Experimental Considerations for Jet Substructure Reconstruction – Present and Future

Peter Loch
University of Arizona
Definitively a way to go…
  • Hadronic W tagging
  • Top tagging
  • Shifting background jet masses out of signal regions

Increasing signal significance
  • Access “boosted” regimes where resolved signal reconstruction becomes problematic
  • Shifting background out of signal regions – e.g., QCD jet masses in pruning
  • Using substructure in fat jets to explore sub-jet correlations – reduction of combinatorics etc.
Understanding the measured subjet features

- Goes surprisingly well...
  - Jet masses
  - Splitting scales
- Comparisons to LO and NLO generators
  - Jet masses
  - Splitting scales
  - N-subjetiness
- MC describes features well

Anti-kt $R = 1.0$, ungroomed (left), trimmed $f=0.05$, $R_{filt} = 0.3$ (right)
Understanding the measured subjet features

- Goes surprisingly well...
  - Jet masses
  - Splitting scales
- Comparisons to LO and NLO generators
  - Jet masses
  - Splitting scales
  - N-subjetiness
- MC describes features well
  - But that may not be a complete surprise – we tuned PS/ISR to e.g. jet shapes...

Anti-kt $R = 1.0$, ungroomed (left), trimmed $f=0.05$, $R_{\text{filt}} = 0.3$ (right)
Intrinsic $k_T$ scale in calorimeters

- Intrinsic transverse momentum “kick” in hadronic showers
  - Order 400-500 MeV, typically
  - Contributes to lateral shower spread and (single hadron) thrust axis distortions – direction resolution in (sub)jets
  - Dissipates energy laterally – pT resolution
  - Note noise cut effect!
Pileup

(A. Schwartzman, talk at "Joint Snowmass-EuCARD/AccNet-HiLumi LHC meeting 'Frontier Capabilities for Hadron Colliders 2013', February 21, 2013)

- **ATLAS LAr calorimeter has a very large integration time relative to bunch spacing:**
  - **Out-of-time** pile-up contributions
  - Bi-polar shape compensates, on average, for both in-time and out-of-time pile-up, but out-of-time effects vary significantly within sub-detectors (eta-dependence)
  - ATLAS needs both in-time and out-of-time pile-up corrections
  - No directly handle on event-by-event out-of-time contribution
    - Cannot reduce out-of-time fluctuations

- **CMS is mostly insensitive to out-of-time pile-up:**
  - 2 time-slices (TS) for integration
- **CMS-like signal shapes**
  - Basically no out-of-time pile-up @ 50 ns bunch crossings
    - Will change @ 25 ns bunch crossings – not too dramatically as likelihood to hit same calorimeter cell in adjacent bunch crossings is low: highly suppressed signal fragments only!
  - Very similar for ATLAS central hadronic TileCal
- **ATLAS liquid argon calorimeter signal shapes**
  - Long charge collection time does not allow full signal integration (400 ns » 50/25 ns)
    - Use bi-polar electronic shaping function – translates initial current into amplitude proportional to (positive) energy and shapes current triangle to integral 0 pulse
    - Online build-in pile-up suppression cell by cell
• Slow signal collection in liquid argon calorimeters
  • ~450 ns @ 1 kV/mm drift time versus 40(20) MHz/25(50) ns bunch crossing time
    • Measure only $I_0 = I(t_0)$ (integrate <25 ns)
  • Applying a fast bi-polar signal shaping
    • Shaping time ~15 ns
    • With well known shape
    • Shaped pulse integral = 0
      • Net average signal contribution from pile-up = 0
    • Need to measure the pulse shape (time sampled readout)
  • Total integration ~25 bunch crossings
    • 23 before signal, 1 signal, 1 after signal

Peter Loch (University of Arizona)
$E_{\text{meas}} = E_{\text{true}} + \Delta E_{\text{PU}}$

$= E_{\text{true}} + \int_{-\infty}^{+\infty} s(t) \langle E \rangle_{\text{bx}} \, dt$

$= E_{\text{true}} + \langle E \rangle_{\text{bx}} \int_{-\infty}^{+\infty} s(t) \, dt$

$\approx E_{\text{true}} + \langle E \rangle_{\text{bx}} \sum_{n=-25}^{25} s(n \cdot \tau_{\text{bx}})$

$\approx E_{\text{true}}$

$\tau_{\text{bx}} = 25 \text{ ns (40 MHz)}$

Integration/summation includes in-time pile-up at $t = 0$, with $s(t = 0) = 1$!
\[ E_{\text{meas}} = E_{\text{true}} + \Delta E_{\text{PU}} \]
\[ = E_{\text{true}} + \langle E \rangle_{\text{bx}} \sum_{n=1}^{13} s(n \cdot \tau_{\text{bx}}) \]
\[ \approx E_{\text{true}} \]

\( \tau_{\text{bx}} = 50 \text{ ns (20 MHz)} \)

Bunch-crossings without collisions

(\( E_{\text{bx}} \approx 0 \))

2011/2012 data taking conditions in ATLAS
\[ E_{\text{meas}} = E_{\text{true}} + \Delta E_{\text{PU}} \]
\[ = E_{\text{true}} + \langle E \rangle_{bx} \sum_{n=0}^{8} s(n \cdot \tau_{bx}) \]
\[ \approx E_{\text{true}} + \delta \langle E \rangle_{bx} \]

\[ \tau_{bx} = 75 \text{ ns (}13.\overline{3} \text{ MHz)} \]

Bunch-crossings without collisions

\( E_{bx} \approx 0 \)
Pile-up Suppression in ATLAS

\[ E_{\text{meas}} = E_{\text{true}} + \Delta E_{\text{PU}} \]

\[ = E_{\text{true}} + \langle E \rangle_{\text{bx}} \sum_{n=0}^{4} s(n \cdot \tau_{\text{bx}}) \]

\[ \approx E_{\text{true}} + \Delta \langle E \rangle_{\text{bx}} \]

\[ \tau_{\text{bx}} = 150 \text{ ns (6.6 MHz)} \]

Bunch-crossings without collisions

\( (E_{\text{bx}} \approx 0) \)
$E_{\text{meas}} = E_{\text{true}} + \Delta E_{\text{PU}}$

$= E_{\text{true}} + \langle E \rangle_{\text{bx}} s(o)$

$\approx E_{\text{true}} + (\Delta \langle E \rangle_{\text{bx}} = \langle E \rangle_{\text{bx}})$

$\tau_{\text{bx}} > 625 \text{ ns (<1.6 MHz)}$

Bunch-crossings without collisions

($E_{\text{bx}} \approx 0$)

2010 data taking conditions in ATLAS
\[ w(\tau_{bx}) = \sum_{n=N_{\min}(\tau_{bx})}^{N_{\max}(\tau_{bx})} s(n \cdot \tau_{bx}) \]

\[ E_{\text{meas}} = E_{\text{true}} + \Delta E_{\text{PU}} \]
\[ = E_{\text{true}} + w(\tau_{bx}) \langle E \rangle_{bx} \]

Pile-up Suppression in ATLAS

decreasing pile-up history

bunch crossing spacing \( \tau_{bx} \) (ns)

increasing pile-up effect on signal
Pile-up Suppression in ATLAS

- Online pile-up suppression
  - Attempt to compensate signal contributions from in-time pile-up by suppression from physics signals in the pile-up history
    - Can only suppress average in-time pile-up contribution in any given calorimeter cell
  - Residual in-time pile-up effects still to be addressed offline
    - Non-gaussian fluctuations in signal from minimum bias events in a given cell – tails due to physics fluctuations and coherent noise effects
  - Distortion of pile-up history
    - LHC bunch structure and super-structure and bunch intensity fluctuations make pile-up history dependent on phase with respect to bunch gaps and intensity profile around triggered event
- Works better in high lumi/high bunch crossing scenario
  - Better cancellation with more stable average energy & finer integration of pulse shape
Pile-up Suppression in ATLAS: Bunch Trains

\[ \tau_{bx} = 25 \text{ ns} \]

best suppression \( \sim 2\% \)

over-compensation due to lack of next bunch-crossing w/collisions
Pile-up Suppression in ATLAS: Bunch Trains

\[ \tau_{bx} = 50 \text{ ns} \]

best suppression \( \sim 5\% \)

\( \text{in-time only} \)

2011 data taking conditions in ATLAS
Pile-up Suppression in ATLAS: Bunch Trains

\[ \tau_{bx} = 175 \text{ ns} \]

best suppression \( \sim 54\% \)

in-time only

Time in bunch train (ns)
Cell level pile-up suppression in 2012

- More local pile-up history information used
  - Bunch history and intensity profiles are known for each triggered event (low resolution, coarse measurement)
  - Pulse shape is known for each (class of) cell
  - Can estimate true weighted pile-up signal history energy pattern
- Suppressed phase dependence and sensitivity to bunch intensity fluctuations
  - Can be applied early in signal processing (online)

\[
E_{\text{meas}} = E_{\text{true}} + \Delta E_{\text{PU}} \\
\approx E_{\text{true}} + \langle E \rangle_{bx} \sum_{n=-1}^{13} s(n \cdot \tau_{bx}) I(n \cdot \tau_{bx})/\langle I \rangle \\
\approx E_{\text{true}}
\]
Pile-up Suppression in ATLAS: Bunch Trains

2012 data taking conditions in ATLAS

τ_{bx} = 50 ns

best suppression \sim 5\%

cell-level correction function
Caveats

- ATLAS cell-by-cell pile-up suppression still has average character
  - Ok to include as much information as possible from cell signal history, but the best one is not available at precision – no direct measurement of previous deposits!
    - Proton bunch intensity fluctuations only reflect changing likelihood of getting any energy into given cell – but not really how much...
  - Only coarse estimate available for intensity fluctuations
  - Cell signal fluctuations scale proportional to $\sqrt{L}$
- Topological cell clusters in ATLAS extract up- and downward fluctuations from pile-up
  - Residual dependence of response on pile-up conditions
  - Acceptance variations dependent on energy flow environment
Design guideline (~1992-95): pile-up noise of same order of magnitude as electronic noise (about true everywhere except FCal – very low electronic noise!)
Schematic view on pile-up – and a naïve way of dealing with it...
Schematic view on pile-up – and a naïve way of dealing with it...

Add pileup at a level corresponding to $\rho$ with local fluctuations $\sigma$
Schematic view on pile-up – and a naïve way of dealing with it...

Subtract baseline and apply "noise cut" (cut method)
Schematic view on pile-up – and a naïve way of dealing with it...

Select signals above noise (filter method)
Schematic view on pile-up – and a naïve way of dealing with it...

Select signals above noise (filter method)
Schematic view on pile-up – and a naïve way of dealing with it...

\[ p_T - \rho A > n\sigma \sqrt{A} \Rightarrow p_T^{\text{corr}} = p_T - \rho A \]

\[ p_T > \rho A + n\sigma \sqrt{A} \]

\( \sim \rho \)
Detector level pile-up suppression (1)

- Pile-up is an experimental condition typically adding energy
  - Ideally to be addressed at the level of signal reconstruction
  - Corrections need to suppress signal baselines (in-time pile-up) and the corresponding global (event-by-event physics) fluctuations
  - Corrections ideally should suppress local fluctuations from signal history (out-of-time pile-up) – provide stable signal at smallest possible (distance/energy) scales with high likelihood to be “true”
- Naïve noise-cut like corrections are likely too hard for substructure
  - Presence of hard signals in a (sub)jet suggest that close-by soft signals are valid/important/”true”... (e.g., generated by small angle radiation...)
  - Group of smallish signals in close proximity may indicate relevant signals – we tried very hard to keep coherent detector noise down!
  - Topological considerations appropriate – jet grooming techniques!

Peter Loch (University of Arizona)
Detector level pile-up suppression (2)

- First step of trimming provides well motivated signal collection
  - Small R jet finder follows topology/transverse energy flow inside jet
  - Connects small signals with close-by larger signals
- Initially constructed small R jets can be evaluated with respect to pile-up baseline and noise
  - Keep only those above threshold without modifying their $p_T$ (selection only)
  - Keep only those above threshold but subtract baseline $p_T$ as given by transverse energy density and small jet area
- Both approaches have been tested at particle level
  - Les Houches 2011 (Soyez, Francavilla, PL,...)
Trimming based pile-up suppression

Start with big jet : $C/A$, $R=1.2$
→ Recluster ($R_{\text{filt}}=0.3$)
  → Keep only sub-jets with $p_T > pA + 5\sigma \sqrt{A}$
  → Keep only subtracted sub-jets with $p_T > 5\sigma \sqrt{A}$

$R_{\text{filt}}$, $n_{\text{filt}}$, $p_T^{\text{cut}}$ being optimised

G. Soffez

WW events
10 PU (14 TeV)
Top mass reconstruction

\[ pp, \sqrt{s} = 10 \text{ TeV} \]
\[ t\bar{t}, 516 < \hat{p}_T / \text{GeV} < 1152 \]
(Pythia 8)
\[ \langle \mu \rangle = 10 \]
(Pythia8 SoftQCD Tune C4)

\[ K_T, R = 1.2 \]
Trimming, \( R_{\text{filt}} = 0.3 \)
\[ p_T^{\text{subj}} > \rho A^{\text{subj}} + 3 \sigma \sqrt{A^{\text{subj}}} \]

Peter Loch (University of Arizona)
Experimental observations

- Jet mass good test bed
  - Effect of in-time pile-up in 2010 data in ATLAS – clean (no out of time pile-up)

- Grooming techniques can mitigate pile-up effects
  - Also change the global mass –scale – dedicated calibration!

- Need to understand bias of grooming to remove pile-up
  - Possible scenarios are first trimming with rho, then grooming with technique optimized for analysis goal (signal/background ratio)
  - Can the resulting spectra be understood from point of phenomenology (possibly) and calculation (?)
Other experimental considerations

- Evaluation of tagger performances
  - It is not obvious that ATLAS and CMS can be directly compared with respect to tagging performance – different effects of pile-up, very different fluctuations in basic signals, different reconstruction cuts...
  - Comparisons and optimizations within one experiment is of course possible

- Can we get the detectors more out of the way?
  - Correction/unfolding of principal observables before substructure analysis?
  - Unfolding hard to imagine (is based on distributions)
  - Correction procedures may suffer from reliable/stable “truth” reference – especially for small signals
  - Clearly unfolding can be done at the level of distributions of scoring variables

- Some concerted effort would be nice...
  - Likely not much manpower/interest in the collaborations, but a few dedicated people will do!
Shower deconstruction deserves a harder look
- Need to understand effect of pile-up/corrections on deconstruction and evaluation variable
- Control of systematics not too obvious – need to find a very well modeled and simulated control region and understand extrapolation to other phase space locations
- Some activity in ATLAS – concerns about large local fluctuations started to be addressed (calorimeter increases parton shower fluctuations)

Qjet
- Interesting approach – but not clear how to establish systematic uncertainties

Focus on substructure observables most stable against loss of small signals
- Increasing luminosity will increase reconstruction thresholds (e.g., minimum pT for tracks, topo-cluster noise thresholds...)
- Evaluation of taggers at particle level should include scenarios with acceptance limitations for individual particles (pT > 500 MeV/1 GeV – magnetic field, dead material, reconstruction thresholds – detectors are intrinsically IR unsafe and collinear instable at some small scale)
Killing observables*

- **Correlations**
  - It seems many observables are just slight variations of the same underlying dynamics – the jet mass and (not entirely independent) pT
  - It would be nice find the set which has the largest de-correlated components – additional information content, increased sensitivity
  - Carefully designed BDTs?

- **Experimental concerns**
  - We cannot really measure e.g. 2-dim correlation matrix – systematics are extremely hard to control (errors in each bin can be 100% correlated)
  - Maybe we can translate uncertainties into groups of nuisance parameters with different correlations
  - Not completely deterministic as nuisance correlations may require assumptions – often only two options (0 and 100%)

*just remembered killing tunes at Les Houches 2011*
It is great to have all these various algorithms, even the ones seemingly using small (distance, kT, energy) scales!

- Need some focus/guidance on best performance for given final state/search goal including some experimental limitations early on (e.g., simple acceptance limitations)
- Can get some realistic estimates before embarking on long and resource heavy projects with potentially little to now gain (e.g., this workshop, BOOST, Les Houches,...)?

Better test beds – simple detector models?

- Some smearing and acceptance models for ATLAS/CMS-like detector (they are not that different in many respects related to jet substructure analysis) – PGD/DELFES etc.
Conclusions

- Need to understand our (CMS/ATLAS) respective signal definitions in increasingly hostile experimental environments (2015 - ?)
  - Keep up with ParticleFlow, TopoCluster performances in high pile-up environments
  - Machine configurations not completely clear yet (lumi, bunch crossing frequency...)
    - ATLAS liquid argon calorimeter likes 25 ns!!
    - But 50 ns may give us effectively more lumi to collect!
  - Performance in both tracking and calorimetry may be significantly altered
    - Sensitivities will change – lowest energy/highest momentum to be reconstructed at sufficient quality
    - Our microscopes get more blurry globally – need to use local constraints to get clearer pictures e.g. inside a jet
- Jet substructure is just emerging as a tool (next talk)
  - Much more important at higher energies and pile-up – despite the experimental challenges!
  - Focus on observables which can be calculated and reconstructed well!
  - Experimental and theoretical work ahead!

Peter Loch (University of Arizona)
Last but not least...

- **BOOST2013 is coming!**
  - August 12-16, 2013 – Flagstaff, Arizona, USA
  - First bulletin is out
  - Registration and webpages are live since March 2013
  - [http://w3atlas.physics.arizona.edu/boost2013/](http://w3atlas.physics.arizona.edu/boost2013/)
Last but not least...

- BOOST2013 is coming!
Backup:

Soft physics tuning
Soft physics tuning

- Description of non-perturbative features of p-p collisions important for e.g. jet shapes and pile-up modeling
  - Parton shower
  - Soft collisions
- Characterized by absence of models
  - Simple parameterization of (soft) radiation patterns observed in data
  - Energy scaling and pT cut-offs
  - Fragmentation and hadronization, FSR from LEP
- Often stressing parameterizations
  - Combined underlying event (PS, MPI) and MinBias (MPI) not always possible at sufficient quality
    - Not in PYTHIA6
    - PYTHIA8 seems to work (ATLAS tune A2)
- Note: we can only tune to unfolded distributions of observables!
- ATLAS performed 4-stage tuning with 2010 data (cleanest!)
  - (1) flavor production in fragmentation functions (LEP data)
  - (2) FSR and fragmentation (LEP data)
  - (3) ISR and primordial kT (ATLAS jet data)
  - (4) MPI tuning (ATLAS jet data for UE, ATLAS MinBias data for Pileup)
- Note that (1)-(4) are UE tunes while (4) can also include a MinBias tune
  - ATLAS provides two different tunes
- MinBias tune validation for 2012
  - Dedicated low mu run
- Both UE and MB tuning are relevant for substructure observables sensitive to small signals
  - Parton shower features
  - Pile-up features
ISR and Primordial kT Tuning (UE)

Before tuning to LHC data

After tuning to LHC data


(ATL-PHYS-PUB-2011-009)

Peter Loch (University of Arizona)
MinBias tuning features: charged multiplicities

\[ p_T^{\text{track}} > 500 \text{ MeV} \]

\[ p_T^{\text{track}} > 100 \text{ MeV} \]
MinBias tuning features: average $p_T$ versus charged multiplicities

$p_T^{\text{track}} > 500$ MeV

$p_T^{\text{track}} > 100$ MeV
Backup:

Clustering and noise suppression
Collect cell into energy “blobs”
- Idea is to collect all cell signals belonging to a given particle into one cluster of cells
  - Basically reconstruct the shower for each particle entering the calorimeter
- Needs algorithm to form energy blobs at the location of the shower signal in the calorimeter
  - Follow the shower-induced cell signal correlations

Extract most significant signal from all calorimeter cells
- Cluster formation uses signal significance as guidance
  - Not the total signal – noise changes from calorimeter region to calorimeter region
- Implicit noise suppression in cluster formation
  - Cluster signals should include least amount of noise
Collect cell into energy “blobs”
- Idea is to collect all cell signals belonging to a given particle into one cluster of cells
  - Basically reconstruct the shower for each particle entering the calorimeter
- Needs algorithm to form energy blobs at the location of the shower signal in the calorimeter
  - Follow the shower-induced cell signal correlations

Extract most significant signal from all calorimeter cells
- Cluster formation uses signal significance as guidance
  - Not the total signal – noise changes from calorimeter region to calorimeter region
- Implicit noise suppression in cluster formation
  - Cluster signals should include least amount of noise
Cluster seeding
- Defined by signal significance above primary threshold
  - Cells above this threshold can seed cluster

Cluster growth
- Defined by signal significance above secondary threshold
  - Cells neighbouring seeds with significance above this threshold drive cluster growths

Cluster signal
- Defined by cells with significance above basic threshold
  - Cells to be considered in cluster energy sums

Use of negative signal cells
- Thresholds are considered for the absolute (unsigned) signal magnitude
  - Large negative signals can seed and grow clusters

Parameters for each stage optimized with testbeam data
- Experimental single pion shower shapes guide cluster algorithm development
  - Clean tuning reference!

Primary threshold
\[ \frac{E_{\text{cell}}}{\sigma_{\text{cell}}} > S, \text{ default } S = 4 \]

Secondary threshold
\[ \frac{E_{\text{cell}}}{\sigma_{\text{cell}}} > N, \text{ default } N = 2 \]

Collecting
\[ \frac{E_{\text{cell}}}{\sigma_{\text{cell}}} > P, \text{ default } P = 2 \]

(note \( S \geq N \geq P \))

Famous “4/2/0” clustering in ATLAS
ATLAS Topological Cell Clustering

- Cluster seeding
  - Defined by signal significance above primary threshold
    - Cells above this threshold can seed cluster
- Cluster growth control
  - Defined by signal significance above secondary threshold
    - Cells neighbouring seeds with significance above this threshold drive cluster growths
- Cluster signal
  - Defined by cells with significance above basic threshold
    - Cells to be considered in cluster energy sums
- Use of negative signal cells
  - Thresholds are considered for the absolute (unsigned) signal magnitude
    - Large negative signals can seed and grow clusters
- Parameters for each stage optimized with testbeam data
  - Experimental single pion shower shapes guide cluster algorithm development
    - Clean tuning reference!
1. Find cell with most significant seed over primary threshold $S$

2. Collect all cells with significance above basic threshold $P$
   - Consider neighbours in three dimensions
     - Defined by (partly) shared area, (partly) shared edge, or shared corner point
     - E.g., 26 neighbours for perfectly cubed volumes of equal size
   - Neighbours can be in other calorimeter regions or even other calorimeter sub-systems
     - Granularity change to be considered in neighbouring definition

3. For all cells neighbouring seeds with signal significance above secondary threshold $N$, collect neighbours of neighbours if their signal significance is above $P$
   - Same rules as for collection around primary seed

4. Continue until cluster does not grow anymore
   - Automatically generate “guard ring” of small signal cells at cluster margin
     - In three dimensions, of course

5. Take next not yet used seed cell and collect next cluster
Clustering Example
Clustering Example

![Graph showing clustering example with signal over noise (SN) levels]
Clustering Example
Clustering Example
Clustering Example
Clustering Example

![Clustering Example Diagram](image-url)
Clustering Example
Clustering Example
Clustering Example
Clustering Example
Clustering Example

[Bar chart showing signal over noise for different clusters.]

- Cluster 1: High signal, low noise
- Cluster 2: Moderate signal, moderate noise
- Cluster 3: Low signal, high noise

Clusters are represented in different colors for easy differentiation.
Clustering Example
Clustering Example
Clustering Example
Clustering Example
Clustering Example
Clustering Example
Clustering Example

[Diagram showing a bar chart with signal over noise (SN ratio) on the y-axis and some data points labeled with numbers 4, 3, 2, 1, and -1 to -5 on the x-axis.]
Clustering Example
Clustering Example
Clustering Example
Large topologically connected regions in calorimeter can lead to large cell clusters

- Lost particle flow structure can introduce problems for jets
  - Infrared safety, in particular
- Need to refine the clustering algorithm
  - Try to match single particle shower shapes better

Splitting the clusters

- Examine spatial cluster signal structure – find local signal maxima
  - “hill and valley” structural analysis in three dimensions
- Split cluster between two maxima
  - In three dimensions, of course!
- Share energy of cells in signal valleys
  - Needs sharing rules – introduces “geometrically” weighted cell energy contribution to cluster signal
- Introduces new tunable parameter
  - Local signal maximum threshold is defined in units of energy, not significance!
Cluster Splitting
Cluster Splitting

E (GeV)

splitting threshold
Cluster Splitting
Splitting & Cell Energy Sharing

• Splitting technique
  • Guided by finest calorimeter granularity
    • Typically in electromagnetic calorimeter
  • Allows to split larger cell signals without signal valley
    • Typically in hadronic calorimeters
Rule for energy sharing:

\[ w_1 = \frac{E_1}{E_1 + rE_2} \]

\[ w_2 = 1 - w_1 \]

\[ r = e^{d_i - d_2} \]

\((d_i)\) is the distance of the cell from the centroid of cluster \(i\)

Each cell can only appear in up to two clusters
Clusters have shapes
  - Geometrical moments and sizes
    - Lateral and longitudinal
  - Tilt of principal axis
    - With respect to direction extrapolation from primary vertex (magnetic field!)
  - Density and compactness measures
    - Cluster energy distribution in cells
  - Energy sharing between calorimeter segments and modules
    - Shower structures
  - Useful for cluster calibration
    - Exploit shape sensitivity to shower character
      - Hadronic versus electromagnetic
Modeled effect of topological clustering on the cell signal significance spectrum, for purposes of illustration here with only the primary (seed) threshold, no secondary (growth) threshold.
Cell Signal Significance Spectrum

no clusters, cells with signal & noise
Cell Signal Significance Spectrum

- No clusters, cells with signal & noise
- No clusters, cells with only noise
Cell Signal Significance Spectrum

- No clusters, cells with signal & noise
- Clustered cells, signal & noise
- No clusters, cells with only noise
- Clustered noise cells

Legend:
- PDF (noise)
- PDF (noise, cluster)
- PDF (total)
- PDF (total, cluster)
- Prob (Signal)
- Prob (signal, cluster)
Cell Signal Significance Spectrum

Probability for "true" signal for all cells

Probability for "true" signal for clustered cells
probability for “true” signal for all cells
probability for “true” signal for clustered cells

Note change of shape of probability density function due to correlations introduced by showering – clustered small signal cells have more likely some true signal because they are in the neighborhood of a cell with significant signal, while cells with the same signal from noise only are more often suppressed!
Significant boost of likelihood that small signals are generated by particles (rather than noise) in clustered cells!
Backup

Forward jets
Some old plots for ATLAS

- The following plots are very old
  - I do not even remember the Pythia version
  - Detector simulation is not official – old Geant3/Gcalor with my own (pretty well described) geometry

- This is no indication of the real detector performance!
  - Just somewhat educational, I thought...
Forward tag jets in Higgs production have average $E_t$ of 30-50 GeV only (decreasing with increasing $\eta$)
ATLAS forward jets and pile-up (really old stuff – 1995 MC)

These plots are strongly affected by the choice of truth reference! Only look at the shape, not the resolution numbers!
pile-up noise is not completely symmetric -> adds signal to jet in FCal!

Jet $p_T$ response versus $R$
Backup
Jet shapes in MB
Core energy fraction vs JVF (Particle level, Pythia8 C4 MB, $\langle \mu \rangle = 30$)

Very preliminary!
Core energy fraction vs Nvtx (Particle level, Pythia8 C4 MB, $<\mu>=30$)

Very preliminary!