IMPROVING IDENTIFICATION OF DIJET RESONANCES AT THE LHC

Brian Shuve
Perimeter Institute for Theoretical Physics
In collaboration with Eder Izaguirre, Itay Yavin
arXiv:1407.7037 & work in progress

Rutgers University
14 October 2014
Summary of LHC Run I

- Entered the era of **precision Higgs physics**

- Still no signs of
  - SUSY
  - Exotic EWSB mechanisms
  - Dark matter
  - ...

---

**Summary of CMS SUSY Results** in SMS framework

<table>
<thead>
<tr>
<th>CMS Preliminary</th>
<th>ICHEP 2014</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed limits</strong></td>
<td><strong>Theory uncertainties not included</strong></td>
</tr>
<tr>
<td><strong>Only a selection of available mass limits</strong></td>
<td><strong>Probe “up to” the quoted mass limit</strong></td>
</tr>
</tbody>
</table>

**Mass scales [GeV]**

- $m_{\text{intermediate}} = x \cdot m_{\text{mother}} + (1-x) \cdot m_{\text{LSP}} = 0$ GeV

---

**Combined**
- $\mu = 1.00 \pm 0.13$

**Untagged**
- $\mu = 0.87 \pm 0.16$

**VBF tagged**
- $\mu = 1.14 \pm 0.27$

**VH tagged**
- $\mu = 0.89 \pm 0.33$

**ttH tagged**
- $\mu = 2.76 \pm 0.99$

---

**19.7 fb$^{-1}$ (8 TeV) + 5.1 fb$^{-1}$ (7 TeV)**

<table>
<thead>
<tr>
<th>CMS Preliminary</th>
<th>$m_{\chi} = 125$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Combined</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Untagged</strong></td>
<td></td>
</tr>
<tr>
<td><strong>VBF tagged</strong></td>
<td></td>
</tr>
<tr>
<td><strong>VH tagged</strong></td>
<td></td>
</tr>
<tr>
<td><strong>ttH tagged</strong></td>
<td></td>
</tr>
</tbody>
</table>

---

**Best fit $\sigma/\sigma_{\text{SM}}$**

- $0$ to $4$
Looking forward to Run II+

- Reach improved due to higher energies
  - Not much to do here: scale up cuts and look for obvious signs of new physics
Looking forward to Run II+

- What if new physics lies around the electroweak scale?
  - Can look very similar to SM physics

- **High integrated luminosity** gives a lot of signal events but....
  - Trigger thresholds are higher
  - SM backgrounds are higher
  - Pile-up is higher

- Cutting out the SM background cuts out most of the signal, too
  - Study BSM physics via precision SM physics

- Requires dedicated strategies to ensure we don’t miss anything
Hadronic resonances

- Ubiquitous in the Standard Model
  - W/Z/H/t
  - SM resonances can also decay leptonically, but suffer from smaller branching fractions
  - Want as many handles as possible on SM rates
  - (Possible) discrepancy in fully leptonic WW cross section

But see:

Meade, Ramani, Zeng
arXiv:1407.4481

Jaiswal, Okui
arXiv:1407.4537

Taken from Curtin, Jaiswal, Meade, arXiv:1206.6888
Hadronic resonances

- Also ubiquitous **beyond** the Standard Model
  - Extended Higgs sectors
  - R-parity-violating supersymmetry
  - Supersymmetric cascade decays
  - Extra dimensions
  - New gauge interactions

![Diagram](threshold)

![Diagram](boosted)
Hadronic resonances

- We are not guaranteed to do better at the LHC
- Extreme example: baryonic Z’

(taken from Dobrescu, Yu arXiv:1306.2629)

- Current approaches:
  - Some searches highly optimized, using sophisticated multivariable techniques (H to bb searches)
  - Others place simple cuts on jet kinematics and do a bump hunt (Z’→WW semileptonic, SM WW+WZ semileptonic,...)
  - Can we do better?
Hadronic resonances

- One method to improve sensitivity: move to new kinematic regime
  - Large boost means resonance decay products merge into “fat jet”
  - QCD jet looks nothing like t/W/Z/H jet - much recent progress on this!
  - Good for highly boosted regime, but suffer large kinematic penalties
  - Extra sensitivity to pile-up, underlying event (~ $R^3$)

- Can we still use some of the same strategies to improve identification of moderately boosted resonances?
  - In this case, the decay products of the resonance are separately resolved
Hadronic resonances

- Past work attempted to combine resolved and substructure searches
  - Gouzevitch et al., arXiv:1303.6636
  - Look at pair production of dijet resonances in 2j, 3j, 4j final states
  - Use substructure when decay products merged, apply similar cuts to resolved jets when not
  - Find that jet substructure observables don’t give much discriminating power away from highly boosted regime

- Are there still ways to separate resonance decay from “hard” QCD splitting?
  - We define a new observable that can be added to existing searches with a factor 2-6 gain in $S/B$
  - Outperforms other possible cuts; includes resolved jet masses
Outline

1. Jet substructure and the highly boosted regime

2. Resonance tagging in the mildly boosted regime

3. Examples
   - SM: WW+WZ
   - SM: V(H→bb)
   - BSM: Z’ → WW

4. Future directions
Jet substructure at high boost

• When an object is highly boosted, its decay products are collimated
  • Can be clustered together into a single, “fat” jet

\[ p_T \gtrsim (\text{few}) \times m_{\text{resonance}} \]

• Dominant background originates from a single QCD parton
Jet substructure at high boost

- The signal typically gives two hard subjets from the decay of a resonance, while the QCD subjets typically come from parton shower
  - Can take either a decomposition approach or energy-flow approach
Decomposition Approach

- Canonical example: BDRS mass-drop tagger (arXiv:0802.2470) (similar mass-drop procedure in HEPTopTagger)

1. Cluster a jet $j$ with the Cambridge/Aachen algorithm ($R = 1.2$)
2. Undo the last step, splitting $j$ into subjets $j_1, j_2$ with $m_{j_1} > m_{j_2}$
3. Discard $j_2$, set $j = j_1$, and continue de-clustering until both:

$$\frac{m_1}{m_j} < 0.67 \quad \text{(mass drop)}$$

$$\frac{\min(p_{T1}^2, p_{T2}^2)}{m_j^2} > 0.09 \quad \text{(symmetric splitting)}$$
Decomposition Approach

- This works because of their respective origins of jet masses
  - QCD masses come largely from sequences of asymmetric, wide-angle splittings
  - Signal masses come from the hard resonance decay

\[ \frac{m_1}{m_j} \ll O(1) \]

\[ \frac{m_1}{m_j} \sim O(1) \]
Energy-flow approach

• Uses inclusive, energy-flow information (algorithm-independent)
  • To compute $N$-subjettiness, define $N$ axes (denoted by Greek letters) and associate each particle $i$ in the jet to the closest axis

$$\tau_N^\beta = \frac{1}{N} \sum_{i \in j} p_{T_i} \Delta R_{i\alpha}^\beta$$

• This gives a measure of how well the radiation is aligned along $N$ axes
  • For dijet resonances, $\tau_2 / \tau_1$ performs better than $\tau_2$
  • Combines information from hard and soft radiation

• Generally, both approaches work well and are complementary
  • See upcoming BOOST 2013 working group report
Moderately boosted resonances
Tagging at moderate boosts

- Often, resonances are produced near threshold, paying a high penalty in signal acceptance for going to the boosted regime
  - Direct \( tt \), diboson, ...

- We can try to recover some of the signal discrimination in the **moderately** boosted regime
  - Still dominated by QCD splittings instead of uncorrelated emission
  - Requiring two separately resolved jets already eliminates most of the backgrounds from soft splittings
  - For further improvement, must now separate relatively hard QCD splittings from the signal

- Use decomposition approach because energy-flow observables are sensitive to all radiation between the jets
  - No longer need large-\( R \) jets
Tagging at moderate boosts

- Analogy of mass drop
  - The lax cut on mass drop from the boosted regime (<0.67) does not veto a hard QCD splitting
  - As jets become more widely separated, the mass drop becomes smaller
  - For background:

\[
\langle m_1^2 \rangle \approx C \frac{\alpha_s}{\pi} R^2 p_{T1}^2
\]

\[
m_{12}^2 \sim p_{T1} p_{T2} \Delta R_{12}^2
\]

\[
\frac{m_1}{m_{12}} \propto \frac{1}{\Delta R_{12}}
\]

- Interpolates between boosted and unboosted regimes
Tagging at moderate boosts

- This suggests that we \textbf{scale} the mass drop cut with $\Delta R$
- Exploit differences between QCD splitting and resonance decay
  - Signal has a mass drop that is more constant in $\Delta R$
  - QCD prefers asymmetric splittings, giving rise to \textbf{larger} $m_1$
  - In many examples with QCD backgrounds, one of the radiated partons is a gluon ($C_A > C_F$), giving rise to \textbf{larger} $m_1$ (on average)

- This motivates a new observable:

\[
\zeta \equiv \frac{m_1}{m_{12}} \Delta R_{12}
\]

$\zeta < \zeta_c$
Tagging at moderate boosts

• Other functional forms could accomplish a similar scaling
  • For example:

\[ \zeta(R_c) = \frac{m_1}{m_{12}} (\Delta R_{12} - R_c) \]

• These types of observables can be very effective at enhancing S/B when added on top of existing searches
  • Outperform other observables we studied
  • Robust performance under simple smearing and with different MC
  • Uses simple, small-\(R\) jet properties
Examples

• SM: WW+WZ
• SM: V(H→bb)
• BSM: Z’ → WW
**WW+WZ Analysis**

- Semileptonic channel is an independent check of the (possible) excess in the fully leptonic channel and an important SM measurement

- Simulate WW+WZ, W+jets events with Madgraph 5
  - Match matrix element to Pythia 6 parton shower using shower-$k_\perp$ scheme
  - Cluster and analyze events with Fastjet 3
  - Validated MC with CMS analysis
  - Include UE but no pile-up (more on this later)

- Use similar cuts as CMS 7 TeV (arXiv: 1210.7544), re-scaled to 13 TeV
  - Two jets with $p_T > 50$ GeV
  - One lepton with $p_T > 25$ GeV
  - MET > 50 GeV
  - $M_T > 50$ GeV
**WW+WZ Analysis**

- After CMS selection cuts:

![Graph showing WW+WZ Analysis](image)

\[ \sqrt{s} = 13 \text{ TeV} \int L dt = 5.0 \text{ fb}^{-1} \]

Events/6 GeV

**Legend:**
- Blue: WW+WZ
- Pink: W+jets

**Axes:**
- X-axis: \( M_{j1j2} \) (GeV)
- Y-axis: Events/6 GeV

**Values:**
- Events at 40 GeV: 100
- Events at 60 GeV: 800
- Events at 80 GeV: 1000
- Events at 100 GeV: 1200
- Events at 120 GeV: 1400
**WW+WZ Analysis**

- After CMS selection AND cut on $\zeta < \zeta_c$

![Diagram showing event distribution with $M_{j1,j2}$ (GeV) on the x-axis and Events 6 GeV on the y-axis, with different colors representing WW, WZ, and W jets.](image)
WW+WZ Analysis

- Gains for different choices of the cut

\[ \epsilon_S(\zeta < \zeta_c) \approx 3\% \]

\[ \epsilon_S(\zeta < \zeta_c) \approx 10\% \]

\[ \epsilon_S(\zeta < \zeta_c) \approx 25\% \]


**WW+WZ Analysis**

- How does this compare to other possible cuts we could have used?
  - Look in $M_{j_1j_2}$ window between 70-100 GeV
  - Use *filtered* version of shape/energy-flow observables

$$y = \frac{p_{T2}^2}{M_{12}^2} \Delta R_{12}^2$$
**WW+WZ Analysis**

- Would this be included in a BDT analysis?
  - Not currently used for SM WW+WZ
  - Seems there is substantial gain that comes from using resolved jet masses, which are not included in most BDT analyses

- Possible worry: jet masses are subject to uncertainties in shower mechanism & reconstruction

![Graph](image)

---

**Figure 1:** A plot of the improvement of signal over background against the signal efficiency for various amounts of smearing applied to the resolved jet mass. Smearing is implemented as a simple random gaussian smearing with a width of 10%, 20%, and 50% of the reconstructed mass.
**WW+WZ Analysis**

- Possible worry: jet masses are subject to uncertainties in shower mechanism & reconstruction
  - Show Pythia 6 vs. Pythia 8 vs. Herwig++

- Zeta performs well and is robust against various uncertainties
Limitations and Caveats

- Our observable gives a significant enhancement in $S/B$ at the cost of a mild reduction in statistical significance
  - Most applicable to searches dominated by systematic uncertainties
  - Will become more relevant for later LHC running

- What about pile-up?
  - Serious challenge facing high-luminosity running
  - We simulated $WW+WZ$ search with $\langle N_{PV}\rangle = 50$, found that a more aggressive form of jet trimming recovered $S/B$ gains to within 10-20%
  - Ongoing work needed for pile-up mitigation of small-$R$ jet masses
  - Our observable only involves small-$R$ jets
Examples

• SM: WW+WZ
• SM: V(H→bb)
• BSM: Z' → WW
**W(H→bb) Analysis**

- ATLAS and CMS have both dijet-mass and multivariate analyses
  - We follow the ATLAS 7+8 TeV analysis (now arXiv:1409.6212)

- Focus on dijet search, associated leptonic W
  - Dominant backgrounds are W+b+jets, tt

- Use same selection cuts as ATLAS
  - One tight lepton, p_T > 25 GeV
  - Exactly 2 b-tagged jets, p_T > 20 GeV (leading jet p_T > 45 GeV)
  - MET > 25 GeV
  - 120 GeV > M_T > 40 GeV
  - Loose selections on ΔR_{bb} as a function of p_T
  - Associate muons with adjacent b-jets to improve mass reconstruction
After ATLAS selection cuts:

- We replicate the ATLAS 7+8 TeV analysis at 13 TeV.
- We look for 2 b-tags, one tight-lepton, and MET.

\[ H \rightarrow b\bar{b} \]

\[ WH \text{ Example (With } H \rightarrow bb \) \]

\[ m_{bb} \quad \text{events} \]

\[ 1 \text{ lep., 2 jets, 2 tags, } p_T^V > 90 \text{ GeV} \]

\[ \sqrt{s} = 13 \text{ TeV } \int L dt = 300 \text{ fb}^{-1} \]

- WH(bb)
- WZ
- top
- W+ jets

\[ m_{bb} \quad \text{events} \]

\[ 10000 \quad 8000 \quad 6000 \quad 4000 \quad 2000 \quad 1000 \]

\[ 50 \quad 100 \quad 150 \quad 200 \quad 250 \]
**W(H→bb) Analysis**

- After ATLAS selection and a cut on the shifted version of $\zeta$:
  - Better at balancing preserving statistics and $S/B$ gain

\[
\zeta(R_c) = \frac{m_{j1}}{m_{j1j2}} (\Delta R_{12} - R_c) < \zeta_c
\]
W(H→bb) Analysis

- Gains for different choices of the cut:
**W(H→bb) Analysis**

- Is our gain just coming from the highly boosted region?
  - BDRS requires $p_{TV} > 200$ GeV

- If we restrict ourselves to the **moderately boosted** regime, $90$ GeV $< p_{TV} < 200$ GeV:
  - We still find an $S/B$ gain of $\sim 2-3$ (reduction of $\sim 25\%$)

- Our observable is effective in a boost range complementary to BDRS and other substructure methods

- Consider inclusion of jet masses in more sophisticated BDT as well
Examples

• SM: WW+WZ
• SM: V(H→bb)
• BSM: Z’ → WW
Z’→WW Analysis

- ATLAS has a search for resonant semileptonic WW/WZ production for masses up to 1 TeV (arXiv:1305.0125)
  - At higher masses, use jet substructure techniques
  - We consider a sequential SM Z’ decaying to WW
  - Dominant background is W+jets

- Use same selection cuts as ATLAS
  - Two jets, at least one with $p_T > 100$ GeV
  - One tight lepton, $p_T > 35$ GeV
  - MET > 40 GeV
  - $p_{TV} > 200$ GeV for each candidate gauge boson
  - $65$ GeV < $m_{jj}$ < 115 GeV
  - Various cuts on $\Delta \phi_{\ell \nu}$
Z'→WW Analysis

- After ATLAS selection cuts:

Note: large systematic uncertainties (~30%)
Z’→WW Analysis

- After ATLAS selection AND cut on $\zeta < \zeta_c$:

![Graph showing $M_{\mu\nu jj}$ distribution with $\sqrt{s} = 7$ TeV, integrated luminosity $Ldt = 4.7$ fb$^{-1}$ after cut $\zeta_c = 0.09$.](image)
$Z' \rightarrow WW$ Analysis

- $S/B$ gains and efficiency change:
Future directions
**Direct resonance production**

- Best bounds come from UA2/Tevatron
- At LHC, hard to pass triggers and discriminate from backgrounds
- Consider **associated production**
  - Provides handle for trigger
  - Gives resonance a (mild) boost

![Diagram of associated production](image)

- Recast of ATLAS techni-rho $W$+dijet search can beat Tevatron by a factor of a few in cross section
- Can we do better with an optimized search?
- What about $\zeta$/some similar observable?
- Work in progress
Multijet resonances

- Jet substructure can also be useful for three-jet resonances, but come at a cost of producing them well above threshold (ex. RPV gluinos in Curtin, Essig, BS arXiv:1210.5523)

- There are already good resolved 3-jet resonance searches (ex. Rutgers gp., CMS analysis arXiv:1311.1799)
  - Already in somewhat boosted regime
Conclusions

- Jet-substructure-inspired observables can improve identification of dijet resonances, even in the moderate boost regime/resolved limit
  - Interpolate between different kinematic regimes
    \[ \zeta \equiv \frac{m_1}{m_{12}} \Delta R_{12} \]  
    (and variations)
- Works well for two important examples of SM hadronic resonances
  - WW+WZ
  - V + (H → bb)
- Also useful in beyond-SM physics searches
  - Z’ → WW
  - Z’ → jj
- Uses standard-radius jets, no optimization for different R
- Let’s find out what LHC13 has in store!
Back-up slides
zeta distribution

\[ \zeta \text{ distribution} \]

\[ W + \text{jets} \]

\[ W + WZ \]

\[ (R_c = 0.2) \]
$\Delta R$ distribution

\begin{align*}
\text{prob.} & \quad \Delta R_{12} \\
\text{WW+WZ} & \quad \text{W+jets}
\end{align*}
zeta distribution

\[ \zeta_c \]

\[ \text{prob.} \]

\[ \text{WH} + Wj + WZ \]

\[ \zeta_c (R_c = 0.2) \]
After $\zeta = 0.11, R_c=0.2$ Cut
zeta distribution

\[ \zeta_c (R_c = 0.2) \]
q/g tagging

Taken from ATLAS q/g tagging study
arXiv:1405.6583