New tools in jet physics:

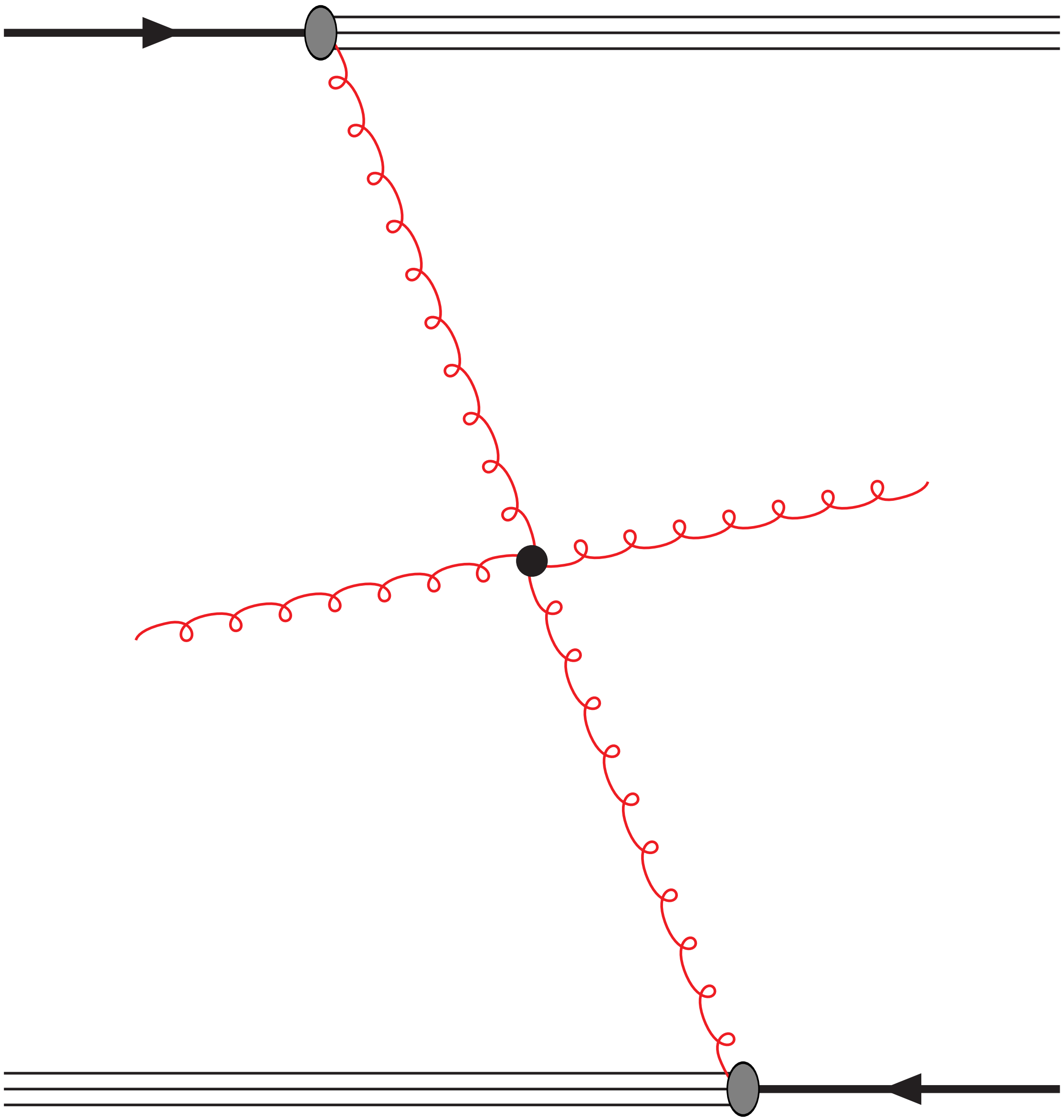
SISCone (new cone algorithm) - jet areas (new concept)

Grégory Soyez

Brookhaven National Laboratory

- in collaboration with Gavin Salam
- paper available as JHEP 05 (2007) 086 [arXiv:0704.0292]
- code available at `http://projects.hepforge.org/siscone`
  FastJet plugin: `http://www.lpthe.jussieu.fr/~salam/fastjet`
- area paper: Matteo Cacciari, Gavin Salam, G.S. arXiV:0802.1188]

General (over)simplified picture

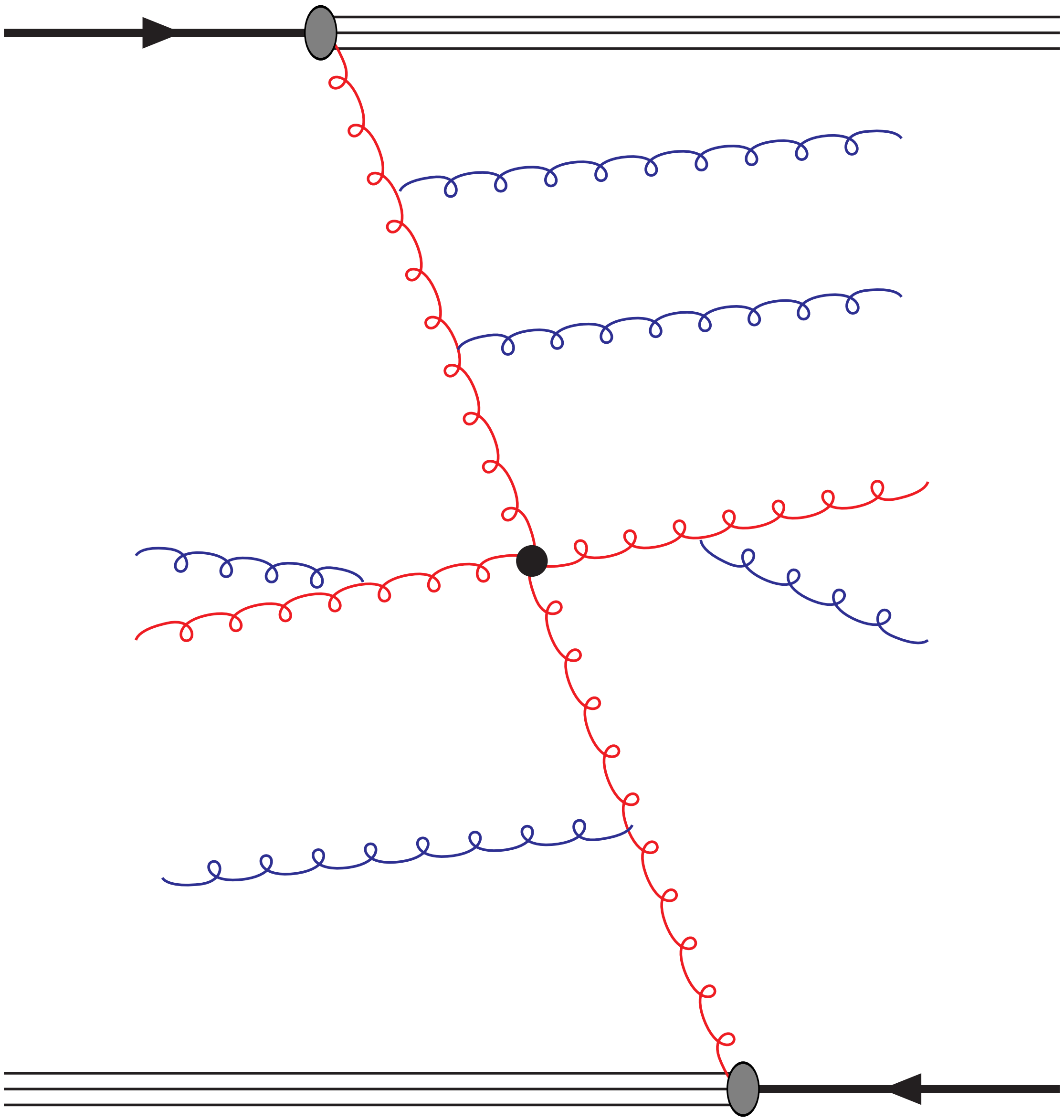


Perturbative level

Hard scattering $2 \to n$ computed exactly at $\mathcal{O}(\alpha_s^p)$

$gg \to gg$, $gg \to ggg$,
$gg \to gggg$,
$gg \to H \to b\bar{b}$,
$gg \to t\bar{t} \to \mu\nu_\mu b\bar{b}q\bar{q}$,
$gg \to Z' \to q\bar{q}$, ...

General (over)simplified picture



Perturbative level

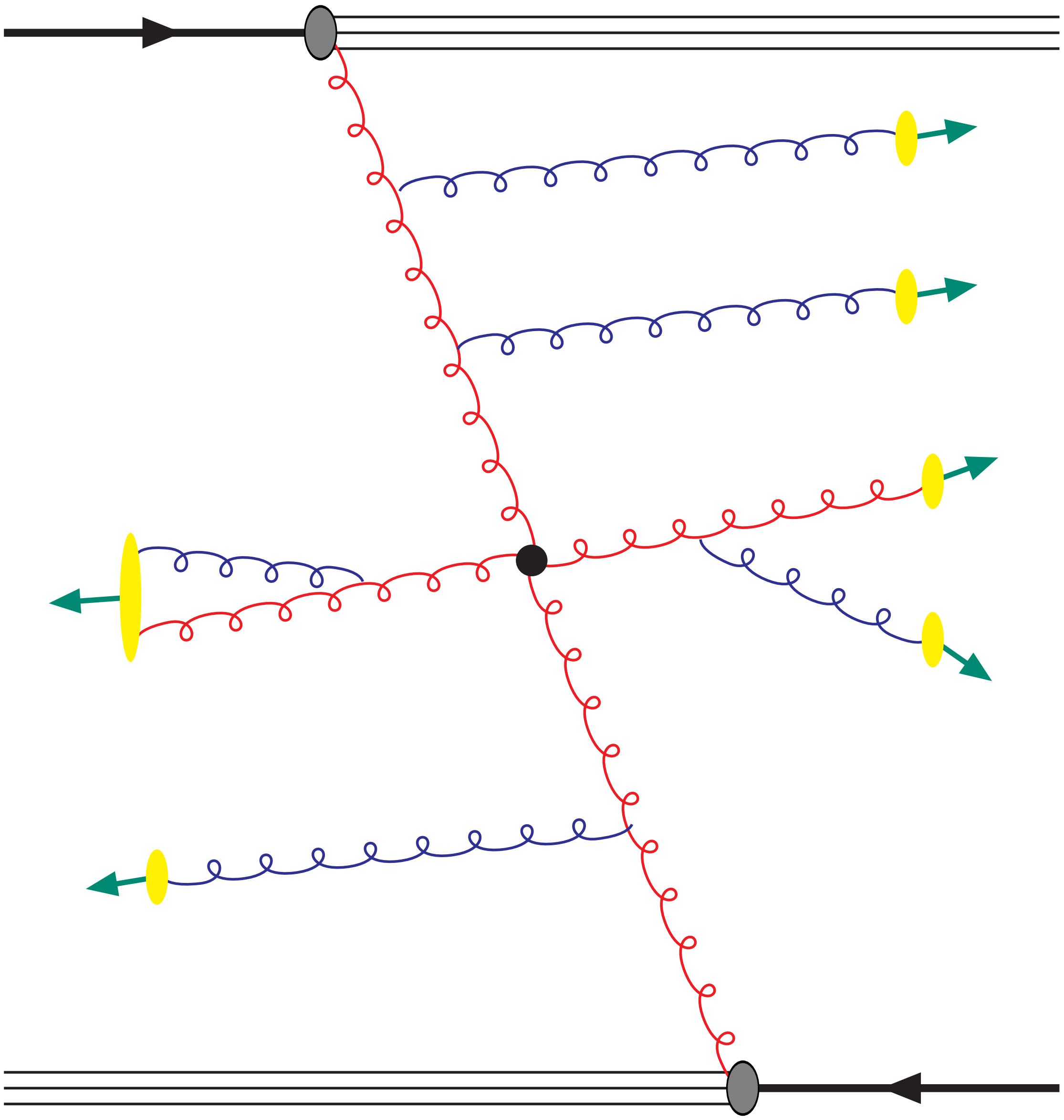
Parton level

$\approx$ collinear divergences

resummation

$\sum_i \alpha_s^i \log^i(p_t^2/\mu^2)$

General (over)simplified picture



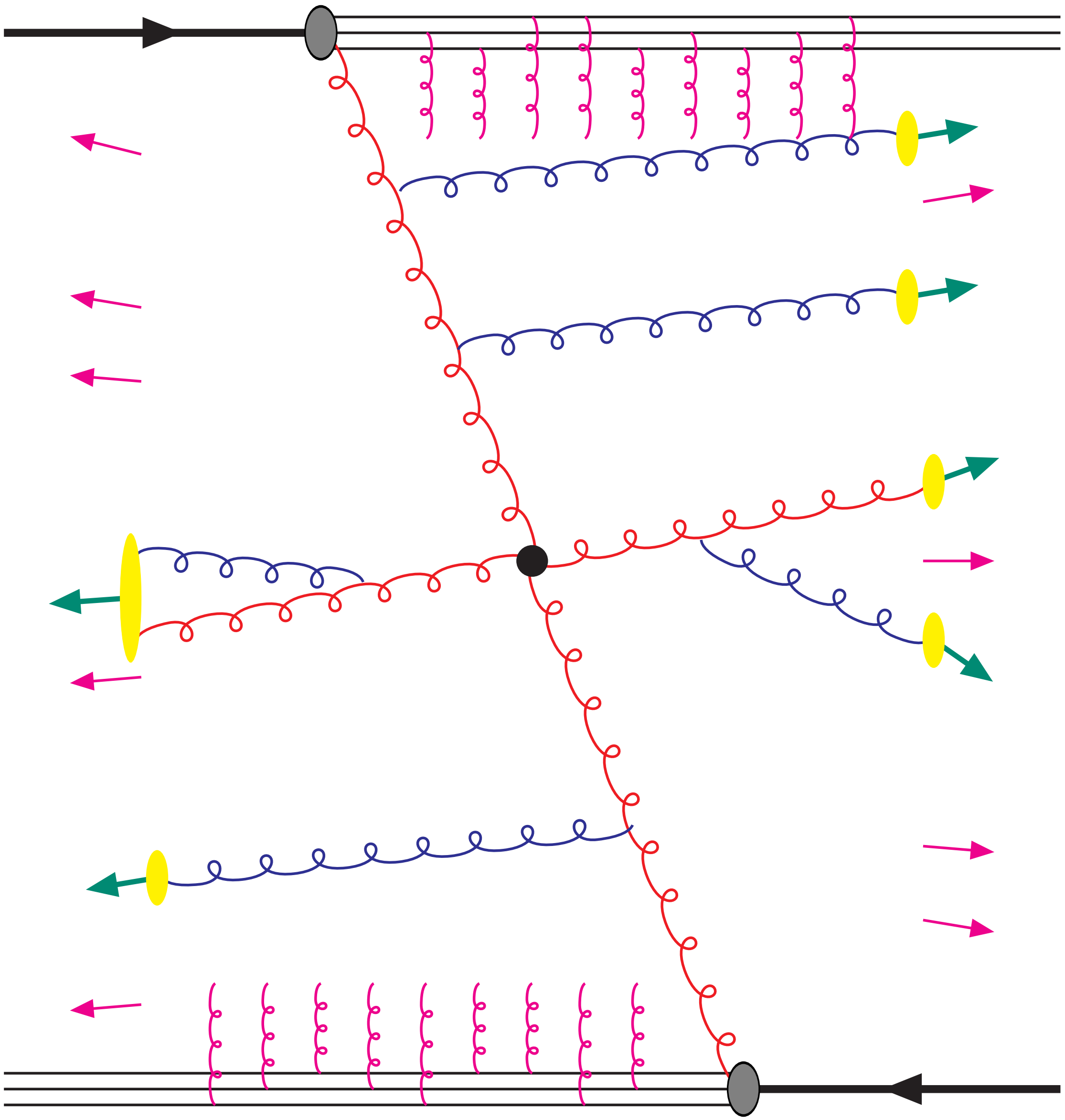
Perturbative level

Parton level

Hadron level

quarks+gluon → hadrons
(various models)

General (over)simplified picture



Perturbative level

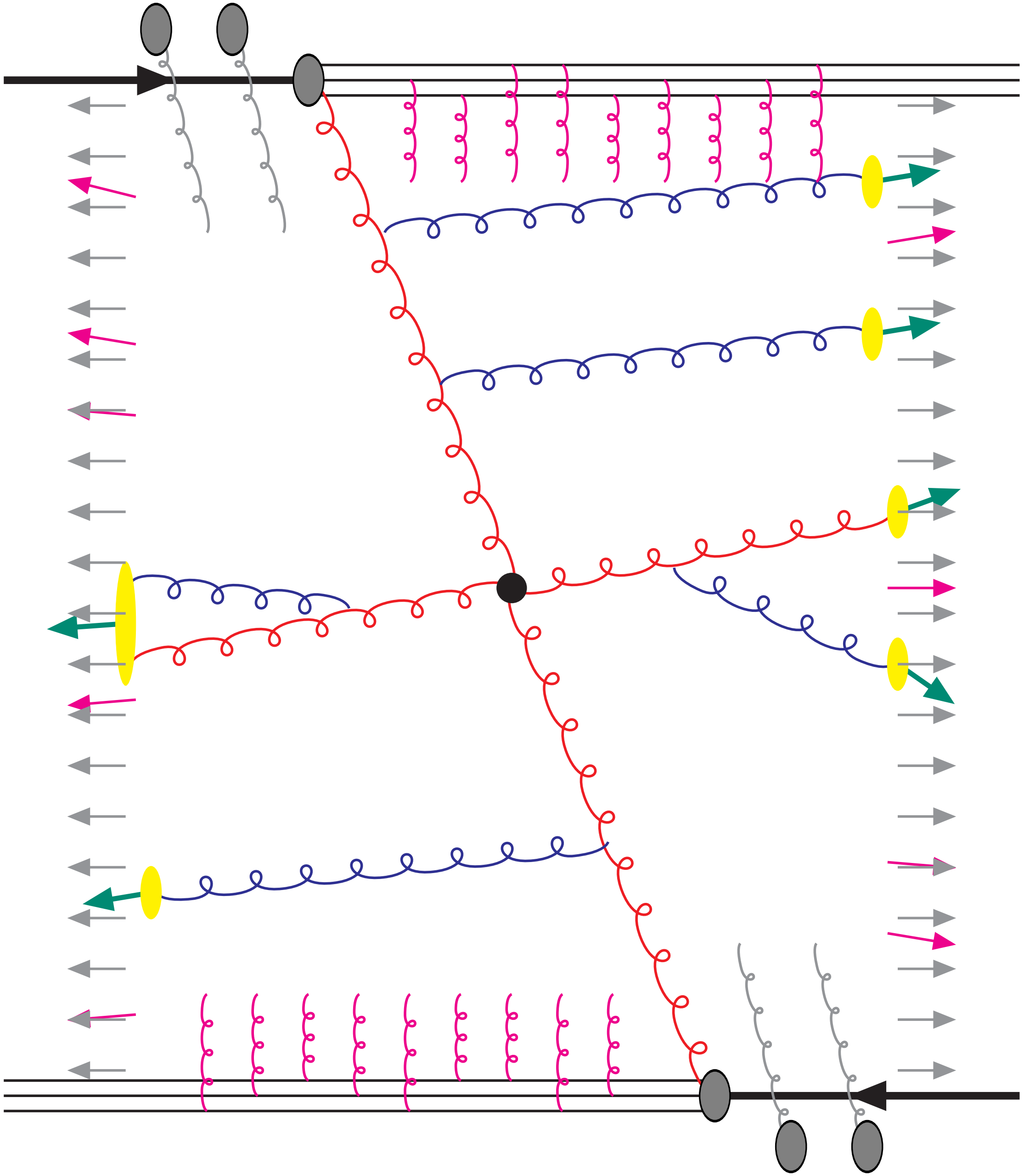
Parton level

Hadron level

+ Underlying event

Multiple interactions
from beam remnants
⇒ soft background

General (over)simplified picture



Perturbative level

Parton level

Hadron level

+ Underlying event

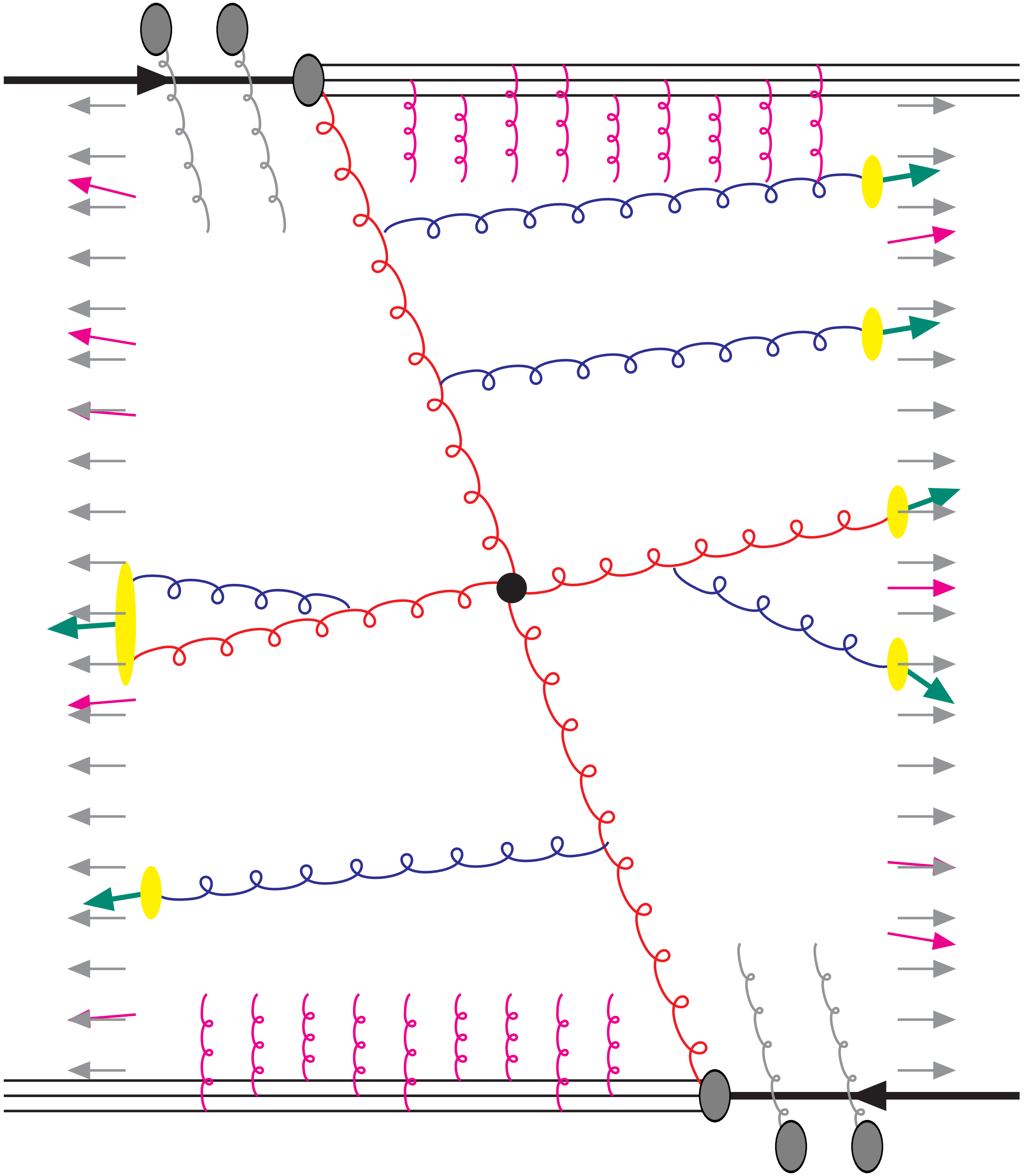
+ Pile up

additionnal pp interactions

$\Rightarrow$ soft background

$\approx$ uniform

General (over)simplified picture



Perturbative level

Parton level

Hadron level

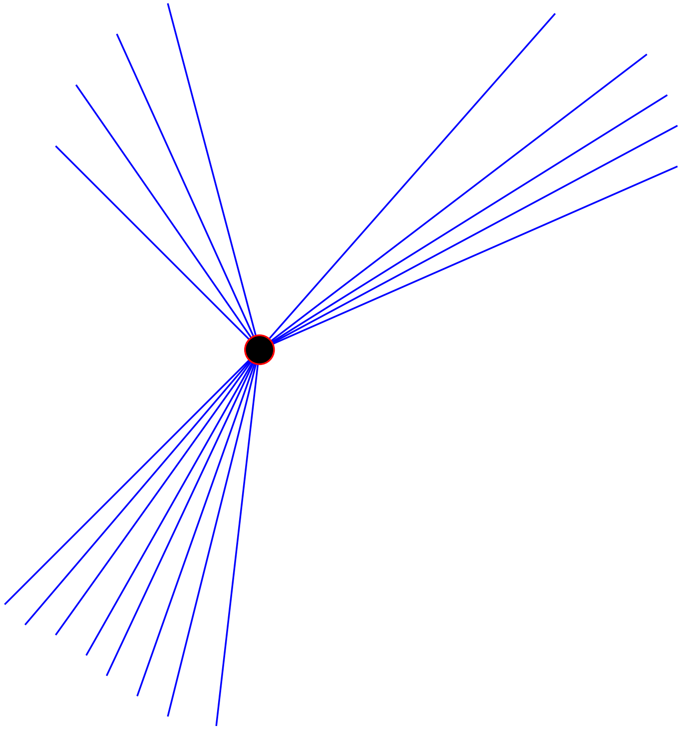
+ Underlying event

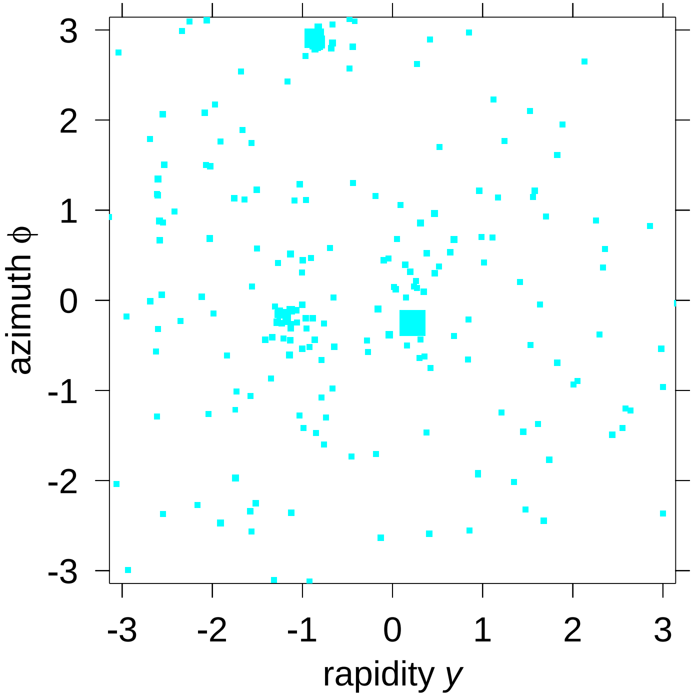
+ Pile up

How to access
the hard scattering?

Why jet algorithms?

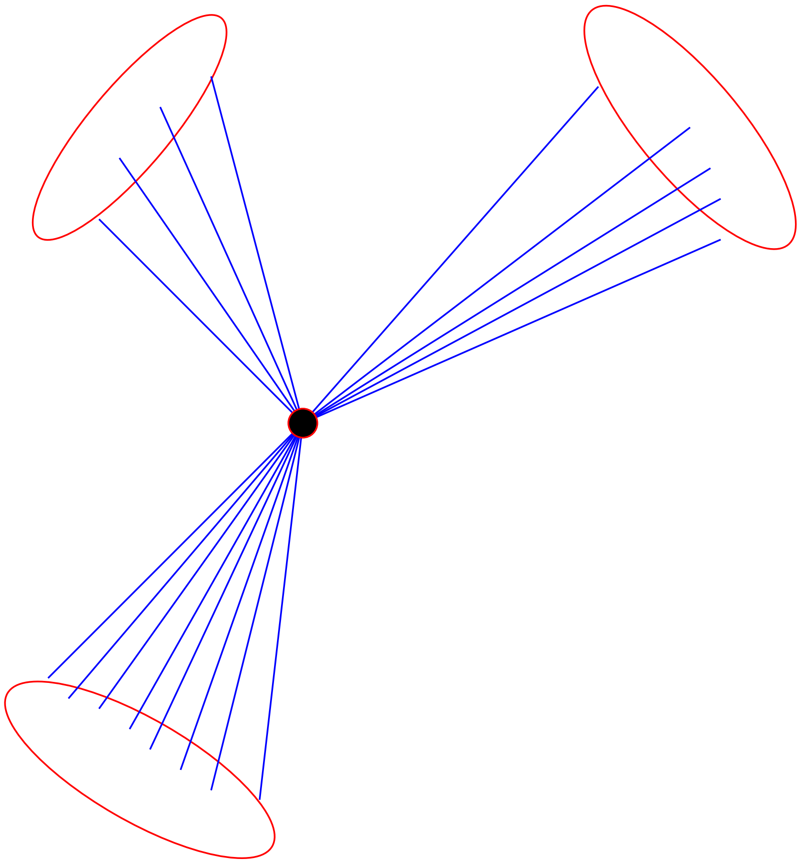
- Given: set of N particles with their 4-momentum

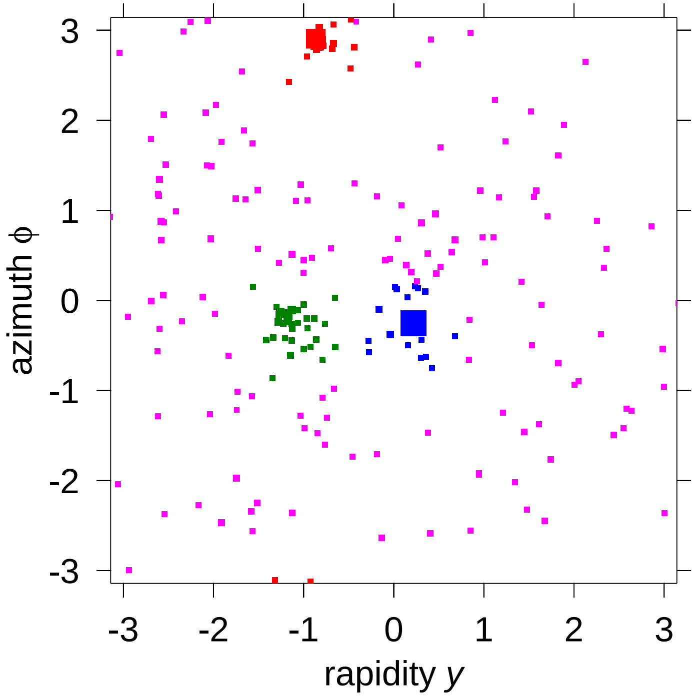




Why jet algorithms?

- Given: set of N particles with their 4-momentum
- Quest: clustering those particles into jets





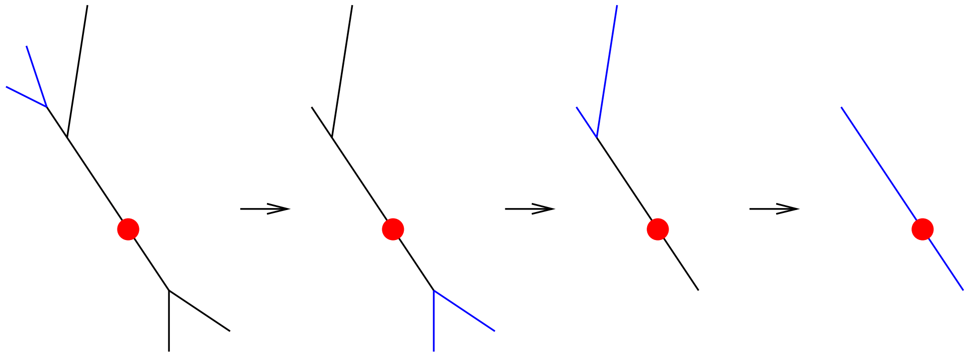
⇒ understand the original, perturbative, particle-level process

"Parton" not well defined ⇒ ambiguity in jet definition

Two classes of algorithms

Class 1: recombination

Successive recombinations of the “closest” pair of particle



- Distance:

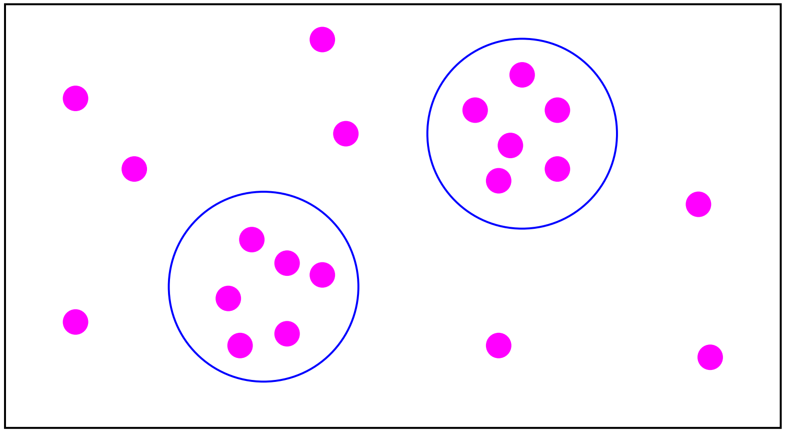
$$
\begin{aligned}
\underline{k_t}: &\quad d_{i,j} = \min(k_{t,i}^2, k_{t,j}^2)(\Delta\phi_{i,j}^2 + \Delta y_{i,j}^2) \\
\underline{\text{Aachen/Cam.}}: &\quad d_{i,j} = \Delta\phi_{i,j}^2 + \Delta y_{i,j}^2
\end{aligned}
$$

- stop when $d_{\text{min}} > R$

Two classes of algorithms

Class 2: cone

Find directions of dominant energy flow



for a cone of fixed radius R in the (y, ϕ) plane:

stable cones such that:

centre of the cone $\equiv$ direction of the total momentum of its particle contents

recombination vs. cone

	Recombination	Cone
Pro's Con's	Perturbative behaviour Sensitivity to radiation	Sensitivity to radiation Perturbative behaviour
Usage	$e^{\pm}e^{\pm}$ or $e^{\pm}p$	pp
	`FastJet`: fast implementation (M. Cacciari, G. Salam, G.S.)	Many: Snowmass, JetClu, PxCone, CDF Midpoint, ...

Outline

- Introduction: jet algorithms in general
- How does the cone work?
  - Generic description
  - Midpoint algorithm: description & IR unsafety
- SISCone: a practical solution
- Physical consequences:
  - Algorithm speed
  - Inclusive jet spectrum
  - Jet mass spectrum in multi-jet events
- Area of a jet
  - Definition and properties
  - Applications

Cone requirements

- Snowmass Accord (FERMILAB, 1990):
  any jet algorithm must satisfy
  1. Can be practically used in experimental analysis
  2. Can be practically used in theoretical computations
  3. Can be defined at any order of the perturbation theory
  4. Yields finite cross-sections at any order
  5. Has a small sensitivity to hadronisation corrections

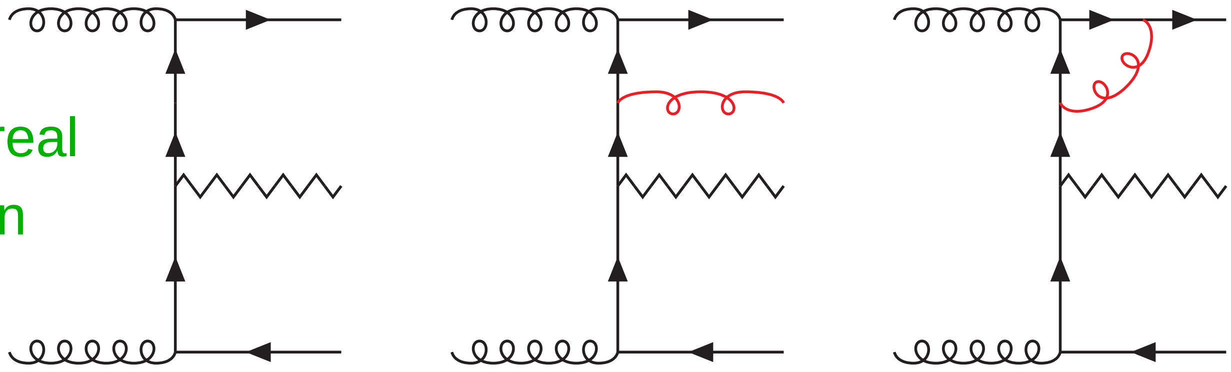
Cone requirements

- Snowmass Accord (FERMILAB, 1990):
  any jet algorithm must satisfy
  1. Can be practically used in experimental analysis
  2. Can be practically used in theoretical computations
  3. Can be defined at any order of the perturbation theory
  4. Yields finite cross-sections at any order
  5. Has a small sensitivity to hadronisation corrections
- Previous cone algorithms:
  - 1, 2 and 4 never satisfied together
  - 5 is unclear (Underlying event and R_{sep} issues)
- This talk: where is the failure + how to fix it.

Infrared Safety

Ellipsis: IR safety, i.e. stability upon emission of soft particles,
is required for perturbative computations to make sense!

IR divergences:
cancellation between real
and virtual SOFT gluon
emissions in QCD



LO NLO, real NLO, virtual

- IF Jet clustering is different in both cases, THEN the cancellation is not done and the result is not consistent with pQCD

  ⇒ Stable cones must not change upon addition of soft particles

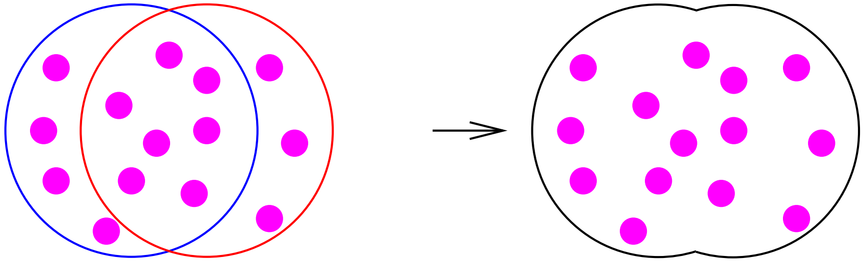
- Note: 100 GeV jet cannot change by adding a 1 GeV particle
  This would break parton/hadron correspondence

Modern cone jet algorithm

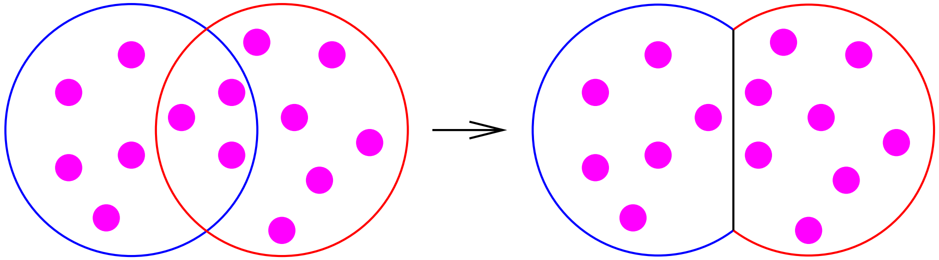
Modern cone jet algorithm (Tevatron Run II type):

- Step 1: find ALL stable cones of radius R
- Step 1': if some of the particles are not in stable cones,
  rerun Step 1 with the remaining ones.
- Step 2: run a split-merge procedure with overlap f
  to deal with overlapping stable cones

$\tilde{p}_{t,\text{shared}} > f\,\tilde{p}_{t,\text{min}}$



$\tilde{p}_{t,\text{shared}} \le f\,\tilde{p}_{t,\text{min}}$



Modern cone jet algorithm

Modern cone jet algorithm (Tevatron Run II type):

- Step 1: find ALL stable cones of radius R
- Step 1': if some of the particles are not in stable cones,
  rerun Step 1 with the remaining ones.
- Step 2: run a split-merge procedure with overlap f
  to deal with overlapping stable cones

Parameters:

- Standard parameters: cone radius R, overlap parameter f
- Additional controls: number of passes n_{pass}, stable cone $p_{t,\text{min}}$ cut-off

Modern cone jet algorithm

Modern cone jet algorithm (Tevatron Run II type):

- Step 1: find ALL stable cones of radius R
- Step 1': if some of the particles are not in stable cones, rerun Step 1 with the remaining ones.
- Step 2: run a split-merge procedure with overlap f to deal with overlapping stable cones

Parameters:

- Standard parameters: cone radius R, overlap parameter f
- Additional controls: number of passes n_{pass}, stable cone $p_{t,\text{min}}$ cut-off

Here, infrared safety means:

adding infinitely soft particles does not modify the stable cones found.

NB.: addition of infinitely soft particles does not modify the set of stable cone, the question is “does it modify the set of stable cones FOUND by our algorithm?”

Typical cone: Midpoint algorithm

Usual seeded method to search stable cones: midpoint cone algorithm

- For an initial seed
  1. sum the momenta of all particles within the cone centred on the seed
  2. use the direction of that momentum as new seed
  3. repeat 1 & 2 until stable state cone reached

Typical cone: Midpoint algorithm

Usual seeded method to search stable cones: midpoint cone algorithm

- For an initial seed
  1. sum the momenta of all particles within the cone centred on the seed
  2. use the direction of that momentum as new seed
  3. repeat 1 & 2 until stable state cone reached
- Sets of seeds:
  1. All particles (above a p_t threshold s) (JetClu)
  2. Midpoints between stable cones found in 1.

Typical cone: Midpoint algorithm

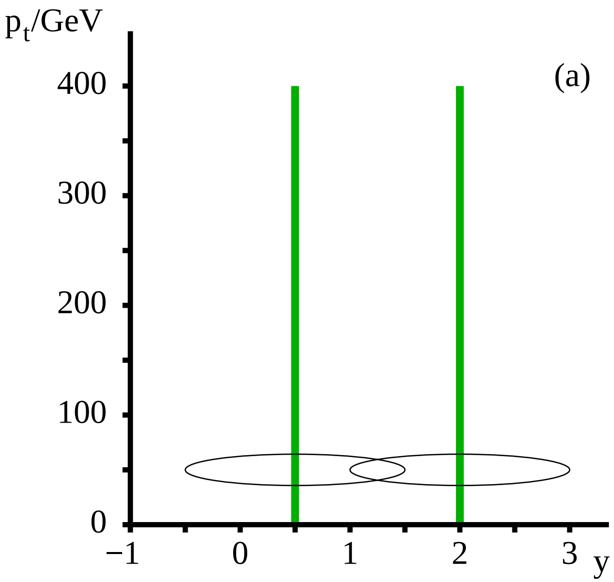
Usual seeded method to search stable cones: midpoint cone algorithm

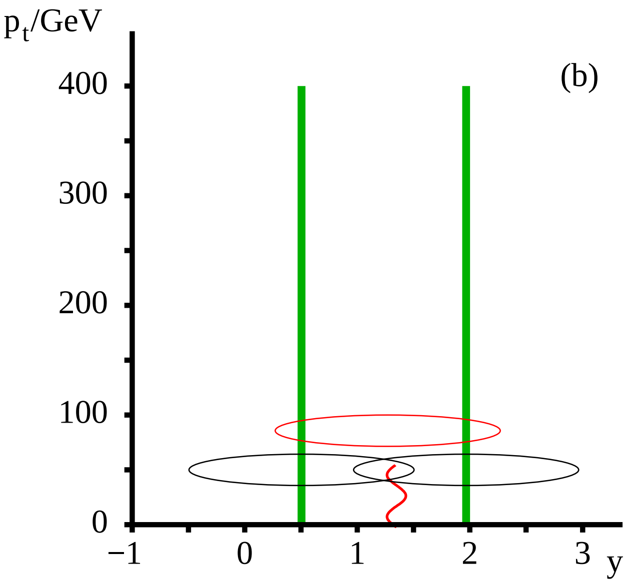
- For an initial seed
  1. sum the momenta of all particles within the cone centred on the seed
  2. use the direction of that momentum as new seed
  3. repeat 1 & 2 until stable state cone reached
- Sets of seeds:
  1. All particles (above a p_t threshold s) (JetClu)
  2. Midpoints between stable cones found in 1.

Problems:

- the p_t threshold s is collinear unsafe
- seeded approach ⇒ stable cones missed ⇒ infrared unsafety

JetClu IR Unsafety (R=1)

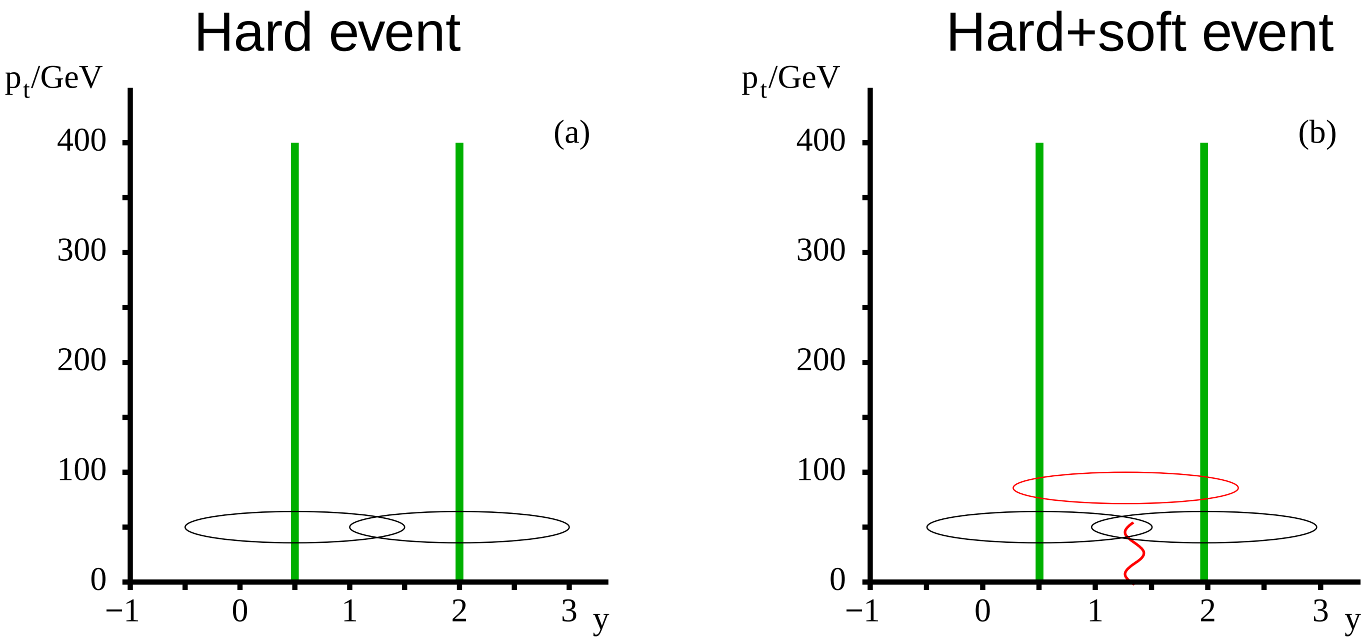

Hard event


Hard+soft event

Stable cones:

JetClu: {1} & {2} {1} & {2} & {1,2}

JetClu IR Unsafety (R=1)



Stable cones:

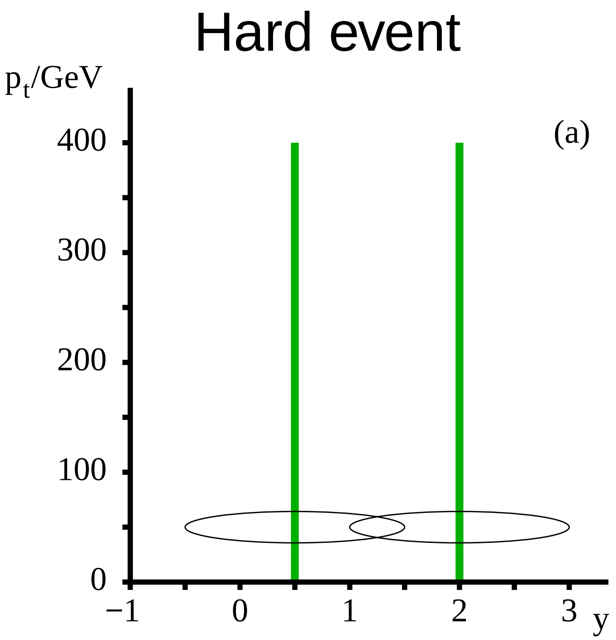
JetClu: {1} & {2} {1} & {2} & {1,2}

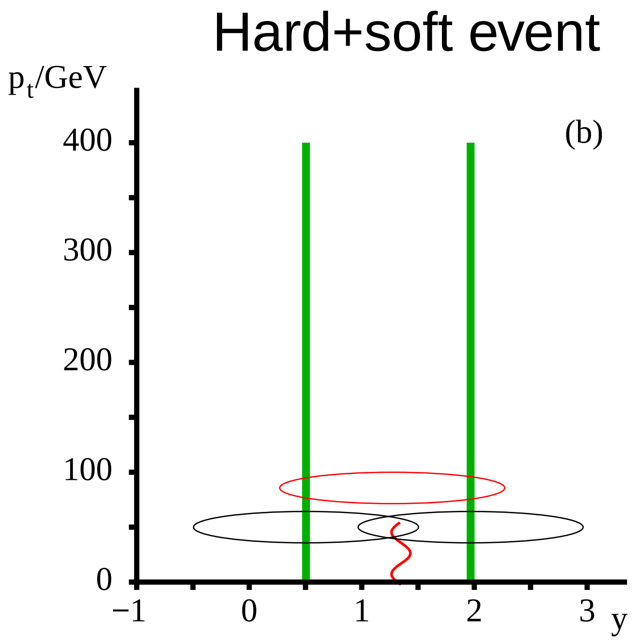
Jets: ($f = 0.5$)

JetClu: {1} & {2} {1,2}

Stable cone missed ⟶ IR unsafety of the JetClu algorithm

JetClu IR Unsafety (R=1)


Hard event

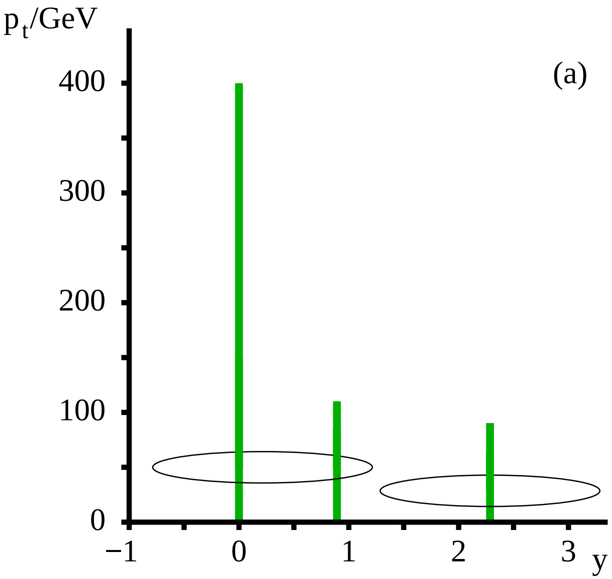

Hard+soft event

	Hard event	Hard+soft event
Stable cones:		
JetClu:	{1} & {2}	{1} & {2} & {1,2}
Midpoint:	{1} & {2} & {1,2}	{1} & {2} & {1,2}
Jets: ($f = 0.5$)		
JetClu:	{1} & {2}	{1,2}
Midpoint:	{1,2}	{1,2}

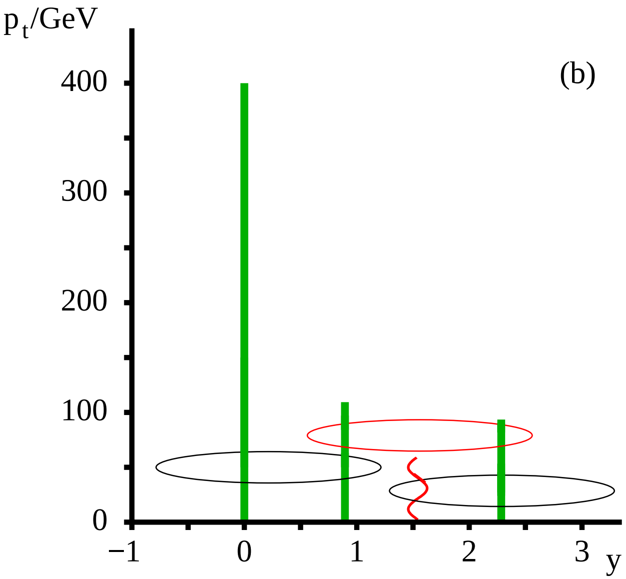
Stable cone missed ⟶ IR unsafety of the JetClu algorithm

Midpoint IR Unsafety (R=1)

Hard event



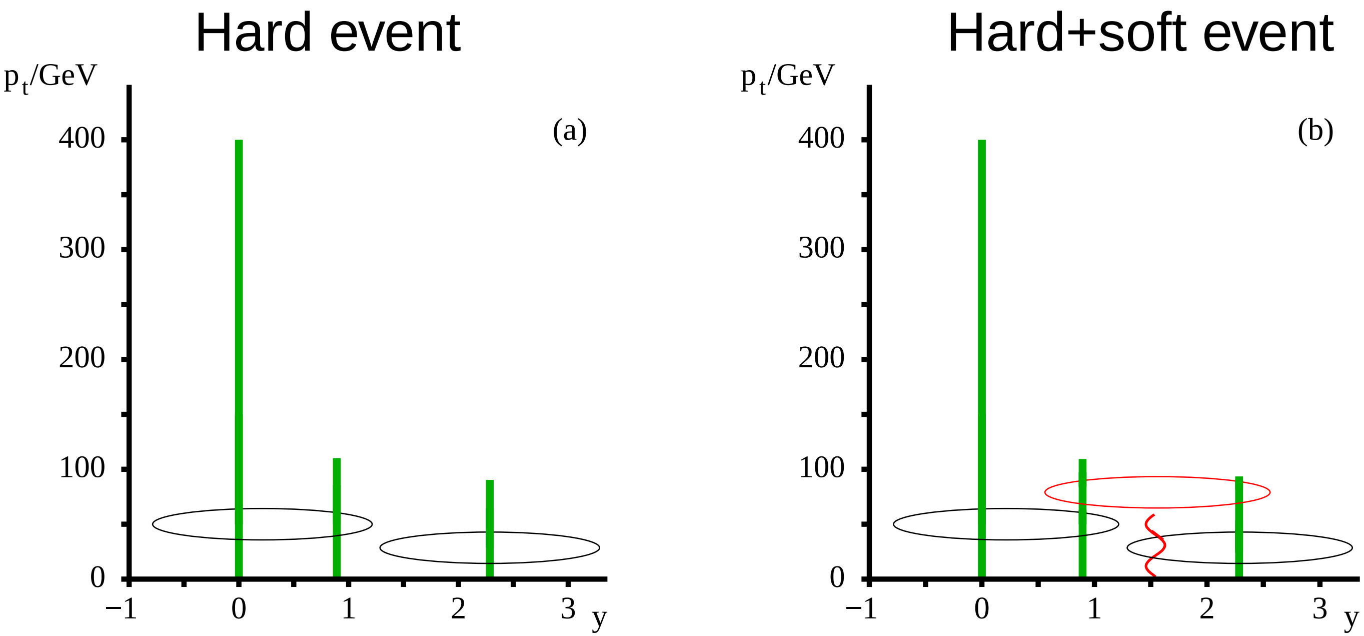
Hard+soft event



Stable cones:

	Hard event	Hard+soft event
Midpoint:	{1,2} & {3}	{1,2} & {3} & {2,3}

Midpoint IR Unsafety (R=1)



Stable cones:

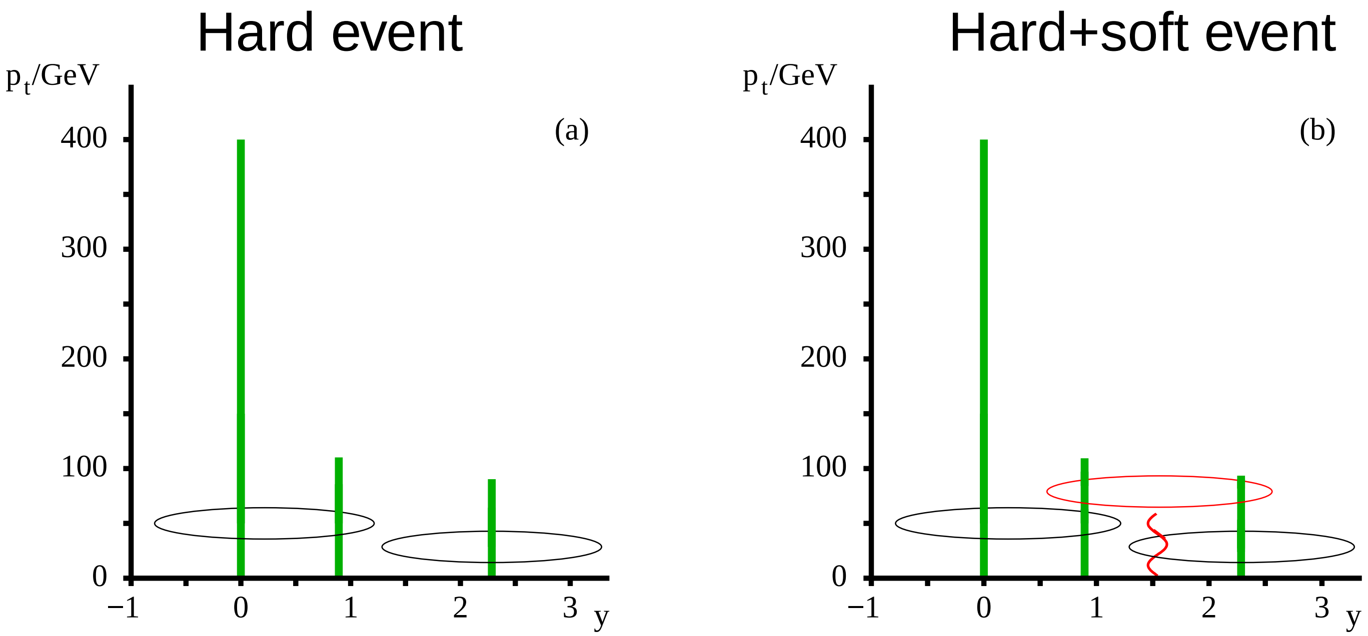
Midpoint: {1,2} & {3}   {1,2} & {3} & {2,3}

Jets: ($f = 0.5$)

Midpoint: {1,2} & {3}   {1,2,3}

Stable cone missed ⟶ IR unsafety of the midpoint algorithm

Midpoint IR Unsafety (R=1)



	Hard event	Hard+soft event
Stable cones:		
Midpoint:	{1,2} & {3}	{1,2} & {3} & {2,3}
Seedless:	{1,2} & {3} & {2,3}	{1,2} & {3} & {2,3}
Jets: ($f = 0.5$)		
Midpoint:	{1,2} & {3}	{1,2,3}
Seedless:	{1,2,3}	{1,2,3}

Stable cone missed $\longrightarrow$ IR unsafety of the midpoint algorithm

is a seedless solution practical?

- Solution: use a seedless approach, find ALL stable cones
- Naive approach: check stability of each subset of particle

is a seedless solution practical?

- Solution: use a seedless approach, find ALL stable cones
- Naive approach: check stability of each subset of particle
  Complexity is $\mathcal{O}\left(N2^N\right)$
  $\Rightarrow$ definitely unrealistic: 10^{17} years for $N = 100$

is a seedless solution practical?

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

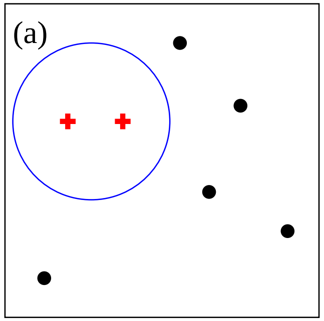
- Naive approach: check stability of each subset of particle
  Complexity is $\mathcal{O}\left(N2^N\right)$
  $\Rightarrow$ definitely unrealistic: 10^{17} years for $N = 100$

- Midpoint complexity:
  - For 1 seed: build and check cone content is $\mathcal{O}\left(N\right)$
  - initially N seeds $\Rightarrow \mathcal{O}\left(N\right)$ stable cones
    $\Rightarrow \mathcal{O}\left(N^2\right)$ new, midpoint, seeds
    $\Rightarrow$ midpoint complexity is $\mathcal{O}\left(N^3\right)$
  - Note: the number of stable cones is $\mathcal{O}\left(N\right)$

The SISCone algorithm

SISCone: seedless solution

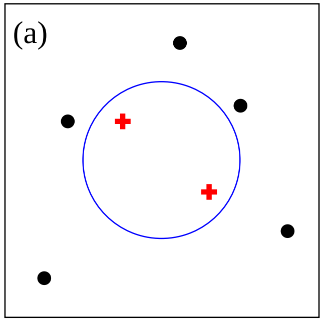
Idea: use geometric arguments



- Enumerate enclosures and check if they are stable

SISCone: seedless solution

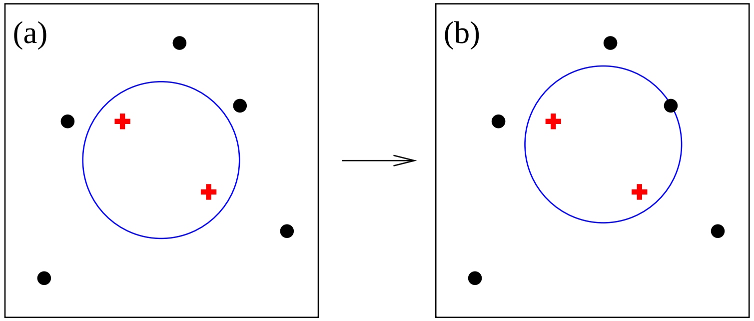
Idea: use geometric arguments



- Enumerate enclosures and check if they are stable

SISCone: seedless solution

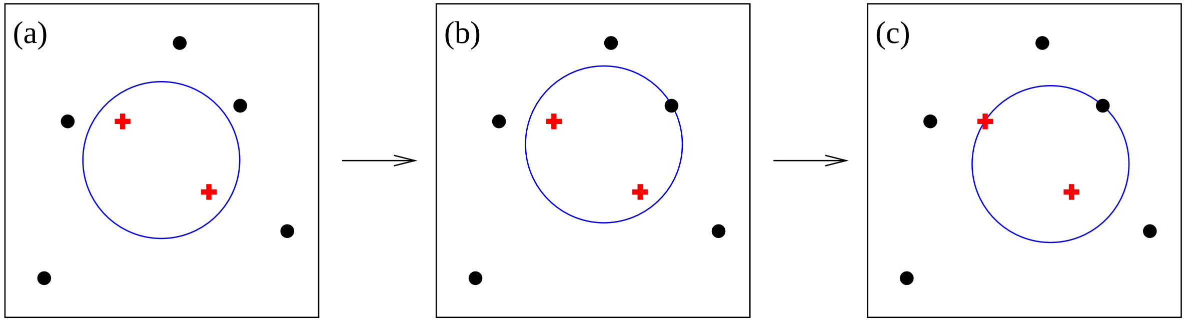
Idea: use geometric arguments



- Enumerate enclosures and check if they are stable
- Each enclosure can be moved (in any direction) until it touches a point

SISCone: seedless solution

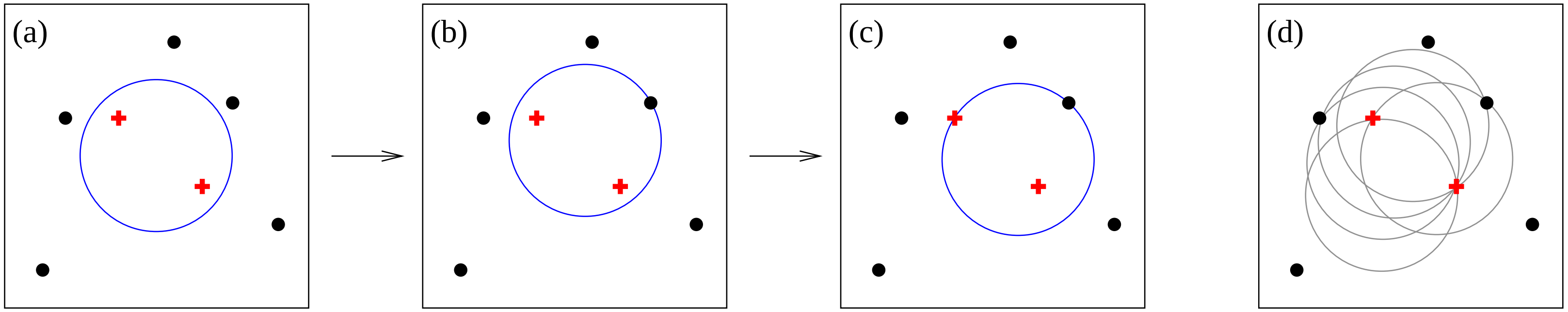
Idea: use geometric arguments



- Enumerate enclosures and check if they are stable
- Each enclosure can be moved (in any direction) until it touches a point
- ... then rotated until it touches a second one

SISCone: seedless solution

Idea: use geometric arguments



- Enumerate enclosures and check if they are stable
- Each enclosure can be moved (in any direction) until it touches a point
- ... then rotated until it touches a second one

$\Rightarrow$ Enumerate all pairs of particles
with 2 circle orientations and 4 possible inclusion/exclusion
$\longrightarrow$ find all enclosures

SISCone: seedless solution

$\Rightarrow$ Enumerate all pairs of particles
with 2 circle orientations and 4 possible inclusion/exclusion
$\longrightarrow$ find all enclosures

Complexity?

- Enumerate all pairs of particles: $\mathcal{O}\left(N^2\right)$
- For each, build content and check stability
  $\Rightarrow \mathcal{O}\left(N^3\right)$

SISCone: seedless solution

$\Rightarrow$ Enumerate all pairs of particles
with 2 circle orientations and 4 possible inclusion/exclusion
$\longrightarrow$ find all enclosures

Complexity?

- Enumerate all pairs of particles: $\mathcal{O}\left(N^2\right)$
- For each, build content and check stability
  $\Rightarrow \mathcal{O}\left(N^3\right)$

Same as midpoint... but we'll use more tricks ...

SISCone: seedless solution

Tricks:

- For all enclosures around a particle, introduce a traversal order
  $\Rightarrow$ avoids recomputing the cone contents at each step
- Only test "border particles" for stability (cost $\mathcal{O}(1)$)
  (q-bit tag + checkxor to keep trace of stability tests)
  $\Rightarrow$ limits the number of full stability test to $\mathcal{O}(N)$
- Total: saves a factor of $\mathcal{O}(N)$ but get a $\mathcal{O}(\log N)$ from the ordering

SISCone: seedless solution

Tricks:

- For all enclosures around a particle, introduce a traversal order
  $\Rightarrow$ avoids recomputing the cone contents at each step

- Only test “border particles” for stability (cost $\mathcal{O}\left(1\right)$)

  (q-bit tag + checkxor to keep trace of stability tests)

  $\Rightarrow$ limits the number of full stability test to $\mathcal{O}\left(N\right)$

- Total: saves a factor of $\mathcal{O}\left(N\right)$ but get a $\mathcal{O}\left(\log N\right)$ from the ordering

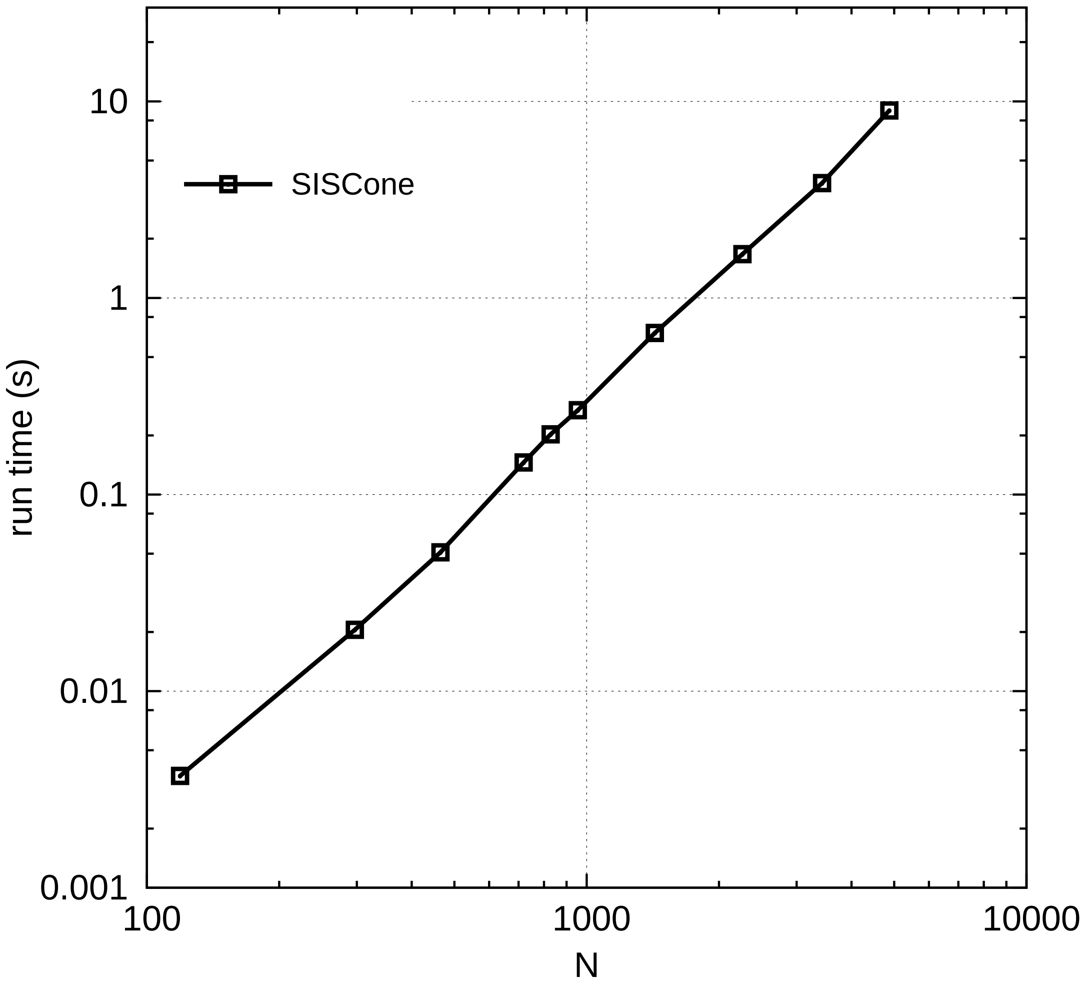
$$\boxed{\boxed{\text{All stable cones found in } \mathcal{O}\left(N^2\log(N)\right)}}$$

$\longrightarrow$ C++ implementation: Seedless Infrared-Safe Cone algorithm (SISCone)

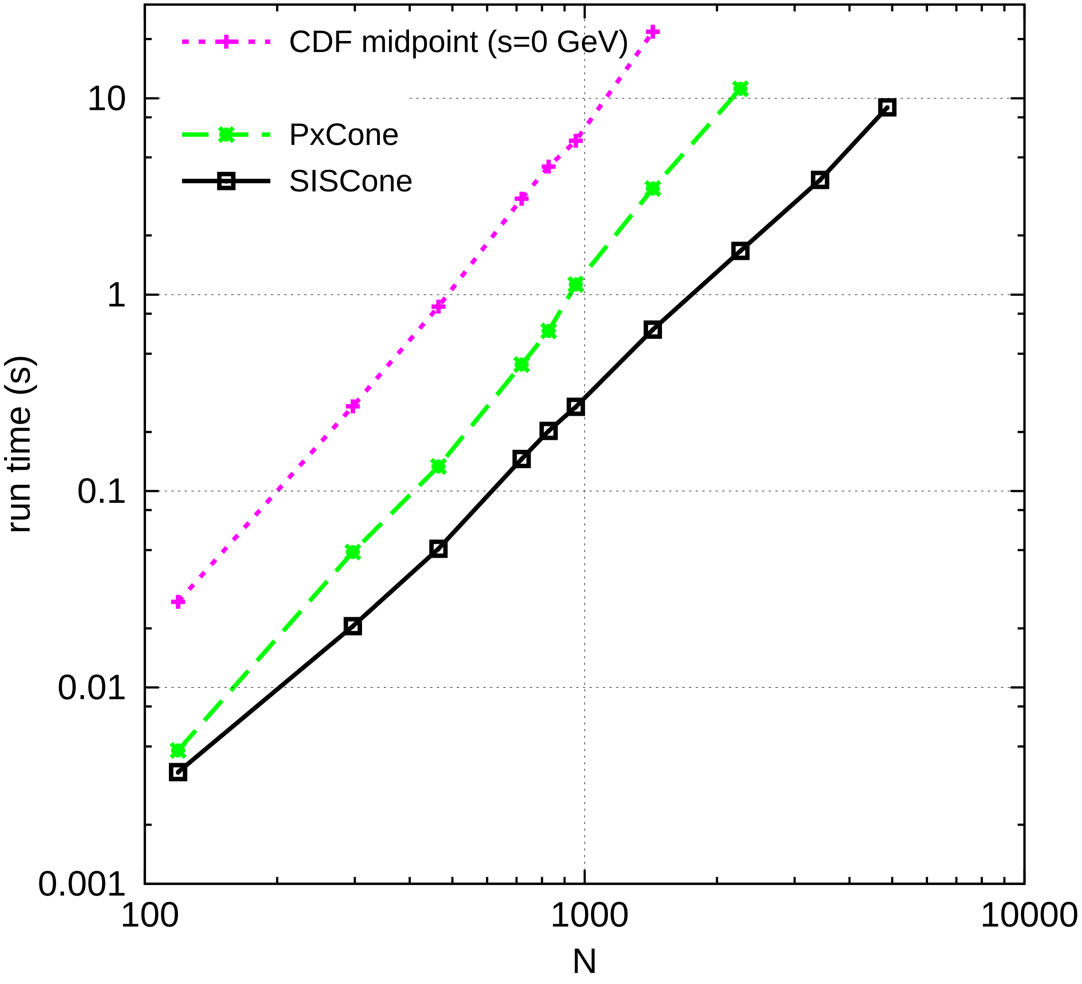
SISCone vs. other cone algorithms

implications of a seedless cone

Speed

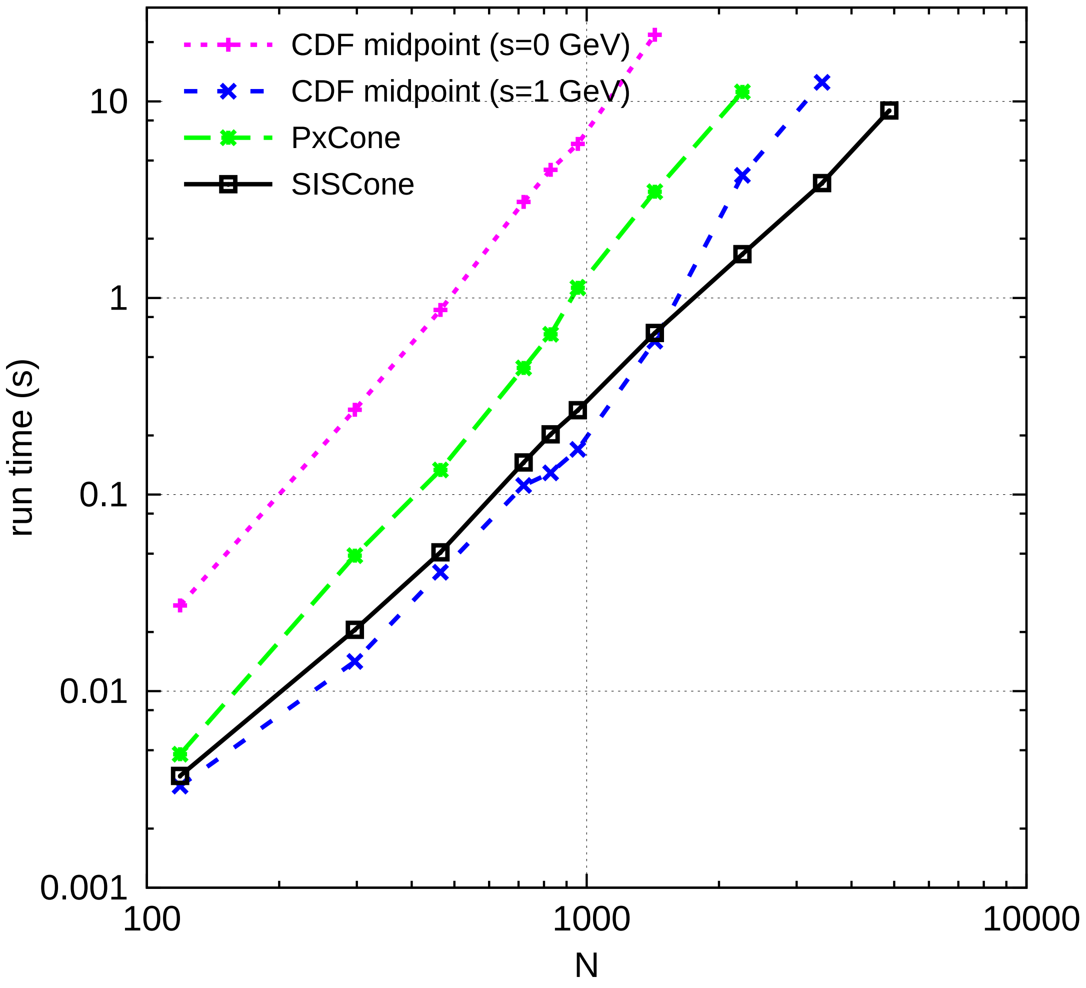


Speed



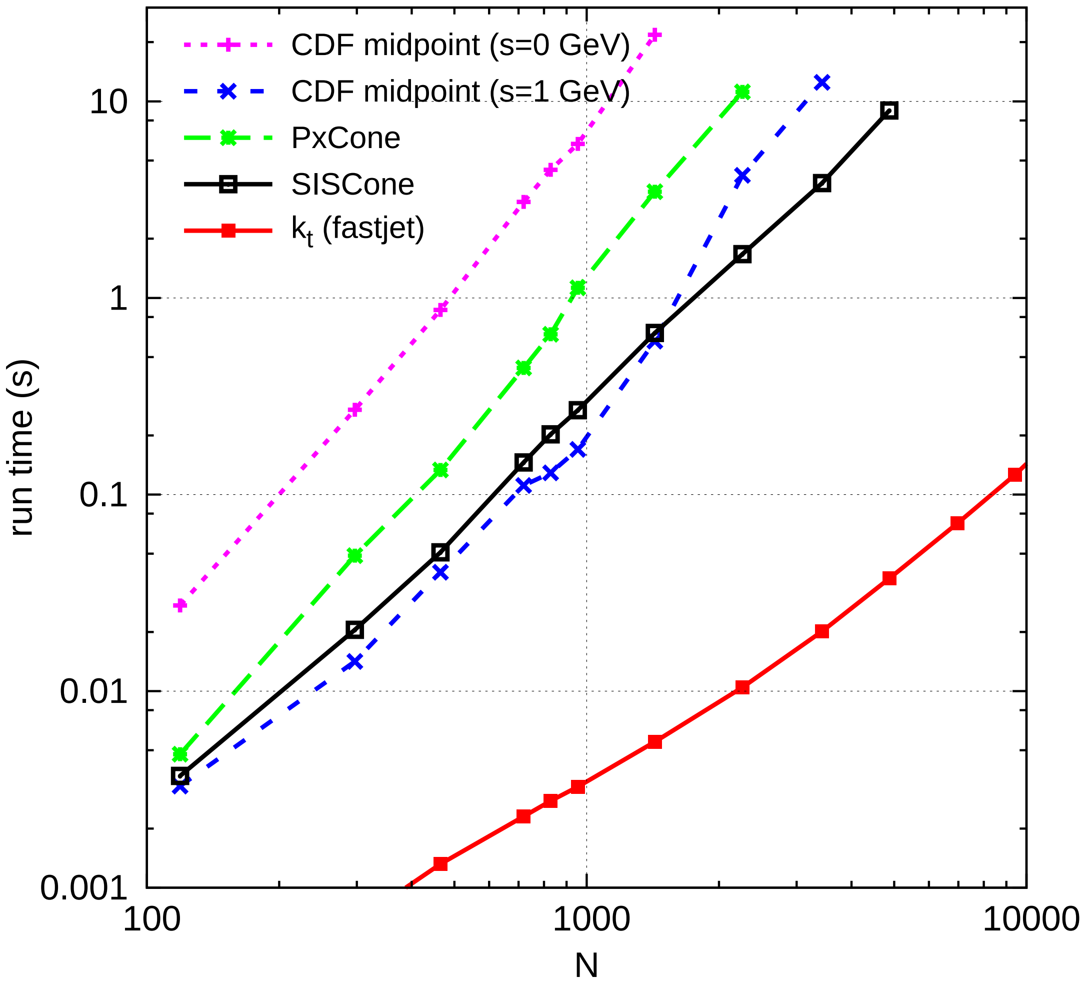
- faster than midpoint with no seed threshold and IR safe

Speed



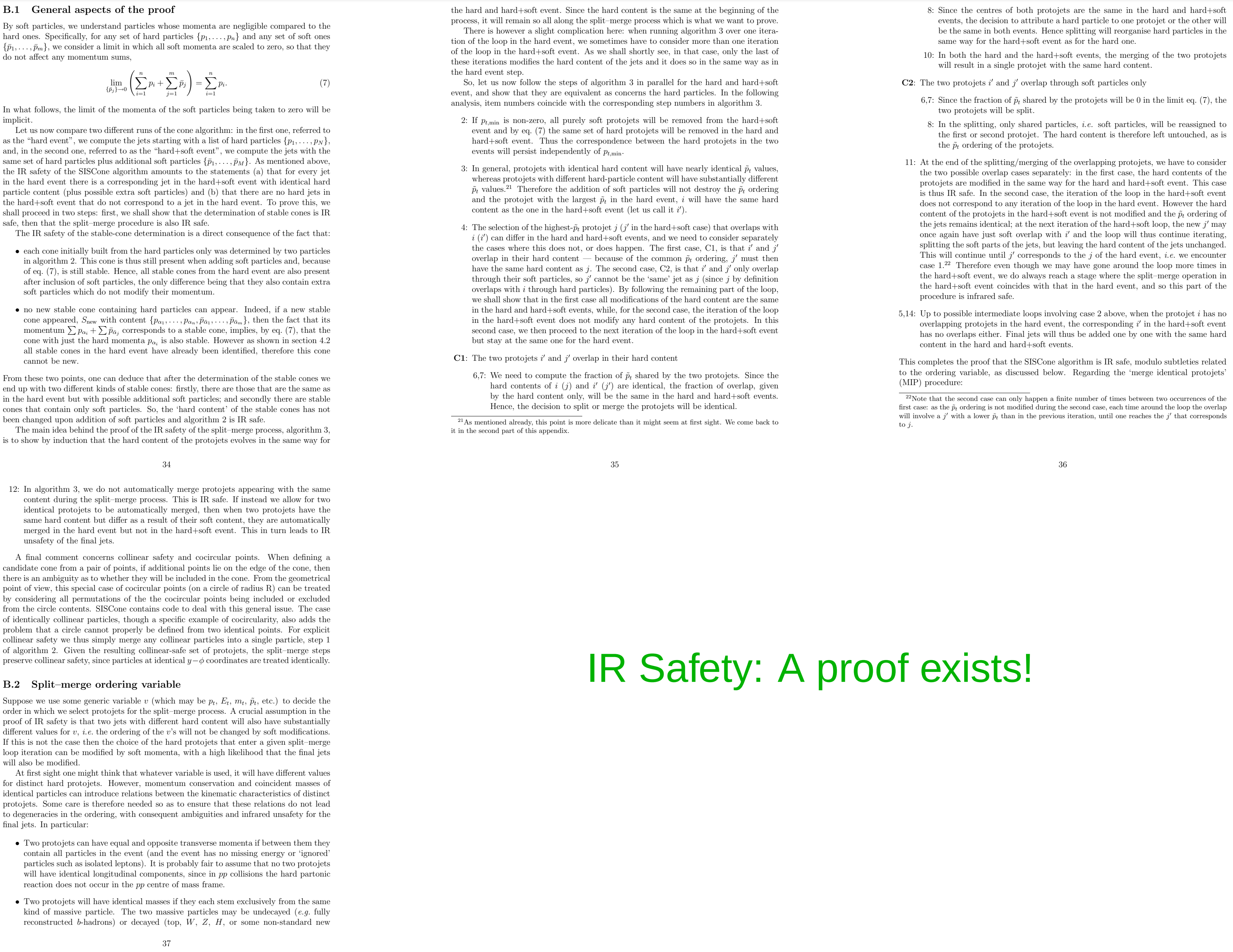
- faster than midpoint with no seed threshold and IR safe
- same as midpoint with 1 GeV seed and collinear safe

Speed



- faster than midpoint with no seed threshold and IR safe
- same as midpoint with 1 GeV seed and collinear safe
- slower that k_t/FastJet affordable for practical usage e.g. at the LHC

SISCone IR safety



B.1 General aspects of the proof

By soft particles, we understand particles whose momenta are negligible compared to the hard ones. Specifically, for any set of hard particles $\{p_1, \dots, p_n\}$ and any set of soft ones $\{\bar{p}_1, \dots, \bar{p}_m\}$, we consider a limit in which all soft momenta are scaled to zero, so that they do not affect any momentum sums,

$$\lim_{\{\bar{p}_j\}\to 0} \left(\sum_{i=1}^{n} p_i + \sum_{j=1}^{m} \bar{p}_j \right) = \sum_{i=1}^{n} p_i. \qquad (7)$$

In what follows, the limit of the momenta of the soft particles being taken to zero will be implicit.

Let us now compare two different runs of the cone algorithm: in the first one, referred to as the “hard event”, we compute the jets starting with a list of hard particles $\{p_1, \dots, p_N\}$, and, in the second one, referred to as the “hard+soft event”, we compute the jets with the same set of hard particles plus additional soft particles $\{\bar{p}_1, \dots, \bar{p}_M\}$. As mentioned above, the IR safety of the SISCone algorithm amounts to the statements (a) that for every jet in the hard event there is a corresponding jet in the hard+soft event with identical hard particle content (plus possible extra soft particles) and (b) that there are no hard jets in the hard+soft event that do not correspond to a jet in the hard event. To prove this, we shall proceed in two steps: first, we shall show that the determination of stable cones is IR safe, then that the split–merge procedure is also IR safe.

The IR safety of the stable-cone determination is a direct consequence of the fact that:

- each cone initially built from the hard particles only was determined by two particles in algorithm 2. This cone is thus still present when adding soft particles and, because of eq. (7), is still stable. Hence, all stable cones from the hard event are also present after inclusion of soft particles, the only difference being that they also contain extra soft particles which do not modify their momentum.
- no new stable cone containing hard particles can appear. Indeed, if a new stable cone appeared, S_{new} with content $\{p_{\alpha_1}, \dots, p_{\alpha_n}, \bar{p}_{\bar{\alpha}_1}, \dots, \bar{p}_{\bar{\alpha}_m}\}$, then the fact that its momentum $\sum p_{\alpha_i} + \sum \bar{p}_{\bar{\alpha}_j}$ corresponds to a stable cone, implies, by eq. (7), that the cone with just the hard momenta p_{α_i} is also stable. However as shown in section 4.2 all stable cones in the hard event have already been identified, therefore this cone cannot be new.

From these two points, one can deduce that after the determination of the stable cones we end up with two different kinds of stable cones: firstly, there are those that are the same as in the hard event but with possible additional soft particles; and secondly there are stable cones that contain only soft particles. So, the ‘hard content’ of the stable cones has not been changed upon addition of soft particles and algorithm 2 is IR safe.

The main idea behind the proof of the IR safety of the split–merge process, algorithm 3, is to show by induction that the hard content of the protojets evolves in the same way for

the hard and hard+soft event. Since the hard content is the same at the beginning of the process, it will remain so all along the split–merge process which is what we want to prove.

There is however a slight complication here: when running algorithm 3 over one iteration of the loop in the hard event, we sometimes have to consider more than one iteration of the loop in the hard+soft event. As we shall shortly see, in that case, only the last of these iterations modifies the hard content of the jets and it does so in the same way as in the hard event step.

So, let us now follow the steps of algorithm 3 in parallel for the hard and hard+soft event, and show that they are equivalent as concerns the hard particles. In the following analysis, item numbers coincide with the corresponding step numbers in algorithm 3.

2: If $p_{t,\min}$ is non-zero, all purely soft protojets will be removed from the hard+soft event and by eq. (7) the same set of hard protojets will be removed in the hard and hard+soft event. Thus the correspondence between the hard protojets in the two events will persist independently of $p_{t,\min}$.

3: In general, protojets with identical hard content will have nearly identical $\tilde{p}_t$ values, whereas protojets with different hard-particle content will have substantially different $\tilde{p}_t$ values.[21] Therefore the addition of soft particles will not destroy the $\tilde{p}_t$ ordering and the protojet with the largest $\tilde{p}_t$ in the hard event, i will have the same hard content as the one in the hard+soft event (let us call it i').

4: The selection of the highest-$\tilde{p}_t$ protojet j (j' in the hard+soft case) that overlaps with i (i') can differ in the hard and hard+soft events, and we need to consider separately the cases where this does not, or does happen. The first case, C1, is that i' and j' overlap in their hard content — because of the common $\tilde{p}_t$ ordering, j' must then have the same hard content as j. The second case, C2, is that i' and j' only overlap through their soft particles, so j' cannot be the ‘same’ jet as j (since j by definition overlaps with i through hard particles). By following the remaining part of the loop, we shall show that in the first case all modifications of the hard content are the same in the hard and hard+soft events, while, for the second case, the iteration of the loop in the hard+soft event does not modify any hard content of the protojets. In this second case, we then proceed to the next iteration of the loop in the hard+soft event but stay at the same one for the hard event.

C1: The two protojets i' and j' overlap in their hard content

  6,7: We need to compute the fraction of $\tilde{p}_t$ shared by the two protojets. Since the hard contents of i (j) and i' (j') are identical, the fraction of overlap, given by the hard content only, will be the same in the hard and hard+soft events. Hence, the decision to split or merge the protojets will be identical.

---

[21] As mentioned already, this point is more delicate than it might seem at first sight. We come back to it in the second part of this appendix.

  8: Since the centres of both protojets are the same in the hard and hard+soft events, the decision to attribute a hard particle to one protojet or the other will be the same in both events. Hence splitting will reorganise hard particles in the same way for the hard+soft event as for the hard one.

  10: In both the hard and the hard+soft events, the merging of the two protojets will result in a single protojet with the same hard content.

C2: The two protojets i' and j' overlap through soft particles only

  6,7: Since the fraction of $\tilde{p}_t$ shared by the protojets will be 0 in the limit eq. (7), the two protojets will be split.

  8: In the splitting, only shared particles, *i.e.* soft particles, will be reassigned to the first or second protojet. The hard content is therefore left untouched, as is the $\tilde{p}_t$ ordering of the protojets.

11: At the end of the splitting/merging of the overlapping protojets, we have to consider the two possible overlap cases separately: in the first case, the hard contents of the protojets are modified in the same way for the hard and hard+soft event. This case is thus IR safe. In the second case, the iteration of the loop in the hard+soft event does not correspond to any iteration of the loop in the hard event. However the hard content of the protojets in the hard+soft event is not modified and the $\tilde{p}_t$ ordering of the jets remains identical; at the next iteration of the hard+soft loop, the new j' may once again have just soft overlap with i' and the loop will thus continue iterating, splitting the soft parts of the jets, but leaving the hard content of the jets unchanged. This will continue until j' corresponds to the j of the hard event, *i.e.* we encounter case 1.[22] Therefore even though we may have gone around the loop more times in the hard+soft event, we do always reach a stage where the split–merge operation in the hard+soft event coincides with that in the hard event, and so this part of the procedure is infrared safe.

5,14: Up to possible intermediate loops involving case 2 above, when the protojet i has no overlapping protojets in the hard event, the corresponding i' in the hard+soft event has no overlaps either. Final jets will thus be added one by one with the same hard content in the hard and hard+soft events.

This completes the proof that the SISCone algorithm is IR safe, modulo subtleties related to the ordering variable, as discussed below. Regarding the ‘merge identical protojets’ (MIP) procedure:

---

[22] Note that the second case can only happen a finite number of times between two occurrences of the first case: as the $\tilde{p}_t$ ordering is not modified during the second case, each time around the loop the overlap will involve a j' with a lower $\tilde{p}_t$ than in the previous iteration, until one reaches the j' that corresponds to j.

12: In algorithm 3, we do not automatically merge protojets appearing with the same content during the split–merge process. This is IR safe. If instead we allow for two identical protojets to be automatically merged, then when two protojets have the same hard content but differ as a result of their soft content, they are automatically merged in the hard event but not in the hard+soft event. This in turn leads to IR unsafety of the final jets.

A final comment concerns collinear safety and cocircular points. When defining a candidate cone from a pair of points, if additional points lie on the edge of the cone, then there is an ambiguity as to whether they will be included in the cone. From the geometrical point of view, this special case of cocircular points (on a circle of radius R) can be treated by considering all permutations of the the cocircular points being included or excluded from the circle contents. SISCone contains code to deal with this general issue. The case of identically collinear particles, though a specific example of cocircularity, also adds the problem that a circle cannot properly be defined from two identical points. For explicit collinear safety we thus simply merge any collinear particles into a single particle, step 1 of algorithm 2. Given the resulting collinear-safe set of protojets, the split–merge steps preserve collinear safety, since particles at identical $y-\phi$ coordinates are treated identically.

B.2 Split–merge ordering variable

Suppose we use some generic variable v (which may be p_t, E_t, m_t, $\tilde{p}_t$, etc.) to decide the order in which we select protojets for the split–merge process. A crucial assumption in the proof of IR safety is that two jets with different hard content will also have substantially different values for v, *i.e.* the ordering of the v's will not be changed by soft modifications. If this is not the case then the choice of the hard protojets that enter a given split–merge loop iteration can be modified by soft momenta, with a high likelihood that the final jets will also be modified.

At first sight one might think that whatever variable is used, it will have different values for distinct hard protojets. However, momentum conservation and coincident masses of identical particles can introduce relations between the kinematic characteristics of distinct protojets. Some care is therefore needed so as to ensure that these relations do not lead to degeneracies in the ordering, with consequent ambiguities and infrared unsafety for the final jets. In particular:

- Two protojets can have equal and opposite transverse momenta if between them they contain all particles in the event (and the event has no missing energy or ‘ignored’ particles such as isolated leptons). It is probably fair to assume that no two protojets will have identical longitudinal components, since in *pp* collisions the hard partonic reaction does not occur in the *pp* centre of mass frame.
- Two protojets will have identical masses if they each stem exclusively from the same kind of massive particle. The two massive particles may be undecayed (*e.g.* fully reconstructed b-hadrons) or decayed (top, W, Z, H, or some non-standard new

IR Safety: A proof exists!

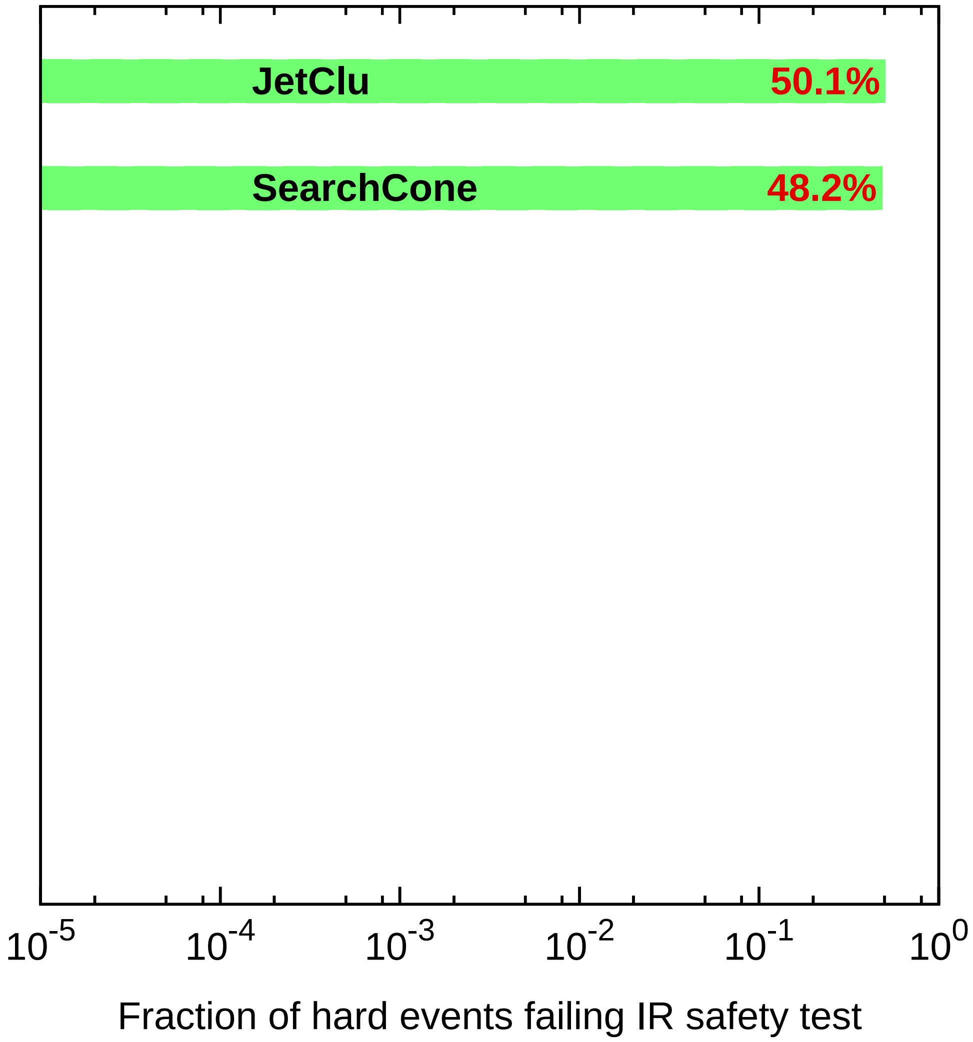
IR Unsafety failure rates

- Hard event: 2-10 particles
- Soft add-on: 1-5 particles
- Run:
  - “hard” only
  - many “hard+soft” trials
  - Search differences

IR Unsafety failure rates

- Hard event: 2-10 particles
- Soft add-on: 1-5 particles
- Run:
  - “hard” only
  - many “hard+soft” trials
  - Search differences

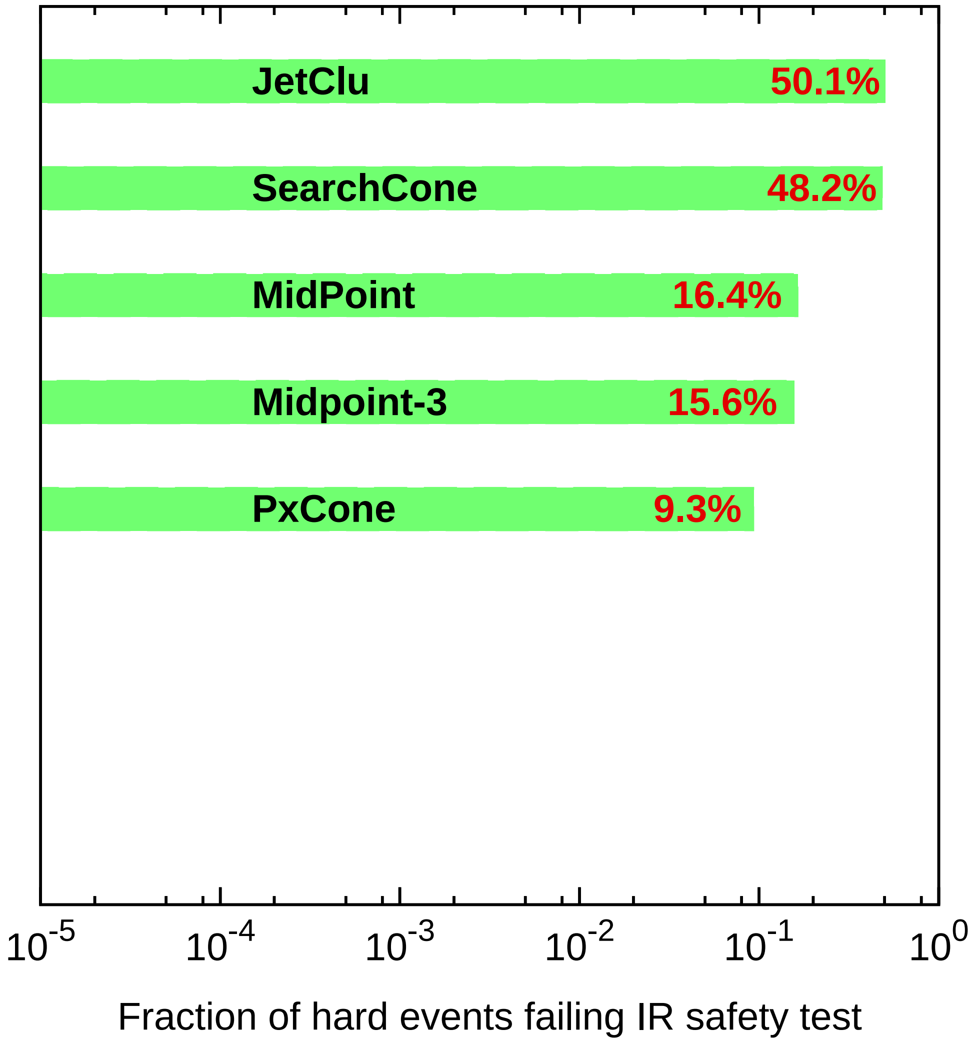
Unsafety level	failure rate
2 hard + 1 soft	$\sim 50\%$



IR Unsafety failure rates

- Hard event: 2-10 particles
- Soft add-on: 1-5 particles
- Run:
  - “hard” only
  - many “hard+soft” trials
  - Search differences

Unsafety level	failure rate
2 hard + 1 soft	$\sim 50\%$
3 hard + 1 soft	$\sim 15\%$

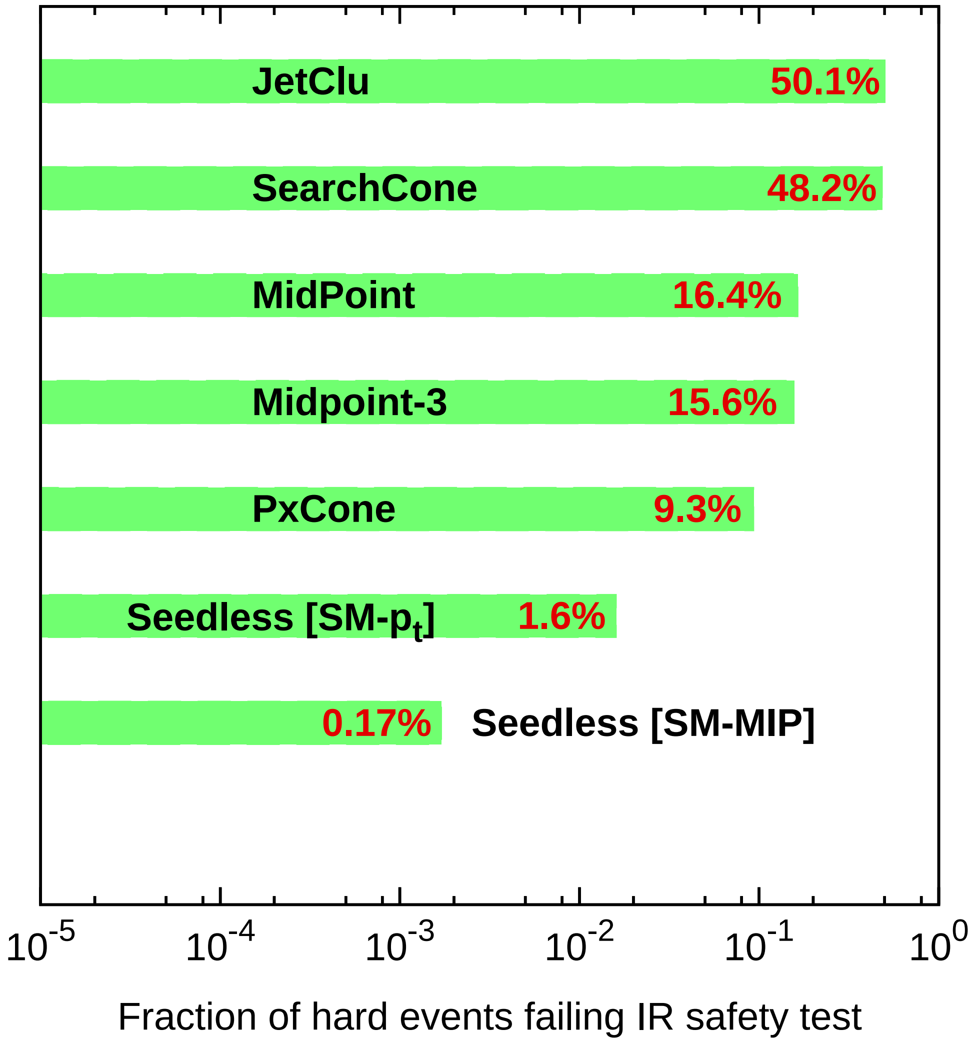


IR Unsafety failure rates

- Hard event: 2-10 particles
- Soft add-on: 1-5 particles
- Run:
    - “hard” only
    - many “hard+soft” trials
    - Search differences

Unsafety level	failure rate
2 hard + 1 soft	$\sim 50\%$
3 hard + 1 soft	$\sim 15\%$

NB: small issues in the split-merge

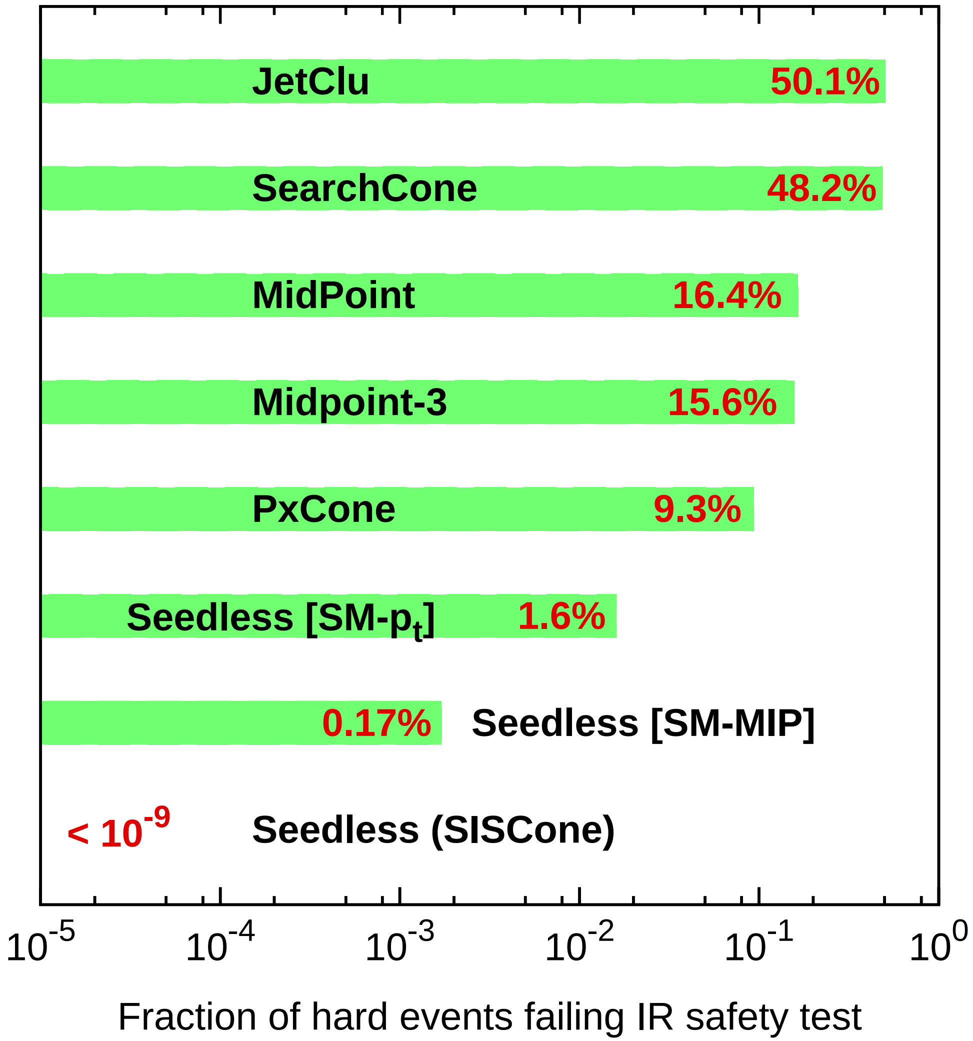


IR Unsafety failure rates

- Hard event: 2-10 particles
- Soft add-on: 1-5 particles
- Run:
  - “hard” only
  - many “hard+soft” trials
  - Search differences

Unsafety level	failure rate
2 hard + 1 soft	$\sim 50\%$
3 hard + 1 soft	$\sim 15\%$
SISCone	IR safe !

NB: small issues in the split-merge



Consequences on observables

Physical impact: SISCone vs. midpoint(s) ?

IR unsafety of midpoint: 3 particles in the same vicinity + 1 to balance p_t
$\Rightarrow$ starts at the $2 \to 4$ level ($\mathcal{O}\left(\alpha_s^4\right)$)

Observable	1st miss cones at	Last meaningful order
Inclusive jet cross section	NNLO	NLO
$W/Z/H$ + 1 jet cross section	NNLO	NLO
3 jet cross section	NLO	LO
$W/Z/H$ + 2 jet cross section	NLO	LO
jet masses in 3 jets	LO	none

Consequences on observables

Physical impact: SISCone vs. midpoint(s) ?

IR unsafety of midpoint: 3 particles in the same vicinity + 1 to balance p_t
$\Rightarrow$ starts at the $2 \to 4$ level ($\mathcal{O}\left(\alpha_s^4\right)$)

Observable	1st miss cones at	Last meaningful order
Inclusive jet cross section	NNLO	NLO
$W/Z/H$ + 1 jet cross section	NNLO	NLO
3 jet cross section	NLO	LO (NLO in NLOJet)
$W/Z/H$ + 2 jet cross section	NLO	LO (NLO in MCFM)
jet masses in 3 jets	LO	none (LO in NLOJet)

The IR-unsafety issue will matter at LHC
+ We do not want the theoretical efforts to be wasted

SISCone vs. other cone algorithms

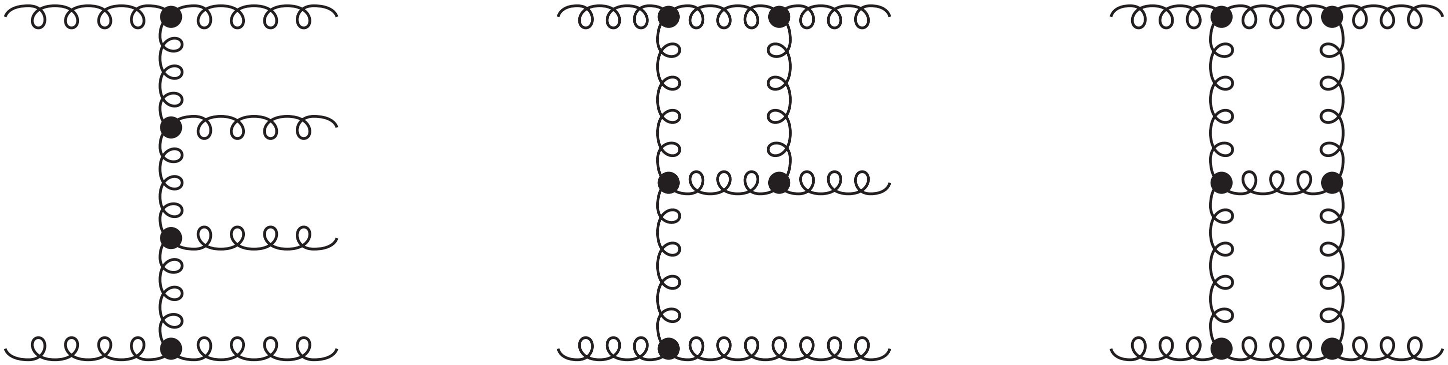
implications of a seedless cone

Inclusive jet spectrum: perturbative exp

SISCone vs. midpoint(s) in inclusive jet spectrum?

- IR unsafety of midpoint: 3 particles in the same vicinity + 1 to balance p_t
  $\Rightarrow$ starts at the $2 \to 4$ level ($\mathcal{O}\left(\alpha_s^4\right)$)

- 3 contributions at this order:
  $2 \to 4$ at LO (tree), $2 \to 3$ at NLO (1 loop) and $2 \to 2$ at NNLO (2 loops)



Inclusive jet spectrum: perturbative exp

SISCone vs. midpoint(s) in inclusive jet spectrum?

- IR unsafety of midpoint: 3 particles in the same vicinity + 1 to balance p_t
  $\Rightarrow$ starts at the $2 \to 4$ level ($\mathcal{O}\left(\alpha_s^4\right)$)

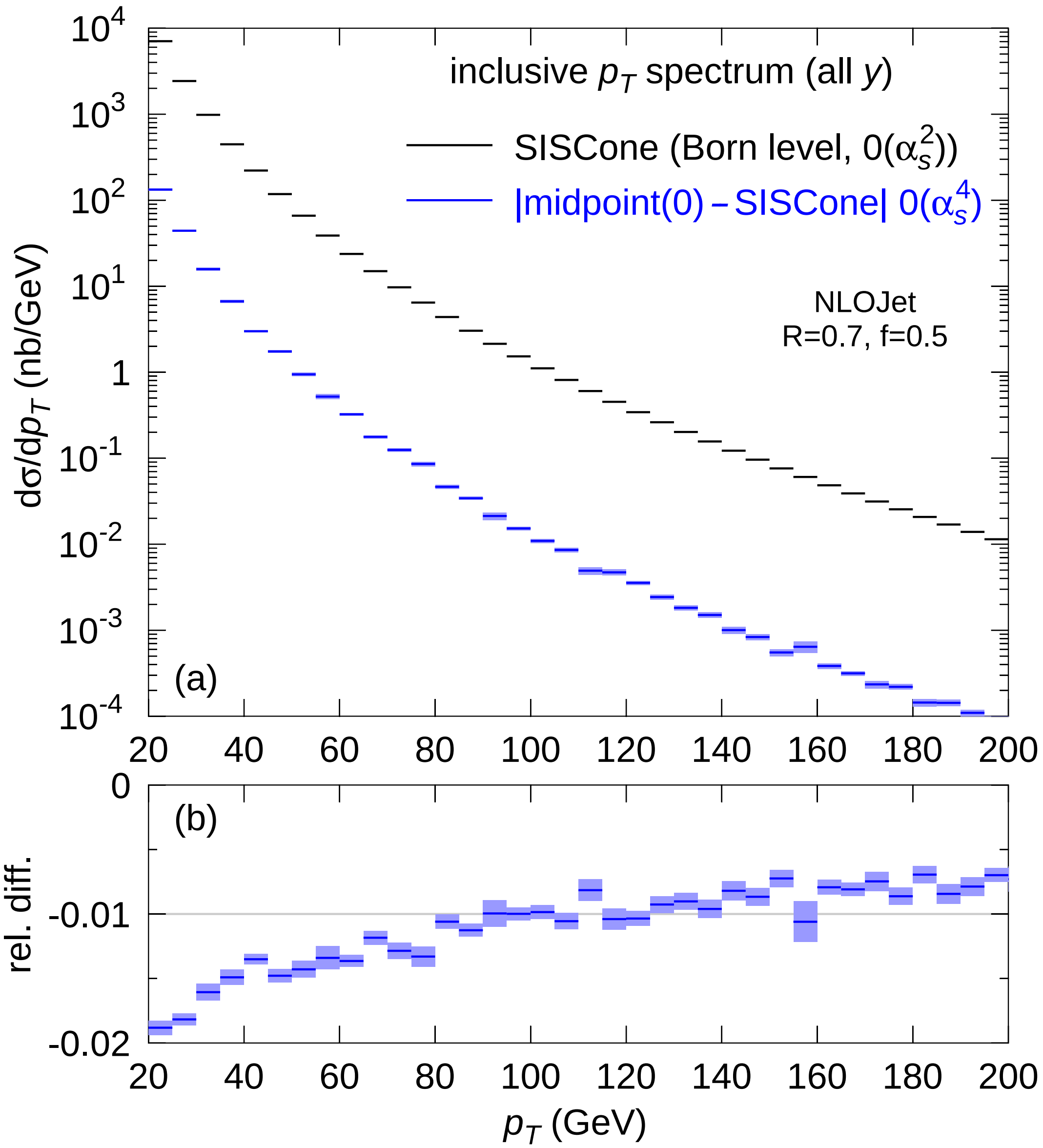
- 3 contributions at this order:
  $2 \to 4$ at LO (tree), $2 \to 3$ at NLO (1 loop) and $2 \to 2$ at NNLO (2 loops)

- $2 \to 4$ at LO is IR divergent
  BUT the difference between SISCone and midpoint(s) in finite since it is 0 at the $2 \to 2$ and $2 \to 3$ levels

$\Rightarrow$ compute $|$SISCone-midpoint$(s)|$ for $2 \to 4$ diagrams

- Compare with the $2 \to 2$ (LO) spectrum to estimate effect

Inclusive jet spectrum: perturbative exp

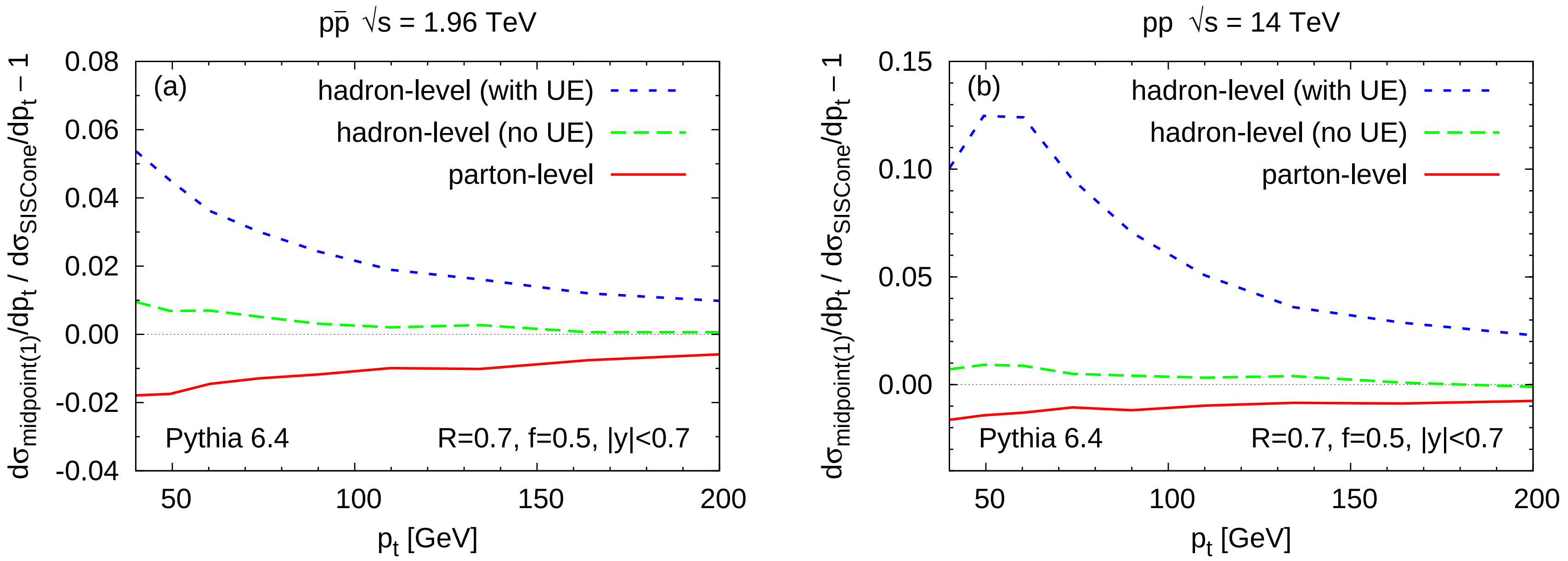


Differences of order 1-2 %

Inclusive jet spectrum: hadron level

Including parton shower, hadronic corrections and/or underlying event:

Ratio midpoint/SISCone-1:

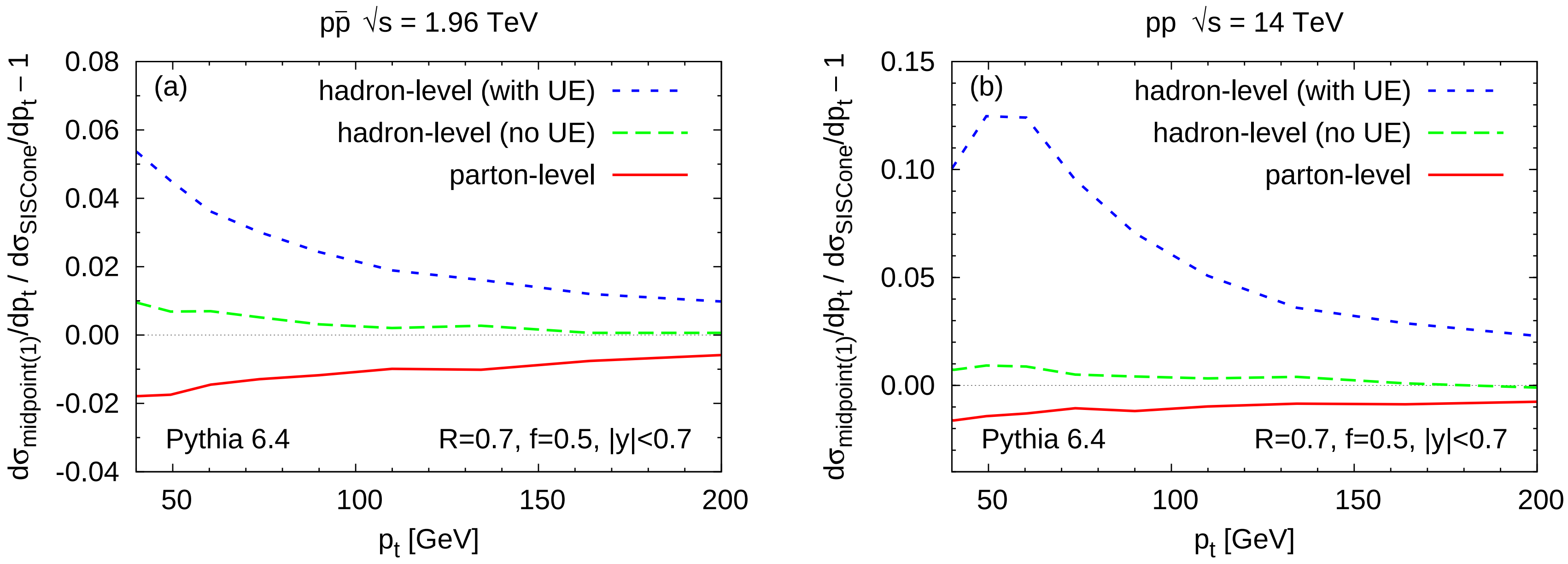


- Differences up to 5% (with a change of sign)
- Raise up to 10% at LHC energy!

Inclusive jet spectrum: hadron level

Including parton shower, hadronic corrections and/or underlying event:

Ratio midpoint/SISCone-1:



- Differences up to 5% (with a change of sign)
- Raise up to 10% at LHC energy!
- Less effect from underlying event in SISCone (i.e. better agreement with parton level)

Jet mass spectrum

Inclusive jet spectrum

$\longrightarrow$ effect at NNLO i.e. $\mathcal{O}\left(\alpha_s^2\right)$ w.r.t. LO

$\Rightarrow$ want to look at more exclusive processes

Example: mass spectrum in 3-jet events (or W/Z/H+2j)

$$\left.\begin{array}{l} 2 \to 2 \text{ has only 2 jets} \\ 2 \to 3 \text{ has zero masses} \end{array}\right\} \Rightarrow \text{ first contribution from } 2 \to 4$$

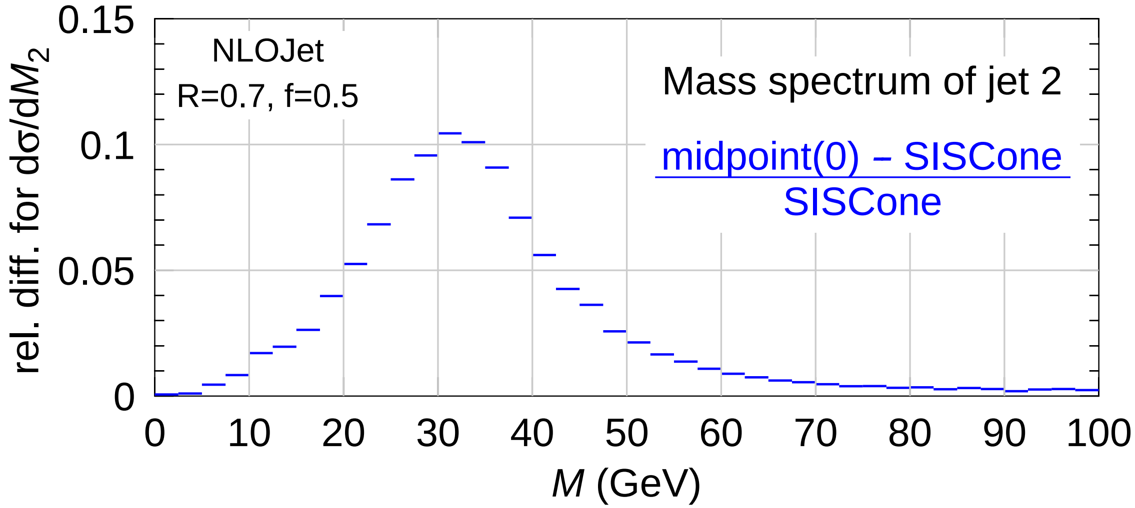
$\Rightarrow$ Expect modifications at LO!

Ratio $\frac{\text{midpoint}-\text{SISCone}}{\text{SISCone}}$ for masses spectra in 3-jet events

cuts: $p_{t,1} \geq 120$ GeV, $p_{t,2} \geq 80$ GeV, $p_{t,3} \geq 40$ GeV

Jet mass spectrum: perturbative level

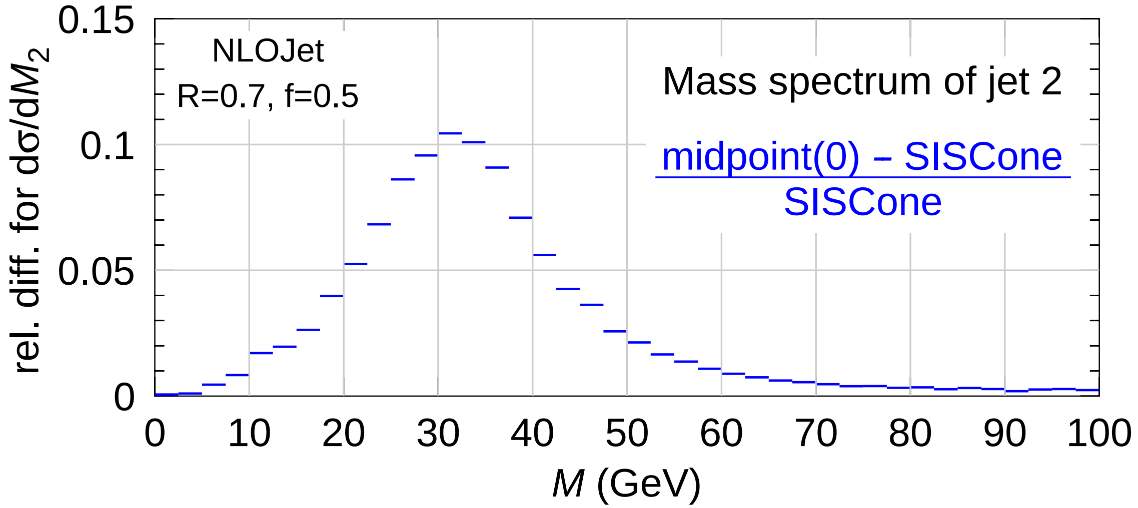
1. Fixed order computation (NLOJet, LO, $2 \to 4$)



Differences up to 10 %

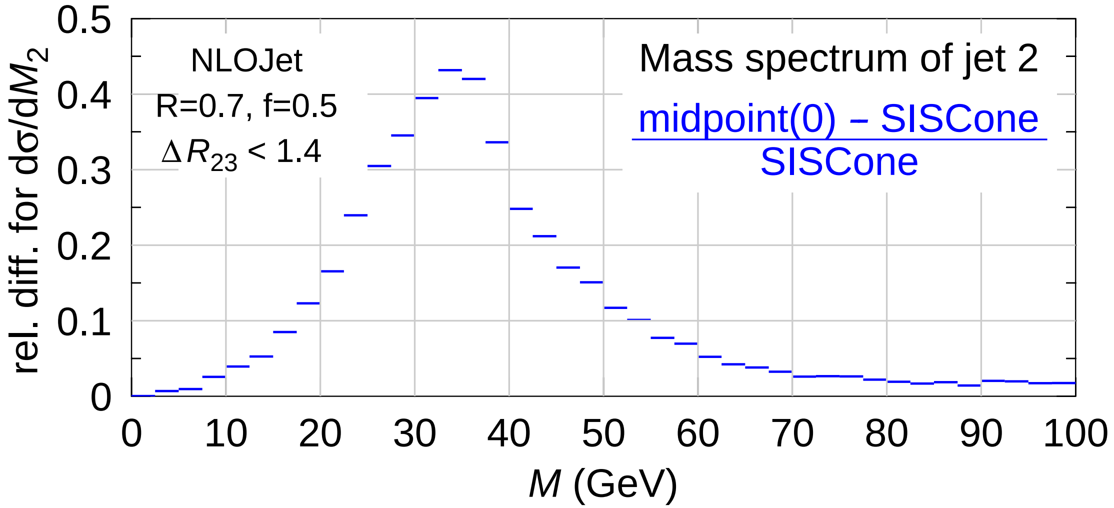
Jet mass spectrum: perturbative level

1. Fixed order computation (NLOJet, LO, $2 \to 4$)



Differences up to 10 %

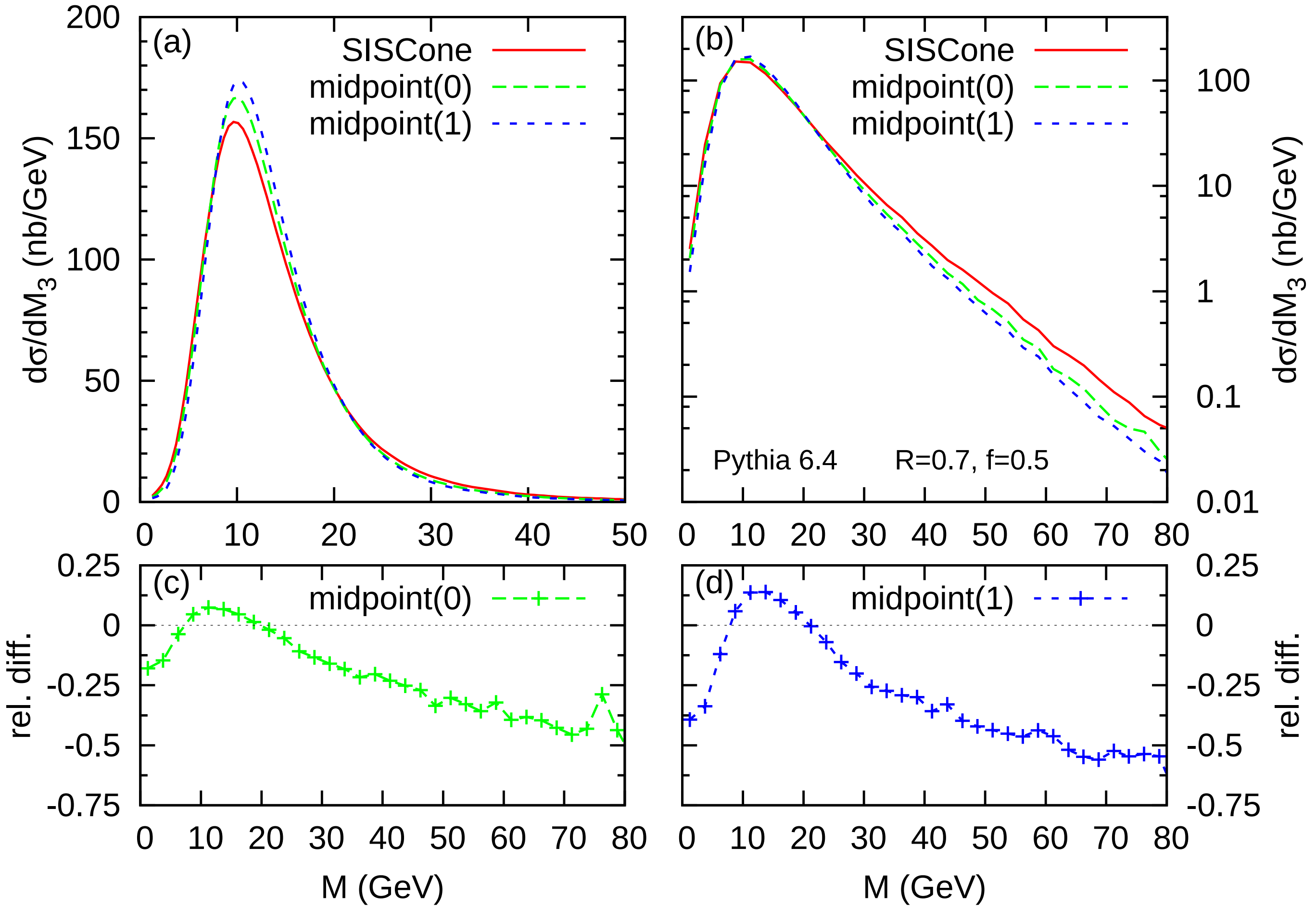
2. Also require jets 2 and 3 within distance $\leq 2R$



Differences up to 40 %

Impact on jet mass spectrum

3. At hadron level (PYTHIA)



- ▷ Differences of order 10 %
- ▷ Larger effects in the tail
- ▷ seed threshold even worse

SISCone vs. recombination-type algorithms

[Les Houches, jet benchmark channels,

M. Cacciarin, J. Rojo, G. Salam, G.S., in preparation]

Les Houches jet benchmarks

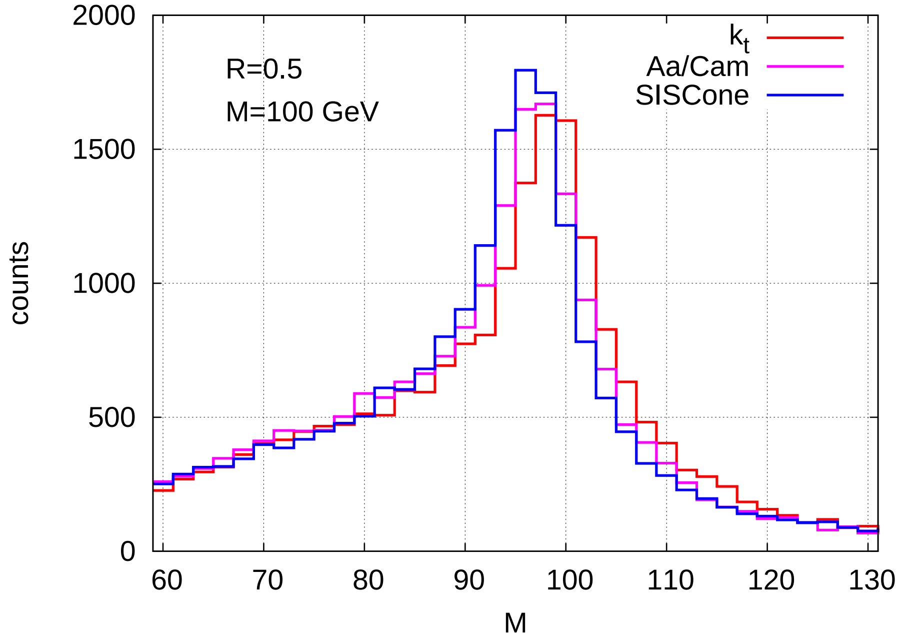
Idea: compare the various algorithms for typical reconstructions, e.g.

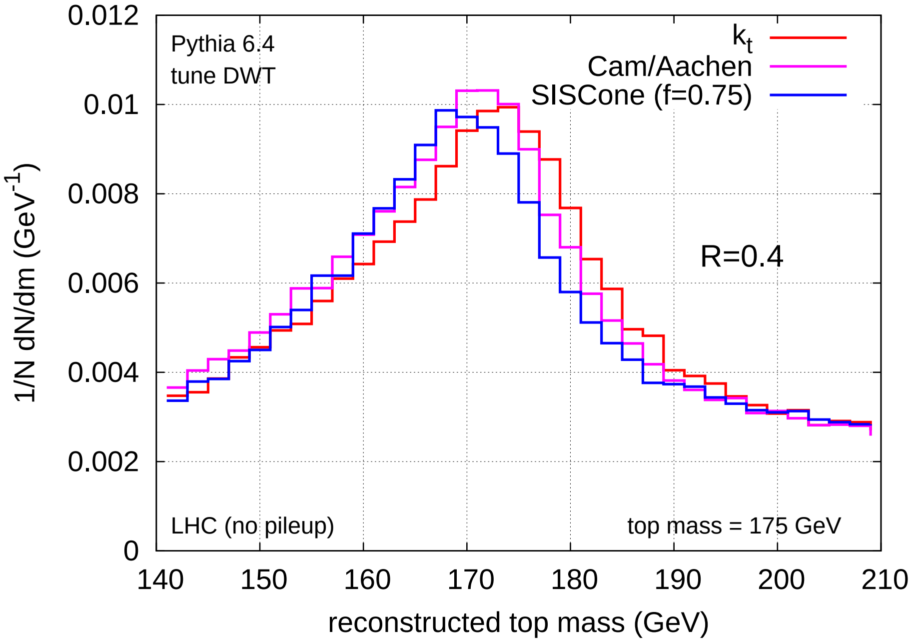
- $Z' \to q\bar{q} \to 2$ jets ($m_{Z'}$ from 100 GeV to 4 TeV)
- $t\bar{t} \to 6$ jets (via $t \to bW^+ \to bq\bar{q}$ and $\bar{t} \to \bar{b}W^- \to \bar{b}q\bar{q}$)

Les Houches jet benchmarks

Idea: compare the various algorithms for typical reconstructions, e.g.

- $Z' \to q\bar{q} \to 2$ jets ($m_{Z'}$ from 100 GeV to 4 TeV)
- $t\bar{t} \to 6$ jets (via $t \to bW^+ \to bq\bar{q}$ and $\bar{t} \to \bar{b}W^- \to \bar{b}q\bar{q}$)



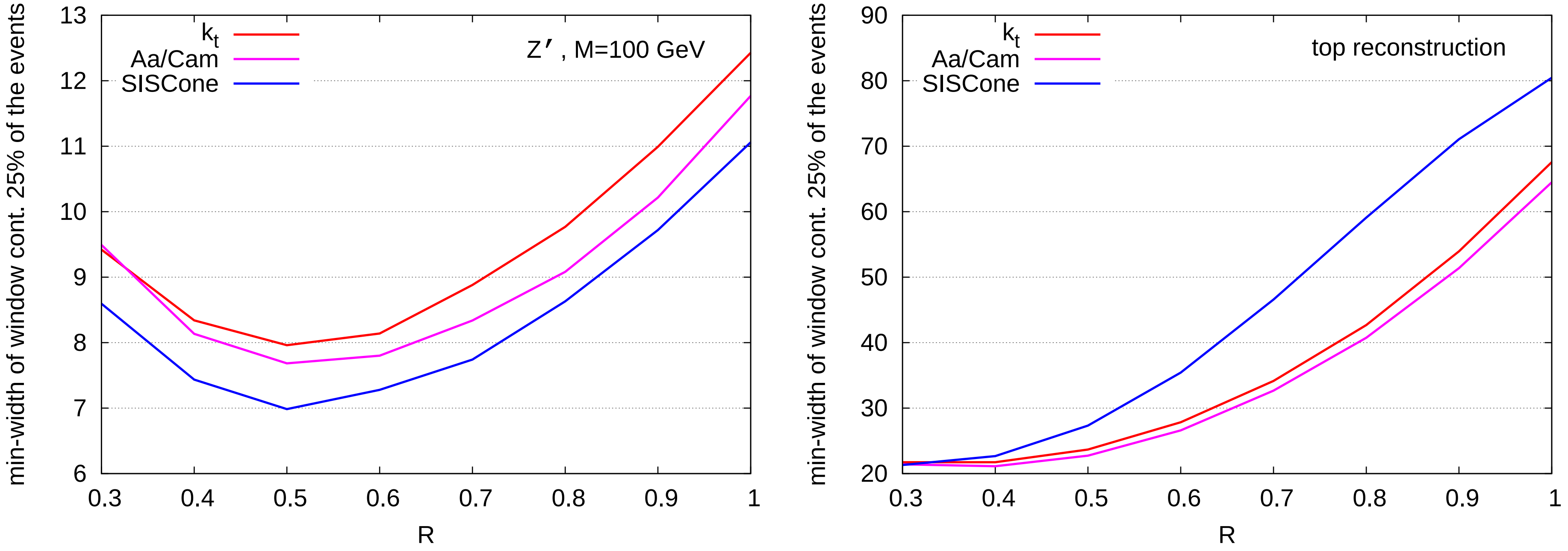


Les Houches jet benchmarks

Idea: compare the various algorithms for typical reconstructions, e.g.

- $Z' \to q\bar{q} \to 2$ jets ($m_{Z'}$ from 100 GeV to 4 TeV)
- $t\bar{t} \to 6$ jets (via $t \to bW^+ \to bq\bar{q}$ and $\bar{t} \to \bar{b}W^- \to \bar{b}q\bar{q}$)

Quality measure: max. width of a window containing 25% of the events



- Z' slightly favours SISCone, $t\bar{t}$ slightly favours k_t/Cam
- The R dependence gives more variations!

SISCone conclusions

- Jets are present everywhere: k_t and cone are widely used
- seeded implementations are IR unsafe (sometimes collinear unsafe)
  IR safety is a prerequisite for perturbative QCD to make sense

We propose a new cone algorithm (SISCone):

- IR safe (and collinear safe)
- as fast as available cone implementations
- has 10% impact on jet mass spectra (can be up to 40%)
- is less affected by underlying events

Jet area

Everyone has an idea of what a jet area is

but can we define that properly?

[M. Cacciari, G. Salam, G.S., arXiv:08021188]

[M. Cacciari, G. Salam, PLB659 (08) 119]

Area definition

- Idea: add infinitely soft particle (ghosts)
  - with IR-safe algorithms such as k_t, Aachen/Cambridge and SISCone, clustering is unchanged
  - look in which jets added particles are catched

Area definition

- Idea: add infinitely soft particle (ghosts)
  - with IR-safe algorithms such as k_t, Aachen/Cambridge and SISCone, clustering is unchanged
  - look in which jets added particles are catched

- Passive area

  add one ghost and look where it ends. repeat to cover the (y, ϕ) plane

Area definition

- Idea: add infinitely soft particle (ghosts)
  - with IR-safe algorithms such as k_t, Aachen/Cambridge and SISCone, clustering is unchanged
  - look in which jets added particles are catched

- Passive area

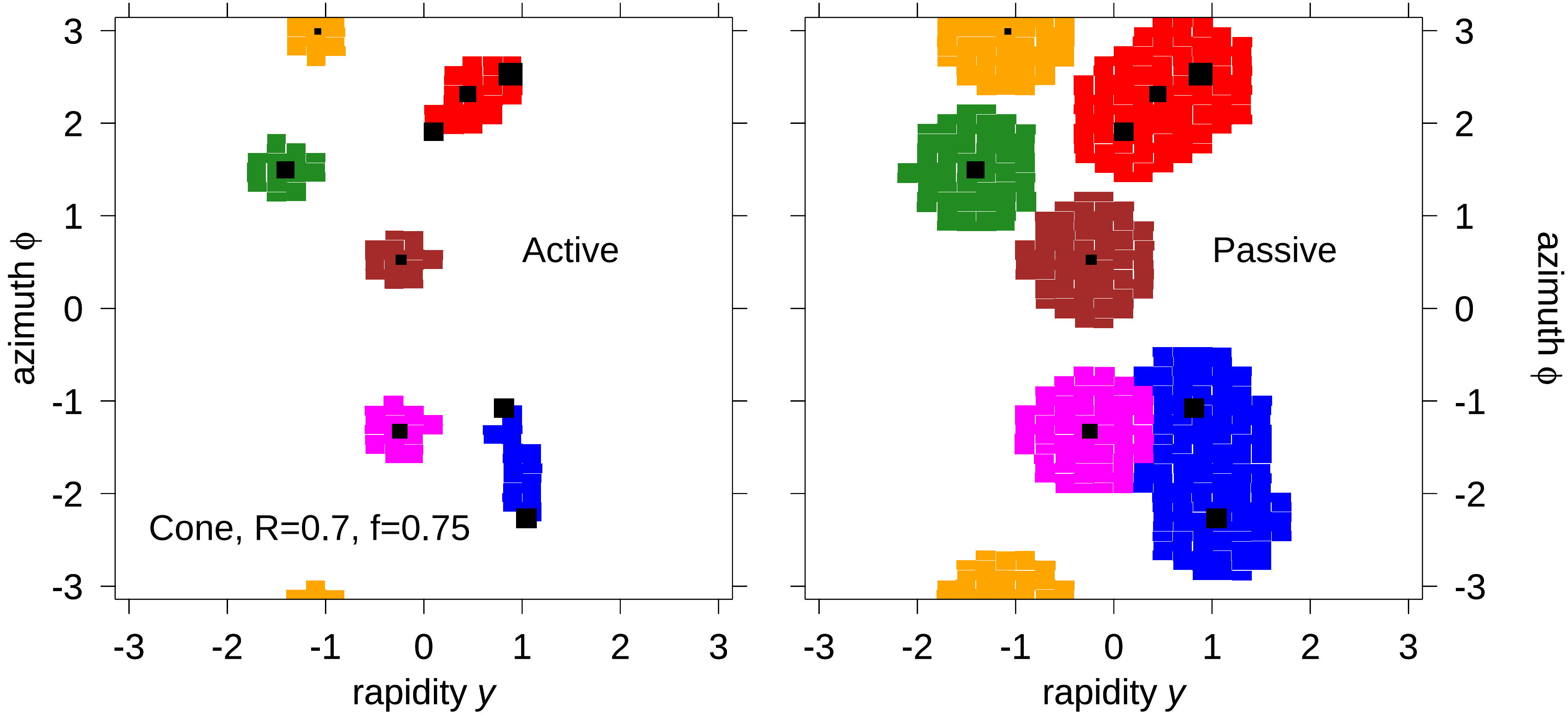
  add one ghost and look where it ends. repeat to cover the (y, ϕ) plane

- Active area

  add a large amount of ghosts and cluster everything
  - also gives purely ghosted jets
  - ghost background $\simeq$ pileup background

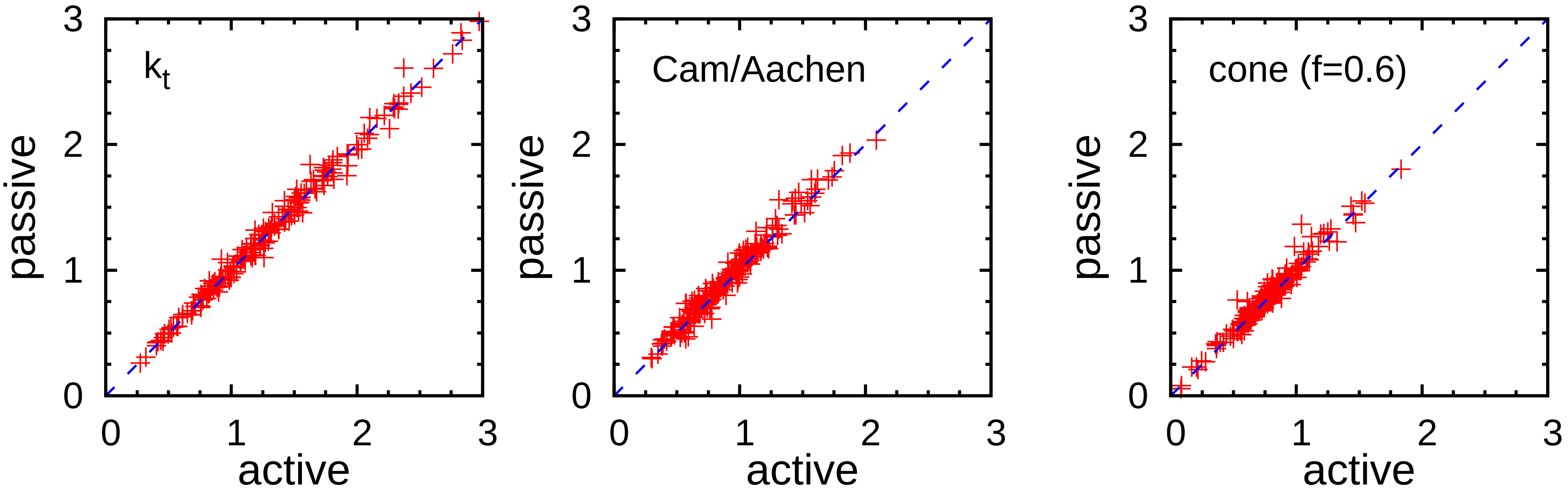
Area definition

- Small N: active area is usually smaller than passive area (especially for the cone)

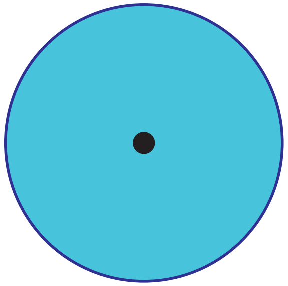
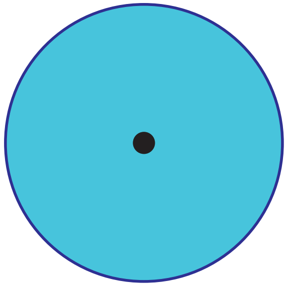
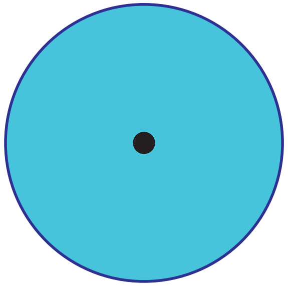


Area definition

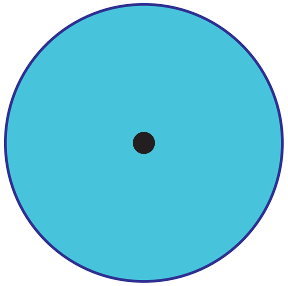
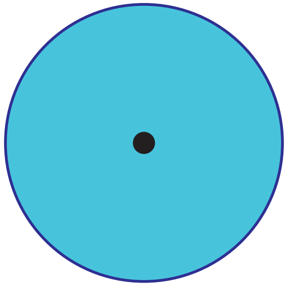
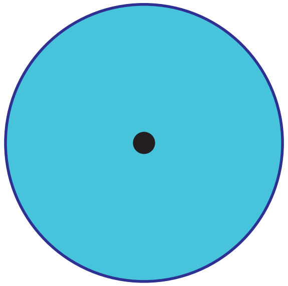
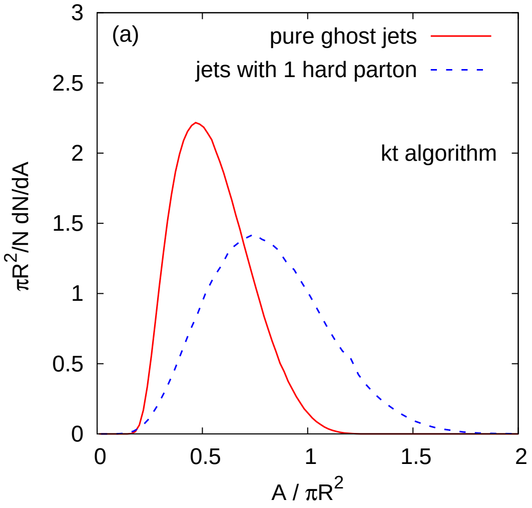
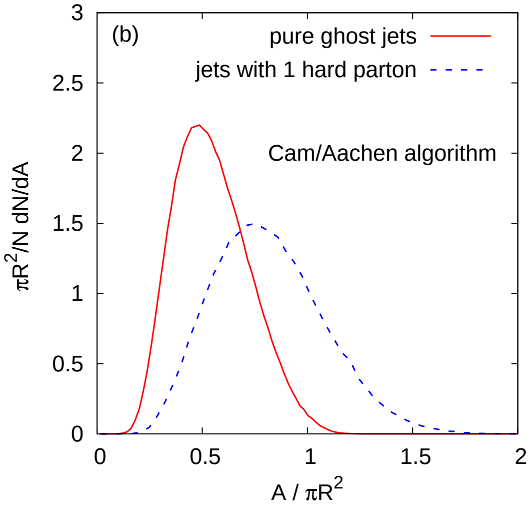
- Small N: active area is usually smaller than passive area (especially for the cone)
- For more dense events (e.g. Pythia with underlying event) they tend to be the same



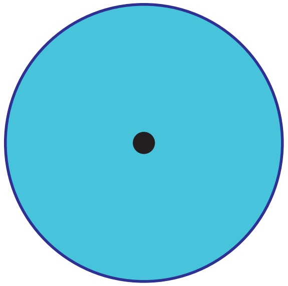
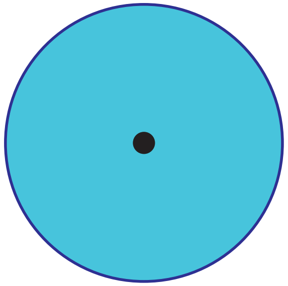
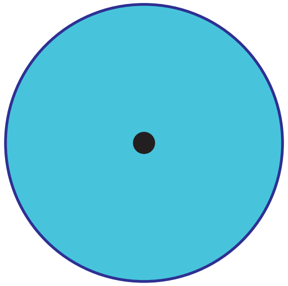
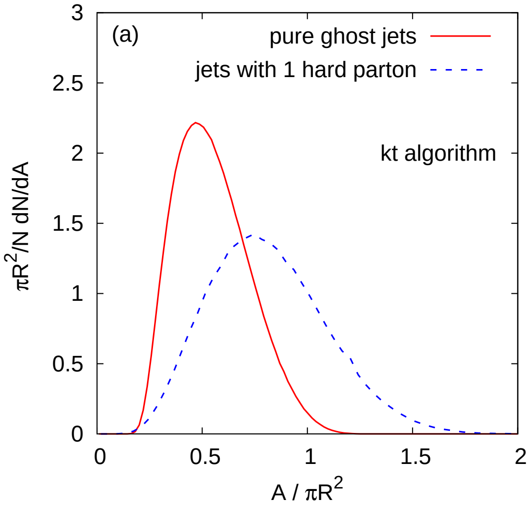
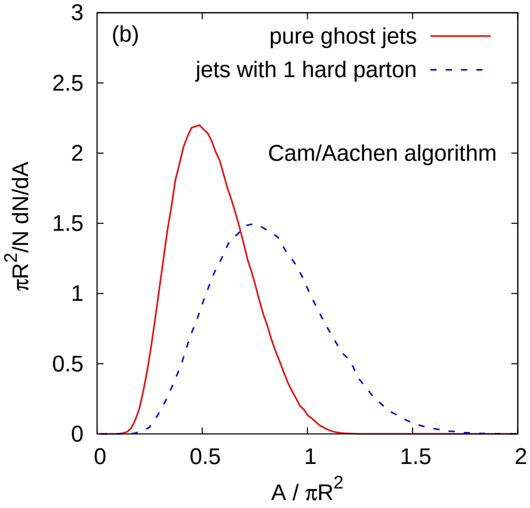
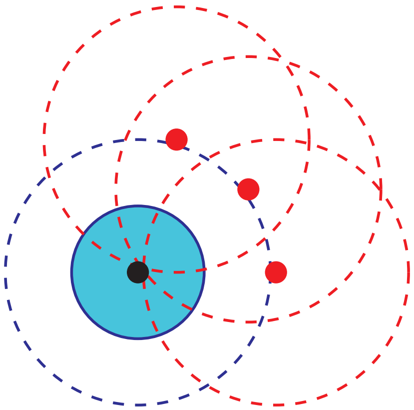
Properties: 1-particle cases

	k_t	Aac/Cam	cone
Passive	 πR^2	 πR^2	 πR^2
Active			

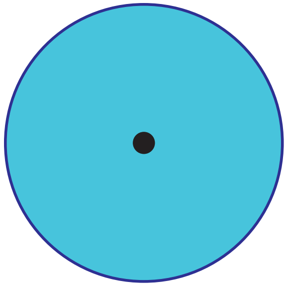
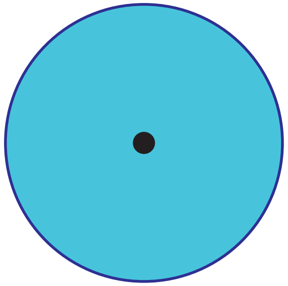
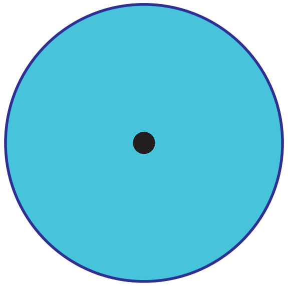
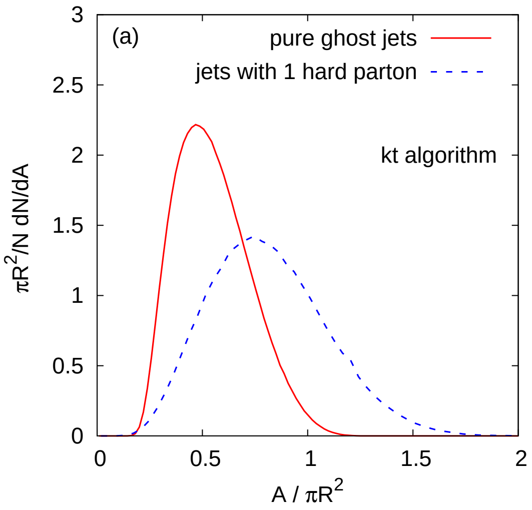
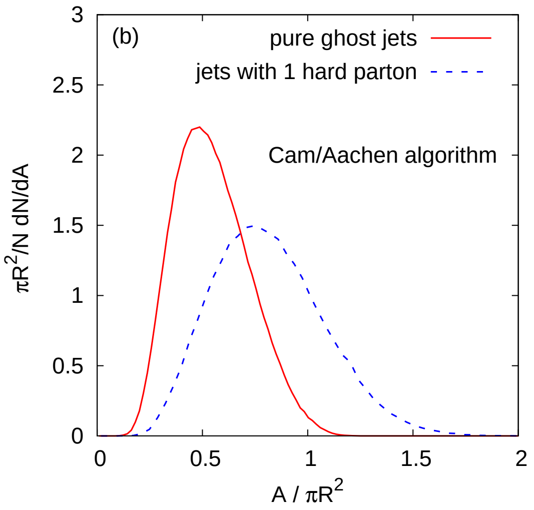
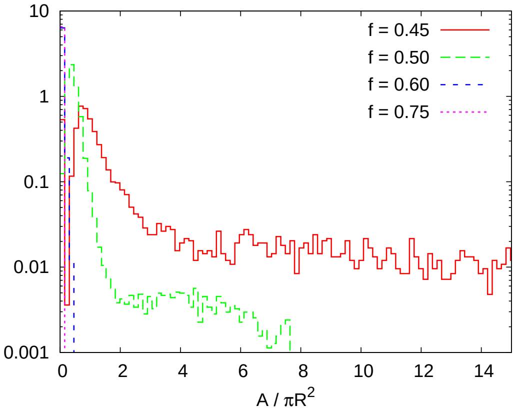
Properties: 1-particle cases

	k_t	Aac/Cam	cone
Passive	 πR^2	 πR^2	 πR^2
Active	 $\frac{A_{\text{hard}}}{\pi R^2} \approx 0.812 \pm 0.277$ $\frac{A_{\text{ghost}}}{\pi R^2} \approx 0.554 \pm 0.174$	 $\frac{A_{\text{hard}}}{\pi R^2} \approx 0.814 \pm 0.261$ $\frac{A_{\text{ghost}}}{\pi R^2} \approx 0.551 \pm 0.176$	

Properties: 1-particle cases

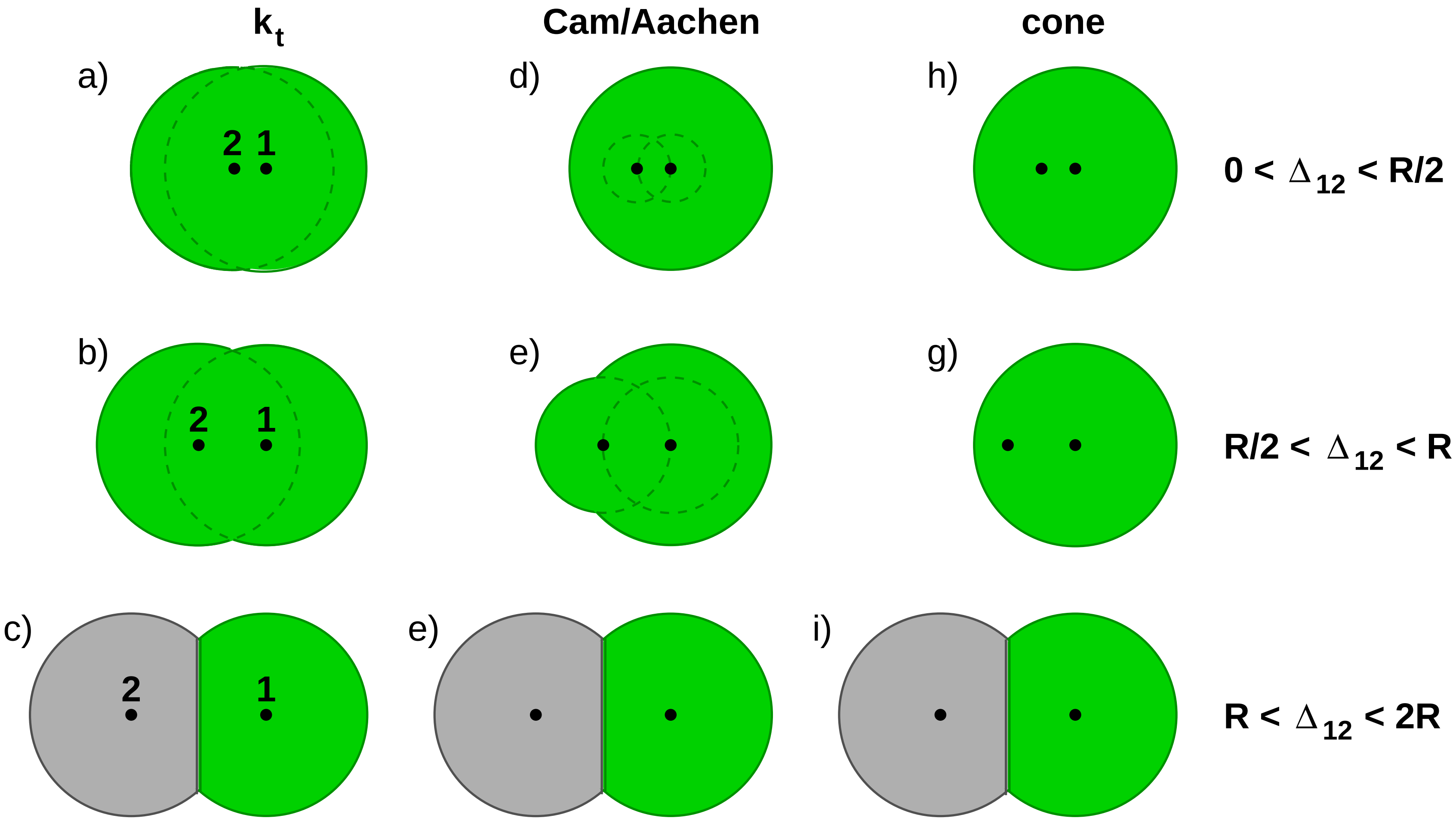
	k_t	Aac/Cam	cone
Passive	 πR^2	 πR^2	 πR^2
Active	 $\frac{A_{\text{hard}}}{\pi R^2} \approx 0.812 \pm 0.277$ $\frac{A_{\text{ghost}}}{\pi R^2} \approx 0.554 \pm 0.174$	 $\frac{A_{\text{hard}}}{\pi R^2} \approx 0.814 \pm 0.261$ $\frac{A_{\text{ghost}}}{\pi R^2} \approx 0.551 \pm 0.176$	 $\frac{A_{\text{hard}}}{\pi R^2} = 0.25$

Properties: 1-particle cases

	k_t	Aac/Cam	cone
Passive	 πR^2	 πR^2	 πR^2
Active	 $\frac{\mathcal{A}_{\text{hard}}}{\pi R^2} \approx 0.812 \pm 0.277$ $\frac{\mathcal{A}_{\text{ghost}}}{\pi R^2} \approx 0.554 \pm 0.174$	 $\frac{\mathcal{A}_{\text{hard}}}{\pi R^2} \approx 0.814 \pm 0.261$ $\frac{\mathcal{A}_{\text{ghost}}}{\pi R^2} \approx 0.551 \pm 0.176$	 $\frac{\mathcal{A}_{\text{hard}}}{\pi R^2} = 0.25$ $\mathcal{A}_{\text{ghost}}$ depends on f possible monster jets!

Properties: 2-particle cases

Passive area: 1 hard particle + 1 soft ($p_{t1} \gg p_{t2}$)



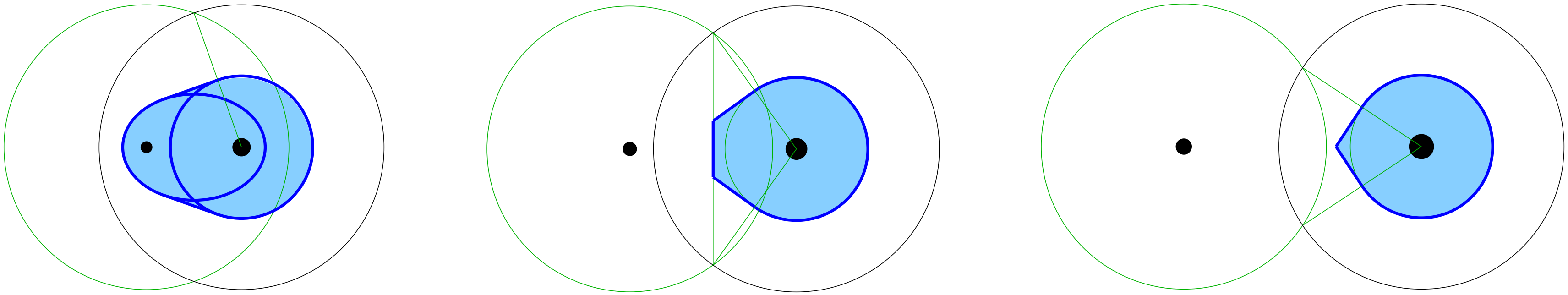
$0 < \Delta_{12} < R/2$

$R/2 < \Delta_{12} < R$

$R < \Delta_{12} < 2R$

2-particle cases

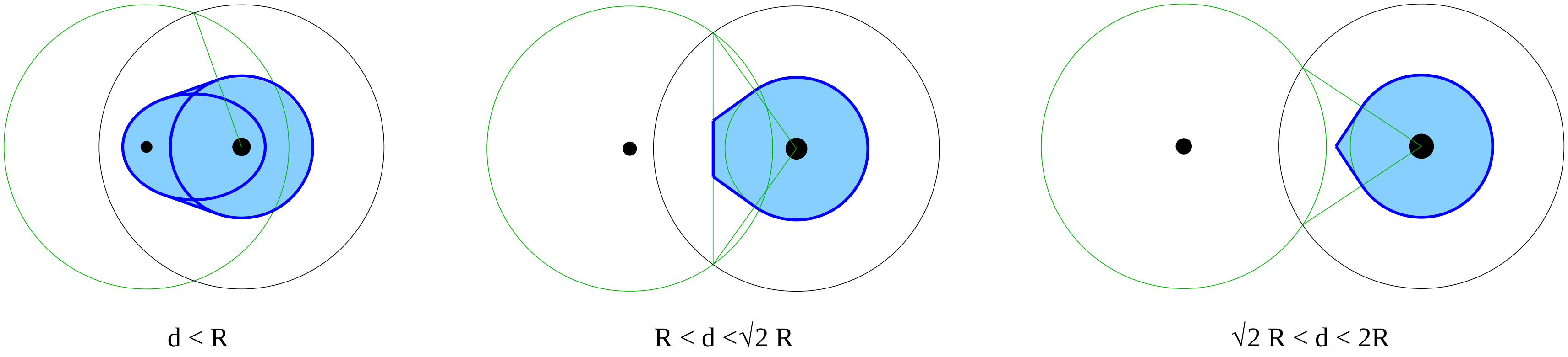
Active area: 1 hard particle + 1 soft: analytic result for cone only



$d < R$ $\quad$ $R < d < \sqrt{2} R$ $\quad$ $\sqrt{2} R < d < 2R$

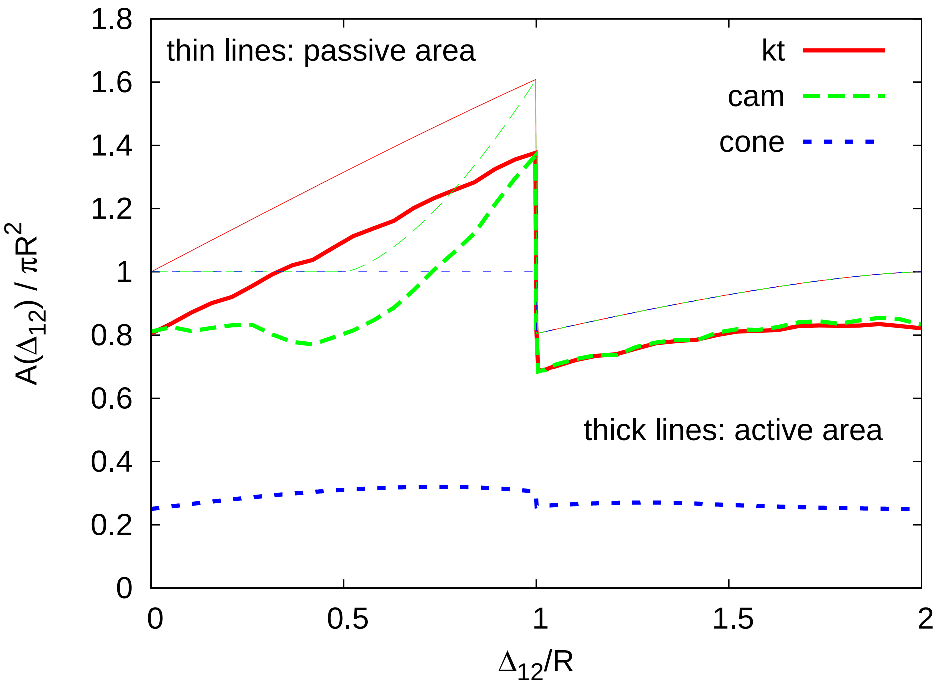
2-particle cases

Active area: 1 hard particle + 1 soft: analytic result for cone only



Alltogether, we have:

- Area $\neq$ cst. πR^2
- Δ_{12} dependence under control



Area scaling violations

QCD probability of emitting a small-angle soft gluon:

$$\frac{dP}{d\Delta_{12} dp_{t,2}} = C_{F,A} \frac{2\alpha_s}{\pi} \frac{1}{\Delta_{12}} \frac{1}{p_{t,2}}$$

Hence the average area is

$$\langle \mathcal{A}(p_{t,1}, R) \rangle \quad = \quad \mathcal{A}_{\text{1hard}}(R) + \int d\Delta \, dp_{t,2} \, \frac{dP}{d\Delta_{12} dp_{t,2}} \left[\mathcal{A}_{\text{hard+1 soft}}(\Delta, R) - \pi R^2 \right]$$

Area scaling violations

QCD probability of emitting a small-angle soft gluon:

$$\frac{dP}{d\Delta_{12} dp_{t,2}} = C_{F,A} \frac{2\alpha_s}{\pi} \frac{1}{\Delta_{12}} \frac{1}{p_{t,2}}$$

Hence the average area is

$$\begin{aligned} \langle \mathcal{A}(p_{t,1}, R) \rangle &= \mathcal{A}_{\text{1hard}}(R) + \int d\Delta \, dp_{t,2} \, \frac{dP}{d\Delta_{12} dp_{t,2}} \left[\mathcal{A}_{\text{hard+1 soft}}(\Delta, R) - \pi R^2\right] \\ &= \frac{C_{F,A}}{\pi b_0} \log\left(\frac{\alpha_s(\Lambda)}{\alpha_s(R p_t)}\right) \pi R^2 \, d \end{aligned}$$

- Scaling violation

Area scaling violations

QCD probability of emitting a small-angle soft gluon:

$$\frac{dP}{d\Delta_{12} dp_{t,2}} = C_{F,A} \frac{2\alpha_s}{\pi} \frac{1}{\Delta_{12}} \frac{1}{p_{t,2}}$$

Hence the average area is

$$\begin{aligned} \langle \mathcal{A}(p_{t,1}, R) \rangle &= \mathcal{A}_{\text{1hard}}(R) + \int d\Delta \, dp_{t,2} \, \frac{dP}{d\Delta_{12} dp_{t,2}} \left[\mathcal{A}_{\text{hard+1 soft}}(\Delta, R) - \pi R^2 \right] \\ &= \frac{C_{F,A}}{\pi b_0} \log\left(\frac{\alpha_s(\Lambda)}{\alpha_s(R p_t)} \right) \pi R^2 \, d \end{aligned}$$

- Scaling violation
- gluon > quark

Area scaling violations

QCD probability of emitting a small-angle soft gluon:

$$\frac{dP}{d\Delta_{12}dp_{t,2}} = C_{F,A}\frac{2\alpha_s}{\pi}\frac{1}{\Delta_{12}}\frac{1}{p_{t,2}}$$

Hence the average area is

$$\begin{aligned}\langle\mathcal{A}(p_{t,1},R)\rangle &= \mathcal{A}_{\text{1hard}}(R) + \int d\Delta\, dp_{t,2}\,\frac{dP}{d\Delta_{12}dp_{t,2}}\left[\mathcal{A}_{\text{hard+1 soft}}(\Delta,R) - \pi R^2\right]\\ &= \frac{C_{F,A}}{\pi b_0}\log\left(\frac{\alpha_s(\Lambda)}{\alpha_s(Rp_t)}\right)\pi R^2\, d\end{aligned}$$

- Scaling violation
- gluon > quark
- with know LO anomalous dimension

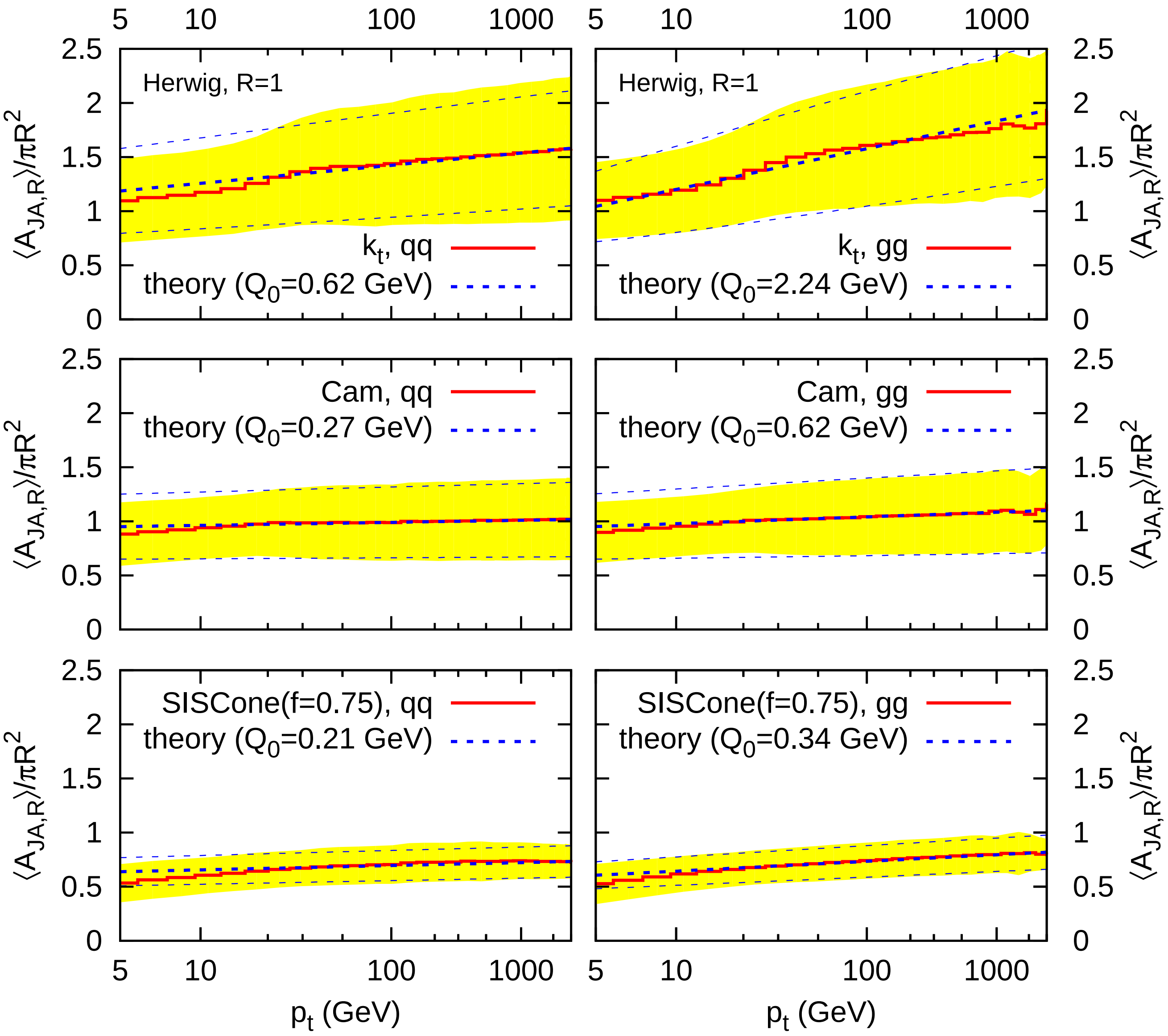
d	passive	active
k_t	0.5638	0.519
Cam	0.07918	0.0865
Cone	-0.06378	0.1246

“Real-life” anomalous dimension

Herwig simulations:

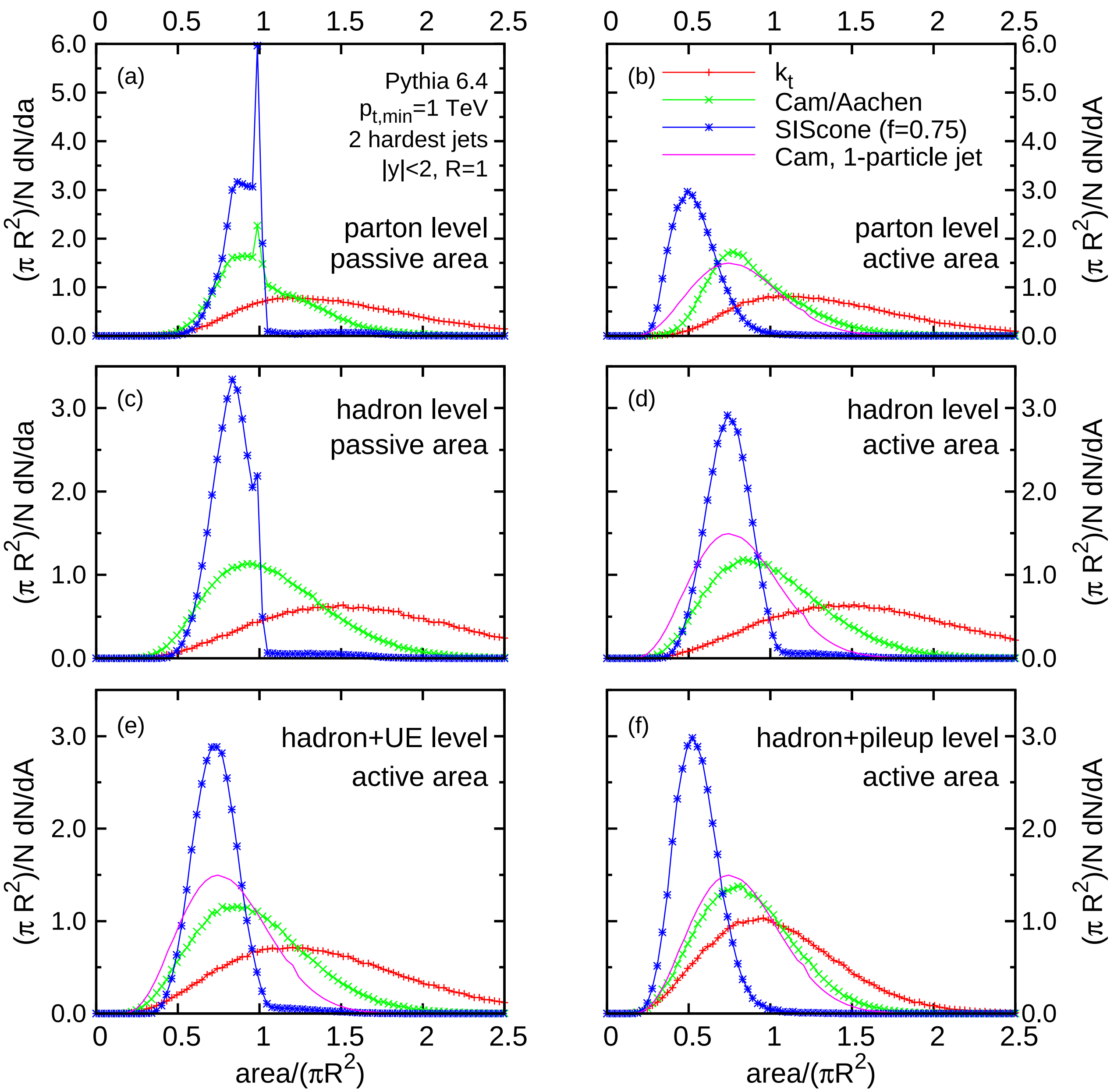
at hadron+UE level:

area vs. p_t of the jet



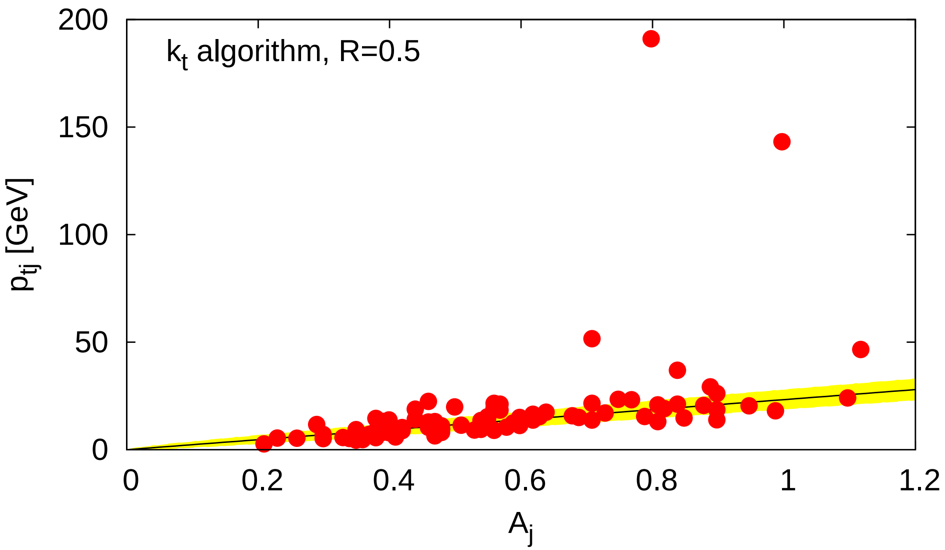
- good agreement with LO predictions
- for flucts. too
- k_t bigger $\Rightarrow$ NLO?

Area histograms



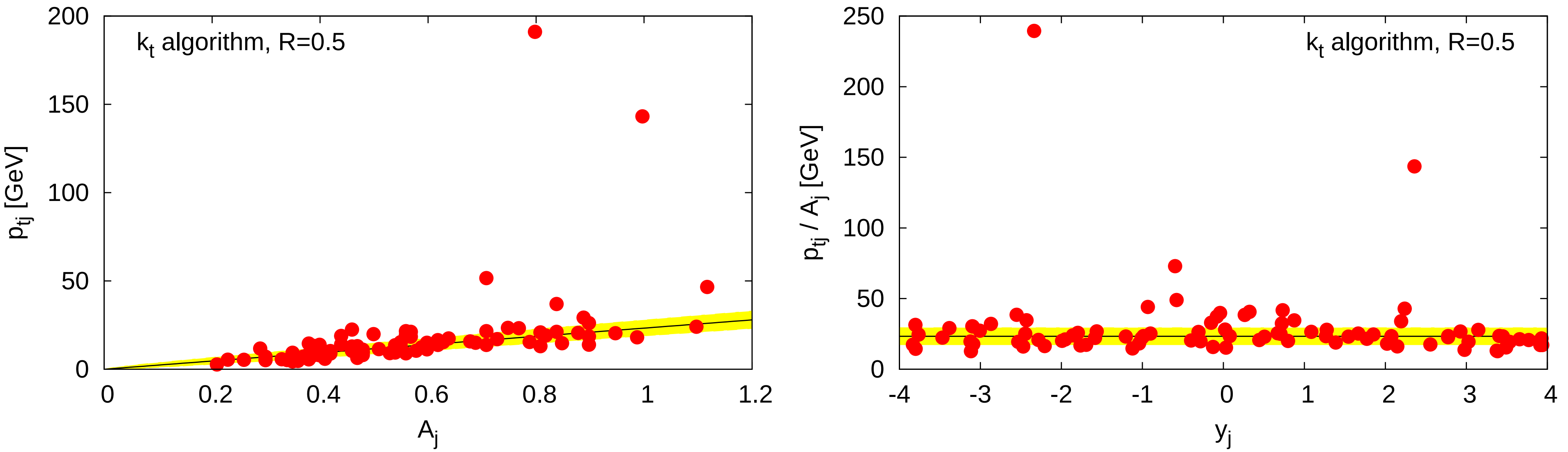
What can area be used for?

Dense event with pile-up:



What can area be used for?

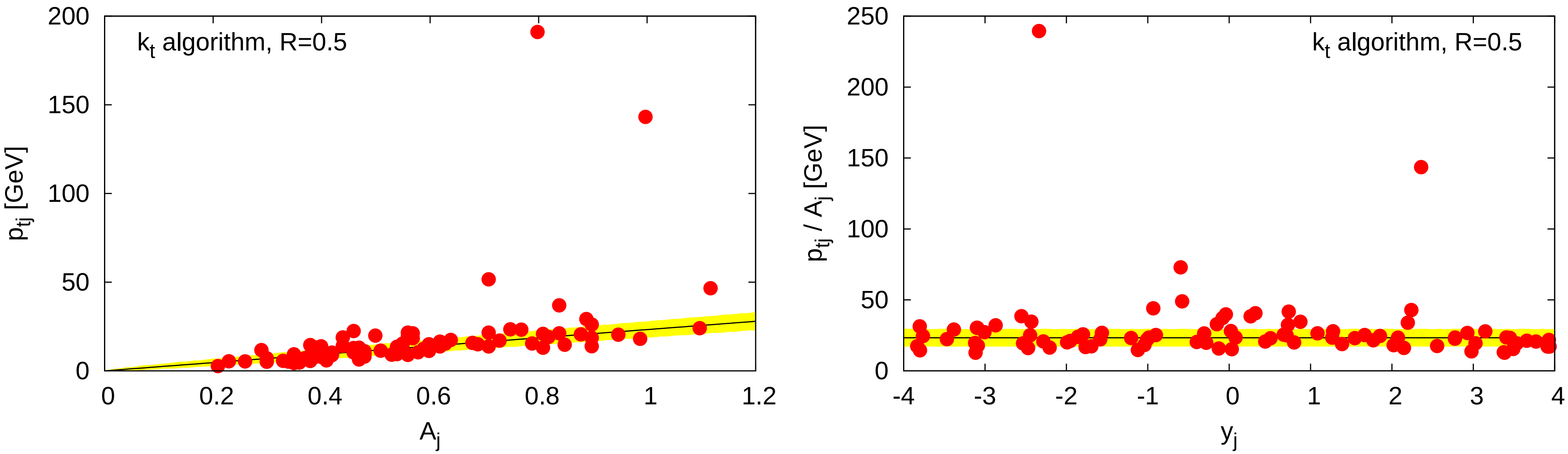
Dense event with pile-up:



- Area $\propto p_t$ of the jet
- p_t/area is constant $\rightarrow$ ρ = median p_t/area

What can area be used for?

Dense event with pile-up:

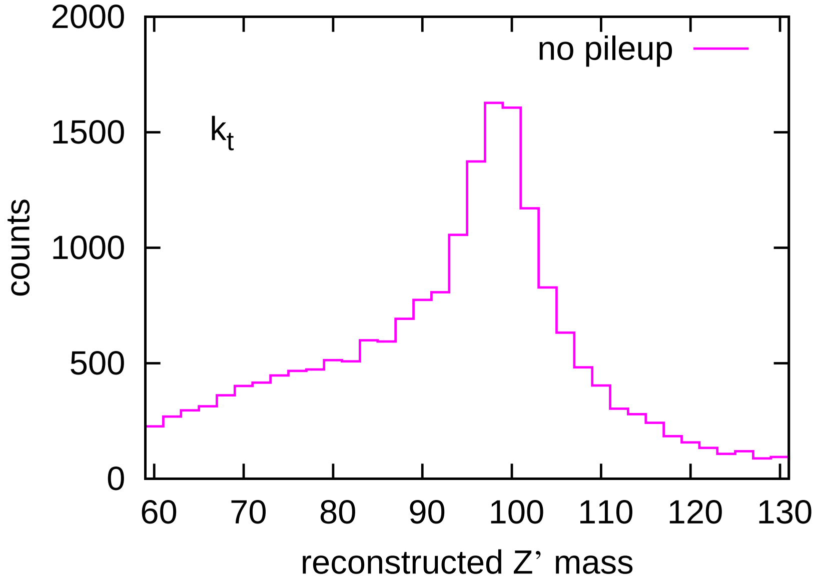


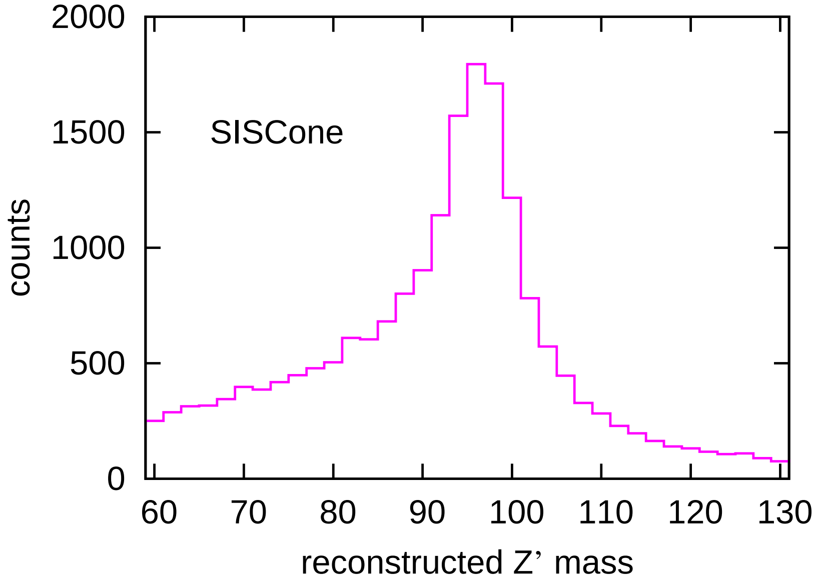
- Area $\propto p_t$ of the jet
- p_t/area is constant $\rightarrow$ ρ = median p_t/area

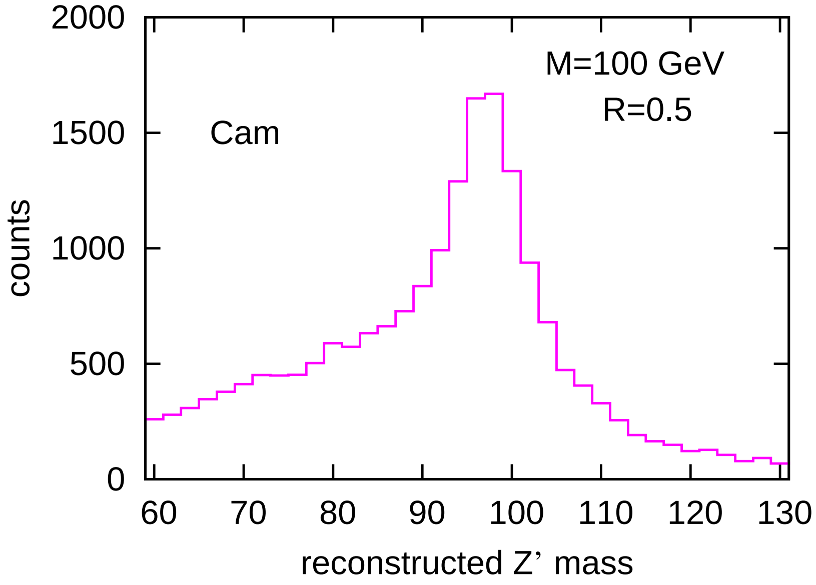
Area can be used to remove pileup pollution
e.g. by removing ρ.area

Subtraction in action

Les Houches jet benchmarks: $Z' \to q\bar{q} \to 2$ jets





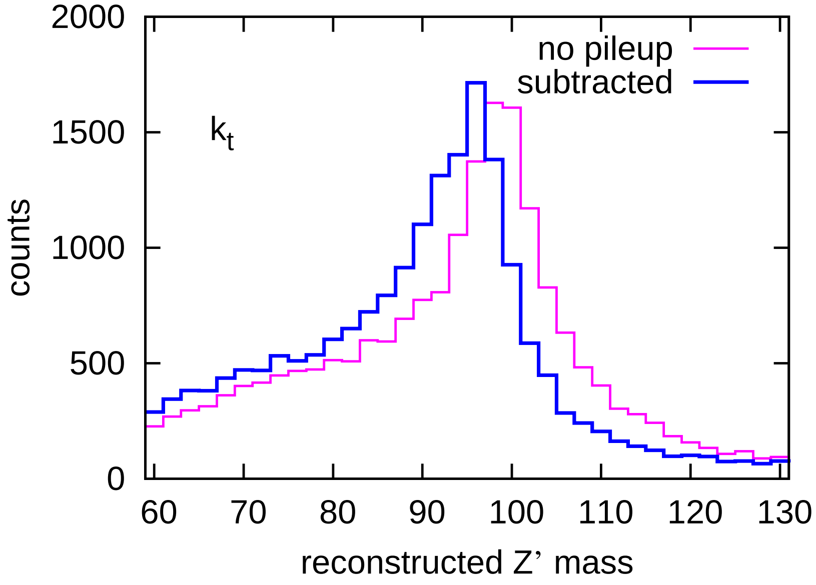


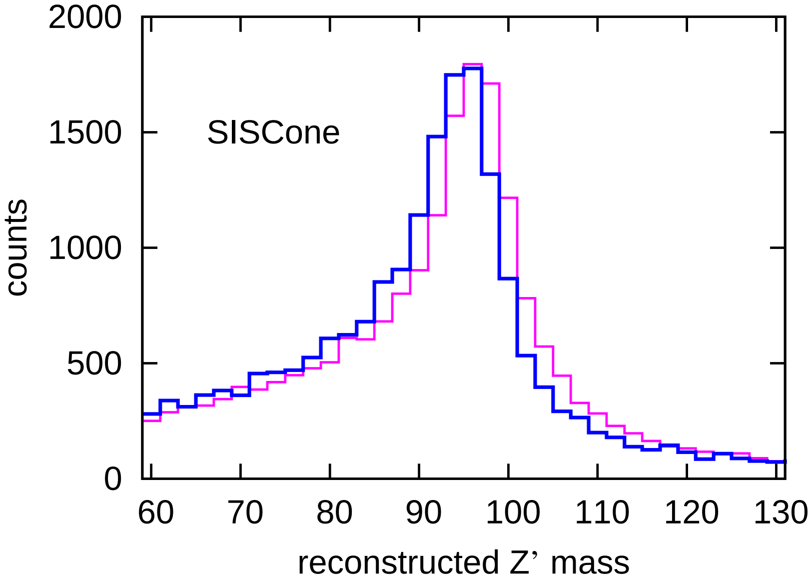
No pileup:

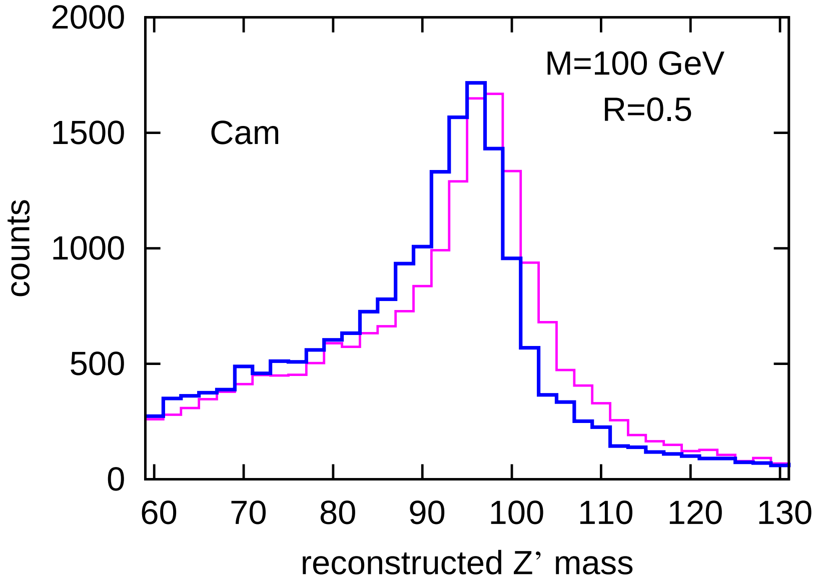
- good result

Subtraction in action

Les Houches jet benchmarks: $Z' \to q\bar{q} \to 2$ jets





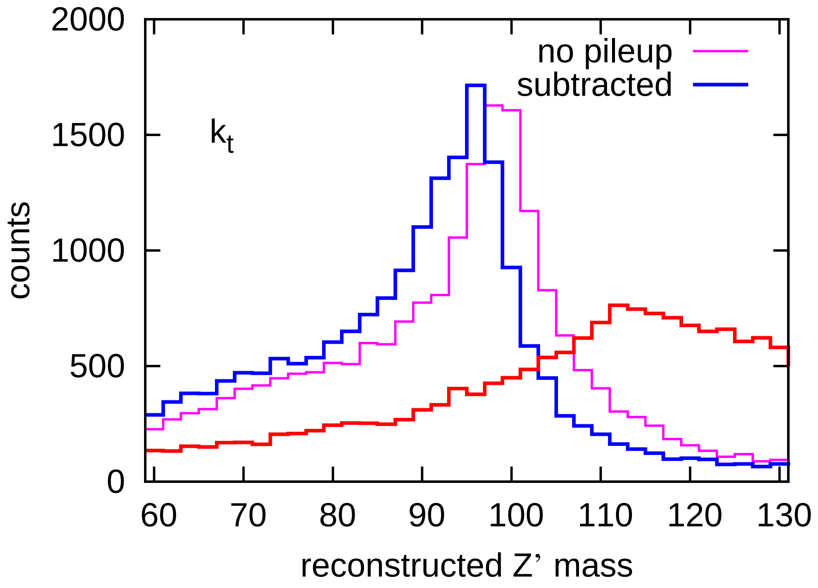


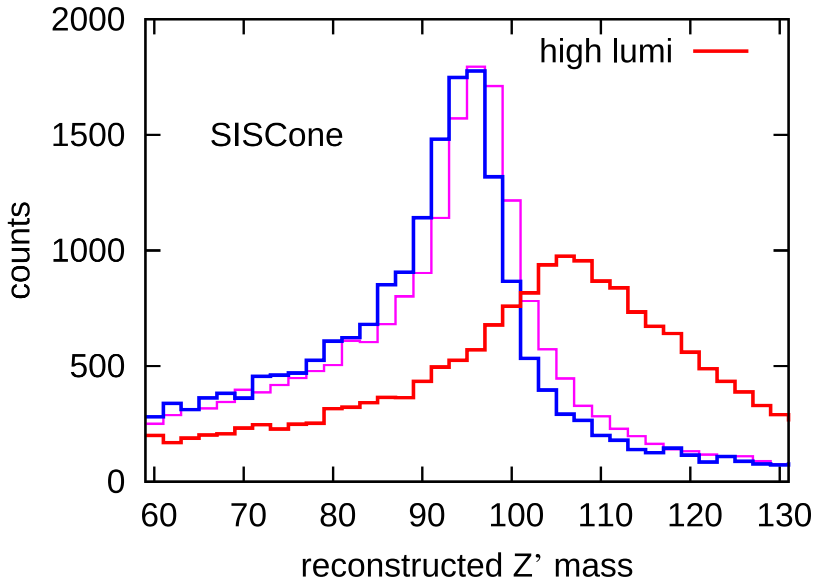
No pileup:

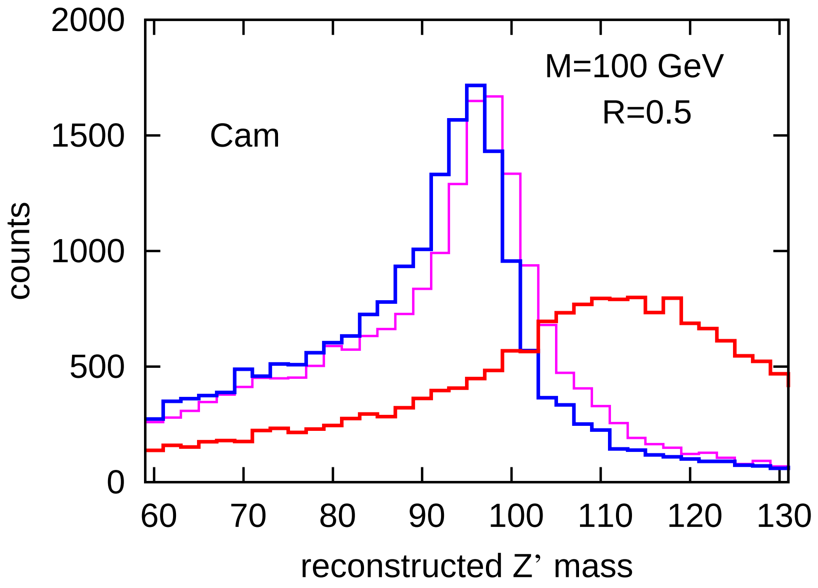
- good result
- no large subtraction effect

Subtraction in action

Les Houches jet benchmarks: $Z' \to q\bar{q} \to 2$ jets





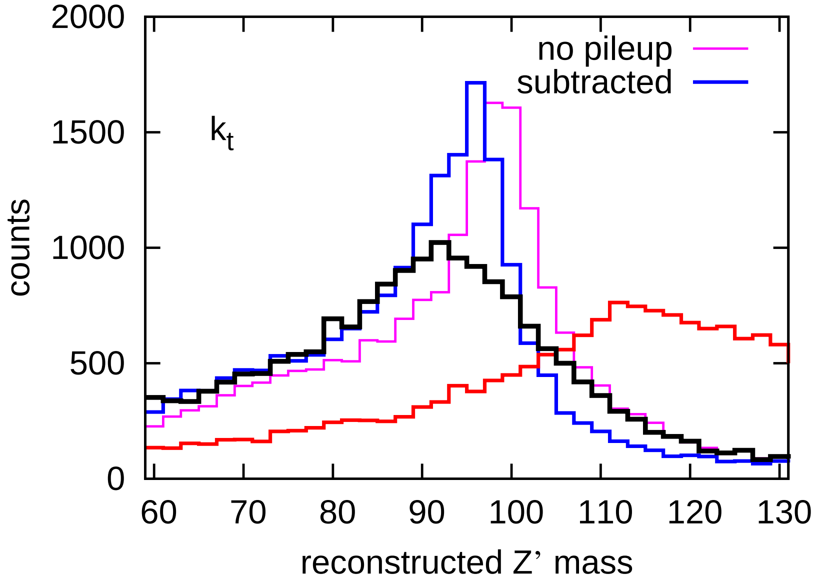


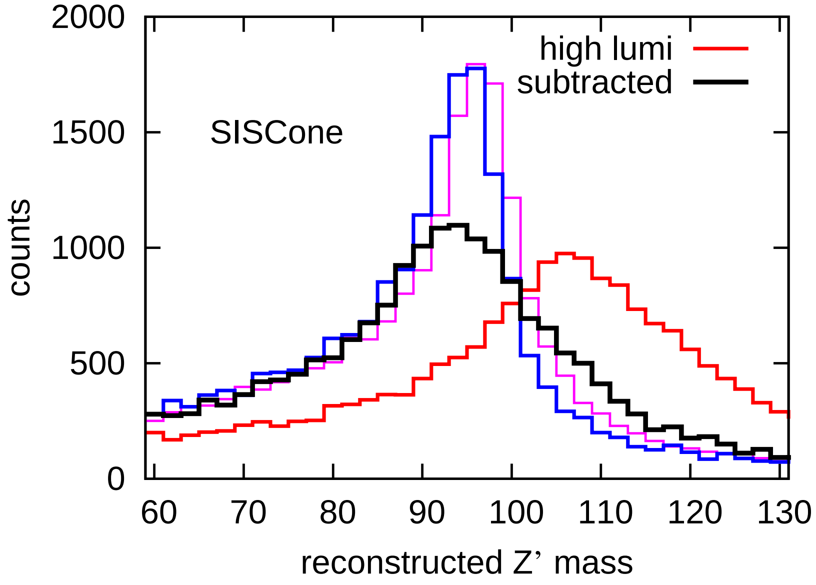
High-luminosity LHC pileup:

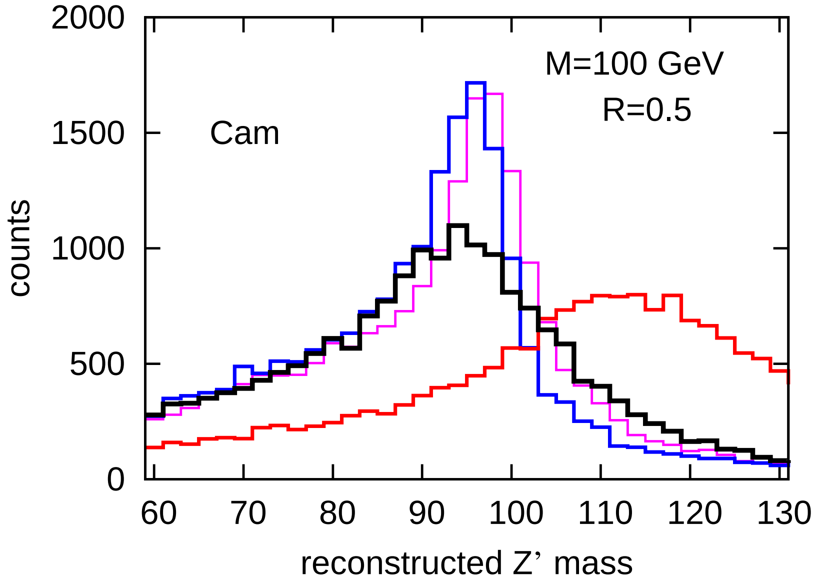
- poor quality

Subtraction in action

Les Houches jet benchmarks: $Z' \to q\bar{q} \to 2$ jets





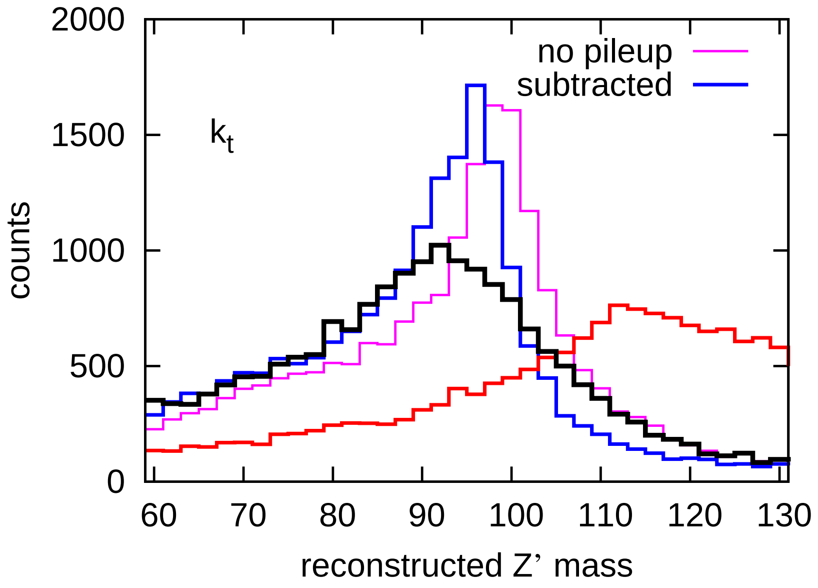


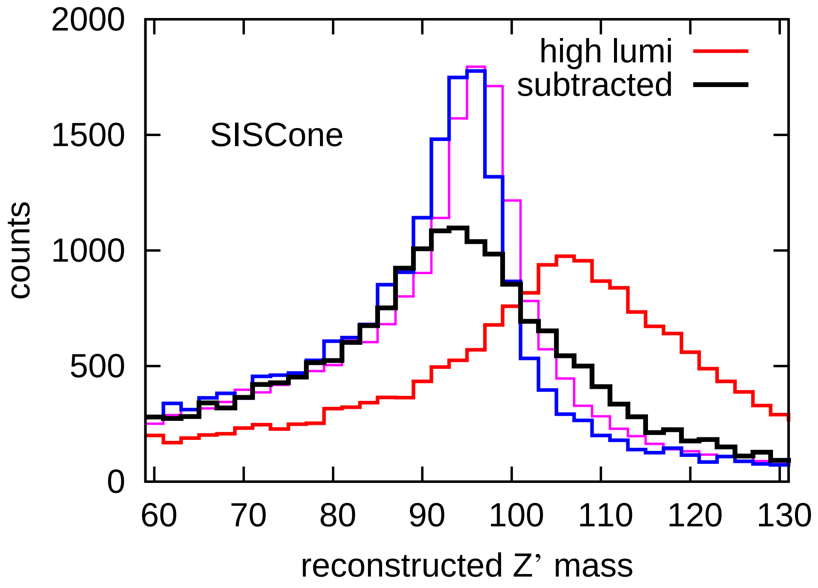
High-luminosity LHC pileup:

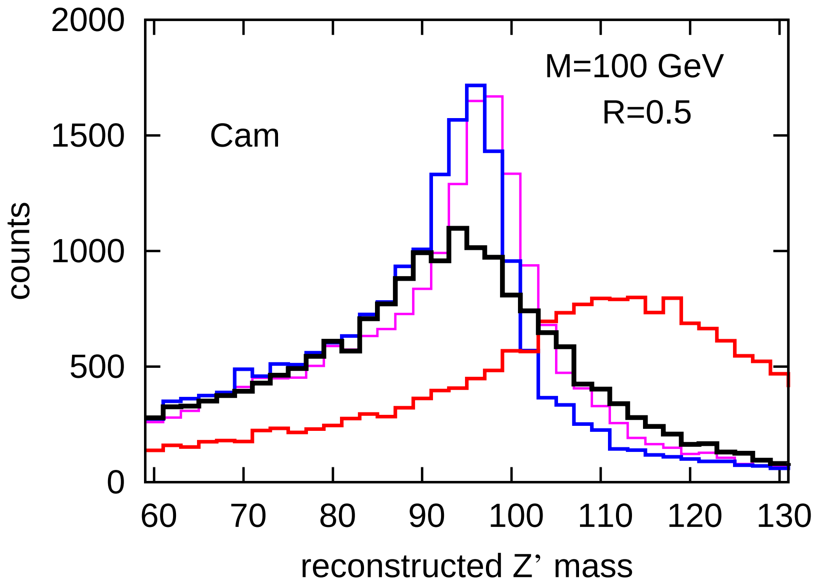
- poor quality
- good after subtraction
- subtraction reduces the width!

Subtraction in action

Les Houches jet benchmarks: $Z' \to q\bar{q} \to 2$ jets







High-luminosity LHC pileup:

- poor quality
- good after subtraction
- subtraction reduces the width!

+Background suppresion in heavy ions!

Conclusions

- SISCone: a new cone jet algorithm
  - first to satisfy requirements of the 90's!
  - mandatory for LHC
  - Get it at `http://projects.hepforge.org/siscone`
    or `http://www.lpthe.jussieu.fr/~salam/fastjet`

Conclusions

- SISCone: a new cone jet algorithm
  - first to satisfy requirements of the 90's!
  - mandatory for LHC
  - Get it at `http://projects.hepforge.org/siscone` or `http://www.lpthe.jussieu.fr/~salam/fastjet`
- TODO: in-depth study of k_t/Cam vs. cone.
  Under study with the "Les Houches jet benchmarks"

Conclusions and perspectives

- SISCone: a new cone jet algorithm
  - first to satisfy requirements of the 90's!
  - mandatory for LHC
  - Get it at `http://projects.hepforge.org/siscone`
    or `http://www.lpthe.jussieu.fr/~salam/fastjet`
- TODO: in-depth study of k_t/Cam vs. cone.
  Under study with the "Les Houches jet benchmarks"

- New concept: the area of a jet (in FastJet 2.3.0)
  - active, passive: definition and basic properties
  - pileup effects subtraction, background subtraction in heavy ions

Conclusions and perspectives

- SISCone: a new cone jet algorithm
  - first to satisfy requirements of the 90's!
  - mandatory for LHC
  - Get it at `http://projects.hepforge.org/siscone`
    or `http://www.lpthe.jussieu.fr/~salam/fastjet`
- TODO: in-depth study of k_t/Cam vs. cone.
  Under study with the "Les Houches jet benchmarks"

- New concept: the area of a jet (in FastJet 2.3.0)
  - active, passive: definition and basic properties
  - pileup effects subtraction, background subtraction in heavy ions
- TODO:
  - anomalous dimension resummation
  - only the beginning...