



Higgs hadronic branching ratio studies for LC and FCCee (incl. PFA performance with ILD-FCC)

Ryu Sawada, ICEPP, University of Tokyo
presenting for

Takumi Seino (UTokyo) : Higgs couplings measurement at ILC

Kanako Watanabe (Iwate U) : $H \rightarrow bs$ search at ILC

Nagomi Dogishi (UTokyo) : $H \rightarrow ss$ at FCC-ee

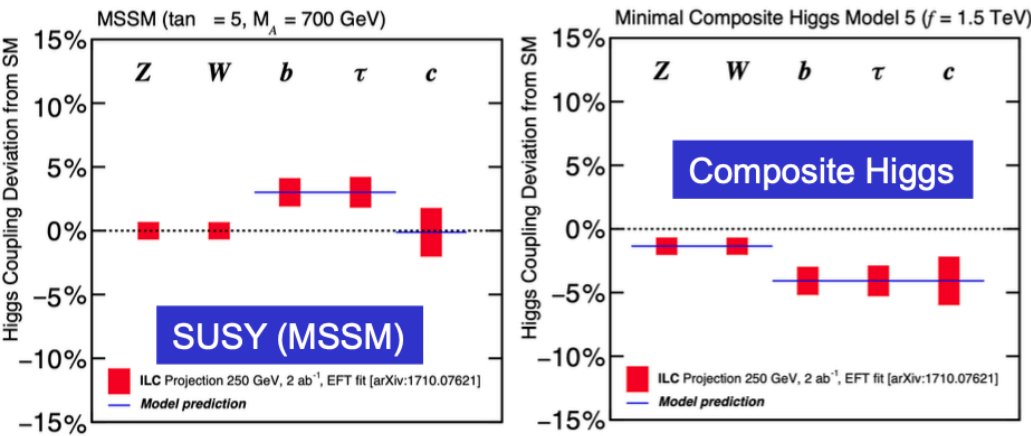
Evaluation of the measurement sensitivity of Higgs boson coupling constants at the ILC

T. Seino^A, H. Ito^B, R. Sugawara^B, T. Suehara^C, R. Hosokawa^B, W. Ootani^C, S. Narita^B

^A *Dept. of Phys. The Univ. of Tokyo*, ^B *Iwate Univ.*, ^C *ICEPP, The Univ. of Tokyo*

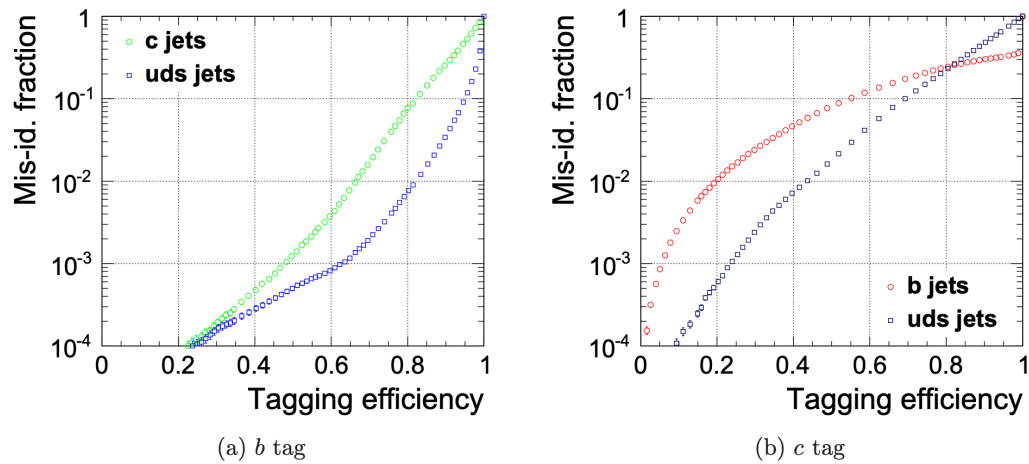
Introduction

- Measurement of Higgs boson couplings is one of the most important goals of the ILC
 - pattern of deviations can be used as probes of BSM models
 - separation of $H \rightarrow b\bar{b}/c\bar{c}/gg/s\bar{s}$ modes is critical
- Coupling to light quark (like c, s) is challenging due to its tiny Branching Ratio and large QCD background



LCFIPlus: old framework for flavor tagging

- [arXiv:1506.08371](https://arxiv.org/abs/1506.08371)
- Vertex finding $\rightarrow$ Jet clustering $\rightarrow$ Flavor tagging
 - GBDT for flavor tagging
 - separate jets into 3 categories: $b/c/\text{light}$



(framework before flavor tagging is also used for following analysis)

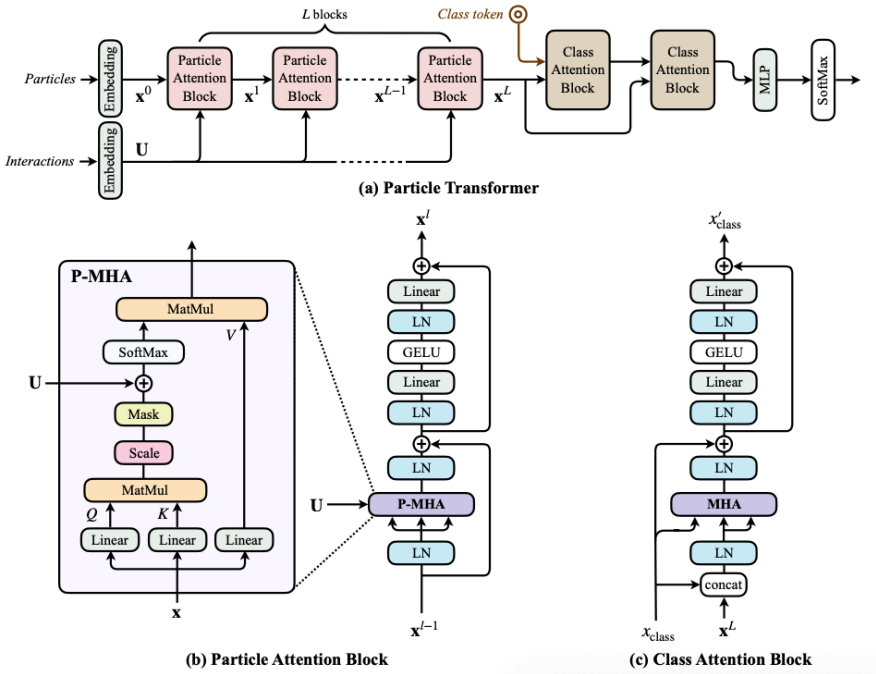
Introduction

Introduction of **Particle Transformer (ParT)** with **Comprehensive PID (CPID)** enabled strange tagging

[arXiv:2410.11322](https://arxiv.org/abs/2410.11322)

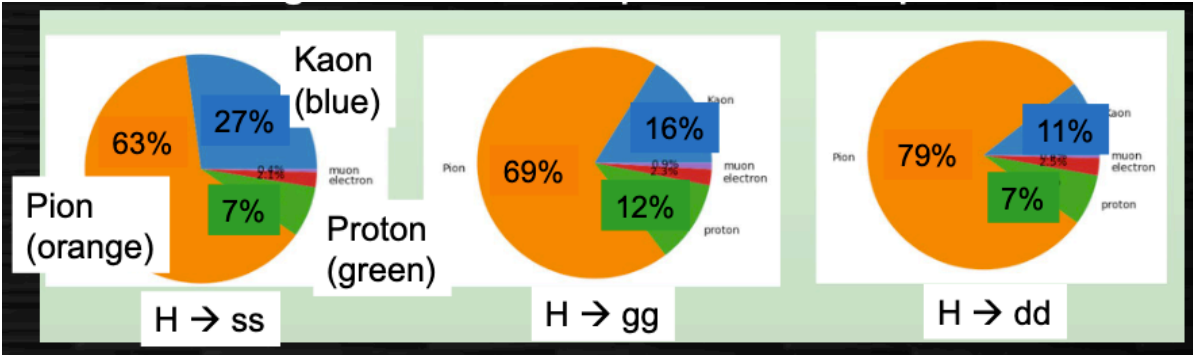
ParT

- Transformer-based architecture for Jet tagging
- Pair-wise variables (interaction, U) are added to plain Transformer encoder



CPID

- Computes hadron ID scores using dE/dx and Time-of-Flight
- Since strange jets have more strange hadrons, identifying Kaons made improvement in discrimination between strange jets and other light jets

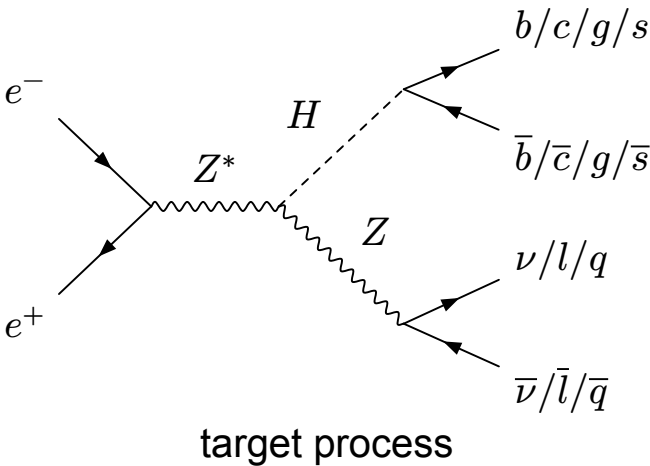


Aim:

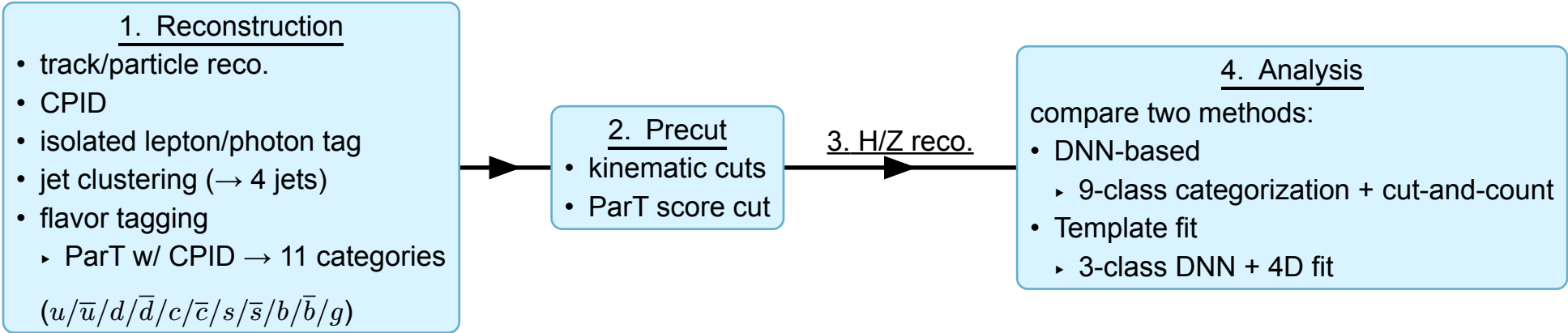
- Evaluate measurement precision for Higgs couplings to b, c, g, s
- Quantify improvements by the upgraded flavor-tagging

Basic configurations

- $\sqrt{s} = 250 \text{ GeV}$
- Detector: ILD
- $P(e^-, e^+) = (-80\%, +30\%), (+80\%, -30\%)$ (eLpR, eRpL, respectively)
- Integrated luminosity 1.8 ab^{-1}
 - 900 fb^{-1} for eLpR, eRpL
- Bkg: possible processes with 2/4 fermions in the final state
- Analysis channel is separated according to Z boson decay modes



Analysis flow

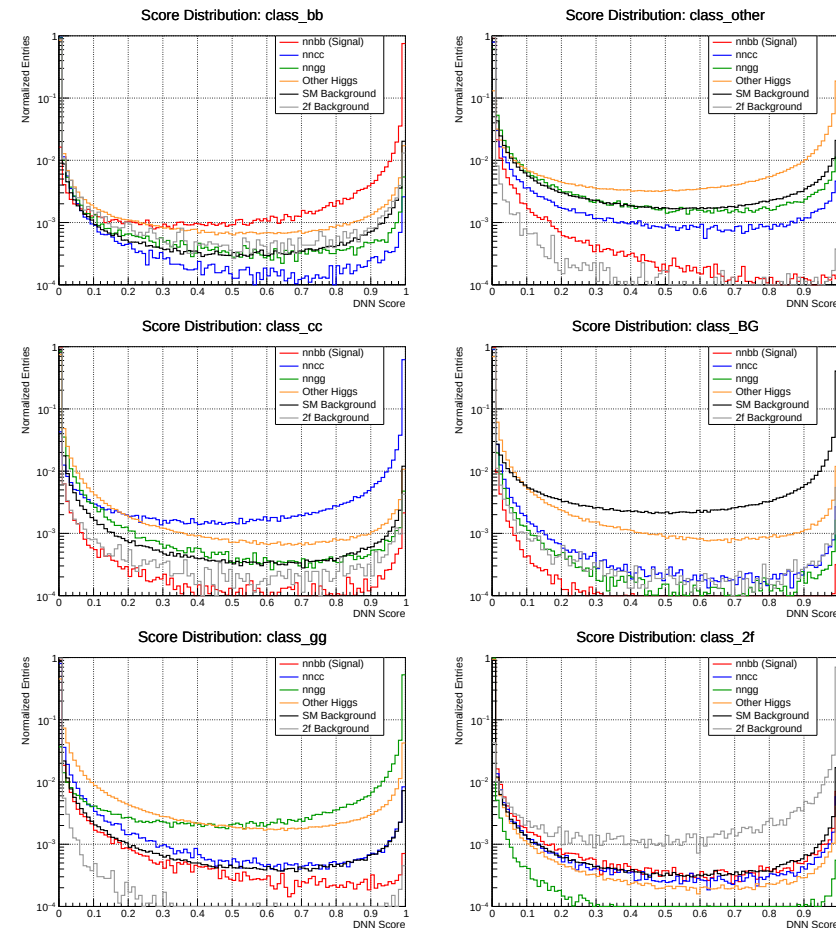


$Z \rightarrow \nu\bar{\nu}/\ell\bar{\ell}$ channel

- used TMVA framework
- categorize events into 6 classes:
 - bb/cc/gg: $H \rightarrow bb/cc/gg$ class
 - otherH: $H \rightarrow ss/\tau\tau/WW$ or ZZ process
 - SM4f: non-Higgs process with 4 fermions in FS
 - SM2f: $Z \rightarrow qq$ process
- after event classification, searched for the best cut condition using DNN score in each category
- finally evaluate the measurement accuracy by likelihood ratio test

$$Z_{\text{comb}} = \frac{\sum_{i=1}^{n_{\text{samp}}} s_i \ln\left(1 + \frac{s_i}{b_i}\right)}{\sqrt{\sum_{i=1}^{n_{\text{samp}}} b_i \left[\ln\left(1 + \frac{s_i}{b_i}\right)\right]^2}}$$

(measurement accuracy $\sigma = 1/Z_{\text{comb}}$)



DNN score plots for each category
(for neutrino channel)

$Z \rightarrow \nu\bar{\nu}/\ell\bar{\ell}$ channel

result table for $Z \rightarrow \nu\nu/\ell\ell$ channel:

	$Z \rightarrow \nu\bar{\nu}$						$Z \rightarrow e^+e^-$						$Z \rightarrow \mu^+\mu^-$					
	$H \rightarrow b\bar{b}$		$H \rightarrow c\bar{c}$		$H \rightarrow gg$		$H \rightarrow b\bar{b}$		$H \rightarrow c\bar{c}$		$H \rightarrow gg$		$H \rightarrow b\bar{b}$		$H \rightarrow c\bar{c}$		$H \rightarrow gg$	
	eLpR	eRpL	eLpR	eRpL	eLpR	eRpL	eLpR	eRpL	eLpR	eRpL	eLpR	eRpL	eLpR	eRpL	eLpR	eRpL	eLpR	eRpL
Sig.	143.5	116.4	14.5	15.1	36.6	33.2	59.4	49.8	6.8	6.3	15.6	13.5	60.1	50.9	7.3	7.3	16.1	14.1
Z_{comb}	184.8		21.0		49.4		77.5		9.2		20.6		78.8		10.4		21.4	
acc. [%]	0.54		4.77		2.02		1.29		10.83		4.86		1.27		9.66		4.67	

comparison with [previous study](#) (dataset: eLpR, $L = 250 \text{ fb}^{-1}$)

	$Z \rightarrow \nu\nu$		$Z \rightarrow e^+e^-$		$Z \rightarrow \mu^+\mu^-$	
	this work	previous work	this work	previous work	this work	previous work
$H \rightarrow b\bar{b}$ [%]	1.32	1.7	3.19	3.8	3.16	3.3
$H \rightarrow c\bar{c}$ [%]	13.1	11.2	28.1	26.8	26.0	22.6
$H \rightarrow gg$ [%]	5.2	13.9	12.2	31.3	11.8	33.0

Across all Z decay modes, performance improved in the $H \rightarrow b\bar{b}/gg$ channels, but degraded in the $H \rightarrow c\bar{c}$ channel.

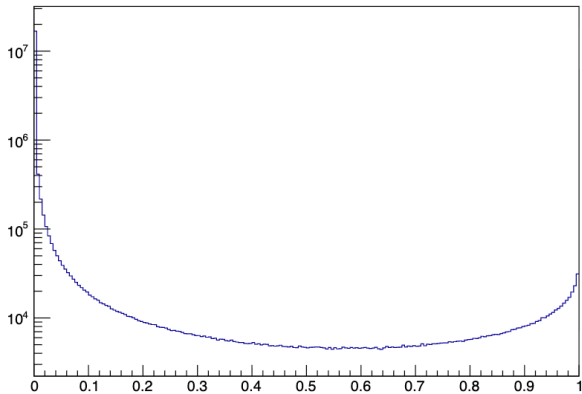
$Z \rightarrow q\bar{q}$ channel

ParT training for 2f/4f separation

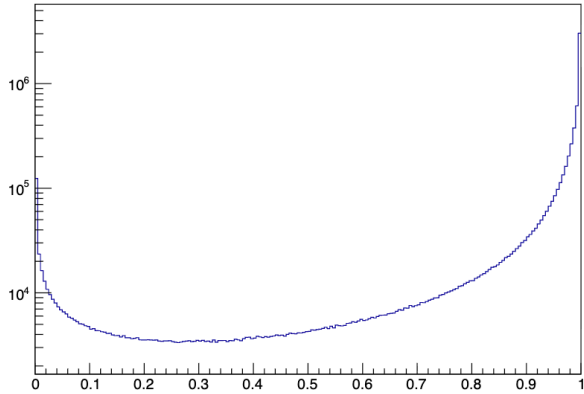
- 2f process largely survived only with y-value cut
- train with 4-momentum and CPID score of all particles
(CPID scores only available for charged particles with energy > 1 GeV)

# used samples	train	val	test
Higgs process	0	0	3.5M
4f non-Higgs process	10M	50k	3.5M
2f process	10M	50k	30M

ParT score (2f process)



ParT score (4f process)



4f-likeness score by ParT

configuration of ParT

layout	head	8
	layer	4
	particle embedding dim.	128
	interaction embedding dim.	64
	activation func.	GELU
strategy	MLP layer	2
	framework	weaver
	optimizer	Ranger
	loss func	cross entropy error
	training epoch	20
hyper parameter	batch size	768
	initial LR	1×10^{-4}

precut: score > 0.1

→ Decreased 2f events down to 5.2%

# events	full	after precut
4f processes	15.2M	15.1M
2f process	67.0M	3.5M

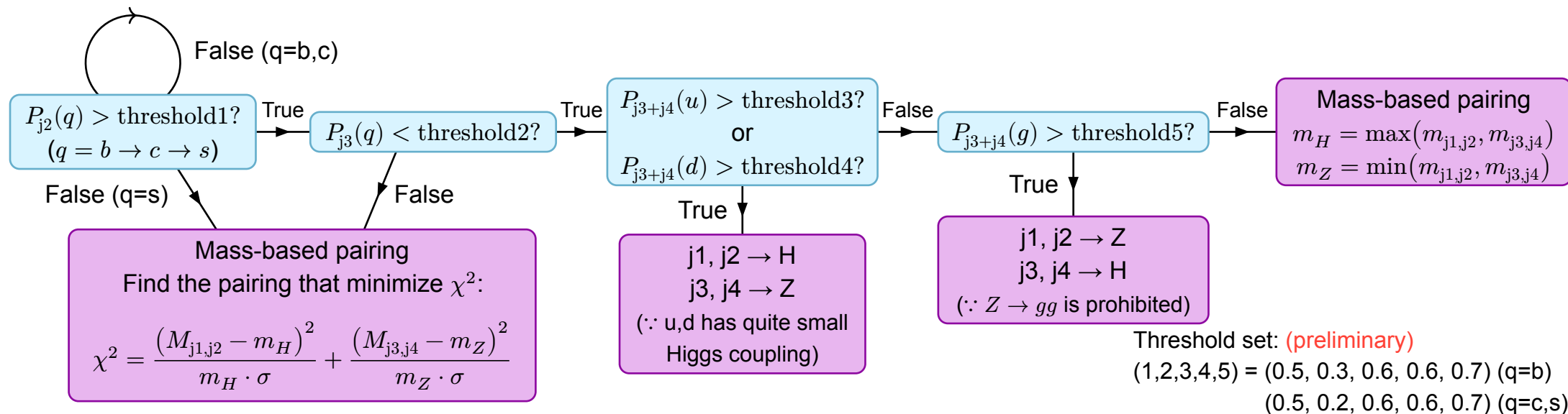
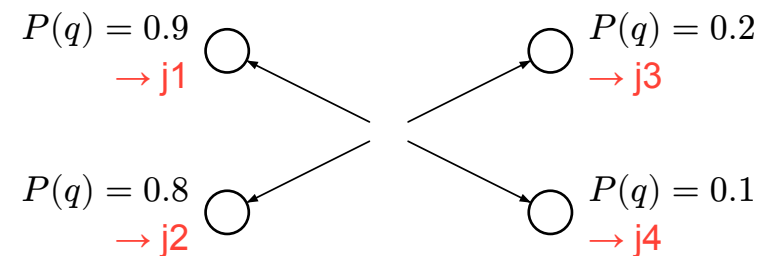
$Z \rightarrow q\bar{q}$ channel

H/Z reconstruction

- Reconstructed particles are clustered into 4jets $\rightarrow$ Perform jet pairing
- Assign jets to H/Z, mainly using flavor tag scores
 - check if 2 pairs ($qq+q'q'$) can be formed
 - if possible,
 - if the rest are u,d,g: Assign H/Z via flavor scores
 - otherwise: rely on mass-based pairing
 - if not, go on to next flavor test
 - when no pair was made, rely on mass-based method(minimizing χ^2)

Definition of $P(q)$: q -tag + $\bar{q}$ -tag

$\rightarrow$ Let “ j_i ” be the jet with the i -th highest $P(q)$ value.

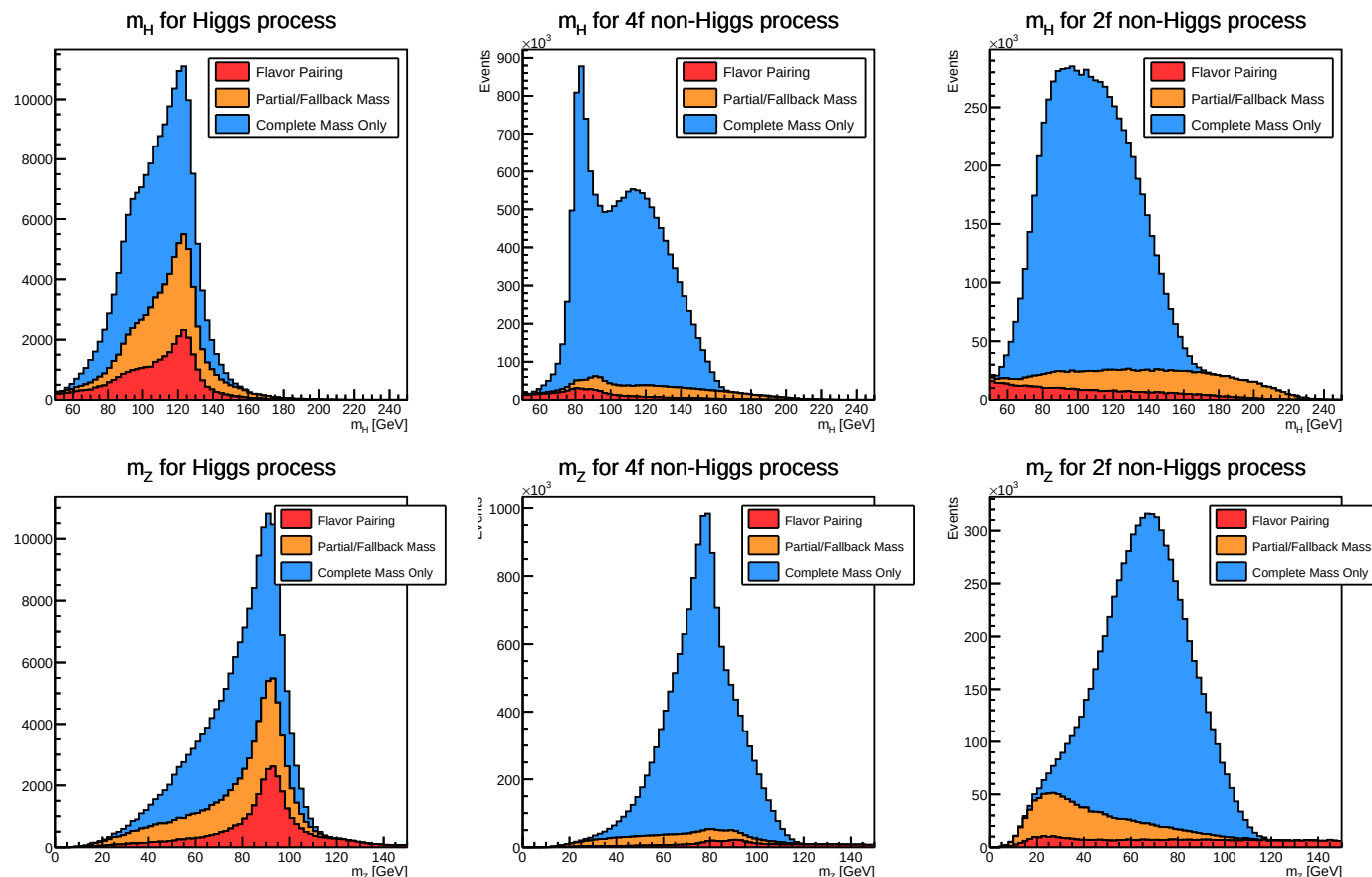


$Z \rightarrow q\bar{q}$ channel

Distribution of reconstructed H/Z mass

- Reconstruction strategy based primarily on flavor-tagging scores
 - remains heavily reliant on mass-based method
- Insufficient background discrimination with this variable

We will investigate improvements via truth jet information



$Z \rightarrow q\bar{q}$ channel

Template fit

- Preselection: ParT score > 0.1 , $100 < m_H/[\text{GeV}] < 140$
- DNN training without flavor-tag variables
 - 3 class: Hqq/4fSMbkg/2fSMbkg
 - $\rightarrow$ cut Hqq score > 0.9

Template Creation

- Calc. **f -likeness** ($f = b, c, g, s$) via:

$$L_f = \frac{f_1 f_2}{f_1 f_2 + (1 - f_1)(1 - f_2)} \quad \left(\text{w/ } \sum_f L_f = 1 \right)$$

where the index 1/2 denotes the 2-jets assigned to the Higgs candidates

- Construct 4D histogram (b, c, s -likeness + DNN score)
- Variable binning:
 - b-likeness: {0.0, 0.2, 0.6, 0.9, 1.0}
 - c-likeness: {0.0, 0.2, 0.99, 0.995, 0.998, 1.0}
 - s-likeness: {0.0, 0.2, 0.6, 0.9, 1.0}
 - DNN score: {0.0, 0.1, 0.5, 0.85, 0.95, 0.99, 0.995, 1.0}

Sensitivity Evaluation

- **Toy MC**: 1M pseudo-experiments performed by generating Poisson fluctuations in each bin
- **Binned Likelihood Fit** on signal strength r_s :

$$f = \sum_s r_s N_{ijkl}^s + N_{ijkl}^{\text{bkg}}$$

(assumed the # bkg processes are well known)

- Measurement accuracy $\sigma = \text{RMS}/\text{Mean}$, extracted from distribution of fitted r_s

$Z \rightarrow q\bar{q}$ channel

Result and Comparison

- Calculated combined result using formula:

$$\frac{1}{\sigma_{\text{comb}}^2} = \frac{1}{\sigma_L^2} + \frac{1}{\sigma_R^2}$$

- 2 previous studies are shown as reference
 - ILC: [H. Ono and A. Miyamoto, \(2013\).](#)
 - eLpR only, 250 fb^{-1}
 - FCC: [A. Del Vecchio et al., \(2025\).](#)
 - no polarisation, 10.8 ab^{-1}

- Better results obtained by Template fitting
- Good improvement for $H \rightarrow b\bar{b}/gg$
- The accuracy for $H \rightarrow c\bar{c}$ is particularly bad
 - only limited improvement from previous study
- Large discrepancy in $H \rightarrow s\bar{s}$:
 - possibly due to differences in simulation conditions (e.g. luminosity, polarisation, fast/full simulation)

acc[%]	template fit			previous work	
	eLpR	eRpL	combine	ILC	FCC
$H \rightarrow b\bar{b}$	0.7	0.7	0.5	1.5	0.3
$H \rightarrow c\bar{c}$	15.4	12.8	9.8	10.2	3.5
$H \rightarrow gg$	8.2	6.5	5.1	13.1	3.1
$H \rightarrow s\bar{s}$	3565	2557	2077		410

Summary and Prospects

Summary

- Evaluated $H \rightarrow b\bar{b}/c\bar{c}/gg/s\bar{s}$ measurement accuracy at the ILC 250 GeV stage
- Results are summarised in the table
 - result for $H \rightarrow b\bar{b}/gg$ largely improved
 - $H \rightarrow c\bar{c}$ process remain worse than expected

summary table

$\sigma_{\text{comb}} [\%]$	$Z \rightarrow \nu\bar{\nu}$	$Z \rightarrow e^+e^-$	$Z \rightarrow \mu^+\mu^-$	$Z \rightarrow q\bar{q}$
$H \rightarrow b\bar{b}$	0.54	1.29	1.27	0.5
$H \rightarrow c\bar{c}$	4.77	10.83	9.66	9.8
$H \rightarrow gg$	2.02	4.86	4.67	5.1
$H \rightarrow s\bar{s}$				2077

Prospect

- Improve jet-pairing logic for $Z \rightarrow q\bar{q}$ channel, using truth-jet information
- Conduct template fitting method also for $Z \rightarrow \nu\bar{\nu}/\ell\bar{\ell}$ channel
 - (template fitting method got better result than only one DNN)
- Investigate why the accuracy for $H \rightarrow c\bar{c}$ was limited

Motivation

Flavor-changing Higgs decays are highly suppressed in the Standard Model and therefore provide an important probe of new physics. The rare $H \rightarrow bs$ decay involves different quark flavors and can provide information on the flavor structure of the Higgs.

Analysis

$$e^+e^- \rightarrow ZH / H \rightarrow bs, Z \rightarrow ll \text{ @ } ILC250GeV$$

- Signal sample generation in the $Z \rightarrow \ell^+\ell^-$ channel
- Event selection using LCFIPlus flavor tagging ($b\bar{b}$, b , and s tags) & Evaluation of the signal efficiency
 - Event generation: WHIZARD
 - Parton shower / hadronization: PYTHIA
 - Flavor tagging : LCFIPlus
 - Detector concept : ILD

Future work

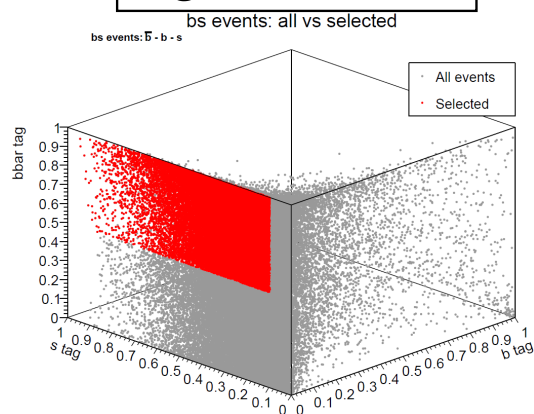
- Generation of signal samples for other Z decay channels
- Addition of other SM background
- Calculation of the signal significance & optimization of the event selection
- Evaluation of the $H \rightarrow bs$ search sensitivity

Remaining events and efficiency after each cut (preliminary)

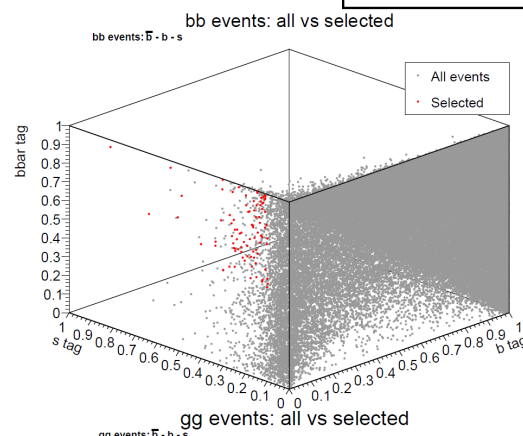
$$H \rightarrow \bar{b}s \quad Z \rightarrow \nu\nu \quad (e^-, e^+) = (-100\%, 0\%)$$

Cut : $b\bar{b}tag[jet0 \text{ or } 1] > 0.50$, $btag[jet 1 \text{ or } 0] < 0.005$, $s tag[jet1 \text{ or } 0] > 0.1$

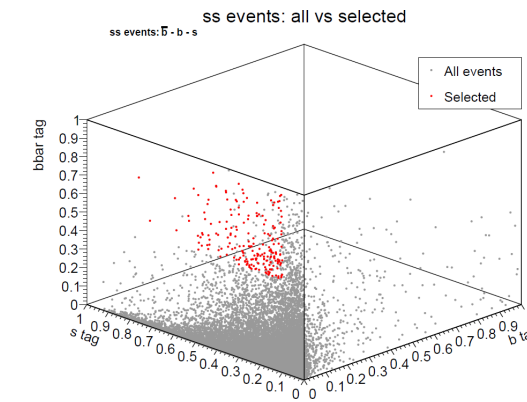
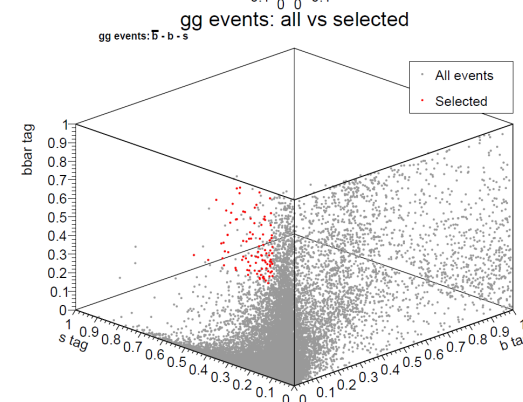
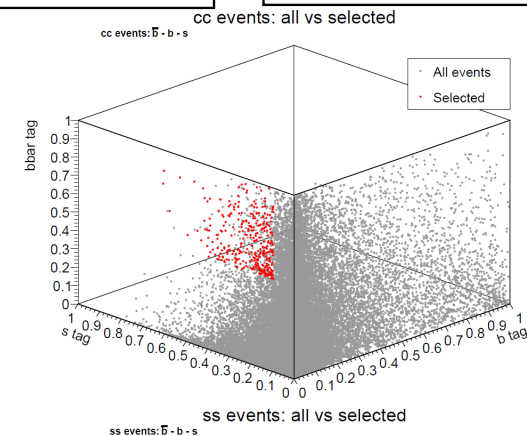
Signal: $H \rightarrow b\bar{s}$



BG: $H \rightarrow b\bar{b}/c\bar{c}/g\bar{g}/s\bar{s}$



gray : All event
red : remained event



Process	Total	$\bar{b}$	$\bar{b} + b$	$\bar{b} + b + s$	Efficiency
$H \rightarrow \bar{b}s$	100000	70017 (70.02%)	47186 (47.19%)	37519 (37.52%)	37.52%
$H \rightarrow b\bar{b}$	66000	50055 (75.84%)	255 (0.39%)	89 (0.13%)	0.13%
$H \rightarrow c\bar{c}$	76000	6071 (7.99%)	1658 (2.18%)	332 (0.44%)	0.44%
$H \rightarrow g\bar{g}$	49200	1770 (3.60%)	286 (0.58%)	98 (0.20%)	0.20%
$H \rightarrow s\bar{s}$	76000	535 (0.70%)	337 (0.44%)	206 (0.27%)	0.27%

Cuts: $\bar{b} > 0.50$, $b < 0.005$, $s > 0.10$

$H \rightarrow s\bar{s}$ at FCC-ee

- New theories, e.g. SFV 2HDM, can enhance the strange Yukawa coupling.
 - Future Higgs Factor is the unique place to allow us to measure strange Yukawa coupling directly.
- A [study](#) with IDEA concept (DELPHES) showed the uncertainty on the strange Yukawa is 110% (ZH)
 - It also suggested the K/π separation up to 30 GeV is important.
- We are investigating how ILD feature, like TPC, and advanced ML techniques can improve the measurement precision with ILD at FCC.
 - Considering to include V0 tracks in the s-tag (a [study](#) using ALEPH data)
 - For a sanity check, we are starting a validation of the FCC-ILD simulation/reconstruction by checking $Z \rightarrow qq$ JER (still on-going)

Summary

- Higgs studies with jet-flavour tagging.
 - Higgs couplings with improved JF tagging.
 - $H \rightarrow bb/gg$ are improved, we are investigating cc
 - $H \rightarrow bs$
 - Signal and BG reduction are evaluated with limited samples.
 - SR optimisation and sensitivity evaluation will be done with complete datasets.
- $H \rightarrow ss$
 - Just started, and currently a sanity check of ILD-FCC is ongoing
 - We will investigate improvements with advanced ML and V0