On strategies for determination and characterization of the underlying event

Sebastian Sapeta

LPTHE, UPMC, CNRS, Paris

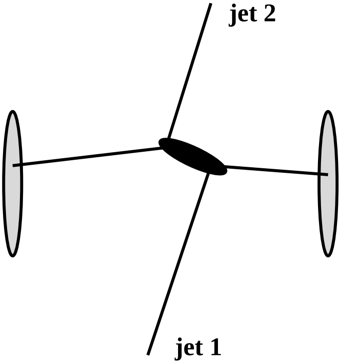
in collaboration with Matteo Cacciari and Gavin Salam*

Rencontres de Moriond, QCD and High Energy Interactions, La Thuile, March 13-20, 2010

*M.Cacciari, G.P.Salam and SS, arXiv:0912.4926

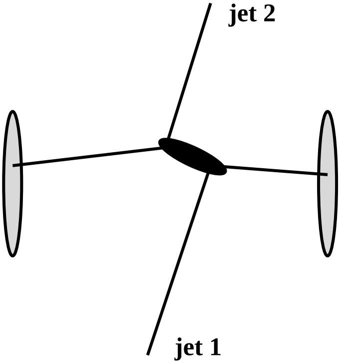
What is the underlying event?

$pp \to jj \;=$



What is the underlying event?

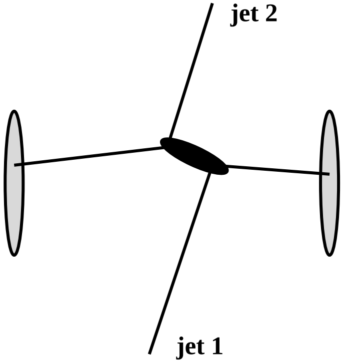
$pp \to jj \ =$



+ beam remnants
+ initial state radiation
+ multiple-parton interactions
+ ...

What is the underlying event?

$pp \to jj \;=$



+ beam remnants
+ initial state radiation
+ multiple-parton interactions
+ ...

hard interaction | **underlying activity**

- these are ingredients of present Monte Carlo models

Problems and questions

Definition of underlying event (UE) is ambiguous ...

- there is only one event with no clear bound between hard part and UE

Problems and questions

Definition of underlying event (UE) is ambiguous ...

- there is only one event with no clear bound between hard part and UE

... and its modeling difficult

- should initial state radiation be counted as part of the underlying event?
- are multiple parton interactions responsible for most of the UE?
- what about correlations? or other mechanisms like BFKL chains?

Problems and questions

Definition of underlying event (UE) is ambiguous ...

- there is only one event with no clear bound between hard part and UE

... and its modeling difficult

- should initial state radiation be counted as part of the underlying event?
- are multiple parton interactions responsible for most of the UE?
- what about correlations? or other mechanisms like BFKL chains?

Therefore, we should be confident that we can measure it well

- this would help constraining, tuning and improving the models

Problems and questions

Definition of underlying event (UE) is ambiguous ...

- there is only one event with no clear bound between hard part and UE

... and its modeling difficult

- should initial state radiation be counted as part of the underlying event?
- are multiple parton interactions responsible for most of the UE?
- what about correlations? or other mechanisms like BFKL chains?

Therefore, we should be confident that we can measure it well

- this would help constraining, tuning and improving the models

This leads us to the following two questions:

1. what do we really measure with existing methods of UE determination?
   $\longrightarrow$ test the methods with toy model
2. which observables are interesting to measure?
   $\longrightarrow$ study UE from Monte Carlo models

What can we measure about UE?

Relevant characteristics of energy flow of UE

- ρ – level of transverse momentum per unit area
- rapidity dependence of ρ
- point-to-point fluctuations within a single event ($\equiv \sigma$)
- fluctuations from event to event
- point-to-point correlations

What can we measure about UE?

Relevant characteristics of energy flow of UE

- ρ – level of transverse momentum per unit area
- rapidity dependence of ρ
- point-to-point fluctuations within a single event ($\equiv \sigma$)
- fluctuations from event to event
- point-to-point correlations

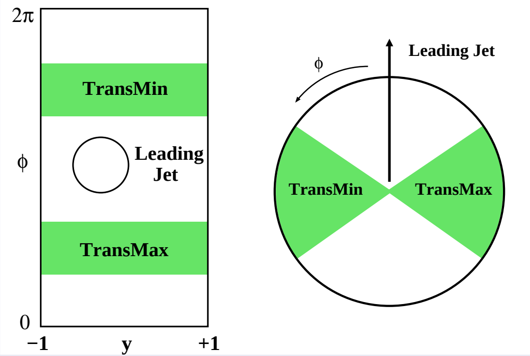
Two existing methods for measuring UE

- traditional approach
- area/median based approach

Traditional approach [Marchesini & Webber (1988), UA1 (1988), Field et al.]

For each event

1. take charged particles with $p_t > 0.5$ GeV and $|y| < 1$
2. cluster with cone jet algorithm with $R = 0.7$ to find the leading jet
3. define typical p_t of UE as $\langle p_t \rangle$ in TransMin, TransMax or TransAv regions



TransAv: $\mathcal{O}(\alpha_s)$

TransMax: $\mathcal{O}(\alpha_s)$

TransMin: $\mathcal{O}(\alpha_s^2)$

- **topological** separation: UE defined as particles entering certain region of (y, ϕ) space

Area/median approach [Cacciari, Salam, Soyez (2008), http://fastjet.fr]

Area/median approach [Cacciari, Salam, Soyez (2008), http://fastjet.fr]

For each event

1. add a dense set of infinitely soft particles, *ghosts*, distributed uniformly in y and ϕ

Area/median approach [Cacciari, Salam, Soyez (2008), http://fastjet.fr]

For each event

1. add a dense set of infinitely soft particles, *ghosts*, distributed uniformly in y and ϕ
2. cluster everything (real particles+ghosts) with an infrared safe jet finding algorithm
   $\longrightarrow$ set of jets ranging from hard to soft

Area/median approach [Cacciari, Salam, Soyez (2008), http://fastjet.fr]

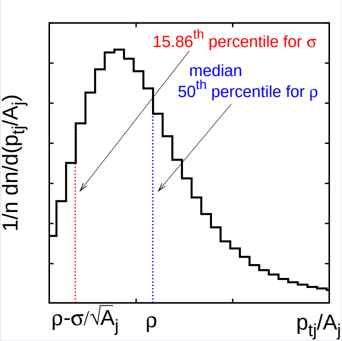
For each event

1. add a dense set of infinitely soft particles, *ghosts*, distributed uniformly in y and ϕ
2. cluster everything (real particles+ghosts) with an infrared safe jet finding algorithm $\longrightarrow$ set of jets ranging from hard to soft
3. determine area (measure of susceptibility to the soft radiation) for each jet

Area/median approach [Cacciari, Salam, Soyez (2008), `http://fastjet.fr`]

For each event

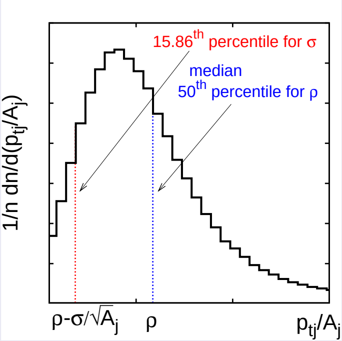
1. add a dense set of infinitely soft particles, *ghosts*, distributed uniformly in y and ϕ
2. cluster everything (real particles+ghosts) with an infrared safe jet finding algorithm $\rightarrow$ set of jets ranging from hard to soft
3. determine area (measure of susceptibility to the soft radiation) for each jet



Area/median approach [Cacciari, Salam, Soyez (2008), http://fastjet.fr]

For each event

1. add a dense set of infinitely soft particles, *ghosts*, distributed uniformly in y and ϕ
2. cluster everything (real particles+ghosts) with an infrared safe jet finding algorithm $\rightarrow$ set of jets ranging from hard to soft
3. determine area (measure of susceptibility to the soft radiation) for each jet
4. from the list of all jets (no cuts required!) determine
$$\rho = \text{median}\left[\left\{\frac{p_{t,j}}{A_j}\right\}\right]$$
and its uncertainty σ
   - median gives a typical value of p_t/A for a given event
   - using median is a way to **dynamically** separate hard and soft parts of the event



Understanding the methods – a toy model study

Two component model: soft UE + hard contamination

soft component (UE)

- take the region of area A in (y, ϕ) space $\longrightarrow$ transverse region (traditional approach) or jet area (area/median approach)
- number of particles in this region, n, given by Poisson distribution with the average $\langle n \rangle$
- single-particle p_t distribution given by
$$\frac{dpt_1}{dp_t} = \frac{1}{\mu} e^{-p_t/\mu}$$
- parameters:
μ – average p_t of particle,
$\nu = \frac{\langle n \rangle}{A}$ – density of particles
- in this model $\rho = \mu\nu$ is the true value of p_t/A of UE

Two component model: soft UE + hard contamination

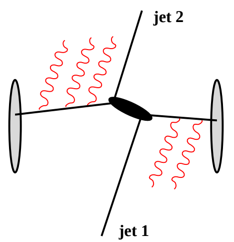
soft component (UE)

- take the region of area A in (y, ϕ) space $\longrightarrow$ transverse region (traditional approach) or jet area (area/median approach)
- number of particles in this region, n, given by Poisson distribution with the average $\langle n \rangle$
- single-particle p_t distribution given by

$$\frac{dpt_1}{dp_t} = \frac{1}{\mu} e^{-p_t/\mu}$$

- parameters:
  μ – average p_t of particle,
  $\nu = \frac{\langle n \rangle}{A}$ – density of particles
- in this model $\rho = \mu\nu$ is the true value of p_t/A of UE

hard component (ISR)



- soft and collinear partons from primary emissions:
  $\frac{dn}{dp_t dy d\phi} \simeq \frac{C_i}{\pi^2} \frac{\alpha_s(p_t)}{p_t}$
- hard scale cut $Q = \frac{1}{2} p_t = 50$ GeV
- partons distributed uniformly in angle and rapidity

Fluctuations in estimation of ρ

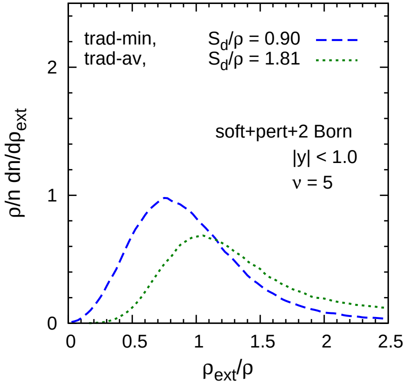
In the toy model: the same ρ distribution used to generate all events

- however, there are event-to-event fluctuations of ρ due to restricted area
- this sets the lower limit for the uncertainty of ρ determination

Fluctuations in estimation of ρ

In the toy model: the same ρ distribution used to generate all events

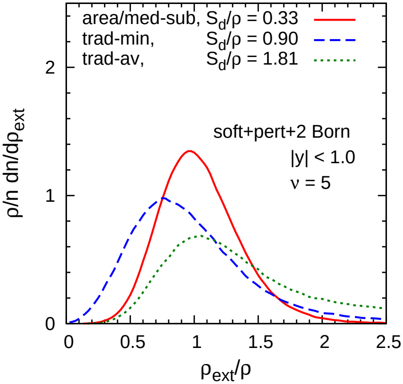
- however, there are event-to-event fluctuations of ρ due to restricted area
- this sets the lower limit for the uncertainty of ρ determination



Fluctuations in estimation of ρ

In the toy model: the same ρ distribution used to generate all events

- however, there are event-to-event fluctuations of ρ due to restricted area
- this sets the lower limit for the uncertainty of ρ determination

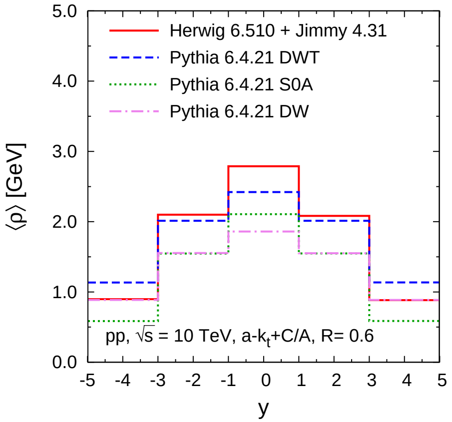


- traditional approach suffers significantly more from the hard contamination $S_d \sim Q$

Approaching real life – Monte Carlo study

Average ρ as a function of y

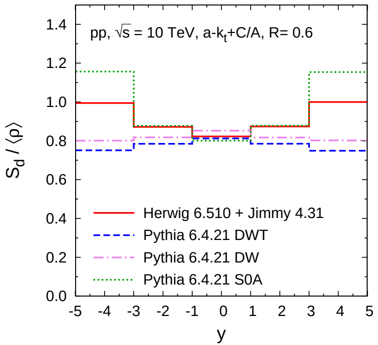
- dijets at the LHC, $\sqrt{s} = 10$ TeV, $p_t > 100$ GeV, $|y| < 4$



- significant y dependence
- strips of $\Delta y{=}2$ sufficient for robust ρ determination

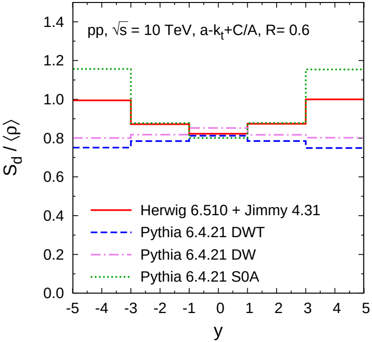
Fluctuations

- from event to event

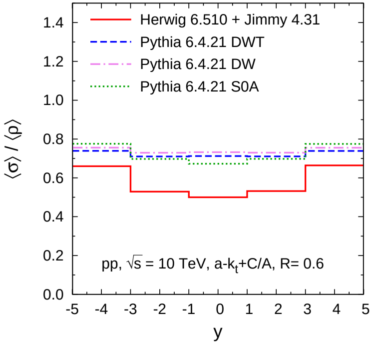


Fluctuations

- from event to event

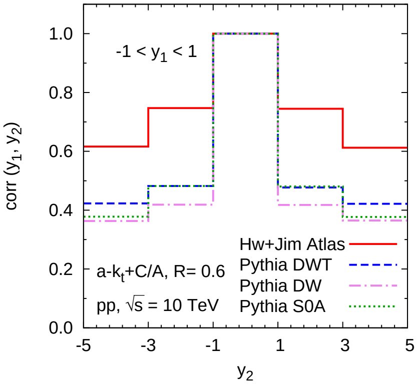


- within an event



- large inter-event and intra-event
- two patterns of rapidity dependence
- sizable difference between Herwig+Jimmy and Pythia

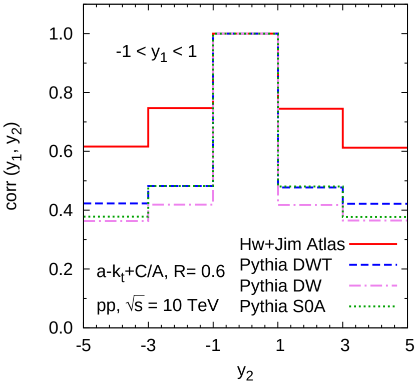
Correlations



$$\mathrm{corr}(y_1, y_2) = \frac{\langle \rho(y_1)\rho(y_2) \rangle - \langle \rho(y_1) \rangle \langle \rho(y_2) \rangle}{S_d(y_1) S_d(y_2)}$$

- y_1, y_2 – rapidity bins of width $\Delta y = 2$
- $\langle \dots \rangle$ – average over many events

Correlations

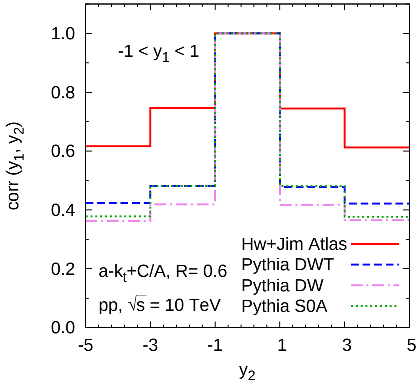


$$\mathrm{corr}(y_1, y_2) = \frac{\langle \rho(y_1)\rho(y_2)\rangle - \langle \rho(y_1)\rangle\langle \rho(y_2)\rangle}{S_d(y_1)S_d(y_2)}$$

- y_1, y_2 – rapidity bins of width $\Delta y = 2$
- $\langle \ldots \rangle$ – average over many events

- significant difference between Herwig + Jimmy and Pythia

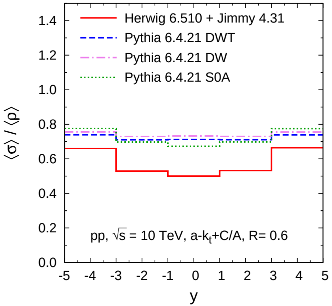
Correlations



$$\mathrm{corr}(y_1, y_2) = \frac{\langle \rho(y_1)\rho(y_2) \rangle - \langle \rho(y_1) \rangle \langle \rho(y_2) \rangle}{S_d(y_1) S_d(y_2)}$$

- y_1, y_2 – rapidity bins of width $\Delta y = 2$
- $\langle \dots \rangle$ – average over many events

- significant difference between Herwig + Jimmy and Pythia
- qualitatively consistent with $\langle \sigma \rangle / \langle \rho \rangle$: smaller fluctuations within event $\Leftrightarrow$ larger correlations



Summary

Measurement of UE is difficult both in principle and in practice

- we have considered a simple toy model to better understand the methods
- both traditional and area/based approach perform comparably well in measuring average quantities
- for event-to-event measurements traditional approach suffers significantly from hard radiation

Summary

Measurement of UE is difficult both in principle and in practice

- we have considered a simple toy model to better understand the methods
- both traditional and area/based approach perform comparably well in measuring average quantities
- for event-to-event measurements traditional approach suffers significantly from hard radiation

The study of UE from MC with the area/median method suggests the set of observables deserving dedicated measurements

- dependence of ρ on rapidity
- fluctuations from event to event (large for all generators/tunes)
- fluctuations within an event, σ, (significant differences between Herwig+Jimmy and Pythia)
- correlations (large differences between Herwig+Jimmy and Pythia)

Summary

Measurement of UE is difficult both in principle and in practice

- we have considered a simple toy model to better understand the methods
- both traditional and area/based approach perform comparably well in measuring average quantities
- for event-to-event measurements traditional approach suffers significantly from hard radiation

The study of UE from MC with the area/median method suggests the set of observables deserving dedicated measurements

- dependence of ρ on rapidity
- fluctuations from event to event (large for all generators/tunes)
- fluctuations within an event, σ, (significant differences between Herwig+Jimmy and Pythia)
- correlations (large differences between Herwig+Jimmy and Pythia)

$\longrightarrow$ for more details see: arXiv:0912.4926

BACKUP SLIDES

Jets have areas

To determine the **active area** of a jet

- supplement a set of physical particles $\{p_i\}$ with an ensemble of dense, infinitely soft and randomly distributed *ghost particles* $\{g_i\}$
- cluster the set $\{p_i, g_i\}$
- compute the active area of a jet J for this specific ensemble of ghosts $\{g_i\}$

$$A(J \mid \{g_i\}) = \frac{\mathcal{N}(J)}{\nu_g},$$

  where $\mathcal{N}(J)$ is the number of ghosts contained in the jet J and ν_g is the number of ghosts per unit area
- average over many ghost ensembles

$$A(J) \equiv \lim_{\nu_g \to \infty} \langle A(J \mid \{g_i\}) \rangle_g$$

Toy model: soft UE (extraction of ρ)

Traditional approach

$$\begin{aligned}\langle\rho_{\rm ext,av}\rangle &= \rho\\ \langle\rho_{\rm ext,min}\rangle &= \rho - \sigma/\sqrt{\pi A_{\rm Trans}}\\ \langle\rho_{\rm ext,max}\rangle &= \rho + \sigma/\sqrt{\pi A_{\rm Trans}}\end{aligned}$$

for $A_{\rm Trans} = 2$ and $\sigma \simeq 0.5 - 0.7\rho$

$$\sigma/\sqrt{\pi A_{\rm Trans}} \simeq 20 - 30\%$$

Toy model: soft UE (extraction of ρ)

Traditional approach

$$\langle \rho_{\text{ext,av}} \rangle = \rho$$
$$\langle \rho_{\text{ext,min}} \rangle = \rho - \sigma / \sqrt{\pi A_{\text{Trans}}}$$
$$\langle \rho_{\text{ext,max}} \rangle = \rho + \sigma / \sqrt{\pi A_{\text{Trans}}}$$

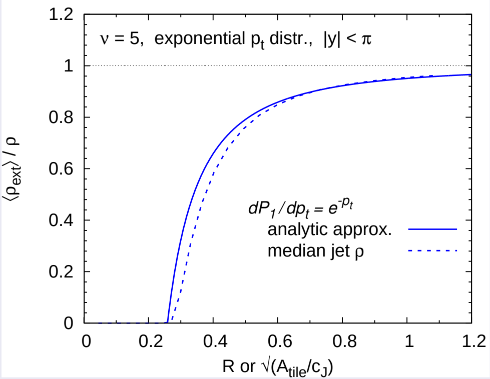
for $A_{\text{Trans}} = 2$ and $\sigma \simeq 0.5 - 0.7\rho$

$$\sigma / \sqrt{\pi A_{\text{Trans}}} \simeq 20 - 30\%$$

Area/median based approach

$$\frac{\langle \rho_{\text{ext}} \rangle}{\rho} \simeq \frac{c_J R^2 \nu - \ln 2}{c_J R^2 \nu - \ln 2 + \frac{1}{2}} \Theta(c_J R^2 \nu - \ln 2)$$

$$c_J \simeq 2 , \quad \nu = \text{particle density}$$



Two component model: biases

Traditional approach

- in the events with at least one perturbative emission the bias of $\rho_{\mathrm{ext,soft}}$ is removed and the bias of $\rho_{\mathrm{ext,hard}}$ dominates

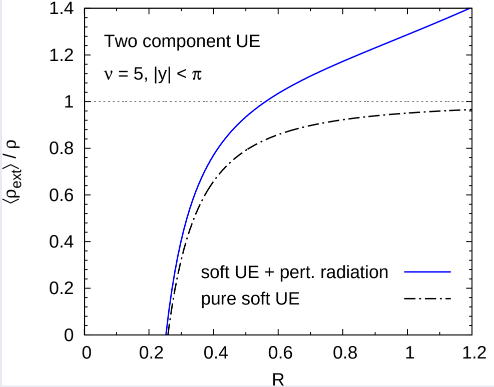
$$\langle \rho_{\mathrm{ext,Av}} \rangle = \rho + \frac{C_i \alpha_s}{\pi^2} Q$$

$$\langle \rho_{\mathrm{ext,Min}} \rangle \simeq \rho - \frac{\sigma \mathcal{P}}{\sqrt{\pi A_{\mathrm{Trans}}}} + 2 \left(\frac{C_i \alpha_s}{\pi^2} \right)^2 A_{\mathrm{Trans}}\, Q$$

$\mathcal{P}$ – fraction of events with perturbative radiation smaller then soft fluctuations

Two component model: biases

Area/median approach



$$\langle \rho_{\rm ext} \rangle \simeq \langle \rho_{\rm ext}^{\rm (soft)} \rangle + \sqrt{\frac{\pi c_J}{2}} \sigma R \frac{\langle n_h \rangle}{A_{\rm tot}}$$

$\langle n_h \rangle$ – number of perturbative part.

σ – measure of fluctuations

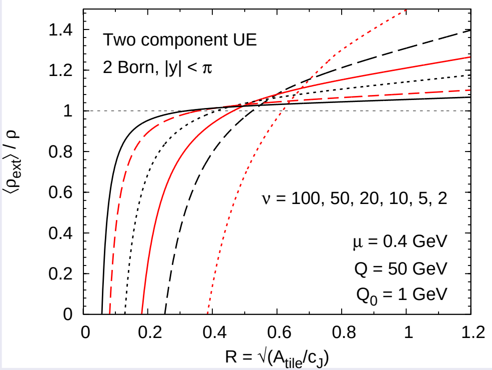
ρ – true value of p_t/A

$$\frac{\langle n_h \rangle}{A_{\rm tot}} \simeq \frac{n_b}{A_{\rm tot}} + \frac{C_i}{\pi^2} \frac{1}{2b_0} \ln \frac{\alpha_s(Q_0)}{\alpha_s(Q)}$$

- the two terms bias $\langle \rho_{\rm ext} \rangle$ in opposite directions
- for $R \simeq 0.5 - 0.6$ (used in most MC analysis of UE) the biases largely cancel
- similar picture and conclusions for σ

Two component model: biases

Area/median approach



Turn-on point:

$$R_{\mathrm{crit}} \simeq 0.41 \cdot \frac{\sigma}{\rho} = 0.41 \cdot \sqrt{\frac{2}{\nu}}$$

Point of zero bias:

$$R_{\text{zero-bias}} \simeq 0.87\, R_{\mathrm{crit}}^{\frac{1}{3}} \left(\frac{C_A}{C_i}\right)^{\frac{1}{3}}$$

- for $R \simeq 0.5 - 0.6$ (used in most MC analysis of UE) the biases largely cancel

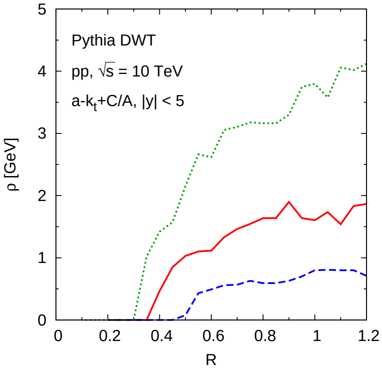
Summary of biases

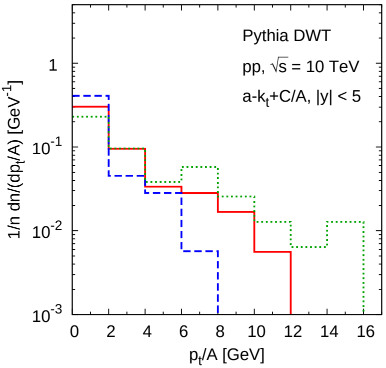
quantity	method	result	numerical value
$\frac{\delta\rho^{(\mathrm{soft})}}{\rho}$	TransAv	0	0
	TransMin	$-\frac{\sigma}{\rho}\frac{\mathcal{P}}{\sqrt{\pi A_{\mathrm{Trans}}}}$	-0.09
	Area/Med*	$-\frac{\sigma^2}{\rho^2}\frac{1}{4c_J R^2}$	-0.14
$\frac{\delta\rho^{(\mathrm{hard})}}{\rho}$	TransAv	$\frac{C_i\alpha_s}{\pi^2}\frac{Q}{\rho}$	0.99
	TransMin	$2\left(\frac{C_i\alpha_s}{\pi^2}\right)^2\frac{A_{\mathrm{Trans}}Q}{\rho}$	0.16
	Area/Med	$\frac{\sigma R}{\rho}\sqrt{\frac{\pi c_J}{2}}\left(\frac{n_b}{A_{\mathrm{tot}}}+\frac{C_i}{\pi^2}\frac{L}{2b_0}\right)$	0.17

Summary of biases (cont.)

quantity	method	result	numerical value
$\frac{S_d^{(\mathrm{soft})}}{\rho}$	TransAv	$\frac{\sigma}{\rho}\sqrt{\frac{1}{2A_{\mathrm{Trans}}}}$	0.31
	TransMin	$\frac{\sigma}{\rho}\sqrt{\frac{\pi-1}{\pi A_{\mathrm{Trans}}}}$	0.36
	Area/Med	$\frac{\sigma}{\rho}\sqrt{\frac{\pi}{2A_{\mathrm{tot}}}}$	0.22
$\frac{S_d^{(\mathrm{hard})}}{\rho}$	TransAv	$\sqrt{\frac{C_i\alpha_s}{4A_{\mathrm{Trans}}\pi^2}}\,\frac{Q}{\rho}$	1.72
	TransMin	$\frac{C_i\alpha_s}{\pi^2\sqrt{2}}\,\frac{Q}{\rho}$	0.70
	Area/Med	$\frac{\sigma R}{\rho}\sqrt{\frac{2\pi c_J}{A_{\mathrm{tot}}}}\left(\frac{n_b}{A_{\mathrm{tot}}}+\frac{C_i}{\pi^2}\frac{L}{2b_0}\right)^{\frac{1}{2}}$	0.19

Comparison of characteristics: toy model vs MC

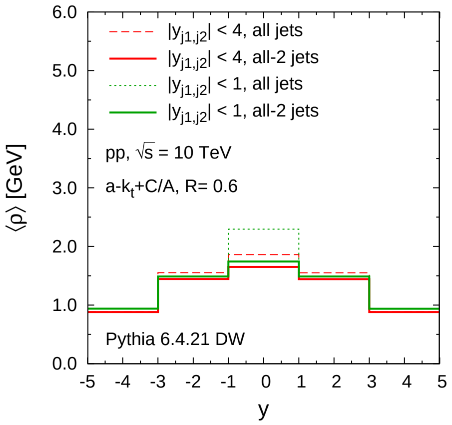




- the pattern for $\rho(R)$ from the toy model present in MC events: (i) turn-on at low R, (ii) linear growth at larger R
- variation in the curves indicative of the inter-event fluctuations
- growth of ρ with R produced by the tails of distributions of p_t/A

Average ρ as a function of y

- dijets at the LHC, $\sqrt{s} = 14$ TeV, $p_{t,\min} = 50$ GeV



Average ρ as a function of y

- dijets at the LHC, $\sqrt{s} = 14$ TeV, $p_{t,\min} = 50$ GeV