

ILD contribution for $H \rightarrow ss$ ECFA focused study

T. Suehara (ICEPP, U. Tokyo)

1 $e^+e^- \rightarrow Zh$ with $h \rightarrow ss$ ($Z \rightarrow$ anything) at $\sqrt{s} = 240..250$ GeV

HtoSS, HTE

Theoretical, phenomenological and MC generator targets

Expanding the BSM interpretations of the studies that have already been performed or developing new simulation-based analyses targeting specific BSM scenarios would enlarge the physics case for strange tagging at future colliders.

- Detailed uncertainty measurements
- BSM models [37];
- BSM models for states including strange quarks, e.g., 2HDM $H^+ \rightarrow cs$ BR $\approx 50\%$;
- $s\bar{s}$ vs. $b\bar{b}$ in BSM models: gain from $s\bar{s}$;
- BSM flavour structure and $h \rightarrow s\bar{s}$ signal.

- extraction of h-ss coupling strength
- BSM!

Target physics observables

Several

- $e^+e^- \rightarrow Zh$
 - $h \rightarrow ss$
 - precision
 - flavour
 - mixing
- $ee \rightarrow Zh$ with $h \rightarrow ss$ at all energies
 - projected precision on BR and diff. x-sec
 - Flavour changing decays

Target analysis techniques

The performed proof-of-concept studies [49, 51] showed that to improve the results there will be a large need for more powerful background rejection techniques as well as a potentially more global approach in the extraction of the Higgs couplings. Two areas of particular interest will be:

- diboson background suppression;
- signal extraction (fit discriminant variables, counting experiments, etc.).

Target methods to be developed

In collaboration with the Reconstruction and Detector groups, the impact from the following features will have to be evaluated when estimating the analysis sensitivity requirements:

- control of strange-tagging related systematic uncertainties;
- reconstruction of in-flight decays, e.g., $K_S^0 \rightarrow \pi^+\pi^-$;
- strangeness-tagging with ML techniques and compared with other methods;
- s vs \bar{s} separation;
- complementarity of particle identification (ID) techniques for (from dN/dx , dE/dx , ToF, RICH);
- understanding the contribution from $g \rightarrow s\bar{s}$ (from single jets) to the overall analysis sensitivity.

- strange tagging
- in-flight decays
- PID techniques
- $g \rightarrow ss$

Target detector performance aspects

The obtained results will inform the community on two crucial aspects:

- dependence of the precision on physics observables on particle ID, strange-tagging, and reconstruction capabilities;
- technology benchmarks for sub-detectors.

Generation and Simulation needs

Full simulation samples will be needed to perform the studies listed above. $\sqrt{s} = 240/250$ GeV and $350/380/550$ GeV are available as indicated in the table. In the years to come, it will be important to iterate with the detector design and fragmentation uncertainties in order to account for more realistic conditions.

- fragmentation uncertainties
- ...

The H → ss workshop

<https://indico.cern.ch/event/1405461/>

THURSDAY, 16 MAY

Time	Topic	Speaker	Duration
22:00 → 22:05	Welcome	Speakers: Valentina Cairo (CERN), John Alison (Carnegie-Mellon University (US)), Karsten Koeneke (Albert Ludwigs Universitaet Freiburg (DE)), Loukas Gouskos (Brown University (US)), Matthew Basso (TRIUMF (CA)), Caterina Vernieri (SLAC National Accelerator Laboratory (US)), Taikan Suehara (ICEPP, The University of Tokyo (JP)), Dr Valerio Dao (Stony Brook University), Yotam Soreq (Technion- Israel Institute of Technology (IL))	5m
22:10 → 22:15	Timelines	Speaker: Jenny List (Deutsches Elektronen-Synchrotron (DE))	5m
22:25 → 22:35	Theory Introduction: Higgs and light flavours	Speaker: Yotam Soreq (Technion- Israel Institute of Technology (IL))	10m
22:40 → 22:55	The challenge of fragmentation modelling	Speaker: Peter Skands (Monash University (AU))	15m
23:00 → 23:15	Quick review of existing studies	Speaker: Matthew Basso (TRIUMF (CA))	15m
23:15 → 23:25	The first step in your studies: where to find the samples [documentation contribution]		10m
23:35 → 23:45	PID reconstruction	Speaker: Ulrich Einhaus	10m
23:50 → 00:00	Detector design	Speaker: Roger Forty (CERN)	10m

🔗 The first step in your studies: where to find the samples

The table below summarises where you can find samples for the various accelerator and detector configurations.

Accelerator	Detector	Samples	Instructions	Info on Membership
ILC	ILD	-----	-----	To ask to join ILD as a guest member, one should write the spokesperson (Ties Behnke) an email asking to join, with some indication of the interests and planned work. Guest membership can be either personal or as an institution (in which case a list of the potential members should be provided). Acceptance is decided on a case-by-case basis by the Executive Team.
FCC	IDEA	Link	Link	To be able to read the files, the user has to subscribe to "fcc-eos-access" e-group. Signing the MoU is highly encouraged, see this page
CEPC	-----	-----	Contact manqi.uan@ihep.ac.cn and zhuyf@ihep.ac.cn	-----

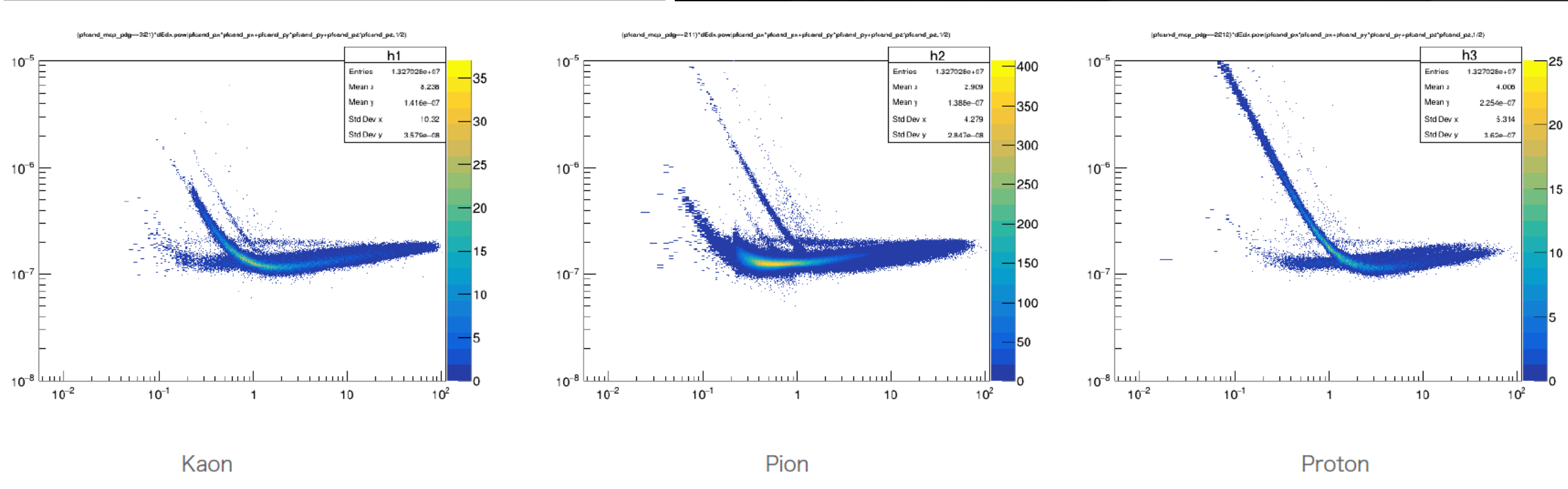
List of active works in ILD

- ILD: Comprehensive $H \rightarrow ss$ paper exists, including discussion on additional RICH on detector arxiv:2203.07535
→ any progress expected?
- ILD: Ongoing work with DNN (ParticleNet/Par. Transformer)
 - Tokyo (R. Tagami, TS), Iwate (R. Sugawara, R. Hosokawa)
 - For us: first result of strange tagging targeted at July (LCWS)
 - Application to $H \rightarrow ss$ in summer (to be presented in ECFA WS)
 - Iwate people are working on reviving previous $H \rightarrow qq$ analysis (to replace flavor tagging and do $H \rightarrow ss$)
- Others?

Progress in strange tag

	s vs c	s vs g	s vs u
0.8 efficiency	0.138	0.288	0.466

Current performance with ParT
(under investigation yet)



dE/dx inside strange jets (separated by MC PID)

Issues / discussion topics

- Particle ID: just applying likelihoodPID seems not too efficient
 - Lepton ID is quite bad, pi/K/p separation is also not very good
 - Maybe having some soft issues
 - Replace by CPID? (ready?) We need to implement LCFIPlus interface
- Inference at Marlin
 - Will try soon in our side (probably using DESY's code)
- dE/dx not very good compared to dN/dx in IDEA
 - They claimed ~ 7 sigma separation
 - We may have dN/dx with pixel TPC?
 - Further better way is to just use every TPC hit (waveform?) to ML to directly obtain PID (but needs extra work...)

Backup

Slides from Paestum workshop

H → ss: target by organizers (1)

Theoretical, phenomenological and MC generator targets

- BSM models predicting deviations in $h \rightarrow s\bar{s}$, e.g. SUSY
- BSM models predicting for example charged Higgs boson with large branching ratios in final states including strange quarks, e.g. 2HDM $H^+ \rightarrow cs$ BR $\approx 50\%$
- $s\bar{s}$ vs. $b\bar{b}$ in BSM models: gain from $s\bar{s}$?
- flavor assumptions in EFTs: decouple 3rd from 1st/2nd family? Partially looked at in the context of the Spontaneous Flavour Violating framework.

No serious model discovered so far?

TBD?

No serious model discovered so far?

Works exist

Target physics observables

- $e^+e^- \rightarrow Zh$ with $h \rightarrow ss$ ($Z \rightarrow$ anything) at $\sqrt{s} = 240/250$ GeV (higher center of mass energies still unexplored)
- projected precision on branching fraction, and differential cross-section in $\cos\theta_s$
- Flavour changing decays are very rare in the SM, for example $\text{BR}(h \rightarrow bs) \simeq 10^{-7}$. NP models, which can be encapsulated by an EFT, allow larger values.

Done with some concepts: to be collected and compared?

TBD?

Target methods to be developed

- charged hadron ID from dN/dx , dE/dx , ToF, RICH, study complementarity in momentum reach.
- reconstruction of in-flight decays, e.g. $K_S^0 \rightarrow \pi^+\pi^-$
- strangeness-tagging
- s vs \bar{s} separation
- control of strange-tagging related systematic uncertainties

Work in progress

TBD?

TBD (some ongoing)?

Some works exist, some ongoing

TBD?

H \rightarrow ss: target by organizers (2)

Target analysis techniques

- diboson background suppression
- signal extraction (fit discriminant variables, counting experiment etc)

Done in ILD study, more to come?

Target detector performance aspects

- dependence of the precision on physics observables on particle ID and reconstruction capabilities

Some works exist, more to come?

MC samples needed

full SM and $e^+e^- \rightarrow f\bar{f}h$ at $\sqrt{s} = 240/250$ GeV and 350...380 GeV available in general samples listed in Section

Existing tools / examples

- similar ILD analysis for $h \rightarrow b\bar{b}/c\bar{c}/s\bar{s}$: https://github.com/ILDAnaSoft/ILDbench_Hbbccgg
- similar SiD analysis ...
- similar CLICdp analysis ...
- similar IDEA analysis ...
- similar CLD analysis ...

Common topics

- Demonstration of technologies
(dE/dx or dN/dx , large-scale picosec timing detectors, RICH)
 - Clarify difference of the performance between detector concepts
- Reconstruction algorithm
 - GNN/Transformer seems promising, dependence on detector performance to be seen
 - Common framework to use ML-based algorithm to analysis
 - Difference between fast and full sim identified
- Physics analysis – comparison of results
- Interpretation (BSM sensitivity?)

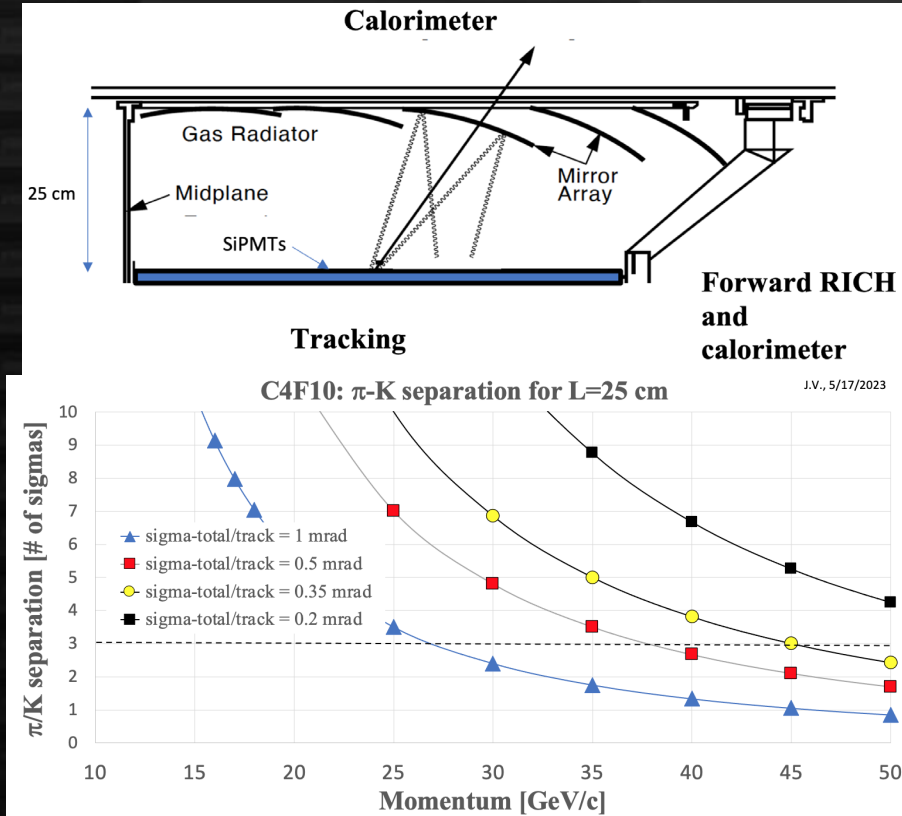
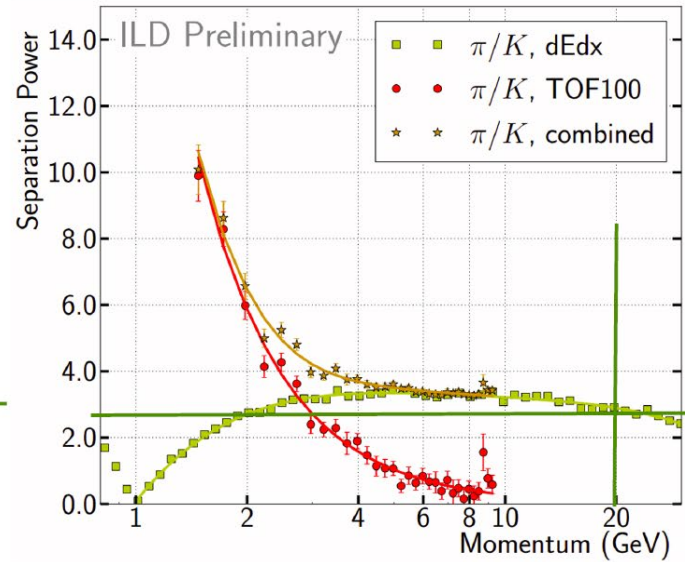
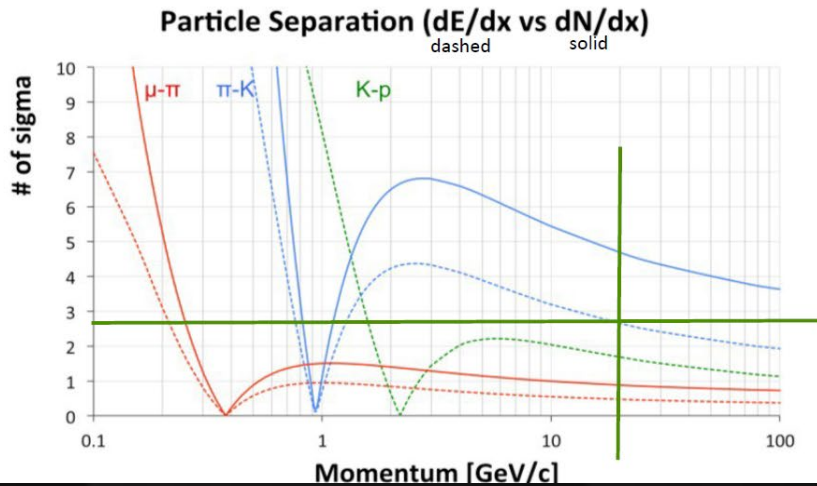
Difference on PID (e.g. ILD and FCCee)

arXiv: 2307.01929

Particle ID is critical

Both IDEA@FCC and ILD@ILC feature a PID detector, a drift chamber or TPC respectively

1912.04601
e2019-900045-4



How the performance should be validated? Detector prototype available?
 Picosec detectors at calorimeter practical? Power consumptions?
 RICH to be included? Impact on detector design/performance?
 Other technologies?

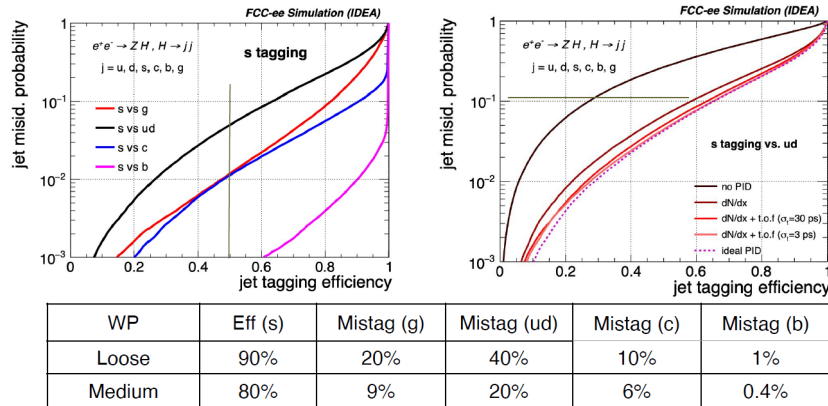
Current status of strange tagging

PRD 101 056019 (2020)
EPJ C 82 646 (2022)
L. Gouskos @FCC week

Strange tagging performance 1/2

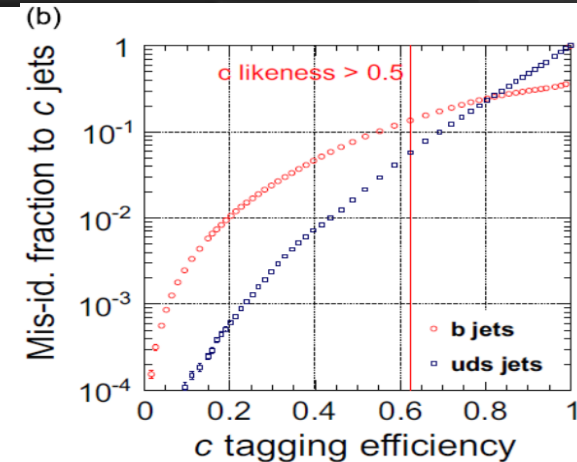
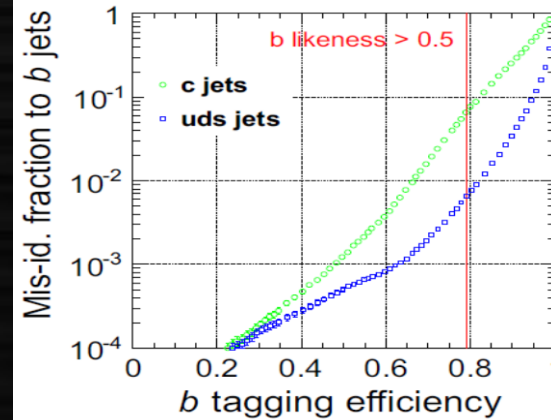
IDEA-like detector and Particle cloud graph neural network (fast sim)

- Both TOF and dN/dx ($3\sigma < 30$ GeV) included as inputs
- No PID to PID with dN/dx \rightarrow at fixed mistag, efficiency doubles



ILD: b/c tagging with Particle Transformer strange tagging to be investigated

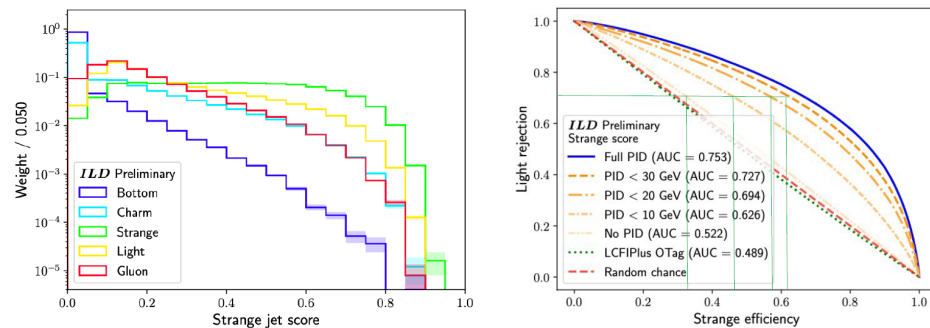
LCFI+



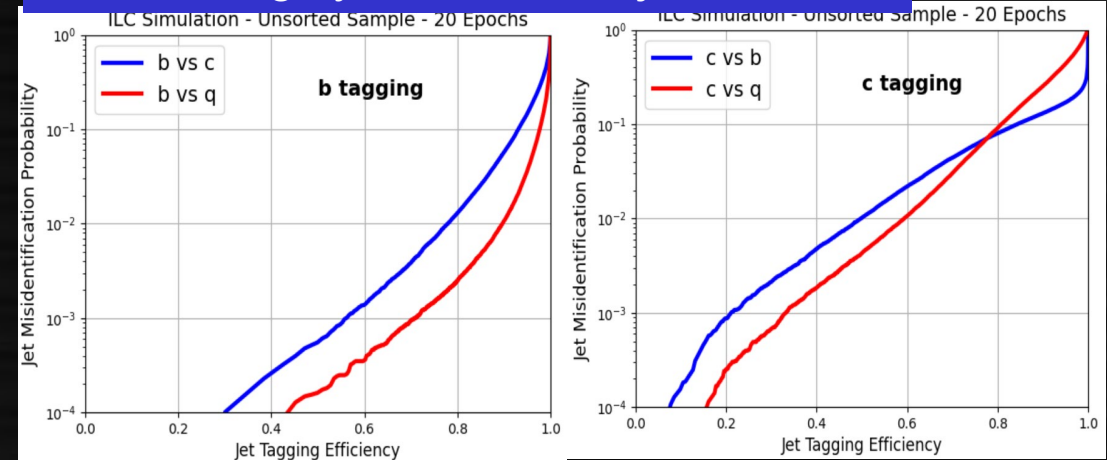
Strange tagging performance 2/2

ILD-like detector with full simulation and Recurrent NN

- Includes PDG-based PID \rightarrow assuming perfect detector capability
- At 50% s-tag efficiency, 90% background rejection
- No PID to PID < 10 (30) GeV \rightarrow at fixed mistag, 1.5x (2x) efficiency



ParT – roughly 10x better rejection ratio



Analyses to be compared

- Strange tag need to be fixed first
 - DNN-based tagger seems to be baseline
 - Need to incorporate into (common?) analysis framework
- Analyses are usually difficult to compare
 - Detector different
 - Simulation details different
 - Analysis method different

Difficult to disentangle those – common framework / analysis would help (to be discussed)

Issues on $H \rightarrow ss$ (apart from strange tag)

- Higgs with other decays ($bb/cc/gg$): main background
 - Need to clarify Z/W background for some channels though
- Separation of $H \rightarrow bb/cc$ is relatively easy
($H \rightarrow bb$: clear signature,
 $H \rightarrow cc$: statistically less demanding)
→ the critical part is discrimination of $H \rightarrow gg/ss$
(if we ignore $H \rightarrow$ light quarks / exotic)
- “Gluon tagging” may be rather essential
 - Included in the current tagger (in part of multiclass)
 - Different from $e^+e^- \rightarrow ss$ (where strange tag is essential)

Investigating more issues

- Jet charge of strange?
- Exotic decay of Higgs like $H \rightarrow bs$?
- Differential cross section of $H \rightarrow ss$?
- Decay of K_0 short?
- Heavy Higgs decay e.g. $H^+ \rightarrow cs$?
- Systematic effects?
 - Esp. serious for DNN-based algorithms

Interpretation?

- Theory colleagues would propose way to go...
 - Sorry, not available today

- Some quasi-personal comments:
 - Physics case for $H \rightarrow ss$ is not very strong
(not the first 2nd generation quark, not easy to reach SM...)
Worthwhile to separate from 1st generation in case of large deviation?
 - How about $H \rightarrow bs$, $H^+ \rightarrow cs$, other exotics?

Caterina's summary

Conclusions and next steps

s-tagging & PID would allow for a complete exploration of the 2nd generation Yukawa couplings

- First simulations with some assumptions on detector performance show promise to test κ_s
- Moving forward we want to:
 - map this into phenomenological targets
 - i.e. BSM models predicting deviations in $h \rightarrow ss$, or $h \rightarrow cs$
 - refine the analysis for $e^+e^- \rightarrow Zh$ with $h \rightarrow ss$ ($Z \rightarrow X$) at 240/250 GeV
 - higher center of mass energies still unexplored
 - study detector benchmarks:
 - the complementarity in momentum reach of charged hadron ID from dN/dx , dE/dx , ToF, RICH
 - reconstruction of in-flight decays, $K^0_S \rightarrow \pi^+\pi^-$
 - strangeness-tagging and $s/s\bar{c}$ separation
 - ***Important to evaluate simultaneously other Higgs benchmarks***

Theoretical, phenomenological and MC generator targets

Expanding the BSM interpretations of the studies that have already been performed or developing new simulation-based analyses targeting specific BSM scenarios would enlarge the physics case for strange tagging at future colliders. In particular, we welcome studies in the following areas:

- Detailed understanding of how to extract the Higgs-strange coupling strength from a $\text{BR}(h \rightarrow s\bar{s})$ measurement, given contributions from Dalitz decays, e.g., $h \rightarrow g^*(\rightarrow s\bar{s})g$ or $h \rightarrow \gamma^*(\rightarrow s\bar{s})\gamma$.
- BSM models predicting deviations in $h \rightarrow s\bar{s}$, e.g., SUSY or composite Higgs — see Refs. [36, 37];
- BSM models predicting, for example, charged Higgs bosons with large branching ratios in final states including strange quarks, e.g., 2HDM $H^+ \rightarrow cs$ $\text{BR} \approx 50\%$;
- $s\bar{s}$ vs. $b\bar{b}$ in BSM models: gain from $s\bar{s}$;
- BSM flavour structure and $h \rightarrow s\bar{s}$ signal.

Target physics observables

Several physics quantities will be investigated:

- $e^+e^- \rightarrow Zh$ with $h \rightarrow ss$ ($Z \rightarrow$ anything) at $\sqrt{s} = 240/250$ GeV (this has been the only target so far, but it will be relevant to explore also higher centre-of-mass energies, which, in turn, enable different Higgs production modes);
- projected precision on the branching fraction and the differential cross-section in $\cos\theta_s$;
- flavour-changing decays are very rare in the SM, for example, $\text{BR}(h \rightarrow bs) \simeq 10^{-7}$. New physics models, which can be encapsulated by an EFT, allow larger values.

Target analysis techniques

The performed proof-of-concept studies [49,51] showed that to improve the results there will be a large need for more powerful background rejection techniques as well as a potentially more global approach in the extraction of the Higgs couplings. Two areas of particular interest will be:

- diboson background suppression;
- signal extraction (fit discriminant variables, counting experiments, etc.).

Target methods to be developed

In collaboration with the Reconstruction and Detector groups, the impact from the following features will have to be evaluated when estimating the analysis sensitivity reach, including:

- control of strange-tagging related systematic uncertainties;
- reconstruction of in-flight decays, e.g., $K_S^0 \rightarrow \pi^+\pi^-$;
- strangeness-tagging with ML techniques and compared with anti- b -tagging techniques;
- s vs \bar{s} separation;
- complementarity of particle identification (ID) techniques for charged hadrons in momentum reach (from dN/dx , dE/dx , ToF, RICH);
- understanding the contribution from $g \rightarrow s\bar{s}$ (from single jets) to strange-tagging performance and analysis sensitivity.

Target detector performance aspects

The obtained results will inform the community on two crucial aspects:

- dependence of the precision on physics observables on particle ID, strange-tagging, and reconstruction capabilities;
- technology benchmarks for sub-detectors.

Generation and Simulation needs

Full simulation samples will be needed to perform the studies listed above. Samples for $e^+e^- \rightarrow f\bar{f}h$ at $\sqrt{s} = 240/250$ GeV and $350/380/550$ GeV are available as indicated in the general samples listed in the motivation. In the years to come, it will be important to iterate with simulation experts on $s\bar{s}$ correlations and fragmentation uncertainties in order to account for more realistic systematic uncertainties.

Existing tools / examples

There are several existing tools and analysis codes available. At the time of writing, this includes examples for ILC and FCC-ee. However, due to ongoing developments, in case you would like to get actively engaged, please contact us directly (see below), such that we can point you to the up-to-date tools and code repositories.