Track-based Measurements of the Underlying Event at 900 GeV & 7 TeV



Gabriel Hare, on behalf of the ATLSA Underlying Event analysis group.

31 May 2010

Thanks to...

- The Early Underlying Event analysis group: Andy Buckley, Sergei Chekanov, Deepak Kar, Arthur Moraes, Jason Nielsen, Jimmy Proudfoot, Arno Straessner, Sebastian Wahrmund, Ben Wynne, Rik Yoshida.
- The many people from the Minimum Bias analysis group whose work provides the basis on which this analysis is built, and who were always willing to answer questions.
- \star The D3PDs, the efficiency corrections, uncertainties and event and track selections are taken with gratitude (but without implication of responsibility) from the Minimum Bias analysis.
- Emily Nurse, who as the soft QCD convener has provided crucial advice and consideration of these results.
- Keith Baker, Stan Lai, and Antonio Zoccoli, who have reviewed these analysis methods, and offered advice for future analyses.
- Borut Kersevan, Stan Lai, Dominique Pallin, Dan Tovey who reviewed the note summarizing this analysis.



The many other people who have contribute directly or indirectly. Especially, all those who have ensured the excellent performance of the ATLAS Inner Detector.

Outline

- Underlying Event: Intro.
- Results: The Underlying Event, and Other Regions
 - Number Densities
 - p_T Densities
 - Standard Deviations
 - Number Density φ Profiles
- Analysis Methods:
 - Track Corrections
 - Event Corrections
 - Correction Factors
- Conclusions



Underlying Event: Intro.

- <u>Goal</u>: Isolate the low-energy QCD contribution to events (in a Minimum Bias sample) that is independent of the principal scattering energy. This contribution will be referred to as the "Underlying Event".
- <u>Method</u>: Assume a Di-Jet structure for events. The φ region that is nearly transverse to the Di-Jets is assumed to be principally filled by the Underlying Event.
- <u>Method</u>: At low energies it is sufficient to use the highest p_T (leading) track T₁, rather than the highest E_T (leading) jet. (The leading track is usually found in the leading jet.)
- <u>Definition</u>: The intervals with $\pi/3 < |\phi| < 2\pi/3$ will be the "Transverse" region that is used to characterize the Underlying Event.
- <u>Definition</u>: The interval within π/3 of the leading track and the region within π/3 opposite the leading track ate the "Toward" and "Away" regions, respectively.



Results

Profile	Abscissa	Ordinate	
Transverse Number Density	Lead Track p _T	Transverse Track Number / Transverse Area	
Away Number Density	Lead Track p _T	Away Track Number / Away Area	
Toward Number Density	Lead Track p⊤	Toward Track Number / Toward Area	
Transverse p⊤ Density	Lead Track p⊤	Transverse Track pT / Transverse Area	
Away p⊤ Density	Lead Track p⊤	Away Track pT / Away Area	
Toward p⊤ Density	Lead Track p⊤	Toward Track p _T / Toward Area	
Standard Deviations	Lead Track p⊤	S.D (Transverse Track Number / Transverse Area)	
	Lead Track p⊤	S.D. (Transverse Track p _T / Transverse Area)	
Transverse Mean Track pT	Transverse Track Number	Event-Mean of Transverse Track p _T	
Away Mean Track pT	Away Track Number	Event-Mean of Away Track pT	
Toward Mean Track pT	Toward Track Number	Event-Mean of Toward Track pT	
Angular Distributions	φ from Lead Track	Number Density of Tracks in φ Bin Interval	





Transverse Number Density



- The number density is higher than predicted by any of the MC tunes.
- The difference is clearer in events with lower lead p_T.
- The difference is more significant at 7 TeV.



All Number Densities



- The **Toward & Away** regions are dominated by jet-like activity, yielding rising number densities.
- In contrast, the number density in the Transverse region appears to be independent of the energy scale defined by the p_T of the leading track.
- Leading track selection bias yields a cross-over of the Toward & Away profiles in both Data and Monte Carlo.



Transverse pt Density



- The higher number density implies a higher p_T density as well.
- These profiles characterize the contribution of the Underlying Event to jets from the hard scatter.





All pt Densities



- The Toward & Away regions have p_T densities that are nearly linear functions of the event energy scale defined by the pT of the leading track.
- The proportionality of the pT densities in the **Toward & Away** regions indicates the extent of the variation in the charged fraction of the total energy in each region.
- At low leading p_T Away ~ 2/3 Toward and contains ~ 1 track.

Standard Deviations





- Event to event p_T fluctuations are on the scale of average values.
- S.D. provides an additional constraint for M.C. tuning.



Transverse Mean Track pt



- The abscissa counts the number of tracks in the Transverse region only.
- The ordinate is the mean p_T of tracks in the Transverse region only.





All Mean Track pt



- The profiles of the mean p_T as a function of the number of tracks are sensitive to the hadronization model.
- As was seen in the Minimum Bias analysis, the mean track p_T at 7 TeV is lower than predicted in high multiplicity regions
- ATLAS TO ATLAS
- The mean track p_T profiles are essentially independent of the energy scale of the collisions.
- The **Transverse** & **Away** profiles match.

Toward

Away

Transverse

Transverse

Angular Distributions



- The number density is for tracks other than the leading track. The plots are reflected, so the same data points appear twice.
- As the p_T of the leading track increases, the development of 'jet-like' region of higher density is observed adjacent to and opposite the leading track.
- The number density is both higher and has a different angular distribution than was predicted.





Analysis Methods



Track Selection

- <u>Selected Tracks</u>: pT > 500 [MeV], |η| < 2.5 [iRad], |d_{0Vtx}| < 1.5 [mm], |z_{0Vtx} · Sin(θ)| < 1.5 [mm], ≥ 1 Pixel Hit, ≥ 6 SCT Hits.
 - These are used to fill profiles.
 - They are also used to characterize the trigger efficiency.
- <u>Beam-Spot Tracks</u>: pT > 500 [MeV], |η| < 2.5 [iRad], |d_{0BS}| < 4 [mm],
 ≥ 1 Pixel Hit, ≥ 6 SCT Hits.
 - These are used to characterize the trigger and vertex reconstruction efficiencies.
- <u>Vertex Tracks</u>: pT > 100 [MeV], |η| < 2.5 [iRad], |d_{0BS}| < 4 [mm], ≥ 1
 Pixel Hit, ≥ 4 SCT Hits, ≥ 6 Si Hits (+1 in Pixel or SCT).
 - 2+ vertex tracks are required to construct a vertex.
 - Vertex tracks are not used to parameterize efficiencies.
- Beam-Spot Tracks are used because:
 - Their efficiencies are presently well characterized due to the p_T > 500 [MeV] requirement
 - The dependency on the vertex reconstruction is avoided by selecting with respect to the beam spot perigee parameter d_{OBS}.



- In addition to the possibility to fail to reconstruct a true charged particle as a track, there are three sources of "fake" tracks that were considered.
- 1. Fake tracks, which are those that cannot be matched to some true charged particle according to some matching criteria.
 - Cone matched if $\Delta R < 0.05$
 - Hit matched if $r_{hit} > 0.55$ with weights 10, 5, 1 for Pixel, SCT, TRT.
 - For tracks with $p_T < 20$ [GeV] the fake fraction is $< 10^{-4}$.
- 2. Secondary tracks resulting from decays (or material interactions) that are identified instead as primary tracks. (Secondary fraction ~ 0.02.)
- Tracks that pass all selection criteria whose matched true particles are outside of the kinematic range, because either p_T is too low or η is too high. (As high as ~0.2, but only at p_T & η edges.)
 - A charged primary stable particle with $\eta_{true} > 2.5$ can be reconstructed if the vertex is displaced towards -z, and will pass the selection criteria if $\eta_{rec} < 2.5$.
- The fractions f_{fake}, f_{sec}, f_{okr} are also estimated as functions of p_T and η, and are compensated for by additional weight factors (1-f_X).



- All of the profiles considered here can be considered as derivations from 2 dimensional distributions, second dimension having bins in the Y-axis values.
 - For the mean transverse pT density the relevant distribution is the sum of the track p_T in the transverse region.
 - For the standard deviation the relevant distribution is the square of the sum of the track p_T in the transverse region.
 - For the mean track pT the relevant distribution is for individual track p_T normalized to the number of tracks.
- In the absence of migration with respect to X-axis bins, and in the absence of event selection bias, the track weights w_{trk} are sufficient to correct <u>only</u> the mean values of the Y-axis distributions.
 - However, the standard deviation cannot be derived from the weighted mean distribution... they are considered separately.
- In the case of the mean track p_T, for a given number of reconstructed tracks, the weights yield a corrected p_T distribution whose normalization is to the corrected number of tracks.
- ATURS OF

 This ensures that the mean track p_T is correct, although it may be associated with the incorrect X-axis bin.

 Including the correction weights discussed above, the aggregation weight associated with each individual track is:

$$w_{trk}(p_{T},\eta) = \frac{1}{\varepsilon_{trk}} * (1 - f_{fake}) * (1 - f_{sec}) * (1 - f_{okr})$$

- The reconstruction characterizations necessarily have some uncertainty σ_{trk} which implies a systematic uncertainty for the corrected distributions.
- This uncertainty is estimated by making three versions of the corrected profiles, one corrected using ϵ_{trk} , and two using $\epsilon_{trk}\pm\sigma_{trk}$.



Event Weighting

- Events are included in the sample if:
 - The Minimum Bias Trigger Scintillator has at least one hit. (This is the L1_MBTS_1 trigger.)
 - At low luminosity this selects events in which inelastic collisions occurred.
 - At least one primary vertex candidate was reconstructed.
 - When more there is more than one candidate primary vertex the vertex with the highest summed |p_T|² of its associated tracks is selected as the event primary vertex.
 - No additional primary vertex candidates with more that 3 associated beam-spot tracks was reconstructed.
 - This selects events without pile-up, while accepting events in which particles decays occurred along the beam-line.
 - There is at least one track with $p_T > 1.0$ [GeV].
 - This reduces the diffractive contribution to < 1% in the simulated events from all of the generators that were considered.
 - Consequently the leading track p_T > 1.0 [GeV]



Event Weighting

- Events with less activity are less likely to yield hits in the MBTS, which would bias the events toward higher multiplicity.
 - The L1_MBTS_1 efficiency $\epsilon_{trig}(n_{BS})$ is ~ 0.99 in events with $n_{BS}=1$, and ~1 in events with $n_{BS}>2$.
- Events with less activity are also less likely to yield a successfully reconstructed vertex.
 - The vertex reconstruction efficiency $\epsilon_{trig}(n_{BS})$ is ~ 0.89 in events with $n_{BS}=1$ and ~1 in events with $n_{BS}>3$.
 - In general, the reconstruction efficiency is reduced if the beamspot tracks have higher η, so the full efficiency characterization is ε_{trig}(n_{BS}, η_{BS}) where η_{BS} is the average beam-spot track η. However this correction is only significant when n_{BS}=1.
- Finally, the requirement that there be at least one track with p_T > 1.0 [GeV] introduces a probability that an event will not be selected when none of the charged stable particles with p_T > 1.0 [GeV] are successfully reconstructed. The associated efficiency ε_{trig} essentially characterizes the probability to find an acceptable leading track.

$$w_{ev} = \frac{1}{\varepsilon_{trig}(n_{BS})} * \frac{1}{\varepsilon_{vert}(n_{BS}, \eta_{BS})} * \frac{1}{\varepsilon_{lead}}$$



Migration Effects

- For the mean and standard deviation of the transverse track p_T density there are only two migration effects, both due to the possibility that the reconstruction will miss the true highest pT charged primary stable (CPS) particle, or that the reconstruction yields a fake highest p_T track.
- 1. If the highest p_T CPS particle is missed, the second highest p_T CPS may be associated with the leading track instead.
 - For the plateau region of the profile this effect is minor. However, in the turn-on region from 1-3 [GeV] the migration yields a mean p_T density that is too high.
- If the highest p_T CPS particle is missed, orientation of the reconstructed event will not be consistent with the orientation of the true event.
 - The effect in this case is that the Transverse region may receive contributions from the Toward & Away regions where there is jet-like activity.
 - This is most significant in the plateau region of the profile, and yields an increase in the mean p_T density.



Correction Factor Method

- The bin multiplier correction factors are derived using the ATLAS MC09 tune of Pythia 6.4.
- The generated events are fully simulated by GEANT 4 using comparable conditions (disabled modules) to those of the runs during which the data would be collected. (Also a comparable misalignment is introduced.)
 - The z-position of the simulated events has a wider distribution than was measured in data, so it was necessary to assign a sample weight $\omega_{ev}(z)$ to the simulated events.
- These simulated events were also used to derive the reconstruction efficiency and fake fractions, (With many cross-checks to distributions in data.) which are used to determine the individual track weights.
- To derive the correction factors, the profiles are derived using the same track & event weighting as is used in the data.
- The bin multiplier is defined to be the ratio of the true value over the reconstructed & corrected value.



• Comparing the derived correction factors to an alternative derived using PhoJet it was found that a flat 2% uncertainty in the correction factors covered the largest discrepancies.

Correction Factor Method

 If there are insufficient statistics the correction factor m_{bin}(x) will not be well estimated.

$$m_{bin}(x) = \frac{v_{true}(x)}{v_{reco}^{corr}(x)}$$

- The uncertainty associated with the numerator $\varepsilon_{\rm true}({\bf x})$ is entirely statistical.
- The statistical uncertainty associated with the denominator $\varepsilon_{reco}^{corr}(x)$ will be larger, since the reconstruction inefficiencies will yield a wider distribution of values.
- The independent uncertainties associated with an unfolding involving limited statistics are estimated as:

$$\varepsilon_{\text{stat}}(\mathbf{x}) = \varepsilon_{\text{true}}(\mathbf{x})$$
$$\varepsilon_{\text{syst}}(\mathbf{x}) = \left| \varepsilon_{\text{true}}(\mathbf{x}) - \varepsilon_{\text{reco}}^{\text{corr}}(\mathbf{x}) * \frac{\mathbf{v}_{\text{true}}(\mathbf{x})}{\mathbf{v}_{\text{reco}}^{\text{corr}}(\mathbf{x})} \right|$$



- In the case of perfect efficiency $\varepsilon_{syst}(x) = 0$.
- In the case of a single outlying measurement $\varepsilon_{reco}^{corr}(x)$ will be large, and so the systematic error will be increased.

Conclusions

- No current Monte Carlo tune provides a well matched description of <u>all</u> of the measurements of the Underlying Event, although the MC09 tune of Pythia already describes many of the characteristics of the Underlying Event.
- The measurements that presently seem to provide the most useful constraints are:
 - The low to intermediate leading p_T range of the number density in the transverse region, preceding the plateau region.
 - The number density as a function of angle from the leading track.
- Measurements at two energy points will enable tunes based on ATLAS measurements to be extrapolated to higher energies.
- The figures (including eps format) & CONF note are available here:
 - <u>https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/</u> <u>ATLAS-CONF-2010-029/</u>



Plans

- Plans for future measurements include (but are not limited to)
 - Topo-cluster measurements
 - Jet oriented event topologies
 - Event Shape oriented event topologies
 - Drell-Yan oriented event topologies
 - Rapidity gap topologies
- Additional profiles will include (but are not limited to)
 - p_T density as a function of ϕ
 - Mean track p_T as a function of leading track p_T
 - Transverse Minimum & Maximum profiles
 - The "Minimum" side of the transverse region contains the lower summed p_T of tracks.
- Analysis methods will use regularized migration inversion, when applicable.
 - These methods require either very well known efficiencies...
 - Or, a migration that is independent of the Monte Carlo model.



Additional Information



Underlying Event: What





- <u>Model dependent</u> contributions include:
- Multiple Parton Interactions (MPI) becomes nearly independent of the hard scatter, when the hard scatter energy > 10 GeV.
- Initial State Radiation (ISR) has an angular distribution that is nearly independent of the hard scatter and can contribute to the transverse region.

Event Generation Models

Pythia 6.4: MC09 Tune	The ATLAS collaboration's tune to CDF data at taken at 630 GeV and 1.8 TeV. Uses pT -ordered shower & color reconnection. The ISR and MPI cutoff scales are tuned separately. Uses the MRST LO* PDF.
Pythia 6.4: Perugia0 Tune	By Peter Skands. Mostly tuned using Tevatron and SppS Minimum-Bias data. Uses p_T -ordered shower and CTEQ5L pdf.
Pythia 6.4: DW Tune	By Rick Field. Maximal ISR, virtuality-ordered shower. To fit data on the di-jet angular distribution measured by the D0 collaboration.
PhoJet	Dual Parton Model based, using pomeron exchange for soft QCD interactions. Incorporates a model for high-mass diffraction dissociation including multi-jet production.



Diffraction in Monte Carlo

- In all of the Monte Carlo models considered, the requirement of at least on track with pT > 1.0 GeV reduced the diffractive contribution to less than 1%.
- Although the identification of "diffractive" events is ambiguous, the requirement on the leading track avoids kinematic regions of where the Monte Carlo models depend on the tuning of their diffractive process models.
 - For the profiles in which the abscissa is the highest track p_T in the event, the track p_T & number densities would be slightly increased if the diffractive components were not included in the Monte Carlo samples.



- The track reconstruction begins in the Pixel tracker (1 hit required) and is extended to the SCT tracker (6 hits required) and when possible is further extended into the TRT tracker.
 - In the case of material interactions (or decays) the extension of a track from the Pixel tracker into the SCT tracker might not succeed.
 - Due principally to material interactions the track efficiency depends on both the p_T and detector-η of the track.
 - Other efficiency dependencies, such as charge, have not yet been fully characterized.
 - In the worst case (low p_T & high η) the efficiency ε_{trk} can be < 0.6, while central tracks with p_T > 10 GeV have efficiencies > 0.9.
- This effect can be compensated for in every case considered here by weighting individual tracks, with a weight factor 1/ε_{trk}(p_T,η).



Track Correction

- Define $P(T,p_T)$ to be the distribution for track momentum p_T , and track number T, with P(T) the p_T distribution normalized to T.
 - A sample for the n event drawn from this distribution yields $\overline{T}[n]$ tracks, where the t track has momentum $p_T[n,t]$.
- Suppose that we are interested in the total p_T of tracks y in (a region of) an event. Using the distribution $P(T,p_T)$ this is simply:

$$M^{1}(y) = \int_{T,p_{T}} p_{T} * P(T,p_{T}) \approx \sum_{n}^{N} \sum_{t}^{\overline{T}} \overline{p_{T}}[n,t]$$

- Suppose that there is p_T dependent track finding efficiency F(pT), and a T dependent vertex finding efficiency G(T).
 - A sample $\overline{F}[t]$ or $\overline{G}[n]$ drawn from an efficiency is $\in \{0, 1\}$.
 - If no corrections are applied the measured distribution converges to G(T)*F(p_T)*P(T,p_T).
- Applying the weight $\hat{F}(\overline{p_T})^{-1}$ to each track, and $\hat{G}(\overline{T})^{-1}$ to each event, the measured distribution converges to a function of the measured track number \overline{T} and the measured momentum p_T normalized to the corrected number of tracks: $\hat{P}(\overline{T}, p_T)$



Track Correction

• Making a measurement of y, which is the total track p_T corrected for the efficiency, is simply a matter of including the correction weights.

$$\mathsf{M}^{1}(\mathsf{y}) \approx \sum_{\mathsf{n}}^{\mathsf{N}} \widehat{\mathsf{G}}(\overline{\mathsf{T}}[\mathsf{n}])^{-1} \sum_{\mathsf{t}}^{\mathsf{T}} \widehat{\mathsf{F}}(\overline{\mathsf{p}}_{\mathsf{T}}[\mathsf{n},\mathsf{t}])^{-1} * \overline{\mathsf{p}}_{\mathsf{T}}[\mathsf{n},\mathsf{t}]$$

• The event-to-event variation of $M^1(y)$ is used in the definition of the statistical error of a measurement of $M^1(y)$. In this case, simply square the result of the weighted sum over t, and weight by $\widehat{G}(T)^{-1}$.

$$\mathsf{M}^{2}(\mathsf{y}) \approx \sum_{\mathsf{n}}^{\mathsf{N}} \widehat{\mathsf{G}}(\overline{\mathsf{T}}[\mathsf{n}])^{-1} \left(\sum_{\mathsf{t}}^{\mathsf{T}} \widehat{\mathsf{F}}(\overline{\mathsf{p}}_{\mathsf{T}}[\mathsf{n},\mathsf{t}])^{-1} * \overline{\mathsf{p}}_{\mathsf{T}}[\mathsf{n},\mathsf{t}]\right)^{2}$$

- In the case of the event-mean track p_T, the weighted sum of p_T is divided by the weighted track count.
- For higher moments of the event track p_T distribution (such as $(p_T)^2$) consider a distribution P(T, $(p_T)^2$) defined as before.



• CONCLUSION: Weighting by 1/ε is correct!

Migration of Lead Track ϕ



- The systematic uncertainty associated with the bin-by-bin unfolding for φ re-orientation is due to the uncertainty in the modeling of the angular correlation between leading & sub-leading tracks.
- The difference between Pythia and PhoJet in the transverse region is equivalent to the difference between Pythia and Data. Consequently, the uncertainty associated with the unfolding is sufficient to account for this model dependency, since it allows for variation in the opposite direction.



Systematic Uncertainties

	Lowest $p_{\rm T}$ bin	Intermediate $p_{\rm T}$ bin	Highest $p_{\rm T}$ bin	
Systematic uncertainty on unfolding				
Difference between PYTHIA and PHOJET		2% (everywhere)		
Statistical uncertainty on PYTHIA unfolding	0.1%	4.0% (7%)	10% (15%)	
Systematic uncertainties from efficiency corrections				
Track reconstruction	5.5%	5%	5%	
Leading track requirement	1.0%	0.1%	< 0.1%	
Trigger and vertex efficiency		< 0.1% (everywhere)		
Total from efficiency corrections	4.5%	5%	5%	
Total systematic uncertainty	5%	7% (9%)	11.5% (16%)	

- These uncertainties are estimated from the transverse p_T density profiles, but are equal or less for all other profiles.
- 900 GeV: the lowest p_T bin refers to 1.0 1.5 GeV, the intermediate p_T bin refers to 4 5 GeV, and the highest p_T bin refers to 9 10 GeV.
- 7 TeV: the lowest p_T bin refers to 1.0 1.5 GeV, the intermediate p_T bin refers to 9 - 10 GeV, and the highest p_T bin refers to 18 - 20 GeV.
- At intermediate and high p_T the lack of statistics in the Monte Carlo samples is the dominant contribution to the systematic uncertainty associated with the corrections.



Away Number Density



• The gradual rise is expected for the number density associated with jet-like activity in the Away region.





Toward Number Density



- The plateau in number density in the Toward region is due to a selection bias.
- In a di-jet event, the jet with the single highest p_T constituent has less energy to be shared among additional hadrons.



Away pt Density



 Multiplying by the area of associated with the Away region, the linear rise corresponds to densities whose total p_T nearly balances that of the leading track alone.





Toward pt Density



 Multiplying by the area of associated with the Toward region, the p_T density is nearly twice what would be found if the leading track were the only track in the region.





Away Mean Track pt



- The abscissa & ordinate refer to the tracks in the Away region only.
- The profiles are very similar to the Transverse region's.





Toward Mean Track pt



- The abscissa & ordinate refer to the tracks in the Toward region only.
- The mean pT in the case of a low number of tracks is constrained by the requirement that the leading track have p_T > 1.0 GeV

