# One Lecture on Jets

## From Physics to Algorithms

Jesse Thaler

8.811 — October 2, 2019

# Outline

Jets in the Big Picture (cartoons)

Dynamics of Jet Formation (more cartoons)

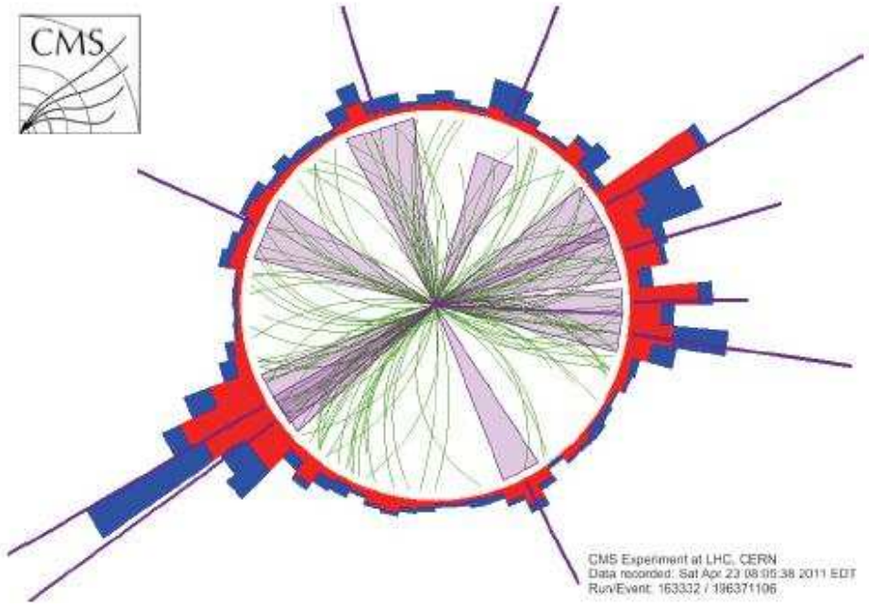
The Shape of Jets (algorithms)

Parting Thoughts (inspiration)

# What is a Jet?

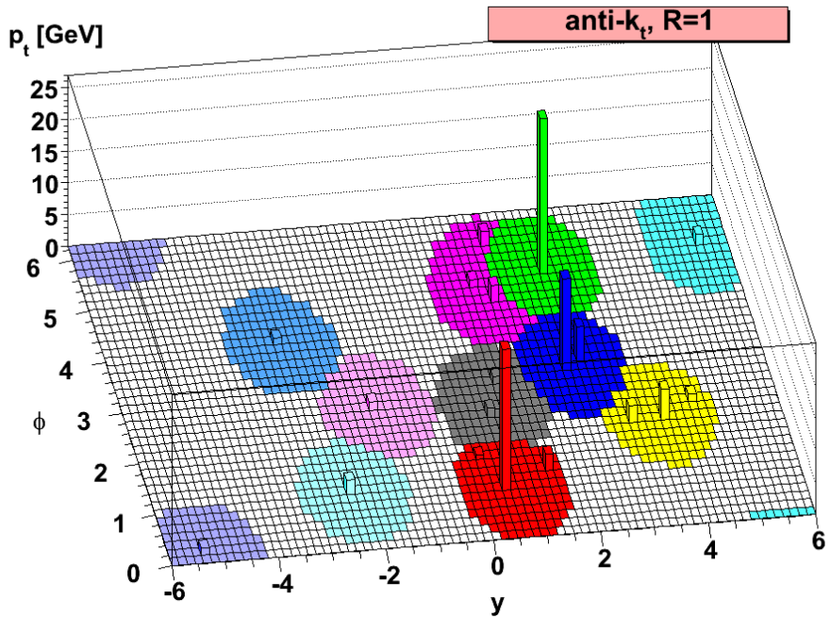
## A physical phenomena:

*Emergent feature of confining gauge theories*




## An analysis technique:

*Method to interpret hadronic final states*

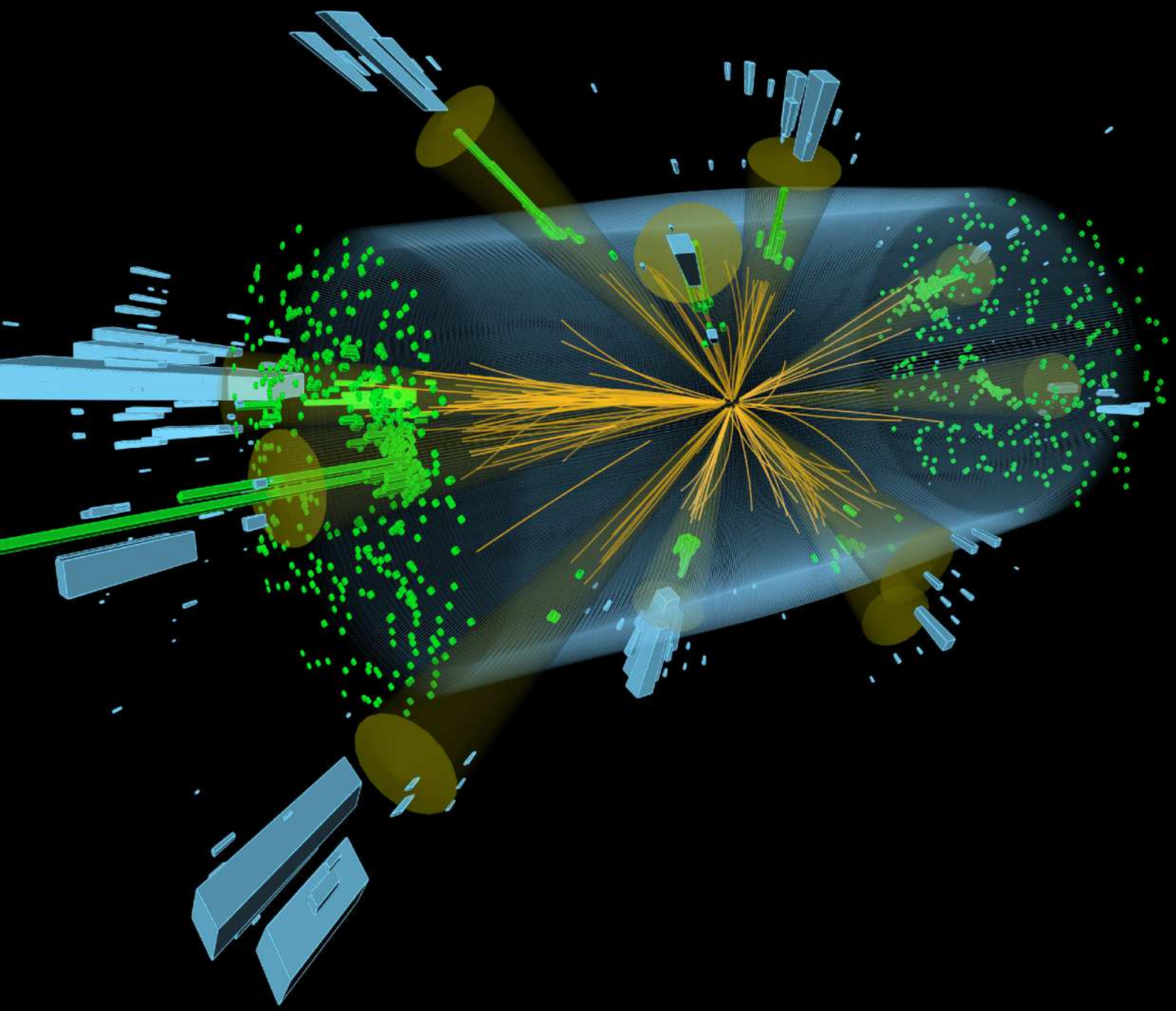



*Freedom to use different analysis strategies for different physical questions*

# *Jets in the Big Picture*

# *Collider Event*

Collection of points in (momentum) space



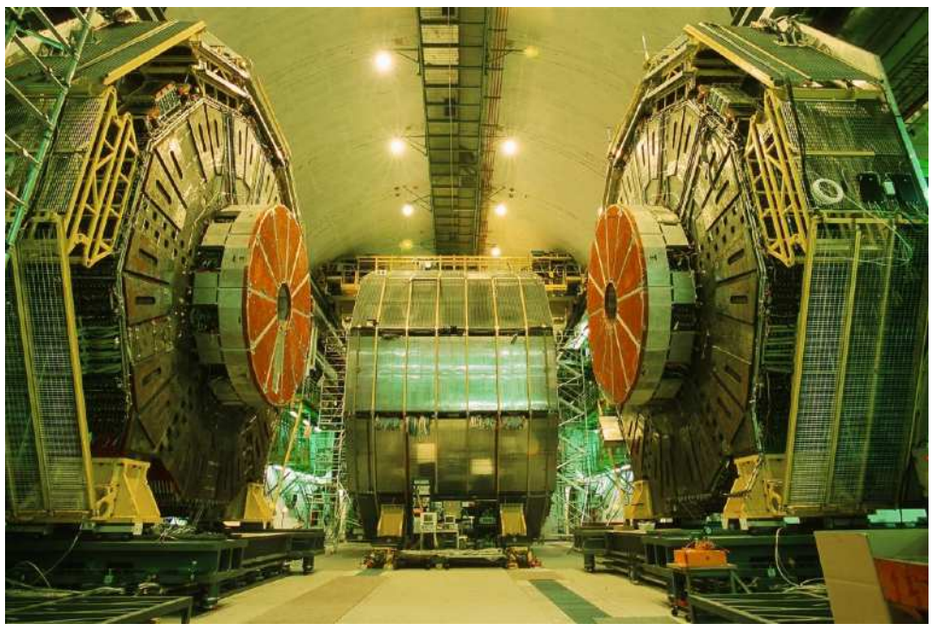
T E H M

| | | |
|---|---|---|
| $\gamma$ | photon | elementary |
| $e^{\pm}$ | electron | elementary |
| $\mu^{\pm}$ | muon | elementary |
| $\pi^{\pm}$ | pion | composite |
| $K^{\pm}$ | kaon | composite |
| $K^0_L$ | K-long | composite |
| $p/\overline{p}$ | proton | composite |
| $n/\overline{n}$ | neutron | composite |

WIRED

JOSHUA BATSON SCIENCE 01.23.15 6:45 AM

# HOW THREE GUYS WITH $10K AND DECADES-OLD DATA ALMOST FOUND THE HIGGS BOSON FIRST

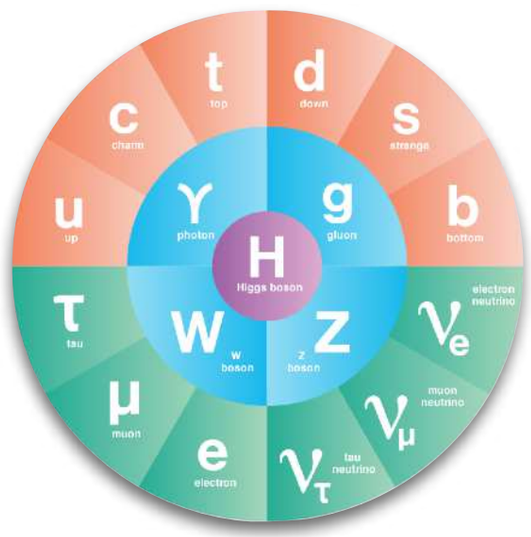


**“Figuring out what happened in a collider is like trying to figure out what your dog ate at the park yesterday. You can find out, but you have to sort through a lot of shit to do it.”**

The Universe
Galaxy
Solar System
Planet
Us
Hairs
Cells
Molecules
Electrons
Protons
The TeV Scale
Grand Unification
Quantum Gravity

$10^{26}$ m $10^{11}$ m $10^{-4}$ m $10^{-19}$ m $10^{-34}$ m

## The Standard Model



Matter/Anti-Matter?

Three Generations?

Strong CP?

Neutrino Masses?

*Compositeness?*
*Extra Dimensions?*
*Supersymmetry?*
*The Unexpected?*

$10^{26}$ m $10^{11}$ m $10^{-4}$ m $10^{-19}$ m $10^{-34}$ m

The Standard Model → ?

Grand Unification?

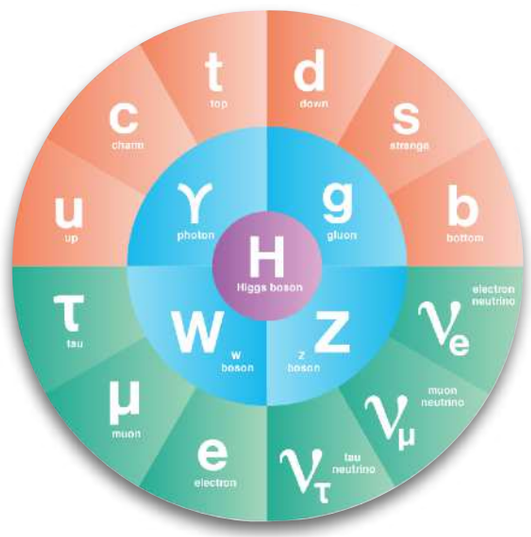
Dark Matter?

Origin of Mass?
Higgs Boson!

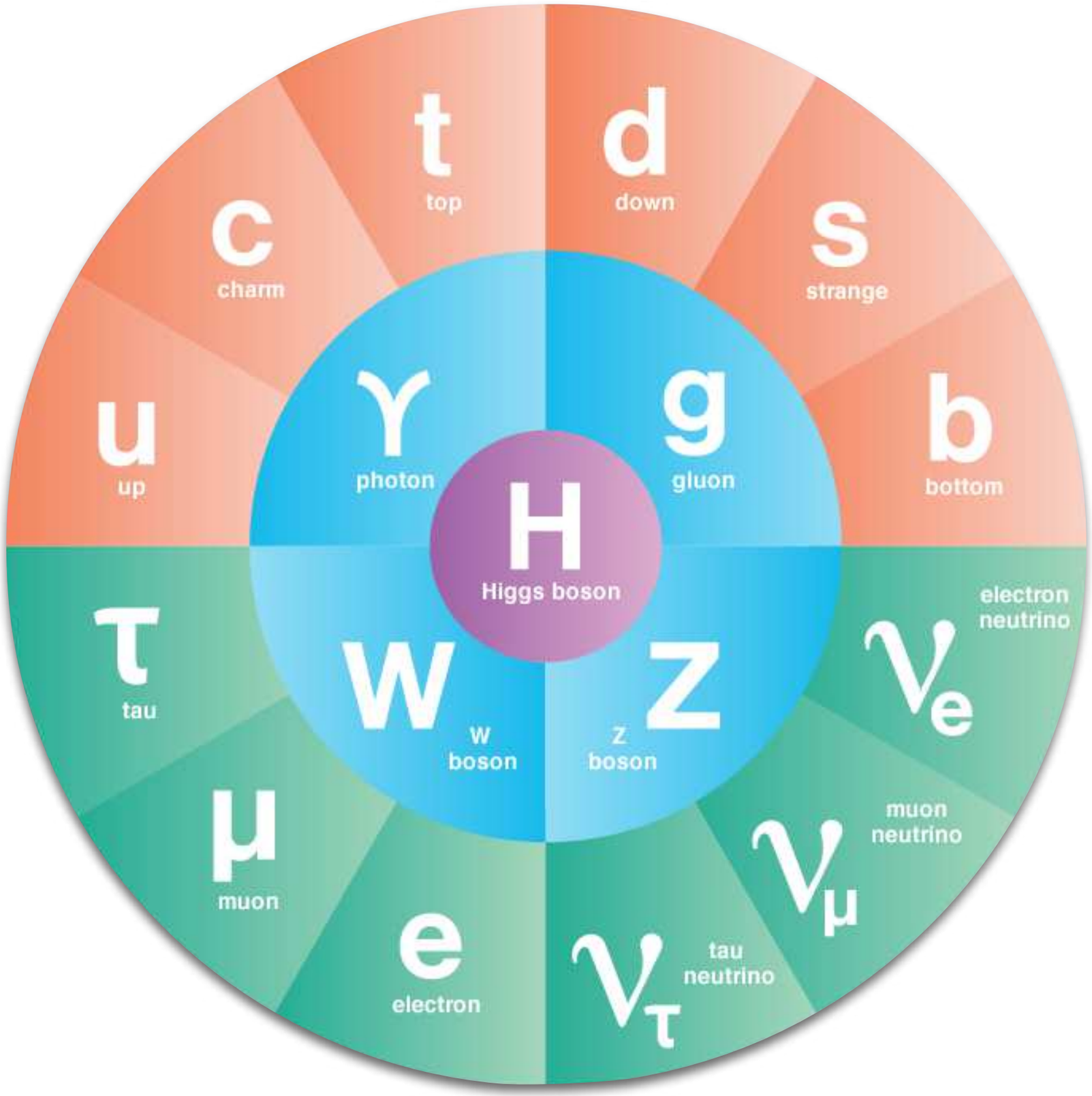
Quantum Gravity?

Dark Energy?

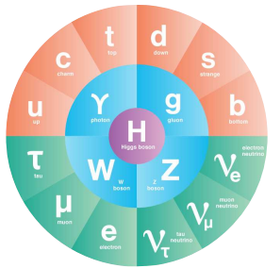
Inflation?




Hierarchy Problem?


t
top
d
down
c
charm
s
strange
u
up
γ
photon
g
gluon
b
bottom
H
Higgs boson
τ
tau
W
W boson
Z
Z boson
electron neutrino
μ
muon
muon neutrino
e
electron
tau neutrino

# Particles as Probes




## Not Jets

γ

$\nu_1$

$\nu_2$

$\nu_3$

e

μ

τ

## Jets

g

u

d

s

c

b

## Multi-Jets (or Not)

W

Z

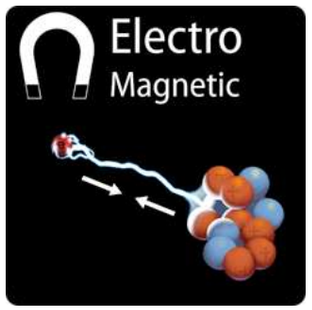
H

t

# Particles and Fundamental Forces


massless spin-2 graviton

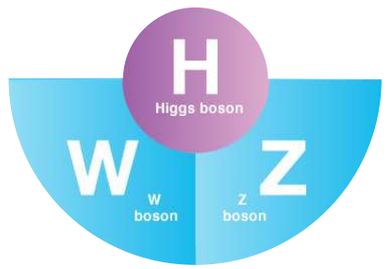

$\Rightarrow$ GR $G_N$



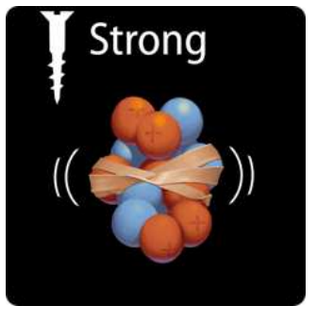

massless spin-1 photon


$\Rightarrow$ QED $\alpha_{EM}$


3 massive spin-1 W/Z bosons
+ massive spin-0 Higgs boson




$\Rightarrow$ Puzzles *16+*




8 massless spin-1 gluons


$\Rightarrow$ QCD $\alpha_S$

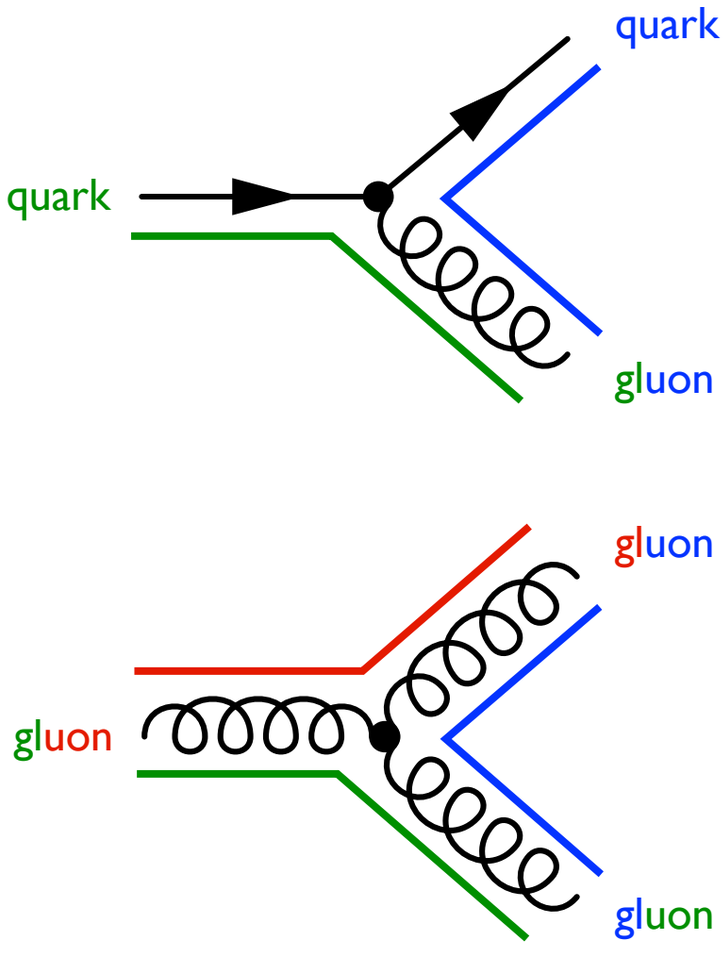
*plus* $\theta_{QCD} \approx 0$

$$\mathcal{L} = \frac{1}{4g^2} G^a_{\mu\nu} G^a_{\mu\nu} + \sum_j \bar{q}_j (i\gamma^\mu D_\mu + m_j) q_j$$

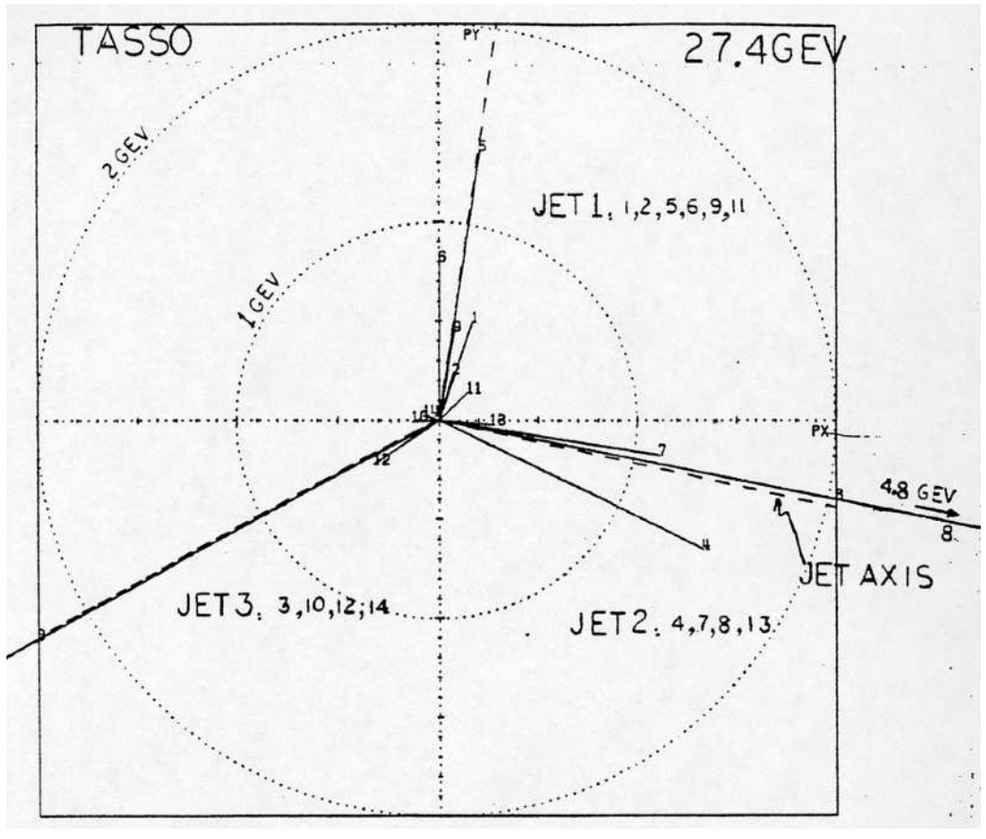
where $G^a_{\mu\nu} \equiv \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + i f_{bc}{}^a A^b_\mu A^c_\nu$

and $D_\mu \equiv \partial_\mu + i t^a A^a_\mu$

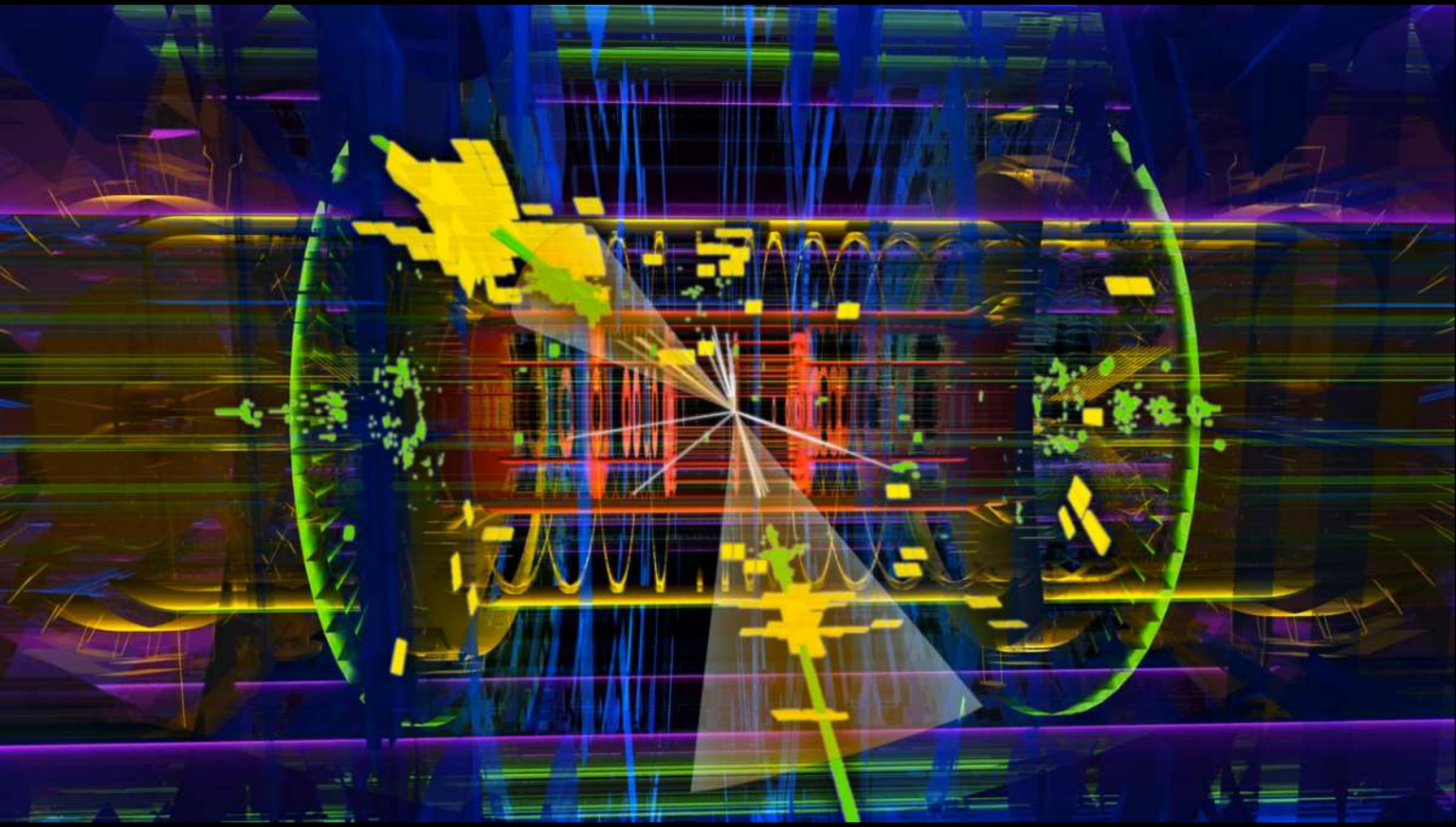
That's it!

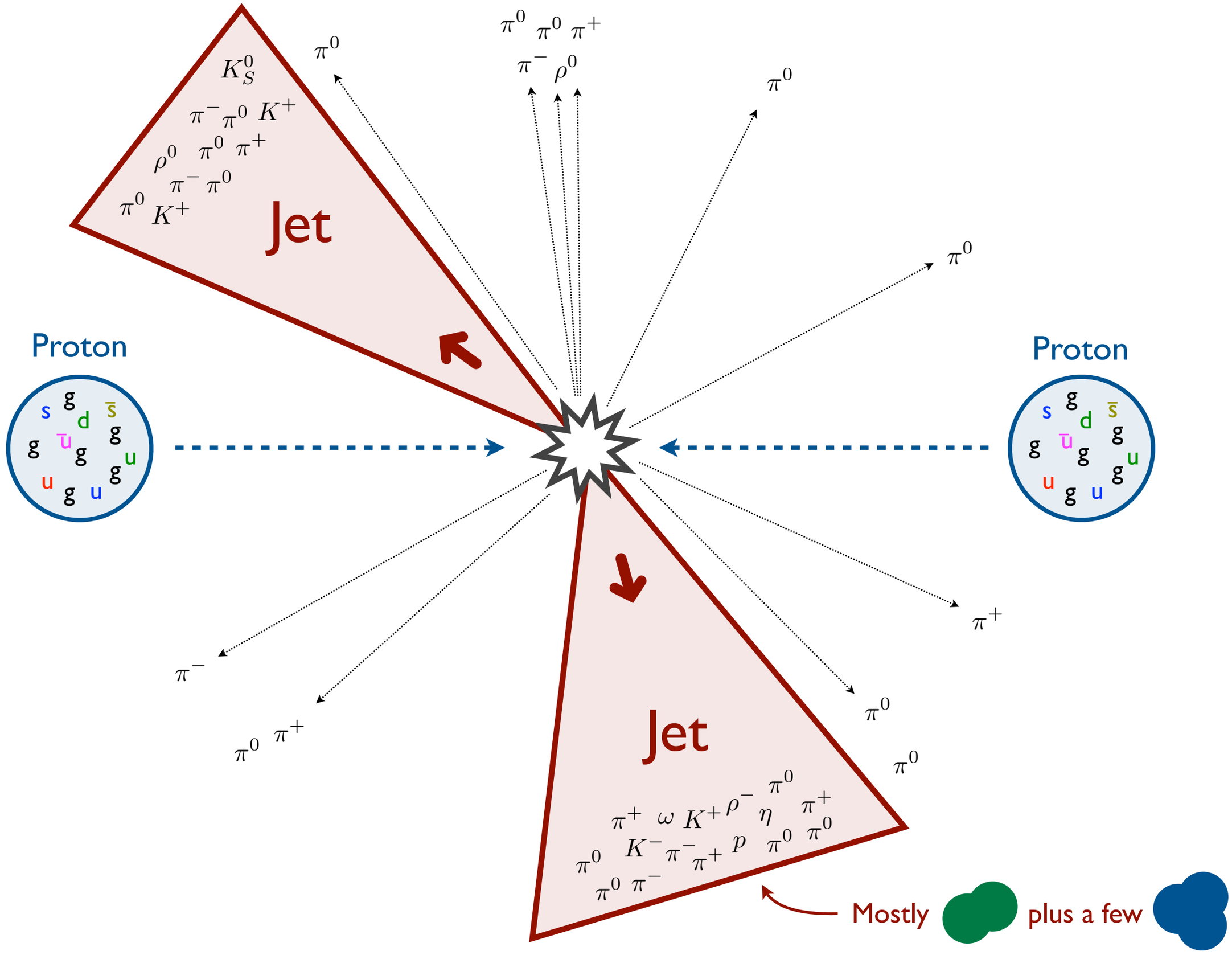



⇒



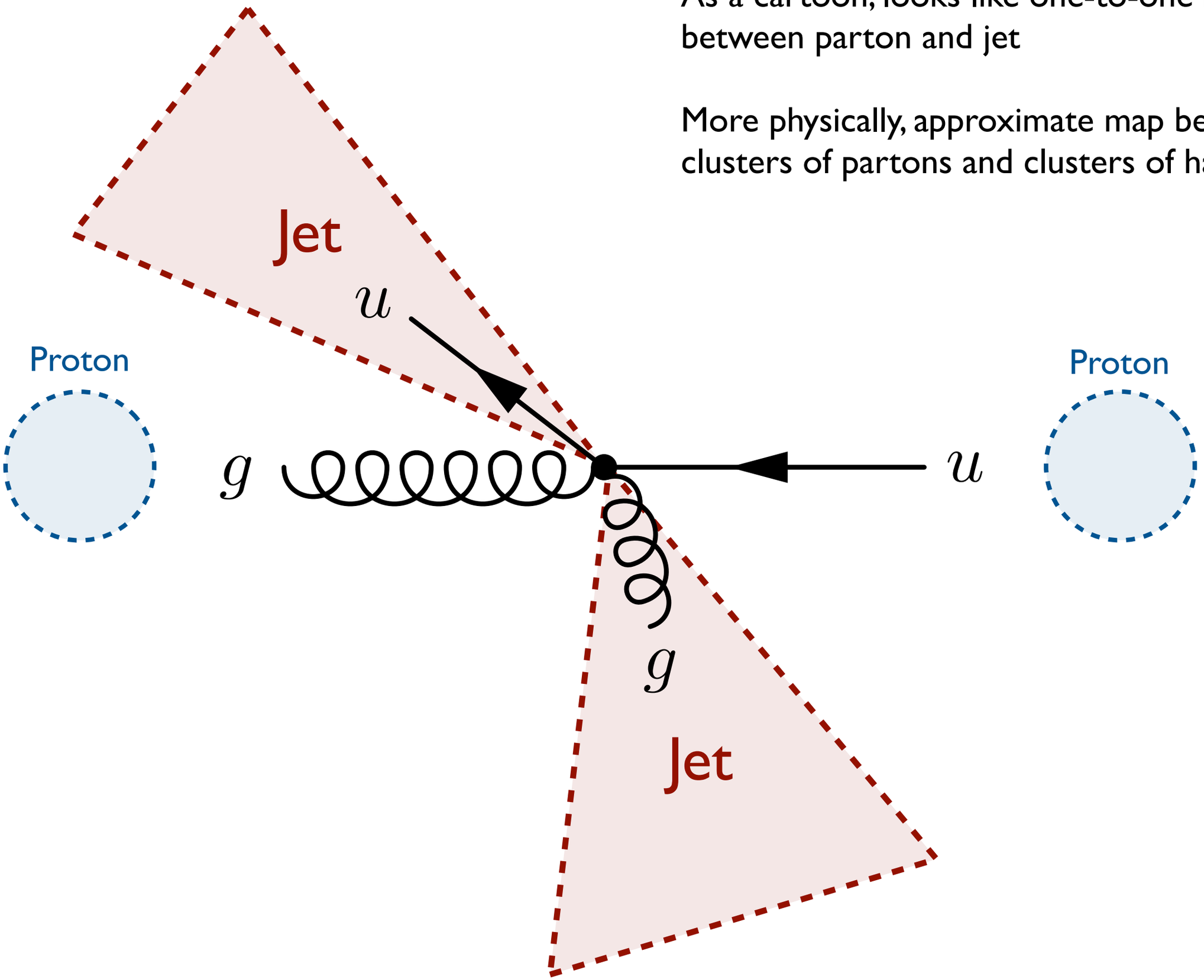

[TASSO @ PETRA @ DESY, 1979]

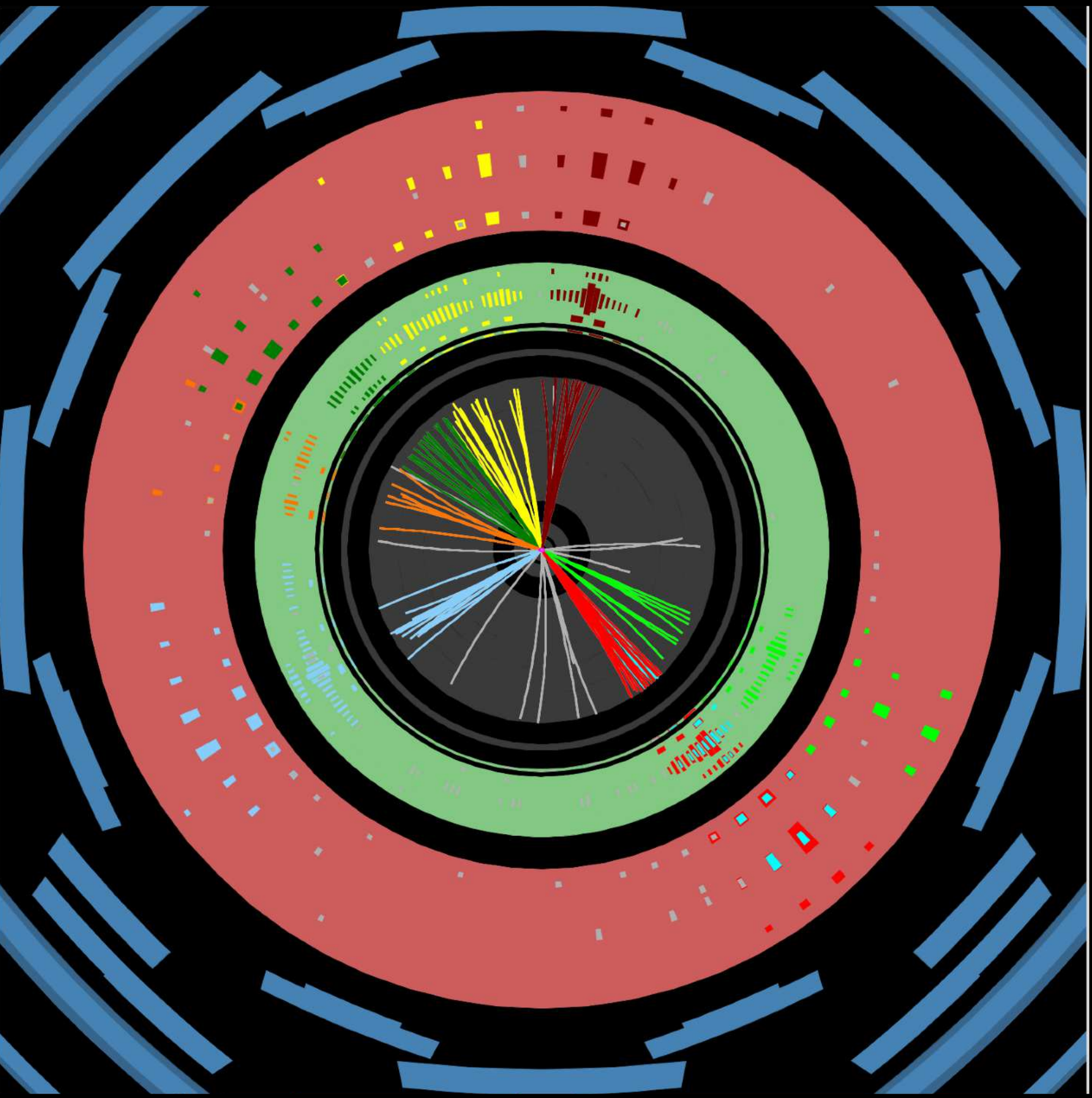


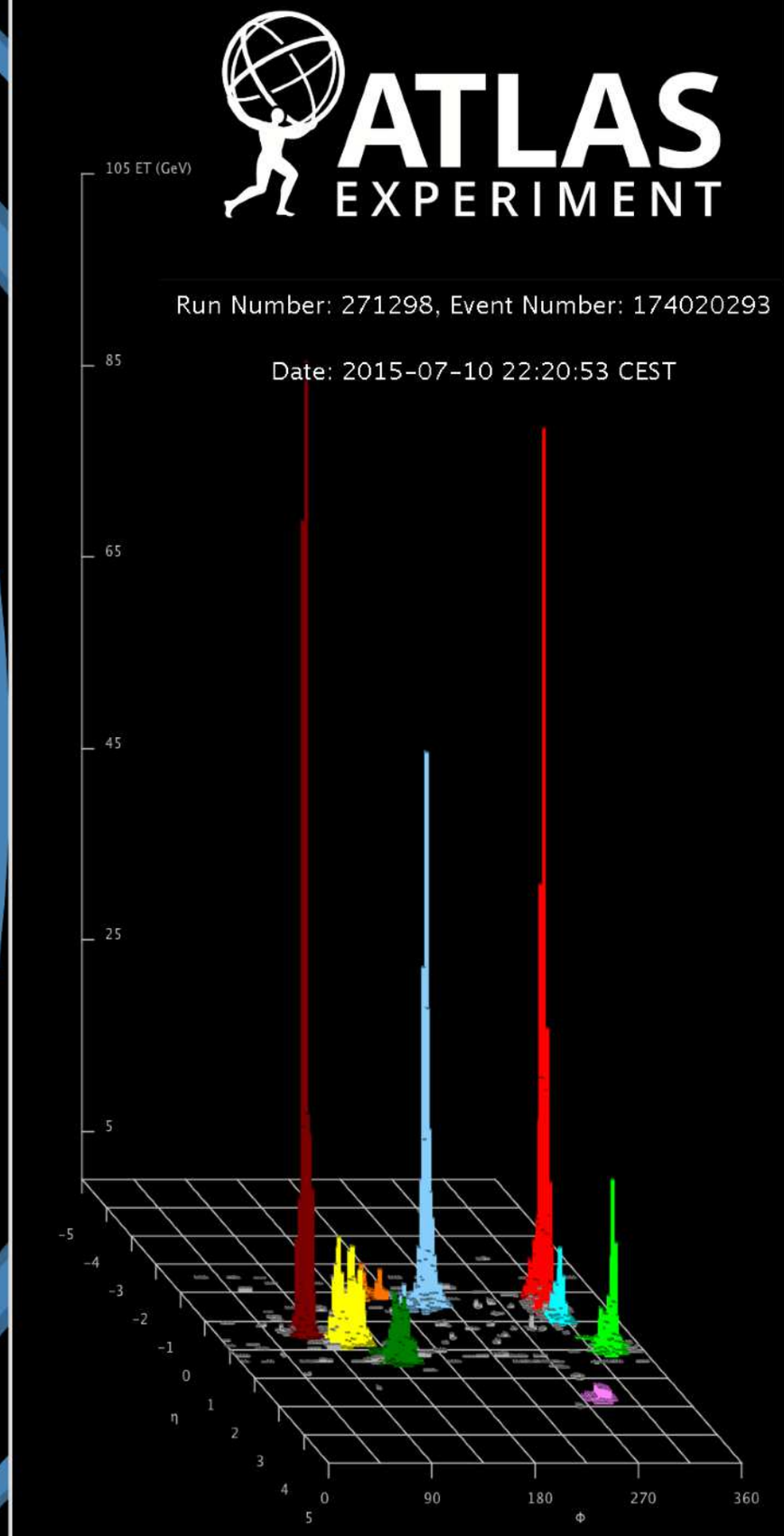

$K_S^0$
$\pi^-\ \pi^0\ K^+$
$\rho^0\ \ \pi^0\ \pi^+$
$\pi^-\ \pi^0$
$\pi^0\ K^+$
Jet
$\pi^0$
$\pi^0\ \pi^0\ \pi^+$
$\pi^-\ \rho^0$
$\pi^0$
$\pi^0$
Proton
Proton
$\pi^-$
$\pi^0\ \pi^+$
$\pi^+$
$\pi^0$
$\pi^0$
$\pi^0$
Jet
$\pi^+\ \omega\ K^+\ \rho^-\ \eta\ \ \pi^+$
$\pi^0\ \ K^-\ \pi^-\ \pi^+\ p\ \ \pi^0\ \pi^0$
$\pi^0\ \pi^-$
Mostly
plus a few

As a cartoon, looks like one-to-one map between parton and jet

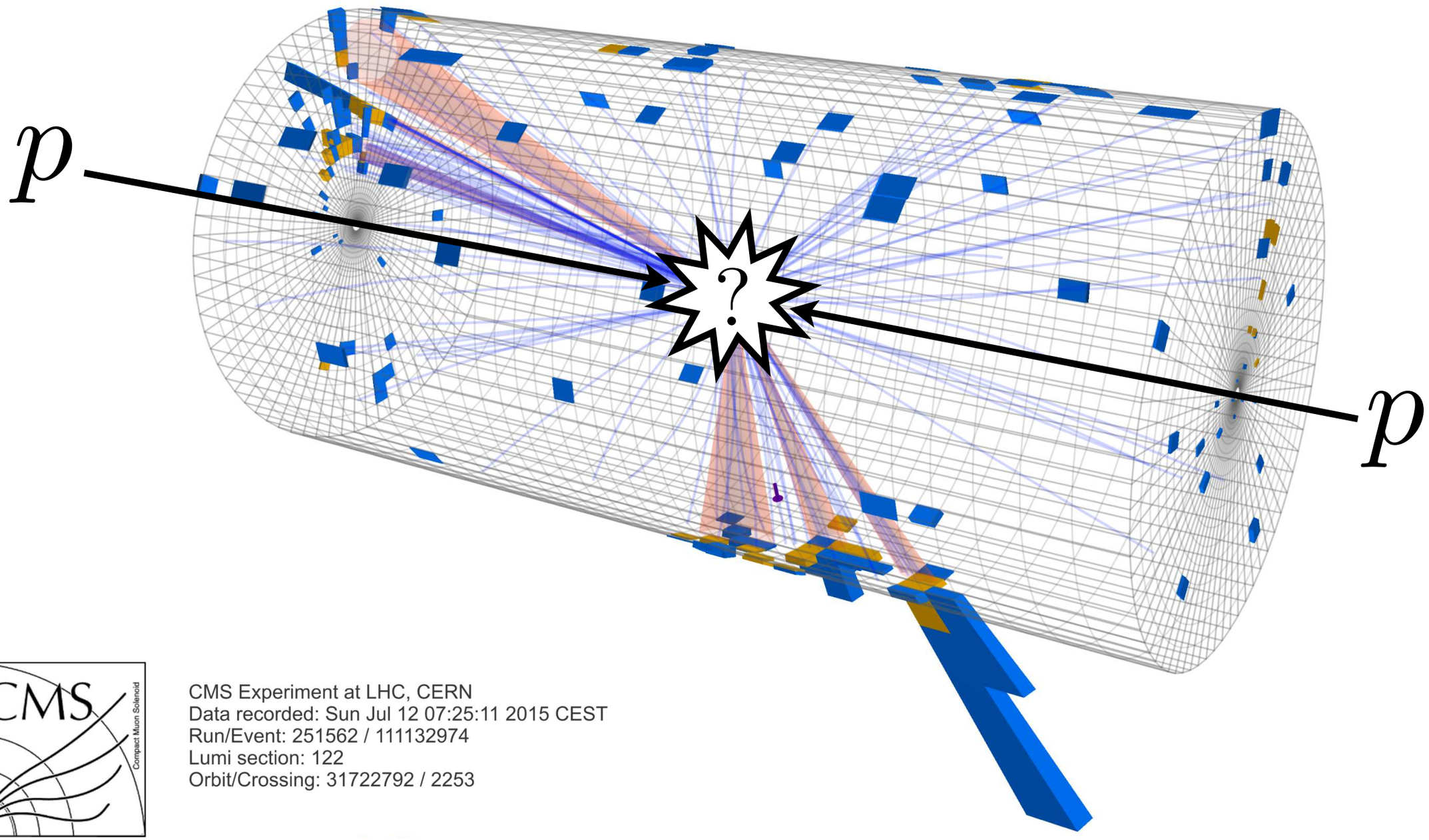
More physically, approximate map between clusters of partons and clusters of hadrons





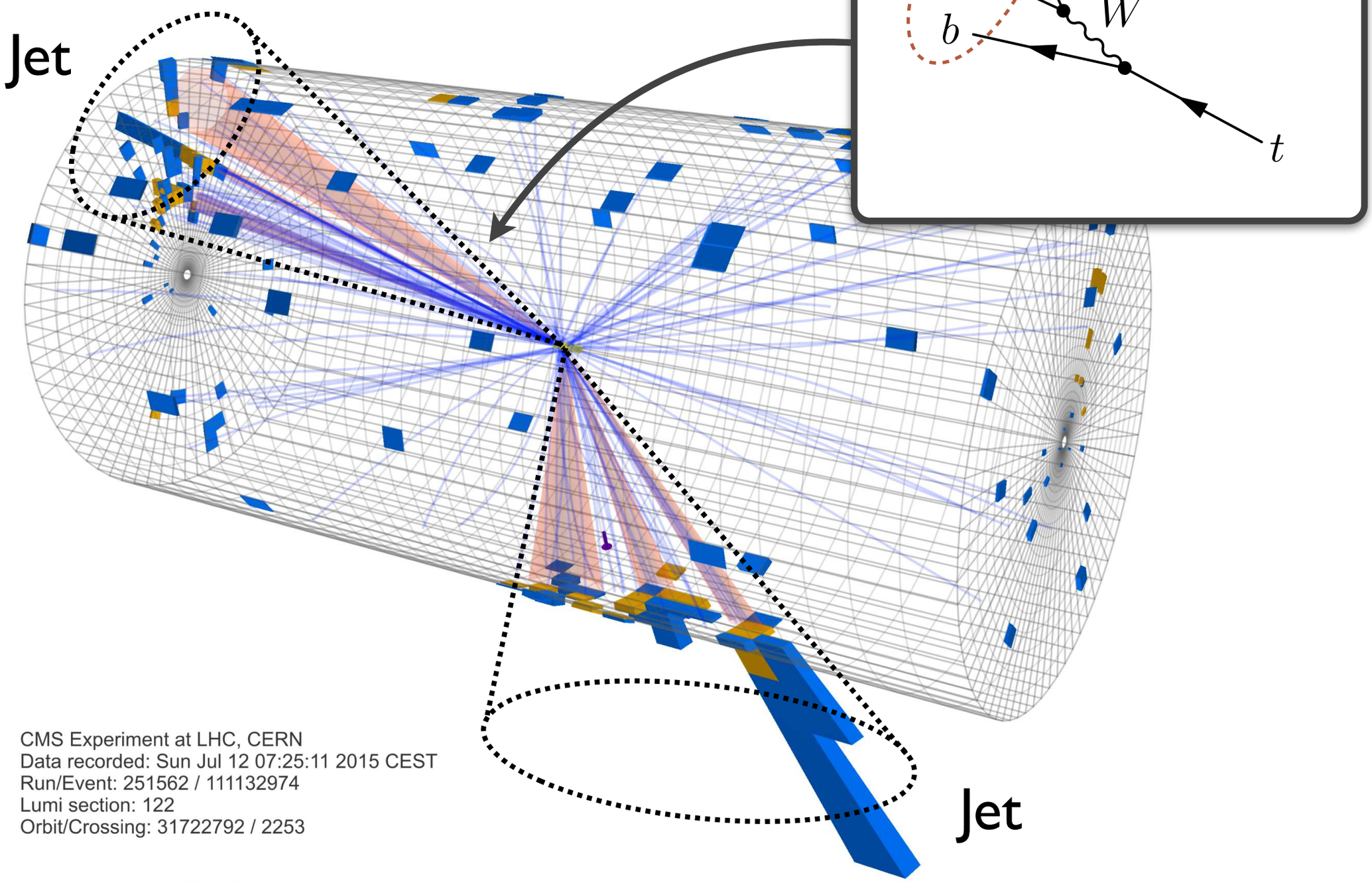

ATLAS
EXPERIMENT
Run Number: 271298, Event Number: 174020293
Date: 2015-07-10 22:20:53 CEST
105 ET (GeV)
85
65
45
25
5
-5
-4
-3
-2
-1
0
1
2
3
4
5
η
0
90
180
270
360
Φ

# The Rise of Jet Substructure




CMS Experiment at LHC, CERN
Data recorded: Sun Jul 12 07:25:11 2015 CEST
Run/Event: 251562 / 111132974
Lumi section: 122
Orbit/Crossing: 31722792 / 2253

# The Rise of Jet Substructure

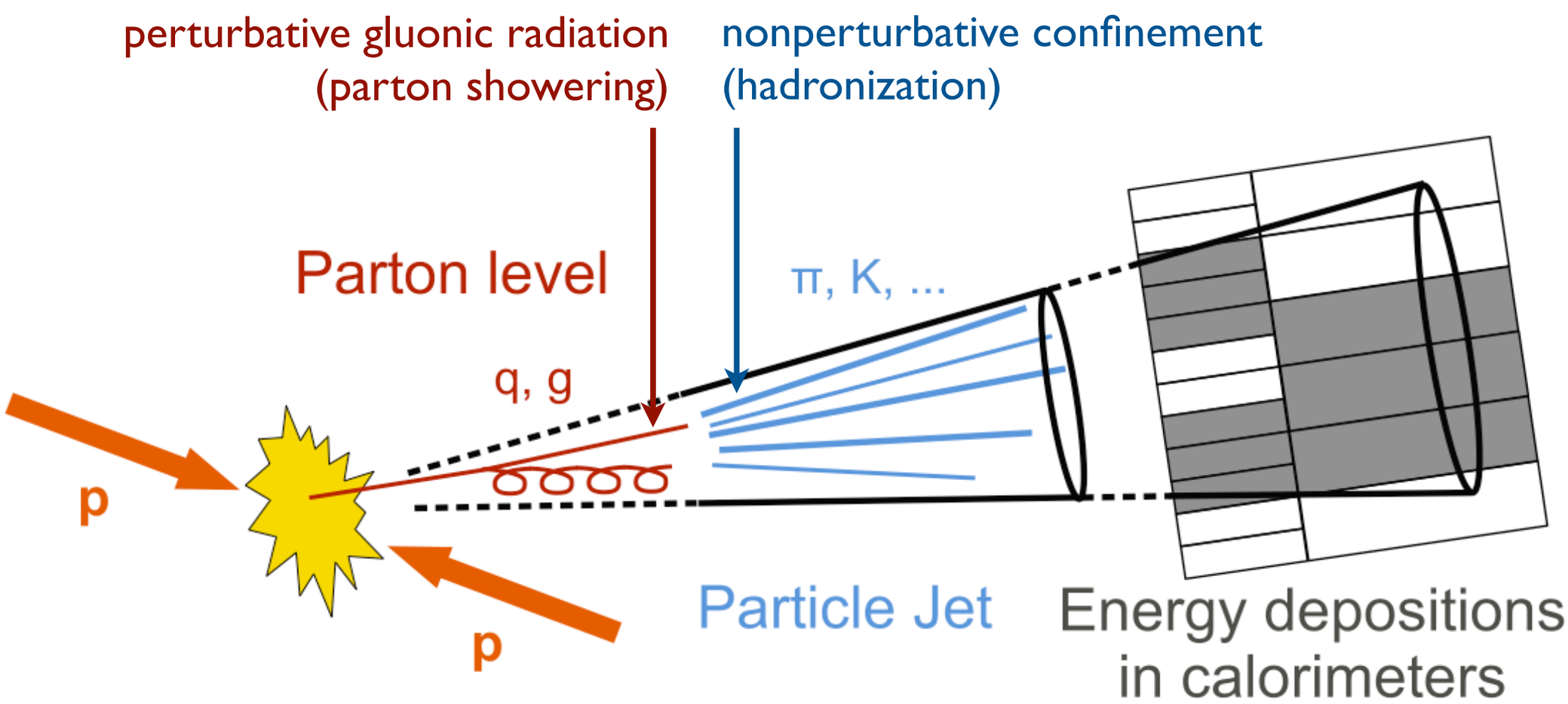


# If you ain't Boostin', you ain't Livin'…


[see reviews in Larkoski, Moult, Nachman, arXiv 2017;
Asquith, Delitzsch, Schmidt, et al., arXiv 2018;
Marzani, Soyez, Spannowsky, LNP 2019]

# *Dynamics of Jet Formation*

# From Partons to Hadrons to Detection

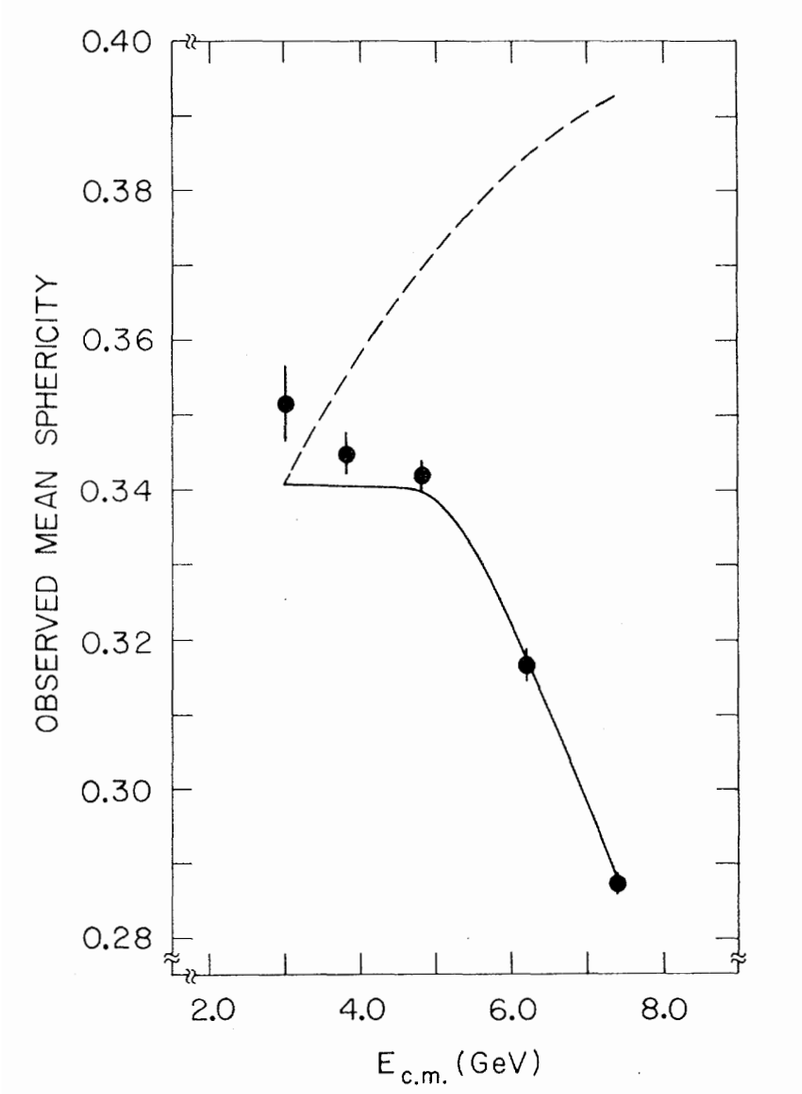



≈ one-to-one @ LHC

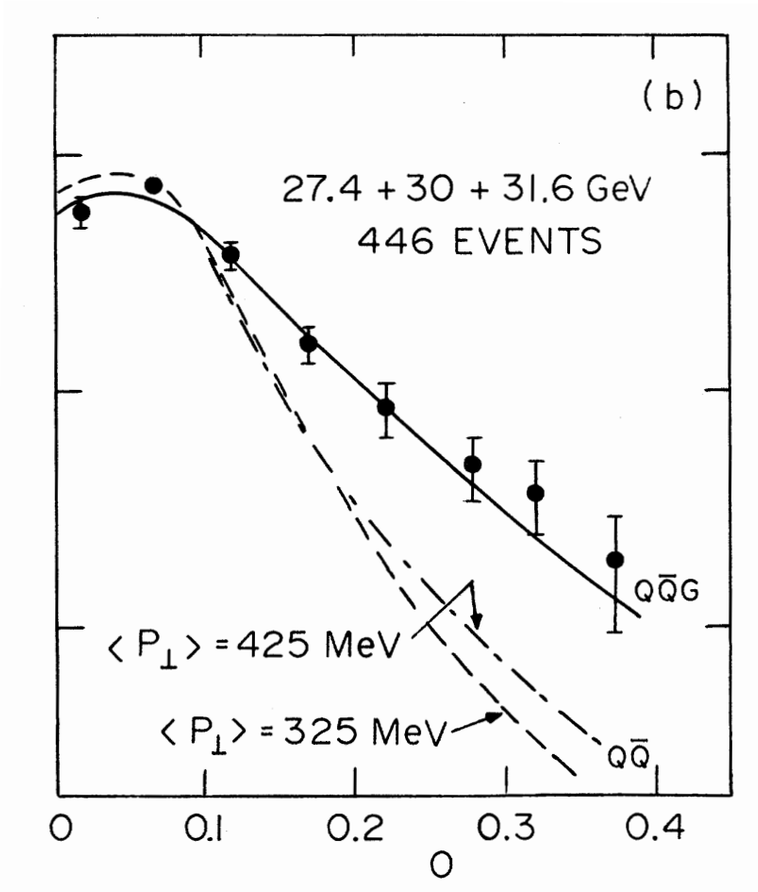
(reality is more complicated, not part of this lecture)

# First Light on Jets

## *Jets @ SPEAR, 1975*

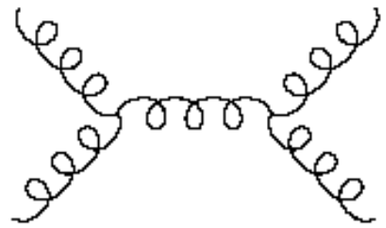


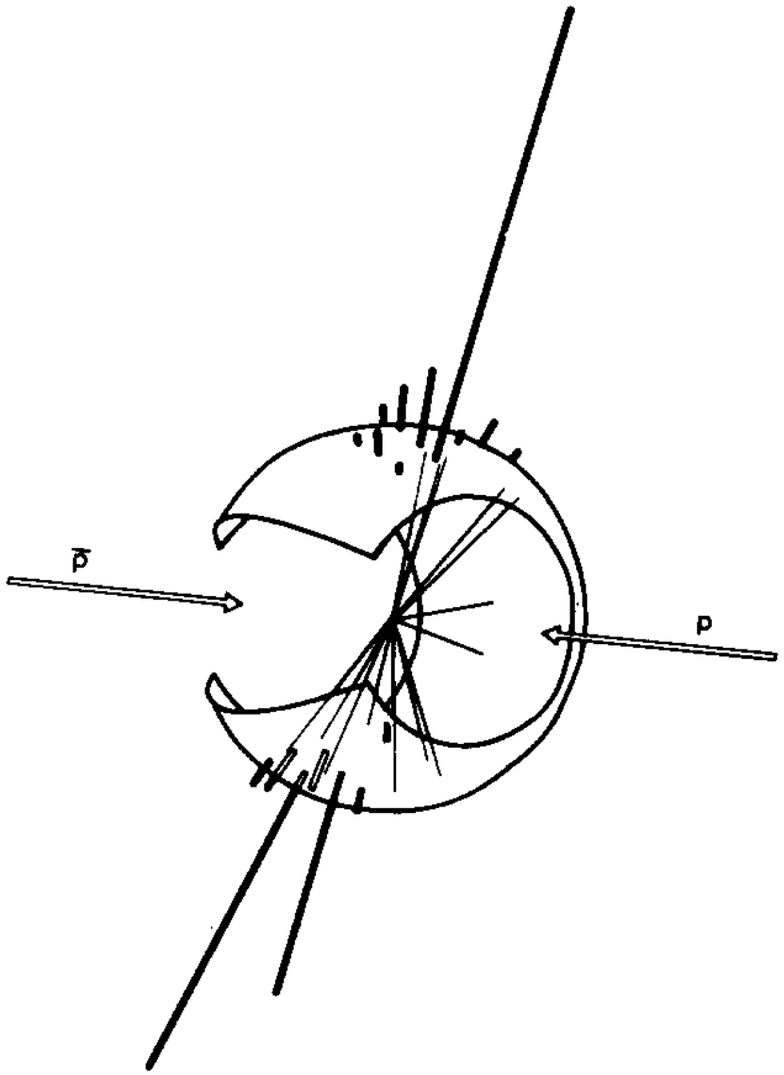

## *Gluons @ PETRA, 1979*

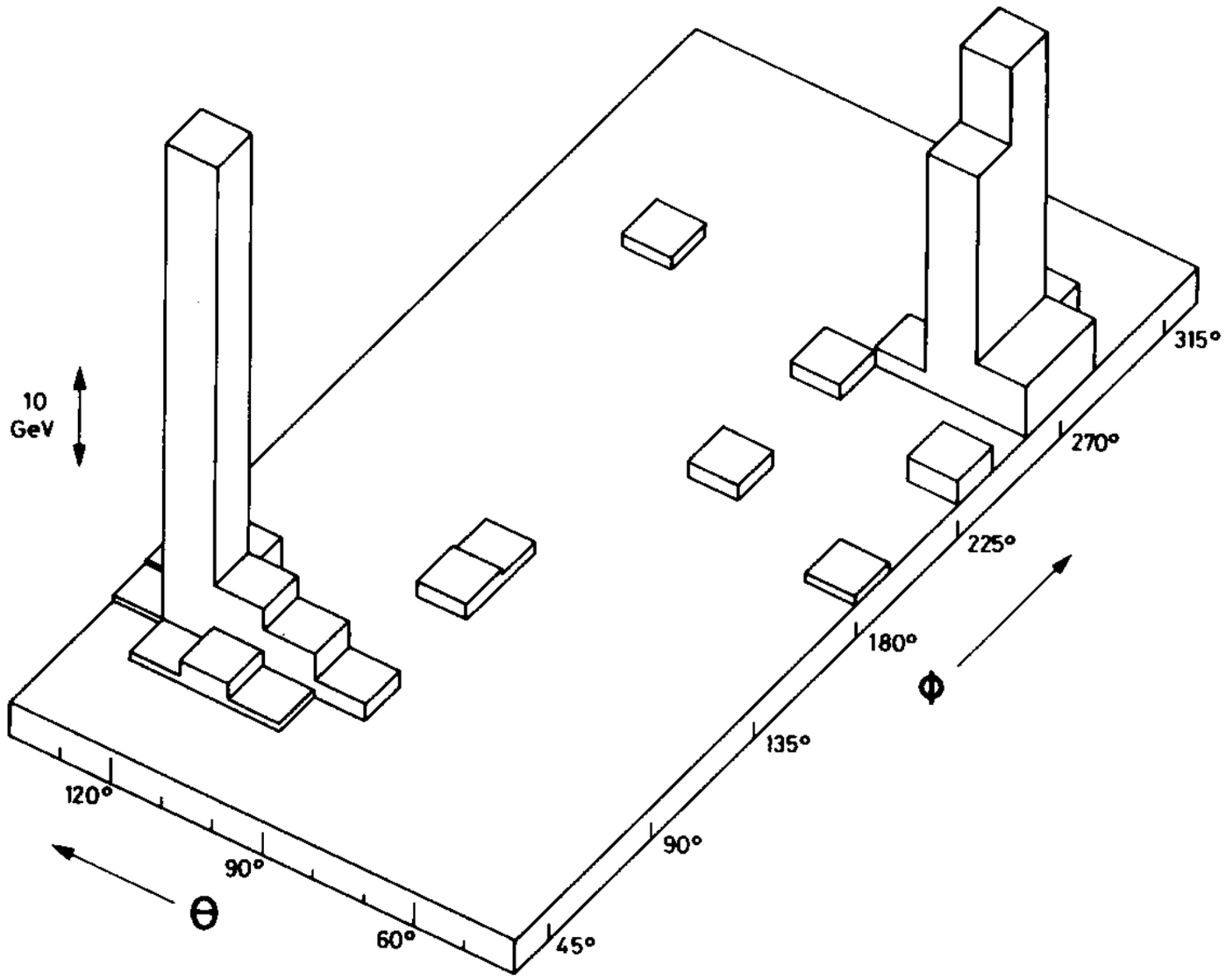


# Four Decades of Jets and QCD

*UA2, 1982*

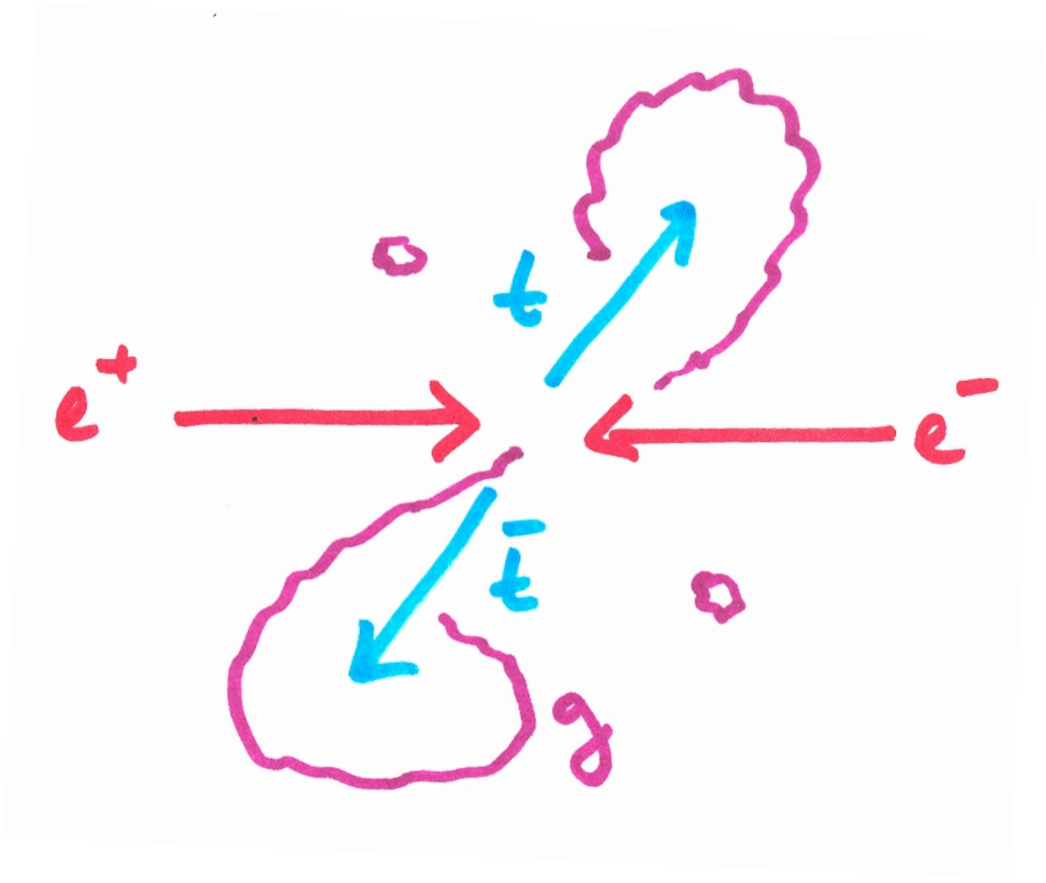




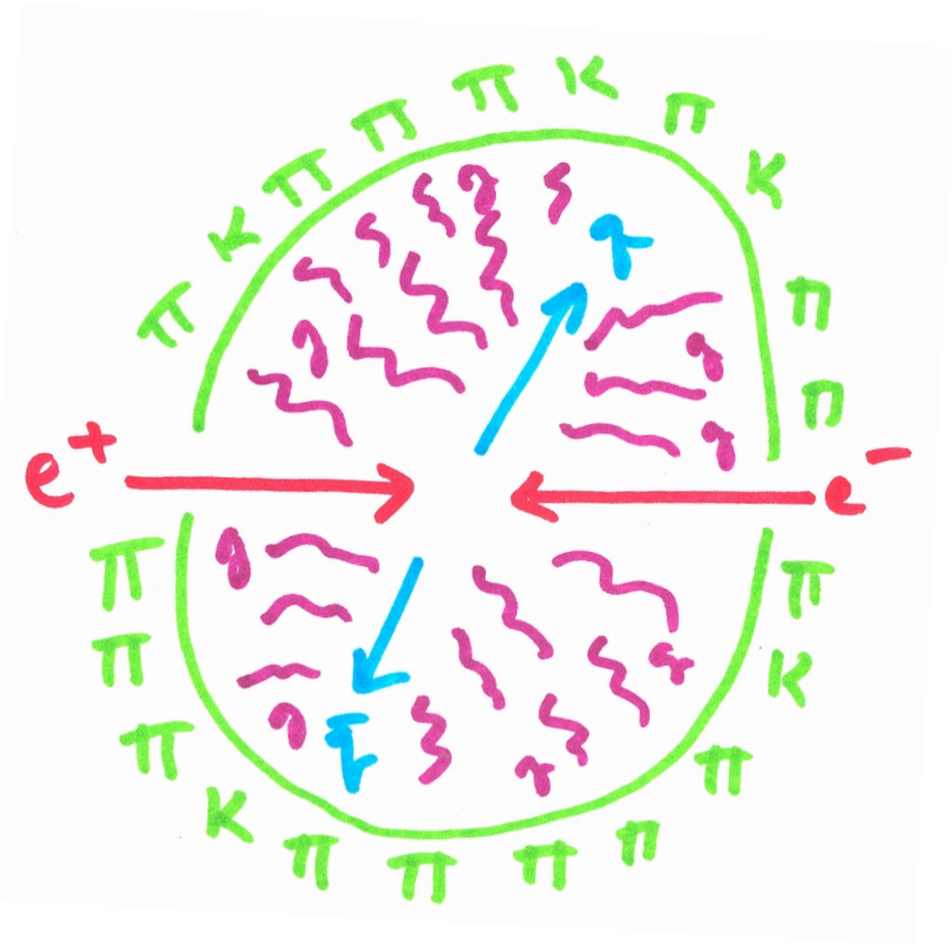
# Jets are not automatic!

## Quirky World
(QCD with only top quark)




Can't break color flux tubes!
Just "toponium" and glueballs!

## Quasi-Conformal World
($\beta \approx 0, g \approx 4\pi$)

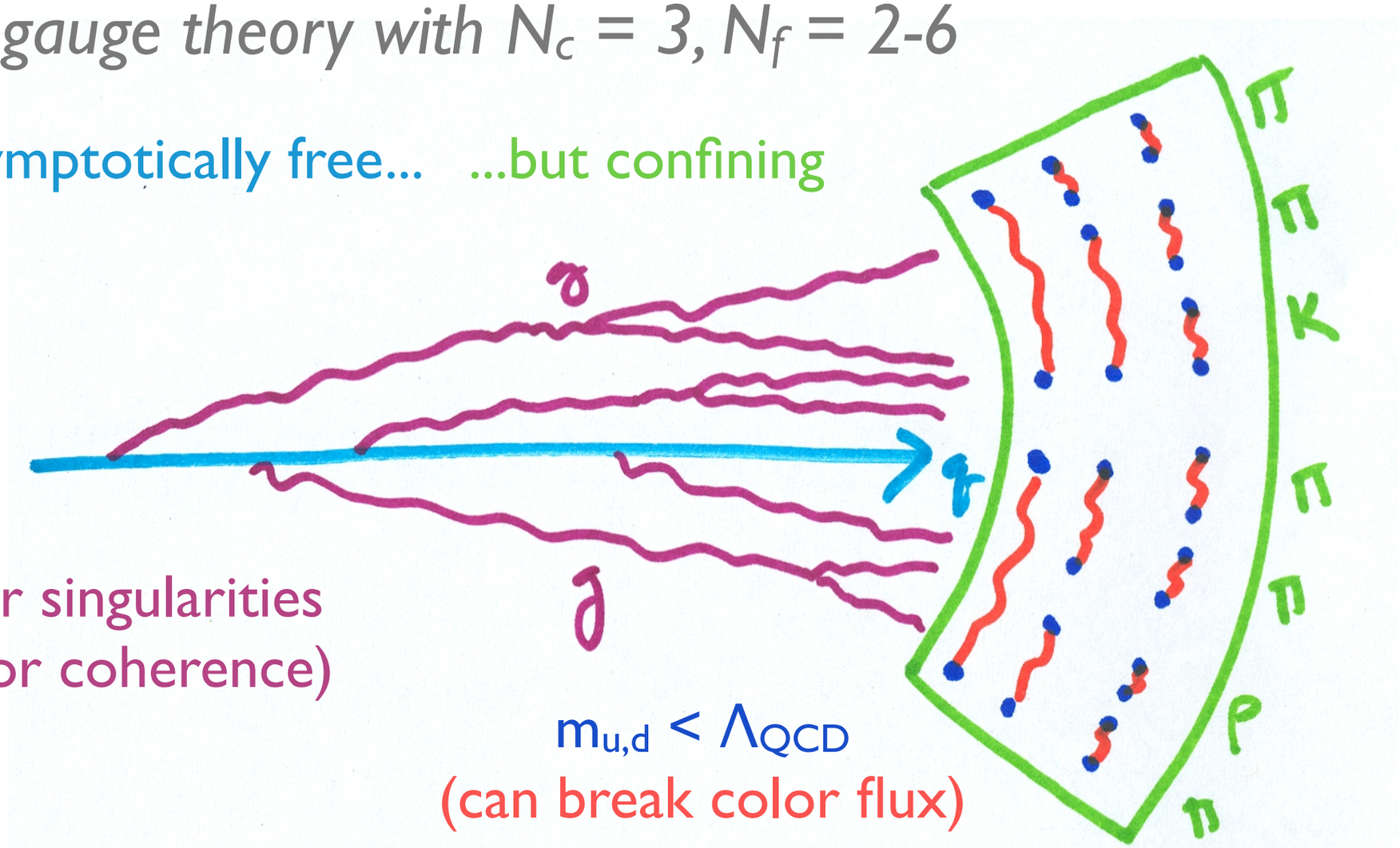



No hierarchy of scales!
All "spherical" events!

# Jets are emergent property of QCD

*Non-abelian gauge theory with $N_c = 3$, $N_f = 2$-6*

Asymptotically free... ...but confining




Collinear singularities
(and color coherence)

$m_{u,d} < \Lambda_{QCD}$
(can break color flux)

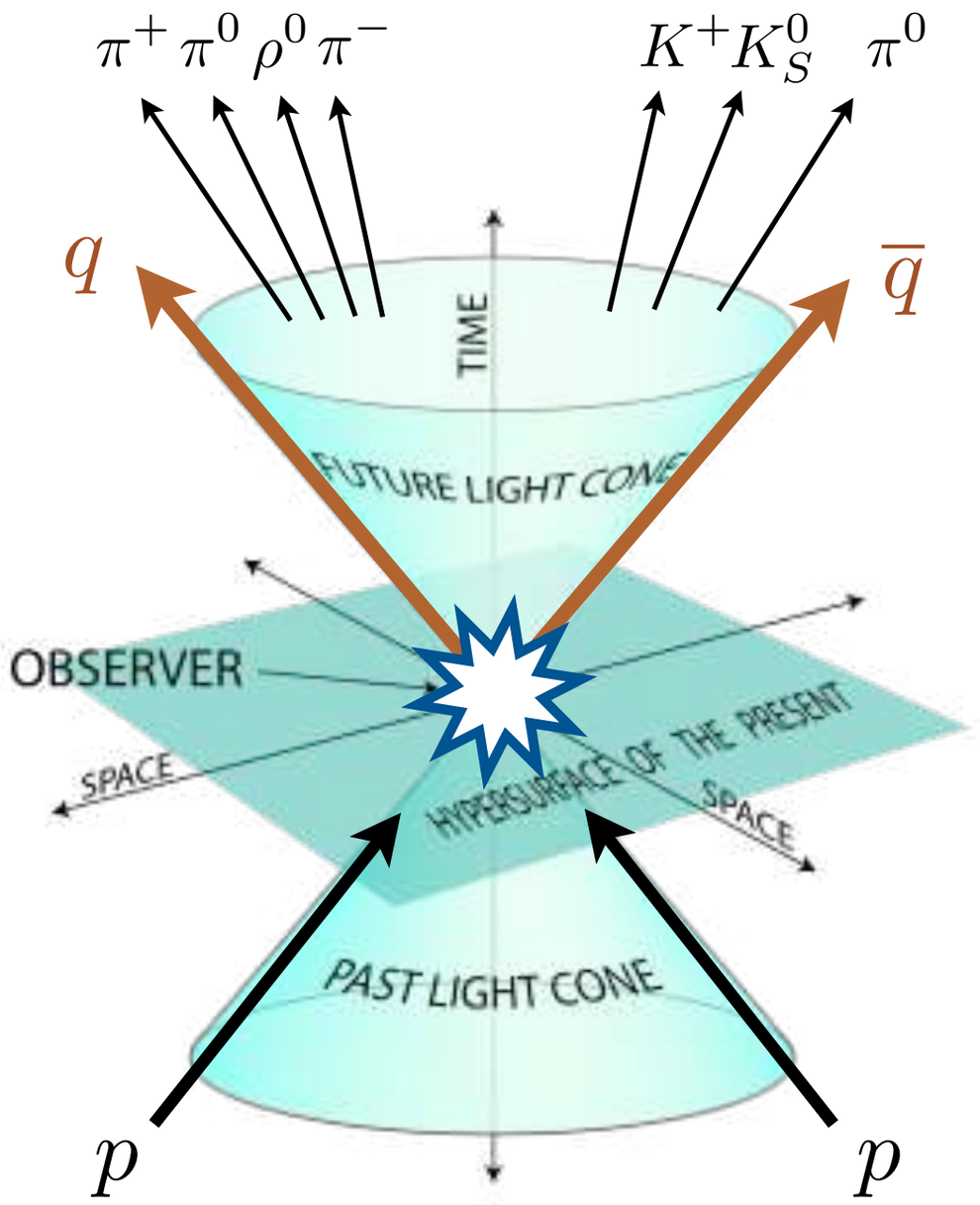
Energy flow in UV
(where partons go)

$\Rightarrow$

Energy flow in IR
(where hadrons go)

Jet = quark/gluon + radiation + ambiguities ($m_J \approx 10\%\ p_{TJ}$)

# Jet formation is a dynamic process



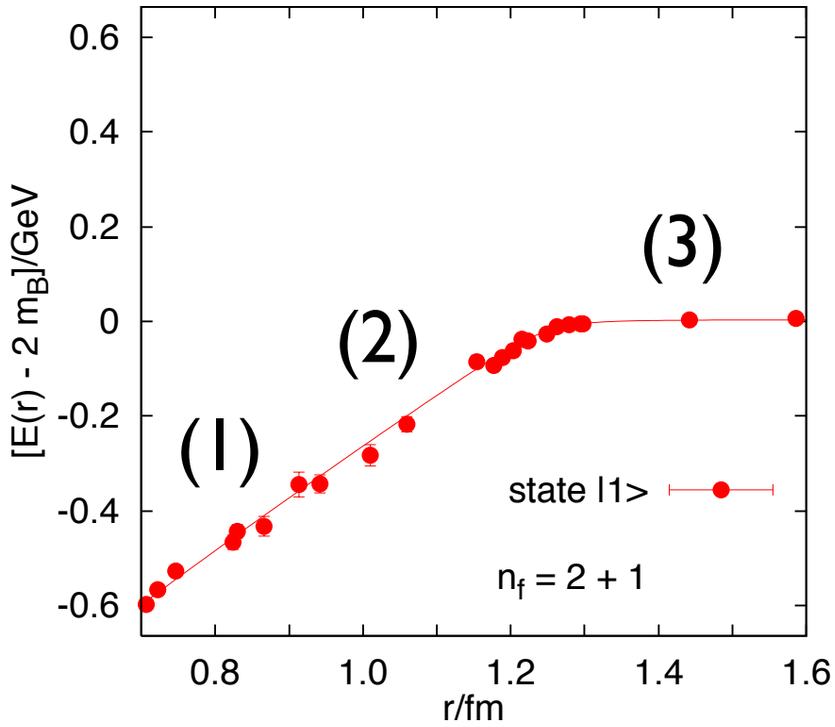

The dynamics of hadronization is one of the deepest mysteries in modern physics

How do short-distance quarks and gluons on the light cone turn into long-distance hadrons off the light cone?

We should be thinking more about this question

For this lecture, you'll have to live with cartoons

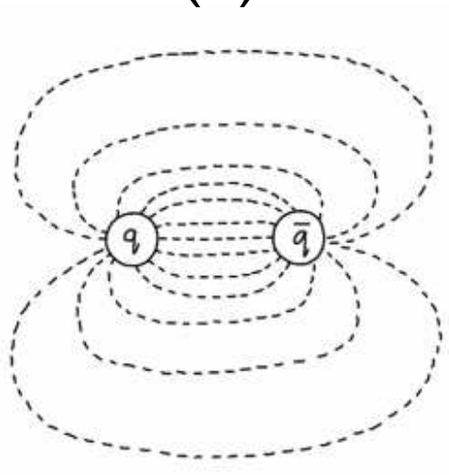
# Confinement/Liberation



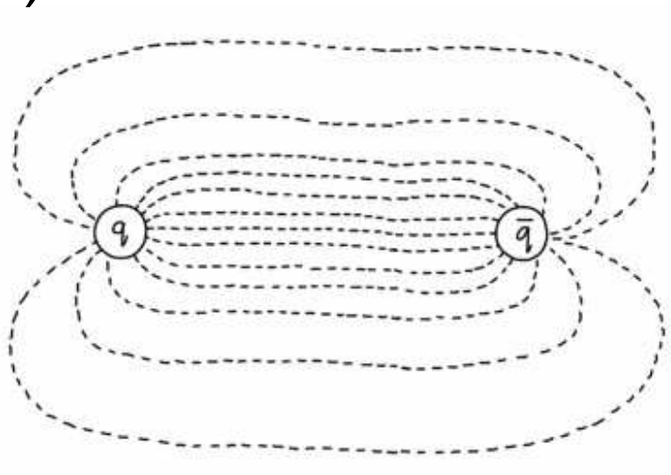

Potential between
two heavy quarks
(from lattice calculation)

[SESAM, 2005]

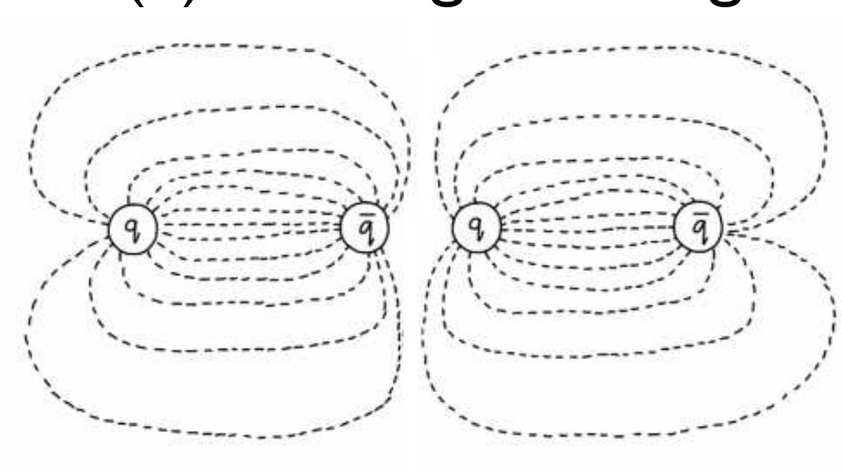
(1)




(2) = linear confinement




(3) = string breaking




[pictures from coffeeshopphysics.com]

# String picture gives jet basics

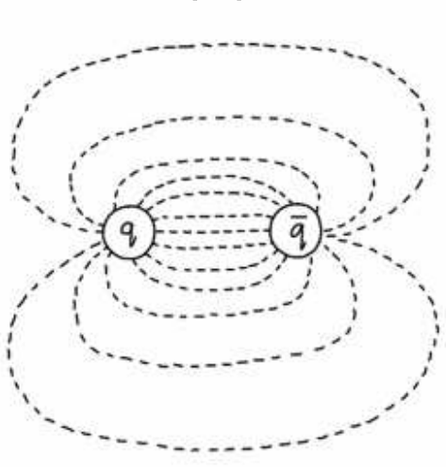
String breaks easily: Quark/gluon direction ≈ Jet direction

String has energy density: Massless quarks/gluons → Massive jet

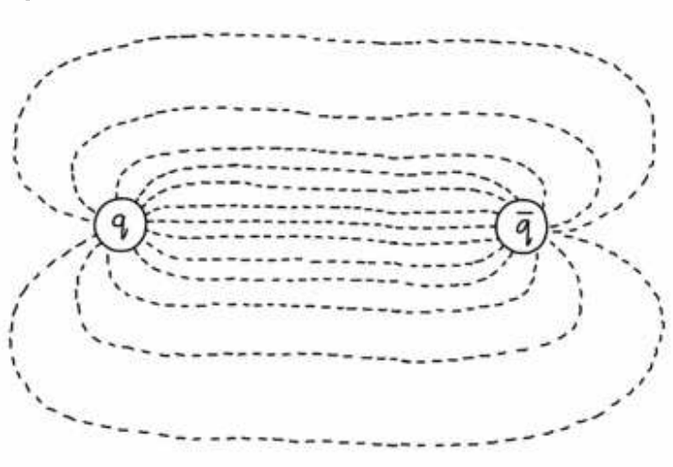
String breaks by popping quarks: Jets are mostly $q\bar{q}$ bound states (mesons)

String is color singlet: Jets are fundamentally ambiguous

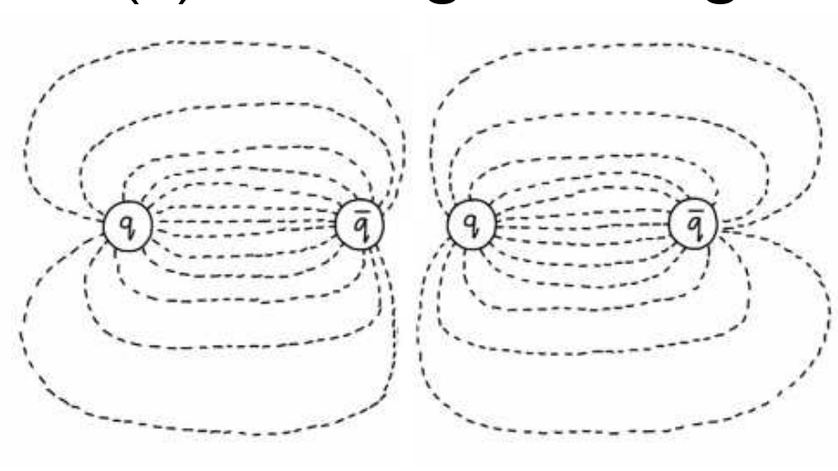
(1)




(2) = linear confinement

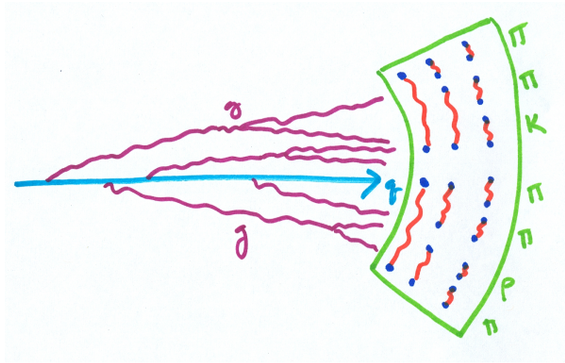


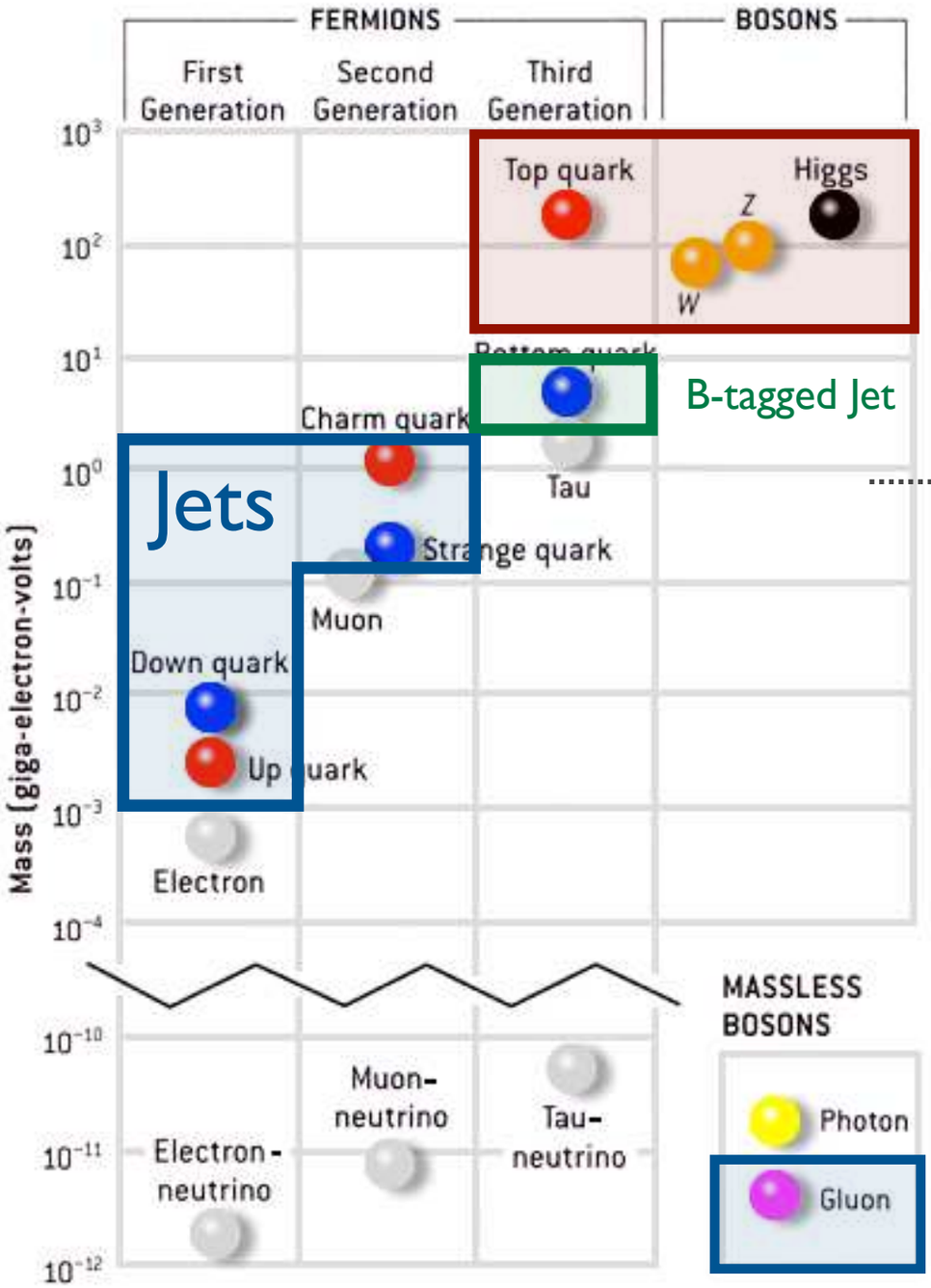

(3) = string breaking

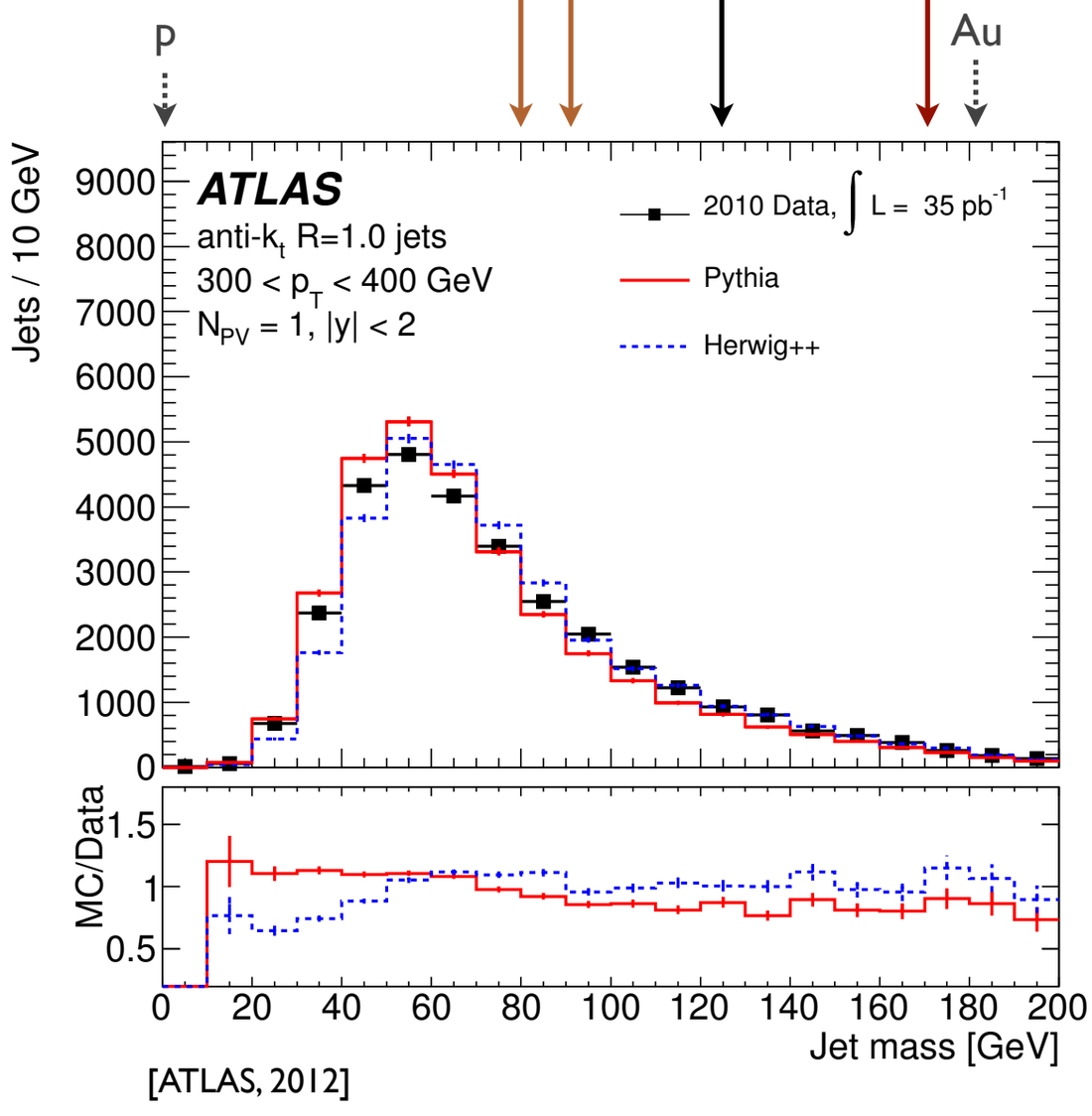



[pictures from coffeeshopphysics.com]

# Yes, jets really are massive

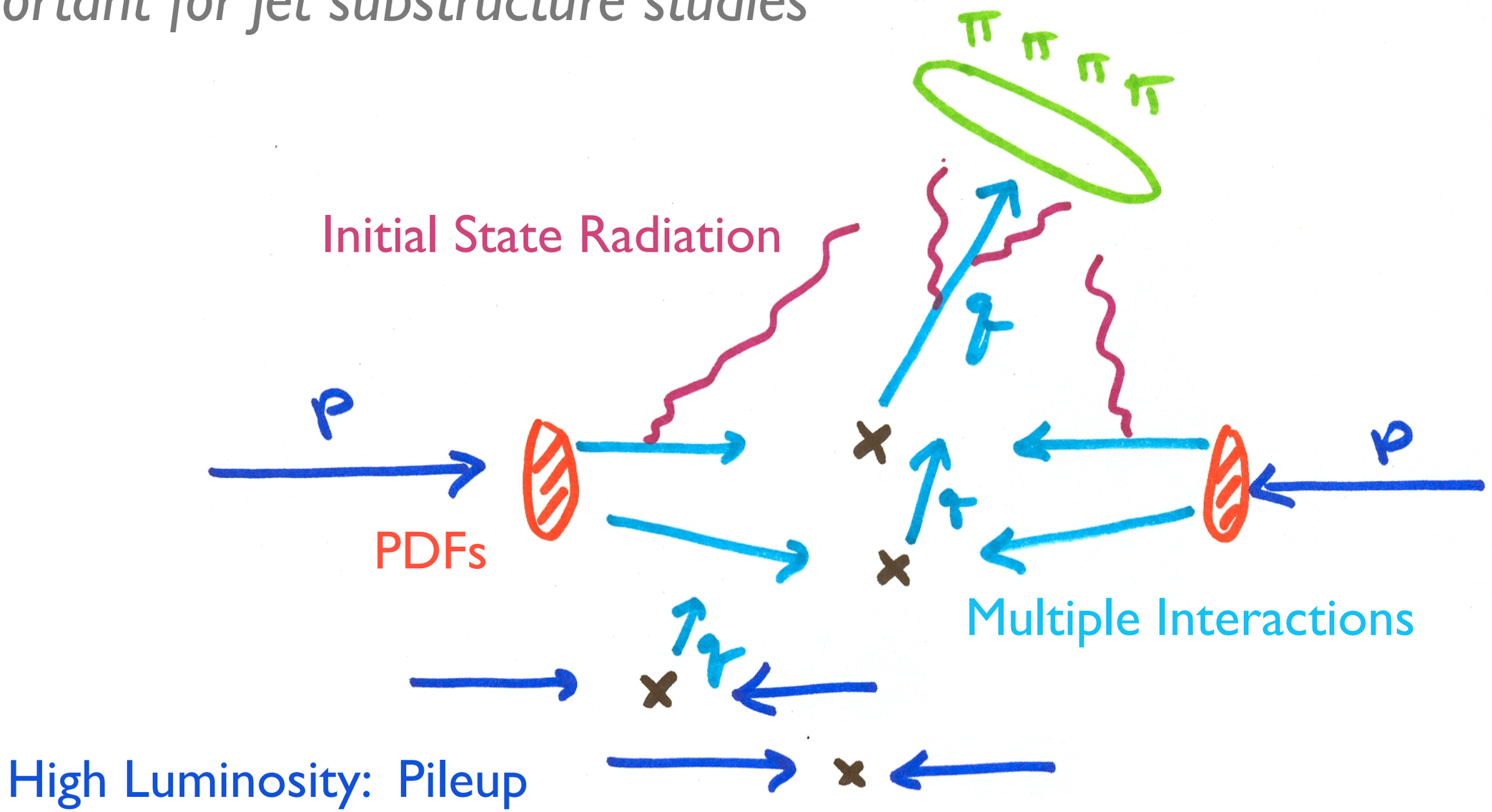




# Messiness is also a property of QCD

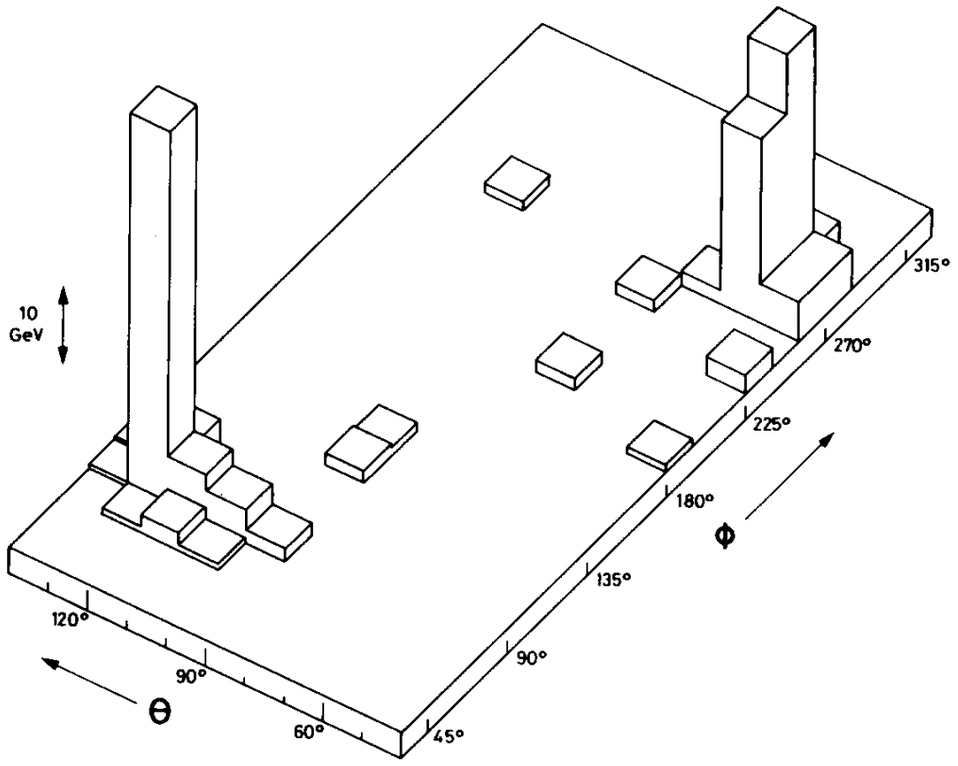
*Important for jet substructure studies*



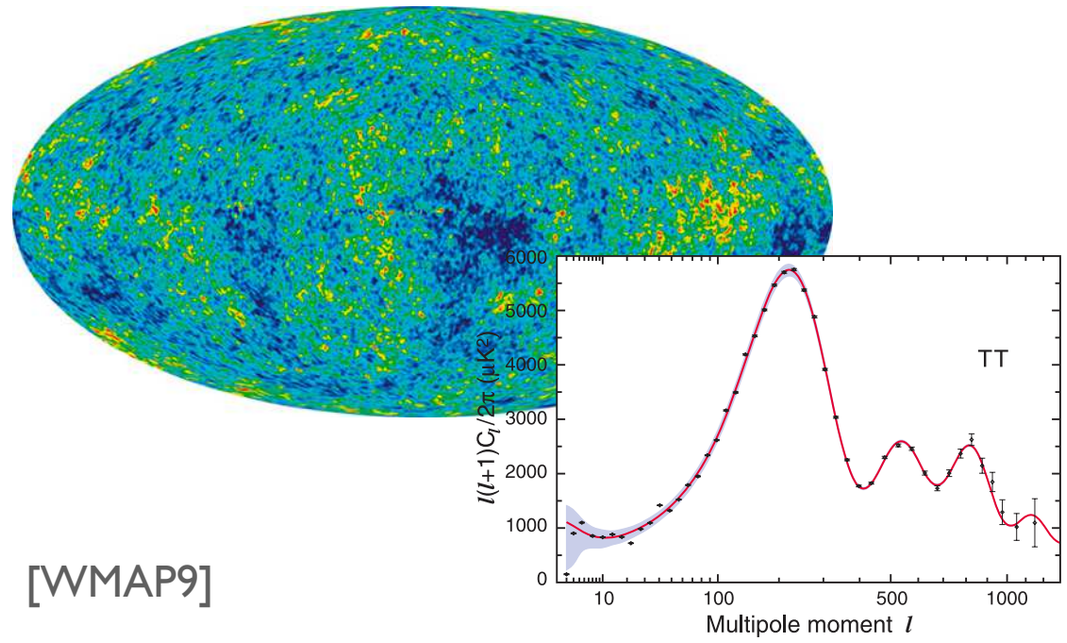

Jet = desired radiation from hard quark/gluon
+ additional contamination

# *The Shape of Jets*

# Identification of Jets




≠




[WMAP9]

By eye, jets are collimated along specific directions
Need "image processing", not "Fourier decomposition"

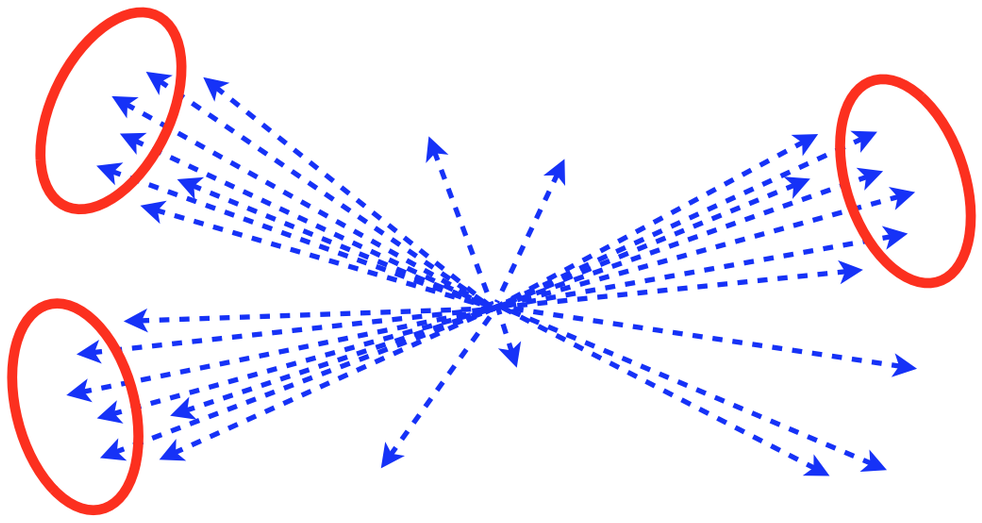
(Though my views on this are evolving…)

# Generic Jet Algorithm

**Inputs:** $\{p_1, p_2, \ldots, p_k\}_{\text{hadrons}}$ $\Rightarrow$ **Outputs:** $\{p_1, p_2, \ldots, p_N\}_{\text{jets}}$

Unless otherwise stated: $\sum_{i \in \text{jet}} p_i = p_{\text{jet}}$

(aka "E-scheme recombination", other schemes also plausible)



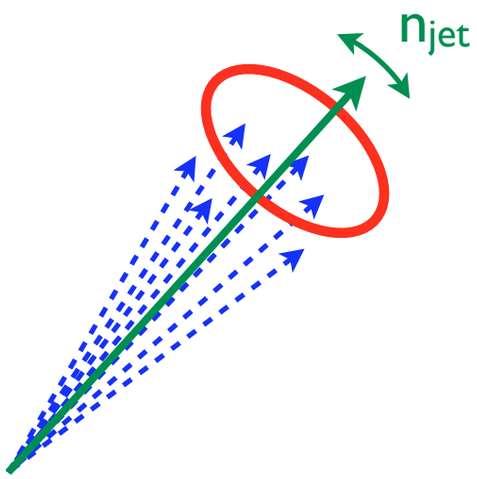
Remember, jets are massive: $p_{\text{jet}}^2 = \left( \sum_{i \in \text{jet}} p_i \right)^2 \geq \sum_{i \in \text{jet}} m_i^2 \geq m_{\text{quark/gluon}}^2$

# Simple Option: Cone Jets

$$\{p_1, p_2, \ldots, p_k\}_{\text{hadrons}} \Rightarrow \{p_1, p_2, \ldots, p_N\}_{\text{jets}}$$

1. Stable Cone:

Collection of hadrons within radius R, such that jet axis ($n_{jet}$) parallel to jet momentum ($p_{jet}$)




2. Cluster that maximizes: $E_{\text{jet}} - \frac{1}{R^2}\frac{m_{\text{jet}}^2}{E_{\text{jet}}}$

3. Axis that minimizes: $\sum_{i \in \text{event}} \min\left\{E_i, \frac{2n \cdot p_i}{R^2}\right\}$

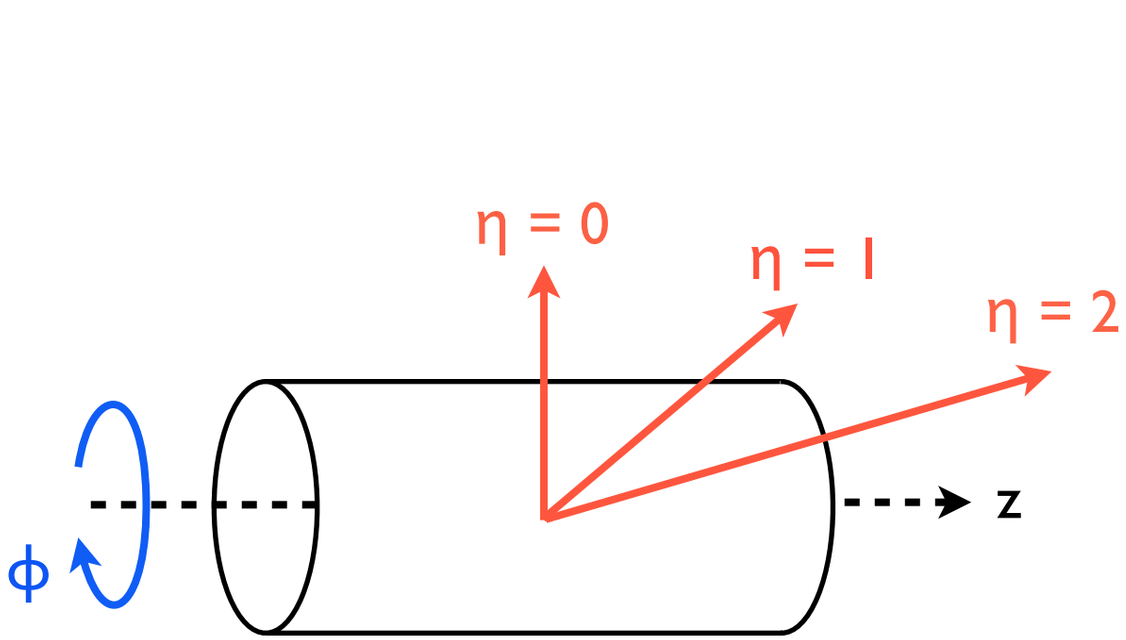
These are all fine (and essentially equivalent) algorithms
Complications when multiple overlapping cones (split/merge headaches)
Not used very often; worth revisiting with modern understanding

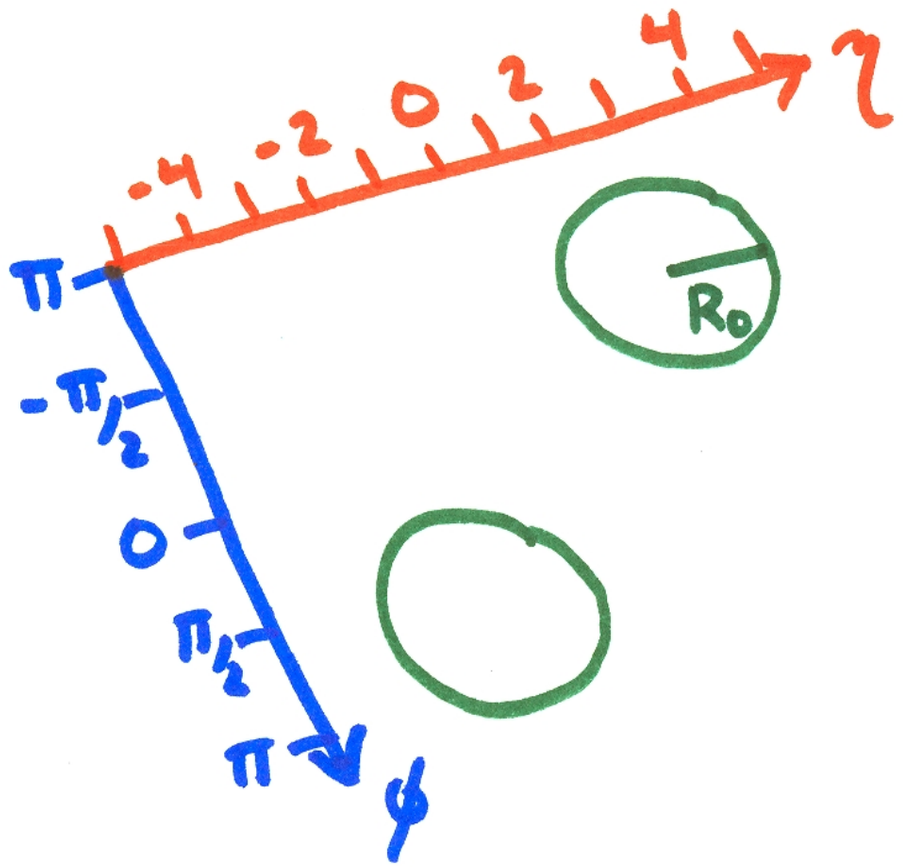
[JDT, PRD 2015; see also Wei, Naik, Harrow, JDT, arXiv 2019]

# Coordinate System for Jets

*Invariant to longitudinal boosts along beam direction*

$$p_T = \sqrt{p_x^2 + p_y^2} \qquad \Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$$



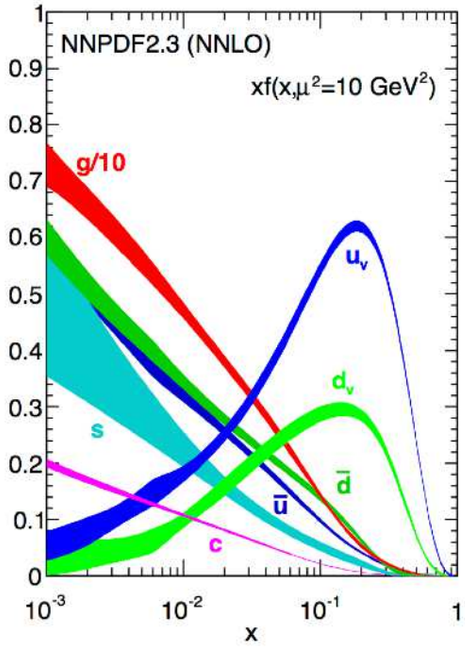

Cone Jet Algorithms: Punch holes of radius $R_0$

# Why $p_T$ & R (instead of E & θ)?

*Invariant to longitudinal boosts along beam direction*

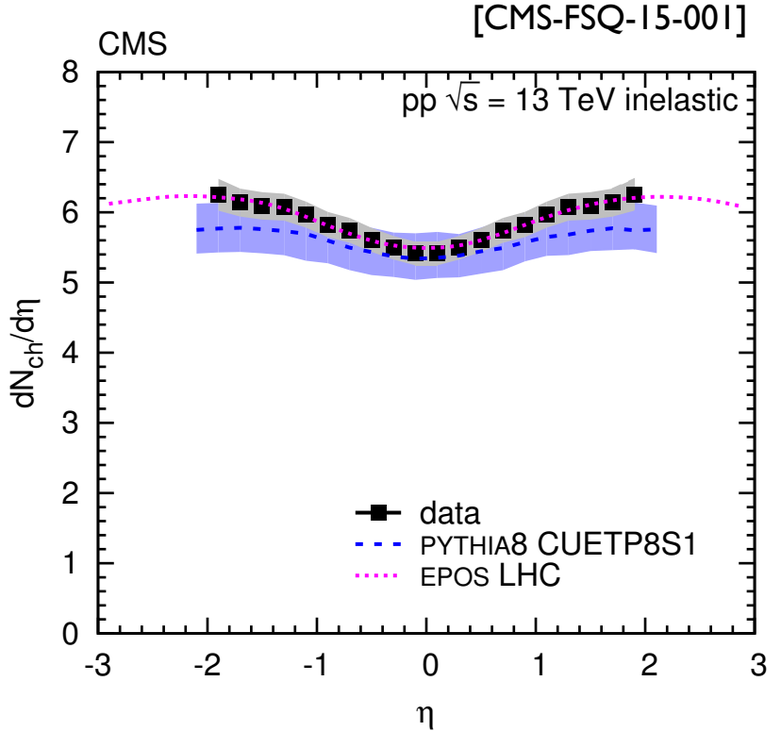
$$p_T = \sqrt{p_x^2 + p_y^2} \qquad \Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$$

Usual Answer: PDFs




Unknown longitudinal momentum fraction of colliding quarks/gluons → longitudinally-boost-invariance

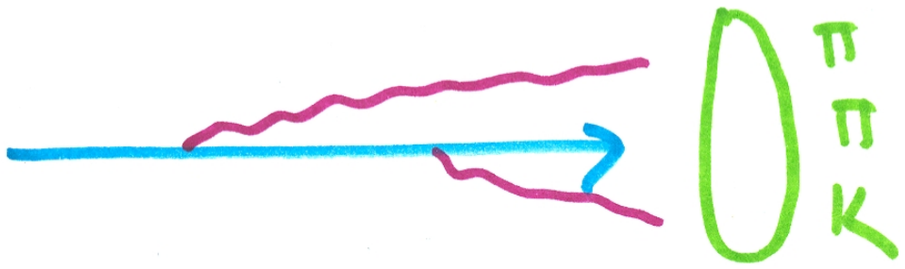
Better Answer: Rapidity Plateau



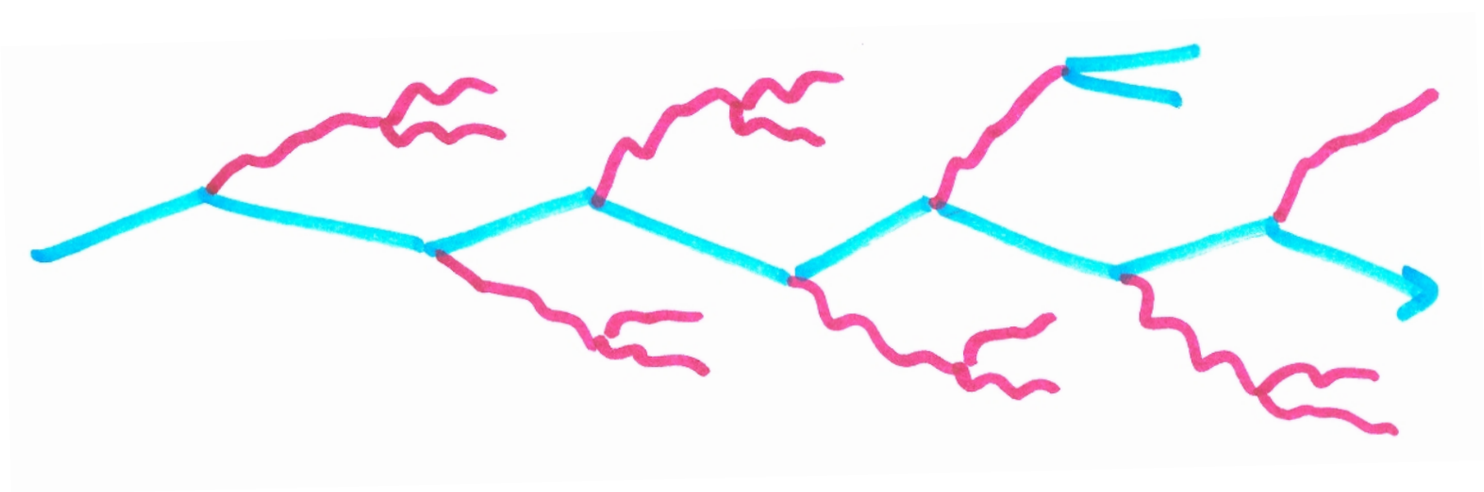

Roughly uniform underlying event as a function of η
A requirement we should probably revisit

# Goal of Jet Algorithm

*Cluster observed hadrons to approximate short-distance partons*




Clustering that mimics fragmentation ↔ Parton shower picture

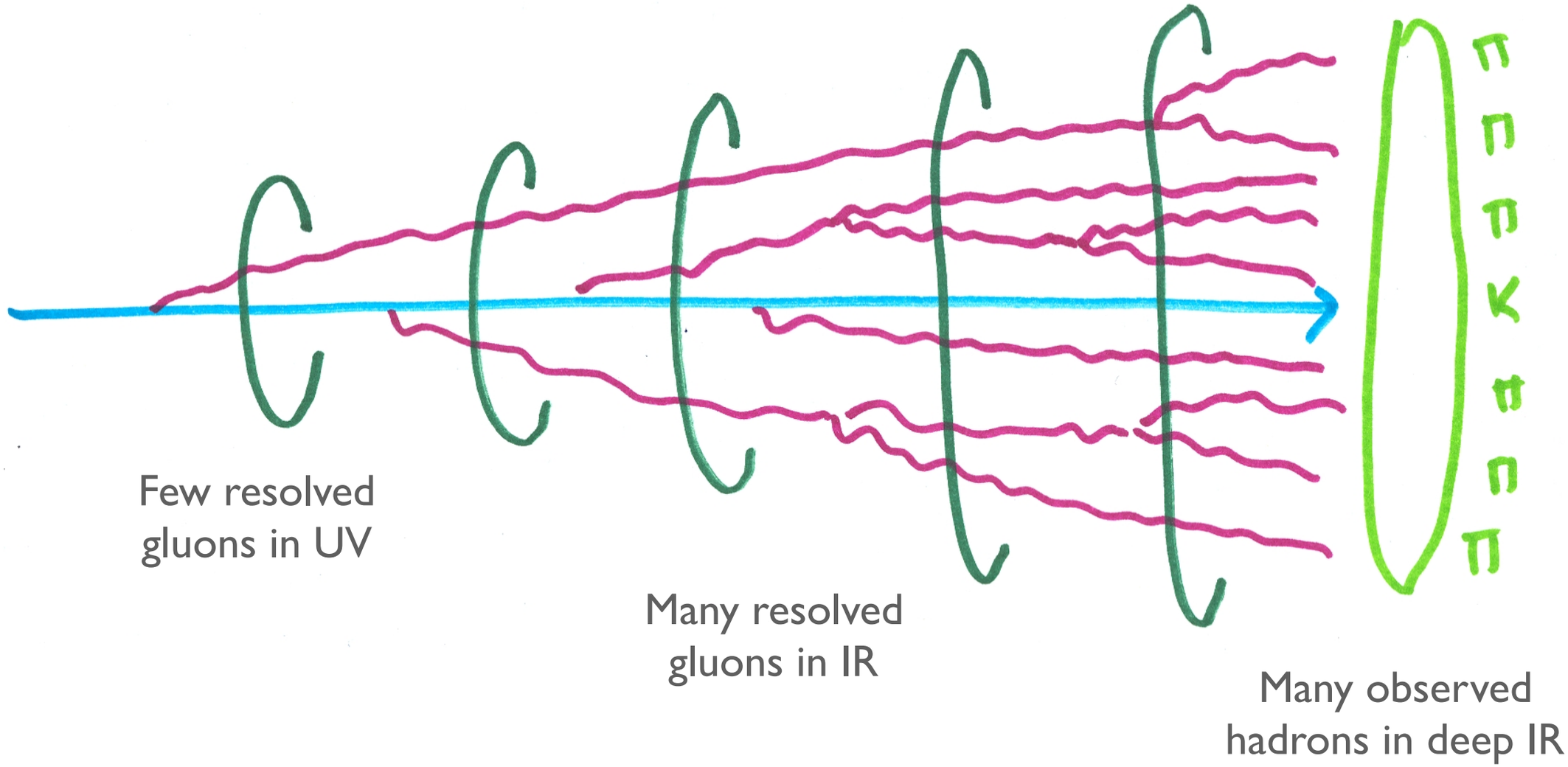


$$\mathrm{d}P_{i\to ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{\mathrm{d}\theta}{\theta} \frac{\mathrm{d}z}{z}$$

Collinear Soft

# “Fractal” Jets

*Intuitively, jets at long distances ≈ jets at short distances*




“Resolved” gluon emissions should not change jet structure
Cone algorithms scale pretty well from UV (partons) to IR (hadrons)
Recursive clustering algorithms attempt to mimic flow from UV to IR

# Recursive Jet Algorithms

*Workhorse of modern jet physics*

$d_{ij}$ — Merge

$d_i$ — Stop

1) Make list of protojets
2) Calculate all $d_{ij}$ and $d_i$
3) If $d_{ij}$ smallest: combine $i + j$ — If $d_i$ smallest: $i$ is finished
4) Repeat until nothing left

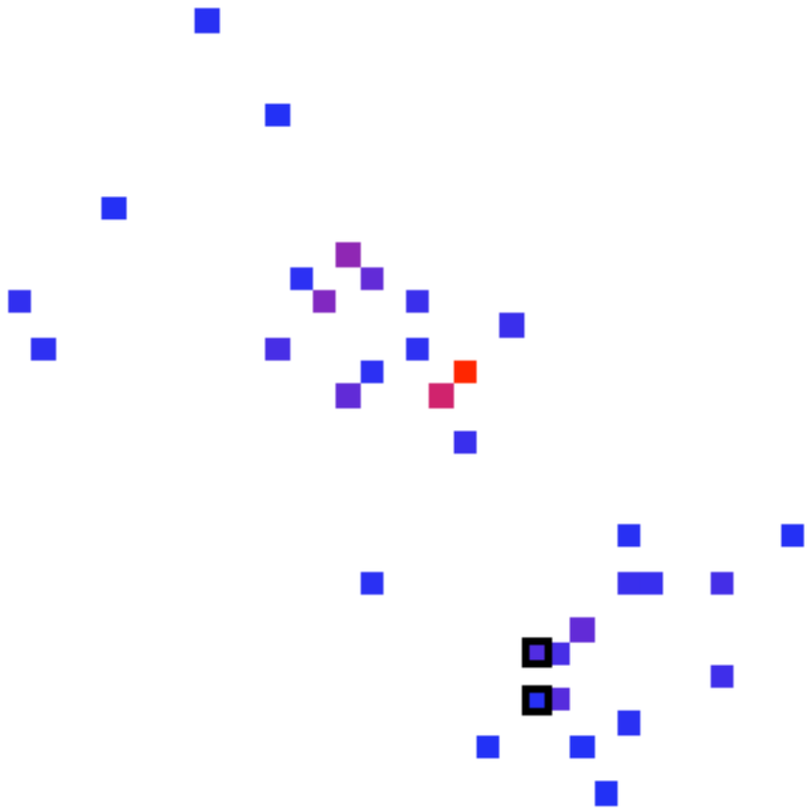
[starting with JADE; Durham, …]

Works the same on calorimeter cells, hadrons & partons

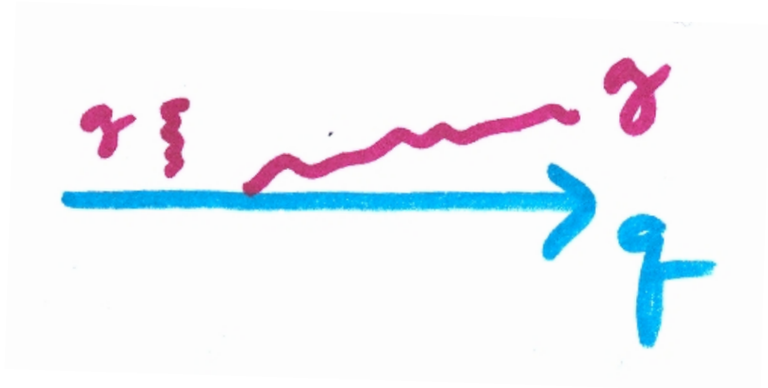
Computationally efficient, flexible, defined to all orders in $\alpha_S$, (mostly) pathology free

# $k_T$ Measure

*Jet built from soft/collinear singularities of QCD*



Animation from Jon Walsh
& Chris Vermilion



$$d_i = p_{Ti}^2$$

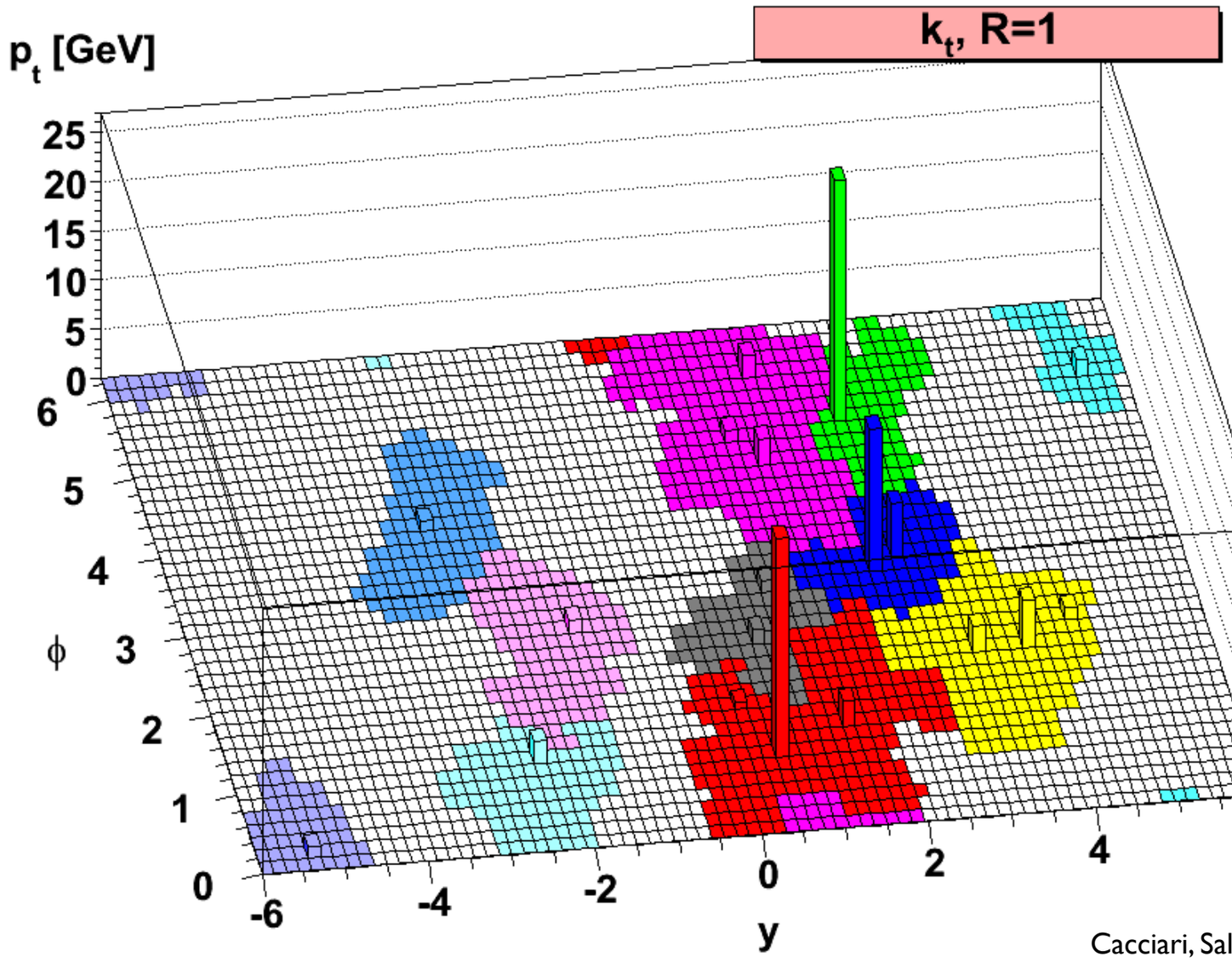
$$d_{ij} = \min\left[p_{Ti}^2, p_{Tj}^2\right] \frac{R_{ij}^2}{R_0^2}$$

Soft Collinear

[Catani, Dokshitzer, Seymour, Webber, 1993;
Ellis, Soper, 1993]

# $k_T$ Measure

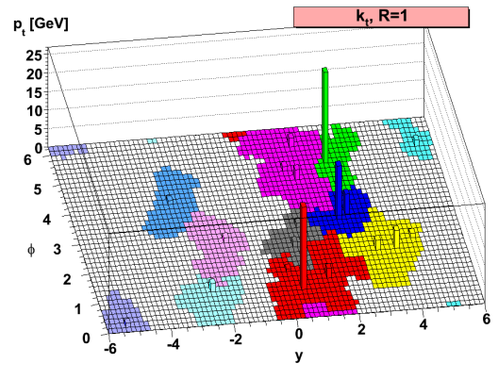
*Jet built from soft/collinear singularities of QCD*



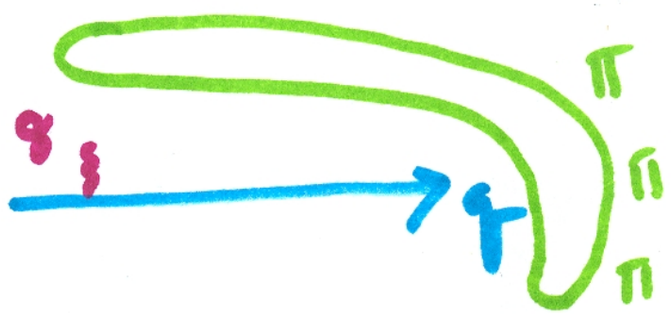

[figure from Cacciari, Salam, Soyez, JHEP 2008]

Funny looking jets, but calculable in pQCD

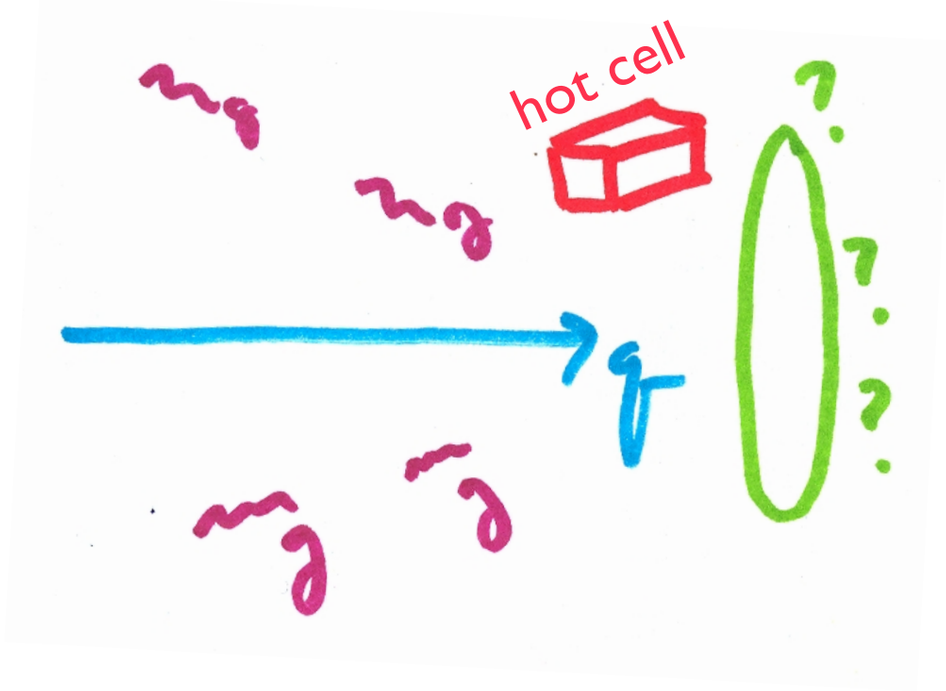
# Amoeba Jets?




QCD has soft singularity...



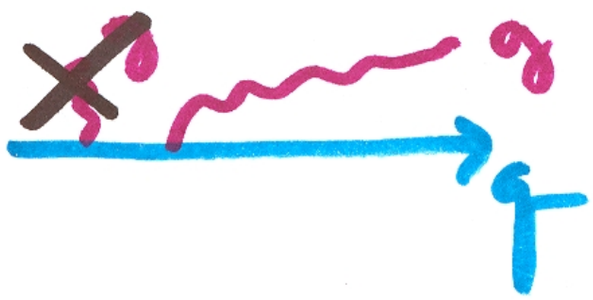

...but ISR/Multiple Interactions/Pileup/Detector Noise also soft




Nothing intrinsically wrong with $k_T$-style clustering,
but (over-)emphasizes soft physics

# Cambridge/Aachen Measure

*Uses only angular information, ignores soft singularity*

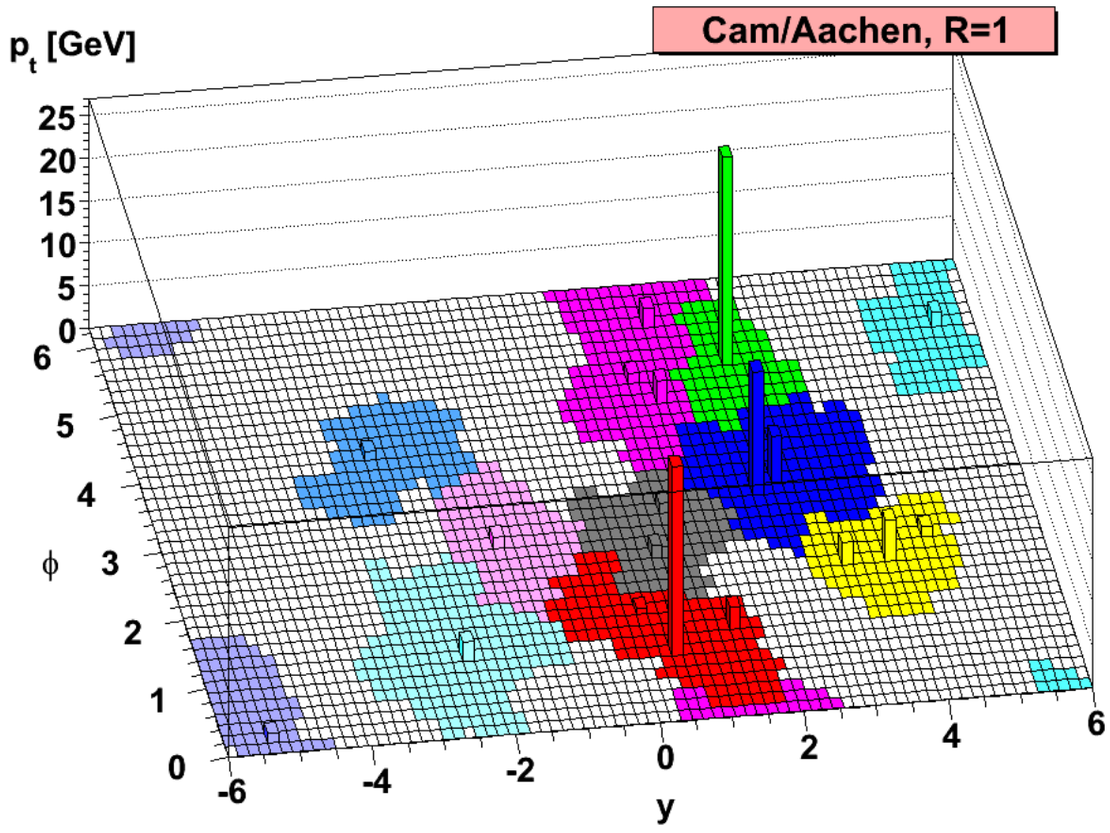


$$d_i = 1$$

$$d_{ij} = \frac{R_{ij}^2}{R_0^2}$$

Just Collinear

[Dokshitzer, Leder, Moretti, Webber, 1997; Wobisch and Wengler, 1998]

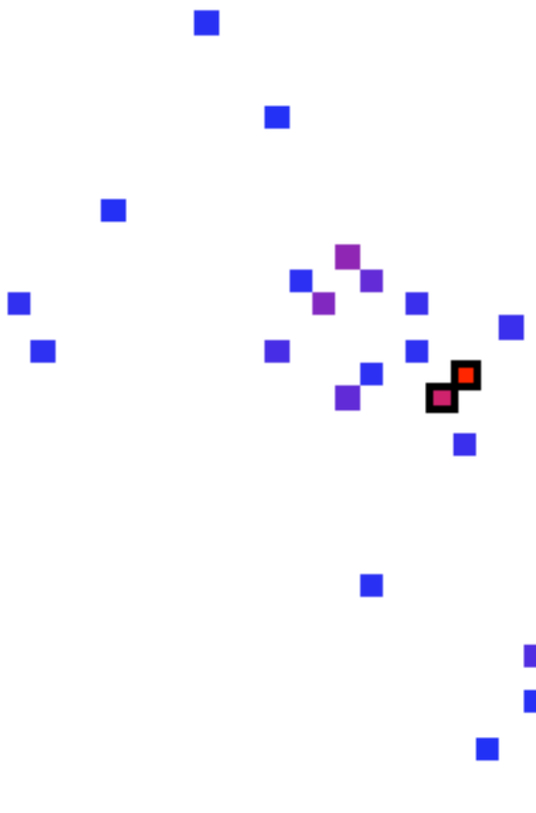



[figure from Cacciari, Salam, Soyez, JHEP 2008]

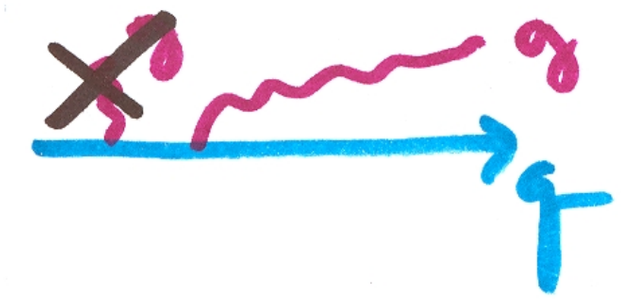
Still amoeba jets…

# Anti-$k_T$ Measure

*"Invert" soft singularity, build jets from hard to soft*



Animation from Jon Walsh
& Chris Vermilion



$$d_i = \frac{1}{p_{Ti}^2}$$

$$d_{ij} = \frac{1}{\max\left[p_{Ti}^2, p_{Tj}^2\right]} \frac{R_{ij}^2}{R_0^2}$$

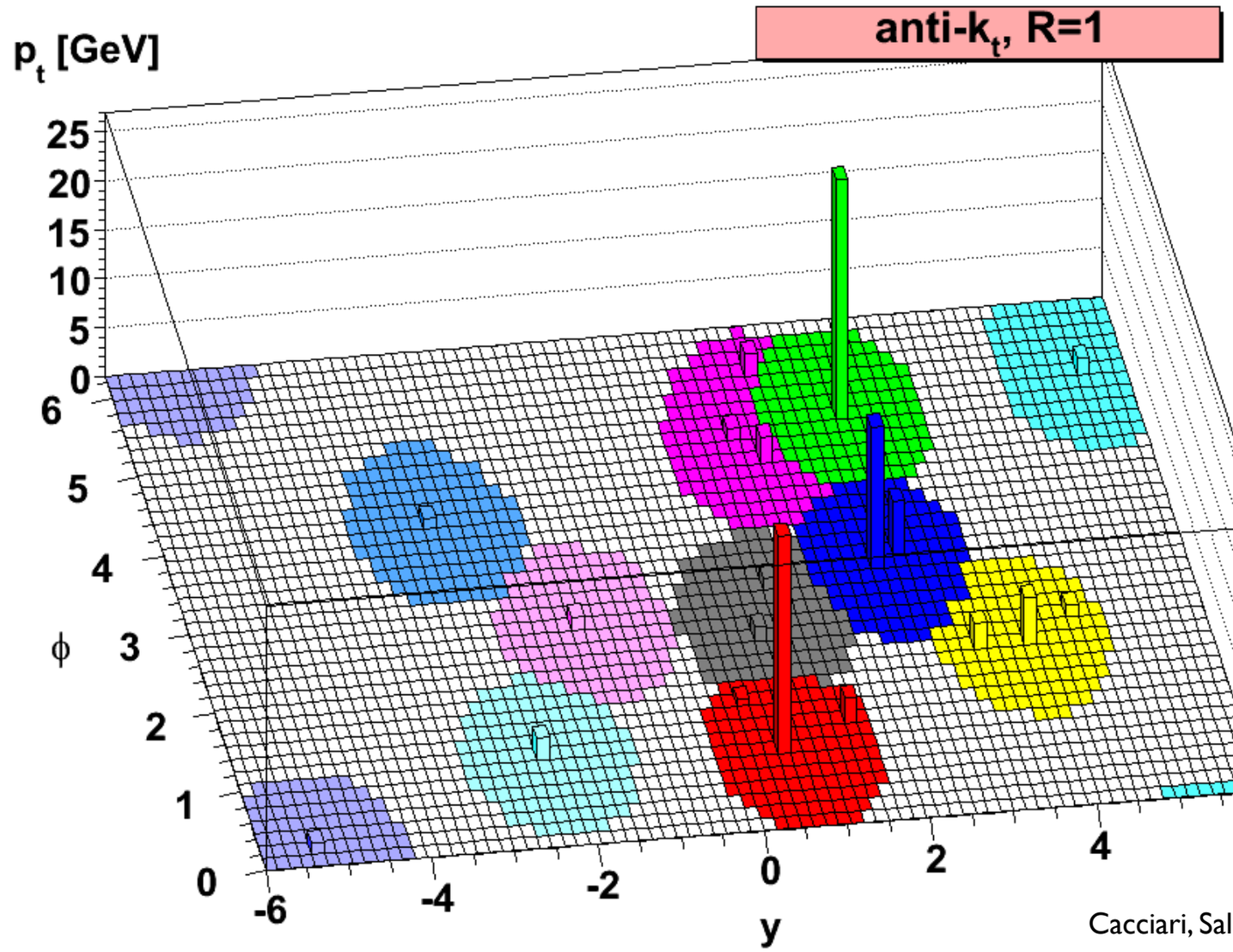
Magic Collinear

[Cacciari, Salam, Soyez, JHEP 2008]

when I was a postdoc @ Berkeley

# Anti-$k_T$ Measure

*Ignore soft singularity, and build jets from hard to soft*

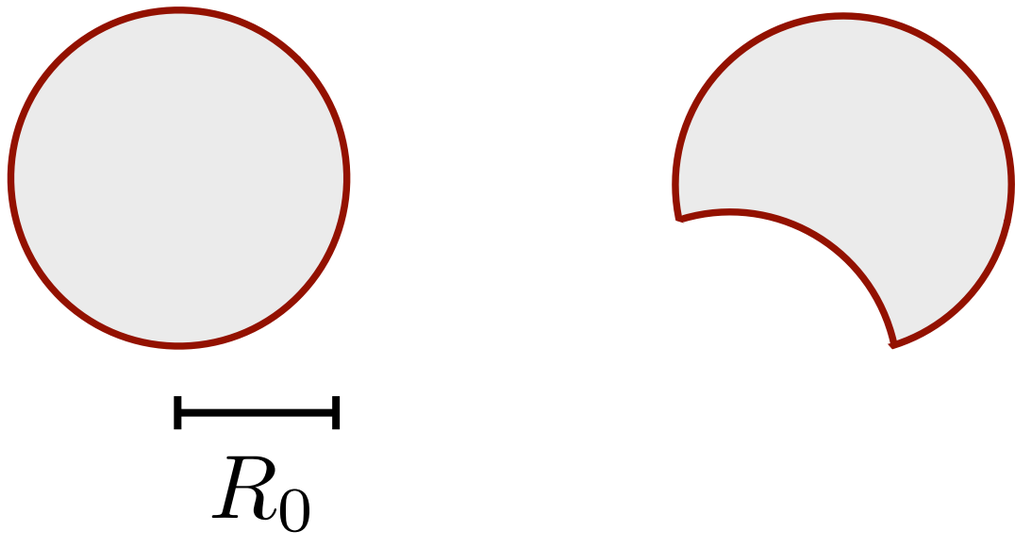



[figure from
Cacciari, Salam, Soyez, JHEP 2008]

Recursive, calculable, soft-insensitive, conical
(essentially equivalent to stable cone finding in eikonal limit)

# The Current Shape of Jets

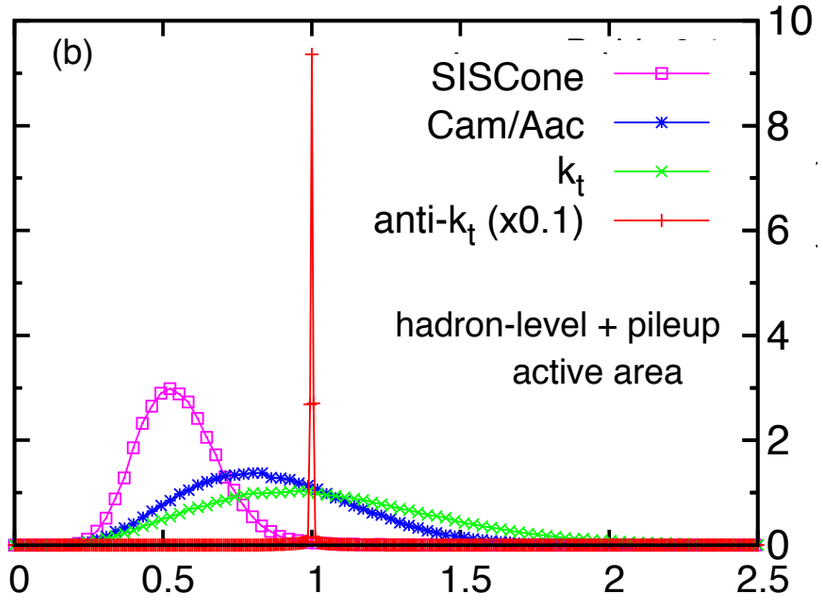
*Almost every analysis at the LHC uses anti-$k_T$*

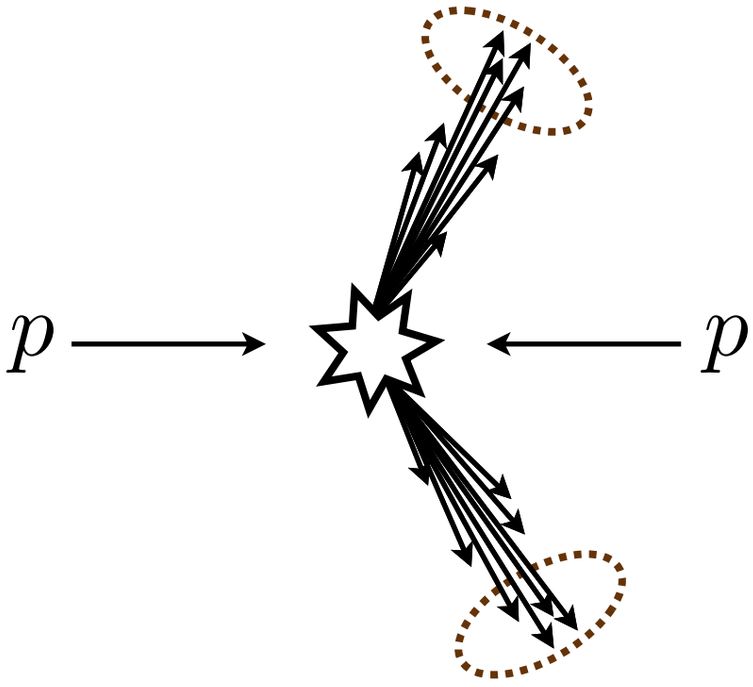


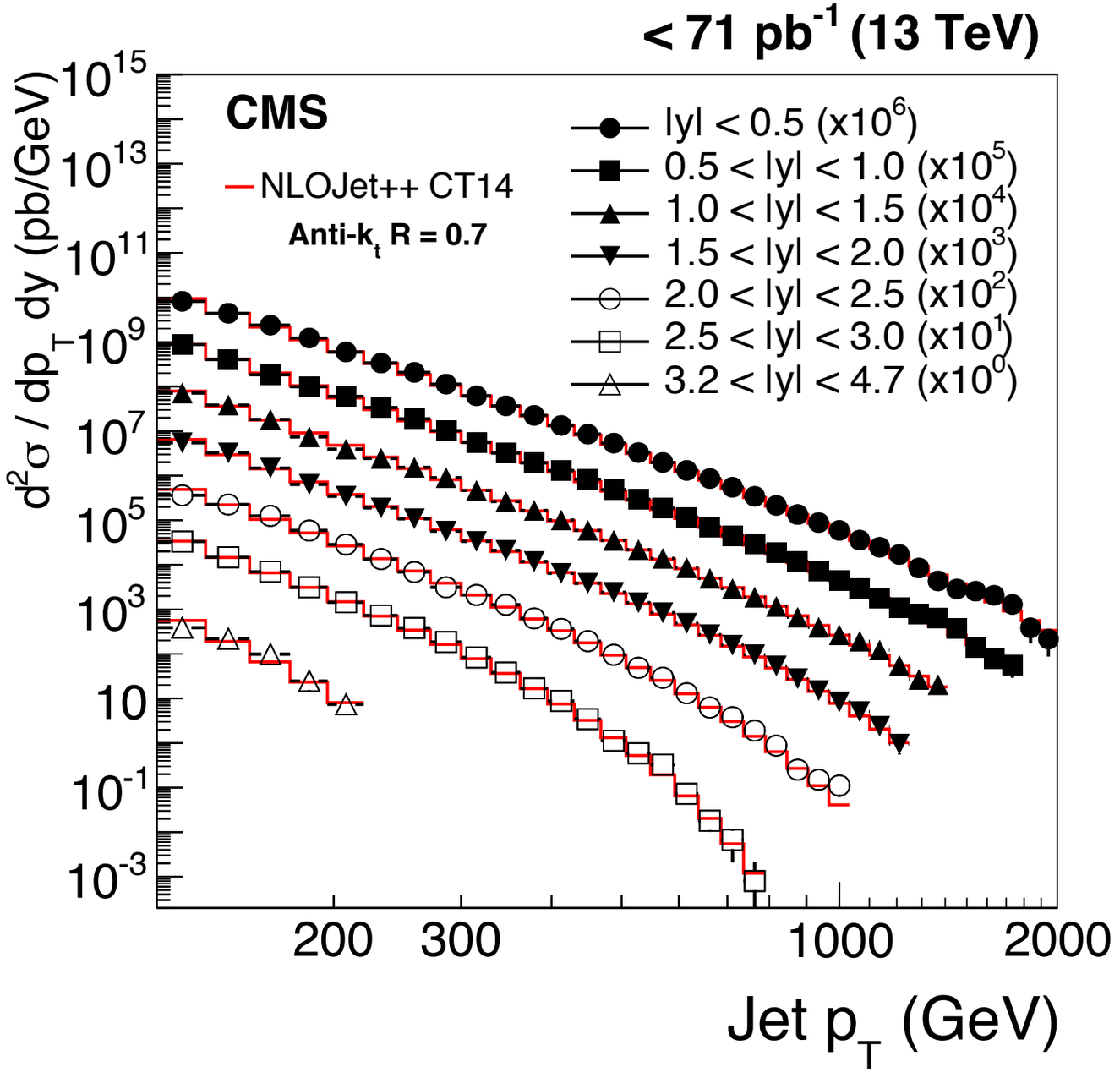

## Nice feature: uniform catchment area

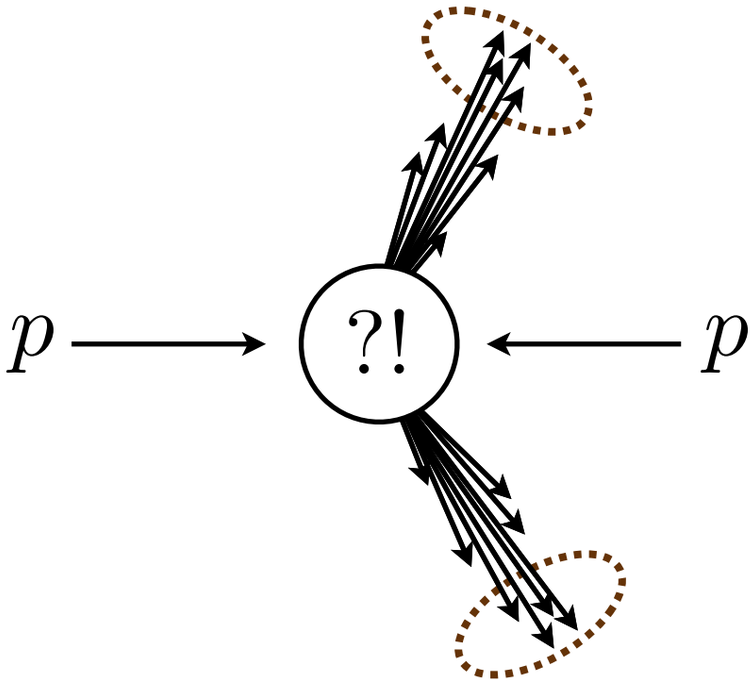
(just like idealized cones!)

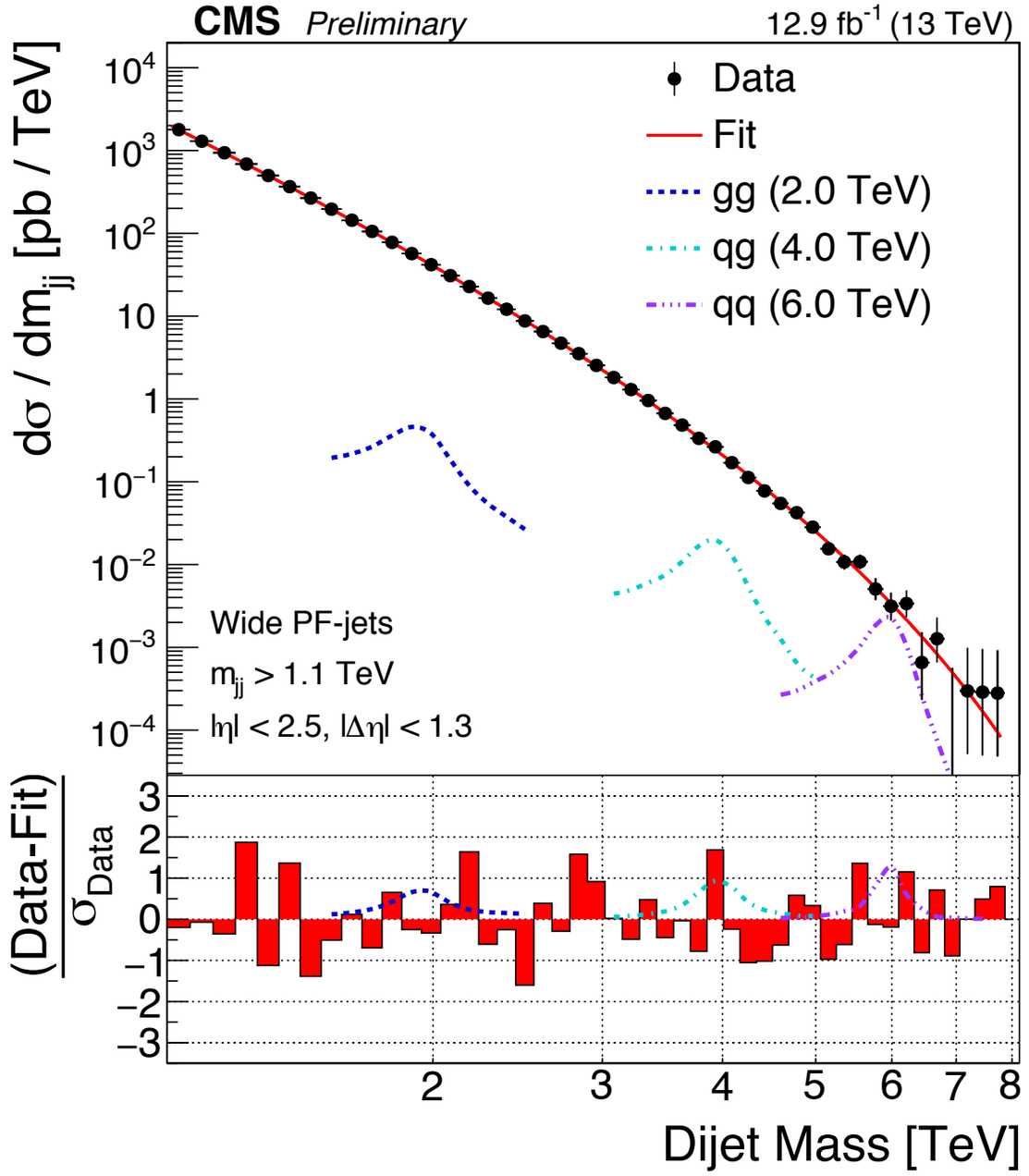
[Cacciari, Salam, Soyez, JHEP 2008, see also JHEP 2008]

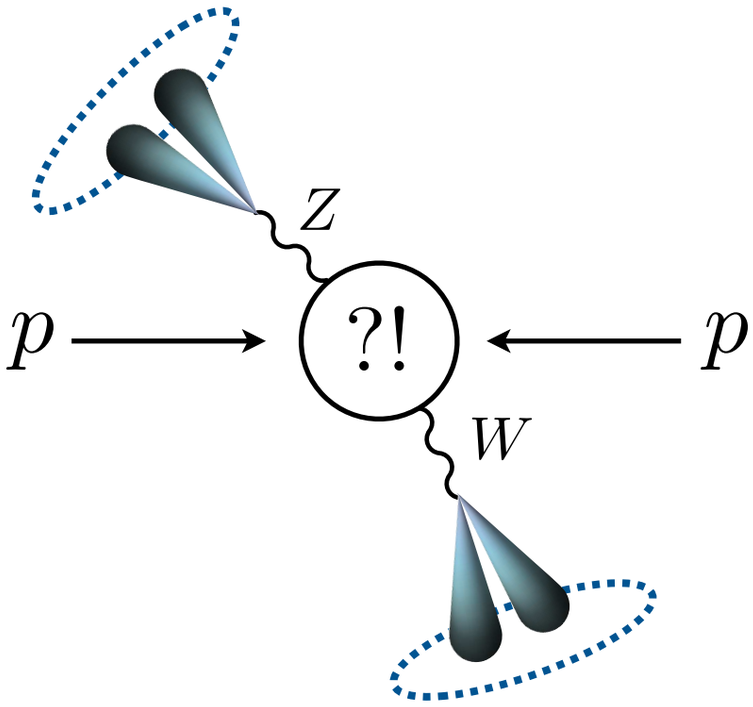


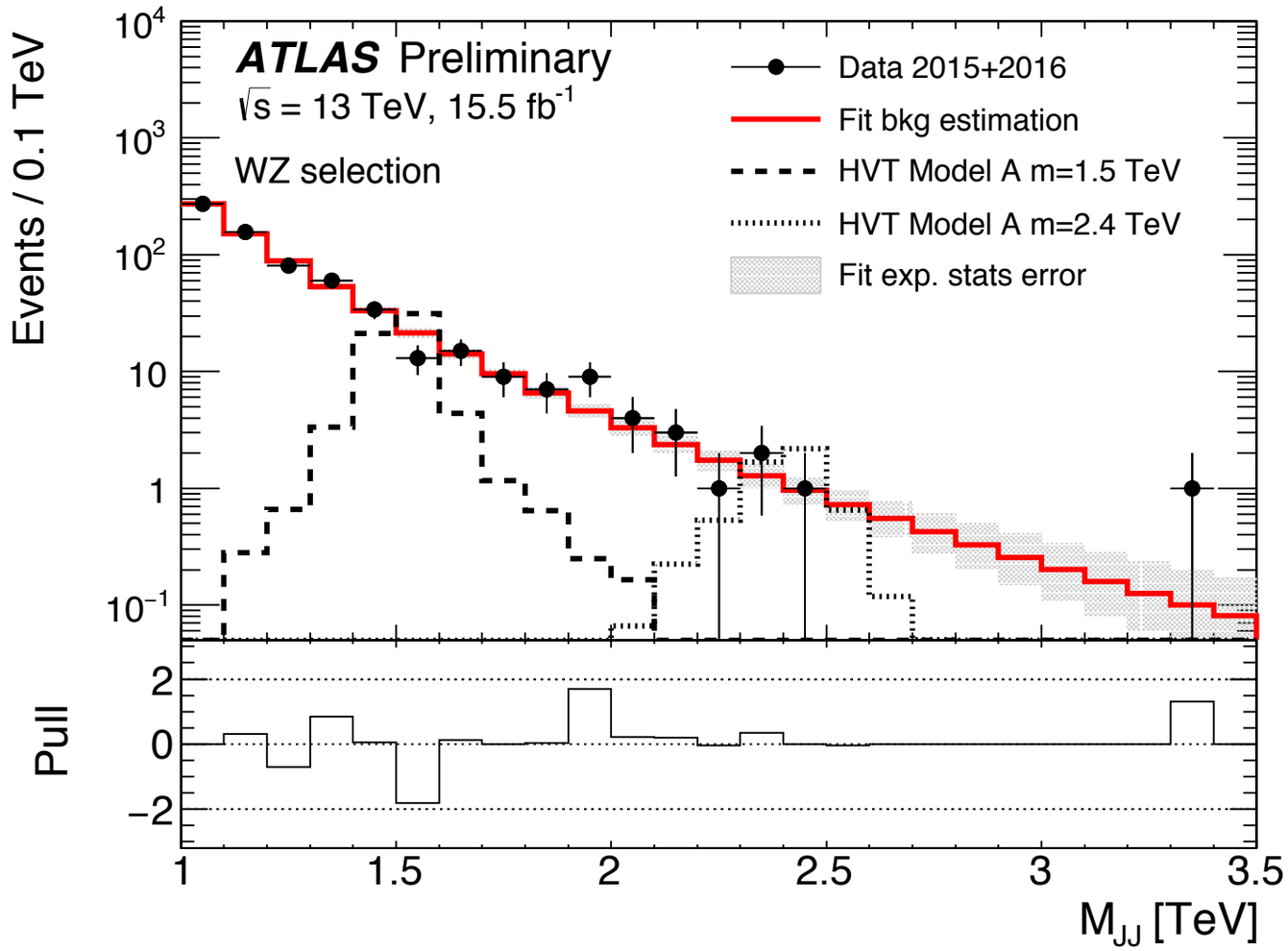

p
p

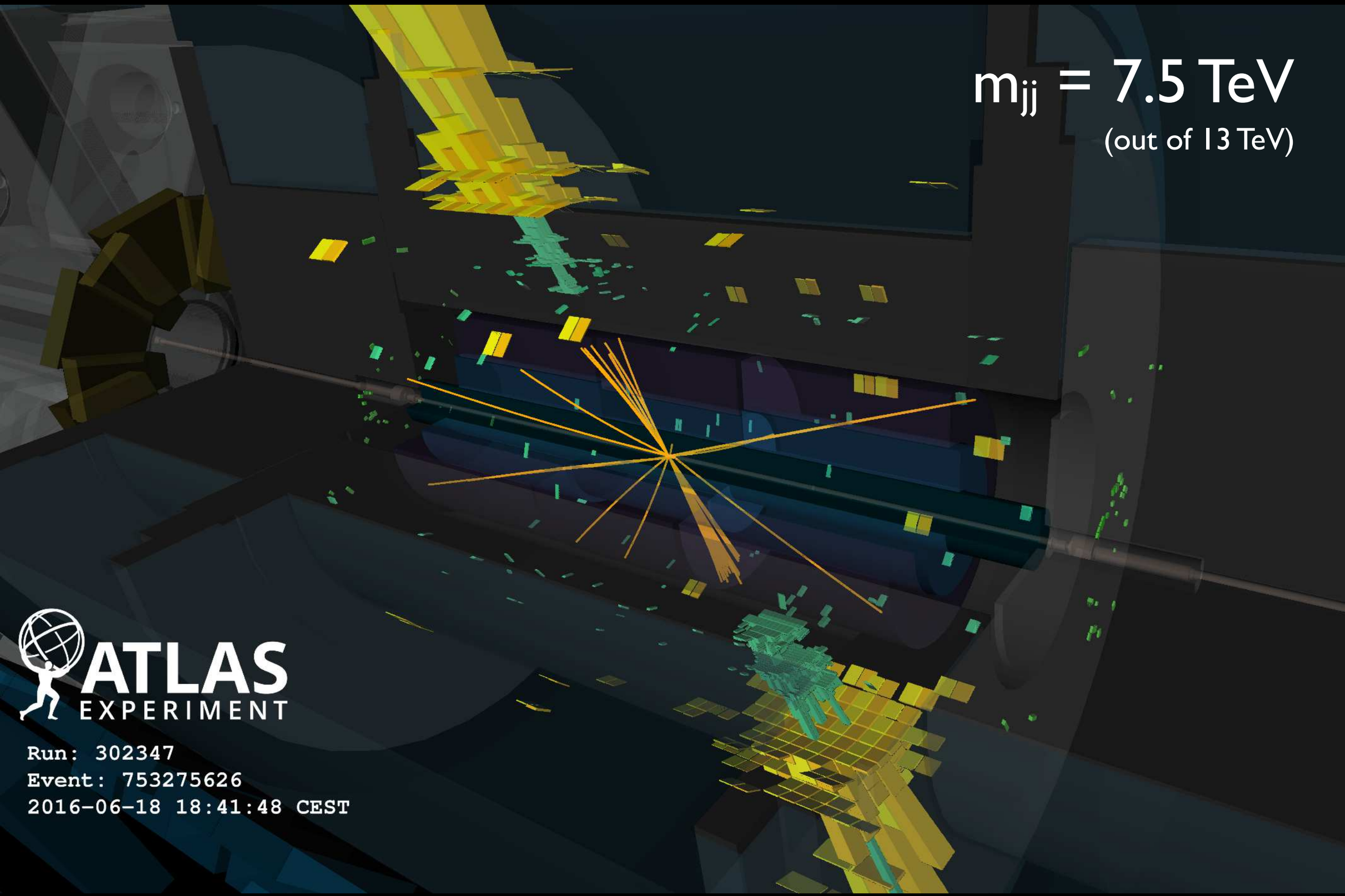

< 71 pb-1 (13 TeV)
CMS
NLOJet++ CT14
Anti-kt R = 0.7
|y| < 0.5 (x10^6)
0.5 < |y| < 1.0 (x10^5)
1.0 < |y| < 1.5 (x10^4)
1.5 < |y| < 2.0 (x10^3)
2.0 < |y| < 2.5 (x10^2)
2.5 < |y| < 3.0 (x10^1)
3.2 < |y| < 4.7 (x10^0)
d^2σ / dp_T dy (pb/GeV)
10^15
10^13
10^11
10^9
10^7
10^5
10^3
10
10^-1
10^-3
200
300
1000
2000
Jet p_T (GeV)


p
?!
p


CMS Preliminary
12.9 fb-1 (13 TeV)
Data
Fit
gg (2.0 TeV)
qg (4.0 TeV)
qq (6.0 TeV)
Wide PF-jets
mjj > 1.1 TeV
|η| < 2.5, |Δη| < 1.3
dσ / dmjj [pb / TeV]
(Data-Fit) / σData
Dijet Mass [TeV]


Z
p
?!
p
W


ATLAS Preliminary
√s = 13 TeV, 15.5 fb-1
WZ selection
Data 2015+2016
Fit bkg estimation
HVT Model A m=1.5 TeV
HVT Model A m=2.4 TeV
Fit exp. stats error
Events / 0.1 TeV
Pull
MJJ [TeV]


$m_{jj}$ = 7.5 TeV
(out of 13 TeV)
ATLAS
EXPERIMENT
Run: 302347
Event: 753275626
2016-06-18 18:41:48 CEST

# *Parting Thoughts*

# Standard lore should be revisited periodically

*It is your job to challenge your elders!*

Longitudinal boost invariance (i.e. $p_T$ & R vs. E & $\theta$)

Manifest energy-momentum conservation

Fixed jet radius

Unique assignment of hadrons to jet clusters

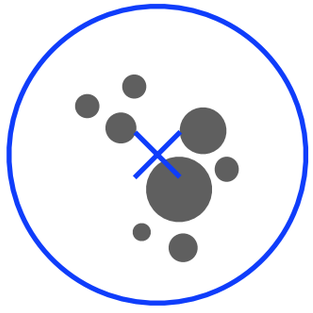
Unique jet reconstruction per event

Hierarchical clustering vs. global optimization

…

# Change the jet axis?

## Jet Momentum Axis



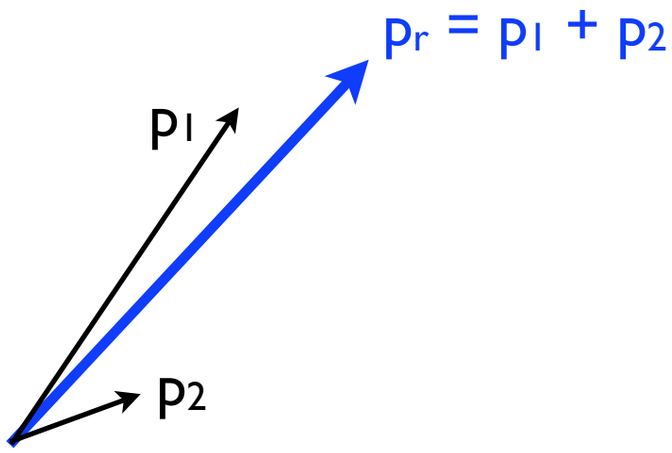
"Mean" ≈ minimizes sum of |distance|$^2$
(sensitive to outliers)

vs.

## Winner-Take-All Axis

"Median" ≈ Minimizes sum of |distance|
(insensitive to outliers)

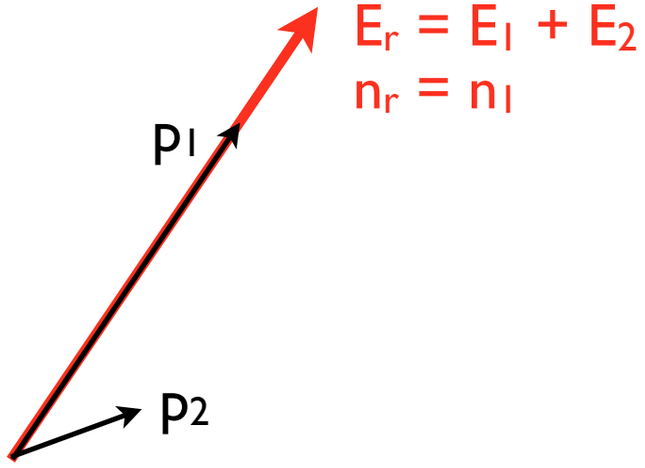
## E-scheme Recombination




⇒

One line of code
in FastJet 3.1

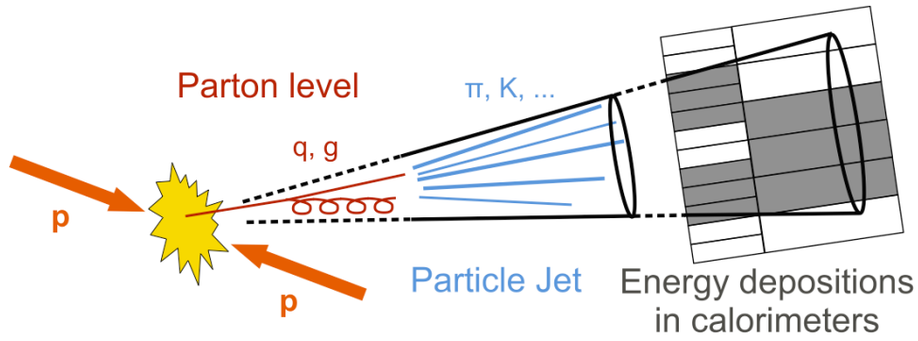
## WTA Recombination




[Bertolini, Chan, JDT, JHEP 2014; Larkoski, Neill, JDT, JHEP 2014; see also JDT, Van Tilburg, JHEP 2012]

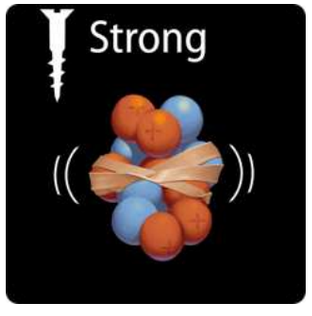
# A Jet/QCD Renaissance

*c. 2008–present*




## LHC

| | *vs. Tevatron* |
|---|---|
| Higher Energy | *≈ x3.5–7* |
| Higher Luminosity | *≈ x10–20* |
| Finer Segmentation | *≈ x5* |




## Theoretical Progress

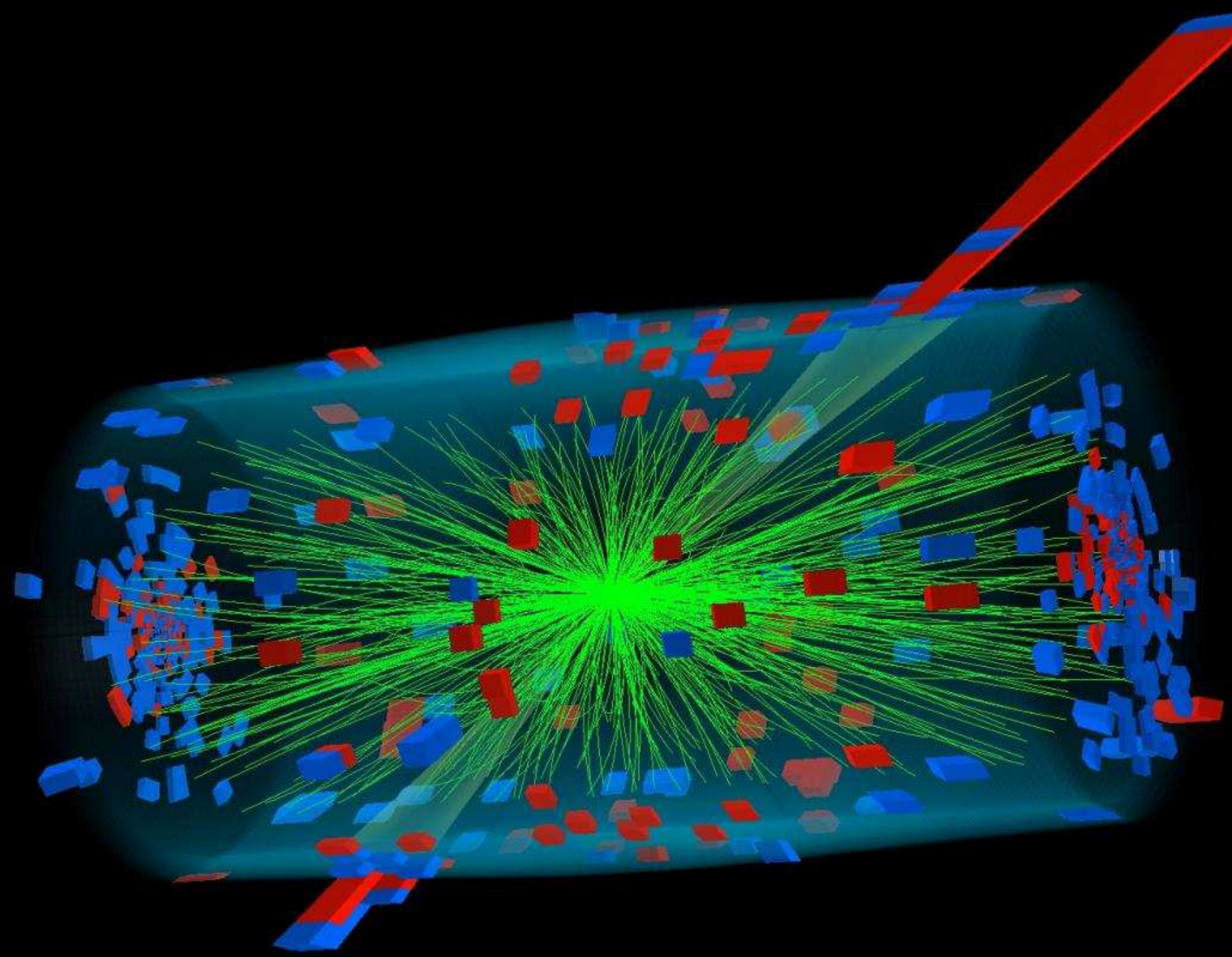
New Jet Algorithms

Loop/Leg/Log Explosion

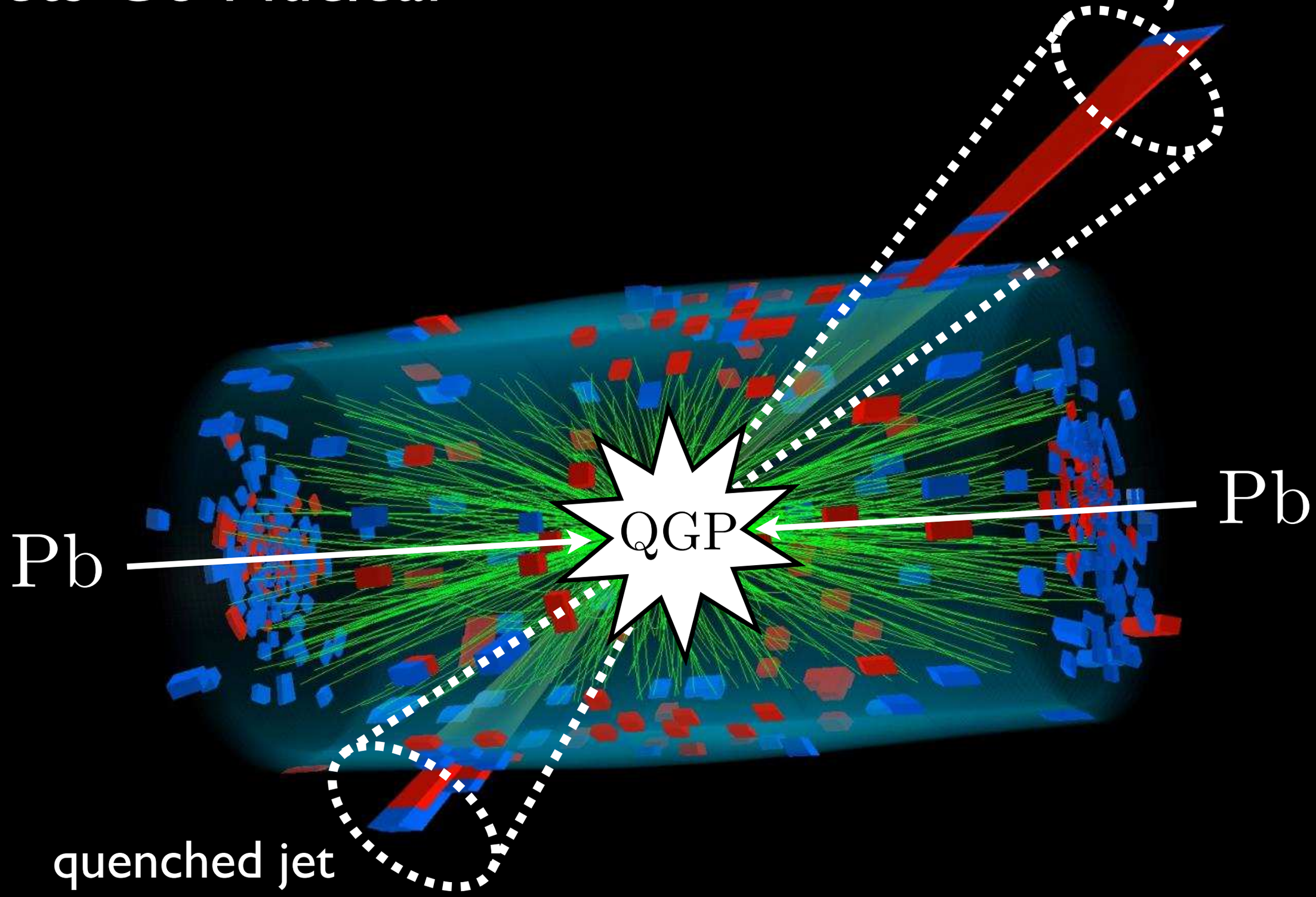
Jet Substructure

[Anti-kT: Cacciari, Salam, Soyez, 2008; see also Delsart, 2006]
[BDRS: Butterworth, Davison, Rubin, Salam, 2008; see also Seymour, 1991, 1994]

# Jets Go Nuclear

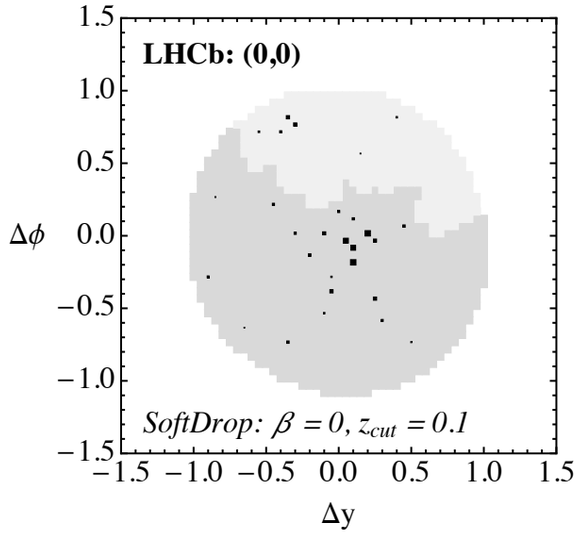


# Jets Go Nuclear

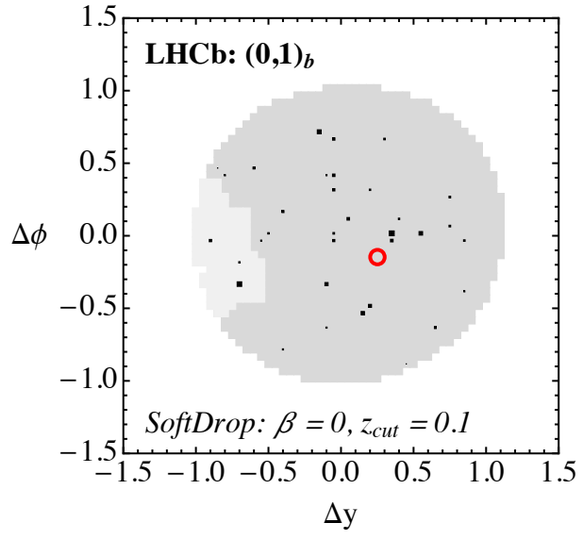


# Jets Get More Flavorful

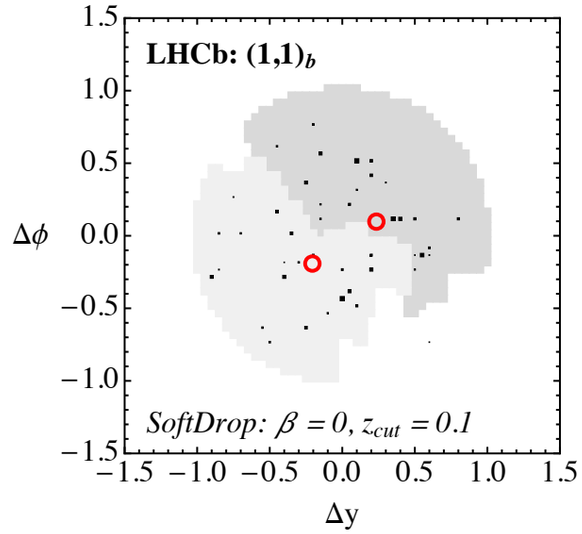
$g \to gg$

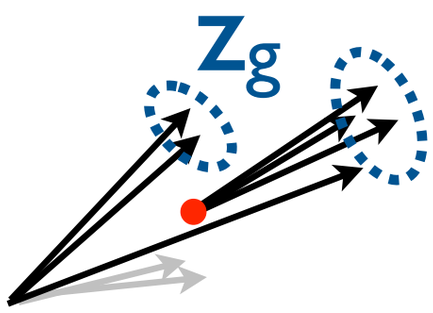


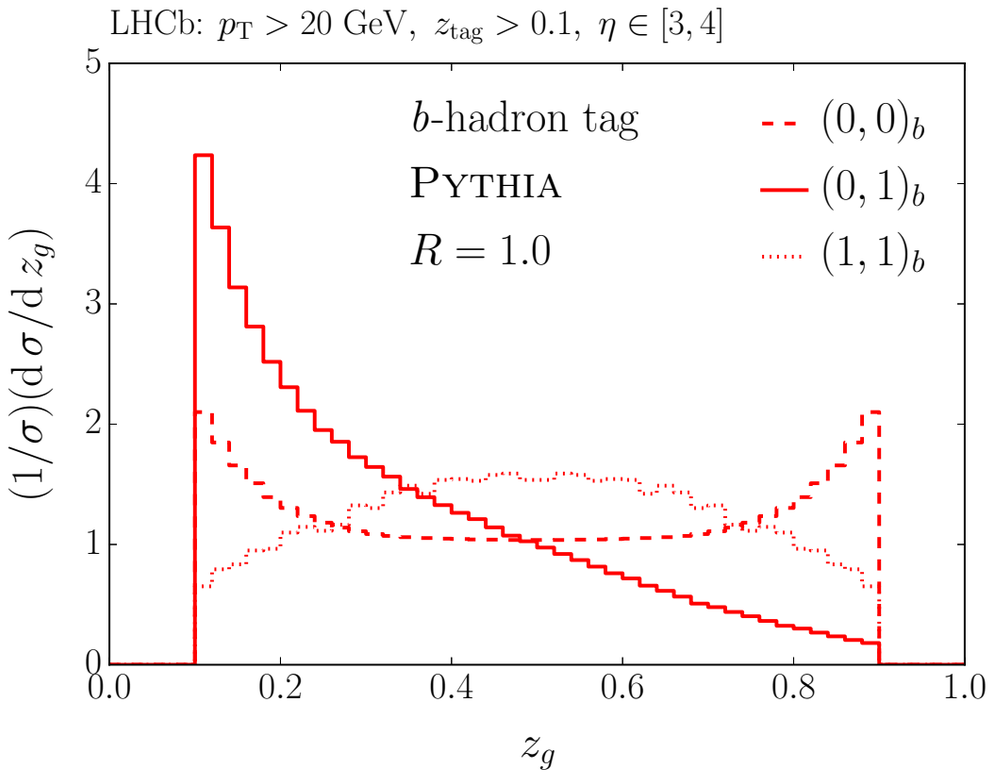

$b \to bg$

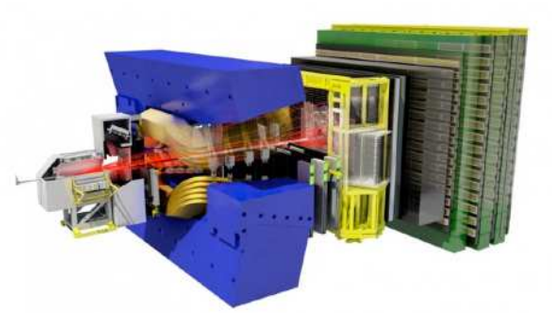



$g \to b\bar{b}$



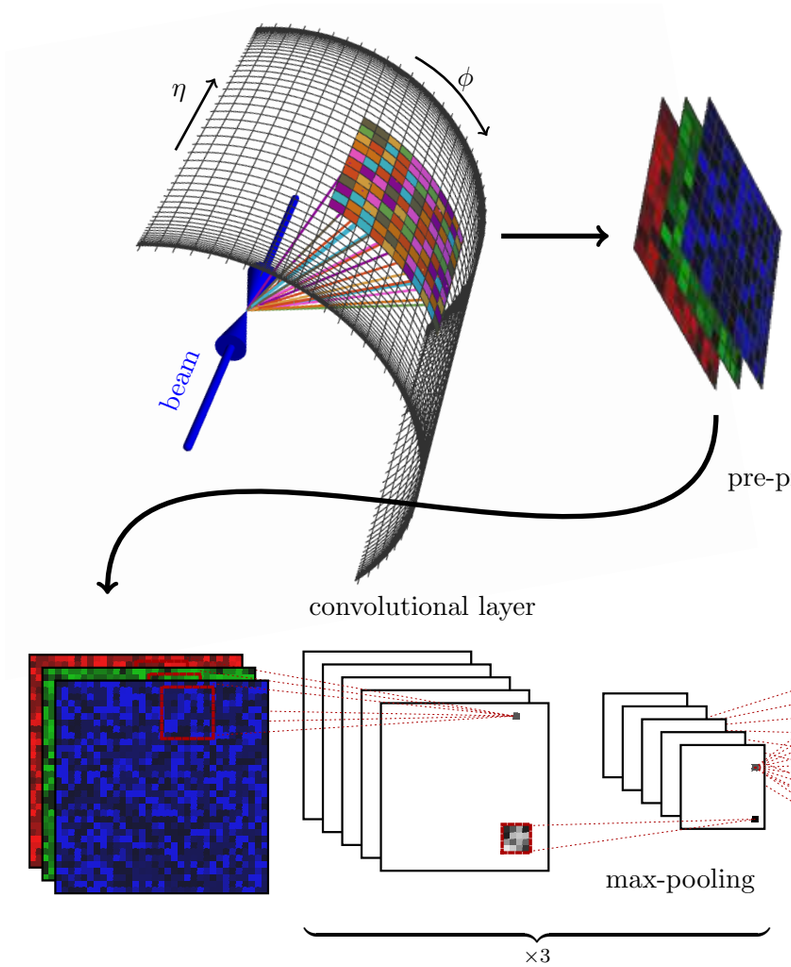



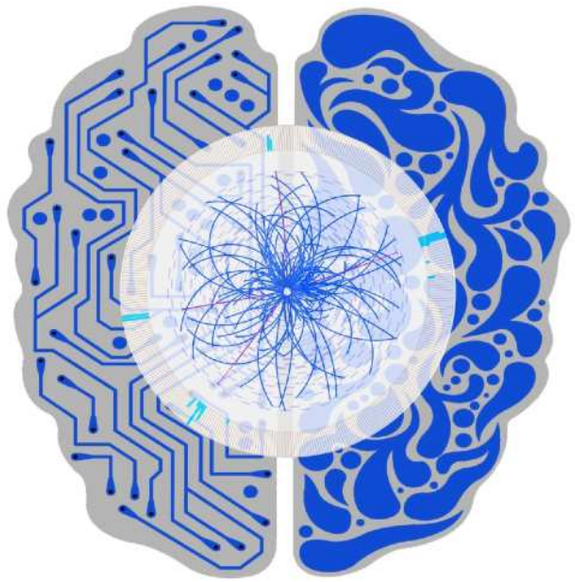
[Ilten, Rodd, JDT, Williams, PRD 2017]

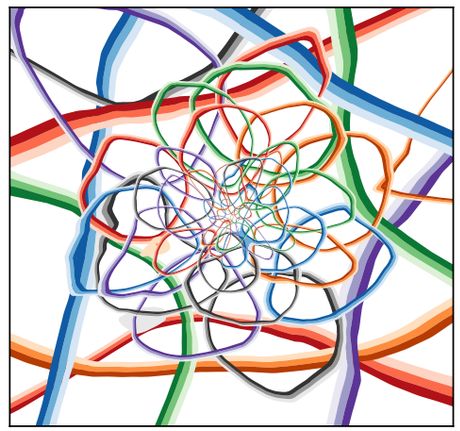
# The Rise of Machine Learning

*c. 2015–present*

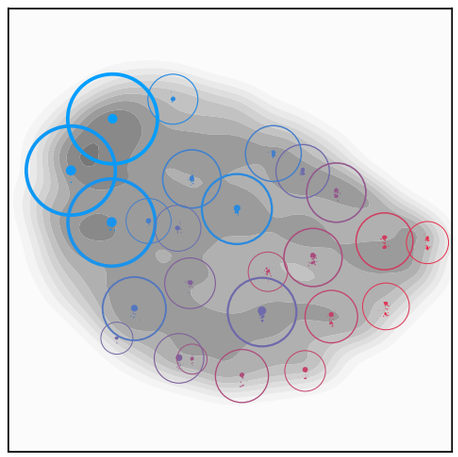



[figure from Komiske, Metodiev, Schwartz, JHEP 2017;
reviews in Larkoski, Moult, Nachman, arXiv 2017; Guest, Cranmer, Whiteson, ARNPS 2018;
ML4Jets Workshop Series, next iteration at NYU, January 2019]

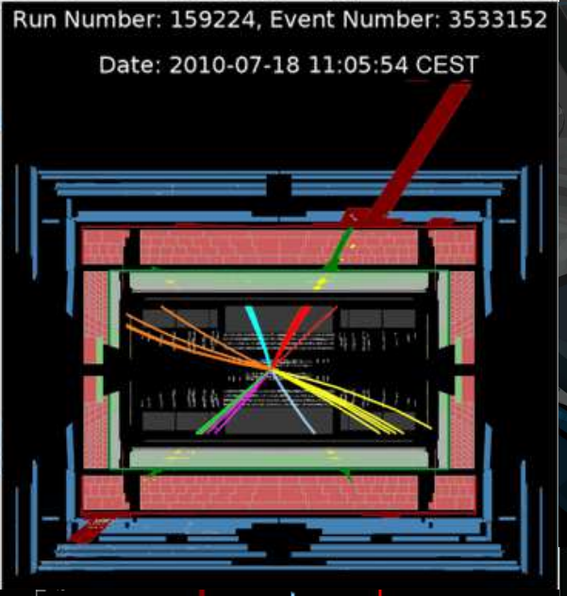


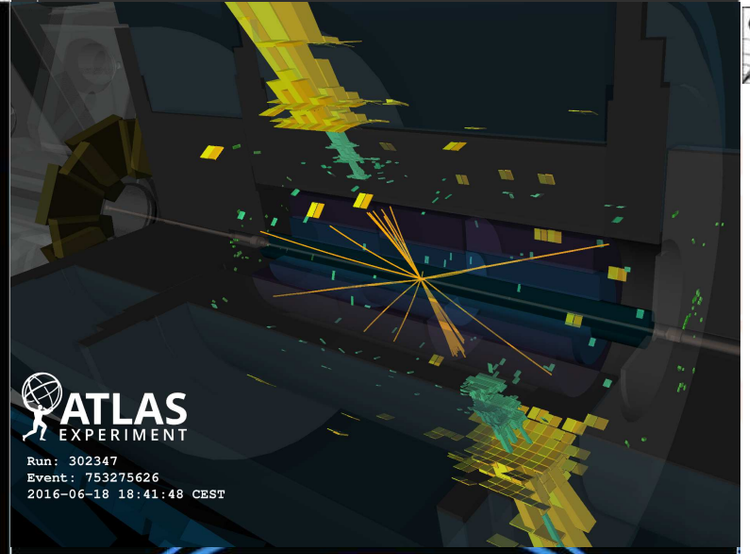


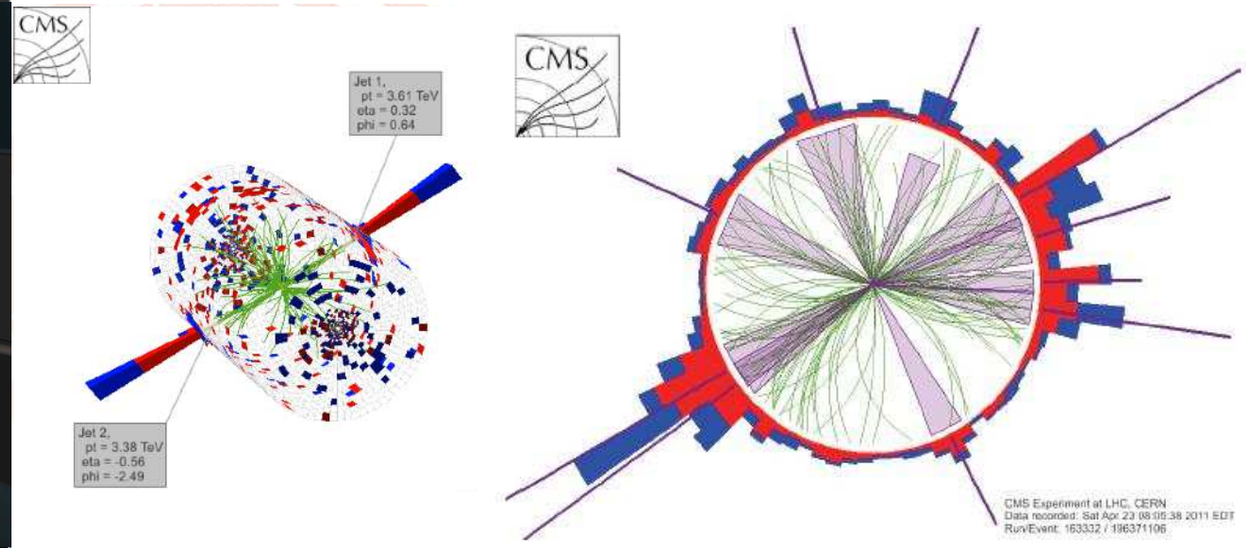
[Komiske, Metodiev, JDT, JHEP 2019]

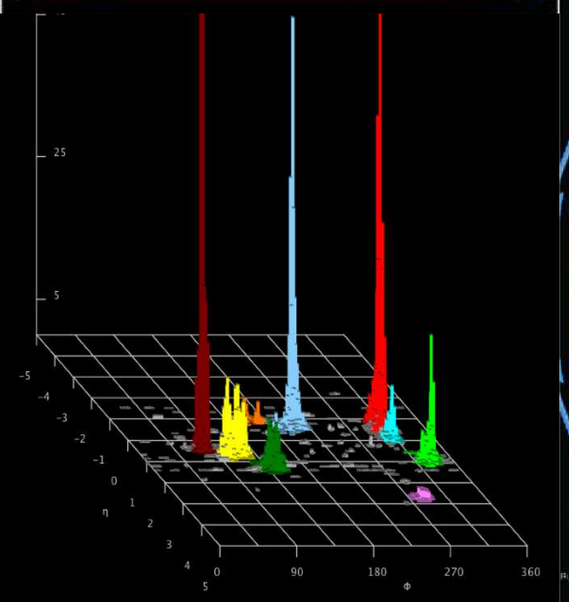


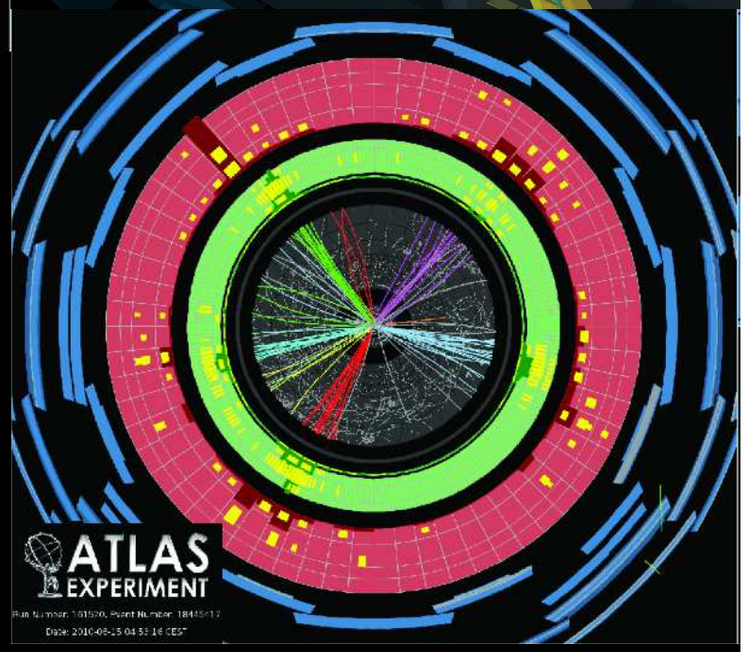
[Komiske, Mastandrea, Metodiev, Naik, JDT, arXiv 2019;
based on Komiske, Metodiev, JDT, PRL 2019]

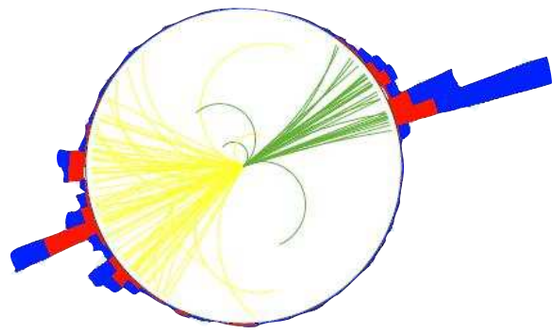

Run Number: 159224, Event Number: 3533152
Date: 2010-07-18 11:05:54 CEST

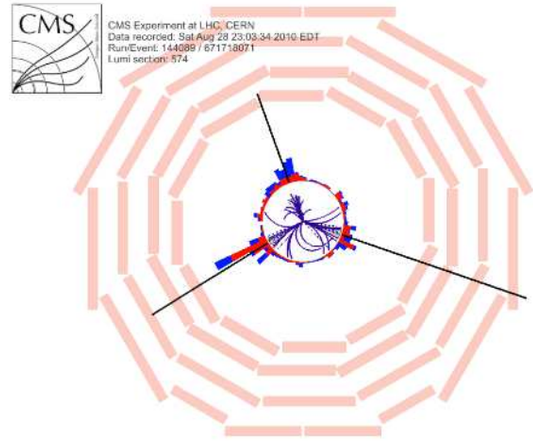

ATLAS
EXPERIMENT
Run: 302347
Event: 753275626
2016-06-18 18:41:48 CEST

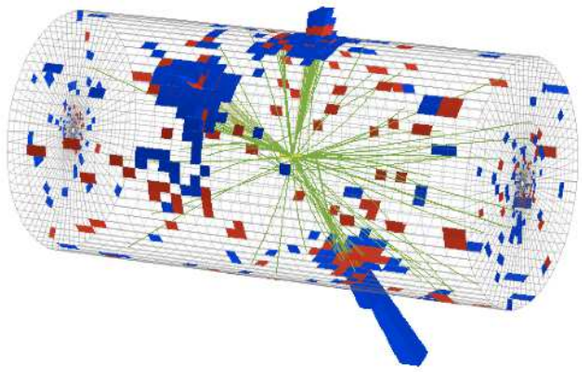

CMS
Jet 1,
pt = 3.61 TeV
eta = 0.32
phi = 0.64
Jet 2,
pt = 3.38 TeV
eta = -0.56
phi = -2.49
CMS
CMS Experiment at LHC, CERN
Data recorded: Sat Apr 23 08:05:38 2011 EDT
Run/Event: 163332 / 196371106

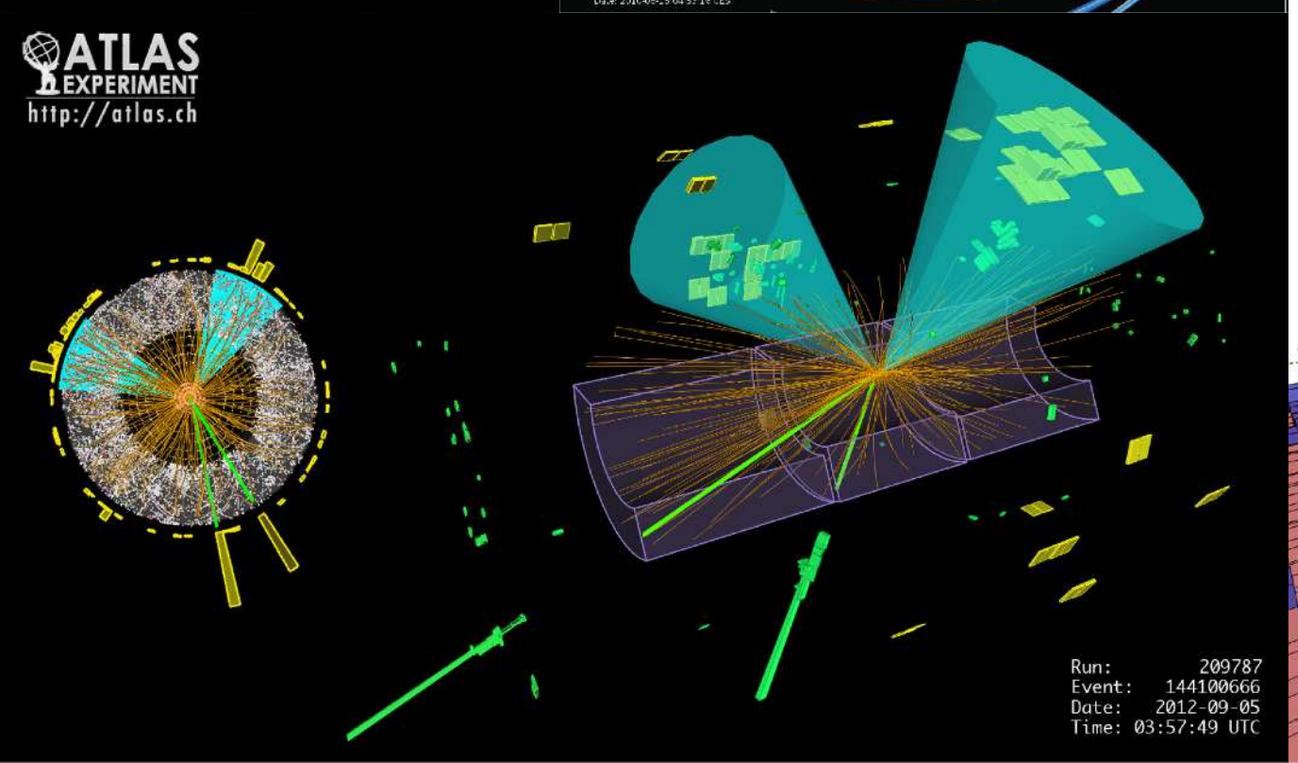

25
5
-5
-4
-3
-2
-1
0
1
2
3
4
η
0
90
180
270
360
Φ

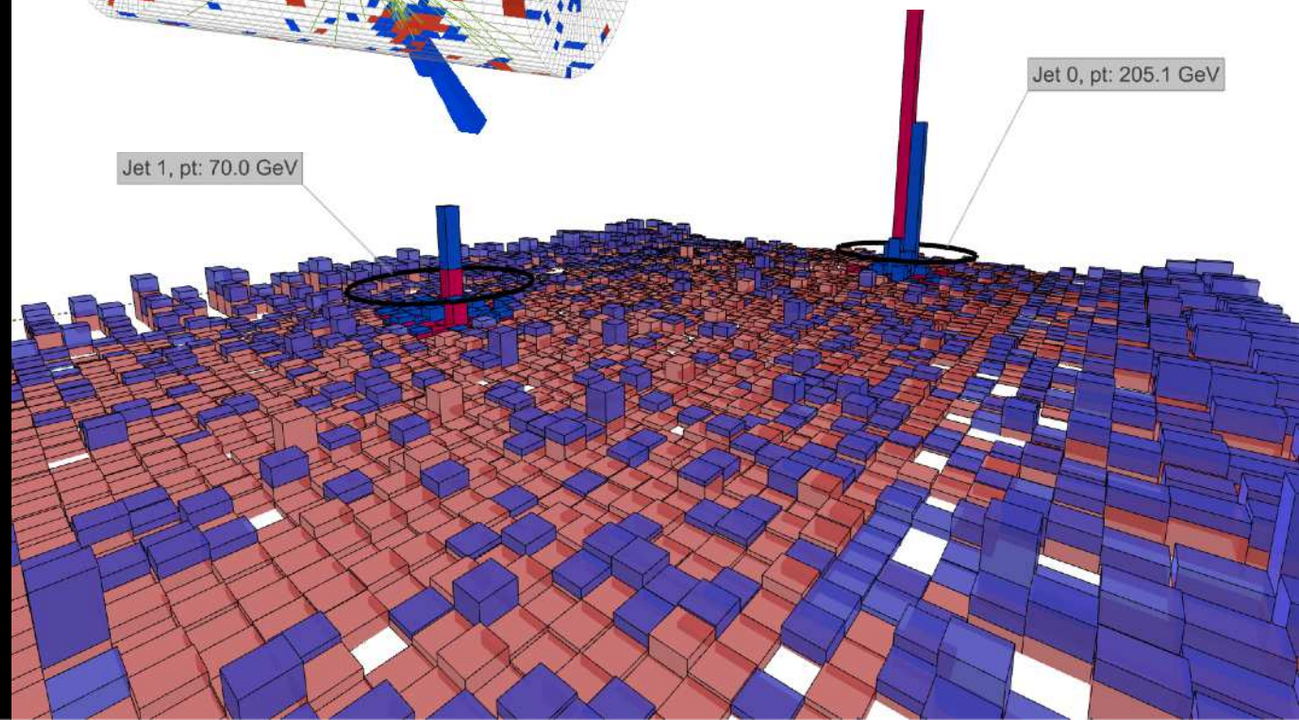

ATLAS
EXPERIMENT




CMS
CMS Experiment at LHC, CERN
Data recorded: Sat Aug 28 23:03:34 2010 EDT
Run/Event: 144089 / 671718071
Lumi section: 574




ATLAS
EXPERIMENT
http://atlas.ch
Run: 209787
Event: 144100666
Date: 2012-09-05
Time: 03:57:49 UTC


Jet 0, pt: 205.1 GeV
Jet 1, pt: 70.0 GeV