Monte Carlo generators as navigation tools for Open**MAPP**

OpenMAPP

Andrzej Siódmok

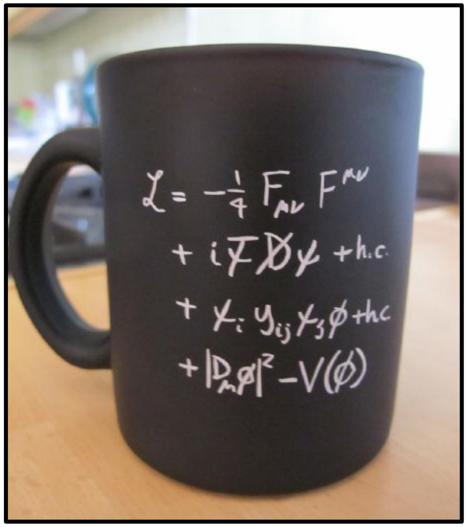
JAGIELLONIAN UNIVERSITY IN KRAKÓW
UCZELNIA BADAWCZA
INICJATYWA DOSKONAŁOŚCI
NARODOWE CENTRUM NAUKI
NCN: 2019/34/E/ST2/00457
H7
MCnet

Standard Model

There is a **huge gap** between a one-line formula of a fundamental theory, like the Lagrangian of the SM, and the experimental reality that it implies

Theory

Standard Model Lagrangian



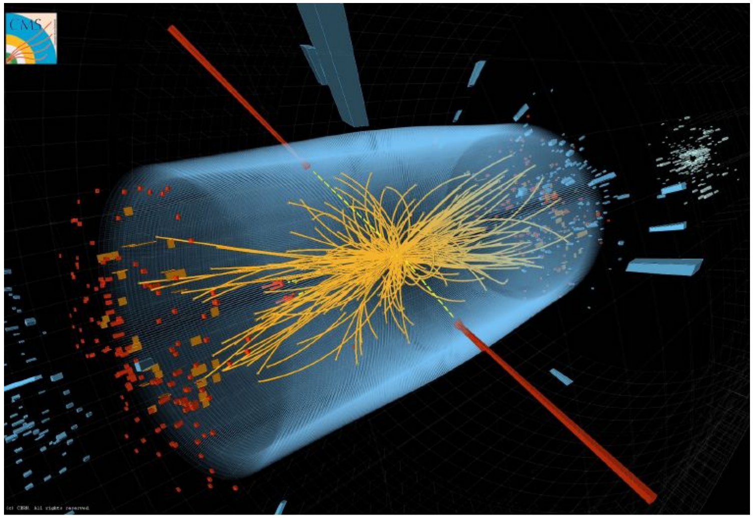
Data makes you smarter

> It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong.

Richard P. Feynman

Experiment

LHC event



Standard Model

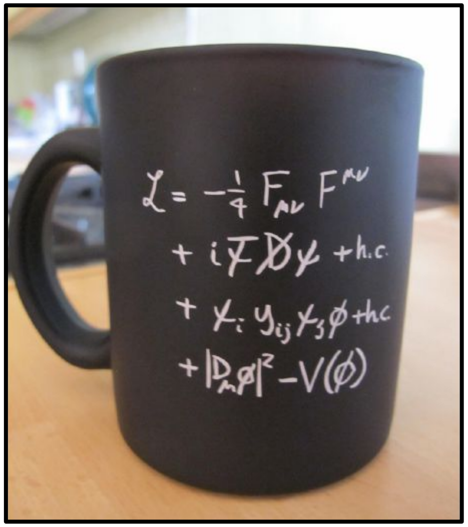
There is a **huge gap** between a one-line formula of a fundamental theory, like the Lagrangian of the SM, and the experimental reality that it implies

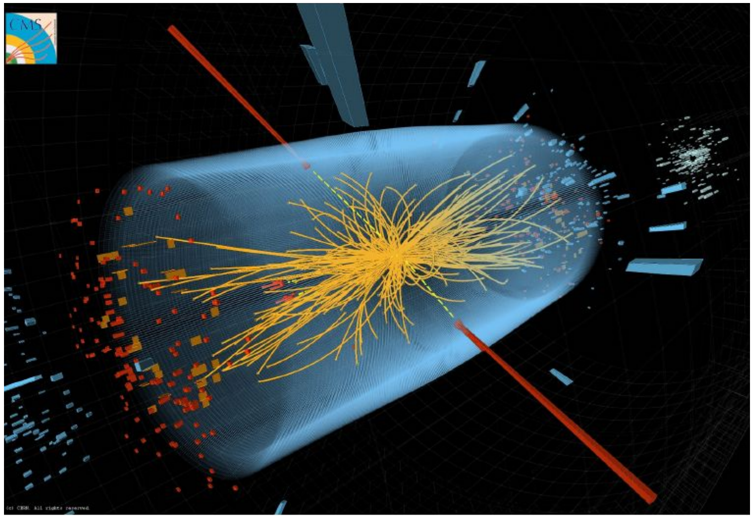
Theory

Standard Model Lagrangian

Experiment

LHC event



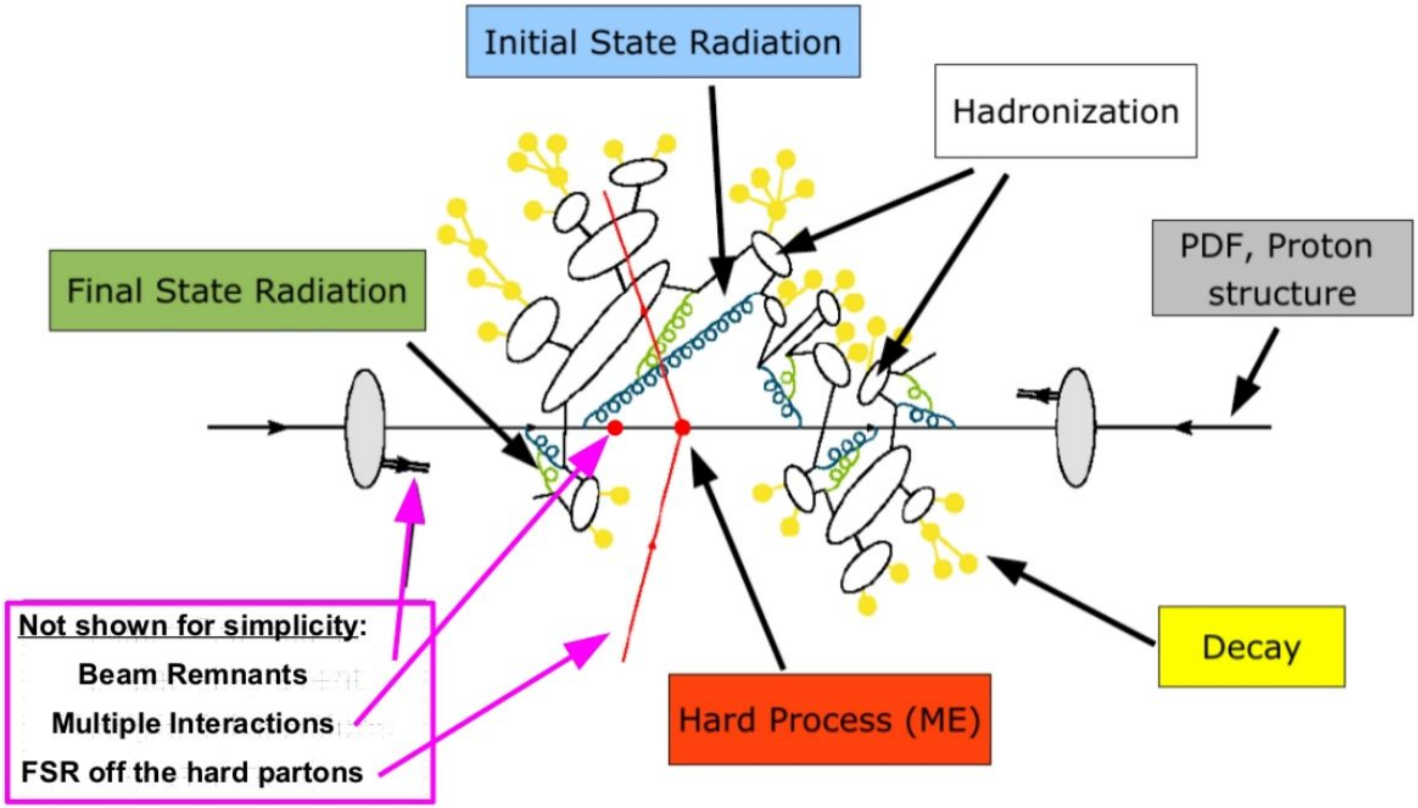


- MC event generators are designed to bridge the that **gap**
- “Virtual collider” ⇒ Direct comparison with data

⇩

Almost all **HEP measurements and discoveries** in the modern era have **relied on MCEG**, most notably the discovery of the Higgs boson.

Published papers by ATLAS, CMS, LHCb: **2252**
Citing at least 1 of 3 existing MCEG: **1888** (**84%**)



taken from Stefan Gieseke©

The general approach is the same in different programs but the models and approximations used are different.

Monte Carlo Generators

Current release series	Hard matrix elements	Shower algorithms	NLO Matching	Multijet merging	MPI	Hadronization	Shower variations
Herwig 7	Internal, libraries, event files	QTilde, Dipoles	Internally automated	Internally automated	Eikonal	Clusters, (Strings)	Yes
Pythia 8	Internal, event files	Pt ordered, DIRE, VINCIA	External	Internal, ME via event files	Interleaved	Strings	Yes
Sherpa 2	Internal, libraries	CSShower, DIRE	Internally automated	Internally automated	Eikonal	Clusters, Strings	Yes

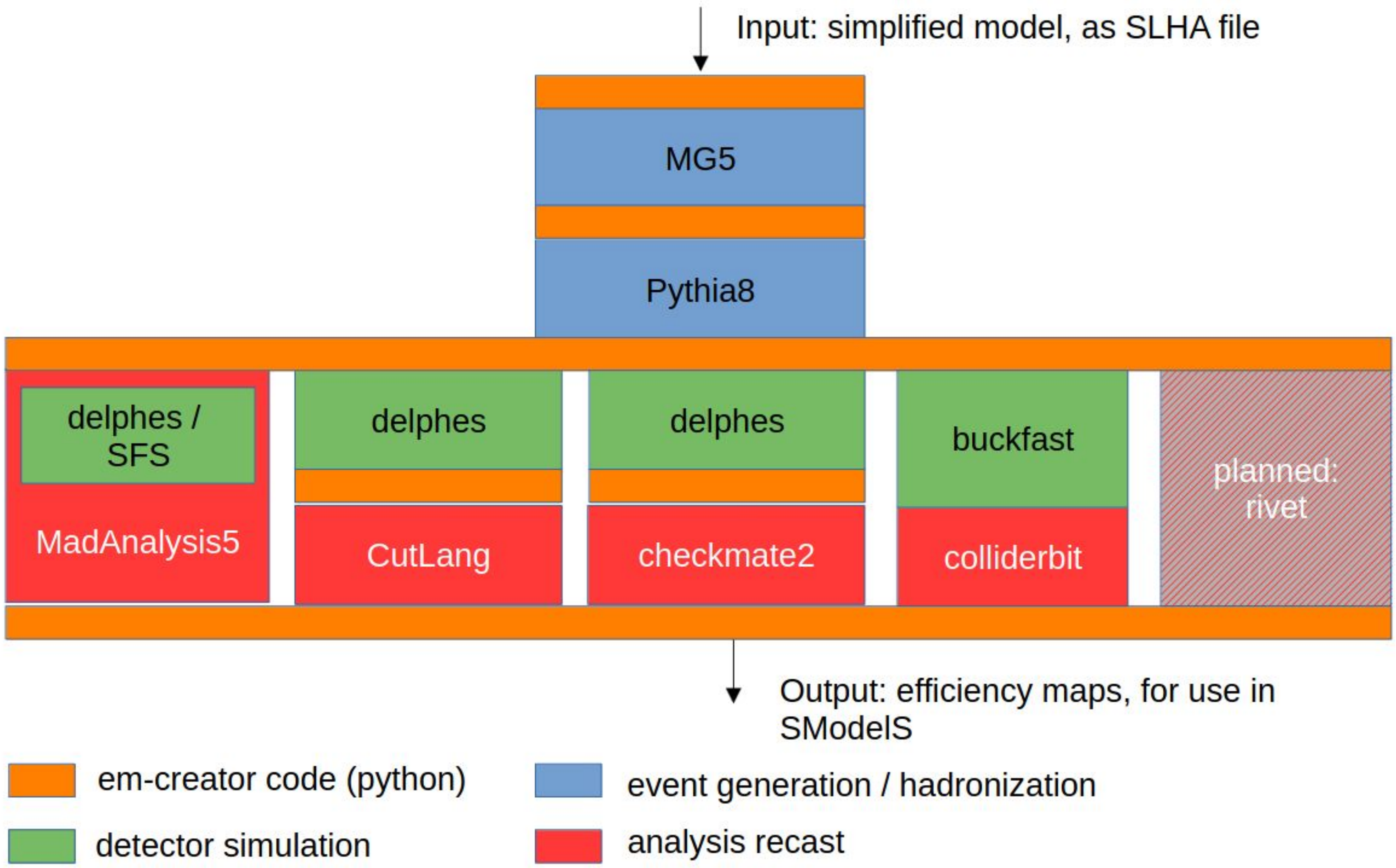
[from S. Platzer]

“But it often happens that the physics simulations provided by the MC generators carry the authority of data itself. They look like data and feel like data, and if one is not careful they are accepted as if they were data.”

J. D. Bjorken

Therefore, any idea that we would like to use on experimental data can be tested on data from generators - a navigational tool.

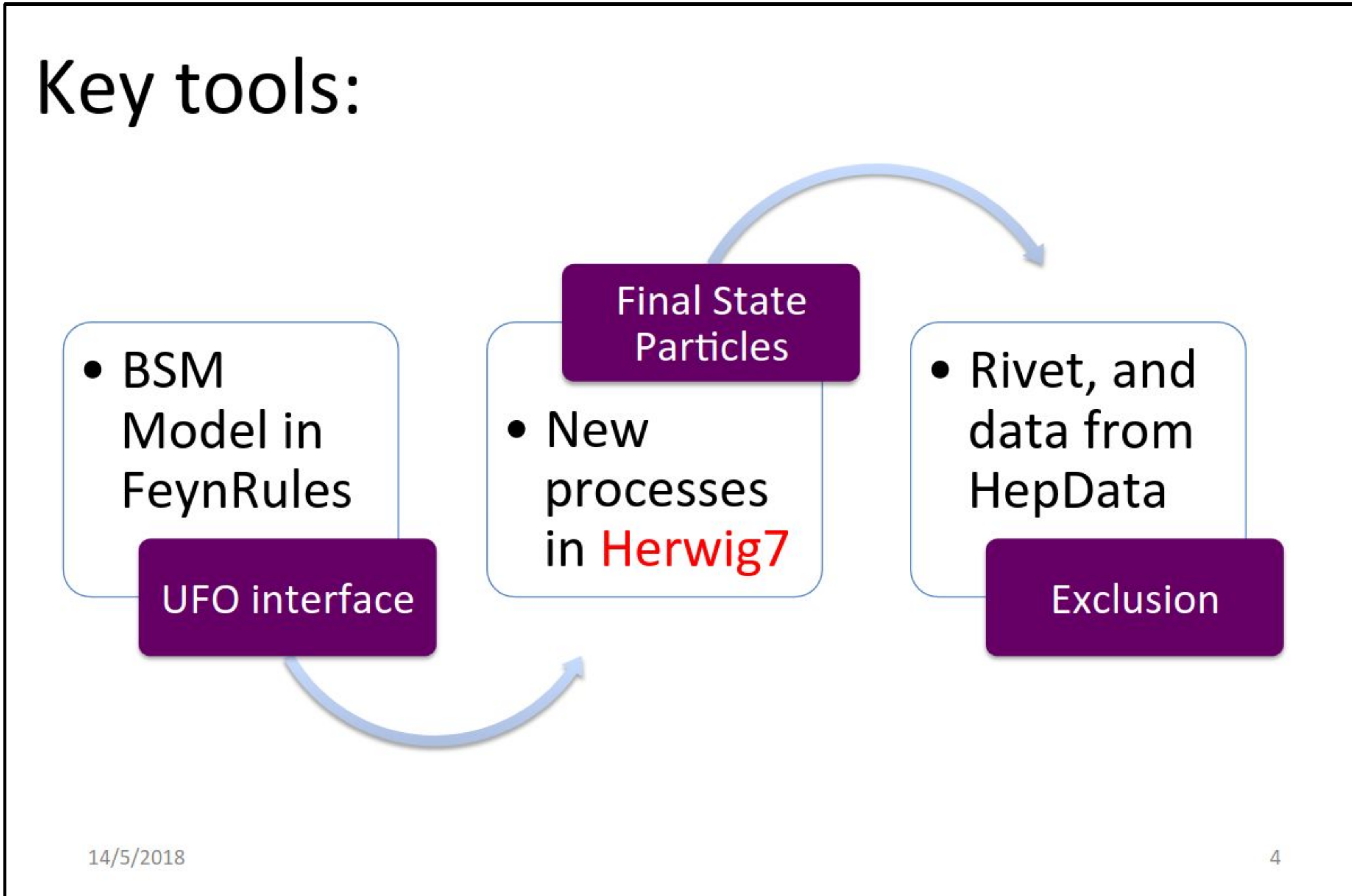
Tool Chains



Wolfgang Waltenberge

Monte Carlo Generators an a navigation tool

Contur Update talk at Jonathan Butterworth at
(Re)interpreting the results of new physics searches at the LHC, CERN, 2018



Monte Carlo Generators an a navigation tool

- For the purposes of this talk, will focus on Gbb model
  - 4 NN SRs (based on parameterized mass points), 3 C&C SRs
  - Gtt performance very similar; Gtb (C&C only!) ~slightly less so.
  - Unless stated otherwise, all MC is Pythia 8.310
    (for easy comparison to Gambit)

T. Procter

- Just one cutflow provided
  CM efficiency for BDT seems to be similar,
  but I suspect large sensitivity to MC details

K. Rolbiecki

(Issues with the pythia code produced by MadGraph, etc etc).

M. Goodsell

- First ADL/CL-related developments targeted within OpenMAPP: HEPData interface, a more generic and robust interface to MC tools.

S. Sekmen

Problems with the description of MC generators settings - we see that in many publication.
Les Houches accord for the generators input (hard to imagine)?
ADL language for the Event Generators?
"HEPDATA/Rivet" for storing of the MC input files or data samples, see MCplots for example?

Underlying Event

Away ▴

<pT> vs Nch

<pT> vs p_T^{lead}

<Nch> vs p_T^{lead}

TransMAX ▾

TransMIN ▾

Toward ▾

Δφ distributions ▾

TransAVE ▾

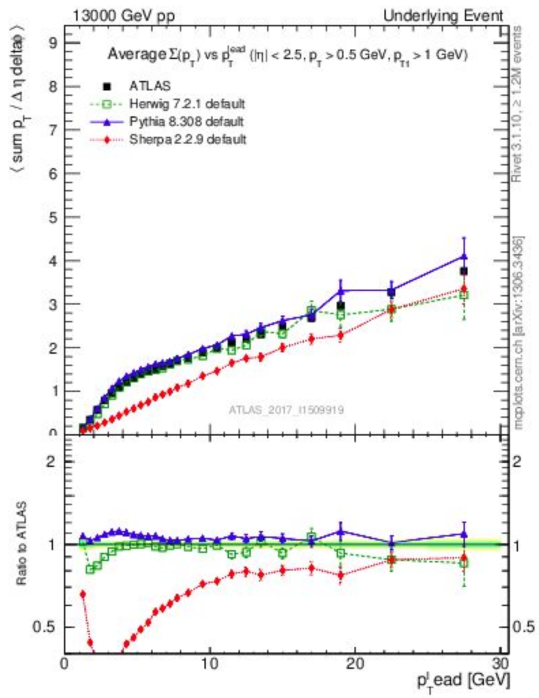
TransDIF ▾

p_T^{lead}

Underlying Event : Away : Σ(pT) vs p_T^{lead}

General-Purpose MCs : Main ▾ | Customize

pp @ 13000 GeV



details

Download as: .pdf .eps .png .script.tgz #

ATLAS experiment: data | article paper

Herwig 7 (Def): data |generator card

Pythia 8 (Def): data |generator card

Sherpa (Def): data |generator card

```
# settings of Pythia 8 wrapper program
Main:numberOfEvents = 100000          ! number of events to generate
Next:numberShowEvent = 0           ! suppress full listing of first events

# random seed
Random:setSeed = on
Random:seed = 7

# Beam parameter settings.
Beams:idA = 2212               ! first beam, p = 2212, pbar = -2212
Beams:idB = 2212               ! second beam, p = 2212, pbar = -2212
Beams:eCM = 13000              ! CM energy of collision


# Minimum Bias process (as taken from one of pythia8 example)
SoftQCD:nonDiffractive = on              ! minimum bias QCD processes
SoftQCD:singleDiffractive = on
SoftQCD:doubleDiffractive = on

# Process setup: min-bias
# Use this for ordinary min-bias (assuming Rivet analysis
# correctly suppresses the diffractive contributions.)
# SoftQCD:all = on # this for min-bias incl diffraction


# Set cuts
# Use this for hard leading-jets in a certain pT window
PhaseSpace:pTHatMin = 0   # min pT
PhaseSpace:pTHatMax = 13000   # max pT

# Use this for hard leading-jets in a certain mHat window
PhaseSpace:mHatMin = 0   # min mHat
PhaseSpace:mHatMax = 13000   # max mHat


# Makes particles with c*tau > 10 mm stable:
ParticleDecays:limitTau0 = On
ParticleDecays:tau0Max = 10.0


# Tune setup:
```

Novel approach to measure quark/gluon jets at the LHC

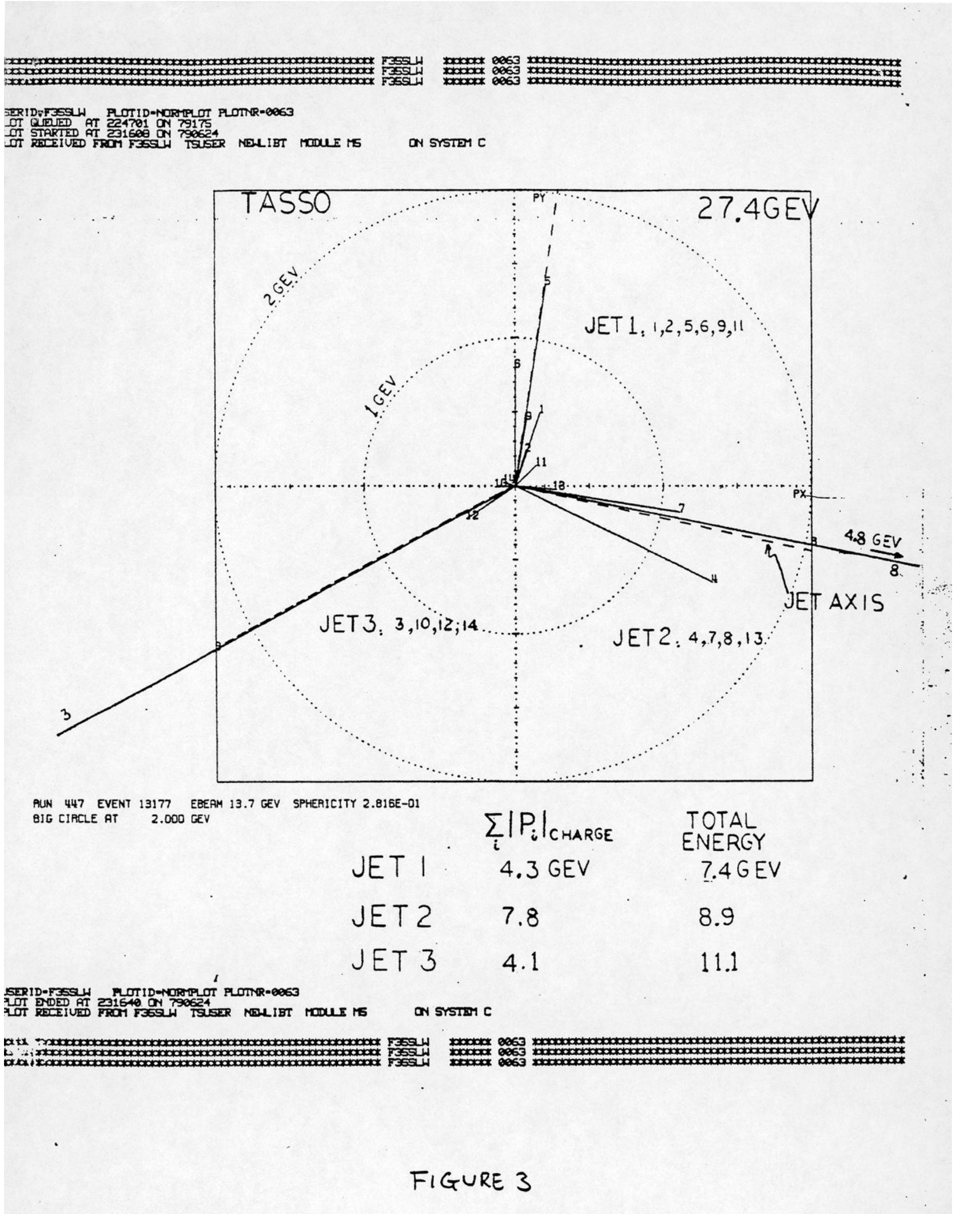
Andrzej Siódmok

in collaboration with

Petr Baron & Micheal Seymour

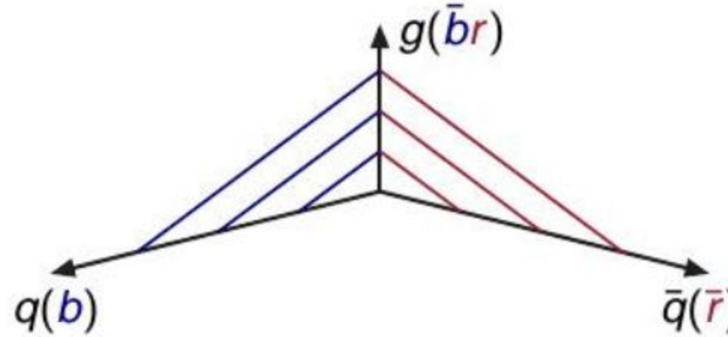
based on Eur.Phys.J.C 84 (2024) 1

History: Discovery of the gluon



This collision event recorded in **1979**, provided the first evidence of the gluon.

Recorded as event 13177 of run 447 of the TASSO experiment at the Deutsches Elektronen-Synchrotron (DESY), the graphic shows three jets of particles produced in an electron-positron collision.



Distinguish Q/G jets as is as old as gluon's discovery

Quark - Gluon Separation in Three Jet Events #1

Hans Peter Nilles (SLAC), K.H. Streng (SLAC) (Aug 1, 1980)

Published in: *Phys.Rev.D* 23 (1981) 1944

pdf links DOI cite 32 citations

A Monte Carlo Program for Quark and Gluon Jet Generation #2

Torbjorn Sjostrand (Lund U., Dept. Theor. Phys.) (Apr 1, 1980)

pdf cite 1 citation

Quark and gluon jet separation: Conventional and neural network methods #2

Z. Fodor (Eotvos U.) (Jul, 1991)

Published in: *Conf.Proc.C* 910725V1 (1991) 438 • Contribution to: Joint International Lepton Photon Symposium at High Energies (15th) and European Physical Society Conference on High-energy Physics, 438

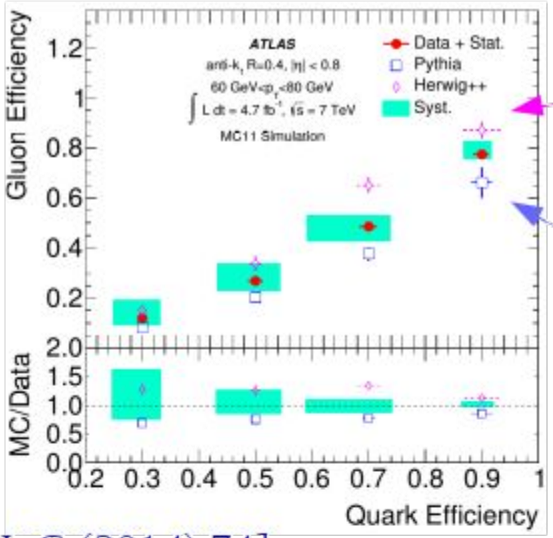
⋮

Quark versus Gluon Jet Tagging Using Charged Particle Multiplicity with the ATLAS Detector #7

ATLAS Collaboration (Apr 11, 2017)

LHC Q/G jet measurement

Efficiency is simply the ratio of the number of jets selected by a discriminant over the total number in the sample.



[ATLAS, Eur. Phys. J. C (2014) 74]

Herwig++ is too pessimistic, Quark and gluon jets looks more the same than in the data.

Pythia is too optimistic, Quark and Gluon jets are too different compared to data.

Conclusion:

"A detailed study of the jet properties reveals that quark-and gluon-jets look more similar to each other in the data than in the Pythia 6 simulation and less similar than in the Herwig++ simulation."

Problem: Q/G jets LHC data show discrepancy with the predictions from MC generators

Why we would like to distinguish Q/G jets?

BSM searches: often signature for a BSM signals: many quark, backgrounds: QCD gluons

- 8-jet Gluino event: $pp \to \tilde{g}\tilde{g}$ and each $\tilde{g}$ decays to 4 quarks:

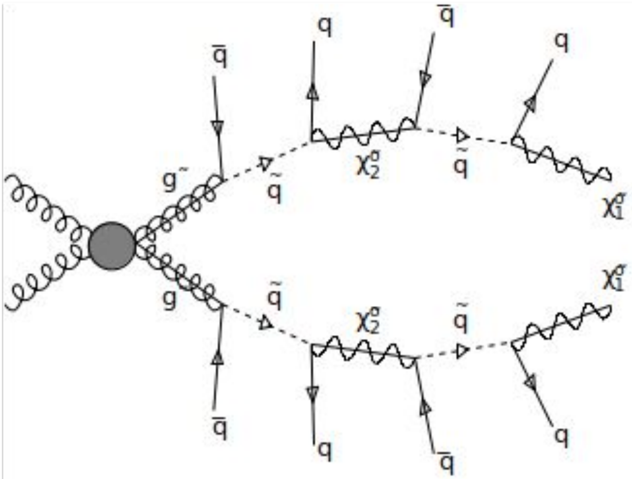
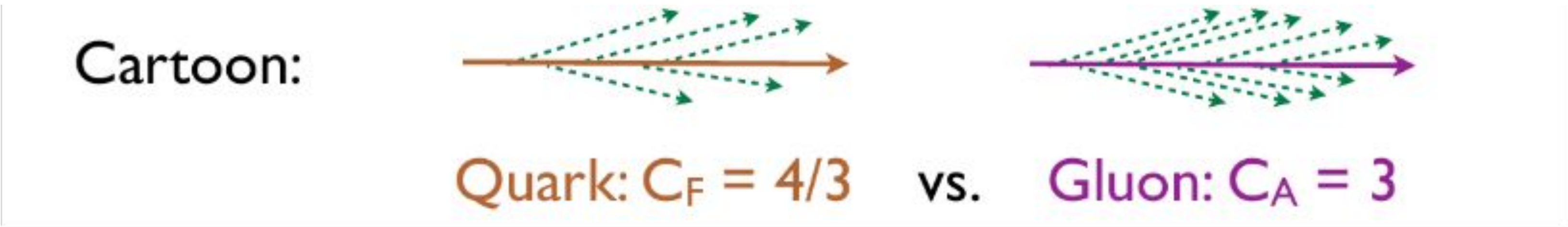


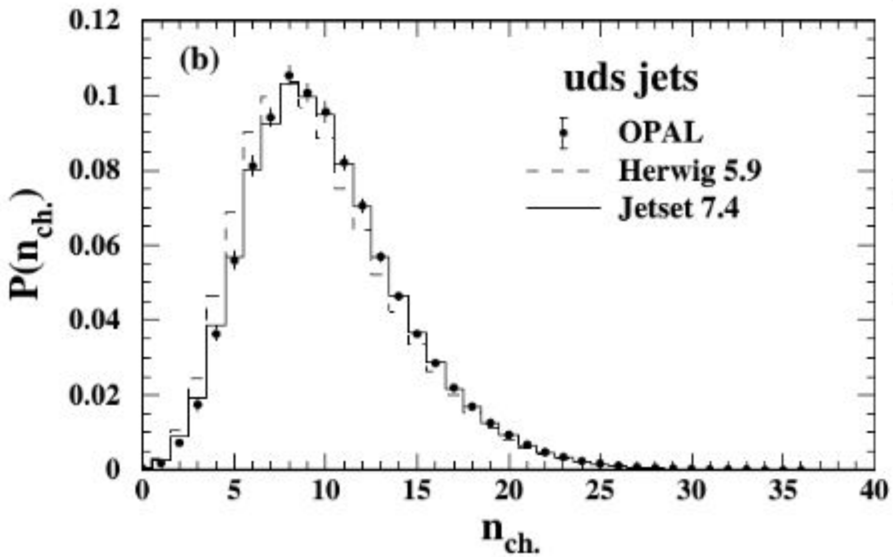
Fig. From J. Gallicchio and M. D. Schwartz, Phys. Rev. Lett.107 (2011)

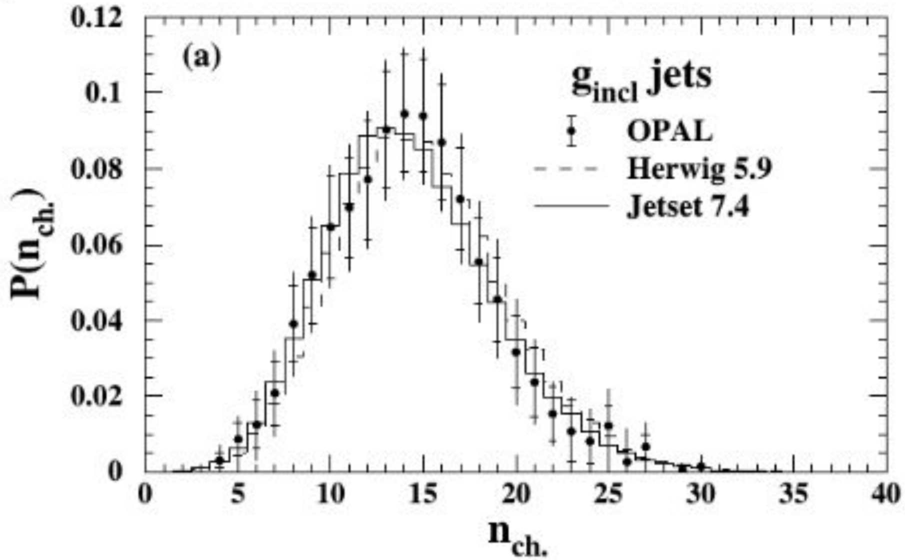
- Higgs $H^+ \to c\bar{s}$ (for charged Higgs mass between τ and t mass)
- Measure Z' coupling to hadrons (or find a leptophobic Z'/W')



Gluon will radiate more, gluon will radiate wider

$$\frac{\langle N_g \rangle}{\langle N_q \rangle} = \frac{C_A}{C_F}$$



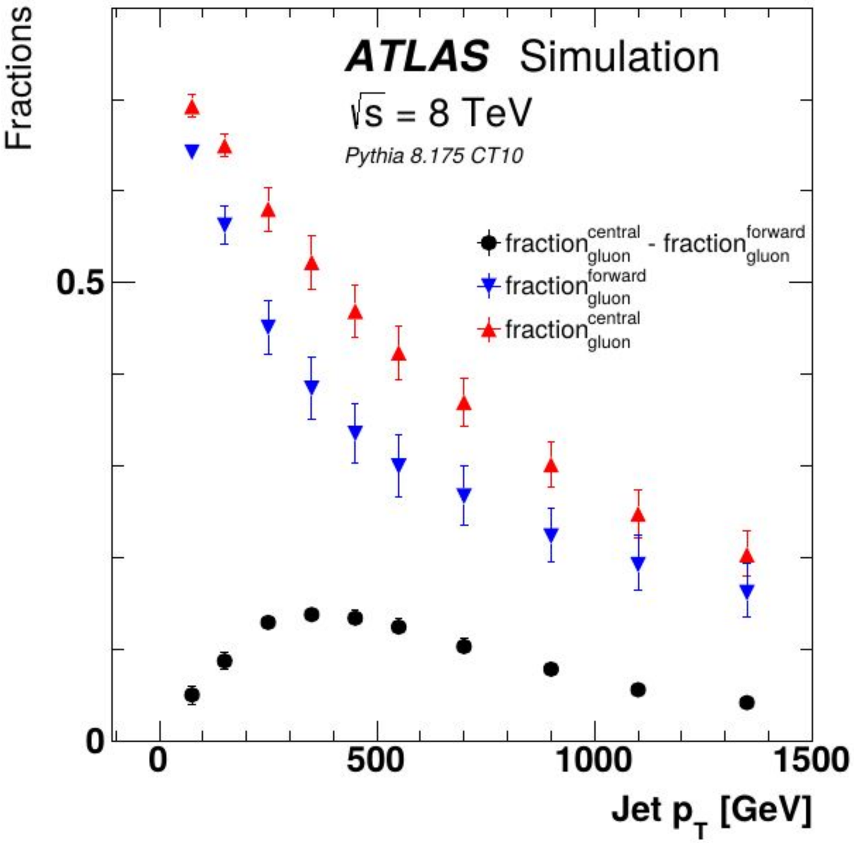


“Multiplicity distributions of gluon and quark jets and tests of QCD analytic predictions” [hep-ex/9708029]

LHC how to define G enhanced sample

Quark versus Gluon Jet Tagging Using Charged Particle Multiplicity with the ATLAS Detector #7

ATLAS Collaboration (Apr 11, 2017)



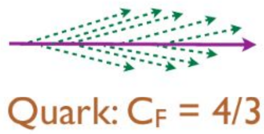
Using phase space cuts, for example:

- Pt - jet transverse momentum
- η - jet rapidity (central/forward)

But then we will have quark and gluon sample jets with different (Pt, η).

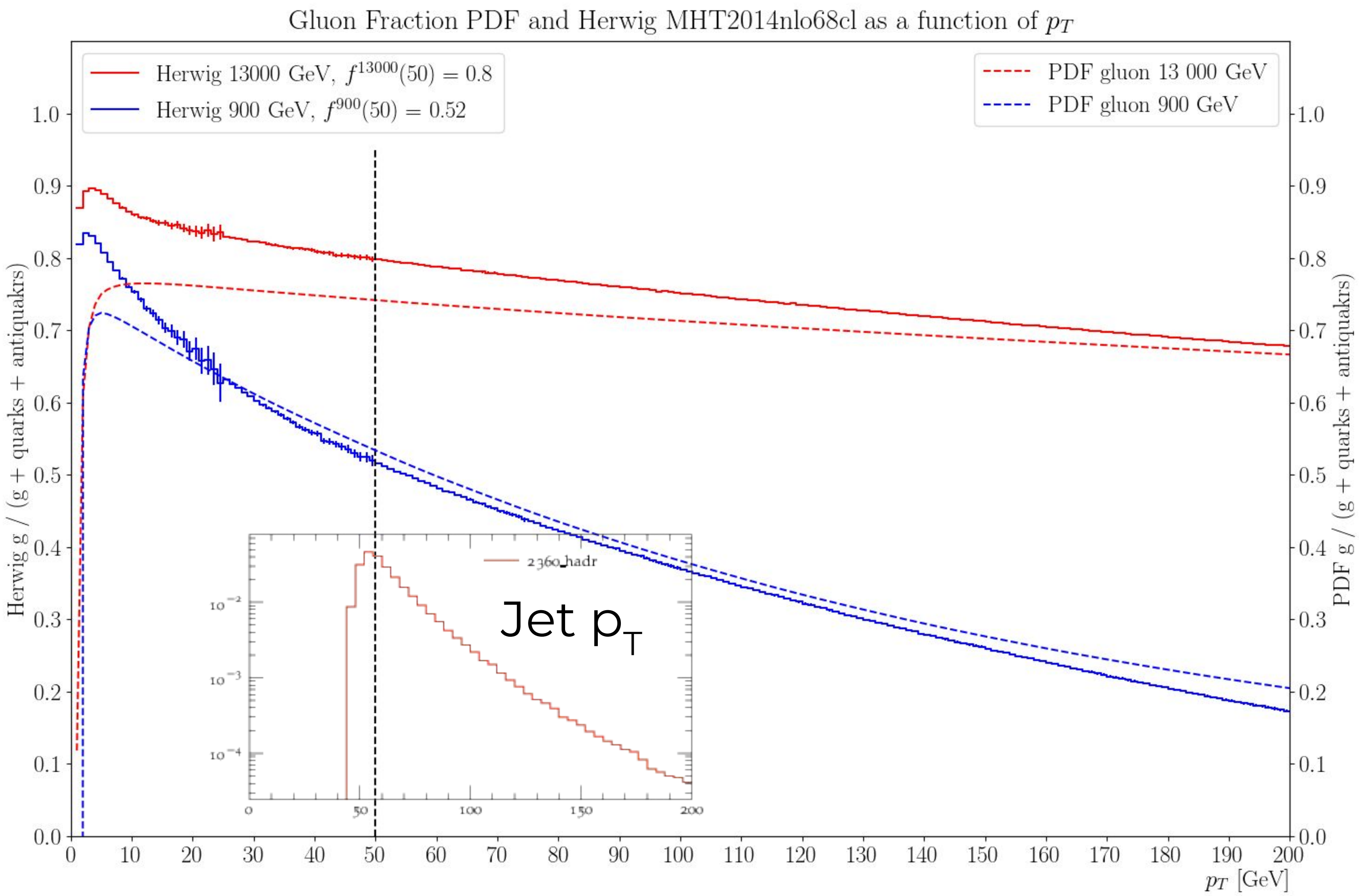
Same Pt Quark and Gluon

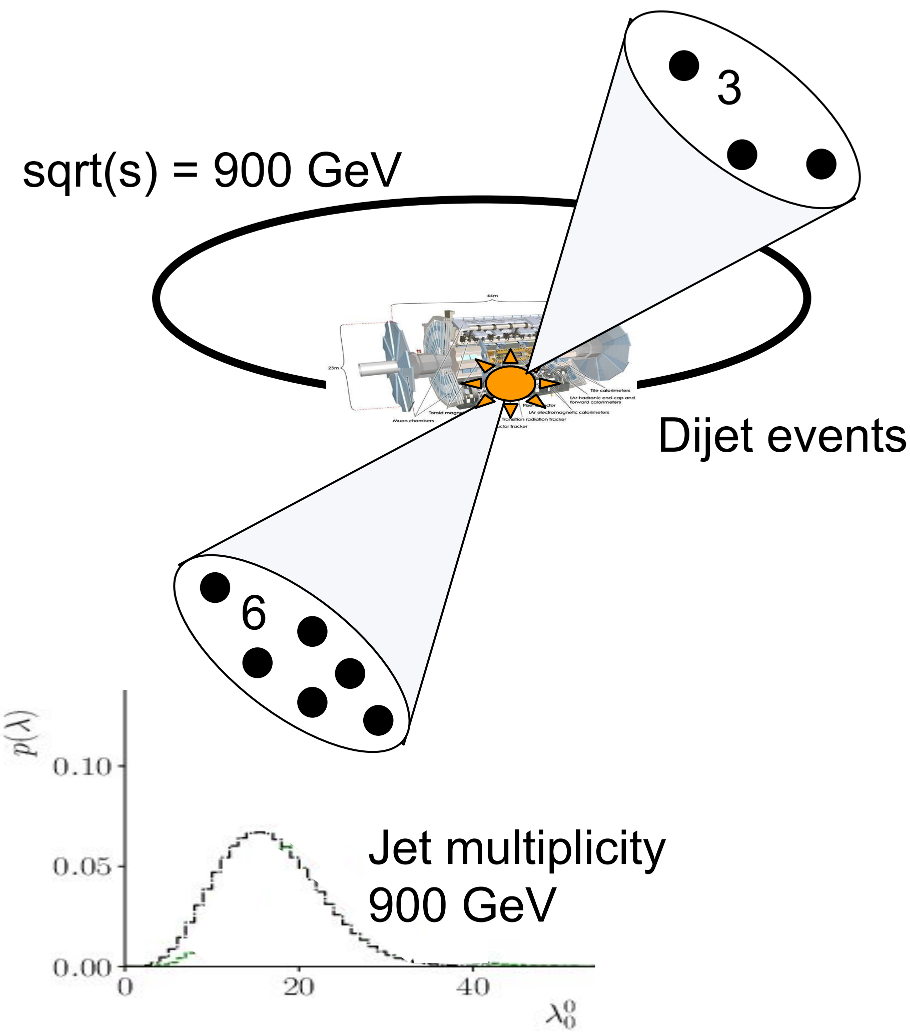
But hight Pt Q will radiate more and look like a G

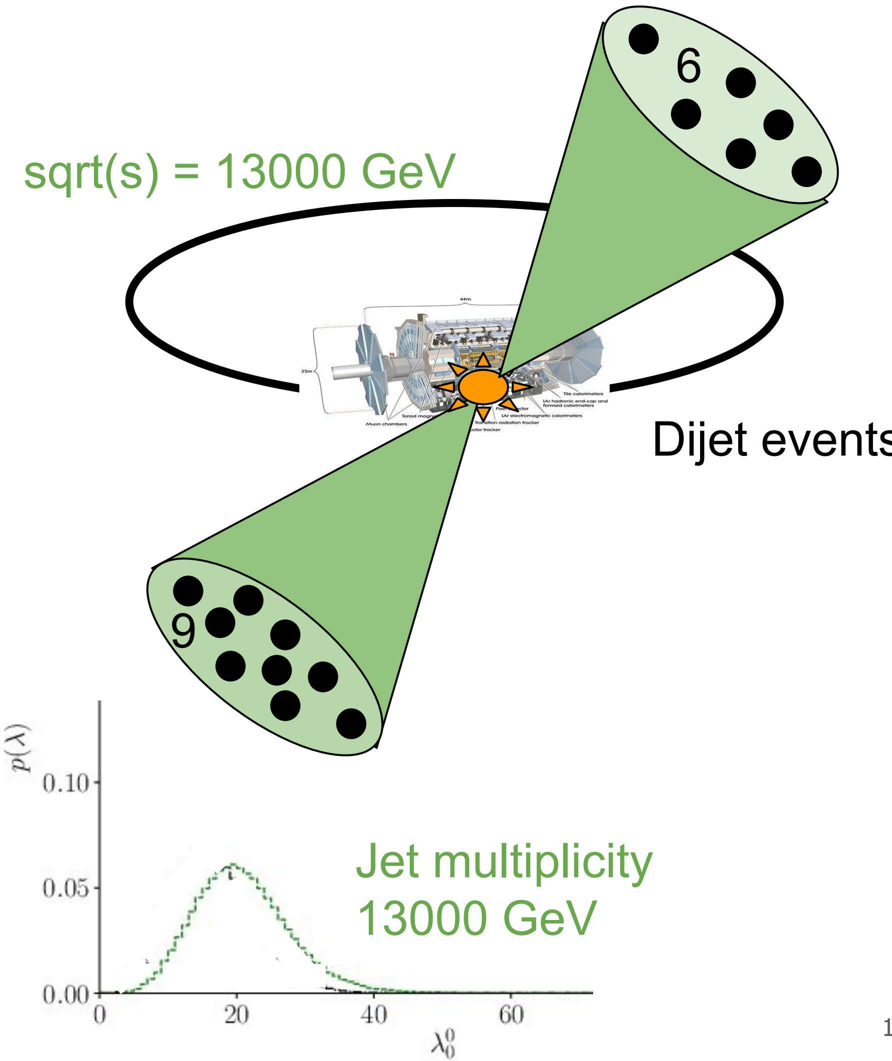


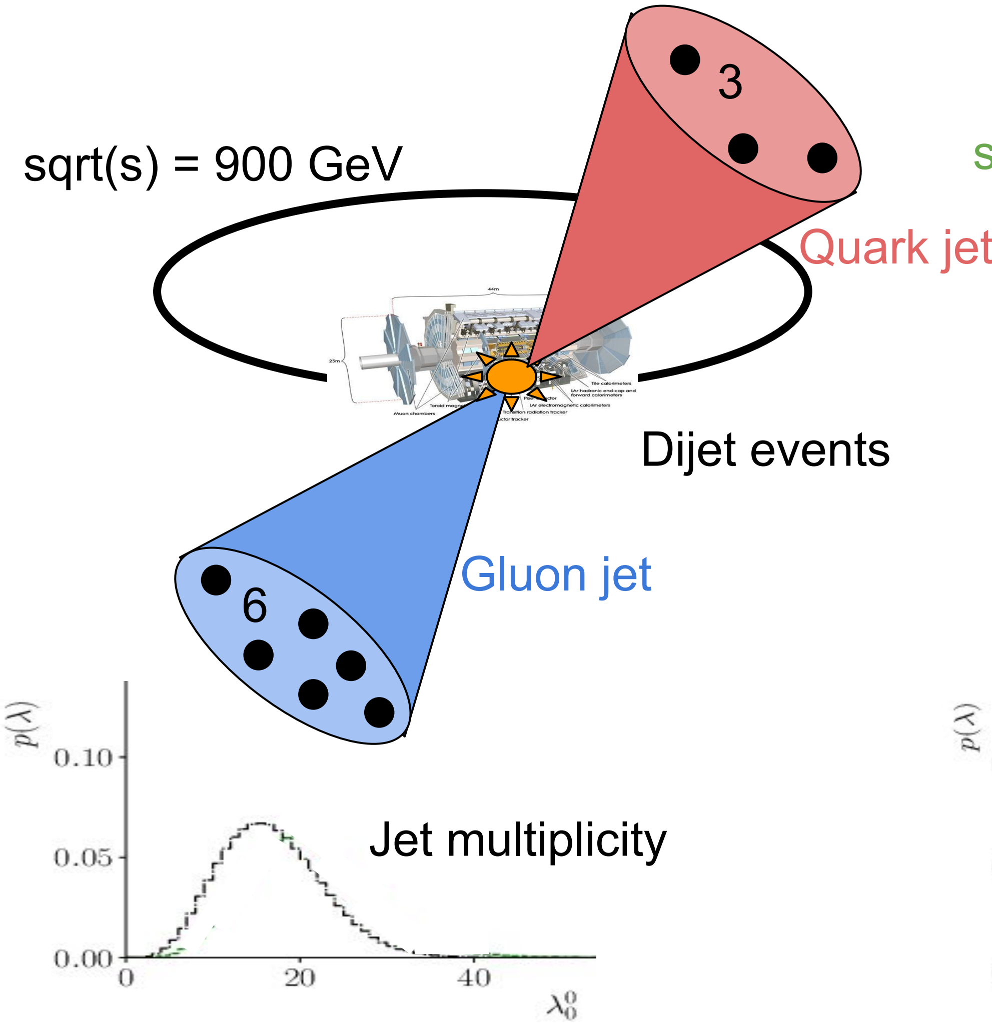
Can we find a way to get enhanced Q/G with the same Pt, η?

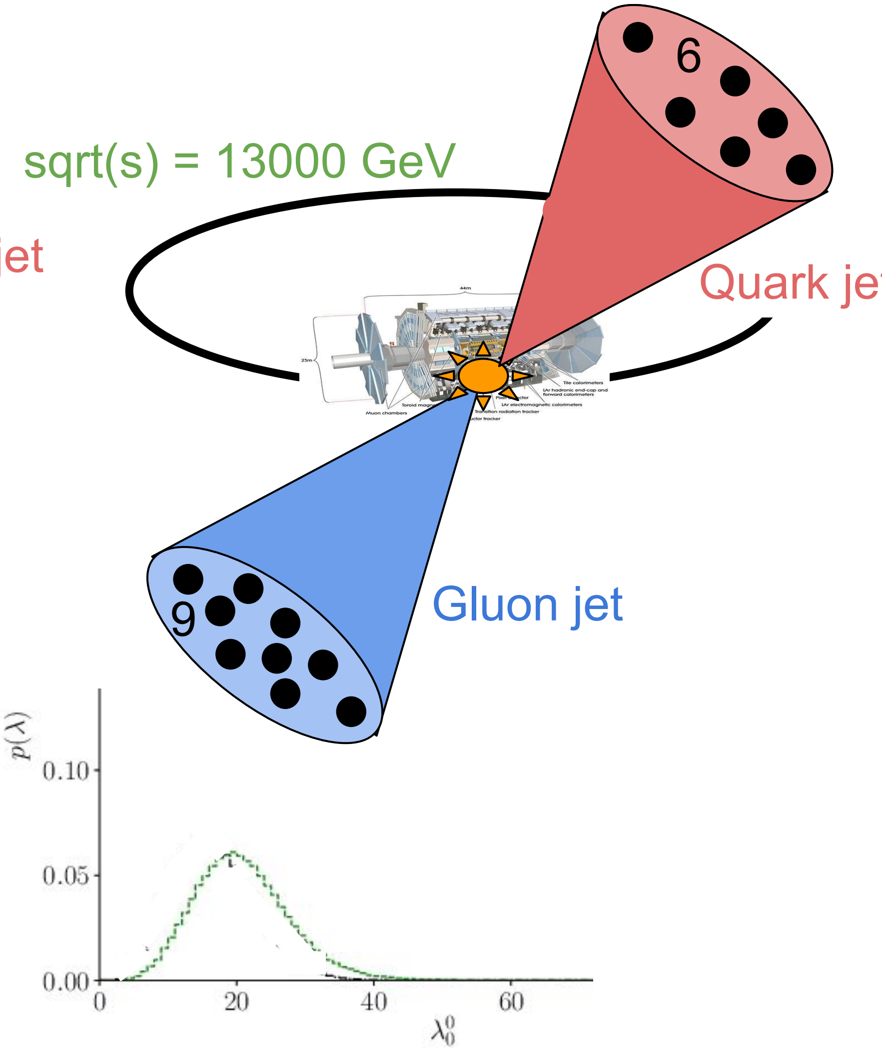
Measurement at different energy




sqrt(s) = 900 GeV
3
6
Dijet events
Jet multiplicity
900 GeV
$p(\lambda)$
0.10
0.05
0.00
0
20
40
λ_0^0

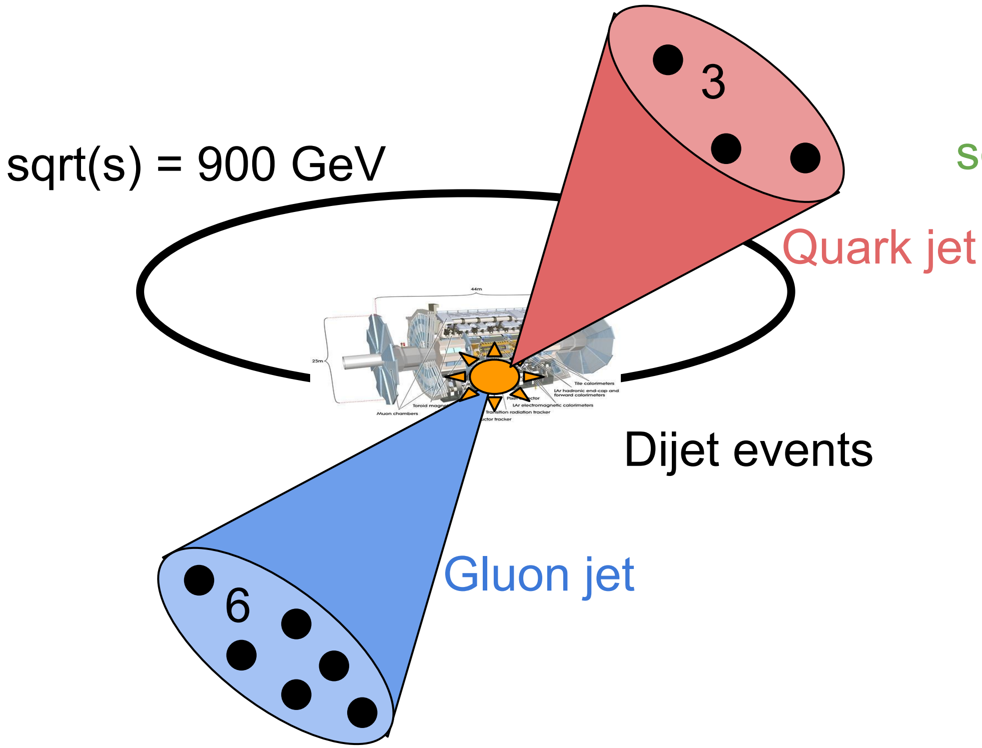

sqrt(s) = 13000 GeV
6
9
Dijet events
Jet multiplicity
13000 GeV
$p(\lambda)$
0.10
0.05
0.00
0
20
40
60
λ_0^0

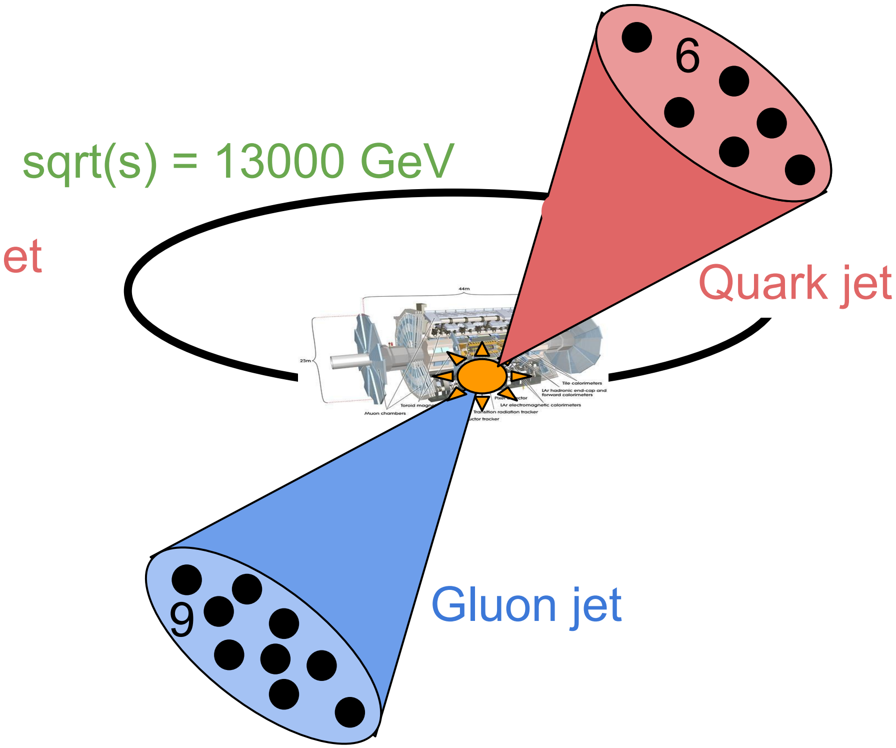




$$\lambda^{900} = f^{900} \lambda_g + (1 - f^{900}) \lambda_q$$

$$\lambda^{13000} = f^{13000} \lambda_g + (1 - f^{13000}) \lambda_q$$



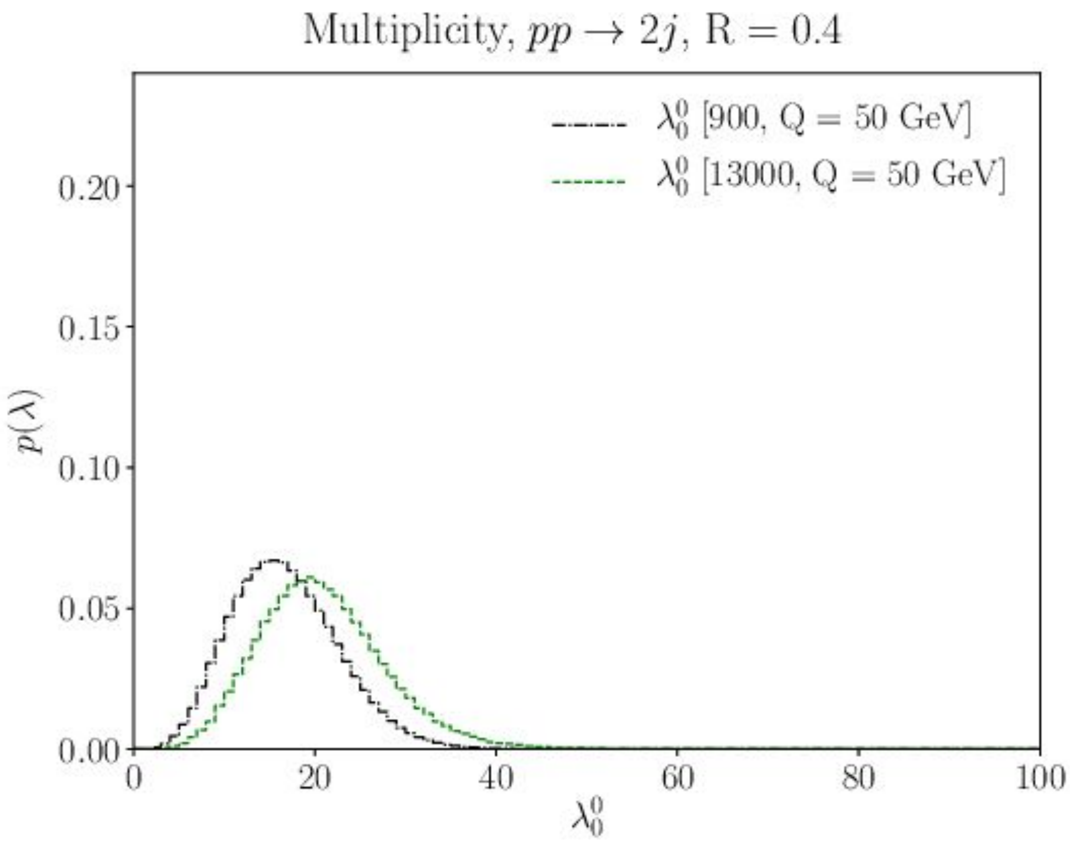


$$\lambda^{900} = f^{900}\lambda_g + (1 - f^{900})\lambda_q$$

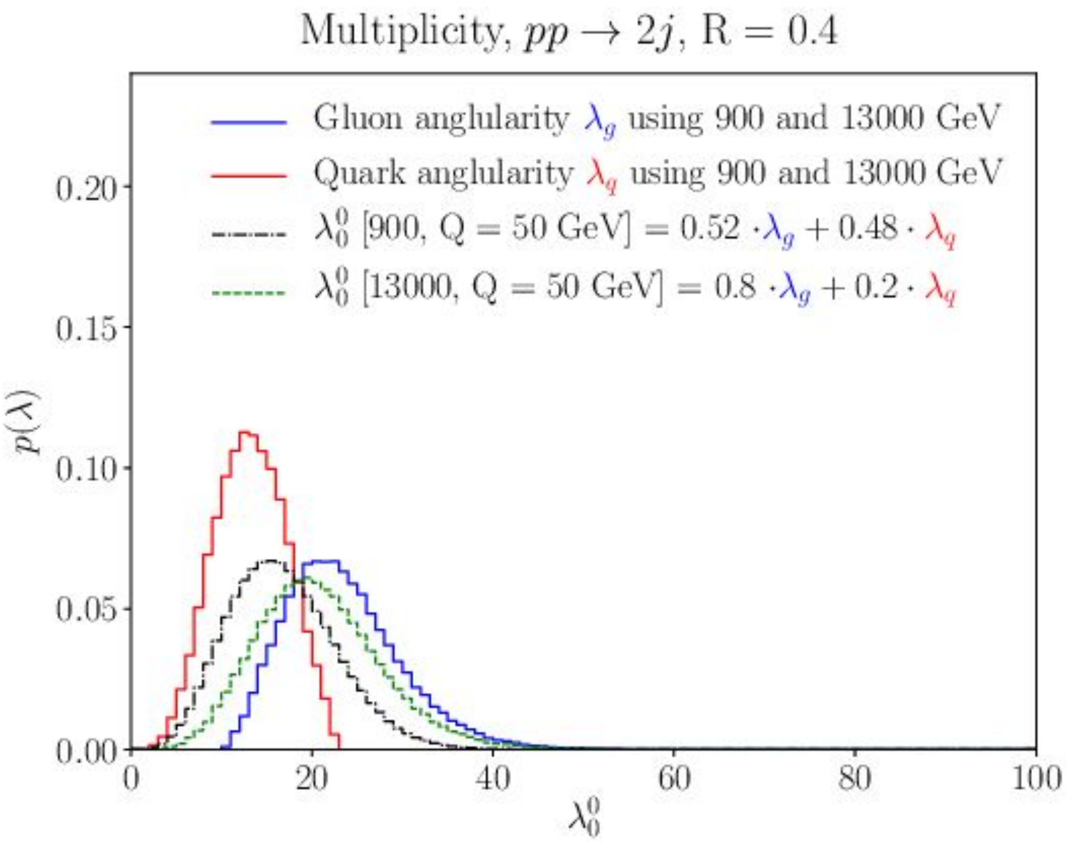
$$\lambda^{13000} = f^{13000}\lambda_g + (1 - f^{13000})\lambda_q$$

$\lambda_q = \frac{f^{900}\lambda^{13000} - f^{13000}\lambda^{900}}{f^{900} - f^{13000}}$	$\lambda_g = \frac{(1 - f^{13000})\lambda^{900} - (1 - f^{900})\lambda^{13000}}{f^{900} - f^{13000}}$

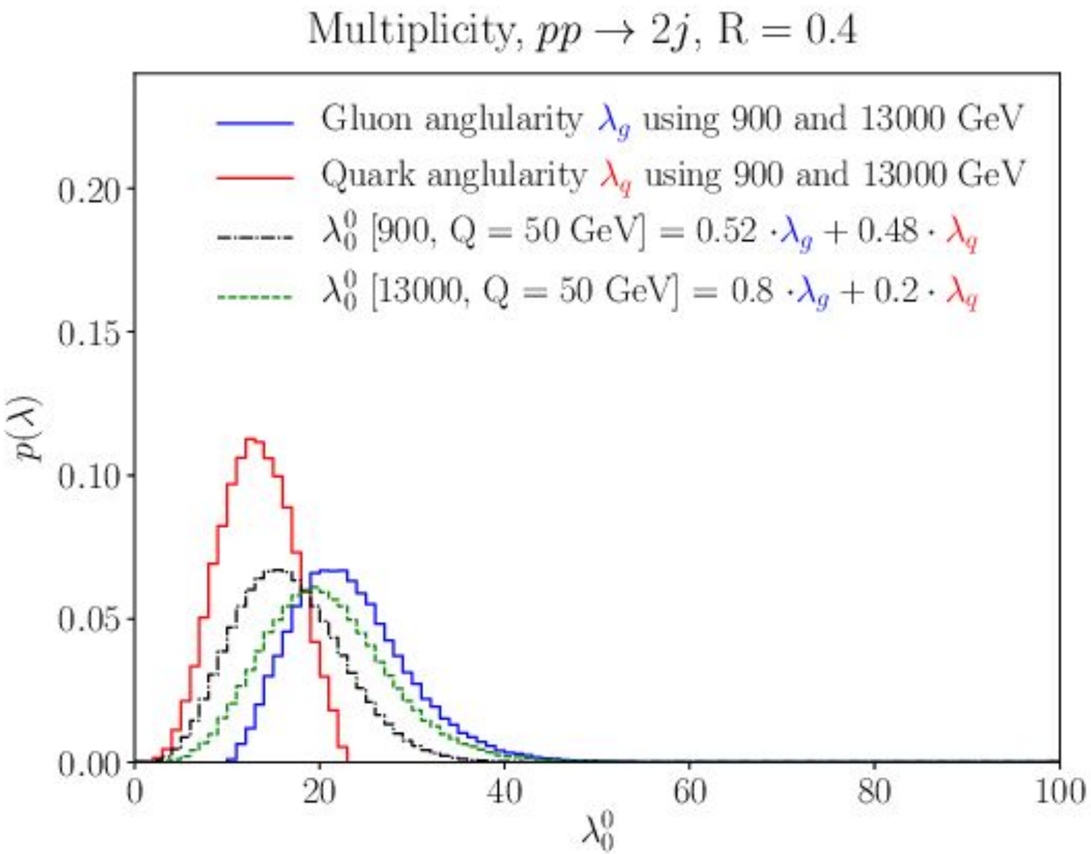
The method in practice: measurement



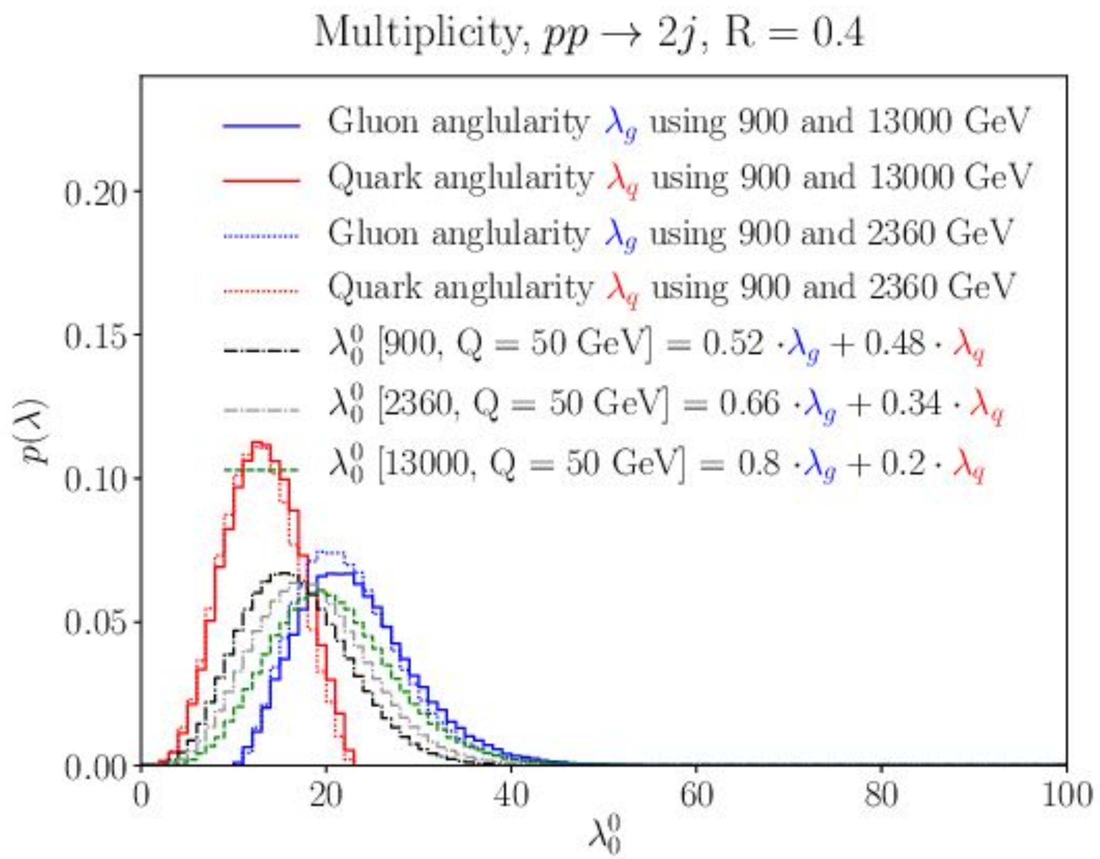
The method in practice: q/g distributions



The method in practice: energy dependence



The method in practice: energy dependence



Results

We considered all combinations of:

– 5 – angularities λ_0^0, $\lambda_{0.5}^1$, λ_1^1, λ_0^2, λ_2^1
– 2 – using groomed (MMDT) / not groomed jets
– 5 – jet radii $R = 0.2, 0.4, 0.6, 0.8, 1.0$
– 4 – regions - dijet average $p_T^{\text{cut}} = 50$ GeV, 100, 200, and 400 GeV

$(p_{T\ \text{lead}} + p_{T\ \text{sublead}})/2 > p_T^{\text{cut}}$ $\qquad p_{T\ \text{sublead}}/p_{T\ \text{lead}} > 0.8$

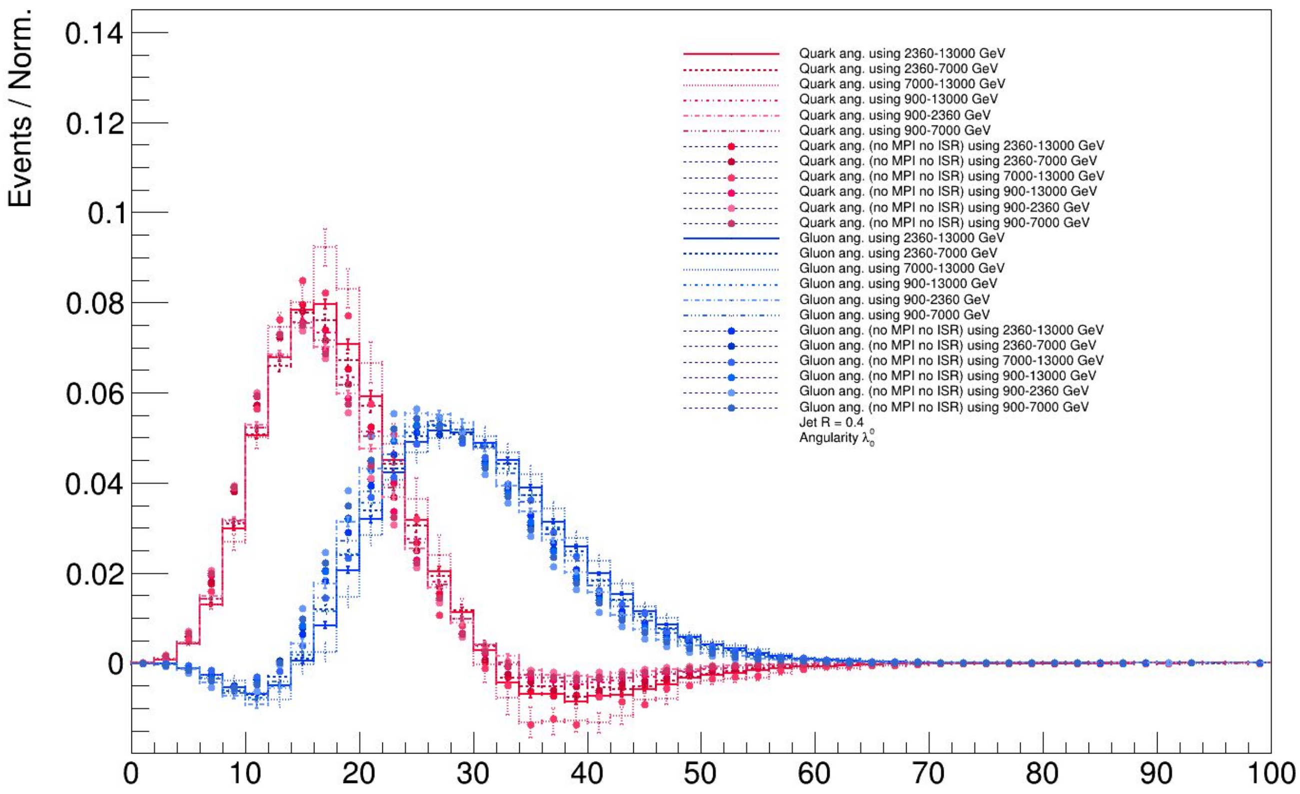
– 2 – quark/gluon
– 2 – MPI and ISR switched on/off
– 6 – energy combinations: 900–2360, 900–7000, 900–13000, 2360–7000, 2360–13000, 7000–13000 GeV
– 2 – event generators Herwig and Pythia

[in total 9600 distributions]

The method in practice: energy dependence

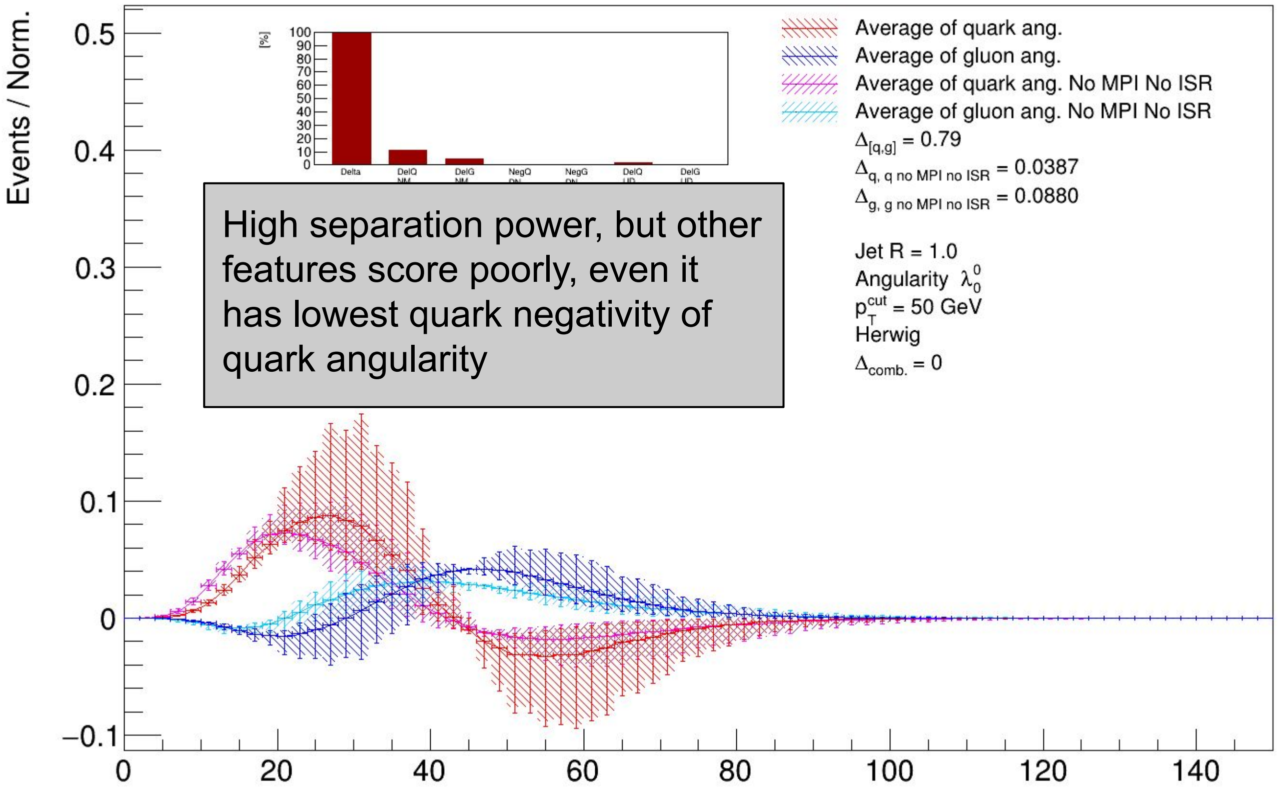
Let's use more 6 energy combinations:

900-2360, 900-7000, 900-13000, 2360-7000, 2360-13000, 7000-13000 GeV

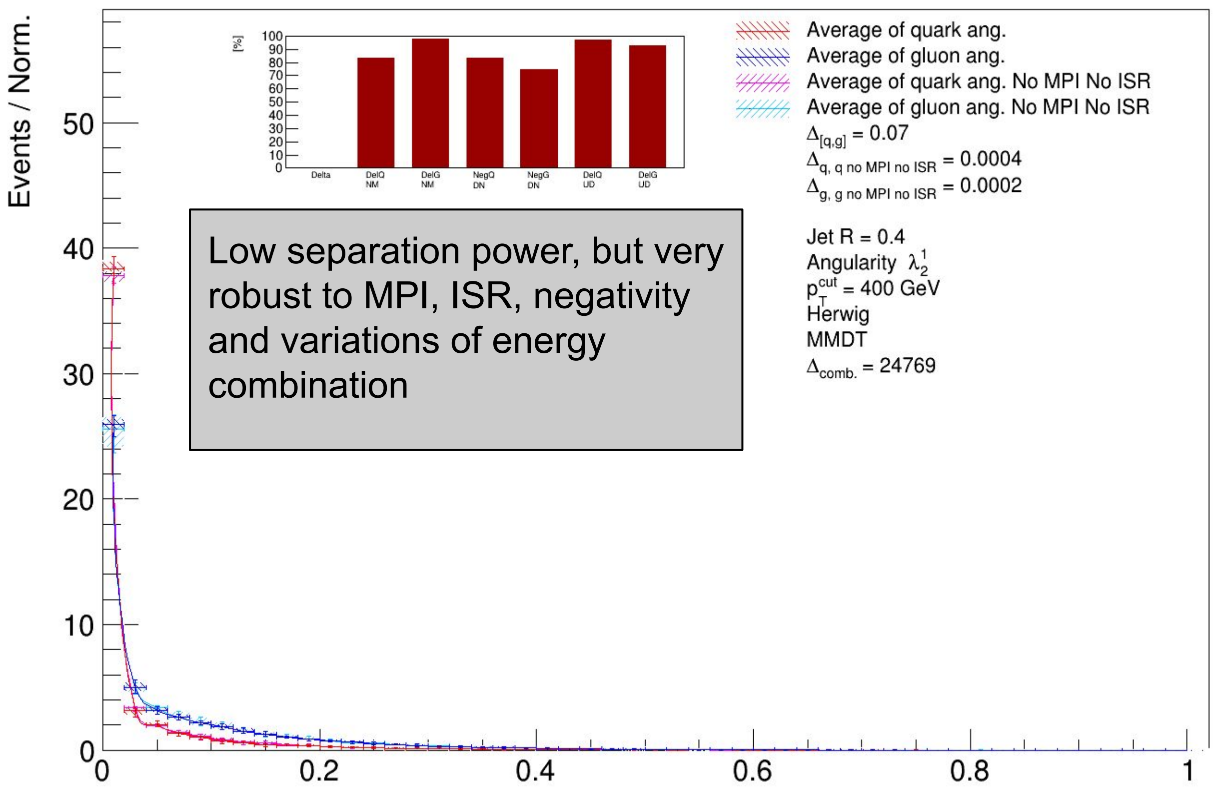


Dotted lines test the **robustness** to Multi Parton Interactions **MPI** and Initial State Radiation **ISR**

Why we looking into other features then separation power (bad examples):



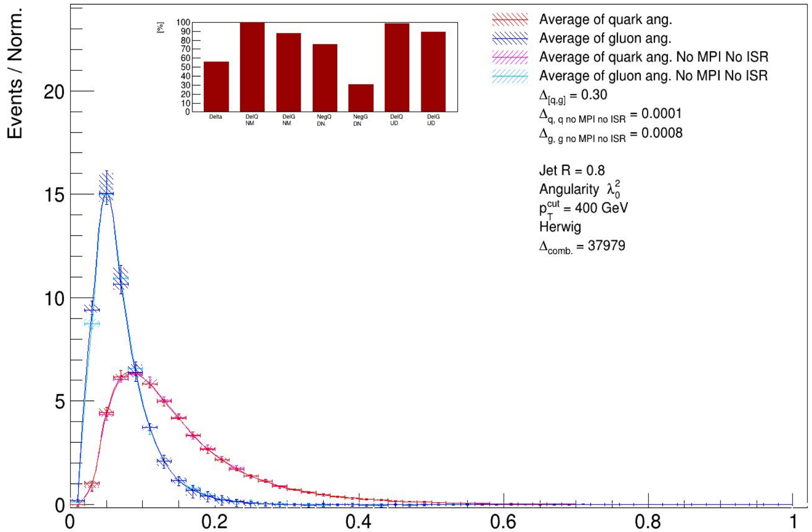
Why we looking into other features then separation power (bad examples):



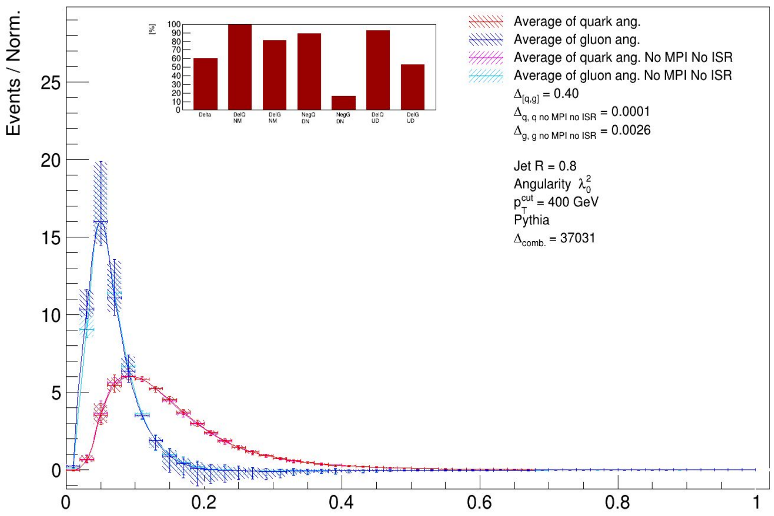
Results - best results

Best performing angularities: p_T^D

Herwig 7

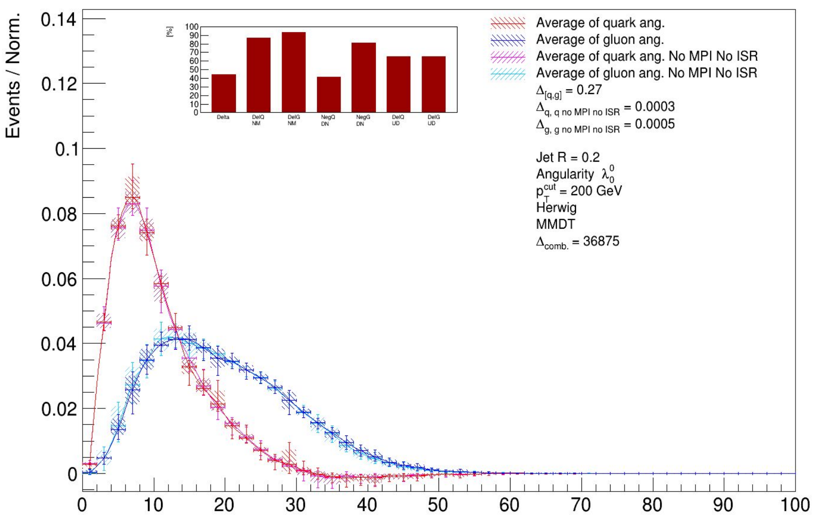


Pythia 8

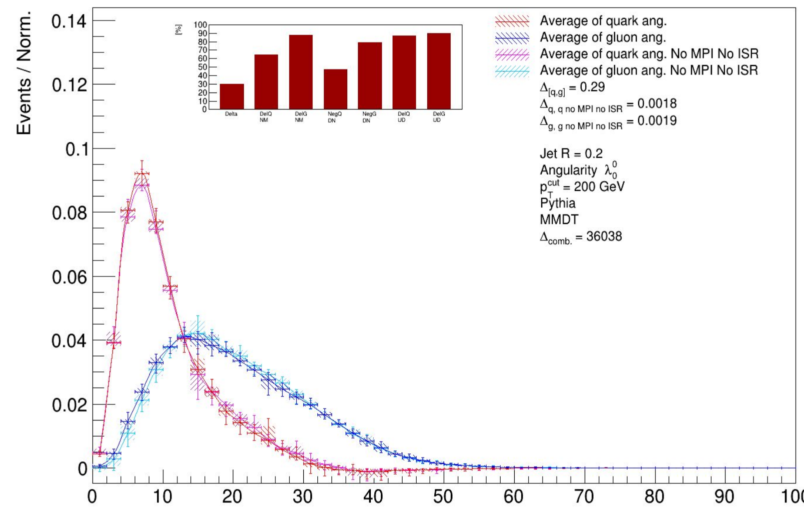


Best performing angularities: Multiplicity

Herwig 7

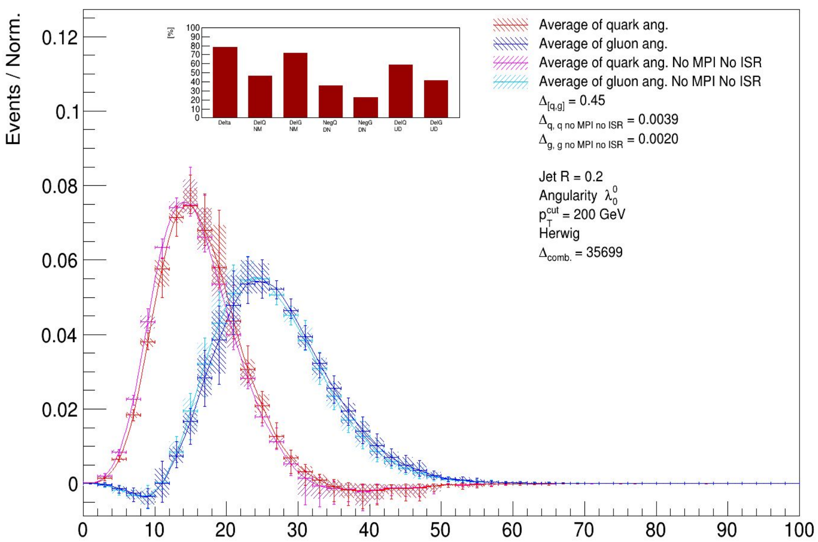


Pythia 8

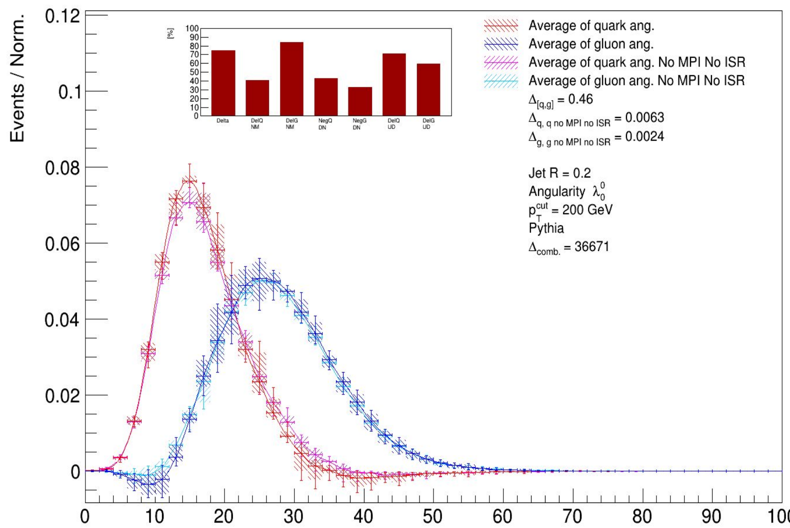


Wild cards (chosen by “eye”): multiplicity

Herwig 7



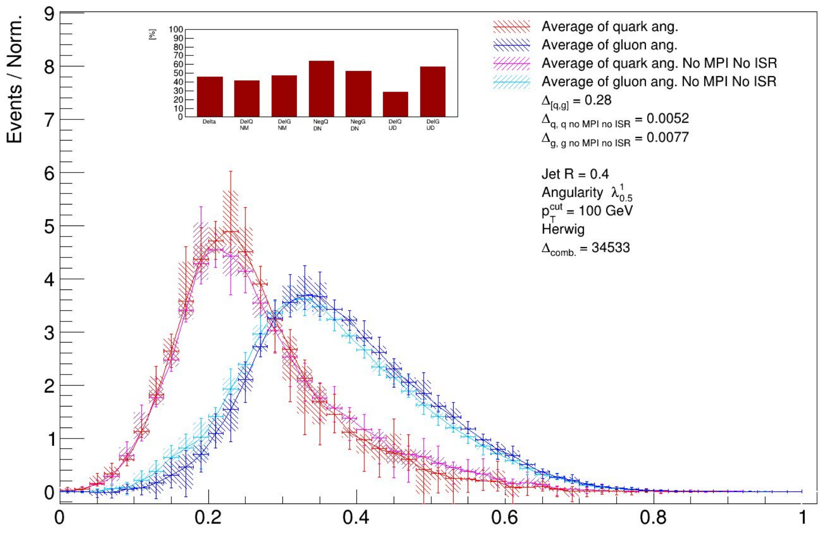
Pythia 8



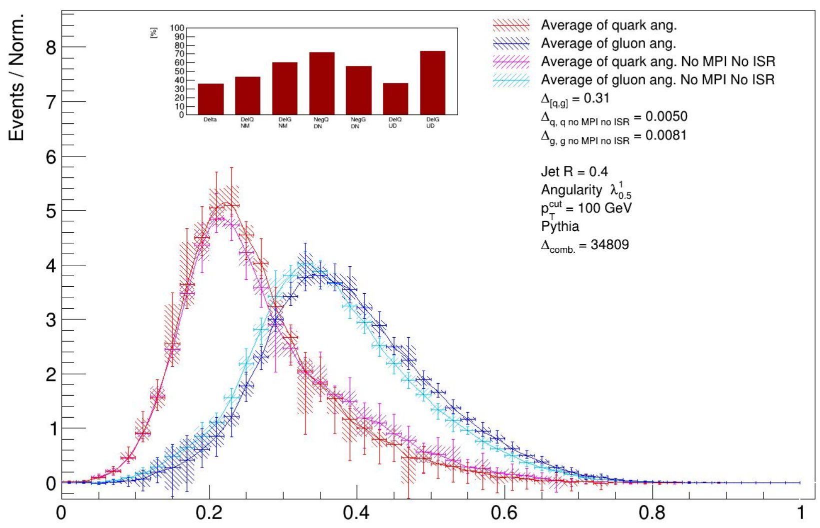
Results - best results

Wild cards (chosen by “eye”): LHA

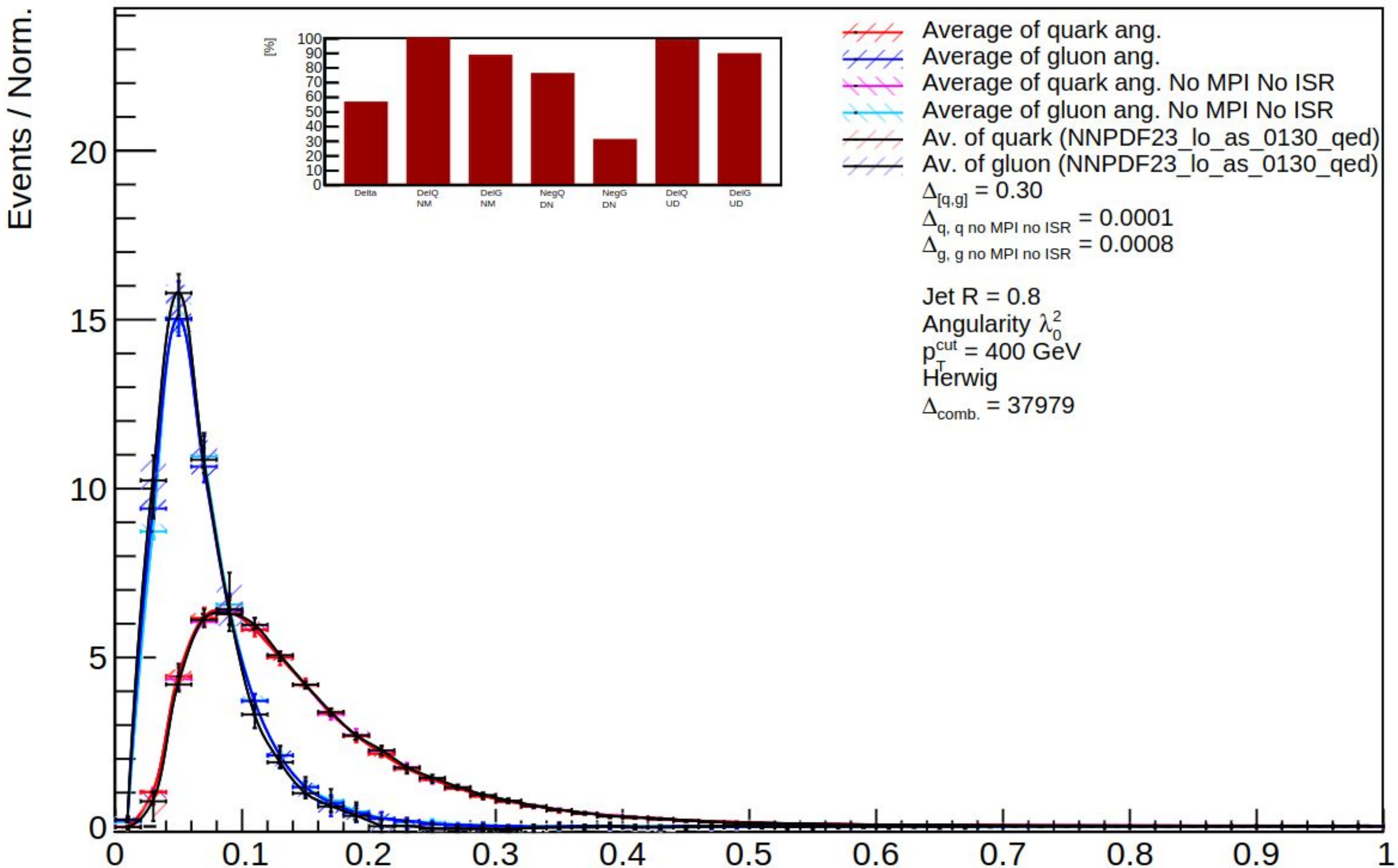
Herwig 7



Pythia 8



Results - PDF dependence



Summary

[see arXiv: 2307.15378 for details]

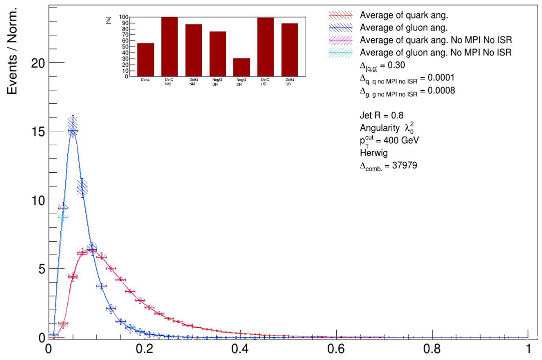
Conclusion:

- We propose a **novel way** to **measure Q/G jets**
- It is **free of the kinematic biases** - problematic in other methods
- Uses **unique** opportunity that LHC ran at **different energies**
- **Best results** are **robust** to Initial State Radiation and MPI

Outlook:

- The **main aim** is to perform the **measurement** at LHC but it looks like experiments are not ready to reanalyse even 7 TeV data! We see how important is legacy data from LEP data (next slides)!
  - We are **happy** to **join** the effort (Petr is member of ATLAS)
  - **Feasible**, seems like ALICE has already measured all what is needed

Thank you for your attention!

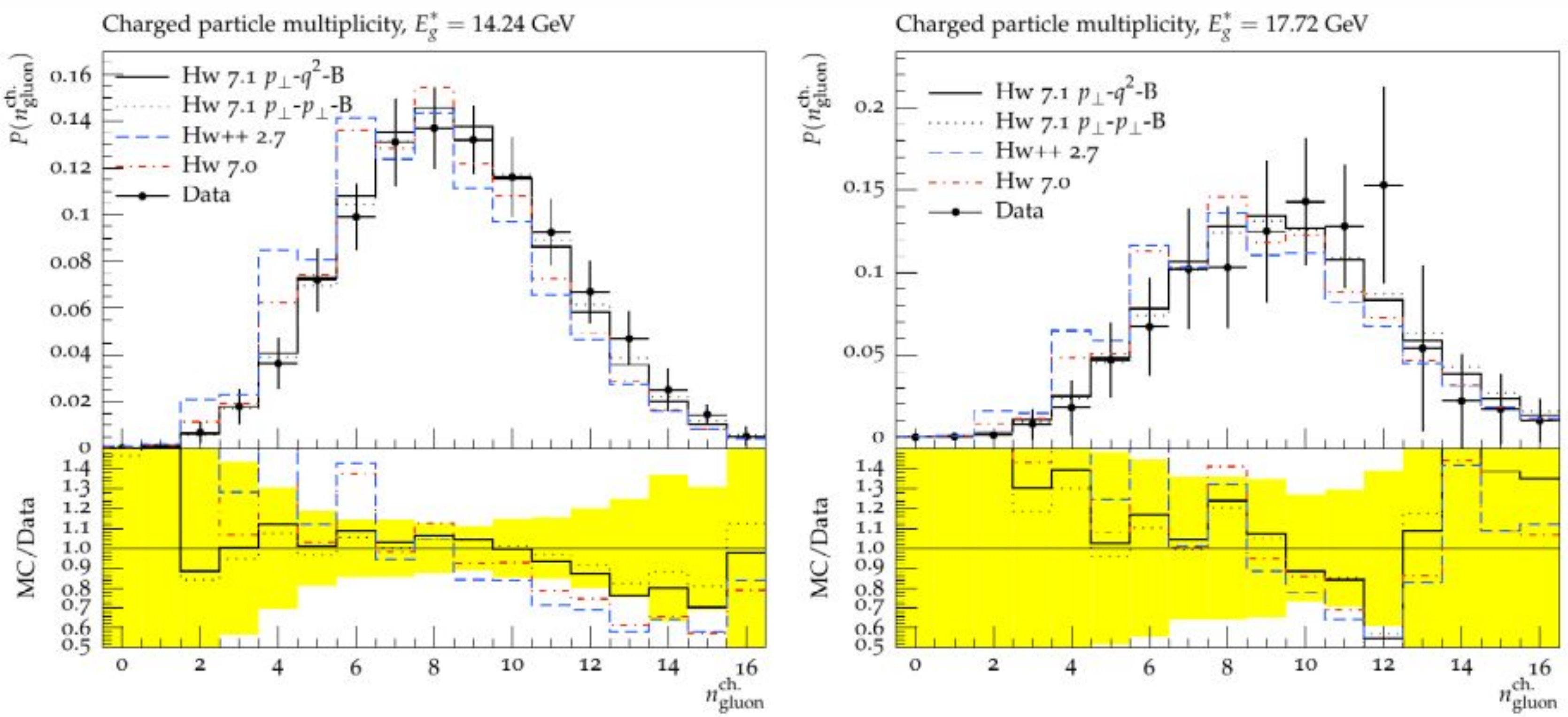


Improving the Simulation of Quark and Gluon Jets with Herwig 7 #17

Daniel Reichelt (Dresden, Tech. U.), Peter Richardson (CERN and Durham U., IPPP), Andrzej Siodmok (Cracow, INP) (Aug 4, 2017)

Published in: *Eur.Phys.J.C* 77 (2017) 12, 876 • e-Print: 1708.01491 [hep-ph]

Multiplicity distribution of charged particles in gluons jets for two different gluon energies.



Data was one of the **key for the improvement** and it is still needed for the progress. However it is hard to measure "clear" q/g samples at the LHC.

Recent progress: Machine learning hadronization

First steps for ML hadronization:

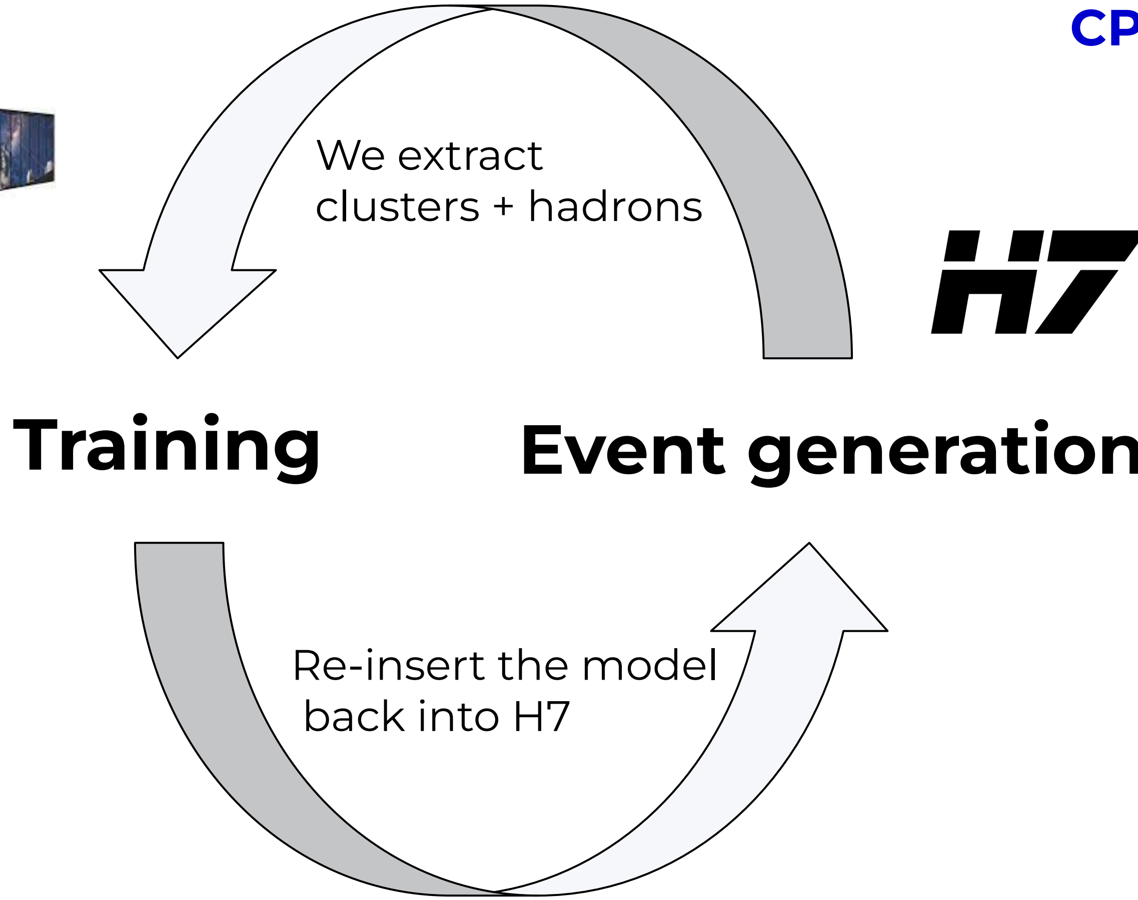
- HADML - [A. Ghosh, Xi. Ju, B. Nachman **AS**, *Phys.Rev.D* 106 (2022) 9]
- MLhad - [P. Ilten, T. Menzo, A. Youssef and J. Zupan, SciPost Phys. 14, 027 (2023)]

	MLhad	HADML
Deep generative model:	Variational Autoencoder	Generative Adversarial Networks
Trained on:	String model	Cluster model
Recent progress:	*“Reweighting Monte Carlo Predictions and Automated Fragmentation Variations in Pythia 8”* [Bierlich, Ilten, Menzo, Mrenna, Szewc, Wilkinson, Youssef, Zupan, 2308.13459] (see Christian’s talk)	*“Fitting a Deep Generative Hadronization Model”* [J. Chan, X. Ju, A. Kania, B. Nachman, V. Sangli and **AS**, JHEP 09 (2023) 084]

Integration into Herwig

GPU

CPU



This then allows us to run a full event generator and produce plots

ALICE - angularites at 5.02 TeV

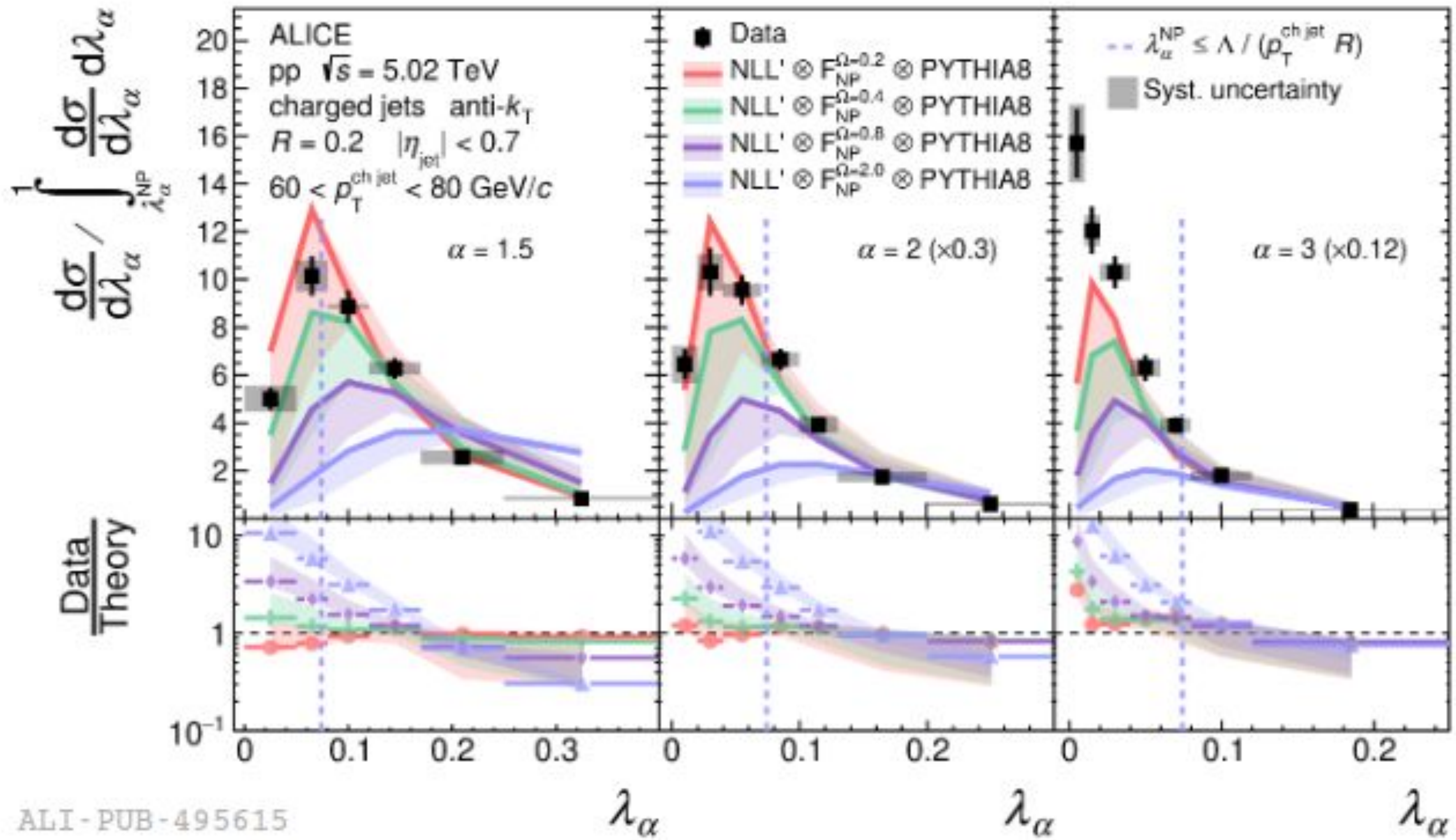


Figure 1 – Example comparison of the ungroomed jet angularities with different values of $\alpha \in \{1.5, 2, 3\}$ to NLL′ SCET calculations convolved with different amounts of smearing with a nonperturbative shape function [4]. Distributions are normalized to the perturbative region, right of the vertical dashed line, which uses $\Lambda = 1$ GeV.

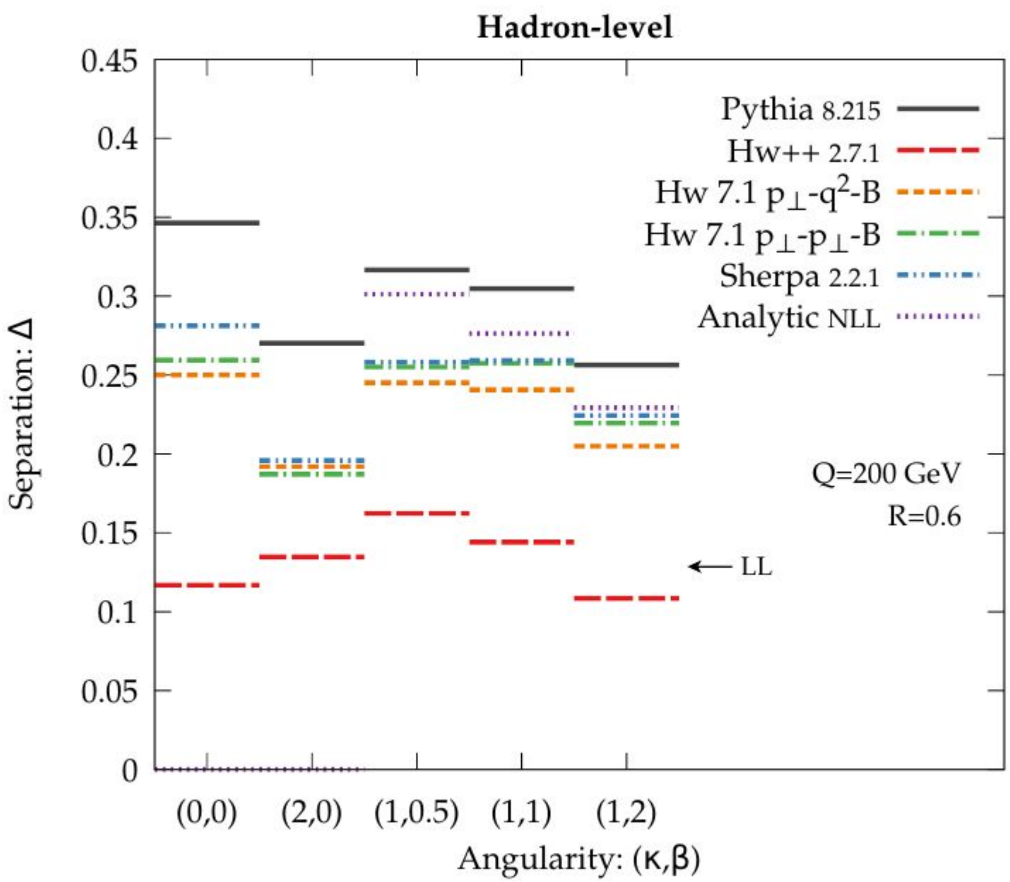
Backup

How we improved simulation of Q/G jets in Herwig

Improving the Simulation of Quark and Gluon Jets with Herwig 7

Daniel Reichelt (Dresden, Tech. U.), Peter Richardson (CERN and Durham U., IPPP), Andrzej Siodmok (Cracow, INP) (Aug 4, 2017)

Published in: *Eur.Phys.J.C* 77 (2017) 12, 876 • e-Print: 1708.01491 [hep-ph]



How we improved simulation of Q/G jets in Herwig

Improving the Simulation of Quark and Gluon Jets with Herwig 7

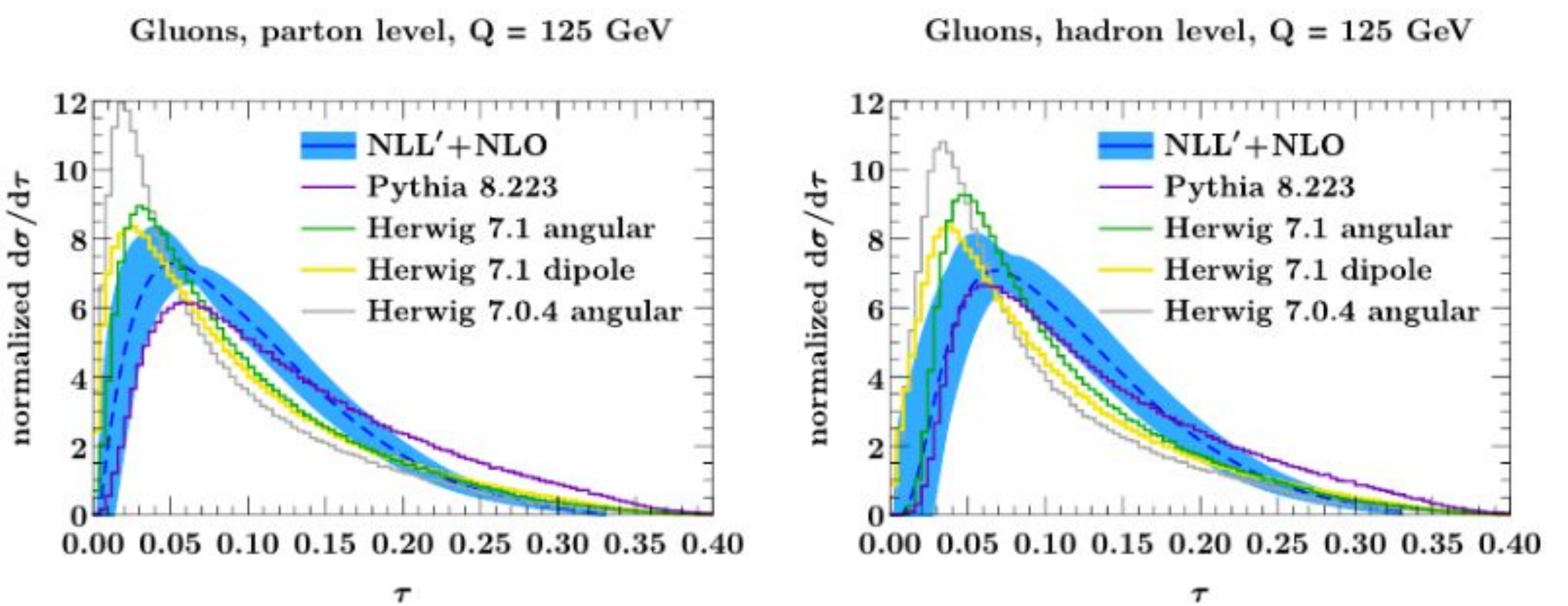
Daniel Reichelt (Dresden, Tech. U.), Peter Richardson (CERN and Durham U., IPPP), Andrzej Siodmok (Cracow, INP) (Aug 4, 2017)

Published in: *Eur.Phys.J.C* 77 (2017) 12, 876 • e-Print: 1708.01491 [hep-ph]

"A case study of quark-gluon discrimination at NNLL0 in comparison to parton showers"

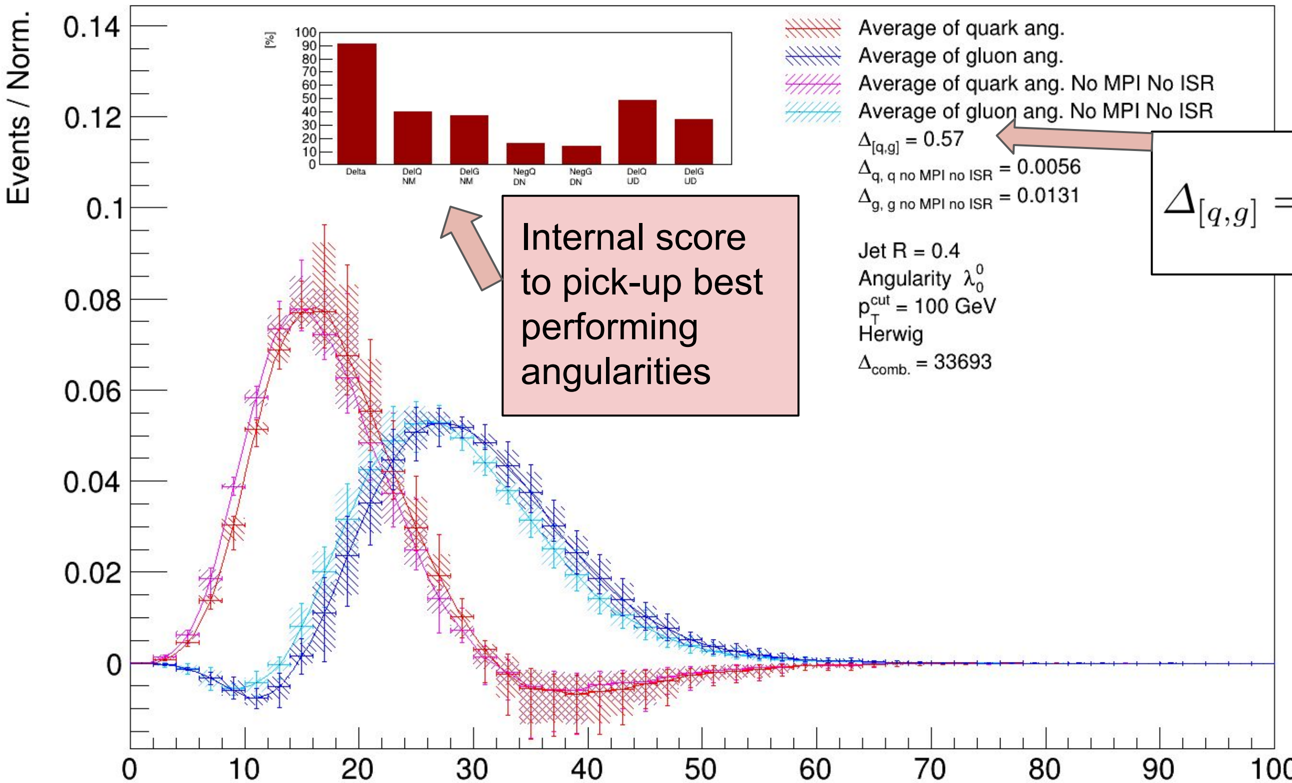
Thrust - similar to general angularity (1,2) but not restricted to particles in a jet.

$$T = \max_{\hat{t}} \frac{\sum_i |\hat{t}\cdot\vec{p}_i|}{\sum_i |\vec{p}_i|}, \quad \tau = 1 - T$$



"This highlights the substantial improvement in the description of gluon jets in the latest version of Herwig"

Simplified averaged plot over 6 energy combinations: - filled area (energy comb. variation),
- ticks - stat. Unc.

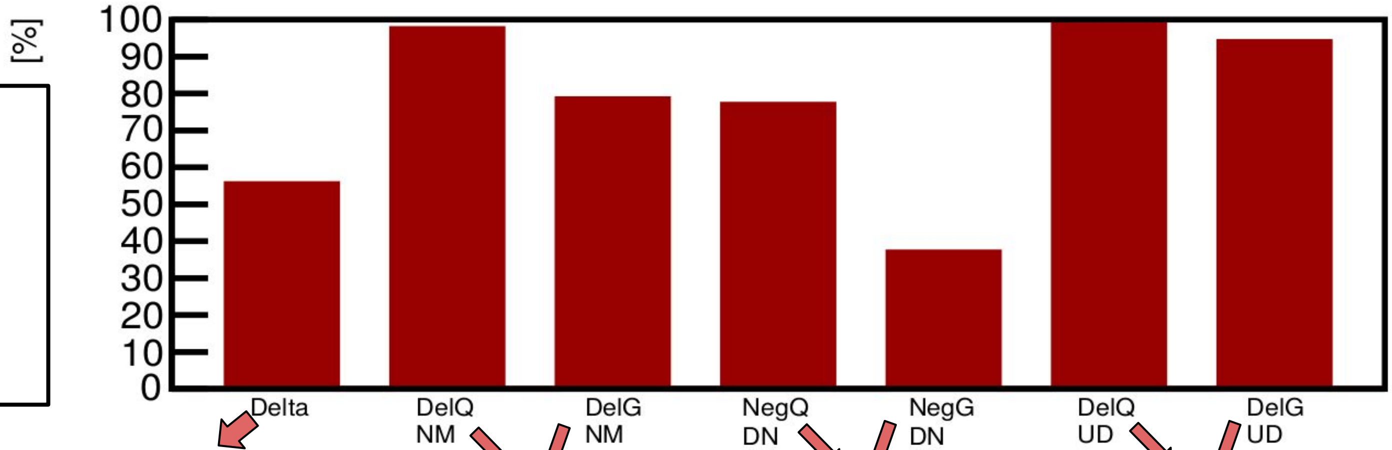


Separation power:

$$\Delta_{[q,g]} = \frac{1}{2} \sum_{i=1}^{N} \frac{(\lambda_{q_i} - \lambda_{g_i})^2}{\lambda_{q_i} + \lambda_{g_i}}$$

[in total 9600 distributions]

Results



Each column represents percentile of given feature of all our studied plots.

In **separation between Q/G jets** about 55 % of variations have lower performance

Separation power between noMPI variations - **robustness to noMPI no ISF**

Negativity (percentage of negative area to whole area) **negative bins**

Separation power to UP and DOWN energy combination variations **robustness to different energy combination used**

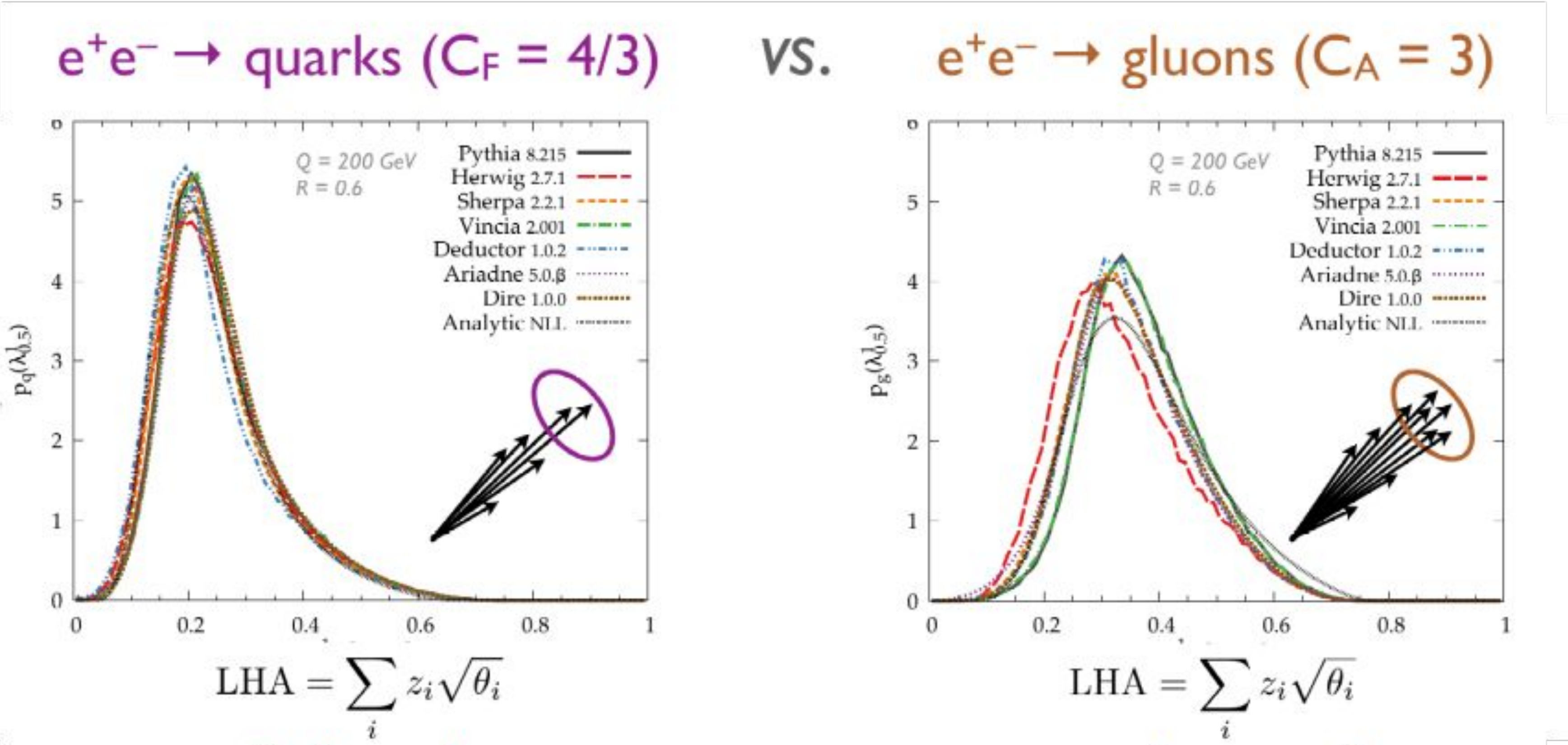
Combining all columns gives us our internal score :

$$\Delta_{\mathrm{comb}} = 1000 \cdot \ln \Big[1 + (\mathrm{Delta})^3 \cdot (\mathrm{DelQ\ NM}) \cdot (\mathrm{DelG\ NM}) \cdot (\mathrm{NegQ\ DN}) \cdot (\mathrm{NegG\ DN}) \cdot (\mathrm{DelQ\ UD}) \cdot (\mathrm{DelG\ UD})\Big]$$

min-max
0-41447

Q/G jet Les Houches study

[Gras, Hoeche, Kar, Larkoski, Lönnblad, Plätzer, AS, Skands, Soyez, Thaler, JHEP 1707 (2017) 091]



Small spread

Constrained by LEP

Large spread

Up to now no e+e- data has been used to constrain it.

What is a Quark Jet?

From lunch/dinner discussions

[slide by Jesse Thaler]

Ill-Defined ↓ Well-Defined	What people sometimes think we mean ↑ What we mean	
		A quark parton
		A Born-level quark parton
		The initiating quark parton in a final state shower
Quark as noun		An eikonal line with baryon number 1/3 and carrying triplet color charge
		A quark operator appearing in a hard matrix element in the context of a factorization theorem
		A parton-level jet object that has been quark-tagged using a soft-safe flavored jet algorithm (automatically collinear safe if you sum constituent flavors)
Quark as adjective		A phase space region (as defined by an unambiguous hadronic fiducial cross section measurement) that yields an enriched sample of quarks (as interpreted by some suitable, though fundamentally ambiguous, criterion)

Ill-Defined → Well-Defined

What we mean → What people sometimes think we mean

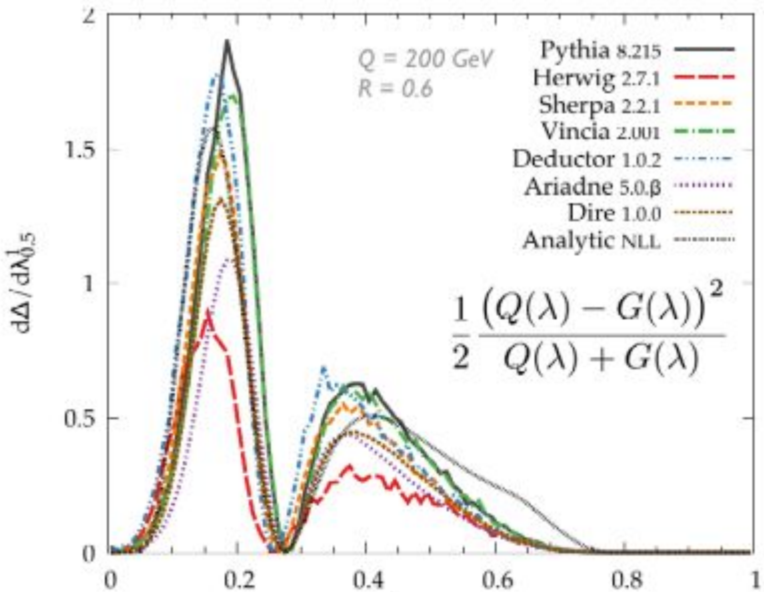
Q/G jet Les Houches study

$$\Delta = \frac{1}{2}\int \mathrm{d}\lambda \, \frac{\big(p_q(\lambda) - p_g(\lambda)\big)^2}{p_q(\lambda) + p_g(\lambda)}$$

$\Delta = 0$ - corresponds to no discrimination power.

$\Delta = 1$ - corresponds to perfect discrimination power.

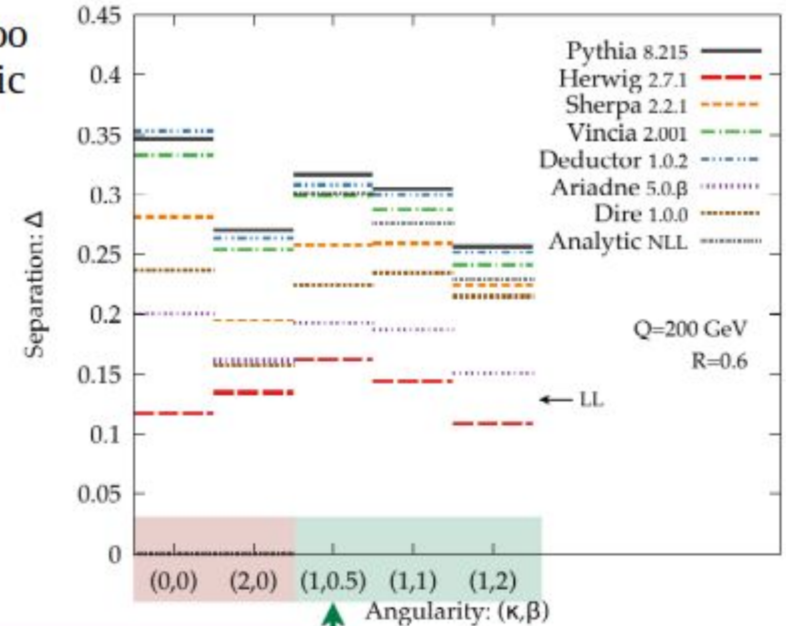
Differential



$$\mathrm{LHA} = \sum_i z_i \sqrt{\theta_i}$$

Pythia too optimistic

Integrated Values



Affects both IRC unsafe and IRC safe observables