QCD in BSM & Higgs

Gabriel Facini
University of Chicago

August 26th 2016
Overview

- The Run 2 (usual) lifetime of a search
- Power of the all hadronic search
- The bumps in the road ahead

As per usual with this sort of talk only a sampling of the possible topics are discussed. SUSY not discussed.

Apologies if your favorite topic is not included. Please mention in discussion afterwards!
Run 1: Success!

- ~500 papers submitted by each ATLAS & CMS
- Wide range of precision measurements
- Extensive searches with 1 or 2 interesting features
- A new particle, confirmation of our understanding of mass
Run 1: Higgs-centric!

All Higgs results are consistent with the Standard Model

Run 2

• Need something unexpected to advance our knowledge

• Do things that have never been done...
  • Probing physics at unexplored energies
  • Unprecedented rates of pp collisions

• Using state-of-the-art tools

• Recipe for discovery: Expect the unexpected
  • \textit{NB: New physics can hide in the uncertainties!}
The lifetime of a Run 2 search
(the simplest example)
Run 2: Searches

- At start, profit from increase in $\sqrt{s}$ to probe higher masses than Run 1
Run 2: Searches

- potential for discovery was huge out of the gate and it was in QCD signatures were we started

- 20/fb @ 8 TeV ~ 0.2/fb @ 13 TeV

  short game - get a quick and robust background estimation and get the results out!
Run 2: Searches

• potential for discovery was huge out of the gate and it was in QCD signatures were we started

• Black holes searches were even more exciting

\[
\frac{d\sigma}{dm_{jj}} = \frac{P_0(1-x)^{P_1}}{x^{P_2+P_3\ln(x)}}
\]

largest concern: smoothness
Run 2: Searches

• potential for discovery was huge out of the gate and it was in $QCD$ signatures were we started

• Huge jump in reach from $\sqrt{s}$

• Limits: $5.6 \rightarrow 6.5$ TeV

expected limit @ 5 TeV: $0.057$ pb
Search Lifetime

• potential for discovery was huge out of the gate and it was in $QCD$ signatures were we started

• mass reach grows like $\log(L)$ [1.5 TeV for factor 50]

• Limits: 6.5–>8.1 TeV

ATLAS

$\sqrt{s}=13$ TeV, 3.6 fb$^{-1}$

- Data
- Background fit
- BumpHunter interval

QBH (BM), $m_h=5.0$ TeV

QBH (BM), $m_h=6.5$ TeV

$\rho$-value = 0.67

Fit Range: 1.1 - 7.1 TeV

$|y^*| < 0.6$

PLB 754 (2016) 302-322

ATLAS

$\sqrt{s}=13$ TeV, 3.6 fb$^{-1}$

$|y^*| < 0.6$

$\sqrt{3.6/0.08}$

expected limit @ 5 TeV: 0.006 pb
potential for discovery was huge out of the gate and it was in QCD signatures were we started

mass reach grows like \( \log(L) \) \[1.5 \text{ TeV for factor 50}\]

- Limits: 8.1–>8.7 TeV

expected limit @ 5 TeV: 0.003 pb

cross section limit shrinks like \( \sqrt{L} \)
Search Lifetime

\[
\text{cross section limit [pb]} \quad \text{log}(L)
\]

- stat limited
- syst limited

this is where precision in background knowledge makes a difference

\[ \sqrt{L} \]

\[ M \text{ [TeV]} \]

psychology - how low can you go??
Besides Bumps

- The simplicity of the bump hunt leaves a weakness
  - insensitivity to non-resonant behavior

- Complementary search needed!

**Underlying Idea:**
- QCD dijet production: mainly t-channel (small scattering angles)
- BSM: more isotropic dijet angular distribution

**Method:**
Angular distributions in dijet system rest frame.

**Comparison:**
- More QCD-like
- More BSM-like

\[ c = e^{y_1} = e^{y_2} = \frac{1}{1 - \cos \Theta^*} \]

**Theoretical and Systematic Uncertainties**

- Contact Interaction (+0.150)
- \( m_{jj} < 800 \text{ GeV} \) (+0.150)
- \( 800 < m_{jj} < 1200 \text{ GeV} \) (+0.075)
- \( 1200 < m_{jj} < 1600 \text{ GeV} \) (+0.025)

**8 TeV Results:**
Insufficient but encouraging!

**Work closely with mass resonance search:**
- Use the same objects, but wider \( y^\ast \) window
- Triggers, performance, analysis framework, MC requests...
- Same blinding strategy in \( |y^\ast| < 0.6 \), our signal region

**Differences:**
- Prediction for SM shape – not a fit
- Sensitive to non-resonant phenomena (benchmark: contact interactions) – complementary discovery potential!

**Need all possible handles on data/MC agreement for first data**

**Use**
- \( |y^\ast| > 0.6 \) region without blinding as control region

We aim for a fast 50 ns analysis!
Besides Bumps

- The simplicity of the bump hunt leaves a weakness
- insensitivity to non-resonant behavior
- Complementary search needed!

Bkgd: Pythia with NLO QCD+EW corrections

\[ \chi = e^{2y^*} = e^{\Delta y} \sim 1 + \cos \Theta^*/1 - \cos \Theta^* \]
Besides Bumps

• The simplicity of the bump hunt leaves a weakness

• insensitivity to non-resonant behavior

• Complementary search needed!

**Bkgd: Pythia with NLO QCD+EW corrections**

\[ \Lambda = \text{scale of new physics} \]

\[ \frac{\sigma}{\sigma_{\text{th}}} \]

\[ \Lambda [\text{TeV}] \]

this is where precision in background knowledge makes a difference
QCD Community

• The foundation of a long and health physics program:

1. Great operations: accelerator and detector

2. Robust object reconstruction / identification

3. Precise *measurements & calculations* are the shoulders of the giants we stand on
QCD Precision

• Three cases where precise knowledge of QCD is/will be the **important** in the long run [personal view]

1. Substructure: How far can the data driven corrections go?

2. MET searches: PDFs and transfer factors

3. H(bb)
1. **Substructure**: How far can the data driven corrections go?
QCD in Searches

• Most analyses: designed to **suppress QCD**:
  
  • Use Leptons, Missing ET, photons
  
  • QCD contribution described by cocktail of MC predictions plus data driven approach

• But…
QCD in Searches

• Most analyses: designed to suppress QCD:
  • Use Leptons, Missing ET, photons
  • QCD contribution described by cocktail of MC predictions plus data driven approach

• But…All Hadronic Searches are the leading edge of search programs i.e....(small sampling)
  • (HVT->)VV->qqqqq
  • (HVT->)VH->qqbb
  • (G->)HH->bbbb

• … very dependent on substructure techniques
Boosted Vs

• Heavy things decaying to W or Z’s

leptons triggers
to suppress
QCD backgrounds

embrace QCD
suppress bkgd rate
with substructure

heavily rely on MC with
(complicated) profile likelihoods

model QCD with fit
validate with MC
Boosted Vs

- Heavy things decaying to W or Z’s leptons triggers

heavily rely on MC with (complicated) profile likelihoods

profile likelihoods facilitate searching in the presence of ignorance of backgrounds
Boosted Vs

• Heavy things decaying to W or Z’s then to jets

• Run 1 excitement

model QCD with fit:

\[
\frac{d\sigma}{dm_{jj}} = \frac{P_0(1 - x)^{P_1}}{x^{P_2 + P_3 \ln(x)}}
\]
Boosted Vs

- Heavy things decaying to W or Z’s then to jets

- Run 1 excitement did not return for Run 2

\[ \frac{d\sigma}{dm_{jj}} = \frac{P_0(1 - x)^{P_1}}{x^{P_2 + P_3 \ln(x)}} \]

model QCD with fit:

Recall:

largest concern: smoothness
Boosted Vs

2-lepton, 1-lepton, 0-lepton (MET, JJ)

low mass
good mass resolution
good to trigger

Run 2 data does not support Run I excess

Run 2 data does not support Run I excess

ATLAS Preliminary

$\sqrt{s} = 13$ TeV, 13.2-15.5 fb$^{-1}$

95% C.L. exclusion limits

HVT model $A g_v=1$

Observed

Expected

qqqq
lvqq
llqq
vvqq

low mass

High signal yield
Controlling Substructure

• Understanding jet substructure is critical for physics above the EW symmetry breaking scale

• How do we control it? Data Driven!
Controlling Substructure

• Understanding jet substructure is critical for physics above the EW symmetry breaking scale

• How do we control it? Data Driven!

N_{trk}: A powerful variable for experimentalists but MC-dependent & non-perturbative

More theoretically robust variables could help?
Controlling Substructure

- Understanding jet substructure is critical for physics above the EW symmetry breaking scale
- How do we control it? Data Driven!

![Graphs showing jet substructure significance with data-driven control methods for ATLAS and CMS.]
Boosted V/Hs

ATLAS Preliminary

95% C.L. exclusion limits

\( \sqrt{s} = 13 \text{ TeV}, 13.2-15.5 \text{ fb}^{-1} \)

HVT model A \( g_y = 1 \)

- qqqq
- lvqq
- llqq
- vvqq
- qqbb

\( \sigma (pp \rightarrow \text{HVT W'}) \)[pb]
Boosted Hs

- The Higgs discovery gave us a new search tool
  - similar to diboson searches, maximal signal yield with hadronic decays (b’s!!!)
- Added complexity of flavor tagging

![Diagram](image)

Soft Drop: \[
\frac{\min[pT_i, pT_j]}{pT_i + pT_j} > z_{\text{cut}} \left( \frac{R_{ij}}{R} \right)^\beta
\]
Boosted Hs

\[ \sigma(gg \rightarrow H) \times BR(H \rightarrow hh) \text{ [pb]} \]

\[ ATLAS \ \sqrt{s} = 8 \text{ TeV}, 20.3 \text{ fb}^{-1} \]

- Observed
- Expected
- WW\gamma\gamma \text{ exp}
- bb\gamma\gamma \text{ exp}
- ± 1σ expected
- ± 2σ expected
- bbbb \text{ exp}

13 TeV summary plot not available … yet
Boosted Hs

CMS employs 2 different data driven background methods profiting from comparison

Here: Expanded ABCD method (alphabet) using untagged events to model tagged events
Is This Enough?

- Do we understand substructure enough to:
  - Check our cuts do not sculpt the background
  - Check our background models are accurate
  - Perform non-resonant searches based on MC driven backgrounds?

Also searches for X>YY→jjjj exist! (covers RPV SUSY)
Is This Enough?

• Do we understand substructure enough to:
  
  • Check our cuts do not sculpt the background
  
  • Check our background models are accurate
    
    • *Perform non-resonant searches based on MC driven backgrounds?*

2012: Open question in preparation for Run 2
Calculations

- Do we understand substructure enough to:
  - Check our cuts do not sculpt the background
  - Check our background models are accurate
  - *Perform non-resonant searches based on MC driven backgrounds?*

2012: Open question in preparation for Run 2
2013: There is hope!

[Dasgupta, Fregoso, Marzani, Salam '13]
[Larkoski, Marzani, Soyez, Thaler '14]
Calculations

Do we understand substructure enough to:

- Check our cuts do not sculpt the background
- Check our background models are accurate
- Perform non-resonant searches based on MC driven backgrounds?

2012: Open question in preparation for Run 2
2013: There is hope!
2016: Precision

[See P. F. Monni, T. Becher talks]
Measurements!? 

- Do we understand substructure enough to:
  - Check our cuts do not sculpt the background
  - Check our background models are smooth
  - Perform non-resonant searches based on MC driven backgrounds?

2012: Open question in preparation for Run 2 
2013: There is hope! 
2016: Precision 

**2017: Measurements!?**
Three cases where precise knowledge of QCD will be the critical in the long run (my guess)

1. Substructure: How far can the data-driven corrections go?

2. MET searches: PDFs and transfer factors

3. After the bump hunt era - the non-resonant searches

Bonus: H(bb) [to stay true to the title of the talk]

2. MET searches: PDFs and transfer factors
MET Based Searches

• Motivation - non-interacting or long life time

• Long lifetime - requires detailed detailed understanding of detector (will not cover here)

• MET based - “MET is hard to model”

• The use of CRs and transfer factors
MET Based Searches

- Motivation: Dark Matter (one of many)

Pie chart showing: Dark Energy 68.3%, Ordinary Matter 4.9%, Dark Matter 26.8%

Diagram showing production at colliders leading to indirect detection and direct detection.
MET Based Searches

• Motivation: Dark Matter (one of many)

Two new particles
Z’ mediator of mass $M_R$
DM candidate $\chi$ of mass $m_\chi$

Two new couplings
coupling $g_{SM}$ of $Z'$ to quarks
coupling $g_{DM}$ of $Z'$ to $\chi$

• Simple model to communicate results of many experiments
• Motivation: Dark Matter (one of many)

ISR

Two new particles
Z’ mediator of mass $M_R$
DM candidate $\chi$ of mass $m_\chi$

Two new couplings
coupling $g_{SM}$ of $Z'$ to quarks
coupling $g_{DM}$ of $Z'$ to $\chi$

• Simple model to communicate results of many experiments
MET Based Searches

- Motivation: Dark Matter (one of many)

  ISR

  Two new particles
  
  Z’ mediator of mass $M_R$
  
  DM candidate $\chi$ of mass $m_\chi$

  Two new couplings
  
  coupling $g_{SM}$ of Z’ to quarks
  
  coupling $g_{DM}$ of Z’ to DM

- Simple model to communicate results of many experiments
Looking for new physics in MET tails

Jet 0,
et = 921.98
teta = -0.463
phi = 2.508

MET 0,
pt = 913.68
teta = 0.000
phi = -0.657
Modeling MET

- Model MET by transferring information from control regions (CRs) to signal regions.
- Possible CRs by replacing $Z \rightarrow \nu \nu$ with:
  - $Z \rightarrow \mu \mu / e e$
  - $W \rightarrow \mu \nu / e \nu$
  - $\gamma$

Irreducible SM Background

CMS Simulation

- $\sigma \frac{d \sigma}{d E_T^{miss}} [1/\text{GeV}]$

- $E_T^{miss}$ ranges from 0 to 500 GeV.

- $m_\chi = 50$ GeV
- $g_q = g_\chi = 1$

- Comparison of $ZZ \rightarrow 4 \nu$ with different masses and production cross-sections.
Modeling MET

- Model MET by transferring information from control regions (CRs) to signal regions.

- Possible CRs by replacing $Z \rightarrow \nu \nu$ with:
  - $Z \rightarrow \mu \mu / ee$
  - $W \rightarrow \mu \nu / ev$
  - $\gamma$  [S. Kallweit, J. M. Linderta, S. Pozzorini]

M. Schönherr, P. Maierhöfer '15

References

5 Conclusions

The observed deviation presented by CMS at Moriond 2015 QCD (also shown in Fig. 6 of arXiv:1111.5206).

• $W$ + single top
• $Z$ (→ $\mu \mu$) + jets control sample.

Where appropriate, the last bin of the signal yields which vary with increasing uncertainty on the background expectations.

The uncertainties on the background expectations.

NLO QCD scale uncertainties are at the level of 10%. On the contrary, the observable is fairly stable in the considered.

Jet multiplicity $dN/dp_T$ at LO and NLO QCD + EW. Such a shift is consistent with preliminary results nu in uncertainty on the background expectations.

The uncertainties on the background expectations.

NLO QCD scale uncertainties at the level of 10%. On the contrary, the observable is fairly stable in the considered.

Jet multiplicity $dN/dp_T$ at LO and NLO QCD + EW. Such a shift is consistent with preliminary results nu in uncertainty on the background expectations.

NLO QCD scale uncertainties at the level of 10%. On the contrary, the observable is fairly stable in the considered.

Jet multiplicity $dN/dp_T$ at LO and NLO QCD + EW. Such a shift is consistent with preliminary results nu in uncertainty on the background expectations.

NLO QCD scale uncertainties at the level of 10%. On the contrary, the observable is fairly stable in the considered.

Jet multiplicity $dN/dp_T$ at LO and NLO QCD + EW. Such a shift is consistent with preliminary results nu in uncertainty on the background expectations.
Transfer Factors

$Z \rightarrow \mu\mu$

1 Control region
100% uncertainty @ 1 TeV

Adapted from P. Harris @ NPKI

G. Facini: QCD@LHC BSM & Higgs
Transfer

$Z \rightarrow \mu \mu$

Factors

$Z \rightarrow e e$

Adapted from P. Harris @ NPKI

2 Control regions
60% uncertainty @ 1 TeV

CMS-EXO-16-010

G. Facini: QCD@LHC BSM & Higgs

Aug 26, 2016
Transfer

3 Control regions
40% uncertainty @ 1 TeV

Factors

Adapted from P. Harris @ NPKI
Z → μμ

W → μν

4 Control regions
30% uncertainty @ 1 TeV

Adapted from P. Harris @ NPKI
Transfer

Z -> μμ

W -> μν

Factors

Z -> ee

W -> ev

5 Control regions
15% uncertainty @ 1 TeV

Adapted from P. Harris @ NPKI

CMS-EXO-16-010
CRs & Transfer Factors

- To maximize the sensitivity CMS uses data to constrain
  - scale systematics for Z & gamma+jets
  - EWK corrections (in a given MET bin)

![Graph showing scale & PDF uncertainty and EWK uncertainty]

- What do the uncertainties look like?
  - EWK Unc
  - Scale & PDF Unc

- Updated unc still too large

- Profiling them in the fit

- Still systematics limited @low MET
- Not systematics limited @ high MET
- → Likely will never be

- Constraints after the fit
- Limited by Theory unc.

Adapted from P. Harris @ NPKI

CMS-EXO-12-055

Limited by Theory unc.

G. Facini: QCD@LHC BSM & Higgs
CRs & Transfer Factors

- To maximize the sensitivity CMS uses data to constrain
  - scale systematics for Z & gamma+jets
  - EWK corrections (in a given MET bin)

What do the uncertainties look like?

EWK Unc
Scale & PDF Unc

Adapted from P. Harris @ NPKI

More simplistic approach taken by ATLAS
Progress!

• Looking forward to the direct impact on these analyses

See: K Mueller, V. A. M. Radescu, L. Harland-Lang, E. Rizv, S. Prestel,
Simplified Dark Matter Model

• For **this coupling**, model is alive only at high $M_{\text{med}}$

substructure analysis here!
An example of the transition from a bump hunt to a less/non-resonant search in the same kinematic regime
Three cases where precise knowledge of QCD will be critical in the long run (my guess)

1. Substructure: How far can the data driven corrections go?
2. MET searches: PDFs and transfer factors
3. H(bb)

QCD is the Key
H(bb)

- We have yet to observe H(bb): BR=58%
- Three ways to go: VH(bb), ttH(bb), VBF
  - concentrating here on the more sensitive 2
- Both *extremely hard analyses* as they sit on top of tremendous & difficult SM backgrounds
- backgrounds:
  - ttH: tt+X
  - VH: every SM process

ATLAS-CONF-2015-044
CMS-PAS-HIG-15-002
ttH(bb)

- To get to

- must fight with

[see S. Pozzorini, K. Lie, GK Krintiras, N. Castro]

Extensively discussed in this conference! Progress on measurement and theory side - great news!
ttH(bb)

- Already systematically limited
  See J. Thomas-Wilsker Talk
tt+jets

• ATLAS & CMS disagreements do not agree!

<table>
<thead>
<tr>
<th>Generators</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMS</td>
</tr>
<tr>
<td>• Powheg+Pythia8 with CUETP8M1 tune (default tt generator for Moriond EW)</td>
</tr>
<tr>
<td>• MG5_aMC@NLO+Pythia8 with FxFx matching (up to 2 extra partons at NLO)</td>
</tr>
<tr>
<td>• Powheg+Herwig++ with EE5C tune</td>
</tr>
<tr>
<td>ATLAS</td>
</tr>
<tr>
<td>• Powheg+Pythia6 with P2012 tune (default tt generator for Moriond EW)</td>
</tr>
<tr>
<td>• MG5_aMC@NLO+Herwig++ with UE-EE-5</td>
</tr>
<tr>
<td>• Powheg+Herwig++ with UE-EE-5</td>
</tr>
<tr>
<td>• Powheg+Pythia8 with A14 tunes (Main31, pThard = 0 and hdamp = mtop)</td>
</tr>
</tbody>
</table>

• Jet pT CMS (ATLAS) 30 GeV (25 GeV)

• ATLAS and CMS have opposite trends in data/MC

CMS-PAS-TOP-16-011

ATLAS-CONF-2015-065
VH(bb)

- Three channels, using leptonic decay of W, Z
  - Sensitivity maximal at high VpT
    - does not reach out to where substructure is profitable (used in searches)
  - Dominant channel is 0-lepton (1-lepton did not profit from increase $\sqrt{s}$)

[see N. Chernyavskaya talk]
VH(bb)

- Three channels, using leptonic decay of W,Z
- Sensitivity maximal at high $V_{pT}$
  - does not reach out to where substructure is profitable (used in searches)
- Dominant channel is 0-lepton (1-lepton did not profit from increase $\sqrt{s}$)

[see N. Chernyavskaya talk]

ATLAS-CONF-2016-091

validated w/ VZ
VH(bb)

- $Z+$Heavy Flavor is the dominant background in the most sensitive region

- $t\bar{t}$bar (2jet) and $W+$Heavy Flavor also very important
Z+Heavy Flavor

- A basic survey of what is out there.

[see S. Uccirati talk]

ATLAS-CONF-2016-046
[see J. Bossio, F. Zhang talk]
Z+Heavy Flavor

Systematics from MC comparisons/variations, data/MC in control regions

EW corrections?
Phase Space

Multivariate analysis (BDT)
Are we modeling properly all regions isolated by this MVA technique?

What can the theory community say about this?
In the long run

What will be the legacy* papers of the LHC?

*including in between runs
Conclusions

• Run 2 is here and we are burning through the data
  • QCD based searches are an important part of searches

• As Run 2 progress the bulk of the phase space for many searches will be **systematically limited**
  • there will always be stat limited tails and new low rate SM processes coming into reach

• *Do we have the patience to put in the hard work needed to beat down those systematic errors?*
  • of course! Already under way

• **The bedrock of this program is precision measurements!**

• But how far do we have to go? As far as we can …
Besides Bumps

\[ \frac{\sigma}{\sigma_{\text{th}}} \]

\[ \eta_{LL} = +1 \]

\[ \eta_{LL} = -1 \]

\[ m_{jj} > 3.4 \text{ TeV} \]

\[ \Lambda [\text{TeV}] \]

**ATLAS**

\[ \sqrt{s} = 13 \text{ TeV}, 3.6 \text{ fb}^{-1}; \eta_{LL} = +1 \]

\[ \sqrt{s} = 13 \text{ TeV}, 15.7 \text{ fb}^{-1}; \eta_{LL} = -1 \]

**ATLAS Preliminary**

- Observed 95% CL_{\alpha} upper limit
- Expected 95% CL_{\alpha} upper limit
- 68% CL_{\alpha} band
- 95% CL_{\alpha} band
DM @ ATLAS

DM Simplified Model Exclusions

ATLAS Internal

August 2016

DM Mass [TeV]

0

0.2

0.4

0.6

0.8

1

1.2

0

1

2

2.5

Mediator Mass [TeV]

DM = 0.25, g_A = 1

Axial-vector mediator, Dirac DM

DM Mass = Mediator Mass × 2

h_c Ω

Thermal relic

< 0.12

Perturbative unitarity
Complementarity
Substructure in 1 slide

- **trimming**: recluster jet constituents with $k_t \, R=R_{\text{sub}}$, drop if $p_T^{\text{sub}}/p_T<f_{\text{cut}}$

- **pruning**: recluster with C/A, killing wide angle with $\Delta R_{12} > R_{\text{cut}} \times 2M/p_T$ and soft with $f_2<Z_{\text{cut}}$

- **mass drop**: de-cluster until significant mass drop $m_{j_1}<um_j$ and not too asymmetric

- **soft drop**: remove soft wide-angle radiation

- **n-subjetiness**: how likely composed of n-subjets

- **D2**: A variation on the ratio of energy correlations which optimizes the separation between one-prong and two-prong decays, in analytical terms
Substructure

Figure 1. The distribution of $\rho = m^2 / (p_t^2 R^2)$ for tagged jets, with three taggers/groomers: trimming, pruning and the mass-drop tagger (MDT). The results have been obtained from Monte Carlo simulation with Pythia 6.425 [17] in the DW tune [38] (virtuality-ordered shower), with a minimum $p_t$ cut in the generation of 3 TeV, for 14 TeV $pp$ collisions, at parton level, including initial and final-state showering, but without the underlying event (multiple interactions). The left-hand plot shows $qq \to qq$ scattering, the right-hand plot $gg \to gg$ scattering. In all cases, the taggers have been applied to the two leading Cambridge/Aachen [39, 40] jets ($R=1$).

The parameters chosen for mass-drop ($y_{cut}=0.09$, $\mu=0.67$), pruning ($z_{cut}=0.1$, $R_{fact}=0.5$) and trimming ($z_{cut}=0.05$, $R_{sub}=0.3$) all correspond to widely-used choices.

The observation is that all three methods are identical to the plain jet mass for $\rho \gtrsim 0.1$. At that point, pruning and MDT have a kink, and in the quark-jet case exhibit a flat distribution below the kink. Trimming has a kink at a lower mass value, and also then becomes flat.

For gluon jets, the kinks appear in the same location, but below the kink there is no flat region. Pruning and trimming then each have an additional transition point, at somewhat smaller $\rho$ values, below which they develop peaks that are reminiscent (but at lower $\rho$) of that of the plain jet mass. Knowing about such features can be crucial, for example in data-driven background estimates, where there is often an implicit assumption of smoothness of background shapes. In this context one observes that for the upper-range of $p_t$'s that the LHC will eventually cover, $p_t \gtrsim 3$ TeV, the lower transition points for pruning and trimming occur precisely in the region of electroweak-scale masses.

To our knowledge the similarities and differences observed in Fig. 1 have not been systematically commented on before, let alone understood. Questions that one can ask include: why do the taggers/groomers have these characteristic shapes for the mass distributions?

At this point, a question arises of whether the LHC experiments are able to accurately measure EW-scale masses for TeV-scale jets. Challenges can arise, for example in terms of the angular resolution of the hadronic calorimeter, which may be relevant with current experimental reconstruction methods. Work in Ref. [41], however, suggests that with full use of information from tracking and electromagnetic calorimetry, which have higher angular resolution, good mass resolution for multi-TeV scale jets may well be possible.

[Dasgupta, Fregoso, Marzani, Salam '13]
ATLAS Substructure

- **Grooming**: to minimize impact of energy deposits from pile-up interactions
  - ATLAS mainly uses “Trimming” (arXiv:0912.1342):
    - re-cluster with $k_t R=0.2$ and remove sub-jets with $p_T^{\text{subj}} / p_T^{\text{jet}} < 0.05$

- **W/Z boson tagging**: [ATL-PHYS-PUB-2015-033]
  - $m_J$ consistent with $m_W/m_Z$ within ±15GeV
    - W and Z windows overlap
  - Sub-structure consistent with two-prong decay
    - Most popular variable: $D_2^{(\beta=1)}$ (arXiv: 1409.6298, 1507.03018)
    - Typical WP: $\varepsilon=50\%$, QCD rejection factor ~50

- **Higgs boson (b-) tagging**: [ATL-CONF-2016-039]
  - Match to anti-$k_t R=0.2$, b-tagged track-jets
CMS Substructure

'Bout That Boost: “V tag”

Jet Pruning

W bosons

QCD Jets

N-subjettiness

W bosons

More pure

Less pure

JME-13-006, JME-16-003, SMP-12-019