Jet Cleansing: Pileup Removal at High Luminosity

David Krohn, **Matthew Low**, Matthew Schwartz, and Lian-Tao Wang

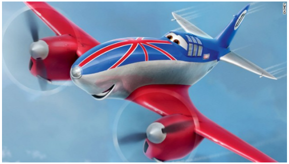
arxiv:1309.4777

Introduction

- Pileup contaminates jets, modifying kinematic quantities and jet shapes.
- Cleansing attempts to return jet to its pre-pileup state.
- Exploits information from charged tracks.
- Infers local information about neutral particles.

Pileup $\Rightarrow$

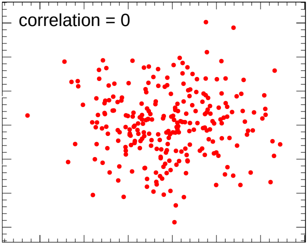
$\Leftarrow$ Cleansing

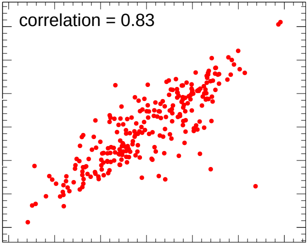


Figures of Merit

- Examples: p_T residual offset, p_T residual fluctuations, mean of observable, etc.
- Want to be sensitive to the structure of the jet.
- We chose linear correlation coefficient ρ_{12}.

$$\rho_{12} = \frac{E[(x_1 - \mu_1)(x_2 - \mu_2)]}{\sigma_1 \sigma_2} \qquad -1 \leq \rho_{12} \leq 1$$





Grooming and Subtraction

Grooming returns a subset of the jet constituents.

- Filtering (+ mass drop) [Butterworth, Davison, Rubin, Salam; arxiv:0802.2470]
- Pruning [Ellis, Vermilion, Walsh; arxiv:0903.5081, arxiv:0912.0033]
- Trimming [Krohn, Thaler, Wang; arxiv:0912.1342]

Subtraction corrects and returns the jet four-vector or shape value.

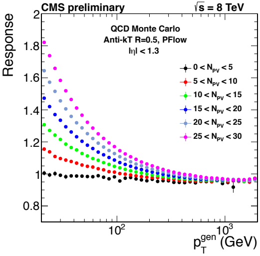
- Area subtraction [Cacciari, Salam; arxiv:0707.1378]

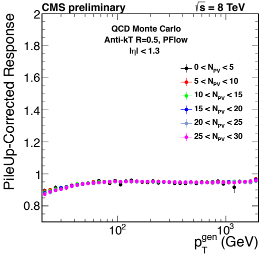
$$p^\mu_{\text{corrected}} = p^\mu - \rho A^\mu \qquad \rho = \text{median}\left[\left\{\frac{p_{T,j}}{A_j}\right\}\right]$$

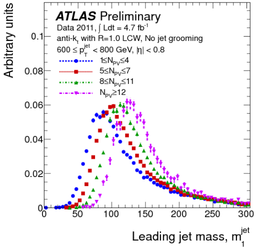
- Shape subtraction [Soyez, Salam, Kim, Dutta, Cacciari; arxiv:1211.2811]

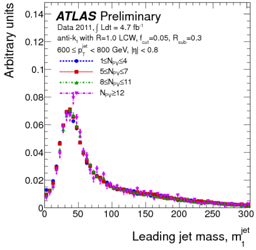
$$V_{\text{corrected}} = V - \rho V^{[1]} + \frac{1}{2}\rho^2 V^{[2]} + \dots$$

Grooming and Subtraction





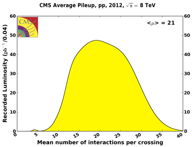


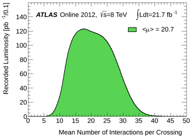


Both grooming and subtraction effectively remove pileup dependence.

LHC Pileup Conditions

Current pileup is $\mu = \langle 21 \rangle$, future pileup may be up to 140.



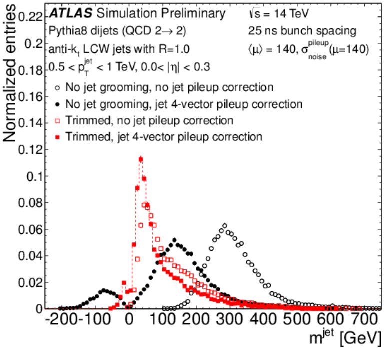


Expected pileup values at the HL-LHC

ATLAS and CMS Collaborations

Abstract

In this brief note the expected mean number of interactions and the length of the luminous region for the HL-LHC is discussed for an instantaneous luminosity of $5 \times 10^{34}/\mathrm{cm}^2/\mathrm{s}$. Its dependence on various parameters is discussed and an uncertainty is estimated. It is found that the mean pileup is about 128, and up to 16% of all bunches will have a pileup of > 140. The length of the luminous region is anticipated to be about 5 cm.



Grooming or subtraction alone doesn't restore distribution.

Jet Cleansing

- Cluster jet constituents into subjets
- Calculate rescaling $\lambda_{\text{sub}}(\text{cells}, \text{tracks}; \text{particles})$ for each subjet
- Rescale constituents in subjet

$$p_i^\mu \to \lambda_{\text{sub}} p_i^\mu \qquad\qquad i \in \text{subjet}$$

- Reassemble jet

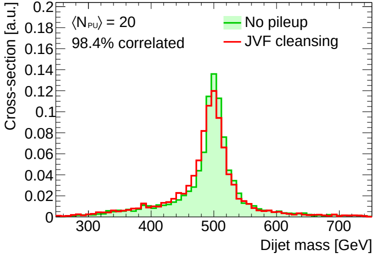
(A) JVF Cleansing

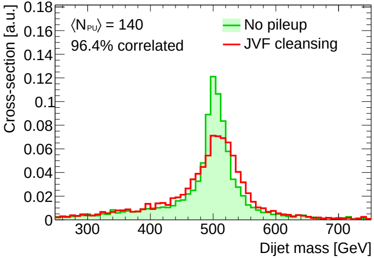
- Estimate scaling by ratio

$$\lambda_{\text{sub}} = \frac{\text{charged } p_T \text{ from LV}}{\text{all charged } p_T}$$

- Using track or particle quantities

$$\lambda_{\text{sub}} = \frac{p_T^{\text{C,LV}}}{p_T^{\text{C,LV}} + p_T^{\text{C,PU}}}$$





* ATLAS has looked at some form of this in the past.

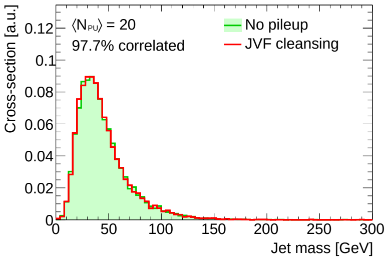
(A) JVF Cleansing

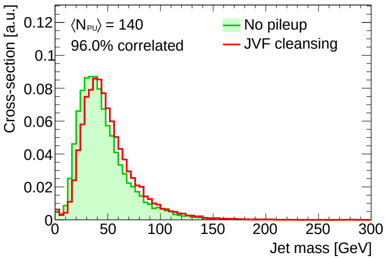
- Estimate scaling by ratio

$$\lambda_{\text{sub}} = \frac{\text{charged } p_T \text{ from LV}}{\text{all charged } p_T}$$

- Using measured quantities

$$\lambda_{\text{sub}} = \frac{p_T^{C,LV}}{p_T^{C,LV} + p_T^{C,PU}}$$





* ATLAS has looked at some form of this in the past.

Parameterization

Useful parameterization of total subjet p_T.

$$p_T = \frac{p_T^{\mathrm{C,PU}}}{\gamma_0} + \frac{p_T^{\mathrm{C,LV}}}{\gamma_1}$$

- p_T = measured calorimeter p_T
- $p_T^{\mathrm{C,PU}}$ = measured track p_T from pileup.
- $p_T^{\mathrm{C,LV}}$ = measured track p_T from the leading vertex.
- γ = charge-to-all p_T ratio.

Can utilize these quantities to estimate neutral components.

(B) Linear Cleansing

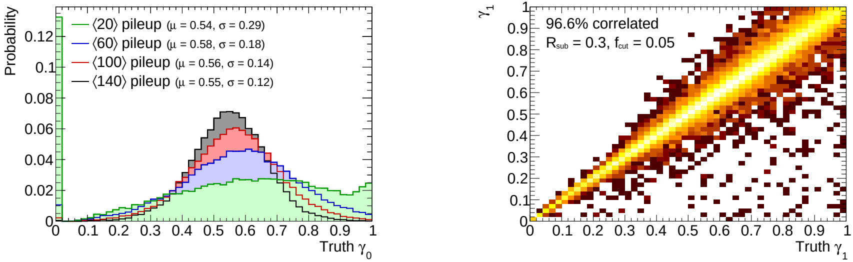
- Estimate scaling by ratio

$$\lambda_{\text{sub}} = \frac{p_T \text{ from LV}}{p_T \text{ from LV} + p_T \text{ from PU}} = \frac{p_T^{\text{C,LV}}/\gamma_1}{p_T^{\text{C,LV}}/\gamma_1 + p_T^{\text{C,PU}}/\gamma_0}$$

- Reduces to

$$\lambda_{\text{sub}} = 1 - \frac{p_T^{\text{C,PU}}}{\gamma_0 p_T}$$

- Estimate charged-to-all ratio in pileup $\gamma_0 =$ constant.



(B) Linear Cleansing

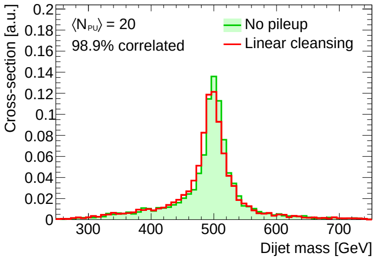
- Estimate scaling by ratio

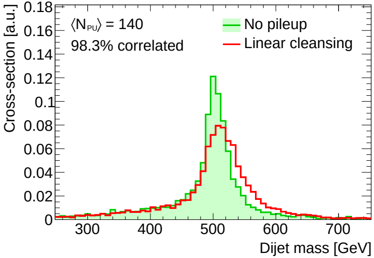
$$\lambda_{\text{sub}} = \frac{p_T \text{ from LV}}{p_T \text{ from LV} + p_T \text{ from PU}} = \frac{p_T^{\text{C,LV}}/\gamma_1}{p_T^{\text{C,LV}}/\gamma_1 + p_T^{\text{C,PU}}/\gamma_0}$$

- Reduces to

$$\lambda_{\text{sub}} = 1 - \frac{p_T^{\text{C,PU}}}{\gamma_0 p_T}$$

- Estimate charged-to-all ratio in pileup $\gamma_0 = \text{constant}$.





(B) Linear Cleansing

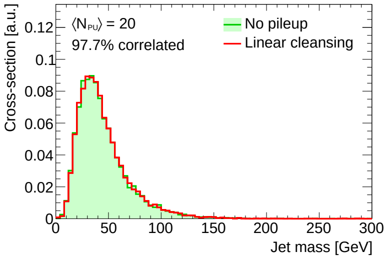
- Estimate scaling by ratio

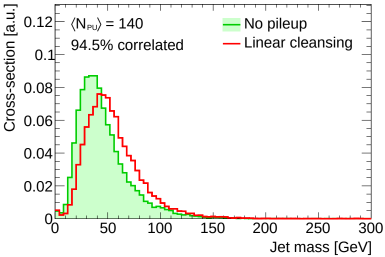
$$\lambda_{\text{sub}} = \frac{p_T \text{ from LV}}{p_T \text{ from LV} + p_T \text{ from PU}} = \frac{p_T^{\text{C,LV}}/\gamma_1}{p_T^{\text{C,LV}}/\gamma_1 + p_T^{\text{C,PU}}/\gamma_0}$$

- Reduces to

$$\lambda_{\text{sub}} = 1 - \frac{p_T^{\text{C,PU}}}{\gamma_0 p_T}$$

- Estimate charged-to-all ratio in pileup $\gamma_0 = \text{constant}$.





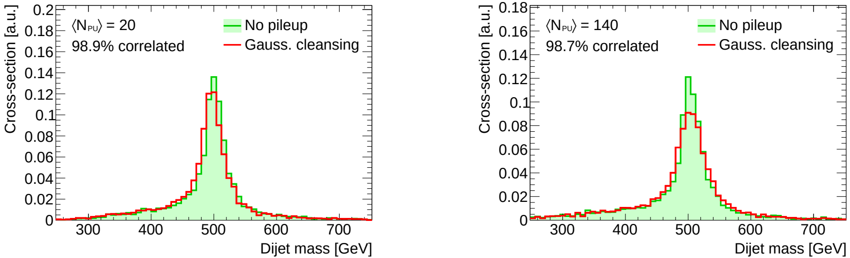
(C) Gaussian Cleansing

- Estimate scaling by ratio

$$\lambda_{\text{sub}} = 1 - \frac{p_T^{\text{C,PU}}}{\gamma_0 p_T}$$

- Find γ_0,γ_1 by maximizing a Gaussian.

$$P(\gamma_0, \gamma_1) \propto \exp\left[-\frac{1}{2}\sum_{i=0,1}\left(\frac{\gamma_i - \overline{\gamma_i}}{\sigma_i}\right)^2\right]$$



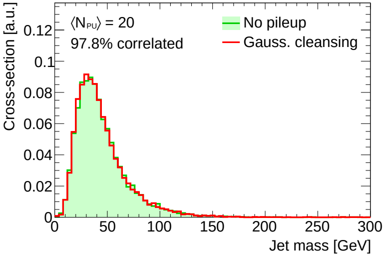
(C) Gaussian Cleansing

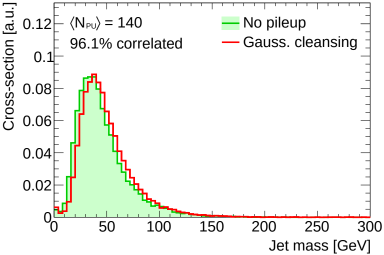
- Estimate scaling by ratio

$$\lambda_{\text{sub}} = 1 - \frac{p_T^{\text{C,PU}}}{\gamma_0 p_T}$$

- Find γ_0,γ_1 by minimizing a Gaussian.

$$P(\gamma_0, \gamma_1) \propto \exp\left[-\frac{1}{2}\sum_{i=0,1}\left(\frac{\gamma_i - \overline{\gamma_i}}{\sigma_i}\right)^2\right]$$





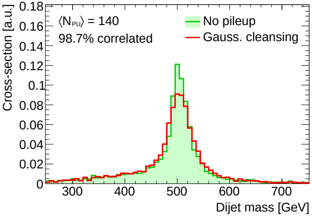
Cleansing + Grooming

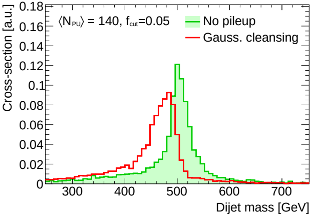
Cleansing interfaces nicely with trimming.

- Cluster jet constituents into subjets
- Calculate rescaling $\lambda_{\text{sub}}(\text{cells}, \text{tracks}; \text{particles})$
- Rescale constituents in subjet

$$p_i^\mu \to \lambda_{\text{sub}} p_i^\mu \qquad\qquad i \in \text{subjet}$$

- Discard subjets with $p_{T,\text{sub}} < f_{\text{cut}} p_{T,\text{jet}}$
- Reassemble jet





Can improve $S/\sqrt{B}$.

Future Directions

- Qcleansing? – Choosing multiple λ_{sub}'s.
- Subjets without subjets? – Alternative to subjets (see Daniele's talk).
- Event cleansing? – Apply cleansing to whole event before jet clustering.

Summary

- Cleansing removes pileup from jets and returns the jet with constituents. Easy to calculate jet shapes.
- Utilizes information from charged tracks to infer local neutral composition.
- Can use charged track information without needing particle flow.
- Interfaces smoothly with existing charged hadron subtraction techniques.

* Plugin available on `FastJet Contrib`

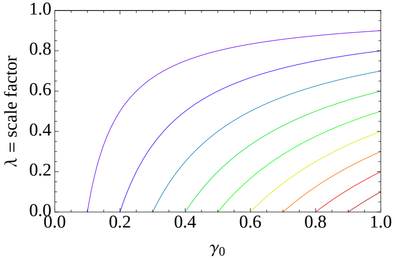
Backup Slides

Simulation Details

- Resonance decaying to light quarks generated with MadGraph5 v1.5.8
- Pythia v8.176 tune 4C, used for pileup and showering.
- Amount of pileup drawn from Poisson.
- FSR, ISR, UE included.
- Charged particles with $p_T < 0.5$ GeV discarded.
- Remaining particles grouped in $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ calo cells.
- Only cells with $E > 1$ GeV are kept.
- Jets with $p_T > 150$ GeV and $|\eta| < 2.5$ after pileup are kept.

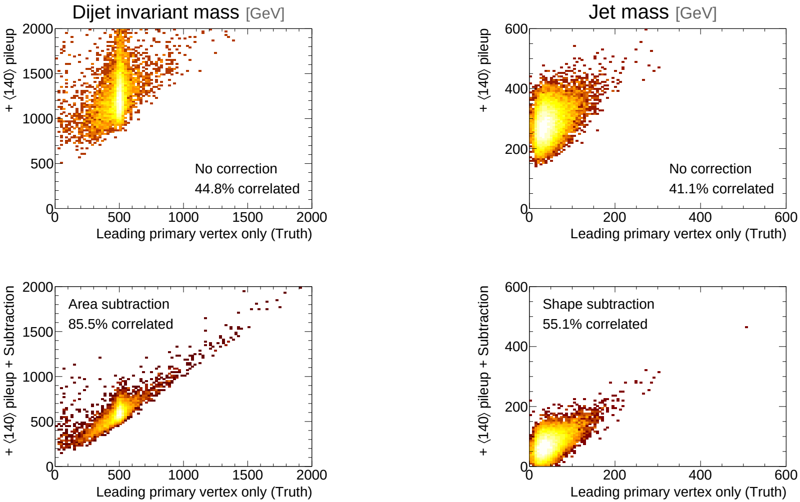
Parameterization

$$p_T = \frac{p_T^{\text{C,PU}}}{\gamma_0} + \frac{p_T^{\text{C,LV}}}{\gamma_1}$$

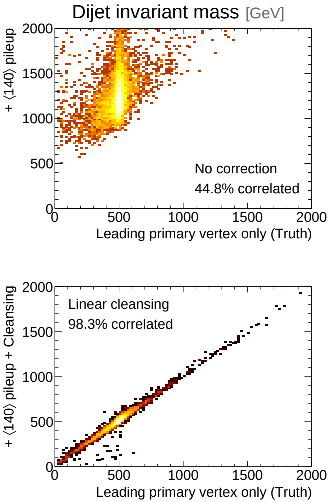


- Colors correspond to different $p_T^{\text{C,PU}}/p_T$ fractions.
- One DOF in parameterization.
- JVF and linear cleansing select choose a γ_0 value.
- Gaussian cleansing puts a Gaussian potential on the line.

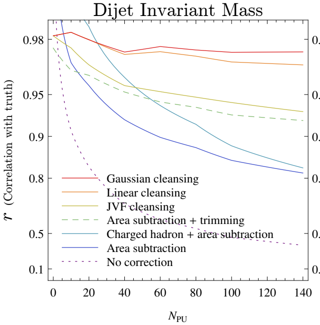
Correlations – Subtraction

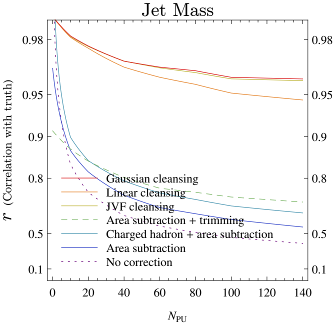


Correlations – Cleansing



Comparisons





Comparisons