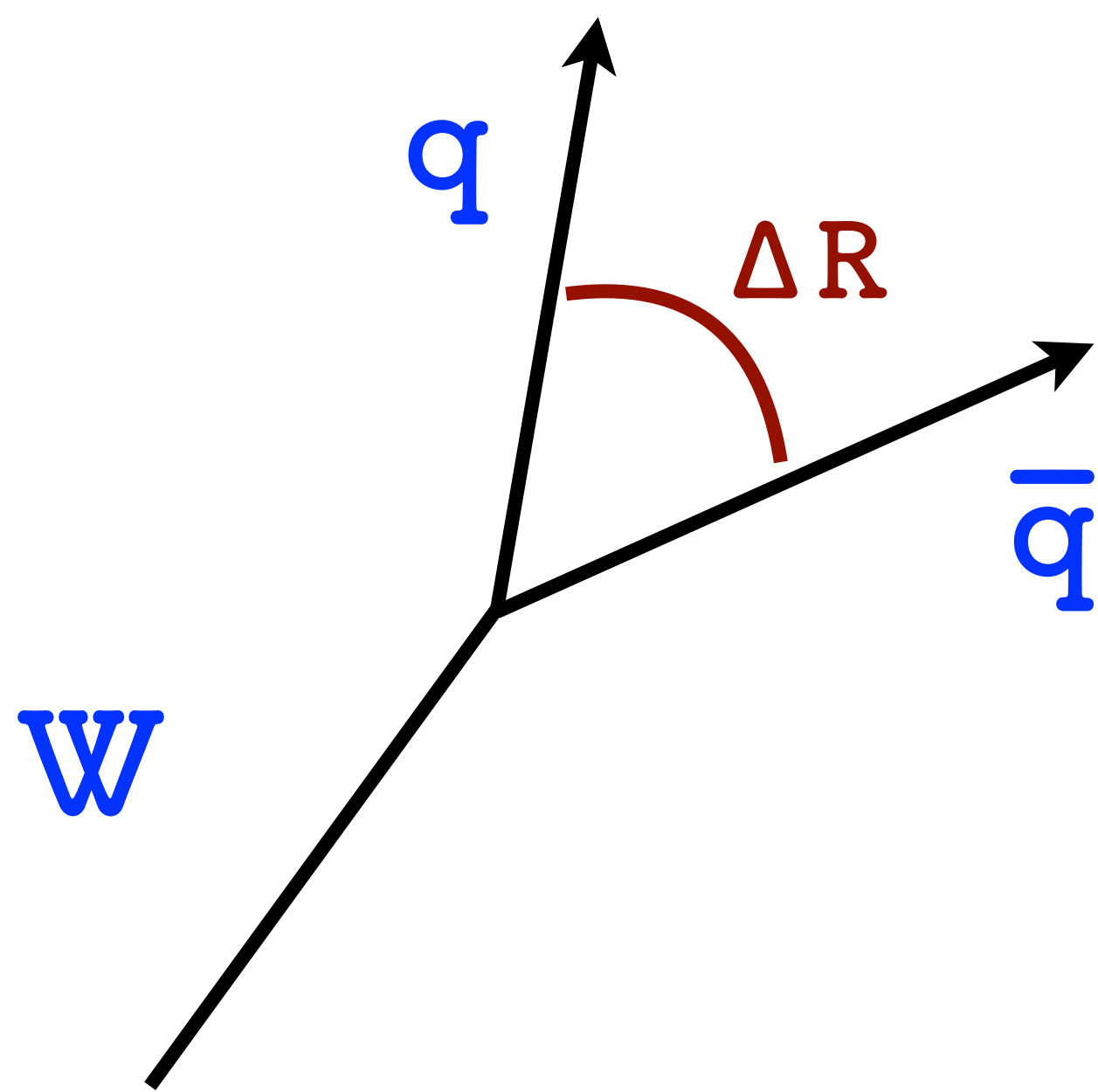


Finding structure in QCD radiation

At LHC energies, interesting heavy objects can be produced with a lot of boost.

Characteristic angular separation

$$\Delta R_{\text{dau}} = 2 m_{\text{mother}} / p_{T,\text{mother}}$$



HIGH MASS RESONANCES

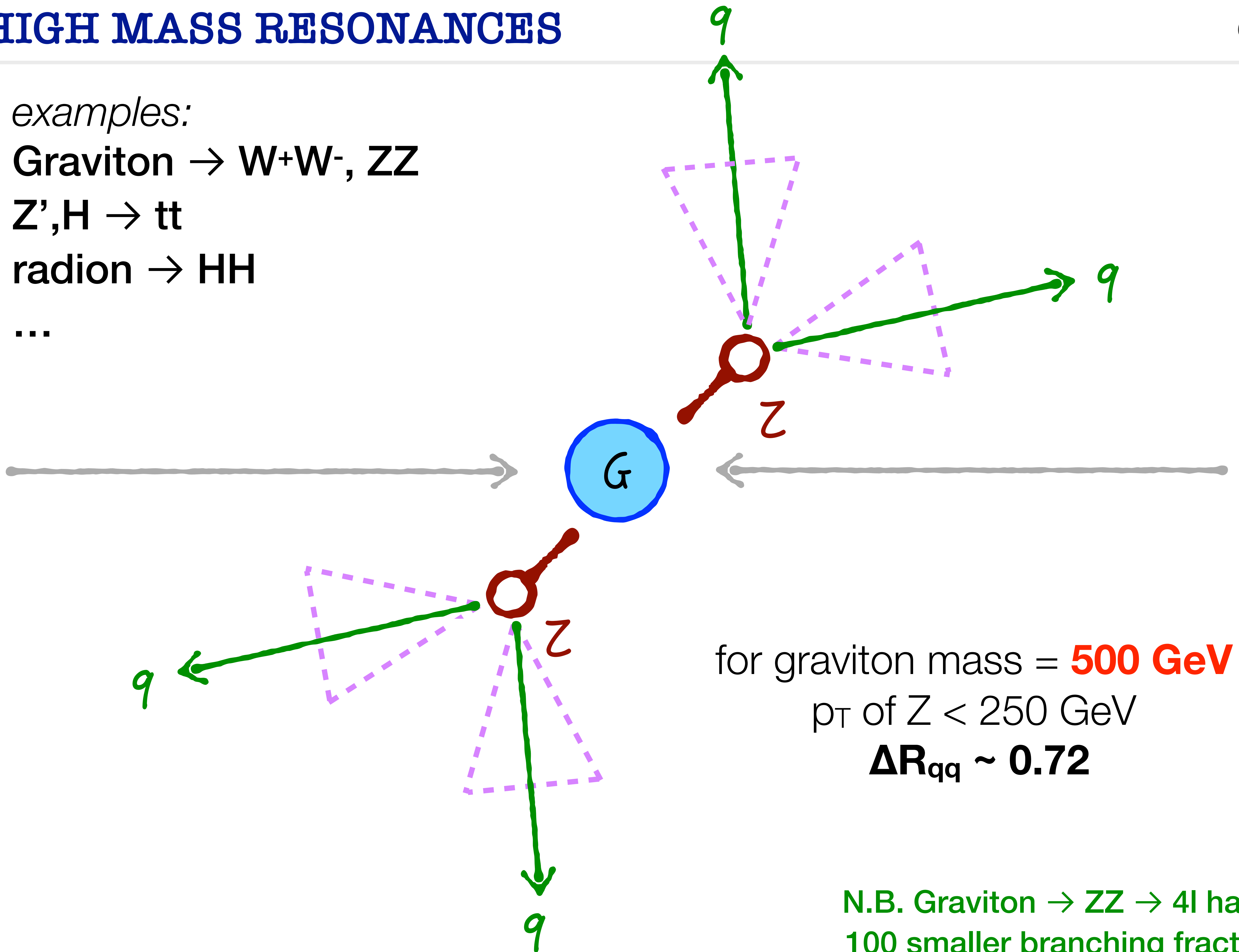
examples:

Graviton $\rightarrow W^+W^-$, ZZ

$Z', H \rightarrow tt$

radion $\rightarrow HH$

...



for graviton mass = **500 GeV**

p_T of $Z < 250$ GeV

$\Delta R_{qq} \sim 0.72$

N.B. Graviton $\rightarrow ZZ \rightarrow 4l$ has a
100 smaller branching fraction

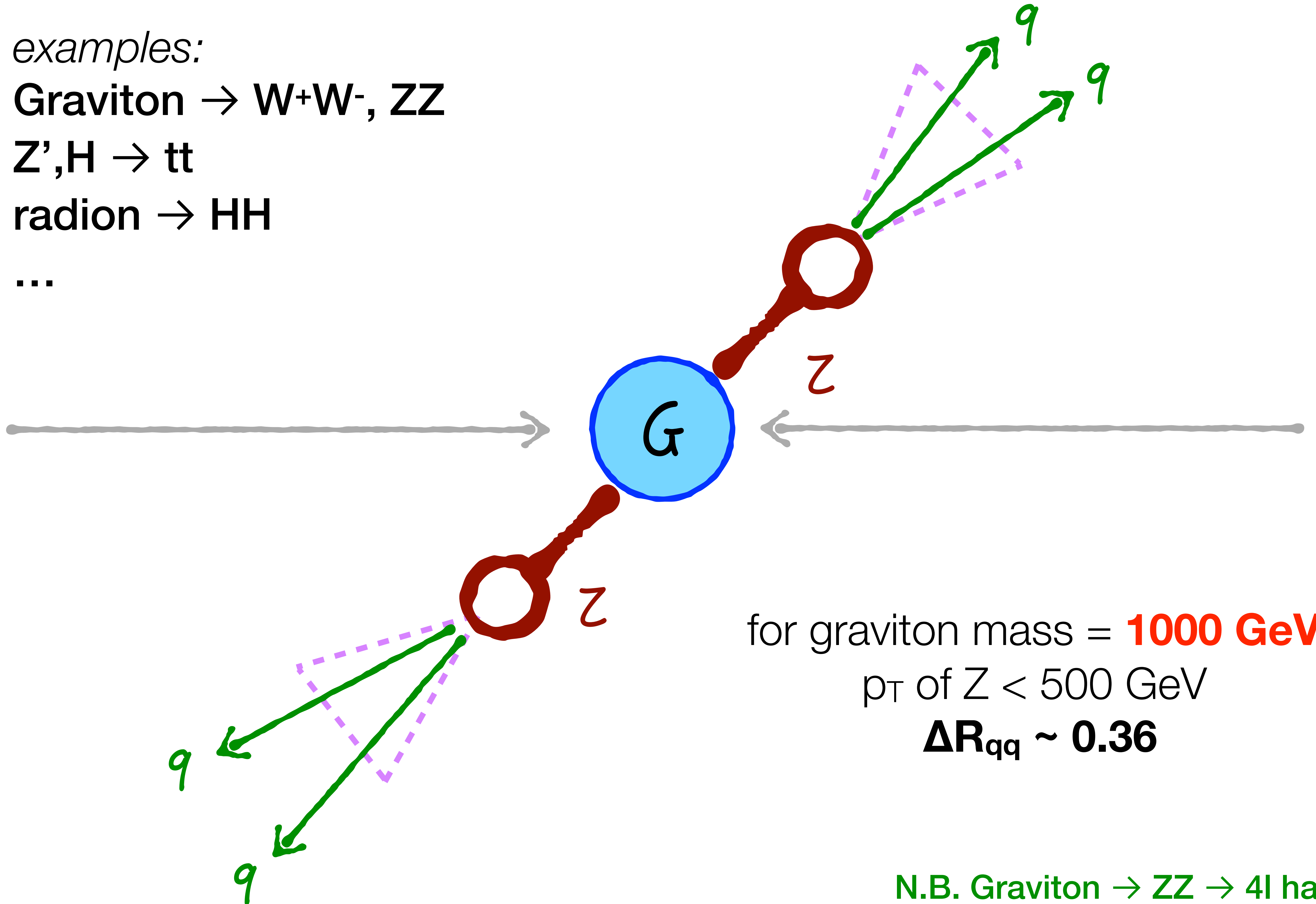
examples:

Graviton $\rightarrow W^+W^-, ZZ$

$Z', H \rightarrow tt$

radion $\rightarrow HH$

...



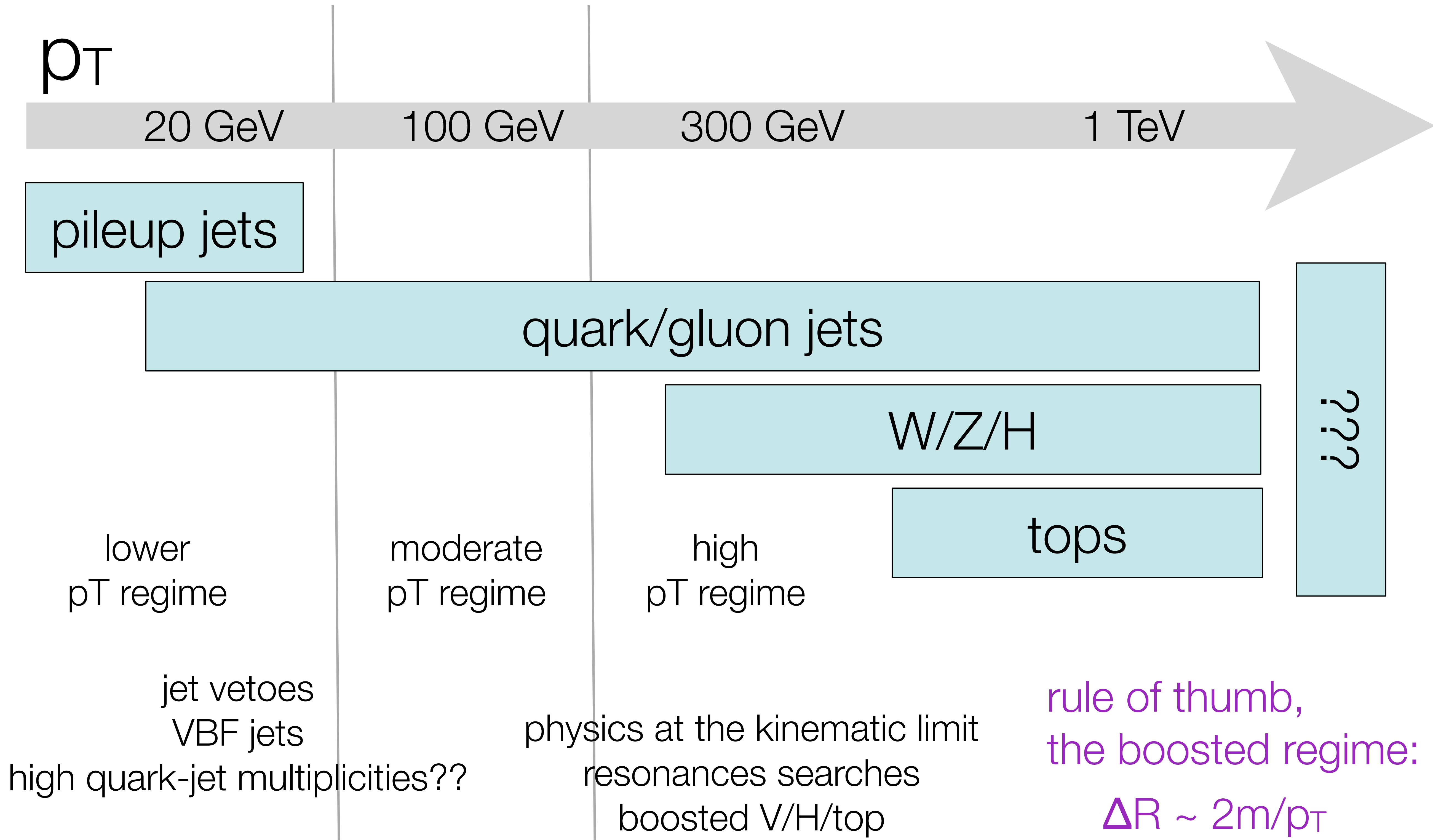
for graviton mass = **1000 GeV**

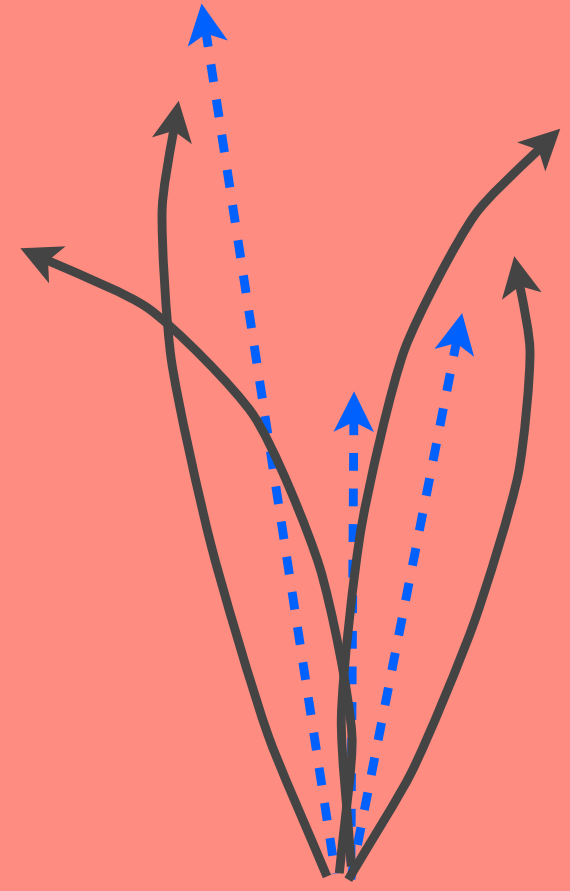
p_T of $Z < 500$ GeV

$\Delta R_{qq} \sim 0.36$

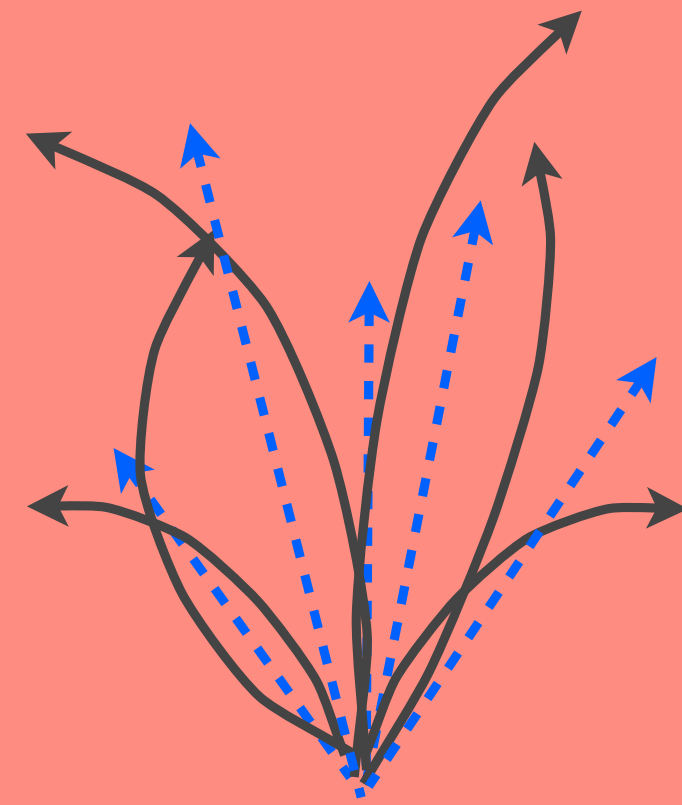
N.B. Graviton $\rightarrow ZZ \rightarrow 4l$ has a
100 smaller branching fraction

JET SUBSTRUCTURE

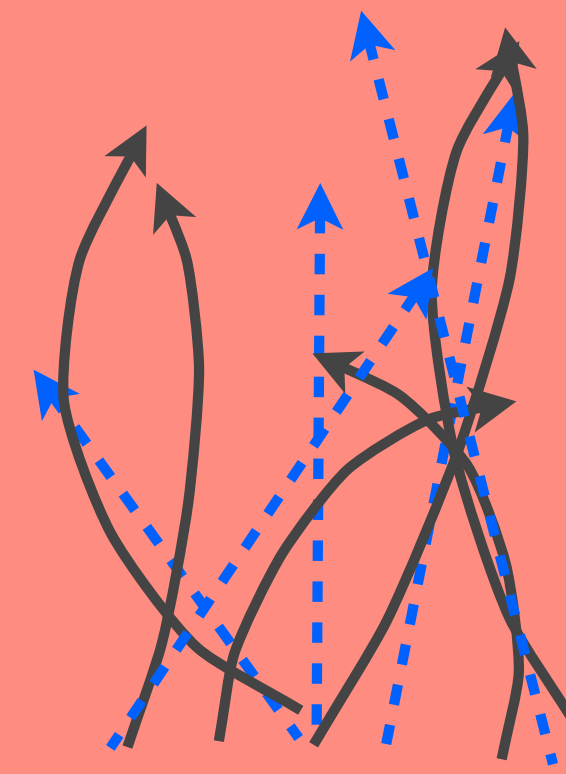




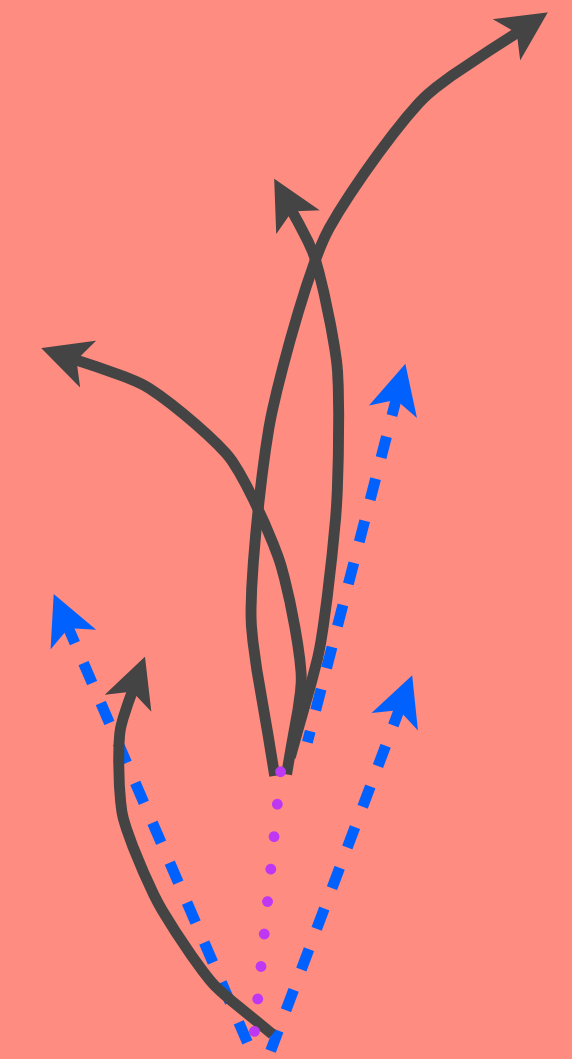
u, d or s jet



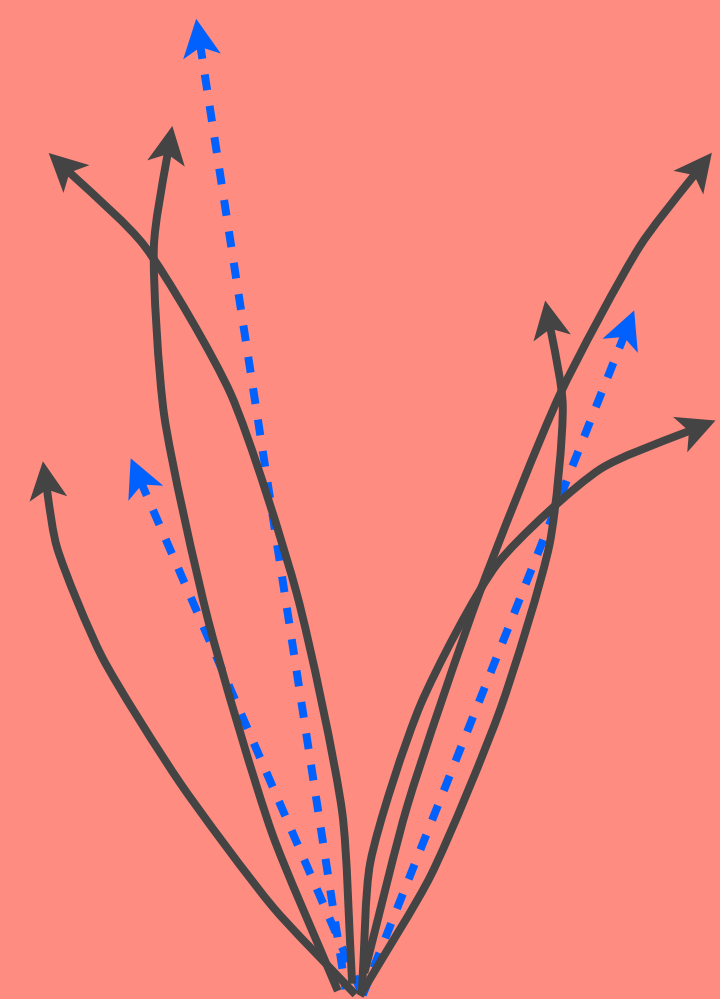
gluon jet



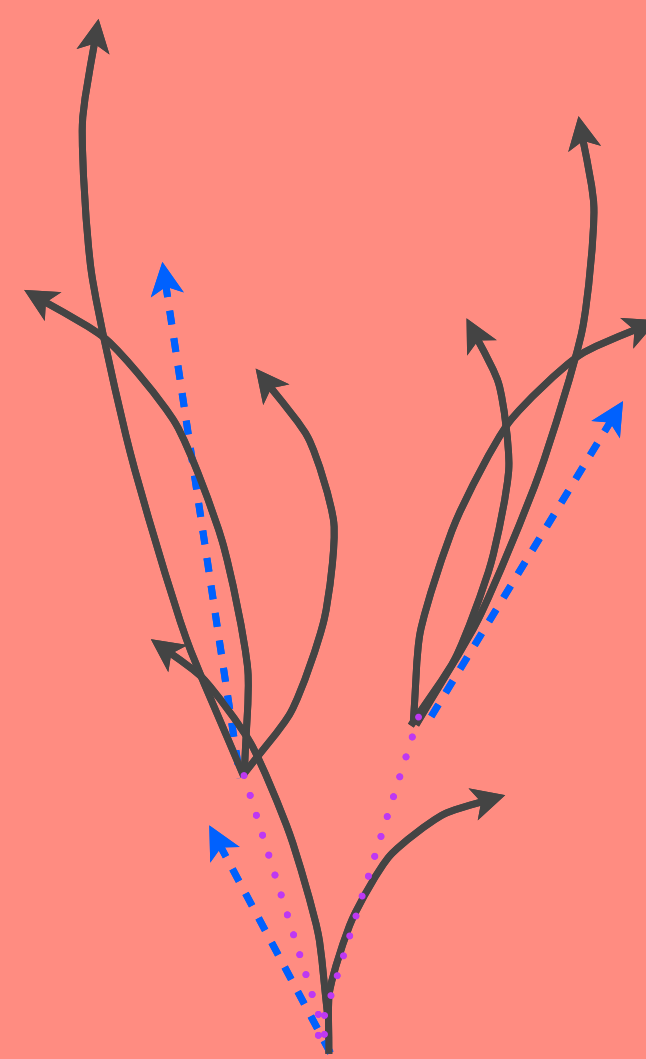
pileup jet



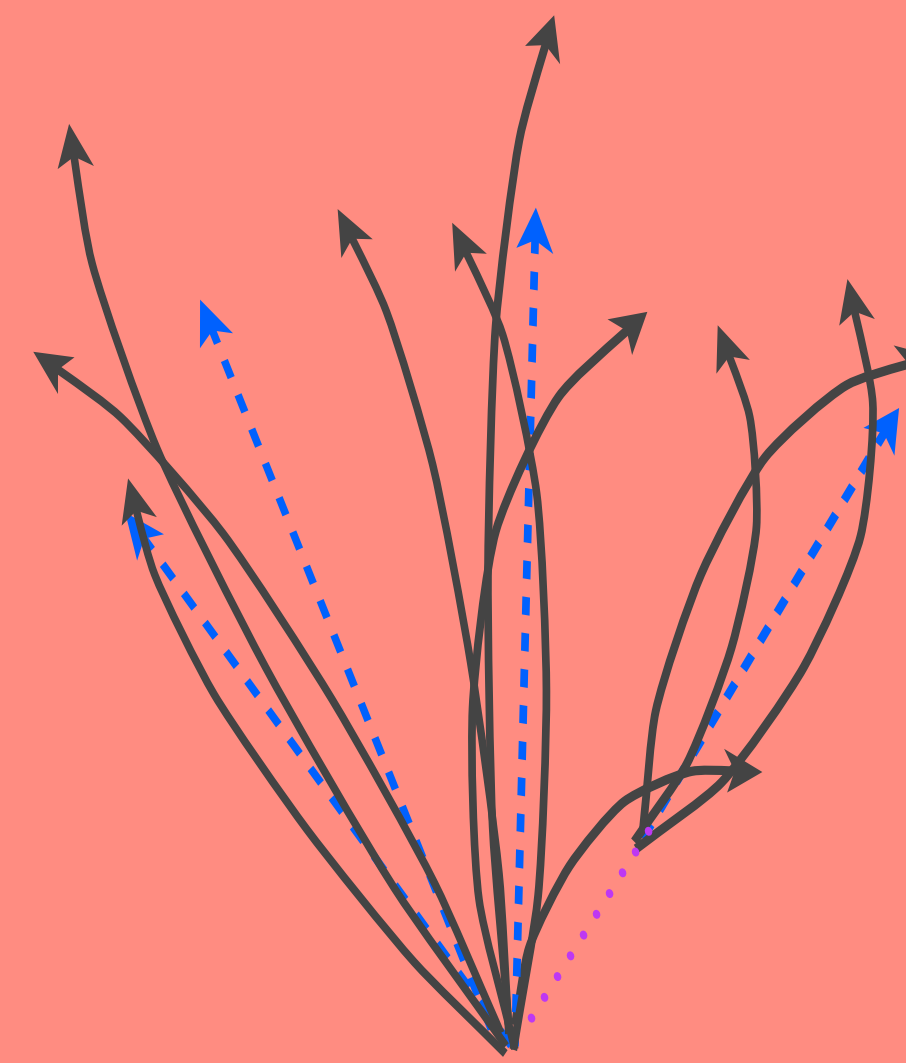
c or b jet



W or Z jet

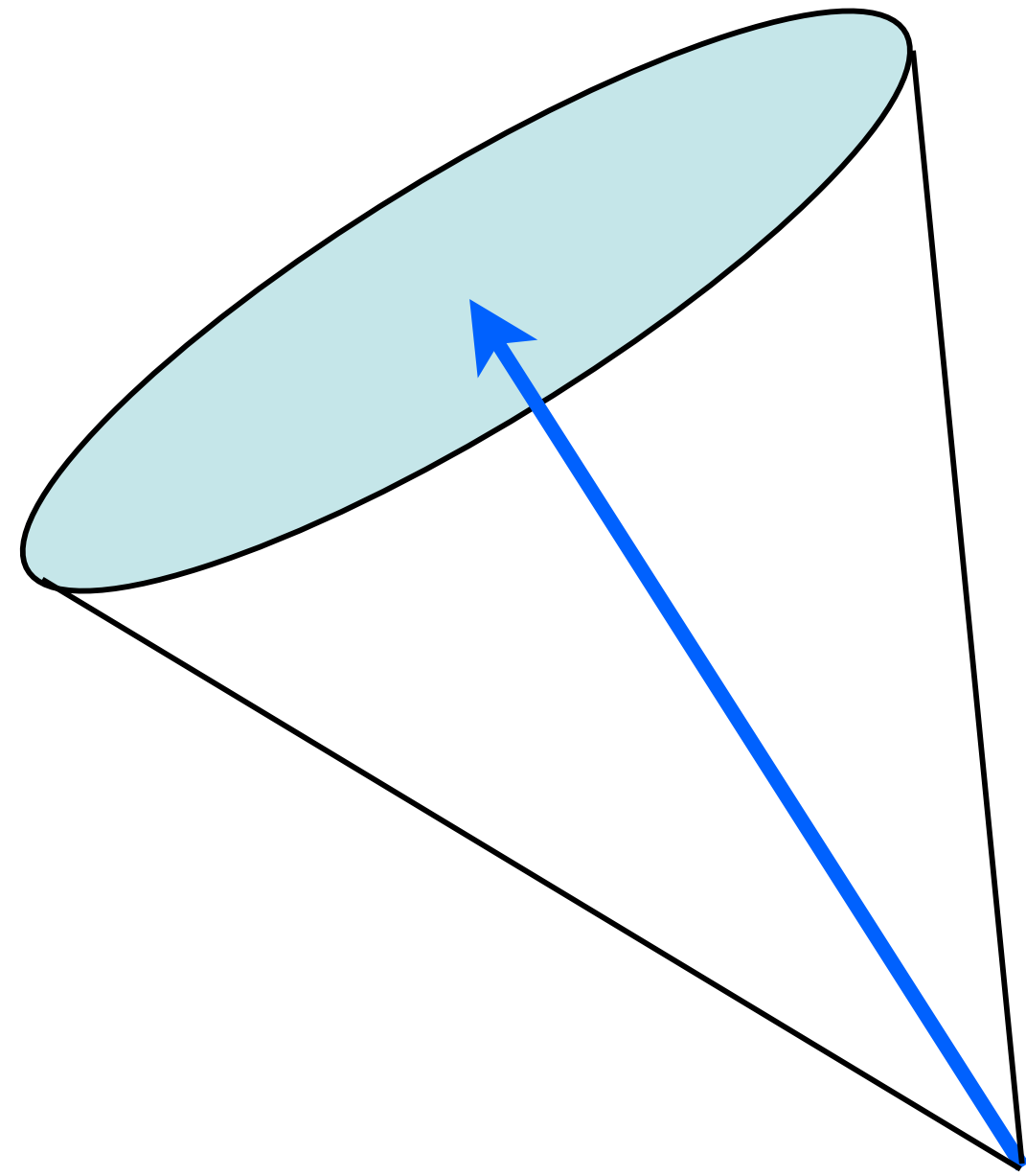


Higgs jet



top jet

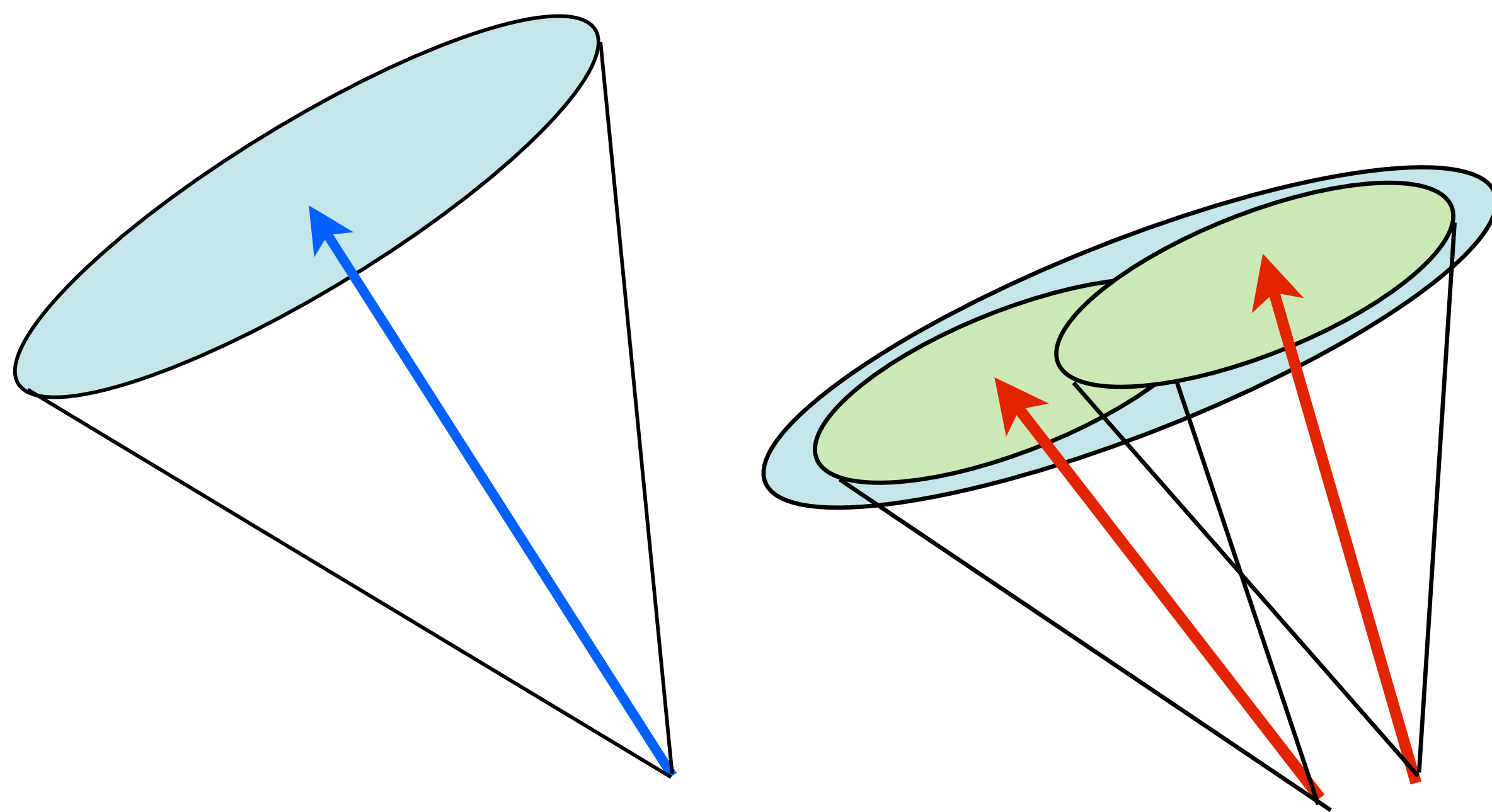
?



udsg/c/b

$\{\eta, \phi, p_T\}$
+
{tracking}

“flavor”-tagging:
b-tagging
c-tagging
uds-tagging



u/ds/g/c/b/W/Z/H/t/pu

quantum numbers:

color charge (quarks vs. gluons)

electric charge

spin

An explosion in the field of jet substructure and properties!

$\{\eta, \phi, p_T\}$
+
{tracking}
+
{m, shapes, subjets}

“flavor”-tagging:

b-tagging

c-tagging

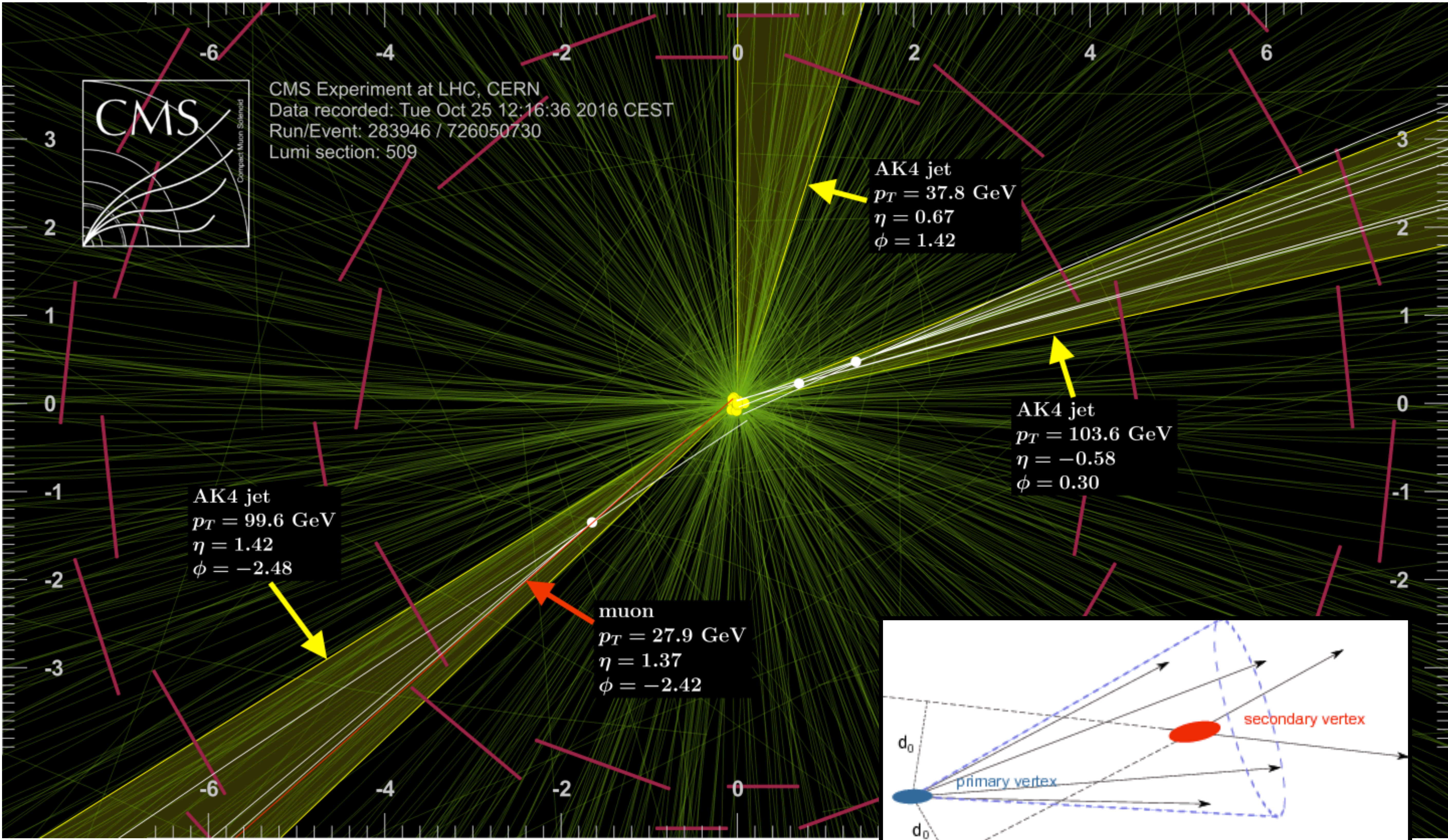
u/ds-tagging

top-tagging

W/Z/H-tagging

pileup-tagging

JETS WITH DISPLACED VERTICES



Displacement $\sim > O(\text{mm})$ scale

p_T, η, ϕ + tracking

mass

4-vector sum of jet constituents

highly sensitive to soft QCD and pileup; grooming can be used to mitigate these dependencies

substructure

several classes: declustering/reclustering, generalized jet shapes and energy flow, statistical interpretation, jet charge

algorithms

some combination of cuts on mass, shapes, tracking
most typical in top tagging

And nowadays ... machine learning too!

p_T, η, ϕ + tracking

mass

4-vector sum of jet constituents

highly sensitive to soft QCD and pileup; grooming can be used to mitigate these dependencies

substructure

several classes: declustering/reclustering, generalized jet shapes and energy flow, statistical interpretation, jet charge

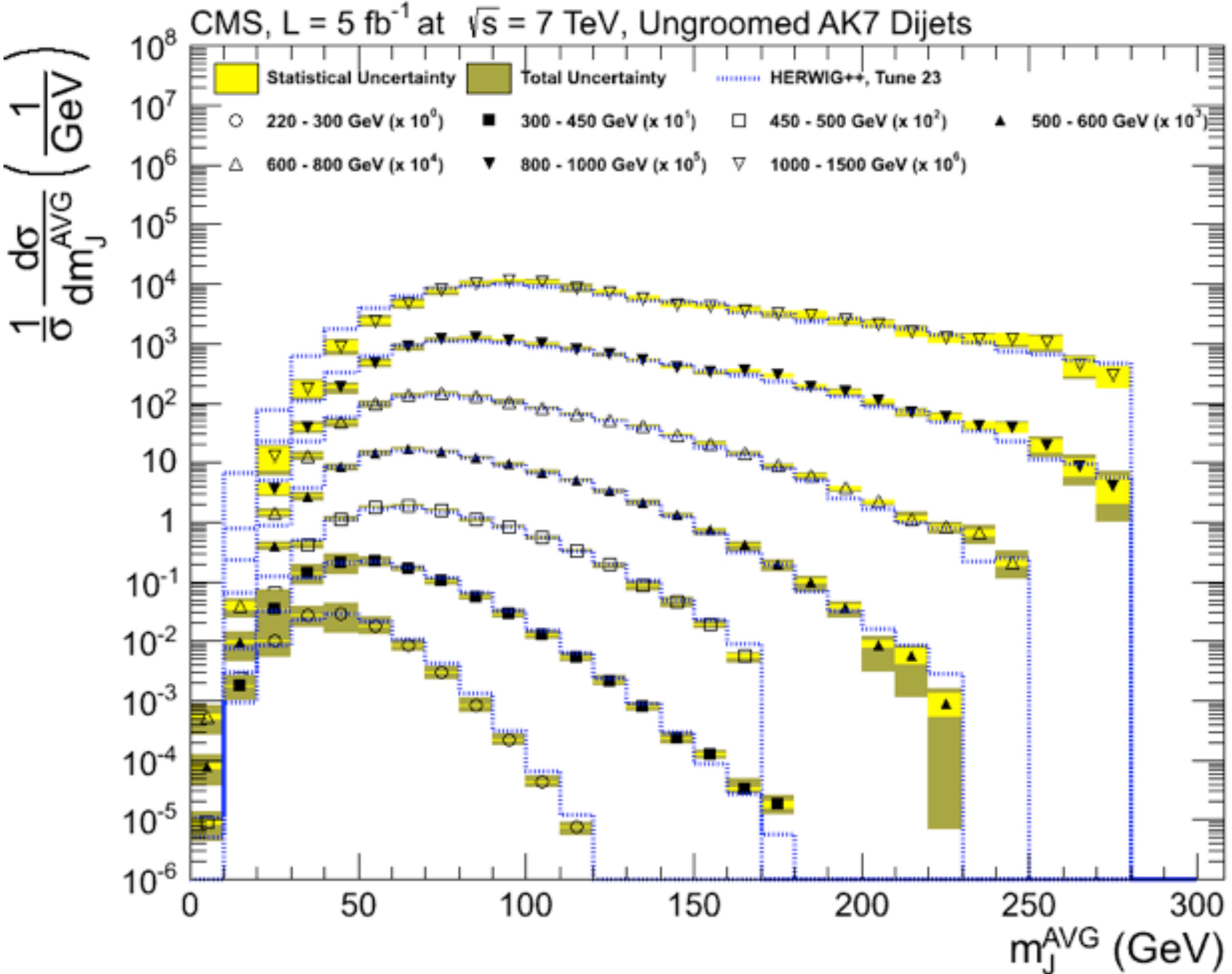
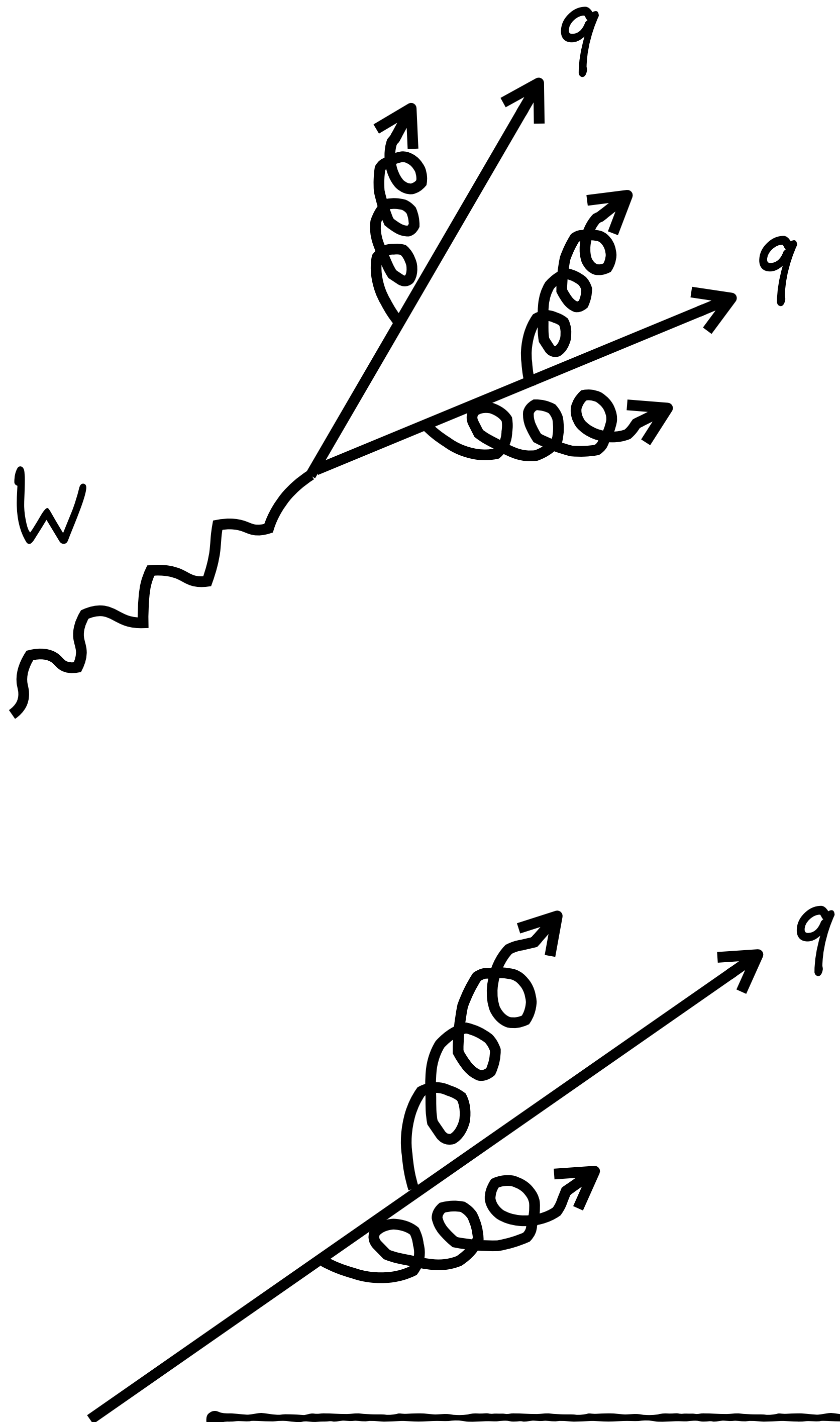
algorithms

some combination of cuts on mass, shapes, tracking
most typical in top tagging

And nowadays ... machine learning too!

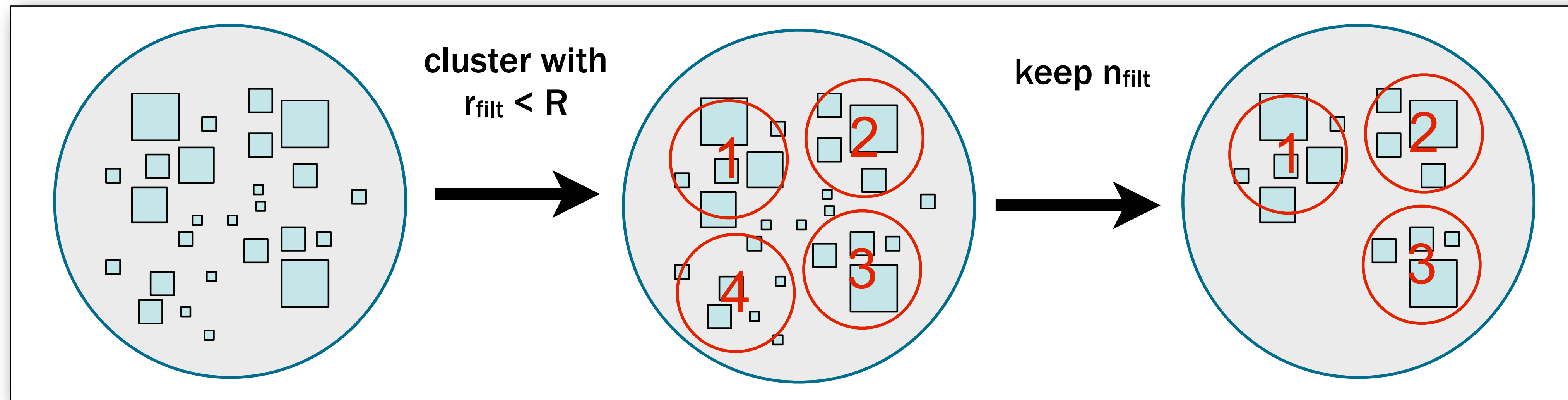
challenge: modeling the QCD
backgrounds is hard!

but jet mass is a perturbative quantity

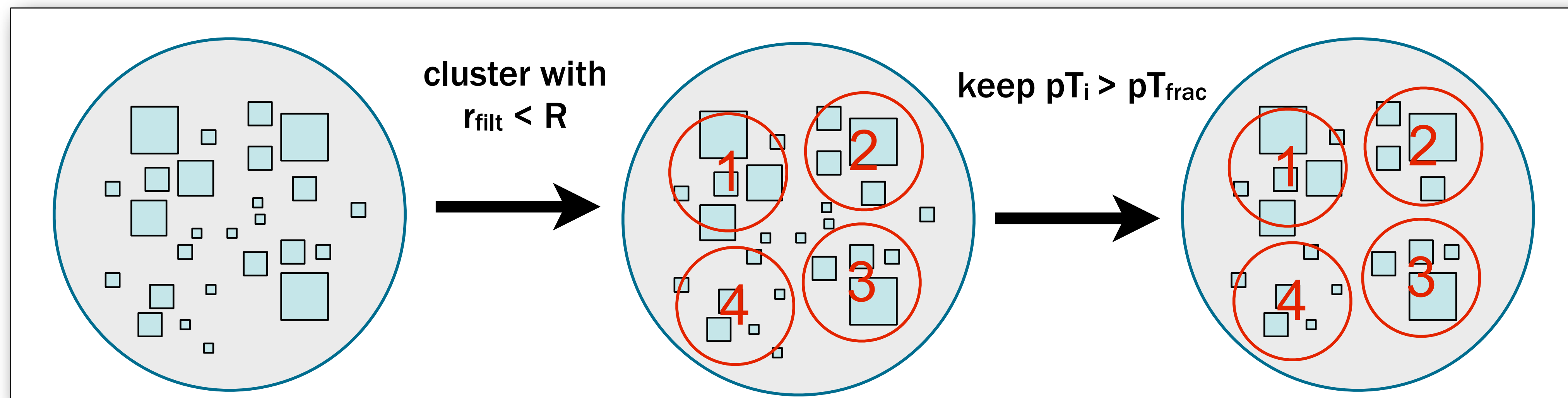


$$\langle M^2 \rangle \simeq \left. \begin{array}{l} \text{quarks: } 0.16 \\ \text{gluons: } 0.37 \end{array} \right\} \times \alpha_s p_t^2 R^2$$

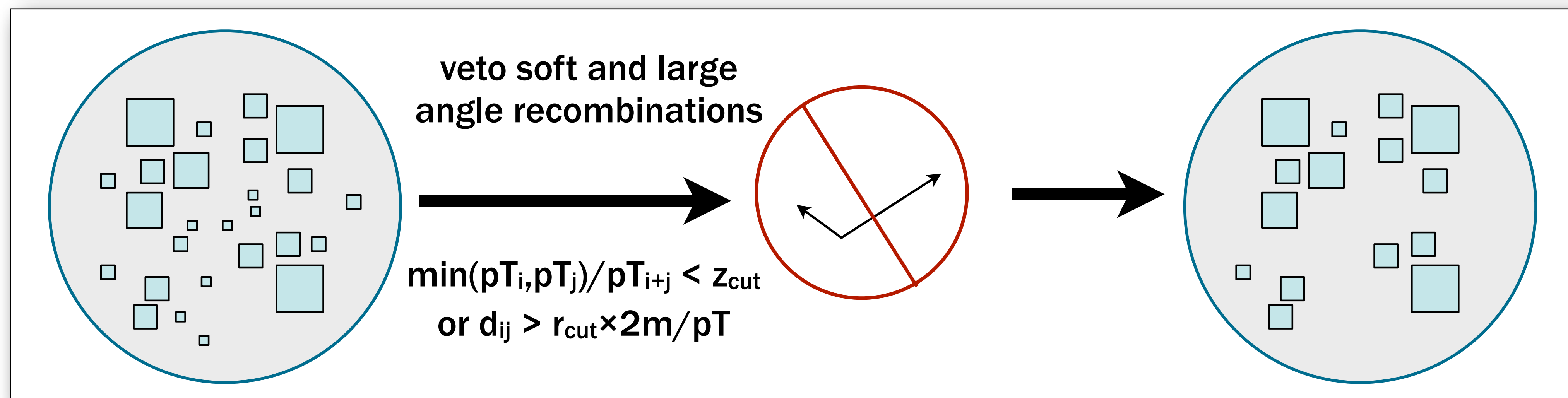
Filtering



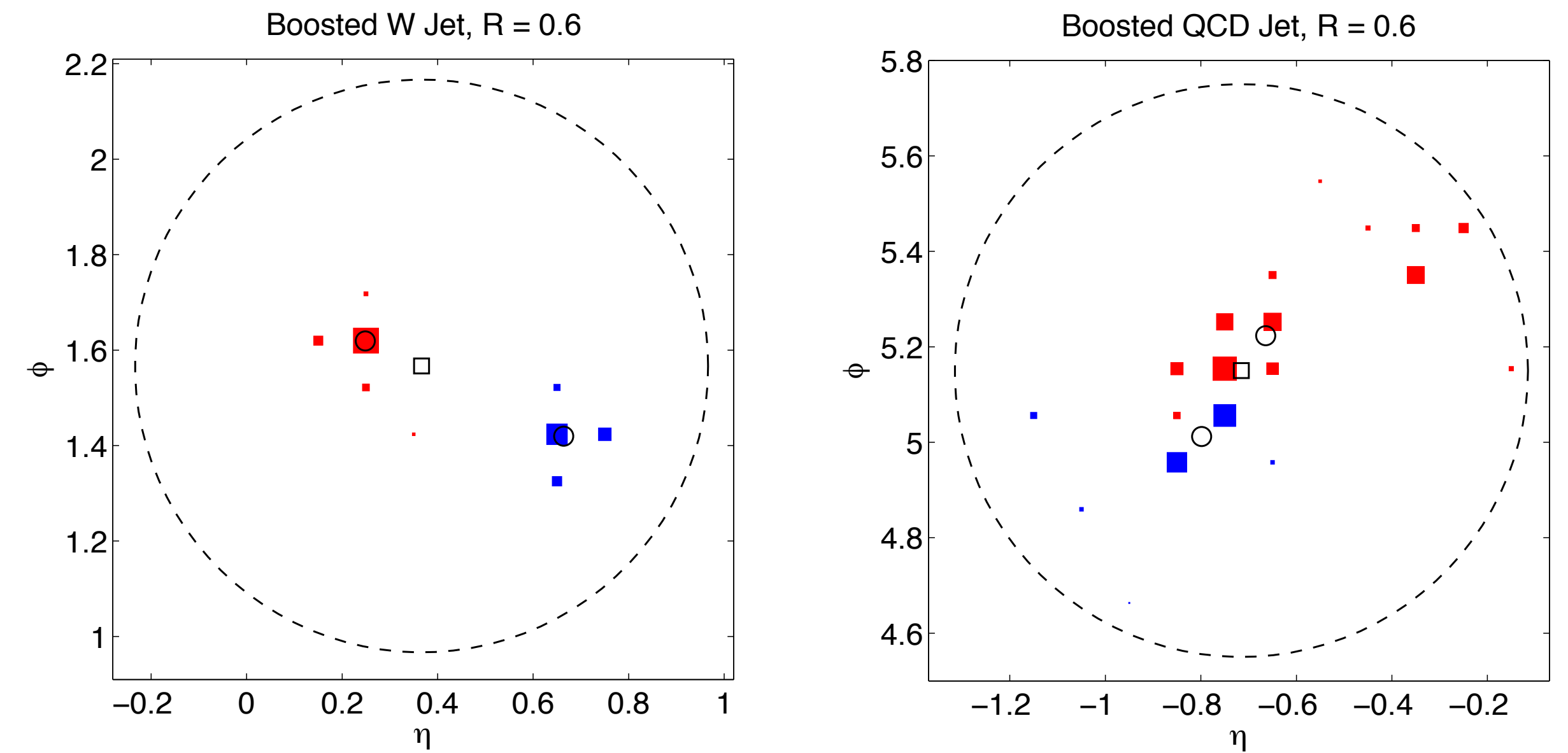
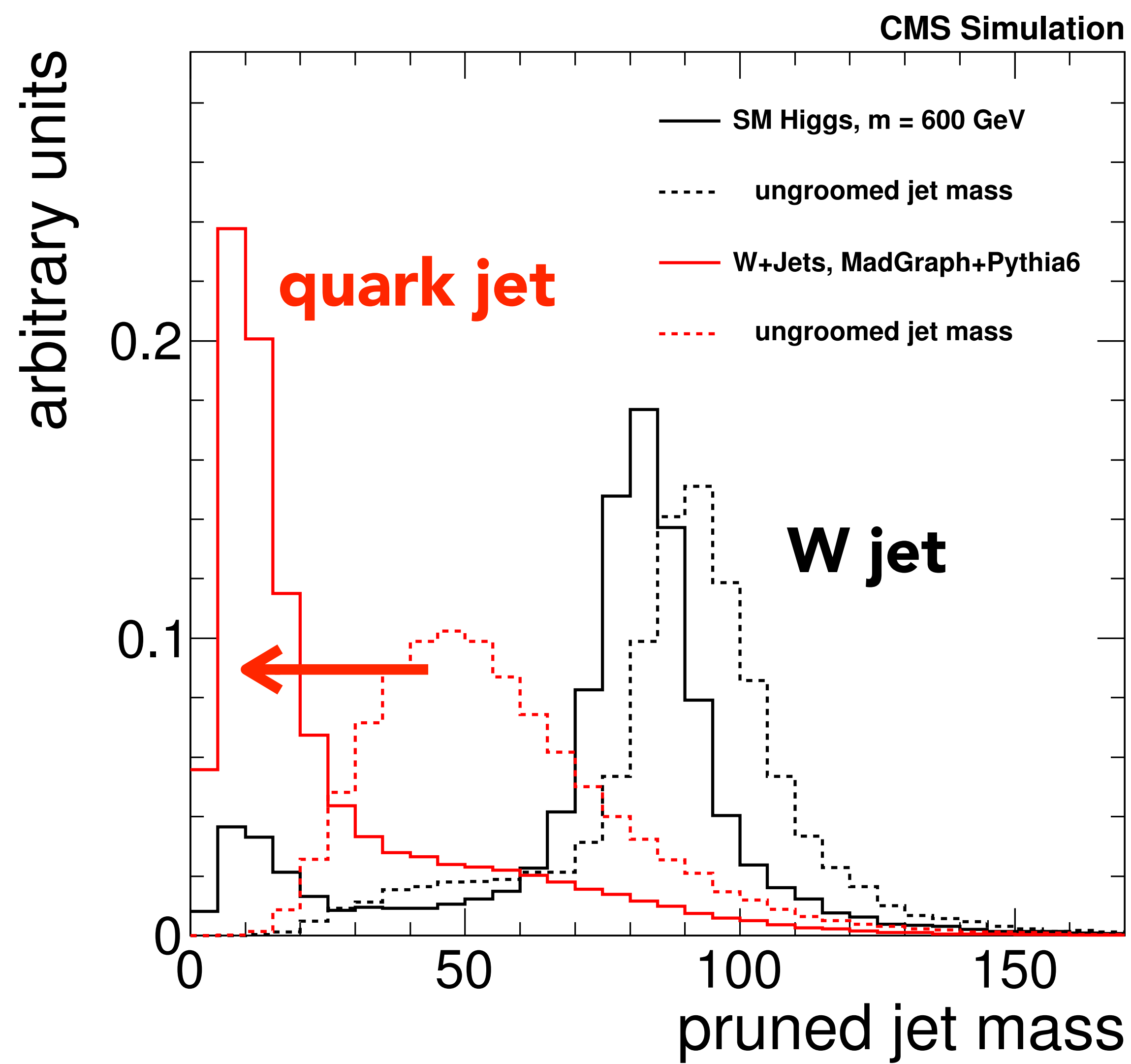
Trimming



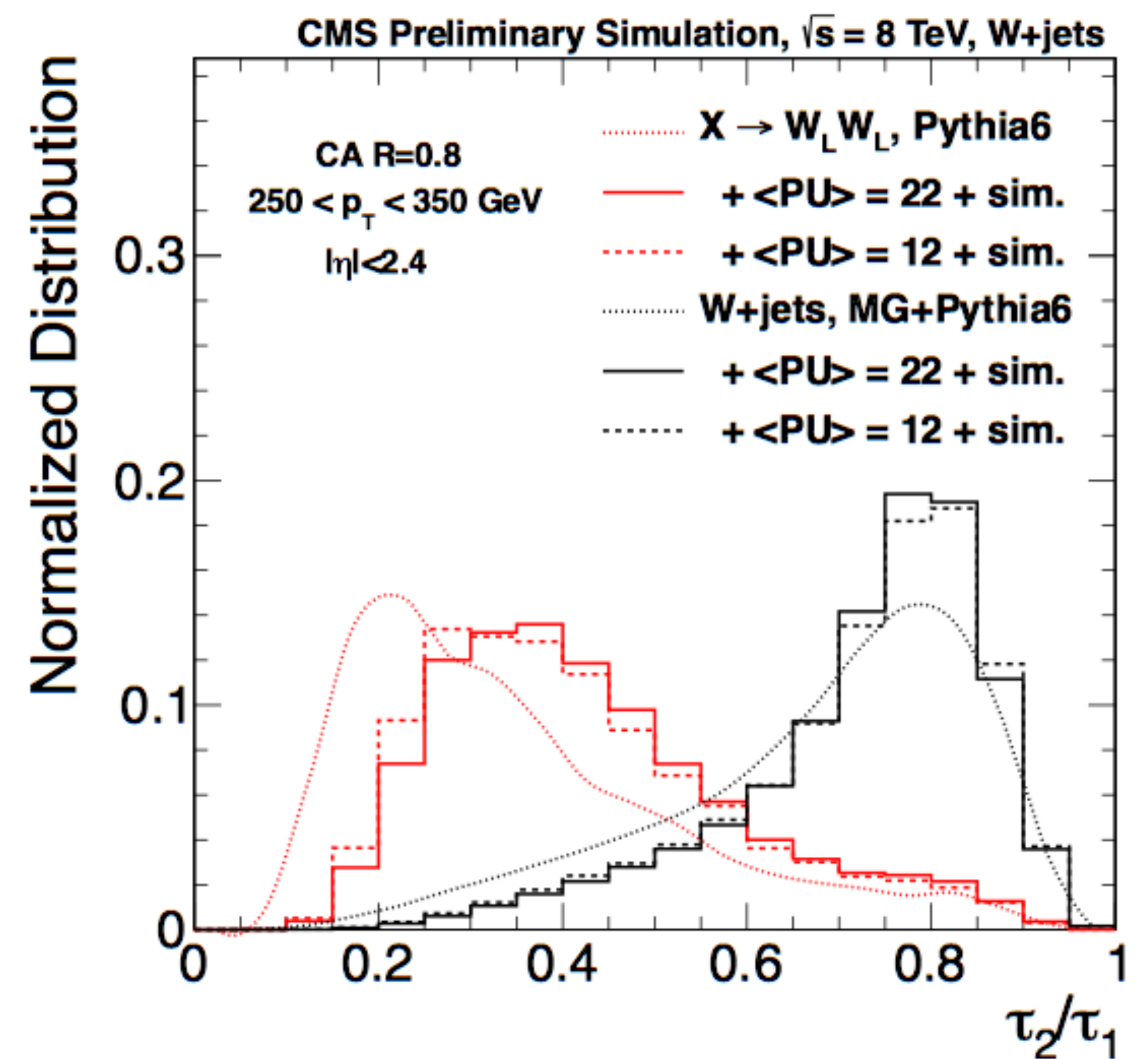
Pruning



jet (groomed) mass:
a very powerful discriminator



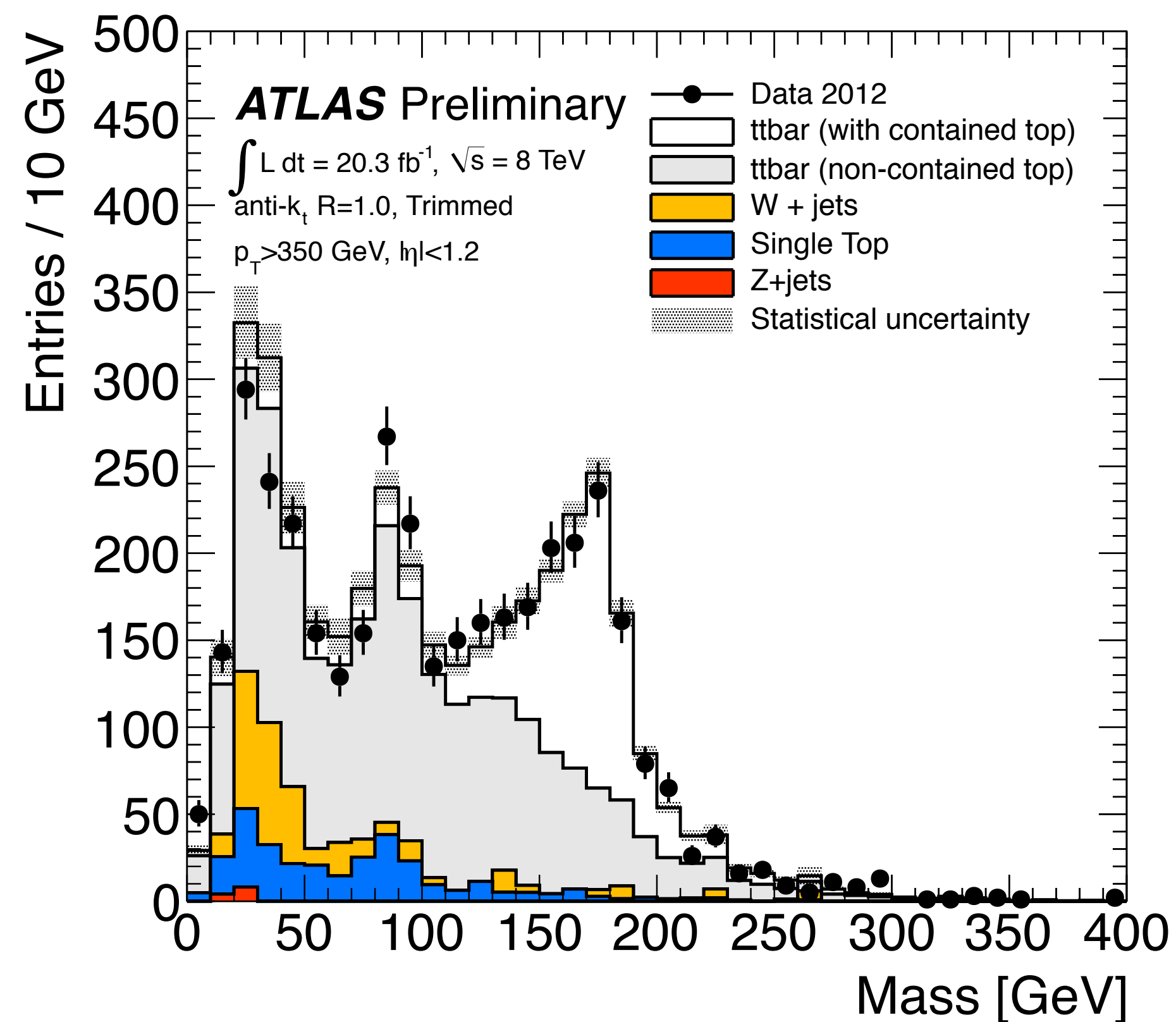
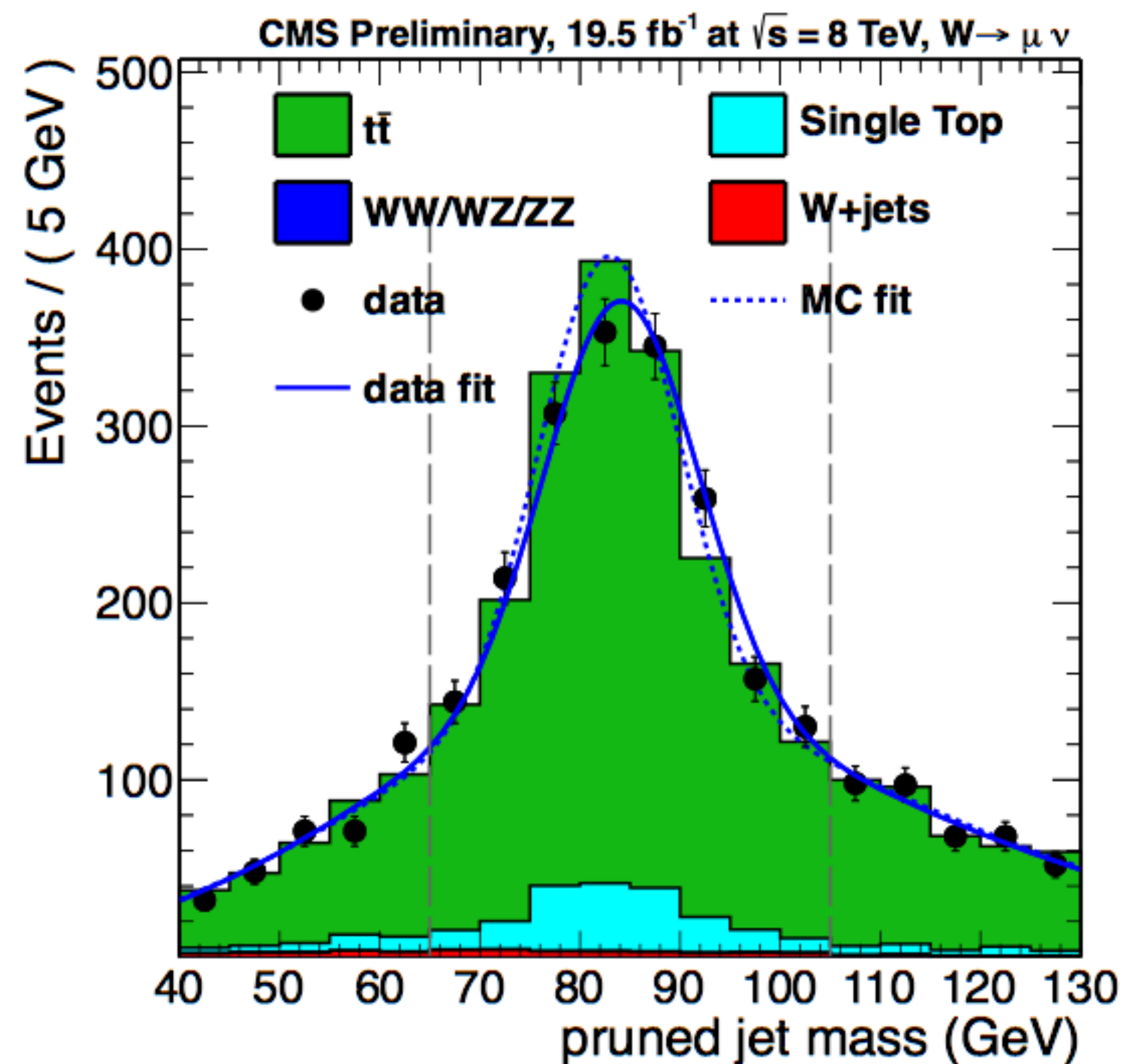
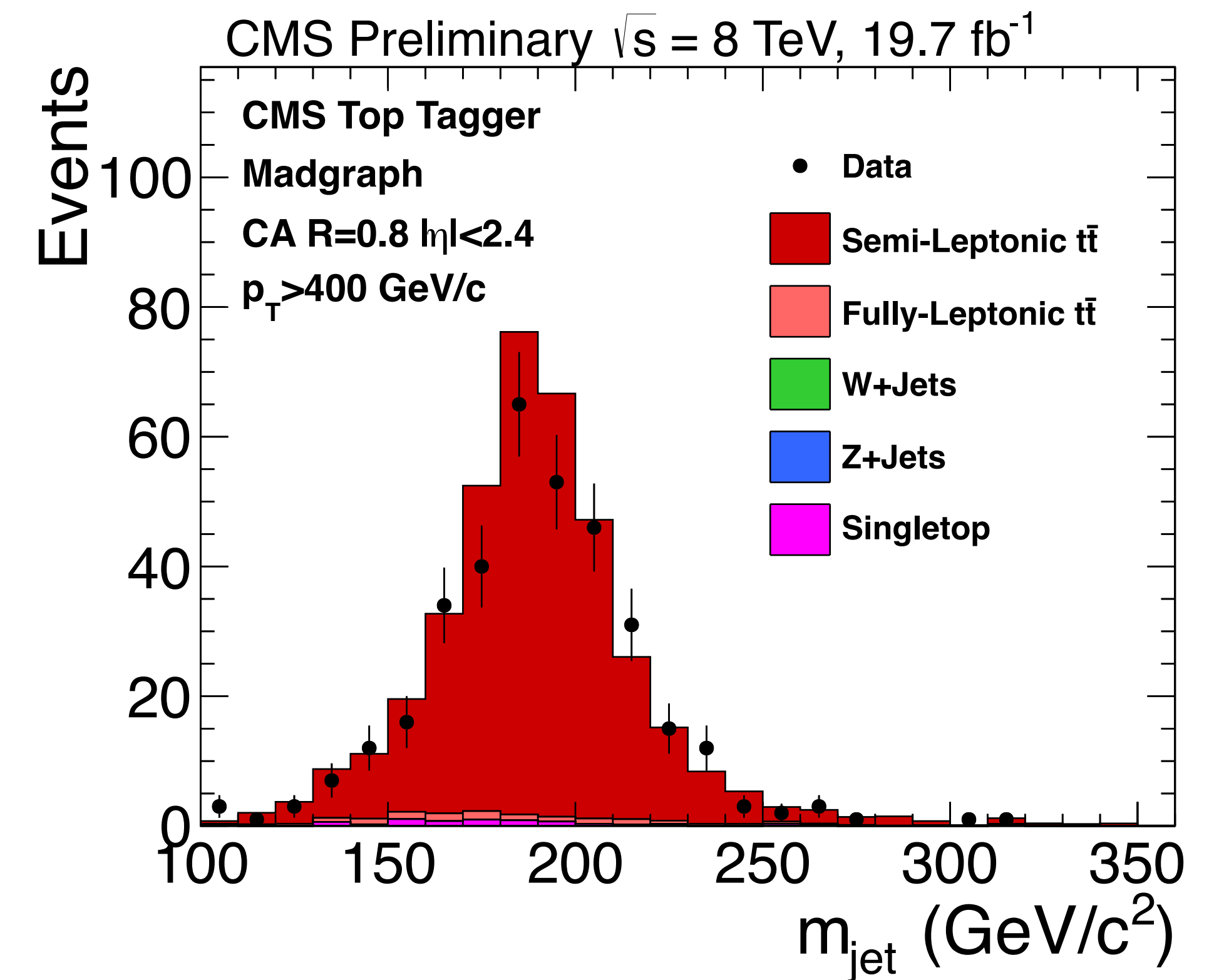
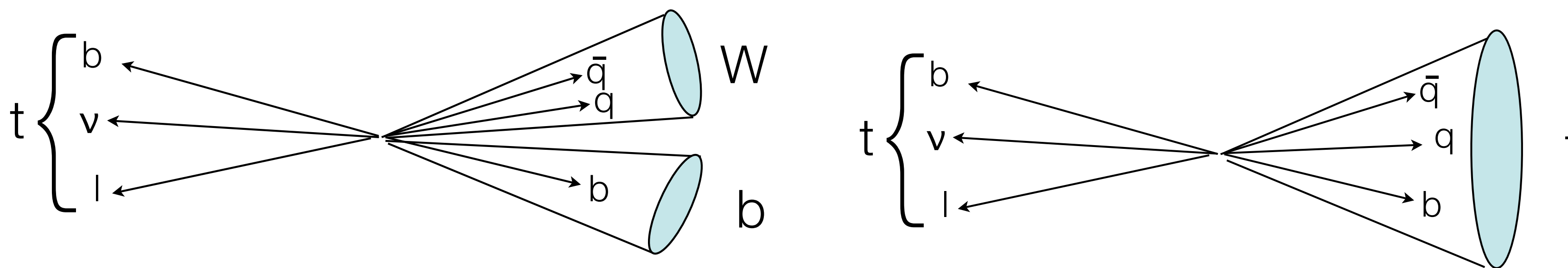
how prong-y are these jets?



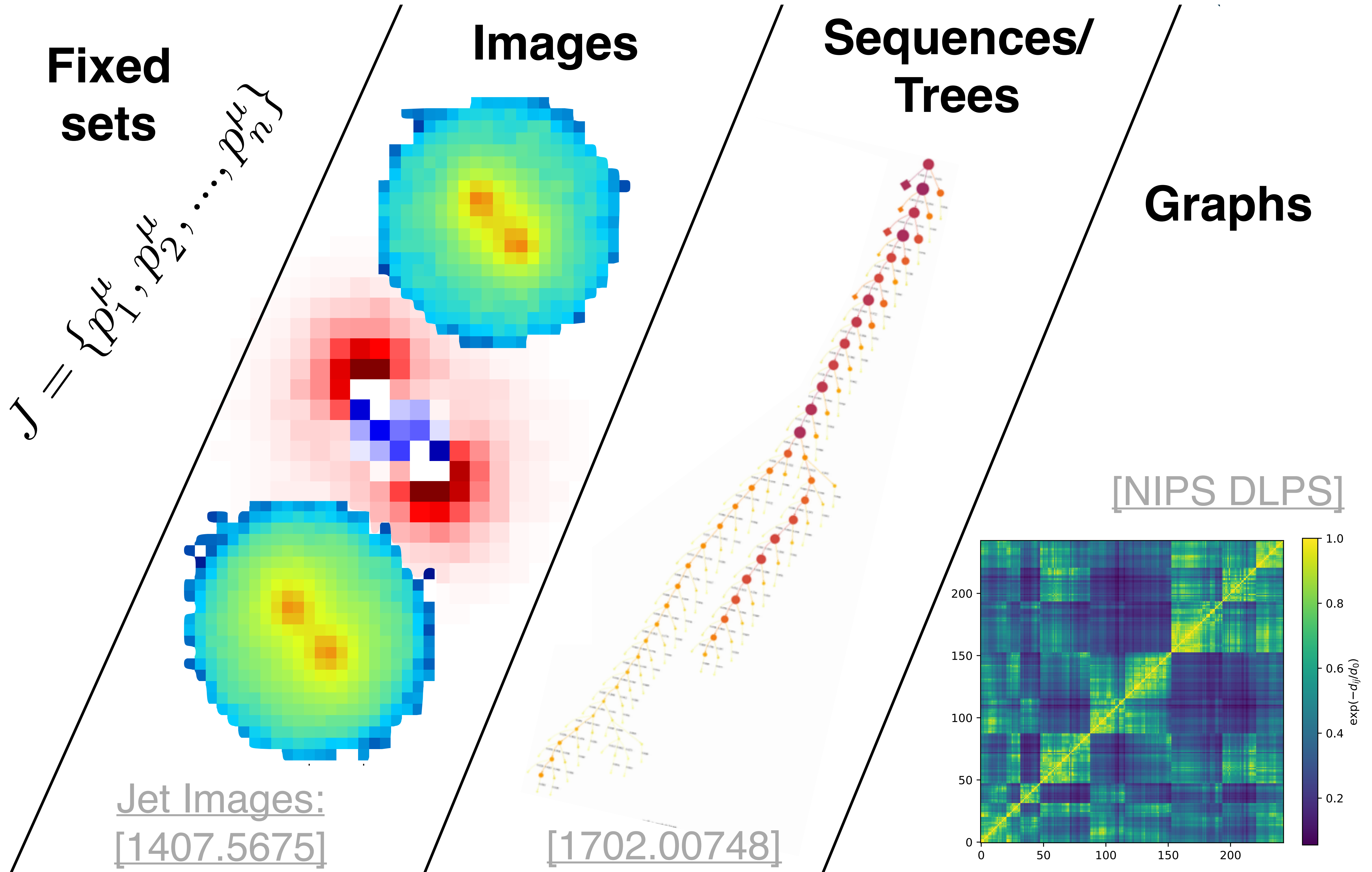
example: N-subjettiness

Example:

Semi-leptonic $t\bar{t}$ events are important for validating tagging techniques of heavy objects



Using events to do
 evaluate tag-and-probe
 efficiency **scale factors**
 and **mass scale/resolution**
 measurements



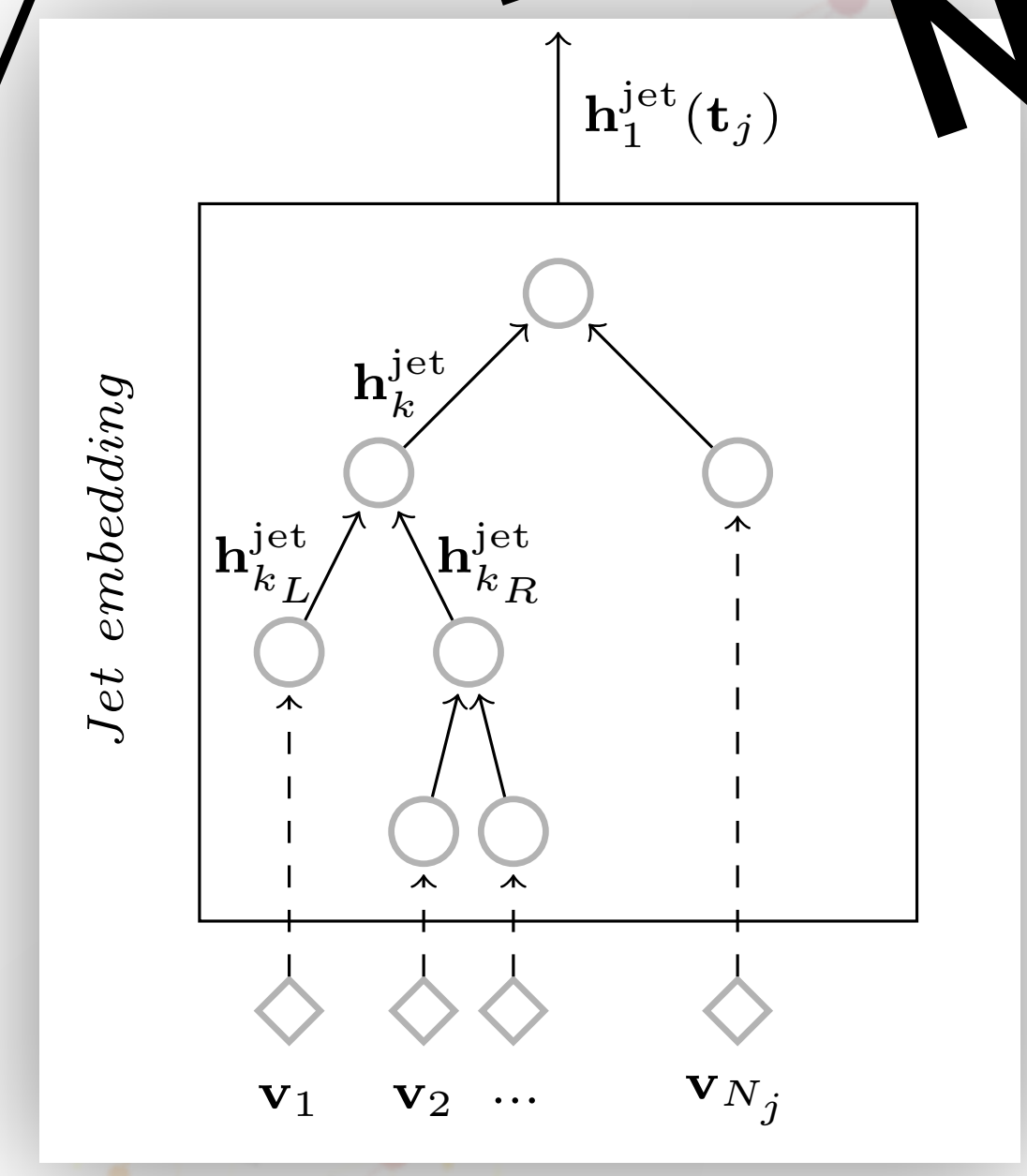
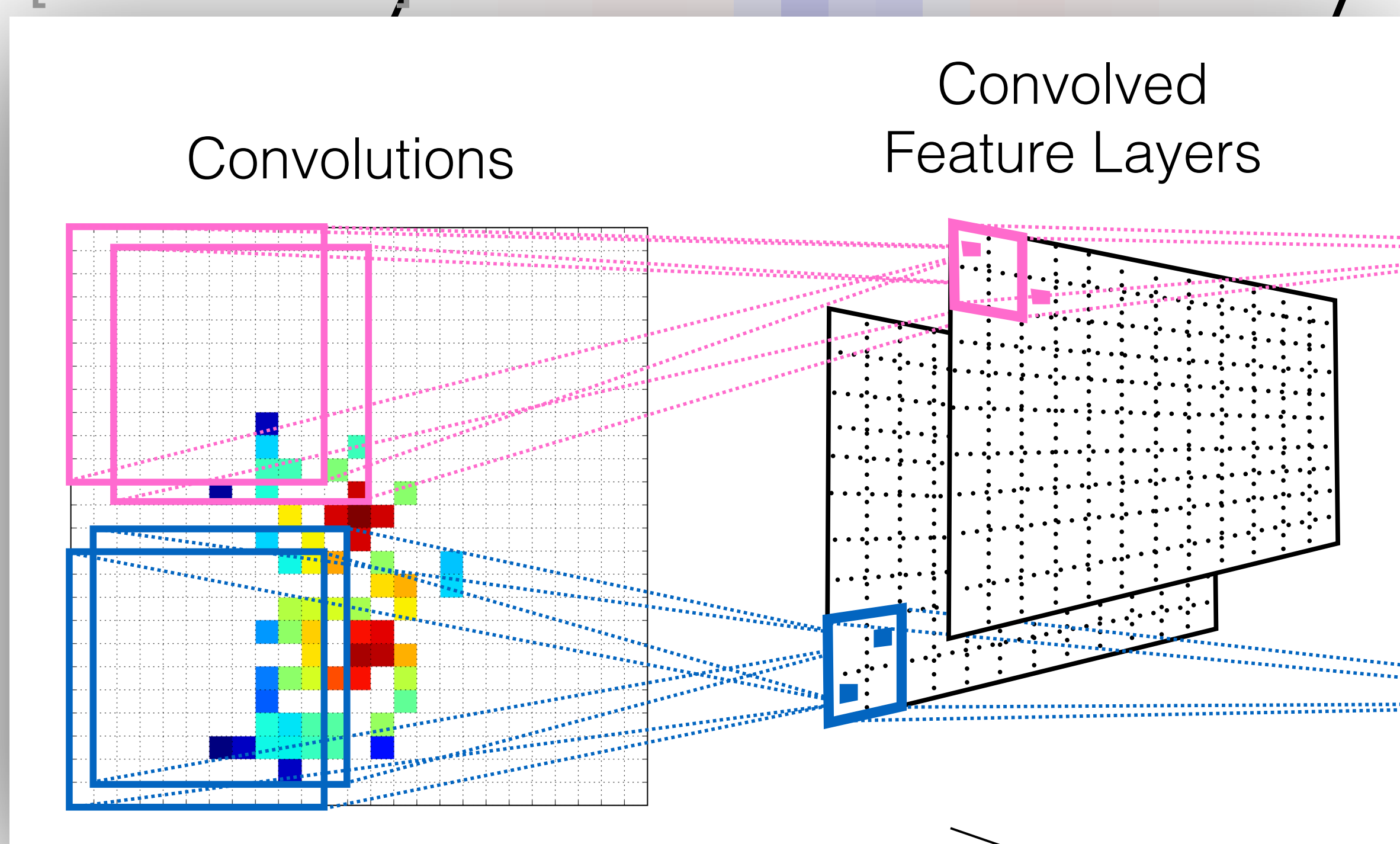
Fixed
**Dense
network**
(...or BDT/SVM/
Fisher/etc.)

Images
**Convolution
Neural
Network
(CNN)**

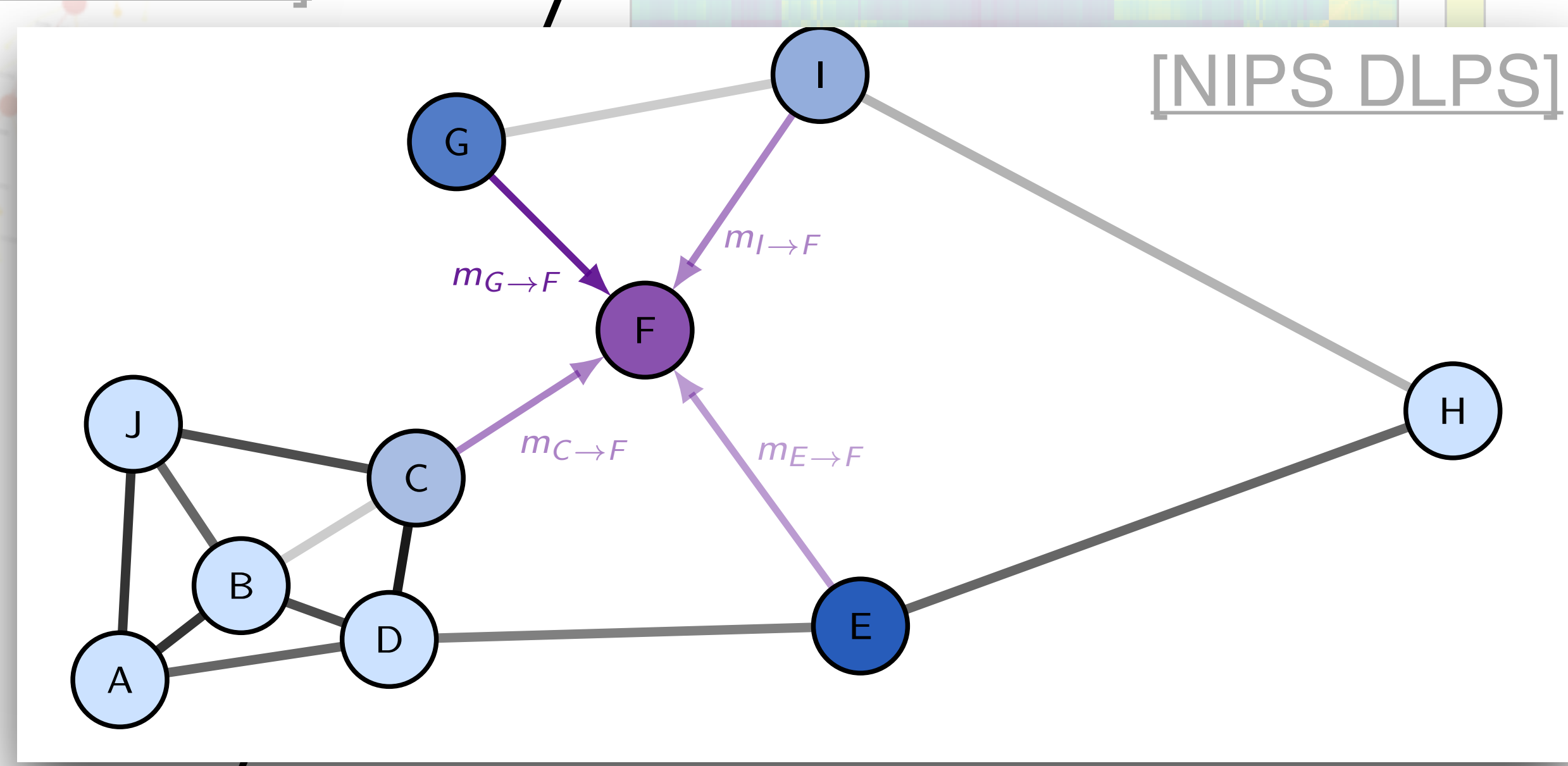
Sequences/
Trees
**Recurrent/
Recursive
NN**

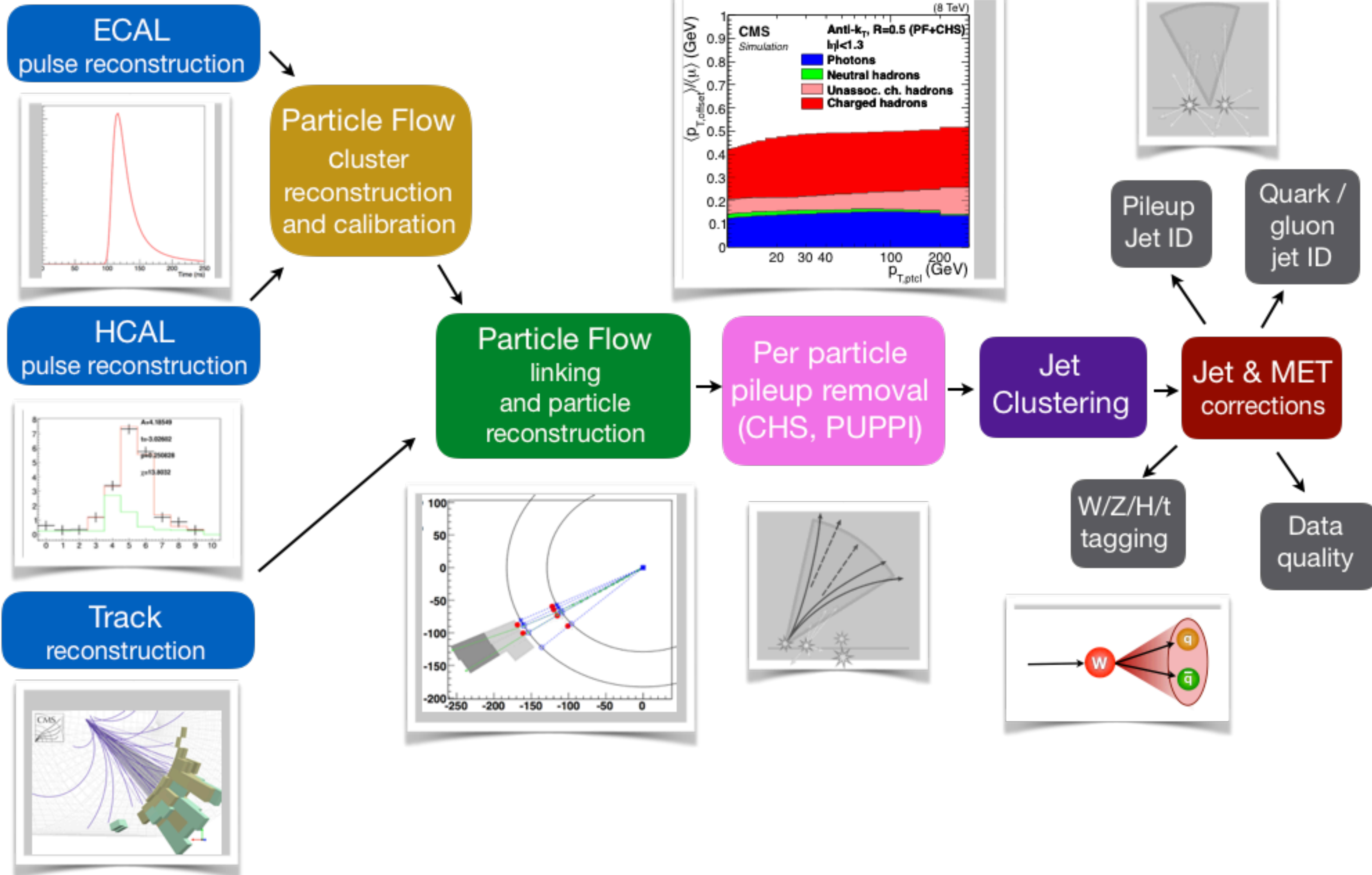
Graphs

**Graph
CNN**



[1702.00748]





searches for new physics at the kinematic limit **require** jet substructure techniques

*jet substructure is **characterizing radiation***

[jets are just an organizing principle]

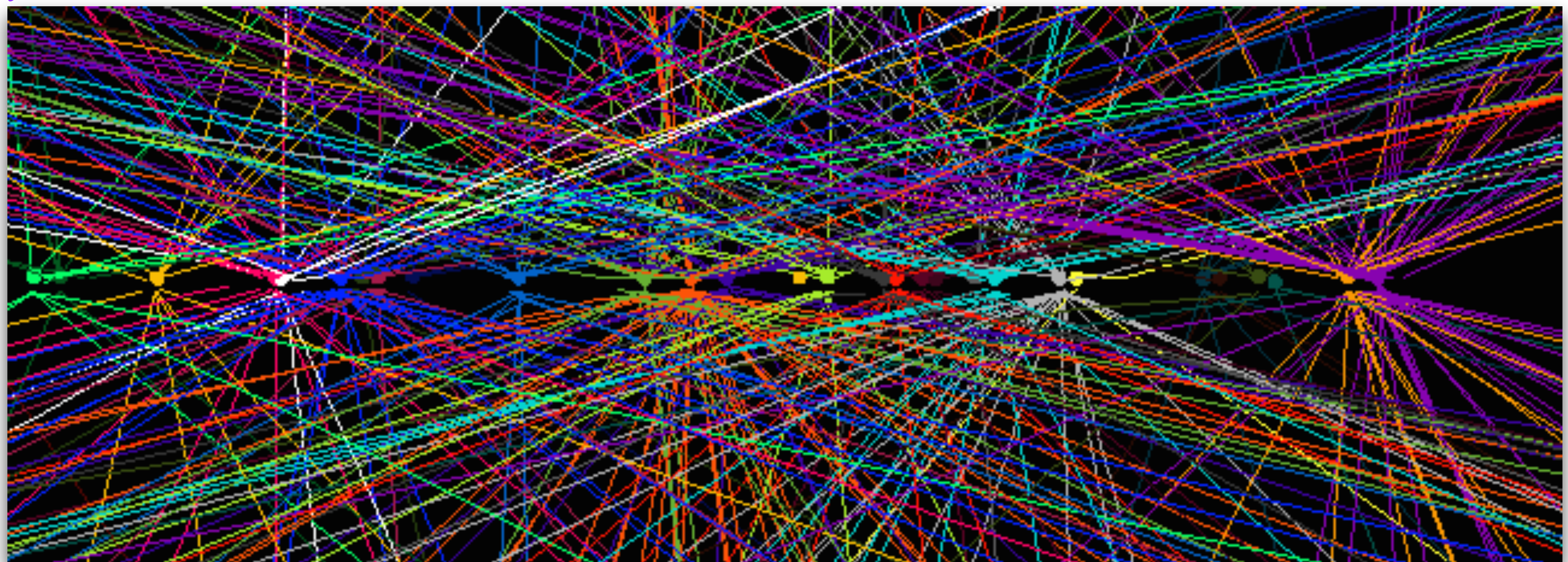
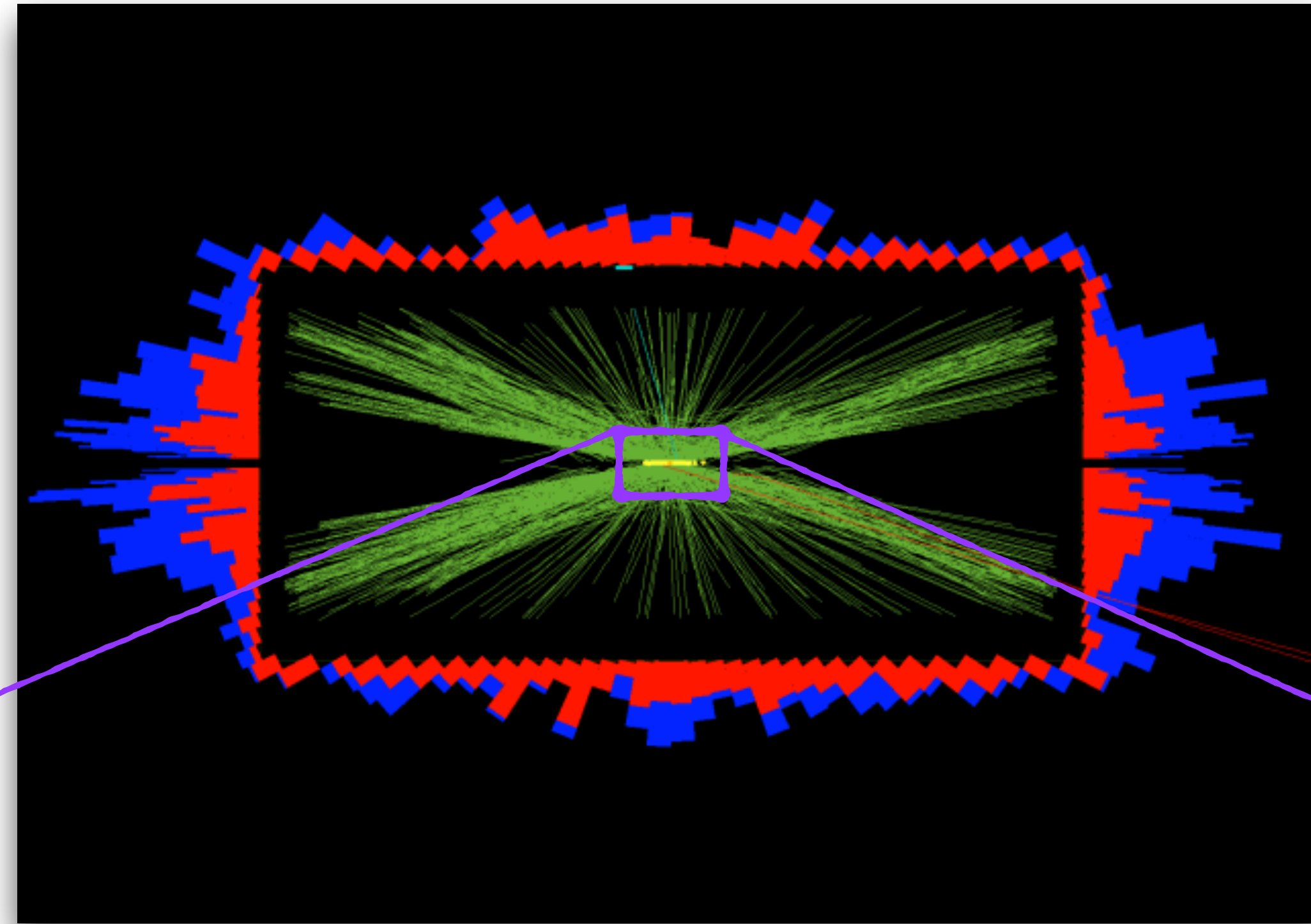
understanding radiation affects everything

e.g. jet substructure \leftrightarrow pileup mitigation

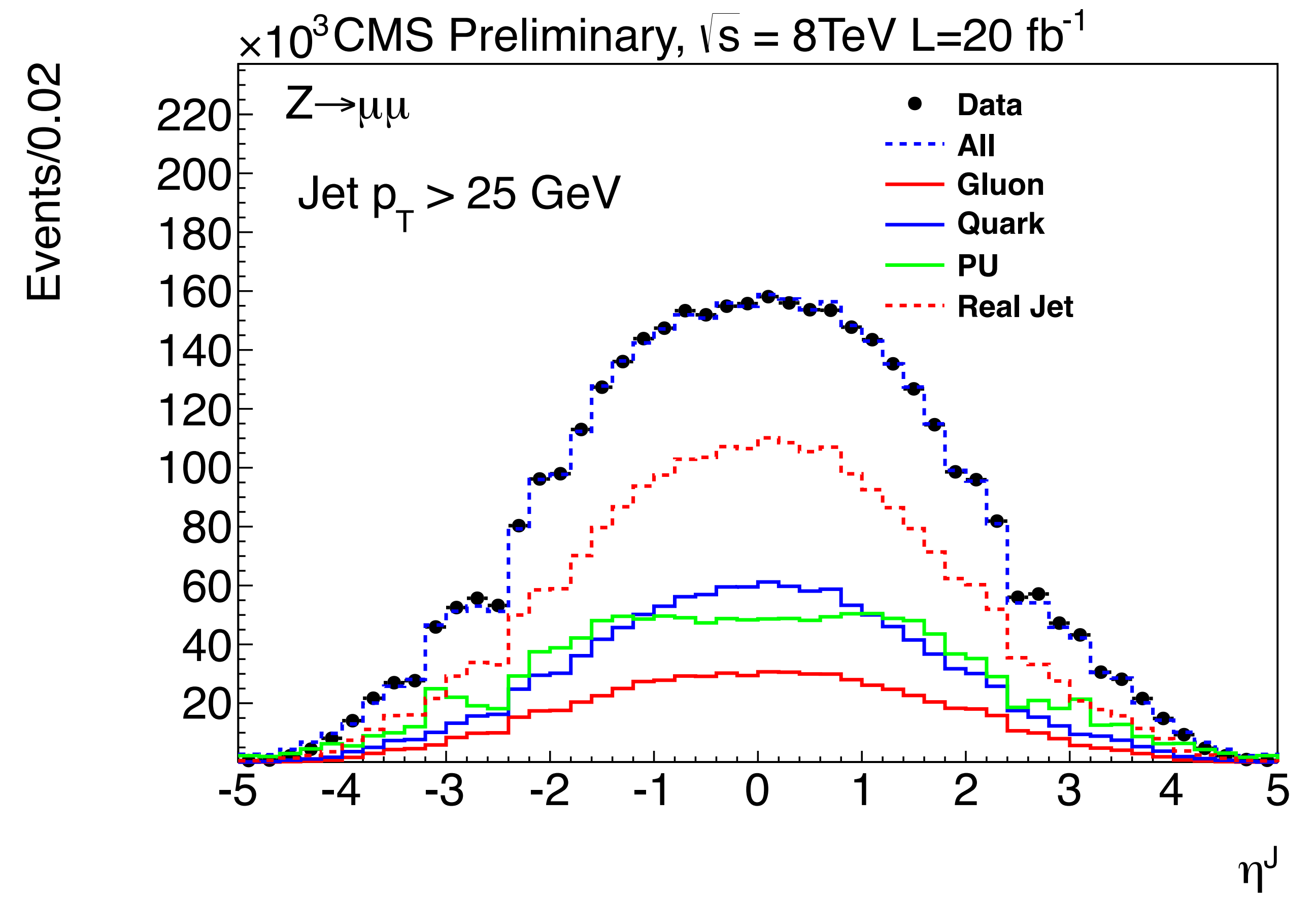
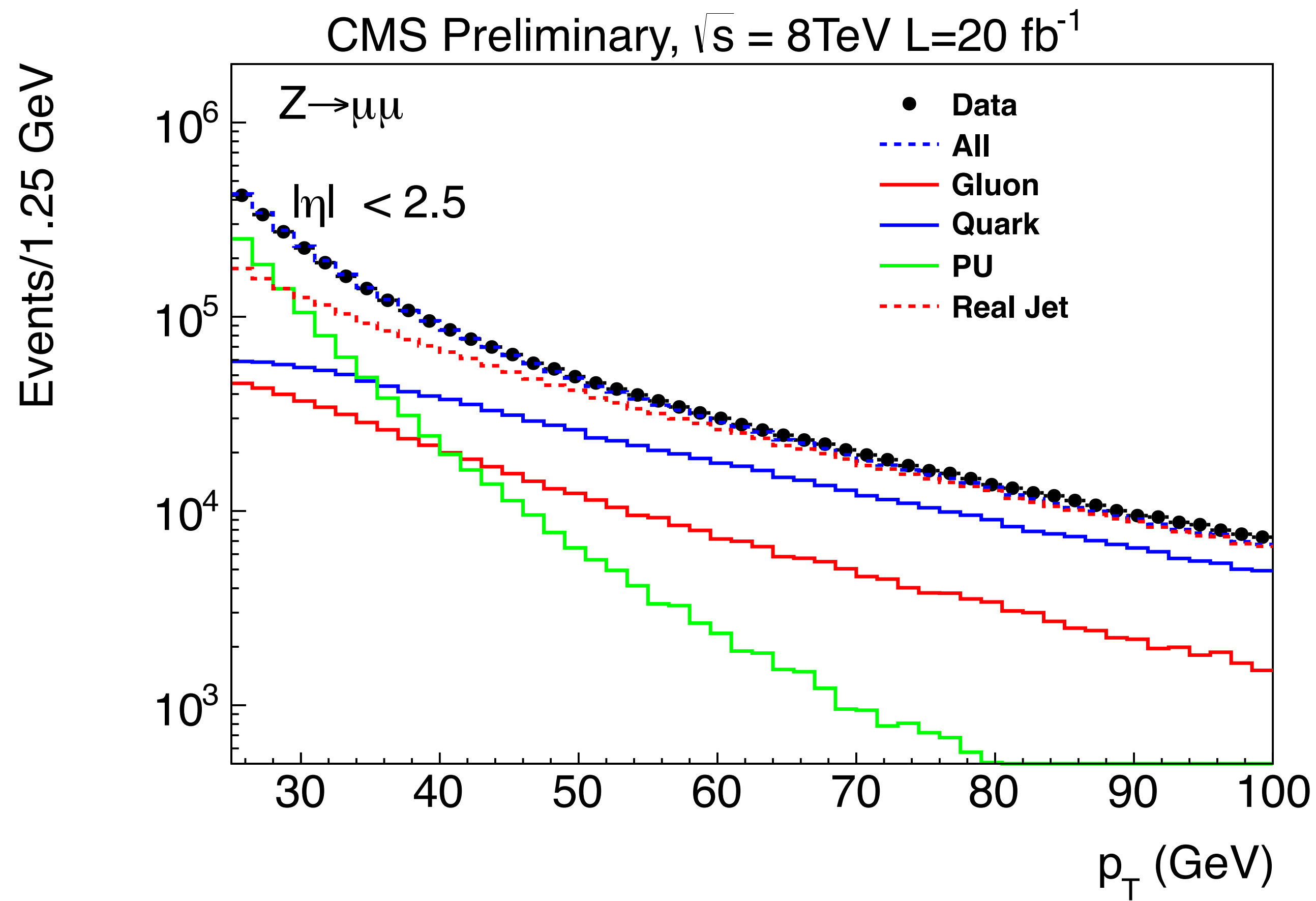
physics at intermediate (Higgs scale) energies are **more affected by pileup**

WHAT IS PILEUP?

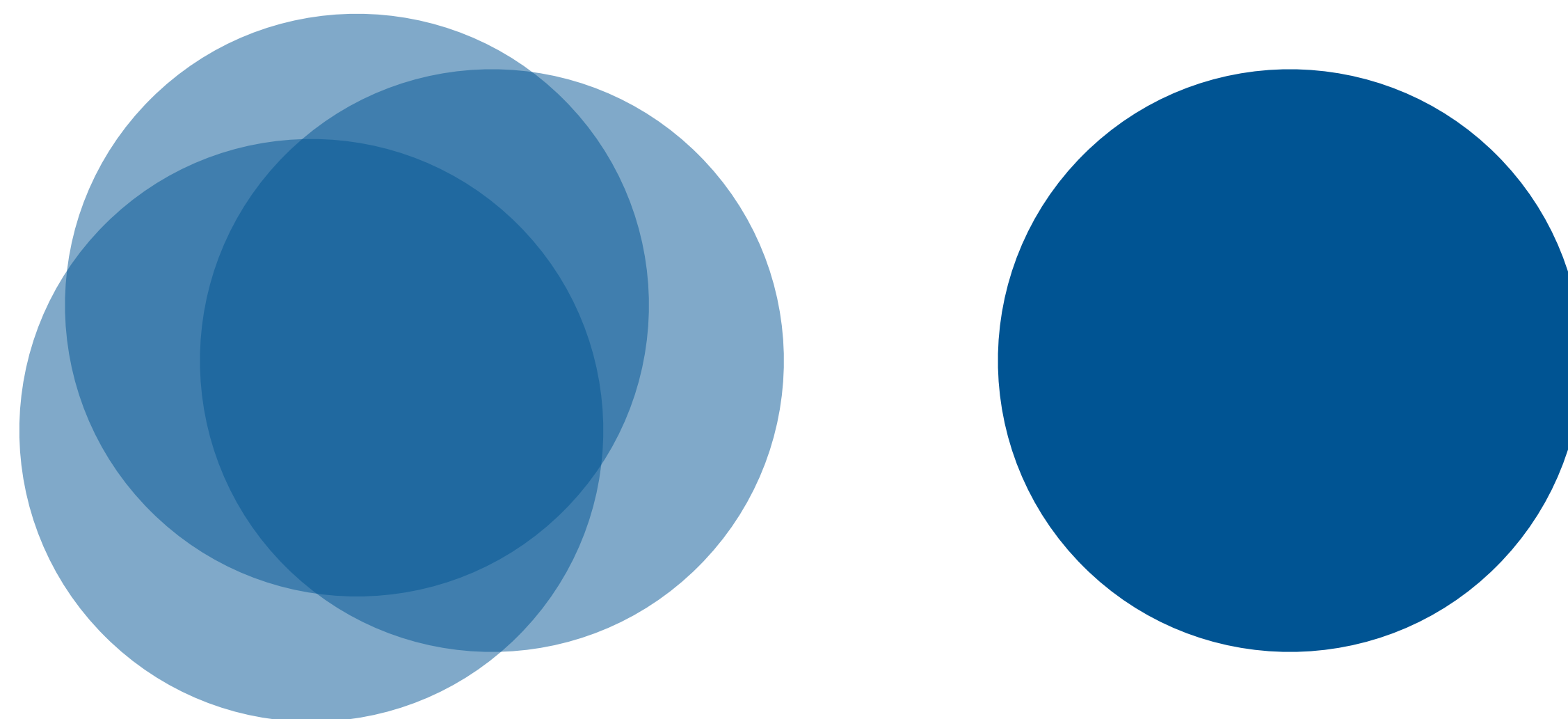
additional pp interactions that occur in each beam crossing because the instantaneous bunch-by-bunch collision luminosity is very high



WHAT IS PILEUP?



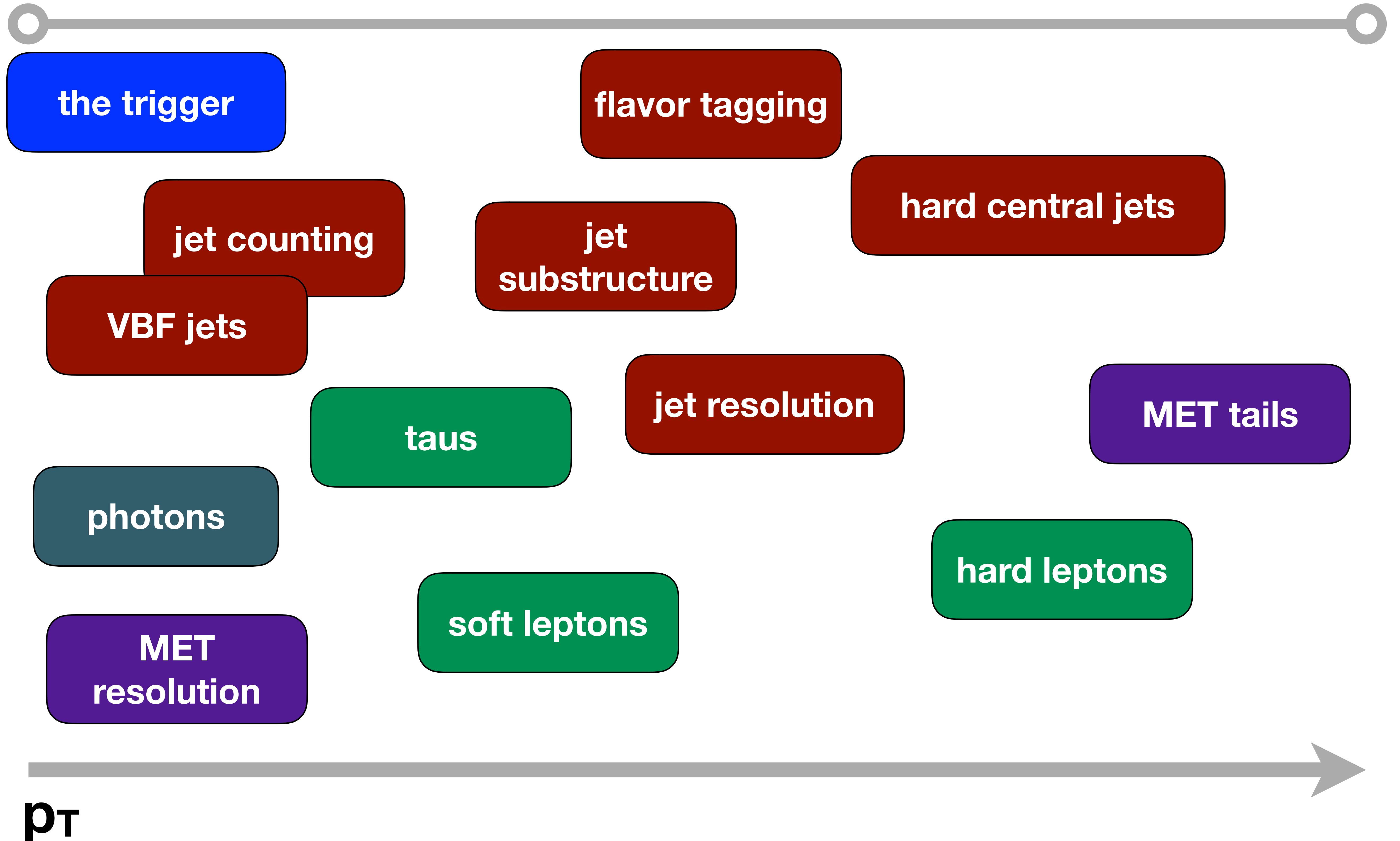
“stochastic” vs. “hard” pileup jets



both contribute to pileup, it's not necessarily either/or

pileup matters

pileup doesn't matter



properties

asymptotic behavior

local shape

tracking/vertexing

precision timing

depth segmentation

techniques

(apologies, not a complete list!)

ρ correction/subtraction

(area, 4-vector, shape, particle)

grooming

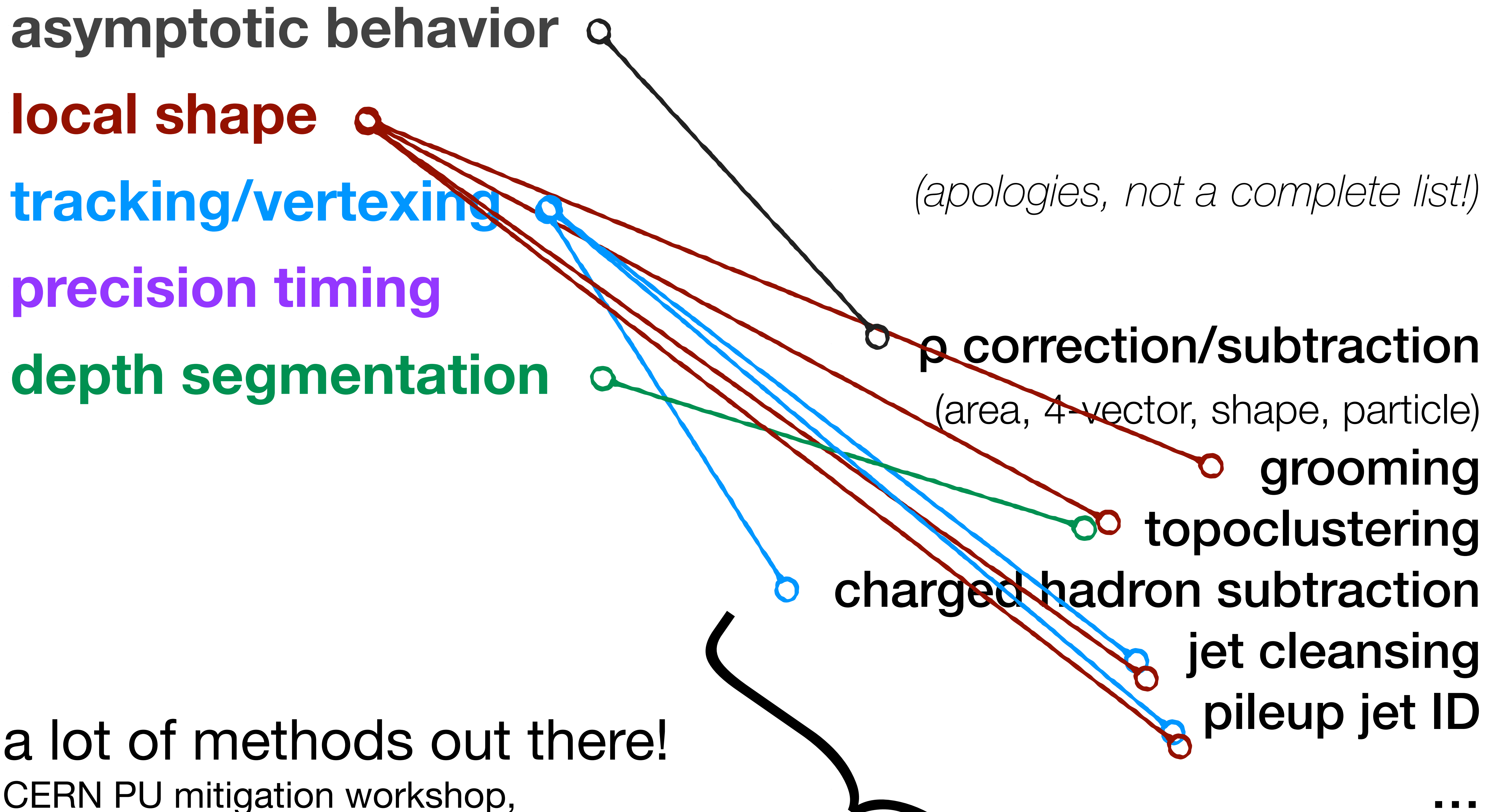
topoclustering

charged hadron subtraction

jet cleansing

pileup jet ID

...



a lot of methods out there!

CERN PU mitigation workshop,
an early exploration of methods
<https://indico.cern.ch/event/306155/>

“ ρ subtraction”

jet pt correction =
median energy density x area

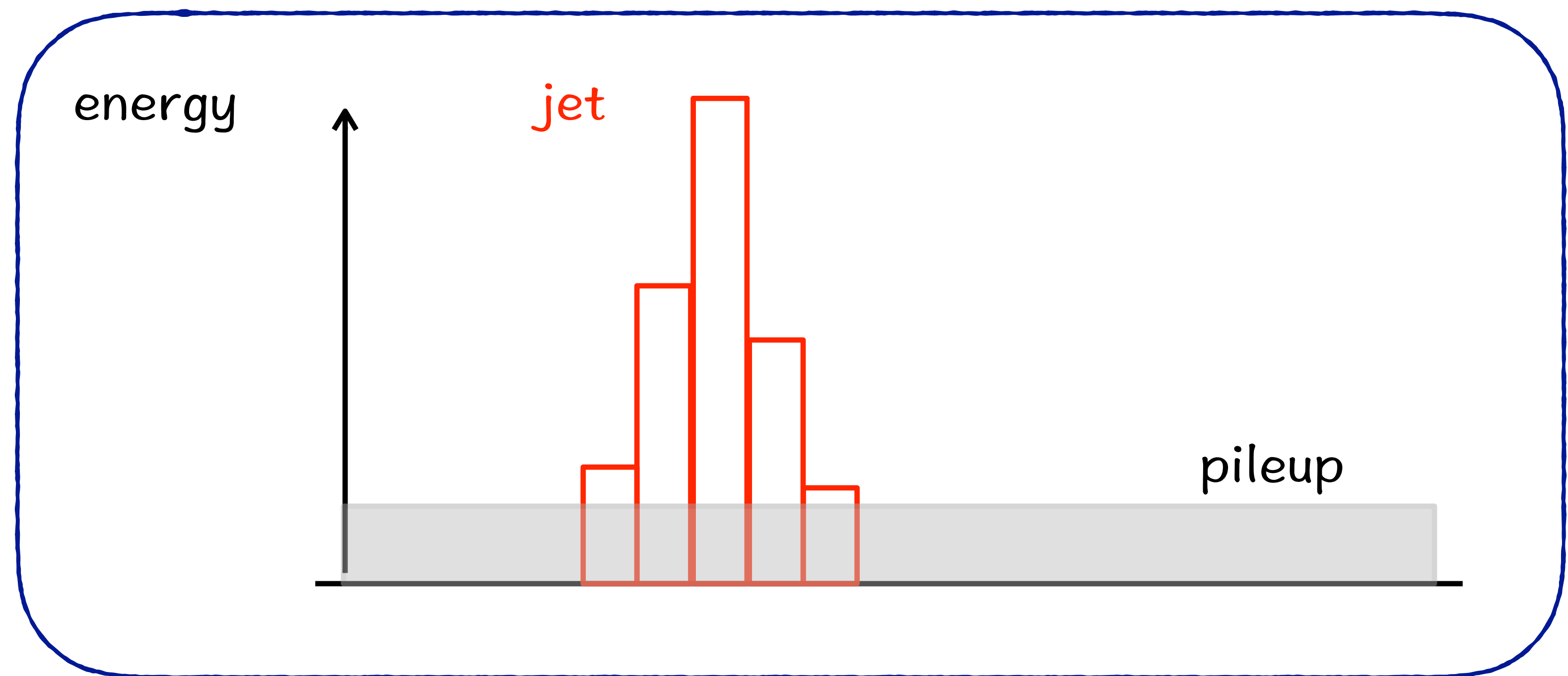
asymptotic behavior

local shape

tracking/vertexing

precision timing

depth segmentation



many variations of this method, including for jet shapes

Modification of the lepton isolation variable in PU

$$I_{\Delta\beta}^{\mu} = \frac{\sum p_T^{\text{CH-PV}} + \max\left(0, \sum p_T^{\text{NH}} + \sum p_T^{\gamma} - \frac{1}{2} \sum p_T^{\text{CH-PU}}\right)}{p_T^{\mu}}$$

Using the charged-to-neutral ratio (2/3 vs. 1/3) and vertexing information

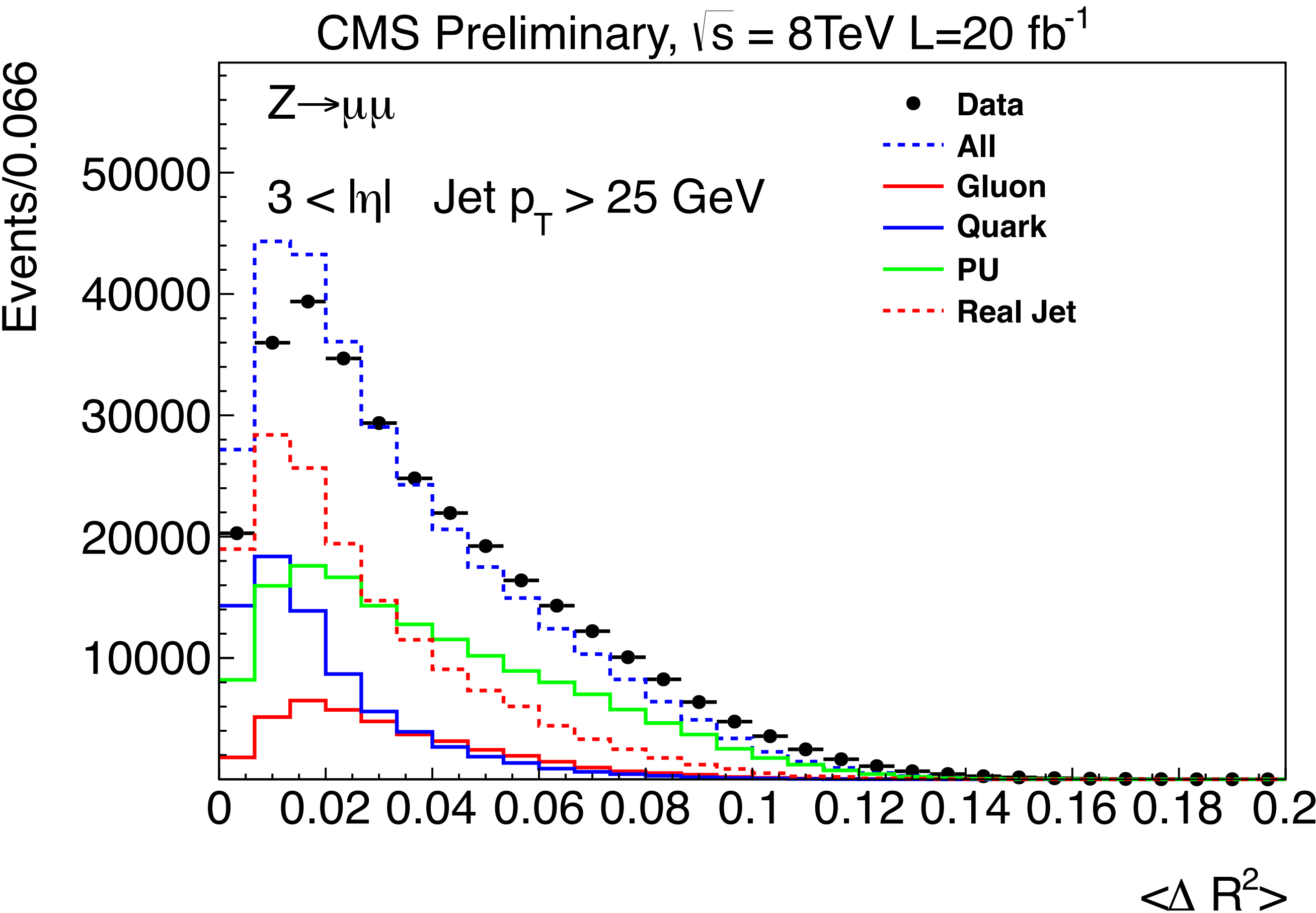
asymptotic behavior

local shape

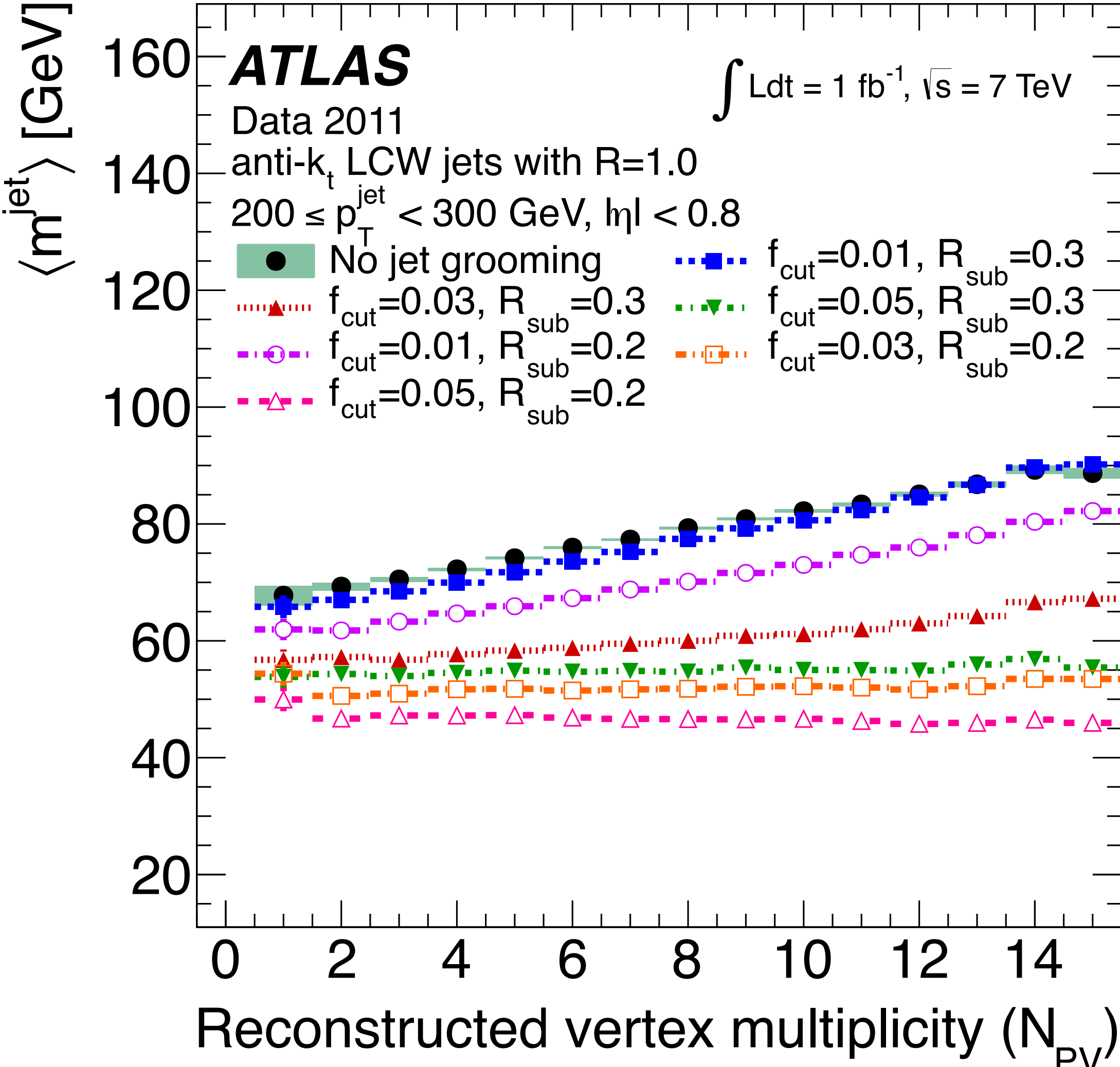
tracking/vertexing

precision timing

depth segmentation



jet grooming, cleans up soft and wide-angle radiation



“jet RMS” of forward pileup jets

asymptotic behavior

local shape

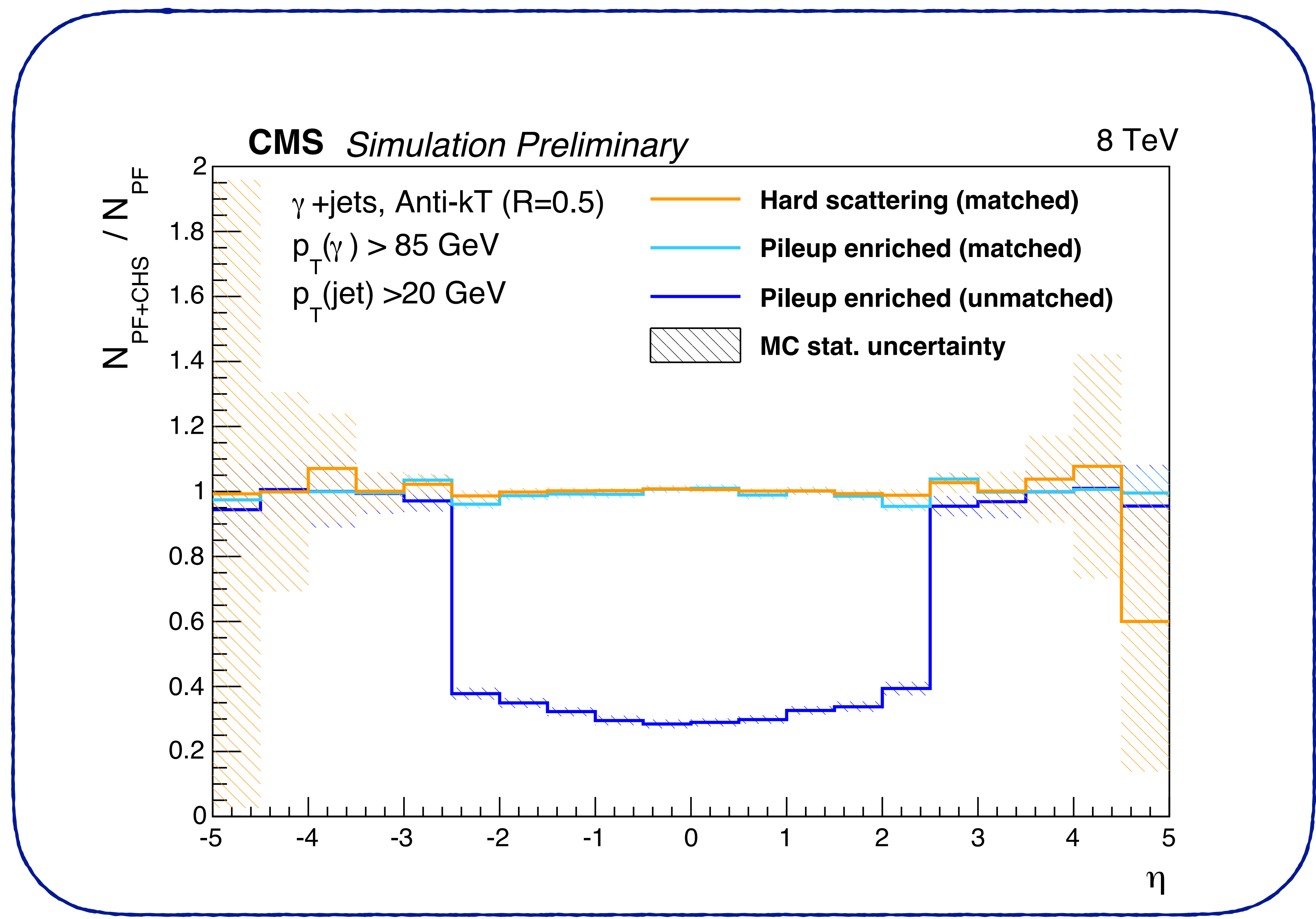
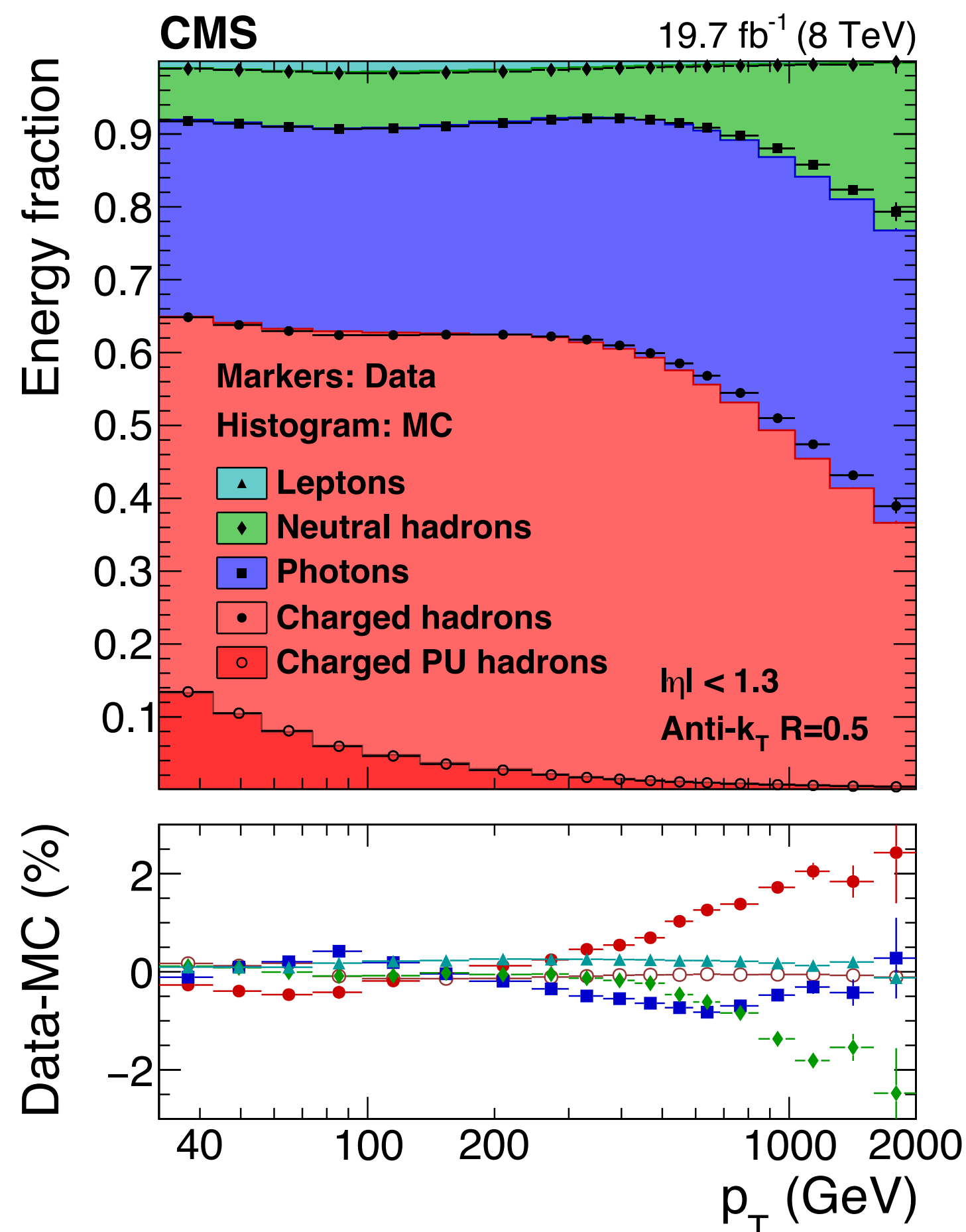
tracking/vertexing

precision timing

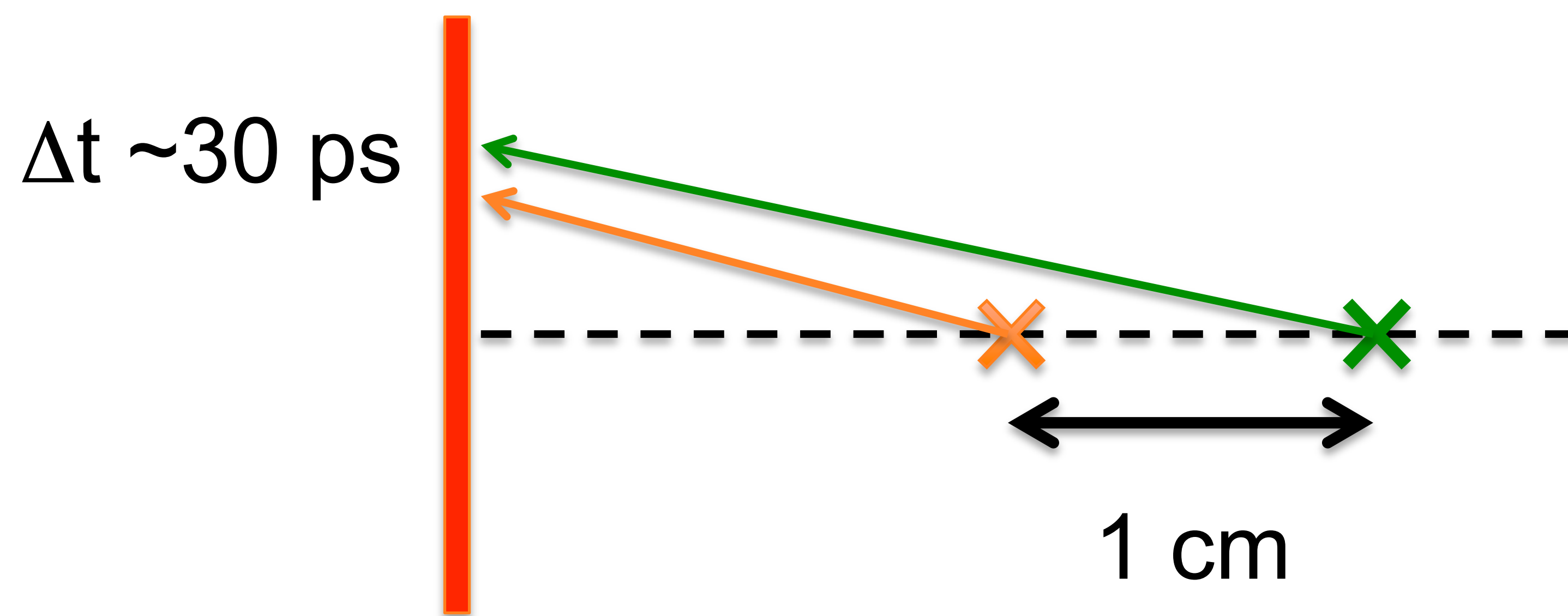
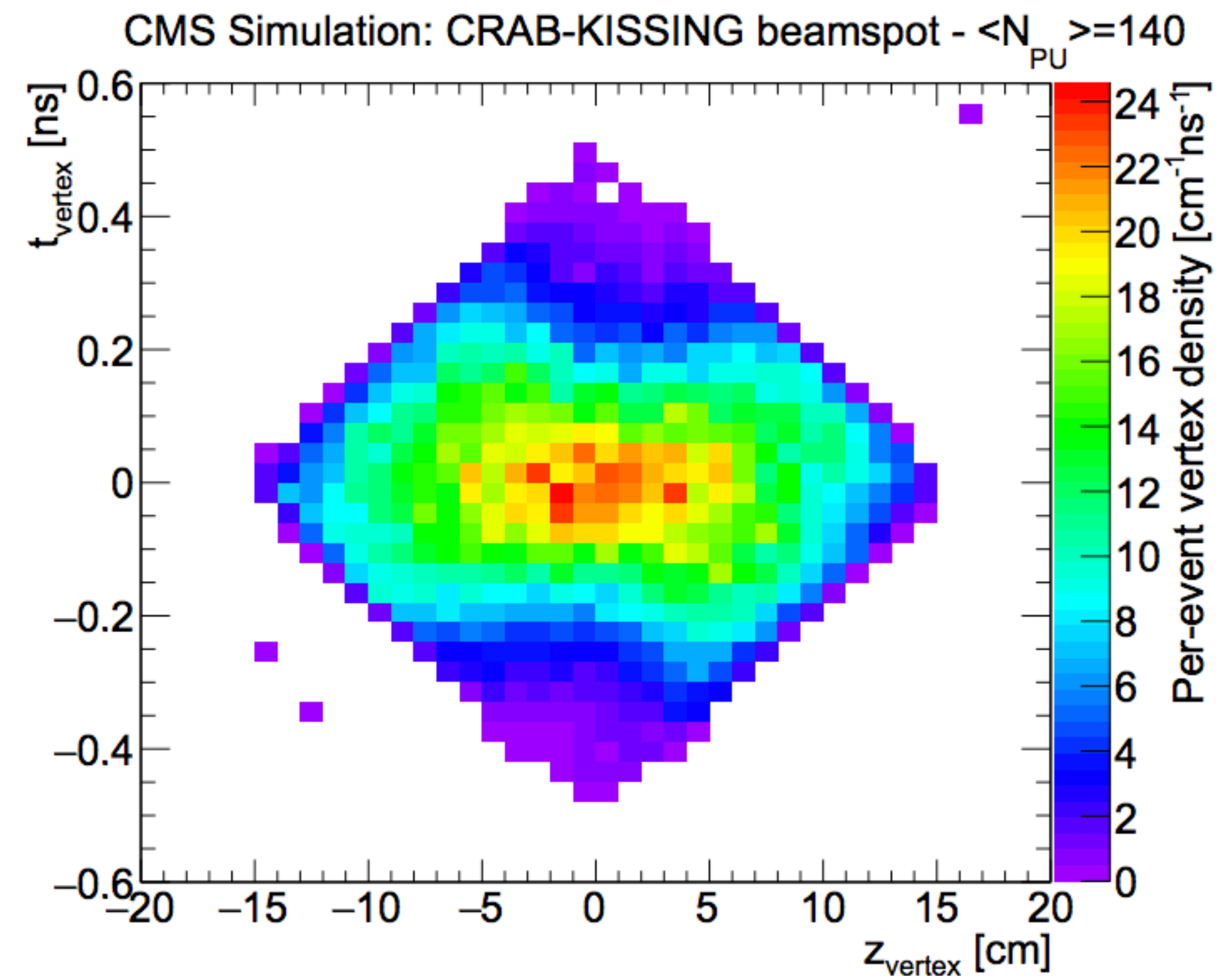
depth segmentation

Charged Hadron Subtraction (CHS)

Falls out naturally from Particle Flow!



asymptotic behavior
local shape
tracking/vertexing
precision timing
depth segmentation



$\sigma_t \sim 30 \text{ ps}$ buys a factor of ~ 10
reduction in effective pileup

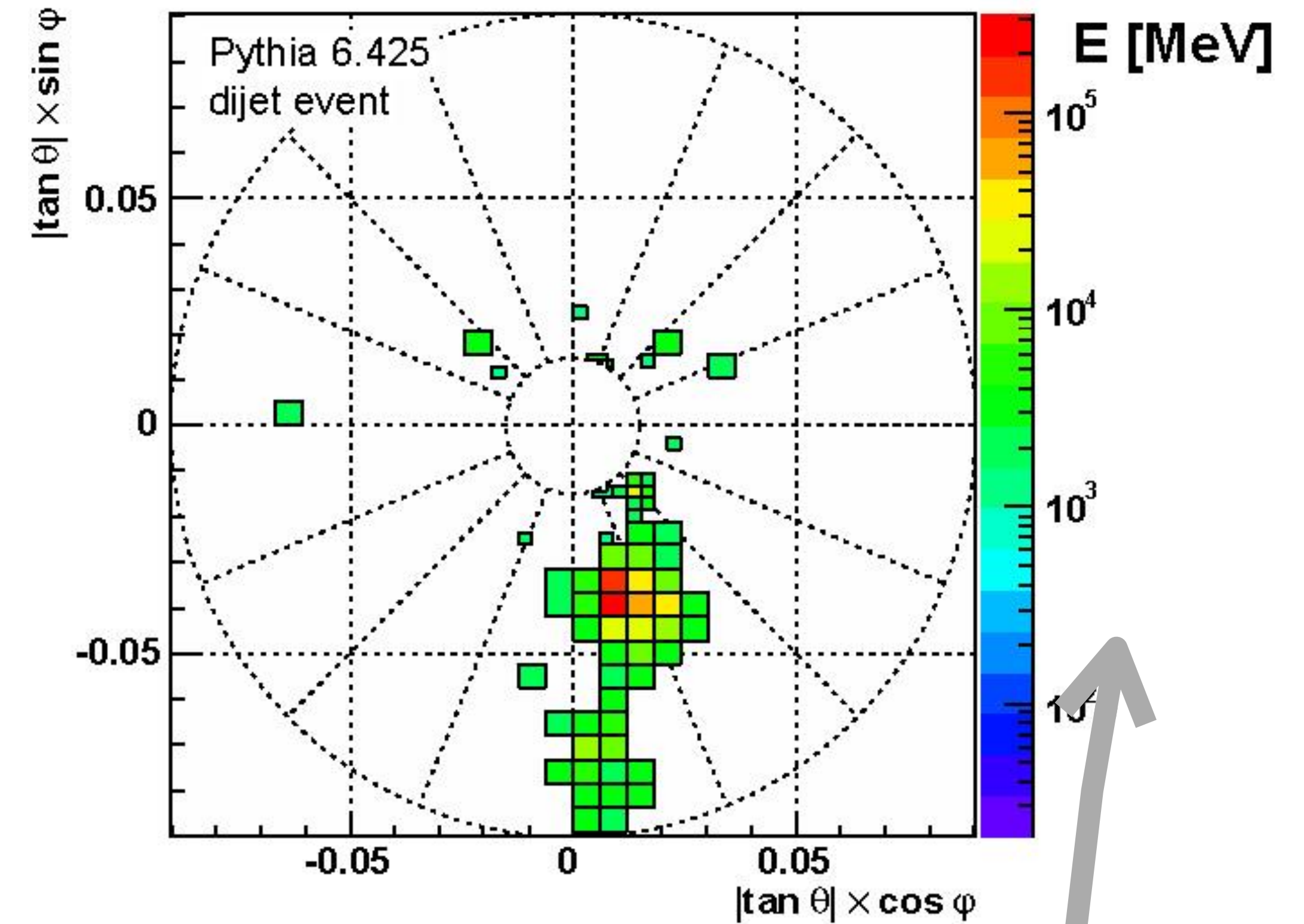
but open questions...
e.g. can we achieve that time
resolution for \sim few GeV photons?

HANDLES ON PILEUP

- asymptotic behavior
- local shape
- tracking/vertexing
- precision timing
- depth segmentation**

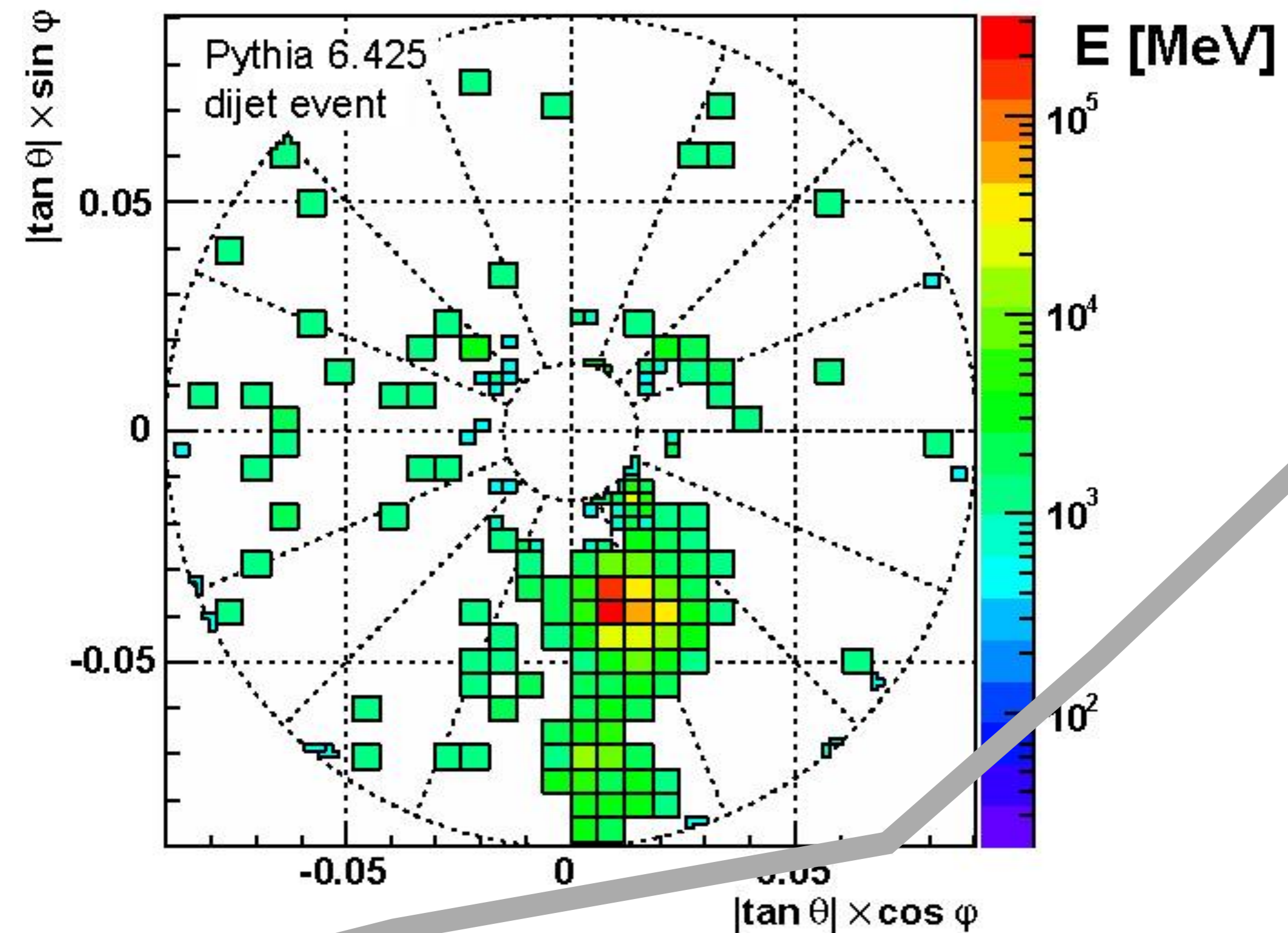
$$\left| \frac{E}{\sigma} \right| > 4$$

ATLAS simulation 2010

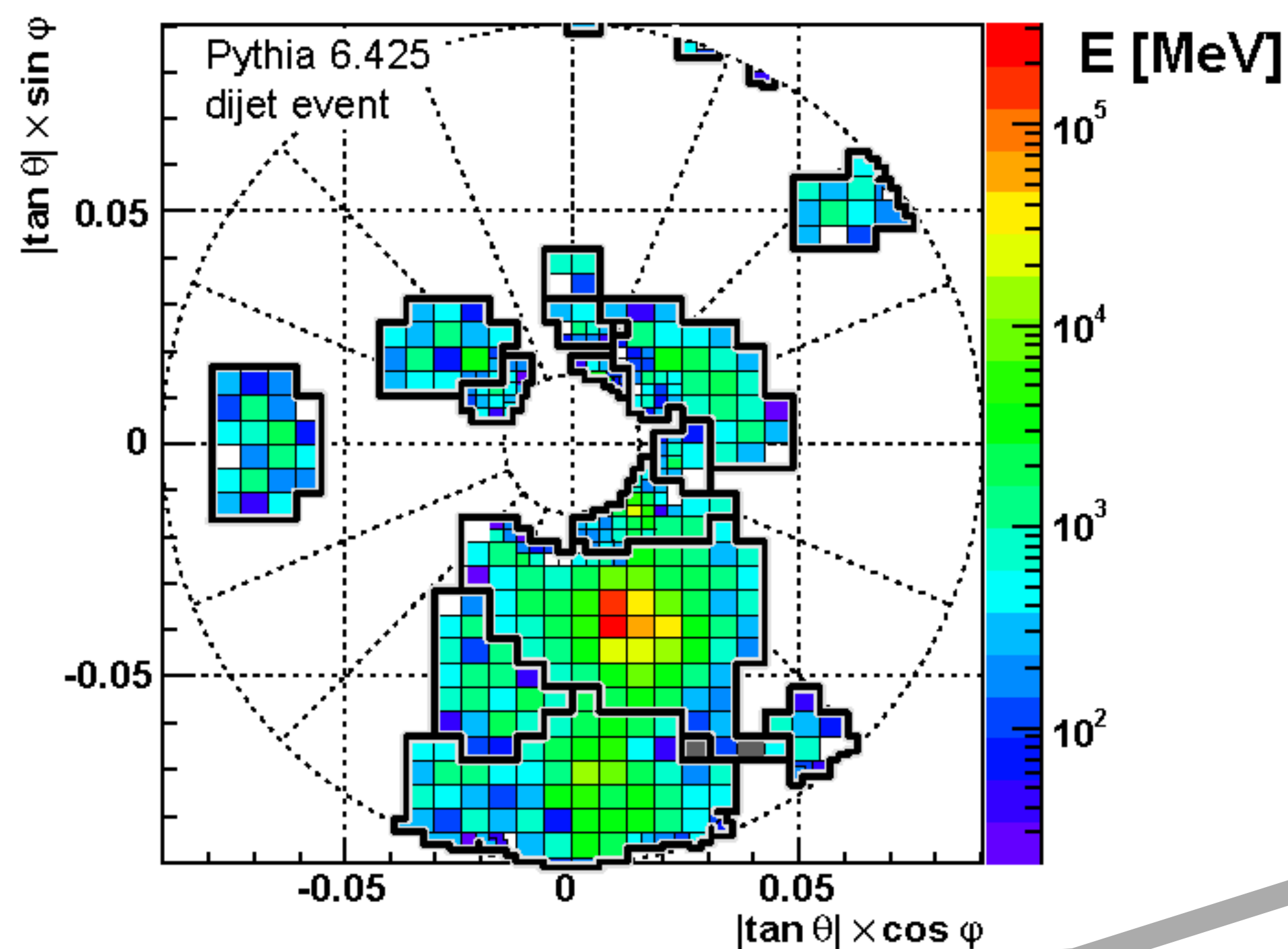


$$\left| \frac{E}{\sigma} \right| > 2$$

ATLAS simulation 2010



ATLAS simulation 2010



clustering uses neighbors in depth too! no longer 2D clustering

Notice that each method that we've described works on a given **physics object**...

each method presented so far also has its downfalls

What if we act on the event building blocks?

e.g. constituents/particles

constituent subtraction, softkiller, PUPPI

hep-ph:1403.3108

hep-ph:1407.0408

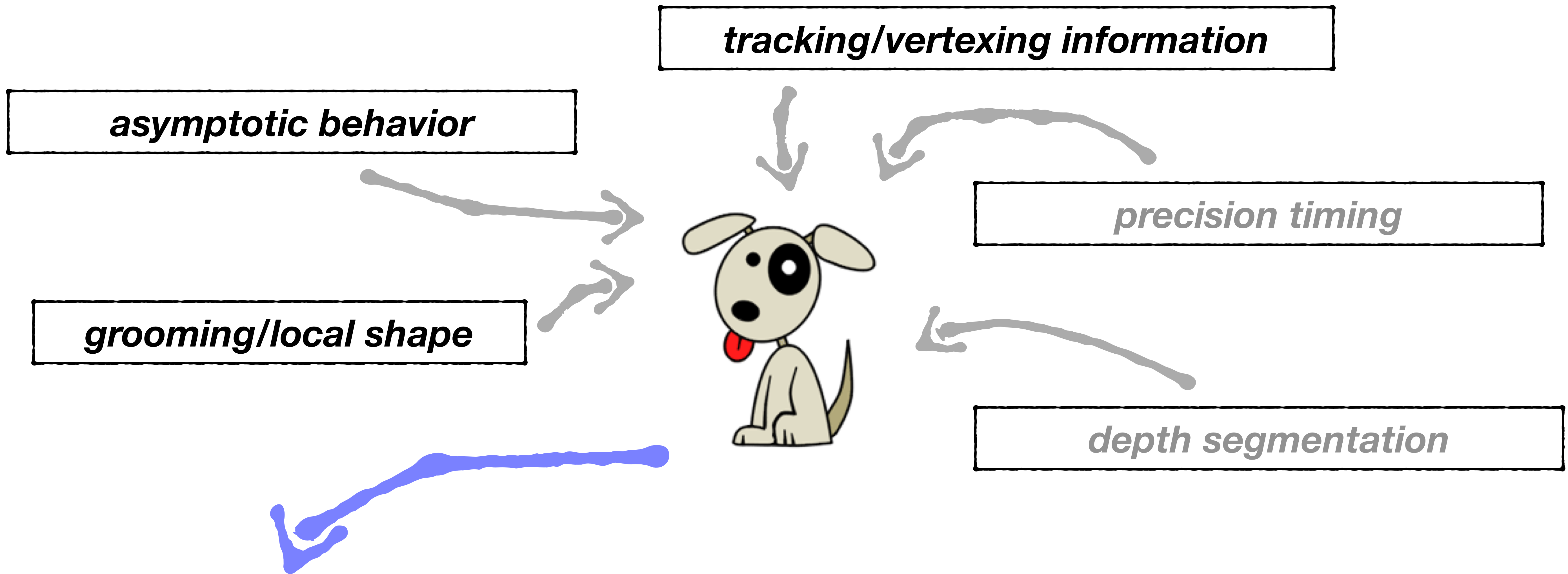
hep-ph:1407.6013

What if we exploit all information possible simultaneously?

asymptotic, local shape, tracking, etc...

What if, you could identify each particle in the event and give the likelihood that it's pileup?

PILEUP PER PARTICLE IDENTIFICATION

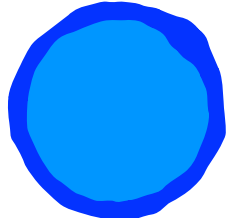
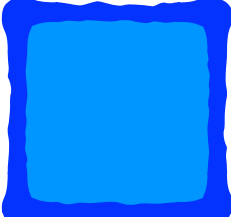
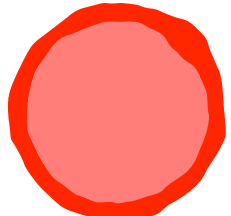

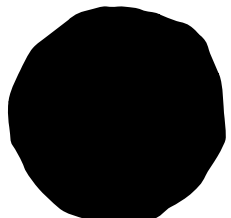
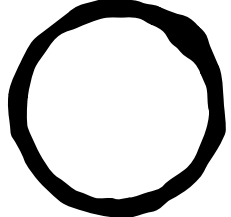


Define on a **per particle** basis, **before jet clustering**, a weight for **how likely** a particle (or jet constituent) is to be from pileup or the leading vertex, then rescale each particle four momentum by that weight

$$\alpha_i^C = \log \left[\sum_{j \in \text{Ch, LV}} \frac{p_{T,j}}{\Delta R_{ij}} \Theta(R_0 - \Delta R_{ij}) \right]$$

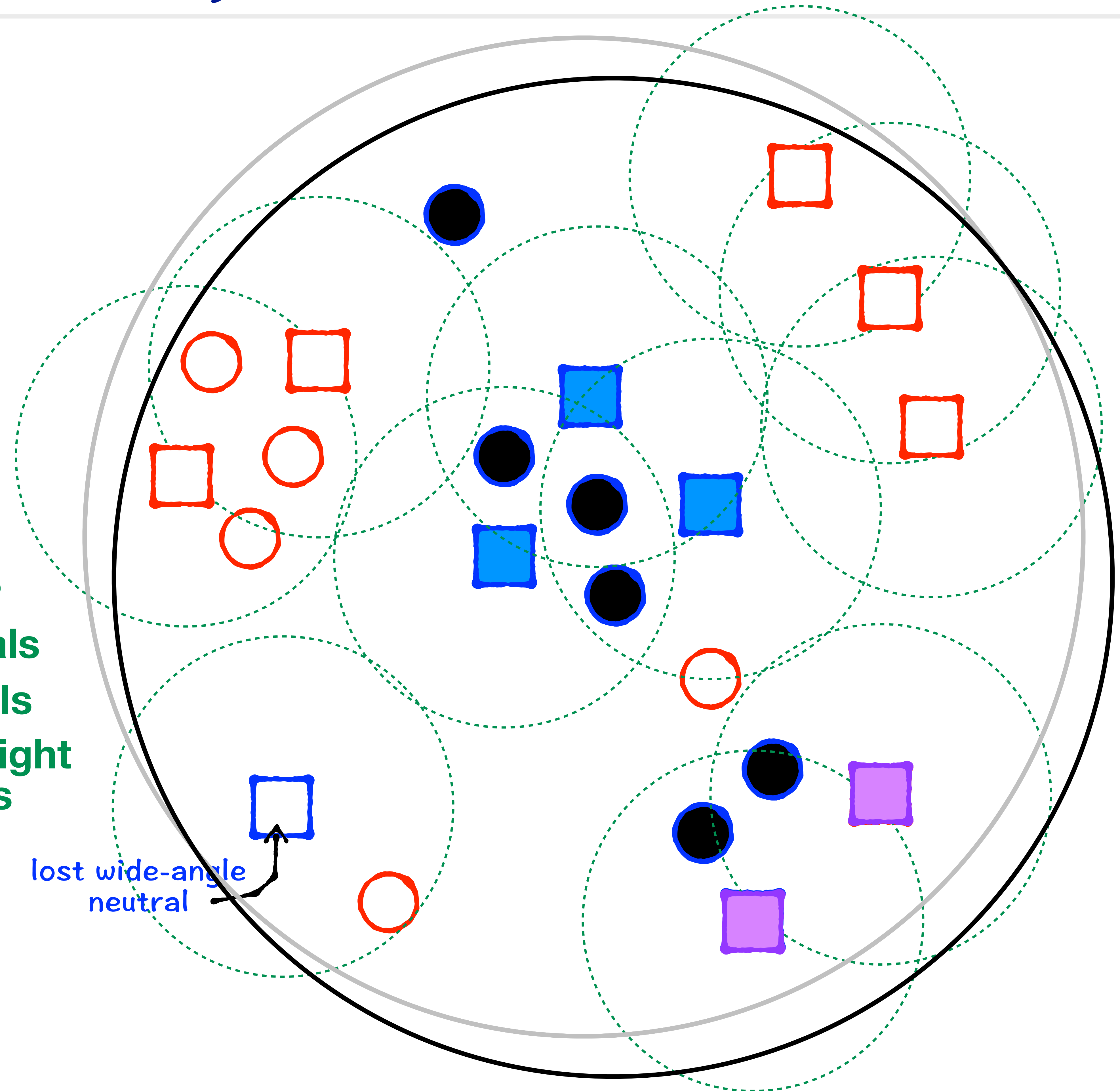
define an α_i per particle; sample the PU α distribution per event; ask how likely particle i is to be pileup

PUPPI (IN CARTOONS)

-  LV charged
-  LV neutral
-  PU charged
-  PU neutral
-  chosen
-  removed

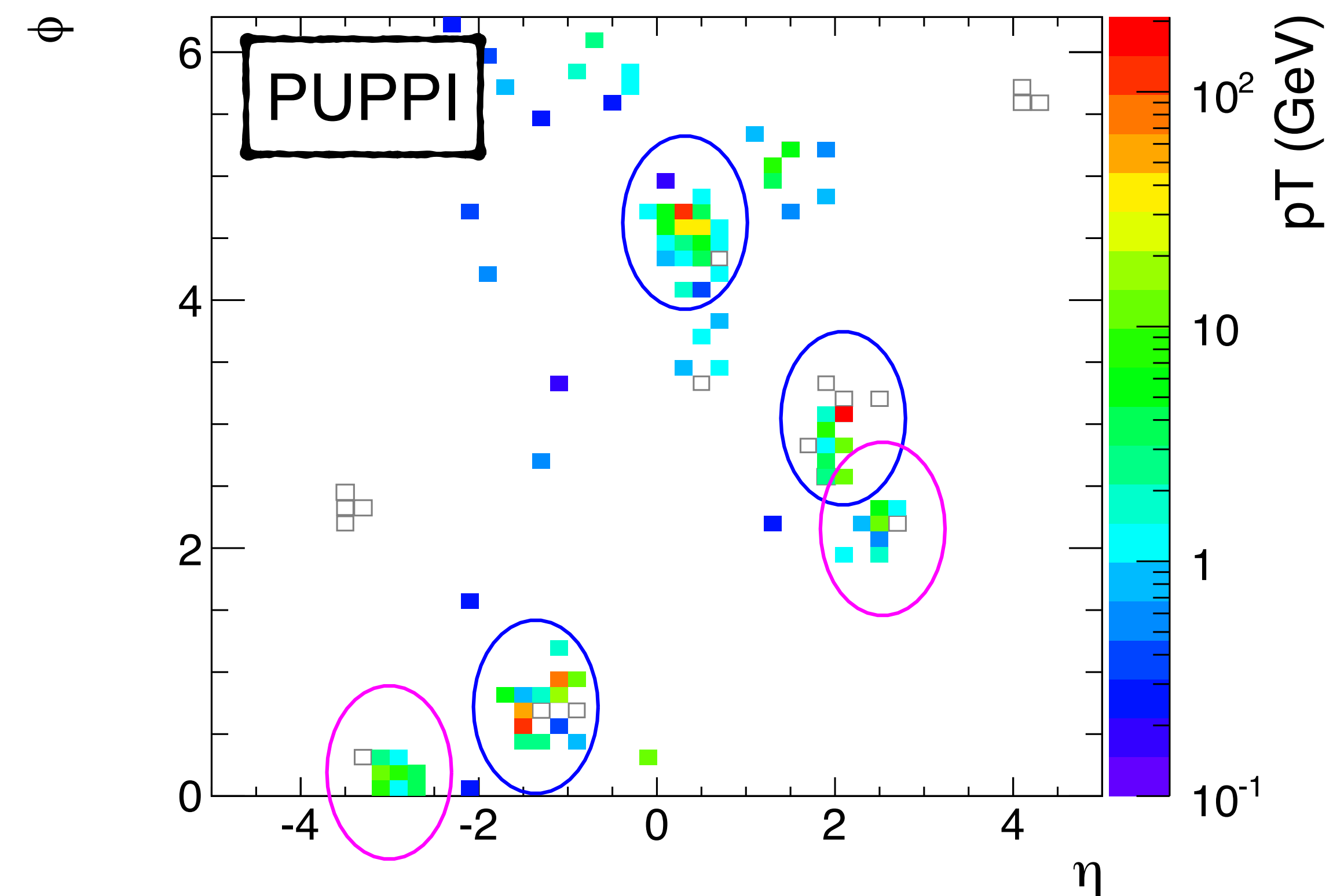
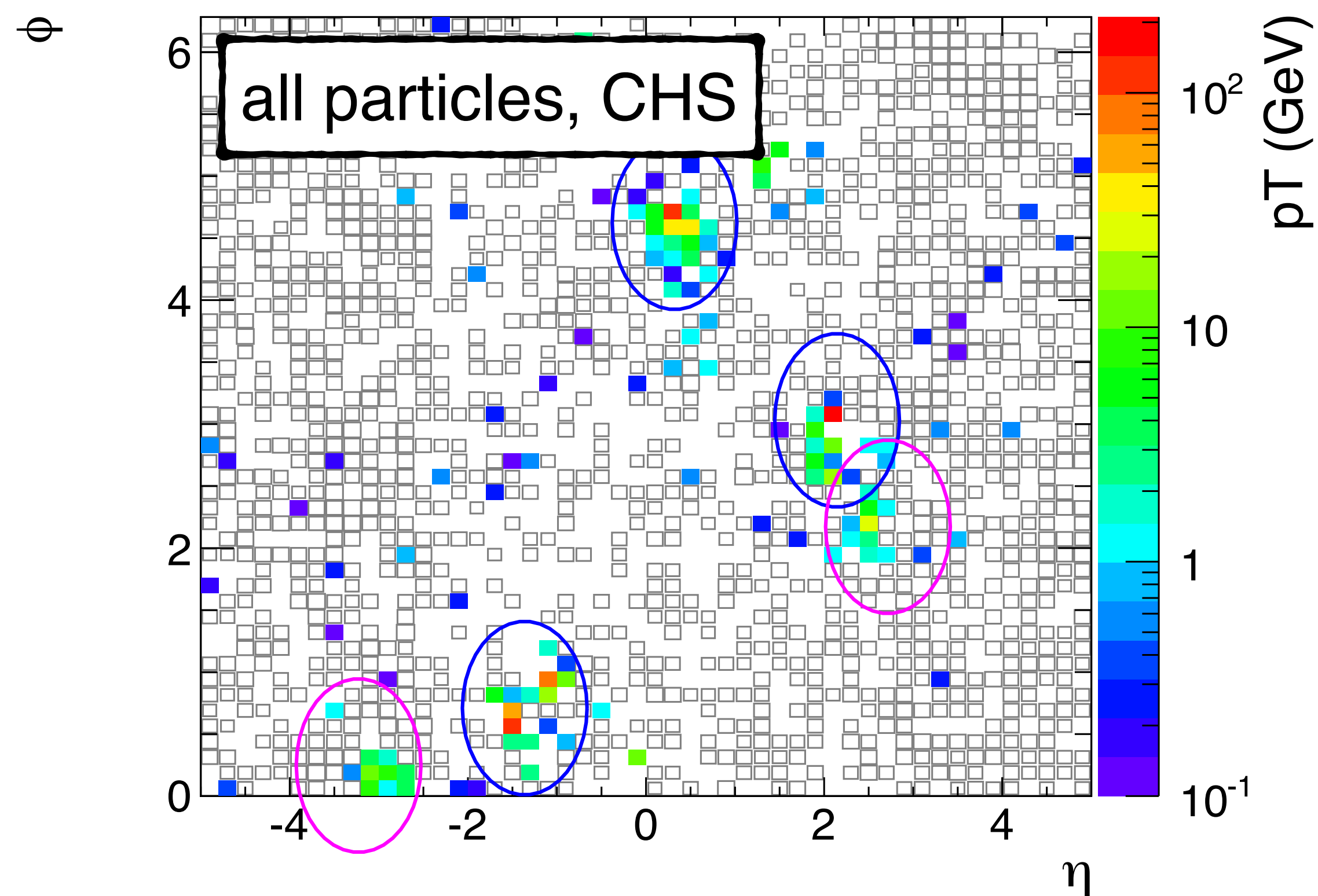
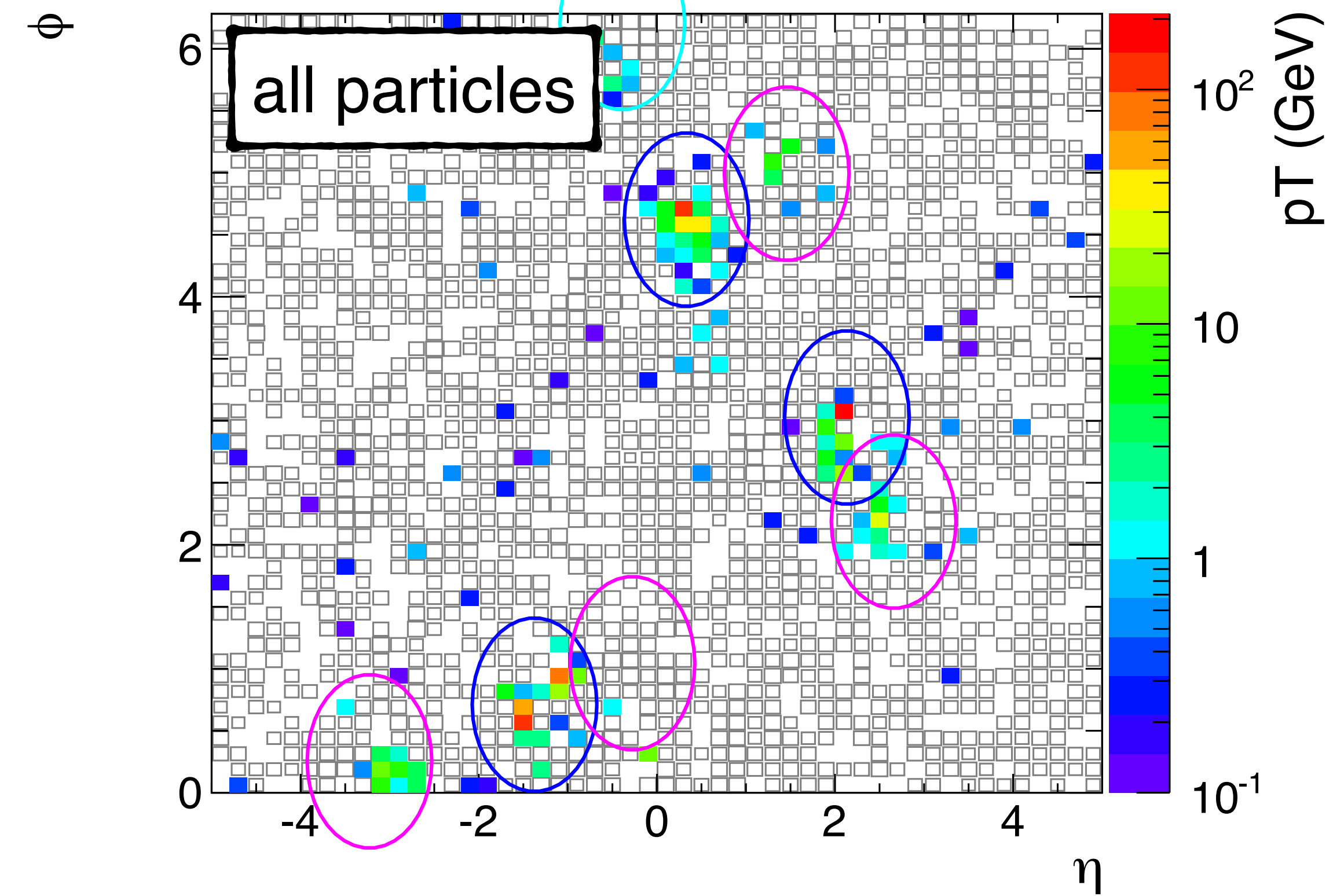
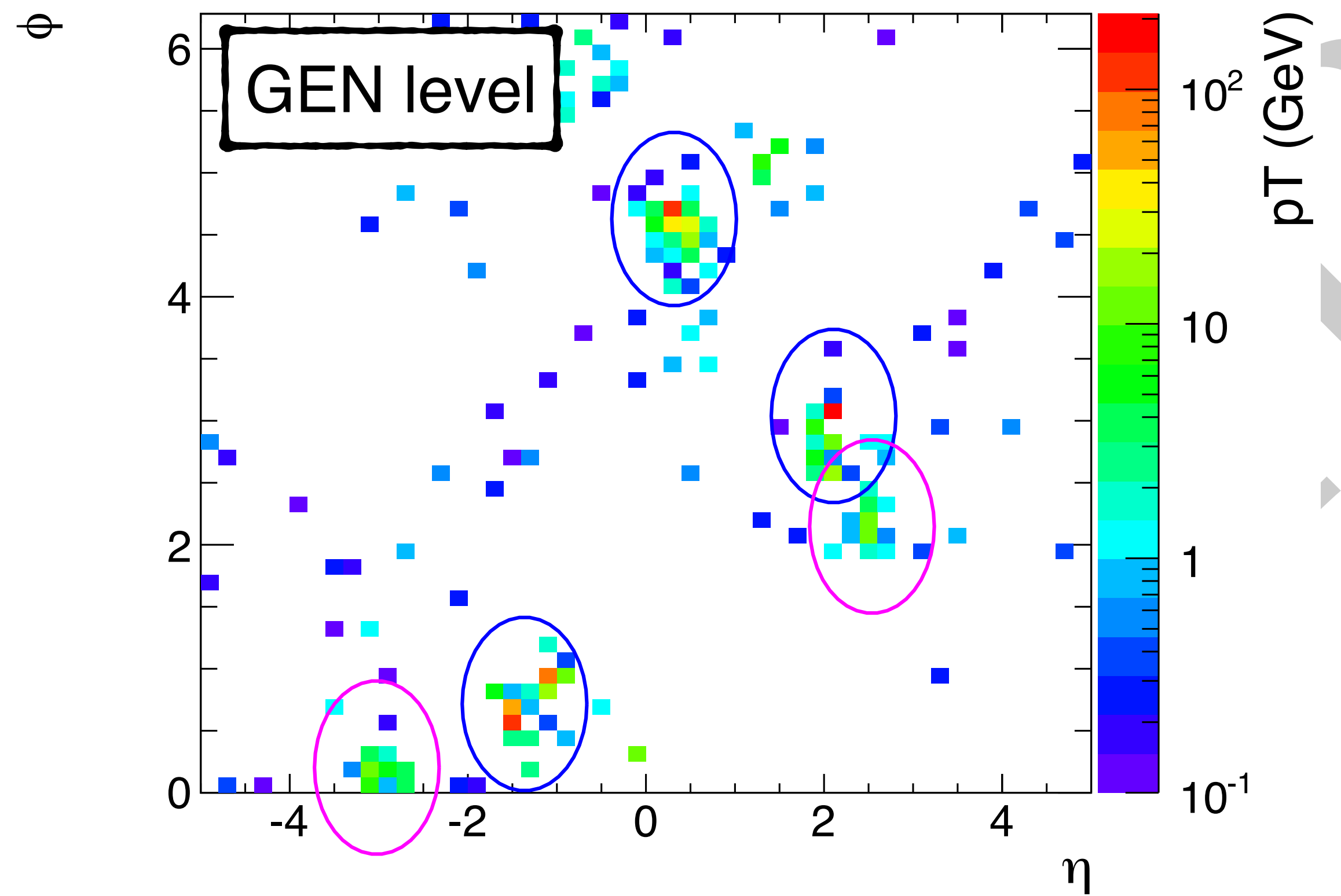
1. use tracking info
2. look around neutrals
3. remove "0" neutrals
4. assign fractional weight to ambiguous cases

recluster event,
new jet!



PILEUP PER PARTICLE ID

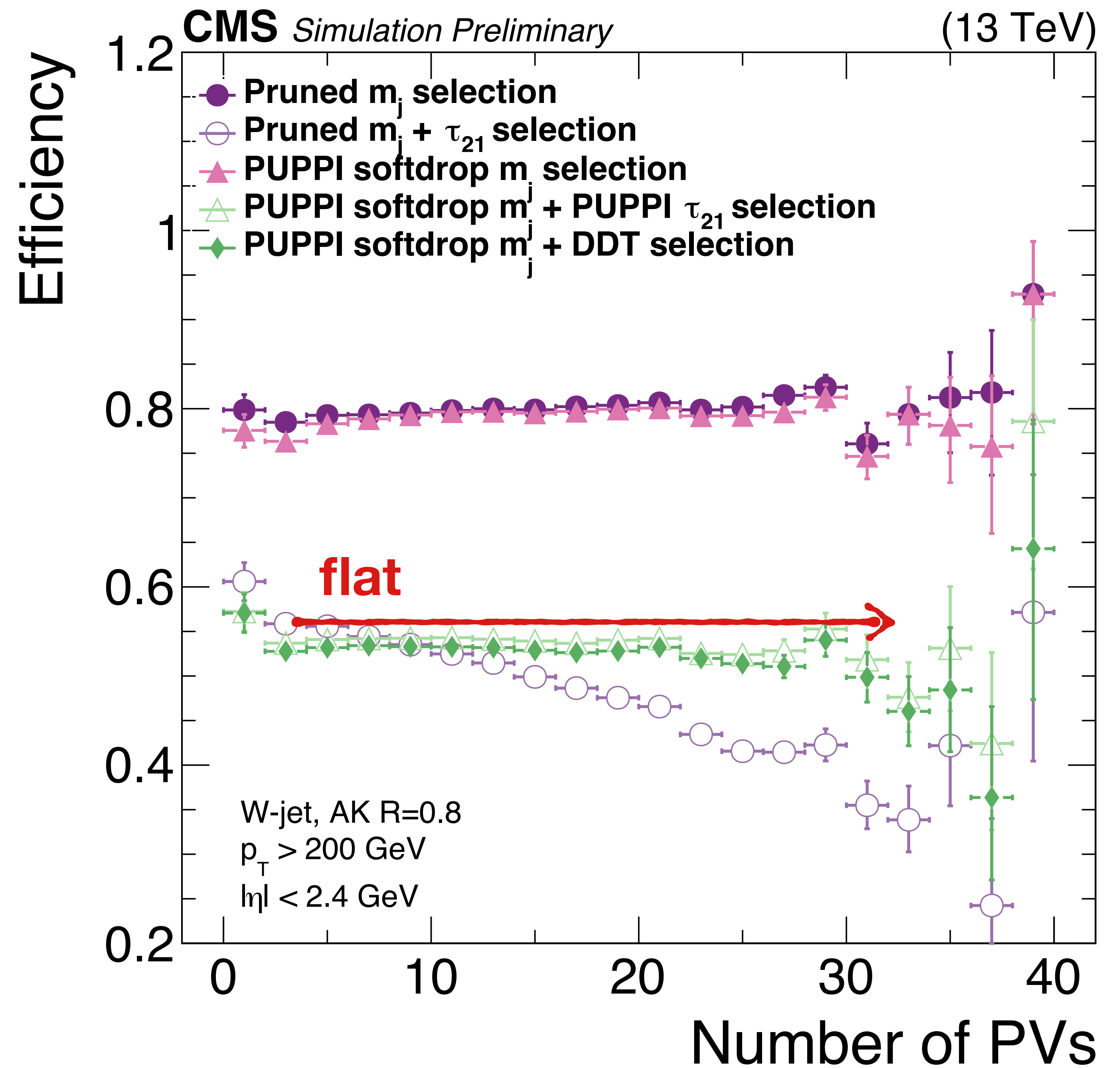
colored cells = process of interest
black cells = pileup



N.B. Particle level studies assuming perfect tracking for $|\eta| < 2.5$

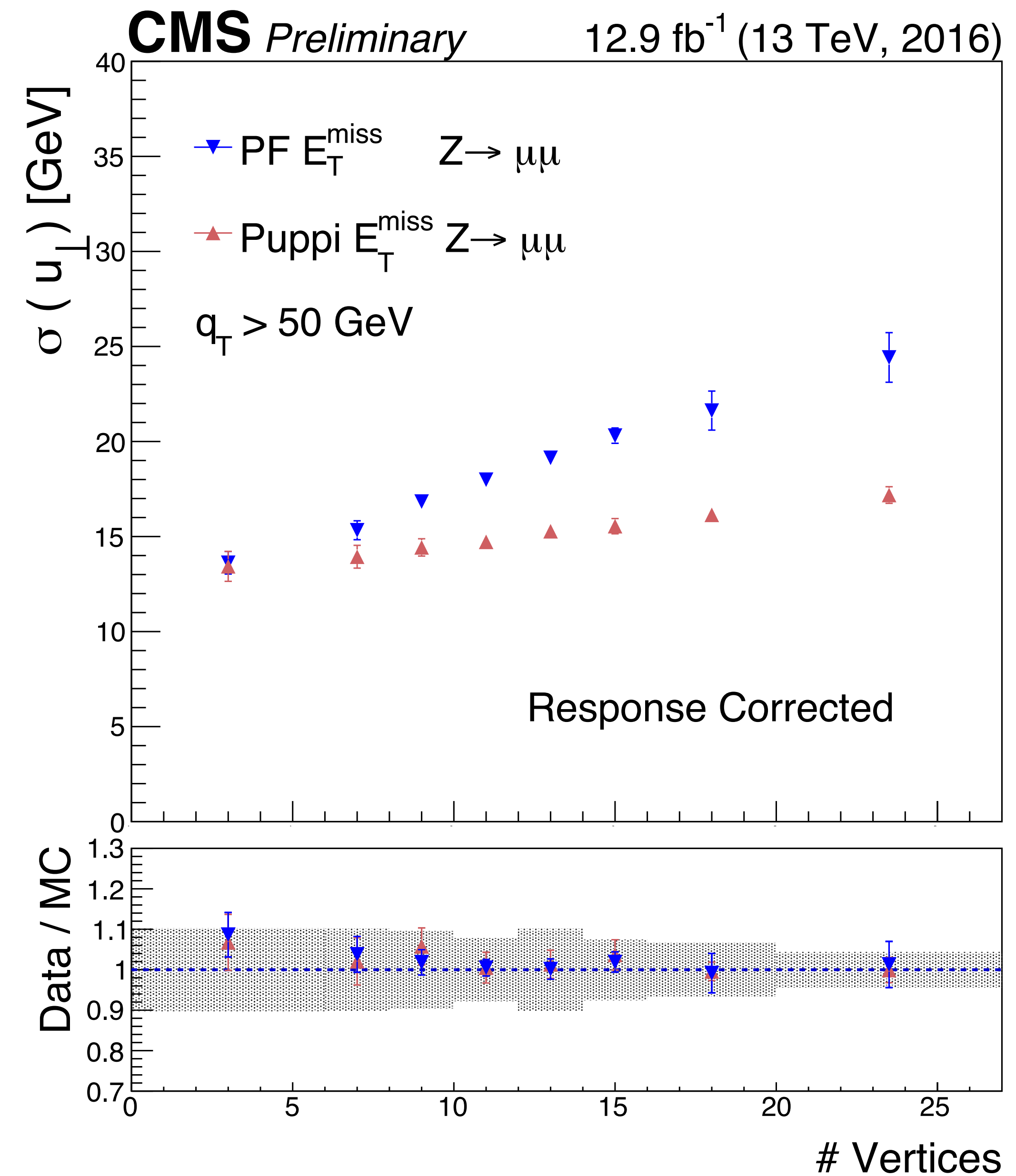
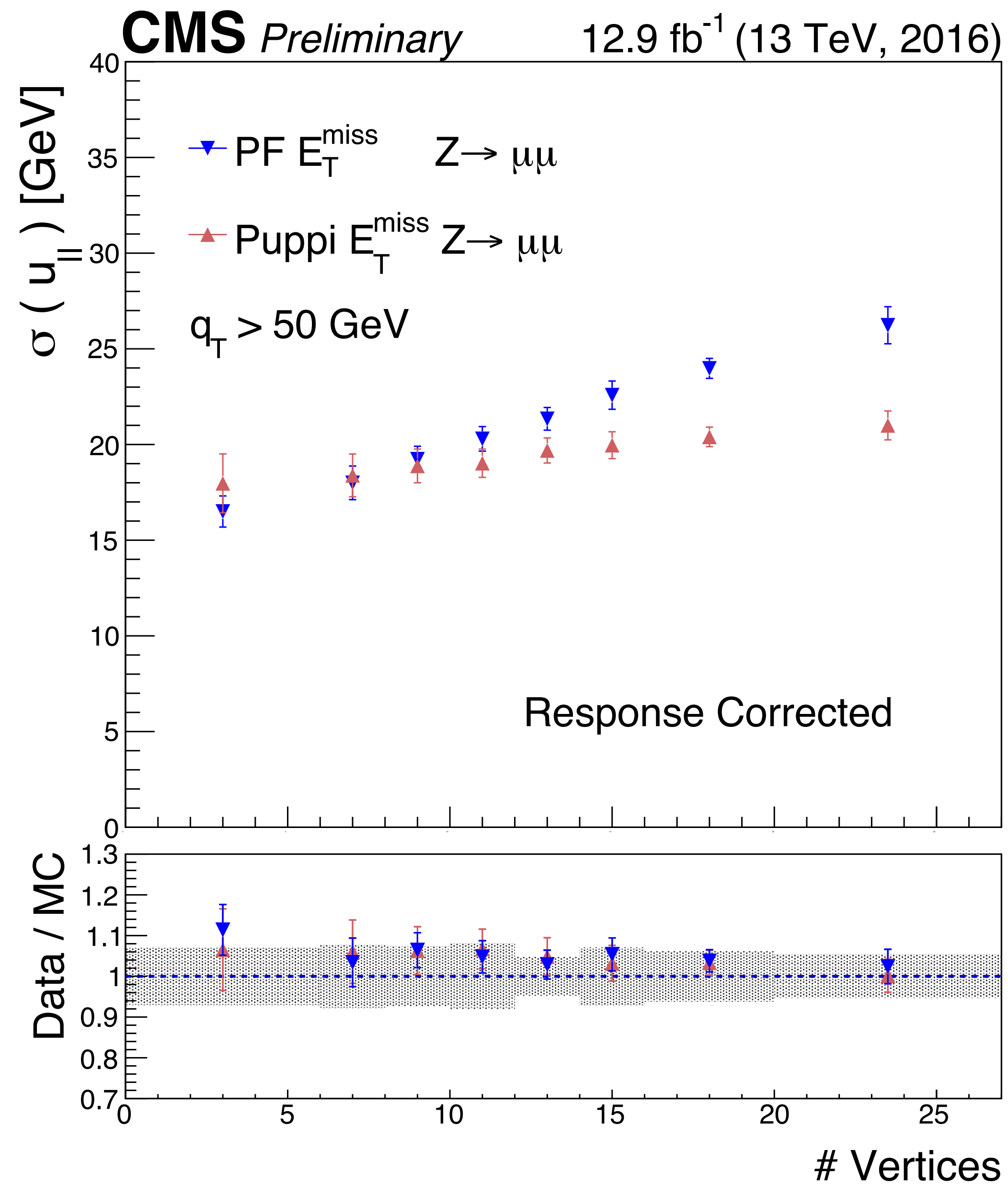
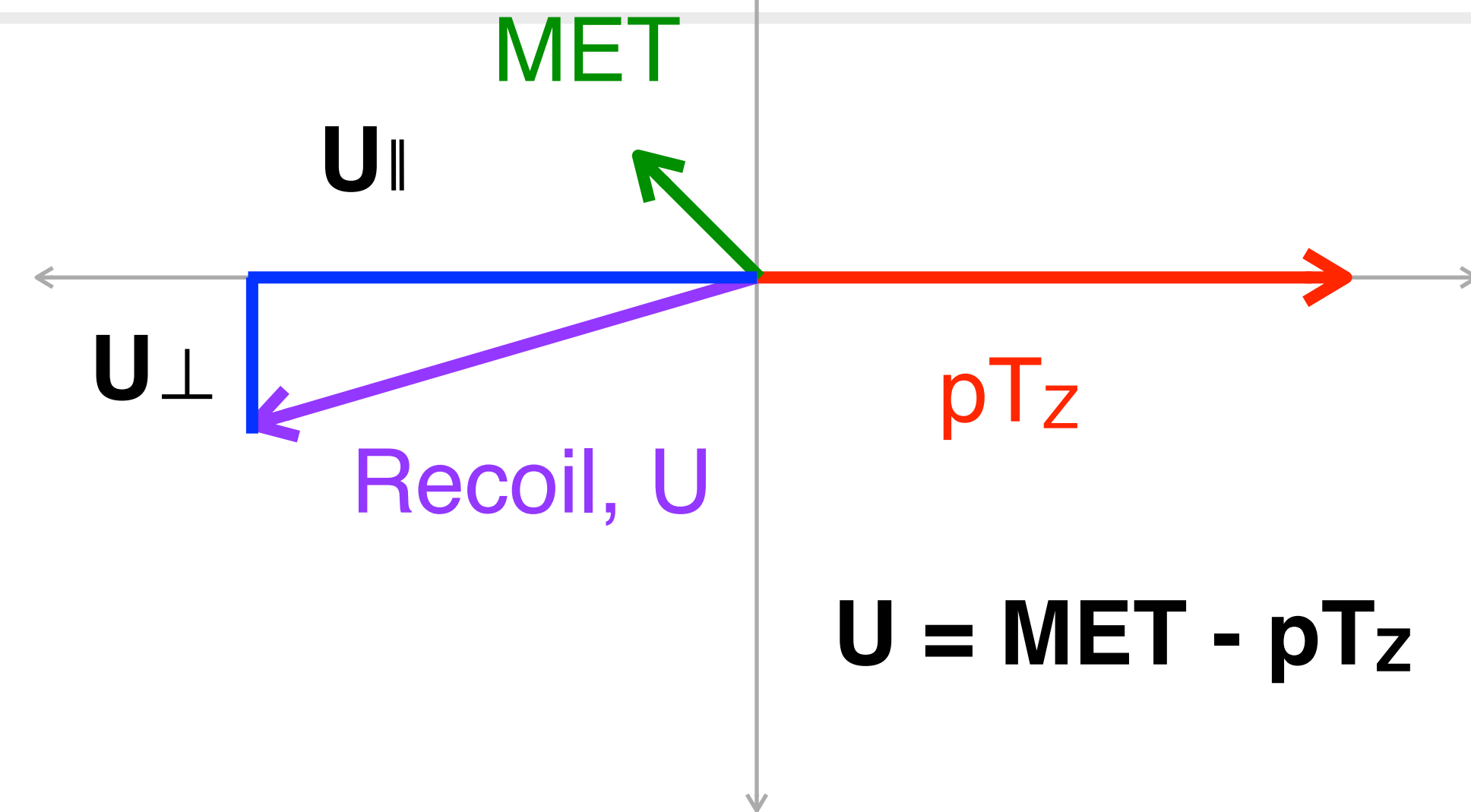
"Classic" use-case for per particle pileup mitigation,
it works for all jet shapes

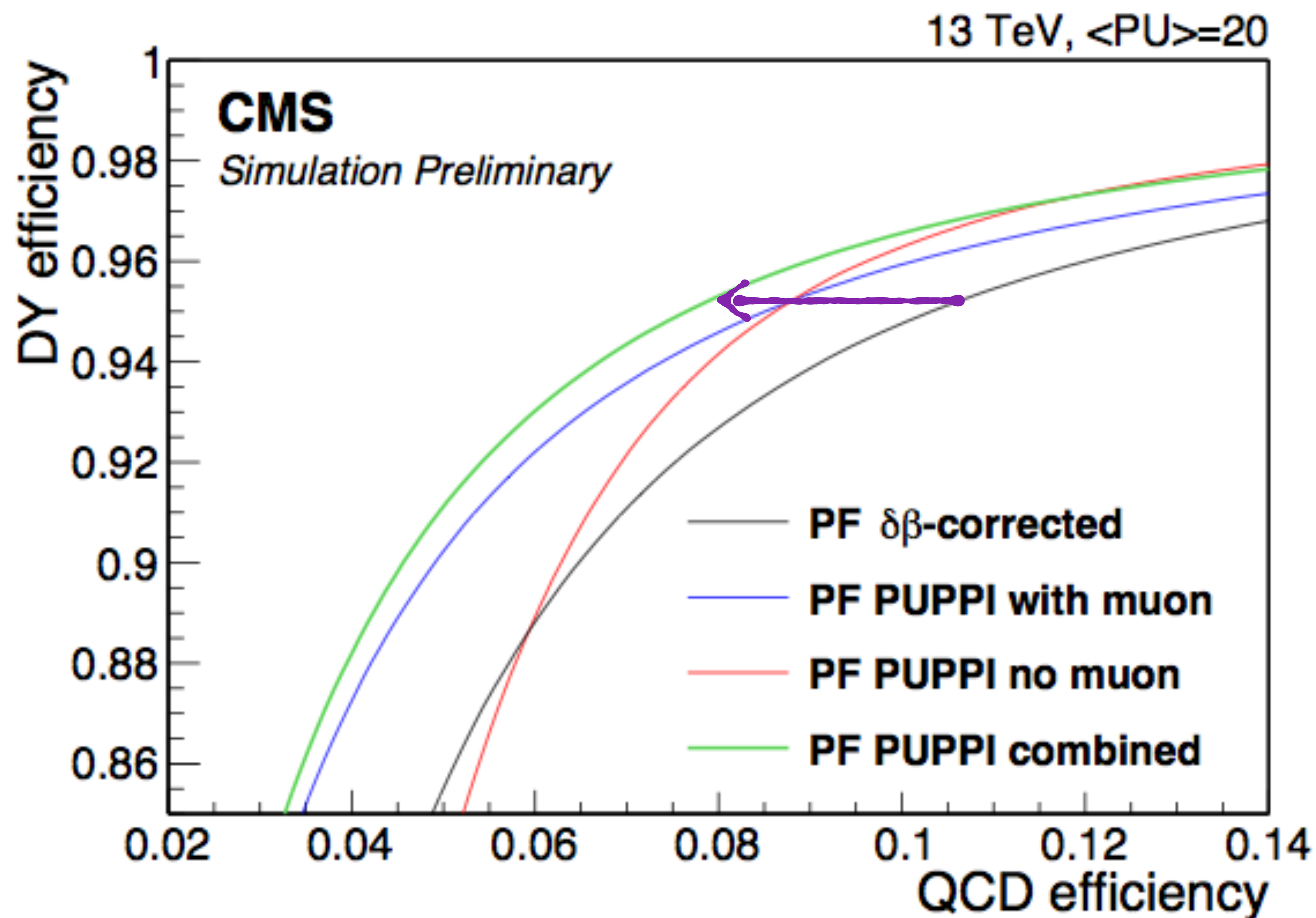
Here, this is the effect of PUPPI on W-tagging shown for PFCHS inputs vs. PUPPI inputs



PUPPI PERFORMANCE

20-30% resolution improvement in the MET resolution @ $N_{PV} \sim 20$ over traditional "PU" corrected MET





25% decrease in backgrounds using per particle uncertainties at 20 PU!

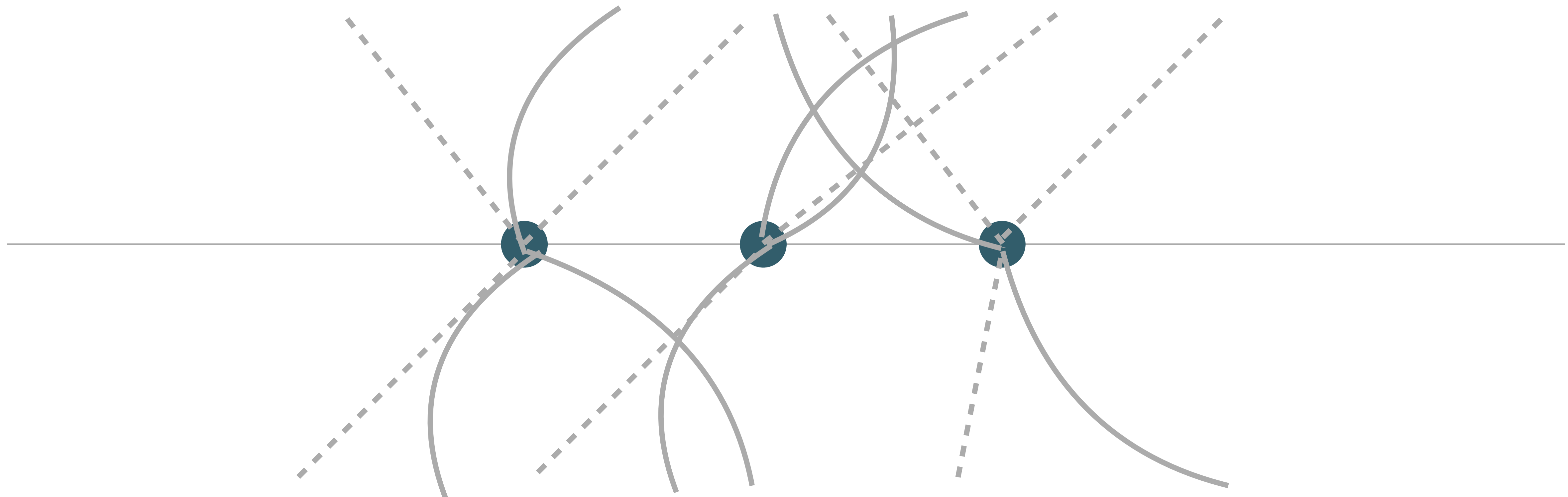
*"combined" curve uses both muon hypotheses
Vs. traditional methods*

What if, you could identify each particle in the event and give the likelihood that it's pileup?

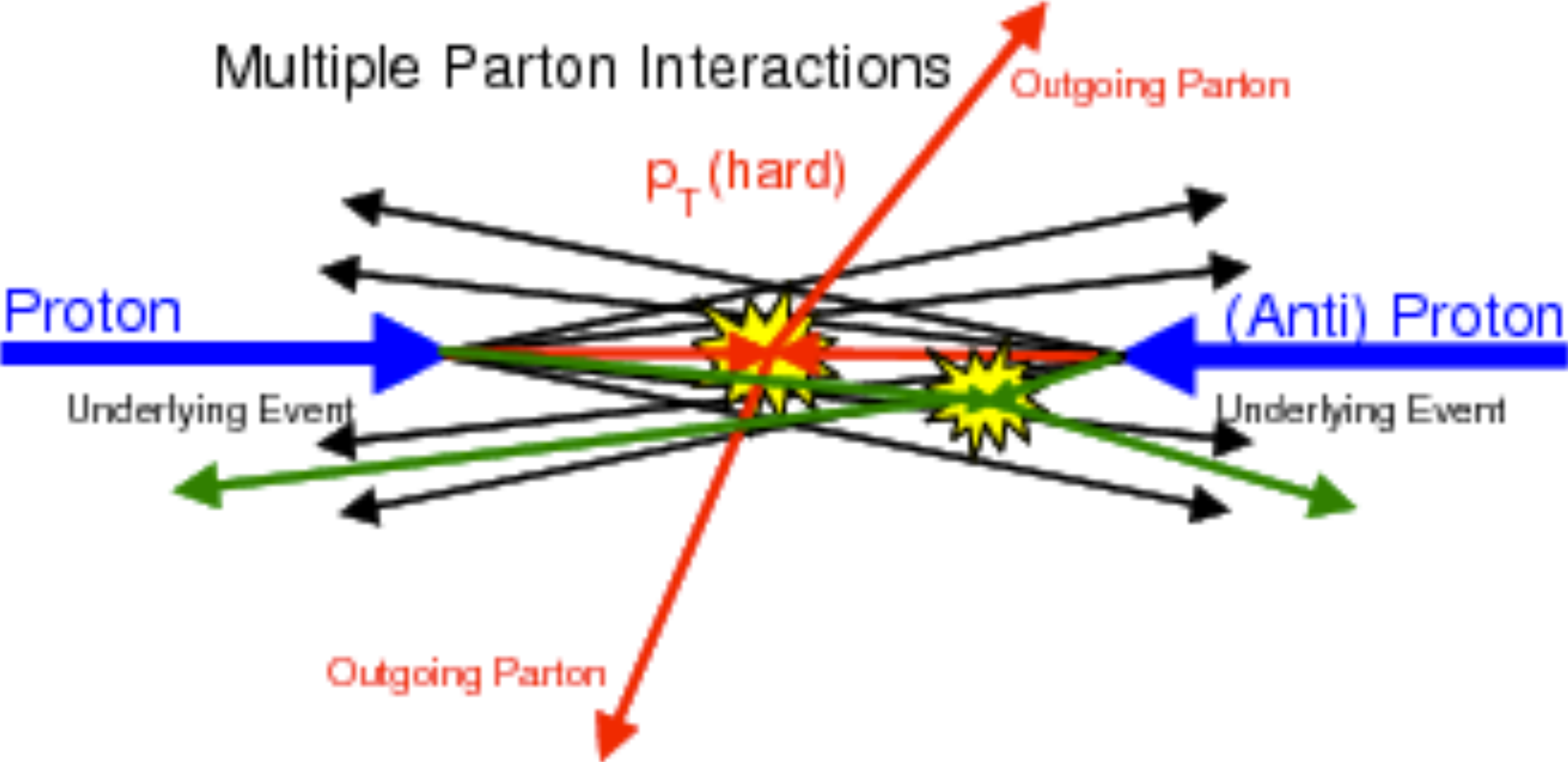


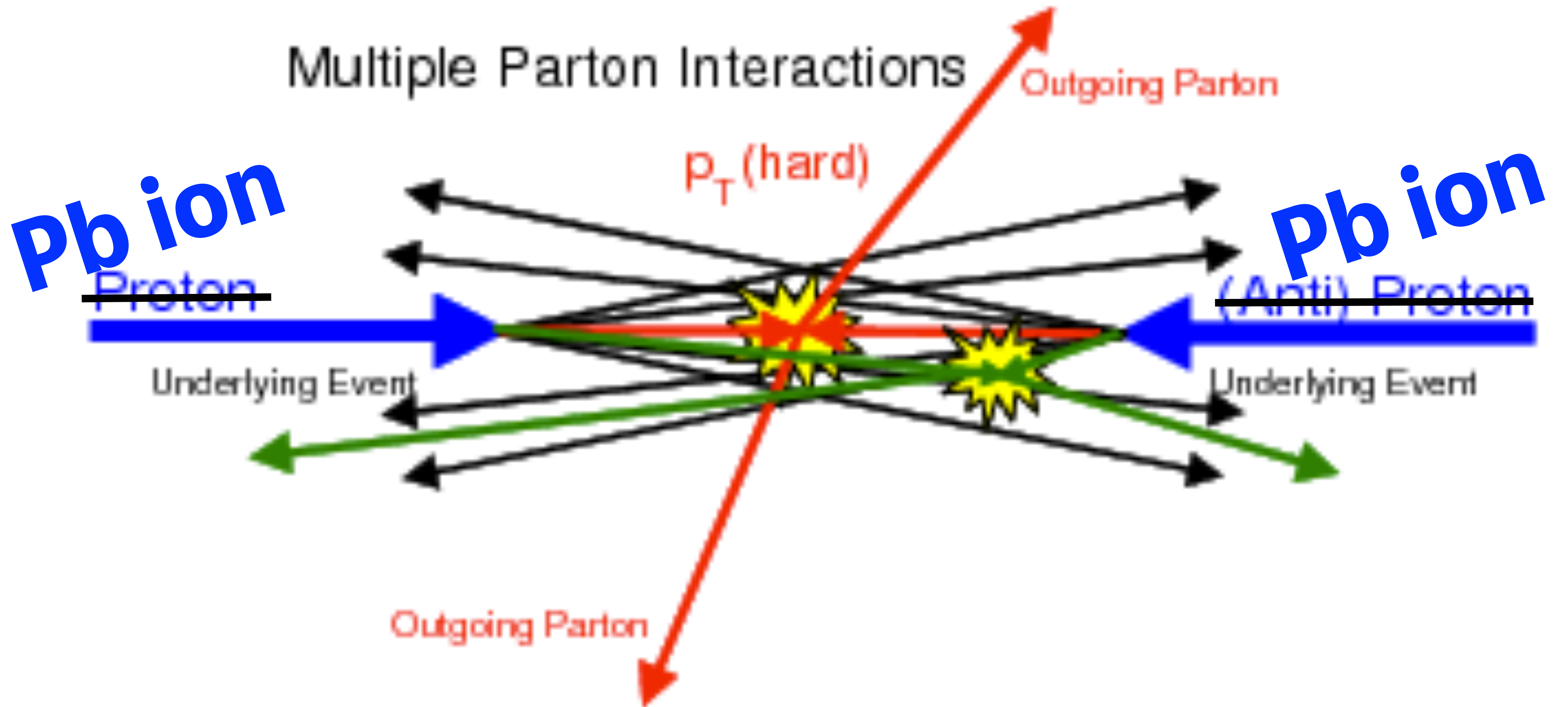
What if, you could identify each particle in the event and give the likelihood that it belongs to a given vertex i ?

a combination of the PUPPI approach and the ATLAS forward vertex jet tagging ideas...



CONTRAST AGAINST UNDERLYING EVENT

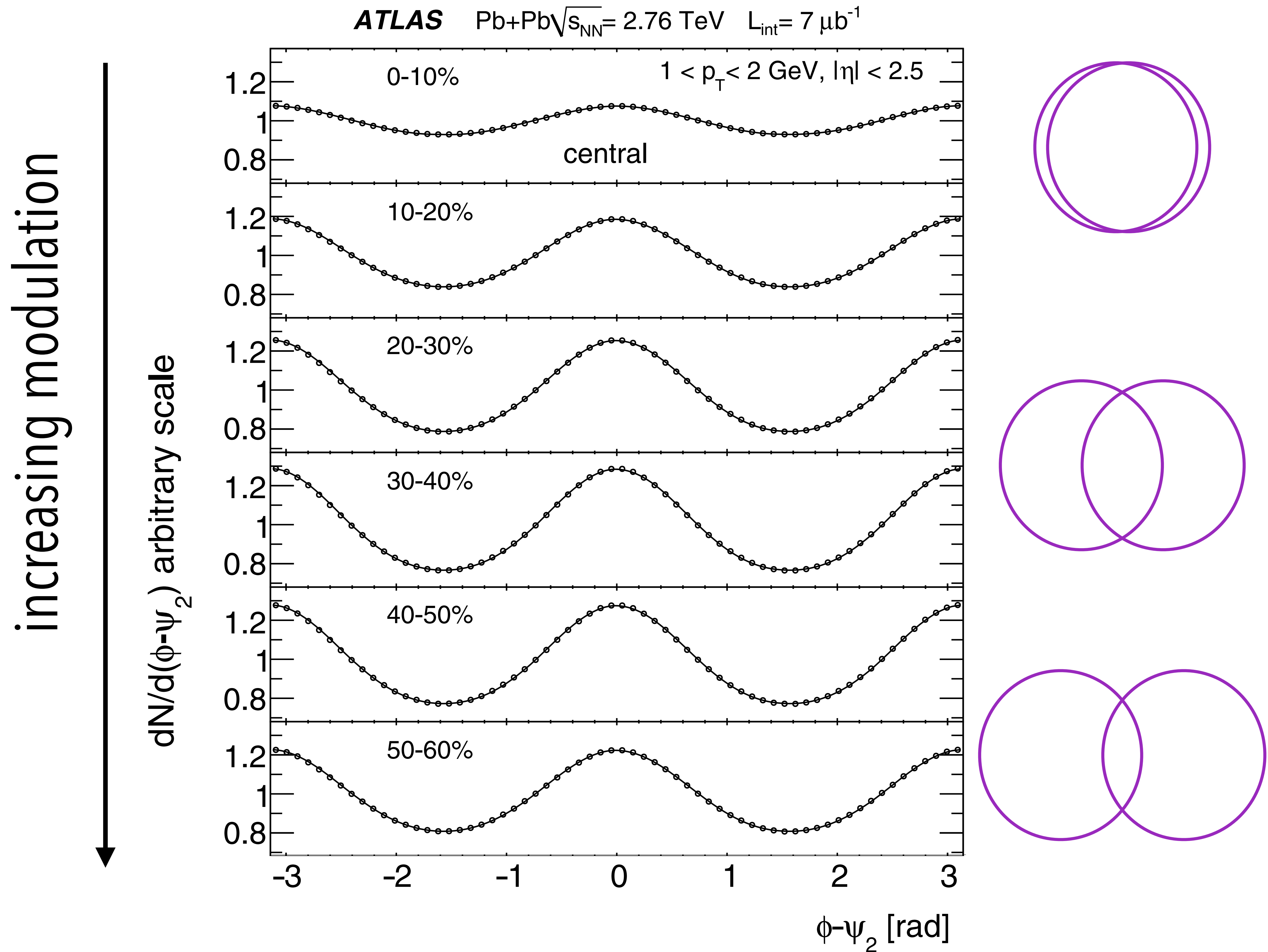




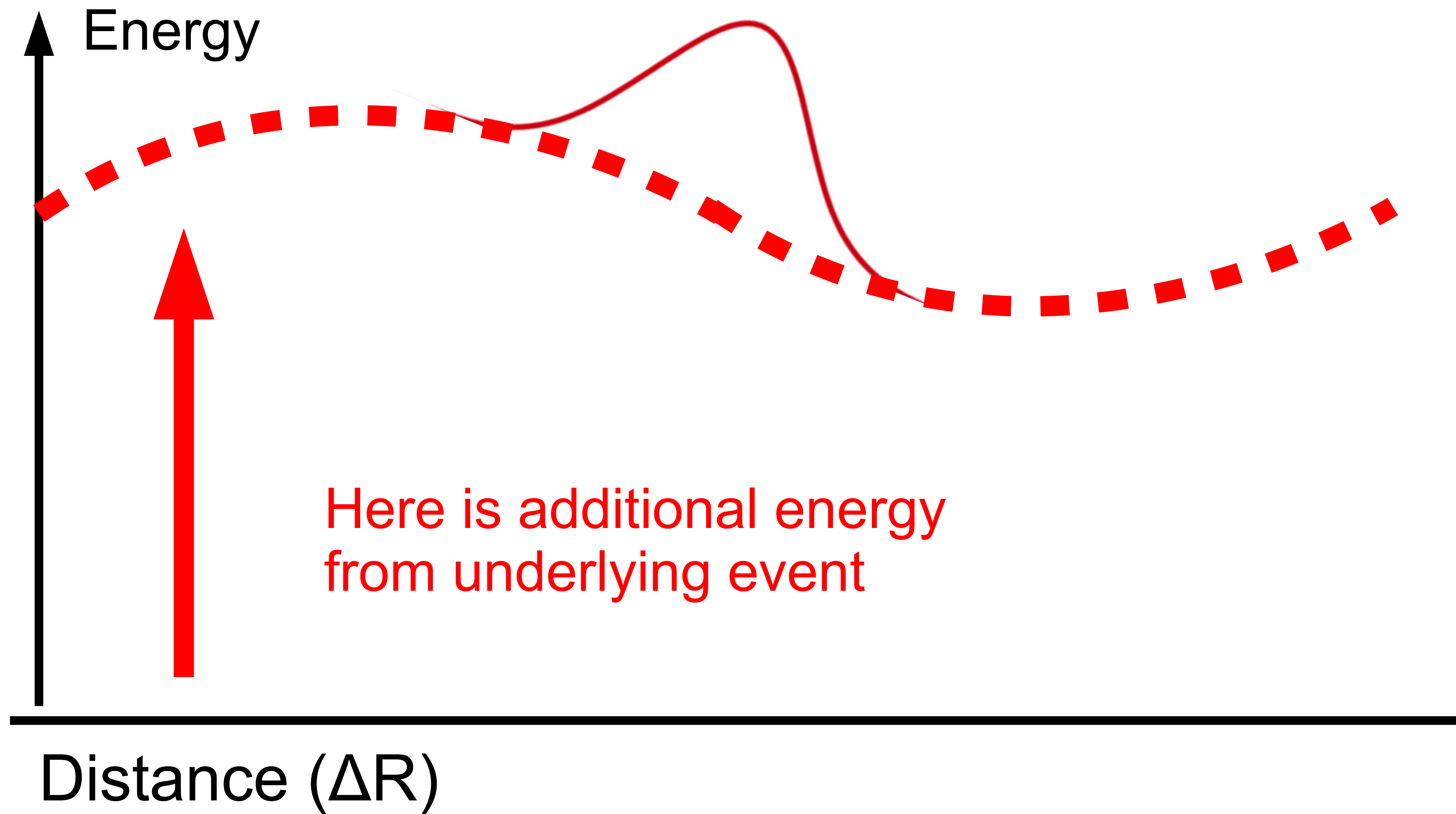
Underlying event in heavy ions

Similar to A LOT of pileup, but without a vertexing handle
... and it has some correlated structure!

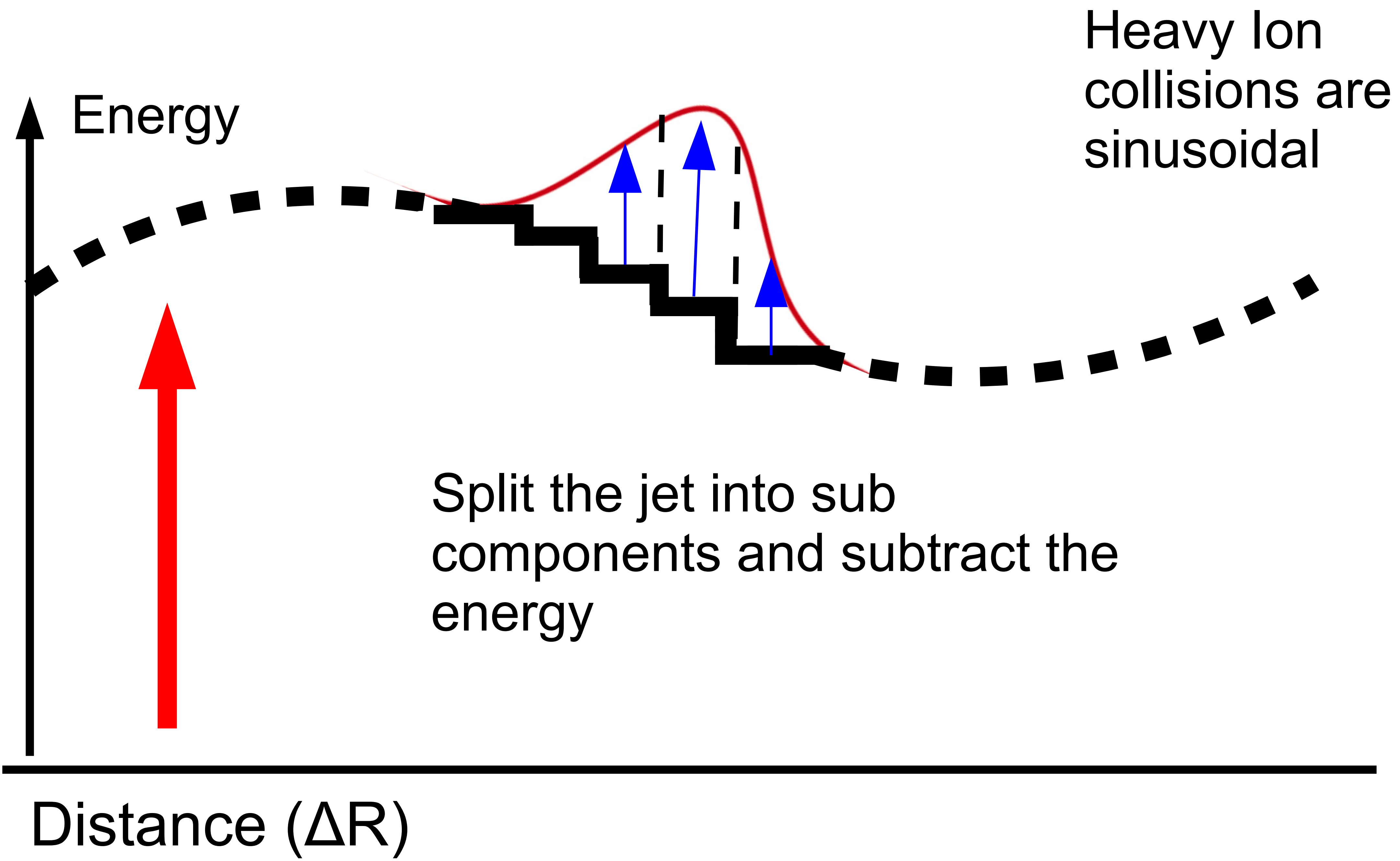
“UNDERLYING EVENT” IN HEAVY IONS

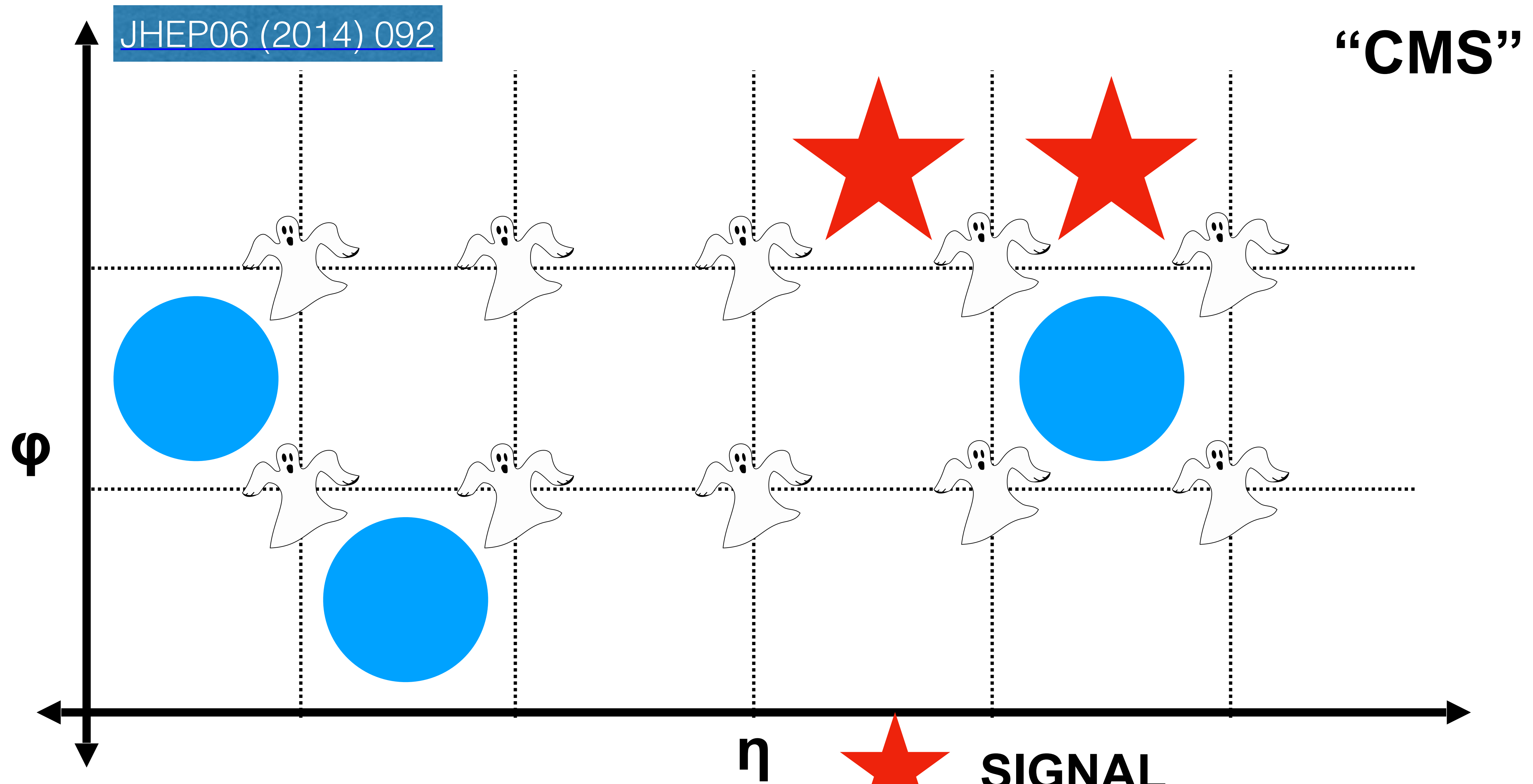


“UNDERLYING EVENT” IN HEAVY IONS

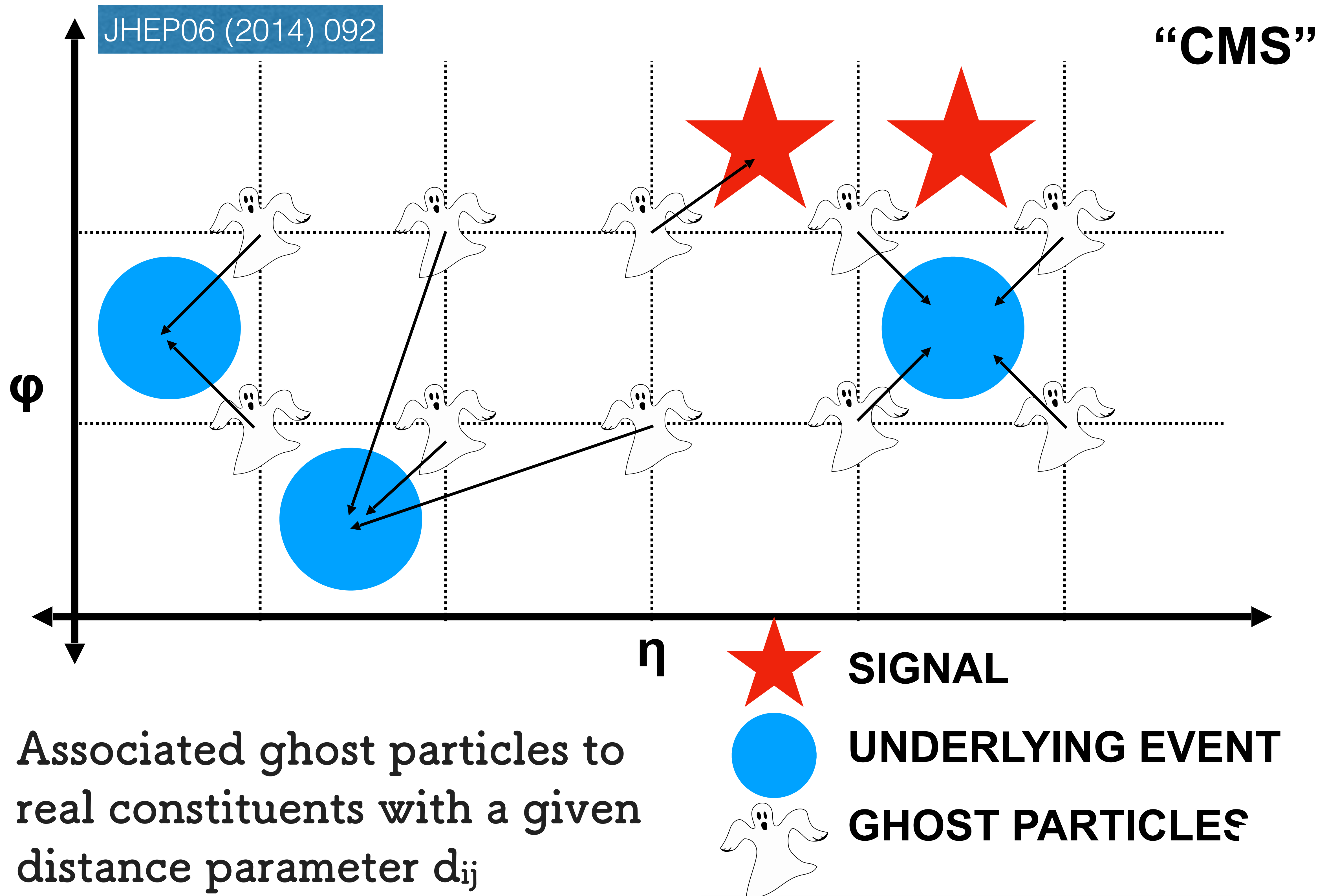


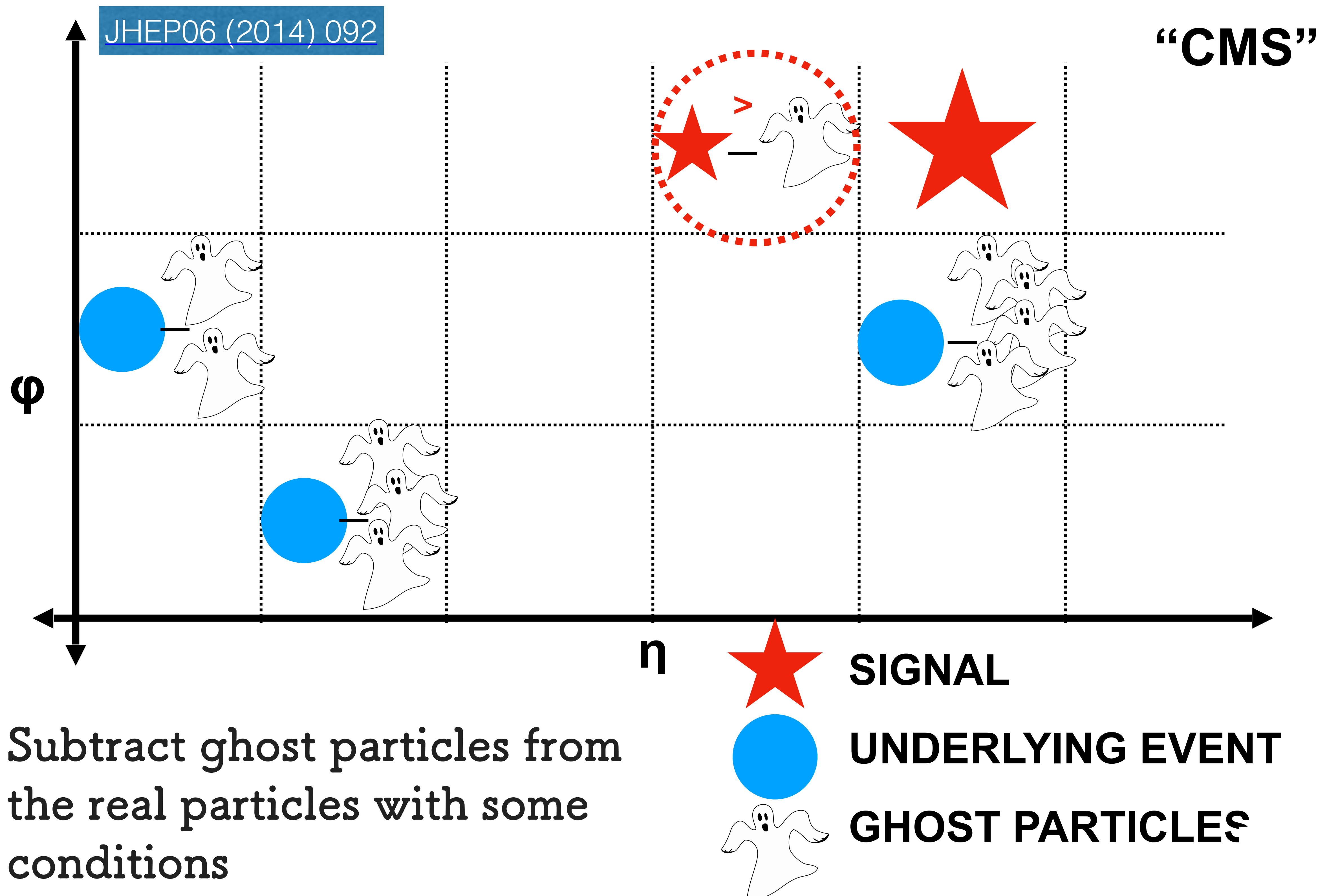
“UNDERLYING EVENT” IN HEAVY IONS

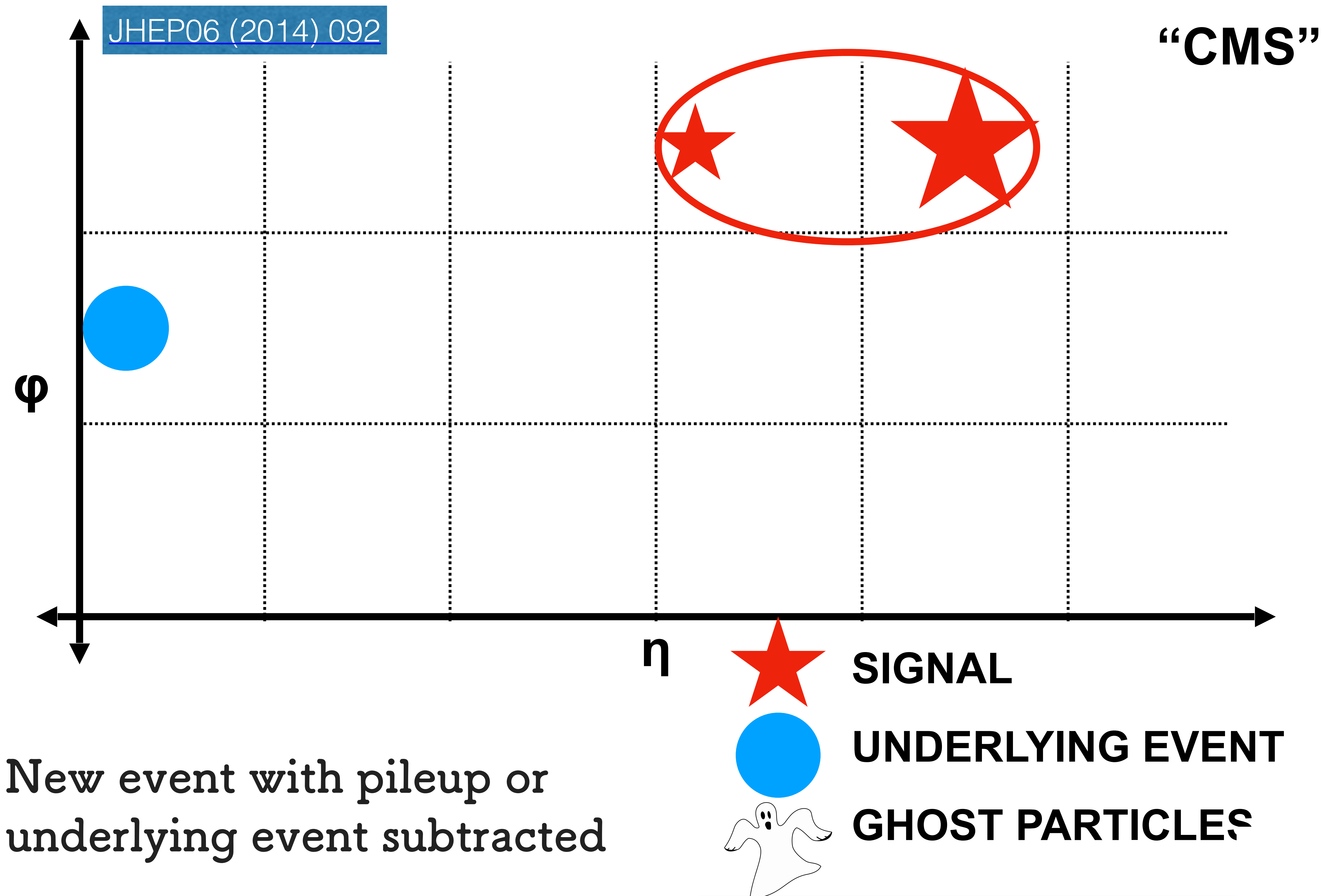




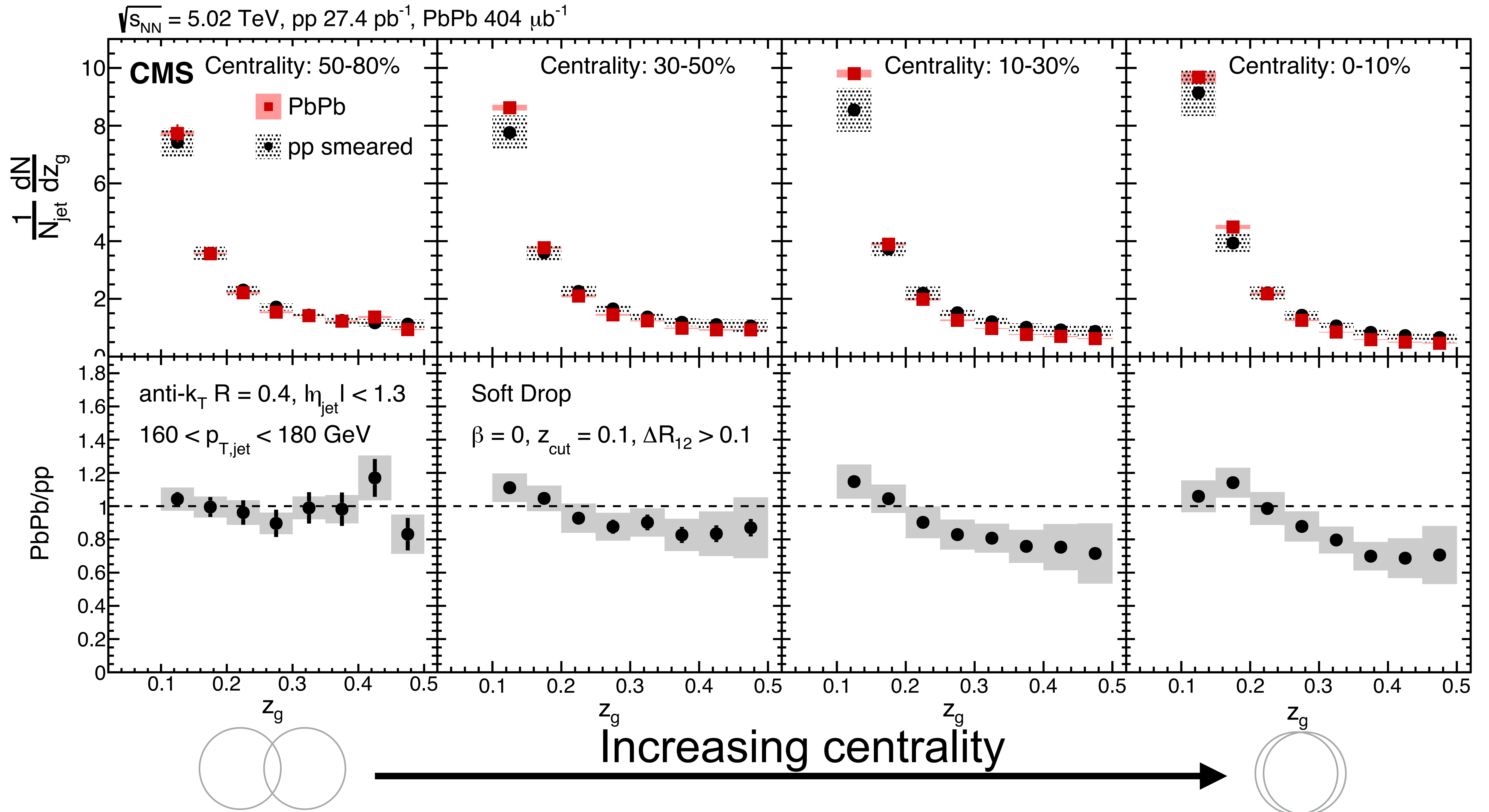
Populate the event with ghost particles of fixed area A_g and $pT_g = \rho \times A_g$







Example: modification of substructure splitting function in HI!

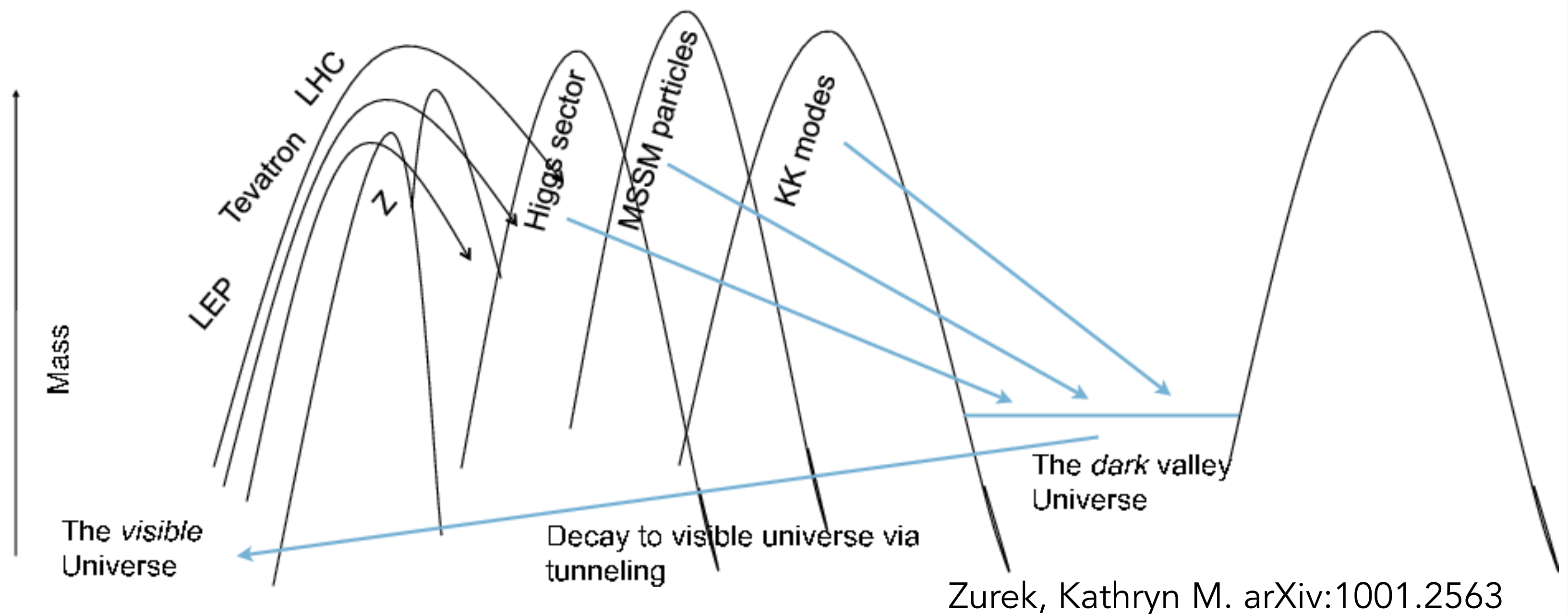


3D. VERY EXOTIC OBJECTS

Long-lived Theoretical Motivations

Including but not limited to:

- Split SUSY
- Baryogenesis
- Twin Higgs
- RPV SUSY
- Emerging Jets
- Semi-visible Jets
- Dark Photons
- GMSB
- Hidden Valley Models



As purely kinematics gains from the LHC diminish exotic decays **continue to indirectly probe higher energy scales**

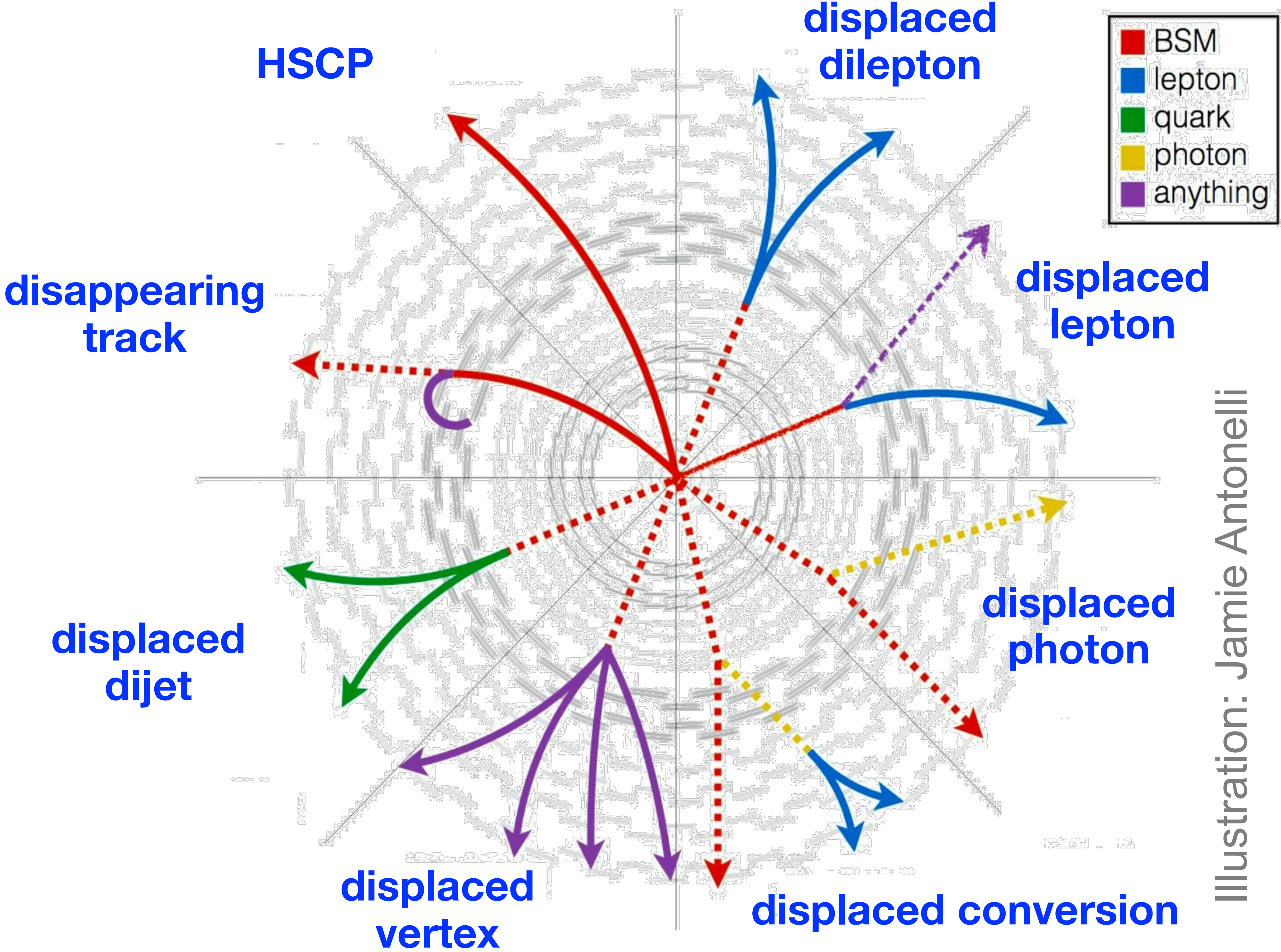
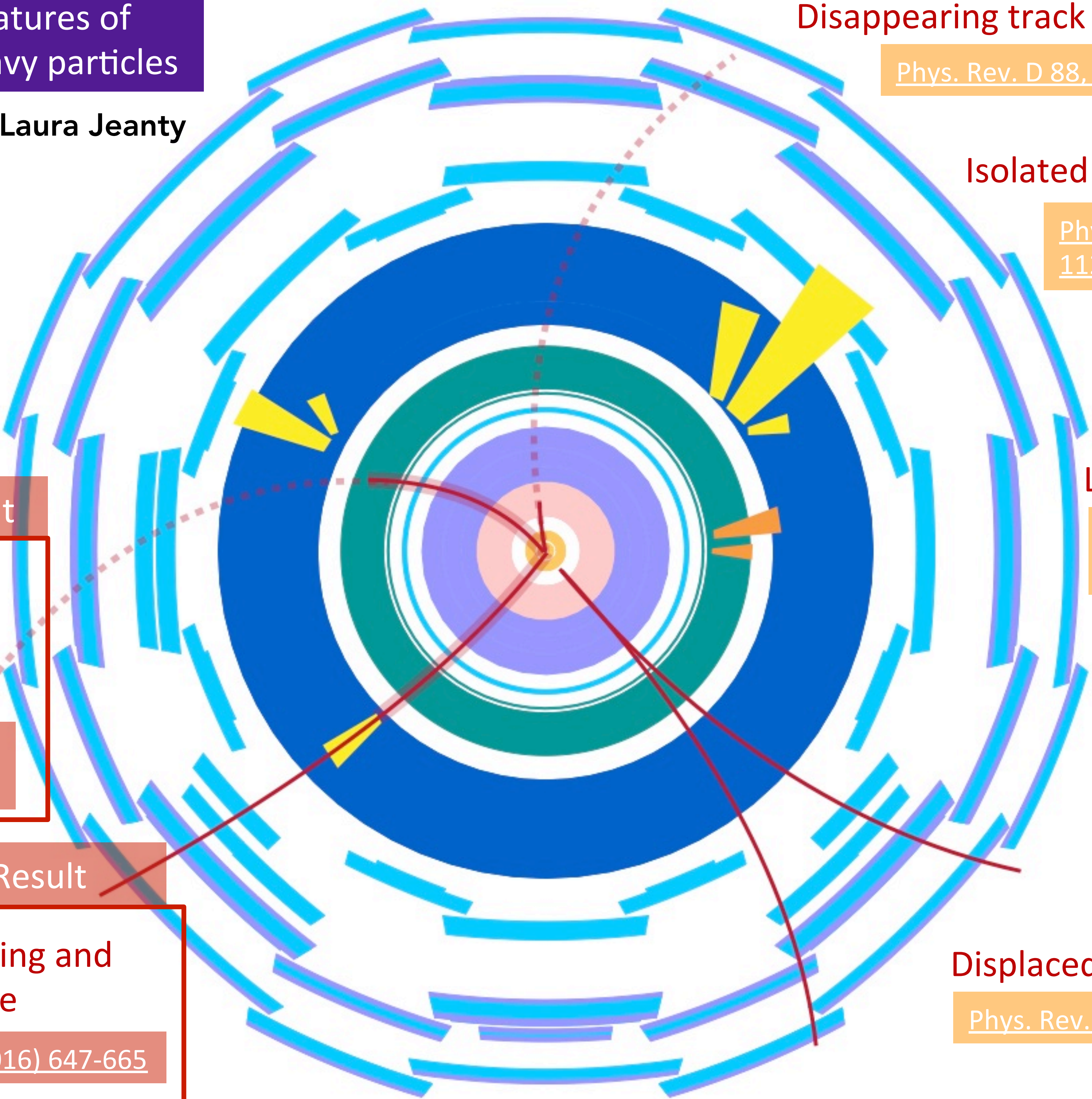


Illustration: Jamie Antonelli

Detector signatures of long-lived heavy particles

Graphic Credit: Laura Jeanty

not shown:
lepton jets



Disappearing track

Phys. Rev. D 88, 112006

Isolated / late jets

Phys. Rev. D 88, 112003

Late photons

Phys. Rev. D 90, 112005

13 TeV Result
Highly ionizing particle
Phys. Rev. D 93, 112015

13 TeV Result
Highly ionizing and slow particle
Phys. Let B (2016) 647-665

Displaced vertex

Phys. Rev. D 92, 072004

A rich variety of signals

Displaced signals at $c\tau > 1\text{mm}$

Reminder: prompt and displaced not exclusive, lifetime distribution $\sim e^{-\tau}$

Out-of-time signals

New tracking, kinked tracks,

Important to remember that we have to pass the trigger

Make sure we save such events!

This can be very non-trivial including new hardware triggers

Use the detector in creative ways!

dE/dX as a powerful discriminator

How can we use timing to improve things?

Often times, this requires developing completely new types of reconstruction algorithms!

WRAPPING UP

My goals for the lectures:

- understand how the design of the detector map into efficient reconstruction of important physics processes
- give basic concept of those reconstruction algorithms
- illustrate examples of how simple reconstruction techniques are built to create composite and complex physics objects

In the landscape of linear luminosity scaling, reconstruction is a great place to improve and extend physics capability

The detectors are more or less fixed; the luminosity is steadily increasing

Room for creativity! Think about novel, interesting, significant physics signals and how you would best detect them.

A fertile area for machine learning applications