VBSCAN Event Reconstruction Workshop Summary

Henning Kirschenmann
Steven Schramm
Senka Duric

Thank you for organizing a very nice workshop :) !

Local Organising Committee:

Magdalena Sławińska

Anna Kaczmarska

Paweł Bruckman de Renstrom

Bartłomiej Żabiński

V-tagging

- Machine learning is promising
- Many observables proposed
- Need to take care of low-level inputs; improve local reconstruction
  - Power of machine learning enhanced by more refined set of inputs
- Decorrelating mass and two-prong taggers
  - important if we want to use mass dependent observables => M(VV) and MT(VV) are sensitive observables to anomalous couplings
- combination of resolved and boosted regime
  - usually analyses so far have been separately optimized for the boosted and resolved regime
  - for the combination studies (differential cross sections) need to be aware that existing tagger outputs don’t have a smooth transition

Polarization

Not part of current state-of-the-art boson tagging techniques

- Optimize taggers for L/T categories?
- DNN-approach:
  - Split up W/Z categories further:
    - W_L/W_T/Z_L/Z_T

Practical problems:

- Can generate signal MC with 100% L/T fraction, but no easy “truth labeling” per jet for generic sample → use a mixture of pure samples when training
- Validation in data?
  - For W, this could be done (with difficulty) using ttbar and V+jets selections
  - For Z, there is no clear way to do this (Z+jets is there, no second sample with diff polarization)

Triggers for VBS

- Leptonic final states should work fine
  - Lepton triggers are typically quite loose
- Efficiency unclear for [all-]hadronic final states/Z→ invisible
  - Triggering these signatures is much more difficult
  - Hadronic final states in particular → huge loss of VBS efficiency using current jet triggers
- We should look into possibile new trigger paths for future data taking
  - How much can be gained from dedicated VBS triggers?

Pileup

- Timing layers/forward tracking will help a lot!
  - However, that's still >5 years away, what can we do before then?
- Mixed constituent-level suppression can help, using only calo information
  - Studied mostly in the central region so far, needs to be investigated in the forward region
  - Simple rho correction over-subtracts in forward region, can hurt calibrations
- Central jet activity veto also impacted by pileup
  - Need to improve central pileup suppression to reduce this
  - Are there other ways to reduce pileup impact on central activity vetoes?

Missing transverse momentum

- Very difficult for VBS, due to forward jet activity
  - Forward pileup jets are very difficult to control (before HL-LHC trackers)
- Can MET be calibrated?
  - Simple scale factors usually make the MET resolution much worse
  - Solution known for one analysis: derive in inclusive Z events, apply to W mass measurement
  - Can this be extended to other topologies?
- MET significance can improve ability to understand whether features are real
  - Accounts for object resolutions
  - May help with VBS due to difficulty of handling forward jets (include their resolutions)

Training on data

- we saw two methods on how to train on data
  - how can we be sure that there is nothing unexpected in the dataset?
    - Other processes entering different regions
    - Different p_T shapes or other important kinematic features
  - can we find another (third) control region for validation?
  - do we have enough statistics to use two control regions, split into training and validation sets?
- More fundamentally, what benefits are expected from training on data?
  - Is MC missing important discriminating features?
- How can we use data-trained classifiers in practice?
  - At some point, we usually want to compare to MC
  - Do we need to derive classification efficiencies for every MC generator we use?
  - Can we combine data-trained classifiers with “transfer learning” in a smart way?

q/g tagging

quark jets are well described with MC due to LEP constraints

gluon jets need to be additionally tuned in MC (no strong LEP constraints)

- LHC experiments should provide measurements of theoretically well defined quantities to perform tuning <= new Les Houches observables were proposed
- lower uncertainty from MC can result in lower theoretical uncertainties in measurements
    - q/g differences are the origin of the limiting uncertainty for many results using jets

Dynamic jet vetos

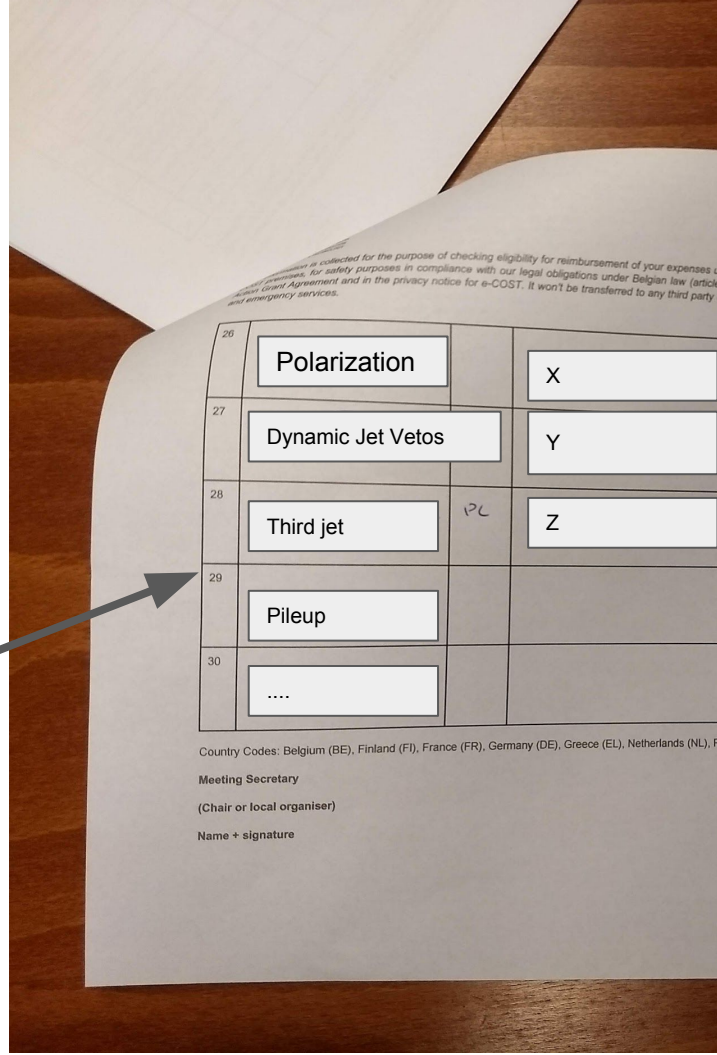
- interesting idea we can try out in CMS and ATLAS VBS analyses
  - Can we use $p_T^{lepton1}$ as a dynamic cut to reject ttbar/etc backgrounds?
  - Seems promising from what was shown (l,l,l,nu final state)
- Is this selection optimal?
  - Probably not, can be optimised for different final states
  - Even for this final state, perhaps average of the three lepton pT or similar may help
- Promising to investigate a unique dynamic veto for each final state

Third jet information

- we saw that using the third jet information could be beneficial to further reduce the background from VBS
  - angular information of the low pT third jet can be used
  - can be used for the VV+3jets category selection and combined with the VV+2jets for the legacy LHC results
  - scale of this third jet not used (beyond pT threshold), maybe track jets are possible?

Conclusion

- identified many areas where work is needed in order to improve the precision of VBS results
  - some defined ideas were presented and discussed
- additional (and expert!) workpower is very welcome in VBScan community :)
- Thanks for already signing up for various tasks on the participant list



Thank you for your participation!
Thanks for the hospitality in Krakow!

Discussion GDoc