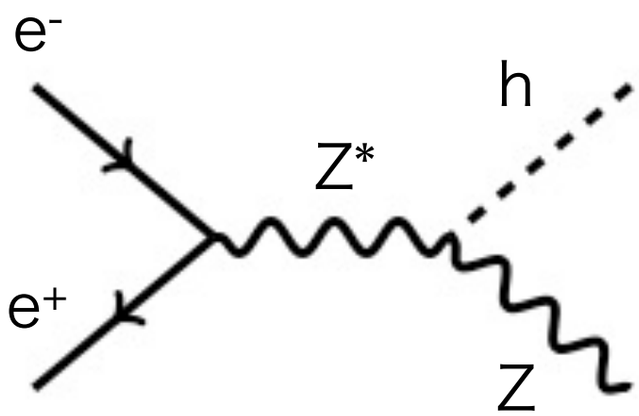
Higgs recoil mass study using ZH->qqH @ 250 GeV ILC

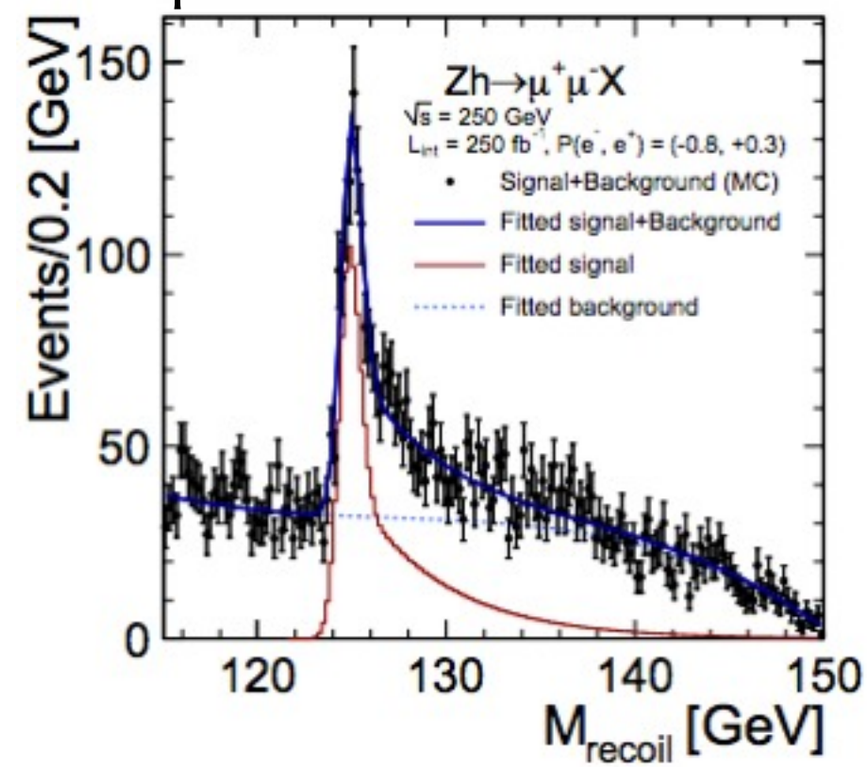
Tatsuhiko Tomita (Kyushu Univ.)
Akiya Miyamoto (KEK), Taikan Suehara (Kyushu Univ.)
& ILC physics working group

Overview - qqH channel



At lepton collider, we can measure Higgs without looking Higgs directly.

-> Model Independent search



The branching ratio of Z -> leptonic is ~3.5% for each generation.

In contrast, the branching ratio of **Z -> hadronic is ~70%**.

- **More statistics**

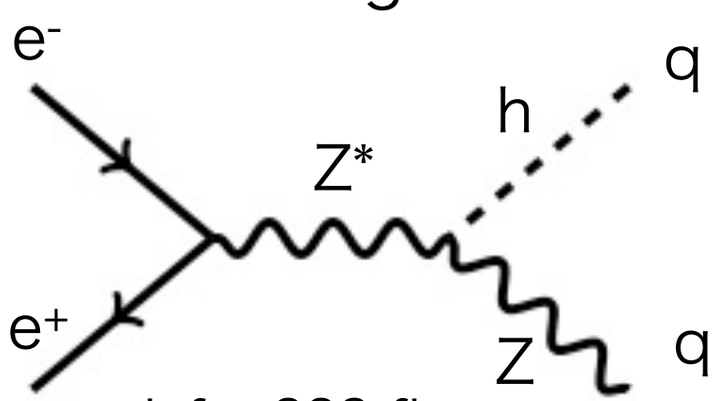
Model independent?

- **More background**

dataset and analysis flow

Higgs mass	Ecm	Luminosity	Polarization	Detector
125 GeV	250 GeV	250 fb^{-1}	left: (-0.8, +0.3) right:(+0.8, -0.3)	ILD_o1_v05 DBD ver.

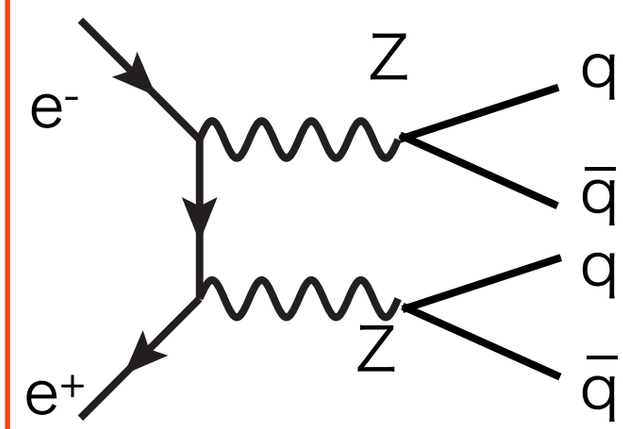
Signal



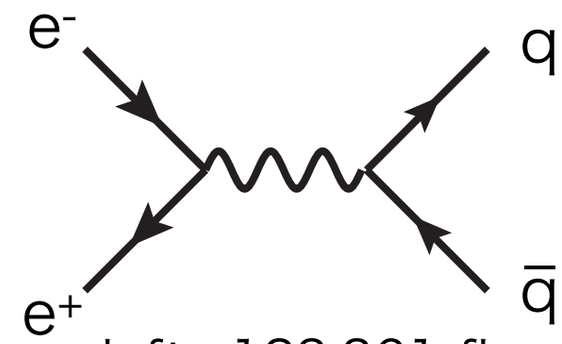
left : 333 fb
right : 234 fb

Semi-leptonic events can also be backgrounds

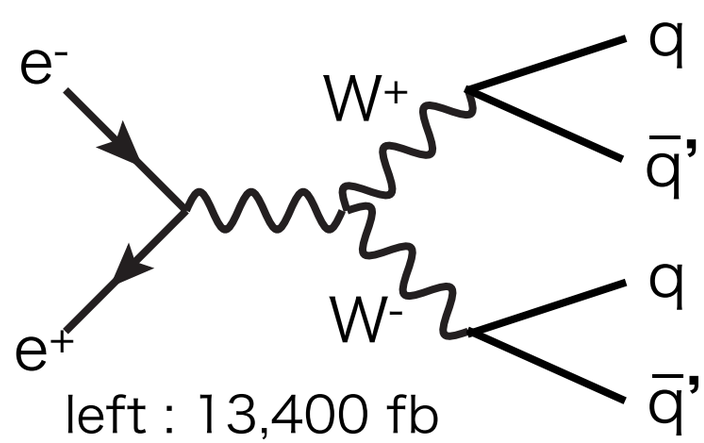
Main background



left : 1,322 fb
right : 685 fb



left : 123,361 fb
right : 77,059 fb



left : 13,400 fb
right : 1,610 fb

dataset and analysis flow

Higgs mass	Ecm	Luminosity	Polarization	Detector
125 GeV	250 GeV	250 fb^{-1}	left: (-0.8, +0.3) right:(+0.8, -0.3)	ILD_o1_v05 DBD ver.

all PFOs → Iso lepton finder → ISR finder → Tau finder → remaining PFOs → Durham jet clustering

remove IsoLep

remove ISR

remove Tau jet

Initial State Radiation Finder

- 30 GeV && | cos θ | > 0.95
- 80 GeV && | cos θ | > 0.85
- charge == 0
- ECAL/(HCAL+ECAL) > 0.9

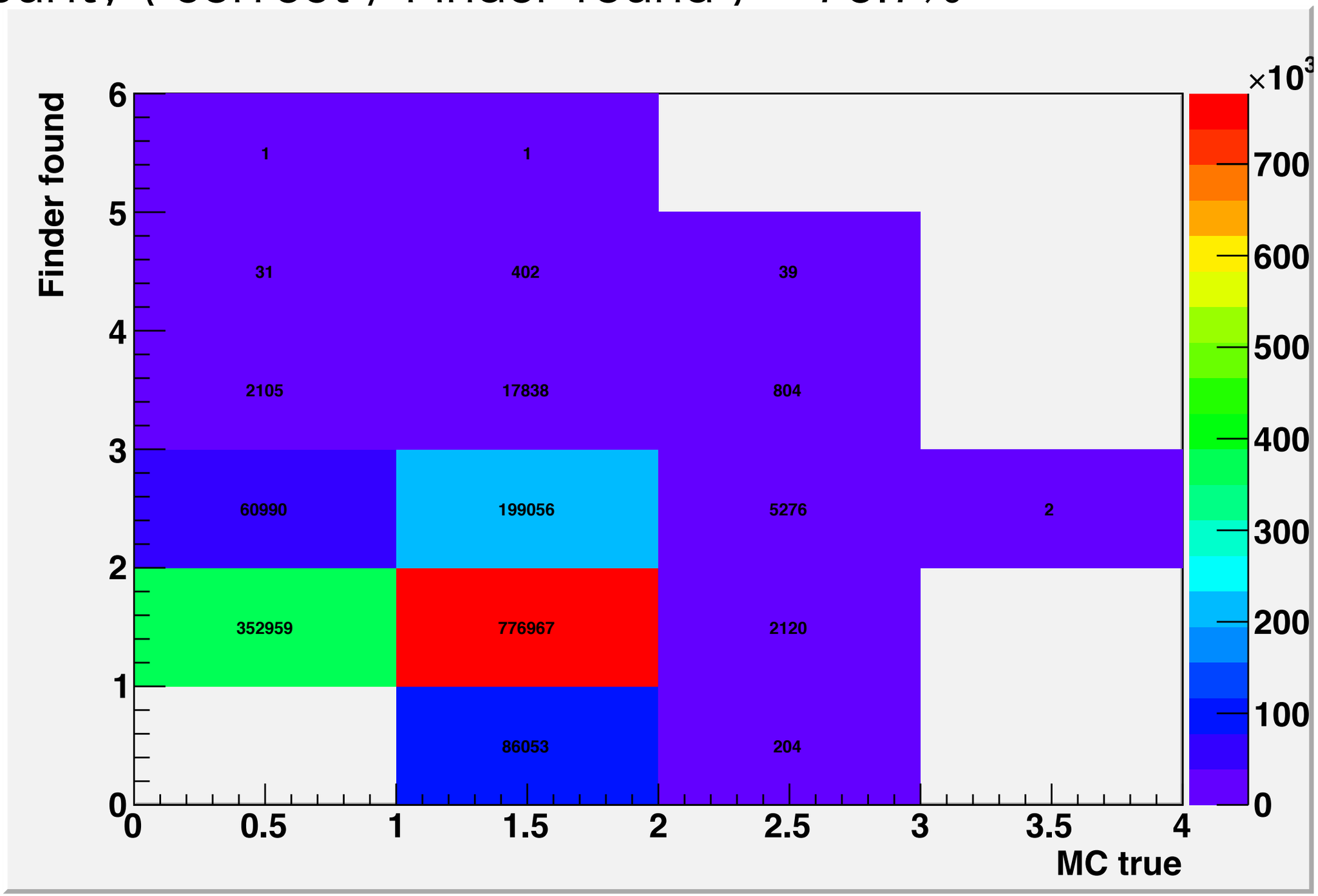
Durham jet clustering

2 jet, 3 jet, 4jet,

5 jet, y-value

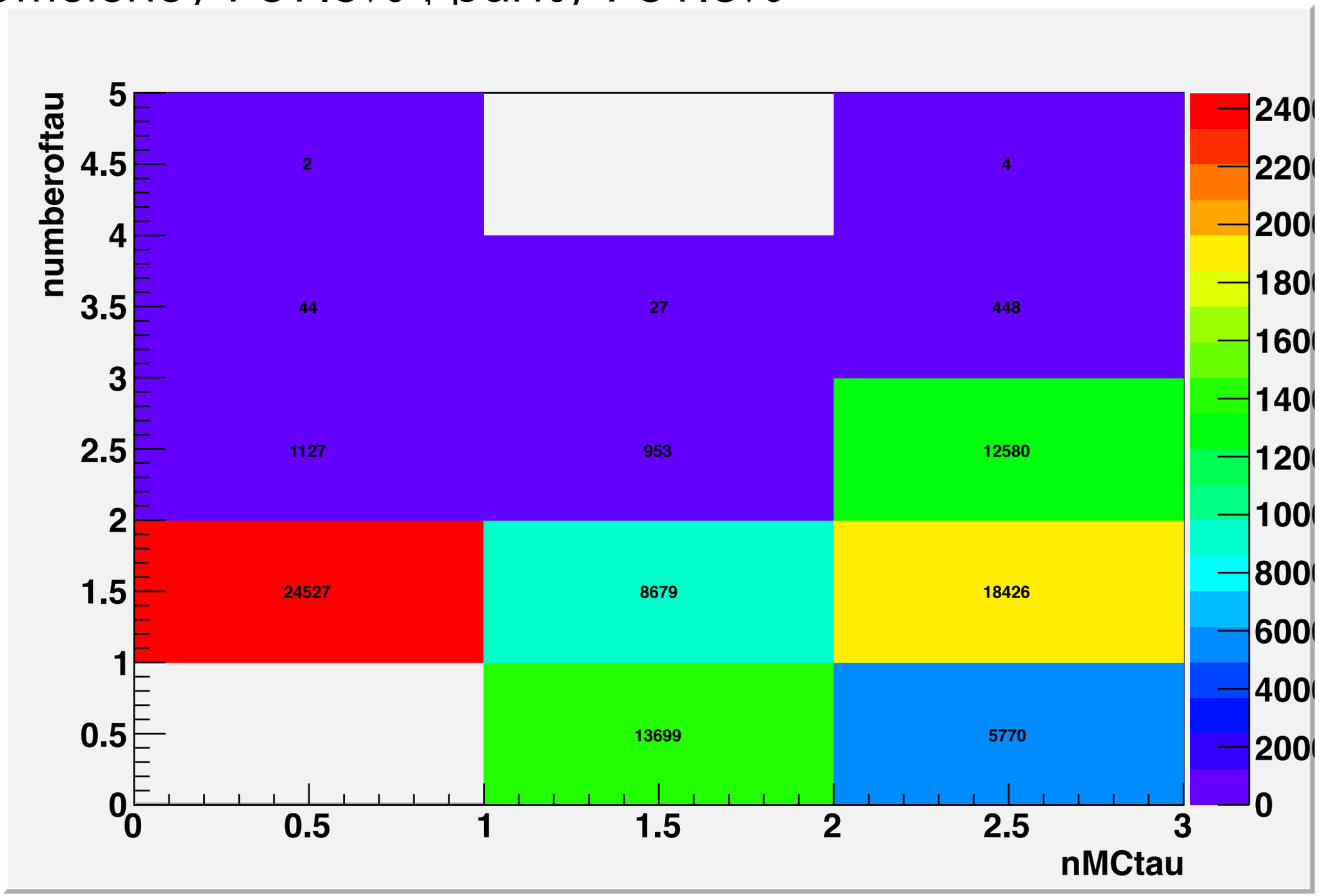
Review of ISR finder

- efficiency (Finder output / MC true) = 92.1%
- purity (correct / Finder found) = 70.7%



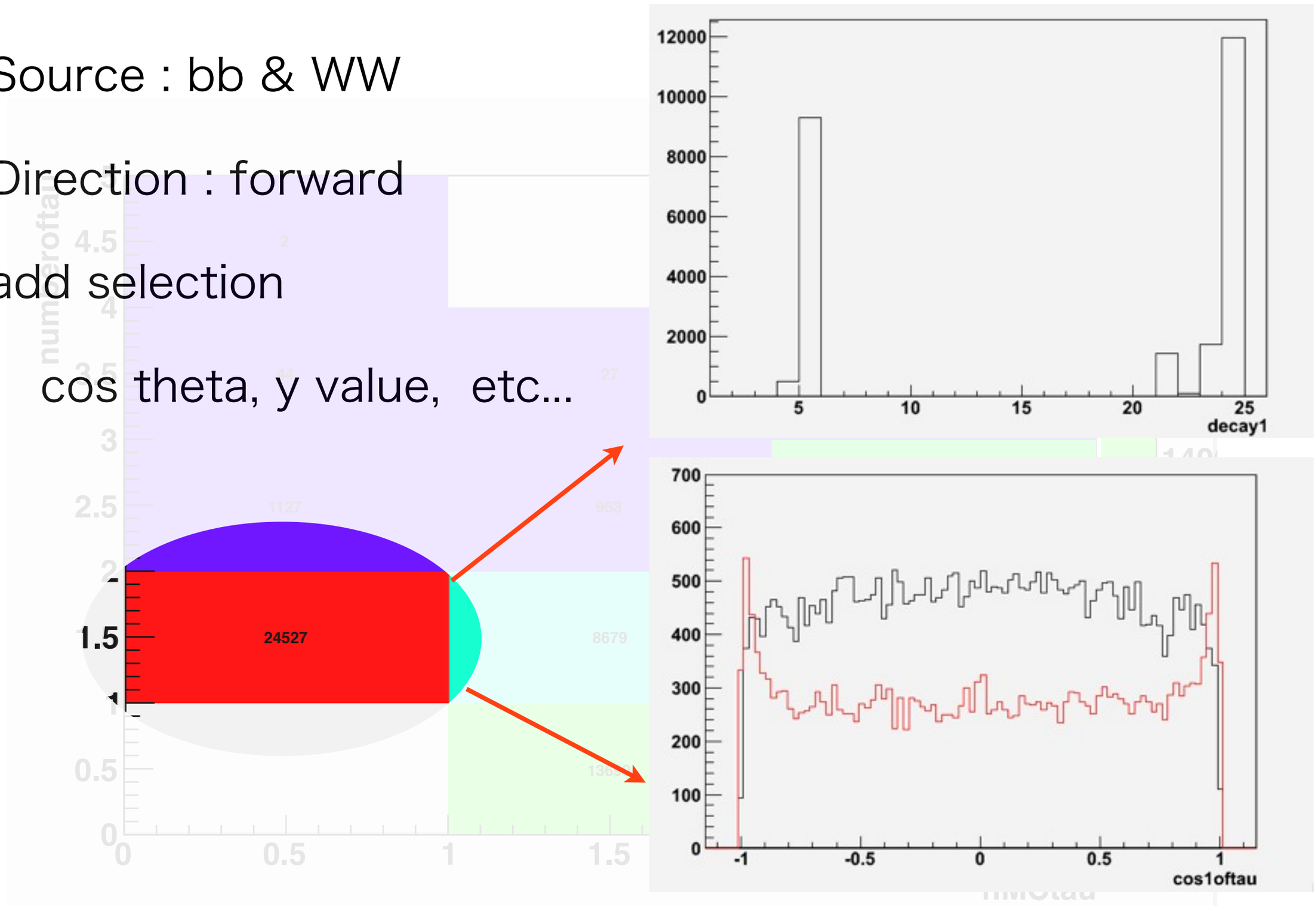
Review of Tau finder

- Target : hadronic tau (lepton can be found in IsoLep finder)
- efficiency : 67.9% , purity : 61.5%



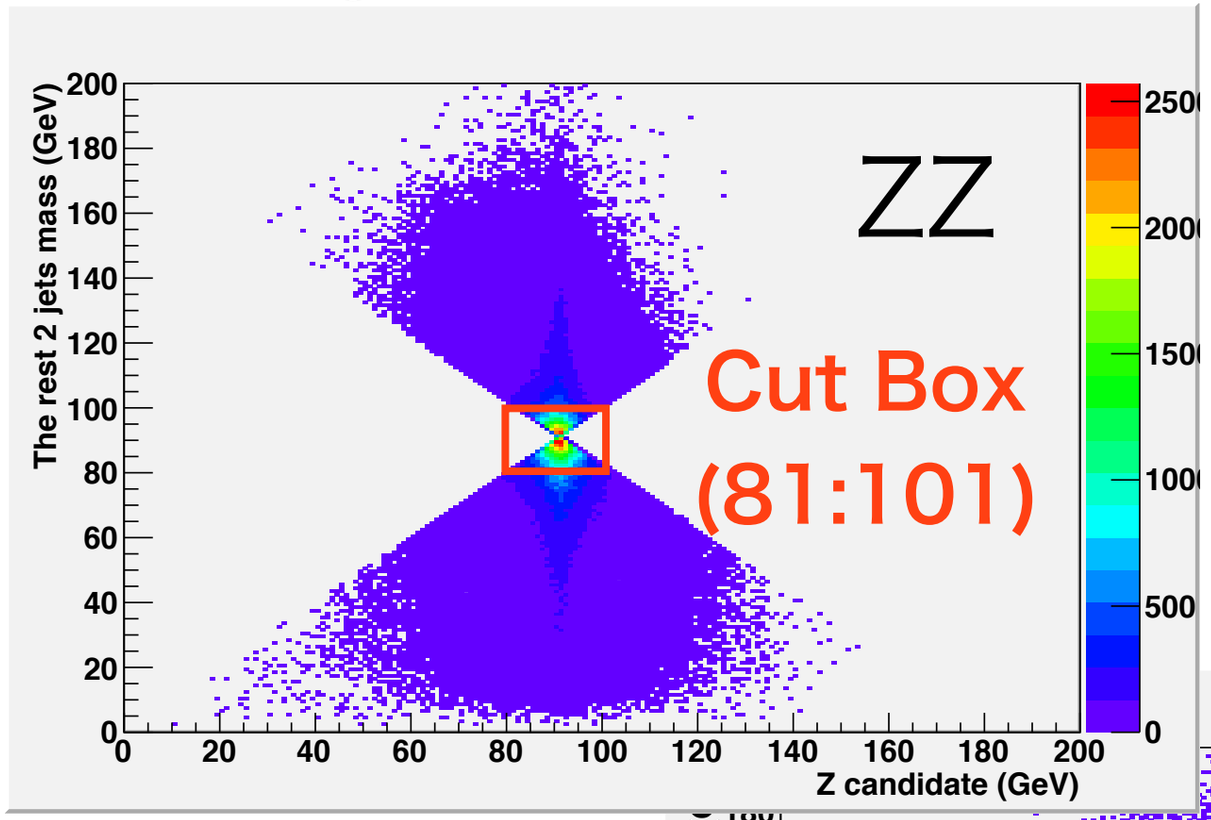
Issue of tau finder

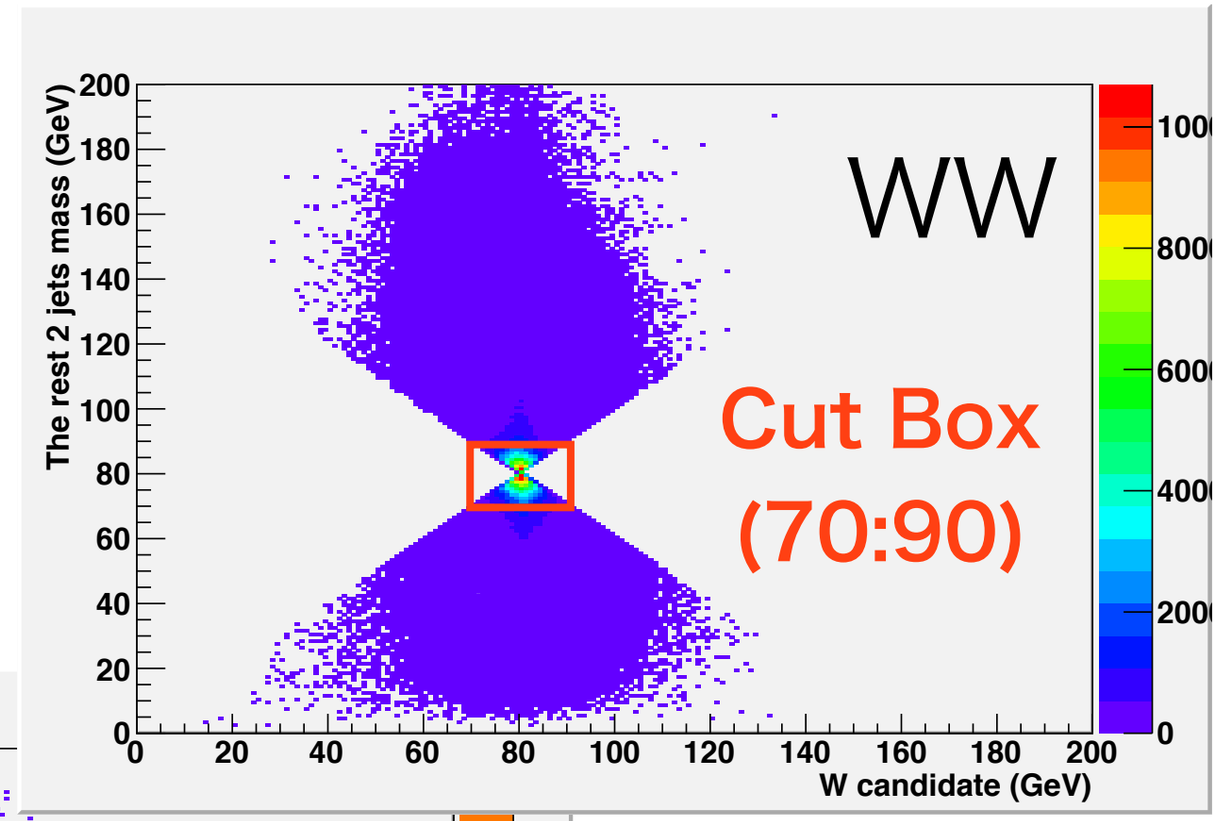
- Finder found == 1 && MC tau ==0 (miss identification)
- Source : bb & WW
- Direction : forward
- add selection
  - cos theta, y value, etc...

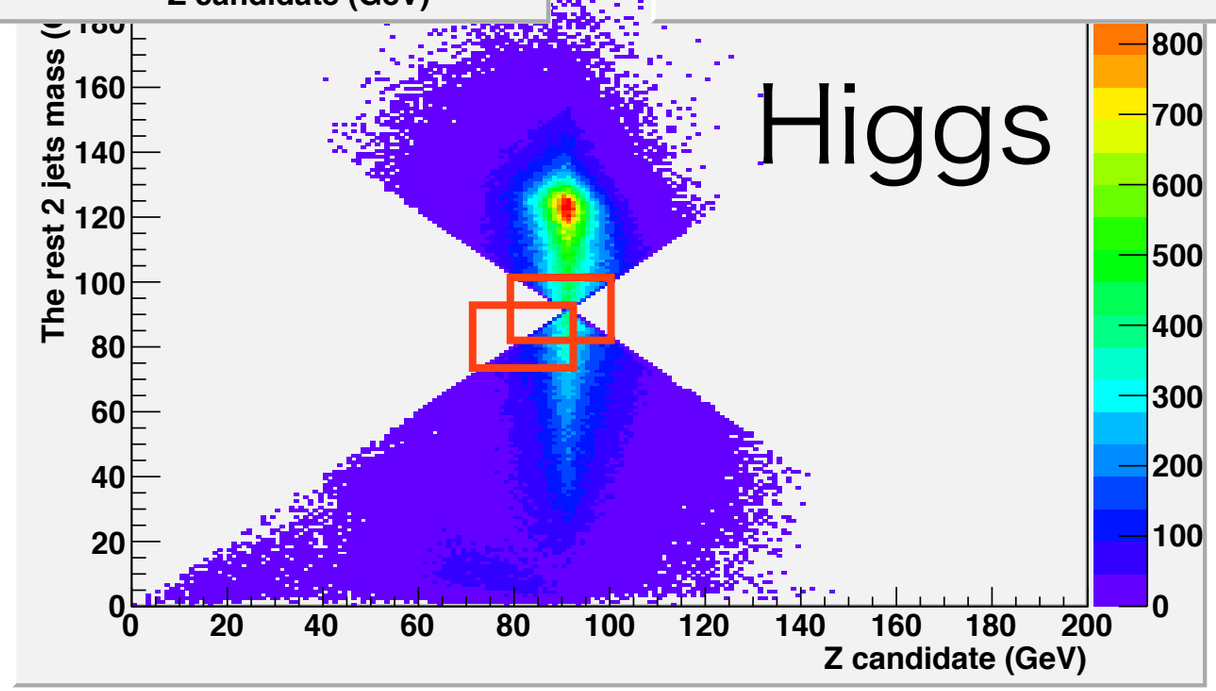


Cuts (1)

- Using forced-4jet clustering,
  Targets : ZZ->hadronic, WW->hadronic

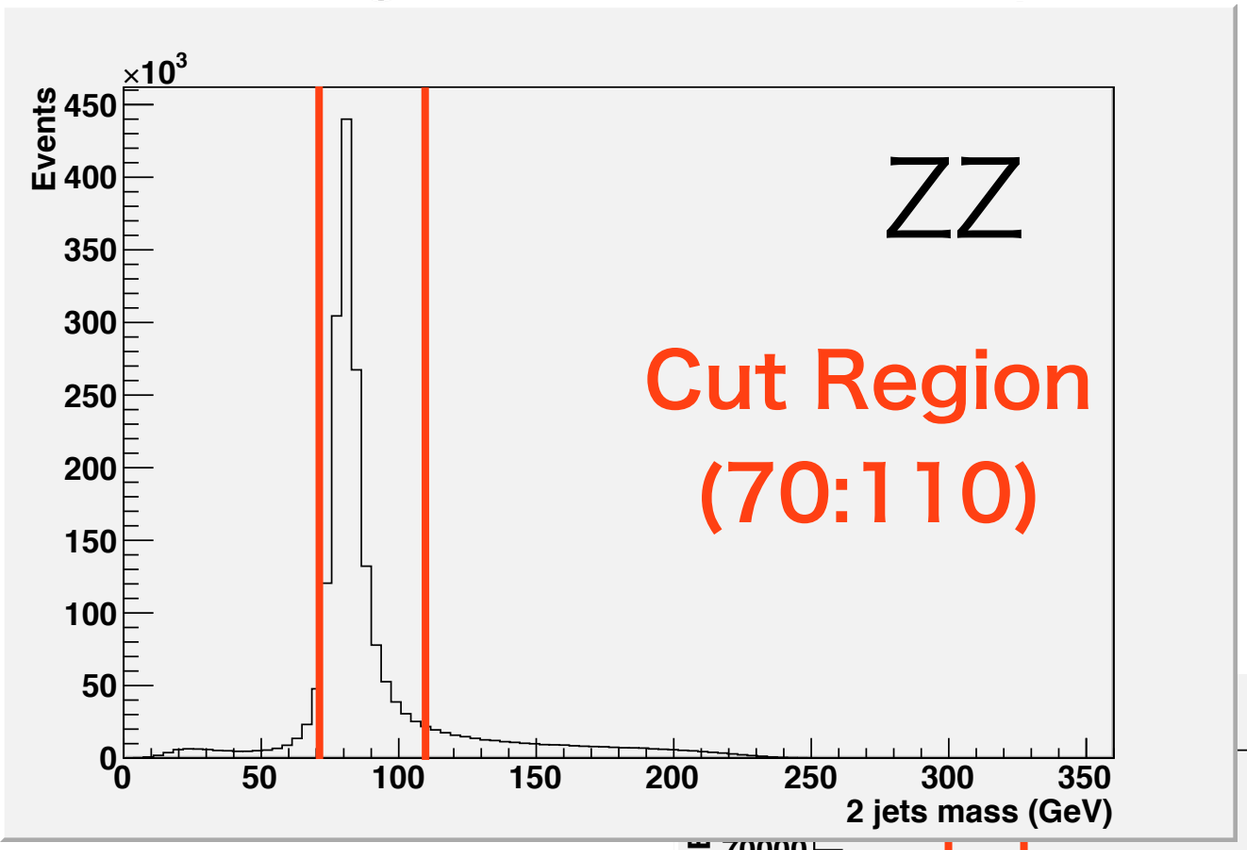


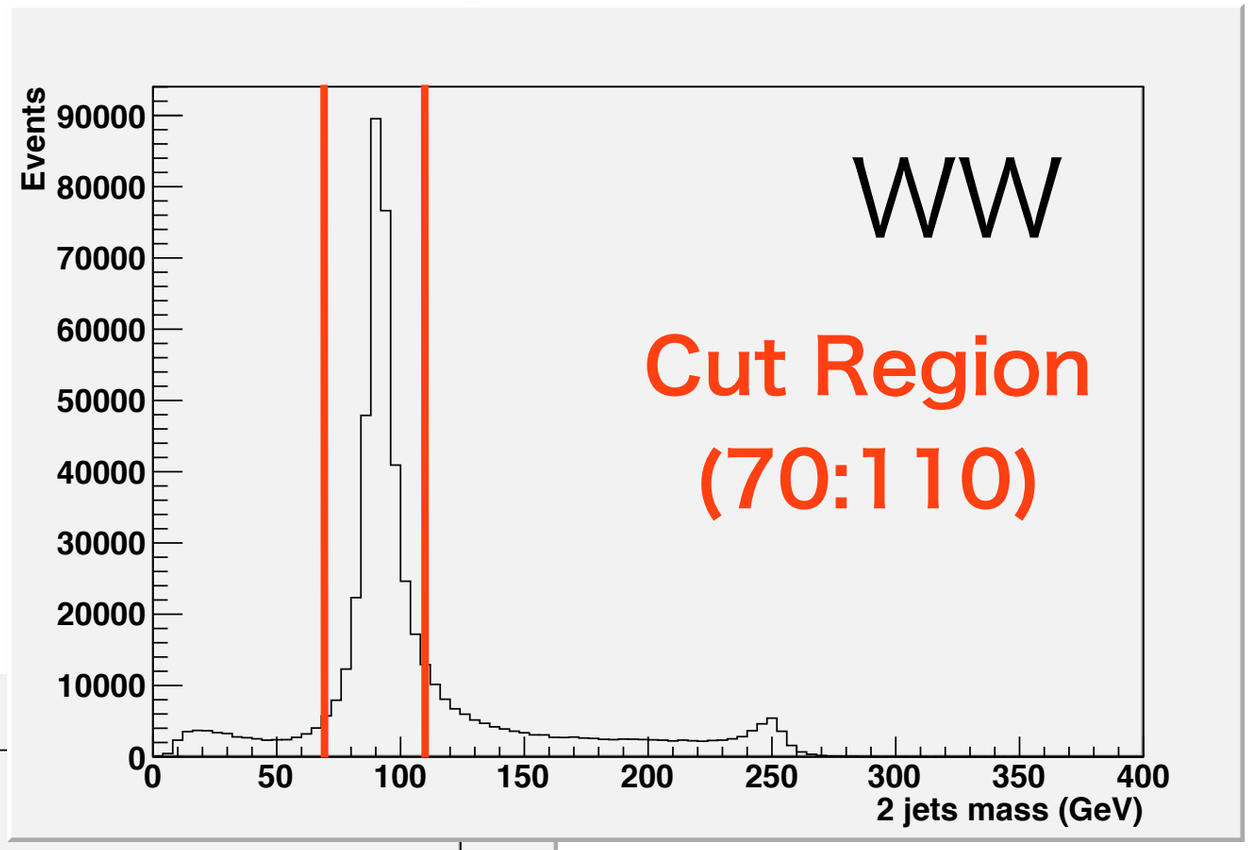


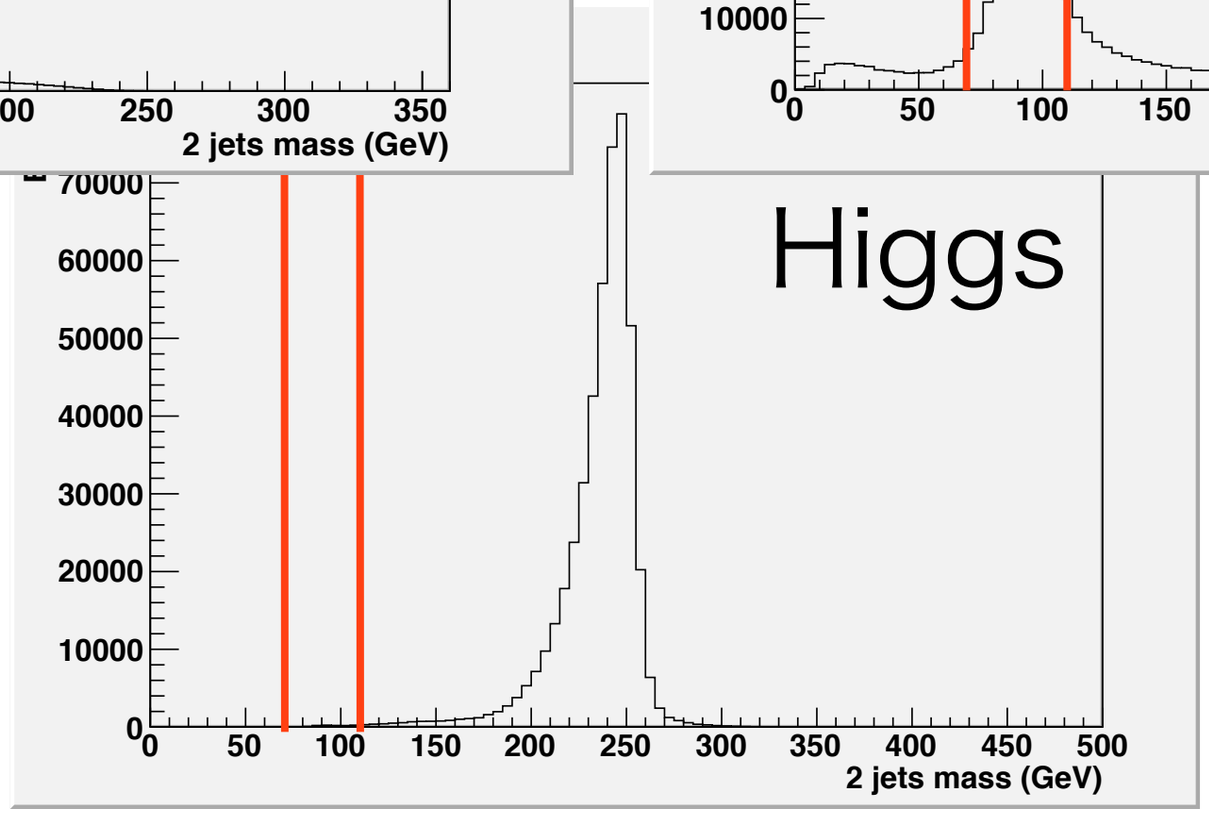


Cuts (2)

- Using forced-2jet clustering,
  Targets : ZZ->semi-leptonic, WW->semi-leptonic etc...

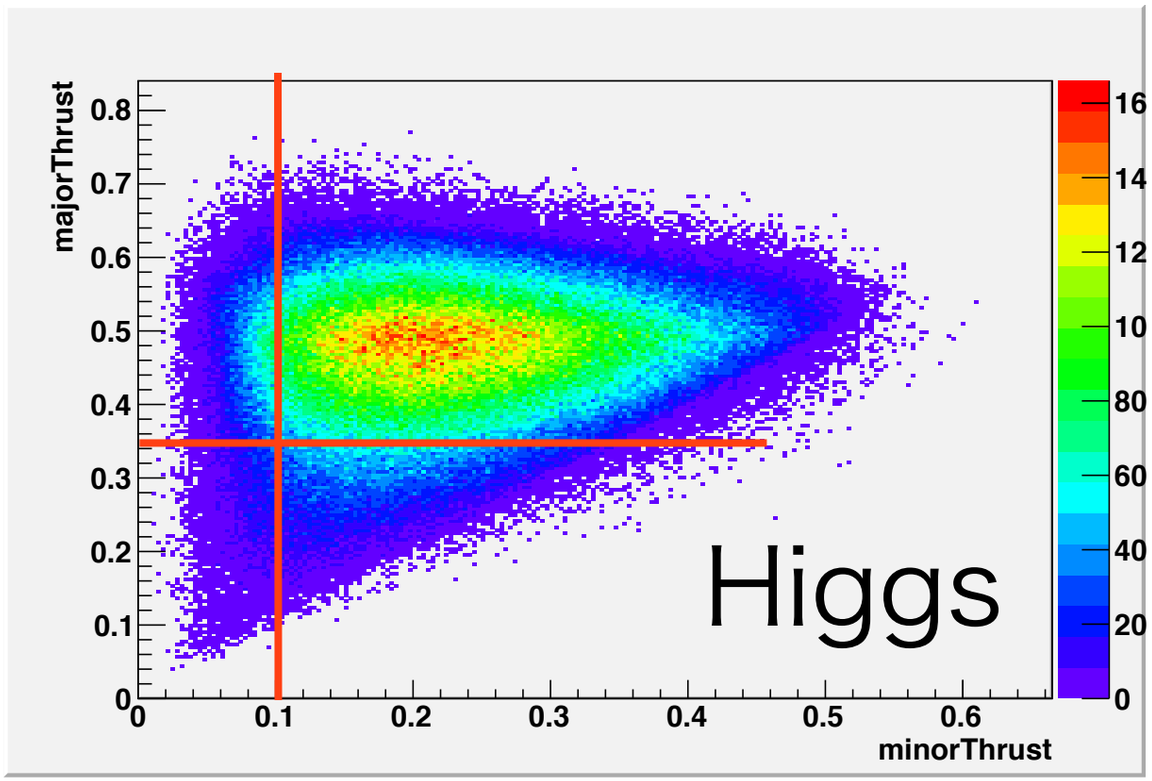


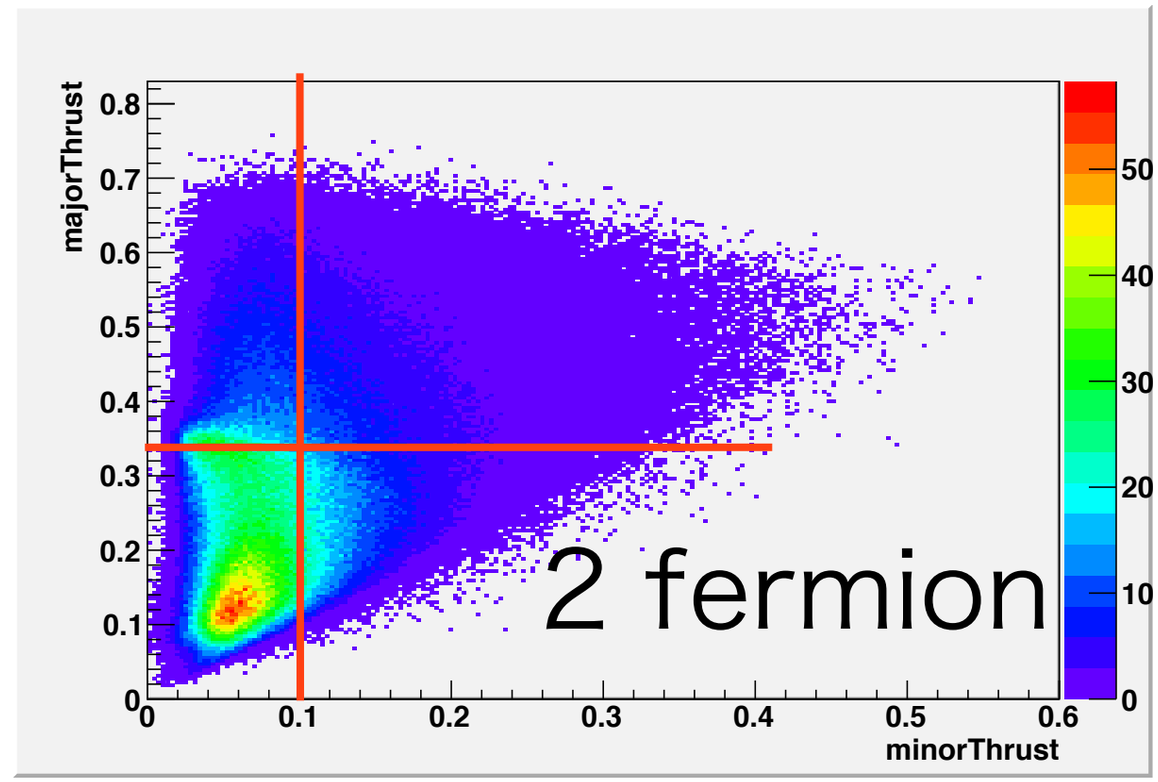




Cuts (3)

- Using major and minor Thrust value,
  Targets : 2 fermion backgrounds





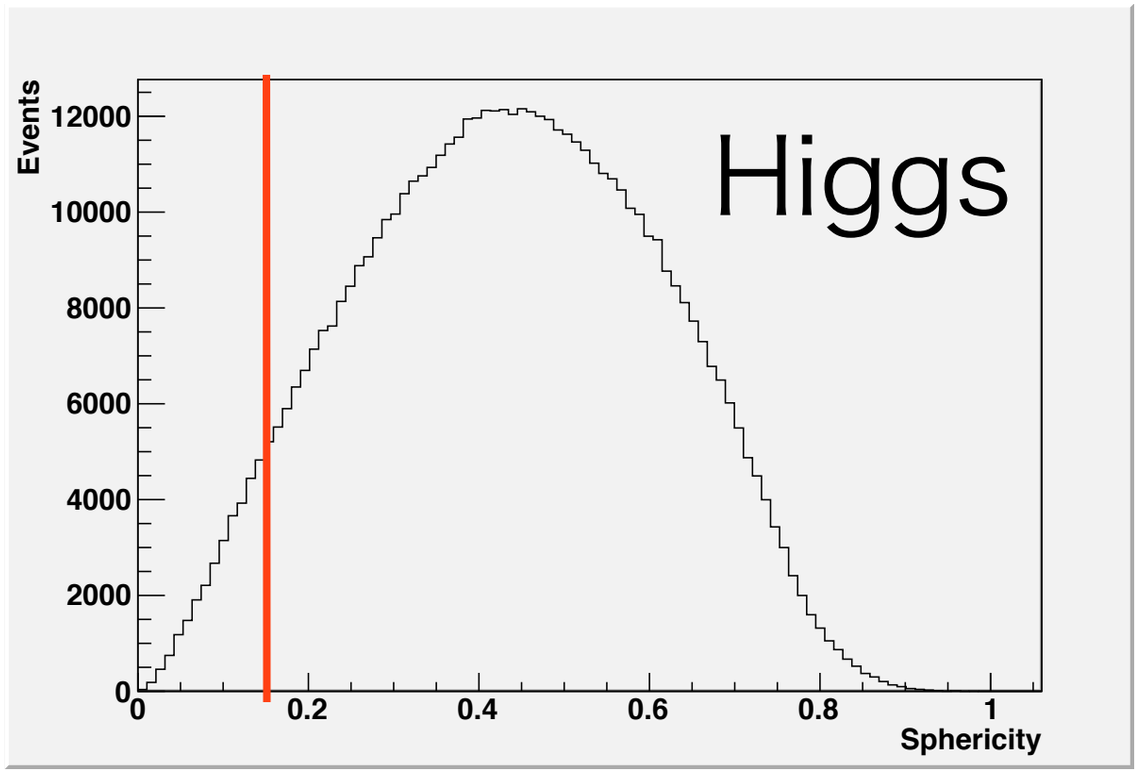
- Cut with majorThrust > 0.35, minorThrust > 0.1

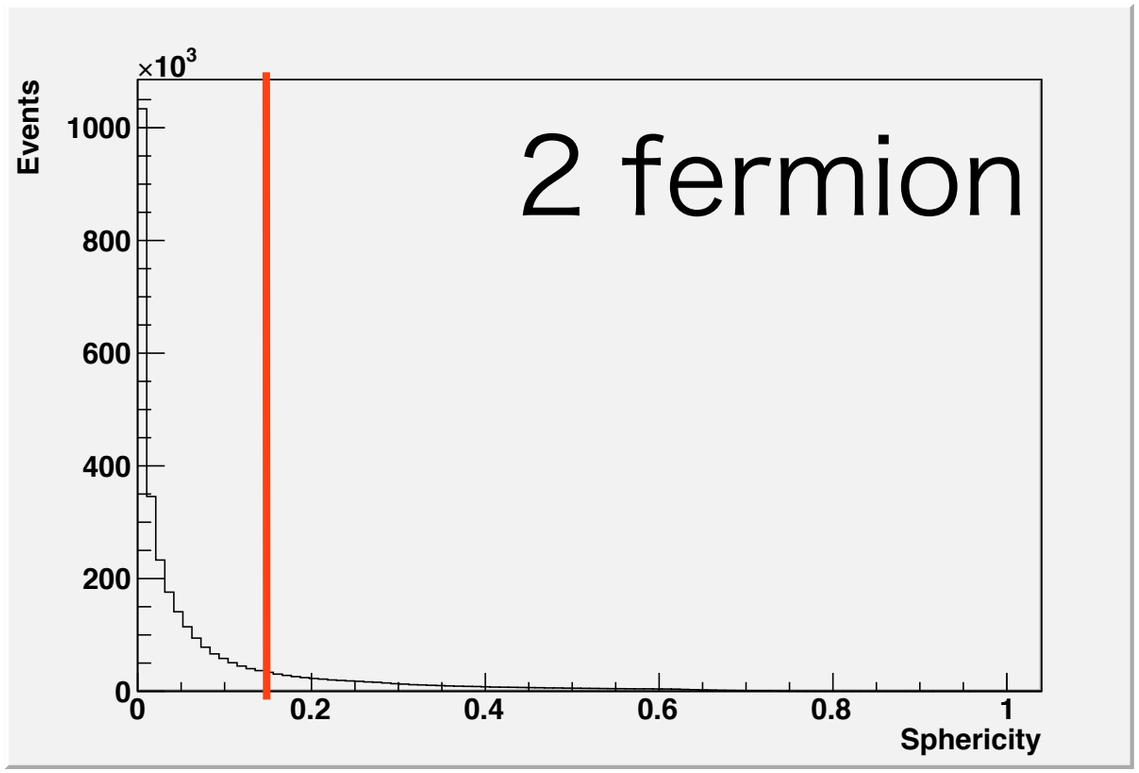
- back-to-back jets -> T~0
  not back-to-back jets -> T~1/2

$$T = \max_{|n|=1} \frac{\sum_i |\, p_i \cdot n \,|}{\sum_i |\, p_i \,|}$$

Cuts (4)

- Using Sphericity value,
  Targets : 2 fermion backgrounds





- Cut with Sphericity > 0.15

- Pencil like events -> S~0
  Spherical events -> S~1/2

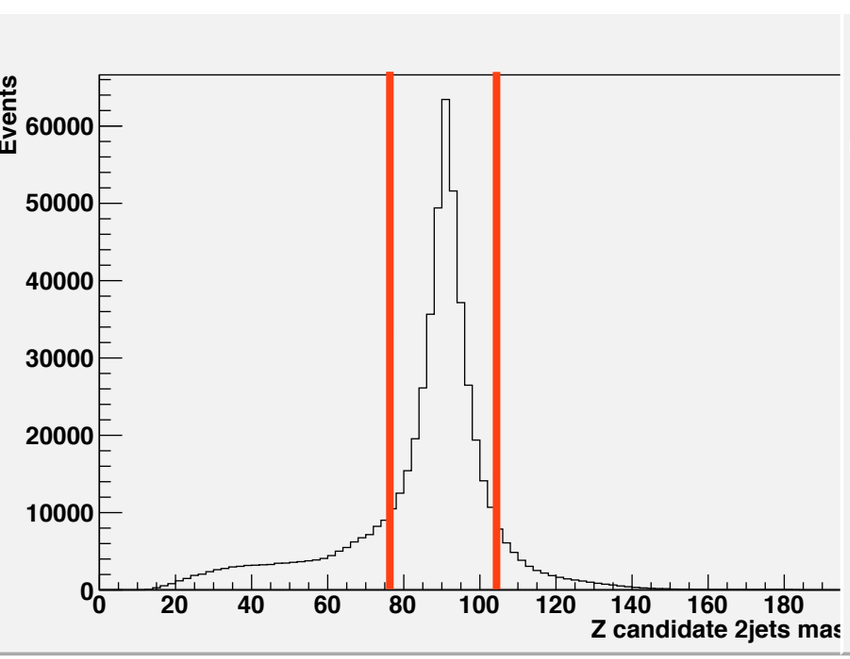
$$S^{ab} = \frac{\sum_i p_i^a p_i^b}{\sum_i p_i^2} \quad a, b = x, y, z$$

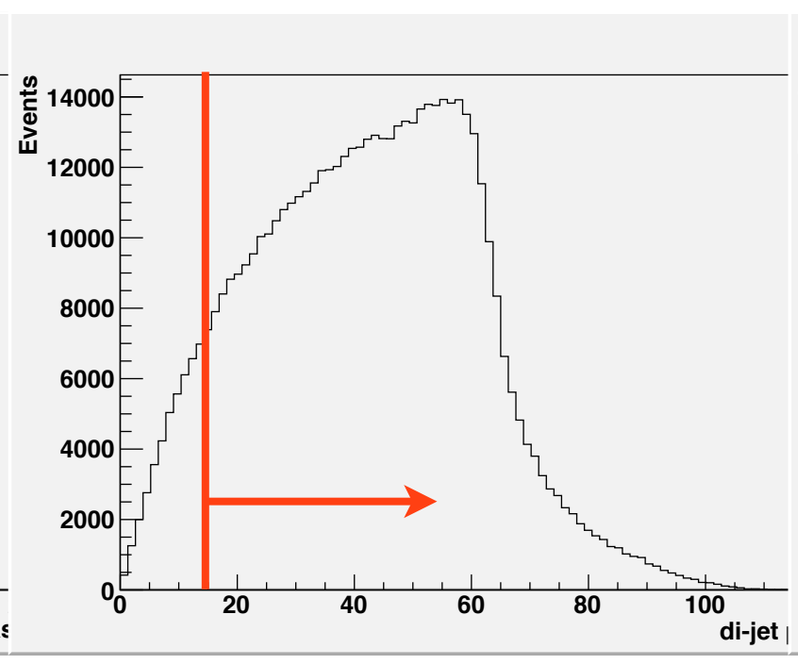
Selections

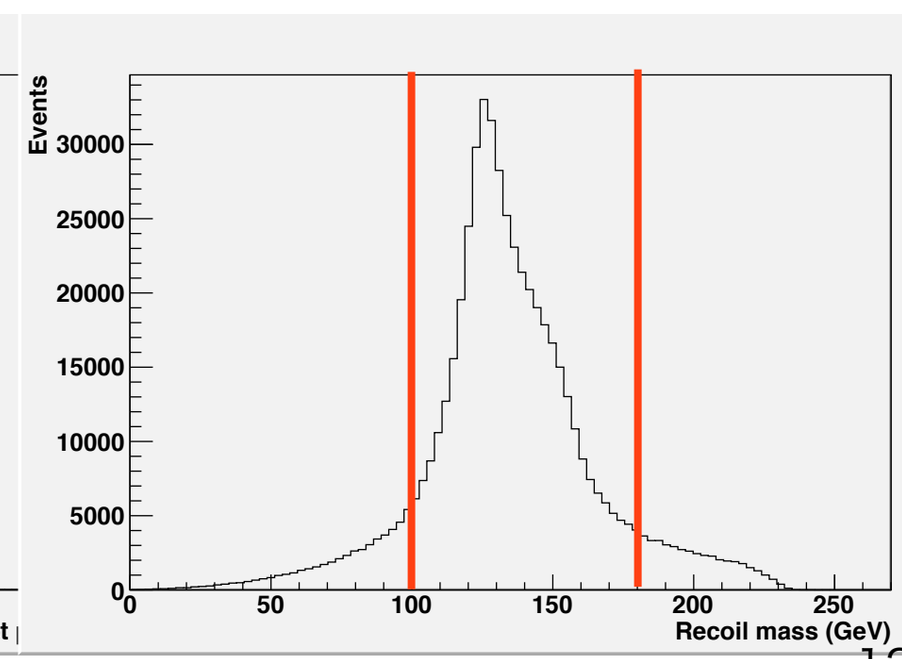
- Using y value clustering for recoil method, (0.0025 fixed)

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

- Z candidate di-jetmass selection (76,106) GeV
- Z candidate di-jetp_T selection > 15 GeV
- Recoil mass selection (100,180) GeV



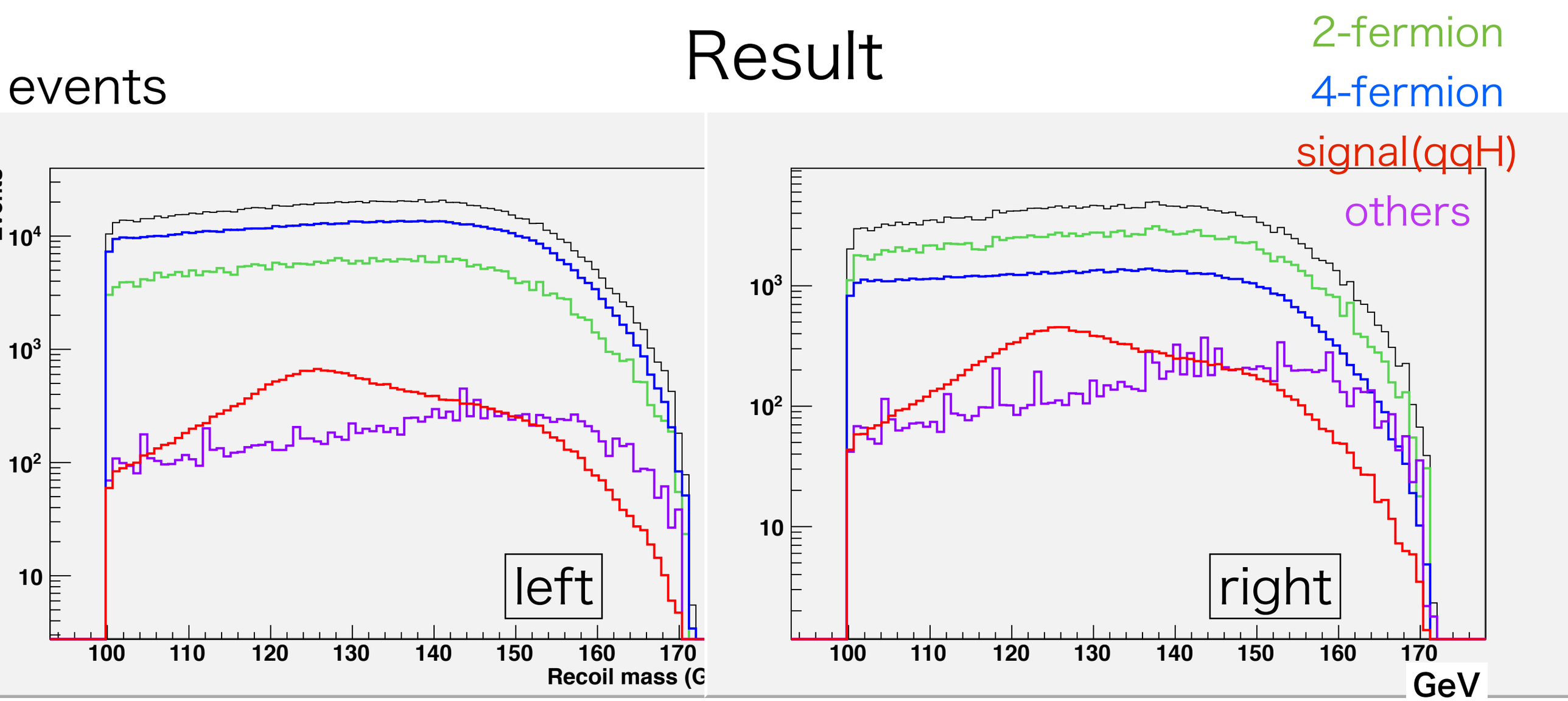




Cuts and Selections table

cuts	qqH	4 fermion	2 fermion	other
left	50,816	9,361,676	19,315,415	216,171,025
right	34,308	1,084,045	12,556,240	222,597,419
box	78.4%	41.1%	66.1%	96.0%
z pt	71.4%	26.1%	22.1%	2.5%
y dijet	53.7%	16.5%	6.3%	0.2%
recoil	49.3%	8.3%	2.2%	0.1%
sphericity	47.1%	6.9%	1.2%	0.06%
thrust	47.0%	6.8%	1.1%	0.05%

Result



significance was calculated in (110,150) GeV

polarization	significance	$\Delta\sigma/\sigma$
left (-0.8, +0.3)	21.5 σ	4.7%
right (+0.8, -0.3)	30.5 σ	3.3%

Model Independent ?

mode	before	after	difference / mean
H->all	549,279	258,159 (47.0%)	-----
H->bb	310,799	138,627 (44.6%)	-5.1%
H->WW(l)	12,622	1,575 (12.5%)	-73.4%
H->WW(sl)	53,607	27,843 (51.9%)	+10.4%
H->WW(h)	56,065	35,926 (64.1%)	+36.4%
H->gg	48,419	30,016 (62.0%)	+31.9%
H->$\tau\ \tau$	35,801	7,989 (22.3%)	-52.6%
H->ZZ	15,103	7,665 (50.8%)	+8.1%
H->cc	14,429	7,367 (51.1%)	+8.7%
H->$\gamma\ \gamma$	2,051	990 (48.3%)	+2.8%

Using 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.

Categories

category	0lep,0tau	0lep,1tau	0lep,≧2tau	1lep,0tau	1lep,≧1tau	≧2lep,≧0tau
H->all 549,279	81.6%	8.1%	2.7%	5.5%	1.3%	0.75%
H->bb 310,799	96.8%	2.8%	0.04%	0.33%	0.01%	~0.0%
H->WW(l) 12,622	8.3%	11.4%	6.9%	24.1%	26.3%	23.0%
H->WW(sl) 53,607	29.7%	19.8%	1.4%	45.4%	3.4%	0.2%
H->WW(h) 56,065	91.9%	7.2%	0.3%	0.5%	0.07%	0.0%
H->gg 48,419	96.6%	3.0%	0.06%	0.3%	0.01%	0.0%
H->$\tau\tau$ 35,801	12.2%	45.7%	35.4%	2.4%	4.2%	0.1%
H->ZZ 15,103	78.2%	8.4%	1.5%	3.2%	2.7%	6.0%
H->cc 14,429	96.3%	3.4%	0.05%	0.3%	0.01%	0.0%
H->$\gamma\gamma$ 2,051	91.3%	5.2%	0.5%	0.7%	0.5%	1.9%

Categories (example 0lep,0tau)

category	0lep,0tau	before	after	difference		w/o categories
H->all	81.6%	448,212	222,362 (49.6%)	----		-----
H->bb	96.8%	300,853	135,195 (44.9%)	-9.5%		-5.1%
H->WW(l)	8.3%	1,048	472 (45.0%)	-9.3%		-73.4%
H->WW(sl)	29.7%	15,921	7,968 (50.0%)	-0.8%		+10.4%
H->WW(h)	91.9%	51,524	33,443 (64.9%)	+30.8%		+36.4%
H->gg	96.6%	46,773	29,159 (62.3%)	+25.6%		+31.9%
H->$\tau\ \tau$	12.2%	4,368	1,658 (38.0%)	-23.4%		-52.6%
H->ZZ	78.2%	11,811	6,248 (52.9%)	+6.6%		+8.1%
H->cc	96.3%	13,895	7,155 (51.5%)	+3.8%		+8.7%
H->$\gamma\ \gamma$	91.3%	1,873	913 (48.7%)	-1.8%		+2.8%

Categories (example 0lep,0tau)

category		difference		w/o categories
H->all	The efficiencies approach mean value with categorization. Especially the event which has lepton and/or tau.	----		-----
H->bb		-9.5%		-5.1%
H->WW(l)		-9.3%		-73.4%
H->WW(sl)		-0.8%		+10.4%
H->WW(h)		+30.8%		+36.4%
H->gg	・Fully hadronic (can add cut)	+25.6%		+31.9%
H->$\tau\ \tau$	・include lepton (should optimize)	-23.4%		-52.6%
H->ZZ		+6.6%		+8.1%
H->cc	bias for gg and WW	+3.8%		+8.7%
H->$\gamma\ \gamma$		-1.8%		+2.8%

Cut optimization (0lep,0tau)

- Recoil mass (100,180) -> (110,150).
- the number of jets < 7
  (more than 7 jets means miss clustering)
- Z candidate 2 jet mass (76,106) -> (60,110)
  expand acceptance
- adjust box cut region.

Categories (example 0lep,0tau)

category	0lep,0tau	before	after	difference		w/o categories
H->all	81.6%	448,212	185,999 (41.5%)	----		-----
H->bb	96.8%	300,853	119,211 (39.6%)	-4.5%		-5.1%
H->WW(l)	8.3%	1,048	429 (40.9%)	+1.4%		-73.4%
H->WW(sl)	29.7%	15,921	5,618 (35.3%)	-14.9%		+10.4%
H->WW(h)	91.9%	51,524	23,773 (46.1%)	+11.1%		+36.4%
H->gg	96.6%	46,773	23,636 (50.5%)	+21.7%		+31.9%
H->$\tau\ \tau$	12.2%	4,368	1,362 (31.2%)	-24.8%		-52.6%
H->ZZ	78.2%	11,811	4,766 (40.4%)	-2.7%		+8.1%
H->cc	96.3%	13,895	6,284 (45.2%)	+8.9%		+8.7%
H->$\gamma\ \gamma$	91.3%	1,873	793 (42.3%)	-1.9%		+2.8%

Categories (example 0lep,0tau)

category		difference		w/o categories
H->all		----		-----
H->bb		-4.5%		-5.1%
H->WW(l)		+1.4%		-73.4%
H->WW(sl)		-14.9%		+10.4%
H->WW(h)		+11.1%		+36.4%
H->gg		+21.7%		+31.9%
H->$\tau\ \tau$		-24.8%		-52.6%
H->ZZ		-2.7%		+8.1%
H->cc		+8.9%		+8.7%
H->$\gamma\ \gamma$		-1.9%		+2.8%

After simple optimization, the difference is reduced effectively.

WW hadronic, gluglu need more effective cut.

NOTE
Do not bundle bb/cc/gg !

Its cut efficiency must be different without consideration.

Summary and Prospects

summary

- The precision of total cross section
  left 5.6% -> 4.7%, right 4.0% -> 3.3% from AWLC.
  (but still not satisfactory.)
- Categorization can reduce difference of efficiency
  especially tautau, WW->leptonic.
- Current cut has bias for gg and WW.

prospects

- Use likelihood to improve statistical precision.
- Equalize the cut efficiency of each Higgs decay mode.
- Improve tau separation by optimizing tau finder.
- Estimate systematic errors.

backup

strategy for resolving efficiency issue -1

To resolve efficiency inconsistent issue, we will categorize events using

- the number of jets (2, 3, 4, and >= 5)
- 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 into 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

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

strategy for resolving efficiency issue -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

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

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 ${\rm BR}_n \cdot \theta_n^i \cdot \frac{\delta\epsilon_n^i}{\epsilon^i}$ << 1 %.

TaJet finder (1)

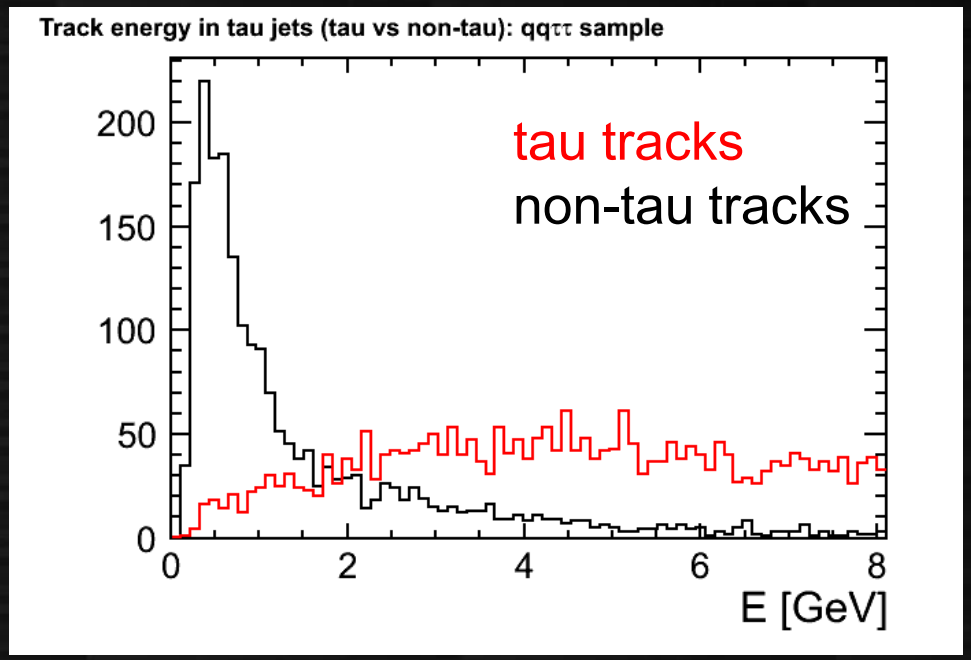
High-purity tau tagging
in presence of jet background

1. Order charged tracks by largest energy
2. Select the first track
3. Combine neighboring particles -> "Tau Jet"
   - Combined mass < 2 GeV && cosθ w.r.t. jet axis > 0.98
4. Tau selection (tuned for rejecting qq background)
   1. Tau Jet energy > 3 GeV
   2. Veto >=3 prong + neutrals (> 1 GeV)
   3. Cone energy ($E_{cone} < 0.1E_{taujet}$) with $\cos\theta_{cone} = 0.9$

ZZ -> qq$\tau\tau$ 250 GeV, 13600 taus	1-prong		3-prong wo/ neutral		3-prong w/ neutral	
	tau	non-tau	tau	non-tau	tau	non-tau
No cut	10326	43286	716	1616	777	4280
$E_{taujet} > 3$	8679	7145	708	1304	742	4244
$E_{cone} < 0.5E_{taujet}$	7170	1009	621	181	681	1813
$E_{cone} < 0.2E_{taujet}$	6455	446	567	64	616	1020
$E_{cone} < 0.1E_{taujet}$	6001	254	527	30	570	620

TaJet finder (2)

5. Jet charge recovery (for better efficiency)
   - Tracks with energy < 2 GeV are detached one by one until tau jet has 1 or 3 tracks and sum charge is +1 or -1
   - Jet is rejected if above condition cannot be satisfied after detaching all < 2 GeV tracks
6. Return to 2. (previous page) with the remaining tracks
   - Stop after all E > 2 GeV tracks have been processed



# tau jets	qqττ	qqln
0	27.1%	47.6%
1	36.3%	46.6%
2	34.0%	5.4%
>3	2.4%	0.3%

purity of tau in qqττ:
94.2% overall
96.5% for # tau jets == 2