ATLAS–CMS Monte Carlo Generators Workshop

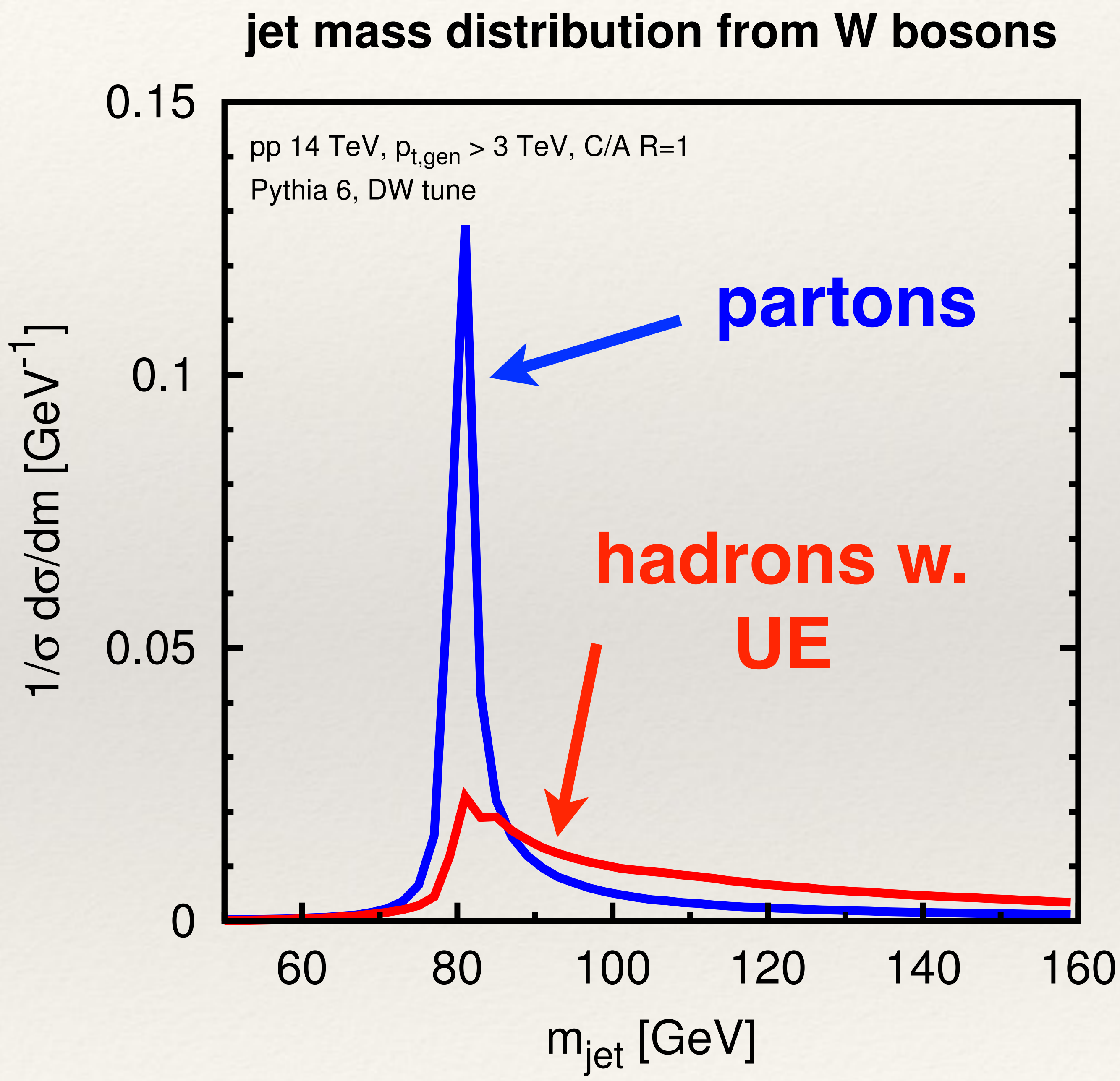
CERN, 11–12 January 2016

Theoretical status and progress of jet substructure

Gavin Salam (CERN)

Principles in use today

#1: the jet mass, a fragile observable.



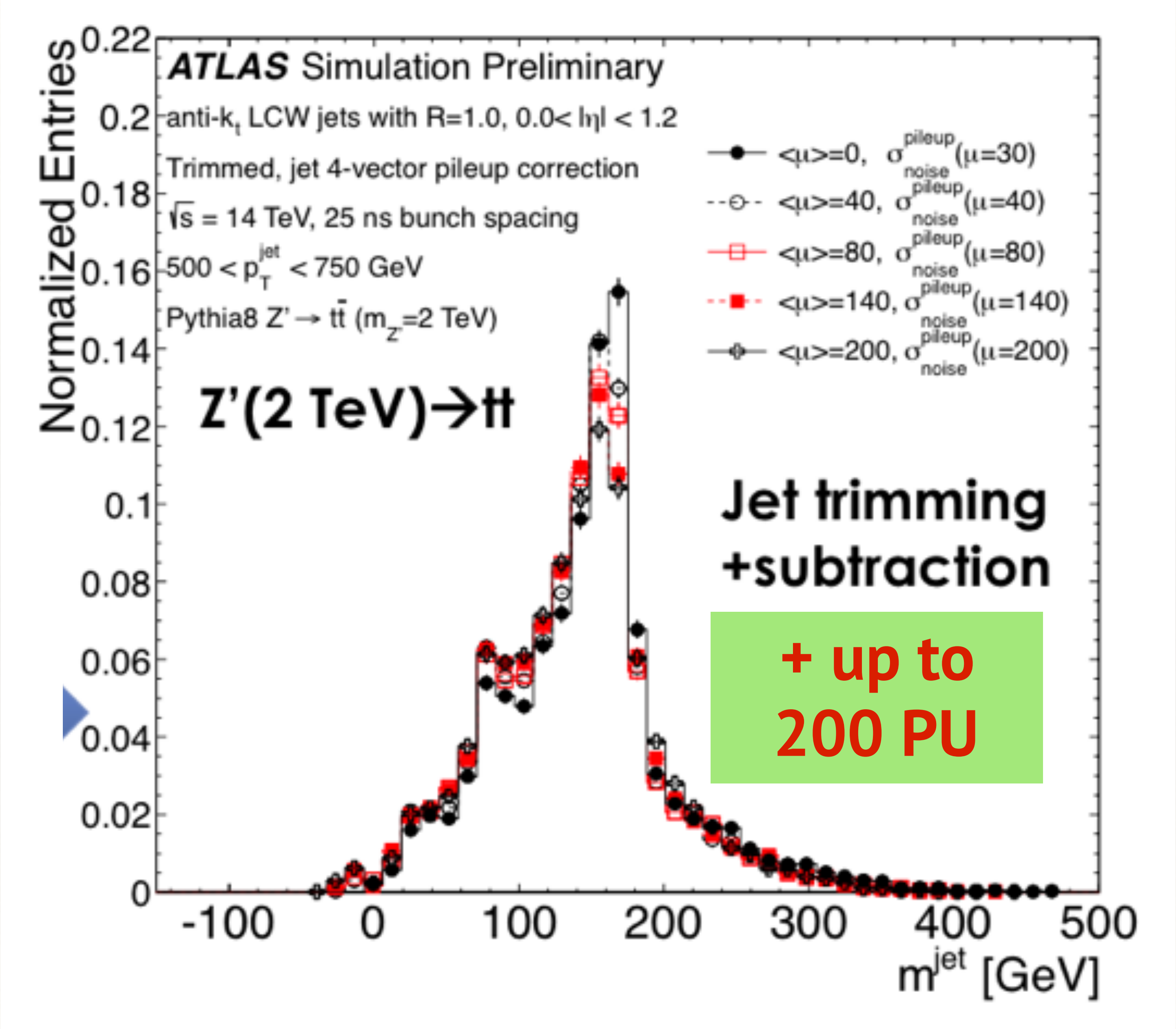
Principles in use today

#1: the jet mass, a fragile observable.

So people usually use a **groomed mass**:

filtering/trimming/pruning

(or you can go to smaller R ~ few x M/p_t)



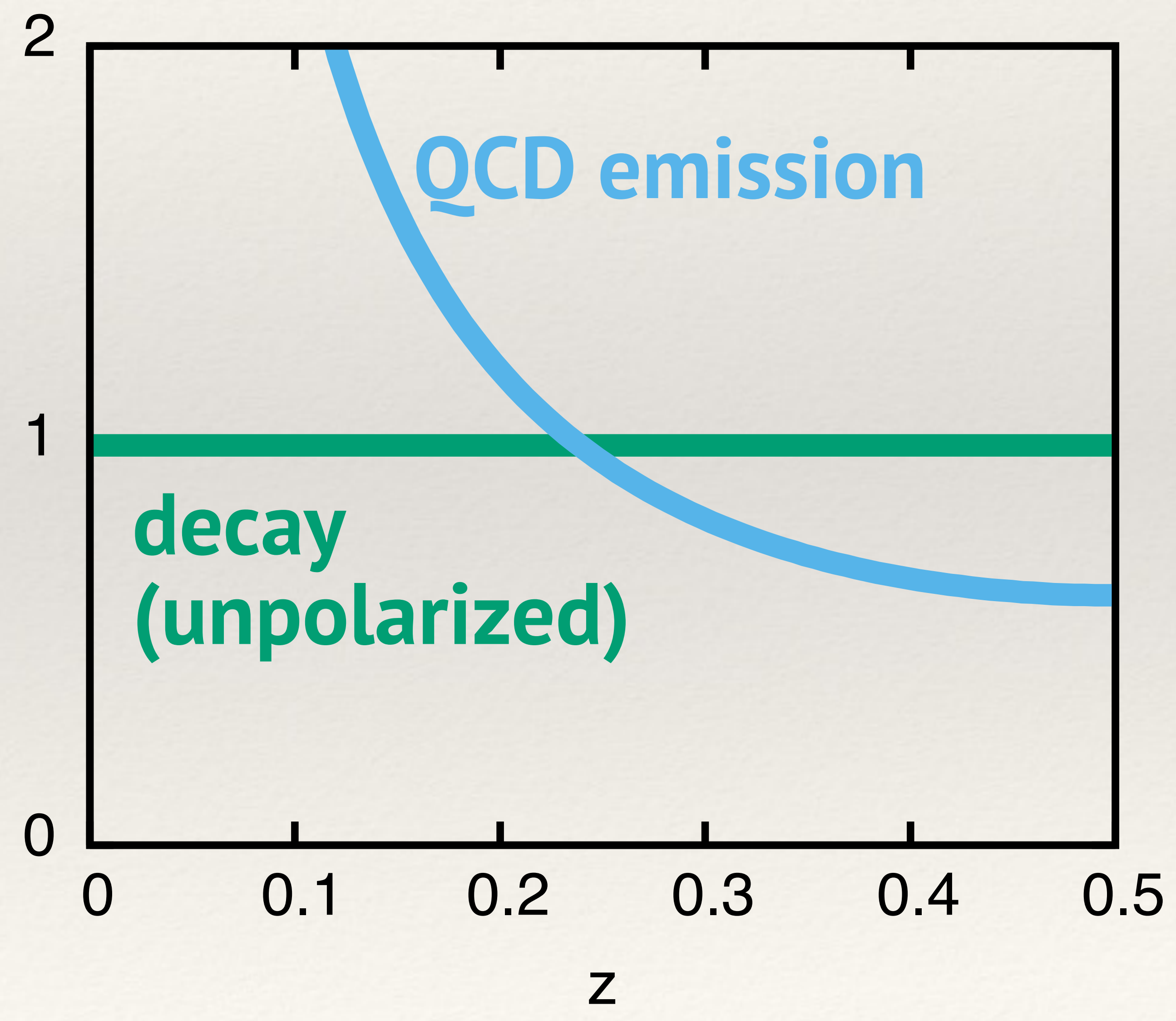
Principles in use today

#2: QCD gluon emission is soft; V/H→qq is not

Identify two-prong structure and cut on "z" (momentum fraction between prongs)

[done by mass-drop taggers/pruning/trimming/soft-drop]

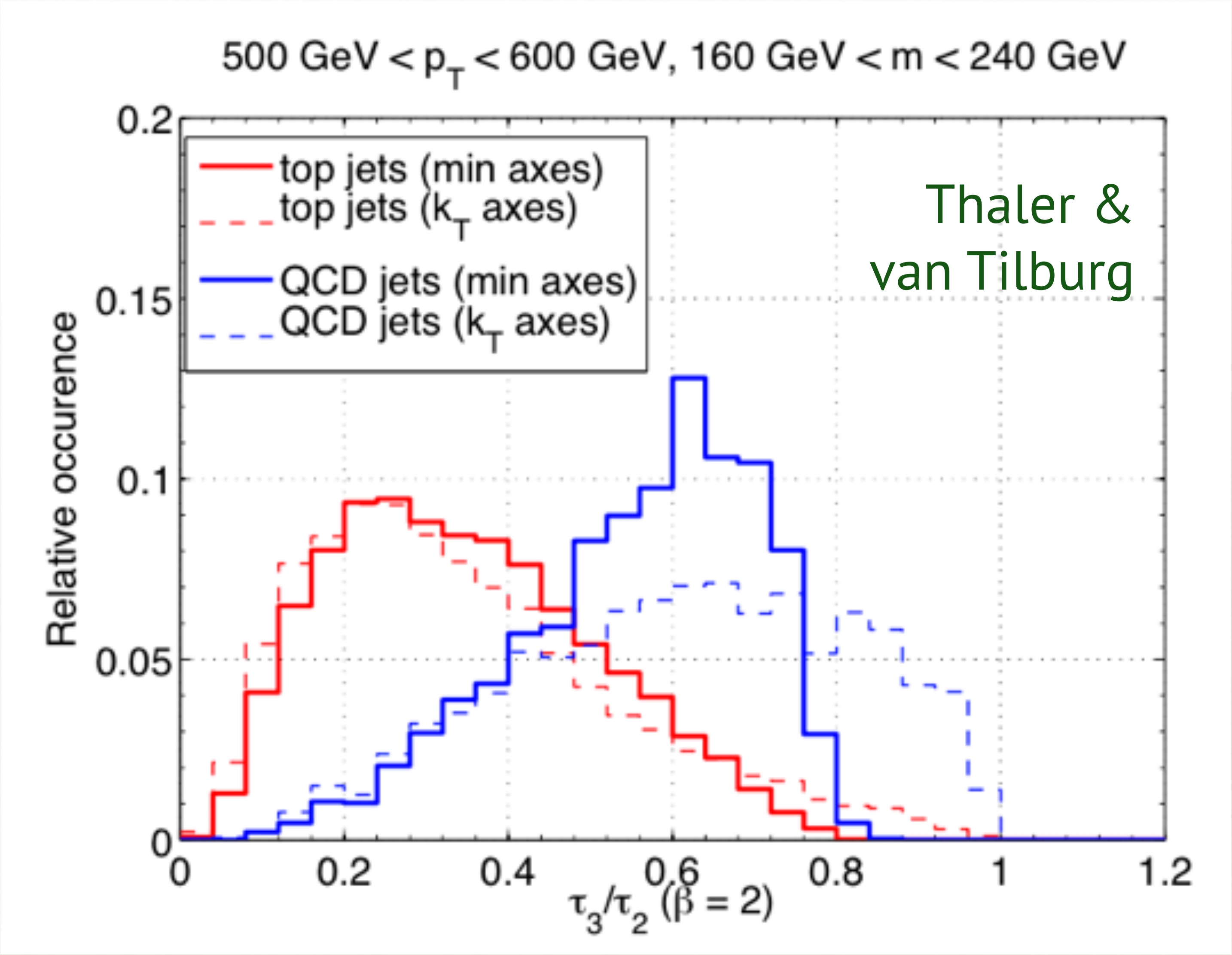
z distribution



Principles in use today

#3: Radiation patterns differ in V/H/top v. QCD

Cut on variables sensitive to deviation from exact n-prong structure, e.g. N-subjettiness

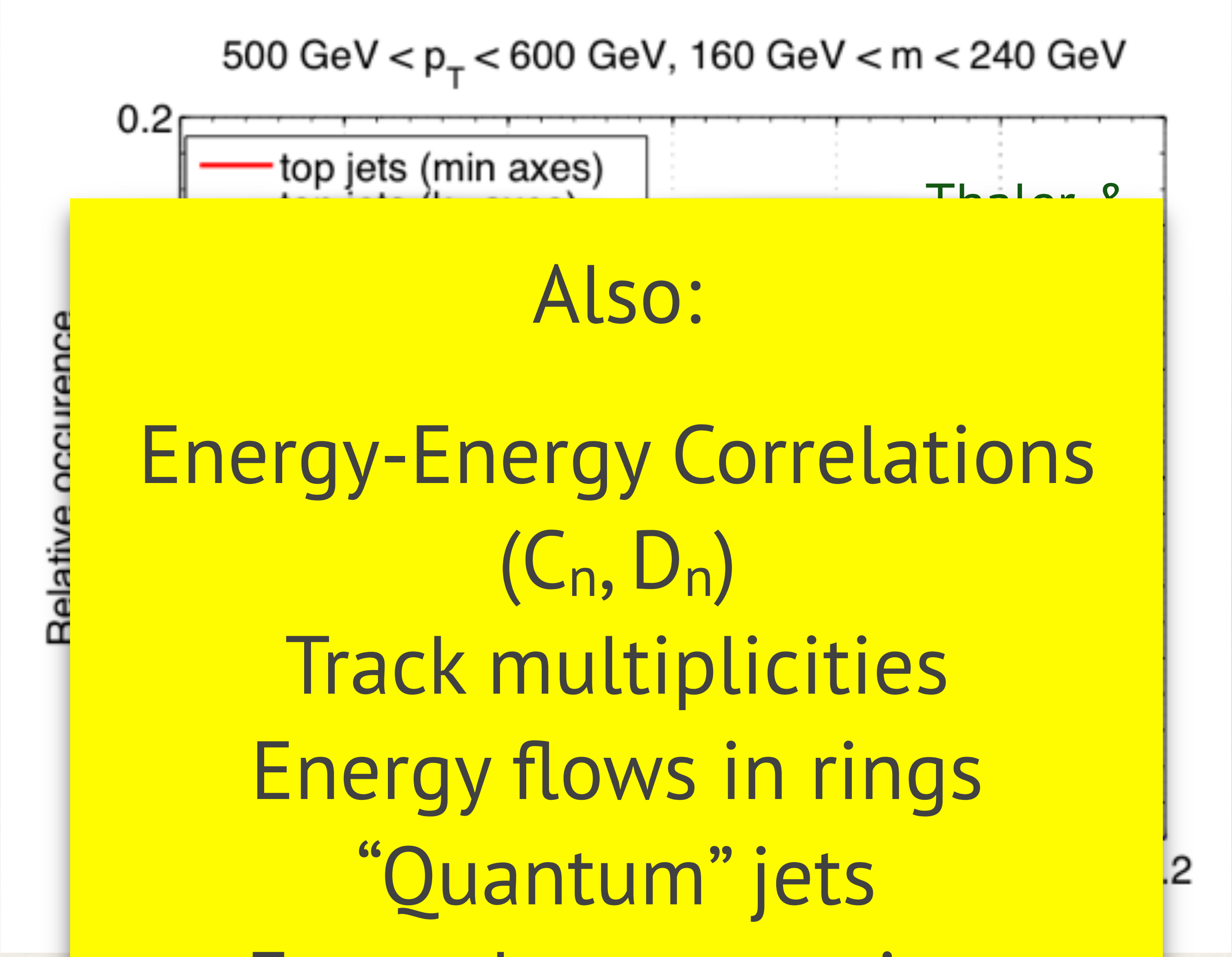


$$\frac{\tau_n}{\tau_{n-1}}; \quad \tau_n = \min_{\text{n axes}} \sum_i p_{ti} \min(\Delta R_{i,\text{axis-1}}, \ldots, \Delta R_{i,\text{axis-n}})$$

Principles in use today

#3: Radiation patterns differ in V/H/top v. QCD

Cut on variables sensitive to deviation from exact n-prong structure, e.g. N-subjettiness



Also:

Energy-Energy Correlations (C_n, D_n)
Track multiplicities
Energy flows in rings
"Quantum" jets
Event deconstruction
etc.

$$\frac{\tau_n}{\tau_{n-1}}; \qquad \tau_n = \min_{\text{n axes}} \sum_i p_{ti} \,\mathrm{m}$$

Computer-vision techniques

1511.05190, de Oliveira, Kagan, Mackey, Nachman, Schwartzman

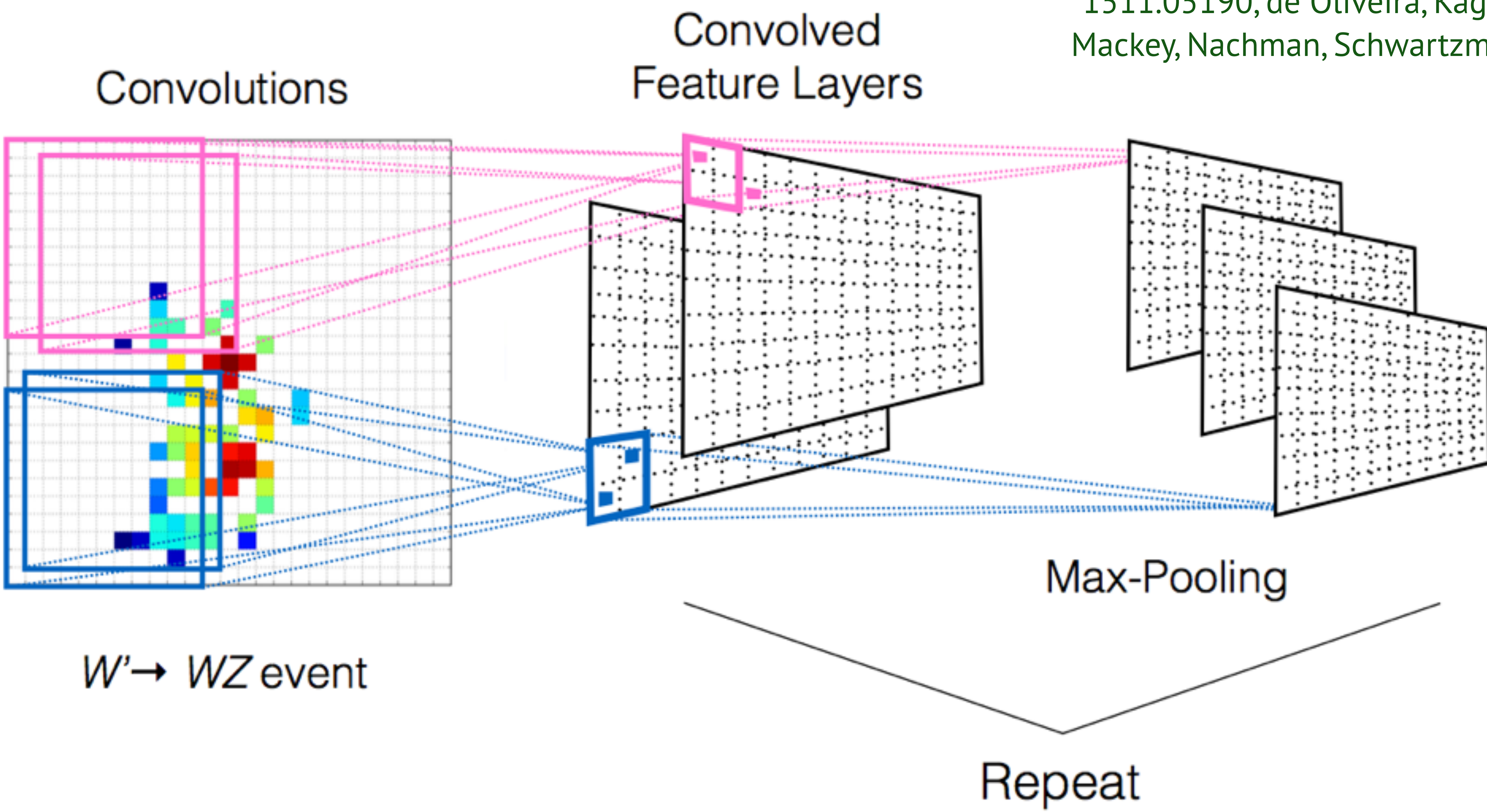
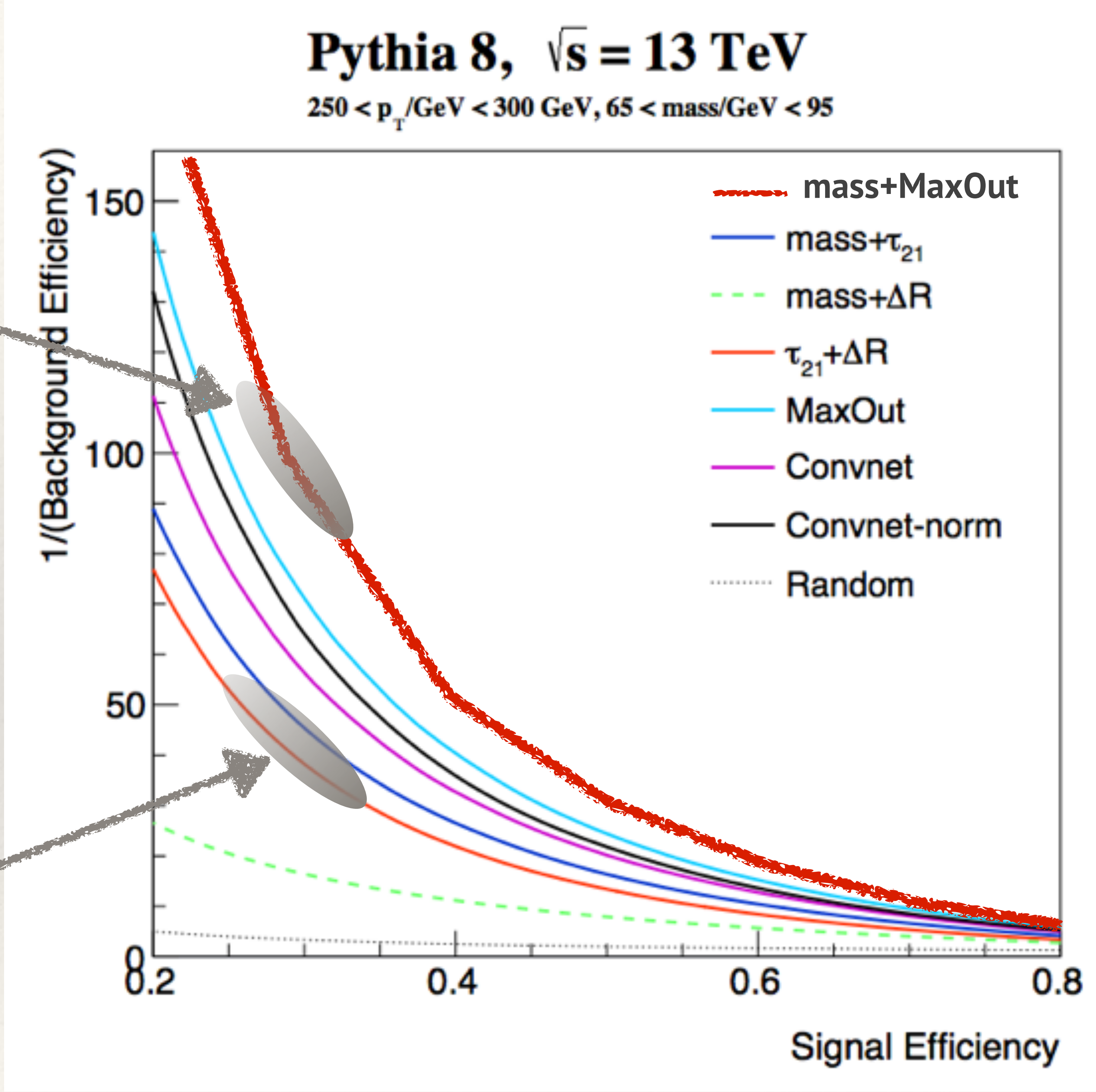


Figure 5: The convolution neural network concept as applied to jet-images.

Computer vision: W tagging

computer vision+trim:
x100 rejection
for 30% efficiency;
x2 better rejection
than "standard"

standard techniques
(trimming & tau21)



adapted from 1511.05190, de Oliveira, Kagan, Mackey, Nachman, Schwartzman

What theory aims?

- Develop more powerful methods for discriminating signal/background
- Understand what physics various "taggers" are actually tagging on:
  - to know whether it's reliably modeled by MCs
  - to know what "features" tagging might induce in data
  - as a guide to developing better tools & for predicting signals & backgrounds

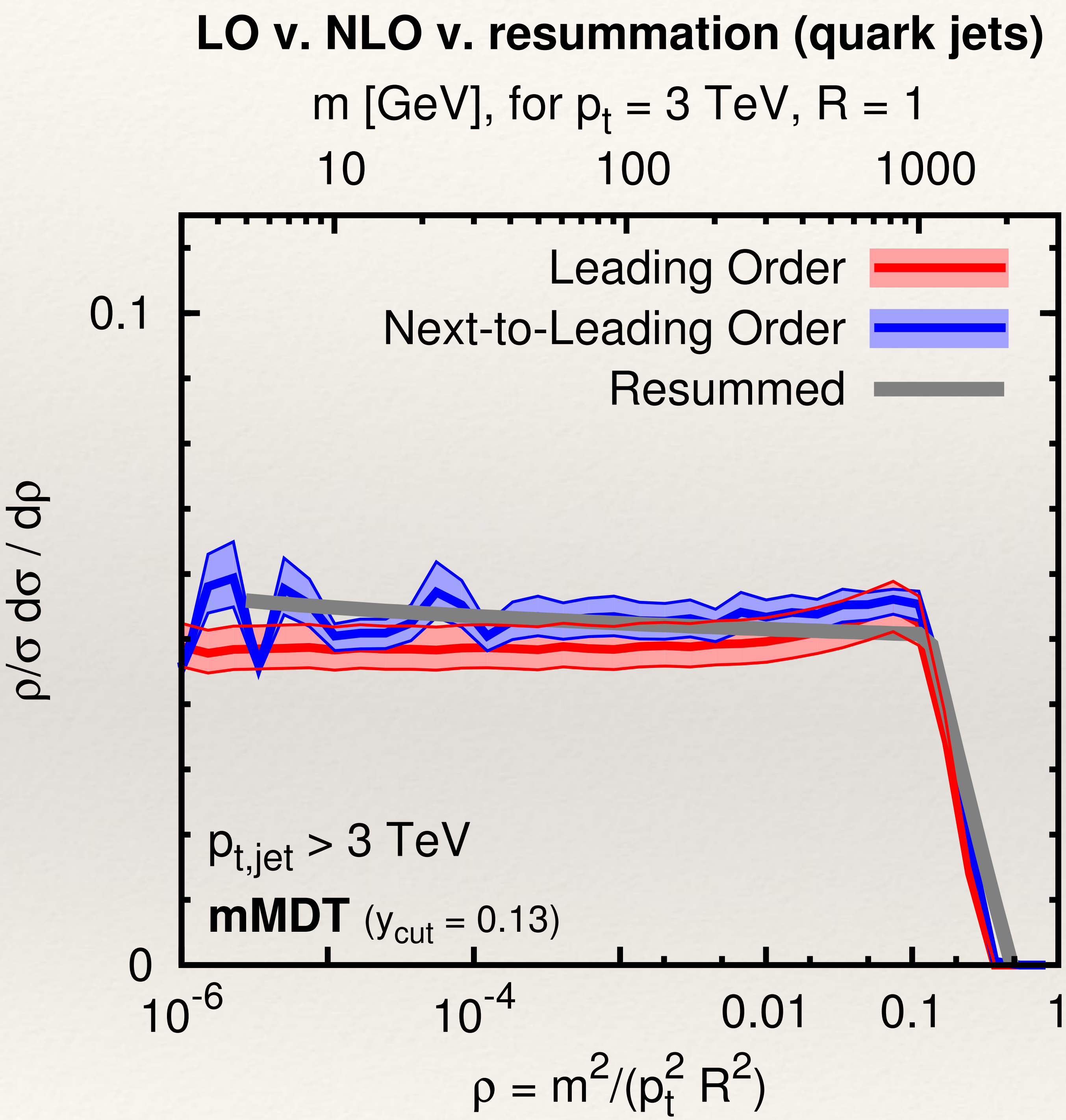
Event generators play key role in testing methods
& theory calculations may teach us things about event generators

Core aspects

- There's always a large ratio of scales, $\mathbf{p_T/m \gg 1}$
- This is the parton shower regime (so most MC comparisons performed w/o merging).
- Some substructure observables (**mMDT, Soft-Drop**[β=0]) special because they pull out **just the hard substructure:**
  - involve only single (collinear) logs of p_T/m
  - not the most discriminating, but probably most robust
  - small hadronisation/UE effects

1307.0007, Dasgupta, Fregoso, Marzani, GPS
1402.2657, Larkoski, Marzani , Soyez & Thaler

mMDT resummation v. fixed order



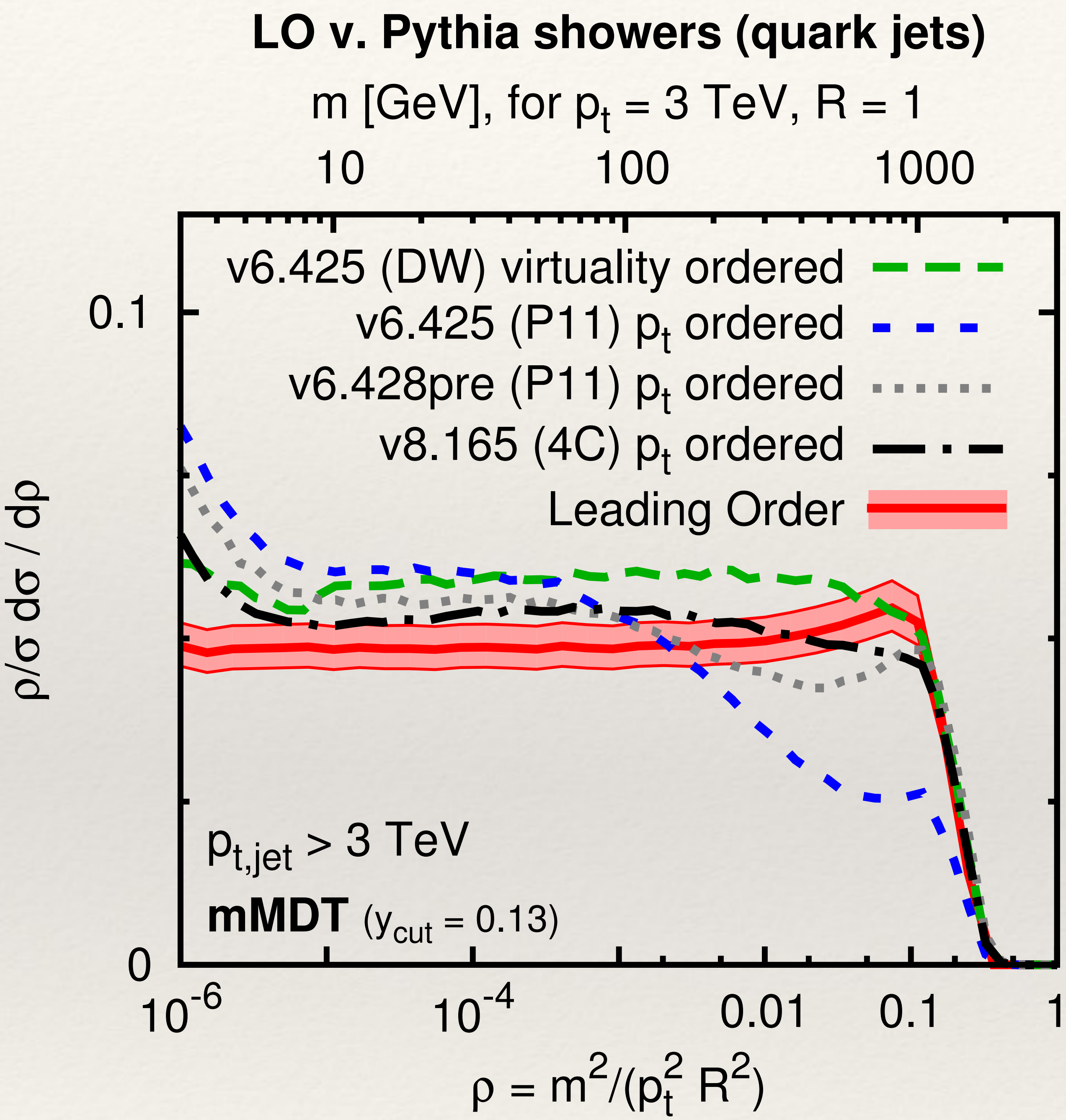
NLO from NLOJet++

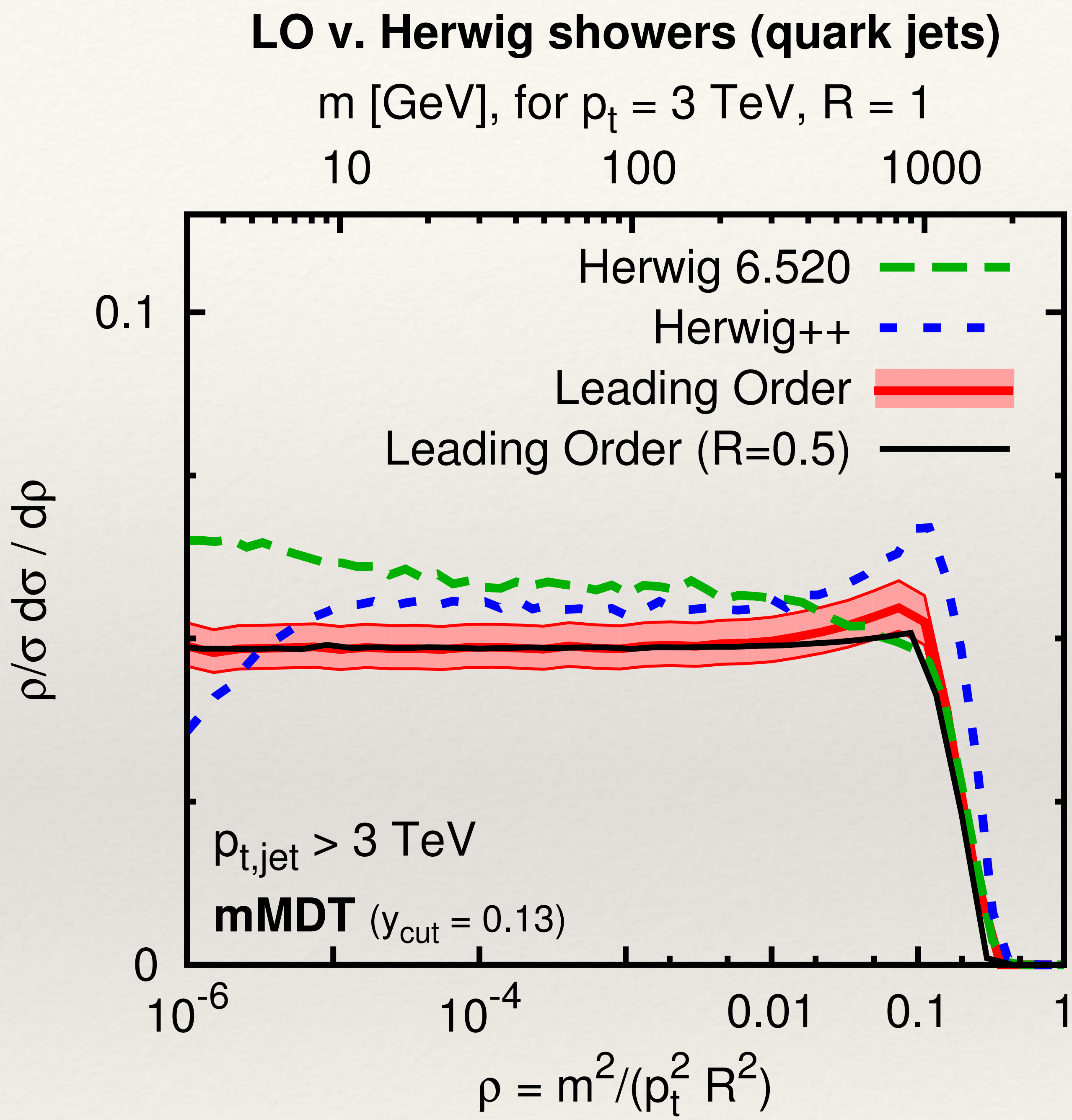
It only has single (collinear) logs → fixed-order is stable over a broader than usual range of scales

(helped by fortuitous cancellation between running coupling and single-log Sudakov)

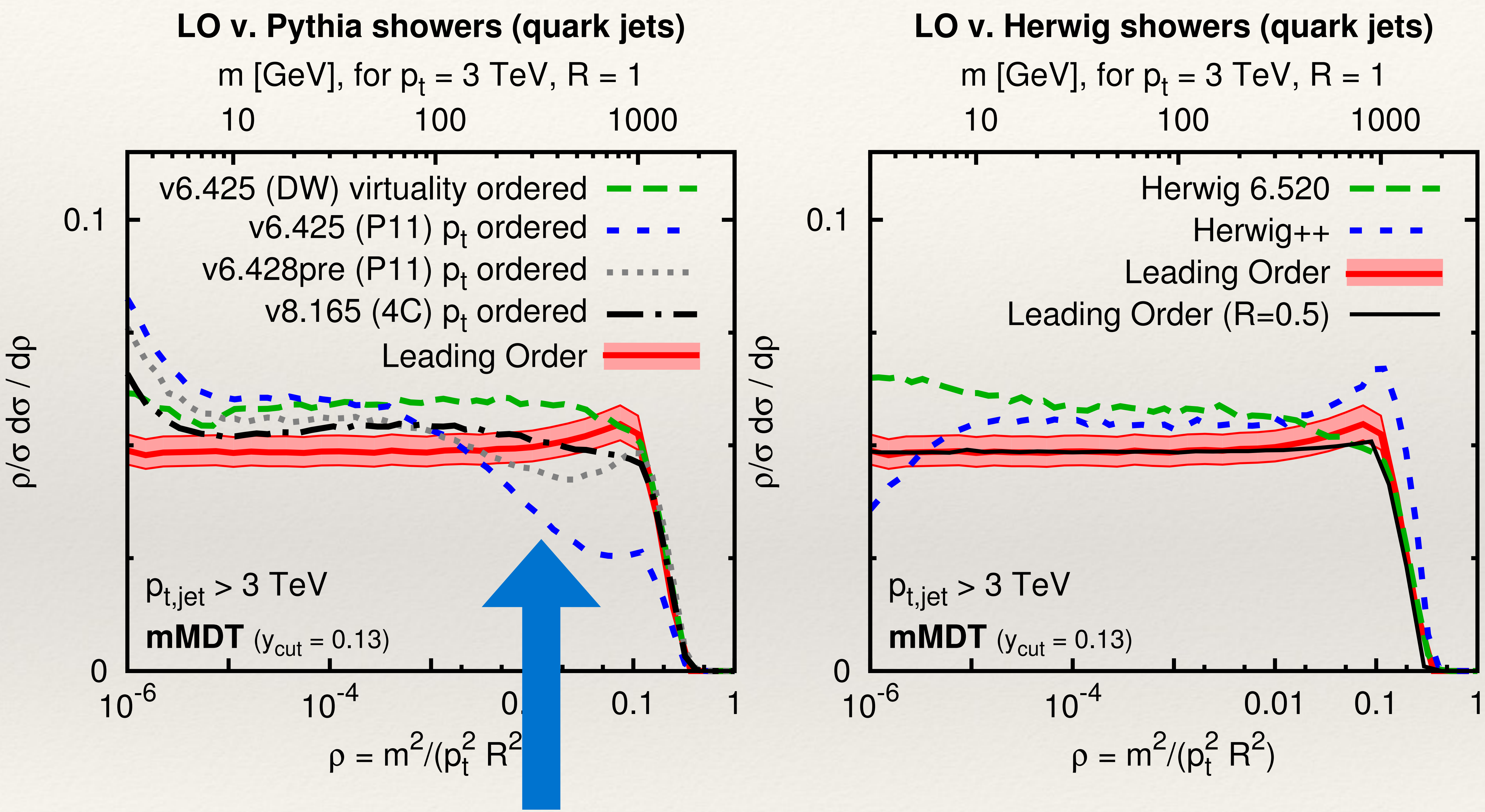
1307.0007, Dasgupta, Fregoso, Marzani, GPS

mMDT: comparing many showers



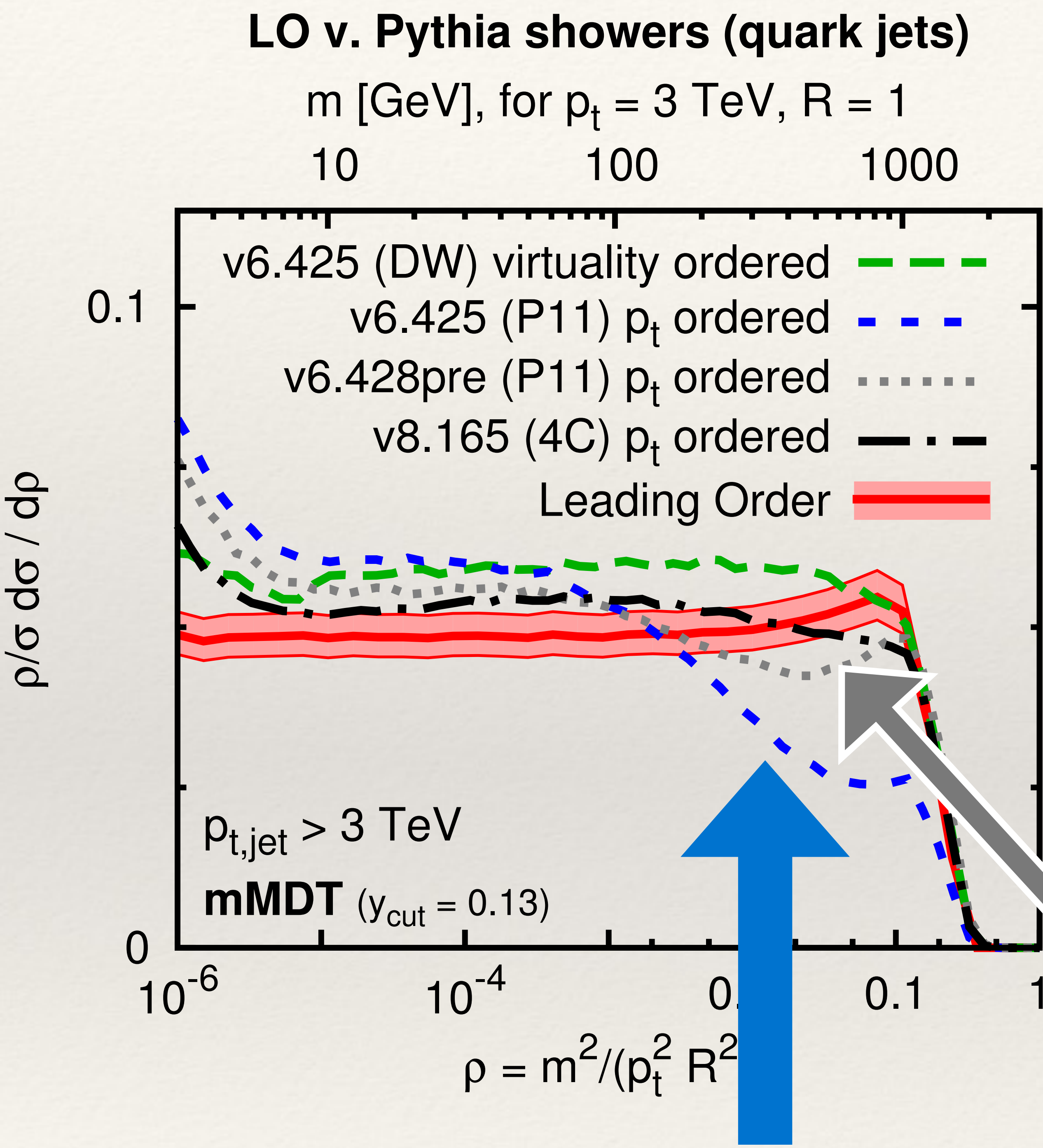


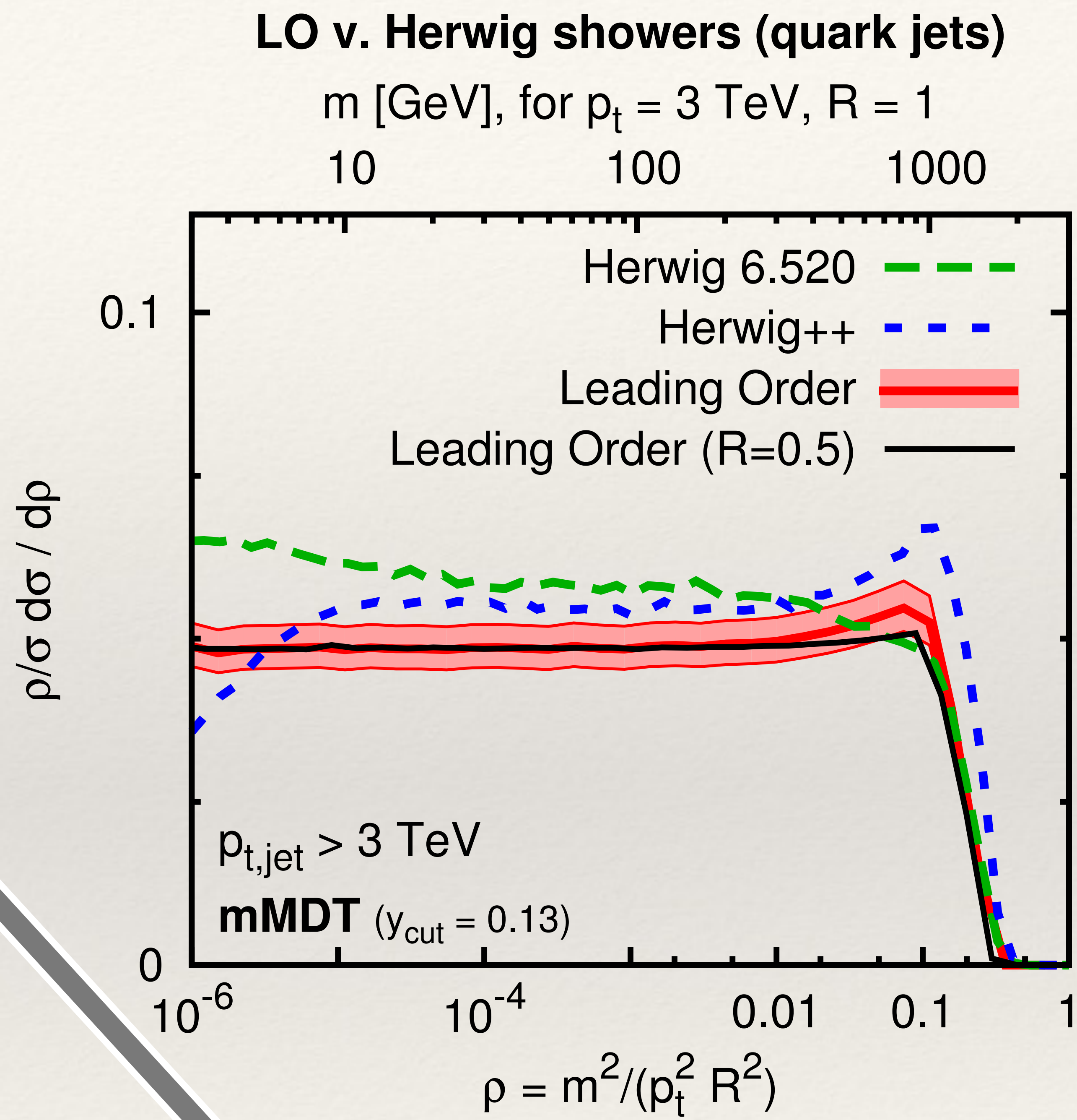
mMDT: comparing many showers



Issue found in Pythia 6 p_t-ordered shower → promptly identified and fixed by Pythia authors!

mMDT: comparing many showers





Issue found in Pythia 6 p_t-ordered shower → promptly identified and fixed by Pythia authors!

Better discrimination?

- Usually involves exploiting different **double log** radiation structures of signal & background, e.g.
  - quark v. gluon prongs
  - singlet v. triplet/octet colour structures
- Best observables are often relatively complex, with multiple parameters (e.g. angular exponent), switches (axis choice)
- First few steps in understanding have been undertaken in past year

Phase space: Lund plane

1512.00516, Dasgupta, Schunk, Soyez

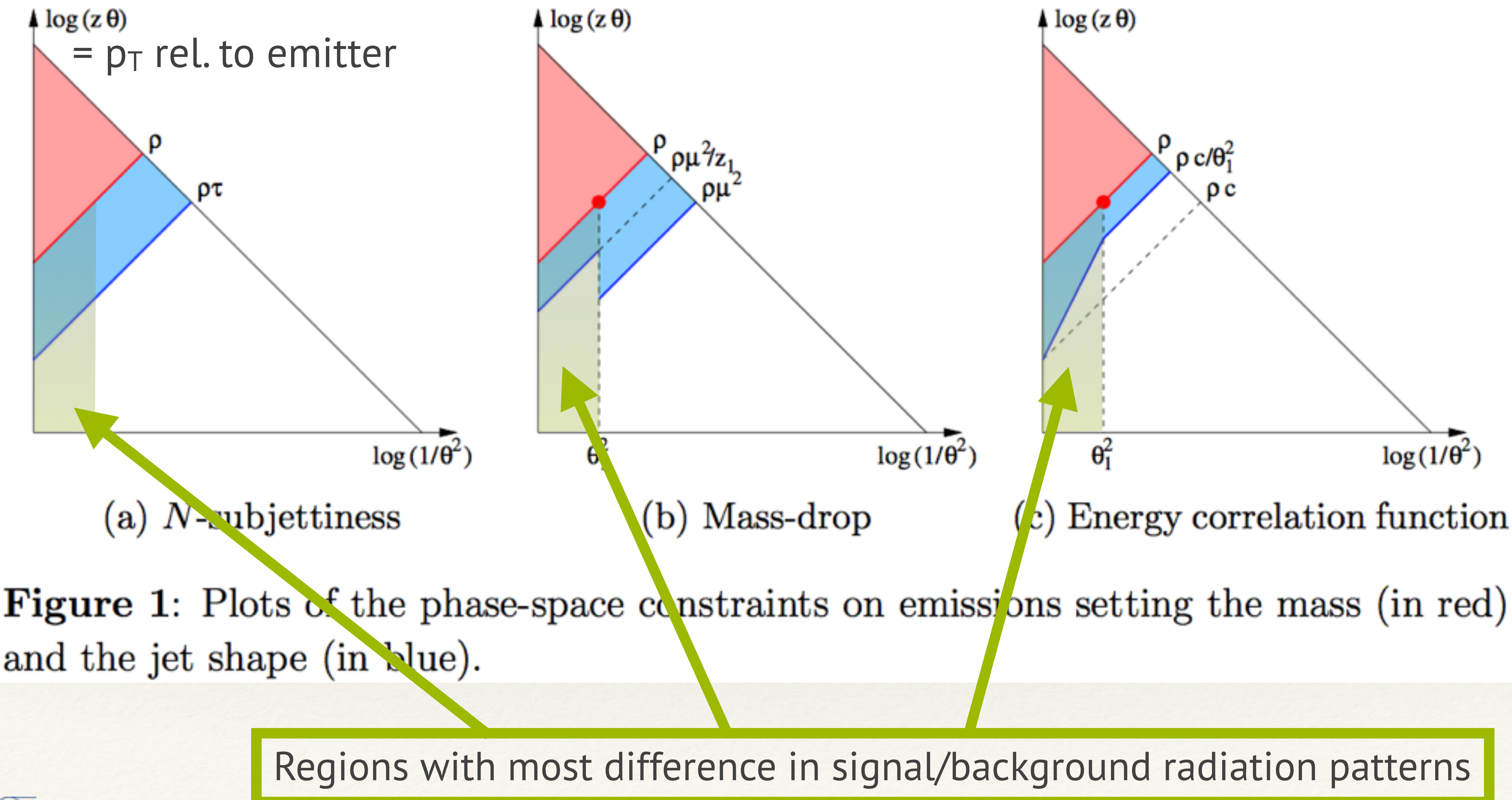
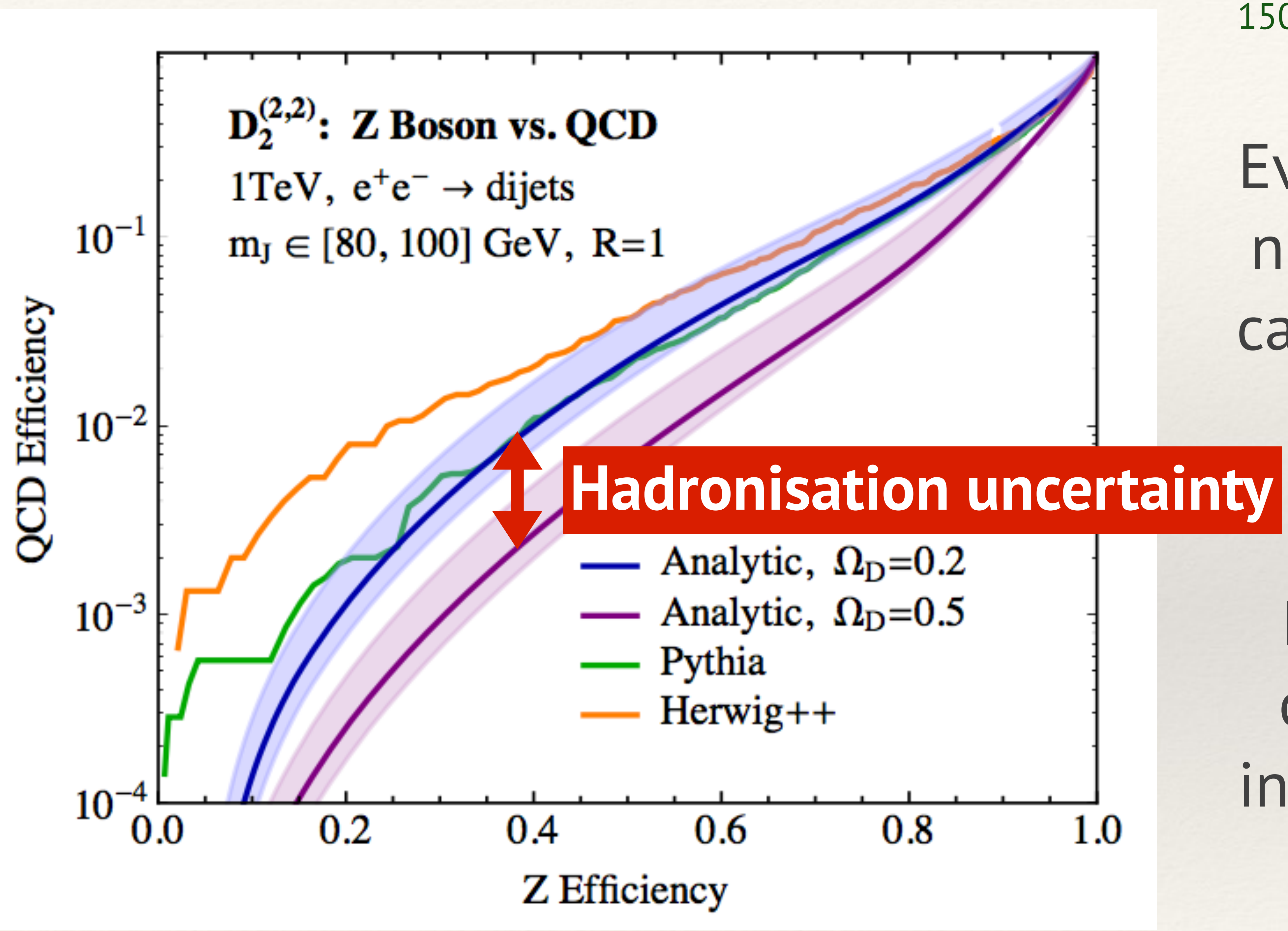


Figure 1: Plots of the phase-space constraints on emissions setting the mass (in red) and the jet shape (in blue).

non-perturbative effects

1507.03018, Larkoski, Moult, Neill



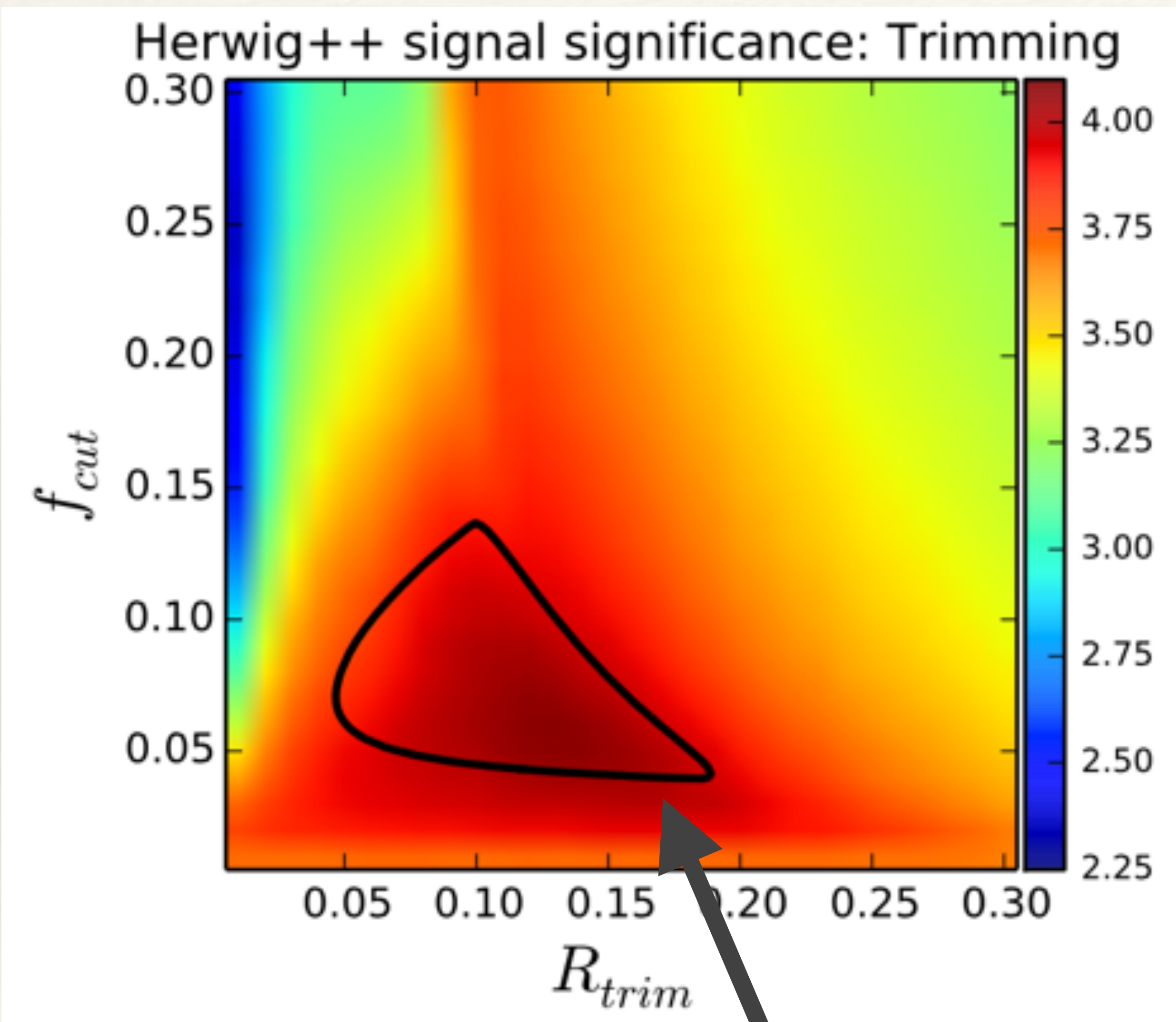
Even at TeV scale, non-pert. effects can be significant

Reason is that one is probing internal structure to O(few GeV)

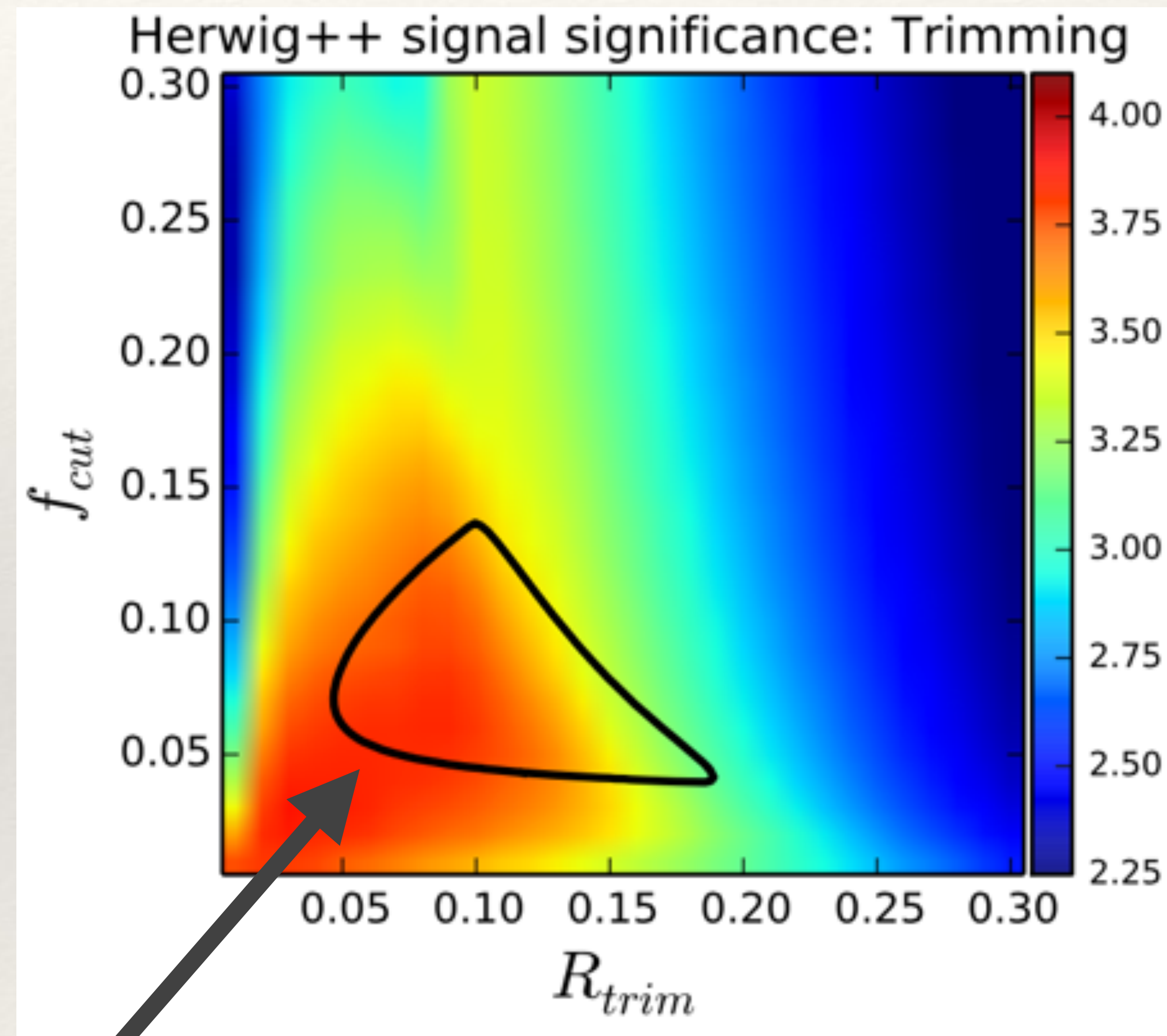
non-perturbative effects

1503.01088, Dasgupta, Powling, Siodmok

Parton level



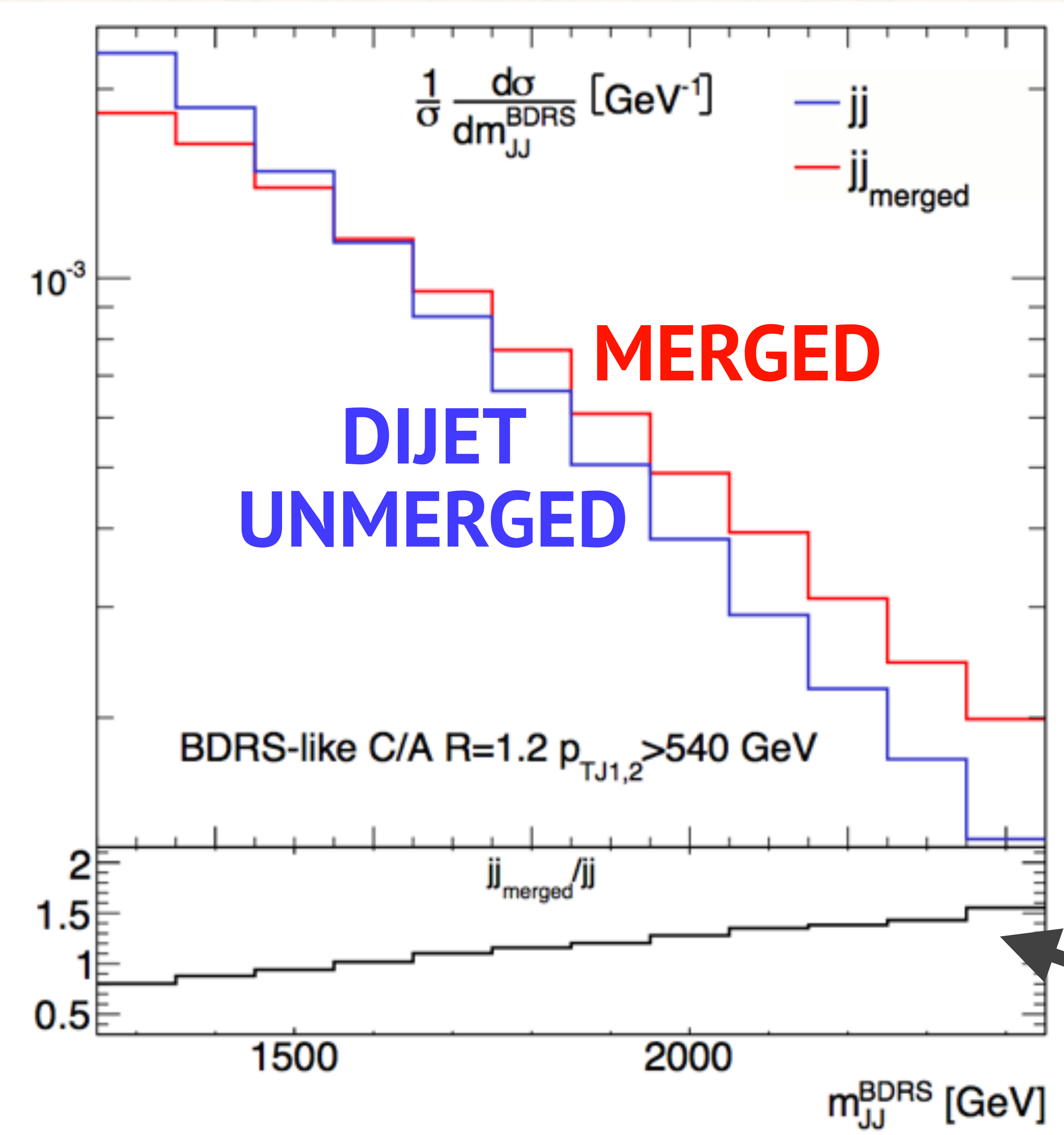
Hadron level



Analytic (parton-level) prediction for optimal parameters

Questions / issues for MC?

Multijet merging?



1508.04162, Gonçalves, Krauss, Spannowsky

Multijet merging & NLOPS seldom used.

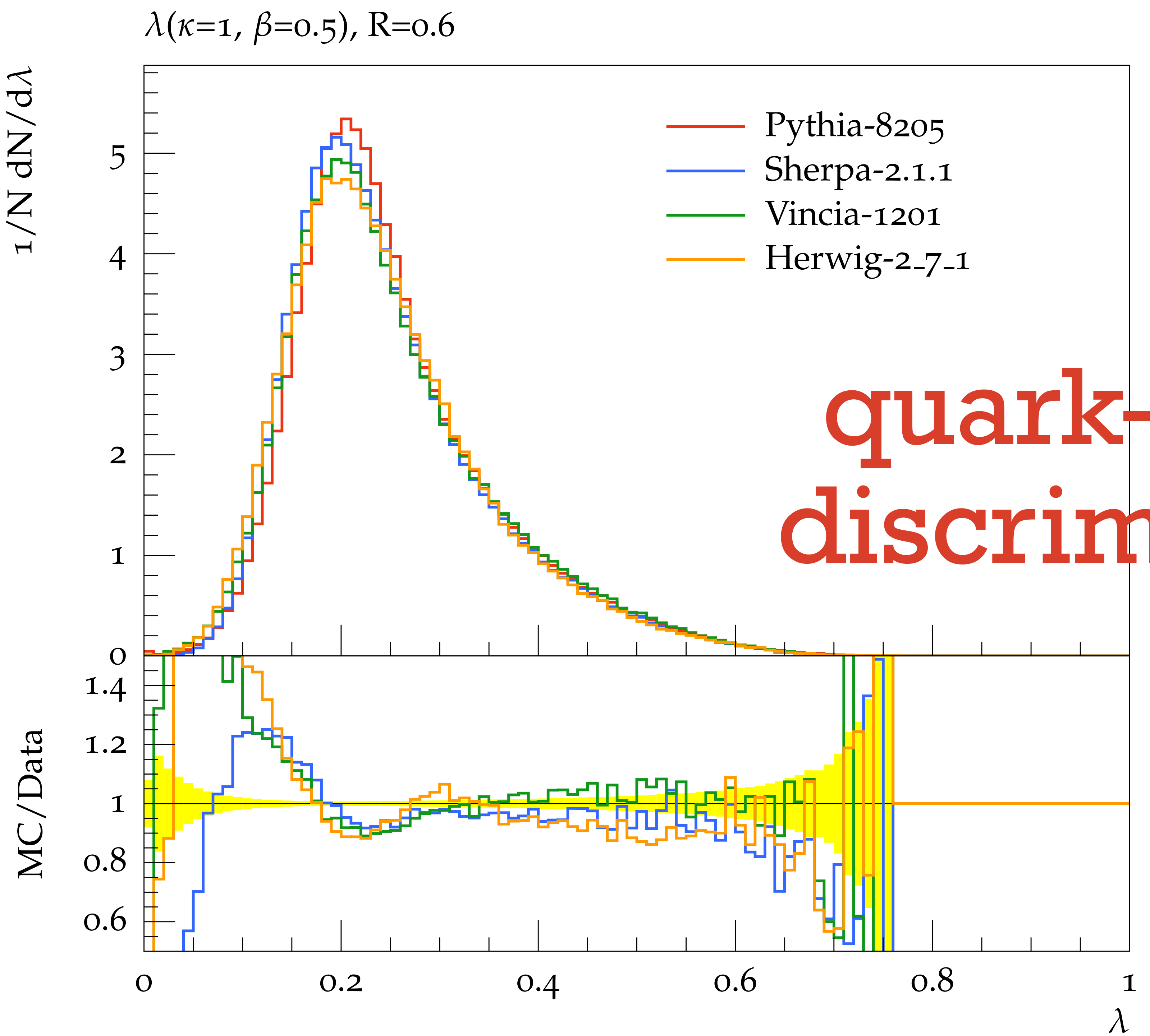
1) often more cumbersome than at low p_T

2) should not matter much since most structure is in jet core, especially at high p_T

Somewhat surprising to see 50% effect at highest p_T – **deserves further study**

Les Houches Angularity: Quarks

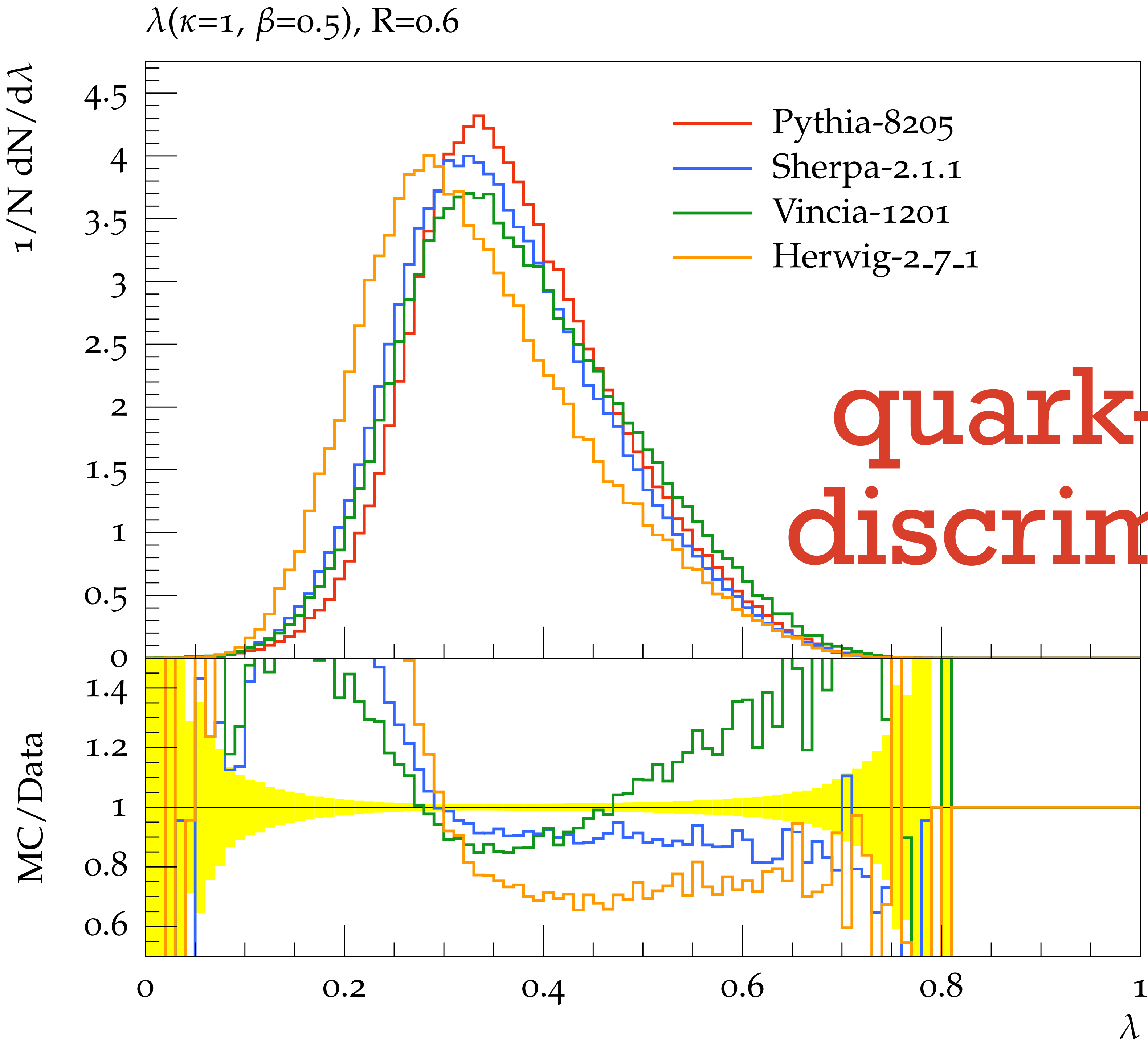
Hadron level, R=0.6, Q=200 GeV



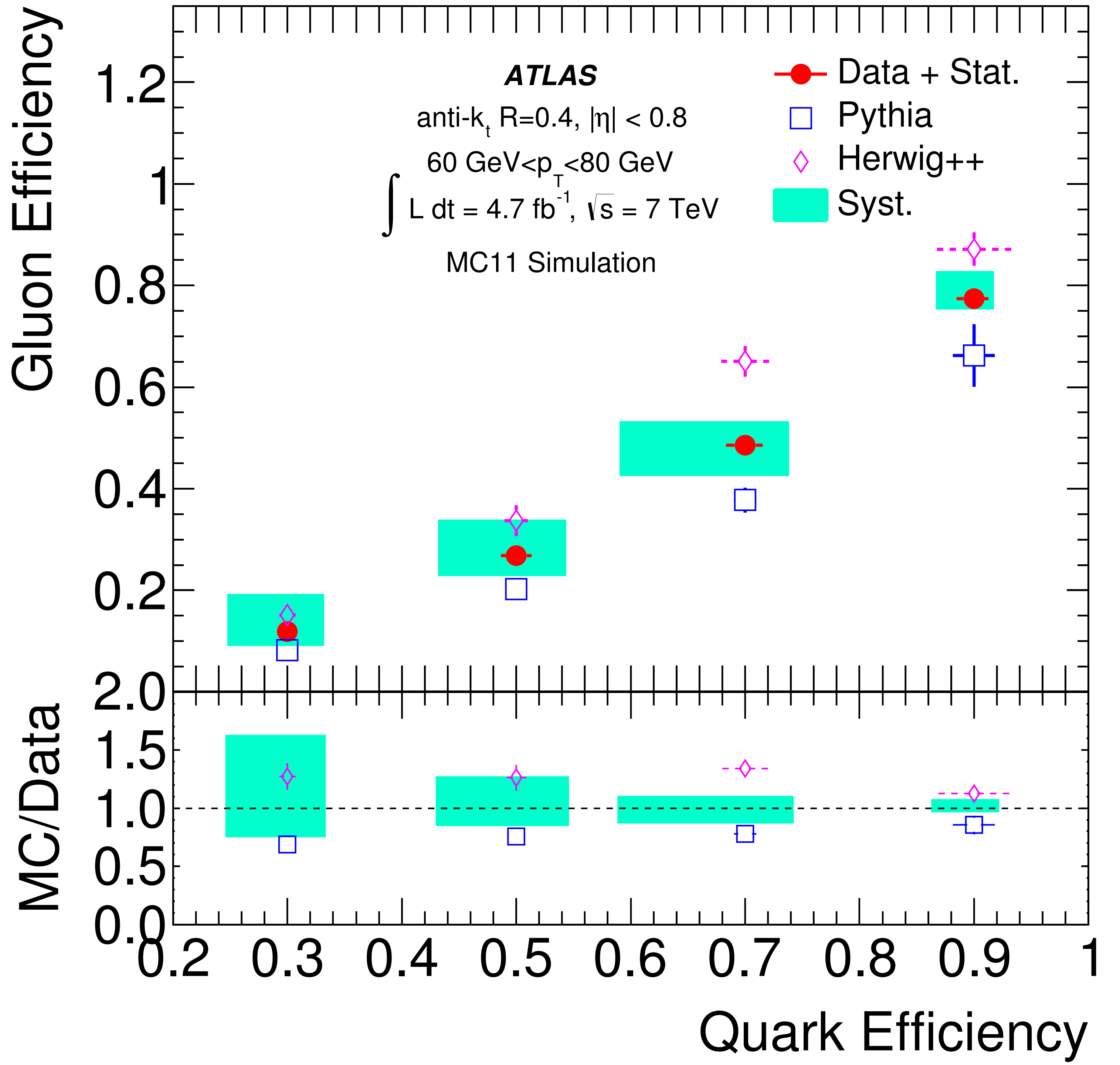
quark–gluon discrimination

Les Houches Angularity: Gluons

Hadron level, R=0.6, Q=200 GeV



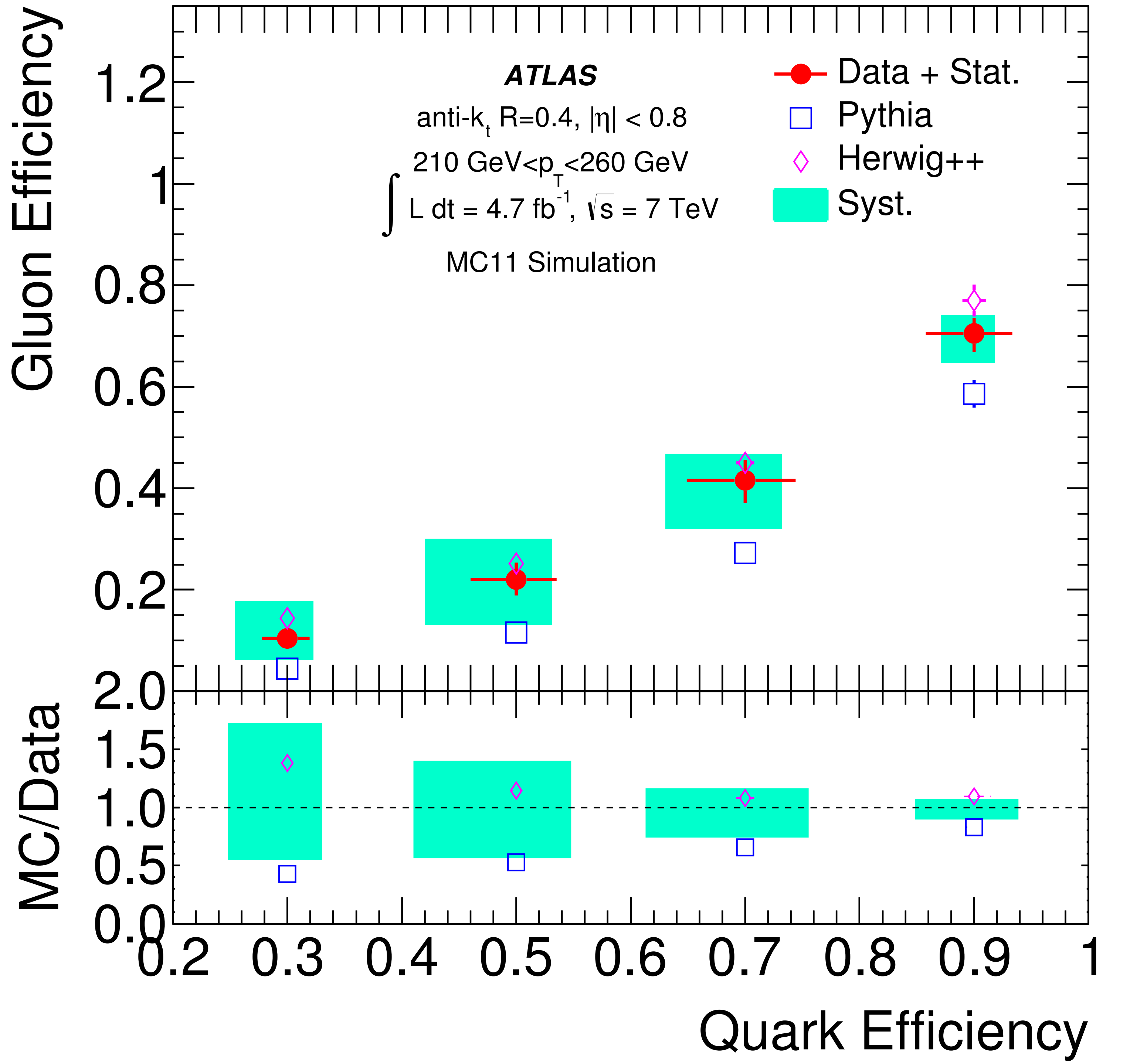
quark–gluon discrimination



Data v. MC

quark-gluon discrimination performance (using track multiplicity & jet width)

"The ability of the tagger to reject gluons at a fixed quark efficiency **is up to a factor of two better in Pythia 6 and up to 50% worse in Herwig++** than in data." [arXiv:1405.6583]



Data v. MC

quark-gluon discrimination performance (using track multiplicity & jet width)

"The ability of the tagger to reject gluons at a fixed quark efficiency **is up to a factor of two better in Pythia 6 and up to 50% worse in Herwig++** than in data." [arXiv:1405.6583]

Conclusions

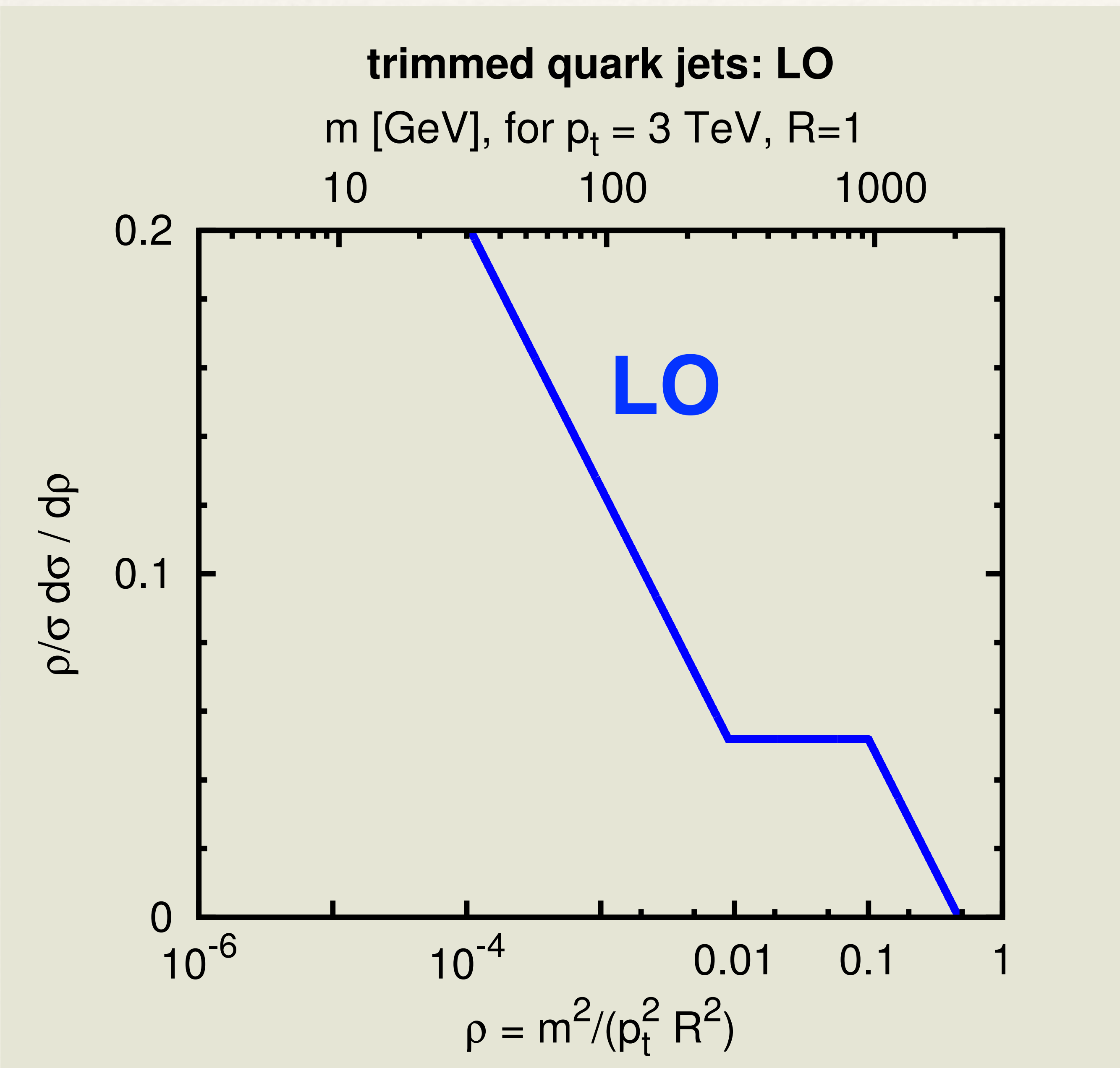
- There is an ongoing transition from "trial and error" to deeper (analytical) understanding of substructure tools, and "maximally-powerful" multivariate techniques
- MCs are essential for guidance, validation and practical applications
- MC-related questions deserving more study:
  - impact & applicability of merging/matching
  - quark/gluon discrimination aspects matter and need to be brought under better control (are there similar questions for singlet/octet discrimination?)

with thanks to
Mrinal Dasgupta,
Gregory Soyez
& Jesse Thaler
for input

Backup

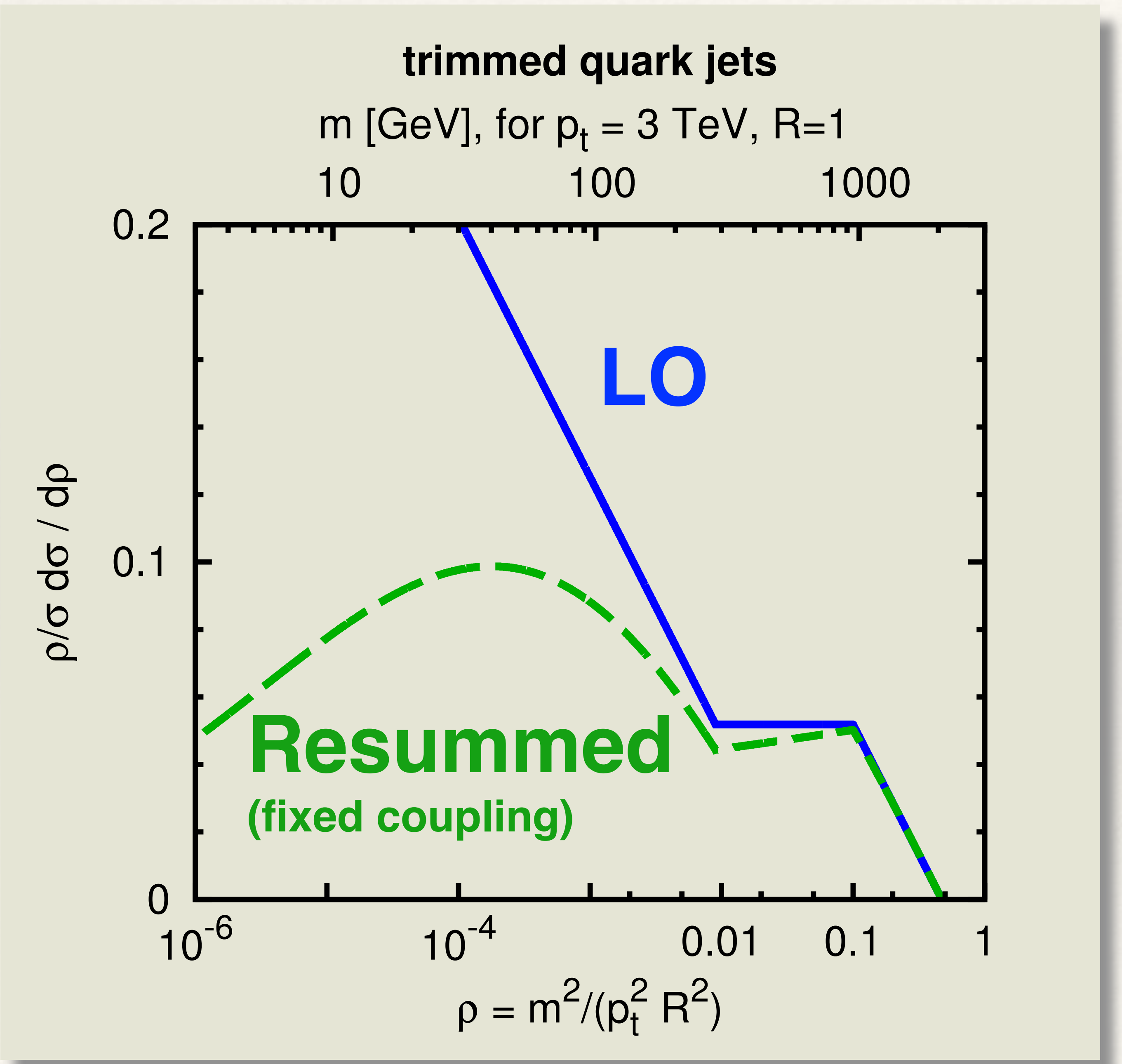
Trimming at all orders

$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho}$$



Trimming at all orders

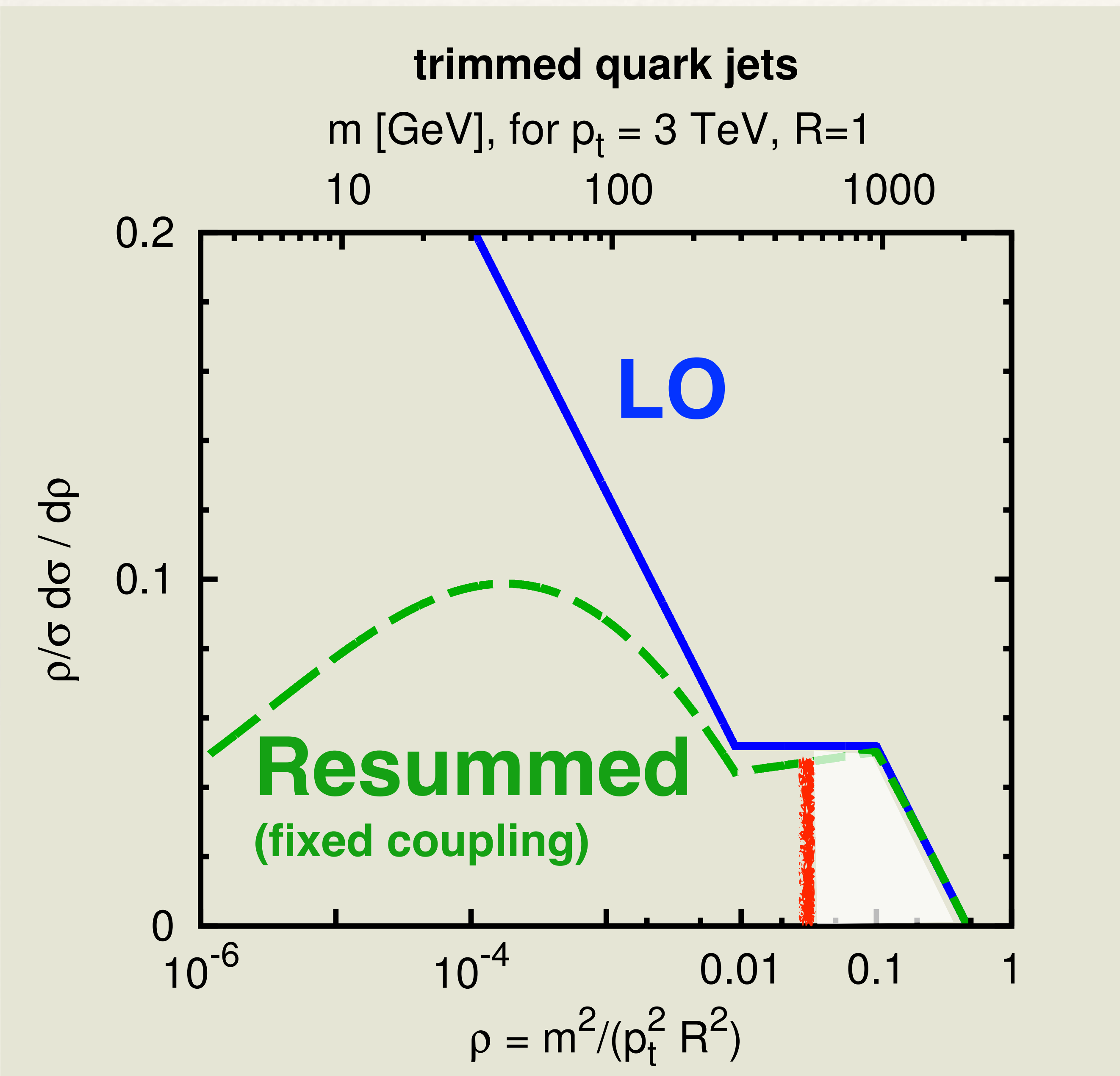
$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho} \exp\left[-\int_\rho d\rho' \frac{1}{\sigma}\frac{d\sigma^{\text{trim,LO}}}{d\rho'}\right]$$

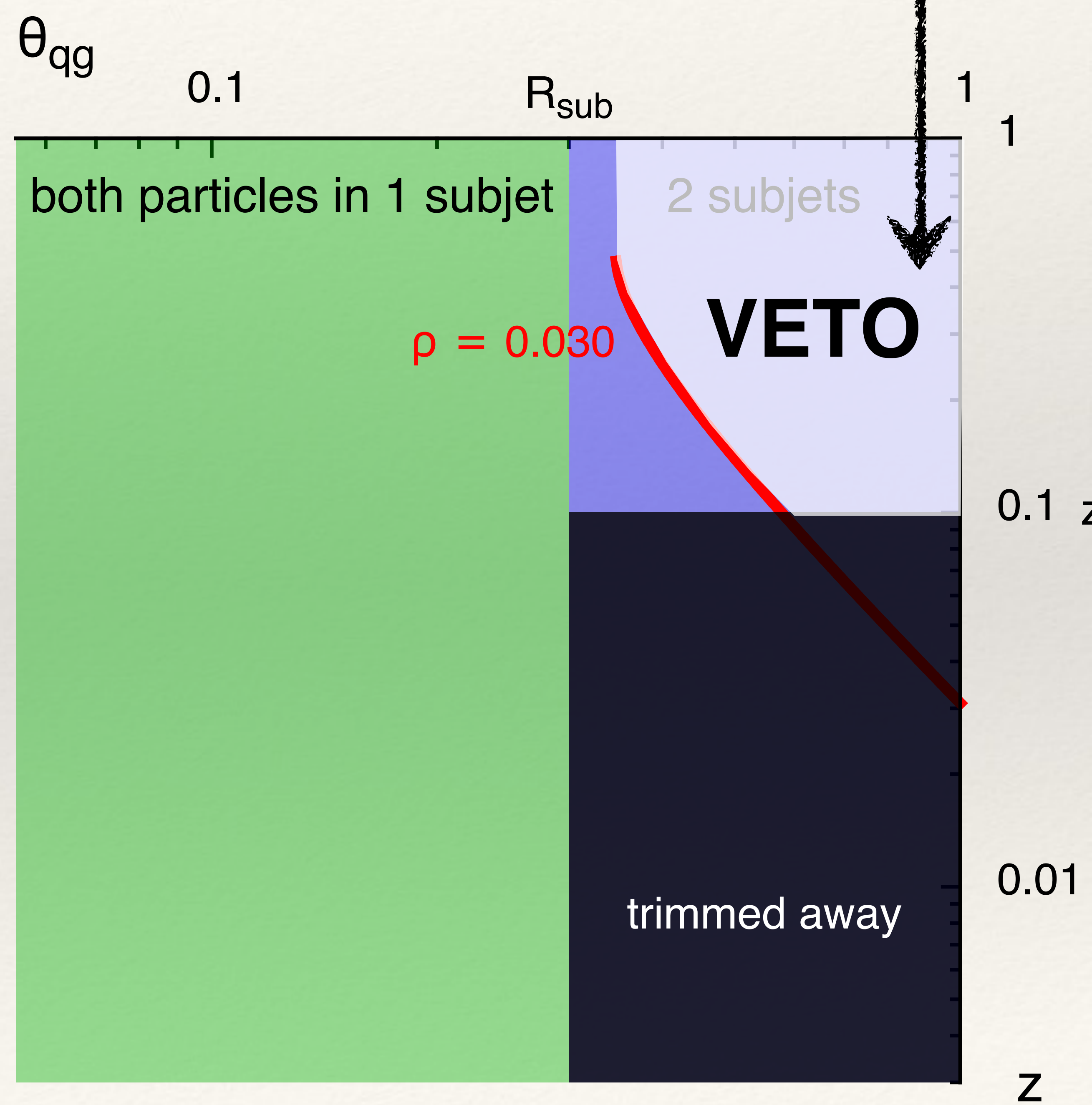


Trimming at all orders

$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho} \exp\left[-\int_\rho d\rho' \frac{1}{\sigma}\frac{d\sigma^{\text{trim,LO}}}{d\rho'}\right]$$

Sudakov

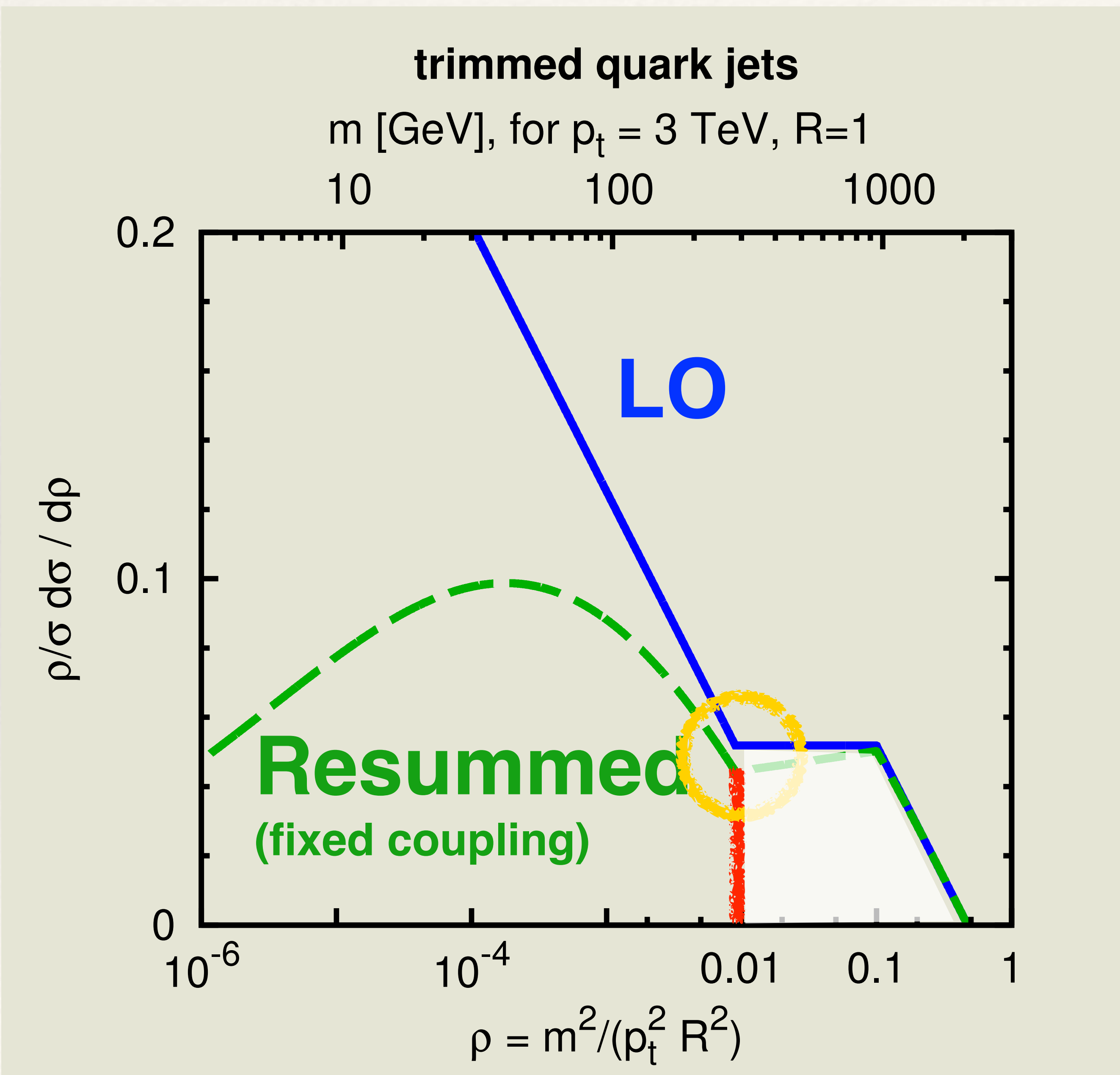


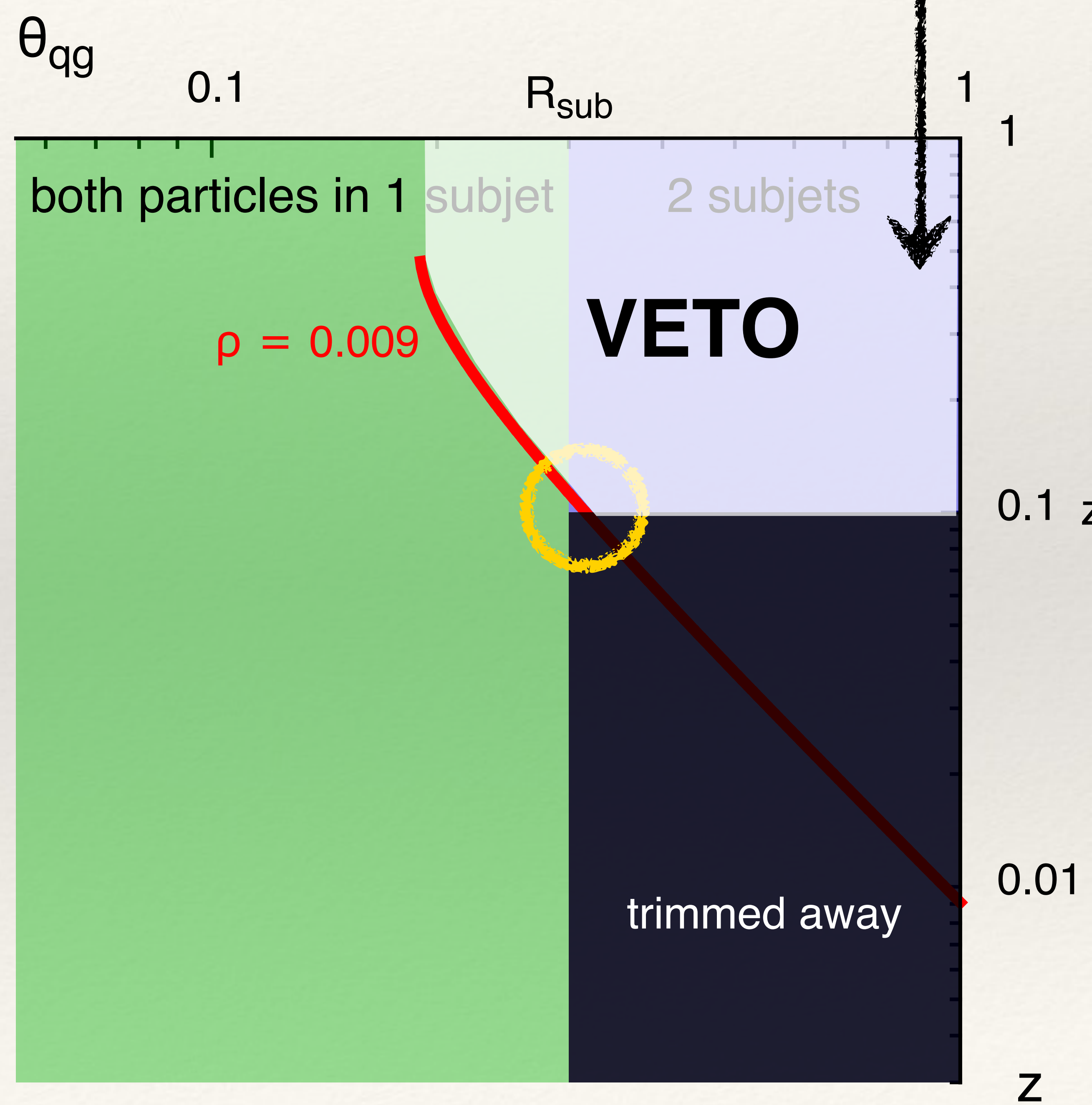


Trimming at all orders

$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho} \exp\left[-\int_\rho d\rho' \frac{1}{\sigma}\frac{d\sigma^{\text{trim,LO}}}{d\rho'}\right]$$

Sudakov

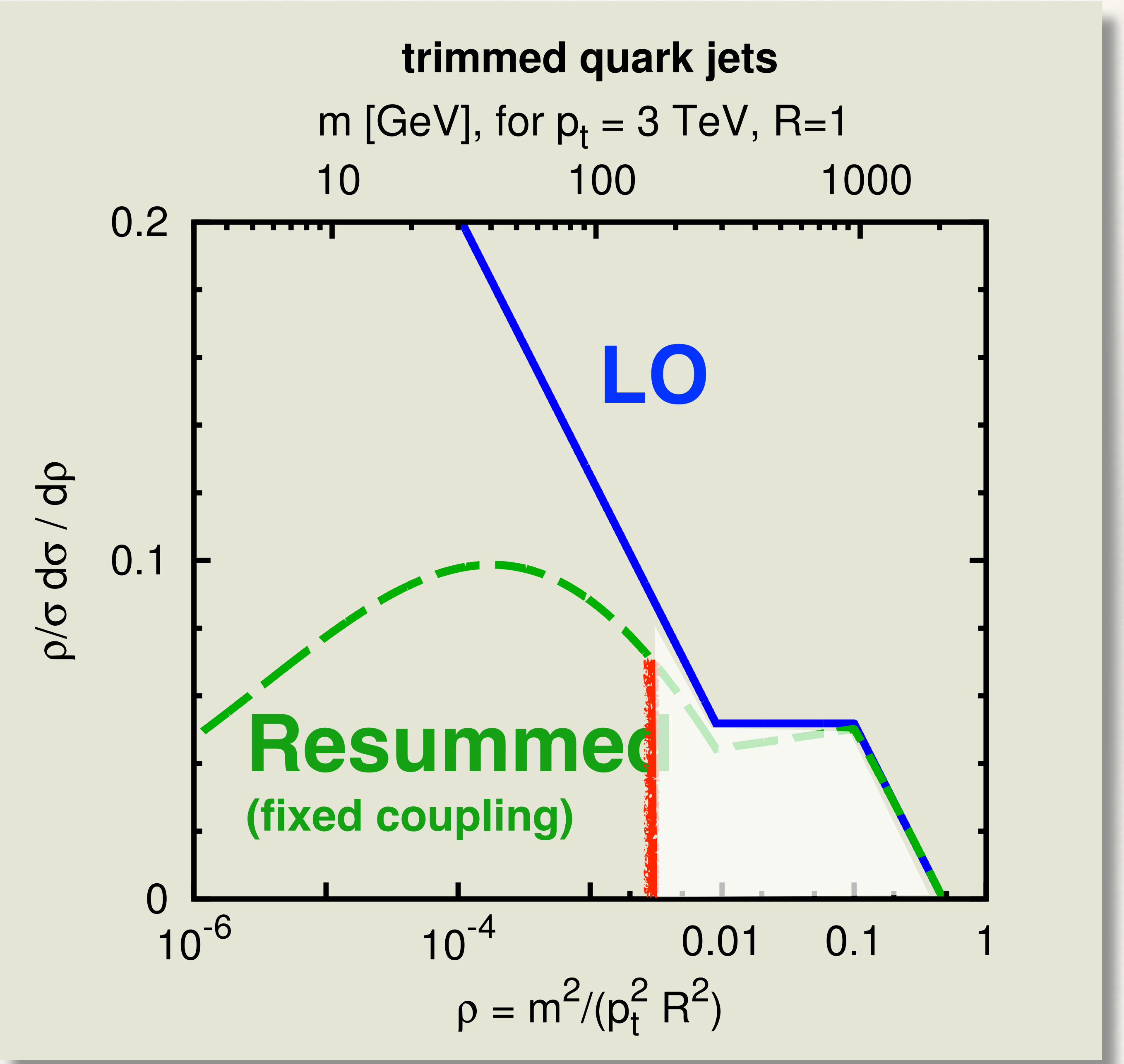


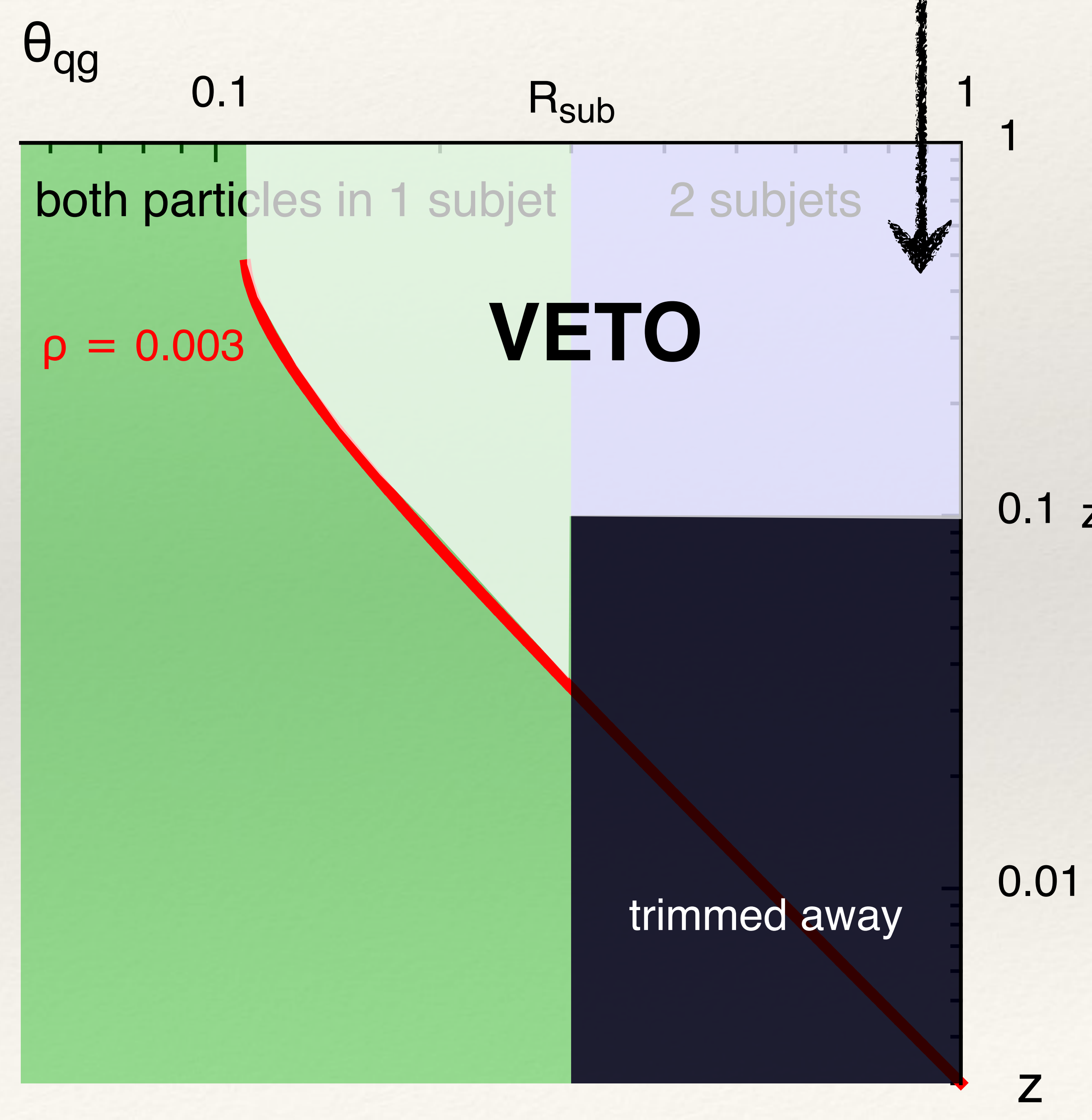


Trimming at all orders

$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho} \exp\left[-\int_\rho d\rho' \frac{1}{\sigma}\frac{d\sigma^{\text{trim,LO}}}{d\rho'}\right]$$

Sudakov

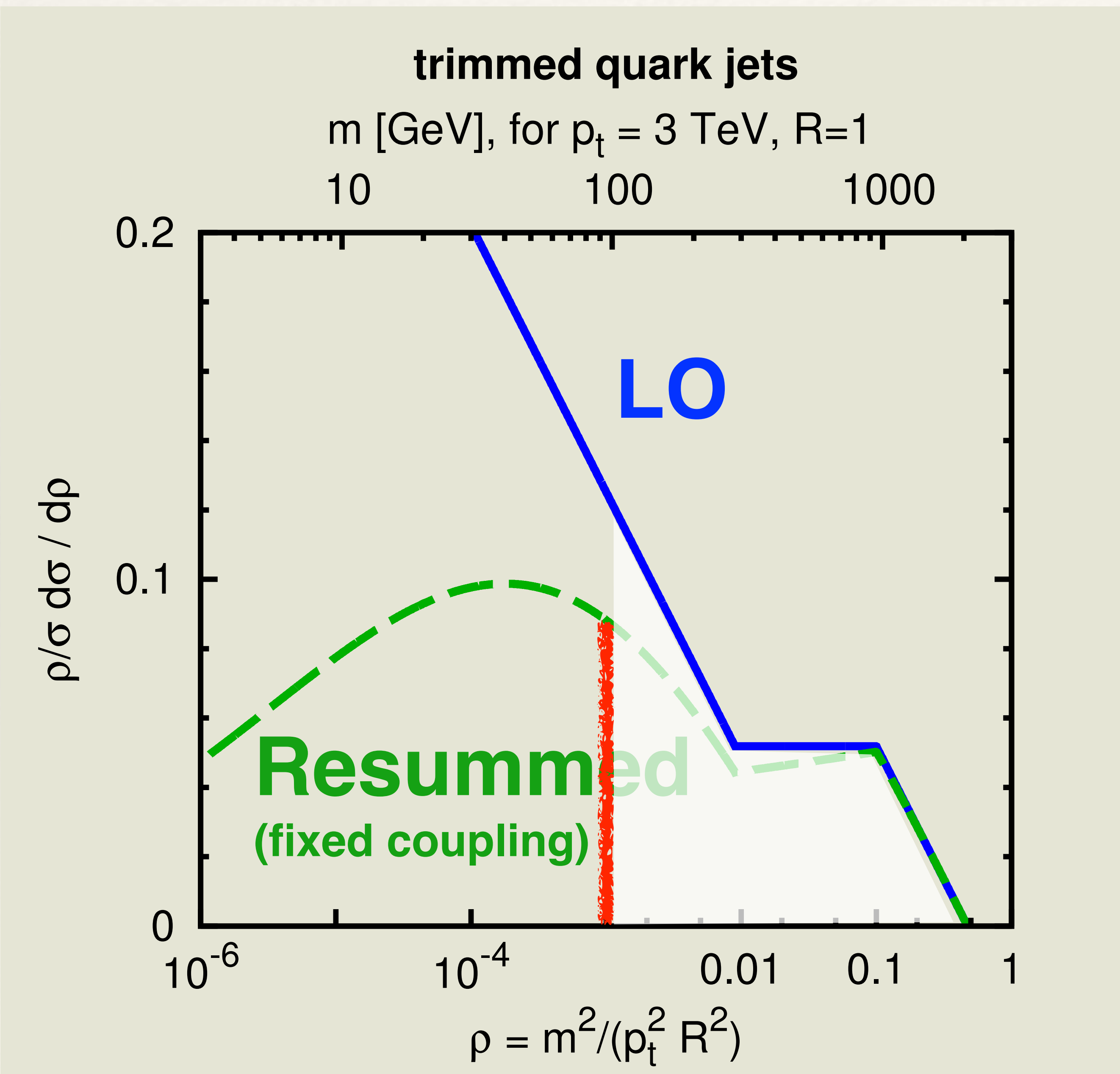


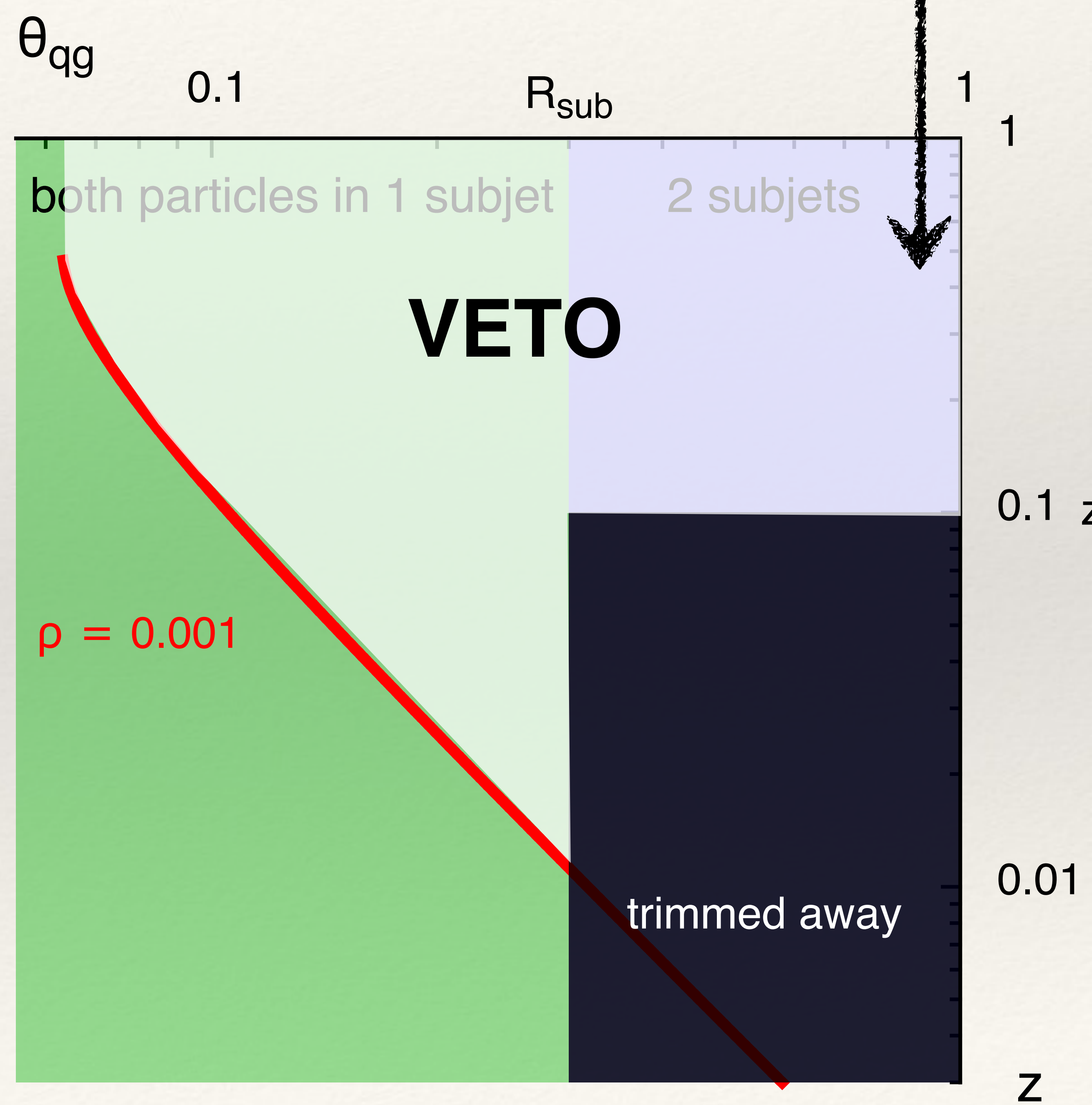


Trimming at all orders

$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho} \exp\left[-\int_{\rho} d\rho' \frac{1}{\sigma}\frac{d\sigma^{\text{trim,LO}}}{d\rho'}\right]$$

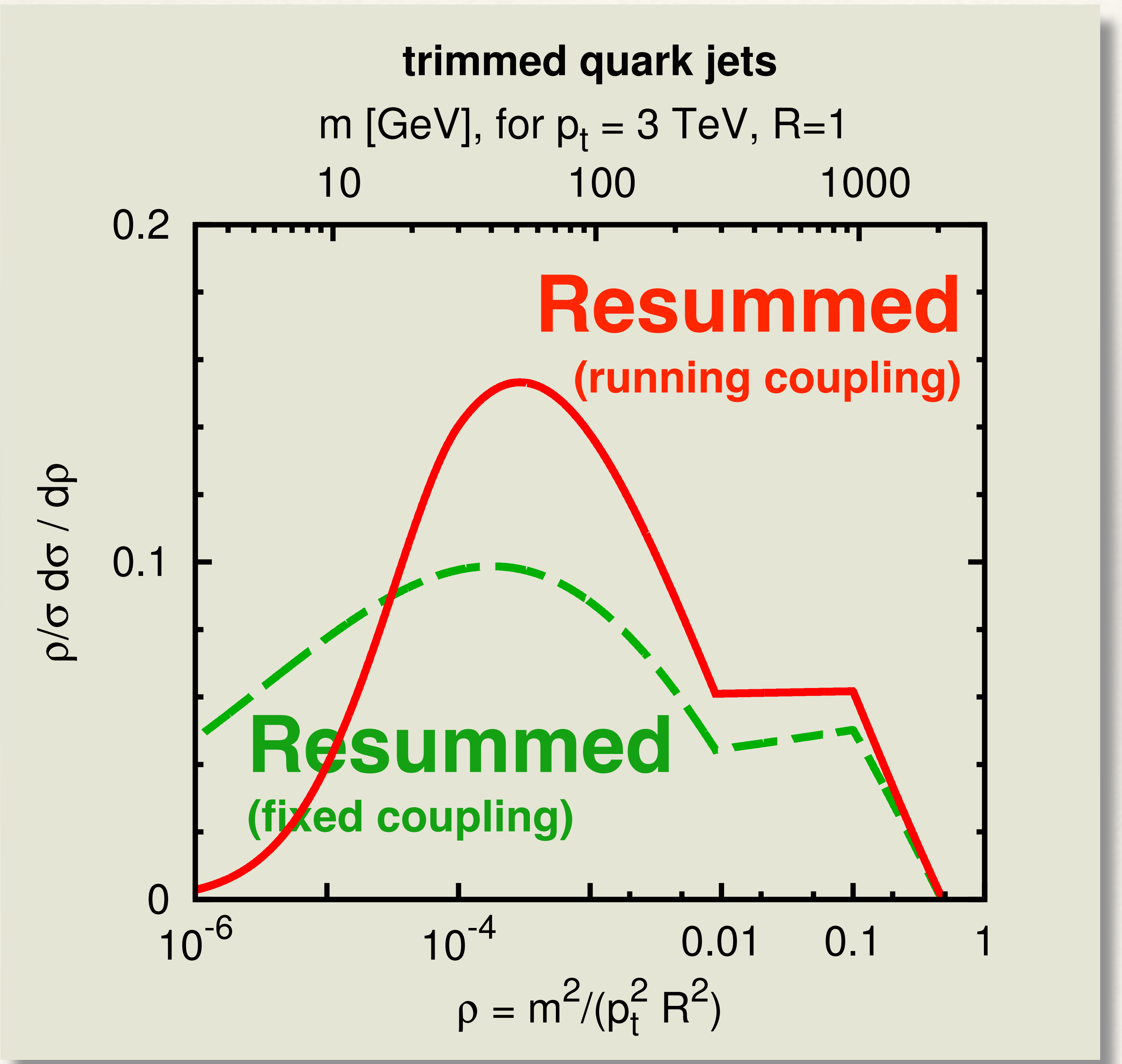
Sudakov





Trimming at all orders

$$\frac{d\sigma^{\text{trim,resum}}}{d\rho} = \frac{d\sigma^{\text{trim,LO}}}{d\rho} \exp\left[-\int_\rho d\rho' \frac{1}{\sigma}\frac{d\sigma^{\text{trim,LO}}}{d\rho'}\right]$$



Full resummation also needs treatment of running coupling