Two motivations

The July 5th problem

07-04-12: find the Higgs boson
07-05-12: measure everything about it
mm-dd-yy: find a new particle
mm-(dd+1)-yy: measure everything about it

Appreciating the
The Standard Model

More than just 27 numbers
Qualitatively new phenomena

Jets
Scaling
Underlying event
Rapidity gaps
Small x physics

Requires precision measurements
Requires precision calculations
Jet substructure

Lots of developments in jet substructure over the past 5 years

- Top-tagging
  - Johns Hopkins Tagger
  - Top template tagger
  - HEP top-tagging
- Jet grooming
  - Filtering
  - Trimming
  - Pruning
- Multijet events
- Pull
- ISR tagging
- Quarks vs Gluons
- Qjets
- Shower deconstruction

- Jet Charge
- N-subjettiness
- Jet cores
- W-tagging
- Dipolarity
- Modified mass drop
- Angularities
- N-point energy correlators
- Semi-classical clustering
  - ...

Many methods tested on data.
Impressive agreement with full simulation.
Applications to BSM (e.g. Z’ resonance searches)
What is a jet?

Energetic quarks and gluons produced

Quarks and gluons “shower” to form jets

Jet algorithms: reconstruct parton momenta

As of 2007: jet=parton

A jet is a 4-vector. Just calibrate it (performance group).
Jets are not just 4-momenta!

- Jets have **substructure**
  - Hard subjets
  - Jet shapes
- Jets have **quantum numbers**
  - Flavor (up/down/strange/charm/bottom)
  - Electric charge
  - Color charge (quark or gluon)
  - Spin (?)
- Jets have **superstructure**
  - Color connections between jets
- Jets are **not partons**
- Jets are not collections of hadrons

2007: A jet is a 4-vector. Just calibrate it.
2013: Jets are **sophisticated emergent phenomena** in the standard model

Let's study them for their own sake!
CASE STUDY: ELECTRIC CHARGE
Jet charge

Can the charge of a jet be measured?

- Could distinguish **up-quark** jets from **down-quark** jets
  - Could help distinguish **up squarks** from **down squarks**

- **W prime** vs **Z prime**

- Many many uses for characterizing new physics (if seen)
Distinguishing charge

Measure the $p_T$-weighted jet charge:

$$Q^i_\kappa = \frac{1}{(p_T^{\text{jet}})^\kappa} \sum_{j \in \text{jet}} Q_j (p_T^j)^\kappa$$

Krohn, Lin, MDS, Waalewijn
Distinguishes $W'$ from $Z'$

Log-likelihood distribution for 1 TeV resonance, various $\kappa$

$2\sigma$ distinction with 30 events

$5\sigma$ discovery with 200 events
Can calibrate with hadronic W's from tops

\[ Q^i_{\kappa} = \frac{1}{(p_T^{\text{jet}})^\kappa} \sum_{j \in \text{jet}} Q_j (p_T^j)^\kappa \]
**2013: measured in data by ATLAS!**

**Theory paper**

Distribution of jet charge for Ws from top pairs

- $\kappa = 0.4$
- $\kappa = 0.7$
- $\kappa = 1$
- $\kappa = 0.3$
- $\kappa = 0.6$
- $\kappa = 1$

**ATLAS Conference note**

![Graph showing jet charge distribution](image)

**arXiv: September, 2012**
Published: May 2013,

11 months

August, 2013
Dijets

Frations for each channel (parton level)

Small x: proton mostly gluon

Larger x: quark-gluon dominates

Valence quarks picked up at large x
Dijets: quarks at large x

2D charges (parton level) for different pT

Fractions (parton level)

Jet charge (hadron level)

September 19, 2013
Matthew Schwartz
Evidence of **valence quarks** in PDFs!

Quark charge measured **without leptons**

-- in pure QCD (dijet) events.
Mean at width are calculable

\[ D_q^h(x, \mu) \text{ Fragmentation function} \]
- Probability that parton q fragments to hadron h with energy fraction x
  \[ E_{\text{hadron}} = xE_{\text{parton}} \]
- Nonperturbative objects with perturbative evolution equations

Moments of fragmentation functions

\[ \tilde{D}_q^h(\nu, \mu) = \int_0^1 dx \ x^\nu D_q^h(x, \mu), \]
(prob. that emission is within jet)

\[ Q_{\kappa}^i = \frac{1}{(p_T^{\text{jet}})^\kappa} \sum_{j \in \text{jet}} Q_j (p_T^j)^\kappa \]

\[ \langle Q_{\kappa}^q \rangle = \frac{1}{16\pi^3} \frac{\tilde{J}_{qq}(E, R, \kappa, \mu)}{J_q(E, R, \mu)} \sum_h Q_h \tilde{D}_q^h(\kappa, \mu) \]

(Prob. of getting jet with E and R)
Mean and width evolution are calculable

- Moments of charge distribution calculable from moments of fragmentation functions

- Evolution of these moments tests precision QCD
  1. Verify dijet charge (2->2 cross sections and PDFs)
  2. Observe new form of scaling violation

Krohn, Lin, MDS, Waalewijn
Contamination

Effect of multiple interactions/pileup not bad
- Tracks from primary interaction vertex part of motivation
- Could be extremely useful tool at high luminosity

\[
\begin{align*}
W' & \text{ vs. } Z', \text{ 50 events} \\
\text{FSR only} & \\
\text{FSR+MI+ISR} & \\
\text{FSR+MI+ISR+trim} & \\
\text{Npileup=10} & \\
\text{Npileup=10 + trim} & \\
\end{align*}
\]
Jet Charge Summary

- $p_T$ weighted jet charge remarkably **useful** at LHC
- Uses **only tracks**
  - Insensitive to pileup
  - Can be used at high luminosity
- Most information in **average** and **width**
  
  $$\langle Q^q_{\kappa} \rangle$$

  $$(\Gamma_{\kappa}^i)^2 = \langle Q^i_{\kappa} \rangle^2 - \langle (Q^i_{\kappa})^2 \rangle$$

- Has been **validated** on W jets from top decays
- Has been **tested** on dijets
  - Quark/Gluon/Flavor content measurable (statistically)
  - Unfolded data will show
    - Tests precision QCD
    - Gluons are at small $x$, valence quarks at large $x$
    - First measurement of scaling violation in charge moments
OTHER IDEAS IN JET SUBSTRUCTURE
Qjets: sample multiple interpretations

Volatility $\mathcal{V} = \frac{\Gamma}{\langle m \rangle}$ is a purely Q-observable

Ellis, Hornig, Krohn, Roy, MDS
Jet Sampling: Qjets for ambiguous/overlapping jets

Kahawala, Krohn, MDS
JHEP 1306 (2013) 006

classical anti-$k_T$
Useful for top/W tagging
Useful for H -> bb
Distinguishes overlapping jets

Interesting standard model physics
Jet is not a parton
Jet is not a collection of hadrons!!

Jets are **sophisticated emergent phenomena** in the standard model

**What is the right way** to think about jets?
Pileup removal will be **ESSENTIAL** for precision QCD at high luminosity

**Existing methods**
- Jet area subtraction
- Jet shape subtraction
- Charged hadron subtraction
- Jet vertex fractioning
- Trimming
- Pruning
- Filtering

**Work well at NPU = 20**

**140 pileup interactions**

- **Jet Cleansing**
  - (new method, Krohn, Low MDS, Wang)

**Area/Shape subtraction**
- Linear cleansing 98.3% correlated
- Shape subtraction 55.1% correlated
- Area subtraction 85.5% correlated

**Existing methods**
- Dijet invariant mass [GeV]
- Jet mass [GeV]
- Leading primary vertex only (Truth)
- No correction
- Area subtraction
- Linear cleansing
- Shape subtraction
- Jet mass [GeV]
- Leading primary vertex only (Truth)
- No correction

**FIG. 2. Correlations for a kinematic variable (dijet mass, jet mass) for events with no leading vertex for 140 pileup interactions.**

- Leading primary vertex only (Truth)
- No correction
- Area subtraction 85.5% correlated
- Linear cleansing 98.3% correlated
- Shape subtraction 55.1% correlated
- 94.5% correlated
- 44.8% correlated
- 41.1% correlated

**FIG. 3. Correlations between events before pileup is added and after 140 pileup interactions.**

- Leading primary vertex only (Truth)
- Jet mass [GeV]
- Leading primary vertex only (Truth)
- No correction
- Linear cleansing 98.3% correlated
- Shape subtraction 55.1% correlated
- 94.5% correlated
- 44.8% correlated
- 41.1% correlated

**TABLE 1. Correlations for a kinematic variable (dijet mass, jet mass) for events with no leading vertex for 140 pileup interactions.**

- Leading primary vertex only (Truth)
- No correction
- Area subtraction 85.5% correlated
- Linear cleansing 98.3% correlated
- Shape subtraction 55.1% correlated
- 94.5% correlated
- 44.8% correlated
- 41.1% correlated
# Quark and gluon tagging

<table>
<thead>
<tr>
<th>Gluon Efficiency % at 50% Quark Acceptance</th>
<th>50 GeV</th>
<th>200 GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Particles</td>
<td>Tracks</td>
</tr>
<tr>
<td>2-Point Moment $\beta=1/5$</td>
<td>P8</td>
<td>H++</td>
</tr>
<tr>
<td>1-Subjettiness $\beta=1/2$</td>
<td>8.7*</td>
<td>17.8*</td>
</tr>
<tr>
<td>2-Subjettiness $\beta=1/2$</td>
<td>9.3</td>
<td>18.5</td>
</tr>
<tr>
<td>3-Subjettiness $\beta=1$</td>
<td>9.2</td>
<td>18.6</td>
</tr>
<tr>
<td>Radial Moment $\beta=1$ (Girth)</td>
<td>9.1</td>
<td>19.3</td>
</tr>
<tr>
<td>Angularity $a=+1$</td>
<td>10.3</td>
<td>20.5</td>
</tr>
<tr>
<td>Det of Covariance Matrix</td>
<td>10.3</td>
<td>20.0</td>
</tr>
<tr>
<td>Track Spread: $\sqrt{&lt;p_T^2&gt;/p_T^jet}$</td>
<td>11.2</td>
<td>21.2</td>
</tr>
<tr>
<td>Track Count</td>
<td>17.7</td>
<td>26.4</td>
</tr>
<tr>
<td>Decluster with $k_T, \Delta R$</td>
<td>15.8</td>
<td>24.5</td>
</tr>
<tr>
<td>Jet $m/p_T$ for $R=0.3$ subjet</td>
<td>13.1</td>
<td>25.9</td>
</tr>
<tr>
<td>Planar Flow</td>
<td>28.7</td>
<td>34.4</td>
</tr>
<tr>
<td>Pull Magnitude</td>
<td>37.0</td>
<td>39.0</td>
</tr>
</tbody>
</table>

| Track Count & Girth                         | 9.9 | 20.1 | 13.4 | 23.2 | 7.1 | 17.3 | 7.7* | 18.7 |
| R=0.3 $m/p_T$ & R=0.7 2-Point $\beta=1/5$   | 7.9* | 17.7 | 12.2* | 22.1 | 5.7 | 14.4* | 8.5 | 17.9 |
| 1-Subj $\beta=1/2$ & R=0.7 2-Point $\beta=1/5$ | 8.5 | 17.3* | 12.9 | 22.1 | 6.0 | 14.6 | 8.6 | 17.7* |
| Girth & R=0.7 2-Point $\beta=1$             | 12.6 | 21.9 | 12.6 | 21.9* | 9.2 | 18.0 | 9.2 | 18.0 |
| 1-Subj $\beta=1/2$ & 3-Subj $\beta=1$      | 8.9 | 18.0 | 14.0 | 23.2 | 5.6* | 15.0 | 8.4 | 18.4 |

| Best Group of 3                              | 7.5 | 17.0 | 11.0 | 20.9 | 4.7 | 14.0 | 6.9 | 16.6 |
| Best Group of 4                              | 7.1 | 16.7 | 10.6 | 20.5 | 4.5 | 13.7 | 6.2 | 16.3 |
| Best Group of 5                              | 6.9 | 16.4 | 10.4 | 20.0 | 4.3 | 13.3 | 6.1 | 15.9 |
Quark and Gluon tagging

- Hard problem: Two equivalence classes
- Discrimination easier at higher $p_T$
- Using all particles works better than just charged tracks
- 80-90% gluon rejection at 50% quark acceptance is realistic
- Pythia gives bigger Q/G difference than Herwig
Data and pythia do not agree
For charged particle multiplicity

Future of Q vs G needs
more data and better theory

<table>
<thead>
<tr>
<th></th>
<th>Pythia 8</th>
<th>Herwig ++</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quark acceptance</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>Gluon acceptance</td>
<td>17%</td>
<td>26%</td>
<td>25%</td>
</tr>
</tbody>
</table>
measuring **Color** flows in jets

**Signal**

$H \rightarrow b\bar{b}$

**Background**

$q\bar{q} \rightarrow Zb\bar{b}$

$gg \rightarrow Zb\bar{b}$
Pull

• Find jets (e.g. anti-$k_T$)
• Construct pull vector ($\sim$ dipole moment)
on radiation in jet

$$\vec{p} = \sum_i \frac{E_T^i |r_i|}{E_T^{jet}} \vec{r}_i$$
CDF dijet excess

- CDF observed an excess of $Wjj$ events in the $jj$ invariant mass region 120-160 GeV with a cross section about 4 pb. (PRL 106:171801 (2011) [arXiv:1104.0699])
- Natural Low Scale Technicolor interpretation. ELM (PRL 106:251803 (2011) [arXiv:1104.0976])

CDF cuts are: exactly one lepton, $l = e, \mu$, with $p_T > 20 GeV$ and $|\eta| < 1.0$; exactly two jets with $p_T > 30 GeV$ and $|\eta| < 2.4$; $\Delta R(l,j) > 0.52$; $p_T(jj) > 40 GeV$; $E/T > 25 GeV$; $M_T(W) > 30 GeV$; $|\Delta \eta(jj)| < 2.5$; $|\Delta \varphi(E/T,j)| > 0.4$.

- $M(p p \rho \omega \pi T Wjj) = 2.4 pb.$

Measure color connections distinguishes s from t channel production
Must validate on Standard Model first
Validate on tops

b-tag

Measure pull

Clean top tag on leptonic side
D0 ruled out color octet W in top decays

Andy Haas and Yvonne Peters, hep-ex:1101.0648
Conclusions

We want to have tools ready and validated before new physics is discovered
- e.g. jet charge
- color connections
- QvG
- …

The July 5th problem

Appreciating the The Standard Model

2007: A jet is a 4-vector. Just calibrate it.
2013: Jets are sophisticated emergent phenomena in the standard model

Lets study them for their own sake!

Examples:
- jet charge
  - shows evidence for quarks and gluons in proton, at different x
  - Mean charge scale-independent to leading order calculable scaling violation
- Qjets/volatility
  - Re-evaluate what a jet is.
  - Wide open field, theoretically and experimentally
QUESTIONS FROM ATLAS
Scale uncertainties

- Several theorists suggest that factorization and normalization scale uncertainties on $W+b$ predictions should be estimated by changing the scales by factors 4 and $\frac{1}{4}$ (instead of the usual 2 and $\frac{1}{2}$). This has become the standard procedure in comparing prediction with past $W+b$ xsec measurements. Is this still justified?
Why should scale variations predict uncertainties?

\[
\sigma \sim \alpha_s(\mu) c_1 + \alpha_s^2(\mu) (c_1 \beta_0 \ln \frac{\mu}{Q} + c_2) + \cdots
\]

No dependence on \(\mu\) if known exactly

Choose \(\mu = Q\) to minimize large logs

\[
\sigma \sim \alpha_s(Q) c_1 + \alpha_s^2(Q) c_2
\]

Suppose \(c_2\) is not known, but \(c_1\) and \(\beta_0\) are. How do we estimate \(c_2\)?

Varying around \(Q/2 < \mu < 2Q\)

\[
\sigma \sim \alpha_s(Q) c_1 \pm \alpha_s^2(Q) c_1 \beta_0 \ln 2
\]

Gives a number of order 1 that appears in the right place where \(c_2\) would in the cross section

Method works well for inclusive single scale observables
**Scale uncertainties**

Method works for inclusive **single scale observables**
No theory justification for most cross sections at LHC

N-jet production, W+jet production, Higgs+W with jet veto

In these cases, we don’t know what $\mu=Q$ means

1. Guess:

\[
\mu = H_T \\
\mu = \sqrt{p_T^2 + m_W^2} \\
\mu = \max\{m_W, E_{\text{jet}}\}
\]

Differences between parameterizations are larger than the individual variations

![Graph showing scale variations at next-to-leading order. The blue south-east stripes show the scale variation of the NLO calculation (called NNLO in fewz) with $\mu=\mu_f=\mu_r=M_W$, as in the atlas paper. The red northeast stripes show the prediction using $\mu_f=\mu_r=\sqrt{M_W^2+p_T^2}$ and the black vertical stripes have $\mu_f$ and $\mu_r$ set to the scales in Eq. (29). Bands correspond to varying $\mu=\mu_f=\mu_r$ by factors of two from these default scales. The increase in the cross section from NLL to N2LL is mostly due to the one-loop constants in the soft and hard functions, as can be seen from the right panel of Figure 3. We have checked how much of a shift the known two-loop jet and soft functions induce and find that it is below a per cent.]

**Comparison with LHC data**

We are now ready to compare to LHC data. We discuss separately the two processes we study, direct photon and $W$ production. For numerical work we use the NNLO MSTW 2008 PDF set and its associated $\alpha_s(M_Z^2)=0.1171$ [45]. We also use $M_W=80.399$ GeV, $\alpha_e=\sin^2\theta_W=0.2226$, $V_{ud}=0.97425$, $V_{us}=0.22543$, $V_{ub}=0.00354$, $V_{cd}=0.22529$, $V_{cs}=0.97342$ and $V_{cb}=0.04128$. [44]
In some cases, we know the origin of the different scales:

- **Jet scale**: \~ jet mass, \~ \text{out of jet energy}
  - High \( p_T \)
  - W boson

- **Hard scale**: \~ \( p_T \)
- **Soft scale**: \~ \text{out of jet energy}

Individual variation show extrema (natural \( \mu_{\text{hard}}, \mu_{\text{jet}}, \mu_{\text{soft}} \) scales, like \( Q \))

When put together \( \mu_{\text{hard}} = \mu_{\text{jet}} = \mu_{\text{soft}} = \mu \) gives NLO

No natural \( \mu \) at NLO (or NNNNLO). **Cannot set all scales equal.**
W + JET at the LHC

Theory vs. ATLAS data
$W^+ + W^-$ (LHC, 7 TeV, 31 pb$^{-1}$)

- - - PDF uncertainties

$N^3$LL + NLO

NLO

$\sigma - \sigma_{NLO}$

$\sigma_{NLO}$

$p_T$ (GeV)

100 150 200 250 300


Public code (PeTeR) for high-$p_T$ W/Z

COMING SOON
Scale setting

My recommendation:

Compare different parameterizations, including all relevant scales, rather than varying each by 2 or 4