



Shout out to Konstantin Matchev, Ann Lee, and Harrison Prosper who are coorganizing a workshop at the Aspen Center for Physics on “Interplay of Fundamental Physics and Machine Learning”!

Weak Supervision for the Strong Force

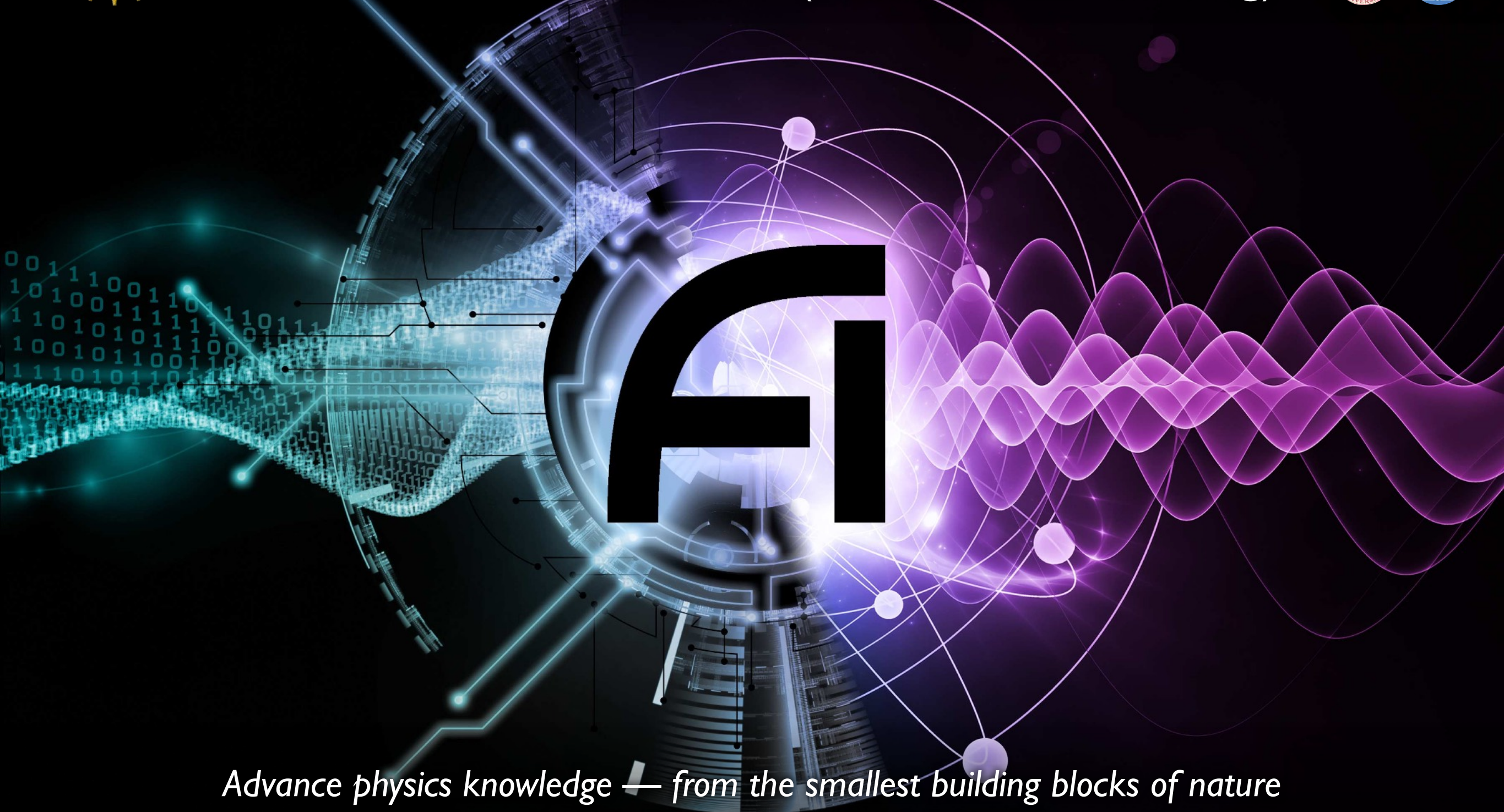
Jesse Thaler



Physics Meets ML — June 1, 2022



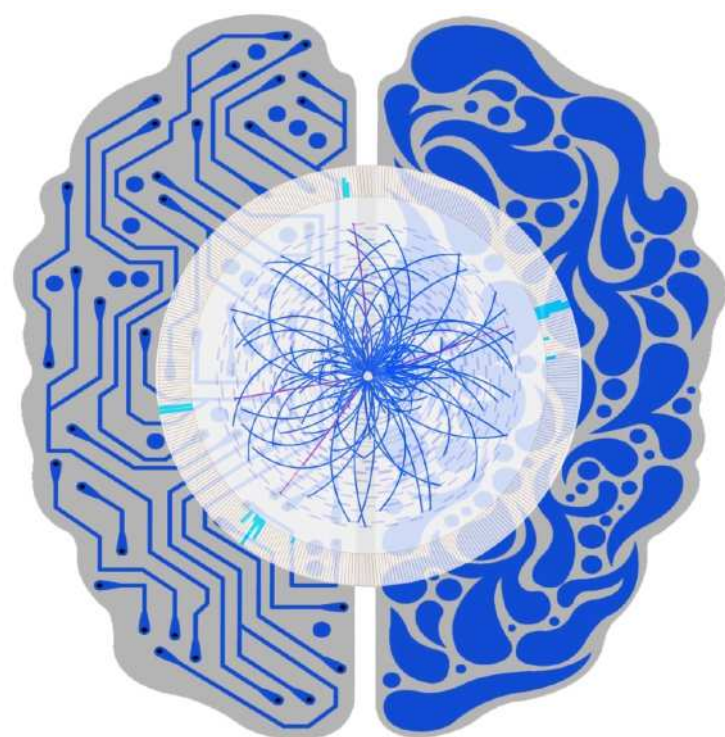
The NSF AI Institute for Artificial Intelligence and Fundamental Interactions (IAIFI /aI-faI/ iaifi.org)



Advance physics knowledge — from the smallest building blocks of nature to the largest structures in the universe — and galvanize AI research innovation



The NSF AI Institute for Artificial Intelligence and Fundamental Interactions (IAIFI /aɪ-faɪ/ iaifi.org)



Infuse physics intelligence into artificial intelligence

Machine learning that incorporates first principles, best practices, and domain knowledge from fundamental physics

Advance physics knowledge — from the smallest building blocks of nature to the largest structures in the universe — and galvanize AI research innovation

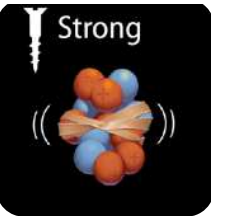


FI IAIFI

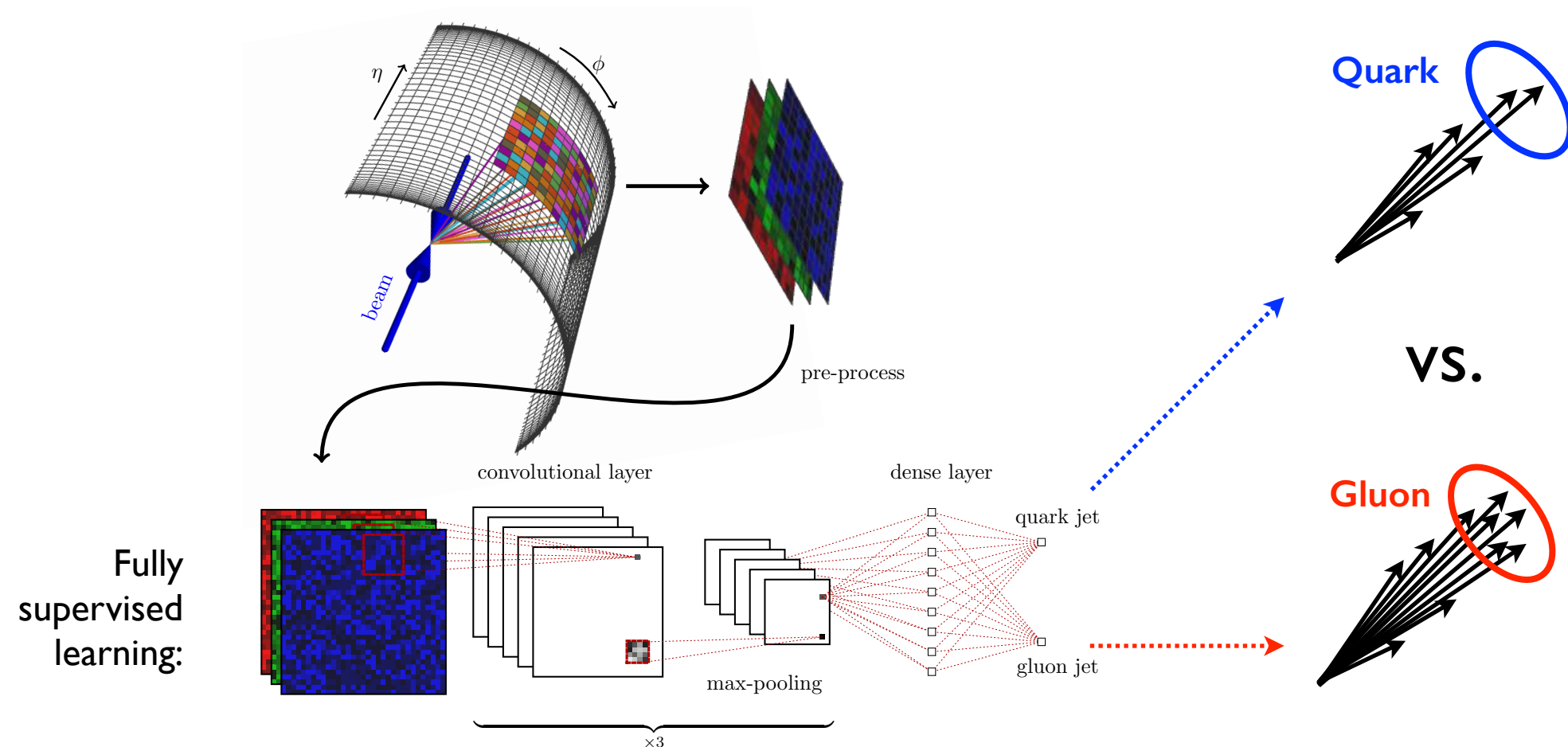
Summer School & Workshop

August 1–9, 2022

Weak Supervision for the Strong Force



An IAFI origin story: Patrick and Eric come to my office at MIT...

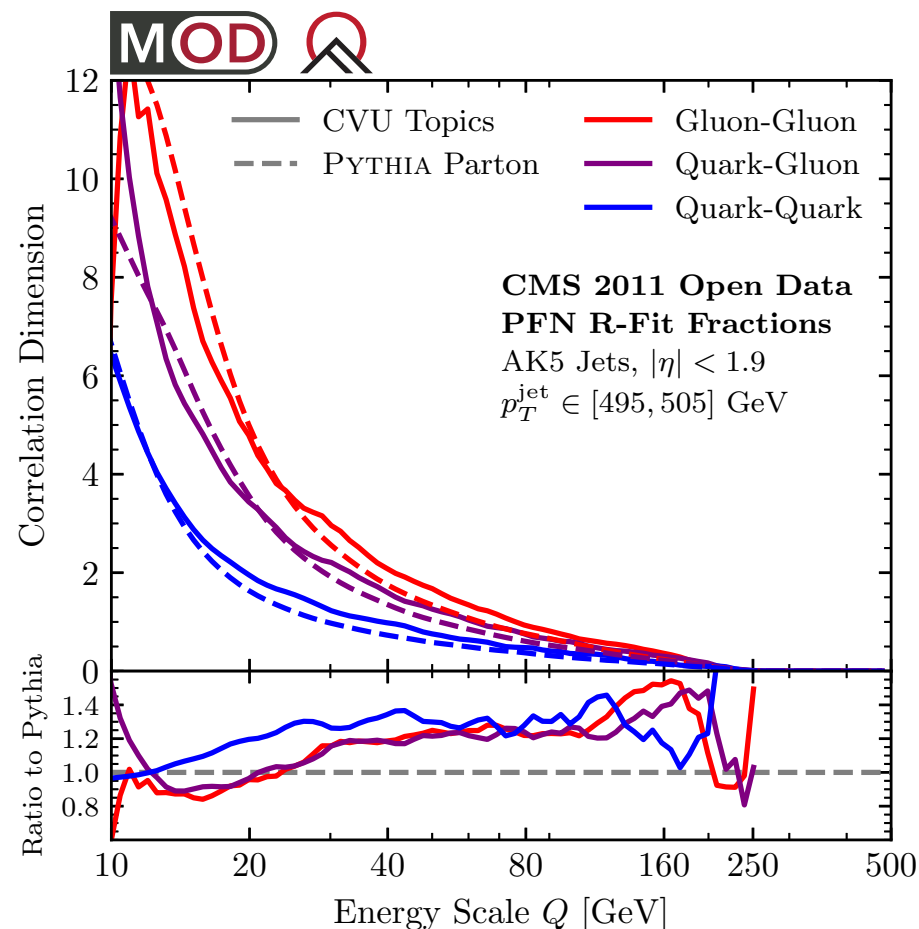


“What do you mean by *quarks* and *gluons*?”

[Komiske, Metodiev, Schwartz, JHEP 2017]



Weak Supervision for the Strong Force



Five years later,
I finally **know an answer!**

*Required weak supervision, topic modeling,
permutation-invariant networks,
simulation-based inference, optimal transport, ...*

[Komiske, Kryhin, JDT, [arXiv 2022](#)]

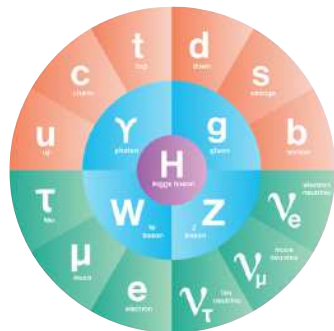


“What do you mean by **quarks** and **gluons**?”

[Komiske, Metodiev, Schwartz, [JHEP 2017](#)]



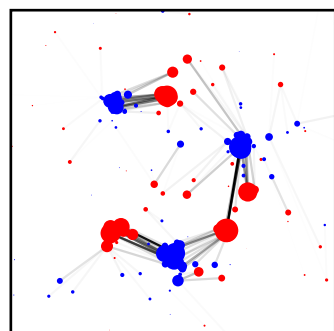
Outline



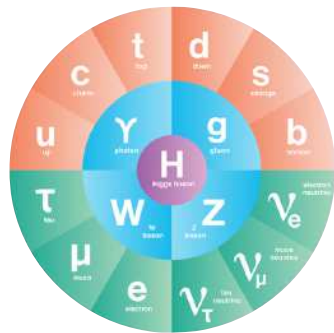
2017: A Quark/Gluon Conundrum



2018: Leveraging Weak Supervision



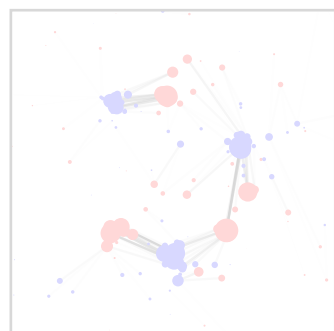
2019-2022: The Strong Force Revisited



2017: A Quark/Gluon Conundrum

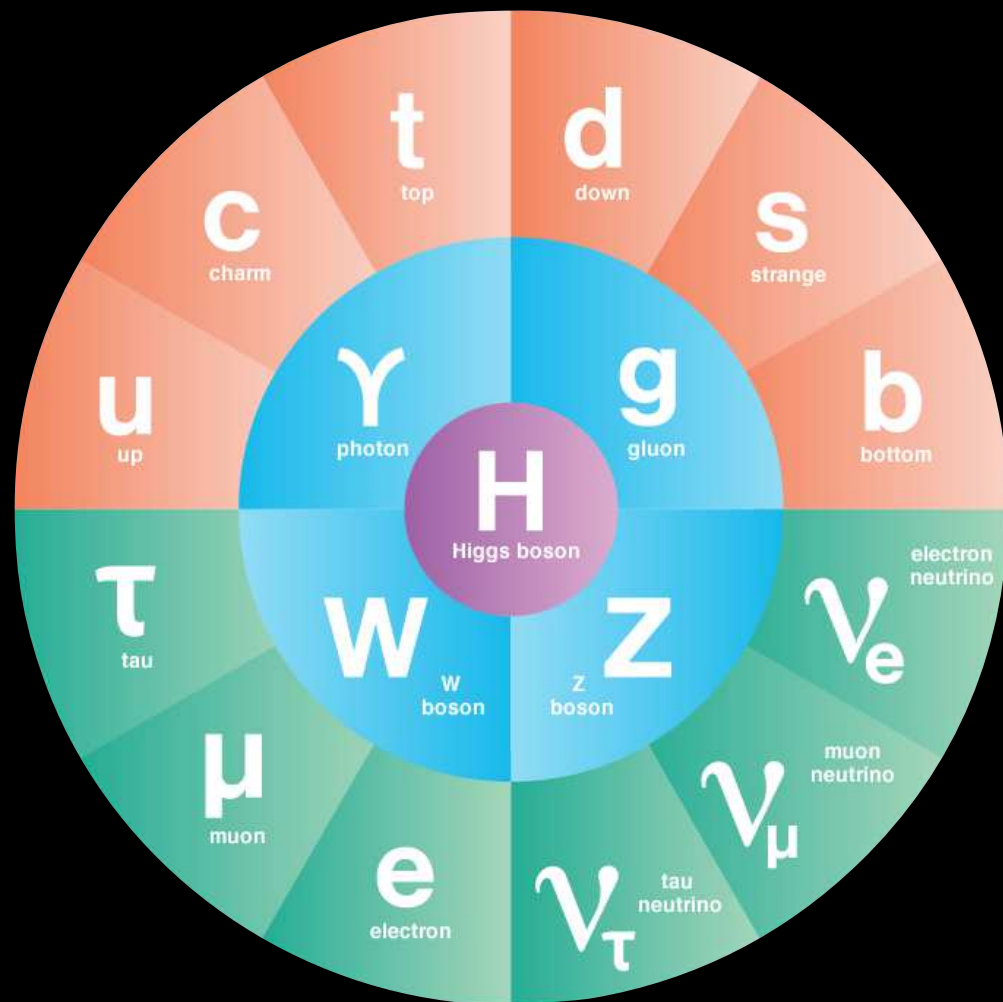


2018: Leveraging Weak Supervision

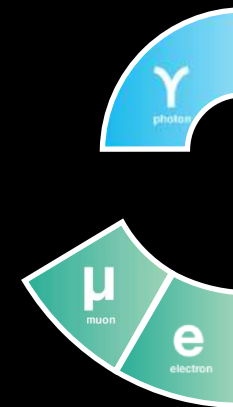
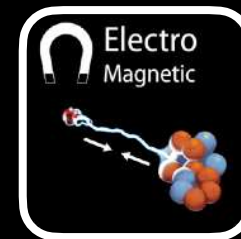
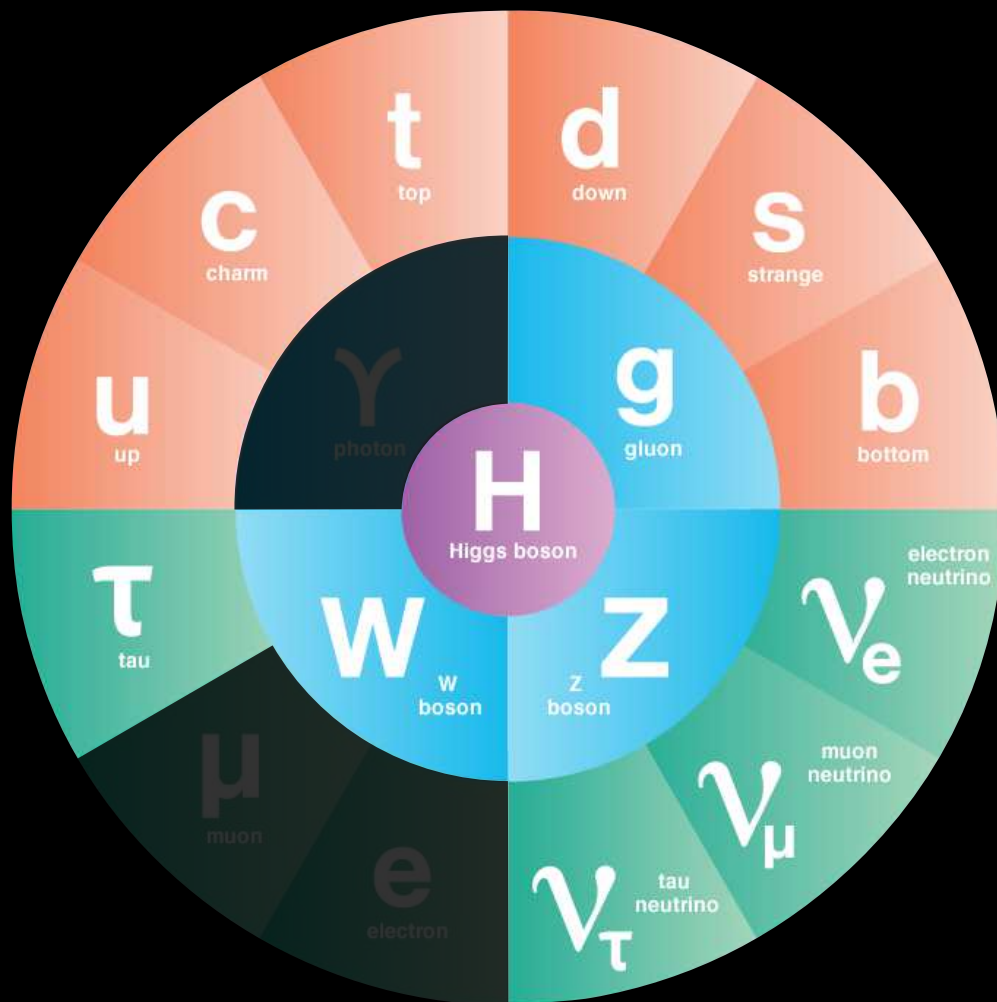


2019-2022: The Strong Force Revisited

Particle Physics 101



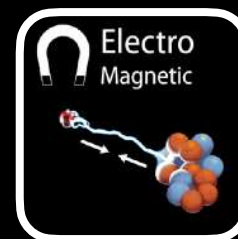
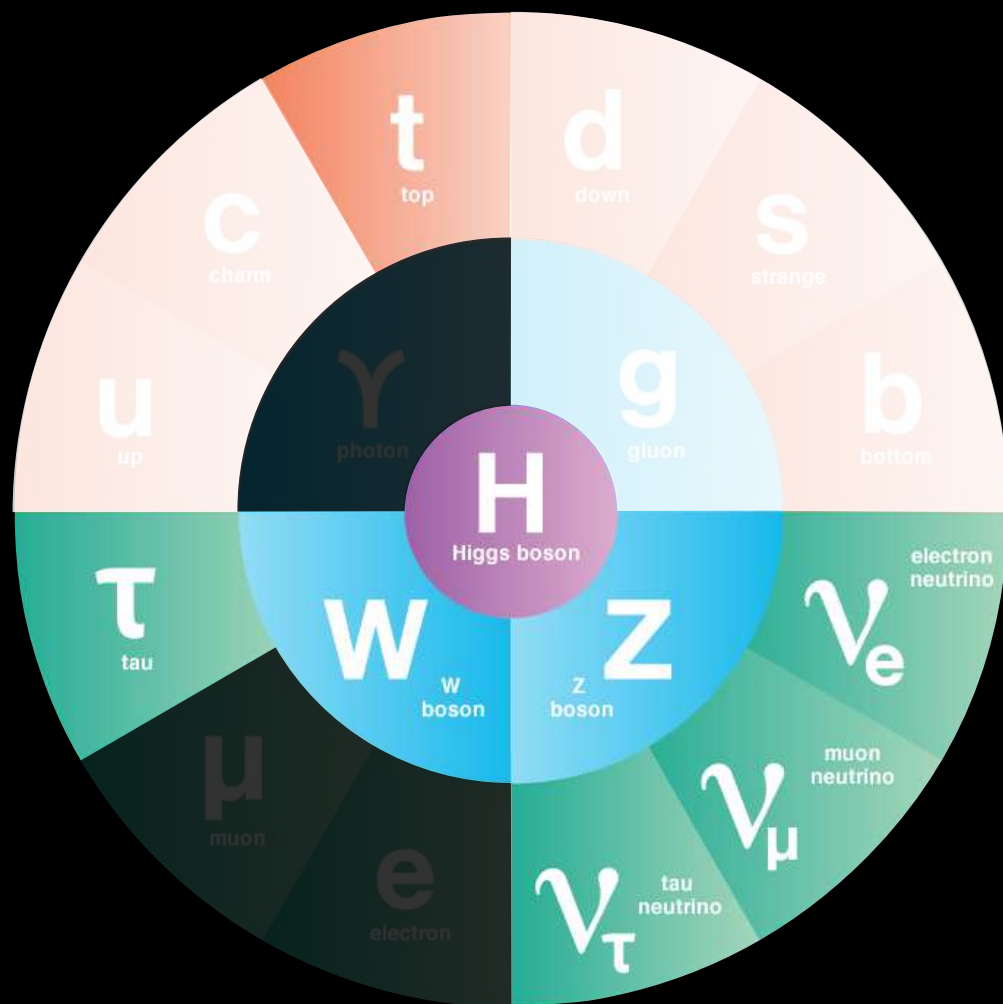
Particle Physics 101



γ photon
 $e^{\pm}$ electron
 $\mu^{\pm}$ muon

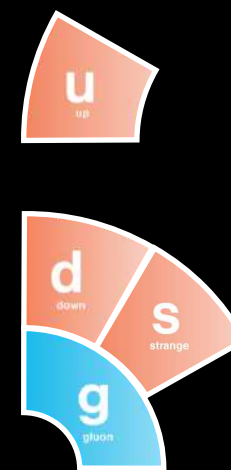
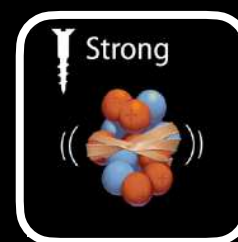
elementary

Particle Physics 101



Strong Force
 $\Rightarrow$ *Confinement*

Quarks
&
Gluons



γ photon

$e^{\pm}$ electron

$\mu^{\pm}$ muon

$\pi^{\pm}$ pion

$K^{\pm}$ kaon

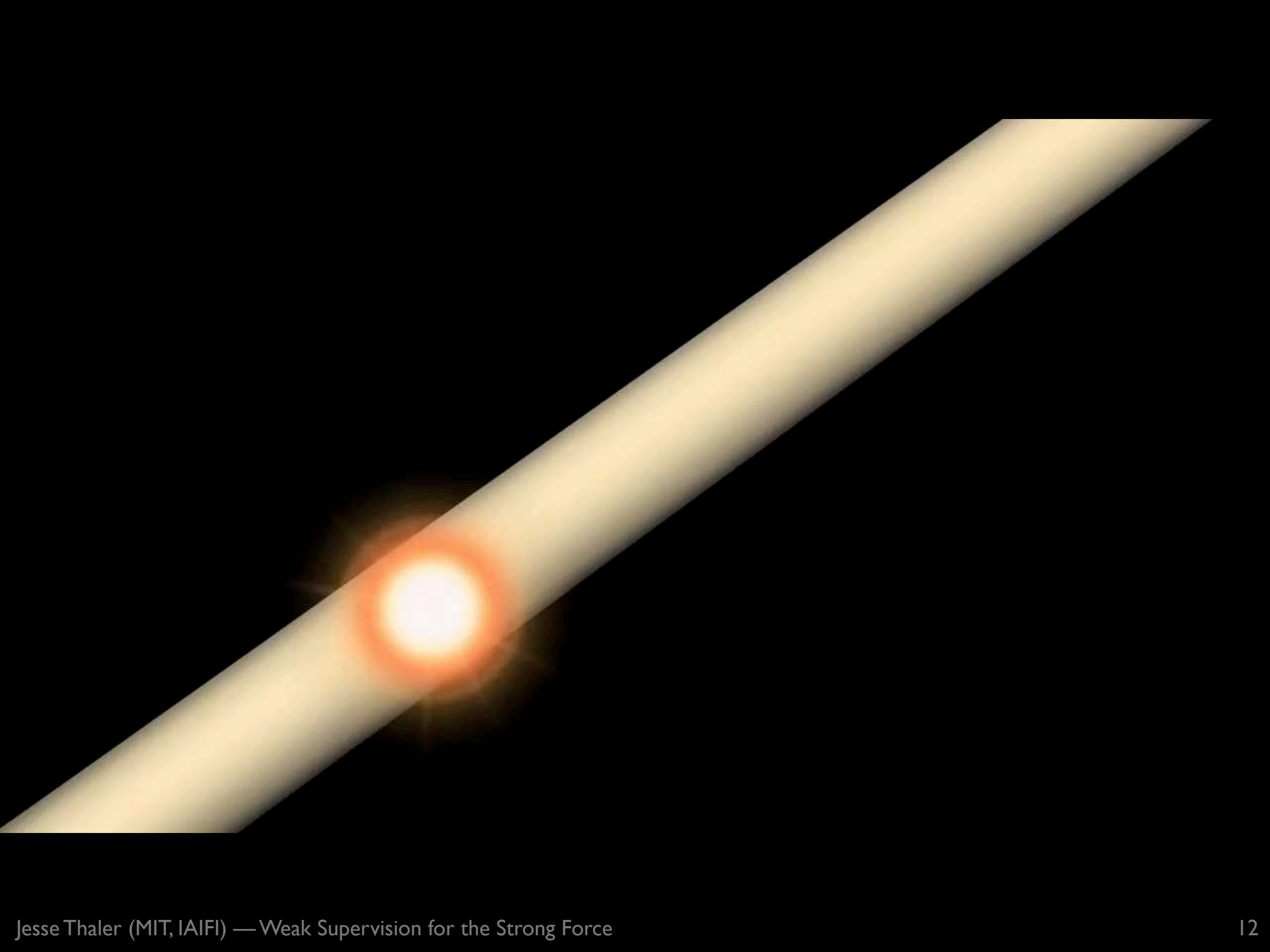
K_L^0 K-long

$p/\bar{p}$ proton

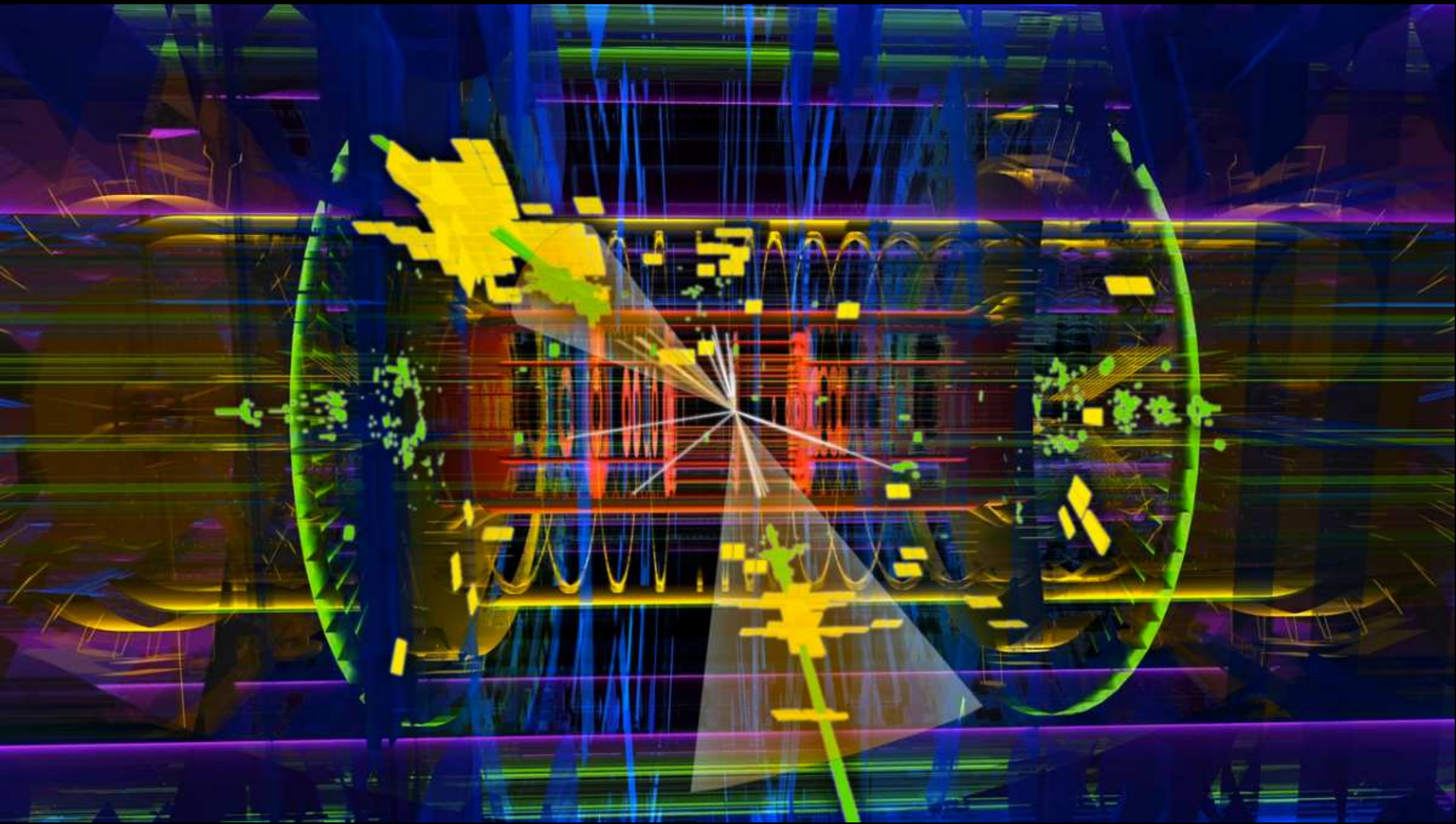
$n/\bar{n}$ neutron

elementary

composite

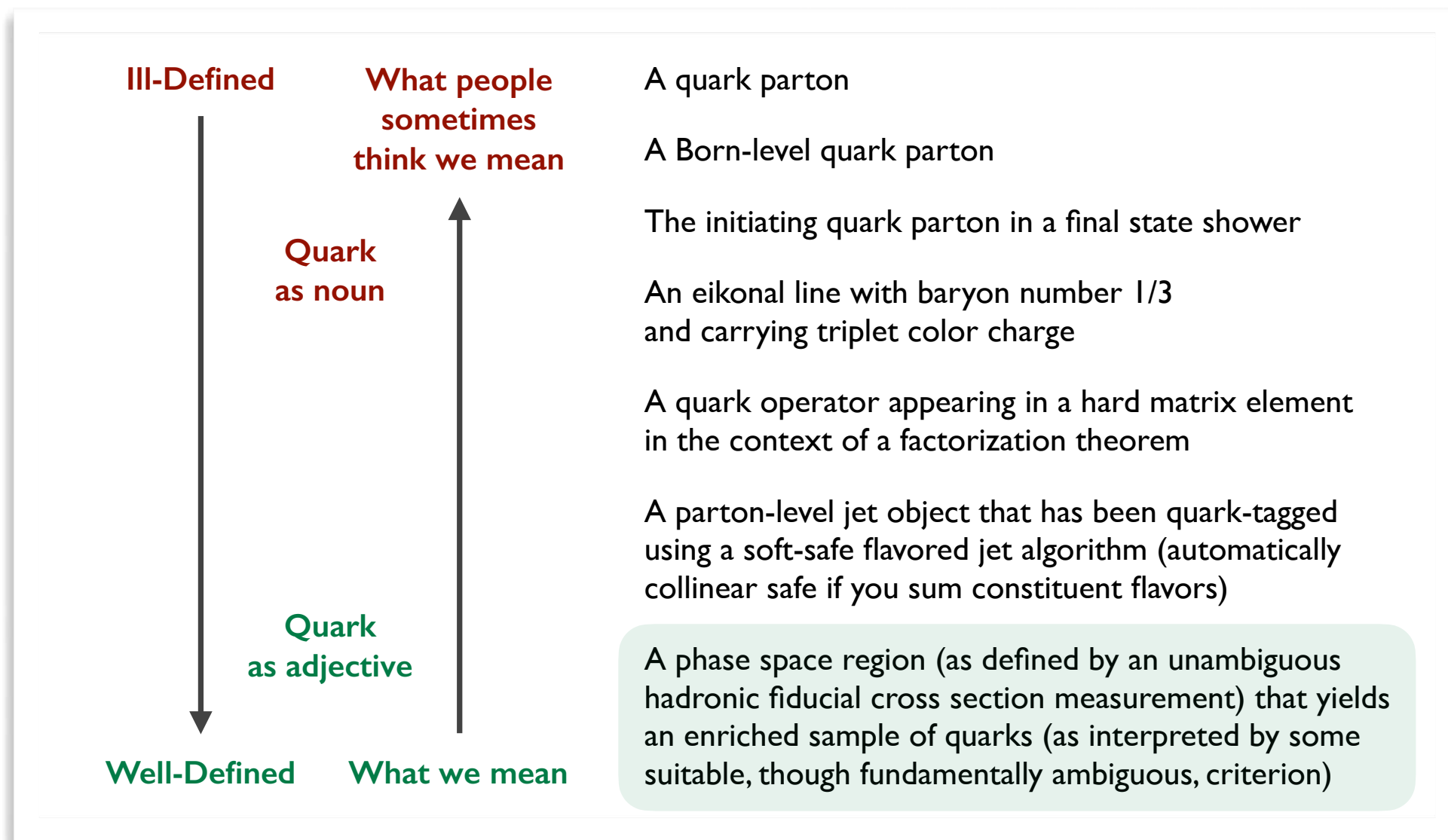






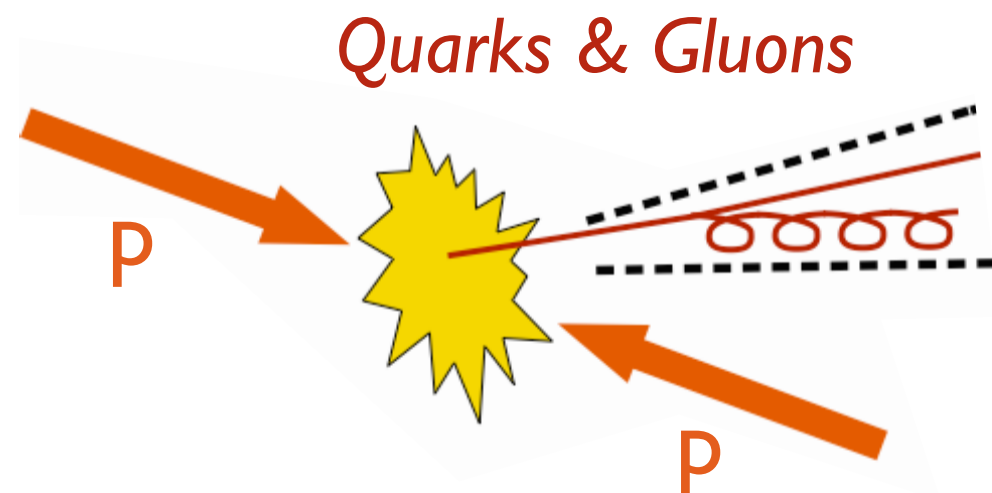
What is a “Quark Jet” vs. “Gluon Jet”?

Quark (color triplet) vs. Gluon (color octet)
But jet constituents are color-singlet hadrons!



[Gras, Hoeche, Kar, Larkoski, Lönnblad, Plätzer, Siódmok, Skands, Soyez, JDT, [JHEP 2017](#); citing long history going back at least to Nilles, Streng, [PRD 1981](#)]

Emphasizing the Challenge



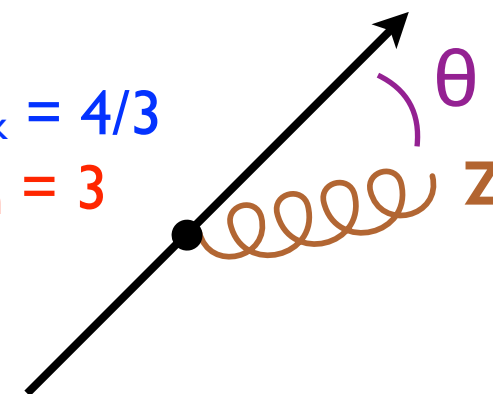
Quarks and gluons are primarily distinguished by their color charge

Altarelli-Parisi Splitting

Core prediction of QCD

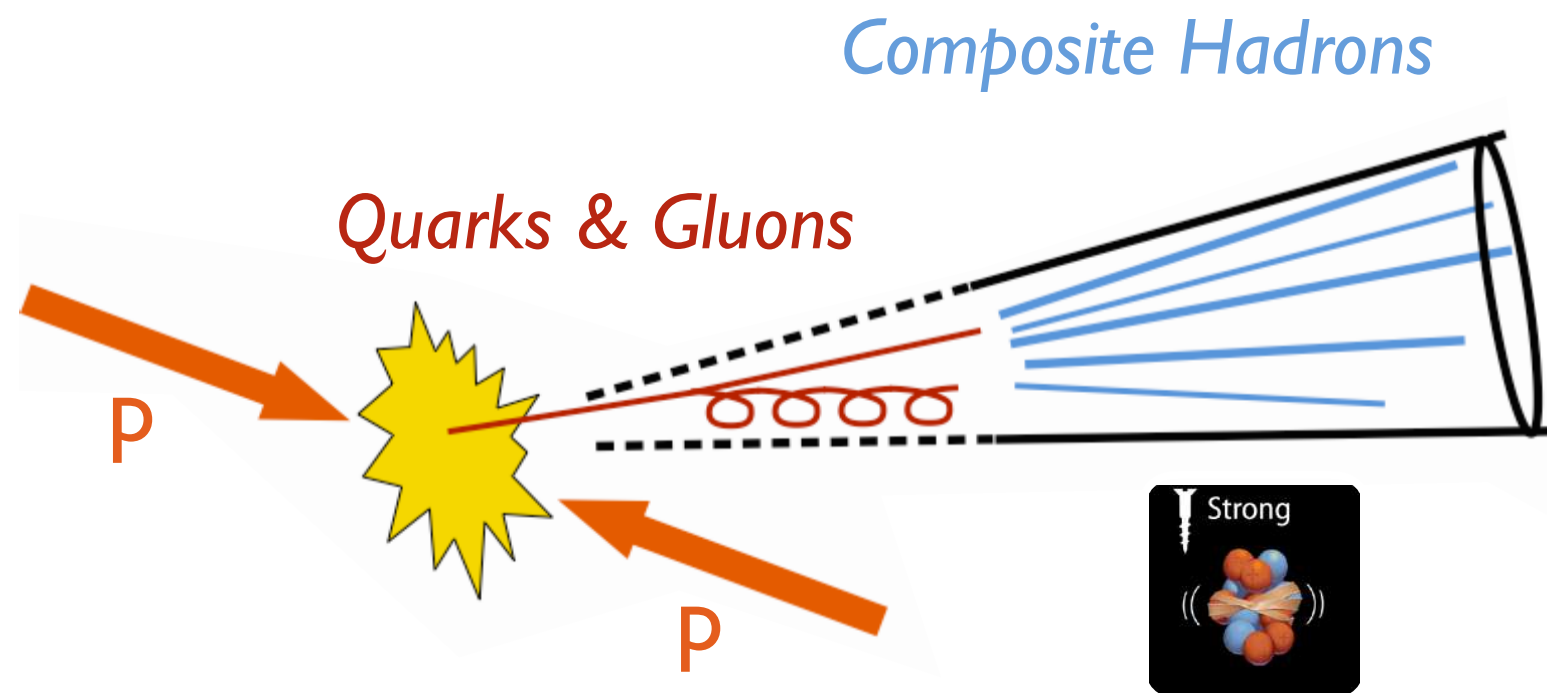
$$C_{\text{quark}} = 4/3$$

$$C_{\text{gluon}} = 3$$



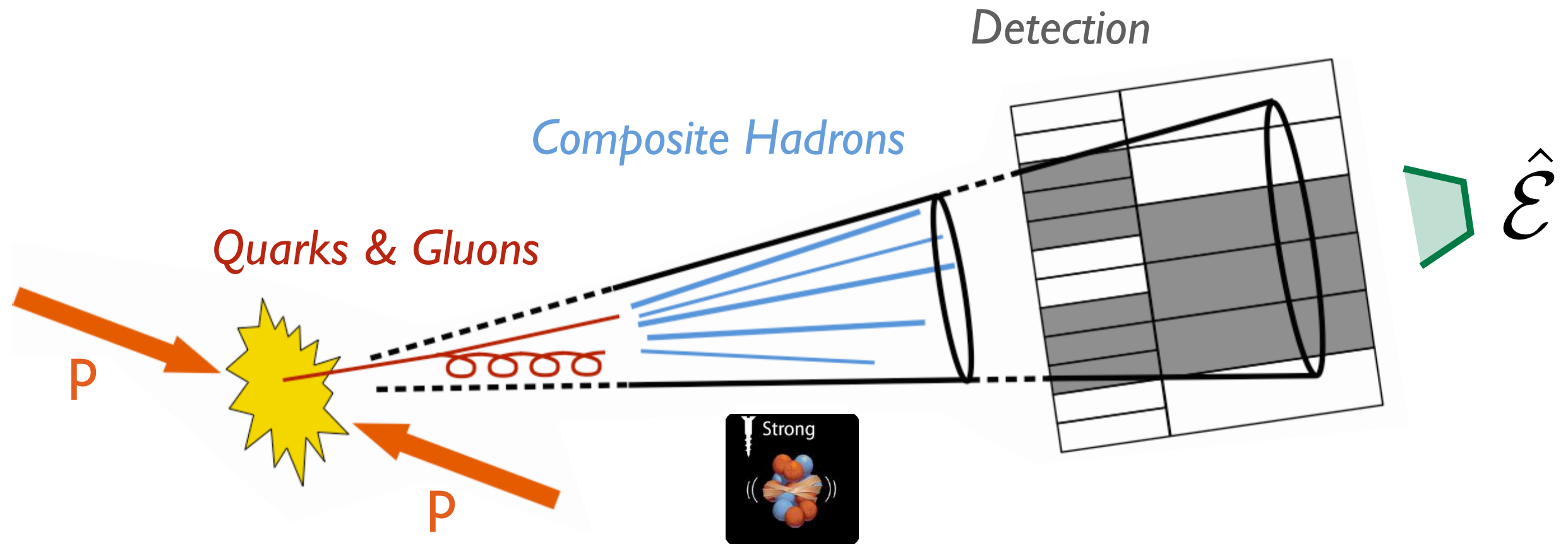
$$dP_{i \rightarrow ig} \simeq \frac{2\alpha_s}{\pi} C_i \underbrace{\frac{d\theta}{\theta}}_{\text{Collinear}} \underbrace{\frac{dz}{z}}_{\text{Soft}}$$

Emphasizing the Challenge



Emphasizing the Challenge

Theory



Energy Flow:

Robust to hadronization and detector effects...
...but blind to direct **quark**/**gluon** information

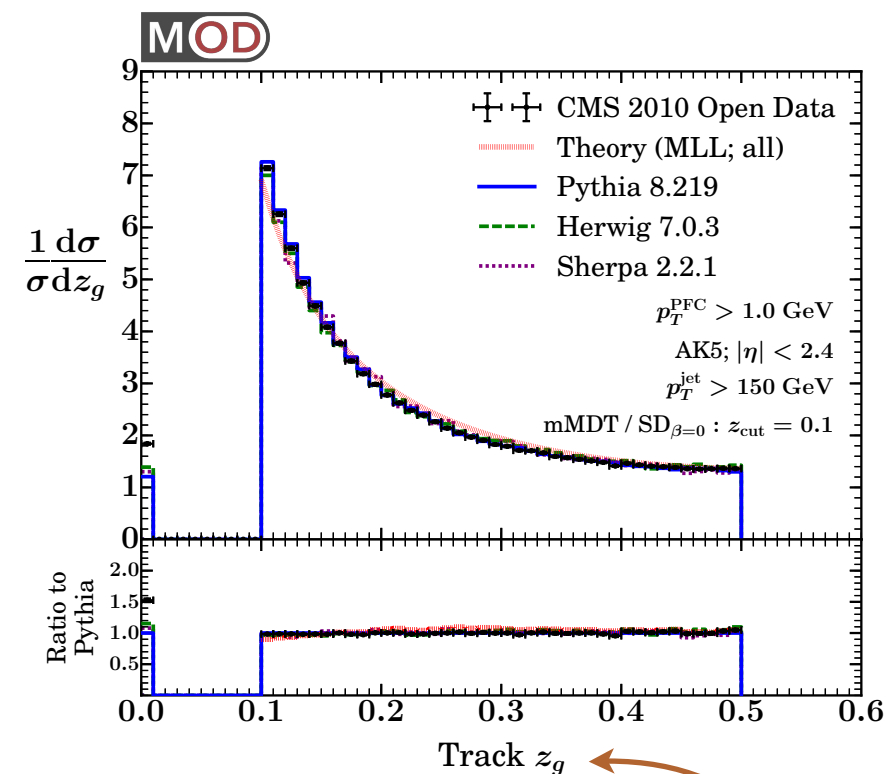
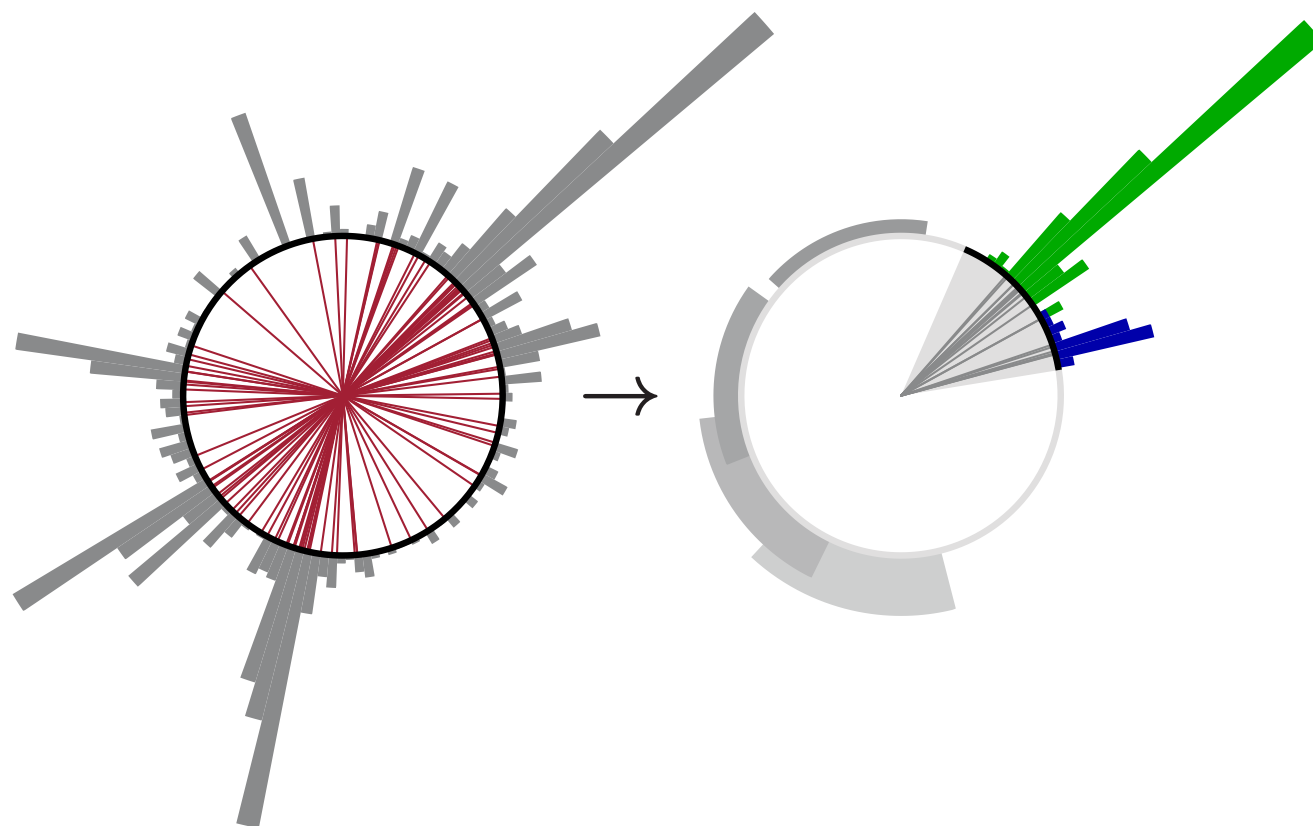
$$\hat{\mathcal{E}} \simeq \lim_{t \rightarrow \infty} \hat{n}_i T^{0i}(t, vt\hat{n})$$

(more about this in backup)

[see e.g. Sveshnikov, Tkachov, [PLB](#) 1996; Hofman, Maldacena, [JHEP](#) 2008; Mateu, Stewart, JDT, [PRD](#) 2013; Belitsky, Hohenegger, Korchemsky, Sokatchev, Zhiboedov, [PRL](#) 2014; Chen, Moul, Zhang, Zhu, [PRD](#) 2020]

Color-Blind Tests of QCD

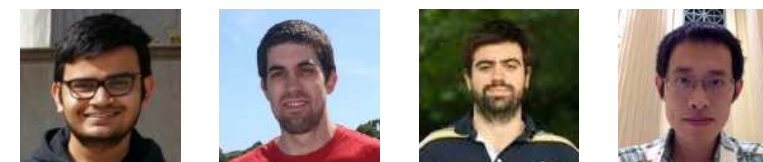
First ever analysis using public LHC data!

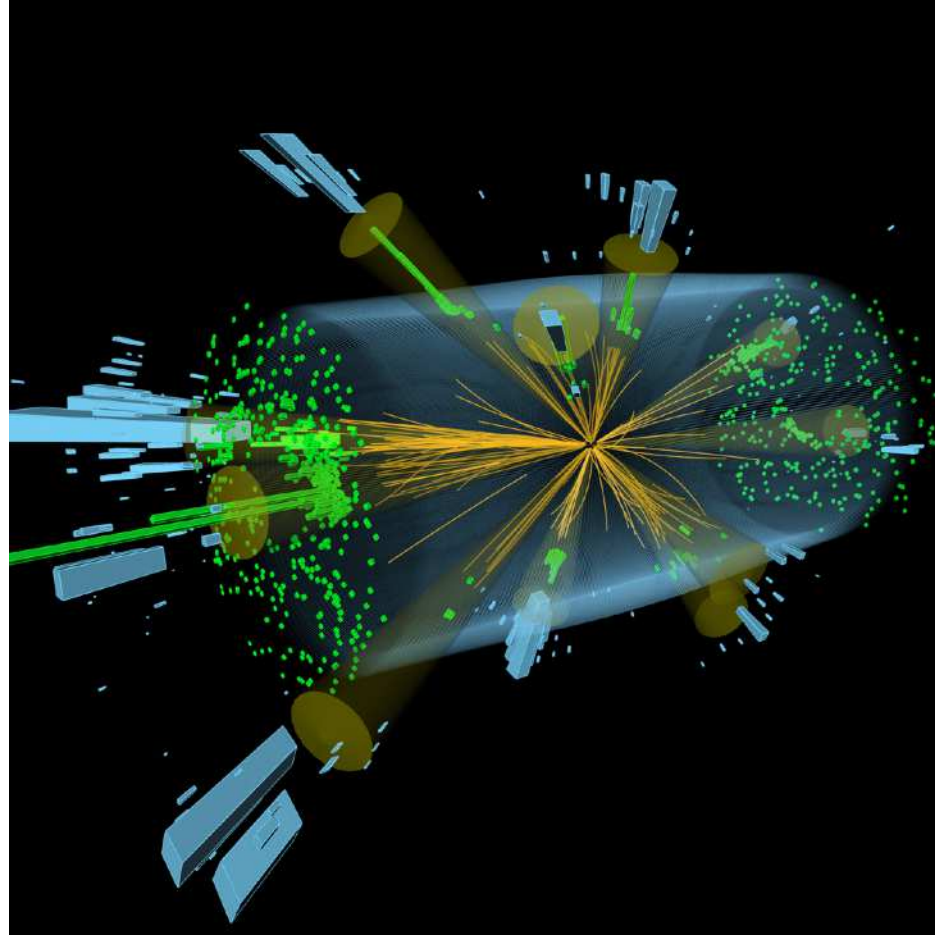


$$dP_{i \rightarrow ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{d\theta}{\theta} \frac{dz}{z}$$

Insensitive to *quark*/*gluon* composition. Can we do the *opposite*?!

[Tripathee, Xue, Larkoski, Marzani, JDT, [PRL 2017](#), [PRD 2017](#);
based on Larkoski, Marzani, JDT, [PRD 2015](#)]





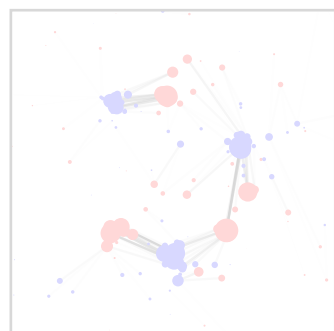
*Jets are manifestation of quarks and gluons,
but there is **no unambiguous way**
to tell a quark jet from a gluon jet*



2017: A Quark/Gluon Conundrum

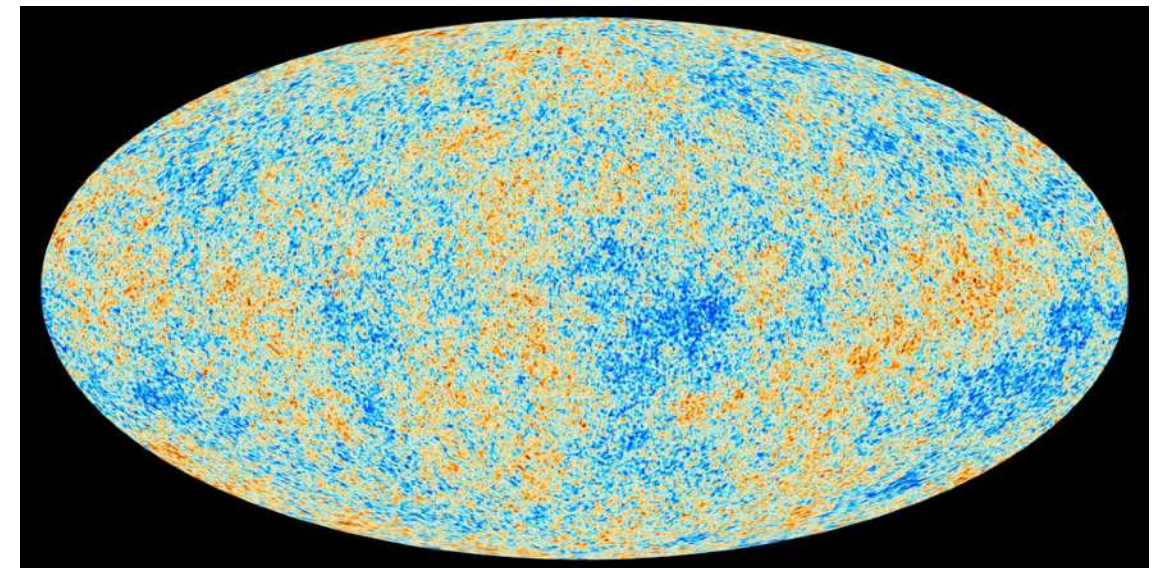
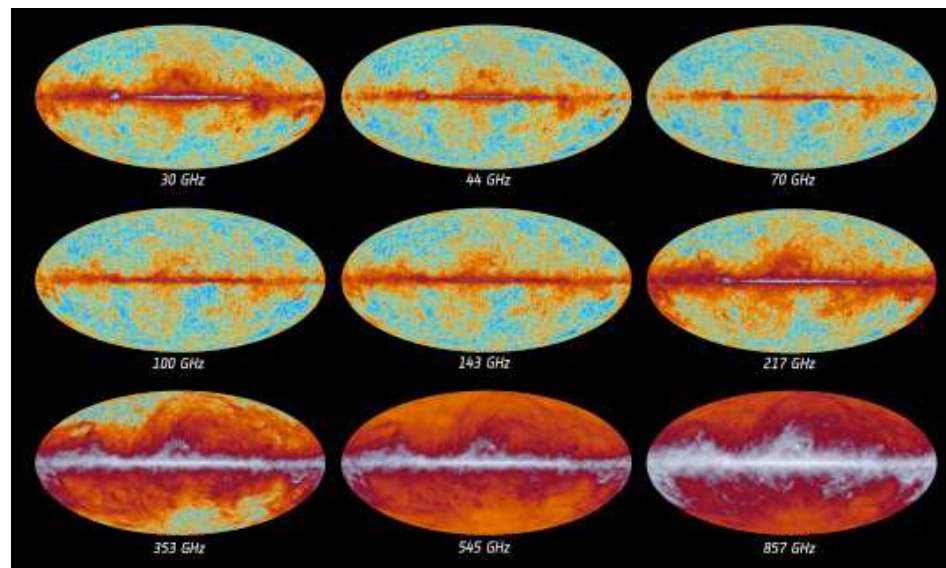
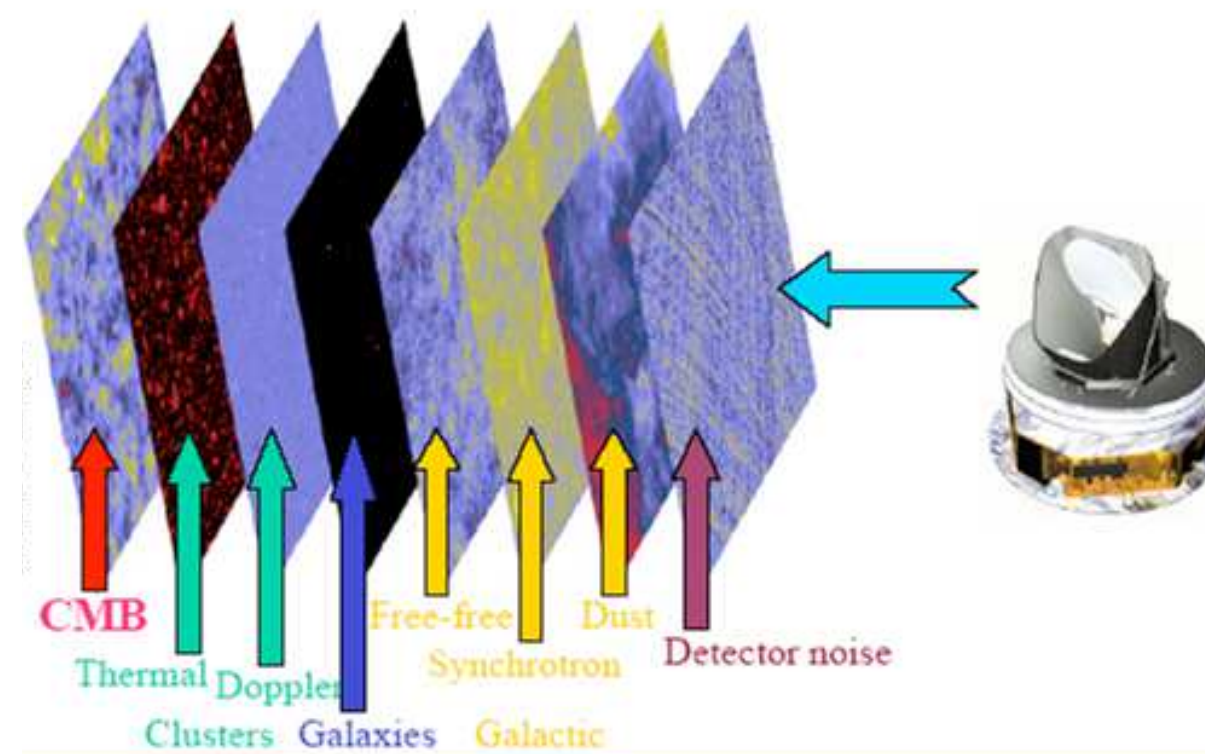


2018: Leveraging Weak Supervision



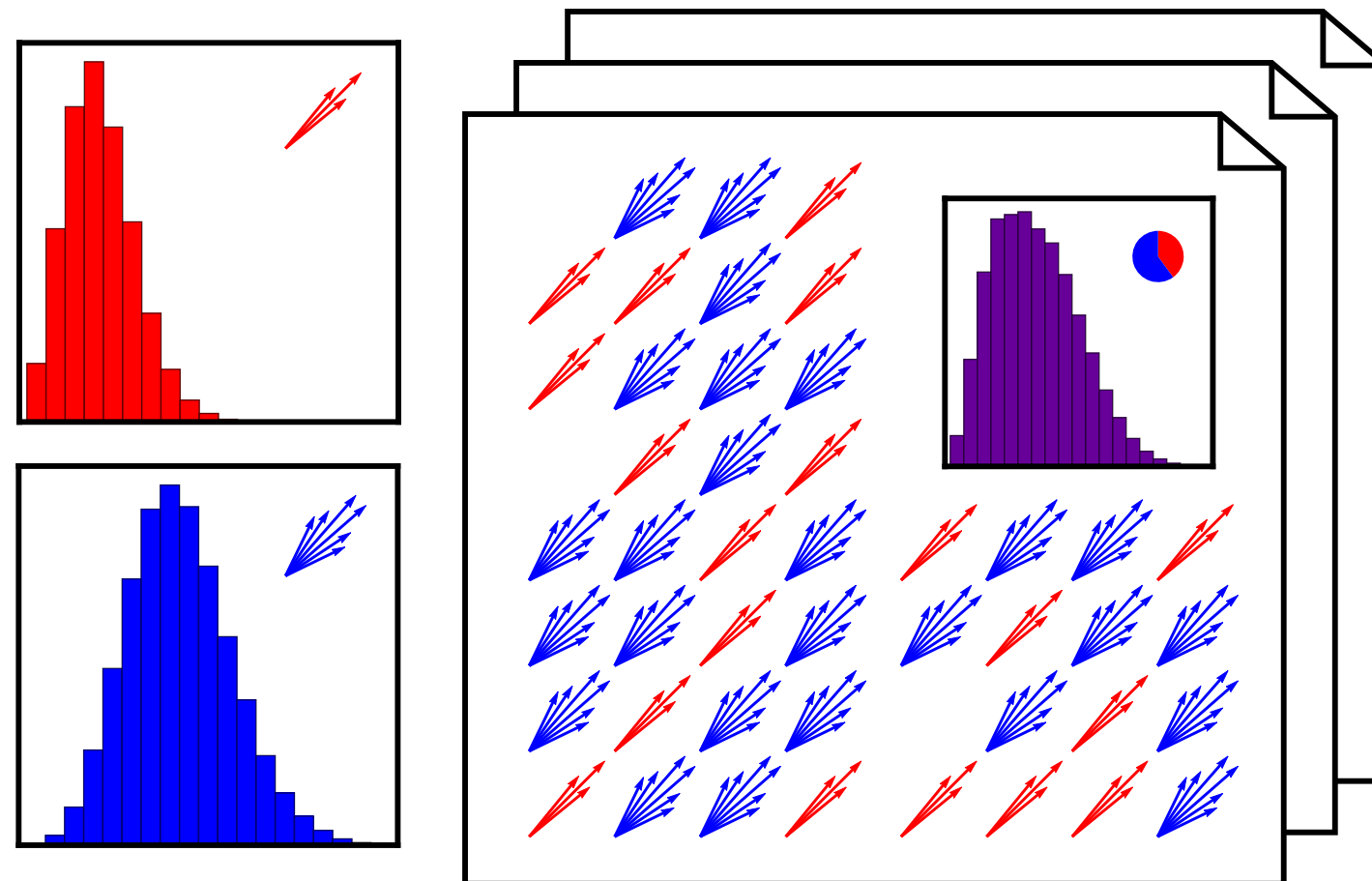
2019-2022: The Strong Force Revisited

Blind Source Separation for Cosmology



Blind Source Separation for Quark/Gluon Jets?

While you can't unambiguously label individual jets, you can extract **quark** and **gluon** distributions from **hadron-level measurements**



In natural language processing, this is known as “**topic modeling**”

[Komiske, Metodiev, JDT, *JHEP* 2018; cf. ATLAS, *PRD* 2019; using Metodiev, Nachman, JDT, *JHEP* 2017; Metodiev, JDT, *PRL* 2018]
[see also Blanchard, Flaska, Handy, Pozzi, Scott, *PLMR* 2013; Katz-Samuels, Blanchard, Scott, *JMLR* 2016]
[Yes, for those of you obsessed with these things, I flipped the color of quark and gluon on this slide.]

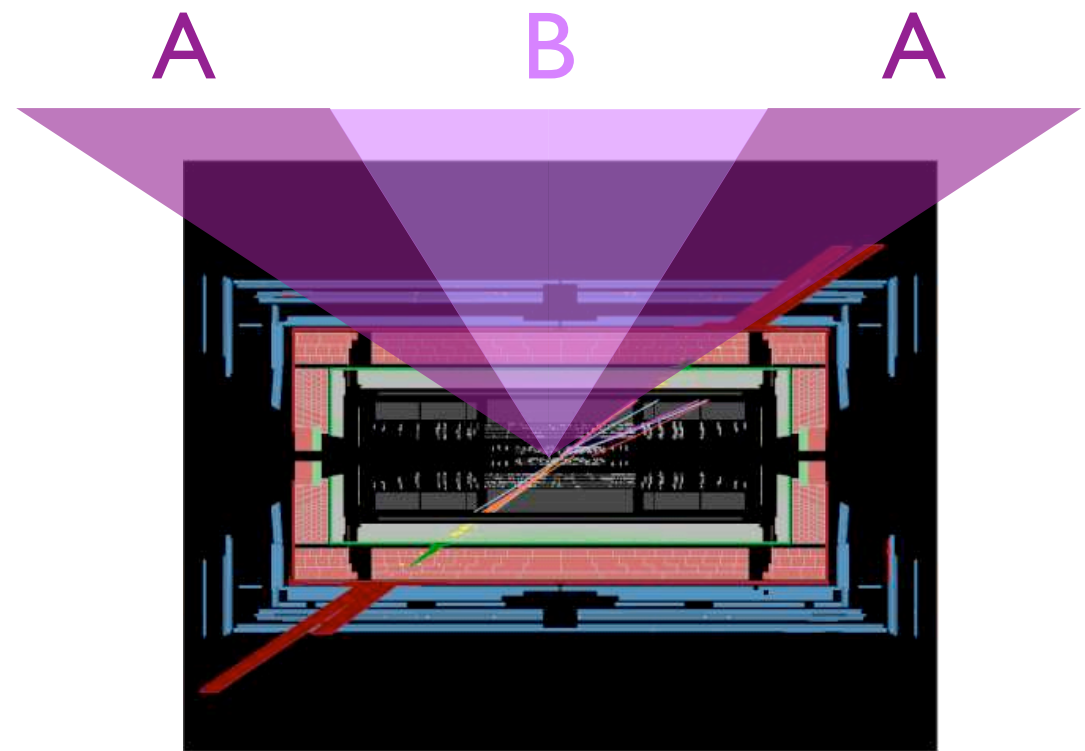
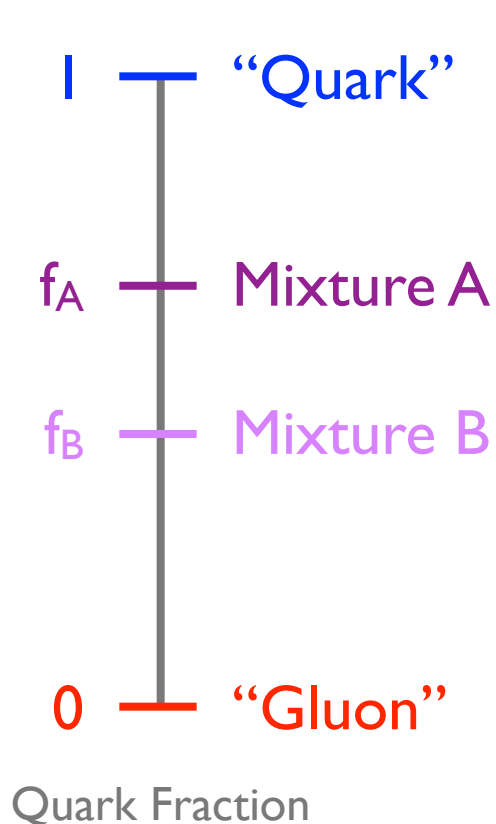


Fundamental Assumption

One can make jet samples that are mixtures of “quarks” and “gluons”

$$p_{\text{mixed}}(\vec{x}) = f_q p_{\text{quark}}(\vec{x}) + (1 - f_q) p_{\text{gluon}}(\vec{x})$$

↳ Jet Features



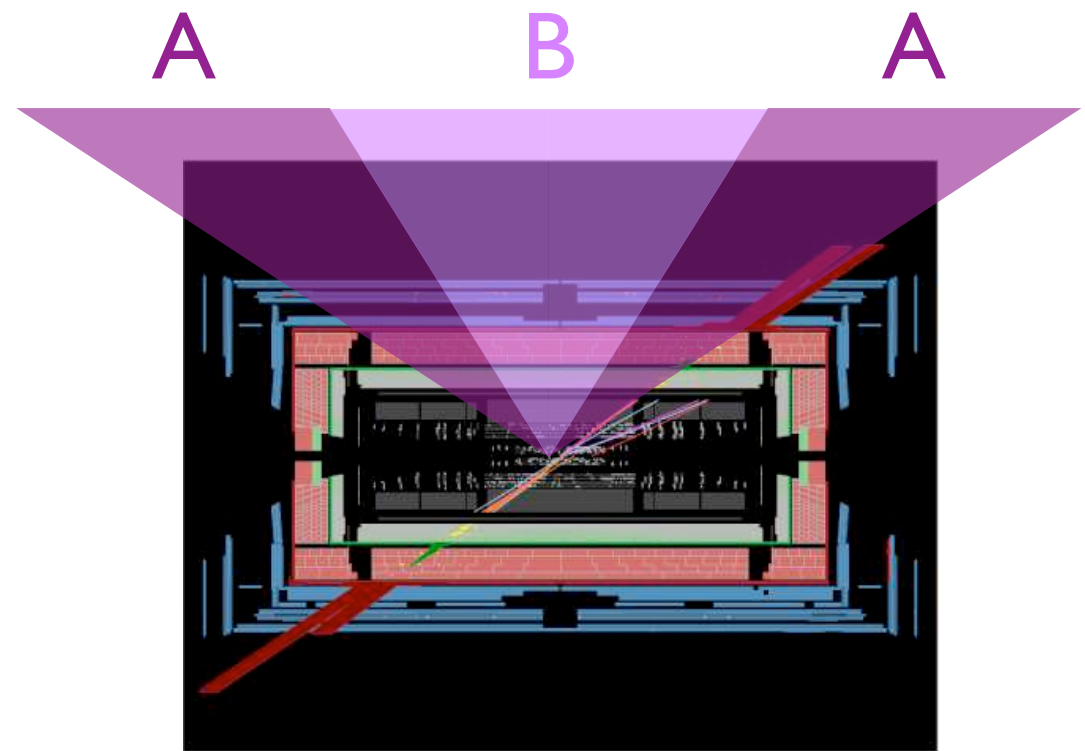
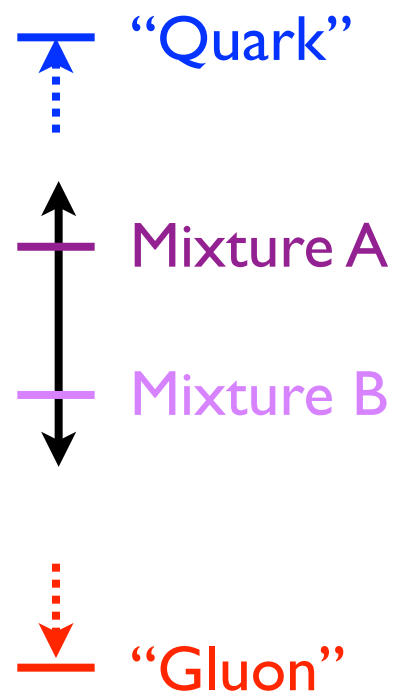
Non-trivial! Assumes A/B mixtures have unbiased jet properties

Fundamental Trick

From A/B mixtures you can define “quarks” and “gluons”

$$p_{\text{quark}}(\vec{x}) = \frac{p_A(\vec{x}) - \kappa_{AB} p_B(\vec{x})}{1 - \kappa_{AB}}$$

$$p_{\text{gluon}}(\vec{x}) = \frac{p_B(\vec{x}) - \kappa_{BA} p_A(\vec{x})}{1 - \kappa_{BA}}$$

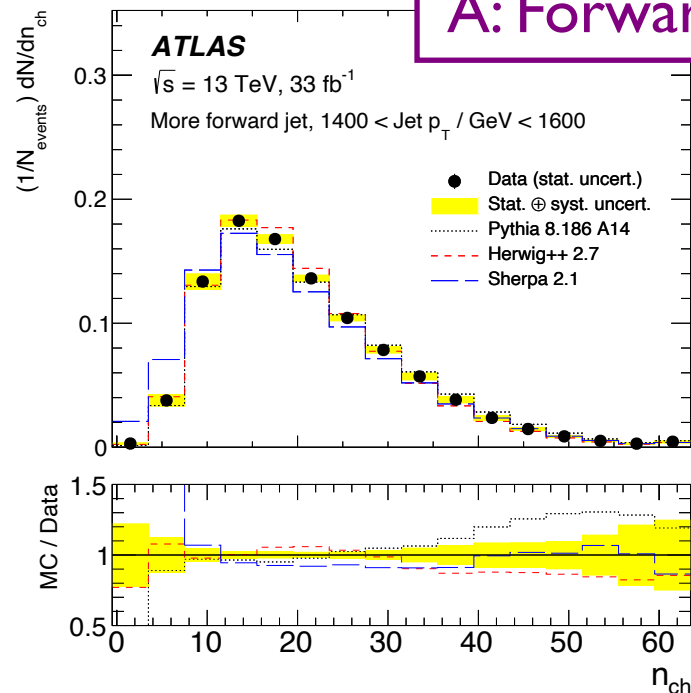


Choosing κ as big as possible yields “mutually irreducible” distributions

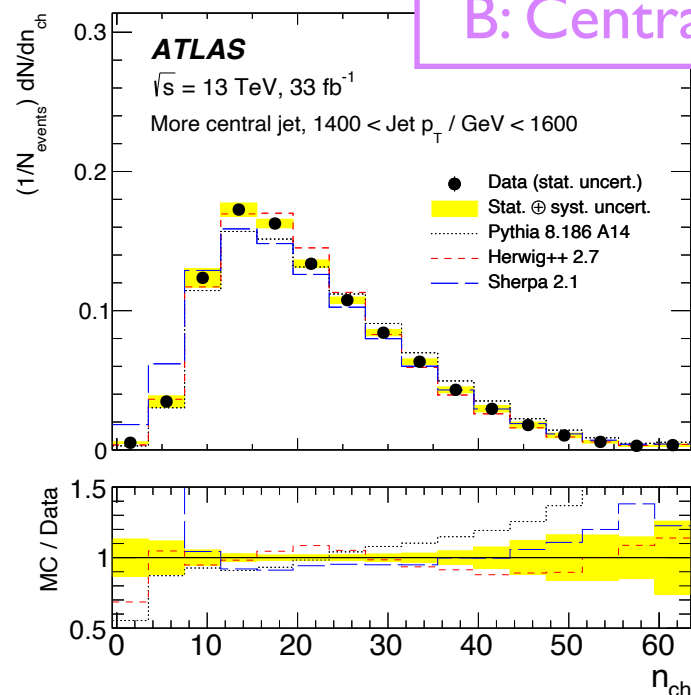
[Katz-Samuels, Blanchard, Scott, [JLMR 2016](#); advocated for in Komiske, Metodiev, JDT, [JHEP 2018](#)]

Jet Topics Result from ATLAS

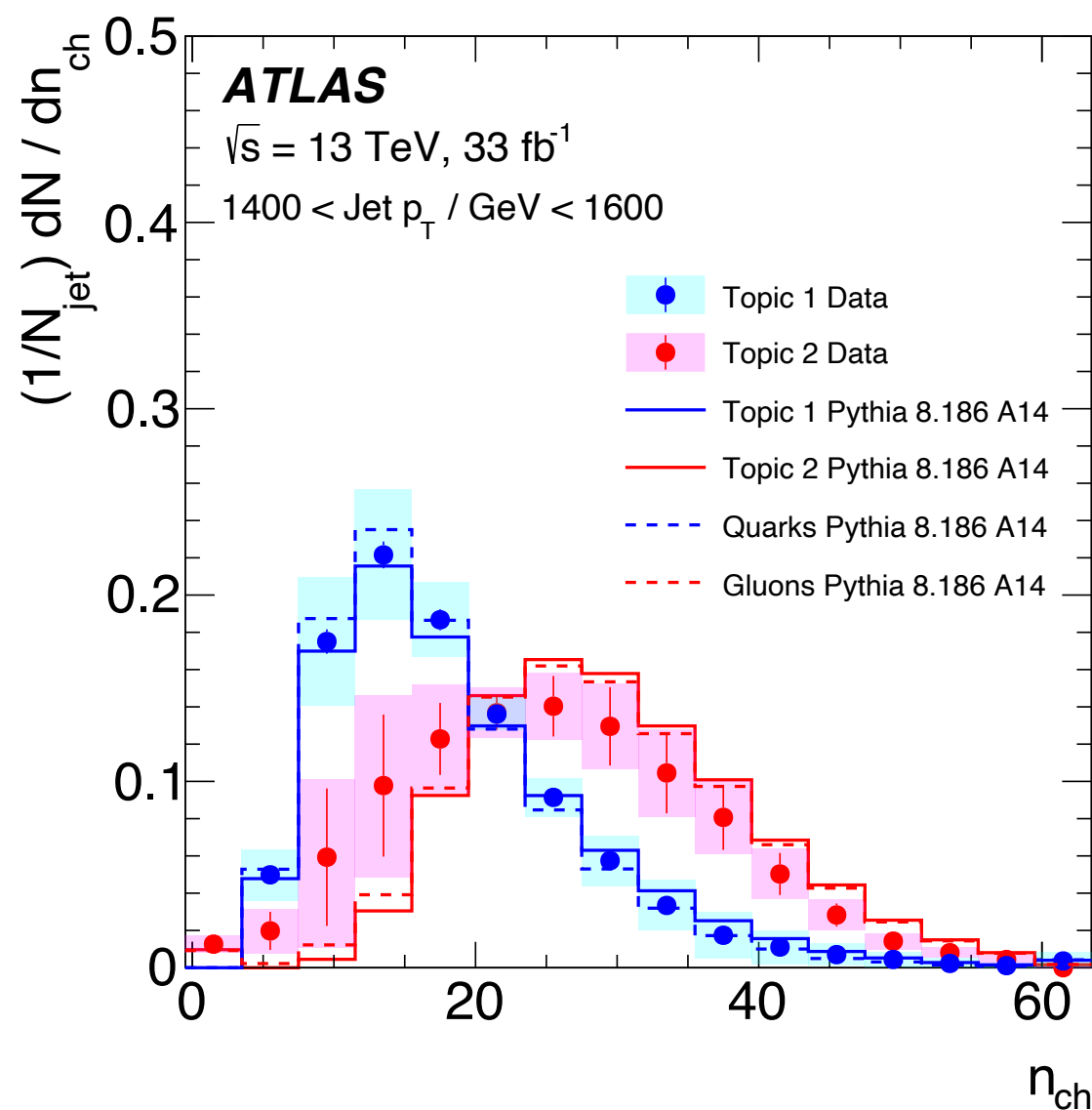
A: Forward



B: Central



Track multiplicity for
“Topic 1” and “Topic 2”



[ATLAS, PRD 2019]

*“Ok, but don’t jet topics
depend on choice of jet features?”*

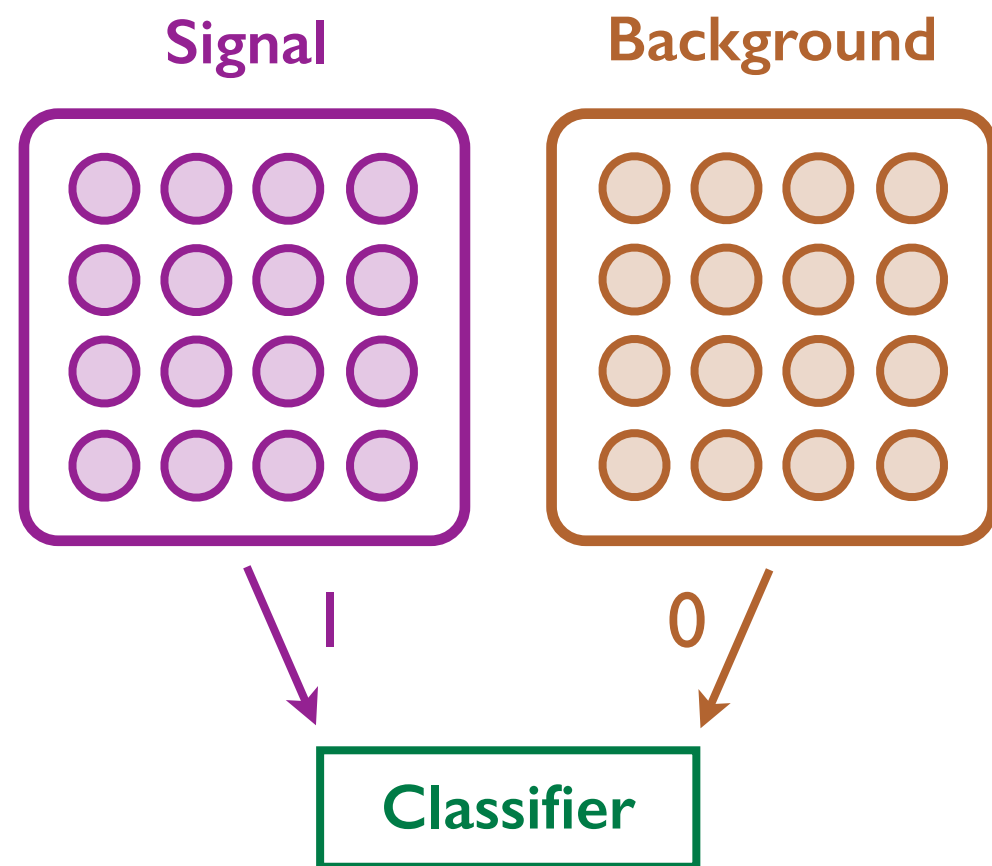


Yes! But to find maximal separability,
we can **leverage weak supervision!**
CWoLa = Classification Without Labels

(More accurately, classification with noisy labels)

Reminder: Fully-Supervised Binary Classification

Training on pure, perfectly labeled examples



Minimize Loss Function

(asymptotically)

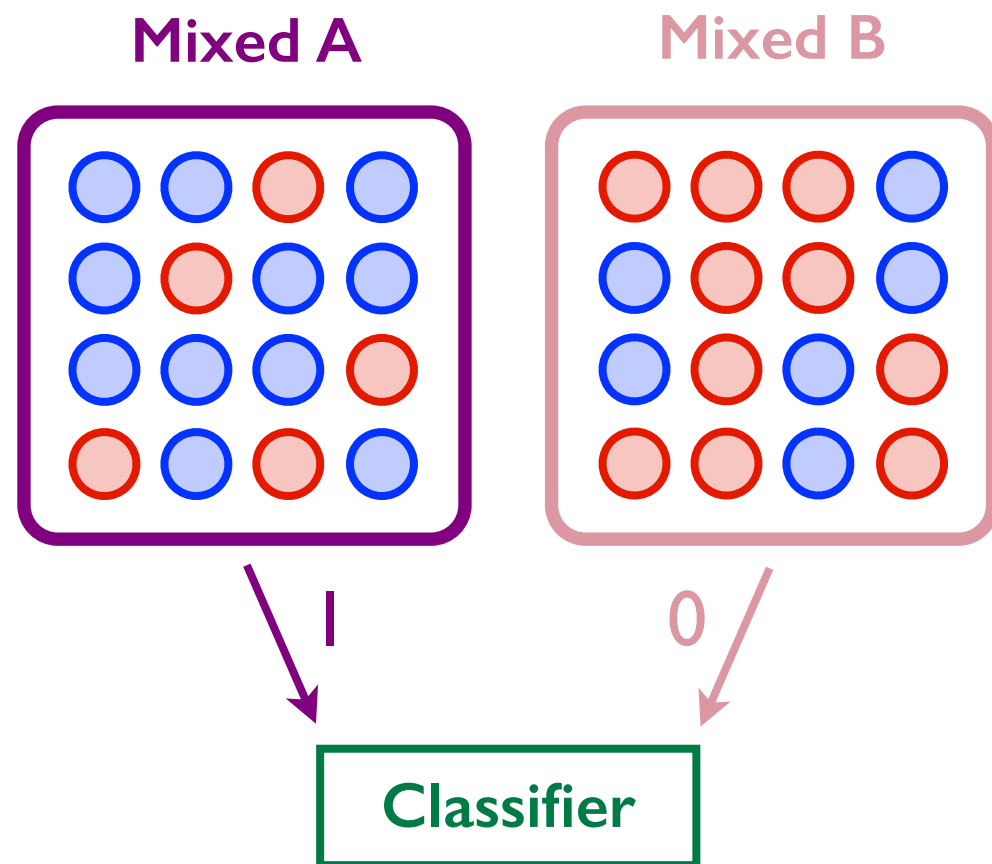
$$h(\vec{x}) = \frac{p_{\text{sig}}(\vec{x})}{p_{\text{sig}}(\vec{x}) + p_{\text{bkgd}}(\vec{x})}$$

Optimal Classifier (Neyman–Pearson)

$$\ell_{\text{MSE}} = \left\langle \left(\underset{\substack{\uparrow \\ \text{Classifier}}}{h(\vec{x})} - \underset{\substack{\uparrow \\ \text{Inputs}}}{1} \right)^2 \right\rangle_{\text{signal}} + \left\langle \left(\underset{\substack{\uparrow \\ \text{Classifier}}}{h(\vec{x})} - \underset{\substack{\uparrow \\ \text{Inputs}}}{0} \right)^2 \right\rangle_{\text{background}}$$

Classification Without Labels (CWoLa)

Train directly on *mixed data*!

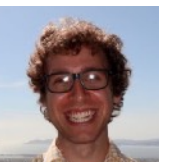


$$h_{\text{mixed}}(\vec{x}) = \frac{p_A(\vec{x})}{p_A(\vec{x}) + p_B(\vec{x})}$$
$$\neq$$
$$h_{\text{pure}}(\vec{x}) = \frac{p_q(\vec{x})}{p_q(\vec{x}) + p_g(\vec{x})}$$

but... $\frac{\partial h_{\text{mixed}}(\vec{x})}{\partial h_{\text{pure}}(\vec{x})} > 0$

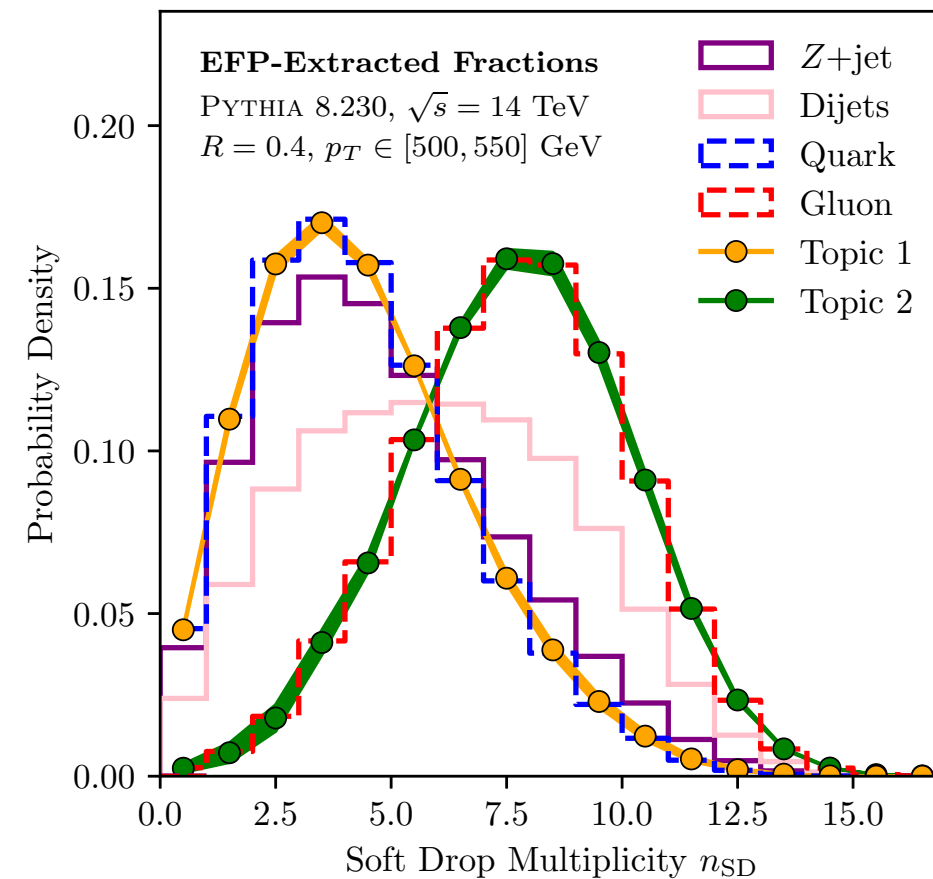
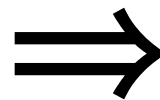
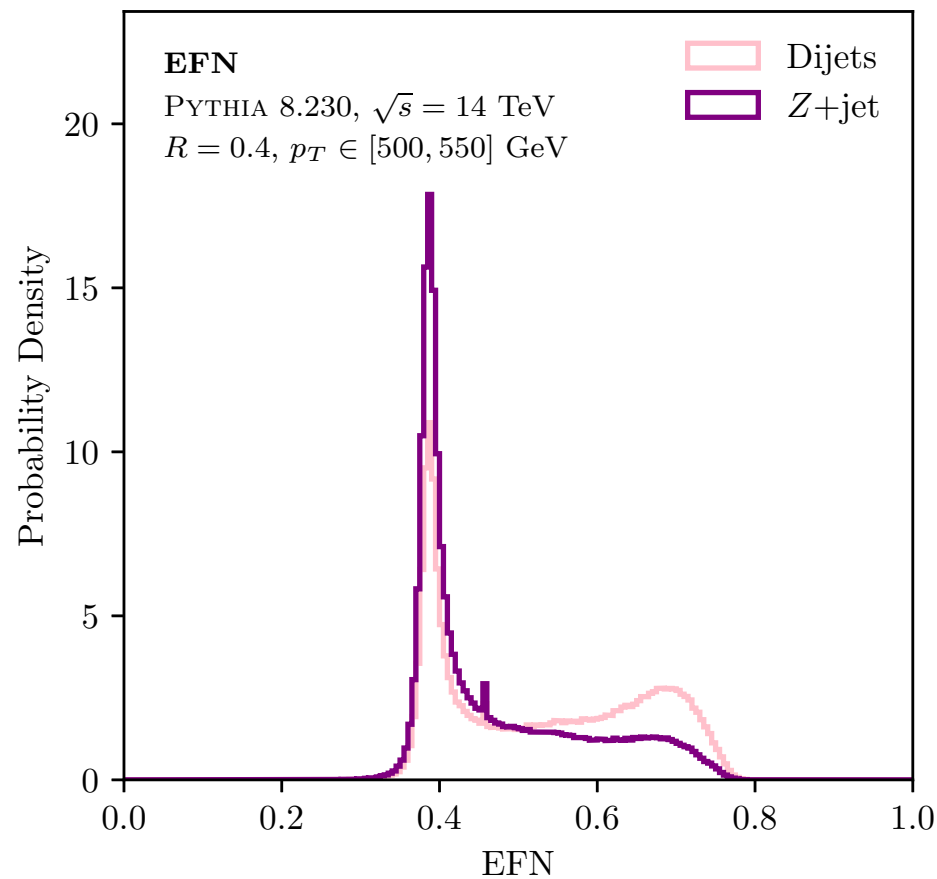
Weak supervision yields *same decision boundaries* as strong supervision!

[Metodiev, Nachman, JDT, [JHEP 2017](#); see also Blanchard, Flaska, Handy, Pozzi, Scott, [PLMR 2013](#);
Cranmer, Pavez, Louppe, [arXiv 2015](#); Dery, Nachman, Rubbo, Schwartzman, [JHEP 2017](#);
Cohen, Freytsis, Ostdiek, [JHEP 2018](#); Komiske, Metodiev, Nachman, Schwartz, [PRD 2018](#); Collins, Howe, Nachman, [PRL 2018](#), [PRD 2019](#)]



The Well-Read CWoLa

Weak supervision meets topic modeling



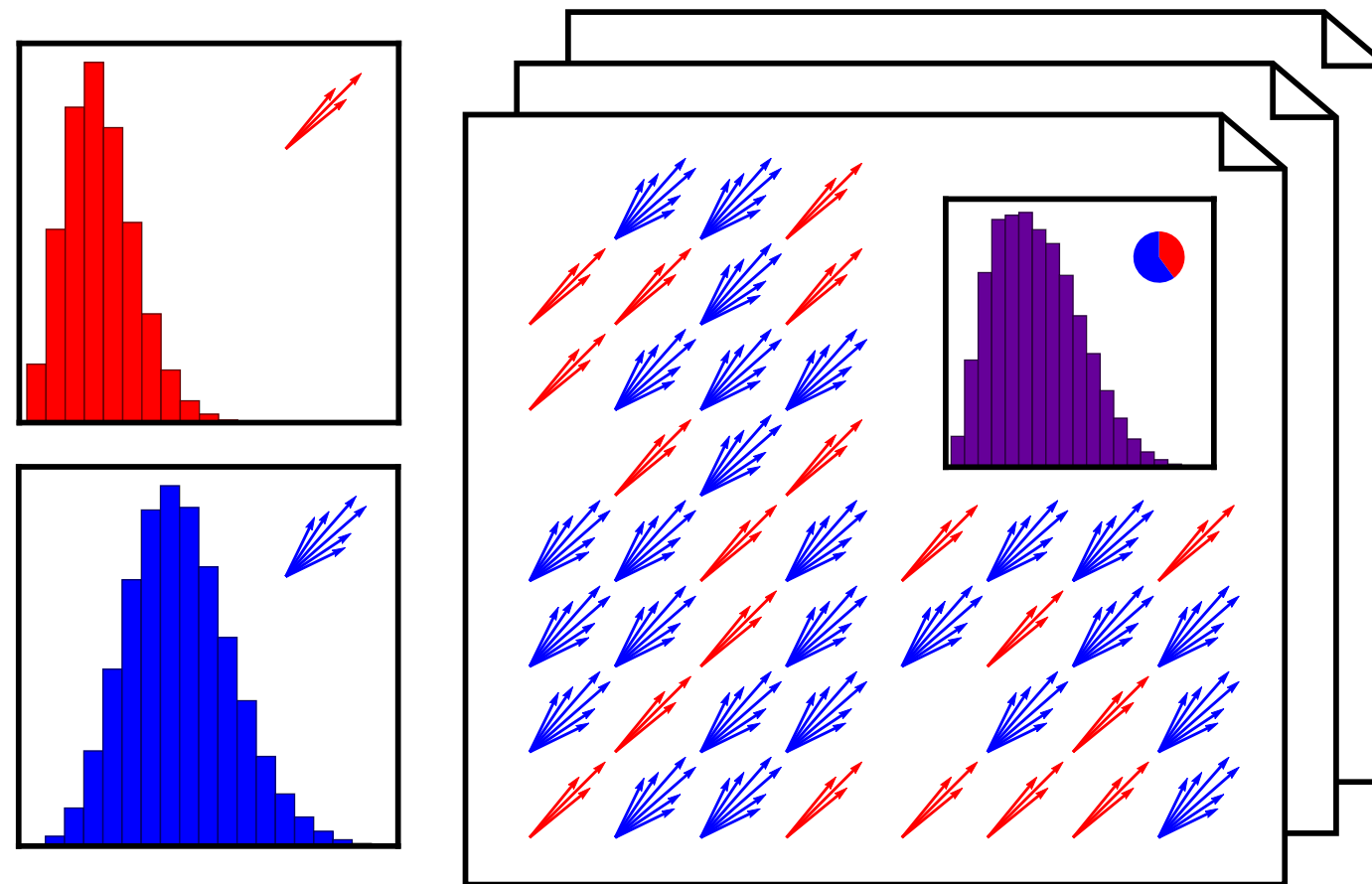
$h_{\text{pure}} \in [0, 1]$

$h_{\text{mixed}} \in \left[\frac{f_g^A}{f_g^A + f_g^B}, \frac{f_q^A}{f_q^A + f_q^B} \right]$

[Komiske, Metodiev, JDT, JHEP 2018]

[using Katz-Samuels, Blanchard, Scott, JLMR 2016; Metodiev, Nachman, JDT, JHEP 2017; Metodiev, JDT, PRL 2018]





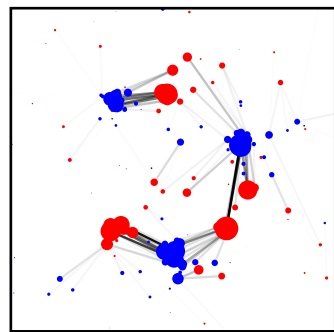
*By assuming jet samples are mixtures of “quarks” and “gluons”, one can **operationally define** jet categories*



2017: A Quark/Gluon Conundrum



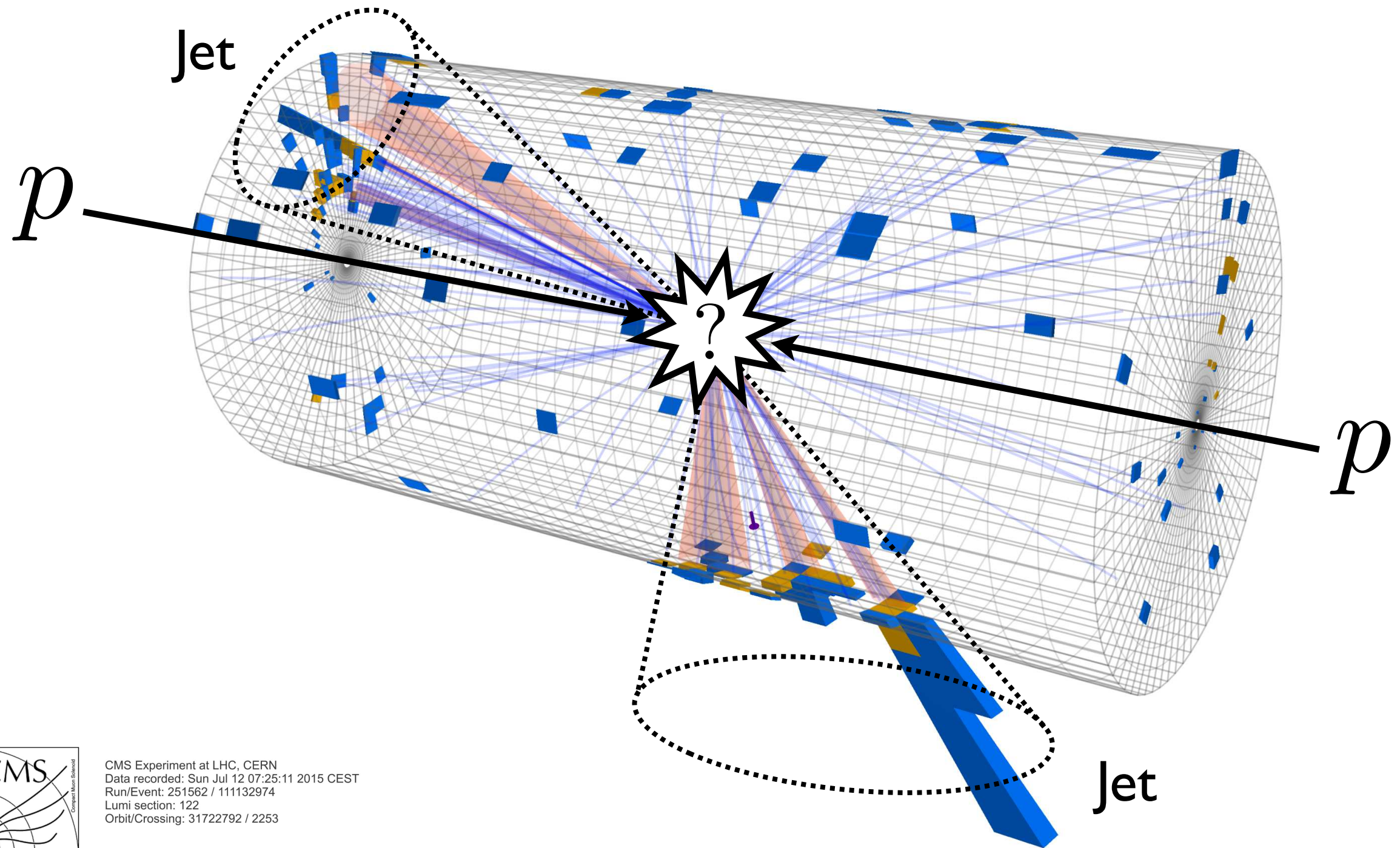
2018: Leveraging Weak Supervision



2019-2022: The Strong Force Revisited

“Collision Course”

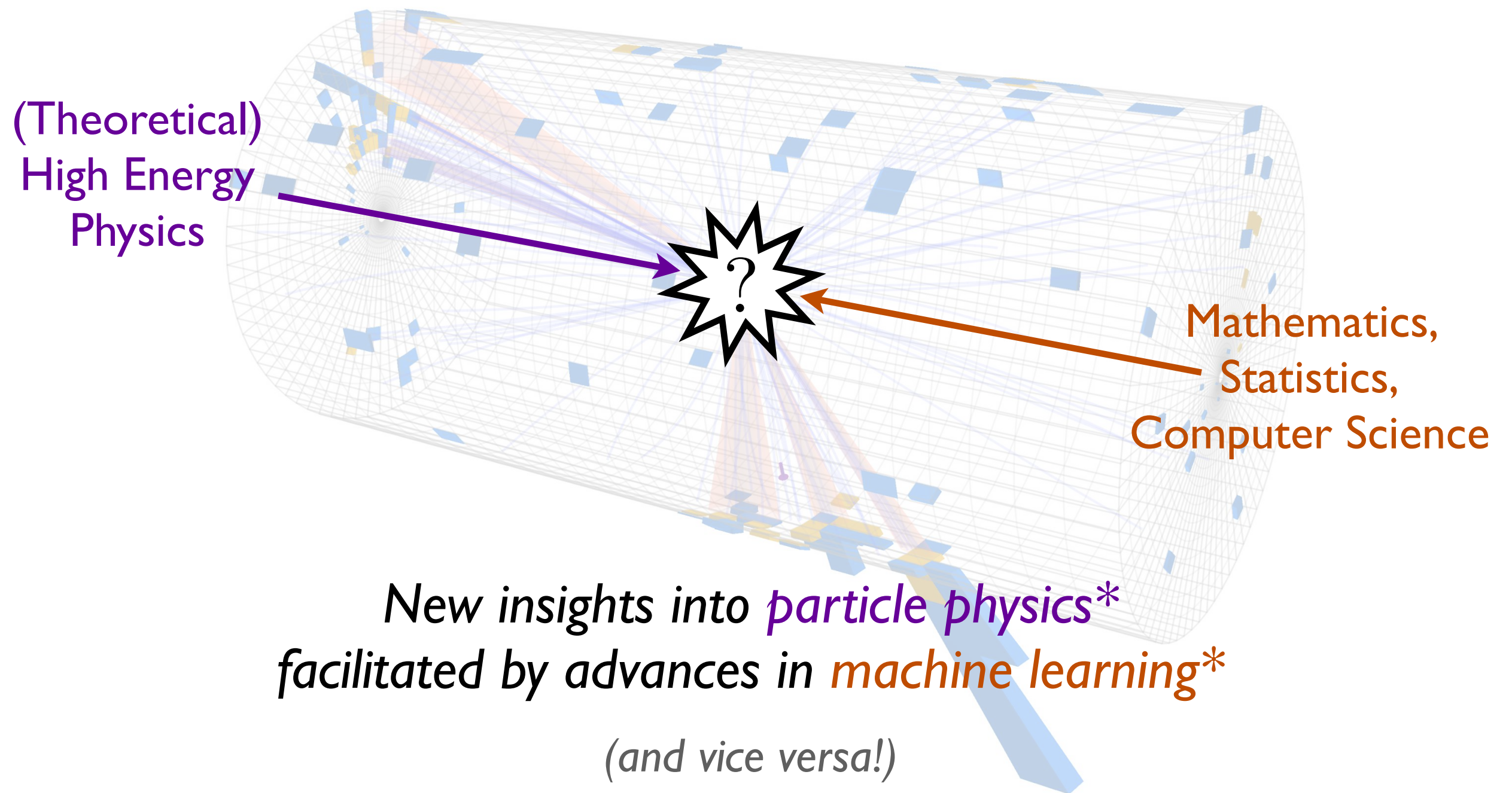
“Theoretical Physics for Machine Learning”
Aspen Center for Physics, January 2019



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

“Collision Course”

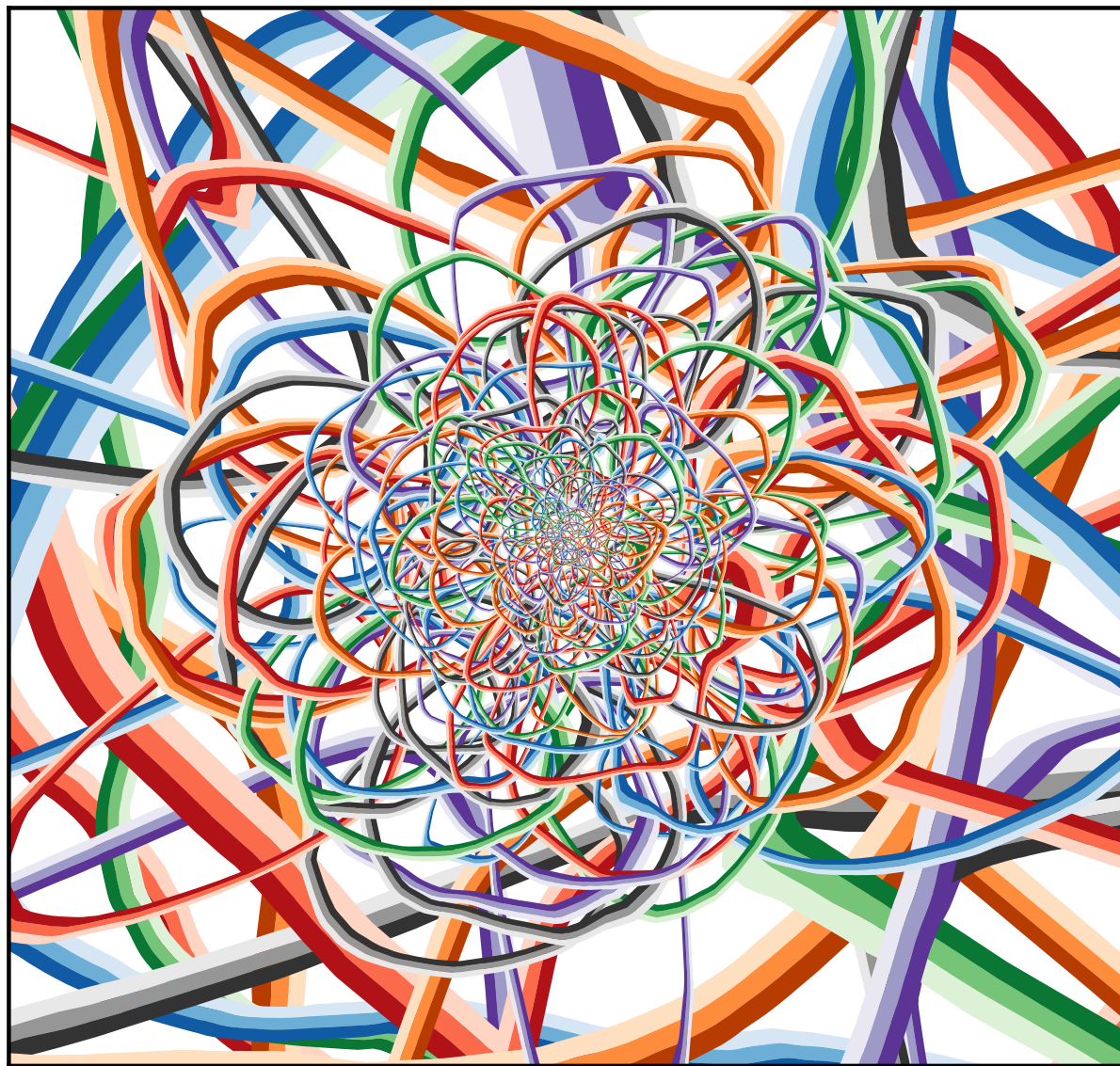
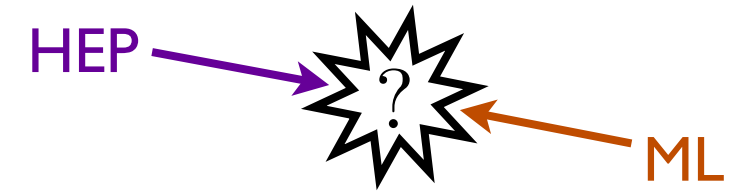
“Theoretical Physics for Machine Learning”
Aspen Center for Physics, January 2019



Cue the Montage!

2019: Point Cloud Learning

Energy Flow Networks



Set-based architecture with
interpretable latent space:

$$S(\mathcal{J}) = F(V_1, V_2, \dots, V_\ell)$$

$$V_a(\mathcal{J}) = \sum_{i \in \mathcal{J}} E_i \Phi_a(\hat{n}_i)$$

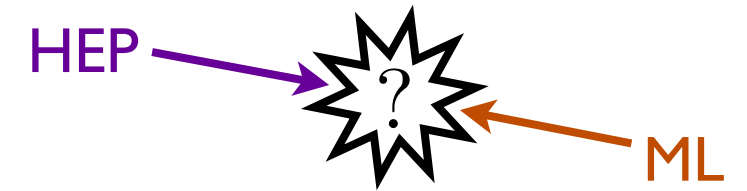
Collinear singularity of QCD!

$$dP_{i \rightarrow ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{d\theta}{\theta} \frac{dz}{z}$$

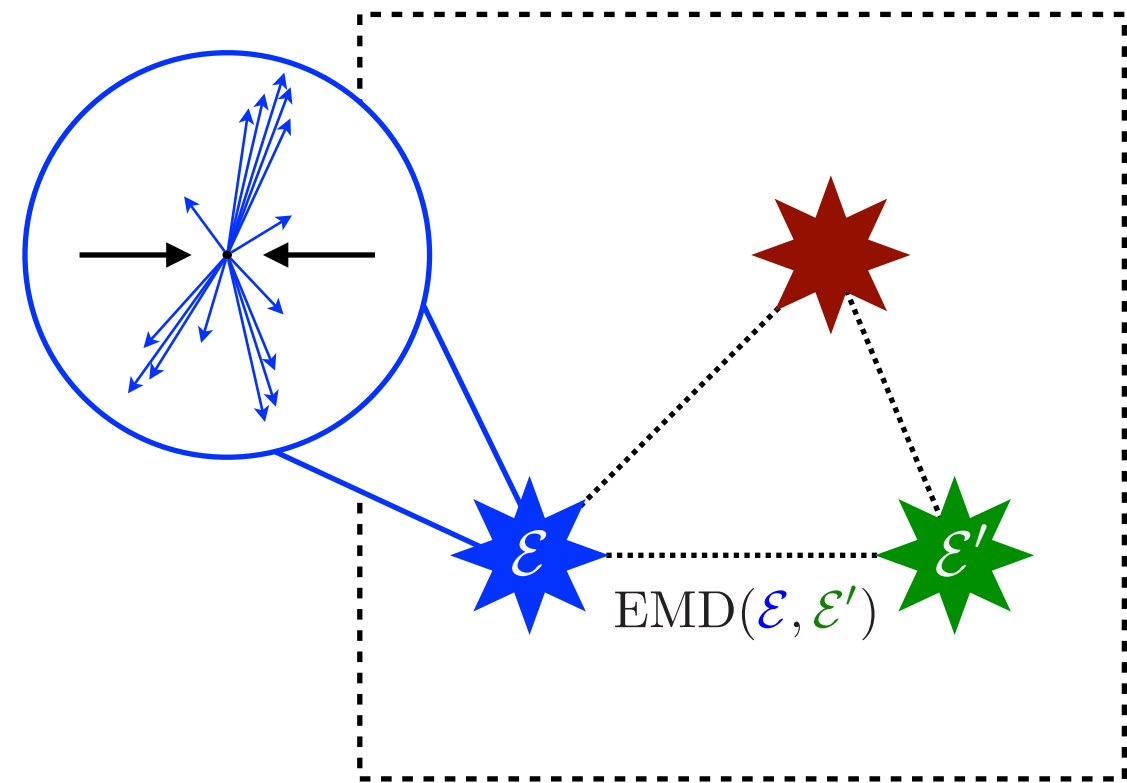
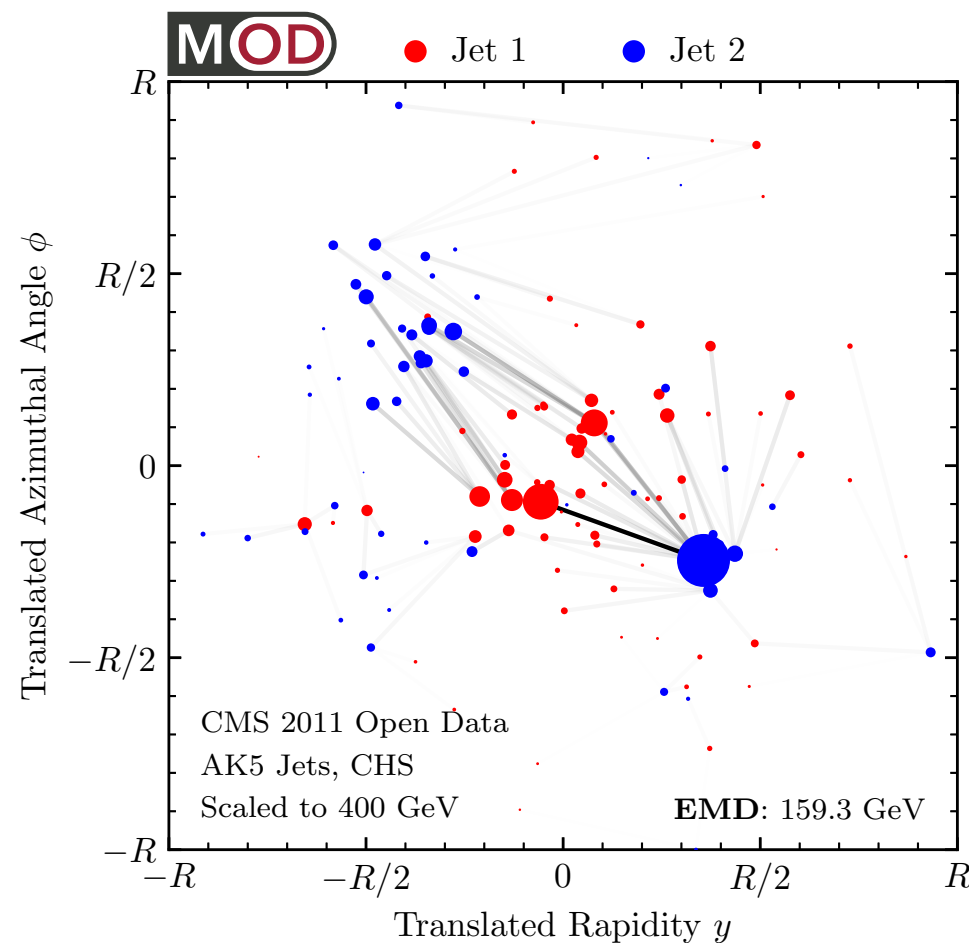
[Komiske, Metodiev, JDT, [JHEP 2019](#); see also Komiske, Metodiev, JDT, [JHEP 2018](#); code at energyflow.network;
special case of Zaheer, Kottur, Ravanbakhsh, Póczos, Salakhutdinov, Smola, [NIPS 2017](#);
other set-based architecture in Qu, Gouskos, [PRD 2020](#); Mikuni, Canelli, [EPJP 2020](#); Dolan, Ore, [arXiv 2020](#)]

2019: Metric Space for Colliders

Energy Mover's Distance



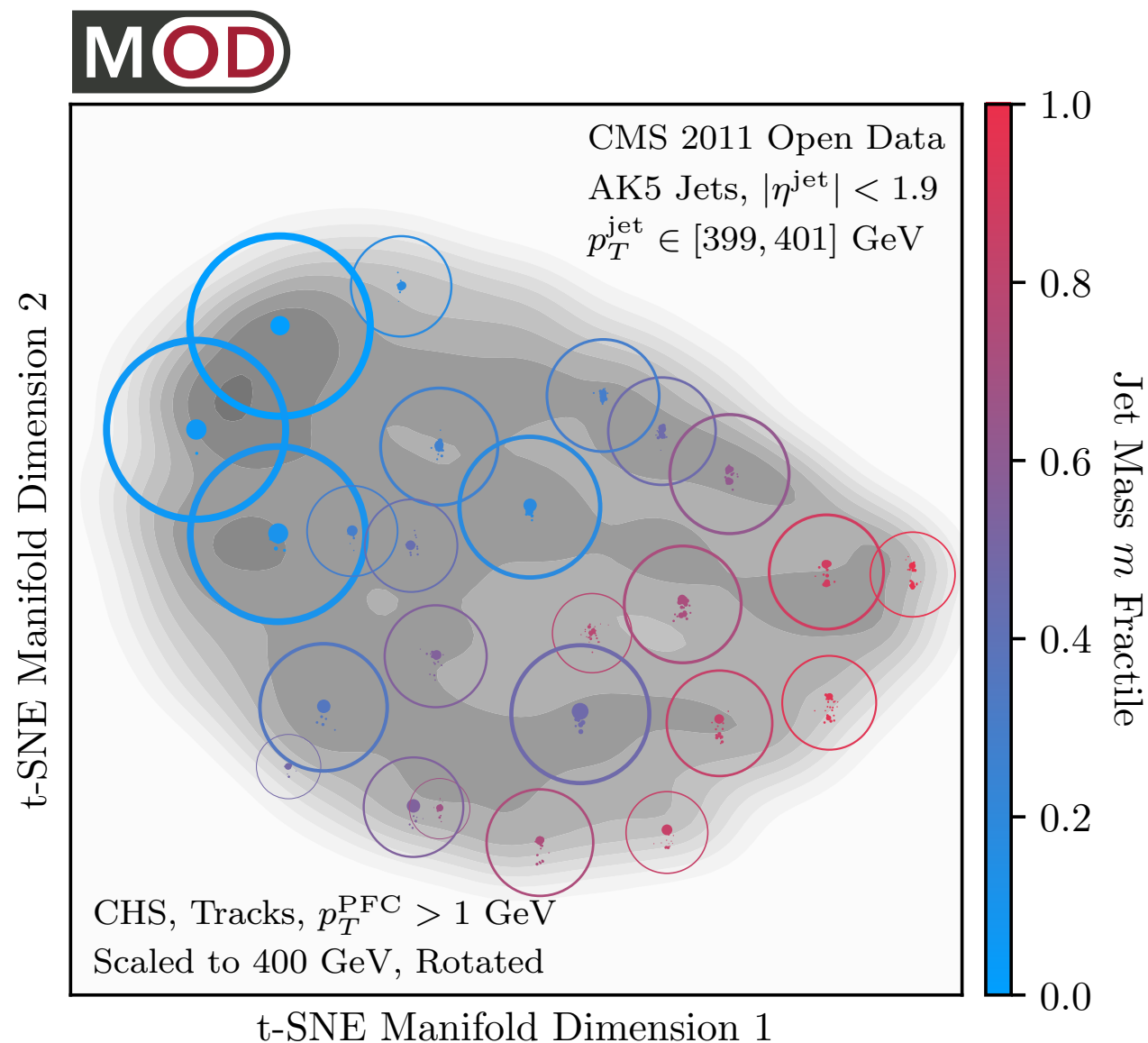
Optimal Transport for Collider Geometry



[Komiske, Metodiev, JDT, [PRL 2019](#); code at Komiske, Metodiev, JDT, [energyflow.network](#); open data study in Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#)]
[based on Peleg, Werman, Rom, [IEEE 1989](#); Rubner, Tomasi, Guibas, [ICCV 1998](#), [ICJV 2000](#); Pele, Werman, [ECCV 2008](#); Pele Taskar, [GSI 2013](#)]
[flavored variant in Crispim Romão, Castro, Milhano, Pedro, Vale, [EPJC 2021](#); linearized and unbalanced transport in Cai, Cheng, Craig, Craig, [PRD 2020](#), [PRD 2022](#)]

2020: Metric Space in Open Data

t-SNE Projection

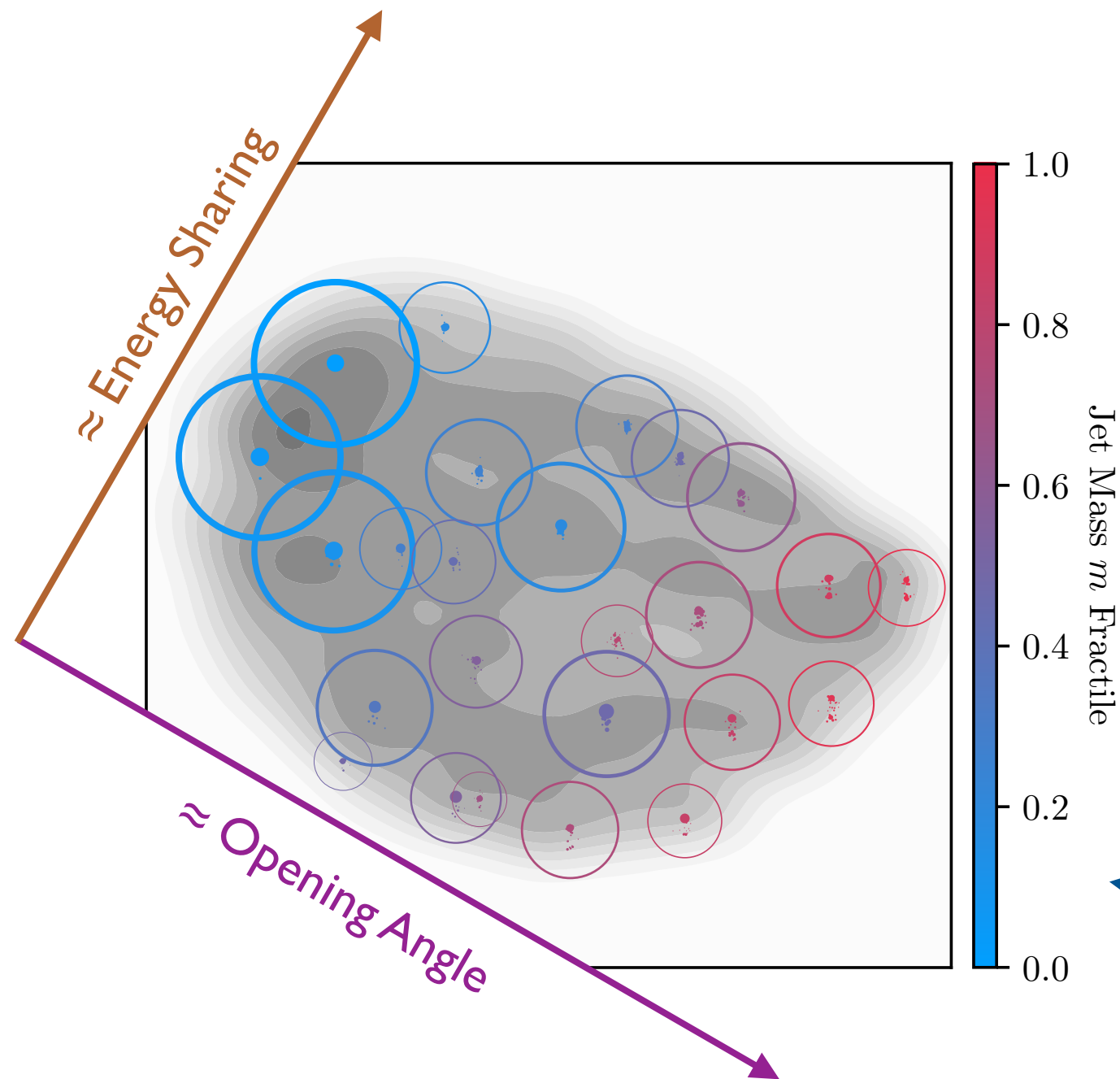


[Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#);
using van der Maaten, Hinton, [JMLR 2008](#); using [CMS Open Data](#)]

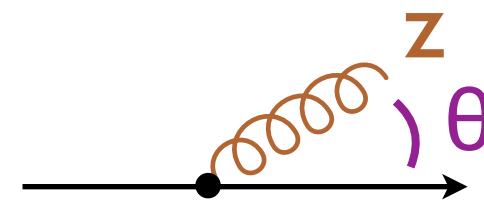


2020: Metric Space in Open Data

t-SNE Projection



Coordinate system of QCD!



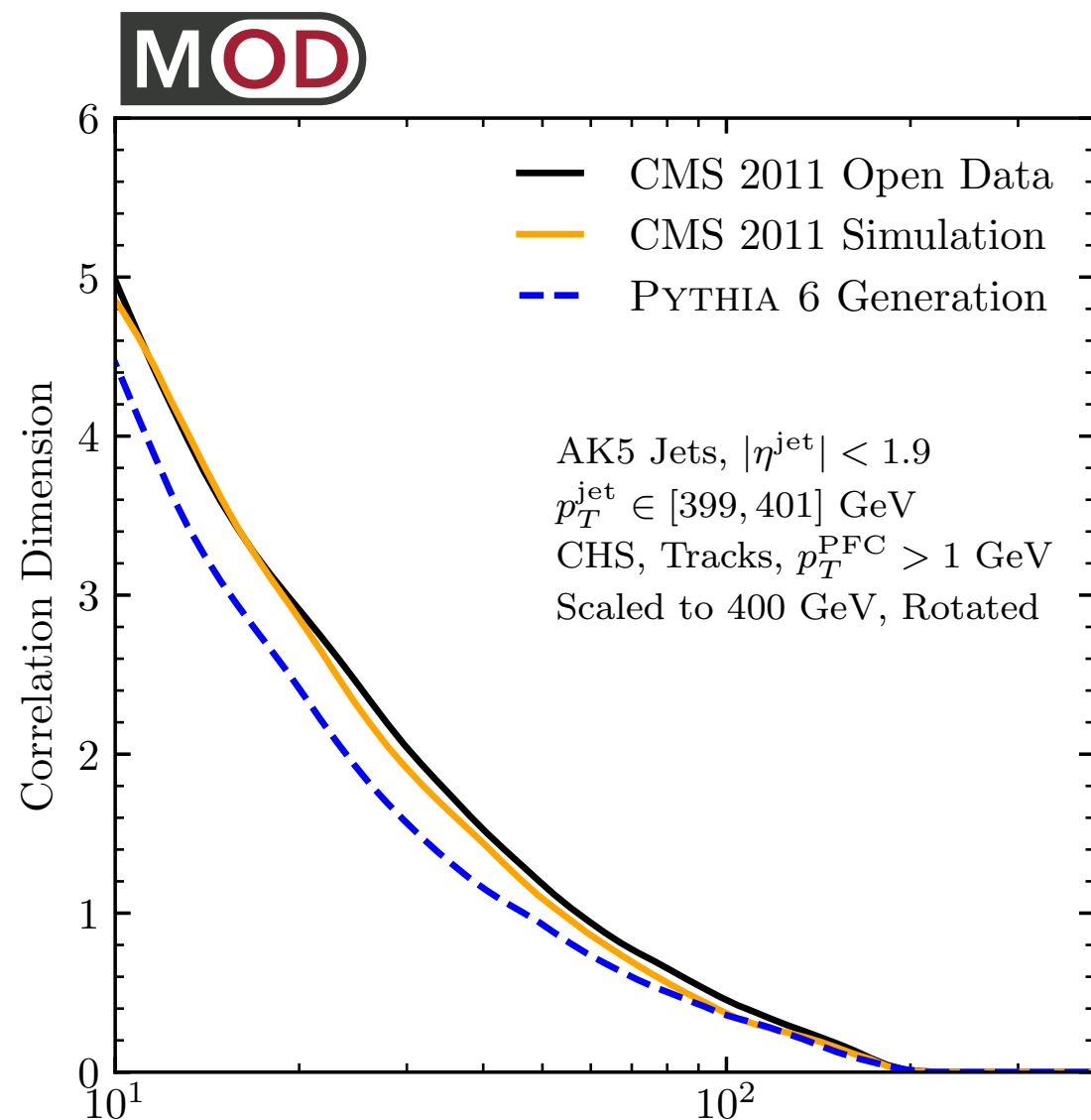
$$dP_{i \rightarrow ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{d\theta}{\theta} \frac{dz}{z}$$

[Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#);
using van der Maaten, Hinton, [JMLR 2008](#); using [CMS Open Data](#)]



2020: Metric Space in Open Data

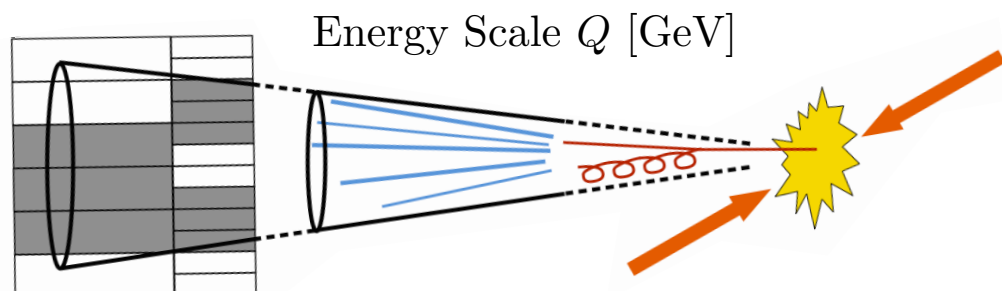
Correlation Dimension



Scaling behavior of QCD!

$$dP_{i \rightarrow ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{d\theta}{\theta} \frac{dz}{z}$$

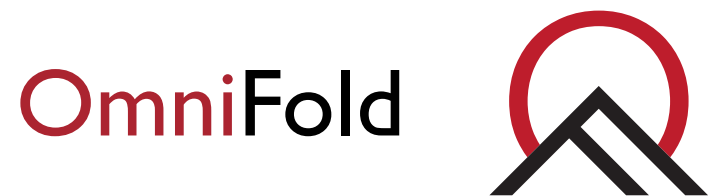
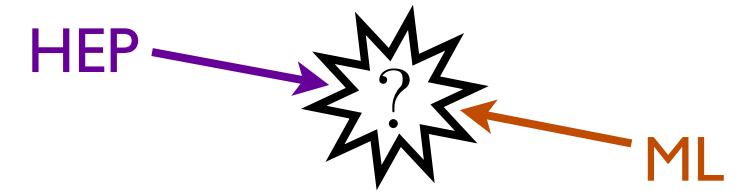
Hmm, detector effects distort underlying physics...



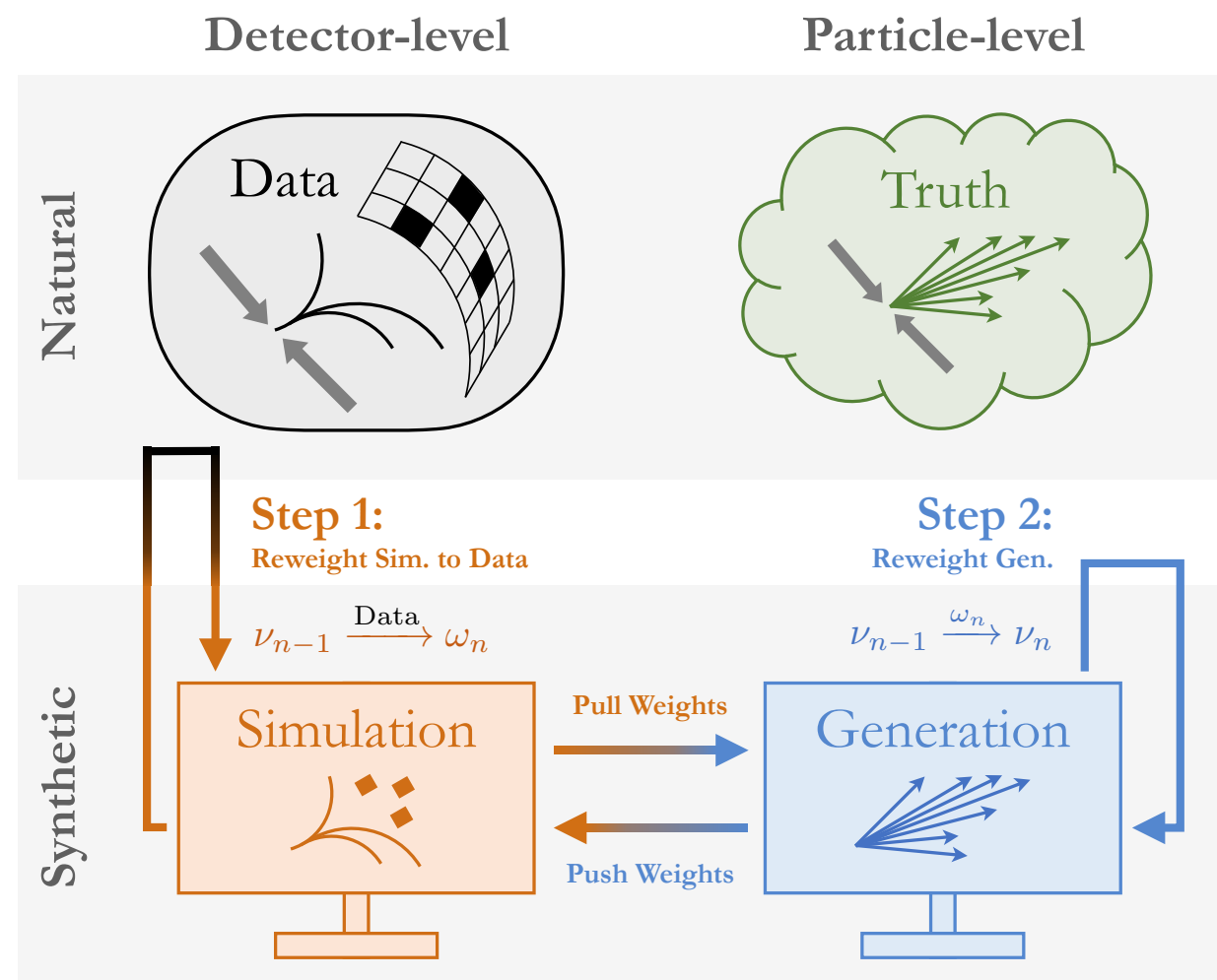
[Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#);
using Grassberger, Procaccia, [PRL 1983](#); Kégl, [NIPS 2002](#); using [CMS Open Data](#)]



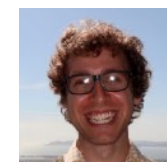
2020: Detector Unfolding



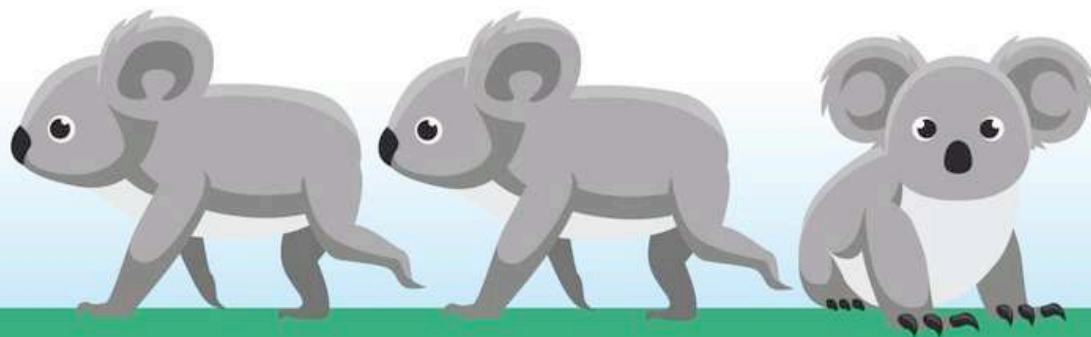
Multi-dimensional unbinned detector corrections
via iterated application of *machine-learned reweighting*



[Andreassen, Komiske, Metodiev, Nachman, JDT, [PRL 2020](#); + Suresh, [ICLR SimDL 2021](#);
Komiske, McCormack, Nachman, [PRD 2021](#); see unfolding comparison in Petr Baron, [APPB 2021](#)]
[see alternative in Bellagente, Butter, Kasieczka, Plehn, Rousselot, Winterhalder, Ardizzone, Köthe, [SciPost 2020](#)]



**Maintain a physical
distance of 1.5 m**



|..... 1.5 Metres|

(or about 3 koalas)

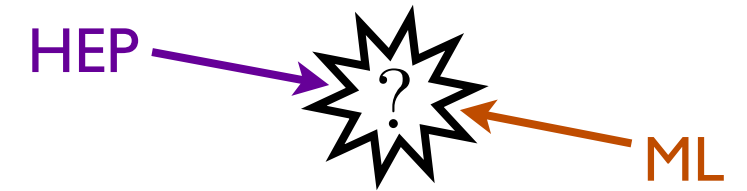


For updates, visit
covid19.act.gov.au

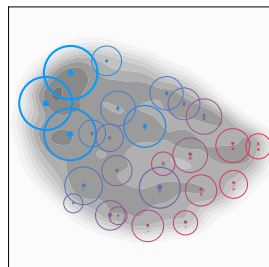
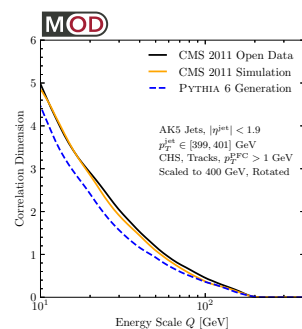


Oh, that...

2021: Launch of IAIFI

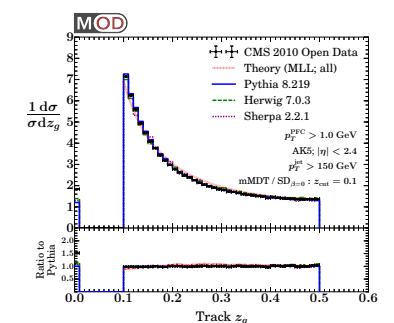
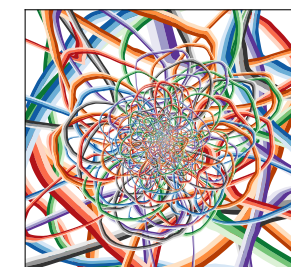


Can we finally put the “AI” in Altarelli-Parisi?



$$dP_{i \rightarrow ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{d\theta}{\theta} \frac{dz}{z}$$

↑

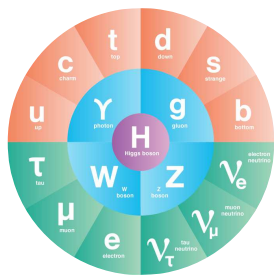


Must separate quarks from gluons!

2022: Putting All the Pieces Together



Quark and Gluon Jets
from the Strong Force



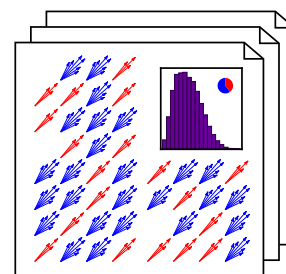
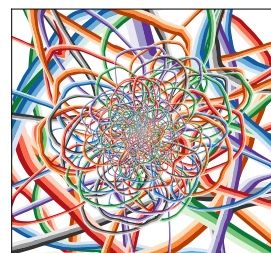
...Confronted with
Public Collider Data...



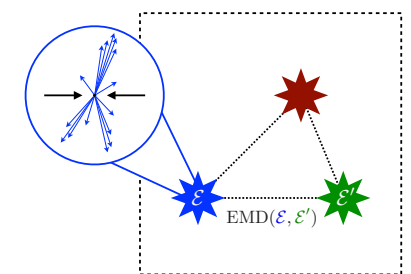
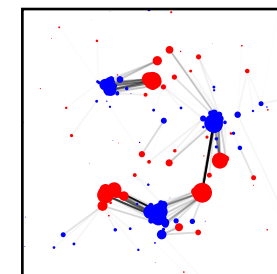
...Unfolded for
Detector Effects...



...Disentangled using Weak Supervision on
Set-Based Classifiers with Topic Modeling...



...Triangulated with
Optimal Transport!



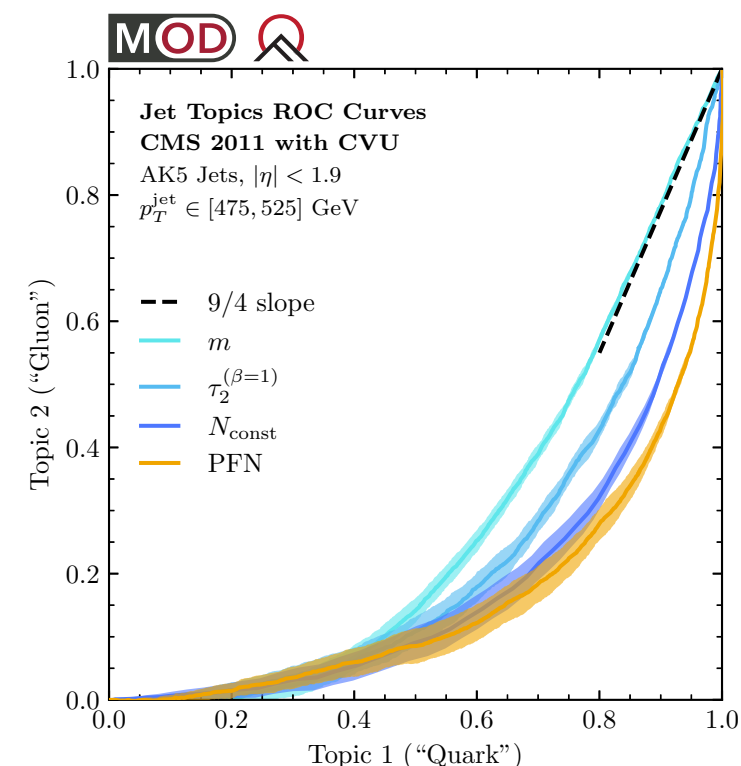
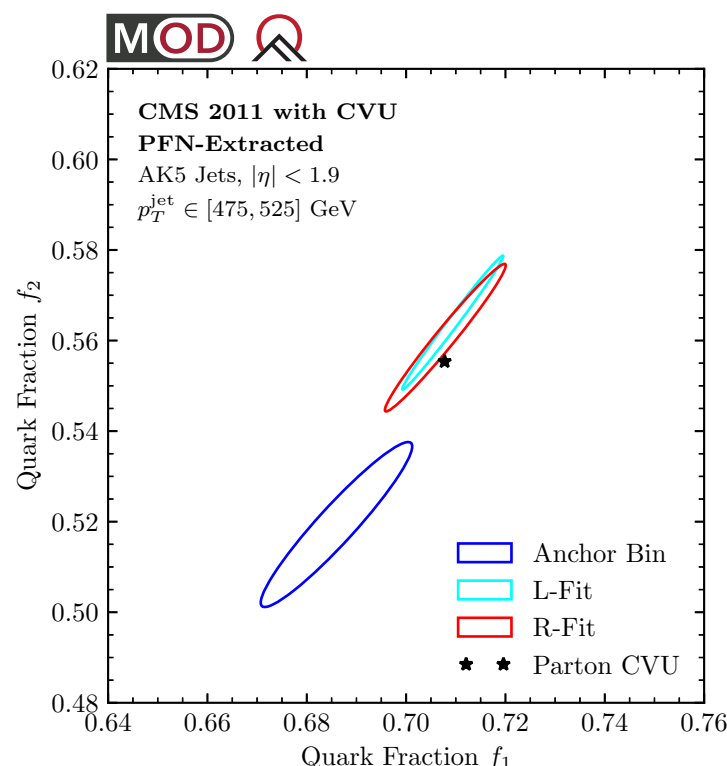
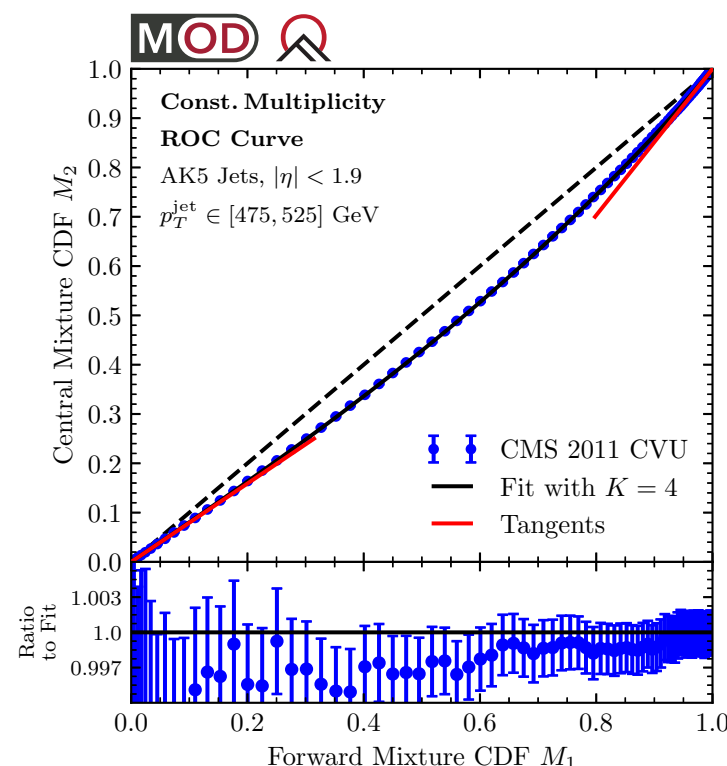
[Komiske, Kryhin, JDT, [arXiv 2022](#); using Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#);
Andreassen, Komiske, Metodiev, Nachman, JDT, [PRL 2020](#); + Suresh, [ICLR SimDL 2021](#);
Metodiev, Nachman, JDT, [JHEP 2017](#); Komiske, Metodiev, JDT, [JHEP 2019](#); Metodiev, JDT, [PRL 2018](#);
Komiske, Metodiev, JDT, [JHEP 2018](#); Komiske, Metodiev, JDT, [PRL 2019](#)]



2022: Putting All the Pieces Together



Plus some technical details as to how we extract reducibility factors, determine correlated uncertainties, and validate the method...

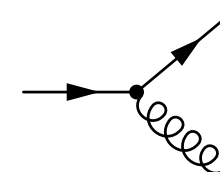
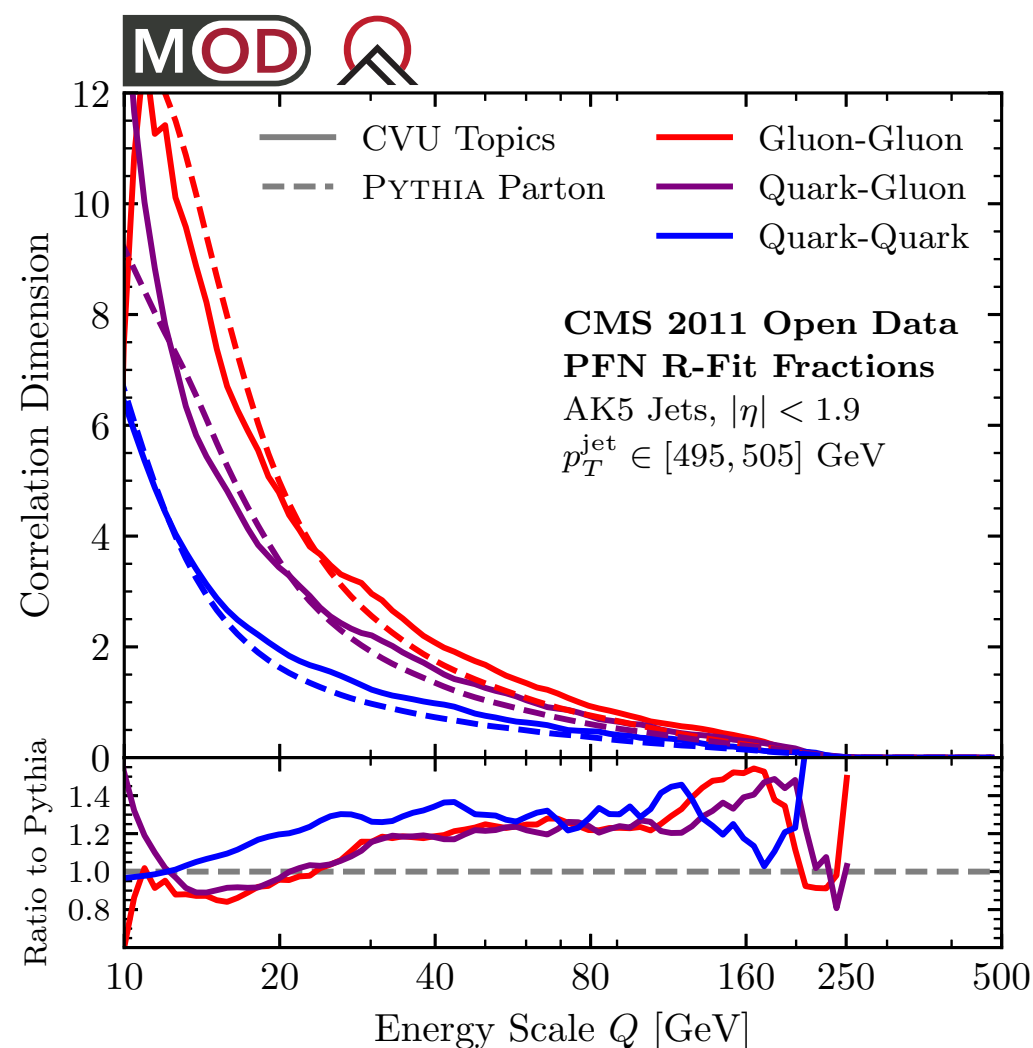


Once you know *quark fractions*, the rest is *linear algebra*

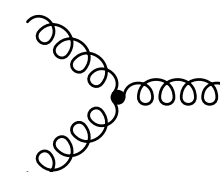
[Komiske, Kryhin, JDT, [arXiv 2022](#); using Komiske, Mastandrea, Metodiev, Naik, JDT, [PRD 2020](#); Andreassen, Komiske, Metodiev, Nachman, JDT, [PRL 2020](#); + Suresh, [ICLR SimDL 2021](#); Metodiev, Nachman, JDT, [JHEP 2017](#); Komiske, Metodiev, JDT, [JHEP 2019](#); Metodiev, JDT, [PRL 2018](#); Komiske, Metodiev, JDT, [JHEP 2018](#); Komiske, Metodiev, JDT, [PRL 2019](#)]



Dimensionality of Quark/Gluon Jets



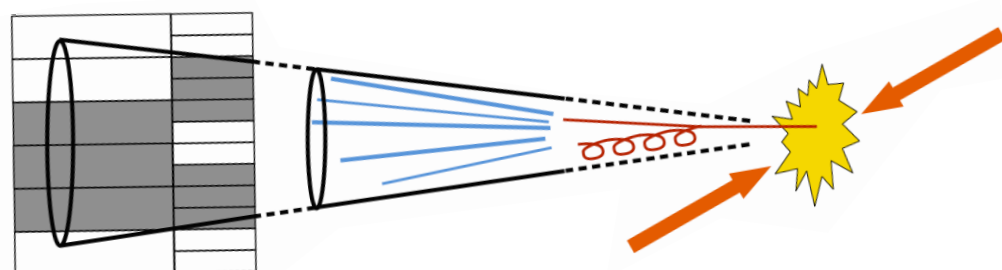
First Principles
QCD Calculation:



$$\dim_{IJ}(Q) \simeq \frac{2\alpha_s}{\pi} (C_I + C_J) \ln \frac{p_T/2}{Q}$$

Color Factors of QCD!

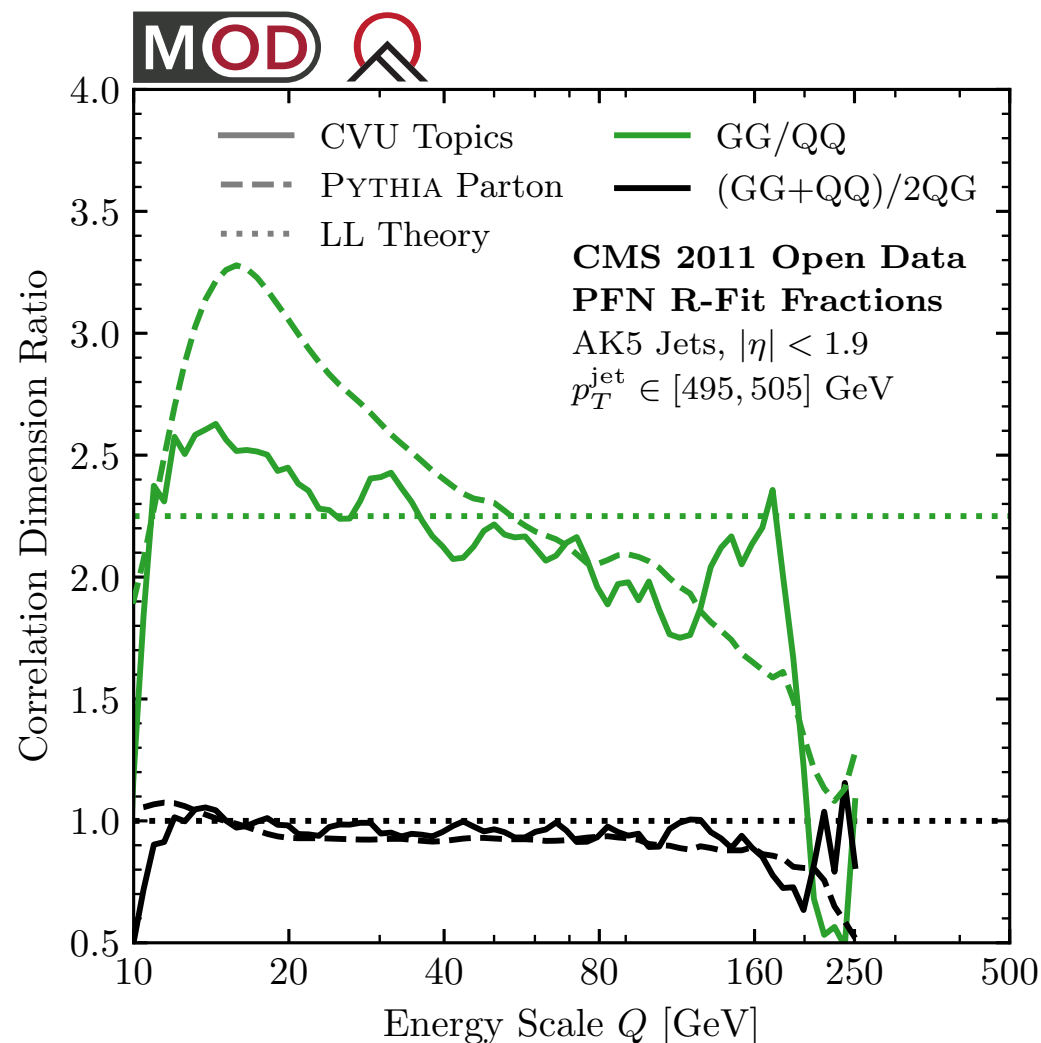
$$dP_{i \rightarrow ig} \simeq \frac{2\alpha_s}{\pi} C_i \frac{d\theta}{\theta} \frac{dz}{z}$$



[Komiske, Kryhin, JDT, arXiv 2022]



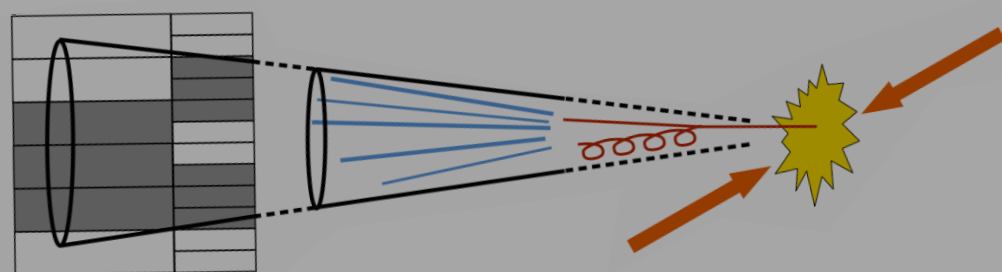
Dimensionality of Quark/Gluon Jets



Fully data-driven extraction of
 quark/gluon color ratio

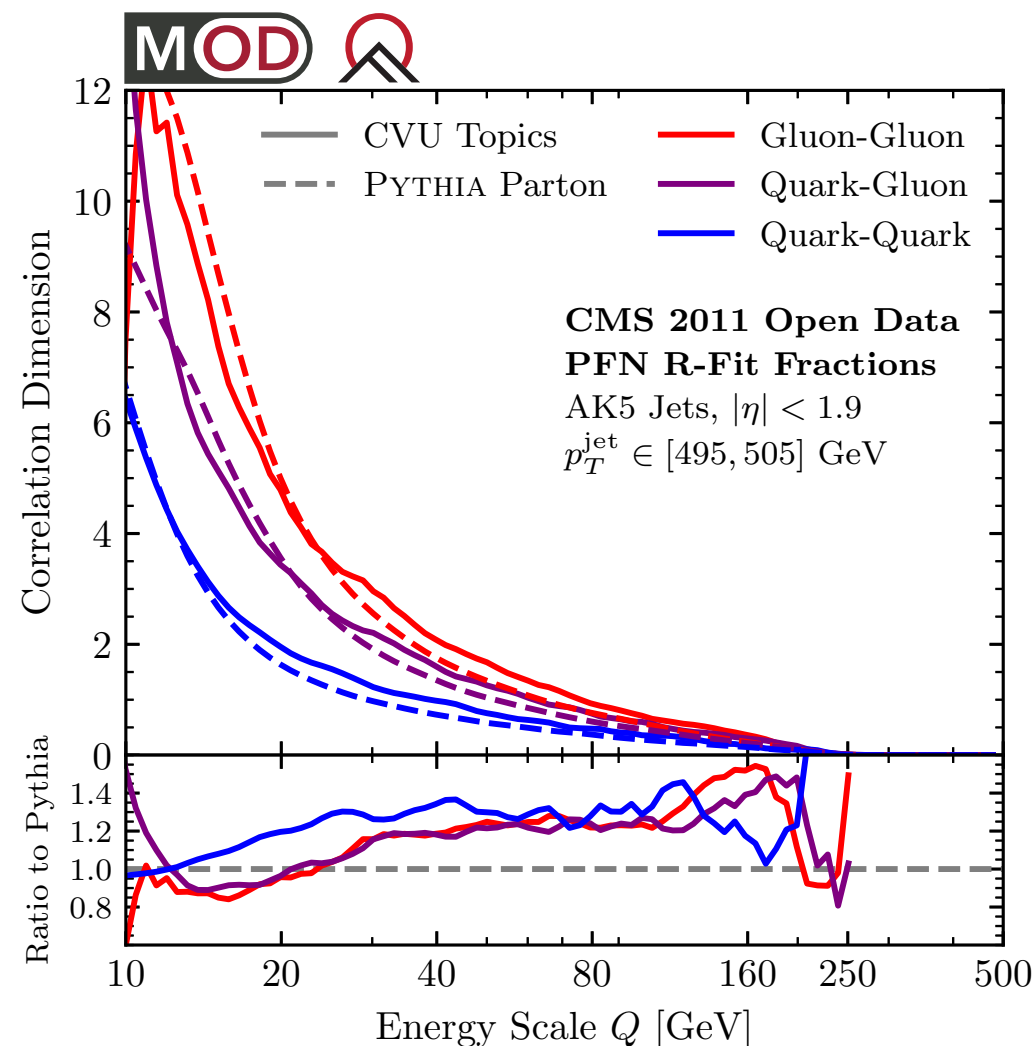
$$\frac{9}{4} = \frac{3}{4/3}$$

$$1 = \frac{(3) + (4/3)}{(3 + 4/3)}$$



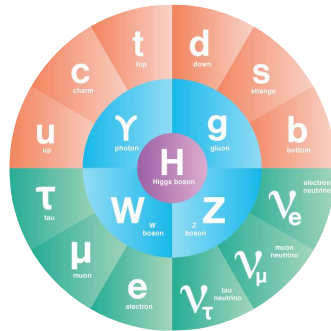
[Komiske, Kryhin, JDT, arXiv 2022]





*We gained new insights into strong force by
 fusing **advances in machine learning** with
 insights from quantum field theory*

Summary



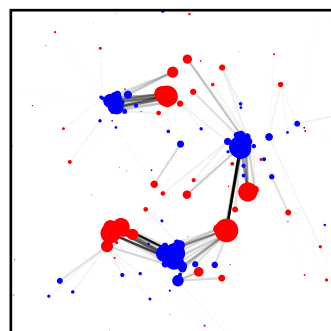
2017: A Quark/Gluon Conundrum

Quark/gluon jets offer an extreme example where fully supervised learning is fundamentally ambiguous



2018: Leveraging Weak Supervision

Assuming that “quark” and “gluon” jet categories exist, one can use machine learning to operationally define them



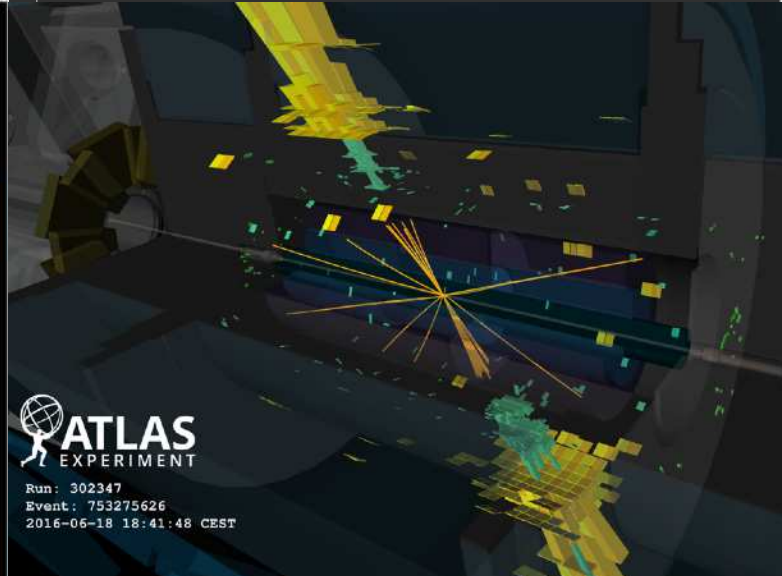
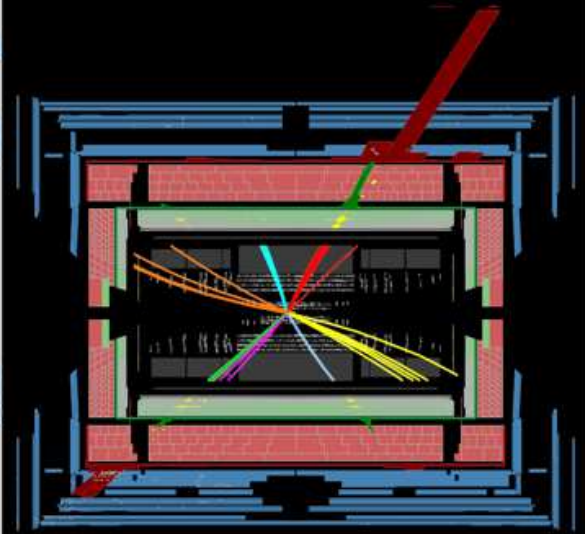
2019-2022: The Strong Force Revisited

Jet physics has crossed an important threshold where machine learning is now yielding insights that go beyond traditional analysis techniques

Backup Slides

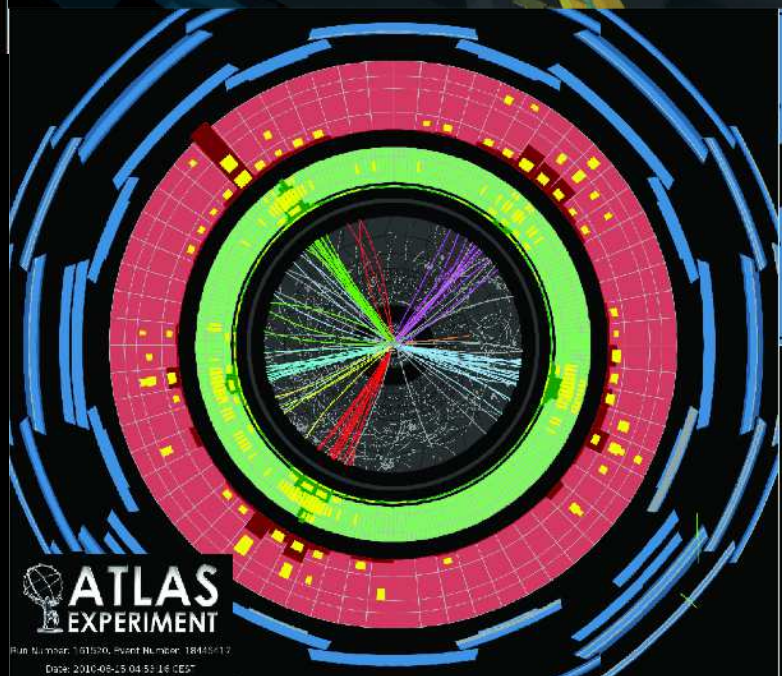
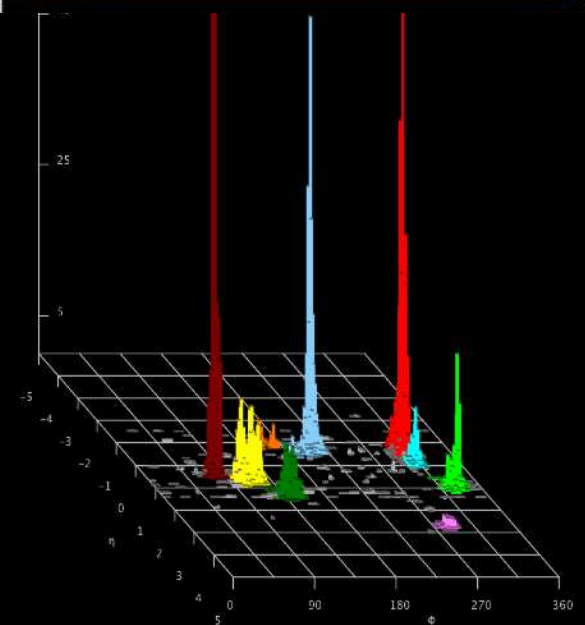
Run Number: 159224, Event Number: 3533152

Date: 2010-07-18 11:05:54 CEST



ATLAS
EXPERIMENT

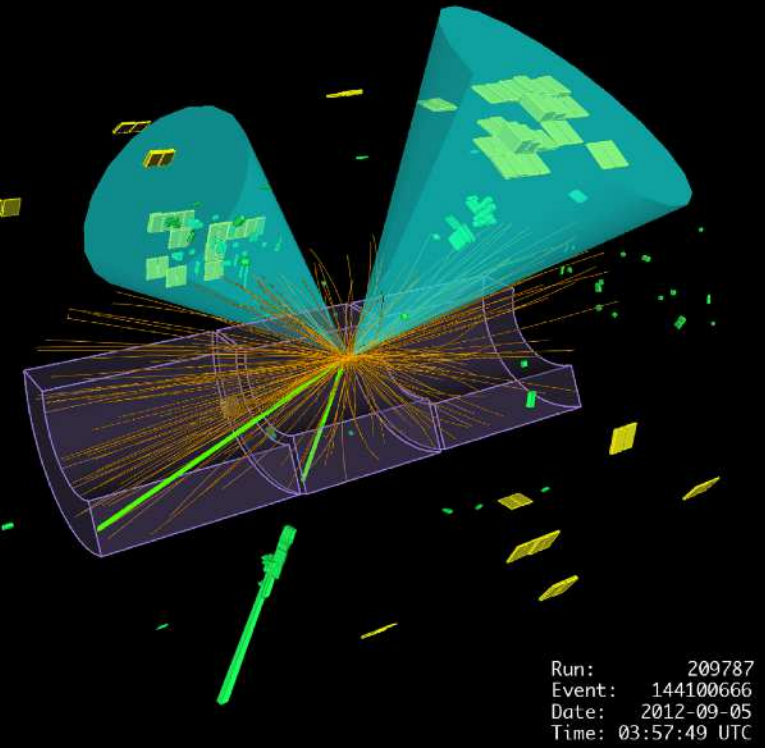
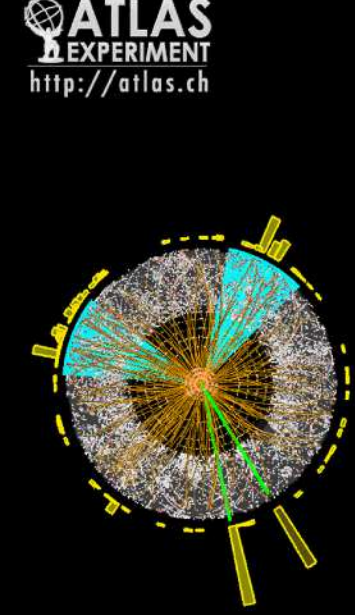
Run: 302347
Event: 753275626
2016-06-18 18:41:48 CEST



ATLAS
EXPERIMENT

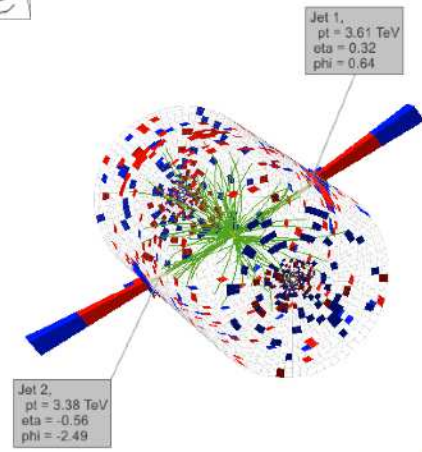
Run Number: 103520, Event Number: 1844512
Date: 2010-08-15 04:53:14 CEST

ATLAS
EXPERIMENT
<http://atlas.ch>



Run: 209787
Event: 144100666
Date: 2012-09-05
Time: 03:57:49 UTC

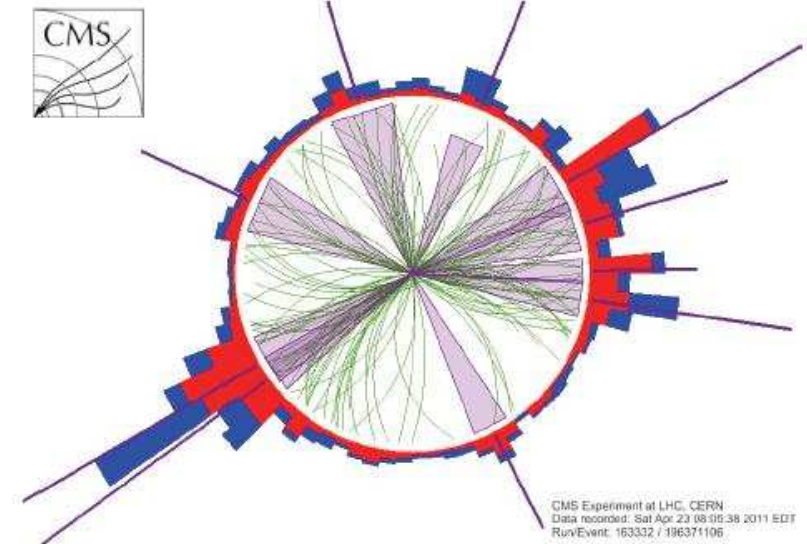
CMS



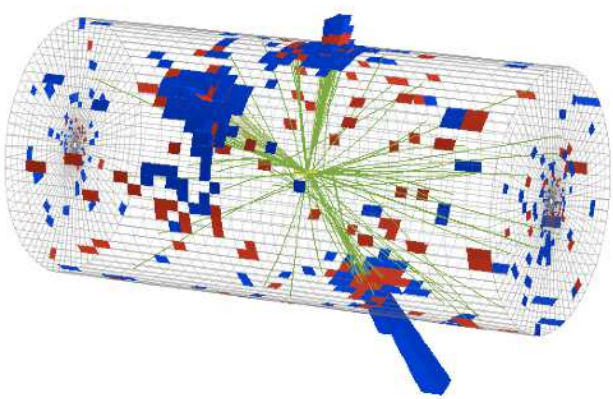
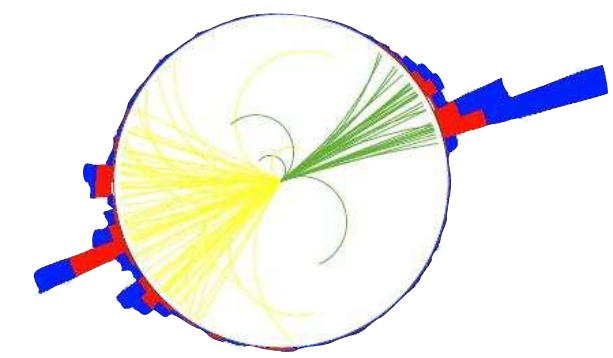
Jet 1,
pt = 3.61 TeV
eta = 0.32
phi = 0.64

Jet 2,
pt = 3.38 TeV
eta = -0.56
phi = -2.49

CMS



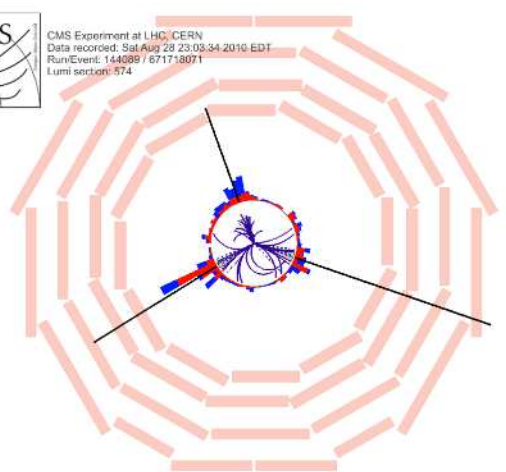
CMS Experiment at LHC, CERN
Data recorded: Sat Apr 23 08:05:38 2011 EDT
Run/Event: 163332 / 196371106



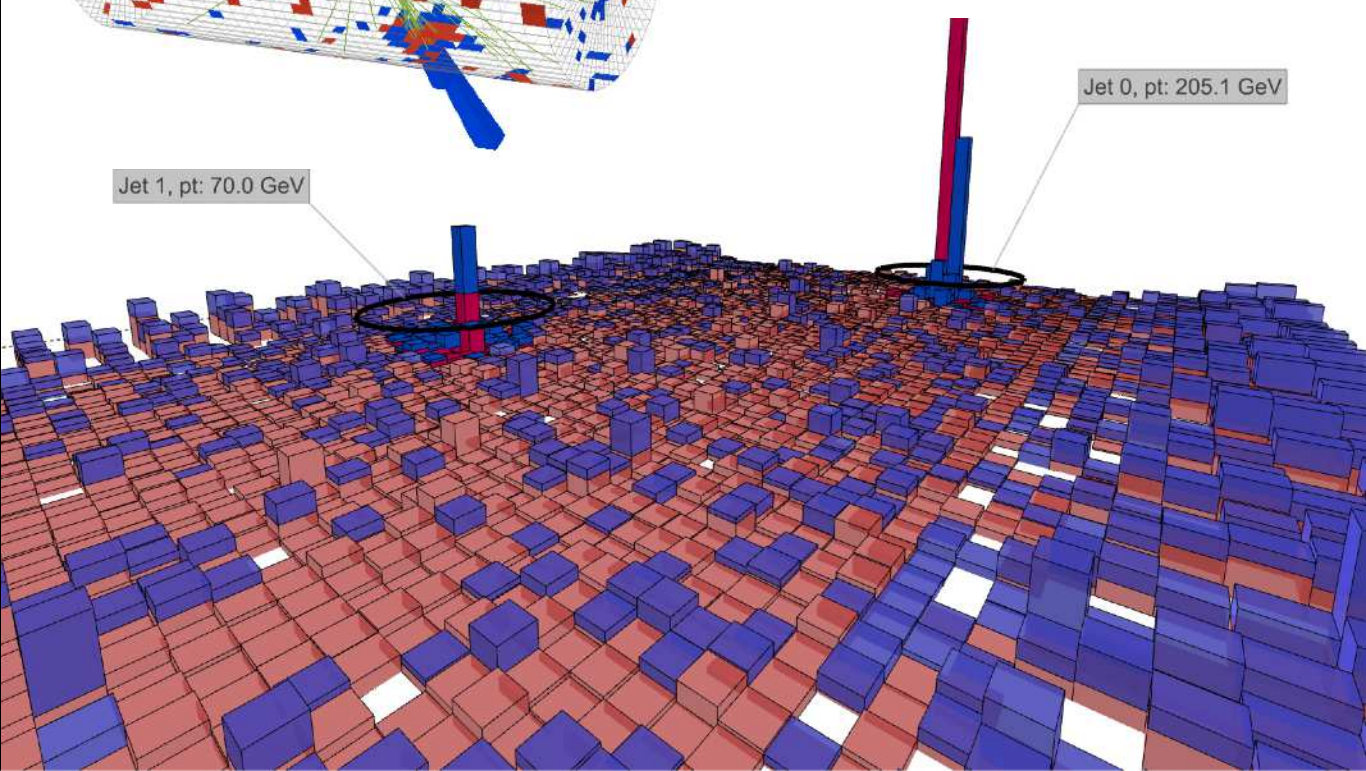
Jet 1, pt: 70.0 GeV

CMS

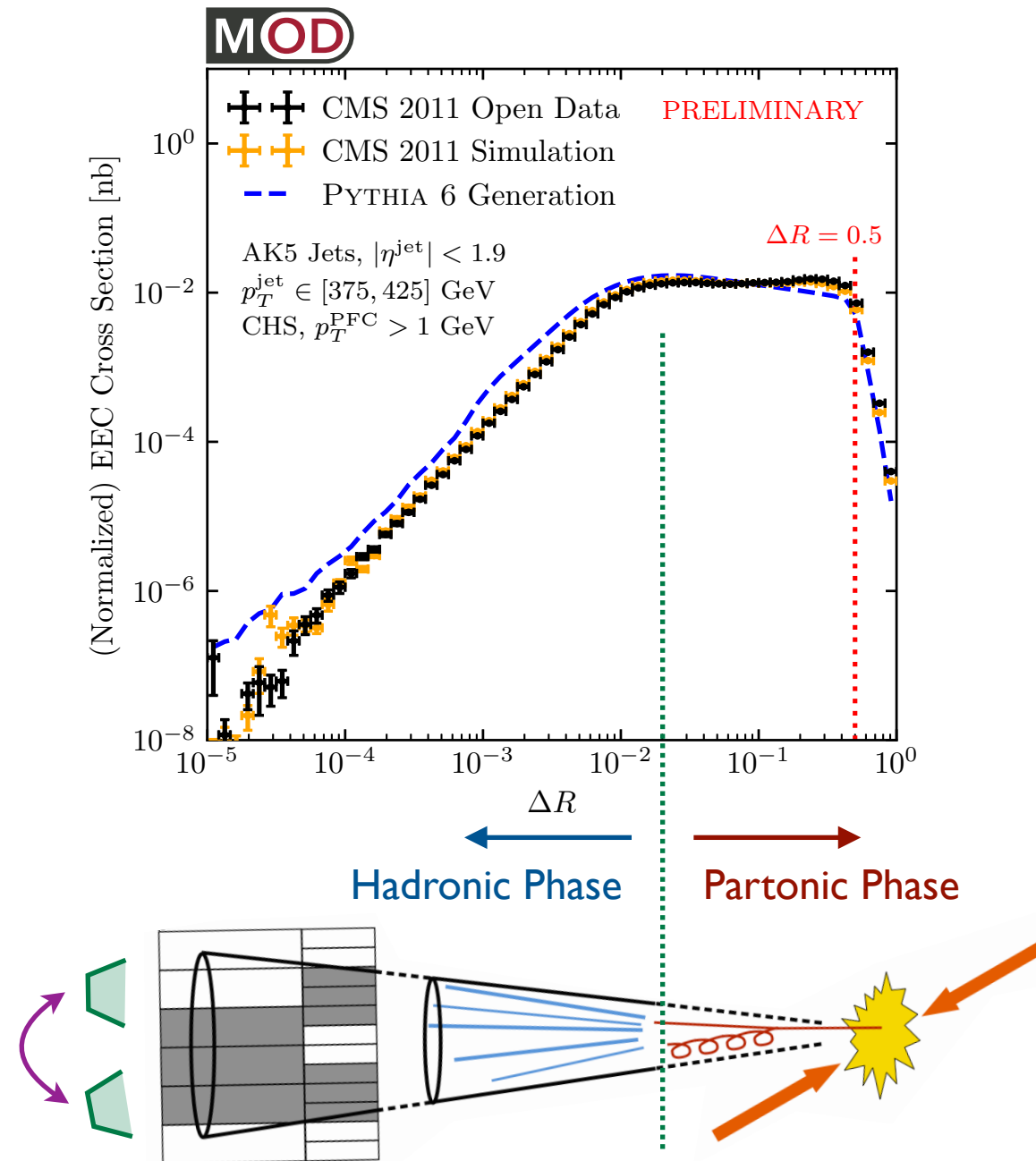
CMS Experiment at LHC, CERN
Data recorded: Sat Aug 28 23:03:34 2010 EDT
Run/Event: 144089 / 671718071
Lumi section: 574



Jet 0, pt: 205.1 GeV



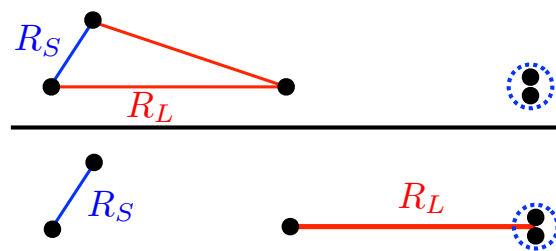
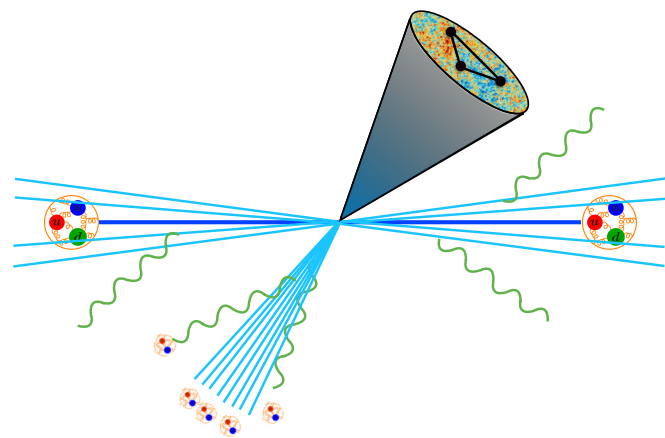
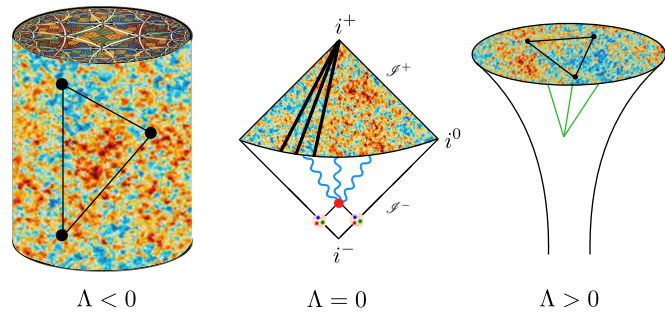
QCD through the Lens of Condensed Matter



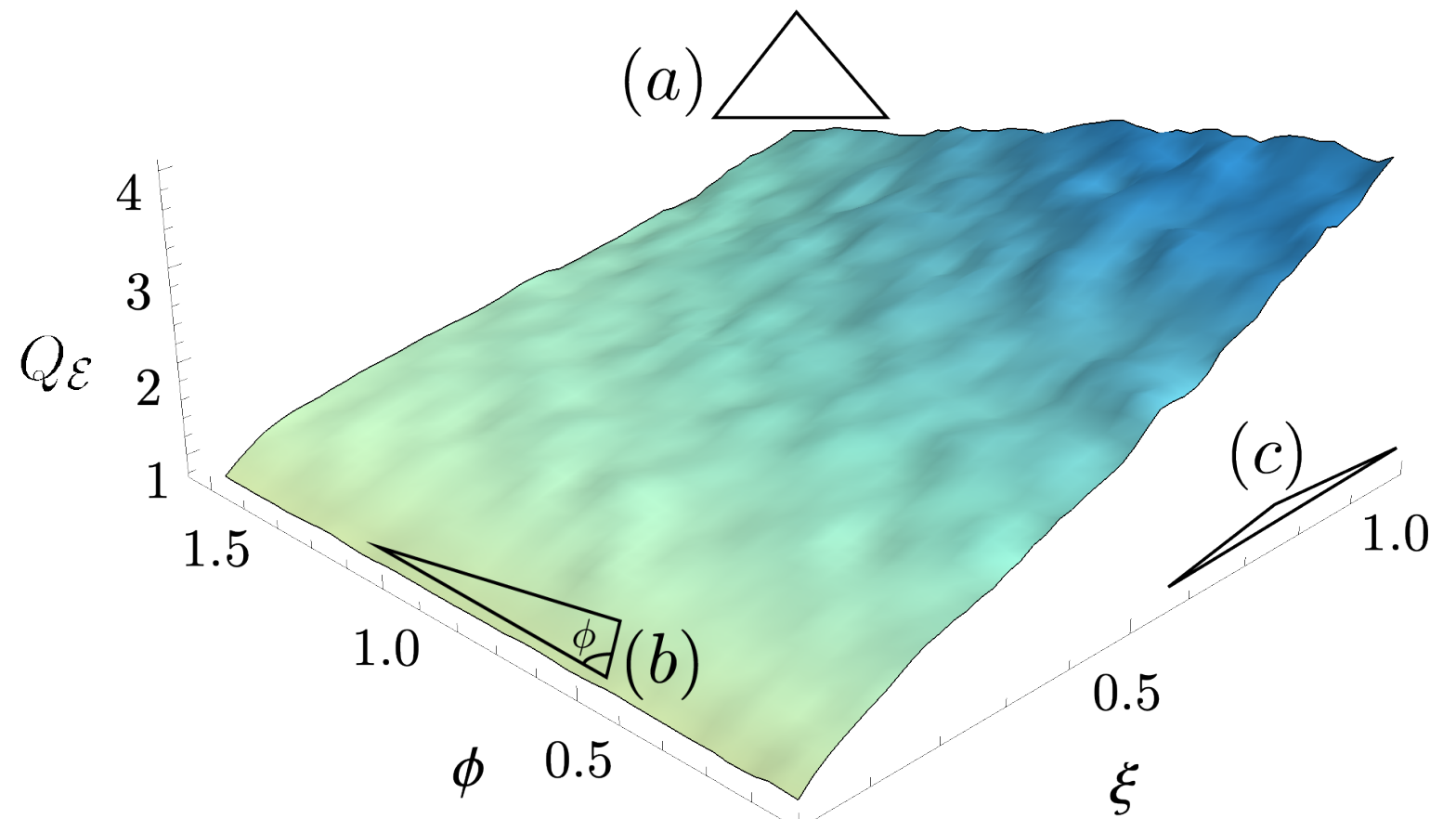
[Komiske, Moul, JDT, Zhu, [arXiv 2022](#); see talks by Moul, [BOOST 2019](#), [BOOST 2020](#)]



QCD through the Lens of Cosmology



CMS Open Data, $R_L \in (0.3, 0.4)$



[Chen, Moulton, JDT, Zhu, [arXiv 2022](#)]





Patrick Komiske



Eric Metodiev



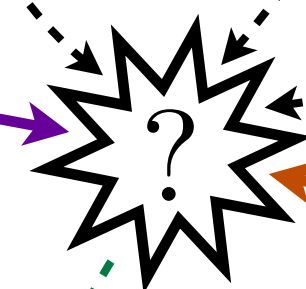
Principles of
Fundamental
Physics

Robustness of Energy Flow

[Komiske, Metodiev, JDT, JHEP 2018]

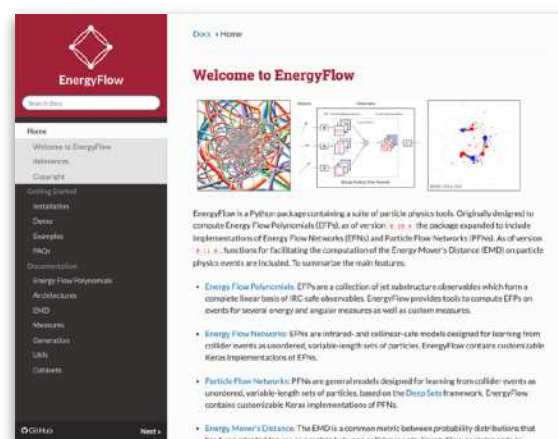
Power of
Artificial
Intelligence

Point Cloud Learning



Energy Flow Networks

<https://energyflow.network/>
[Komiske, Metodiev, JDT, JHEP 2019]

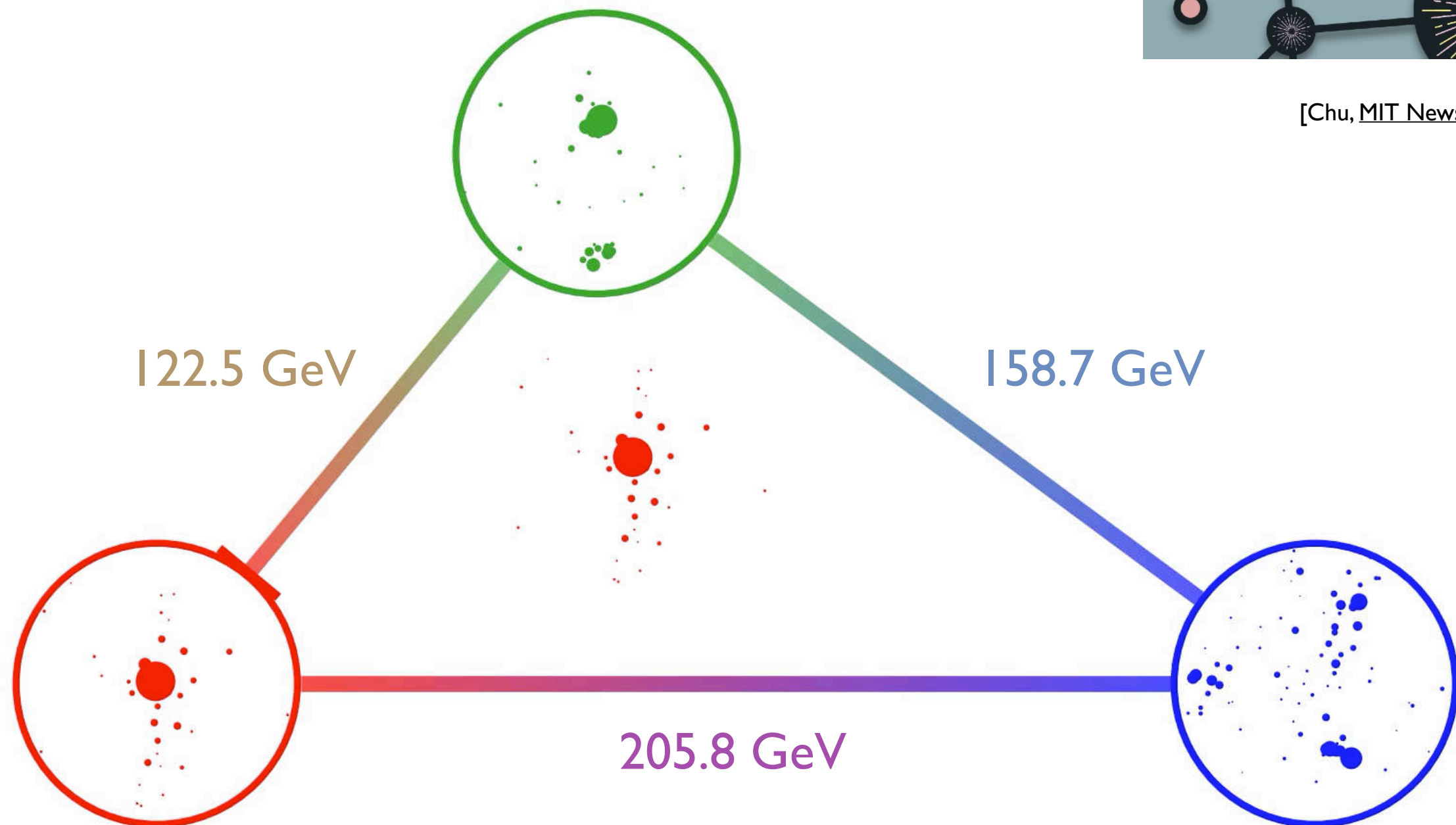


[Zaheer, Kottur, Ravanbakhsh, Póczos,
Salakhutdinov, Smola, NIPS 2017]

Triangulating the Space of Jets

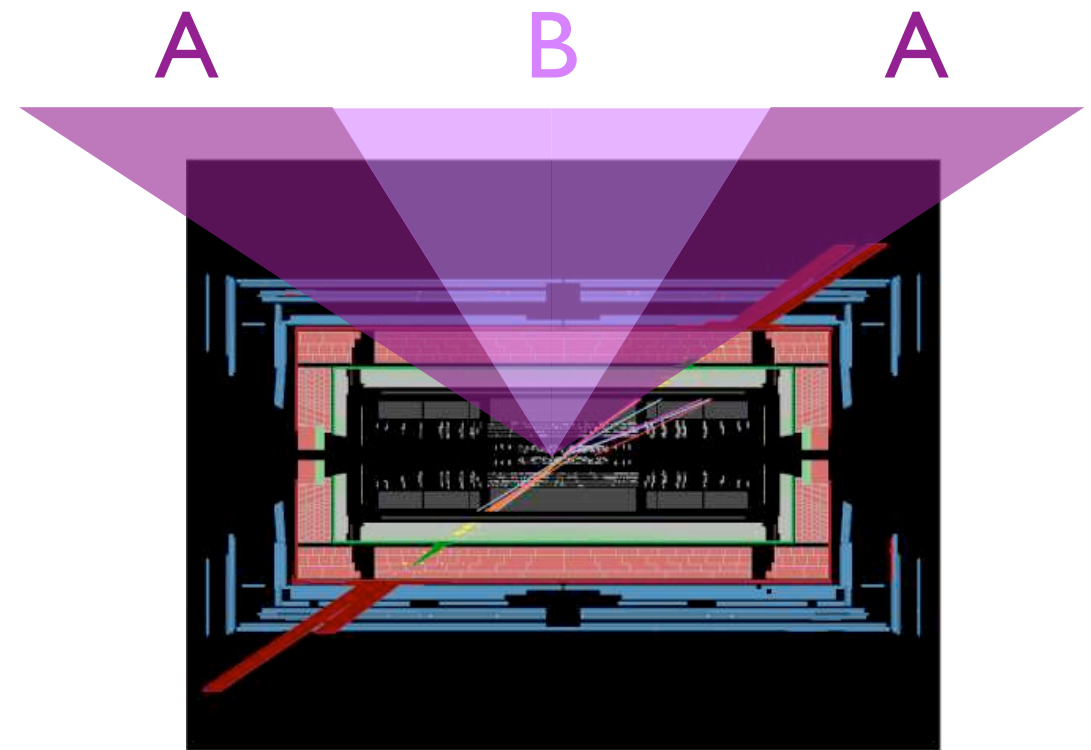
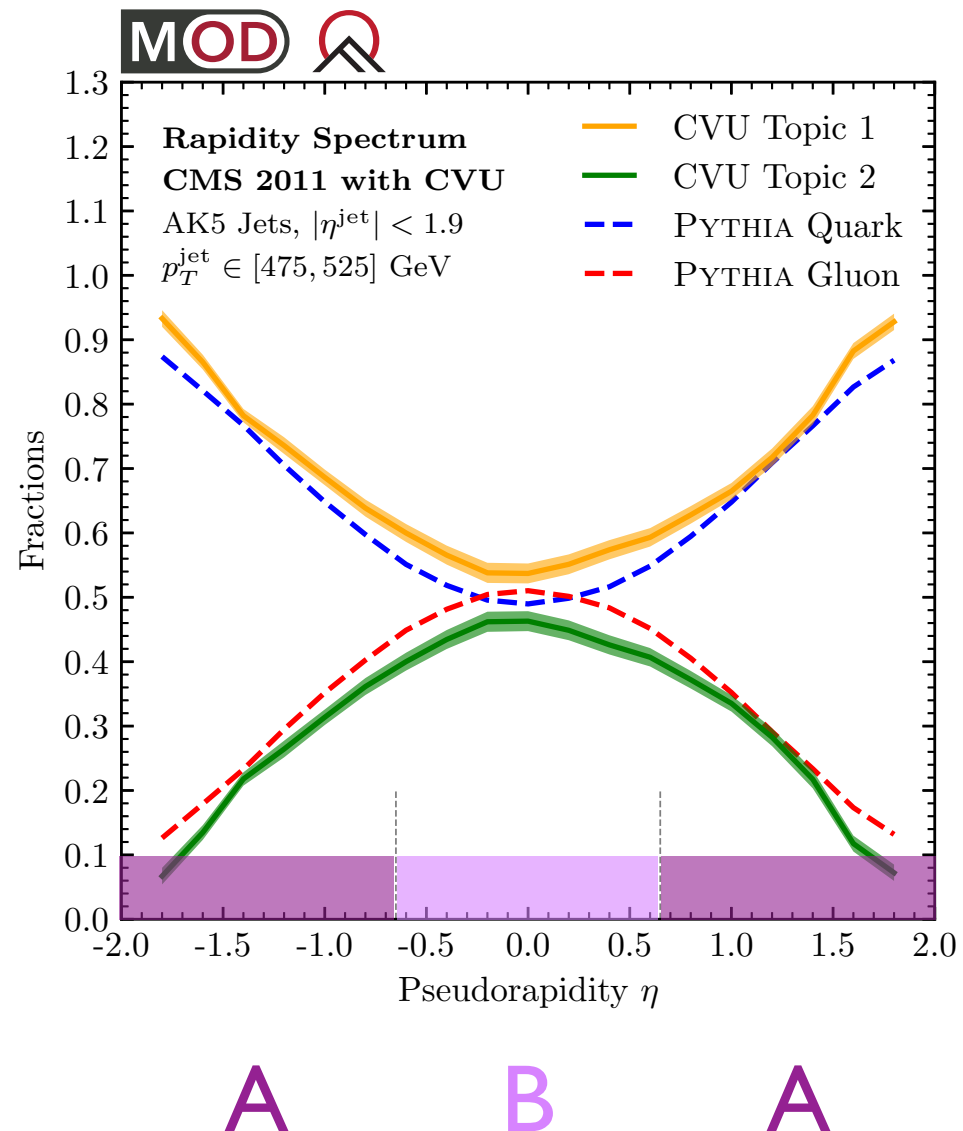


[Chu, [MIT News July 2019](#)]



[Komiske, Metodiev, JDT, [PRL 2019](#); code at Komiske, Metodiev, JDT, [energyflow.network](#);
see alternative graph network approach in Mullin, Pacey, Parker, White, Williams, [JHEP 2021](#)]

Rapidity-Dependent Quark/Gluon Fraction

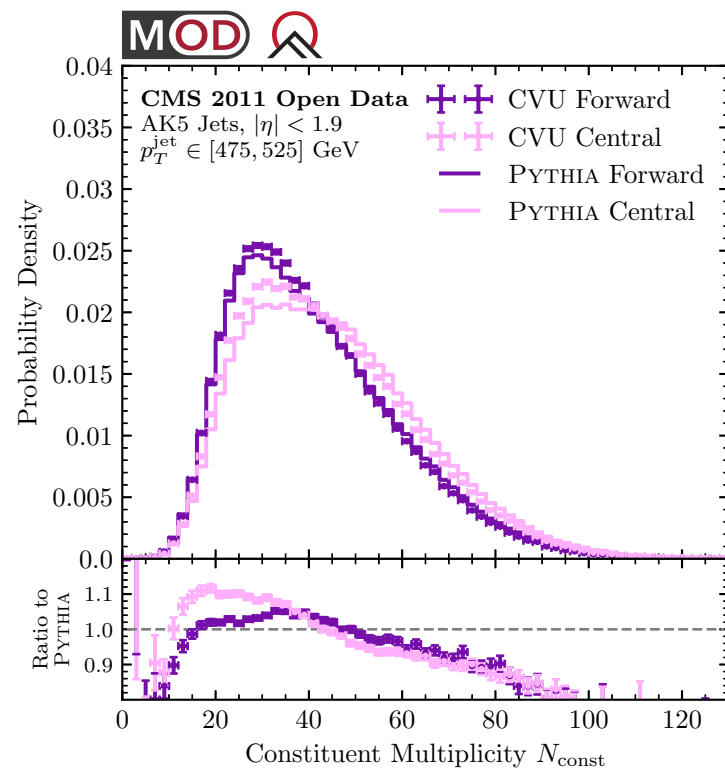


[Komiske, Kryhin, JDT, [arXiv 2022](#)]

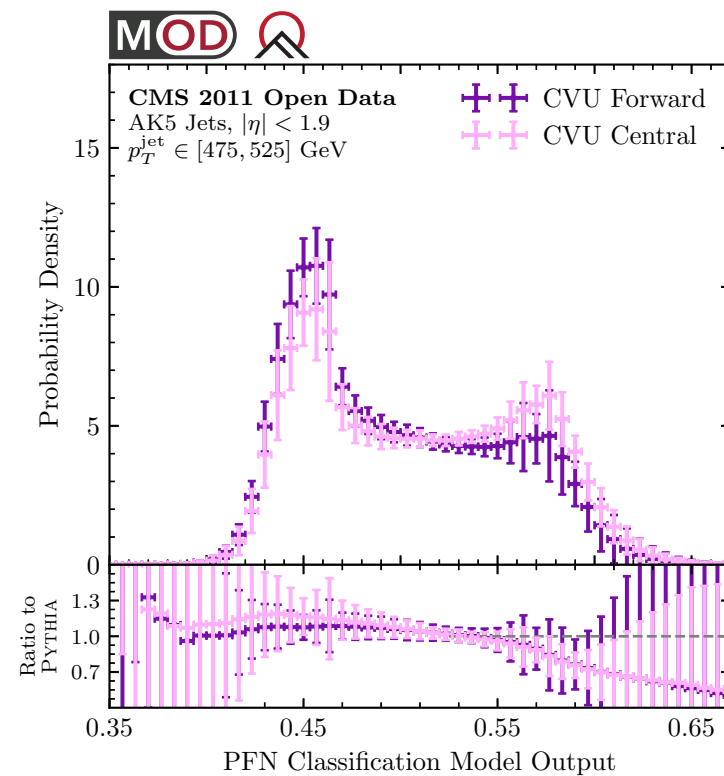


Forward/Central Differences

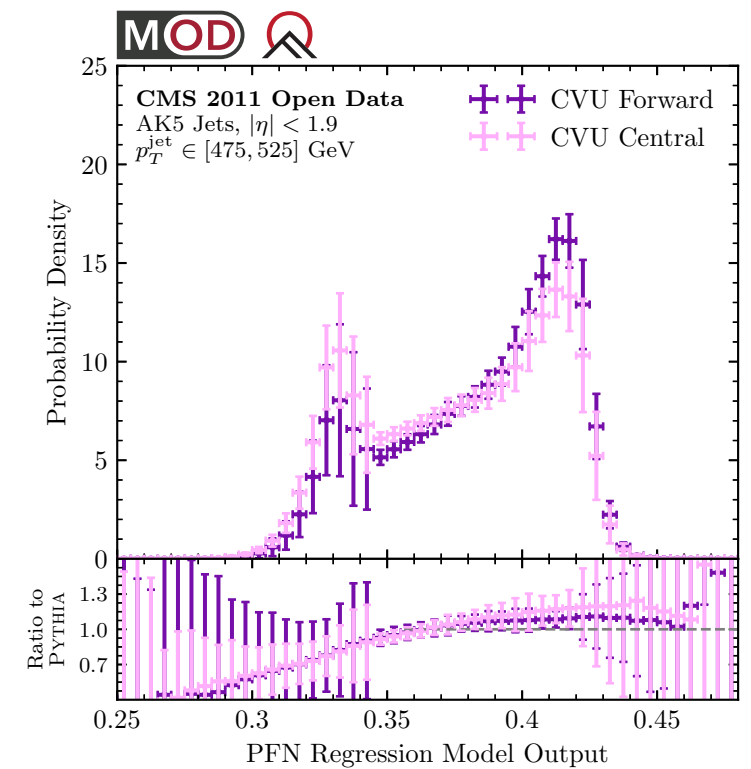
Multiplicity



Weak Classification



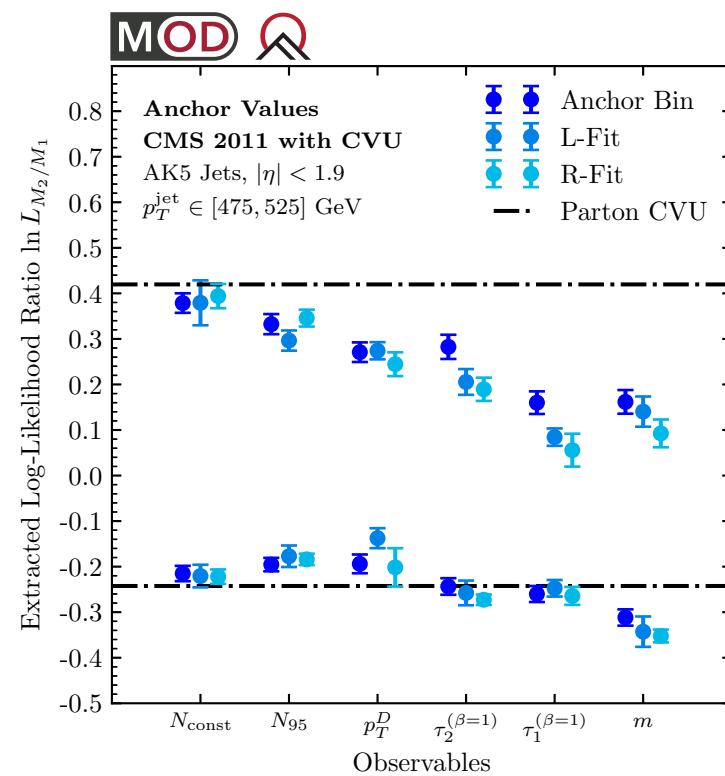
Weak Regression



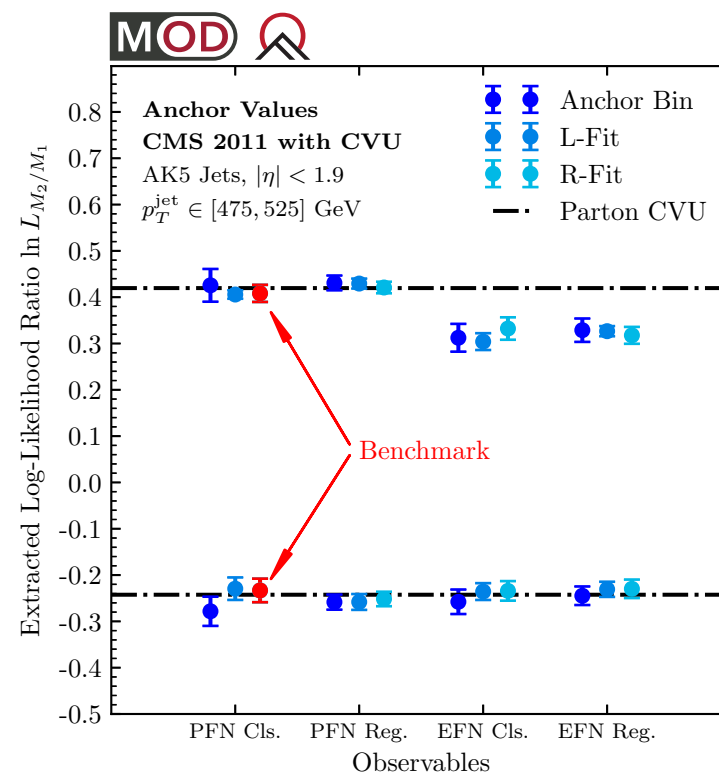
[Komiske, Kryhin, JDT, [arXiv 2022](#)]

Extracted Fractions with Different Observables

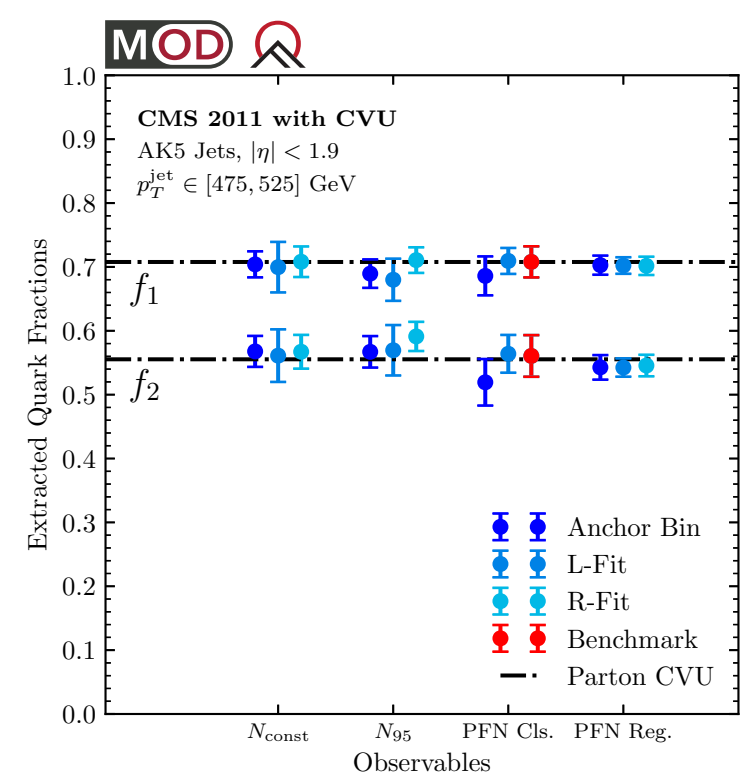
Traditional Observables



Weak Supervision

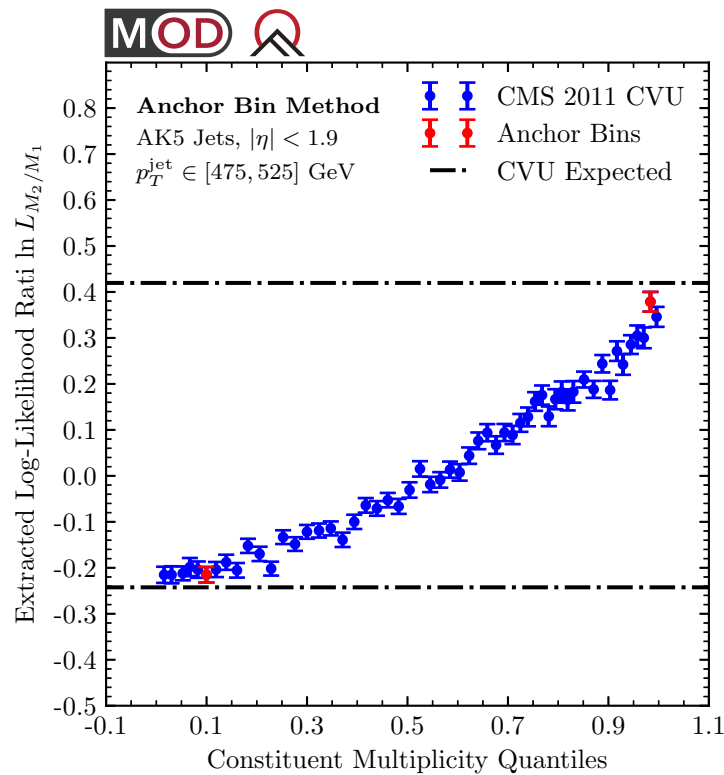


Quark Fractions

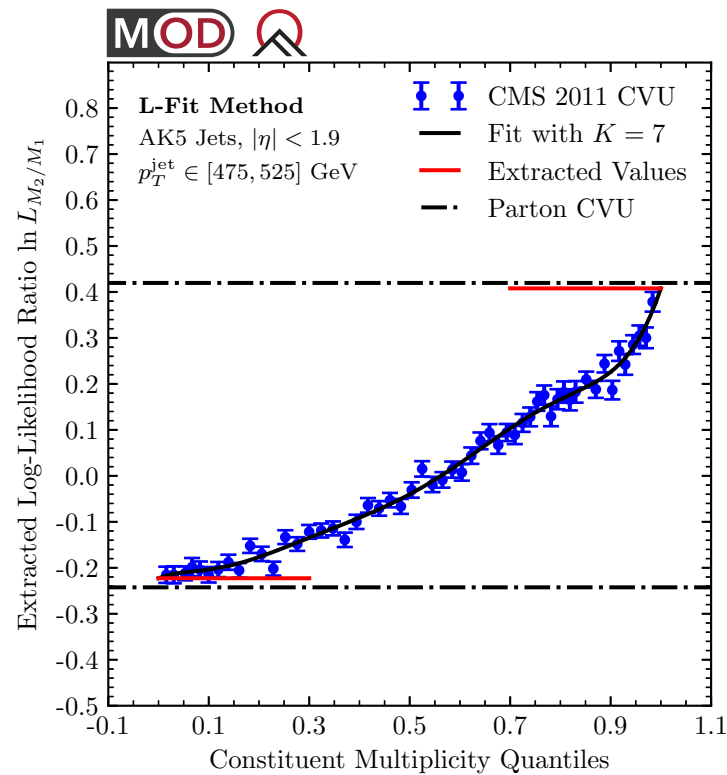


Three Different Strategies for K

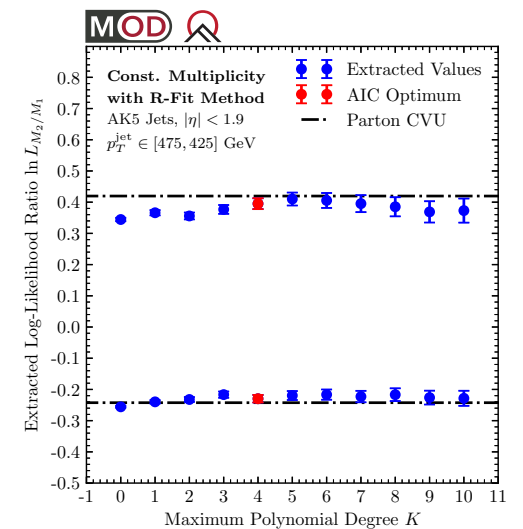
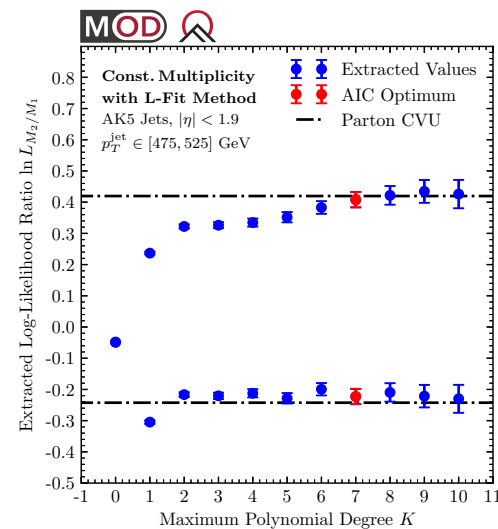
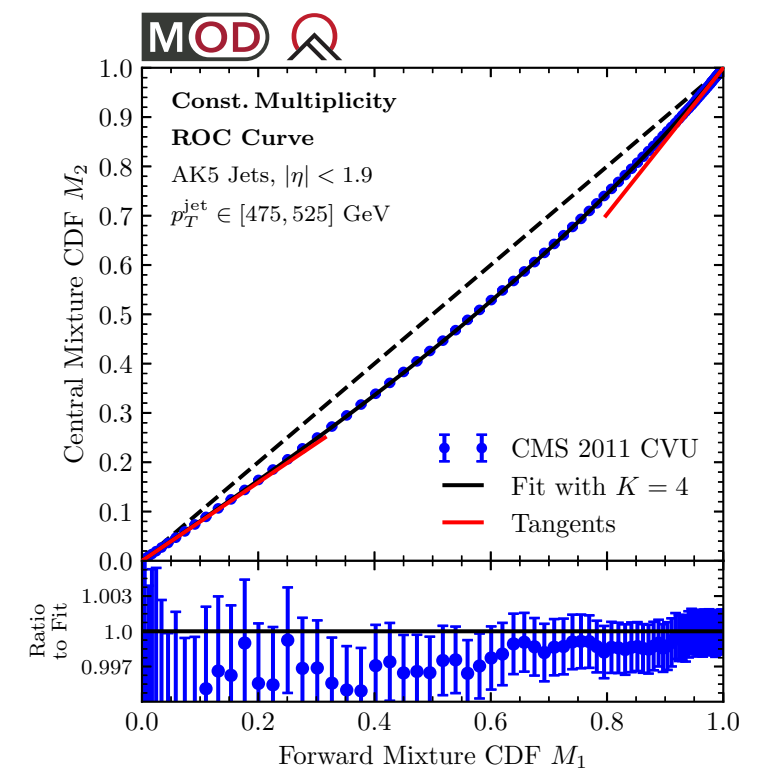
Anchor Bin



Likelihood Ratio Fit

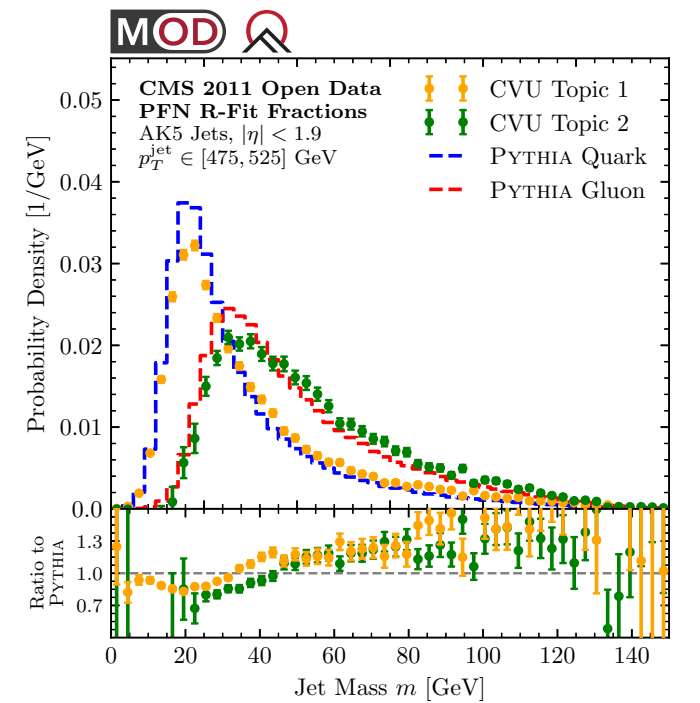
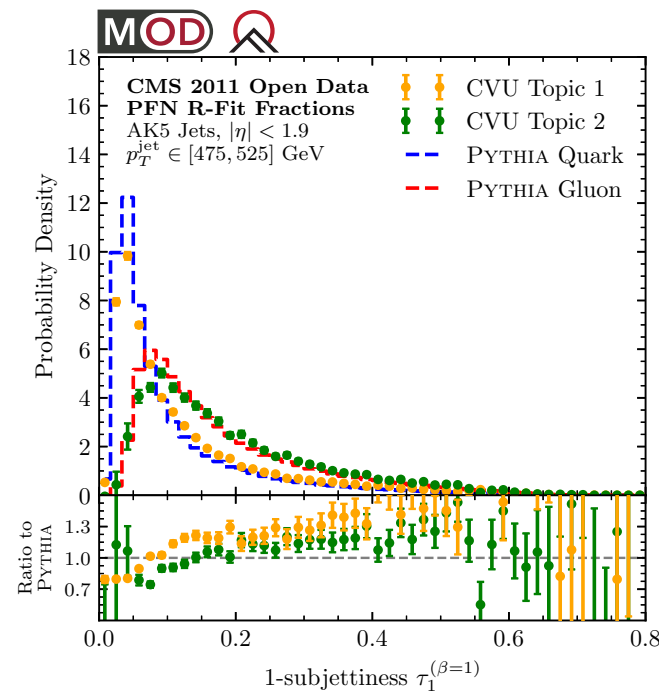
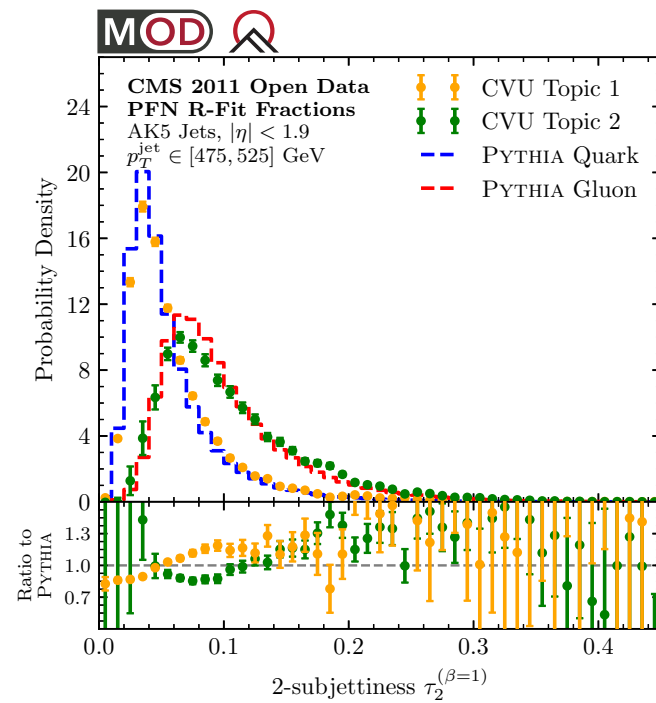
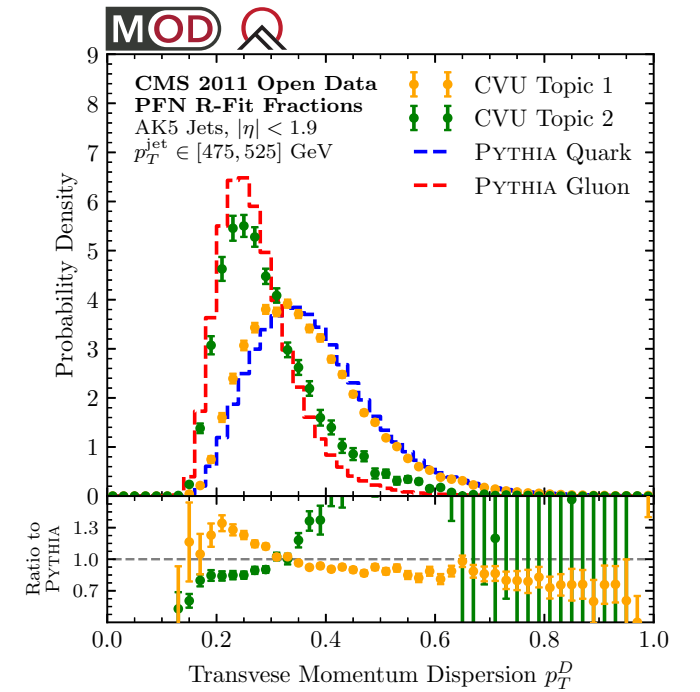
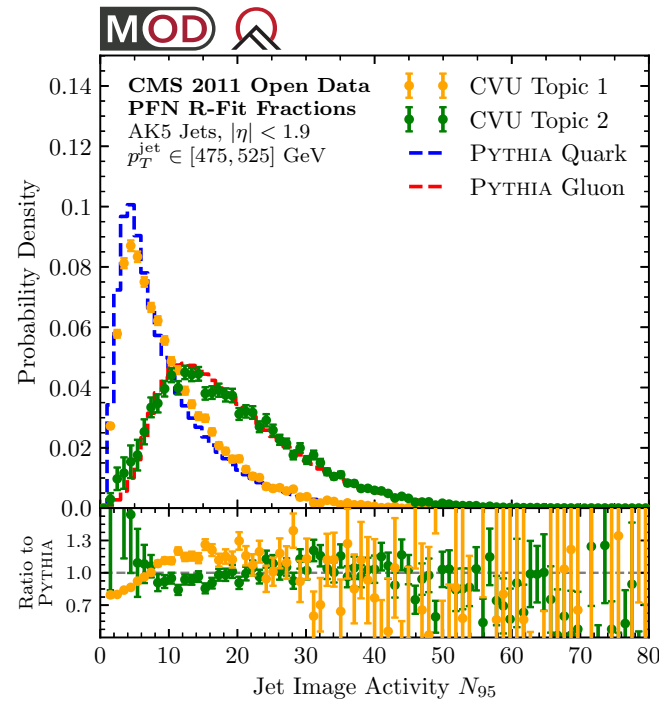
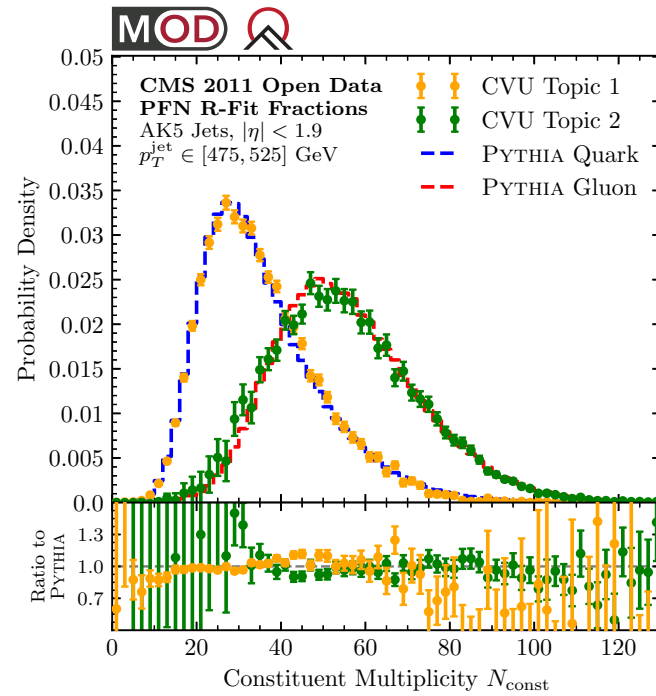


ROC Curve Fit



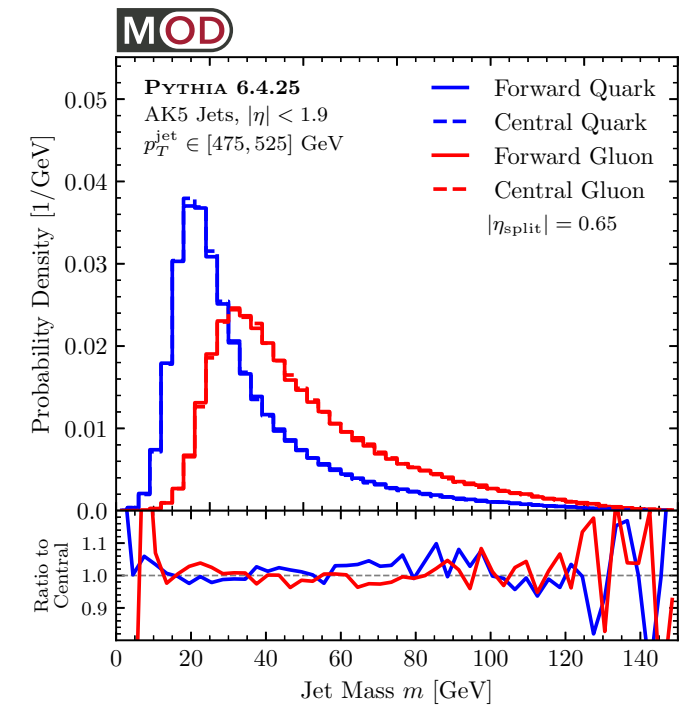
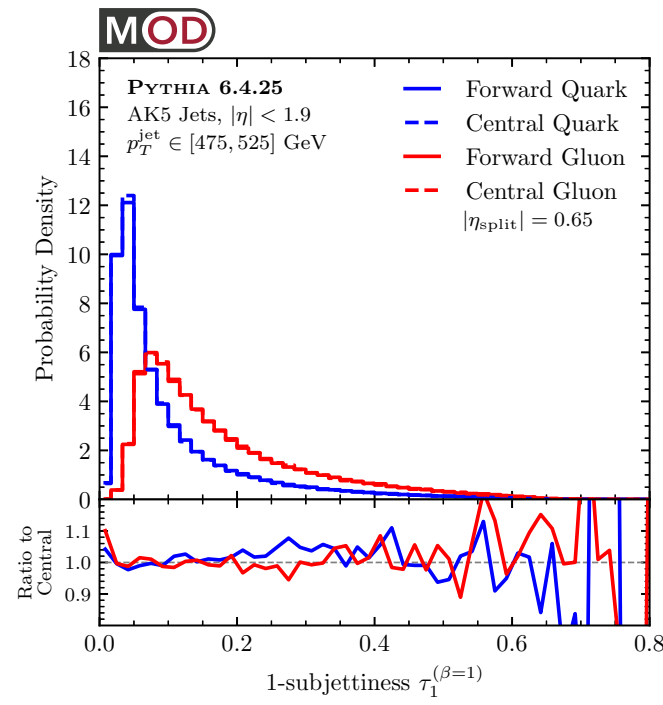
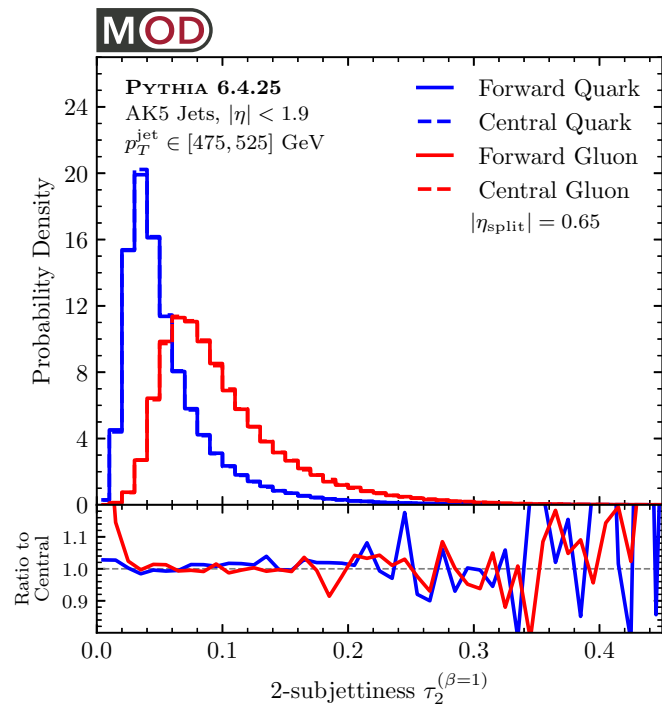
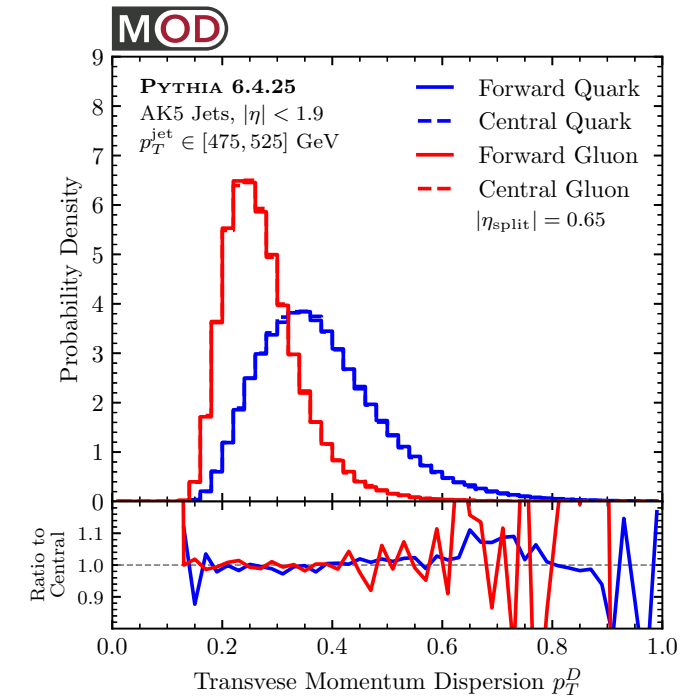
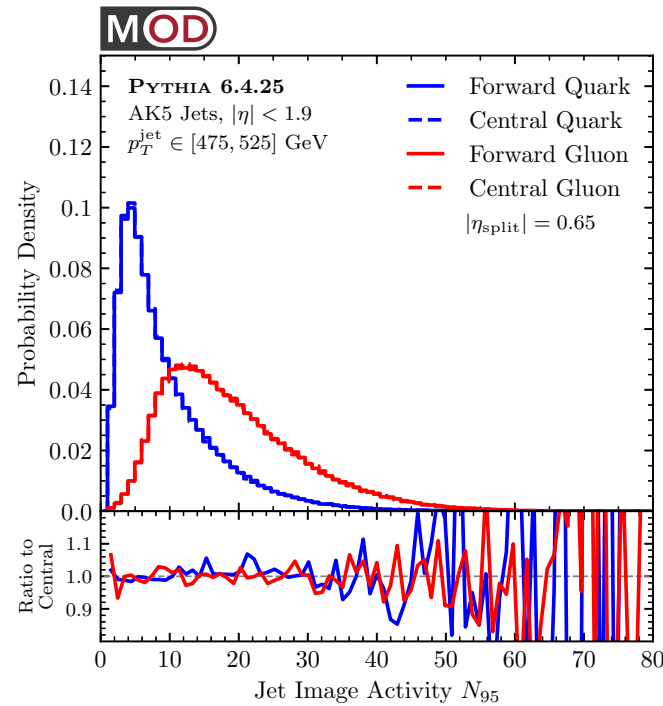
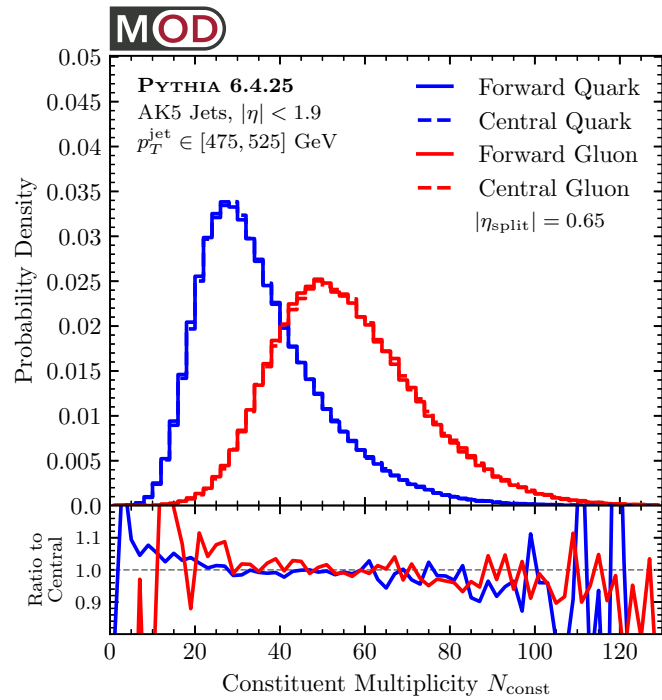
[Komiske, Kryhin, JDT, arXiv 2022]

Six Traditional Observables



[Komiske, Kryhin, JDT, arXiv 2022]

Testing Sample Independence



[Komiske, Kryhin, JDT, arXiv 2022]

Unfolding Checks

