Quark jet rates and quark gluon discrimination in multi-jet final states

Yasuhiro Sakaki
(KAIST)

1807.01421

ICHESP 2018 @Seoul, 4-11/7/2018
• No clear sign of BSM at the LHC
• Need to examine final states more precisely
• Final states are categorized by inclusive variables
  \[\text{N(jets), N(leptons), MET, } H_T, \ldots\]
• Categories containing jets encounter a huge QCD background
• As increasing N(jets), kinematics and MC validation become more complicated
• LHC is jets production machine. We want to examine precisely even such multi-jet final state.
Multi-jet final state and New physics

- **No accurate simulation** for the large jet multiplicity background due to the absence of higher-order, huge number of diagrams...

  ➡ **data-driven analysis**

inclusive variable:

ex) $H_T = \sum p_T, \text{jet}$

$\int L\, dt = 3.0\, \text{fb}^{-1}$

$\sqrt{s} = 13\, \text{TeV}$

$n_{\text{jet}} \geq 3$

$\text{SR}$

$\text{CR}$

$H_T [\text{TeV}]$

... up to $> 8\text{jet}$
Multi-jet final state and New physics

- Accurate simulation for the large jet multiplicity background does not exist due to the absence of higher-order, huge number of diagrams...
Jet substructure techniques

- Quark/Gluon discrimination

- It works well even with small-R (even with $R<0.4$)

- QCD radiation is approximately scale invariant
Large enhancement of S/B for a signal that predicts the number of quark jets which is different from what the QCD background does

Let's see how many quark jets are included in the QCD background.
• $R_{n,m}^i$ : Probability that $i$ emits $n$ jets in which $m$ quark jets are contained

• Generating functional:

$$\Phi_i(p, t) = \sum_{n=1}^{\infty} \sum_{m=0}^{n} u^n v^m R_{n,m}^i(p, t)$$

$$\Phi_i(p, t) = u \nu_i \Delta_i(p, t) + \sum_k \int_{p_0/p}^1 dz \int_{t_0}^t dt' \frac{\Delta_i(p, t)}{\Delta_i(p, t')} \mathcal{P}_{i \rightarrow jk} \Phi_j(p, t') \Phi_k(zp, t')$$
Solutions

\[
\Phi_i(p, t) = \sum_{n=1}^{\infty} \sum_{m=0}^{n} u^n v^m R_{n,m}^i(p, t)
\]

\[
\Phi_q = u^1 v^1 e^{-(1-u) a_q L^2} e^{S_q} \]

\[
\Phi_g = u^1 v^0 e^{-(1-u) a_g L^2} e^{-(1-u v^2) a_{qg} L} e^{S_g}
\]

0-th order

DL

SL(g→qq)

subsequent emissions

(gluon jet)

(quark jet)

(both)

hard process

q/g

subsequent emissions
• evolution equation

\[ \Psi_i(x, t) = \Pi_i(x, t) + \sum_k \int_x^1 \frac{dx'}{x'} \int_{t_0}^t \frac{dt'}{t'} \frac{\Pi_i(p, t)}{\Pi_i(p, t')} \times \frac{f_k(x', t)}{f_i(x, t)} \mathcal{P}_{k\rightarrow ij} \Psi_k(x', t') \Phi_j((x' - x) p_b, t') \]

• solutions

\[ \Psi_q = e^{-(1-u)a_q k_q \lambda} e^{-(1-uv)c_q a_{qq} \lambda} e^{S_g[f_{q/q}]} \quad \cdots \]
\[ \Psi_g = e^{-(1-u)a_g k_g \lambda} e^{-(1-uv)c_g a_{q} \lambda} e^{S_g[f_{g/g}]} \quad \cdots \]

(ASI) (quark jet) (subsequent emissions (both))

(GLU)
# of quark jets

- Increase of gluon jet (double-log), quark jet (single-log)
- QCD multi-jets background is composed of few valence quark jets and many gluon jets
- W/Z/gamma + jets are also available
- It would be useful for MC tuning and development

\[ \sqrt{s} = 2 \text{ TeV} \]
Expected improvement of S/B

\[
\frac{S}{B} \propto \frac{\varepsilon_S}{\varepsilon_B} \sim \left( \frac{\varepsilon_q}{\varepsilon_g} \right)^{N_q^{\text{signal}} - N_q^{\text{QCD-bkg}}}
\]

\[\sqrt{s} = \Lambda_{\text{new}}\]

\[N_q^{\text{signal}}\]

all jets are quark jets

\[
\begin{aligned}
\epsilon_g & = 0.1 \\
\epsilon_g & = 0.2 \\
\epsilon_g & = 0.3 \\
\epsilon_g & = 0.4 \\
\epsilon_q & = 0.5 \\
\epsilon_q & = 0.6 \\
\epsilon_q & = 0.7 \\
\epsilon_q & = 0.8 \\
\epsilon_q & = 0.9 \\
\end{aligned}
\]

\[
\begin{aligned}
N_{\text{jets}} & = 3 \\
N_{\text{jets}} & = 4 \\
N_{\text{jets}} & = 5 \\
N_{\text{jets}} & = 6 \\
N_{\text{jets}} & = 7 \\
N_{\text{jets}} & = 8 \\
N_{\text{jets}} & = 9 \\
N_{\text{jets}} & = 10 \\
\end{aligned}
\]

background: QCD multijet
signal: quark jet dominant
\[
\epsilon_S = 0.4
\]
MC analysis

QCD jets \( \overrightarrow{\text{BDT}} \sim (1, 0, 0, 0, 0, ...), \ (1, 1, 0, 0, 0, ...) \)

signal \( \overrightarrow{\text{BDT}} \sim (1, 1, 1, 1, 1, ...) \)

- BDT distance: \( d = \sqrt{|\overrightarrow{\text{BDT}}|^2} \)

- We can estimate # of background of each bins by data-driven extrapolations

\[
\frac{N(d > d_{\text{cut}})}{N}
\]

\( H_T[\text{TeV}] \)

\( N_{\text{jet}} \geq 6 \)
We can estimate the number of background of each bin by data-driven extrapolations.

\[ N(d < d_{\text{cut}}) \]
\[ N(d > d_{\text{cut}}) \]

- QCD jets: \( \overrightarrow{\text{BDT}} \sim (1, 0, 0, 0, 0, \ldots), (1, 1, 0, 0, 0, \ldots) \)
- Signal: \( \overrightarrow{\text{BDT}} \sim (1, 1, 1, 1, 1, \ldots) \)

- BDT distance: \( d = \sqrt{\overrightarrow{\text{BDT}}^2} \)

\[ \frac{N(d > d_{\text{cut}})}{N} \]
Enhancement of S/B with $d$

$$S/B \sim \frac{\sigma_S}{\sigma_B} \times \left( \frac{\epsilon_S}{\epsilon_B} \right) \times \left( \frac{\epsilon_S}{\epsilon_B} \right) \times \left( \frac{\epsilon_S}{\epsilon_B} \right)$$

selection cut

jet substructure cut ($d$ cut)

$H_T$ cut

**toy-signal:** $gg/q\bar{q} \rightarrow XX$, $X \rightarrow N$-quarks

After imposing $H_T$ cut. Fixed at $\epsilon_S = 0.4$

- **N=2**
  - $n_X = 2$ ($u\bar{u} \rightarrow 4q$)
  - $\epsilon_S = 0.4$

- **N=3**
  - $n_X = 3$ ($u\bar{u} \rightarrow 6q$)
  - $\epsilon_S = 0.4$

- **N=4**
  - $n_X = 4$ ($u\bar{u} \rightarrow 8q$)
  - $\epsilon_S = 0.4$

- **N=5**
  - $n_X = 5$ ($u\bar{u} \rightarrow 10q$)
  - $\epsilon_S = 0.4$

• Large enhancement of S/B
Summary

• Quark/gluon discrimination can be maximally utilized for BSM searches in multi-jet final states.

• Quark and gluon jet fraction in QCD multi-jet background was estimated.

• Introducing a variable for the data-driven analysis in multi-jet final states, we checked the large improvement of S/B using the variable.
Matrix element correction

\[ \langle N_{\text{quark-jets}} \rangle \]

\[ H_T>2\text{TeV}, \sqrt{s} = 14\text{TeV} \]

- **Born**
- + up to 1 parton
- + up to 2 partons

\( N_{\text{jets}} \):

\[ 2, 3, 4, 5, 6, 7, 8 \]
• Jet substructure technique was established well as top/W/Z/H tagging tools (2-, 3-prong structure)

• Quark/Gluon discrimination is also available (1-prong structure)

<table>
<thead>
<tr>
<th>jet substructure</th>
<th>formed by</th>
<th>R (jet radius)</th>
</tr>
</thead>
<tbody>
<tr>
<td>top W Z H</td>
<td>EW</td>
<td>~1.0 (fat jet)</td>
</tr>
<tr>
<td>quark gluon</td>
<td>QCD</td>
<td>0.4</td>
</tr>
</tbody>
</table>

• QCD radiation is approximately scale invariant

• Quark/Gluon discrimination works well even small-R (even for R<0.4)
A whole generating functional for a matrix element is given by a product of FSR and ISR generating functionals.

\[ W = \Psi_1 \Psi_2 \left( \prod_k \Phi_k \right) \]

\[ W(t) = \Psi_1 \Psi_2 \left( \prod_k \Phi_k \right) \]

\[ R_{n,m}(t) = \frac{1}{n! m!} \frac{\delta^n}{\partial u^n} \frac{\delta^m}{\partial v^m} W(t) \bigg|_{u=v=0} \]

A whole generating functional for a matrix element is given by a product of FSR and ISR generating functionals.
Jet rates

- $R_n(t)$: Probability that an event has $n$ jet
- Studied well. Contribute to understanding of QCD
- and used...

![Graph showing jet rates and distributions for OPAL and ATLAS experiments.](image)
How to measure quark jet rates in multi-jet final state

• Make quark/gluon jet templates for a variable from di-jet, Z/γ+jet sample

• Measurable, if the QCD jet substructure is universal (It depends on only $p_T$ and rapidity, not # of jet)

• Many applications are conceivable
MC comparison

- MC (Herwig++)
- Analytic (LL + g→qq + subsequent + ISR)

\[ \sqrt{s} = 1 \text{TeV} \]

- Increment of jet (leading, LL), quark jet (sub-leading, NLL)
- QCD jets background is composed of 1 or 2 valence quark jets and many gluon jets