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

LCWS14, 10/06/2014-10/10/2014

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

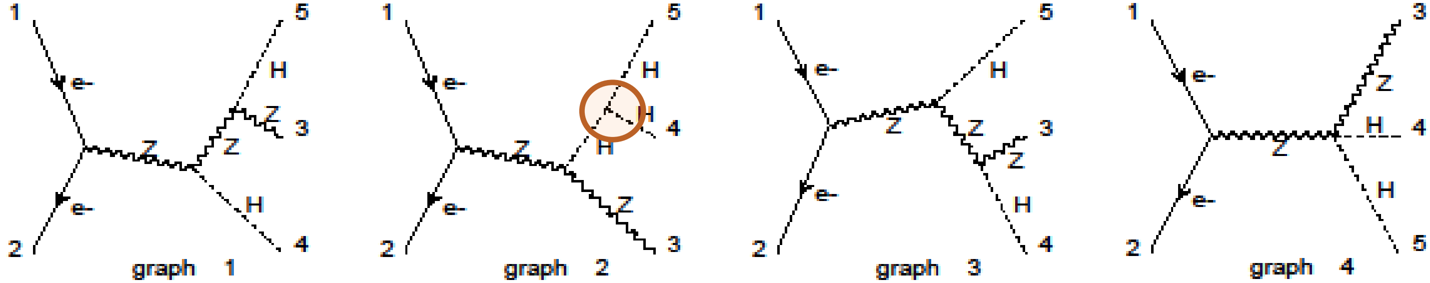
$SM:\quad \lambda = \frac{m_H^2}{2v^2} \quad v \sim 246 GeV$

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

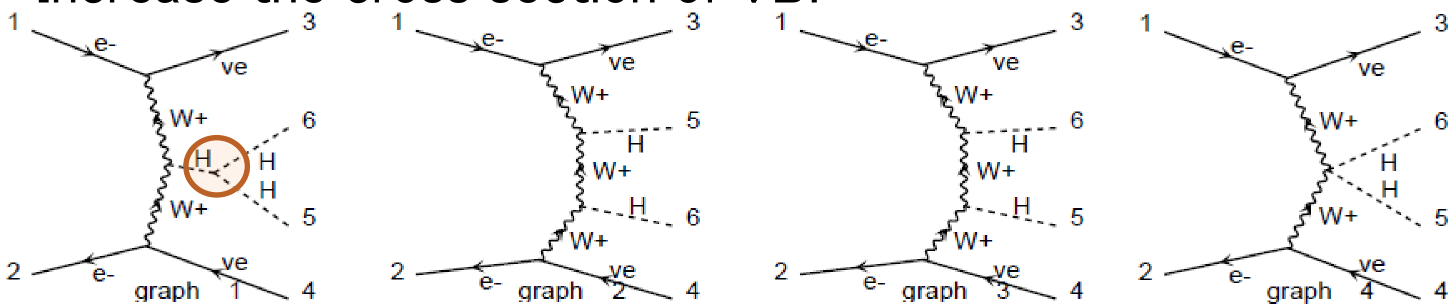
SIGNAL EVENTS

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



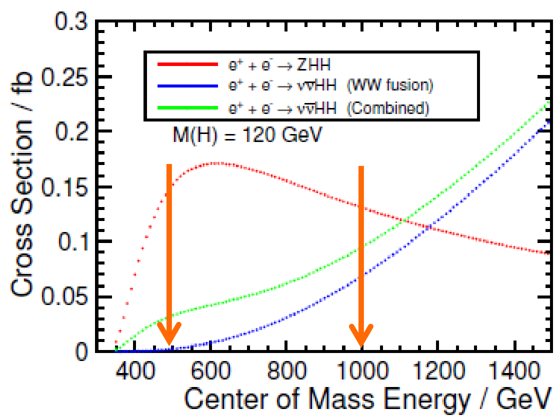
Signal: 2 Irreducible B.G.: 1, 3, 4

- 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

- Higgs decay modes:
  - HH→(bb)(bb): golden channel thanks to b-tag
  - HH→(bb)(WW): improve the final result



Components for better results

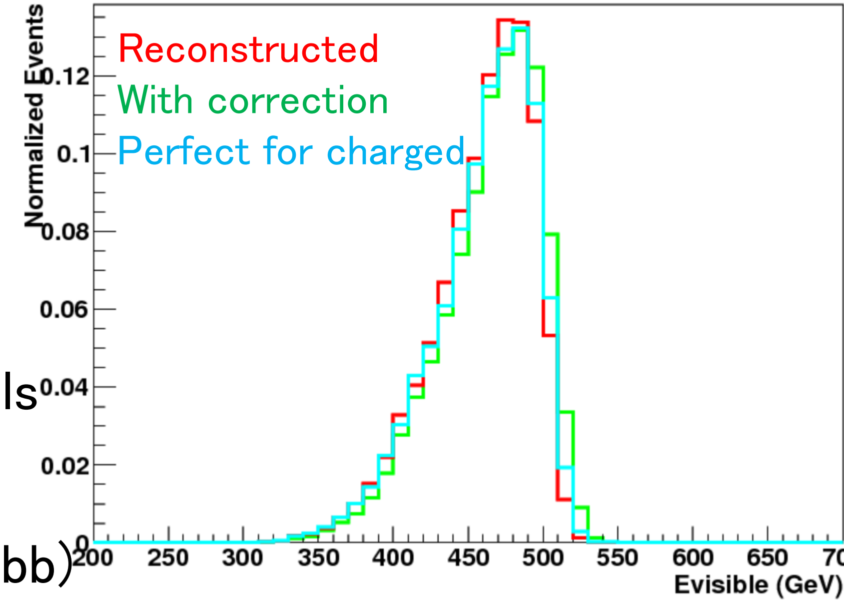
- Basic components for better sensitivity
  - Lepton ID: Isolated leptons can be identified well, and very good fake suppression

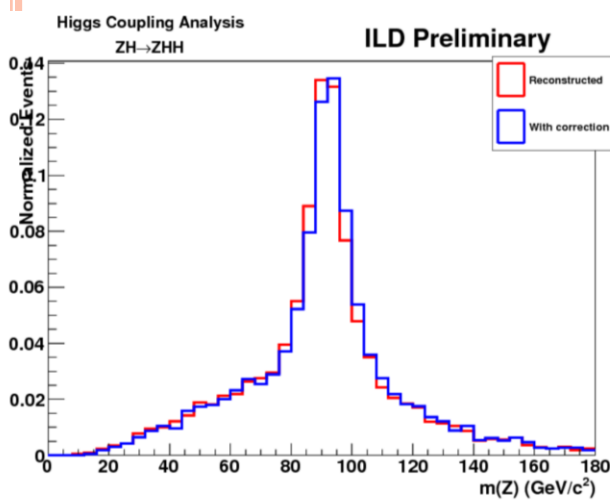
    →many idea have been introduced
  - B-tagging: better b-tagging algorithm provides better background suppression
  - Jet pairing: good jet pairing can obtain good kinematic variables, which leads to good background suppression
  - Good energy & momentum resolution: of course, but limited by the detector performance

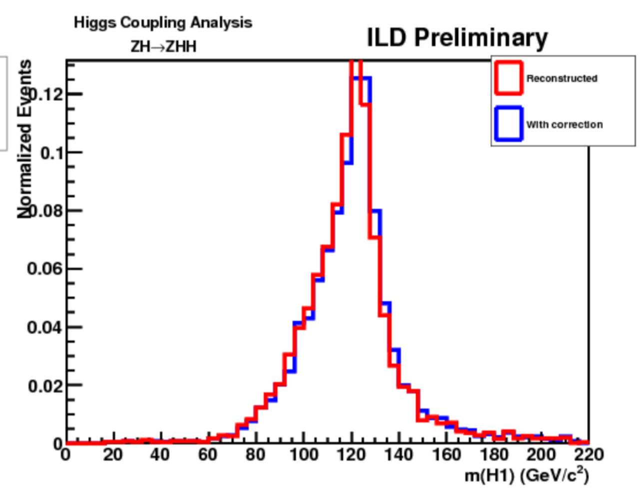
    →particle ID will be the key to energy correction
  - Jet clustering: jet reconstruction is the key to the analysis, but it is difficult
- All the components are related each other

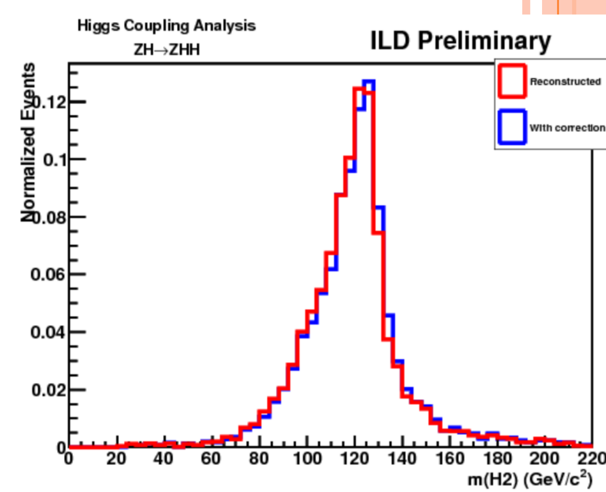
Track energy correction

- Track energies are corrected using momentum & mass
  - Using particle ID to identify tracks
- Visible energy
  - Using qqHH→qq(bb)(bb)
  - So far, overestimated due to misID
  - Correction effect is small due to neutrals
- Mass distribution
  - Checking Z(Z→qq, q is light) and H(H→bb)
  - Jet matching with MC truth is applied
  - Effect is small too due to neutrals









Jet Pairing using Bayesian approach

- Bayesian probability – posterior probability when x is given

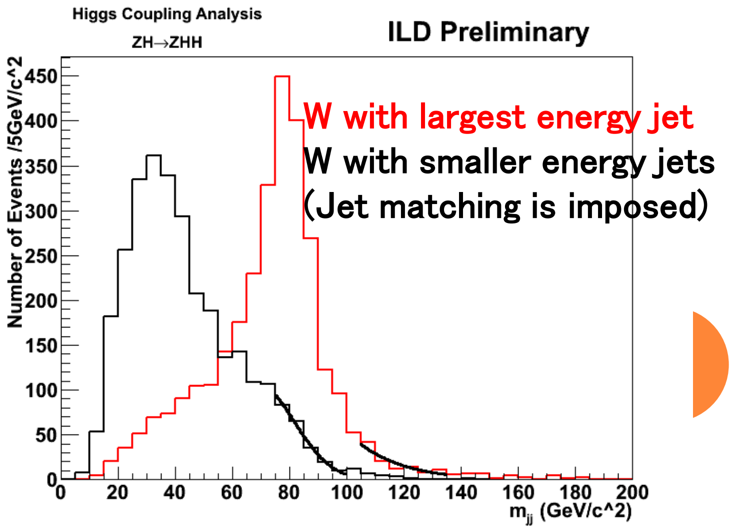
$$P(A|x) = \frac{P(x|A)\cdot P(A)}{P(x)}$$

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

P(A): prior probability of class A

P(x): probability of variable 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 & problem

- 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!

Looks hopeful, but･･･

- 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･･･

- Can we get better result?

Trying Kinematic Fitter

- Determining the kinematics globally in the events
  - Distort the event kinematics to meet the constraint in specific process
  - Estimate how much is a event likely to the specific process?
  - Mass resolution will be improved by using χ^2 minimization

- First trial: ZHH→(bb)(bb)(WW*)→(bb)(bb)(l ν jj) kinematic fitter
  - Constraints:

$$m(bb) = m_Z$$

$$Max\big(m(l\nu), m(jj)\big) = m_W$$

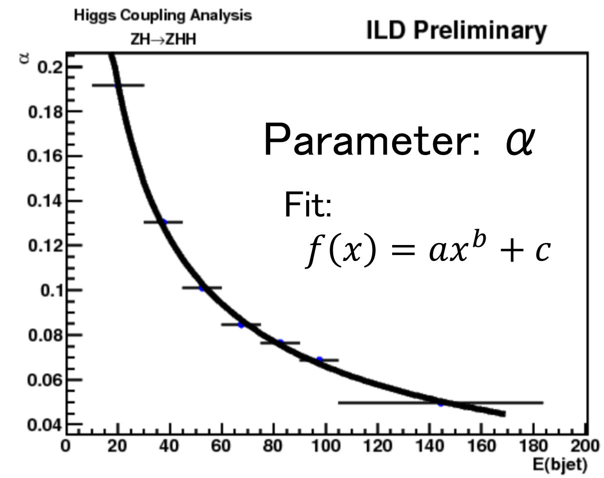
$$m(bb) = m(l\nu jj)$$

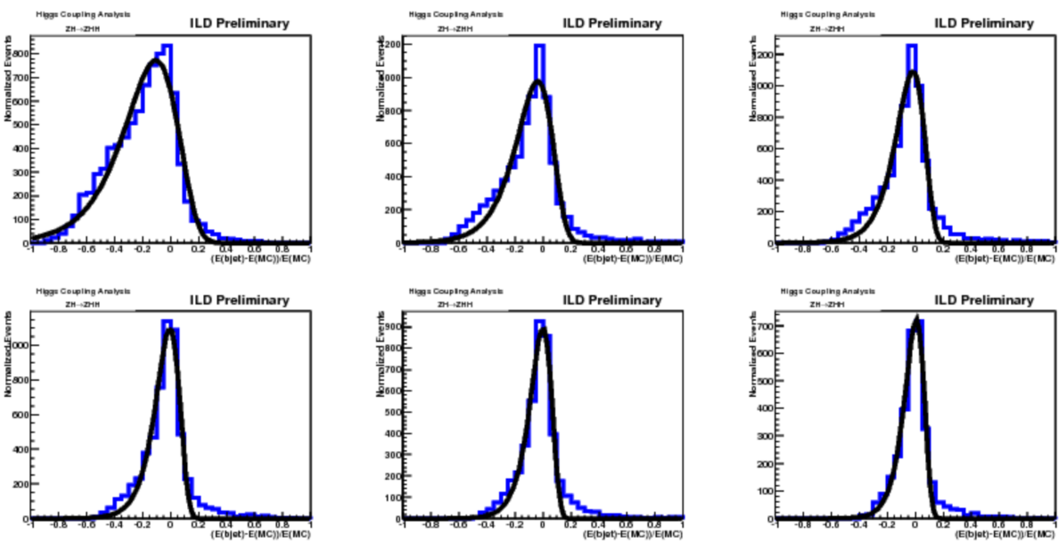
$$E(H) + E(Z) + E(jj) + E(l\nu) = \sqrt{s}$$

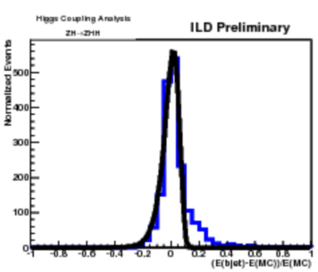
$$\overrightarrow{p_H} + \overrightarrow{p_Z} + \overrightarrow{p_{jj}} + \overrightarrow{p_{l\nu}} = \vec{0}$$

Jet energy resolution

- Most critical factor which degrades mass resolution is jet energy resolution
  - So it is necessary to include this effect into Kinematic fitter
  - Jet energy resolution has energy dependence of jets
    - Parameterize fit parameters with jet energy



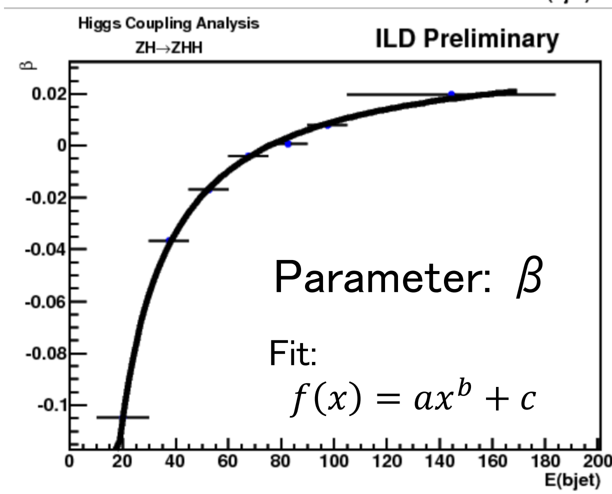




Fit: Gumbel dist.

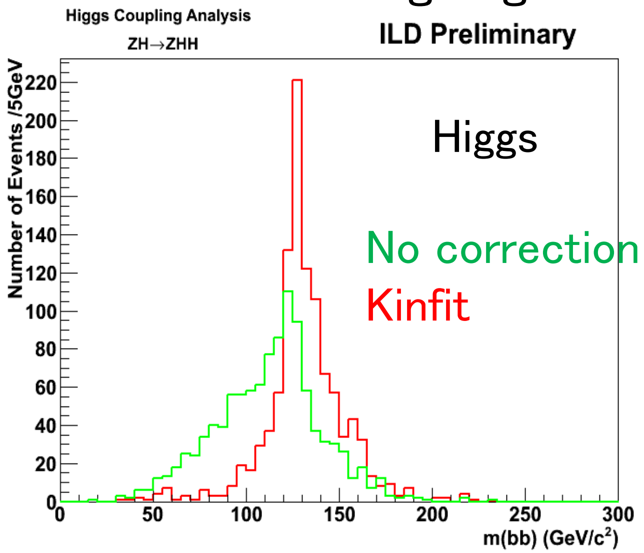
$$f(x) = \frac{1}{\beta}\exp\left(\frac{x-\alpha}{\beta}\right)\exp(-\exp(\frac{x-\alpha}{\beta}))$$

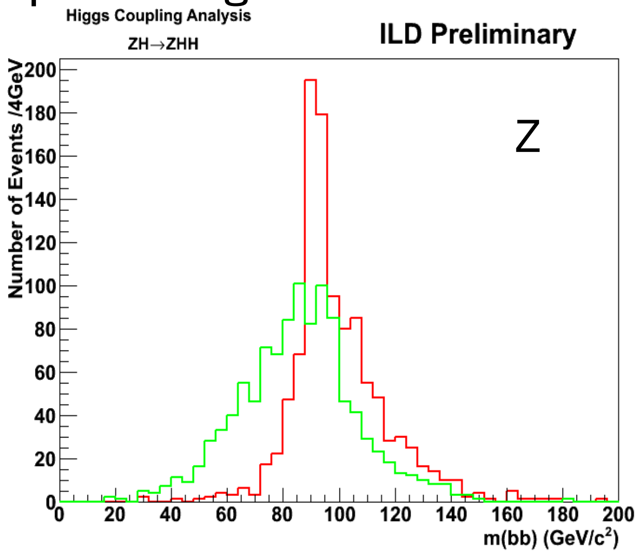
To include asymmetric energy resolution



Performance check

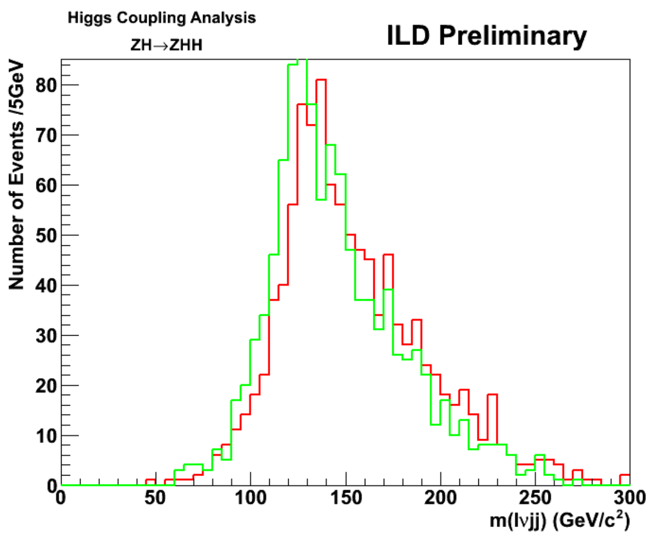
- Higgs mass(H→bb) & Z mass distribution
  - Mass resolution is going better! →promising





  - Higgs(H→WW*→l ν jj)
    - No resolution improvement···

    →under investigation



Possibility of Flavor tagging improvement

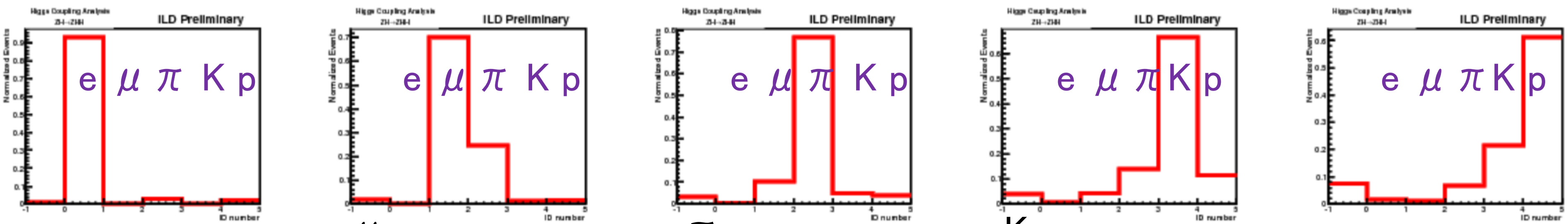
- For flavor tagging improvement
  - Vertex mass is the key to separate heavy/light flavor vertex
  - Many pi0s will escape from B/D vertex → checked that using MC truth
  - Mass resolution will be degrade due to escaping neutrals
  - Is there possibility to recover pi0s which escape from vertices?

- Building π^0 finder – many components are necessary
  - Gamma finder – using shower profile in calorimeters
  - π^0 finder – solving gamma pairing
  - Vertex finder – which vertex is the π^0 coming?

- Find vertex of pi0s:
  - Very difficult to identify vertex – depends on detector configuration
  - Making the best of decay kinematics
  - Using TMVA to find pi0 candidates from the vertex
  - Details: my talks@software/simulation session(Thursday)

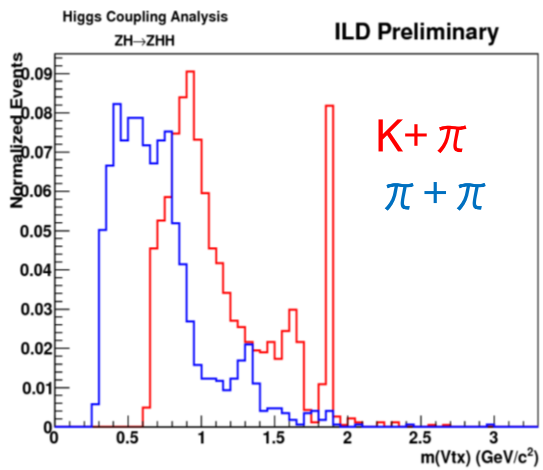
A CLUE FOR FLAVOR TAGGING IMPROVEMENT

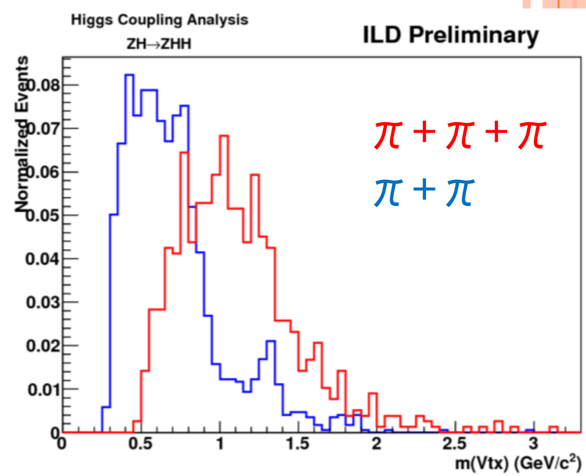
- We have Particle ID! – it is the key
  - Using TPC dE/dx and shower profile in calorimeters
  - ID efficiency using Particle ID



e μ π K p

- From PID – Different vertex patterns have different vertex mass patterns
- examples
  - 1) K+π vs. π+π
  - 2) π+π vs. π+π+π

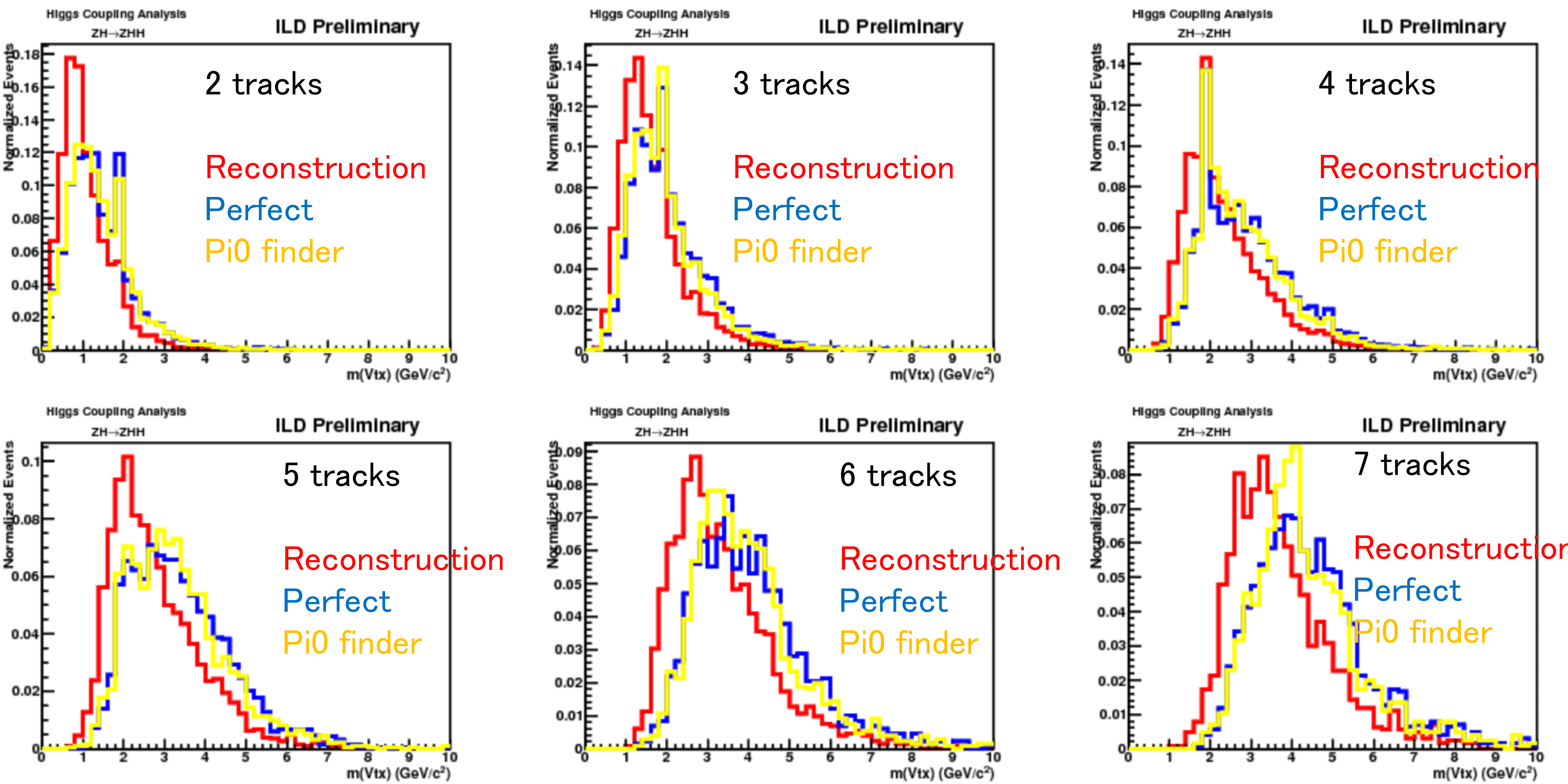




- Can good and general classifier(MVA) to separate pi0s from vertices and pi0s from IP be constructed?

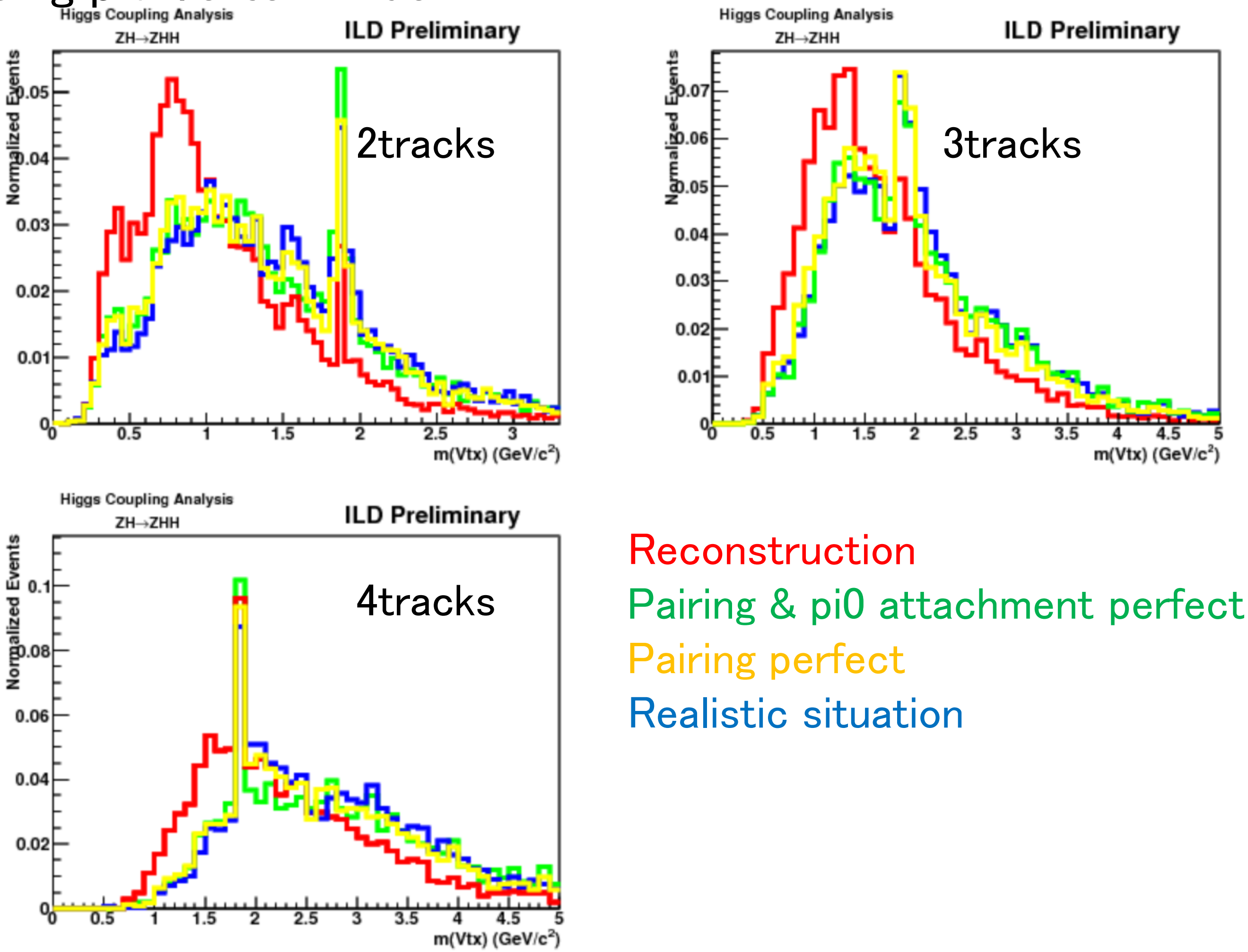
VTX MASS RECOVERY USING PI0 VERTEX FINDER

- Vtx mass distributions for each vertex pattern(ntrk)
  - When a jet has 1 vertex(secondary) inside itself
  - Gamma pairing is perfect within pi0 reconstruction capacity



Most realistic situation

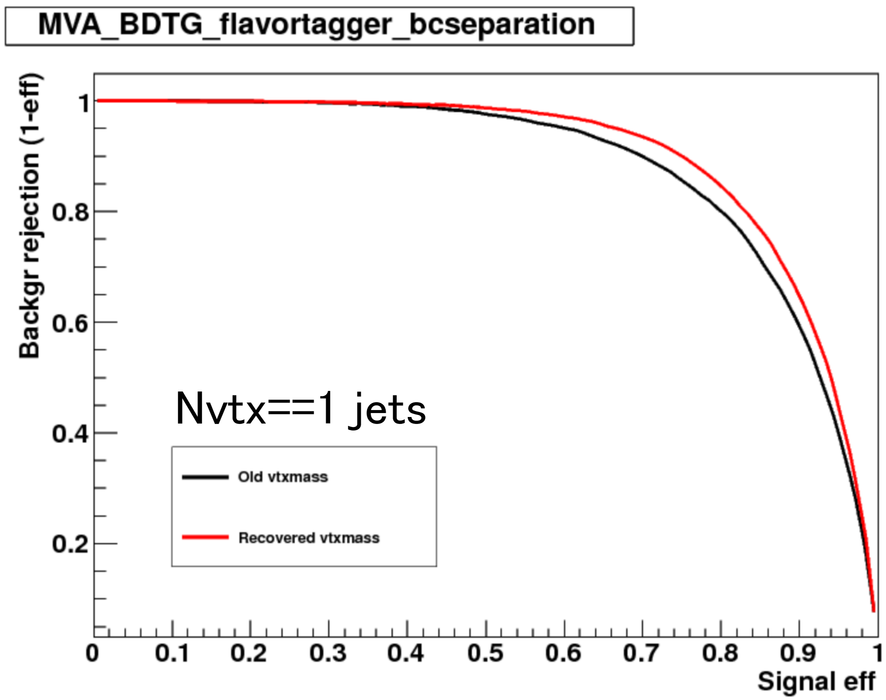
- Pi0 is reconstructed from PFO gamma candidates
  - Using gamma finder
  - Using pi0 reconstruction
  - Using pi0 vertex finder

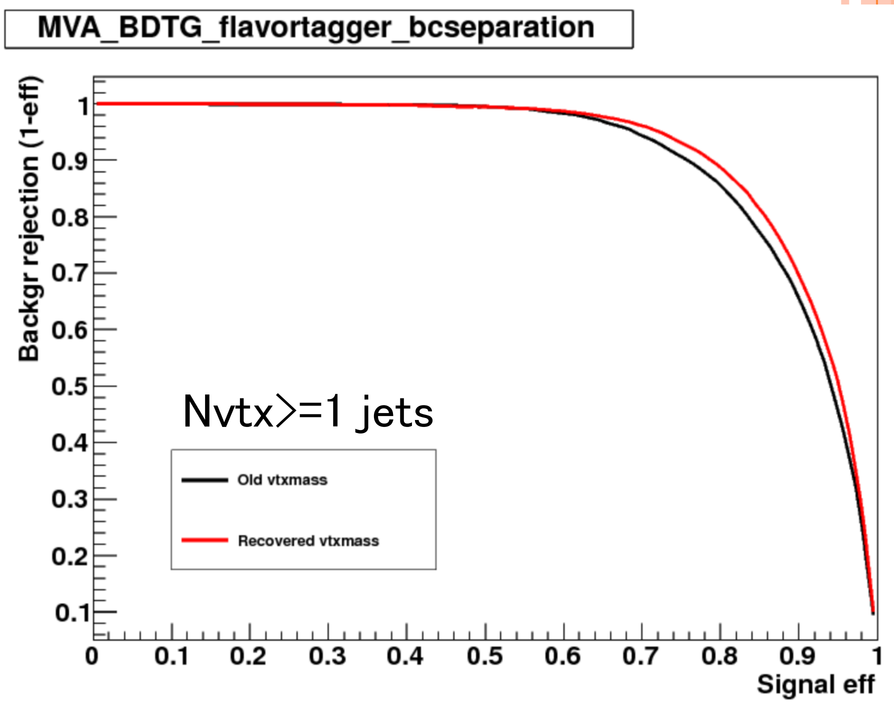


Reconstruction
Pairing & pi0 attachment perfect
Pairing perfect
Realistic situation

Vertex mass recovery effect on flavor tagging

- Construct a "toy" flavor tagger
  - Variables are obtained from LCFIPlus
  - Input variable selection is too primitive!
  - Only vertex mass is replaced to recovered vertex mass
  - Compare with ROC curve





- For more precise study, need to step into LCFIPlus

Summary and plan

- Higgs self coupling analysis using the events with $H \rightarrow WW^*$ is ongoing.
  - Improvement of basic analysis components is necessary
  - Effect of track energy correction using particle ID is small, but going to good direction
  - Kinematic fitting will be a good tool for mass resolution improvement
  - There is hope for flavor tagging improvement!

- **Plan:**
  - Optimization of kinematic fitter & performance check of signal & backgrounds
  - Flavor tagging study inside LCFIPlus
  - Study of flavor tagging improvement in 0 vertex jet case

  - Finally, incorporate all the improvements and update the self-coupling result!

BACKUPS

Analysis strategy for HH→(bb)(WW)

- Classify the events with Z and W decays:

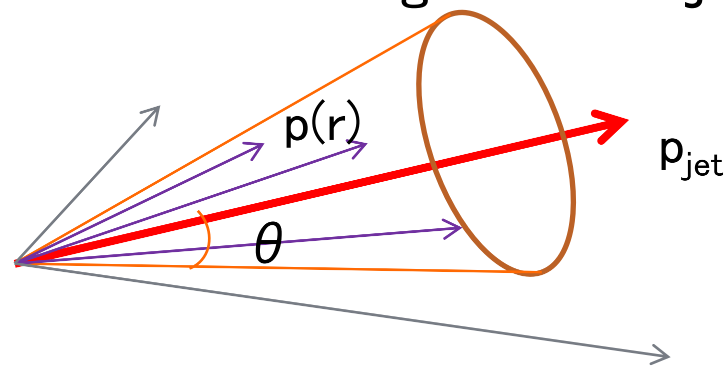
@500GeV	WW→(qq)(qq)	WW→(qq)(l ν)
Z→bb	8jets	Lepton+6jets
Z→cc	8jets	Lepton+6jets
Z→ll	Dilepton+6jets	Trilepton+4jets

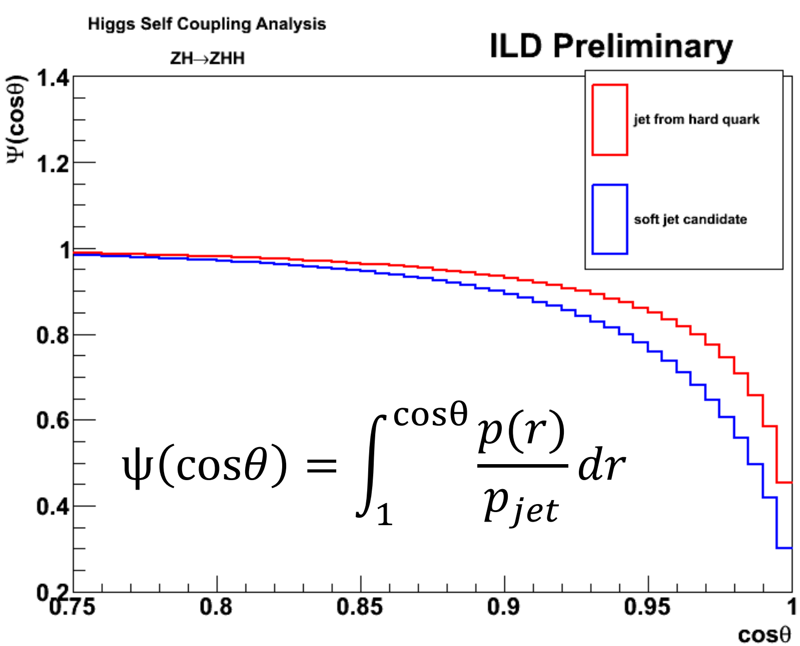
@1TeV	WW→(qq)(qq)	WW→(qq)(l ν)
Z→bb	8jets	Lepton+6jets
Z→ll	Dilepton+6jets	N/A
$\nu\nu$ HH	6jets (+missing)	N/A

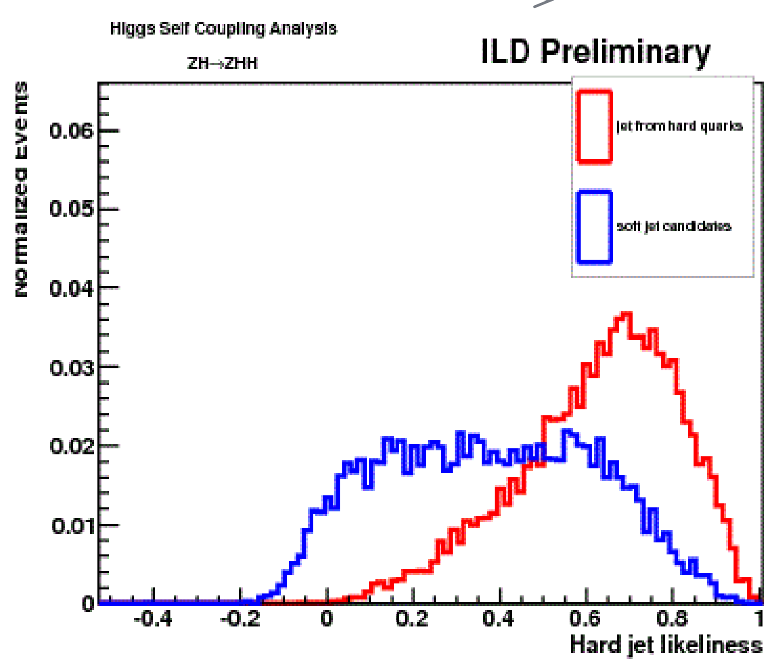
- **Z decays into heavy flavor pair or lepton pair mainly**
  - Need flavor tagger or clean Z mass distribution to reject huge backgrounds
- Number of b jet candidates in the event and number of leptons can form exclusive samples
  - Number of b-tagging available: up to 4
    - Basically, 2 or 4 b-tagged jets events can be used
    - c-tagging is also available
  - Number of leptons: from 0 to 3

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

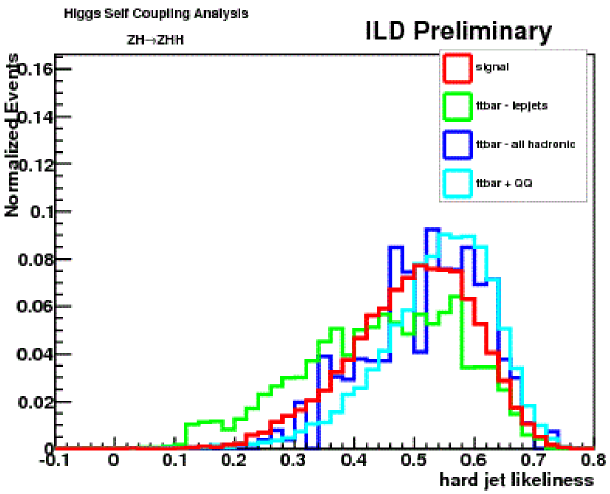


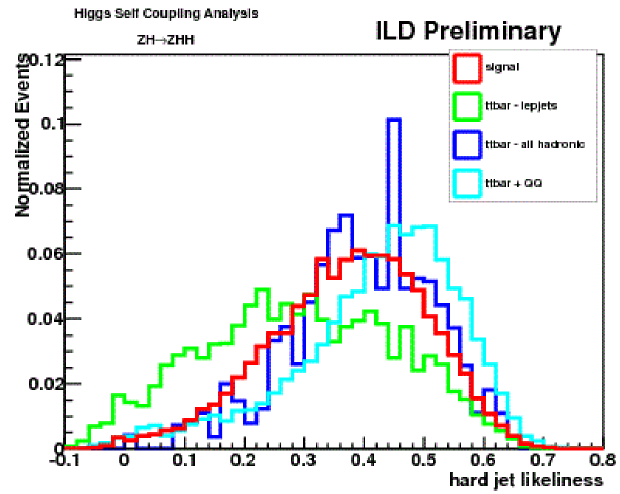




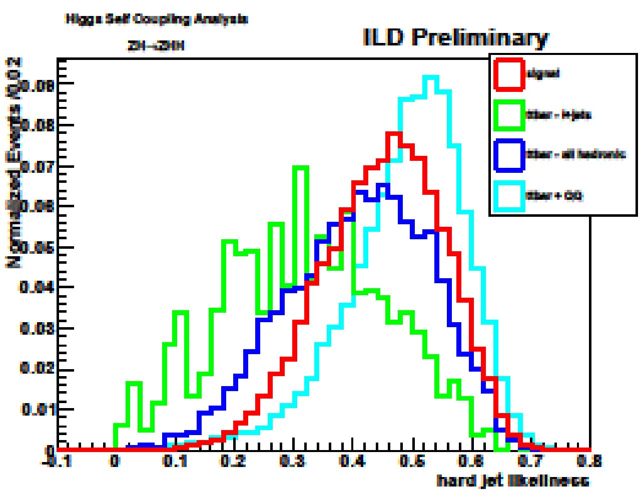
Check the performance

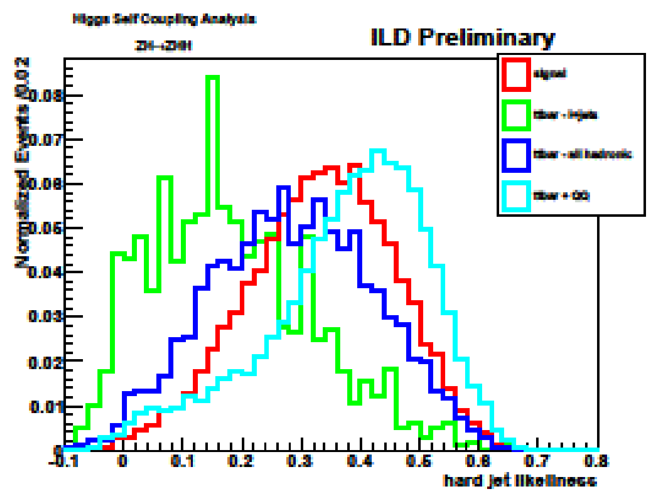
- Check the jets with small hard jet likeliness – signal vs. ttbar
- For 6jets





- For 8iets

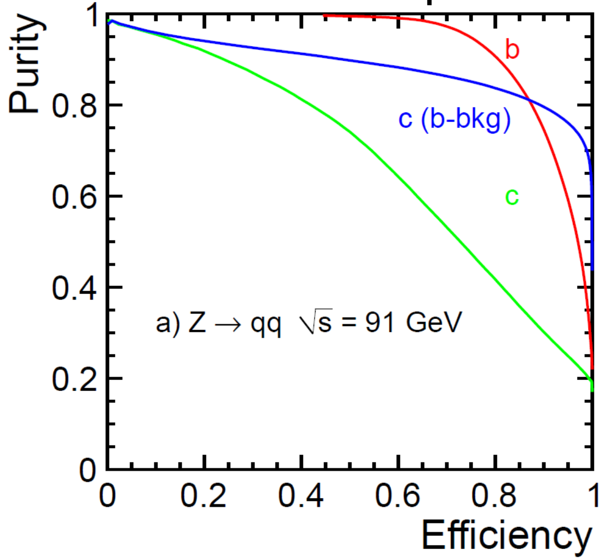


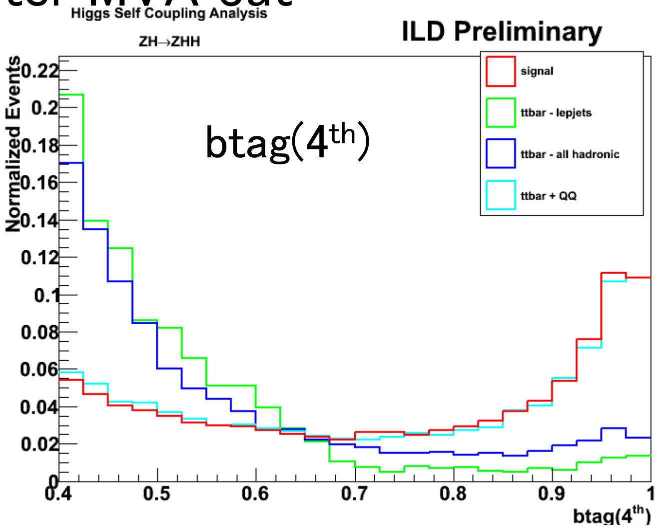


FLAVOR TAGGING

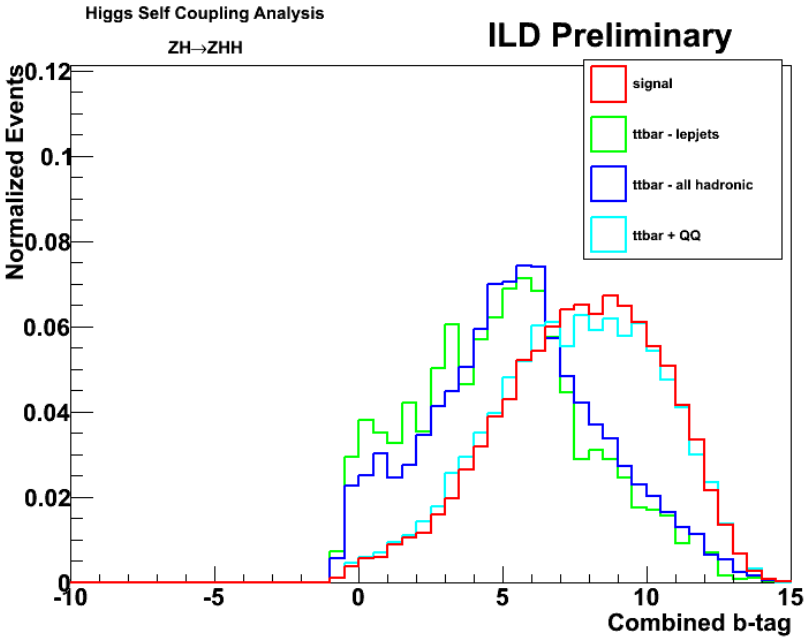
Using LCFIPlus

- b candidate is set >0.4
- Final b-likeliness is optimized after MVA cut





- Introduce combined b-tagging
  - After solving the jet pairing
  - $b(Combined) = \log(\frac{b_1 b_2}{(1-b_1)(1-b_2)})$
  - Use as an input variable for MVA



Background Rejection

- Multi Variate Analysis @500GeV
  - Some cuts are implemented before MVA to tighten the input variable space – jet energy, χ^2, visible energy, (Z mass)
  - Background rejection strategy : rejecting small backgrounds first and then rejecting main background
    - Tighten the variable space when rejecting main backgrounds
  - e.g. all hadronic case:

8jets MVA
- ttbar+QQ
- ttbar+Z
- ttbar+H

→

leptonic
- ttbar(lep+jets, dilepton)
- ZZ+H, ZZZ

→

ttbar hadronic
- All hadronic

- Multi Variate Analysis @1TeV
  - Same strategy as the case of 500GeV
    - VBF process

ZZ+X MVA
- ZZH
- ZZZ

qq

→

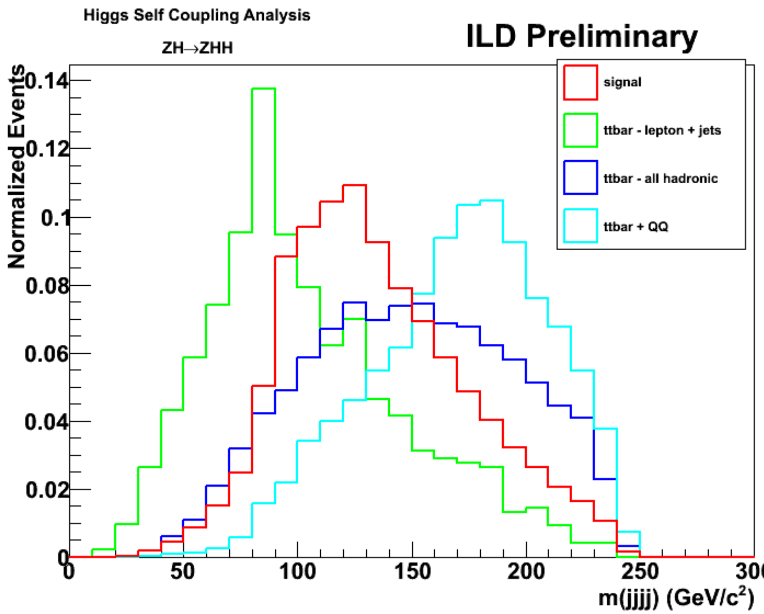
ttbar+X
- ttbar+QQ
- ttbar+Z
- ttbar+H

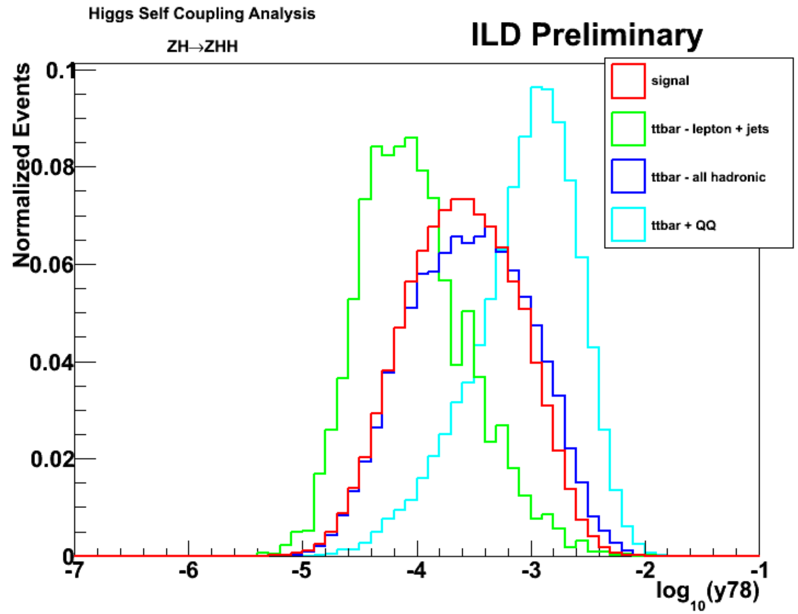
→

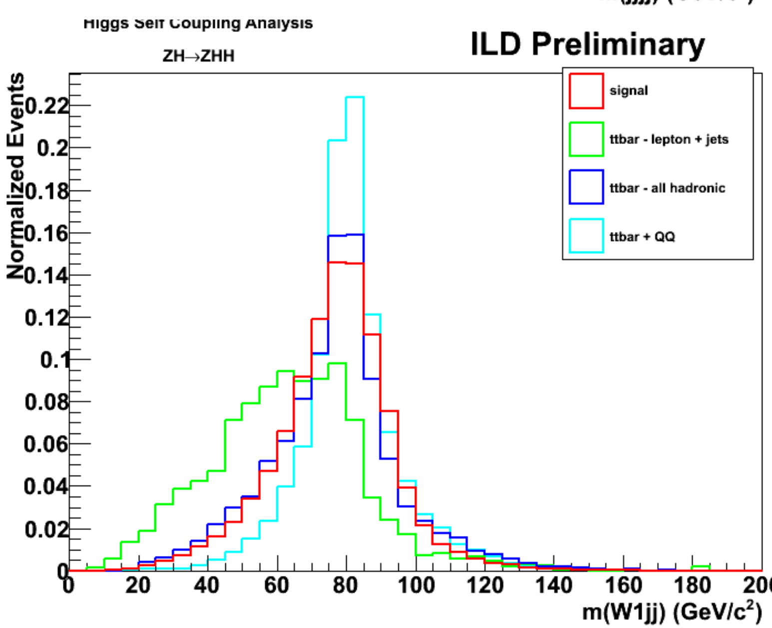
ttbar
- ttbar
- ZWW

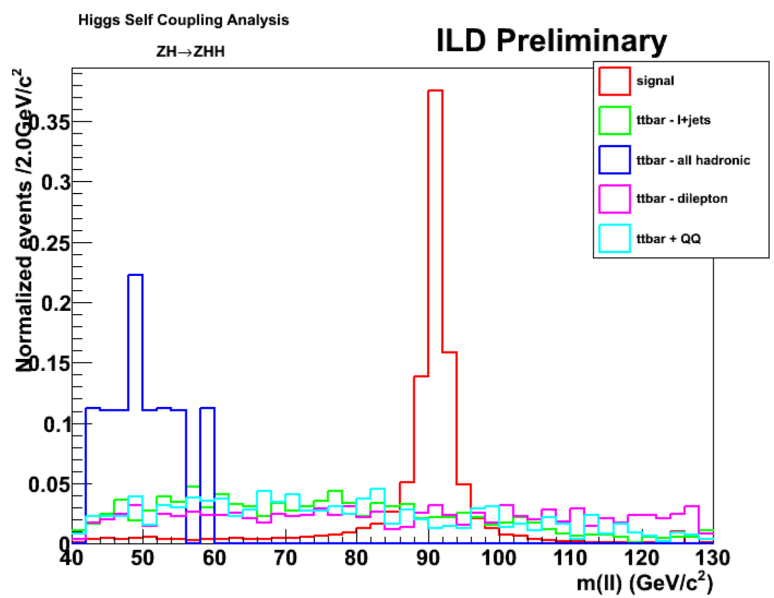
SOME KINEMATIC VARIABLES USED FOR MVA

Very powerful variable @500GeV: m(jjjj), m(l ν jj)



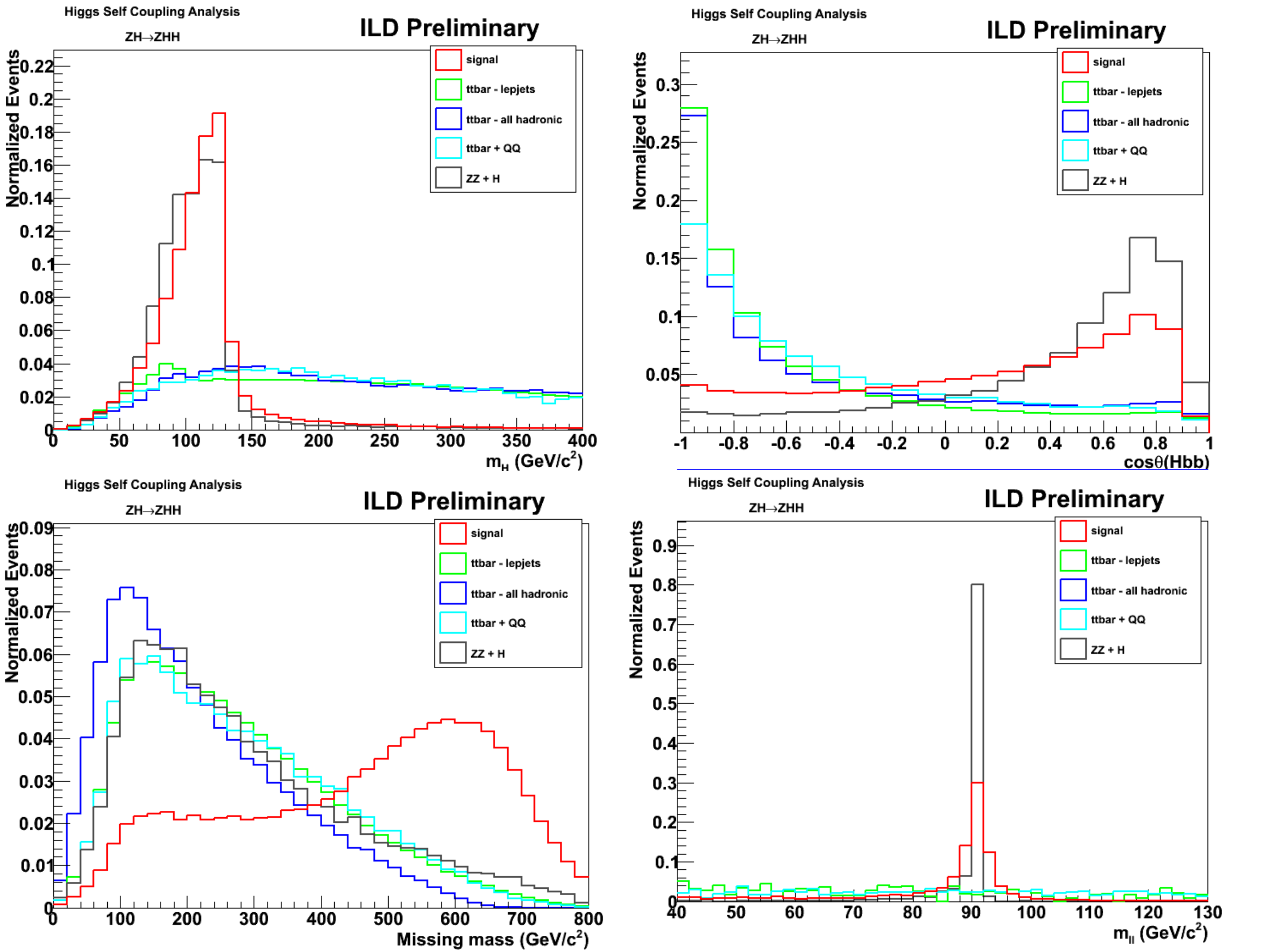






Some kinematic variables used for MVA

Very powerful variable @1TeV: m_H, $\cos\theta$ (Hbb)



Non–simple variables used for MVA

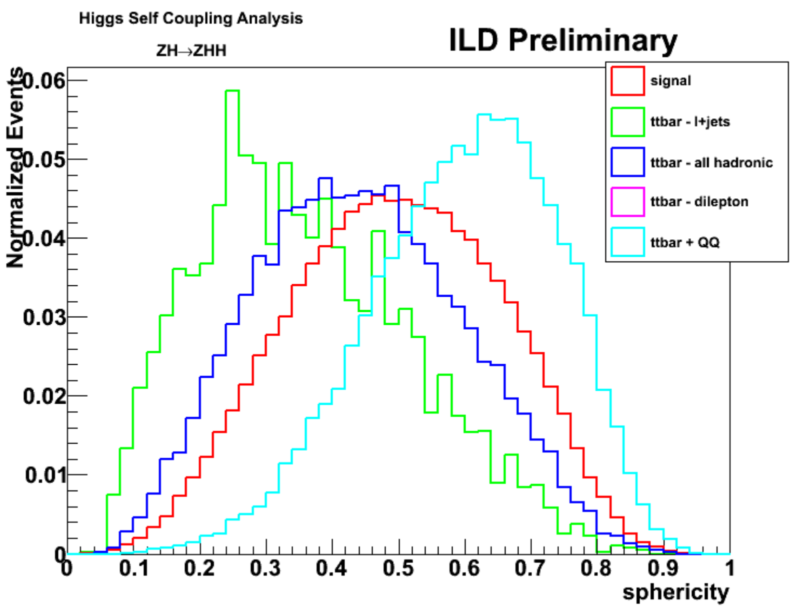
Sphericity and aplanarity

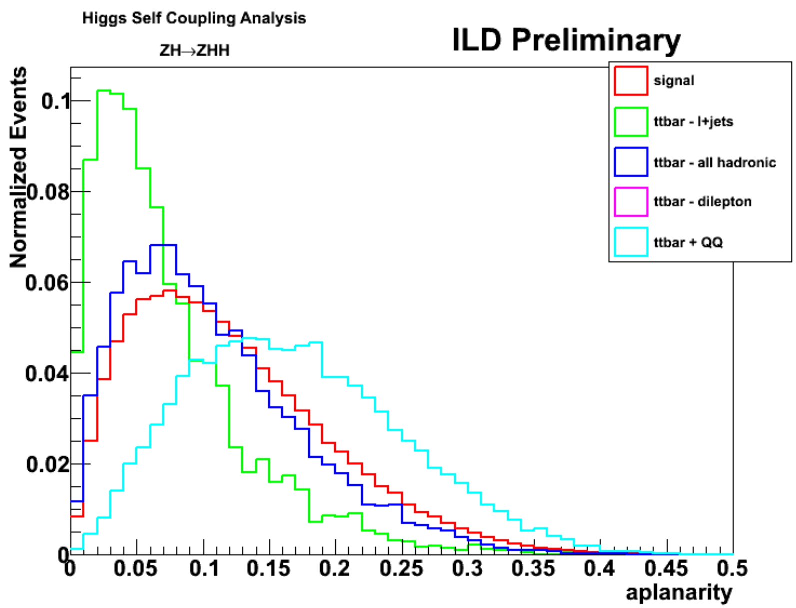
- Eigenvalue combinations of sphericity tensor:

$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\mathbf{p}_i|^2},$$

eigenvalues: $\lambda_1 > \lambda_2 > \lambda_3$

- Sphericity: $S = \frac{3}{2}(\lambda_2 + \lambda_3)$
- Aplanarity: $A = \frac{3}{2}\lambda_3$
- Indicates whether the event is 2–jets like or isotropic



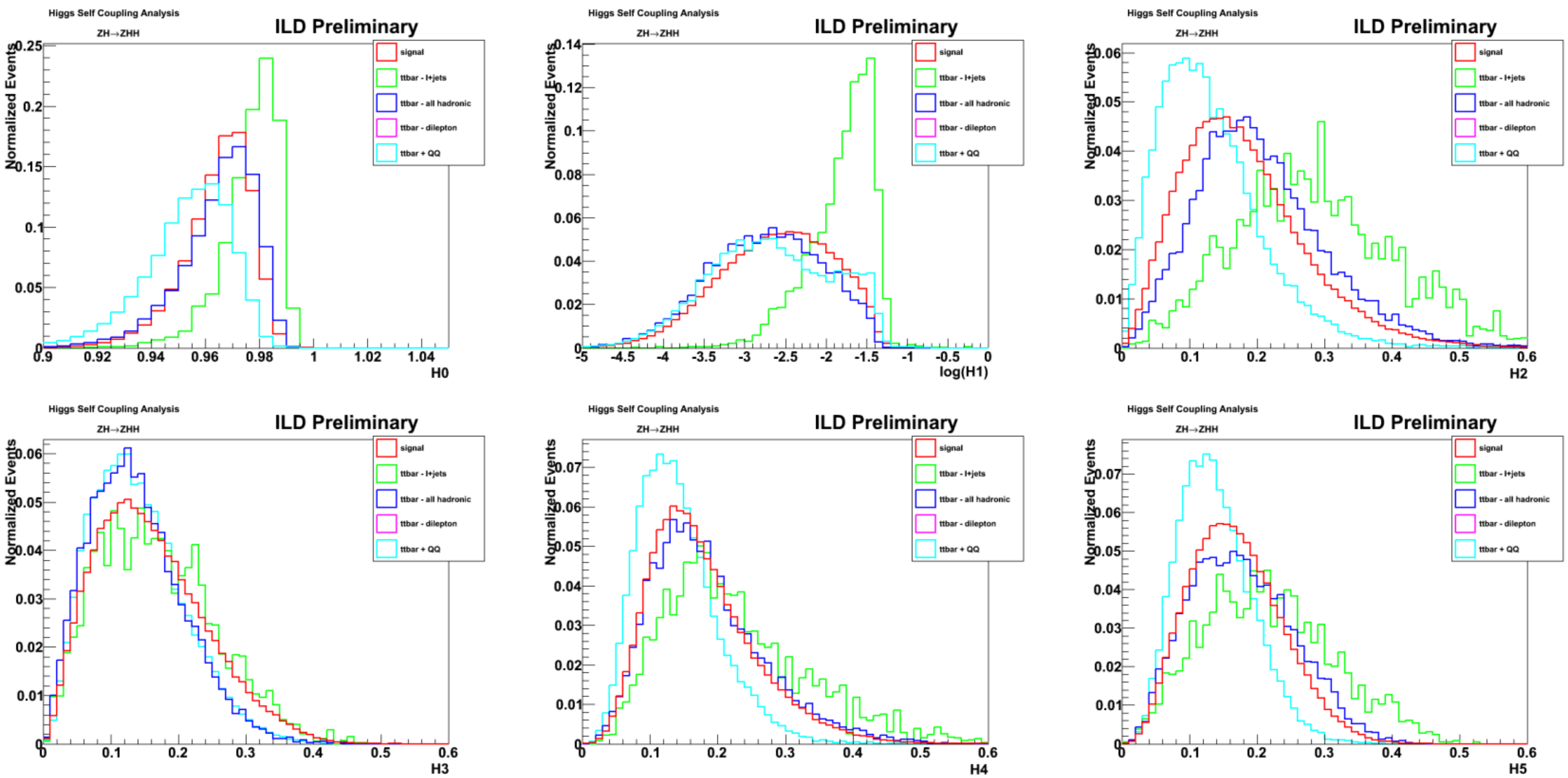


NON–SIMPLE VARIABLES USED FOR MVA

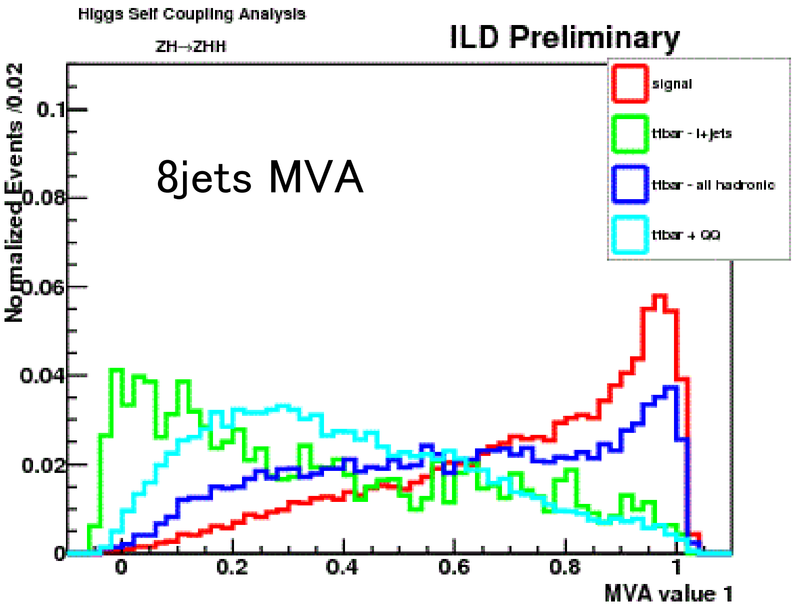
Fox–wolfram moments

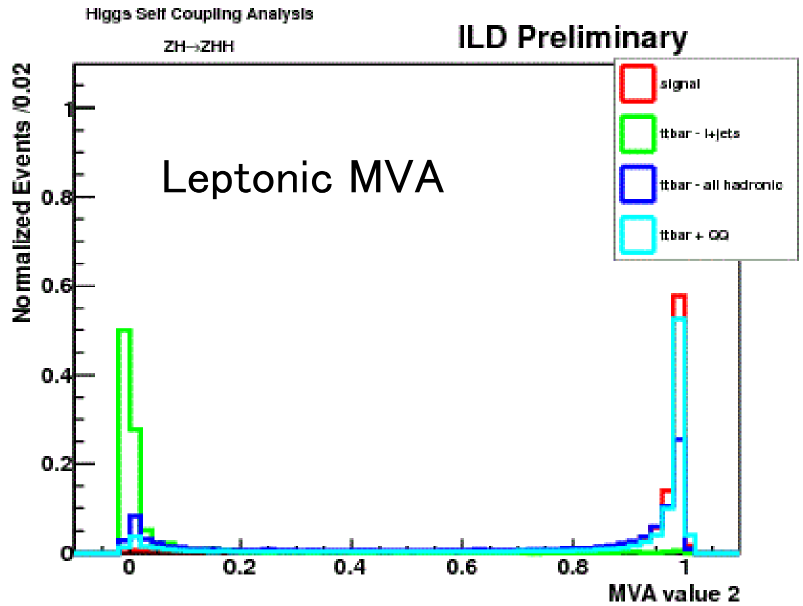
$$H_l = \sum_{i,j} \frac{|\mathbf{p}_i||\mathbf{p}_j|}{E_{\mathrm{vis}}^2} P_l(\cos\theta_{ij}) \;,$$

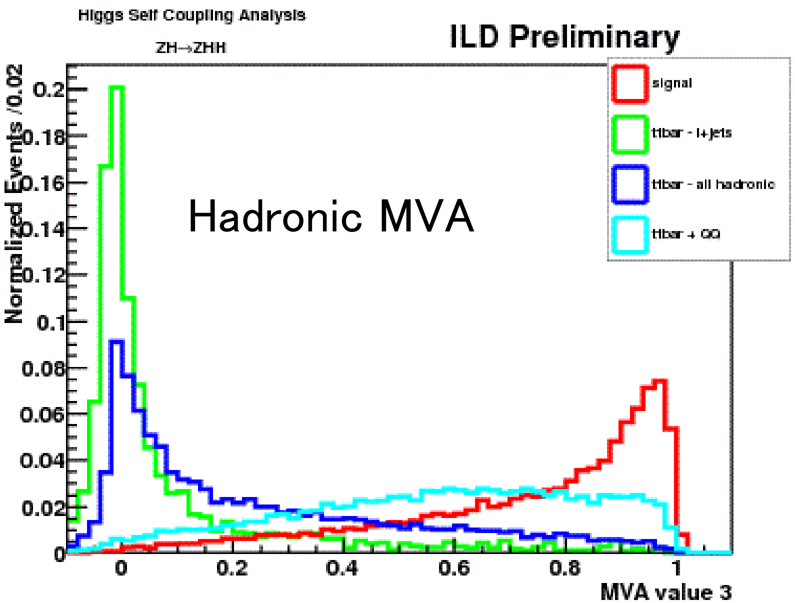
- P_l is Legendre polynomials
- Those moments characterize the structures of 2jets, 3jets, or isotropic events



MVA outputs examples(allhadronic@500GeV)



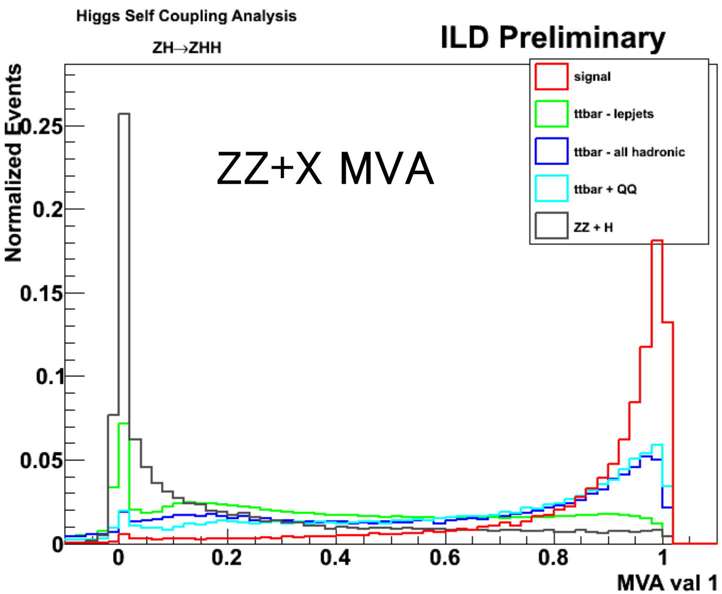


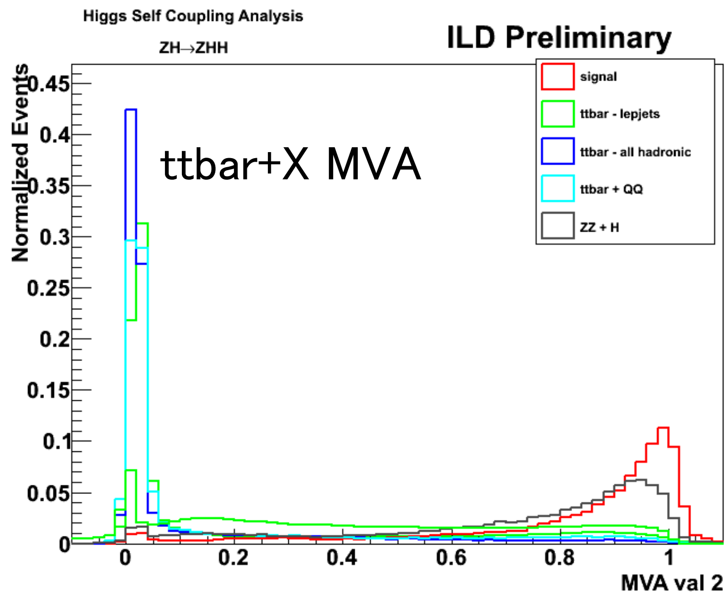


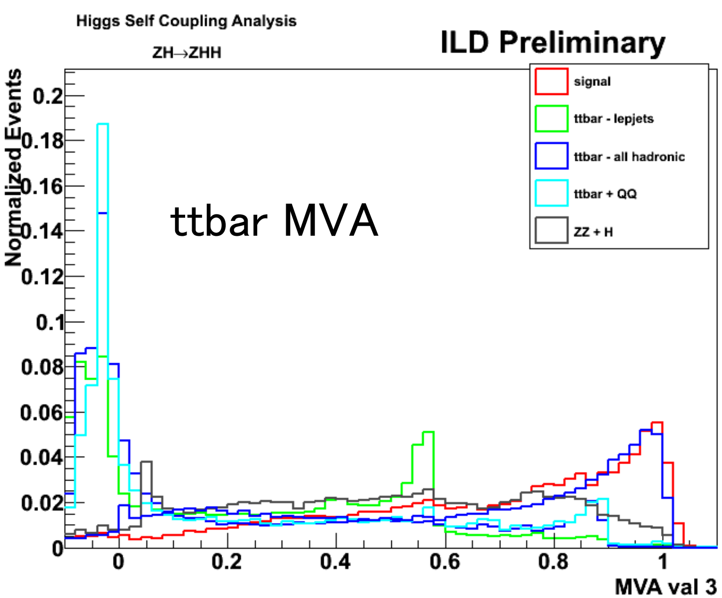
cut of MVA:

- MVA8jets>0.08
- MVAlep>0.02
- MVAhad>0.74

MVA OUTPUTS EXAMPLES(N NHH@1TEV)







cut of MVA:

MVAZZX>0.34

MVAttbarX>0.86

MVAttbar>0.91

Sensitivity@500GeV

- HH→(bb)(WW)
  - As mentioned, categorized with decay tipes of Z and W boson
    - Z→bb, cc or ll
  - b-tagging strategy – introduce looser b-tag category
    - 4-btag & 3-btag
  - E_{CM}=500GeV, L=2ab-1
  - **Significance ～1.91 σ**

Modes	Z decay	b tag	Signal	Background	Significance
All hadronic	Z→bb	4btag	15.20	87.52	1.50 σ
		3btag	19.43	3099.49	0.35 σ
	Z→cc		11.29	366.13	0.58 σ
Lepton + jets	Z→bb		1.65	17.62	0.38 σ
	Z→cc		1.50	819.61	0.05 σ
Dilepton	Z→ll		2.24	8.44	0.69 σ
Trilepton	Z→ll		1.05	2.60	0.55 σ
Combined					1.91 σ

Sensitivity@1TeV

- HH→(bb)(WW)
  - As mentioned, categorized with decay types of Z and W boson
    - Z→bb and ll, VBF channel
  - b-tagging strategy – fully used the b-tagging for each category
  - E_{CM}=1TeV, L=2ab−1
  - **Significance ～2.80σ**

Modes	Z decay	Signal	Background	Significance
All hadronic	Z→bb	17.15	48.17	2.12σ
Lepton + jets	Z→bb	1.16	9.24	0.36σ
Dilepton	Z→ll	1.03	14.30	0.26σ
6jets+ Missing	No Z, νν HH	6.90	8.24	1.77σ
Combined				2.80σ

Summary and plan

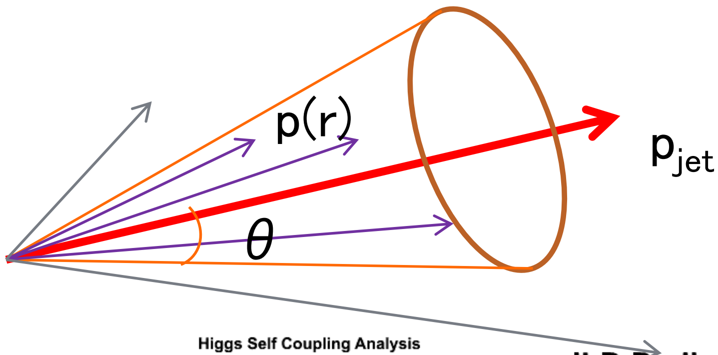
- Higgs self coupling analysis using the events with $H \rightarrow WW^*$ is ongoing.
  - Multi variate analysis to reject the backgrounds
  - Total sensitivity @500GeV is ~1.91 σ
  - Total sensitivity @1TeV is ~2.80 σ

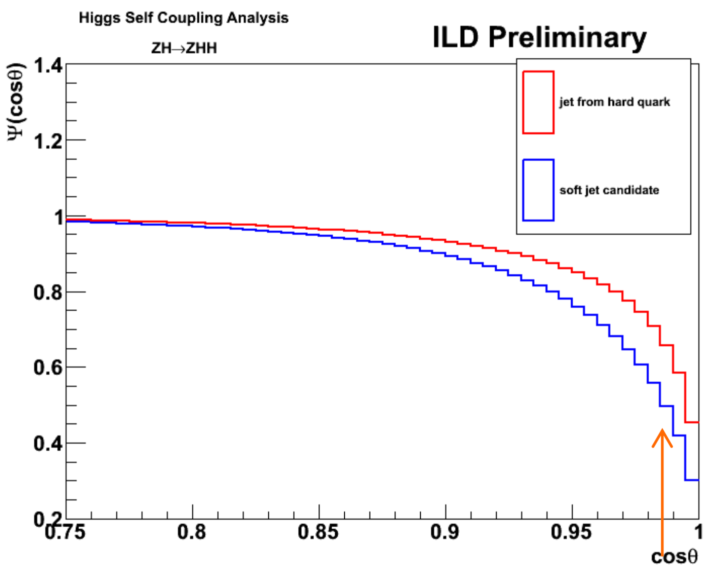
- **Plan:**
  - Start to combine with golden channel and estimate the Higgs self coupling
  - Full simulation @1TeV
  - Optimize b-tagging strategy
    - Forming looser b-tag category
  - Improvement of basic components for the analysis
    - Lepton ID
    - b-tagging
    - Jet energy correction
    - Jet clustering

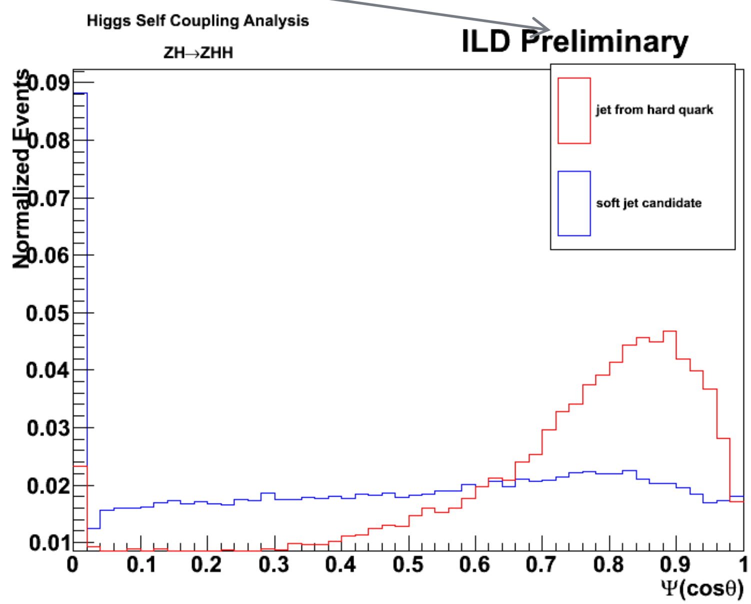
SOFT JET FINDING

- Soft jet finding may be available for the events with extra jets not coming from hard process quarks
  - e.g. 8 jets requirement to ttbar hadronic events(6 jets from hard quarks)
- Traditional jet shape indicates the same tendency as hadron collider analysis

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



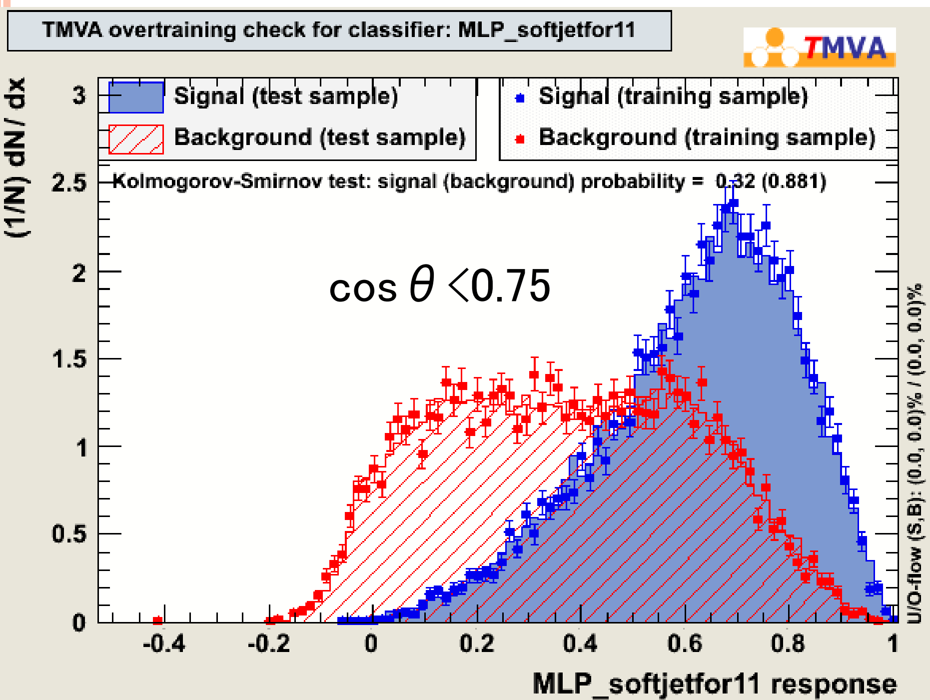


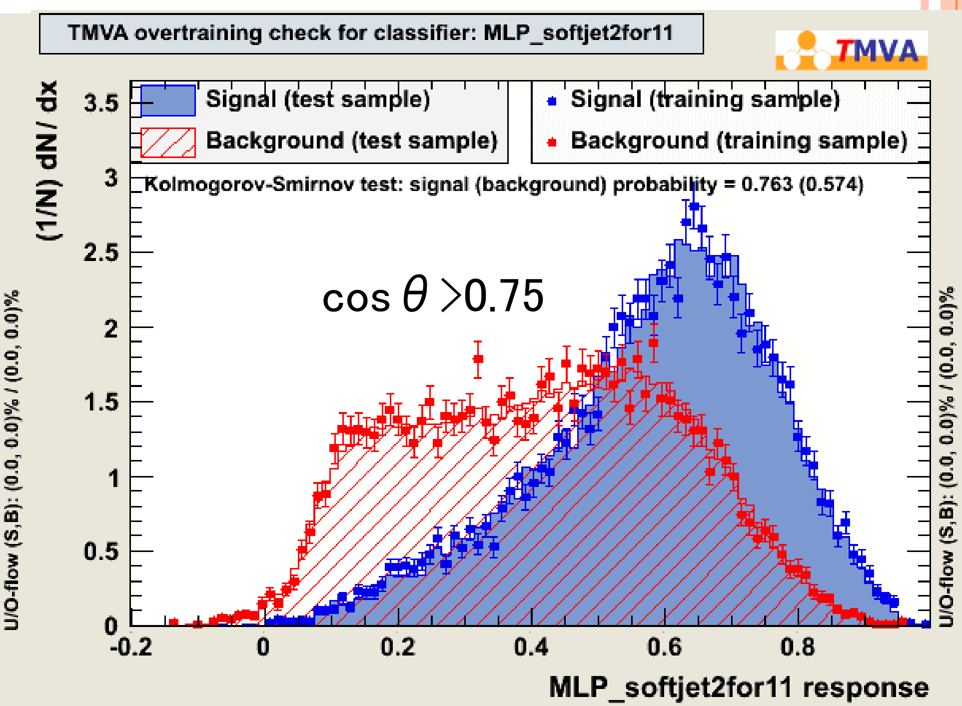


Soft jet finding

- 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
$\chi 2$	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