BOOST 2021 Discussion Points

🎉 *Jet-vetting … without jet-setting!* ✈️

Links:
Conference webpage : https://indico.cern.ch/e/BOOST2021
Gather.Town : https://gather.town/app/teTVyx3VkBjYNe42/BOOST%202021

Contributions

5 : QCD physics measurements at the LHCb experiment
LHCb Collaboration

- Can jet substructure techniques help in tagging jets to identify heavy flavor? Ideally what techniques would you want and that would be complementary to the LHCb detector?
- Are there physics analyses at ATLAS or CMS with jets that could help resolve the intrinsic charm question?
- Javier D.: Slide 8: What are the inputs to the BDT? Would a deep learning approach based on low-level inputs possibly improve the discrimination?
- Javier D.: Slide 18: How is the c-tagging efficiency and systematic uncertainty measured in data?
- Javier D.: Slide 19: What are the other implications of intrinsic charm contribution (besides the relative Higgs production rates mentioned earlier)? Are there any models that predict a y-dependent effect? How much could the global PDF analysis change the results?

---

6 : Ab Initio Coupling of Jets to Collective Flow in the Opacity Expansion Approach
Andrey Sadofyev

- How general and extendable is this formalism? Is it straightforward to calculate multiple-point correlation observables on a jet in medium, like the energy-energy correlator?

---

9 : Jet Angularities in Z+jet production at the LHC
Oleh Fedkevych

- Can you define what you mean by “NLL” accuracy? What logarithms do you correctly resum to NLL? Do you capture all logarithms of the angularities that depend on the jet radius, or zcut, or non-global logarithms?

---

11 : The soft drop momentum sharing fraction zg beyond leading-logarithmic accuracy
https://videos.cern.ch/record/2776143

Pedro Cal

- At NLL accuracy, is it true that the zg distribution is the same with any choice of auxiliary observable? What is the NLL distribution of zg if the regulating observable is, say, the jet mass? Why is Rg the “preferred” auxiliary observable?

---

12 : Understanding top tagging with N-subjettiness and prong finding
Jack Helliwell

- How do the various procedures sculpt the top mass distribution? In particular, mMDT grooming on the top quark seems very strong; is the mass peak still around 175 GeV or is it moved to significantly lower masses or smeared out beyond the ungroomed width?

---

14 : Time-reversal odd side of a jet
Wai Kin Lai

- Do we need an EIC to study the T-odd jet function? Could it be probed at LHCb, for instance, where the detector is designed to capture forward physics?

---

15 : Spin-sensitive jet observables and their resummation
Alexander Karlberg

- I don’t quite follow the reasoning why the D-parameter isn’t sensitive to all-orders effects, but the spin correlation observables you define are. The all-orders effects in these spin correlation observables seem very small compared to fixed-order (see slide 9). Can you explain a bit more what you mean here?
- On Slide 5, you say the words “EPR like”. Do you have any ideas (even in principle) for how to perform a Bell type test in this case when one is only measuring energy flow in the detector (in standard Bell test one should measure two non-commuting observables.) In the context of cosmology, Maldacena made a toy model where one could overcome this issue. Are there toy models also in the collider context?
- You mention there are non-perturbative spin effects to be studied in the future. Do you expect these non-perturbative spin effects similar to the so-called T-odd jet effects (see talks by Wai Kin Lai) or the Collins effect of hadrons?

---

17 : Measurements of jet energy correlations and jet fragmentation, including heavy flavour, using the ATLAS detector
Javier Llorente o.b.o. the ATLAS Collaboration

- Is there any hope of measuring the TEEC on particles, rather than jets? Even just on charged particles, the endpoints of the cross section (especially near 0 angle) contain very interesting information about the structure of QCD.
- Javier D.: Slide 10: How is the systematic for the parametric function choice assigned? By trying different functions and using the difference? Or by profiling a discrete nuisance parameter [arXiv:1408.6865]?
- Javier D.: Slide 15: In the lower-right hand plot, is each histogram normalized independently? So the message is there is a greater fraction of g->bb in the data than predicted by Pythia?
- Javier D.: How can the shower tuning be improved by taking these g->bb measurements into account? Would that have a detrimental effect on other observables?

---

18 : Testing the Standard Model in boosted top quark production with the ATLAS experiment at the LHC
ATLAS Collaboration

- The comparison of simulation with the prediction for the lightly groomed top mass distribution is extremely encouraging. What do you see as the main limitations of the measurement? Will there be sufficient statistics for highly-boosted top quarks with Run 3 data?

---

21 : Boosting sensitivity in searches for Dark Matter and heavy resonances with ATLAS
Josu Cantero o.b.o. the ATLAS Collaboration

- For the DNN that was constructed for the ttbar resonance search (and perhaps other searches too), the input observables to the DNN are in general already strongly correlated. This can potentially mean that the DNN learns idiosyncratic information of Monte Carlos for tagging that is difficult to understand physically (if at all). Are there efforts to use less correlated observables (pronginess bases, EFPs, exclusive subjets, etc.) that would still ensure a broad coverage of possible jet topologies on phase space?
- Santeri: Are there trigger-related limitations/challenges in any of these analyses? Or do any of them rely on some non-standard/interesting "boosted" trigger algorithms?

---

22 : Jet and MET reconstruction and calibration in ATLAS
Christopher Young o.b.o. the ATLAS Collaboration

- So is the future of jet/ETmiss calibration just machine learning? Are there known outstanding physics problems that need to be addressed? On the other hand, the energy scale uncertainties that you showed were at the 1% level or even less in some cases. Can you hope to do any better than that?
- Are you able to tune MC to the measured single pion E/p and confirm that this explains the residuals measured with dijet, Z+jet and gamma+jet? The qualitative results on p.12 are very interesting, and it would be great to see the quantitative agreement with full tuning of MC.
- Have you considered directly measuring response of individual jet flavors (ud, s, c, b, g) in data? Flavor uncertainty on p.7 dominates at 30-400 GeV. Or do you expect ML for pi+ ID to

reduce flavor differences between quarks and gluons enough to make these uncertainties negligible?
-> Measuring the response in data is tricky but is being tried. Z+jet and gamma+jet have similar fractions giving little lever arm, and unbiased topological selections are tricky. But we are trying.

---

23 : Tagging boosted decays using large-radius jets with ATLAS
Yicong Huang o.b.o. the ATLAS Collaboration

- Are all taggers moving to machine learning now? Do the few-variable taggers demonstrate that there is physics that we are missing to understand the taggers? In your opinion, can theorists help develop new taggers?

---

25 : Performance and calibration of boosted H(bb) tagging, including applications in physics analysis with ATLAS
ATLAS Collaboration

- On high-pT jets that have been double b-tagged, what is the limit on resolution of additional emissions in the jet? For example, would it be possible to study in detail the distribution of radiation that lies between the two b subjets? This could have interesting consequences for the color representation of the larger jet.
- Loukas: Slide 5; the efficiency and mistag rate SF / subjet are shown up to pT~200-300 GeV, where in many applications the relevant pT region is higher. Are there SFs derived beyond the shown pT-range, or those extrapolated to higher values, and how?
- Loukas: Slide 14: How do the g->bb SFs compare to the Z->bb ?
- Santeri: Slide 14: What are the dominant sources of systematics for the X->bb SFs?
- Javier D.: Slide 18: Which analyses do you expect to utilize these taggers for Run 2 and Run 3?
- Javier D.: Slide 15: How pure is the ttbar control region?

---

26 : New Physics searches in boosted diboson final states at CMS
CMS Collaboration, Santeri Laurila

- With extensive use of machine learning for diboson searches, how amenable are these searches to recasting or for setting bounds on a different, but related physics searches?
  - Follow up: is there a plan to release HEPData (or other sources of additional information) with the performance of each of the ML-based taggers/algorithms for specific kinds of resonant signals, to allow for re-interpretations?

- Slides 7-8, for the X->aa->4b analysis:
  - Your average mass plot (slide 7) shows a deficit of data for three bins around 40-60 GeV immediately followed by an excess for two bins around 60-80 GeV. It seems like instead of being purely a statistical fluctuation, it may be that the background may be estimated with the

wrong slope in that region. Since you use an analytical function to determine your transfer factor when estimating the QCD background, were other functions considered, and how were the systematic effects of the choice of parameterisation taken into account in your background estimation?
- You mention a "2D fit to dijet mass and the average jet mass (mj1+mj2)/2" and then go on to show multiple exclusion limit plots with a fixed average jet mass and variable dijet mass. Was the fitting done in slices of these variables? Do you foresee having a full 2D plot with exclusion limits, including the evaluation of the limits in between the mass points for which you have MC signal samples available? (For example, by using techniques to estimate an intermediate mass point between two generated points).

- Max: s9: “DeepAK8-MD provides 2.5x sensitivity gain w.r.t. Double-b” → but what about with respect to a single-b? Double-b is very inefficient at highest invariant masses just because of track collimation in b-tagging, etc.

---

27 : Searches for heavy BSM particles resulting in boosted final states at CMS
CMS Collaboration

- Some of these searches seem to be approaching an absolute upper bound in searches for resonances at the LHC. How much further can these searches be pushed, even with Run 3 or HL-LHC data? What is the next frontier for BSM searches using boosted final states? Are there rare, but low mass, processes that can be probed?

---

30 : Jet substructure measurements at CMS
CMS Collaboration, Alejandro Espinoza

- The comparison of data with the newer generators was rather shocking. Is there any feedback to the Monte Carlo collaborations to use this data for tuning?
- Max: For your main metric of comparison you’re using the mean of the λ variables. Did you try to quantify the impact of the width, or even the tails? Often these matter more for taggers since we are cutting into these distributions.
- Max: Why does the shower/hadronization uncertainty bounce around so much?
- Max: is there anything meaningful to the flip of the simulation/data ratio for the p_T^D variable?

---

31 : Recent Developments in Jet Reconstruction and Calibration for LHC Runs 2 and 3 in CMS
CMS Collaboration, Andrea Malara

- What, if anything, can theorists do to help with calibration going forward? Previously, grooming, new jet algorithms, etc., had been developed, but they are now implemented in experiment and validated. Is the future in your opinion just ML, or are there still some outstanding physics issues to be resolved?

    - In my opinion, a good understanding of the generator level properties is always crucial. An example could be seen in the different modeling of fragmentation and UE for different generators.
- Chris: On slide 22/47 you show the in situ techniques. The last Z+jet points are at 600 GeV and there are 2 y+jet points above that but they are both below the linear fit which extends to the TeV region. How confident are you have a linear extrapolation is justified?
    - Answered online
- Chris: Slide 24/47 the pile-up uncertainty in the JEC is large at very high pT - why is this?
    - Answered online
- Slide 10/47: Presumably the CHS performance depends on how the PU vertex association is done, is there any expectation that CHS may improve in the future as a result of improvements to track-to-vertex association techniques etc?
    - Answered online
- Slide 12/47: You show a plot of drell-yan events, was this the main topology used to study the PU Jet ID, CHS and Puppi or did you also look at a jettier topology like dijet events?
- Slide 23/47: Why does the eta binning change between years?
    - A finer binning than the one showed is generally used. Bins are merged to gain statistics, as long as the SF are similar. Due to the increased radiation damage, a finer binning was needed to describe better the shape of the SFs.
- Max: s8, does the regression use no grooming? How does the peak at 0 from softdrop go away? Or does the regression somehow just correct everything at 0 to the correct value? This sounds like it could have weird effects on backgrounds, etc.
    - Also fine to wait for the dedicated talk
- Max: s9, where does the improved mass resolution come from? Is this just the effect of having a much “smaller” jet area, getting rid of the UE/PU on the outside? R=1.2 C/A seems like a poor comparison in this case, as it’s clearly much too large compared to the XCone jets
    - The comparison is done on the same jet area between C/A (Run1) and XCONE (Run2). The improvement is intrinsic in the subjet reconstruction of the two methods, as one is based on the angle and the other on the n-jettiness.

---

33 : Boosted jet tagging in CMS
CMS Collaboration, Huilin Qu

- In the discussion of mass sculpting, you mentioned that soft drop can push H > bb jets to have very small mass if one prong is removed. One way to control this is to change the parameters of the groomer or to just use machine learning. It seems that CMS just did the latter. So, are there outstanding physics issues that theorists can help with, or is it all machine learning now?
- Max: What exactly is the DNN doing to correct for the missed prong (referred to above?)? If you’re just multiplying up a very low reco mass to a high target, that’s a huge correction factor, and applying it to background would do very strange things, wouldn’t it?
    - As a followup-- are these jets even useful? They won’t pass any other substructure variables because you haven’t really restored the prong, and just applied a large correction factor, right?
- Max: why does the regression target the pole mass, instead of the corresponding truth jet mass? In the case that particles are out of cone, or otherwise missed, you have to make huge extrapolations with the regression. This could be very final state/density dependent, for

example. This is similar to some of the reasons why ATLAS/CMS calibrate jets to the truth jet scale, not quark (though there are additional reasons for that).

---

34 : Dynamical grooming beyond the leading-log approximation
Paul Caucal

- The non-exponentiation of dynamical grooming is intriguing. Does this potentially present a limit to the accuracy to which these observables could be calculated? Or, do you see a way forward to systematically improve the resummation, like with a factorization theorem?

---

36 : Jet substructure measurements in pp and heavy-ion collisions with ALICE
James Mulligan

- Grooming exclusively on track jets is subtle and interesting. Obviously, a jet composed exclusively of tracks does not contain the total energy/momentum of the jet, and so imposing the soft drop grooming constraint may bias the resulting distribution and correspondingly its interpretation. Do you have a sense for the size of charged to neutral effects for something like the zg distribution?
- What are, in your view, the next steps concerning jet substructure in heavy ions in order to probe the microscopic properties of the QGP?

---

37 : Measurement of splittings along a jet shower in sqrt(s)=200 GeV pp collisions at STAR
STAR Collaboration

- It's very interesting to see that the zg distribution is basically flat in the smallest Rg bin (Rg < 0.15). Is this because for these low pT jets at small Rg, you are actually sensitive to hadronic decays, and not the perturbative parton shower, or even "non-perturbative emissions"?
- Matt : there is a lot of jitter in the unfolded data points (slide 12), which seems to be larger than the error bands on these plots. Do these indicate the full uncertainty on the unfolded distributions?

---

41 : Safely Eating Junk: Pileup and Infrared Radiation Annihilation
Samuel Alipour-fard

- I'm not sure I follow why the discontinuity you describe is responsible for the spread in hadronization effects in slide 7. Further, the discontinuity indeed produces a cusp in the mMDT groomed mass distribution at fixed-order, but that is smeared by all-orders effects. So, what is wrong with a cusp/discontinuity?
- Matt : it looks like the RSS distributions on slide 17/18 have much smaller features in the region at low-mass dominated by NP effects (hadronisation). Have you checked whether different generators give more similar predictions for RSS than for SD in this poorly-modelled region?

- Matt : on slide 23, your calibration may have some deeper ramifications : what happens if you shift your background jets by the same factor? We've seen that the response for Q and q/g jets in ATLAS can behave rather differently, making it difficult to calibrate jets which have been groomed with certain algorithms like recursive soft-drop (2009.04986 Sec. 4.3).
- (I should say in advance that I really like this work/ talk, but this was the talk that I felt best to ask this question, so...)There are two arguments for the importance of grooming presented in your talk. One is experimental, i.e. for resolution for W-tagging (on this I very much agree with you, and grooming is of course very important from this perspective. Your work is nice progress in this direction). The second is for simplifying theoretical calculations (this one I am highly suspicious of). It is nice that your groomer gives a smooth behavior. The plot you show comparing Piranha with the kinked SD mass is very beautiful. However, another way of continuously weighting soft radiation is to simply use a correlator of energy flow observables. This can be viewed as weighting soft radiation linearly with E. It is simply impossible to argue that from a theoretical perspective the groomed observable is simpler. This is backed up by the wealth of beautiful analytic results for energy correlators in various theories, and their connections to the underlying physics of these theories (at weak coupling, strong coupling, in string theory, in CFT, whatever one wants), and the lack of such results for groomed observables (Larkoski's LO calculation being the only such analytic result, which is most certainly lacking any analytic beauty, and provides zero insight into QFT (this is not the fault of Larkoski)). This of course does not mean it is not interesting, but it means that one must really convince me that you are getting out some physics that I cannot get more easily from other more beautiful observables. If there is, then I am more than happy to give up on elegance, and I think it would be of great use to explain what one is learning physics-wise from such observables very clearly (particularly to explain to more formal theorists outside of the jet substructure community) e.g. for top mass, there is a clear reason, because there is a mass scale and so one wants a mass type observable but groomed. If there is not, then I am confused about this very roundabout way of constructing soft insensitive observables, and I prefer to stick to things that are beautiful.
- How robust is adjusting z-cut by hand? Specifically, if z-cut is adjusted for measuring the W-mass (or to account for underlying event in a particular case), is it possible to use this same z-cut elsewhere with PIRANHA-RSS?
- Slide 23, could you (in general) comment on the comparison between RSS and other recursive grooming methods (eg. recursive soft-drop)?
- Simone: Very nice idea and nice presentation! I'm wondering if you can say a few more words on the analytic understanding of RSS in comparison to soft-drop. Is the all-order structure similar?
- Adi: How does one describe the hadronization corrections that originate from the change in the outcome of the soft drop test for a subjet that is at the verge of failing/passing soft drop, now that the *groomer no longer has a sharp cutoff. Is this kind of correction even relevant now?

---

42 : Pure Quark and Gluon Observables in Collinear Drop
Xiaojun Yao

- It seems to me that the pure quark/gluon observables are defined deep in the infrared only. Is this true? If so, I don't follow how the shape function analysis is valid. Don't non-perturbative effects just dominate where the observables are defined?

- One small clarification: which MC generator are you using for the plots that you are showing? I think I got confused with your last bullet of the summary, since you mentioned that you need to study more MC.
- Max: As an ignorant experimentalist who may want to use something like this in data: what is the key improvement here compared to older studies of pure quark/gluon jet samples, like https://arxiv.org/abs/1104.1175 ? Is it that your work and observables is motivated by proper calculations and groomed observables, instead of something a bit more ad-hoc and purely MC based?
- Adi: How much does the shape function contributes relative to the perturbative contribution could depend rather sensitively on where you freeze the collinear-soft function scale (mu = 1 or 1.5). Have you checked that the sensitivity to this scale is reduced when taking ratios?
- Did you do any studies on how these techniques might respond in the presence of background or noise? (Such as Pile up) It strikes me as very hard experimentally to obtain the kinds of purity this technique seems to count on.

---

44 : Momentum and charge correlations within jets : a new observable to probe nonperturbative aspects of jet evolution
Mriganka Mouli Mondal

- How do neutral particles affect the interpretation of this observable? I guess it makes sense to focus on charged particles because of the high precision of trackers and relative ease for particle identification, but it seems simplistic to me to ignore neutral particles in hadron formation. Is the effect of neutral particles expected to be just some overall scaling from isospin or SU(3) flavor?
- Simone: Pythia and Herwig have different hadronisation models but also different parton showers. It would interesting to repeat your study with Sherpa, which would allow you to switch between different hadronisation models, while keeping the same parton shower.

---

46 : Probing Spin Interference with Energy Correlators
Ian Moult

- So do the spin correlations vanish in a conformal, supersymmetric theory, like N = 4? Also, could (or have) these spin effects be used to study hadronic physics? For example, would it be possible to probe the spin of a J/Psi or Lambda from its decay products with the EEEC?
- Max: why is g->bb particularly interesting? Just as a source of fairly pure gluons, or does the splitting to a heavy quark enhance the energy cut off you are describing?
- Max: and just to understand what this looks like in ATLAS or CMS: you're proposing essentially a particular energy-energy correlation measurement within jets? Or at the full event level, with jets as inputs?
- A very naive question; as you are able to probe helicity states, can you probe strong CP violation in g->gg and g->ggg? (maybe if we know the initial nature of the jet)

---

53 : Extending Precision Perturbative QCD with Track Functions

Yibei Li

- The track function moments are universal which is great for predictivity. However, given one data set, you must fit the distribution for the non-perturbative parameters, which is not (necessarily) predictive. What is the strategy for fitting/extracting the track function moments and then applying the fit values to other data? For example, is LEP data and then perturbative running sufficient for predictions on tracks at the LHC?
- Can you tell us something about the size of \Delta in QCD?
- In slide 8, for what you call “delta-function observable”, it has an infinite sum over track functions. Does this really mean you need infinitely many?

---

54 : Semi-supervised machine learning for pileup per particle identification at the LHC with Graph Neural Networks
Yongbin Feng

- On slide 14, it appears that the biggest difference between the PUPPI and GNN results is that the GNN removes neutral particles that are relatively isolated from charged particles. Is this actually an expected physical effect? Further, could you just include an isolation criteria into PUPPI to improve its performance?
- It’s definitely interesting that you are able to transfer the training from charged particles to neutral particles, but this also means you are unable to make use of features other than the kinematics and the spatial correlations. Do you see no prospect for associating labels to neutral PFOs in CMS, so as to make use of calo-specific information?
- Matt : how significant are the differences in performance on slide 12? If the differences are insignificant, would you use the Gated NoHybrid model anyway, or does another one offer advantages which are not captured by these metrics?
- Matt : on slide 16, the GNN and PUPPI weights look completely different. Do you understand what the GNN learning that PUPPI missed? Can we bring some new insight from this study back to improve our non-ML-based techniques?

---

55 : Towards precision αs and top mass extraction using soft drop jet mass
Aditya Pathak

- The mMDT groomer or soft drop with beta = 0 is a single logarithmic observable, and so alpha_s sets the slope of the distribution in the resummation regime. To measure the slope more precisely, one would want to extend the perturbative resummation regime as large as possible. You’ve plotted jets with pT > 600 GeV; could more sensitivity to alpha_s be gained by considering higher pT jets, with, say, pT > 1 TeV?
    - > Adi: Great question, yes, if you consider higher pT jets > 1 TeV the ultimate uncertainty for beta = 0 does go down to 1%. However, the NNLL uncertainties are still large and there isn’t much difference compared to pT > 600 GeV. I’ll upload plots for beta = 0 for the mixed case in backup slides :)

---

95 : Jet tagging in the Lund plane with graph networks
Frederic Alexandre Dreyer

- The LundNets (3 or 5) seem to contain much more information than actually exists in a jet. For instance, a jet with fixed total momentum with N particles has a phase space that is 3N-4 dimensional. However, the information that the LundNet5 keeps is 5N-5 dimensional (I think) and the universal approximation theorem says that any data that spans the 3N-4 dimensional phase space is all possible information that can be useful. So why is it helpful to include a significant amount of redundant data in the neural net? On the other hand, your resilience studies show that this redundant information might not be beneficial. Can you provide a rule-of-thumb for what information in jets is itself robust and resilient and therefore can usefully be used as input to the neural net?
- Would it be interesting to try to decompose the gains in the LundNet training wrt ParticleNet into the graph-building (use of ClusterSequence) vs the feature generation? E.g. by using raw kinematic info on the nodes rather than the Lund plane features.
- Matt : the exploration of resilience you showed is really nice, it would be interesting to know how W/top tagging performance changes if you do something similar with the npQCD emissions. Have you taken a look at that?
- Do the LundNets lend themselves generically to being used on groomed jets? Perhaps for SD and related grooming methods, this is more straightforward?
- You showed a nice plot of performance vs resilience to NP effects in the case of W tagging, do you have a similar plot for top tagging?
- Loukas: What is not entirely clear to me is how one can include additional information beyond particles in LundNet, like secondary vertices that are useful in b/c tagging. Do you have any thoughts on this?

Panel Discussion Points

- *What should the BOOST community try to accomplish / consider before we meet next year in Hamburg?*
- *As an experimentalist, where would you like to have more information/help from theorists?*
- *As a theorist, what are you missing from the experiments?*
- *Where do you see the biggest (not yet exploited) potential for ML-based techniques in the context of jet substructure?*
- *Where do you think the next big ideas in boosted physics will come from? What are the biggest outstanding problems of jets, in your opinion?*
- *How can one best facilitate an interaction with other areas of theory?*