Toward precision physics with jet substructure

Marat Freytsis
U. of Oregon ⟶ IAS

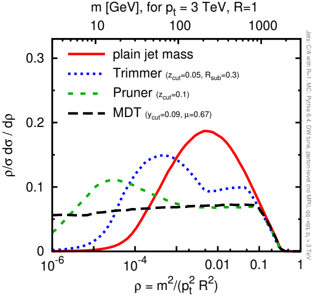
Università di Genova/INFN – Sezione di Genova
September 26, 2018

MDCCCLXXVI

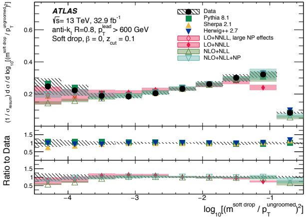
MF, Joshua Lin, Ian Moult, Ben Nachman [arXiv:1807.10768],
MF, Phil Harris, Andreas Hinzmann, Ian Moult, Nhan Tran,
Caterina Vernieri [arXiv:1807.07454]
+ ongoing work

From searches to precision

- Detailed LHC jet measurements
  ⟶ explosion of new ideas and observable
- Originally mostly focused on signals for searches
- Better theory understanding has made us realize many of these observables are interesting in their own right
- Detailed study of observables
  ⟶ start to really understand underlying theory



[arXiv:1307.0007]



[arXiv:1711.08341]

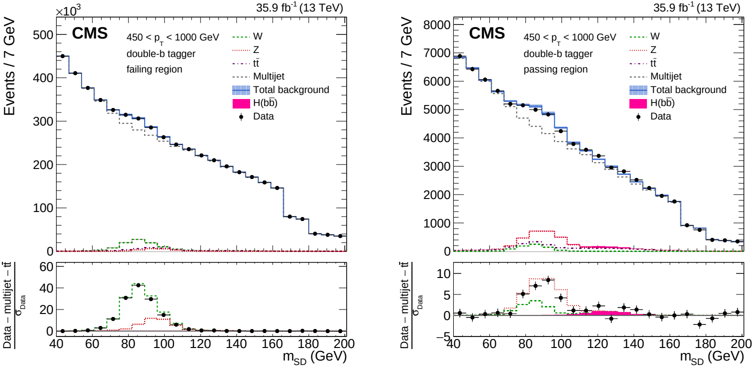
SM understanding from understanding jets

- We are now in a position to start using jet observables as probes of the SM itself
- Both experimental and theoretical understanding will have to advance to make full use of this opportunity
- But some early steps can already be taken

Plan

- Introduction
- Boosting $H \to b\bar{b}$ with machine learning
  - Method and architecture
  - Tagging boosted Higgs
  - Performance and projected constraints
- Prospects for an all-hadronic W mass measurement
  - General method
  - Statistical uncertainties
  - Systematic uncertainties
- Concluding thoughts

Boosted H on the verge



[arXiv:1709.05543]

Search based on double b-tagging and m_{groomed} event selection

More sensitivity through more substructure

- Unlike most background $b\bar{b}$, $H \to b\bar{b}$ is a color singlet
- 'Traditional' measurement of color flow via jet pull

[arXiv:1001.5027]

$$\mathbf{r}_i = (\Delta y_i, \Delta \phi_i), \qquad \mathbf{t} = \sum_{i \in \text{jet}} \frac{p_{T,i} |r_i|}{p_T^{\text{jet}}} \mathbf{r}_i$$

- More modern variable built out of N-subjetiness

[arXiv:1710.01305]

$$\beta_3 = \left(\tau_1^{(0.5)}\right)^a \left(\tau_1^{(1)}\right)^b \left(\tau_1^{(2)}\right)^c \left(\tau_2^{(1)}\right)^d \left(\tau_2^{(2)}\right)^e$$

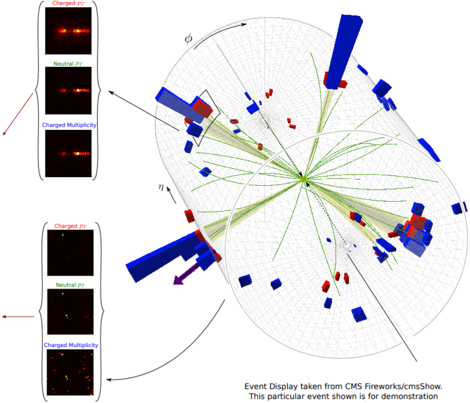
$$\tau_N^{(\beta)} = \sum_{i \in \text{jet}} \frac{p_{T,i}}{p_T^{\text{jet}}} \min\{R_{1i}^\beta, \ldots R_{Ni}^\beta\}$$

- Can we do better?

Taking the next step with machine learning

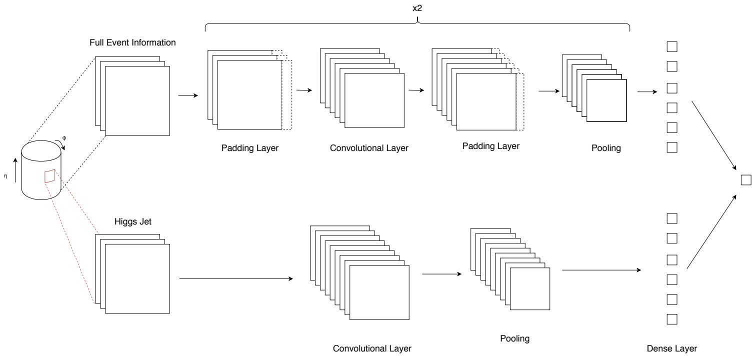
Three Color Channel Jet Images
(arxiv 1612.01551)

MADGRAPH5_aMC@NLO 2.6.2 -> PYTHIA 8.226
Signal : pp > Hj [QCD], Hjj[QCD]   Background : pp > jj,jjj,jjjj



Event Display taken from CMS Fireworks/cmsShow.
This particular event shown is for demonstration
purposes only.

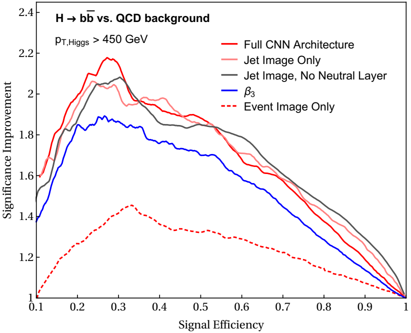
Two-stream convolutional neural network



- Split event into individual jets images and full event image
- Add padding layers to event image to account for ϕ invariance
- Smear core of jets inside event image to remove substructure

Tagger performance

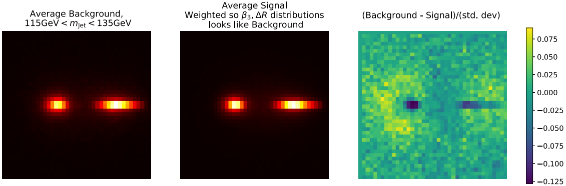
Significance ($= \epsilon_S / \sqrt{\epsilon_B}$) *improvement*



- Binned likelihood fit in 7 GeV bins for $p_T > 450$ GeV
- CNNs outperform best single human-built variables (β_3)
- Removing neutral layer still leads to noticeable improvement
  - method is pileup resistant

Tagger performance

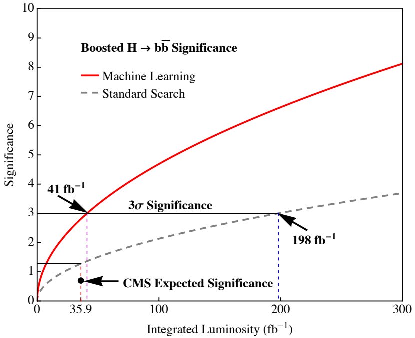
Are we learning anything new?



- Reweight all events by β_3^{-1} to remove correlated info
- Retrain network
- Remaining radiation pattern shows uncorrelated sensitivity remains

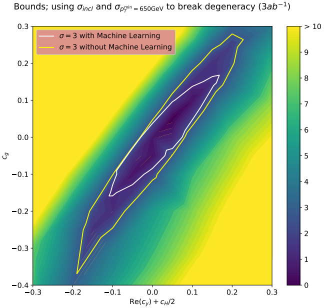
Tagger performance

Detection with integrated luminosity



- Observation possible with data already recorded
- Both discovery and measurement by the end of Run III
- $O(1)$ of the significance is coming from event image stream
  - Color singlet/octet information exists at sizable R away from jets

Constraining the SM EFT



Modifying $pp \to H$

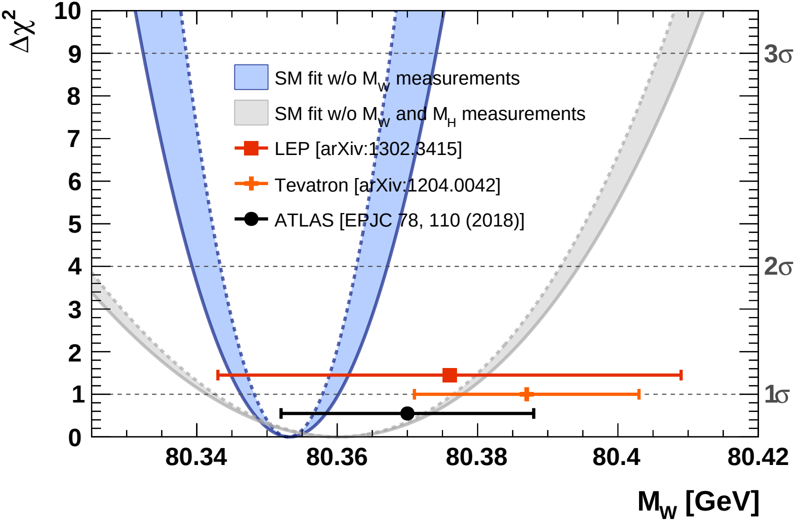
$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \left(c_y \frac{y_t}{v^2} |H|^2 \bar{Q}_L \tilde{H} t_R + \text{h.c.} \right) + c_H \frac{1}{2v^2} \partial_\mu |H|^2 \partial^\mu |H|^2$$
$$+ c_g \frac{\alpha_s}{12\pi v^2} |H|^2 G^a_{\mu\nu} G^{a\mu\nu} + \tilde{c}_g \frac{\alpha_s}{8\pi v^2} |H|^2 G^a_{\mu\nu} \tilde{G}^{a\mu\nu}.$$

- Degeneracy in total rate only broken at high p_T
- Constraints with ML will be better than global average without Higgs by factor of ~ 2–3

Plan

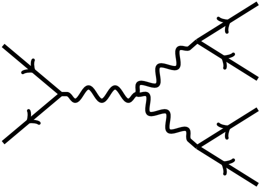
- Introduction
- Boosting $H \to b\bar{b}$ with machine learning
    - Method and architecture
    - Tagging boosted Higgs
    - Performance and projected constraints
- Prospects for an all-hadronic W mass measurement
    - General method
    - Statistical uncertainties
    - Systematic uncertainties
- Concluding thoughts

Current state of m_W



Gfitter, [arXiv:1803.01853]

Previous measurements

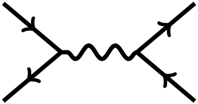


$$\Delta m_W^{(\ell\nu qq)} = \pm 54 \pm 25\,\mathrm{MeV}$$

$$\Delta m_W^{(4q)} = \pm 70 \pm 28 \pm 28\,\mathrm{MeV}$$

stat. + syst. (+ FSI)

ALEPH, [arXiv:hep-ex/060511]



$$\Delta m_W^{(\ell\nu)} = \pm 7 \pm 11 \pm 14\,\mathrm{MeV}$$

stat. + exp. syst. + mod. syst.

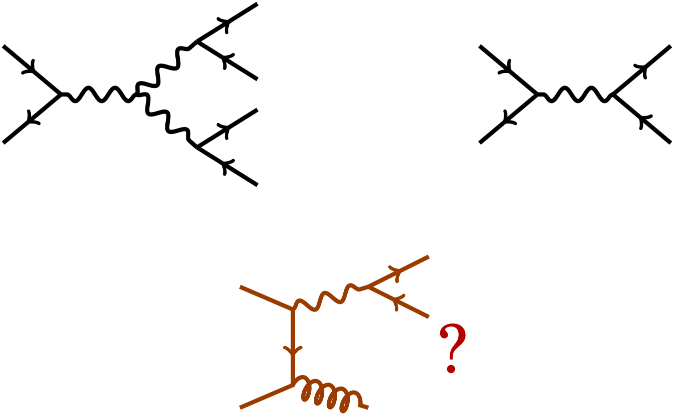
ATLAS, [arXiv:1701.07240]

also CDF, [arXiv:1203.0275]

D0, [arXiv:1203.0293]

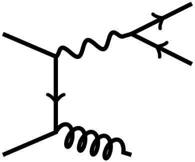
The all-jet final state?

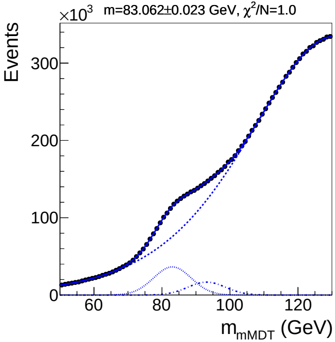
(HL)-LHC edition



hard? ⟷ crazy?

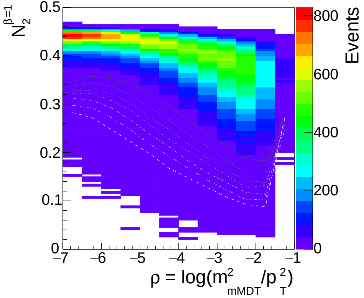
Signal and background

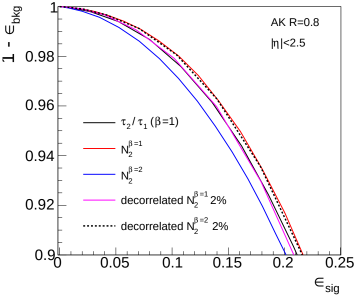




- $W+$jets, $Z+$jets, QCD multijets, $t\bar{t}$, single t
- MadGraph with simple detector simulation tuned to current ATLAS/CMS jet substructure performance
- Pseudodata corresponding to HL-LHC luminosity

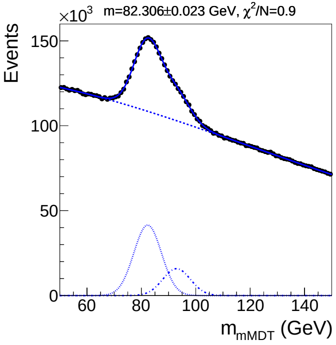
Choice of tagger



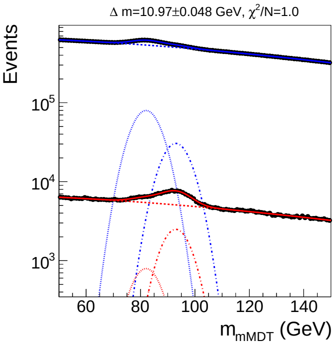


- Flatten background by decorrelating jet substructure selection from m_j [arXiv:1603.00027]
- Small effect on signal efficiency, better control of background

Extracting of W and Z mass peaks



$p_T > 500\,\mathrm{GeV}, N_2^{\beta=1} 2\,\%$



double b tag

- Enriched sample of Z bosons with double b tag
- Measure Δm_{WZ} so that many experimental systematics cancel

Statistical uncertainty

- Assume current detector performance and triggers
- Statistical precision for m_W

Selection	Int. luminosity	σ_{m_W} [MeV]
decorrelated $N_2^{\beta=1} 1\%, p_T > 500\,\mathrm{GeV}$	$300\,\mathrm{fb}^{-1}$	75
decorrelated $N_2^{\beta=1} 1\%, p_T > 500\,\mathrm{GeV}$	$3000\,\mathrm{fb}^{-1}$	23

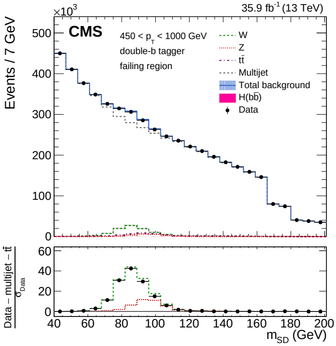
- Statistical precision for Δm_{WZ}

Selection	Int. luminosity	σ_{m_W} [MeV]
decorrelated $N_2^{\beta=1} 2\%, p_T > 500\,\mathrm{GeV}$	$300\,\mathrm{fb}^{-1}$	171
decorrelated $N_2^{\beta=1} 5\%, p_T > 500\,\mathrm{GeV}$	$3000\,\mathrm{fb}^{-1}$	48

  - Limited by $Z \to b\bar{b}$ cross section

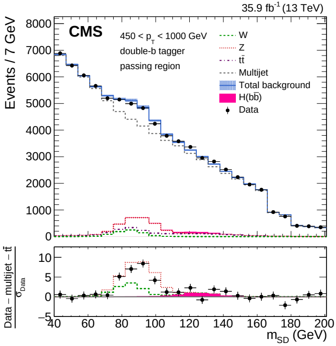
Where the LHC stands now

no double *b* tag



$\Delta m \sim 200\,\mathrm{MeV}$

double *b* tag

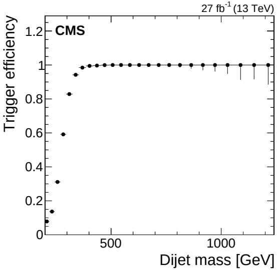


$\Delta m \sim 1\,\mathrm{GeV}$

CMS, [arXiv:1709.05543]

Trigger strategy

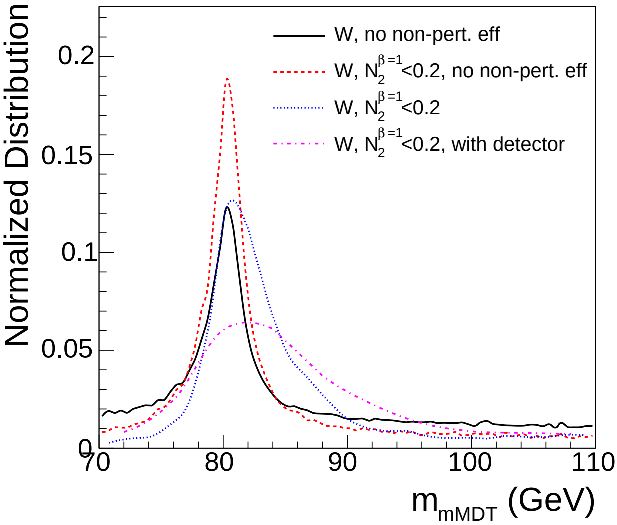
- Current trigger threshold for ATLAS/CMS $p_T \gtrsim 500\,\mathrm{GeV}$
- Alternative approaches storing lower size events at higher rates allows going to $p_T \gtrsim 200\,\mathrm{GeV}$
- Assume substructure evaluated at L1 or HLT level at HL-LHC



CMS, [arXiv:1806.00843]

Strategy	Selection	Int. luminosity	σ_{m_W} [MeV]
measure m_W	decorrelated $\mathrm{N}_2^{\beta=1}$ 1%, $p_T > 500\,\mathrm{GeV}$	$3000\,\mathrm{fb}^{-1}$	23
measure m_W	decorrelated $\mathrm{N}_2^{\beta=1}$ 1%, $p_T > 400\,\mathrm{GeV}$	$3000\,\mathrm{fb}^{-1}$	21
measure m_W	decorrelated $\mathrm{N}_2^{\beta=1}$ 2%, $p_T > 300\,\mathrm{GeV}$	$3000\,\mathrm{fb}^{-1}$	13
measure $m_Z - m_W$	decorrelated $\mathrm{N}_2^{\beta=1}$ 5%, $p_T > 500\,\mathrm{GeV}$	$3000\,\mathrm{fb}^{-1}$	48
measure $m_Z - m_W$	decorrelated $\mathrm{N}_2^{\beta=1}$ 5%, $p_T > 400\,\mathrm{GeV}$	$3000\,\mathrm{fb}^{-1}$	40
measure $m_Z - m_W$	decorrelated $\mathrm{N}_2^{\beta=1}$ 5%, $p_T > 200\,\mathrm{GeV}$	$3000\,\mathrm{fb}^{-1}$	32

Systematic uncertainties



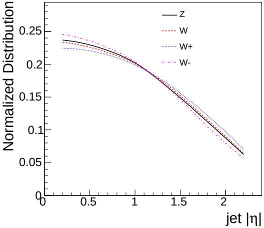
Experimental uncertainties

- Assume particle-flow reconstruction, evaluating systematic effects separately on charged particles, γ (and π^0), and neutral hadrons
- Estimate precision of energy scale calibration needed to achieve $\Delta m_W < 10\,\mathrm{MeV}$

Effect	Understanding needed for $\sigma_{m_W} = 10\,\mathrm{MeV}$	Typical current precision
Charged particle energy scale	0.03%	0.05%
Photon (and π^0) energy scale	0.06%	0.1%
Neutral hadron energy scale	0.1%	1%
200 pileup interactions	1.4%	1%

- Uncertainties cancel when measuring Δm_{WZ}
  - Residual effects from hadronization model affecting $W \to q\bar{q}'$ vs. $Z \to b\bar{b}$ jet response (more below)

Perturbative effects



$p_T > 300\,\mathrm{GeV}$

Effect	Size of effect	Understanding needed for $\sigma_{m_W} = 10$ MeV
NLO QCD	8 MeV	✓
NLO EW	1 MeV	✓
NLO PDF	1 MeV	✓
$N_2^{\beta=1} < 0.2$ selection	200 MeV	5%

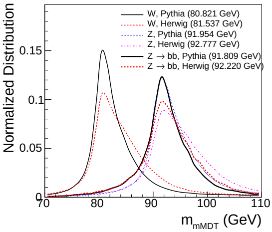
- Prediction of W boson kinematics not a limiting factor in hadronic final state
- Need prediction at 5 % level of how much substructure selection changes m_W

Nonperturbative effects

- Disabling non-perturbative effects (MPI and hadronization in PYTHIA8) to estimate size of effect on both m_W and Δm_{WZ}
  - 10 times smaller for Δm_{WZ} than for m_W
- Comparing $Z \to q\bar{q}$ and $Z \to b\bar{b}$ mass peaks to estimate size of hadronization effects on Δm_{WZ}

Quantity	Effect	Size of effect	Understanding needed for $\sigma_{m_W} = 10\,\text{MeV}$
m_W	$N_2^{\beta=1} < 0.2$ selection	310 MeV	3%
m_W	non-pert. corrections	1100 MeV	0.9%
Δm_{WZ}	non-pert. corrections	110 MeV	9%
m_Z	$Z \to q\bar{q}$ vs. $Z \to b\bar{b}$	140 MeV	7%

Where MC generators stand now

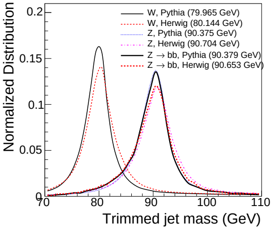


- Estimate current understanding of convolution of perturbative and nonperturbative effects by comparing Pythia8 and Herwig++
- Depends on grooming algorithm and substructure selection

- $\delta m_W^{\mathrm{had}} \sim 200\text{–}1000\,\mathrm{MeV}$
- $\delta m_{WZ}^{\mathrm{had}} \sim 50\text{–}500\,\mathrm{MeV}$
- $\delta m_Z^{b\bar{b}} \sim 50\text{–}500\,\mathrm{MeV}$

Where MC generators stand now

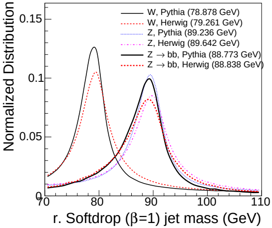
- Estimate current understanding of convolution of perturbative and nonperturbative effects by comparing Pythia8 and Herwig++
- Depends on grooming algorithm and substructure selection



- $\delta m_W^{\text{had}} \sim 200\text{–}1000\,\text{MeV}$
- $\delta m_{WZ}^{\text{had}} \sim 50\text{–}500\,\text{MeV}$
- $\delta m_Z^{b\bar{b}} \sim 50\text{–}500\,\text{MeV}$

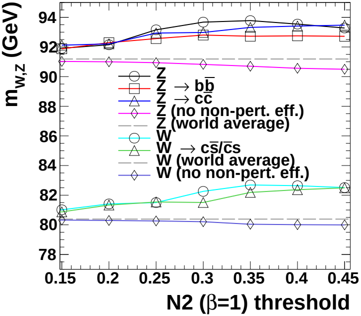
Where MC generators stand now

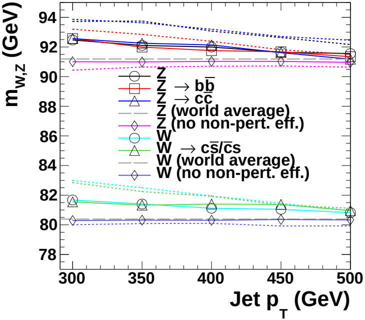
- Estimate current understanding of convolution of perturbative and nonperturbative effects by comparing Pythia8 and Herwig++
- Depends on grooming algorithm and substructure selection



- $\delta m_W^{\text{had}} \sim 200\text{–}1000\,\text{MeV}$
- $\delta m_{WZ}^{\text{had}} \sim 50\text{–}500\,\text{MeV}$
- $\delta m_Z^{b\bar{b}} \sim 50\text{–}500\,\text{MeV}$

Constraining nonperturbative effects





- Nonperturbative effects strongly reduced by substructure selection and at high jet p_T
- Pythia–Herwig difference for Δm_{WZ} reduced to 10–50 MeV at $p_T > 500\,\mathrm{GeV}$
- Differential measurement of Δm_{WZ} vs. p_T and substructure promising to contain nonperturbative effects

Discussion

- The leading theoretical task will be an extraction of nonperturbative corrections, either from other data or self-consistently with mass measurement itself
  - W boson groomed N_2 and groomed mass (a color singlet)
    - *cf.* groomed D_2 [arXiv:1708.06760,1710.00014,1710.06859]
    - *cf.* groomed m_t [arXiv:1708.02586]
  - A statement on universality of nonperturbative corrections for hadronic W and Z decays
- Measurement of m_W peak interesting in itself, since it can help to better understand hadronization of boosted W/Z bosons, supporting searches
- HE-LHC would allow access to even higher p_T with smaller nonpertrubative effects

Plan

- Introduction
- Boosting $H \to b\bar{b}$ with machine learning
  - Method and architecture
  - Tagging boosted Higgs
  - Performance and projected constraints
- Prospects for an all-hadronic W mass measurement
  - General method
  - Statistical uncertainties
  - Systematic uncertainties
- Concluding thoughts

Conclusions

- Using modern ML techniques high-p_T $H \to b\bar{b}$ is already on the verge of od discovery
  - Can soon be turned into measurement and precision constraint opportunity
  - Our two-stream NN approach suggests there is more information in color flow waiting to be used
- Hadronic m_W measurement could avoid experimental systematic uncertainties related to measurement of $\not{E}_T$ and theoretical uncertainties related to m_T
  - Measurement of Δm_{WZ} more feasible than m_W itself
  - New trigger strategies needed to reach statistical uncertainty of 30 MeV with 3000 fb^{-1} of HL-LHC data
  - Measurement limited by the understanding of nonperturbative contributions the the masses of $W \to q\bar{q}'$ and $Z \to b\bar{b}$
- In both cases, much work still to be done

Thank you!