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?

$pp \rightarrow jj = \text{jet 1} \rightarrow \text{jet 2}$
What is the underlying event?

\[ pp \rightarrow jj = \]

- jet 1
- jet 2
- beam remnants
- initial state radiation
- multiple-parton interactions
- ...

Sebastian Sapeta (LPTHE, Paris)
What is the underlying event?

$pp \rightarrow jj = \text{hard interaction}$

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

these are ingredients of present Monte Carlo models
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?
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?
   → test the methods with toy model
2. which observables are interesting to measure?
   → study UE from Monte Carlo models
What can we measure about UE?

Relevant characteristics of energy flow of UE

- $\rho$ – level of transverse momentum per unit area
- Rapidity dependence of $\rho$
- 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

- $\rho$ – level of transverse momentum per unit area
- Rapidity dependence of $\rho$
- 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

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

- **TransMin:** $O(\alpha_s^2)$
- **TransMax:** $O(\alpha_s)$
- **TransAv:** $O(\alpha_s)$

▶ **topological** separation: UE defined as particles entering certain region of $(y, \phi)$ space
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 $\phi$
For each event

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

2. cluster everything (real particles + ghosts) with an infrared safe jet finding algorithm
   $\rightarrow$ 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 $\phi$

2. cluster everything (real particles+ghosts) with an infrared safe jet finding algorithm → 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 $\phi$

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

\[
\frac{1}{n} \frac{dn}{dp_{tj}/A_j}
\]

15.86\text{th} percentile for $\sigma$

median

50\text{th} percentile for $\rho$

Sebastian Sapeta (LPTHE, Paris)

On strategies for determination and characterization of the underlying event
For each event

1. add a dense set of infinitely soft particles, *ghosts*, distributed uniformly in \( y \) and \( \phi \)

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 \( \sigma \)

- 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

\[
\frac{1}{n} \frac{dn}{d\left(\frac{p_{t,j}}{A_j}\right)}
\]

\[\begin{align*}
\text{15.86}^{\text{th}} \text{ percentile for } \sigma \\
\text{median} \\
\text{50}^{\text{th}} \text{ percentile for } \rho 
\end{align*}\]
Understanding the methods
– a toy model study
Two component model: soft UE + hard contamination

soft component (UE)

- take the region of area $A$ in $(y, \phi)$ space $\rightarrow$ 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

$$\frac{dp_{t_1}}{dp_t} = \frac{1}{\mu} e^{-p_t/\mu}$$

- parameters:
  - $\mu$ – 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

<table>
<thead>
<tr>
<th>Soft component (UE)</th>
<th>Hard component (ISR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>▶ take the region of area $A$ in $(y, \phi)$ space → transverse region (traditional approach) or jet area (area/median approach)</td>
<td>▶ soft and collinear partons from primary emissions: $\frac{dn}{dp_t dy dy} \sim \frac{C_i}{\pi^2} \frac{\alpha_s(p_t)}{p_t}$</td>
</tr>
<tr>
<td>▶ number of particles in this region, $n$, given by Poisson distribution with the average $\langle n \rangle$</td>
<td>▶ hard scale cut $Q = \frac{1}{2} p_t = 50$ GeV</td>
</tr>
<tr>
<td>▶ single-particle $p_t$ distribution given by $\frac{dpt_1}{dp_t} = \frac{1}{\mu} e^{-p_t/\mu}$</td>
<td>▶ partons distributed uniformly in angle and rapidity</td>
</tr>
<tr>
<td>▶ parameters: $\mu$ – average $p_t$ of particle, $\nu = \frac{\langle n \rangle}{A}$ – density of particles</td>
<td></td>
</tr>
<tr>
<td>▶ in this model $\rho = \mu \nu$ is the true value of $p_t/A$ of UE</td>
<td></td>
</tr>
</tbody>
</table>
In the toy model: the same $\rho$ distribution used to generate all events
- however, there are event-to-event fluctuations of $\rho$ due to restricted area
- this sets the lower limit for the uncertainty of $\rho$ determination
Fluctuations in estimation of $\rho$

In the toy model: the same $\rho$ distribution used to generate all events

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

![Graph showing fluctuations in estimation of $\rho$](attachment:image.png)

- $S_d/\rho = 0.90$
- $S_d/\rho = 1.81$

Sebastian Sapeta (LPTHE, Paris)

On strategies for determination and characterization of the underlying event
Fluctuations in estimation of $\rho$

In the toy model: the same $\rho$ distribution used to generate all events

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

$\nu = 5$

$S_d/\rho = 0.33$

$S_d/\rho = 0.90$

$S_d/\rho = 1.81$

traditional approach suffers significantly more from the hard contamination

$S_d \sim Q$
Approaching real life
– Monte Carlo study
Average $\rho$ as a function of $y$

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

![Graph showing $\langle \rho \rangle$ as a function of $y$ for different event generators.]

- significant $y$ dependence
- strips of $\Delta y = 2$ sufficient for robust $\rho$ determination
Fluctuations from event to event

\[ S_d/\langle \rho \rangle \]

\( pp, \sqrt{s} = 10 \text{ TeV}, \, a_{k_t+C/A}, R=0.6 \)

- Herwig 6.510 + Jimmy 4.31
- Pythia 6.4.21 DWT
- Pythia 6.4.21 DW
- Pythia 6.4.21 S0A
Fluctuations

- from event to event

\[ S_d / \langle \varphi \rangle \]

\[ \langle \sigma \rangle / \langle \rho \rangle \]

- within an event

\[ pp, \sqrt{s} = 10 \text{ TeV}, a-k_t+C/A, R=0.6 \]

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

\[ \text{corr}(y_1, y_2) = \frac{\langle \rho(y_1)\rho(y_2) \rangle - \langle \rho(y_1) \rangle \langle \rho(y_2) \rangle}{S_d(y_1)S_d(y_2)} \]

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

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

\[ \text{corr}(y_1, y_2) = \frac{\langle \rho(y_1) \rho(y_2) \rangle - \langle \rho(y_1) \rangle \langle \rho(y_2) \rangle}{S_d(y_1) S_d(y_2)} \]

\( y_1, y_2 \) – 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 \( \Leftrightarrow \) 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
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 $\rho$ on rapidity
- fluctuations from event to event (large for all generators/tunes)
- fluctuations within an event, $\sigma$, (significant differences between Herwig+Jimmy and Pythia)
- correlations (large differences between Herwig+Jimmy and Pythia)
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 $\rho$ on rapidity
- fluctuations from event to event (large for all generators/tunes)
- fluctuations within an event, $\sigma$, (significant differences between Herwig+Jimmy and Pythia)
- correlations (large differences between Herwig+Jimmy and Pythia)

→ for more details see: arXiv:0912.4926
BACKUP SLIDES
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 \( \nu_g \) 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 $\rho$)

#### Traditional approach

\[
\begin{align*}
\langle \rho_{\text{ext,av}} \rangle &= \rho \\
\langle \rho_{\text{ext,min}} \rangle &= \rho - \sigma/\sqrt{\pi A_{\text{Trans}}} \\
\langle \rho_{\text{ext,max}} \rangle &= \rho + \sigma/\sqrt{\pi A_{\text{Trans}}}
\end{align*}
\]

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

\[
\sigma/\sqrt{\pi A_{\text{Trans}}} \approx 20 - 30\%
\]
Toy model: soft UE (extraction of $\rho$)

### Traditional approach

\[
\begin{align*}
\langle \rho_{\text{ext, av}} \rangle &= \rho \\
\langle \rho_{\text{ext, min}} \rangle &= \rho - \sigma / \sqrt{\pi A_{\text{Trans}}} \\
\langle \rho_{\text{ext, max}} \rangle &= \rho + \sigma / \sqrt{\pi A_{\text{Trans}}}
\end{align*}
\]

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

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

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

\[\begin{align*}
\langle \rho_{\text{ext}} \rangle / \rho &= 1.2 \\
\langle \rho_{\text{ext}} \rangle / \rho &= 1 \\
\langle \rho_{\text{ext}} \rangle / \rho &= 0.8 \\
\langle \rho_{\text{ext}} \rangle / \rho &= 0.6 \\
\langle \rho_{\text{ext}} \rangle / \rho &= 0.4 \\
\langle \rho_{\text{ext}} \rangle / \rho &= 0.2 \\
\langle \rho_{\text{ext}} \rangle / \rho &= 0
\end{align*}\]

\[\begin{align*}
\text{d}P_1 / \text{d}p_t &= e^{p_t} \\
\text{analytic approx.} \\
\text{median jet } \rho
\end{align*}\]

\[\begin{align*}
R \text{ or } \sqrt{(A_{\text{tile}} / c_J)}
\end{align*}\]
Two component model: biases

Traditional approach

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

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

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

$\mathcal{P}$ – fraction of events with perturbative radiation smaller than soft fluctuations
Two component model: biases

Area/median approach

\[
\langle \rho_{\text{ext}} \rangle \simeq \langle \rho_{\text{ext}}^{(\text{soft})} \rangle + \sqrt{\frac{\pi c J}{2}} \sigma R \frac{\langle n_h \rangle}{A_{\text{tot}}}
\]

\[
\langle n_h \rangle - \text{number of perturbative part.} \quad \sigma - \text{measure of fluctuations} \quad \rho - \text{true value of } p_t/A
\]

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

- the two terms bias \( \langle \rho_{\text{ext}} \rangle \) in opposite directions
- for \( R \approx 0.5 - 0.6 \) (used in most MC analysis of UE) the biases largely cancel
- similar picture and conclusions for \( \sigma \)
Two component model: biases

Area/median approach

Two component UE
2 Born, $|y| < \pi$

$\langle \rho_{\text{ext}} \rangle / \rho$

$R = \sqrt{A_{\text{tile}}/c_{\text{J}}}$

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 $R \simeq 0.5 - 0.6$ (used in most MC analysis of UE) the biases largely cancel
<table>
<thead>
<tr>
<th>quantity</th>
<th>method</th>
<th>result</th>
<th>numerical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\frac{\delta \rho^{(\text{soft})}}{\rho})</td>
<td>TransAv</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(\frac{\delta \rho^{(\text{hard})}}{\rho})</td>
<td>TransMin</td>
<td>(-\frac{\sigma}{\rho^2} \frac{P}{\sqrt{\pi} A_{\text{Trans}}})</td>
<td>(-0.09)</td>
</tr>
<tr>
<td>Area/Med*</td>
<td>TransAv</td>
<td>(\frac{C_i \alpha_s Q}{\pi^2 \rho})</td>
<td>0.99</td>
</tr>
<tr>
<td>Area/Med</td>
<td>TransMin</td>
<td>(2 \left( \frac{C_i \alpha_s}{\pi^2} \right)^2 \frac{A_{\text{Trans}} Q}{\rho})</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>Area/Med</td>
<td>(\frac{\sigma R}{\rho} \sqrt{\frac{\pi c_J}{2}} \left( \frac{n_b}{A_{\text{tot}}} + \frac{C_i L}{\pi^2 2b_0} \right))</td>
<td>0.17</td>
</tr>
<tr>
<td>quantity</td>
<td>method</td>
<td>result</td>
<td>numerical value</td>
</tr>
<tr>
<td>------------------</td>
<td>------------</td>
<td>------------------------------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>$S^{(soft)}_d$</td>
<td>TransAv</td>
<td>$\frac{\sigma}{\rho} \sqrt{\frac{1}{2A_{\text{Trans}}}}$</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>TransMin</td>
<td>$\frac{\sigma}{\rho} \sqrt{\frac{\pi - 1}{\pi A_{\text{Trans}}}}$</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Area/Med</td>
<td>$\frac{\sigma}{\rho} \sqrt{\frac{\pi}{2A_{\text{tot}}}}$</td>
<td>0.22</td>
</tr>
<tr>
<td>$S^{(hard)}_d$</td>
<td>TransAv</td>
<td>$\sqrt{\frac{C_i \alpha_s}{4A_{\text{Trans}} \pi^2}} \frac{Q}{\rho}$</td>
<td>1.72</td>
</tr>
<tr>
<td></td>
<td>TransMin</td>
<td>$\frac{C_i \alpha_s}{\pi^2 \sqrt{2}} \frac{Q}{\rho}$</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>Area/Med</td>
<td>$\frac{\sigma R}{\rho} \sqrt{\frac{2\pi c_J}{A_{\text{tot}}}} \left( \frac{n_b}{A_{\text{tot}}} + \frac{C_i L}{\pi^2 \sqrt{2} b_0} \right)^{\frac{1}{2}}$</td>
<td>0.19</td>
</tr>
</tbody>
</table>
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 $\rho$ with $R$ produced by the tails of distributions of $p_t/A$
Average $\rho$ as a function of $y$

- dijets at the LHC, $\sqrt{s} = 14$ TeV, $p_{t,\text{min}} = 50$ GeV
Average $\rho$ as a function of $y$

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