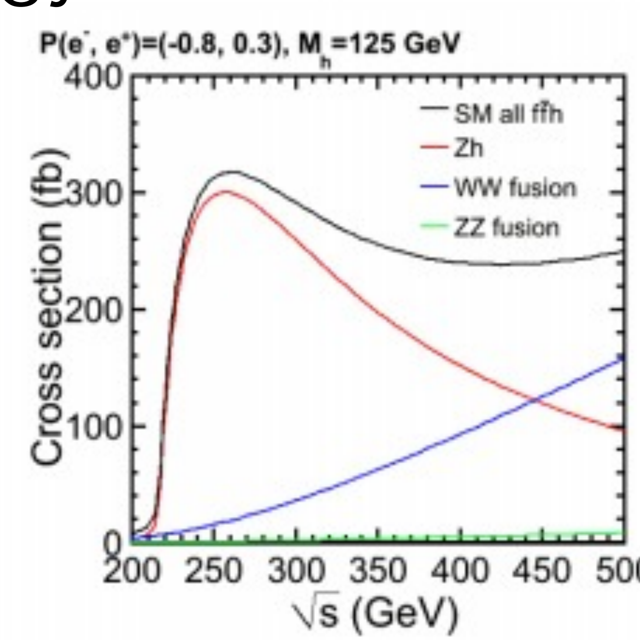
Hadronic recoil mass study using ZH -> qqH (ILC @ 250 GeV)

Tatsuhiko Tomita
Kyushu University
Taikan Suehara (Kyushu Univ.)

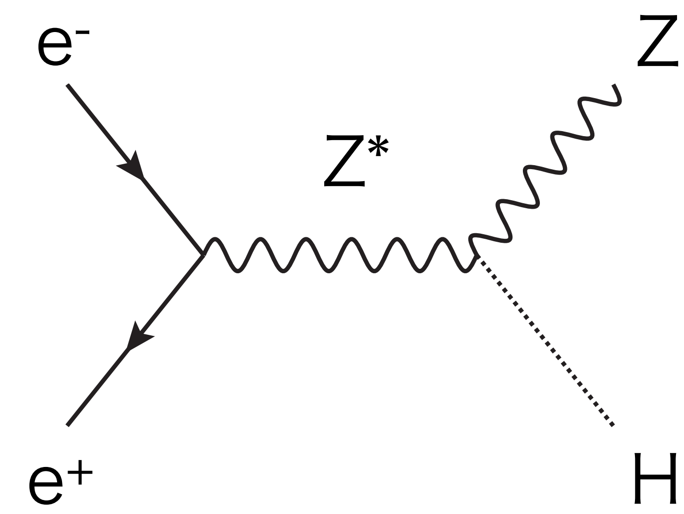
九州大学
KYUSHU UNIVERSITY

Motivation

- One of the most important measurements at 250 GeV ILC is measurement of coupling constant with Higgs boson.
- In ILC, g^2_{ZZH} can be measured by σ_{ZH} directly.
- Using this g^2_{ZZH}, we can calculate some other couplings.
- So, it is important to measure σ_{tot} of ZH production with high precision !
- 250 GeV is one of the most suitable energy to measure σ_{tot} of ZH production.

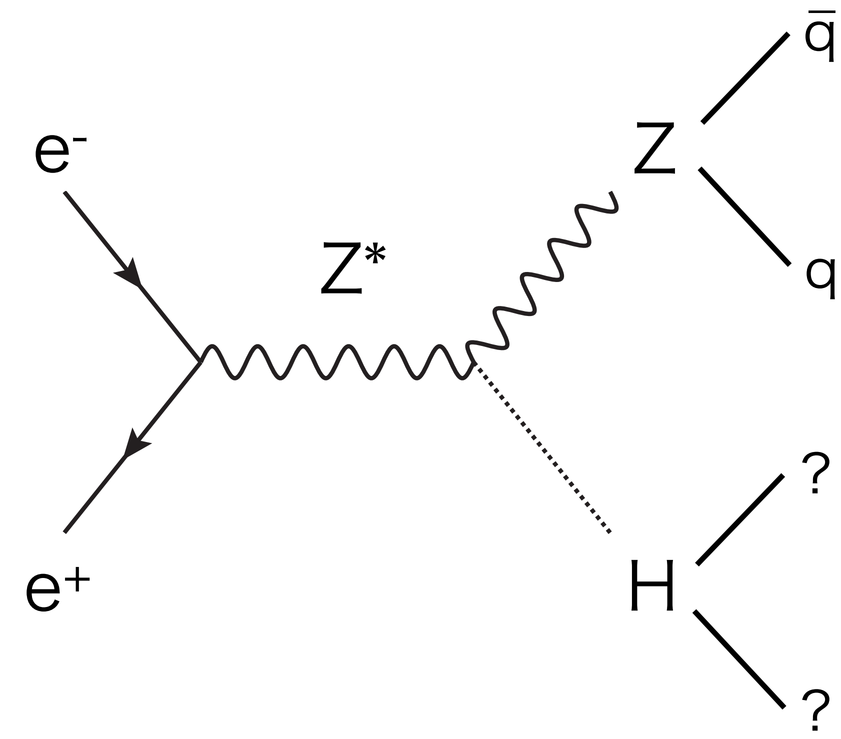


ZH production



- Major Higgs production process at 250 GeV.
- Z boson will decay to …
  - charged leptons (e, μ, τ) - total ~10 %
  - neutrinos - total ~20 %
  - hadrons - **total ~70 %**
- Leptonic decay channel is useful for mass measurement.
  → High precision of mass measurement ~30 MeV.
  σ_{tot} measurement is also good ($\delta\sigma/\sigma$ ~2.6 %).
  But, statistics is limited. (only ~3.4% each lepton generation.)
- Hadronic decay channel has **large statistics**.
  → σ_{tot} measurement is promising in hadronic channel.
  The problems are model dependency and large background.

Recoil method in hadronic channel



- Using 4-momentum conservation,

$$m^2_{\mathrm{recoil}} = (\sqrt{s} - E_Z)^2 - \mid \vec{p}_Z \mid^2$$

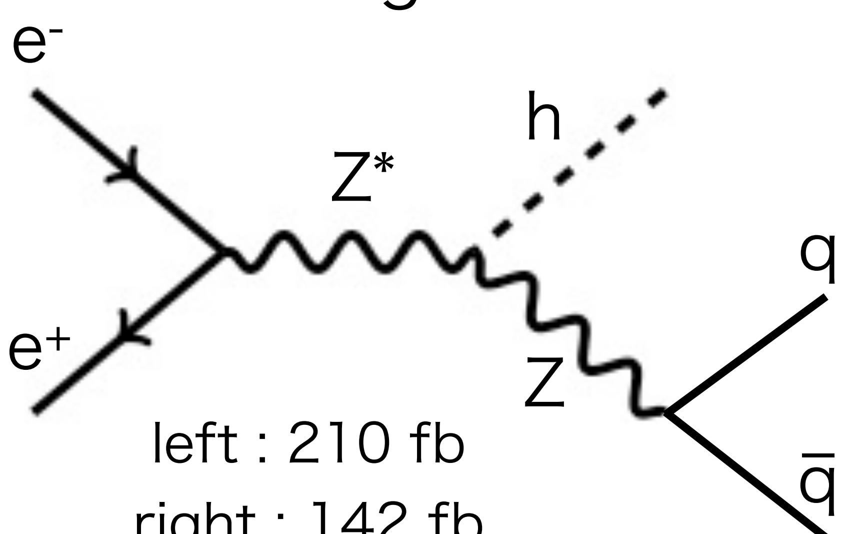
(recoil method)

- Lepton channel has good S/N.
- Hadron channel has worse S/N, but **large statistics**.
  → Uncertainty of jet energy, clustering, etc …
- Is it possible to use hadronic channel of ZH production ?

Data samples

Higgs mass	E_{CM}	Luminosity	Polarization	Detector
125 GeV	250 GeV	250 fb^{-1}	left: (-0.8, +0.3) right:(+0.8, -0.3)	ILD_DBD ver.

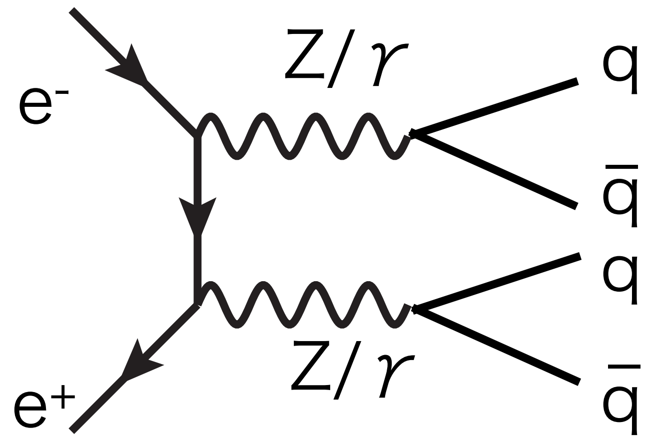
signal



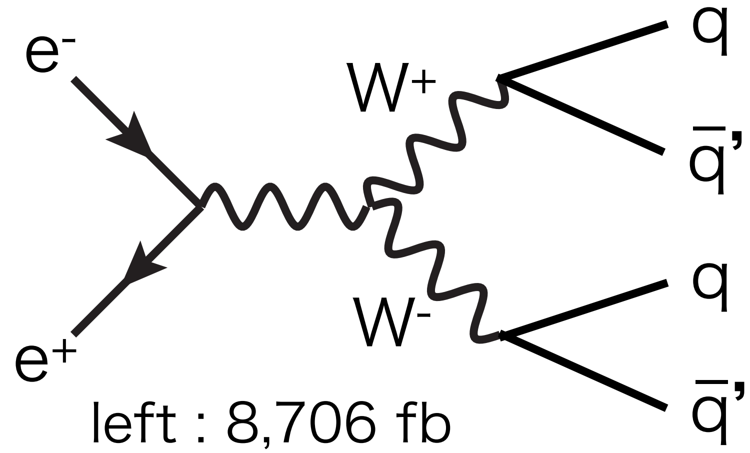
left : 210 fb
right : 142 fb

semi-leptonic events
are also considerable BG.

Major background

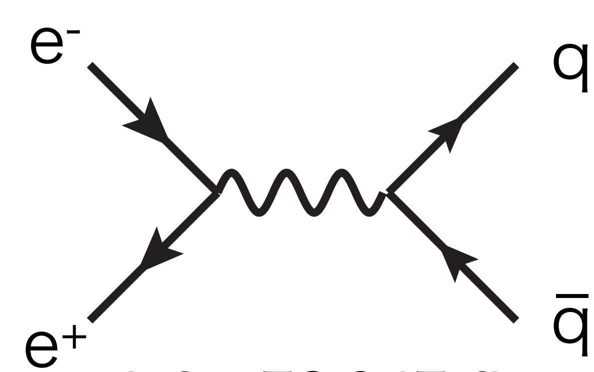


left : 841 fb
right : 403 fb



left : 8,706 fb
right : 600 fb

4 fermion



left : 78,047 fb
right : 46,214 fb

2 fermion

Analysis flow

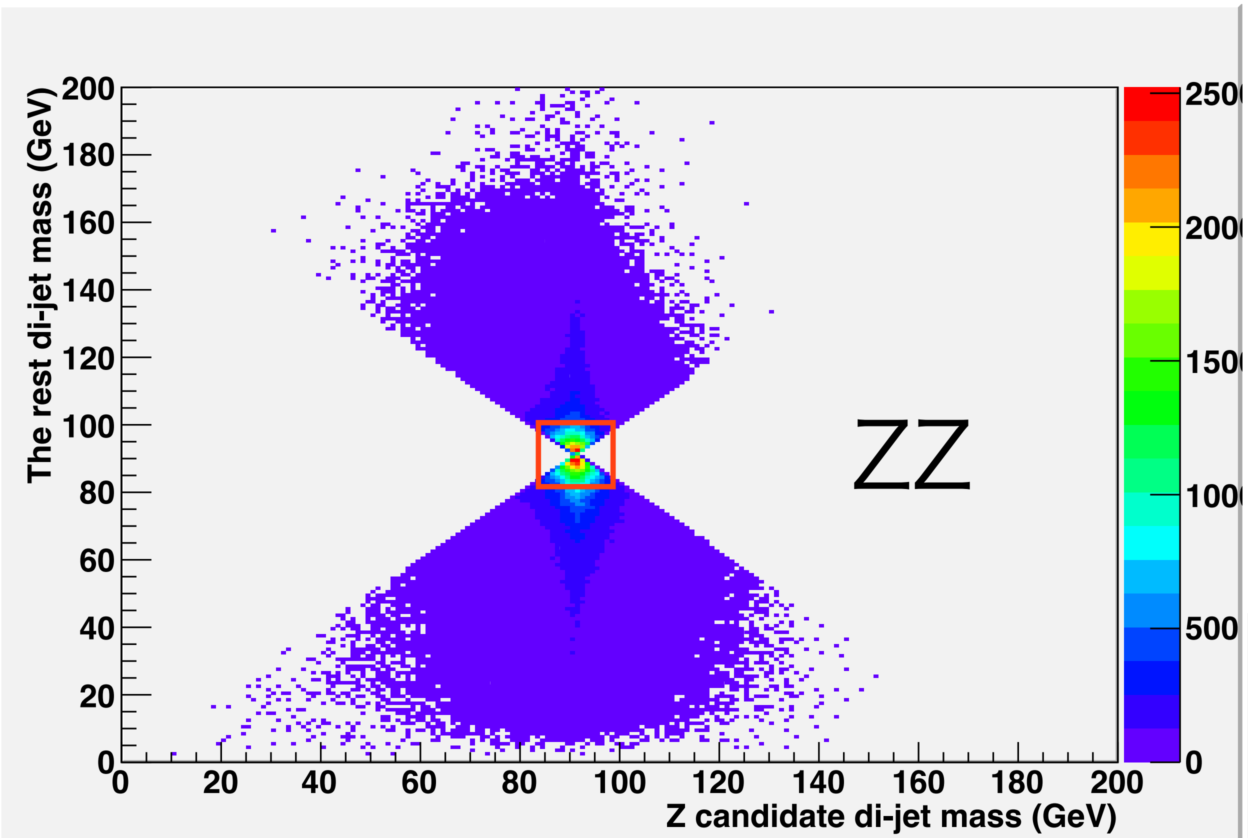
- To improve jet clustering,
  - Initial state radiation
  - Isolated lepton
  - Hadronic tau jet were removed from events.
- Durham jet clustering was applied to the remaining particles.

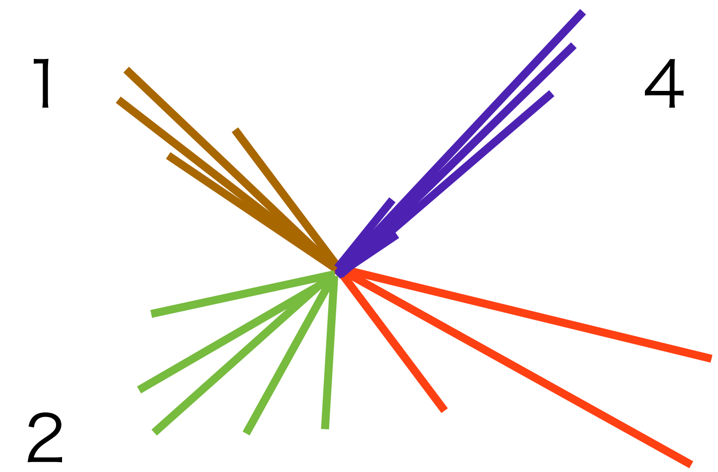
$$y = \frac{2\mathrm{min}(E_i^2, E_j^2)(1 - \cos\theta_{ij})}{Q^2}$$

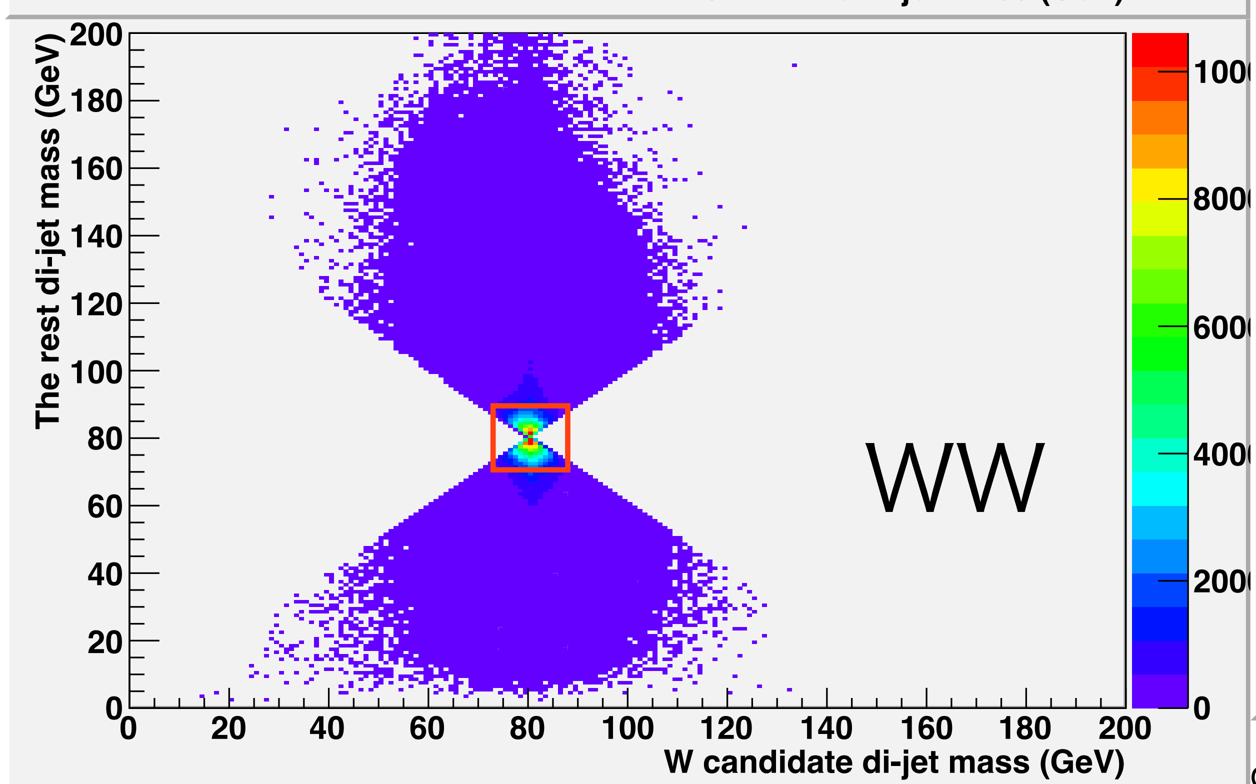
- Forced 4 jet clustering, y threshold clustering were used. (y = 0.0025)

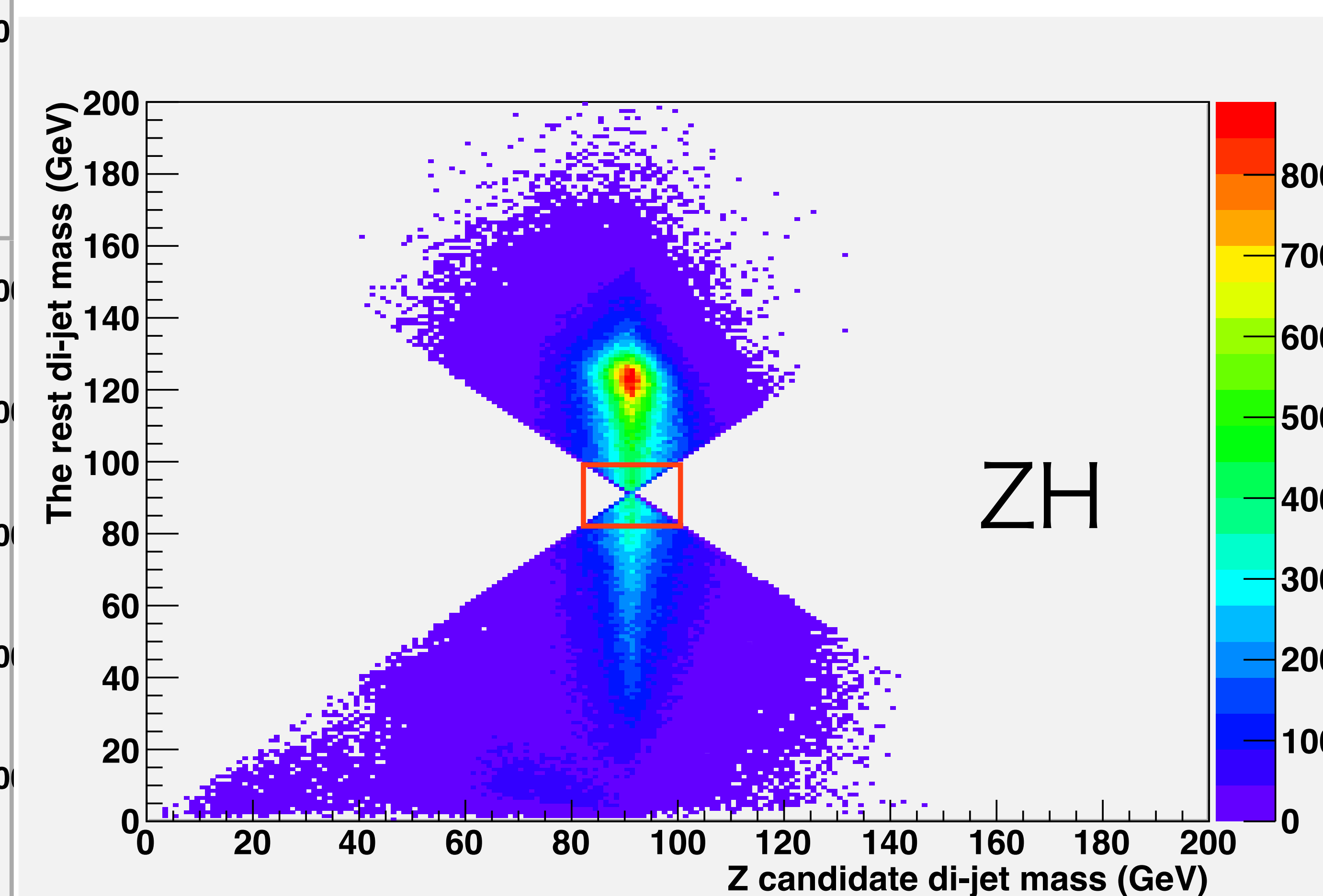
Background estimation - 1

- For 4 fermion backgrounds : forced 4 jet clustering



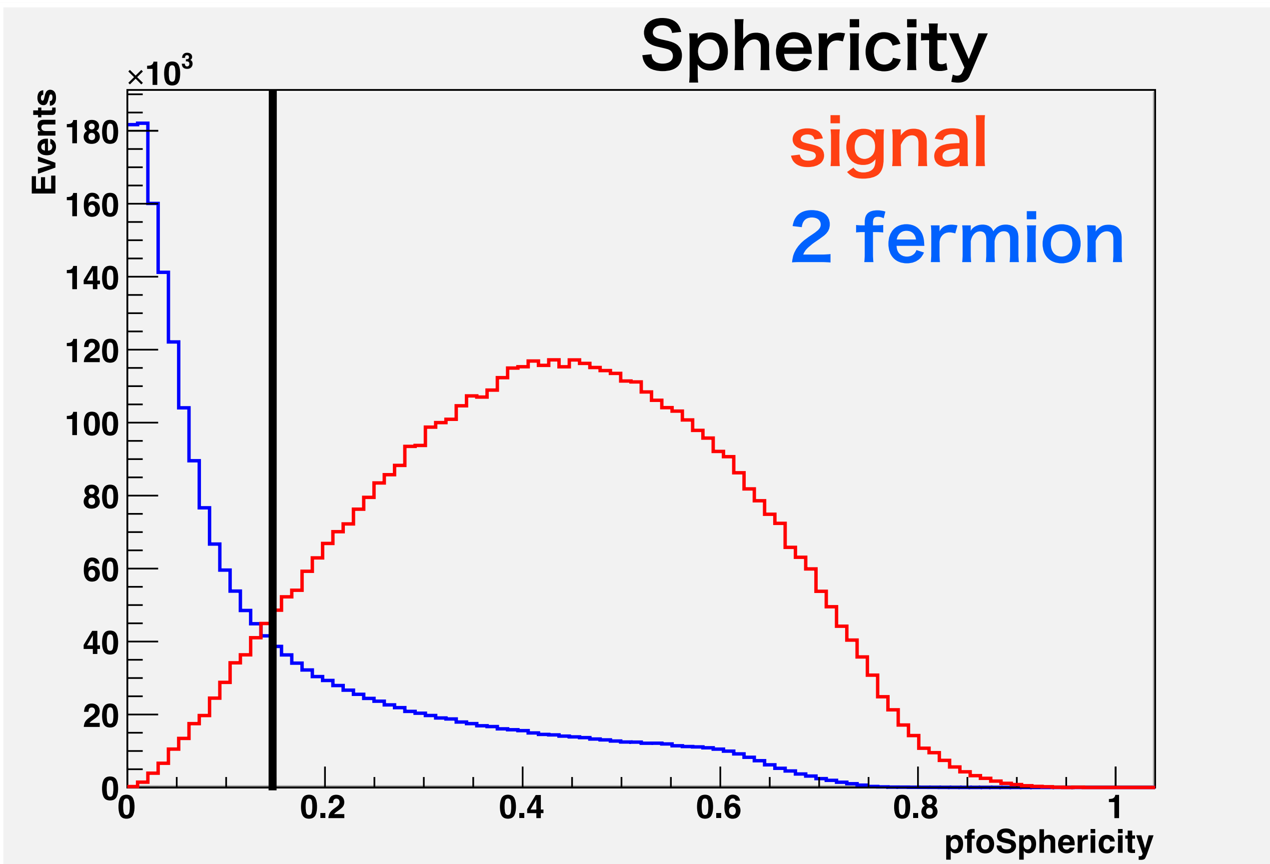


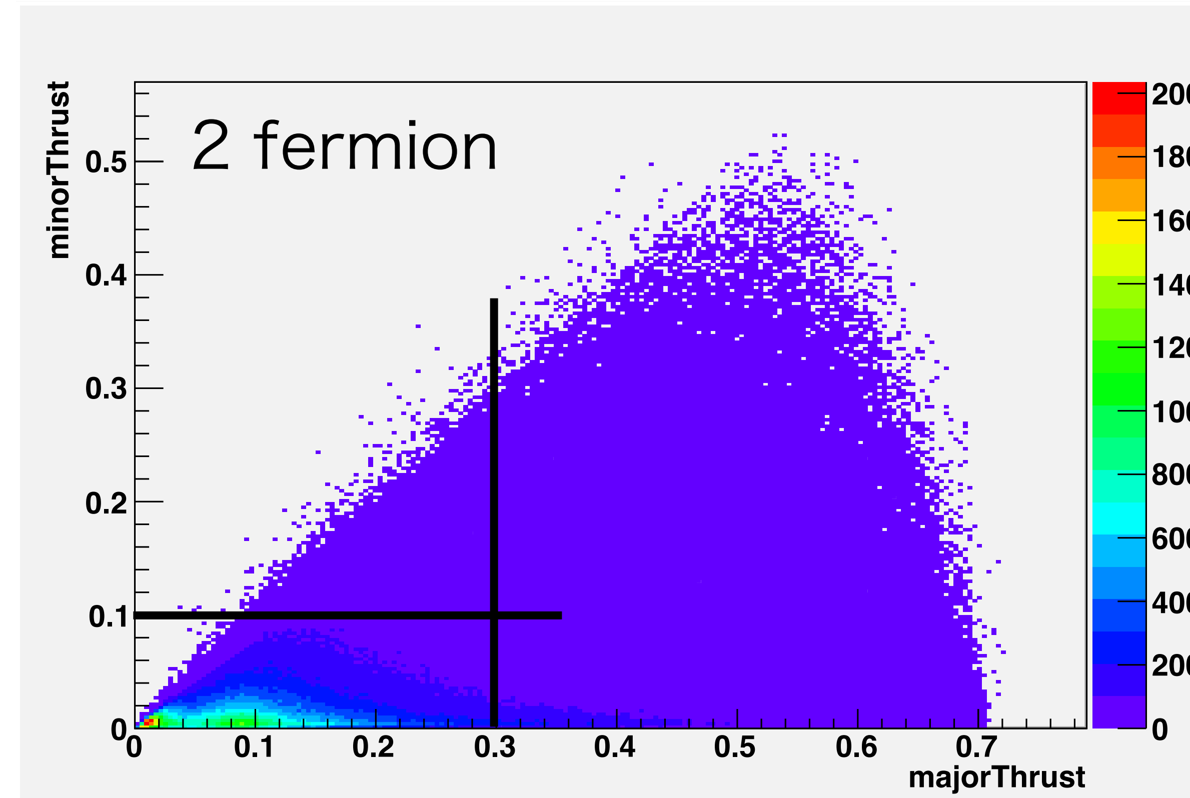


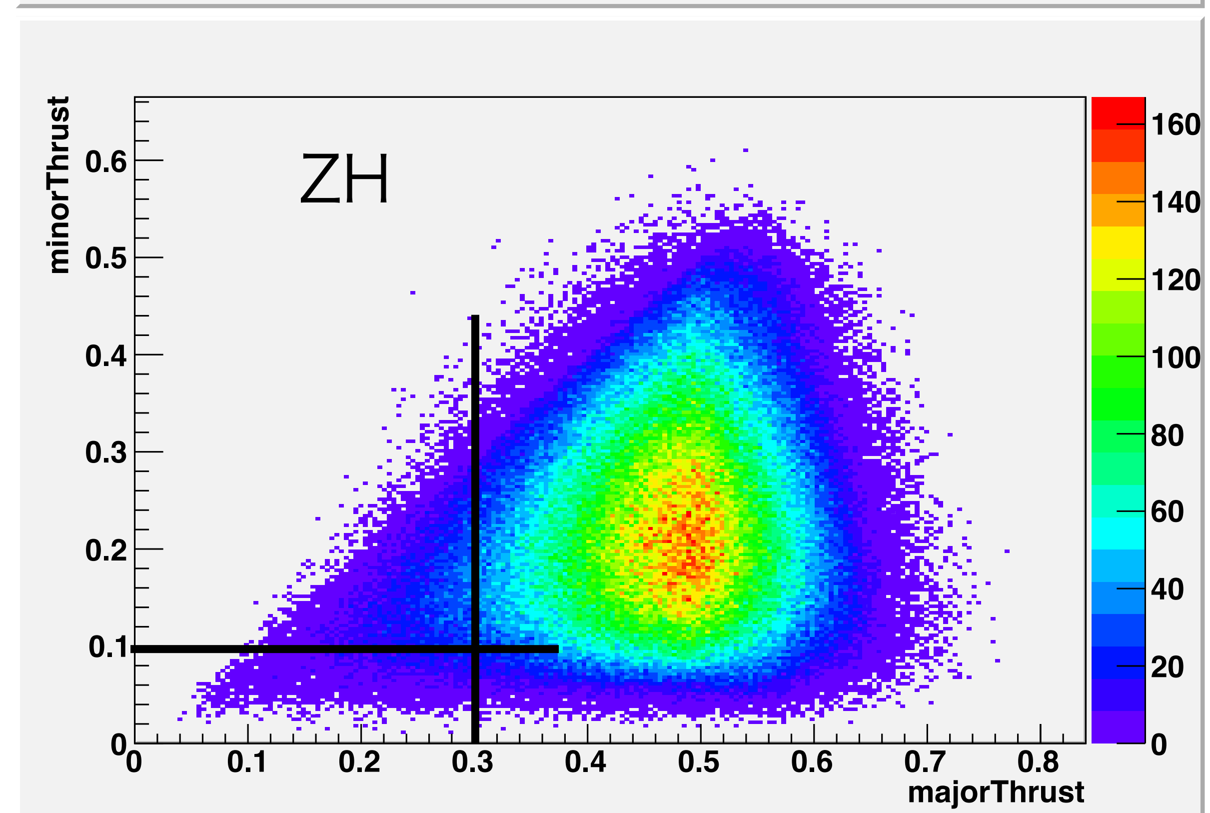


Background estimation - 2

- For 2 fermion background : Sphericity, Thrust (major, minor)



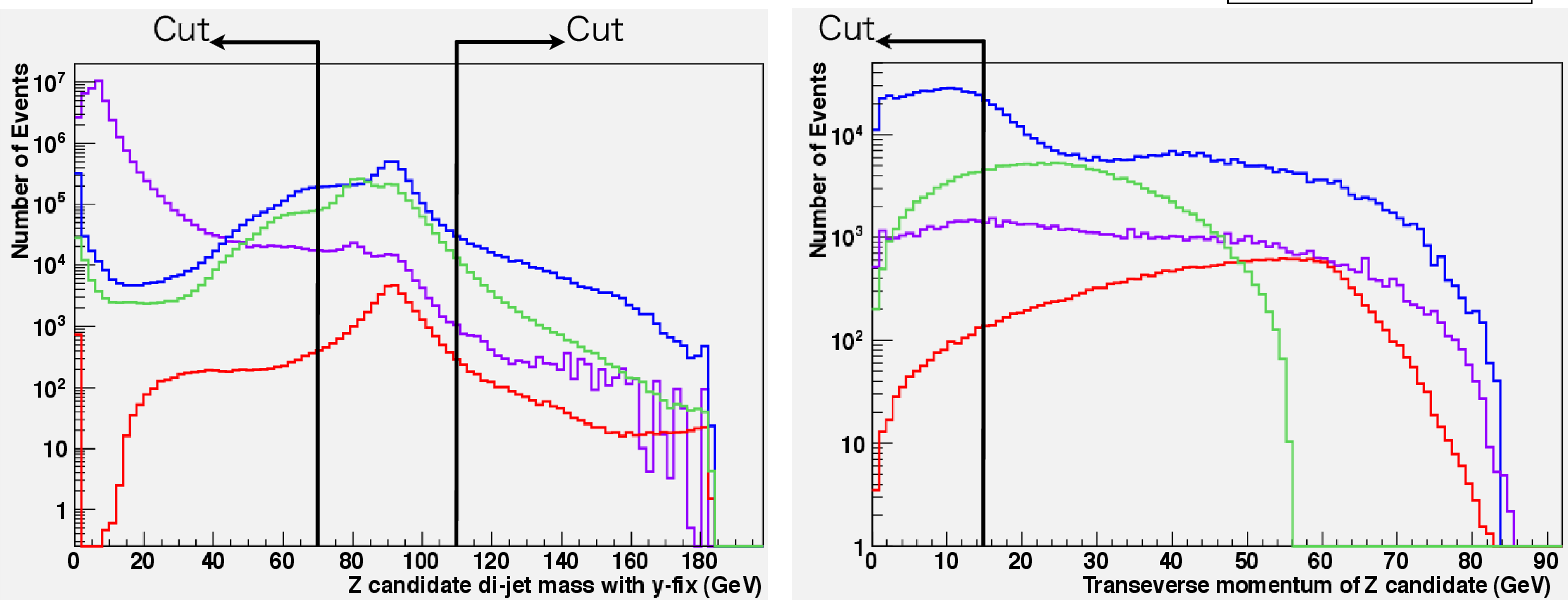




Background estimation - 3

- by y threshold clustering.
- Reconstructed Z mass and Z p_T were used.

signal
2 fermion
4 fermion
others

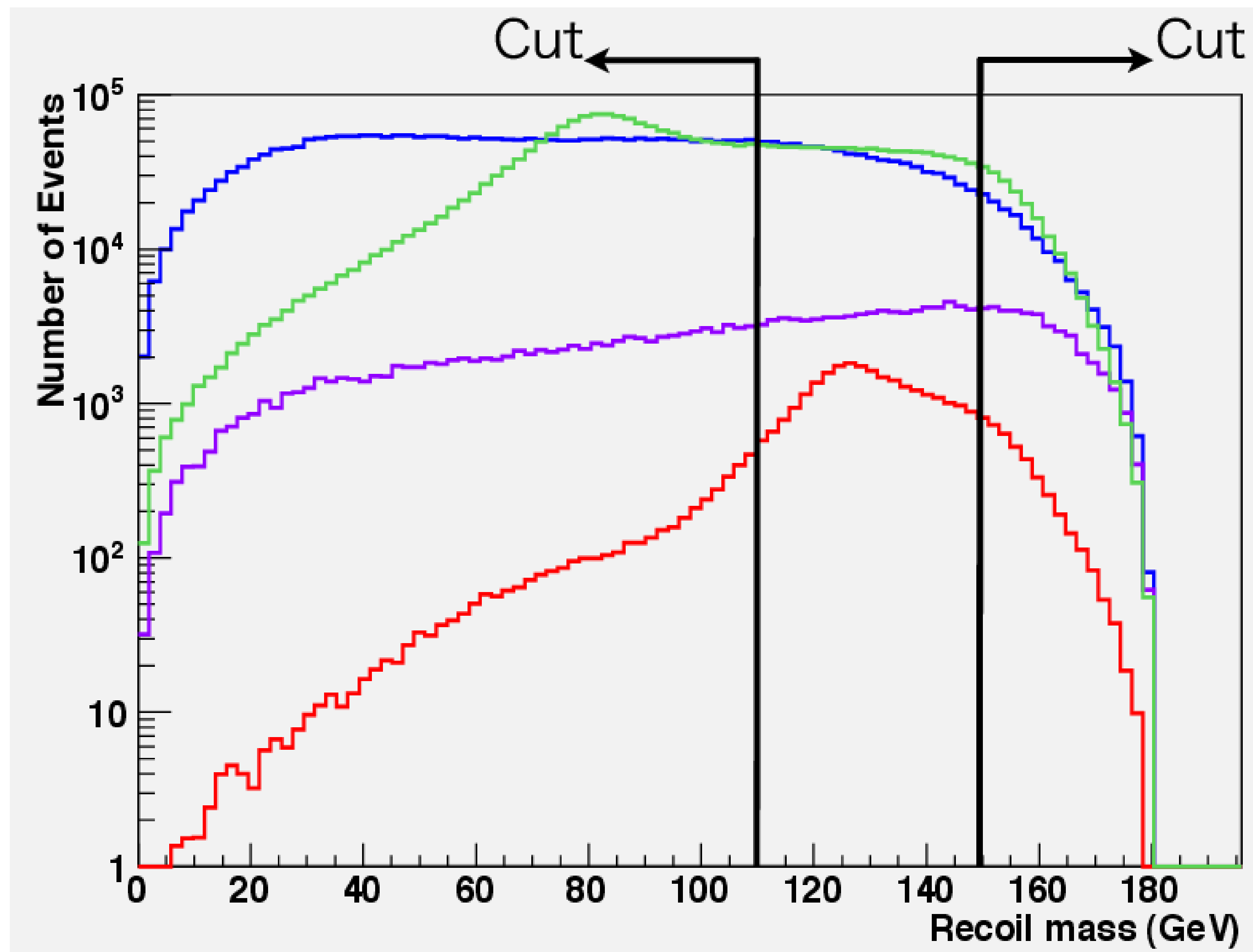


Background estimation - 4

- Recoil mass distribution cut.

signal
2 fermion
4 fermion
others

- Recoil mass was calculated by using 4 momentum of Z_{Rec} with y threshold clustering.



Cuts table

cuts	signal	4 fermion	2 fermion	others
left	50,816	9,361,676	19,315,415	216,171,025
right	34,308	1,084,045	12,556,240	222,597,419
4 fermion cut	82.8%	61.8%	97.9%	99.9%
2 fermion cut	78.5%	39.5%	33.6%	25.8%
Z mass Z p_T Recoil	46.2%	8.3%	1.4%	0.3%

Statistical precision after cuts

- After applying cuts, significance and stat. precision were calculated.

polarization	significance	stat. precision
left (-0.8, +0.3)	20.9 σ	4.8%
right (+0.8, -0.3)	28.0 σ	3.6%

- Stat. precision is about 1% worse than lepton channel (2.6%).

- Right handed polarization is better than left handed one.
  → W background can be suppressed with right polarization.

Cut efficiency for each Higgs decay branch

- In order not to depend on decay mode of Higgs boson, the events should be survived equally after cuts.

mode	After cuts (%)	diff./mean
H->all	46.2%	-----
H->bb (57.7%)	43.3%	-6.3%
H->WW(leptonic) (2.3%)	45.3%	-2.0%
H->WW(semi-leptonic) (9.5%)	46.9%	+1.4%
H->WW(hadronic) (9.8%)	54.4%	+17.7%
H->gg (8.6%)	55.2%	+19.5%
H->$\tau\tau$ (6.3%)	45.3%	-2.1%
H->ZZ (2.6%)	48.6%	+5.1%
H->cc (2.9%)	47.1%	+1.8%
H->$\gamma\gamma$ (0.2%)	43.8%	-5.2%

- WW(hadronic), gg have quite large inconsistency from mean…

Strategy to reduce inconsistency (Categorization)

- Categorization is a powerful tool to reduce difference of efficiency among Higgs decay modes.
  - Categorize events using number of jets, leptons, taus, etc.
  - Minimize the difference of efficiency in each category (decay modes with too small fraction in the category is negligible.)
  - Calculate partial cross section from each category
  - Combine all cross section from categories to get the total cross section of ZH production.

Categorization - 1

To resolve efficiency inconsistent issue, we will categorize events using

- the number of tau jets (0, 1, and >= 2)
- the number of isolated lepton (0, 1, and >= 2)

$$\mathrm{N}^i = \sum_n \sigma_{\mathrm{tot}} \cdot \mathrm{BR}_n \cdot \theta_n^i \cdot \epsilon_n^i$$

$$\mathrm{n} = (\mathrm{b}, \mathrm{W}, \mathrm{g}, \tau, ...)$$

N^i is a number of events in category i , σ_{tot} is total cross section, BR_n is Higgs decay branching ratio, θ_n^i is fraction in category i , ϵ_n^i is cut efficiency for category i .

If the cut efficiency of each decay mode can be assumed to be the same as ϵ^i($=\epsilon_n^i$).

$$\frac{\mathrm{N}^i}{\epsilon^i} = \sigma_{\mathrm{tot}} \sum_n \mathrm{BR}_n \cdot \theta_n^i$$

Then we can get

$$\sum_i \frac{\mathrm{N}^i}{\epsilon^i} = \sigma_{\mathrm{tot}} \sum_n \sum_i \mathrm{BR}_n \cdot \theta_n^i = \sigma_{\mathrm{tot}}$$

Categorization - 2

If the cut efficiency is not exactly the same,
we should consider the systematic effect caused by the difference.

$$\delta\epsilon_n^i = \epsilon_n^i - \epsilon^i$$

And the cross section is

$$\sigma_{\mathrm{tot}} = \frac{\sum_i \frac{\mathrm{N}^i}{\epsilon^i}}{1 + \sum_n \sum_i \mathrm{BR}_n \cdot \theta_n^i \cdot \frac{\delta\epsilon_n^i}{\epsilon^i}}$$

We want to keep systematic uncertainty is less than 1 % to do model independent analysis.

If we don't assume any models, we should keep $\theta_n^i \cdot \frac{\delta\epsilon_n^i}{\epsilon^i}$ << 1 %.

If we can assume SM like Higgs, we should keep $\mathrm{BR}_n \cdot \theta_n^i \cdot \frac{\delta\epsilon_n^i}{\epsilon^i}$ << 1 %.

Cut efficiency after categorization

mode	After cuts (%)	Before optimization	After optimization (square sum)
H->all	46.2%	-----	-----
H->bb (57.7%)	43.3%	-6.3%	±0.6%
H->WW(leptonic) (2.3%)	45.3%	-2.0%	±1.6%
H->WW(semi lep) (9.5%)	46.9%	+1.4%	±3.0%
H->WW(hadronic) (9.8%)	54.4%	+17.7%	±1.5%
H->gg (8.6%)	55.2%	+19.5%	±5.5%
H->$\tau\tau$ (6.3%)	45.3%	-2.1%	±1.9%
H->ZZ (2.6%)	48.6%	+5.1%	±1.5%
H->cc (2.9%)	47.1%	+1.8%	±4.0%
H->$\gamma\gamma$ (0.2%)	43.8%	-5.2%	±4.2%

- After optimization, diff./mean is at most 5.5 %.
- Need check the impact of this inconsistency.

Statistical precision after categorization

polarization	significance	stat. precision
left (-0.8, +0.3)	40.3 σ	2.5%
right (+0.8, -0.3)	44.6 σ	2.2%

- After reducing cut difference from mean value, the stat. precision recalculated with categories.
- In this case, 2.2 % stat. precision with right polarization.

Systematic uncertainty

- The uncertainty of the Higgs branching ratio was studied.
- Changed each Higgs branching ratio 3 times of the measurement accuracy in ILC 250 GeV.
  (ex. White Paper : H->bb 1.2 % <=> This study : H->bb 3.6 %)

$$\sigma_{tot} = \frac{N_{\rm eve}}{\mathcal{L} \times \epsilon}$$ (keep total N_{eve})

$b\bar{b}$ +3.6%	210.23 fb	141.57 fb	~0.0%	−0.0%
$b\bar{b}$ −3.6%	210.10 fb	141.61 fb	~0.0%	+0.0%
$c\bar{c}$ +24.9%	210.11 fb	141.55 fb	~0.0%	~0.0%
$c\bar{c}$ −24.9%	210.22 fb	141.63 fb	~0.0%	~0.0%
gg +21.0%	210.07 fb	141.59 fb	~0.0%	~0.0%
gg −21.0%	210.26 fb	141.59 fb	~0.0%	~0.0%
WW +19.2%	210.07 fb	141.63 fb	~0.0%	~0.0%
WW −19.2%	210.26 fb	141.55 fb	~0.0%	~0.0%
$\tau\tau$ +12.6%	210.28 fb	141.66 fb	+0.1%	+0.1%
$\tau\tau$ −12.6%	210.05 fb	141.51 fb	−0.1%	−0.1%
ZZ + 57.0%	210.33 fb	141.71 fb	+0.1%	+0.1%
ZZ − 57.0%	210.00 fb	141.47 fb	−0.1%	−0.1%
$\gamma\gamma$ +100.5%	210.52 fb	141.86 fb	+0.2%	+0.2%
$\gamma\gamma$ +100.5%	209.83 fb	141.32 fb	−0.2%	−0.2%

- The effect of Higgs branching ratio is at most ± 0.2 %.

Statistical precision using MVA

- Using TMVA (BDT and Likelihood), the stat. precision is…
(Input variables : Z mass, Z p_T, Recoil mass, and Sphericity)

polarization	significance	stat. precision
left (-0.8, +0.3)	54.9 σ	1.8%
right (+0.8, -0.3)	60.8 σ	1.6%

- Achieved less than 2 % by using TMVA.
→ systematic uncertainty and model dependency should be considered.

- Input variables are not optimized, just some of variables which were used cut method.

Summary

- Using hadronic channel of e^-e^+ -> ZH,
  we can measure the total cross section of ZH with 2.2 %
  accuracy in right handed polarization.

- Some bias of the cuts were observed, we need to reduce this inconsistency with more optimized categorization.

- Systematic uncertainty of Higgs branching ratio is about 0.2 %.
  It is much smaller than stat. precision.

- MVA can improve stat. precision up to 1.6 %.
  Systematic uncertainty and model dependency should be studied.

Future plans

- Consider other systematic uncertainties, such as jet clustering, flavor tagging, background estimation and so on.
- Optimize categorization (more divided, more model independent cut...)
- Consider systematic uncertainty and model dependency of MVA.
- Apply same method to 350 GeV case.

backup slides

Sphericity, Thrust

$$S^{ab} = \frac{\sum_i p_i^a p_i^b}{\sum_i p_i^2} \quad a, b = x, y, z$$

$$T_{major} = \max_{|\vec{n}'|=1, \vec{n}'\cdot\vec{n}=0} \frac{\sum_i |\ \vec{p}_i \cdot \vec{n'}\ |}{\sum_i |\ \vec{p}_i\ |}$$

$$T_{minor} = \frac{\sum_i |\ \vec{p}_i \cdot \vec{n}''\ |}{\sum_i |\ \vec{p}_i\ |} \quad \text{with } \vec{n}'' \cdot \vec{n} = \vec{n}'' \cdot \vec{n}' = 0$$

diff./mean of each category

category	0lep,0tau btag	0lep,0tau no btag	0lep,1tau	0lep,1tau	0lep,≧2tau	1lep,0tau	1lep,≧1tau	≧2lep, ≧0tau
H->all	——	——	——	——	——	——	——	——
H->bb	-0.6%	-0.1%	-0.2%	-0.03%	~0.0%	-0.05%	~0.0%	~0.0%
H->WW(l)	+0.5%	+0.9%	~0.0%	-0.3%	-0.2%	-0.9%	-0.8%	-0.05%
WW(sl)	-0.1%	-2.7%	-0.2%	-0.9%	-0.2%	+0.7%	-0.2%	+0.02%
WW(h)	+0.7%	+0.9%	+0.9%	+0.1%	+0.03%	-0.07%	~0.0%	~0.0%
H->gg	+4.1%	+3.7%	-0.2%	+0.08%	~0.0%	-0.05%	~0.0%	~0.0%
H->$\tau\tau$	-0.3%	-1.7%	-0.3%	+0.5%	-0.3%	+0.02%	-0.2%	-0.02%
H->ZZ	+1.2%	-0.2%	-0.6%	+0.3%	-0.1%	-0.3%	+0.4%	+0.4%
H->cc	-3.8%	+1.1%	-0.4%	-0.2%	~0.0%	-0.08%	~0.0%	~0.0%
H->$\gamma\gamma$	+0.2%	-4.0%	+1.0%	-0.1%	+0.1%	-0.4%	+0.2%	+0.6%

Fraction to each category

category	0lep,0tau btag	0lep,0tau no b	0lep,1tau E_{vis}>180	0lep,1tau E_{vis}≦180	0lep, ≧2tau	1lep,0tau	1lep, ≧1tau	≧2lep, ≧0tau
H->all 549,279	60.2%	21.6%	3.5%	4.6%	2.7%	5.5%	1.3%	0.75%
H->bb 57.7%	92.0%	4.8%	2.3%	0.5%	0.04%	0.33%	0.01%	~0.0%
H->WW(l) 2.3%	2.2%	6.1%	0.04%	11.4%	6.9%	24.1%	26.3%	23.0%
WW(sl) 9.5%	7.5%	22.2%	8.9%	10.9%	1.4%	45.4%	3.4%	0.2%
WW(h) 9.8%	25.4%	66.5%	6.8%	0.4%	0.3%	0.5%	0.07%	0.0%
H->gg 8.6%	26.9%	69.8%	2.7%	3.0%	0.06%	0.3%	0.01%	0.0%
H->$\tau\tau$ 6.3%	3.9%	8.4%	2.8%	42.9%	35.4%	2.4%	4.2%	0.1%
H->ZZ 2.6%	34.4%	43.8%	5.0%	3.4%	1.5%	3.2%	2.7%	6.0%
H->cc 2.9%	28.3%	68.0%	2.9%	0.5%	0.05%	0.3%	0.01%	0.0%
H->$\gamma\gamma$ 0.2%	25.3%	65.7%	3.1%	2.1%	0.5%	0.7%	0.5%	1.9%

Hadronic | with Tau | Leptonic