Higgs Branching Ratio
H→bb, cc, gg
for new detector models at ILD

Masakazu Kurata, Ryo Yonamine, Hiroaki Ono

LCWS2018

10/24/2018

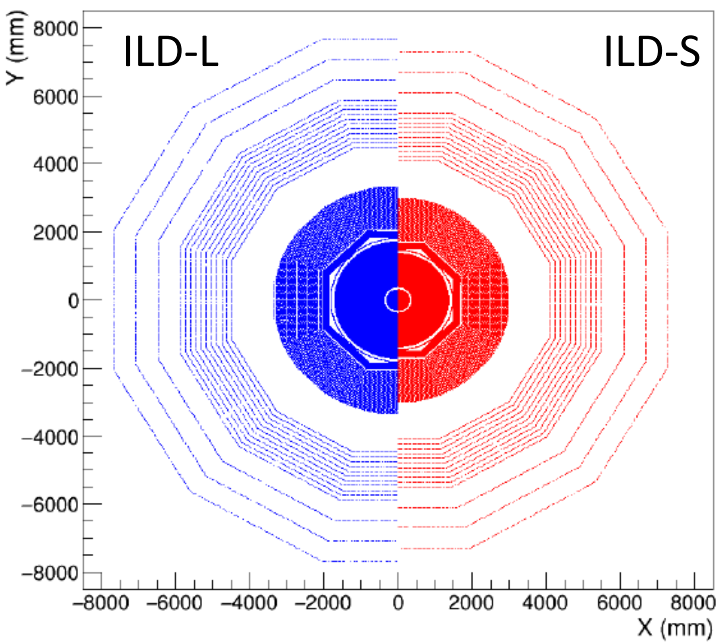
Introduction

- Start to re-optimize new detector models using physics benchmarks
- Higgs analyses are one of the most important benchmark processes for detector optimization at ILD
- H→bb, cc, gg measurement:
  - Flavor tagging is indispensable for jet flavor separation
  - Try@500GeV, focus on ννH process
  - Based on Hiroaki's 1TeV analysis in DBD era

  - So far, do not consider separation of ZH and VBF
    - But, considering it for future plan

Optimization with 2 detector models

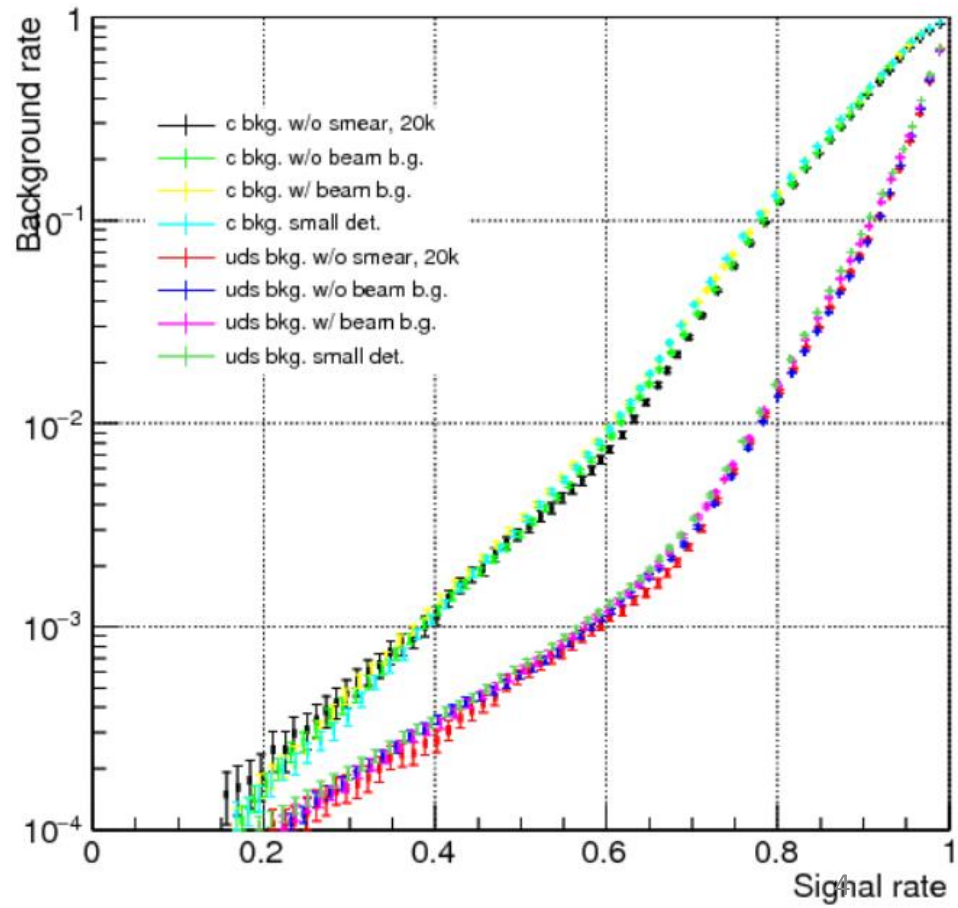
- Re-optimize ILD detector
  - Revisit optimization of cost and detector performance



Detector models	ILD-L	ILD-S
B-field	3.5T	4T
VTX inner radius	1.6cm	1.6cm
TPC inner radius	33cm	33cm
TPC outer radius	180cm	146cm
TPC length (z/2)	235cm	235cm
Inner ECAL radius	184cm	150cm
Outer ECAL radius	202.5cm	168.5cm
Inner HCAL radius	206cm	172cm
Outer HCAL radius	335cm	301cm
Coil inner radius	344cm	310cm

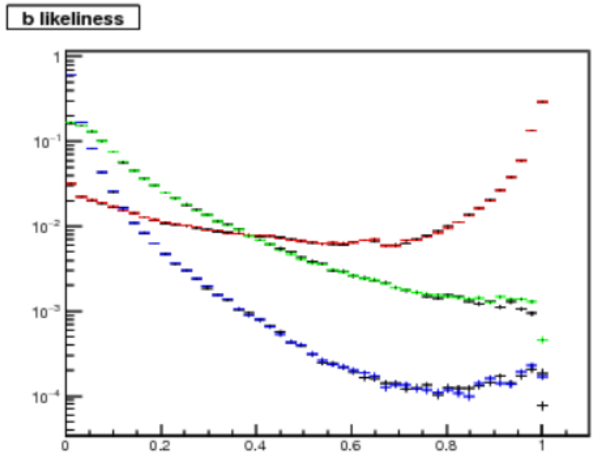
LCFIPlus summary: Comparison among any situation

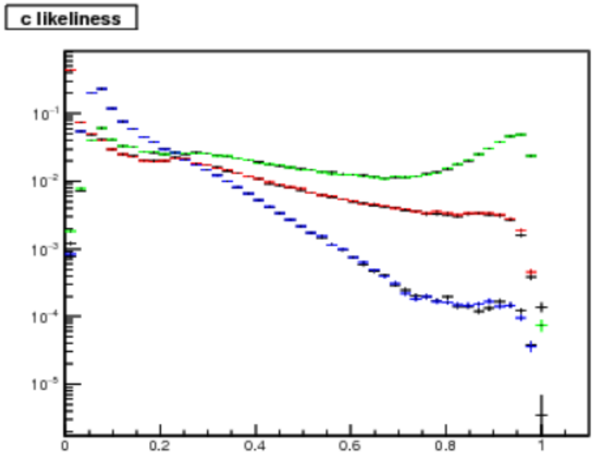
- Ryo made great efforts to resolve internal problems in LCFIPlus for new verson of iLCsoft
- Use 6f flavor tag samples: bbbbbb, cccccc, qqqqqq
- Tune MVA training to recover performance to those of DBD era
- Almost same performance as DBD era
- Over training?

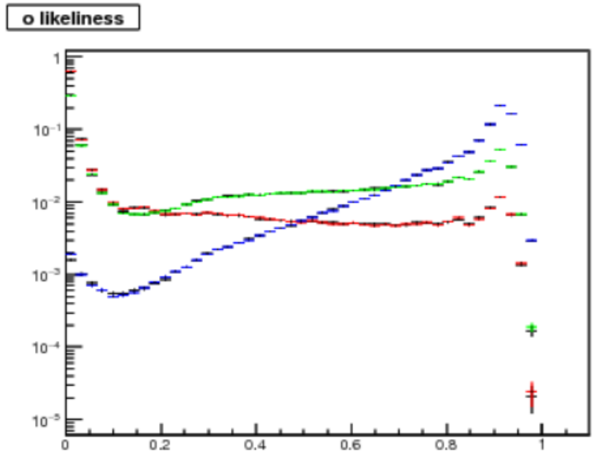


Over training check

- Check MVA output: LCFIPlus provides 3 MVA outputs
  - w/ beam background







Black dots: training
Color dots: test
red: b jets
green: c jets
blue: light flavor

Log scale

No over training

Status of analysis

- E_{CM}=500GeV
- Luminosity: $4ab^{-1}$

Polarization	(e-, e+)=(-,+)	(+,-)	(+,+)	(-,-)
Luminosity(fb^{-1})	1800	1800	200	200

- Analysis flow

Isolated Lepton ID
Look for electron and muon

Jet Clustering
Durham with beam background rejection
2 jet clustering

Event selection
Reject backgrounds
Preselection
MVA

- Signal and backgrounds
  - Signal: use nnH→nnbb, nncc, nngg
  - Backgrounds: 2f, 4f, 5f, 6f, aa, ZH, nnh→nnWW

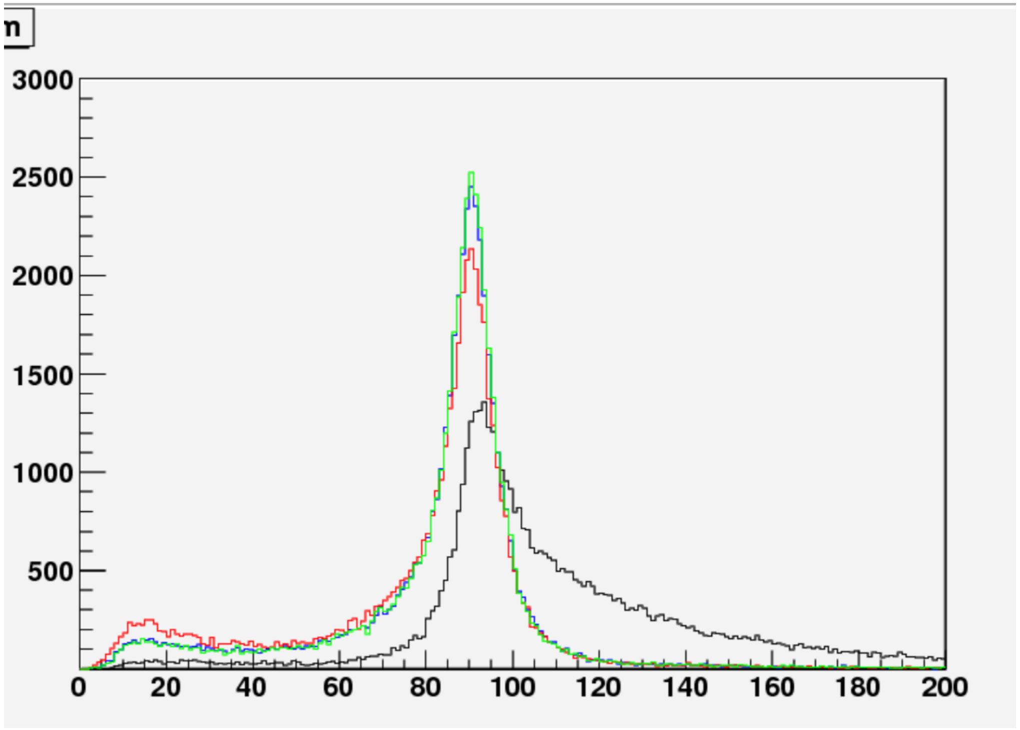
Beam background rejection

- $y_{ij} = \frac{\min(E_i^2, E_j^2)(1-\cos\theta)}{E_{vis}^2}$, $y_{beam} = \frac{2E_i^2\alpha^2(1-\cos\theta)}{E_{vis}^2}$

  α: beam rejection parameter

  smaller→ beam rejection becomes stronger

  - Particle i with $y_{ij}>y_{beam}$ is discarded



νν̄Z@500GeV(DBD)

- 2 jet clustering
- Parameters are tuned for better result

w/o beam b.g. rejection

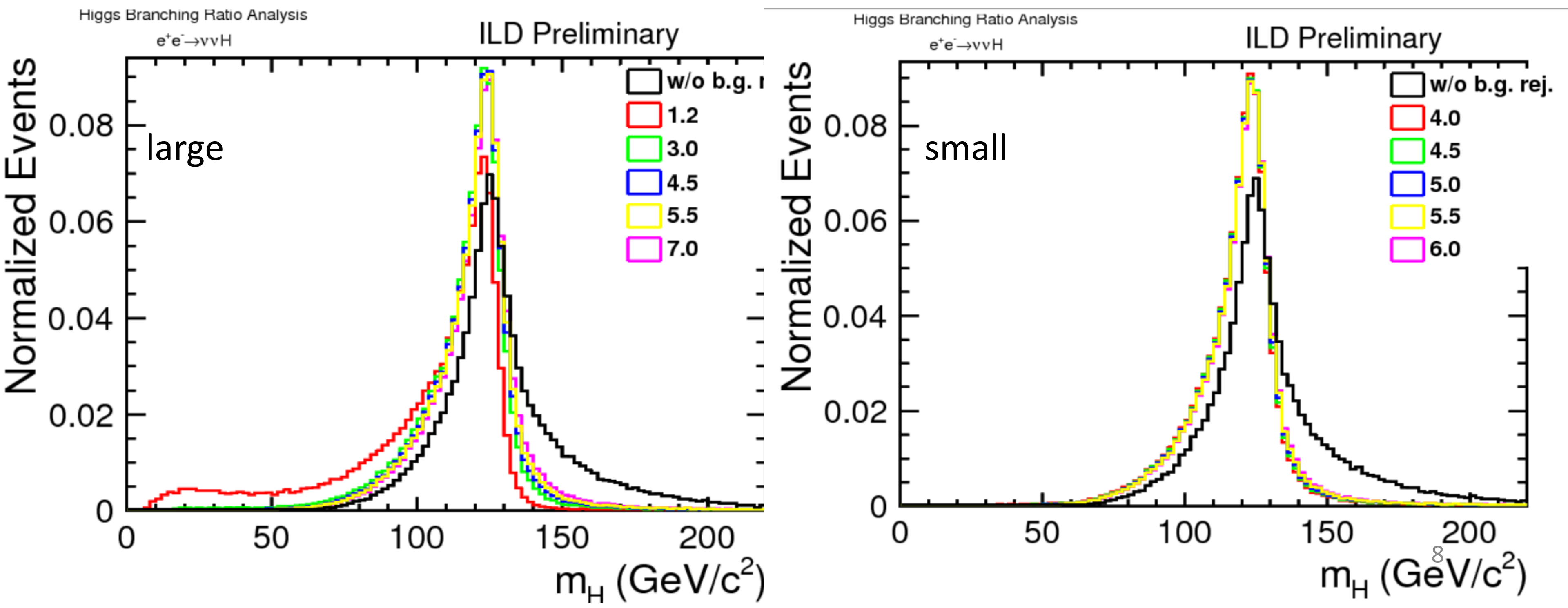
Kt

Durham

Valencia

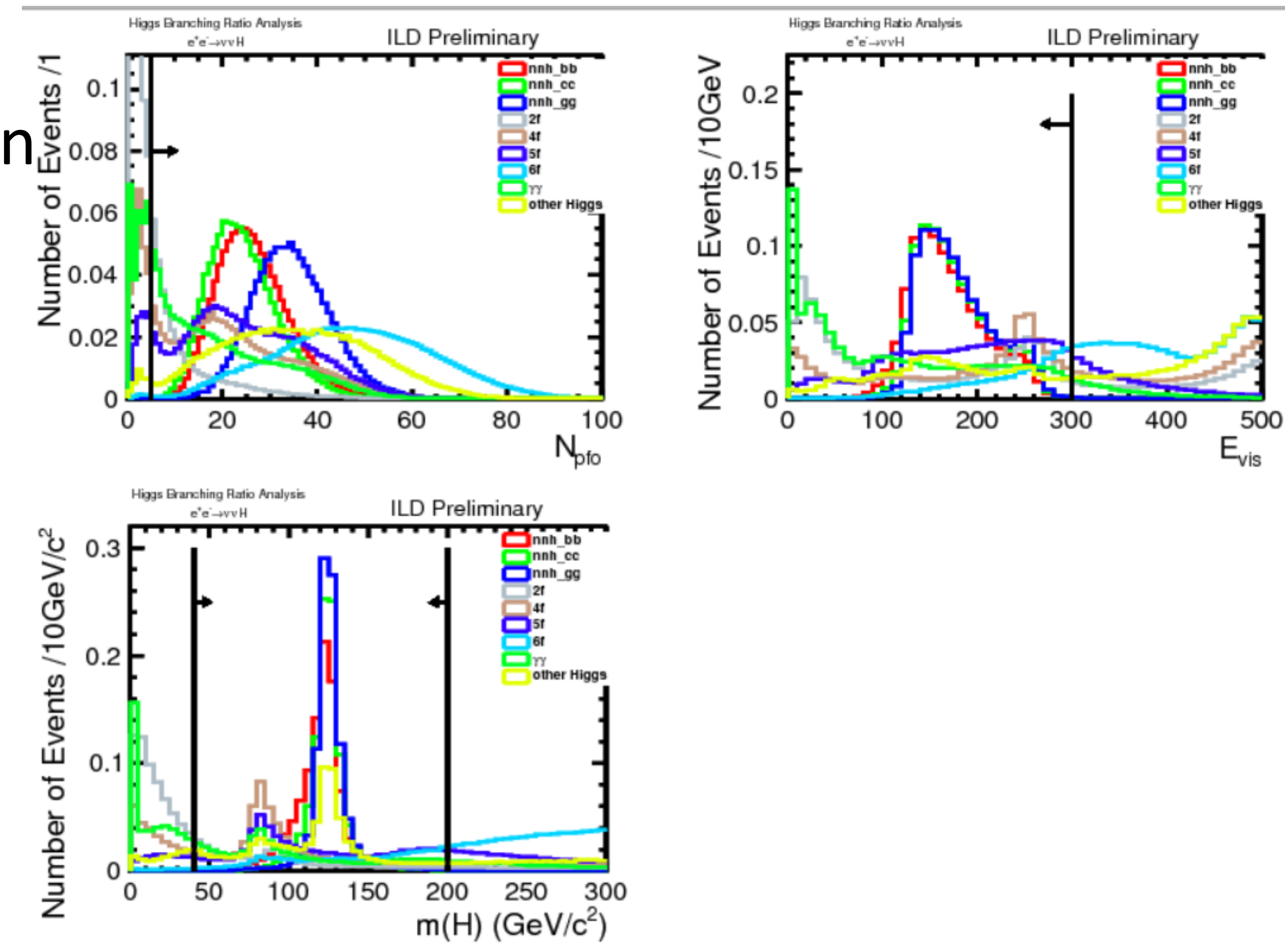
Apply to nnH events

- Apply Durham beam b.g. rejection to nnH→nnbb events
- Parameter scan to make Higgs mass distribution better
  - 5.5 seems best for both detector models



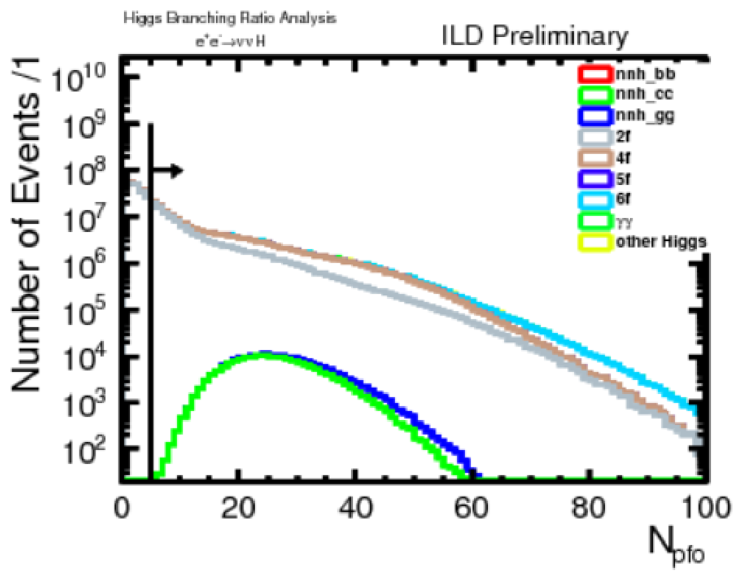
Preselection

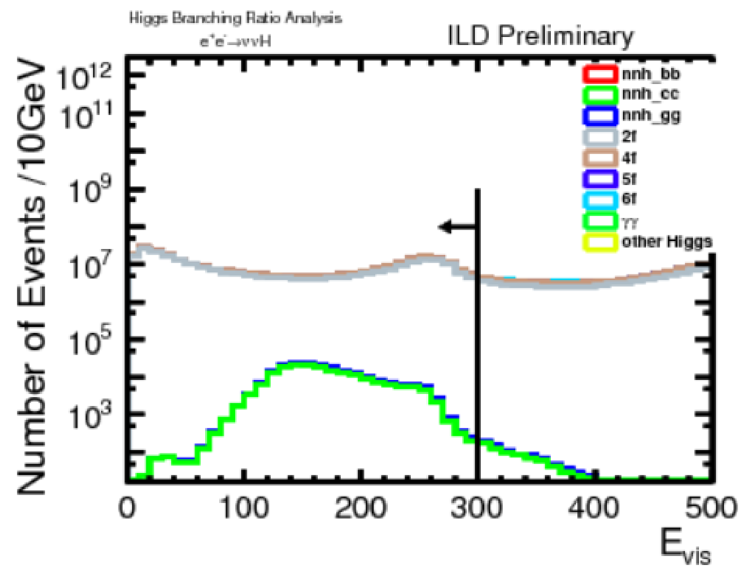
- Just reject backgrounds which are trivial ones
  - Signal tail part of each variable is rejected

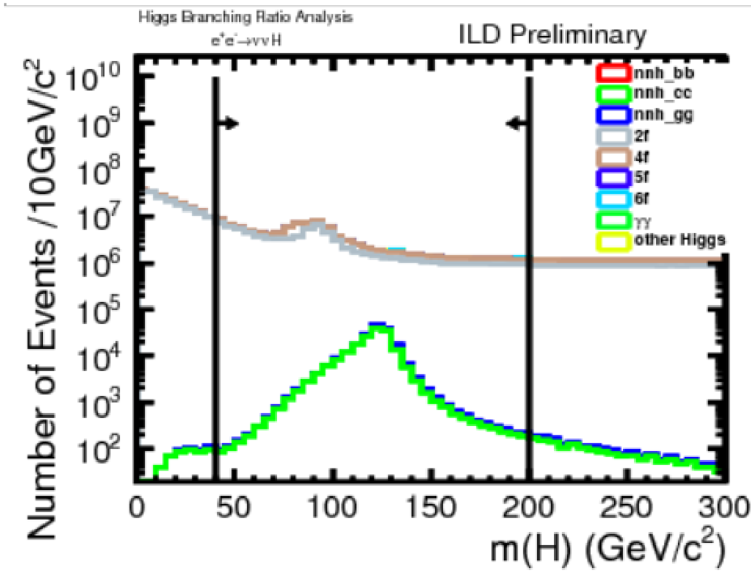


- No Isolated lepton
- $N_{jets}=2$
- $N_{pfo} \geqq 5$
- $E_{vis} \leqq 300$
- $40 \leqq m_{2jets} \leqq 200$
- Reject NaN

Stacked plots







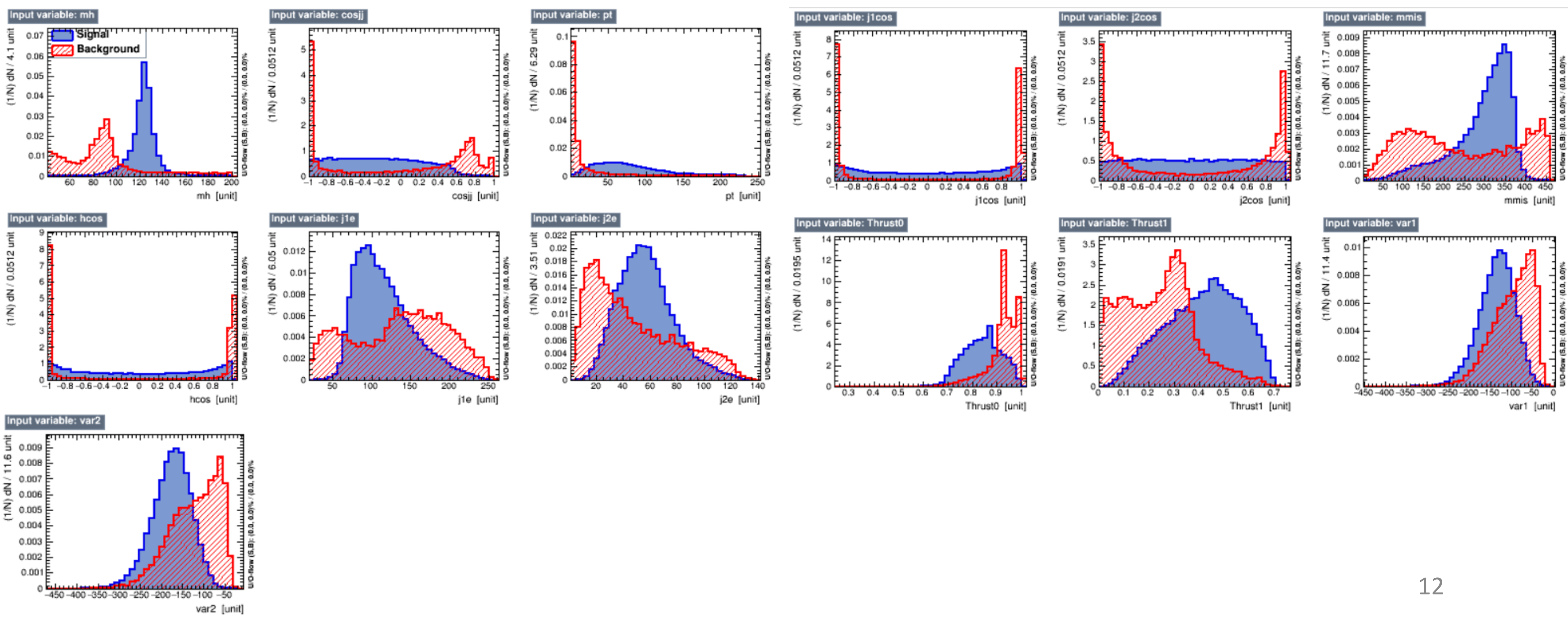
Multivariate Analysis

- Expect better background rejection efficiency than cut based
- Need to reject other Higgs processes as much as SM backgrounds

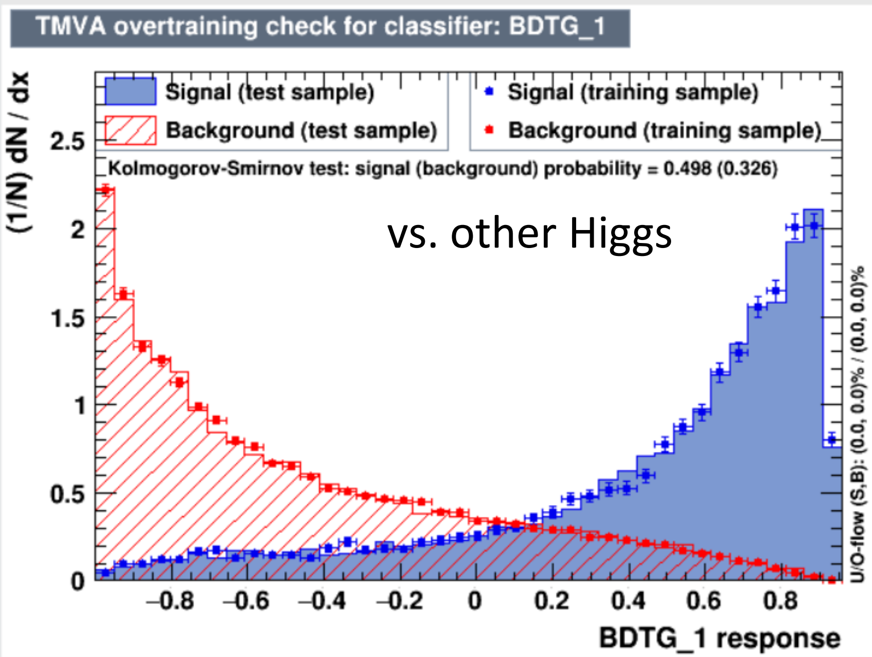
- Important to separate both SM backgrounds and other Higgs processes for backgrounds
- Use Binary Classification
  - Due to the phase space difference of each background component
  - Signal vs.: other Higgs, 2f&4f, 5f&6f&aa
  - 3kinds of classifier trained

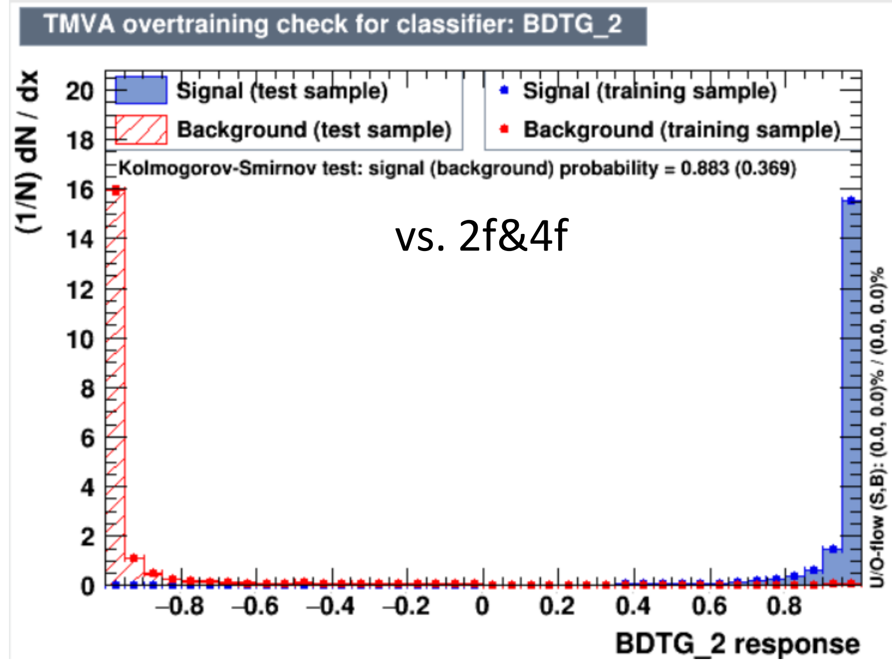
Input variables

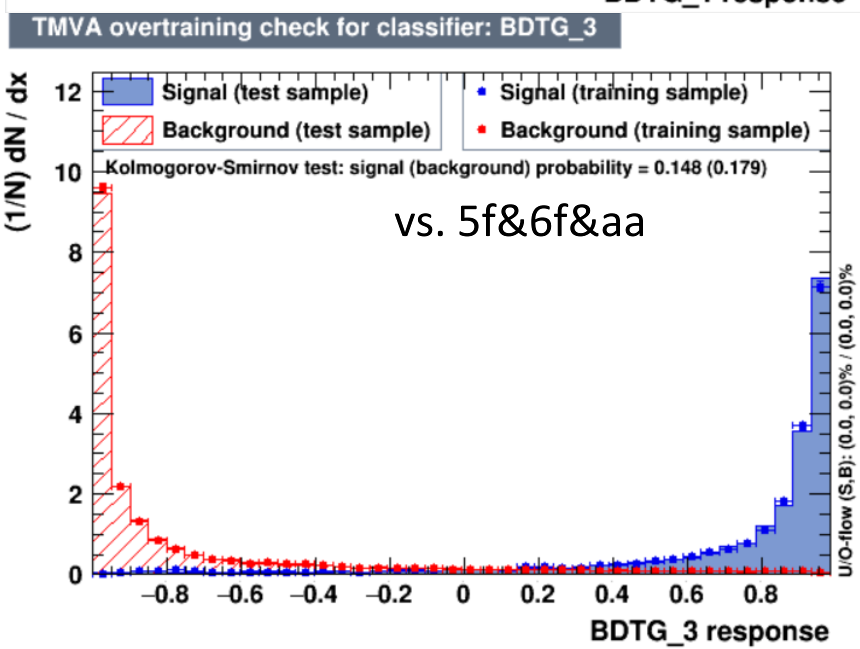
- Need to suppress bias between H→bb,cc and H→gg
- Reduce difference by combining variables
  - m_{2jets}, $\cos\theta_H$, $\cos\theta_{jj}$, E_{j1}, $\cos\theta_{j1}$, E_{j2}, $\cos\theta_{j2}$, m_{miss}, Pt, Principal Thrust, Major Thrust, Log(y23)*n_{pfo}, Log(y34)*n_{pfo}
- Example: signal vs. 2f&4f



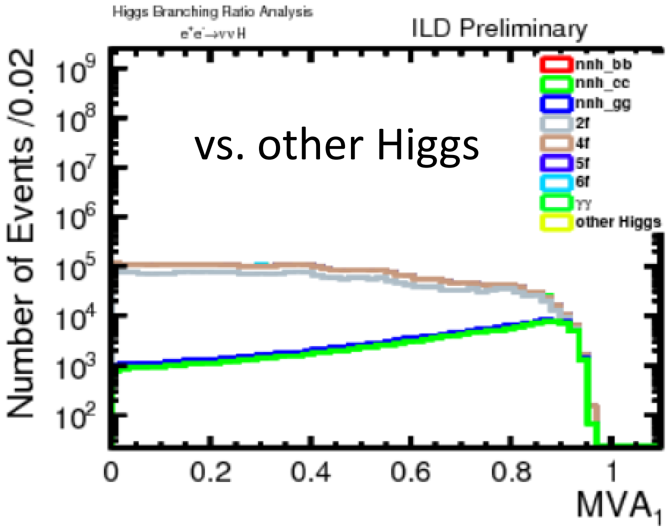
MVA output

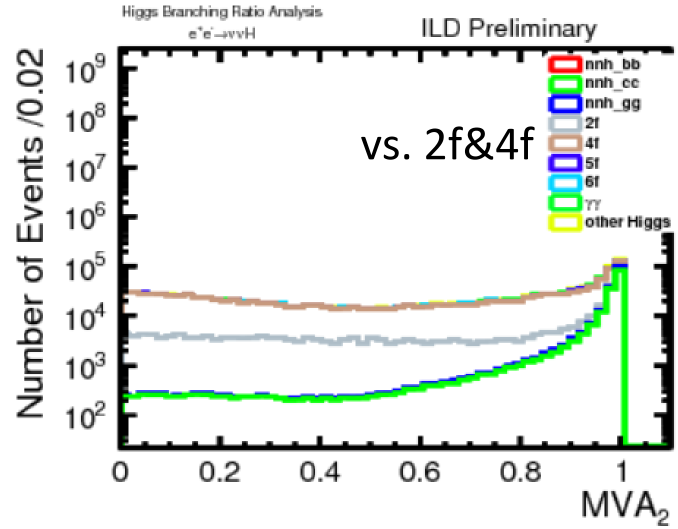

TMVA overtraining check for classifier: BDTG_1
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability = 0.498 (0.326)
vs. other Higgs
(1/N) dN / dx
BDTG_1 response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%

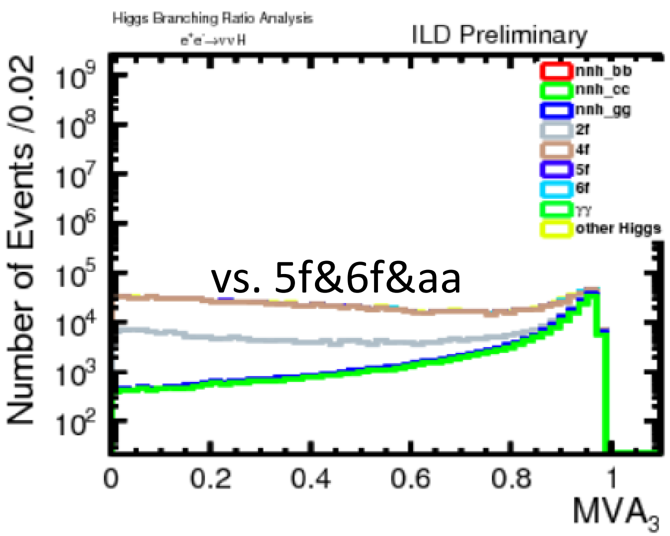

TMVA overtraining check for classifier: BDTG_2
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability = 0.883 (0.369)
vs. 2f&4f
(1/N) dN / dx
BDTG_2 response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%


TMVA overtraining check for classifier: BDTG_3
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability = 0.148 (0.179)
vs. 5f&6f&aa
(1/N) dN / dx
BDTG_3 response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%

Stacked







- Determine operation point:
  MVA_1>x.xx && MVA_2>y.yy && MVA_3>z.zz

Cut table

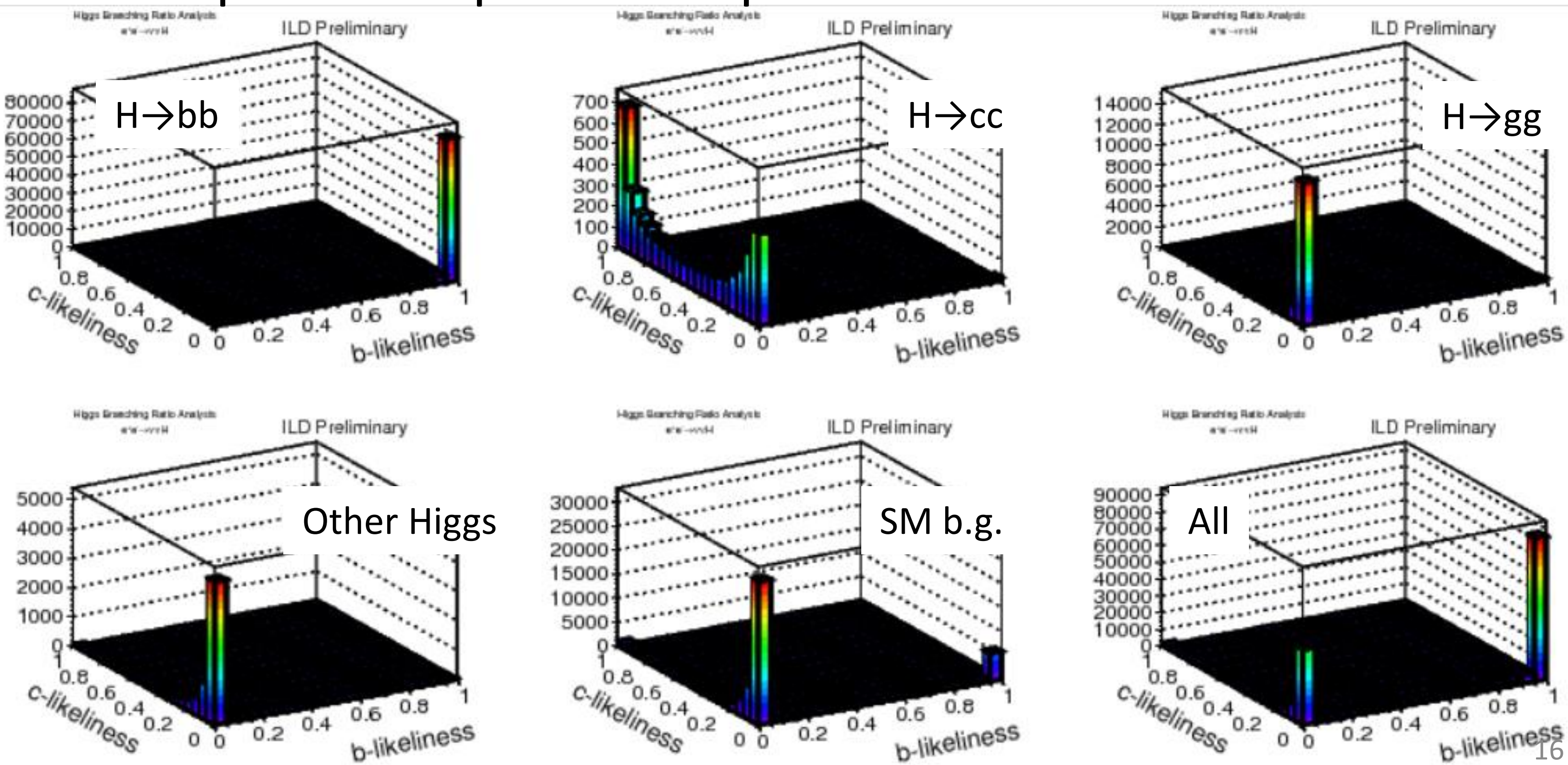
- (e-, e+)=(-,+) polarization, $L=1800fb^{-1}$

process	H→bb	H→cc	H→gg	2f	4f	5f	6f	γγ	Other Higgs
No cut	176726	8202	26189	3.67e+08	6.29e+07	129014	1.77e+06	457688	226855
$N_{lep}=0$	171857	8008	25932	2.90e+08	4.25e+07	41116	1.04e+06	282545	160037
$N_{jet}=2$	171857	8008	25932	2.84e+08	4.23e+07	41101	1.04e+06	267152	160024
$N_{pfo} \geqq 5$	171167	7969	25802	1.87e+08	2.01e+07	29384	116391	243952	54780
$E_{vis} \leqq 300$	169450	7898	25554	4.84e+07	1.16e+07	19510	44885	101863	46512
$40 \leqq m_{2jets} \leqq 200$	169448	7898	25554	3.12e+07	9.09e+06	17927	43969	92427	46182
MVA	105697	4956	17736	2650.2	49601	207.6	1197.1	538.1	9741.5
Efficiency	0.598	0.604	0.677	7.22e-06	7.89e-04	1.61e-03	6.76e-04	1.18e-03	0.043
Hiroaki's 1TeV	0.350	0.373	0.359						

Significance: 292.9

Template fit

- Toy MC to extract the measurement precision of H→bb, cc, gg
  - Using flavor tag templates for fit according to Poisson statistics
  - 3-D template of b, c, bc likeliness $$\mathrm{N}_{\mathrm{ijk}}^{\mathrm{template}} = \sum_{\mathrm{s=bb,\ cc,\ gg}} \mathrm{r_s} \cdot \mathrm{N}_{\mathrm{ijk}}^{\mathrm{s}} + \mathrm{N}_{\mathrm{ijk}}^{\mathrm{bkg}},$$
  - 5000 pseudo-experiments performed



Very preliminary results so far

- Large detector, all the polarization

Process(-,+)	H→bb	H→cc	H→gg
Precision(%)	0.39	3.59	1.59

Process(+,-)	H→bb	H→cc	H→gg
Precision(%)	0.98	10.0	4.61

Process(-,-)	H→bb	H→cc	H→gg
Precision(%)	1.35	15.5	6.73

Process(+,+)	H→bb	H→cc	H→gg
Precision(%)	2.68	29.8	12.4

- Scale to 500fb^{-1} (3000exp.)

Process(-,+)	H→bb	H→cc	H→gg
Precision(%)	0.70	6.73	2.88

TDR

Process(-,+)	H→bb	H→cc	H→gg
Precision(%)	0.60	5.2	5.0

- Small detector, all the polarization

Process(-,+)	H→bb	H→cc	H→gg
Precision(%)	0.38	3.42	1.67

Process(+,-)	H→bb	H→cc	H→gg
Precision(%)	0.96	10.0	4.34

Process(-,-)	H→bb	H→cc	H→gg
Precision(%)	1.47	13.8	5.84

Process(+,+)	H→bb	H→cc	H→gg
Precision(%)	2.69	29.4	12.5

Summary & Prospects

- We can get the results of both detectors & all the polarizations
  - Still preliminary, but small detector seems better performance than large???
- Need to optimize more for better results
  - Check why Small det. is better than Large
  - Need to investigate templates for fitting
    - Num. of bins
    - Variable bin size necessary
- Go to note for ILD Design Report
  - Basically, follow Hiroaki's 1TeV note in DBD era

Backups

Significance

	(-,+)	(+,-)	(-,-)	(+,+)
Large	292.9	101.3	69.5	36.3
Small	290.8	104.8	70.2	36.8

- Even if better significance, precision is worse…
- Large(-,+)

process	H→bb	H→cc	H→gg	2f	4f	5f	6f	γγ	Other Higgs
MVA	105697	4956	17736	2650.2	49601	207.6	1197.1	538.1	9741.5
Efficiency	0.598	0.604	0.677	7.22e-06	7.89e-04	1.61e-03	6.76e-04	1.18e-03	0.043

- Small(-,+)

process	H→bb	H→cc	H→gg	2f	4f	5f	6f	γγ	Other Higgs
MVA	106989	4999	17441	2428.03	55239	245.2	1298.3	756.6	9178.4
Efficiency	0.605	0.609	0.665	6.61e-06	8.78e-04	1.90e-03	7.38e-04	1.70e-03	0.040

Apply weight file from small to large

- Apply MVA weight files of small det. to large det.

Process(-,+)	H→bb	H→cc	H→gg
Precision(%)	0.40	3.53	1.64

- nominal

Process(-,+)	H→bb	H→cc	H→gg
Precision(%)	0.39	3.59	1.59

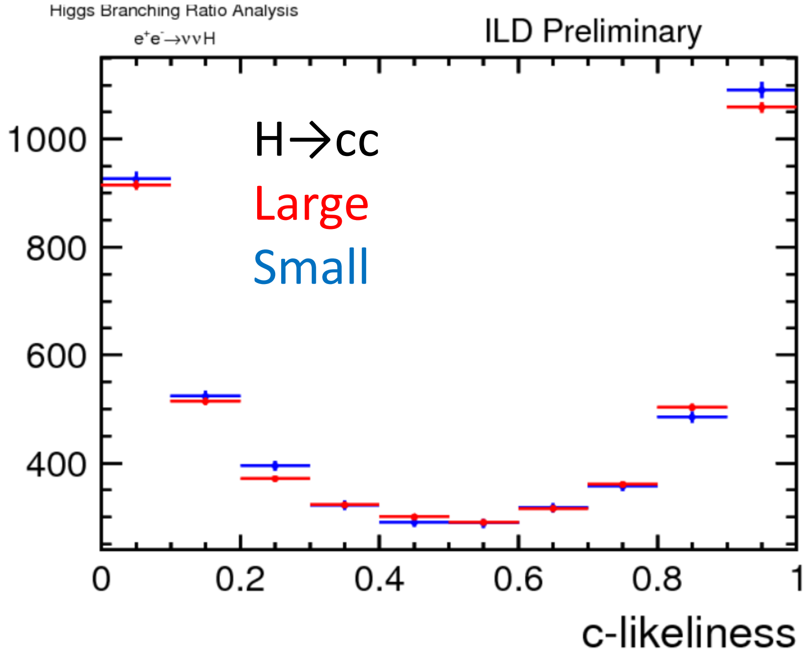
Num. of pseudo experiments

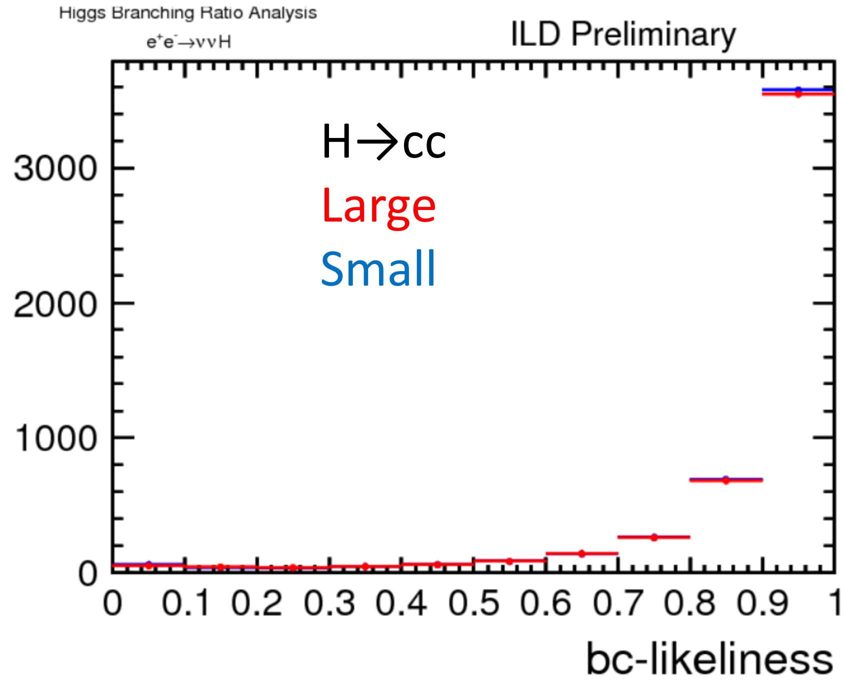
Large, (-,+)	5000	10000	20000
H→bb	0.39	0.40	0.43
H→cc	3.59	3.56	3.61
H→gg	1.59	1.60	1.60

Small, (-,+)	5000	10000	20000
H→bb	0.38	0.40	0.41
H→cc	3.42	3.39	3.37
H→gg	1.47	1.66	1.66

- In small det., H→cc systematically better than large

Shape of H→cc


Higgs Branching Ratio Analysis
e+e-→ννH
ILD Preliminary
H→cc
Large
Small
1000
800
600
400
0
0.2
0.4
0.6
0.8
1
c-likeliness


Higgs Branching Ratio Analysis
e+e-→ννH
ILD Preliminary
H→cc
Large
Small
3000
2000
1000
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
bc-likeliness

Impact of flavor tagging on precision measurement

- We do the pseudo experiments with (artificially) perfect & pessimistic case
  - Other operation point?

- Artificial efficiency
  - Perfect

input/output	b jet	c jet	l jet
b jet	1.0	0.0	0.0
c jet	0.0	1.0	0.0
l jet	0.0	0.0	1.0

Pessimistic

input/output	b jet	c jet	l jet
b jet	0.50	0.25	0.25
c jet	0.25	0.50	0.25
l jet	0.25	0.25	0.50

- Do the same procedure for the precision measurement

Comparison between pseudo experiments

- Large detector, (-,+) polarization, $L=1800fb^{-1}$

- Perfect

Process(-,+)	H→bb	H→cc	H→gg
Precision(%)	0.38	1.53	1.74

- LCFIPlus

Process(-,+)	H→bb	H→cc	H→gg
Precision(%)	0.39	3.59	1.59

- Pessimistic

Process(-,+)	H→bb	H→cc	H→gg
Precision(%)	0.92	12.2	6.04

- Better flavor tagging is very important for better precision
  - Especially, better c-tagging is important