Strange Tagging and $H \rightarrow s\bar{s}$ with ILD@ILC

Valentina Maria Martina Cairo(SLAC)
Matthew Basso(U. of Toronto)

on behalf of:
Strange quark as a probe for new physics in the Higgs Sector

M.J. Basso[a], V.M.M. Cairo[b], U. Heintz[c], J. Luo[c], M. Narain[c], R.S. Orr[a], A. Schwartzman[b], J. Strube[d], D. Su[b], T. Tanabe[e], E. Usai[c], C. Vernieri[b], C. Young[b]

(a) University of Toronto, Toronto ON – Canada
(b) SLAC National Accelerator Laboratory, Stanford CA – USA
(c) Brown University, Providence RI – USA
(d) University of Oregon, Eugene OR – USA
(e) High Energy Accelerator Research Organization, Tsukuba – Japan

Big Thank You to Junping Tian & Shin-ichi Kawada for all their help with ILD sample production!

FCC-ee
Physics Meeting

October 25th 2021

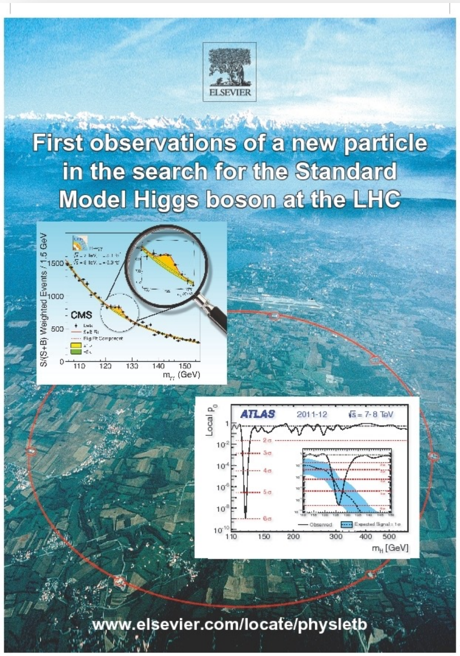
U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

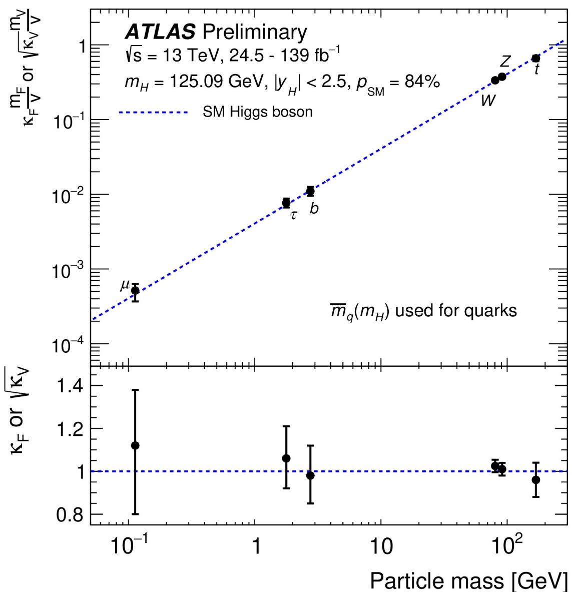
The Higgs Puzzle

SLAC

Past

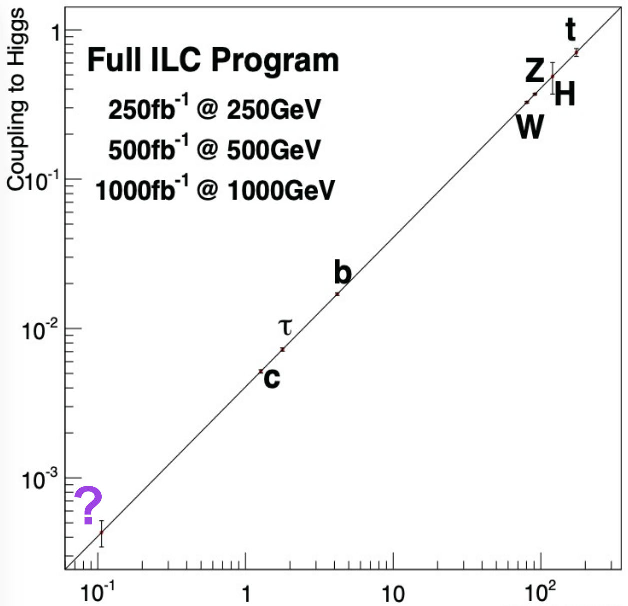


Present

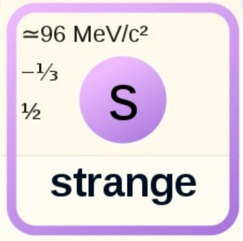


ATLAS-CONF-2020-027

Future



https://arxiv.org/pdf/1310.0763.pdf



Our LoI for Snowmass2021

- ***Strange Quark as a probe for new physics in the Higgs Sector***

> *"More specifically, in the context of Snowmass 2021, we propose to study the feasibility of the measurement of Higgs boson couplings to light quarks, in particular to strange quarks, as of paramount importance to complete the understanding of the Higgs sector. The emphasis will be put on future lepton colliders since the branching ratio for h → ss is below the level of 10-3 [6] in the SM and the measurement requires a large number of Higgs bosons in a very clean environment, but important information on the usage of advanced 4D tracking capabilities can also be learned in the HL-LHC context. This study strongly aims at motivating the development of strange tagging techniques and at providing requirements to future tracking algorithms and timing detectors performance."*

- Somewhat related to the Instrumentation Frontier LoI on 4D Tracking

The ILC case

- **ILC Study Questions for Snowmass 2021**: https://arxiv.org/pdf/2007.03650.pdf

At this moment, we see an opportunity for the ILC to actually be constructed for operation in the 2030's. This makes it especially important today to understand and evaluate the ILC capabilities.

5 Questions about general e^+e^- event analysis

11. Particle ID. The SiD and ILD detectors do not have dedicated subdetectors for particle ID; however, they can identify kaons by dE/dx measurement or by track timing. Identification of kaons is important for many aspects of e^+e^- physics, including the vertex charge measurement described in the previous question and the reconstruction of τ leptons from their final states. It is then interesting to consider how kaon identification could be improved, for example, using the new timing detector technologies being developed for HL-LHC, and the implications for e^+e^- physics measurements. [miniDST].

12. Strange quark tagging. It would be useful to be able to tag strange quark jets in e^+e^- processes. Studies of $Z \to s\bar{s}$ from the LEP/SLC era can be found in in [51, 52]; these use dedicated RICH/CRID particle ID detectors. More recent proposals for strange taggers, for general purpose detectors, are given in [53, 54]. These strategies can surely be improved. [miniDST]

The ILC case

- **ILC Study Questions for Snowmass 2021**: https://arxiv.org/pdf/2007.03650.pdf

At this moment, we see an opportunity for the ILC to actually be constructed for operation in the 2030's. This makes it especially important today to understand and evaluate the ILC capabilities.

6 Questions about Higgs boson physics: $e^+e^- \to Zh$

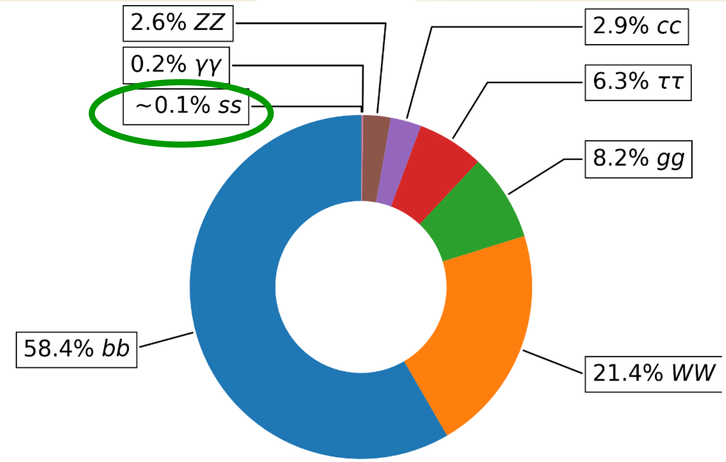
7. Higgs decays to 2 jets. At e^+e^- colliders, Higgs decays to all hadronic modes can be observed directly. Current studies of $h \to b\bar{b}, gg, c\bar{c}$ (*e.g.*, [62]) date from the era before deep learning, and before the understanding of q/g jet separation gained from LHC. What, now, is the optimum method for separating these three decay modes. What systematic errors can be achieved? [miniDST]

8. Higgs decays to light quarks. One can add to the previous question the possibility of Higgs decay to $s\bar{s}$, $d\bar{d}$, $u\bar{u}$. What limits on these modes can be achieved? Can $h \to s\bar{s}$, with $BR = 10^{-4}$ in the SM, be observed? A theoretical study on the discrimination of light quark and gluon initiated jets can be found in [63]. The possibility of observing $h \to s\bar{s}$ is discussed in [54] and in question #12 of Sec. 5. [miniDST]

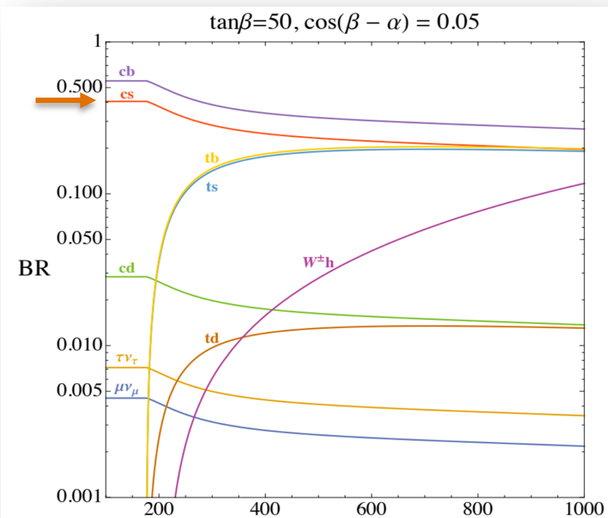
13. Flavor-violating Higgs decays. What limits can be placed on $h \to \tau\mu$, $h \to bs$, and other flavor-violating fermion combinations? [miniDST]

Goals

- Derive sensitivity to **Higgs strange Yukawa coupling at ILC**
- Develop a **strange tagger** using **ILD@ILC** and apply the tagger to a direct $SM\ H \rightarrow ss$ or $BSM\ H \rightarrow cs$ analysis
    - $H \rightarrow ss$: likely out of experimental reach unless enhanced by BSM
    - $H \rightarrow cs$: BSM models allow for the 1st & 2nd generation fermion masses to be an additional source of EW symmetry breaking
        - Charged heavy Higgs can undergo flavour violating decays (e.g., cs)
            - both **s/c-tagging** can help here
- Provide **inputs to detector instrumentation**



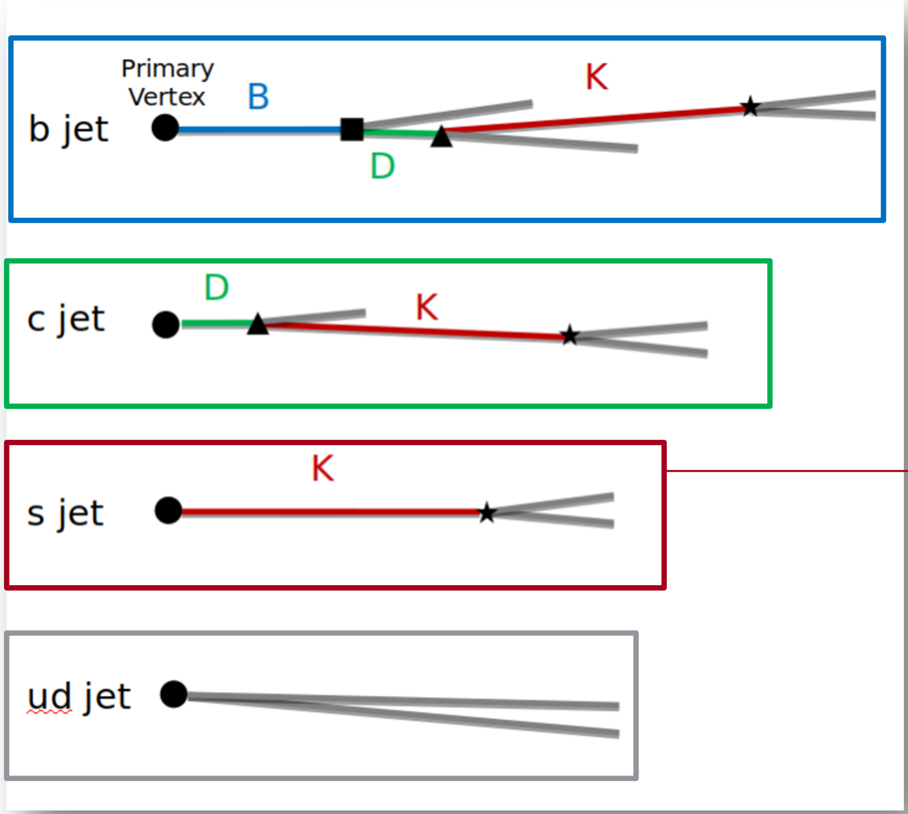
$\sqrt{s}$ = 13 TeV, m_H = 125 GeV



1610.02398.

Experimental Handles for Flavour Tagging

T. Tanabe's presentation



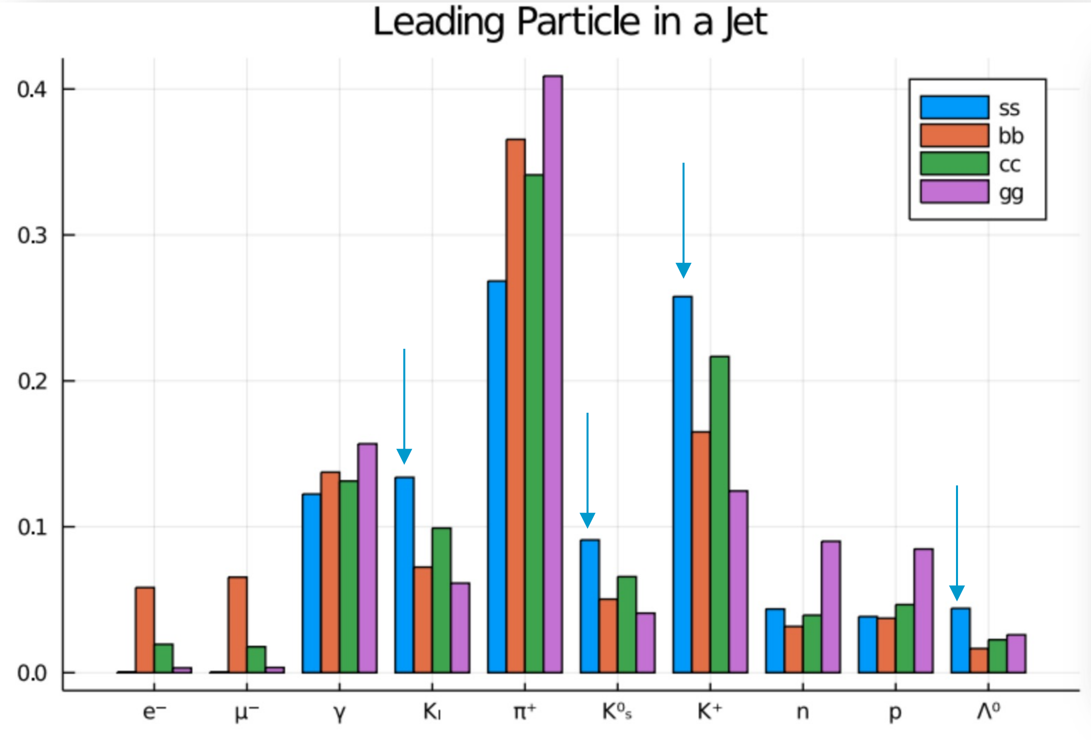
	# of secondary vertices (excluding V^0)	# of strange hadrons ($K^\pm$, K_L^0, K_S^0, Λ^0)
b	2	≥1
c	1	≥1
s	0	≥1
ud	0	0

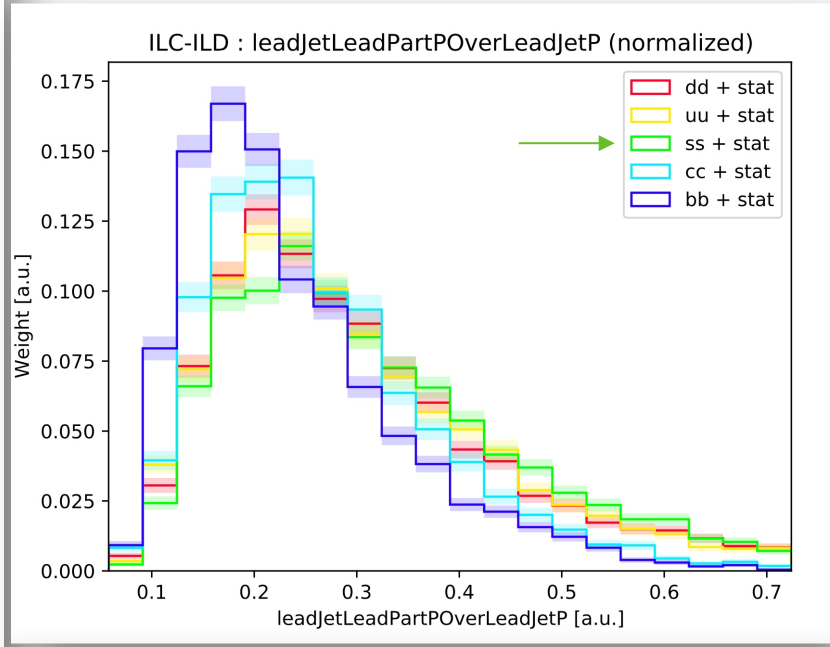
Strange Hadron reconstruction

- $K^\pm$ [PID]
- $K_S^0 \rightarrow \pi^+\pi^-$ [Vertex] (BF ~69.2%)
- $\Lambda^0 \rightarrow p\pi^-$ [Vertex] (BF ~64%)
- K_L^0 [Particle Flow]

Experimental Handles for Strange Tagging

SLAC

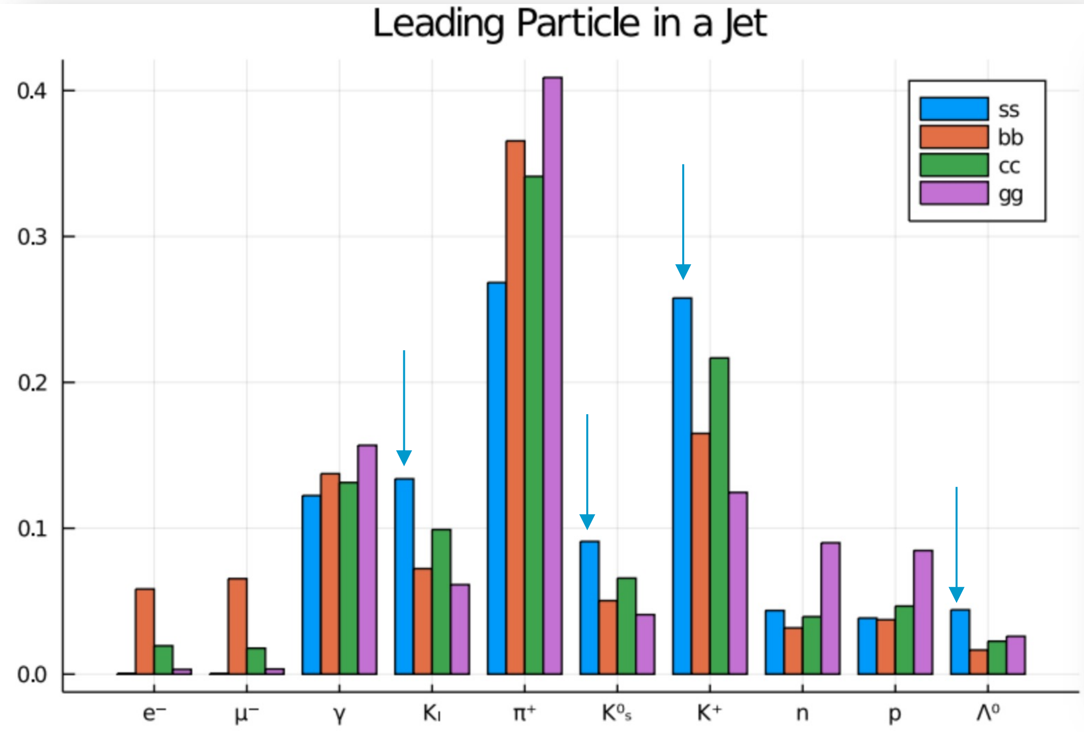


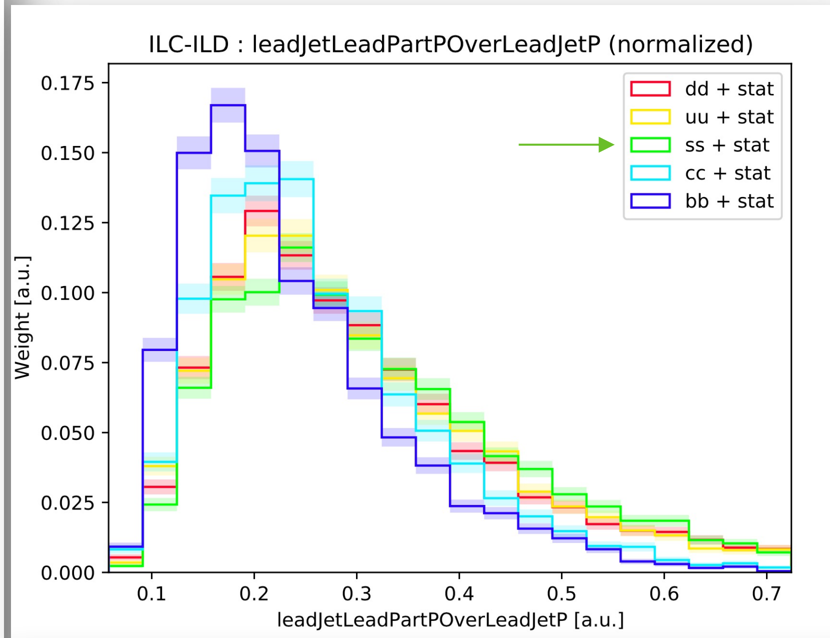


J. Strube's studies

Need **K/pi discrimination** over the momentum range approximately (0.2-0.7) x 0.5 x 125 ≅ **12 to 50 GeV**

Experimental Handles for Strange Tagging



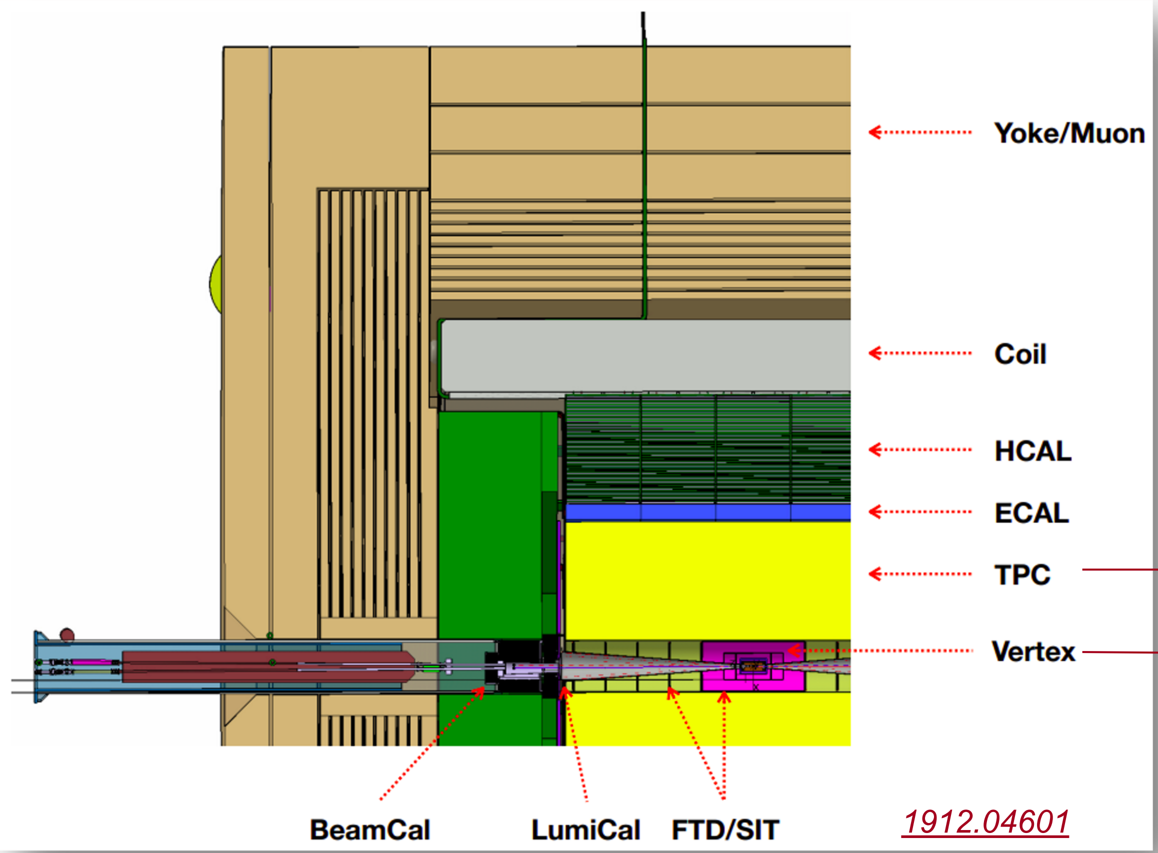


J. Strube's studies

I will not cover the PID aspects in detail today, but I invite you to join our talk at **ILCX2021** on Wednesday 27th Oct. 2021

International Large Detector (ILD) @ ILC

ILD @ ILC: proposed detector layout



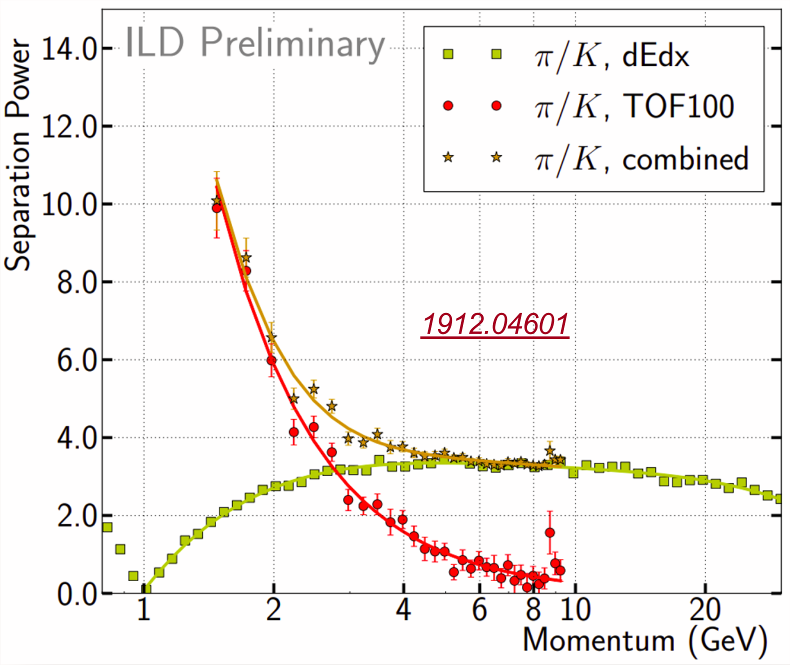
Tracking/calorimetry contained in 3.5 T field

continuous tracking with inner/outer Si layers

3 double-layer pixel

ILD @ ILC: performance

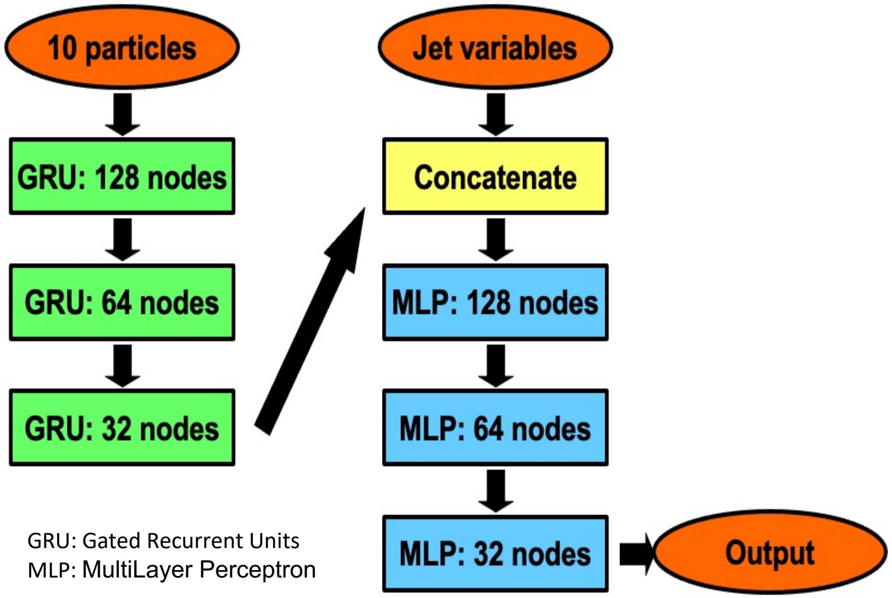
- IP resolution O(few μm)
  - Pertinent to secondary vertexing and, in turn, b/c-tagging
- **For strange vs up/down ("light") quark tagging, need kaon tagging**
  - TPC provides dE/dx, Si detectors on either side of TPC provide time-of-flight (TOF) measurement
  - TOF works best at low p (< 10 GeV), expect dE/dx to work better for kaon tagging (where p > 10 GeV)
- ILD provides BDT scores for b/c-taggers and an *other* ("o") tagger per jet



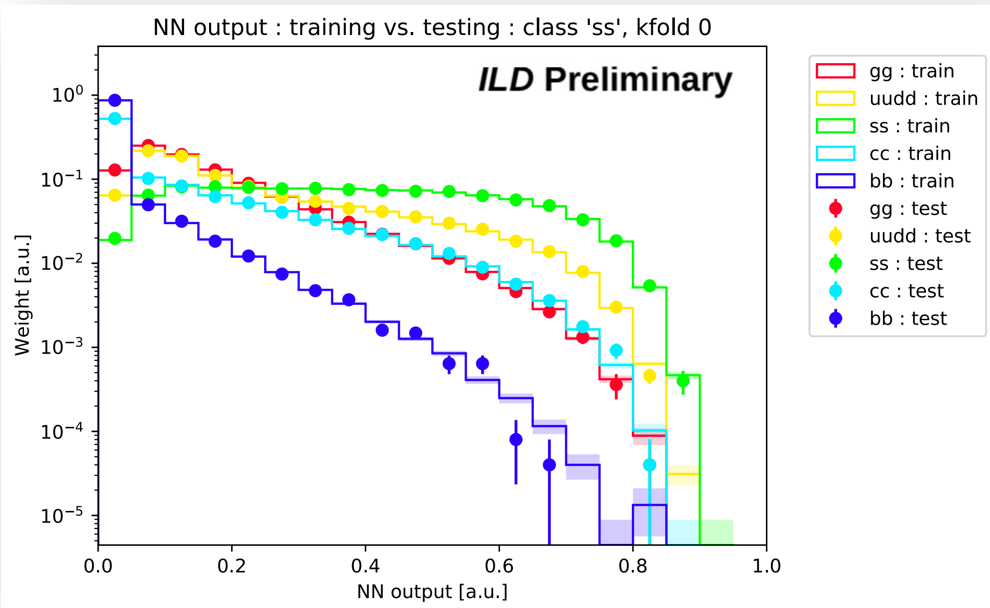
Strange Tagger

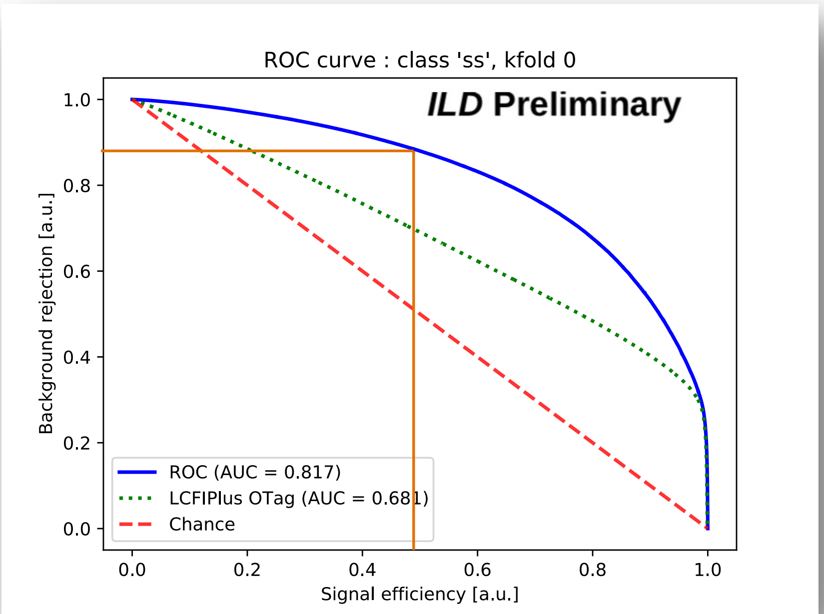
Multiclassifier Tagger: architecture & inputs

- Use a neural network-based tagger for classifying jets by flavour
- Train on ILD-reconstructed $(Z \to inv)(H \to qq/gg)$ samples
- Use **per-jet level inputs** as well as variables on the **10 leading particles** in each jet:
  - **Jets:**
    - momentum p, pseudorapidity η, polar angle φ, mass m, b/c/o-tagger scores, category, Nparticles
  - **Particles:**
    - p, η, φ, m, charge, truth electron/muon/pion/kaon/proton likelihoods (0 or 1, **using PDG ID** – "kaons" include K_S^0, $K^{+/-}$, and Λ)



Performance: s and u/d jets





- Separation of *s* and *u*/*d* is **possible** with using truth likelihoods
- Also **good discrimination of *s* jets from *g* jets** – here, $N_{particles}$ is powerful
- At **50% strange tagging efficiency**, we have **90% background rejection** over **70%** for LCFIPlus Otag (more ROC curves in back-up and LCWS2021 talk)

$H \rightarrow s\bar{s}$ analysis

σ_H @ $\sqrt{250}\mathrm{GeV}{\sim}200$ fb

- 2000 fb^{-1} collected by the ILC after 10 years
- → 400k Higgs out of which only about 40 will decay to strange quarks

Analysis overview

- Performed on same $H \to qq/gg$ samples (500K events per flavour) as well as $Z \to qq$ and $ZZ \to qqqq$ samples (~1M events each)
  - currently missing W-fusion signal (~10x smaller xs) and WW background
- Scale BR[$H \to cc$] by ratio of *s*/*c* quark mass ratio squared: BR[$H \to ss$] ~ 2E-4
- Perform simple cut-based analysis
  - Kinematic selection uses
    - **Jet quantities**:
      - leading/subleading jet momenta, p_j; dijet mass, M_{jj}; dijet energy, E_{jj}
    - **Missing 4-vector quantities**:
      - mass, M_{miss}; angular separation, $\Delta R_{jj,miss} = \sqrt{(\Delta\varphi_{jj,miss}^2 + \Delta\eta_{jj,miss}^2)}$
    - **Leading/subleading jet *b*/*c*-tagger scores**
    - **Number of Particle Flow Objects** (PFOs):
      - per event, *N*PFOs/event; per jet, *N*PFOs/jet
- Multiply cross sections by integrated luminosity of **2000 fb^{-1}** to yield events at **sqrt(s)=250 GeV**
  - Could consider adding the 500 GeV int. lumi but it implies additional sample production, not easy for the time being

Analysis cuts

- Preliminary selection:
  - Leading and subleading jet momenta, $p_j > 30$ GeV
  - Dijet mass, $M_{jj} \in [120, 140]$ GeV
  - Dijet energy, $E_{jj} \in [125, 160]$ GeV
  - Missing mass, $M_{\text{miss}} \in [75, 120]$ GeV
  - Angular separation, $\Delta R_{jj,\text{miss}} = \sqrt{(\Delta\phi_{jj,\text{miss}}{}^2 + \Delta\eta_{jj,\text{miss}}{}^2)} < 4$
    - During a recent ILD meeting, suggestion to use angular variable between jets as well, not added yet
  - Leading and subleading LCFIPlus tagger scores, $\text{score}^b{}_j < 0.2$ && $\text{score}^c{}_j < 0.35$
  - Number of PFOs per event, $N_{\text{PFOs}}/\text{event} \in [30, 60]$
  - Number of PFOs per jet, $N_{\text{PFOs}}/\text{jet} \in [10, 40]$

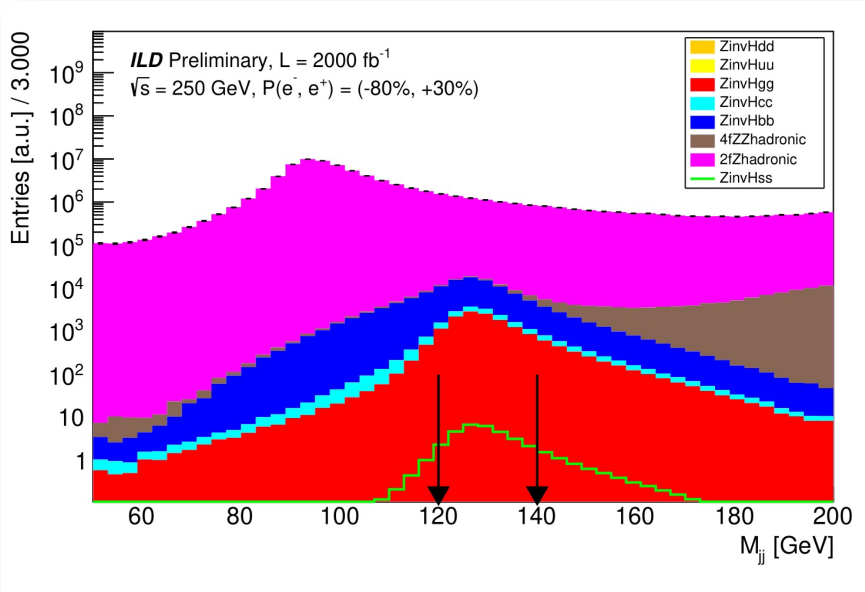
Suggested also to look at the scale at which the event goes from 2->3 / 3->4 jets) to reduce 4-jet events from eg WW, ZZ

Cutflow

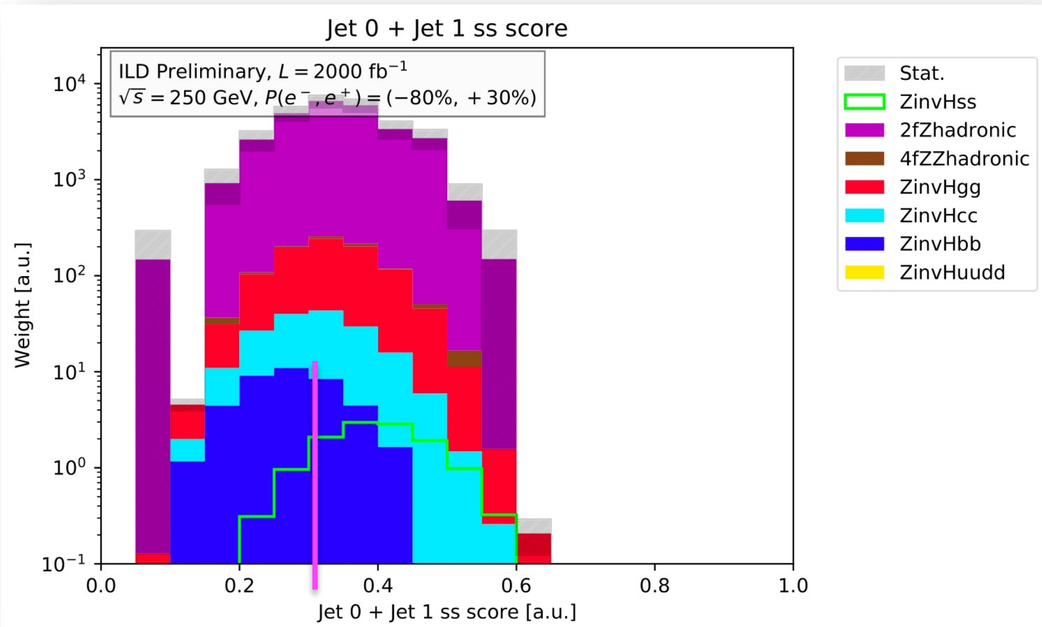
$\textbf{\textit{ILD}}$ Preliminary, $\mathcal{L} = 2000$ fb^{-1}, $\sqrt{s} = 250$ GeV, $P(e^-, e^+) = (-80\%, +30\%)$

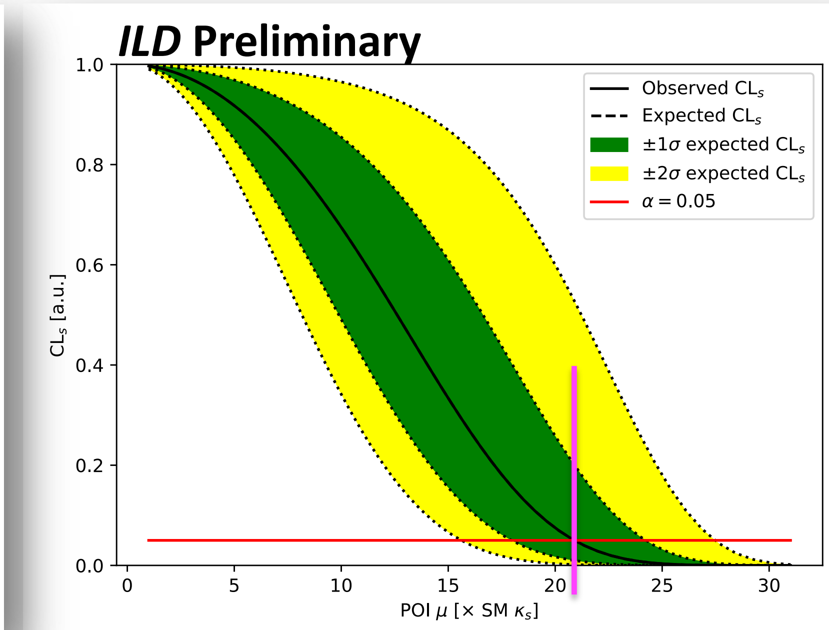
	$(H \to s\bar{s})(Z \to \nu\nu)$	$(H \to gg)(Z \to \nu\nu)$	$(H \to u\bar{u}/d\bar{d})(Z \to \nu\nu)$	$(H \to c\bar{c})(Z \to \nu\nu)$	$(H \to b\bar{b})(Z \to \nu\nu)$	$Z \to q\bar{q}$	$ZZ \to q\bar{q}q\bar{q}$	Sig. eff.	Bkg. eff.
No cut	42.65 ± 0.06	17254.17 ± 24.41	0.59 ± 0.0	5858.77 ± 8.29	116168.67 ± 164.29	$176876516.6 \pm 161411.64$	1342206.08 ± 1338.33	1.00e+00	1.00e+00
No leptons	42.55 ± 0.06	17225.89 ± 24.39	0.59 ± 0.0	5846.08 ± 8.28	115535.31 ± 163.84	$175328405.19 \pm 160703.71$	1335436.33 ± 1334.95	9.98e-01	9.91e-01
≥ 2 jets	42.55 ± 0.06	17225.89 ± 24.39	0.59 ± 0.0	5846.08 ± 8.28	115535.31 ± 163.84	$175328405.19 \pm 160703.71$	1335436.33 ± 1334.95	9.98e-01	9.91e-01
$p_{j0}, p_{j1} > 30$ GeV	39.46 ± 0.06	16424.08 ± 23.81	0.55 ± 0.0	5619.05 ± 8.12	109492.68 ± 159.5	$131310044.43 \pm 139074.89$	1331247.44 ± 1332.86	9.25e-01	7.44e-01
$M_{jj} \in [120, 140]$ GeV	29.75 ± 0.05	12459.56 ± 20.74	0.42 ± 0.0	3883.41 ± 6.75	63849.78 ± 121.8	7424895.55 ± 33070.82	8041.49 ± 103.59	6.97e-01	4.21e-02
$E_{jj} \in [125, 160]$ GeV	29.62 ± 0.05	12401.25 ± 20.69	0.42 ± 0.0	3862.38 ± 6.73	63407.65 ± 121.38	4027593.77 ± 24356.93	6111.86 ± 90.31	6.94e-01	2.31e-02
$M_{\text{miss}} \in [75, 120]$ GeV	27.56 ± 0.05	11614.11 ± 20.02	0.39 ± 0.0	3612.75 ± 6.51	59551.31 ± 117.63	867590.51 ± 11304.65	2105.79 ± 53.01	6.46e-01	5.30e-03
$\Delta R_{jj,\text{miss}} < 4$	23.82 ± 0.05	10039.07 ± 18.62	0.34 ± 0.0	3124.94 ± 6.05	51512.9 ± 109.4	151865.16 ± 4729.65	1537.31 ± 45.29	5.58e-01	1.22e-03
$\text{score}^b/\text{jet} < 0.2$	22.2 ± 0.04	8593.49 ± 17.22	0.32 ± 0.0	1917.39 ± 4.74	551.1 ± 11.32	88968.53 ± 3620.08	689.92 ± 30.34	5.20e-01	5.65e-04
$\text{score}^c/\text{jet} < 0.35$	20.72 ± 0.04	7745.04 ± 16.35	0.3 ± 0.0	302.77 ± 1.88	179.83 ± 6.46	73060.25 ± 3280.5	548.47 ± 27.05	4.86e-01	4.59e-04
$N_{\text{PFOs}}/\text{event} \in [30, 60]$	13.93 ± 0.03	854.7 ± 5.43	0.2 ± 0.0	146.28 ± 1.31	44.14 ± 3.2	33584.15 ± 2224.16	64.05 ± 9.25	3.27e-01	1.95e-04
$N_{\text{PFOs}}/\text{jet} \in [10, 40]$	12.53 ± 0.03	778.96 ± 5.19	0.18 ± 0.0	136.34 ± 1.26	39.96 ± 3.05	26955.7 ± 1992.62	56.05 ± 8.65	2.94e-01	1.57e-04

- Big **decrease** in signal eff. at M_{jj} cut
  - Powerful against $Z \to qq$
- **Net result: ~30% signal efficiency, 0.016% background efficiency**
  - *Sanity check: $H \to bb$ s/b = 0.00065 @ No cut – comparable to T. Ogawa's thesis*



Limits con coupling strength modifier





Cut on (0.5x) sum of strange scores for leading and subleading jets **>0.3**, generated limits for modifier to SM BR

95% upper confidence bounds at **~21 x SM *κs***

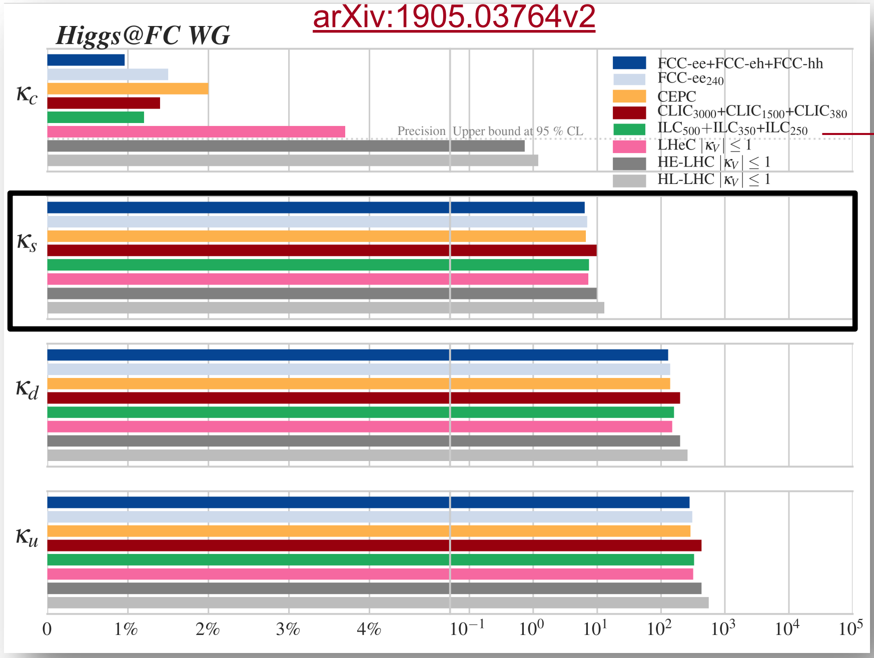
(in the kappa framework, ***κs***^{2} is the modifier to BR[$H \rightarrow ss$])

Asymptotic significance ~ 0.1σ (see extra slides)

Discussion and Outlook (1)

- Discovery measurement seems ***unlikely*** – looking at tagger with truth PID
  - For 30% signal efficiency, need **10,000x** better background rejection
  - Set limits on coupling strength modifier κ_s at **O(20) x SM prediction** using 2000 fb-1 of data at $\sqrt{s}$ = 250 GeV
- Sensitivity is ***limited*** but more promising than previous projections derived from unitarity bounds or exclude Higgs decays

arXiv:1905.03764v2



Based on all the ILC stats and indirect measurements, while we considered only the 250 GeV scenario and, for the first time, a direct analysis!

Discussion and Outlook (2)

- Gains would come from **reducing the $Z \rightarrow qq$ background**
  - Will exploit angular variables between jets
  - More statistics are available for the $Z \rightarrow qq$ and $Z \rightarrow qqqq$ backgrounds
  - Samples being produced for $Z(ll)H(qq)$

- In contact with Whizard Generator experts to try and provide prospects for **BSM 2HDM $H \rightarrow cs$ or $H(125) \rightarrow bs$ decays**

- Studies based on ILD samples, but tagger and analysis strategy as well as information on detector capabilities is of **vast applicability to other e+e- machines**
  - You are welcome to join the ILCX discussion!
  - We did notice the recent presentation on strange tagging @ FCC-ee

- Work being documented in a paper as part of **Snowmass 2021**

Thanks for your attention!

Valentina Cairo

Extra Slides

ILC https://linearcollider.org/files/images/pdf/Executive%20Summary.pdf

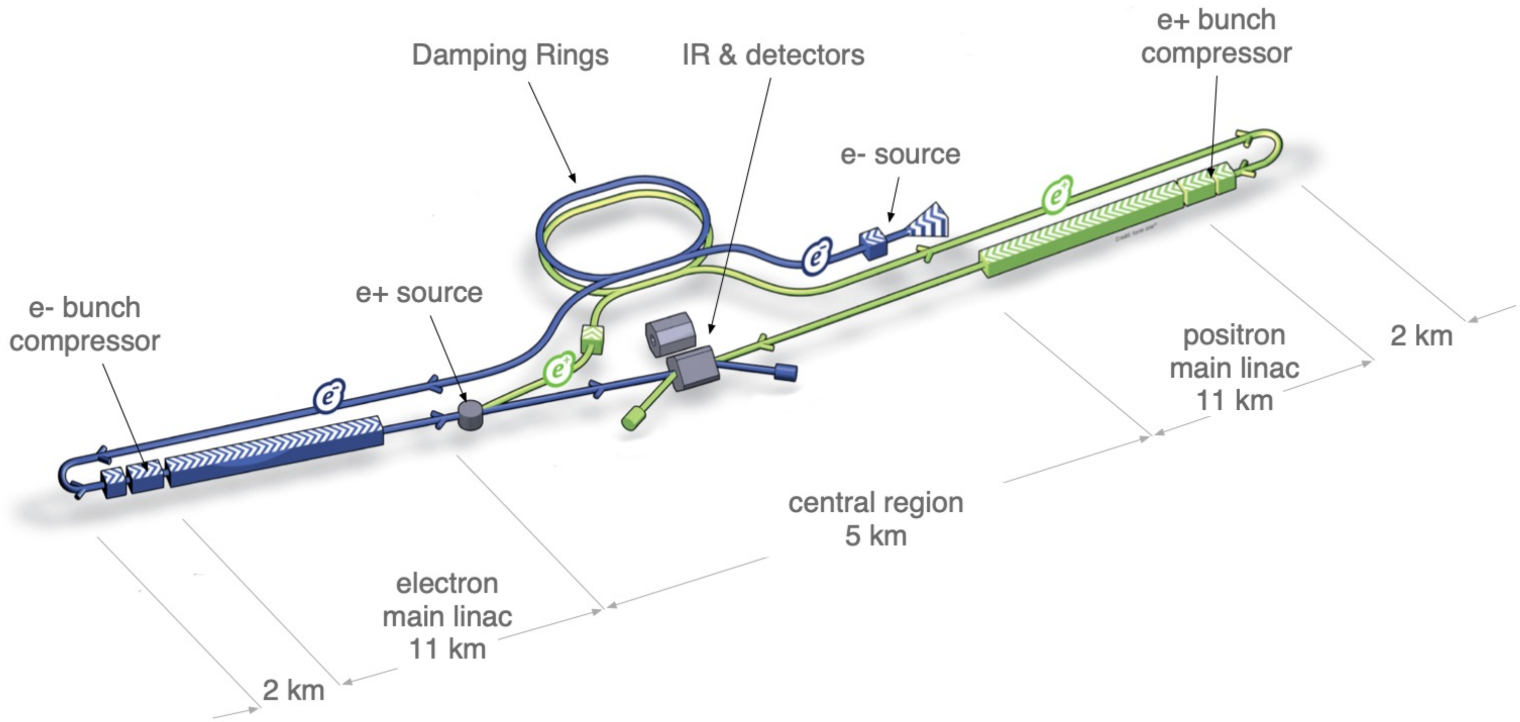


Figure 3.1. Schematic layout of the ILC, indicating all the major subsystems (not to scale).

ILC https://arxiv.org/pdf/1908.11299.pdf

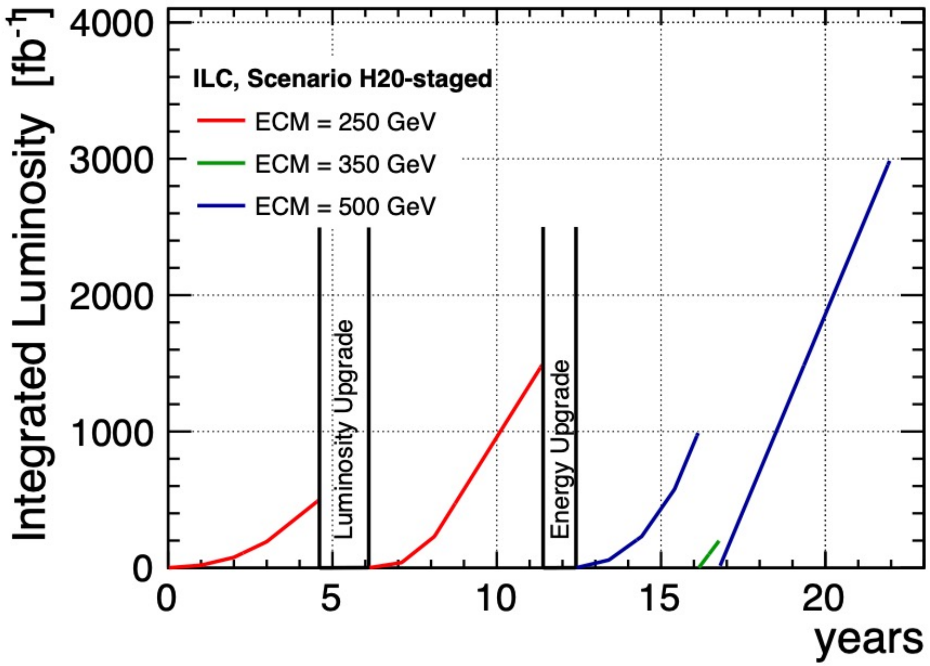


Figure 1: The nominal 22-year running program for the staged ILC, starting operation at 250 GeV [8].

ILC https://arxiv.org/pdf/1908.11299.pdf

The current proposed run plan for the ILC raises the CM energy in stages, with runs at 250 GeV, 350 GeV, and 500 GeV. It is also possible to run the ILC at the Z pole with minimal modification. By lengthening the ILC tunnel, improving the gradient of the superconducting RF cavities, or a combination of these, it is possible to run the ILC at a CM energy of 1 TeV. The luminosity of a linear collider naturally rises approximately linearly with CM energy, making it easier to acquire larger luminosity samples as the energy is increased.

	E_{CM} (GeV)	$\int \mathcal{L}$ (fb^{-1})	fraction with $\mathrm{sign}(P(e^-), P(e^+)) =$ (−+)	(+−)	(−−)	(++)
ILC250	250	2000	45%	45%	5%	5%
ILC350	350	200	67.5%	22.5%	5%	5%
ILC500	500	4000	40%	40%	10%	10%
GigaZ	91.19	100	40%	40%	10%	10%
ILC1000	1000	8000	40%	40%	10%	10%

Table 1: CM energy, integrated luminosity, and polarisation fractions for the stages of ILC discussed in this report. In all cases, the magnitude of the e^- polarisation is taken to be 80% and the magnitude of the e^+ polarisation is taken to be 30%, except that, at ILC1000, 20% e^+ polarisation was used in the studies quoted.

ILC https://linearcollider.org/files/images/pdf/Executive%20Summary.pdf

Table 2.1
Major physics processes to be studied by the ILC at various energies. The table indicates the various Standard Model reactions that will be accessed at increasing collider energies, and the major physics goals of the study of these reactions. A reaction listed at a given energy will of course be studied at all higher energies.

Energy	Reaction	Physics Goal
91 GeV	$e^+e^- \to Z$	ultra-precision electroweak
160 GeV	$e^+e^- \to WW$	ultra-precision W mass
250 GeV	$e^+e^- \to Zh$	precision Higgs couplings
350–400 GeV	$e^+e^- \to t\bar{t}$	top quark mass and couplings
	$e^+e^- \to WW$	precision W couplings
	$e^+e^- \to \nu\bar{\nu}h$	precision Higgs couplings
500 GeV	$e^+e^- \to f\bar{f}$	precision search for Z'
	$e^+e^- \to t\bar{t}h$	Higgs coupling to top
	$e^+e^- \to Zhh$	Higgs self-coupling
	$e^+e^- \to \tilde{\chi}\tilde{\chi}$	search for supersymmetry
	$e^+e^- \to AH, H^+H^-$	search for extended Higgs states
700–1000 GeV	$e^+e^- \to \nu\bar{\nu}hh$	Higgs self-coupling
	$e^+e^- \to \nu\bar{\nu}VV$	composite Higgs sector
	$e^+e^- \to \nu\bar{\nu}t\bar{t}$	composite Higgs and top
	$e^+e^- \to \tilde{t}\tilde{t}^*$	search for supersymmetry

Timeline

https://arxiv.org/pdf/1905.03764.pdf

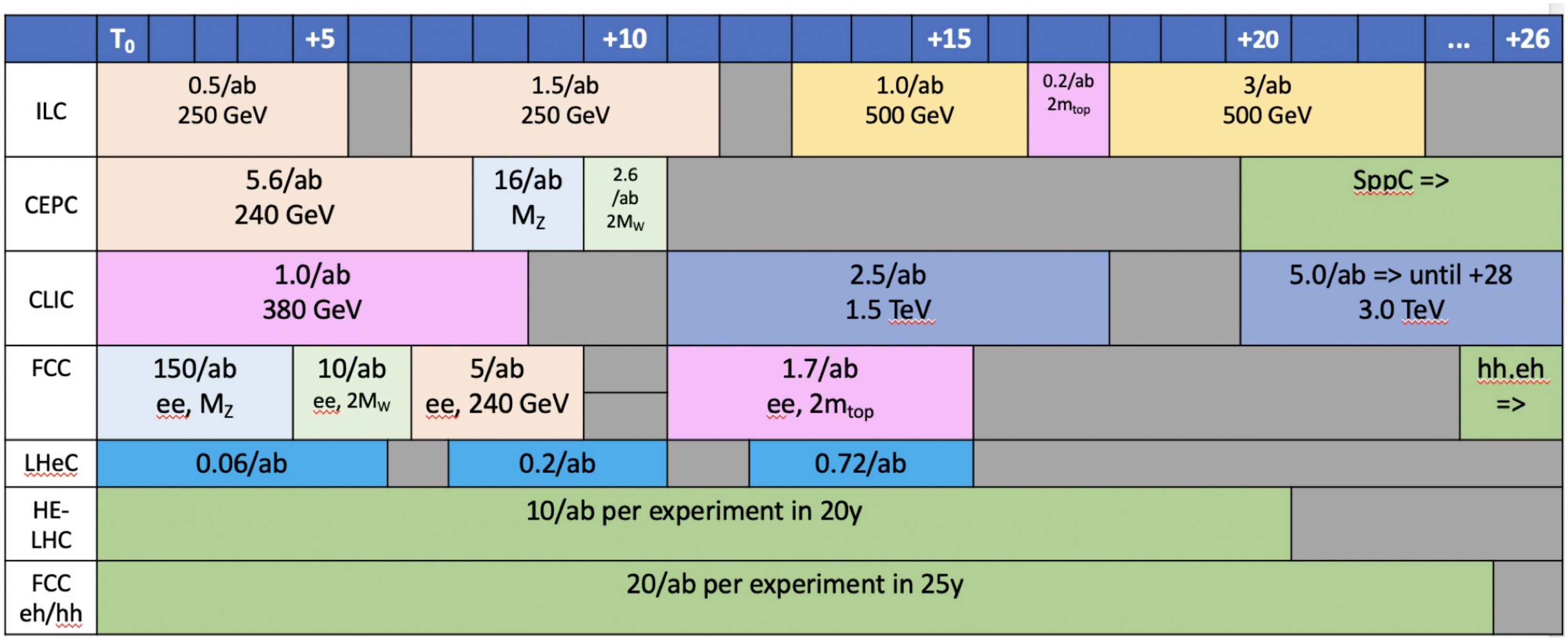
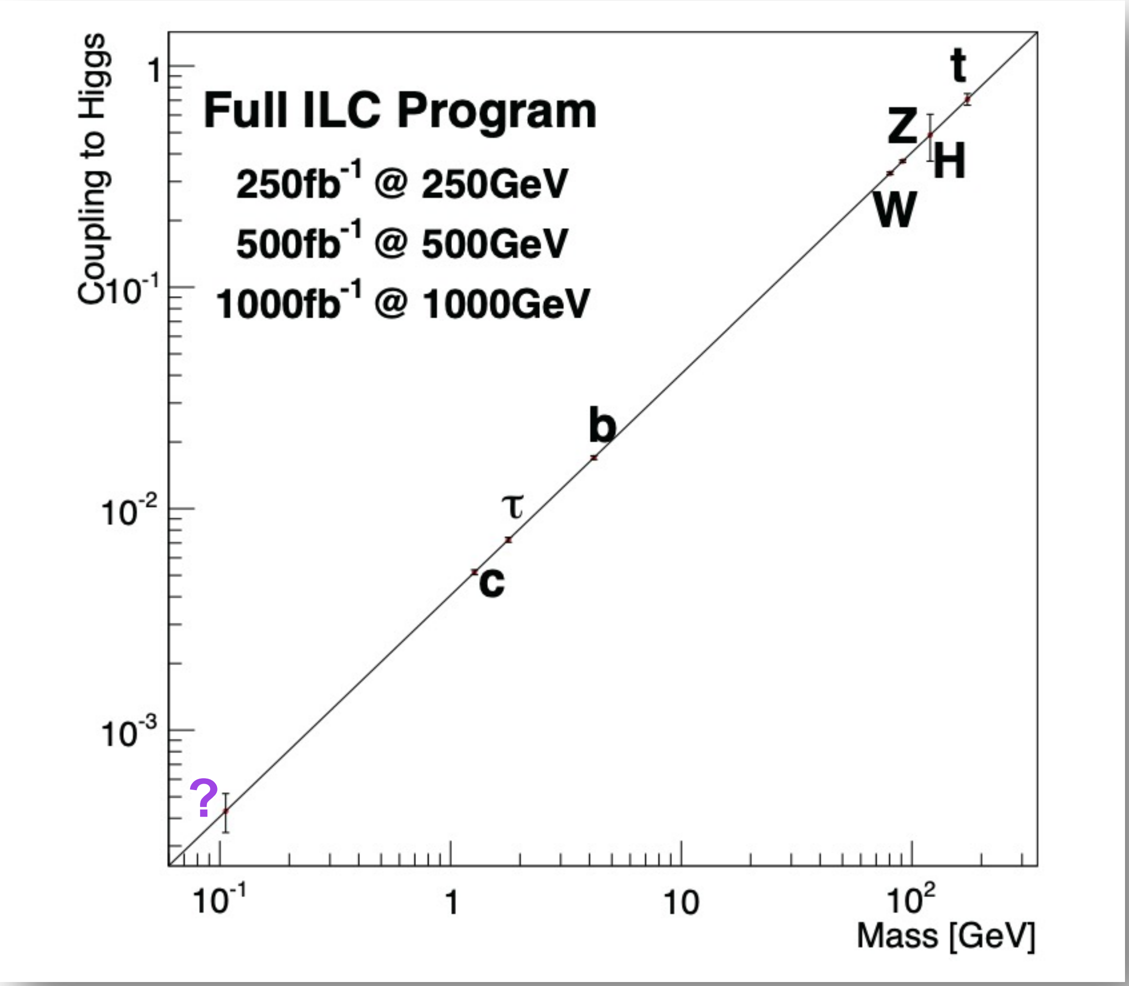
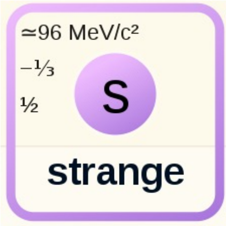


Figure 1. Time line of various collider projects starting at time T_0 as submitted to the European Strategy Update process. Some possible extensions beyond these baseline run plans have been discussed and are presented in more detail in Appendix F. For the clarification of the meaning of a year of running, see the caption of Table 1. Figure 13 in Appendix C shows an alternative version of this figure using the earliest possible start date (i.e. the calendar date of T_0) given by the proponents.

Higgs couplings

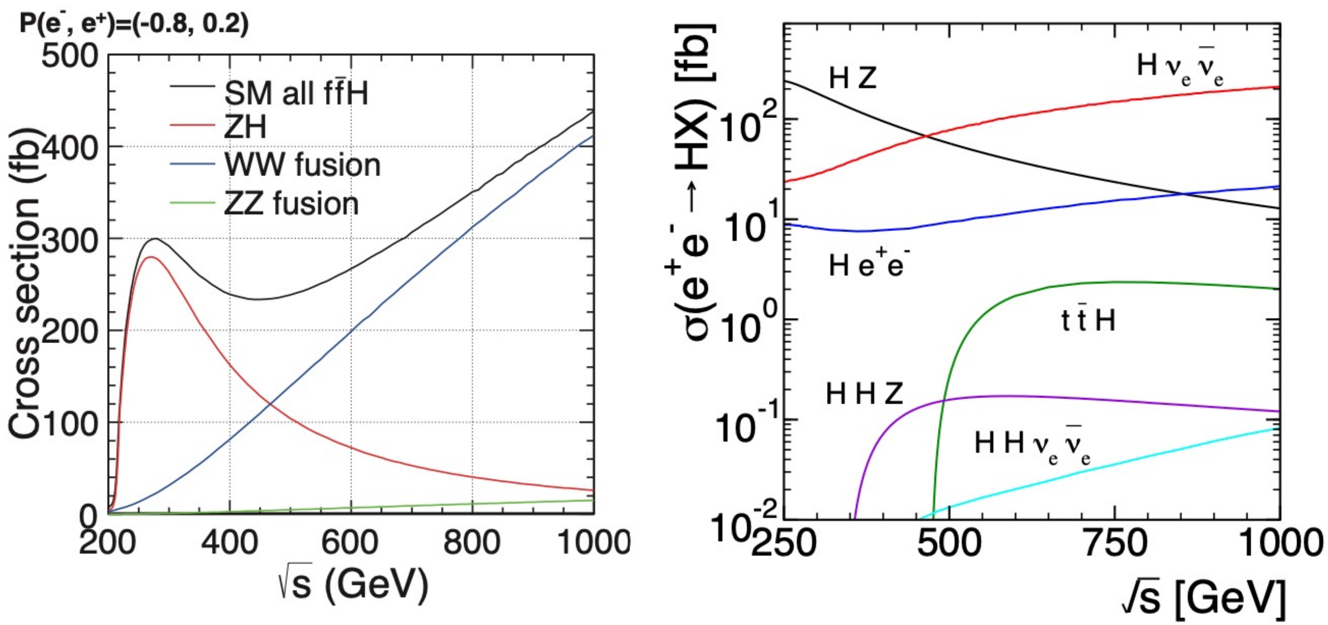
https://arxiv.org/pdf/1310.0763.pdf





Production Cross Section

https://arxiv.org/pdf/1310.0763.pdf



~200fb

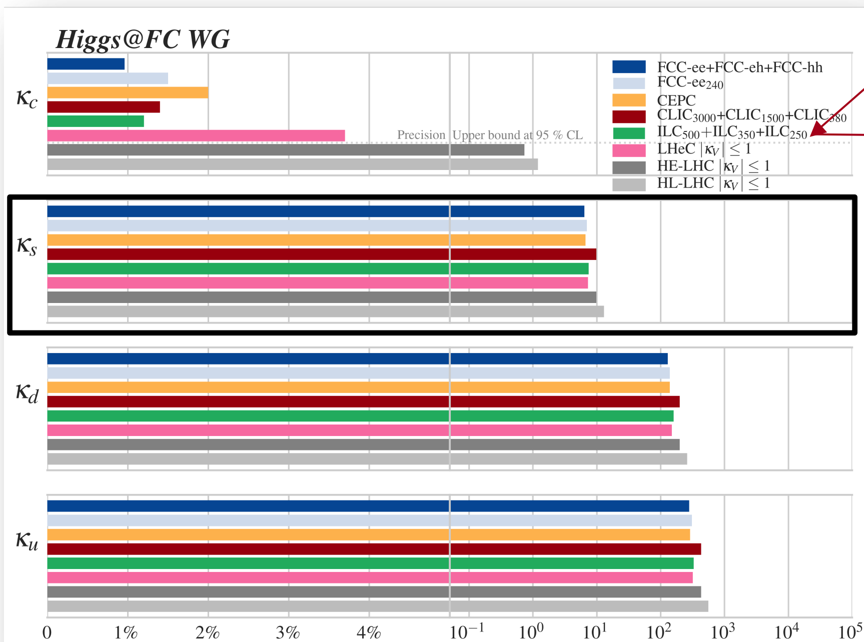
If we consider 2000 fb-1 data after 10 years, we have 400k Higgs out of which only 40 will decay to ssbar

Figure 1.4. (Left)The production cross sections of the Higgs boson with the mass of 125 GeV at the ILC as a function of the collision energy $\sqrt{s}$. Polarization of the electron beam (80%) and the positron beam (20%) is assumed. (Right) The cross sections of the production processes $e^+e^- \to hZ$, $e^+e^- \to H\nu_e\bar{\nu}_e$, $e^+e^- \to He^+e^-$, $e^+e^- \to t\bar{t}H$, $e^+e^- \to HHZ$ and $e^+e^- \to HH\nu_e\bar{\nu}_e$ as a function of the collision energy for the mass of 125 GeV. No polarization is assumed for the initial electron and positron beams.

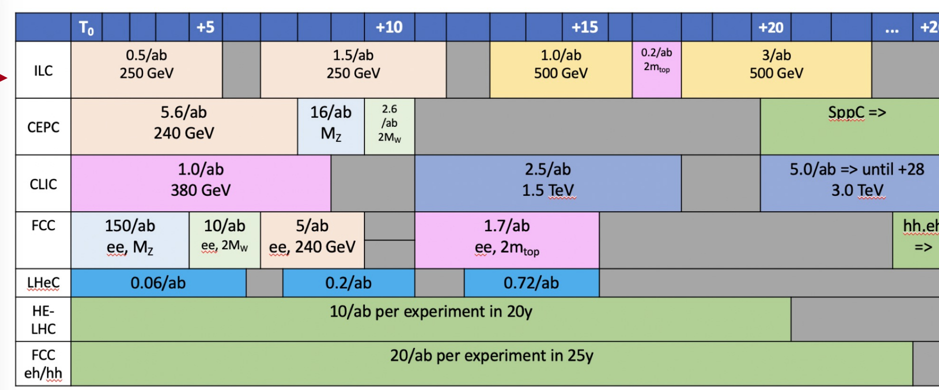
Comparison with existing projections

- Discovery measurement seems ***unlikely***, as expected, even after using truth info in the tagger
- So we set limits at O(10) xSM

arXiv:1905.03764v2

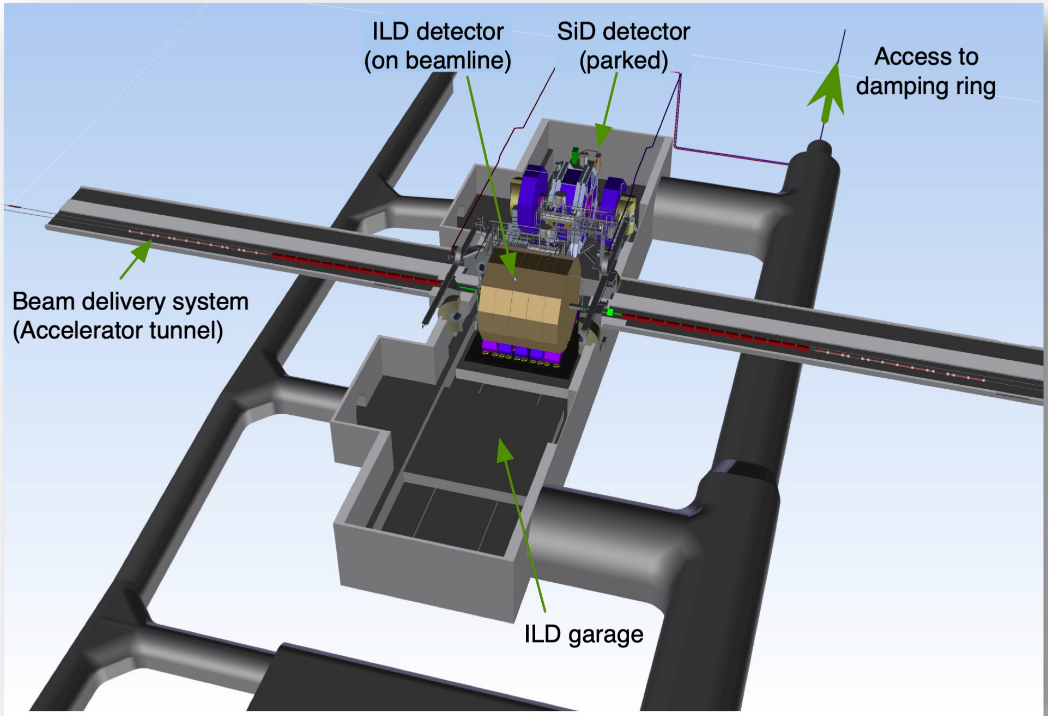


Using all ILC stats, while we use only the first 10 years!



ILD & SiD

https://linearcollider.org/files/images/pdf/Executive%20Summary.pdf

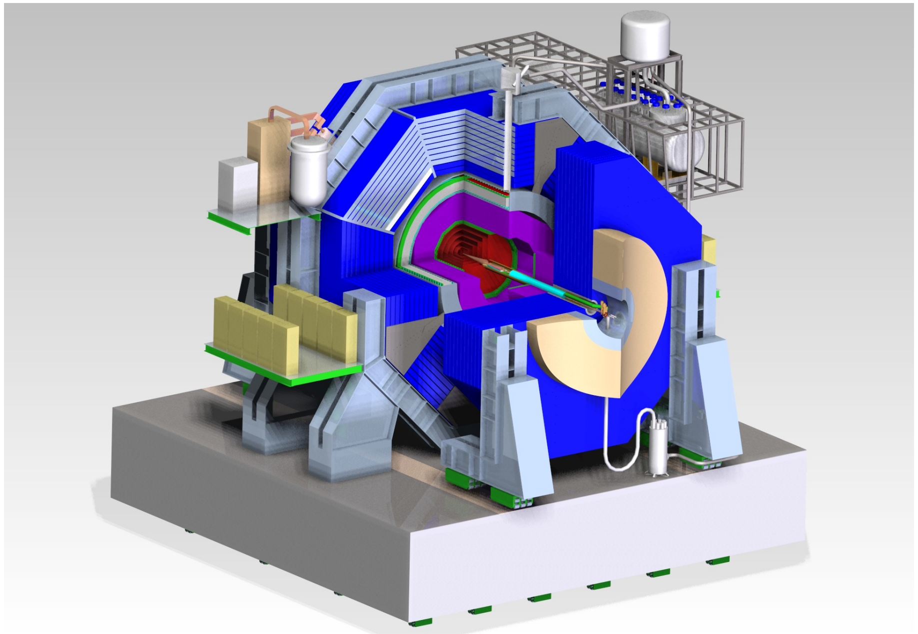


- **SiD** is a compact, cost-constrained detector made possible with a **5 Tesla** magnetic field and silicon tracking. Silicon enables time-stamping on single bunch crossings to provide robust performance, derived from immunity to spurious background bursts. The highly granular calorimeter is optimised for particle-flow analysis.
- The **ILD** group has designed a large detector with robust and stable performance over a wide range of energies. The concept uses a tracking system based on a continuous-readout time-projection chamber combined with silicon tracking for excellent efficiency and robust pattern-recognition performance. A granular calorimeter system contained inside a **3.5 T** magnetic field provides very good particle-flow reconstruction.

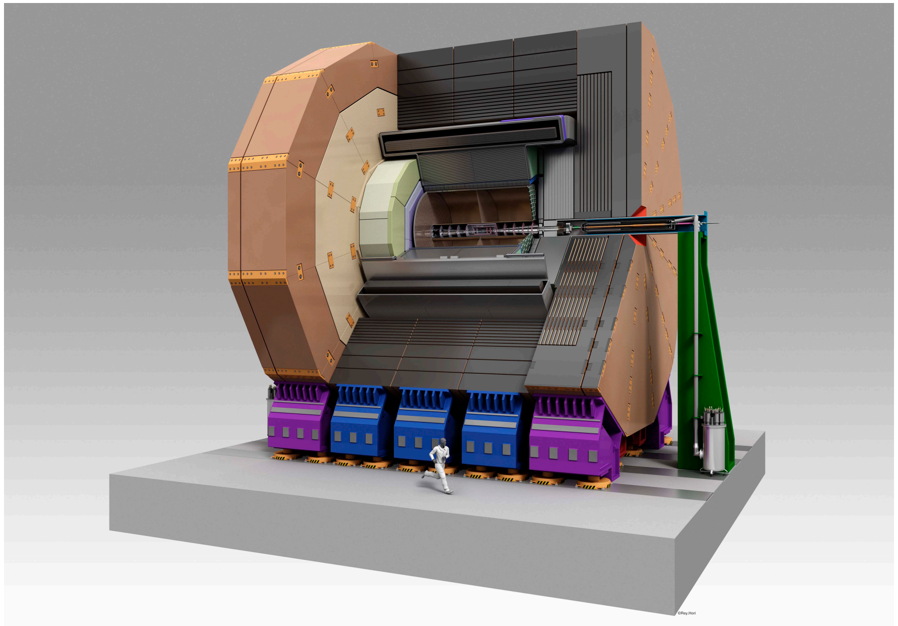
ILD & SiD

https://linearcollider.org/files/images/pdf/Executive%20Summary.pdf

SiD

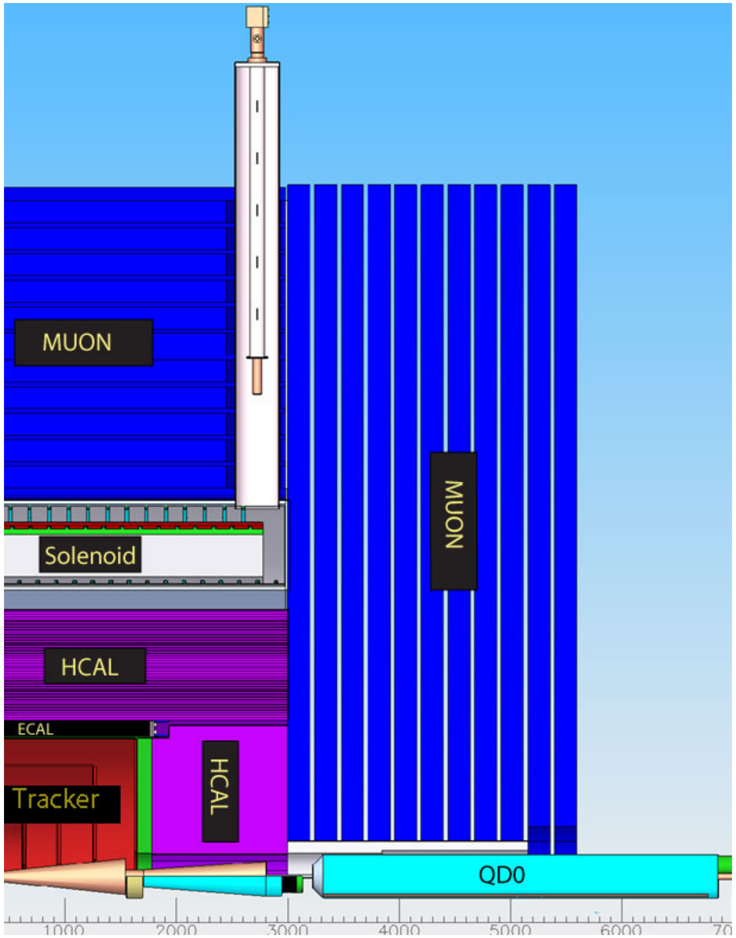


ILD

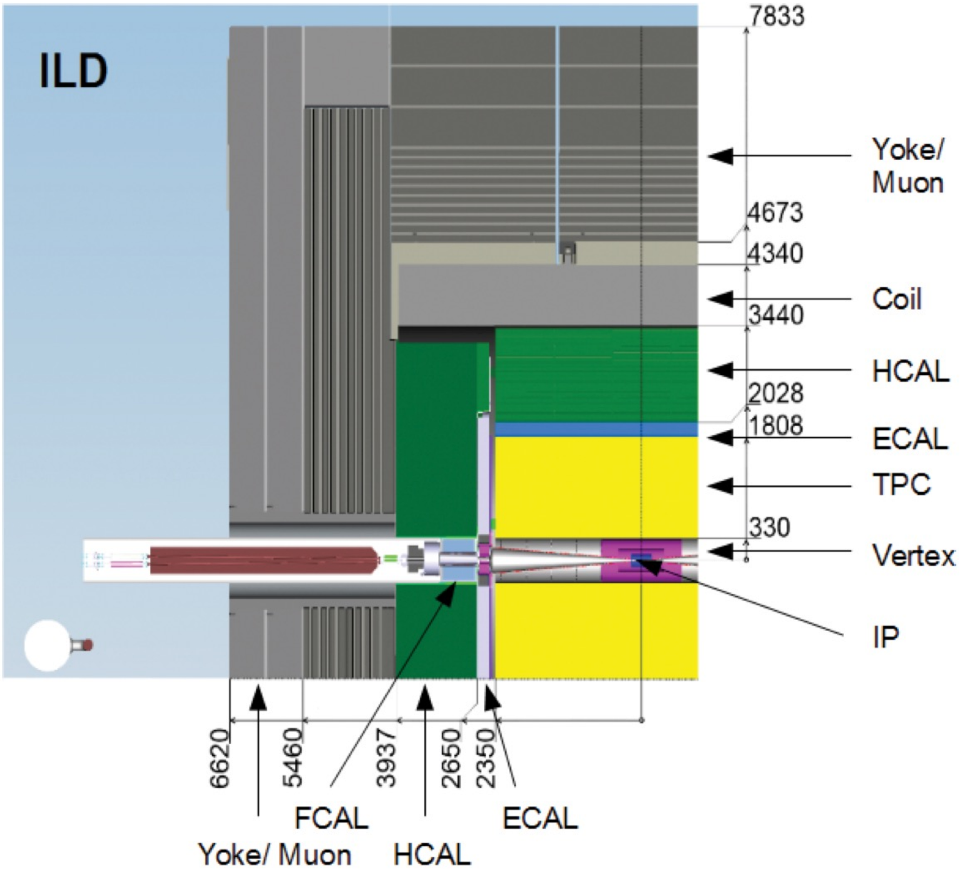


ILD & SiD https://linearcollider.org/files/images/pdf/Executive%20Summary.pdf

SiD



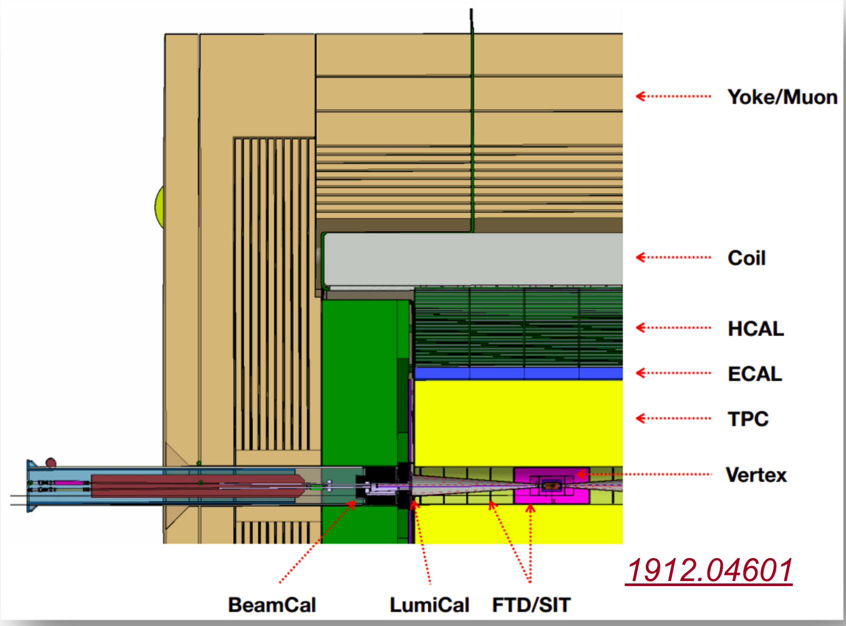
ILD



ILD @ ILC: proposed detector layout

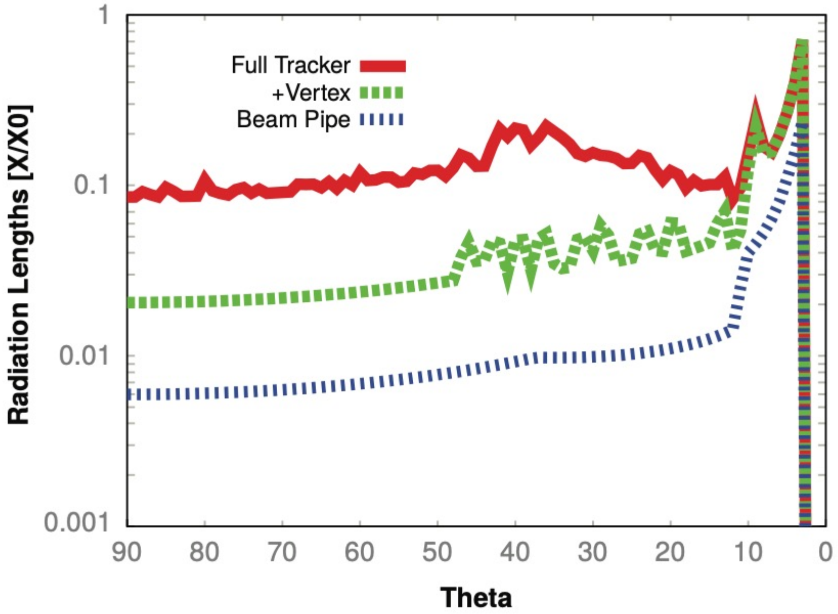
International Large Detector

- 3 double-layer pixel detectors for vertexing
- Time projection chamber (TPC) for tracking with inner/outer Si layers
  - Low material assists in low-p tracking
- High granularity sampling calorimeters for particle flow reconstruction
  - Challenge is reconstructing neutral hadrons
  - Precise EM/hadronic design still under study
- Tracking/calorimetry contained in 3.5 T field

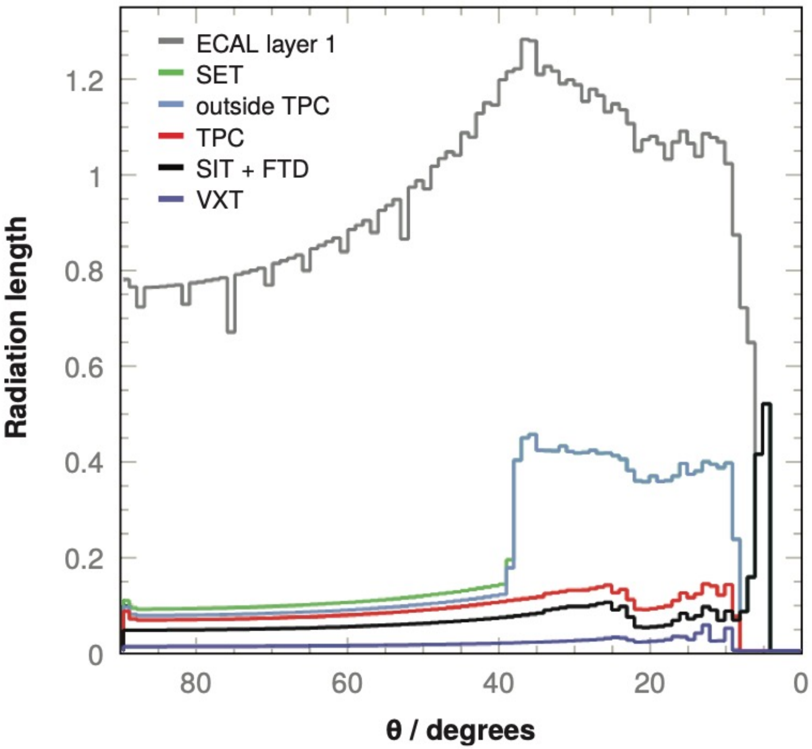


ILD & SiD https://linearcollider.org/files/images/pdf/Executive%20Summary.pdf

SiD

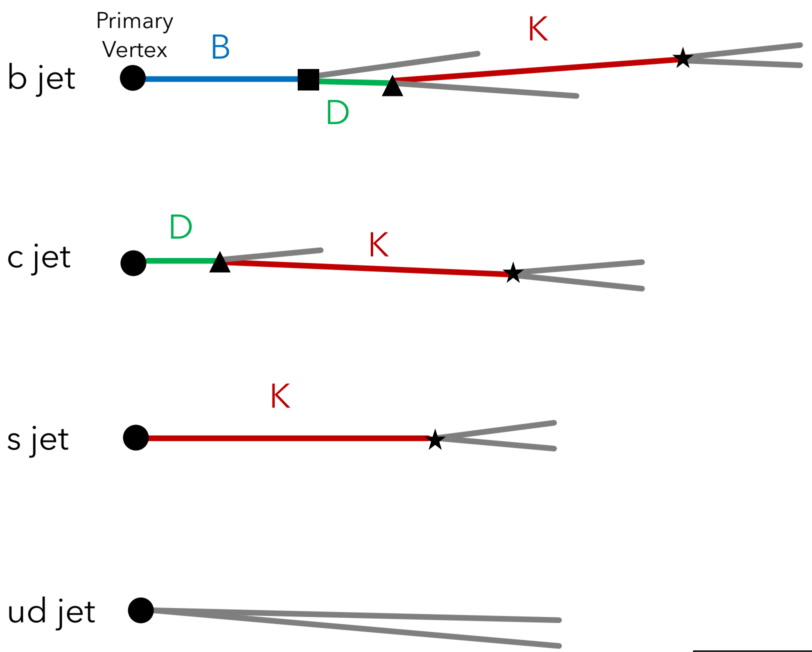


ILD



Jet Types

Categories



Under perfect reconstruction, we only need to count:

	# of secondary vertices (excluding V^0)	# of strange hadrons ($K^\pm$, K_L^0, K_S^0, Λ^0)
b	2	≥1
c	1	≥1
s	0	≥1
ud	0	0

In reality: missing tracks / fake tracks → migration

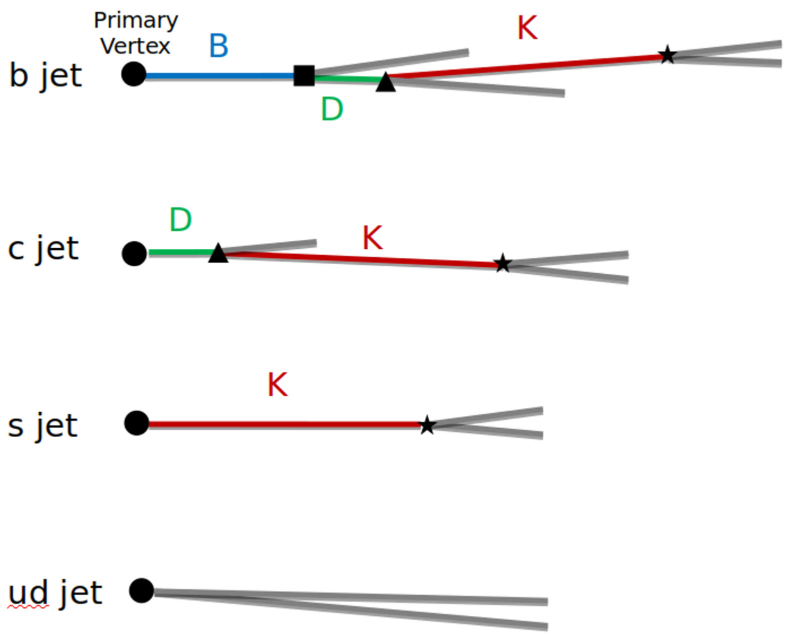
→ These provide natural categories for multivariate classification.

Taken from Slide 5 of Tomohiko Tanabe's 2020/11/24 presentation.

4

Jet Types

Discriminants



Charged Kaon track
- Zero track impact parameter w.r.t. primary vertex
- Momentum fraction relative to the jet momentum carried by the leading Kaon
  - (Longitudinal vs transverse components?)

V⁰ (K_S^0,Λ^0)
- Vertex momentum & displacement must point in the same direction
- Mean vertex distance smaller compared to b/c

+ the usual b/c discriminants (vertex mass, impact parameter for all tracks, etc.)

Remember to normalize the discriminants to make them boost invariant (as much as possible)

Taken from Slide 5 of Tomohiko Tanabe's 2020/11/24 presentation.

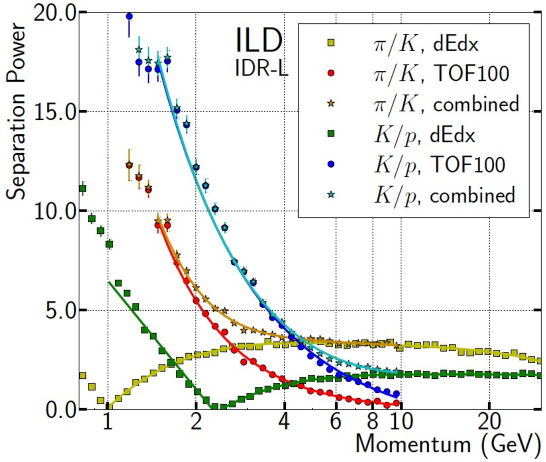
5

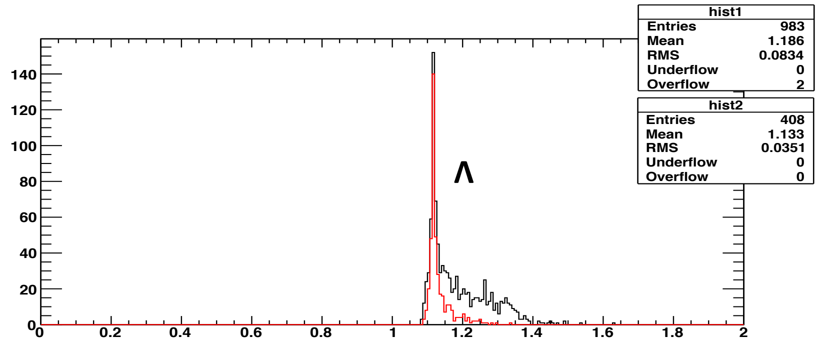
Jet Types

Strange Hadron Reconstruction/Selection

Strange Hadron reconstruction
- $K^{\pm}$ [PID]
- $K_S^0 \rightarrow \pi^+\pi^-$ [Vertex] (BF ~69.2%)
- $\Lambda^0 \rightarrow p\pi^-$ [Vertex] (BF ~64%)
- K_L^0 [Particle Flow]

OPEN QUESTIONS
- Pion/Kaon separation:
  - With and without timing
- K_S & Λ reconstruction: efficiency/purity?
  - Does proton tag help Λ reconstruction?
- Neutron/K_L separation (probably not)
- Best strange hadron selection:
  - Lists ordered by purity, or likelihood in terms of PID/mass?
- *[Do we want perfect (cheated) reconstruction?]*





6

Taken from Slide 5 of Tomohiko Tanabe's 2020/11/24 presentation.

Experimental Handles for Strange Tagging

- The goal is to **discriminate strange jets from u/d jets** (discrimination from c/b happens through c/b tagging)
- The challenge is that **strange hadrons** are certainly **produced from the fragmentation of strange quarks**, but they are **also produced frequently in the fragmentation of u/d quarks** (typically with a lower fraction of the jet's transverse momentum x)
  - Identical QCD and electromagnetic interactions, but different hadronization and subsequent decay processes
  - The main idea behind strange taggers is that strange quarks mostly hadronize to prompt kaons that carry a large fraction of the jet momentum
  - A strange-quark jet contains on average a **higher ratio of neutral kaon energy** (more energy in the HCal) **to neutral pion energy** (more energy in the ECal) wrt a down-quark jet
    - One of the main handles used in arxiv:2003.09517v1 (presented in EF2)

- It is clear that **PID capabilities** to discriminate between kaons and pions would be very helpful in building a solid strange tagger
  - Lessons can be learned from the SLD experience in tagging strange jets from Z decays (Cherenkov detector with k/pi separation at all momenta)
  - PID has been considered also in arxiv:1811.09636 (see extra slides for some details)

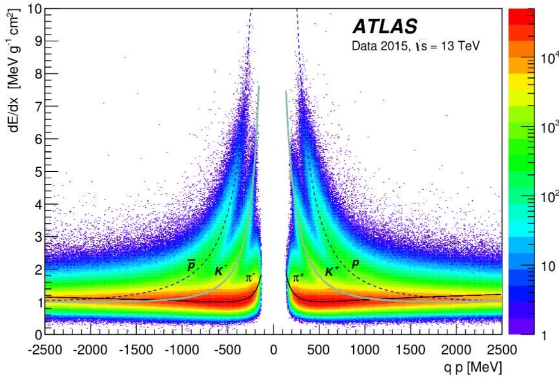
Particle Identification

- **Existing strange tagging** studies suffer from **low efficiency and very large mis-tag probability** from u and d quarks, even when using sophisticated machine learning algorithms

- To complement existing studies (more on this in the next slides), we thought we would put **more emphasis on exploiting Particle Identification**

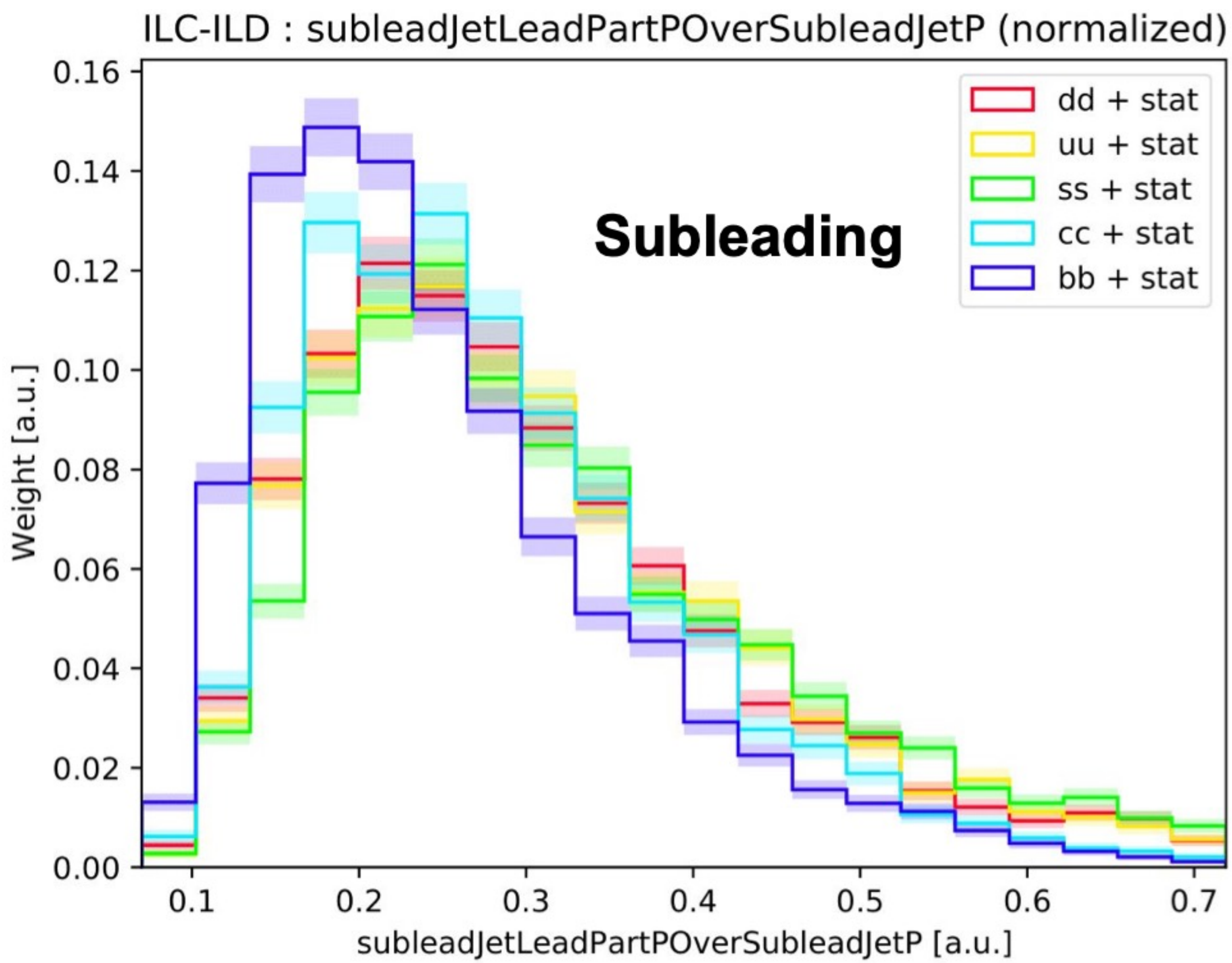
  - This implies looking at new detector concepts
  - Current general purpose detectors use the well known **d*E*/dx dependence on βγ**, but this only allows to get to good PID up to ~1 GeV
  - Alternatively, as foreseen for the HL-LHC detectors, **timing information** can be used to deduce a **velocity** that, in combination with the standard measurement of **momentum from track curvature in the magnetic field**, yields a measure of the **charged particle mass.**
  - Another very effective way to achieve particle identification is through **Cherenkov detectors**, as done in the ALICE and LHCb experiments at the LHC

Phys. Rev. D 93, 112015 (2016)



In either scenario, **kaons with at least 20% of the jet's transverse momentum x** will have to be identified in order for this to be relevant for strange tagging.

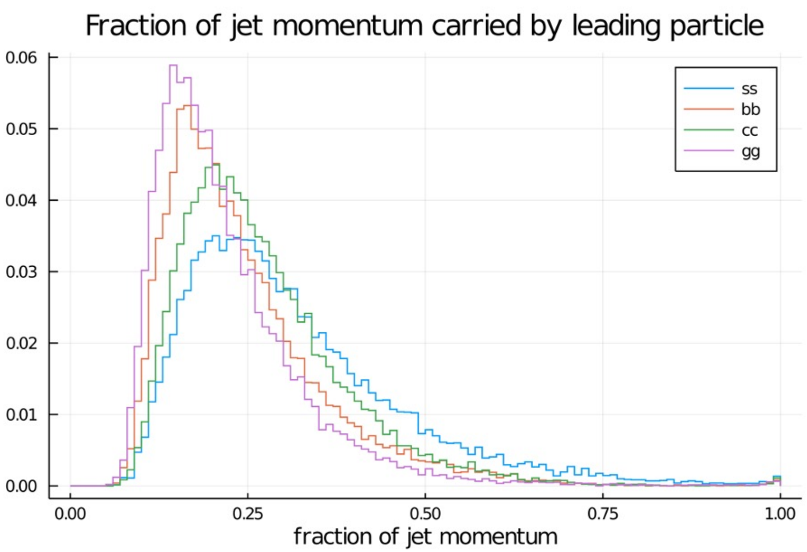
ILD @ ILC: performance

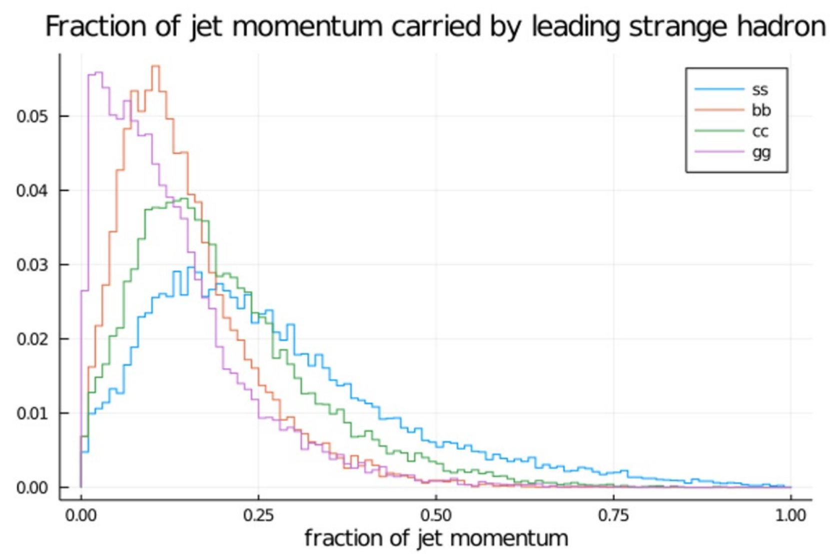


J. Strube's studies:
https://indico.slac.stanford.edu/event/6674/#2-summary-of-existing-studies

Fragmentation properties

Using perfect PID (MCParticle)





Pythia standalone generation

https://indico.slac.stanford.edu/event/6674/contributions/1581/attachments/737/2070/201124_Strange_Tagging_Meeting_Physics_Studies.pdf#page=13

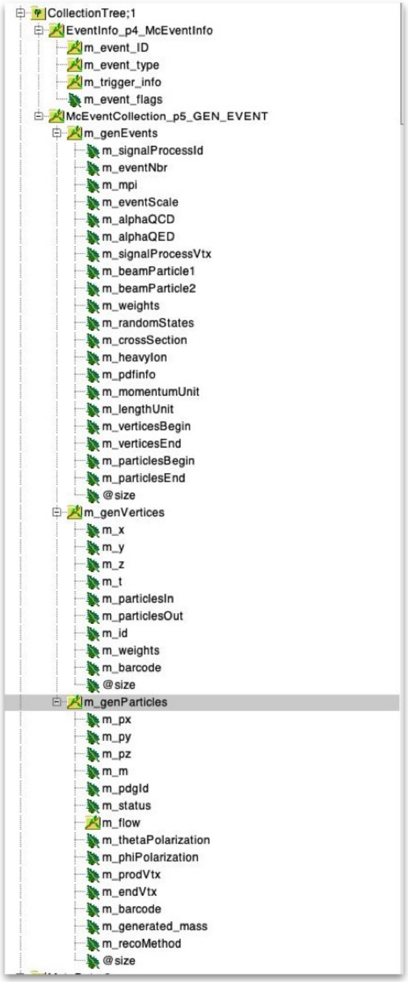
(in the meanwhile...)

Standalone Py8 generation

```
include('Pythia8_i/Pythia8_A14_NNPDF23LO_EvtGen_Common.py')

genSeq.Pythia8.Commands += [ 'HiggsSM:all = on',
                              '25:m0 = 125.',
                              '25:onMode = off',    switches off all H0 decay modes
                              '25:onIfAll = 3 -3']  switches back on the strange decays
```

- Used a small Rivet routine to analyse the generated events (thanks to the ATLAS expert D.Kar!)
  - The goal would be to understand simply how different an H->ss event is from any other H->jj at generation level, before even getting to detector limitations
  - Other existing ATLAS Internal work:
    - Flavor of leading hadrons in strange jets (Pythia 8):
      - ≈ 39% up/down
      - ≈ 61% strange
      - ≈ 0.5% charm
      - ≈ 0.05% bottom
    - The most discriminant variable remains the fraction of energy deposited in the EM calorimeters that is larger for d-jets
    - This actually showed up today on the arxiv :-(https://arxiv.org/pdf/2011.10736.pdf



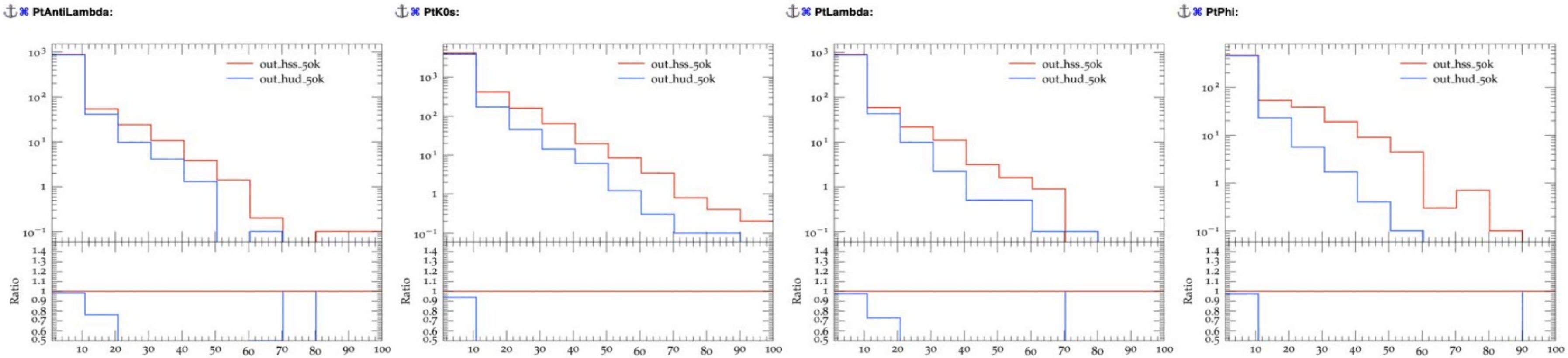
13

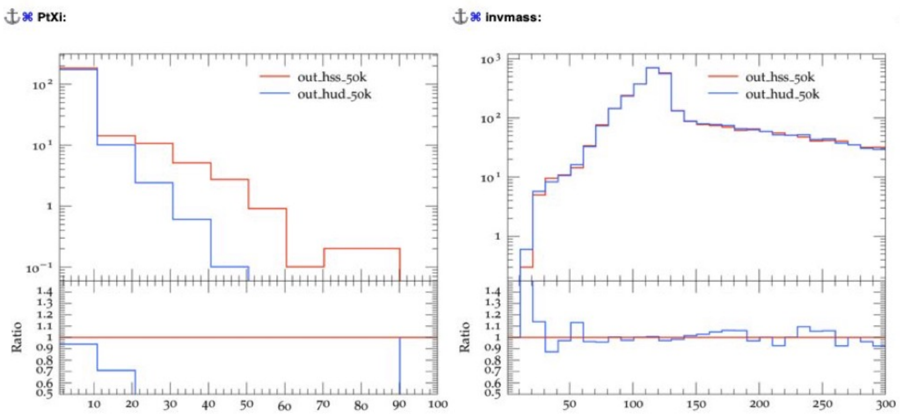
Pythia standalone generation

https://indico.slac.stanford.edu/event/6674/contributions/1581/attachments/737/2070/201124_Strange_Tagging_Meeting_Physics_Studies.pdf#page=13

(in the meanwhile...)

Standalone Py8 generation





Looking at the the pT of the strange hadrons in H->ss and H->u/d events (di-jet mass window 100<mjj<150)

Strange hadrons in H->ss are generally more energetic than in H->u/d events, as expected

14

Existing study by J. Duarte-Campderros, G. Perez, M. Schlaffer, and A. Soffer

https://arxiv.org/pdf/1811.09636.pdf

In this letter we propose using a strangeness tagger, inspired by Z → ss measurements at the DELPHI [24] and SLD [25] experiments, for probing the Higgs coupling to the strange quark via the h → ss decay.

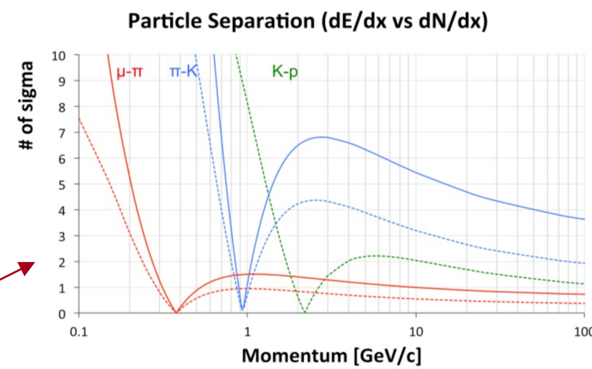
We limit the discussion to lepton colliders, having in mind the International Linear Collider (ILC) [26], the Future Circular Collider in electron mode (FCC-ee) [27], and the Circular Electron Positron Collider (CEPC) [28]. These proposed colliders would run at a center-of-mass energy $\sqrt{s}$ = 250 GeV, where the cross section for associated Higgs production, $e^+e^- \to Zh$, peaks at $\sigma \approx 210$ fb for unpolarized beams [29, 30]. We therefore adopt this scenario for this letter.

hadrons remains. To exploit this, we define a new jet-flavor variable,

$$J_F = \frac{\sum_H \mathbf{p}_H \cdot \hat{\mathbf{s}} R_H}{\sum_H \mathbf{p}_H \cdot \hat{\mathbf{s}}}. \qquad (4)$$

Here, the sum is over all hadrons inside the jet, $\mathbf{p}_H$ is the momentum vector of the hadron H, R_H is its quantum number or numbers in the flavor representation of interest, and $\hat{\mathbf{s}}$ is the normalized jet axis. In our case of a Higgs boson produced approximately at rest and undergoing a $h \to jj$ decay, $|\sum \mathbf{p}_H| \approx m_h/2$, so that the denominator is nonzero, and J_F is well defined. Here and in the following we use $h \to jj$ to denote a Higgs decay into a final state of two gluons or a quark–anti-quark pair.

We further take into account the possibility that the detector has a particle identification (PID) capability, to discriminate pions from kaons. For concreteness we adopt the PID capability of the IDEA drift chamber with cluster counting, which can separate pion from kaon tracks by more than 5 standard deviations in the relevant momentum range [38].



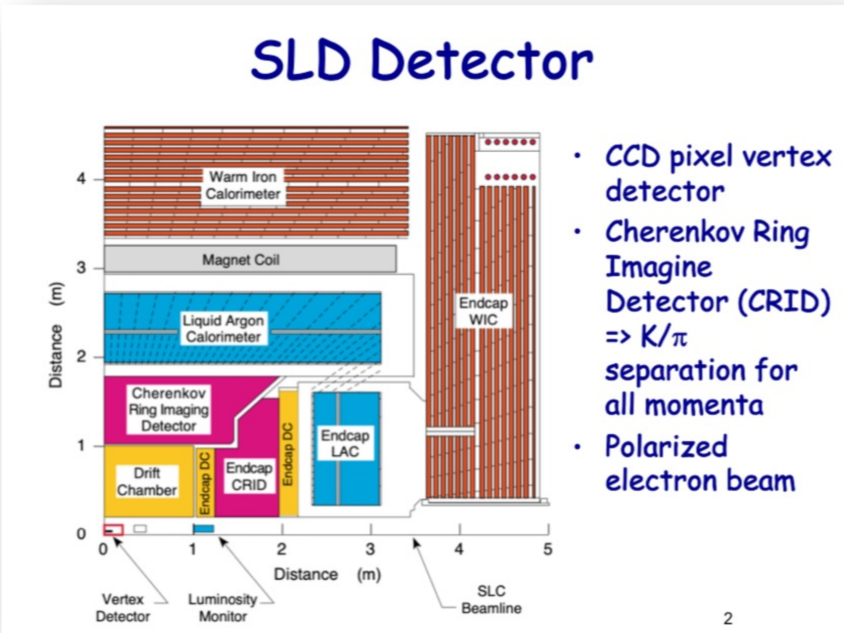
https://link.springer.com/article/10.1140/epjst/e2019-900045-4

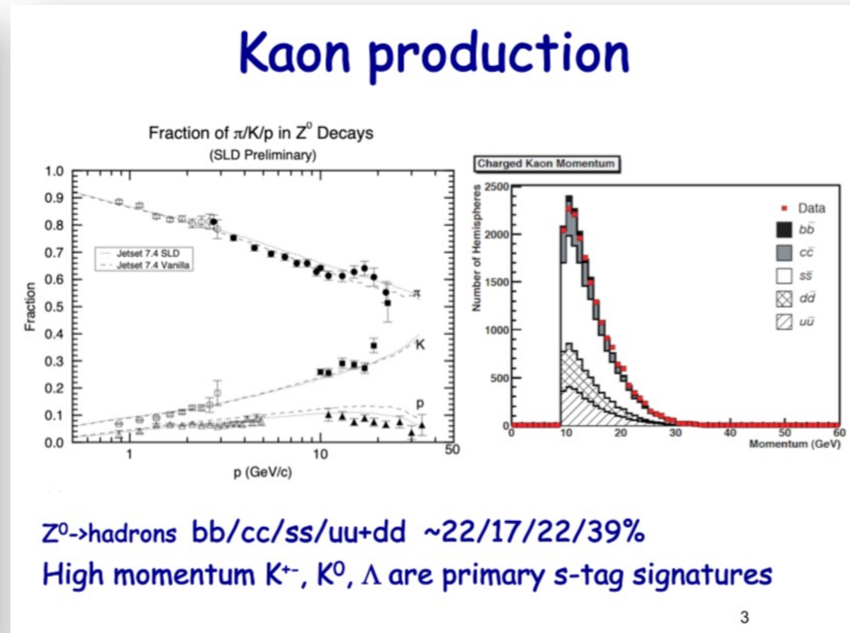
We obtain an upper limit on the signal strength of μss < 14 and < 7 for integrated luminosities of 5 and 20 ab^{-1}, respectively. The limit is weakened to μss < 60 for an integrated luminosity of 250 fb^{-1}.

Can these limits be improved (and eventually make this decay mode accessible) with new detector features? How can we better complement these studies? What would be a nice flavor violating Higgs decay benchmark to target?

SLD

https://indico.slac.stanford.edu/event/6617/contributions/1443/attachments/683/1978/s-tag-SLD.pdf



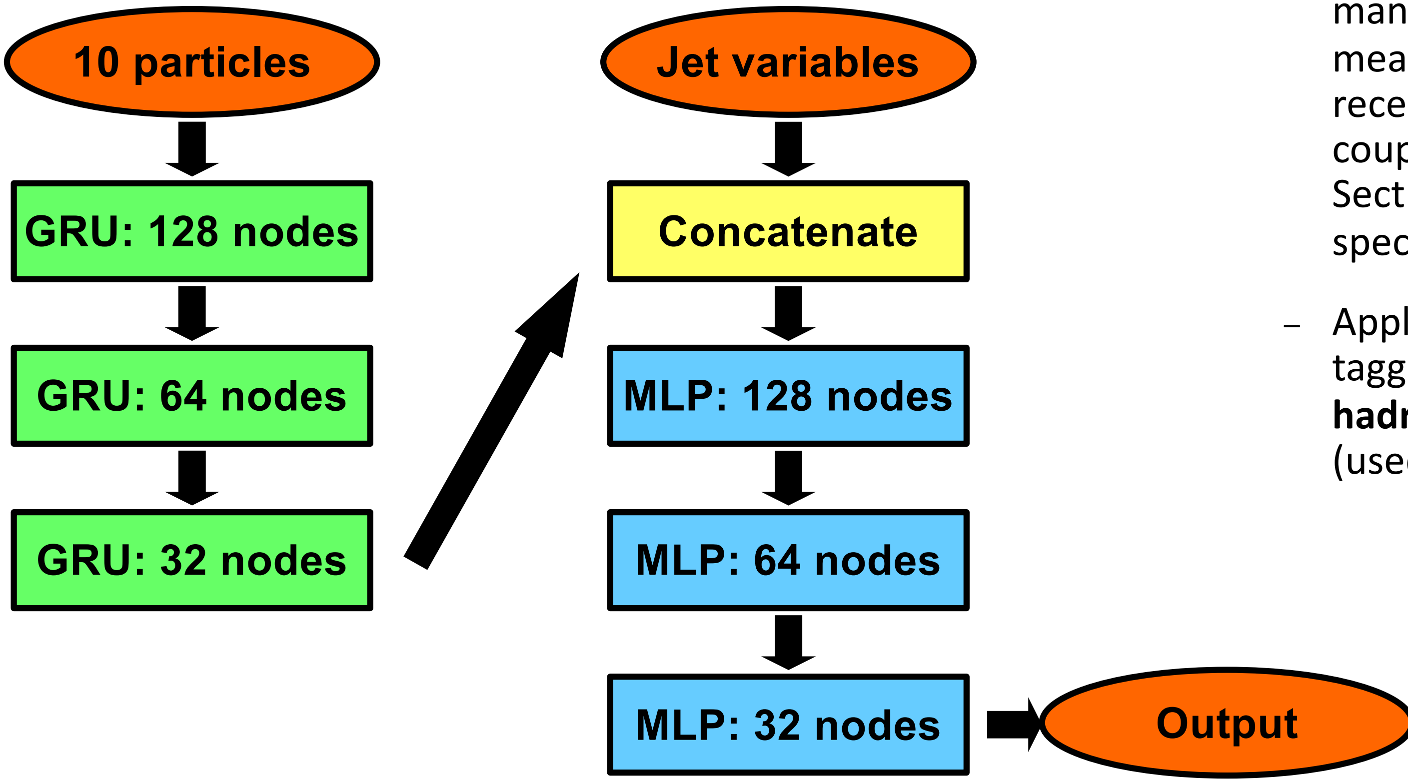


Multiclassifier tagger and inputs

- Use a ***multi-classifier*** **tagger**, which assigns probabilities to the possible flavours of a jet simultaneously
- Train on **ILD-reconstructed *H->qq/gg* samples** (*qq = uu*, *dd*, *ss*, *cc*, *bb*) with √*s* = 250 GeV and $P_L[e^-]$ = -100% and $P_R[e^+]$ = +100%
  - *Unskimmed*, except for $N_{jets} \geq 2$, $N_{leptons} = 0$, and truth *H*→*qq/gg* cuts

- Use **per-jet level inputs** as well as variables on the **10 leading particles** in each jet (with kinematics re-defined relative to the jet axis and re-normalized relative to jet momentum)
  - Jets:
    - momentum *p*, pseudorapidity *η*, polar angle *ϕ*, mass *m*, *b*/*c*-tagger scores, $N_{particles}$
  - Particles:
    - *p*, *η*, *ϕ*, *m*, charge, **truth** electron/muon/pion/kaon(including KS0, K+/-, and Λ)/proton likelihoods (0 or 1, using PDG ID **–** ***dE/dx*** **and TOF likelihoods in ILD samples have a bug – not used in current analysis, opted for truth info instead**)

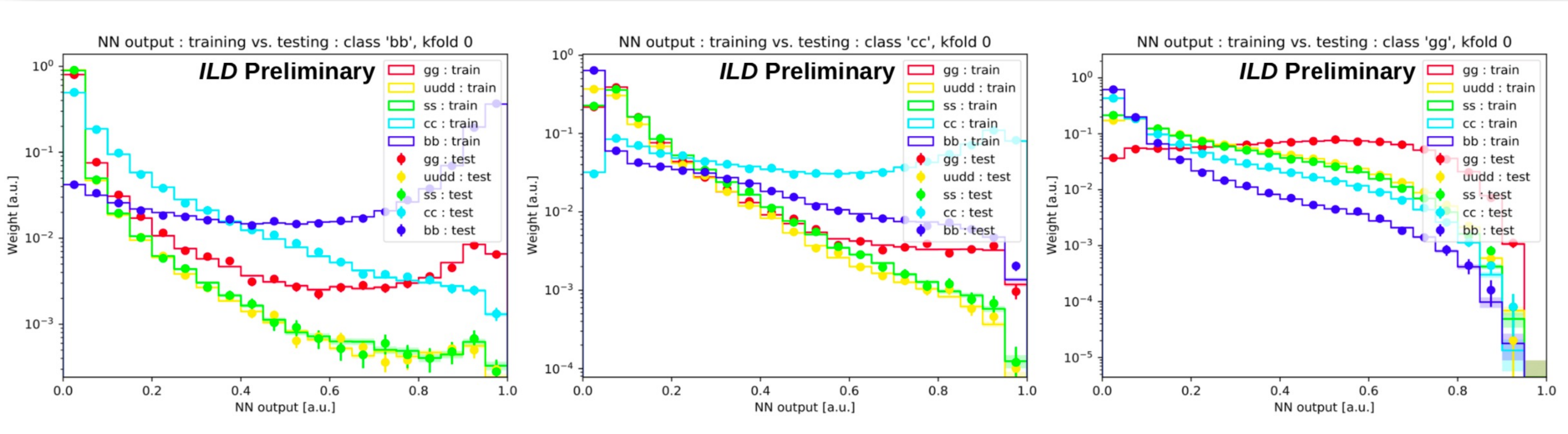
Multiclassifier Tagger: architecture & inputs



- Architecture shows up in many different HEP measurement scenarios (e.g., recent ATLAS $H \rightarrow ZZ \rightarrow 4\ell$ couplings measurement, see Section 5.2 of 2004.03447); specifically
- Applied even to strange tagging performance at **hadron** colliders 2011.10736 (used LSTMs instead of GRUs)

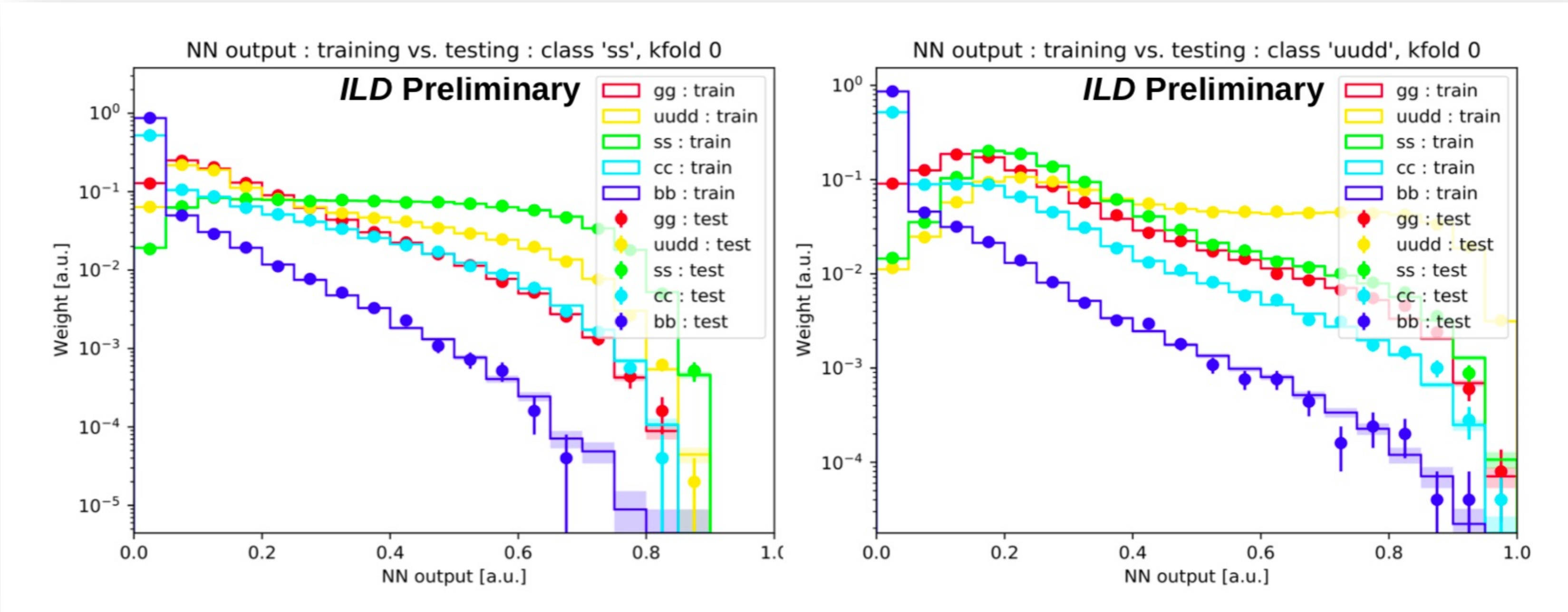
GRU: Gated Recurrent Units
MLP: MultiLayer Perceptron

Performance: b, c, and g jets



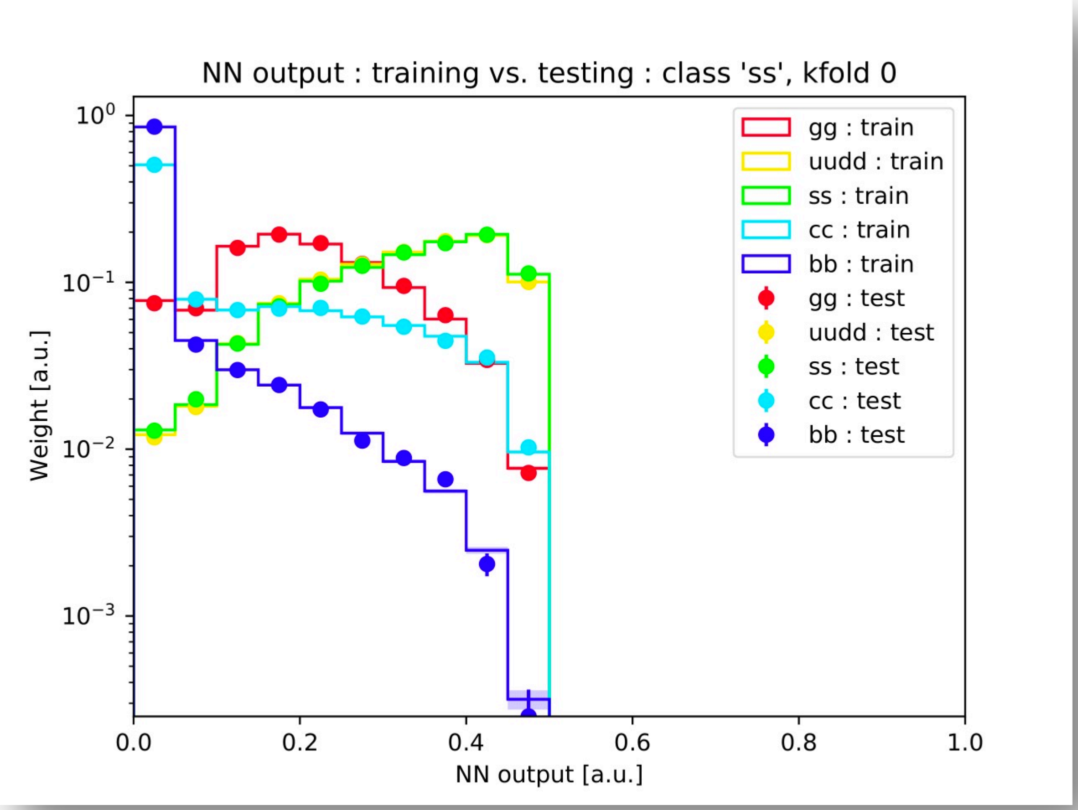
- MVA likely returning *b*/*c*-tagger scores – should do just as well or better than input BDT scores
- Reasonable discrimination of gluon jets

Performance: s and u/d jets



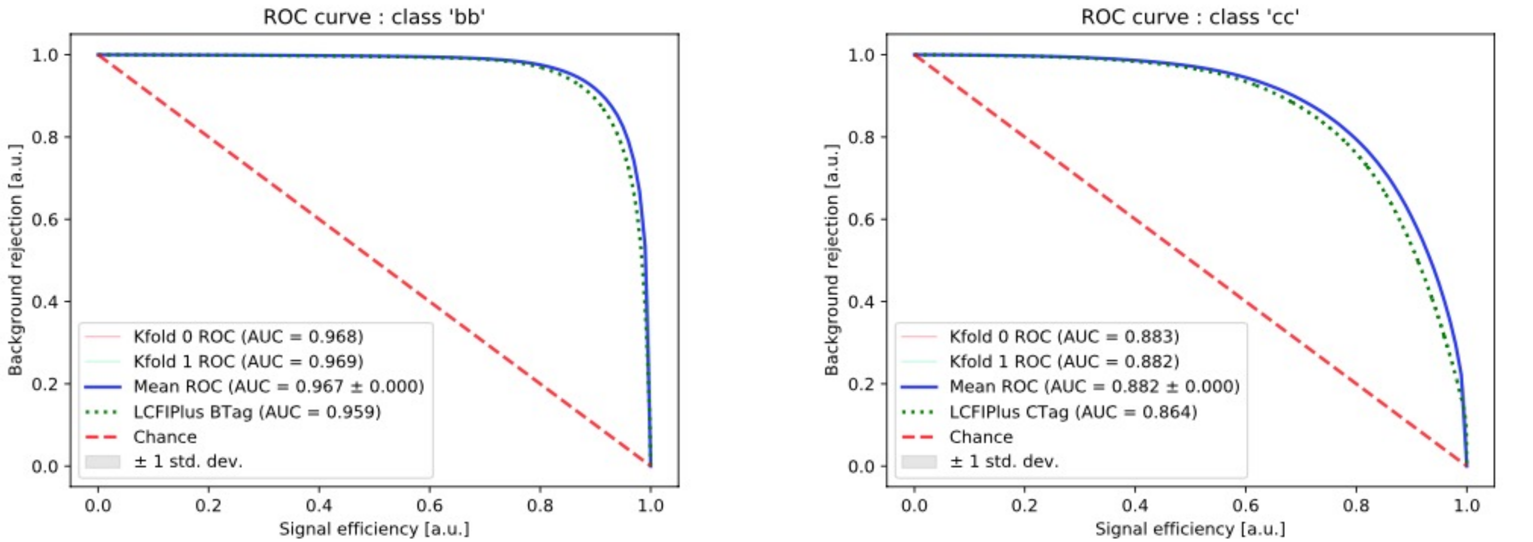
- Separation of *s* and *u*/*d* is **possible** with using truth likelihoods
- Also **good discrimination of *s* jets from *g* jets** – here, $N_{particles}$ is powerful
- At 50% strange tagging efficiency, we have **90%** background rejection over **70%** for LCFIPlus Otag (more ROC curves in back-up and LCWS2021 talk)

Performance: s jets with no PID



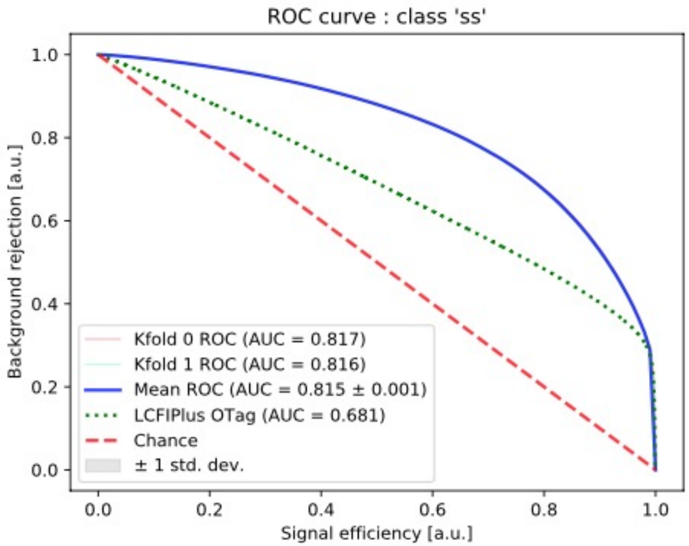
- Old version of the tagger with no PID
- No discrimination between s and u/d jets can be achieved

ROC curves: b and c jets

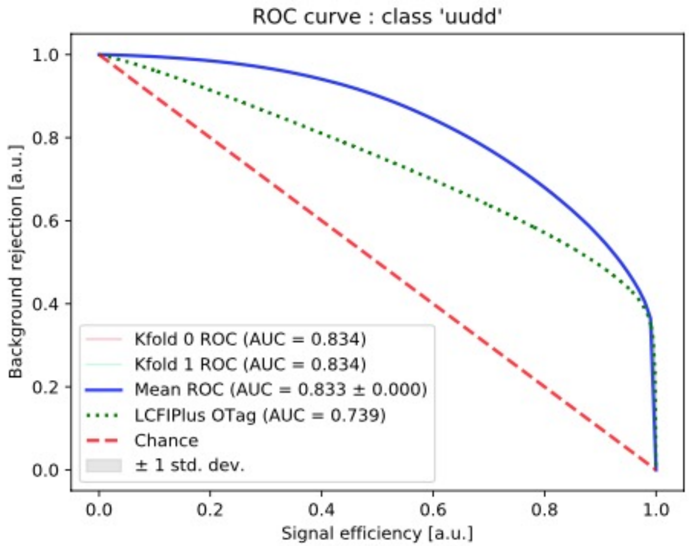


(a) b-jet score (b) c-jet score

ROC curves: light jets

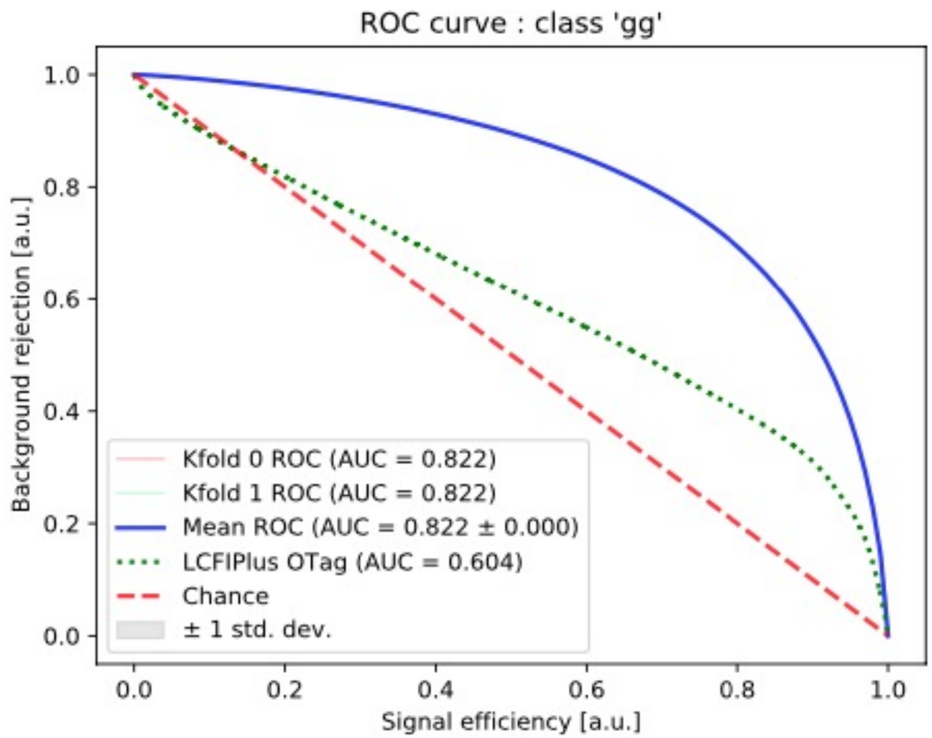


(c) s-jet score



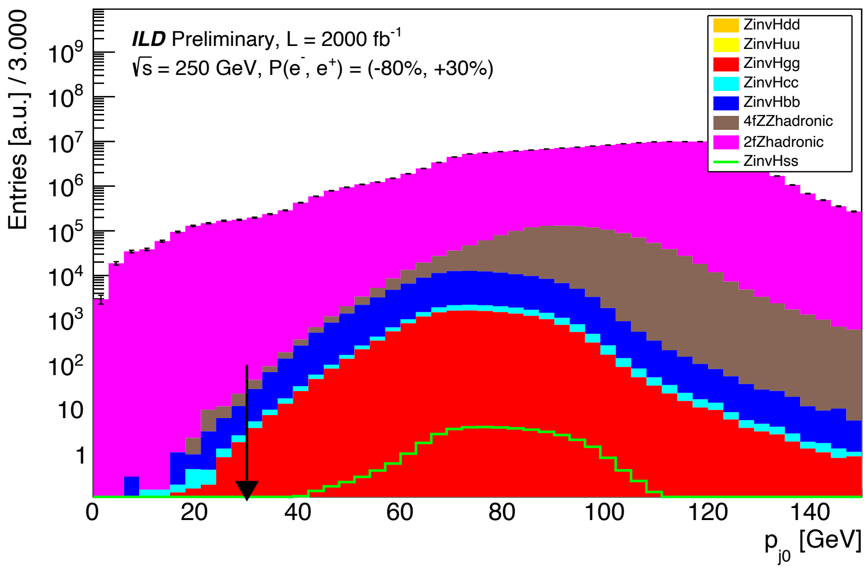
(d) u/d-jet score

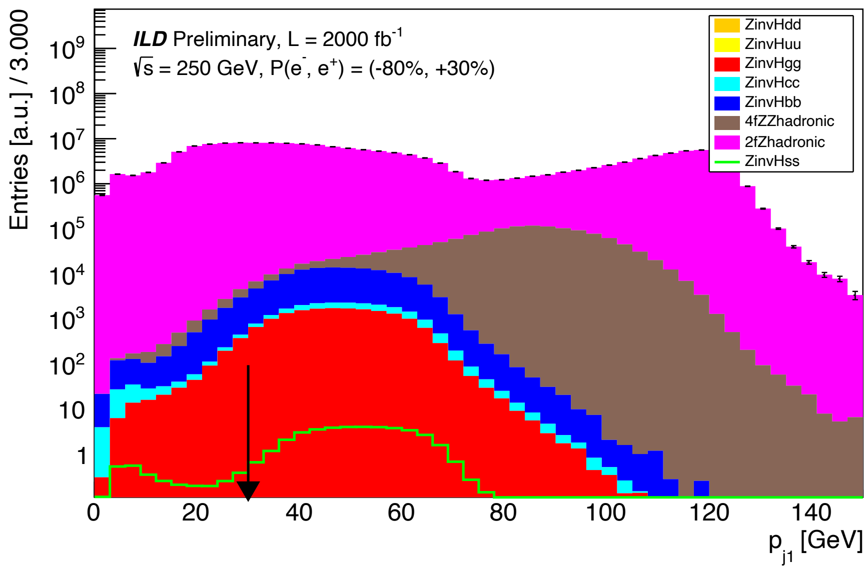
ROC curves: gluon jets



(e) g-jet score

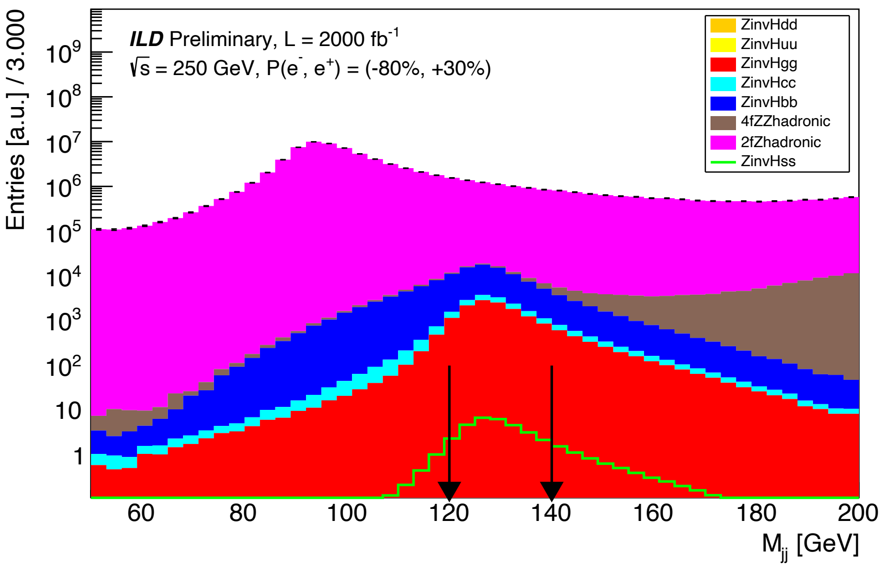
Histograms: pj0 and pj1

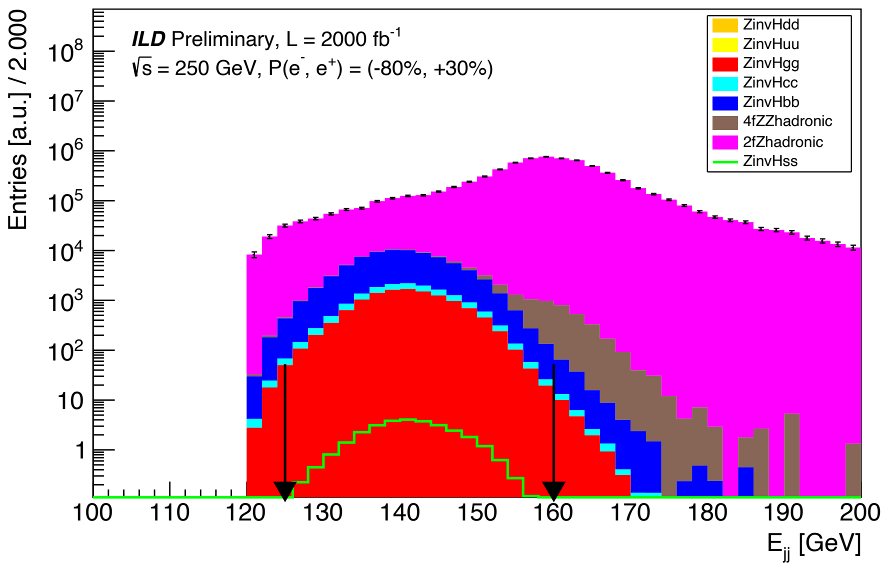




****Unstacked green line is signal****

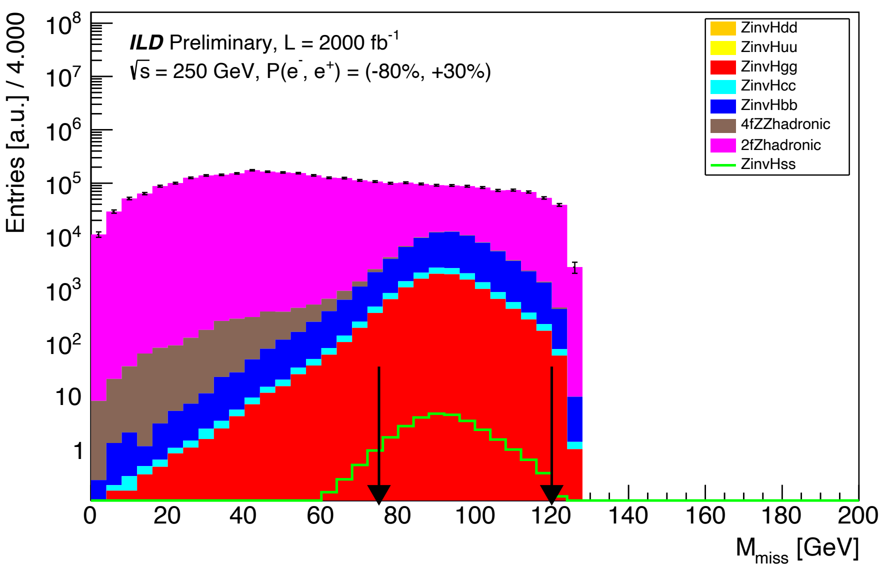
Histograms: Mjj and Ejj

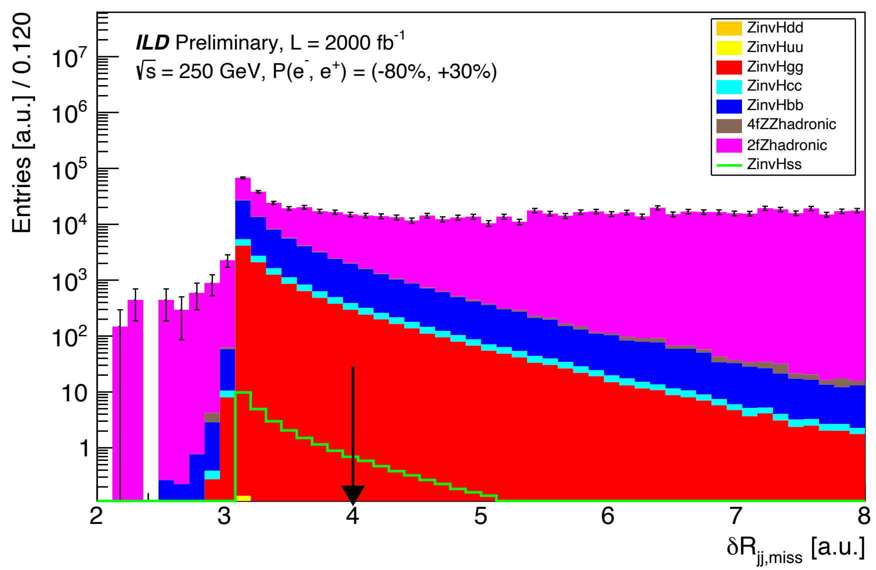




****Unstacked green line is signal****

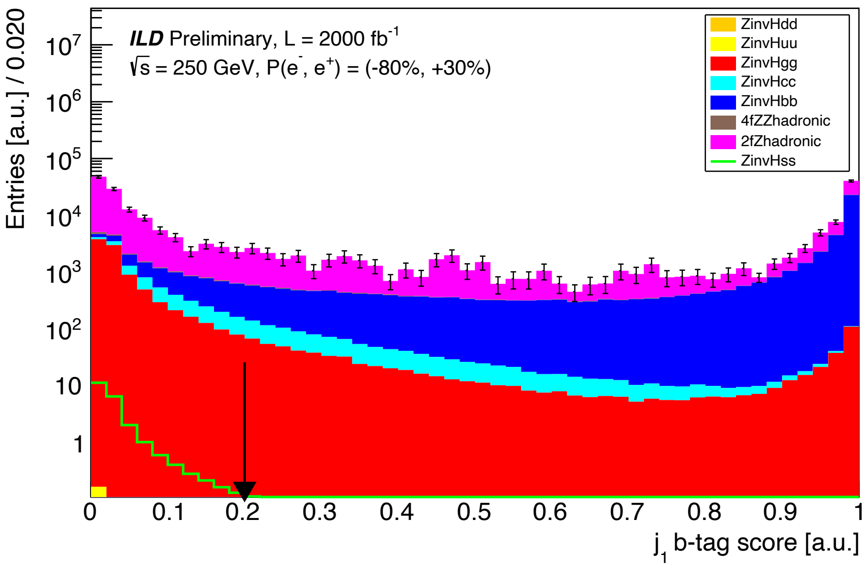
Histograms: Mmiss and ΔRjj,miss

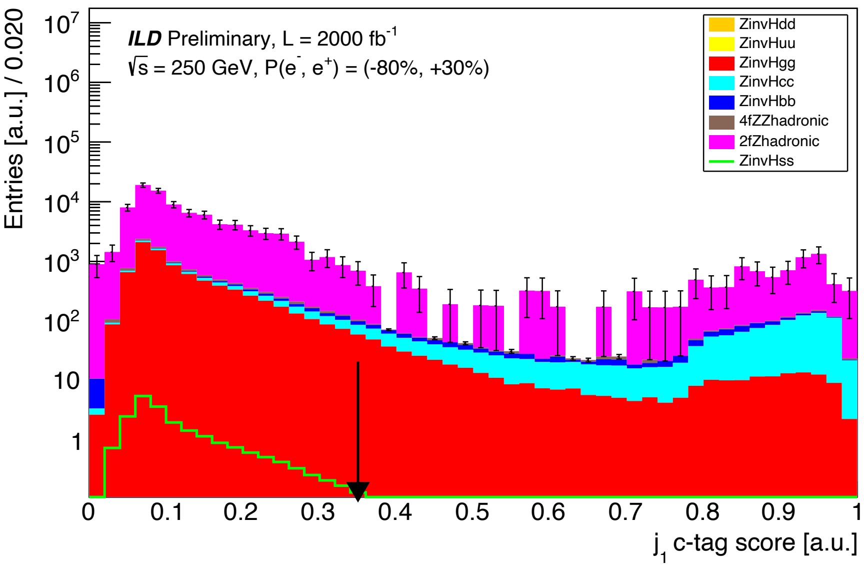




****Unstacked green line is signal****

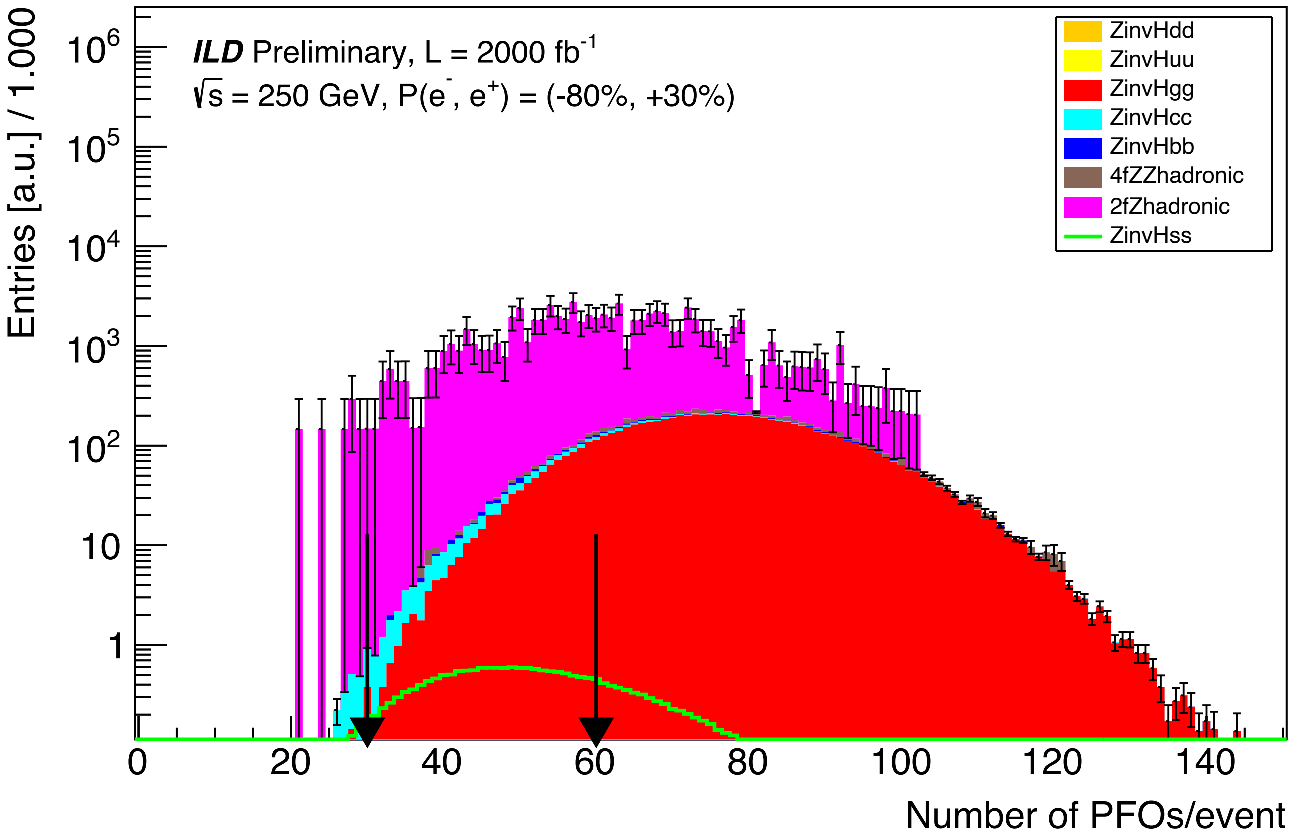
Histograms: b- & c-tagger scores



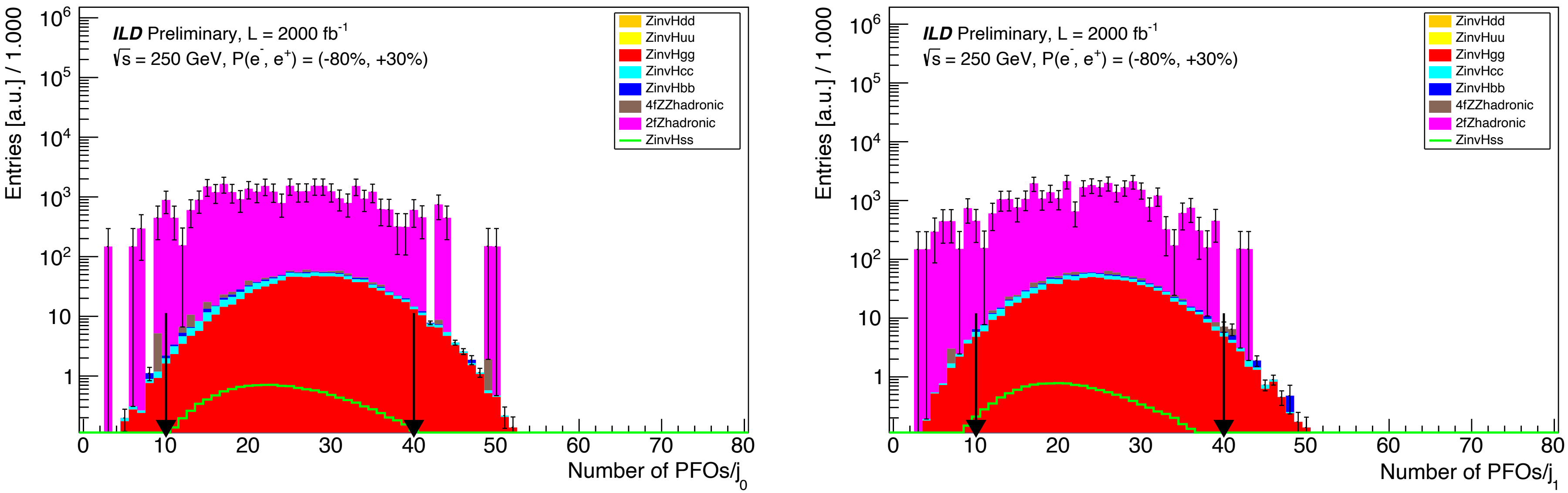


****Unstacked green line is signal****

Histograms: NPFOs/event



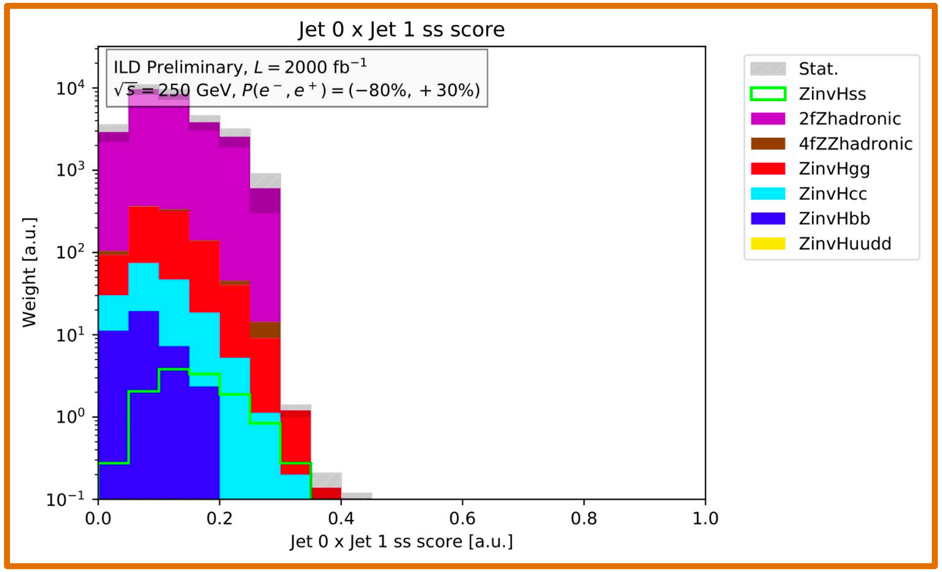
Histograms: NPFOs/jet

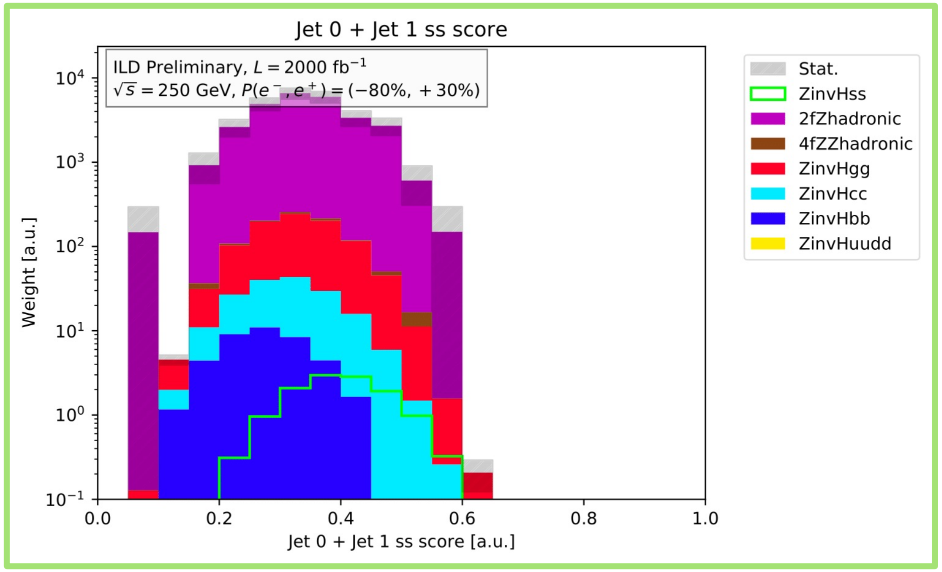


Unstacked green line is signal

Signal discriminant

- Using the **product** or the (0.5x) **sum** of leading and sub-leading strange scores as a discriminant



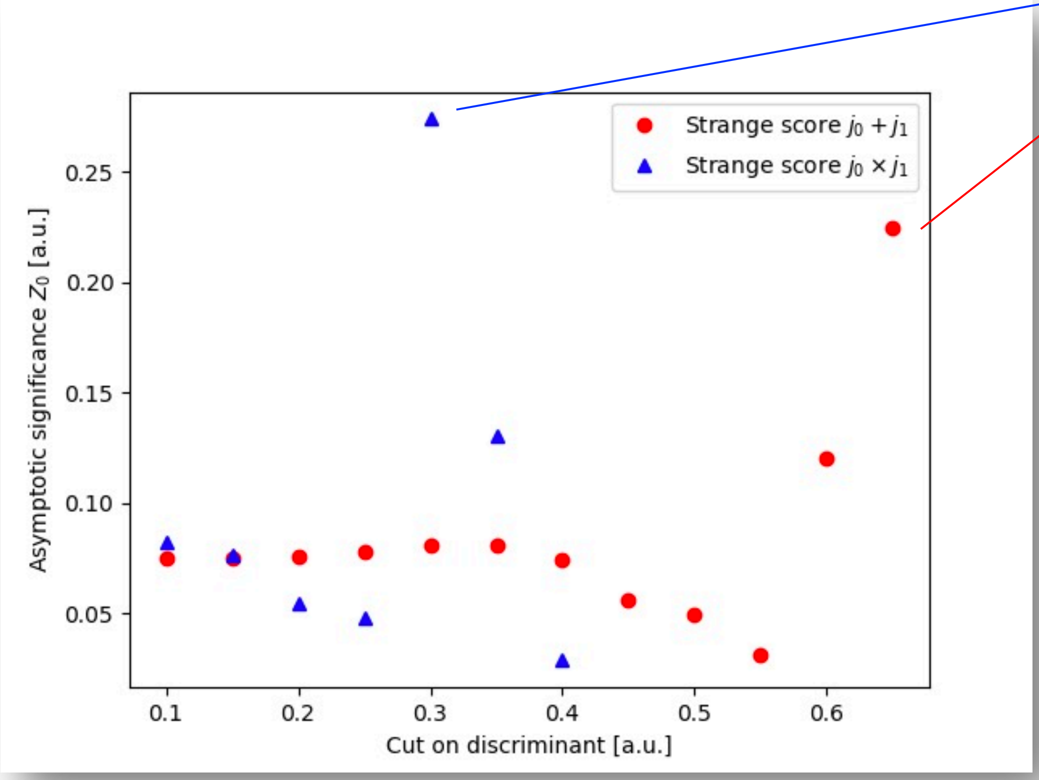


Signal discriminant (2)

- Yields for different cuts:
  - Using asymptotic significance assuming Asimov data (neglecting MC stats):

$$Z_0 = \sqrt{(2 * ((s + b) * \ln(1 + s / b) - s))}$$

Outliers, s < 1 and b ~ O(1)



Better sensitivity achieved from the sum of the score

Signal discriminant (2)

j0 x j1:
Cut: 0.2
s : 5.5654785096412525
b : 5253.939919094857
Z0 : 0.07676854538565478
Cut: 0.25
s : 3.9403779269196093
b : 3728.830329093824
Z0 : 0.06451713959512251
Cut: 0.3
s : 2.7415917753241956
b : 1644.0106210620702
Z0 : 0.06759737625722617
Cut: 0.35
s : 1.7592644195538014
b : 1042.245556926145
Z0 : 0.054478354764755384
Cut: 0.4
s : 1.0887995637021959
b : 594.2361149458384
Z0 : 0.04465148158251319
Cut: 0.45
s : 0.566939894342795
b : 443.92903776872777
Z0 : 0.026902202768505058
Cut: 0.5
s : 0.2524129586527124
b : 148.01879303174542
Z0 : 0.020741007866406074
Cut: 0.55
s : 0.07083354517817497
b : 147.72465044426963
Z0 : 0.00582743974151281
Cut: 0.6
s : 0.021902027539908886
b : 0.013714998960494995
Z0 : 0.15548956770016534

j0 + j1:
Cut: 0.2
s : 12.242391029838473
b : 24256.373724540405
Z0 : 0.07859895706194787
Cut: 0.25
s : 11.75890307186637
b : 20571.76969071827
Z0 : 0.08197654633491663
Cut: 0.3
s : 10.947859286679886
b : 17024.09735365923
Z0 : 0.08389780949113125
Cut: 0.35
s : 9.810693963780068
b : 13326.963469873936
Z0 : 0.08497298080156246
Cut: 0.4
s : 8.271826807060279
b : 9660.45653458523
Z0 : 0.08414738944817839
Cut: 0.45
s : 6.574709334061481
b : 6314.407373362772
Z0 : 0.08272464629857522
Cut: 0.5
s : 4.666403093840927
b : 4183.467504632149
Z0 : 0.07213289177636732
Cut: 0.55
s : 3.065837964299135
b : 2088.717378884584
Z0 : 0.06706611758544637
Cut: 0.6
s : 1.8046592578757554
b : 894.4982530597899
Z0 : 0.06031974977244867
Cut: 0.65
s : 0.9168394939042628
b : 445.62633642762626
Z0 : 0.043416925775766405
Cut: 0.7
s : 0.3047961258562282
b : 148.42523148232203
Z0 : 0.025009616775977003

Analysis overview

- Analysis performed on the same flavour tag samples as for training (500K events per flavour) as well as 2f_Z_hadronic and 4f_ZZ_hadronic samples (1.5M events each) – currently missing W-fusion signal (~10x smaller xs) and WW background
  - Cross sections assume $\sqrt{s}$ = 250 GeV and $P_L[e^-]$ = -80%, $P_R[e^+]$ = +30%
  - Accordingly, use the cross sections decorated onto the miniDSTs and multiply by BR[$H\rightarrow$inv] * BR[$H\rightarrow qq/gg$], BR[$Z\rightarrow$had], or BR[$Z\rightarrow$had]^2
    - N.B.: BR[$H\rightarrow ss$], BR[$H\rightarrow uu$], and BR[$H\rightarrow dd$] **aren't available**, so we take BR[$H\rightarrow cc$] and scale using **ratios of quark masses squared**
      - BR[$H\rightarrow ss$] ~ 2E-4, BR[$H\rightarrow uu$] ~ 2E-6, BR[$H\rightarrow dd$] ~ 5E-7
  - Multiply cross sections by integrated luminosity of **2000 fb^{-1}** to yield events
    - Could consider adding the 500 GeV int. lumi but it implies additional sample production, not easy for the time being

e+e- cross sections

Table 2: Cross-sections and number of generated MC samples on the Higgs production processes and the major SM background processes for both $\sqrt{s} = 250$ and 500 GeV. The cross-sections given in the table are set to be each operation beam polarization states: P$(e^-, e^+) = (-80\%, +30\%)$ and P$(e^-, e^+) = (+80\%, -30\%)$, whereas the number of MC samples are given with fully beam polarization states: P(e^-, e^+)= $P^L_{e^-} P^R_{e^+} = (-100\%, +100\%)$. The eeH(s) and eeH(t) denote the s-channel ZH process and the t-channel ZZ-fusion processes. $2f\_l$ and $2f\_h$ in the table indicate that the final state has a lepton pair such as charged leptons or neutrinos, and a quark pair like $u\bar{u}$, $d\bar{d}$ except $t\bar{t}$. $4f\_l$ and $4f\_h$ are the same indication with $2f\_l$ or $2f\_h$, that means a final state has two lepton pairs or two quark pairs. $4f\_sl$ shows that a final state has a lepton pair and a quark pair. At $\sqrt{s} = 500$ GeV $6f$ is included in the SM backgrounds, where possible diagrams of 6 fermions in a final state are considered such as $t\bar{t}$ and a fermion pair with two W bosons and two fermion pairs with the Z boson.

$\sqrt{s}$=250 GeV	operation polarization		fully polarization			
	Cross-section (fb)		MC sample			
P(e^-, e^+)	$(-80\%, +30\%)$	$(+80\%, -30\%)$	$P^L_{e^-} P^R_{e^+}$	$P^R_{e^-} P^L_{e^+}$	$P^L_{e^-} P^L_{e^+}$	$P^R_{e^-} P^R_{e^+}$
eeH(s)	10.7	7.14	$4.00 \cdot 10^4$	$1.00 \cdot 10^4$	0	0
eeH(t)	0.71	0.52	$1.00 \cdot 10^4$	$1.00 \cdot 10^4$	3992	3992
$\mu\mu H$	10.4	7.03	$4.00 \cdot 10^4$	$1.00 \cdot 10^4$	0	0
qqH	210.2	141.9	$5.45 \cdot 10^5$	$2.94 \cdot 10^5$	0	0
$\nu\nu H$ (s)	61.6	41.6	$12.8 \cdot 10^4$	$6.50 \cdot 10^4$	0	0
$\nu\nu H$ (t)	15.4	0.93	$12.8 \cdot 10^4$	$6.50 \cdot 10^4$	0	0
$2f\_l$	$3.82 \cdot 10^4$	$3.49 \cdot 10^4$	$2.63 \cdot 10^6$	$2.13 \cdot 10^6$	$5.03 \cdot 10^5$	$5.03 \cdot 10^5$
$2f\_h$	$7.80 \cdot 10^4$	$4.62 \cdot 10^4$	$1.75 \cdot 10^6$	$1.43 \cdot 10^6$	0	0
$4f\_l$	$6.03 \cdot 10^3$	$1.47 \cdot 10^3$	$2.25 \cdot 10^6$	$9.80 \cdot 10^4$	$2.73 \cdot 10^5$	$2.73 \cdot 10^5$
$4f\_sl$	$1.84 \cdot 10^4$	$2.06 \cdot 10^3$	$4.04 \cdot 10^6$	$3.56 \cdot 10^5$	$9.78 \cdot 10^4$	$9.78 \cdot 10^4$
$4f\_h$	$1.68 \cdot 10^4$	$1.57 \cdot 10^3$	$2.38 \cdot 10^6$	$2.42 \cdot 10^5$	0	0

Table 2, taken from page 62 of Tomohisa Ogawa's thesis

H→bb analysis: histograms

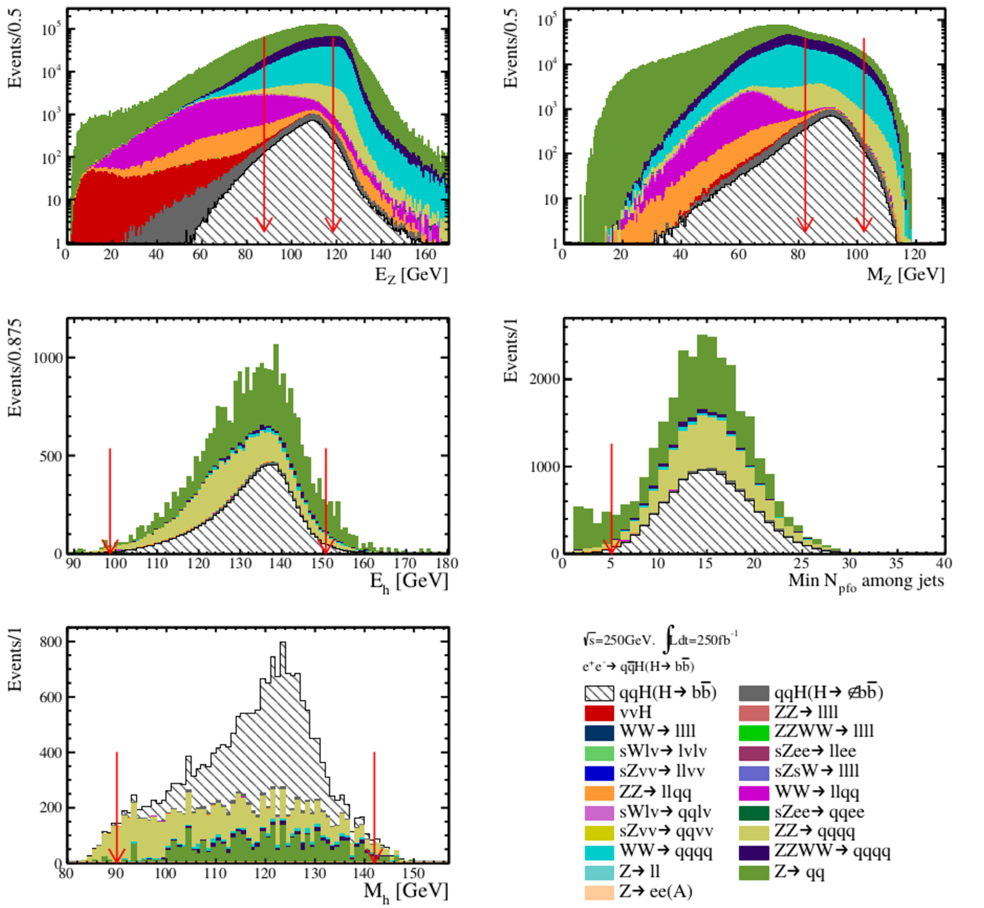


Figure 66: The distributions show each observable used for the background suppression assuming 250 fb^{-1} with P(e^-, e^+)= (-80%,+30%). The explanation of the observables are given in the text. Red arrows on each plot indicate the cut values applied to each observable as the background suppression.

Figure 66, taken from page 87 of Tomohisa Ogawa's thesis

H→bb analysis: cutflow

Table 4, taken from page 89 of Tomohisa Ogawa's thesis

Table 4: The expected number of remaining signal and background events after each cut for the $Zh \to q\bar{q}b\bar{b}$ at $\sqrt{s}$=250 GeV, with with both of the beam polarization states: P(e^-, e^+)= (-80%,+30%) and (+80%,-30%). The integrated luminosity of 250 fb^{-1} is assumed. The signal efficiency ϵ and significance S_{sig} are also given in the table.

$\sqrt{s}$=250 GeV P(e^-, e^+)= (-80%,+30%)

Cut variables	$q\bar{q}b\bar{b}$	ϵ	$q\bar{q}H(H \notin b\bar{b})$	$2f$	$4f$	S_{sig}
No cut	30372	100	22175	$2.9 \cdot 10^7$	$1.02 \cdot 10^7$	-
$N_{isolep} = 0$	30314	99.8	17492	$2.6 \cdot 10^7$	$6.9 \cdot 10^6$	5.28
$N_{pfo} \in [55,\ 170]$	30218	99.5	15141	$6.0 \cdot 10^6$	$4.4 \cdot 10^6$	9.37
$E_Z \in [87.75,\ 118.50]$ GeV	25712	84.7	11365	$3.3 \cdot 10^6$	$2.8 \cdot 10^6$	10.35
$M_Z \in [82.29,\ 102.29]$ GeV	18658	61.4	7572	$3.8 \cdot 10^5$	$1.0 \cdot 10^6$	15.62
b-$tag \in [1.25,\ 2.0]$	11203	36.9	381	9364	8454	65.76
$E_H \in [98.67,\ 150.67]$ GeV	10909	35.9	368	8242	7998	66.21
Min $N_{pfo} \in [5,\ 40]$	10841	35.7	358	6932	7792	67.81
$-\log y_{32} \in [0.5, 3.62]$	10409	34.3	349	3917	7453	70.53
$-\log y_{43} \in [1.8, 5.52]$	10065	33.2	346	2921	7027	71.15
thrust $T \in [0.5,\ 0.89]$	9966	32.8	345	2520	7004	71.39
$M_H \in [90,\ 142]$ GeV	9907	32.6	335	2419	6382	72.43

Meeting the ILD community

Inputs and outputs

- <u>Outputs</u>: could imagine the network provides bottom, charm, strange, and light output scores
  - **Multiclassifier** provides more freedom for output class
- <u>Jets</u>: p4, ILD tagger scores (b-, c-, o-, and category?), ...
  - **Anything else which is sensible/useful to include?**
- <u>Tracks (jet constituent particles)</u>: p4, momentum / jet momentum, dE/dx (+ uncertainty?), different PID likelihoods, ...
  - **Anything else?**

We really invite you to read this!

More in Jan's talk!

Tagger architecture(s)

- Possible architectures from the literature include:
  - "Maximum performance of strange-jet tagging at **hadron** colliders" (2011.10736 – published in November 2020)
    - {Recurrent neural network for track inputs} + {jet inputs} -> Concatenate -> multilayer perceptron (MLP) -> output
    - Could also use MLP on the jet inputs prior to concatenation
  - "ParticleNet: Jet Tagging via Particle Clouds" (1902.08570)
    - Proposed for flavour tagging at FCC-ee (see talk here)
    - *Complex*: represent particles in jet as a graph and apply EdgeConv (1801.07829) units to relationships between a given particle and its nearest neighbours

We found out in the meeting that dE/dx is bugged in the current version of the ILD ntuples! ☹

Maximum performance of strange tagging at colliders: https://arxiv.org/pdf/2011.10736.pdf

Name	Selection criteria	Input variables
Universal detector	$\tau > 0$	E, η, ϕ, m (4-momentum) r_0, η_0, ϕ_0 (origin) τ (lifetime in lab system) q (charge)
Universal detector excluding beampipe	$\tau > 0$ $r_f > 10\,\text{mm}$	E, η, ϕ, m (4-momentum), r_i, η_i, ϕ_i (initial measurement), τ (lifetime in lab system), q (charge)
Infinite tracker	$\tau > 0$ $r_f > 10\,\text{mm}$ $q \neq 0$	p, η, ϕ (4-momentum minus mass) r_i, η_i, ϕ_i (initial measurement) τ (lifetime in lab system) q (charge)
Finite tracker	$\tau > 0$ $r_f > 10\,\text{mm}$ $q \neq 0$ $r_0 < 1\,\text{m}$	p, η, ϕ (4-momentum minus mass) r_i, η_i, ϕ_i (initial measurement) τ (lifetime in lab system) q (charge)
Cherenkov tracker	$\tau > 0$ $r_f > 10\,\text{mm}$ $q \neq 0$ $r_0 < 1\,\text{m}$	p, η, ϕ, m (4-momentum) r_i, η_i, ϕ_i (initial measurement) τ (lifetime in lab system) q (charge)
Calorimeter without ECAL/HCAL separation	$\tau > 0$ $r_0 < 1\,\text{m}$ $r_f > 1\,\text{m}$ no ν	E, η, ϕ (3-momentum)
Calorimeter with ECAL/HCAL separation	$\tau > 0$ $r_0 < 1\,\text{m}$ $r_f > 1\,\text{m}$ no ν	E, η, ϕ (3-momentum), particle category (γ/e, μ, other)
Finite tracker, no $0 \rightarrow +-$ decays	$\tau > 0$ $r_f > 10\,\text{mm}$ $r_0 < 1\,\text{m}$ $q \neq 0$ no charged particles from neutral decays	p, η, ϕ, m (4-momentum) r_i, η_i, ϕ_i (initial measurement) τ (lifetime in lab system) q (charge)
Finite tracker, only $0 \rightarrow +-$ decays	$\tau > 0$ $r_f > 10\,\text{mm}$ $r_0 < 1\,\text{m}$ $q \neq 0$ only charged particles from neutral decays	p, η, ϕ, m (4-momentum) r_i, η_i, ϕ_i (initial measurement) τ (lifetime in lab system) q (charge)

Table 1. List of all considered detector scenarios as a combination of ideal detector components. The second column shows the selection requirements imposed on the particles used as input to the neural networks, where τ means the lifetime of the particles, r_0 is the radial distance between the primary vertex and the point where the particle is created, and r_f is the radial distance between the primary vertex and the decay vertex. The third column describes the variables that are used as inputs to the neural network. If the variable carries a subscript 0, it refers to the spacepoint of creation, and if it carries a subscript i, it refers to the spacepoint of initial measurement.

– 6 –

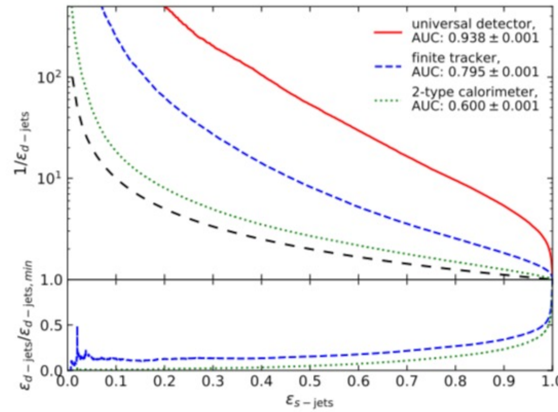


Figure 14. ROC curves illustrating the classification power of the neural networks in the universal-detector, finite-tracker, and ECAL/HCAL-separated calorimeter ("2-type calorimeter") scenarios. The signal is composed of s-jets, and the background is composed of d-jets. The dashed line illustrates a ROC curve for the case of no separation. The ratio beneath the ROC curves shows the efficiency for d-jets in one scenario ($\varepsilon_{d\text{-jets}}$) divided by the efficiency for d-jets in the scenario with the best separation power shown in the ROC curve ($\varepsilon_{d\text{-jets,min}}$). The efficiencies are evaluated on the test sample and the uncertainty in the area under the curve is the statistical uncertainty associated with that sample.

- Assuming an ideal detector that can perfectly measure all jet constituents ("universal collider detector"), s- and d-jets can be separated well. This means that the fragmentation of s- and d-jets shows promising differences that may be explored in an s-tagging algorithm, but that the maximum achievable performance of an s-tagger is by far not as good as for example achieved for b-tagging algorithms (see back-up for comparison)
- The comparison also shows that the information measured in a perfect tracker may be much more valuable for s-tagging than the energy deposits measured in electromagnetic and hadronic calorimeters.
- Interestingly, the addition of an ideal Cherenkov detector to the tracking scenario does not yield a large improvement.

Possible FCC Collaborators

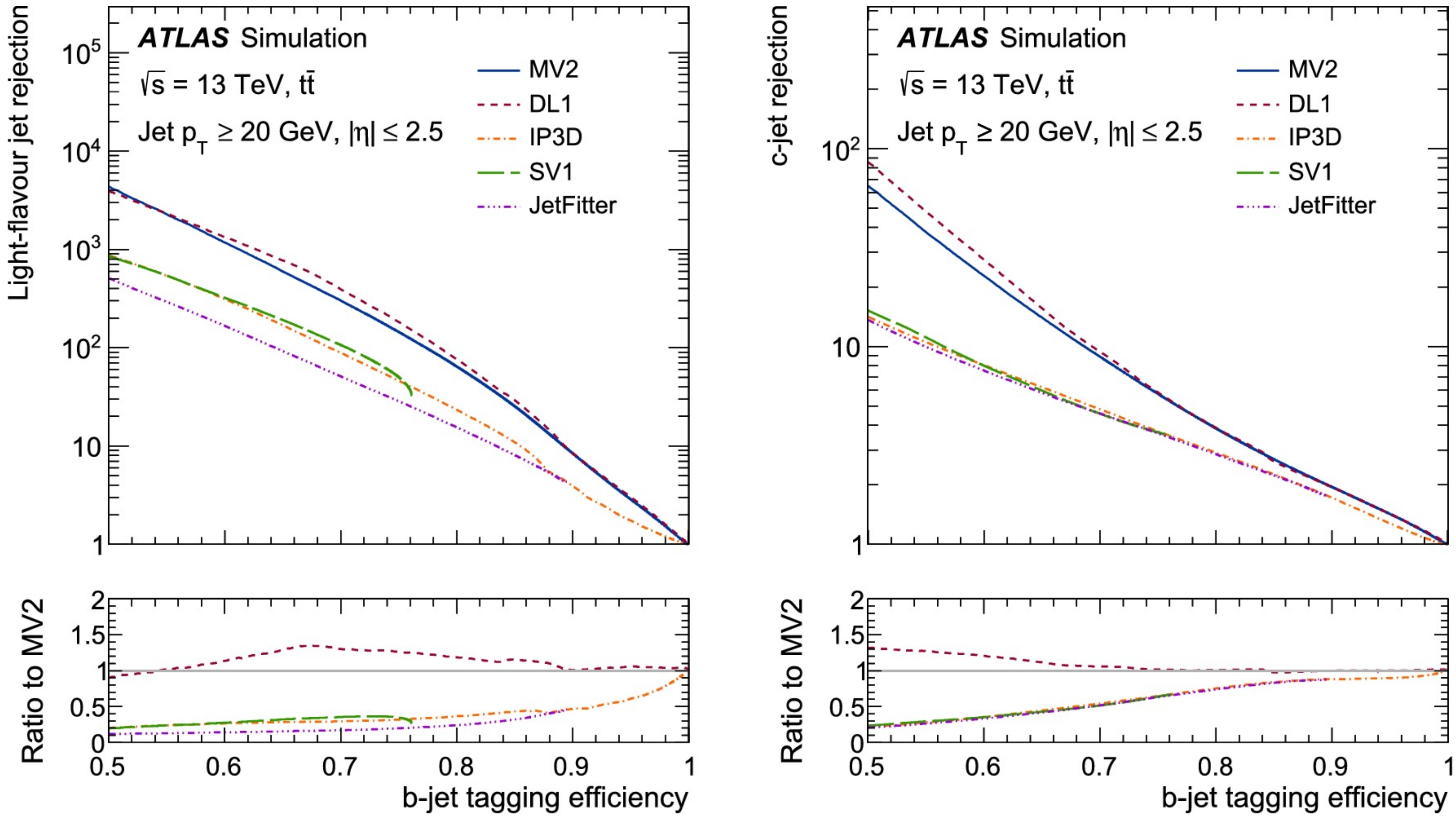
- David D'Enterria gave a talk during one of the EF1 meetings on ***Electron Yukawa from s-channel in e^+e^- → Higgs production at FCC-ee*** and during the talk he mentioned that he was working also on H->ss, so we got in touch with him to explore possible collaborations

- Their focus would be on the exclusive **H->phi+gamma** decay, rather than the full h->ssbar with jet reconstruction
  - *One expects a handful of such rare decay events with the ~1.5 million Higgs expected at the* ***FCC-ee***
  - *This direct decay interferes with the (more probable)* ***H->gamma gamma*->gamma phi channel****, and one needs to disentangle the dependence of the yields on k_gamma and k_s (the k_gamma coupling should be known with good accuracy...).*
  - *There are phenomenological studies for the LHC* (in fact we has cited the one from the ATLAS Collaboration), *but neither of us recalled them for e+e-.*
  - He proposed to take a closer look at it and try to estimate the actual sensitivity to k_s
  - If potentially relevant, and if we are interested in that channel, we carry out together a simulation analysis...
    - Main signal and background samples needed would be the ones mentioned above

Possible FCC Collaborators

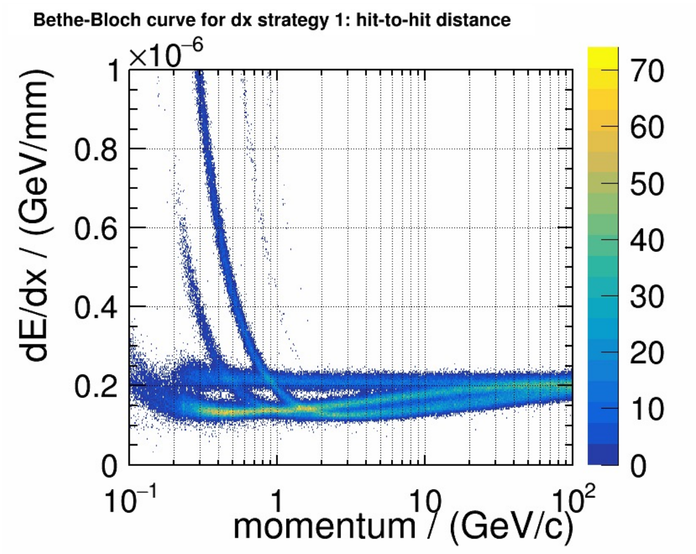
- Recent Updates from David:
- For orientation, in order to obtain bounds coming close to the SM kappa_s/kappa_b ~ 0.02 expectation, one needs **H->phi+gamma** measurements with a 1% uncertainty (corresponding to -0.04 < kappa_s/kappa_b < 0.08).
- With 1.5e6 Higgs expected at the FCC-ee and a BR(H->phi+gamma) = 2.3e-6, we only expect 3.5 signal events (on top of probably small backgrounds).
- So, any measurement of the decay will have, at least, a 50% statistical uncertainty. This would imply to set limits about 2 < kappa_s/kappa_b < 4, i.e. **more than 100 times the SM prediction**...
- Summary: **No strong motivation right now on running a simulation for this rare final state**. But it's worth to **quote this generic result in** a couple of lines **in any document that may be produced**, because people keep asking.
- David is happy to produce those lines if needed

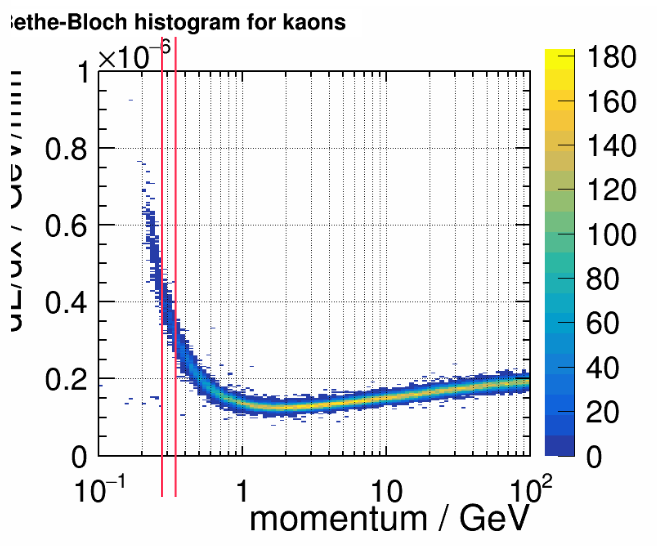
https://link.springer.com/article/10.1140/epjc/s10052-019-7450-8/figures/2



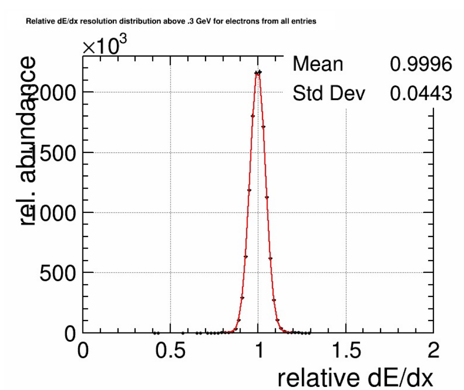
(a) (b)

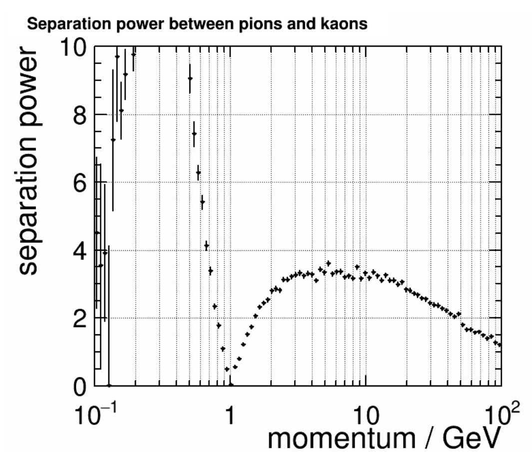
https://agenda.linearcollider.org/event/8067/contributions/43101/attachments/34181/52634/2019_01_10_LCTPC_Col_Meeting.pdf



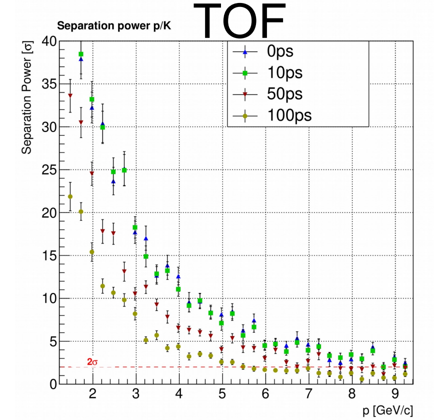


Single particles





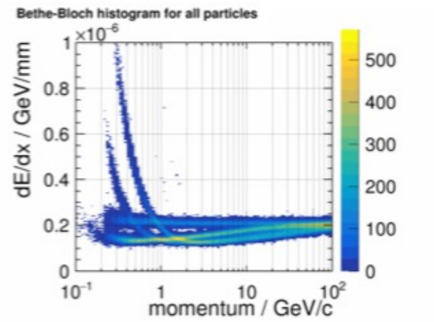
TOF

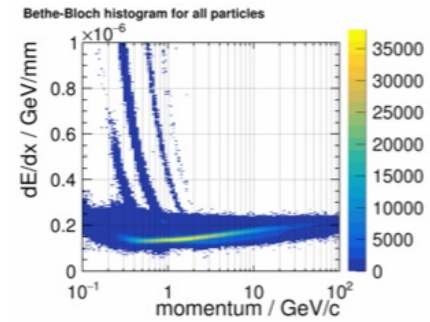


https://agenda.linearcollider.org/event/8067/contributions/43101/attachments/34181/52634/2019_01_10_LCTPC_Col_Meeting.pdf

SLAC

Comparison: dE/dx for single particles vs. 6f-$t\bar{t}$ events





	l5 single	l5 6f-tt	s5 single	s5 6f-tt
electrons	4.3 %	4.5 %	5.3 %	5.4 %
muons	4.5 %	4.8 %	5.4 %	5.7 %
pions	4.5 %	4.6 %	5.5 %	5.6 %
kaons	4.6 %	4.7 %	5.5 %	5.7 %
protons	4.6 %	4.7 %	5.5 %	5.7 %

ILD

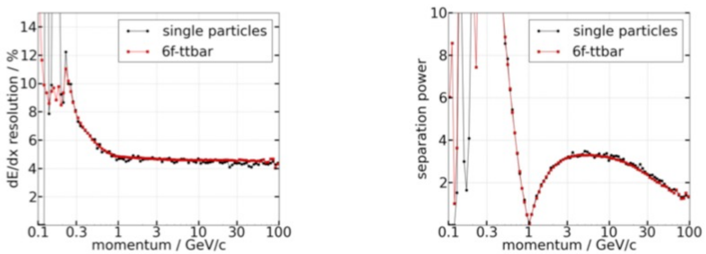
Uli Einhaus | PID with dE/dx and TOF at ILD | 10.01.2019 | Page 9

DESY.

https://agenda.linearcollider.org/event/8067/contributions/43101/attachments/34181/52634/2019_01_10_LCTPC_Col_Meeting.pdf

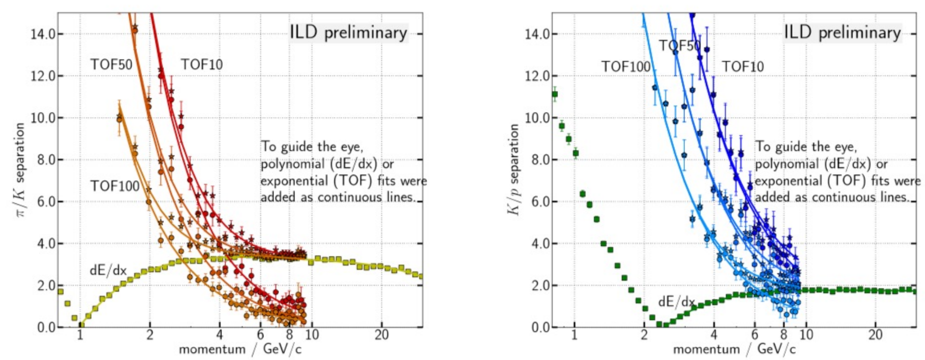
SLAC

Resolution & π/K-separation in single comp. to $t\bar{t}$



ILD

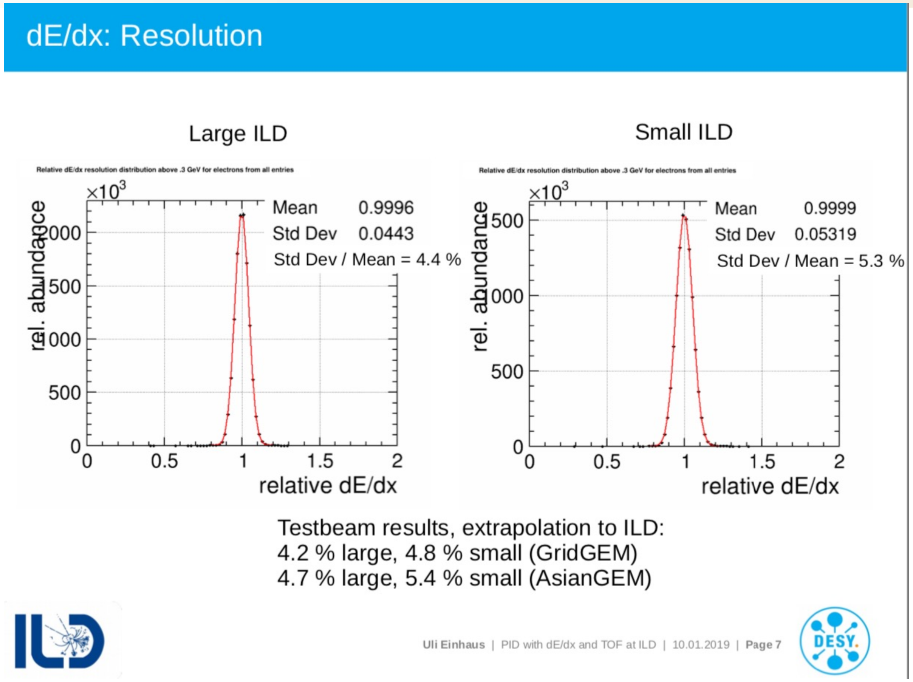
Uli Einhaus | PID with dE/dx and TOF at ILD | 10.01.2019 | Page 10

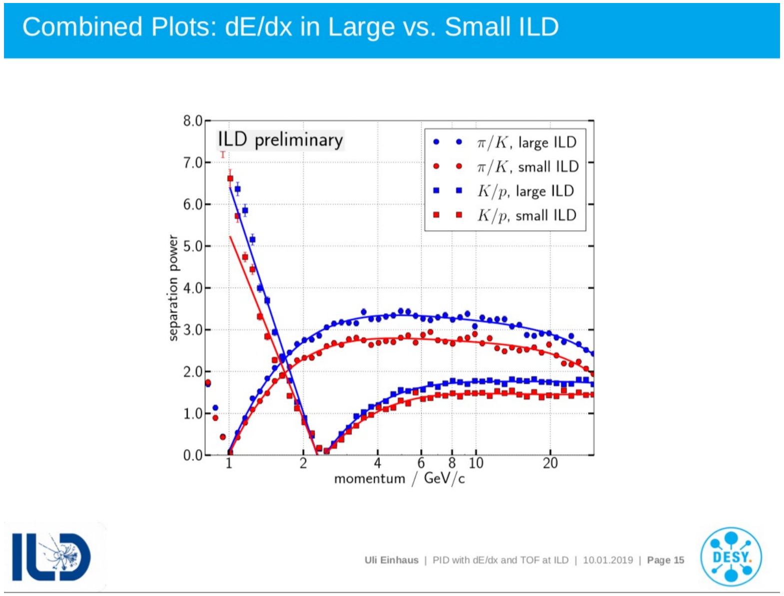


ILD

Uli Einhaus | PID with dE/dx and TOF at ILD | 10.01.2019 | Page 19

DESY.


dE/dx: Resolution
Large ILD
Small ILD
Mean 0.9996
Std Dev 0.0443
Std Dev / Mean = 4.4 %
Mean 0.9999
Std Dev 0.05319
Std Dev / Mean = 5.3 %
rel. abundance
relative dE/dx
Testbeam results, extrapolation to ILD:
4.2 % large, 4.8 % small (GridGEM)
4.7 % large, 5.4 % small (AsianGEM)
Uli Einhaus | PID with dE/dx and TOF at ILD | 10.01.2019 | Page 7
DESY.


Combined Plots: dE/dx in Large vs. Small ILD
ILD preliminary
π/K, large ILD
π/K, small ILD
K/p, large ILD
K/p, small ILD
separation power
momentum / GeV/c
Uli Einhaus | PID with dE/dx and TOF at ILD | 10.01.2019 | Page 15
DESY.

SLAC

Sukeerthi Dharani

Particle Identification using Time of Flight

Goal: Use arrival time at ECAL to determine particle ID

$$\beta = \frac{l_{\text{track}}}{t_{\text{arrival}}}$$

l_{track}: From momentum & curve in B field

t_{arrival}: Time of first hit from 10 closest hits in ECal

- Test on $t\bar{t}$ events
- Charged Particles : p, κ, pi, μ, e
- **Time resolution:** 0 ps,10 ps, 50 ps
- Cuts: filter only particles hitting barrel, omit particles that spiral inside TPC
- Processors: First hit & Closest hits

DESY. | Particle Identification using Time of Flight | Sukeerthi Dharani | June 13, 2018 | Page 2/8 https://agenda.linearcollider.org/event/7972/contributions/41753/attachments/33290/50855/ParticleID-TOF.pdf

SLAC

Sukeerthi Dharani

Particle Identification using Time Of Flight

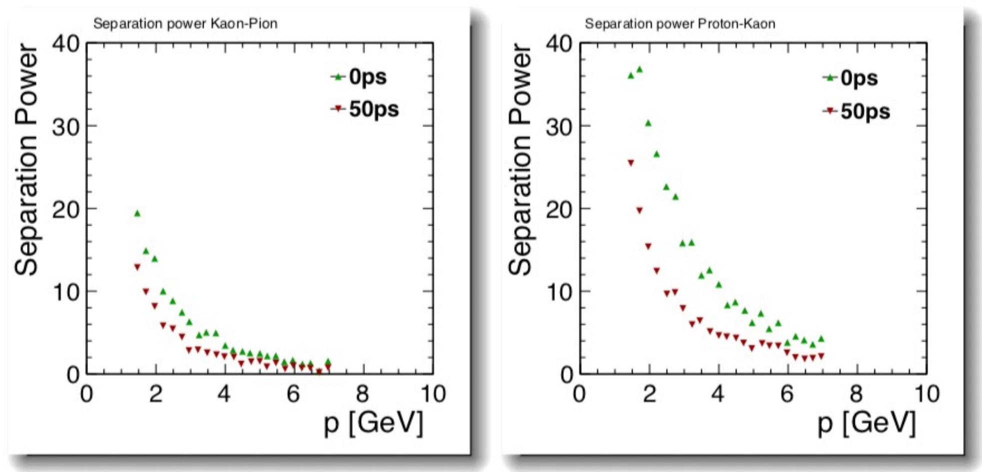
Particle-ID for low-p hadrons

Separation power (between particle i and j):

Gaussian fit $\Rightarrow$ μ_i: Mean for particle type i

σ_i: Std. dev. for particle type i

$$S = \frac{|\mu_i - \mu_j|}{\sqrt{(\sigma_i^2 + \sigma_j^2)/2}}$$

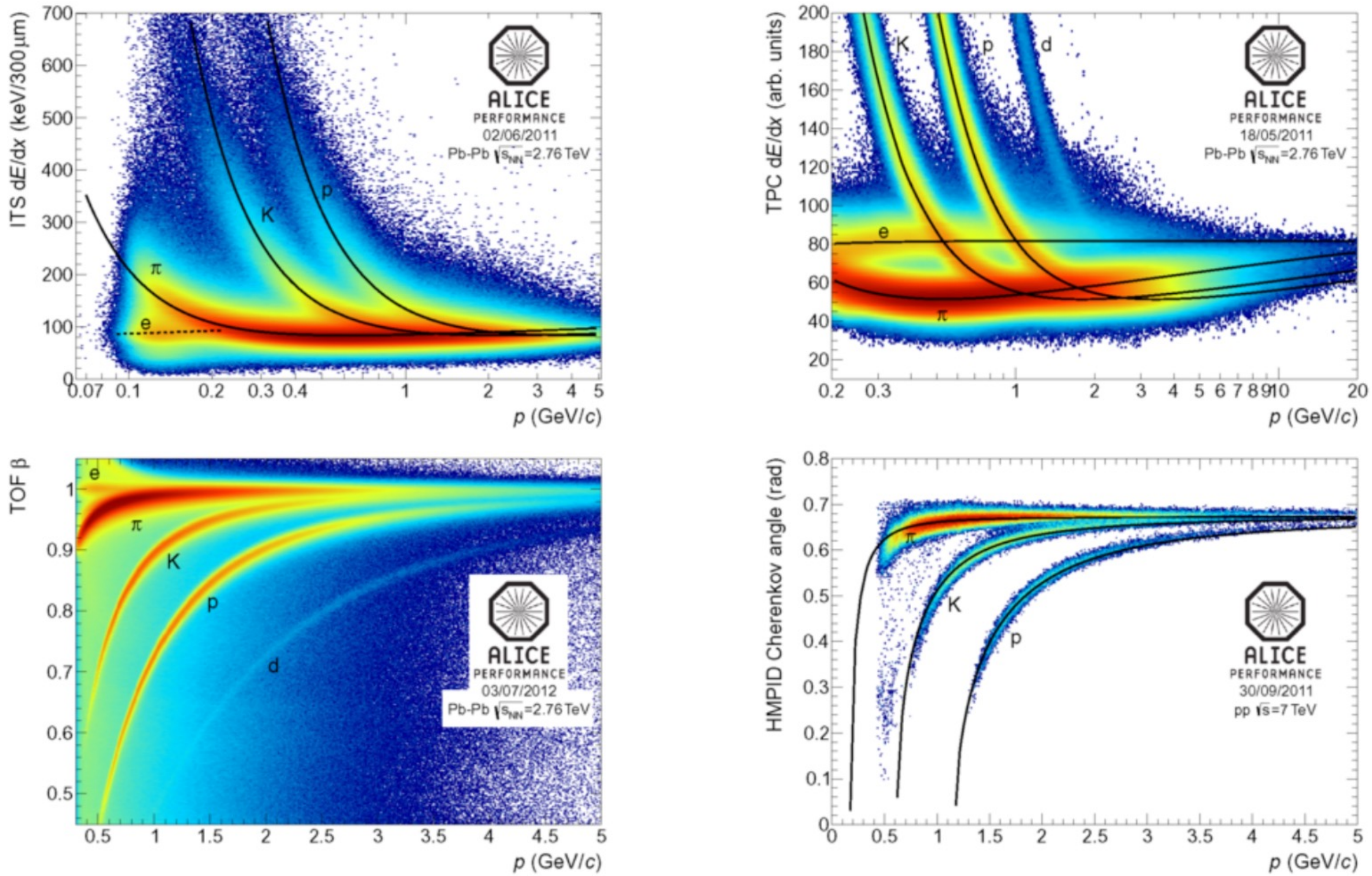


$\Rightarrow$ TOF usable for **low-p hadron ID** $\longrightarrow K - p$ up to 6GeV @ **50ps single hit resolution**

$\longrightarrow K - \pi$ up to 3.5GeV

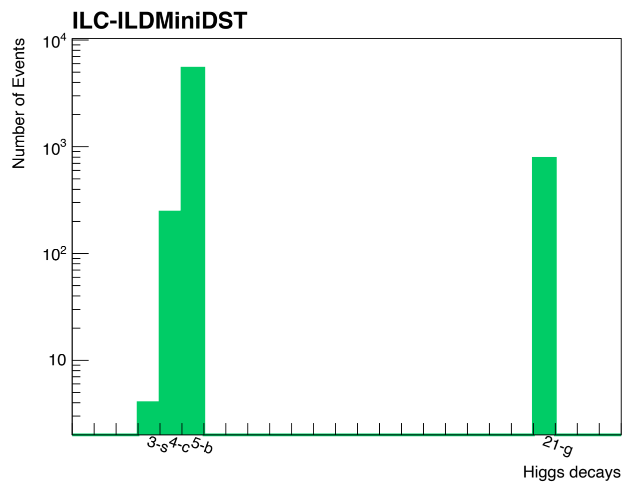
https://www.researchgate.net/publication/257190213_Identified_charged_hadron_production_measured_with_ALICE_at_the_LHC

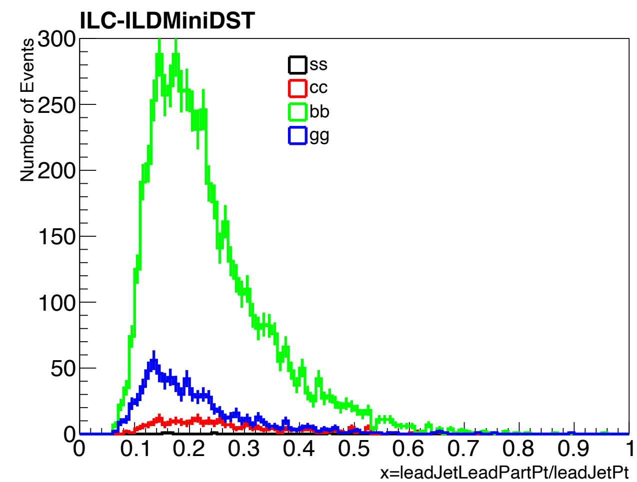
M. Ivanov / Nuclear Physics A 904–905 (2013) 162c–169c



Some more details

```
root [12] tree->Scan("leadJetLeadPartPdgId", "leadJetLeadPartPdgId>3000")
************************
   Row   leadJetLe *
************************
    474      3122 *   Λ
   1204      3122 *
   1797      3122 *
   2027      3122 *
   2116      3322 *   Ξ0
   2223      3122 *
   2567      3112 *
   2860      3122 *
   2994      3122 *
   3889      3122 *
   3930      3122 *
   4593      3222 *   Σ+
   5143      3122 *
   5148      3122 *
   5315      3122 *
   5346      3122 *
   5759      3122 *
   6264      3122 *
************************
==> 18 selected entries
```






```
root [16] tree->Scan("leadJetLeadPartPdgId", "higgsDecayPdgId==3")
************************
   Row   leadJetLe *
************************
   1180       211 *  π+
   1911       211 *
   4446       321 *  K +
   5532      -211 *
************************
==> 4 selected entries
```

BR

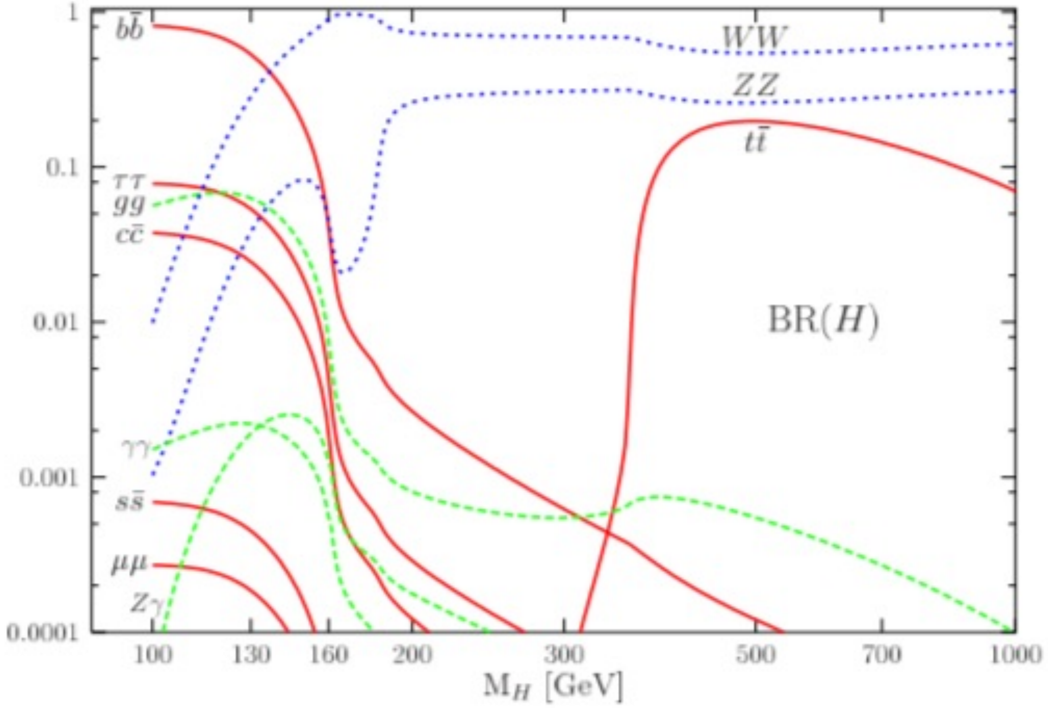


Fig. 1 From [5]: The decay branching ratios of the SM Higgs boson as a function of its mass.

[5] "*Electroweak Symmetry Breaking at the LHC*", A. Djouadi, R.M. Godbole, https://link.springer.com/chapter/10.1007%2F978-81-8489-295-6_5, https://arxiv.org/abs/0901.2030