Jets and jet substructure

Gavin Salam (CERN)

with extensive use of material by Matteo Cacciari and Gregory Soyez

> CFHEP April 2014







Gavin Salam (CERN)



And they've been used and studied at every collider since





And they've been used and studied at every collider since







Gavin Salam (CERN)

Jets and jet substructure (1)

Find all papers by ATLAS and CMS

				578 records found
eportnumber:CERN-PH-EP (collaboration:atlas or collaboration:CMS)		Brief format 🔹	Search Easy Search	
ind j "Phys.Rev.Lett.,105*" :: more			Advanced o	(bar on)
Sort by:	Display results:			
latest first 🔹 desc. 💌 – or rank by – 💌	25 results 💌 single list 💌			
lo exact match found for cern-ph-ep, using cern ph ep instead				
HEP 464 records found 1 - 25	jump to record: 1			Search took 0.04 seconds.

 Measurement of the W+W- cross section in pp collisions at sqrt(s) = 7 TeV and limits on anomalous WW gamma and WWZ couplings CMS Collaboration (Serguei Chatrchyan (Yerevan Phys. Inst.) et al.). Jun 5, 2013. CMS-SMP-12-005, CERN-PH-EP-2013-075 e-Print: <u>arXiv:1306.1126</u> [hep-ex] | <u>PDF</u>

References | BibTeX | LaTeX(US) | LaTeX(EU) | Harvmac | EndNote ADS Abstract Service

Detailed record

2. Measurement of the hadronic activity in events with a Z and two jets and extraction of the cross section for the electroweak production of a Z with two jets in pp collisions at \sqrt{s} = 7 TeV

CMS Collaboration (Serguei Chatrchyan (Yerevan Phys. Inst.) et al.). May 31, 2013. CMS-FSQ-12-019, CERN-PH-EP-2013-060 e-Print: arXiv:1305.7389 [hep-ex] | PDF

References | BibTeX | LaTeX(US) | LaTeX(EU) | Harvmac | EndNote ADS Abstract Service

Detailed record

Pull out those that refer to one widely used jet-alg



No exact match found for cern-ph-ep, using cern ph ep instead...

HEP	272 records found 1 - 25 ▶ jump to record: 1	Search took 0.04 seconds.
-----	--	---------------------------

 Measurement of the W+W- cross section in pp collisions at sqrt(s) = 7 TeV and limits on anomalous WW gamma and WWZ couplings CMS Collaboration (Serguei Chatrchyan (Yerevan Phys. Inst.) et al.). Jun 5, 2013. CMS-SMP-12-005, CERN-PH-EP-2013-075 e-Print: <u>arXiv:1306.1126</u> [hep-ex] | PDF

References | BibTeX | LaTeX(US) | LaTeX(EU) | Harvmac | EndNote ADS Abstract Service

Detailed record

2. Measurement of the hadronic activity in events with a Z and two jets and extraction of the cross section for the electroweak production of a Z with two jets in pp collisions at \sqrt{s} = 7 TeV

CMS Collaboration (Serguei Chatrchyan (Yerevan Phys. Inst.) et al.). May 31, 2013. CMS-FSQ-12-019, CERN-PH-EP-2013-060 e-Print: arXiv:1305.7389 [hep-ex] | PDF

References | BibTeX | LaTeX(US) | LaTeX(EU) | Harvmac | EndNote ADS Abstract Service

Detailed record

Jet usage at the LHC







e jets?



e jets?



Reconstructing jets

tion



Projection to jets should be resilient to QCD effects





2 clear jets



2 clear jets

3 jets?

Reconstructing jets is an ambiguous task





2 clear jets

3 jets? or 4 jets?

Make a choice: specify a jet definition



- Which particles do you put together into a same jet?
- How do you recombine their momenta (4-momentum sum is the obvious choice, right?)

"Jet [definitions] are legal contracts between theorists and experimentalists" -- MJ Tannenbaum

They're also a way of organising the information in an event 1000's of particles per events, up to 20.000,000 events per second

Jets and jet substructure (1)



Cluster analysis

From Wikipedia, the free encyclopedia

Cluster analysis or clustering is the task of grouping a set of objects in such a

way that objects in the same group (called **cluster**) are more similar (in some sense or another) to each other than to those in other groups (clusters). It is a main task of exploratory data mining, and a common technique for statistical data analysis used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, and bioinformatics.

Contents [hide]

- 1 Clusters and clusterings
- 2 Clustering algorithms
 - 2.1 Connectivity based clustering (hierarchical clustering)
 - 2.2 Centroid-based clustering
 - 2.3 Distribution-based clustering
 - 2.4 Density-based clustering
 - 2.5 Newer developments
- 3 Evaluation of clustering results
 - 3.1 Internal evaluation
 - 3.2 External evaluation
- **4** Applications
- 5 See also
 - 5.1 Related topics
 - 5.2 Related methods
- 6 References

There is no objectively "correct" clustering algorithm, but [...] "clustering is in the eye of the beholder."[1] The most appropriate clustering algorithm for a particular problem often needs to be chosen experimentally, unless there is a mathematical reason to prefer one cluster model over another.

Partitioning / centroid-based clustering [cone algorithms]

k-means

Example of a partitional algorithm

- I) Choose K centroids at random
- 2) Assign objects to closest centroid, forming K clusters
- 3) Calculate centroid (mean of distances) of each cluster, update centroids
- 4) Check if an object in a cluster is closer to another centroid. Reallocate in case.
- 5) Repeat from step 3 until no object changes cluster anymore.



Step I (random centroids)



Step 2 (allocate objects)



Step 3 (move centroids)



Step 5 (end of iteration)

k-means

Example of a partitional algorithm

- I) Choose K centroids at random
- 2) Assign objects to closest centroid, forming K clusters
- 3) Calculate centroid (mean of distances) of each cluster, update centroids
- Check if an object in a cluster is closer to another centroid. Reallocate in case.
- 5) Repeat from step 3 until no object changes cluster anymore.



Step I (random centroids)



Step 2 (allocate objects)

Step 3 (move centroids)

Step 5 (end of iteration)

One of the main shortcomings:

result of final convergence can be highly sensitive to choice of initial seeds. Also, the concept of 'mean distance' (to calculate the centroid) must be defined.

Matteo Cacciari - LPTHE

MadGraph School - May 2013 - Beijing

Garrett-Mayer

ш

Figures by



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

Notes





One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

Notes



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

<u>Notes</u>



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

<u>Notes</u>



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

<u>Notes</u>



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

<u>Notes</u>



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

Notes



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

Notes



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

Notes


One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event





































One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

Notes





One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

Notes



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

<u>Notes</u>



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

<u>Notes</u>



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

<u>Notes</u>



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

Notes



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

Notes



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

<u>Notes</u>



One of the simplest of the cone algs e.g. CMS iterative cone

- Take hardest particle as seed for cone axis
- Draw cone around seed
- Sum the momenta use as new seed direction, iterate until stable
- Convert contents into a "jet" and remove from event

Notes








































































Collinear splitting can modify the final hard jets The algorithm is **collinear unsafe**



Invalidates perturbation theory

The full cross section that you measure in experiment should correspond to an expression looking roughly as follows:

$$\sigma^{full} = \sigma_{LO} \left(1 + \alpha_{s}c_{1} + \alpha_{s}^{2}c_{2} + \alpha_{s}^{3}c_{3} + \ldots + \mathcal{O}\left(\frac{\Lambda_{QCD}}{p_{t}}\right) \right)$$

A perturbative series

plus a non-perturbative contribution, suppressed by a power of Λ_{QCD}/p_t

We don't have the technology to calculate the full series or the non-perturbative part. Typically, one might "just" calculate next-to-leading order

 $\sigma^{NLO} = \sigma_{LO} \left(1 + \alpha_{\rm s} c_1 \right)$

The point to perturbation theory is that the $c_2 \alpha_s^2$, etc. terms are small compared to the ones you have calculated — *hence (e.g.) NLO should be a good approximation*.

Real life does not have infinities, but pert. infinity leaves a real-life trace

$$\alpha_{\rm s}^2 + \alpha_{\rm s}^3 + \alpha_{\rm s}^4 \times \infty \to \alpha_{\rm s}^2 + \alpha_{\rm s}^3 + \alpha_{\rm s}^4 \times \ln p_t / \Lambda \to \alpha_{\rm s}^2 + \underbrace{\alpha_{\rm s}^3 + \alpha_{\rm s}^3}_{\text{BOTH WASTER}}$$

Among consequences of IR unsafety:

	Last meaningful order			
	JetClu, ATLAS	MidPoint	CMS it. cone	Known at
	CONE [IC-SM]	[IC _{mp} -SM]	[IC-PR]	
Inclusive jets	LO	NLO	NLO	$NLO (\rightarrow NNLO)$
W/Z + 1 jet	LO	NLO	NLO	NLO $(\rightarrow NNLO)$
3 jets	none	LO	LO	NLO [nlojet++]
W/Z + 2 jets	none	LO	LO	NLO [MCFM]
$m_{ m jet}$ in $2j + X$	none	none	none	NLO [Blackhat/Rocket/]

NB: 50,000,000/£/CHF/ \in investment in NLO

Multi-jet contexts much more sensitive: ubiquitous at LHC And LHC relies on QCD for background double-checks extraction of cross sections, extraction of parameters

Real life does not have infinities, but pert. infinity leaves a real-life trace

$$\alpha_{\rm s}^2 + \alpha_{\rm s}^3 + \alpha_{\rm s}^4 \times \infty \to \alpha_{\rm s}^2 + \alpha_{\rm s}^3 + \alpha_{\rm s}^4 \times \ln p_t / \Lambda \to \alpha_{\rm s}^2 + \underbrace{\alpha_{\rm s}^3 + \alpha_{\rm s}^3}_{\text{BOTH WASTER}}$$

Among consequences of IR unsafety:

	Last meaningful order			
	JetClu, ATLAS	MidPoint	CMS it. cone	Known at
	cone [IC-SM]	[IC _{mp} -SM]	[IC-PR]	
Inclusive jets	LO	NLO	NLO	NLO (\rightarrow NNLO)
W/Z+1 jet	LO	NLO	NLO	NLO (\rightarrow NNLO)
3 jets	none	LO	LO	NLO [nlojet++]
W/Z + 2 jets	none	LO	LO	NLO [MCFM]
$m_{ m jet}$ in $2j + X$	none	none	none	NLO [Blackhat/Rocket/]

NB: 50,000,000 $/\pounds/CHF/\in$ investment in NLO

Multi-jet contexts much more sensitive: **ubiquitous at LHC** And LHC relies on QCD for background double-checks extraction of cross sections, extraction of paramete₈₀

Real life does not have infinities, but pert. infinity leaves a real-life trace

$$\alpha_{s}^{2} + \alpha_{s}^{3} + \alpha_{s}^{4} \times \infty \rightarrow \alpha_{s}^{2} + \alpha_{s}^{3} + \alpha_{s}^{4} \times \ln p_{t}/\Lambda \rightarrow \alpha_{s}^{2} + \underbrace{\alpha_{s}^{3} + \alpha_{s}^{3}}_{\text{BOTH WASTE}}$$

Among consequences of IR unsafety:

	Last meaningful order			
	JetClu, ATLAS	MidPoint	CMS it. cone	Known at
	cone [IC-SM]	[IC _{mp} -SM]	[IC-PR]	
Inclusive jets	LO	NLO	NLO	NLO (\rightarrow NNLO)
W/Z+1 jet	LO	NLO	NLO	NLO (\rightarrow NNLO)
3 jets	none	LO	LO	NLO [nlojet++]
W/Z + 2 jets	none	LO	LO	NLO [MCFM]
$m_{ m jet}$ in $2j + X$	none	none	none	NLO [Blackhat/Rocket/]

NB: 50,000,000 $/\pounds/CHF/\in$ investment in NLO

Multi-jet contexts much more sensitive: **ubiquitous at LHC** And LHC relies on QCD for background double-checks extraction of cross sections, extraction of parameters₈₁

Real life does not have infinities, but pert. infinity leaves a real-life trace

$$\alpha_{s}^{2} + \alpha_{s}^{3} + \alpha_{s}^{4} \times \infty \to \alpha_{s}^{2} + \alpha_{s}^{3} + \alpha_{s}^{4} \times \ln p_{t}/\Lambda \to \alpha_{s}^{2} + \underbrace{\alpha_{s}^{3} + \alpha_{s}^{3}}_{\text{BOTH WASTED}}$$

Among cons	sequ	ences of IR unsa	fety:			
	-	There is a way problem	y of resol for cone	ving this kind algorithms	d <mark>Of</mark> Known	at
Inclusive jet	S					\rightarrow NNLO)
W/Z + 1 j			SISCon	NLO		$\rightarrow NNLO)$
3 jets $W/Z + 2$ j		[Seedless	Infrared	Safe Cone]		nlojet++] MCFM]
$m_{\rm jet}$ in $2j$ +	$\cdot X$					Blackhat/Rocket/]
But, in the end, this is not the solution that the LHC experiments chose, so let's Spend our time on the one they did						ent in NLO
		And LF		QCD for backgro	ound do	uble-checks
		CALIA	CHOIL OF CLOS	5 Sections, extra		parameters ₈₁

Hierarchical clustering [sequential recombination algorithms]
Sequential recombination algorithms (e+e-)

[aka hierarchical agglomerative clustering]

- Define a distance measure d_{ij} between all pairs of particles.
- Recombine pair with smallest d_{ij}
- Repeat until all $d_{ij} > d_{cut}$

The algorithm's general behaviour is governed by how one defines the interpair distances d_{ij}

d_{cut} is a **resolution parameter**, which governs how whether jets are coarse or fine objects

JADE algorithm – the original seq. rec. algorithm

$$d_{ij} = m_{ij}^2 \simeq 2E_i E_j (1 - \cos \theta_{ij})$$

The most obvious choice? But does not give sensible behaviour

[see blackboard]

Use QCD divergences to help decide distance measure:

The stronger the divergence between a pair of particles, the more likely it is they should be associated with each other. *[see blackboard]*

k_t / Durham algorithm – the most widely used at LEP

$$d_{ij} = 2\min(E_i^2, E_j^2)(1 - \cos\theta_{ij})$$

Catani, Dokshitzer, Olsson, Turnock & Webber '91

<t/> kt/Durham in action



What changes at a hadron collider?

• You have (unseen) beams, introduce a "beam distance"

$$d_{iB} = 2E_i^2(1 - \cos\theta_{iB})$$

squared trans. mom. wrt beam

• You want to use longitudinally invariant variables, i.e. p_t , rapidity (*y*) and azimuth (ϕ)

 $d_{ij} = \min(p_{ti}^2, p_{tj}^2) \Delta R_{ij}^2, \quad \Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ $d_{iB} = p_{ti}^2$

exclusive kt algorithm

still just one parameter *d_{cut}* Catani, Dokshitzer, Seymour & Webber 1993

kt alg. at hadron colliders (2)

Two parameters, *R* and *p*_{*t*,*min*}

(These are the two parameters in essentially every widely used hadron-collider jet algorithm)

$$d_{ij} = \min(p_{ti}^2, p_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad \Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

Reformulate algorithm

- 1. Find smallest of d_{ij} , d_{iB}
- 2. If *ij*, recombine them
- 3. If *iB*, call i a jet and remove from list of particles
- 4. repeat from step 1 until no particles left
 - Only use jets with $p_t > p_{t,min}$

Inclusive kt algorithm S.D. Ellis & Soper, 1993

In what way do the inclusive and exclusive variants' behaviours differ? [see blackwhiteboard]

Is one "right", the other wrong"? A priori no, maybe we'll come back to this later.

$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0





kt alg.: Find smallest of

 $d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers



kt alg.: Find smallest of

 $d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers





$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers



kt alg.: Find smallest of

$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers





$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- ▶ if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers



kt alg.: Find smallest of

$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers



kt alg.: Find smallest of

$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers



kt alg.: Find smallest of

 $d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers





$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers



kt alg.: Find smallest of

$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0



kt alg.: Find smallest of

$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- ► if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0



kt alg.: Find smallest of

$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0



kt alg.: Find smallest of

$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0



kt alg.: Find smallest of

 $d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0





$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0



kt alg.: Find smallest of

 $d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0





$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0



kt alg.: Find smallest of

 $d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers





$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- ► if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers



kt alg.: Find smallest of

$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- ▶ if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers





$$d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$$

- ► If *d_{ij}* recombine
- ► if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0 ϕ assumed 0 for all towers



kt alg.: Find smallest of

 $d_{ij} = \min(k_{ti}^2, k_{tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^2$

- ► If *d_{ij}* recombine
- if d_{iB} , *i* is a jet

Example clustering with k_t algorithm, R = 1.0

<u>INT V. CONPES</u>



(Some) cone algorithms give circular jets in $y - \phi$ plane Much appreciated by experiments e.g. for acceptance corrections

k_t jets are **irregular**

Because soft junk clusters together first:

 $d_{ij} = \min(k_{ti}^2, k_{tj}^2) \Delta R_{ij}^2$ Regularly held against k_t



The kt algorithms form one of several "families" of sequential recombination jet algorithm

Others differ in

1. the choice of distance measure between pairs of particles [i.e. relative priority given to soft and collinear divergences]

2. using $3 \rightarrow 2$ clustering rather than $2 \rightarrow 1$ [ARCLUS; not used at hadron colliders, so won't discuss it more]

Cambridge/Aachen: the simplest of hadron-collider algorithms

- Recombine pair of objects closest in ΔR_{ij}
- Repeat until all $\Delta R_{ij} > R$ remaining objects are jets

Dokshitzer, Leder, Moretti, Webber '97 (Cambridge): more involved e+e- form Wobisch & Wengler '99 (Aachen): simple inclusive hadron-collider form One still applies a p_{t,min} cut to the jets, as for inclusive k_t

> C/A privileges the collinear divergence of QCD; it 'ignores' the soft one

anti-kt

Anti-kt: formulated similarly to inclusive kt, but with

$$d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$$

Cacciari, GPS & Soyez '08 [+Delsart unpublished]

Anti-kt privileges the collinear divergence of QCD and disfavours clustering between pairs of soft particles

Most pairwise clusterings involve at least one hard particle

Anti-kt in action

Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Anti-kt in action

Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$


Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Clustering grows around hard cores $d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$



Clustering grows around hard cores

$$d_{ij} = \frac{1}{\max(p_{ti}^2, p_{tj}^2)} \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = \frac{1}{p_{ti}^2}$$



Anti-kt gives cone-like jets without using stable cones Is it really only about the "circularity" of the jets' boundaries?















p_t/GeV kt clustering, R=1 3 y

p_t/GeV kt clustering, R=1 Т 3 y





















Anti-kt experimental performance

Efficiency for finding detector jet that matches particle jet



As good as, or better than all previous experimentallyfavoured algorithms.

Essentially because anti-kt has linear response to soft particles.

And it's also infrared and collinear safe (as needed for theory calcs).

[see blackboard for explanation of plot]

Anti-kt experimental performance

Efficiency for finding detector jet that matches particle jet



Infrared & Collinear safe algorithms

kt	sequential recombination $d_{ij} = min(k_{ti}^2, k_{tj}^2)\Delta R_{ij}^2/R^2$ hierarchical in relative p_t	Catani et al '91 Ellis, Soper '93	NInN
Cambridge/ Aachen	sequential recombination $d_{ij} = \Delta R_{ij}^2 / R^2$ hierarchical in angle	Dokshitzer et al '97 Wengler, Wobish '98	NInN
anti-k _t	sequential recombination $d_{ij} = min(k_{ti}^2, k_{tj}^2)\Delta R_{ij}^2/R^2$ gives perfectly conical hard jets	MC, Salam, Soyez '08 (Delsart)	N ^{3/2}
SISCone	Seedless iterative cone with split-merge gives 'economical' jets	Salam, Soyez '07	N ² InN

Infrared & Collinear safe algorithms



Infrared & Collinear safe algorithms

kt		sequential recombination $d_{ij} = min(k_{ti}^2, k_{tj}^2)\Delta R_{ij}^2/R^2$ hierarchical in relative p_t	Catani et al '91 Ellis, Soper '93	NInN
Cambrid	ge/	sequential recombination $d_{ii} = \Lambda R_{ii}^2/R^2$	Dokshitzer et al '97	NInN
Aache		Advertisment		
anti-k		All are available in FastJet, <u>http://fastjet.fr</u> (as well as many IRC unsafe ones)		
		Other (recent) software		
SISCo	S	Slow lot (in Pythics)		
	Slo	SlowJet (in Pythia8) SlowJet (mathematica code by J. Ruderman)		

http://fastjet.fr/

// specify a jet definition JetDefinition jet_def(antikt_algorithm, double R);

jet_algorithm can be any one of the four IRC safe algorithms, or also most of the old IRC-unsafe ones, for legacy purposes
http://fastjet.fr/

// specify a jet definition JetDefinition jet_def(antikt_algorithm, double R);

jet_algorithm can be any one of the four IRC safe algorithms, or also most of the old IRC-unsafe ones, for legacy purposes

// specify the input particles
vector<PseudoJet> input_particles = . . .;

http://fastjet.fr/

// specify a jet definition JetDefinition jet_def(antikt_algorithm, double R);

jet_algorithm can be any one of the four IRC safe algorithms, or also most of the old IRC-unsafe ones, for legacy purposes

// specify the input particles
vector<PseudoJet> input_particles = . . .;

// create a ClusterSequence
ClusterSequence cs(input_particles, jet_def);

http://fastjet.fr/

// specify a jet definition JetDefinition jet_def(antikt_algorithm, double R);

jet_algorithm can be any one of the four IRC safe algorithms, or also most of the old IRC-unsafe ones, for legacy purposes

```
// specify the input particles
vector<PseudoJet> input_particles = . . .;
```

```
// create a ClusterSequence
ClusterSequence cs(input_particles, jet_def);
```

```
// extract the jets
vector<PseudoJet> jets = sorted_by_pt(cs.inclusive(jets));
```

```
// pt of hardest jet
double pt_hardest = jets[0].pt();
```

```
// constituents of hardest jet
vector<PseudoJet> constituents = jets[0].constituents();
```

FastJet speed

Time needed to cluster an event with N particles



Gavin Salam (CERN)

CFHEP, April 2014 138