Higgs Self Coupling Analysis using the events containing H→WW* decay

Masakazu Kurata, Tomohiko Tanabe
The University of Tokyo
Junping Tian, Keisuke Fujii
KEK
Taikan Suehara
Kyushu University

LCWS13, 11/11/2013-11/15/2013

Introduction

- Measuring the Higgs self coupling is the key point to prove the electroweak symmetry breaking mechanism
  - Higgs potential in SM:

$$V = \lambda v^2 H^2 + \lambda v H^3 + \frac{1}{4}\lambda H^4$$

Mass term

Trilinear coupling

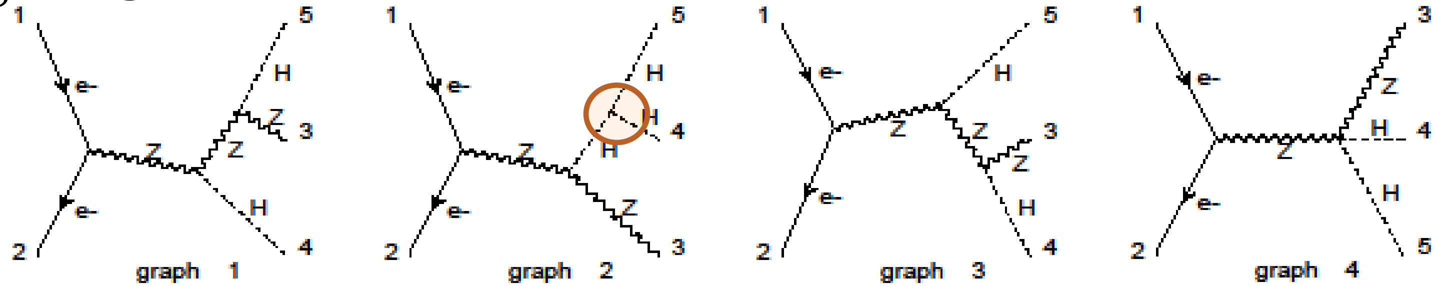
SM: $\lambda = \frac{m_H^2}{2v^2}$ $v \sim 246 GeV$

Quartic coupling
→difficult to measure

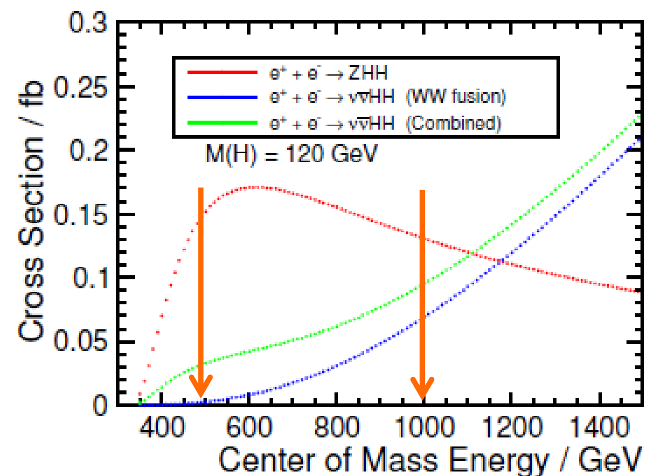
- Observing two Higgs bosons in the event is the only way to measure the self coupling
- Accurate test of the coupling may lead to the extended nature of Higgs sector → may go to new physics
- Our goal is to observe and measure the Higgs self coupling first

Signal Events

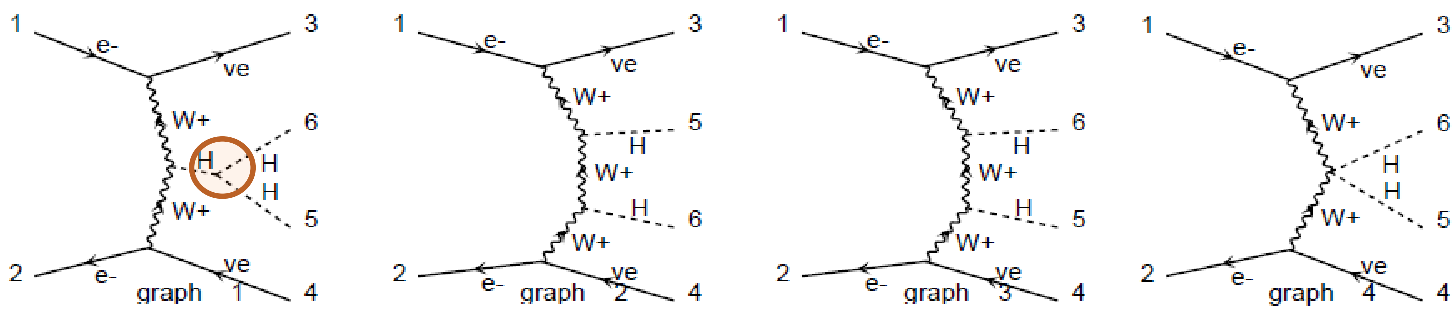
- Signal@500GeV - $e^+e^- \to Z^* \to ZH \to ZHH$ can be used



Signal: 2 Irreducible B.G.: 1, 3,4



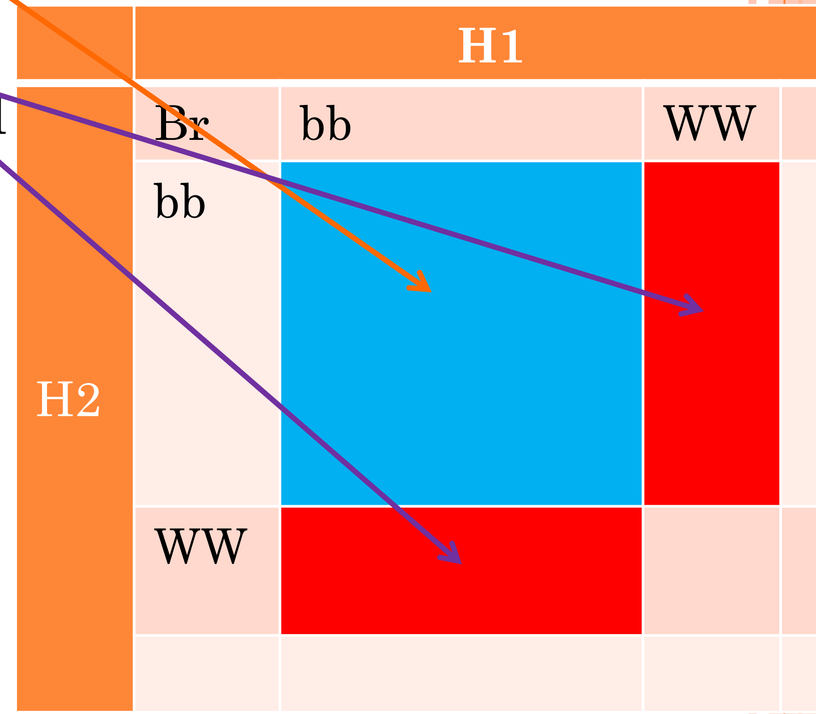
- Signal@1TeV – VBF $e^+e^- \to \nu\nu HH$ channel is opened
  - Increase the cross section of VBF



Signal:1 Irreducible B.G.: 2, 3,4

Status of this analysis

- Golden channel: Z(bb)(bb)
  - b-tagging can suppress the backgrounds
- Requirement of **H→WW*** decay
  - Same cross section as golden channel
  - Contribution of the total sensitivity
  - Need to reduce the backgrounds using kinematics of the events



But…

- Disadvantage of the analysis
  - Num. of b-tagging available is fewer than the golden channel
  - Background cross section is increased at higher energy
    - ttbar+X, which is difficult to reject from Higgs Strahlung process

Backgrounds and simulation

- Backgrounds considered:

B.G. Process	Feature	Basic idea for rejection
ttbar **ZWW**	Huge number of events	Flavor tagging Kinematics topology Difference of the final states
ttbar + X	b-jet rich in the final states Similar final states	Kinematics topology
Triple boson • ZZ + H • ZZZ	Small cross section b-jet rich in the final states	Kinematics topology Difference of the final states

S/B ~ 1/3000@500GeV, 1/1000@1TeV

- Simulation

	500GeV	1TeV
Polarization (e,p)	(-0.8,+0.3)	(-0.8,+0.2)
m_H(GeV/c^2)	125	125
simulator	Full with DBD	Fast

process	σ(fb)	σ(fb)
Signal(inclusive)	0.2	0.3
ttbar & ZWW	581.8	264.9
ttbar + QQ	0.83	5.74
ttbar + Z	0.98	7.81
ttbar + H	0.14	3.22
ZZ + H	0.77	0.54
ZZZ	1.83	1.62

- This talk, only 500GeV is shown

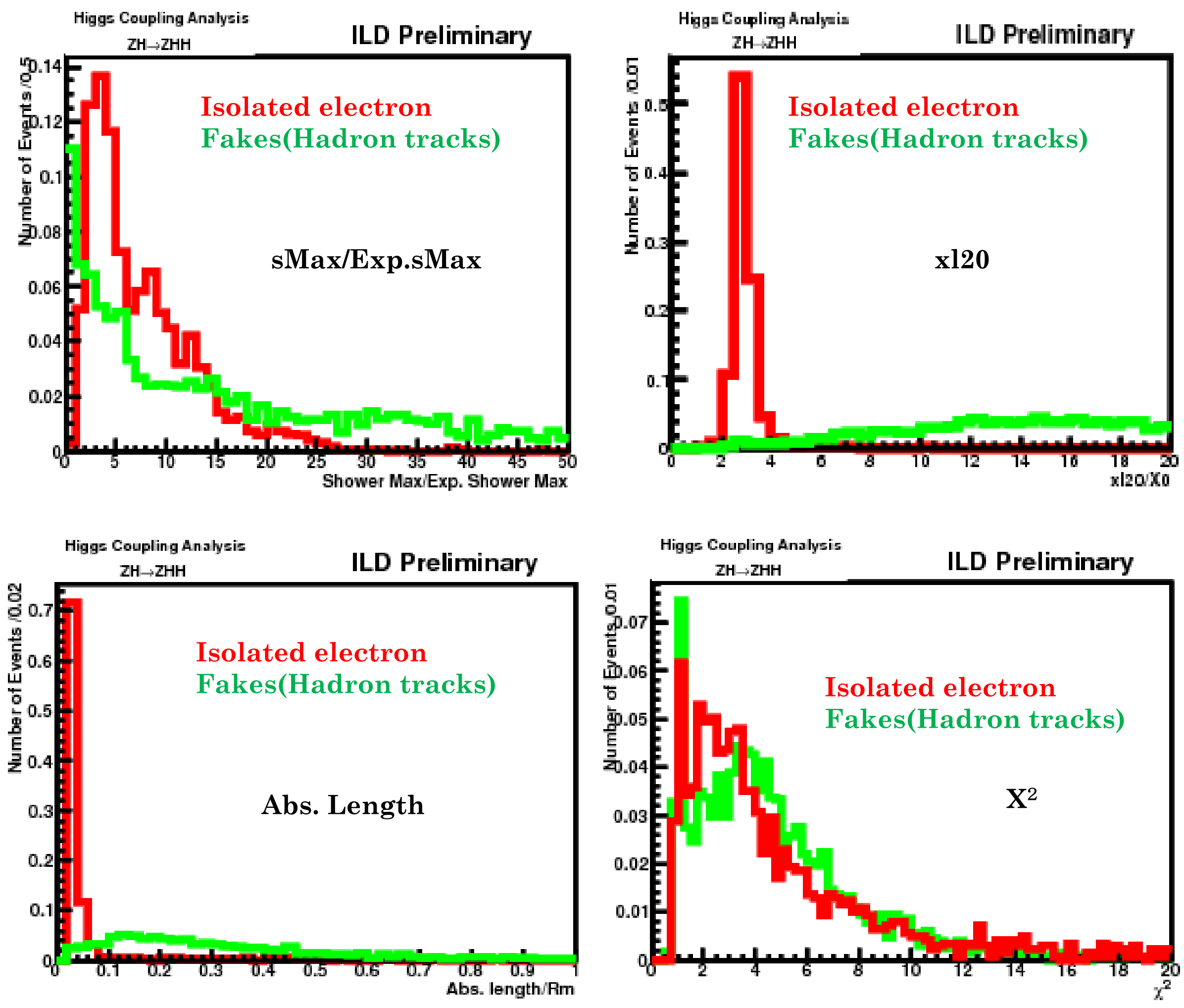
Components for better results

- Basic components for better sensitivity
  - Lepton ID: Isolated leptons can be identified well, and very good fake suppression
  - B-tagging: better b-tagging algorithm provides better background suppression
  - Jet pairing: good jet pairing can obtain good kinematic variables, which leads to good background suppression
  - Good energy & momentum resolution: of course, but limited by the detector performance

    →particle ID will be the key to energy correction
  - Jet clustering: jet reconstruction is the key to the analysis, but it is difficult
- All the components are related each other

Lepton ID improvement

- Shower profile
  - Shower development is different between EM/muon/Hadron tracks
  - Extract the shower profile information of longitudinal/ transverse components
  - Longitudinal



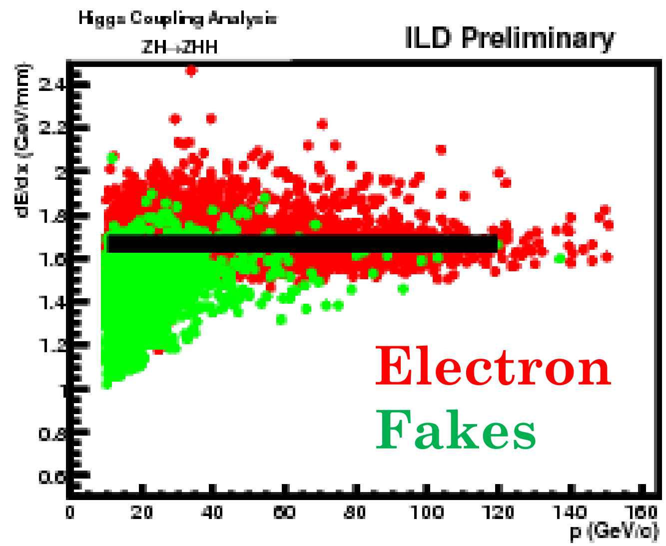
  - Transverse &χ^2

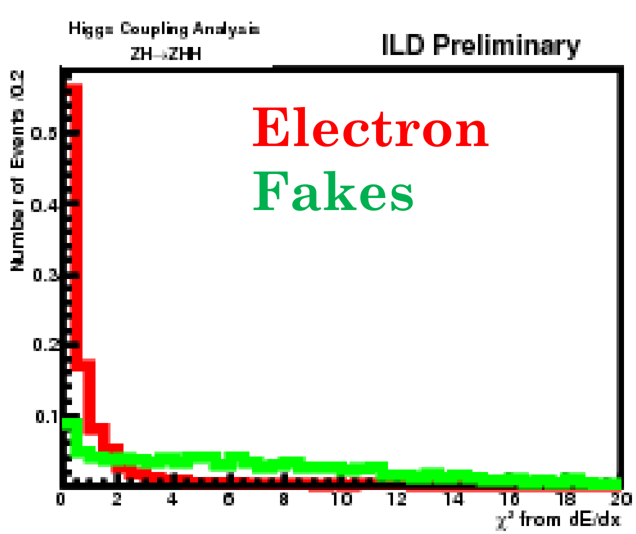
dE/dx

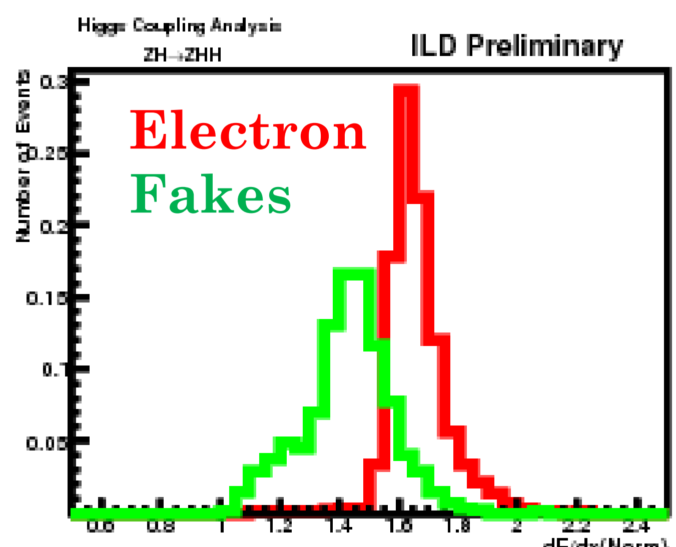
- Normalized dE/dx
  - Hadron tracks have low dE/dx value
  - Exp. mean with electron hypothesis is almost constant
- dE/dx distribution(1D)
  - Looks some difference →good for leptonID?
- X2 distribution with electron hypothesis:

$$\chi^2 = \left(\frac{\frac{dE}{dx} - \frac{dE}{dx}_{exp}}{\sigma}\right)^2$$

5% error imposed







Lepton ID

- Lepton ID is based on likelihood
  - Lepton selection imposing just one cut of likelihood output
- Target is to find the leptons from W boson as Higgs daughter
  - In some case, lepton energy is so small
  - Form general lepton ID to make the analysis easier
    - Want to apply it to Z lepton finding too

- Likelihood definition:
  - Isolated lepton likeliness

    $L = \frac{s}{s+b} \quad s = \prod s_i \quad b = \prod b_i,$

    s_i:pdfs of signal variables b_i:pdfs of background variables
  - Weight factor is introduced to make the efficiency improved

    taking log: $\log s = \sum w_i \log s_i$

Lepton ID for single lepton

- Signal detection efficiency – set almost same sensitivity
  - Signal is HH→(bb)(WW*)→(bb)(lvjj)
  - Old means without shower profile & dE/dx

method	Cut based	Likelihood_old	Likelihood_new
Signal(%)	98.1	98.1	97.8

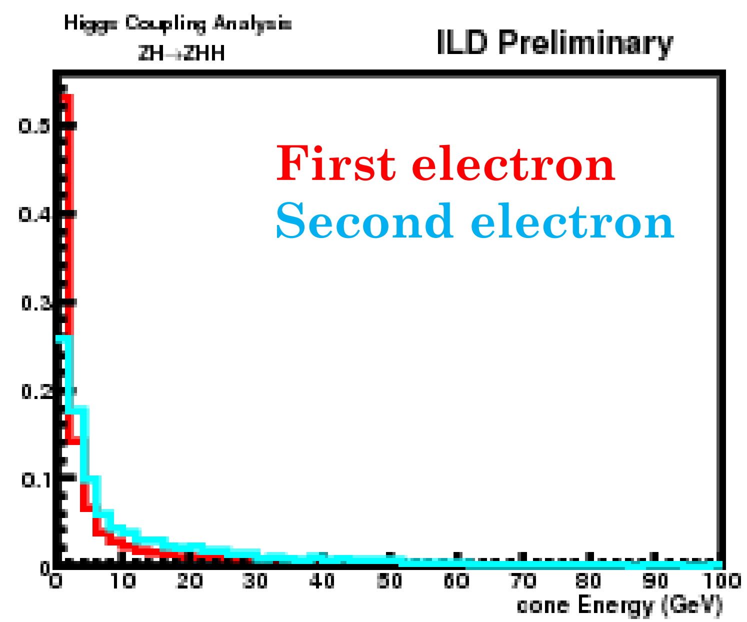
- Background rejection efficiency

method	Cut based	Likelihood_old	Likelihood_new
ttbar – lep+jets(%)	62.2	-	**62.4**
ttbar – allhad(%)	7.9	3.1	**2.3**
ttbar – dilepton(%)	47.2	-	**17.9**
HH→(bb)(bb) (%)	-	2.3	**1.0**

  - Note: lepton energy threshold is loosened on likelihood_new
    - From E(lep)>15GeV → E(lep)>10GeV

Lepton ID for Z→ll decay

- Second lepton(esp. electron) finding is difficult
  - As num. of tracks is large in the event, lepton might be buried in the tracks…
  - Change cone energy p.d.f. for second lepton
  - Loosen the operation point for second lepton



- Z→ll finding – preliminary result
  - Signal efficiency

signal	Z→ee, μμ
HH→(bb)(WW*)(%)	90.1
HH→(bb)(bb)(%)	89.6

  - Background rejection efficiency –use ttbar samples

method	Lep+jets	allhad	dilepton
Cut base(%)	0.79	0.071	17.3
Likelihood_new(%)	0.60	0.050	20.5

Jet Pairing using Bayesian approach

- Bayesian probability – posterior probability when x is given

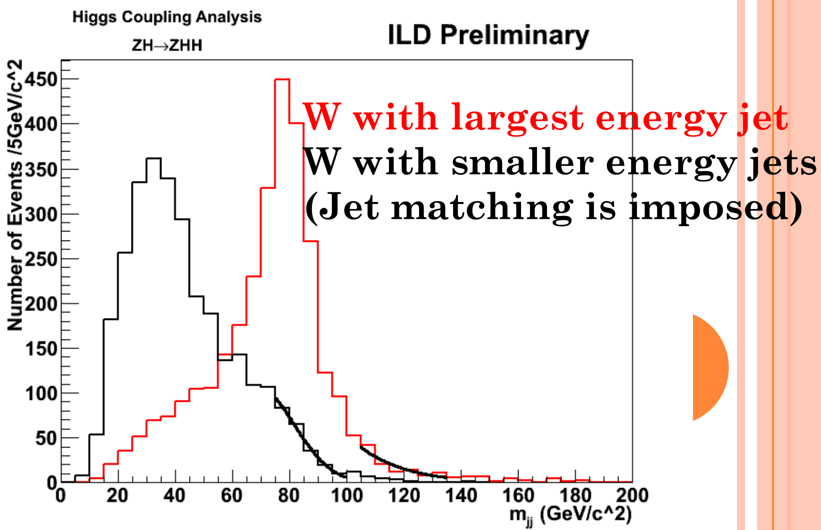
$$P(A|x) = \frac{P(x|A) \cdot P(A)}{P(x)}$$

P(x | A): likelihood(probability when x is given from class A)

P(A): prior probability of class A

P(x): probability of x (sum of all the classes' p.d.f.)

- Bayesian classifier – regard x as the element of class A,
  - When P(A | x) is largest of all the classes
    - e.g. x belongs to A when P(A | x)>P(B | x), P(A | x)>P(C | x), etc.
- Likelihood – introduce angle information
- In WW*→jjjj case, combination is 3
  - Jet with large energy tends to come from on-shell W



Preliminary results

- WW→jjjj pairing case
  - Also check maximum likelihood using LDA
  - $\chi^2 = -2\log BW(m(j1j2)|m_W, \Gamma_W)$

Pairing type	χ^2	Just likelihood	Naïve Bayes
True positive(%)	60.2	70.1	**74.7**

  - Good improvement can be obtained!

- ZH→(bb)(bb) case
  - $\chi^2 = \frac{(m_1-m_Z)^2}{\sigma_Z^2} + \frac{(m_2-m_H)^2}{\sigma_H^2}$

Pairing type	χ^2	Just likelihood	Naïve Bayes
True positive(%)	56.6	59.8	**59.8**

  - Improve slightly thanks to the angle information
  - But, need more improvement…
    - No improvement even if using naïve Bayes…

RESULTS

- Background rejection – Multivariate analysis
  - Re-train the classifier
  - Form some classifiers with different kinematic topologies
- Significance estimate:
  - Calculate $\frac{S}{\sqrt{S+B}}$
  - $L = 2ab^{-1}$
  - Fully used b-tagging and tightest b-tag condition

Category	signal	background	significance
Lepton + 6jets	2.07	20.65	**0.41 σ**
Dilepton + 6jets	2.94	9.73	**0.83σ**

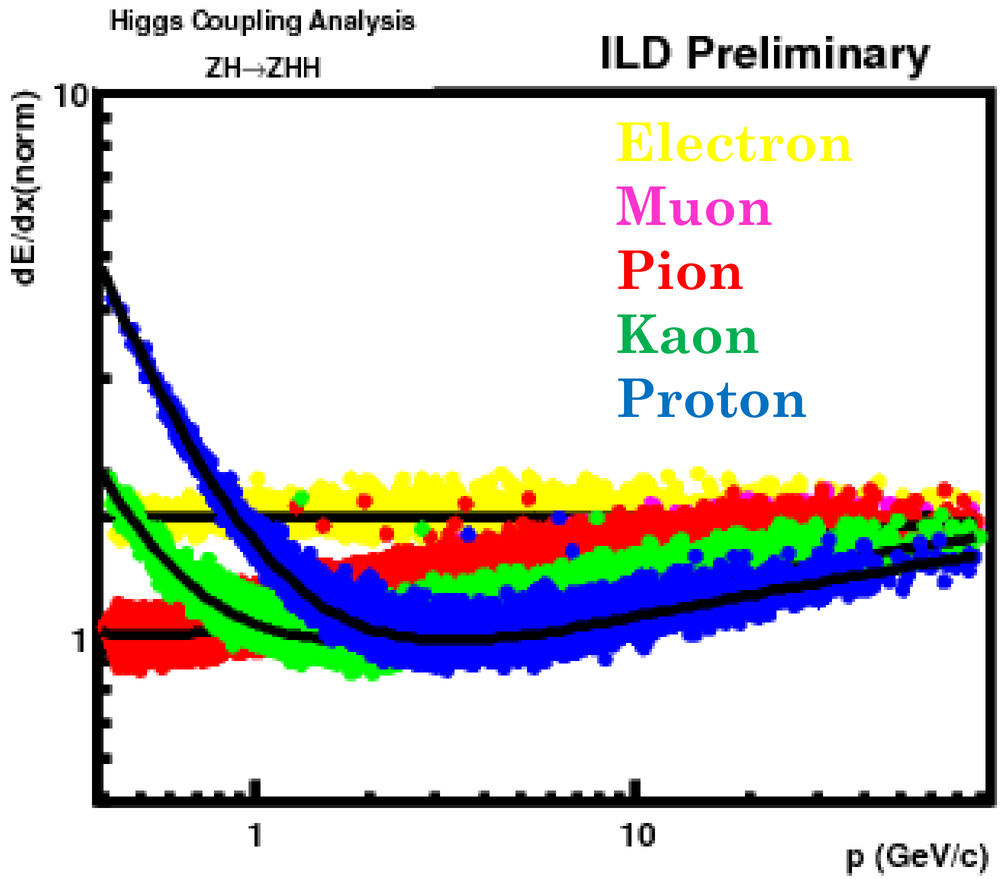
- Comparison with the old results

Category	LCWS13	AWLC14
Lepton + 6jets	0.38σ	**0.41σ**
Dilepton + 6jets	0.68σ	**0.83σ**

  - Updated analysis provides 8&19% improvement!

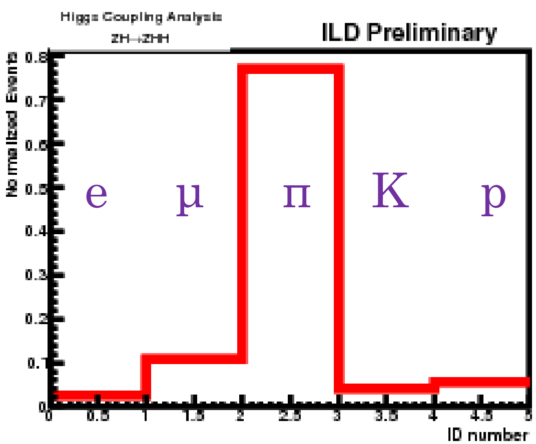
PROSPECTS FOR FURTHER IMPROVEMENT

- dE/dx
  - Checking from qqHH sample –tracks in the jet
  - There are many tracks located in the range where dE/dx can work!

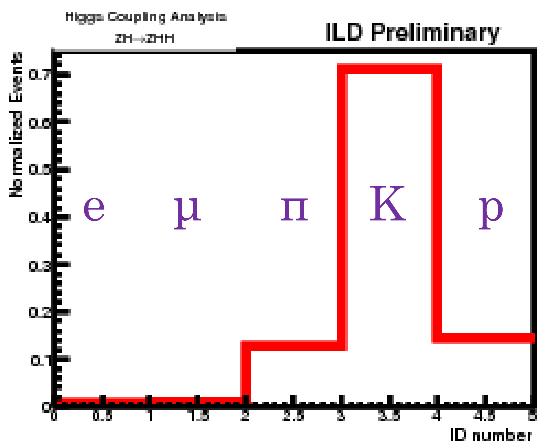


PROSPECTS FOR FURTHER IMPROVEMENT

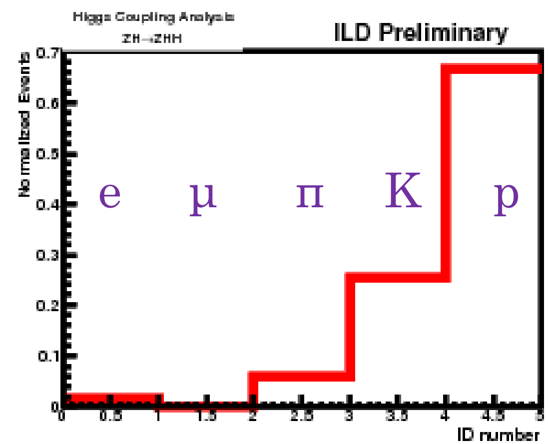
- Particle ID based analysis
  - Using dE/dx and shower profile, particle ID can be performed
    - Hadron ID efficiency



π



K



p

- Can apply to many part of the analysis components
  - Track energy correction?
  - B-tagging?
  - Hope for jet clustering improvement?

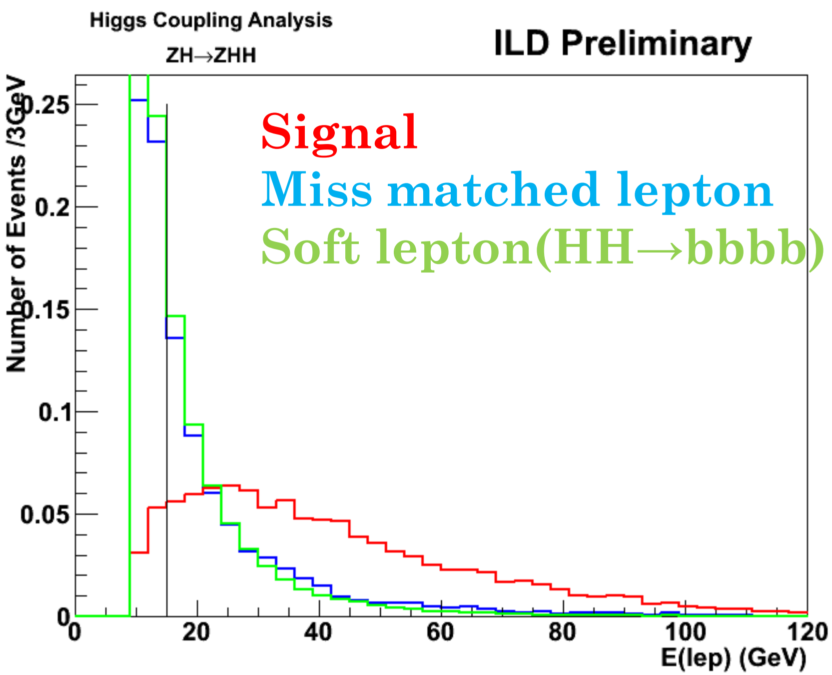
Summary and plan

- Higgs self coupling analysis using the events with H→WW* is ongoing.
  - Shower profile & dE/dx information are included for lepton ID
  - Jet pairing based on the Bayesian approach
  - Lepton related analysis obtained 8-19% improvement from old analysis
- **Plan:**
  - Start to combine with golden channel and estimate the Higgs self coupling
  - Full simulation @1TeV
  - Optimize b-tagging strategy
    - Forming looser b-tag category
  - Construct the Particle ID strategy and optimize it
    - Try to correct track energy
    - B-tagging improvement
    - Jet clustering?

BACKUPS

EVENT SELECTION

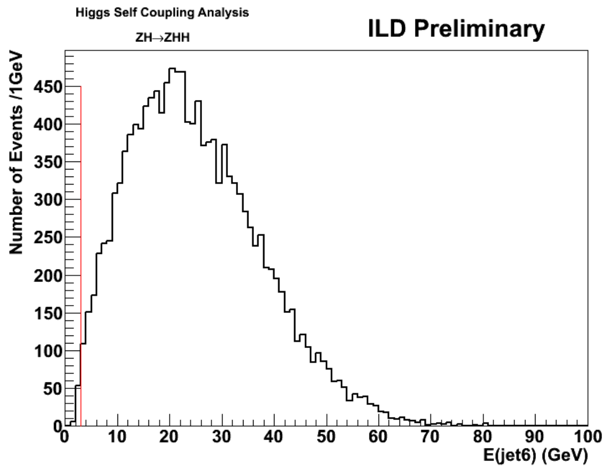
- Lepton selection
  - E(lep)>15GeV is required to suppress soft leptons
  - Dividing into 3 samples:
    - one lepton for lepton + jets
    - two leptons for dilepton → opposite charge & same flavor
    - three leptons for trilepton → looking for the lepton pair from Z
    - tight lepton veto for all hadronic

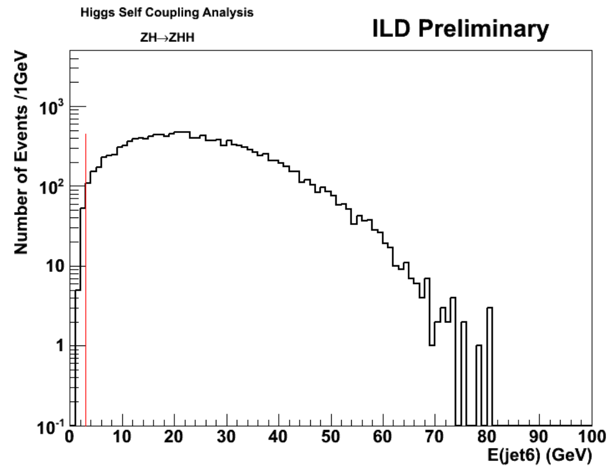


Lepton matching is required:
cosθ>0.9

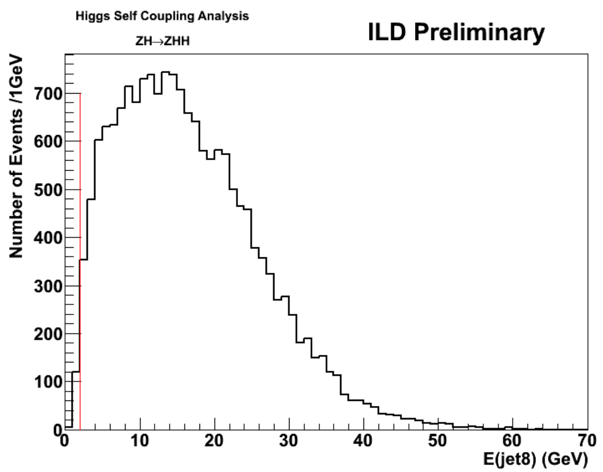
Event selection

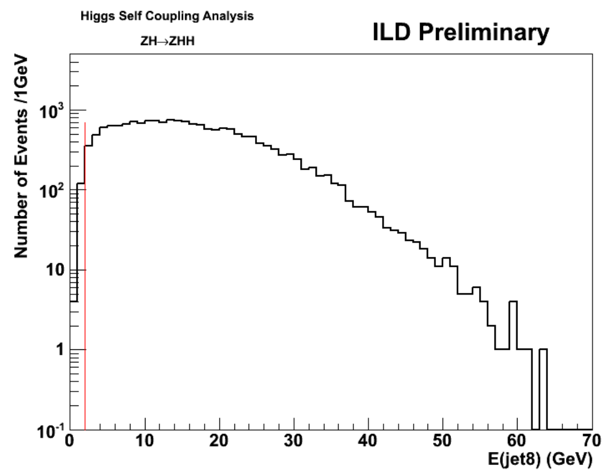
- jet selection
  - requiring 4jets, 6 jets or 8 jets using LCFIPlus & Durham
  - E(jet)>3GeV is required to care the lowest energy jet for **6 jets** case
  -





  - E(jet)>2GeV is required to care the lowest energy jet for **8 jets** case



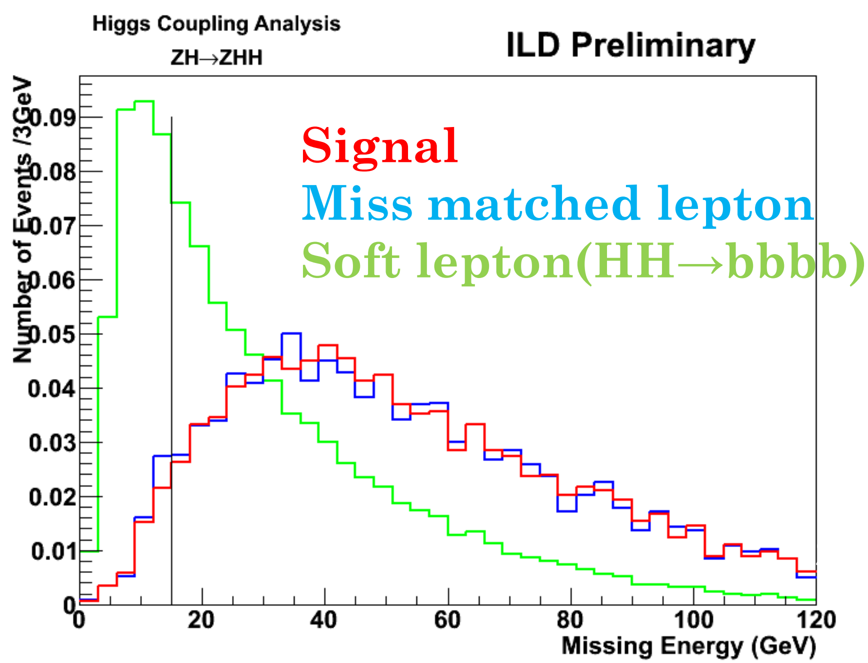


Event Selection

- Detection efficiency after the jet energy cut:

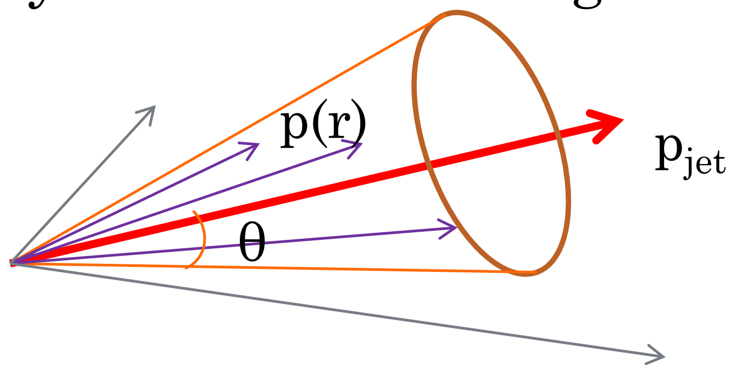
signal	4jets requirement	6jets requirement	8jets requirement
Efficiency(%)	99.0	99.4	99.6

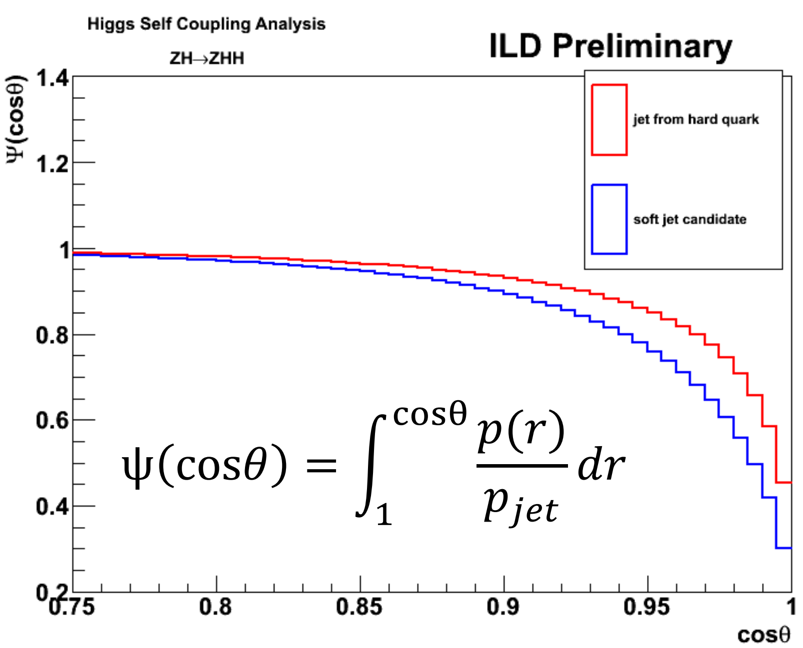
- Missing momentum
  - lepton + jets: P(Miss)>15GeV/c to suppress HH→bbbb events
  - All hadronic & dilepton: P(Miss)<80 GeV/c required
  - Trilepton: P(Miss)<150GeV/c to gain the acceptance



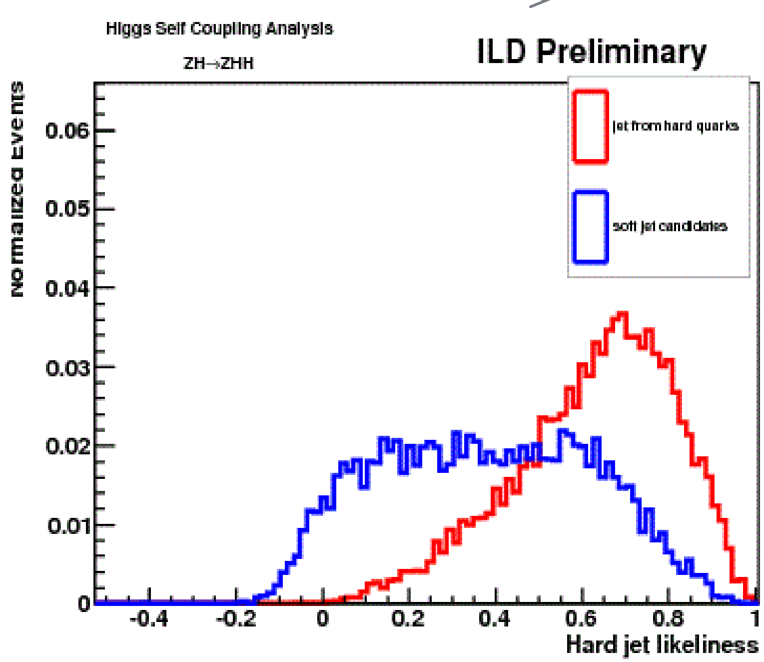
Soft jet finding

- Tracks in the gluon jets spread wider than those in quark jets(e.g. analyses on hadron collider)
  - Traditional jet shape can be a good estimator
- Using Multivariate Analysis and estimating the hard jet likeliness for each jet



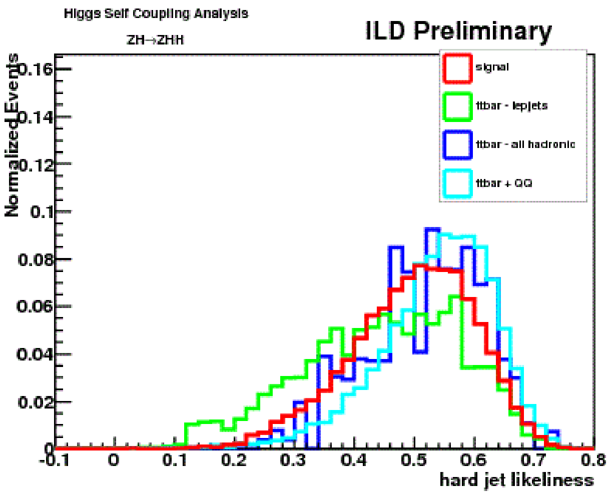


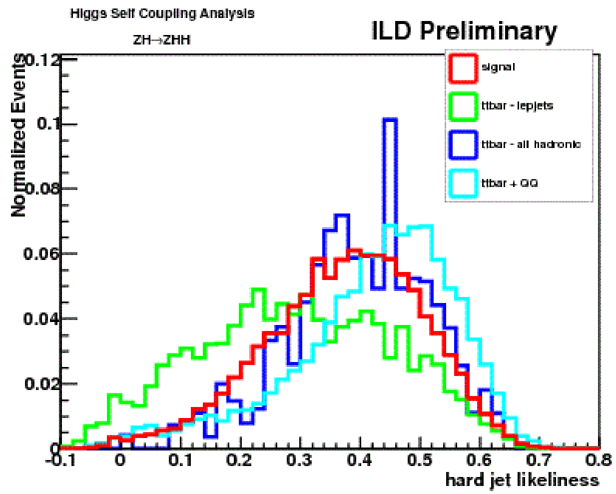
$$\psi(\cos\theta) = \int_{1}^{\cos\theta} \frac{p(r)}{p_{jet}} dr$$



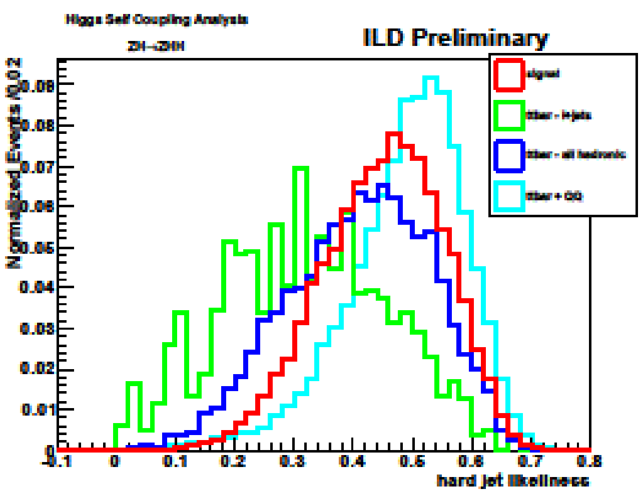
CHECK THE PERFORMANCE

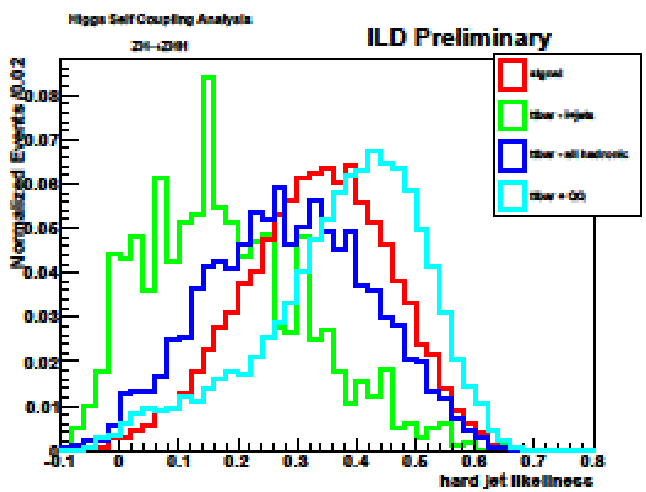
- Check the jets with small hard jet likeliness – signal vs. ttbar
- For 6jets





- For 8jets

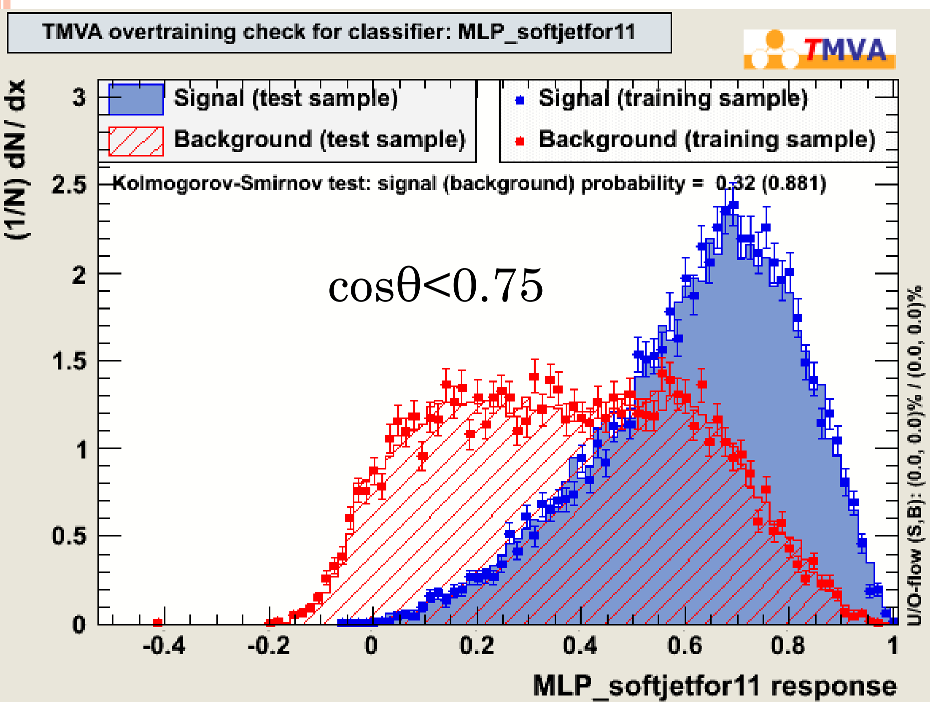


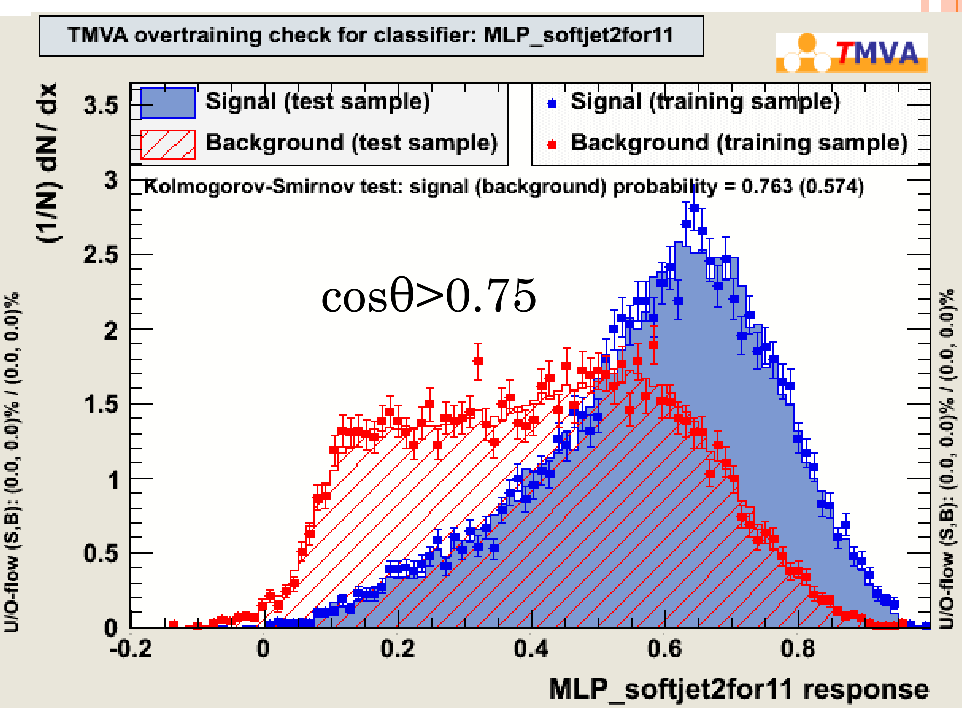


Soft jet finding

- Hard jet likeliness is introduced
  - Using MVA to form it
  - Analysis samples are divided into 2 based on the angle with the nearest jet

  →large shared area for both jets deteriorate the traditional jet shape
- Use the likeliness for the input of background rejection MVA or simple cut of backgrounds





Reduction table

All hadronic

- Final b-tagging: btag(3)>0.92 && btag(4)>0.44
- HH→bbbb contamination is 5.41 events

process	signal	ttbar	tt + QQ	tt+Z	tt + H	ZZ + H	ZZZ
expected	354.00	1.16×10^6	1660.00	3307.00	280.00	1540.00	3660.00
preselection	49.47	2462.09	79.11	76.25	38.32	87.22	70.72
Jet energy	47.92	1970.58	77.62	74.98	37.96	72.88	57.28
$\chi2$	44.32	1353.38	64.57	62.41	34.02	61.60	48.16
Visible energy	44.23	1326.19	64.31	62.00	33.92	61.18	47.90
NN for 8 jets	36.51	1011.92	36.37	34.37	16.38	51.59	47.90
NN for ttbar	20.53 (9.85)	302.59	26.44	25.17	13.07	21.71	9.00
b-tagging	14.92 (5.41)	87.54	17.54	16.42	9.13	16.10	6.03