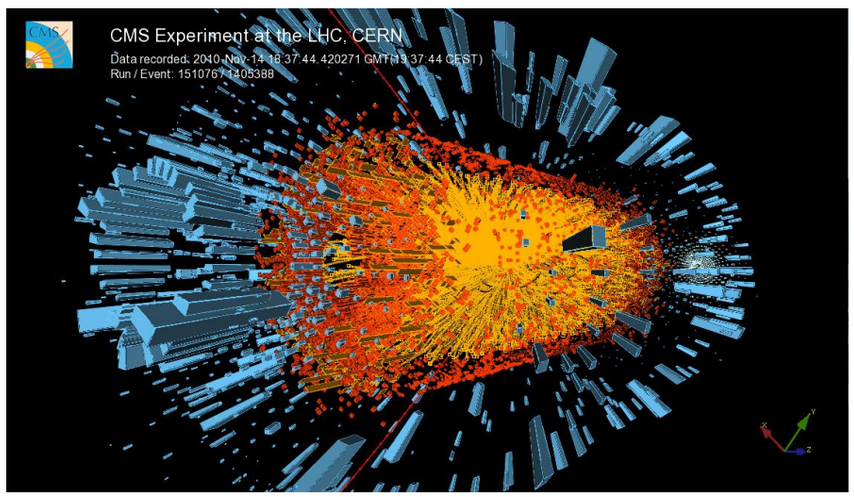
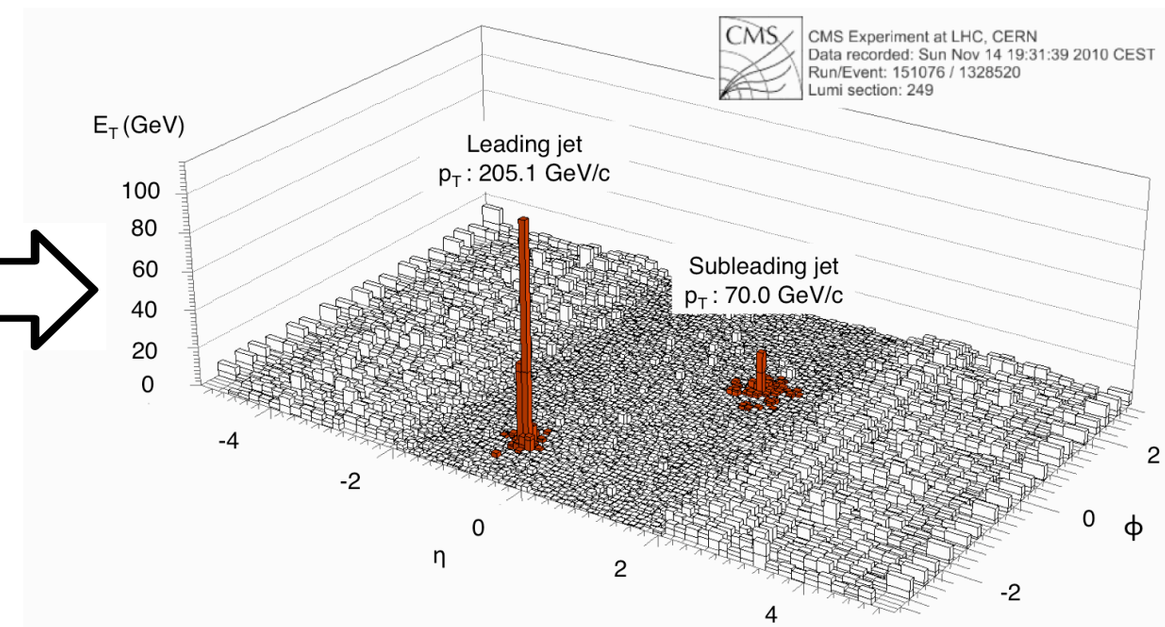
JETS AND BACKGROUND SUBTRACTION





Alba Soto-Ontoso

[Yacine Mehtar-Tani, ASO, Marta Verweij arXiv:1904.12815]

RHIC/AGS Users Meeting

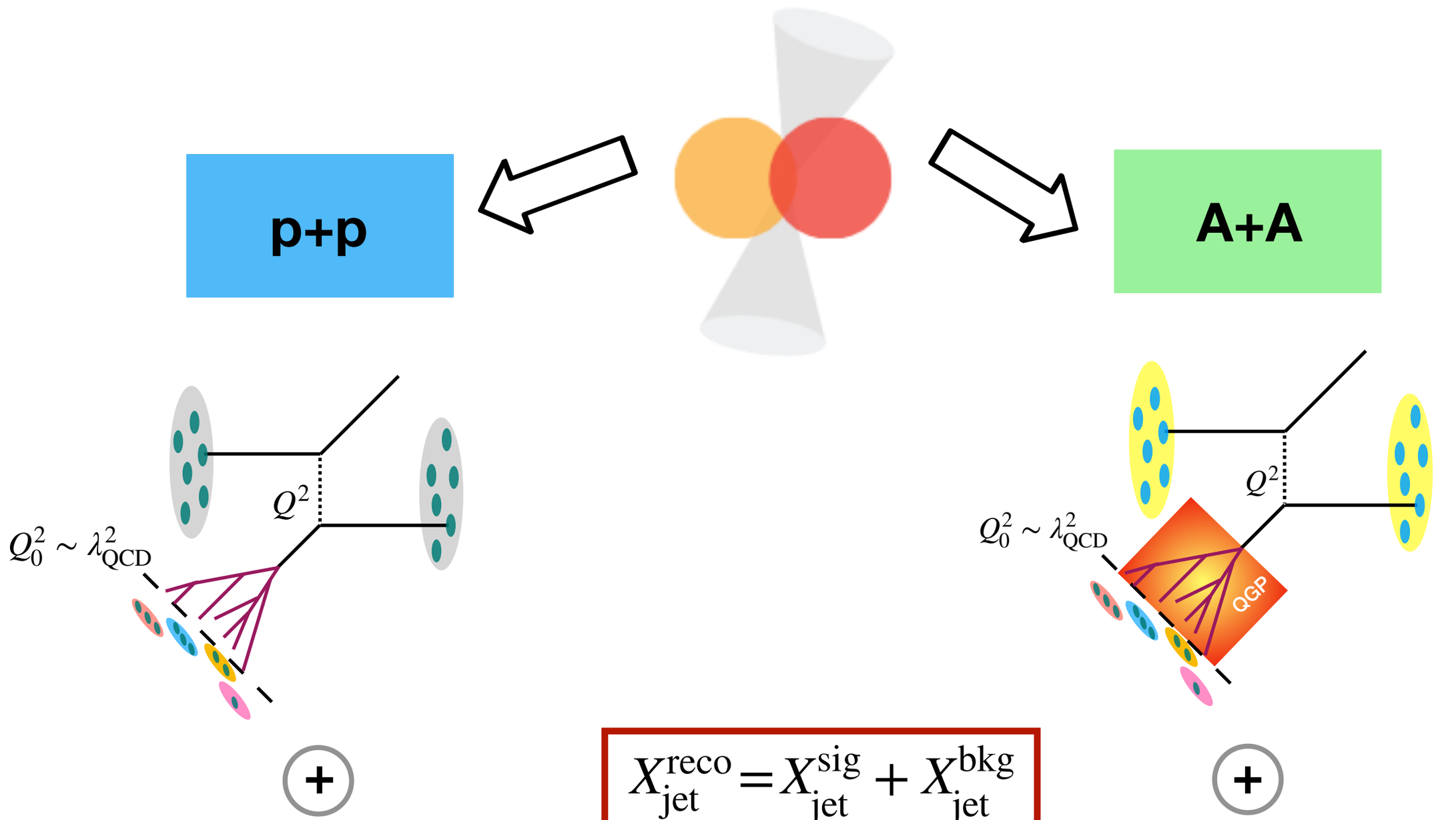
BNL, 4th June, 2019

BROOKHAVEN NATIONAL LABORATORY

Center for Frontiers in Nuclear Science

Jet reconstruction analyses

Cluster all particles in the event with k_T / anti-k_T / Cambridge-Aachen



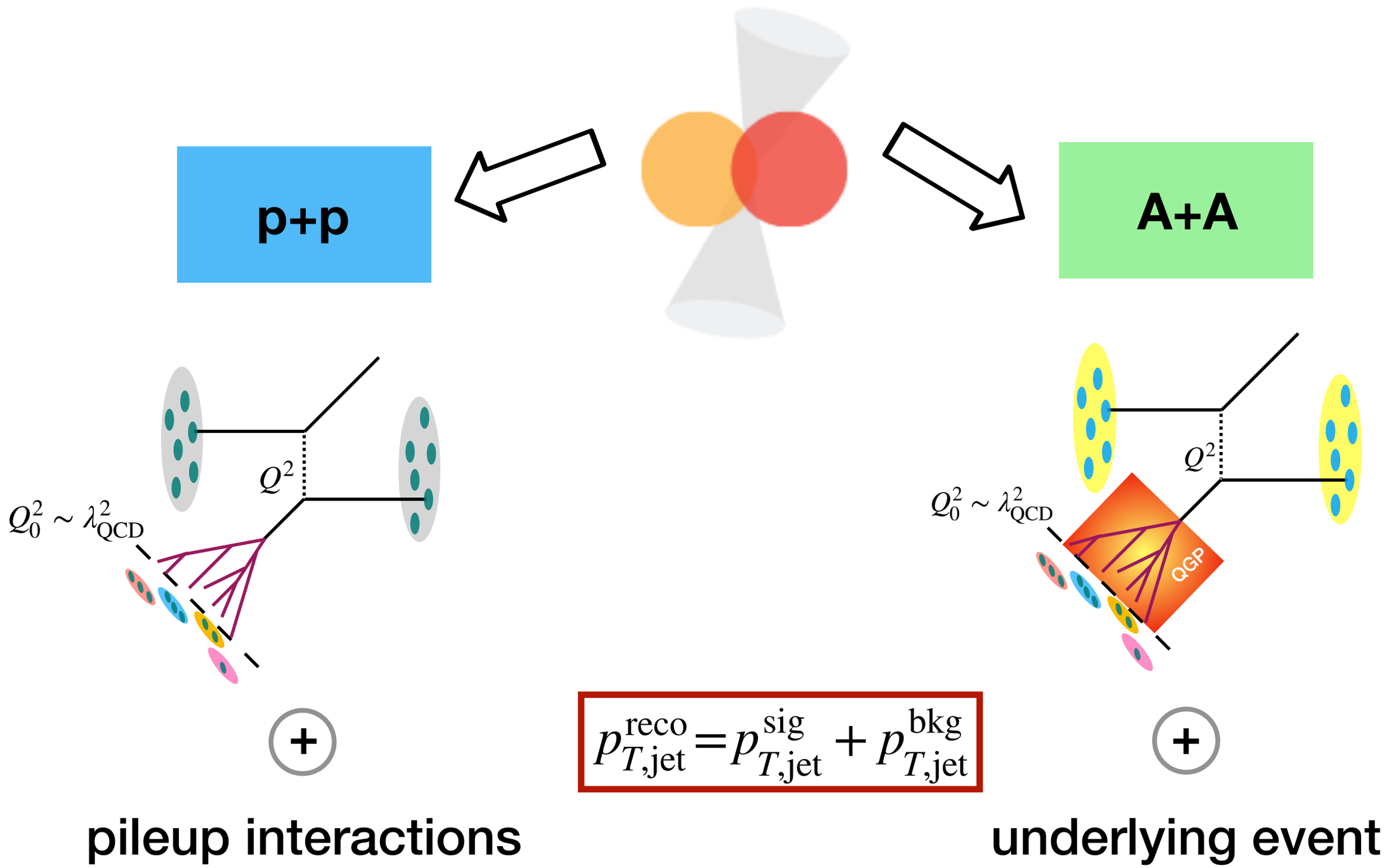
$$X_{\text{jet}}^{\text{reco}} = X_{\text{jet}}^{\text{sig}} + X_{\text{jet}}^{\text{bkg}}$$

pileup interactions

underlying event

Jet reconstruction analyses

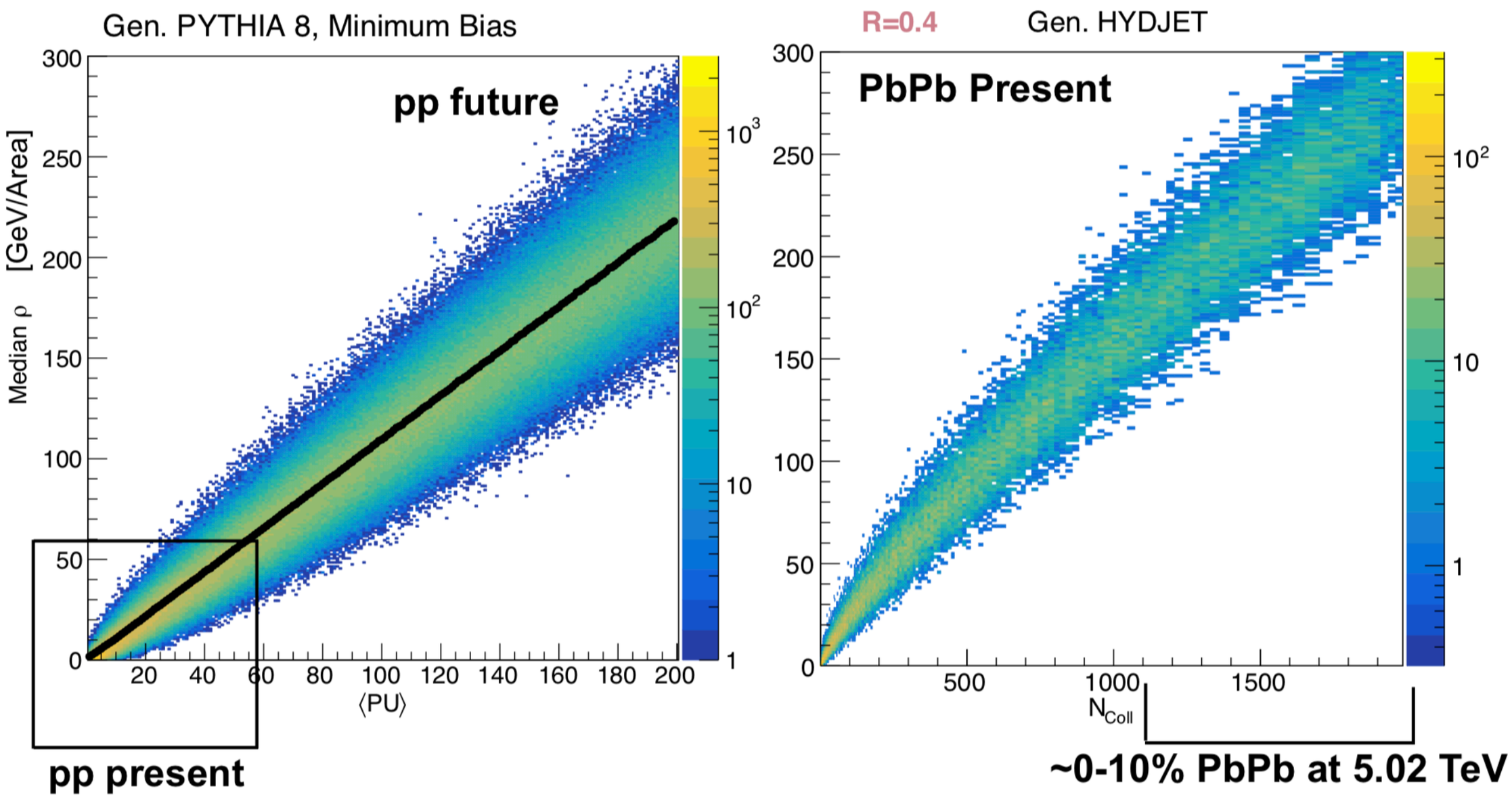
Cluster all particles in the event with k_T / anti-k_T / Cambridge-Aachen



$$p_{T,\mathrm{jet}}^{\mathrm{reco}} = p_{T,\mathrm{jet}}^{\mathrm{sig}} + p_{T,\mathrm{jet}}^{\mathrm{bkg}}$$

HL-LHC and HIC background contamination

Background momentum density per unit area $\rho = \frac{p_T^{\mathrm{bkg}}}{A}$



[C.McGinn Talk at "The Definition of Jets in Large Background"]

A two-step process

Background estimation $\xrightarrow{p_T^{\mathrm{bkg}}}$ **Subtraction method**

- Background fluctuates **event-by-event.**
- Fluctuations make any subtraction method inaccurate by definition.

- It should work best when the background is known exactly.
- Can be done at the particle level on the full **event** or **jet-by-jet.**

How to mitigate background fluctuations?

How to remove it?

Simulation setup for LHC

Background

- Randomly distributed in (η, ϕ)
- Thermal distribution with

This talk $\begin{cases} |\eta| < 3, \\ \langle N_{\text{part}} \rangle = 7000, 12000 \\ \mu \equiv \langle p_{\text{T}} \rangle = 1.2, 0.7 \text{ GeV/c} \end{cases}$

Signal

- Pythia8 dijet-event

This talk: $\langle \hat{p}_{\text{T}} \rangle = 100 \text{ GeV}$

- JEWEL w/ and w/o energy loss (no recoil)

This talk: $\langle \hat{p}_{\text{T}} \rangle = 150 \text{ GeV}$
$\sqrt{s} = 5.02 \text{ TeV}$
$0 - 10\,\%$

PYTHIA8 + Thermal background == p+p collision @ HL-LHC

JEWEL w/ E-loss + Thermal background == Pb+Pb collision @LHC

NEW: Simulation setup for RHIC

Background

- Randomly distributed in (η, ϕ)
- Thermal distribution with

This talk $\begin{cases} |\eta| < 3, \\ \langle N_{\text{part}} \rangle = 600 \\ \mu \equiv \langle p_{\text{T}} \rangle = 0.6 \text{ GeV/c} \end{cases}$

Signal

- JEWEL w/ and w/o energy loss (no recoil)

This talk: $\langle \hat{p}_{\text{T}} \rangle = 10 \text{ GeV}$

$\sqrt{s} = 200 \text{ GeV}$

$0 - 5\,\%$

JEWEL w/ E-loss + Thermal background == Au+Au collision @RHIC

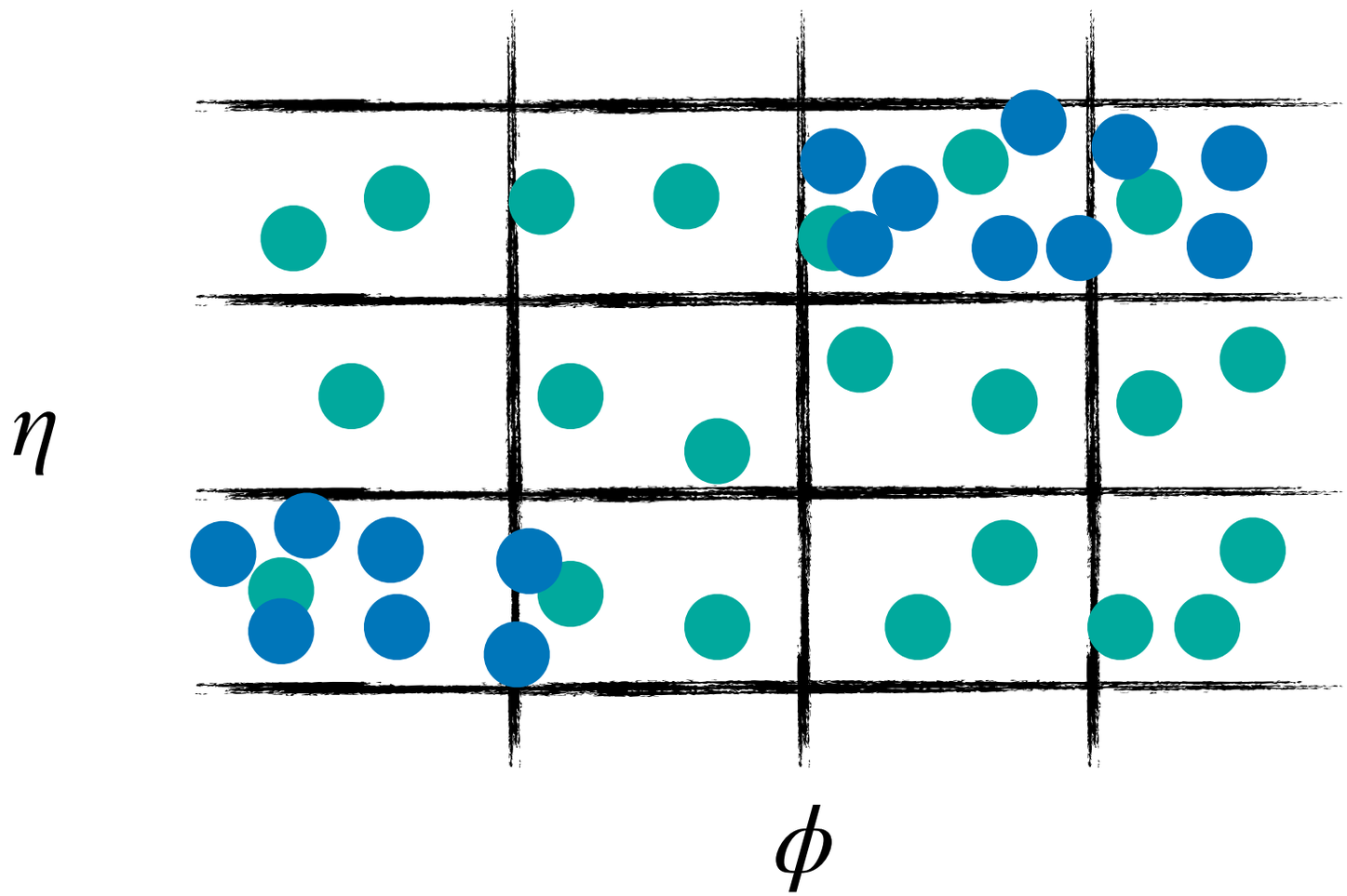
Samples available at **https://github.com/JetQuenchingTools/JetToyHI**

BACKGROUND ESTIMATION METHODS

Area-median: an unbiased method

[M. Cacciari, G.Salam Phys. Lett. B659 (2008) 119–126]

Cluster all particles (sig+bkg) in the event with **k_T algorithm**

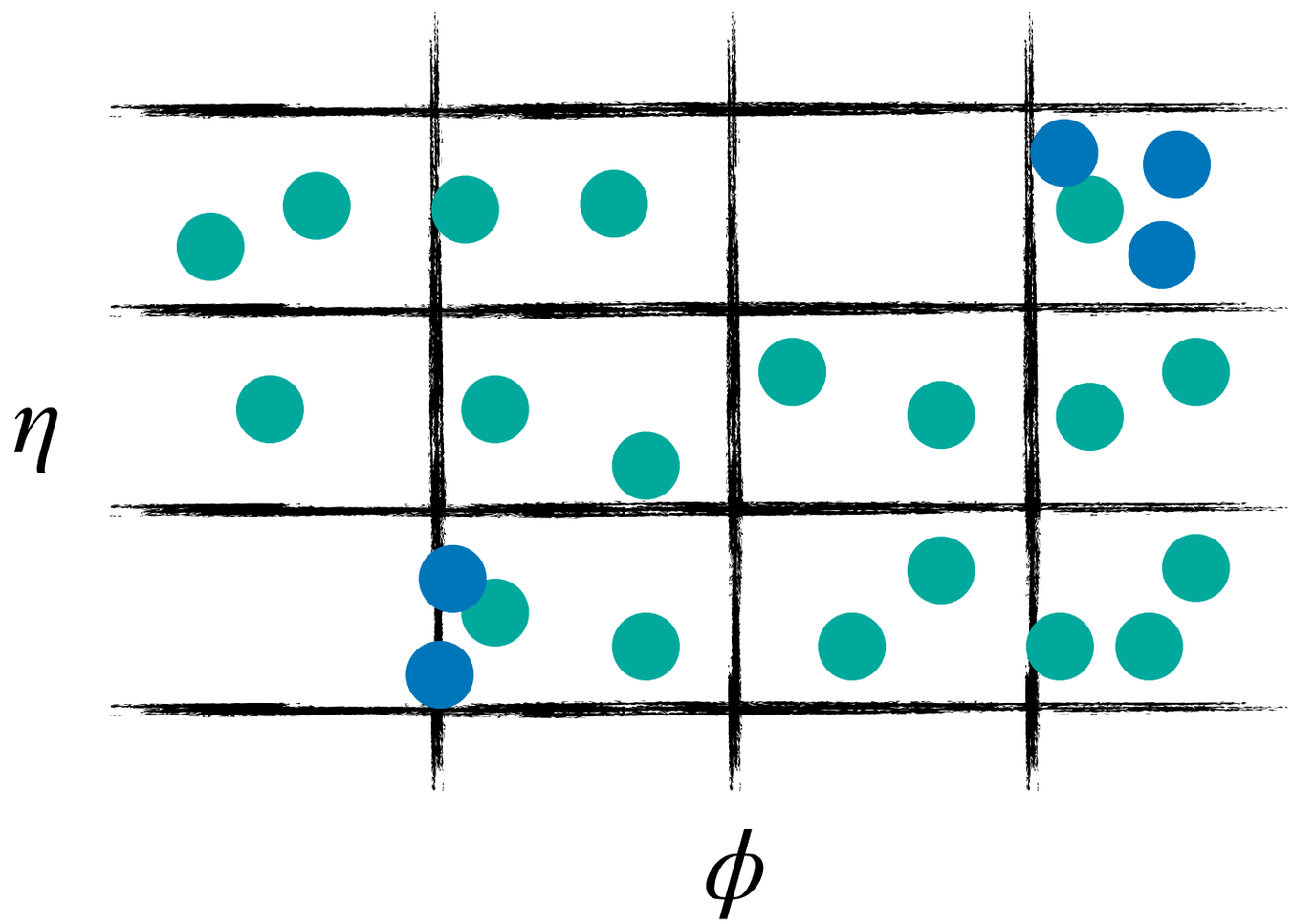


Area-median: an unbiased method

[M. Cacciari, G.Salam Phys. Lett. B659 (2008) 119–126]

Cluster all particles (sig+bkg) in the event with **k_T algorithm**

Remove the **2 hardest** jets from the sample



For every patch, i, compute

$$\rho_i = \frac{p_{\mathrm{T},i}}{A_i}$$

Area-median: an unbiased method

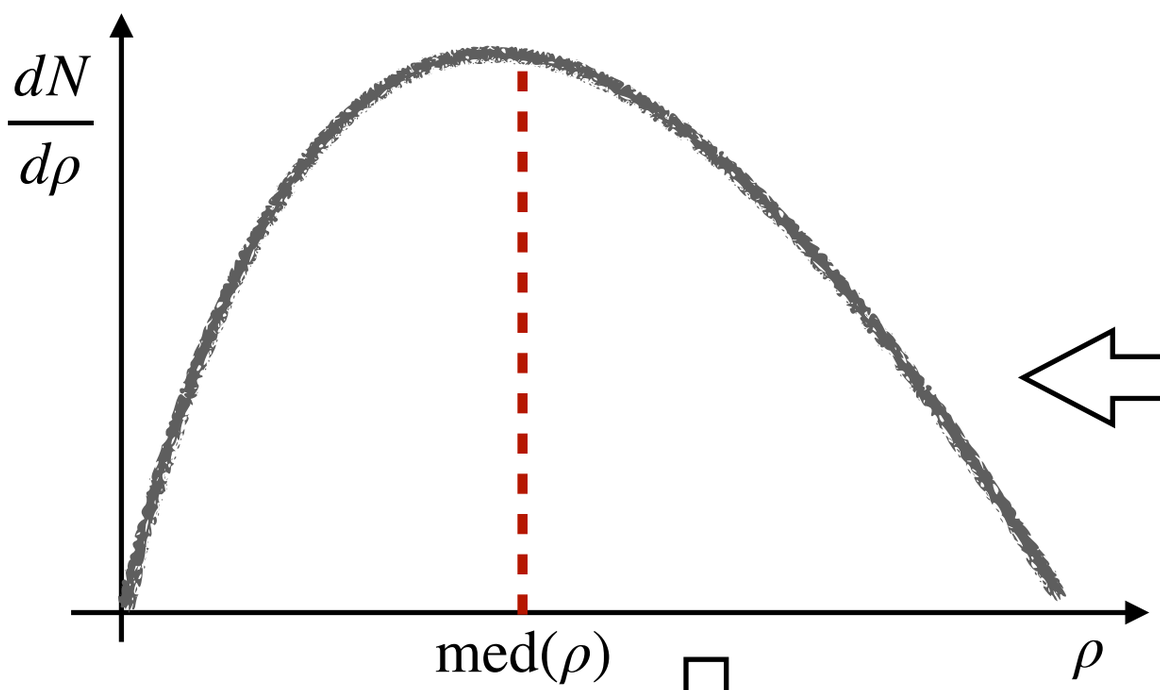
[M. Cacciari, G.Salam Phys. Lett. B659 (2008) 119–126]

Cluster all particles (sig+bkg) in the event with **k_T algorithm**

Remove the **2 hardest** jets from the sample

For every patch, i, compute

$$\rho_i = \frac{p_{\mathrm{T},i}}{A_i}$$



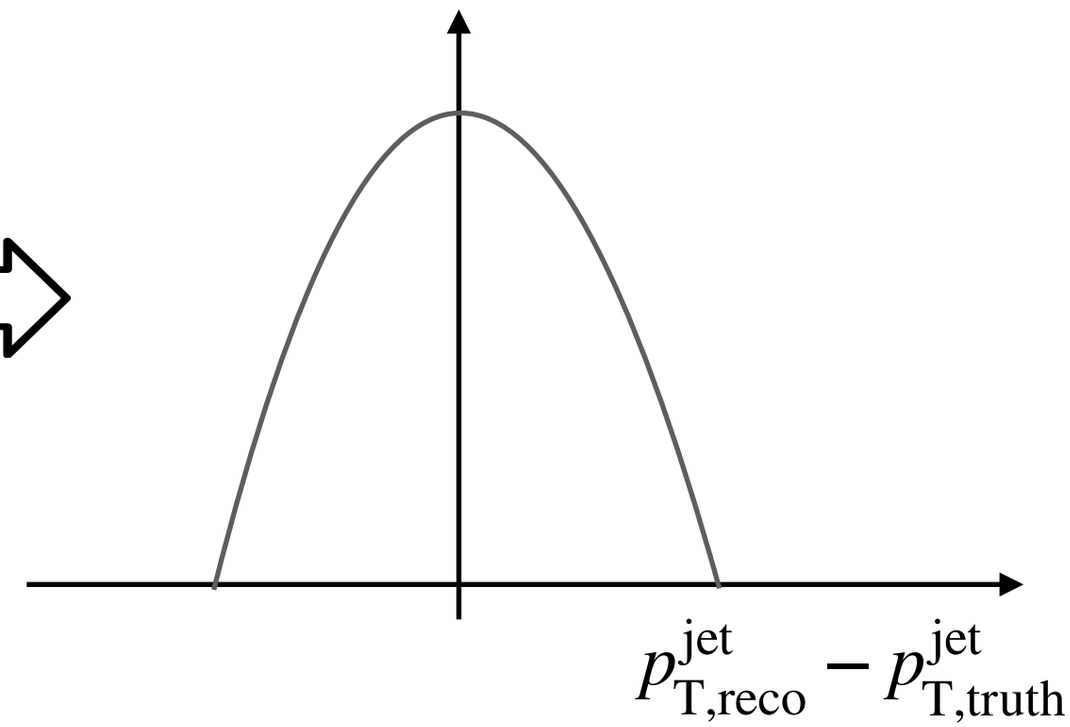
Cluster all particles (sig+bkg) in the event with **anti-k_T algorithm**

$$p_{\mathrm{T,reco}}^{\mathrm{jet}} = p_{\mathrm{T,raw}}^{\mathrm{jet}} - \mathrm{med}(\rho) A_{\mathrm{raw}}^{\mathrm{jet}}$$

Area-median: an unbiased method

[M. Cacciari, G.Salam Phys. Lett. B659 (2008) 119–126]

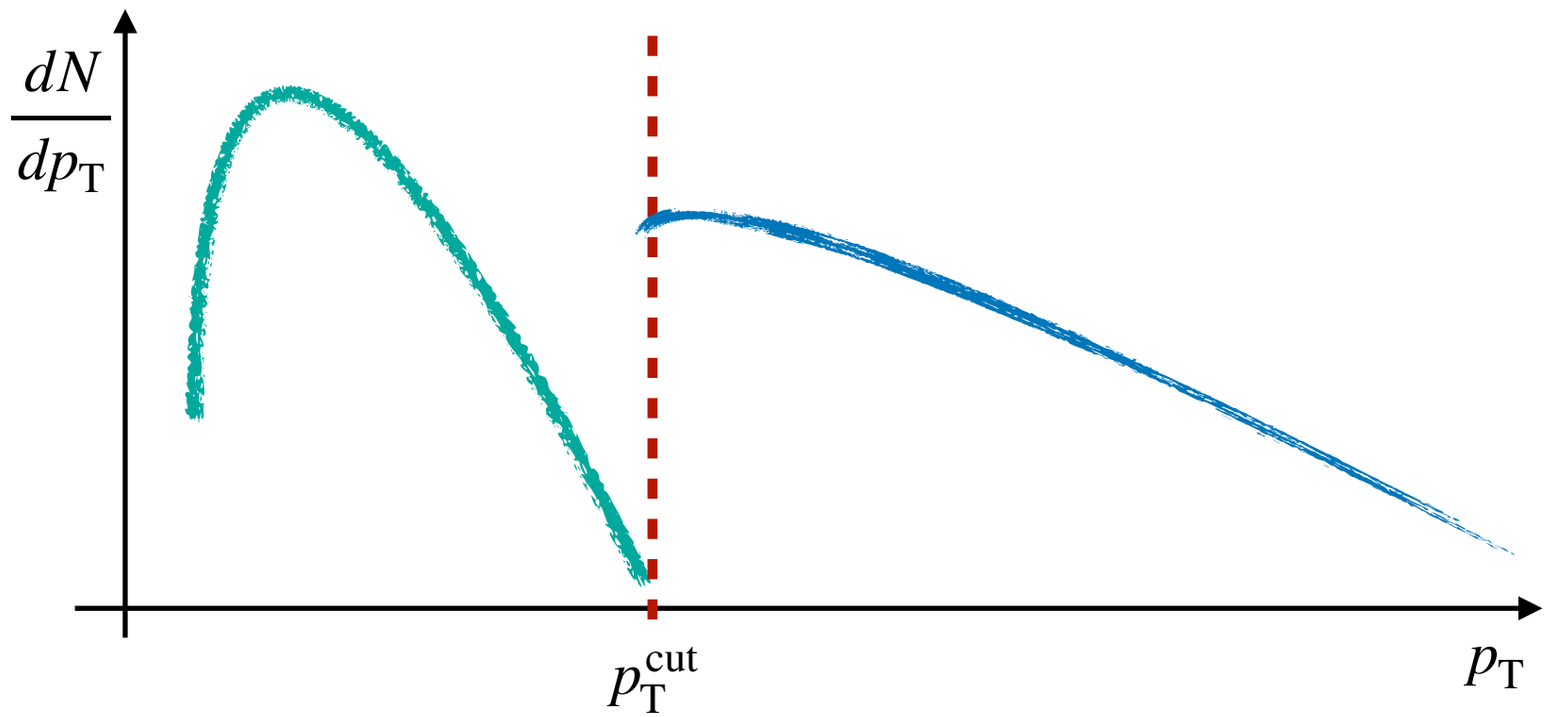
$$p_{\mathrm{T,reco}}^{\mathrm{jet}} = p_{\mathrm{T,raw}}^{\mathrm{jet}} - \mathrm{med}(\rho) A_{\mathrm{raw}}^{\mathrm{jet}}$$

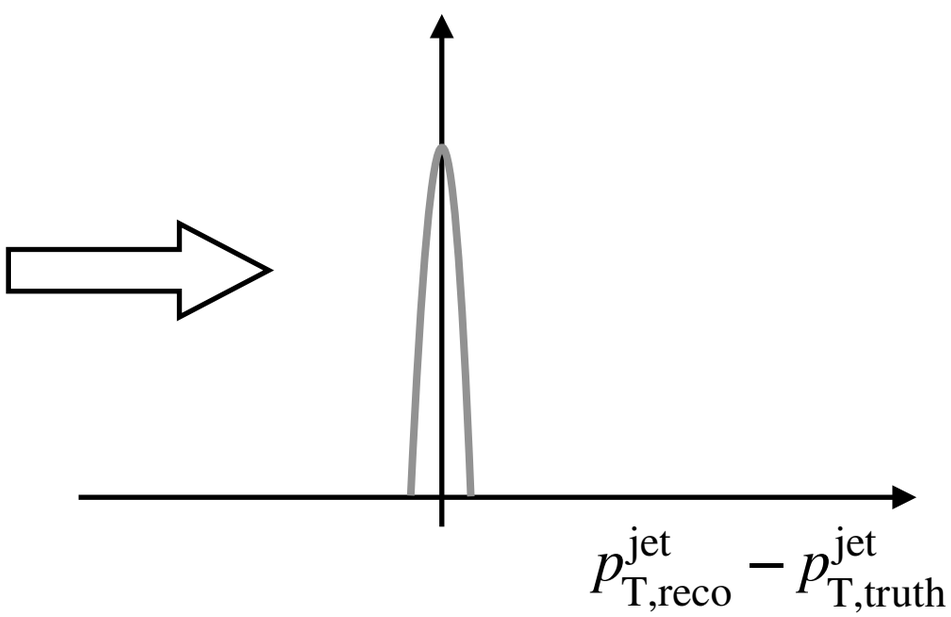


- Data-driven, unbiased, correct $\langle p_{\mathrm{T,reco}}^{\mathrm{jet}} - p_{\mathrm{T,truth}}^{\mathrm{jet}} \rangle$
- No control over background fluctuations

SoftKiller: introducing a p_T-cut

[M. Cacciari et al. Eur.Phys.J. C75 (2015) no.2, 59]





If there is no signal below the p_T-cut, exact handle on background fluctuations

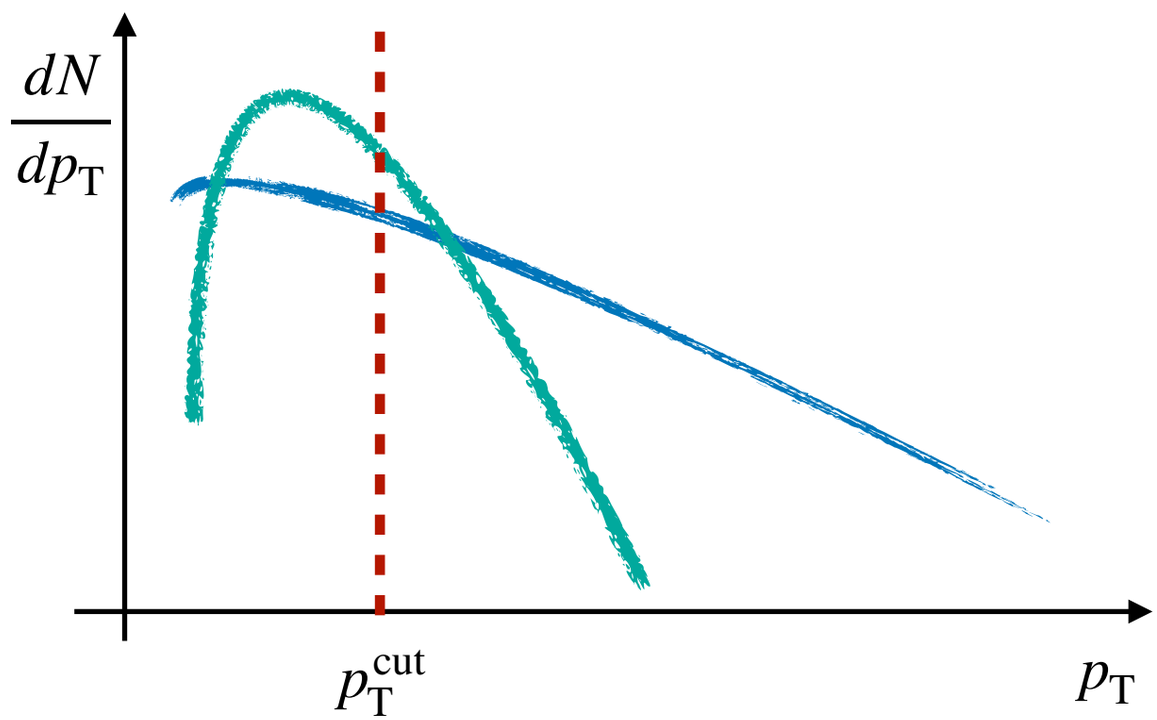
$$\langle p_{T,reco}^{jet} - p_{T,truth}^{jet} \rangle = 0$$

$$\sigma(p_{T,reco}^{jet} - p_{T,truth}^{jet}) = 0$$

SoftKiller: introducing a p_T-cut

[M. Cacciari et al. Eur.Phys.J. C75 (2015) no.2, 59]

In reality, signal and background overlap in momentum space



How to find the value of p_T-cut?

SoftKiller: introducing a p_T-cut

[M. Cacciari et al. Eur.Phys.J. C75 (2015) no.2, 59]

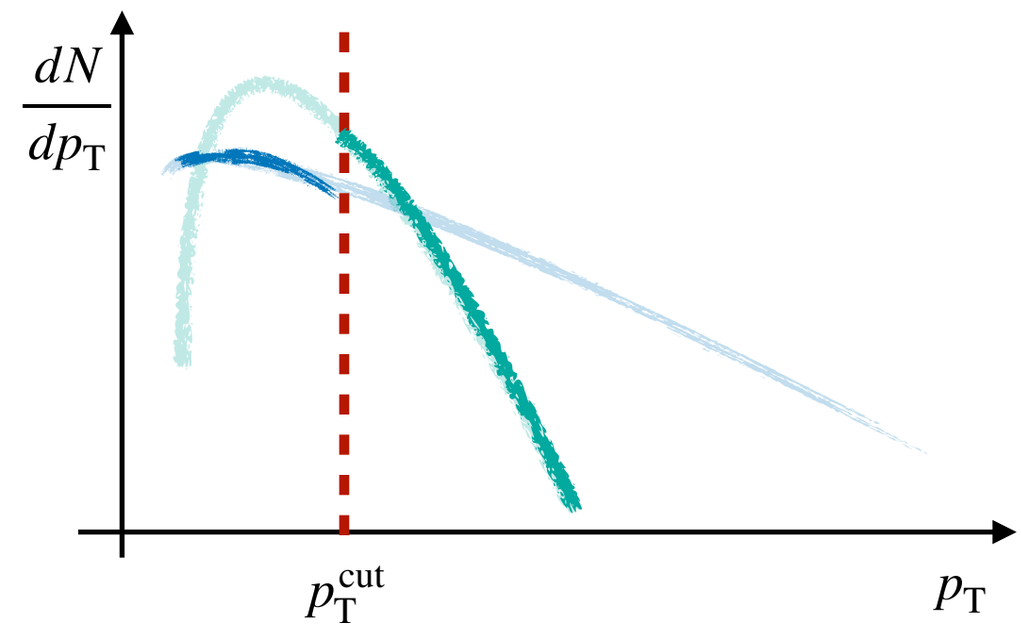
- Choose p_T-cut such that $p_{T,<}^{\mathrm{sig}}$ and $p_{T,>}^{\mathrm{bkg}}$ balance each other

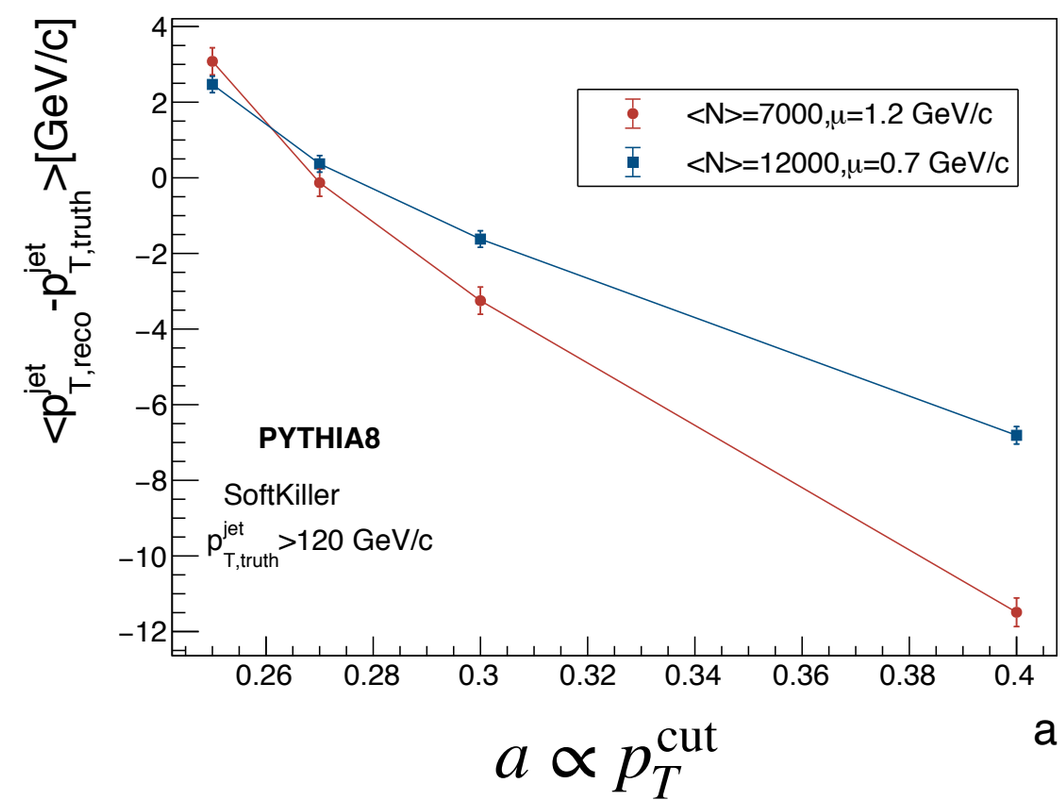
$$\langle p_{T,\mathrm{reco}}^{\mathrm{jet}} - p_{T,\mathrm{truth}}^{\mathrm{jet}} \rangle \sim 0$$

$$p_{T,\mathrm{reco}}^{\mathrm{jet}} = p_{T,\mathrm{raw}}^{\mathrm{jet}} - \sum_{i}^{p_{T,i}<p_T^{\mathrm{cut}}} p_{T,i}$$

- But…this actually requires a large degree of fine tuning

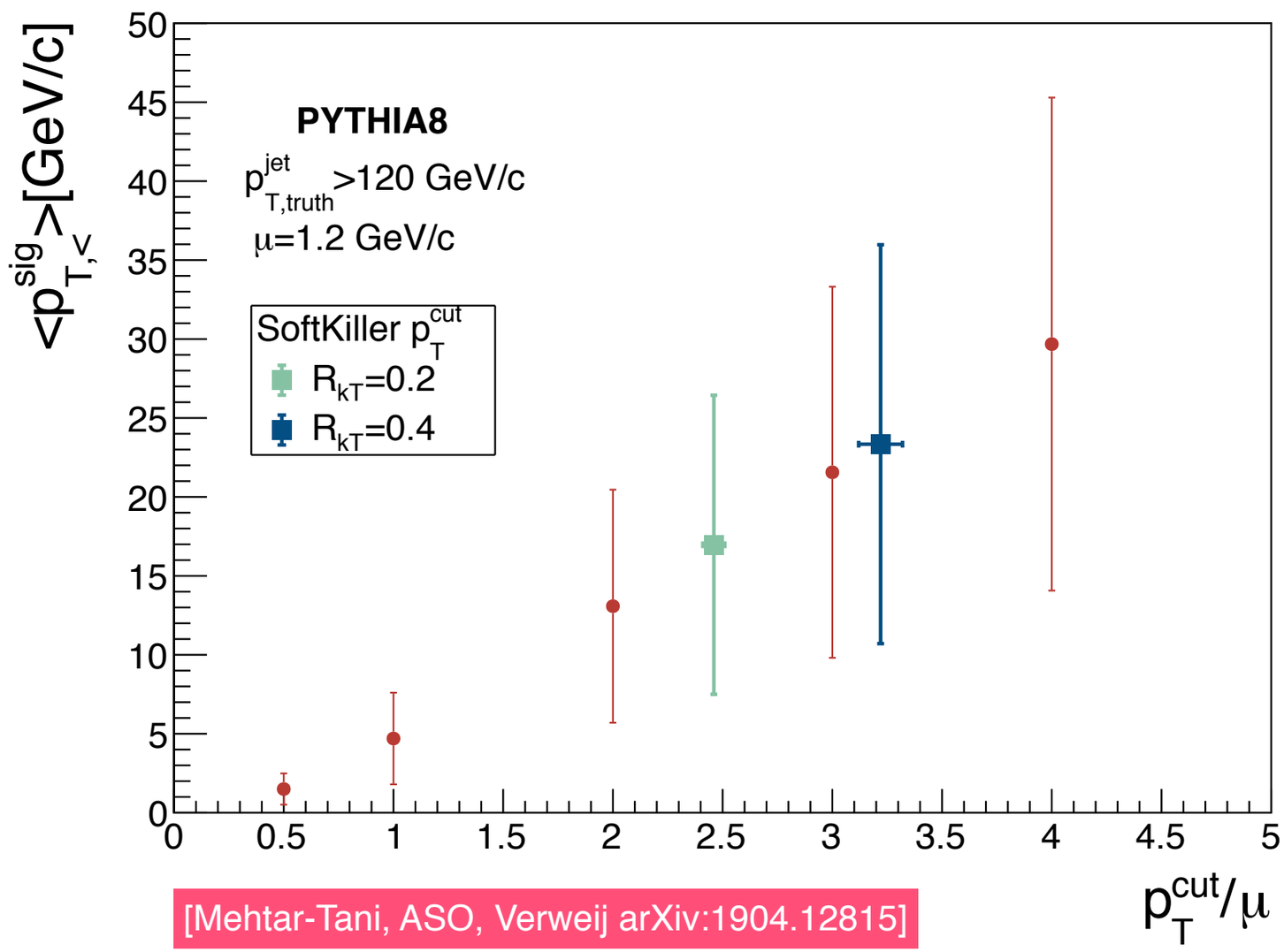
"SoftKiller R^{adj}"





Impact of signal contamination

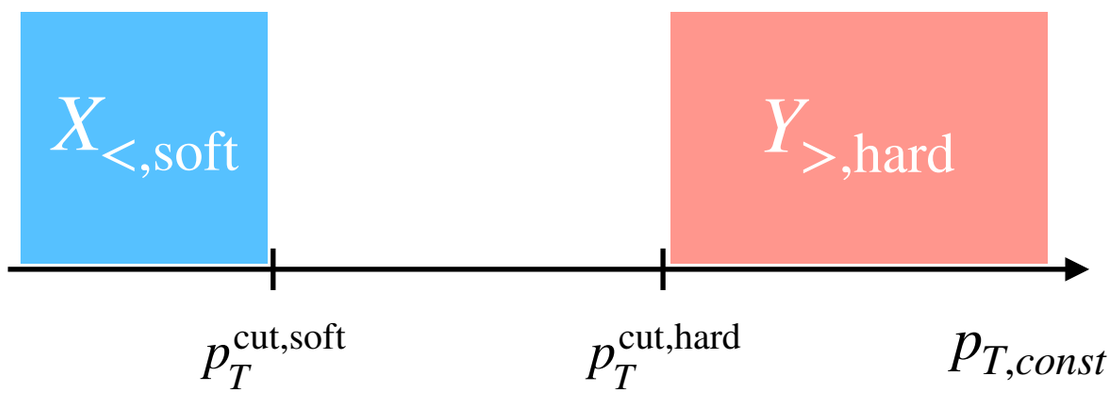
Any background estimator that relies on a soft momentum cut is sensitive to QCD radiation below this threshold

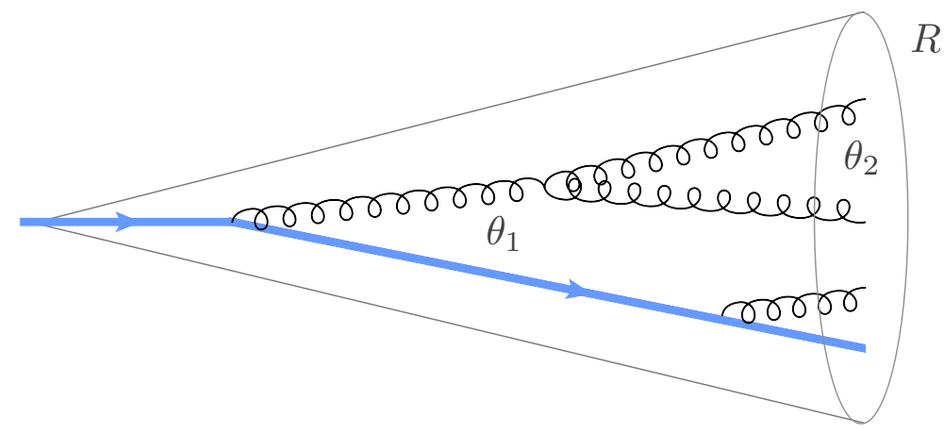


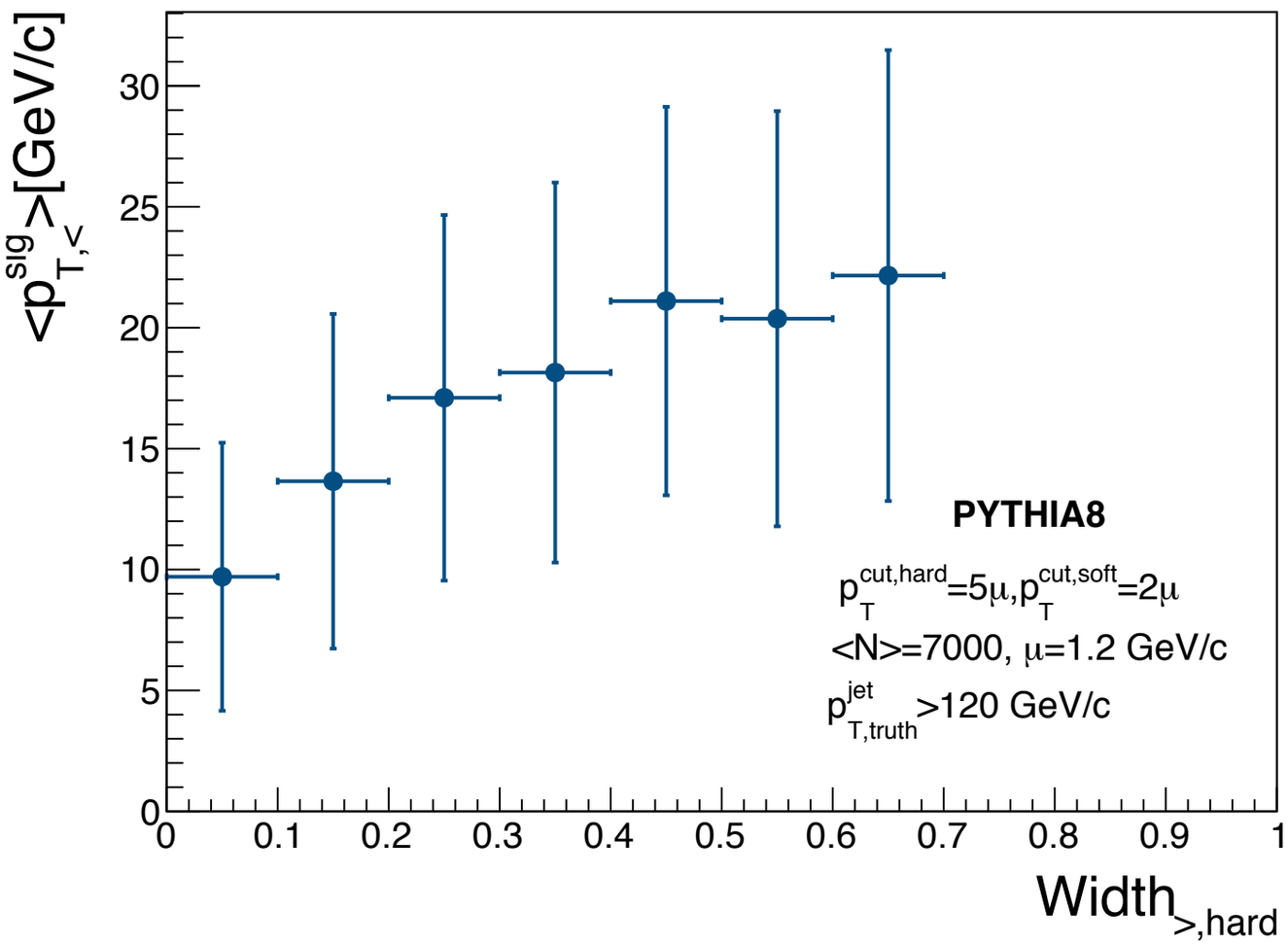
[Mehtar-Tani, ASO, Verweij arXiv:1904.12815]

$$p^{\mathrm{sig}}_{T,<} = \sum_{i \in \mathrm{sig}}^{p_{T,i} < p^{\mathrm{cut}}_{\mathrm{T}}} p_{T,i}$$

Hard-soft correlations within a QCD jet



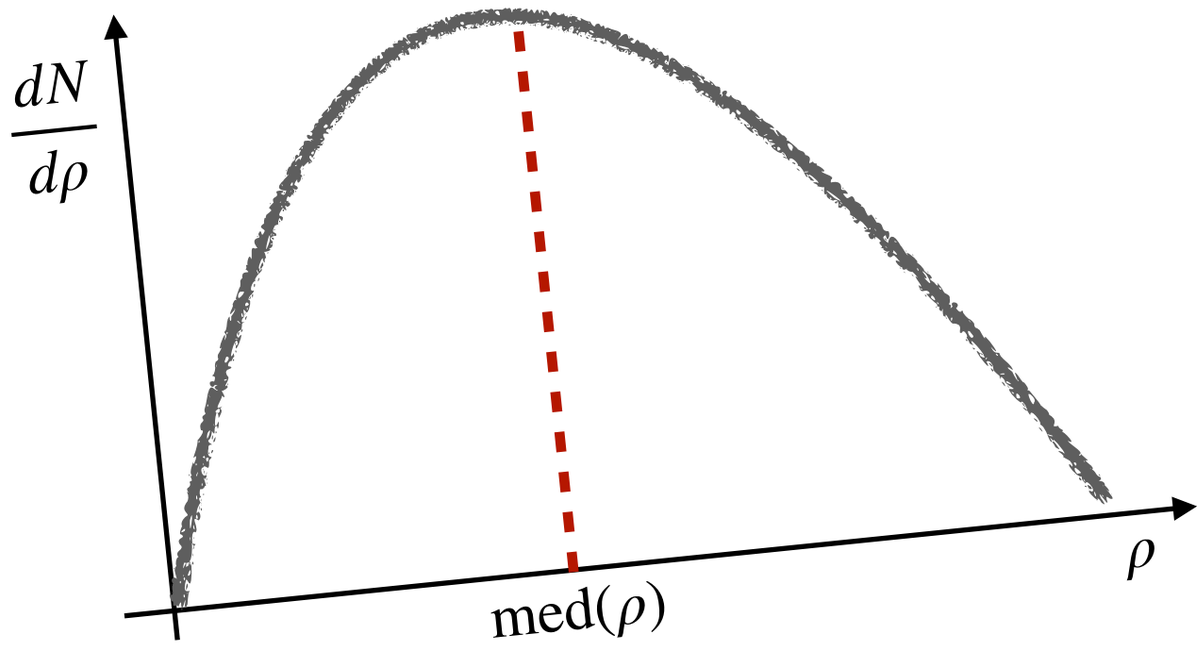


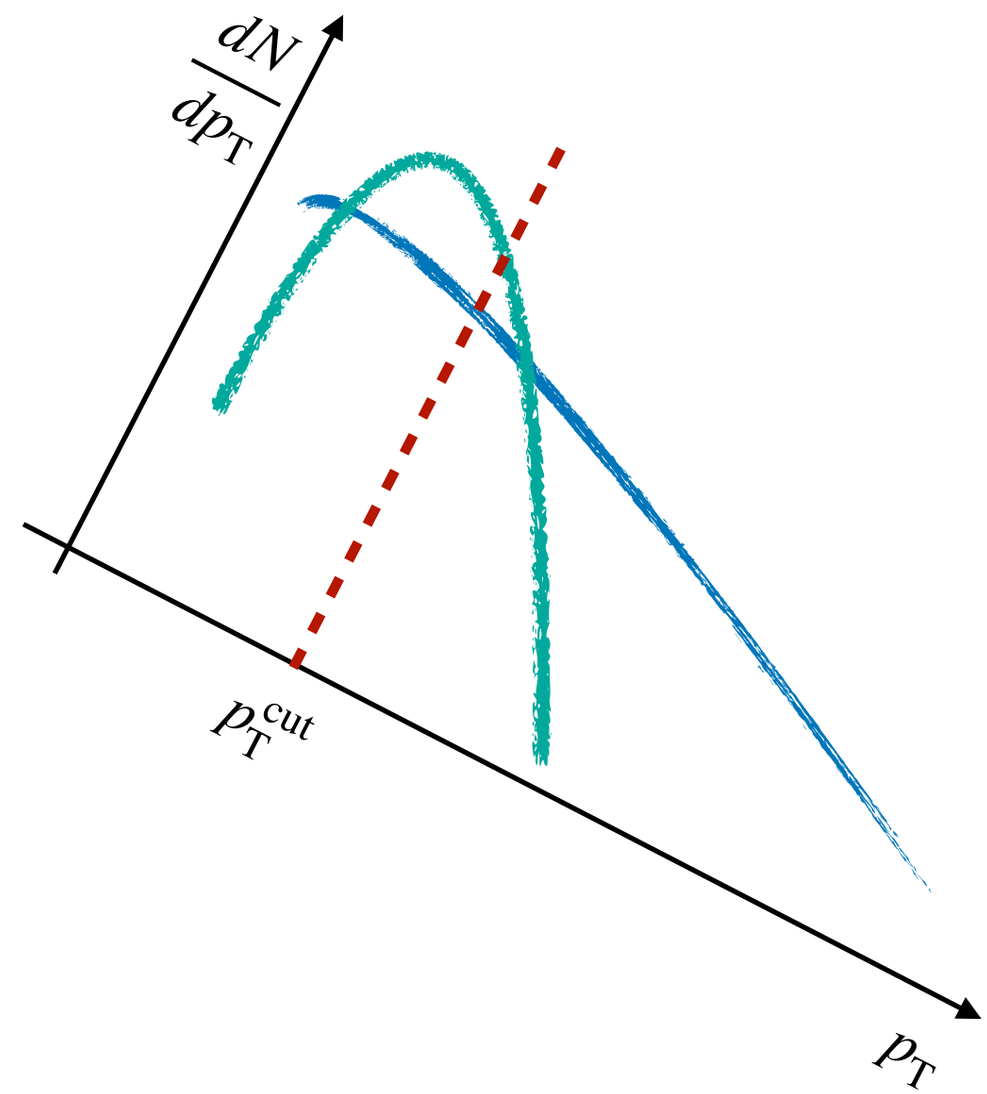


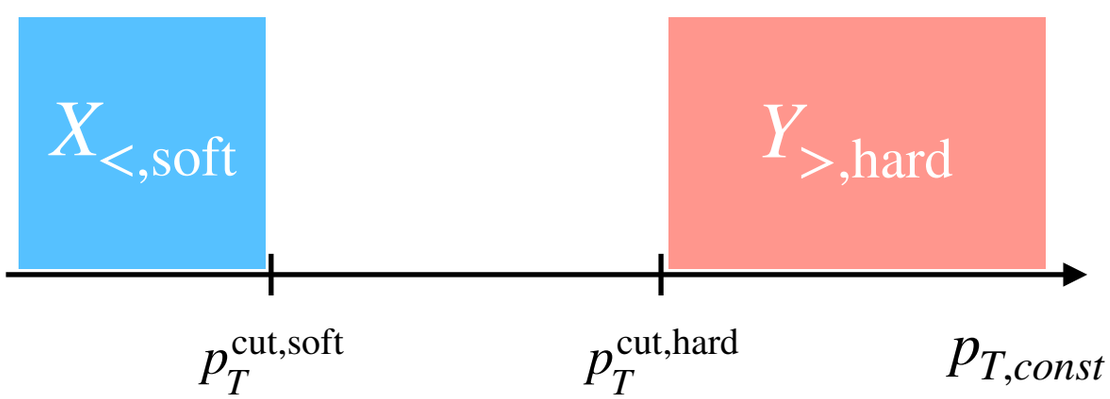
Signal fluctuations in the soft p_T-region can be mitigated by knowing the width in the hard sector. What about other metrics? Machine learning?

The ρ-correction approach

[Mehtar-Tani, ASO, Verweij arXiv:1904.12815]







The ρ-correction approach

[Mehtar-Tani, ASO, Verweij arXiv:1904.12815]

1 Cluster all particles (sig+bkg) in the event with **k_T algorithm, R=0.4** ⇒ Remove the **2 hardest** patches from the sample

For each patch, $\rho_> = \frac{1}{A} \sum_i^{p_T > p_T^{\mathrm{cut,soft}}} p_{T,i}$ ⇒ $\mathrm{med}(\rho_>)$

2 Re-cluster the whole event (sig+bkg) in the event with **anti-k_T algorithm, R=0.4**

For each jet, $p_{T,<} = \sum_i^{p_T < p_T^{\mathrm{cut,soft}}} p_{T,i}$

3 For a given $p_T^{\mathrm{cut,soft}}$ measure $\langle p_{T,<}^{\mathrm{sig}} \rangle$ in low pileup proton-proton collisions

The ρ-correction approach

[Mehtar-Tani, ASO, Verweij arXiv:1904.12815]

$$p_{\mathrm{T,reco}}^{\mathrm{jet}} = p_{\mathrm{T,raw}}^{\mathrm{jet}} - \mathrm{med}(\rho_{>})A_{\mathrm{raw}}^{\mathrm{jet}} - p_{T,<} + \langle p_{T,<}^{\mathrm{sig}} \rangle$$

- Neglecting $\langle p_{T,<}^{\mathrm{sig}} \rangle$-term, $\begin{cases} p_{\mathrm{T}}^{\mathrm{cut,soft}} = 0 : \text{Area-median} \\ p_{\mathrm{T}}^{\mathrm{cut,soft}} = p_{\mathrm{T}}^{\mathrm{cut,SK}} : \text{SoftKiller } R_{kT} \end{cases}$

- Role of $\langle p_{T,<}^{\mathrm{sig}} \rangle$: correct for the signal contamination below $p_{\mathrm{T}}^{\mathrm{cut,soft}}$
  - If measured inclusively, ρ-correction
  - If measured as a function of $\mathrm{width}_{>,\mathrm{hard}}$, ρ-correction HS_{corr}

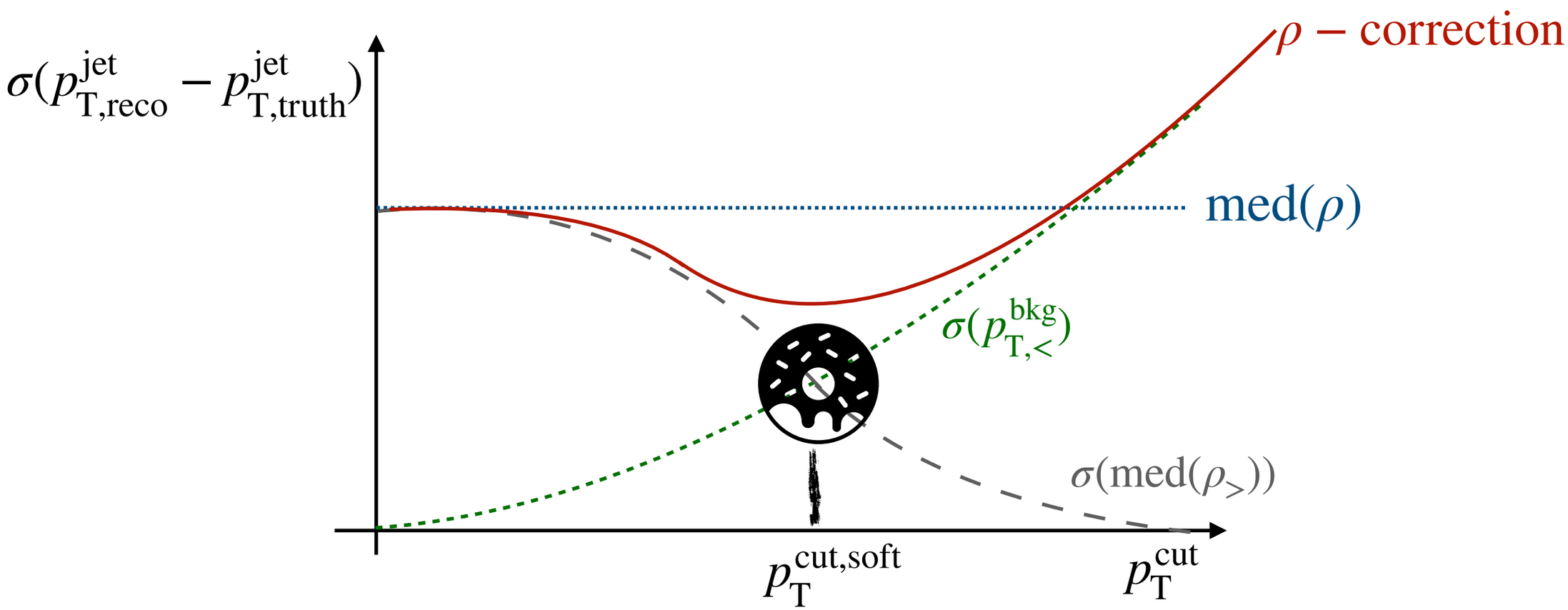
Value of $p_T^{\mathrm{cut,hard}}$? Large enough to be in the background-free region

$$p_T^{\mathrm{cut,hard}} = 5\mu$$ (μ : average momentum of bkg constituents)

The ρ-correction approach

[Mehtar-Tani, ASO, Verweij arXiv:1904.12815]

- Prescription to choose $p_T^{\mathrm{cut,soft}}$: minimize $\sigma(p_{\mathrm{T,reco}}^{\mathrm{jet}} - p_{\mathrm{T,truth}}^{\mathrm{jet}})$

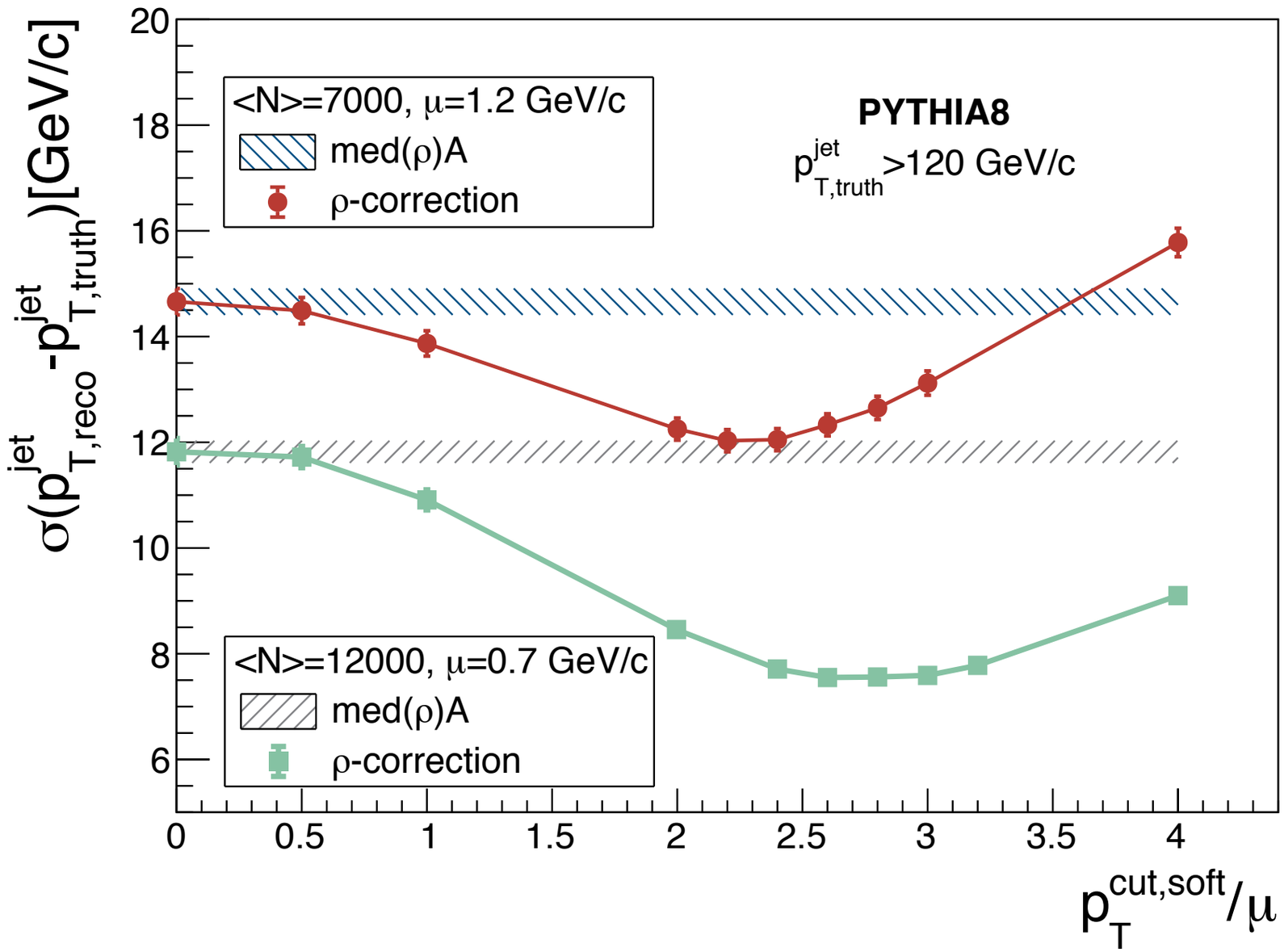


$$p_{\mathrm{T,reco}}^{\mathrm{jet}} = p_{\mathrm{T,raw}}^{\mathrm{jet}} - \mathrm{med}(\rho_>) A_{\mathrm{raw}}^{\mathrm{jet}} - p_{T,<} + \langle p_{T,<}^{\mathrm{sig}} \rangle$$

RESULTS

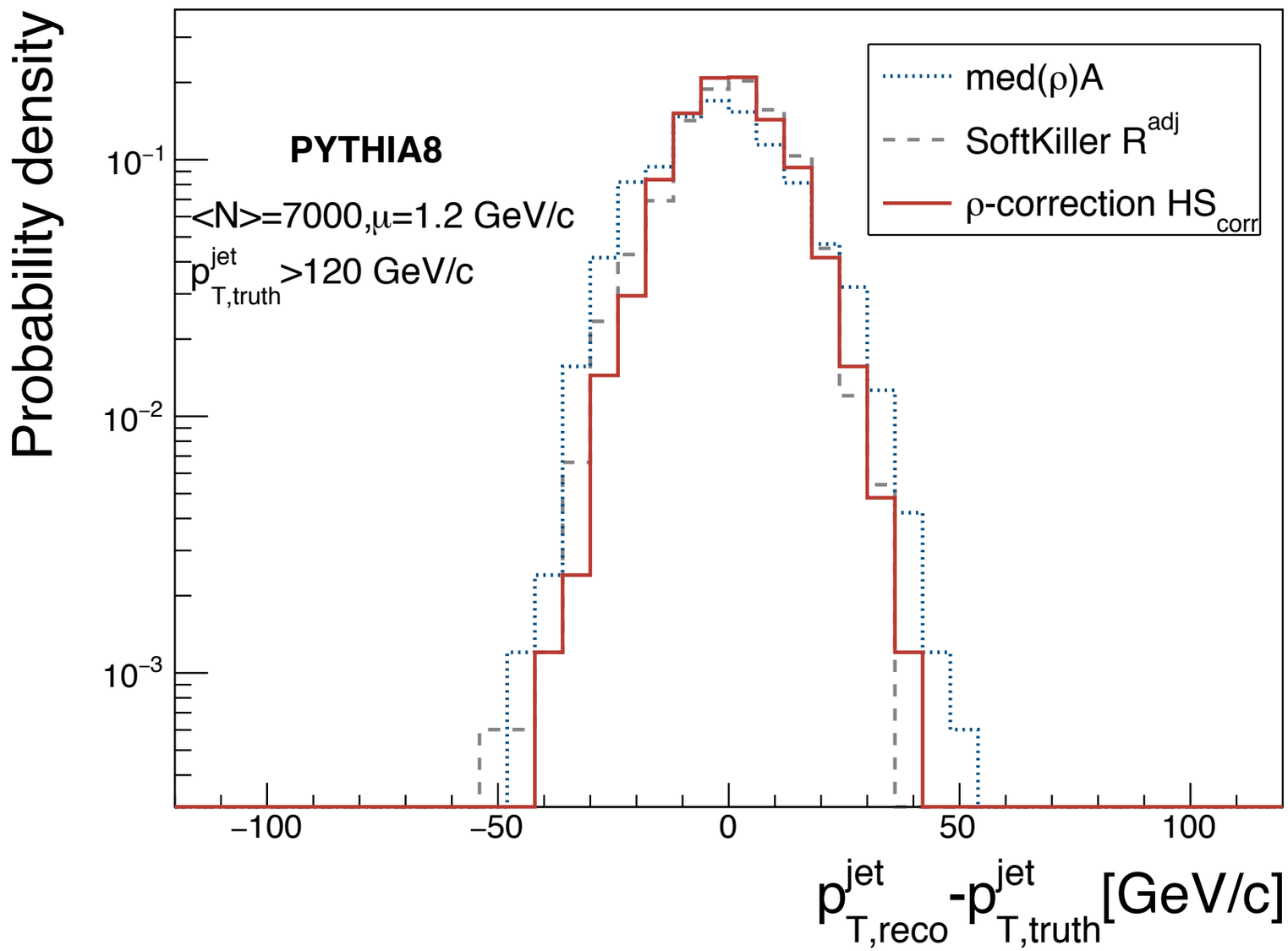
[Mehtar-Tani, ASO, Verweij arXiv:1904.12815]

p+p@HL-LHC



Plateau-like region around the optimal $p_T^{cut,soft}$.
No need to fine-tune

p+p@HL-LHC



ρ-correction HS_{corr}

- 23% improvement w.r.t area-median
- 6% improvement w.r.t SoftKiller R^{adj}

[GeV/c]	Mean	Standard deviation
med(ρ)A	-1.51±0.36	14.66±0.25
SoftKiller R^{adj}	0.29±0.29	12.10±0.21
ρ-correction	0.37±0.31	12.03 ± 0.21
ρ-correction $\mathrm{HS_{corr}}$	0.49±0.28	11.39 ±0.21

Pb+Pb@LHC

(rough) **Assumption**: energy loss is a higher-order effect.

⇨ $p_{\mathrm{T}}^{\mathrm{cut,soft}}$ minimizes $\sigma(p_{\mathrm{T,reco}}^{\mathrm{jet}} - p_{\mathrm{T,truth}}^{\mathrm{jet}})$ in JEWEL w/o E-loss

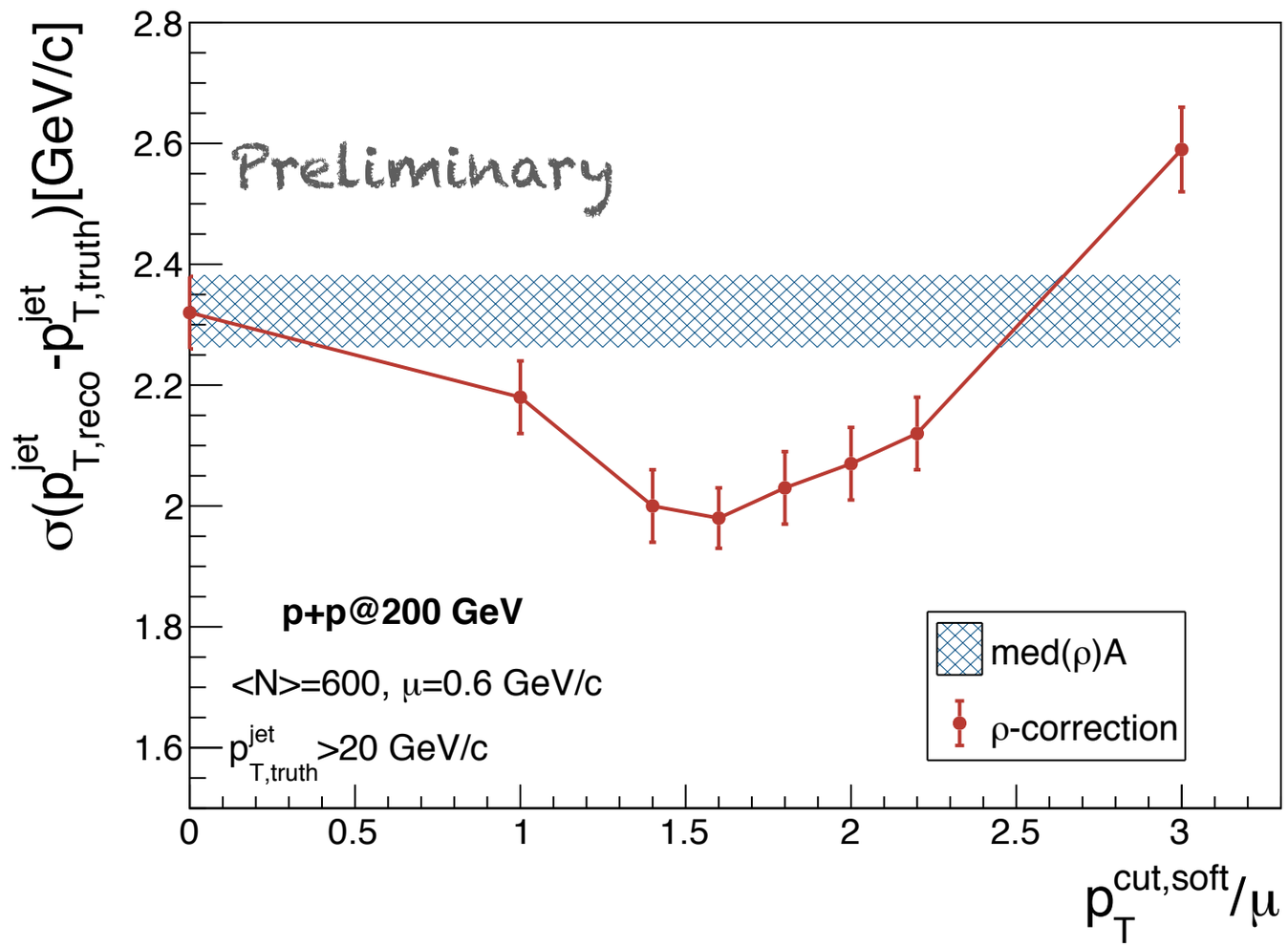
⇨ Apply $\langle p_{T,<}^{\mathrm{sig}} \rangle$ as measured in low pileup p+p

[GeV/c]	Mean	Standard deviation	Δ_{truth}
med(ρ)A	-0.61±0.17	14.73±0.12	10.1
SoftKiller R^{adj}	2.90±0.12	10.91± 0.09	7.55
ρ-correction	2.96±0.12	10.92± 0.09	7.6
ρ-correction $\mathrm{HS_{corr}}$	2.04±0.12	10.34± 0.08	6.9

$[-\Delta_{\mathrm{truth}}, \Delta_{\mathrm{truth}}]$: interval in $p_{\mathrm{T,reco}}^{\mathrm{jet}} - p_{\mathrm{T,truth}}^{\mathrm{jet}}$ centered around 0 and containing 1/2 jets

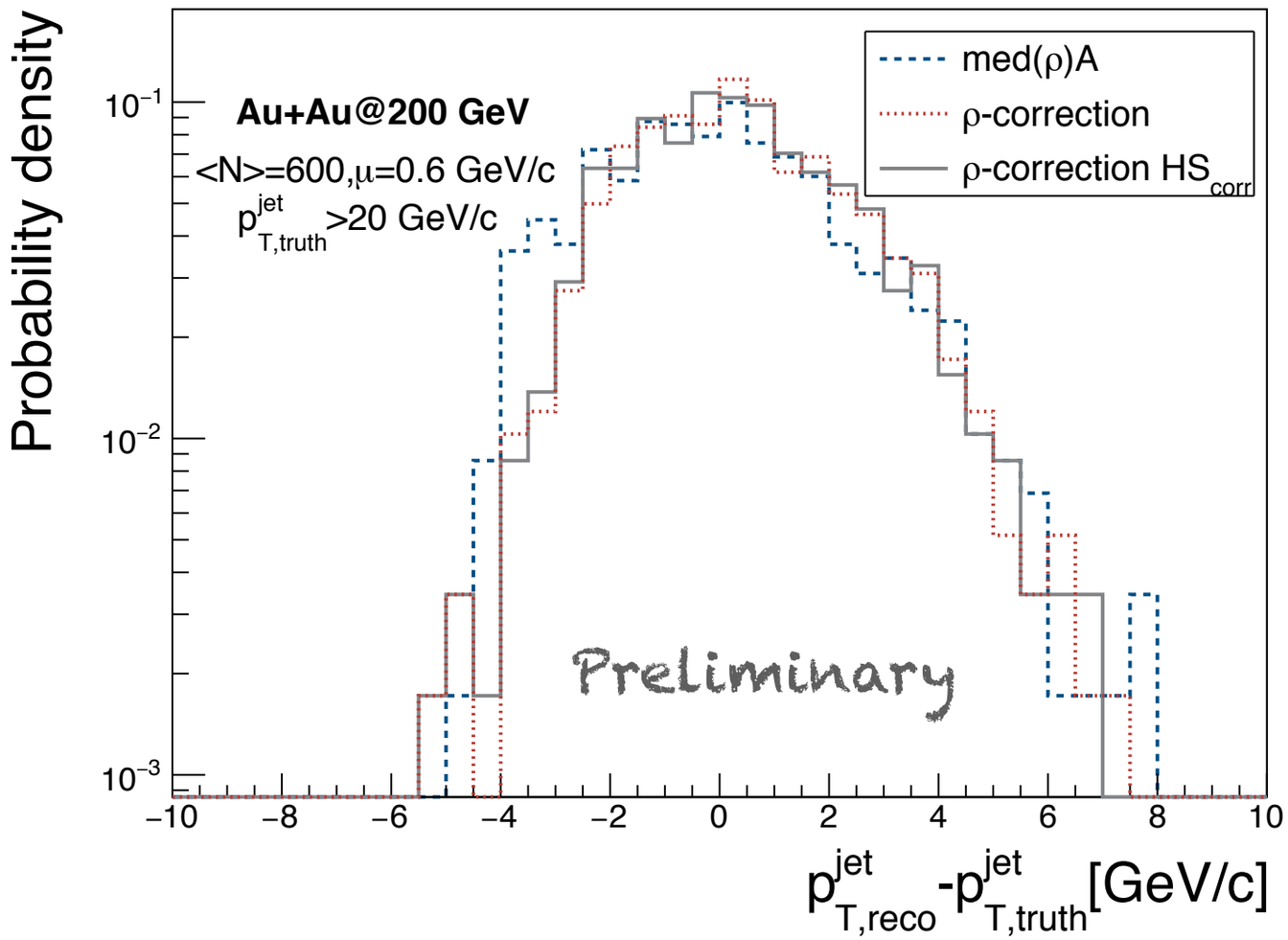
Better resolution but no way of controlling the bias due to in-medium modifications

NEW: p+p@RHIC



Method	$\langle p^{\mathrm{jet}}_{\mathrm{T,reco}} - p^{\mathrm{jet}}_{\mathrm{T,truth}} \rangle$ [GeV/c]	$\sigma(p^{\mathrm{jet}}_{\mathrm{T,reco}} - p^{\mathrm{jet}}_{\mathrm{T,truth}})$ [GeV/c]
Area-median	-0.04 ± 0.09	2.32 ± 0.06
ρ-correction	0.13 ± 0.08	1.98 ± 0.05
ρ-correction HS_{corr}	0.12 ± 0.07	1.97 ± 0.05

NEW: Au+Au@RHIC



Method	$\langle p^{\mathrm{jet}}_{\mathrm{T,reco}} - p^{\mathrm{jet}}_{\mathrm{T,truth}} \rangle$ [GeV/c]	$\sigma(p^{\mathrm{jet}}_{\mathrm{T,reco}} - p^{\mathrm{jet}}_{\mathrm{T,truth}})$ [GeV/c]
Area-median	0.05 ± 0.09	2.28 ± 0.06
ρ-correction	0.39 ± 0.08	2.02 ± 0.06
ρ-correction HS$_{\mathrm{corr}}$	0.37 ± 0.07	2.02 ± 0.06

Wrap-up

- Background subtraction essential for jet studies both in HL-LHC and heavy ions

- ρ-correction is a new method based on:

  - Choice of $p_{\mathrm{T}}^{\mathrm{cut,soft}}$ to minimize $\sigma(p_{\mathrm{T,reco}}^{\mathrm{jet}} - p_{\mathrm{T,truth}}^{\mathrm{jet}})$

  - $\langle p_{\mathrm{T,reco}}^{\mathrm{jet}} - p_{\mathrm{T,truth}}^{\mathrm{jet}} \rangle$ under control via an experimentally accesible quantity $\langle p_{\mathrm{T},<}^{\mathrm{sig}} \rangle$

  - Introduce QCD hard-soft correlations as a tool to improve $\sigma(p_{\mathrm{T,reco}}^{\mathrm{jet}} - p_{\mathrm{T,truth}}^{\mathrm{jet}})$

  - To be combined with a subtraction method
    [Y.Mehtar-Tani, ASO, M. Verweij in preparation]

Keep tuned and find more on: https://github.com/aontoso/JetToyHI