Jets and QCD — lecture 4

Matteo Cacciari[1,2] and Gavin Salam[3,4,1]

[1] LPTHE (CNRS/UPMC)
[2] Université Denis Diderot (Paris 7)
[3] CERN, PH-TH
[4] Princeton University

Focus Week at the GGI Workshop
High energy QCD after the start of the LHC
Florence, Italy, 12-16 September 2011

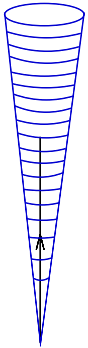
Towards an understanding of jets

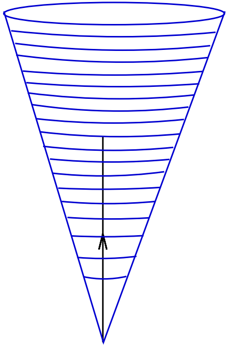
How a jet is and isn't like a parton — quantitatively

And how this relationship is affected by the jet radius

Small v. large jet radius (R) ≡ HSBC

Small jet radius | **Large jet radius**

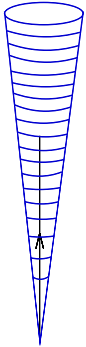




single parton @ LO: **jet radius irrelevant**

Small v. large jet radius (R) ≡ HSBC

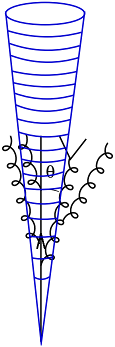
Small jet radius **Large jet radius**

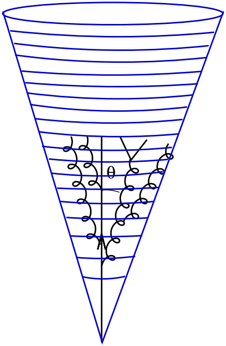


single pa

Small v. large jet radius (R) $\equiv$ HSBC

Small jet radius **Large jet radius**

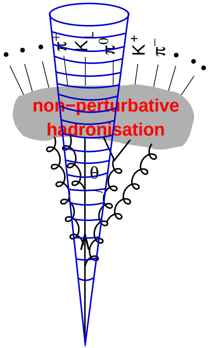




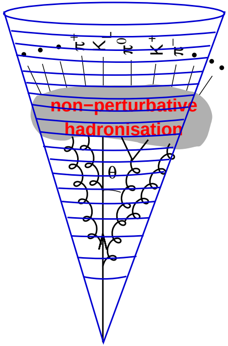
perturbative fragmentation: **large jet radius better**
(it captures more)

Small v. large jet radius (R) $\equiv$ HSBC

Small jet radius



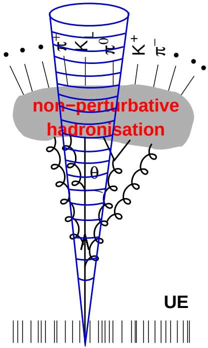
Large jet radius



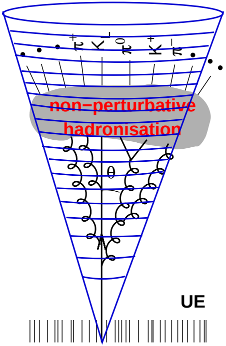
non-perturbative fragmentation: **large jet radius better**
(it captures more)

Small v. large jet radius (R) $\equiv$ HSBC

Small jet radius



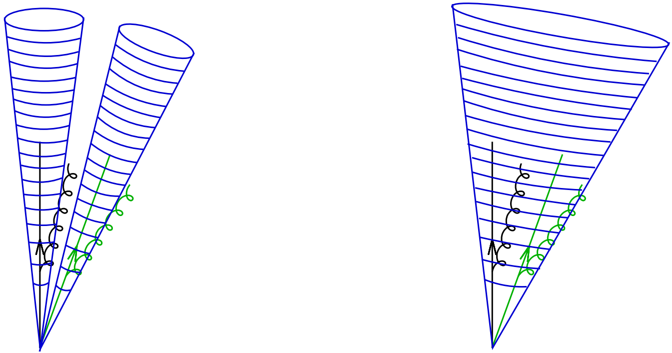
Large jet radius



underlying ev. & pileup "noise": **small jet radius better**
(it captures less)

Small v. large jet radius (R) ≡ HSBC

Small jet radius | **Large jet radius**



multi-hard-parton events: **small jet radius better**

(it resolves partons more effectively)

Parton p_t v. jet p_t

3 physical effects:

1. Gluon radiation from the parton
2. Hadronisation
3. Underlying Event

One important consideration:

Whether the parton is a quark or a gluon

[quarks radiate with colour factor $C_F = 4/3$
gluons radiate with colour factor $C_A = 3$]

Jet p_t v. parton p_t: perturbatively?

The question's dangerous: a "parton" is an ambiguous concept

Three limits can help you:

- Threshold limit — e.g. de Florian & Vogelsang '07
- Parton from color-neutral object decay (Z')
- Small-R (radius) limit for jet

One simple result (small-R limit)

$$\frac{\langle p_{t,jet} - p_{t,parton}\rangle}{p_t} = \frac{\alpha_s}{\pi} \ln R \times \begin{cases} 1.01\,C_F & \textit{quarks} \\ 0.94\,C_A + 0.07\,n_f & \textit{gluons} \end{cases} + \mathcal{O}\left(\alpha_s\right)$$

only $\mathcal{O}\left(\alpha_s\right)$ depends on algorithm & process
cf. Dasgupta, Magnea & GPS '07

Jet p_t v. parton p_t: hadronisation?

Hadronisation: the "parton-shower" → hadrons transition

Method:

- "infrared finite α_s" — à la Dokshitzer & Webber '95
- **prediction** based on e^+e^- event shape data
- could have been deduced from old work — Korchemsky & Sterman '95, Seymour '97

Main result

$$\langle p_{t,jet} - p_{t,parton-shower} \rangle \simeq -\frac{0.4\ \text{GeV}}{R} \times \begin{cases} C_F & quarks \\ C_A & gluons \end{cases}$$

cf. Dasgupta, Magnea & GPS '07

coefficient holds for anti-k_t; see Dasgupta & Delenda '09 for k_t alg.

Underlying Event (UE)

"Naive" prediction (UE $\simeq$ colour dipole between pp):

$$\Delta p_t \simeq 0.4 \text{ GeV} \times \frac{R^2}{2} \times \begin{cases} C_F & q\bar{q} \text{ dipole} \\ C_A & \text{gluon dipole} \end{cases}$$

Modern Monte Carlo tunes tell you ($\sqrt{s} = 7$ TeV):

$$\Delta p_t \simeq \mathbf{8\ GeV} \times \frac{R^2}{2} \simeq 1.2 \text{ GeV} \times (\pi R^2)$$

This big coefficient motivates special effort to understand interplay between jet algorithm and UE: "jet areas"

How does coefficient depend on algorithm?

How does it depend on jet p_t? How does it fluctuate?

cf. Cacciari, GPS & Soyez '08

Underlying Event (UE)

"Naive" prediction (UE $\simeq$ colour dipole between pp):

$$\Delta p_t \simeq 0.4 \text{ GeV} \times \frac{R^2}{2} \times \begin{cases} C_F & q\bar{q} \text{ dipole} \\ C_A & \text{gluon dipole} \end{cases}$$

Modern Monte Carlo tunes tell you ($\sqrt{s} = 7$ TeV):

$$\Delta p_t \simeq \mathbf{8\ GeV} \times \frac{R^2}{2} \simeq 1.2 \text{ GeV} \times (\pi R^2)$$

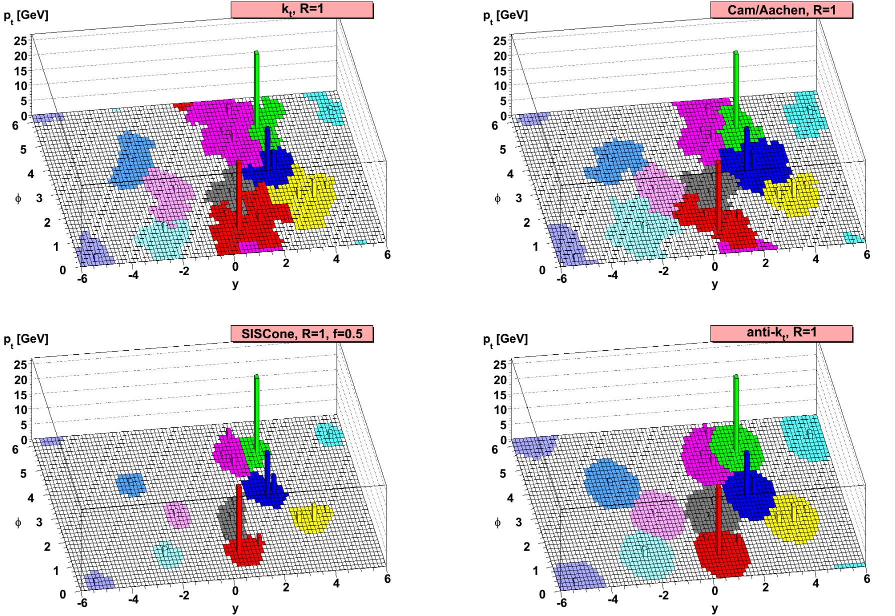
This big coefficient motivates special effort to understand interplay between jet algorithm and UE: "jet areas"

How does coefficient depend on algorithm?

How does it depend on jet p_t? How does it fluctuate?

cf. Cacciari, GPS & Soyez '08

Jet contours – visualised



Using our understanding to help discover a dijet resonance, $q\bar{q} \to X \to q\bar{q}$.

What R is best for an isolated jet?

E.g. to reconstruct $m_X \sim (p_{tq} + p_{t\bar{q}})$

PT radiation:

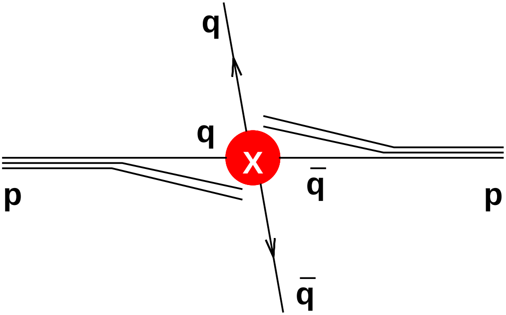
$$q: \quad \langle \Delta p_t \rangle \simeq \frac{\alpha_s C_F}{\pi} p_t \ln R$$

Hadronisation:

$$q: \quad \langle \Delta p_t \rangle \simeq -\frac{C_F}{R} \cdot 0.4 \text{ GeV}$$

Underlying event:

$$q, g: \quad \langle \Delta p_t \rangle \simeq \frac{R^2}{2} \cdot 2.5\text{—}15 \text{ GeV}$$



Minimise fluctuations in p_t

Use crude approximation:

$$\langle \Delta p_t^2 \rangle \simeq \langle \Delta p_t \rangle^2$$

in small-R limit (!)

NB: full calc, correct fluct: Soyez '10

What R is best for an isolated jet?

PT radiation:

$$q: \quad \langle \Delta p_t \rangle \simeq \frac{\alpha_s C_F}{\pi} p_t \ln R$$

Hadronisation:

$$q: \quad \langle \Delta p_t \rangle \simeq -\frac{C_F}{R} \cdot 0.4 \text{ GeV}$$

Underlying event:

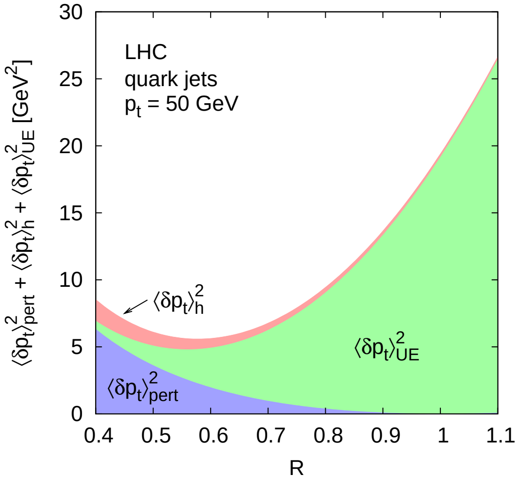
$$q, g: \quad \langle \Delta p_t \rangle \simeq \frac{R^2}{2} \cdot 2.5\text{—}15 \text{ GeV}$$

Minimise fluctuations in p_t

Use crude approximation:

$$\langle \Delta p_t^2 \rangle \simeq \langle \Delta p_t \rangle^2$$

50 GeV quark jet



in small-R limit (!)

NB: full calc, correct fluct: Soyez '10

What R is best for an isolated jet?

PT radiation:

$$q: \quad \langle \Delta p_t \rangle \simeq \frac{\alpha_s C_F}{\pi} p_t \ln R$$

Hadronisation:

$$q: \quad \langle \Delta p_t \rangle \simeq -\frac{C_F}{R} \cdot 0.4 \text{ GeV}$$

Underlying event:

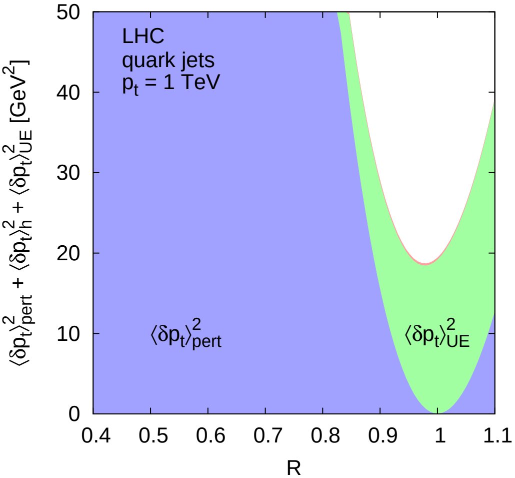
$$q, g: \quad \langle \Delta p_t \rangle \simeq \frac{R^2}{2} \cdot 2.5\text{—}15 \text{ GeV}$$

Minimise fluctuations in p_t

Use crude approximation:

$$\langle \Delta p_t^2 \rangle \simeq \langle \Delta p_t \rangle^2$$

1 TeV quark jet



in small-R limit (!)

NB: full calc, correct fluct: Soyez '10

What R is best for an isolated jet?

PT radiation:

$$q: \quad \langle \Delta p_t \rangle \simeq \frac{\alpha_s C_F}{\pi} p_t \ln R$$

Hadr

$q:$

At low p_t, small R limits relative impact of UE

At high p_t, perturbative effects dominate over non-perturbative $\to R_{best} \sim 1$.

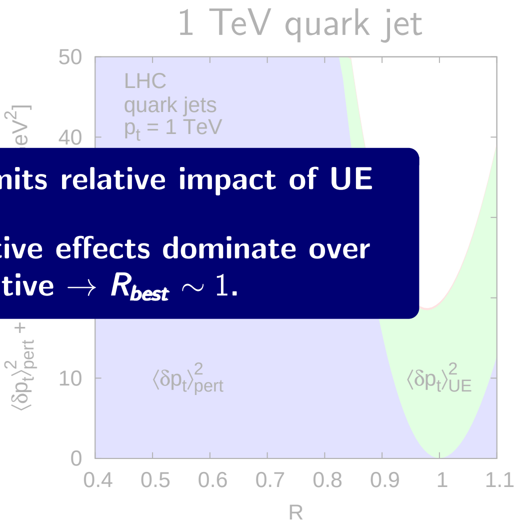
Underlying event:

$$q, g: \quad \langle \Delta p_t \rangle \simeq \frac{R^2}{2} \cdot 2.5\text{—}15 \text{ GeV}$$

Minimise fluctuations in p_t

Use crude approximation:

$$\langle \Delta p_t^2 \rangle \simeq \langle \Delta p_t \rangle^2$$

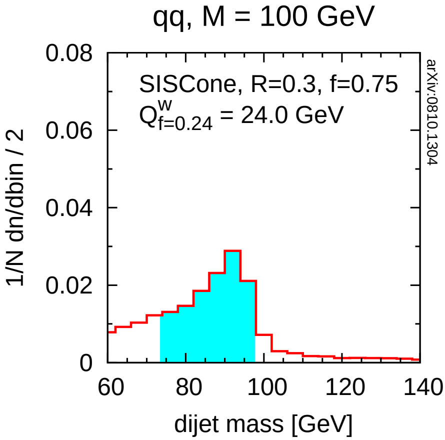


in small-R limit (!)

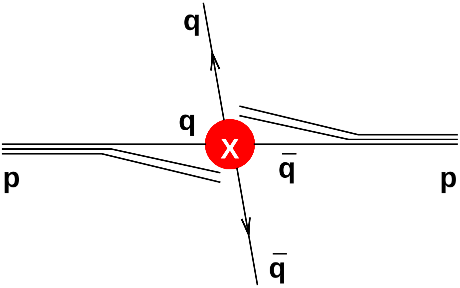
NB: full calc, correct fluct: Soyez '10

Dijet mass: scan over R [Pythia 6.4]

R = 0.3

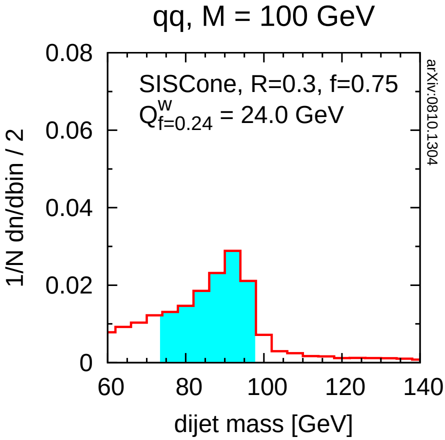


Resonance X → dijets

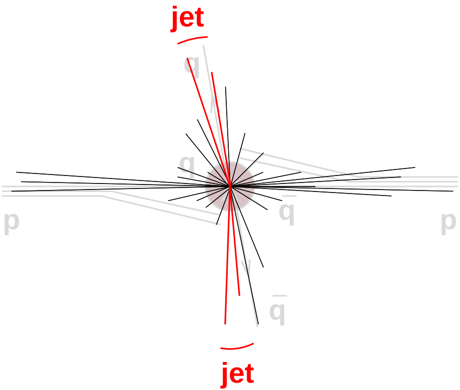


Dijet mass: scan over *R* [Pythia 6.4]

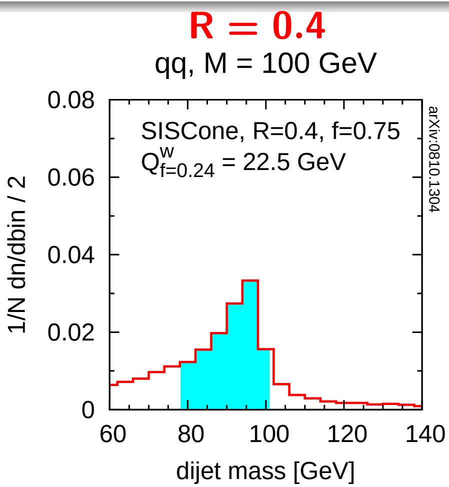
R = 0.3

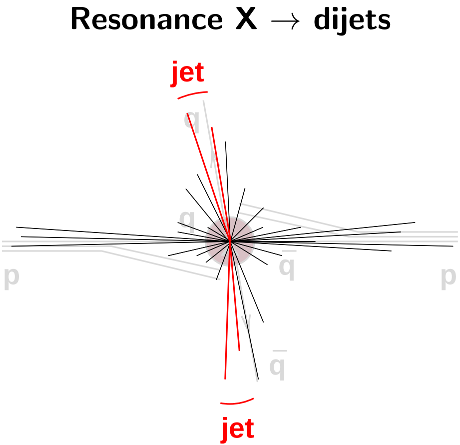


Resonance X → dijets



Dijet mass: scan over *R* [Pythia 6.4]

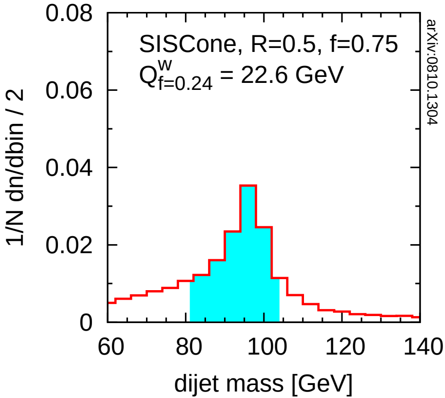




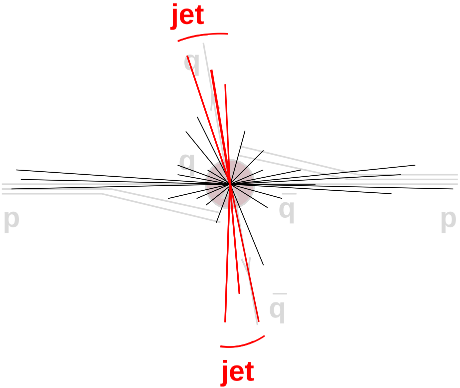
Dijet mass: scan over R [Pythia 6.4]

R = 0.5

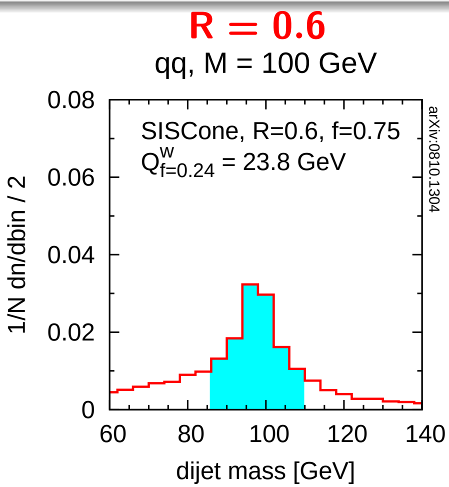
qq, M = 100 GeV

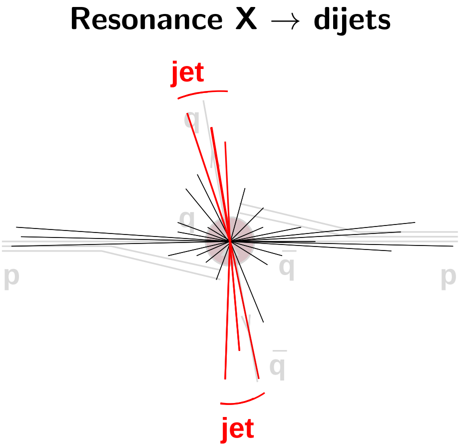


Resonance X → dijets

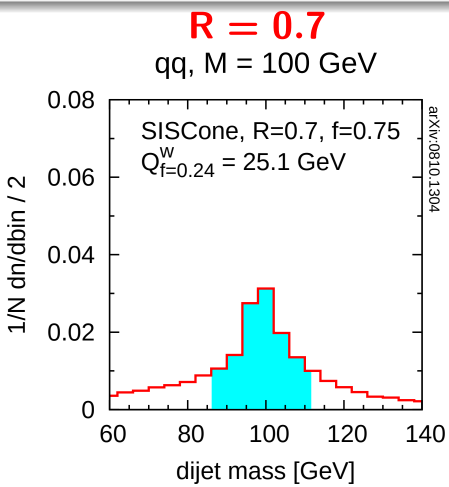


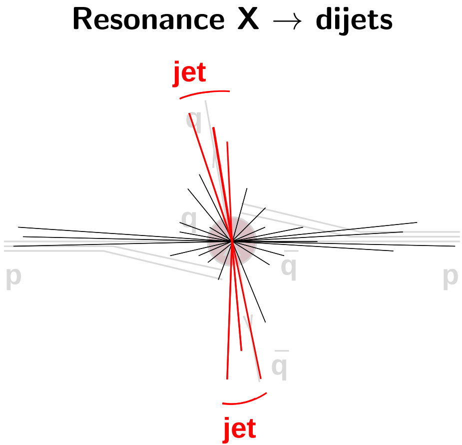
Dijet mass: scan over R [Pythia 6.4]



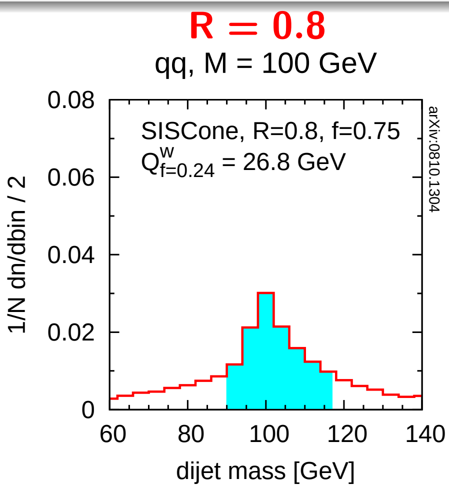


Dijet mass: scan over R [Pythia 6.4]

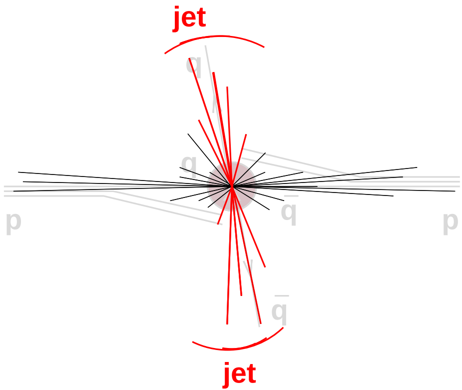




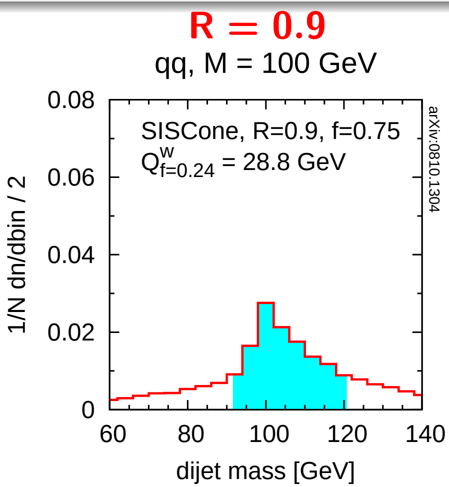
Dijet mass: scan over R [Pythia 6.4]



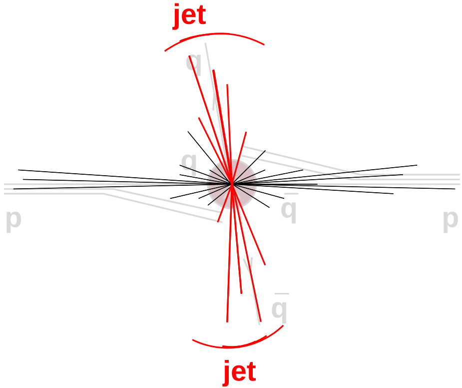
Resonance X → dijets



Dijet mass: scan over *R* [Pythia 6.4]

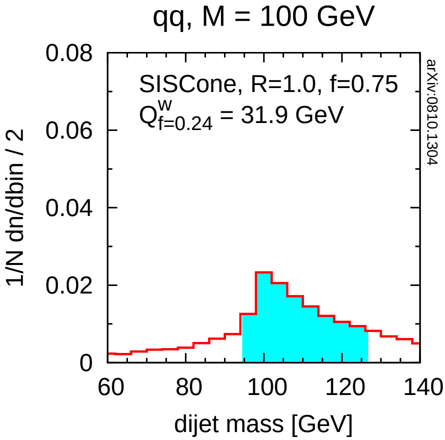


Resonance X → dijets

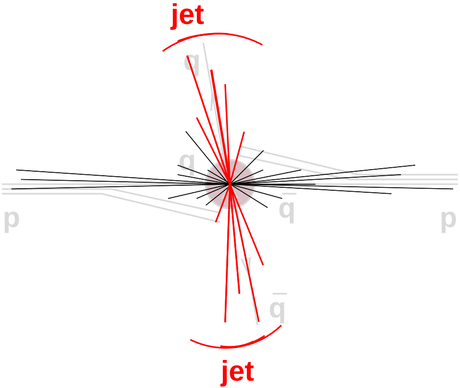


Dijet mass: scan over *R* [Pythia 6.4]

R = 1.0

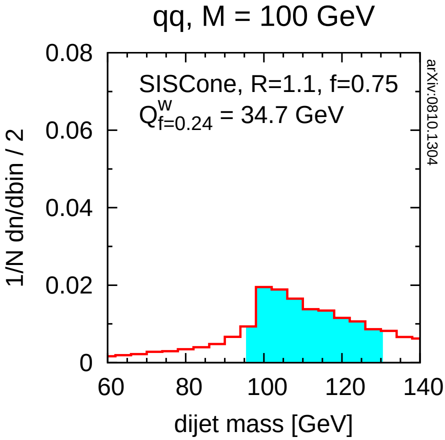


Resonance X → dijets

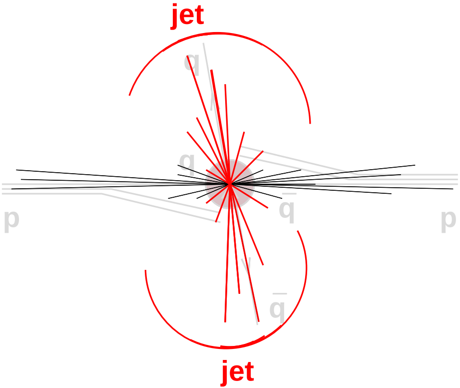


Dijet mass: scan over R [Pythia 6.4]

R = 1.1



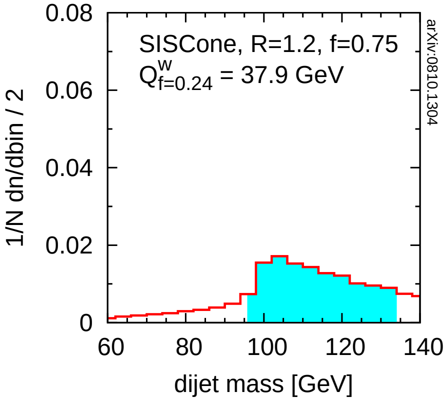
Resonance X → dijets



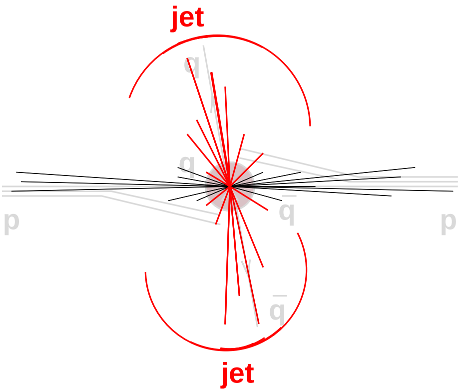
Dijet mass: scan over *R* [Pythia 6.4]

R = 1.2

qq, M = 100 GeV



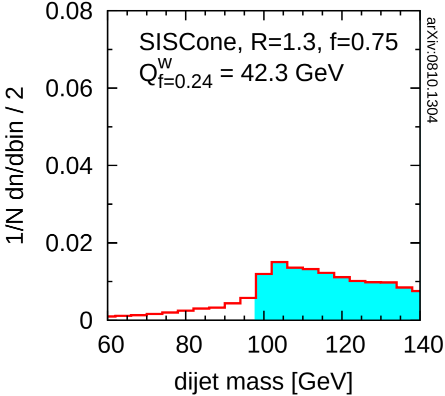
Resonance X → dijets



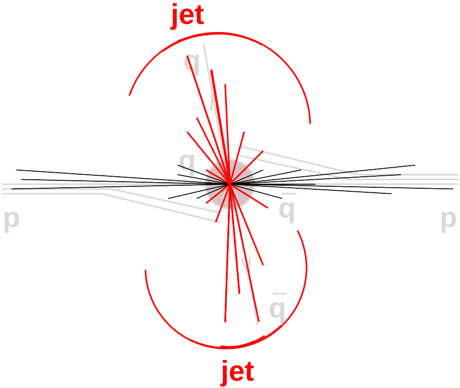
Dijet mass: scan over R [Pythia 6.4]

R = 1.3

qq, M = 100 GeV

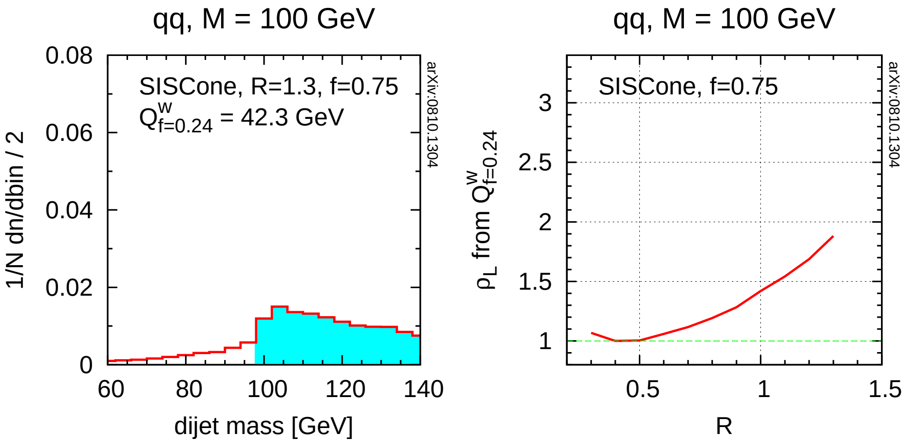


Resonance X → dijets



Dijet mass: scan over R [Pythia 6.4]

R = 1.3

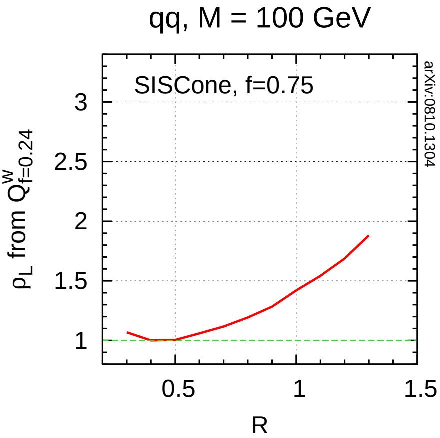


After scanning, summarise "quality" v. R. Minimum $\equiv$ BEST

picture not so different from crude analytical estimate

Scan through $q\bar{q}$ mass values

$\mathbf{m_{qq} = 100}$ GeV



Best R is at minimum of curve

- Best R depends strongly on mass of system
- Increases with mass

can reproduce this anayltically
Soyez '10

Message received by CMS: they combine all $R = 0.5$ jets ($p_t >$ 10 GeV) within $\Delta R = 1.1$ of two hardest to improve resolution.

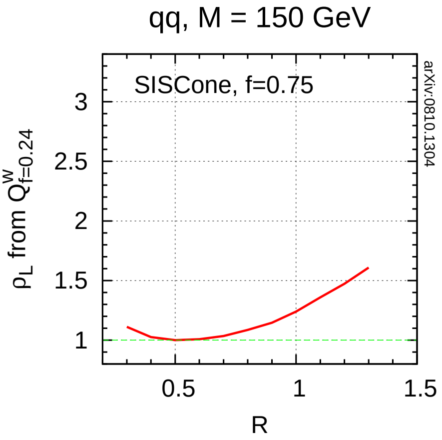
ATLAS '11 still just use $R = 0.6$

NB: 100,000 plots for various jet algorithms, narrow qq and gg resonances from `http://quality.fastjet.fr` Cacciari, Rojo, GPS & Soyez '08

Other related work: Krohn, Thaler & Wang '09

Scan through $q\bar{q}$ mass values

$\mathbf{m_{qq} = 150}$ GeV



Best R is at minimum of curve

- Best R depends strongly on mass of system
- Increases with mass

can reproduce this anayltically
Soyez '10

Message received by CMS: they combine all $R = 0.5$ jets ($p_t > 10$ GeV) within $\Delta R = 1.1$ of two hardest to improve resolution.

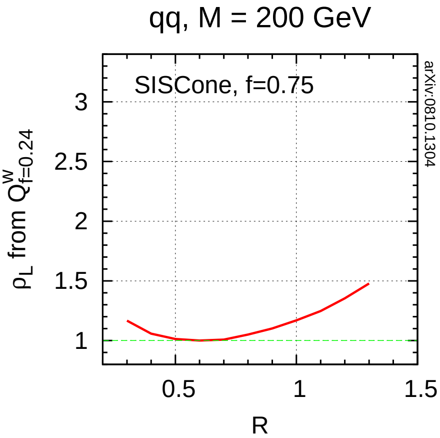
ATLAS '11 still just use $R = 0.6$

NB: 100,000 plots for various jet algorithms, narrow qq and gg resonances from `http://quality.fastjet.fr` Cacciari, Rojo, GPS & Soyez '08

Other related work: Krohn, Thaler & Wang '09

Scan through $q\bar{q}$ mass values

$\mathbf{m_{qq} = 200\ GeV}$



Best R is at minimum of curve

- Best R depends strongly on mass of system
- Increases with mass

can reproduce this anayltically
Soyez '10

Message received by CMS: they combine all $R = 0.5$ jets ($p_t > 10$ GeV) within $\Delta R = 1.1$ of two hardest to improve resolution.

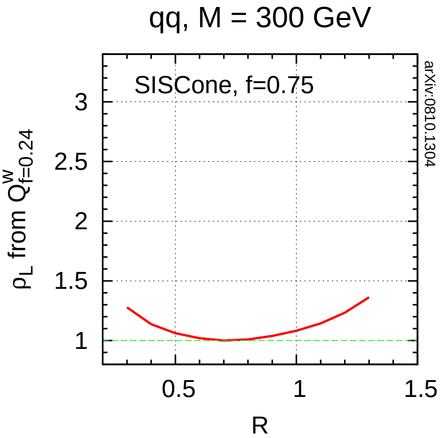
ATLAS '11 still just use $R = 0.6$

NB: 100,000 plots for various jet algorithms, narrow qq and gg resonances from `http://quality.fastjet.fr` Cacciari, Rojo, GPS & Soyez '08

Other related work: Krohn, Thaler & Wang '09

Scan through $q\bar{q}$ mass values

$m_{qq} = 300$ GeV



Best R is at minimum of curve

- Best R depends strongly on mass of system
- Increases with mass

can reproduce this anayltically
Soyez '10

Message received by CMS: they combine all $R = 0.5$ jets ($p_t >$ 10 GeV) within $\Delta R = 1.1$ of two hardest to improve resolution.

ATLAS '11 still just use $R = 0.6$

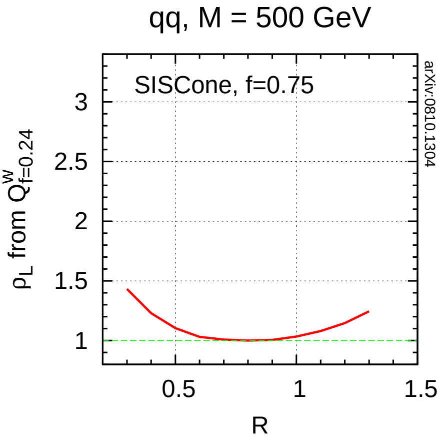
NB: 100,000 plots for various jet algorithms, narrow qq and gg resonances from `http://quality.fastjet.fr`

Cacciari, Rojo, GPS & Soyez '08

Other related work: Krohn, Thaler & Wang '09

Scan through $q\bar{q}$ mass values

$\mathbf{m_{qq} = 500}$ GeV



Best R is at minimum of curve

- Best R depends strongly on mass of system
- Increases with mass

can reproduce this anayltically

Soyez '10

Message received by CMS: they combine all $R = 0.5$ jets ($p_t >$ 10 GeV) within $\Delta R = 1.1$ of two hardest to improve resolution.

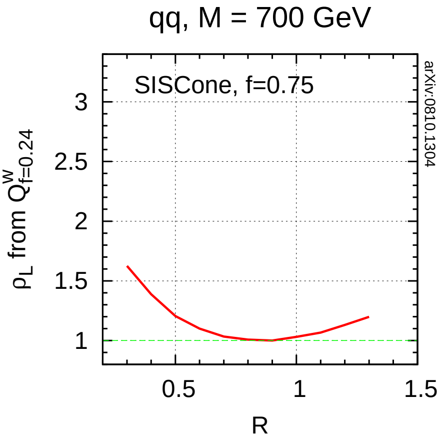
ATLAS '11 still just use $R = 0.6$

NB: 100,000 plots for various jet algorithms, narrow qq and gg resonances from `http://quality.fastjet.fr` Cacciari, Rojo, GPS & Soyez '08

Other related work: Krohn, Thaler & Wang '09

Scan through $q\bar{q}$ mass values

$\mathbf{m_{qq} = 700\ GeV}$



Best R is at minimum of curve

- Best R depends strongly on mass of system
- Increases with mass

can reproduce this anayltically Soyez '10

Message received by CMS: they combine all $R = 0.5$ jets ($p_t >$ 10 GeV) within $\Delta R = 1.1$ of two hardest to improve resolution.

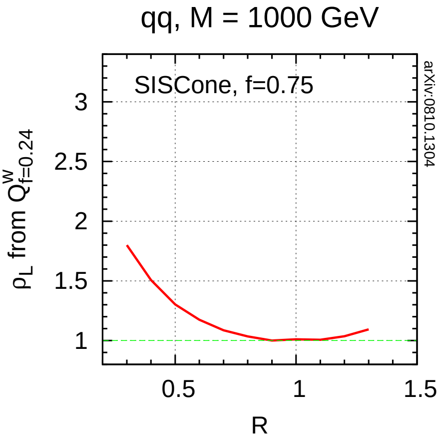
ATLAS '11 still just use $R = 0.6$

NB: 100,000 plots for various jet algorithms, narrow qq and gg resonances from `http://quality.fastjet.fr` Cacciari, Rojo, GPS & Soyez '08

Other related work: Krohn, Thaler & Wang '09

Scan through $q\bar{q}$ mass values

$m_{qq} = 1000$ GeV



Best R is at minimum of curve

- Best R depends strongly on mass of system
- Increases with mass

can reproduce this anayltically

Soyez '10

Message received by CMS: they combine all $R = 0.5$ jets ($p_t > 10$ GeV) within $\Delta R = 1.1$ of two hardest to improve resolution.

ATLAS '11 still just use $R = 0.6$

NB: 100,000 plots for various jet algorithms, narrow qq and gg resonances from `http://quality.fastjet.fr`

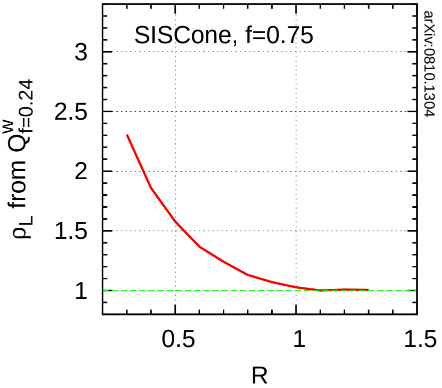
Cacciari, Rojo, GPS & Soyez '08

Other related work: Krohn, Thaler & Wang '09

Scan through $q\bar{q}$ mass values

$m_{qq} = 2000$ GeV

qq, M = 2000 GeV



Best R is at minimum of curve

- Best R depends strongly on mass of system
- Increases with mass

can reproduce this anayltically
Soyez '10

Message received by CMS: they combine all $R = 0.5$ jets ($p_t > 10$ GeV) within $\Delta R = 1.1$ of two hardest to improve resolution.

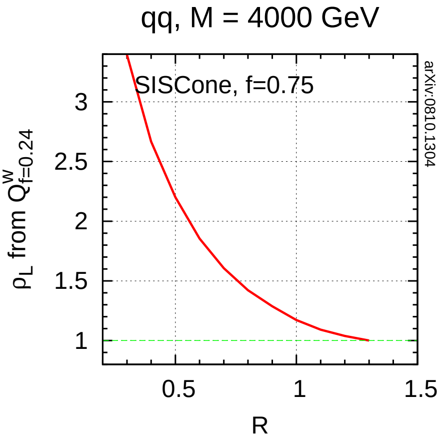
ATLAS '11 still just use $R = 0.6$

NB: 100,000 plots for various jet algorithms, narrow qq and gg resonances from `http://quality.fastjet.fr` Cacciari, Rojo, GPS & Soyez '08

Other related work: Krohn, Thaler & Wang '09

Scan through $q\bar{q}$ mass values

$m_{qq} = 4000$ GeV



Best R is at minimum of curve

- Best R depends strongly on mass of system
- Increases with mass

can reproduce this anayltically

Soyez '10

Message received by CMS: they combine all $R = 0.5$ jets ($p_t > 10$ GeV) within $\Delta R = 1.1$ of two hardest to improve resolution.

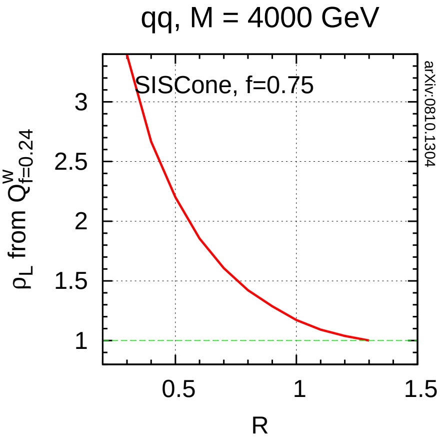
ATLAS '11 still just use $R = 0.6$

NB: 100,000 plots for various jet algorithms, narrow qq and gg resonances from `http://quality.fastjet.fr` Cacciari, Rojo, GPS & Soyez '08

Other related work: Krohn, Thaler & Wang '09

Scan through $q\bar{q}$ mass values

$\mathbf{m_{qq} = 4000}$ GeV



Best R is at minimum of curve

- Best R depends strongly on mass of system
- Increases with mass

can reproduce this anayltically
Soyez '10

Message received by CMS: they combine all $R = 0.5$ jets ($p_t > 10$ GeV) within $\Delta R = 1.1$ of two hardest to improve resolution.

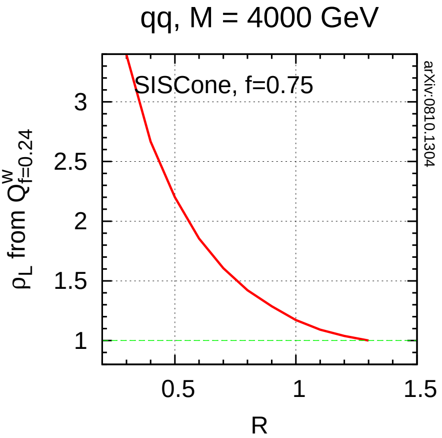
ATLAS '11 still just use $R = 0.6$

NB: 100,000 plots for various jet algorithms, narrow qq and gg resonances from `http://quality.fastjet.fr` Cacciari, Rojo, GPS & Soyez '08

Other related work: Krohn, Thaler & Wang '09

Scan through $q\bar{q}$ mass values

$m_{qq} = 4000$ GeV



Best R is at minimum of curve

- Best R depends strongly on mass of system
- Increases with mass

can reproduce this anayltically

Soyez '10

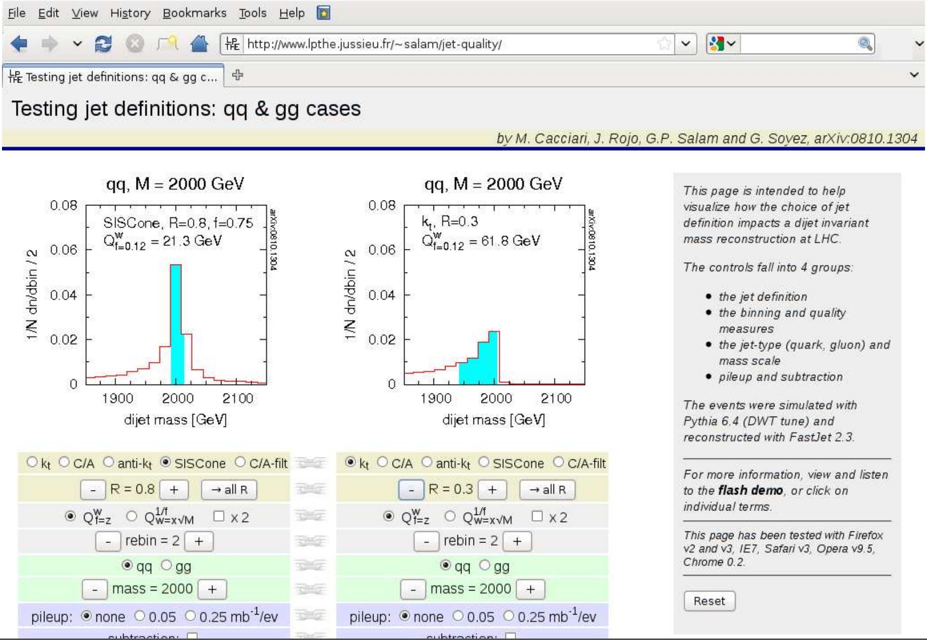
Message received by CMS: they combine all $R = 0.5$ jets ($p_t > 10$ GeV) within $\Delta R = 1.1$ of two hardest to improve resolution.

ATLAS '11 still just use $R = 0.6$

NB: 100,000 plots for various jet algorithms, narrow qq and gg resonances from `http://quality.fastjet.fr` Cacciari, Rojo, GPS & Soyez '08

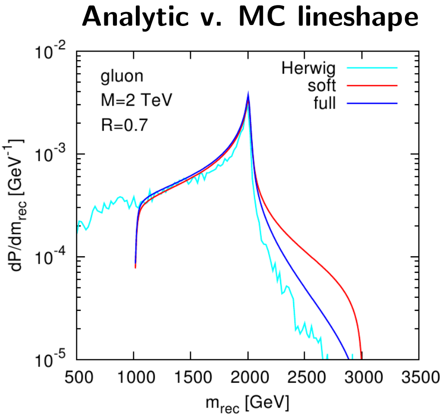
Other related work: Krohn, Thaler & Wang '09

http://quality.fastjet.fr/



Analytic quality estimates

Soyez '10

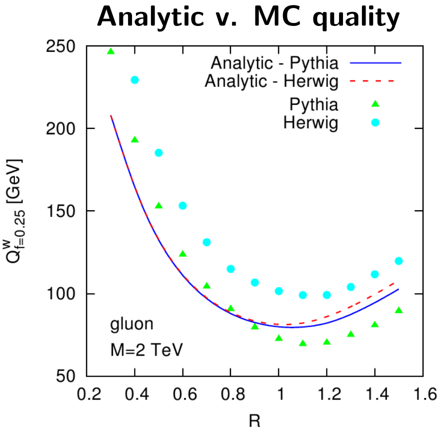


Perturbatively resum resonance "line-shape", convolute with model for non-perturbative effects.

etermine "quality" of line-shape from the analytic results, as a function of jet radius R

Analytic quality estimates

Soyez '10

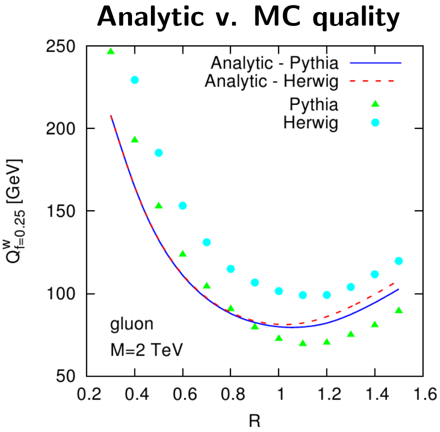


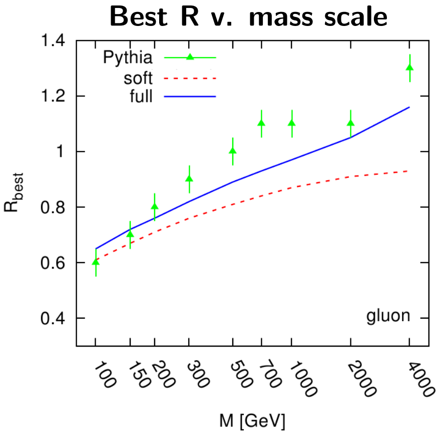
Perturbatively resum resonance "line-shape", convolute with model for non-perturbative effects.

Determine "quality" of line-shape from the analytic results, as a function of jet radius R

Analytic quality estimates

Soyez '10





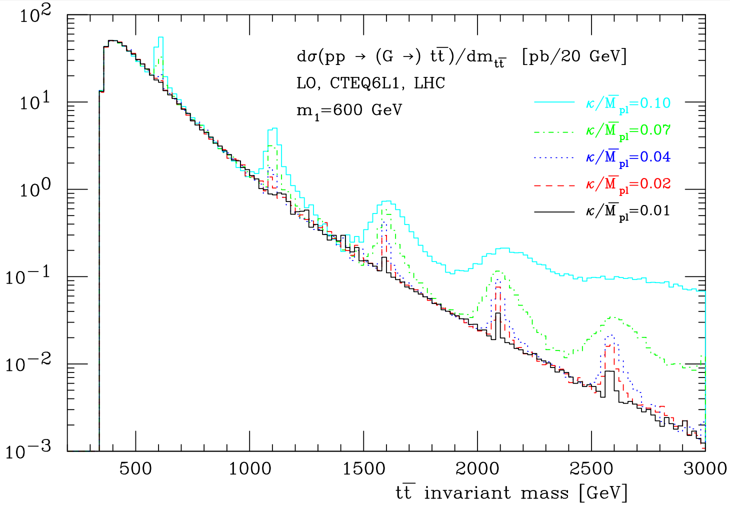
Fat jets

boosted massive hadronically decaying objects

E.g. when a known particle, W, Z or a top $\rightarrow$ a single jet
or a new particle, Higgs, gluino, neutralino $\rightarrow$ a single jet

This will be common for electroweak-scale objects at LHC:
$m_W, m_t \ll 14$ TeV

E.g. $X \to t\bar{t}$ resonances of varying difficulty

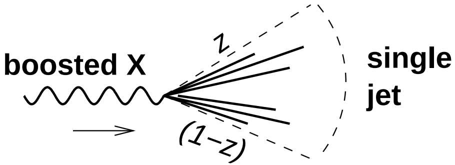


RS KK resonances $\to t\bar{t}$, from Frederix & Maltoni, 0712.2355

NB: QCD dijet spectrum is $\sim 10^3$ times $t\bar{t}$

Boosted massive particles, e.g.: EW bosons

Hadronically decaying EW boson at high $p_t \neq$ two jets



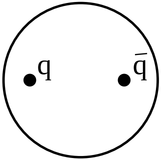
$$R \gtrsim \frac{m}{p_t} \frac{1}{\sqrt{z(1-z)}}$$

Rules of thumb:

$m = 100$ GeV, $p_t = 500$ GeV

- $R < \dfrac{2m}{p_t}$: always resolve **two** jets — $R < 0.4$
- $R \gtrsim \dfrac{3m}{p_t}$: resolve **one** jet in ∼75% of cases ($\frac{1}{8} < z < \frac{7}{8}$) — $R \gtrsim 0.6$

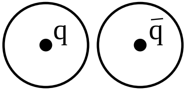
Boosted ID strategies



Select on the jet mass with one large (cone) jet

Can be subject to large bkgds

[high-p_t jets have significant masses]

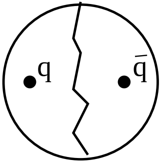


Choose a small jet size (R) so as to resolve two jets

Easier to reject background

if you actually see substructure

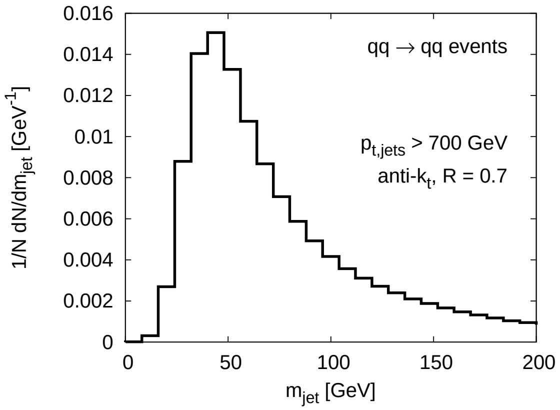
[NB: must manually put in "right" radius]



Take a large jet and split it in two

Let jet algorithm establish correct division

Jet masses



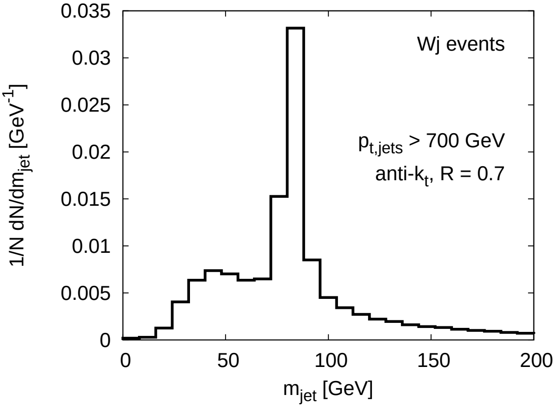
Look at jet mass distribution for two leading jets in

- $qq \to qq$ events
- $pp \to W +$ jet events
- a mixture of the two

In roughly sensible proportions

Jet mass gives clear sign of massive particles inside the jet;

Jet masses



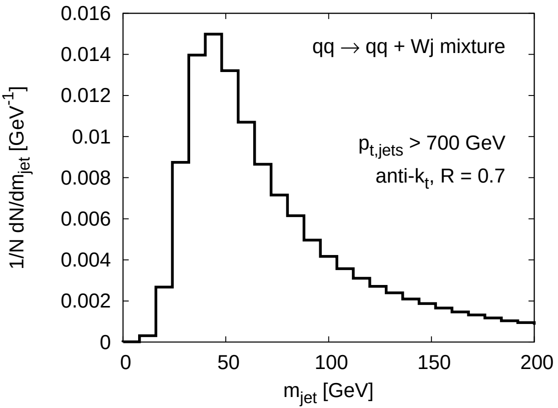
Look at jet mass distribution for two leading jets in

- $qq \to qq$ events
- $pp \to W +$ jet events
- a mixture of the two

In roughly sensible proportions

Jet mass gives clear sign of massive particles inside the jet;

Jet masses



Look at jet mass distribution for two leading jets in

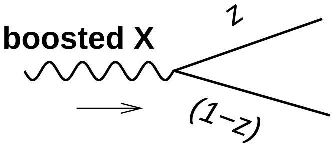
- $qq \to qq$ events
- $pp \to W +$ jet events
- a mixture of the two

In roughly sensible proportions

Jet mass gives clear sign of massive particles inside the jet; but QCD jets are massive too — must learn to reject them

QCD principle: soft divergence

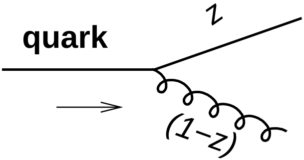
Signal



Splitting probability for Higgs:

$$P(z) \propto 1$$

Background



Splitting probability for quark:

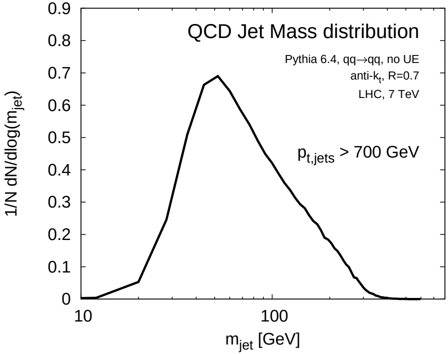
$$P(z) \propto \frac{1+z^2}{1-z}$$

$1/(1-z)$ divergence enhances background

Remove divergence in bkdg with cut on z
Can choose cut analytically so as to maximise $S/\sqrt{B}$

Originally: cut on (related) k_t-distance
Butterworth, Cox & Forshaw '02

Inside the jet mass



QCD jet mass distribution has the approximate

$$\frac{dN}{d\ln m} \sim \alpha_s \ln \frac{p_t R}{m} \times \text{Sudakov}$$

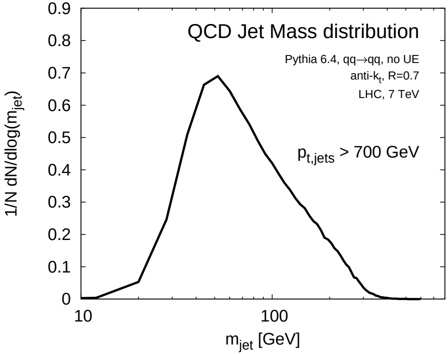
Work from '80s and '90s
+ Almeida et al '08

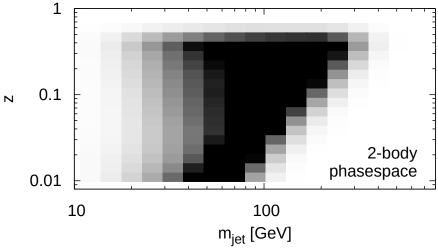
The logarithm comes from integral over soft divergence of QCD:

$$\int_{\frac{m^2}{p_t^2 R^2}}^{\frac{1}{2}} \frac{dz}{z}$$

A hard cut on z reduces QCD background & simplifies its shape

Inside the jet mass





QCD jet mass distribution has the approximate

$$\frac{dN}{d\ln m} \sim \alpha_s \ln \frac{p_t R}{m} \times \text{Sudakov}$$

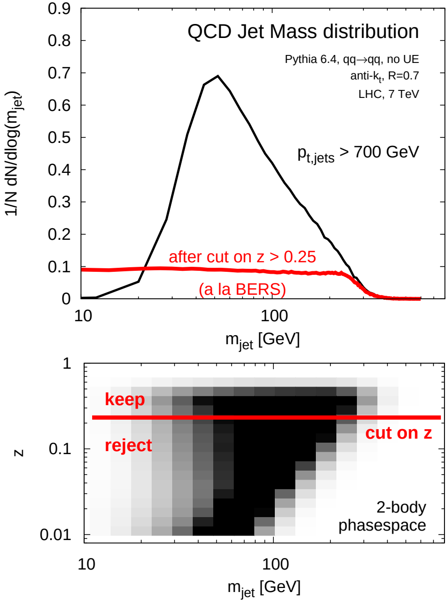
Work from '80s and '90s
+ Almeida et al '08

The logarithm comes from integral over soft divergence of QCD:

$$\int_{\frac{m^2}{p_t^2 R^2}}^{\frac{1}{2}} \frac{dz}{z}$$

A hard cut on z reduces QCD background & simplifies its shape

Inside the jet mass



QCD jet mass distribution has the approximate

$$\frac{dN}{d\ln m} \sim \alpha_s \ln \frac{p_t R}{m} \times \text{Sudakov}$$

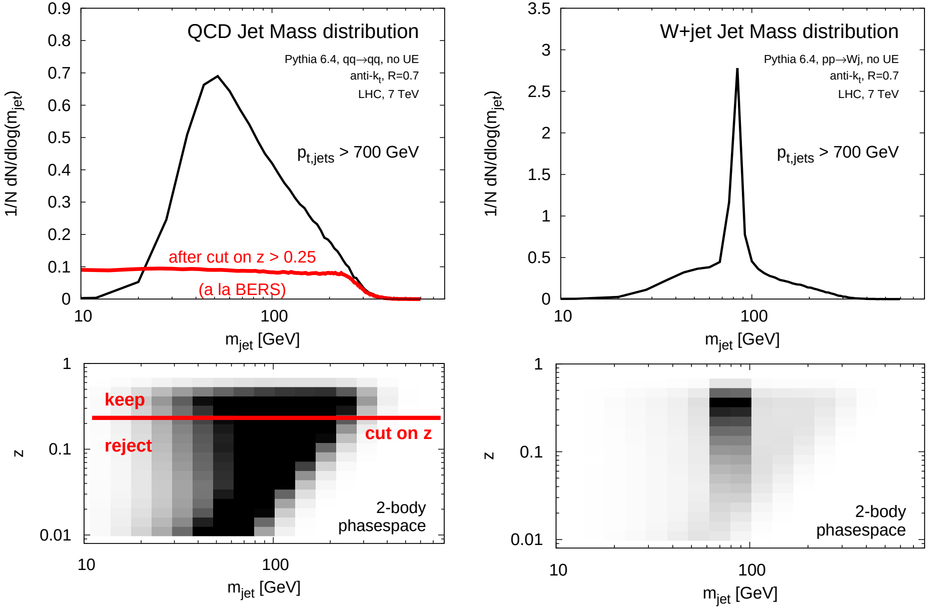
Work from '80s and '90s
+ Almeida et al '08

The logarithm comes from integral over soft divergence of QCD:

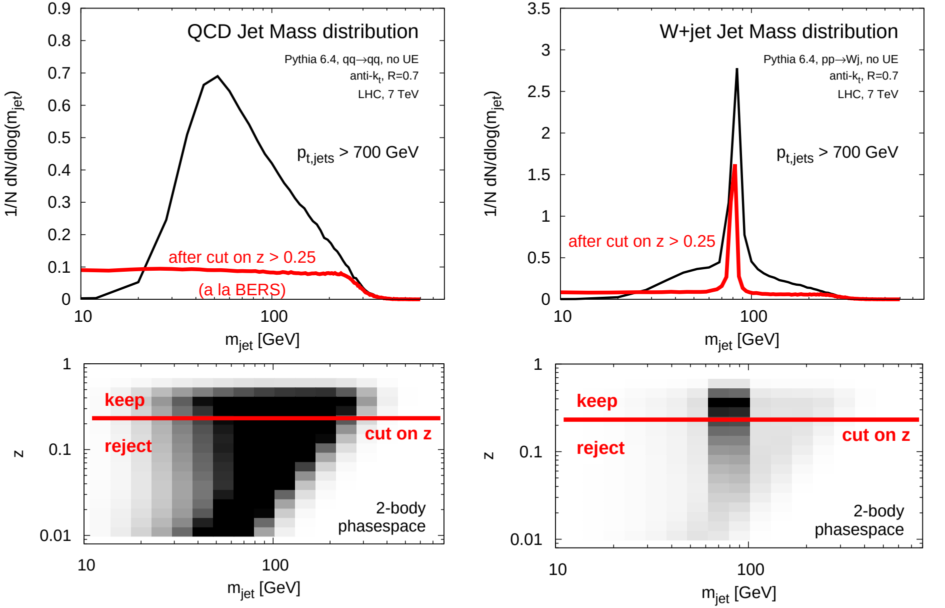
$$\int_{\frac{m^2}{p_t^2 R^2}}^{\frac{1}{2}} \frac{dz}{z}$$

A hard cut on z reduces QCD background & simplifies its shape

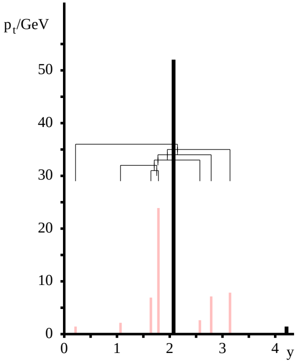
Inside the jet mass



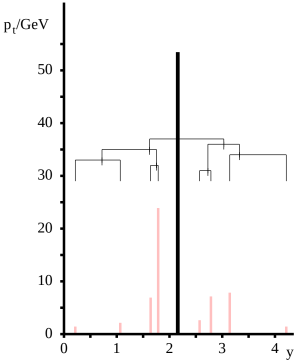
Inside the jet mass

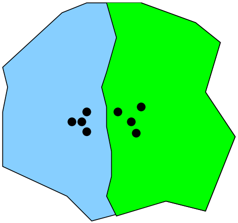


anti-k_t algorithm

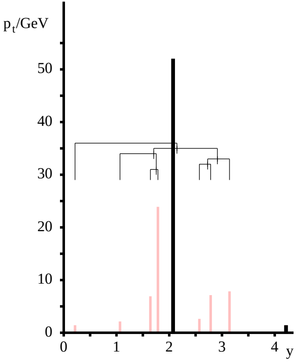


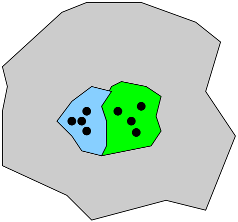
k_t algorithm





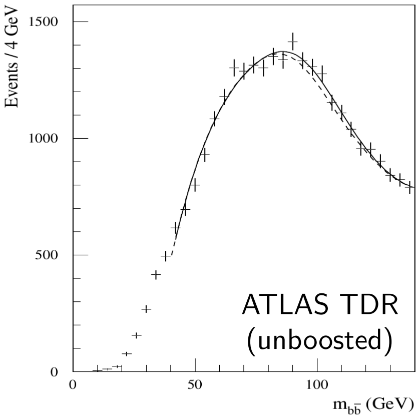
Cambridge/Aachen

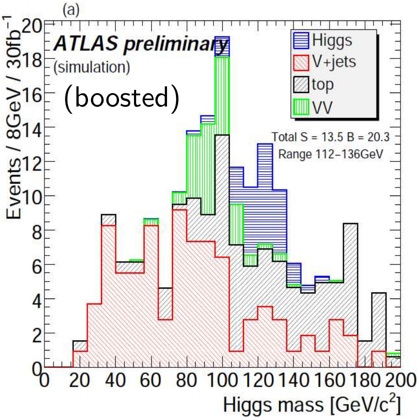




Example improvement from boosted regime

Search for main decay of light Higgs boson in W/Z+H, H $\to$ b$\bar{b}$



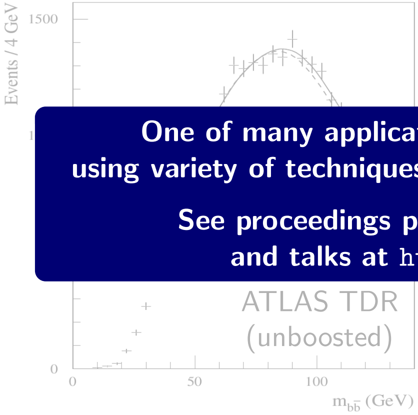


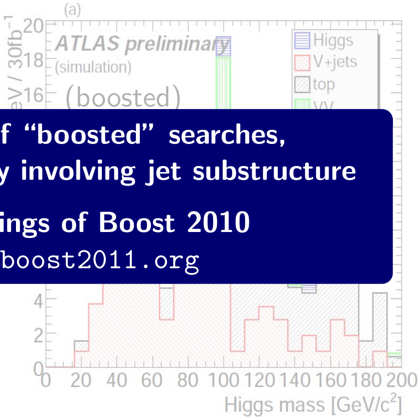
restricting search to $p_{tH} > 200$ GeV

using the method from Butterworth, Davison, Rubin & GPS '08

Example improvement from boosted regime

Search for main decay of light Higgs boson in W/Z+H, H $\to b\bar{b}$





One of many applications of "boosted" searches, using variety of techniques, many involving jet substructure

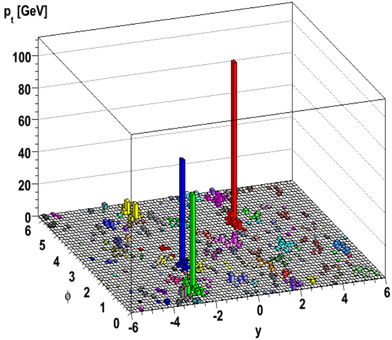
See proceedings proceedings of Boost 2010 and talks at `http://boost2011.org`

restricting search to $p_{tH} > 200$ GeV

using the method from Butterworth, Davison, Rubin & GPS '08

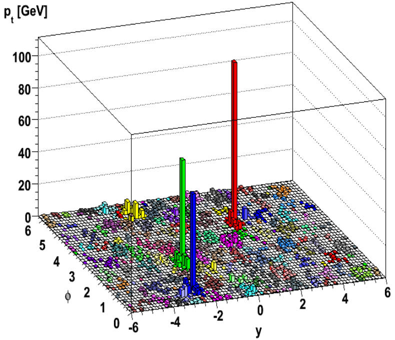
Noise removal from jets

Plain pythia event

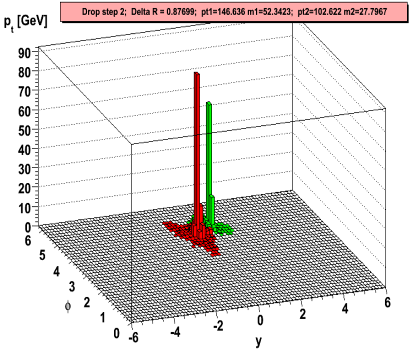


Noise removal from jets

Plain pythia event + 10 pileup



Noise removal from jets



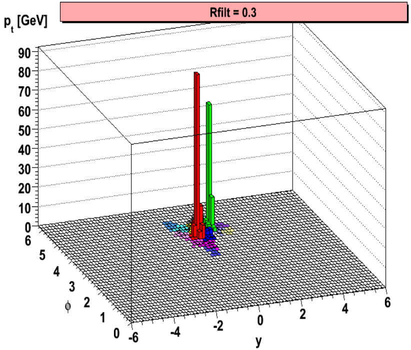
Key idea:

- Look at jet on smaller angular scale
- Discard its softer parts

- Filtering — Butterworth et al '08
- Pruning — Ellis, Vermillion and Walsh '09
- Trimming — Krohn, Thaler & Wang '09

[With earlier methods by Seymour '93 and Kodolova et al '07; Rubin '10 for filtering optimisation; also Soper & Spannowsky '10, '11]

Noise removal from jets



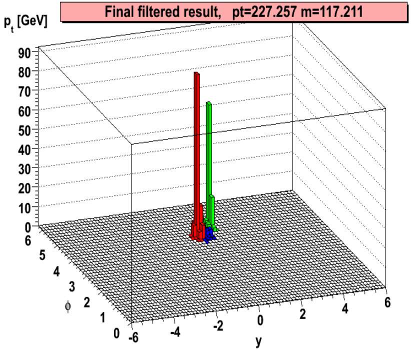
Key idea:

- Look at jet on smaller angular scale
- Discard its softer parts

- Filtering — Butterworth et al '08
- Pruning — Ellis, Vermillion and Walsh '09
- Trimming — Krohn, Thaler & Wang '09

[With earlier methods by Seymour '93 and Kodolova et al '07; Rubin '10 for filtering optimisation; also Soper & Spannowsky '10, '11]

Noise removal from jets



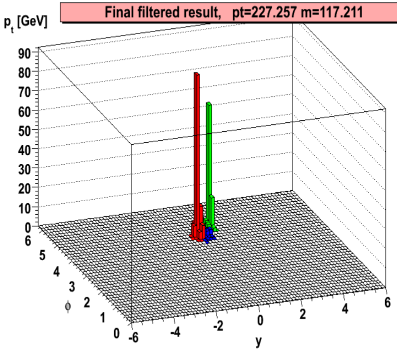
Key idea:

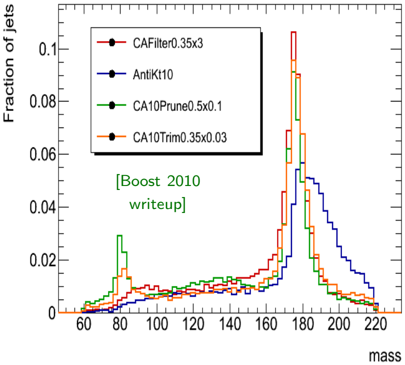
- Look at jet on smaller angular scale
- Discard its softer parts

- Filtering — Butterworth et al '08
- Pruning — Ellis, Vermillion and Walsh '09
- Trimming — Krohn, Thaler & Wang '09

[With earlier methods by Seymour '93 and Kodolova et al '07; Rubin '10 for filtering optimisation; also Soper & Spannowsky '10, '11]

Noise removal from jets

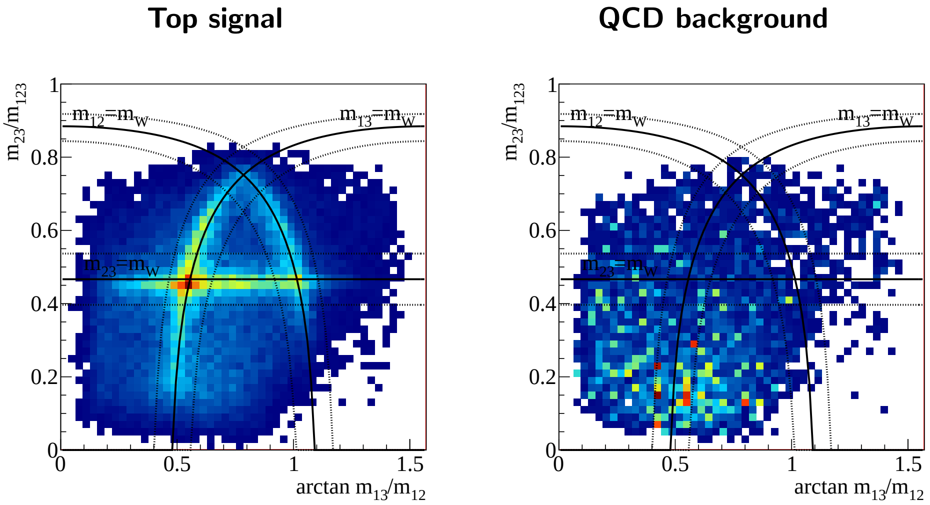




- Filtering — Butterworth et al '08
- Pruning — Ellis, Vermillion and Walsh '09
- Trimming — Krohn, Thaler & Wang '09

[With earlier methods by Seymour '93 and Kodolova et al '07; Rubin '10 for filtering optimisation; also Soper & Spannowsky '10, '11]

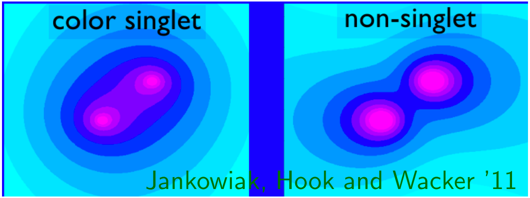
$t \to bjj$ decays — multiple kinematic variables



Using Dalitz-like plots to pull out the W **without using *b* tagging**

Plehn, Spannowsky, Takeuchi & Zerwas '10

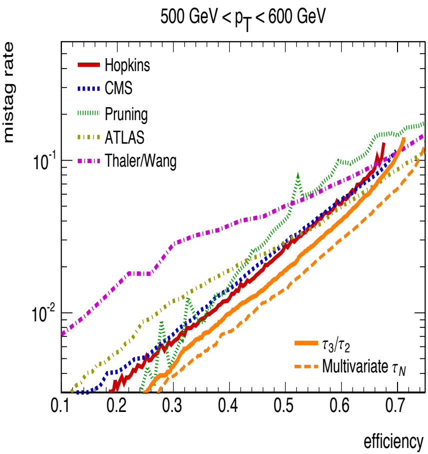
Exploiting energy-flow beyond LO structure



Background (e.g. $g \to gg$) and signal (e.g. $W \to q\bar{q}$) often have different colour structure $\to$ **different radiation patterns**.

- Pull (non-boosted context) — Gallicchio & Schwartz '10
- N-subjettiness — Jihun Kim '10; Thaler & Van Tilburg '10
- "Buried Higgs" light singlets — Falkowski et al '10; Chen et al '10
- Boosted decision trees — Cui & Schwartz '10
- Dipolarity, applied to HEPTopTagger — Jankowiak, Hook and Wacker '11
- Jet deconstruction — Soper & Spannowsky '11
- Template method beyond LO — Almeida et al '11
- ...

Jet shape variables (here for top tagging)



Thaler & van Tilburg '11
cf. also J.-H. Kim '10 for Higgs

Early proposals include planar flow (3- v. 2-body structure of top decay) Thaler & Wang '08 Almeida et al '08

Recent try: N-subjettiness. Break jet into subjets $1, \ldots, N$

$$\tau_N = \frac{1}{p_{t,jet}} \sum_i p_{ti} \min(R_{i1}, \ldots, R_{iN})$$

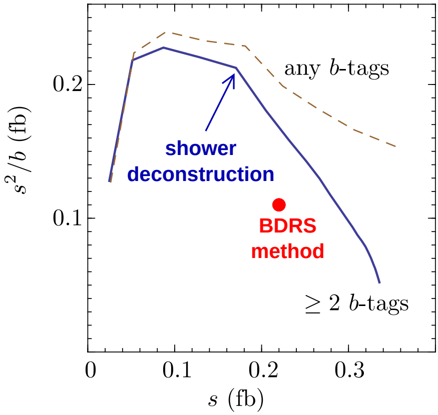
N-pronged decay: cut on mass &

$$\frac{\tau_N}{\tau_{N-1}}$$

Combines constraints on LO structure (energy sharing among prongs) and higher-order rad^n (from quarks in signal v. gluons in bkgd)

Shower deconstruction for ZH

Matrix-element method on steroids



For each event estimate the probability that event is signal-like or background like.

Break event into many mini-jets; use Monte-Carlo type Sudakovs and splitting functions to get estimate of multi-parton matrix element for S & B hypotheses.

Intelligently combines full info about LO splitting, radiation, b-tags, etc.

Soper & Spannowsky '11

cf. also multivariate (BDT) type methods from Cui & Schwartz '10

Closing remarks

The study of jets is an increasingly broad subject.

Driven by the ubiquity of jets in LHC studies, the greater flexibility of LHC detectors relative to Tevatron ones, and by the broad dynamic range of LHC.

[jets from 20 GeV to several TeV — from below to far above EW scale; UE $\sim$ 1 GeV per unit area, pileup $10-20$ GeV, HIC $100-300$ GeV]

Many basic ideas ideas have been worked out.

But intense recent activity suggests that there is still scope for significant further progress.

Supplementary material

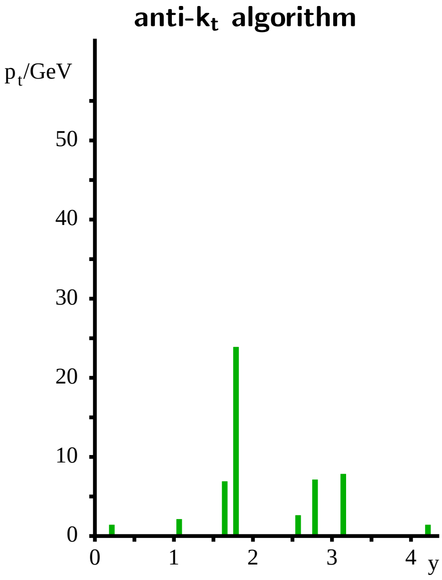
Jet algorithm properties: summary

	k_t	Cam/Aachen	anti-k_t	SISCone
reach	R	R	R	$(1 + \frac{p_{t2}}{p_{t1}})R$
$\Delta p_{t,PT} \simeq \frac{\alpha_s C_i}{\pi} \times$	$\ln R$	$\ln R$	$\ln R$	$\ln 1.35R$
$\Delta p_{t,hadr} \simeq -\frac{0.4\ \text{GeV}\, C_i}{R} \times$	0.7	?	1	?
area $= \pi R^2 \times$	0.81 ± 0.28	0.81 ± 0.26	1	0.25
$+\pi R^2 \frac{C_i}{\pi b_0} \ln \frac{\alpha_s(Q_0)}{\alpha_s(R p_t)} \times$	0.52 ± 0.41	0.08 ± 0.19	0	0.12 ± 0.07

In words:

- k_t: area fluctuates a lot, depends on p_t (bad for UE)
- Cam/Aachen: area fluctuates somewhat, depends less on p_t
- anti-k_t: area is constant (circular jets)
- SISCone: reaches far for hard radiation (good for resolution, bad for multijets), area is smaller (good for UE)

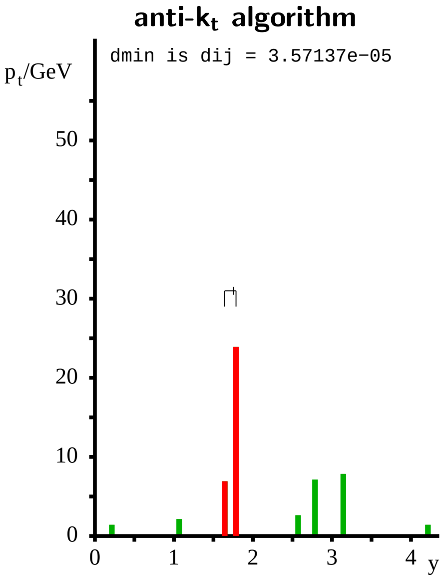
Identifying jet substructure: try out anti-k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

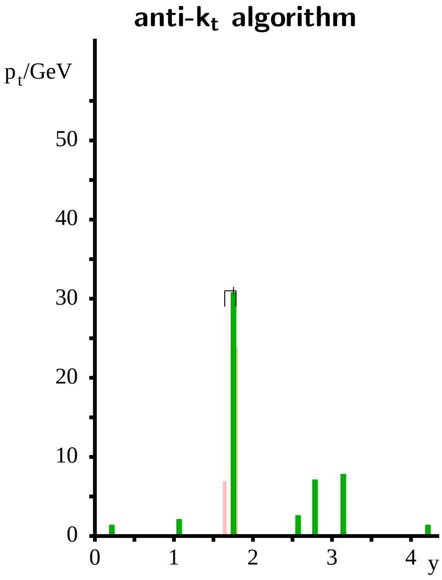
Identifying jet substructure: try out anti-k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

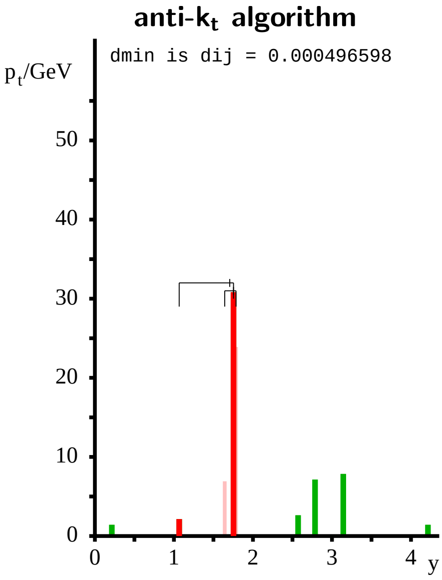
Identifying jet substructure: try out anti-k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

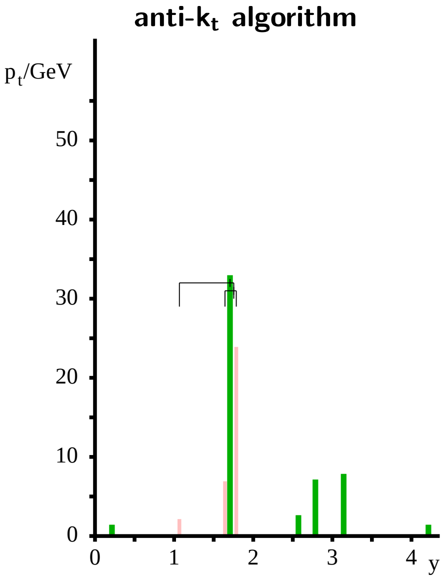
Identifying jet substructure: try out anti-k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

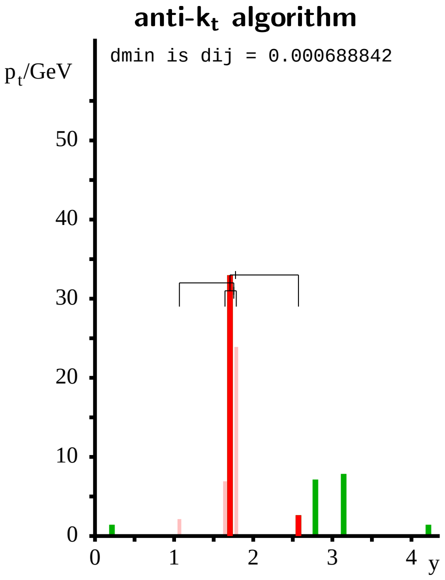
Identifying jet substructure: try out anti-k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

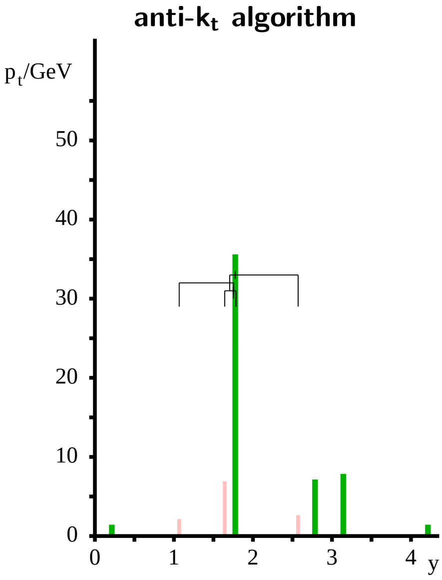
Identifying jet substructure: try out anti-k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

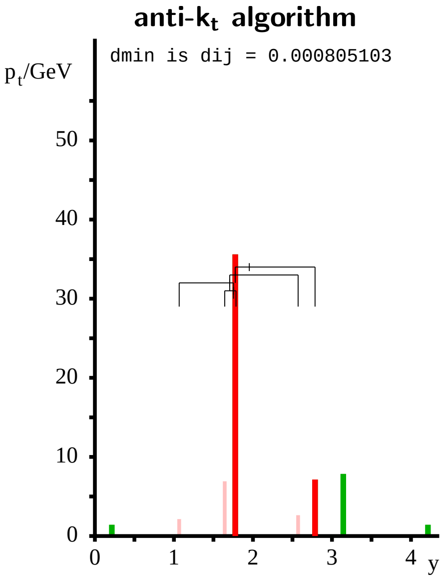
Identifying jet substructure: try out anti-k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

Identifying jet substructure: try out anti-k_t

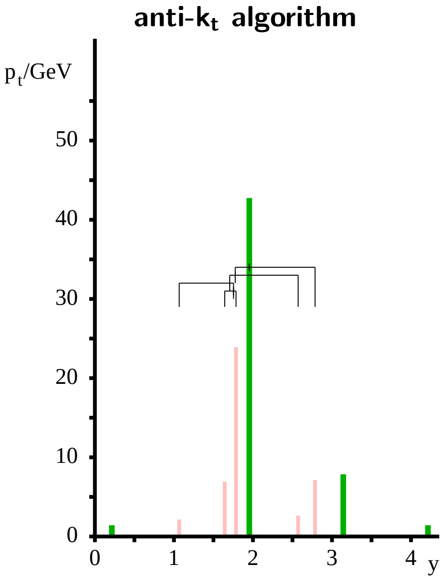


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

Anti-k_t gradually makes its way through the secondary blob → no clear identification of substructure associated with 2nd parton.

Identifying jet substructure: try out anti-k_t

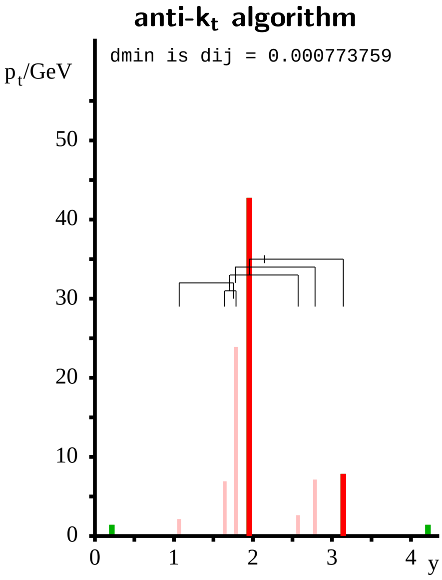


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

Anti-k_t gradually makes its way through the secondary blob $\rightarrow$ no clear identification of substructure associated with 2nd parton.

Identifying jet substructure: try out anti-k_t

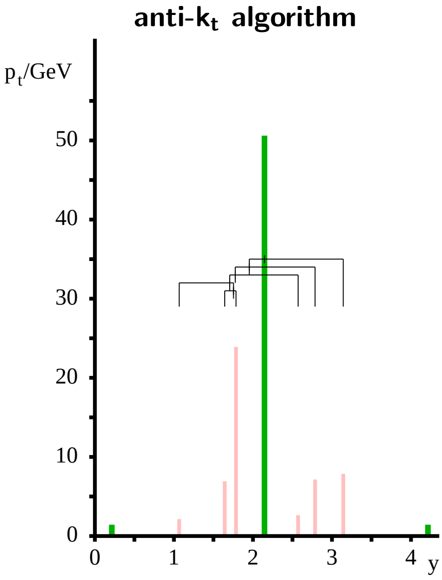


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

Anti-k_t gradually makes its way through the secondary blob → no clear identification of substructure associated with 2nd parton.

Identifying jet substructure: try out anti-k_t

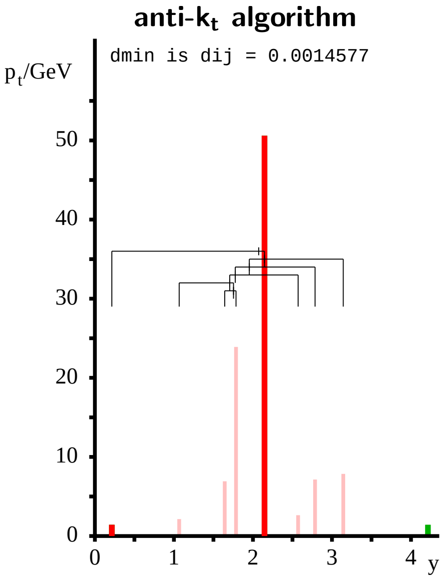


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

Anti-k_t gradually makes its way through the secondary blob $\rightarrow$ no clear identification of substructure associated with 2nd parton.

Identifying jet substructure: try out anti-k_t

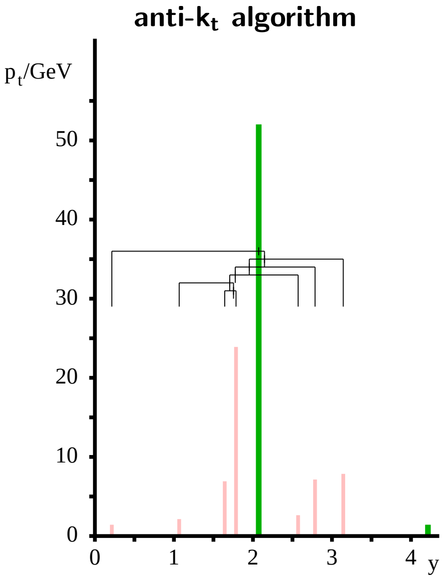


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

Anti-k_t gradually makes its way through the secondary blob → no clear identification of substructure associated with 2nd parton.

Identifying jet substructure: try out anti-k_t

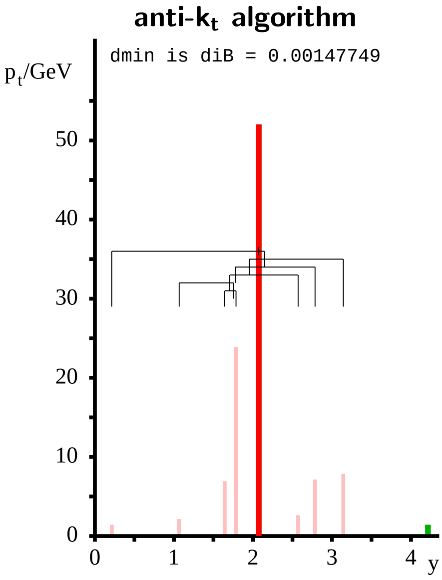


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

Anti-k_t gradually makes its way through the secondary blob → no clear identification of substructure associated with 2nd parton.

Identifying jet substructure: try out anti-k_t

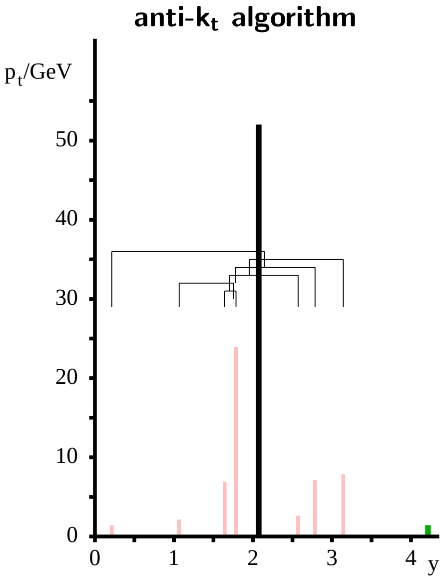


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

Anti-k_t gradually makes its way through the secondary blob $\rightarrow$ no clear identification of substructure associated with 2nd parton.

Identifying jet substructure: try out anti-k_t

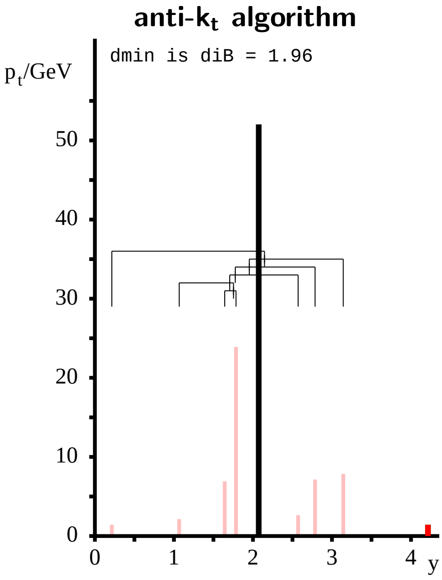


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

Anti-k_t gradually makes its way through the secondary blob $\rightarrow$ no clear identification of substructure associated with 2nd parton.

Identifying jet substructure: try out anti-k_t

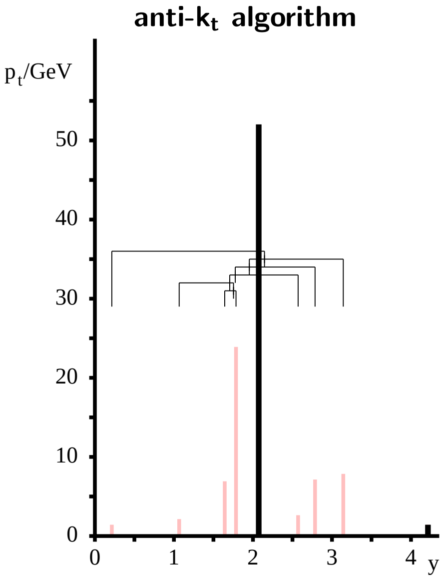


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

Anti-k_t gradually makes its way through the secondary blob → no clear identification of substructure associated with 2nd parton.

Identifying jet substructure: try out anti-k_t

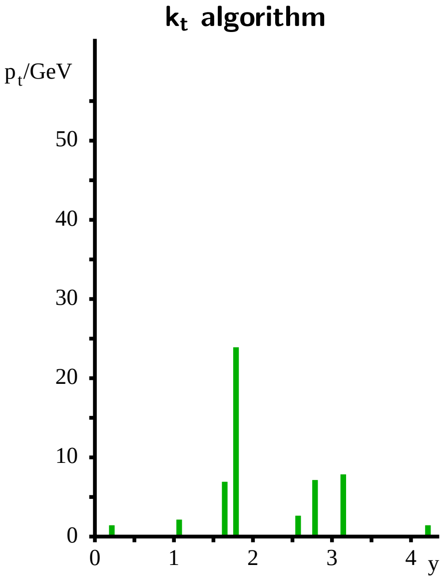


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

Anti-k_t gradually makes its way through the secondary blob $\rightarrow$ no clear identification of substructure associated with 2nd parton.

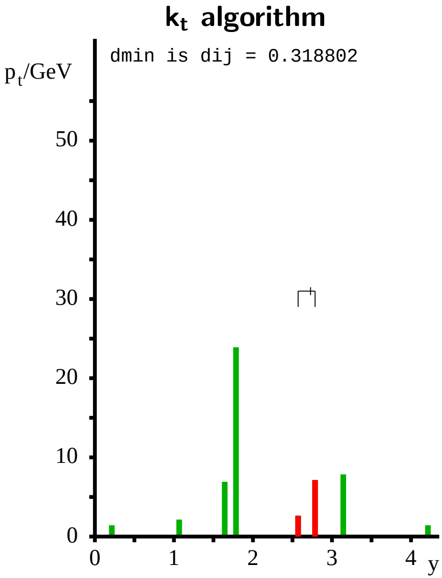
Identifying jet substructure: try out k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

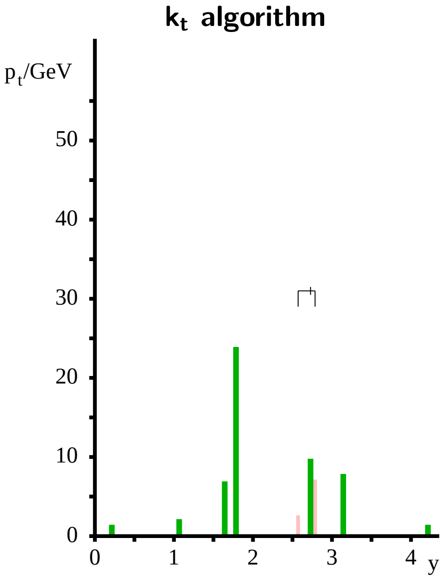
Identifying jet substructure: try out k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

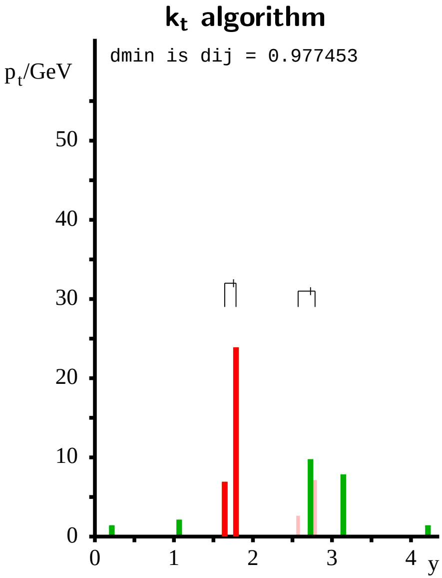
Identifying jet substructure: try out k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

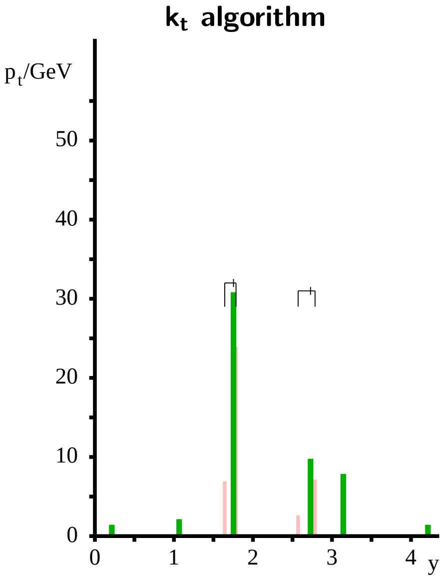
Identifying jet substructure: try out k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

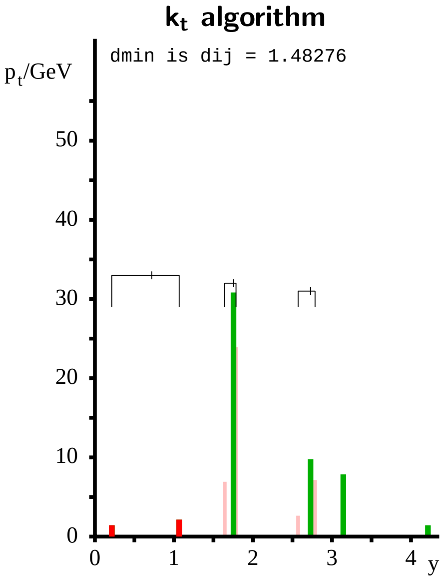
Identifying jet substructure: try out k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

Identifying jet substructure: try out k_t

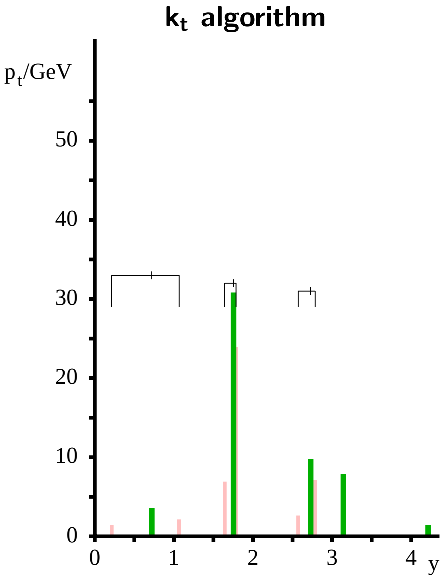


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

k_t clusters soft "junk" early on in the clustering

Identifying jet substructure: try out k_t

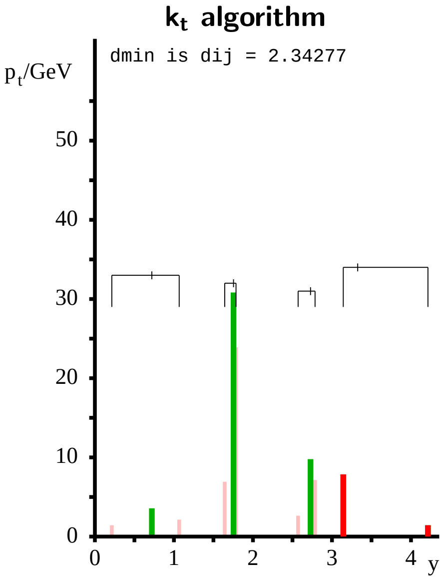


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

k_t clusters soft "junk" early on in the clustering

Identifying jet substructure: try out k_t

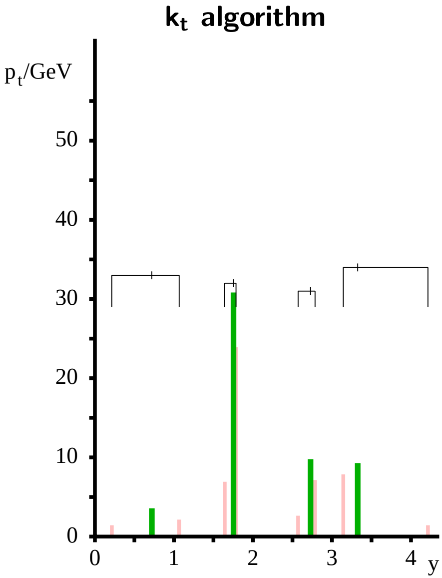


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

k_t clusters soft "junk" early on in the clustering

Identifying jet substructure: try out k_t

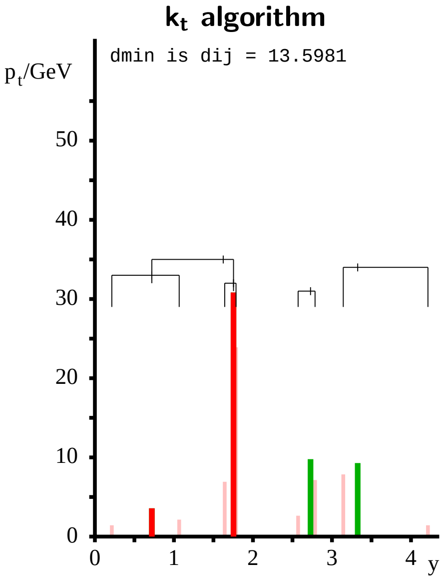


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

k_t clusters soft "junk" early on in the clustering

Identifying jet substructure: try out k_t

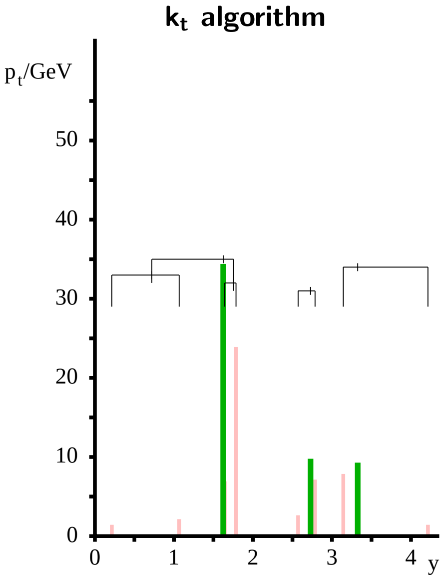


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

k_t clusters soft "junk" early on in the clustering

Identifying jet substructure: try out k_t

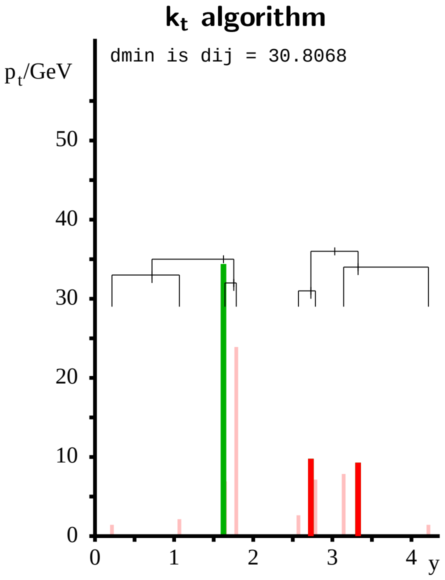


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

k_t clusters soft "junk" early on in the clustering

Identifying jet substructure: try out k_t

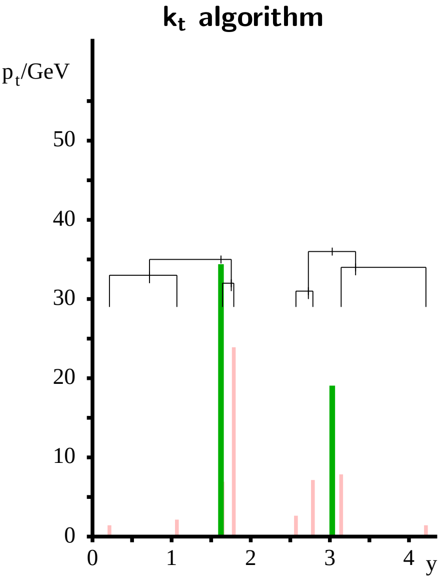


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

k_t clusters soft "junk" early on in the clustering

Identifying jet substructure: try out k_t

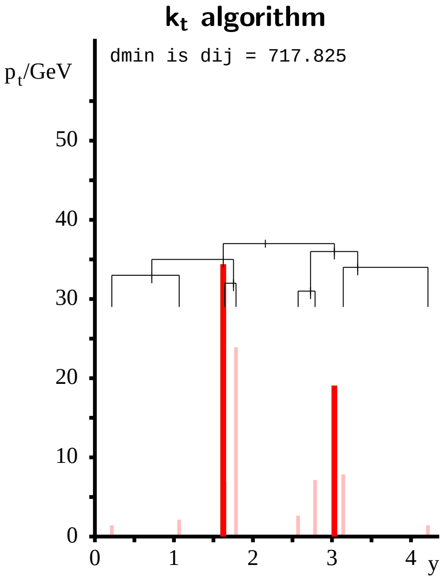


How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

k_t clusters soft "junk" early on in the clustering

Identifying jet substructure: try out k_t



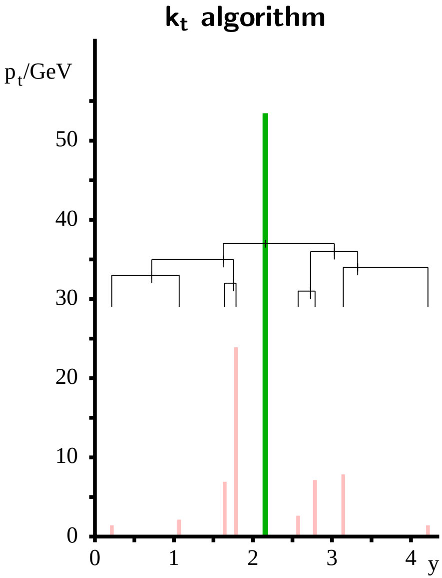
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

k_t clusters soft "junk" early on in the clustering

Its last step is to merge two hard pieces. Easily undone to identify underlying kinematics

Identifying jet substructure: try out k_t



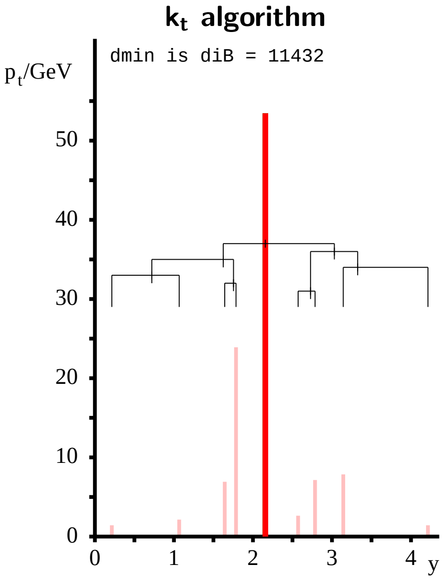
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

k_t clusters soft "junk" early on in the clustering

Its last step is to merge two hard pieces. Easily undone to identify underlying kinematics

Identifying jet substructure: try out k_t



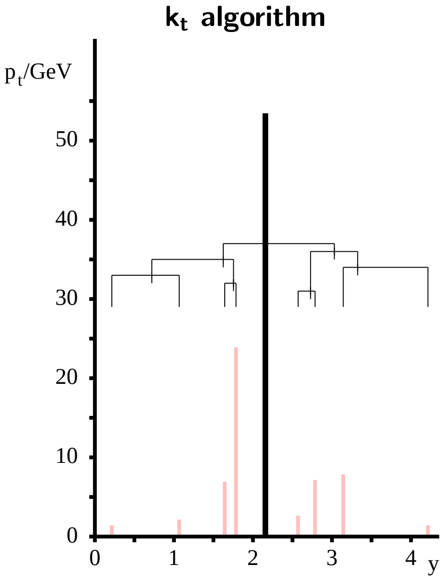
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

k_t clusters soft "junk" early on in the clustering

Its last step is to merge two hard pieces. Easily undone to identify underlying kinematics

Identifying jet substructure: try out k_t



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

This is crucial for identifying the kinematic variables of the partons in the jet (e.g. z).

k_t clusters soft "junk" early on in the clustering

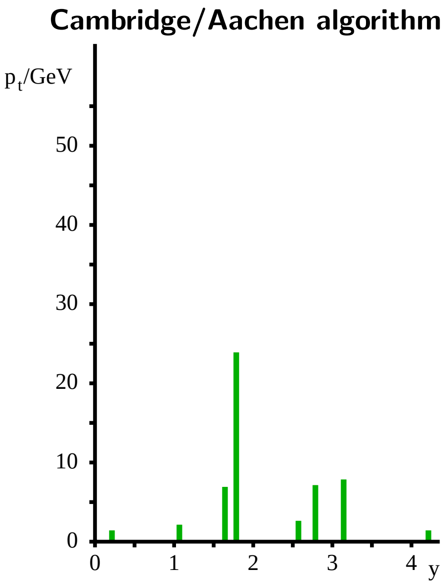
Its last step is to merge two hard pieces. Easily undone to identify underlying kinematics

This meant it was the first algorithm to be used for jet substructure.

Seymour '93

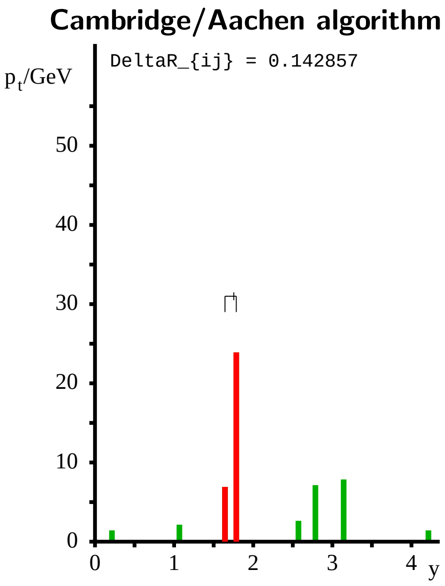
Butterworth, Cox & Forshaw '02

Identifying jet substructure: Cam/Aachen



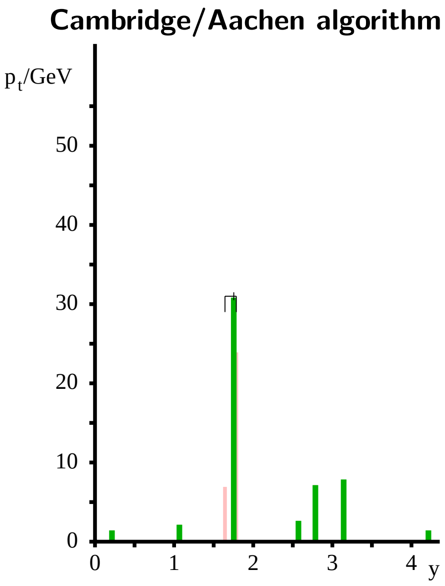
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

Identifying jet substructure: Cam/Aachen



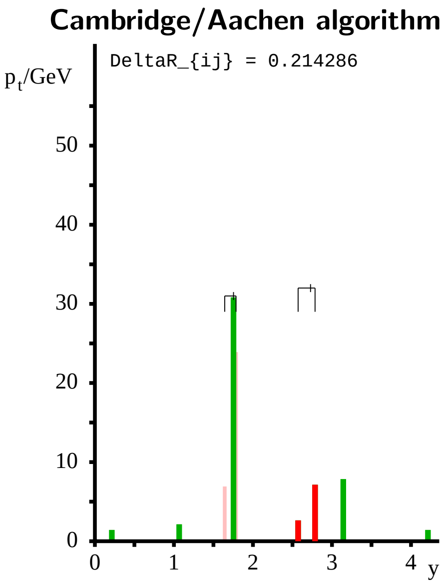
How well can an algorithm identify the “blobs” of energy inside a jet that come from different partons?

Identifying jet substructure: Cam/Aachen



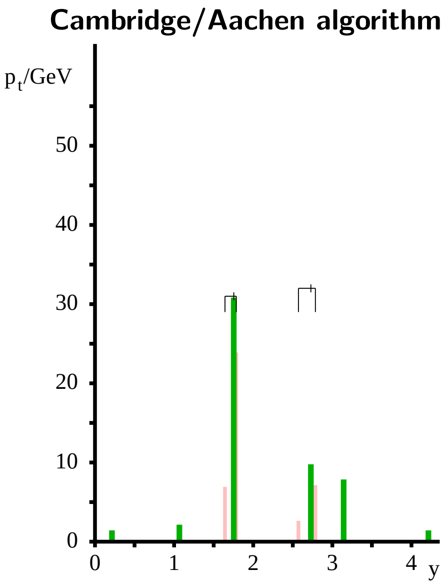
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

Identifying jet substructure: Cam/Aachen



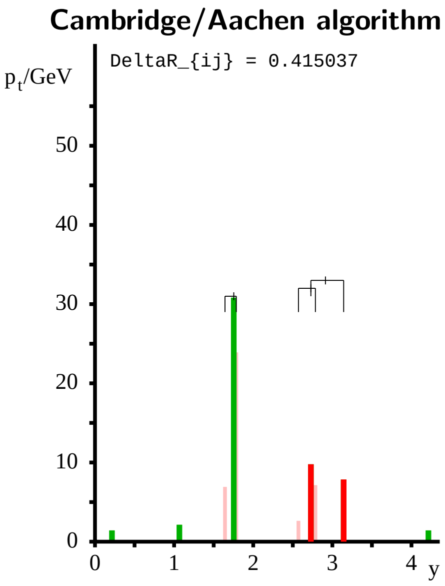
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

Identifying jet substructure: Cam/Aachen



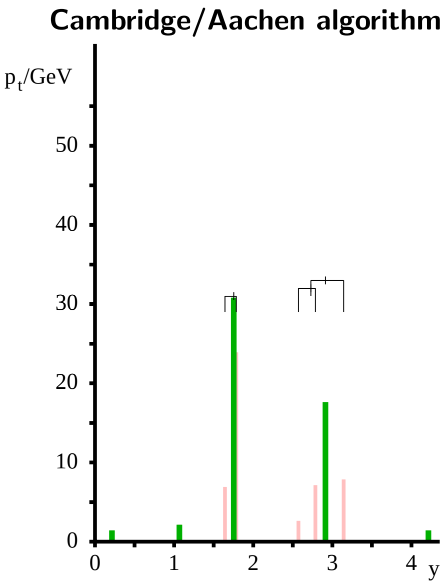
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

Identifying jet substructure: Cam/Aachen



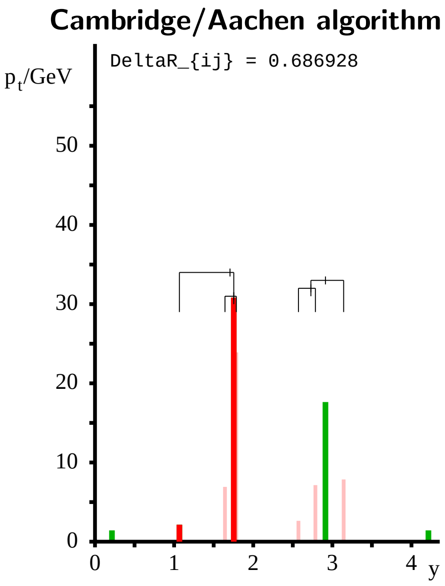
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

Identifying jet substructure: Cam/Aachen



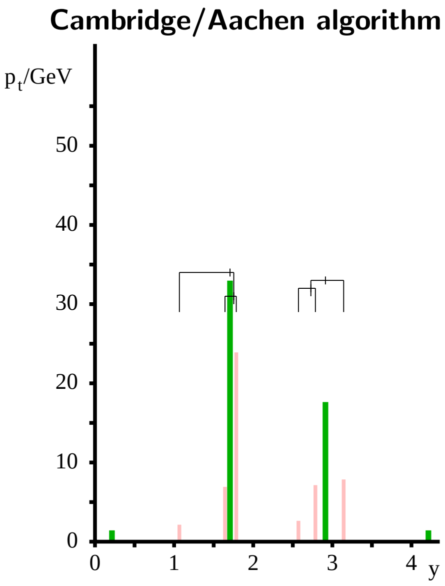
How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

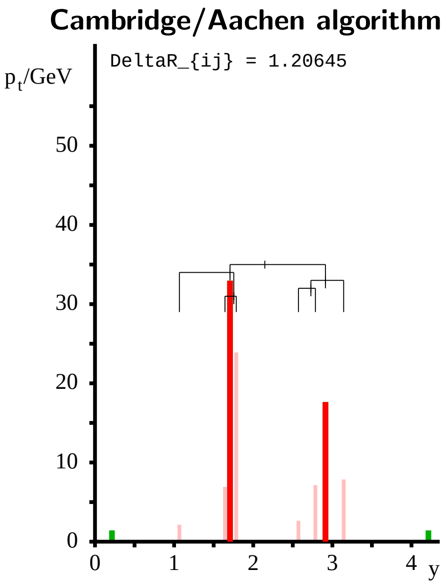
Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A identifies two hard blobs with limited soft contamination

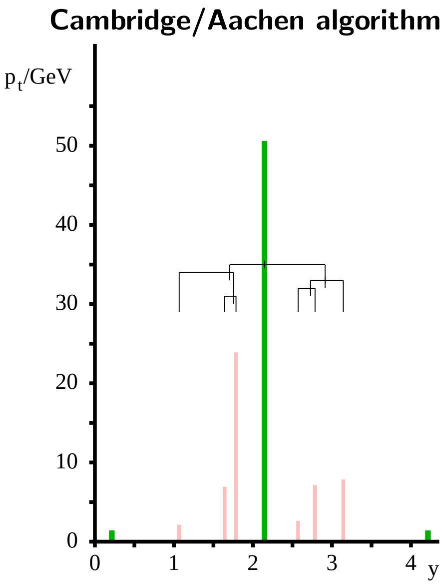
Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A identifies two hard blobs with limited soft contamination, joins them

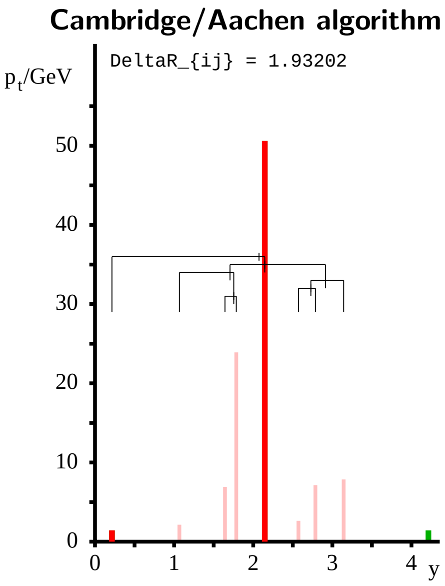
Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A identifies two hard blobs with limited soft contamination, joins them

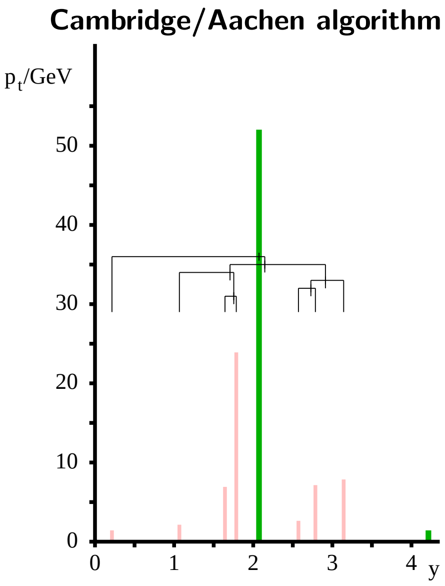
Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A identifies two hard blobs with limited soft contamination, joins them, and then adds in remaining soft junk

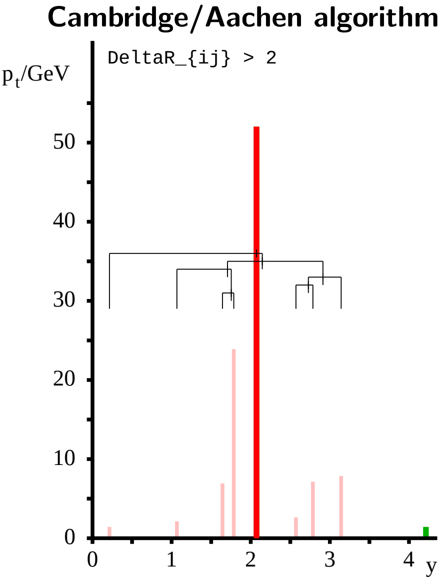
Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A identifies two hard blobs with limited soft contamination, joins them, and then adds in remaining soft junk

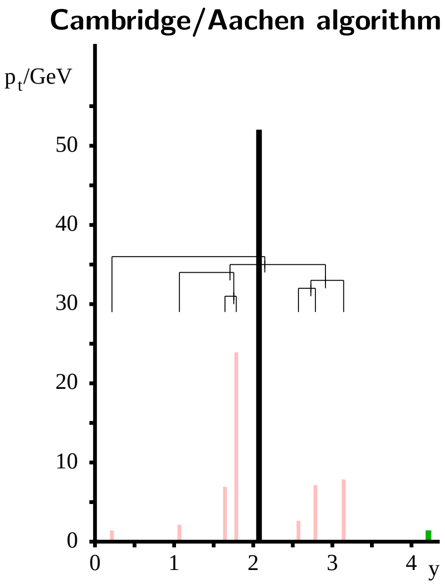
Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A identifies two hard blobs with limited soft contamination, joins them, and then adds in remaining soft junk

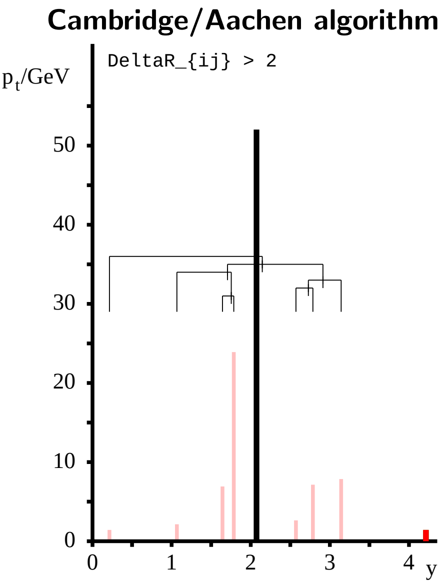
Identifying jet substructure: Cam/Aachen



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

C/A identifies two hard blobs with limited soft contamination, joins them, and then adds in remaining soft junk

Identifying jet substructure: Cam/Aachen

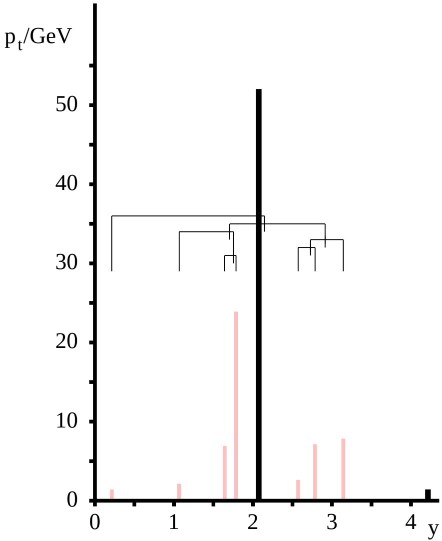


How well can an algorithm identify the “blobs” of energy inside a jet that come from different partons?

C/A identifies two hard blobs with limited soft contamination, joins them, and then adds in remaining soft junk

Identifying jet substructure: Cam/Aachen

Cambridge/Aachen algorithm



How well can an algorithm identify the "blobs" of energy inside a jet that come from different partons?

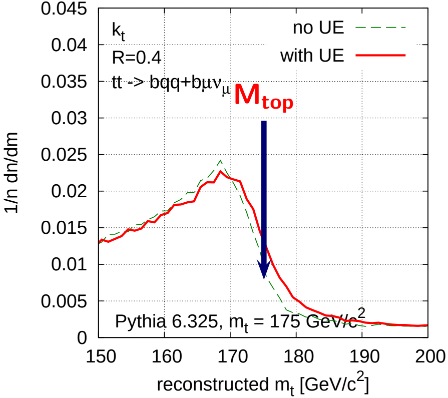
C/A identifies two hard blobs with limited soft contamination, joins them, and then adds in remaining soft junk

The interesting substructure is buried inside the clustering sequence — it's less contamined by soft junk, but needs to be pulled out with special techniques

Butterworth, Davison, Rubin & GPS '08
Kaplan, Schwartz, Reherman & Tweedie '08
Butterworth, Ellis, Rubin & GPS '09
Ellis, Vermilion & Walsh '09

A more complex example: top reconstruction

Robustness: M_{top} varies with R?



<u>Game: measure top mass to 1 GeV</u>

example for Tevatron

$m_t = 175$ GeV

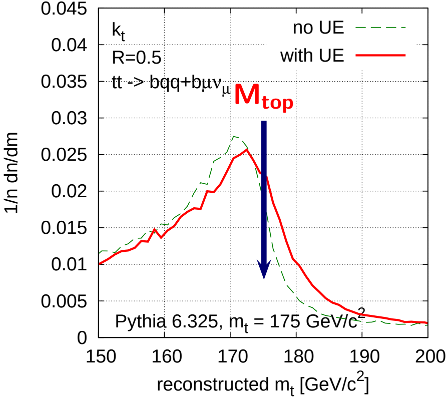
- Small *R*: lose 6 GeV to PT radiation and hadronisation, UE and pileup irrelevant
- Large *R*: hadronisation and PT radiation leave mass at ~ 175 GeV, UE adds $2-4$ GeV.

Is the final top mass (after W jet-energy-scale and Monte Carlo unfolding) independent of R used to measure jets?

Flexibility in jet finding gives powerful cross-check of systematic effects

cf. Seymour & Tevlin '06

Robustness: M_{top} varies with R?



Game: measure top mass to 1 GeV

example for Tevatron

$m_t = 175$ GeV

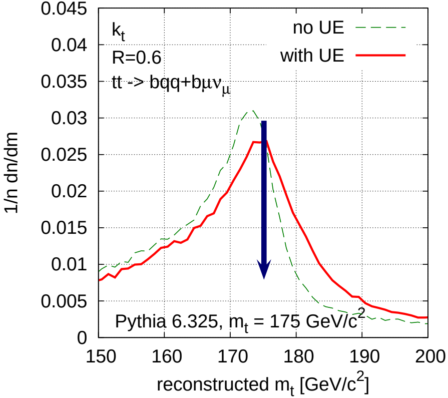
- Small R: lose 6 GeV to PT radiation and hadronisation, UE and pileup irrelevant
- Large R: hadronisation and PT radiation leave mass at ~ 175 GeV, UE adds $2-4$ GeV.

Is the final top mass (after W jet-energy-scale and Monte Carlo unfolding) independent of R used to measure jets?

Flexibility in jet finding gives powerful cross-check of systematic effects

cf. Seymour & Tevlin '06

Robustness: M_{top} varies with R?



Game: measure top mass to 1 GeV

example for Tevatron

$m_t = 175$ GeV

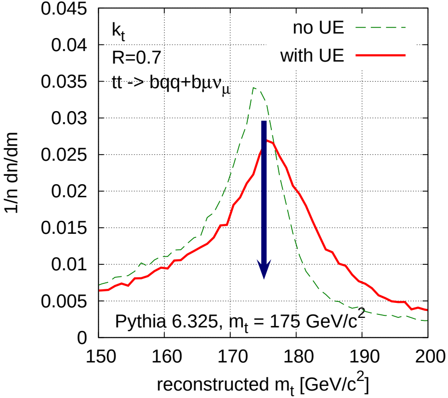
- Small R: lose 6 GeV to PT radiation and hadronisation, UE and pileup irrelevant
- Large R: hadronisation and PT radiation leave mass at ~ 175 GeV, UE adds $2 - 4$ GeV.

Is the final top mass (after W jet-energy-scale and Monte Carlo unfolding) independent of R used to measure jets?

Flexibility in jet finding gives powerful cross-check of systematic effects

cf. Seymour & Tevlin '06

Robustness: M_{top} varies with R?



Game: measure top mass to 1 GeV

example for Tevatron

$m_t = 175$ GeV

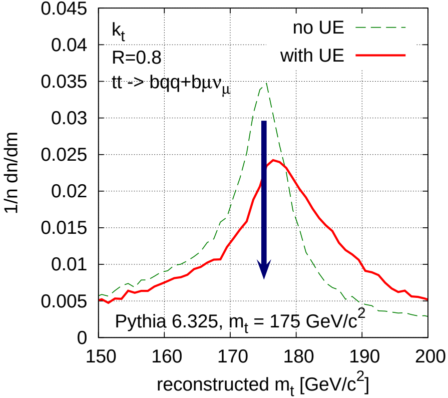
- Small R: lose 6 GeV to PT radiation and hadronisation, UE and pileup irrelevant
- Large R: hadronisation and PT radiation leave mass at ~ 175 GeV, UE adds $2-4$ GeV.

Is the final top mass (after W jet-energy-scale and Monte Carlo unfolding) independent of R used to measure jets?

Flexibility in jet finding gives powerful cross-check of systematic effects

cf. Seymour & Tevlin '06

Robustness: M_{top} varies with R?



Game: measure top mass to 1 GeV

example for Tevatron

$m_t = 175$ GeV

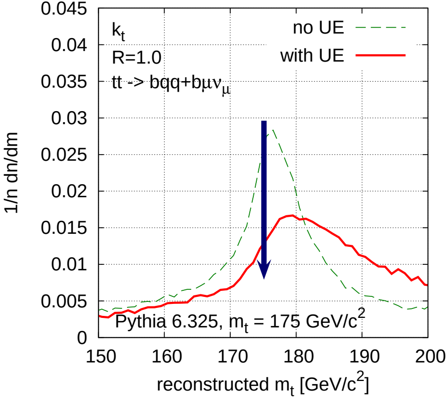
- Small R: lose 6 GeV to PT radiation and hadronisation, UE and pileup irrelevant
- Large R: hadronisation and PT radiation leave mass at ~ 175 GeV, UE adds $2-4$ GeV.

Is the final top mass (after W jet-energy-scale and Monte Carlo unfolding) independent of R used to measure jets?

Flexibility in jet finding gives powerful cross-check of systematic effects

cf. Seymour & Tevlin '06

Robustness: M_{top} varies with R?



Game: measure top mass to 1 GeV

example for Tevatron

$m_t = 175$ GeV

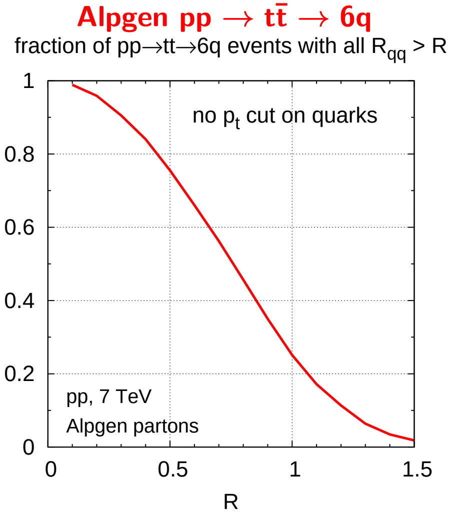
- Small R: lose 6 GeV to PT radiation and hadronisation, UE and pileup irrelevant
- Large R: hadronisation and PT radiation leave mass at ~ 175 GeV, UE adds $2 - 4$ GeV.

Is the final top mass (after W jet-energy-scale and Monte Carlo unfolding) independent of R used to measure jets?

Flexibility in jet finding gives powerful cross-check of systematic effects

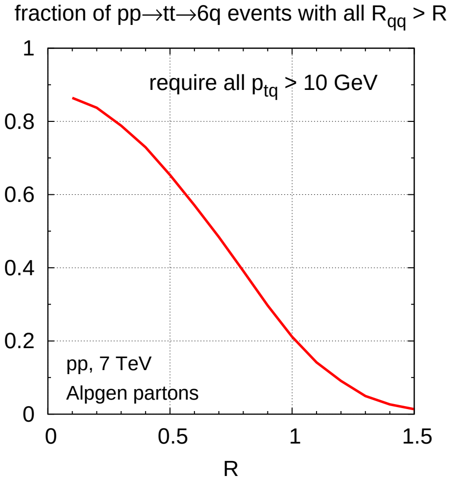
cf. Seymour & Tevlin '06

6 partons v. 6 jets?



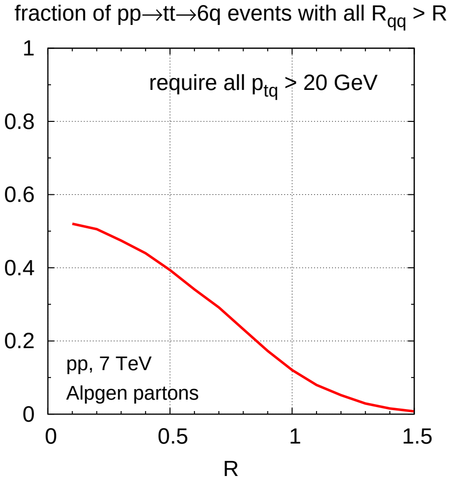
6 partons v. 6 jets?

Alpgen $pp \rightarrow t\bar{t} \rightarrow 6q$



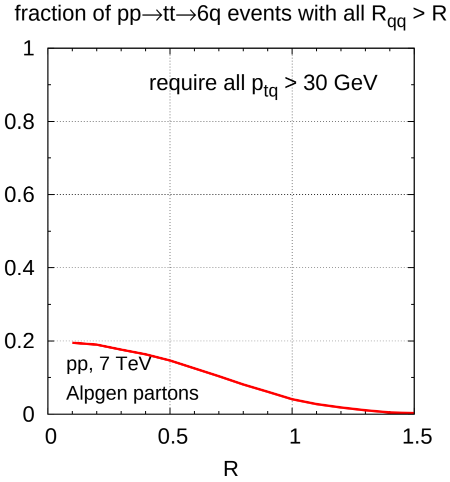
6 partons v. 6 jets?

Alpgen pp $\rightarrow t\bar{t} \rightarrow$ 6q



6 partons v. 6 jets?

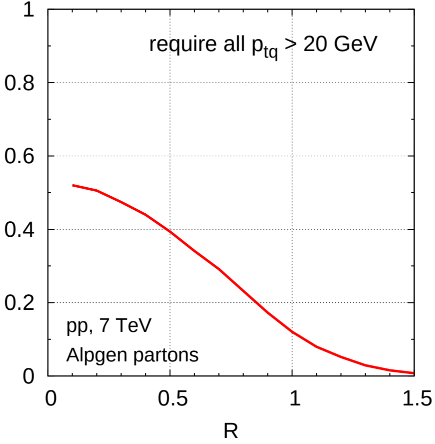
Alpgen $pp \to t\bar{t} \to 6q$



6 partons v. 6 jets?

Alpgen pp → $t\bar{t}$ → 6q

fraction of pp→tt→6q events with all R_{qq} > R



Herwig pp → $t\bar{t}$ → hadrons

Distribution of number of jets