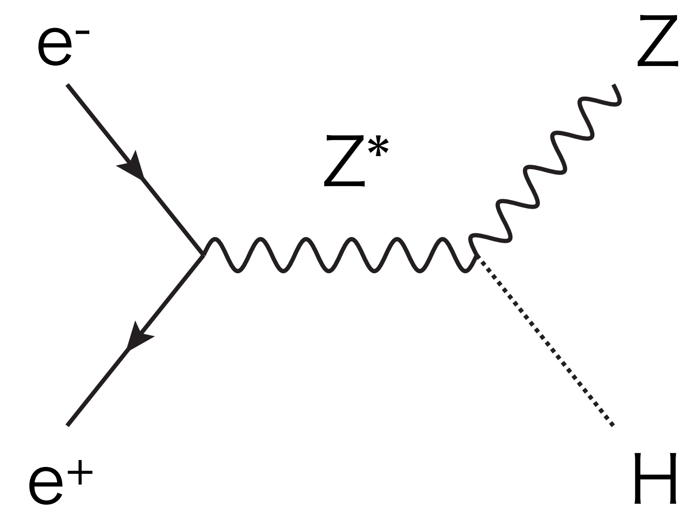
Hadronic recoil mass study using ZH -> qqH (ILC @ 250 GeV)

Tatsuhiko Tomita
Kyushu University

九州大学

ZH production

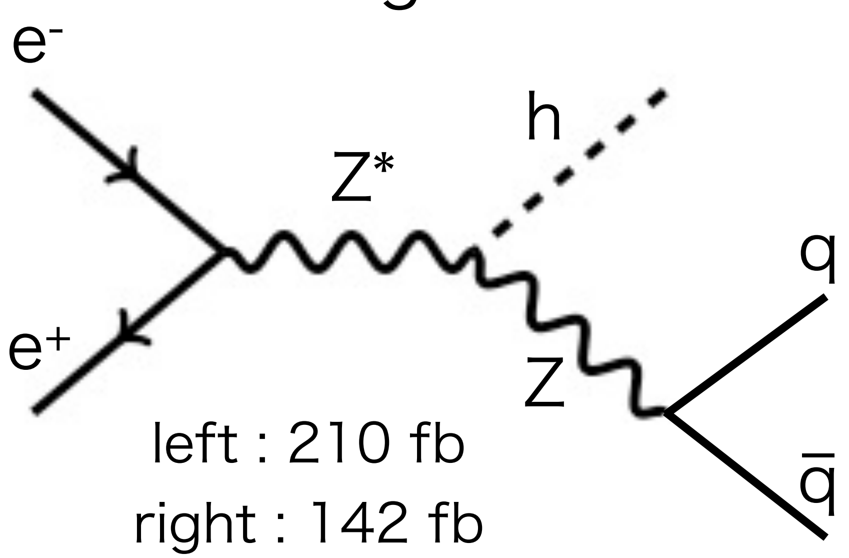


- Major Higgs production process at 250 GeV.
- Z boson will decay to …
  - charged leptons (e, μ, τ) - total ~10 %
  - neutrinos - total ~20 %
  - hadrons - **total ~70 %**
- Leptonic decay channel is useful for mass measurement.
  → High precision of mass measurement ~30 MeV.
  σ_{tot} measurement is 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.

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.

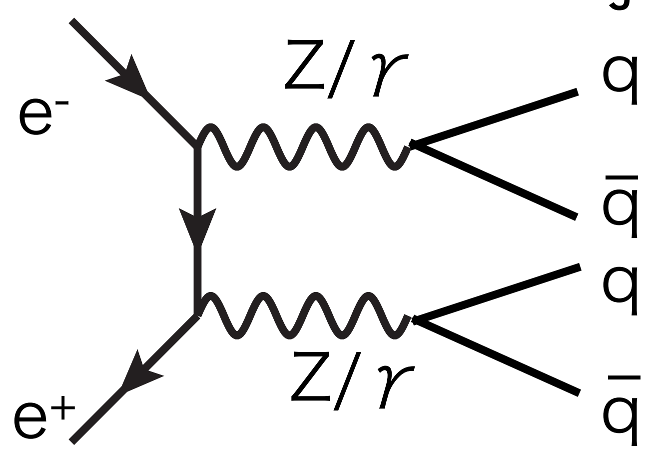
signal



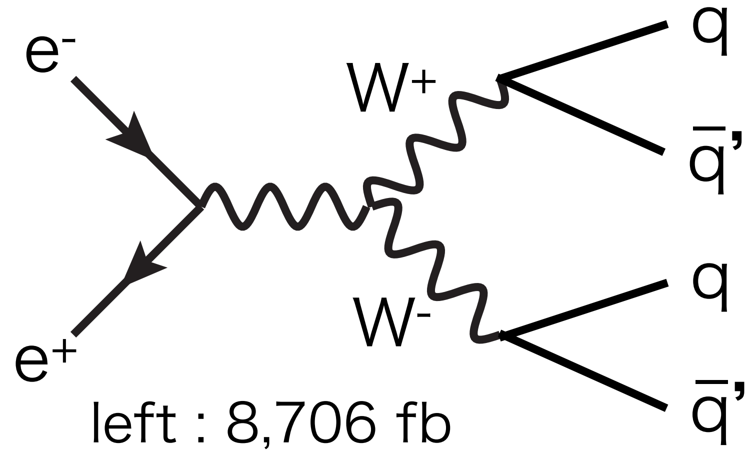
left : 210 fb
right : 142 fb

semi-leptonic events are also considerable BG.

Major background

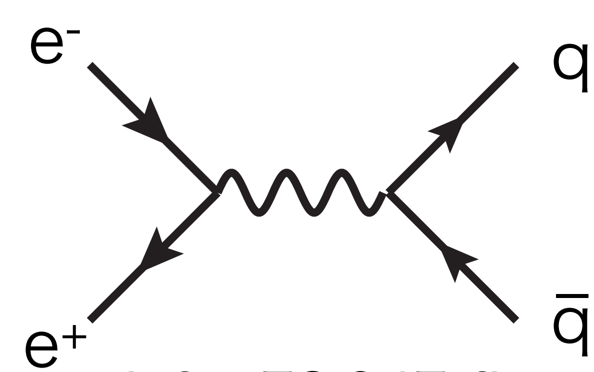


left : 841 fb
right : 403 fb



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 events.

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

- Forced 4 jet clustering, y threshold clustering were used.

Cut variables

- ZZ, WW mass cut, using forced 4 jet clustering.
  (the region of ± 10 GeV from Z(or W) mass was cut)
- Sphericity and Thrust cut.
  Sphericity was calculated with all particles.
  Thrust was calculated with the particles w/o ISR, IsoLep, Taujet.
- Reconstructed Z mass cut, using y value fixed clustering.
  (y = 0.0025)
- Reconstructed Z pT cut, using y value fixed clustering.
- Recoil mass cut, using y value fixed 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
ZZ, WW mass cut	82.8%	61.8%	97.9%	99.9%
Sphericity Thrust cut	78.5%	39.5%	33.6%	25.8%
Z mass Z p_T Recoil	46.2%	8.3%	1.4%	0.3%

Cut efficiency for each Higgs decay branch

- In order not to depend 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

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

The difference of cut efficiency is defined,

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

And then, 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 categorization	After categorization (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 categorization and optimization of cut, diff./mean is at most 5.5 %.
- Need to 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 the difference from mean cut efficiency, the stat. precision calculated with each categories.
- In this case, 2.2 % stat. precision with right polarization.

Systematic uncertainty

- The uncertainty of the Higgs branching ratio was studied.
- Changed Higgs branching ratio with 5 %.
  (ex. H->bb +5%, the others - a few %)

$$\sigma_{tot} = \frac{N_{\mathrm{eve}}}{\mathcal{L} \times \epsilon}$$

(keep total N_{eve})

(∵ keep total BR = 1)

bb + 5% (57.7->62.7)	210.27	141.51	+0.1%	-0.1%
bb - 5% (57.7->52.7)	210.06	141.67	-0.1%	+0.1%
cc + 5% (2.9->7.9)	209.07	140.84	-0.5%	-0.5%
cc - 5% (2.9->0.0)	210.77	142.00	+0.3%	+0.3%
gg + 5% (8.6->13.6)	209.95	141.63	-0.1%	~0.0%
gg - 5% (8.6->3.6)	210.38	141.56	+0.1%	~0.0%
WW + 5% (21.6->26.6)	210.01	141.61	-0.1%	~0.0%
WW - 5% (21.6->16.6)	210.15	141.46	~0.0%	-0.1%
tau + 5% (6.3->11.3)	210.40	141.73	+0.1%	+0.1%
tau - 5% (6.3->1.3)	209.93	141.44	-0.1%	-0.1%
ZZ + 5% (2.6->7.6)	210.50	141.86	+0.2%	+0.2%
ZZ - 5% (2.6->0.0)	210.09	141.51	~0.0%	-0.1%
$\gamma\gamma$ + 5% (0.2->5.2)	218.57	148.18	+4.0%	+4.7%
$\gamma\gamma$ - 5% (0.2->0.0)	209.83	141.32	-0.2%	-0.2%

- The effect of Higgs branching ratio is at most ± 0.5 % .

Sensitivity

- Using current stat. precision (2.2 % with right polarization), the sensitivity was calculated. (Changed one of the Higgs branch, the others didn't change)

$$\text{deviation} = \frac{\sigma_{tot}^{\text{SIM}} / \sigma_{tot}^{\text{SM}}}{\text{stat. precision}}$$

H→ $b\bar{b}$ (−0.8,+0.3)	± 1.7%
H→ $b\bar{b}$ (+0.8,−0.3)	± 1.6%
H→ $c\bar{c}$ (−0.8,+0.3)	± 4.0%
H→ $c\bar{c}$ (+0.8,−0.3)	± 3.2%
H→ gg (−0.8,+0.3)	± 3.7%
H→ gg (+0.8,−0.3)	± 3.0%
H→ WW (−0.8,+0.3)	± 3.7%
H→ WW (+0.8,−0.3)	± 2.7%
H→ $\tau\tau$ (−0.8,+0.3)	± 0.9%
H→ $\tau\tau$ (+0.8,−0.3)	± 0.8%
H→ ZZ (−0.8,+0.3)	± 3.4%
H→ ZZ (+0.8,−0.3)	± 3.1%
H→ $\gamma\gamma$ (−0.8,+0.3)	± 3.8%
H→ $\gamma\gamma$ (+0.8,−0.3)	± 3.1%

1 σ deviation from SM like Higgs.

% means "relative" value from SM like Higgs.

- bb, τ τ can be detected with small difference from SM Higgs.

Statistical precision using MVA

- Using TMVA (BDTG 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 0.5 %.
  It is much smaller than stat. precision.
- MVA can improve stat. precision up to 1.6 %.
  Systematic uncertainty and model dependency should be studied.

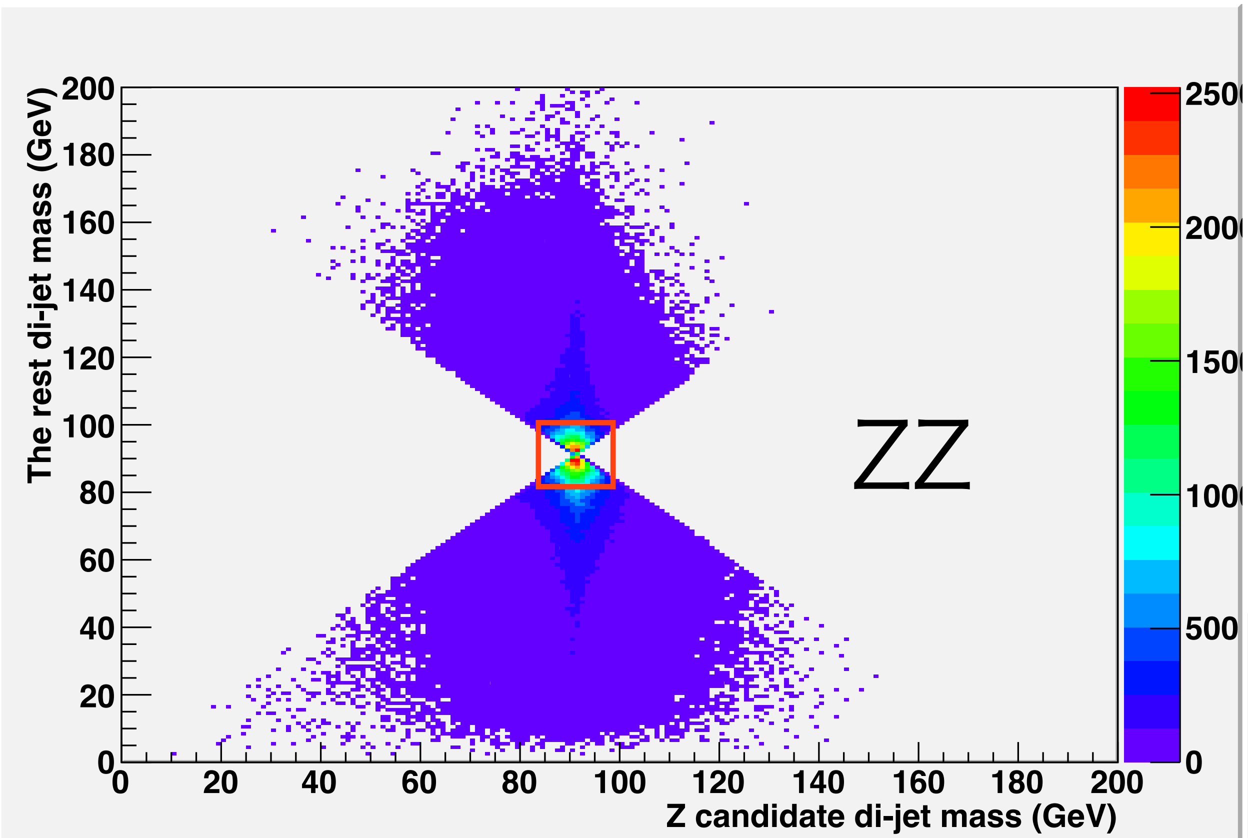
Future plans

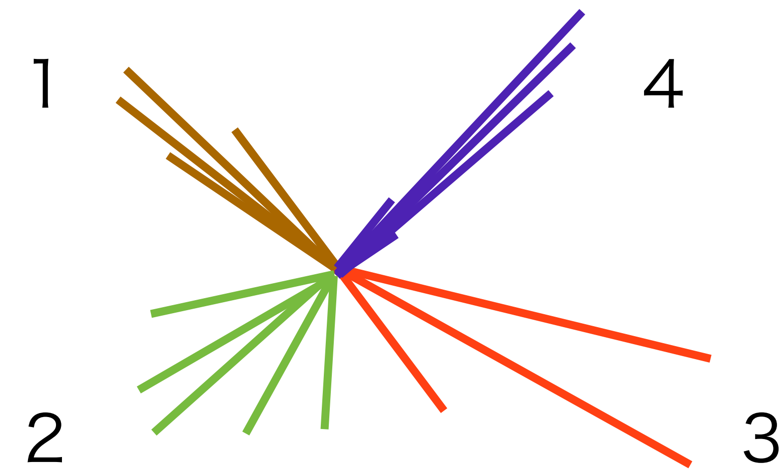
- Consider other systematic uncertainties, such as jet clustering, flavor tagging, background estimation and so on.
- Optimize categorization (more divided, more model independent cut...)
- Consider systematics and model dependency of MVA.
- Apply the same method to 350 GeV case.

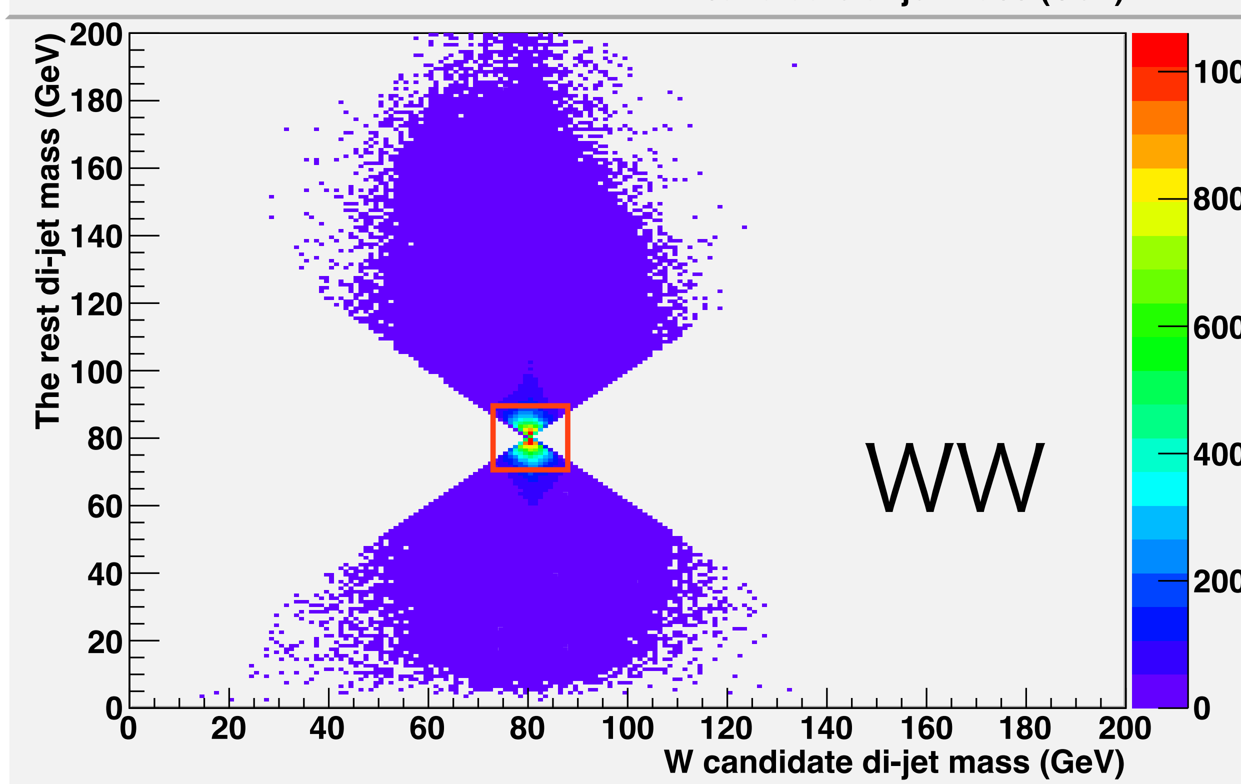
backup slides

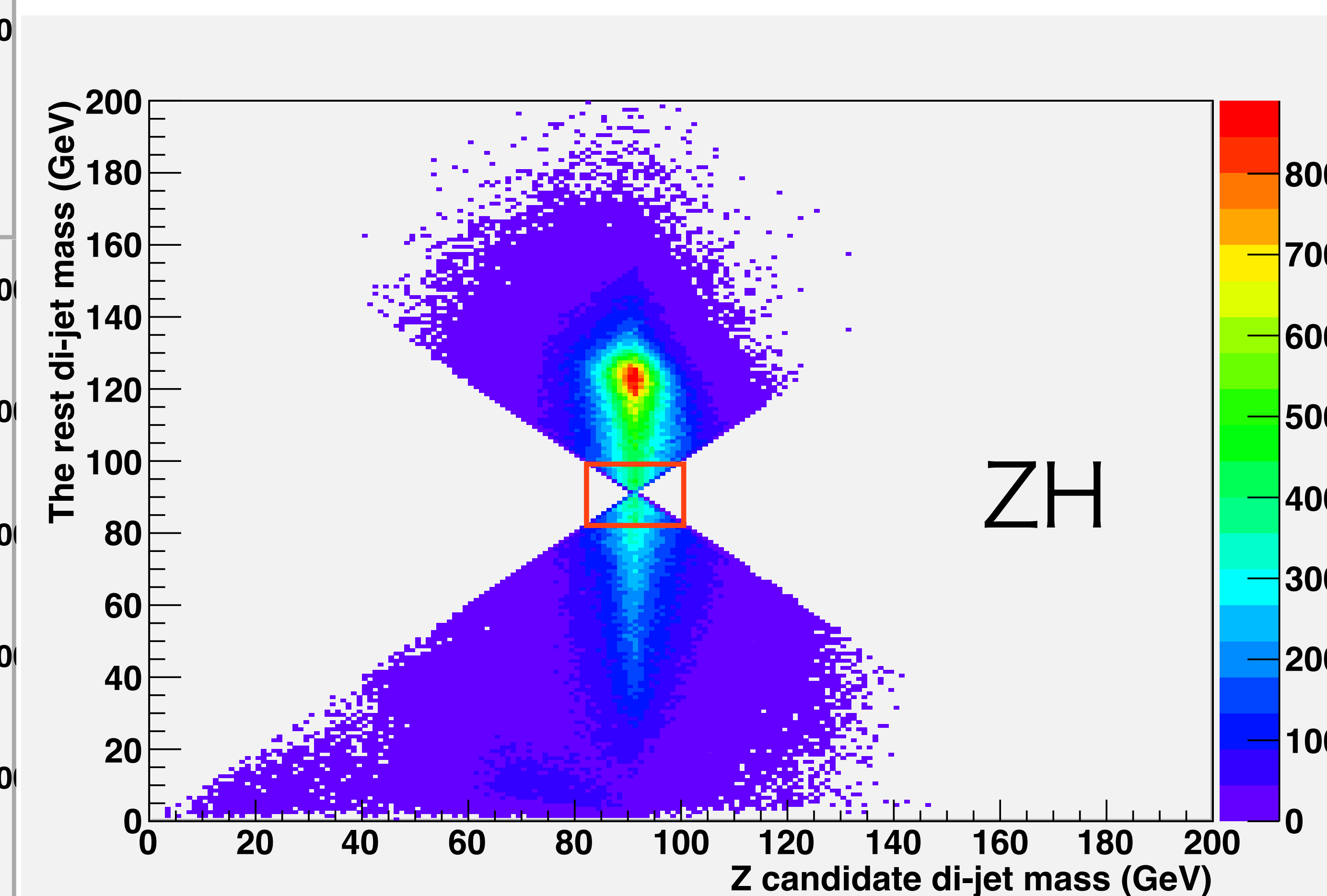
Background estimation - 1

- For 4 fermion backgrounds : forced 4 jet clustering



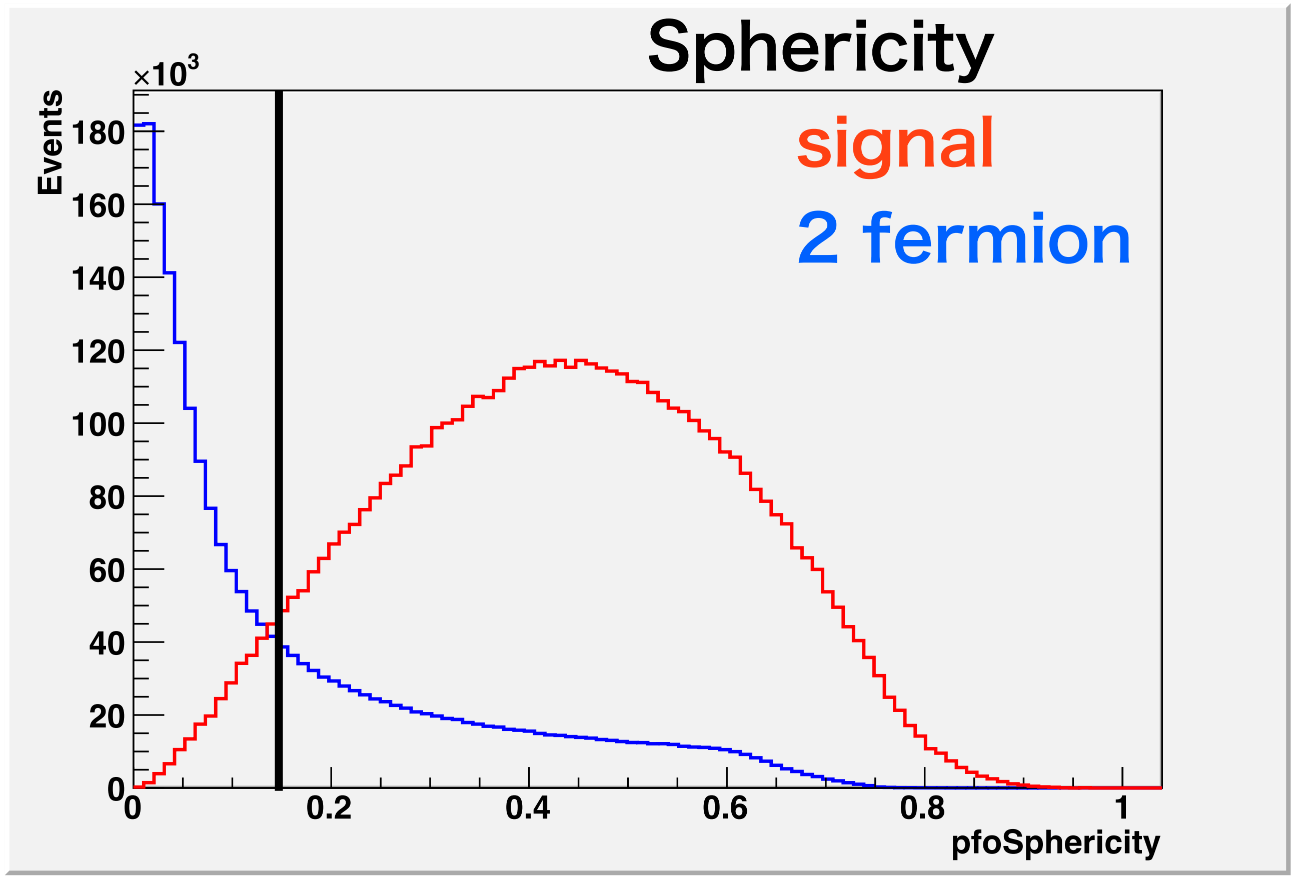


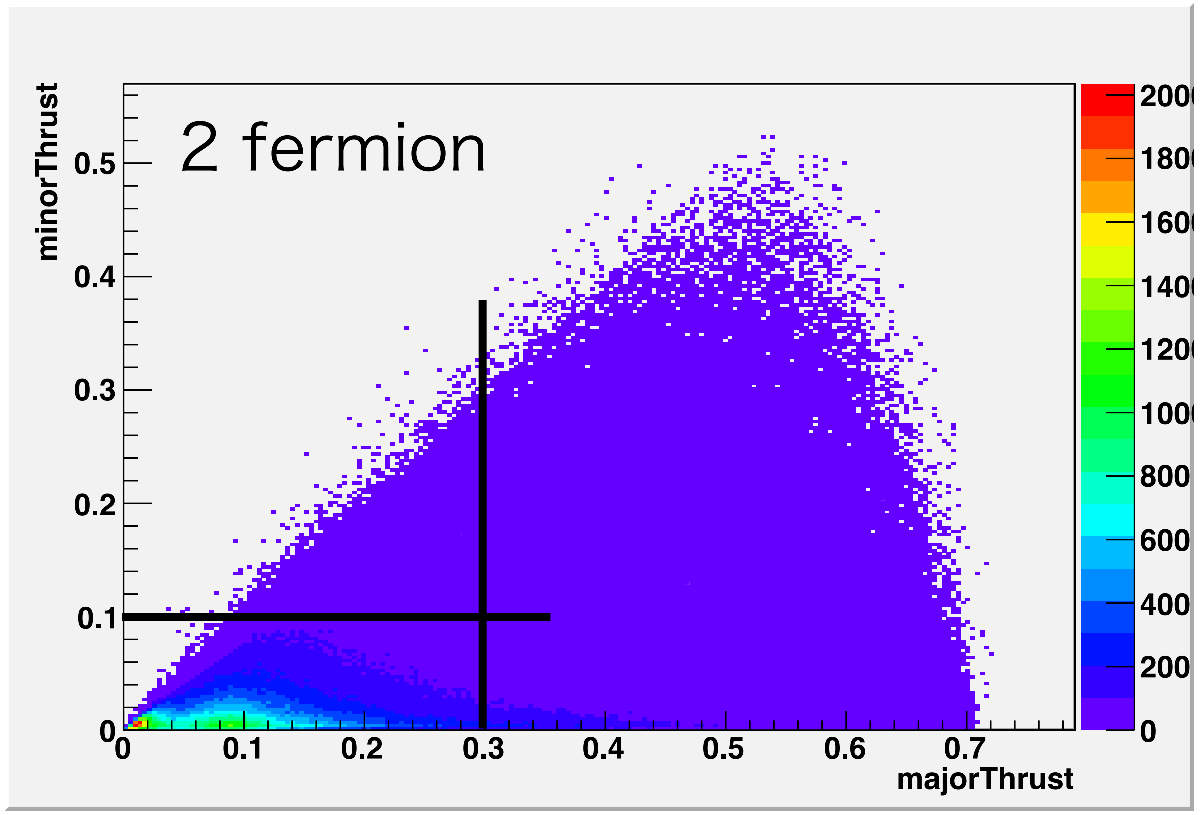


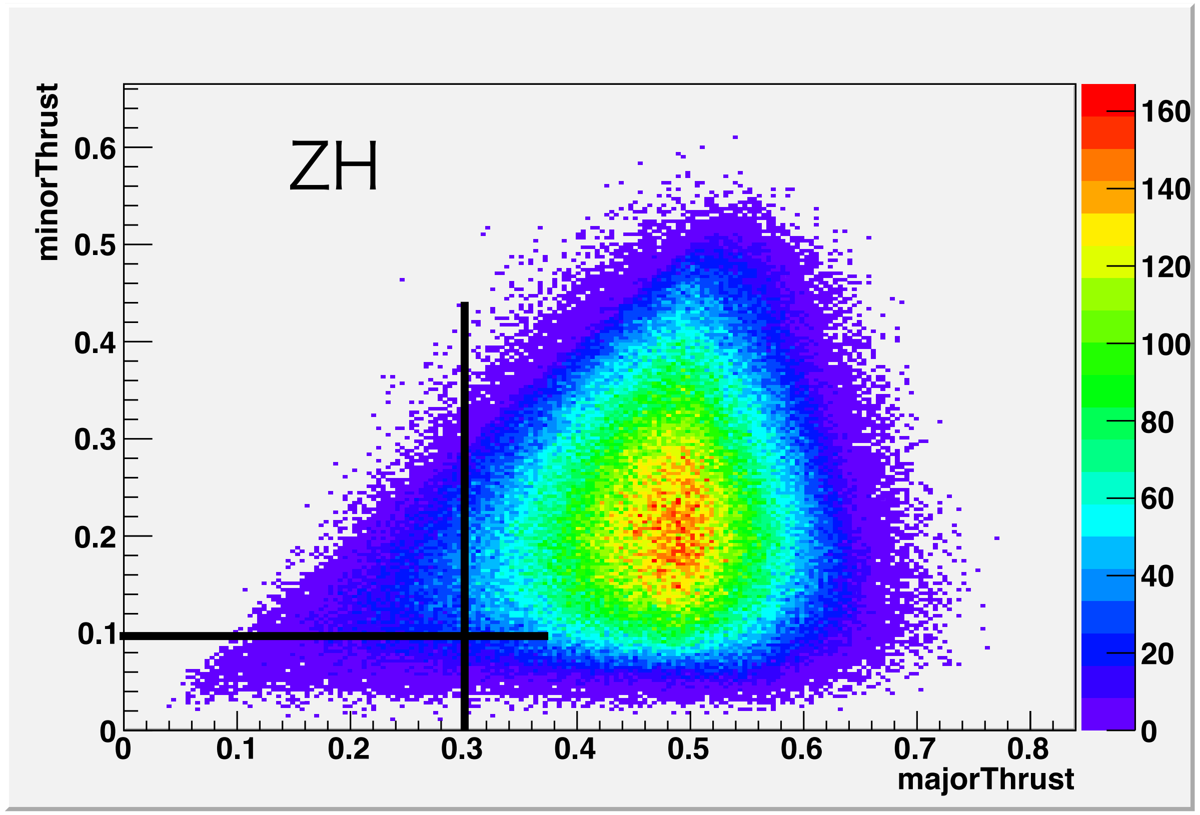


Background estimation - 2

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



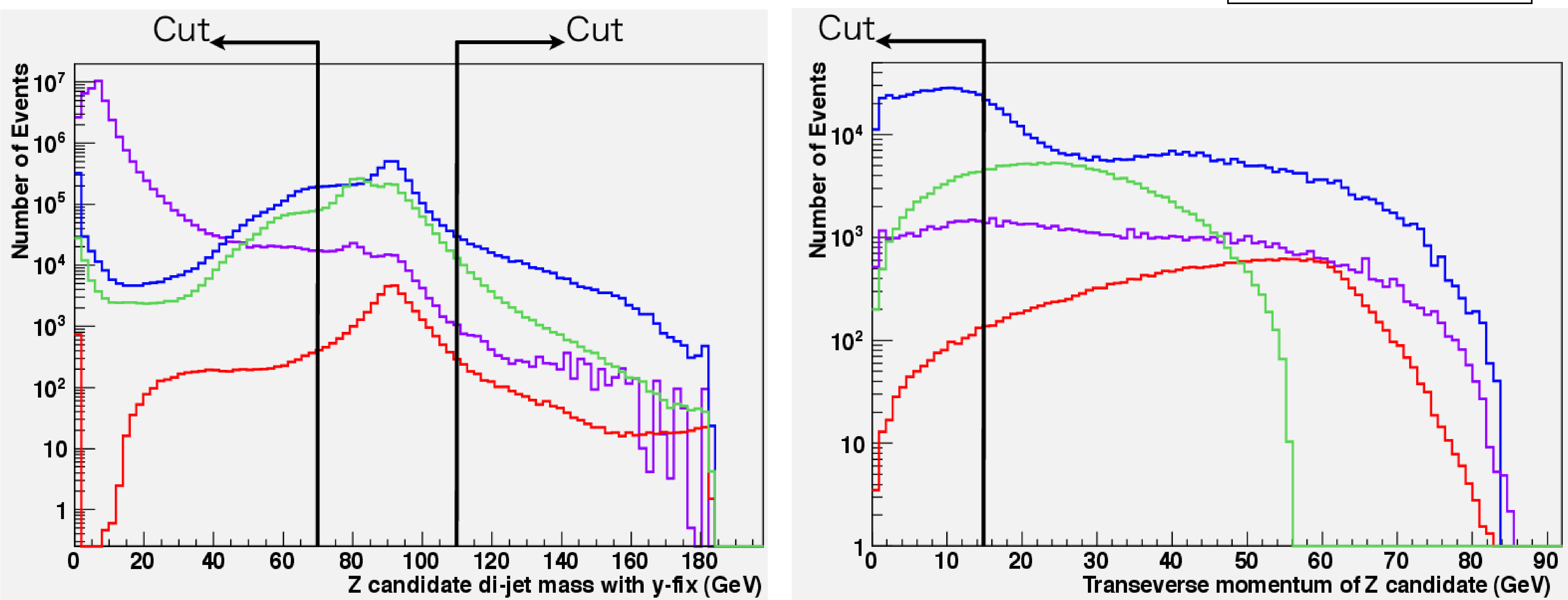




Background estimation - 3

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

signal
2 fermion
4 fermion
others

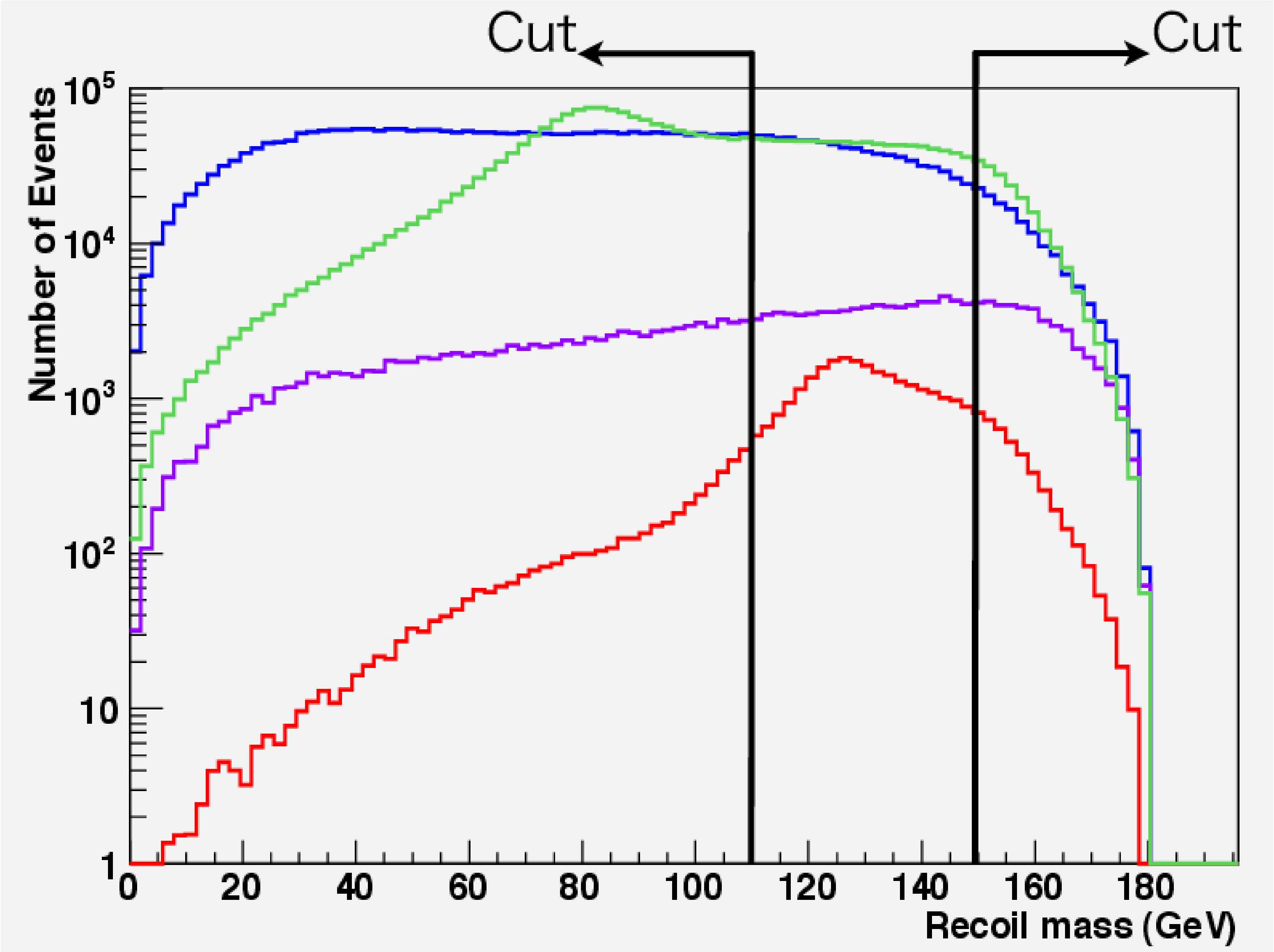


Background estimation - 4

- 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.



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 $\epsilon^i (= \epsilon_n^i)$.

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

Then we can get

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

fraction into 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%

hadrons | with Tau | with IsoLep

After optimization of cut

category	0lep,0tau btag	0lep,0tau no btag	0lep,1tau	0lep,1tau	0lep,≧2tau	1lep,0tau	1lep,≧1tau	≧2lep, ≧0tau	sum
H->all	——	——	——	——	——	——	——	——	
H->bb	-0.6%	-0.1%	-0.2%	-0.03%	~0.0%	-0.05%	~0.0%	~0.0%	±0.6%
H->WW(l)	+0.5%	+0.9%	~0.0%	-0.3%	-0.2%	-0.9%	-0.8%	-0.05%	±1.6%
WW(sl)	-0.1%	-2.7%	-0.2%	-0.9%	-0.2%	+0.7%	-0.2%	+0.02%	±3.0%
WW(h)	+0.7%	+0.9%	+0.9%	+0.1%	+0.03%	-0.07%	~0.0%	~0.0%	±1.5%
H->gg	+4.1%	+3.7%	-0.2%	+0.08%	~0.0%	-0.05%	~0.0%	~0.0%	±5.5%
H->$\tau\tau$	-0.3%	-1.7%	-0.3%	+0.5%	-0.3%	+0.02%	-0.2%	-0.02%	±1.9%
H->ZZ	+1.2%	-0.2%	-0.6%	+0.3%	-0.1%	-0.3%	+0.4%	+0.4%	±1.5%
H->cc	-3.8%	+1.1%	-0.4%	-0.2%	~0.0%	-0.08%	~0.0%	~0.0%	±4.0%
H->$\gamma\gamma$	+0.2%	-4.0%	+1.0%	-0.1%	+0.1%	-0.4%	+0.2%	+0.6%	±4.2%