PileUp Per Particle Id

Daniele Bertolini
MIT
Philip Harris
CERN
Matthew Low
University of Chicago
Nhan Tran
Fermilab/LPC

work soon to appear

Mitigation of pileup effects at the LHC
May 16, 2014

PUPPI

Daniele Bertolini
MIT
Philip Harris
CERN
Matthew Low
University of Chicago
Nhan Tran
Fermilab/LPC

work soon to appear

Mitigation of pileup effects at the LHC
May 16, 2014

outline

- **what is PUPPI?**
- defining the algorithm
- setup and results
- discussion and extensions

handles on pileup

global event metric

local shape

tracking/vertexing

precision timing

depth segmentation

(apologies, not a complete list!)

ρ correction/subtraction

(area, 4-vector, shape, particle)

grooming

topoclustering

charged hadron subtraction

jet cleansing

pileup jet ID

...

handles on pileup

global event metric

local shape

tracking/vertexing

precision timing

depth segmentation

(apologies, not a complete list!)

ρ correction/subtraction

(area, 4-vector, shape, particle)

grooming

topoclustering

charged hadron subtraction

jet cleansing

pileup jet ID

...

handles on pileup

global event metric

local shape

tracking/vertexing

precision timing

depth segmentation

(apologies, not a complete list!)

ρ correction/subtraction

(area, 4-vector, shape, particle)

grooming

topoclustering

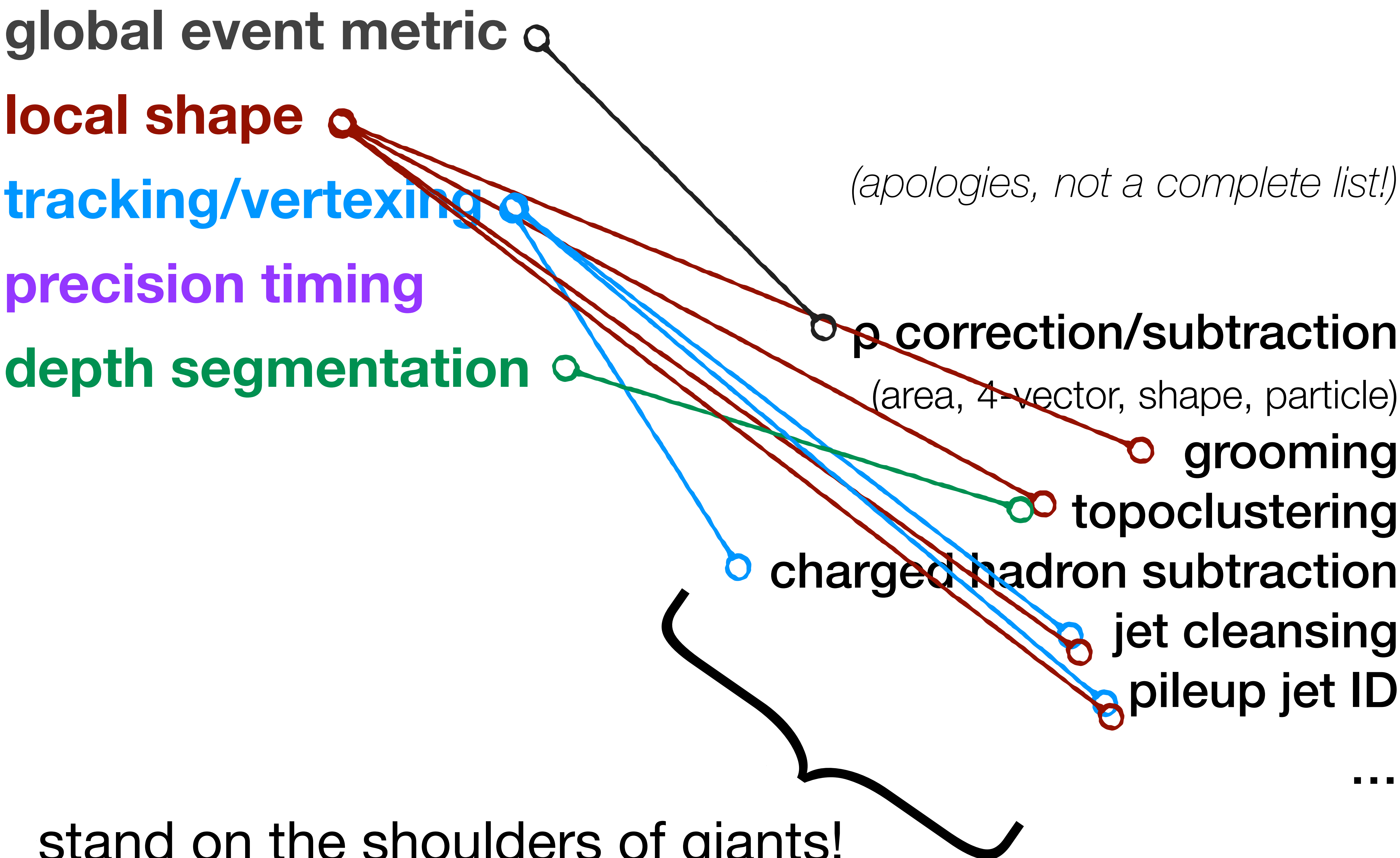
charged hadron subtraction

jet cleansing

pileup jet ID

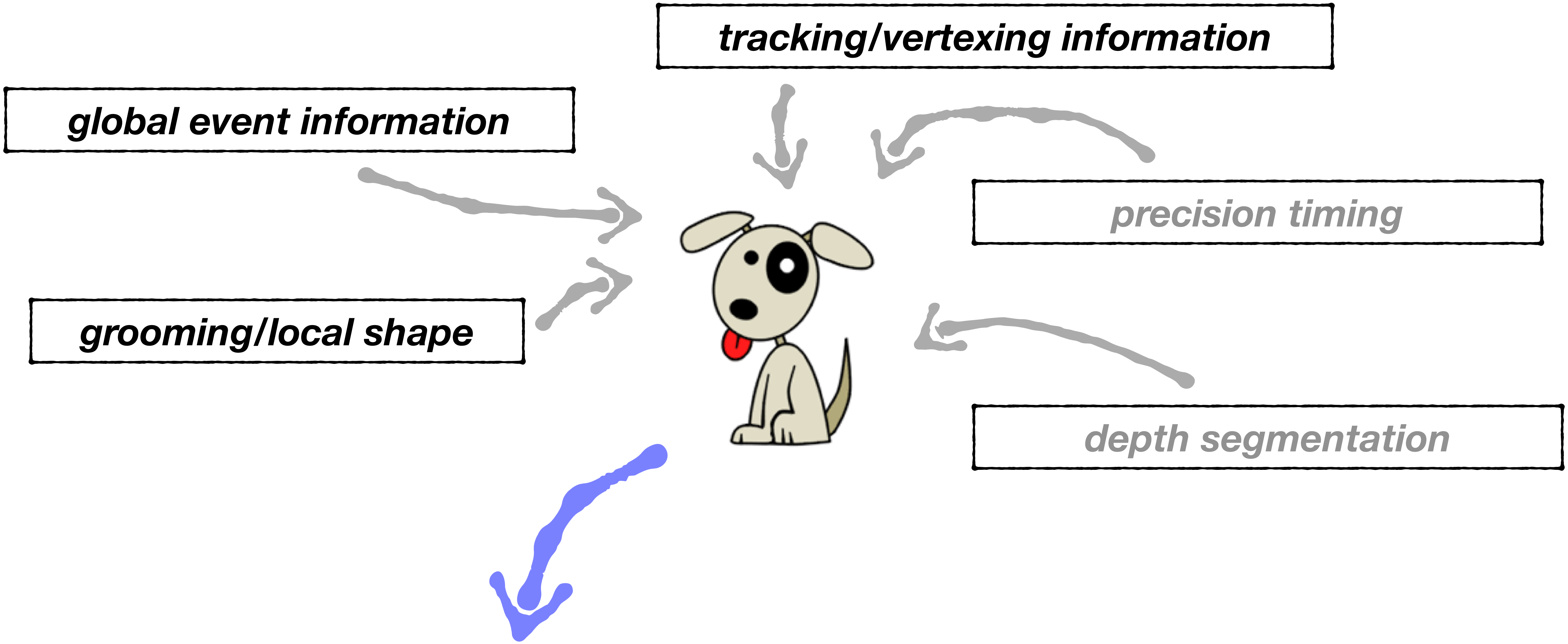
…

handles on pileup



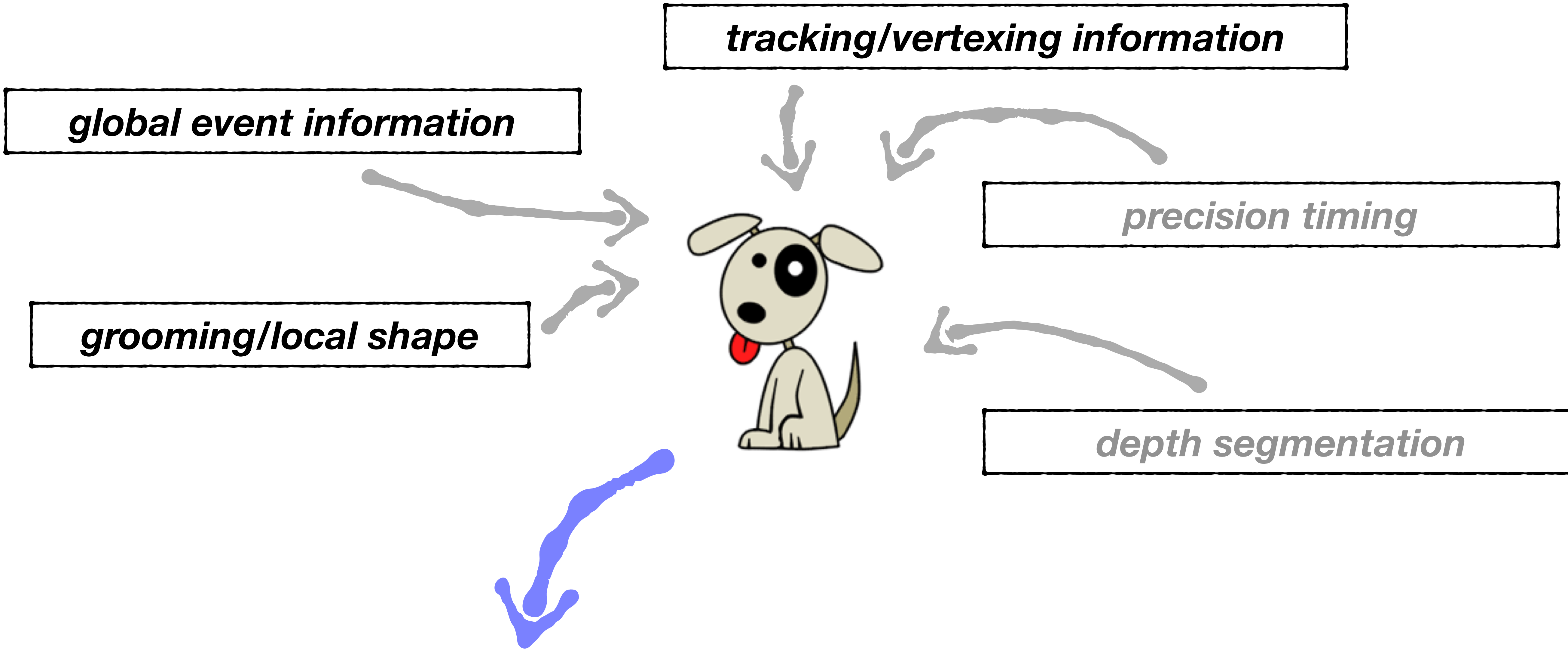
stand on the shoulders of giants!

PUPPI



following the "jets without jets" paradigm…

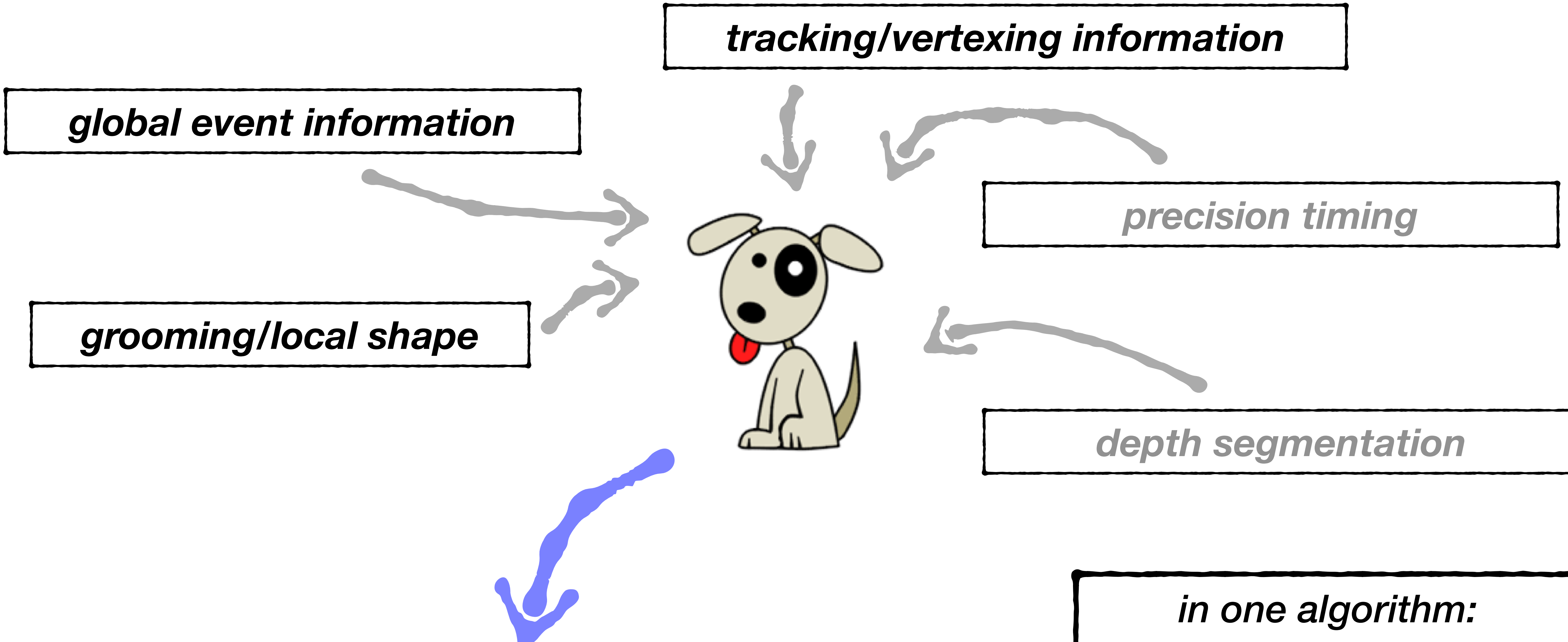
PUPPI



following the "jets without jets" paradigm…

Define on a ***per particle*** basis, ***before jet clustering***, a weight for ***how likely*** a particle (or jet constituent) is to be from pileup or the leading vertex, then rescale each particle four momentum by that likelihood

PUPPI



following the "jets without jets" paradigm…

Define on a ***per particle*** basis, ***before jet clustering***, a weight for ***how likely*** a particle (or jet constituent) is to be from pileup or the leading vertex, then rescale each particle four momentum by that likelihood

in one algorithm:

correct the jet pT

correct the jet mass/ shapes

perform pileup jet ID

classify particles for MET determination

outline

- what is PUPPI?
- **defining the algorithm**
- setup and results
- discussion and extensions

PUPPI overview

for a particle *i* with nearby particles *j*

[1] define a local metric, α, that differs between pileup (PU) and leading vertex (LV)

Let's assume something similar to **Particle Flow** inputs (*not a requirement*):
neutral hadrons
charged hadrons from LV
charged hadrons from PU

PUPPI overview

for a particle *i* with nearby particles *j*

[1] define a local metric, α, that differs between pileup (PU) and leading vertex (LV)

[2] using tracking information (e.g. charged particles) "sample" the event, define unique distributions of α for PU and LV

Let's assume something similar to **Particle Flow** inputs (*not a requirement*):
neutral hadrons
charged hadrons from LV
charged hadrons from PU

PUPPI overview

for a particle *i* with nearby particles *j*

Let's assume something similar to **Particle Flow** inputs (*not a requirement*):
neutral hadrons
charged hadrons from LV
charged hadrons from PU

[1] define a local metric, α, that differs between pileup (PU) and leading vertex (LV)

[2] using tracking information (e.g. charged particles) "sample" the event, define unique distributions of α for PU and LV

[3] for the neutrals, ask "how PU-like is α for this particle?", compute a weight for how un-PU-like (or LV-like) it is

PUPPI overview

for a particle *i* with nearby particles *j*

Let's assume something similar to **Particle Flow inputs** (*not a requirement*):
neutral hadrons
charged hadrons from LV
charged hadrons from PU

[1] define a local metric, α, that differs between pileup (PU) and leading vertex (LV)

[2] using tracking information (e.g. charged particles) "sample" the event, define unique distributions of α for PU and LV

[3] for the neutrals, ask "how PU-like is α for this particle?", compute a weight for how un-PU-like (or LV-like) it is

[4] reweight the four-vector of the particle by this weight, then proceed to cluster the event as usual

[1] define a metric

there are many possibilities, we have tested many and find these to be near-optimal

example: 2-body system, for a particle **i**, what does particle **j** tell us?

$$\frac{p_{T,j}}{\Delta R_{ij}}$$

⬆ for harder, collinear particles

⬇ for softer, wide angle particles

[1] define a metric

there are many possibilities, we have tested many and find these to be near-optimal

example: 2-body system, for a particle **i**, what does particle **j** tell us?

$$\frac{p_{T,j}}{\Delta R_{ij}}$$

⬆ for harder, collinear particles

⬇ for softer, wide angle particles

define **2** metrics, with and without using tracking information

$$\alpha_i^{C} = \log\left[\sum_{j\in \mathrm{Ch,LV}} \frac{p_{T,j}}{\Delta R_{ij}} \Theta(R_0 - \Delta R_{ij})\right]$$

$$\alpha_i^{F} = \log\left[\sum_{j\in \mathrm{all}} \frac{p_{T,j}}{\Delta R_{ij}} \Theta(R_0 - \Delta R_{ij})\right]$$

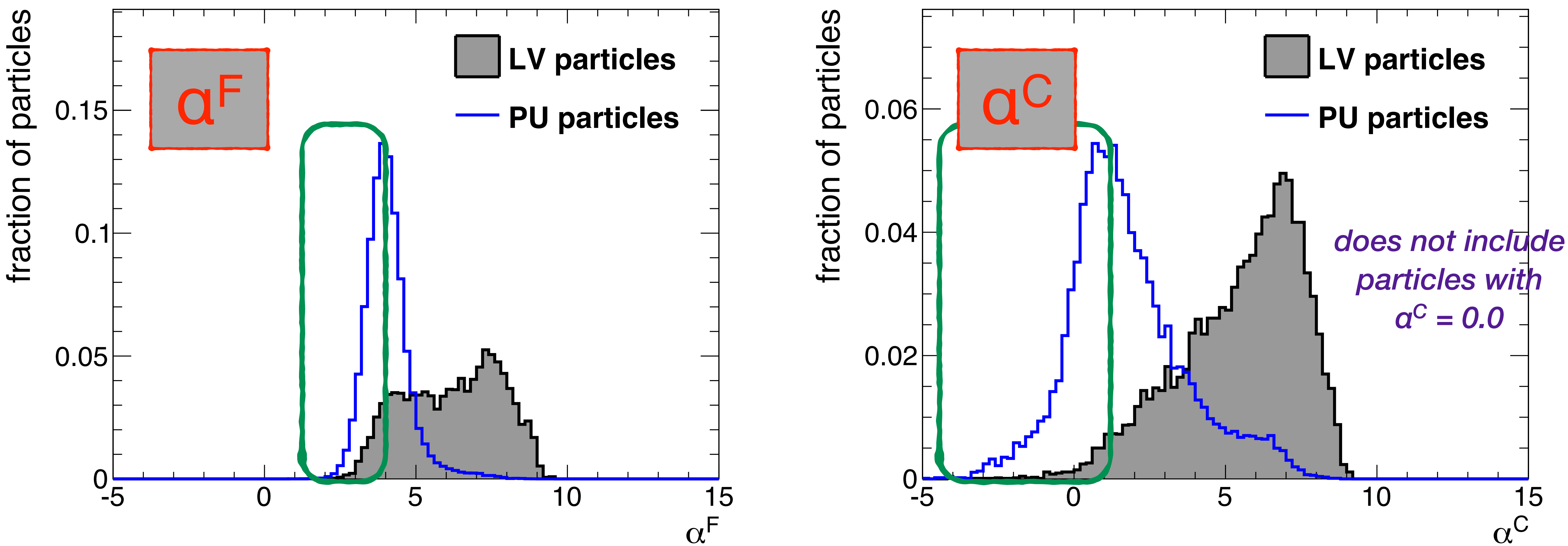
R_0 is the radius around particle i which is considered

the log regularizes the scale for event sampling (next slides)

“C” = charged or central, “F” = full or forward

[2] sample the event

charged particles with $p_T > 1$ GeV, populated over 200 events



sample the **charged** particles in the event,
get separate distributions of **α** for LV and PU particles

Determine the **median and (left-side) RMS values** for the charged PU

(good) assumption: α for charged PU and neutral PU are similar

need to extrapolate α^F to the forward region (η-dependence)

** dijet events, more details on simulation later

[3] compute weight

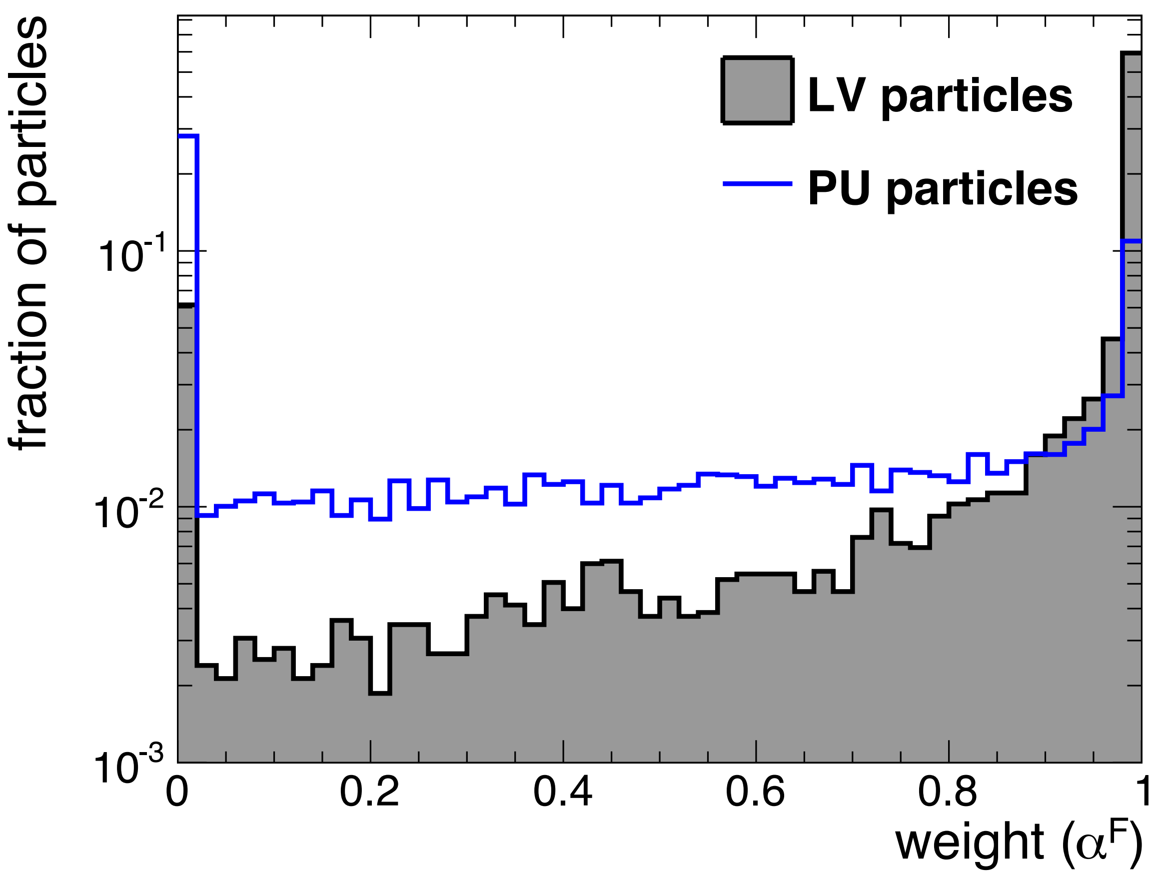
Use a ***χ^2 approximation*** to derive a ***probability of pileup***

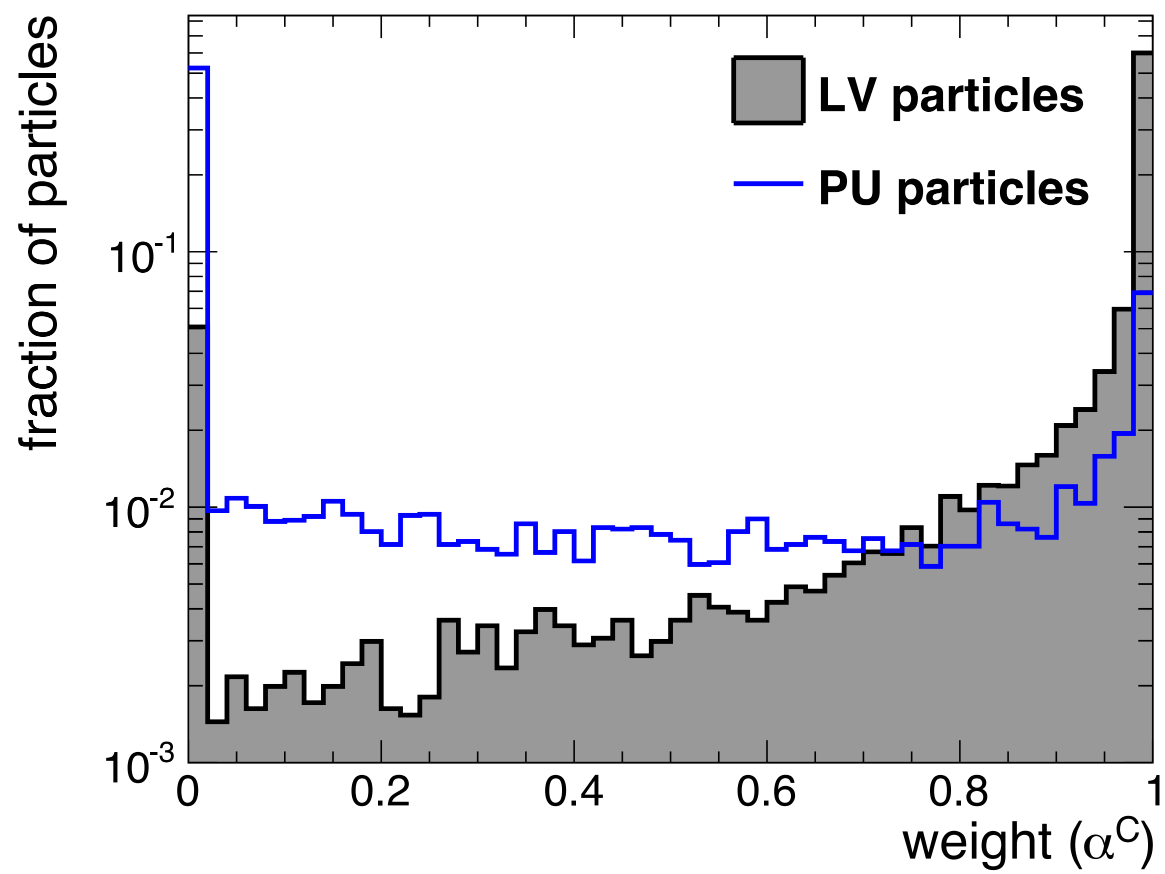
$$\chi_i^2 = \text{sgn}(\alpha_i - \bar{\alpha}_{\text{PU}}) \times \frac{(\alpha_i - \bar{\alpha}_{\text{PU}})^2}{\sigma_{\text{PU}}^2}$$

Use χ^2 CDF to determine probability, p_i:
weight, $w_i = 1 - p_i$

signed information because we know LV particles lie to the right of the PU distribution

grooming and local shape
global event metrics
tracking/vertexing information





χ^2 construction allows for additional information, use χ^2 with N DOF

[4] recompute 4-vectors

step 1: *assuming perfect tracking* for charged particles, set $\mathbf{w_{ChLV} = 1}$ and $\mathbf{w_{ChPU} = 0}$.

step 2: for neutrals, re-weight 4-vectors, $\mathbf{w_i \times p_i}$

step 3: cut on neutrals with $\mathbf{w_i < w_{cut}}$ and $\mathbf{w_i * p_{Ti} < \beta_{cut}}$

$\mathbf{w_{cut}}$ and $\mathbf{\beta_{cut}}$ are tunable parameters

this study: $w_{cut} = 0.001$

$\beta_{cut} = f(n_{PV}) \approx 0.7e\text{-}2 \times n_{pu} + 0.1$ GeV (central)

$\beta_{cut} = f(n_{PV}) \approx 1.1e\text{-}2 \times n_{pu} + 0.2$ GeV (forward)

$\mathbf{R_0}$, PUPPI cone size, is a tunable parameter

this study: $R_0 = 0.3$

outline

what is PUPPI?

defining the algorithm

setup and results

discussion and extensions

‘experimental’ setup

dijet events, Pythia 8.176

flat pT spectrum for populating higher p_T bins

pileup scenarios, n_{PU} = 20-140, every 15 n_{PU}

not Poisson-distributed

tracker exists to $|\eta| < 2.5$, all particles out to $|\eta| < 5$

assume perfect tracking to determine if particles are LV or PU

neutrals are reconstructed in cells ($\Delta\eta\times\Delta\phi$ = 0.1×0.1)

cut on neutrals, $p_T > 0.1$ GeV

input collections

LV

only the particles from the leading vertex of interest

PF: *like particle flow*

all charged and neutral particles

PFCHS: *like particle flow with charged hadron subtraction*

all charged and neutral particles except for charged particles from PU in the tracking volume

PUPPI: *like particle flow*

particles after the PUPPI algorithm

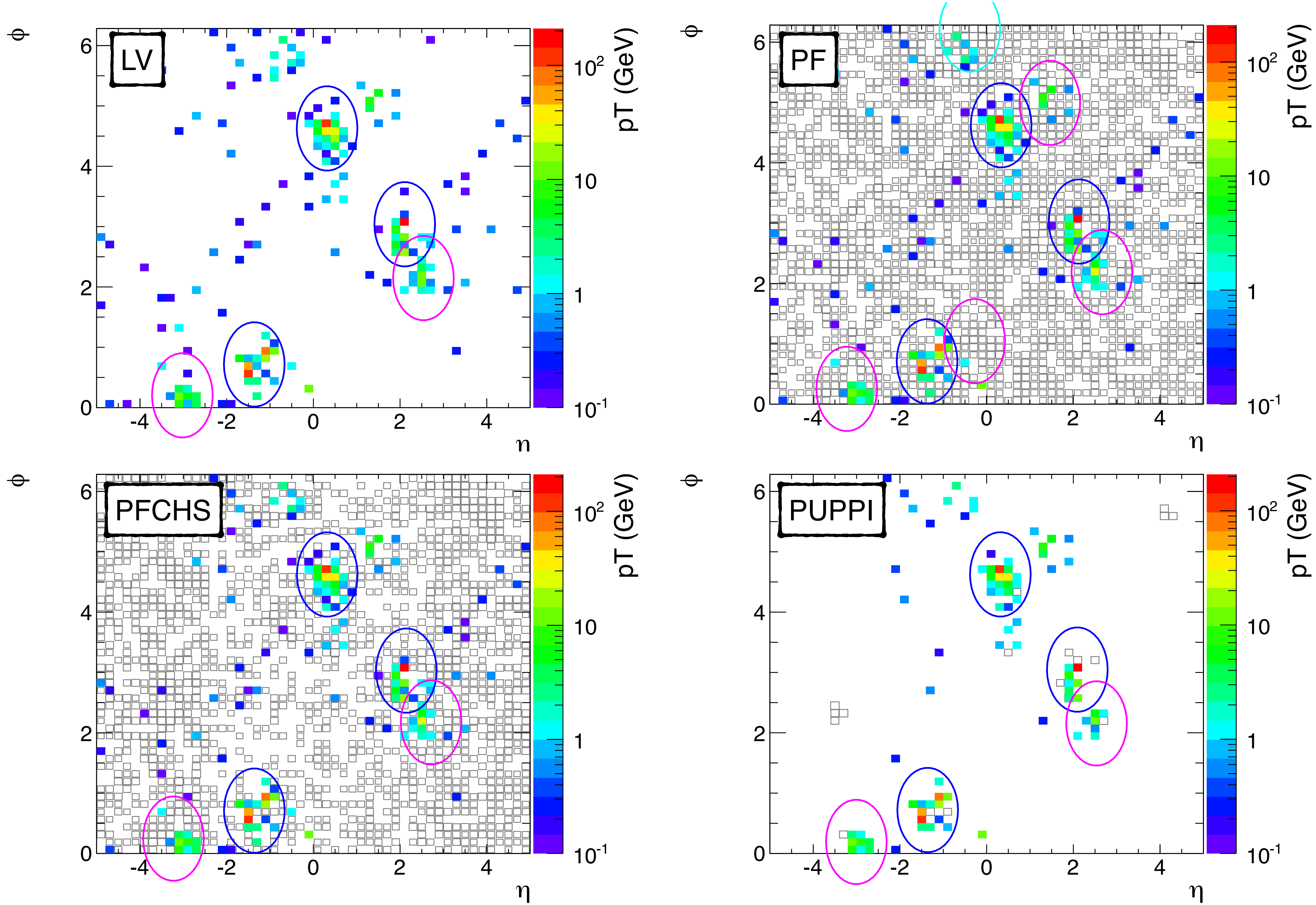
PF and **PFCHS** jet collections are corrected using fastjet 4-vector grid ρ subtraction

PFCHS uses its own ρ inside tracker volume and PF ρ outside

event displays

colored cells = LV process, black cells = pileup

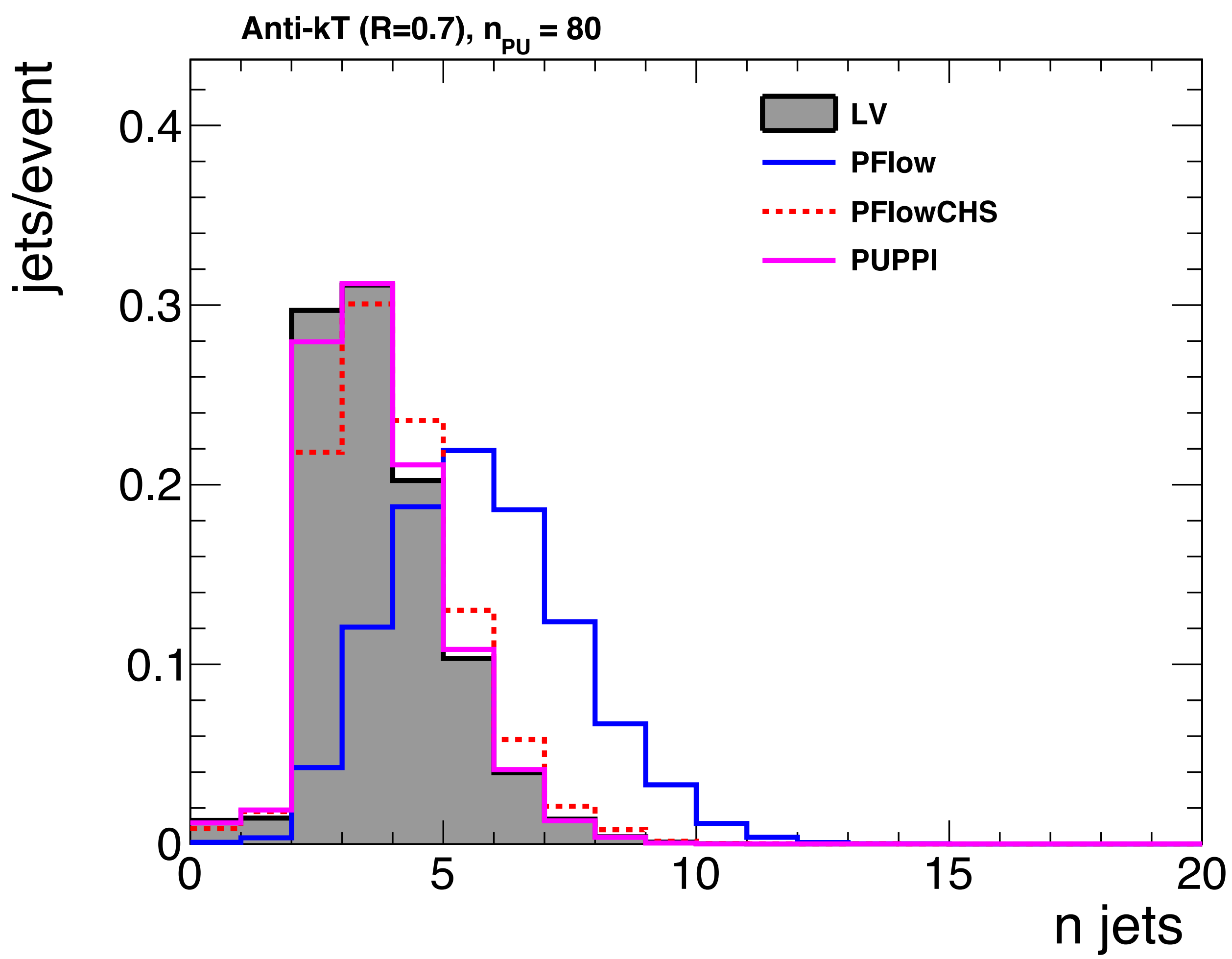
n_{PU} = 80 scenario

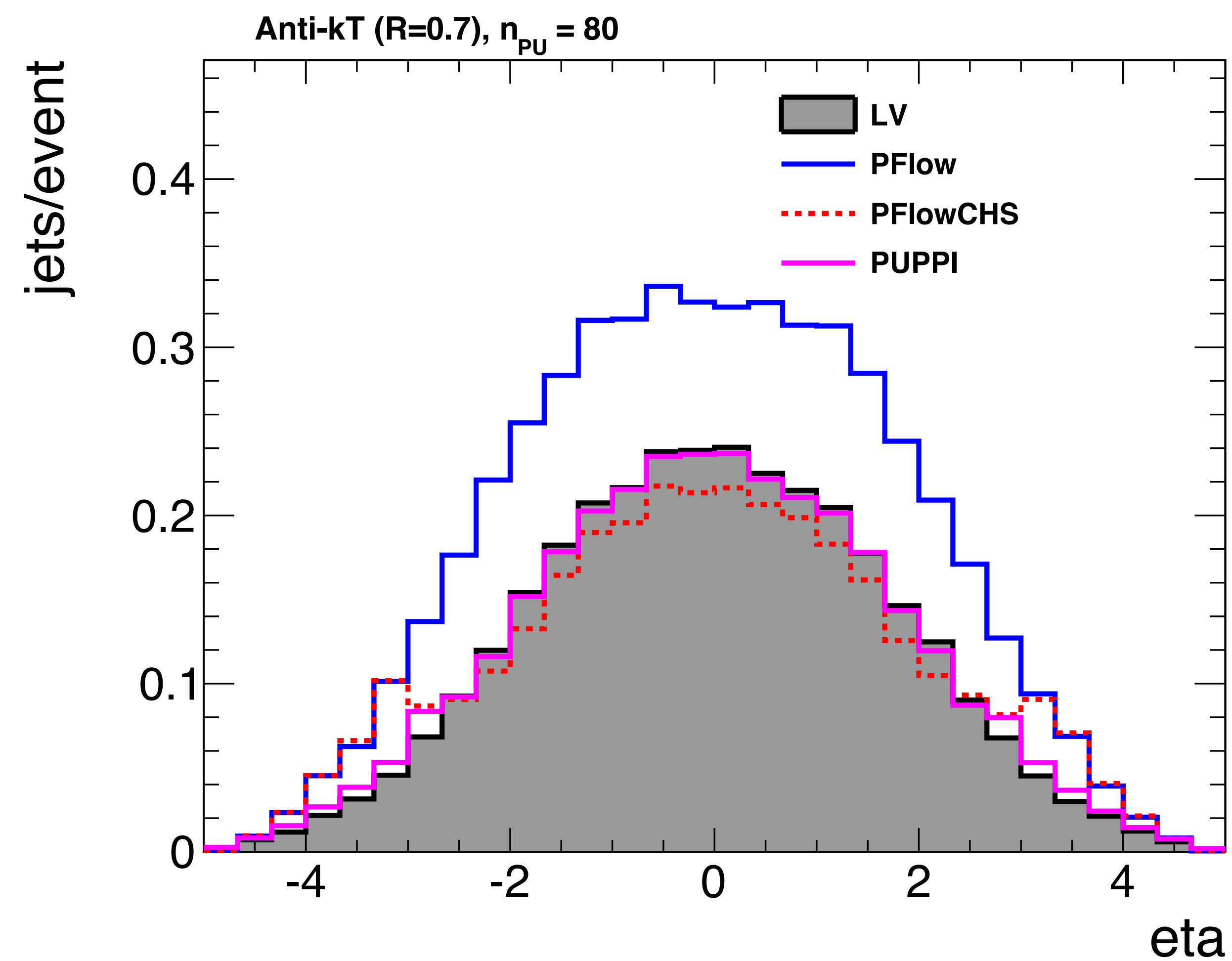


jet kinematics

n_{PU} = 80 scenario

Number of jets with pT > 25 GeV per event





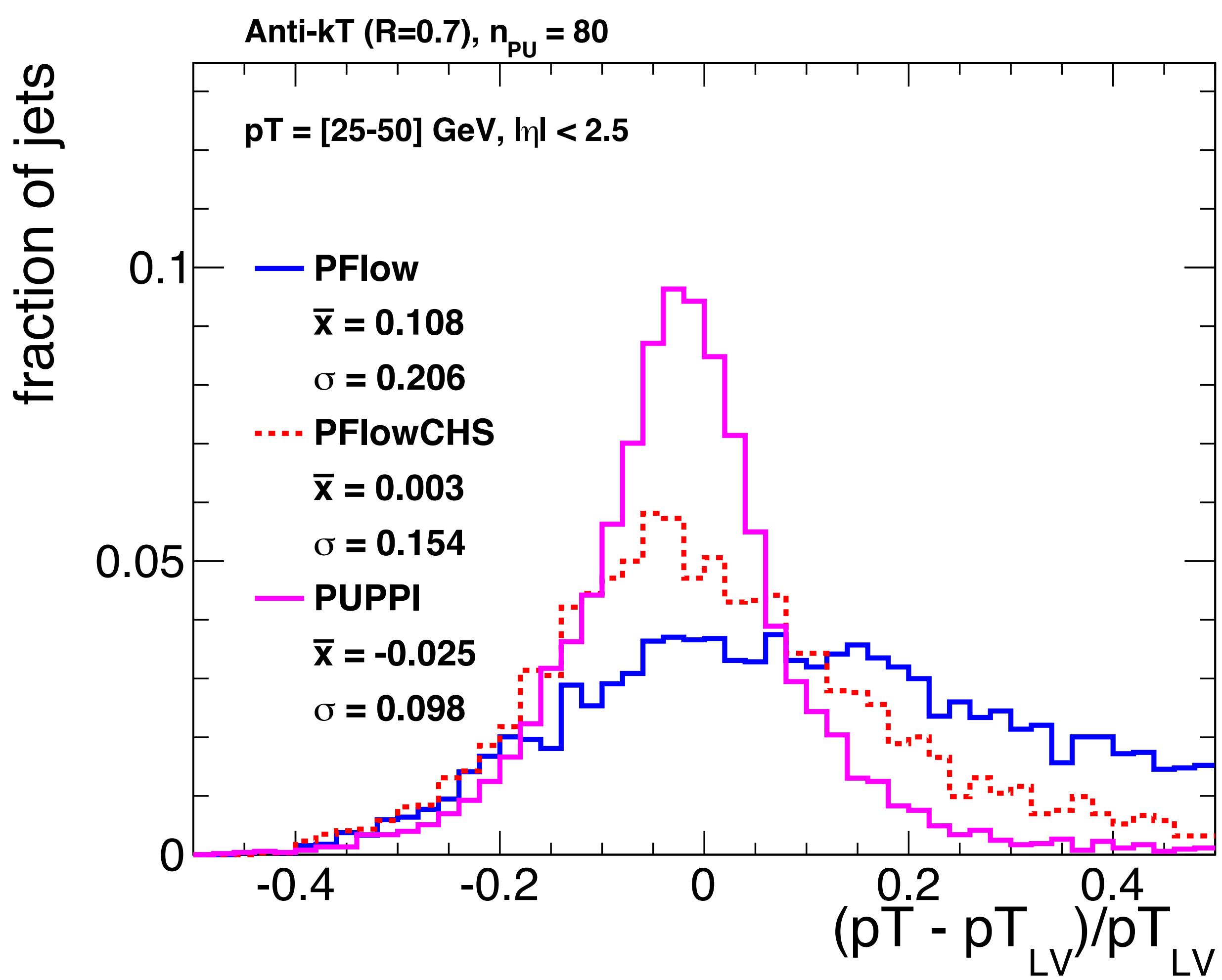
PUPPI performs pileup jet identification reducing the amount of spurious jets formed from overlapping pileup events

jet p_T resolution

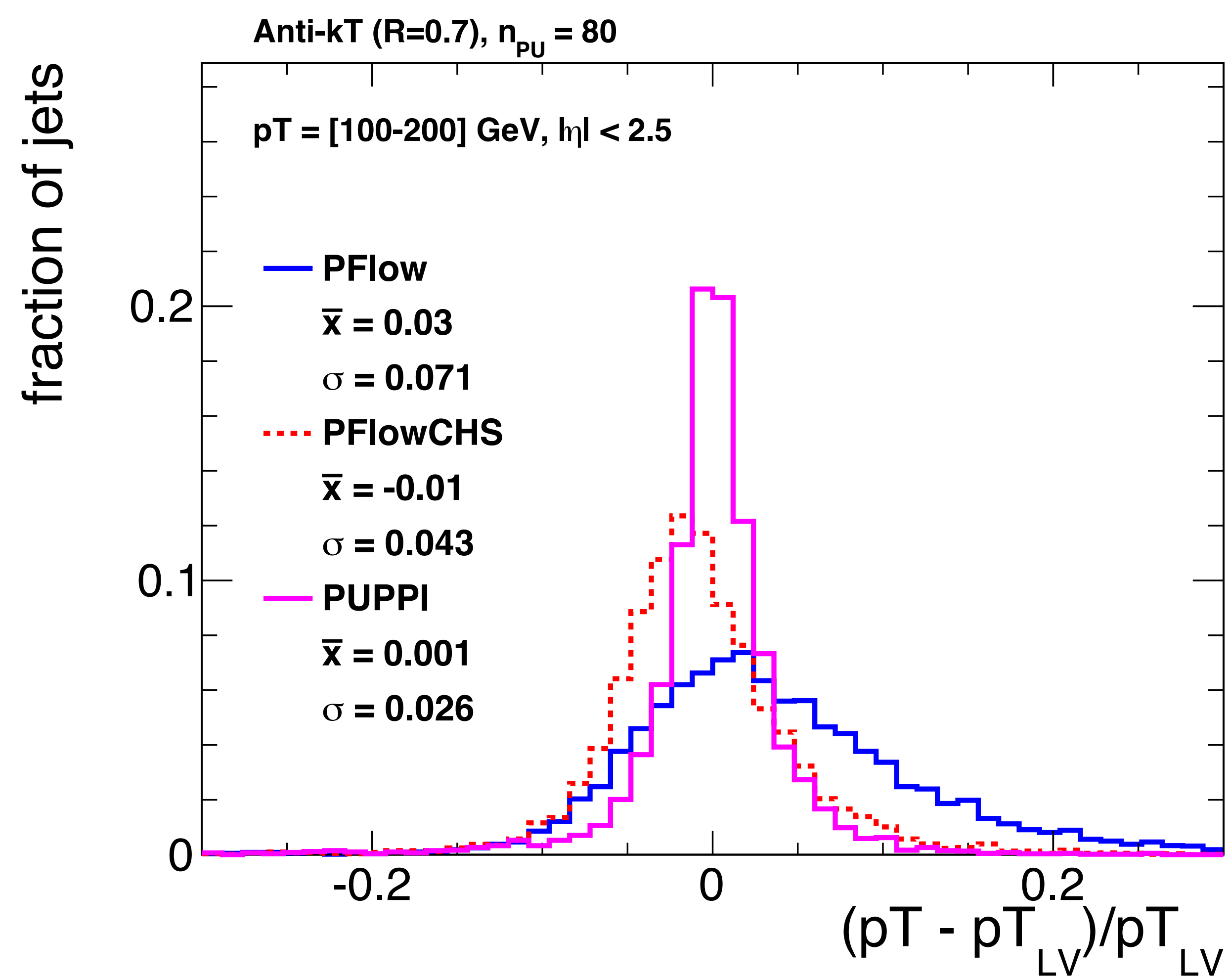
n_{PU} = 80 scenario

match PF, PFCHS, PUPPI jets to LV jets
and compare the jet p_T resolution

lower p_T jets (25-50 GeV)

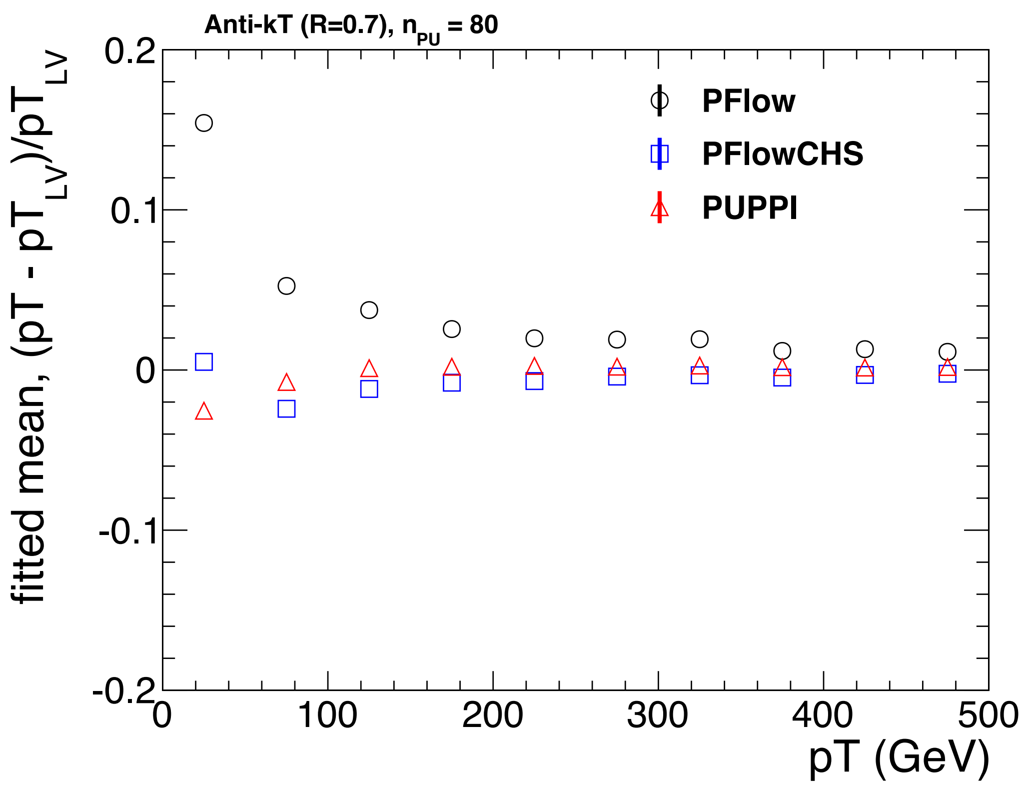


higher p_T jets (100-200 GeV)

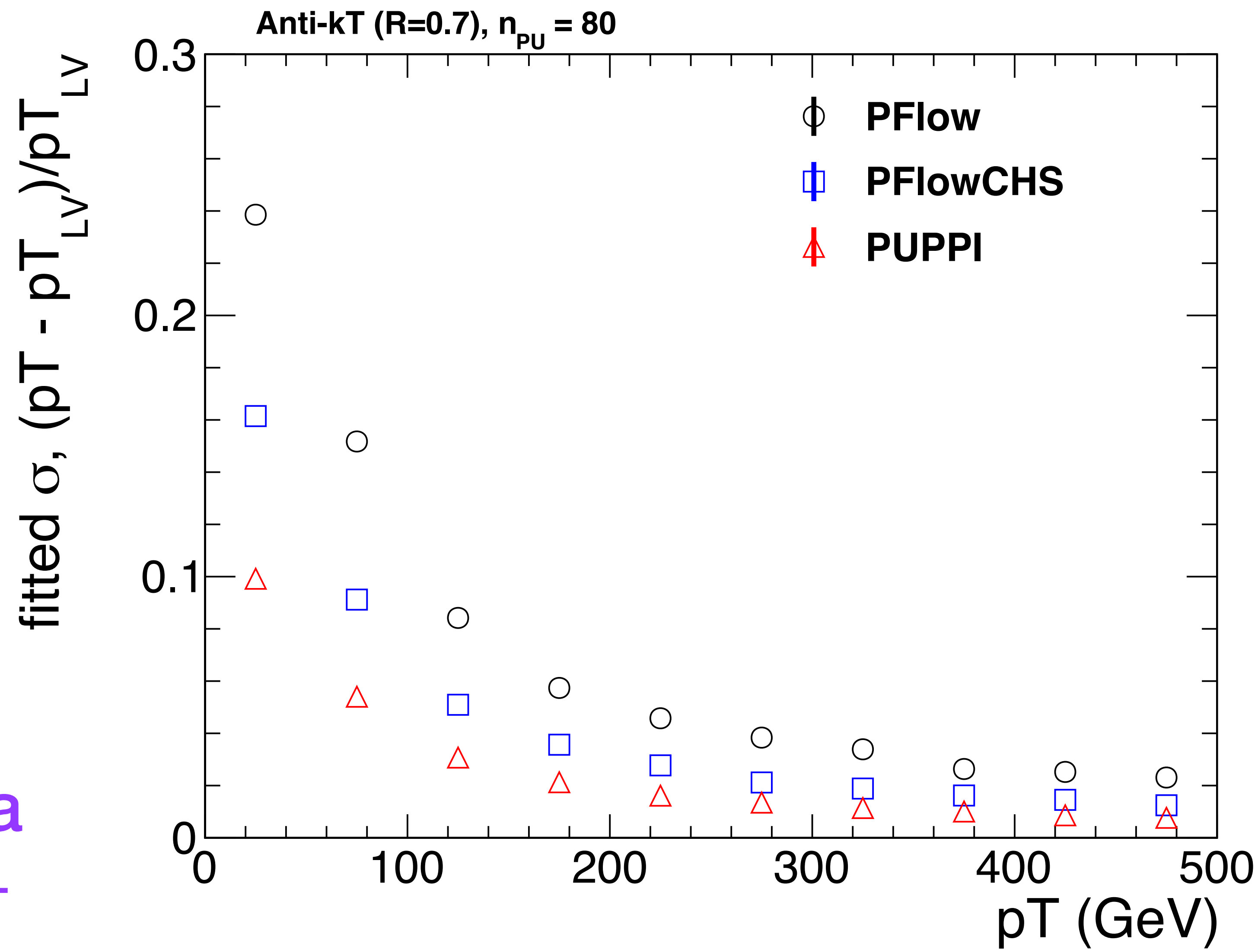


jet p_T resolution vs. p_T

n_{PU} = 80 scenario



jet p_T scale as a function of LV jet p_T

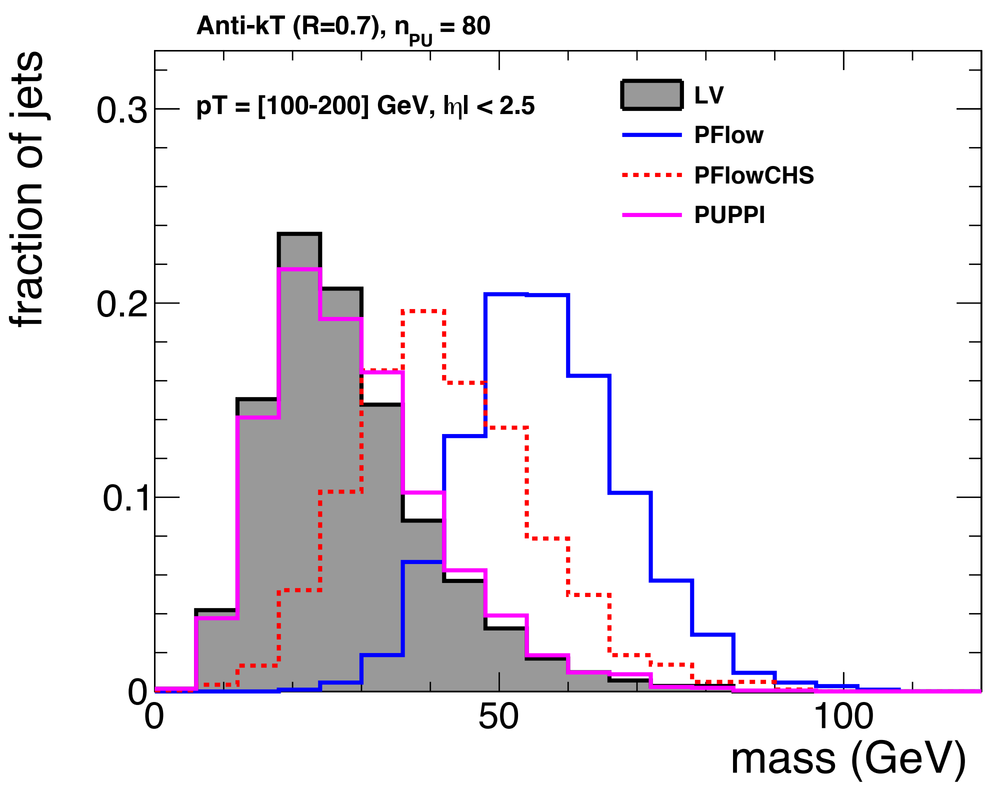


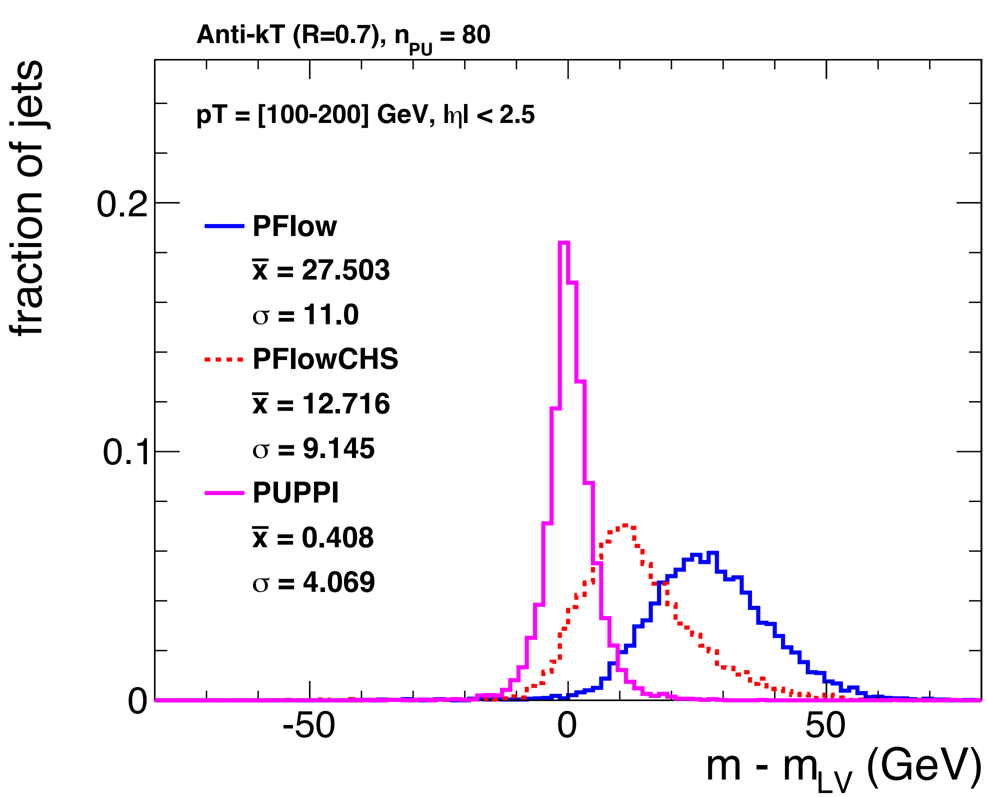
jet p_T resolution as a function of LV jet p_T

jet mass

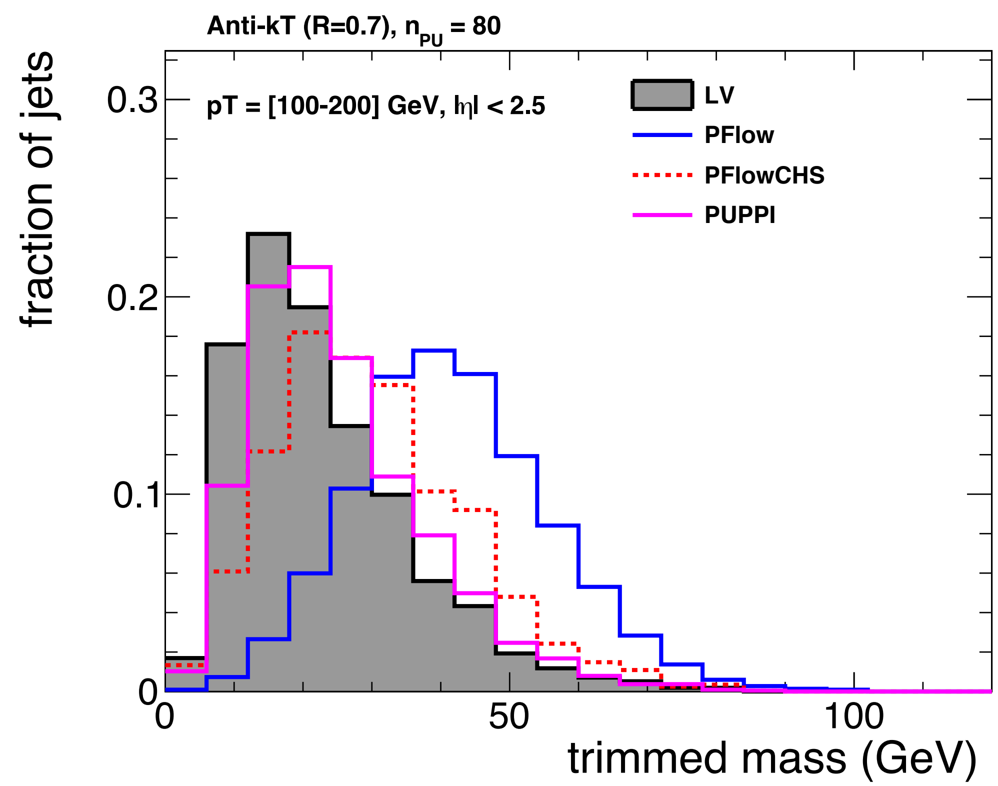
n_{PU} = 80 scenario

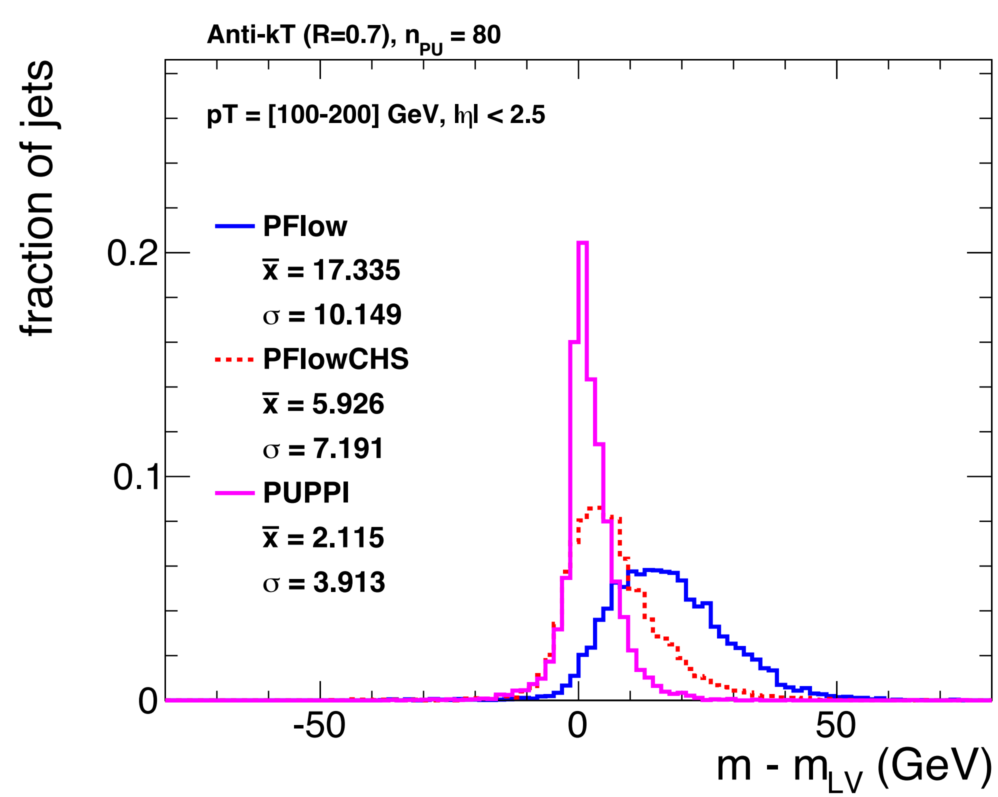
jet mass





trimmed jet mass

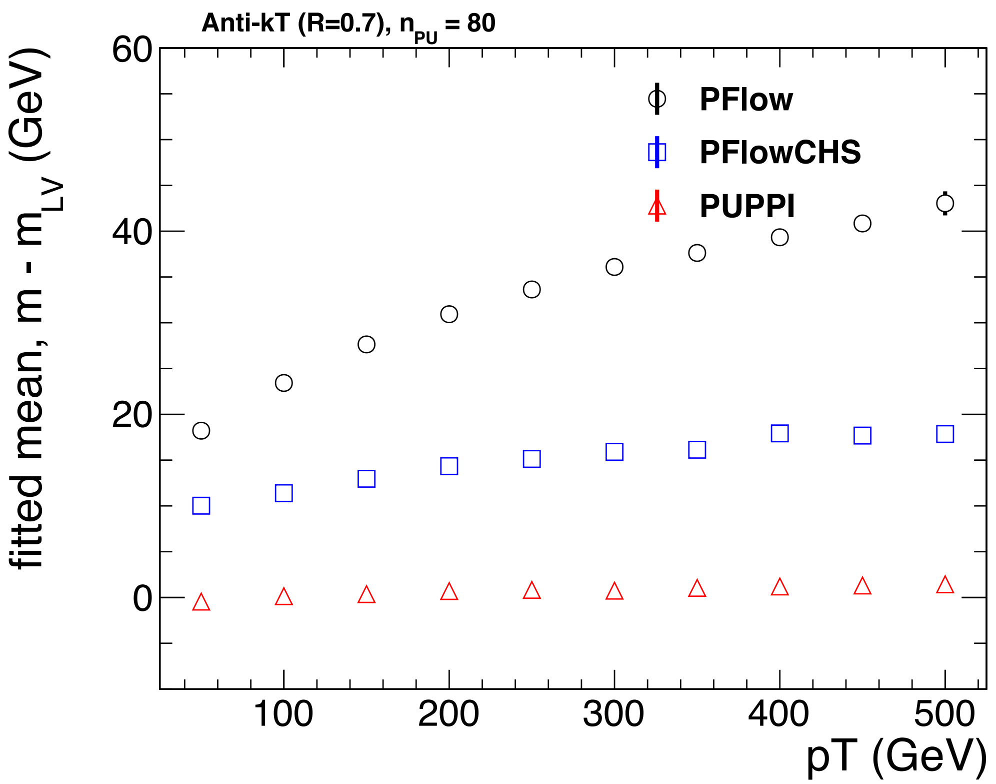




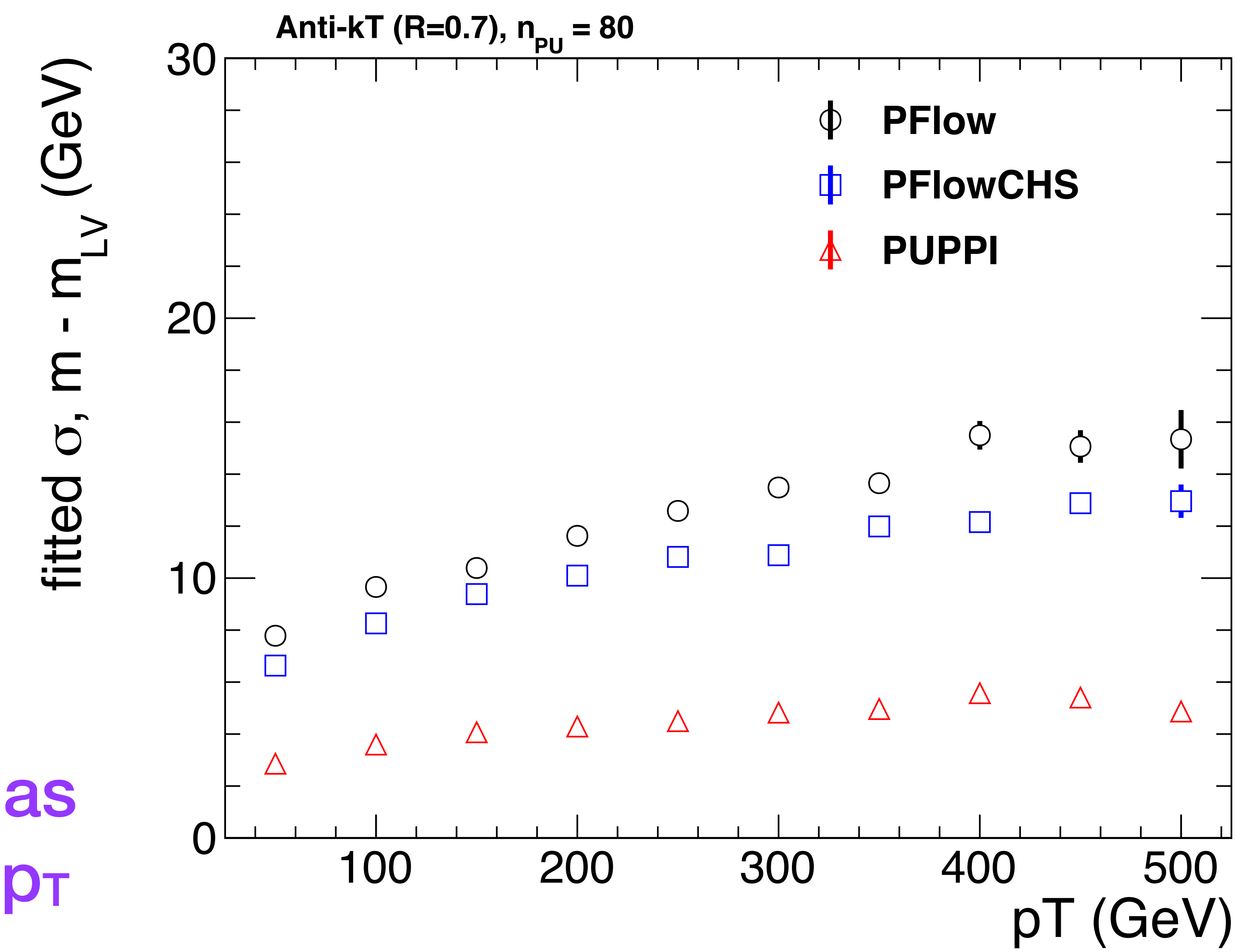
PF and PFCHS jets corrected doing trimming, then ρ 4-vector subtraction

jet mass resolution

n_{PU} = 80 scenario

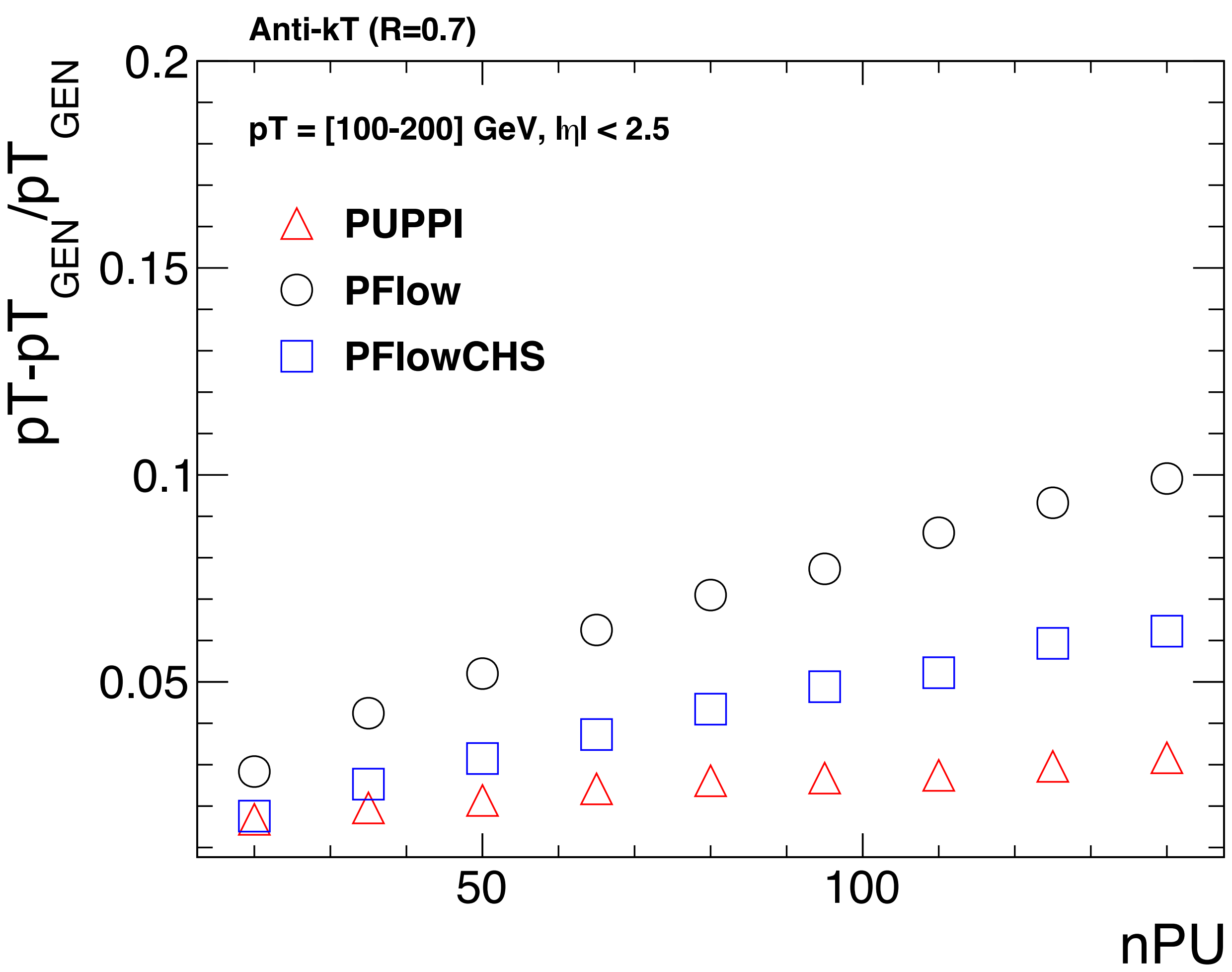


jet mass scale as a function of LV jet p_T

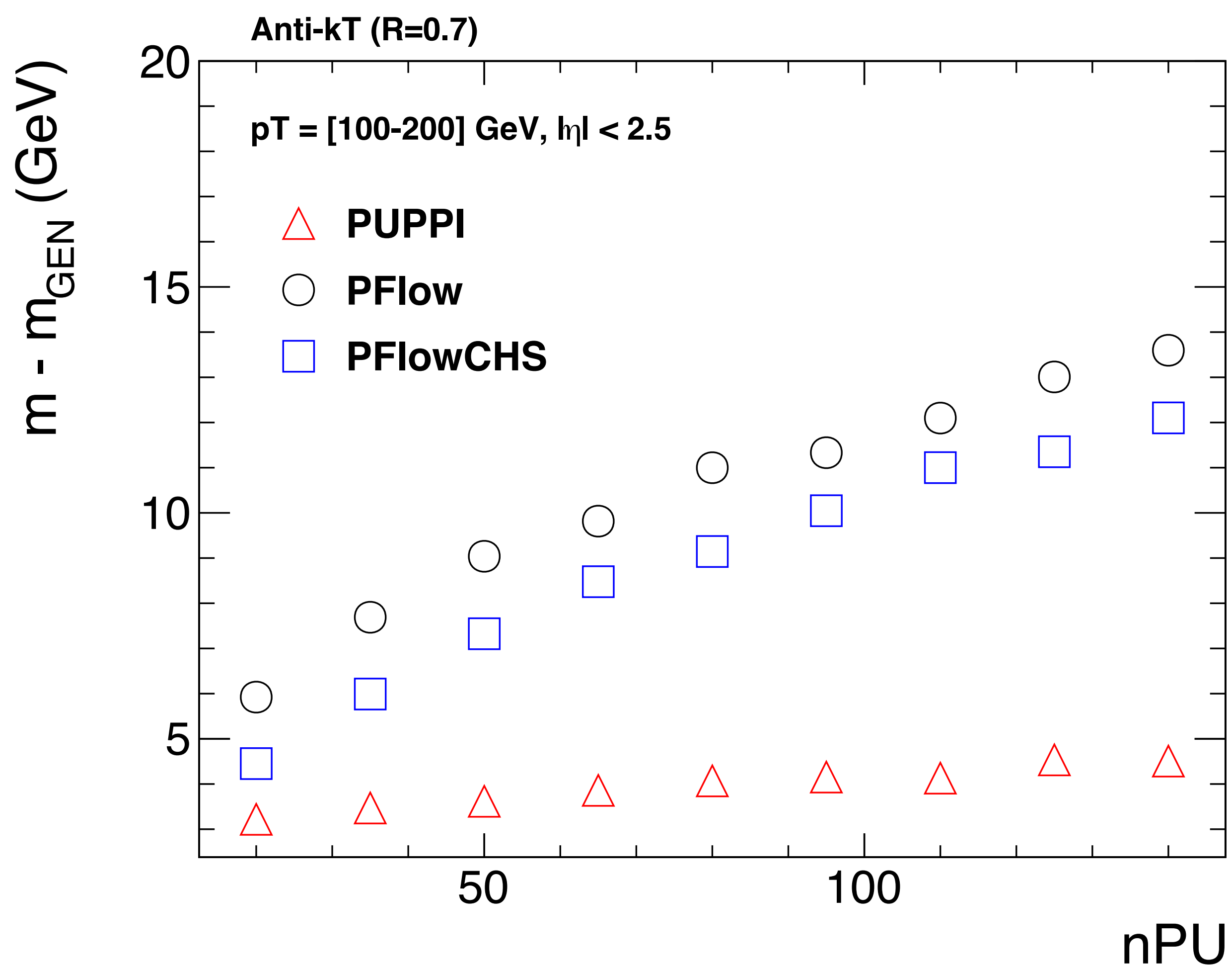


jet mass resolution as a function of LV jet p_T

results vs. PU



jet p_T resolution as a function of n_{PU}

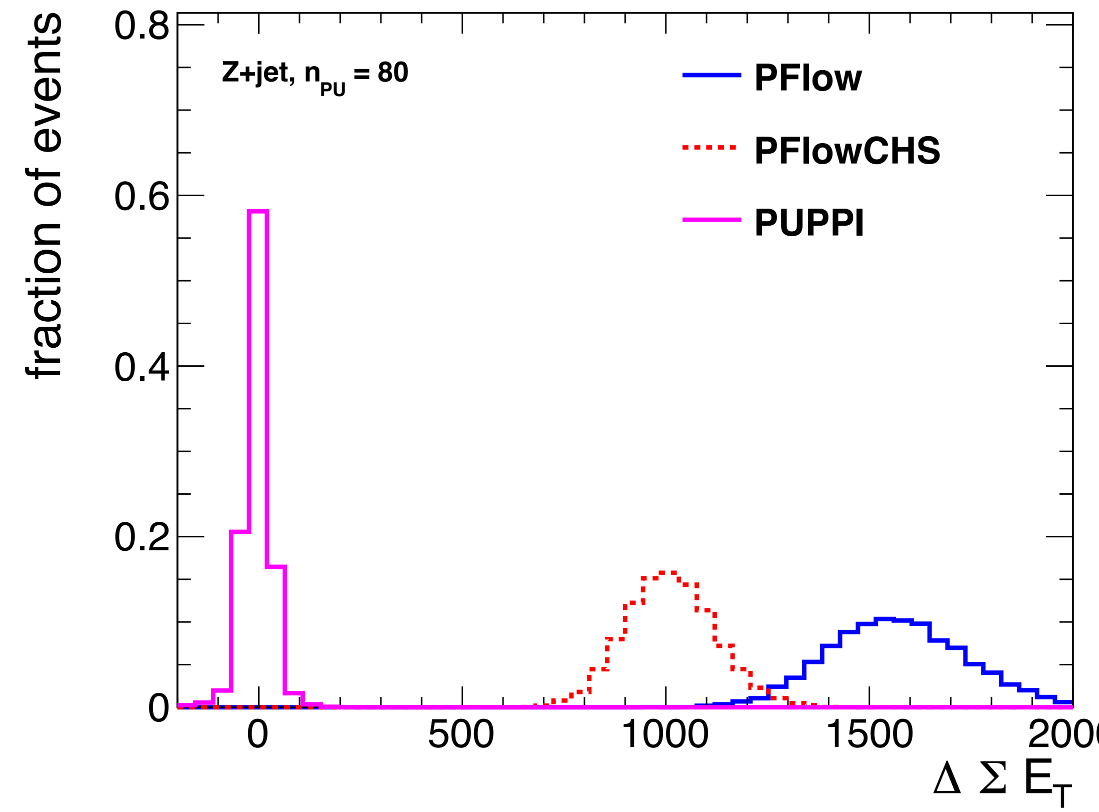


jet mass resolution as a function of n_{PU}

outline

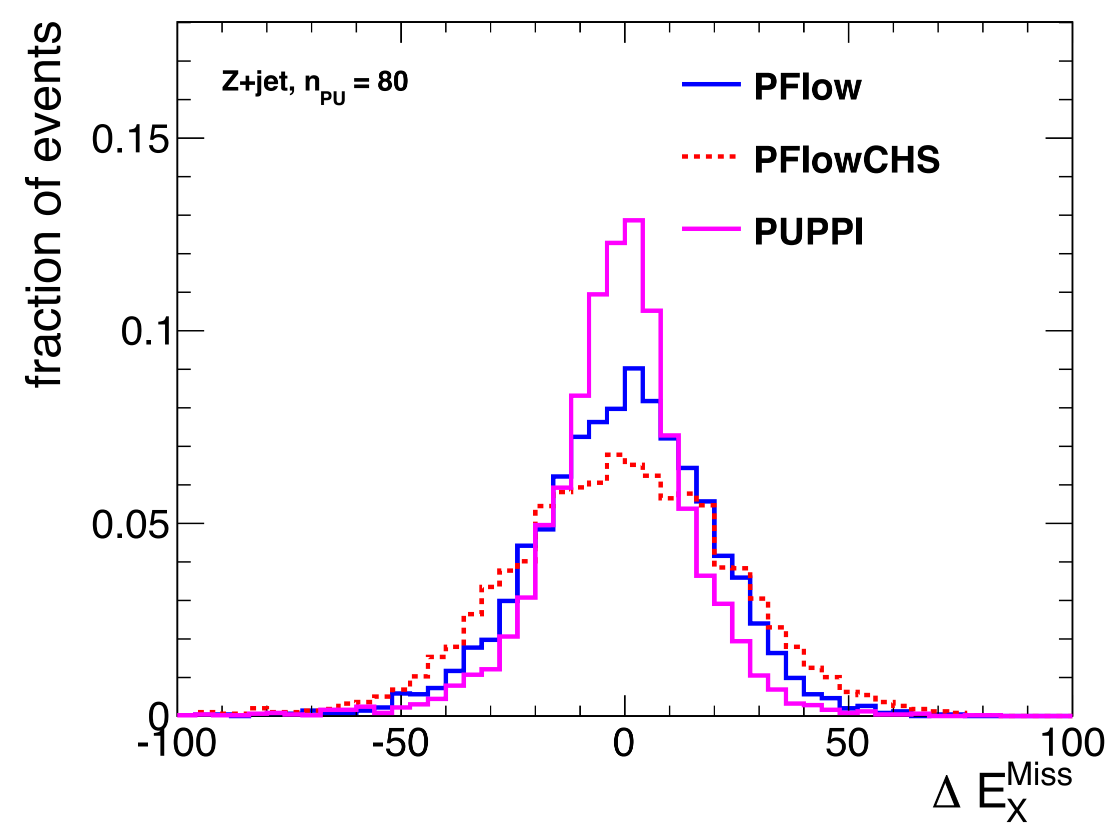
- what is PUPPI?
- defining the algorithm
- setup and results
- discussion and extensions

discussion and extensions



iPUPPI

isolation of leptons and photons using PUPPI-weight particles should have improved stability in high pileup environments



PUPP$\not{E}_T$

missing transverse energy (MET) determination using input particles from PUPPI algorithm should improve MET resolution and tails

discussion and extensions

not an algorithm set-in-stone, but rather a framework

- takes some ideas from existing algorithms, uses as much information as possible
- **classify** particles (or smallest detector units) ***before* clustering**
  - **can be used for lepton isolation and MET determination**
- can combine **experimental** and **theoretical** considerations
  - include probabilistic information about tracking, depth segmentation and timing

needs more rigorous comparisons against other methods

- *looking forward to working sessions!*

For the future…

would like to understand the definition of the metrics **more rigorously from the QCD** point-of-view

while α is IRC-unsafe under collinear splitting of particle i, the weight is IRC-safe

consider algorithm **outside of the PF framework**

i.e. in topoclustering, consider nearby tracks

extend jet shape studies beyond jet mass

the "M" word: MVA particle classification using several metrics shows modest improvement

summary

a new algorithm is presented for pileup mitigation at the LHC

aims to classify particles before jet clustering

- using local information, global event metrics and charged particle tracking
- can be extended to use further experimental information

showing good performance for jet p_T and mass resolution, comparing to ρ-subtracted PF and PFCHS

- *need to exercise in detector environment!*

not the final story, algorithm can be tuned for different experiments or improved further

thanks!

thank you to many people for helpful discussions!

We'd like to thank in particular Jeff Berryhill, Dinko Ferencek, Andrew Larkoski, David Krohn, Jesse Thaler, Lian Tao Wang + many CMS colleagues