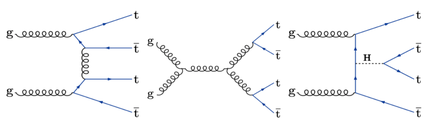
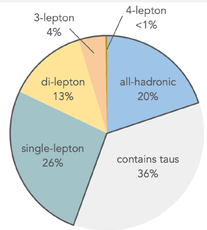
Standard model four-top quark production in the all-hadronic final state at 13 TeV with the CMS experiment

Melissa Quinnan on behalf of the CMS Collaboration

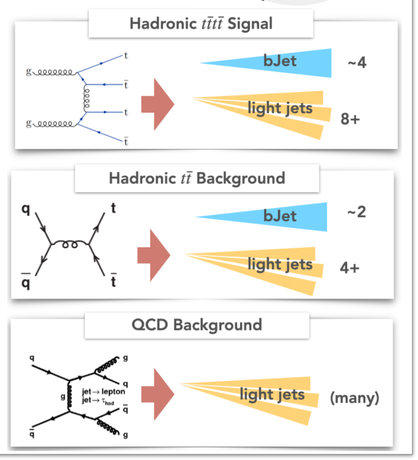
UC San Diego · UCSB · Carnegie Mellon University · Korea University

Introduction



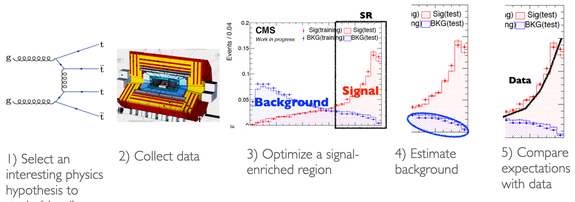


- Rare SM process with cross section of 12fb at 13 TeV
- Enhancement of cross section predicted by some BSM physics (ex 2HDMs)
- ATLAS Run II multi-channel combined results: observed (expected) significance of 4.7(2.6)[1]
- CMS Run II results for same-sign dilepton channel: observed (expected) significance of 2.6(2.7)[2]
- First all-hadronic analysis
- Included in Run II CMS result in multiple final states[3]



Analysis Strategy

- BDT discriminates signal vs. background
- Maximum likelihood fit to BDT in signal region (SR)

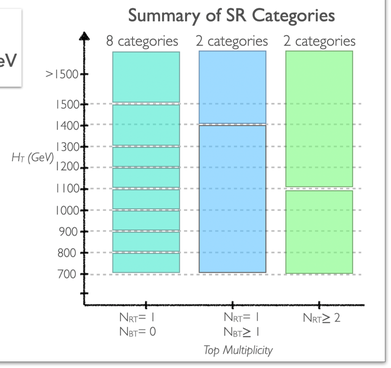


Do results agree with SM predictions?

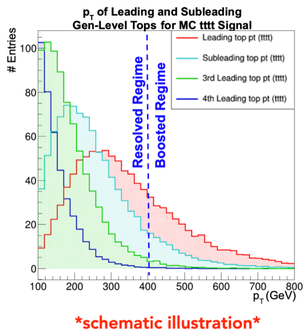
Event Selection

Baseline Selection:
N(leptons)=0, N(Jets)≥9, N(bJets)≥3, H_T≥700 GeV

- Events selected to reduce major ttbar+QCD backgrounds
- 12 SR categories split by boosted/resolved top multiplicity (N_{BT}/N_{RT}) and H_T (scalar sum of jet transverse momentum)

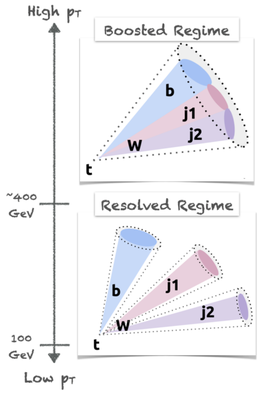


Top Tagging



schematic illustration

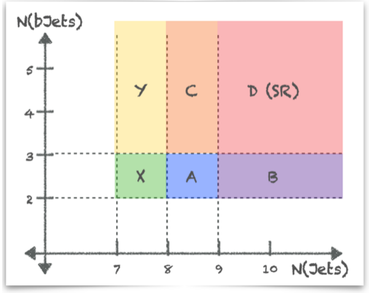
- Define both boosted and resolved hadronic top decays
- Boosted tops use standard DeepAK8[4] tagger
- Resolved tops use custom tagger
- BDT tagger based on 2016 stop analysis[5]



- Significantly reduces computing resources

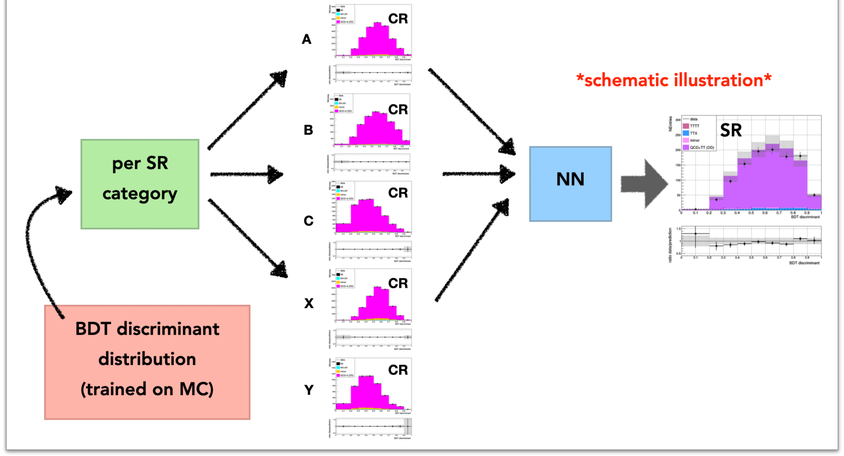
Background Estimation

- BDT distribution of ttX (ttH+ttW+ttZ) and other minor backgrounds from simulation
- Major backgrounds (QCD+ttbar) use data-driven (DD) methods
- DD yields estimated from data yields in 5 control regions (CRs) using extended ABCD method[6]
- Shape of DD BDT from NN (neural autoregressive flow[7])
- NN transforms input BDT histograms (ttbar MC) to match target (data) trained on same 5 CRs
- Predicted distributions normalized to extended ABCD yields

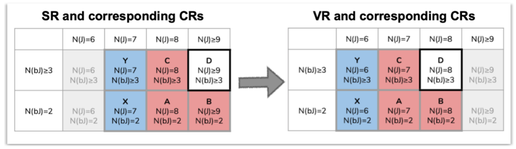


$$D(SR) \approx \left(\frac{B \cdot C}{A}\right)\left(\frac{C \cdot X}{A \cdot Y}\right)$$

ABCD term · higher order term

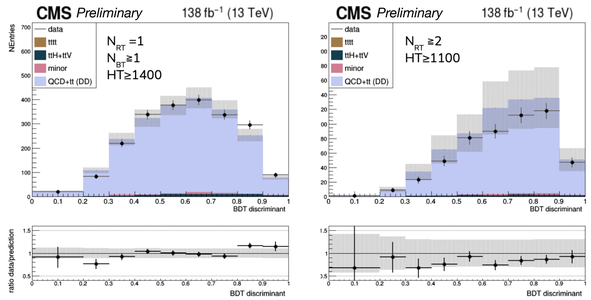


Background Validation & Systematics



- Validate DD background in validation region (VR)
- Normalization and shape uncertainties based on disagreement in VR

- VR uncertainties applied to corresponding SR categories
- Main systematics statistics and VR uncertainties
- No theoretical uncertainties on modeling of ttbar

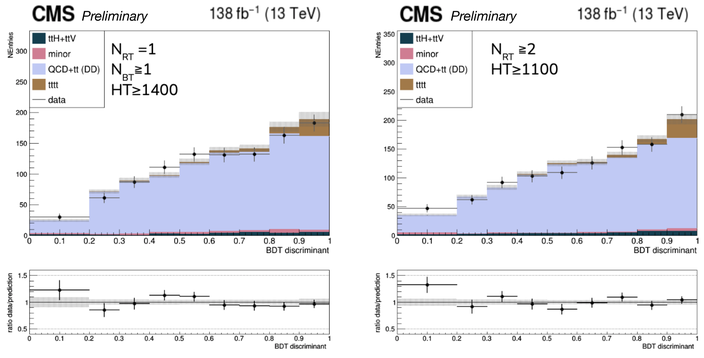


BDT discriminant distributions in most sensitive VR bins in Run II

Results

- Observe large signal excess, albeit with large uncertainties
- First all-hadronic four-top analysis
- Showcases novel ML tools that are potentially applicable to other analyses
- Motivates future SM and BSM four-top analyses

observed (expected) significance:
2.46(0.42)



Example post-fit S+B BDT discriminant distributions in most sensitive SR bins in Run II

References

1. ATLAS Collaboration. arXiv:2007.14858
2. CMS Collaboration. arXiv:1908.06463
3. CMS Collaboration. TOP-21-005
4. CMS Collaboration. CMS-PAS-JME-18-002
5. CMS Collaboration. arXiv:1707.03316
6. S. Choi. arXiv:1906.1083, arXiv:2008.0363
7. Huang, Krueger, Lacoste, Courville. *Neural Autoregressive Flows.* arXiv:1804.00779

melissa.quinnan@cern.ch