Tagging Quark/Gluon Initiated Jets at ATLAS

Boston Jets 2014

Maximilian Swiatlowski

21 January, 2014
History and Motivation

- Quark-initiated and gluon initiated jets have long been known to have different properties
  - Well measured at PETRA, SLAC, LEP, others
- Two papers from Schwartz and Gallicchio in 2011, along with previous efforts in ATLAS, led to a push for creating and commissioning a quark-gluon tagger at ATLAS
  - Theory paper (here) investigated the best variables to use to train a tagger, in parallel to our own efforts
- Many potential applications in searches for new physics and standard model measurements
  - Separate hadronically decaying bosons from gluon dominated backgrounds (diboson searches, Higgs, etc.), improve discrimination in dijet searches, monojet characterization, many more
Today’s Talk

• Today, we are showing for the first time the full results of the 2011 ATLAS q/g tagger
• Results not yet publicly available: will be published in a paper (very) soon

1. Constructing of a Tagger
   Variable Selection
   Purified Samples
   Defining the Likelihood

2. Tagging Performance
   Systematic Uncertainties
   Overview of Performance
   Angularities
Constructing a Tagger
Variable Selection

- Important to choose pileup robust variables: **use only tracking**
- Need strong performance across wide range of $p_T$: $n_{trk}$ has best performance at highest, track width better at low
- EEC variables have good separation as well– but have systematic issues we will describe later
- We use a **likelihood** combining $n_{trk}$ and track width
Template Methods

- Significant data/MC disagreement for the input variables required the use of a data-driven template technique

\[\text{quarks} \quad 60\% \quad + \quad 40\% \quad = \quad \text{gluons} \quad \text{and} \quad \text{dijet}\]

- Take percentages from MC, measure $\gamma + \text{jet}$ and dijet in data: solve for quark and gluon distributions in data
- More information on method in [backup]
MC-labeled distributions in $\gamma$+jet and dijets agree very well with templates derived in MC
- Disagreement at low $p_T$ will be discussed at length soon
- Gives us confidence that the algorithm is doing something sensible
Templates with Data

- But data disagrees with Pythia in $n_{trk}$, leading to worse separation than expected
- Track Width has better agreement, though not good at high $p_T$
Purified Samples

- Are the data templates correct? How can we test these derived shapes?
- Define **topological/kinematic regions** where jets are more likely to be quark-initiated or gluon-initiated
  - Trijet sample, with $\zeta = |\eta_3| - |\eta_1| - |\eta_2| < 0$ is gluon-like
  - $\gamma + 2\text{jet}$ sample, with $\xi = \eta_{\text{jet1}} \times \eta_{\gamma} + \Delta R(\text{jet2}, \gamma) < 1$ is quark-like
  - See arXiv:1104.1175 for more details
- These regions have purity of $\sim 90\%$– good regions for **validation of templates**!
  - Not enough statistics to derive 2D templates, but enough to be useful for validation
• Shapes from topologically purified samples *generally agree* with extracted templates to 1 $\sigma$
Pure Shapes: Gluons

- **Similar levels of agreement** with gluon shapes
- **Completely independent** data samples verify our template shapes
• Define \( L = \frac{q}{q + g} \) separately in data and MC

• Immediately can see that while **shapes are similar**, performance is **much worse** in data
• **Significantly reduced** performance in data
  • But enough to still make something useful!

• We will define a tagger at 4 operating points: 0.3, 0.5, 0.7, and 0.9 quark efficiency
Tagging Performance
Systematic Uncertainties

- Many different sources of error considered for the tagger:
  1. PDF variations— affect $q/g$ fractions
  2. $\gamma$ purity— affects data input
  3. Heavy flavor shapes/fraction— affects MC inputs
  4. Madgraph/Pythia fraction differences— affect $q/g$ fractions
  5. Non-closure/ sample dependence— affects data inputs
Systematics Summary

- Here, show **breakdown of systematics** for 50% quark-like o.p.
- **Sample dependence** is by far the largest effect
  - Quarks/gluons from $\gamma$+jet do not look exactly like quarks/gluons from dijets (in both Pythia and Herwig)
  - Need to understand this effect to apply this tagger to other topologies
- More operating points, and details on non-closure, in backup
Overview of Performance

For measuring performance, we will show several different tests together:

- **Red points** indicate performance of data tagger, tested on data templates
  - **Red lines** on those points indicate statistical uncertainties
- **Teal band** indicates systematic uncertainties
- **Blue points** indicate performance of pythia tagger, tested on pythia templates
- **Magenta points** indicate performance of data tagger, tested on **pure data samples**
Gluon Efficiency vs. Quark Efficiency

- Purified samples show slightly worse gluon efficiency than data, but agreement within $1\sigma$
- Data shows worse performance than MC—generally greater than $1\sigma$ disagreement
Left shows 50% quark point, right shows 70%

Results are consistent across $p_T$: purified samples measurement generally agree with data, but MC significantly overestimates performance

Other operating points in backup
Angularities

• New class of variables, called “Energy Correlation Angularities,” described in arXiv:1305.0007

• Interestingly: possible to show that these variables contain maximum discriminatory power between q/g

• Defined with free parameter $\beta$:

$$\text{Ang} = \frac{\sum_i \sum_j p_{T,i} p_{T,j} (\Delta R(i,j))^\beta}{(\sum p_{T,i})^2}$$  \hspace{1cm} (1)

• How does gluon efficiency change with $\beta$, and how large are the systematics?
Angularity Performance

- **NB:** 1 - Gluon Efficiency shown
- **Significant differences** between data and MC performance, and **systematics are larger** than for the likelihood
  - Sample dependence is **very large** for angularities, at least with $\beta < 1$
  - $\beta = 0.2$ is slightly optimal in MC, but difficult to tell trend in data
Conclusions
Summary

- Much effort has gone into studying the properties of quark and gluon initiated jets (see existing conf note, ATLAS-CONF-2012-138)
- Since then, work has focused on deriving a tagger, calibrating it to data, and determining the systematics
- Data/MC disagreements make tagger derivation difficult—use templates from data
- Systematics need to be carefully assessed—large sample dependencies observed
  - Angularities in particular seem sensitive to these effects
- Final paper, with all these results and more, should be out soon!
Thank You For Your Attention!
Backup
Defining Quark/Gluon Initiated Jets

- Need to use a consistent definition across generators for defining a quark/gluon initiated jet
- We use: “a jet is defined by the flavor of the highest energy parton inside the jet”
  - This labelling is studied in Madgraph to determine how often it matches the Matrix Element: 95 – 99% of the time
Extracting Templates

- Goal: to better understand quark/gluon shapes in data, extrapolate data to 100% purity with fractions from MC
- Ideally, solve for $q/g$ on bin-per-bin basis from:

$$h^{\gamma+j} = P_Q q + P^\gamma_j g$$
$$h^{dijet} = P_Q^{dijet} q + P_G^{dijet} g$$

- But, need to account for $b$ and $c$ fractions (for now, taken from MC):

$$h^{\gamma+jet} = P_Q^{\gamma+jet} q + P_G^{\gamma+jet} g + P_B^b + P_C^c$$
$$h^{dijet} = P_Q^{dijet} q + P_G^{dijet} g + P_B^b + P_C^c$$

- Then, compare pure data shapes to pure MC shapes (used for training tagger)
• Results are consistent across $p_T$: purified samples measurement generally agree with data, but MC significantly overestimates performance
Gluon Efficiency vs. Quark Efficiency, with Herwig++

- Herwig++ generally agrees with data better: sometimes even under-predicts performance
• Similar effects as at other operating points: largest here at low efficiency
• Similar effects as at other operating points
Systematics Summary: 90% Operating Point

- Similar effects as at other operating points
Nonclosure: 30% Operating Point

- Breakdown of Pythia/Herwig++ disagreements with their respective templates
Nonclosure: 50% Operating Point

- Breakdown of Pythia/Herwig++ disagreements with their respective templates
Nonclosure: 70% Operating Point

- Breakdown of Pythia/Herwig++ disagreements with their respective templates
Nonclosure: 90% Operating Point

- Breakdown of Pythia/Herwig++ disagreements with their respective templates

ATLAS Simulation Work in Progress
Error Percentage on Quark Efficiency
anti-k, $R=0.4$, $|\eta| < 0.8$
Efficiency Point 0.90
MC11 Simulation, $\sqrt{s} = 7$ TeV

ATLAS Simulation Work in Progress
Error Percentage on Gluon Efficiency
anti-k, $R=0.4$, $|\eta| < 0.8$
Efficiency Point 0.90
MC11 Simulation, $\sqrt{s} = 7$ TeV