Defining jets at the dawn of the LHC

Grégory Soyez

CERN

In collaboration with Gavin Salam, Matteo Cacciari and Juan Rojo

EPFL, Lausanne — January 11 2010
Plan

Jet algorithms and jet definitions
- basic ideas: why jets? recombinations and cones
- failures of the 20th-century cone algorithms
- new algorithms without the failures

More advanced topics: how to better use the tools we have?
- jet areas: tool for pileup subtraction
- new generation of algorithms
- optimal choice (for kinematic reconstructions)
QCD probability for gluon emission (angle $\theta$ and $\perp$-mom. $k_t$):

$$dP \propto \alpha_s \frac{d\theta}{\theta} \frac{dk_t}{k_t}$$

Two divergences:

- collinear
- soft

Divergences cancelled by virtual corrections
Motivation: why jets

Collinear divergence ⇒ QCD produces “jetty” showers

Example: LEP (OPAL) events

“Jets” ≡ bunch of collimated particles ⇐ hard partons
Motivation: why jets

Collinear divergence $\Rightarrow$ QCD produces “jetty” showers

“Jets” $\equiv$ bunch of collimated particles $\Leftrightarrow$ hard partons

BUT

- a “parton” is an ambiguous concept (NLO)
- “collinear” has some arbitrariness

\[ \begin{aligned}
\text{2 jets} & \quad \text{3 jets} & \quad ? \text{ jets}
\end{aligned} \]
Motivation: why jets

Collinear divergence ⇒ QCD produces “jetty” showers

“Jets” ≡ bunch of collimated particles \( \cong \) hard partons

In practice: use of a jet definition

Particles \( \{ p_i \} \) \rightarrow \text{jet definition} \rightarrow \text{jets} \( \{ j_k \} \)

Jet algorithm: the recipe (insufficient!)
Jet definition: algorithm + the parameters
**Recombination:**
- $k_t$ algorithm
- Cambridge/Aachen alg.

**Cone:**
- CDF JetClu
- CDF MidPoint
- D0 (run II) Cone
- PxCone
- ATLAS Cone
- CMS Iterative Cone
- PyCell/CellJet
- GetJet
Recombination:
- $k_t$ algorithm
- Cambridge/Aachen alg.

**Idea:** undo the showering

Successively
- find the closest pair of particles
- recombine them

**Distance:**

$k_t$:
$$d_{i,j} = \min (k_{t,i}^2, k_{t,j}^2) (\Delta \phi_{i,j}^2 + \Delta y_{i,j}^2)$$

Cam/Aachen:
$$d_{i,j} = \Delta \phi_{i,j}^2 + \Delta y_{i,j}^2$$

stop at a distance $R$
20th century jet algorithms

**Idea:** dominant flow of energy

**Stable cone (radius $R$):**
sum of particles in the cone points towards the cone centre

All these are **iterative cones**:
- start from a **seed**
- iterate until stable

seeds = \{particles, midpoints\}

Jet $\equiv$ stable cone
modulo overlapping

**Cone:**
- CDF JetClu
- CDF MidPoint
- D0 (run II) Cone
- PxCone
- ATLAS Cone
- CMS Iterative Cone
- PyCell/CellJet
- GetJet
**20th century jet algorithms**

Cone with **split-merge**

Split/merge if the overlap is smaller/larger than a **threshold** $f$

**Cone:**
- CDF JetClu
- CDF MidPoint
- D0 (run II) Cone
- PxCone
- ATLAS Cone
- CMS Iterative Cone
- PyCell/CellJet
- GetJet
Cone with **progressive removal**

Successively
- iterate from hardest particle
- call that a jet (remove particles)

**Basic property:**
- hard circular jets

**Cone:**
- CDF JetClu
- CDF MidPoint
- D0 (run II) Cone
- PxCone
- ATLAS Cone
- CMS Iterative Cone
- PyCell/CellJet
- GetJet
20th century jet algorithms

Recombination:
- $k_t$ algorithm
- Cambridge/Aachen alg.

✓ perturbative behaviour

Cone:
- CDF JetClu
- CDF MidPoint
- D0 (run II) Cone
- PxCone
- ATLAS Cone
- CMS Iterative Cone
- PyCell/CellJet
- GetJet

✓ UE sensitivity
21st century: how does that picture change?
Ingredient: QCD soft and collinear divergencies

\[ \infty \quad \infty \quad \infty \]

\( \infty \) (from soft gluons) cancel (inclusive x-section)
QCD divergences

Ingredient: QCD soft and collinear divergencies

- Consider an extra (NLO) soft gluon
- Assume LO gives 2 jets $\Rightarrow$ NLO(virt) gives 2 jets
**QCD divergences**

Ingredient: QCD soft and collinear divergencies

Consider an extra (NLO) soft gluon

- Assume LO gives 2 jets $\Rightarrow$ NLO(virt) gives 2 jets
- NLO(real) gives 2 jets $\Rightarrow \infty$ cancel $\Rightarrow$ finite jet cross-section
QCD divergences

Ingredient: QCD soft and collinear divergencies

Consider an extra (NLO) soft gluon

Assume LO gives 2 jets $\Rightarrow$ NLO(virt) gives 2 jets

NLO(real) gives 2 jets $\Rightarrow \infty$ cancel $\Rightarrow$ finite jet cross-section

NLO(real) gives 1 jets $\Rightarrow \infty$ do not cancel $\Rightarrow$ infinite jet x-section
**QCD divergences**

Ingredient: QCD soft and collinear divergencies

For pQCD to make sense, the (hard) jets should not change when

- one has a soft emission *i.e.* adds a very soft gluon
- one has a collinear splitting
  *i.e.* replaces one parton by two at the same place \((\eta, \phi)\)

[SNOWMASS Accords, Fermilab, 1990]
IR (un)safety? JetClu and Atlas Cone
IR (un)safety? JetClu and Atlas Cone

Stable cones found
A soft gluon changed the number of jets

⇒ IR unsafety of JetClu and the ATLAS Cone
A soft gluon changed the number of jets
⇒ IR unsafety of JetClu and the ATLAS Cone

Fixed by MidPoint

[Blazey et al., 00]
IR (un)safety? MidPoint
Stable cones found
A soft gluon changed the number of jets

⇒ **IR unsafety of MidPoint** (1 order in $\alpha_s$ later than JetClu)
**Solution**: be sure to find all stable cones

*SISCones*: Seedless Infrared-Safe Cone algorithm

[http://projects.hepforge.org/siscone](http://projects.hepforge.org/siscone)

[G.Salam, G.S., 07]

Idea: enumerate enclosures by enumerating pairs of particles
Collinear (un)safety? the CMS iterative cone
A colinear splitting changed the number of jets

⇒ Collinear unsafety of the CMS iterative cone
Come back to recombination-type algorithms:

\[ d_{ij} = \min(k_{t,i}^{2p}, k_{t,j}^{2p}) \left( \Delta \phi_{ij}^2 + \Delta \eta_{ij}^2 \right) \]

- \( p = 1 \): \( k_t \) algorithm
- \( p = 0 \): Aachen/Cambridge algorithm
Come back to recombination-type algorithms:

\[ d_{ij} = \min(k_{t,i}^{2p}, k_{t,j}^{2p}) \left( \Delta \phi_{ij}^2 + \Delta \eta_{ij}^2 \right) \]

- \( p = 1 \): \( k_t \) algorithm
- \( p = 0 \): Aachen/Cambridge algorithm
- \( p = -1 \): anti-\( k_t \) algorithm \[M.Cacciari, G.Salam, G.S., 08\]

Why should that be related to the iterative cone ?!? 

- “large \( k_t \) \( \Rightarrow \) small distance” 
  \( i.e. \) hard partons “eat” everything up to a distance \( R \)
  \( i.e. \) circular/regular jets, jet borders unmodified by soft radiation
- infrared and collinear safe
21st century jet finders

Recombination:
- $k_t$ algorithm
- Cambridge/Aachen algorithm
- anti-$k_t$ algorithm

Cone:
- CDF JetClu
- CDF MidPoint
- D0 (run II) Cone
- PxCone
- ATLAS Cone
- CMS Iterative Cone
- PyCell/CellJet
- GetJet
- SISConce
Recombination:
- $k_t$ algorithm
- Cambridge/Aachen alg.
- anti-$k_t$ algorithm

Cone:
- CDF JetClu
- CDF MidPoint
- D0 (run II) Cone
- PxCone
- ATLAS Cone
- CMS Iterative Cone
- PyCell/CellJet
- GetJet
- SISCones

4 available safe algorithms
21st century jet finders

Recombination:
- $k_t$ algorithm
- Cambridge/Aachen alg.
- anti-$k_t$ algorithm

Cone:
- CDF JetClu
- CDF MidPoint
- D0 (run II) Cone
- PxCone
- A TLAS Cone
- CMS Iterative Cone
- SISCone

All those algorithms (and much more) implemented (efficiently) in FastJet
### 21st century jet finders

<table>
<thead>
<tr>
<th><strong>Recombination:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_t$ algorithm</td>
</tr>
<tr>
<td>Cambridge/Aachen alg.</td>
</tr>
<tr>
<td>anti-$k_t$ algorithm</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Cone:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>CDF JetClu</td>
</tr>
<tr>
<td>CDF MidPoint</td>
</tr>
<tr>
<td>D0 (run II) Cone</td>
</tr>
<tr>
<td>PxCone</td>
</tr>
<tr>
<td>GetJet</td>
</tr>
<tr>
<td>SISCones</td>
</tr>
</tbody>
</table>

---

anti-$k_t$ adopted as the default jet algorithm by both CMS and ATLAS

---

safe algorithms
When does IRC safety matters?

Take e.g. the MidPoint cone

\[ \alpha_s^2 \times \ldots + \alpha_s^3 \times \ldots + \alpha_s^4 \times \ldots + \alpha_s^5 \times \ldots + \ldots \]

- QCD expansion (one \( \alpha_s \) can be replaced by \( \alpha_{EW} \))
When does IRC safety matters?

Take e.g. the MidPoint cone

\[ \alpha_s^2 \times \ldots + \alpha_s^3 \times \ldots + \alpha_s^4 \times \ldots + \alpha_s^5 \times \log\left(\frac{p_t}{\Lambda_{QCD}}\right) \ldots + \ldots \]

- QCD expansion (one \(\alpha_s\) can be replaced by \(\alpha_{EW}\))
- IRC unsafety (regulated at the hadronic scale \(\sim \Lambda_{QCD}\))
When does IRC safety matters?

Take *e.g.* the MidPoint cone

\[
\begin{align*}
\text{2 particles} & \quad \alpha_s^2 \times \ldots + \alpha_s^3 \times \ldots + \\
\text{3 particles} & \quad \alpha_s^4 \times \ldots + \\
\text{4 particles} & \quad \alpha_s^5 \times \log\left(\frac{p_t}{\Lambda_{QCD}}\right) \ldots + \\
\text{4 particles + 1 soft} & \quad \text{cannot be trusted}
\end{align*}
\]

- QCD expansion (one \(\alpha_s\) can be replaced by \(\alpha_{EW}\))
- IRC unsafety (regulated at the hadronic scale \(\sim \Lambda_{QCD}\))
- \(\alpha_s \log\left(\frac{p_t}{\Lambda_{QCD}}\right) \sim 1\)
- **last meaningful order** = \(\alpha_s^3\) or \(\alpha_{EW} \alpha_s^2\)
When does IRC safety matters?

Take e.g. the MidPoint cone

\[ \alpha_s^2 \times \ldots + \alpha_s^3 \times \ldots + \alpha_s^4 \times \ldots + \alpha_s^5 \times \log \left( \frac{p_t}{\Lambda_{QCD}} \right) \ldots + \ldots \]

- QCD expansion (one \( \alpha_s \) can be replaced by \( \alpha_{EW} \))
- IRC unsafety (regulated at the hadronic scale \( \sim \Lambda_{QCD} \))
- \( \alpha_s \log \left( \frac{p_t}{\Lambda_{QCD}} \right) \sim 1 \)
- last meaningful order = \( \alpha_s^3 \) or \( \alpha_{EW} \alpha_s^2 \)
- same argument for the Iterative Cone
- 1 order worse for JetClu or the ATLAS cone
Physical impact

MidPoint/CMS iterative cone unsafe at $\mathcal{O}(\alpha_s^4)$ (or $\mathcal{O}(\alpha_{ew}\alpha_s^3)$)

<table>
<thead>
<tr>
<th>Physical observable</th>
<th>JetClu/ATLAS c.</th>
<th>MidPoint/CMS it. c.</th>
<th>SISCone/recomb.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inclusive jet cross section</td>
<td>LO</td>
<td>NLO</td>
<td>any</td>
</tr>
<tr>
<td>3-jet cross section</td>
<td>none</td>
<td>LO</td>
<td>any</td>
</tr>
<tr>
<td>$W/Z/H + 2$ jet cross sect.</td>
<td>none</td>
<td>LO</td>
<td>any</td>
</tr>
<tr>
<td>jet masses in 3 jets</td>
<td>none</td>
<td>none</td>
<td>any</td>
</tr>
</tbody>
</table>

Example: (Midpoint-SISCone)/SISCone

- Incl. cross-section: a few %
- Masses in 3-jet events: $\sim 45\%$
Physical impact

MidPoint/CMS iterative cone unsafe at $\mathcal{O}(\alpha_s^4)$ (or $\mathcal{O}(\alpha_{ew}\alpha_s^3)$)

<table>
<thead>
<tr>
<th>Physical observable</th>
<th>IRC-safe until</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>JetClu/ATLAS c.</td>
</tr>
<tr>
<td>Inclusive jet cross section</td>
<td>LO</td>
</tr>
<tr>
<td>3-jet cross section</td>
<td>none</td>
</tr>
<tr>
<td>$W/Z/H + 2$ jet cross sect.</td>
<td>none</td>
</tr>
<tr>
<td>jet masses in 3 jets</td>
<td>none</td>
</tr>
</tbody>
</table>

Huge effort ($\sim$ 50 M€) to compute processes in pQCD

Note:
- arXiv:0903.0814: $W + 2$ jets vs. LO QCD using CDF JetClu
- arXiv:0903.1748: $Z + 2$ jets vs. NLO QCD using the D0runII cone
- arXiv:0903.1801: $Z + 2$ jets vs. NLO QCD using the CMS iterative cone
## Summary of IRC safe algorithms

<table>
<thead>
<tr>
<th></th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_t$</td>
<td>matches QCD branchings</td>
<td>large UE sensitivity</td>
</tr>
<tr>
<td>Cam/Aa</td>
<td>good QCD behaviour</td>
<td>poor UE sensitivity</td>
</tr>
<tr>
<td></td>
<td>easily look at $\neq$ scales</td>
<td></td>
</tr>
<tr>
<td>anti-$k_t$</td>
<td>easy algorithm</td>
<td>poor UE sensitivity</td>
</tr>
<tr>
<td></td>
<td>easy calibration</td>
<td></td>
</tr>
<tr>
<td>SISCones</td>
<td>small UE sensitivity</td>
<td>poor QCD behaviour</td>
</tr>
</tbody>
</table>
We (finally) have a good set of tools

Can we do better?
A growing list

Many ideas and applications:

✓ jet areas and background subtraction
   \[ \rightarrow \text{UE, pileup, heavy-ion background subtraction} \]

✓ jet substructure and filtering
   \[ \rightarrow \text{see below} \]

✓ “best” jet definition
   \[ \rightarrow \text{kinematic dijet reconstruction} \]

✓ boosted objects tagging
   \[ \rightarrow H \rightarrow b\bar{b}, t, \tilde{\chi}_0^1 \rightarrow qqq, \ldots \]

I will cover the first three (see e.g. Gavin Salam’s talk here for the 4th)
New idea #1: filtering
Filtering

cluster with Cambridge/Aachen (R)

$\mathbf{p}_t$ [GeV]
Filtering

![Diagram](image)

- cluster with Cambridge/Aachen(R)
- for each jet
cluster with Cambridge/Aachen(R)

for each jet

recluster with Cambridge/Aachen(R/2)
Filtering

- Filtered

cluster with
Cambridge/Aachen(R)

for each jet

recluster with
Cambridge/Aachen(R/2)

keep the 2 hardest subjets
Filtering

- cluster with Cambridge/Aachen(R)
  - for each jet
    - recluster with Cambridge/Aachen(R/2)
    - keep the 2 hardest subjets

Idea:
- keep perturbational radiation
- remove UE

Proven useful for boosted jet $H \rightarrow b\bar{b}$ tagging

[J.Butterworth, A.Davison, M.Rubin, G.Salam, 08]

Proven useful for kinematic reconstructions

[M.Cacciari, J.Rojo, G.Salam, GS, 08]
New idea #2: jet definition optimisation
Optimisation: underlying idea

Competition between

- catching perturbative radiation

\[
\langle \delta p_t \rangle \propto - \int_R \frac{d\theta}{\theta} \sim - \log(1/R)
\]

- not catching soft background radiation (underlying event)

\[
\langle \delta p_t \rangle \sim \text{Soft contents} \propto \text{jet area} \sim R^2
\]

the coefficients depend on the algorithm
Optimisation: underlying idea

Competition between

- catching perturbative radiation
- not catching soft background radiation (underlying event)

Out-of-cone radiation:

\[ \langle \delta p_t \rangle \propto - \int_{R}^{} \frac{d\theta}{\theta} \sim - \log(1/R) \]

What is the optimal jet definition (algo+R!)?

\[ \langle \delta p_t \rangle \sim \text{Soft contents} \propto \text{jet area} \sim R^2 \]

the coefficients depend on the algorithm
Example process to illustrate various effects:

$$Z' \rightarrow q\bar{q} \rightarrow 2 \text{ jets}$$

- $M_{Z'}$ can be varied (between 100 GeV and 4 TeV)
- Also valid for $H \rightarrow gg$ to study gluon jets
- Reconstruction method:
  - get the 2 hardest jets: $j_1$ and $j_2$
  - reconstruct the $Z'$: $m_{Z'} = (j_1 + j_2)^2$

  Look how the mass peak is reconstructed
- Also $t\bar{t}$ with full hadronic decay for multijet tests
Measure of the jet reconstruction efficiency:

- Forget about measures related to parton-jet matching
- Forget about fits depending on the shape of the peak

⇒ maximise the signal over background ratio \( \frac{S}{\sqrt{B}} \)

A narrower peak is better.
Measure of the jet reconstruction efficiency:

- Forget about measures related to parton-jet matching.
- Forget about fits depending on the shape of the peak.
⇒ maximise the signal over background ratio ($S/\sqrt{B}$) for a narrower peak is better.

<table>
<thead>
<tr>
<th>dijet mass [GeV]</th>
<th>$1/N , dN/dbin$</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>0.01</td>
</tr>
<tr>
<td>100</td>
<td>0.02</td>
</tr>
<tr>
<td>120</td>
<td>0.03</td>
</tr>
</tbody>
</table>

SISCone $Q^w_{f=0.12}$ (GeV)

- $k_t, R=Q^w_{f=0.1}$
- $R_{best}$

qq 100 GeV

- $f=0.75$

- p. 23
Assuming a constant background,

quality measure $\rightarrow$ effective luminosity ratio

$$\rho_{\mathcal{L}}(\text{JD}_2/\text{JD}_1) = \frac{\mathcal{L} \text{ needed with } \text{JD}_2}{\mathcal{L} \text{ needed with } \text{JD}_1} = \frac{Q_{f=z}^{w}(\text{JD}_2)}{Q_{f=z}^{w}(\text{JD}_1)}$$

e.g. $\rho_{\mathcal{L}}(\text{JD}_2/\text{JD}_1) = 2$

$\Leftrightarrow$ JD$_2$ requires 2 times the integrated luminosity of JD$_1$

to achieve the same discriminative power.

Note: results cross-checked with 2 different definitions of the quality measure
SISCon and C/A+filt. do slightly better than $k_t$, C/A or anti-$k_t$
**Optimisation: best definition**

- SIS Cone and C/A+filt. do slightly better than $k_t$, C/A or anti-$k_t$

- $M \uparrow \Rightarrow R_{\text{best}} \uparrow$ (and $R_{\text{best}}(g) > R_{\text{best}}(q)$)

![Graph showing the comparison between different methods](image-url)
Using a single jet definition for all processes may cost a factor $\sim 2$ in time for early discoveries at the LHC.
Using a single jet definition for all processes may cost a factor ~ 2 in time for early discoveries at the LHC

[Graphs and diagrams depicting jet definition optimizations for various scenarios, with links to more information on the given website.]
New idea #3: jet area and soft background subtraction
Jet areas

Area ≡ region where the jet catches soft particles

**Recipe:** add infinitely soft particles (aka *ghosts*) and see in which jet they are clustered

**2 methods:**
- Passive area: add one ghost at a time and repeat many times
- Active area: add a set of ghosts and cluster once

**Idea:** ghost ≈ background particle

⇒ active area ≈ uniform background
passive area ≈ pointlike background

**Notes:**
- passive = active for large multiplicities
- require an IR-safe algorithm!
- generic/universal definition (e.g. independent of a calorimeter)
Jet area: examples

**Example**: active area for a simple event

\[ k_t \quad \text{anti-} k_t \]

one ghost at every grid cell
Example: perturbative expansion of areas (at order $\alpha_s$)

$$\langle A(p_t, R) \rangle = A_0 + \frac{C_{F,A}}{b_0\pi} \pi R^2 d \log \left( \frac{\alpha_s(Q_0)}{\alpha_s(Rp_t)} \right)$$

- area $\neq \pi R^2$, area $\neq$ const.
- coefficients computable

<table>
<thead>
<tr>
<th></th>
<th>$A_0/\langle \pi R^2 \rangle$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>passive</td>
<td>active</td>
</tr>
<tr>
<td>$k_t$</td>
<td>1</td>
<td>0.81</td>
</tr>
<tr>
<td>Cam/Aachen</td>
<td>1</td>
<td>0.81</td>
</tr>
<tr>
<td>anti-$k_t$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SISCone</td>
<td>1</td>
<td>$1/4$</td>
</tr>
</tbody>
</table>

$Q_0 \equiv IR$ regulator $\propto$ background density
**Pileup subtraction** *(for uniform backgrounds)*

**Basic idea:** [M.Cacciari, G.Salam, 08]

\[ p_{t,\text{subtracted}} = p_{t,\text{jet}} - \rho_{\text{pileup}} \times \text{Area}_{\text{jet}} \]

- **Jet area:** [M.Cacciari, G.Salam, G.S., 08]
  - region where the jet catches infinitely soft particles (active/passive)
  - analytic control and understanding in pQCD

- **Pileup density per unit area:** \( \rho_{\text{pileup}} \)
  - e.g. estimated from the median of \( p_{t,\text{jet}} / \text{Area}_{\text{jet}} \)

- \( p_{t,\text{jet}} \)
- \( \text{Area}_{\text{jet}} \)
- \( \eta \)

![Graph](image-url)
**Pileup subtraction** *(for uniform backgrounds)*

**Basic idea:** [M. Cacciari, G. Salam, 08]

\[ p_{t,\text{subtracted}} = p_{t,\text{jet}} - \rho_{\text{pileup}} \times \text{Area}_{\text{jet}} \]

- **Jet area:** [M. Cacciari, G. Salam, G. S., 08]
  - region where the jet catches infinitely soft particles (active/passive)
  - analytic control and understanding in pQCD

- **Pileup density per unit area:** \( \rho_{\text{pileup}} \)
  - e.g. estimated from the median of \( \frac{p_{t,\text{jet}}}{\text{Area}_{\text{jet}}} \)

---

- P_{t,\text{jet}} / \text{Area}_{\text{jet}}

- \( \eta \) median

- hard jets

- background jets
**Pileup subtraction** *(for uniform backgrounds)*

**Basic idea:** [M.Cacciari, G.Salam, 08]

\[
pt_{\text{subtracted}} = pt_{\text{jet}} - \rho_{\text{pileup}} \times \text{Area}_{\text{jet}}
\]

- **Jet area:** [M.Cacciari, G.Salam, G.S., 08]
  - region where the jet catches infinitely soft particles (active/passive)
  - analytic control and understanding in pQCD

- **Pileup density per unit area:** \(\rho_{\text{pileup}}\)
  - e.g. estimated from the median of \(pt_{\text{jet}}/\text{Area}_{\text{jet}}\)

implemented in FastJet on an event-by-event basis
**Effect on dijet reconstruction**

Pileup unsubtracted

<table>
<thead>
<tr>
<th>1/N dN/dm (GeV⁻¹)</th>
<th>reconstructed Z' mass (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.07</td>
<td>29.5 GeV</td>
</tr>
<tr>
<td>0.06</td>
<td>21.0 GeV</td>
</tr>
<tr>
<td>0.05</td>
<td>17.7 GeV</td>
</tr>
</tbody>
</table>

✓ position reasonable
✓ dispersion reduced (thanks to the event-by-event approach)
✓ used by STAR for the first jet analysis in heavy-ions

pileup subtracted

<table>
<thead>
<tr>
<th>1/N dN/dm (GeV⁻¹)</th>
<th>reconstructed Z' mass (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.07</td>
<td>21.0 GeV</td>
</tr>
<tr>
<td>0.06</td>
<td>17.7 GeV</td>
</tr>
</tbody>
</table>

width = 29.5 GeV
width = 21.0 GeV
width = 17.7 GeV
Message #1:
Use infrared-and-collinear-safe algorithms

| ATLAS Cone | CDF/D0 MidPoint | $\rightarrow$ | SISCon | $\sqrt{\text{fast}}$ | $\sqrt{\text{safe}}$ |
| CMS Lt. Cone | $\rightarrow$ | anti-$k_t$ | $\sqrt{\text{fast}}$ | $\sqrt{\text{safe}}$ |

Important to benefit fully from pQCD multilegs/multiloops calculations
**Summary (2)**

**Message #2:**
- correct tools $\Rightarrow$ new ideas, new concepts
- $\Rightarrow$ new generation of jet definitions

- jet areas $\longrightarrow$ pileup and HI background subtraction
- jet substructure improves reconstruction (Higgs, top, SUSY, ...)

**Message #3:**
- keep some flexibility in the jet definition choice

- optimisation $\longrightarrow$ luminosity gains for LHC searches
- different approaches $\longrightarrow$ better understanding of HI collisions
backup slides
**The SISCones search for stable cones**

- **Solution**: use a seedless approach, find ALL stable cones

- **Naive approach**: check stability of each subset of particle
The SISCone search for stable cones

- **Solution**: use a seedless approach, find **ALL** stable cones

- **Naive approach**: check stability of each subset of particle
  Complexity is $O(N^{2N})$
  ⇒ definitely unrealistic: $10^{17}$ years for $N = 100$

- **Midpoint complexity**: $O(N^3)$
The SISCone search for stable cones

**Solution**: use a seedless approach, find **ALL** stable cones

**Midpoint complexity**: $\mathcal{O}(N^3)$

**Idea**: use geometric arguments

- Each enclosure can be moved (in any dir.) until it touches a point
- ... then rotated until it touches a second one

\[ \Rightarrow \text{Enumerate all pairs of particles} \]
\[ \quad \text{with 2 circle orientations and 4 possible inclusion/exclusion} \]
\[ \quad \longrightarrow \text{find all enclosures} \]
The SISCones search for stable cones

- **Solution**: use a seedless approach, find ALL stable cones

- **Midpoint complexity**: $O(N^3)$

**Idea**: use geometric arguments

⇒ Enumerate all pairs of particles

with 2 circle orientations and 4 possible inclusion/exclusion

→ find all enclosures

- **Complexity**: $O(N^3)$, with improvements: $O(N^2 \log(N))$

→ [C++ implementation: Seedless Infrared-Safe Cone algorithm (SISCones)]

G.Salam, G.S., JHEP 04 (2007) 086; http://projects.hepforge.org/siscone

NB.: also available from FastJet

[M.Cacciari, G.Salam, G.S.]; http://www.fastjet.fr
Recombination algorithms very fast

SISCones not slower than Midpoint (even with a 1 GeV seed threshold)

[M. Cacciari, G. Salam, 06]
A technical point: Back-reaction

Additional soft background has 2 effects:

- Throw soft particles in the hard jet: dealt with by subtraction
- Modify the hard scattering (back-reaction)
  - can be pointlike or diffuse
  - gain:

  \[
  \begin{align*}
  \text{no medium: } p_t &= p_{t1} \\
  \text{medium: } p_t &= p_{t1} + p_{t2} + p_{tm}
  \end{align*}
  \]

  \[
  \begin{align*}
  \text{no medium: } p_t &= p_{t1} + p_{t2} \\
  \text{medium: } p_t &= p_{t1} + p_{tm}
  \end{align*}
  \]
A technical point: Back-reaction

Additional soft background has 2 effects:

- Throw soft particles in the hard jet: dealt with by subtraction
- Modify the hard scattering (back-reaction)
  - can be pointlike or diffuse
  - tractable analytically (similar to areas)
  - $k_t \gtrsim$ Cambridge $>$ SIScone $\gg$ anti-$k_t$

![Graph showing $\Delta p_t$ distribution with different jet definitions](image)
Example: application to HI collisions

$pp +$ pileup

$AA$

 siscone (R=0.7)  

 antikt (R=0.6)
Framework for study

- **Hard event**: Pythia(v6.4) or Pythia(v6.4)+PyQuen(v1.5)
- **Background**: Hydjet(v1.5) (others under study)
- **Analysis**: FastJet(v2.4)
  - Ideally: smallest $\Delta p_t$ shift, smallest $\Delta p_t$ dispersion
- **Note**: in what follows, $R$ fixed to 0.4
Framework for study

- Hard event (quenched or unquenched)
- + Background event

Hard jets

\[ \Delta p_t \] average dispersion

- Cluster subtract
- Hard jets

Generic trends under control
Final numbers may change

Analysis: FastJet(v2.4)
Ideally: smallest \( \Delta p_t \) shift, smallest \( \Delta p_t \) dispersion

Note: in what follows, \( R \) fixed to 0.4

[M.Cacciari, J.Rojo, G.Salam, GS, in prep.]
Idea #1: use a local range to compute $\rho_{\text{bkg}}$

- Fluctuating background
  $\rightarrow$ determine the background density $\rho_{\text{bkg}}$
  from jets in the vicinity of the jet we want to subtract

- Exclude the hardest jets from the determination of $\rho_{\text{bkg}}$
  $\Rightarrow$ reduce the bias in the computation median
Effect of choosing a local range

- effect $\sim 0.5$-1 GeV
- differences between local ranges $\rightarrow$ uncertainty
- for limited acceptance, global range $\approx$ local range
- analytic control would be nice
Results: RHIC kinematics

- average $p_t$ shift:
  - anti-$k_t$ and C/A+filt. Ok
Results: RHIC kinematics

- average $p_t$ shift:
  - anti-$k_t$ and C/A+filt. Ok

- $p_t$ shift dispersion:
  - C/A+filt. better
Results: RHIC kinematics

- average $p_t$ shift:
  - anti-$k_t$ and C/A+filt. Ok
- $p_t$ shift dispersion:
  - C/A+filt. better
- watch out C/A+filt. average:
  - back-reaction compensated
Results: RHIC kinematics – quenching

Performances not much affected by quenching (need more models)
Results: LHC kinematics

- average $p_t$ shift:
  - anti-$k_t$ and C/A+filt. Ok
Results: LHC kinematics

- average $p_t$ shift:
  - anti-$k_t$ and C/A+filt. Ok

- $p_t$ shift dispersion:
  - C/A+filt. better
  - anti-$k_t$ Ok
Large quenching effect but anti-$k_t$'s rigidity plays for it

**LHC, unquenched |y|<2.4, R=0.4 Donut(R,3R) range**

**LHC, quenched Donut(R,3R) range**

**Preliminary**

Cam/Aachen

anti-$k_t$

Cam+filt.

**$p_t$ shift [GeV]**

**$p_t,hard$ [GeV]**

**$p_t$ dispersion [GeV]**

**$p_t,hard$ [GeV]**