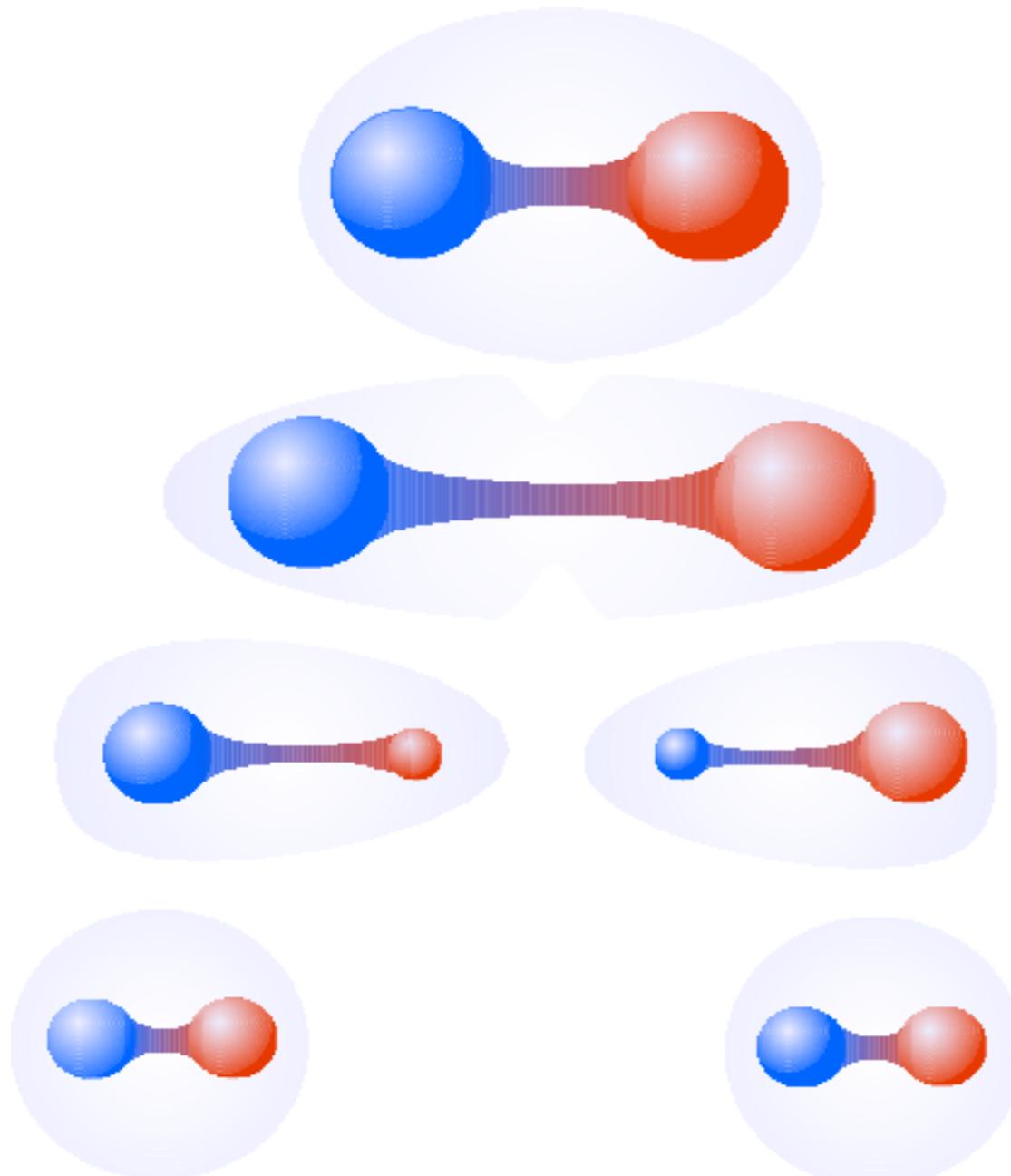


# QCD

- So what happens if you try to, say, pull apart a pion?



As you try to break  
the quarks apart,  
a quantum pair of  
quarks is pulled from  
the vacuum, and you  
get two hadrons  
instead of one!

**Quarks are confined!**

# QCD

- In terms of the running coupling constants, the quarks and gluons see MORE charge at low energy

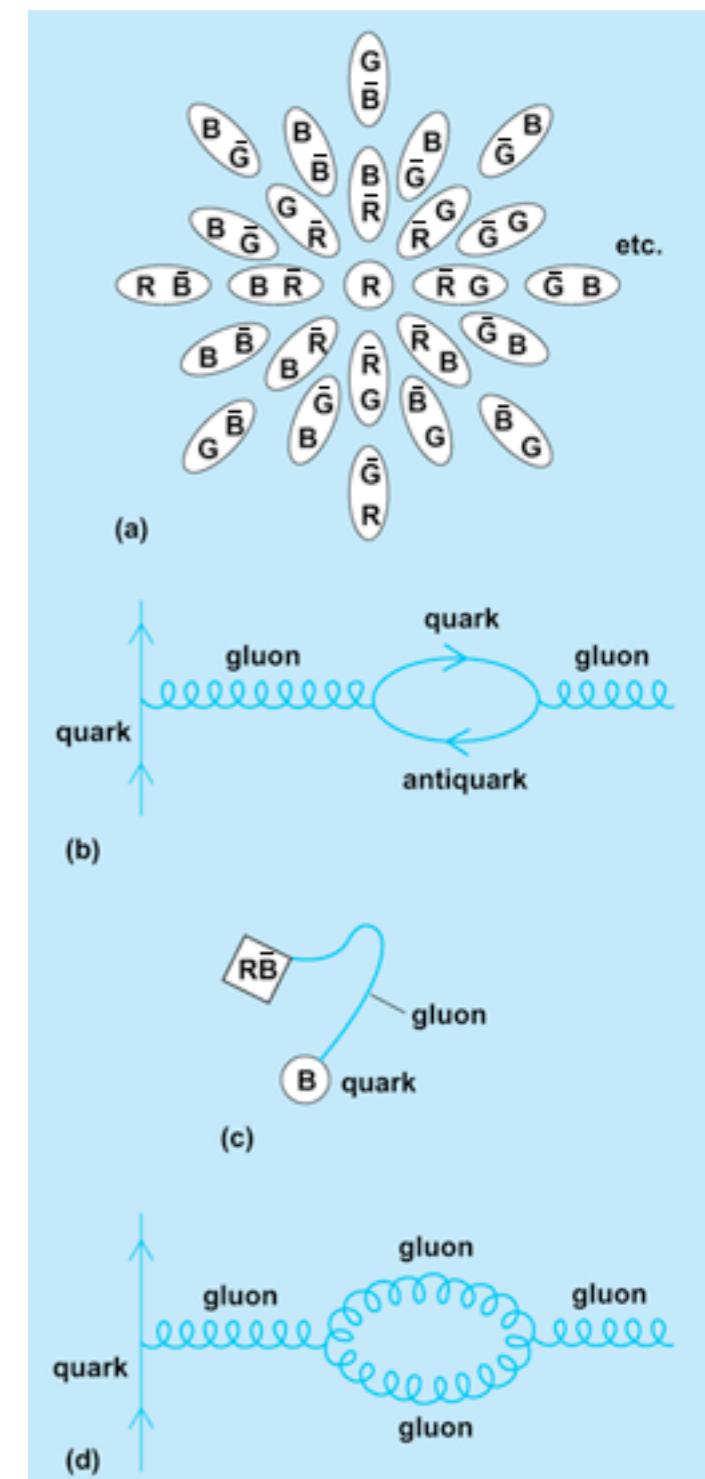
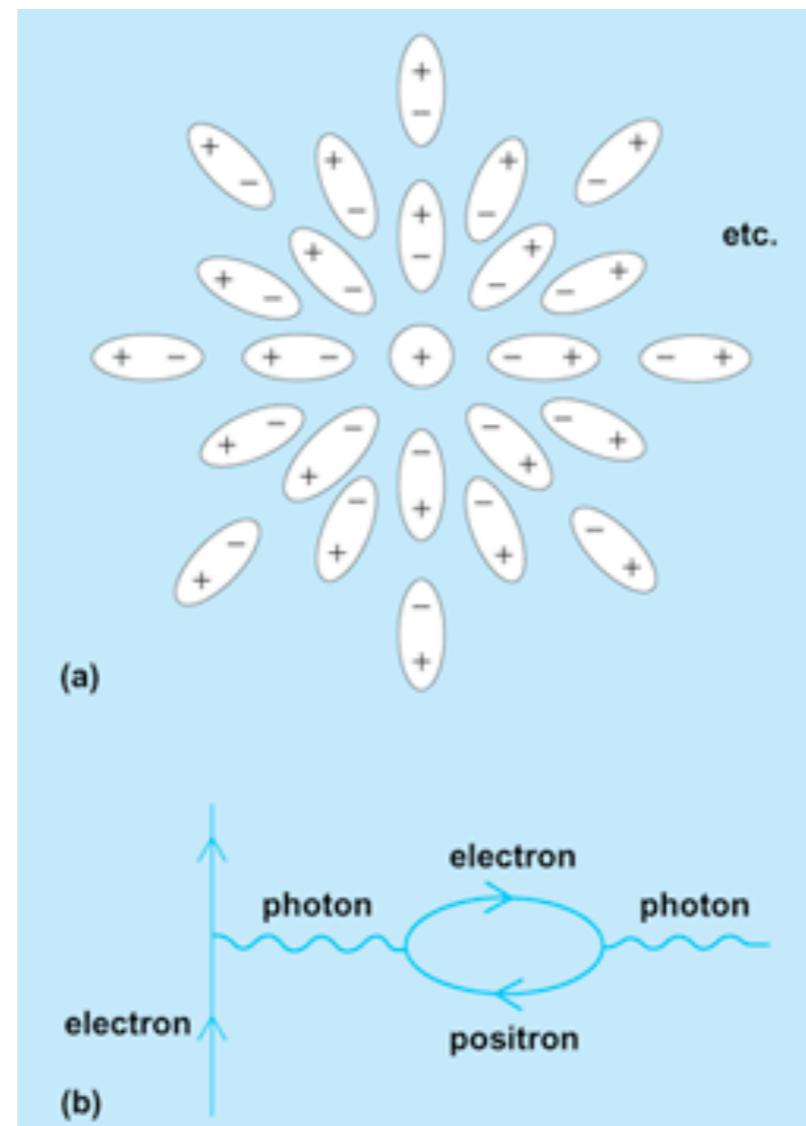
QCD

- This is “antiscreening”

<http://www.accessscience.com/overflow.aspx?SearchInputText=Quantum+chromodynamics&ContentTypeSelect=10&term=Quantum+chromodynamics&rootID=796322>

QED

- This is different from QED, which sees LESS charge at low energy



# QCD

- In terms of the running coupling constants, the quarks and gluons see MORE charge at low energy

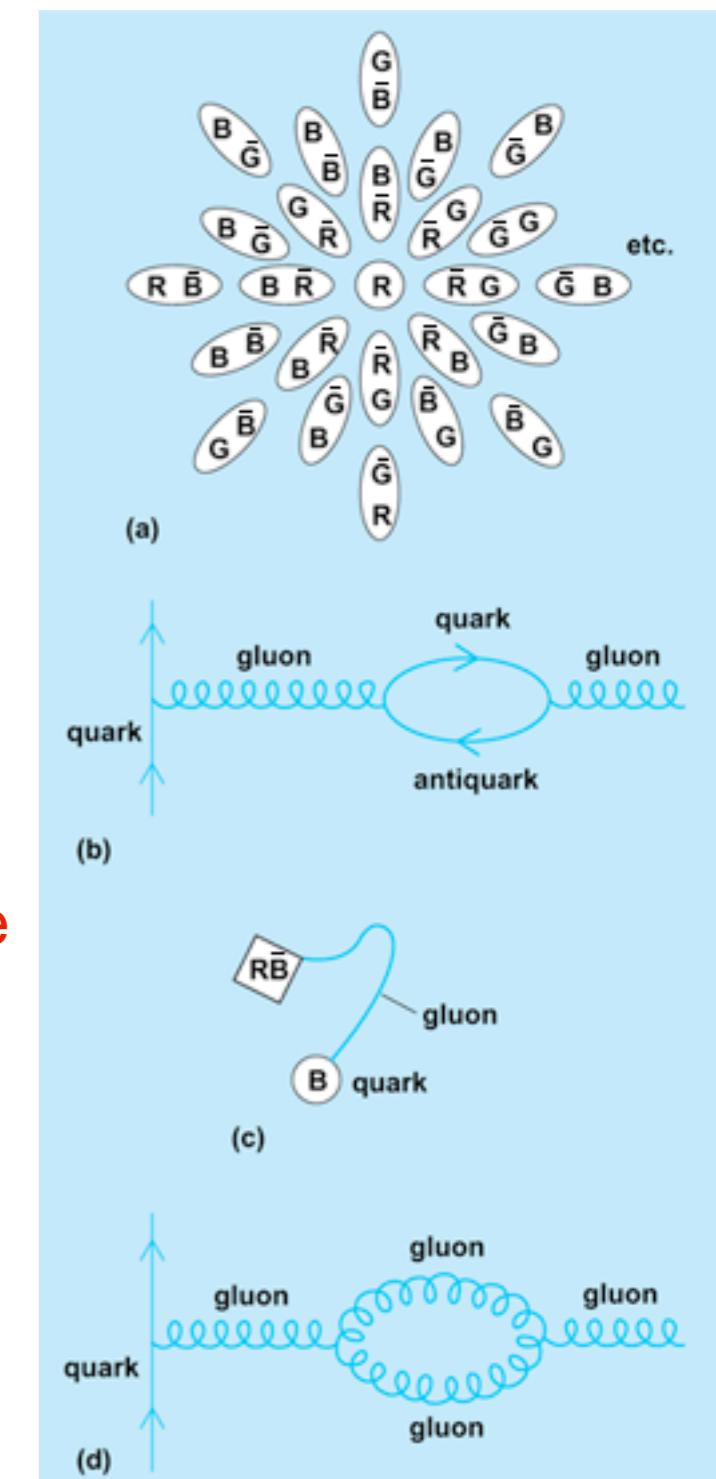
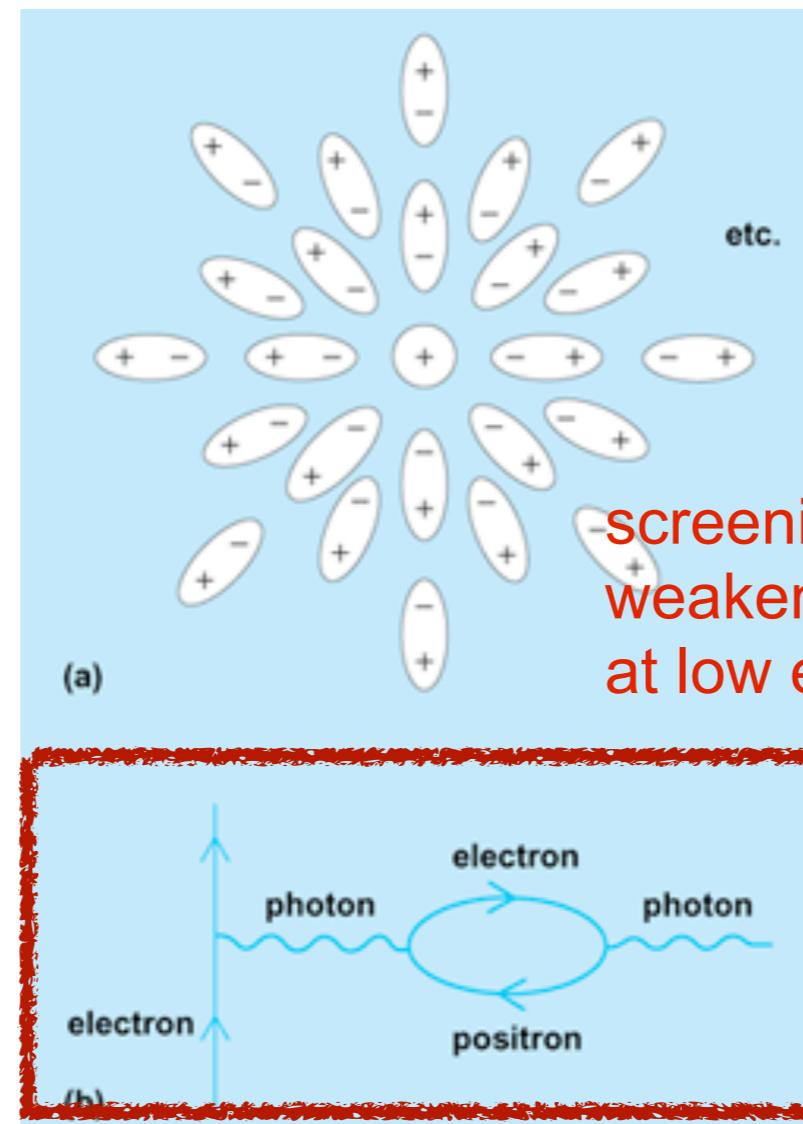
QCD

- This is “antiscreening”

<http://www.accessscience.com/overflow.aspx?SearchInputText=Quantum+chromodynamics&ContentTypeSelect=10&term=Quantum+chromodynamics&rootID=796322>

QED

- This is different from QED, which sees LESS charge at low energy



# QCD

- In terms of the running coupling constants, the quarks and gluons see MORE charge at low energy

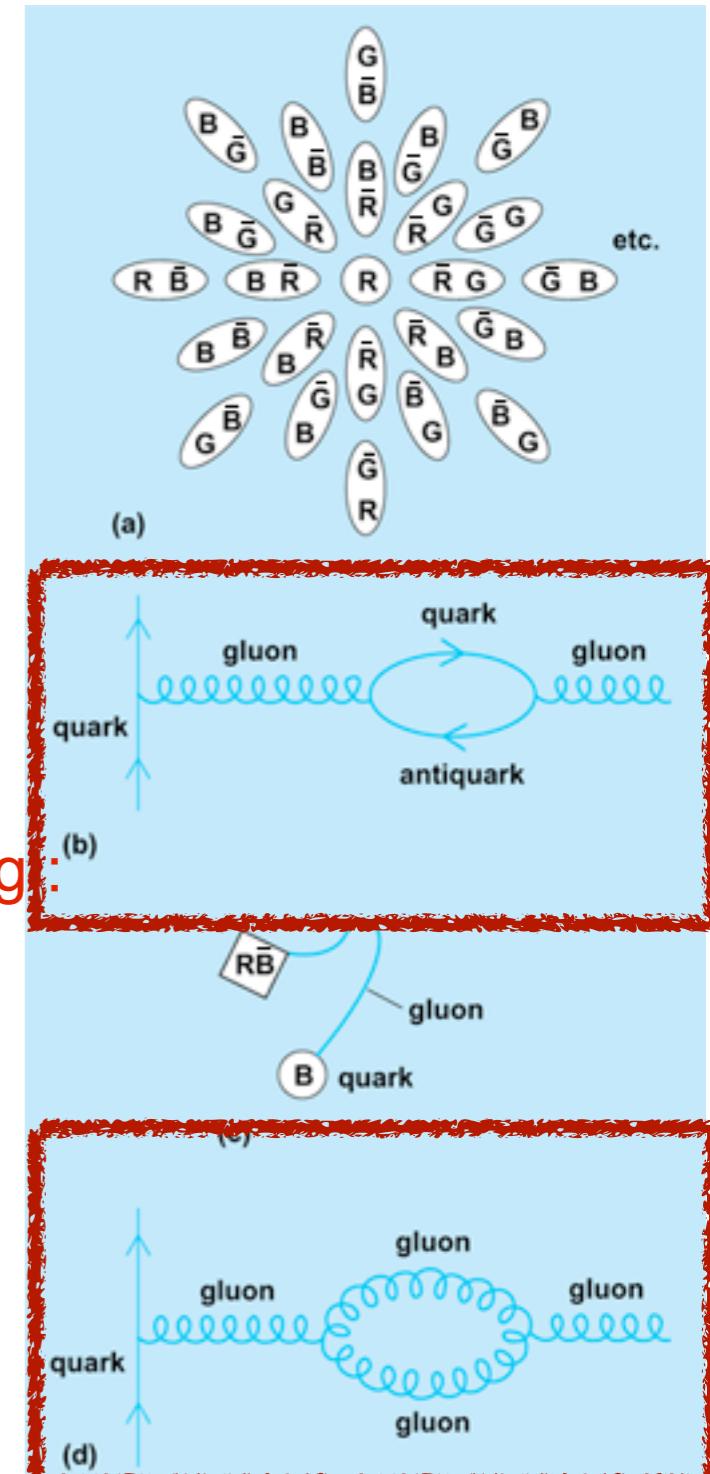
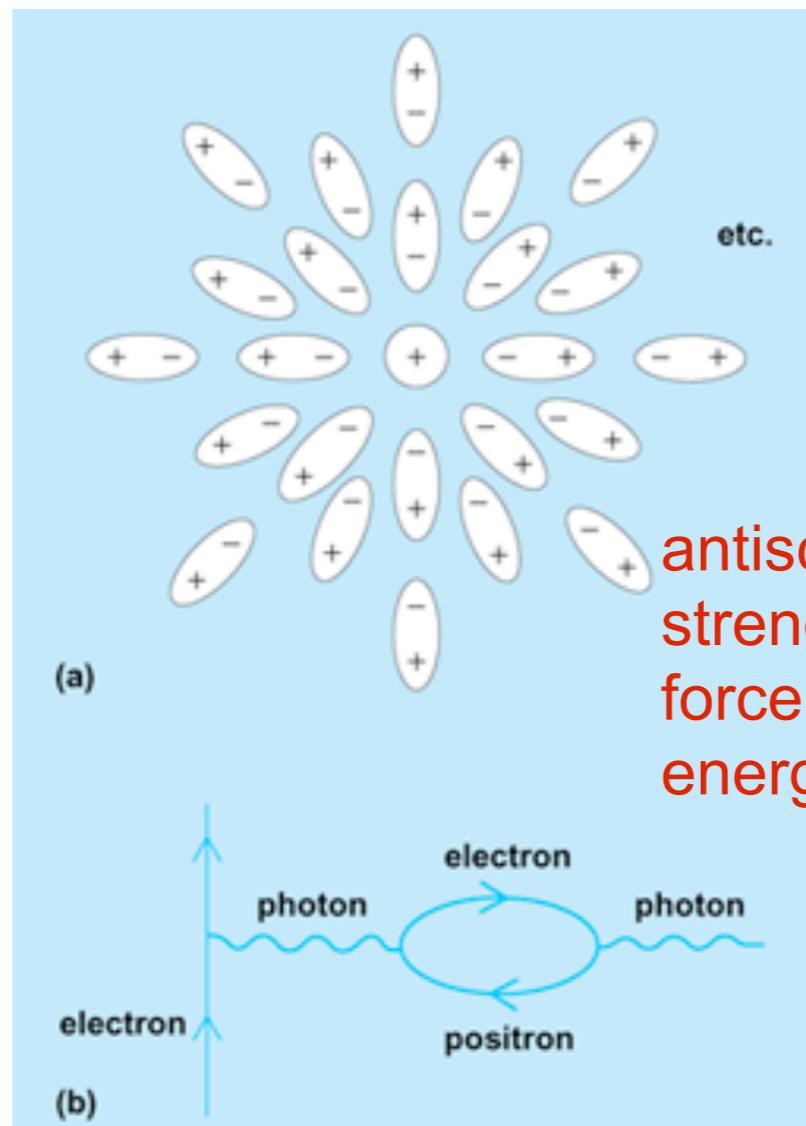
QCD

- This is “antiscreening”

<http://www.accessscience.com/overflow.aspx?SearchInputText=Quantum+chromodynamics&ContentTypeSelect=10&term=Quantum+chromodynamics&rootID=796322>

QED

- This is different from QED, which sees LESS charge at low energy



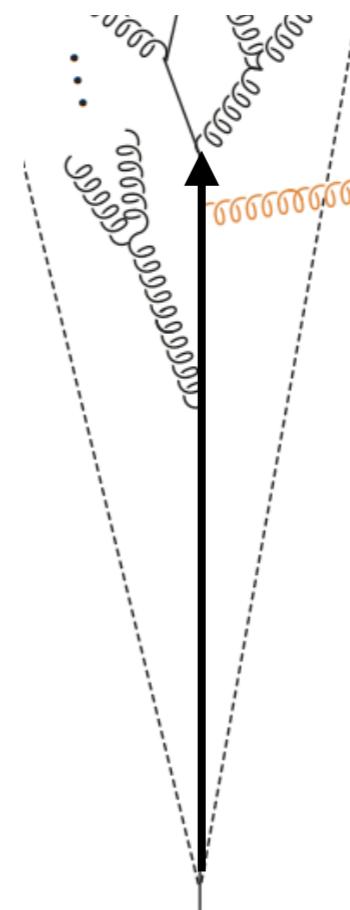
# Jets

- “Free” quark produced at the LHC...



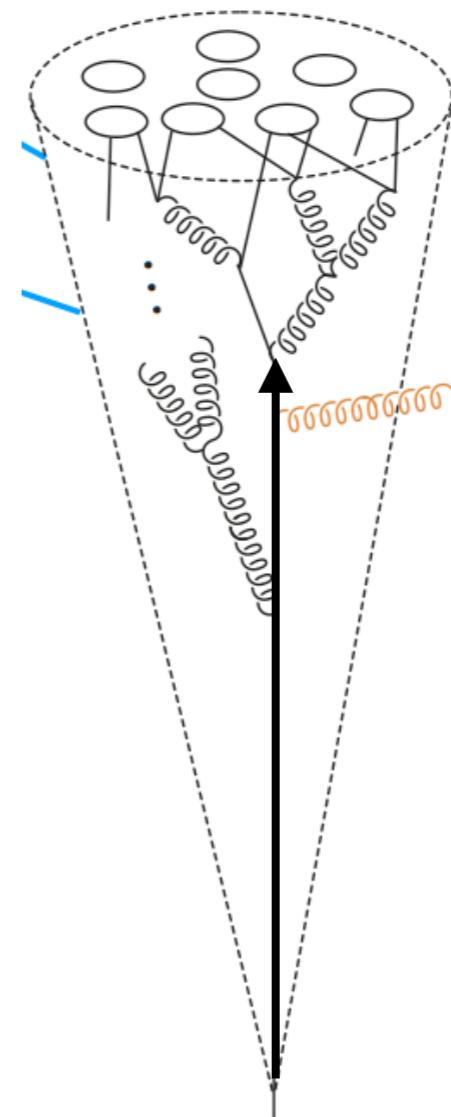
# Jets

- “Free” quark produced at the LHC... asymptotic freedom makes it radiate collinearly



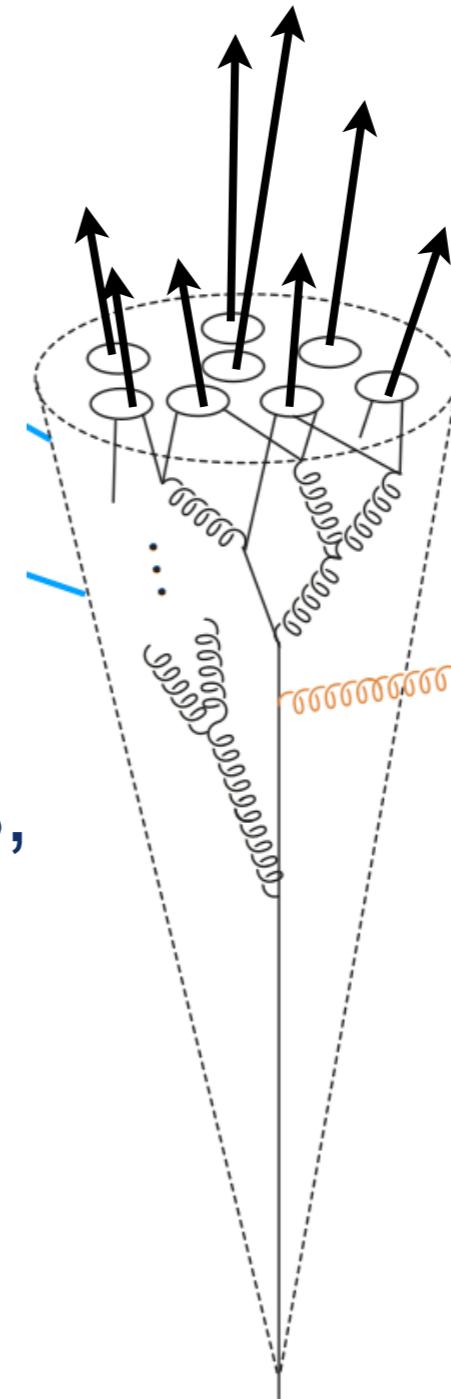
# Jets

- “Free” quark produced at the LHC... asymptotic freedom makes it radiate collinearly
- Quarks combine with other quarks from the quantum vacuum



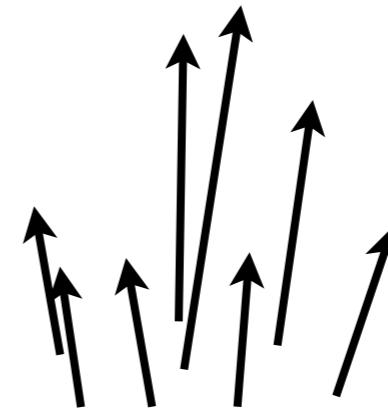
# Jets

- “Free” quark produced at the LHC... asymptotic freedom makes it radiate collinearly
- Quarks combine with other quarks from the quantum vacuum
- These fragment into observable particles like pions, kaons, protons, etc



# Jets

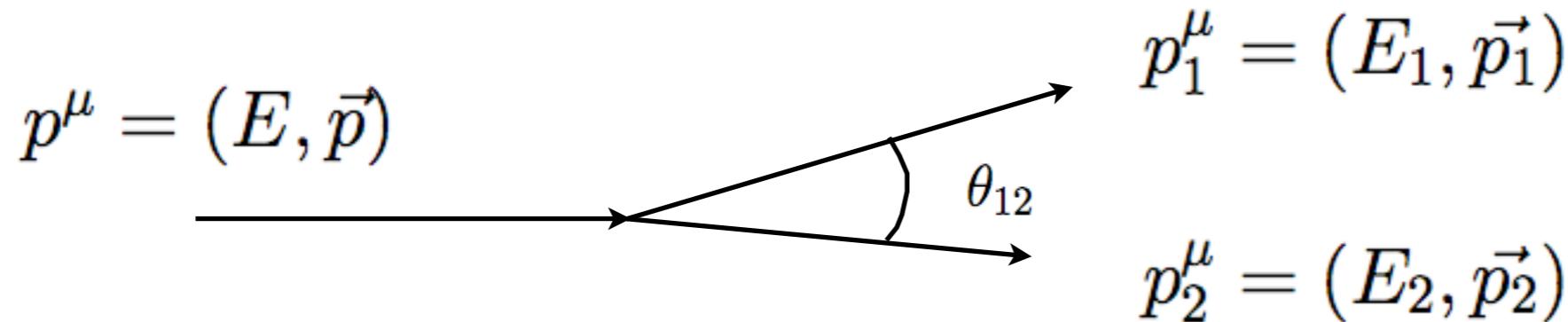
- “Free” quark produced at the LHC... asymptotic freedom makes it radiate collinearly
- Quarks combine with other quarks from the quantum vacuum
- These fragment into observable particles like pions, kaons, protons, etc
- We actually detect those remnants, not the original quark



Use “data science” algorithms  
to back these out!

(PS HEP has been doing  
data science since WAAYY  
before it was cool)

# Boosted Particles



$$\begin{aligned} p^\mu p_\mu &= (p_1 + p_2)^\mu (p_1 + p_2)_\mu \\ m^2 &= (E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2) \cdot (\vec{p}_1 + \vec{p}_2) \\ m^2 &\approx 2E_1 E_2 (1 - \cos(\theta_{12})) \end{aligned}$$

Assume  $E_1 = E_2 = E/2$

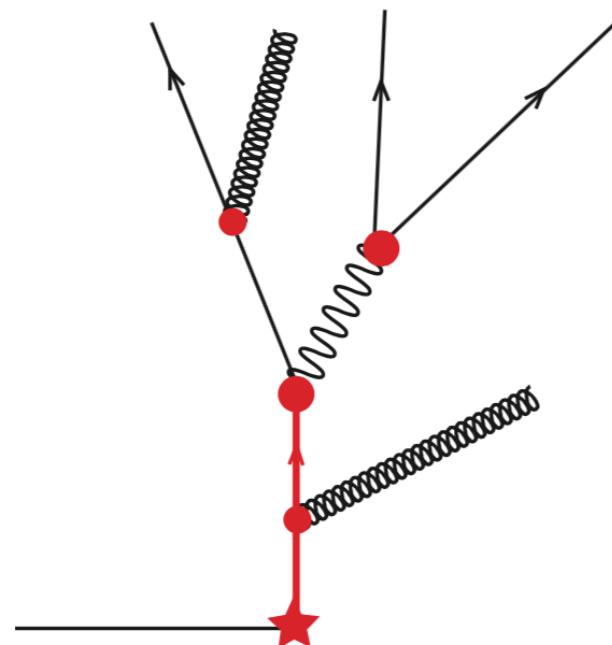
$$\begin{aligned} m^2 &\approx E^2 (1 - \cos(\theta_{12})) \\ \cos(\theta_{12}) &\approx 1 - \frac{m^2}{E^2} \approx 1 - \frac{1}{\gamma^2} \end{aligned}$$

$$\theta_{12} \approx \frac{2m}{E} = \frac{2}{\gamma}$$

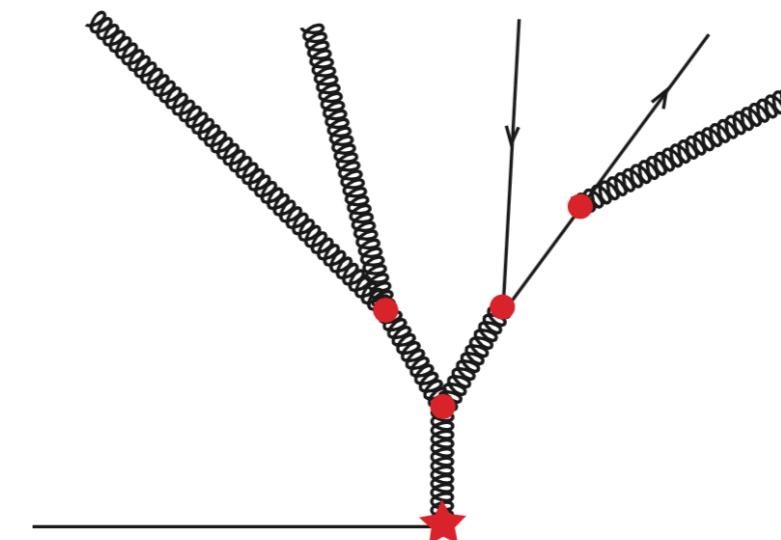
Both Jets and Boosted SM particles result in collimated sprays of particles, need to classify!  
(More later)

# Undergrad Physics → Groundbreaking Research

- Goal: Classify sprays of strongly interacting particles (“jets”):



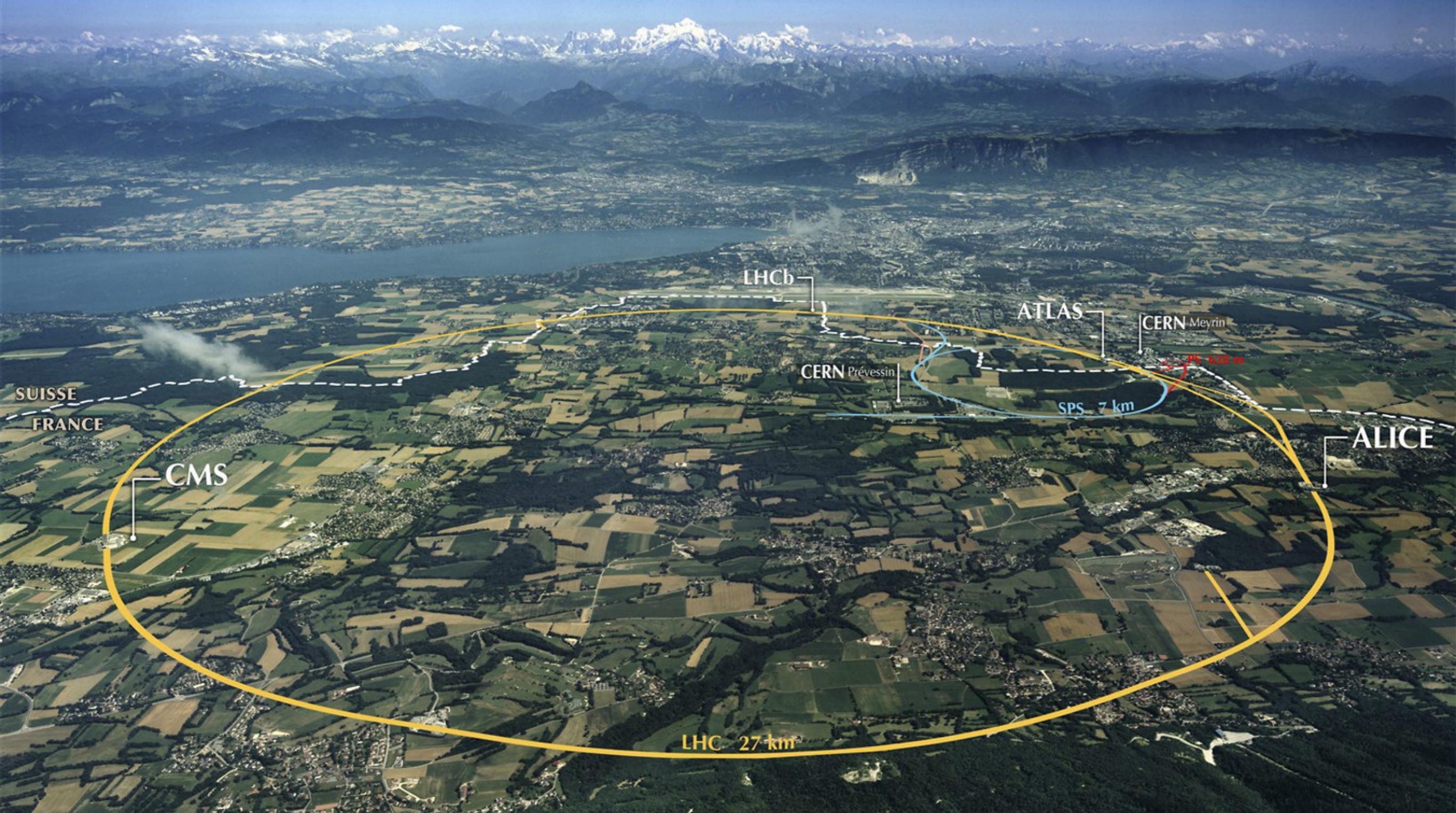
Category A:  
Heavy particles  
decaying into  
jets



Category B:  
Light particles  
decaying into  
jets

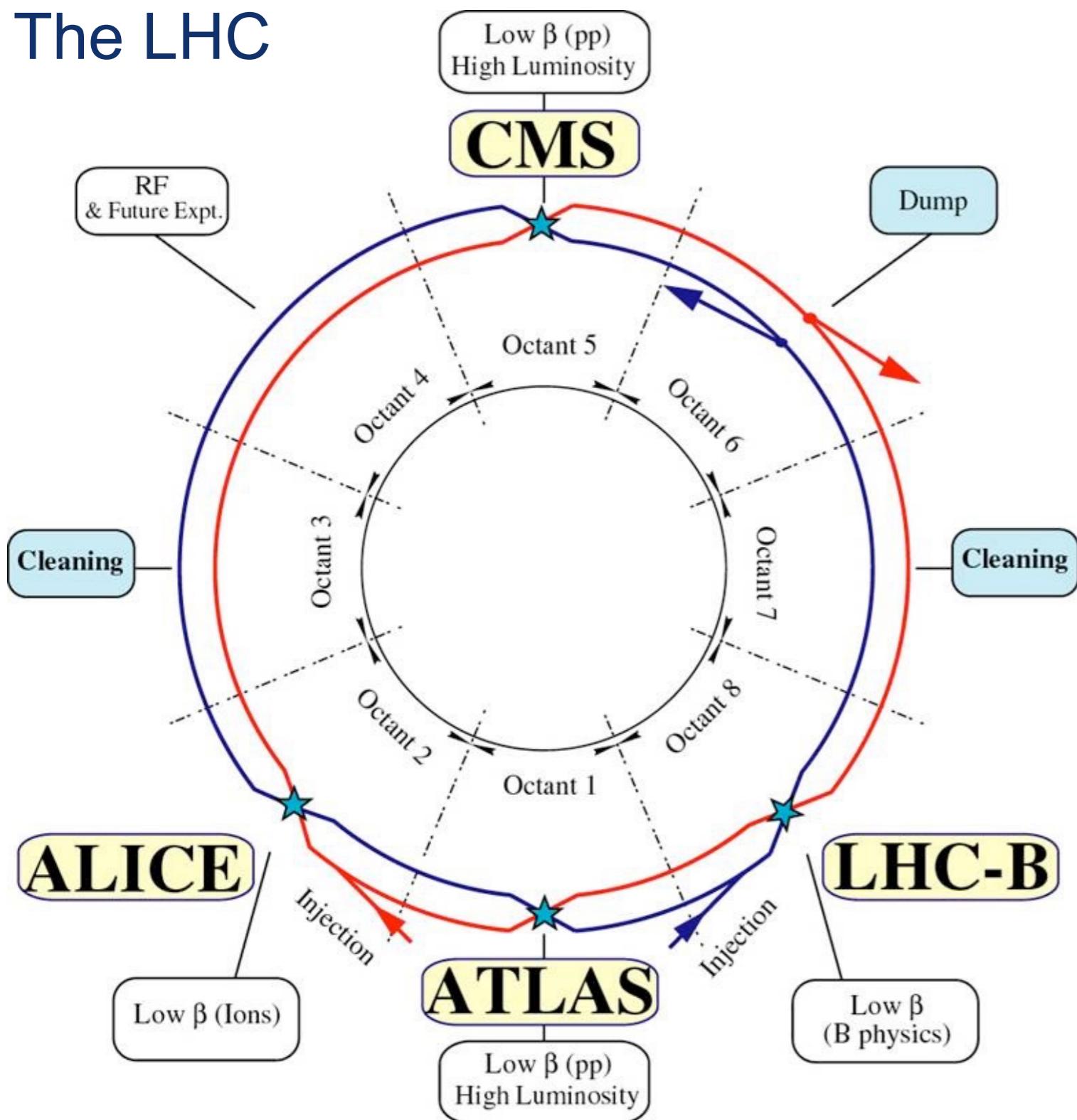
How do we get there?

# The Large Hadron Collider



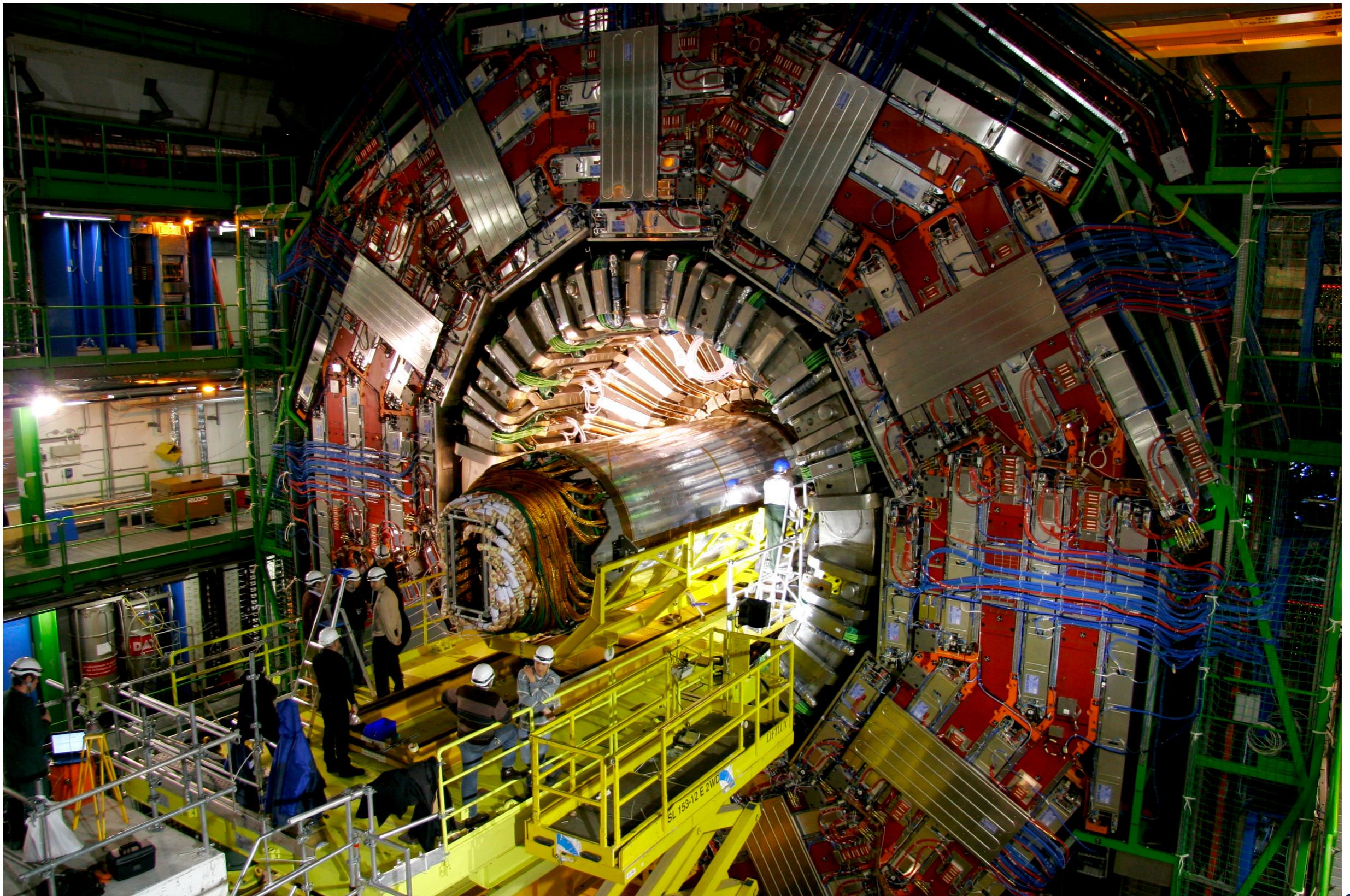
# The Large Hadron Collider

## The LHC

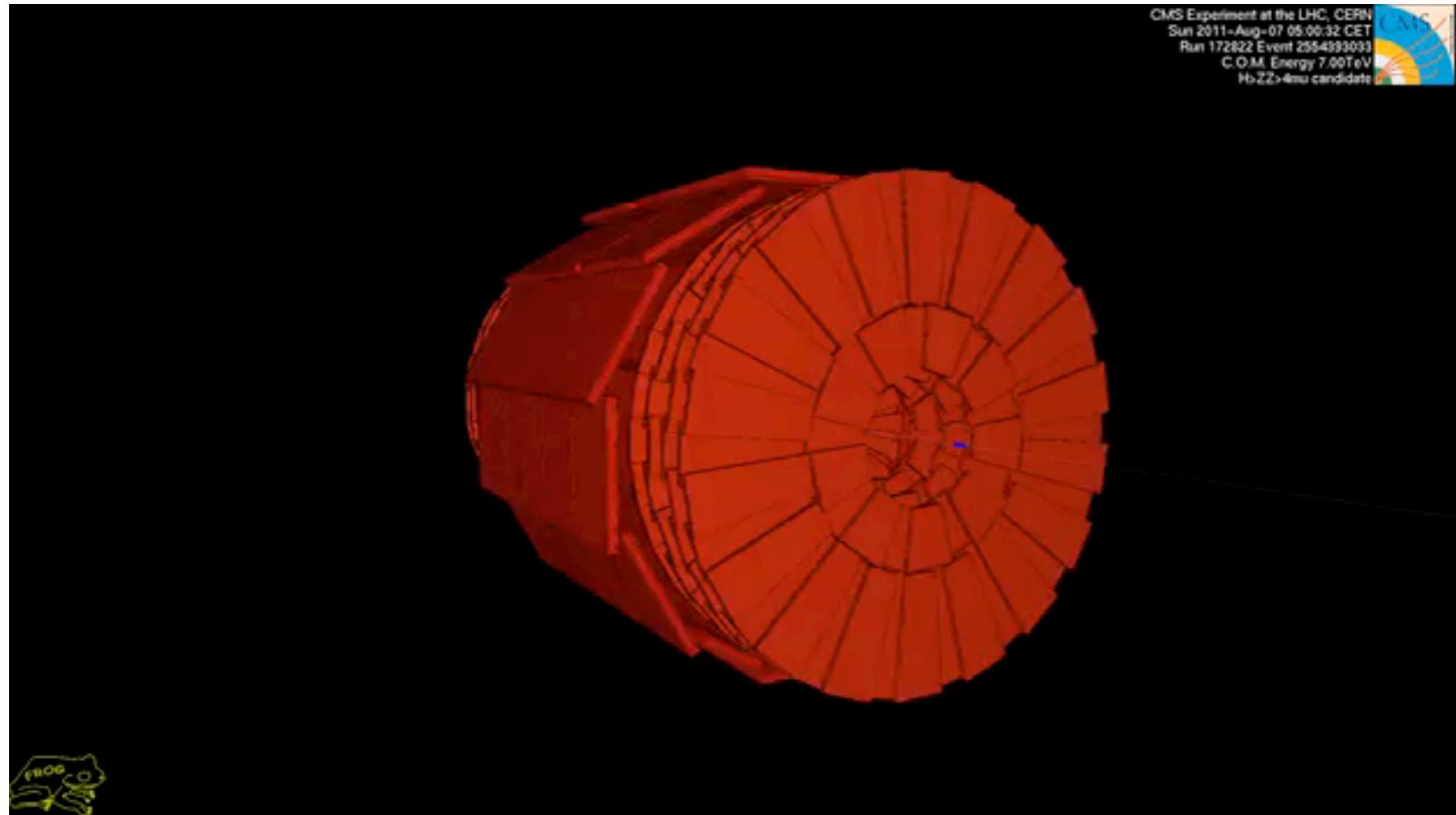


- Geneva, Switzerland and surrounding areas in France
- 27 kilometer circumference
- 50-175 meters underground
- Proton-on-proton collisions
  - 7 TeV : 2010-2011
  - 8 TeV : 2012
  - 13 TeV : 2015-2018
  - >13 TeV: 2021+

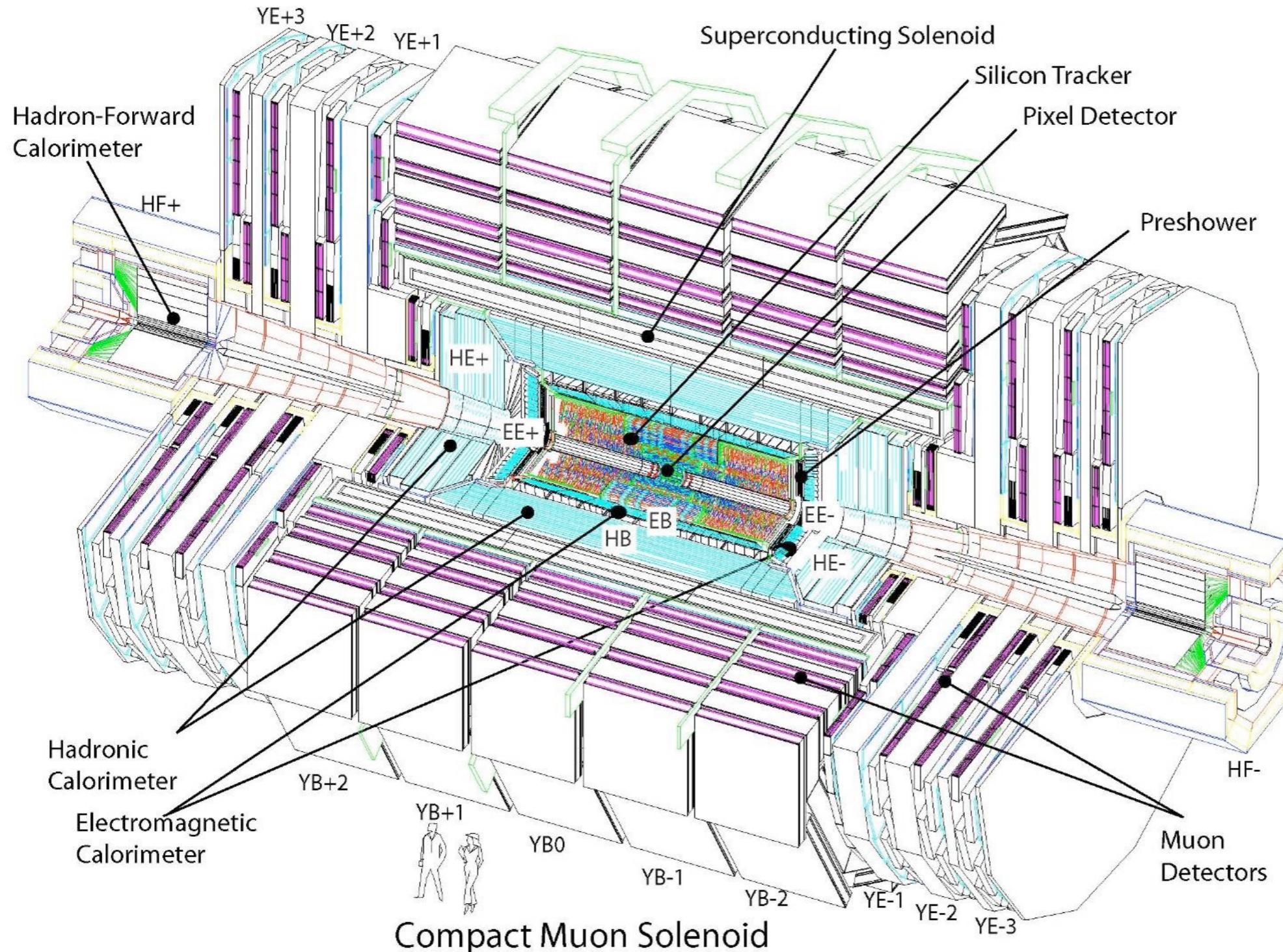
# The Compact Muon Solenoid



# The Compact Muon Solenoid



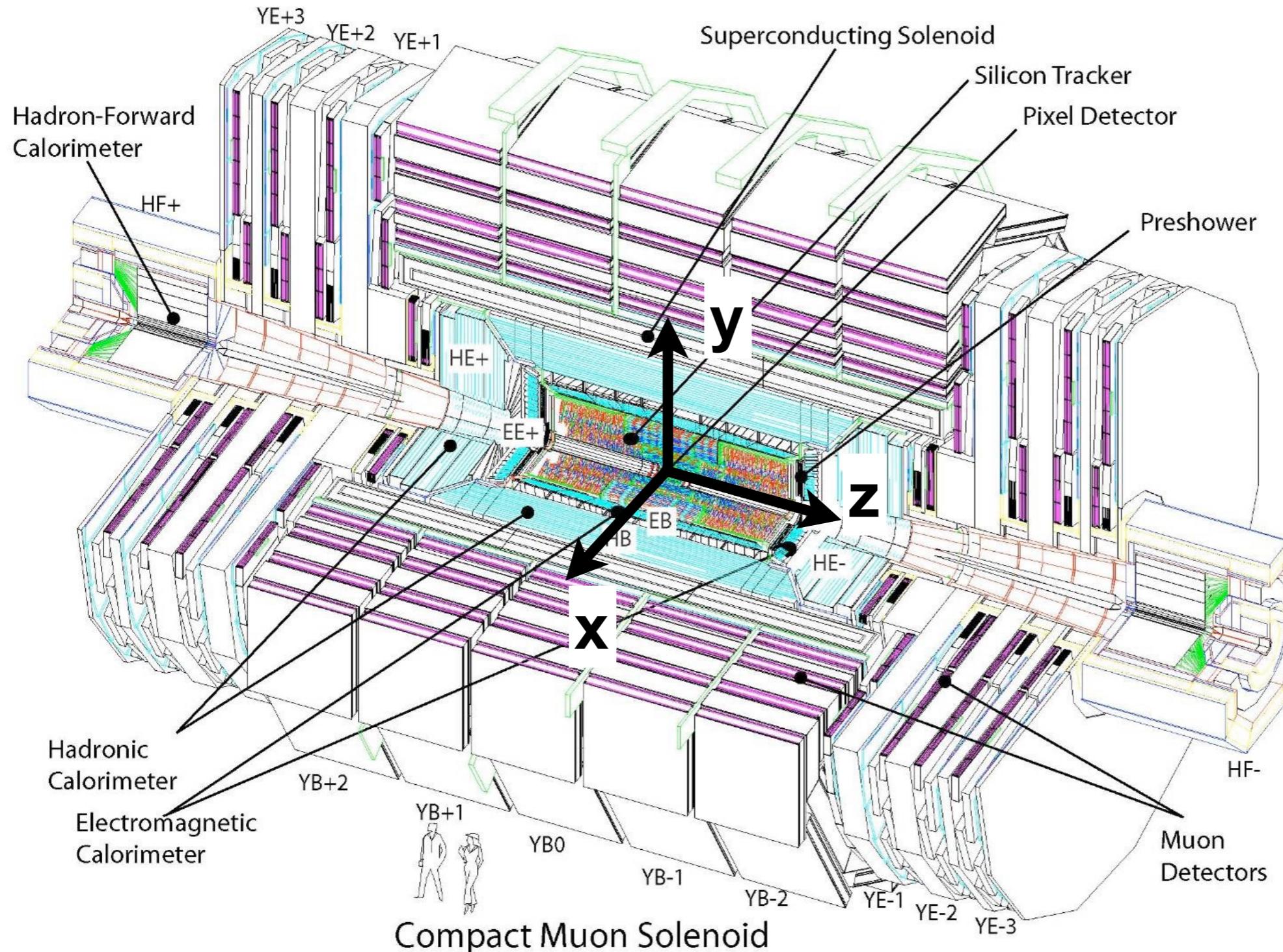
# The Compact Muon Solenoid



$\varphi$  : Azimuthal angle around z-axis

$\theta$  : polar angle from z-axis

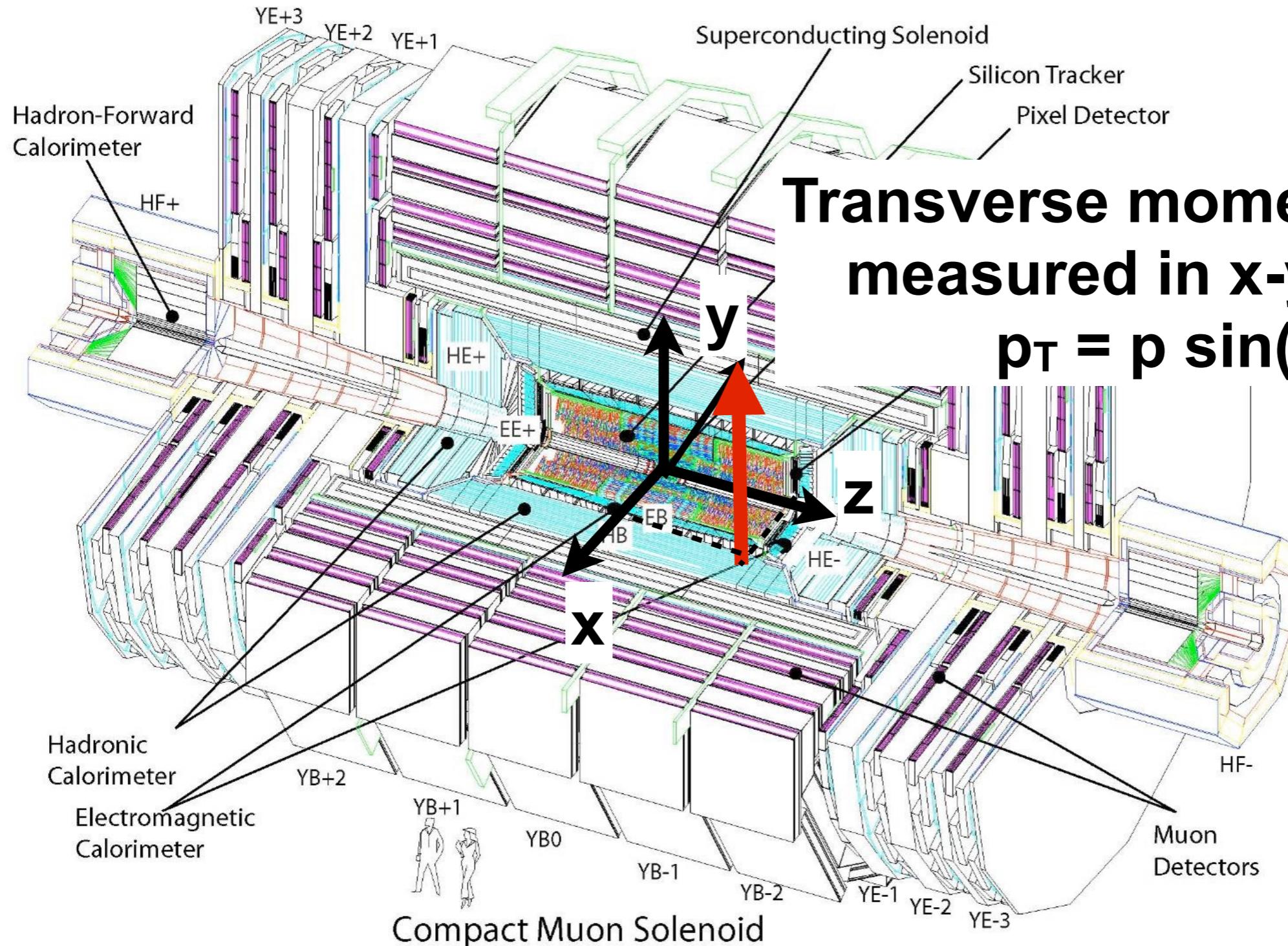
# The Compact Muon Solenoid



$\varphi$  : Azimuthal angle around z-axis

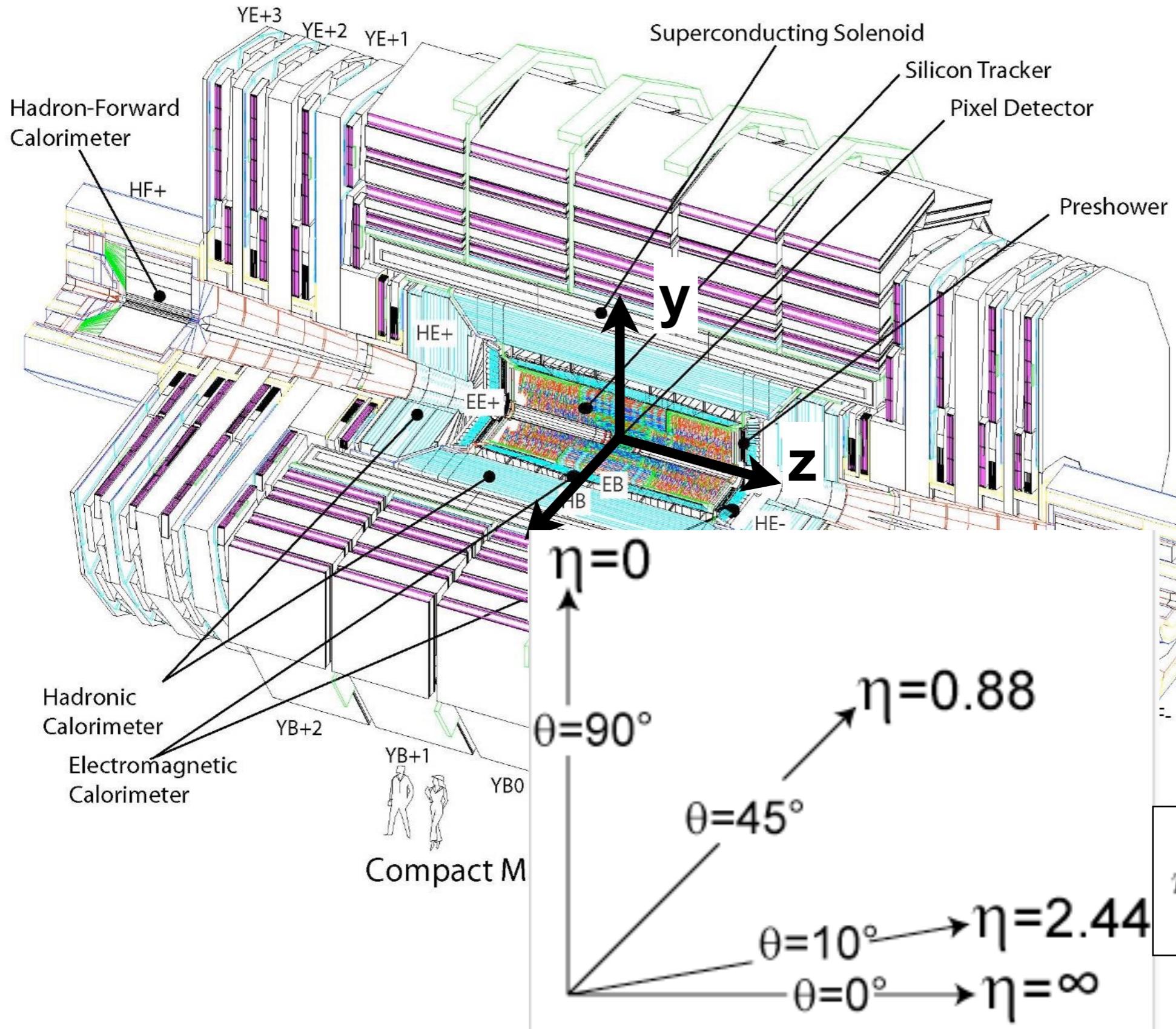
$\theta$  : polar angle from z-axis

# The Compact Muon Solenoid



**Transverse momentum:  $p_T$   
measured in x-y plane**  
 $p_T = p \sin(\theta)$

# The Compact Muon Solenoid



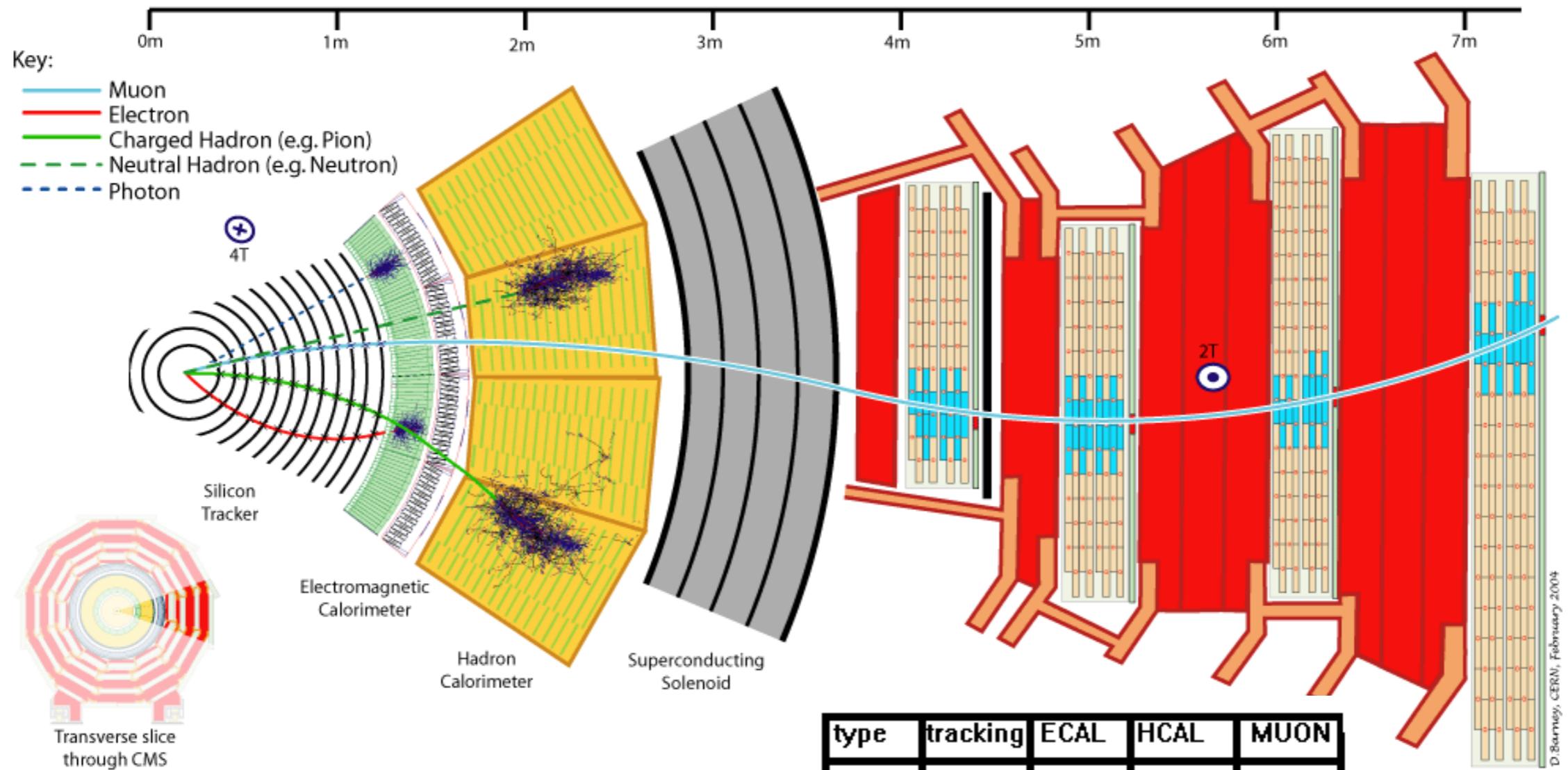
$\varphi$  : Azimuthal angle around z-axis

$\theta$  : polar angle from z-axis

$\eta$  : pseudo-rapidity, invariant under Lorentz boosts

$$\eta = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right]$$

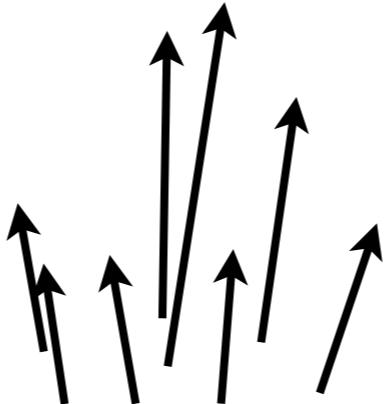
# The Compact Muon Solenoid



type	tracking	ECAL	HCAL	MUON
$\gamma$		→		
e	→	→		
$\mu$	→			→
Jet		→	→	
$E_T$ miss				

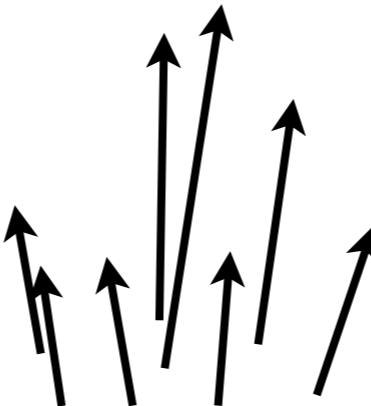
D.Barney, CERN, February 2004

# Clustering



- Clustering: grouping  $N$  particles into  $k$  jets
  - Hey! Sounds like a data science algorithm!

# Clustering



- Clustering: grouping  $N$  ~~particles~~ into  $k$  ~~jets~~  
points in space      partitions

# Clustering

- Goal: Given  $N$  points in space, associate to  $k$  partitions

- Many applications:

- Data classification

- Galaxy clustering

- Jet clustering

- Sociology

- Social networks



(a) Original points.



(b) Two clusters.



(c) Four clusters.

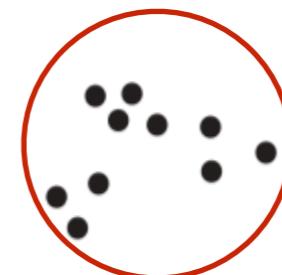


(d) Six clusters.

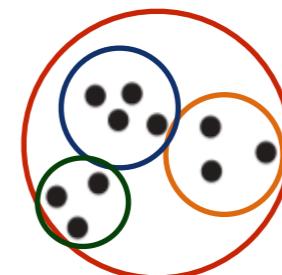
# Clustering

- Uses:
  - Summarize / compress “spatial” information
  - Find nearest neighbors
- Types of clustering algorithms:
  - Exclusive:
    - Partitional: divide into k exclusive categories
    - Hierarchical: can have sub-clusters
  - Non-exclusive:
    - Add same element to more than one cluster
  - Fuzzy:
    - Weight elements according to cluster
- Each can be either complete or partial

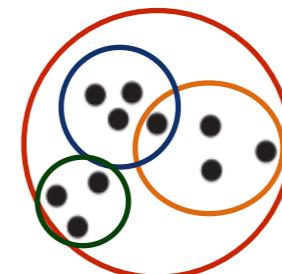
Partitional



Hierarchical

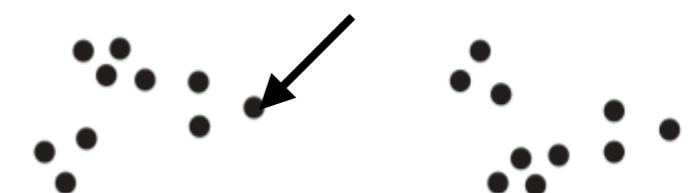


Non-exclusive



Fuzzy

$$w_{11}=0.99, w_{12}=0.01$$



# Clustering

- Popular categories :
  - K-means
    - Partition into k clusters using mean central values (usually exclusively)
  - Agglomerative hierarchical clustering
    - Pair individual elements into clusters given some distance metric
  - Density based scan
    - Considers low-density regions to be noise, not exclusive clustering

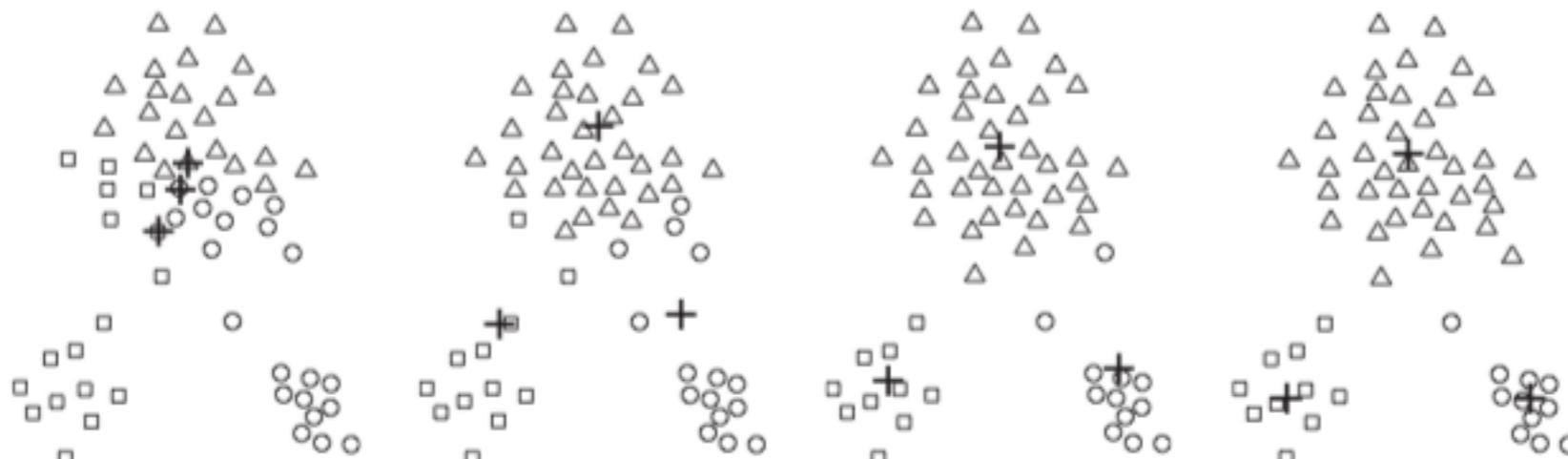
# Clustering

- K-means algorithm:

---

**Algorithm 8.1** Basic K-means algorithm.

- 
- 1: Select  $K$  points as initial centroids.
  - 2: **repeat**
  - 3:   Form  $K$  clusters by assigning each point to its closest centroid.
  - 4:   Recompute the centroid of each cluster.
  - 5: **until** Centroids do not change.
- 



(a) Iteration 1.

(b) Iteration 2.

(c) Iteration 3.

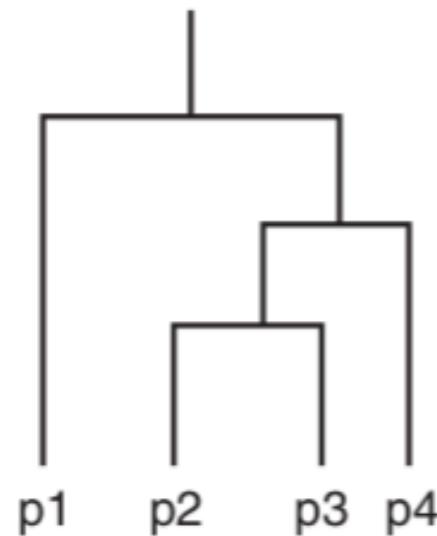
(d) Iteration 4.

# Clustering

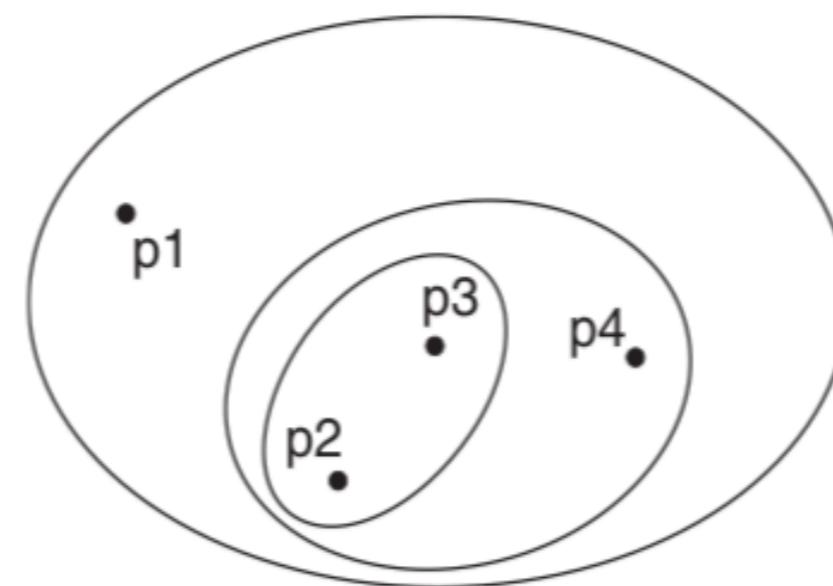
- Computational complexity is ~linear in product of:
  - Number of points
  - Number of “dimensions” (or attributes)
  - Number of clusters
  - Number of iterations to converge
- Shortcomings of k-means:
  - As in all minimization routines, danger of local minima
    - Need heuristic methods to avoid them
  - Can result in empty clusters if initialized poorly
  - Outliers have disproportionate impact
- Can try to split and merge centroids to mitigate these!

# Clustering

- Hierarchical clustering:
  - Agglomerative (“bottom up”):
    - Start with individual constituents, merge until criteria met
  - Divisive (“top down”):
    - Start with conglomerate, split until criteria met (or you get to individual constituents)
  - Represent by a tree (“dendrogram”) or Venn diagram (“nested cluster diagram”):Basically the same,  
but in reverse



(a) Dendrogram.



(b) Nested cluster diagram.

# Clustering

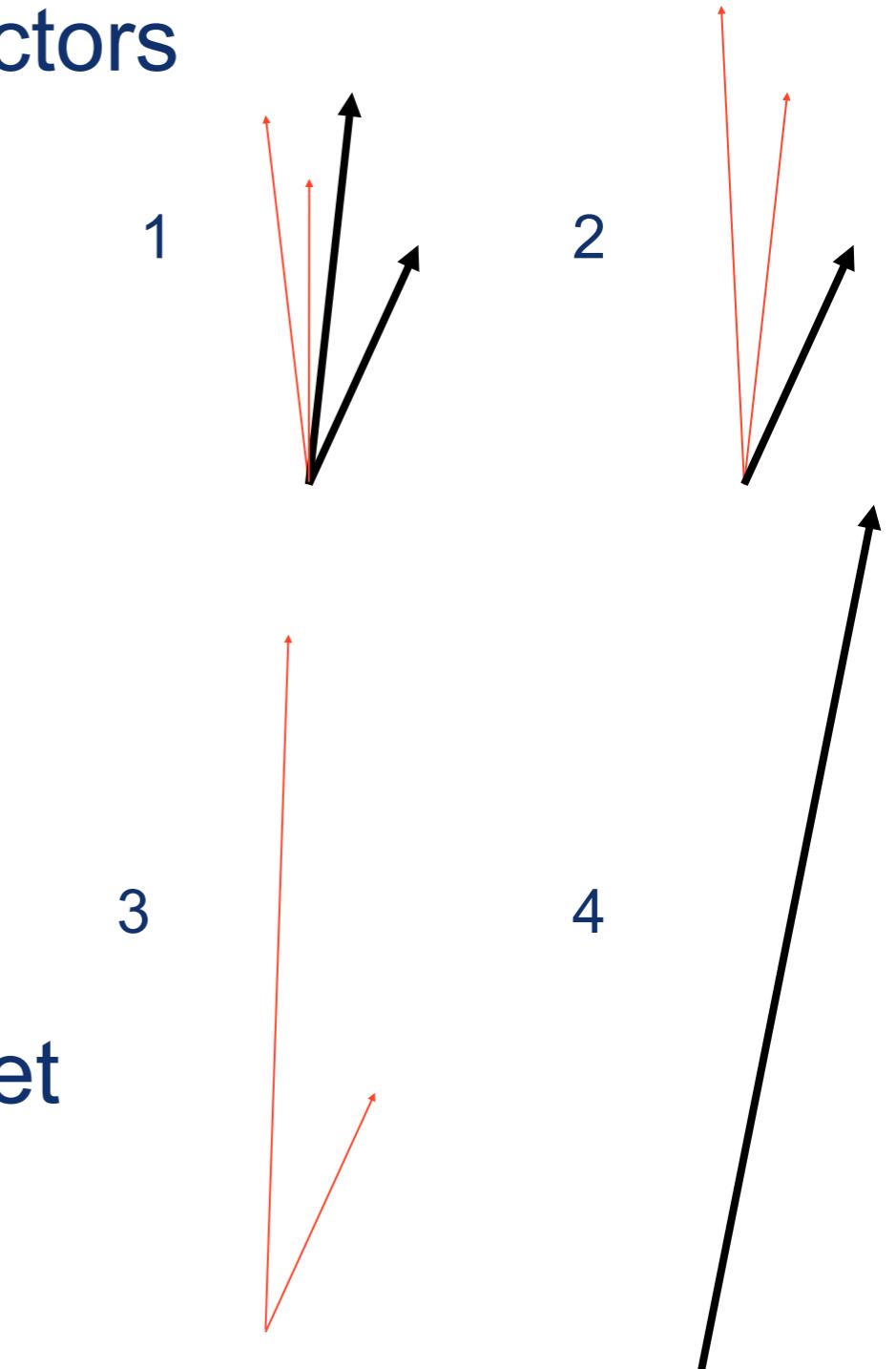
- Pairwise examination of input 4-vectors
- Calculate  $d_{ij}$

$$d_{ij} = \min(k_{ti}^n, k_{tj}^n) \Delta R_{ij}^2 / R^2$$

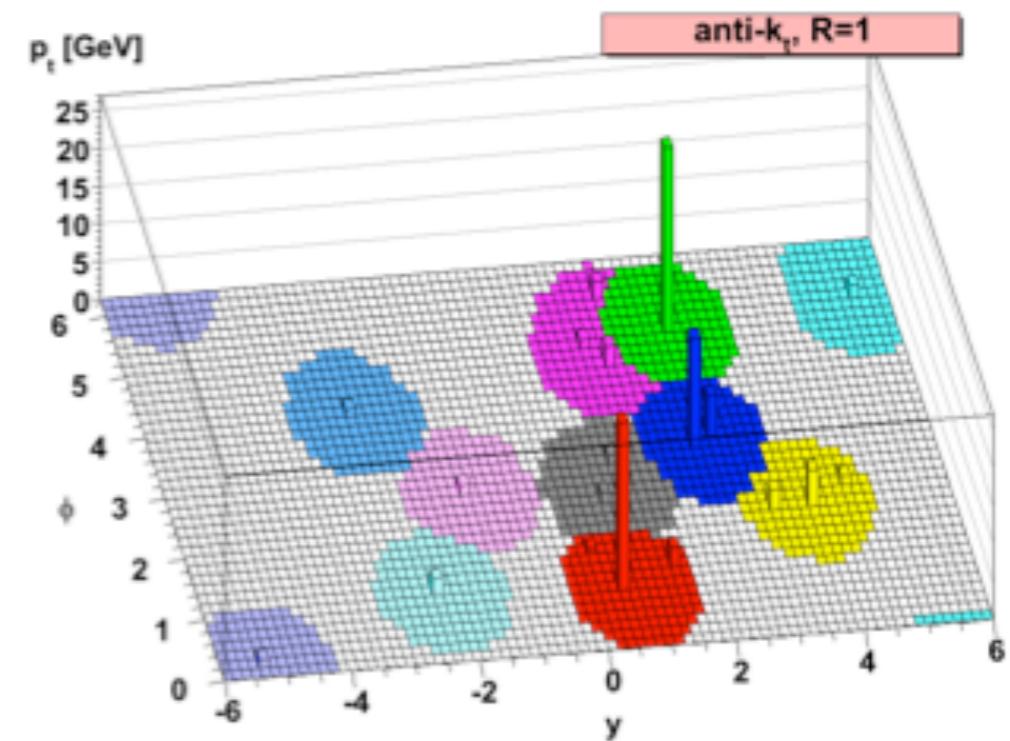
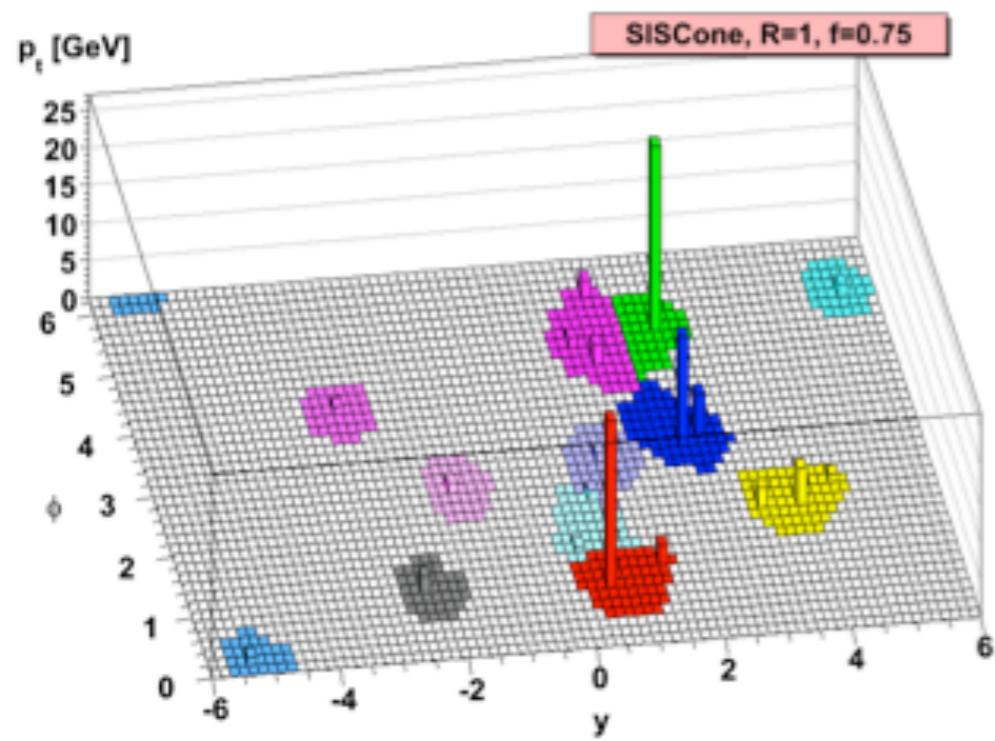
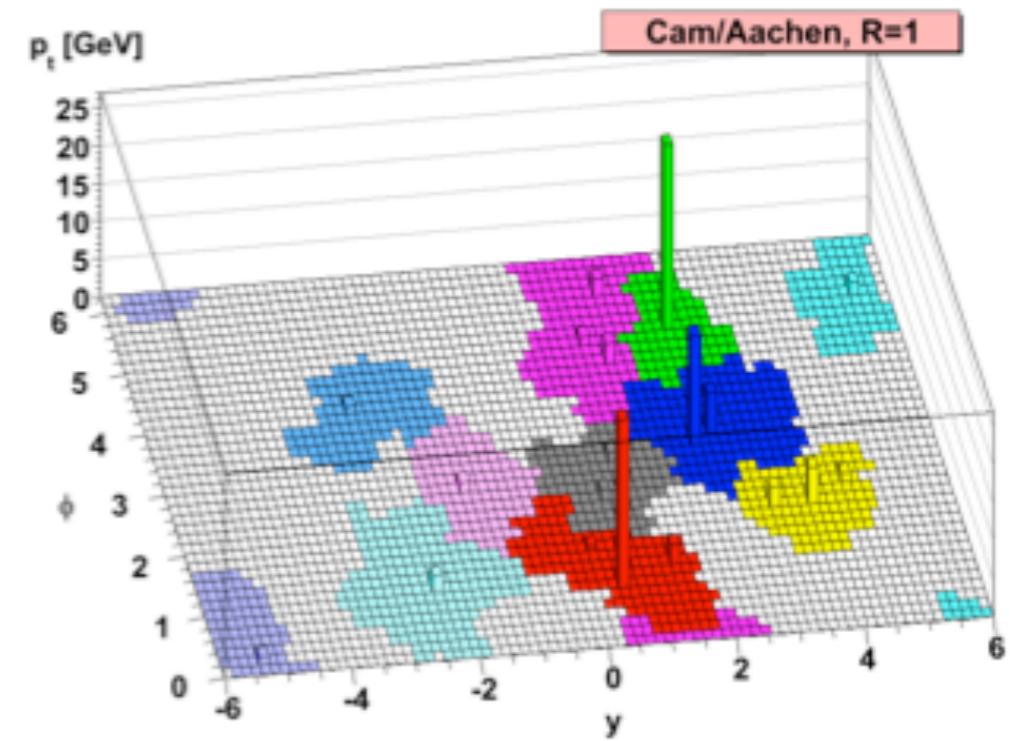
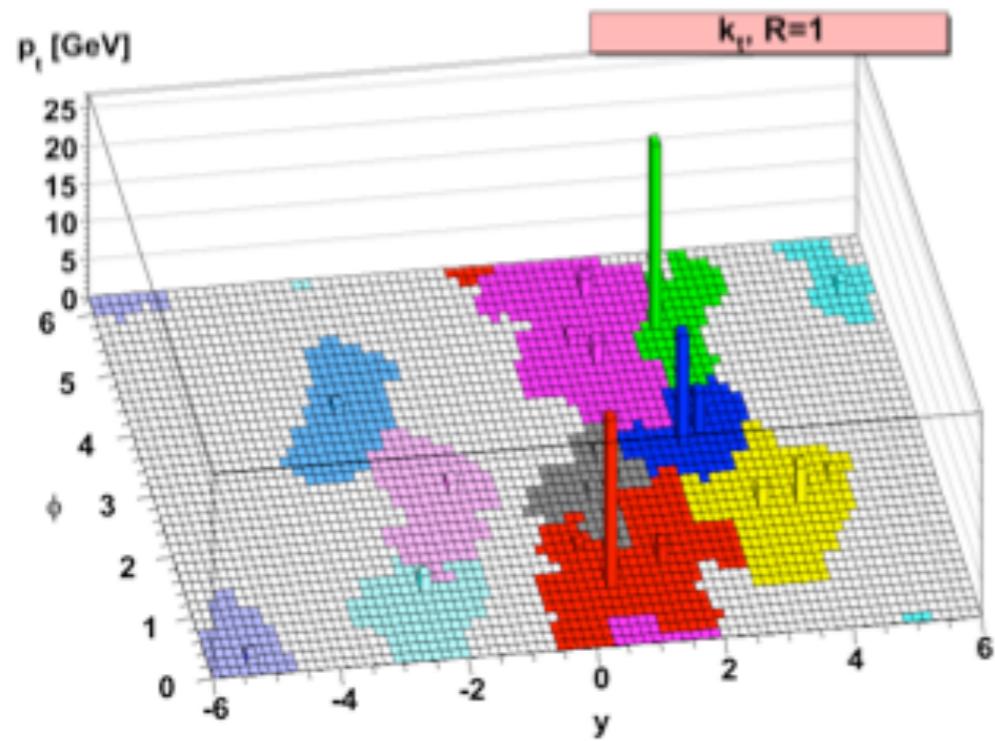
- Also find the “beam distance”

$$d_{iB} = k_{T,i}^n$$

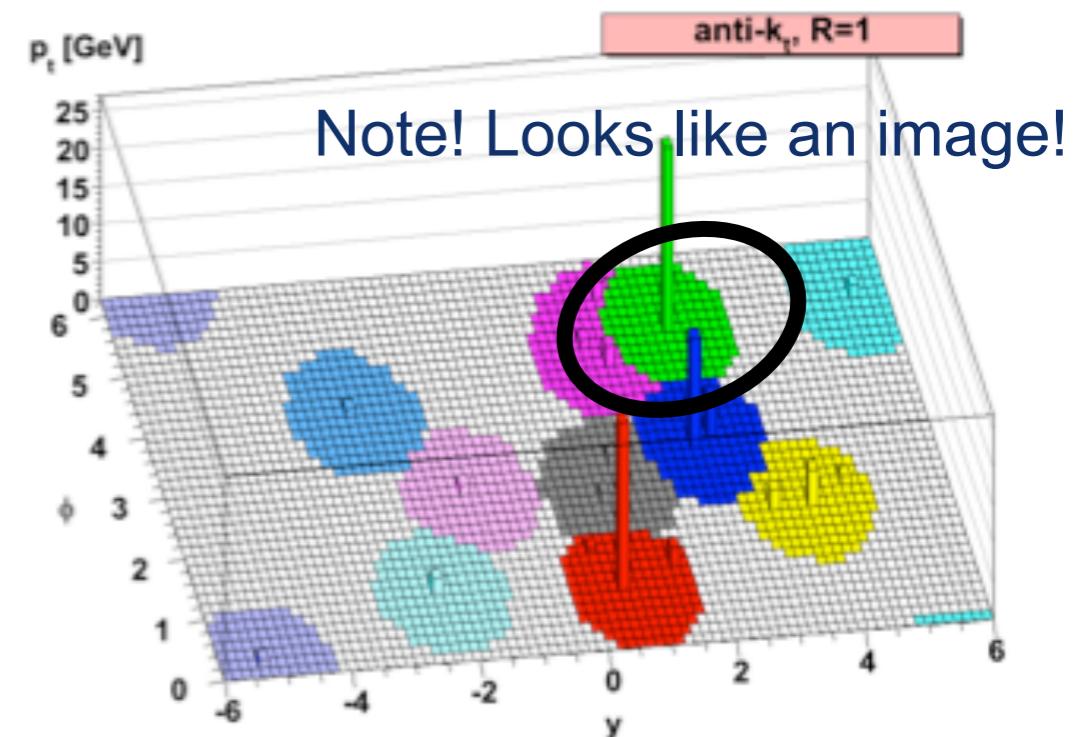
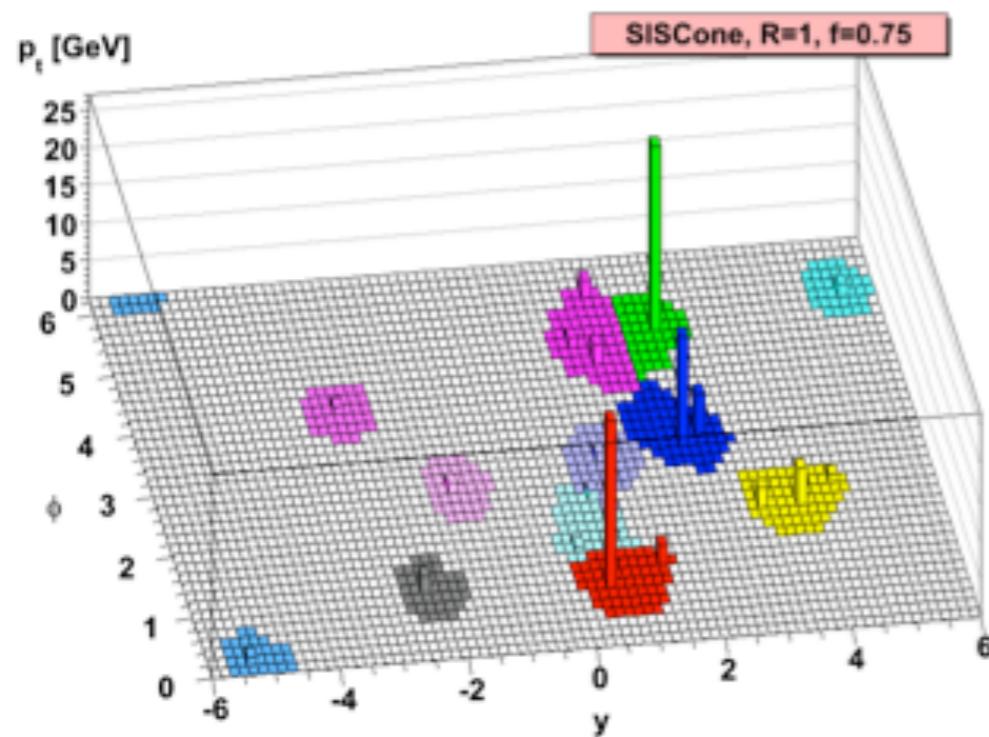
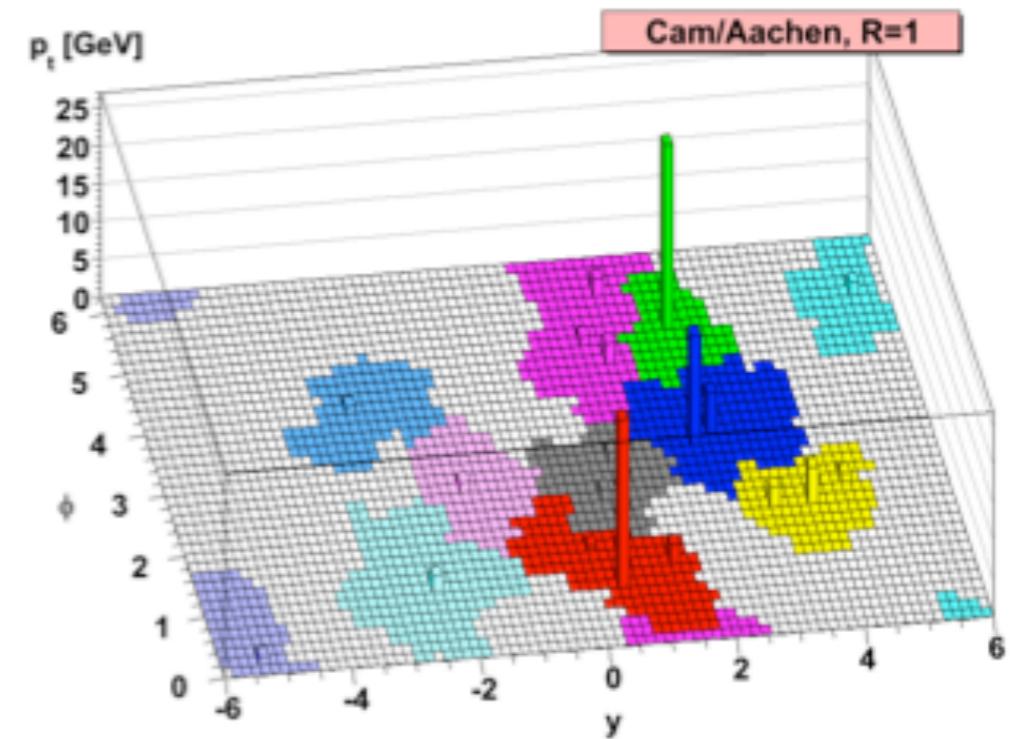
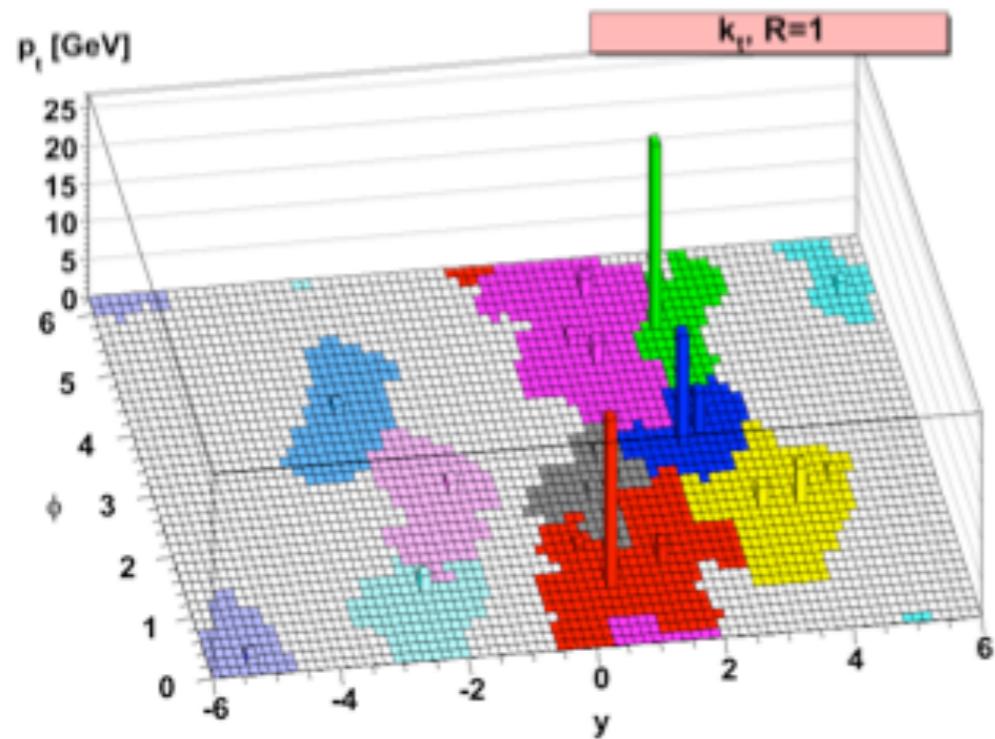
- Find min of all  $d_{ij}$  and  $d_{iB}$ 
  - If min is a  $d_{ij}$ , merge and iterate
  - If min is a  $d_{iB}$ , classify as a final jet
- Continue until list is exhausted



# Clustering

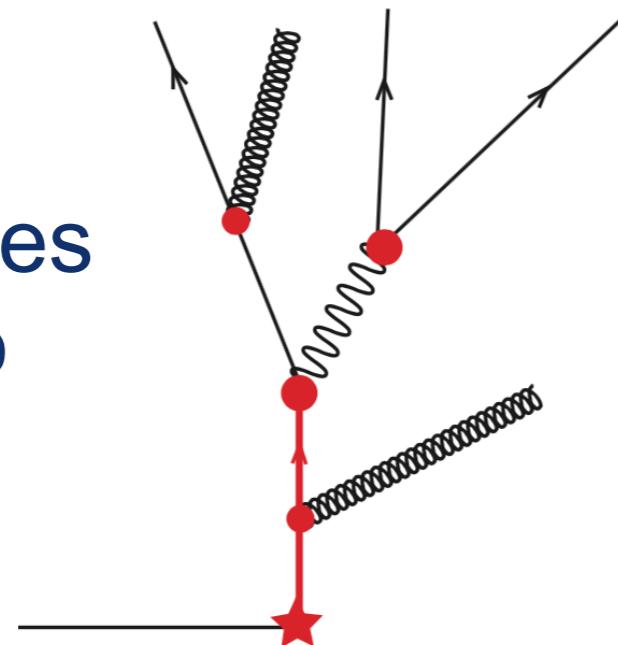


# Clustering

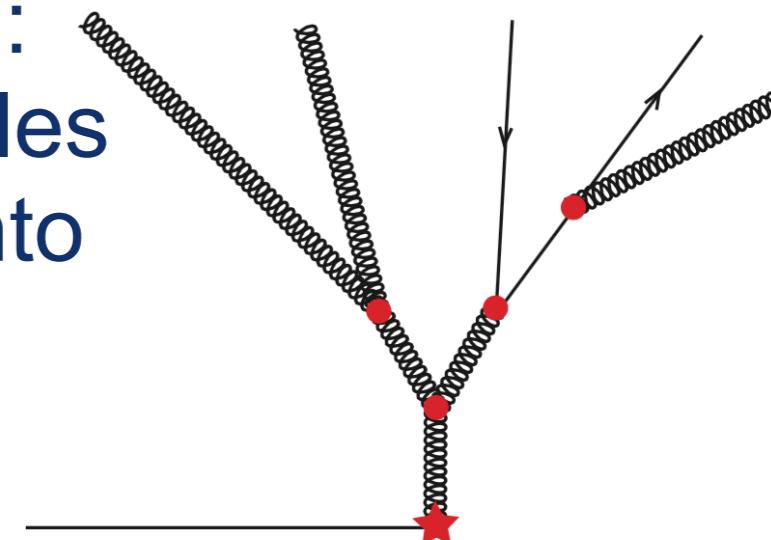


# Classification : “High Level” / “Expert” Information

Category A:  
Heavy particles  
decaying into  
jets



Category B:  
Light particles  
decaying into  
jets



Mass

Driven by particle mass  
(W: 80 GeV, Z: 90 GeV,  
H: 125 GeV, t: 173 GeV)

Particle content

Often decay to b quarks

Shape

Normally 2- or 3-pronged

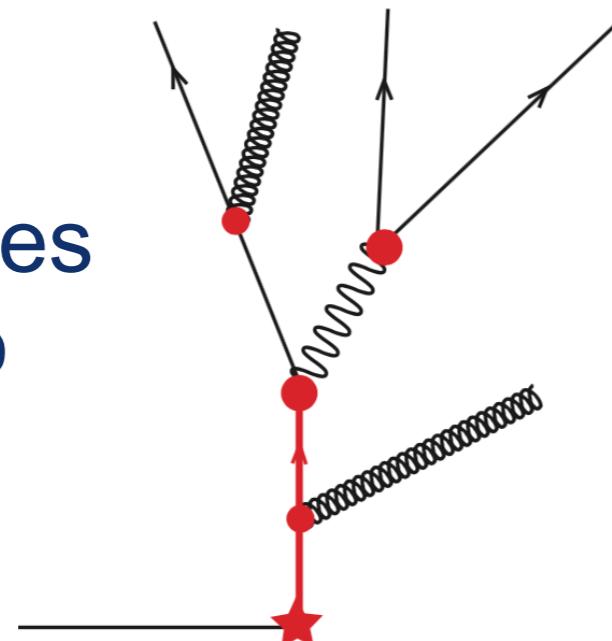
Driven by QCD radiation

Sometimes decay to b quarks

Normally 1-pronged

# Classification : “High Level” / “Expert” Information

Category A:  
Heavy particles  
decaying into  
jets



Mass

Driven by particle mass  
(W: 80 GeV, Z: 90 GeV,  
H: 125 GeV, t: 173 GeV)

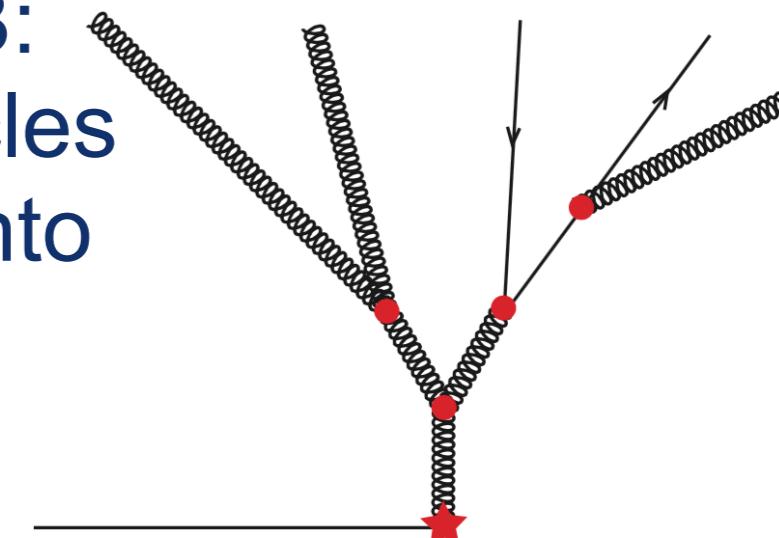
Particle content

Often decay to b quarks

Shape

Normally 2- or 3-pronged

Category B:  
Light particles  
decaying into  
jets



Classify  
by “grooming”  
soft radiation

Classify  
by displaced  
particles

Driven by QCD radiation

Sometimes decay to b quarks

Classify  
by shape

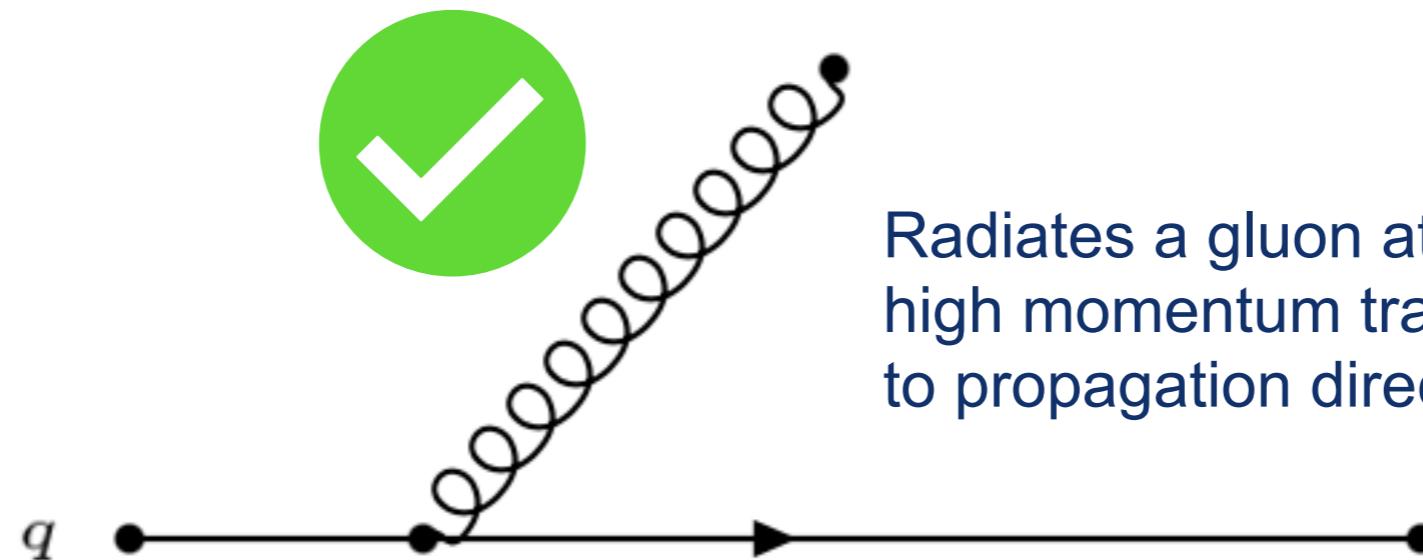
Normally 1-pronged

# Grooming soft radiation

Propagating quark through space



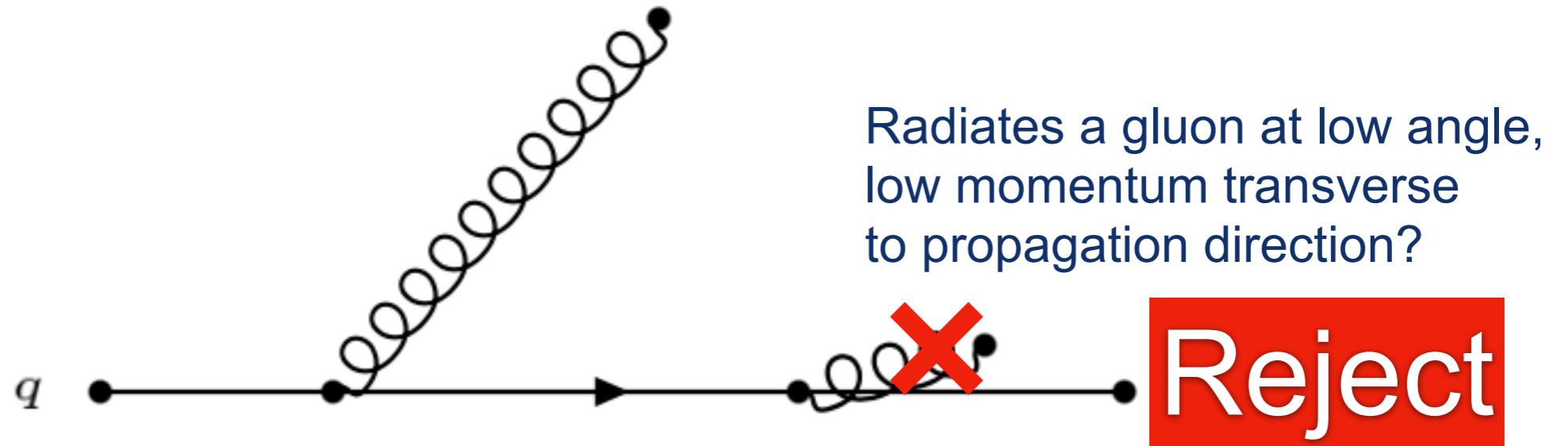
# Grooming soft radiation



Radiates a gluon at wide angle,  
high momentum transverse  
to propagation direction?

OK

# Grooming soft radiation



# Grooming soft radiation

