Characterizing the underlying event in hadron-hadron collisions

Sebastian Sapeta

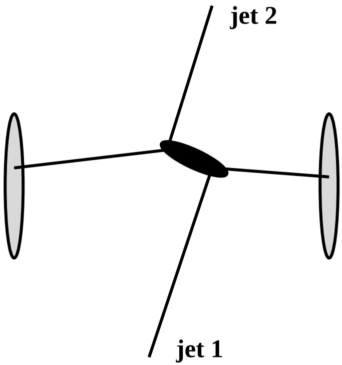
LPTHE, UPMC, CNRS, Paris

in collaboration with Matteo Cacciari and Gavin Salam[§]

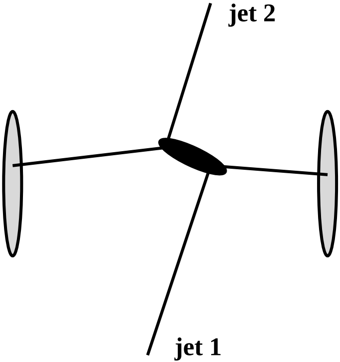
Workshop on Hadron-Hadron & Cosmic-Ray Interactions at multi-TeV Energies
ECT - Trento, Nov 29th - Dec 3rd, 2010*

[§]M.Cacciari, G.P.Salam and SS, JHEP 1004 (2010) 065

What is the underlying event?

$pp \rightarrow jj \;=$ 

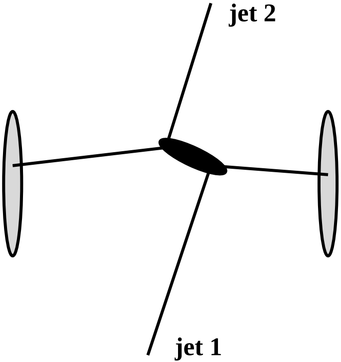
What is the underlying event?

$pp \to jj \;=$ 

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

What is the underlying event?

$pp \longrightarrow jj \ =$



+ 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

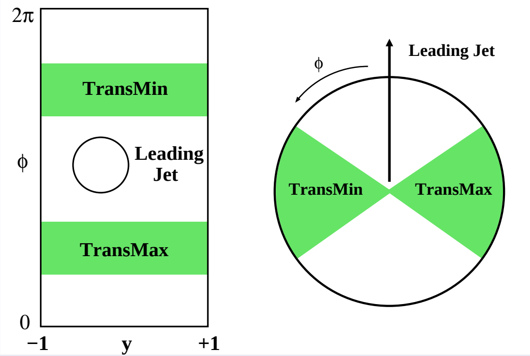
Two existing methods for measuring UE

- traditional approach
- jet 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 $|y| < y_{\max}$
2. cluster with cone jet algorithm with $R = 0.5 - 0.7$ 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

Parenthesis: Sequential recombination jet algorithms

Distance measure between pair of particles d_{ij} and between particle and beam d_{iB}

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta_{ij}^2}{R^2}\,, \qquad\qquad d_{iB} = k_{ti}^{2p}\,,$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is a geometrical distance in the (y, ϕ) plane

Parenthesis: Sequential recombination jet algorithms

Distance measure between pair of particles d_{ij} and between particle and beam d_{iB}

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}\,, \qquad d_{iB} = k_{ti}^{2p}\,,$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is a geometrical distance in the (y, ϕ) plane

- identify the smallest of d_{ij} and d_{iB}
- recombine particles i and j (if d_{iB} is the smallest call i a jet and remove it)
- recalculate distances and repeat the procedure until no entries are left

Parenthesis: Sequential recombination jet algorithms

Distance measure between pair of particles d_{ij} and between particle and beam d_{iB}

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}\,, \qquad d_{iB} = k_{ti}^{2p}\,,$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is a geometrical distance in the (y, ϕ) plane

- identify the smallest of d_{ij} and d_{iB}
- recombine particles i and j (if d_{iB} is the smallest call i a jet and remove it)
- recalculate distances and repeat the procedure until no entries are left

$p = 1$: k_t algorithm
[Catani, Dokshitzer, Seymour, Webber (1993); Ellis, Soper (1993)]

$p = 0$: Cambridge/Aachen algorithm
[Dokshitzer, Leder, Moretti, Webber (1997); Wobisch, Wengler (1999)]

$p = -1$: anti-k_t algorithm
[Cacciari, Salam, Soyez (2008)]

Parenthesis: Jets have areas [Cacciari, Salam, Soyez (2008)]

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

- supplement a set of physical particles $\{p_i\}$ with an ensemble of dense, infinitely soft, 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

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

Parenthesis: Jets have areas [Cacciari, Salam, Soyez (2008)]

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

- supplement a set of physical particles $\{p_i\}$ with an ensemble of dense, infinitely soft, 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 $\{g_i\}$

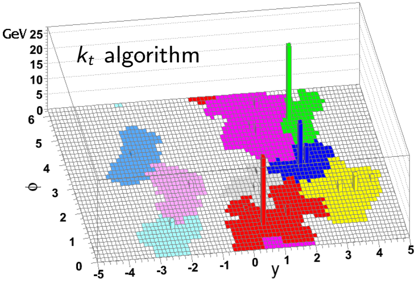
$$A(J \mid \{g_i\}) = \frac{\mathcal{N}(J)}{\nu_g},$$

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

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

http://fastjet.fr



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

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

For each event

1. cluster particles with an infrared safe jet finding algorithm (all particles are clustered so we have set of jets ranging from hard to soft)

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

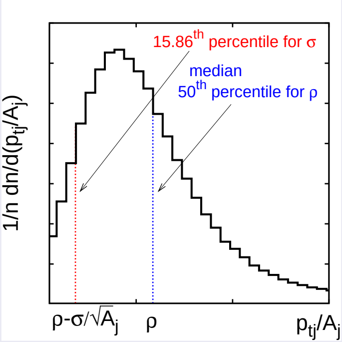
For each event

1. cluster particles with an infrared safe jet finding algorithm (all particles are clustered so we have set of jets ranging from hard to soft)
2. 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



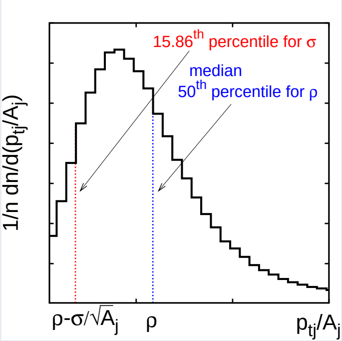
Area/median approach [Cacciari, Salam, Soyez (2008), `http://fastjet.fr`]

For each event

1. cluster particles with an infrared safe jet finding algorithm (all particles are clustered so we have set of jets ranging from hard to soft)
2. 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



- ρ may be used e.g. to correct hard jet transverse momentum

$$p_{t,j}^{(\text{sub})} = p_{t,j} - \rho\, A_j \pm \sigma\sqrt{A_j}$$

since jet area measures the jet susceptibility to the soft radiation

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

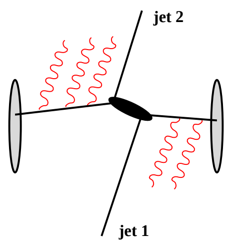
soft component (UE)

- 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)

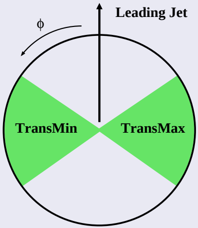


- soft and collinear partons from primary emissions:
  $$\frac{dn}{dp_t dy d\phi} \simeq \frac{C_i}{\pi^2} \frac{\alpha_s(p_t)}{p_t}$$
- hard scale cut $Q = \frac{1}{2} p_t = 50$ GeV
- partons distributed uniformly in angle and rapidity

Two component model: biases

Traditional approach

- TransAv and TransMin variants



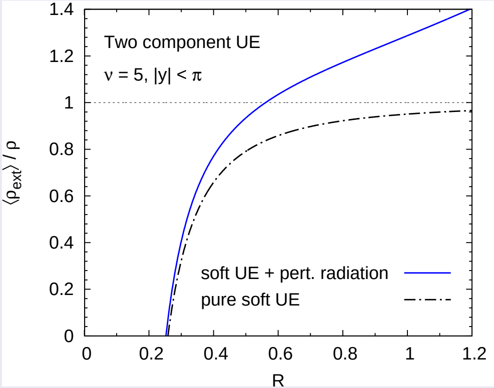
$$\langle \rho_{\text{ext,Av}} \rangle = \rho + \frac{C_i \alpha_s}{\pi^2} Q$$

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

$\mathcal{P}$ – 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

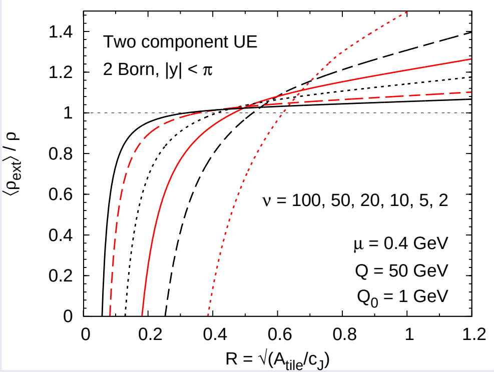
ρ – true value of p_t/A

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

- 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:

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

Point of zero bias:

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

- for $R \simeq 0.5 - 0.6$ the uncertainty always stays below 20% for a broad range of particle densities ν

Fluctuations in estimation of ρ

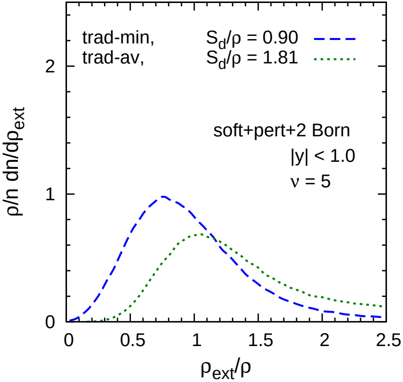
In the toy model: the same ρ distribution used to generate all events

- 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

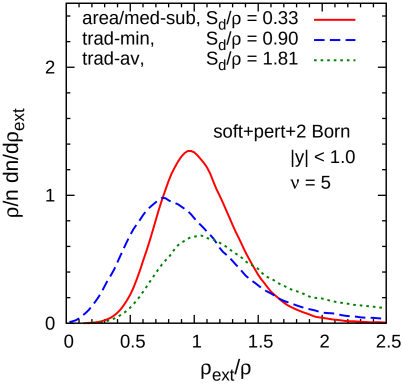
- however, there are event-to-event fluctuations of ρ due to restricted area
- this sets the lower limit for the uncertainty of ρ determination



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

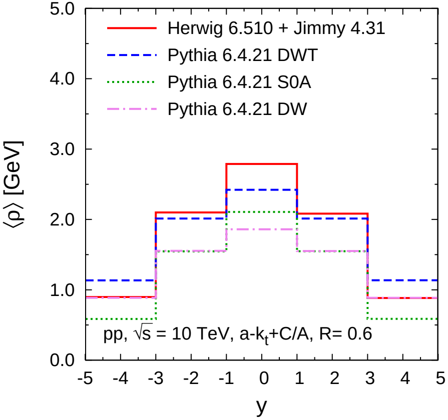


- traditional approach suffers significantly more from the hard contamination $S_d \sim Q$

Approaching real life – Monte Carlo study

Average ρ as a function of y

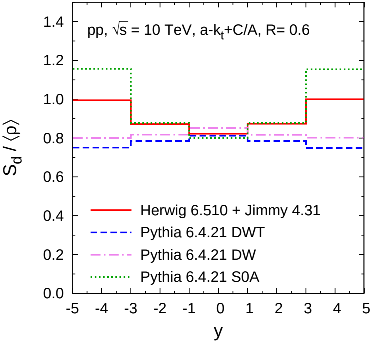
- dijets at the LHC, $\sqrt{s} = 10$ TeV, $p_t > 100$ GeV, $|y| < 4$



- significant y dependence
- strips of Δy=2 sufficient for robust ρ determination

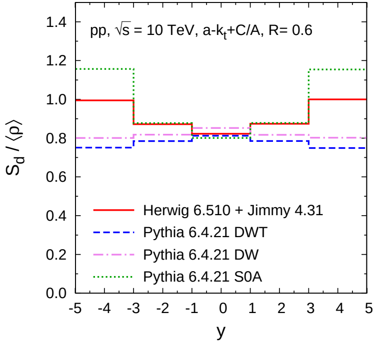
Fluctuations

- from event to event

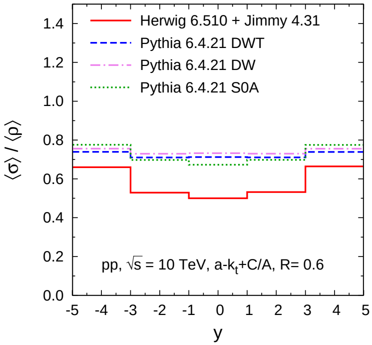


Fluctuations

- from event to event

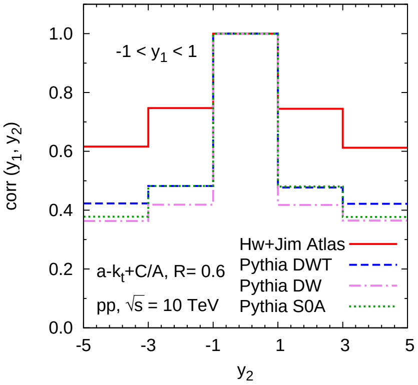


- within an event



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

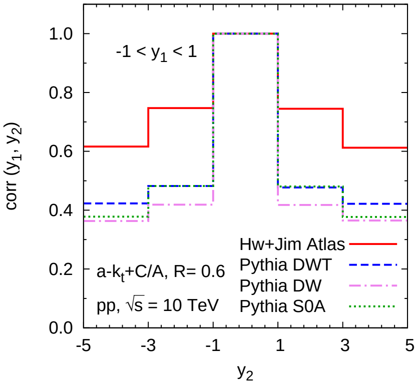
Correlations



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

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

Correlations

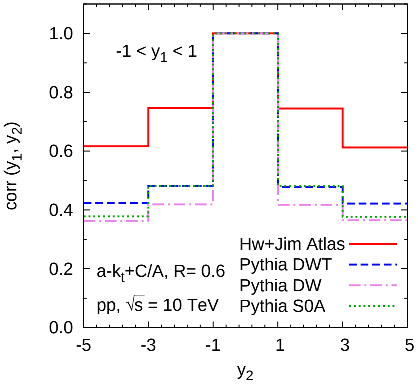


$$\mathrm{corr}(y_1, y_2) = \frac{\langle \rho(y_1)\rho(y_2)\rangle - \langle \rho(y_1)\rangle\langle \rho(y_2)\rangle}{S_d(y_1)S_d(y_2)}$$

- 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

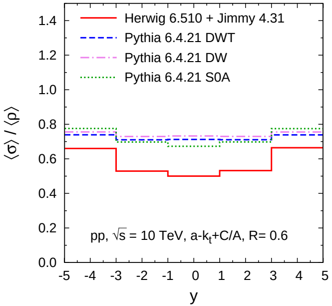
Correlations



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

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

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



CMS analysis at $\sqrt{s} = 0.9$ TeV

Experimental conditions

- only charged tracks
- $p_t > 0.3$ GeV
- $|\eta| < 2.3$

CMS analysis at $\sqrt{s} = 0.9$ TeV

Experimental conditions

- only charged tracks
- $p_t > 0.3$ GeV
- $|\eta| < 2.3$

$\longrightarrow$ **very low multiplicity**

CMS analysis at $\sqrt{s} = 0.9$ TeV

Experimental conditions

- only charged tracks
- $p_t > 0.3$ GeV
- $|\eta| < 2.3$

$\longrightarrow$ **very low multiplicity**

Modification of $\rho \longrightarrow \rho'$

$$\rho' = \underset{j \in \text{physical jets}}{\text{median}} \left[\left\{ \frac{p_{t,j}}{A_j} \right\} \right] \cdot \mathcal{C}$$

$$\mathcal{C} = \frac{\sum_{j \in \text{physical jets}} A_j}{A_{\text{tot}}}$$

CMS analysis at $\sqrt{s} = 0.9$ TeV

Experimental conditions

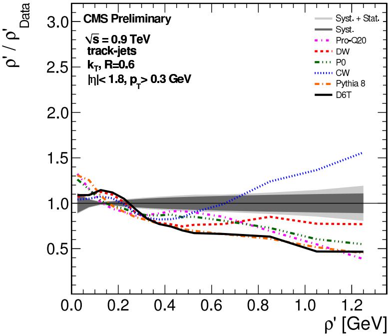
- only charged tracks
- $p_t > 0.3$ GeV
- $|\eta| < 2.3$

$\longrightarrow$ **very low multiplicity**

Modification of $\rho \to \rho'$

$$\rho' = \underset{j \in \text{physical jets}}{\text{median}} \left[\left\{ \frac{p_{t,j}}{A_j} \right\} \right] \cdot \mathcal{C}$$

$$\mathcal{C} = \frac{\sum_{j \in \text{physical jets}} A_j}{A_{\text{tot}}}$$



- can discriminate between UE tunes even at extreme conditions
- none of the tunes describes data
- looking forward to the 7 TeV results (also for other observables)

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)