The cookbook for jets in heavy-ion collisions?

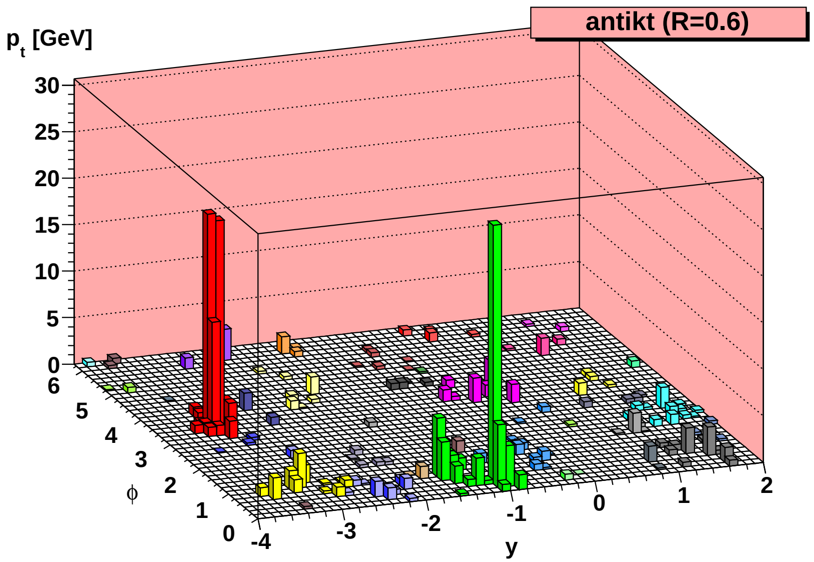
Grégory Soyez

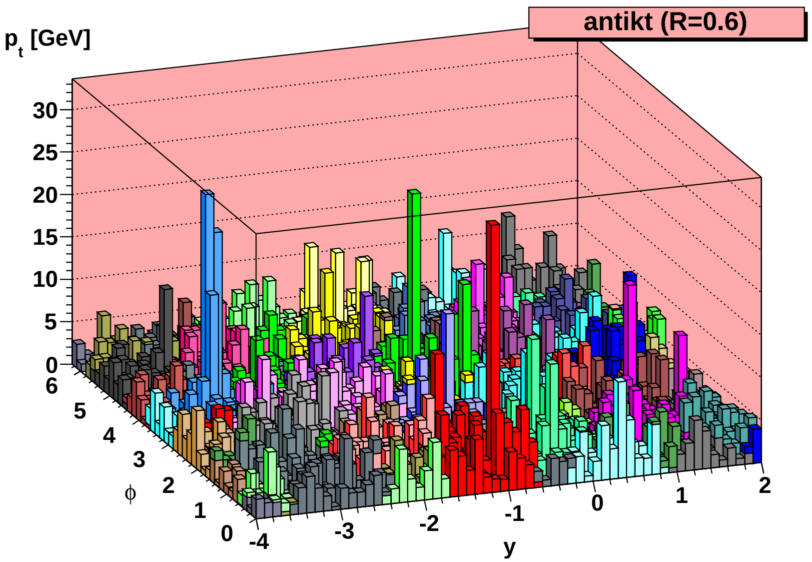
IPhT, CEA Saclay

In collaboration with Gavin Salam, Matteo Cacciari and Juan Rojo

Heavy Ion meeting — IPN, Orsay — Nov 26 2010

How to “see” jets in a soft background





Valid for many backgrounds

- UE in pp (~ 1 GeV)
- pileup in pp (~ 10 GeV)
- UE in AA (~ 100 GeV)

(Hopefully) for everyone

- Standard method
- New hints
- comments for experts

Central formula

One basic formula for **background subtraction for a single event**

$$p_{t,\mathrm{jet}}^{(\mathrm{sub})} = p_{t,\mathrm{jet}} - \rho_{\mathrm{bkg}} A_{\mathrm{jet}}$$

assumes that the background is uniform

3 things needed:

- Define a jet
- Define the area of a jet
- Obtain ρ_{bkg}, the background p_t density per unit area

[Cacciari, Salam, 07]

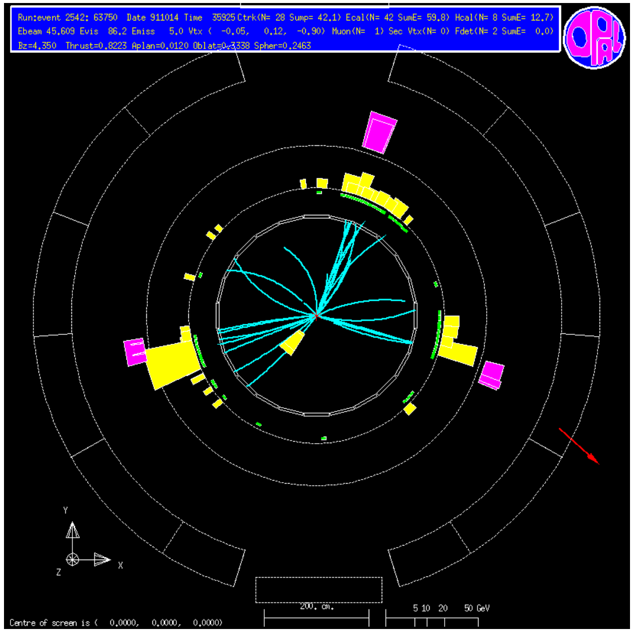
$$p_{t,\text{jet}}^{(\text{sub})} = p_{t,\text{jet}} - \rho_{\text{bkg}} A_{\text{jet}}$$

3 things needed:

- Define a jet
- Define the area of a jet
- Obtain ρ_{bkg}, the background p_t density per unit area

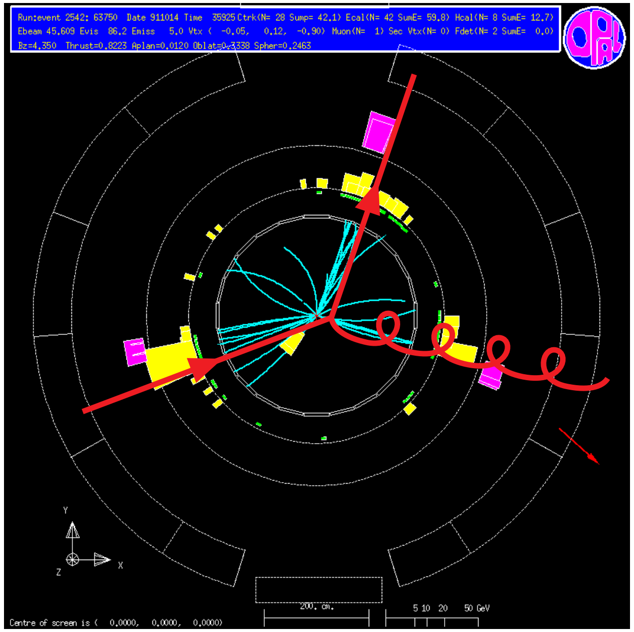
Jet definitions

"Jets" $\equiv$ bunch of collimated particles $\cong$ hard partons



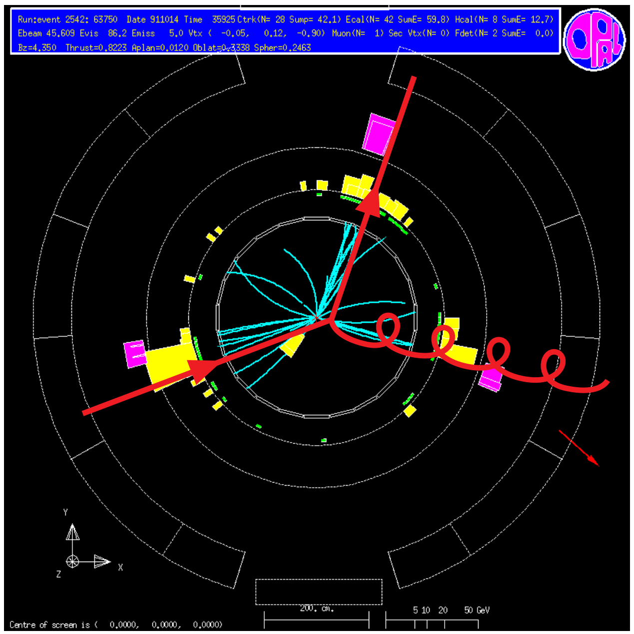
Jet definitions

"Jets" $\equiv$ bunch of collimated particles $\cong$ hard partons



Jet definitions

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



In practice: use a <u>jet definition</u>

$$\text{particles } \{p_i\} \xrightarrow[\textbf{definition}]{\textbf{jet}} \text{jets } \{j_k\}$$

algorithm: the recipe (insufficient!)

definition: algorithm + params

Jet=hadron is too simplistic: NLO? What opening for “collimated”?

Examples of jet definitions

- Recombination: successively recombine the closest pair

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

  Stop at distance R

  - $p = 1$: k_t algorithm (very close to QCD)

    [Catani, Dokshitzer, Seymour, Webber, 93]

  - $p = 0$: Cambridge/Aachen (C/A) algorithm (substructure studies)

    [Dokshitzer, Leder, Moretti, Webber, 93]

  - $p = -1$: anti-k_t algorithm (the default at the LHC)

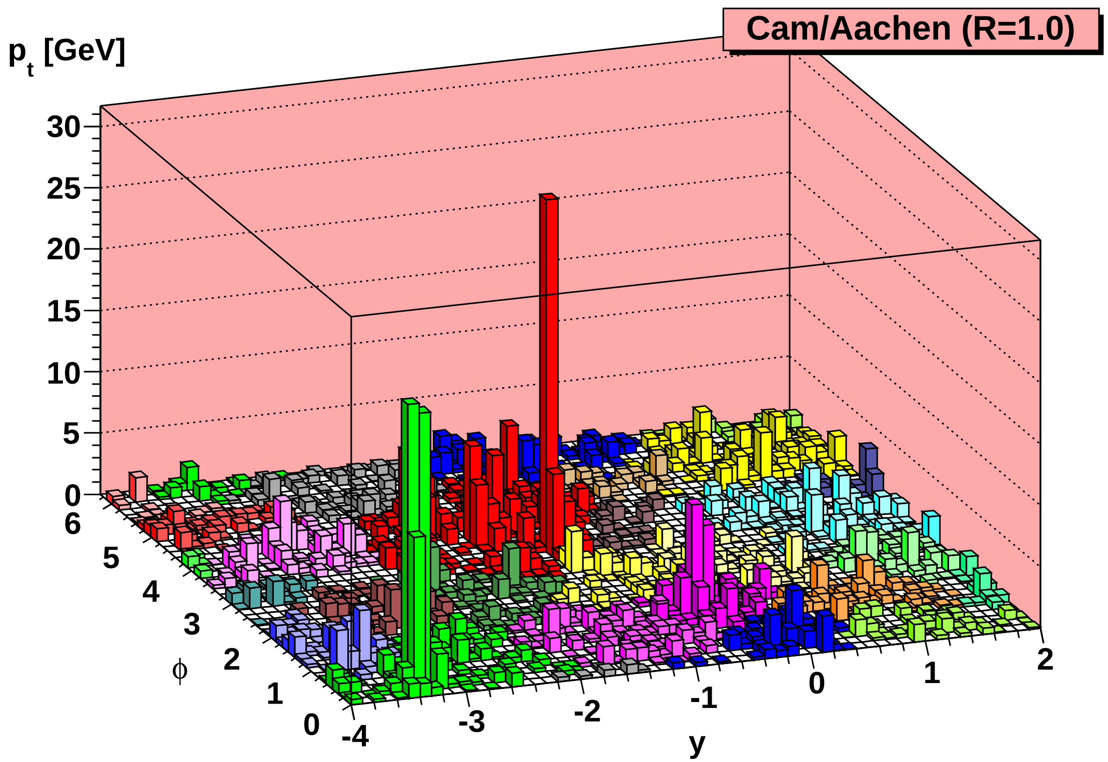
    [Cacciari, Salam, GS, 08]

- Cone: $\approx$ flow of energy in a cone (of fixed R) centred on the cone centre: SISCone

  [Salam, GS, 07]

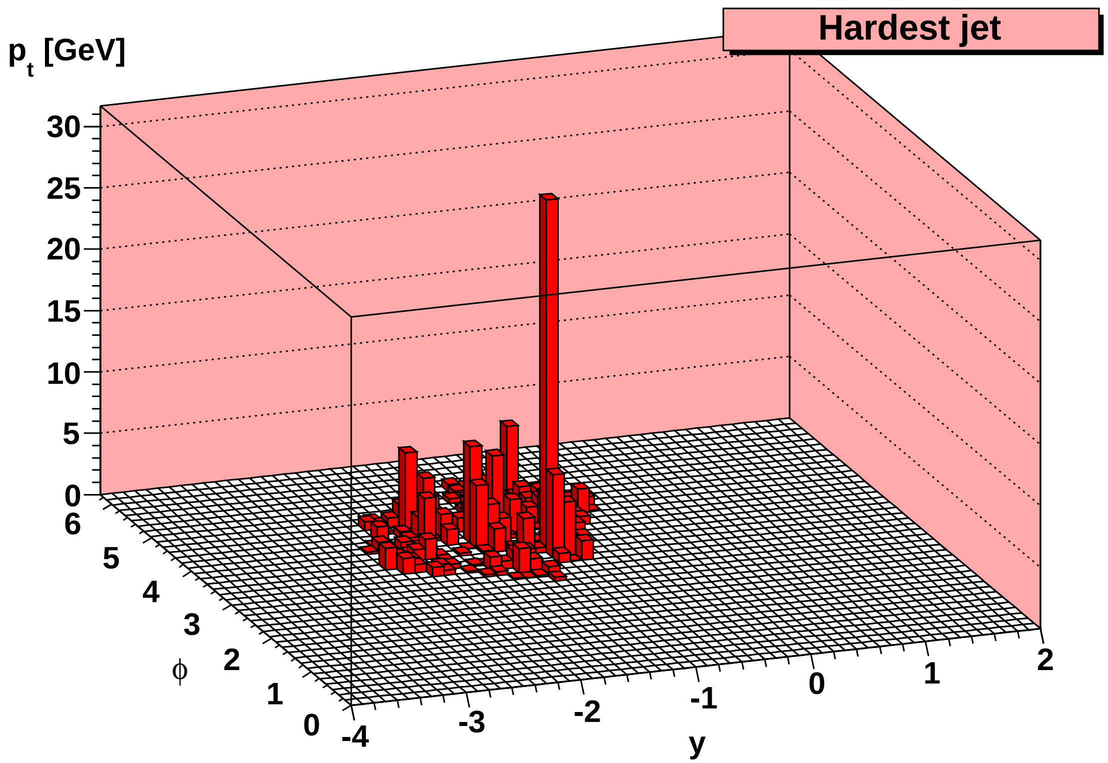
Finite perturbative cross-section: only consider infrared-and-collinear-safe algorithms

New suggestion #1: Filtering



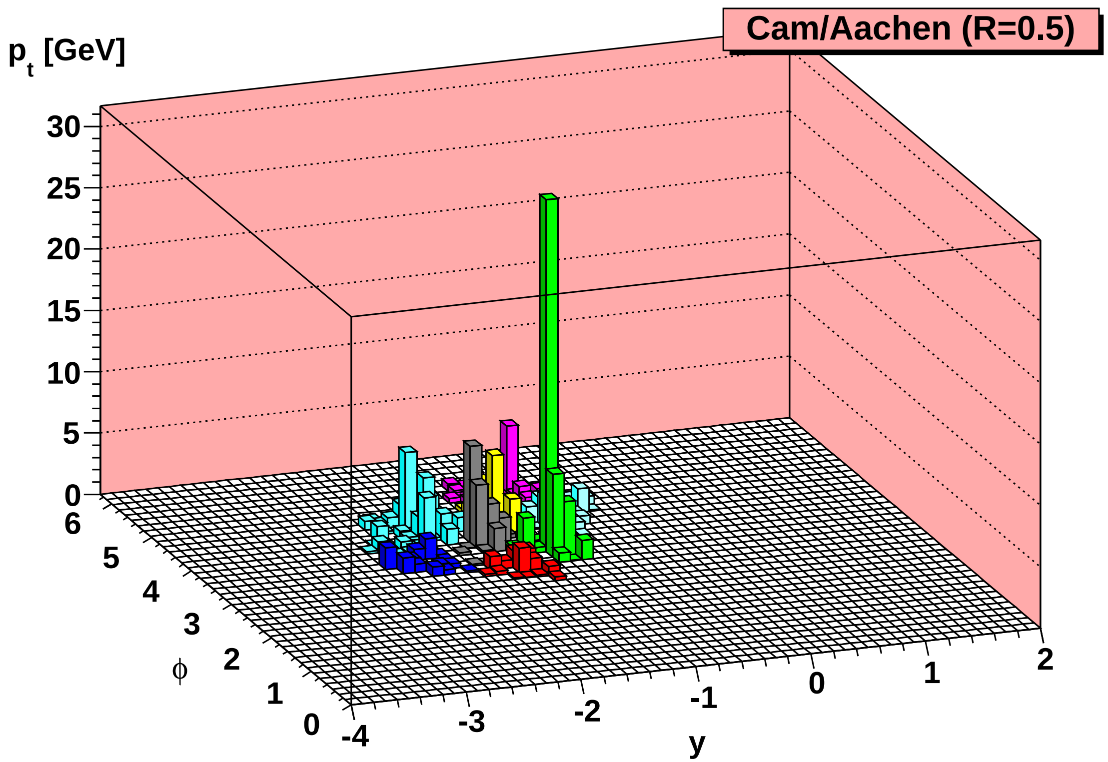
- cluster with Cambridge/Aachen(R)

New suggestion #1: Filtering



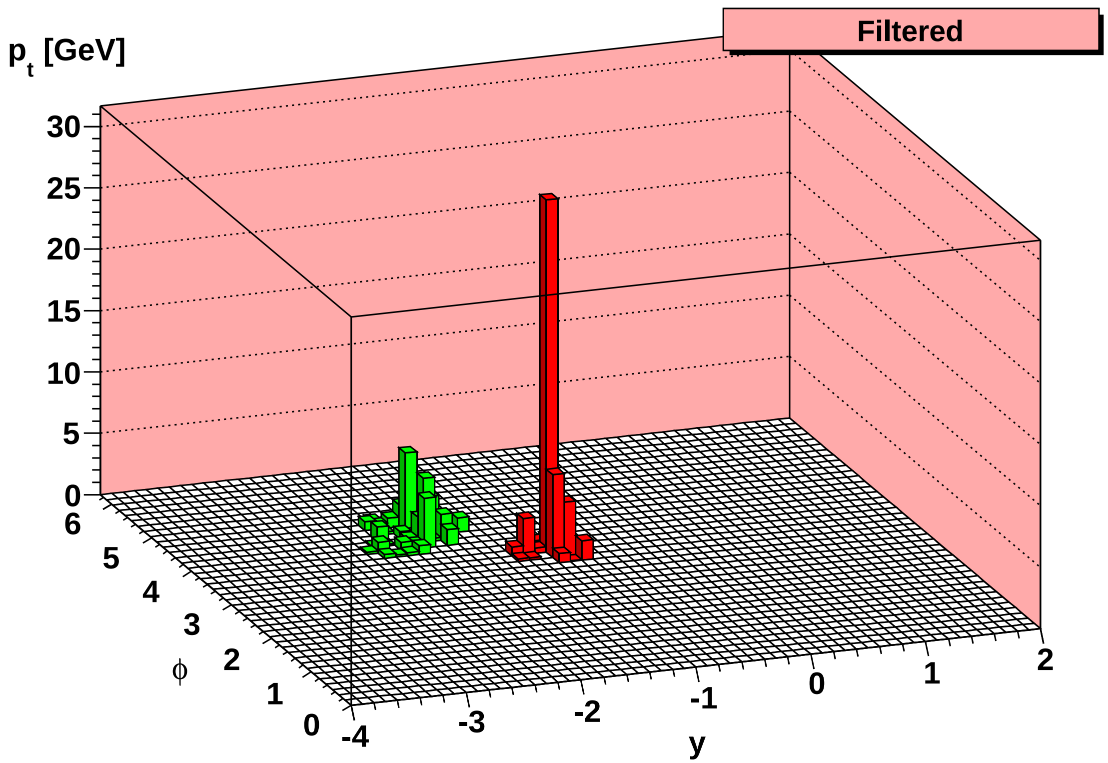
- cluster with Cambridge/Aachen(R)
- for each jet

New suggestion #1: Filtering



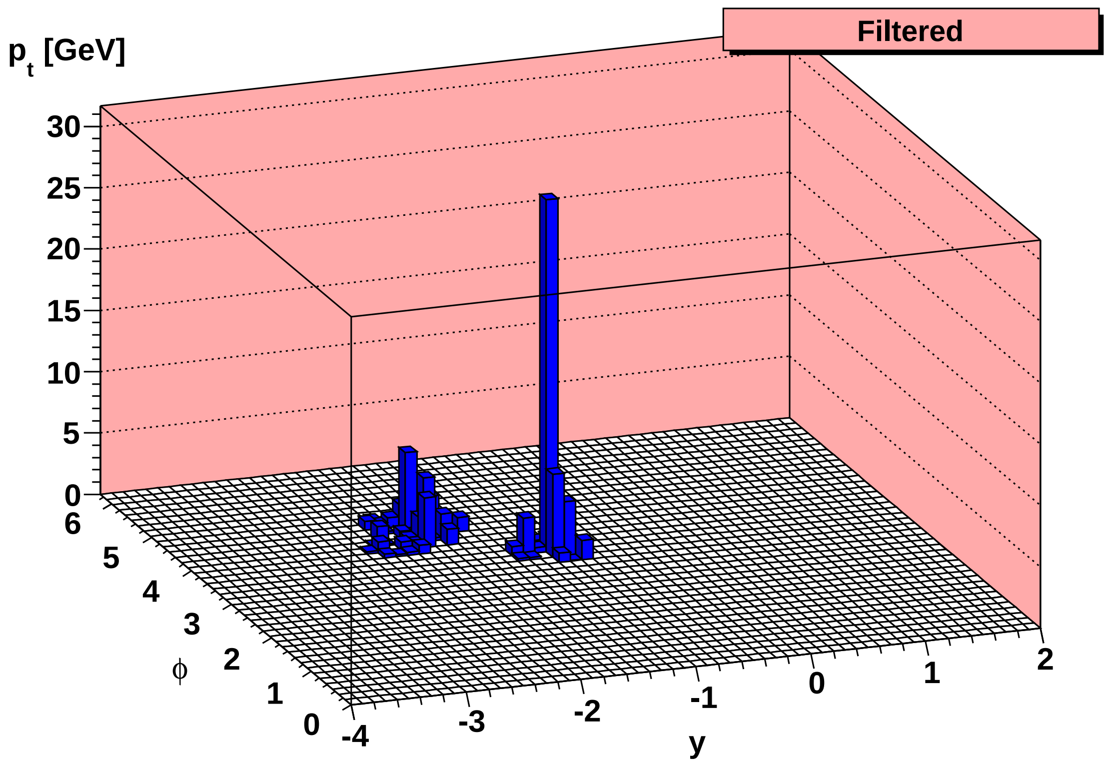
- cluster with Cambridge/Aachen(R)
- for each jet
  - recluster with Cambridge/Aachen(R/2)

New suggestion #1: Filtering



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

New suggestion #1: Filtering



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

Idea:

✓ keep perturb. radiation

✓ remove UE

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

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

- Proven useful for kinematic reconstructions

  [M.Cacciari, J.Rojo, G.Salam, GS, 08]

$$p_{t,\mathrm{jet}}^{(\mathrm{sub})} = p_{t,\mathrm{jet}} - \rho_{\mathrm{bkg}} A_{\mathrm{jet}}$$

3 things needed:

- Define a jet
- Define the area of a jet
- Obtain ρ_{bkg}, the background p_t density per unit area

Area definitions

[M.Cacciari, G.Salam, GS, 08]

Area $\equiv$ region where the jet catches soft particles

- Recipe: add a dense coverage of infinitely soft particles (**ghosts**)

  (active) area = region where a jet catches the ghosts

- Idea: ghost $\approx$ background particle

  $\Rightarrow$ area where catching ghost $\equiv$ area where catching background

- Advantages:
  - generic/universal definition (*e.g.* independent of a calorimeter)
  - allow for analytic computations

- Notes for experts:
  - put ghosts up to at least $y_{\rm jet,max} + R$
  - preferably use a “4-vector” definition of the area (sum ghost 4-momenta)
  - require an IRC-safe algorithm!
  - alternative: passive area (equivalent for large multiplicities)
  - Better handling with active_area_explicit_ghosts

$$p_{t,\text{jet}}^{(\text{sub})} = p_{t,\text{jet}} - \rho_{\text{bkg}} A_{\text{jet}}$$

3 things needed:

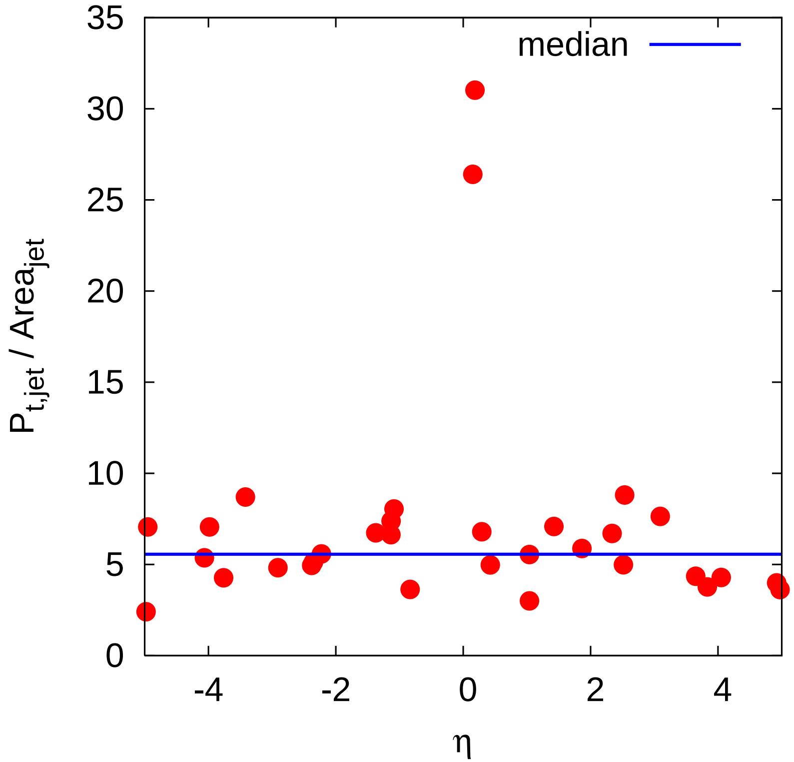
- Define a jet
- Define the area of a jet
- Obtain ρ_{bkg}, the background p_t density per unit area

Example: $\rho_{\rm bkg}$ from jets

Recipe for estimating $\rho_{\rm bkg}$:

- Cluster with k_t of C/A with "radius" R_ρ
- Estimate $\rho_{\rm bkg}$ using

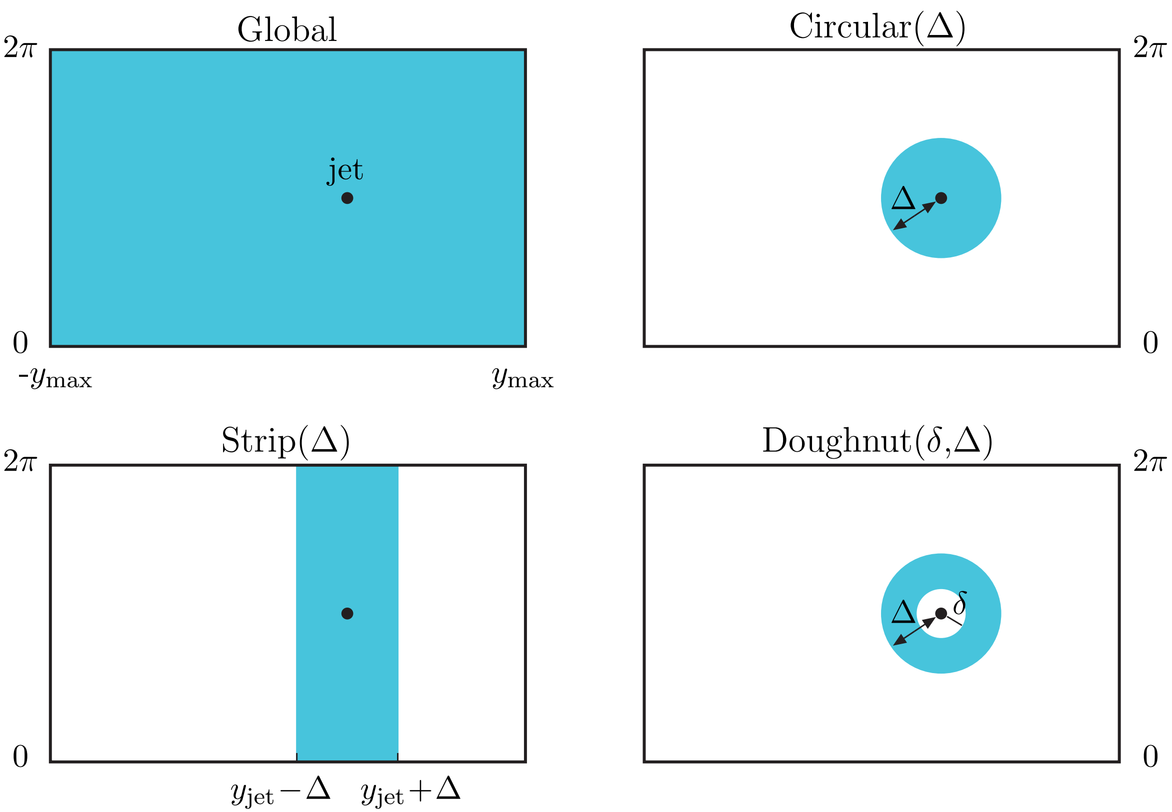
$$\rho_{\rm bkg} = \underset{j\in \text{jets}}{\text{median}} \left\{ \frac{p_{t,j}}{A_j} \right\}$$



- Notes for experts
  - Other algorithms produce unwanted jets with small area
  - Typically, R_ρ between 0.3 and 0.6 is OK (I'll take 0.5)

New suggestion #2: Use a local range

Fluctuating background (*e.g.* rapidity dependence) $\rightarrow$ **local range**



$$\rho_{\text{bkg}}(j) = \\ = \underset{j' \in \mathcal{R}(j)}{\text{median}} \left\{ \frac{p_{t,j'}}{A_{j'}} \right\}$$

Also:
exclude the n (typically 2) hardest jets in the event

Notes for experts:

- Limited acceptance $\equiv$ local range
- Put ghosts at least up to $|y_{\text{jet,max}}| + \Delta + R$...
- ... but for limited acceptance: put ghosts over the acceptance and keep all jets (tbc)

Subtraction uncertainties

- Background fluctuations: (inside an event!)

$$p_{t,\mathrm{jet}} = p_{t,\mathrm{jet}}^{\mathrm{hard}} + \rho_{\mathrm{bkg}} A_{\mathrm{jet}} \pm \sigma_{\mathrm{bkg}} \sqrt{A_{\mathrm{jet}}}$$

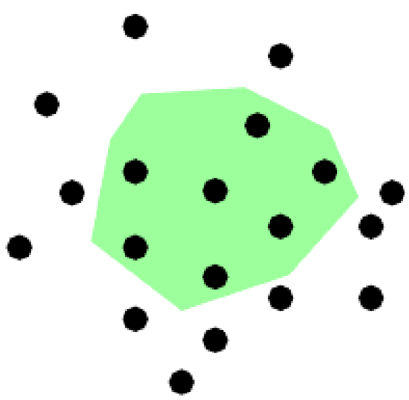
Subtraction uncertainties

- Background fluctuations: (inside an event!)

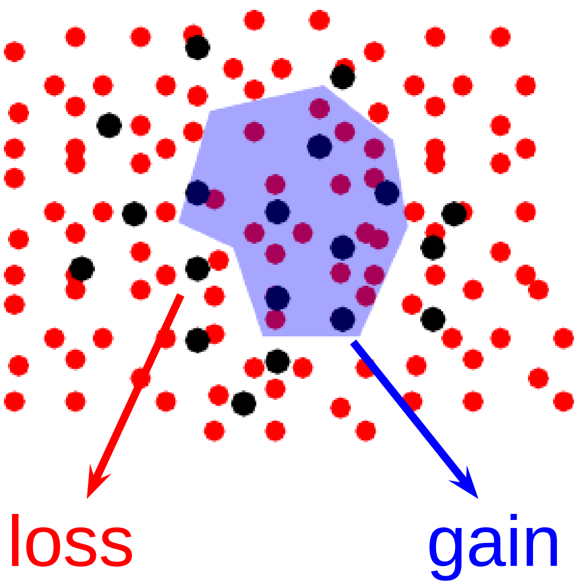
$$p_{t,\mathrm{jet}} = p_{t,\mathrm{jet}}^{\mathrm{hard}} + \rho_{\mathrm{bkg}} A_{\mathrm{jet}} \pm \sigma_{\mathrm{bkg}} \sqrt{A_{\mathrm{jet}}}$$

- Back-reaction:

No background



With background

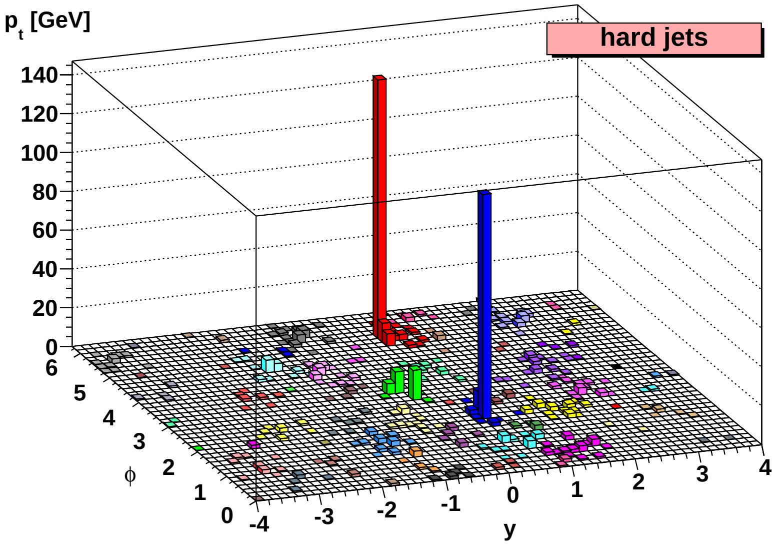


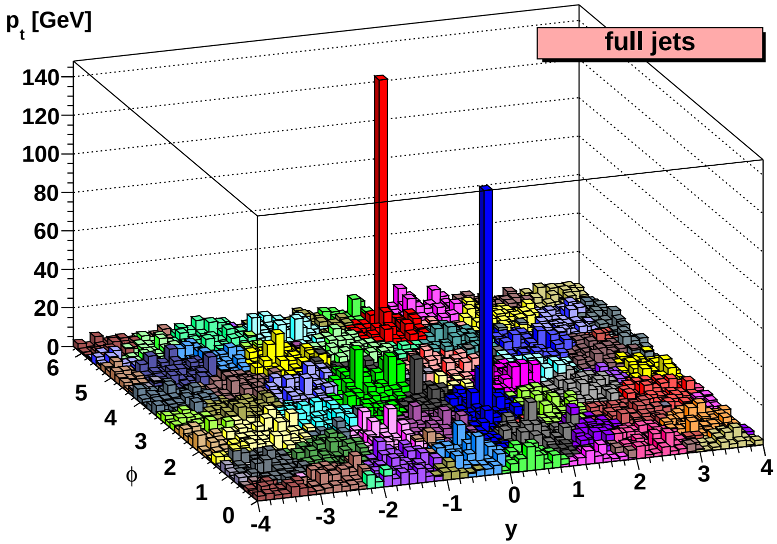
Subtraction efficiency: what precision may we hope for?

[Cacciari, Rojo, Salam, GS, 1010.1759]

Framework for study

Hard (Pythia) event | ... embedded in AA (Hydjet) event



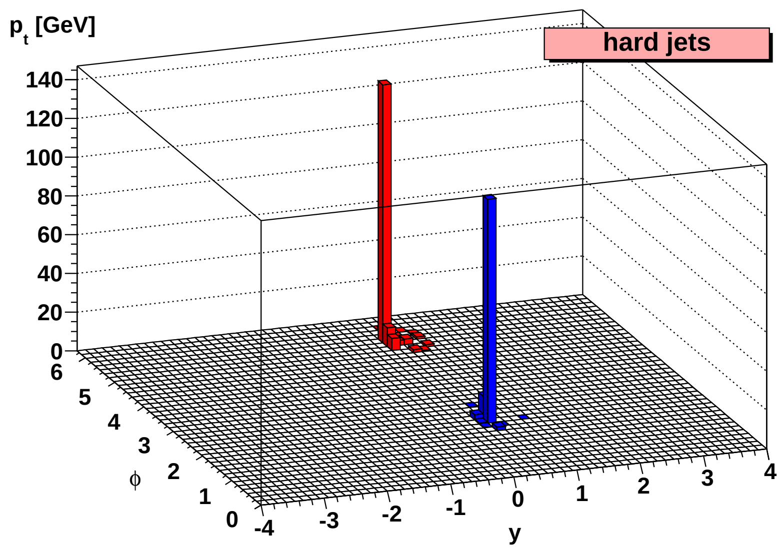


- Get the jets and apply subtraction in both cases ($R = 0.4$)

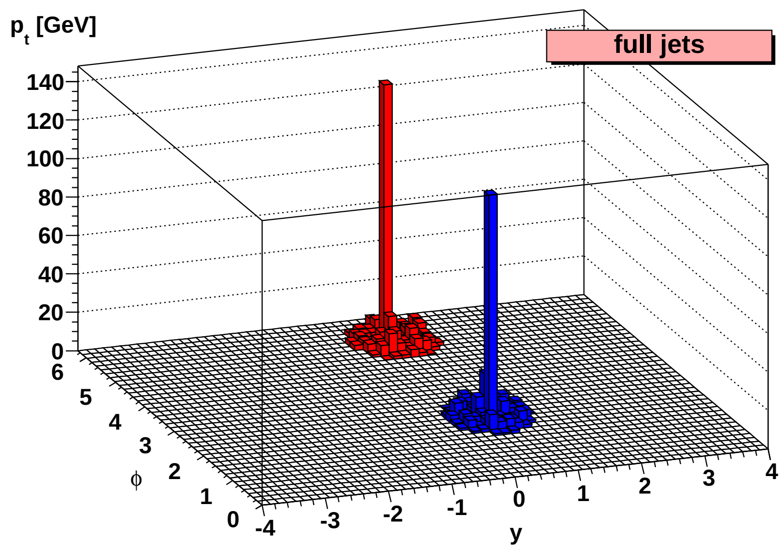
  test different algs. and ρ-estimation ranges

Framework for study

Hard (Pythia) event



... embedded in AA (Hydjet) event



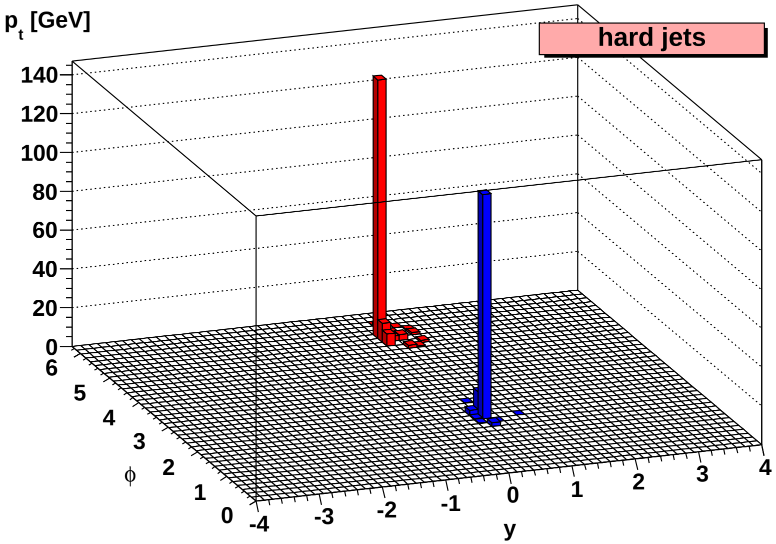
- Get the jets and apply subtraction in both cases ($R = 0.4$)
- Match the 2 hardest jets

at least 50% of the hard contents recovered after embedding

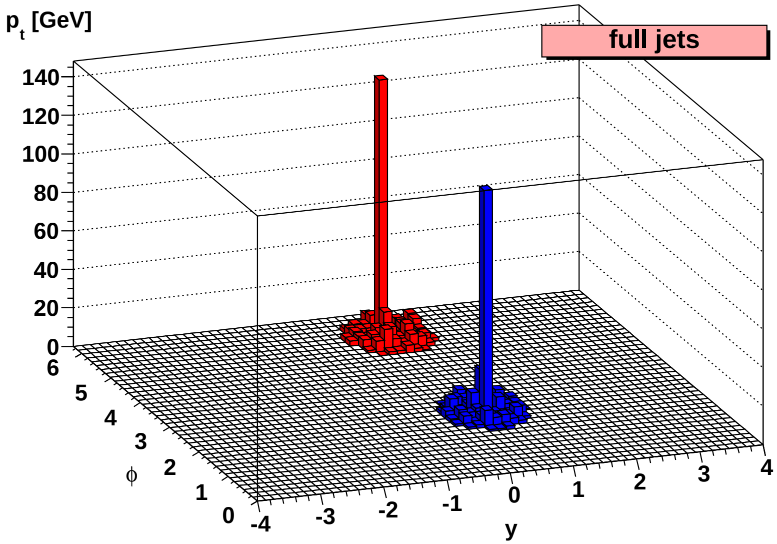
Efficiencies $\geq$ 95%

Framework for study

Hard (Pythia) event



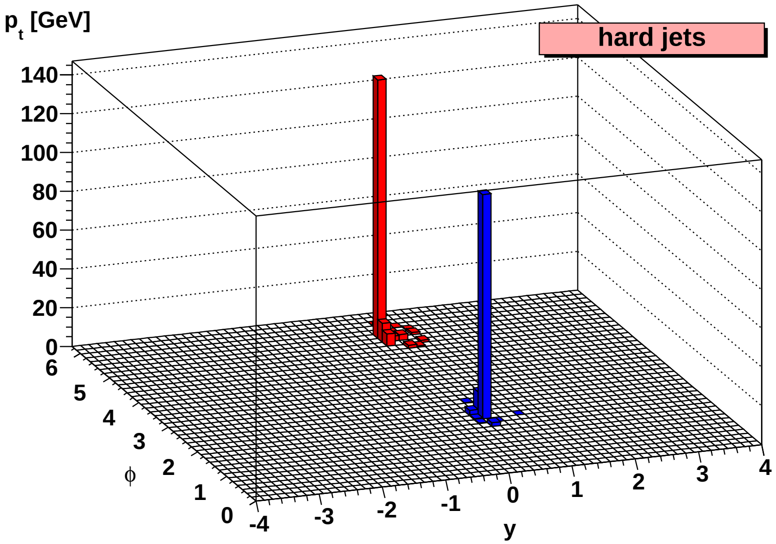
... embedded in AA (Hydjet) event



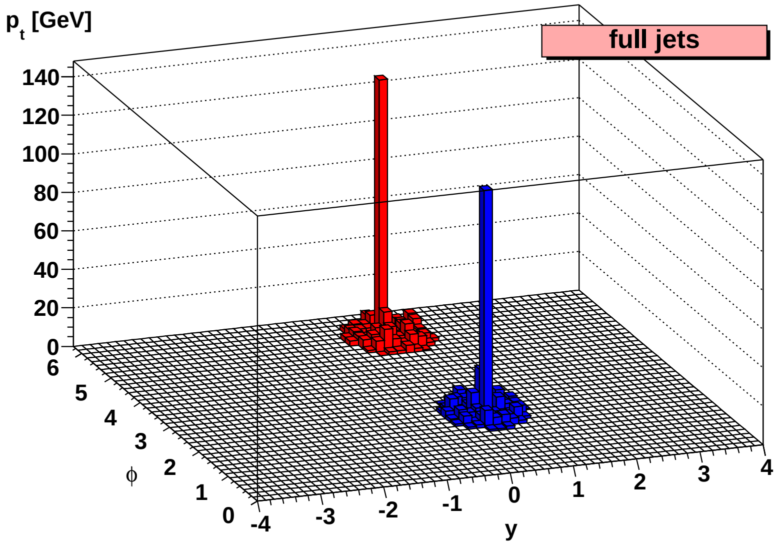
- Get the jets and apply subtraction in both cases ($R = 0.4$)
- Match the 2 hardest jets
- Subtraction quality: $\Delta p_t = p_t^{AA,\mathrm{sub}} - p_t^{pp,\mathrm{sub}}$
- Study $\langle \Delta p_t \rangle$ and $\sigma_{\Delta p_t}$

Framework for study

Hard (Pythia) event



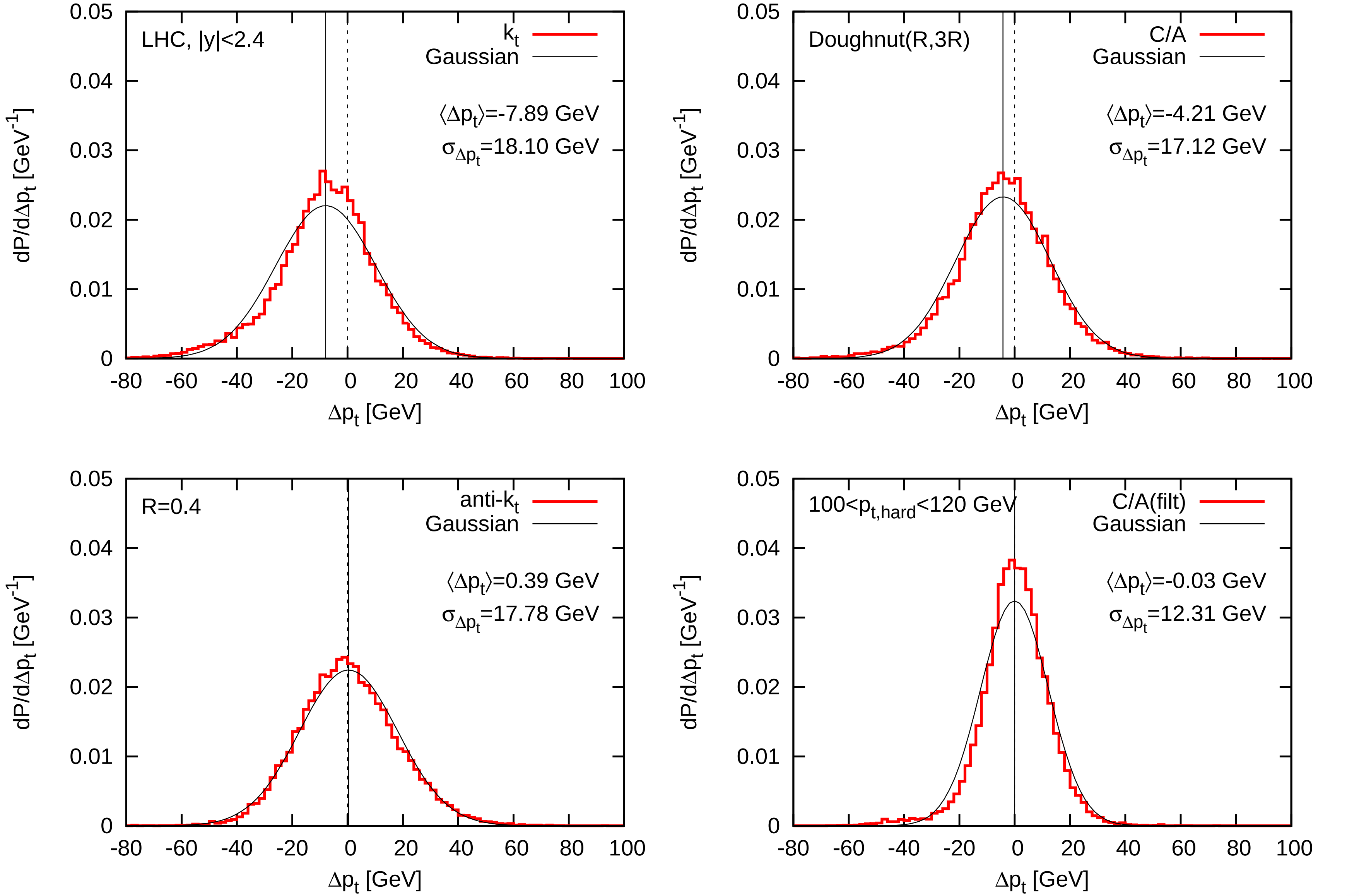
... embedded in AA (Hydjet) event



- Get the jets and apply subtraction in both cases ($R = 0.4$)
- Match the 2 hardest jets
- Subtraction quality: $\Delta p_t = p_t^{AA,\mathrm{sub}} - p_t^{pp,\mathrm{sub}}$
- Study $\langle \Delta p_t \rangle$ and $\sigma_{\Delta p_t}$
- Used: Pythia(v6.4), Hydjet(v1.6) + PyQuen(v1.5) + FastJet(v2.4.2)

  Cross-check: Q-Pythia, Hydjet++

Δp_t distributions



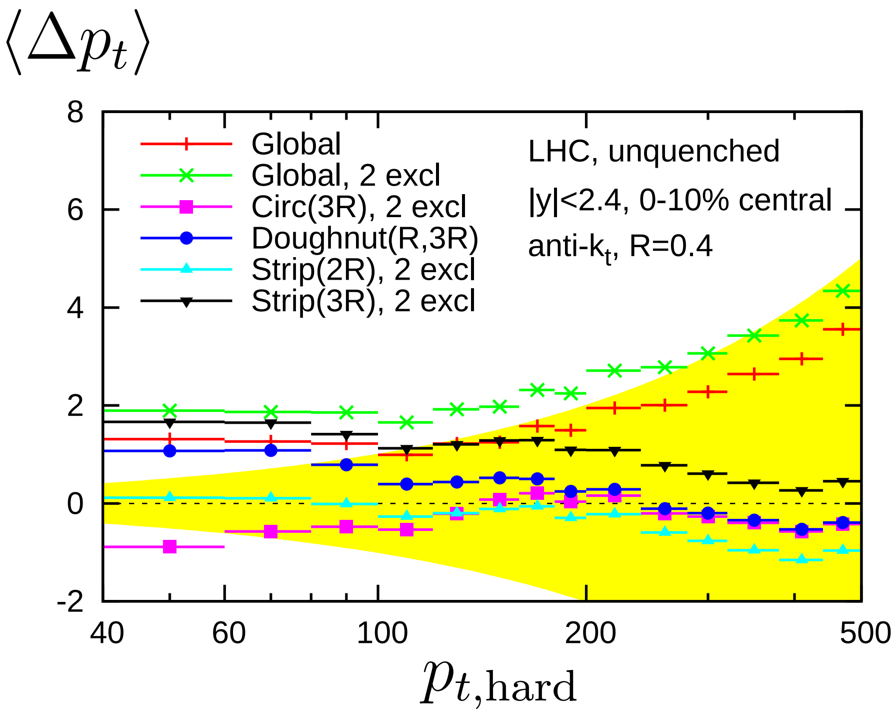
Gaussian enough for this talk: focus on $\langle \Delta p_t \rangle$ and $\sigma_{\Delta p_t}$

Effect of choosing a local range

Number of jets in a range

range	area	n_{jets}
Circ(2R)	$4\pi R^2$	4.5
Circ(3R)	$9\pi R^2$	10
Donut(R,2R)	$3\pi R^2$	3.5
Donut(R,3R)	$8\pi R^2$	9
Strip(2R)	$4\pi R$	11

$(R = 0.4, R_\rho = 0.5)$

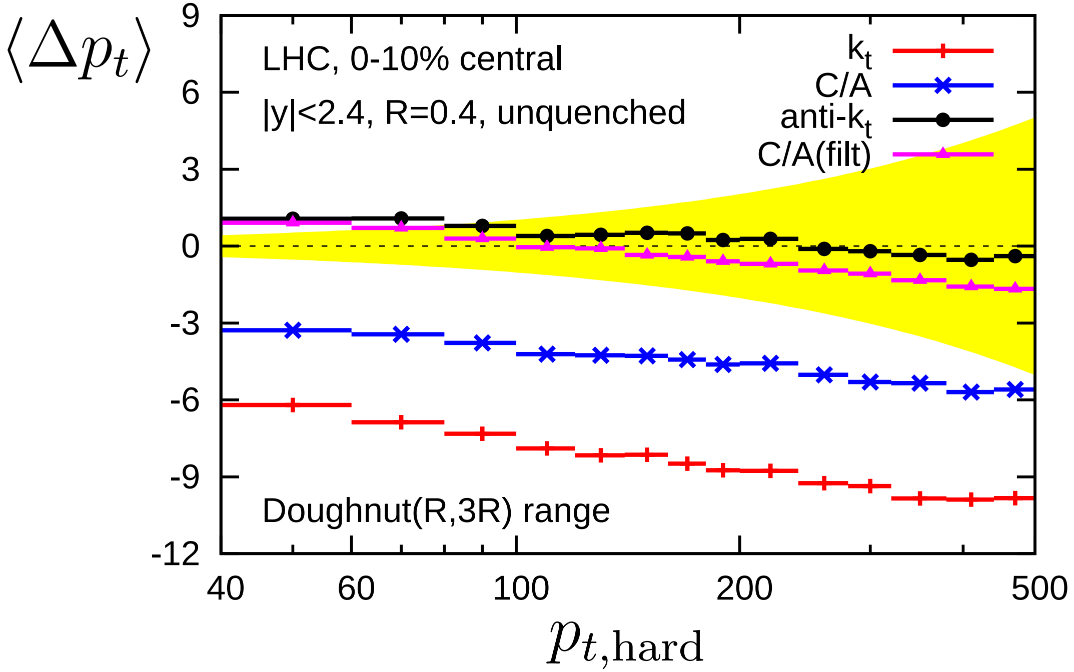


- rule of thumb: at least 8 jets needed to estimate ρ
- global range: OK with limited accept., worsen for larger
- different ranges $\longrightarrow$ estimate of the uncertainty

Note for experts: Analytic estimate show that at least 8 jets
$\Rightarrow$ less than 10% of $\sigma_{\Delta p_t}$ due to ρ misestimation

Differences between algorithms

- *Average shift*: preference for anti-k_t and C/A+filt$^{(*)}$



1-2 GeV precision for a contamination of $\sim$ 150 GeV!

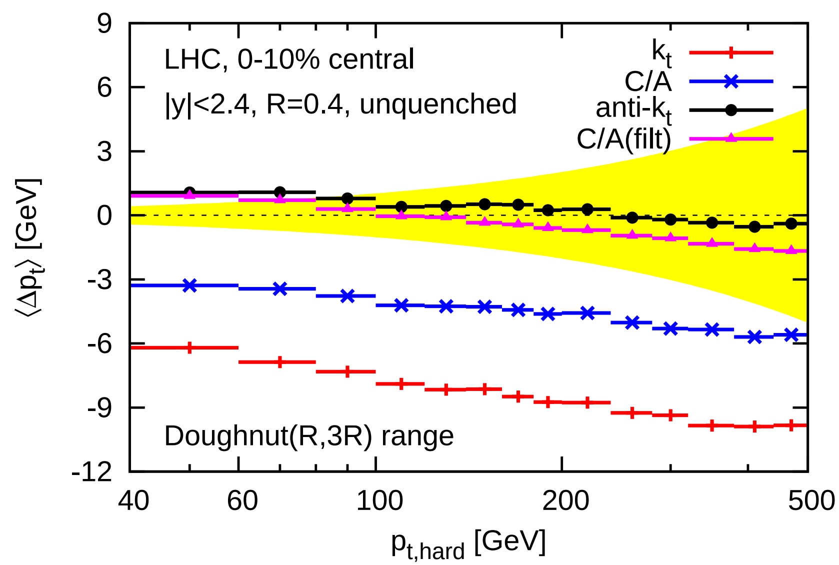
Notes for experts:

- C/A & k_t: offset due to *back-reaction*
- $^{(*)}$ C/A+filt: watch out: cancellation between back-reaction and filtering bias

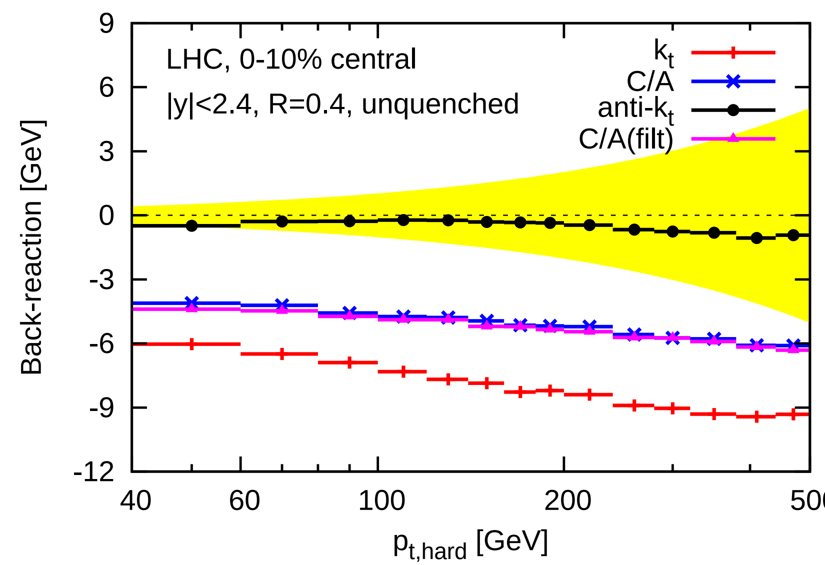
Differences between algorithms

- *Average shift*: preference for anti-k_t and C/A+filt

$\langle \Delta p_t \rangle$



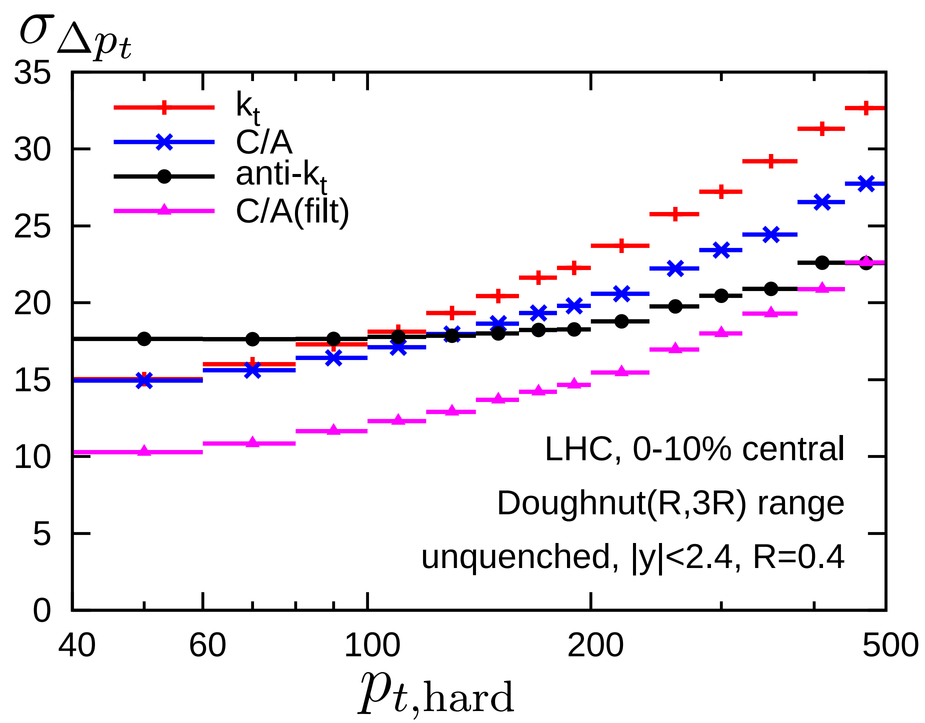
Back-reaction



  - diff. between k_t, C/A and anti-k_t from back-reaction
  - (fortuitous) cancellation with a filtering bias for C/A(filt)

Differences between algorithms

- *Average shift*: preference for anti-k_t and C/A+filt
- *Dispersion*: preference for C/A+filtering

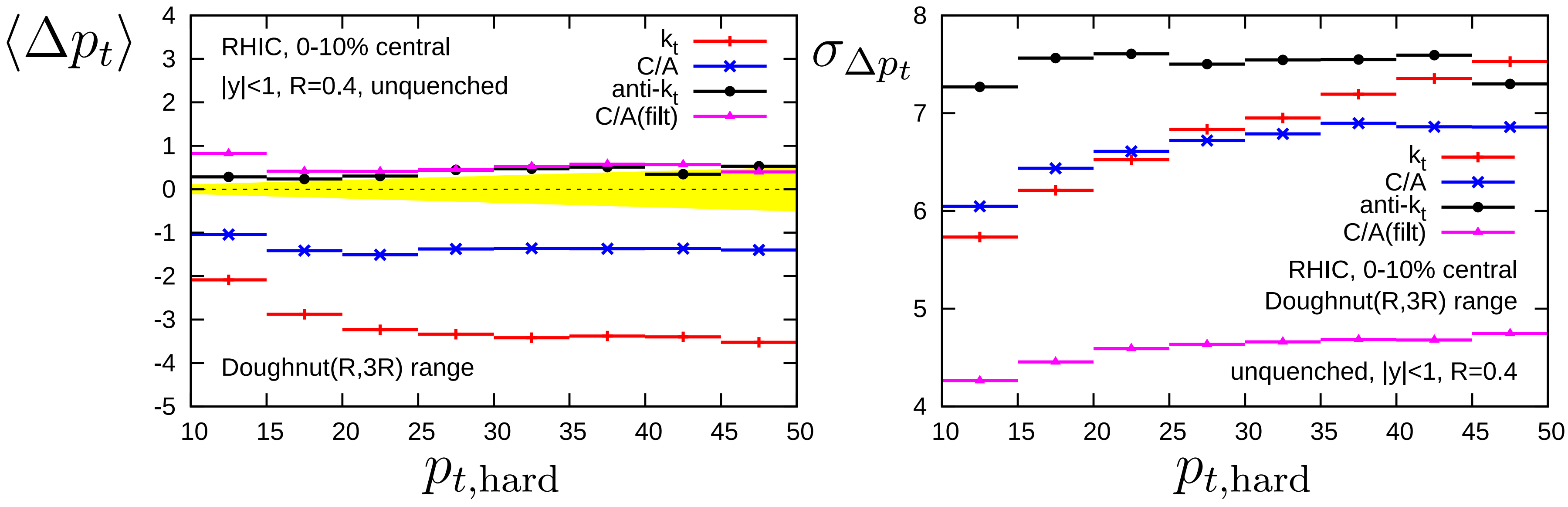


$$A_{\mathrm{C/A+filt}} \approx \frac{1}{2} A_{\mathrm{C/A}} \quad \Rightarrow \quad \sigma_{\Delta p_t}^{(\mathrm{C/A+filt})} \approx \frac{1}{\sqrt{2}} \sigma_{\Delta p_t}^{(\mathrm{C/A})}$$

Differences between algorithms

- *Average shift*: preference for anti-k_t and C/A+filt
- *Dispersion*: preference for C/A+filtering

- Same conclusions for the RHIC



- No subtraction bias due to quenching (at most a 2% effect at the LHC)
- Valid for non-central collisions (smaller background but v_2)

  (with a residual ϕ dependence!)

Example: inclusive jet cross-section

Original hard spectrum:

$$\frac{d\sigma^{pp}}{dp_t} = \lambda\sigma_0\, e^{-\lambda p_t}$$

In the background, after subtraction

$$\begin{aligned} \frac{d\sigma^{AA,\mathrm{sub}}}{dp_t} &= \frac{d\sigma^{pp}}{dp_t} \otimes \mathrm{Gaussian}(\langle \Delta p_t \rangle, \sigma_{\Delta p_t}) \\ &= \frac{d\sigma^{pp}}{dp_t} \exp\left(\lambda \langle \Delta p_t \rangle + \frac{\lambda^2 \sigma^2_{\Delta p_t}}{2} \right) \end{aligned}$$

In practice, we have $\lambda \approx 0.3$ GeV^{-1} for RHIC, 0.05 for the LHC

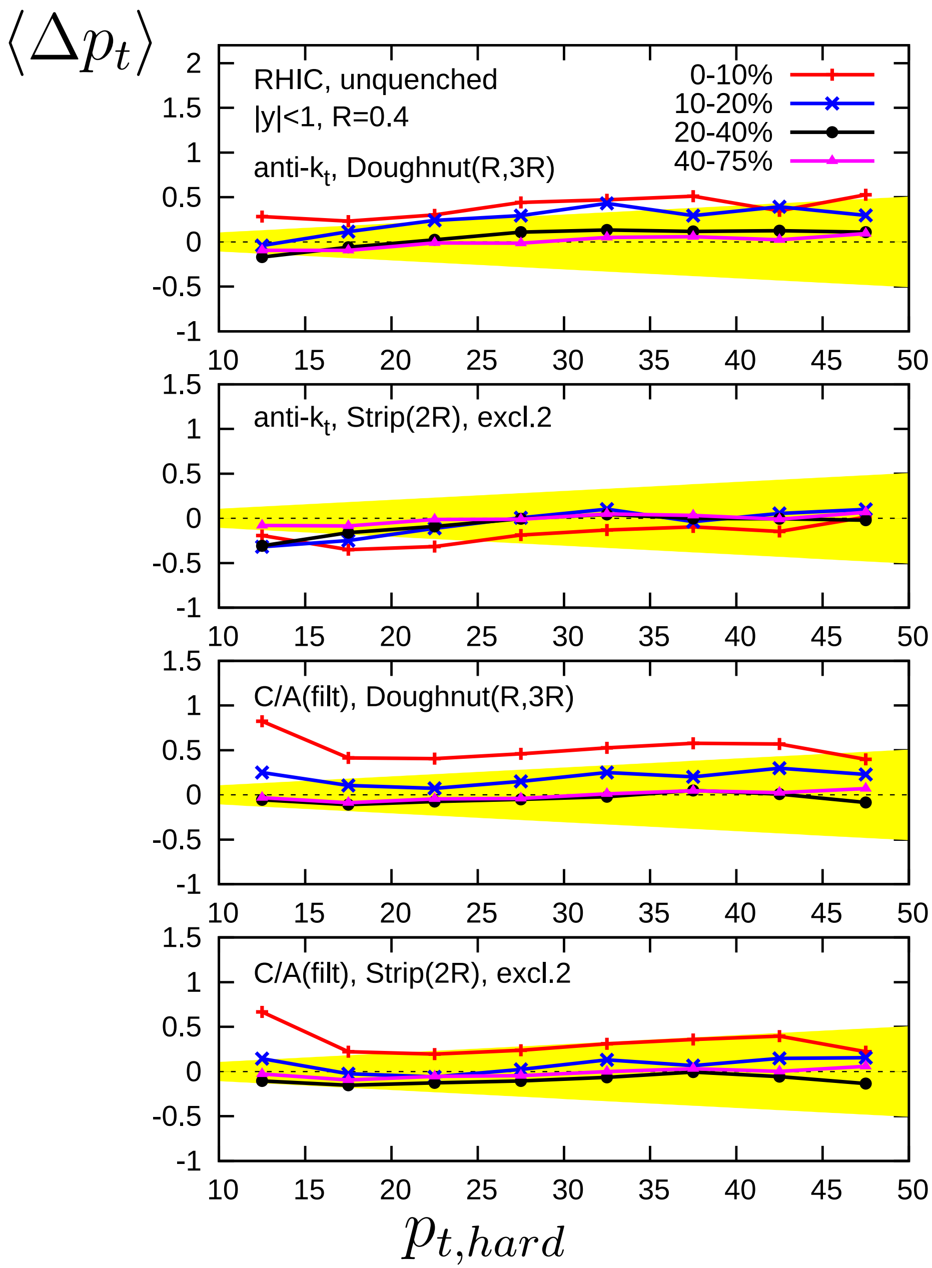
	RHIC			LHC		
$R = 0.4$	$\langle \Delta p_t \rangle$	$\sigma_{\Delta p_t}$	$\frac{d\sigma^{AA,\mathrm{sub}}/dp_t}{d\sigma^{pp}/dp_t}$	$\langle \Delta p_t \rangle$	$\sigma_{\Delta p_t}$	$\frac{d\sigma^{AA,\mathrm{sub}}/dp_t}{d\sigma^{pp}/dp_t}$
anti-k_t	0	7.5	12	0	18	1.5
C/A+filt	0	4.8	3	0	13	1.2

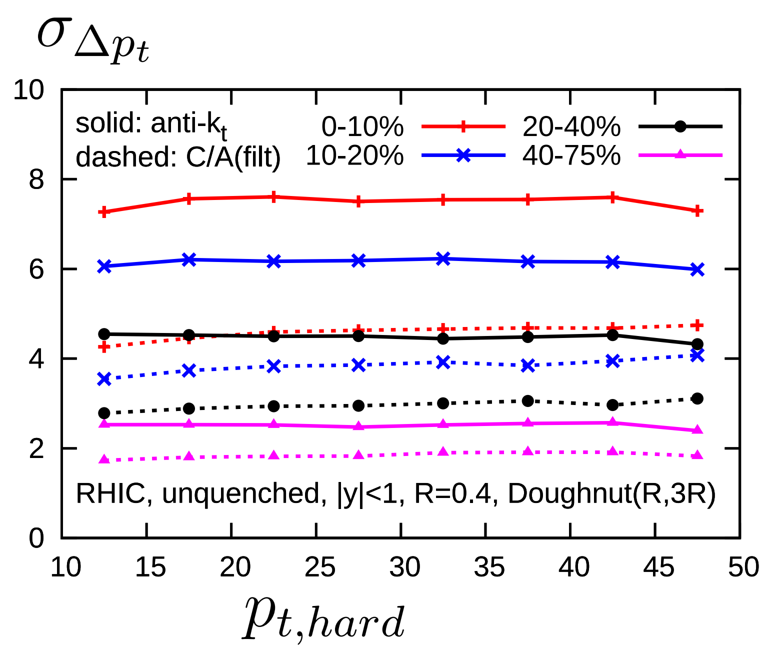
Summary

- The recipe: $p_{t,\text{jet}}^{(\text{sub})} = p_{t,\text{jet}} - \rho_{\text{bkg}} A_{\text{jet}}$
  - Define a jet: use an IRC-safe one
  - Define the area of a jet: ghost-based active area
  - Obtain ρ_{bkg}, the background p_t density per unit area: median of $\{p_{t,j}/A_j\}$
- New hints:
  1. Use filtering: reduce sensitivity to background (smaller $\sigma_{\Delta p_t}$)
  2. Use local ranges:
     handle non-uniform backgrounds (+ estimate subtraction error)
- Efficiency:
  - At least ≈ 8 jets in a local range
  - anti-k_t and C/A+filt give $\langle \Delta p_t \rangle \approx 0$ ($\langle \Delta p_t \rangle / p_t \lesssim 1\%$)
  - C/A+filt has reduced $\sigma_{\Delta p_t}$ (helps unfolding espacially at RHIC)

Further investigation

Centrality dependence

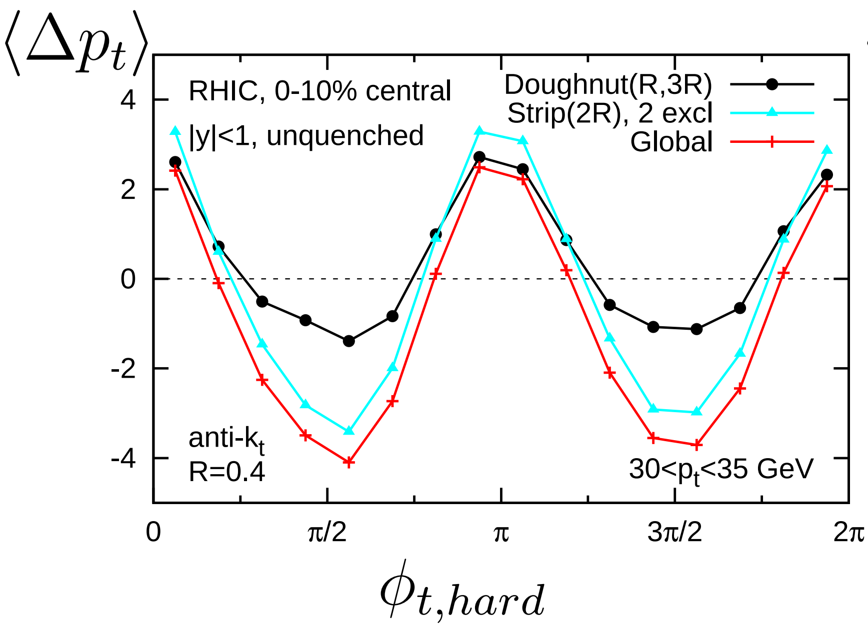


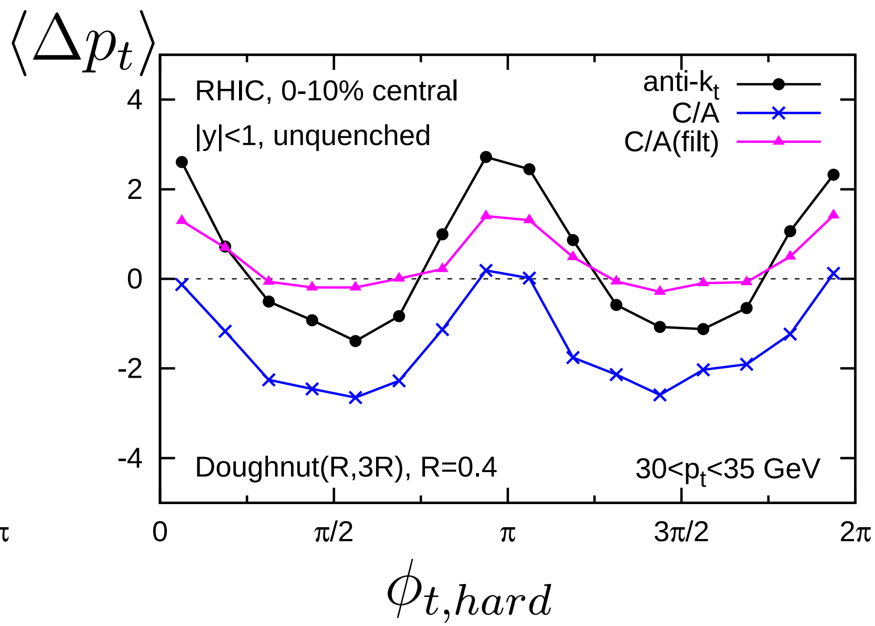


- $\langle \Delta p_t \rangle \approx 0$
- $\sigma_{\Delta p_t}$ decreases with centrality

Centrality dependence (2)

Look at $\langle \Delta p_t \rangle$ as a function of ϕ (wrt. reaction plane)





- the Doughnut range has less variations (less ϕ averaging!)
- C/A+filt has less variations (reduced area/dispersion)

Fakes

Issue: background fluctuations may “mimic” a hard jet without a “real hard origin”

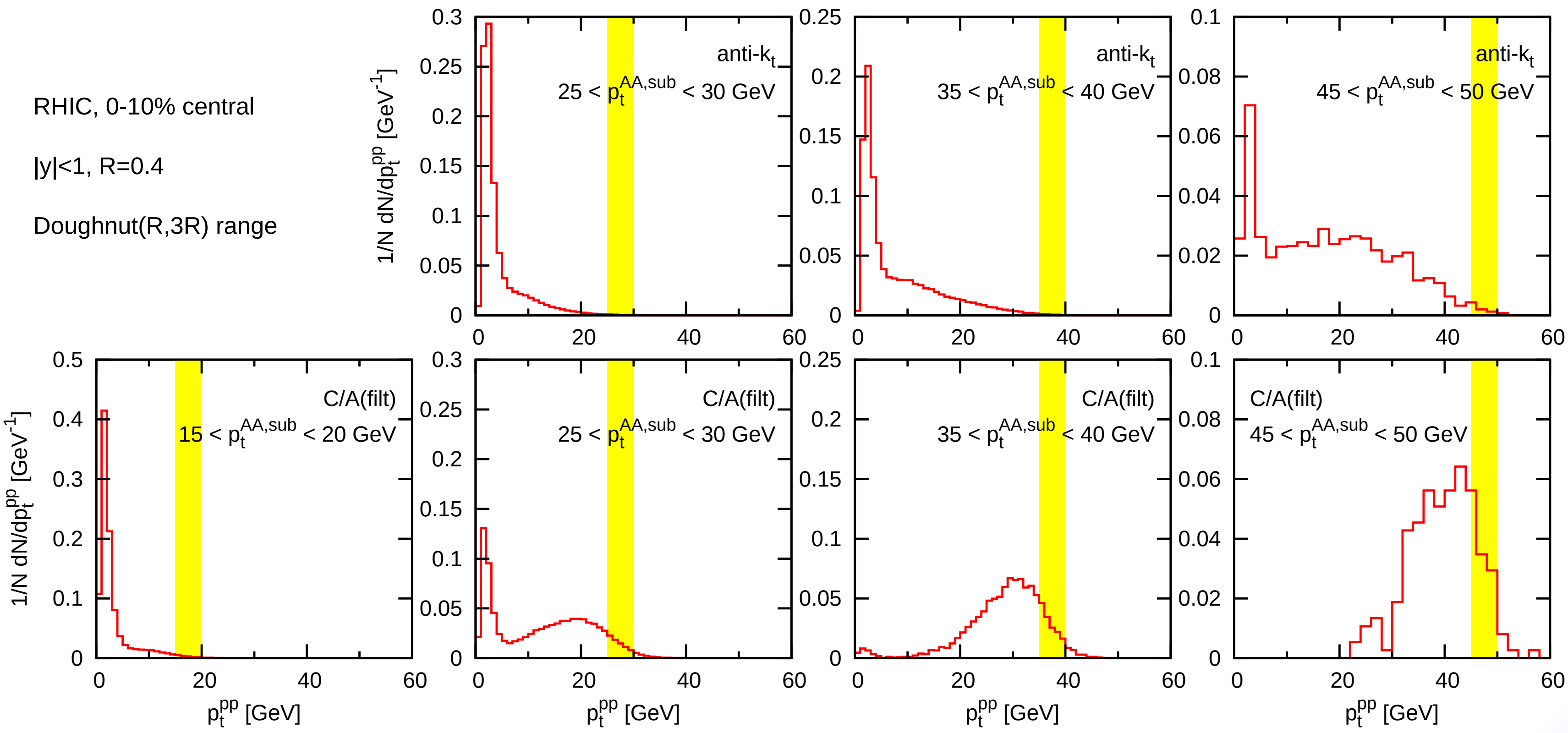
- separation hard/soft is unclear/model-dependent
- there is a large number of semi-hard jets

$\Rightarrow$ the question is more: “if we observe a jet with a given reconstructed p_t, is the original semi-hard jet softish or close to p_t?”

- Inclusive spectrum: fakes $\rightarrow$ unfolding
- Exclusive (*e.g.* Dijets): hard-soft separation clearer

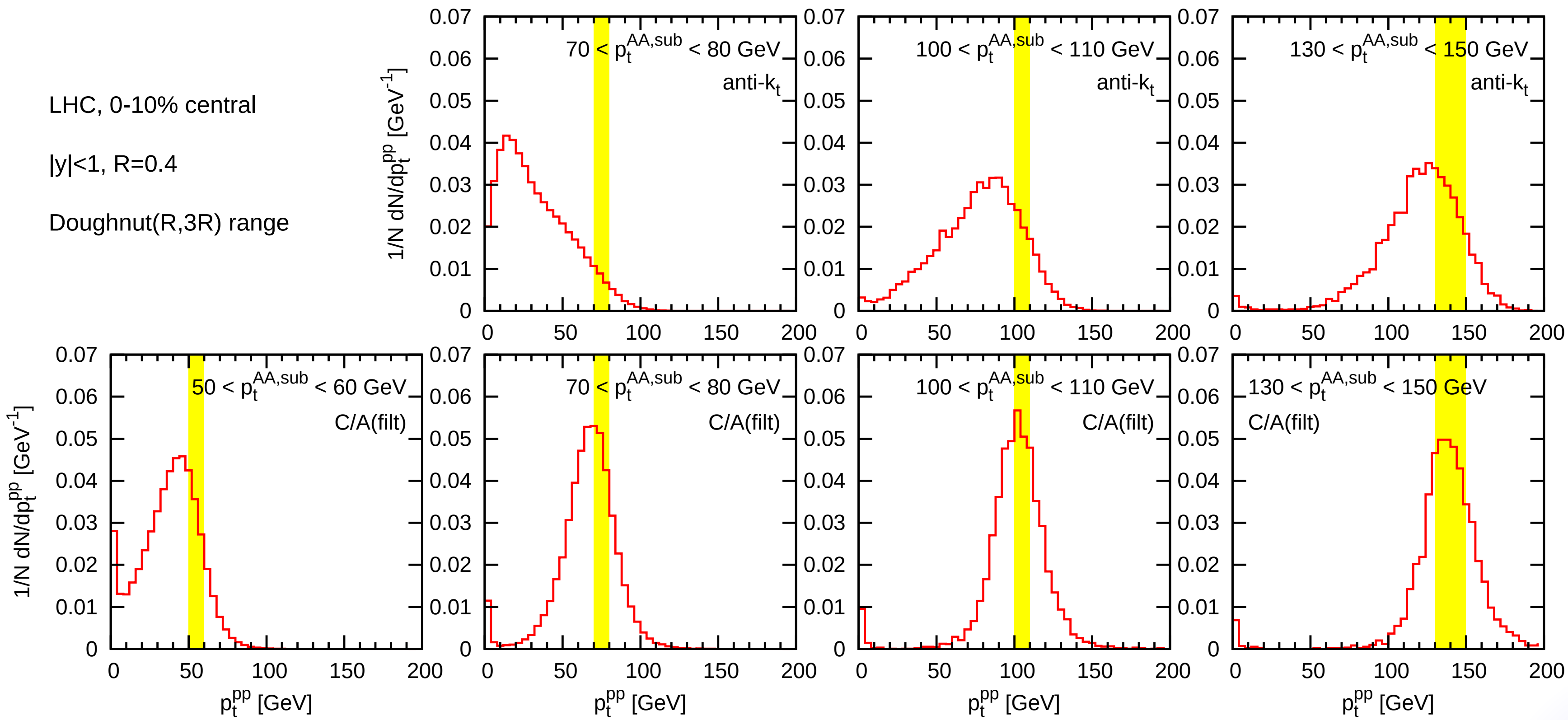
Fakes – origin (RHIC)

- HYDJET: HYDRO + many semi-hard pp event
- take each of the of the full (subtracted) jets
- find the corresponding semi-hard pp jet



Fakes – origin (LHC)

- HYDJET: HYDRO + many semi-hard pp event
- take each of the of the full (subtracted) jets
- find the corresponding semi-hard pp jet

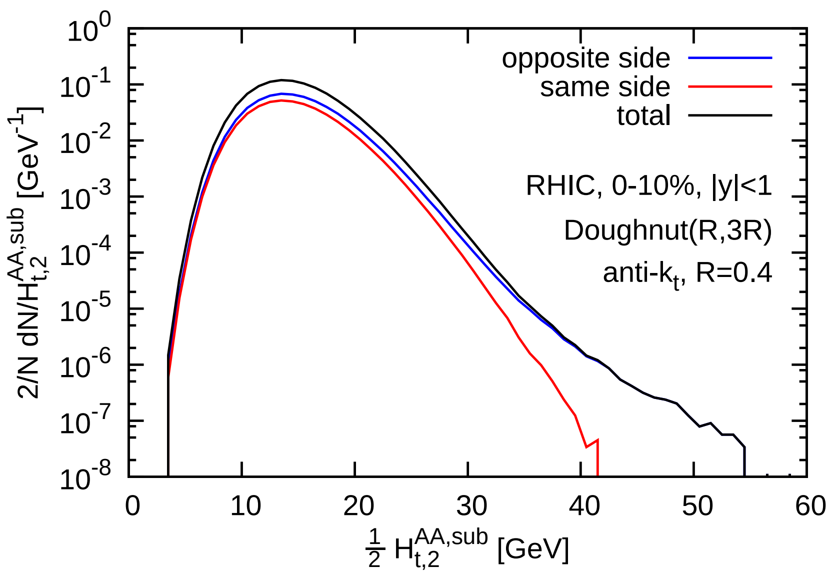


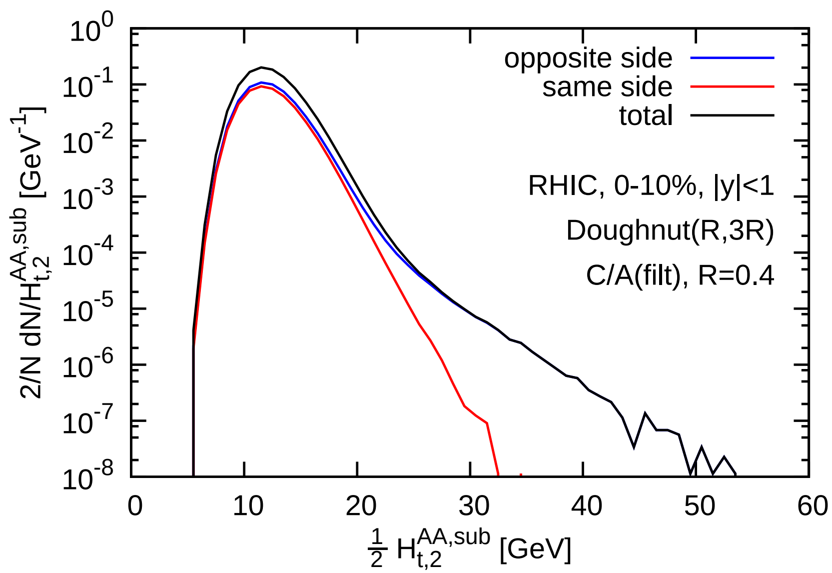
Fakes – dijet case (RHIC)

- Take the 2 hardest jets in a HYDJET event

$$H_{T,2} = p_{t,1}^{AA,sub} + p_{t,2}^{AA,sub}$$

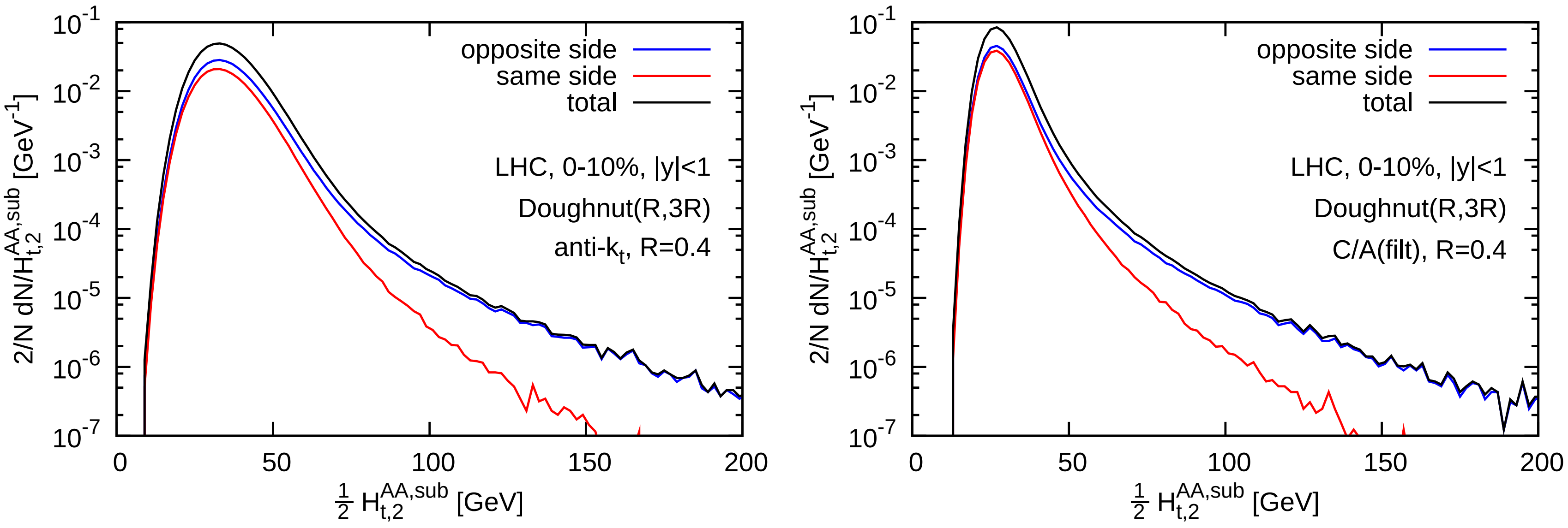
- split into same-side and opposite-side
- data-driven estimate of “fakes”





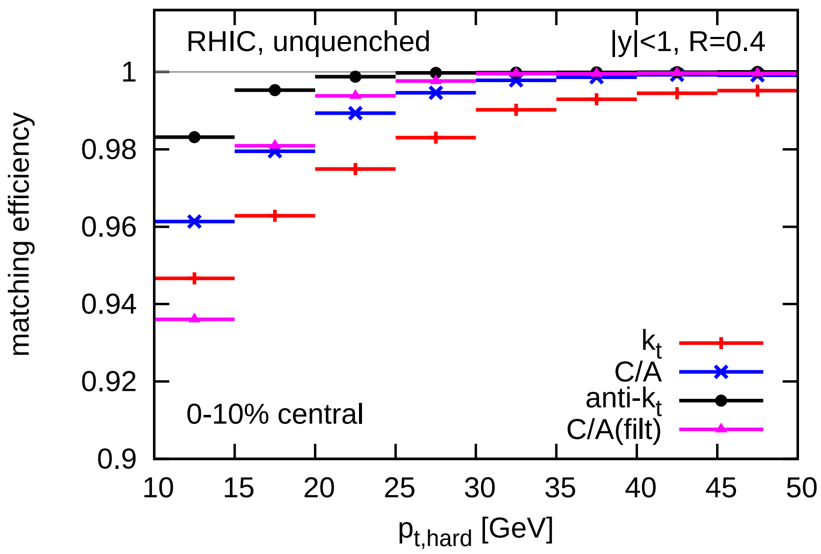
Fakes – dijet case (LHC)

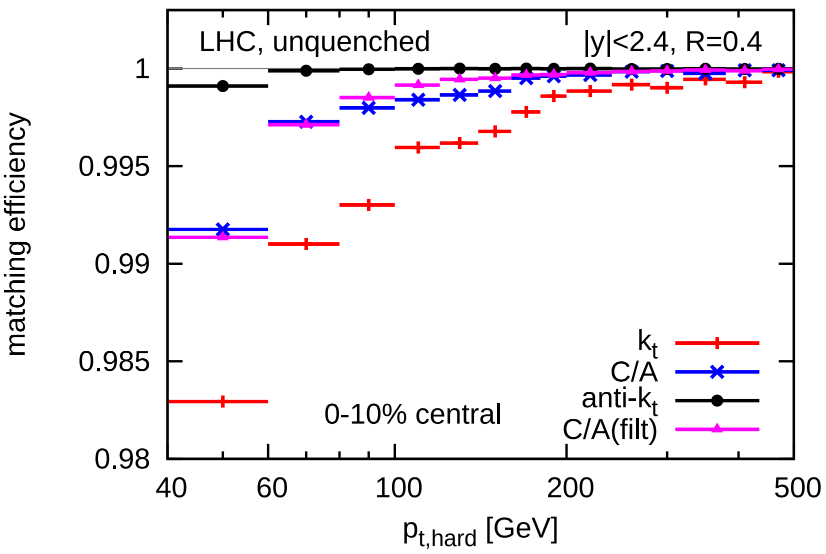
- Take the 2 hardest jets in a HYDJET event
- split into same-side and opposite-side
- data-driven estimate of “fakes”



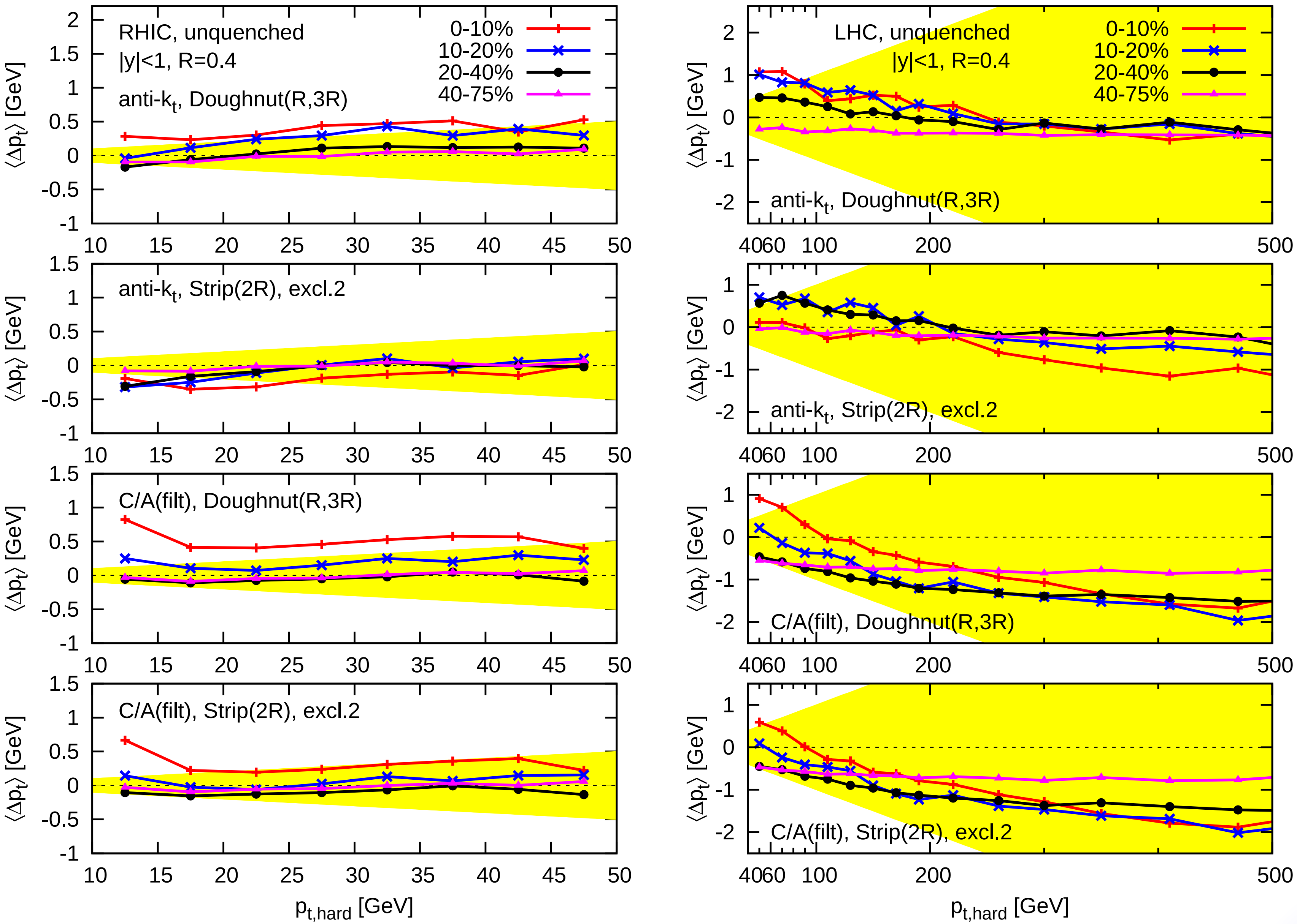
Backup slides

Efficiency

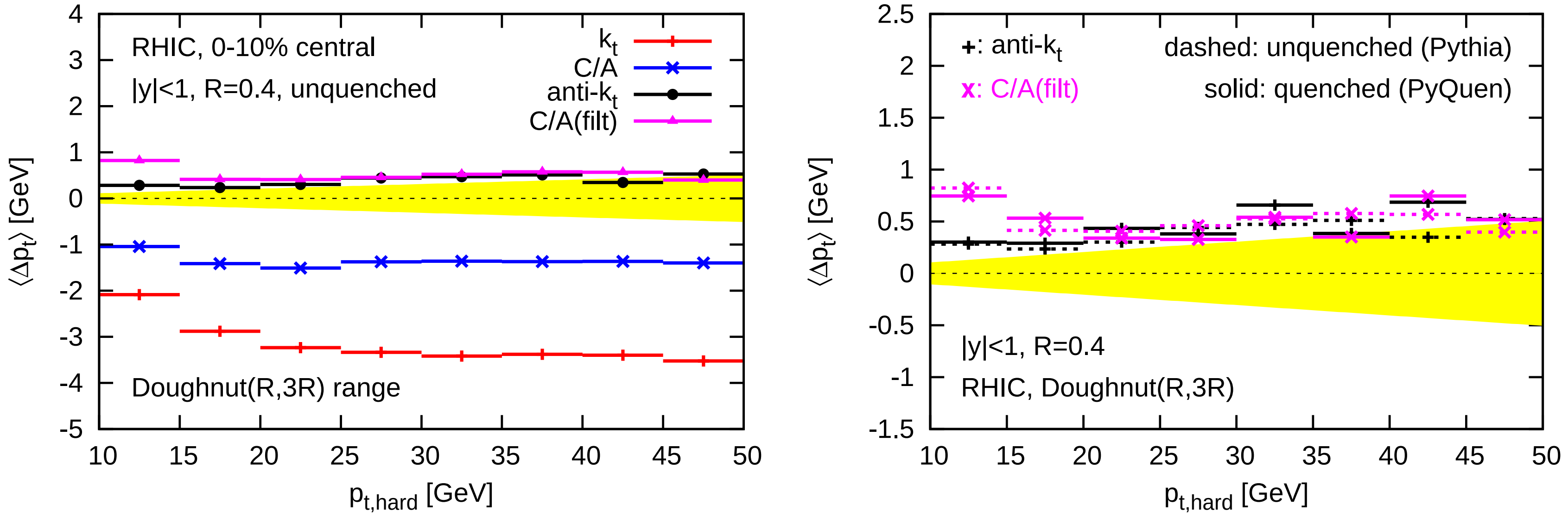

RHIC, unquenched
|y|<1, R=0.4
0-10% central
matching efficiency
$p_{t,hard}$ [GeV]
k_t
C/A
anti-k_t
C/A(filt)
1
0.98
0.96
0.94
0.92
0.9
10
15
20
25
30
35
40
45
50


LHC, unquenched
|y|<2.4, R=0.4
0-10% central
matching efficiency
$p_{t,hard}$ [GeV]
k_t
C/A
anti-k_t
C/A(filt)
1
0.995
0.99
0.985
0.98
40
60
100
200
500

Centrality dependence: LHC



Quenching effects - RHIC



Quenching effects - LHC