Machine Learning and Heavy Flavor Jets in Heavy Ion Collisions

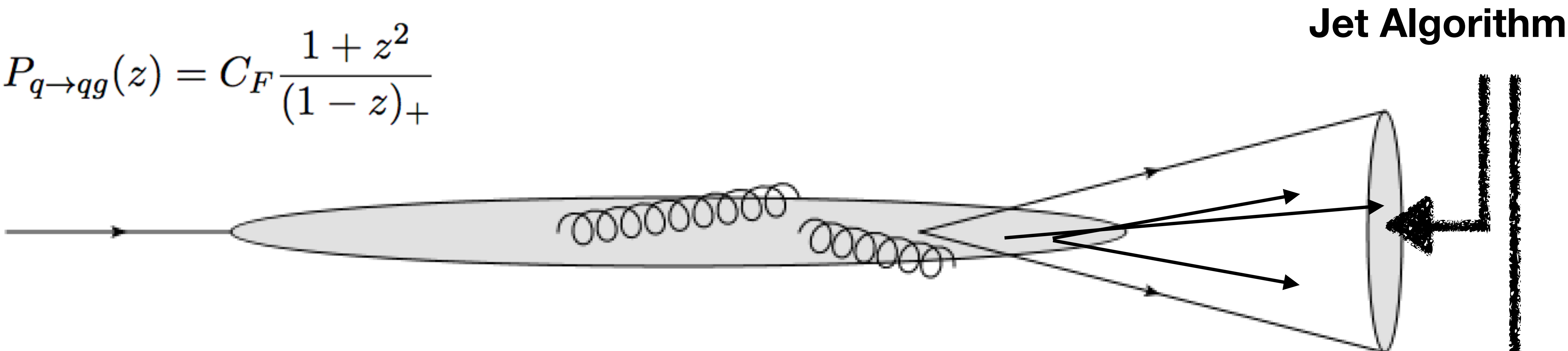
Raghav Kunnawalkam Elayavalli
Yale/BNL

Summer student lectures @ Prague
26-28 June 2022

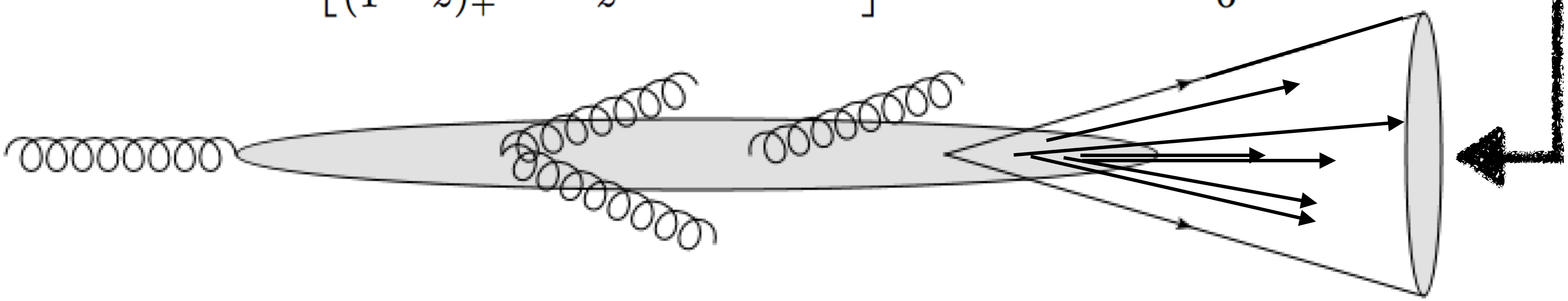
raghavke.me

Quark and Gluon Jets

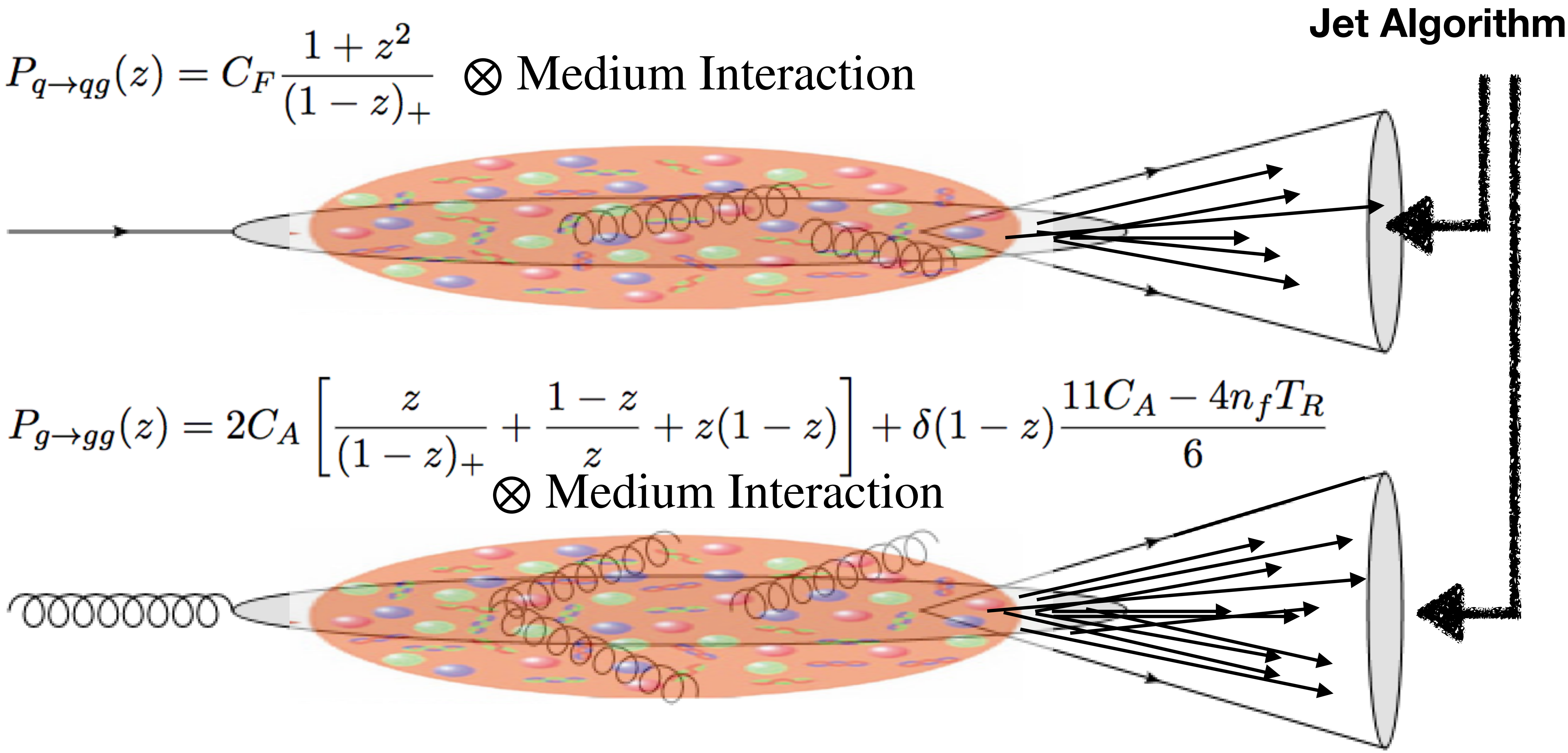
$$P_{q\to qg}(z) = C_F \frac{1+z^2}{(1-z)_+}$$



$$P_{g\to gg}(z) = 2C_A\left[\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z)\right] + \delta(1-z)\frac{11C_A - 4n_f T_R}{6}$$



Quark and Gluon Jets

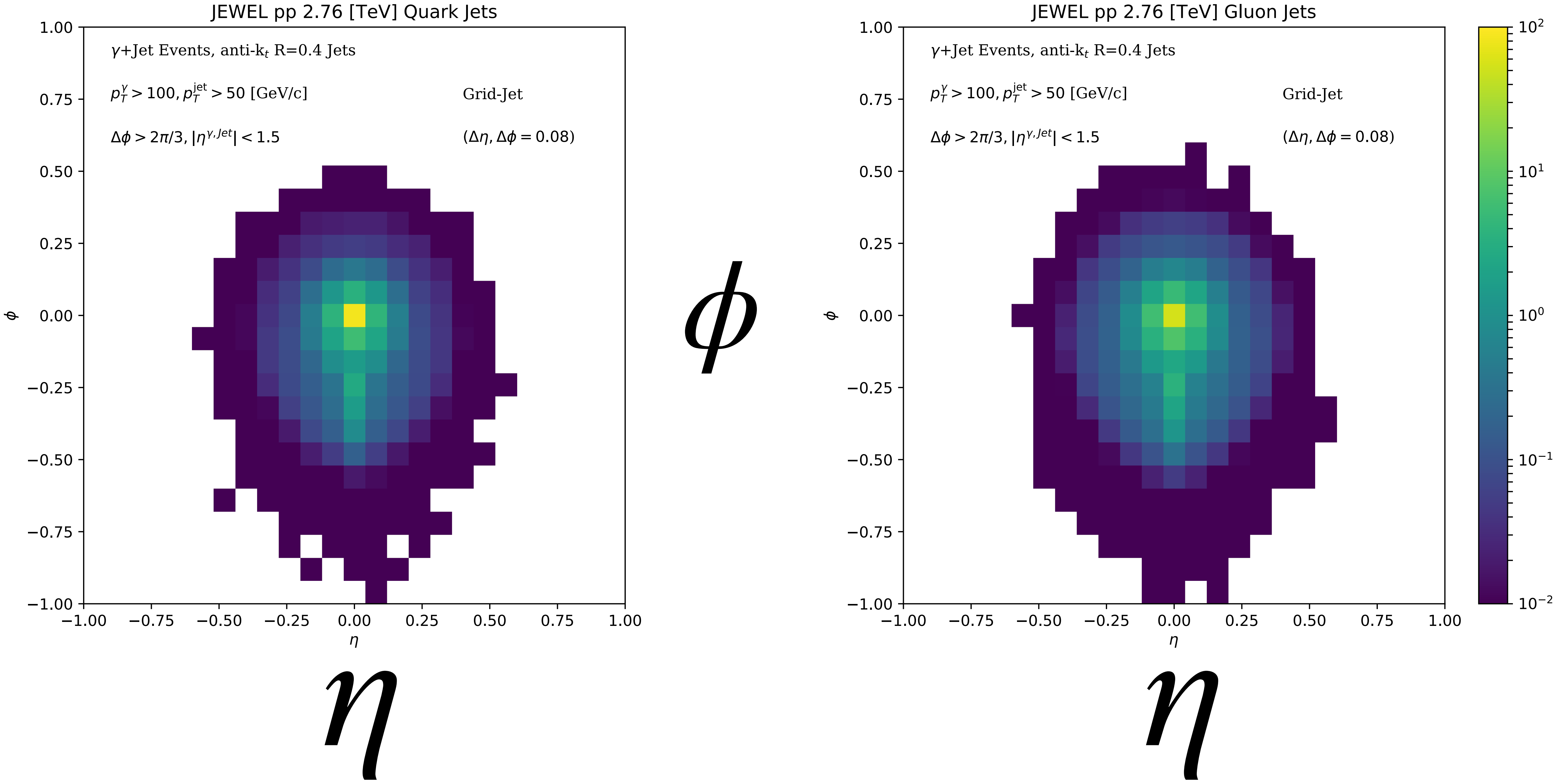


Effects of the QGP on jet propagation manifests via modifications to jet energy and jet sub-structure

Jet Images

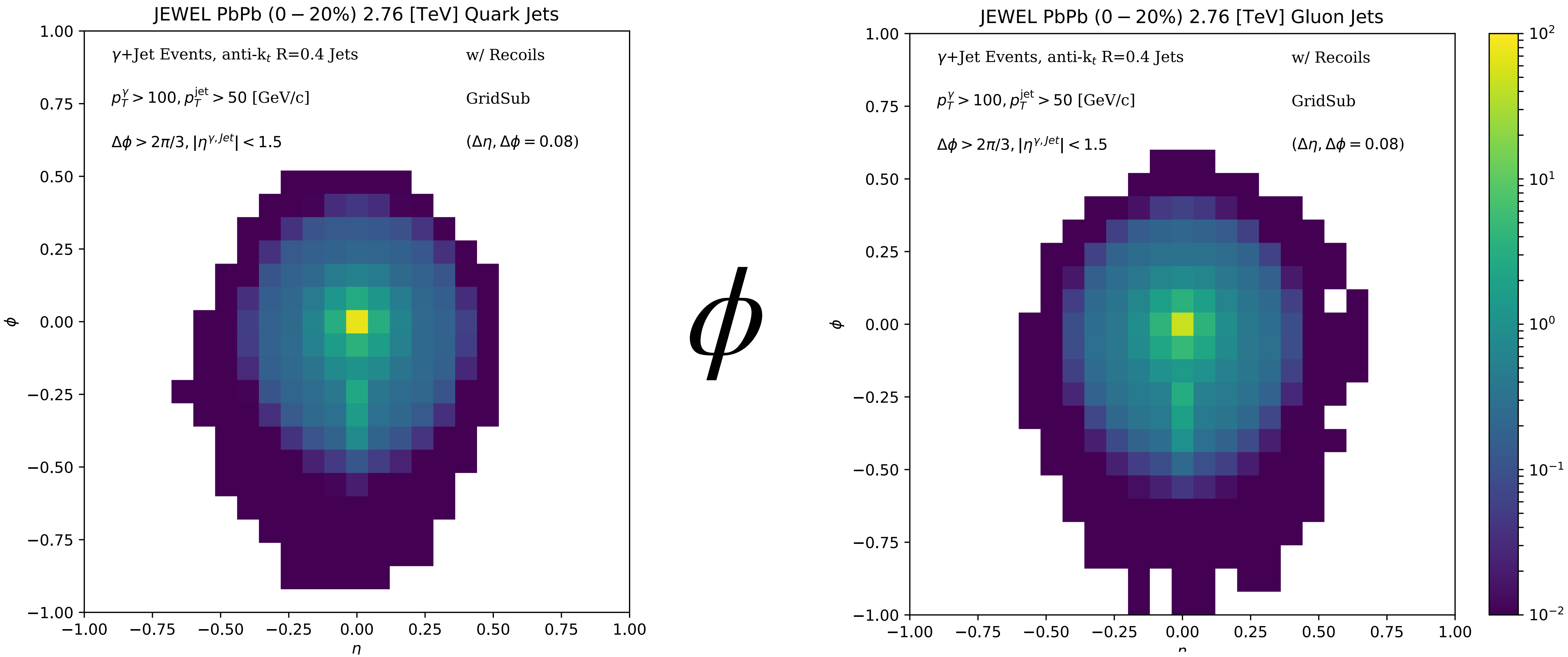
PP Q-Jet → **PP G-Jet**

QCD Color factor
G-Jets are broader



Jet Images

Quenched Q-Jet → **Quenched G-Jet**

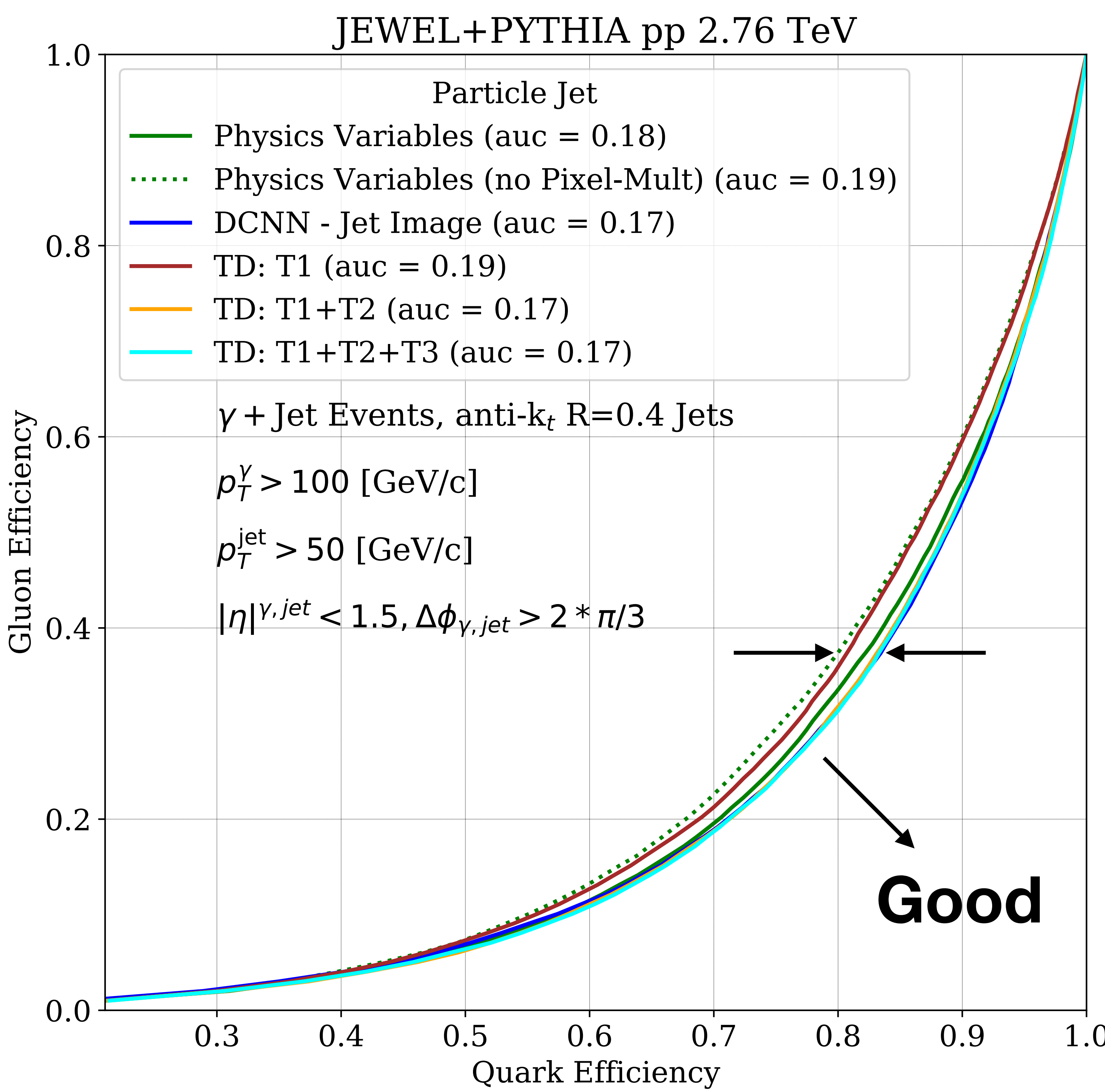


ϕ

η η

Quenched Q-Jet looks similar to G-Jet
G-Jet Still broader

ROC curve for pp Particle Jets

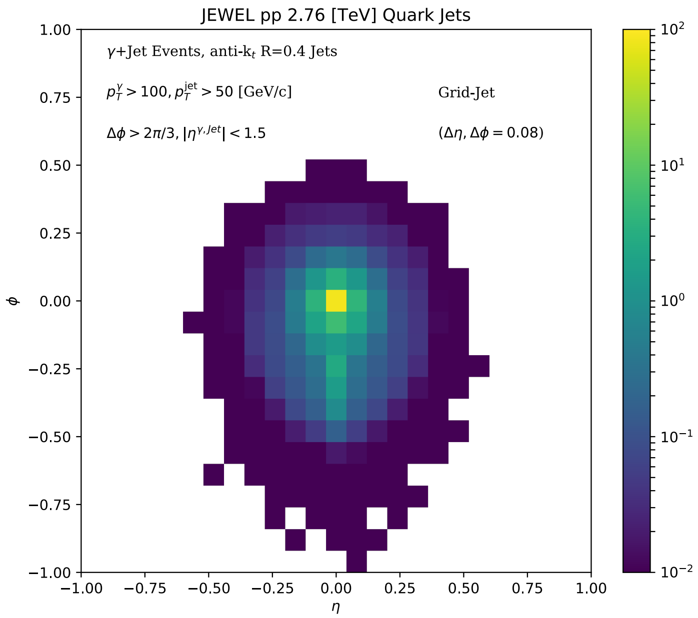


All methods are relatively close to each other

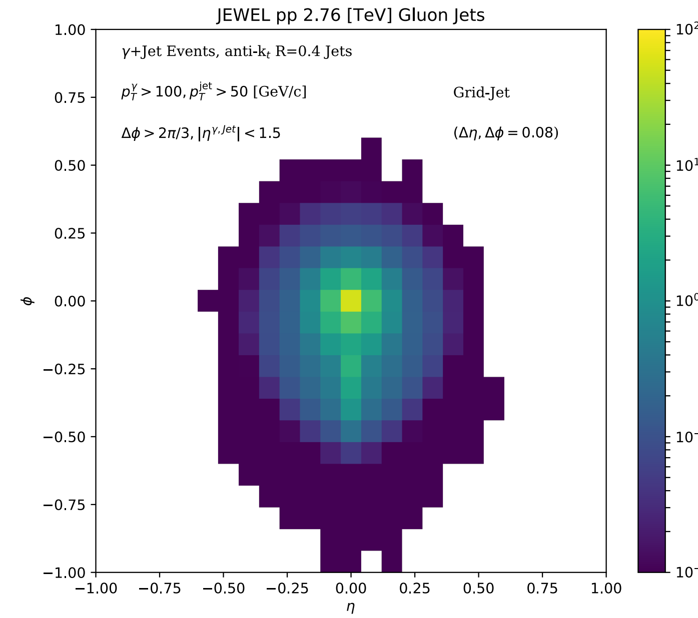
How about Quenched jets?

Jet Images

PP Q-Jet



PP G-Jet



QCD Color factor
G-Jets are broader

JEWEL

Pros

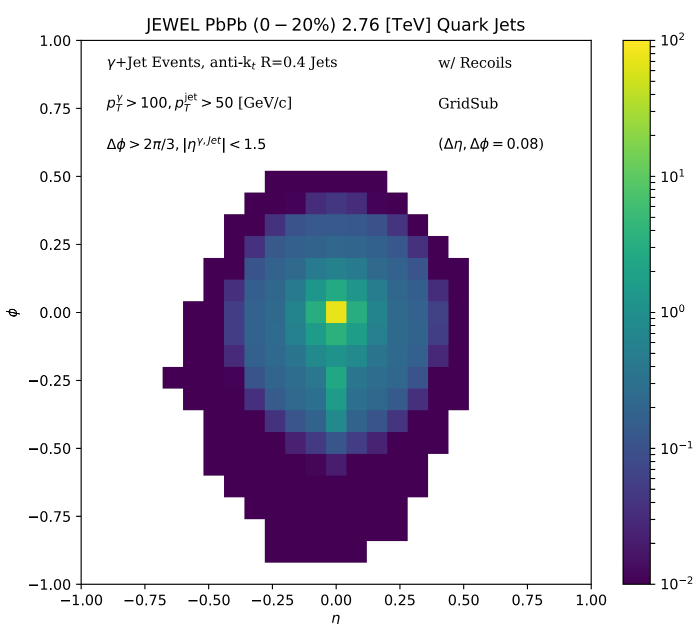
- Image representation should contain all info.
- Current State of the art - easy to implement

Cons

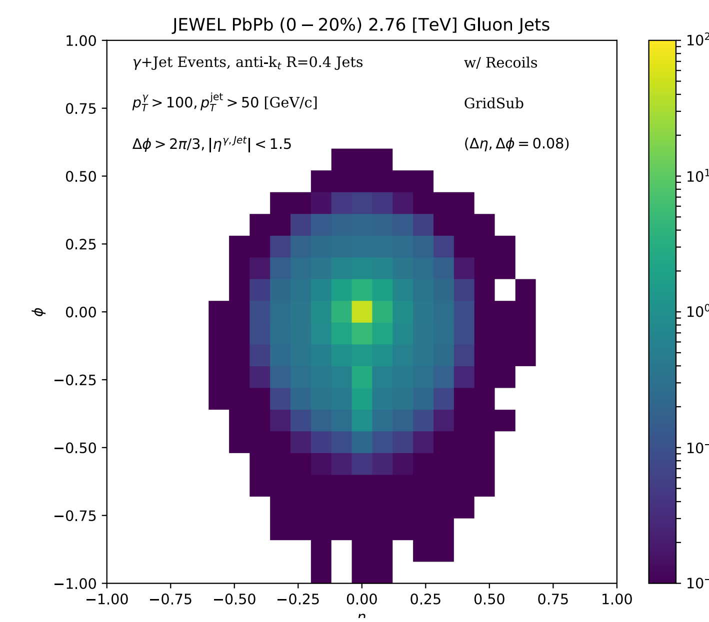
- classification in non-physics basis
- Best case scenario - no fluctuating background!

JEWEL

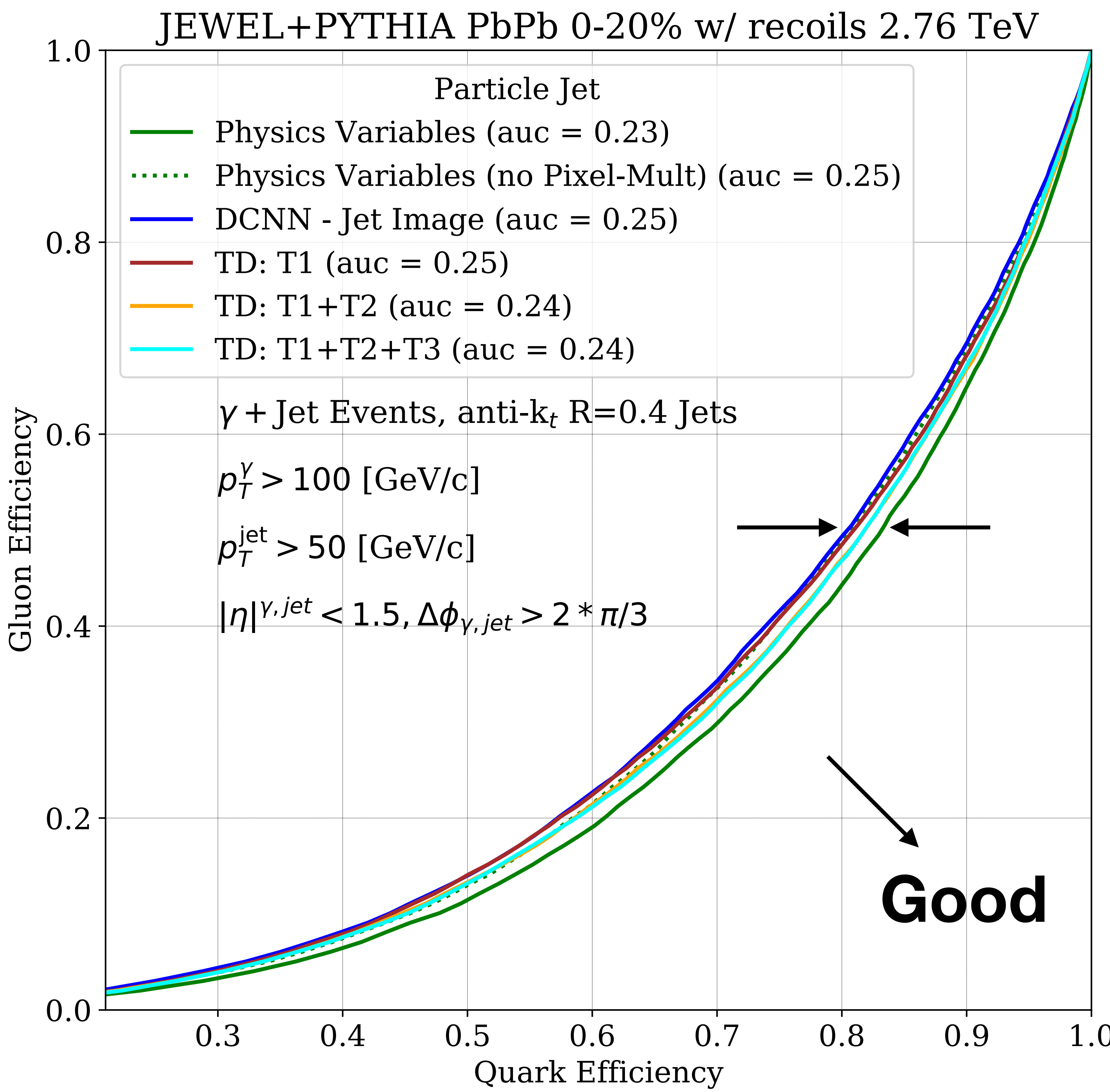
Quenched Q-Jet



Quenched G-Jet



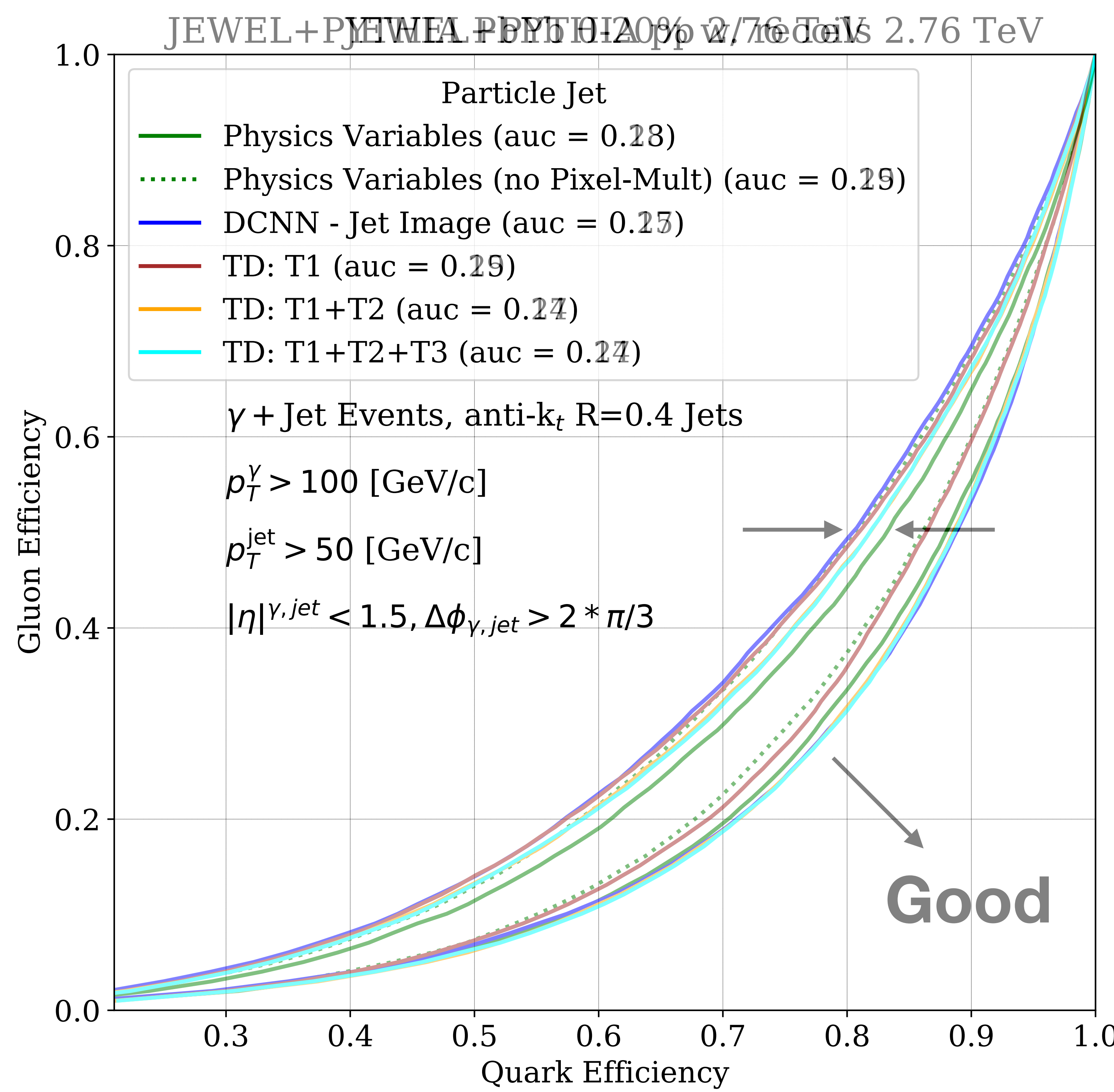
ROC curve for Quenched PbPb Particle Jets



All methods are relatively close to each other

Performance reduces!

ROC curve for Quenched PbPb Particle Jets



All methods are relatively close to each other

Performance reduces!

Recap - Classifiers

- Classifiers in our field are mostly supervised - with a potential built-in bias (utilize it!)
- There are many different ways to represent jets - information content is available to be exploited
- Quenched quark jets look like gluon jets!

Lets regress the truth!

The basics

What is machine learning?

Why are these tools useful in high energy colliders?

How to quantify performance?

Physics with ML

Classifier - Can select Heavy-Flavor or Quark vs Gluons

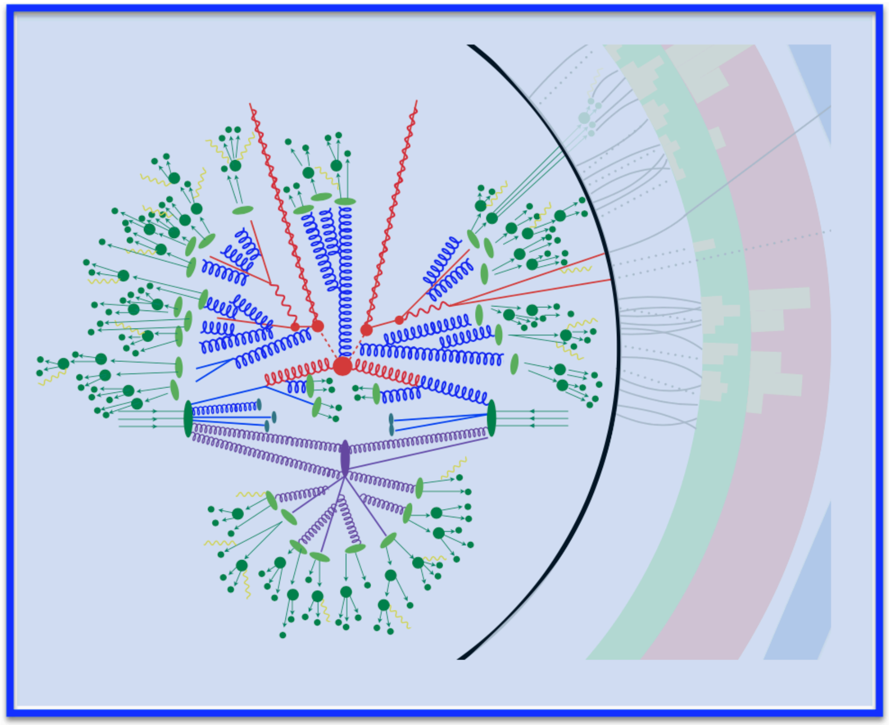
Regressor - multi-dimensional correction and unfolding

Generator - learn underlying physics of MC generators

Regression in HEP

- Correction procedures for energy scales and resolutions
- Multi-dimensional unfolding techniques

Want this ↓



Measure this ↓

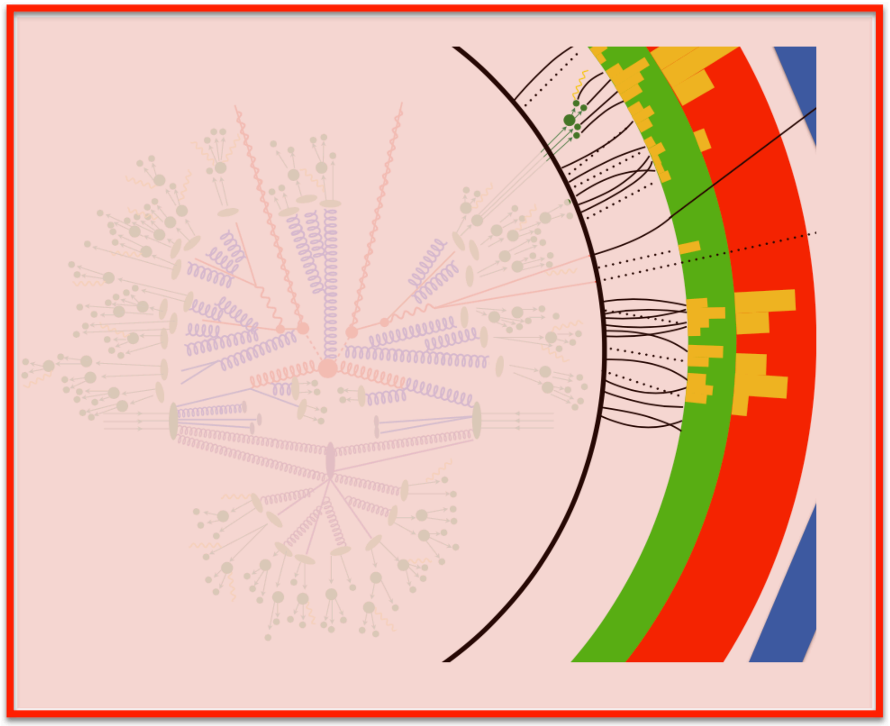
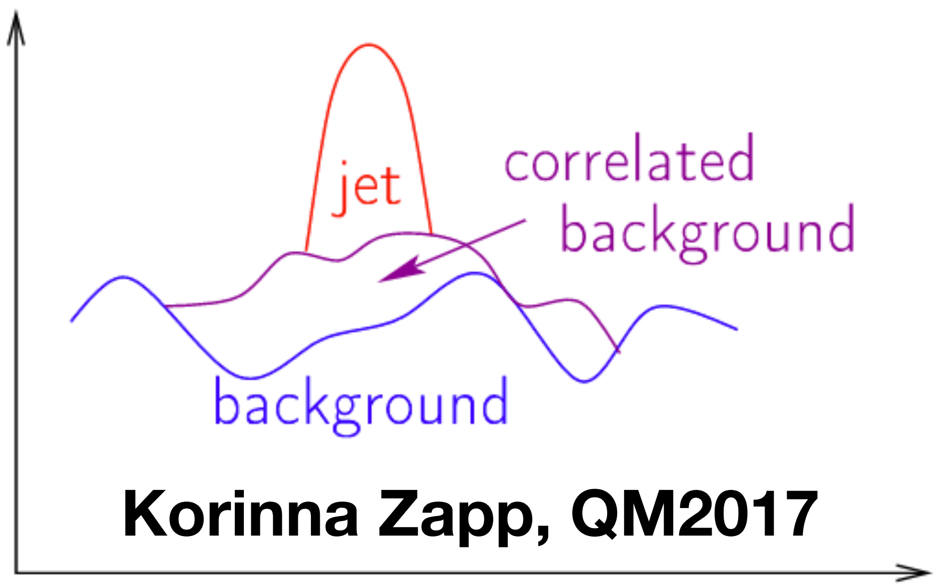
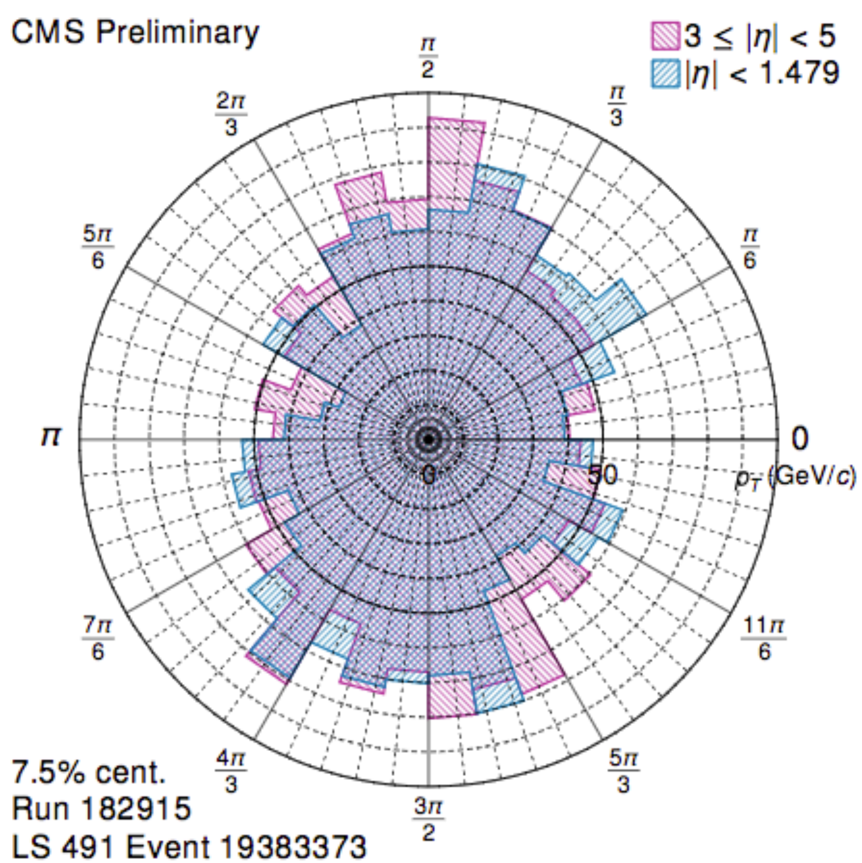


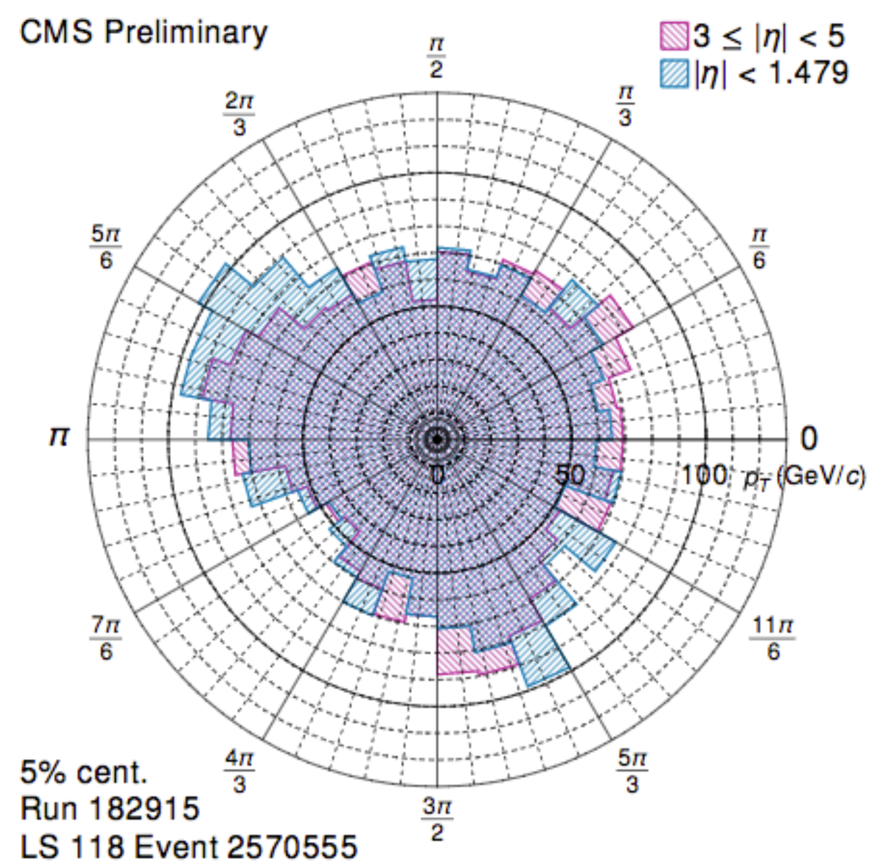
Image credit: Ben Nachman

Impact of the heavy ion background



Underlying event has flow, fluctuations and is correlated with the jet (like a wake)





ALICE method of ML based subtraction

Hannah Bossi (Yale) RHIC/AUM 2021

Training (PYTHIA fragmentation)

Train on "hybrid event" created by embedding PYTHIA jets into Pb-Pb Background

Shallow neural network

Testing

Apply ML estimator to hybrid events not used in training.

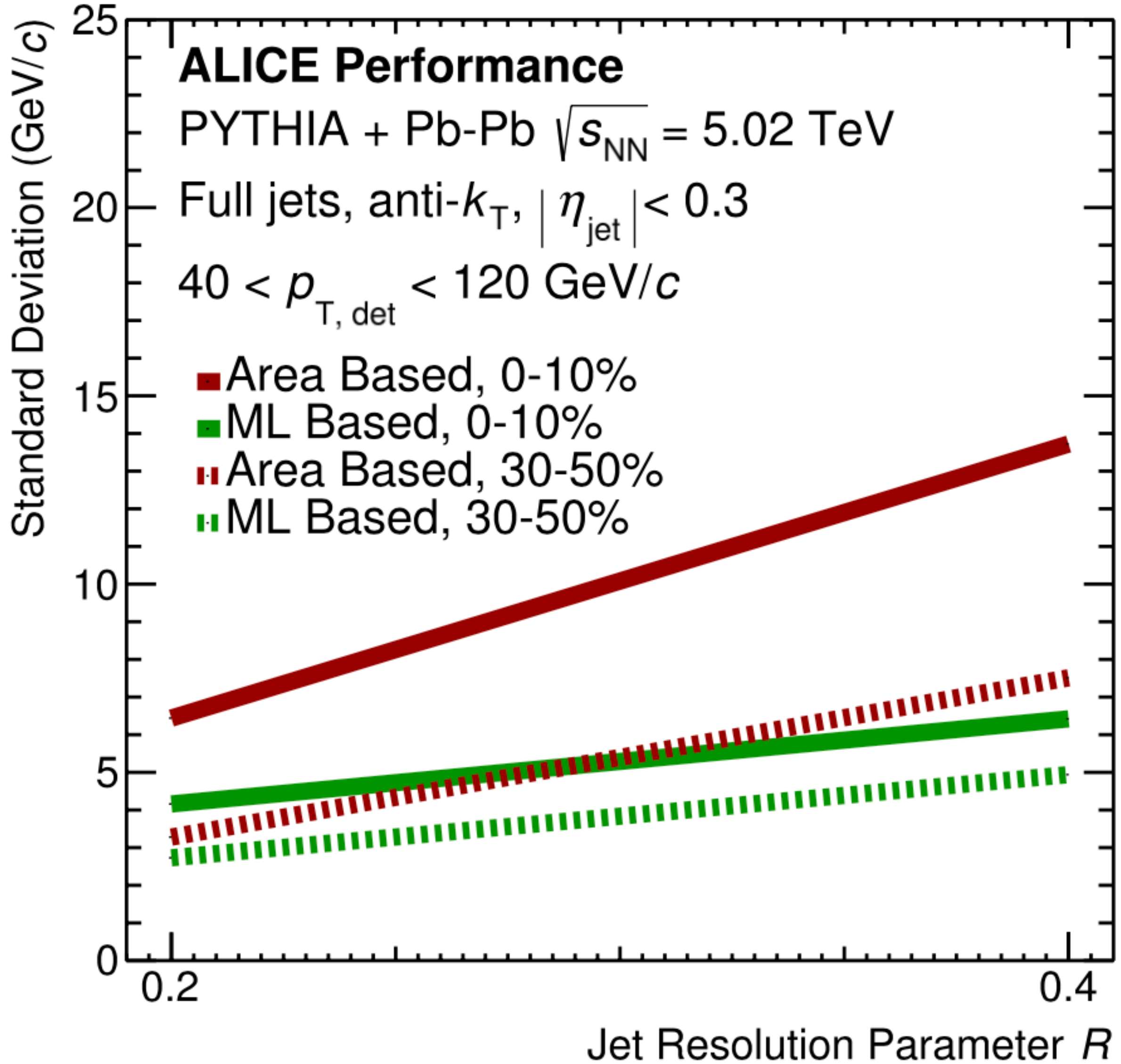
Key is that this background is *realistic.*

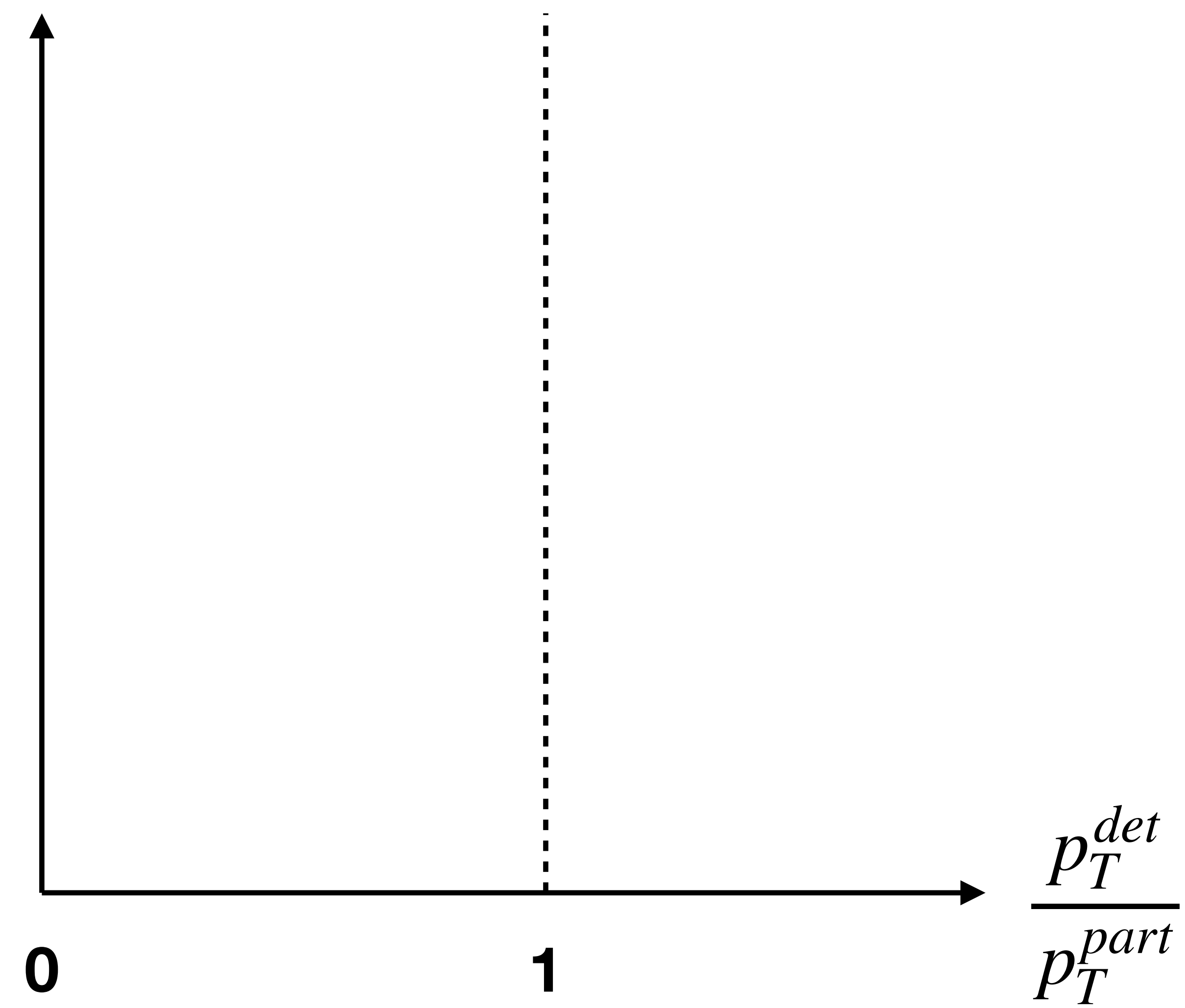
Simple question, relatively simple network can get a short clear answer!

Do we get back the signal we put in?

ML corrector in action

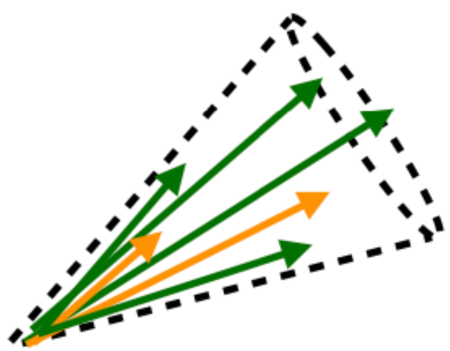
Hannah Bossi (Yale) RHIC/AUM 2021





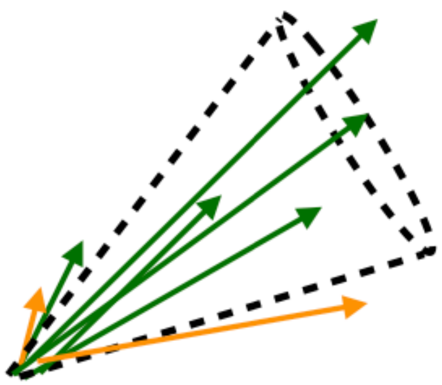
- Significantly less jet energy resolution with the ML based method along with first ever estimate of impact of truth shape 'bias' in correction

ML corrector in action



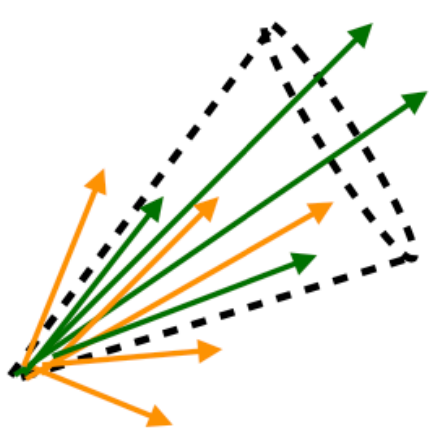
Fractional In-Cone

→ Learning on constituents introduces a bias towards PYTHIA fragmentation!

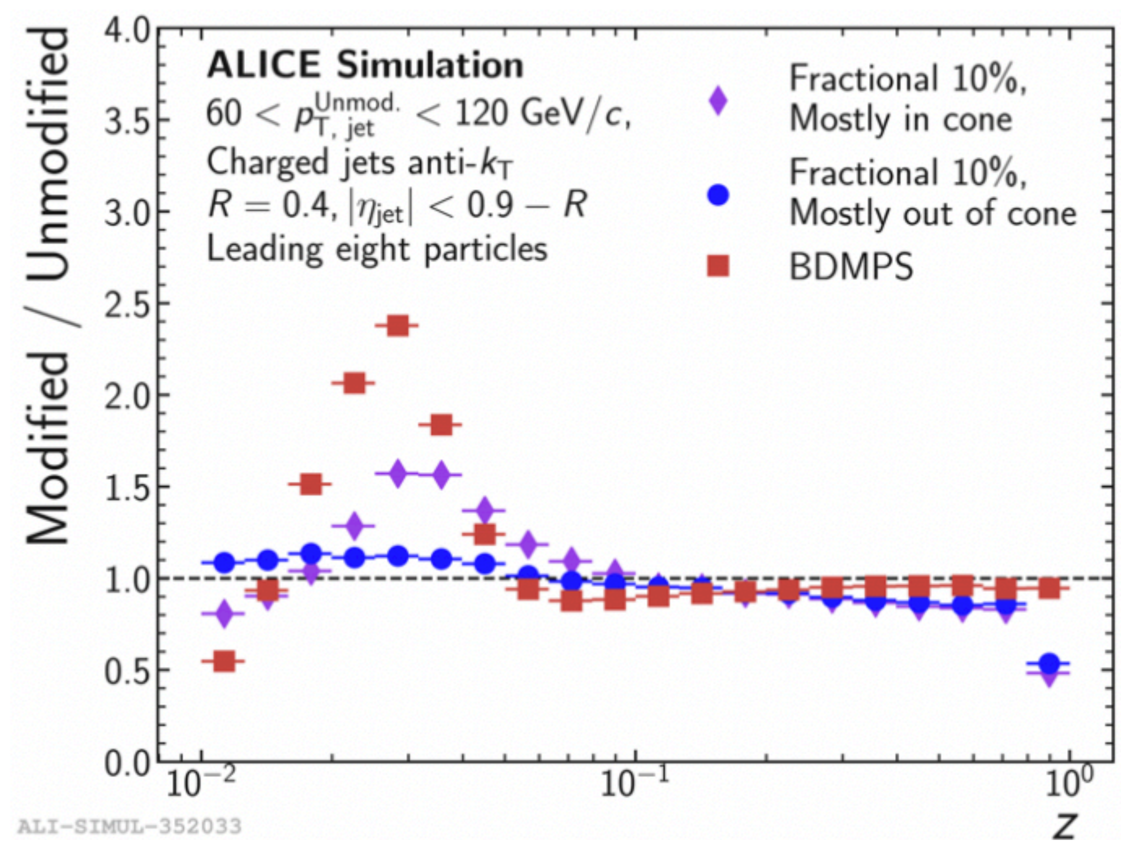


Fractional Out-of-Cone

→ Modify PYTHIA jets to change the fragmentation.

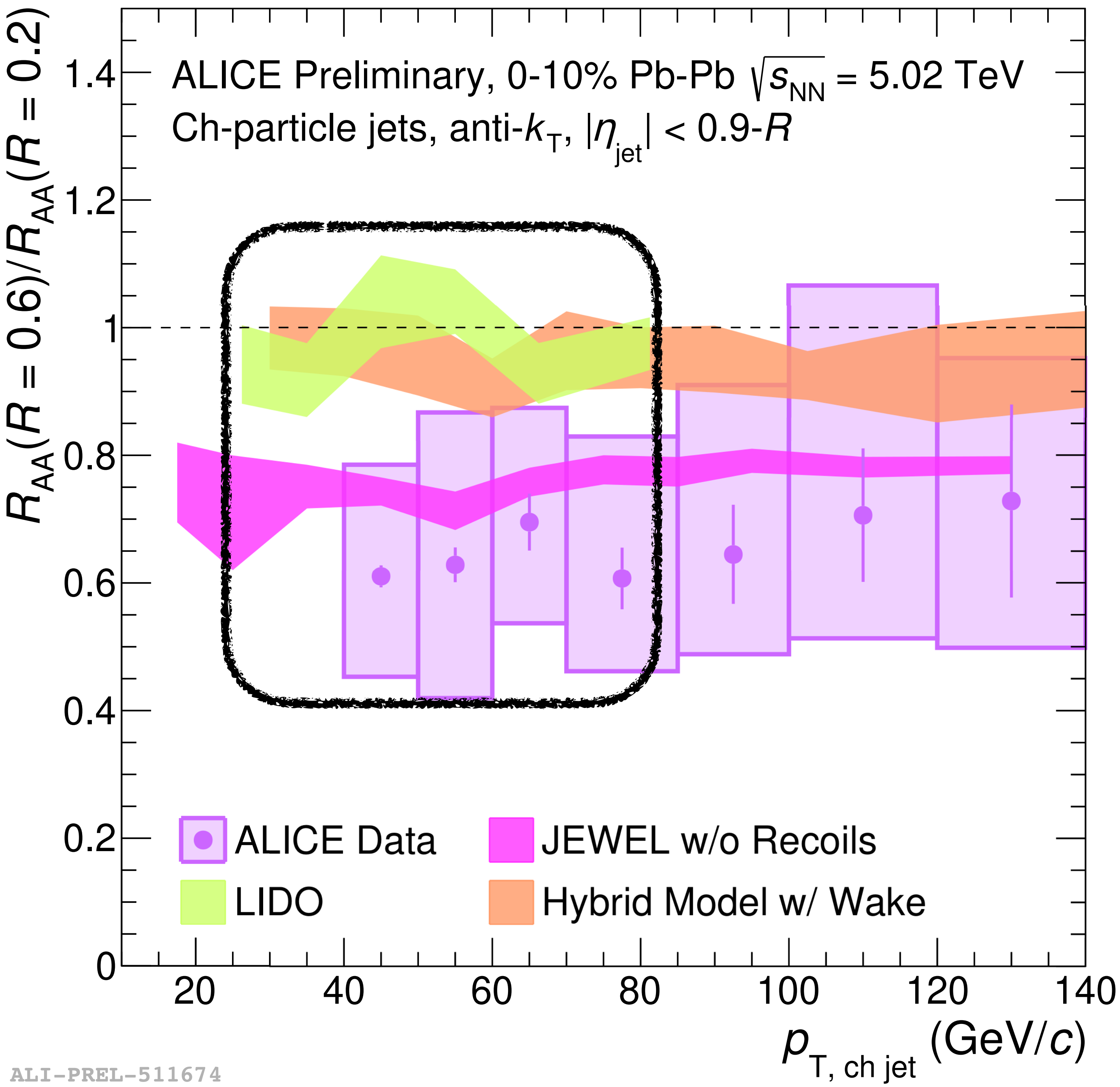


BDMPS- Motivated



Hannah Bossi (Yale) RHIC/AUM 2021

What did it enable us to do?

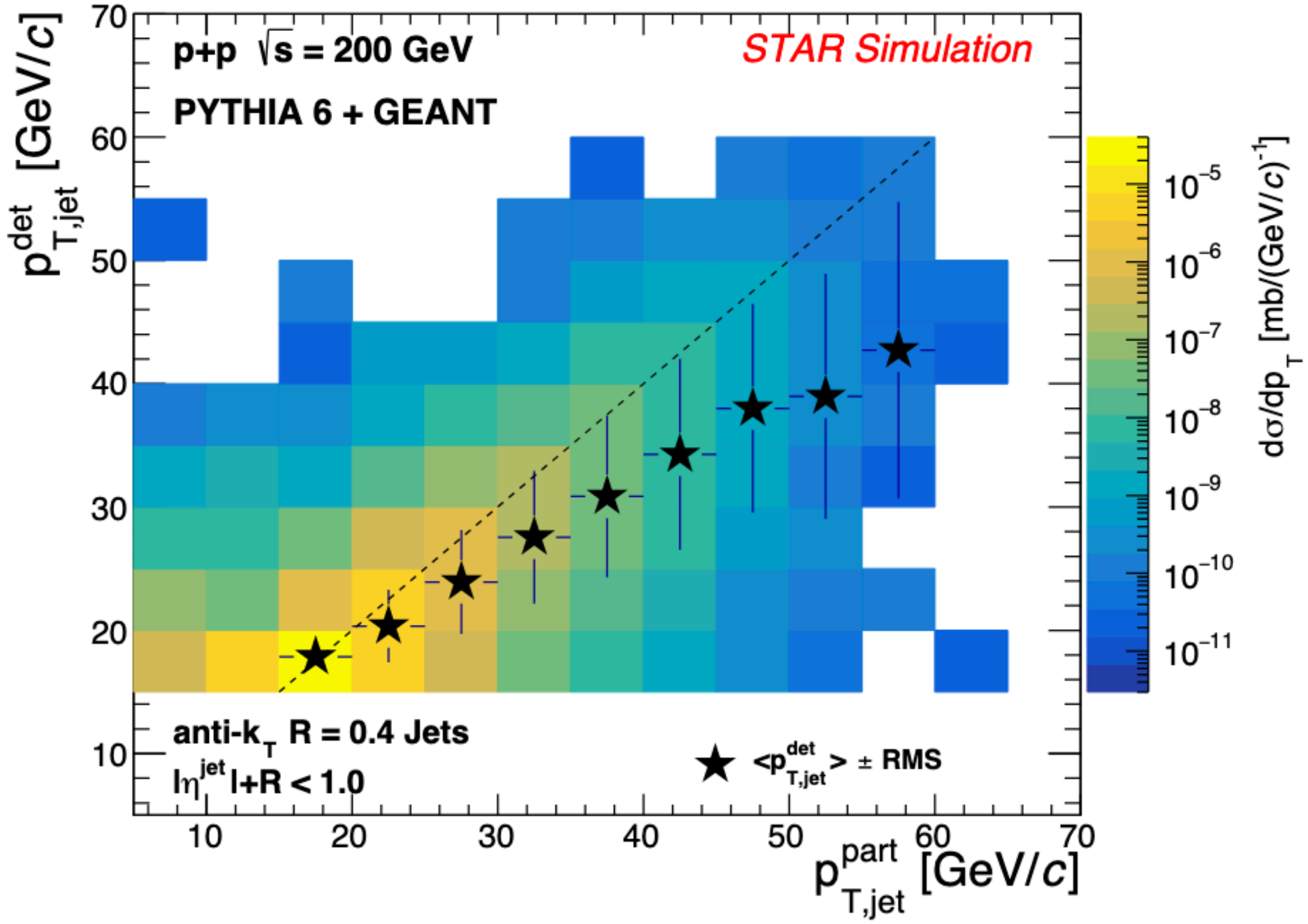


- Extend our measurements to lower momentum range where the impact of the background is large
- Reduced uncertainties key to making a potentially tantalizing statement about radial dependence of energy loss

Unfolding - a quick primer

Corrections for Detector Resolution

Response Matrix



For a given generator jet p_T - the probability get reconstructed at a certain p_T

Two separate methods

- Bayesian
- Single Value Decomposition

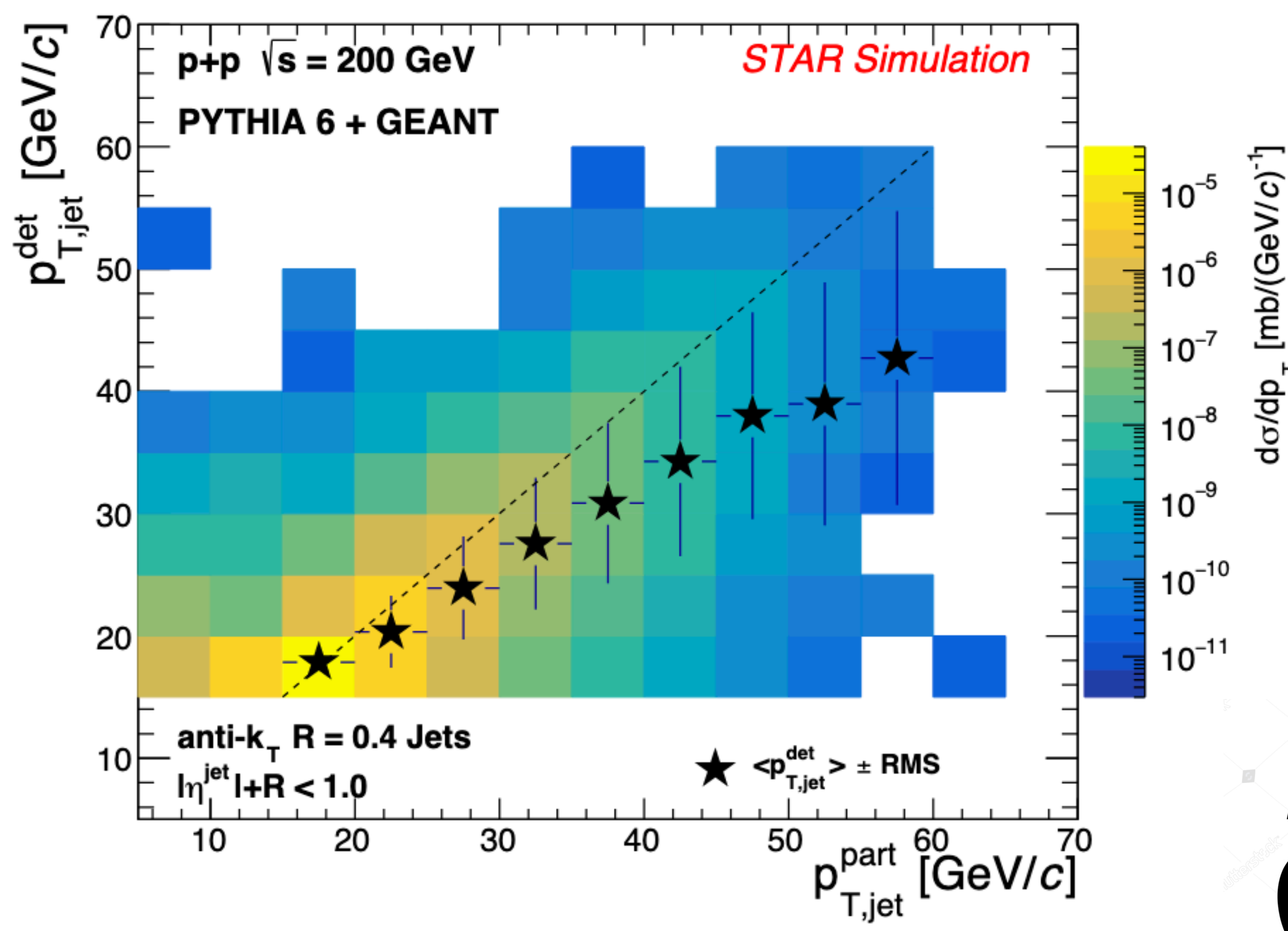
Based on RooUnfold Package

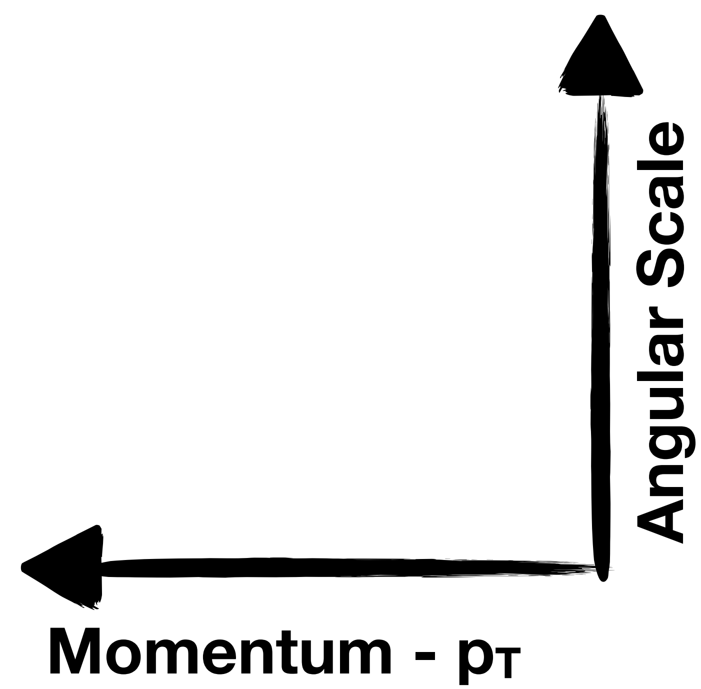
After unfolding - can directly compare with theory calculations

Unfolding - a quick primer

Corrections for Detector Resolution

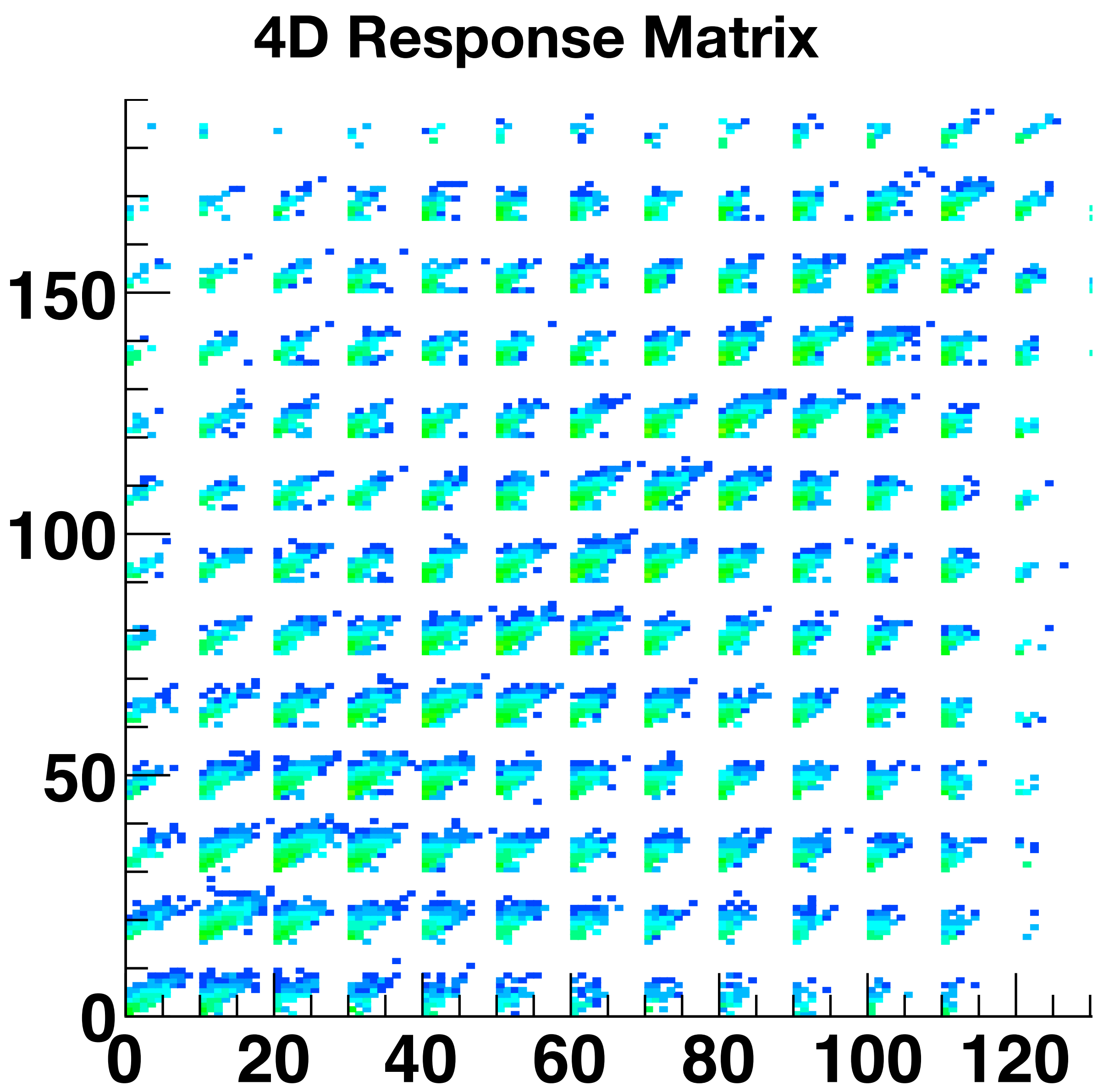
Response Matrix





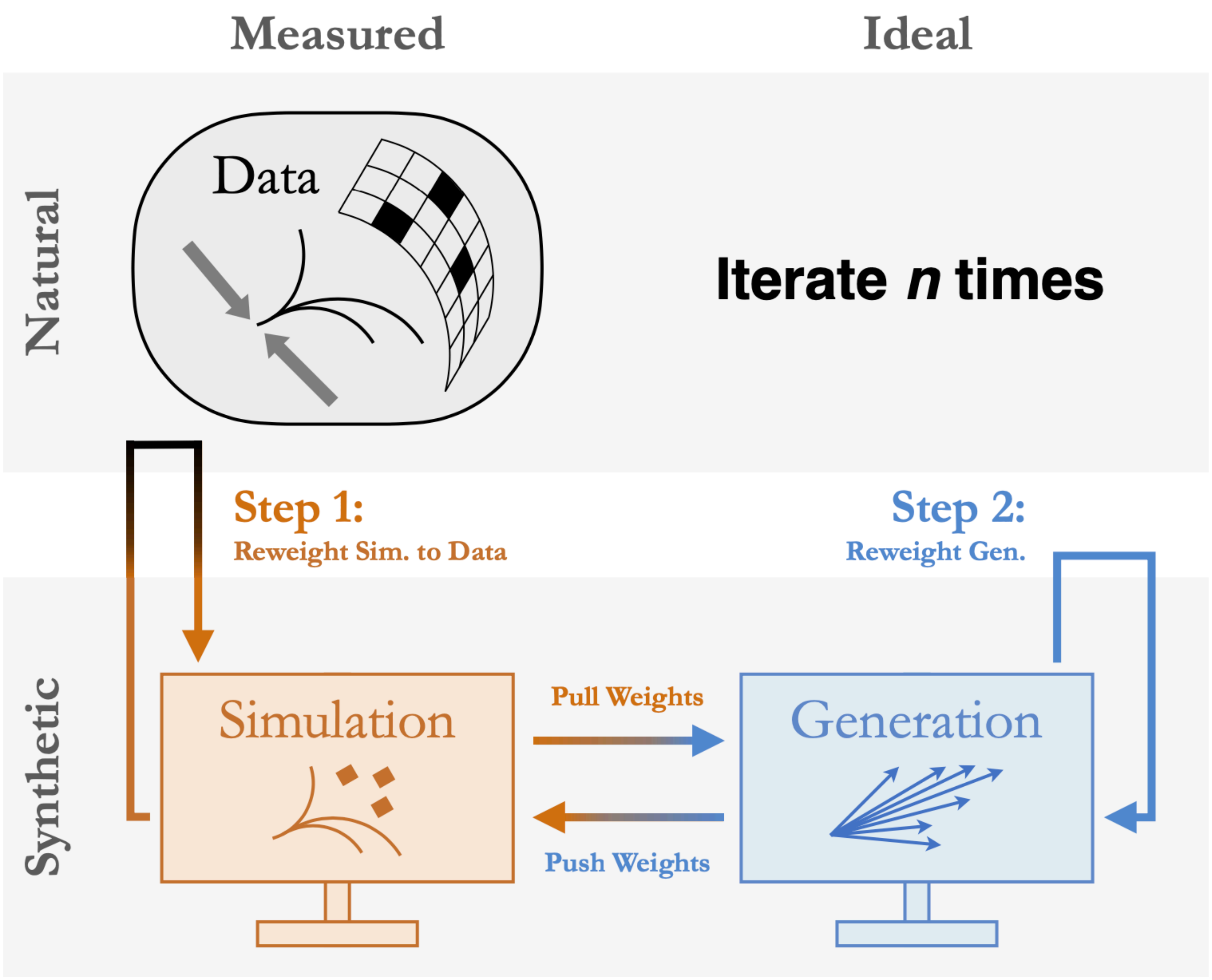
What are the dimensions of the response matrix for correlated observables?

Unfolding multi-dimensions with standard method



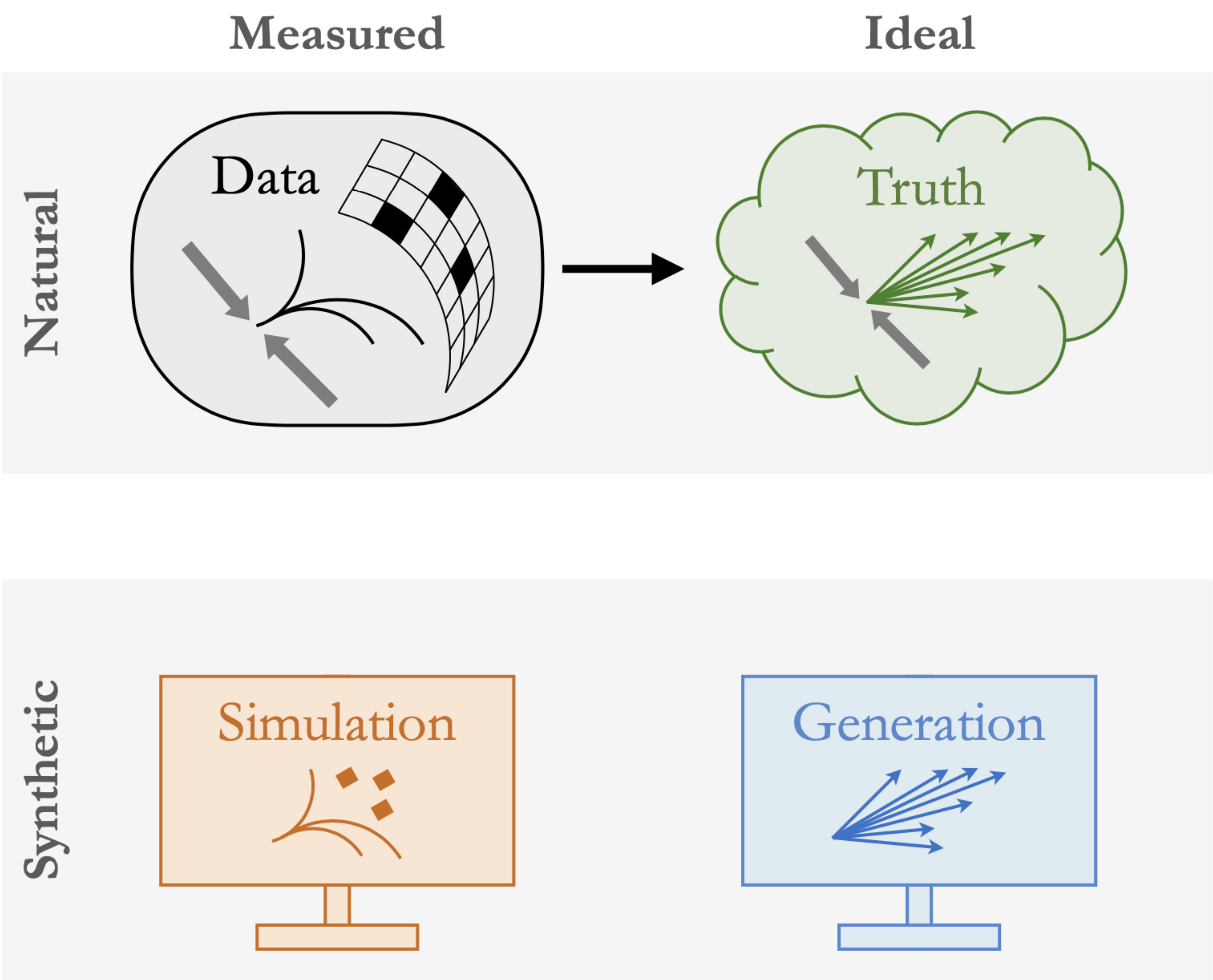
- Lets unfold correlation between jet mass and pT
- Unfolding 2D observable jet pT requires 4D response matrix
- Increases dependence on statistics and prior shape variations

MultiFold (Omnifold)



Ben Nachman (LBL)

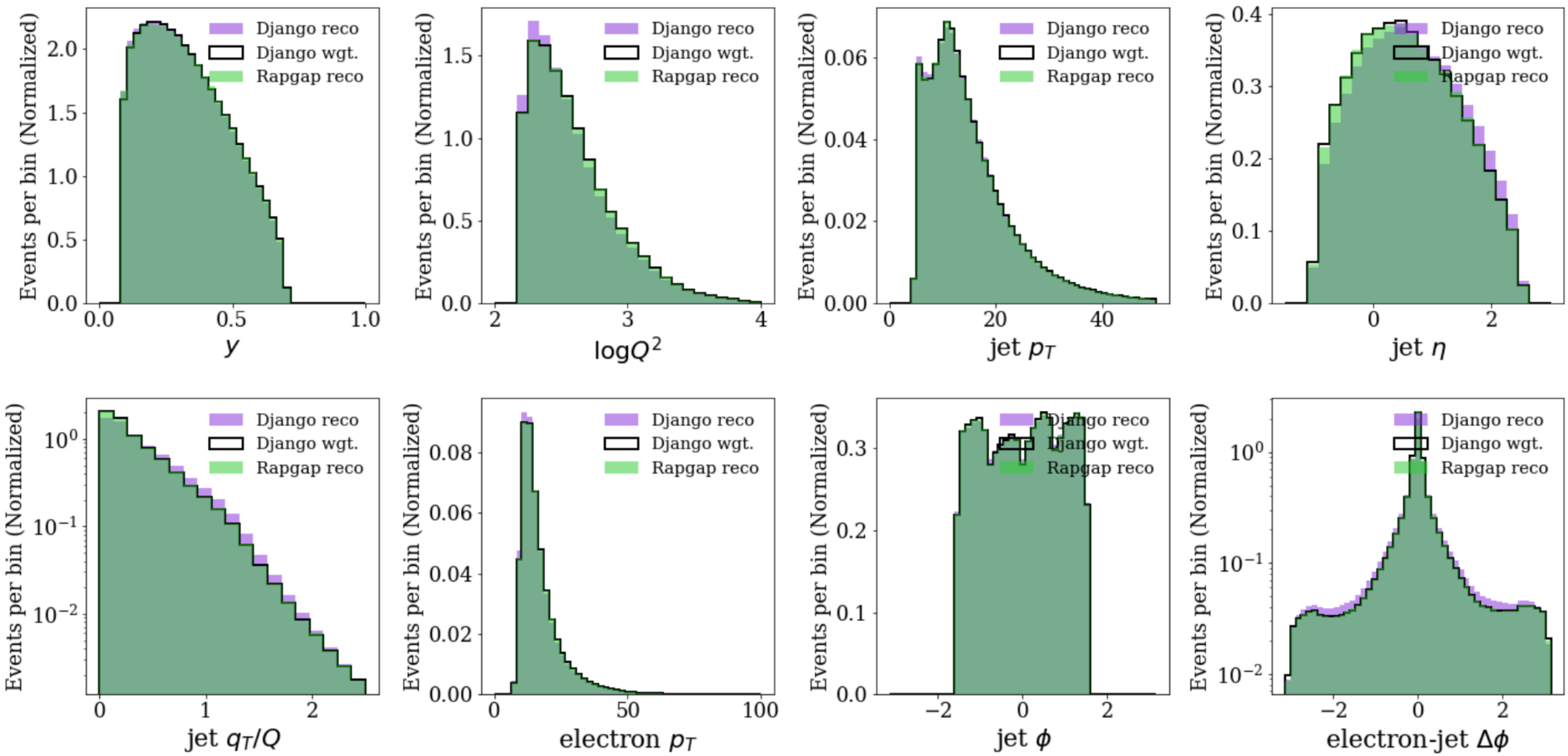
MultiFold (Omnifold)



Ben Nachman (LBL)

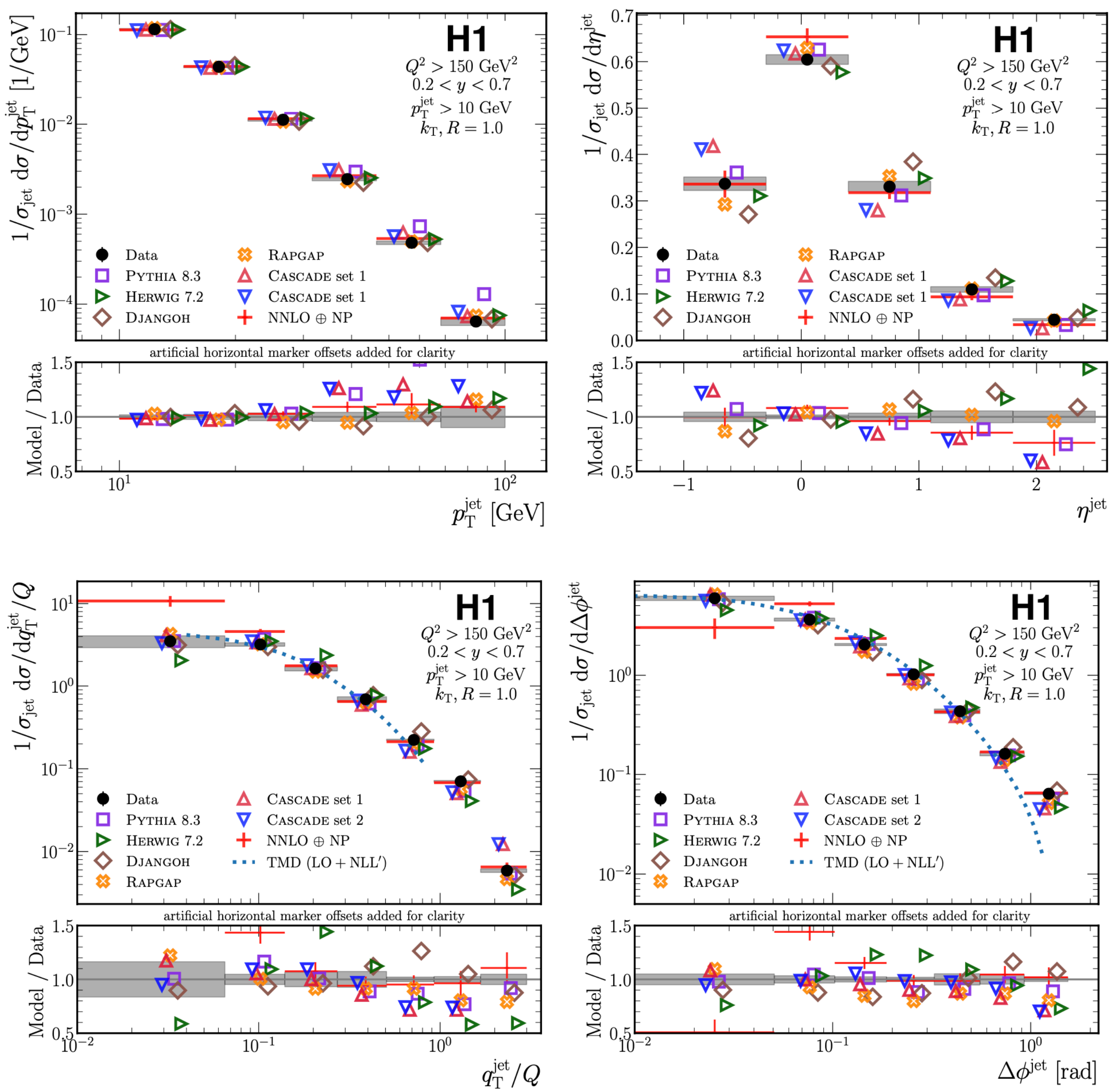
Unfolding closure tests using two different MC samples

All of these distributions are simultaneously reweighted!



Ben Nachman (LBL)

What you get at the end?



Ben Nachman (LBL)

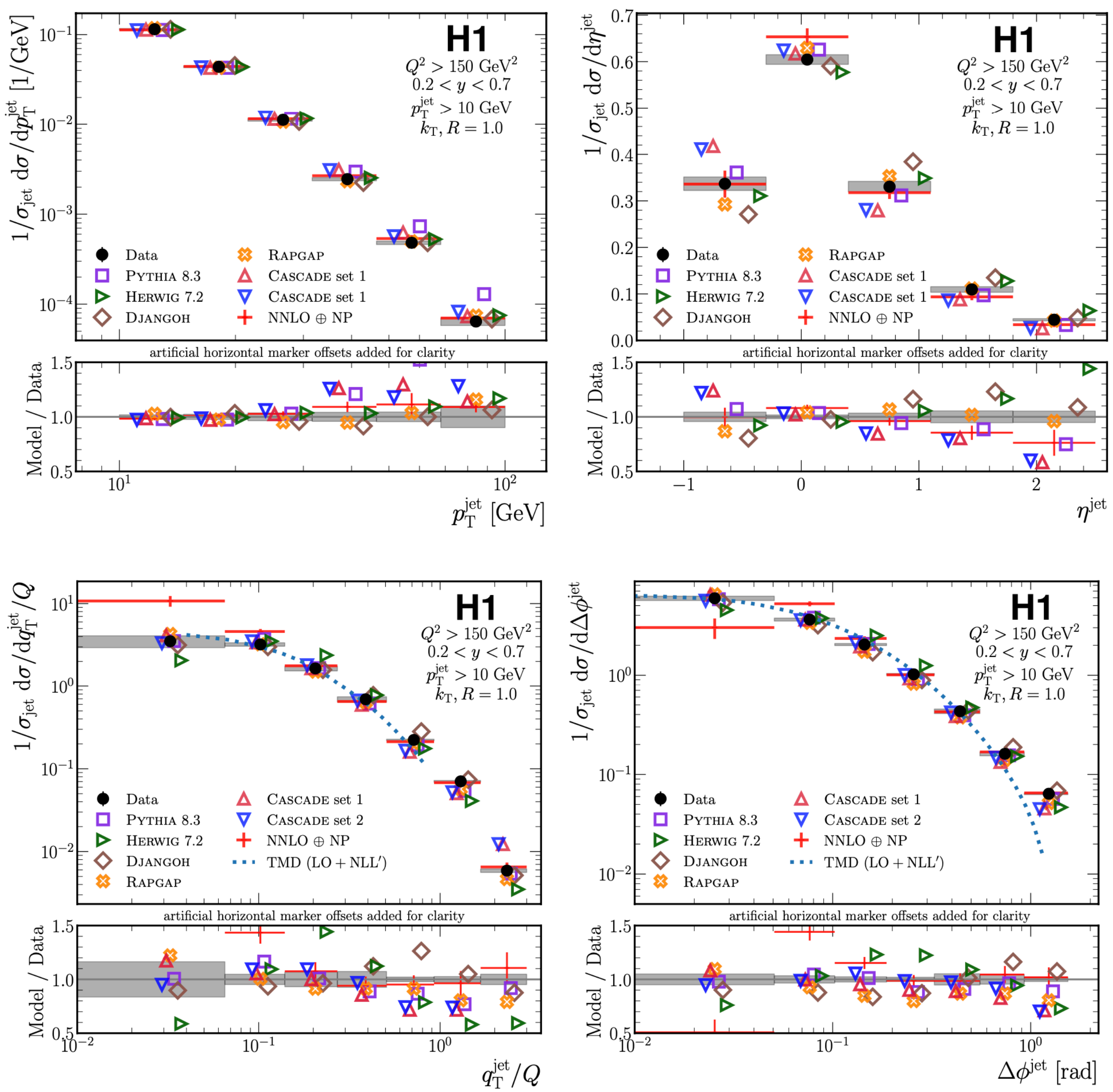
- Multi-dimensional measurements 'un-binned'

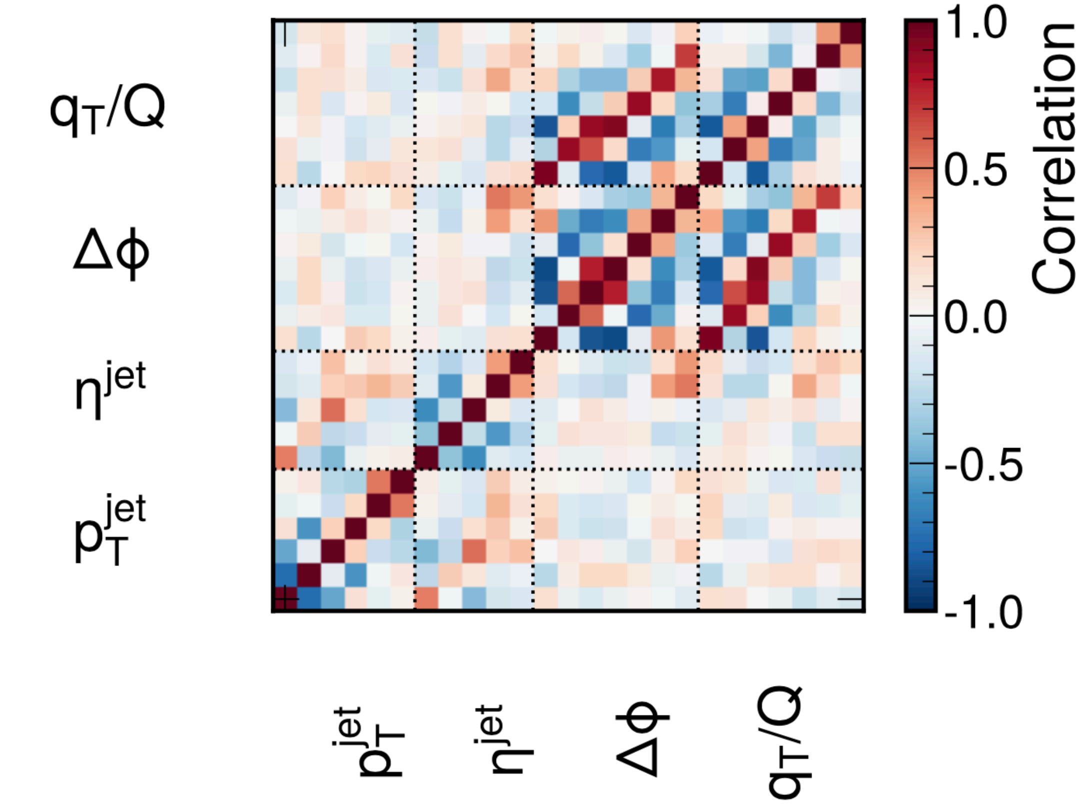
If you unfold

$p_T, \eta, \Delta\phi, q_i/Q, \ldots$

what do you get for free?

What you get at the end?





Ben Nachman (LBL)

- Multi-dimensional measurements 'un-binned'
- You get the correlations for 'free'

The basics

What is machine learning?

Why are these tools useful in high energy colliders?

How to quantify performance?

Physics with ML

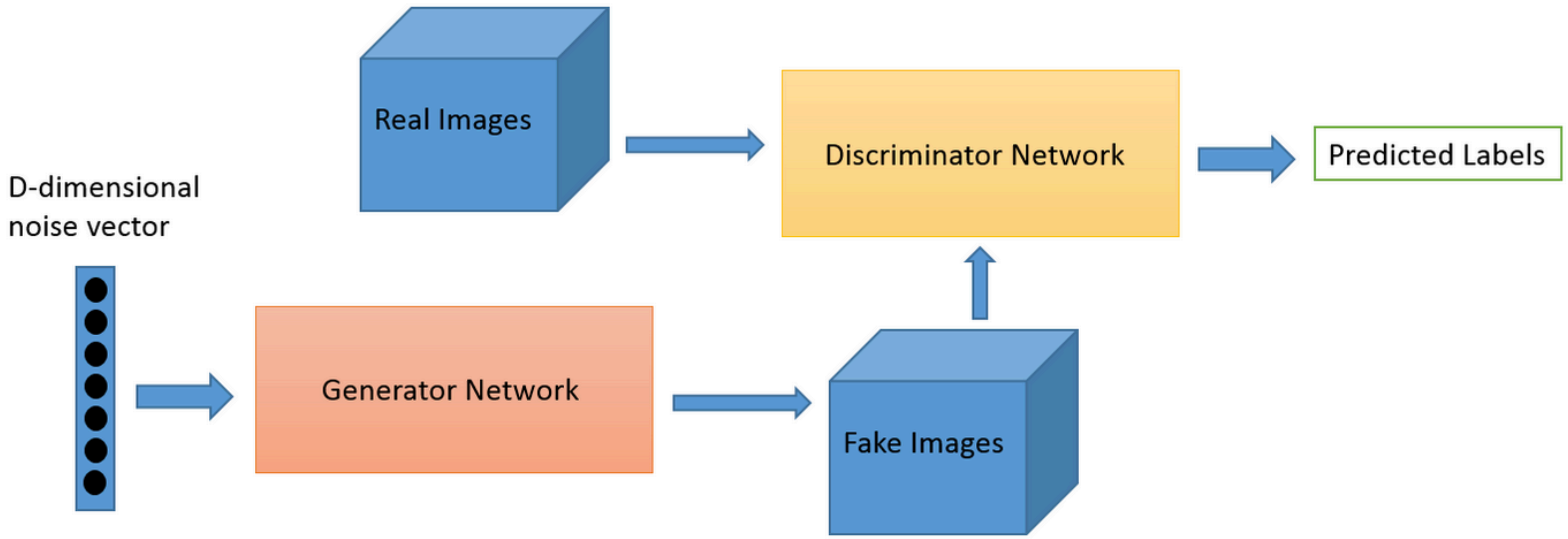
Classifier - Can select Heavy-Flavor or Quark vs Gluons

Regressor - multi-dimensional correction and unfolding

Generator - learn underlying physics of MC generators

Lets ask the AI to learn physics (or something..?)

- Given a particle-by-particle, event-by-event distribution of quantities - can a model early the intricacies of the generation?
- Enter Generative-Adversarial-Networks (GAN) - playing one network vs another



Credit: O'Reilly

https://skymind.ai/wiki/generative-adversarial-network-gan

A few things GANs can do!

Generate Faces!

Nearest training set

Ian Goodfellow et. al, 1406.2661

A few things GANs can do!

latent space arithmetic : Reduce images to its inherent hidden representation (same-dimensions) so we can perform mathematical operations!

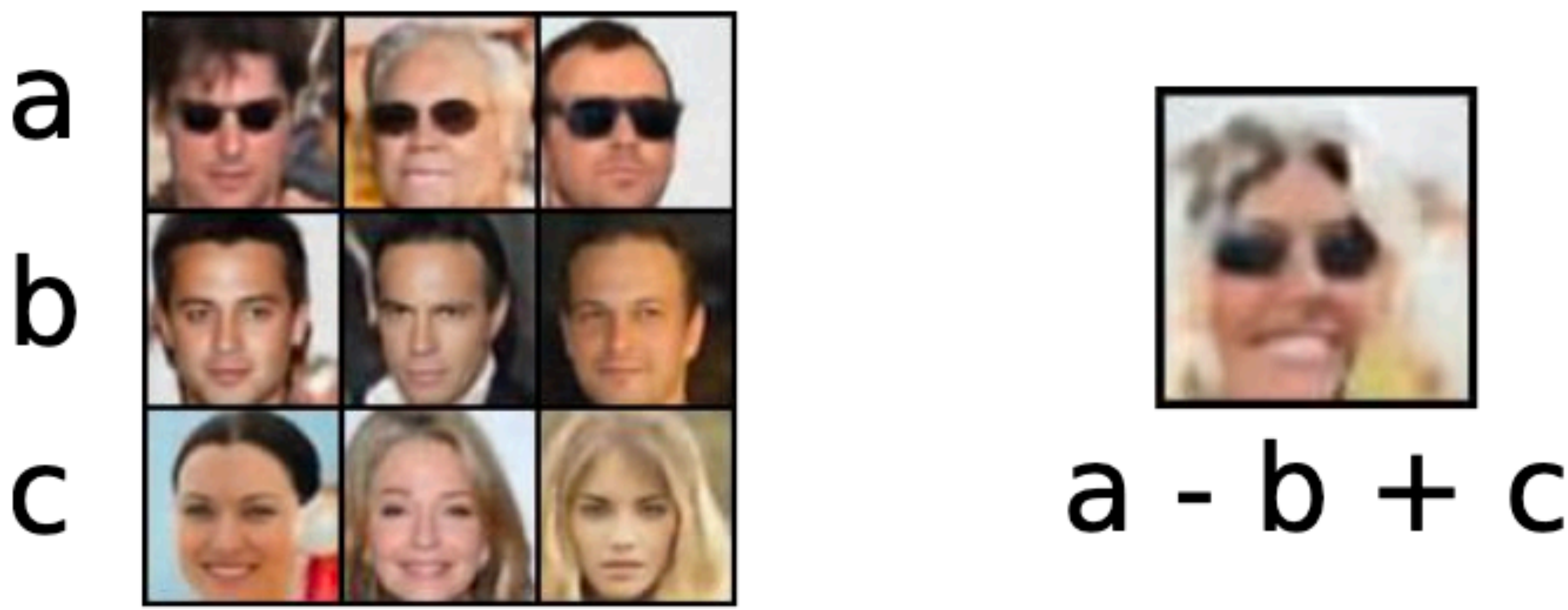
a - b + c

Piotr Bojanowski et. al, 1707.05776

Facebook AI

A few things GANs can do!

latent space arithmetic : Reduce images to its inherent hidden representation (same-dimensions) so we can perform mathematical operations!



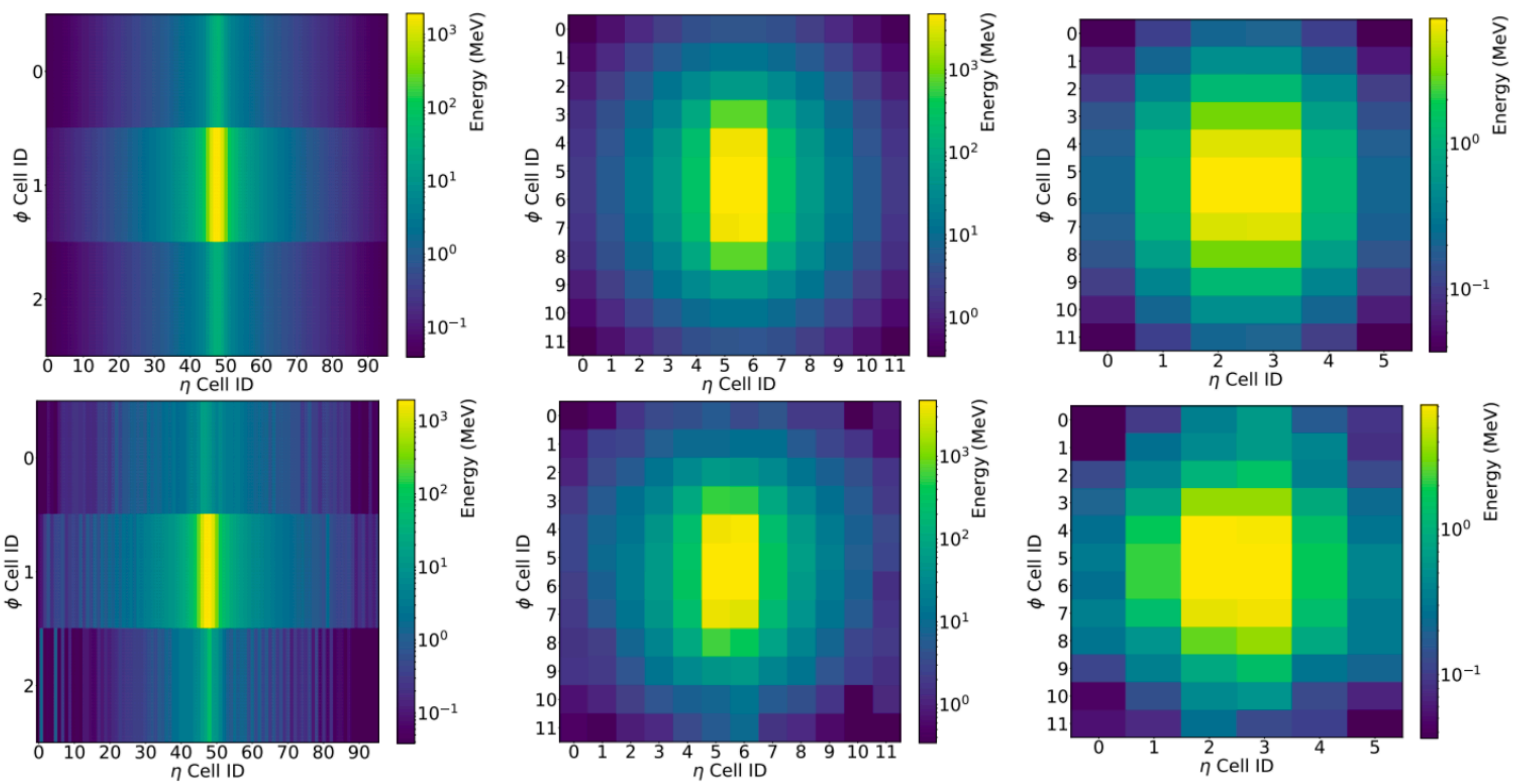
Piotr Bojanowski et. al, 1707.05776

Facebook AI

Simulating ATLAS segmented calorimeter

GEANT

CaloGAN



Pagnini M, Nachman B, Olivera L
Phys. Rev. D 97, 014021 (2018)

Michela Pagnini (Yale, LBNL), ML4Jets17

Individual positron showers and generated nearest neighbors

GEANT 1st layer deposition

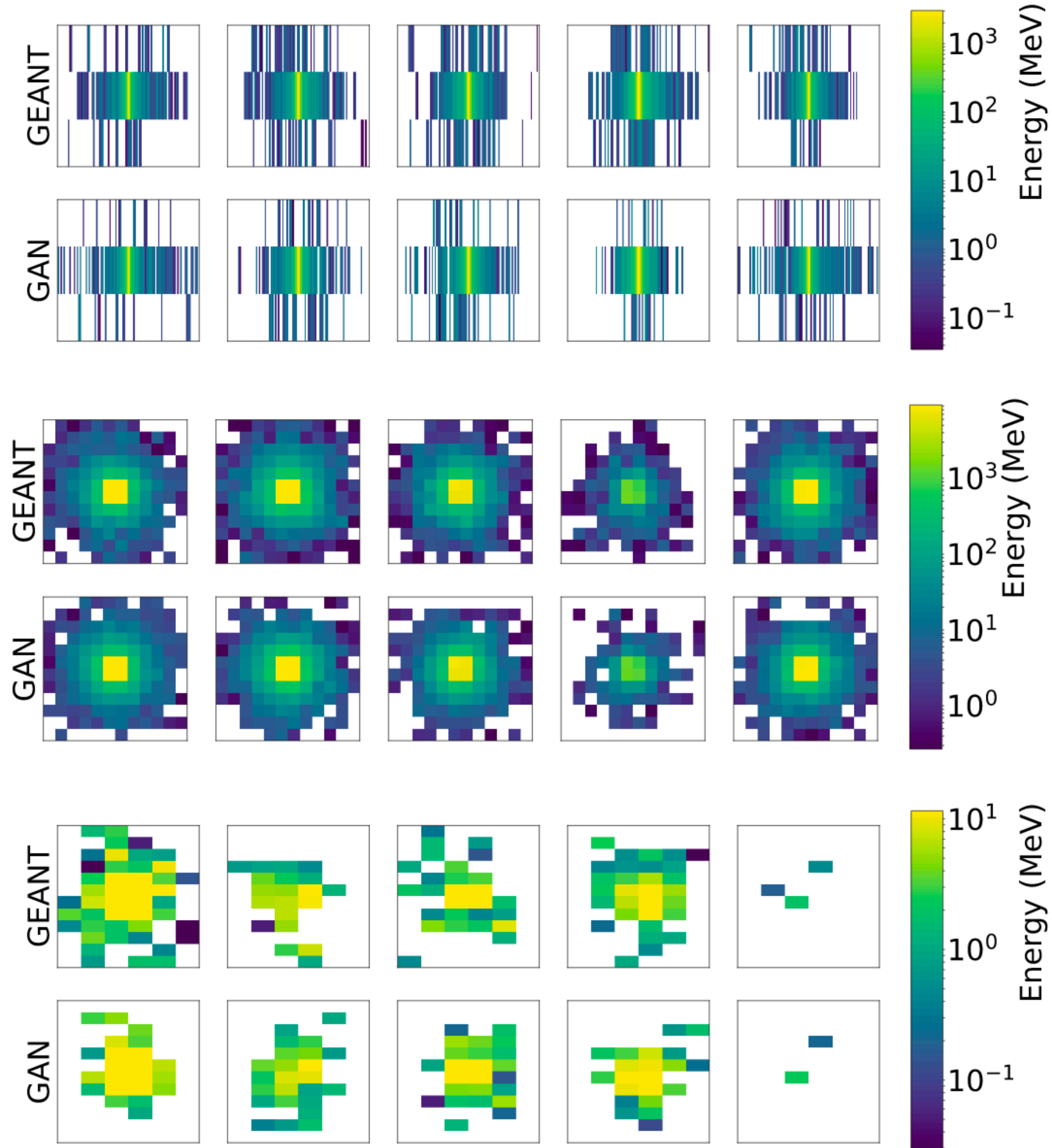
CaloGAN 1st layer deposition

GEANT 2nd layer deposition

CaloGAN 2nd layer deposition

GEANT 3rd layer deposition

CaloGAN 3rd layer deposition



Michela Pagnini (Facebook AI), ML4Jets17

Individual positron showers and generated nearest neighbors

Generation Method	Hardware	Batch Size	milliseconds/shower
GEANT4	CPU	N/A	1772
CALOGAN	CPU	1	13.1
		10	5.11
		128	2.19
		1024	2.03
	GPU	1	14.5
		4	3.68
		128	0.021
		512	0.014
		1024	0.012

Up to a 100,000x speed-up!

CaloGAN 3rd layer deposition

Michela Pagnini (Facebook AI), ML4Jets17

The basics

What is machine learning?

Why are these tools useful in high energy colliders?

How to quantify performance?

Physics with ML

Classifier - Can select Heavy-Flavor or Quark vs Gluons

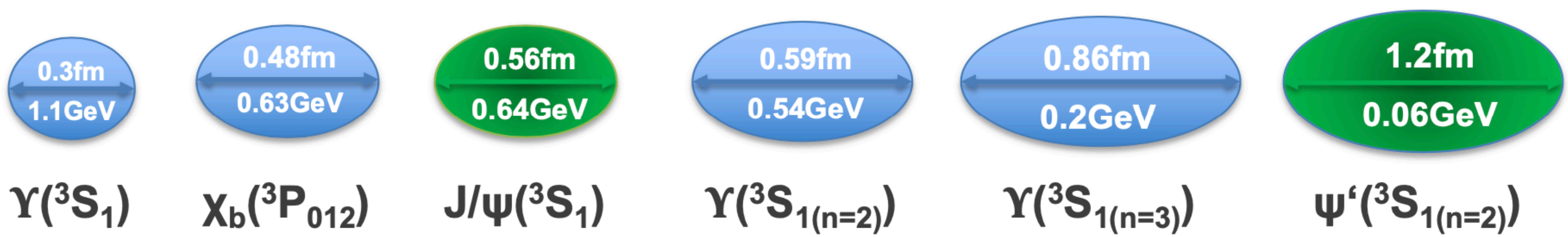
Regressor - multi-dimensional correction and unfolding

Generator - learn underlying physics of MC generators

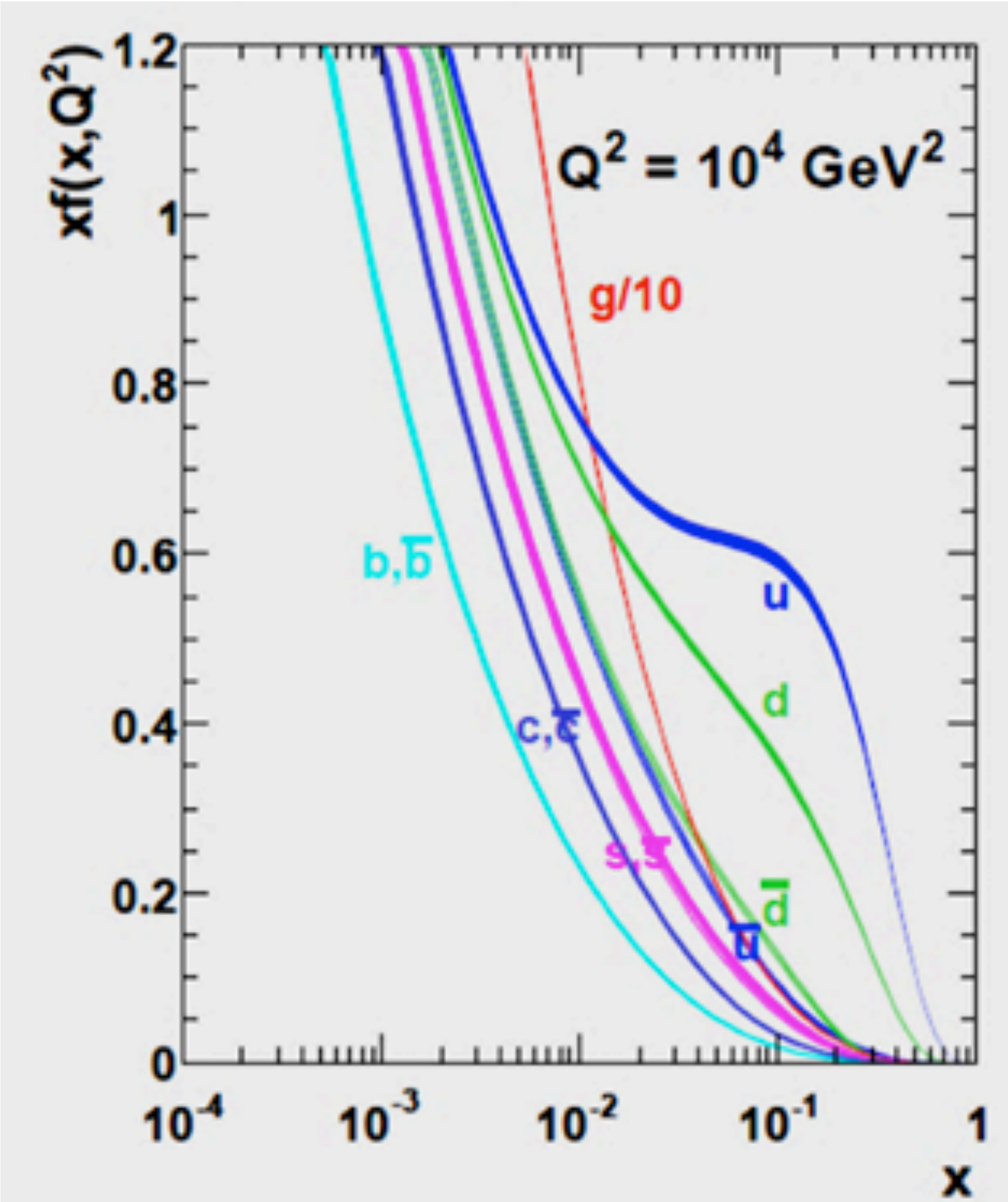
Heavy Flavor

- Experimentally - can typically mean one of two things - charm, bottom or top quarks, or the hadrons resulting from those quarks

Bound states of heavy flavor quark-antiquark

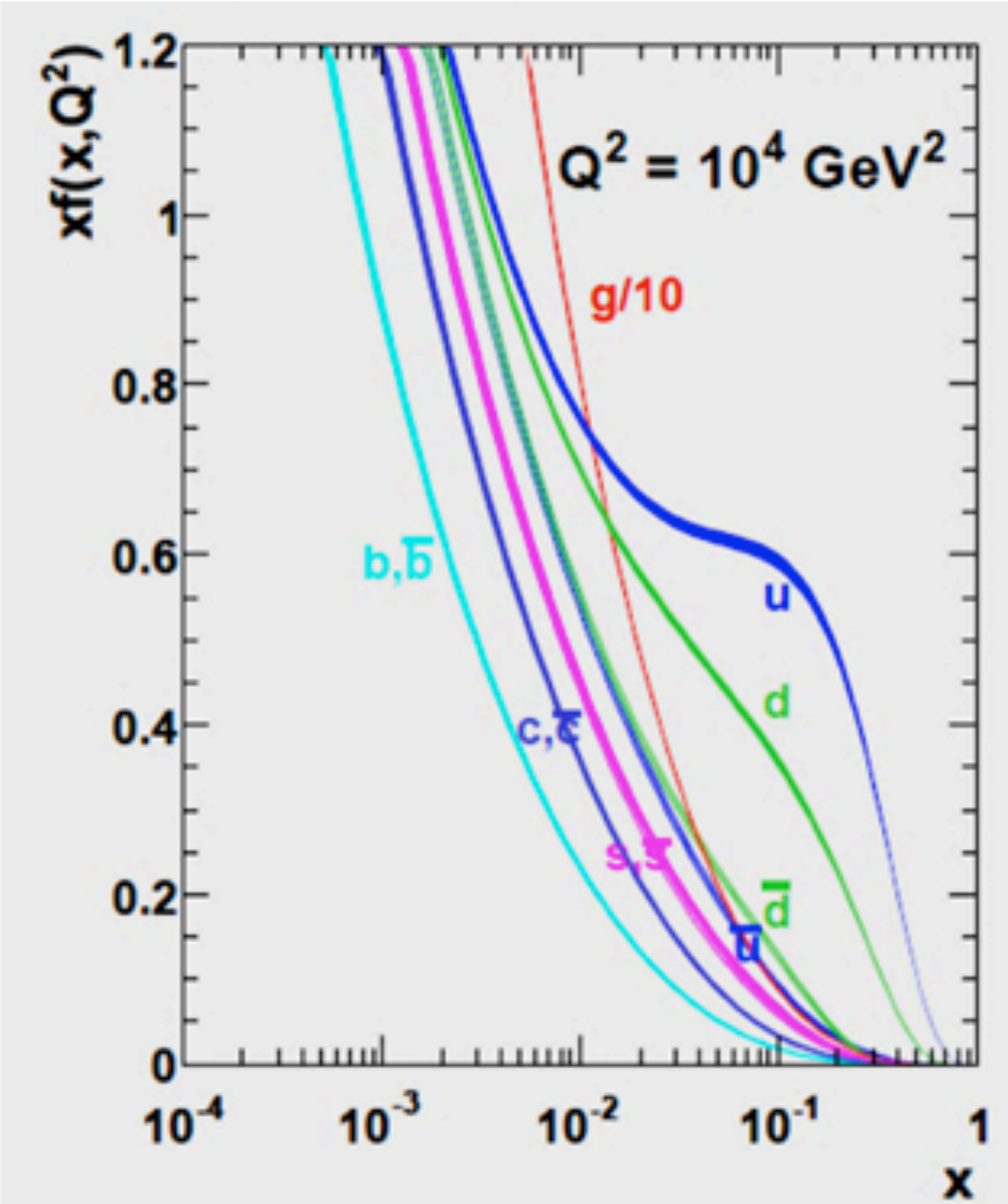


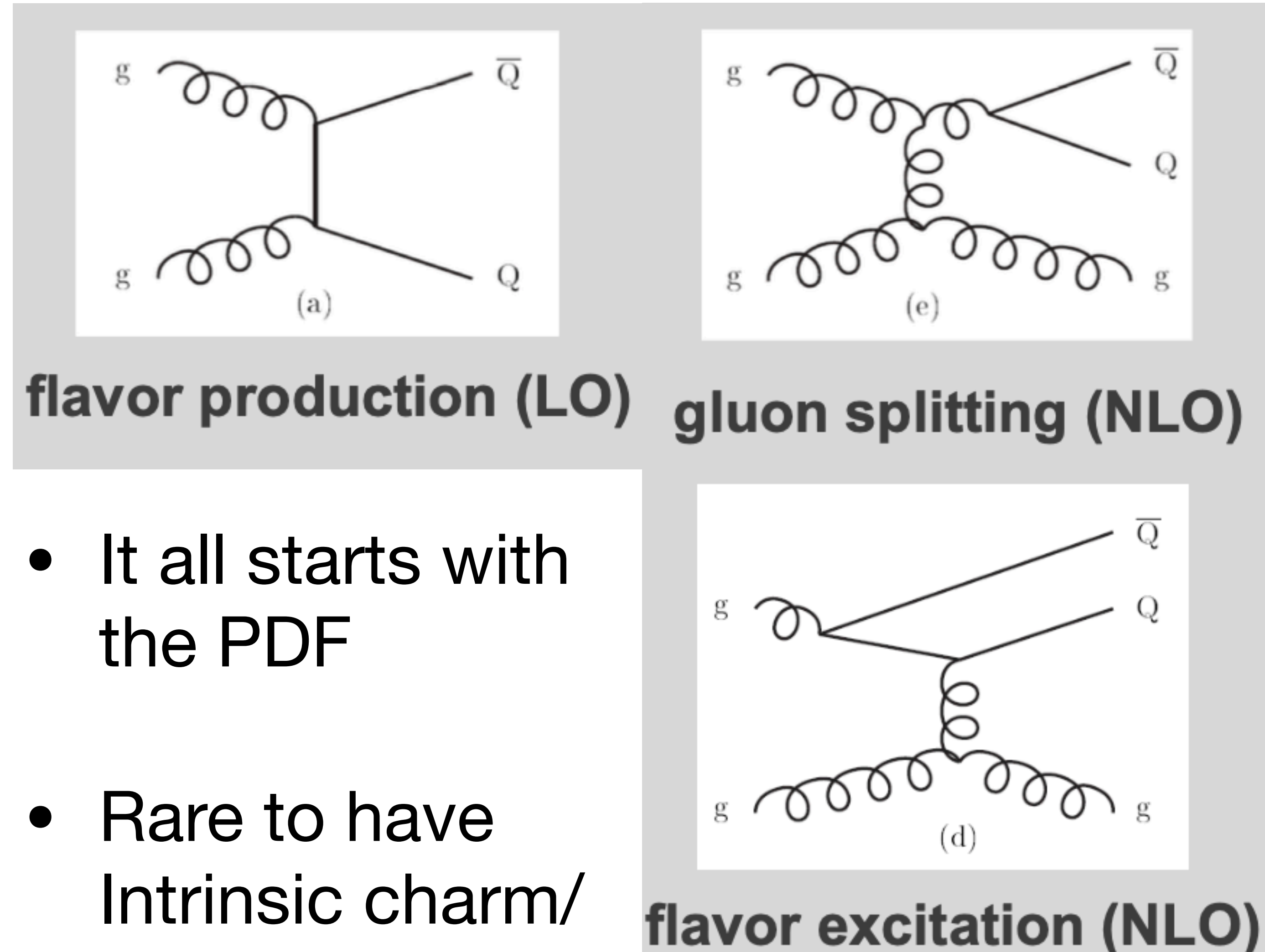
Producing heavy flavor



- It all starts with the PDF
- Rare to have Intrinsic charm/ bottom in a proton

Producing heavy flavor

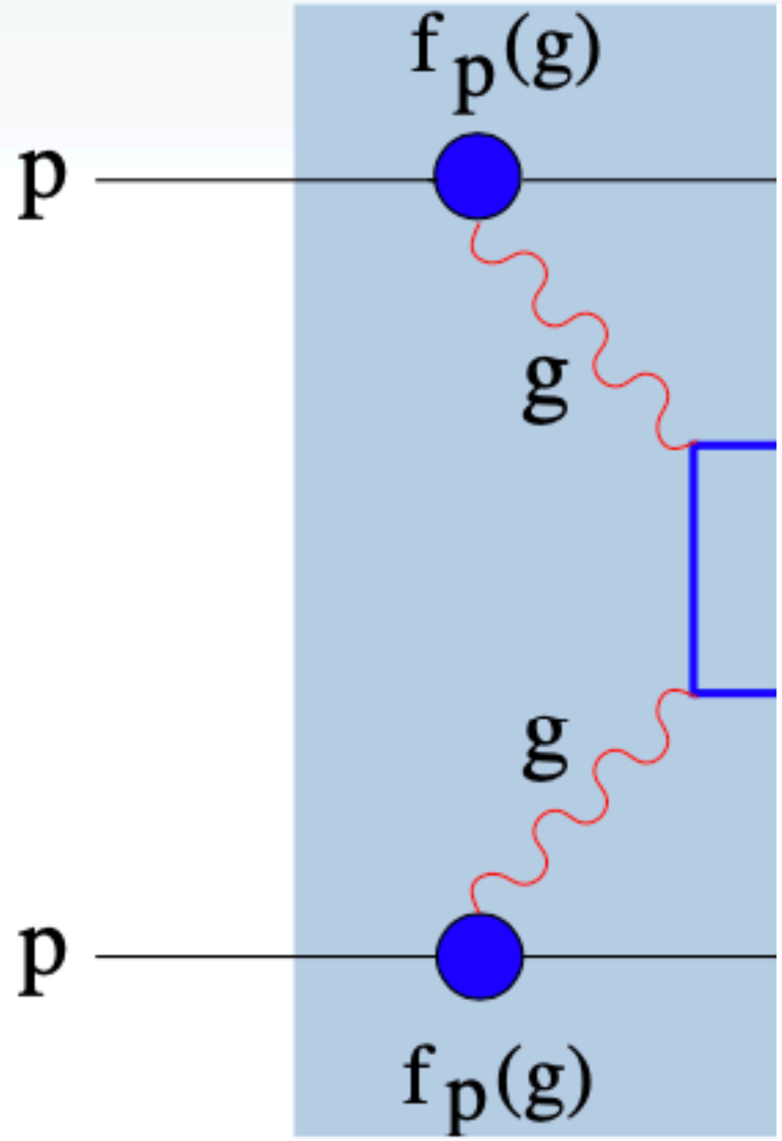




- It all starts with the PDF
- Rare to have Intrinsic charm/ bottom in a proton

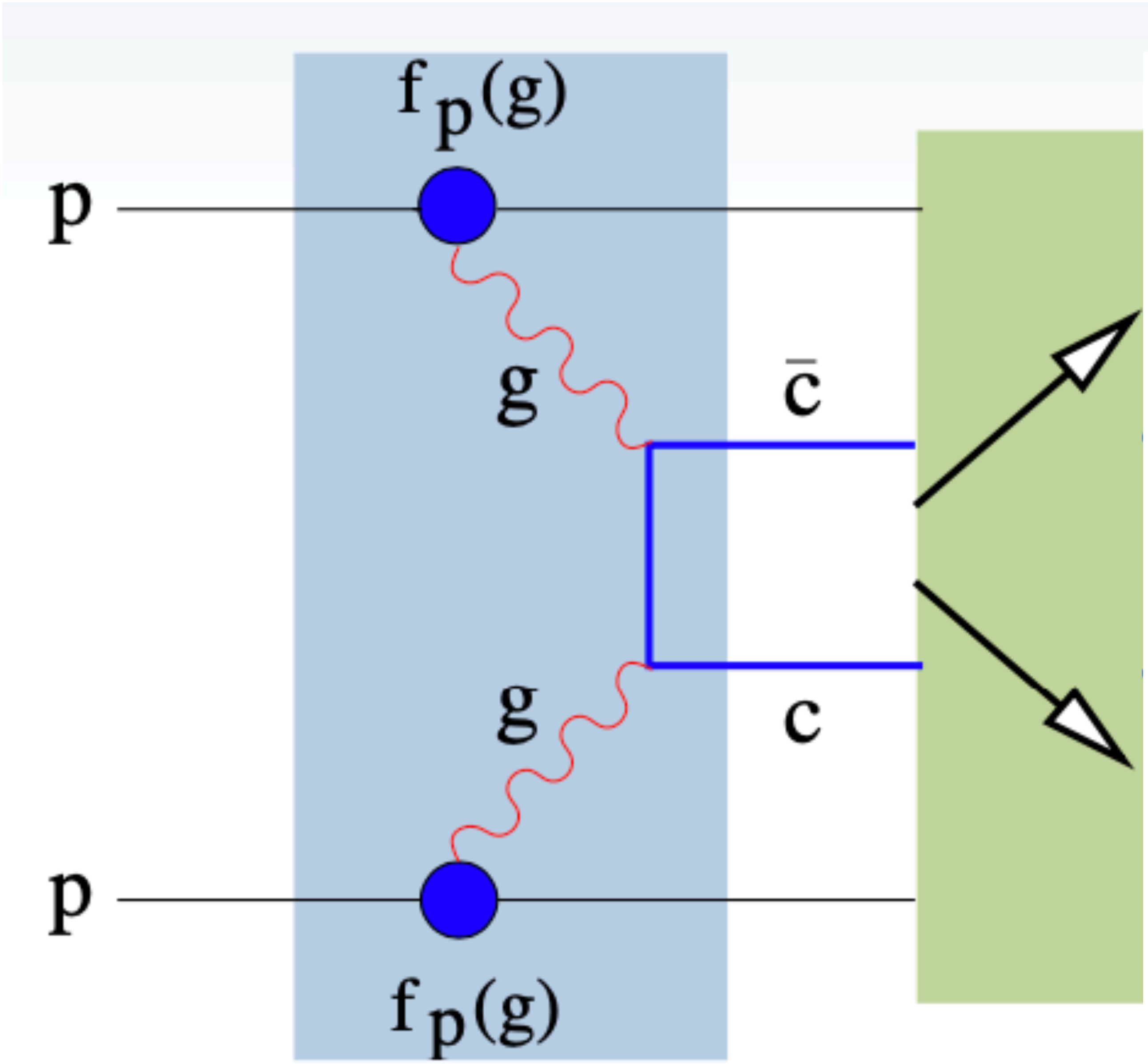
- Heavy quarks production suppressed compared to inclusive quark/gluon - requires high statistics first and foremost

Factorization in heavy flavor



$$d\sigma^{pp\to gg+X} \overset{Q\gg\Lambda_{QCD}}{\approx} \sum_{n,X'} f_g^p(Q^2) \otimes f_g^p(Q^2)$$

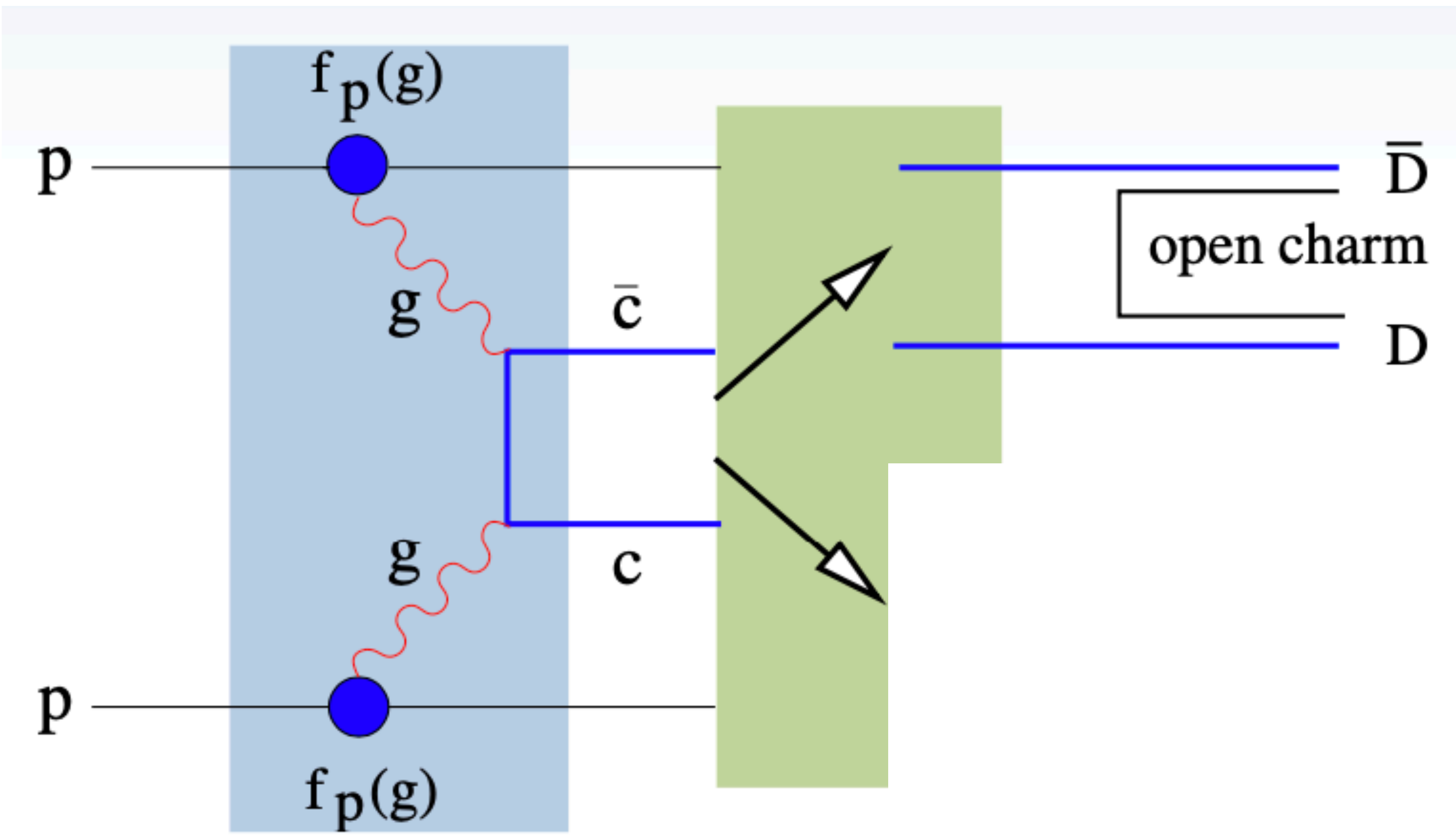
Factorization in heavy flavor



Additional contribution in the $c\bar{c}$ production

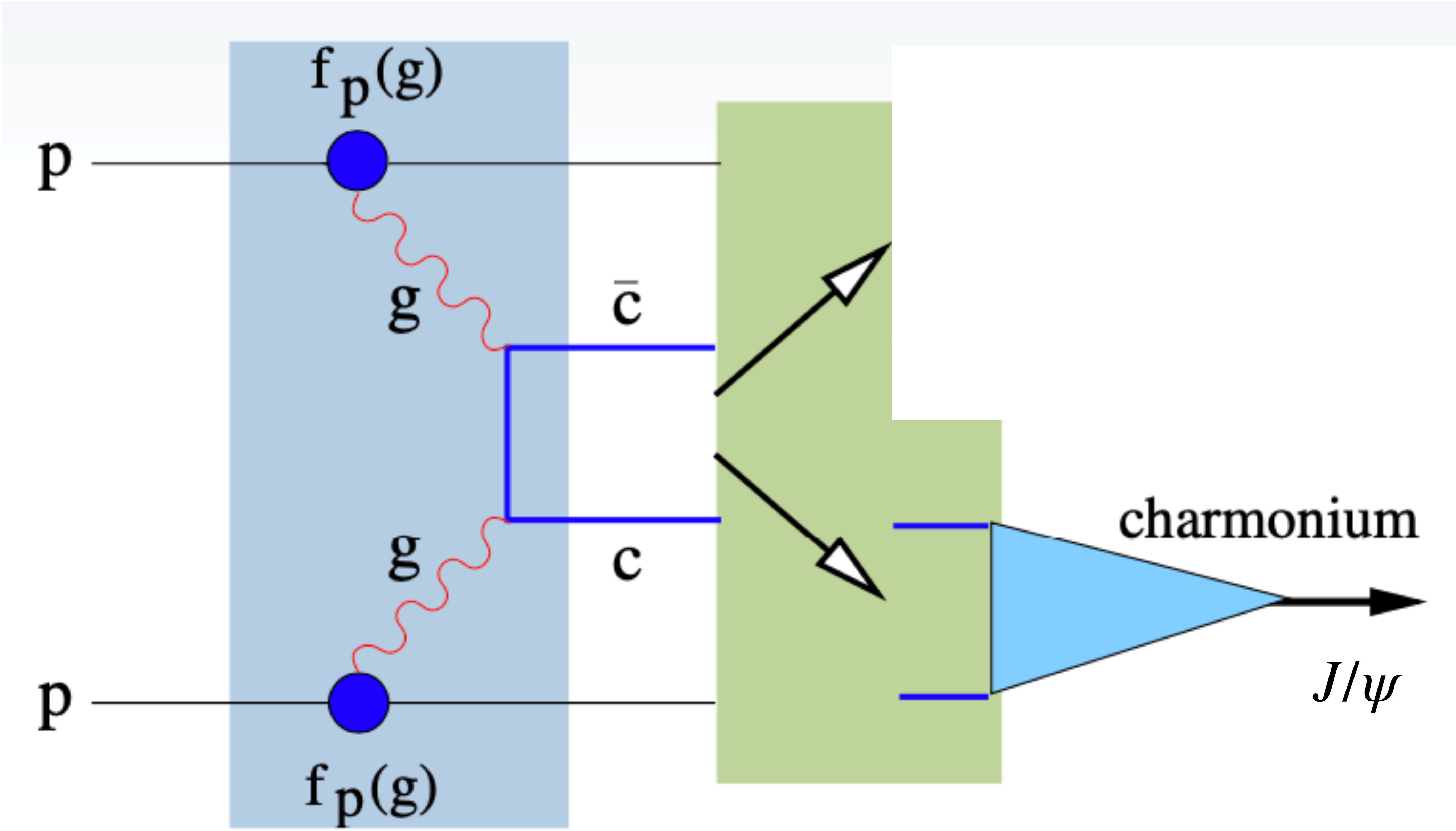
$$d\sigma^{pp\to c\bar{c}+X} \overset{Q\gg\Lambda_{QCD}}{\approx} \sum_{n,X'} f_g^p(Q^2) \otimes f_g^p(Q^2) \otimes d\sigma_n^{gg\to c\bar{c}+X'}$$

Factorization in heavy flavor



- Charm quark hadronizes in vacuum and turns into open charm hadron (D^0, Λ_c etc…)

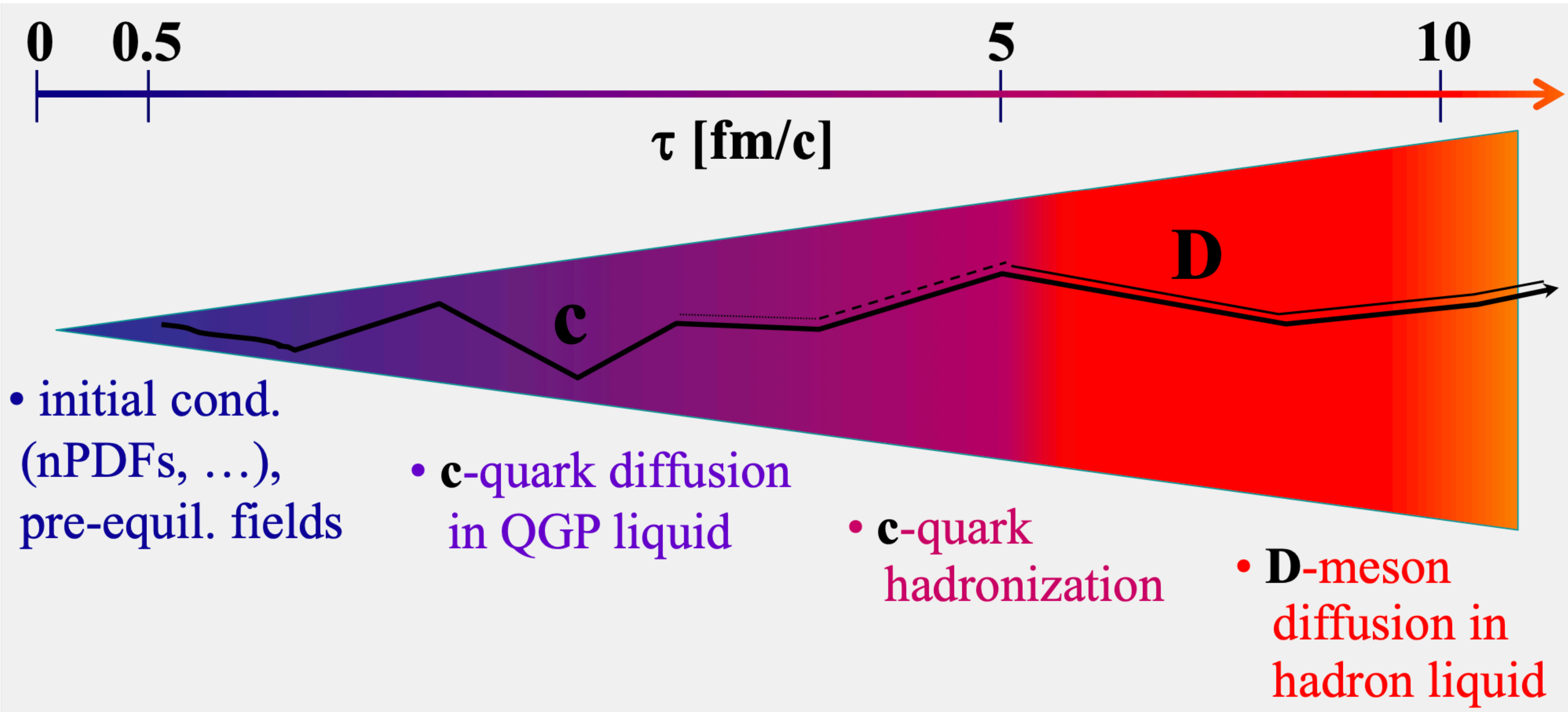
Factorization in heavy flavor



Slide from Alexander Rothkopf

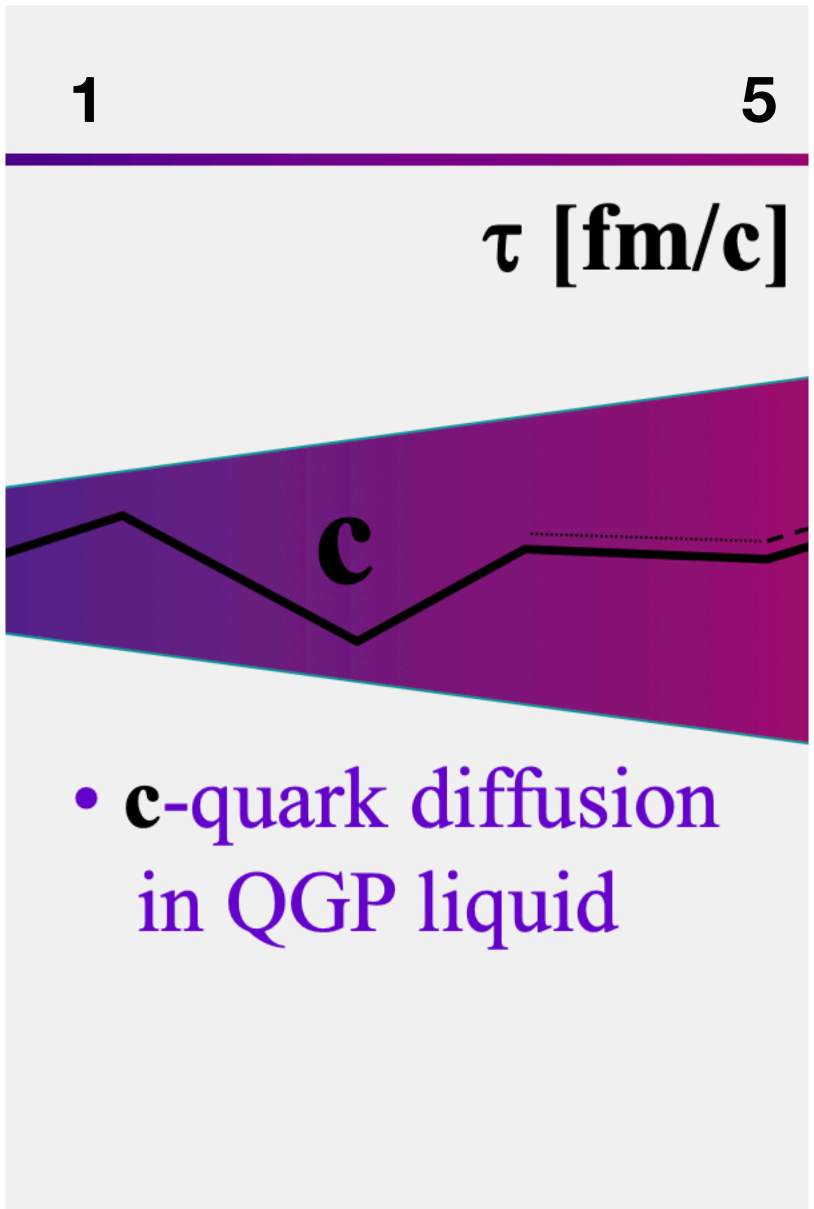
- $c\bar{c}$ bound states J/ψ, $\Psi(2S)$ etc… could potentially also be created and carry away a majority of the quark's energy

Why do heavy flavor in heavy ions

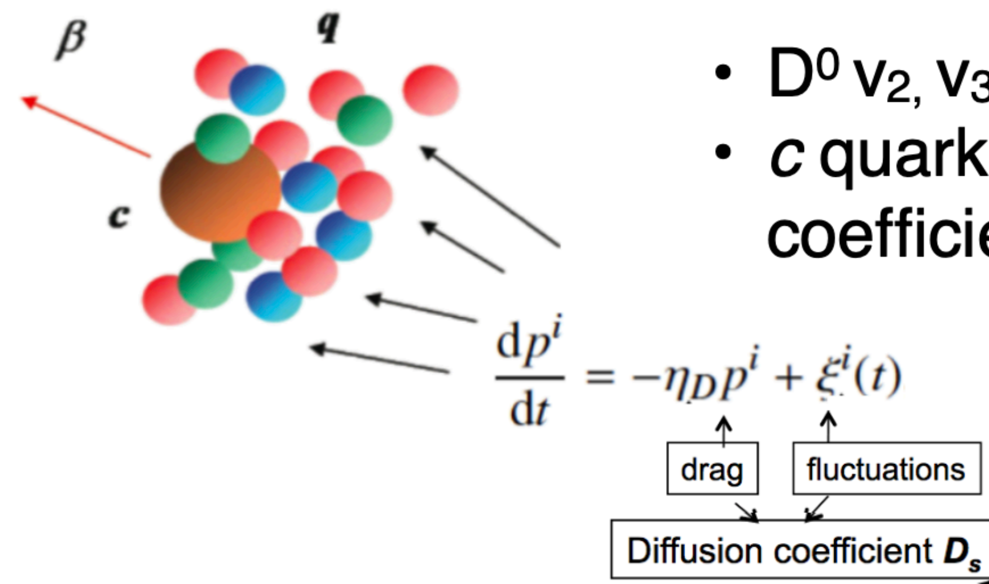


- Each segment deals with potential physics question

Why do heavy flavor in heavy ions

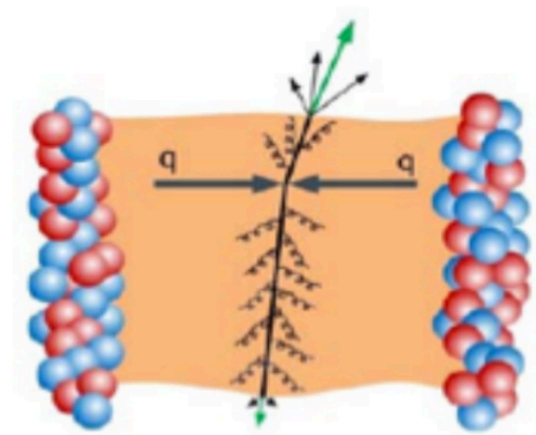


Transport



- D^0 v_2, v_3
- *c* quark diffusion coefficient in QGP

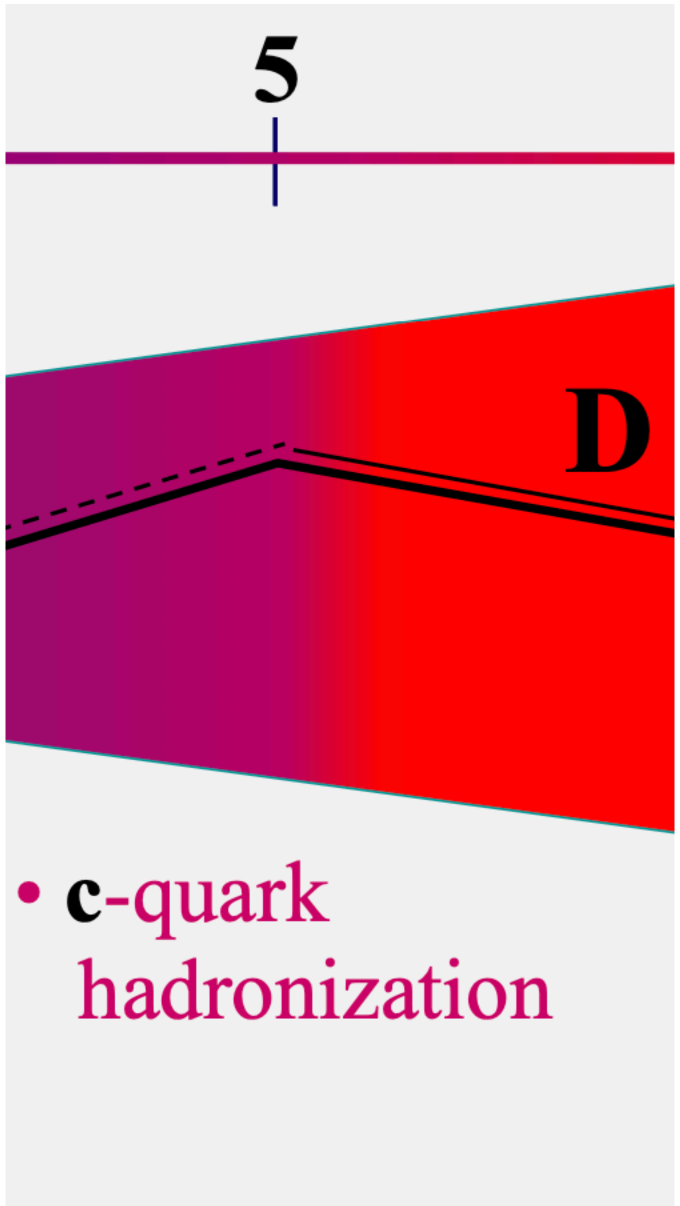
Energy Loss



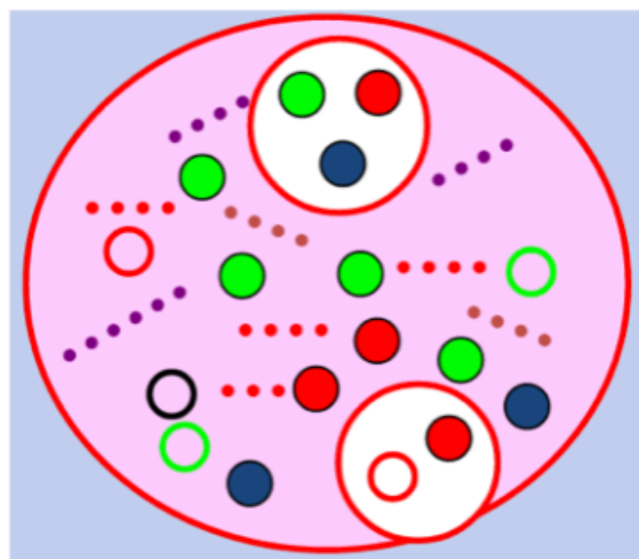
- D, B meson R_{AA} and R_{CP}
- Collisional and radiative energy loss

Mechanism of charm/bottom interactions with the QGP

Why do heavy flavor in heavy ions



Hadronization

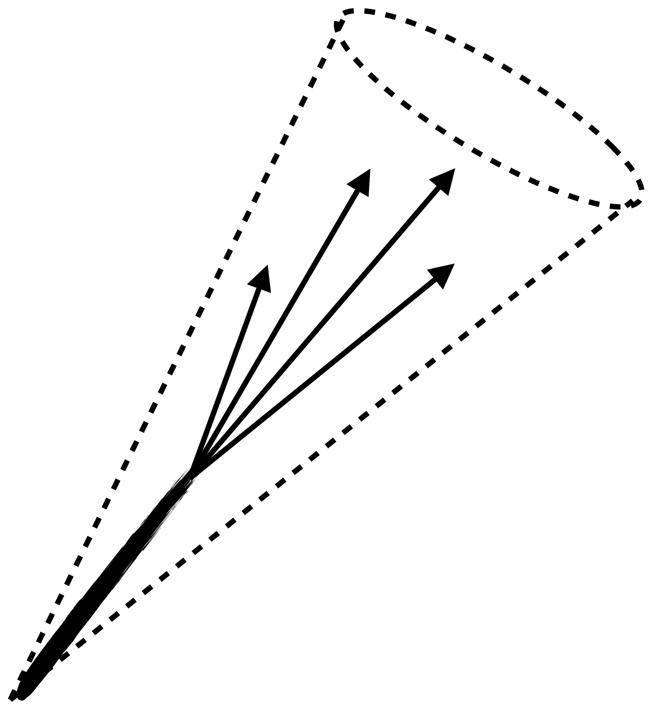


?

- Λ_c, D_s production
- Coalescence?
- Ideal probes as total c quark is fixed at initial scatterings

Perturbative to non-perturbative transition w/ a mass scale

Two sets of objects



$c\,\bar{c}$

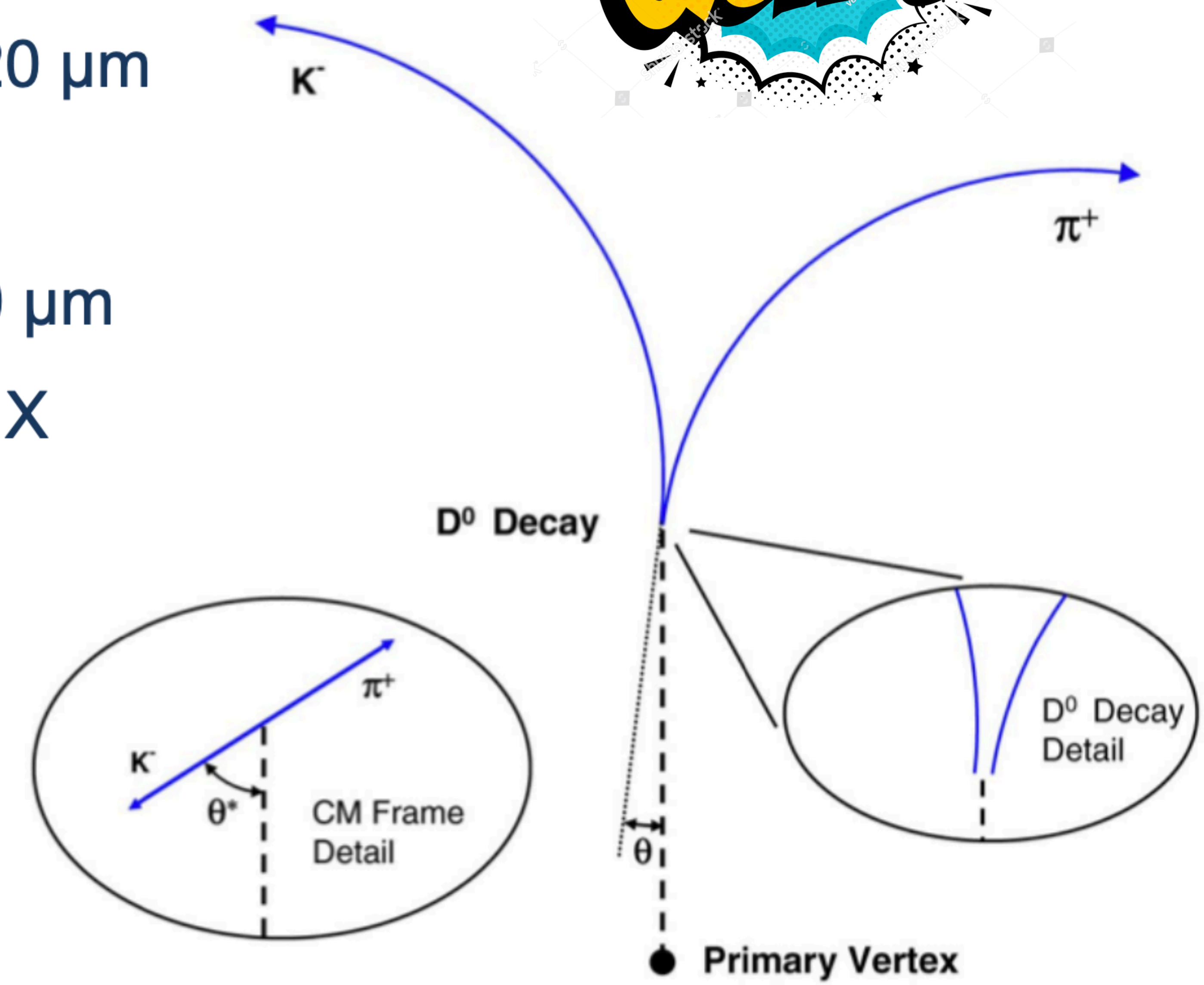
$c\,\bar{u}$

How do we detect or reconstruct them?

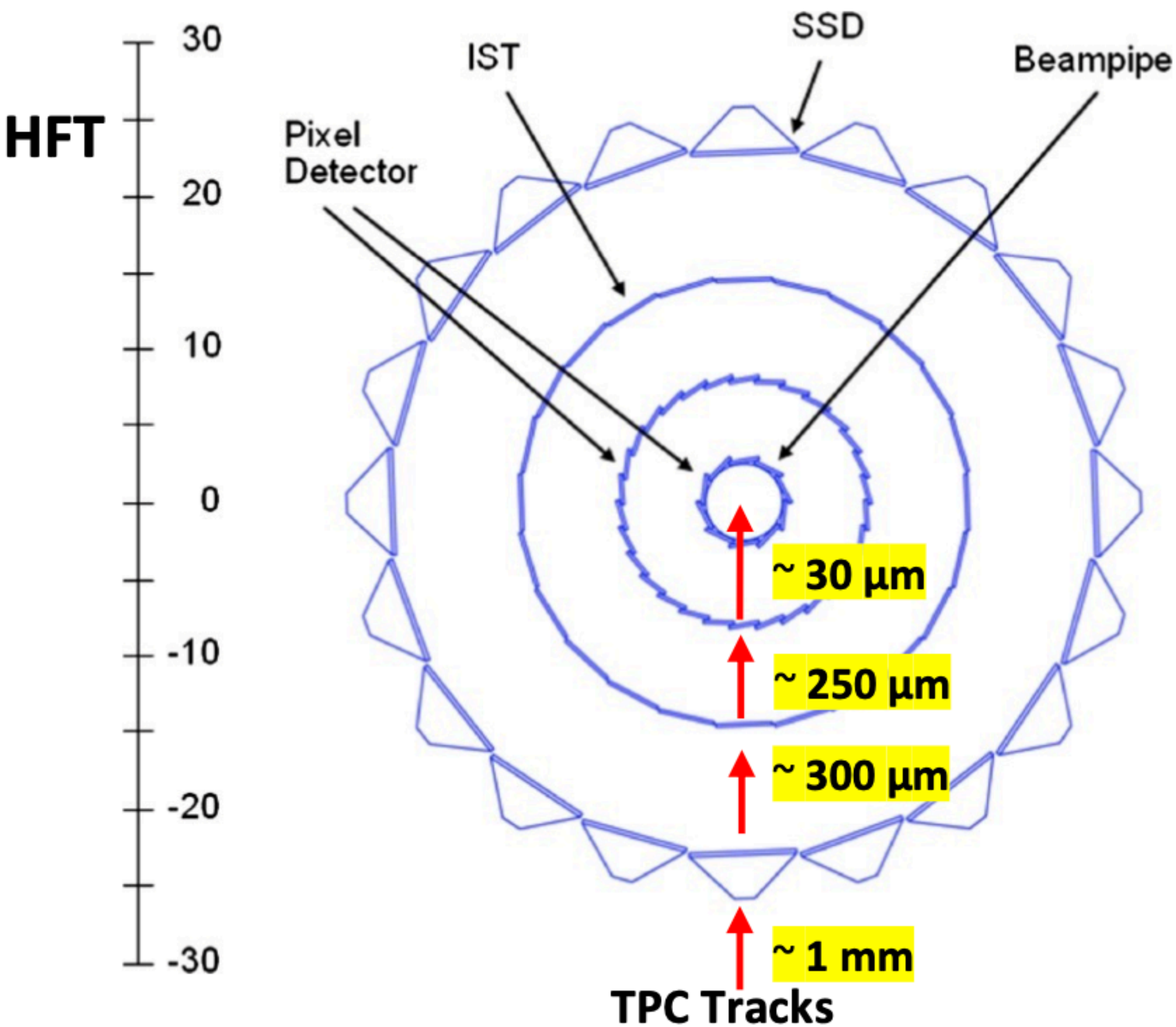
Reconstruct HF hadrons in experiment - Vertex trackers

- $D^0 \rightarrow K^- \pi^+$
  - BR = 3.83 % cτ ~ 120 μm
- $\Lambda_c^+ \rightarrow p\, K^- \pi^+$
  - BR = 5.0 % cτ ~ 60 μm
- B mesons → J/ψ + X or e + X
  - cτ ~ 500 μm

- Reconstruct the decay daughters tracks, extrapolate its curvature to the primary vertex and calculate the distance of closest approach

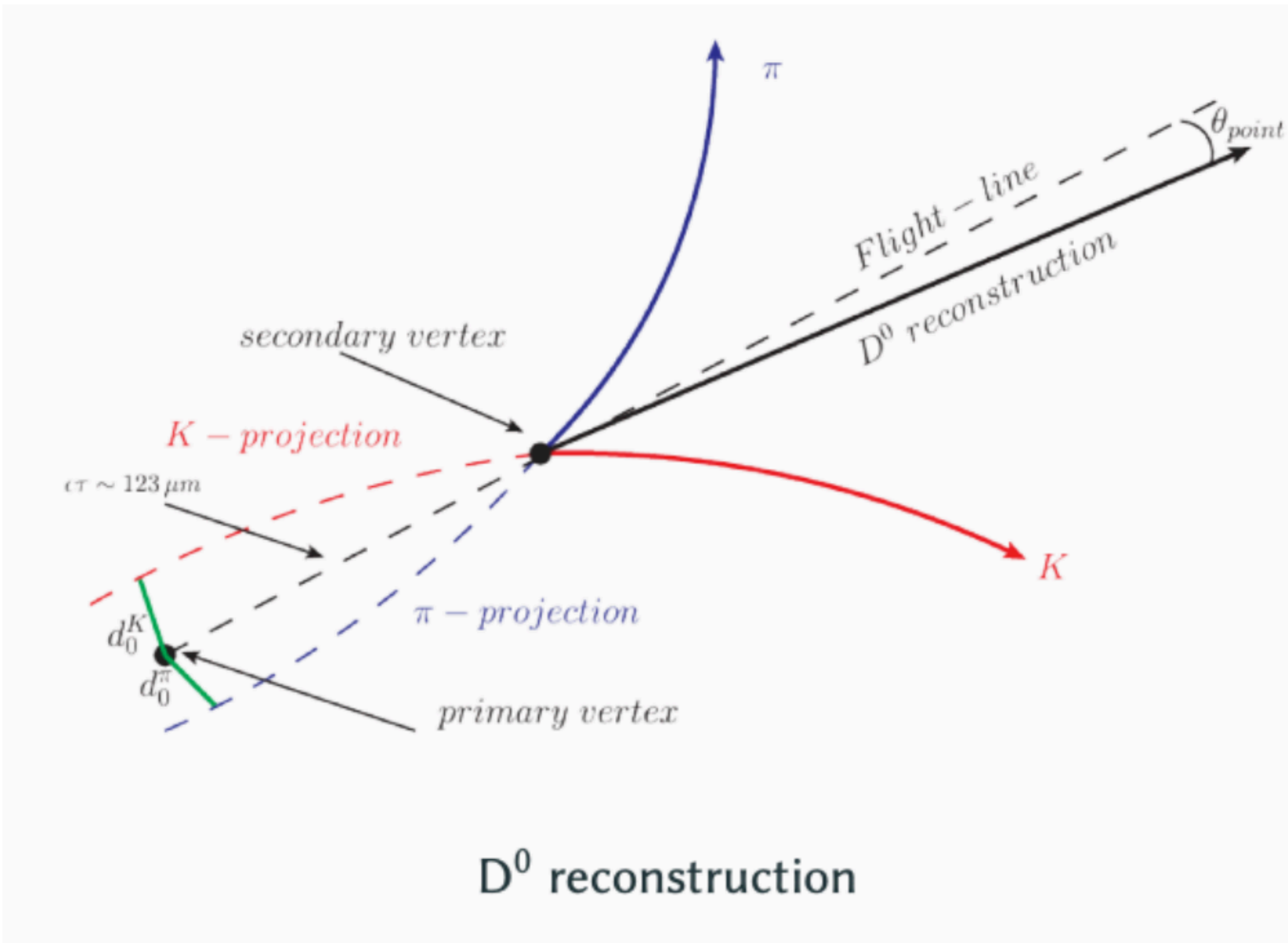


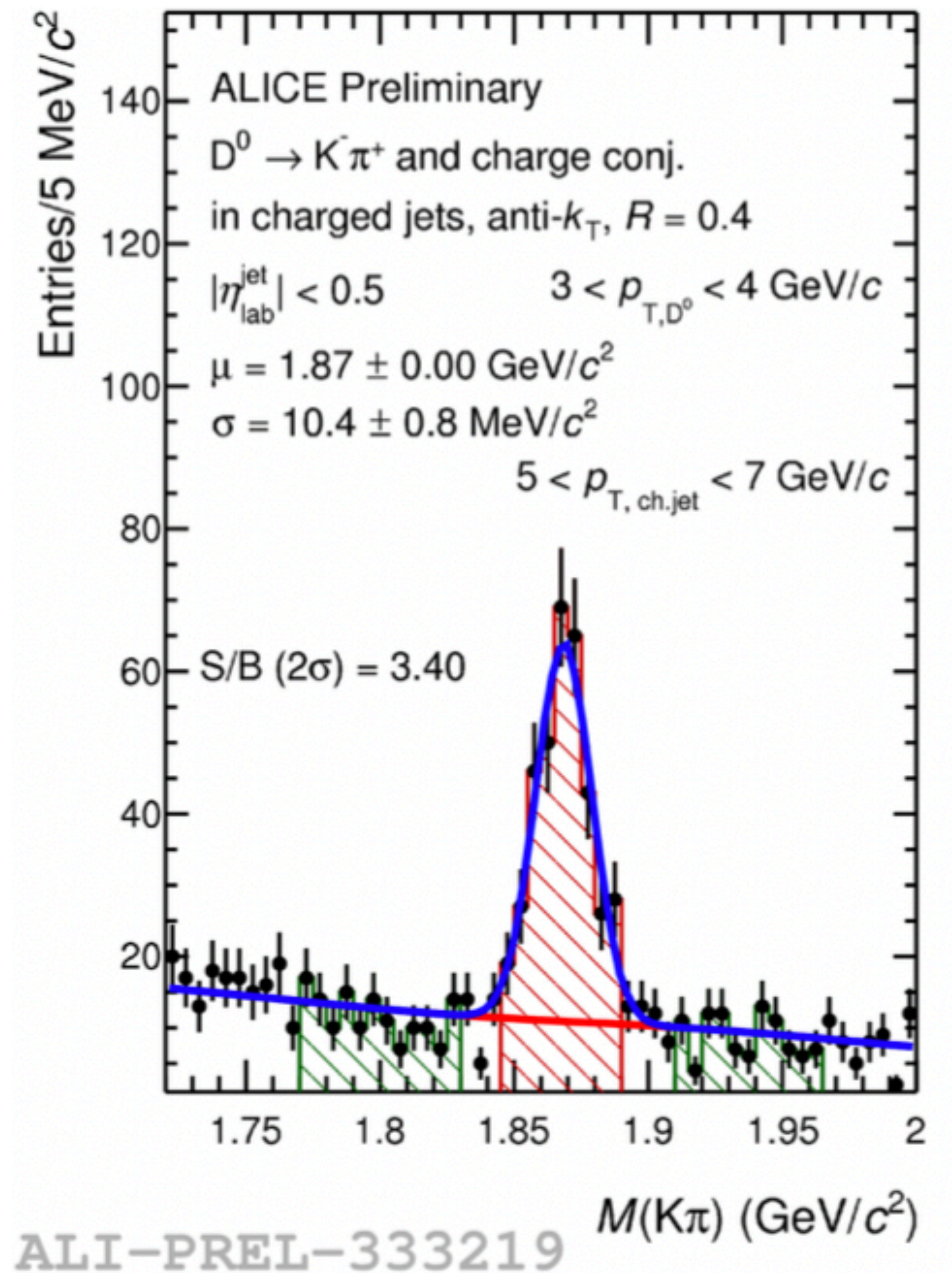
STAR Heavy Flavor Tracker



- Installed and run in STAR during 2014-2016
- 2 layers of pixel detector followed by a silicon tracker and a strip detector

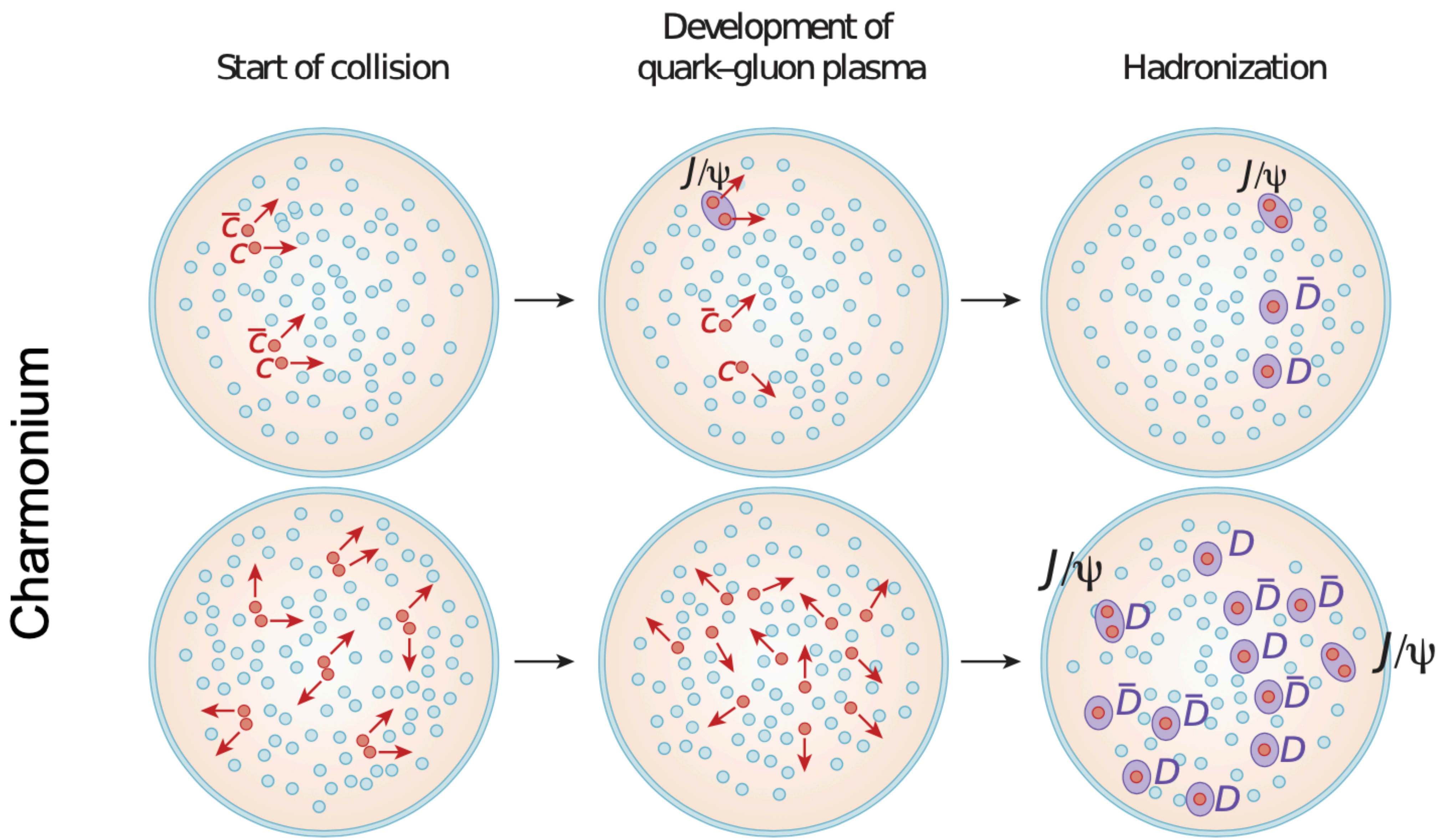
D^0 reconstruction





- Fit to the invariant mass of the charged pairs
- Extract the signal yield after subtracting background

Different production mechanisms



Significant melting at at RHIC

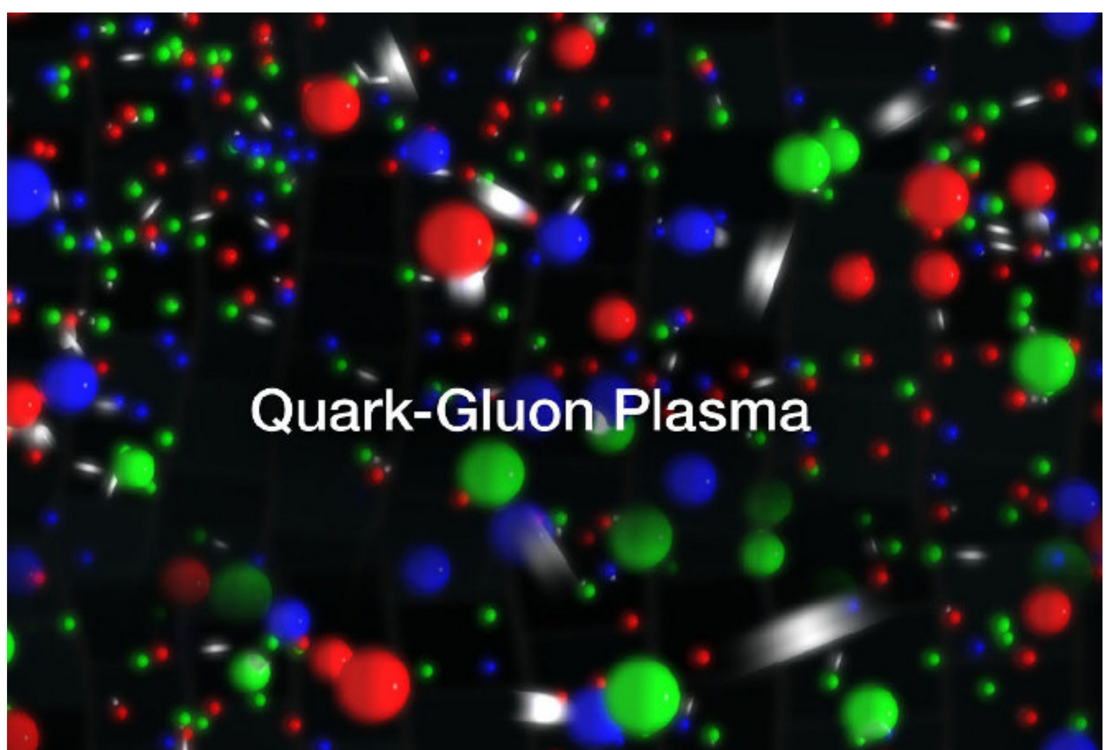
Significant melting and recombination at LHC

Central AA collisions	RHIC 200 GeV	LHC 5 TeV
N_{ccbar}/event	~10	~115

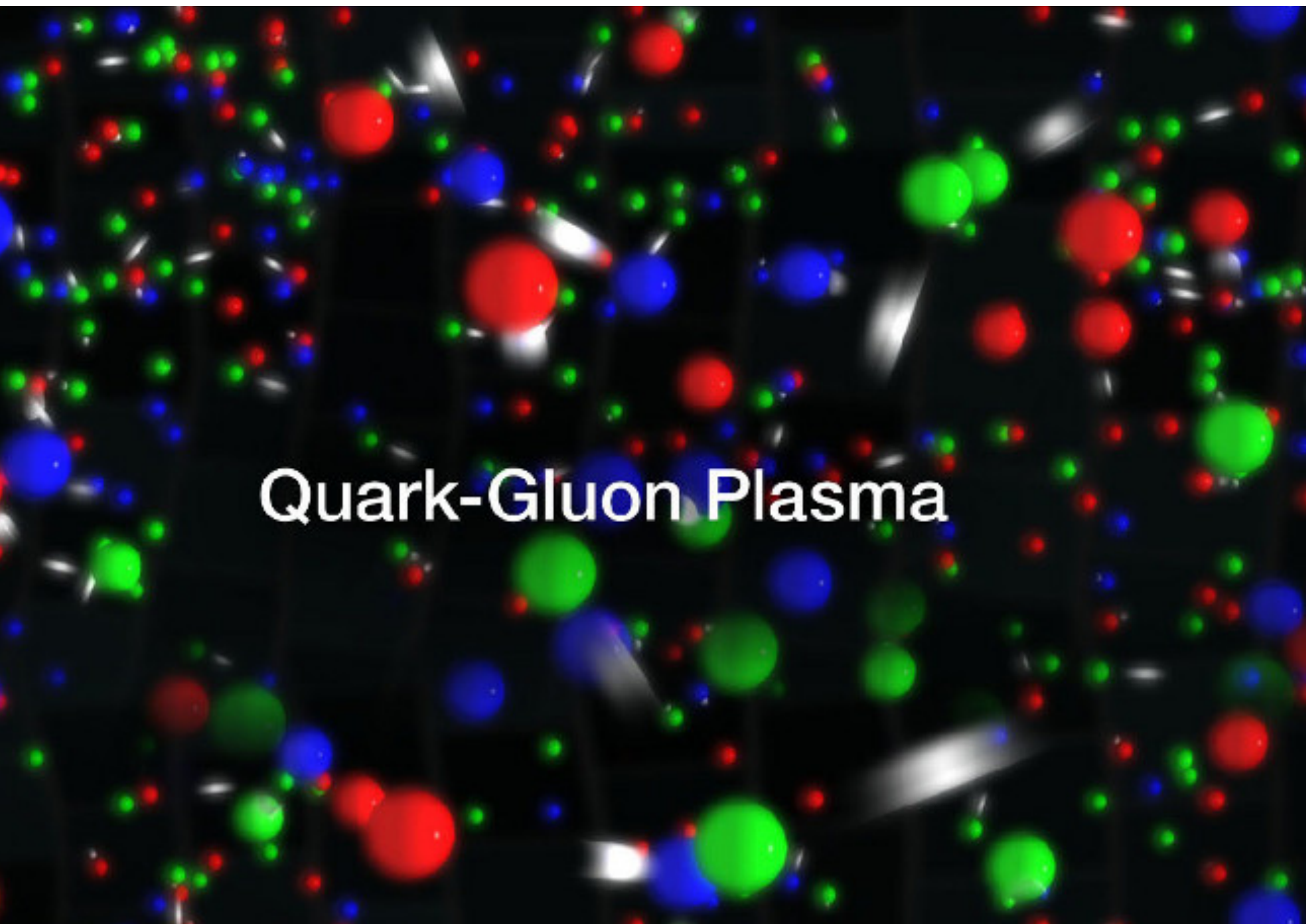
Slide from Alexander Rothkopf

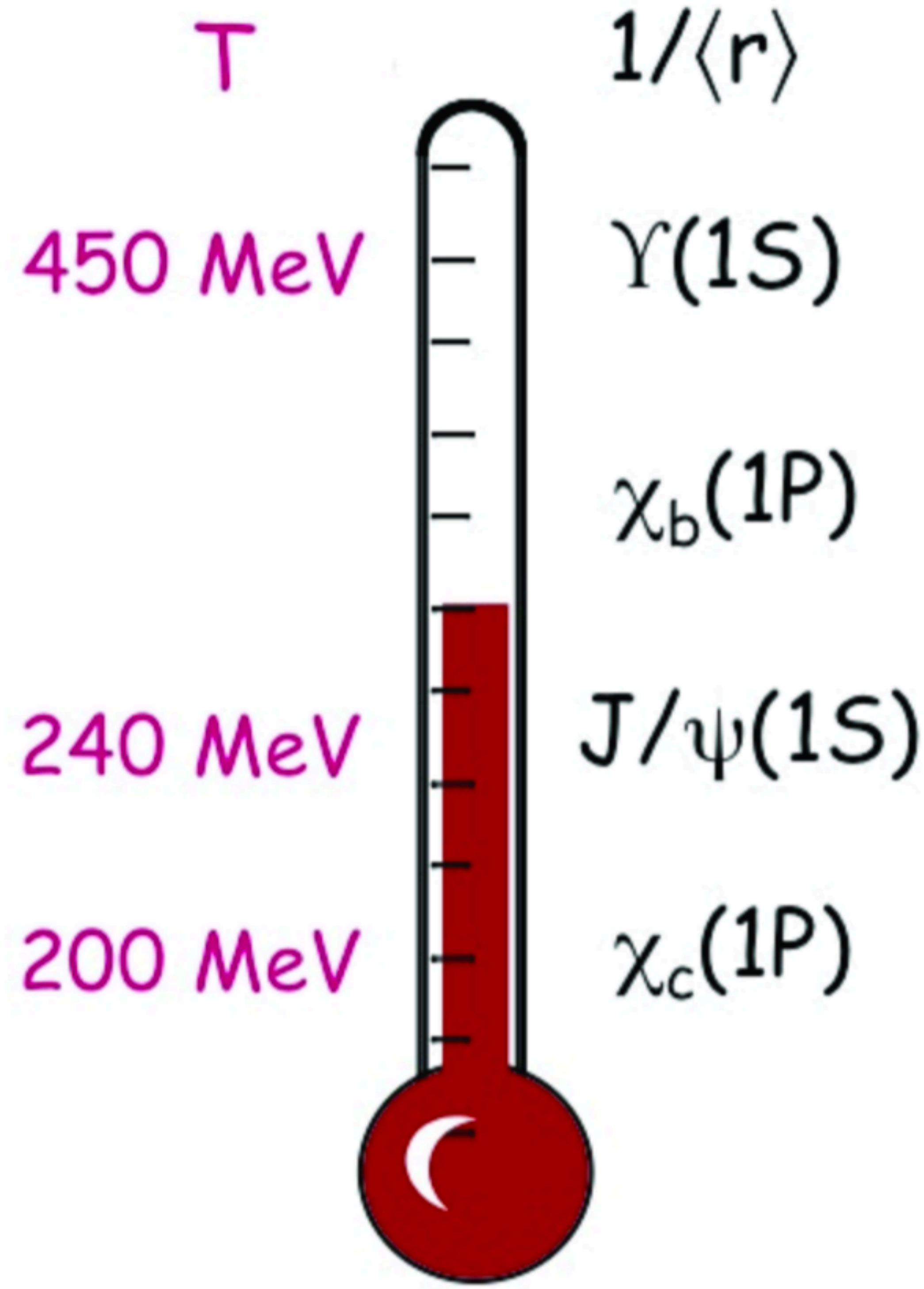
Why are they considered a thermometer

0.59fm	0.86fm	1.2fm	
0.54GeV	0.2GeV	0.06GeV	Binding Energy
$\Upsilon(^3S_{1(n=2)})$	$\Upsilon(^3S_{1(n=3)})$	$\psi'(^3S_{1(n=2)})$	



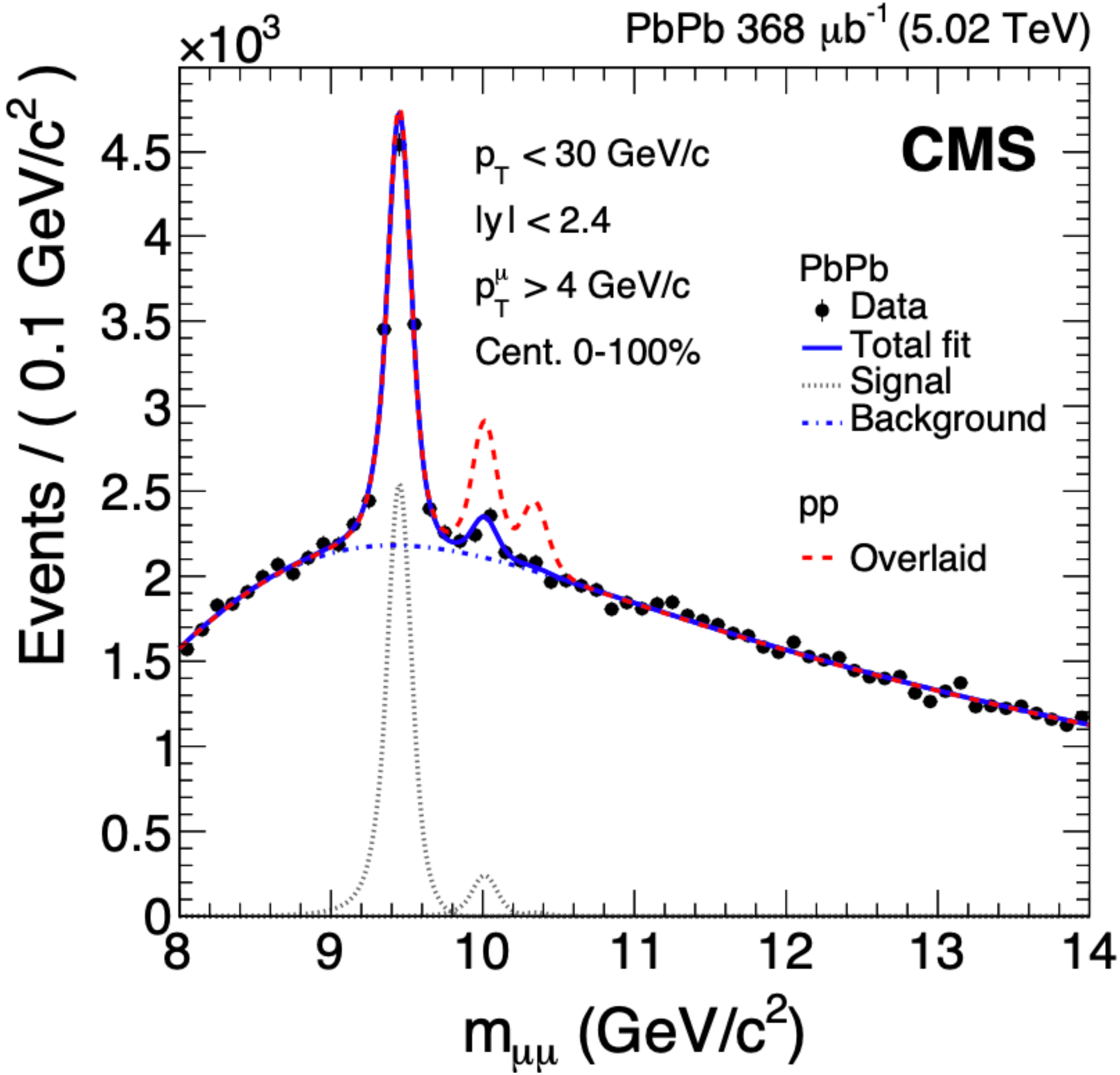
Temperature of the plasma



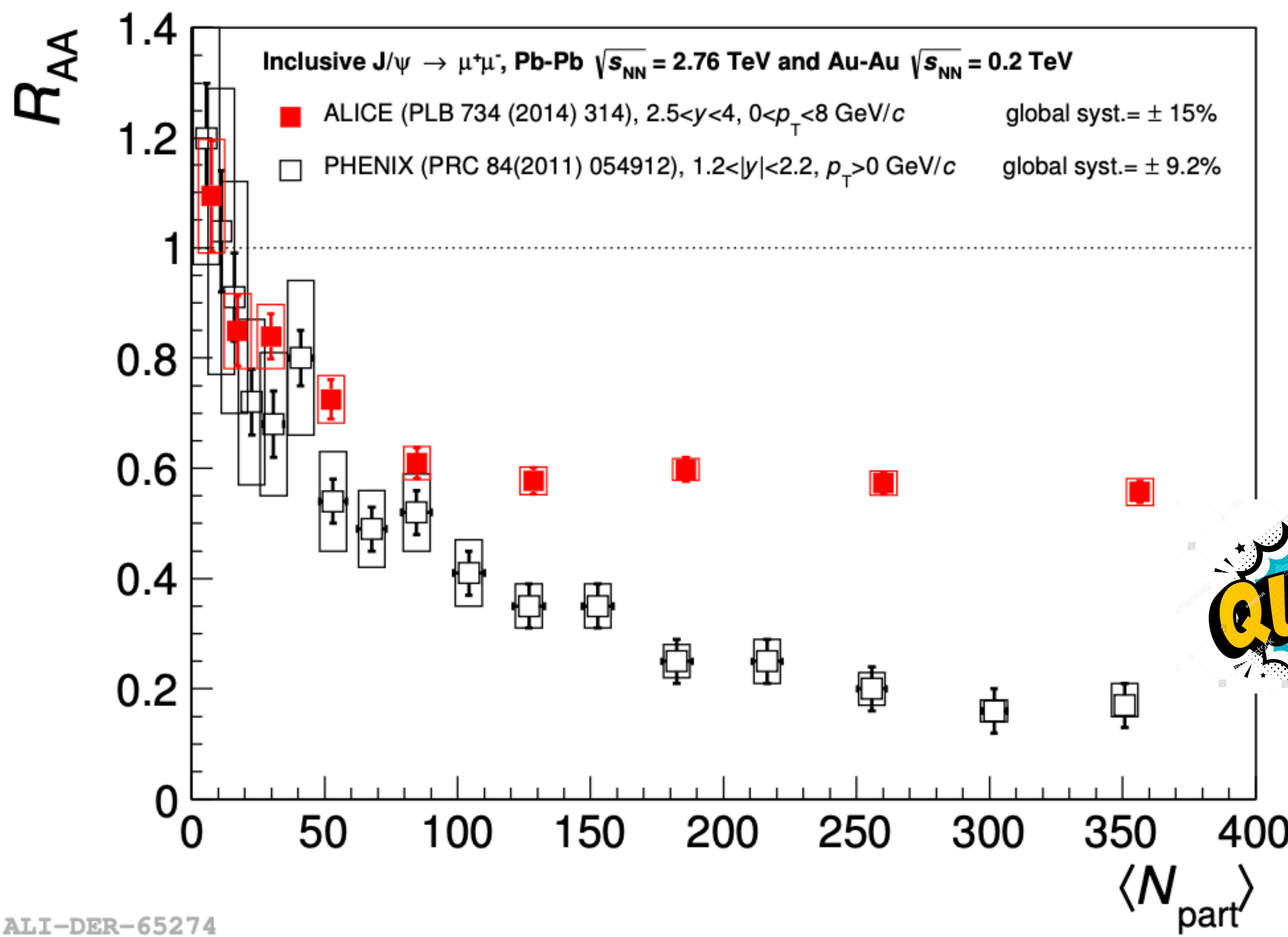


A. Mócsy, P. Petreczky, M. Strickland
Int. J. of Mod. Phys. A
Vol. 28 (2013) 1340012

- Cartoon picture of temperature dependence of the yield



- Significantly smaller RAA for higher Upsilon states compared to expectation!



- Significantly smaller RAA for J/Psi at RHIC compared to LHC

Recap - 2

- Several physics topics are accessible with heavy flavor mesons from initial to final state
- Basic measurements are studies of the yield as compared to proton-proton and between heavy ions
- ‘See’ signature of sequential melting and evidence for recombination

Now on to HF jets!

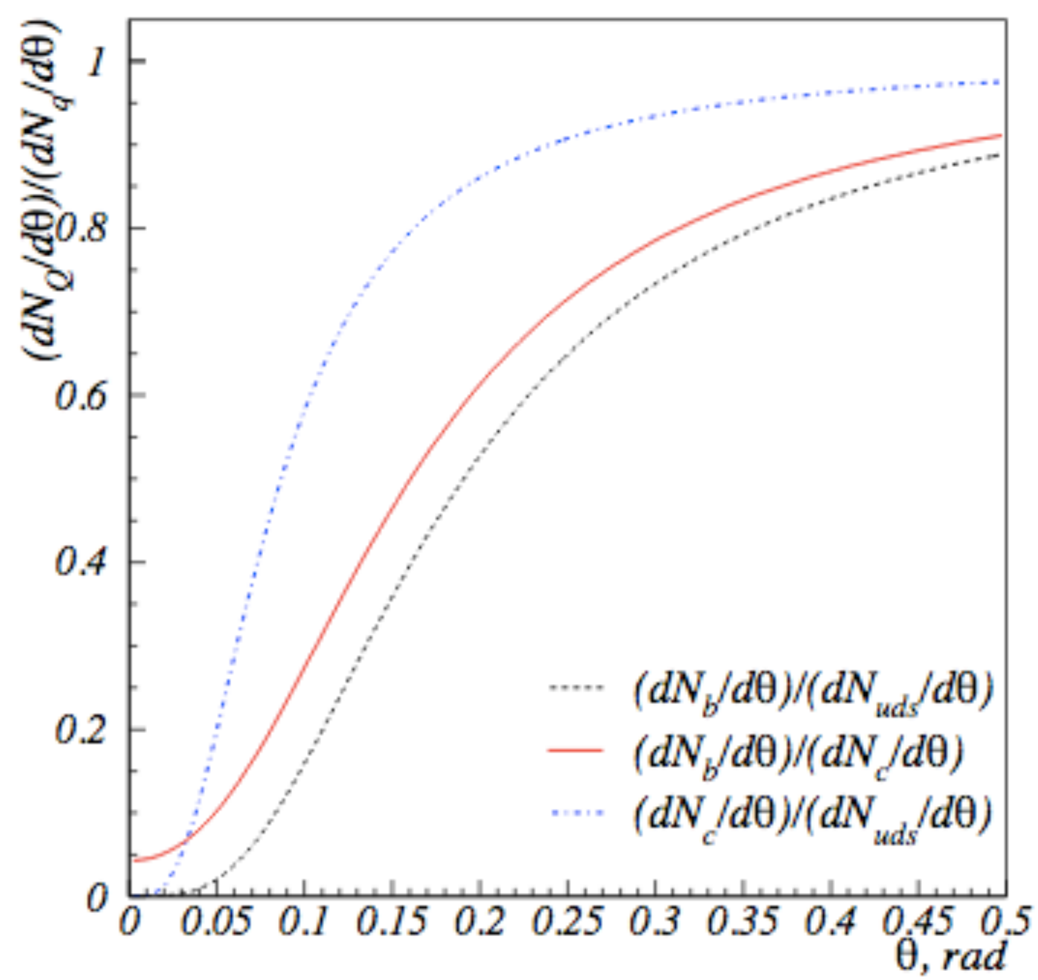
Physics of heavy flavor jets in vacuum

Gluon radiation by a particle of mass m and energy E is suppressed within a cone of angular size m/E around the emitter

$$\frac{\frac{dN_Q}{d\theta}}{\frac{dN_q}{d\theta}} \propto \frac{\theta^4}{(\theta^2 + \theta_0^2)^2}$$

$$\theta_0 = \frac{m_Q}{E_Q}$$

Battaglia et al,
DELPHI-2004-037 CONF 712



Dead-cone effect

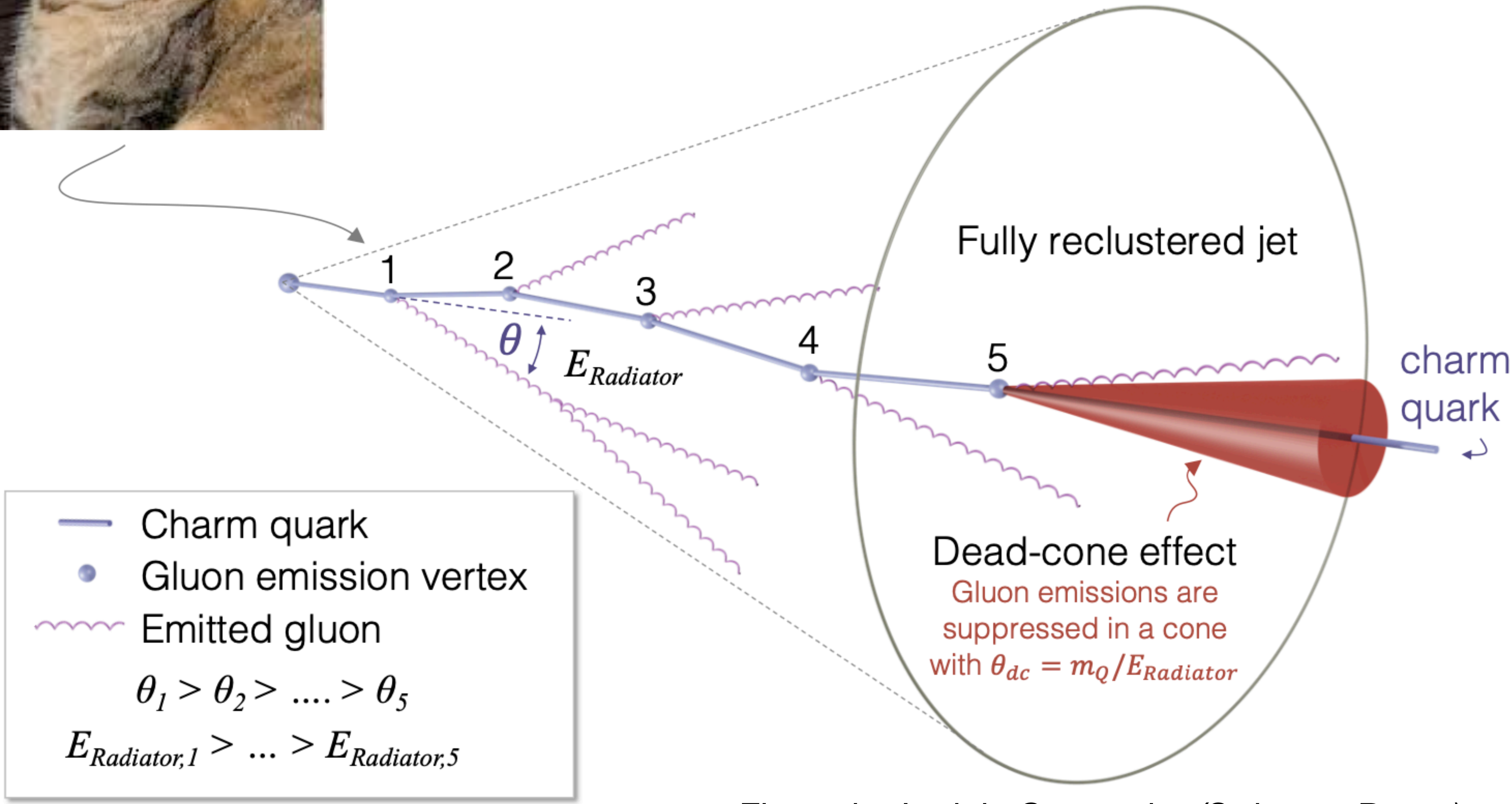
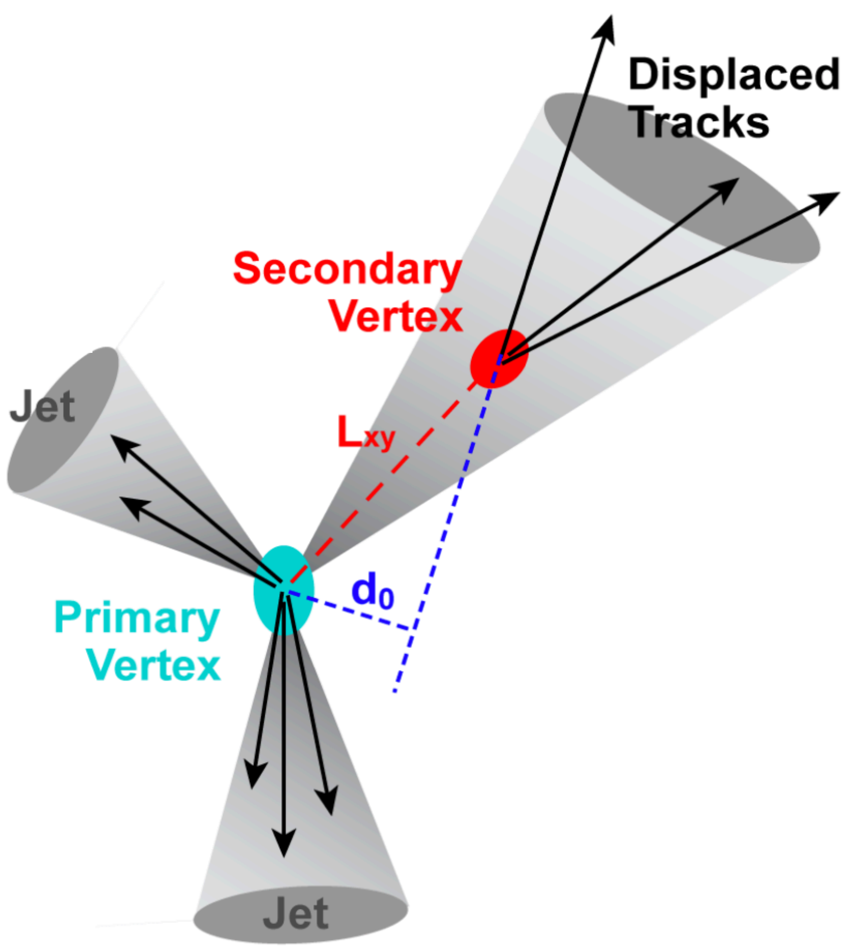


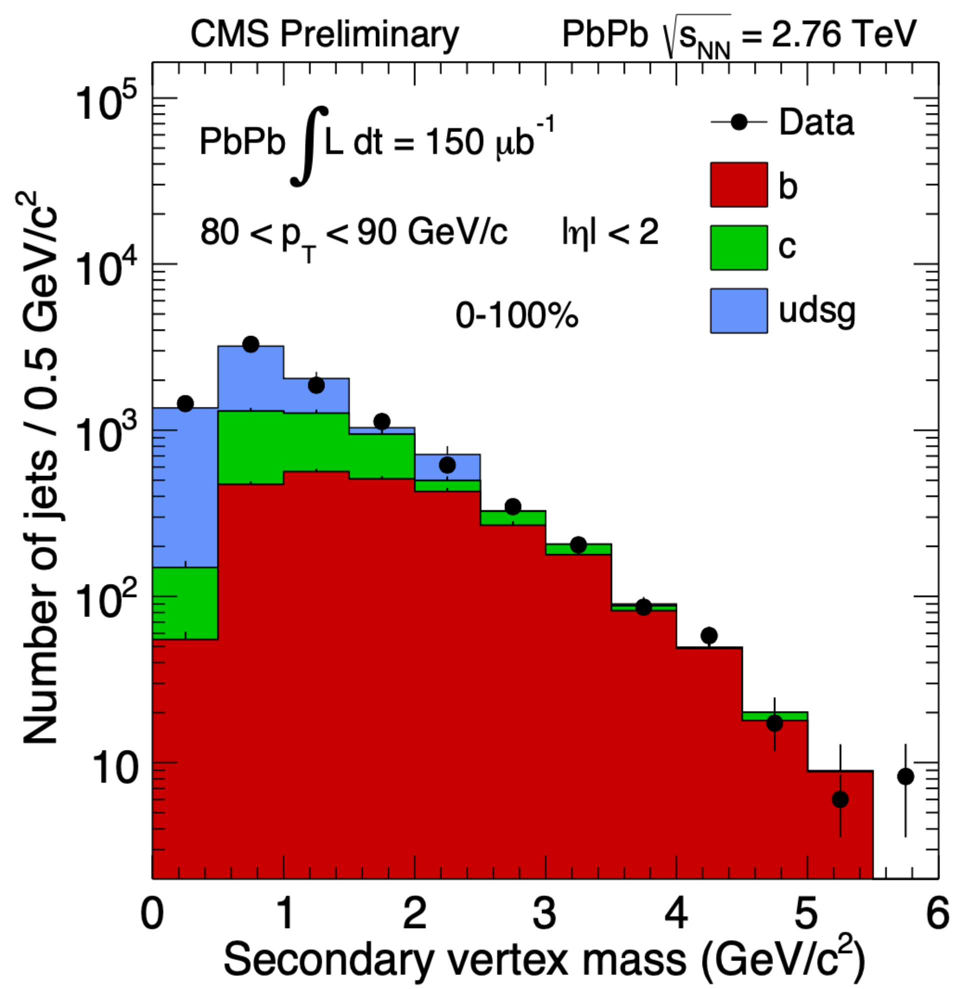
Figure by Leticia Cunqueiro (Spienza, Rome)

How did we select HF jets in the past?

- Identify the secondary vertex within the jet

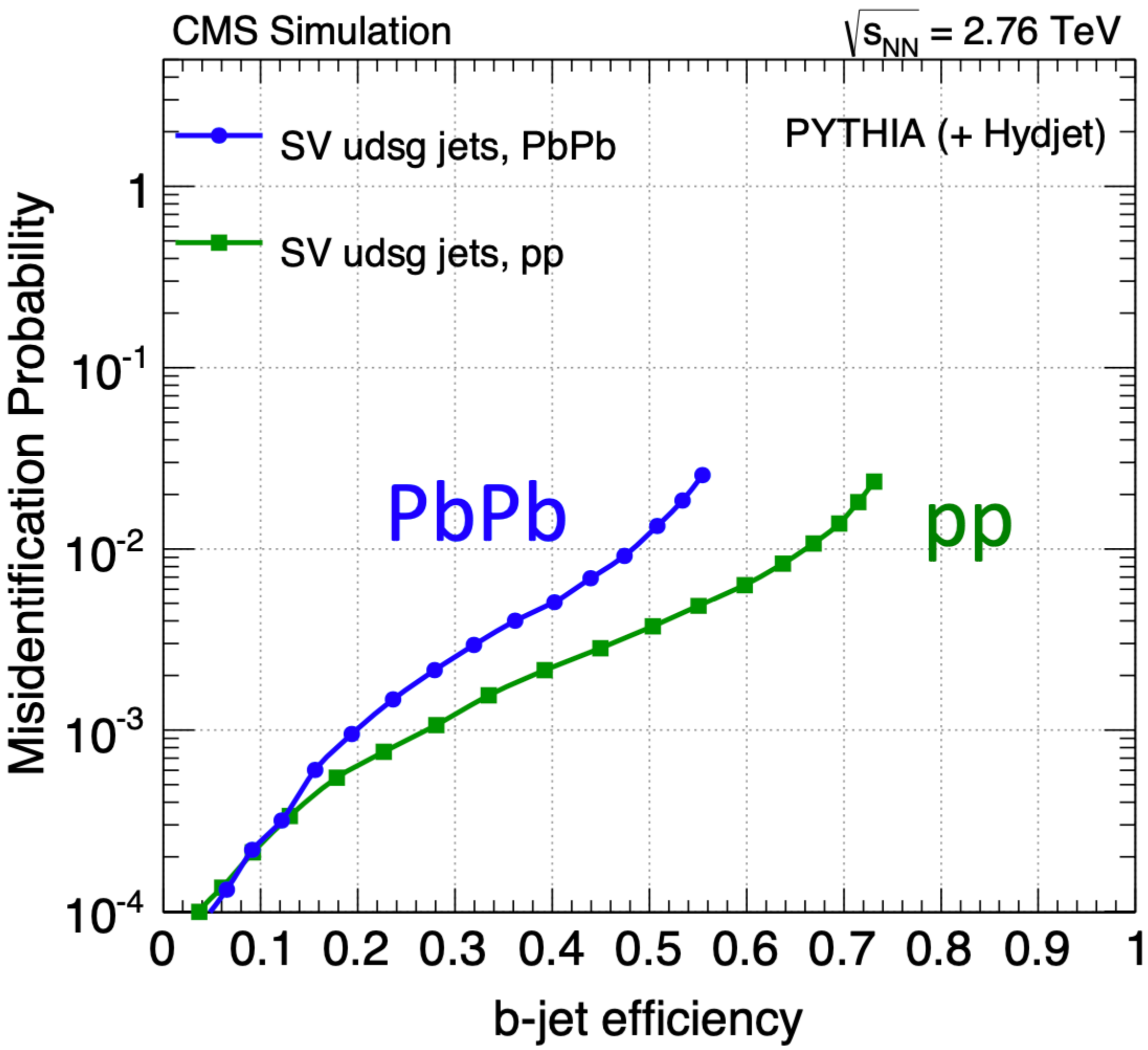


- Extract the fraction from secondary vertex mass

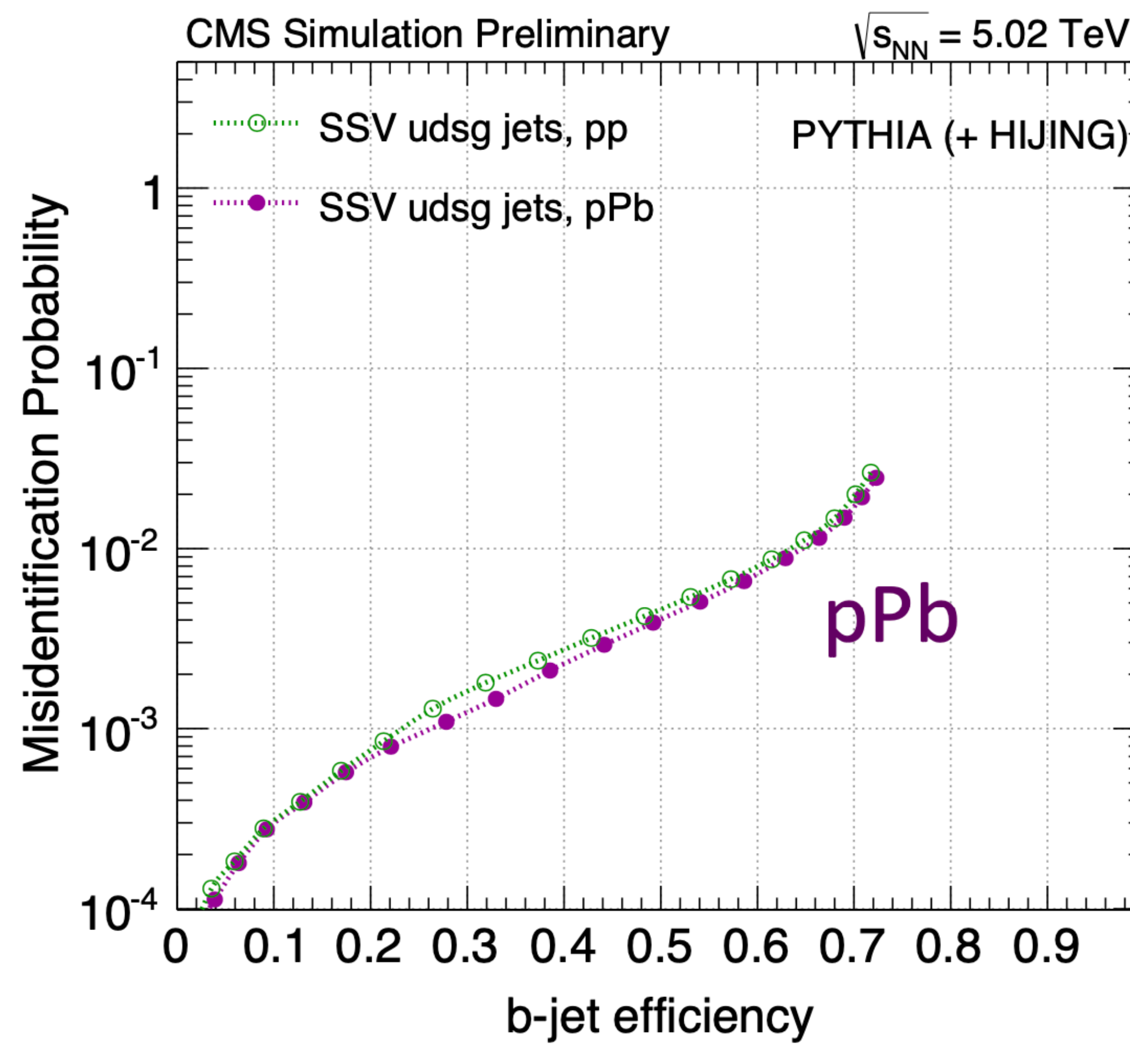


Performance in the past

Performance in PbPb



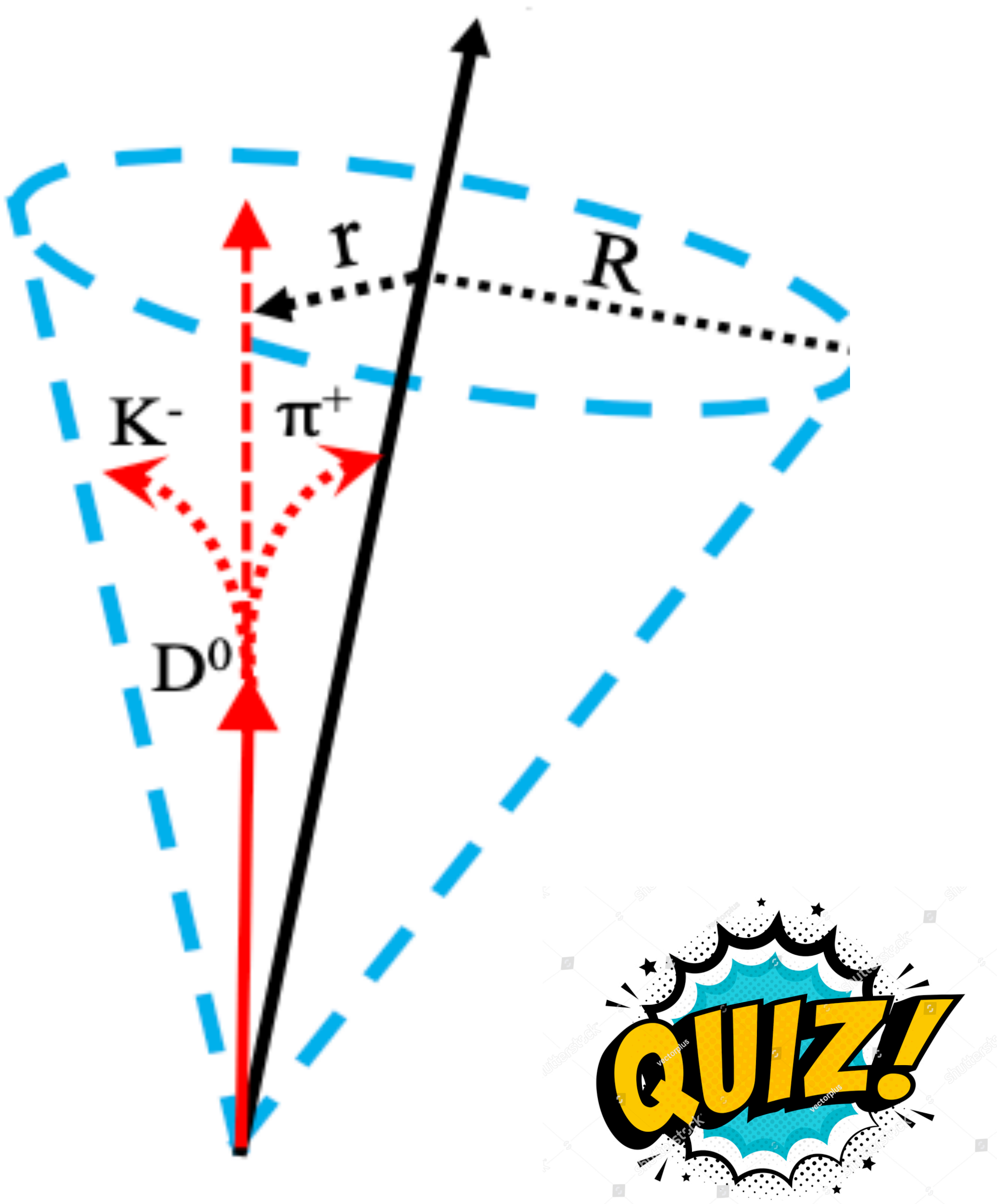
Performance in pPb



- B-jet efficiency plotted against probability of misidentifying a light jets as a b-jet

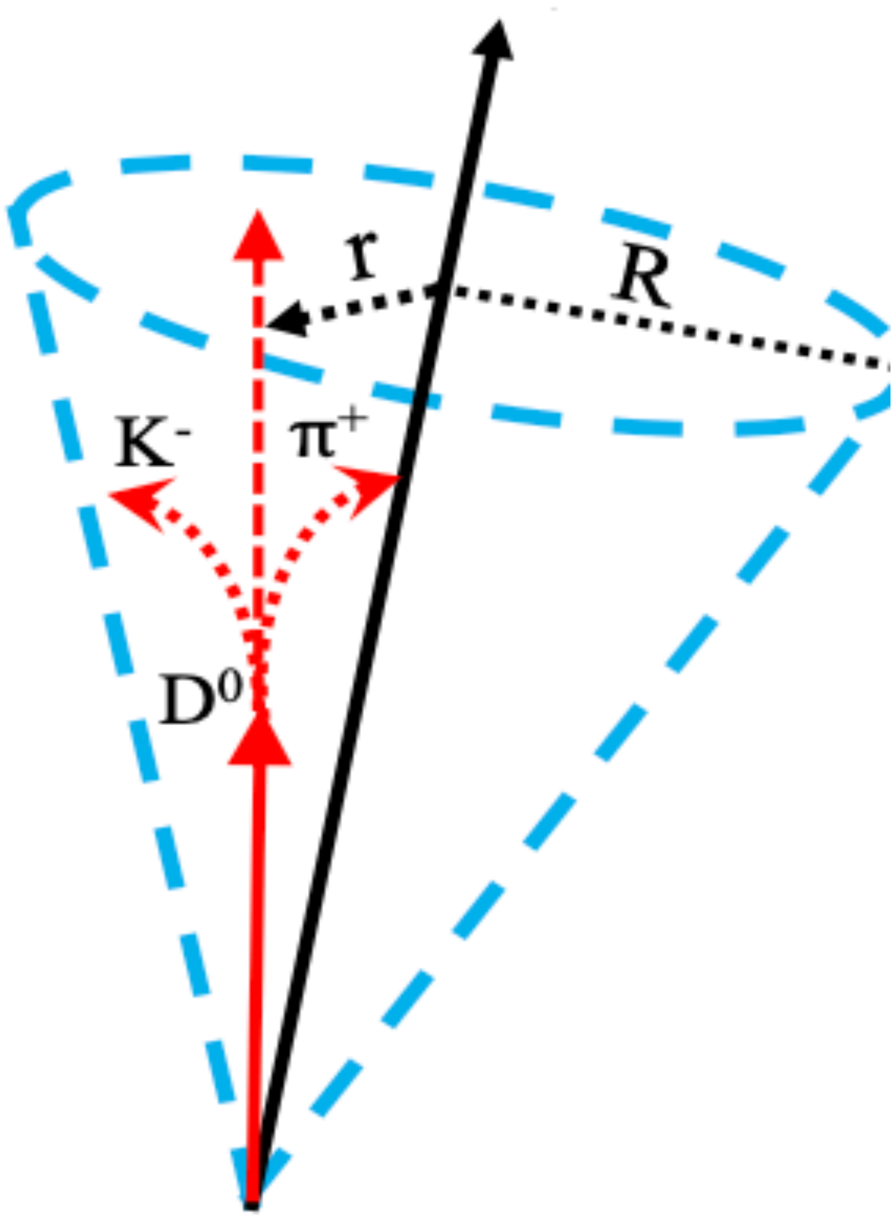
Kurt Jung, QM 2014

How we select HF jets now!



- Meson tagging!
- Enables a clean Monte-Carlo bias free selection of jets which include heavy-quark content
- Are these all HF jets?

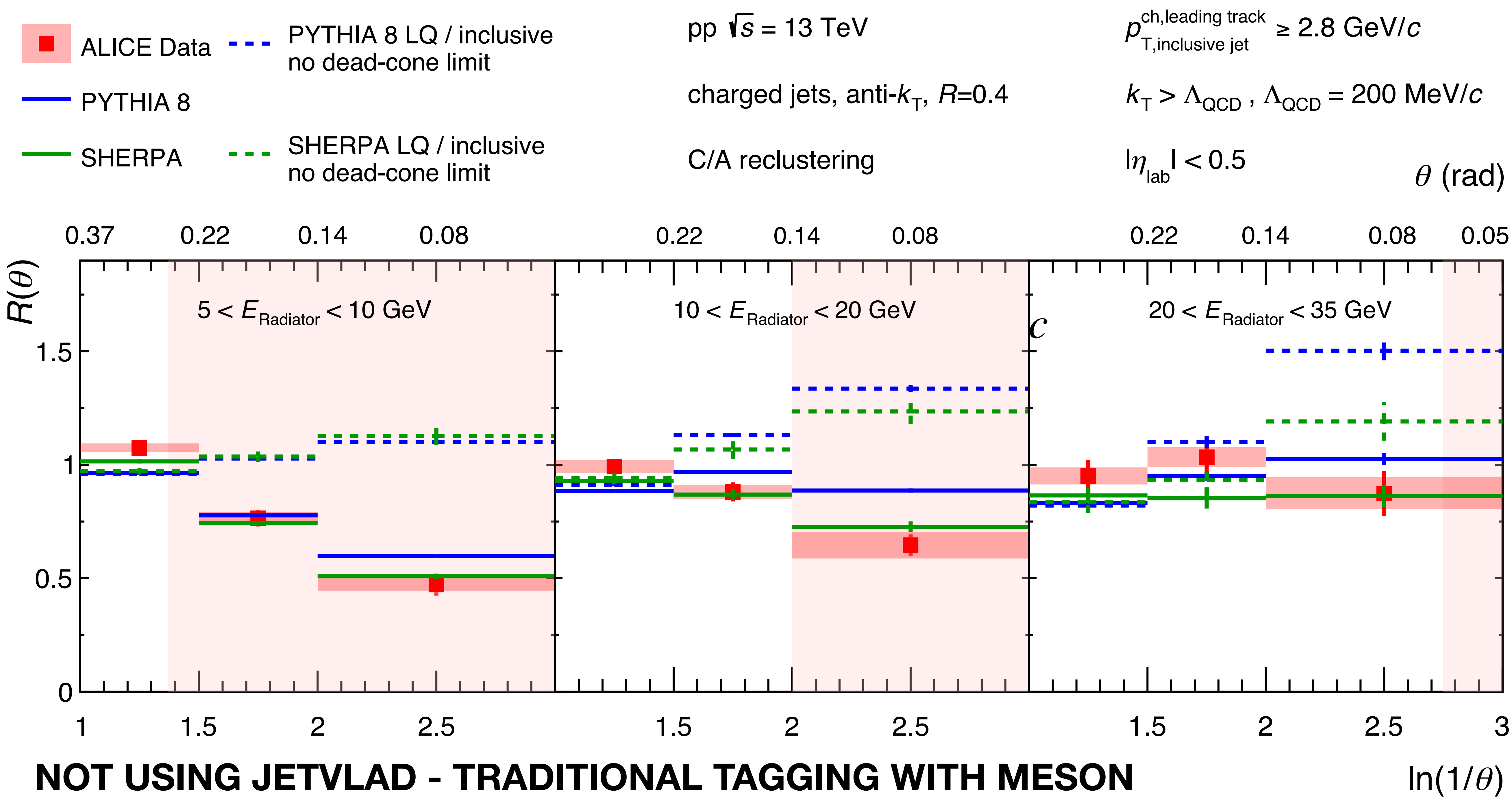
How we select HF jets now!



- Meson tagging!
- Enables a clean Monte-Carlo bias free selection of jets which include heavy-quark content
- Are these all HF jets?
- Enables a study of heavy quark radiation patterns in the QGP for an impactful measurement with early sPHENIX data

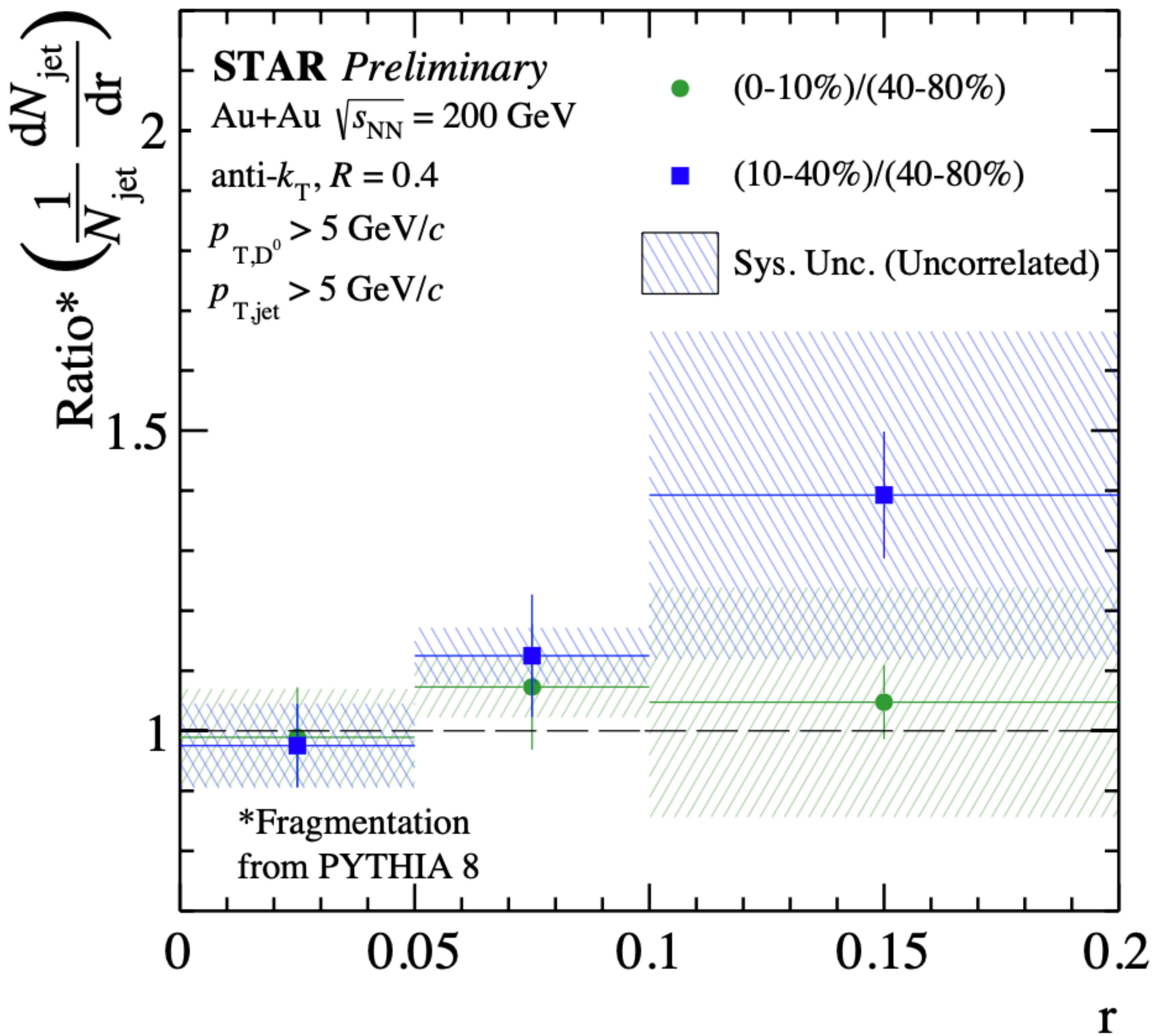
Realistic tagging - Heavy flavor mesons in jet

ALICE 2106.05713

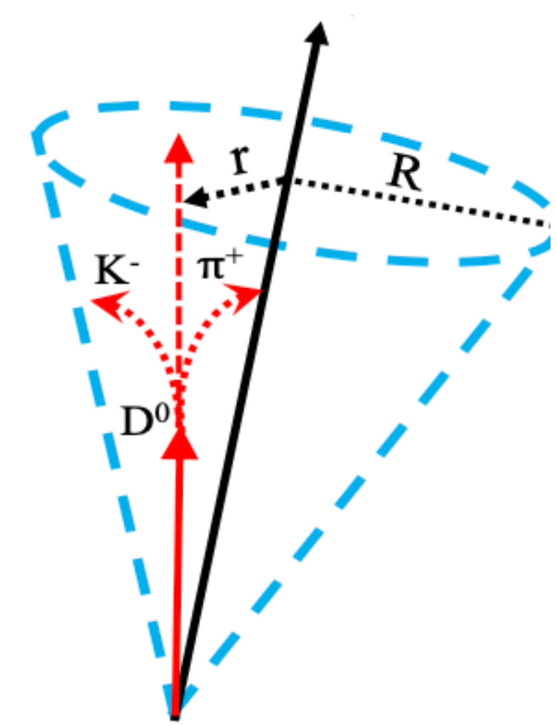


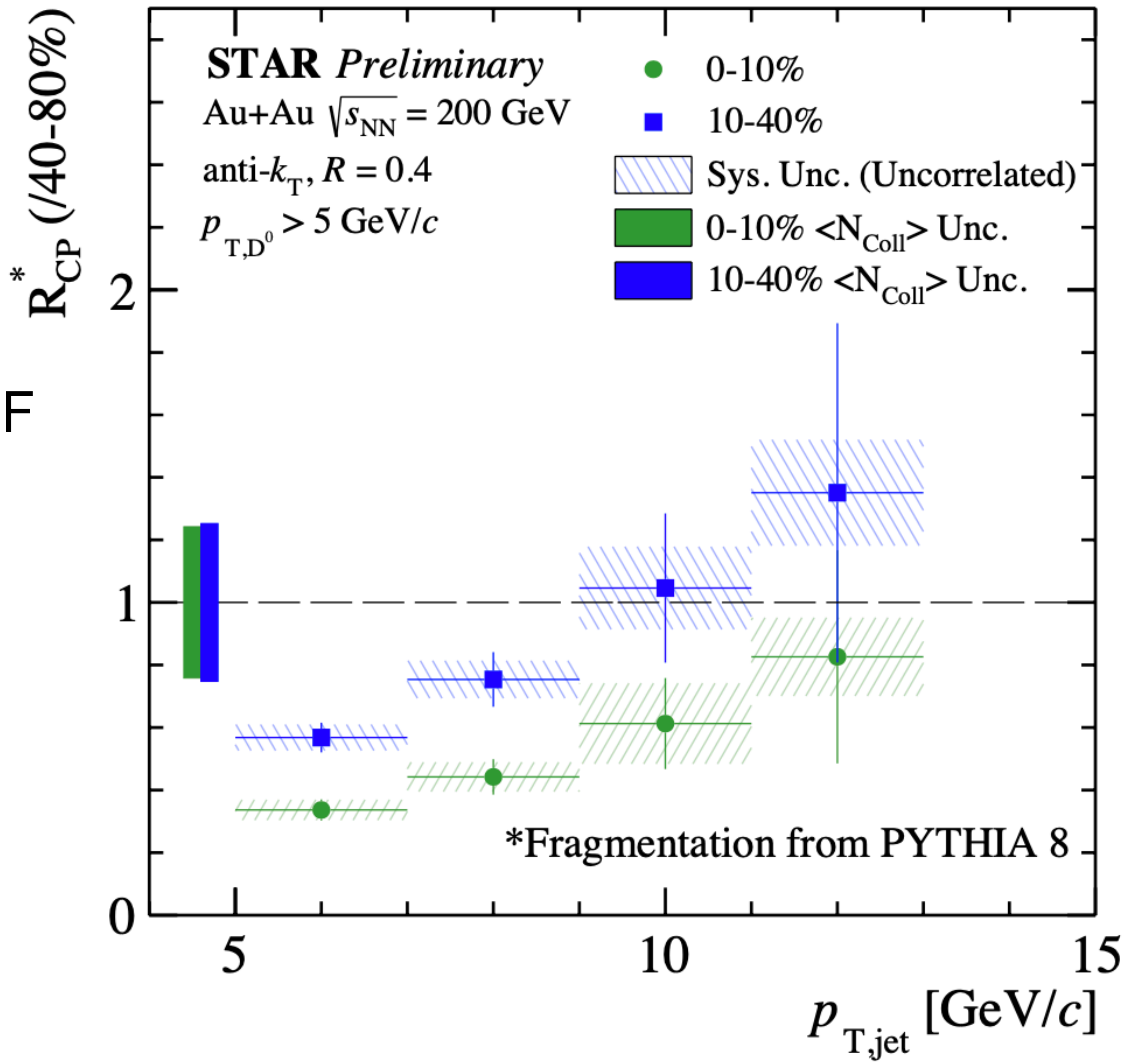
NOT USING JETVLAD - TRADITIONAL TAGGING WITH MESON

Where are the D0s produced within the jet?

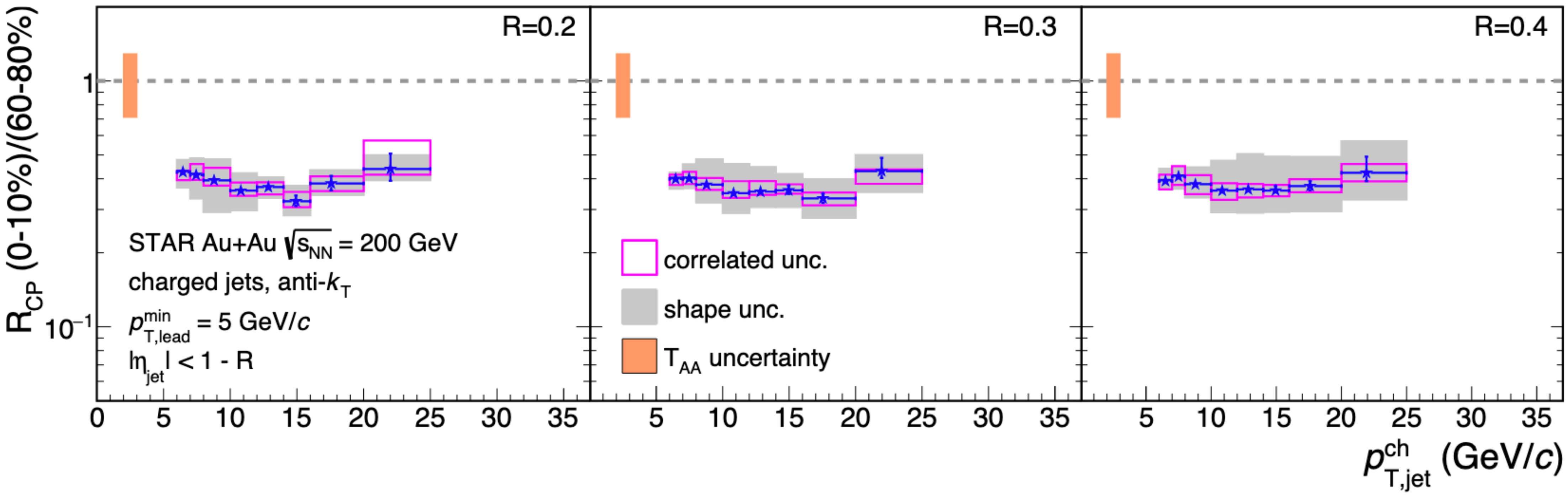


@ RHIC



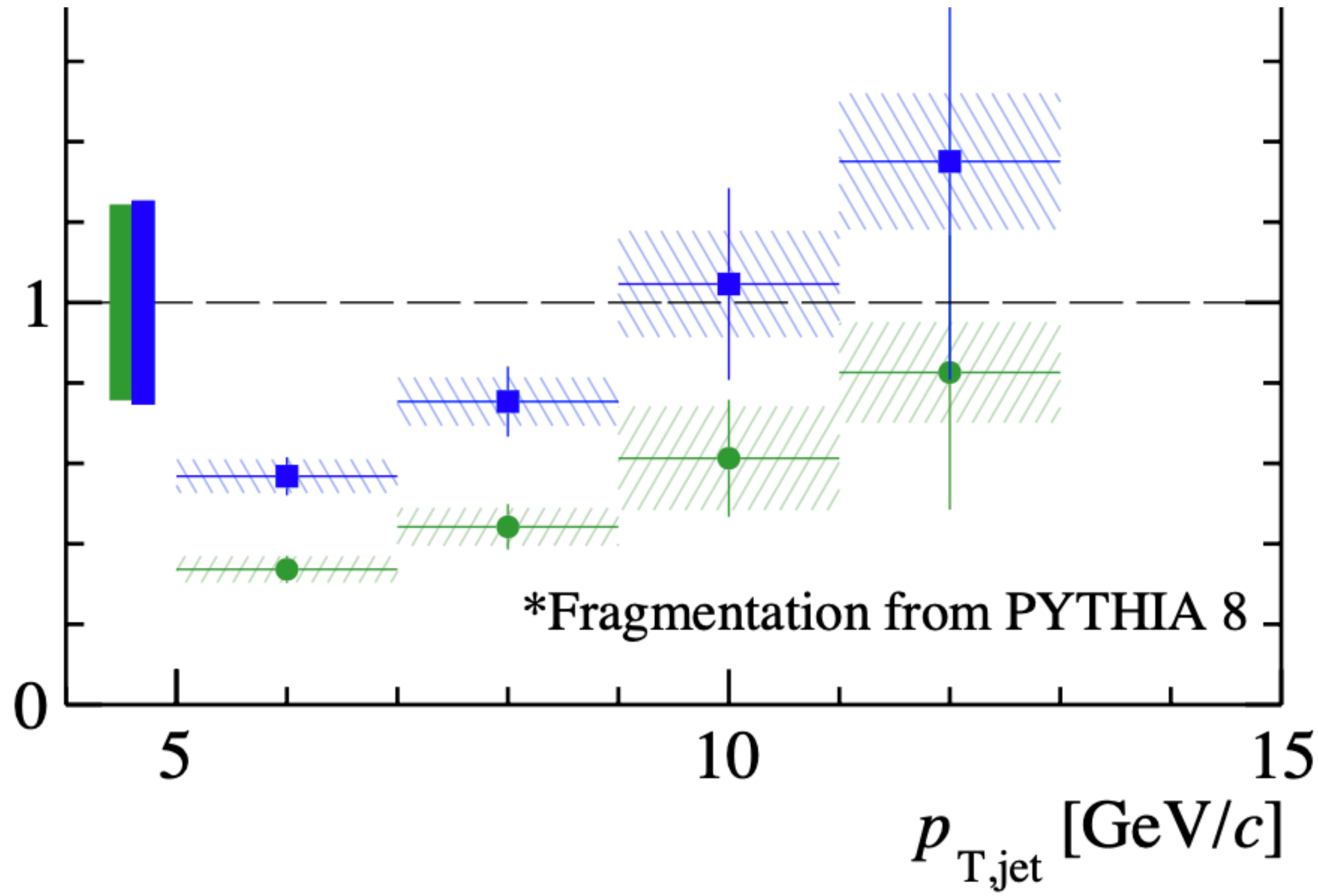


- Suppression of HF jets at RHIC
- Potential slope in RCP
- How different are they from light flavor jets?

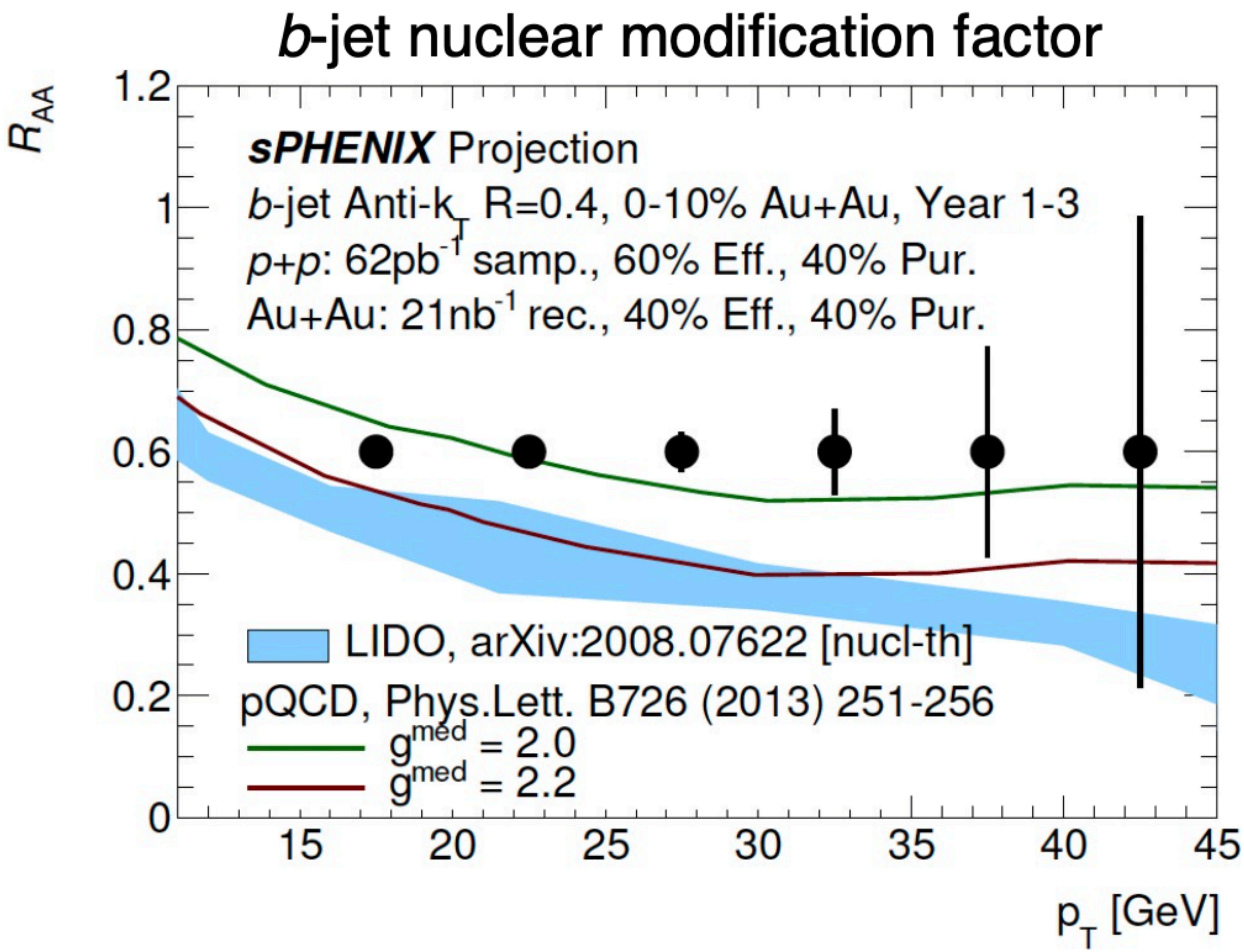


- Lower momentum are similar between the two

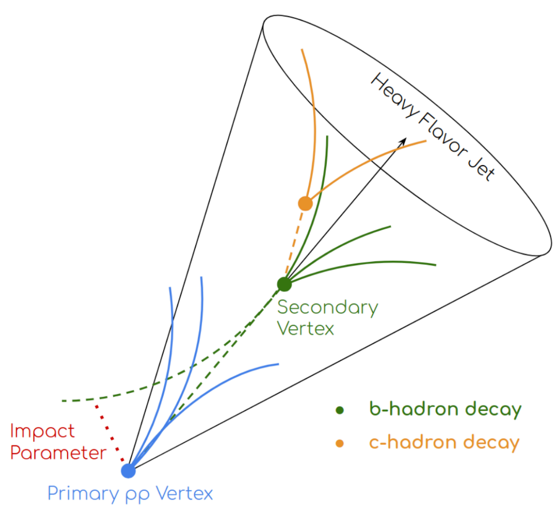
- Tantalizing hints in the data - but nothing conclusive since the uncertainties are large



Looking forward to sPHENIX



Machine Learning!

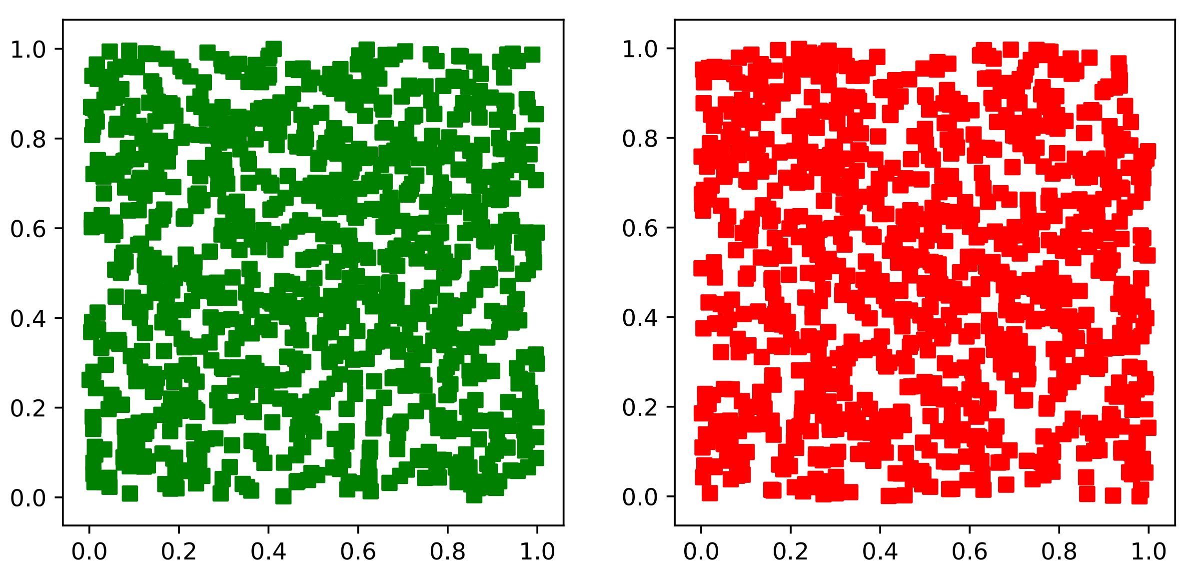


- PYTHIA 8.235 dijet sample for light (u, d, s, g) and heavy (c, b)
- $\hat{p}_T \in [8,17], [13-22], [18-27], [23-42]$
- Particle decays are limited to 2000 mm in x, y and 600 mm in z
- Dataset split into $80:10:10$ for training, testing and validation

Jets are tagged based on the initiator parton

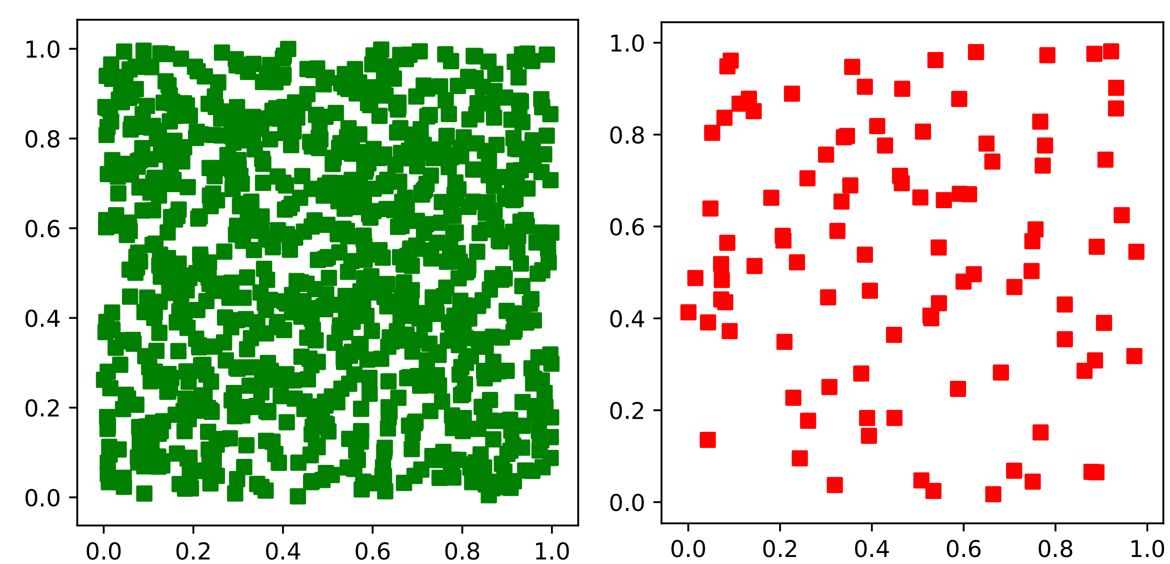
Balanced Sample (2M events)

Light Jets　　HF Jets



Hard QCD Sample (4M events)

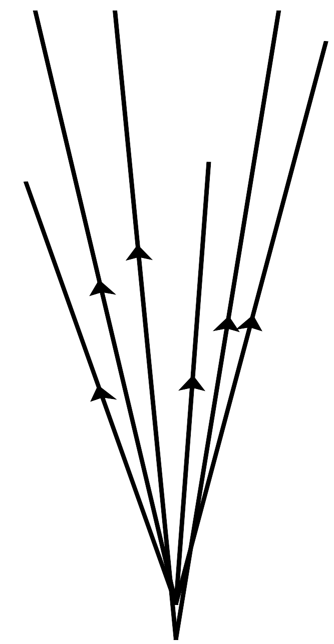
Light Jets　　HF Jets



Tagging Heavy-Flavor Jets

Machine Learning

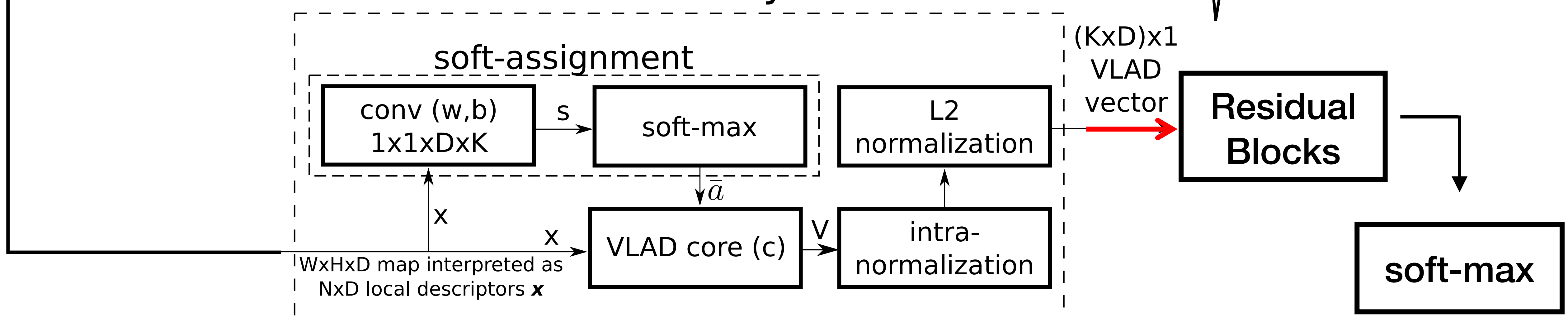
JetVLAD @ RHIC



$$\mathcal{J} = \{(p_{T,i}, \eta_i, \phi_i, \dots)\}_{i=1}^{n} ,$$

$$V(:,k) = \sum_{i=1}^{N} \frac{e^{w_k^T x_i + b_k}}{\sum_{k}' e^{w_{k'}^T x_i + b_{k'}}} (x_i - c_k)$$

NetVLAD layer



Vector of Locally Aggregated Descriptors

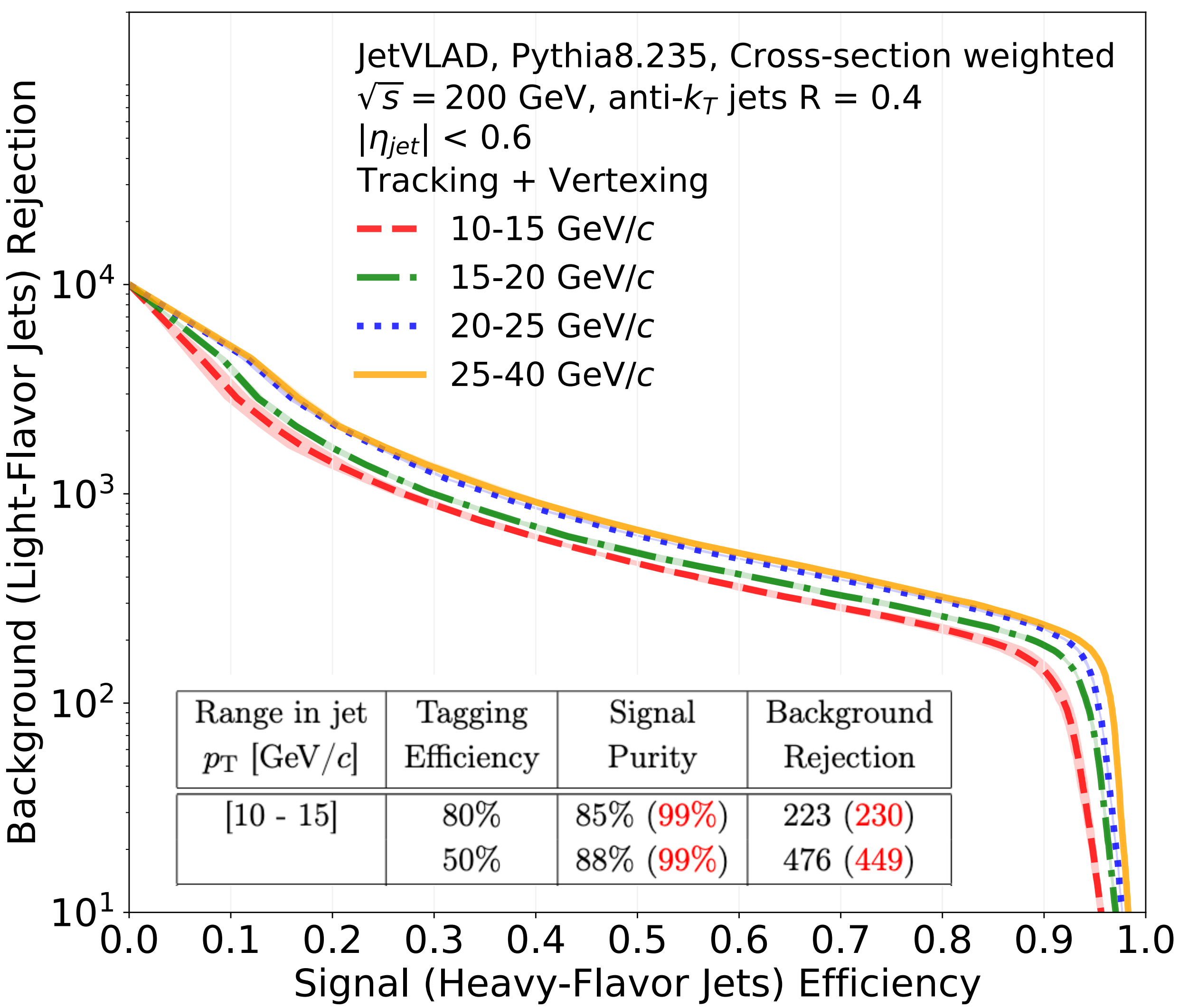
Ponimatkin, et. al
JINST (2021) 2005.01842

Total of 111608 trainable parameters

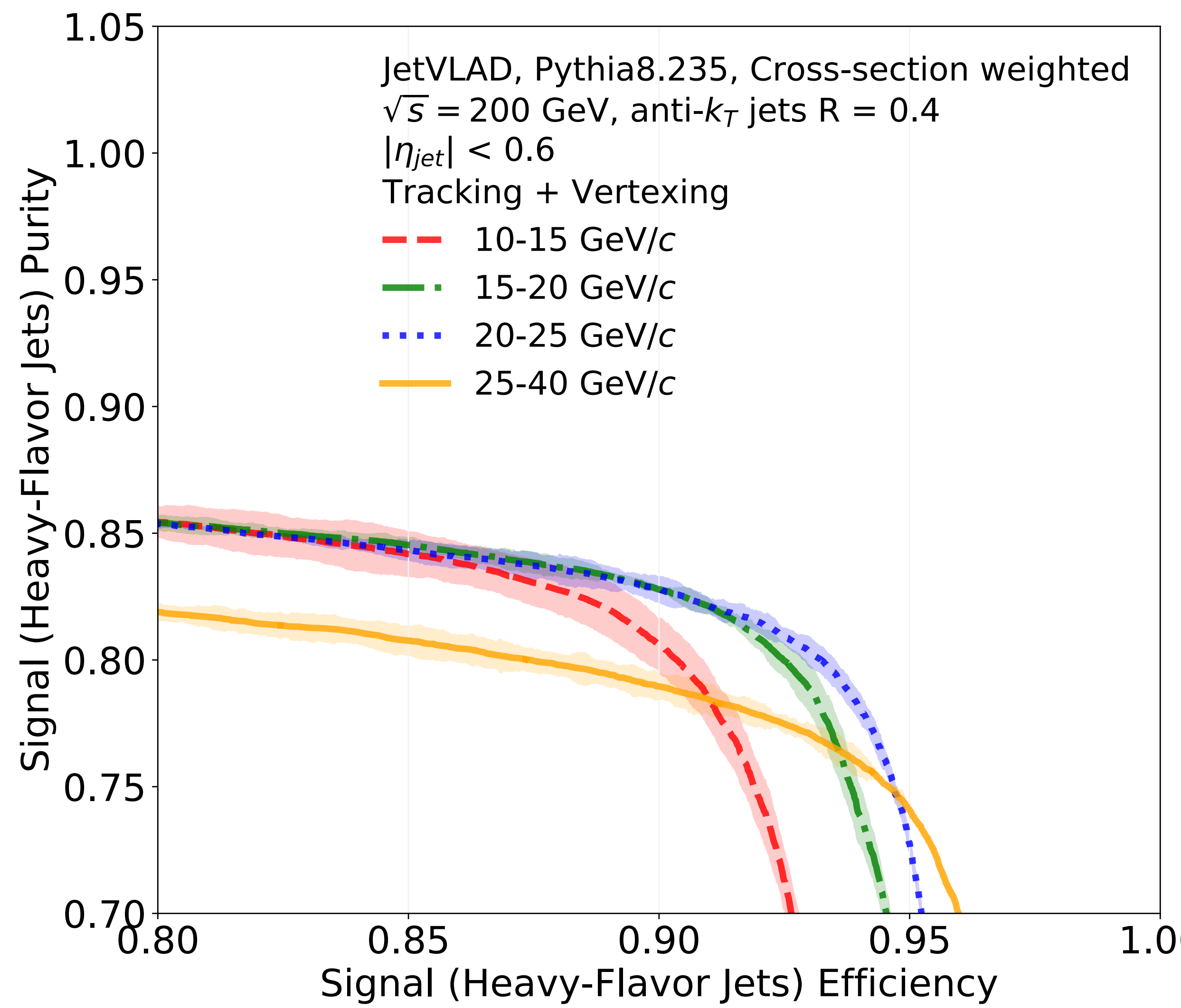
D - Depth
K - # Clusters

JetVLAD @ RHIC

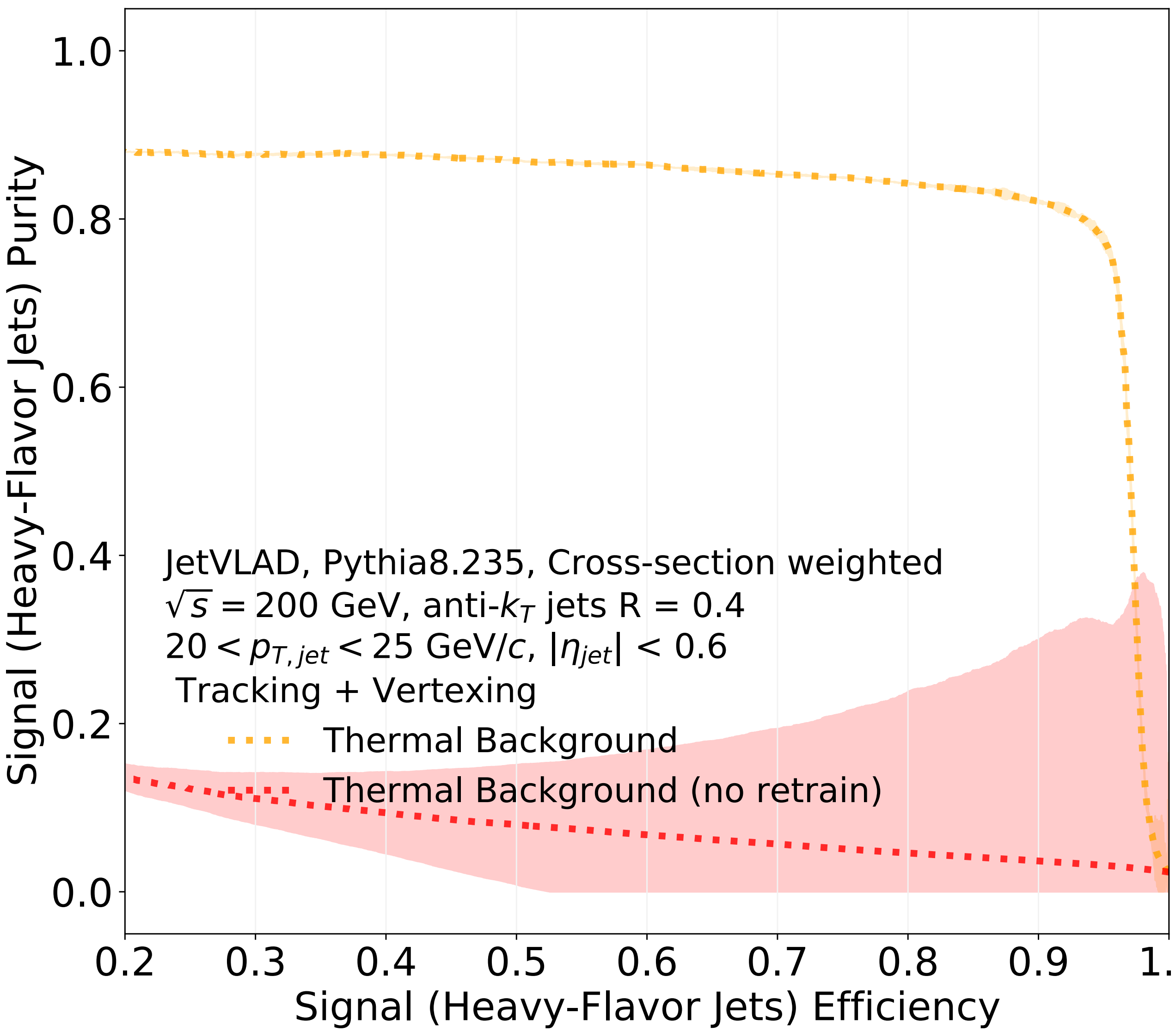
- With increasing jet momenta, at fixed efficiency (80%), we increase background rejection, but purity reduces — points to interesting kinematic effects - fragmentation differences for higher pT HF jets

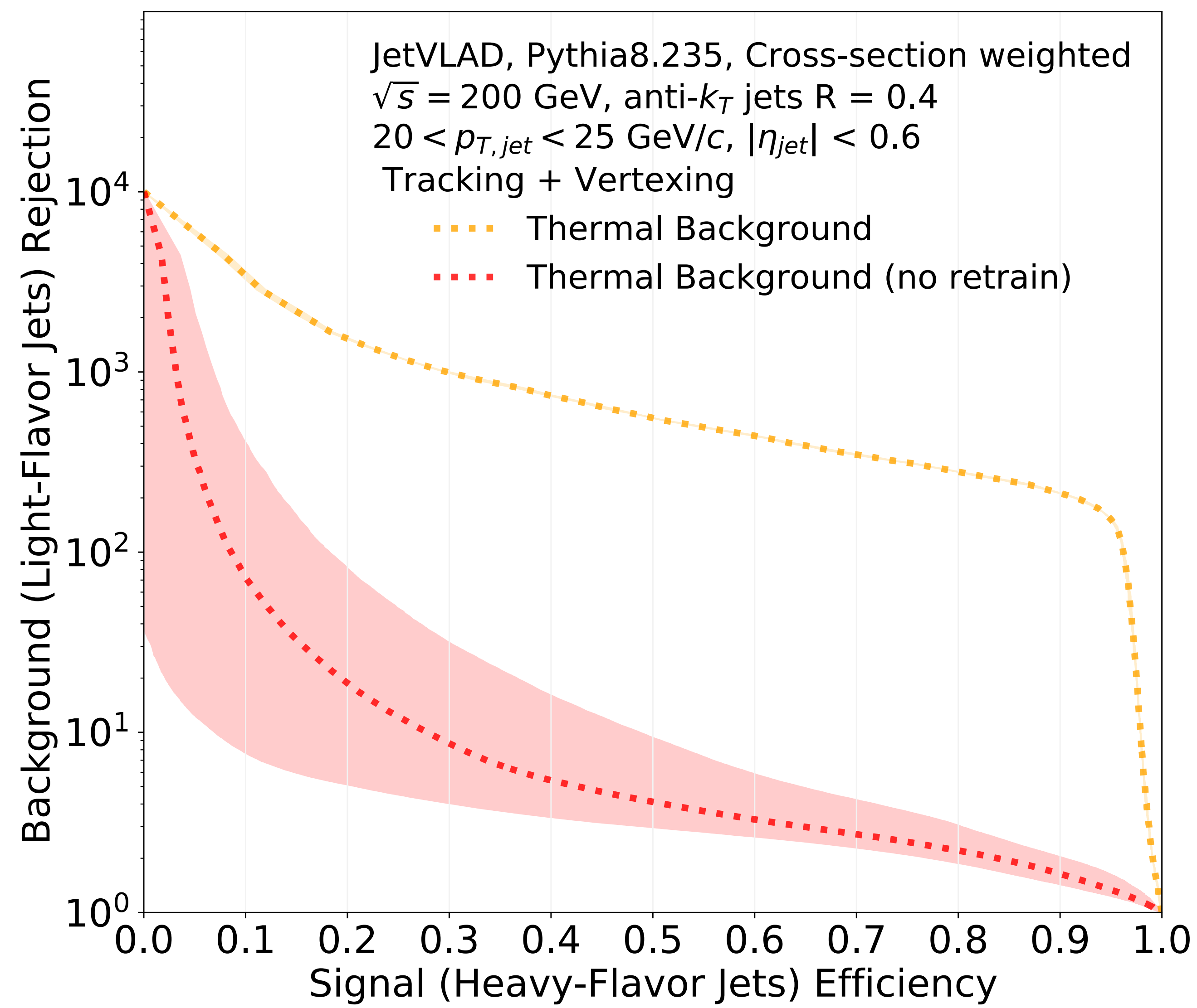


Range in jet p_T [GeV/c]	Tagging Efficiency	Signal Purity	Background Rejection
[10 - 15]	80%	85% (99%)	223 (230)
	50%	88% (99%)	476 (449)



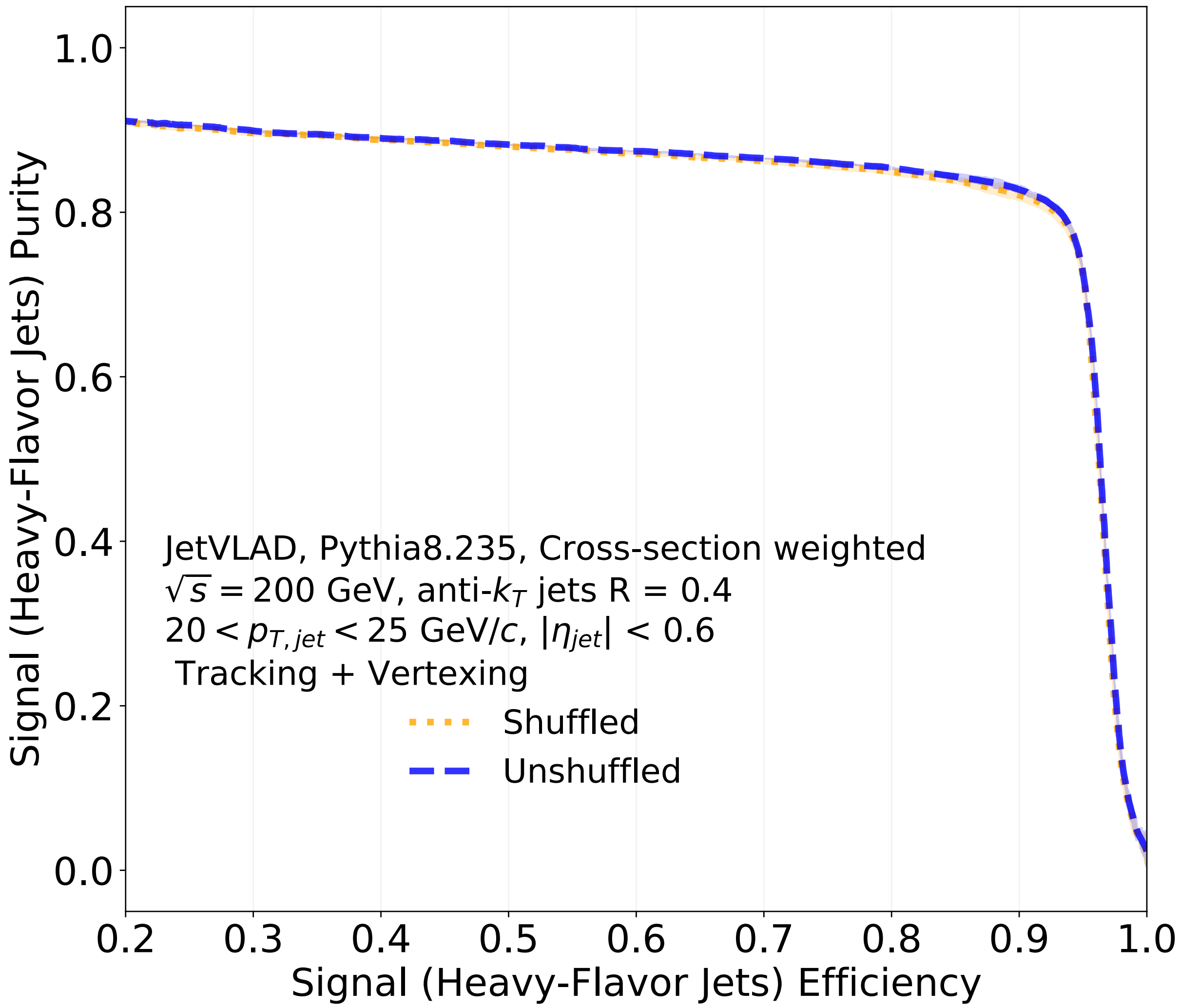
What is the impact of the heavy ion background?

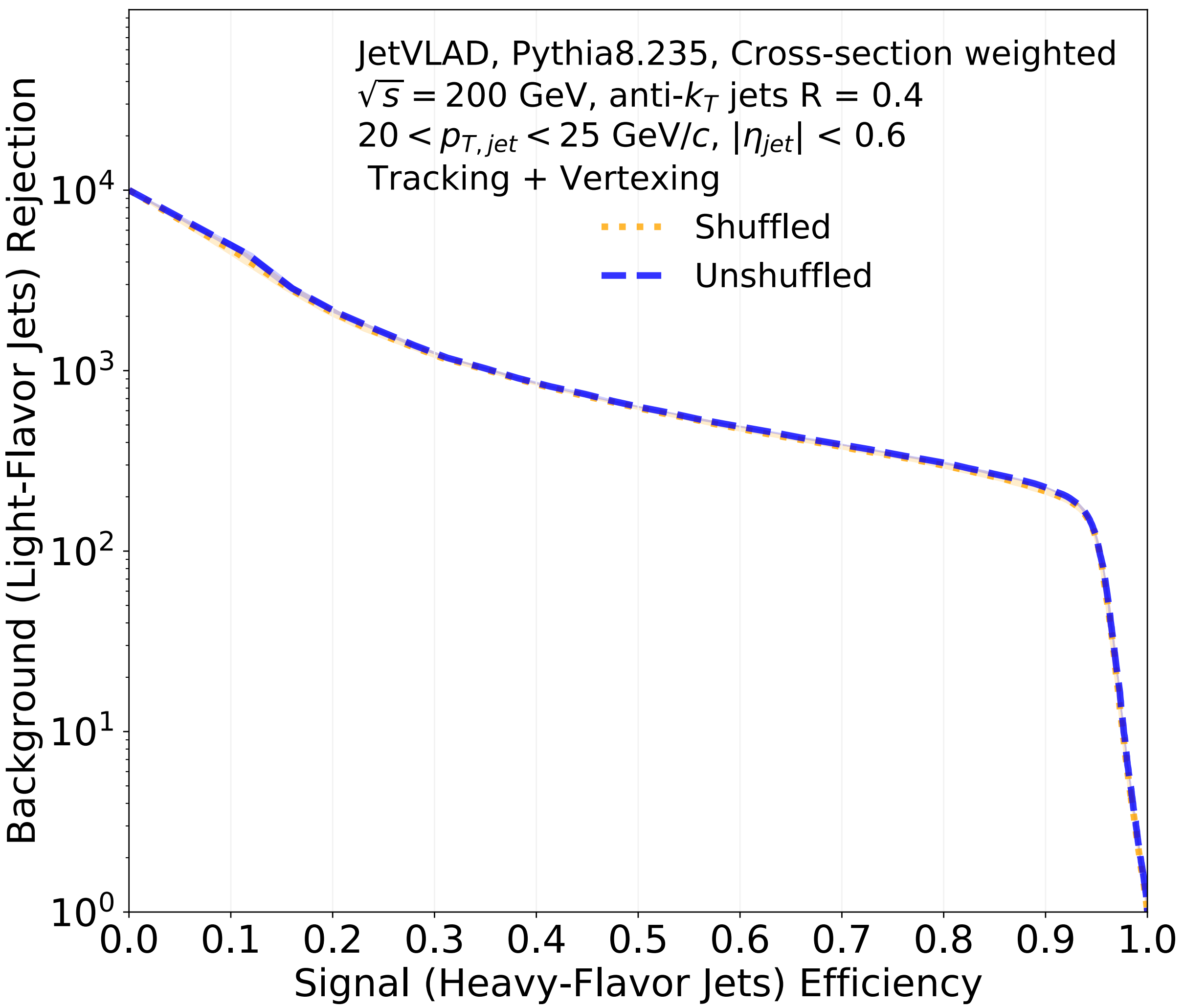




- Vacuum pre-trained model is completely swamped by the background! BUT retraining fixes the issue!

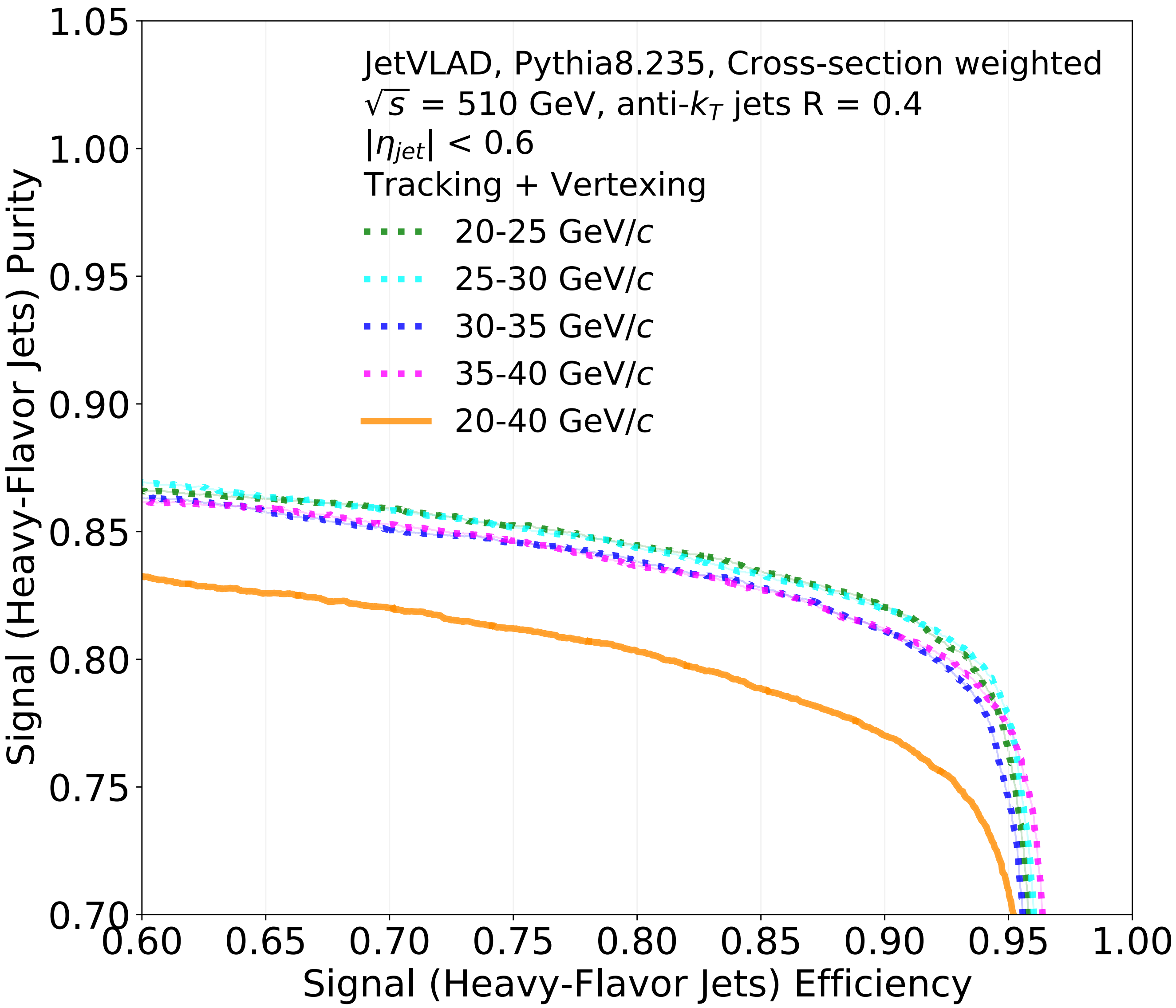
Is performance dependent on an ordering of inputs?

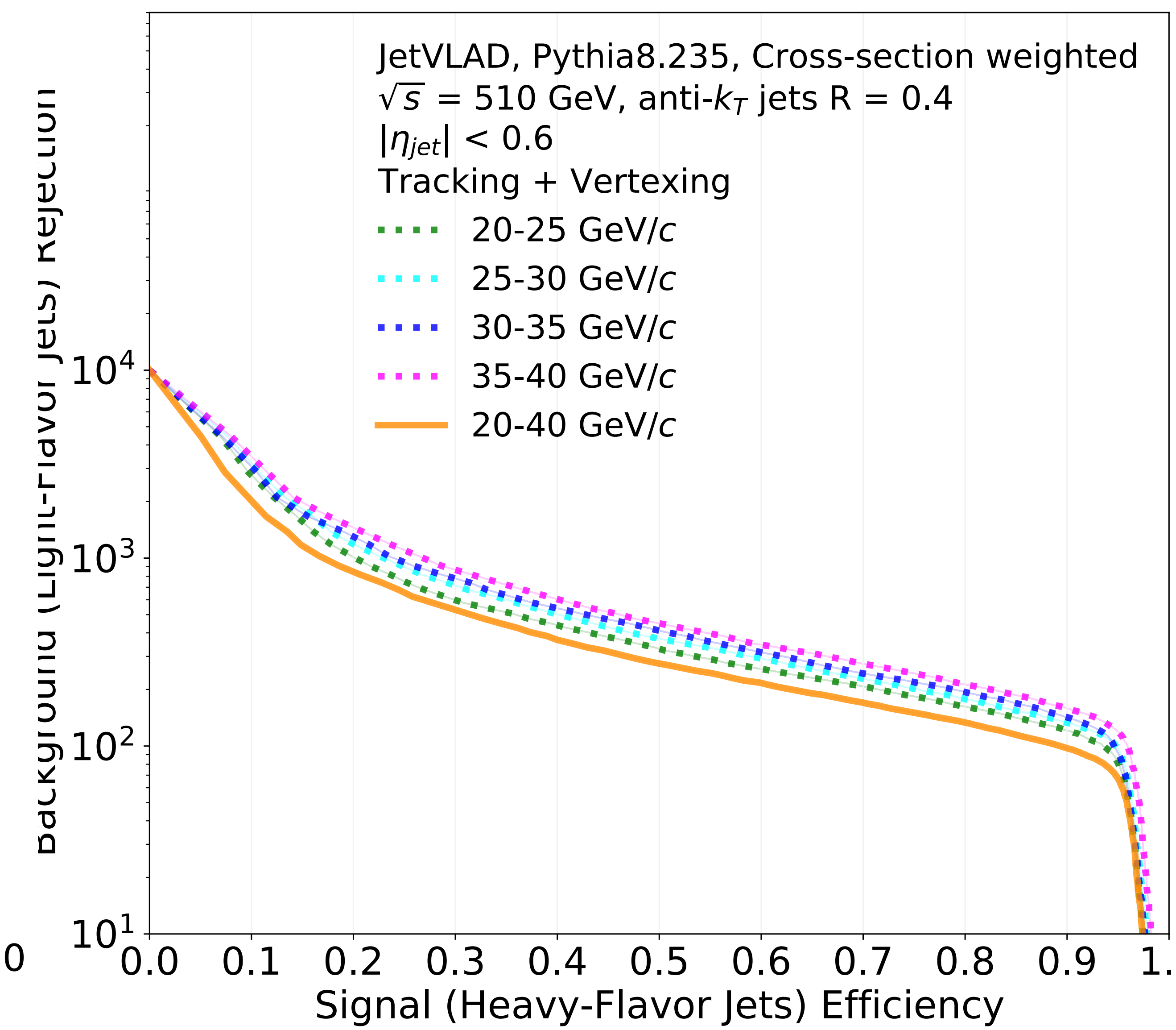




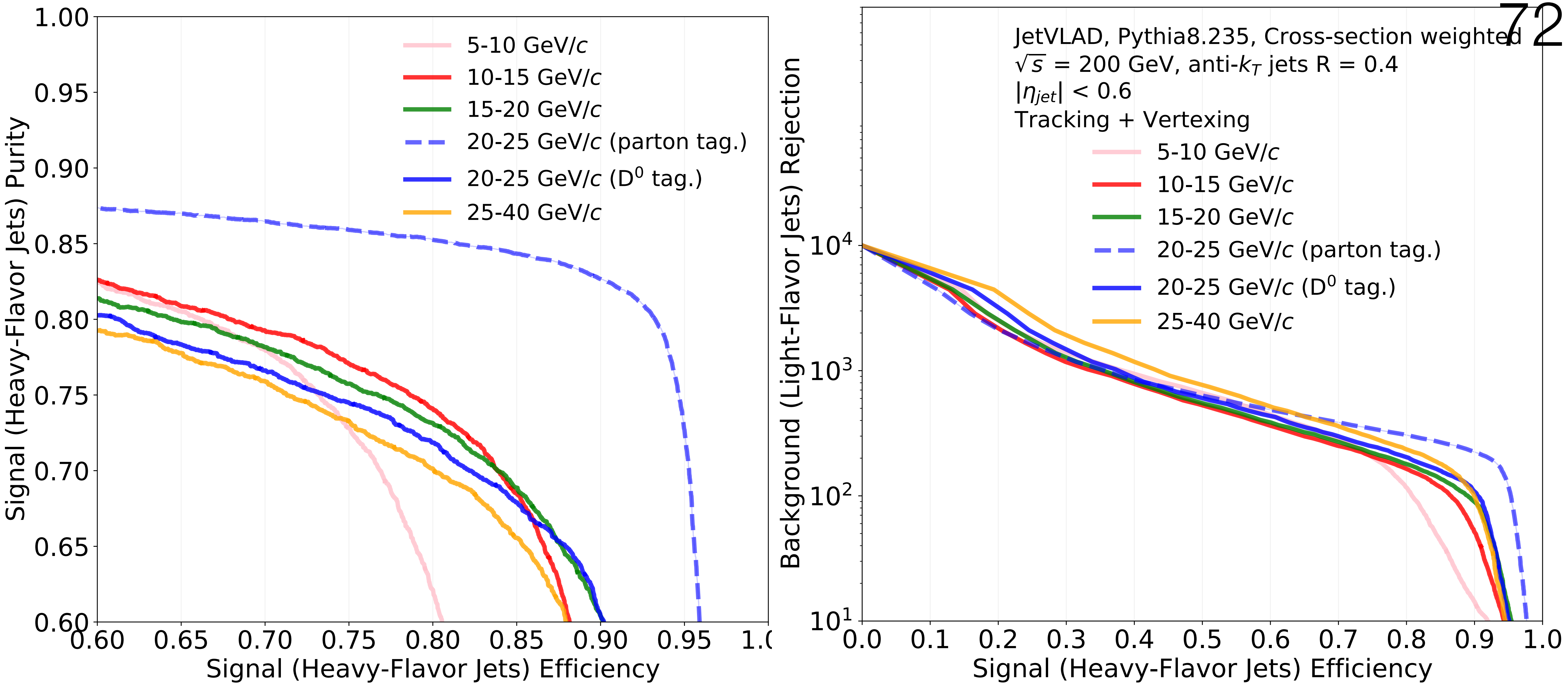
- Take the jet constituents and shuffle their order within the jet (keeping the 4-mom fixed) - no effect at all!

Any dependence on the jet momenta binning?





- The narrower the jet momentum range the better - larger bins result in varying admixtures of signal and background leading to greater overlap in the latent space



- Inputs are still daughter particles, except signal jets are tagged based on the fact that there is a D0 in them
- Overall we observe a reduction of ~15% purity at fixed efficiency (80%) with background rejection unaffected!

Conclusions

- Hard Probes - Jets and Heavy Flavor
- Produced right at the moment of collision and traverse, observe and interact with the plasma
- Useful for both extending fundamental QCD into the non-perturbative regime and transport properties of the plasma

LUX ET VERITAS

Wright Laboratory

Brookhaven National Laboratory

VANDERBILT UNIVERSITY

raghav.ke@vanderbilt.edu