Highlights from High pT at LHC

Andrea Rizzi                      University and INFN Pisa

Trento, 12th Sept 2017
Which topic to cover with this “high pt” talk?

- V+jets → yesterday
- Top → tomorrow
- BSM → friday
- So I’ll cover Higgs and some di-boson

Are Higgs and di-boson not “strong enough” for this workshop?

- Higgs to bb and ttH
- Boosted topologies of boson hadronic decays
- VBF signatures (ok they are ewk “high pt processes” but with important “soft QCD” effects)
- HH final states
- In general backgrounds matters, and they are QCD (V+jets, top)
Outline

- Standard Model VH(bb) recent results
  - Details of the analysis
- Boosted Higgs
- VBF signatures
- di-boson
- HH (resonant and non resonant)
The Higgs boson discovery was one of the major results of LHC Run1.

ATLAS and CMS measured different decay and production modes with results $>3\sigma$ in most channels.

What was left after Run 1?
- bb decay mode
- ttH and VH associated productions

So from Run 2 we would expect:
- VH, with H→bb
- ttH in the various decay modes
- Higgs to muons?

We surf the wave... but top does too!
- And it is heavier!
VH

- Dominant VH production at LHC via
  - $qq \to V^* \to VH$

- Production via $ggZH$ only few %

- In the inclusive phase space

- The picture changes in the typical VH(bb) analysis phase space
  - $p_T \, V > \sim 100-150$ GeV

- Needed simulation of $ggZH$
LHC VH(bb) search

- Signal topology overview
  - Two b-tagged jets
  - 0,1,2 isolated leptons
  - or large missing energy

- Event features for S vs B separation:
  - dijet invariant mass
  - dijet system transverse momentum
  - Additional hadronic activity

- Main backgrounds:
  - V+jet
  - Ttbar & single top
  - diboson
  - QCD multi jet production
Di-jet mass

- B-jet energy is typically badly measured due to leptons+neutrinos in the decay
- Corrections can improve Mbb resolution
  - Additional constraint in $Z \rightarrow ll$ mode as the event is fully reconstructed
**V+jet background**

- Vector+jet is one of the dominant backgrounds
- Several difficulties in modelling this background
  - Important EWK and QCD corrections to the V-pt spectrum (see arxiv:1511.08692)
  - Experimental tags of 2 b-jets can have several origins (hard bb, gluon splitting, experimental mistags)
- Apply relevant corrections
  - Data to simulation (corners of phase space are not predicted as nicely as inclusive)
  - LO to NLO (when NLO fully simulated samples not available)
- Take into account V_pt uncertainties
  - V_pt is a crucial variable to separate signal from background
Background control regions

- Vector pt (and other important variables) checked after all corrections in dedicated control regions
- ttbar control region
- V+light control region
- V+heavy control region

**Z → ll b-enriched**

**Z → nunu b-enriched**

**Z → ll b-enriched**
Signal vs background separation

- Several features can be combined to separate S-B
- Multivariate analysis based on Boosted Decision Tree used to maximize sensitivity

<table>
<thead>
<tr>
<th>Variable</th>
<th>Channels utilizing</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M(jj)$: dijet invariant mass</td>
<td>All</td>
</tr>
<tr>
<td>$p_T(jj)$: dijet transverse momentum</td>
<td>All</td>
</tr>
<tr>
<td>$p_T(V)$: vector boson transverse momentum</td>
<td>All</td>
</tr>
<tr>
<td>$CMVA_{\text{max}}$: value of CMVA for the Higgs boson daughter with largest CSV value</td>
<td>2-lepton, 0-lepton</td>
</tr>
<tr>
<td>$CMVA_{\text{min}}$: value of CMVA for the Higgs boson daughter with second largest CSV value</td>
<td>All</td>
</tr>
<tr>
<td>$CMVA_{\text{add}}$: value of CMVA for the additional jet with largest CSV value</td>
<td>0-lepton</td>
</tr>
<tr>
<td>$\Delta\phi(V,H)$: azimuthal angle between $V$ and dijet</td>
<td>All</td>
</tr>
<tr>
<td>$p_T(j)$: transverse momentum of each Higgs boson daughter</td>
<td>2-lepton, 0-lepton</td>
</tr>
<tr>
<td>$p_T(\text{add.})$: transverse momentum of leading additional jet</td>
<td>0-lepton</td>
</tr>
<tr>
<td>$</td>
<td>\Delta\eta(jj)</td>
</tr>
<tr>
<td>$\Delta R(jj)$: distance in $\eta$--$\phi$ between Higgs boson daughters</td>
<td>2-lepton</td>
</tr>
<tr>
<td>$N_{\text{add}}$: number of additional jets</td>
<td>1-lepton, 2-lepton</td>
</tr>
<tr>
<td>N.B. definition slightly different per channel</td>
<td></td>
</tr>
<tr>
<td>$p_T(jj)/p_T(V)$: $p_T$ balance between Higgs boson candidate and vector boson</td>
<td>2-lepton</td>
</tr>
<tr>
<td>$M_t$: reconstructed top quark mass</td>
<td></td>
</tr>
<tr>
<td>$\Delta\phi(E^{\text{miss}}, \ell)$: azimuthal angle between $E^{\text{miss}}$ and lepton</td>
<td>1-lepton</td>
</tr>
<tr>
<td>$E^{\text{miss}}$: missing transverse energy</td>
<td>1-lepton</td>
</tr>
<tr>
<td>$m_T(W)$: $W$ transverse mass</td>
<td></td>
</tr>
<tr>
<td>$\Delta\phi(E^{\text{miss}}, \text{jet.})$: azimuthal angle between $E^{\text{miss}}$ and the closest jet with $p_T &gt; 30$ GeV</td>
<td>0-lepton</td>
</tr>
</tbody>
</table>

N.B. definition slightly different per channel
BDT? Can we trust BDT?

- Are BDT analysis reliable? (the problem is not the BDT)
  - The question should rather be, are our predictions of the backgrounds reliable
  - Can we really model all uncertainties?

- In general any background prediction is reliable only up to some given number of data events (real detector simulation is hard to model, qft calculations are limited in precision, etc.. etc...)
  - Typical “shape analysis” fits the data with signal and background models having some “uncertainty” parameterized as a scalar “nuisance parameter”
  - We need to be sure the background model have enough freedom in the fit

- So what? … cross check “candle” analysis!
Diboson: $V+Z(bb)$

- Try to measure $VZ$ (with $Z \rightarrow bb$) instead of $VH$ using the very same analysis technique.
- Larger cross section
  - Observe with 5 sigma
- Expected theory value already cross checked in leptonic final states
- Main differences
  - Different peak value of the dijet mass (but reasonably close)
  - Peaking non b-jet signal (but b-tag is well studied in other processes)
  - Di-boson is a background for Higgs, but higgs is less relevant background for di-boson (assume SM higgs x-sec)
Very nice agreement in control regions

Signal enriched bins compatible with SM predictions
Evidence of higgs to bb

Above 3 sigma in both experiments!
VBF Hbb

- Exploit VBF features to categorize events (with a mass-independent BDT)

- Forward tag: quark vs gluon discrimination

- Rapidity gap, m_\(jj\) mass, soft activity

- Fit (wide)peak on a smooth multi-jet QCD background

- Validate with Z→bb (VBF “pre-selection” but no VBF-BDT requirements)
**ttH, H→ bb**

- Similar techniques in ATLAS and CMS
  - Categorize in N-jets and N-bjets, define CR/SR
  - Dominated by irreducible tt+bb background (no “mass peak”)
  - Difficulties in modelling collinear/soft bb production

- Systematics are the largest contribution to the final uncertainty
How about ggH, $H \rightarrow bb$?

- Can we measure $H \rightarrow bb$ also in gluon fusion production?
  - QCD “$bb$” production is several orders of magnitude higher
  - How about the “boosted” regime?

- Jet substructures techniques tested in the past years
  - Observe hadronic final state of $V$ bosons and top
  - Mature for “measurements”
  - Recent studies on subjets b-tagging to distinguish from $V$-bosons

- New techniques to keep systematic uncertainty under control
  - Decorrelated Taggers

- All ready for boosted $H \rightarrow bb$?
Dedicated bb-tagger

- Multivariate bb-in-a-jet tagger exploiting substructure information (Nsubjettiness axis)
- Measure performance on data with two-muons-in-jet sample (as a proxy for g → bb)
Boosted Higgs to $bb$

**Anti b-tag (control region)**

- CMS Preliminary
- $450 < p_T < 1000$ GeV
- $\text{double-b tag} < 0.9$
- Backgrounds: W, Z, $t\bar{t}$, Multijet, Total Background
- Signal: H(bb), Data

**B-tag (signal region)**

- CMS Preliminary
- $450 < p_T < 1000$ GeV
- $\text{double-b tag} > 0.9$
- Backgrounds: W, Z, $t\bar{t}$, Multijet, Total Background
- Signal: H(bb), Data
Boosted $H \to bb$ results

- Sensitivity at 1 sigma level with 36/fb
- To be continued at high luminosity....
ttH - multileptons

- Decay of H via W/Z
  - Low branching ratios, but cleaner states
  - Look for >3 leptons or 2 same-sign lept
- Backgrounds: tt+vector boson, tt+jets, diboson
- Reaching 2/3 sigma sensitivity
- First evidence of ttH?

ATLAS data

ATLAS: $\mu = 2.5 \pm 1.3 - 1.1$
Various multilepton di-boson analysis repeated at 13 TeV

- Nice agreement with NNLO QCD
- Boosted analysis performed in final state with lepton+fat-jet
- Allow to measure $W$ pt in WW channel
- Probe high pt region where all lepton analysis are limited by $\sigma \times \text{BR}$
VBF/VBS and alike

- Same experimental techniques as for VBF Higgs tagging
- Associated production of Z, W, diboson with two jets
- Hard process is typically purely EWK
- QCD play a crucial role in
  - Identification of the two q-jets
  - Background modeling
  - Underlying event activity
- V+jets is the dominant background
- Rapidity gap distinguish S vs B
- Quard vs Gluon discriminator
- Soft-activity variables (e.g. HT of additional jets)
Results of “VBF” analyses

- **EWK Zjj production**

\[ \sigma_{EW} (\ell\ell jj) = 552 \pm 19 \text{ (stat)} \pm 55 \text{ (syst)} \text{ fb} \]

- **EWK diboson production (VBS):**

**Control region**

**Signal Region**
Another interesting VBF

- $H \rightarrow \mu\mu$
- VBF signature is the most sensitive and have largest S/B
- Need precise prediction of V+jets to use multivariate VBF techniques

<table>
<thead>
<tr>
<th></th>
<th>$S$</th>
<th>$B$</th>
<th>$S/\sqrt{B}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central low $p_T^{\mu\mu}$</td>
<td>11</td>
<td>8000</td>
<td>0.12</td>
</tr>
<tr>
<td>Noncentral low $p_T^{\mu\mu}$</td>
<td>32</td>
<td>38000</td>
<td>0.16</td>
</tr>
<tr>
<td>Central medium $p_T^{\mu\mu}$</td>
<td>23</td>
<td>6400</td>
<td>0.29</td>
</tr>
<tr>
<td>Noncentral medium $p_T^{\mu\mu}$</td>
<td>66</td>
<td>31000</td>
<td>0.37</td>
</tr>
<tr>
<td>Central high $p_T^{\mu\mu}$</td>
<td>16</td>
<td>3300</td>
<td>0.28</td>
</tr>
<tr>
<td>Noncentral high $p_T^{\mu\mu}$</td>
<td>40</td>
<td>13000</td>
<td>0.35</td>
</tr>
</tbody>
</table>

VBF loose: $S/\sqrt{B} = 0.21$
VBF tight: $S/\sqrt{B} = 0.38$
Higgs boson pairs interesting both for SM and exotics

- SM sensitivity is still far away
- bb+bb mode has largest BR
- Best sensitivity for $m_{res} > 500$ GeV
HH results

- Resonant analysis: low mass – resolved, high mass – boosted

Non-resonant SM, ATLAS 4b:

\[ \sigma(pp \rightarrow hh \rightarrow b\bar{b}b\bar{b}) < 330 \text{ fb} \]

30x the SM prediction:

\[ \sigma(pp \rightarrow hh \rightarrow b\bar{b}b\bar{b}) = 11.3^{+0.9}_{-1.0} \text{ fb} \]
Conclusions

- Many measurements ongoing at high-pt scale in Run2
- Current understanding of hard and soft QCD processes is often a limiting factor
- Vector plus jet is an everywhere background
  - Lot of theory calculation available in large regions of the phase space with NNLO accuracy
  - Large uncertainties in heavy flavour (gluon splitting) realm
- Data driven background models often (always) used for multijet processes
- Long term targets/more precise measurements will require perfect understanding of key kinematic distributions (perhaps it is not just a matter of how many Ns in front of LO)
- Stay tuned for $H \rightarrow \mu\mu$
backup
Hbb mass
CMS EWK measurements vs. Theory

- **qqW**: $0.84 \pm 0.08 \pm 0.18$ (19.3 fb$^{-1}$)
- **qqZ**: $0.93 \pm 0.14 \pm 0.32$ (5.0 fb$^{-1}$)
- **qqZ**: $0.84 \pm 0.07 \pm 0.19$ (19.7 fb$^{-1}$)
- **qqZ**: $1.02 \pm 0.03 \pm 0.10$ (35.9 fb$^{-1}$)
- **γγ→WW**: $1.74 \pm 0.00 \pm 0.74$ (19.7 fb$^{-1}$)
- **qqWγ**: $1.77 \pm 0.67 \pm 0.56$ (19.7 fb$^{-1}$)
- **ss WW**: $0.69 \pm 0.38 \pm 0.18$ (19.4 fb$^{-1}$)
- **ss WW**: $0.90 \pm 0.16 \pm 0.08$ (35.9 fb$^{-1}$)
- **qqZγ**: $1.48 \pm 0.65 \pm 0.48$ (19.7 fb$^{-1}$)
- **qqZZ**: $1.38 \pm 0.64 \pm 0.38$ (35.9 fb$^{-1}$)

All results at: http://cern.ch/go/pNj7
July 2017

Production Cross Section, $\sigma$ [pb]

- 7 TeV CMS measurement ($L \leq 5.0$ fb$^{-1}$)
- 8 TeV CMS measurement ($L \leq 19.6$ fb$^{-1}$)
- 13 TeV CMS measurement ($L \leq 35.9$ fb$^{-1}$)
- Theory prediction

All results at: http://cern.ch/go/pNj7

Fiducial W, Z and H searches with W$\rightarrow$lv, Z$\rightarrow$ll, H$\rightarrow$$\gamma\gamma$ and kinematic selection
VBF Hbb BDT

![Graph showing data and MC predictions for VBF Hbb BDT]
Boosted Hbb

CMS Simulation Preliminary 35.9 fb⁻¹ (13 TeV)

Events vs. $p_T$ leading-jet $m_{SD}$ (GeV)

- QCD
- W(qq)+jets
- Z(qq)+jets
- $t\bar{t}$+jets
- single-{$t$}
- VV(4q)
- ggH($bb$)

$s/\sqrt{b} \sim 0.1$

double-$b$ tagged

$s/\sqrt{b} \sim 1$
Jet Substructure

- Measures the degree to which a jet can be considered as composed of N prongs
- Energy correlation functions are sensitive to N-point correlations in a jet
  - A 2-pronged jet will have $e_3 < e_2$

\[
N_2^\beta = \frac{2e_3^\beta}{(1e_2^\beta)^2}
\]

\[
1e_2^\beta = \sum_{1 \leq i < j \leq n_j} z_i z_j \Delta R_{ij}^\beta
\]

\[
2e_3^\beta = \sum_{1 \leq i < j < k \leq n_j} z_i z_j z_k \min \left\{ \Delta R_{ij}^\beta \Delta R_{ik}^\beta, \Delta R_{ij}^\beta \Delta R_{jk}^\beta, \Delta R_{ik}^\beta \Delta R_{jk}^\beta \right\}
\]
Jet Substructure

$N^1_2$ sculpts jet mass distribution

Multijets background

Cut