QCD-aware Jet Labeling

BOOST 2015

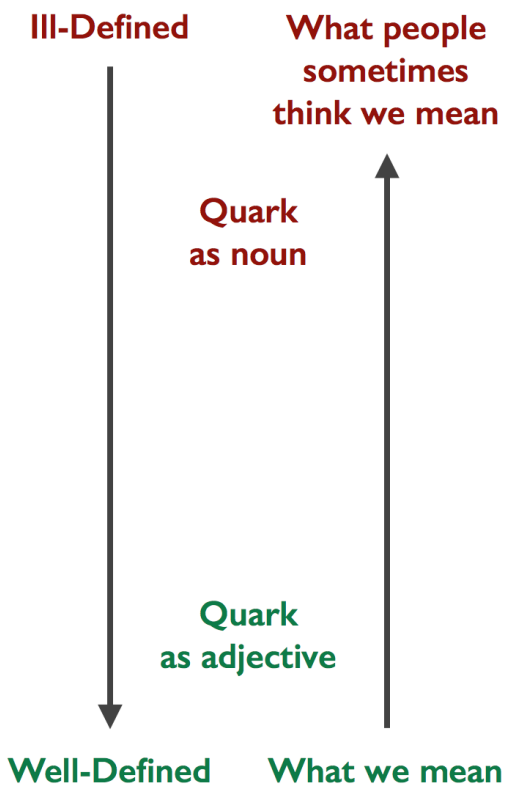
Andy Buckley, Chris Pollard

2015 08 13

What is a quark jet?

What is a Quark Jet?

From lunch/dinner discussions



A quark parton

A Born-level quark parton

The initiating quark parton in a final state shower

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)

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)

J Thaler & co., Les Houches 2015

Labeling jets

These issues notwithstanding, we'd like to find a way to categorize quark- and gluon-like jets to evaluate our ability to discriminate them.

What do we want out of a labeling algorithm?

- Parton shower agnostic
- Matrix-element generator agnostic
- "Simple"

The QCD-aware algorithm

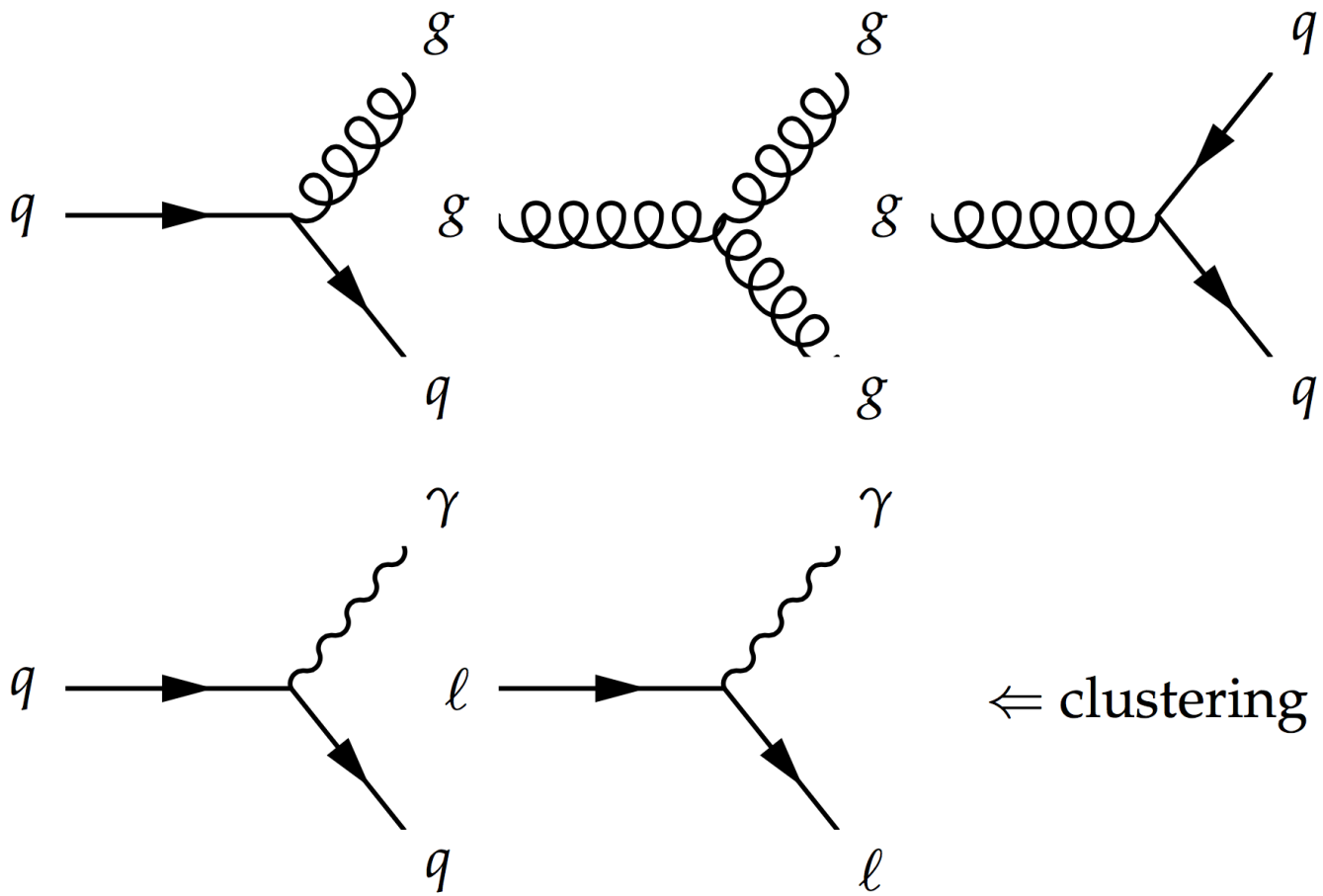
Here's what we came up with and tried...

1. Cluster the event *final* partons into jets ("parton jets") with two additions to the usual algorithms:
   - only allow clustering steps that follow $1 \to 2$ QCD or QED Feynman rules;
   - track the flavor of each pseudo-jet based on the Feynman rule used to create it.
2. For each particle jet in the event, label it with the flavor of the corresponding parton jet.

The parton clustering is fairly simple to realize for the k_t family of clustering algorithms: set $d_{ij} \to \infty$ for disallowed recombinations.

It's also similar to the "bland" flavor algorithm outlined in hep-ph/0601139 (Banfi, Salam, Zanderighi).

Allowed Feynman rules



Details

There are of course some details to be worked out (not exhaustive):

- What parton jet definitions should be used?
  (we used k_t parton jets and anti-k_t particle jets; $R = 0.4$ for *all* algorithms)
- What kinematic cuts should be applied?
  (parton, particle jet $p_T >$ 5, 25 GeV, respectively)
- How should we match particle- and parton jets?
  (closest ghost associated parton jet within $\Delta R < 0.2$)

"Performance"

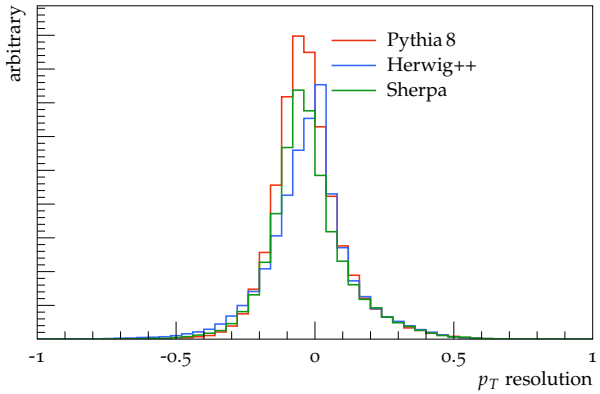
The "performance" of the labeling+matching scheme is somewhat ill-defined; we will compare to the "max-p_T" parton label scheme, but...

$\Delta p_T/p_T = \left(p_T^{\text{particle jet}} - p_T^{\text{parton jet}}\right)/p_T^{\text{particle jet}}$ is used to gauge how well the parton jet matches the particle jet.

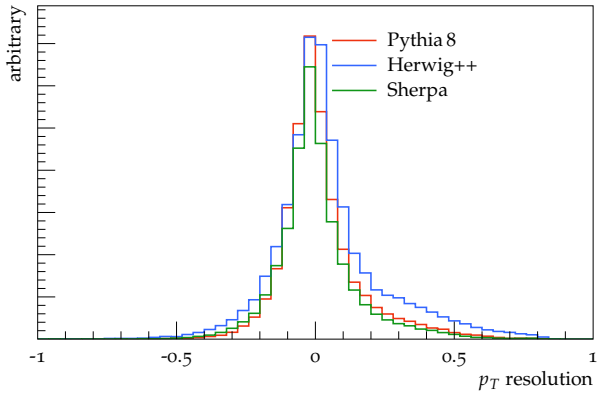
We will look at inclusive jet and γ+jet spectra from 14 TeV collisions under various conditions.

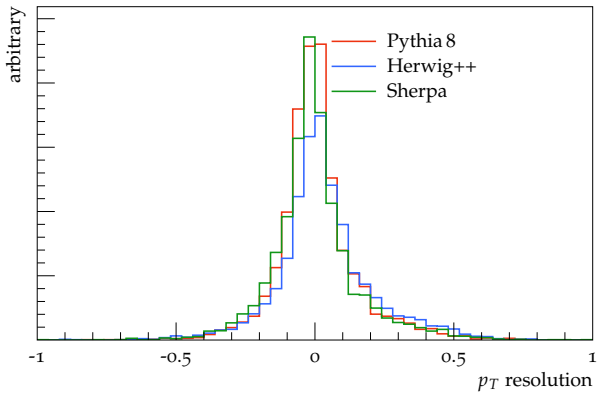
Generator dependence: inclusive jets

gluon labeled jets

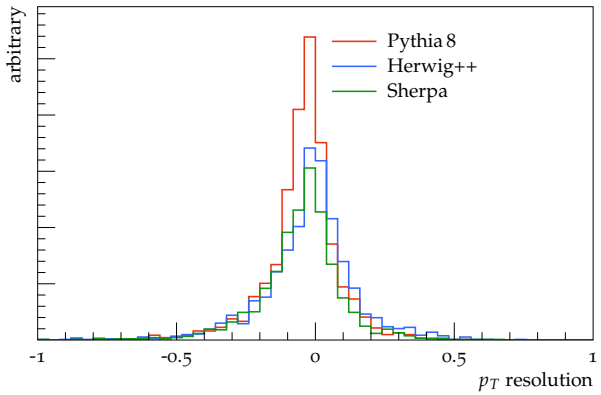


light-quark labeled jets





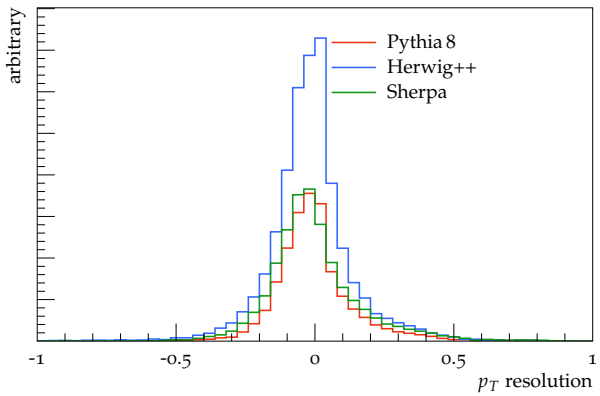
charm-quark labeled jets



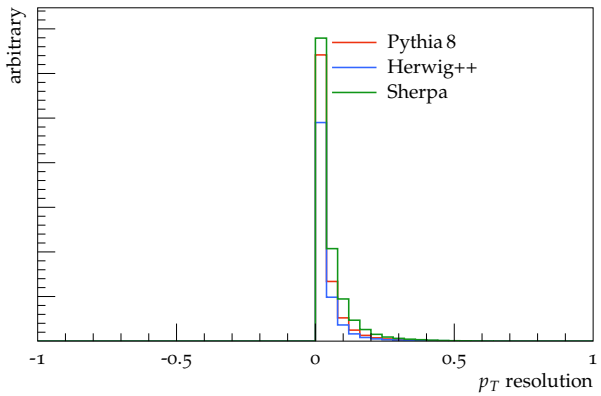
bottom-quark labeled jets

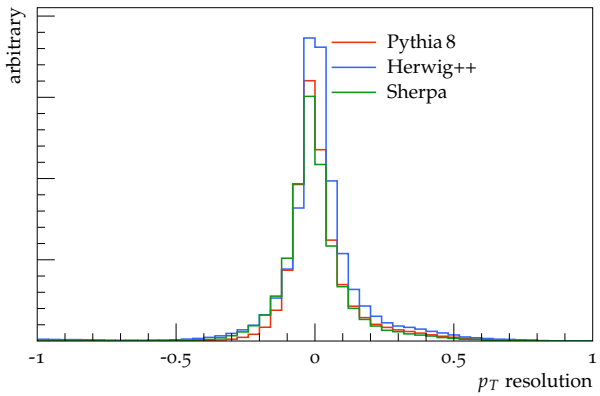
Generator dependence: γ+jets

gluon labeled jets



photon labeled jets





light-quark labeled jets

N.B.: due to allowed Feynman rules parton jets have a photon label *only if the photon is the sole constituent*.

Comparison with max-p_T label

pythia inclusive jets

k_t label \ max-p_T label	none	g	q	c	b	γ
none	-	1.1	0.1	-	-	-
g	-	53.5	1.1	-	-	-
q	-	15.7	20.5	-	-	-
c	-	1.5	-	3.6	-	-
b	-	0.4	-	-	2.1	-
γ	-	-	-	-	-	-

pythia γ+jet

k_t label \ max-p_T label	none	g	q	c	b	γ
none	-	-	0.2	-	-	0.9
g	-	3.1	1.8	0.2	-	0.4
q	-	0.9	27.5	-	-	3.4
c	-	-	-	6.4	-	0.2
b	-	-	-	-	0.9	-
γ	-	-	-	-	-	53.8

Comparison across generators and jet definitions

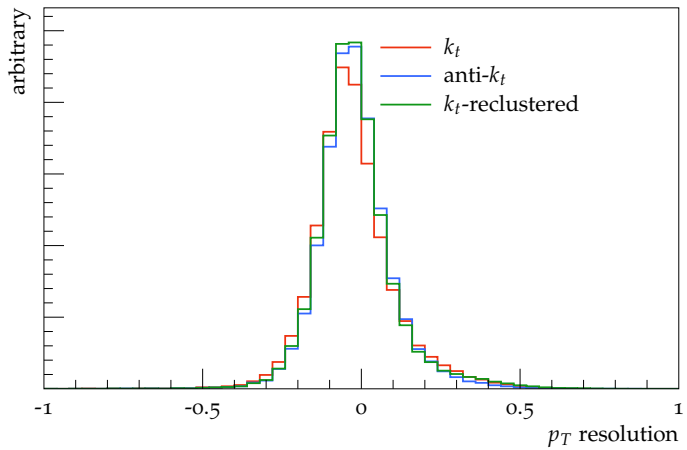
Scheme	Generator	Jets q/g	γ + jet γ/g	q/g
Max-p_T	Pythia 8	0.38	17.2	10.5
	Herwig++	0.33	7.7	4.8
	Sherpa	0.55	21.0	9.6
k_T	Pythia 8	0.80	10.4	8.2
	Herwig++	1.17	3.6	4.6
	Sherpa	0.85	10.5	7.5
anti-k_T	Pythia 8	0.79	10.2	8.3
	Herwig++	1.74	3.2	4.5
	Sherpa	0.86	10.2	7.5
Reclustered	Pythia 8	0.77	10.1	8.0
	Herwig++	1.36	3.5	4.8
	Sherpa	0.83	10.1	7.3

Algorithm dependence

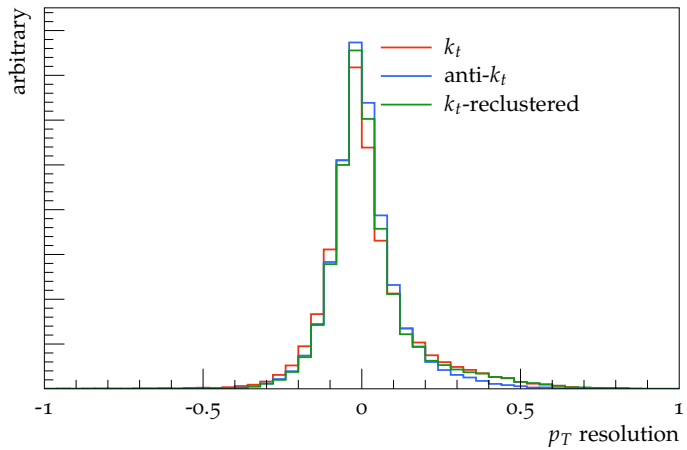
So far we've only discussed labeling with QCD-aware jets using the k_t distance measure.
Two other options for labeling anti-k_t particle jets, for example:

- use anti-k_t QCD-aware parton jets for labeling
- ghost-associate final partons to anti-k_t particle jets then cluster associated partons with the QCD-aware k_t algorithm ("k_t-reclustered")

gluon labeled jets



light-quark labeled jets



pythia inclusive jets

Dependence on ME parton multiplicity

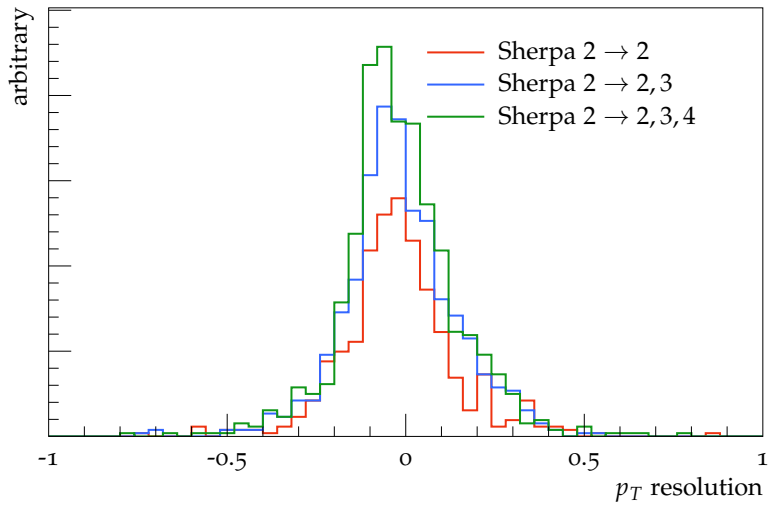
How dependent is the label on the number of matrix element partons?

Here we show only the 3^{rd} hardest jet in the event for inclusive jets from Sherpa $2 \to 2, 3, 4$ parton matrix elements

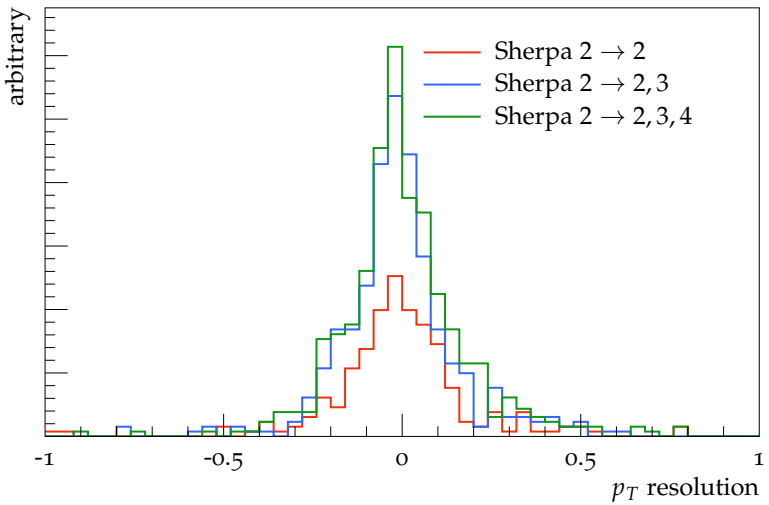
ME	$N_{\mathrm{j}3}/N_{\mathrm{j}3}^{2\to2}$	Gluon	Light quark	Light parton	Unlabeled
$2 \to 2$	1.00	62.7%	27.0%	89.6%	2.3%
$2 \to 3$	1.59	56.4%	31.4%	88.3%	2.9%
$2 \to 4$	1.79	58.3%	31.9%	90.2%	2.6%

Dependence on ME parton multiplicity

ME	$N_{j3}/N_{j3}^{2\to2}$	Gluon	Light quark	Light parton	Unlabeled
$2 \to 2$	1.00	62.7%	27.0%	89.6%	2.3%
$2 \to 3$	1.59	56.4%	31.4%	88.3%	2.9%
$2 \to 4$	1.79	58.3%	31.9%	90.2%	2.6%



gluon labeled jets (3^{rd} hardest jet)



light-quark labeled jets (3^{rd} hardest jet)

Sensitivity to MPI model

We expect the glu- and quark-iness of jets to be somewhat sensitive to MPI, but...

with MPI

Scheme	Generator	Jets q/g	$\gamma+$jet γ/g	$\gamma+$jet q/g
Max-p_T	Pythia 8	0.38	17.2	10.5
	Herwig++	0.33	7.7	4.8
	Sherpa	0.55	21.0	9.6
k_T	Pythia 8	0.80	10.4	8.2
	Herwig++	1.17	3.6	4.6
	Sherpa	0.85	10.5	7.5
anti-k_T	Pythia 8	0.79	10.2	8.3
	Herwig++	1.74	3.2	4.5
	Sherpa	0.86	10.2	7.5
Reclustered	Pythia 8	0.77	10.1	8.0
	Herwig++	1.36	3.5	4.8
	Sherpa	0.83	10.1	7.3

no MPI

Scheme	Generator	Jets q/g	$\gamma+$jet γ/g	$\gamma+$jet q/g
Max-p_T	Pythia 8	0.39	15.4	9.5
	Herwig++	0.33	18.3	11.4
	Sherpa	0.57	13.4	7.0
k_T	Pythia 8	0.65	11.8	7.6
	Herwig++	0.68	11.2	8.0
	Sherpa	0.73	13.0	7.0
anti-k_T	Pythia 8	0.65	11.7	7.6
	Herwig++	0.93	11.0	8.1
	Sherpa	0.74	12.9	7.0
Reclustered	Pythia 8	0.64	11.5	7.5
	Herwig++	0.80	11.0	8.2
	Sherpa	0.73	12.7	6.9

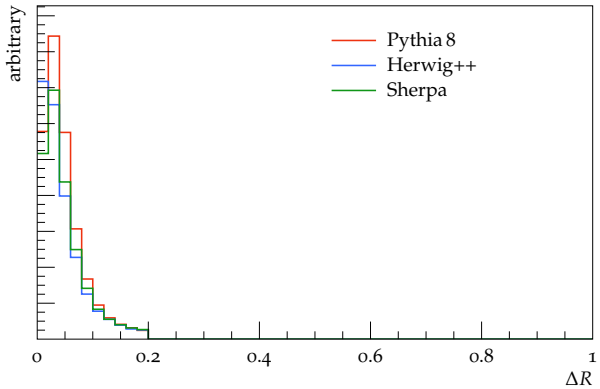
Conclusions and future work

- QCD-aware: simple modification to generic jet algorithms that can be run to cluster partons into flavor-ful jets
- We can then match parton jets to "real" jets for labels
- We'd also like to try this on more complicated final states like W, Z, H, top
- Still some outstanding questions: can we understand/control effect of soft quarks from MPI (especially $\mathrm{H{+}{+}}$)?
- More info at arXiv:1507.00508
- Plugin available in fj-contrib; release in the near future.

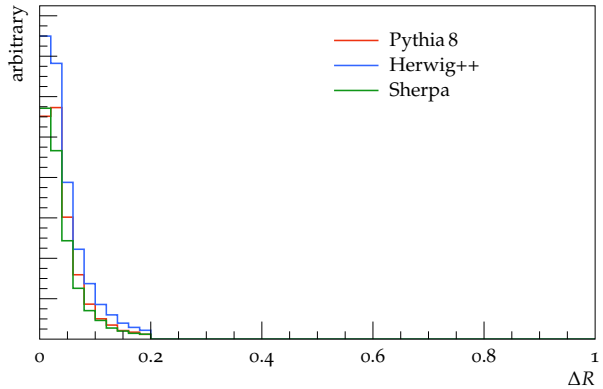
Additional Material

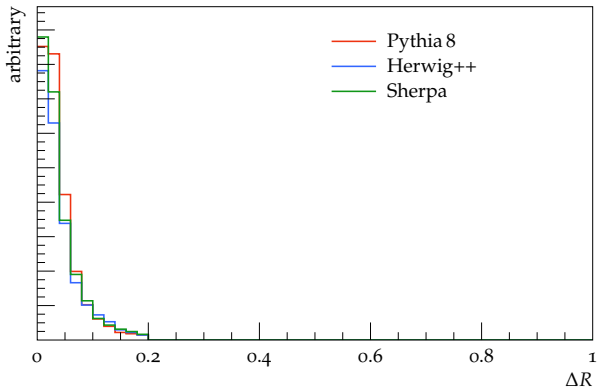
Generator dependence: inclusive jets

gluon labeled jets

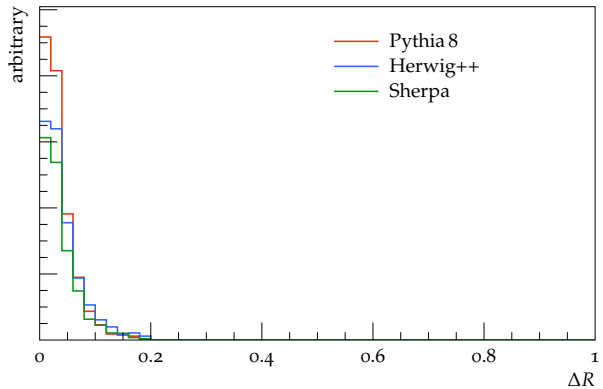


light-quark labeled jets





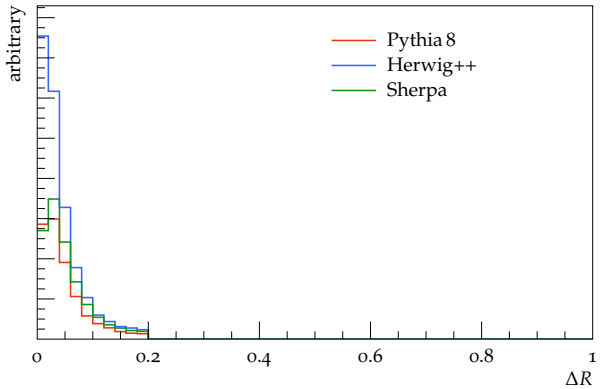
charm-quark labeled jets



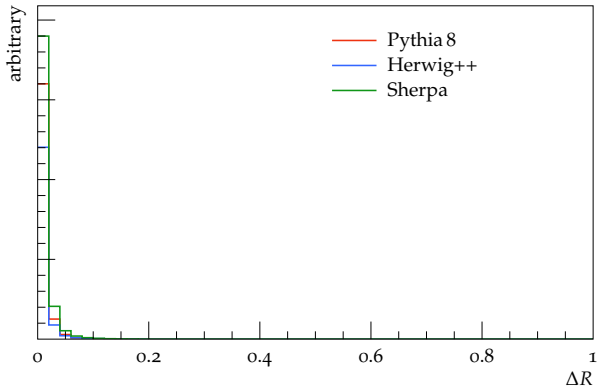
bottom-quark labeled jets

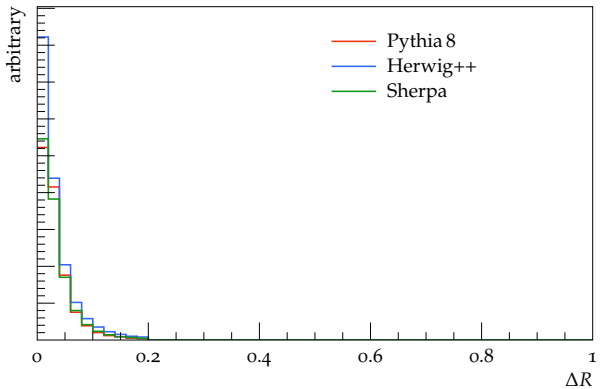
Generator dependence: γ+jets

gluon labeled jets



photon labeled jets



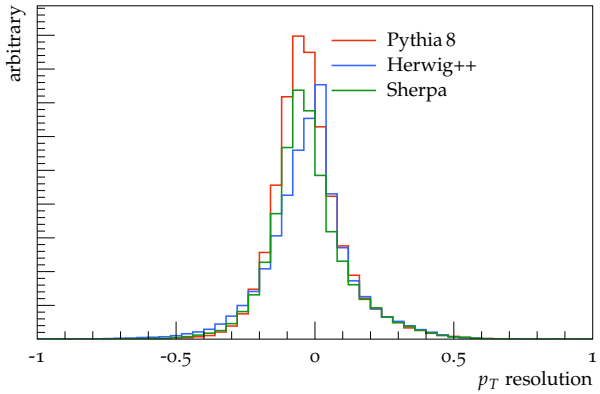


light-quark labeled jets

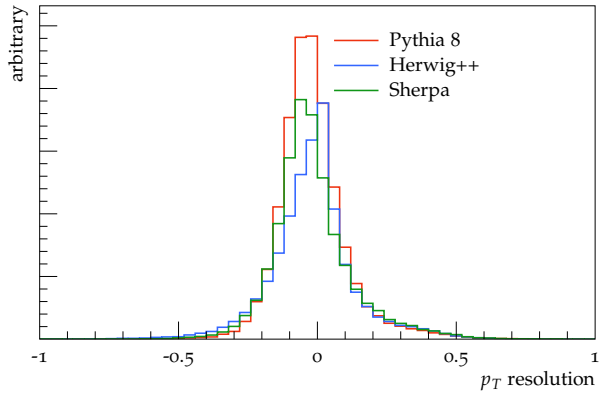
N.B.: due to allowed Feynman rules parton jets have a photon label *only if the photon is the sole constituent*.

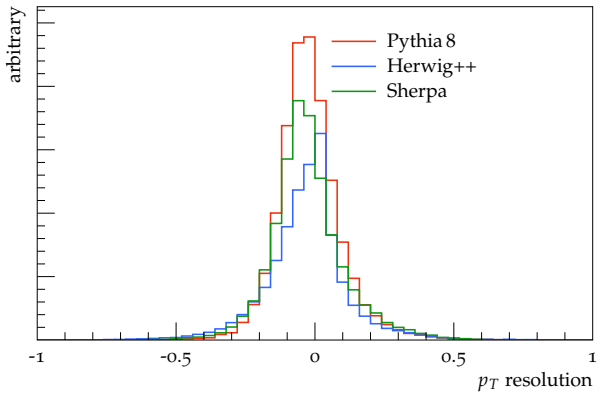
Generator and algorithm dependence: inclusive jets

k_t parton jets



k_t-reclustered parton jets



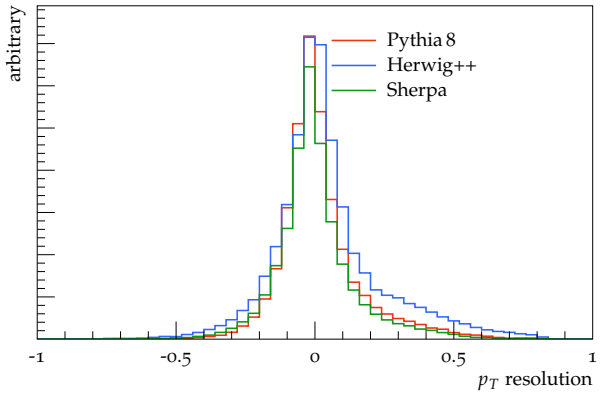


anti-k_t parton jets

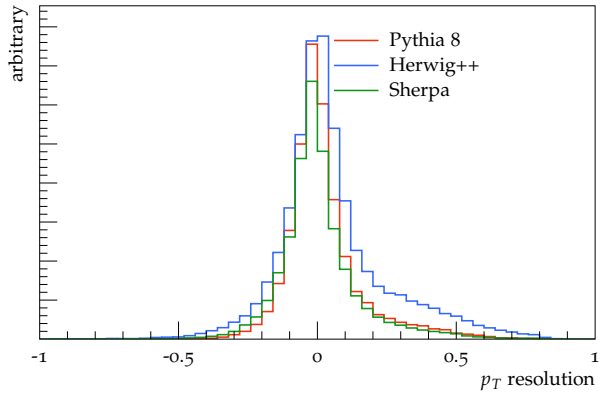
gluon labeled jets

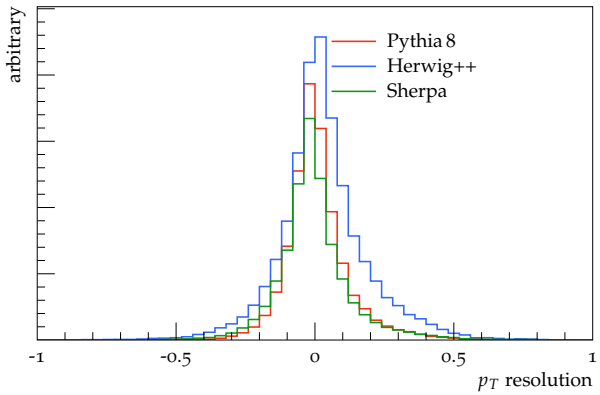
Generator and algorithm dependence: inclusive jets

k_t parton jets



k_t-reclustered parton jets





anti-k_t parton jets

light-quark labeled jets

Comparison across algorithms

anti-k_t label \ k_t label	none	g	q	c	b	γ
none	0.8	0.1	-	-	-	-
g	0.3	52.5	2.5	0.1	-	-
q	0.1	2.1	33.7	-	-	-
c	-	-	-	5.0	-	-
b	-	-	-	-	2.5	-
γ	-	-	-	-	-	-

k_t label \ k_t-reclustered label	none	g	q	c	b	γ
none	-	0.9	0.3	-	-	-
g	-	53.0	1.6	-	-	-
q	-	2.5	33.8	-	-	-
c	-	0.1	-	5.0	-	-
b	-	-	-	-	2.5	-
γ	-	-	-	-	-	-

Comparison across algorithms and MPI

MPI on

k_t label \ max-p_T label	none	g	q	c	b	γ
none	-	0.2	0.3	-	-	0.6
g	-	5.6	2.7	0.2	-	2.8
q	-	4.3	25.6	0.2	-	14.7
c	-	0.2	-	6.1	-	0.3
b	-	-	-	-	0.9	-
γ	-	-	-	-	-	34.8

MPI off

k_t label \ max-p_T label	none	g	q	c	b	γ
none	-	-	0.2	-	-	0.9
g	-	3.1	1.8	0.2	-	0.4
q	-	0.9	27.5	-	-	3.4
c	-	-	-	6.4	-	0.2
b	-	-	-	-	0.9	-
γ	-	-	-	-	-	53.8