

Status report

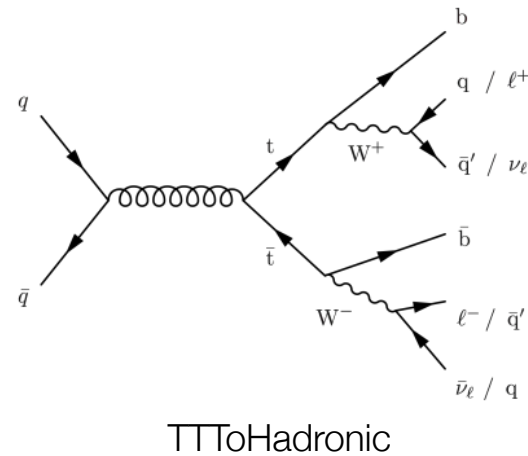
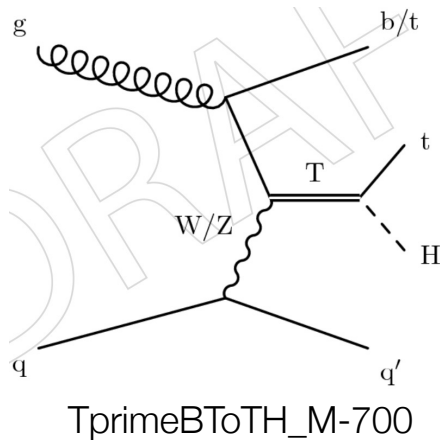
August 26, 2022

Jieun Choi

HYU / IP2I

NN in analysis

Target process

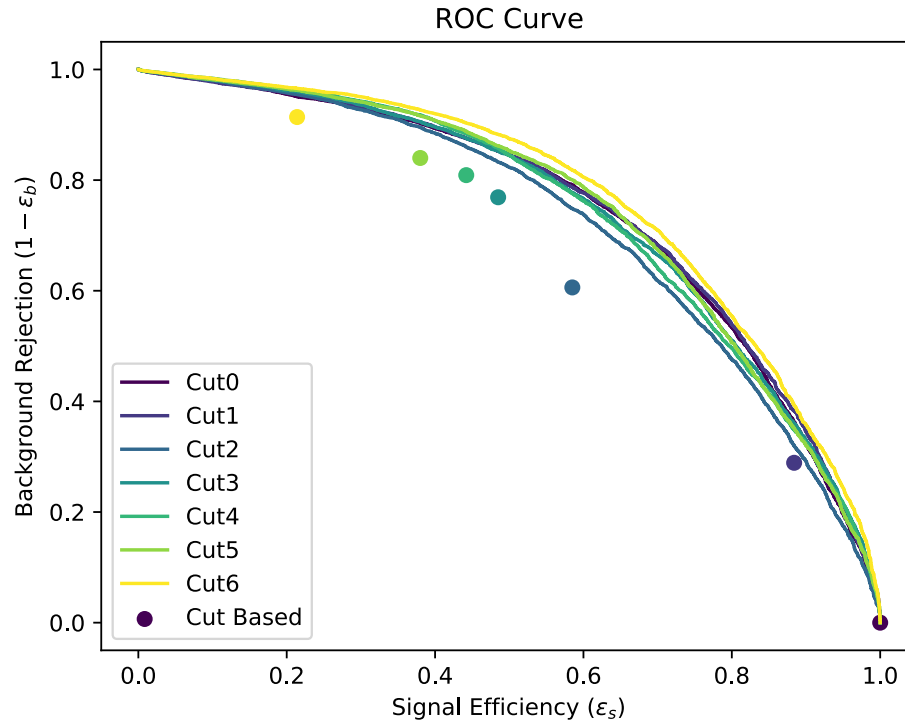


Check the importance of cuts via NN

- Variables which were used for Cut0
 - {eta, energy/chi2mass, b-disc.} of (b-tagged) jets, goodHT, Chi2_min, 2nd Top Mass, Higgs Mass
 - njets, nbjets
- Cut1: RelHT
- Cut2: Chi2_Max
- Cut3: Chi2_dRHbb
- Cut4: Chi2_min_H
- Cut5: Chi2_dRWjj
- Cut6: Chi2_dRbW
- Candidates (not tested yet): Chi2_min_W, Chi2_min_Top, mindR_dRbb, mindR_mbb, w_mass, top_mass, chi2_mass, w+H_mass, mass(top-H / top+H), mass(2ndTop+2ndW / H), Chi2_dRHTop, Chi2_dEtaWH, Chi2_dPhiHTop, pT (2ndTop - top / 2ndTop), pT (H-top / Tprime), pT (T'/H - T'/top)

NN in step by step

DNN Structure:
3 layers with 100 nodes
Dropout: 0.1
Activation: relu+sigmoid
Optimizer: Adam
Loss: binary_crossentropy
Batch size: 2048



Cut step	AUC (1)	AUC (2)	AUC (3)
Cut 0	0.7520	0.7462	0.7588
Cut 1	0.7544	0.7500	0.7351
Cut 2	0.7418	0.7236	0.7222
Cut 3	0.7530	0.7446	0.7316
Cut 4	0.7627	0.7419	0.7559
Cut 5	0.7624	0.7488	0.7671
Cut 6	0.7648	0.7693	0.7566

AUC: Area Under Curve

Details

- Training strategy is the same: (with TprimeBToTH_M-700 2018)
 - Train on odd numbered event after DNN selection (HLT+6 jets+3 bjets)
 - Evaluate on even numbered event after cut 0
- Compared performances by adding input variables used in cutBased without cutting itself
- Performance fluctuates $\sim AUC \pm 0.02$ (tested for 3 trainings per each cuts (Table 1~3))
 - Depends on computing resources / random seed in the initializer
 - Once a model is determined, performance does not change
 - Not much differences observed from adding a few variables
 - General performance tend to go up by adding cut 0 to cut 6, but drop down at **Cut2 (Chi2_Max)**

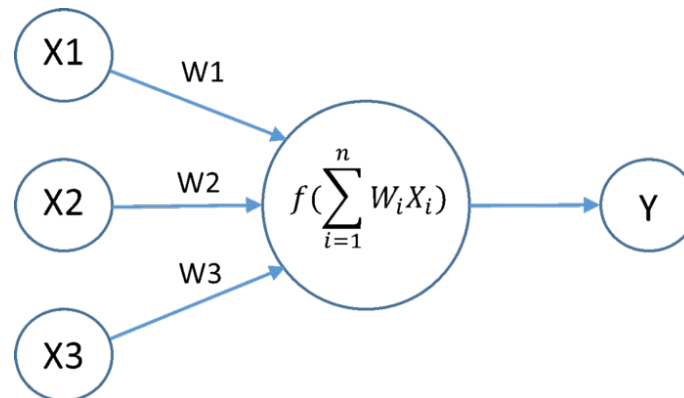
Taylor expansion for feature importance

Motivation

- Initial question: What are the input features with the largest impact on the NN output nodes?
- Extract average gradient for each input features
- Will be able to "see" how much each variable "effects" on training model

Methodology

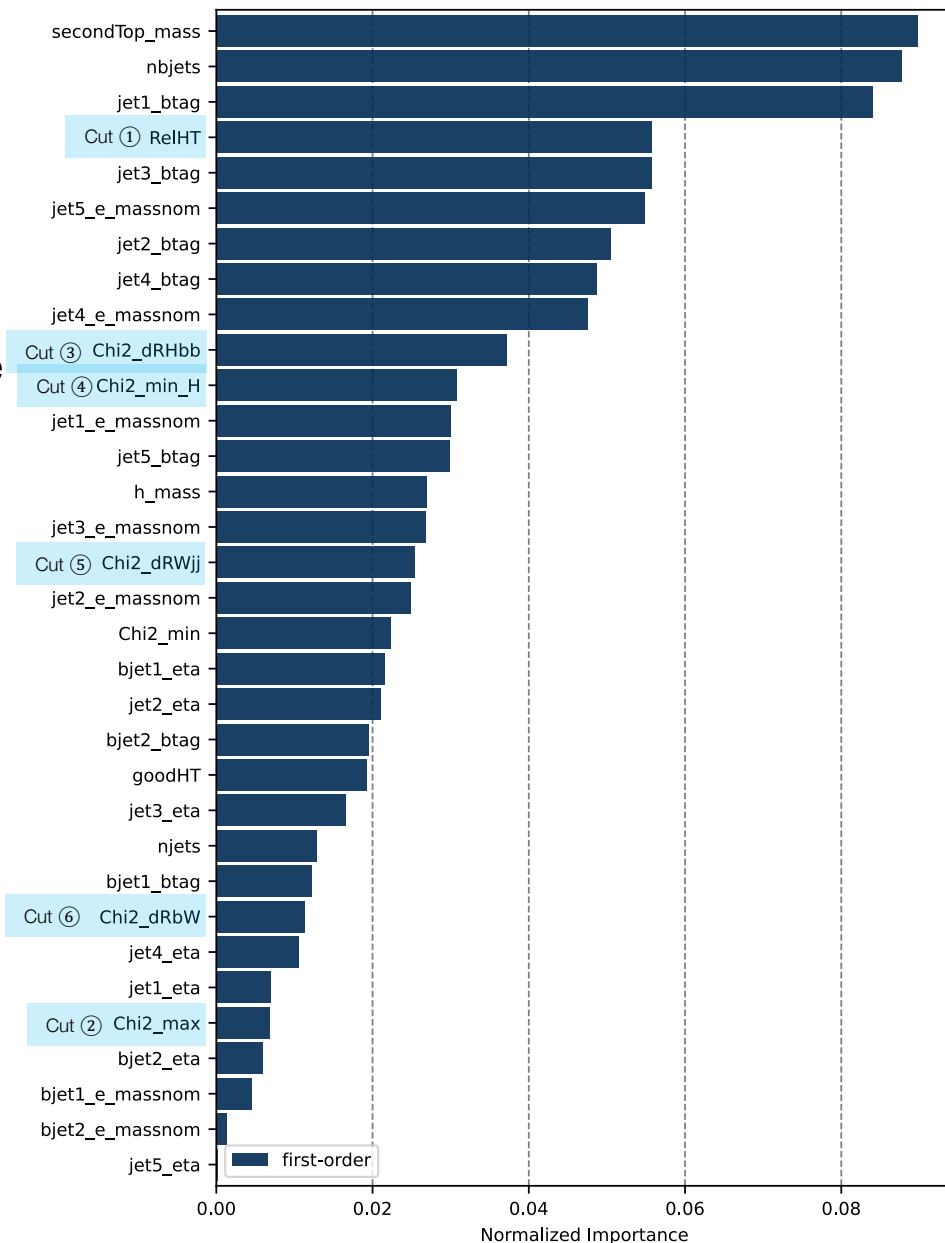
- Taylor expansion of the output function at the minima (model)
- [Tensorflow.GradientTape\(\)](#)
 - Allow us to record the history of operations applied to target input features
 - Calculate gradients of output(node) w.r.t. inputs(event)
- 1st order: Physical location of feature/marginal distributions – weight w_i for x_i
- 2nd order: Curvature of NN output function - correlations across two features: Gradient of each element of the source w.r.t target – weight w_{ij} for $x_i * x_j$



Taylor expansion for feature importance

Result

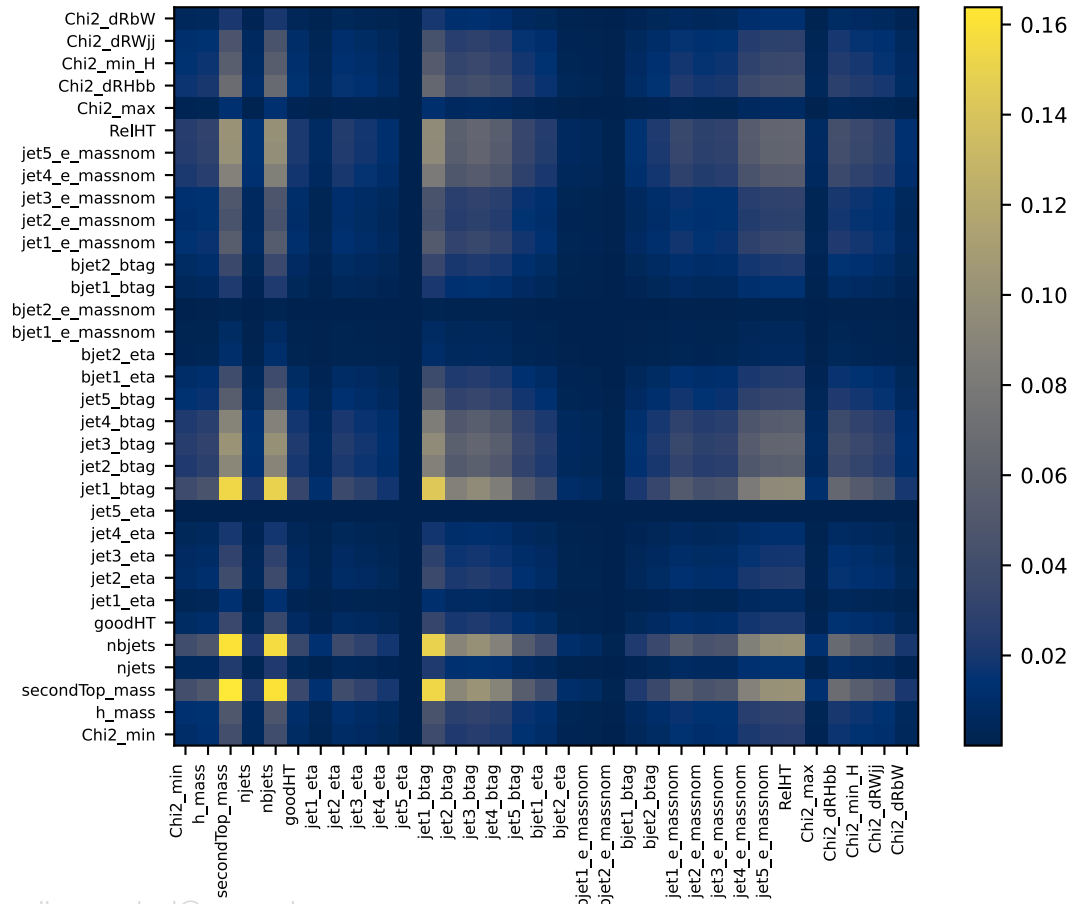
- feature importance in evaluation set
- Importance is normalized by 1
- 1st order gradient
 - weight
- Cut2 (Chi2_max) has lower importance than other cuts
- Cut1 (RelHT) has higher importance than others
- Will check for the other variable candidates



Taylor expansion for feature importance

Result

- feature importance in evaluation set
- Importance is normalized by 1
- 2nd order gradient
 - Hessian matrix
 - Correlation between features: how much a variable influence the other variables for the output function
- Will check for the other variable candidates



Rate in HLT path

Motivation

- LHC Beam line is displaced a bit → will be fixed soon
- Triggering as displaced jet would cause higher rate in b-tag path while no effect in photon

Target path

- HLT_Photon30EB_TightID_TightIso
- HLT_PFHT400_FivePFJet_100_100_60_30_30_DoublePFBTagDeepJet_4p5
- HLT_PFHT400_FivePFJet_100_100_60_30_30_DoublePFBTagDeepCSV_4p5
- HLT_PFHT400_FivePFJet_120_120_60_30_30_DoublePFBTagDeepJet_4p5
- HLT_PFHT400_FivePFJet_120_120_60_30_30_DoublePFBTagDeepCSV_4p5

Investigate recent runs (checked at 28.07.2022)

- Run [356323](#), [356321](#) (PU ~ 37, Fill [8059](#))
- Run [356316](#) (PU ~27, Fill [8058](#)) (CTPPS, CTPPS_TOT, GEM out)
- Run [356309](#) (PU ~34, Fill [8057](#))

Path/Run (Rate in Hz)	356323	356321	356316	356309
Photon30EB_TightID_TightIso	23.78	24.07	14.05	23.09
PFHT400_FivePFJet_100_100_60_30_30_*DeepJet	4.71	4.88	2.59	4.19
PFHT400_FivePFJet_100_100_60_30_30_*DeepCSV	8.82	8.75	4.63	7.76
PFHT400_FivePFJet_120_120_60_30_30_*DeepJet	3.10	3.25	1.72	2.89
PFHT400_FivePFJet_120_120_60_30_30_*DeepCSV	6.12	6.08	3.23	5.42

Rate in HLT path

Motivation

- LHC Beam line is displaced a bit → will be fixed soon
- Triggering as displaced jet would cause higher rate in b-tag path while no effect in photon

Target path

- HLT_Photon30EB_TightID_TightIso
- HLT_PFHT400_FivePFJet_100_100_60_30_30_DoublePFBTagDeepJet_4p5
- HLT_PFHT400_FivePFJet_100_100_60_30_30_DoublePFBTagDeepCSV_4p5
- HLT_PFHT400_FivePFJet_120_120_60_30_30_DoublePFBTagDeepJet_4p5
- HLT_PFHT400_FivePFJet_120_120_60_30_30_DoublePFBTagDeepCSV_4p5

Investigate recent runs (checked at 25.08.2022)

- Run [357542](#) (PU ~47, Fill [8132](#))
- Run [357329](#) (PU ~50, Fill [8118](#))

Path/Run (Rate in Hz)	357329	357542
Photon30EB_TightID_TightIso	10.12	10.82
PFHT400_FivePFJet_100_100_60_30_30_*DeepJet	19.21	23.10
PFHT400_FivePFJet_100_100_60_30_30_*DeepCSV	47.32	66.48
PFHT400_FivePFJet_120_120_60_30_30_*DeepJet	12.94	15.74
PFHT400_FivePFJet_120_120_60_30_30_*DeepCSV	33.38	46.78

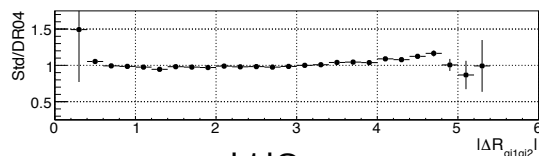
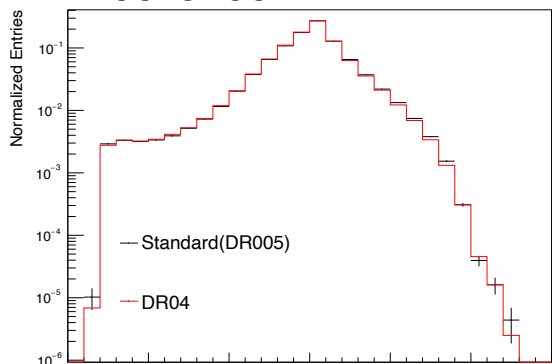
Comments for JERC meeting

Comparison between Gjets standard and Gjets DR cut

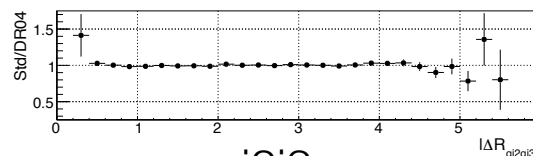
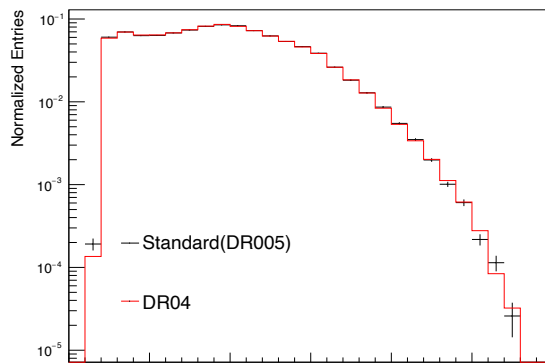
- 2nd and 3rd jet dR
- HT100To200

Recophoton == 1
RecoJet >=1

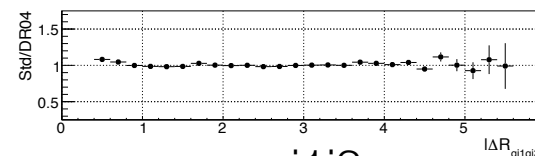
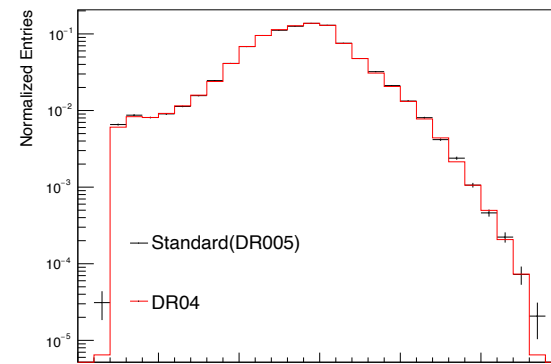
GEN



j1j2

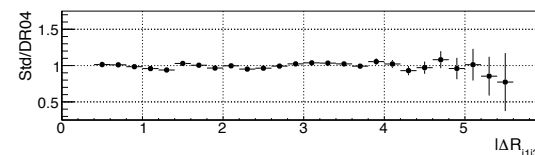
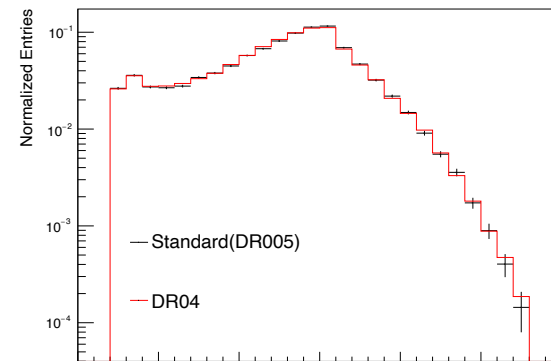
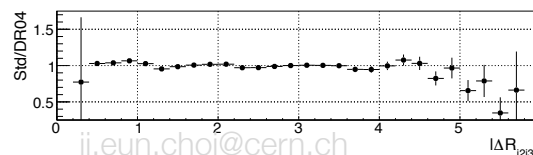
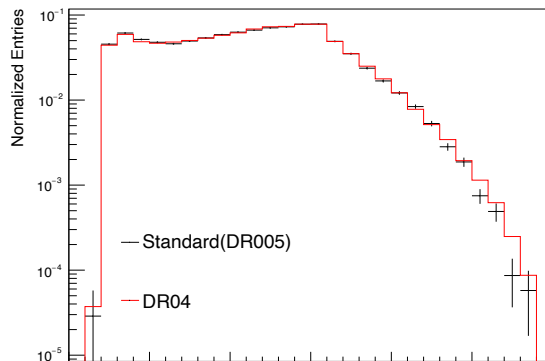
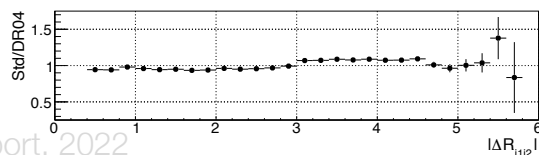
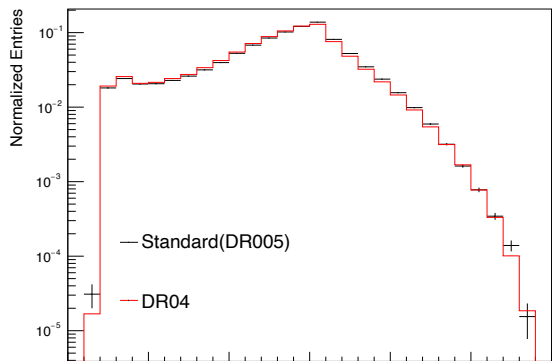


j2j3



j1j3

RECO



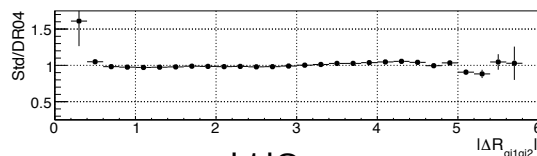
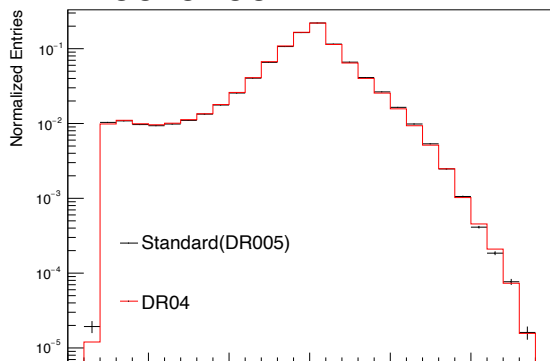
Comments for JERC meeting

Comparison between Gjets standard and Gjets DR cut

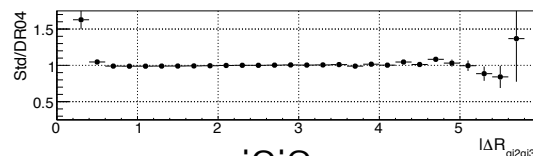
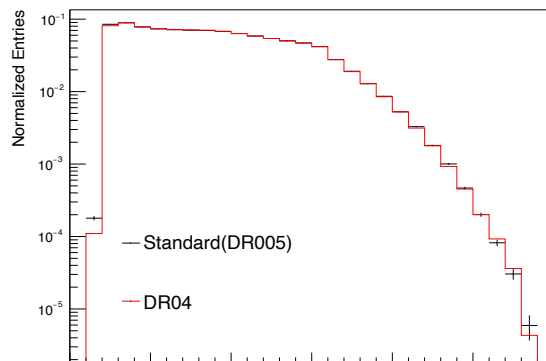
- 2nd and 3rd jet dR
- HT200To400

Recophoton == 1
RecoJet >=1

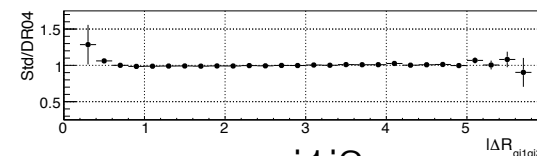
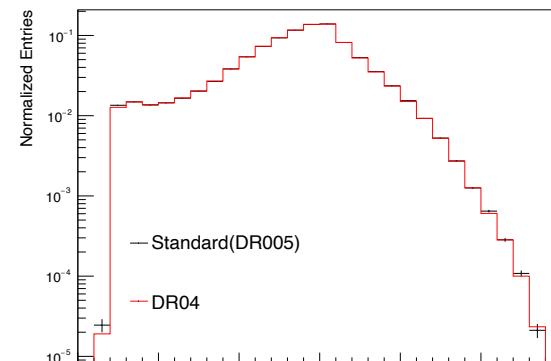
GEN



j1j2

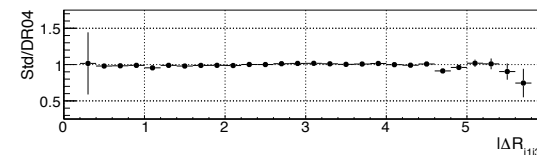
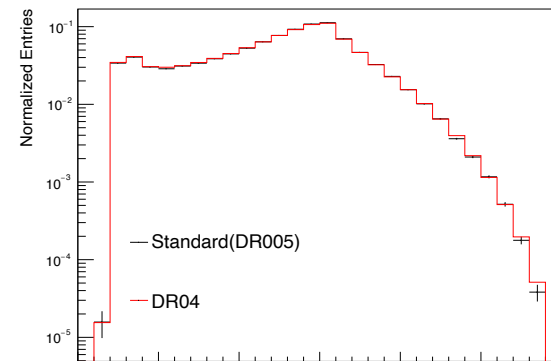
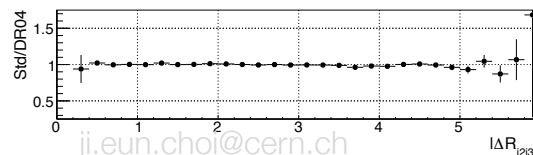
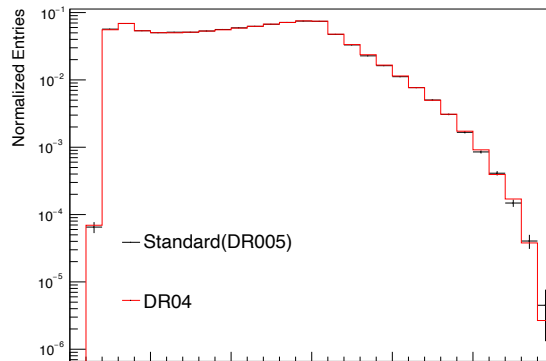
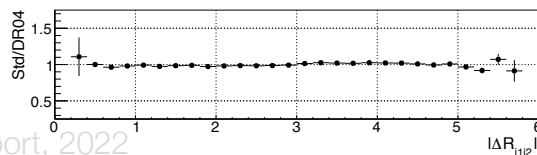
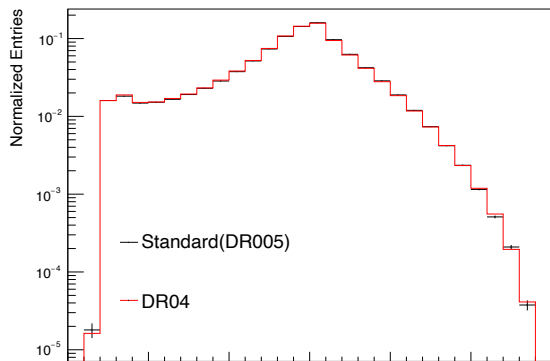


j2j3



j1j3

RECO



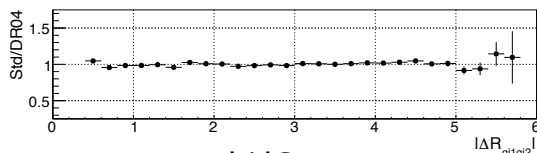
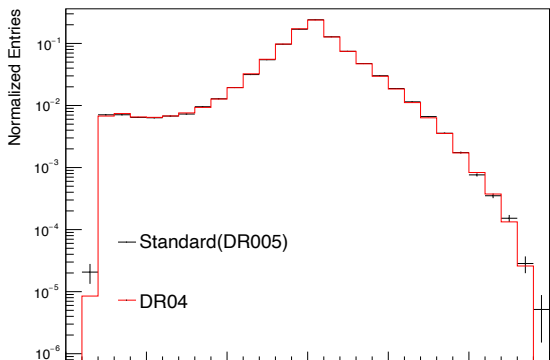
Comments for JERC meeting

Comparison between Gjets standard and Gjets DR cut

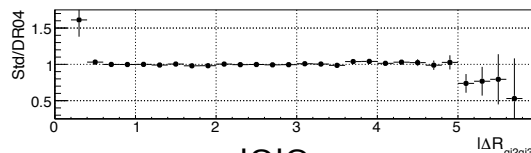
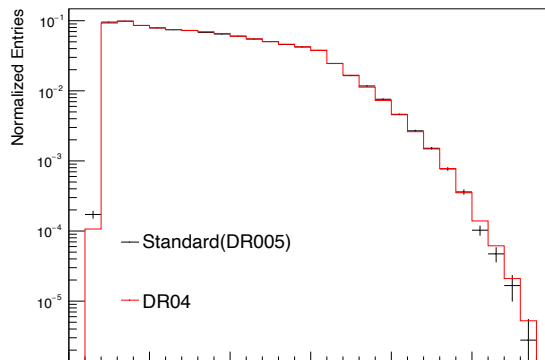
- 2nd and 3rd jet dR
- HT400To600

Recophoton == 1
RecoJet >=1

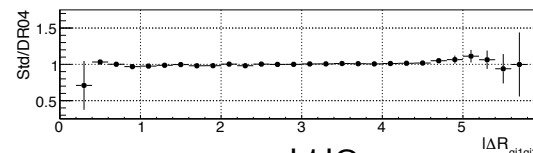
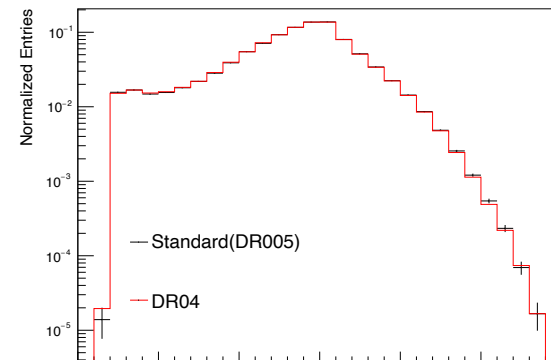
GEN



j1j2

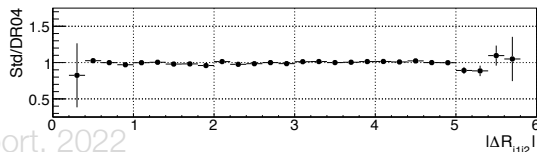
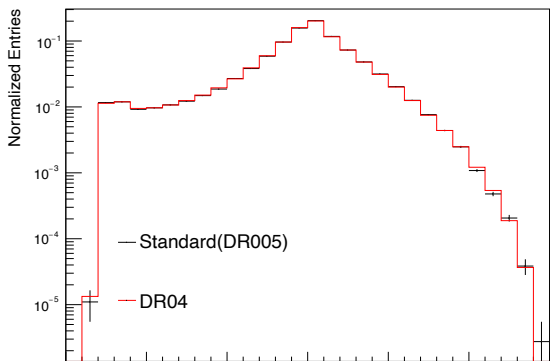


j2j3

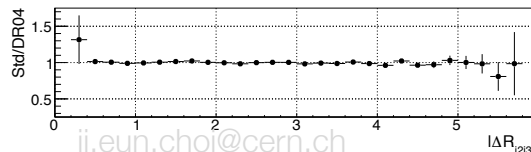
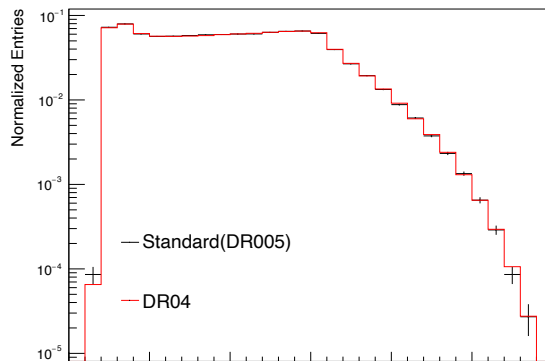


j1j3

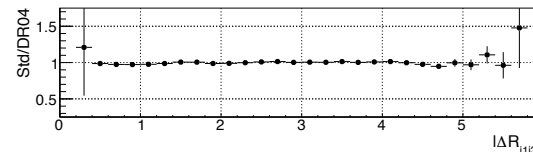
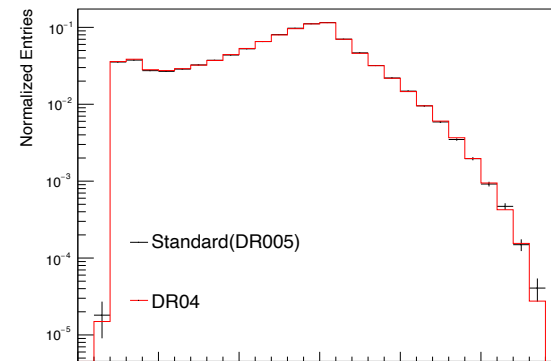
RECO



j1j2



j2j3



j1j3

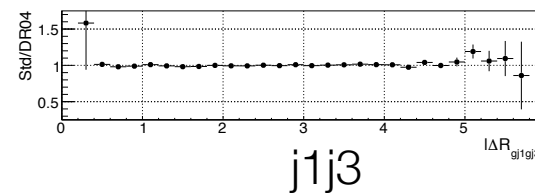
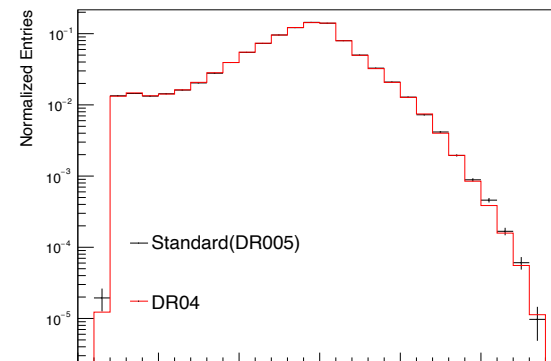
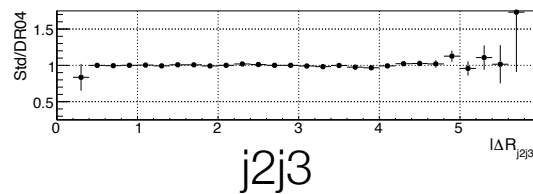
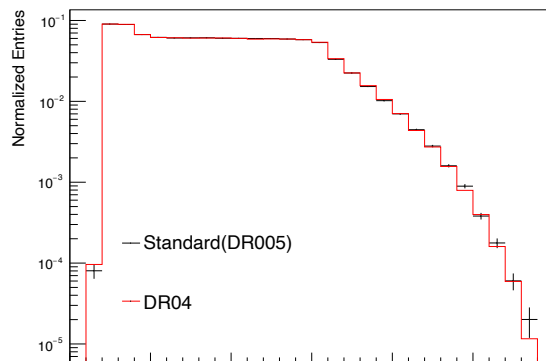
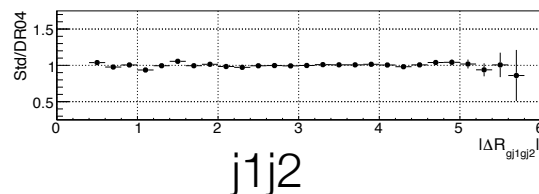
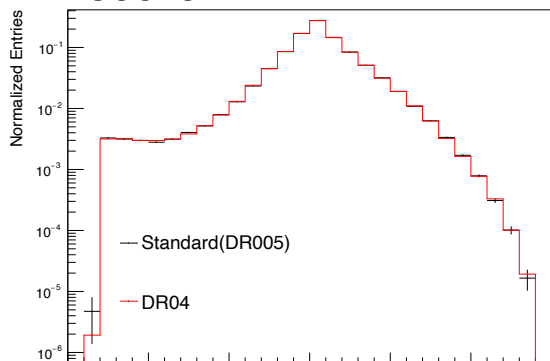
Comments for JERC meeting

Recophoton == 1
RecoJet >= 1

Comparison between Gjets standard and Gjets DR cut

- 2nd and 3rd jet dR
- HT600ToInf

GEN



RECO

