Progress in defining jets

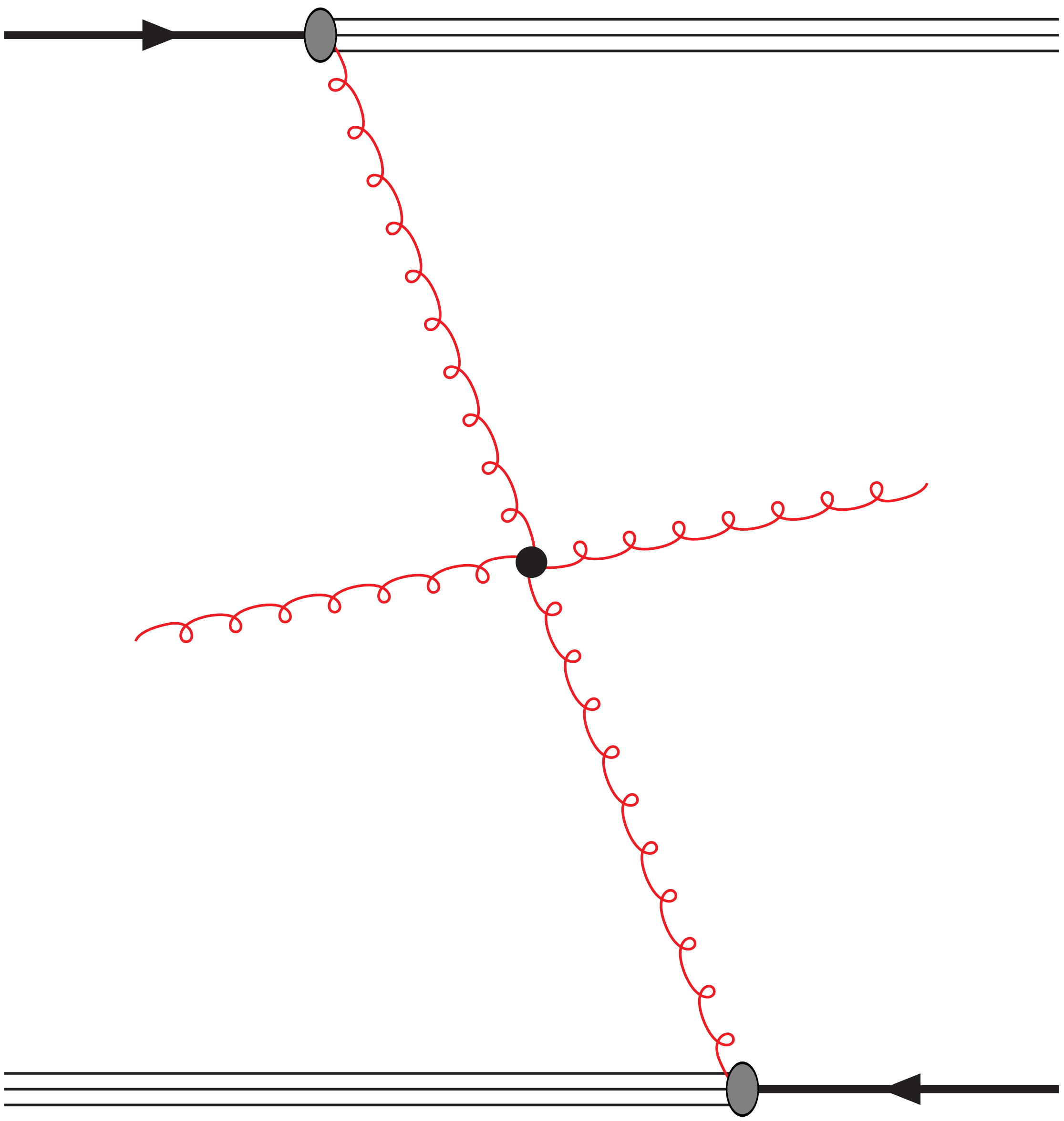
Grégory Soyez

Brookhaven National Laboratory

in collaboration with G. Salam, M. Cacciari

arXiv:0704:0292, arXiv:0802:1188, arXiv:0802.1189

General (over)simplified picture

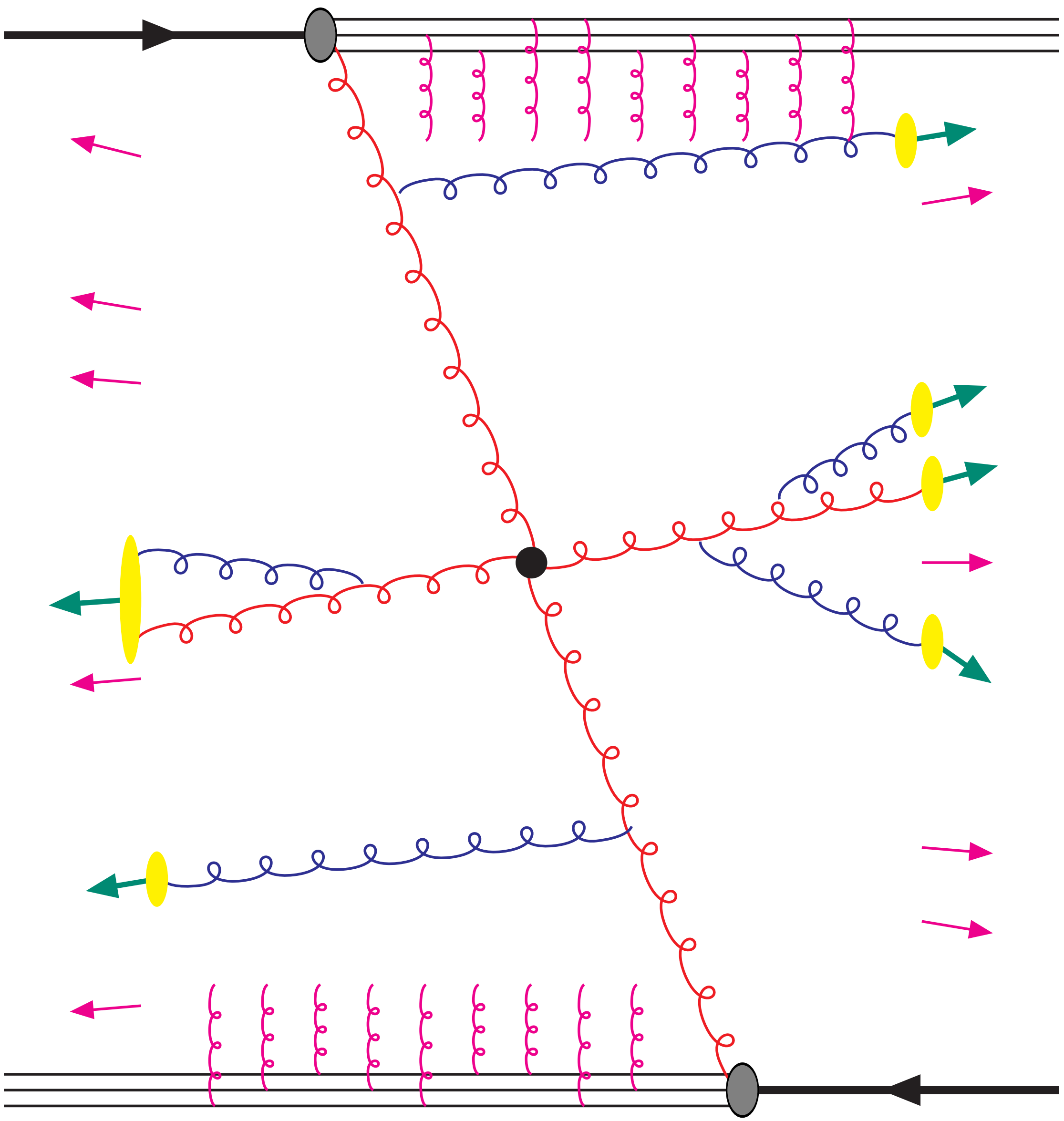


Hard scattering ($2 \to n$)

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



Hard scattering ($2 \to n$)

Parton level

$\approx$ resummed collinear div.

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

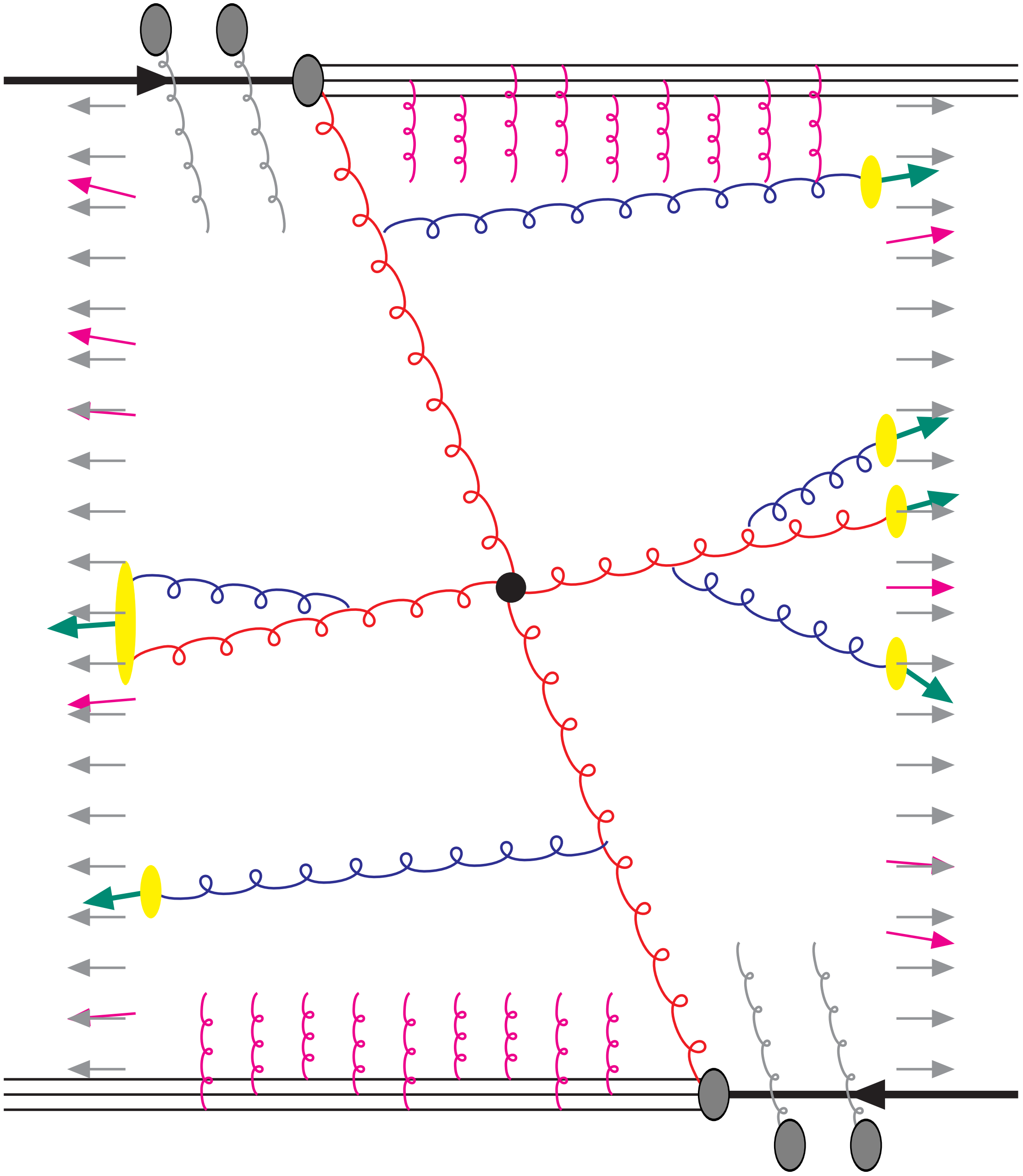
Hadron level: hadronisation

Underlying event

beam remnants interactions

$\Rightarrow$ soft background

General (over)simplified picture



Hard scattering $(2 \to n)$

Parton level

$\approx$ resummed collinear div.

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

Hadron level: hadronisation

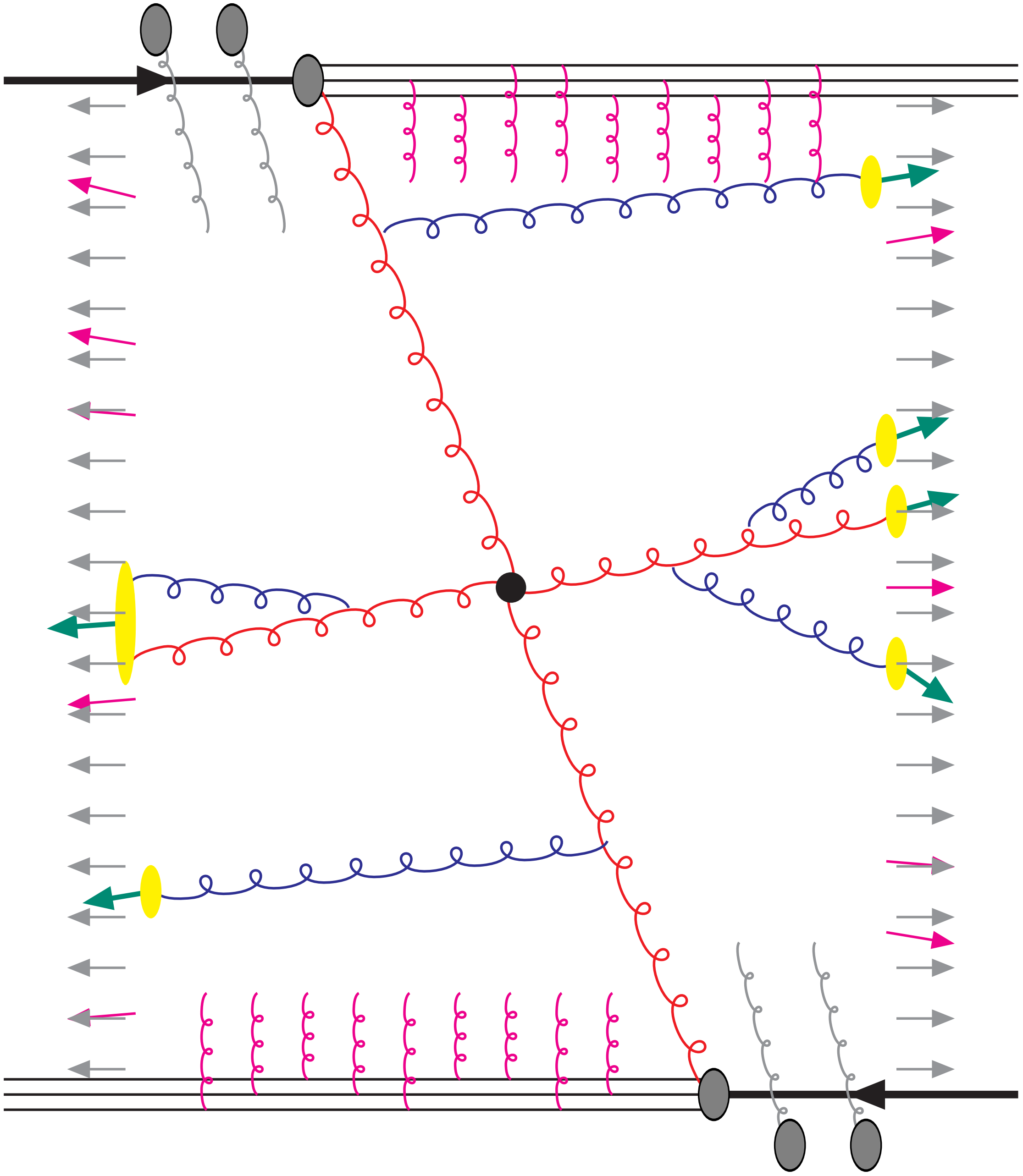
Underlying event

beam remnants interactions

$\Rightarrow$ soft background

Pileup

$\approx$ uniform soft background



Hard scattering ($2 \to n$)

Parton level

$\approx$ resummed collinear div.

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

Hadron level: hadronisation

Underlying event

beam remnants interactions

$\Rightarrow$ soft background

Pileup

$\approx$ uniform soft background

How to access the hard scattering?

Plan

Define jets:

- Idea: a bunch of collimated partons/particles
- But: partons are ambiguous

  $\Rightarrow$ different *jet definitions*

  $\Rightarrow$ jets are really what they're defined to be

Plan

Define jets:

- Idea: a bunch of collimated partons/particles
- But: partons are ambiguous
  $\Rightarrow$ different *jet definitions*
  $\Rightarrow$ jets are really what they're defined to be

Outline:

- Part 1: Failures to meet the fundamental requirements
  $\Rightarrow$ introduce *new algorithms*
- Part 2: Deal with the soft background contamination (pileup or HI background) $\Rightarrow$ introduce *jet areas* (how to control and use them)

Note: Jet definition = algorithm + parameters
Also valid for jets in heavy-ion collisions

The main recipes (1/2)

Algorithm Class 1: *successive recombinations*

- define a distance d_{ij} between any pair of particles
- repeatedly recombine the 2 closest "particles" into one
- stop when they are all more than R apart

Common distance choices:

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

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

The main recipes (2/2)

BROOKHAVEN NATIONAL LABORATORY

Algorithm Class 2: *cone*

> **stable cones** (radius R) such that:
> the total momentum of its contents points in the direction of its centre

Seeded (iterative) approaches: iterate from an initial position until stable

- seed = initial particle
- seed = midpoint between stable cones found at first step

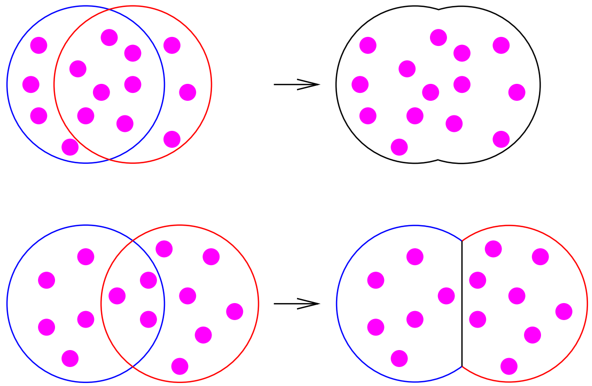
Algorithm Class 2: *cone*

stable cones (radius R) such that:
the total momentum of its contents points in the direction of its centre

Class 2(a): cone with split-merge (ex.: JetClu, Atlas, MidPoint):

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

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



Class 2(b): cone with progressive removal (ex.: Iterative Cone)

- iterate from the hardest seed
- remove the stable cone as a jet and start again $\Rightarrow$ "regular/circular" jets

20th century jet finders

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

Part 1
21st century: towards a solid toolkit

1990: fixing the rules

SNOWMASS accords, Tevatron 1990 (i.e. old!):

Several important properties that should be met by a jet definition are [3]:

1. Simple to implement in an experimental analysis;
2. Simple to implement in the theoretical calculation;
3. Defined at any order of perturbation theory;
4. Yields finite cross section at any order of perturbation theory;
5. Yields a cross section that is relatively insensitive to hadronization.

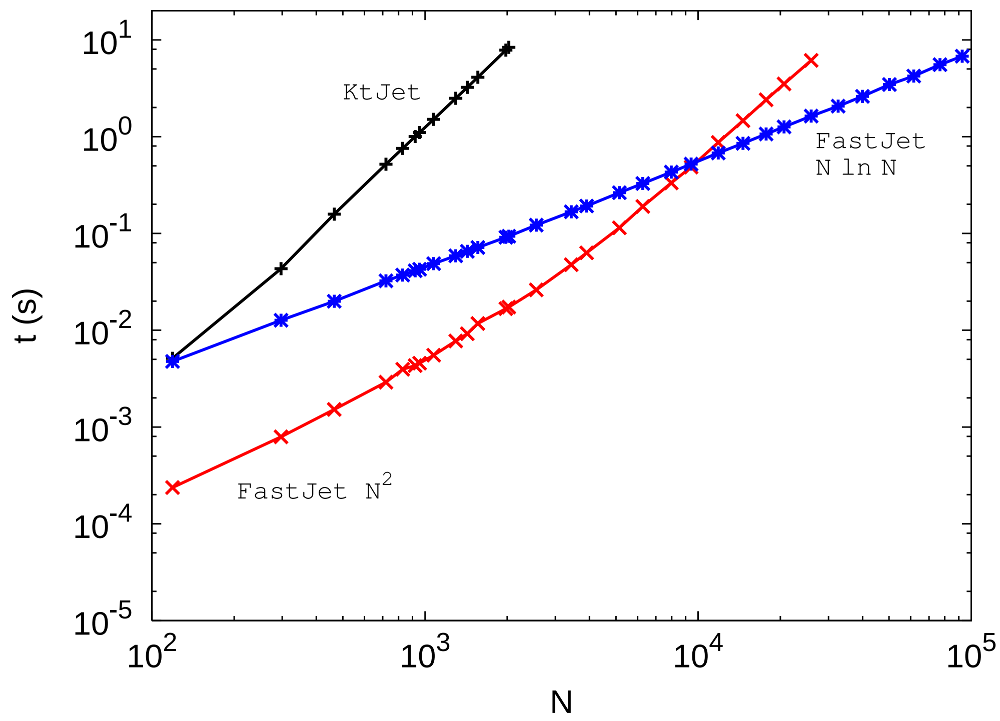
i.e. usable by theoreticians (*e.g.* finite perturbative results)

and experimentalists (*e.g.* fast enough, not much UE sensitivity)

Speed improvement

[M. Cacciari, G. Salam, 06]

- Speeding up the k_t and Cam/Aachen algorithms
  using computational-geometry techniques: $\mathcal{O}\left(N^3\right) \to \mathcal{O}\left(N \log N\right)$



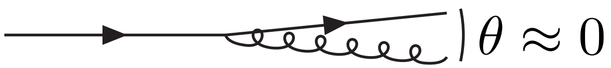
- C++ implementation in `FastJet`
  `http://www.fastjet.fr` (M. Cacciari, G. Salam, G.S.)

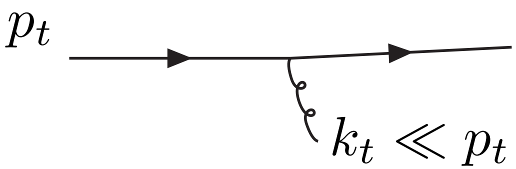
QCD divergences

QCD probability for gluon bremsstrahlung at angle θ and $\perp$-mom. k_t:

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

Two divergences:





Collinear

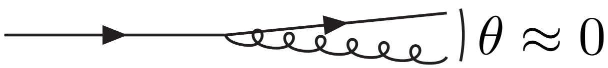
Soft

QCD divergences

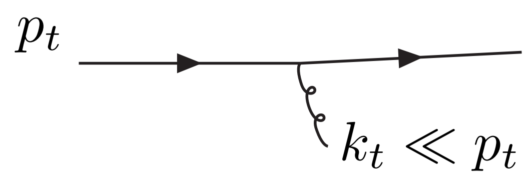
QCD probability for gluon bremsstrahlung at angle θ and $\perp$-mom. k_t:

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

Two divergences:



Collinear

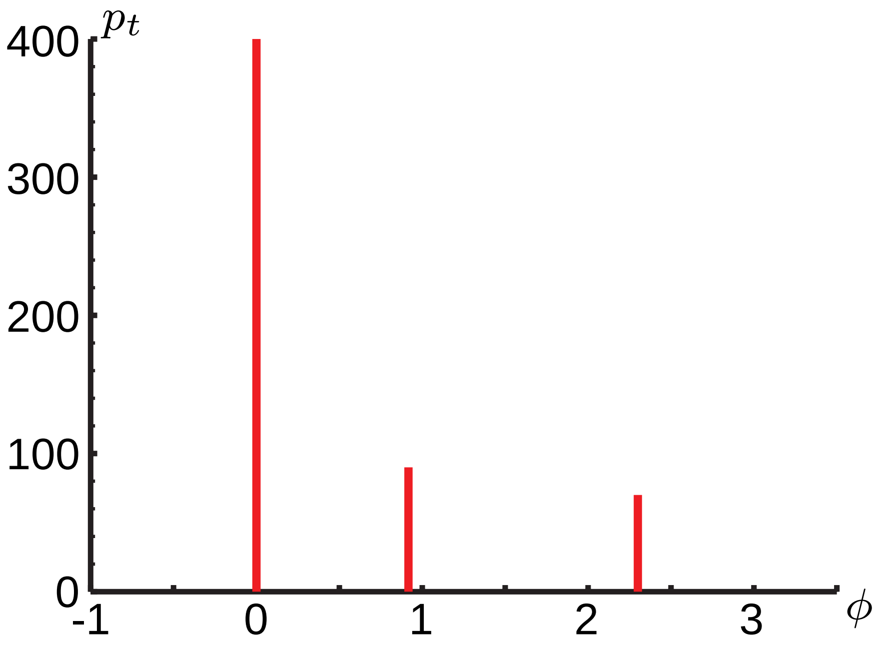


Soft

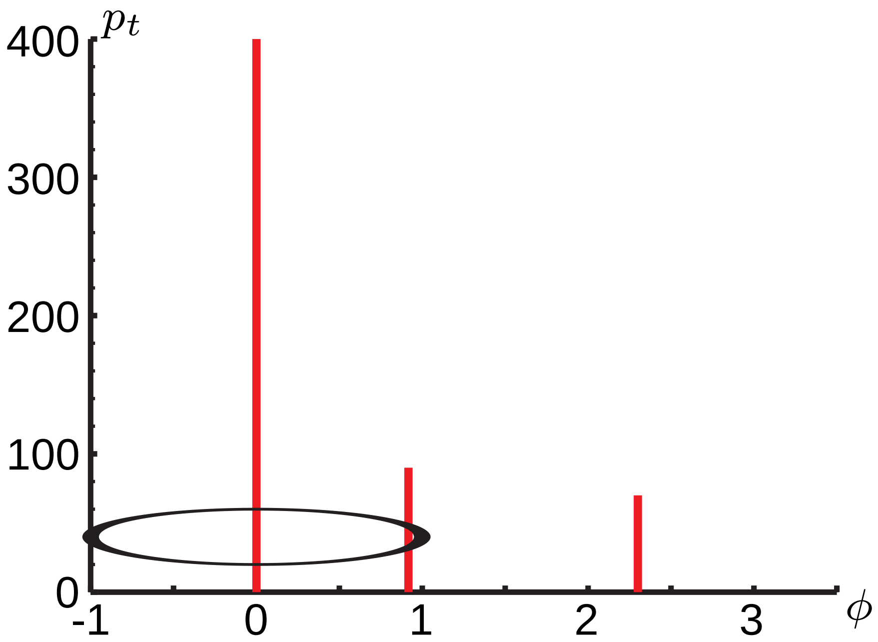
For pQCD to make sense, the (hard) jets should not change when

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

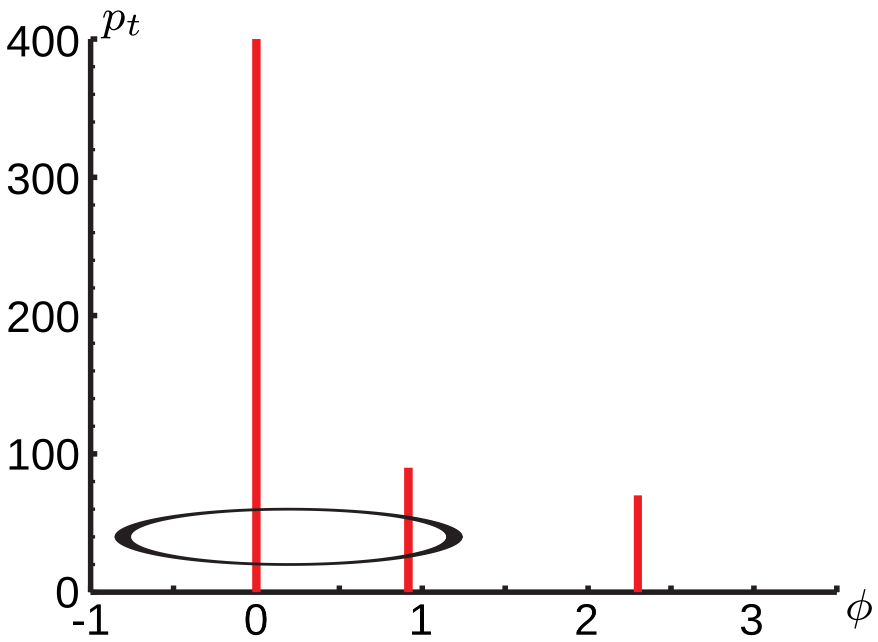
IR unsafety of the Midpoint alg



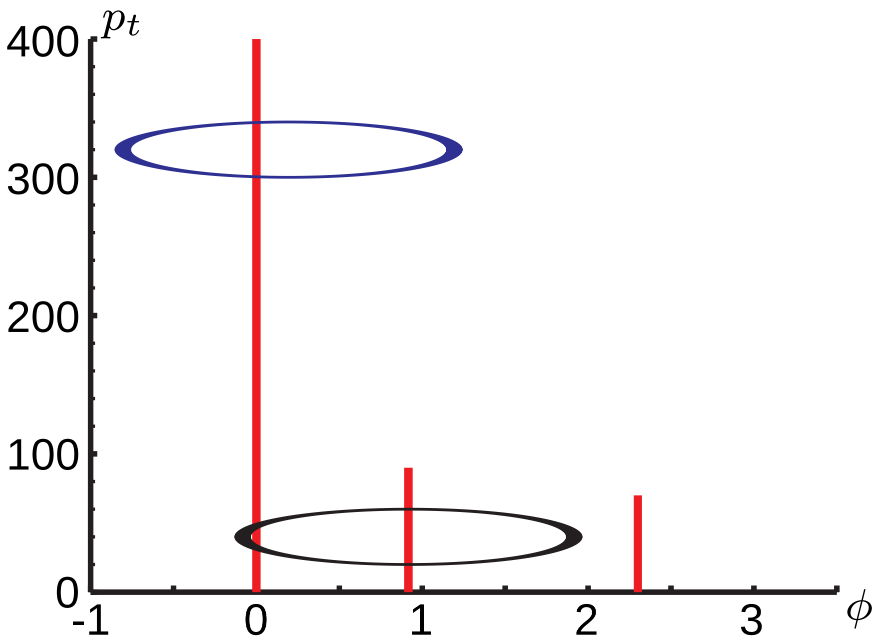
IR unsafety of the Midpoint alg



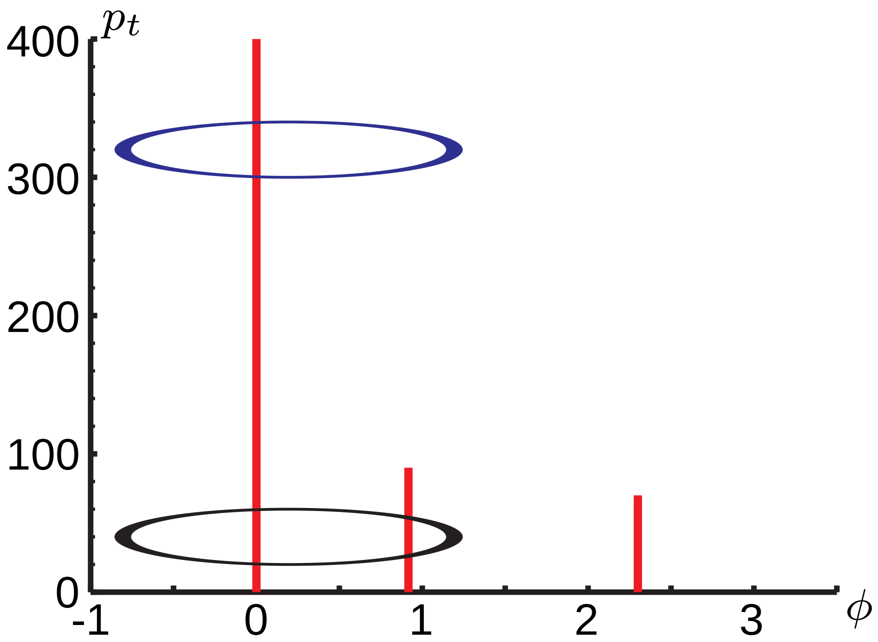
IR unsafety of the Midpoint alg



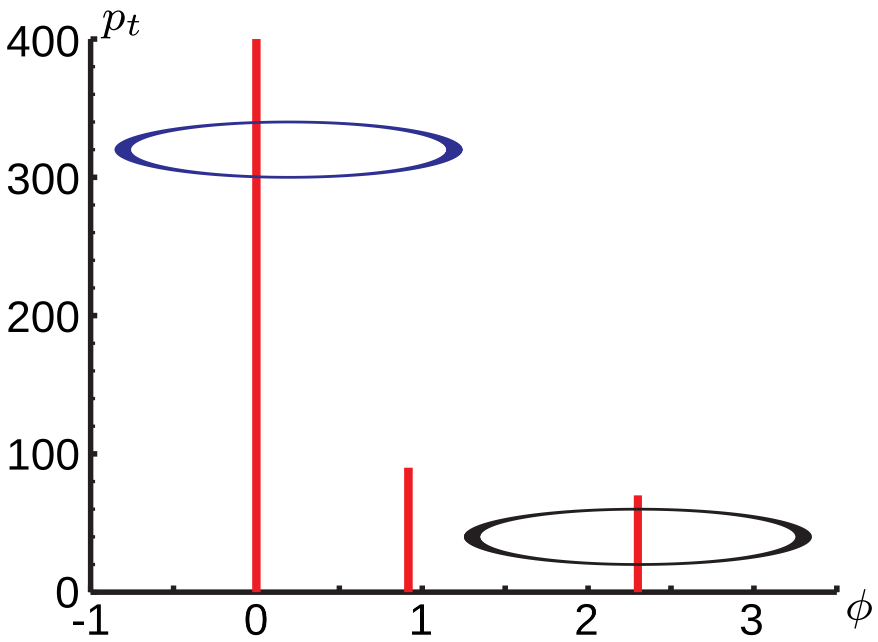
IR unsafety of the Midpoint alg



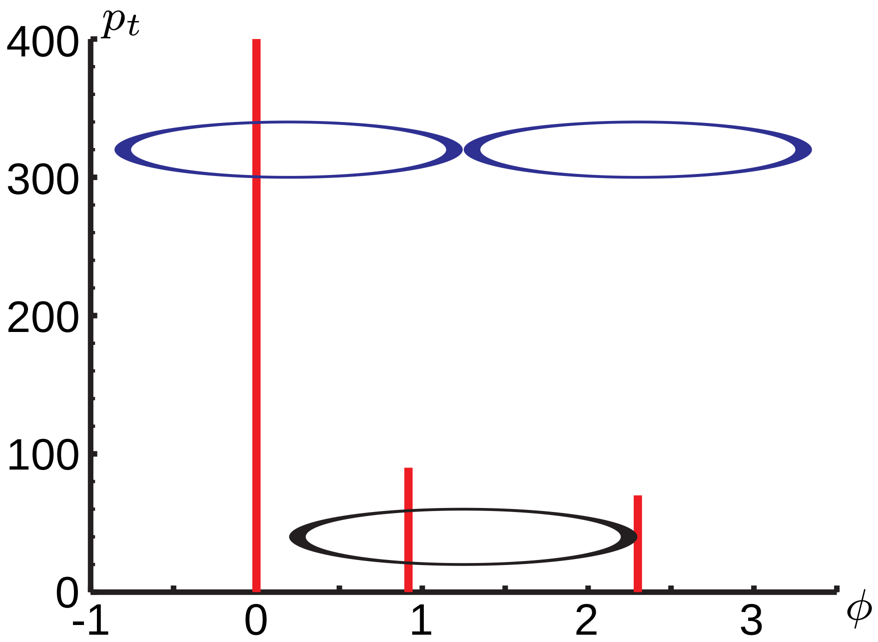
IR unsafety of the Midpoint alg



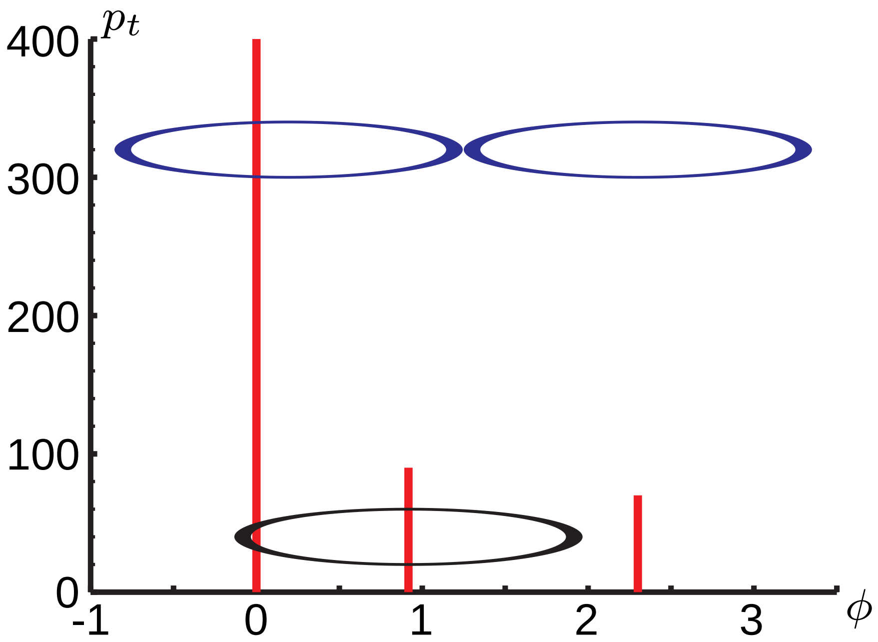
IR unsafety of the Midpoint alg



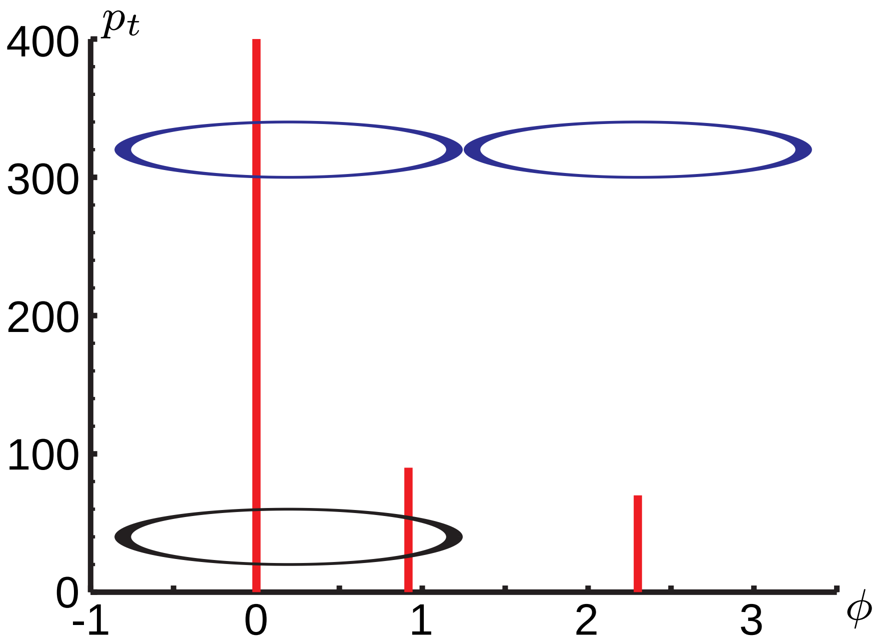
IR unsafety of the Midpoint alg



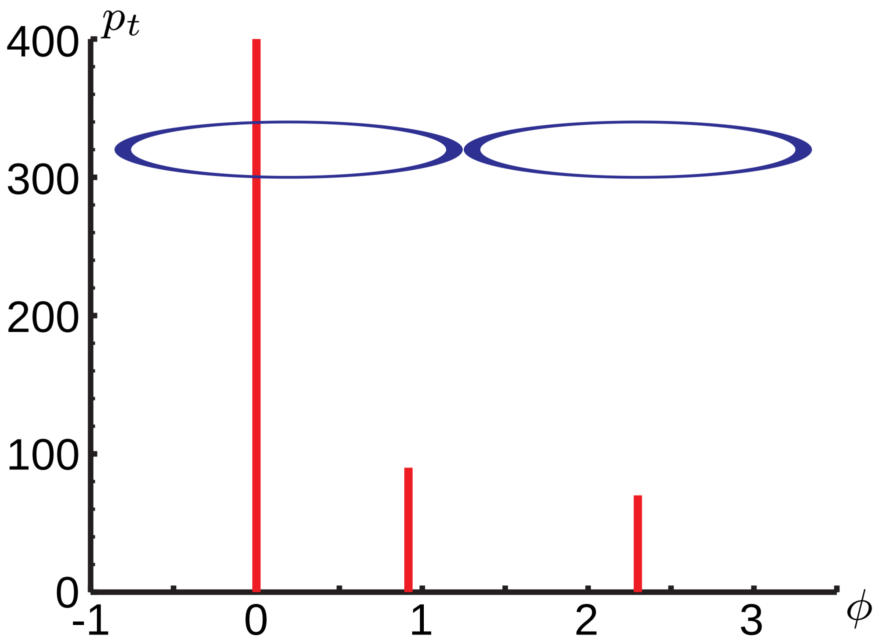
IR unsafety of the Midpoint alg

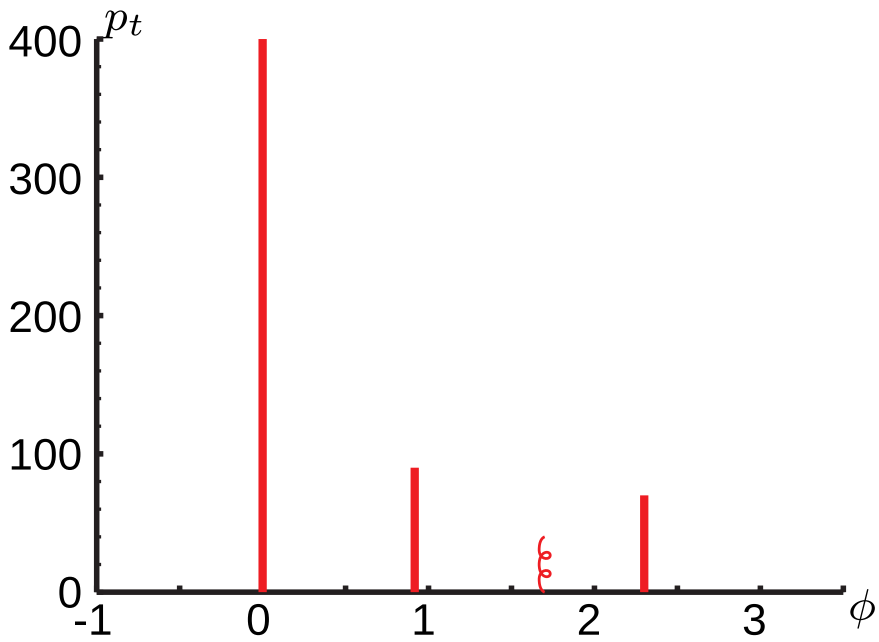


IR unsafety of the Midpoint alg

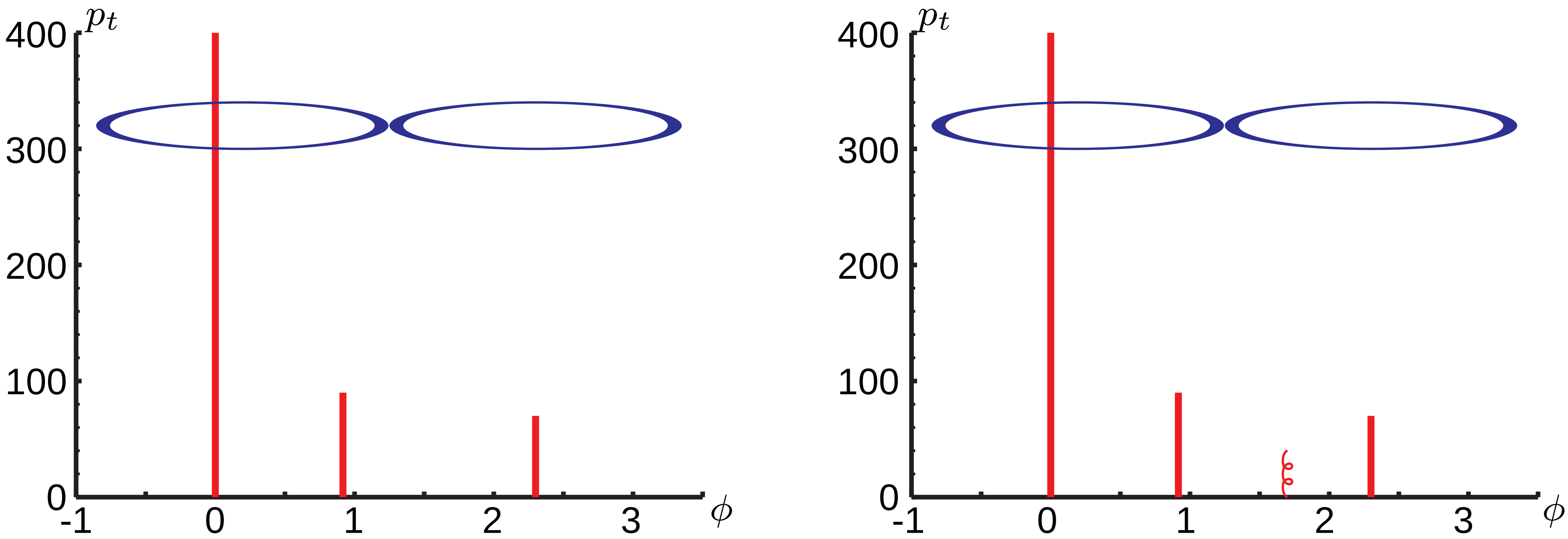


IR unsafety of the Midpoint alg

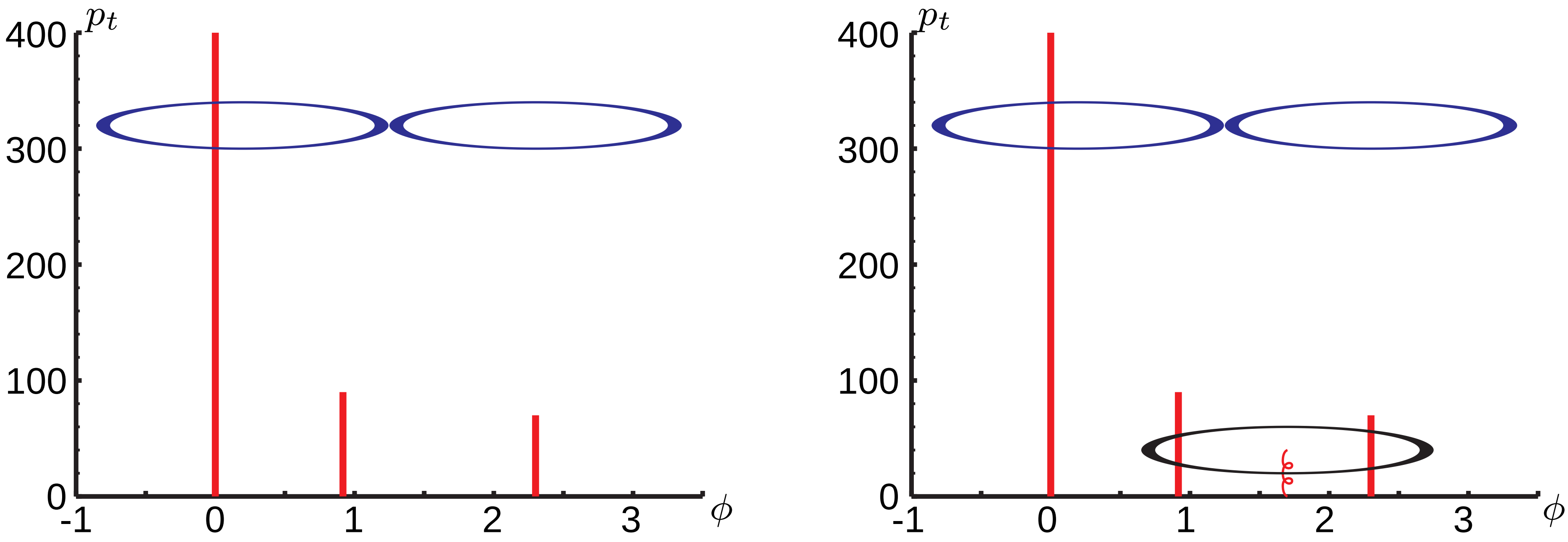




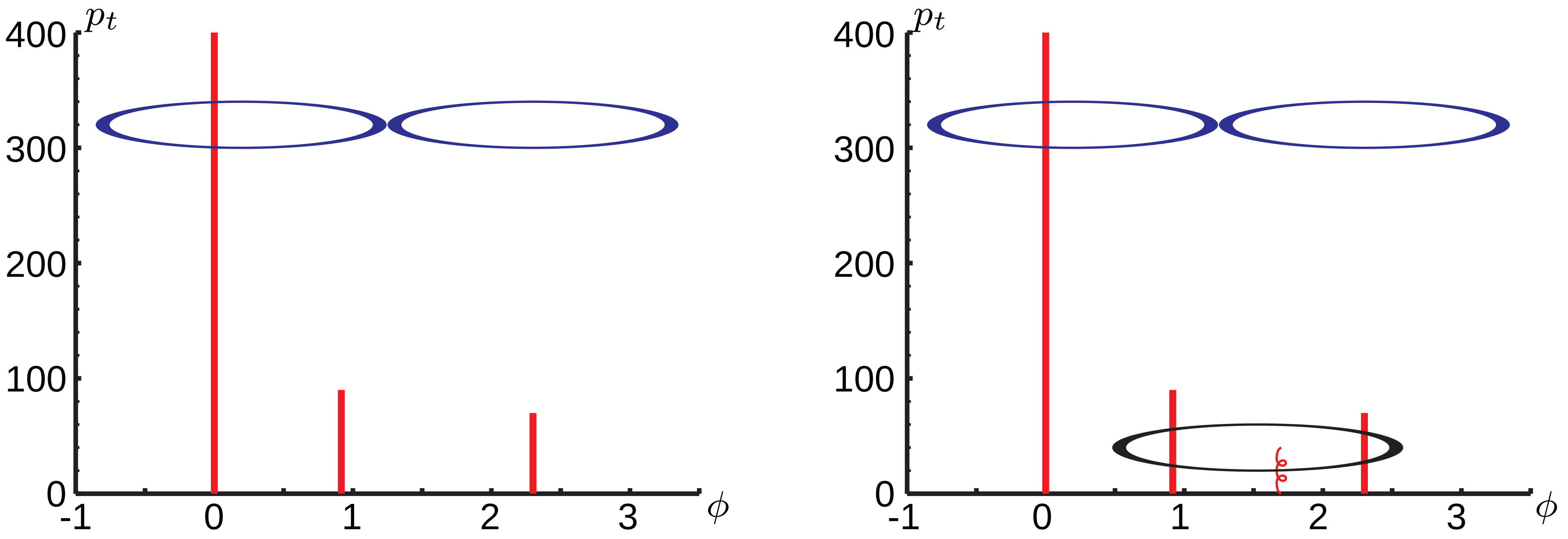
IR unsafety of the Midpoint alg



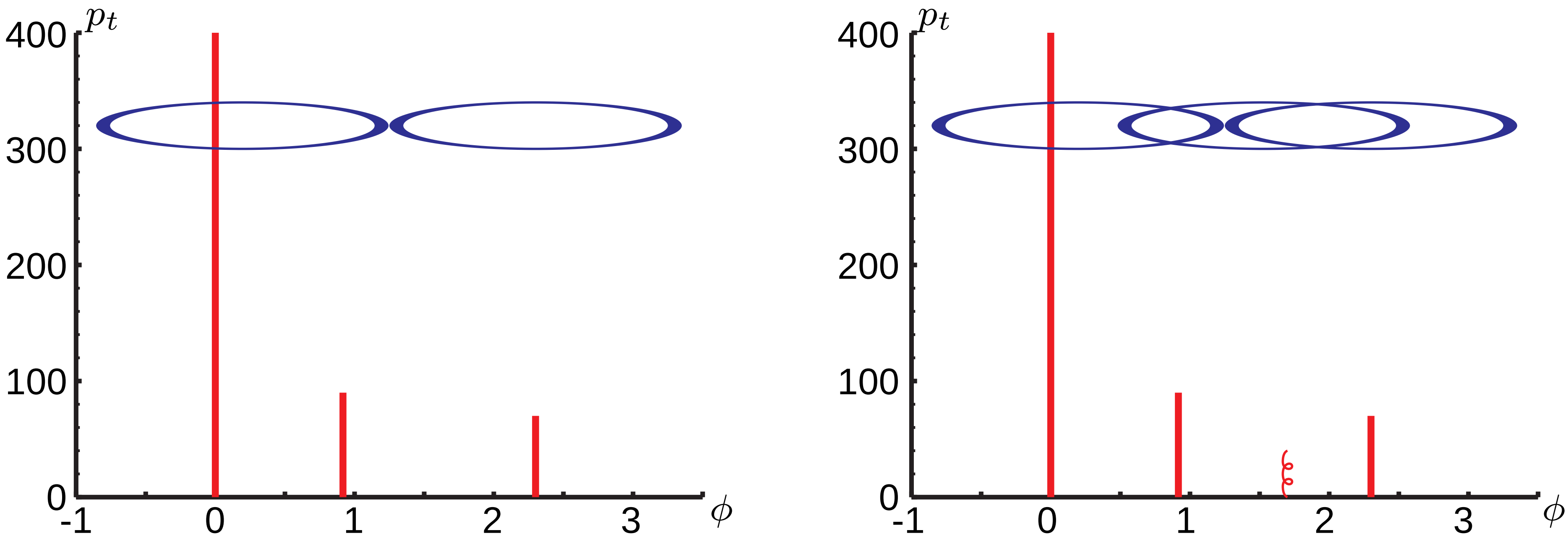
IR unsafety of the Midpoint alg



IR unsafety of the Midpoint alg



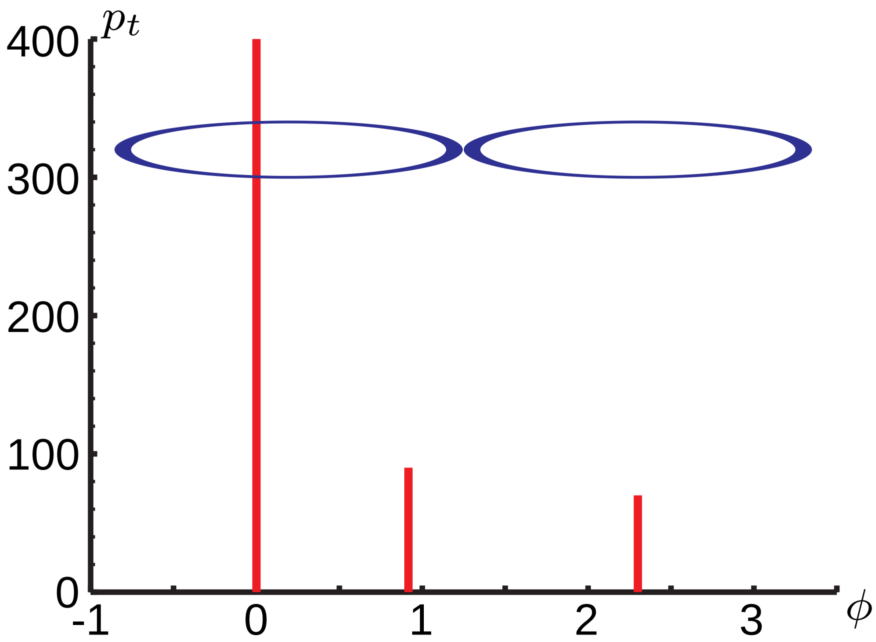
IR unsafety of the Midpoint alg

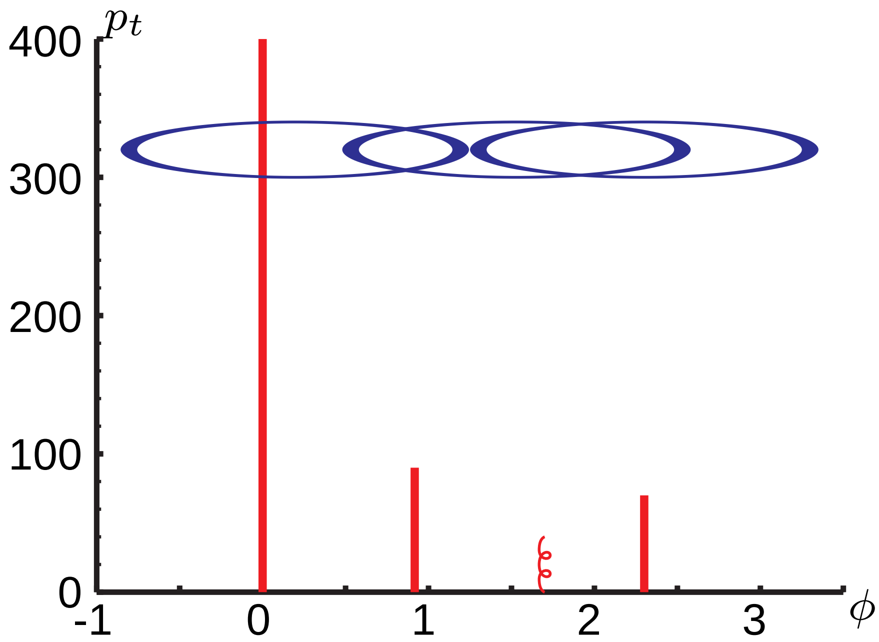


Stable cones:

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

IR unsafety of the Midpoint alg





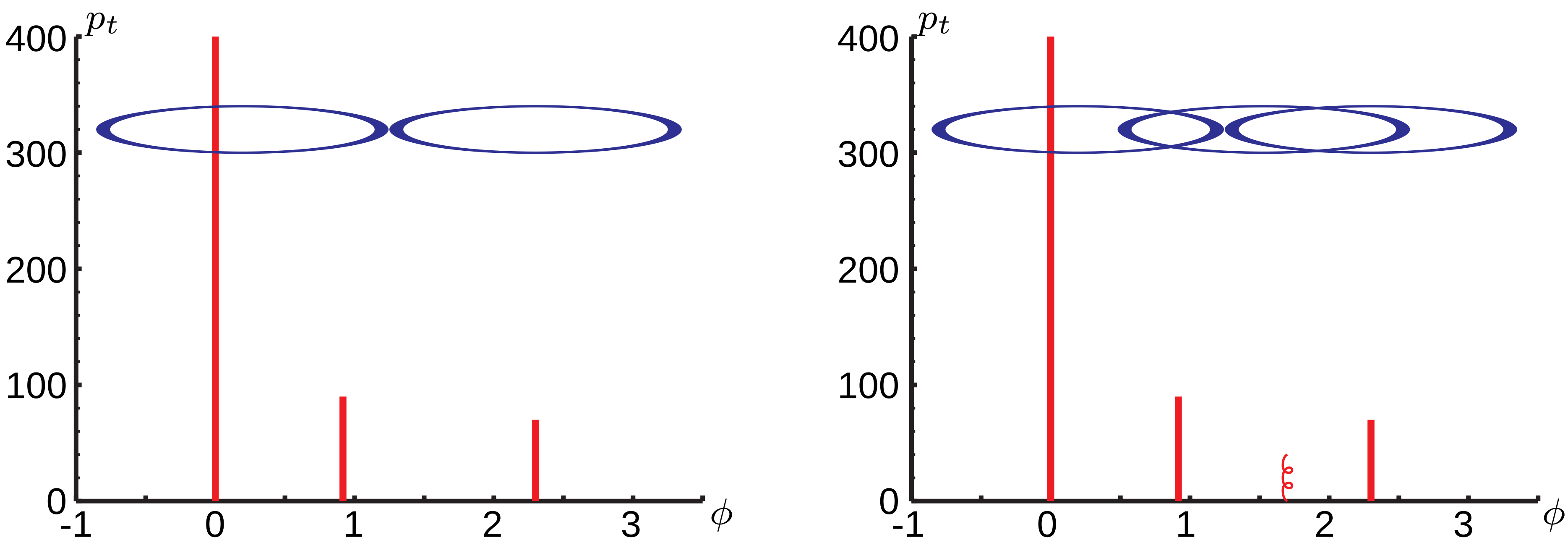
Stable cones:

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

Jets: ($f = 0.5$)

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

IR unsafety of the Midpoint alg



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

Solution: SISCone

Solution: use a seedless approach, find ALL stable cones

Solution: SISCone

Solution: use a seedless approach, find ALL stable cones

Issue: watch your speed!

- check each subset of particle: $\mathcal{O}\left(N2^N\right)$

  irrealistically slow: $\sim 10^{17}$ years for $N = 100$
- midpoint: $\mathcal{O}\left(N^3\right)$
- New solution: use geometry again $\Rightarrow \mathcal{O}\left(N^2 \log(N)\right)$

  Idea: enumerate enclosures by enumerating pairs of particles

Solution: SISCone

Solution: use a seedless approach, find ALL stable cones

Issue: watch your speed!

- check each subset of particle: $\mathcal{O}\left(N2^N\right)$
  irrealistically slow: $\sim 10^{17}$ years for $N = 100$
- midpoint: $\mathcal{O}\left(N^3\right)$
- New solution: use geometry again $\Rightarrow \mathcal{O}\left(N^2 \log(N)\right)$
  Idea: enumerate enclosures by enumerating pairs of particles

$\longrightarrow$ SISCone: Seedless Infrared-Safe Cone algorithm

http://projects.hepforge.org/siscone [G.Salam, G.S., 07]

Notes:

- first cone to be infrared-and-collinear safe
- also available from FastJet

Physical impact

Midpoint IR unsafe at $\mathcal{O}(\alpha_s^4)$ (or $\mathcal{O}(\alpha_{\rm ew}\alpha_s^3)$)

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

Huge theoretical effort to compute multileg/NLO processes

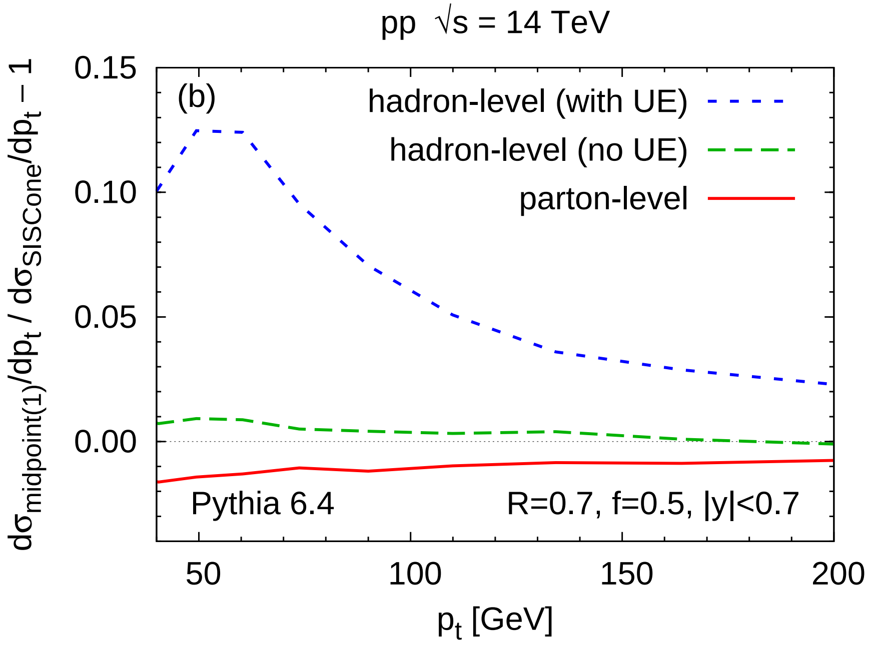
That can be wasted by using unappropriate tools.

Physical impact

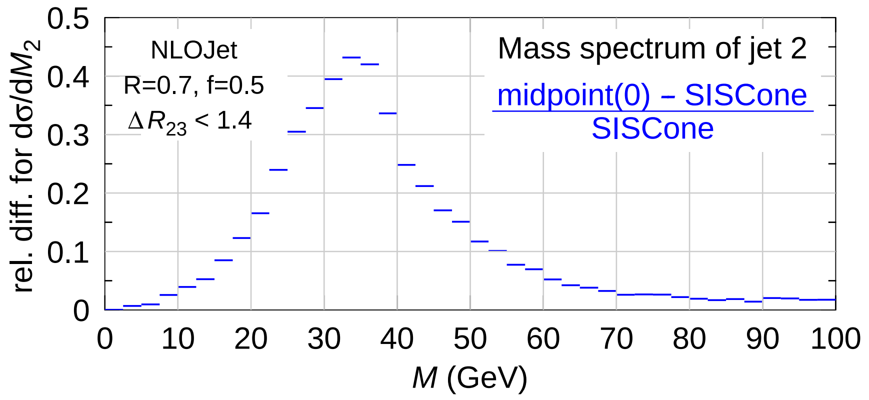
Midpoint IR unsafe at $\mathcal{O}(\alpha_s^4)$ (or $\mathcal{O}(\alpha_{\rm ew}\alpha_s^3)$)

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

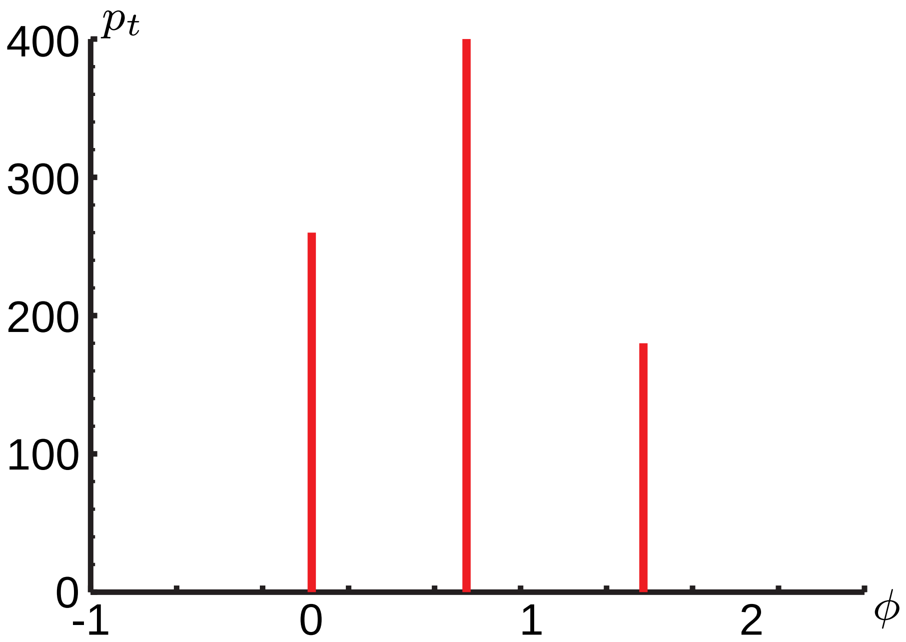
Example: (Midpoint-SISCone)/SISCone



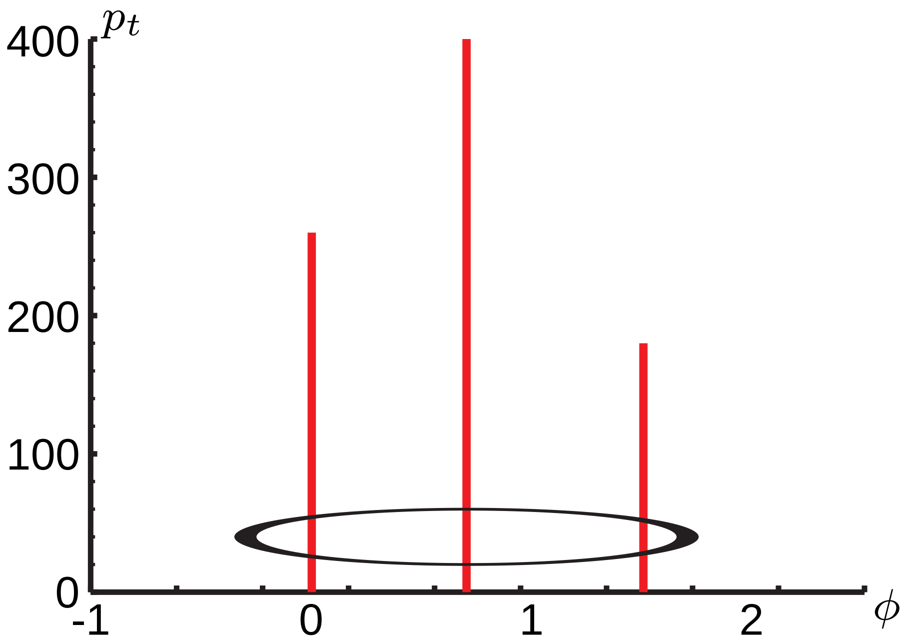
- Incl. cross-section: a few %
- Masses in 3-jet events: $\sim 45\%$



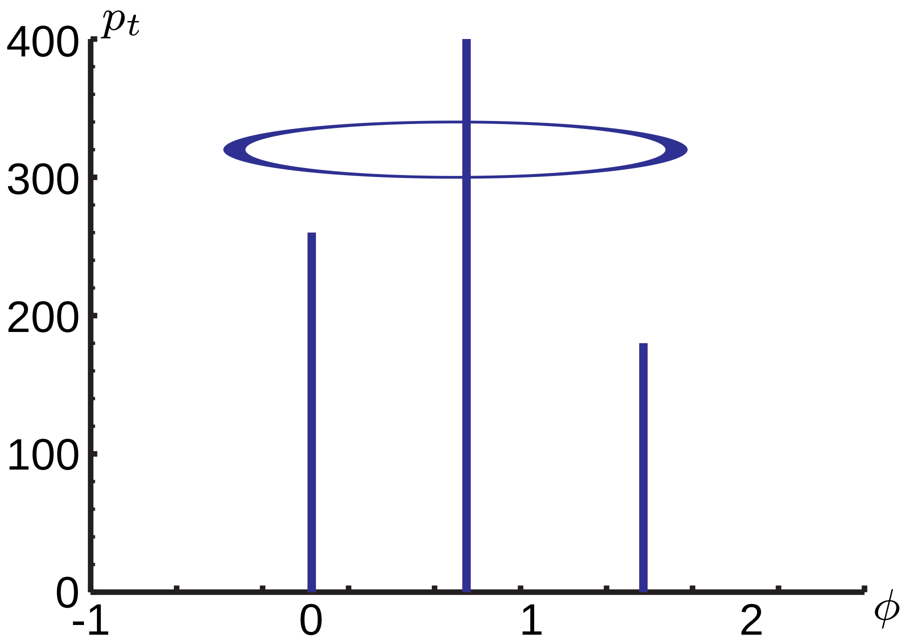
Coll. unsafety of the iterative cone



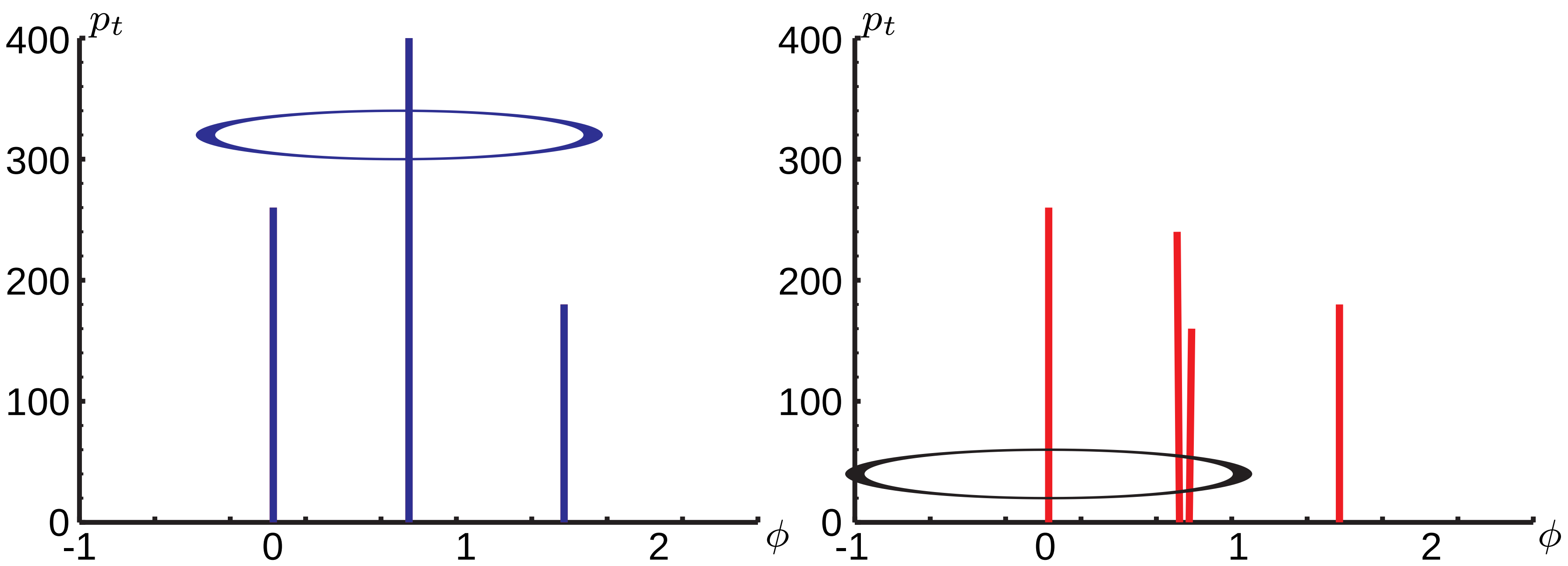
Coll. unsafety of the iterative cone



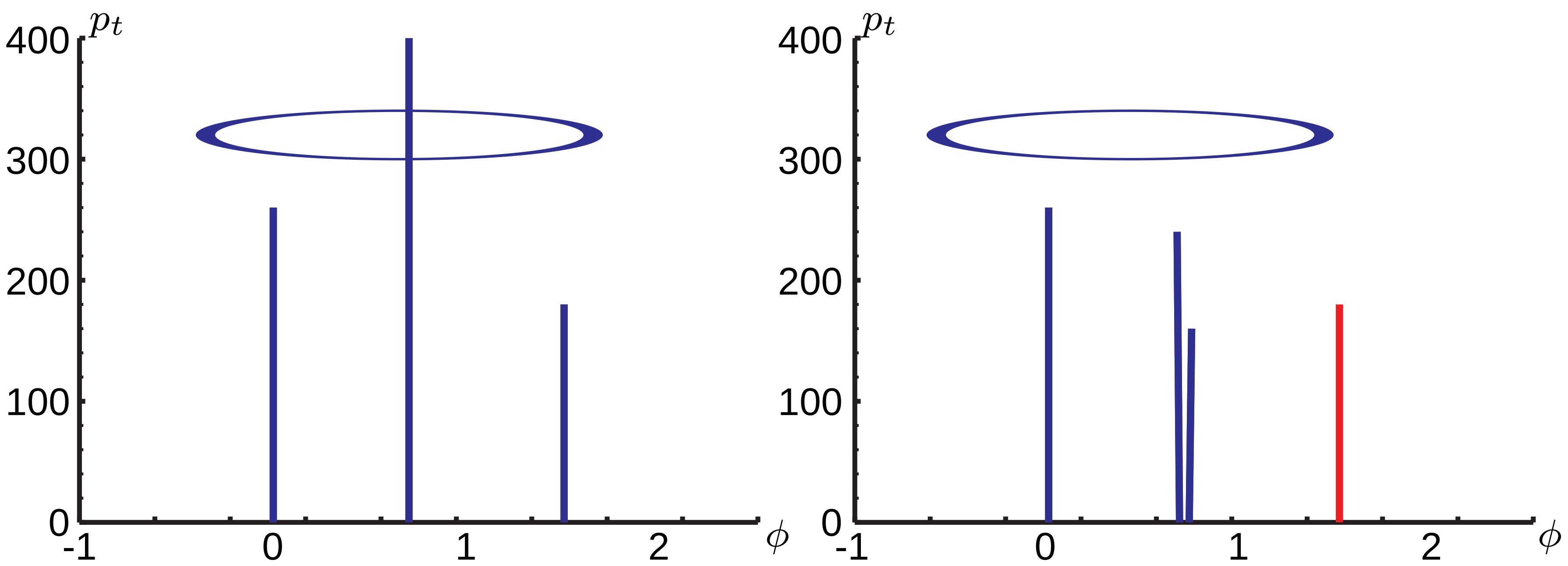
Coll. unsafety of the iterative cone



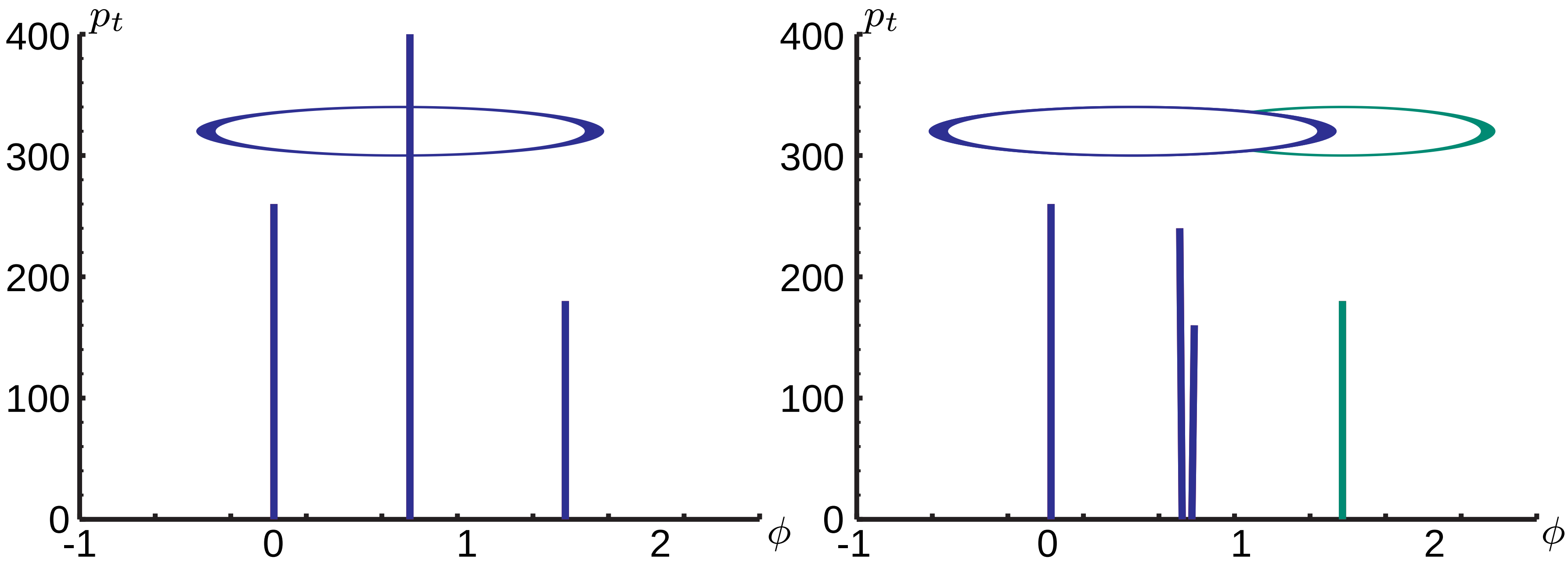
Coll. unsafety of the iterative cone



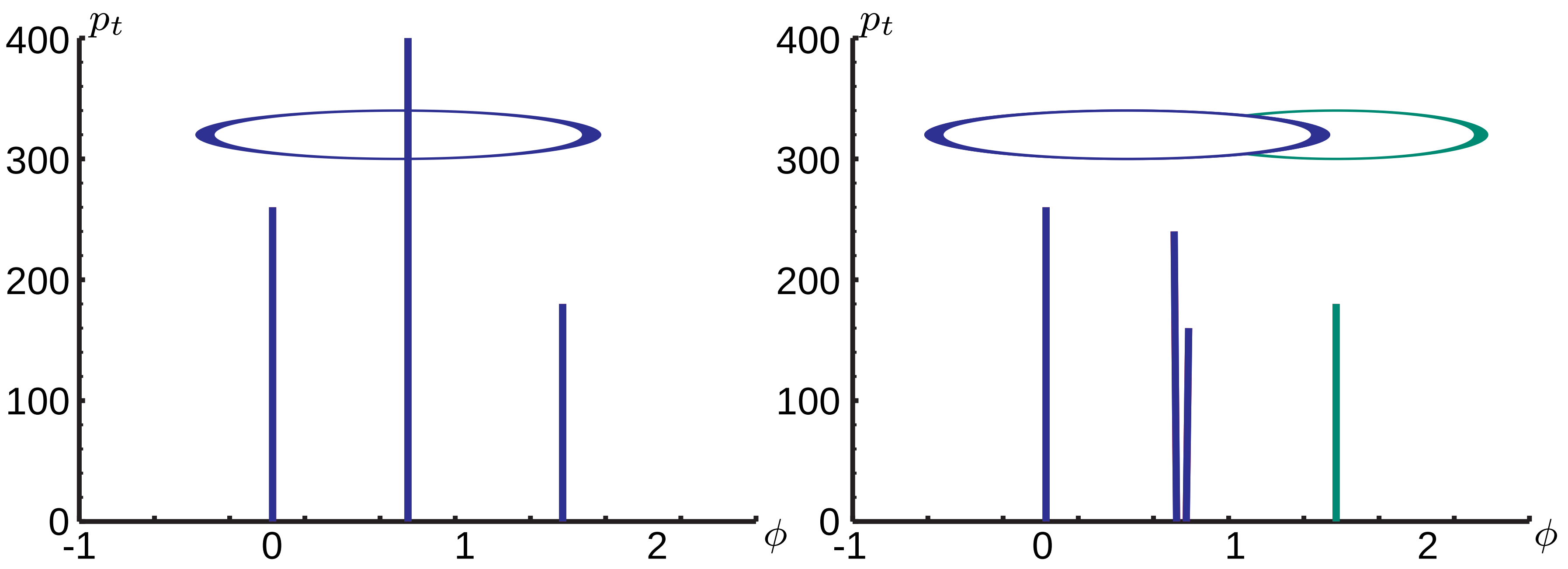
Coll. unsafety of the iterative cone



Coll. unsafety of the iterative cone



Coll. unsafety of the iterative cone



- Before collinear spliting: 1 jet
- After collinear spliting: 2 jets

$\longrightarrow$ **collinear unsafety of the iterative cone algorithm**

Come back to recombination-type algorithms:

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

- $p = 1$: k_t algorithm
- $p = 0$: Aachen/Cambridge algorithm

Anti-k_t

Come back to recombination-type algorithms:

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

- $p = 1$: k_t algorithm
- $p = 0$: Aachen/Cambridge algorithm
- $p = -1$: anti-k_t algorithm [M.Cacciari, G.Salam, G.S.,JHEP 04 (08) 063]

Anti-k_t

Come back to recombination-type algorithms:

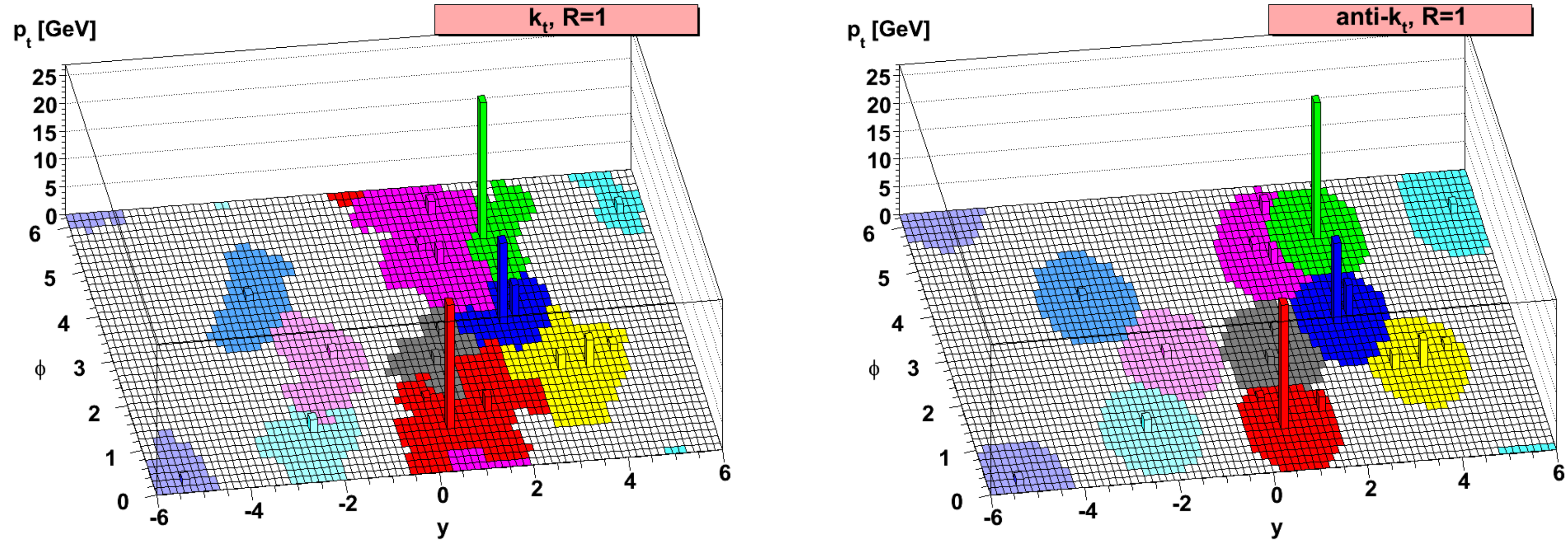
$$d_{ij} = \min(k_{t,i}^{2p}, k_{t,j}^{2p})\ (\Delta\phi_{ij}^2 + \Delta\eta_{ij}^2)$$

- $p = 1$: k_t algorithm
- $p = 0$: Aachen/Cambridge algorithm
- $p = -1$: anti-k_t algorithm [M.Cacciari, G.Salam, G.S.,JHEP 04 (08) 063]

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

Hard event + homogeneous soft background



- anti-k_t is soft-resilient
- fast implementation as for k_t and Cambrigde/Aachen
- more later in this talk...

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
- ATLAS Cone
- CMS Iterative Cone
- PyCell/CellJet
- GetJet
- SISCone

4 available
safe algorithms

All accessible from FastJet

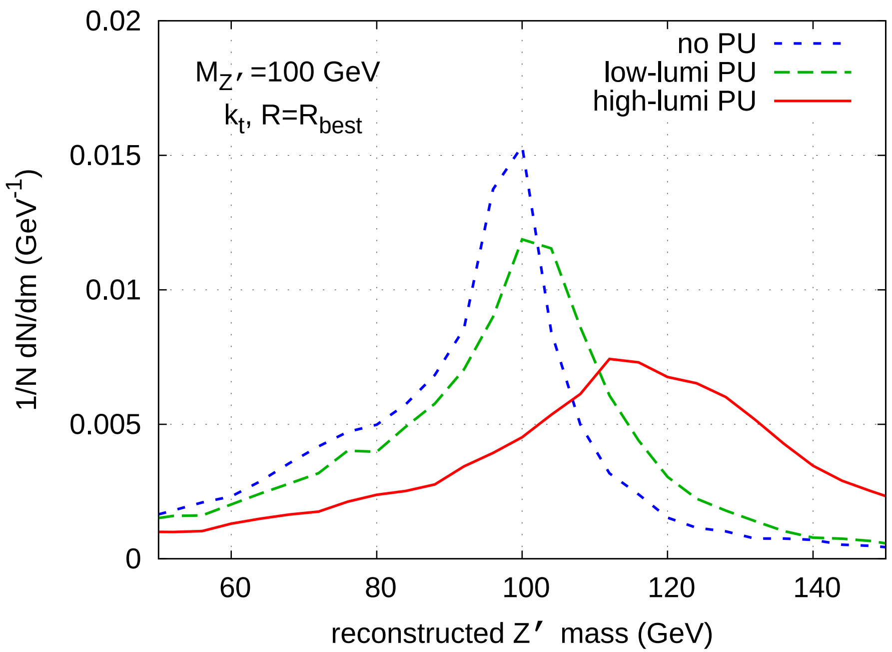
Part 2
Jets areas
Dealing with the background

[M.Cacciari,G.Salam,G.S., 07-08]

Need for subtraction

Pileup ≈ uniform soft background that shifts jets to higher p_t

Example: reconstructions of a 100 GeV Z' through $Z' \to q\bar{q} \to$ 2 jets



low-lumi = $2\ 10^{33}$ cm^{-1} s^{-1}
high-lumi = 10^{34} cm^{-1} s^{-1}

- peak shifted (because of pileup particles adding into the jets
- peak smeared (due to pileup fluctuations)

Area definition

- Idea: add soft particles (ghosts) and look in which jets they are caught

jet area = region where it catches ghosts

Area definition

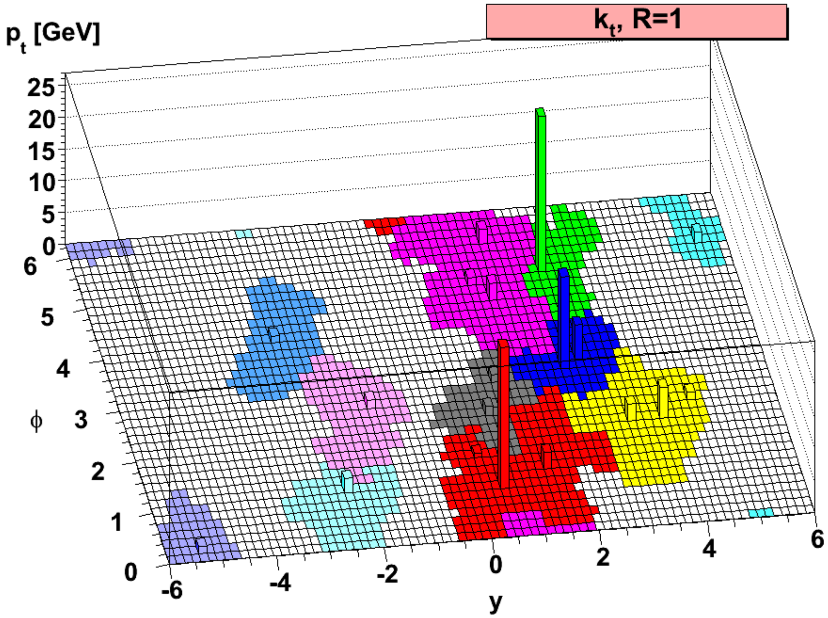
- Idea: add soft particles (ghosts) and look in which jets they are caught

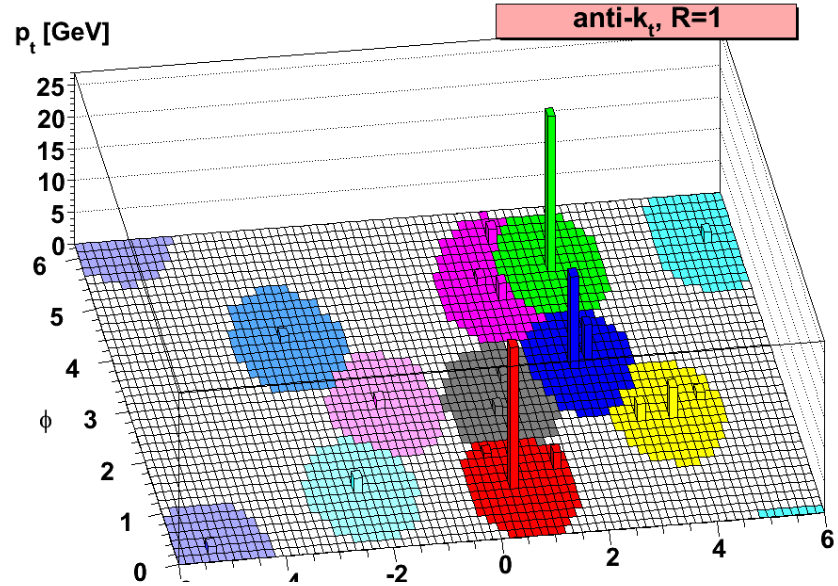
jet area = region where it catches ghosts

- 2 definitions
  - Passive area
    add one ghost and look where it ends. repeat to cover the (y, ϕ) plane
    mimics pointlike background
  - Active area
    add a large amount of ghosts and cluster everything
    mimics uniform background
    also gives purely ghosted jets
- Both practical and tractable analytically

Examples

BROOKHAVEN NATIONAL LABORATORY



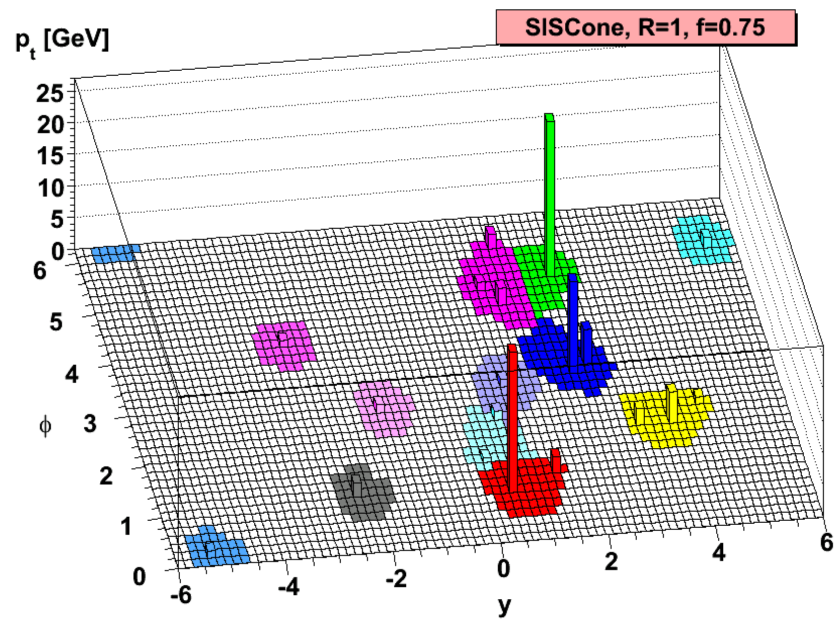


Notes: (active area)

- k_t: irregular “boundaries”
- anti-k_t: regular “boundaries”
- SISCone: smaller area

  risks of *monster jets* with $f = 0.5$

  $\Rightarrow$ use $f = 0.75$



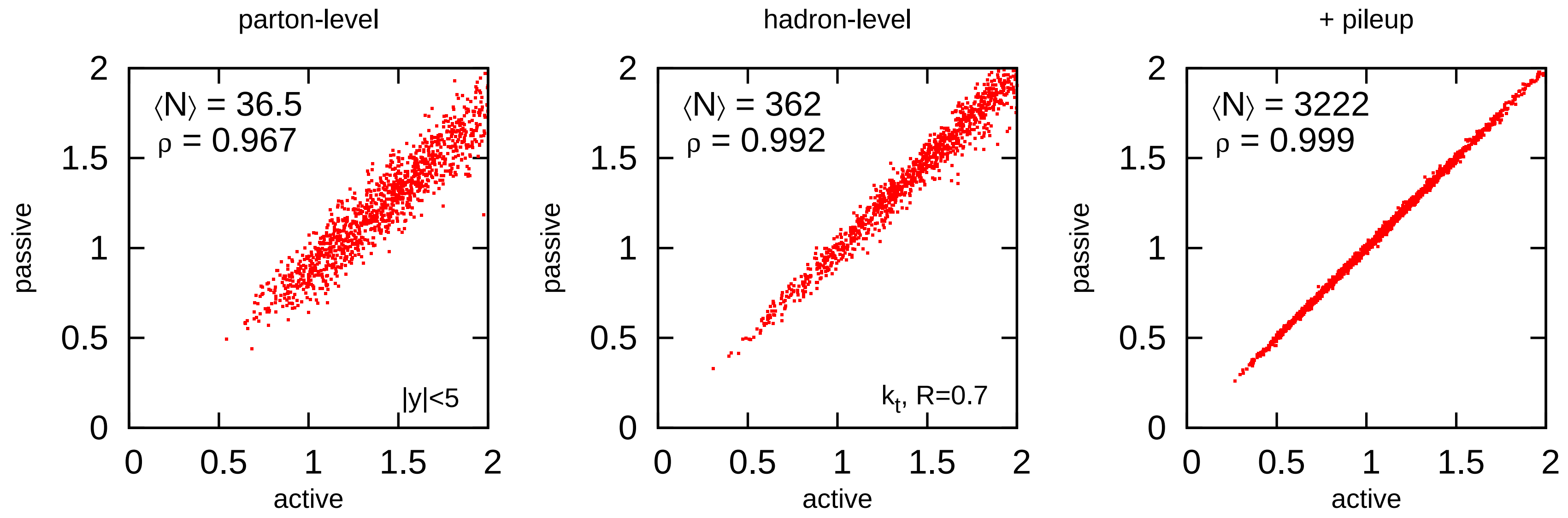
Density dependence

Dilute event:

active $\neq$ passive

Dense event

active $\simeq$ passive



Area scaling violations

Tractable analytically

Example: Average jet area (0^{th} order: 1 particle, 1^{st} order: 2 particles) :

$$\begin{aligned}\langle \mathcal{A}(p_{t,1}, R)\rangle &= \mathcal{A}_{1\text{hard}}(R) + \int d\Delta\, dp_{t,2} \frac{dP}{d\Delta_{12} dp_{t,2}} \left[\mathcal{A}_{\text{hard+1 soft}}(\Delta, R) - \pi R^2\right] \\ &= \mathcal{A}_{1\text{hard}}(R) + \frac{C_{F,A}}{\pi b_0} \log\left(\frac{\alpha_s(Q_0)}{\alpha_s(R p_t)}\right) \pi R^2\, d\end{aligned}$$

- Scaling violation (area $\neq \pi R^2$)
- gluon > quark
- know LO anomalous dimension

	passive		active	
	$\frac{\mathcal{A}_1}{\pi R^2}$	d	$\frac{\mathcal{A}_1}{\pi R^2}$	d
k_t	1	0.5638	0.81	0.519
Cam	1	0.07918	0.81	0.0865
SISCone	1	-0.06378	1/4	0.1246
anti-k_t	1	0	1	0

Area scaling violations

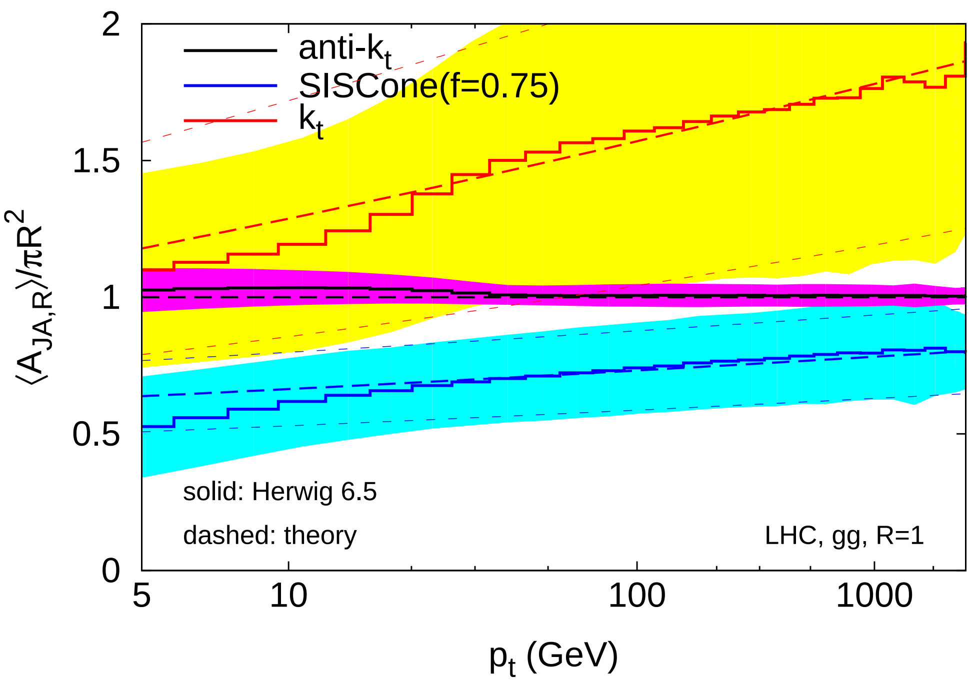
Tractable analytically

Example: Average jet area ($0^{\rm th}$ order: 1 particle, $1^{\rm st}$ order: 2 particles) :

$$\begin{aligned}\langle \mathcal{A}(p_{t,1}, R)\rangle &= \mathcal{A}_{1\text{hard}}(R) + \int d\Delta\, dp_{t,2}\, \frac{dP}{d\Delta_{12} dp_{t,2}} \left[\mathcal{A}_{\text{hard+1 soft}}(\Delta, R) - \pi R^2\right] \\ &= \mathcal{A}_{1\text{hard}}(R) + \frac{C_{F,A}}{\pi b_0} \log\left(\frac{\alpha_s(Q_0)}{\alpha_s(Rp_t)}\right) \pi R^2\, d\end{aligned}$$

Note the Q_0 in the scaling violations:

- $Q_0 \equiv$ IR cut-off
- vacuum $Q_0 \propto \Lambda_{QCD}$,
  with background $Q_0 \propto \rho_{\rm bkg}$
- good agreement with MC



Pileup subtraction

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

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

Pileup subtraction

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

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

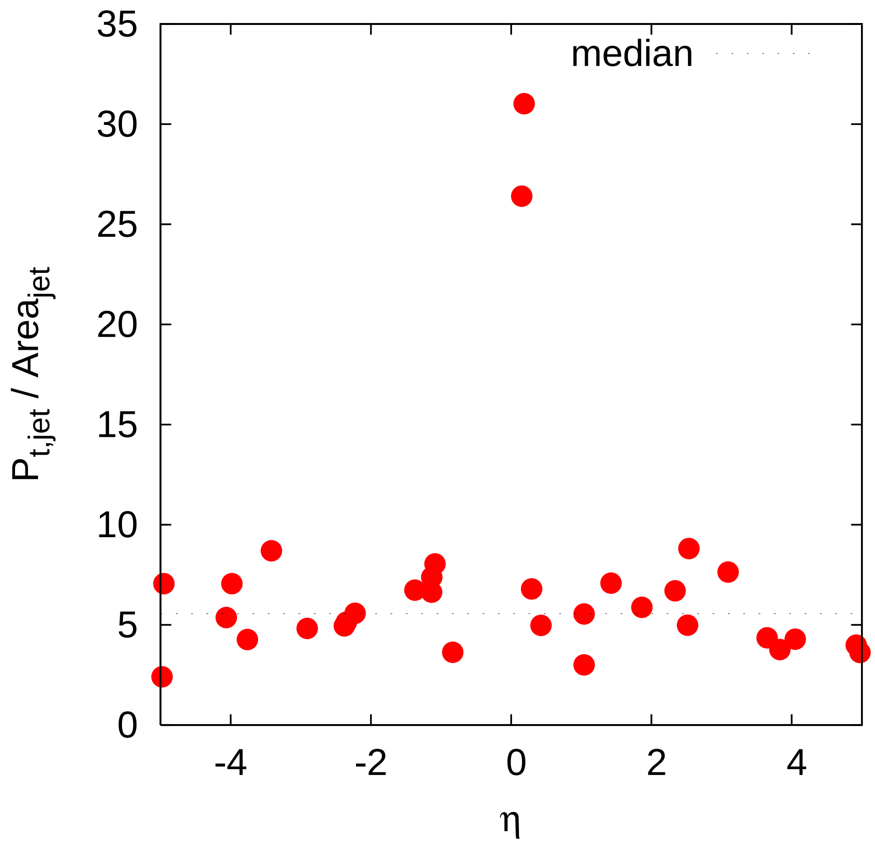
- Jet area: [M.Cacciari, G.Salam, G.S., 08]
  - region where the jet catches infinitely soft particles (active/passive)
  - tractable analytically in pQCD

Pileup subtraction

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)
  - tractable analytically in pQCD

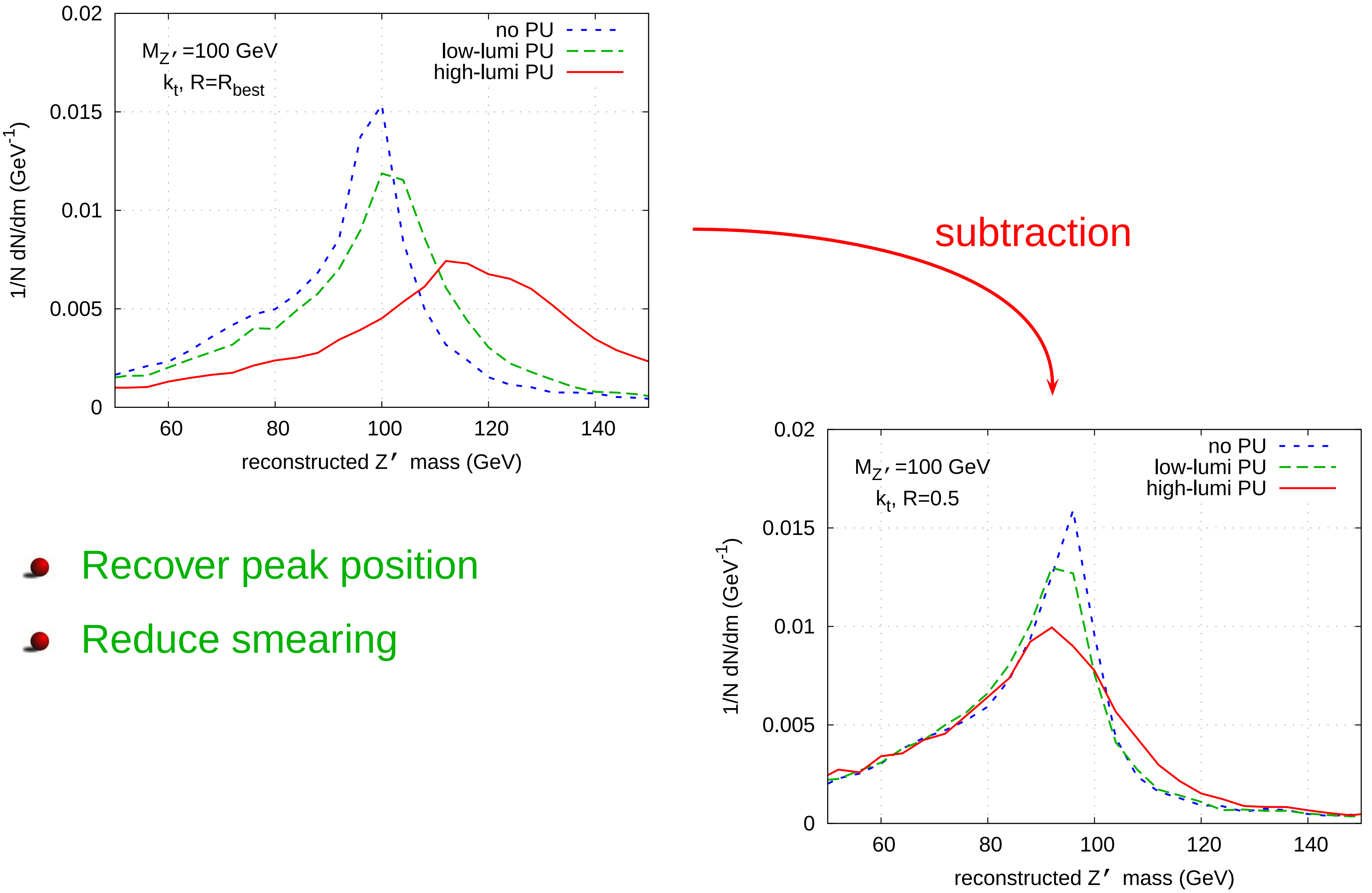


- Pileup density per unit area: ρ_{pileup}

  e.g. estimated from the median of $p_{t,\text{jet}}/\text{Area}_{\text{jet}}$

implemented in FastJet
on an event-by-event basis

Subtraction at work

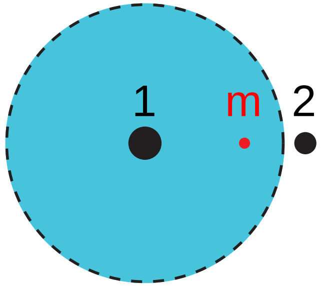


- Recover peak position
- Reduce smearing

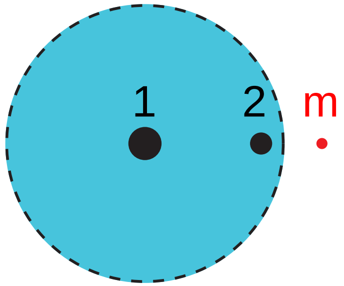
Back-reaction

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: p_2 gained when adding p_m



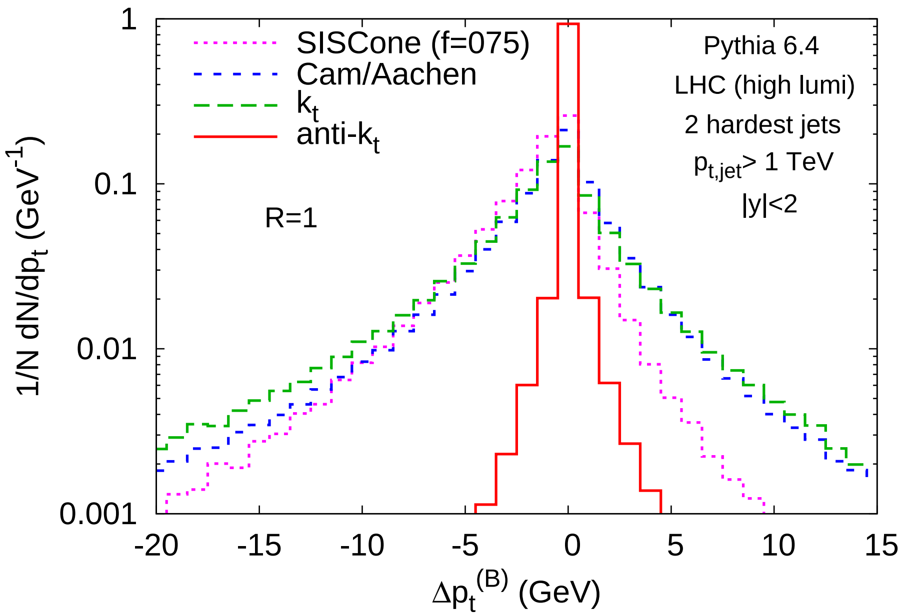
  - loss: p_2 lost when adding p_m



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



Conclusions

- Use acceptable jet algorithms
  - 4 (fast-enough) IRC-safe algorithms:
    k_t, Cambridge/Aachen, anti-k_t and SISCone
  - IRC safety will matter at the LHC

- Definition of “the *area* of a jet”
  - defined as the region where it catched soft background
  - tractable analytically
  - can be use to subtract soft-background contamination
  - same technique holds for heavy-ion background
    some refinements are under study