Higgs Analysis II
Top-Yukawa and Higgs Self-coupling

Masakazu Kurata, Tomohiko Tanabe
The University of Tokyo
Junping Tian, Keisuke Fujii
KEK
Claude-Fabienne Durig、Jenny List
DESY
Taikan Suehara, Yuji Sudo
Kyushu University

Tokusui workshop, 12/08/2015-12/10/2015

Components for better results (in General)

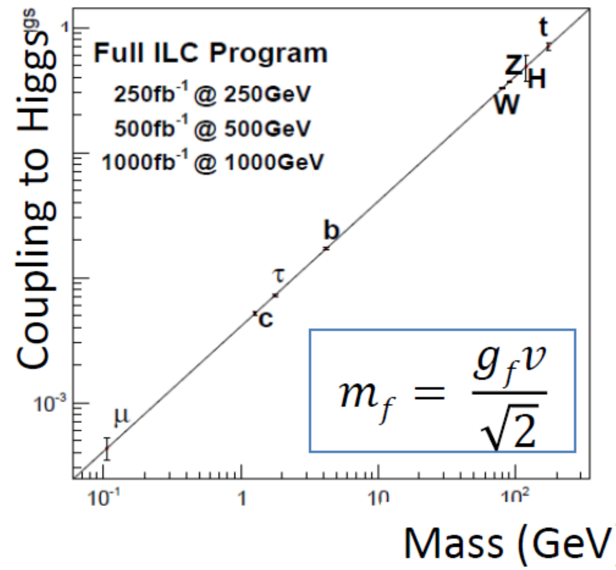
- Basic components for better sensitivity
  - Lepton ID: Isolated leptons can be identified well, and very good fake suppression
    →many idea have been introduced
  - B-tagging: better b-tagging algorithm provides better background suppression
  - Jet pairing: good jet pairing can obtain good kinematic variables, which leads to good background suppression
    - Using likelihood jet pairing for ttH analysis
  - 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
    - Some people start to study
  - Good background rejection: of course main theme in analyses
    - Of course, MVA will be a main tool
    - →Focus on Kinematic fitter for Higgs self-coupling analysis

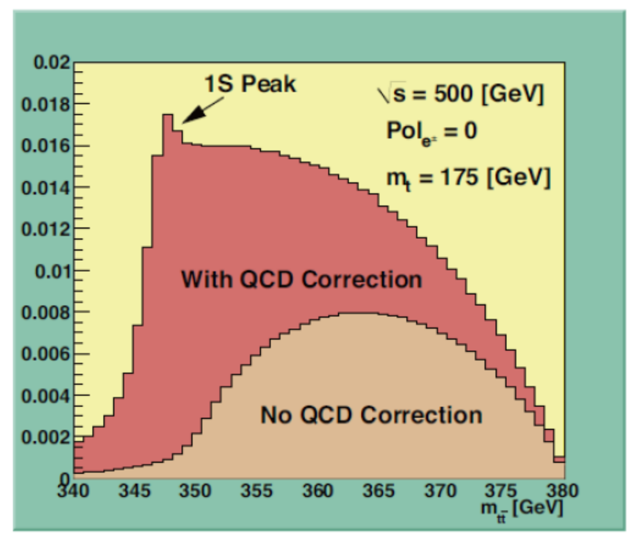
- All the components are related each other

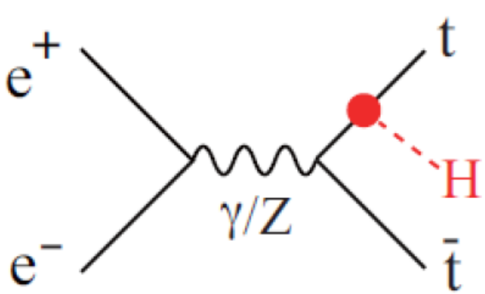
Introduction of top–Yukawa analysis

- We are working on tth study
  - M_h=125 GeV.
  - Polarization : $(Pe^-,Pe^+)=(\mp0.8,\pm0.3)$
  - Integrated luminosity 500 fb^{-1} and H20
  - ILD full simulation
- ttbar cross section is increased around ttbar threshold by ttbar bound-state effect
- tth cross section is enhanced
- ttZ cross section is also increased

We can directly measure the top-Yukawa coupling via tth channel at $\sqrt{s}$ = 500 GeV.







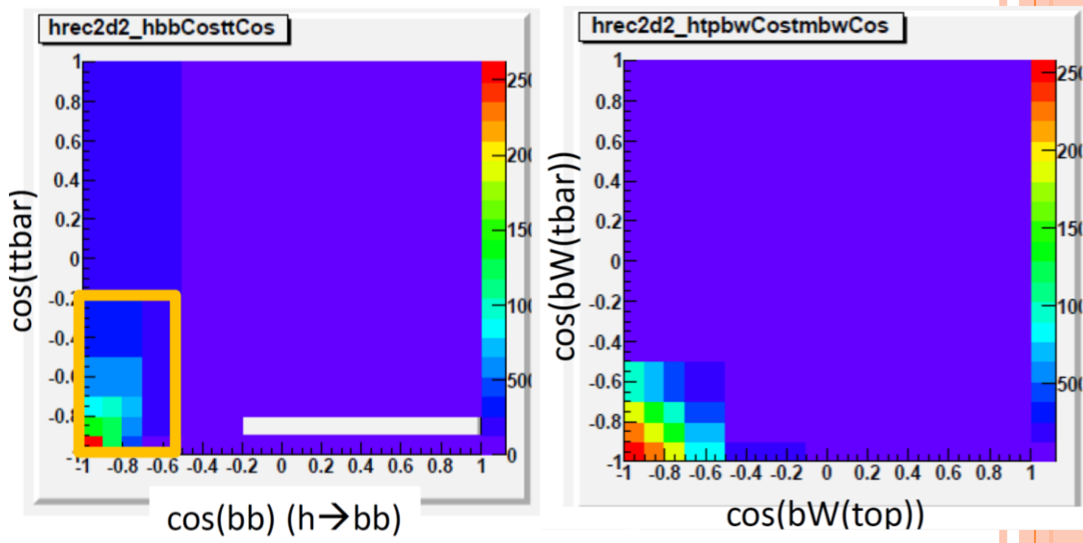
Jet pairing using likelihood method

- Jet pairing including not only mass constraint, but also kinematics will improve the pair–matching efficiency
  - So, should use Likelihood method
- Likelihood templates can be obtained from signal events
- Use some kinematic variables as well as mass constraint

Template examples: use 2D likelihood for correlation

- 8jets

* 2D likelihood template
cos(tt), cos(bb(higgs))
cos(bW(anti-top)), cos(bW(top))
top1 mass, W1 mass
top2mass, W2 mass

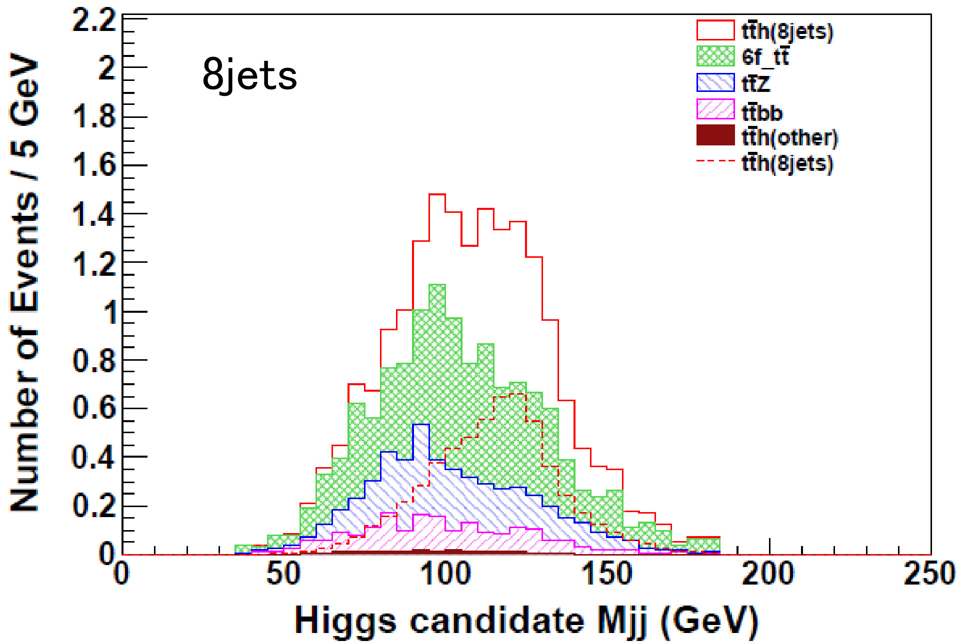


- ～10% improvement on a pair–matching efficiency for each signal

  8jets, 1lepton+6jets, 2leptons+4jets

RESULTS

- 3 exclusive category
  - 8jets
  - 1lepton + 6jets
  - 2leptons + 4jets
- Event selection
  - Cut based, not MVA



- e.g.) 8jets mode
- Significance $S/\sqrt{S+B}$ for each int. luminosity

Integrated Lumi. (fb^{-1})	(-0.8,+0.3)	(+0.8,-0.3)
500	2.11	1.42
200	1.33	0.90
1400	3.53	2.38
Mjj range (GeV)	$85<M_{jj}<175$	$75<M_{jj}<175$

Expectation of running scenario and energy

- H–20 scenario: $S/\sqrt{S+B}$=6.6, $\left|\frac{\Delta g_t}{g_t}\right| = 7.9\%$

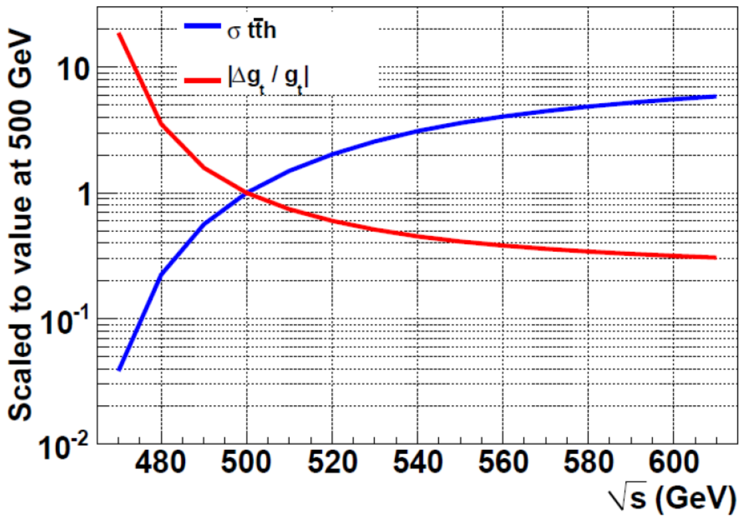
	Integrated Lumi. (fb^{-1})	(-0.8,+0.3)	(+0.8,-0.3)	combined (-0.8,+0.3)&(+0.8,-0.3)
	500	3.05	2.07	3.68
H20	200	1.92	1.30	2.32
	1400	5.11	3.46	6.18
	200&1400	5.46	3.69	6.61

- Energy: more than twice better for coupling precision@550GeV!

ttH 200+1400 fb^{-1}

$\sqrt{s}$	$S/\sqrt{S+B}$	$\|\Delta g_t/g_t\|$ %
490	4.22	12.2
500	6.61	7.85
510	8.86	5.86
520	10.9	4.76
530	12.7	4.06
540	14.4	3.60
550	15.8	3.28
600	20.5	2.53

σ_{ttH} = 0.4088 fb @ 500 GeV



Higgs Self–Coupling

- 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

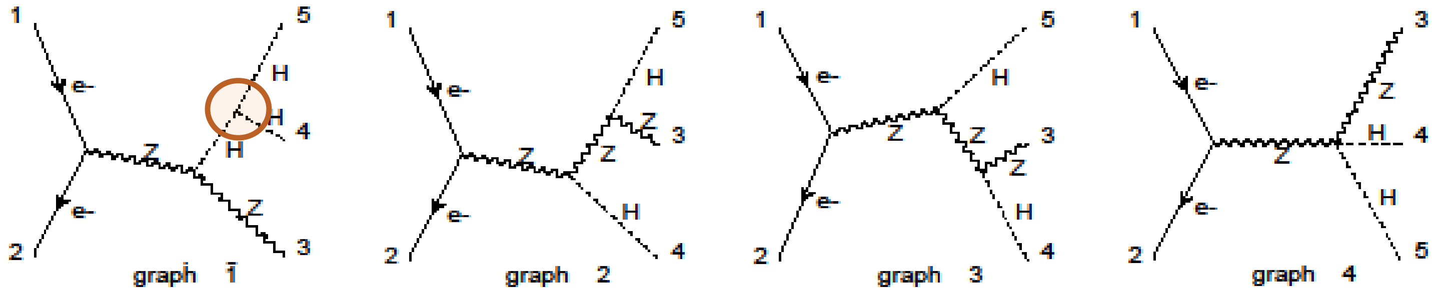
$$SM: \quad \lambda = \frac{m_H^2}{2v^2} \quad v \sim 246 GeV$$

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

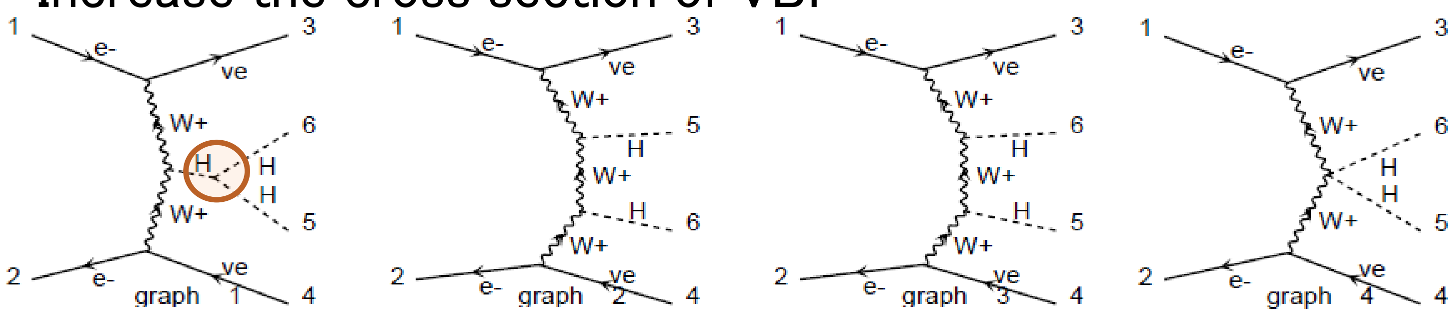
Signal Events

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



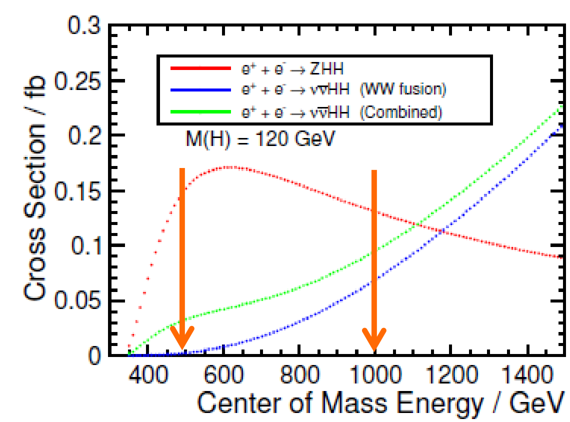
Signal: 1 Irreducible B.G.: 2, 3, 4

- Signal@1TeV – VBF $e^+e^- \rightarrow \nu\nu HH$ 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

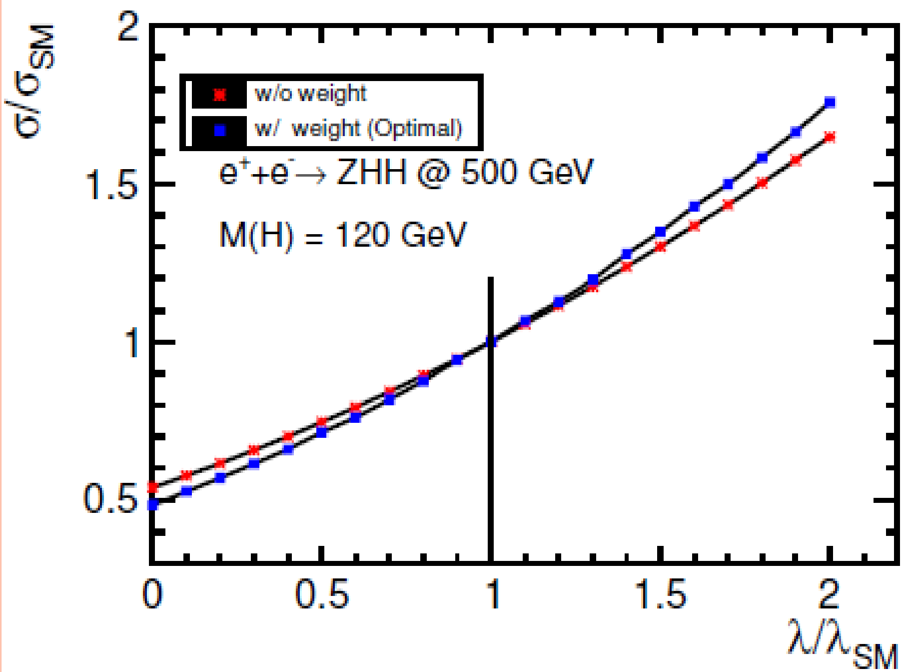


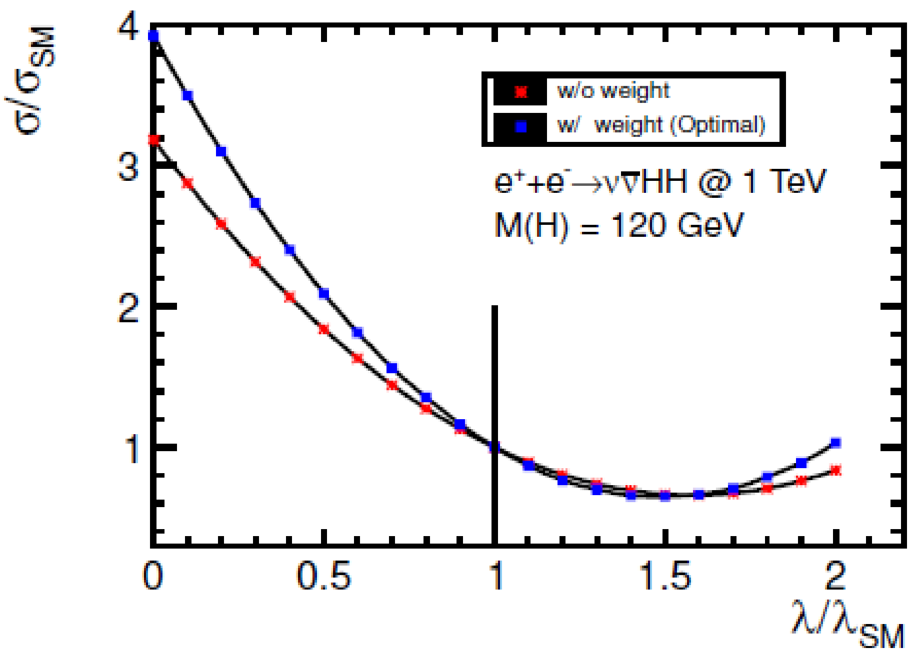
Irreducible background effect on measurement

- Irreducible backgrounds cause interference with signal process:
  - If no interference, $\sigma \propto \lambda^2 \rightarrow \frac{\Delta\lambda}{\lambda} = 0.5 \frac{\Delta\sigma}{\sigma}$
  - Due to the interference: $\sigma = \lambda^2 S + \lambda I + B$
    - I: interference term B: background term

$$\rightarrow \frac{\Delta\lambda}{\lambda} = 1.62 \frac{\Delta\sigma}{\sigma}\ (@500\mathrm{GeV}),\ \frac{\Delta\lambda}{\lambda} = 0.73 \frac{\Delta\sigma}{\sigma}\ (@1\mathrm{TeV})$$

  - Huge degradation of self-coupling measurement precision…



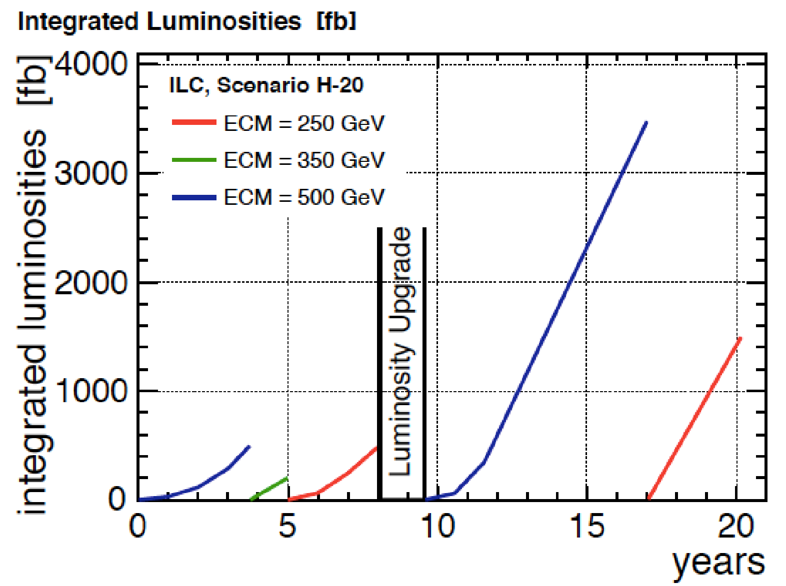


Expectation from The Running Scenario

- Expectation so far:
  - In H−20 scenario, we will be able to obtain 4ab^{-1} @500GeV as a full dataset
  - Precision of Higgs self-coupling $\Delta\lambda/\lambda$:

Luminosity	500GeV	1TeV
500fb−1	77%	−
4ab−1	**27%**	−
2ab−1	−	16%

  - ～30%@500GeV in full ILC program
  - 10% precision→5ab^{-1}@1TeV



ECM	Initial phase	Lumi upgrade	total
250 GeV	500 fb-1	1500 fb-1	**2000 fb-1**
350 GeV	200 fb-1		**200 fb-1**
500 GeV	500 fb-1	3500 fb-1	**4000 fb-1**
runtime	8.1 yrs	10.6 yrs	**20.2 yrs**

- It is very challenging analysis!
  - We need to explore the possibility of better result
  - First, 500GeV(baseline!) improvement is necessary
  - 1TeV improvement is also essential for good precision

Kinematic Fitting - ISR Treatment

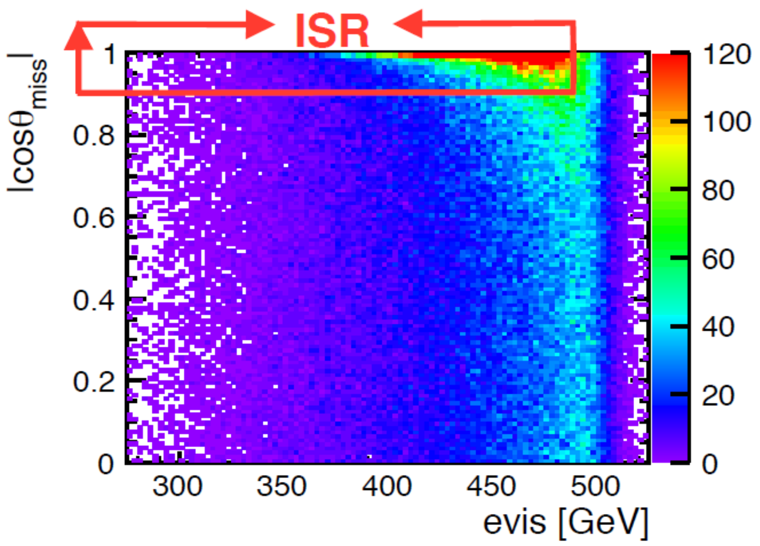
so far: $\sum \left(E_i, \vec{P}_i\right) = \left(\sqrt{s}, \vec{0}\right)$

ISR and beamstrahlung in samples

$$\sum \left(E_i, \vec{P}_i\right) = \left(\sqrt{s} - E_\gamma^{ISR}, -\vec{P}_\gamma^{ISR}\right)$$

considering ISR in fit (ISRPhotonFitobject)

- works well for light jets
  (diploma thesis M.Beckmann)
- in b/c jets E_{miss} due to $\nu \rightarrow$ special handling



ISR not considered in fit:

- energy assigned to jets (E, $\vec{P}$ conserved)

Problem: events with ISR

- larger fitted 4-momenta of jets
- bias to large masses

ISR considered in fit:

- certain amount of energy assigned to ISR

Problem: events without ISR

- "fake" ISR, energy missing to jets
- bias to small masses

Correct treatment of ISR in fit on events by events basis using ISR characteristics

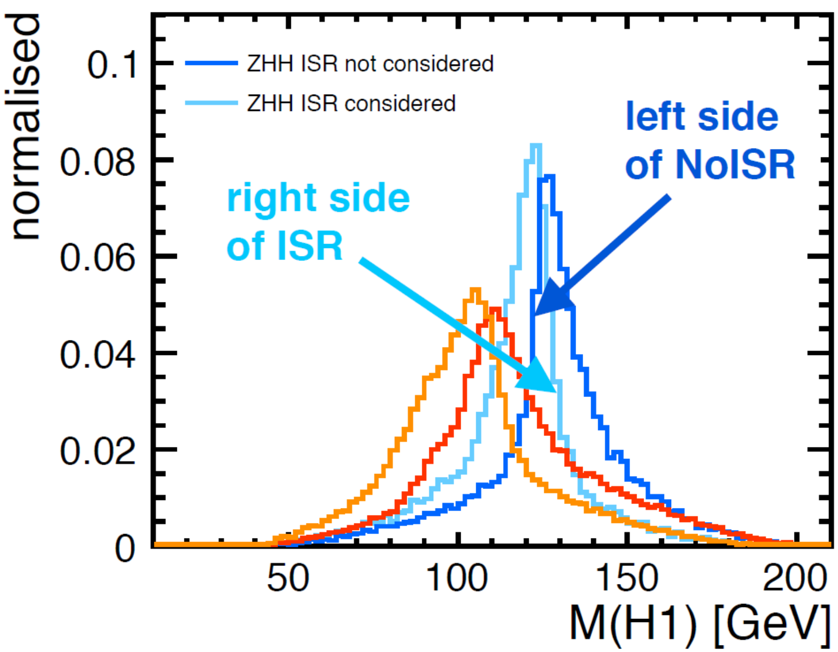
STRATEGY FOR ISR TREATMENT

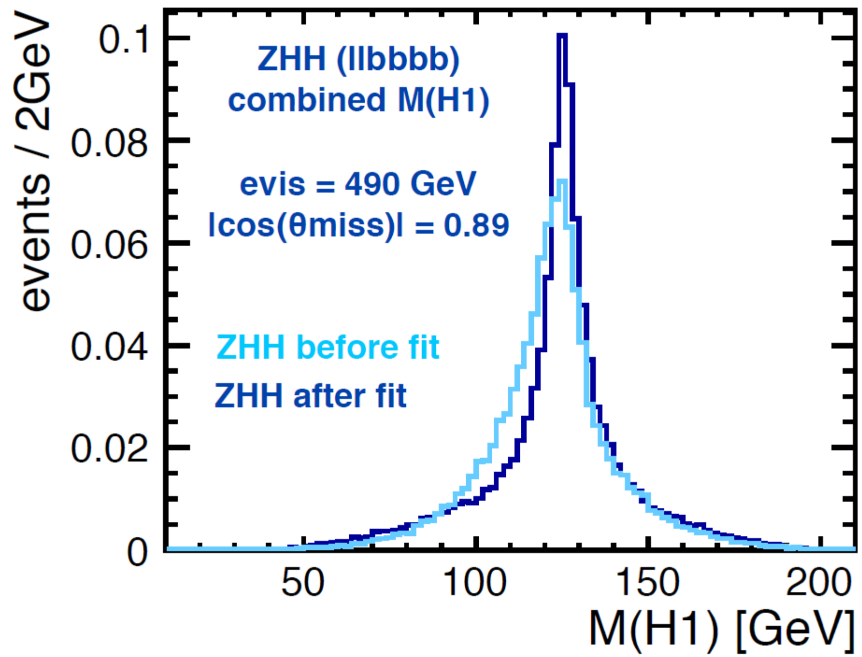
events need to be selected by their amount of missing energy and $|\cos(\theta_{miss})|$ values

➤ divide events into two categories

ISR $\quad evis < X$ GeV && $|\cos(\theta_{miss})| > 0.YY$

no ISR $\quad evis > X$ GeV || $|\cos(\theta_{miss})| < 0.YY$





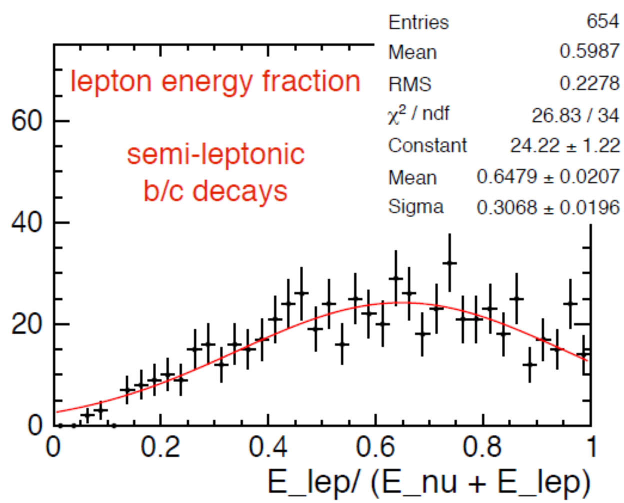
Mass resolution is significantly improved
Using corrected ISR treatment

20% improvement in llHH mode due to usage of χ^2 in neural nets!

further improvement by using new mass reconstruction in neural nets expected

Impact of Semi-leptonic b/c Decays

if lepton in jet, correct missing energy from neutrino: $E_{jet}^{corr} = E_{jet} + E_{\nu} = E_{jet} + x \cdot E_{lep}$



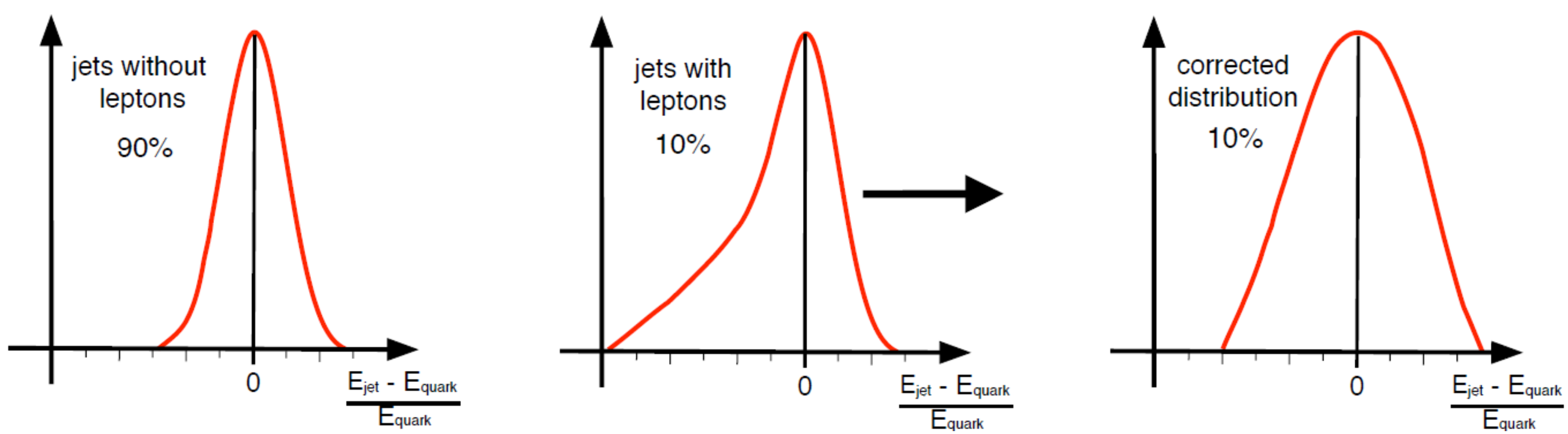
$$x = \frac{E_{lep}}{(E_{lep} + E_{\nu})} \Leftrightarrow E_{\nu} = \left(\frac{1}{x} - 1\right) E_{lep}$$

$$E_{jet}^{corr} = E_{jet} + \left(\frac{1}{\langle x\rangle} - 1\right) E_{lep}$$

$$\boxed{\Delta E_{jet,corr}^2 = \Delta E_{jet}^2 + \left(\frac{\Delta\langle x\rangle}{\langle x\rangle^2}\right)^2 E_{lep}^2} + \underbrace{\cancel{\left(\frac{1}{\langle x\rangle} - 1\right) \Delta E_{lep}^2}}_{\text{track momentum resolution!}} \to 0$$

plot: $\langle x\rangle = 0.65 \rightarrow E_{jet}^{corr} = E_{jet} + 0.54 E_{lep}$

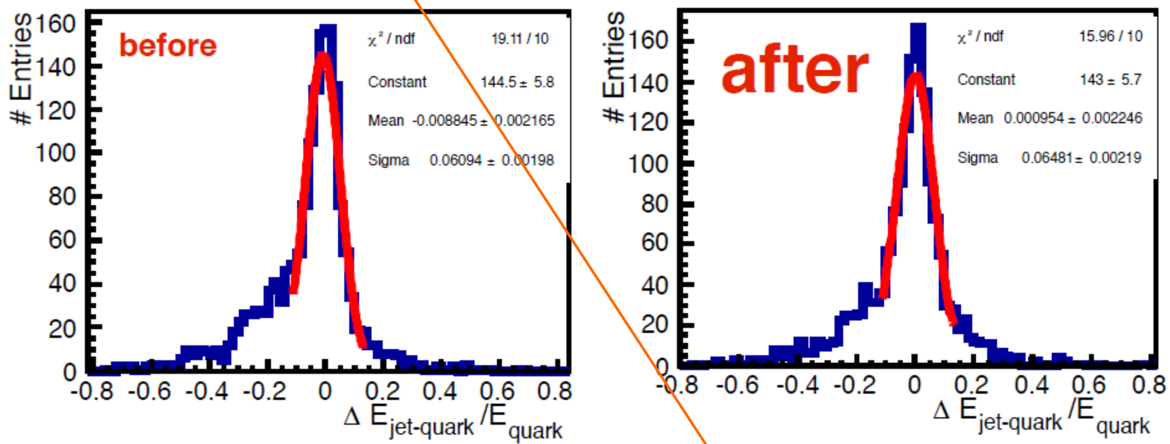
$\Delta\langle x\rangle = 0.31 \rightarrow \Delta E_{jet,corr}^2 = \Delta E_{jet}^2 + (0.73 E_{lep})^2$



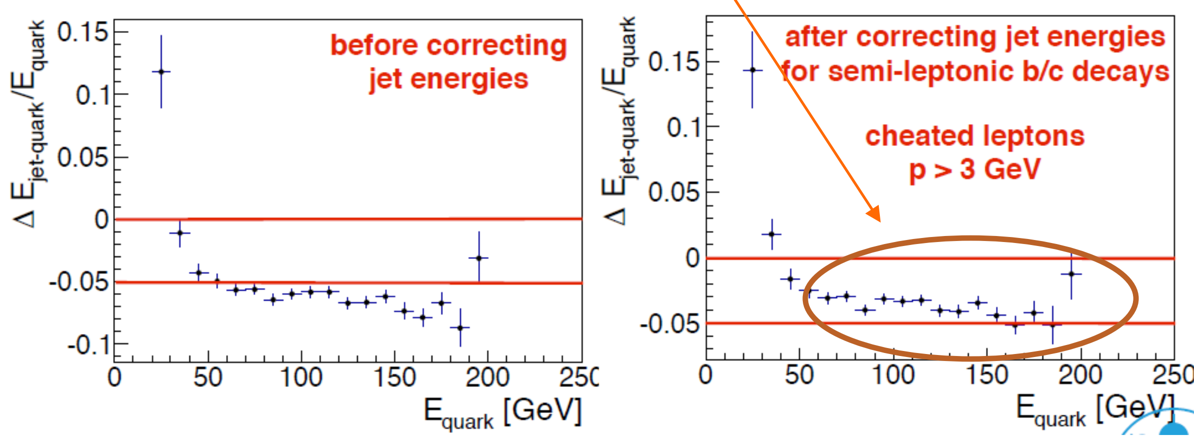
JER IMPROVEMENT FOR SEMI-LEPTONIC BJETS

- Jet energy correction for semi-leptonic bjets
  - Jet energy resolution becomes better

→provides better Kinematic Fitter!



  - Analysis ongoing



remark: samples with standard DBD reconstruction → ParticleID tools not included
following studies with cheated e/μ, with p > 3 GeV

Kinematic Fitter for HH→(bb)(WW) @500GeV

- 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 (or −2log(likelihood)) minimization

- Trying kinematic fitter to all hadronic events
  - Largest cross section
  - Difficult to reject backgrounds due to disadvantage if b-tagging
  - Is Kinematic fitter good tool for background rejection?

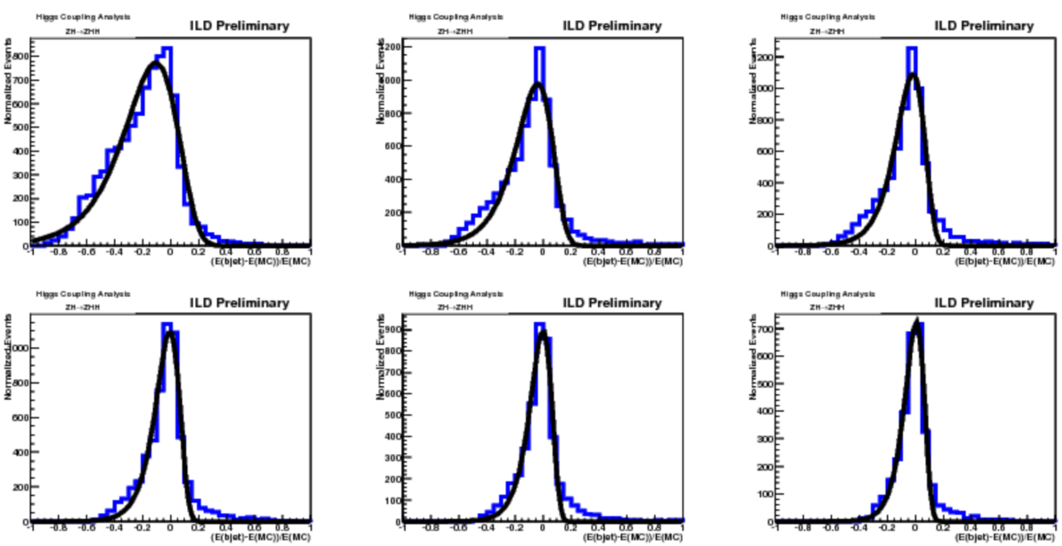
- ZHH→(bb)(bb)(WW*)→(bb)(bb)(jjjj) kinematic fitter
  - Constraints:

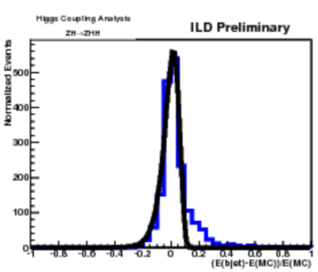
$$m(bb) = m_Z$$
$$Max(m1(jj), m2(jj)) = m_W$$
$$m(bb) = m(\text{jjjj})$$
$$E(H) + E(Z) + E(jj) + E(jj) = \sqrt{s}$$
$$\overrightarrow{p_H} + \overrightarrow{p_Z} + \overrightarrow{p_{jj}} + \overrightarrow{p_{jj}} = \vec{0}$$

  - No ISR effect is included…

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
  - e.g.) bjet energy resolution

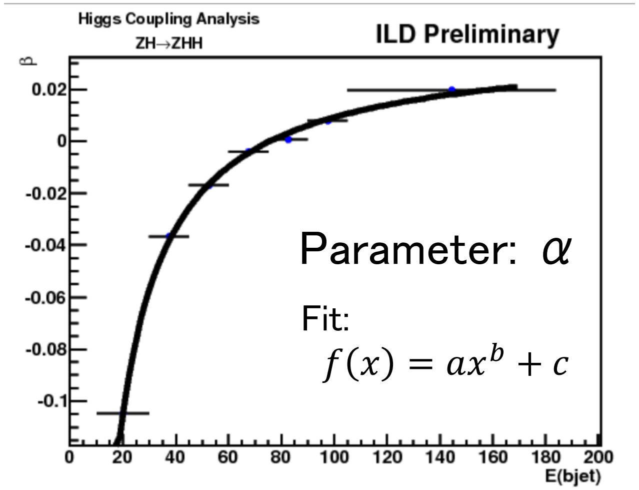


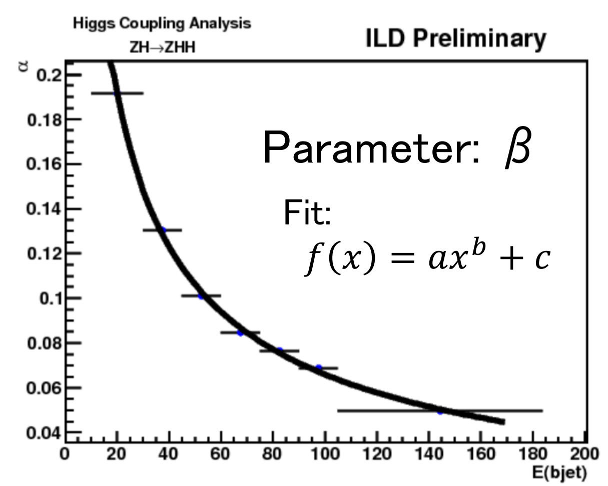


Fit: Gumbel dist.

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

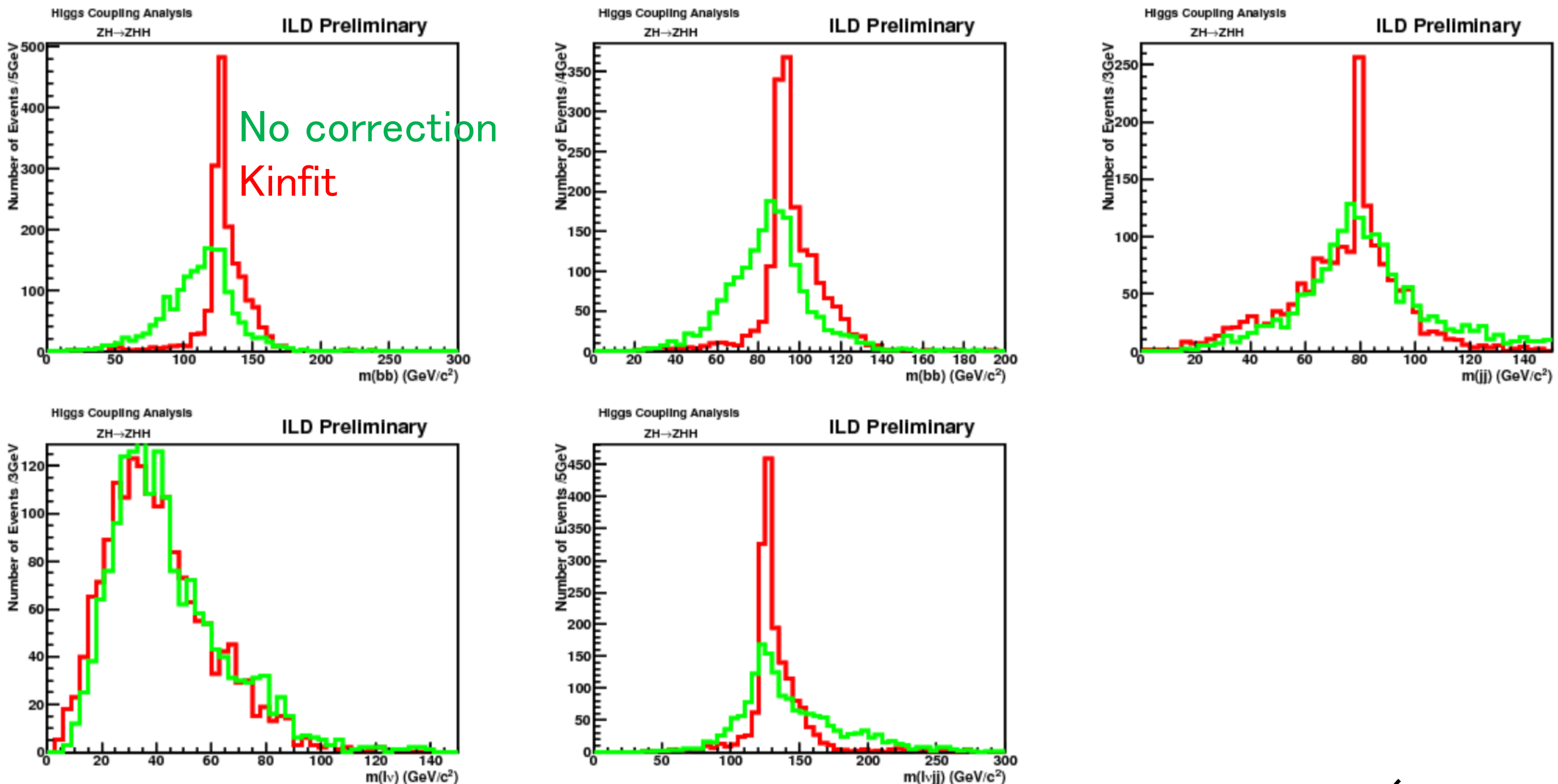
To include asymmetric energy resolution





Performance check

- Check each resonance distribution:
  - Mass resolution is going better! →promising



- Better mass resolution provides better MVA classifier(even if backgrounds come in signal mass region)··· →same in ILC?
  - c.f.) @CDF, 15% mass resolution improvement→10% improvement of sensitivity for Higgs search

PRELIMINARY RESULTS@500GEV

- Higgs-strahlung process, ZHH→(bb)(bb)(WW)
- 4 b-tagging or 3 b-tagging
- Calculate the significance, $\frac{s}{\sqrt{s+b}}$
- $L=4ab^{-1}$(H-20 scenario)

Category	b-tag	Signal	Background	Significance
8jets	4	28.44	174.86	2.00σ
8jets	3	18.57	925.69	0.60σ
Lepton+ 6jets	4	3.44	29.12	0.60σ
Lepton+ 6jets	3	2.08	7.17	0.68σ

16% improvement

5% improvement

- Kinematic fitter provides 5-16% improvement
  - 8jets, 4-btag doesn't have improvement so far…

  →under investigation
  - Need to Include ISR effect

Summary and Global summary

- Top-Yukawa study ongoing:
  - Likelihood method will provide ~10% improvement on jet matching efficiency
  - H-20 scenario: $S/\sqrt{S+B} = 6.61$, $|\Delta g_t/g_t| = 7.85\%$
  - Will more than twice better @550GeV
  - MVA will provide 10~20% improvement
- Higgs Self-Coupling analysis is ongoing.
  - Kinematic fitting will be a good tool for mass resolution improvement
  - Kinematic fitter provides 20% improvement for HH→(bb)(bb), 5%-16% improvement for some HH→(bb)(WW) modes
  - Include ISR effect and semi-leptonic B decay effect in Kinematic fitter will give more improvement
- **Global Summary:**
  - Basic analysis components need improvement
    - We already have had many improvements of analysis components in individual study level
    - Especially, flavor tagging will become better
  - Jet clustering is the last key to obtain better result

→better jet energy resolution gives us better kinematics in an event

  - Finally, incorporating all the improvements and updating the result are necessary!

BACKUPS

Summary for HH→(bb)(bb)

- Kinematic fitter provides better background rejection:
  - Improvement is ～20% using fitting $\chi 2$
  - b/c jet require special ISR treatment
  - Optimize the strategy for good mass resolution
- $\nu\nu$HH and qqHH studies are ongoing
  - Same as llHH
  - Kinematic fit is challenging
- Semi-leptonic b/c decays:
  - Correct missing energy from semi-leptonic decay in b/c jets
  - Results are promising – better JER!
  - Particle will help for this correction!
  - Investigation ongoing
  - Treatment of semi-leptonic b/c decays in $\nu\nu$HH is interesting

Kinematic Fitting - ISR Treatment

identify events using ISR characteristics

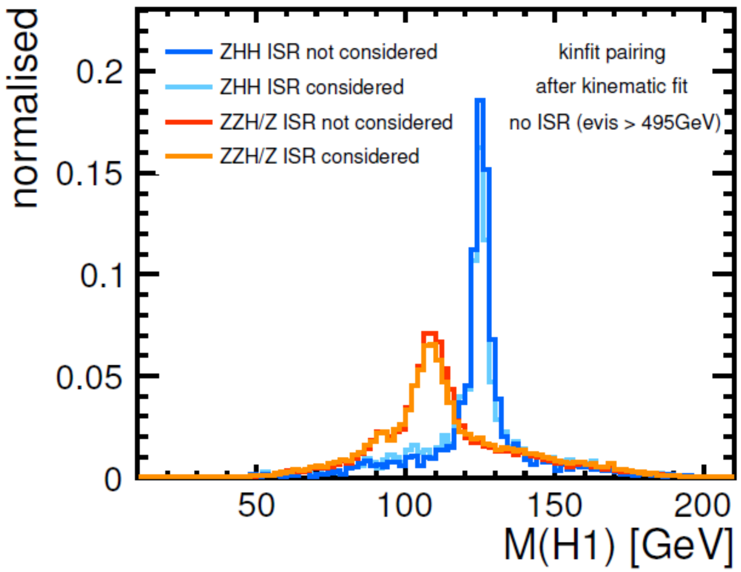
→ significant missing energy

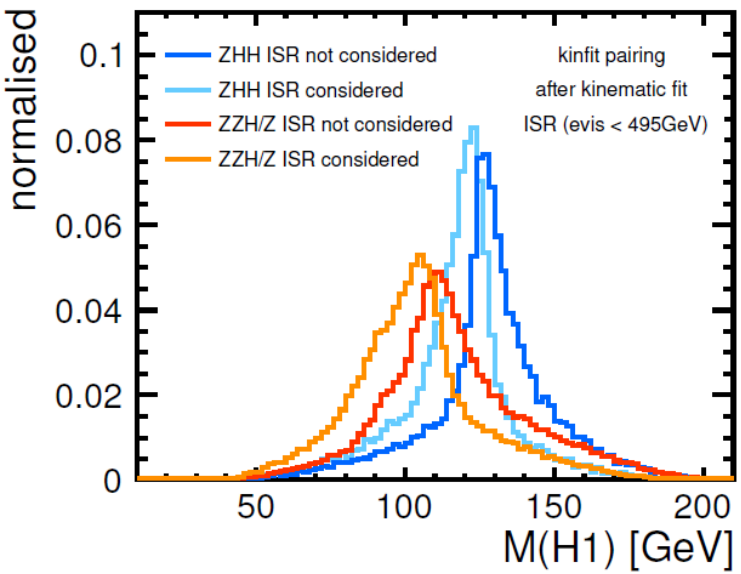
→ large $|\cos(\theta_{miss})|$

- **events w/o significant E_{miss}**
  - → independent of ISR consideration in fit
  - → independent of $|\cos(\theta_{miss})|$
- **events w/ significant E_{miss}**
  - → dependent on ISR consideration in fit
  - → dependent on $|\cos(\theta_{miss})|$

E_{miss} not enough to categorise events into with and without ISR

→ information on $|\cos(\theta_{miss})|$ needed





Kinematic Fitting - ISR Treatment

identify events using ISR characteristics

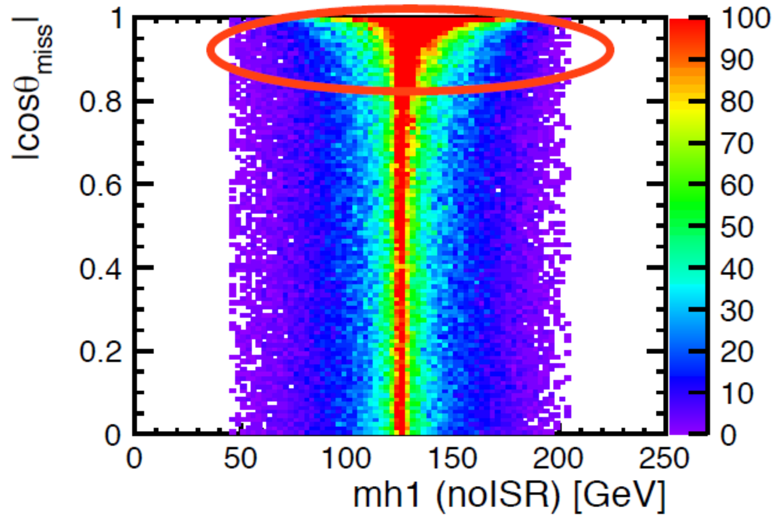
→ significant missing energy

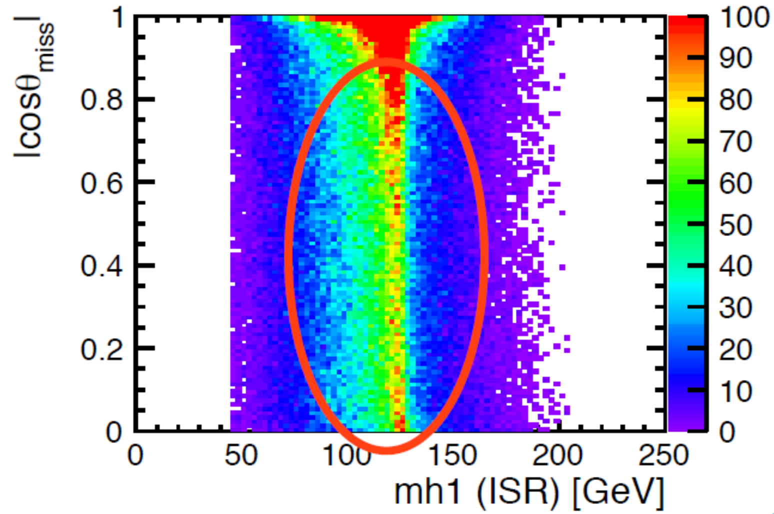
→ large $|\cos(\theta_{miss})|$

- **ISR not considered in fit**
  - → **small** $|\cos(\theta_{miss})|$: narrow around 125 GeV
  - → **large** $|\cos(\theta_{miss})|$: bias to large M(H1)

- **ISR considered in fit**
  - → **small** $|\cos(\theta_{miss})|$: bias to small M(H1)
  - → **large** $|\cos(\theta_{miss})|$: narrow around 125 GeV

events need to be selected by their amount of missing energy and $|\cos(\theta_{miss})|$ values





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

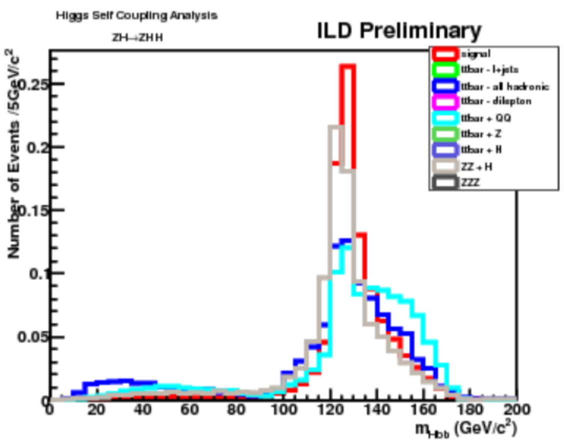
process	σ (fb)	σ (fb)
Signal(inclusive)	0.2	0.3
ttbar & ZWW	581.8	264.9
ttbar + QQ	0.83	5.74
ttbar + Z	0.98	7.81
ttbar + H	0.14	3.22
Z, $\nu\ \nu$ + ZH	0.77	2.70
Z, $\nu\ \nu$ + ZZ	1.83	14.01

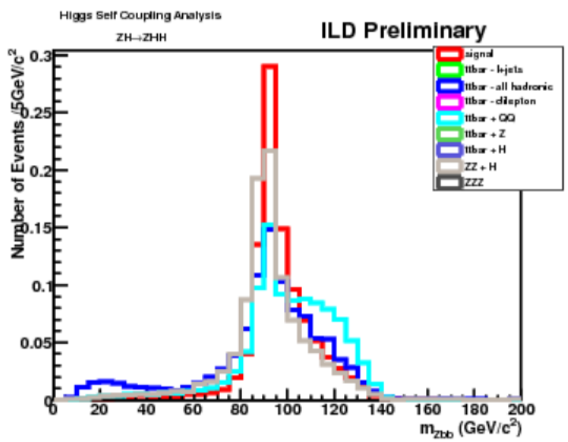
Simulation

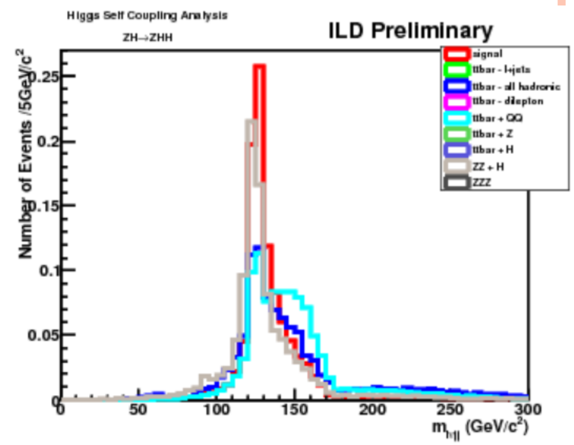
	500GeV	1TeV
Polarization (e,p)	(-0.8,+0.3)	(-0.8,+0.2)
$m_H(GeV/c^2)$	125	125
simulator	Full(DBD)	Full(DBD)

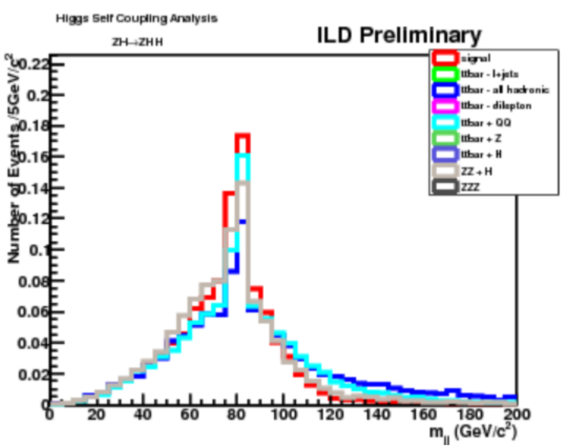
Comparison between signal and backgrounds

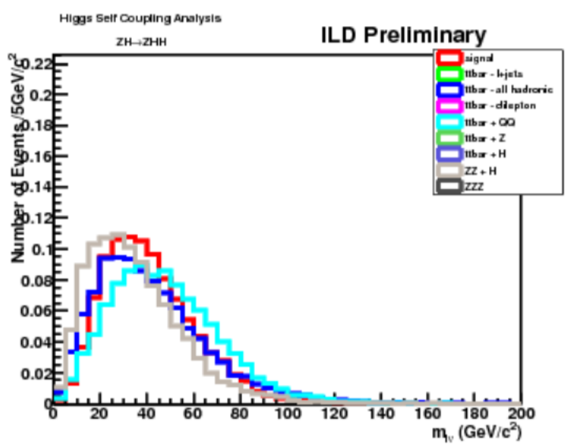
- Higgs mass(H→bb) & Z mass distribution
  - How are mass distributions of backgrounds?
  - ZZH background is hard to reject?
  - Top related backgrounds will be separated well











Signal
ttbar allhadronic
ttbar+QQ
ZZ+H

B-Tagging for HH→(bb)(WW)

- Trying to gain the total acceptance
  - Make b-tagging loosen and categorize using b-tagging condition

- ZHH→(bb)(bb)(WW): maximum number of b-tagging available is **4**

- Making 3-btag category:
  - 4 b-tag caterory: 4 jets with b-likeliness>0.4(0.35) in an event
  - 3 b-tag category: 3 jets with b-likeliness>0.4(0.35) in an event

- We can apply this categorization to:
  - ZHH→(bb)(bb)(l ν jj): Lepton+6jetys
  - ZHH→(bb)(bb)(jjjj): 8jets

Comparison with old results

- Compare with old results
- Calculate the significance, $\frac{s}{\sqrt{s+b}}$
- For comparison, L=2ab^{-1}

Category	b-tag	Old result	New result
8jets	4	1.50 σ (LCWS13)	1.41 σ (Now)
8jets	3	0.35 σ (LCWS13)	0.41 σ (Now)
Lepton+ 6jets	4	0.41 σ (LCWS14)	0.43 σ (ALCW15)
Lepton+ 6jets	3	N/A	0.48 σ (Now)

← Due to leptonID
← 16% improvement
← 5% improvement
← New!

- Why does 8jets 4 b-tag result become worse?
  - Due to lepton ID improvement?

  →more ttbar all hadronic events move into this category

  →become more difficult to reject ttbar backgrounds…

  Under investigation to recover the significance

- Kinematic Fitter provides 5–16% improvement for those modes

ongoing: investigation of how to see colour correlation

Seeing in Color: Jet Superstructure

arXiv:1001.5027

Jason Gallicchio and Matthew D. Schwartz
Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

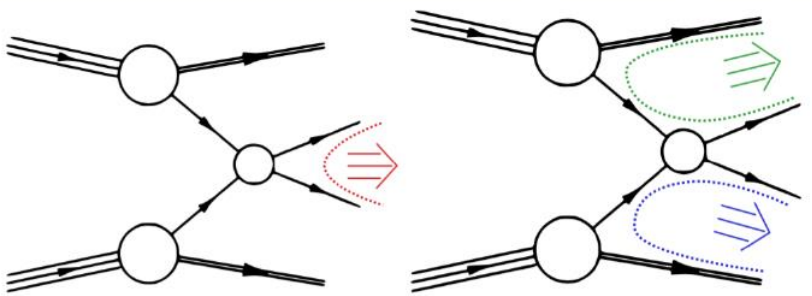
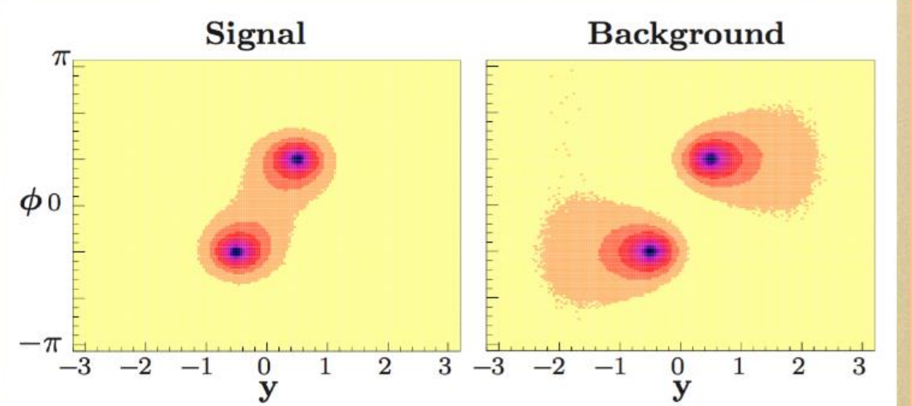


FIG. 1: Possible color connections for signal ($pp \to H \to b\bar{b}$) and for background ($pp \to g \to b\bar{b}$).



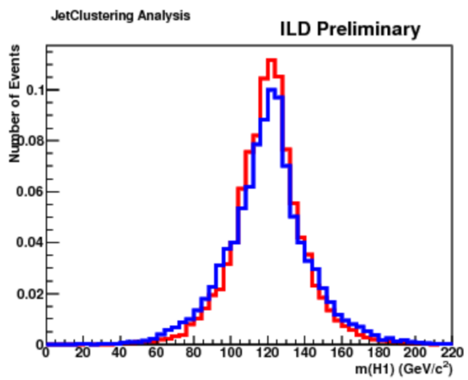
introduce a new variable, *jet pull*

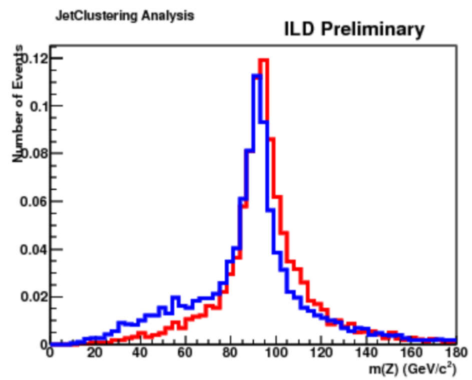
$$\vec{t} = \sum_{i \in \text{jet}} \frac{p_T^i \, |r_i|}{p_T^{\text{jet}}} \, \vec{r}_i \, .$$

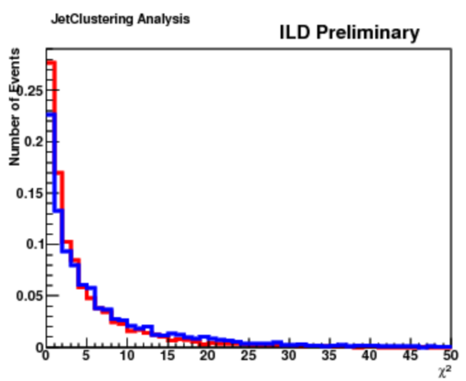
i stands for constituents of jet; ri is vector from jet axis to particle i in (rapidity, phi) plane

Realistic case for better jet clustering study

- Trying realistic case for analysis qqHH→qq(bb)(bb)
  - 6 jet clustering
  - Btag>0.35 for 4 jets, no jet energy constraint
  - Mass constraint of Higgs and Z is imposed(χ 2)
  - No MCTruth information is used







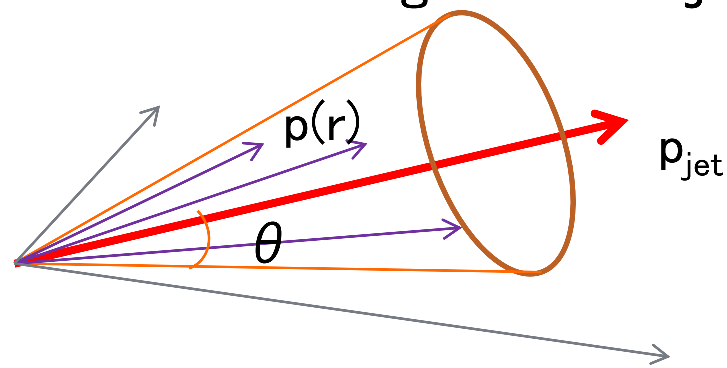
Durham
Self-Organized Mapping

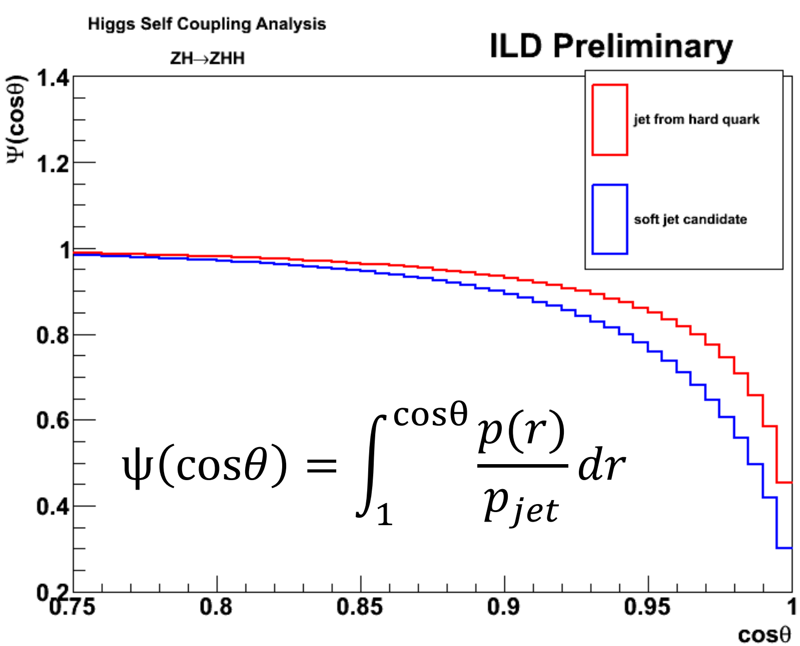
- Num. of events
  - b-tagging becomes better thanks to clustering?

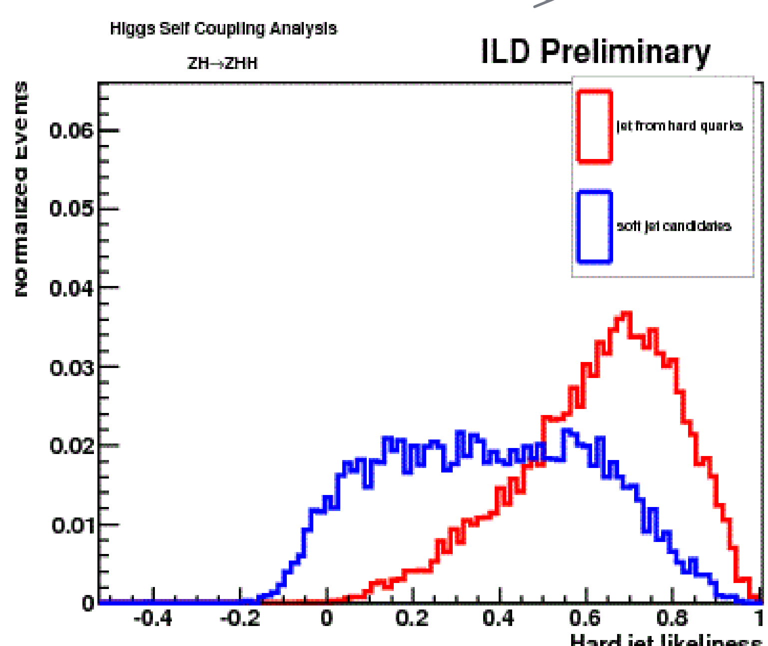
Method	Durham	SOM
Num. of events	5743	6075

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



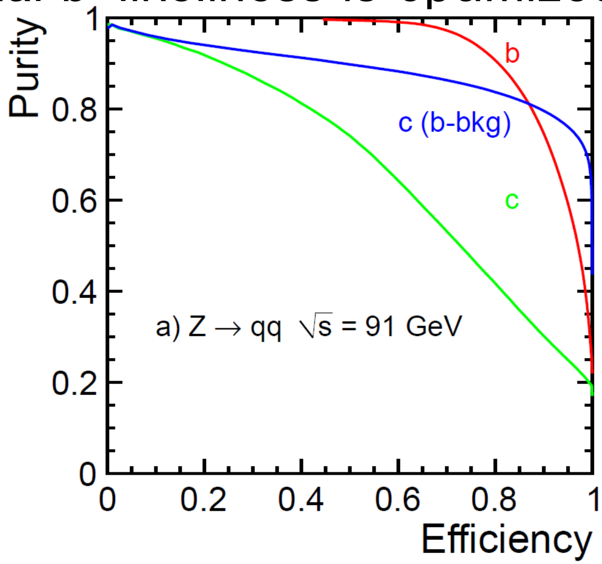


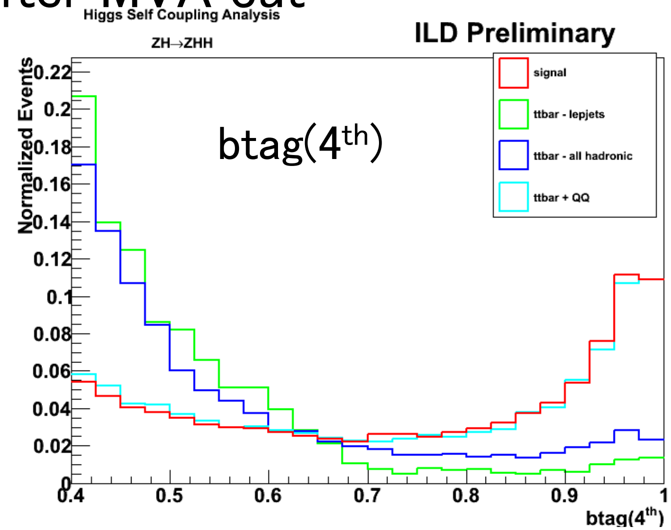


Flavor Tagging

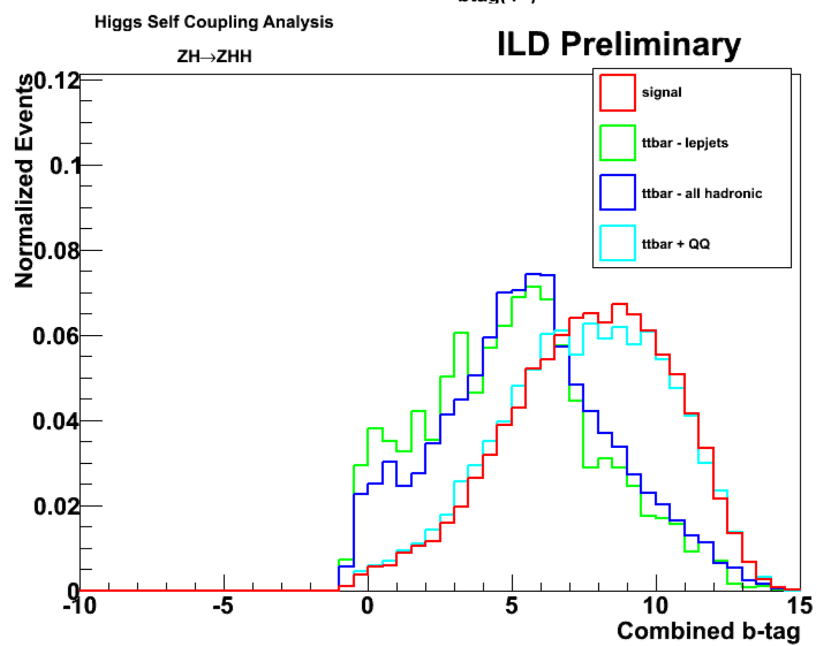
Using LCFIPlus

- 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 → **leptonic** → **ttbar hadronic**

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

ZZ+X MVA → **ttbar+X** → **ttbar**

ZZ+X MVA
- ZZH
- ZZZ

qq

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

ttbar
- ttbar
- ZWW