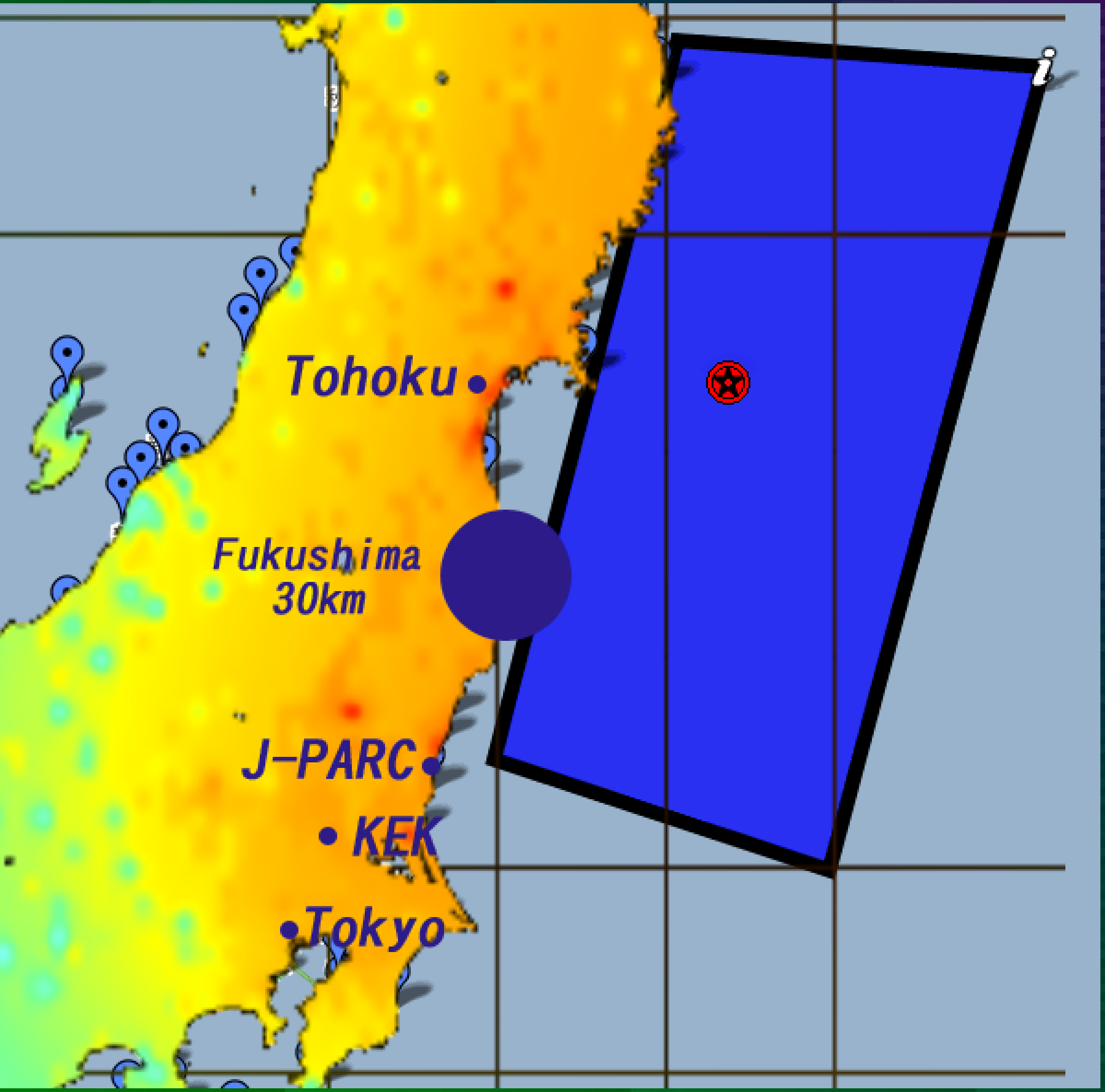
Status of LCFIVertex: vertex finding, jet clustering and flavor-tagging neural network

Taikan Suehara, Tomohiko Tanabe

(The Univ. of Tokyo)

The earthquake



The M9.0 earthquake gives…

- Some damage in Tohoku & KEK (worse in J-PARC)
- Nearly no damage in Tokyo

None of our facility are affected by the Fukushima accident

but…

…and severe electric power shortage in Tokyo and KEK…

- Electricity in KEK stopped for several days
  - Now partially recovered
- Severe power shortage happened in Tepco area (including Tokyo & KEK)
  - Tepco operated Fukushima plants
  - ‘Planned periodic power outage’ now
  - All power-consuming facilities must be off now
    - Including KEK computing facilities
      -> We cannot access most of MC data in KEK!
  - Now trying to use GRID…

Topics

- Neural-net optimization (Tomohiko)
  - Reconstruction of LCFI framework
  - Optimization of input variables
  - Results
- New jet clustering for multi-b environment(me)
  - Vertex finder without jet clustering
  - Good vertex selection
  - Jet clustering with vertex information
  - Results

Status of LCFIVertex I: Progress in vertex finding, flavor-tagging neural net framework

Tomohiko Tanabe

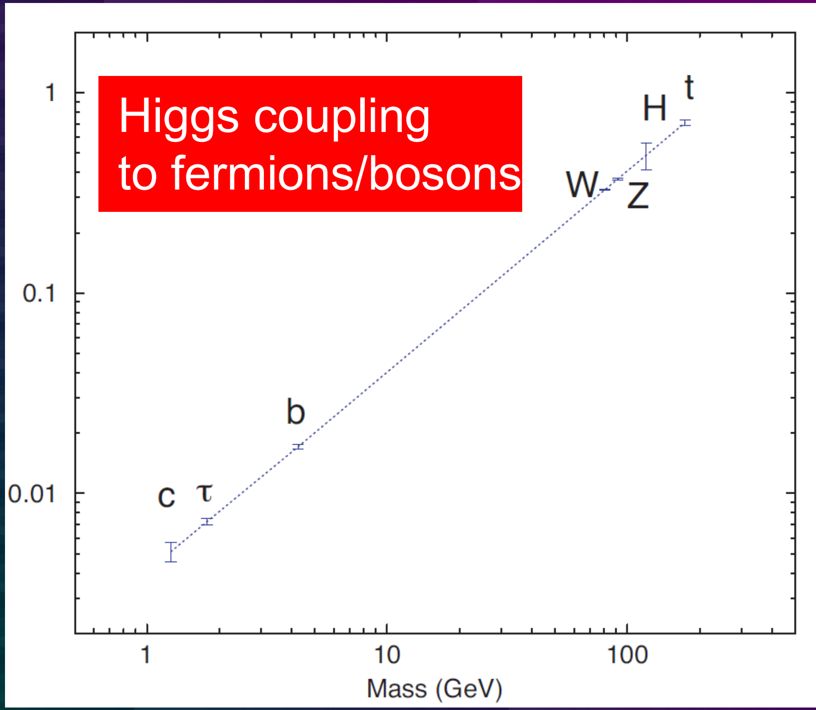
Taikan Suehara, Satoru Yamashita
(The Univ. of Tokyo)

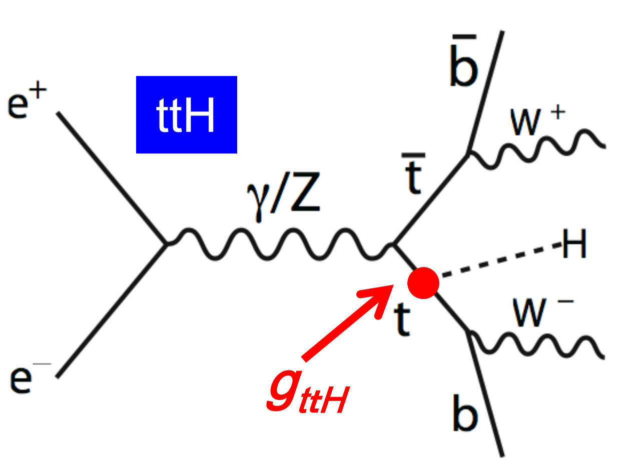
in this talk

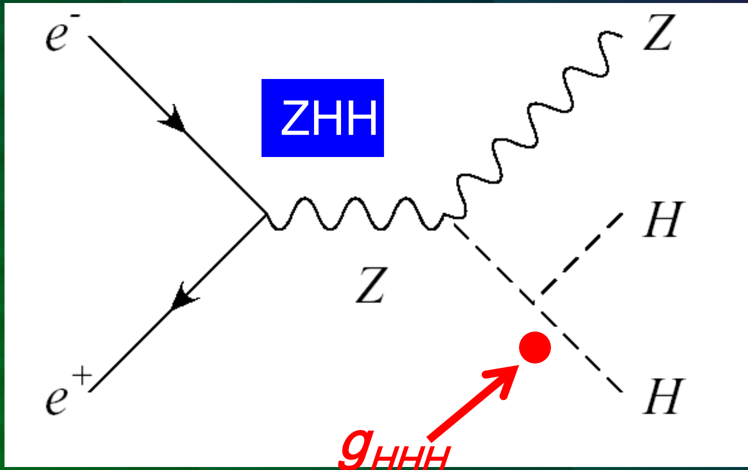
- introduction
- vertex finding improvements
  - efficiency increase in the near-IP region
- flavor tagging improvements
  - new variables
  - improved signal/background ratio

motivation

- Many important physics processes have multiple heavy flavor jets
  - Higgs BF : H -> bb, H -> cc
  - Higgs self-coupling : ZHH -> qqbbbb
  - top-Yukawa : ttH -> bWbWbb
  - top physics : tt -> bWbW



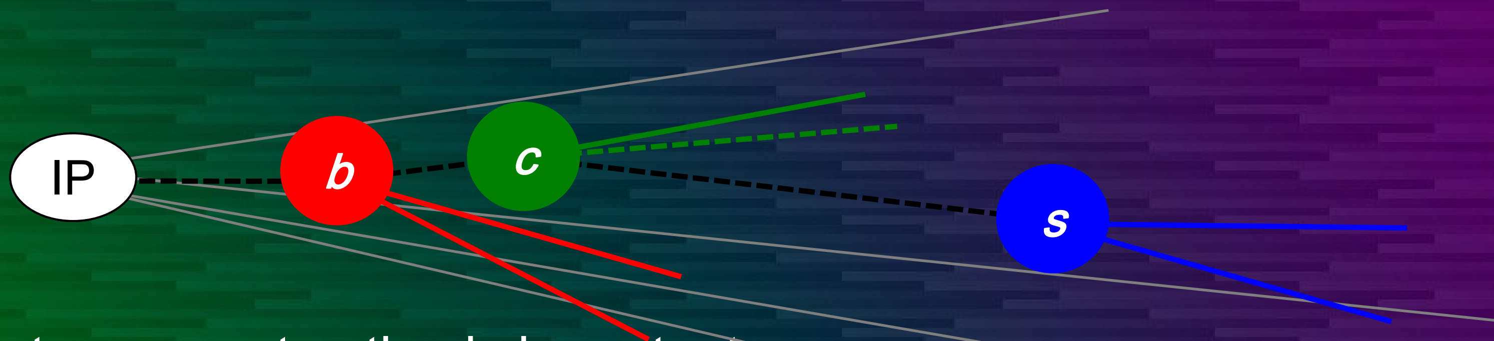




Z→qq (70%)
ll (30%)
W→qq (65%)
lν (35%)
H→bb (65%)
(m_H=120GeV)

ideal flavor tagging

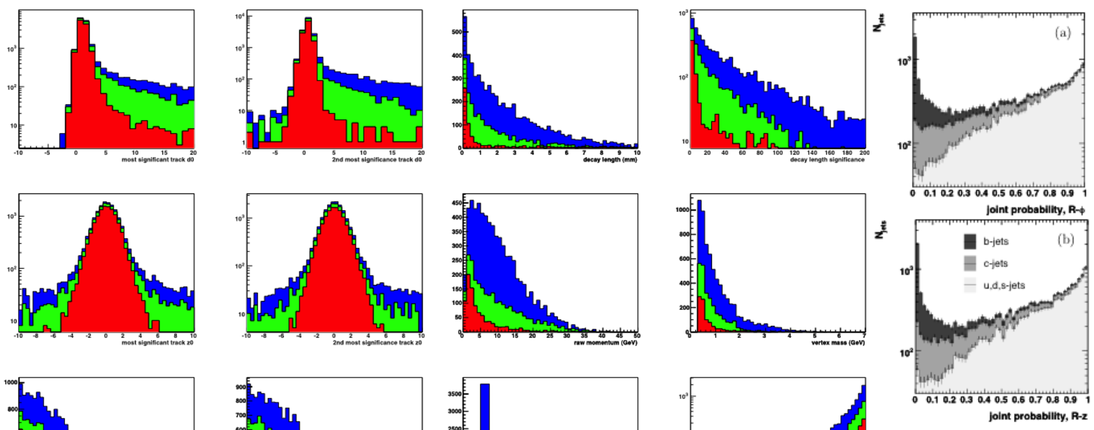
- reconstruct the entire decay chain (b -> c -> s) in a jet



  - vertex reconstruction is important
    - but vertices cannot be made from a single track
      - use track measurements (impact parameter)
  - presence of neutral particles
    - missing correction by using p_T
  - lepton ID: energetic/isolated leptons is a sign of heavy quark decays
- key is variable combination
  - likelihood, multivariate analysis (e.g. ANN)
  - event categorization (discrete variables)

LCFIVertex

LCFIVertex is a collection of algorithms for flavor tagging and parton charge identification.



input variables (tracks)

input variables (vertex)

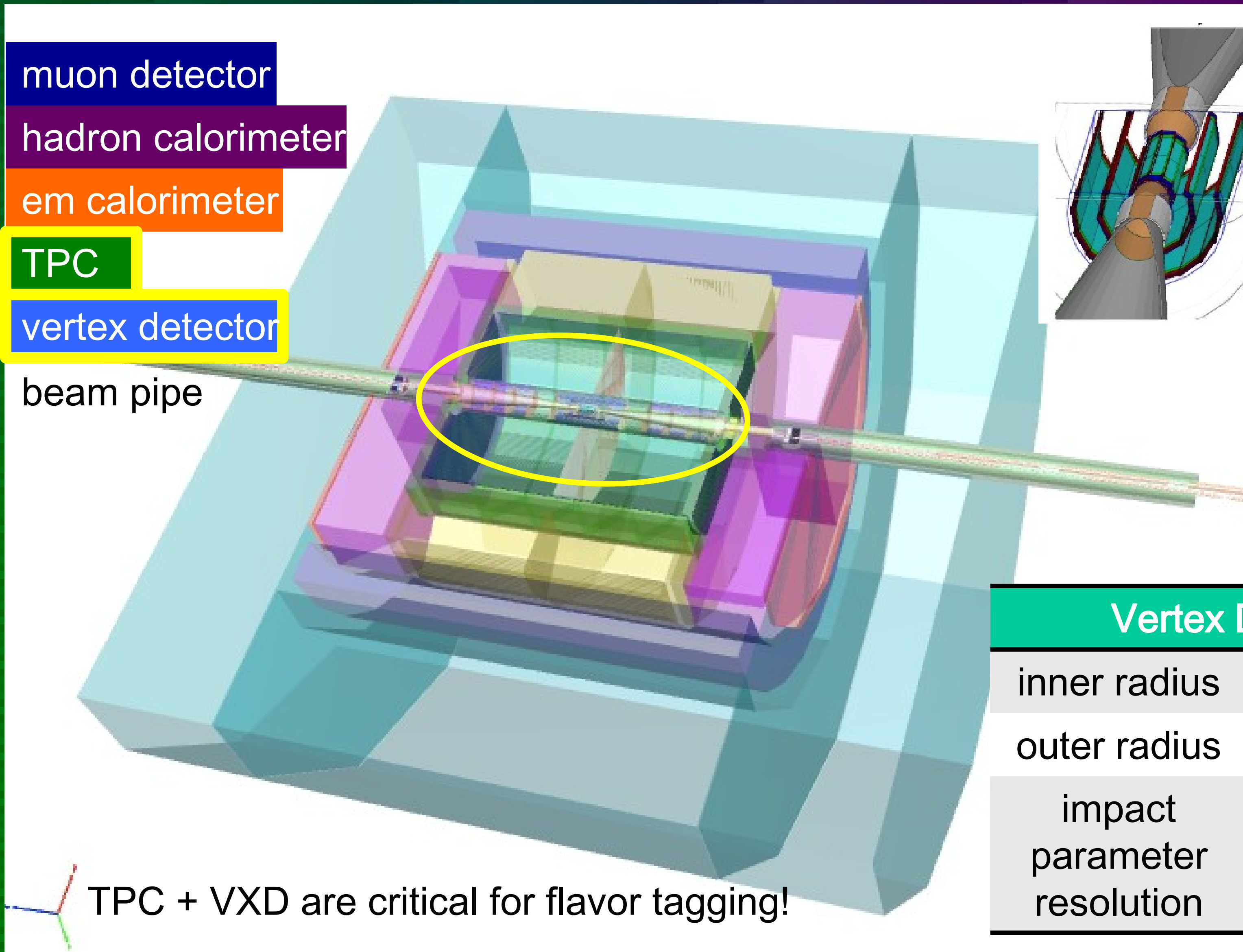
joint prob.

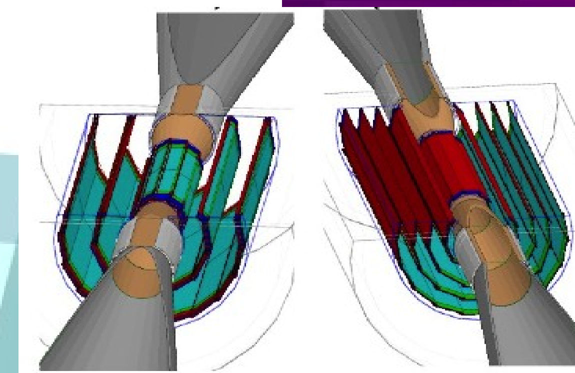
neural net

b-likeness
c-likeness
parton charge

LCFI Collaboration: NIM A 610 (573) [arxiv:0908.3019]

ILD Detector





TPC + VXD are critical for flavor tagging!

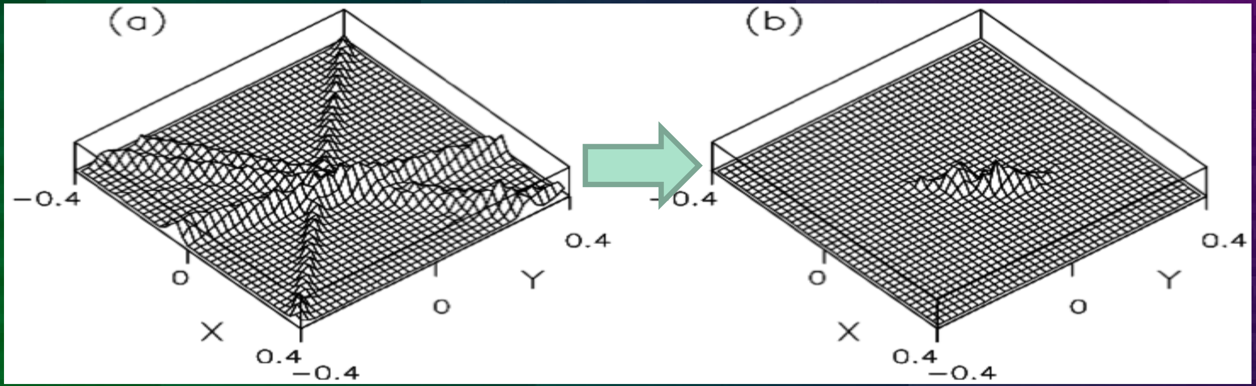
Vertex Detector	
inner radius	15 mm
outer radius	60 mm
impact parameter resolution	< 5 mm (high momentum)

vertex finding algorithms

- topological vertex finder (ZVTOP)

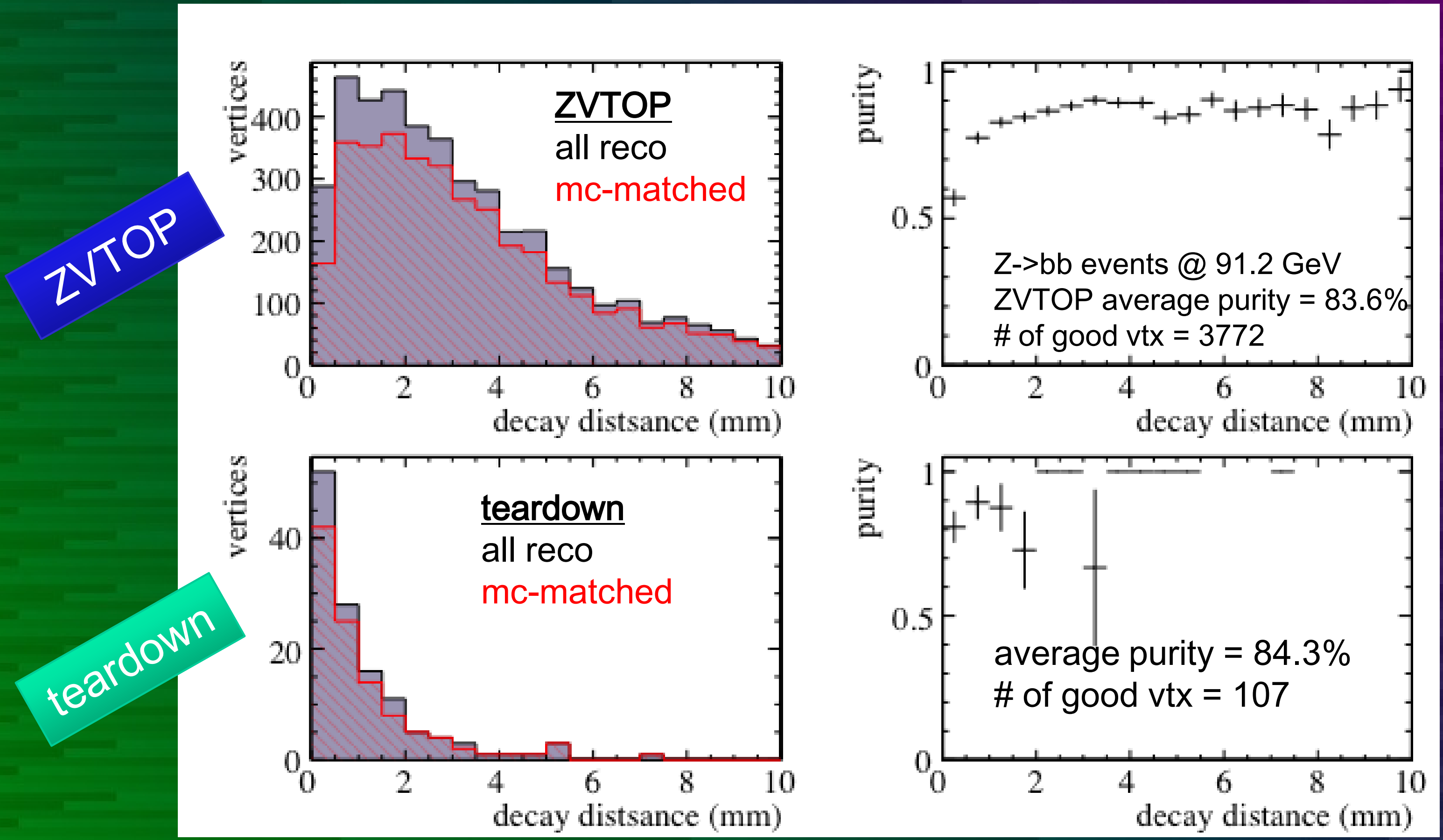
track probability tubes

vertex function: overlapping tubes -> high values



  - can find vertices for arbitrary topology with any number of tracks（must be separated from the primary vertex）
  - it takes lots of CPU time due to having to evaluate the vertex function at many points in space
- tear-down
  - start from a set of tracks, remove tracks which are inconsistent (large chi-squared contribution)
  - if the primary tracks are properly removed, vertices can be found with high efficiency
- build-up
  - using track pairs as seed, attach other tracks
  - good seeds lead to good vertices

ZVTOP + teardown



ZVTOP is not good at vertices near the IP.
Teardown is performed using tracks not used by ZVTOP.
After quality cuts (dist, chi2, ntrks), we find +3% efficiency increase for the same purity.

rejection of V^0 particles

- despite having V^0 taggers in the Marlin reconstruction chain, our vertex finders still find V^0's (K_S, Lambda, conversions) for two-track vertices
- we apply the following cuts to reject V^0's (reduce uds contamination):
  - cut on the angle θ between the vertex displacement from IP and the V^0 direction
  - mass requirements
  - K_S: $\cos\theta > 0.999$ & mass 15 MeV within PDG value
  - Lambda: $\cos\theta > 0.99995$ & mass 20 MeV within PDG value
  - conversions: $\cos\theta > 0.99995$ & less than 10 MeV for conversion mass, where the mass is geometrically corrected so that it is calculated using the track dip angles

$$m^2_{\mathrm{conv}} = 2|\vec{p}_1||\vec{p}_2|(1 - \cos\Delta\lambda_{12})$$

	before cut	after cut
K_S	3205	623
Lambda	1482	371
conversions	2544	278
other two-track reco vertices	30747	30333

multivariate analysis

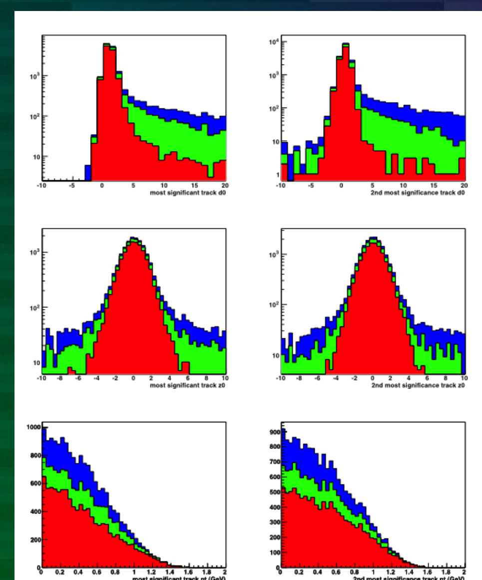
- LCFI's original neural net implementation is difficult to maintain/extend
- we replaced it by TMVA (ANN and BDT)
  - a series of scripts has been written to perform the following tasks:
    - normalize input variables
    - split the data into training and testing samples
    - separate samples into several categories for individual training by the TMVA classifiers
    - combining output of the individual classifiers
    - evaluate the performance using the testing samples

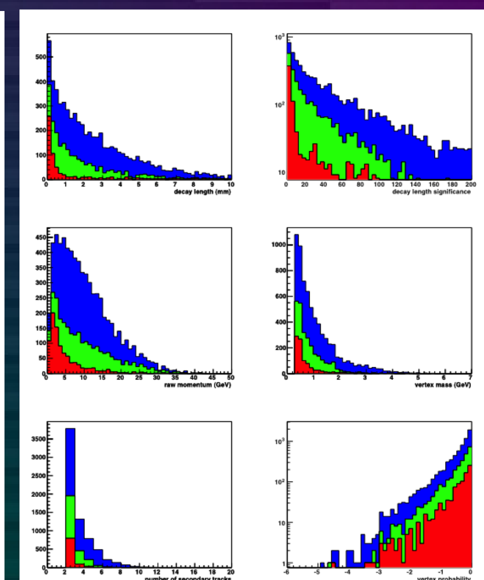
LCFI input variables

- previously in LCFI:
  - three categories, trained independently:
    - # vertex = 0
    - # vertex = 1
    - # vertex >= 2
  - for # vertex = 0 (8 variables):
    - d_0 impact parameter (1)
    - d_0 impact parameter (2)
    - z_0 impact parameter (1)
    - z_0 impact parameter (2)
    - track momentum (1)
    - track momentum (2)
    - d_0 joint probability
    - z_0 joint probability
  - for # vertex = 1, >=2 (8 variables):
    - d_0 joint probability
    - z_0 joint probability
    - vertex decay length
    - vertex decay length significance
    - vertex momentum
    - pt-corrected vertex mass
    - vertex multiplicity
    - vertex probability from the fitter

(1) and (2) indicate the most and second most significant track.

"joint probability" – probability that a track comes from the IP, computed *a priori* using the distribution of impact parameter significance (separately for d_0 and z_0), multiplied for all tracks in the jet





input variables

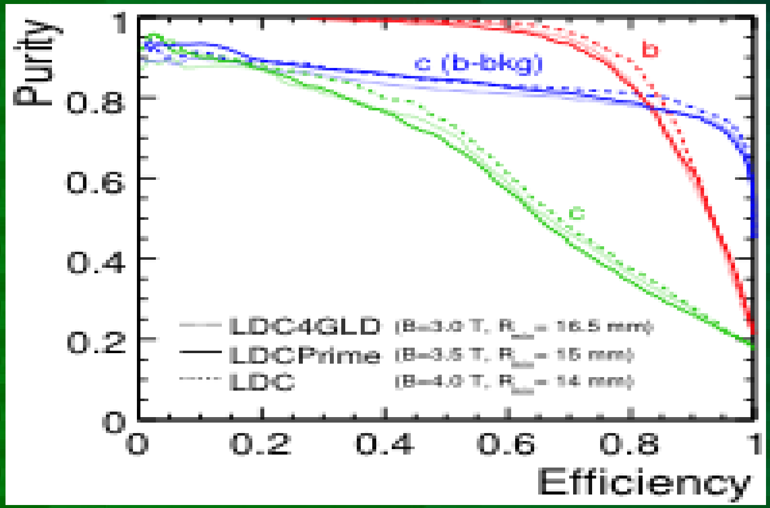
- additional variables
  - boosted sphericity
  - vertex displacement/momentum angle
  - vertex mass (not pt corrected)
  - mass, momentum, decay distance, decay significance for the first and second (if found) vertices
  - distance, significance, and displacement/momentum angle between the first and second vertices

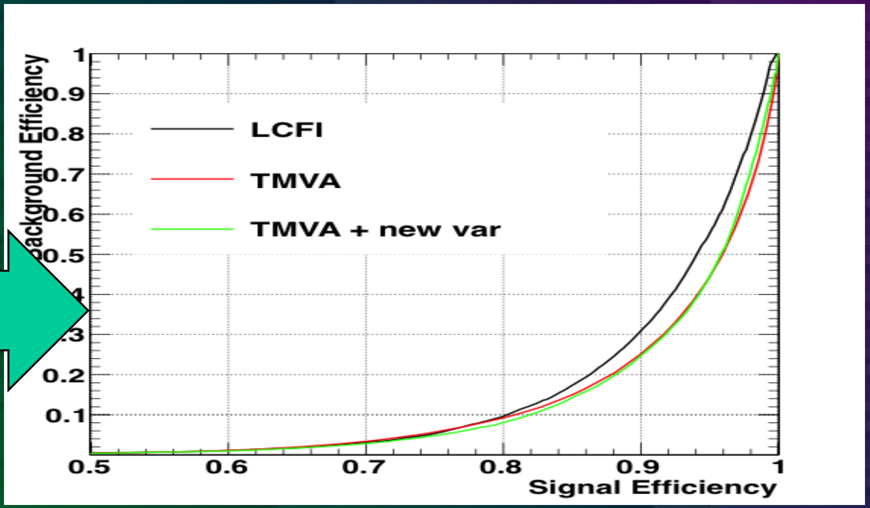
new input variables

- the new variables are trained by BDT
  - for # vertex = 0 (9 variables):
    - d_0 impact parameter (1)
    - d_0 impact parameter (2)
    - z_0 impact parameter (1)
    - z_0 impact parameter (2)
    - track momentum (1)
    - track momentum (2)
    - d_0 joint probability
    - z_0 joint probability
    - boosted sphericity
  - for # vertex = 1 (17 variables):
    - d_0 impact parameter (1)
    - d_0 impact parameter (2)
    - z_0 impact parameter (1)
    - z_0 impact parameter (2)
    - track momentum (1)
    - track momentum (2)
    - d_0 joint probability
    - z_0 joint probability
    - boosted sphericity
    - vertex decay length
    - vertex decay length significance
    - vertex momentum
    - vertex mass (pt-corrected)
    - vertex mass (not pt-corrected)
    - vertex multiplicity
    - vertex probability from the fitter
    - vertex disp/momentum angle
  - for # vertex >= 2 (29 variables):
    - d_0 impact parameter (1)
    - d_0 impact parameter (2)
    - z_0 impact parameter (1)
    - z_0 impact parameter (2)
    - track momentum (1)
    - track momentum (2)
    - d_0 joint probability
    - z_0 joint probability
    - boosted sphericity
    - vertex #1 decay length
    - vertex #2 decay length
    - distance between vertex #1 & #2
    - vertex #1 decay length significance
    - vertex #2 decay length significance
    - separation significance between vertex #1 & #2
    - vertex #1 momentum
    - vertex #2 momentum
    - vertex momentum (combined)
    - vertex #1 mass (not pt-corrected)
    - vertex #2 mass (not pt-corrected)
    - vertex mass (combined, pt-corrected)
    - vertex #1 multiplicity
    - vertex #2 multiplicity
    - vertex multiplicity (combined)
    - vertex probability from the fitter
    - vertex #1 disp/momentum angle
    - vertex #2 disp/momentum angle
    - vertex #1/#2 disp/momentum angle
    - vertex #1/#2 angle

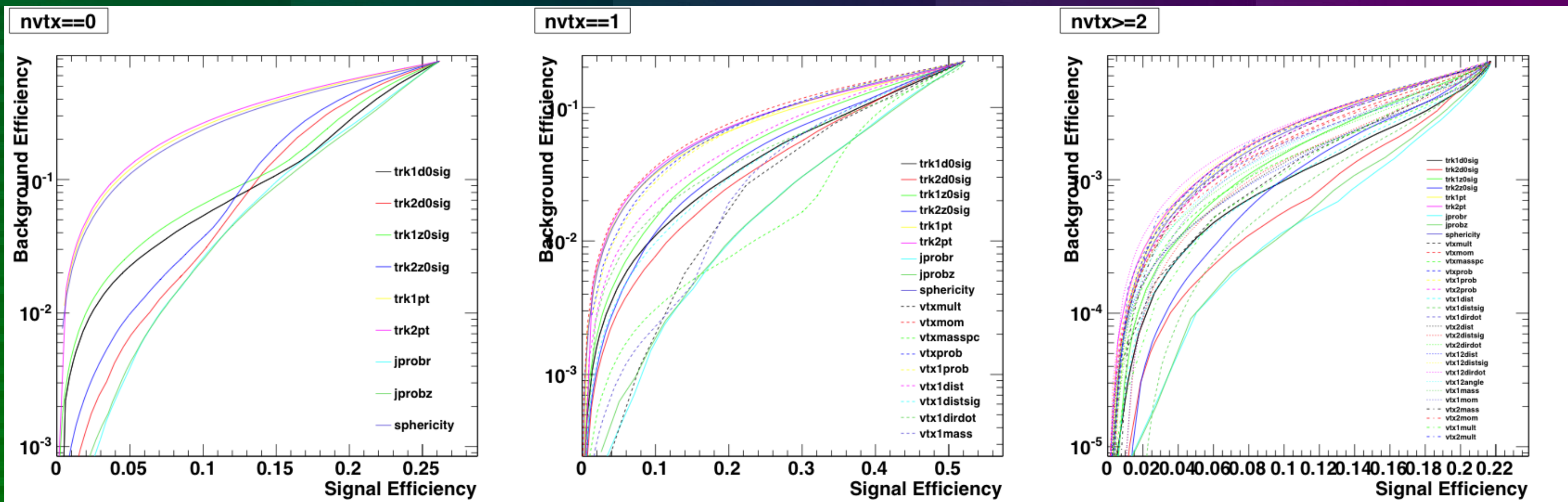
evaluating classifier response

- LoI flavor-tagging evaluation produced purity-efficiency plots
  - but this depends on the fraction of heavy jets, which changes from sample to sample
    - BF(Z->bb)=15%, BF($H_{120\ GeV}$->bb)=68%
- better to use a fraction-independent measure: evaluate using background efficiency versus signal efficiency instead



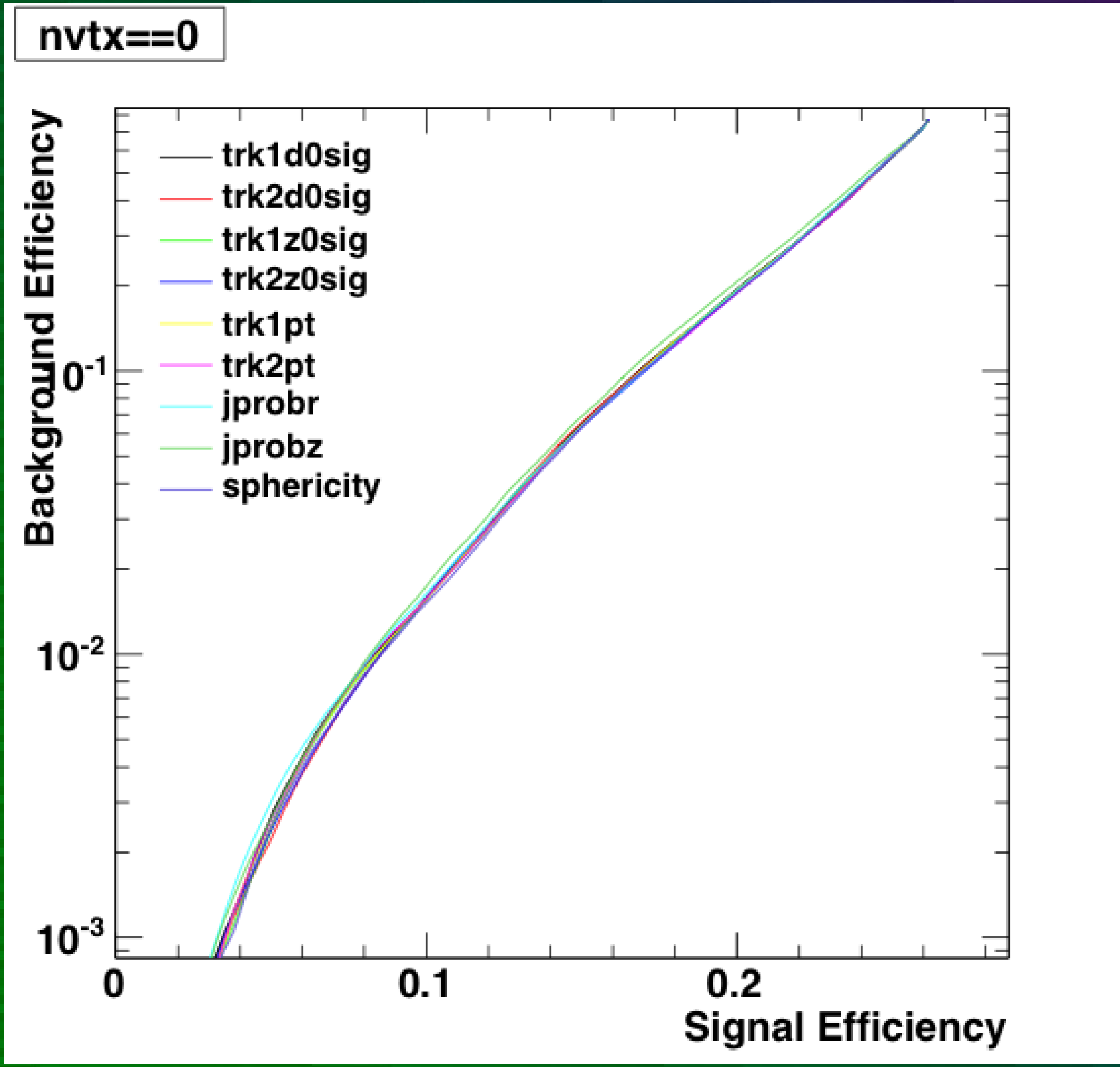


variable ranking



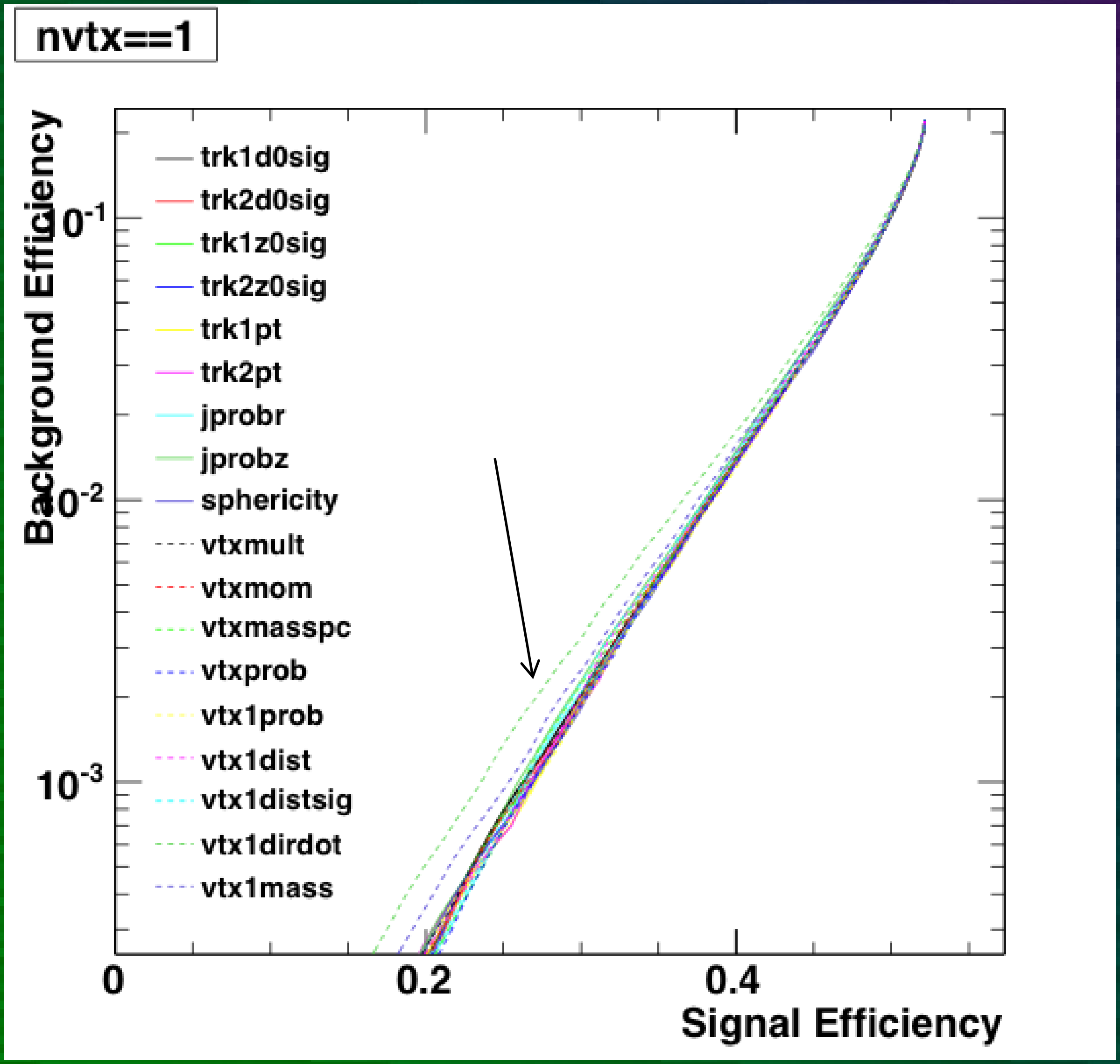
- single variable ranking shows the most useful variables (on their own)
  - joint probabilities, vertex mass
- however, any other uncorrelated variable can help; these plots do not show this effect

variable ranking (correlation)



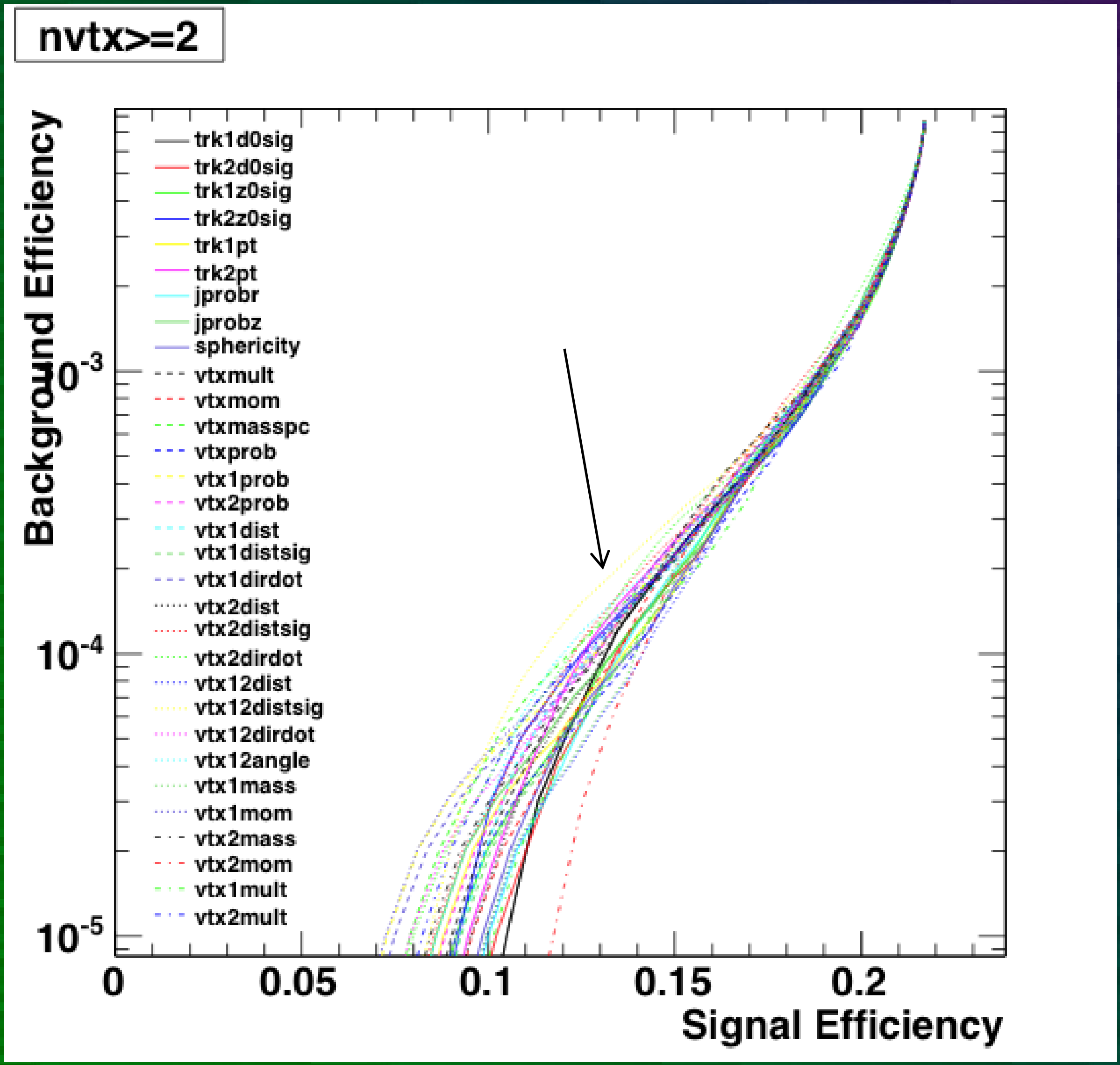
- result of training after removing a variable
- this shows how "unique" this variable is in terms of uncorrelated classifying power
- significantly worse performance after removing the variable shows that it's effective
- for nvtx=0, joint probabilities (both d_0 & z_0) are the most powerful as expected

variable ranking (correlation)



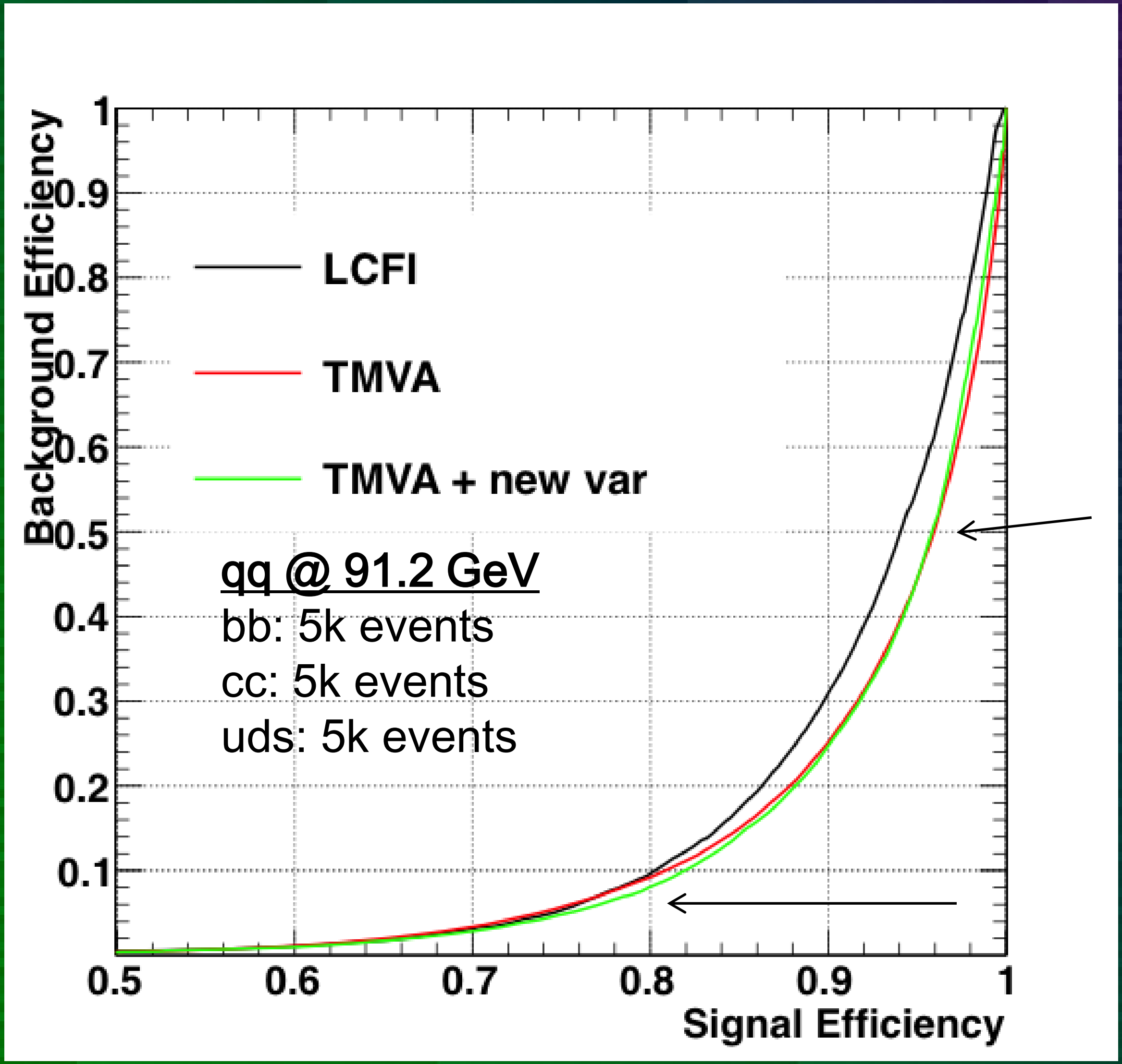
- for nvtx=1, the most effective variables are:
  - displacement/momentum angle of the vertex
  - uncorrected mass of the vertex
- newly added variables are shown to be effective!!!

variable ranking (correlation)



- for nvtx>=2, the most effective variables are:
    - separation significance between the 1st and 2nd vertices
- AGAIN: newly added variables are shown to be effective!!!

results

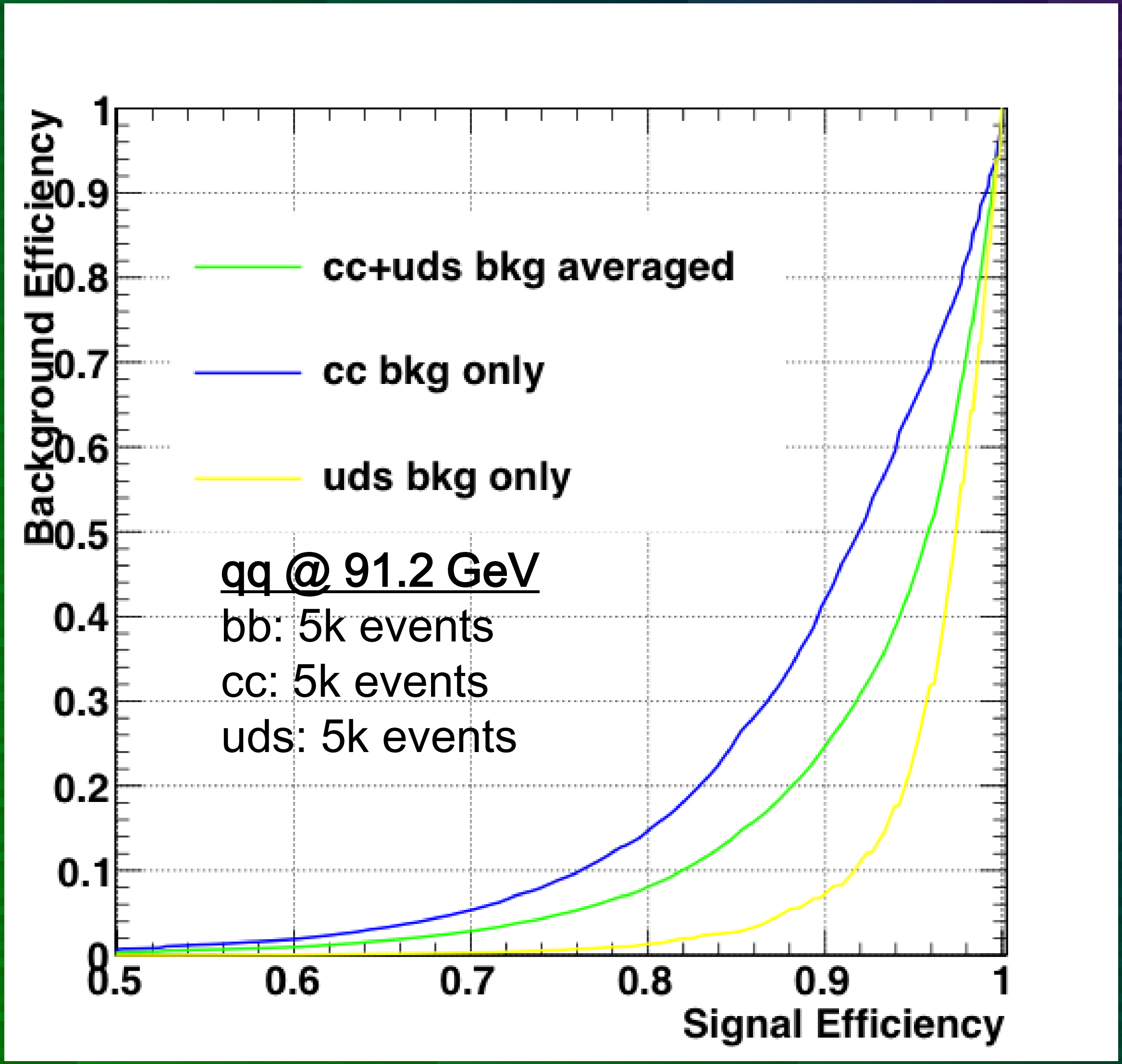


background = cc & uds mixed equally

there is already improvement merely by switching to TMVA (signal eff > 0.8)

more improvement by adding new variables (signal eff > 0.75)

results



uds rejection is pretty good

need more work on charm

summary (part I)

- we have a mature analysis framework to optimize flavor tagging
  - improvements over LCFI: ability to add arbitrary number of categories and input variables, choice of training method (via TMVA)
- performance for qq @ 91.2 GeV has been improved
- next:
  - leptons!
  - dedicated charm tagging
  - optimization at higher jet energies
    - ZHH jets, qq @ 500 GeV
    - qq @ CLIC energies
    - samples have been provided by A. Timoce and S. Posse (thanks, will look at them asap)

Status of LCFIVertex II: Application to a many-jet environment with a new jet clustering algorithm

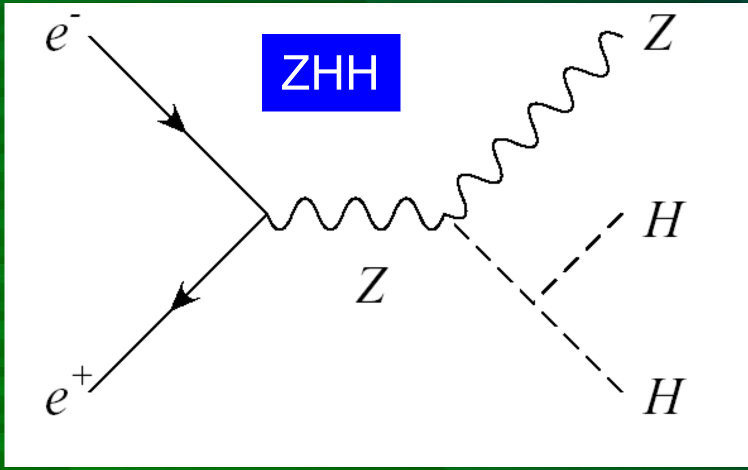
Taikan Suehara

Tomohiko Tanabe, Satoru Yamashita
(The Univ. of Tokyo)

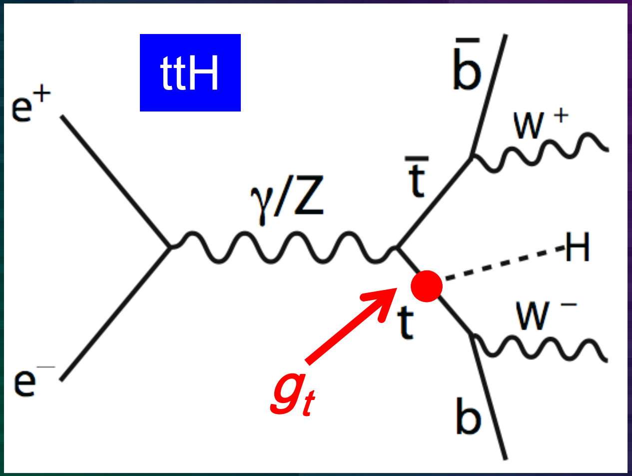
Multi-jet flavor analysis

- Many LC physics targets have
  - Many jets (4, 6, 8, …)
  - Many b jets (from H->bb, t->bW etc.)

➡ b-tagging in multi-jet environment



qqbbbb (32%) qqbbWW(14%)
ννbbbb (9%) etc. (120GeV Higgs)



bbbbqqqq(31%)
bbbbqqlν(31%) etc.

B-tagging in ZHH (Z->qq)

Y. Takubo, ALCPG09

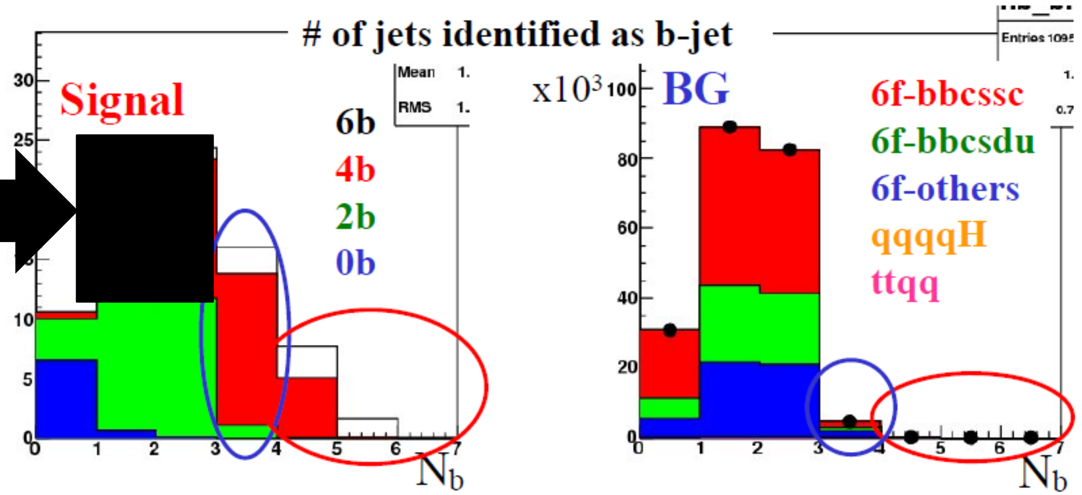
Selection of event samples

BG can be rejected effectively by using events with $N_b \geq 3$.

	No cut	$N_b = 3$	$N_b \geq 4$
Signal	79	15.9(0.20)	9.5(0.12)
BG	207,144	4663(0.02)	147(7 x 10^{-4})

σ = 0.16 fb

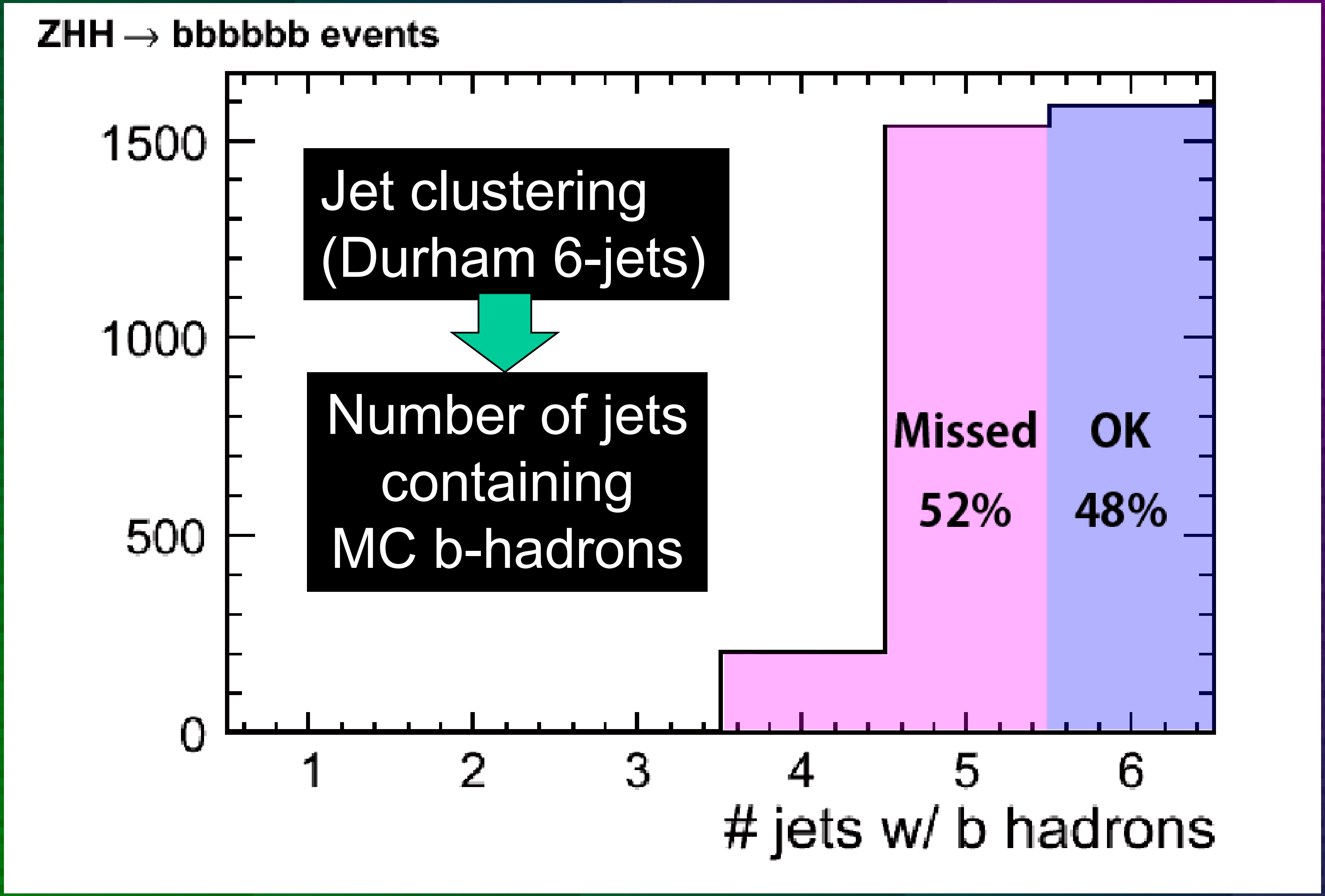
Events with $N_b = 3$ and $N_b \geq 4$ were selected as analysis samples.



Should be rescued…

b-tagging efficiency seems to be bad…

'# of b jets' in ZHH



# of b-jets is reduced due to mis-jet-clustering.

The new jet clustering

Find b- (or c-) vertices

↓

Cluster around the vertices
Distribute vertices to separate jets

The procedure

Tracks

1. Vertex finding

Vertices

2. Vertex selection

Good vertices

muons → 3. Vertex combination

Combine b- and c-vertices

Combined vertices

PFOs → 4.Jet clustering

1. Vertex finding (1)

- Original jet finder based on “build-up” method
  - ZVTOP cannot be used without tuning
    - It’s designed to be used after jet clustering
    - Too many fakes – without “jet-direction” parameter

- IP tracks are firstly removed (tear-down)
- Vertices are calculated for every track-pair
  - Calculate nearest points of two helices
    - Geometric calculation for the start point
    - Minuit minimization using track error-matrices

1. Vertex finding (2)

- Pre-selection
  - Mass < 10 GeV (B: ~5 GeV)
  - Momentum & vertex pos: not opposite to IP
  - Vertex mass < energy of either track
    - This selection is very effective for dropping fakes
  - Vertex distance to IP > 0.3 mm
  - Track chi2 to the vertex < 25
- Associate more tracks to passed vertices
  - Using same criteria as above
- Sort & select obtained vertices by probability
  - Associated (# tracks >=3) vertices are prioritized
  - Example in next slide…

1. Vertex finding (3)

- Example: 5 vertices are found

Vertex #	Tracks included	Probability
1	1,2,3	0.9
2	2,4,5	0.4
3	3,4	0.8
4	5,6	0.6
5	6,7	0.5

1. Vertex finding (3)

- Example

Vertex #	Tracks included	Probability	
1	1,2,3	0.9	Adopted!
2	~~2~~,4,5	0.4 -> 0.6	
3	~~2~~,4	0.8	Removed!
4	5,6	0.7	
5	4,7	0.5	

1. Vertex finding (3)

- Example

Vertex #	Tracks included	Probability
1	1,2,3	0.9
4	5,6	0.7
2	4,5	0.6
5	4,7	0.5

Adopted!

1. Vertex finding (3)

- Example

Vertex #	Tracks included	Probability	
1	1,2,3	0.9	Adopted!
4	5,6	0.7	Adopted!
2	4,5	0.6	Removed!
5	4,7	0.5	

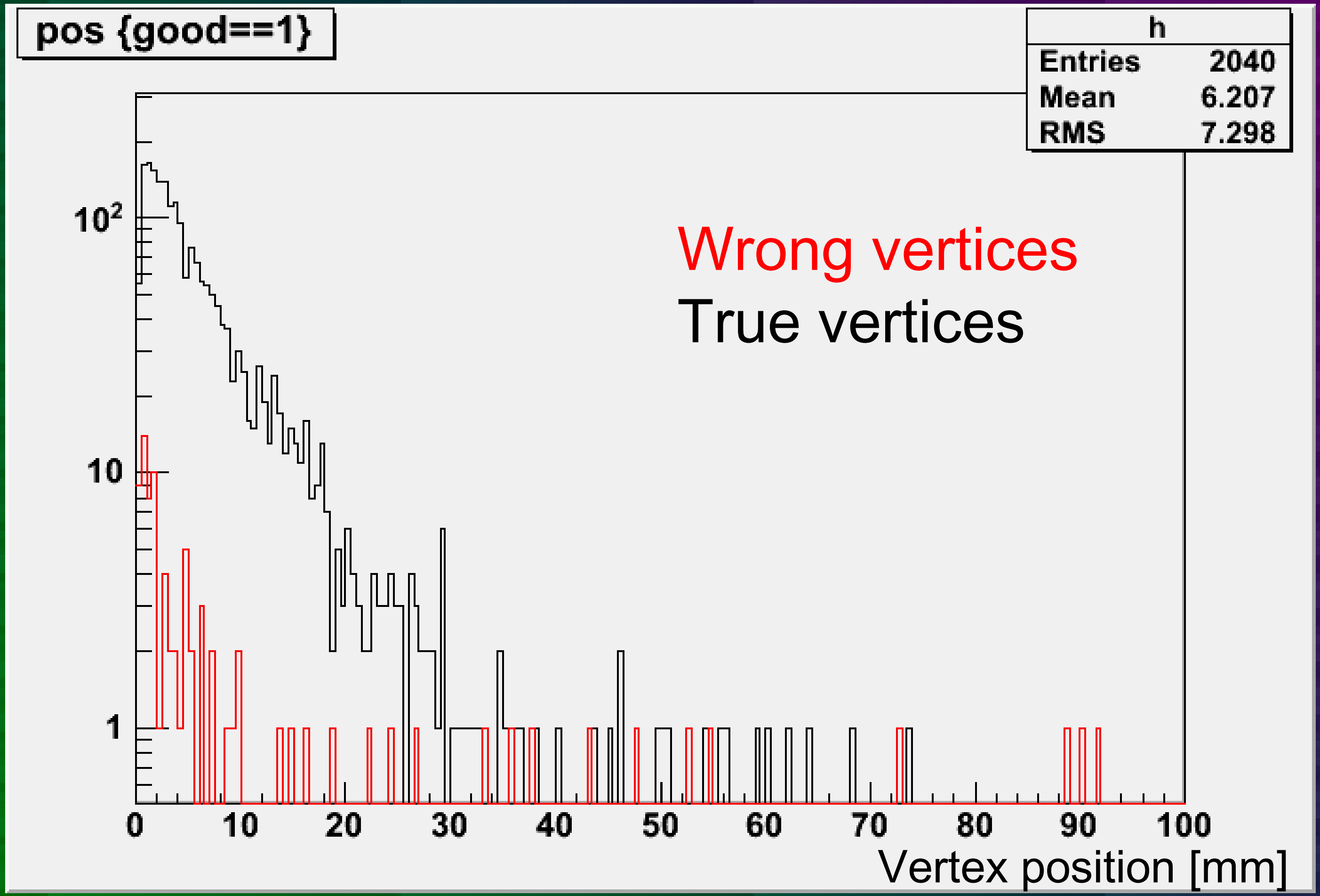
1. Vertex finding (3)

- Example

Vertex #	Tracks included	Probability	
1	1,2,3	0.9	Adopted!
4	5,6	0.7	Adopted!
5	4,7	0.5	Adopted!

Finally, three vertices are adopted.

Vertices in ZHH -> bbbbbb



In 347 bbbbbb events:

Good = 2040, bad (not from B-semistable) = 90

2. Vertex selection & muons

- Vertex selection
  - K0 vertices are removed (mass +- 10 GeV)
  - Vertex position > 30 mm are removed
    - Mostly s-vertices
- Secondary muons
  - Following tracks are treated as same as vertices
    - With muon hit
      - Currently > 50 MeV energy deposit
    - Impact parameter (> 5 sigma & not too much)
    - Ecal, Hcal energy deposit

Selection performance

- Vertex selection (347 bbbbbb events)

	Good vtx	Bad vtx	Purity
No cut	2040	90	96%
K0 and pos cut	1960	61	97%

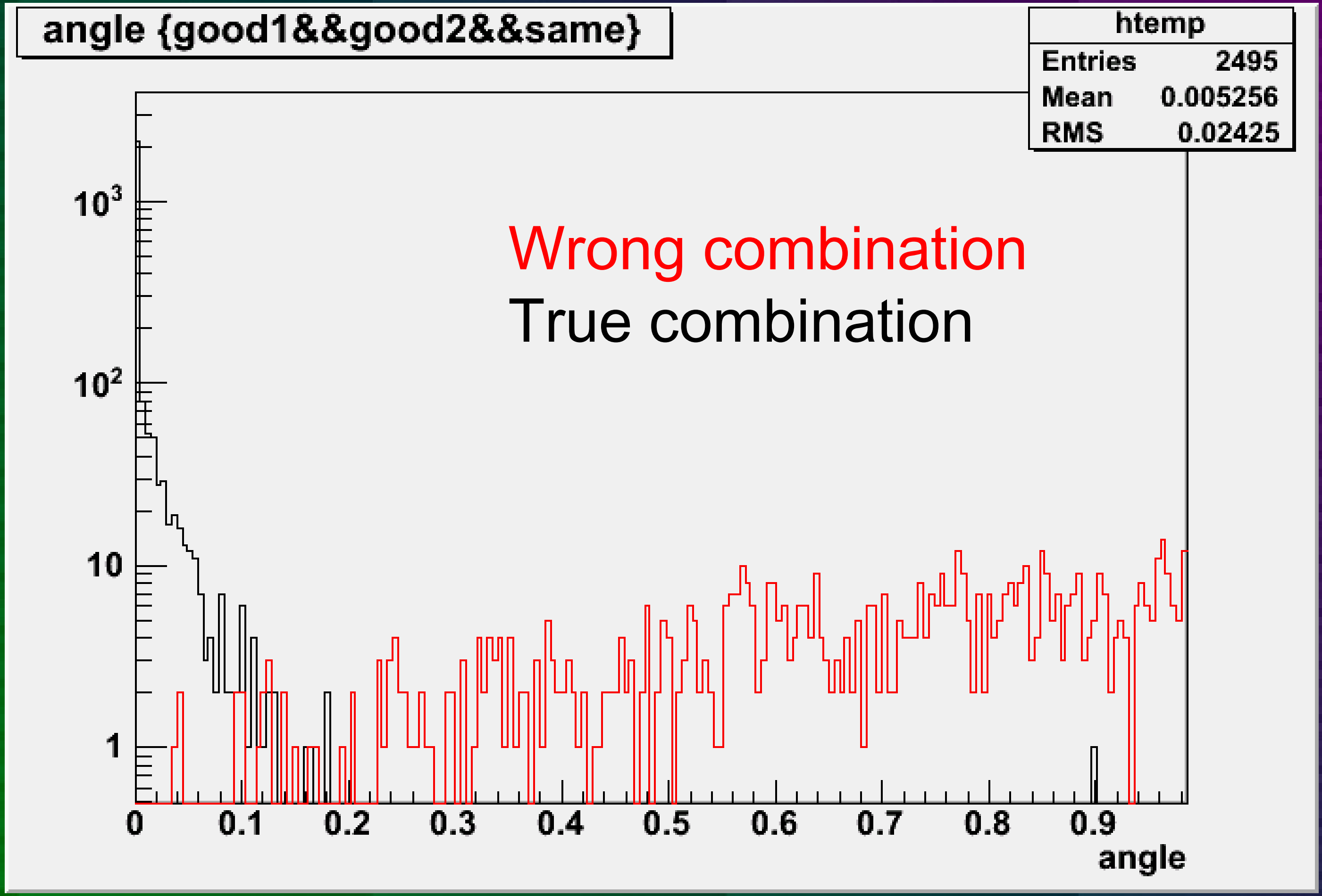
Optimized for efficiency (bad contains partially bad)

- Lepton selection (347 bbbbbb events)

	Secondary μ	Other μ	Others	Purity
No cut	430	585	23168	1.8%
Muon hit > 50 MeV	267	23	49	79%
All cuts	178	4	5	95%

Optimized for purity (not so good efficiency)

3. Vertex combination



Opening angle < 0.2 rad is treated to be the same vertices
(< 0.3 rad if either is muon)

4. Jet clustering

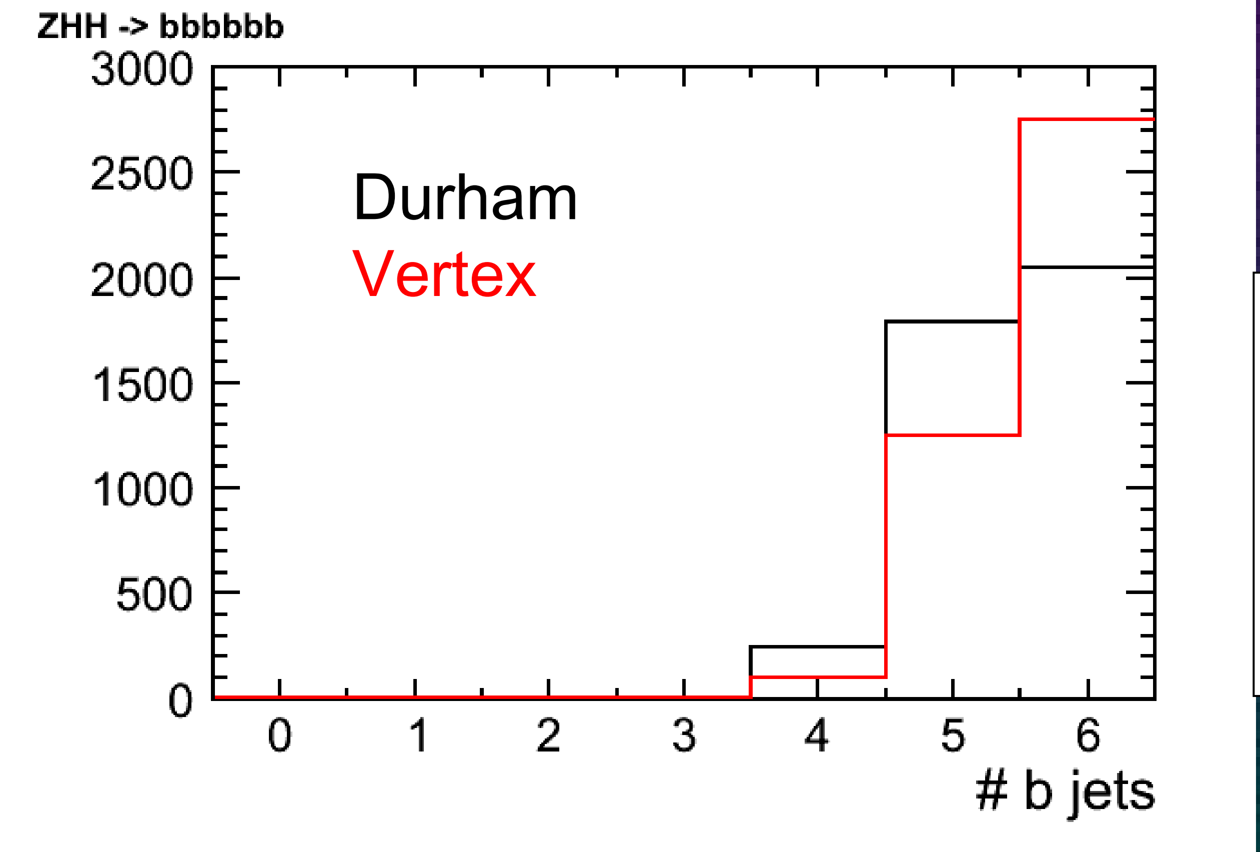
- Collecting PFOs around each vertex
  - PFOs with opening angle < 0.2 rad is associated to the vertex
- Ordinary Durham method for rest of PFOs
  - Combined from the least y-value

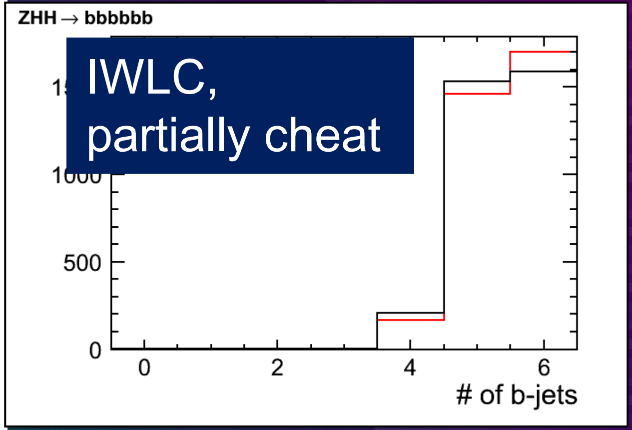
$$y = \frac{2\min(E_1, E_2)^2(1 - \cos\theta_{ij})}{Q^2}$$

Durham y-measure

  - Jet candidates with vertices are never combined
    - y-value with such combination is set to be infinite

Number of b jets in bbbbbb

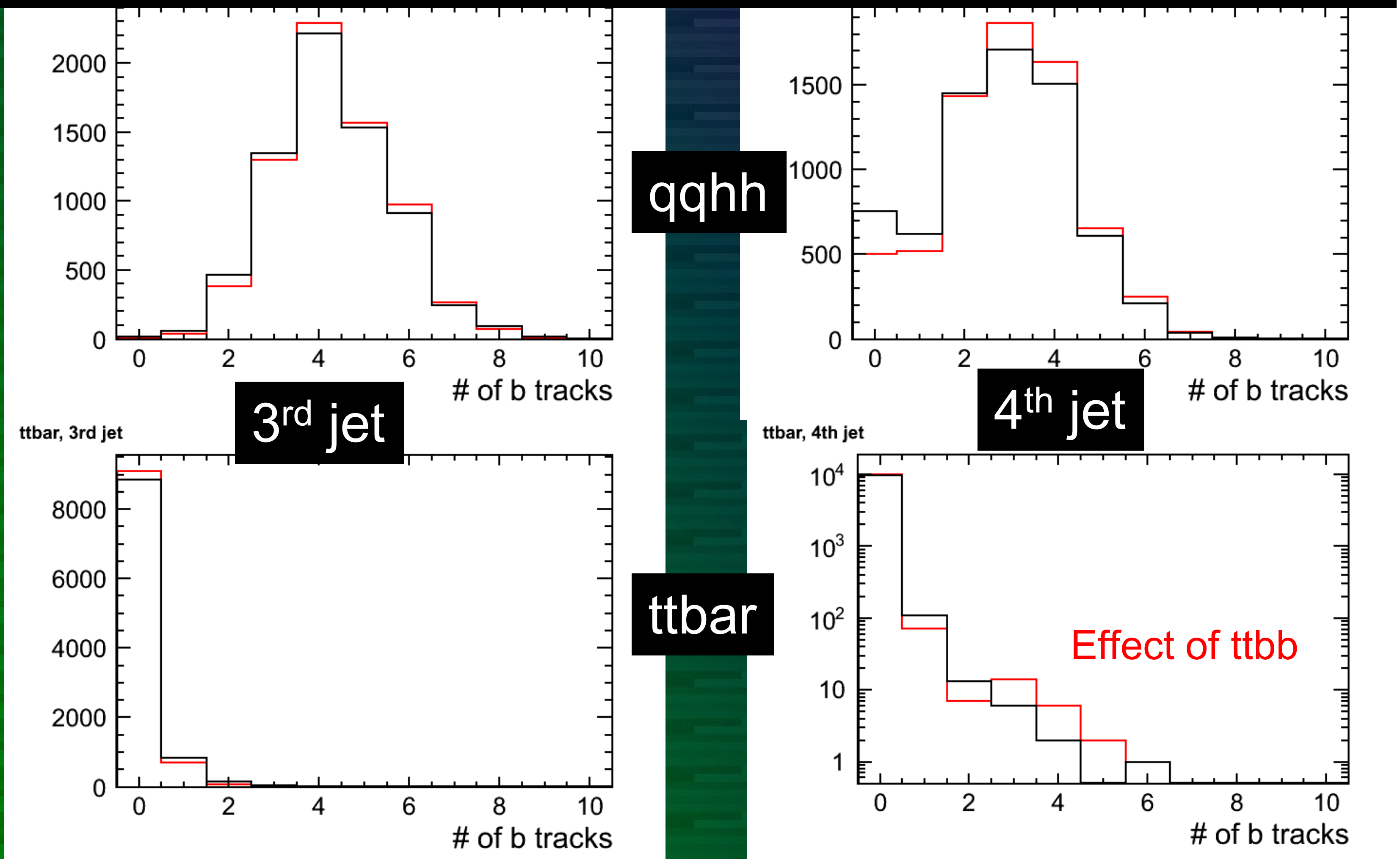




significant improvement seen!
All jets including b – 52% -> 66%

of b-hadron tracks

Jets are sorted by descending order of # of b-hadron tracks



ttbar rejection improved

qqhh acceptance improved

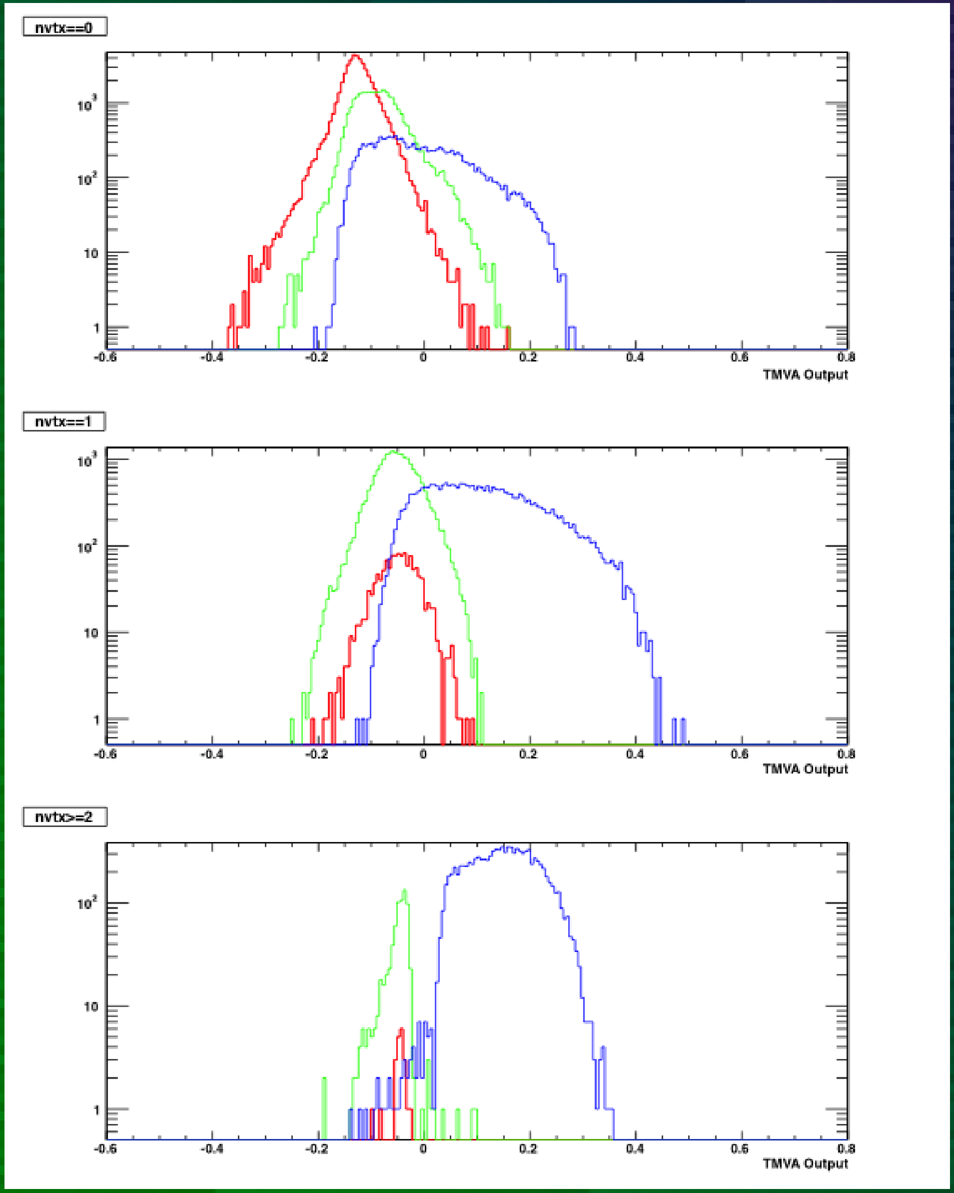
B-tagging performance

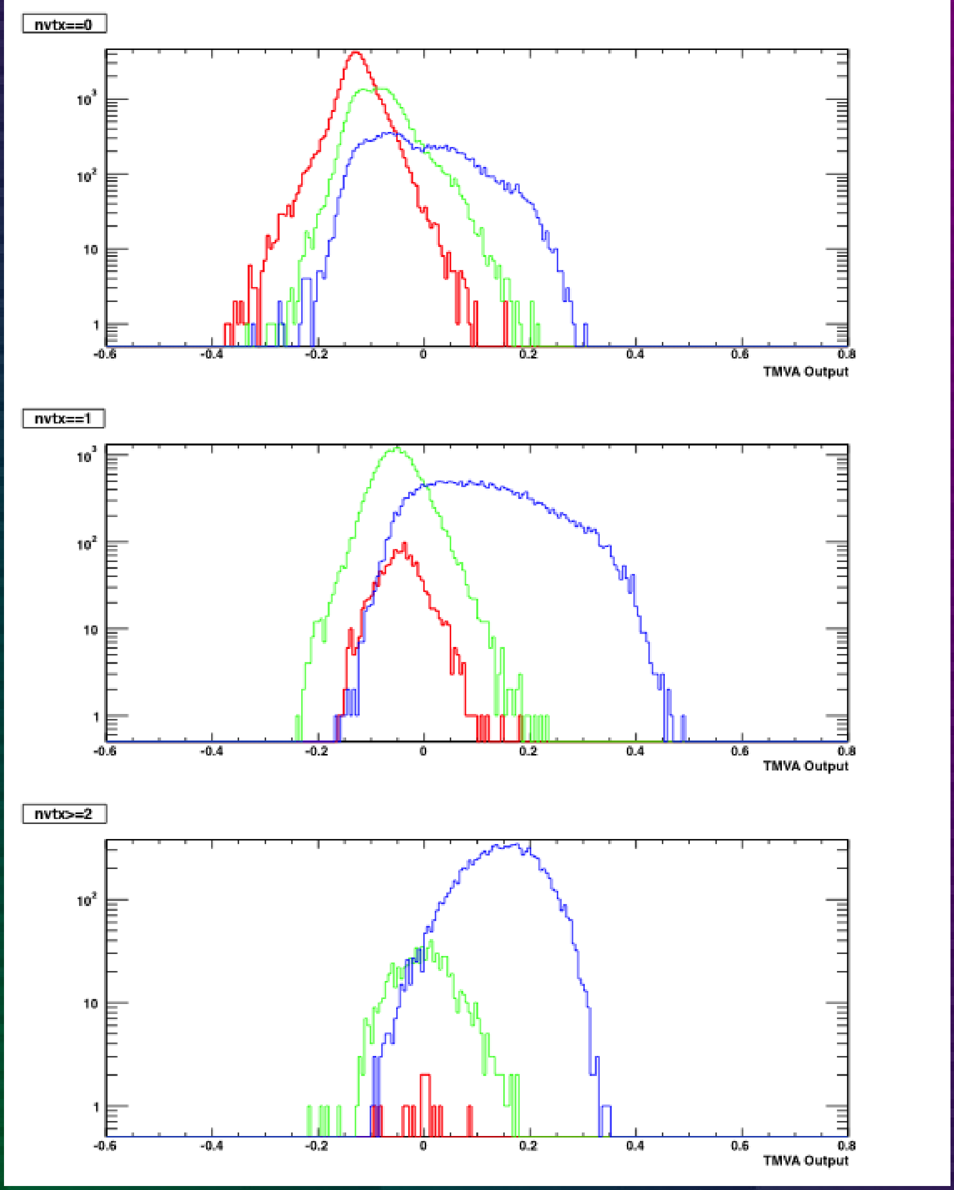
Data buried in KEK servers…

Summary for all

- Flavor tagging
  - New framework of BDT/ANN developed
  - New variables examined
  - Performance at Z-pole improved
  - c-tagging & CLIC study about to start
- Jet clustering
  - New jet clustering with vertex developed
  - Performance in ZHH/ttbar improved in MC-truth level
  - b-tagging performance will be checked

backup





BDT classifier response of training sample

BDT classifier response of testing sample