# *Jet Substructure in top/Higgs*

Gregor Kasieczka on behalf of ATLAS/CMS

*Benasque Workshop*

*2018-05-29*

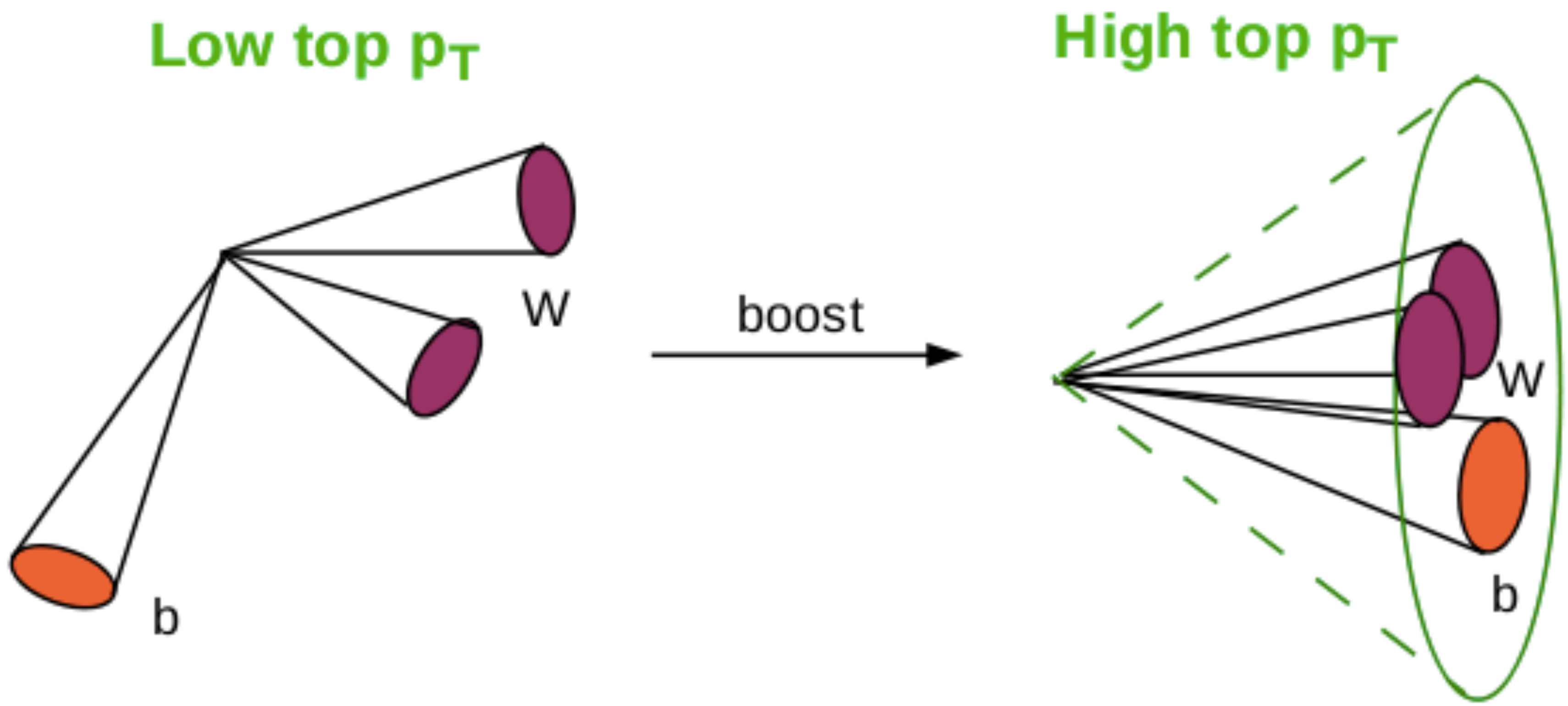
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Emmy Noether-Programm
Deutsche Forschungsgemeinschaft
DFG

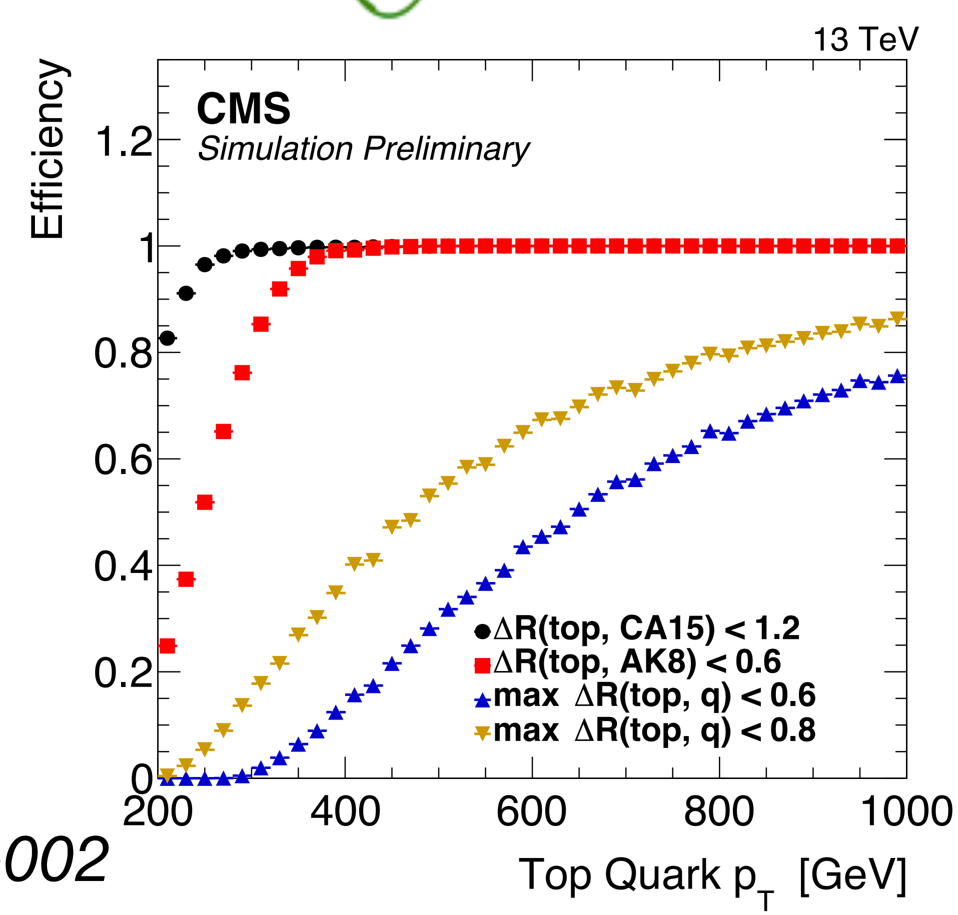
Bundesministerium für Bildung und Forschung

# Jet Substructure

# The (very) basics

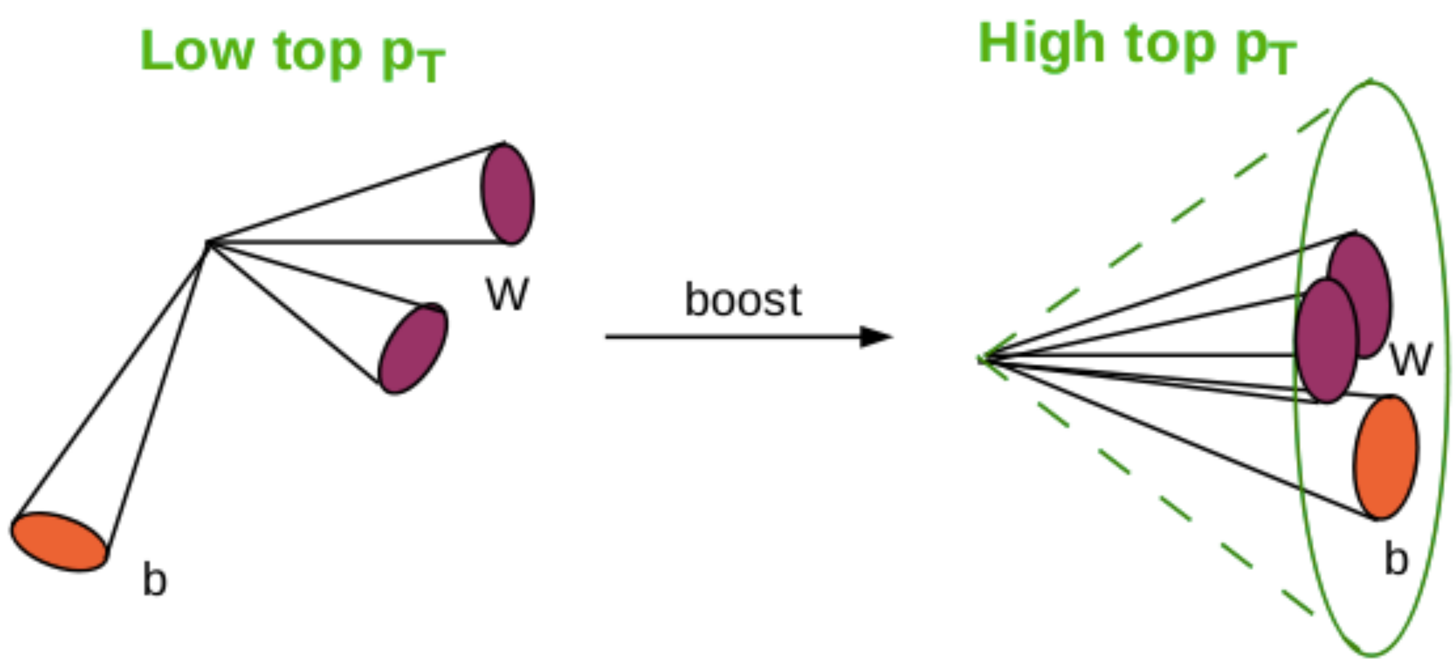



- Hadronically decaying top/Higgs
- Contained in one (large-R) jet
- How to distinguish from light quark/gluon/W/Z/.. jets?




*CMS PAS JME-15-002*

# Jet Substructure




- Hadronically decaying top/Higgs/W/Z
- Contained in one (large-R) jet
- How to distinguish from light quark/gluon jets (and from each other)
- For new physics searches (and SM studies)

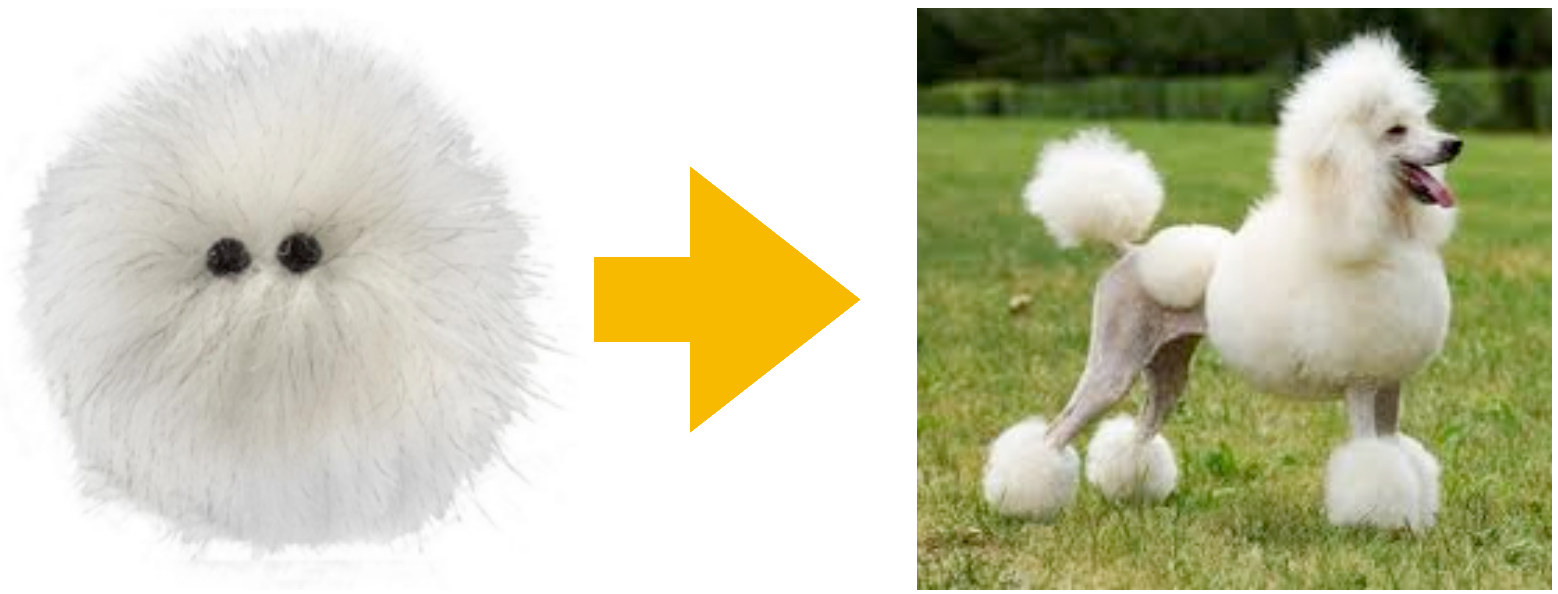
*Some Classical solutions:*

- Mass
  *Calculate after removing pile-up/soft radiation (eg mMDT/softdrop or pruning)*
- Centers of hard radiation
  *n-subjettiness or energy correlation functions*
- Flavour
  *b tagging of large-R jets or subjets*
- Soft substructure
  *Color connection*
- Inclusive reconstruction
  *HEPTopTagger V2, HOTVR*
- Other substructure variables
  *Shower deconstruction, template tagger, …*

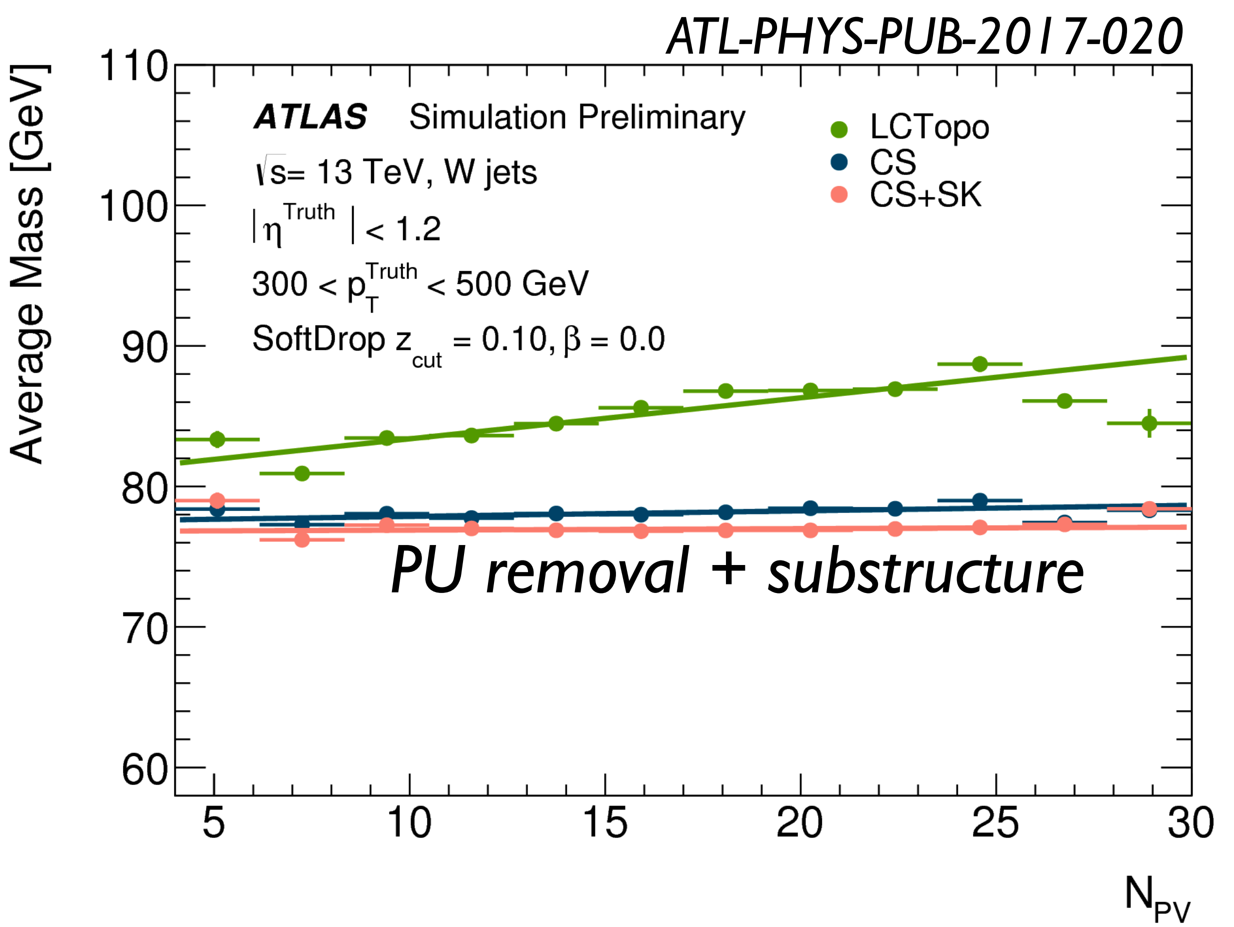
*Towards an Understanding of the Correlations in Jet Substructure*
D Adams et al (BOOST 2013 Participants), Eur.Phys.J. C75
*Top Tagging,* T Plehn, M Spannowksy, J.Phys. G39 (2012) 083001
*Boosted Top Tagging Method Overview,* GK, Proc. Top2017

***H(bb) tagging in flavour talk!***

# Jet Grooming

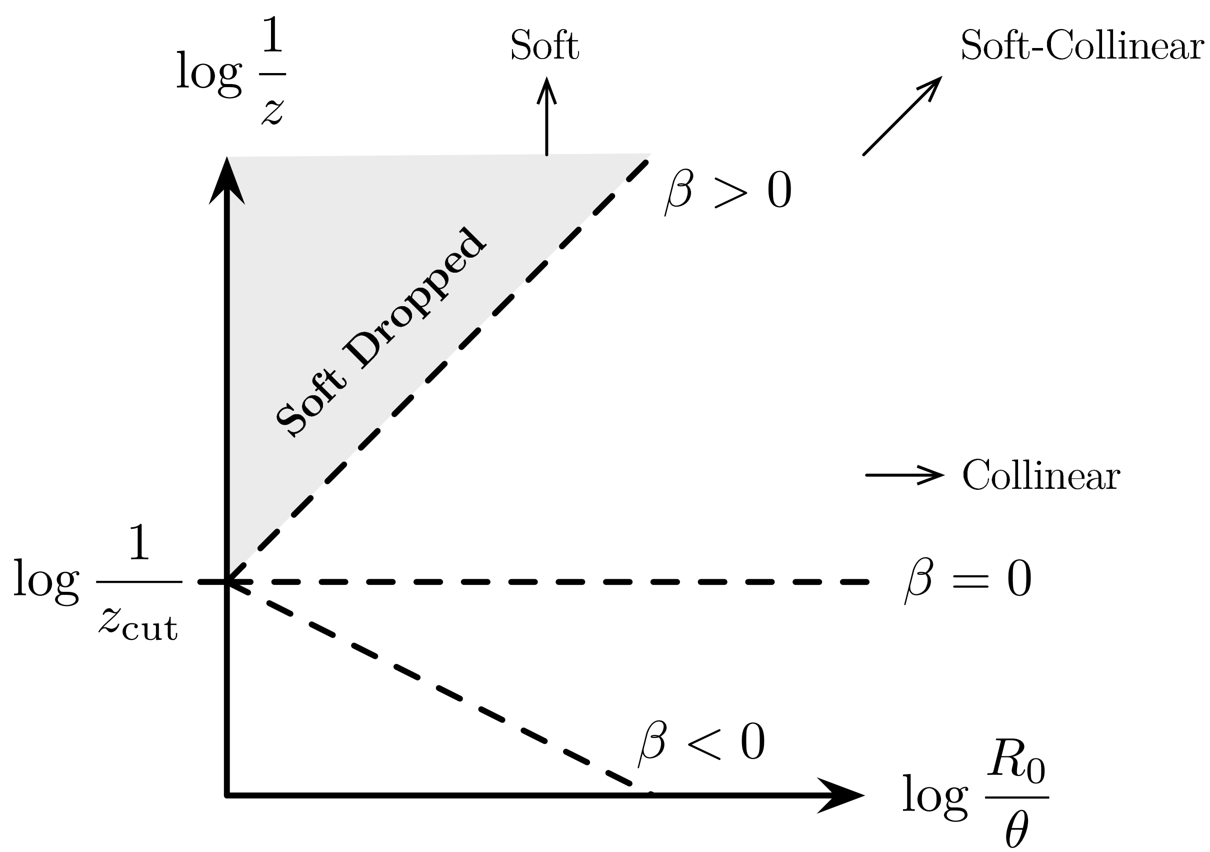


- Remove
  - soft radiation
  - underlying event
  - pile up
- from jet to access top mass



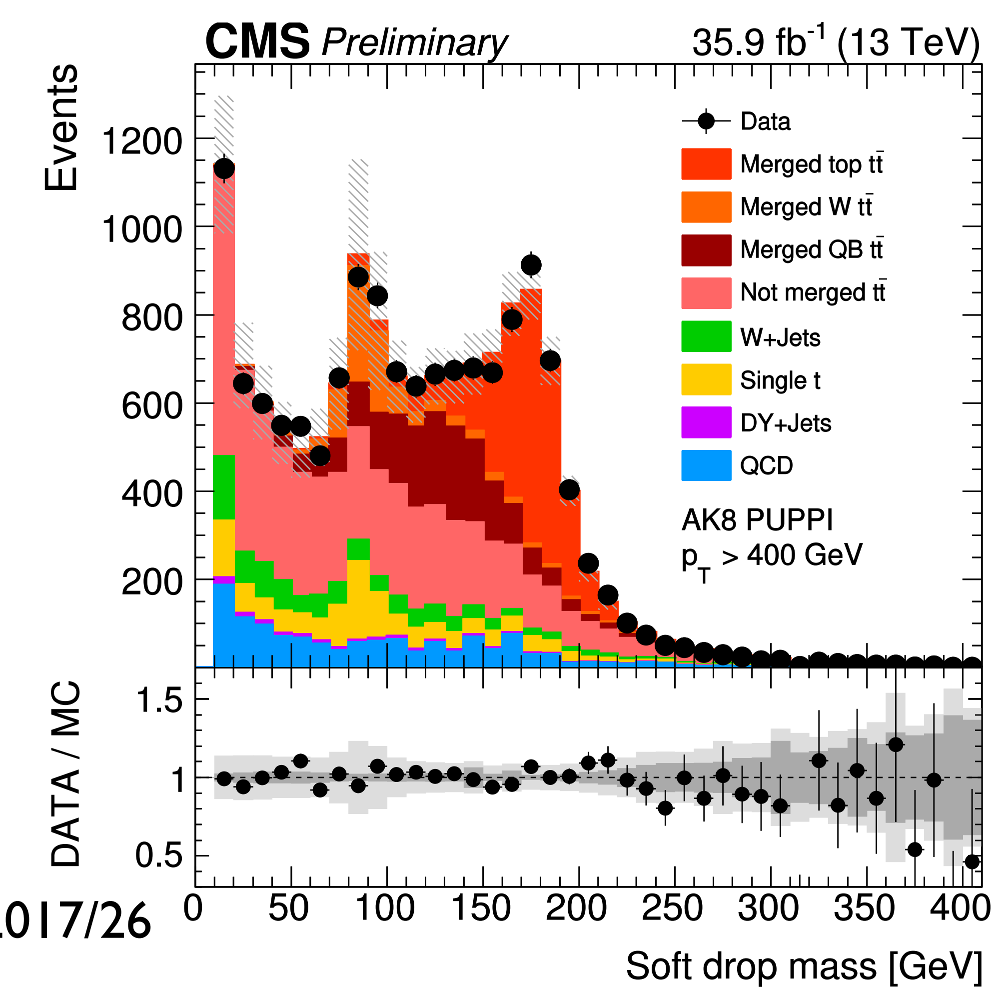
# mMDT
# Softdrop

*Towards an understanding of jet substructure*
M Dasgupta, A Fregoso, S Marzani, G Salam
JHEP 1309 029
Soft Drop
A Larkoski, S Marzani, G Soyez, J Thaler
JHEP 1405 146
*Factorization for groomed jet substructure beyond the next-to-leading logarithm*
C Frye, AJ Larkoski, MD Schwartz, K Yan
JHEP 1607 064




- Find hard substructure using step-wise unclustering
- No pure soft divergences
- Analytically calculable to high precision

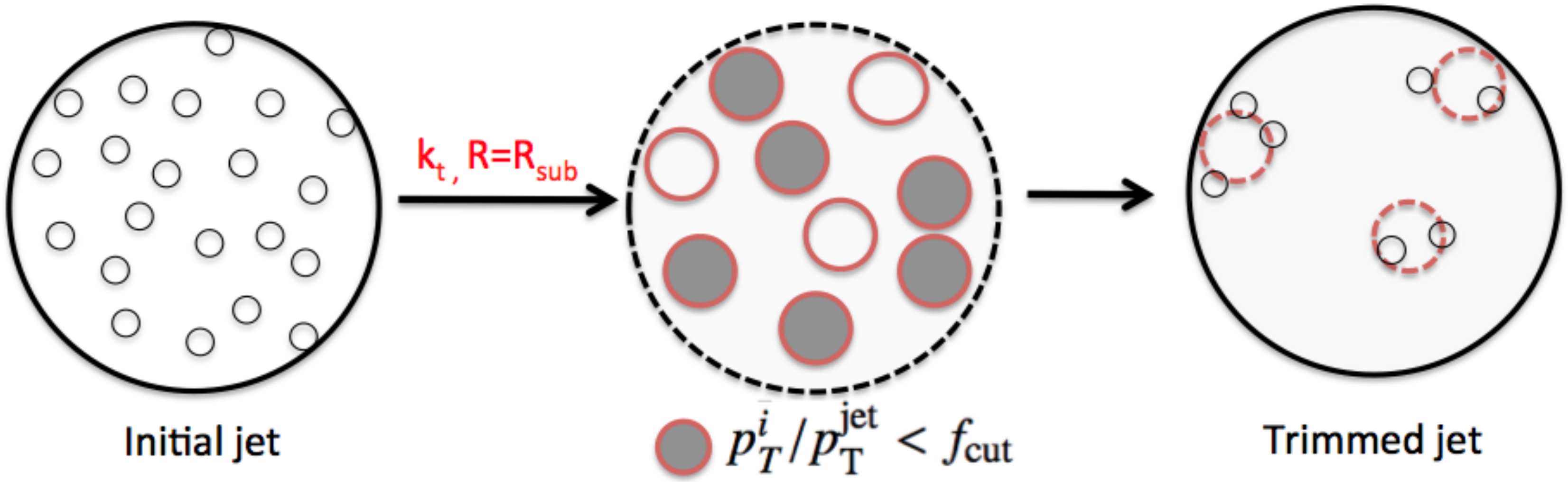
$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{\rm cut} \left( \frac{\Delta R_{12}}{R_0} \right)^{\beta}$$



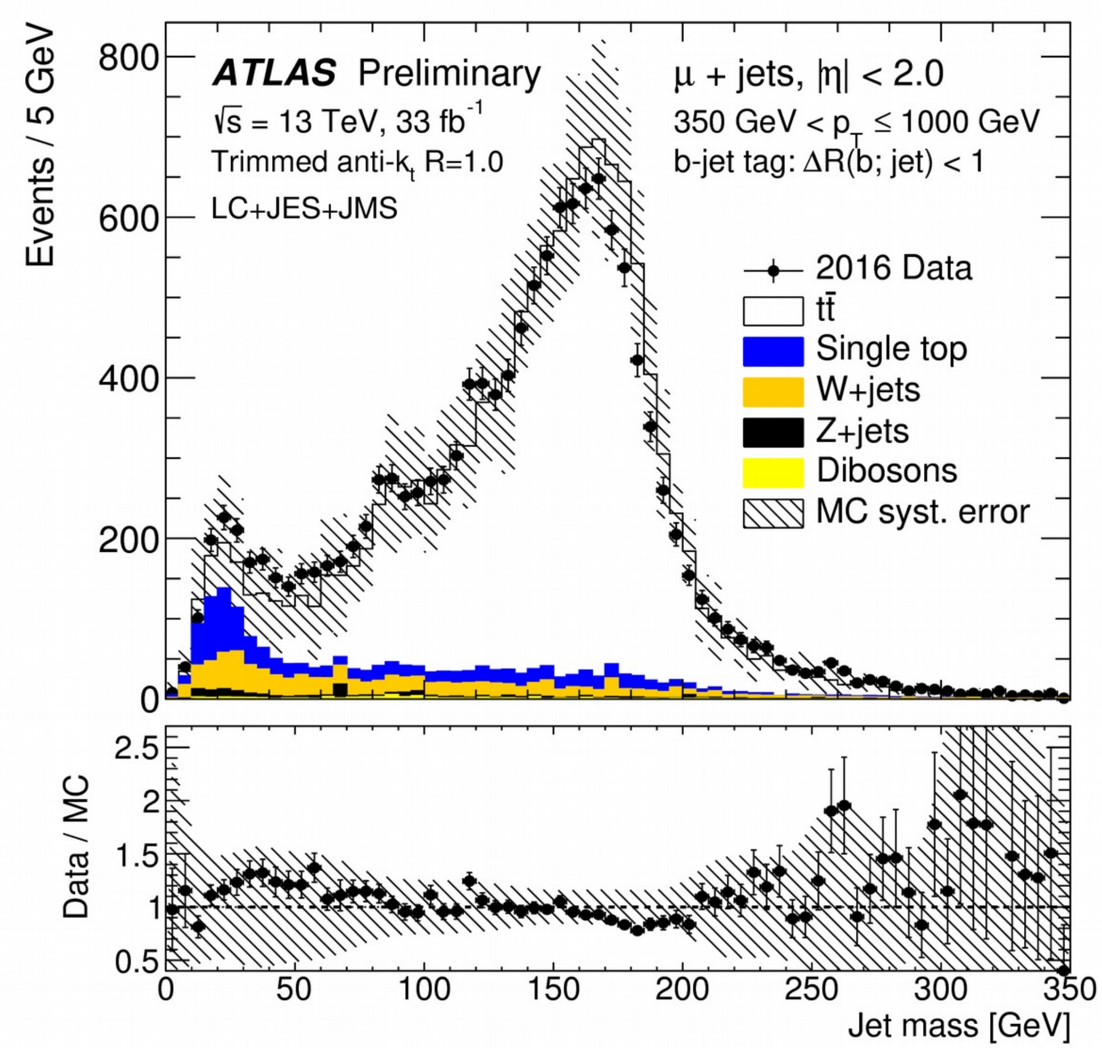

CMS DP-2017/26

# Trimming

*Jet Trimming*
D Krohn, J Thaler, LT Wang
JHEP 1002 084
*In-situ measurements of large-radius jet reconstruction performance*
ATLAS-CONF-2017-063




- Recluster constituents with R=0.2
- Remove subjets with less than 5% of jet $p_T$
- ATLAS Default



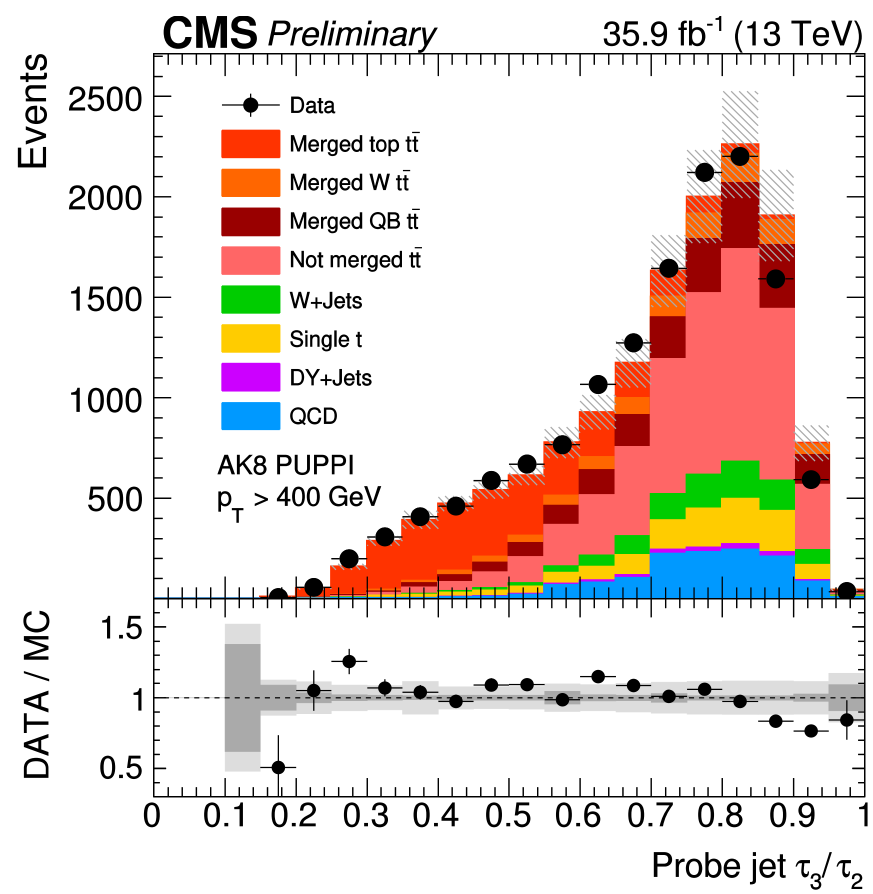
*Identifying Boosted Objects with N-subjettiness*
J Thaler, KV Tilburg,
JHEP 1103 015

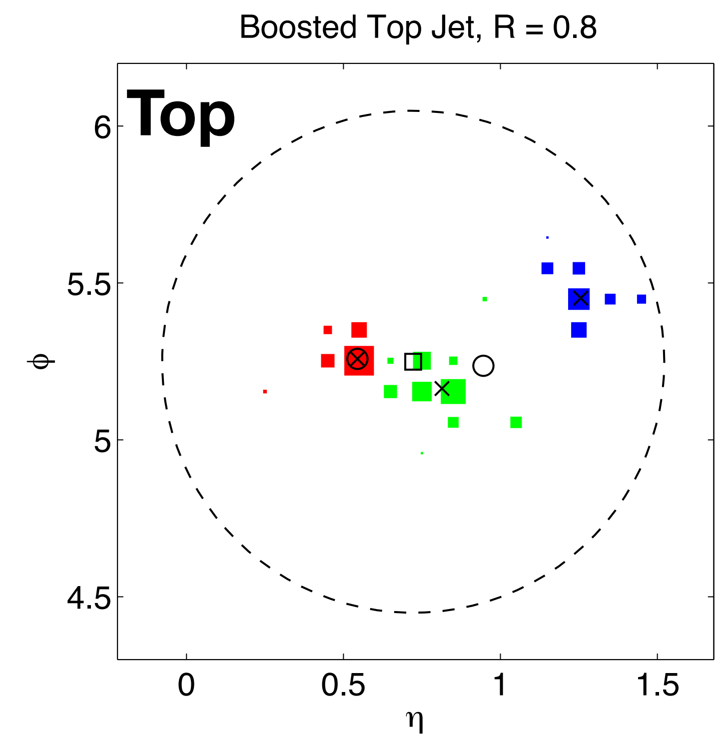
*Dichroic subjettiness ratios to distinguish colour flows in boosted boson tagging*
G Salam, L Schunk, G Soyez
JHEP 1703 022

# n-Subjettiness

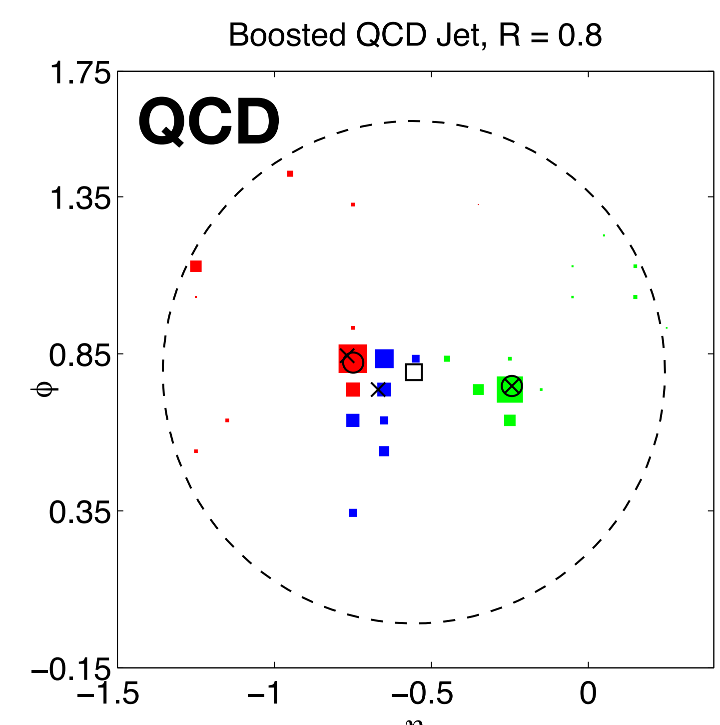
$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\}$$

- n-subjettiness: Small when compatible with n-prong substructure
- Used for top-tagging: $\frac{\tau_3}{\tau_2}$




(b)




- Recent ideas:
  - Dichroic n-subjettiness = ratio of n-subjettiness with different grooming (JHEP 1703 022)
  - Use for jet clustering (XCone: JHEP 1511 07)

*Energy Correlation Functions for Jet Substructure*
A Larkoski, GP Salam, J Thaler
*JHEP 1306 108*

*New Angles on Energy Correlation Functions*
I Moult, L Necib, J Thaler
JHEP 1612 153

# Energy Correlation Functions

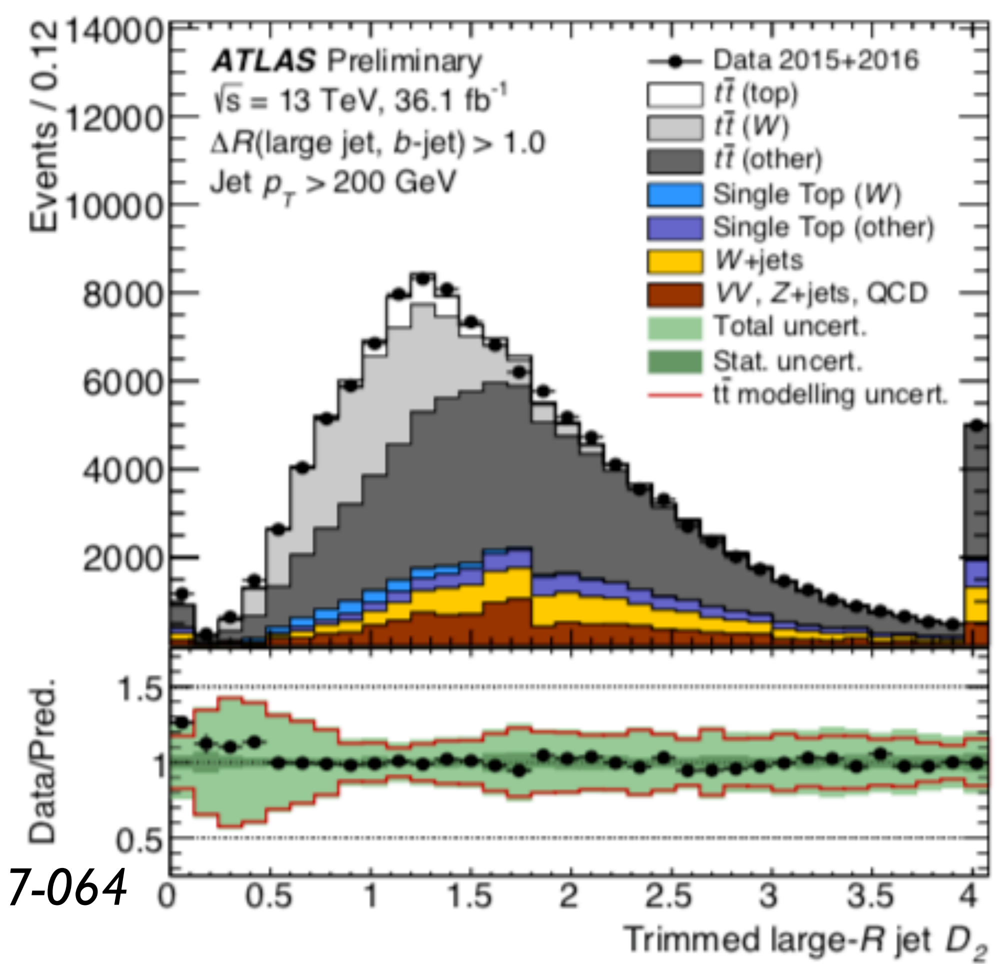
$$ {}_{v}e_{n}^{(\beta)} = \sum_{1\leq i_1<i_2<\cdots<i_n\leq n_J} z_{i_1} z_{i_2} \ldots z_{i_n} \prod_{m=1}^{v} \overset{(m)}{\min_{s<t\in\{i_1,i_2,\ldots,i_n\}}} \left\{\theta_{st}^{\beta}\right\} $$

angles

particles

$$ D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3} $$

- Energy correlation functions
- Replacing n-subjettiness
  for heavy-resonance identification
- Wide range of other uses



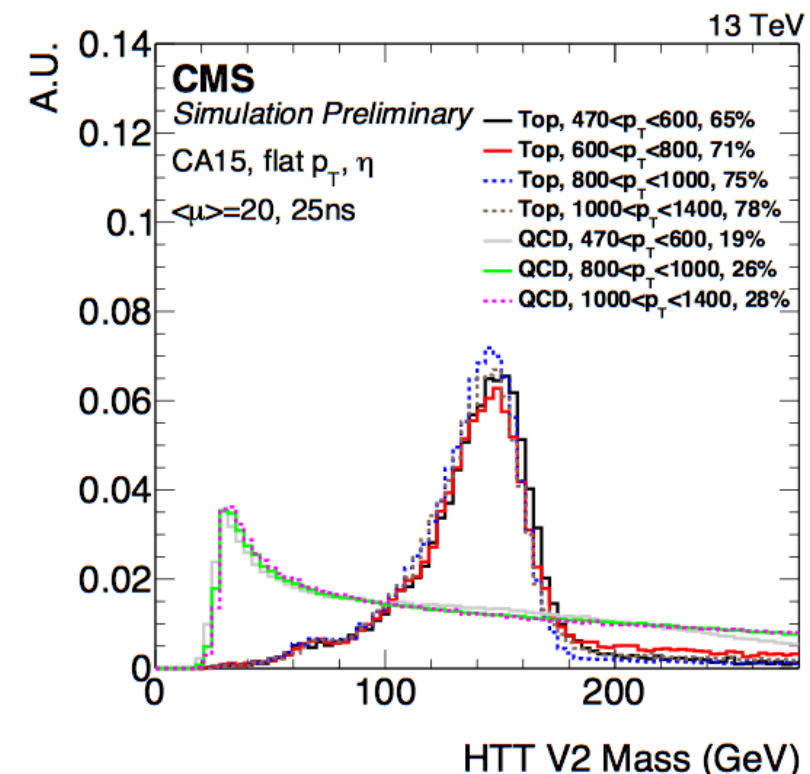

ATLAS-CONF-2017-064

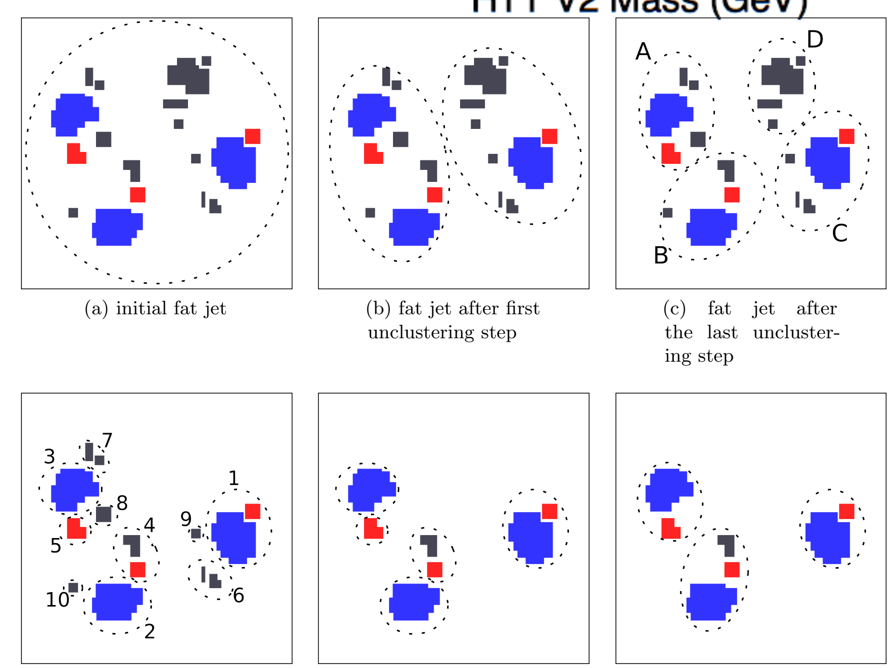
# Inclusive Taggers

# HEPTopTagger (V2)

*Resonance Searches with an Updated Top Tagger*
*GK, T Plehn, T Schell, T Strebler, GP. Salam*
*JHEP 1506 203*

- OptimalR-Algorithm:
  - Start with C/A, R=1.5 *seed* fat-jet
  - Perform unclustering to identify *small fat-jets* with R=0.5 to R=1.5 (in steps of 0.1) and run HEPTopTagger on each of them
  - Calculate: $R_{min}$ = Smallest cone size for which the mass differs by less than 20% from the mass at R=1.5
  - Get: $R_{opt, calc}$ ($p_T$). Result of fitting $R_{opt}$ as function of $p_T$ for signal jets
  - Output observables:
    - Top candidate mass: $m(R=R_{opt})$
    - W / top mass ratio: $f_W(R=R_{opt})$
    - $R_{opt}$ difference: $R_{opt}$ - $R_{opt, calc}$ ($p_T$)




(a) initial fat jet

(b) fat jet after first unclustering step

(c) fat jet after the last unclustering step

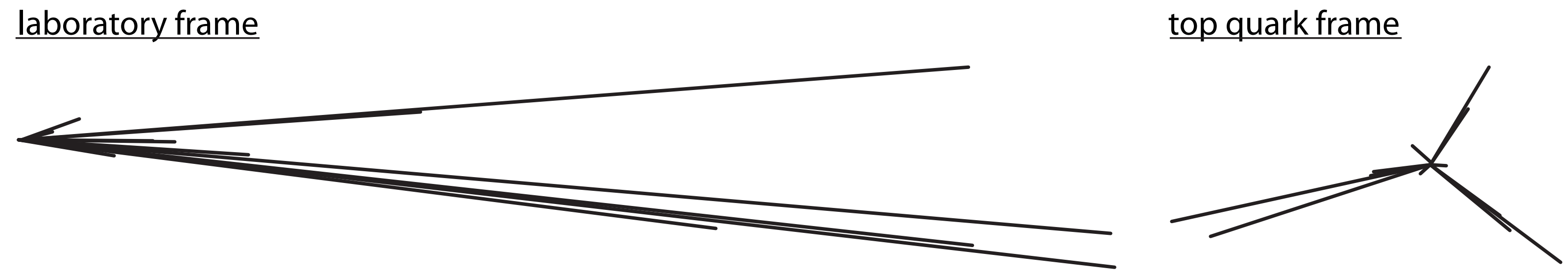
(d) re-clustered constituents of the triplet of jets A, B and C from the previous step using a smaller distance parameter

(e) the five leading subjets found in the previous step

(f) constituents of the five jets from the previous step re-clustered into exactly three jets for testing mass ratios

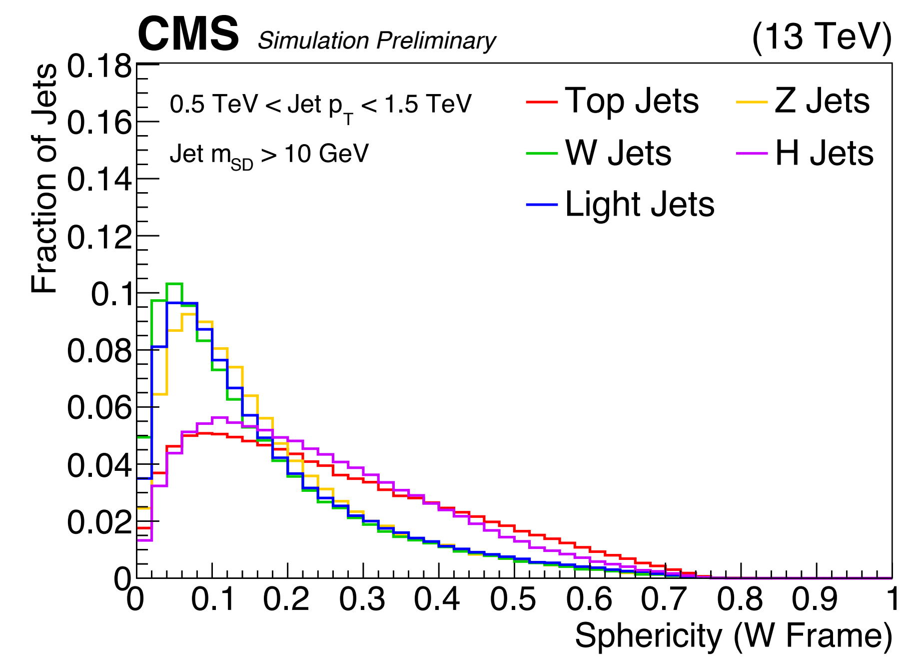
# BEST

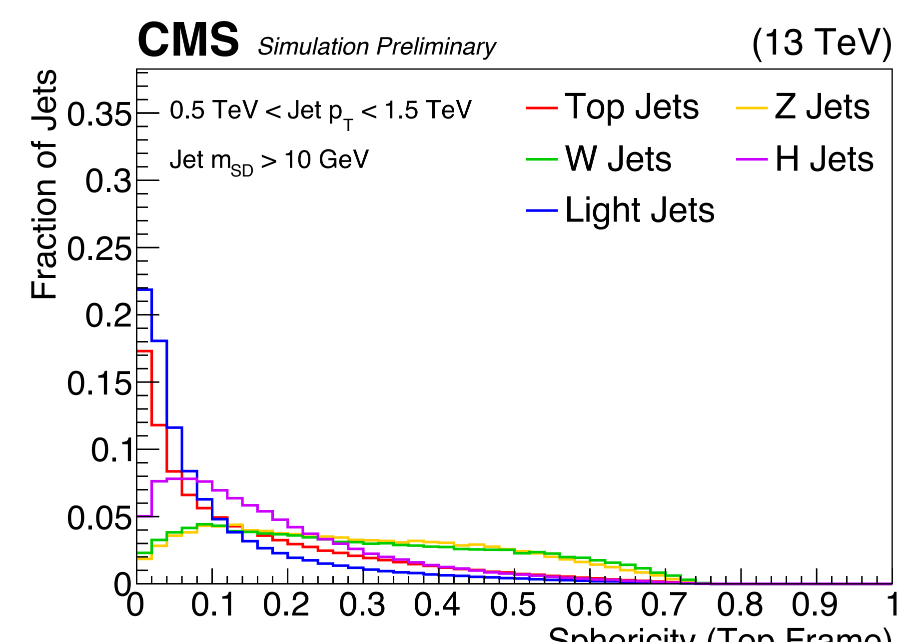
*JS Conway, R Bhaskar, RD Erbacher, J Pilot, Identification of High-Momentum Top Quarks, Higgs Bosons, and W and Z Bosons Using Boosted Event Shapes, PRD 94*




- Boosted Event Shape Tagger (BEST)
- Boost jet constituents individually into each reference frame corresponding to particle origin hypothesis (t, W, Z, H)
- Calculate angular distributions in boosted frame
- Fox-Wolfram Moments, Sphericity, Aplanarity, Isotropy, Thrust, ..
- Use NN for simultaneous classification

$$S^{\alpha,\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\vec{p}_i|^2}$$

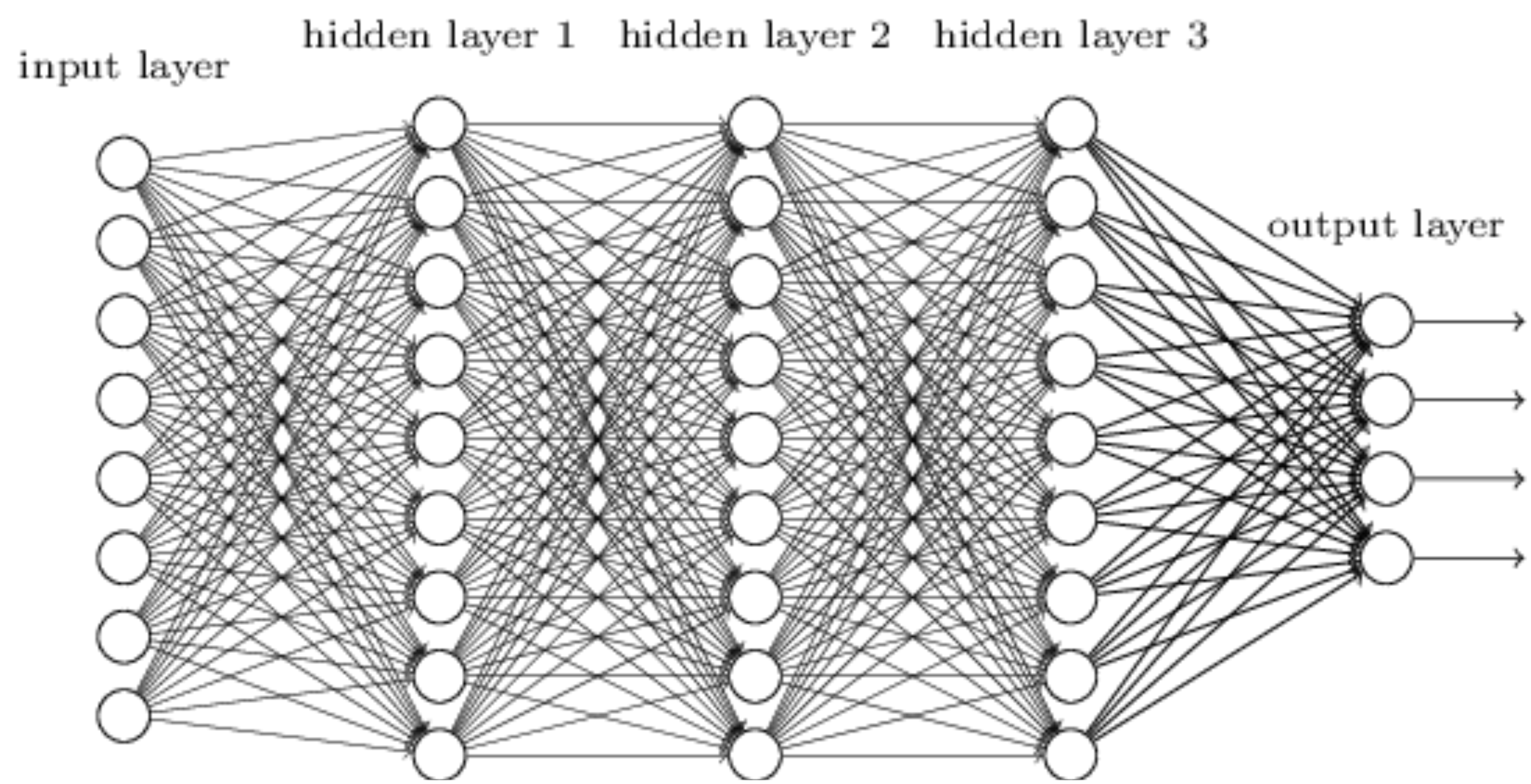



CMS DP-2017/27

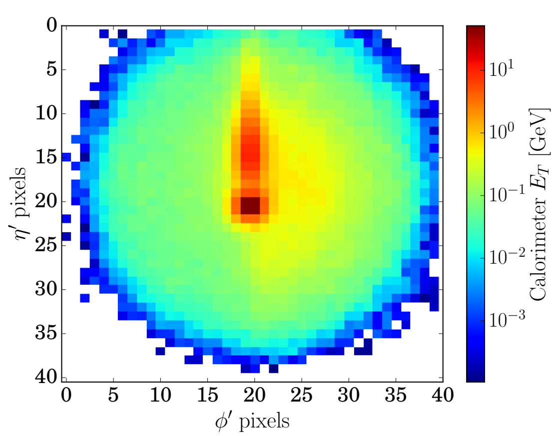
# Rise of the (tagging) machines

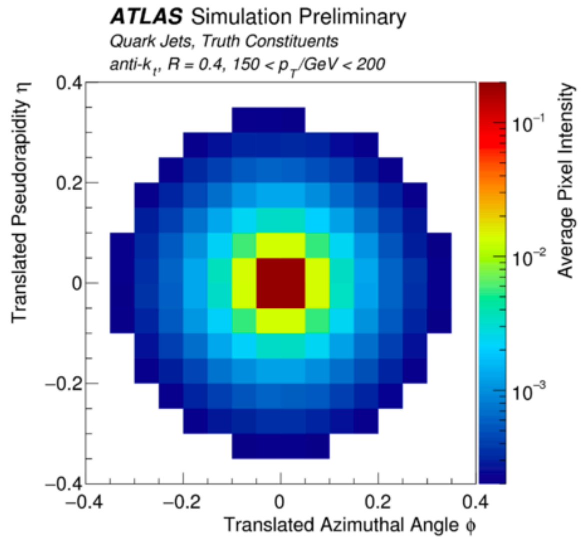
- *Use some representation of a jet (image, list of constituents,..) to train a deep neural network classifier on MC*

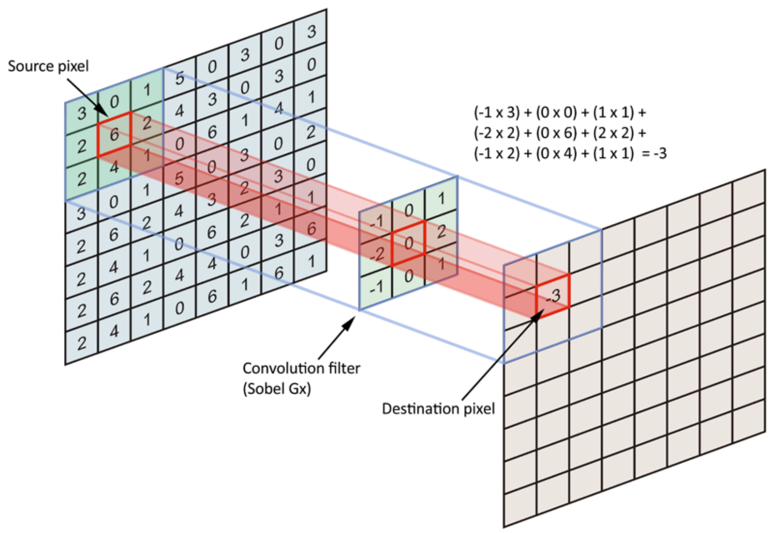



- *Powerful improvement of tagging performance. But will it help ttH/tHq?*

# Images

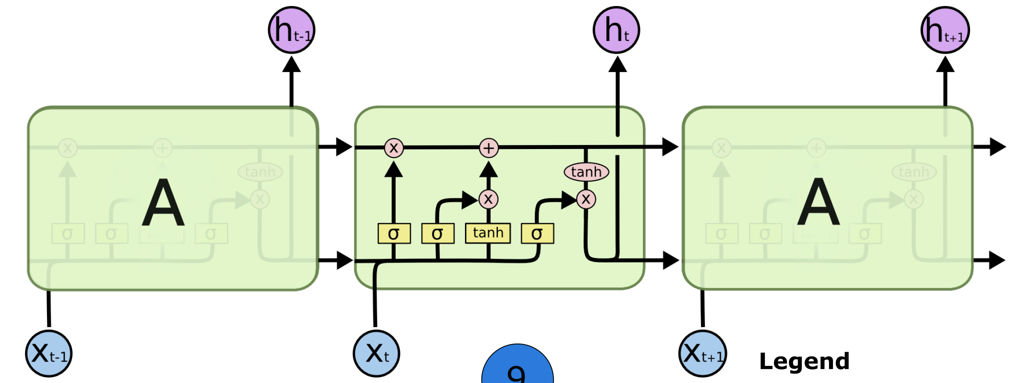


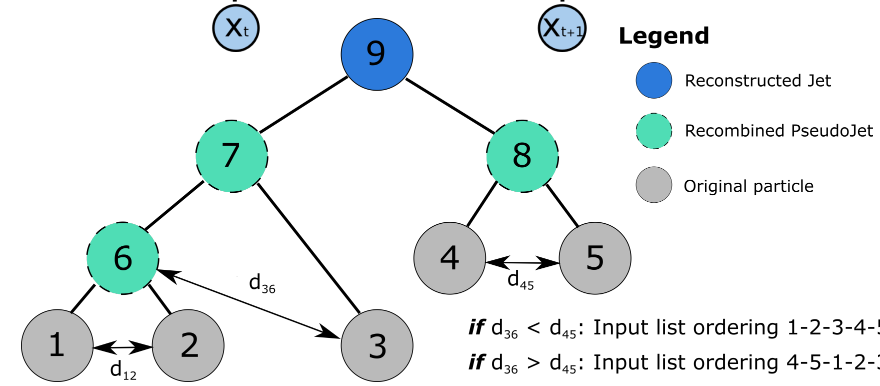




*Deep-learning Top Taggers or The End of QCD?*
*GK, Tilman Plehn, Michael Russell, Torben Schell*
*JHEP 05 (2017) 006*
*Deep learning in color: towards automated quark/gluon jet discrimination*
PT Komiske, EM Metodiev, MD Schwartz
JHEP 01 (2017) 110
*Jet-Images: Computer Vision Inspired Techniques for Jet Tagging*
J Cogan, M Kagan, E Strauss, A Schwartzman
*arXiv:1407.5675*
*Jet-Images – Deep Learning Edition*
Ld Oliveira, M Kagan, L Mackey, B Nachman, A Schwartzman
JHEP 1607 069
Quark and gluon tagging with Jet Images in ATLAS, ATL-PHYS-PUB-2017-017

# Recursive

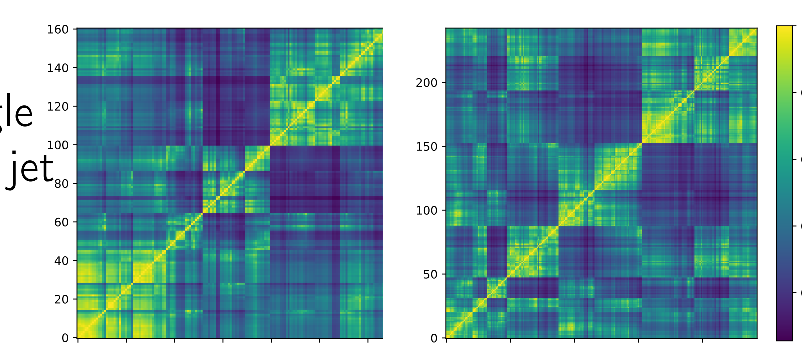



*Long Short-Term Memory (LSTM) networks with jet constituents for boosted top tagging at the LHC*
S Egan, W Fedorko, A Lister, J Pearkes, C Gay
arXiv: 1711.09059
*QCD-Aware Recursive Neural Networks for Jet Physics* G Louppe, K Cho, C Becot, K Cranmer
arXiv:1702.00748

# Other

Single QCD jet



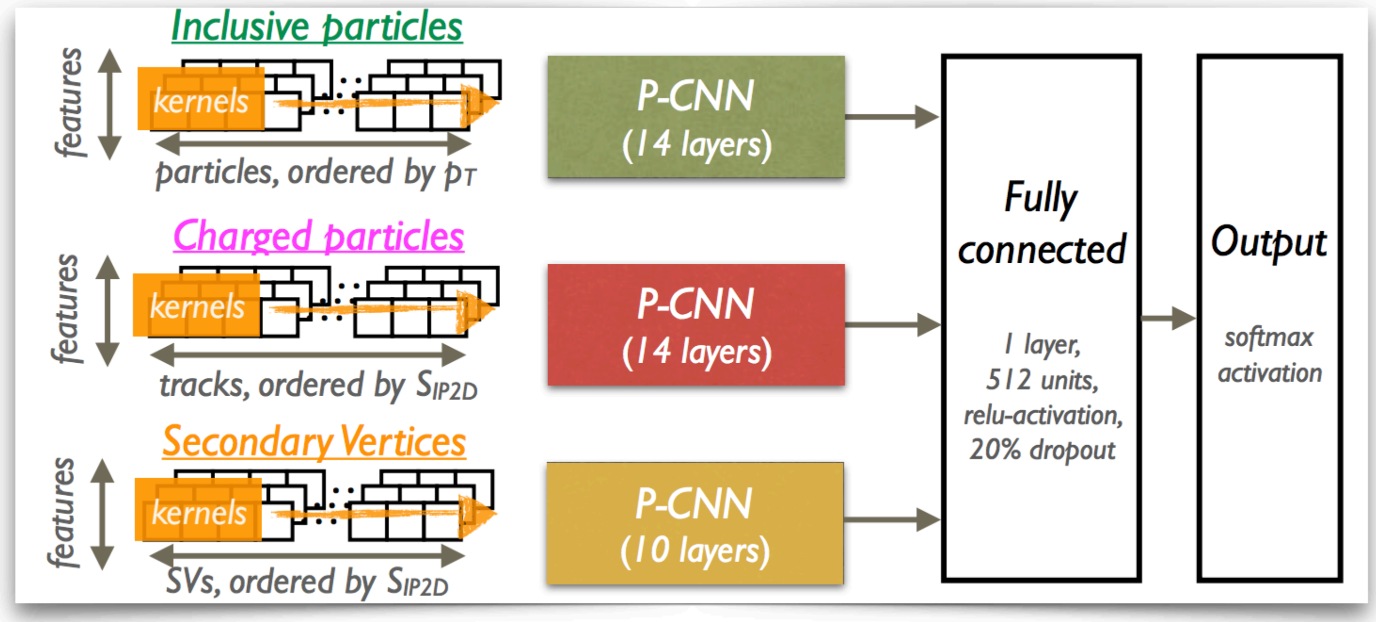
Single $W$ jet

*Neural Message Passing for Jet Physics I Henrion et al*
*Procs. of the Deep Learning for Physical Sciences Workshop at NIPS (2017)*
*Deep-learning Top Taggers & No End to QCD A Butter, GK, T Plehn, M Russell*
1707.08966

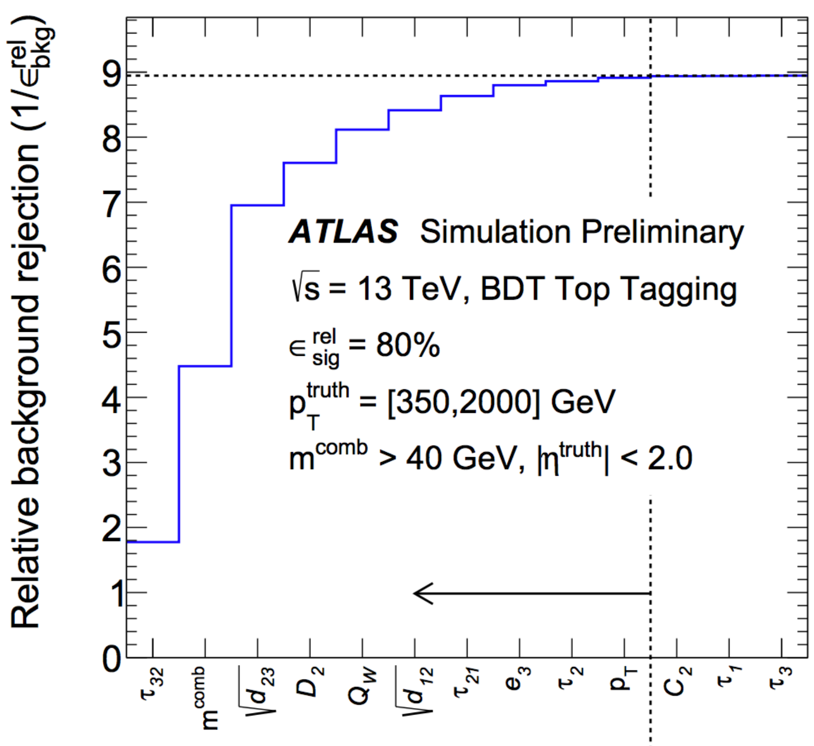
# Studies by ATLAS & CMS

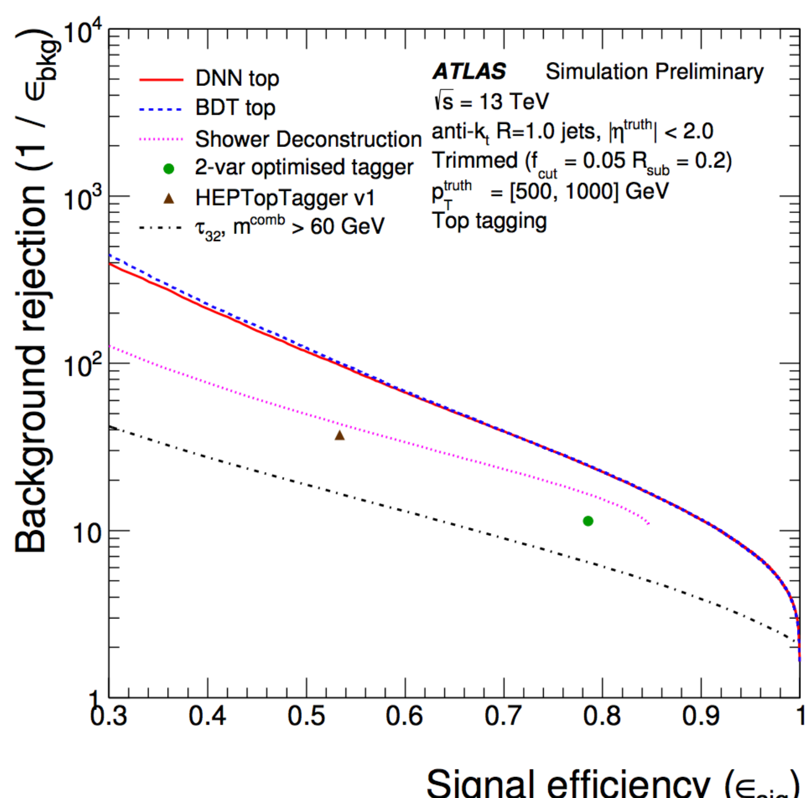
*ATLAS-CONF-2017-064*
*CMS DP 2017-049*

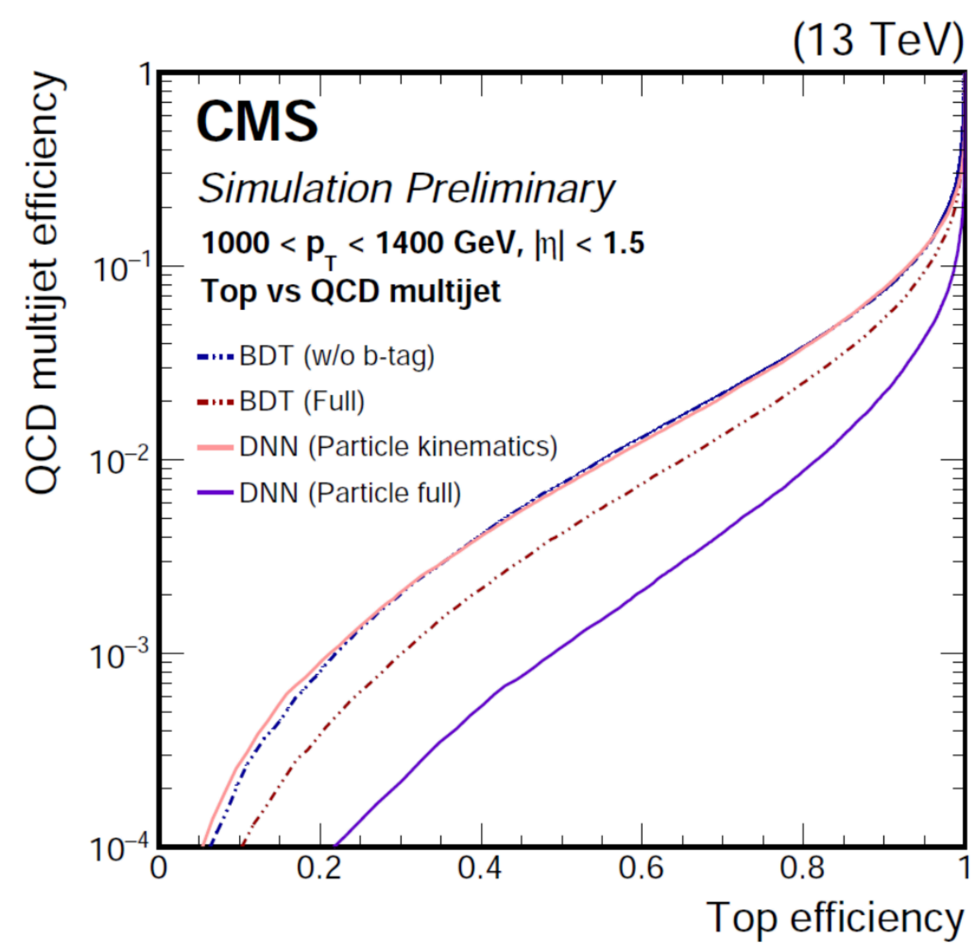
| Observable | Variable | Used For | Reference |
|---|---|---|---|
| Jet mass | $m^{\mathrm{comb}}$ | top, $W$ | [35] |
| Energy Correlation Ratios | $ECF_1$, $ECF_2$, $ECF_3$ <br> $C_2$, $D_2$ | top, $W$ | [41, 42] |
| N-subjettiness | $\tau_1$, $\tau_2$, $\tau_3$ <br> $\tau_{21}$, $\tau_{32}$ | top, $W$ | [43, 44] |
| Center of Mass Observables | Fox Wolfram ($R_2^{\mathrm{FW}}$) | $W$ | [45] |
| Splitting Measures | $Z_{\mathrm{CUT}}$ <br> $\sqrt{d_{12}}$, $\sqrt{d_{23}}$ | $W$ <br> top, $W$ | [46] <br> [47] |
| Planar Flow | $\mathcal{P}$ | $W$ | [48] |
| Angularity | $a_3$ | $W$ | [49] |
| Aplanarity | $A$ | $W$ | [50] |
| KtDR | $KtDR$ | $W$ | [51] |
| Qw | $Q_w$ | top | [46] |




*(DNN=fully connected)*

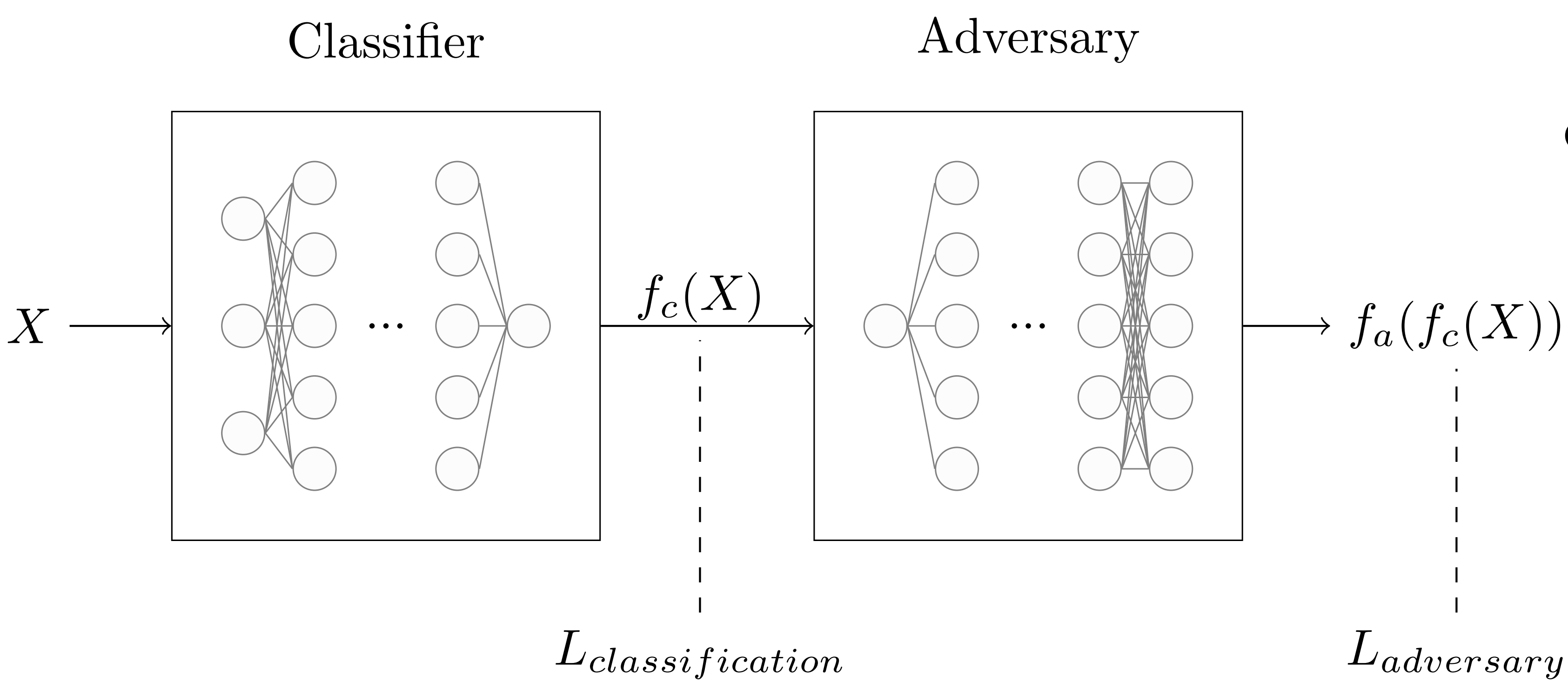


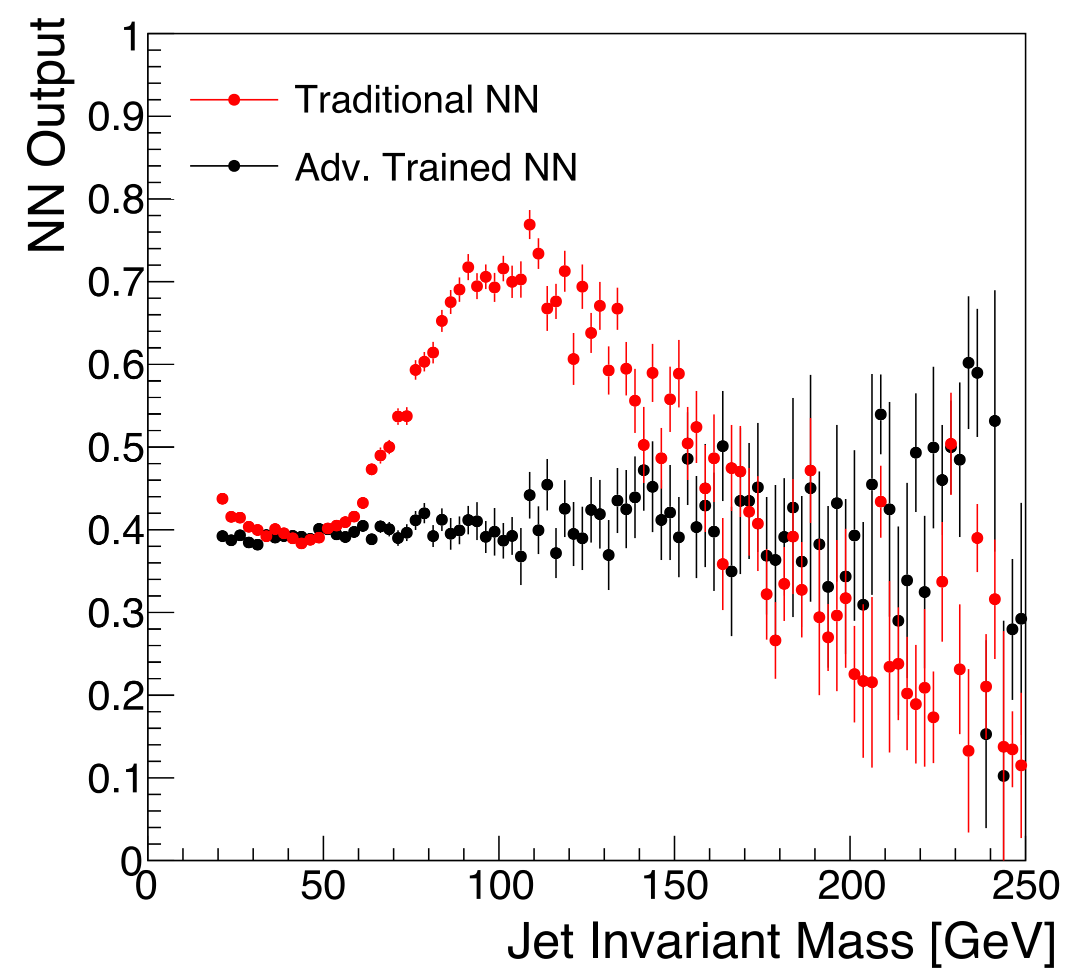
# Removing Correlations

*Decorrelated Jet Substructure Tagging using Adversarial Neural Networks*
C Shimmin, P Sadowski, P Baldi, E Weik, D Whiteson, E Goul, A Søgaard 1703.03507

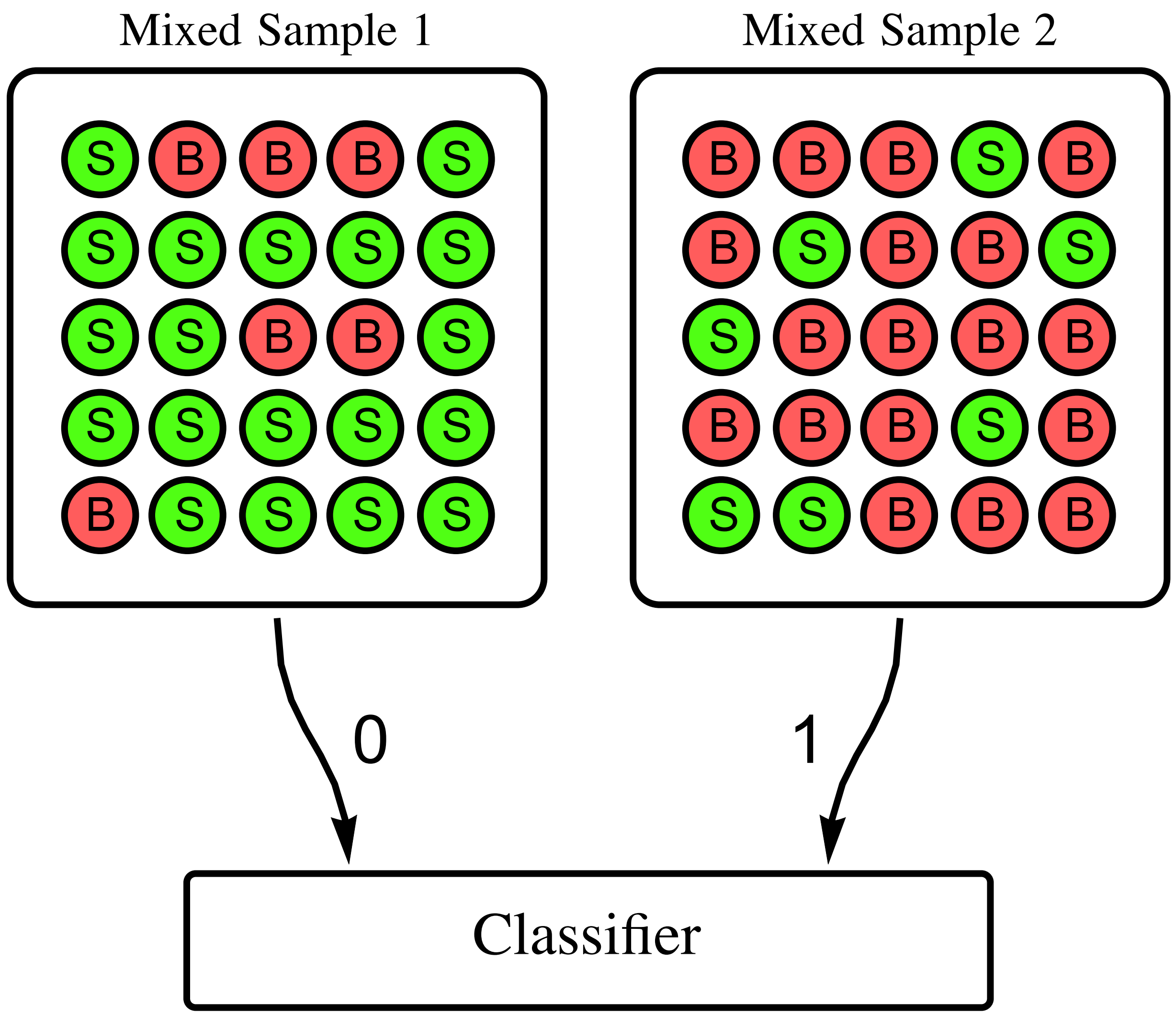


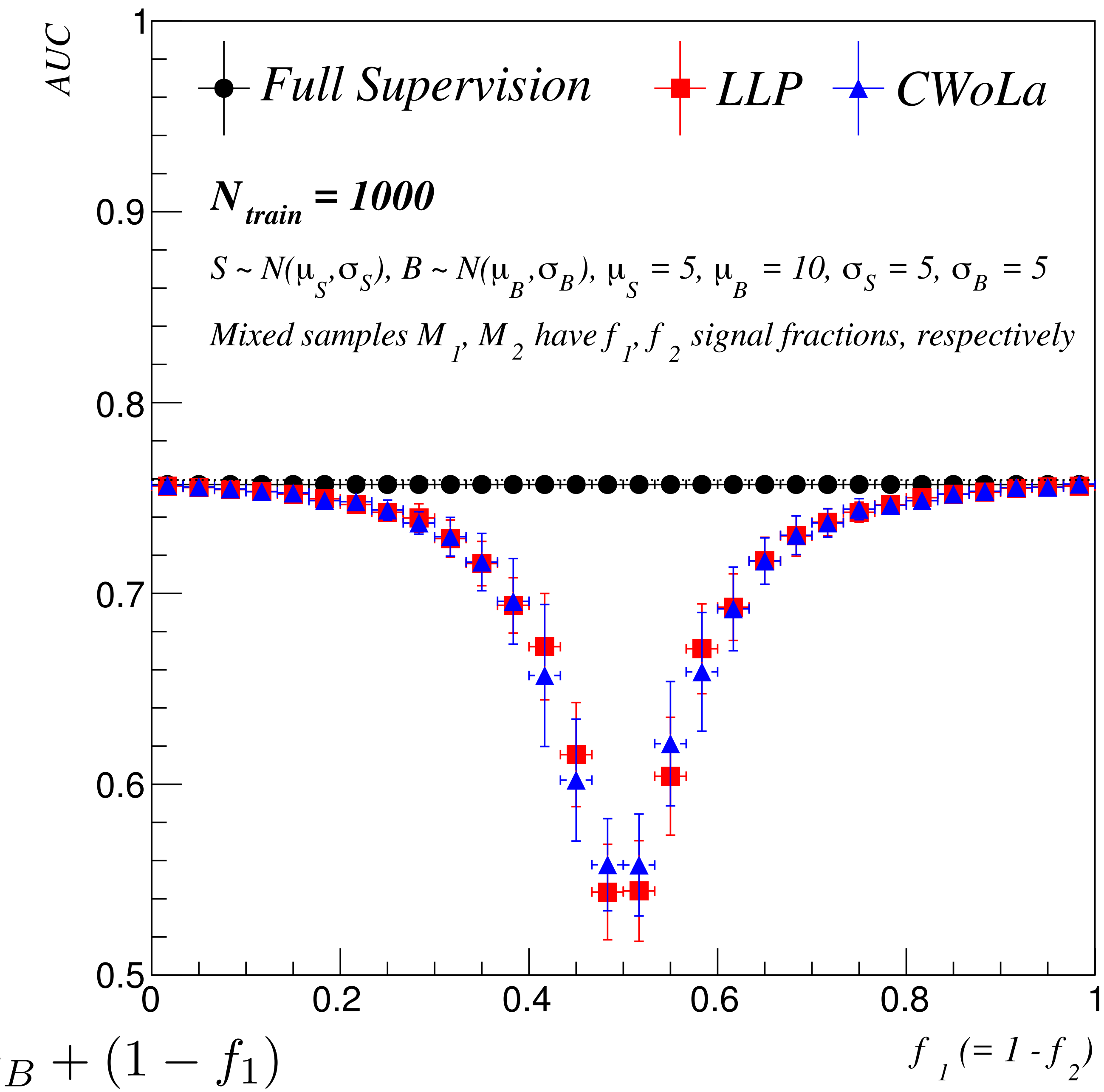

$$L_{\text{tagger}} = L_{\text{classification}} - \lambda L_{\text{adversary}}$$

- Classifier:
  - Distinguish Z' from QCD
- Adversary:
  - Infer jet mass
- **Trade-off discrimination power and stability**



# No Labels




$$L_{M_1/M_2} = \frac{p_{M_1}}{p_{M_2}} = \frac{f_1\, p_S + (1-f_1)\, p_B}{f_2\, p_S + (1-f_2)\, p_B} = \frac{f_1\, L_{S/B} + (1-f_1)}{f_2\, L_{S/B} + (1-f_2)}$$

*Distinguishing mixed samples is equivalent to signal/background classification!*

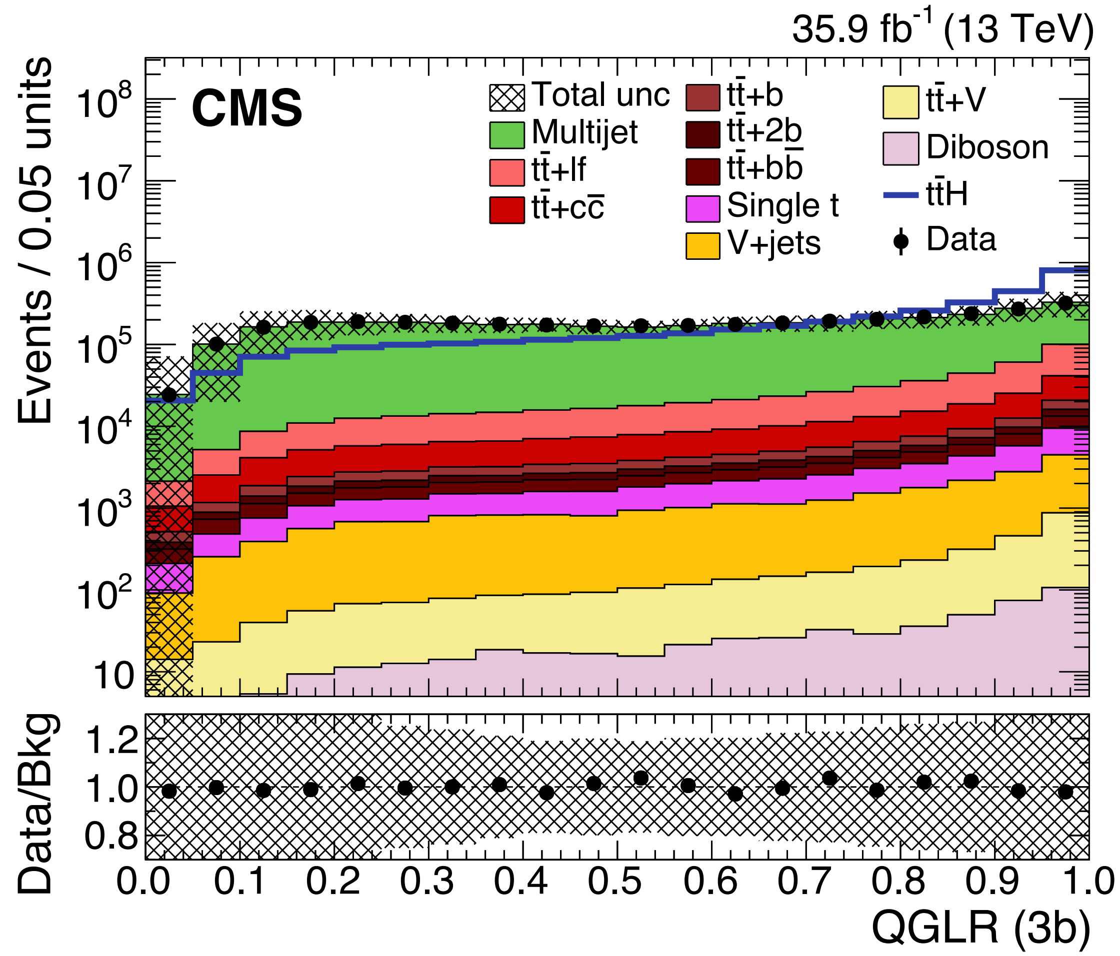
*Weakly Supervised Classification in High Energy Physics*
LM Dery, B Nachman, F Rubbo, A Schwartzman, 1702.00414
*Learning to Classify from Impure Samples*
PT Komiske, EM Metodiev, B Nachman, MD Schwartz, 1801.10158
*Classification without labels: Learning from mixed samples in high energy physics*, EM Metodiev, B Nachman, J Thaler, 1708.02949

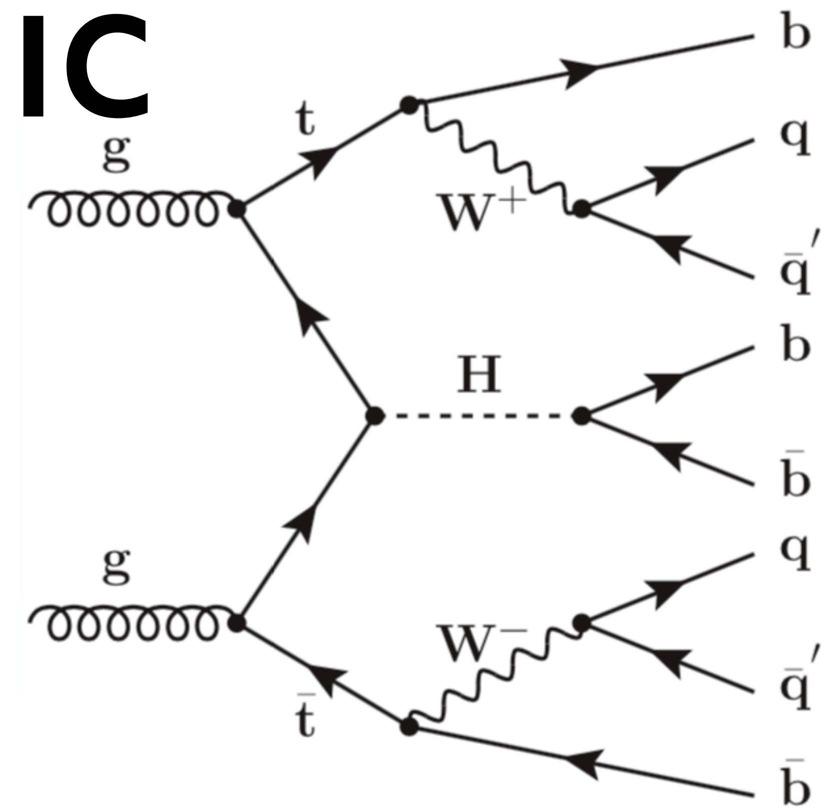
# Application

# ttH - Fully Hadronic

CMS-HIG-17-022
(last ATLAS result for ttH(bb) all hadronic was JHEP 1605 160
wo/ substructure)

| Category | |
|---|---|
| 7 jets, | 3 b jets |
| 8 jets, | 3 b jets |
| $\geq$9 jets, | 3 b jets |
| 7 jets, | $\geq$4 b jets |
| 8 jets, | $\geq$4 b jets |
| $\geq$9 jets, | $\geq$4 b jets |




**Signal Region: At least three bs + high QG likelihood**

- per-jet quark/gluon discrimination
  - Track multiplicity
  - pT distribution of constituents
  - Spatial profile

$$L(N_q, N_g) = \sum_{\text{perm}} \left( \prod_{k=i_1}^{i_{N_q}} f_q(\zeta_k) \prod_{m=i_{N_q+1}}^{i_{N_q+N_g}} f_g(\zeta_m) \right)$$

*Event based likelihood*

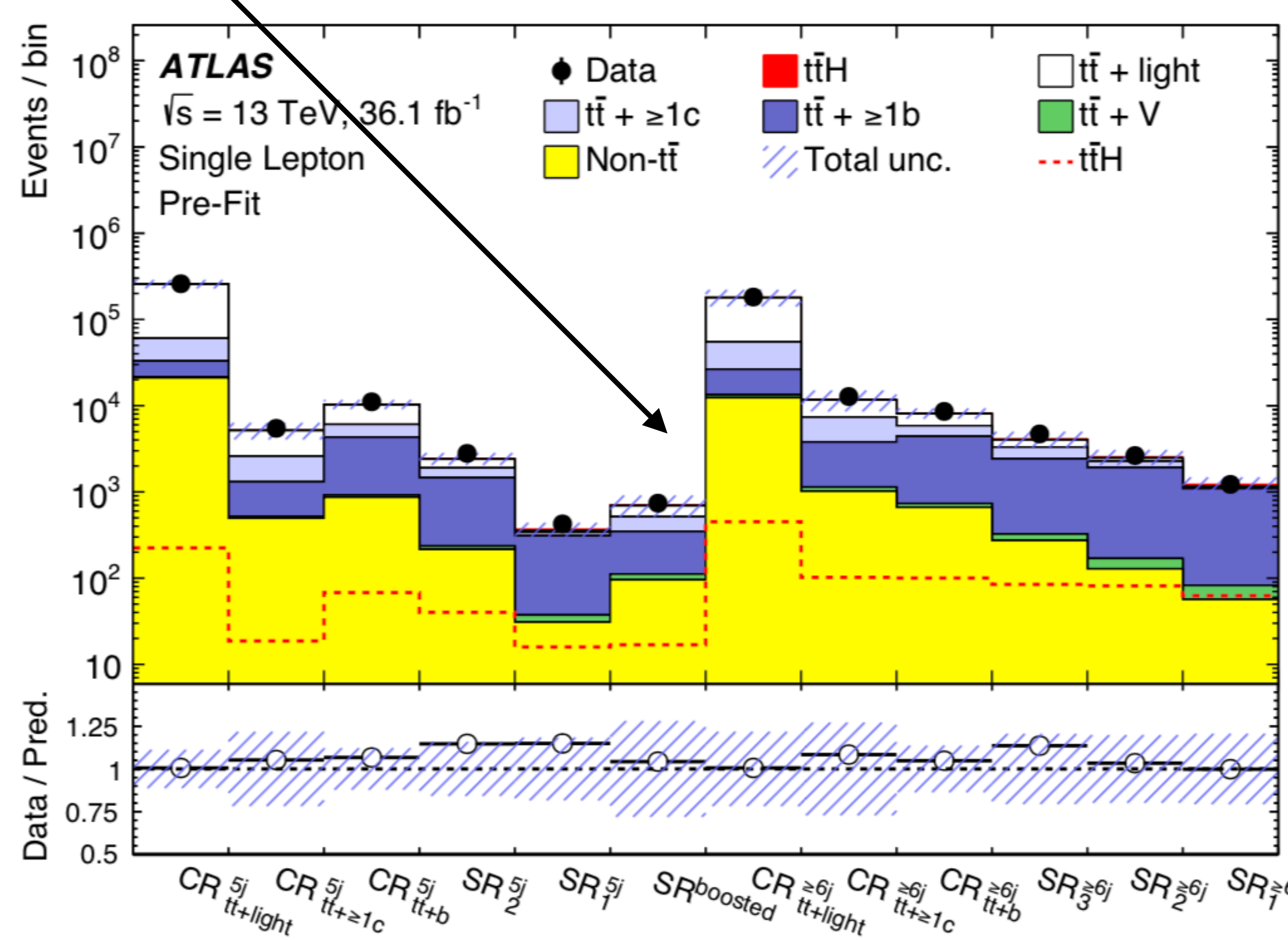
# ttH - with Leptons (ATLAS)

arXiv: 1712.08895

- Start with standard jets:
  - topoclusters, R=0.4
  - pT > 25 GeV, |eta| < 2.5, jet vertex tag (JVT)
  - b-tag using MV2c10
- Boosted reconstruction:
  - Re-cluster **jets** with Anti-Kt R=1.0
    - Remove if mass < 50 GeV
  - Look for Higgs candidates
    - pT > 200 GeV, at least 2 two b-jets
    - Tie breaker:
      Choose highest sum of b-tag scores
  - Look for top candidates in remaining jets
    - pT > 250 GeV, exactly one b-tagged jet, at least one non-tagged
    - Tie breaker: Choose highest mass

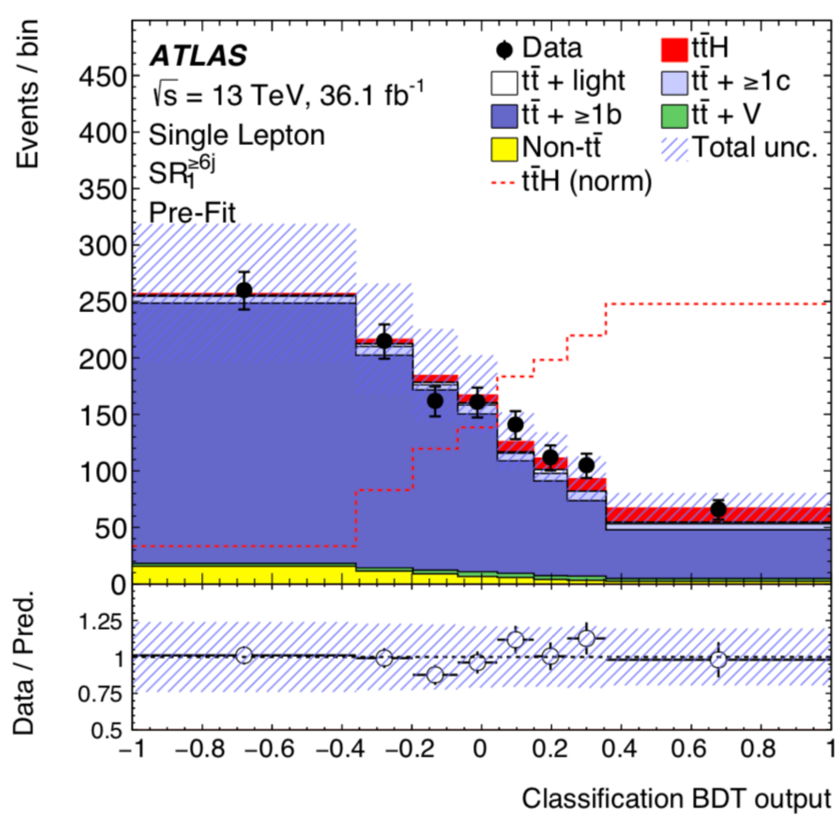
- Boosted **category**:
  - Single lepton
  - (at least) one Higgs & one Top
  - one additional b (outside Higgs/Top)

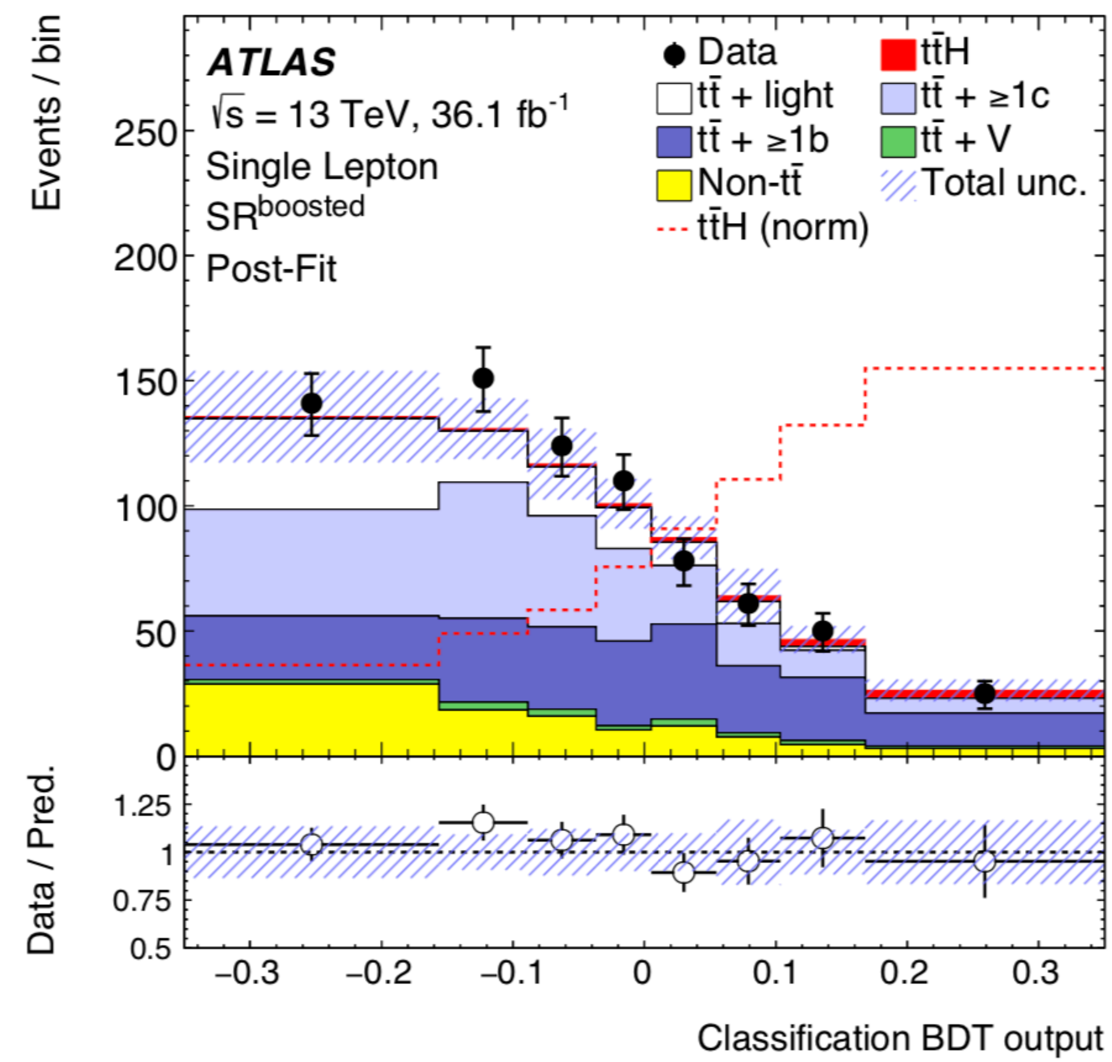
*Boosted*



# Classification BDT

| Variable | Definition |
|---|---|
| Variables from jet reclustering | |
| $\Delta R_{H,t}$ | $\Delta R$ between the Higgs-boson and top-quark candidates |
| $\Delta R_{t,b^{\text{add}}}$ | $\Delta R$ between the top-quark candidate and additional $b$-jet |
| $\Delta R_{H,b^{\text{add}}}$ | $\Delta R$ between the Higgs-boson candidate and additional $b$-jet |
| $\Delta R_{H,\ell}$ | $\Delta R$ between the Higgs-boson candidate and lepton |
| $m_{\text{Higgs candidate}}$ | Higgs-boson candidate mass |
| $\sqrt{d_{12}}$ | Top-quark candidate first splitting scale [100] |
| Variables from $b$-tagging | |
| $w_{b\text{-tag}}$ | Sum of $b$-tagging discriminants of all $b$-jets |
| $w^{\text{add}}_{b\text{-tag}}/w_{b\text{-tag}}$ | Ratio of sum of $b$-tagging discriminants of additional $b$-jets to all $b$-jets |



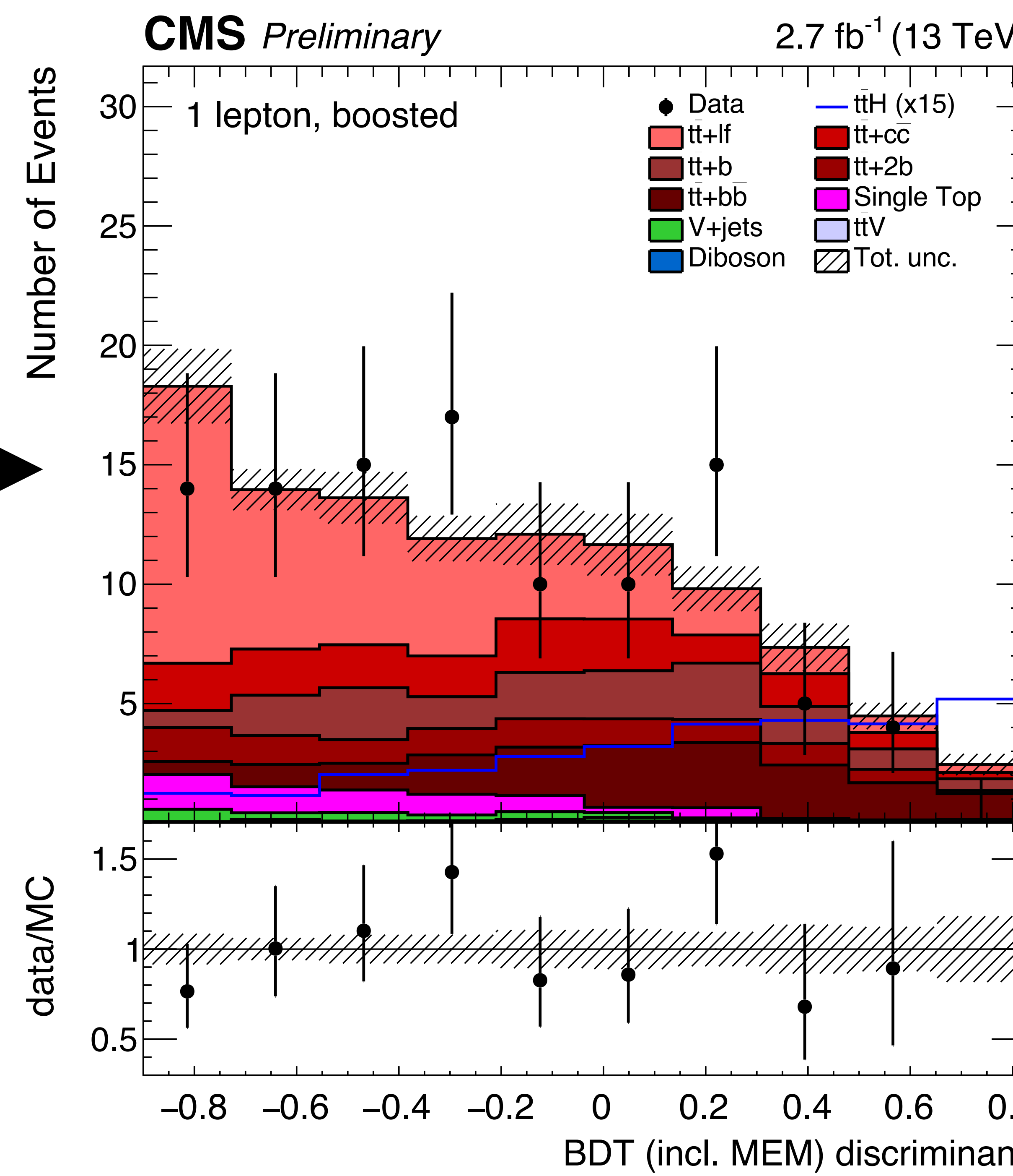



# ttH - with Leptons (CMS)

HIG-16-004
(2.7 fb-1, 2015 data)

- Independent clustering (CA, R=1.5)
  - HEPTopTagger for top tagging
  - Subjet-filtering for Higgs

| $\geq 4$ jets, $\geq 2$ b-tags boosted |
| --- |
| avg $\Delta R$(tag,tag) |
| $\tau_2/\tau_1$ of Higgs cand. |
| third-highest CSV |
| fourth-highest CSV |
| $\Delta\eta$(top,Higgs) |
| aplanarity |
| m(Higgs, di-filterjet) |
| min $\Delta R$(tag,tag) |
| avg CSV (all) |
| MEM discriminator (using subjets) |
| b-tagging likelihood ratio |



# CMS/ATLAS contrast

- ATLAS:
  - Building AK10 Top/Higgs candidates out of AK4 jets
- CMS
  - separate clustering into CA15

- Reclustering allows *re-cycling* jet energy corrections, simplify analysis
- Potentially higher reach of separate clustering for very boosted events

# Summary and Conclusions

- Boosted jet substructure methods have become a *default* analysis tool
- Also used (sparingly) in Higgs/Top analyses
  - Will become more important for differential boosted measurements
- Interplay with deep learning progress

*Thank you!*