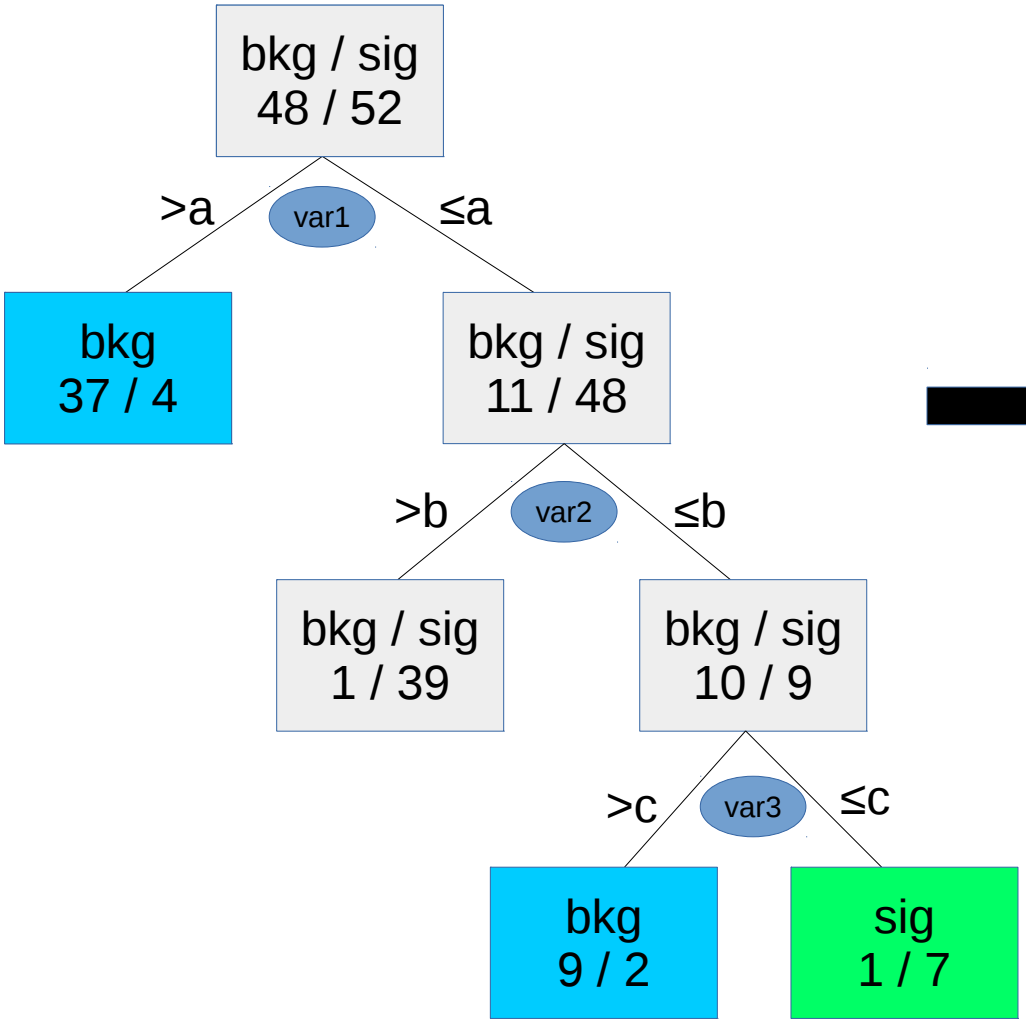
QCD Rejection in fully hadronic ttH with BDT training

DT (Decision Tree) Example



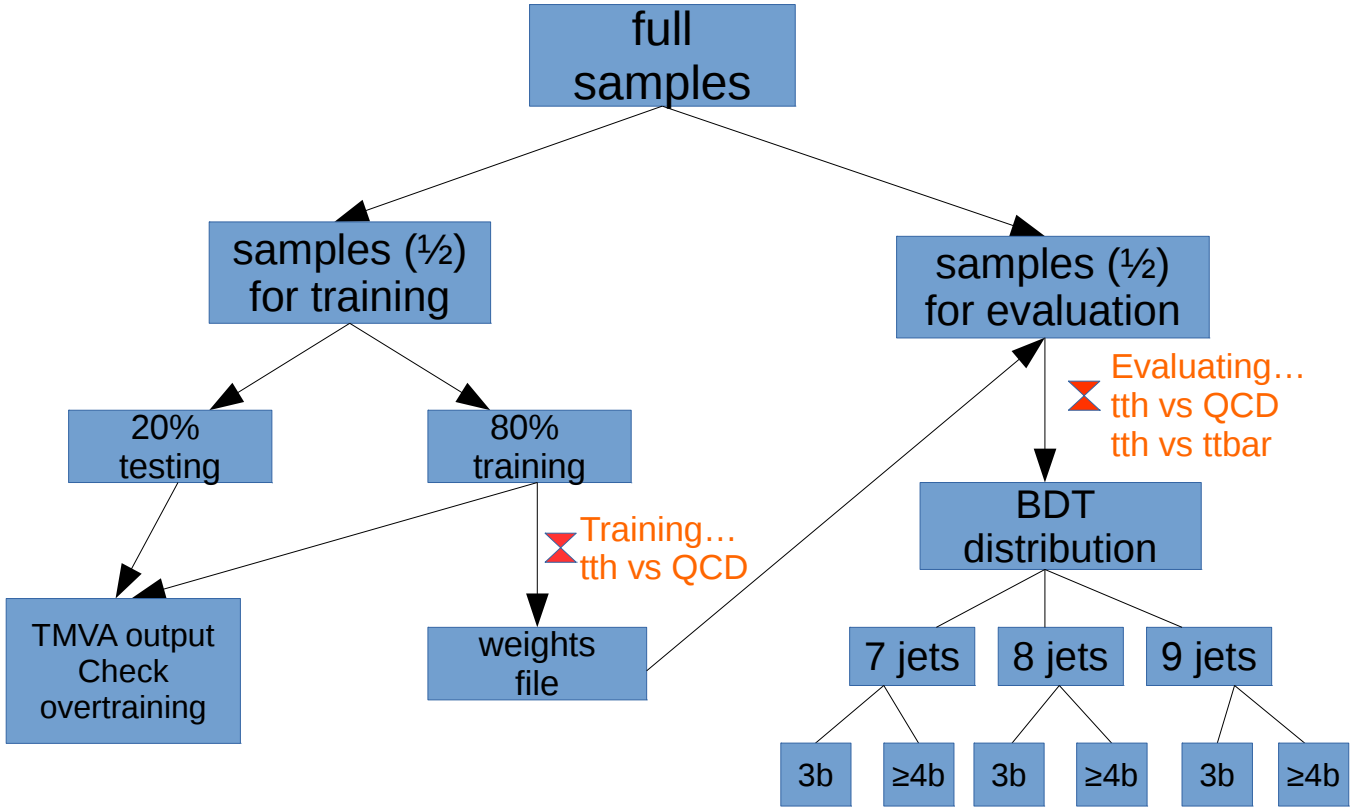
BDT (Boosted Decision Tree)

→ build many DTs
O(100) ~ O(1000),
so that the weighted
average will be insensitive
to fluctuations.

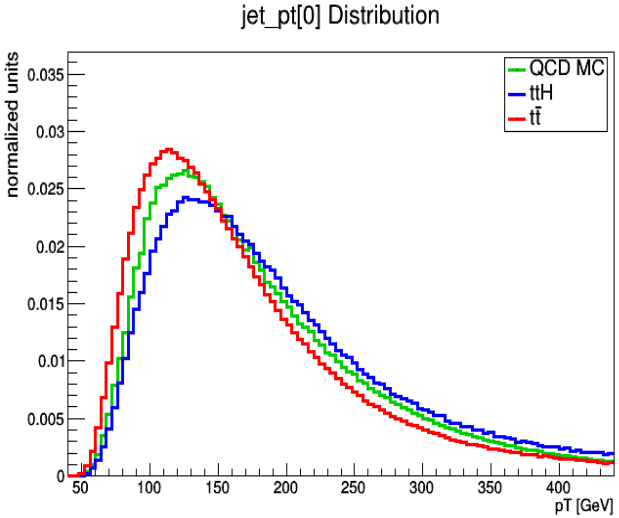
→ implement algorithm
A*daBoost*
(adaptive boosting)

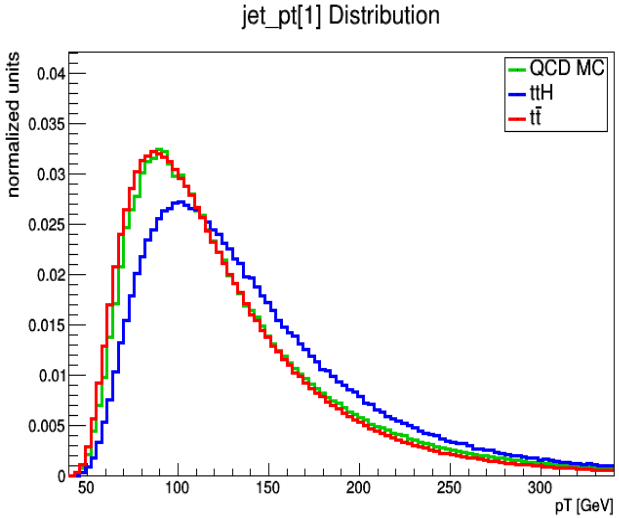
Training samples & procedures

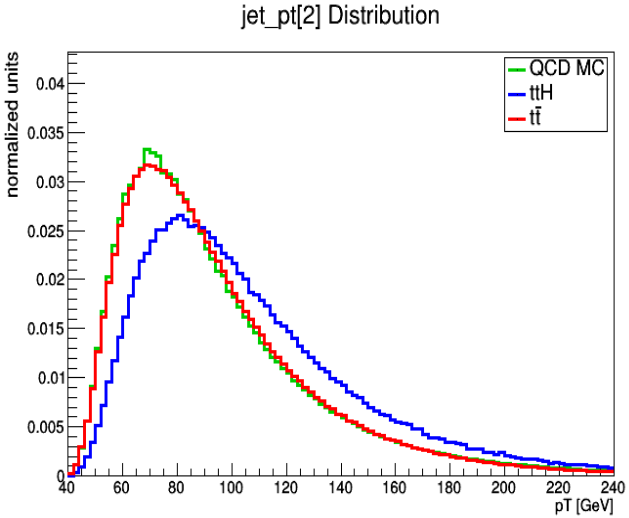
- Signal :
  - ttH : had_V24_4__ttHTobb_M125_13TeV_powheg_pythia8
- Background :
  - ttbar : had_V24_4__TT_TuneCUETP8M1_13TeV-powheg-pythia8
  - QCD Multijets : had_V24_4__QCD_HT300to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8, ~_HT500to700_~, … , ~_HT2000toInf_~

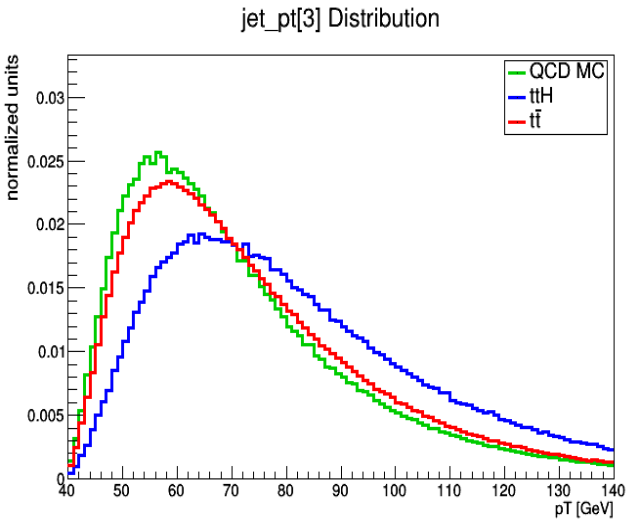


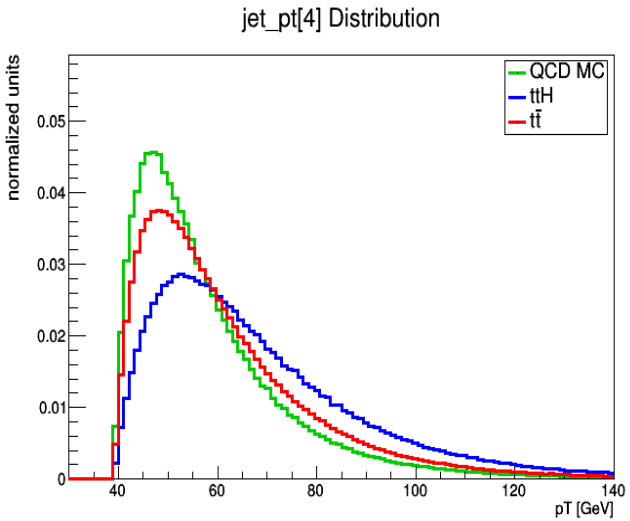
Training Variables: jet_pt[0], jet_pt[1] … jet_pt[5]

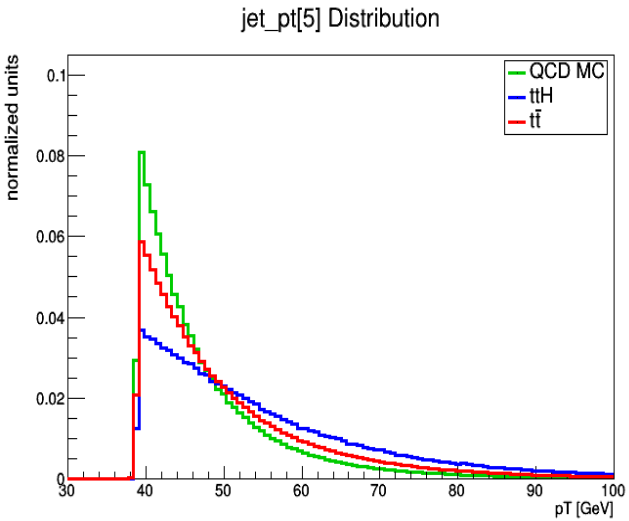




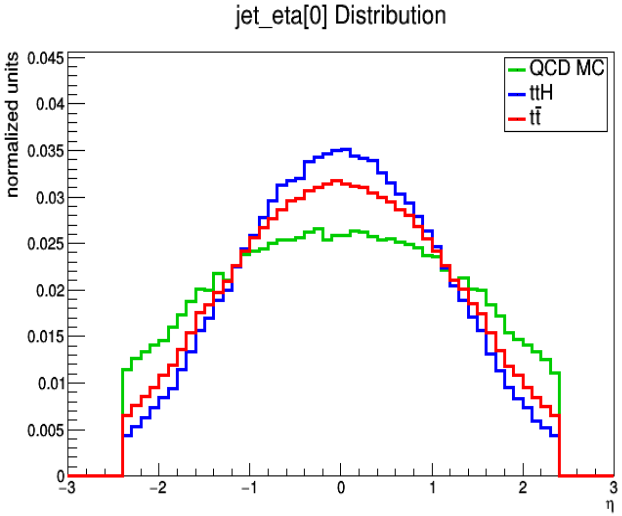


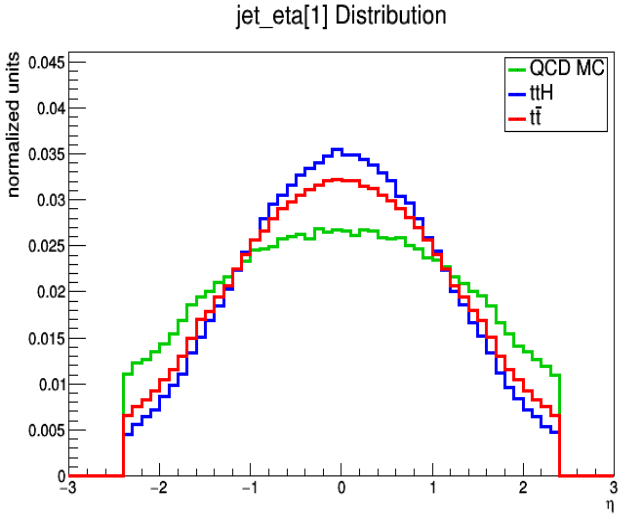


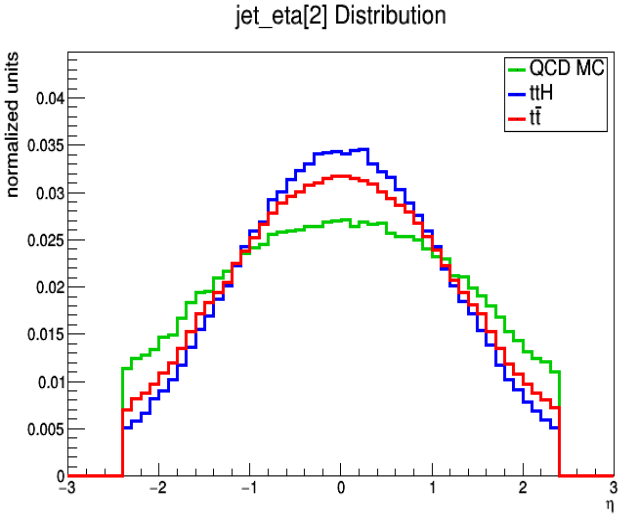


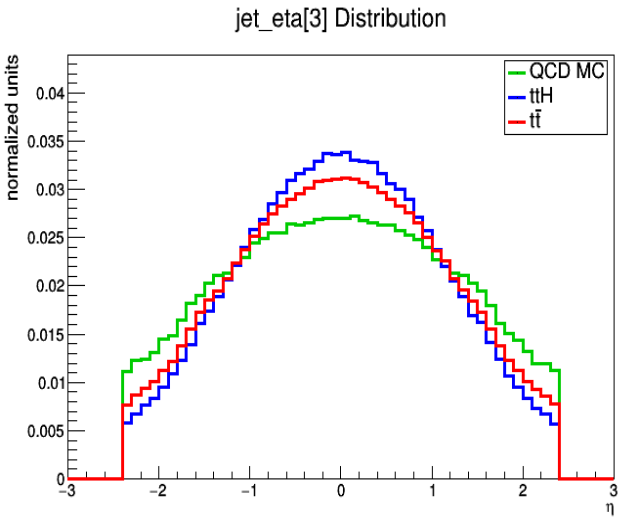


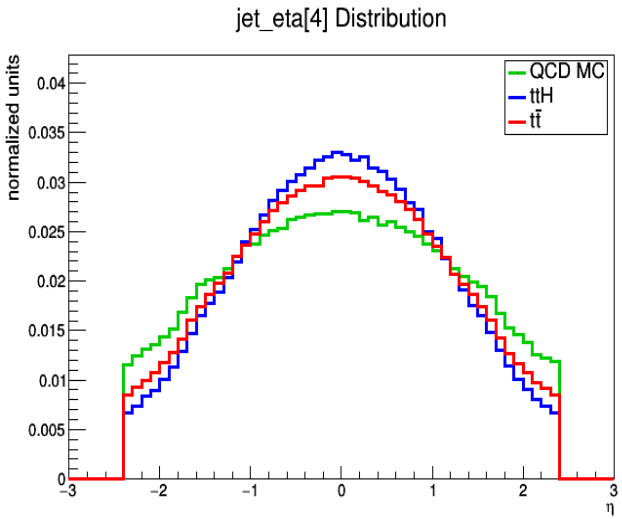
Training Variables: jet_eta[0] ... jet_eta[5]

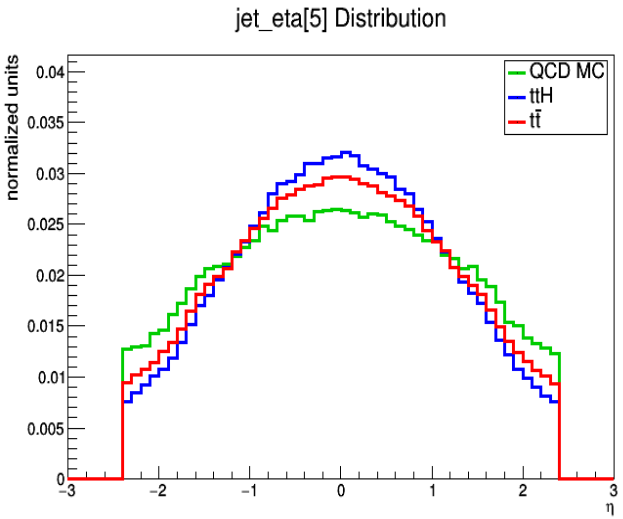




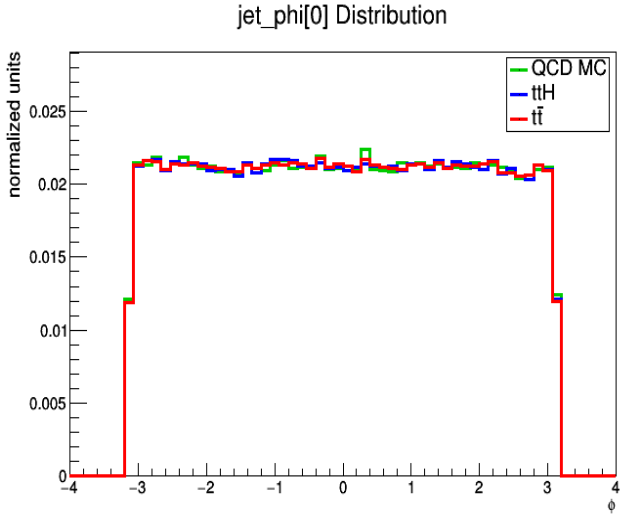


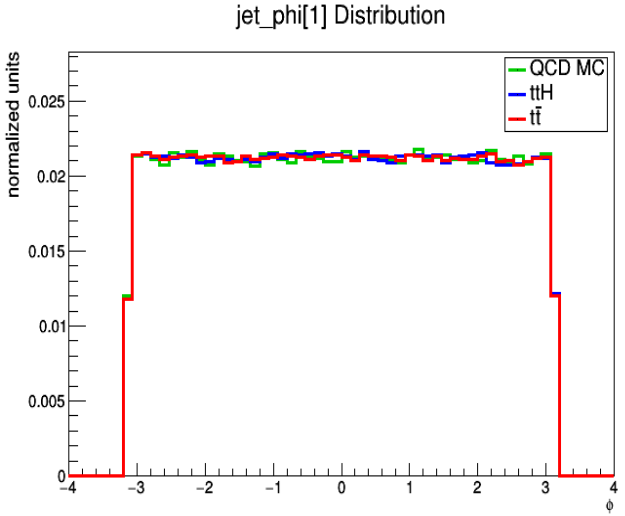


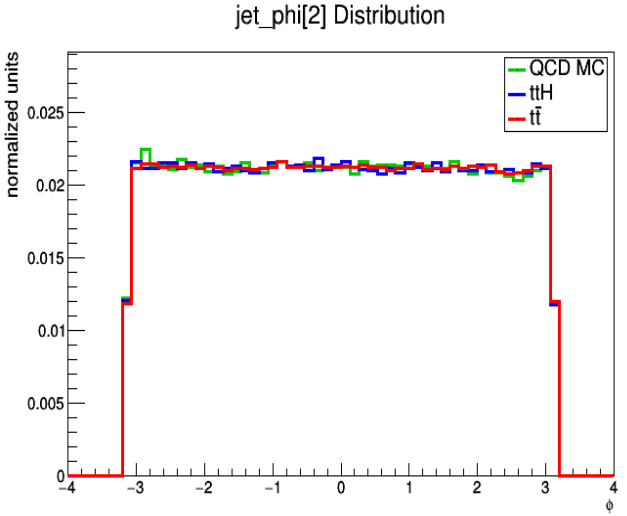


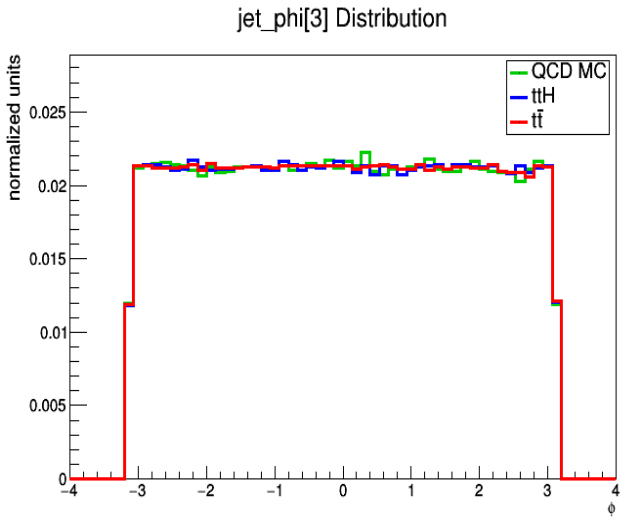


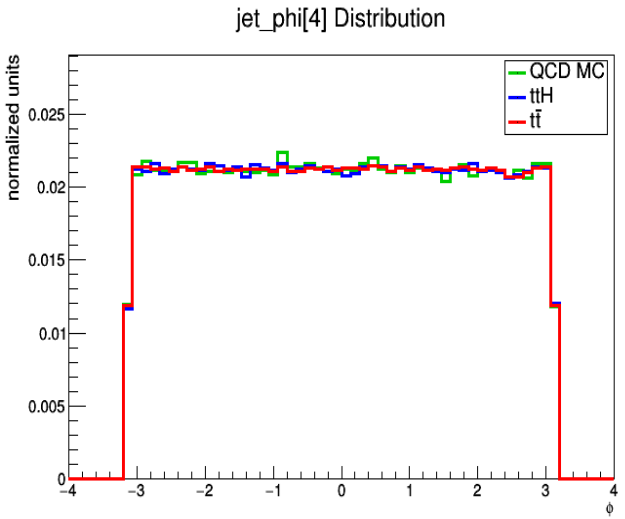
Training Variables: jet_phi[0] … jet_phi[5]

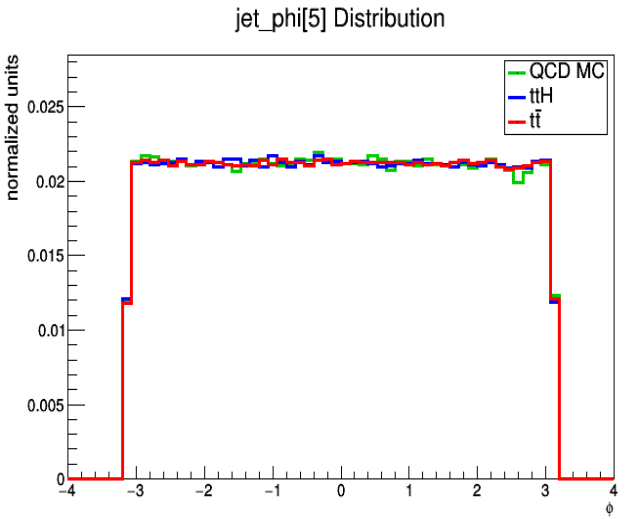




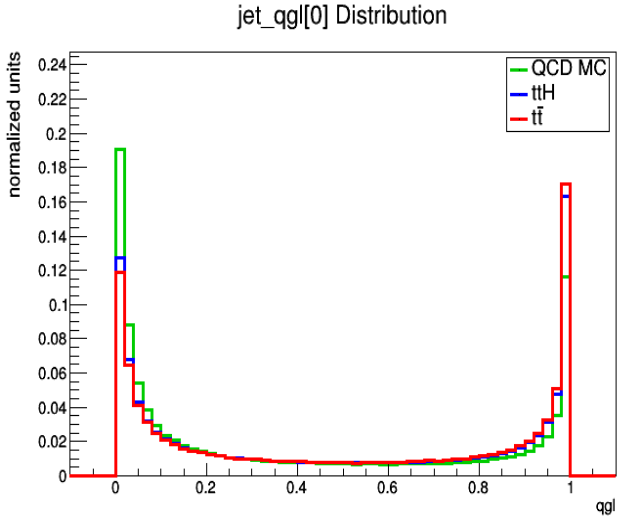


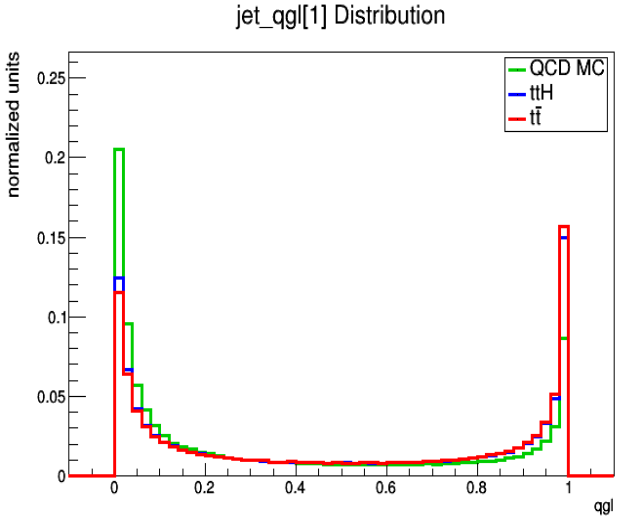


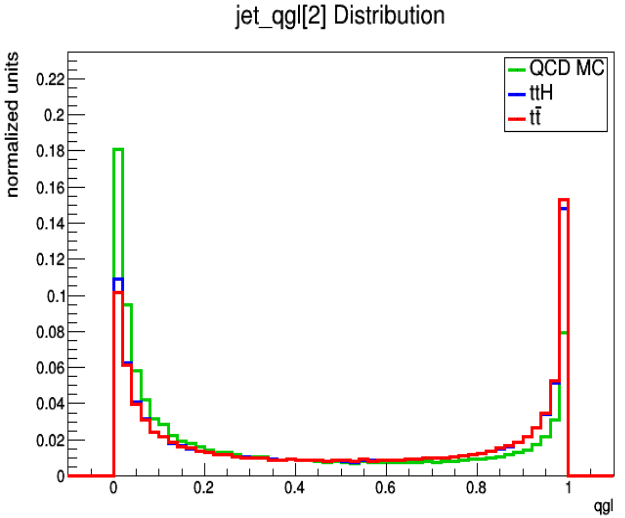


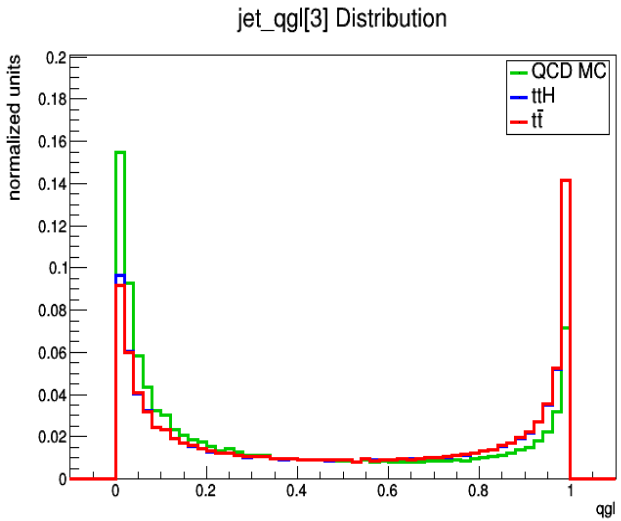


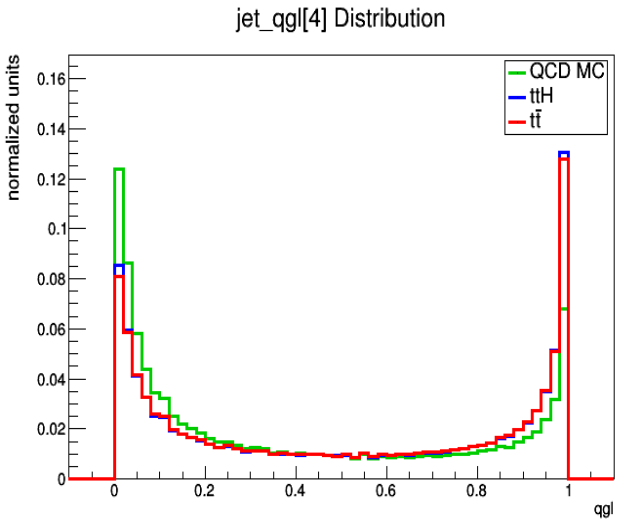
Training Variables: jet_qgl[0] ... jet_qgl[5]

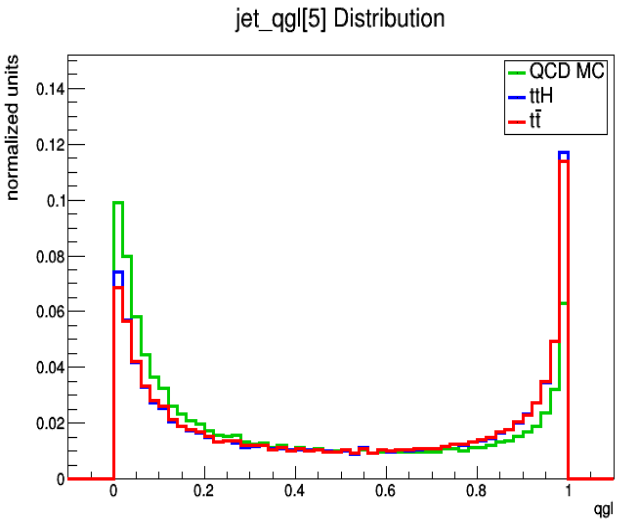




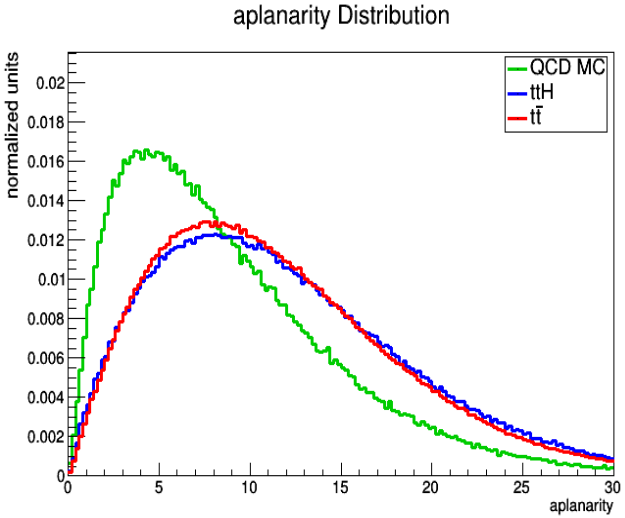


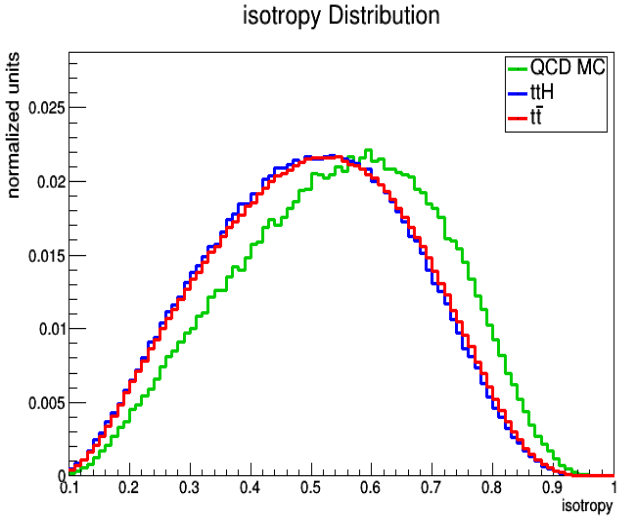


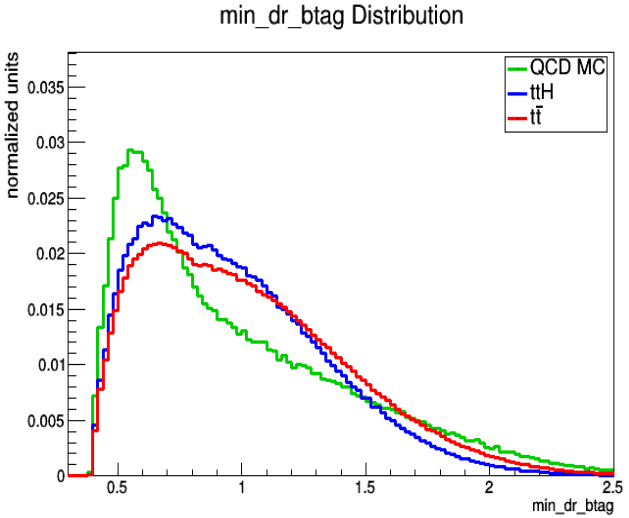


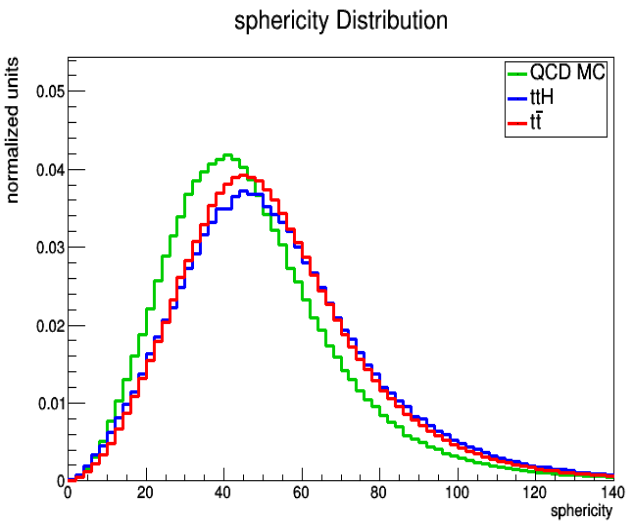


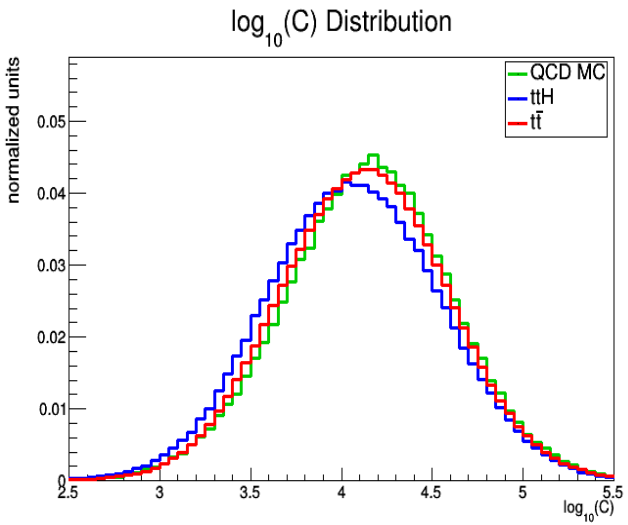
Training Variables: aplanarity, isotropy, min_dr_btag, sphericity, log(C), log(D)

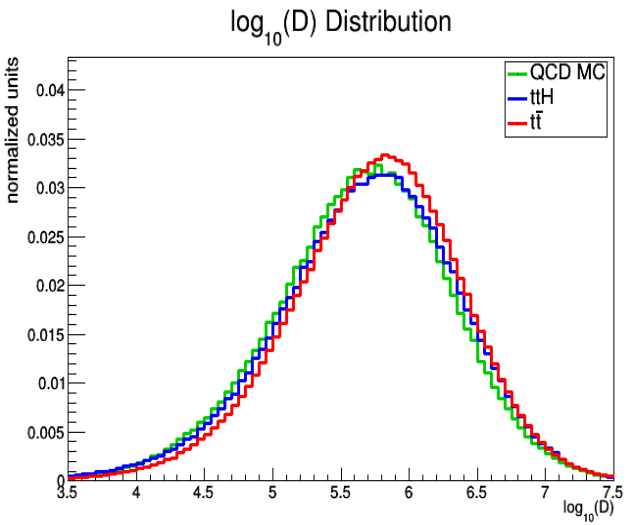




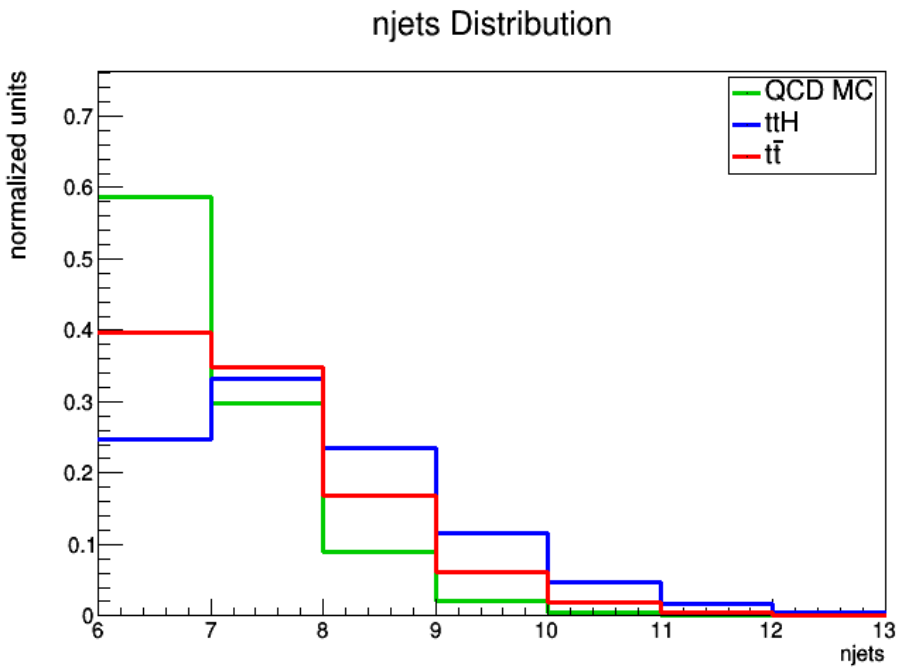


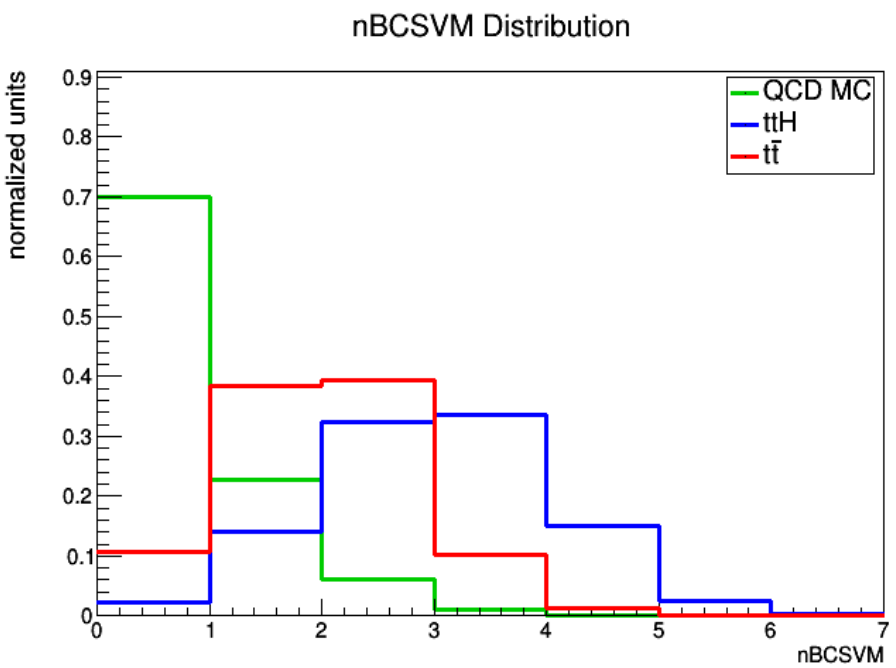




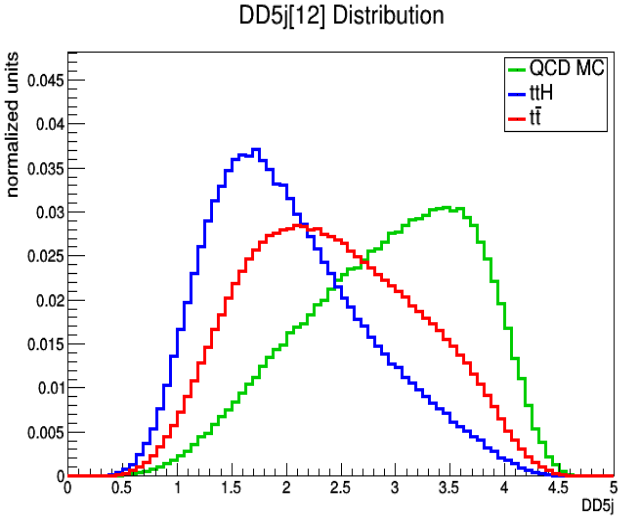


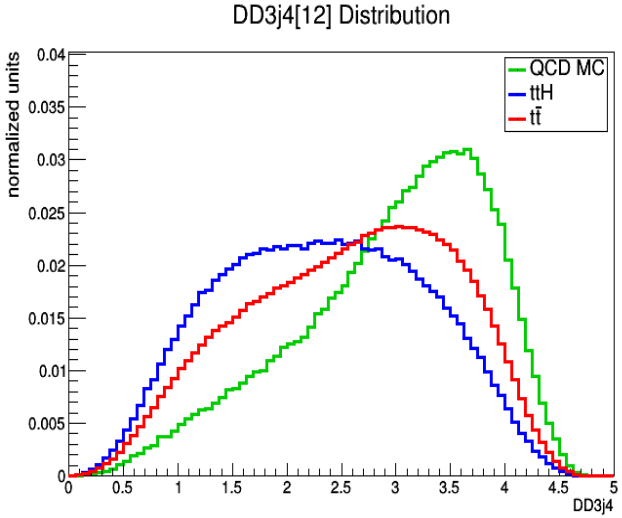
Training Variables: njets, nBCSVM

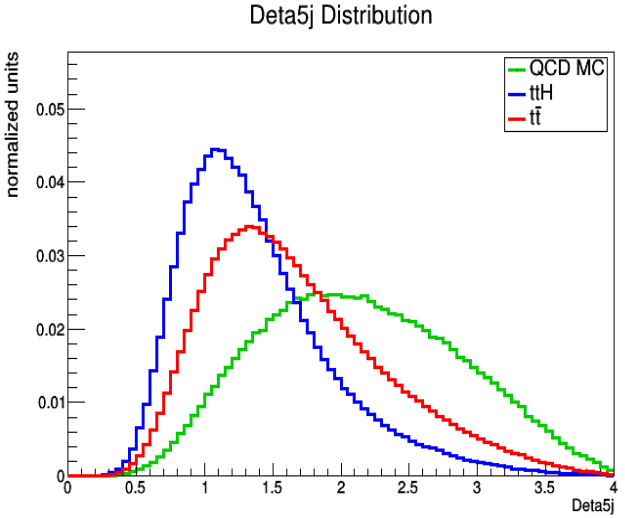


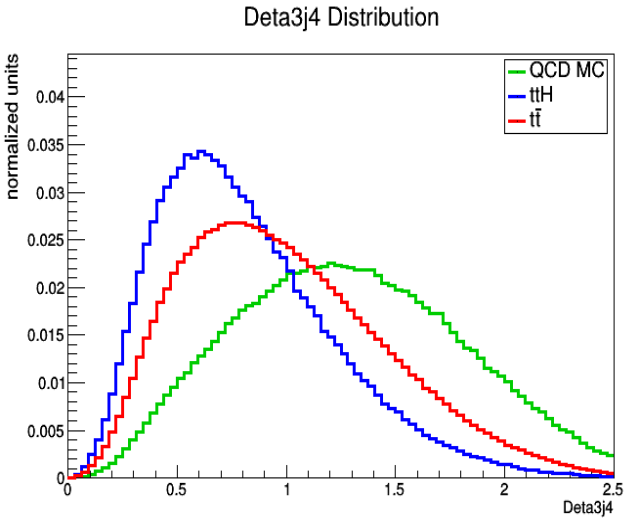


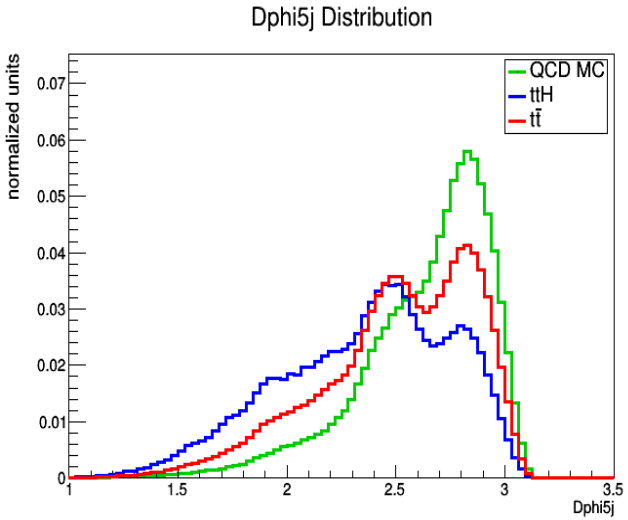
Training Variables: DD5j[12], DD3j4[12], Deta5j, Deta3j4, Dphi5j, Dphi4j5

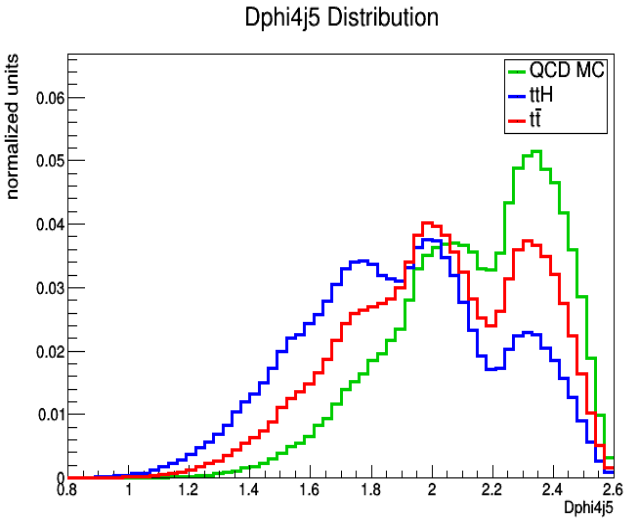




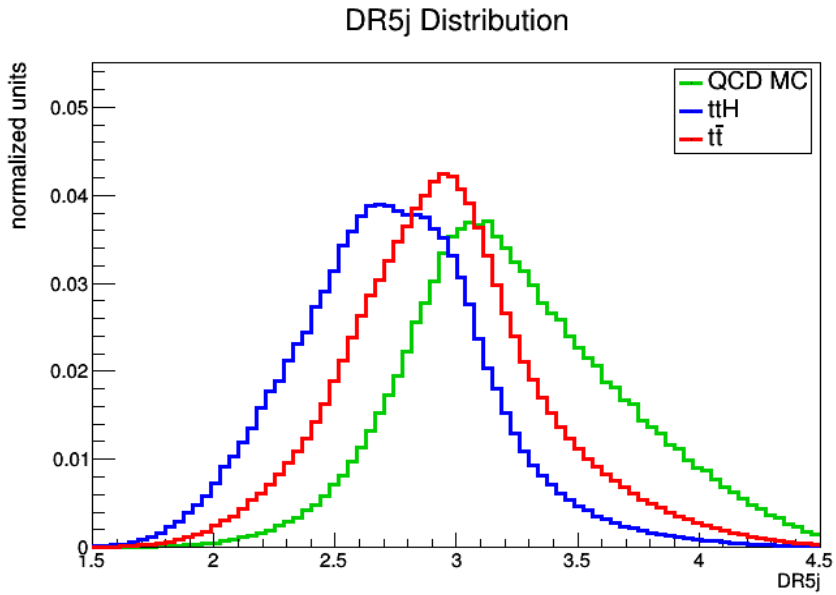


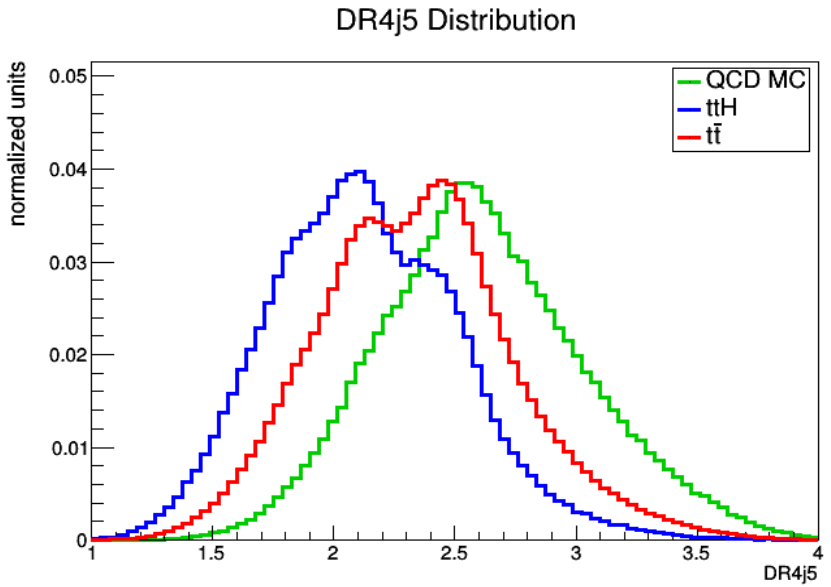


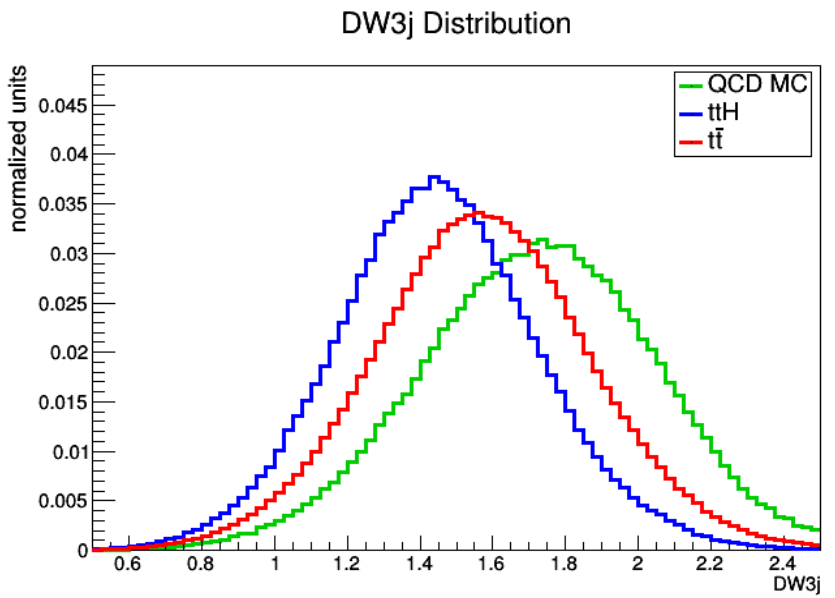


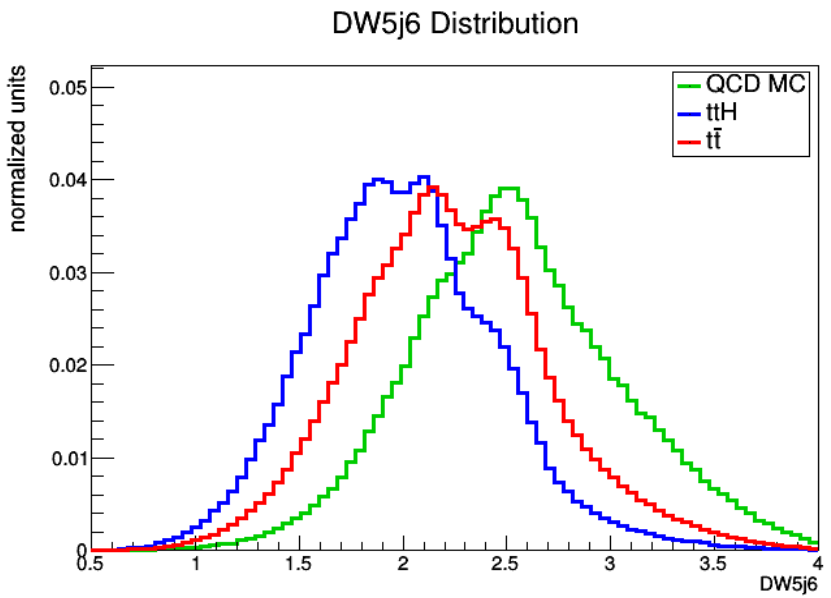


Training Variables: DR5j, DR4j5, DW3j, DW5j6









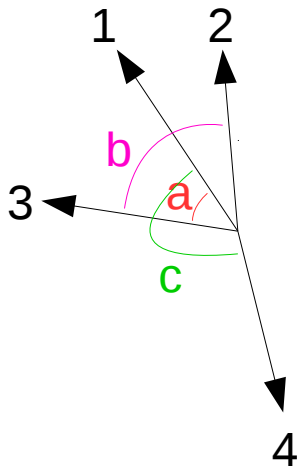
Distance variables

Example 1:

→ Dphi2j : the average Δφ of the 2nd closest jet to each jet in an event.

→ in this case:
j1: Δφ = a, j2: Δφ = b,
j3: Δφ = b, j4: Δφ = c
Dphi2j = ¼(a + b + b + c)

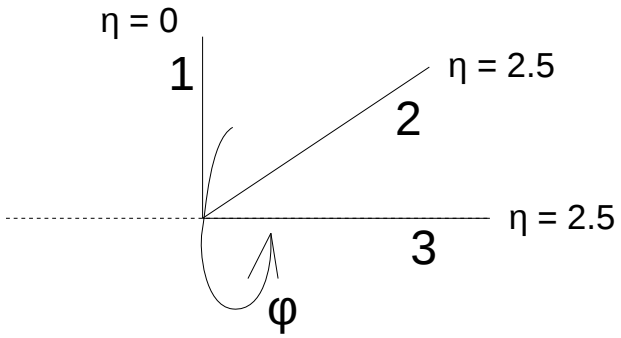
→ Therefore, Dphi4j5 : the average Δφ of the 4th closest jet to each jet of the first 5 jets in an event.



Example 2:

→ DR = √($\Delta\varphi^2 + \Delta\eta^2$)

→ DD & DW are similar to DR, but are used to penalize Δφ when Δη is very small

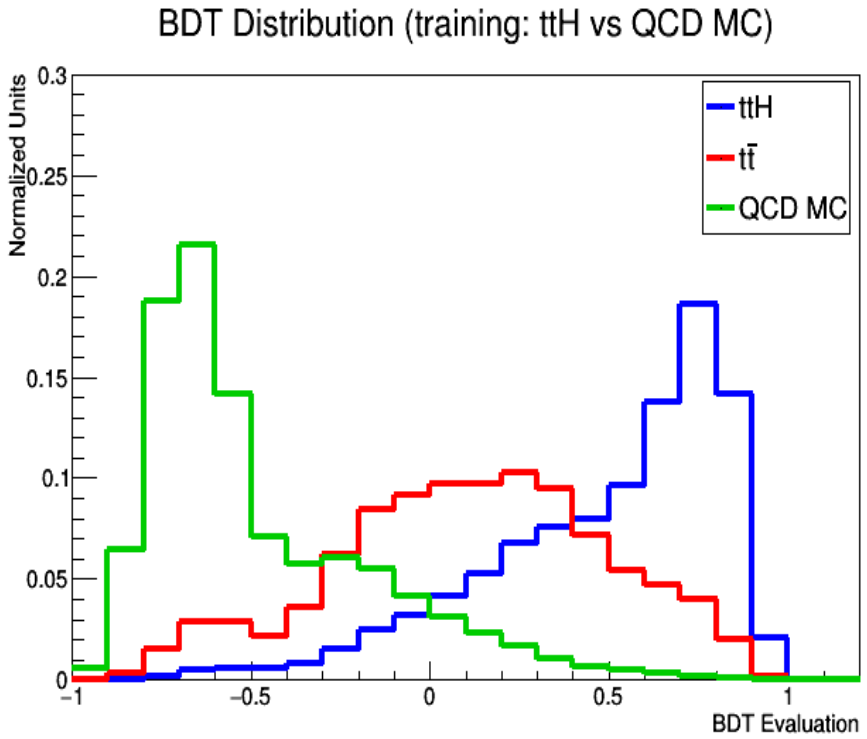


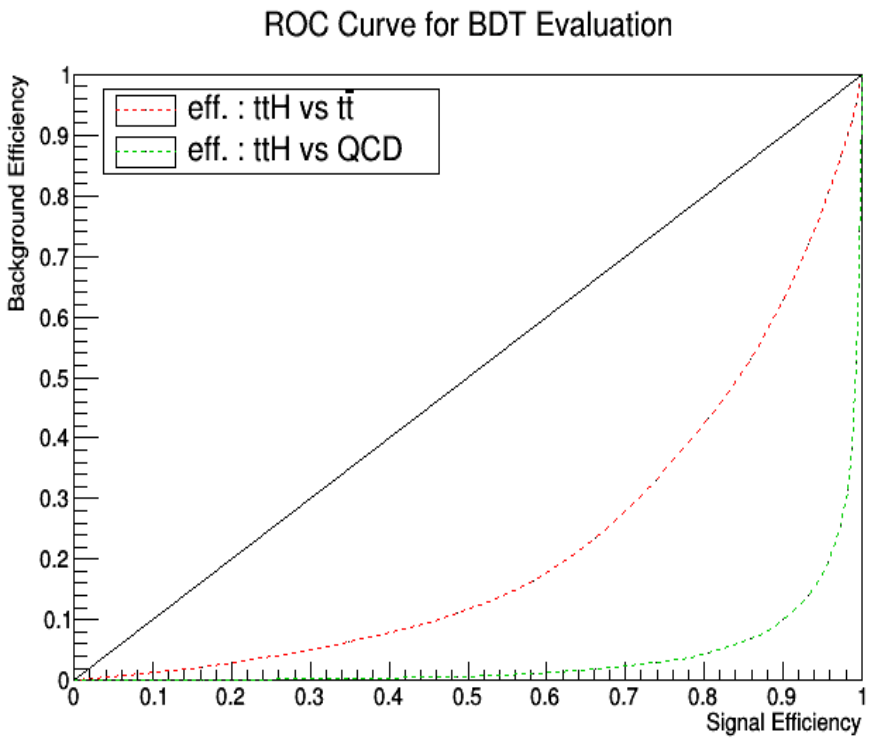
BDT Training & Evaluation (1st attempt)

Input variables(16) :

- jet_pt[0] … jet_pt[3], jet_eta[0] … jet_eta[3]
- aplanarity, isotropy, sphericity, min_dr_btag, log(C), log(D)
- njets, nBCSVM

However, we can't rely on njets & nBCSVM since later on we will categorize BDT distribution based on these two variables.

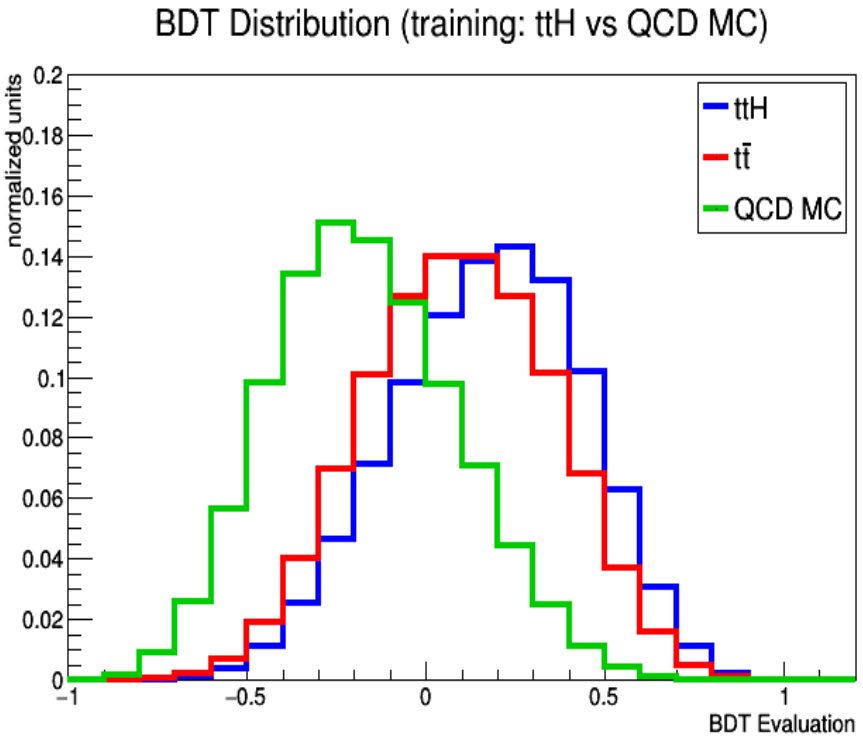


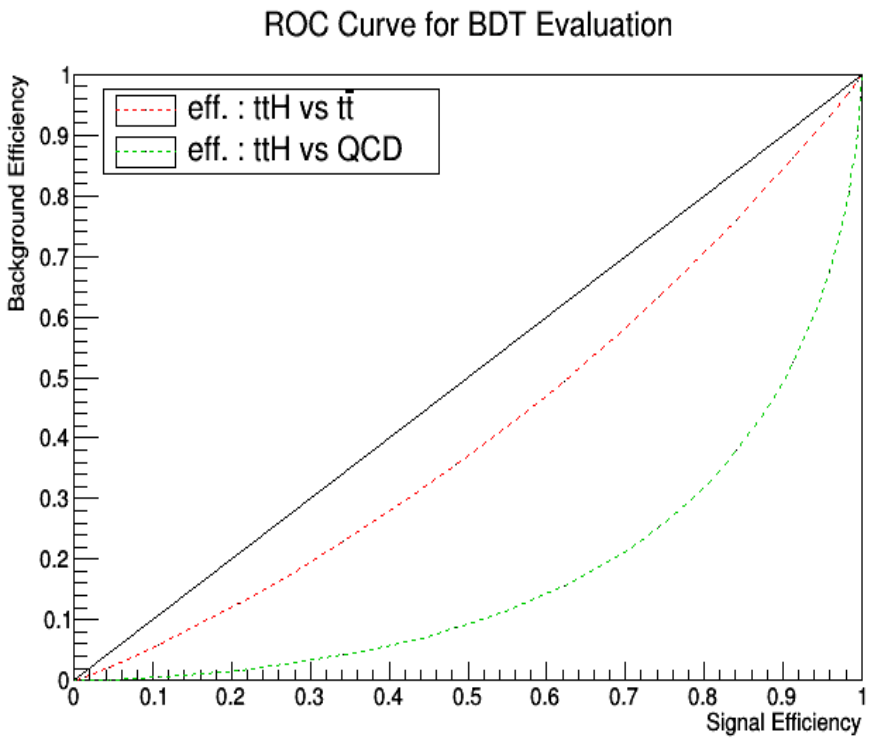


BDT Training & Evaluation (2nd attempt)

Input variables(30) : → njets & nBCSVM are restricted

- jet_pt[0] … jet_pt[5], jet_eta[0] … jet_eta[5]
- jet_phi[0] … jet_phi[5], jet_qgl[0] … jet_qgl[5]
- aplanarity, isotropy, sphericity, min_dr_btag, log(C), log(D)





BDT Training & Evaluation (3rd attempt)

Input variables(40) :

Previous variables(30) :

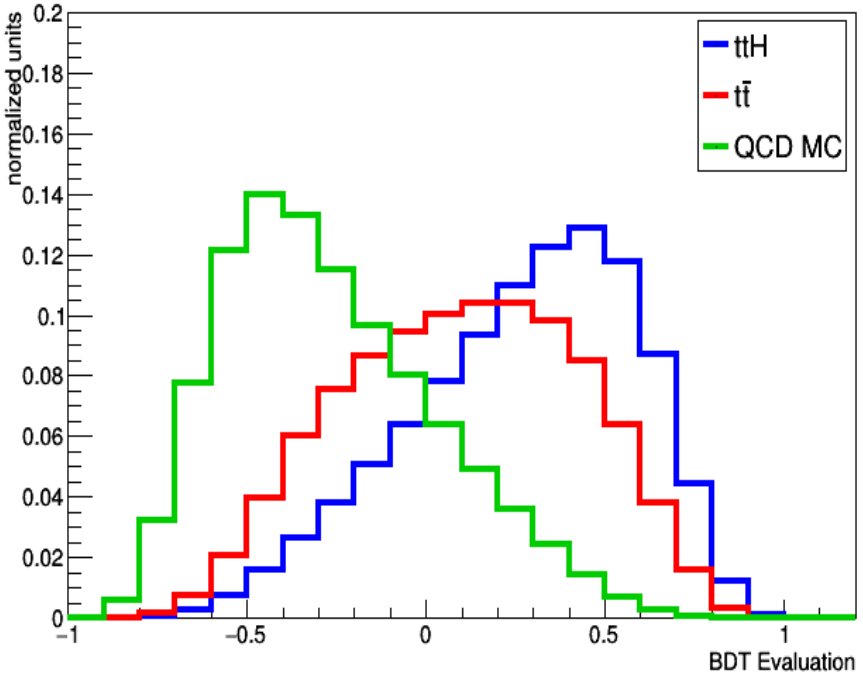
- jet_pt[0] … jet_pt[5], jet_eta[0] … jet_eta[5]
- jet_phi[0] … jet_phi[5], jet_qgl[0] … jet_qgl[5]
- aplanarity, isotropy, sphericity, min_dr_btag, log(C), log(D)

+

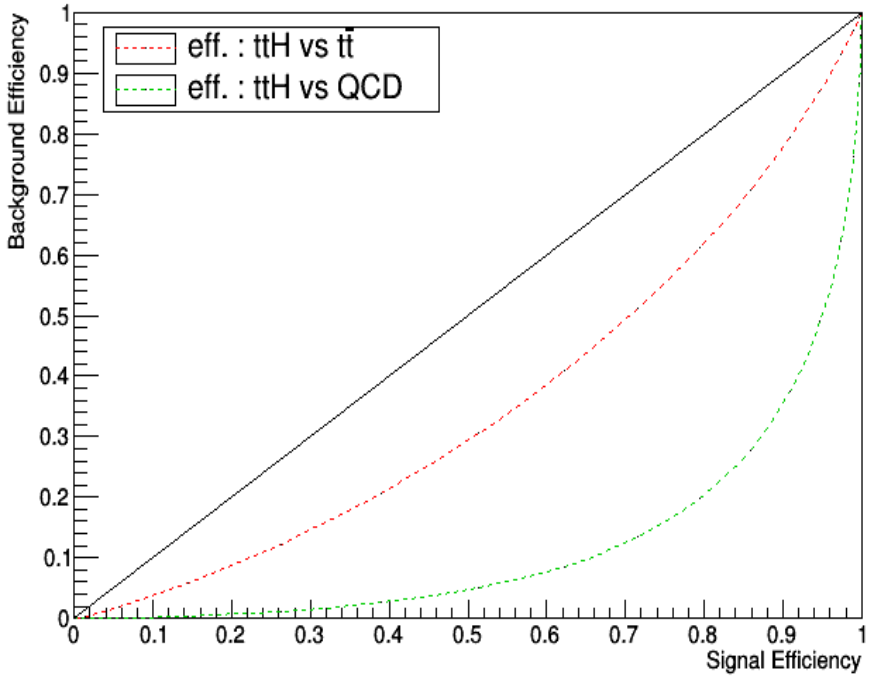
Distance variables(10):

- DD5j[12], DD3j4[12]
- Deta5j, Deta3j4, Dphi5j, Dphi4j5, DR5j, DR4j5, DW3j, DW5j6

BDT Distribution (training: ttH vs QCD MC)



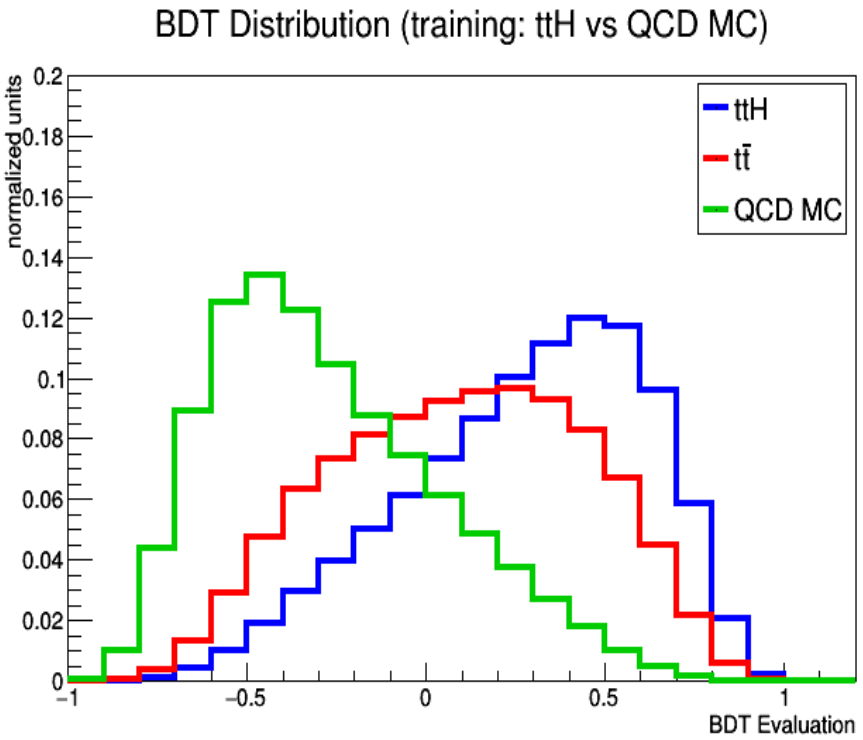
ROC Curve for BDT Evaluation

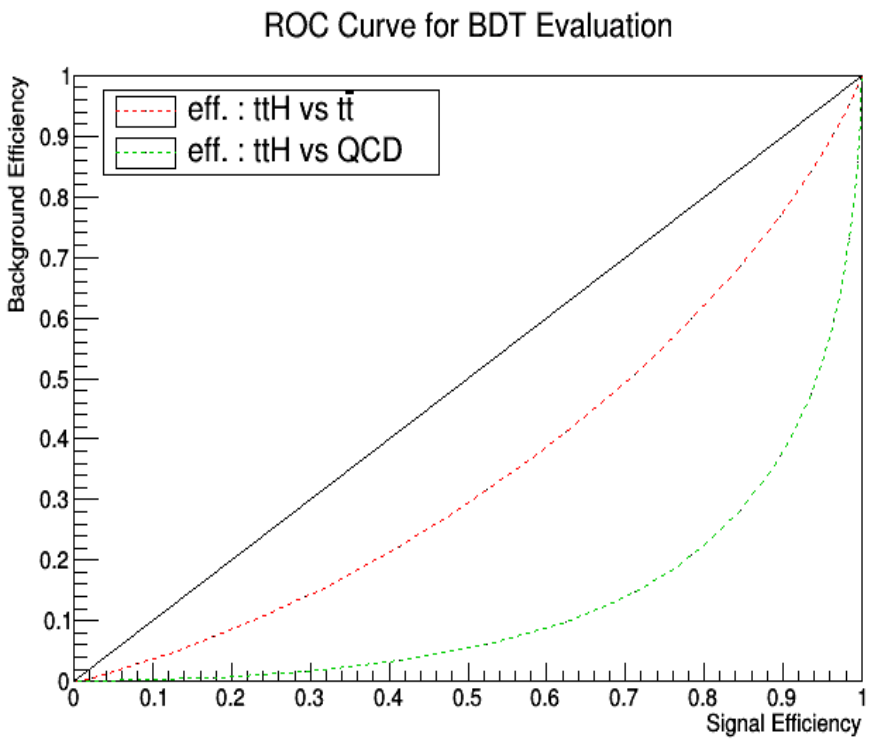


BDT Training & Evaluation (4th attempt)

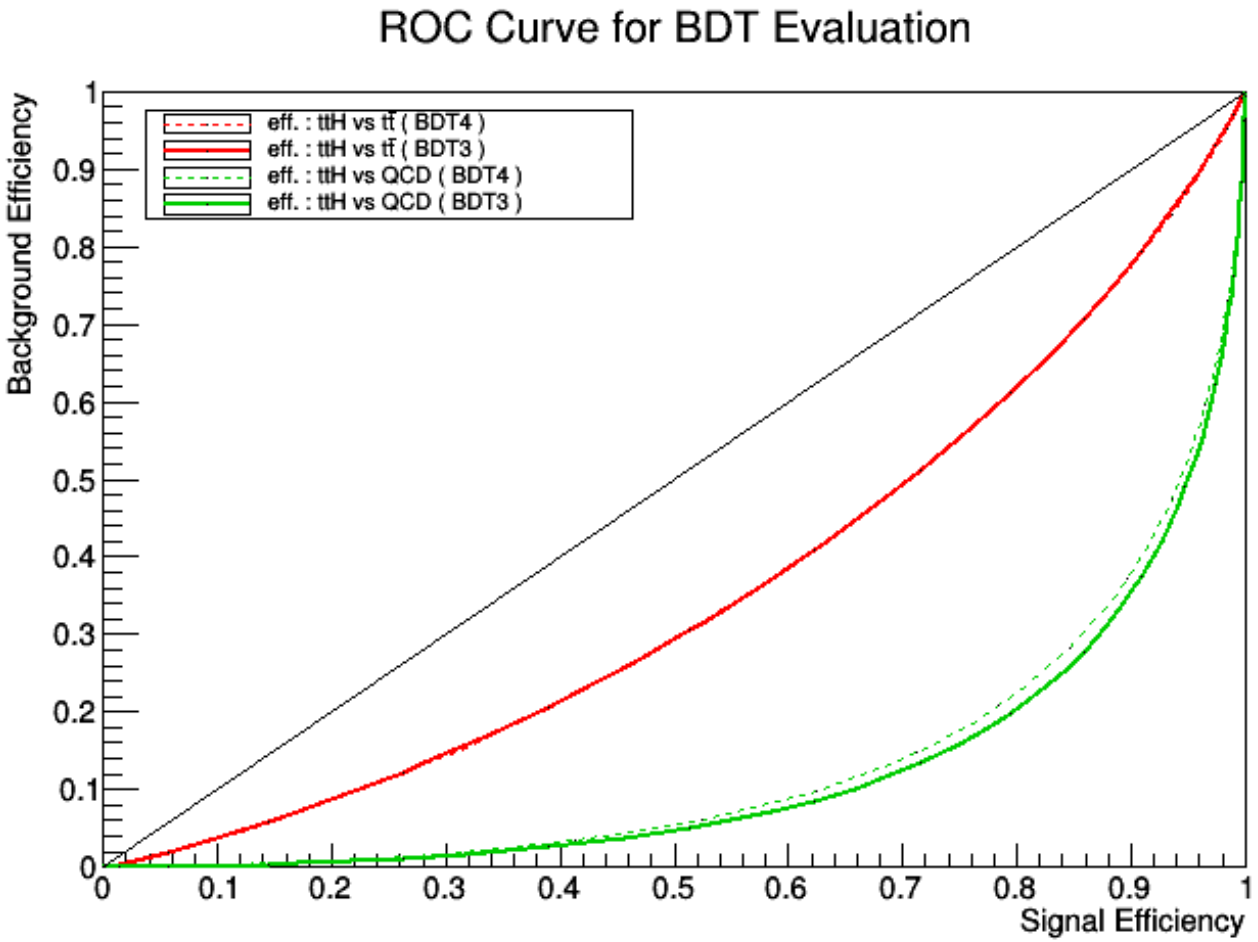
Input variables(28) : →variables that don't show too much discrimination are restricted

- jet_pt[0] … jet_pt[3], jet_eta[0] … jet_eta[3]
- jet_qgl[0] … jet_qgl[3]
- aplanarity, isotropy, sphericity, min_dr_btag, log(C), log(D)
- DD5j[12], DD3j4[12]
- Deta5j, Deta3j4, Dphi5j, Dphi4j5, DR5j, DR4j5, DW3j, DW5j6



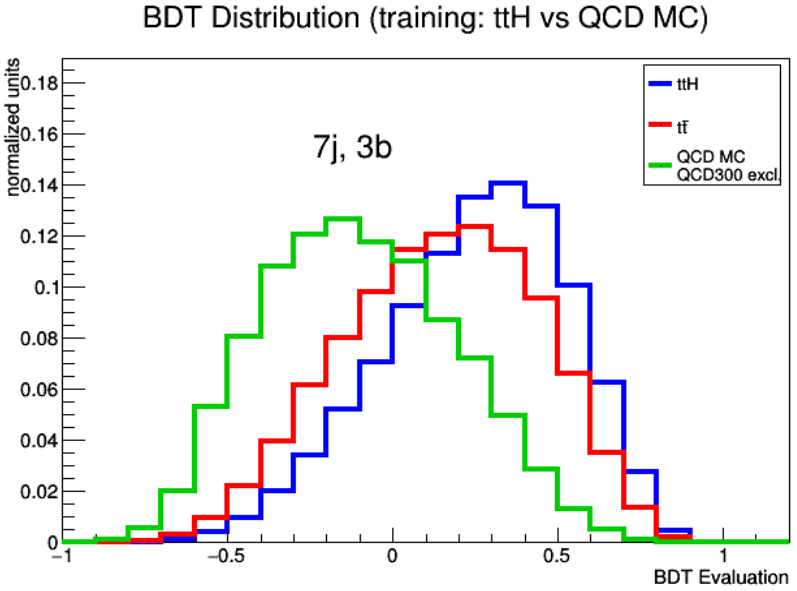


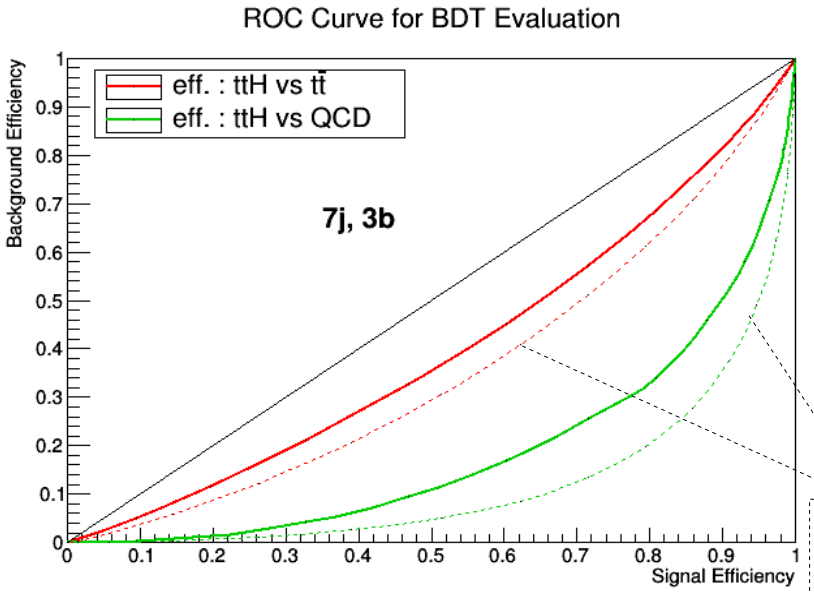
Comparison between the 3rd & the 4th attempts



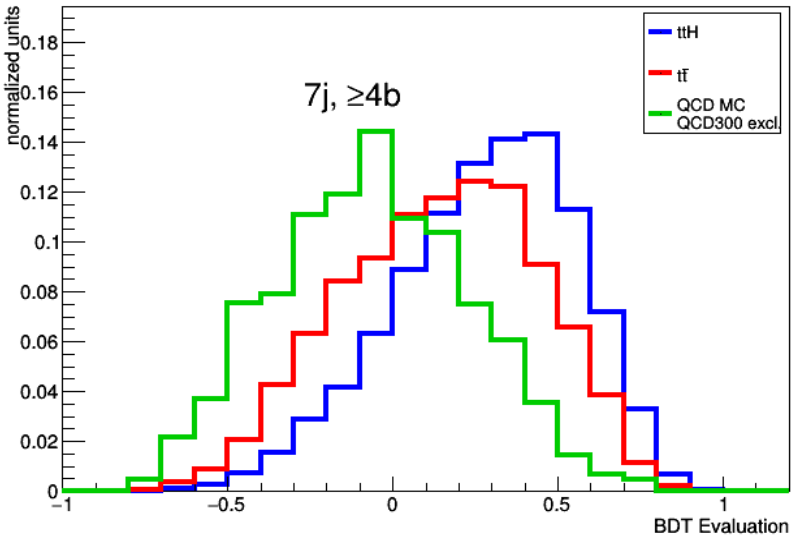
The full training provides the best performance

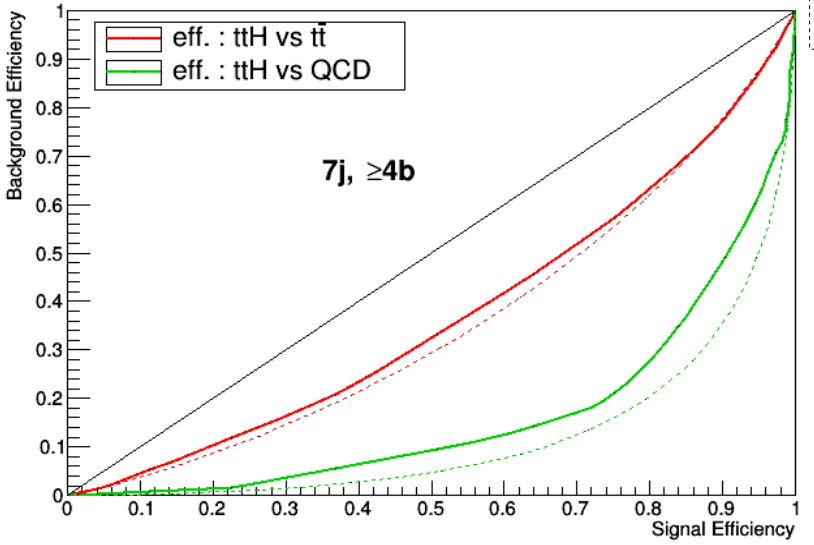
Categorization for the full training
BDT distribution when njets = 7



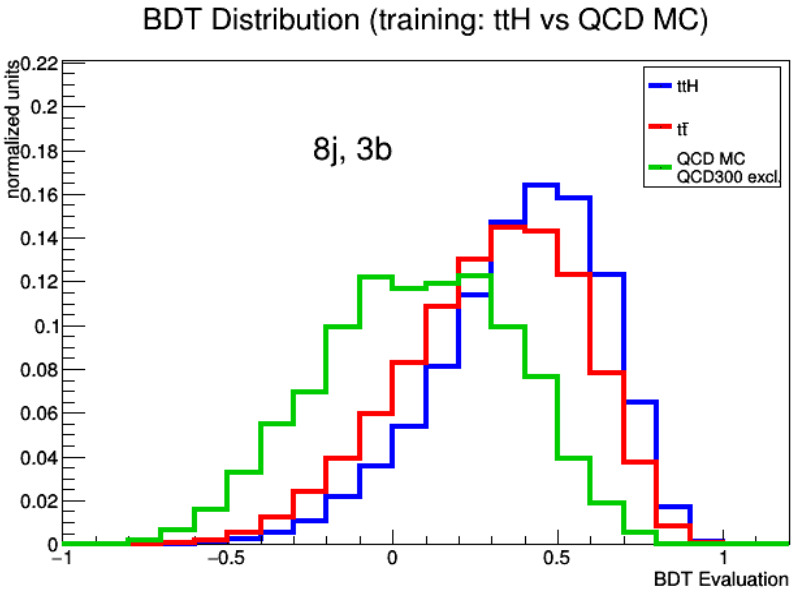


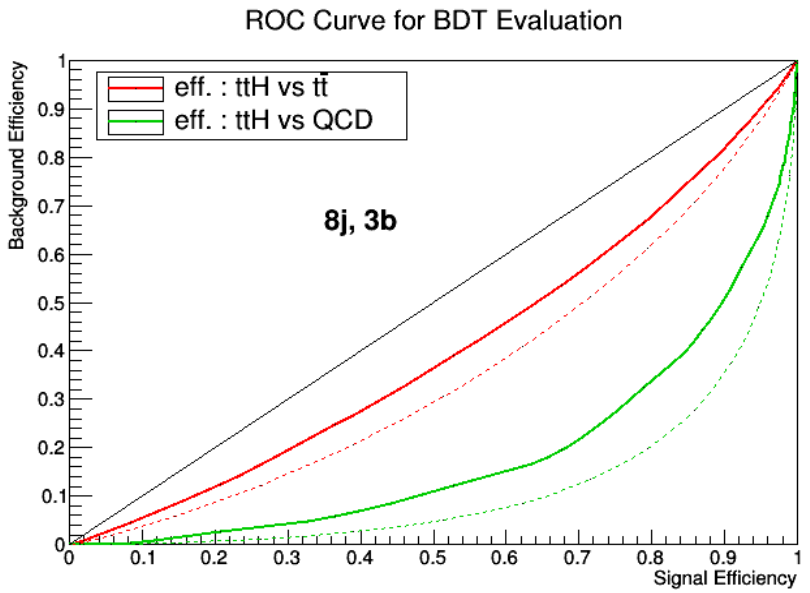
ref. ROC: without any selections in njets & nBCSVM

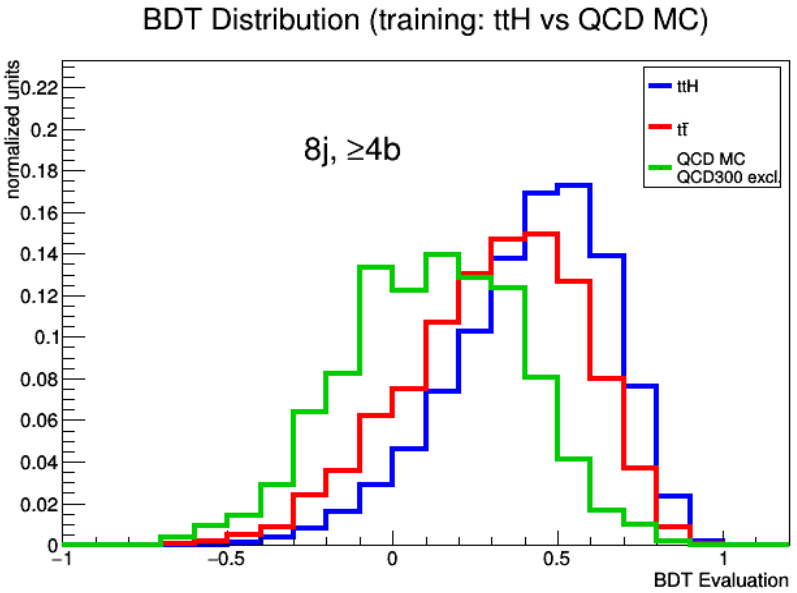


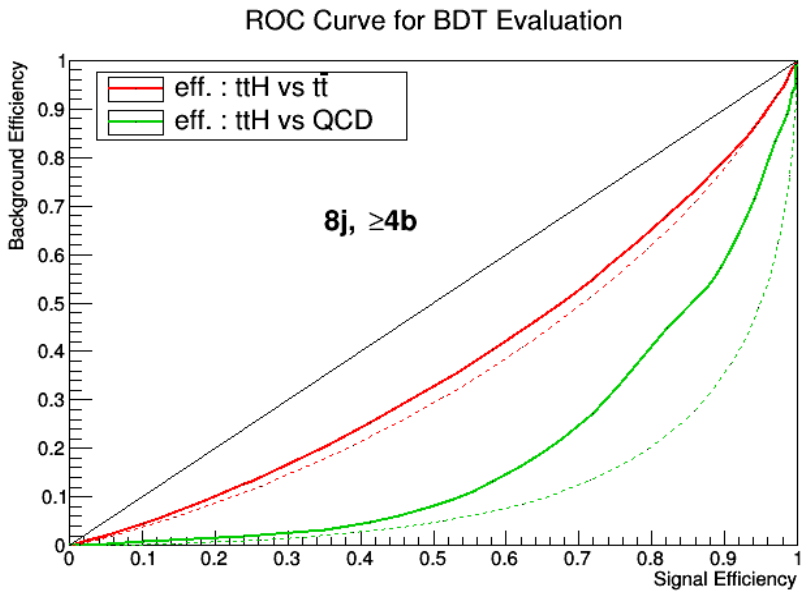


Categorization for the full training
BDT distribution when njets = 8









Categorization for the full training
BDT distribution when njets = 9