Strange quark tagging with ILD to search for new physics in the Higgs sector

ILC WG3 Physics Meeting – September 16, 2021

Matthew Basso (Toronto) and **Valentina Cairo** (SLAC),
on behalf of everyone on the Snowmass 2021 LoI

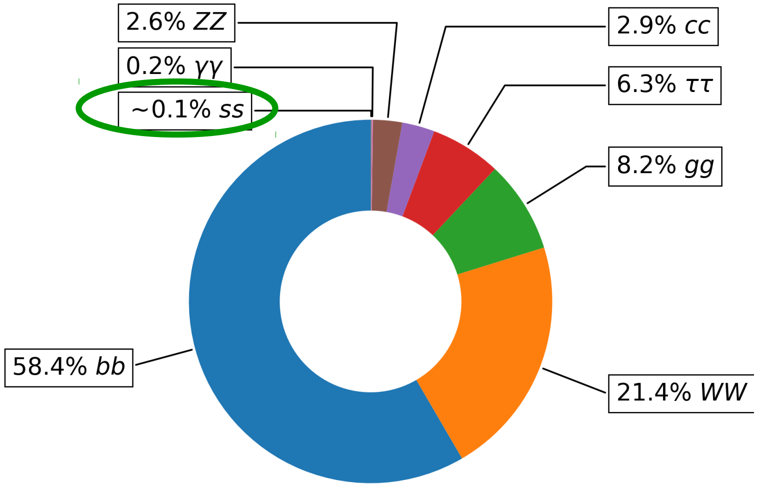
Overview

- Submitted a Letter of Interest as part of Snowmass 2021
  - Basic goal: develop a strange tagger using ILD@ILC and apply the tagger to a simple SM $H \rightarrow ss$ or BSM $H \rightarrow cs$ analysis
    - In line with ILC Snowmass 2021 study questions (2007.03650)
    - Interplay with the instrumentation: strange tagging capabilities strongly depend on the detector (e.g., PID)

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. Schwarzman[(b)], D. Su[(b)], E. Usai[(c)], C. Vernieri[(b)], C. Young[(b)]

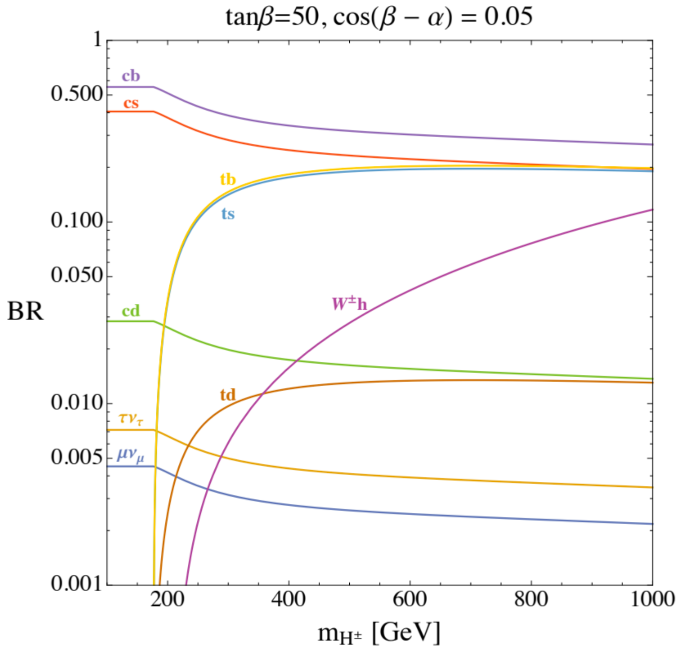
(a) University of Toronto, Toronto ON - Canada
(b) SLAC National Accelerator Laboratory, Stanford CA - USA
(c) Brown University, Providence RI - USA



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

$H\to ss$ and $H\to cs$

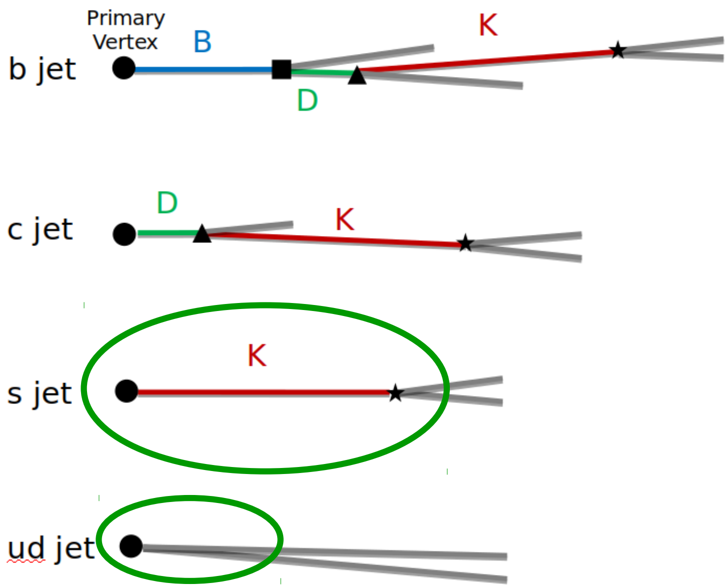
- $H\to ss$: likely to remain out of experimental reach unless enhanced relative to SM expectations
- $H\to cs$: some BSM models allow for the 1st & 2nd generation fermion masses to be an additional source of EW symmetry breaking, resulting in a "SM" Higgs doublet (125 GeV) and a "heavy" Higgs doublet
  - See 1610.02398 for instance
  - Predicts an ***enhancement*** to Higgs cross section
  - Charged heavy Higgs can undergo flavour violating decays (e.g., *cs*) – *s*/*c*-tagging can help here



Charged heavy Higgs branching ratios. Taken from Fig. 6 of 1610.02398.

Different jet types, pictorially

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.)

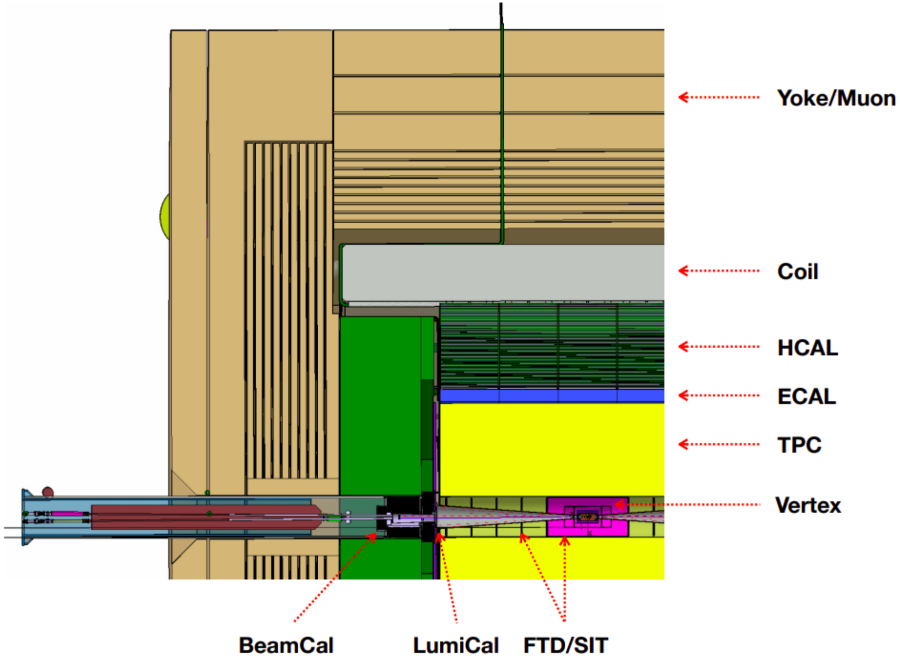
Remember to normalize the discriminants to make them boost invariant (as much as possible)

5

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

ILD@ILC

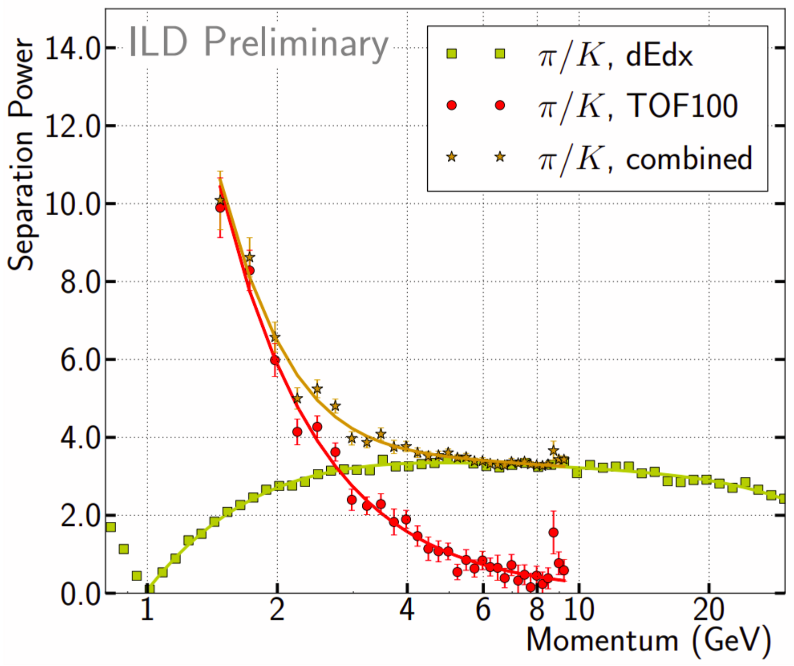
- The ILD detector
  - Detector overview: 1912.04601
  - 3 double-layer pixel detectors for vertexing
  - Time projection chamber (TPC) for tracking with inner/outer Si layers
    - Low material assists in low-*p* tracking
  - High granularity sampling calorimeters for particle flow reconstruction
    - Challenge is reconstructing neutral hadrons
    - Precise EM/hadronic design still under study
  - Tracking/calorimetry contained in 3.5 T field



ILD detector quadrant. Taken from Fig. 1 of 1912.04601.

Flavour tagging requirements

- Good impact parameter resolution, secondary vertexing
  - Pertinent to *b*/*c*-tagging
- For strange versus up/down ("light") quark tagging, there's a need for **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 already provides BDT scores for *b*/*c*-taggers and an other ("*o*") tagger per jet – these can be utilized



ILD separation power for pions and kaons using dE/dx and TOF (100 ps resolution). Taken from Fig. 3 of 1912.04601.

Strange jet tagger

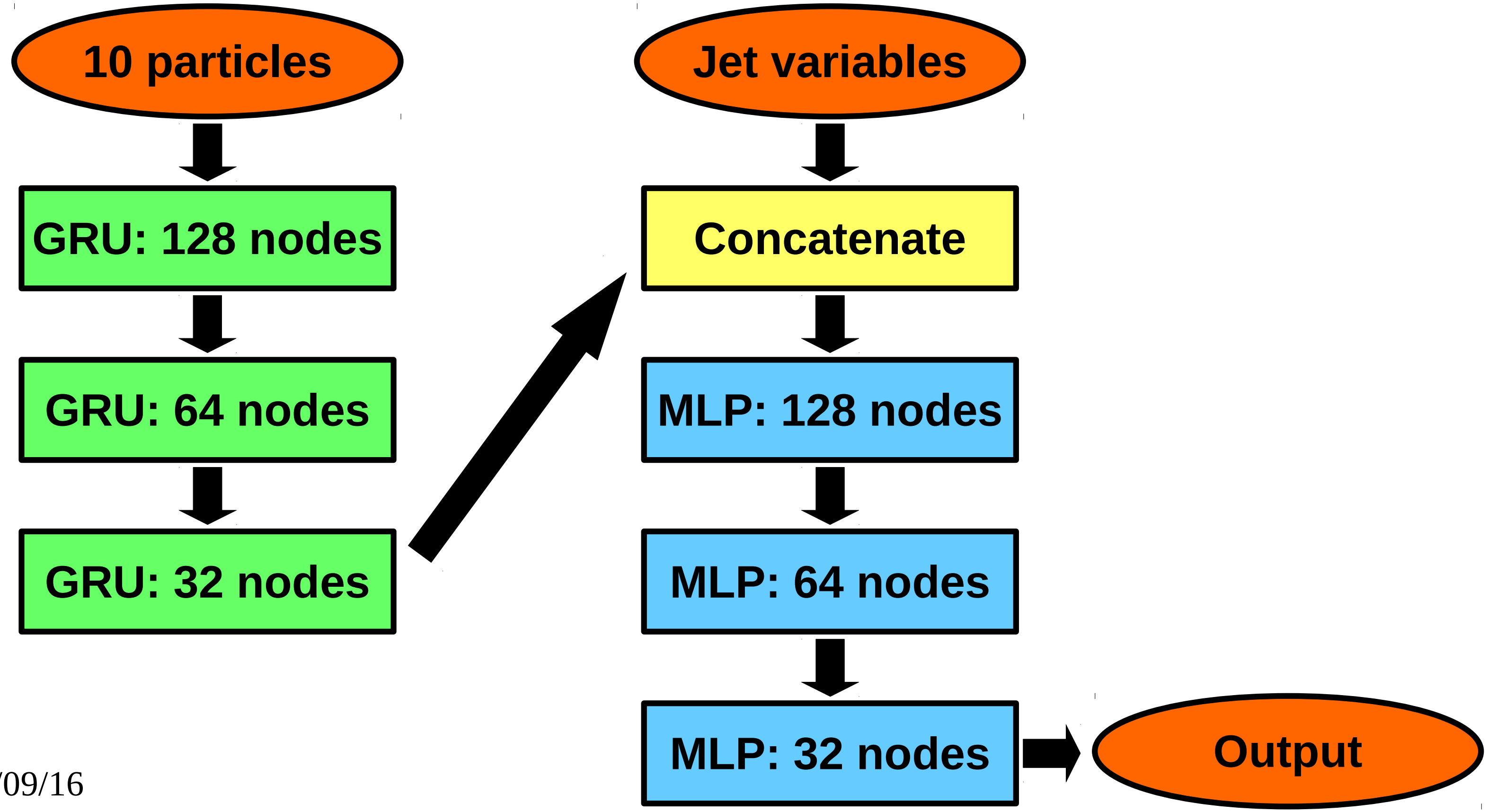
Multiclassifier tagger and inputs

- Use a *multiclassifier* tagger, which assigns probabilities to the possible flavours of a jet simultaneously
- Train on ILD-reconstructed $H \rightarrow qq/gg$ samples ($qq = uu$, dd, ss, cc, bb) with $\sqrt{s} = 250$ GeV and $P_L[e^-] = -100\%$ and $P_R[e^+] = +100\%$
  - *Unskimmed*, except for $N_{jets} \geq 2$, $N_{leptons} = 0$, and truth $H \rightarrow qq/gg$ 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*/*o*-tagger scores, category, $N_{particles}$
  - Particles: p, η, ϕ, m, charge, **truth** electron/muon/pion/kaon/proton likelihoods (0 or 1, using PDG ID – dE/dx and TOF likelihoods not yet fixed in samples)
    - "Kaons": K_S^0 (310), $K^{+/-}$ (321), Λ (3122)

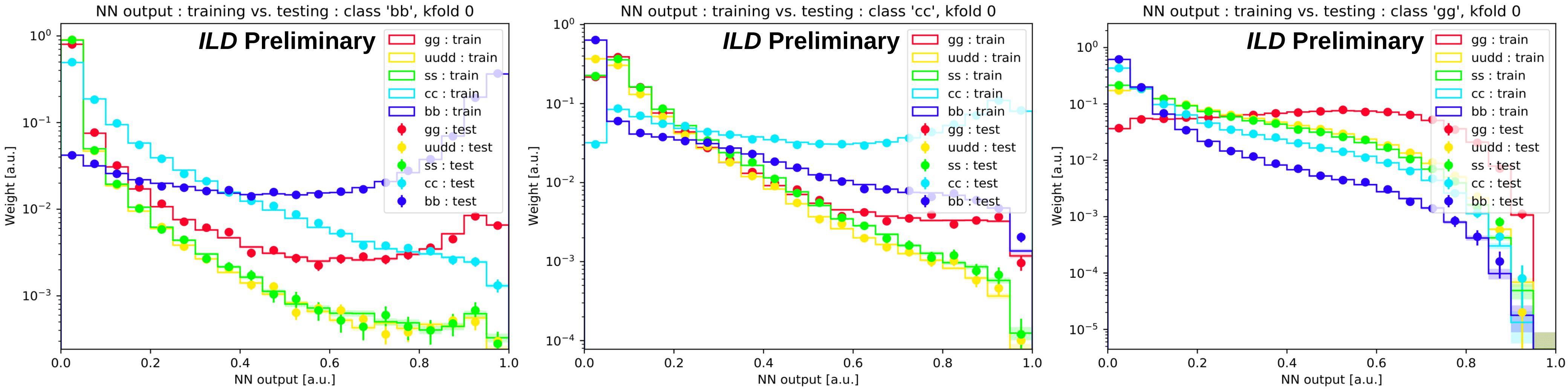
Tagger architecture

- We picked a relatively simple neural network (NN) architecture
  - 3 layer (128→64→32 nodes) recurrent neural network (using gated recurrent units) for particle-level inputs – then concatenated with jet-level inputs and fed into a 3 layer (128→64→32 nodes) multilayer perceptron
  - Architecture shows up in many different HEP measurement scenarios (e.g., recent ATLAS $H \to ZZ \to 4\ell$ couplings measurement, see Section 5.2 of 2004.03447); specifically, applied even to strange tagging performance at **hadron** colliders (used LSTMs instead of GRUs)
    - "Maximum performance of strange-jet tagging at hadron colliders" (2011.10736)

Tagger architecture: pictorially

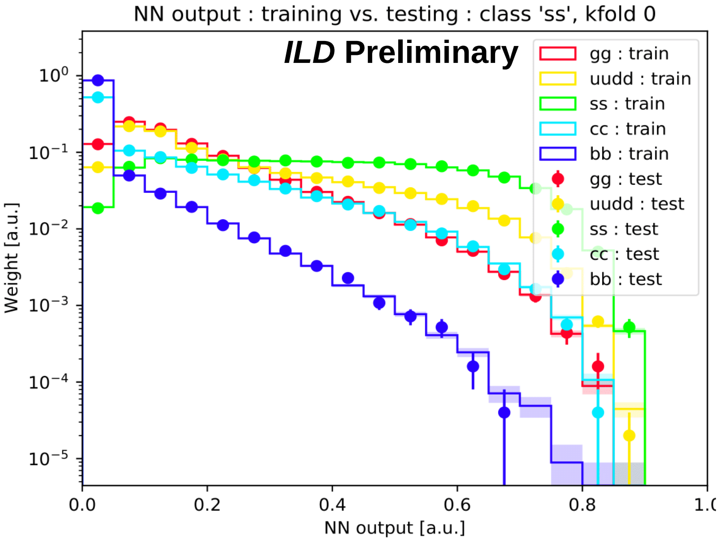


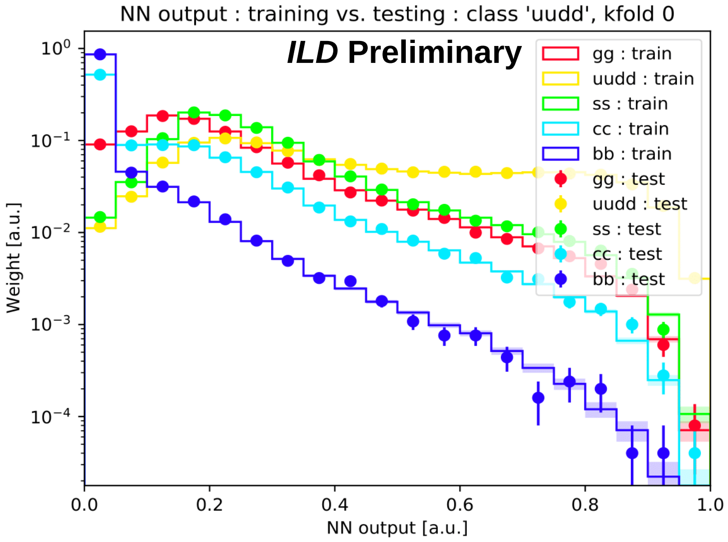
Performance: *b*, *c*, and *g* jets



- 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
- At 50% strange tagging efficiency, we have **90%** background rejection over **70%** for LCFIPlus Otag (see ROC curves in Backup and LCWS2021 talk)

$H \to ss$ analysis

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 (~1M events each)
  - 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

Analysis cuts

- Preliminary selection:
  - Leading and subleading jet momenta, $p_j > 30$ GeV
  - Dijet mass, $M_{jj} \epsilon$ [120, 140] GeV
  - Dijet energy, $E_{jj} \epsilon$ [125, 160] GeV
  - Missing mass, $M_{\text{miss}} \epsilon$ [75, 120] GeV
  - Angular separation, $\Delta R_{jj,\text{miss}} = \sqrt{(\Delta\phi_{jj,\text{miss}}{}^2 + \Delta\eta_{jj,\text{miss}}{}^2)} < 4$
  - Leading and subleading LCFIPlus tagger scores, $\text{score}^b{}_j < 0.2$ && $\text{score}^c{}_j < 0.35$
  - Number of PFOs per event, N_{PFOs}/event ϵ [30, 60]
  - Number of PFOs per jet, N_{PFOs}/jet ϵ [10, 40]

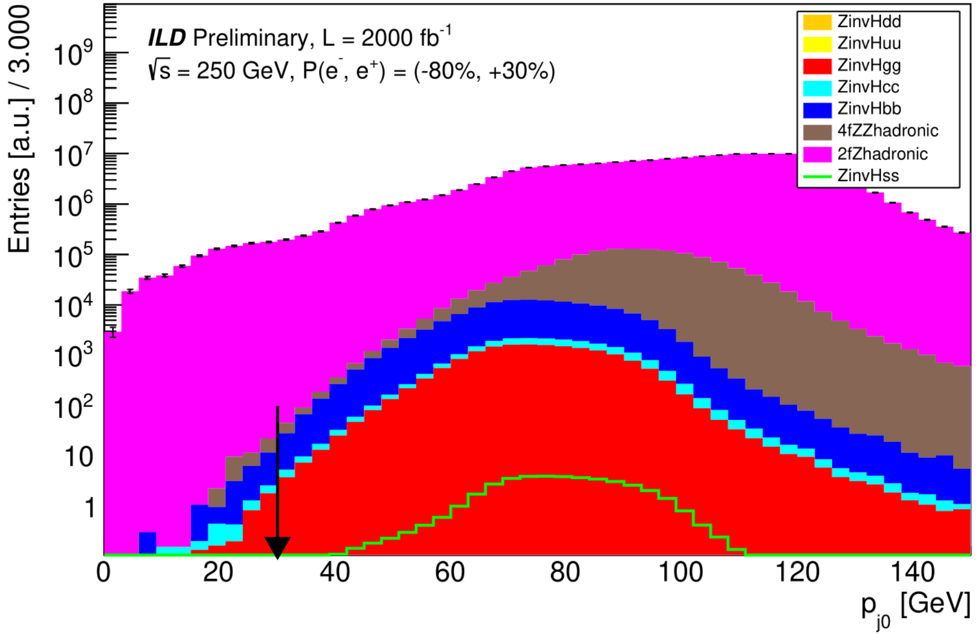
Cutflow

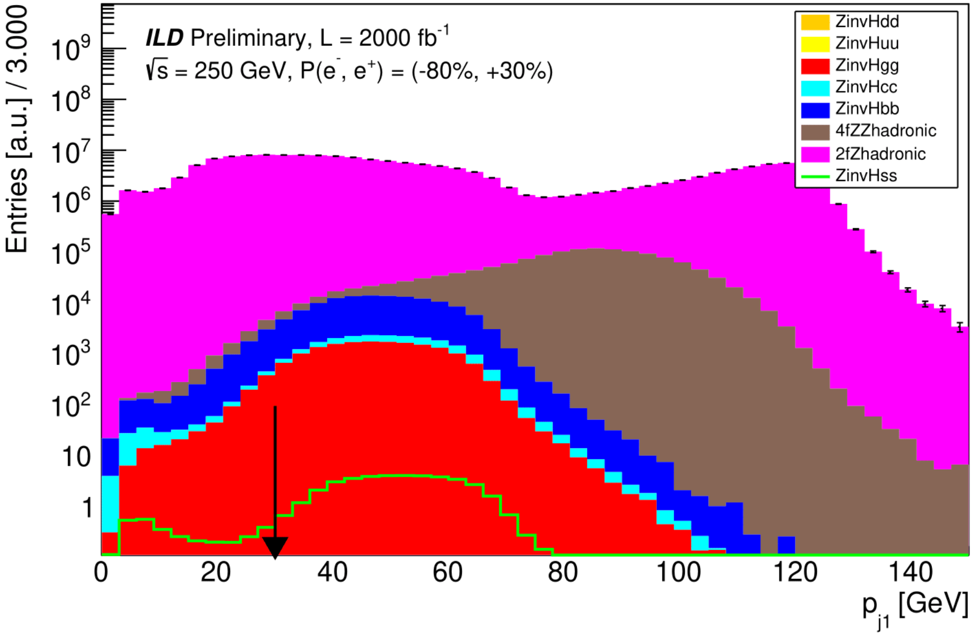
ILD Preliminary, $\mathcal{L} = 2000\ \text{fb}^{-1}$, $\sqrt{s} = 250$ GeV, $P(e^-, e^+) = (-80\%, +30\%)$

	$(H \to s\bar{s})(Z \to \nu\nu)$	$(H \to gg)(Z \to \nu\nu)$	$(H \to u\bar{u}/d\bar{d})(Z \to \nu\nu)$	$(H \to c\bar{c})(Z \to \nu\nu)$	$(H \to b\bar{b})(Z \to \nu\nu)$	$Z \to q\bar{q}$	$ZZ \to q\bar{q}q\bar{q}$	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

- We use Durham2Jets, so we always have 2 jets
- Largest **decrease** in signal efficiency at M_{jj} cut
  - Provides one of the strongest handles on reducing 2f_Z_hadronic, however
- Net result: 29% signal efficiency, 0.016% background efficiency
  - *H→bb s/b* ~ 0.0007 @ No cut can be compared to that in the 4-jet channel, 0.00077 (Backup)

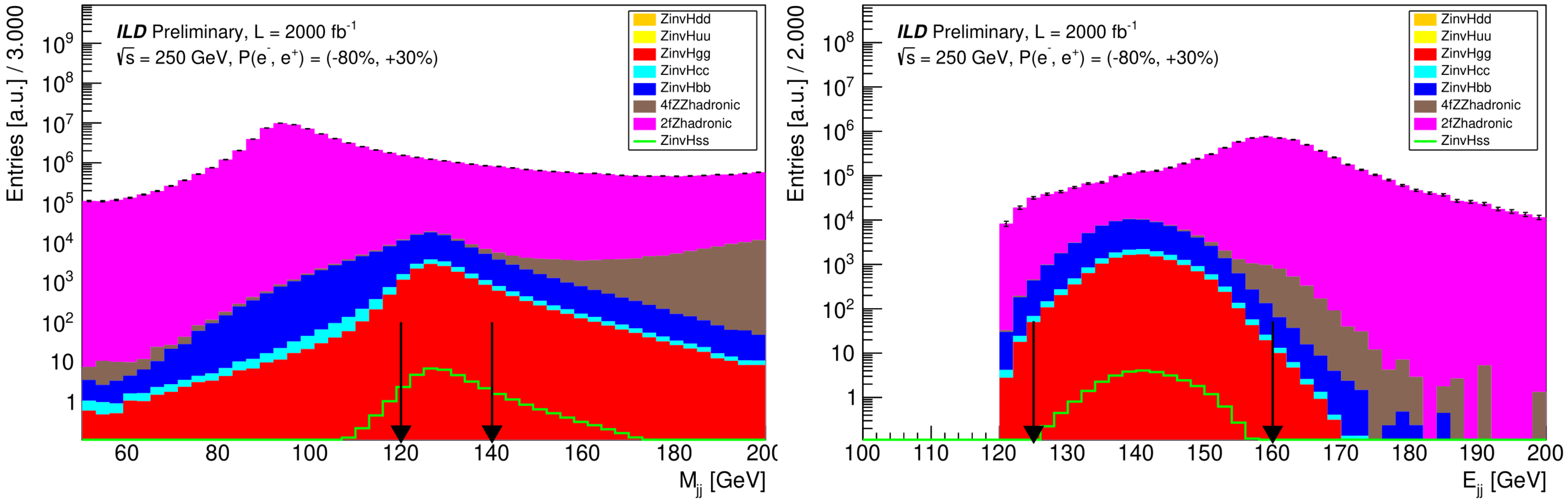
Histograms: p_{j0} and p_{j1}





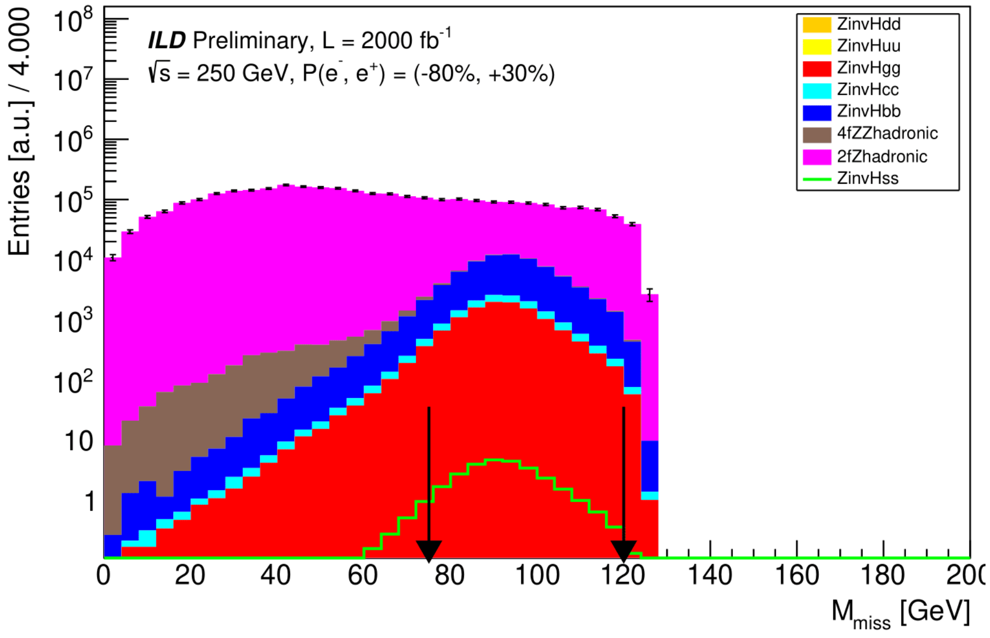
****Unstacked green line is signal****

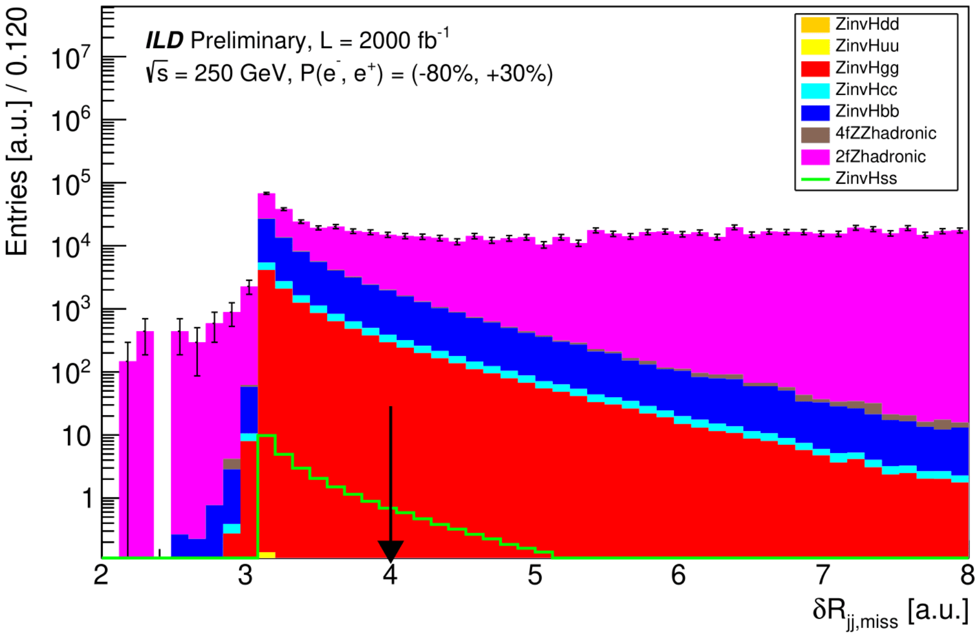
Histograms: M_{jj} and E_{jj}



Unstacked green line is signal

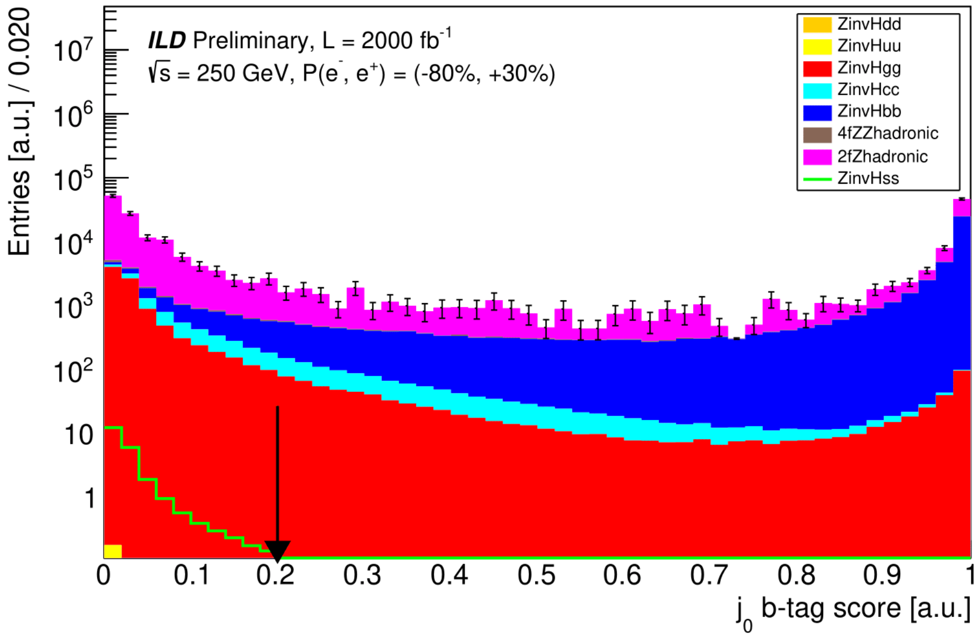
Histograms: M_{miss} and $\Delta R_{jj,miss}$

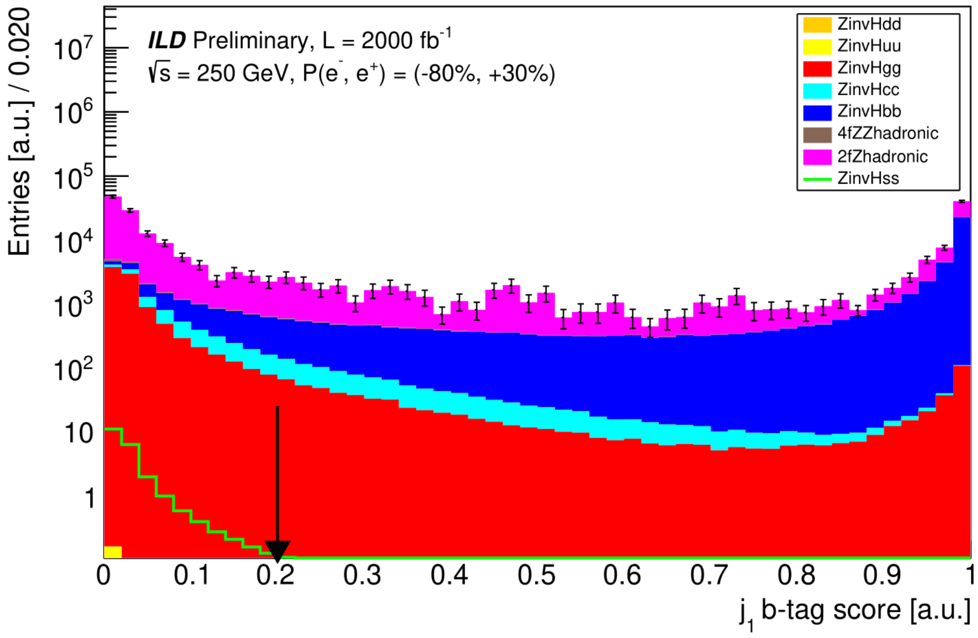




****Unstacked green line is signal****

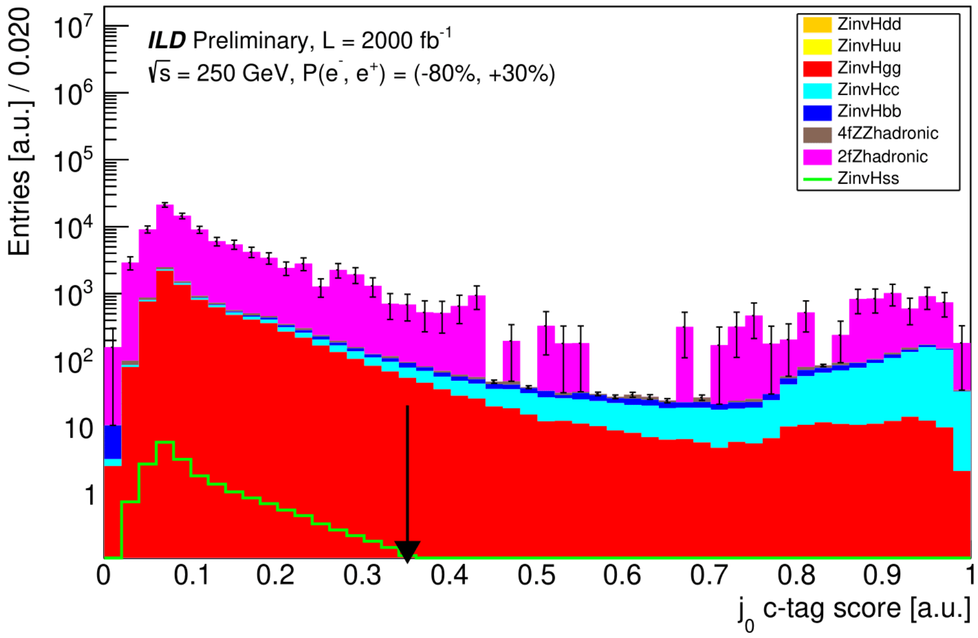
Histograms: *b*-tagger scores

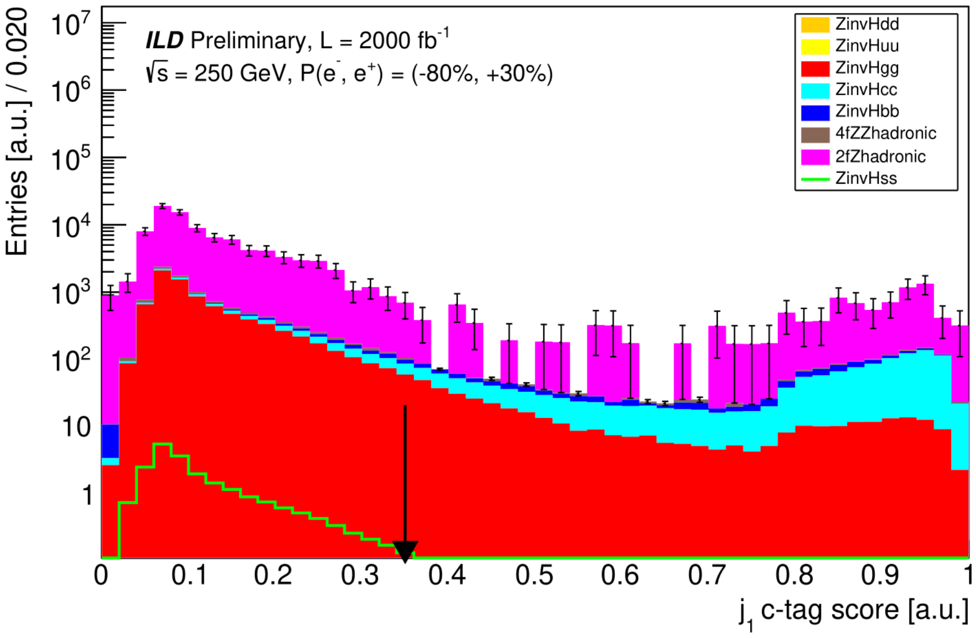




****Unstacked green line is signal****

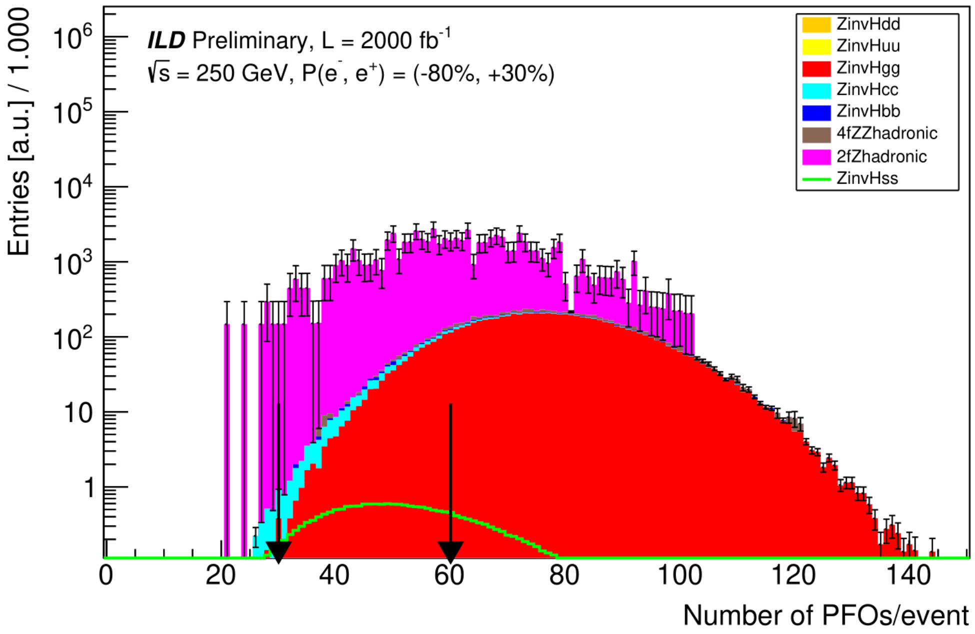
Histograms: *c*-tagger scores





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

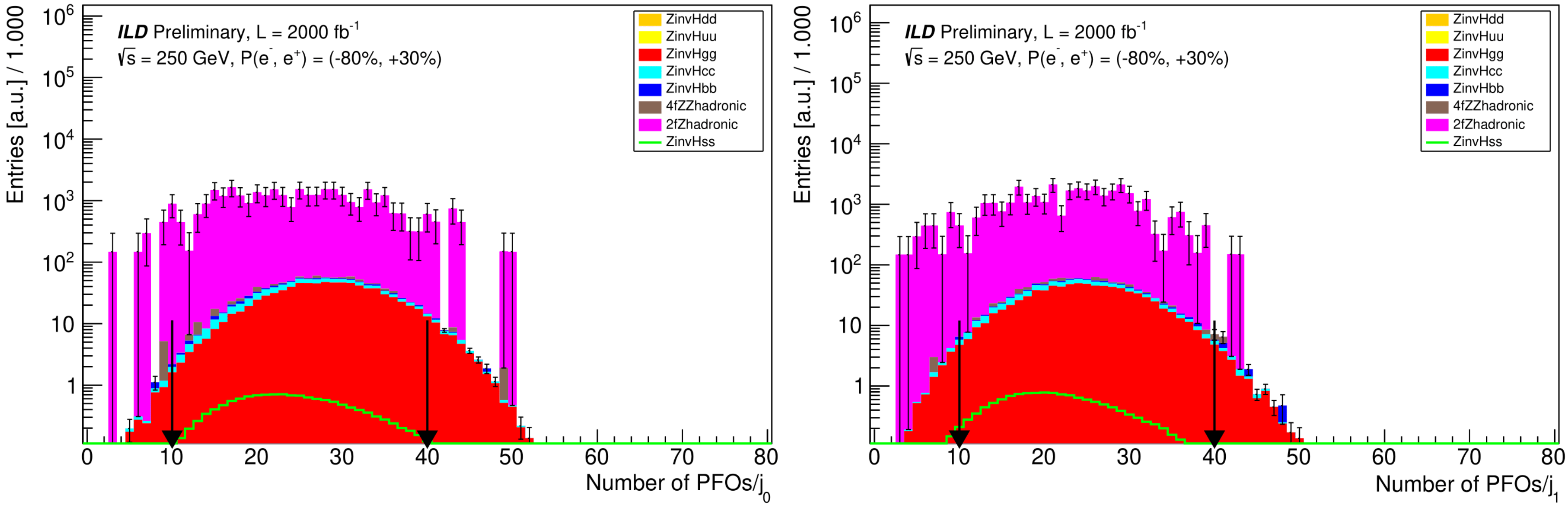
Histograms: N_{PFOs}/event



Should we soften this cut to keep more signal? 67% efficiency for signal, but 11% efficiency for *H→gg*

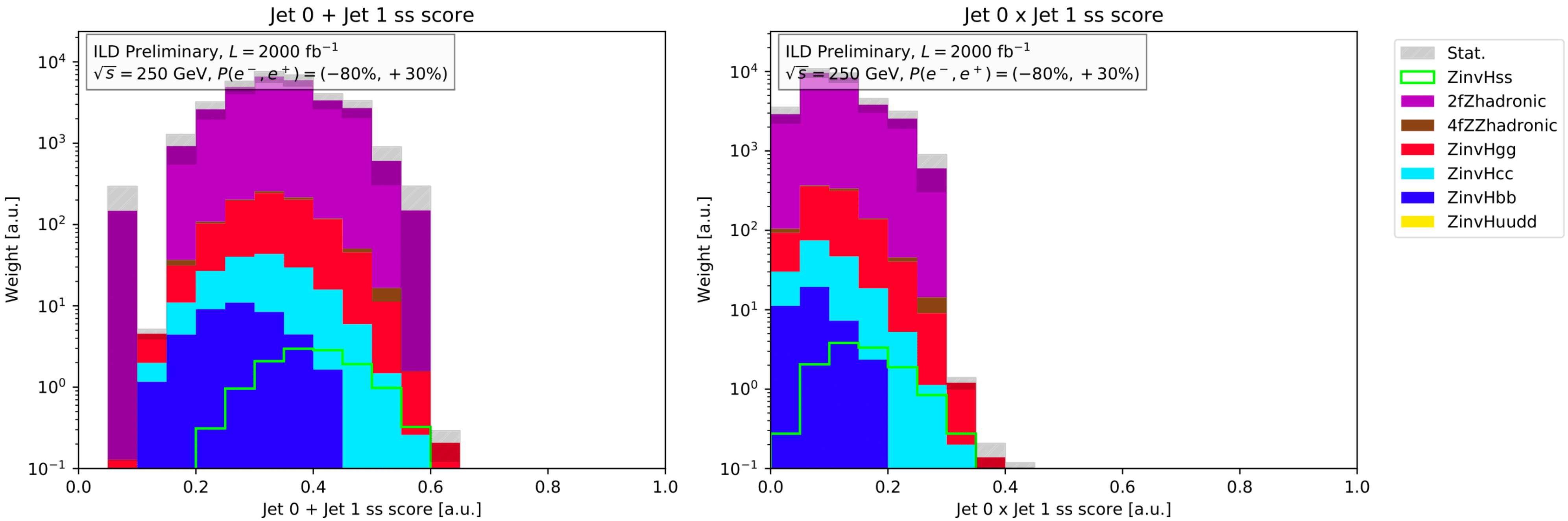
****Unstacked green line is signal****

Histograms: N_{PFOs}/jet



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

Signal discriminant(s)

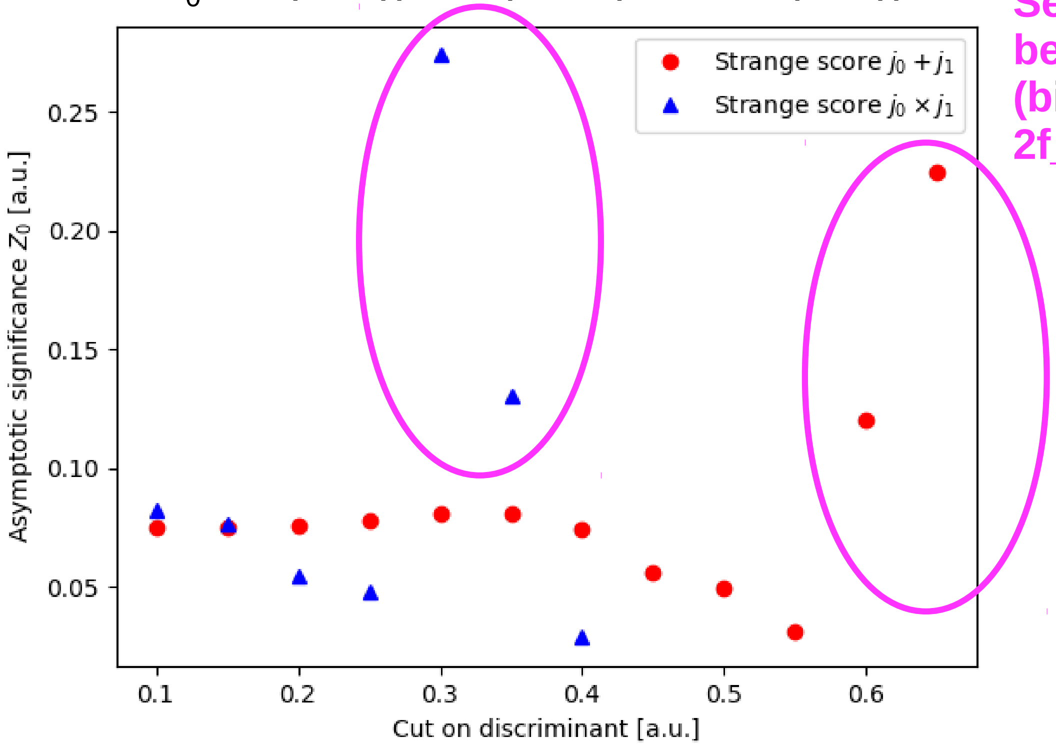


We can look at sum of leading and subleading jet strange scores (right) or product (left)

Signal discriminant (2)

Using asymptotic significance assuming Asimov data (neglecting MC stats):

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



Sensitivity peaks only because s < 1 and b ~ O(1) (bins not populated by 2f_Z_hadronic)

Discussion

- Generally, cutting on the sum of the strange jet scores seems more stable
  - Discovery measurement seems ***unlikely***, even after using truth info in the tagger
  - We will convert the results into limits on the Higgs to strange BR as the “money” plot (neglecting stat uncertainties on 2f_Z_hadronic? Would likely dominate the result, and these can probably be reduced with enlarged samples)
- Any gains in the analysis would come from reducing the 2f_Z_hadronic background
  - Tricky, as even now we only keep <10 raw events per 15,000 raw events when processing this background (similar for 4f_ZZ_hadronic)
  - As a suggestion from the ILD community, quantities like $\Delta\phi_{jj}$ or p_T^j should help

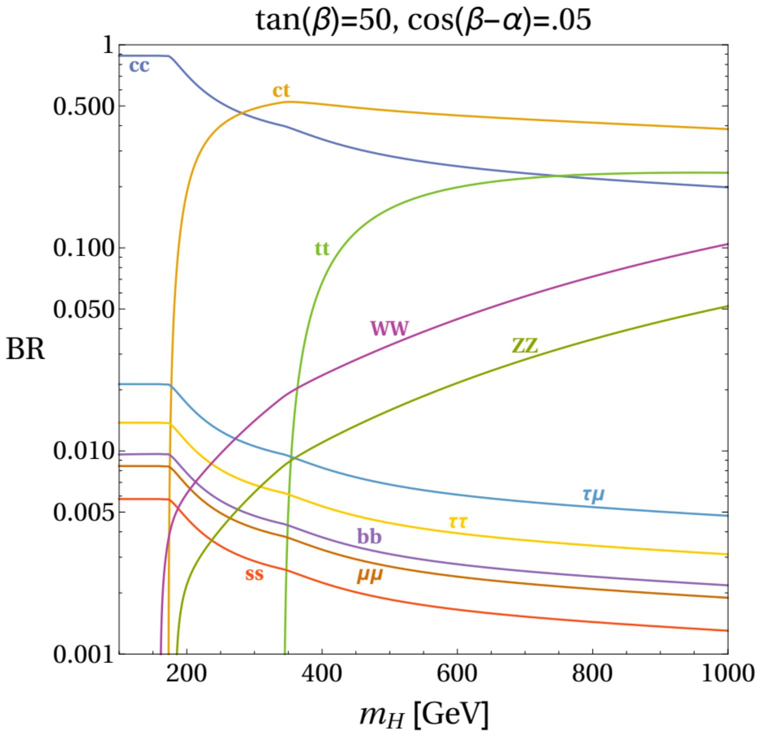
Summary

- Presented progress towards strange tagging with ILD and a $H \rightarrow ss$ analysis
  - Sensitivity is ***limited*** – we are looking at the best case scenario in terms of tagger (save other architectures), analysis could better optimized more, however
  - We are also happy to scale the results to other scenarios
  - More statistics for both 2f_Z_hadronic and 4f_ZZ_hadronic would be beneficial – at the same time, having MC for a 2HDM $H \rightarrow cs$ decay could provide prospects for BSM scenarios with enhanced yields
    - *Ongoing discussion on whether to enlarge samples statistics*
  - Looking to write up what has been done as a contribution to Snowmass 2021

Questions?

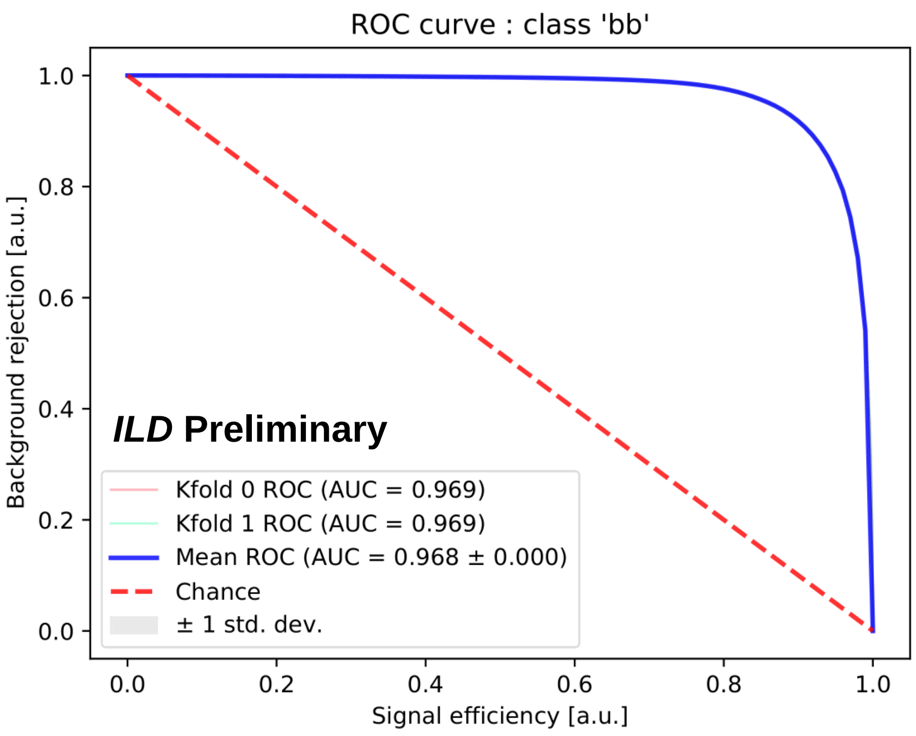
Backup

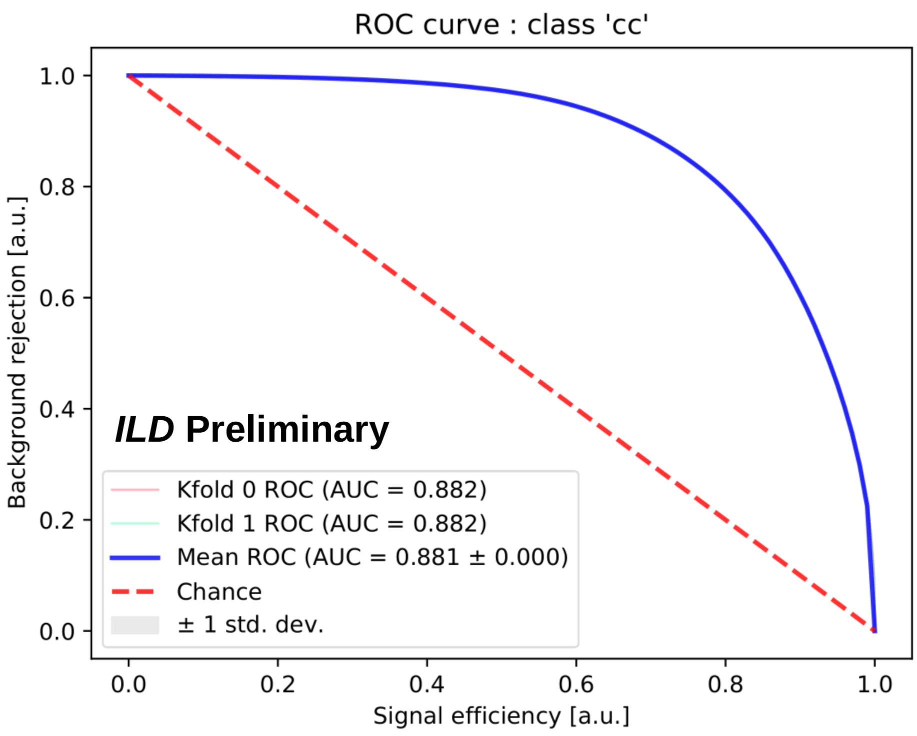
Neutral heavy Higgs BRs



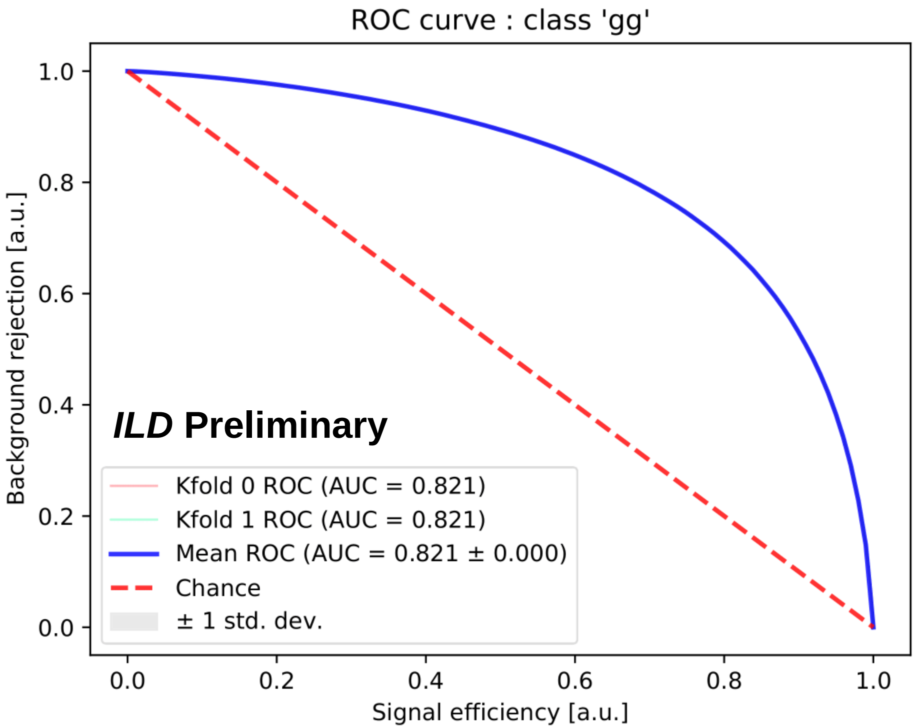
Neutral heavy Higgs branching ratios. Taken from Fig. 3 of 1610.02398.

ROC curves: *b* and *c* jets

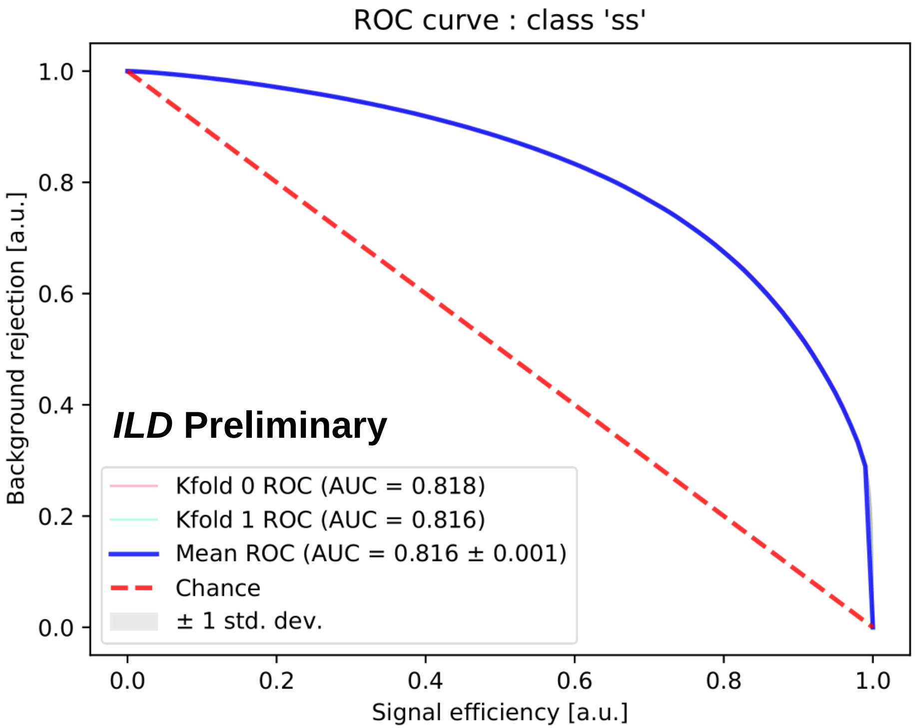


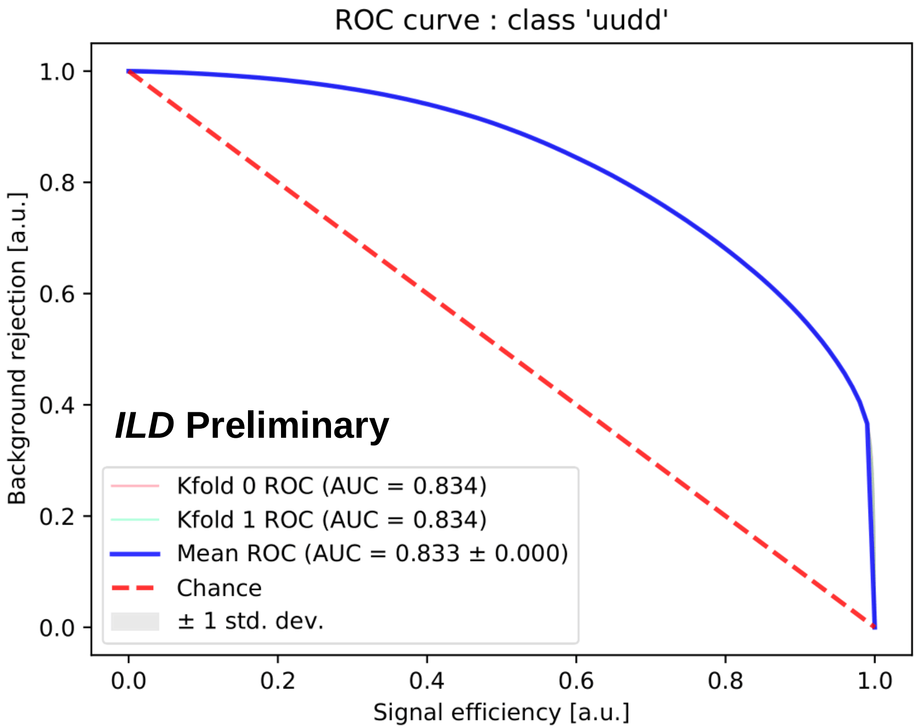


ROC curves: *g* jets



ROC curves: *s* and *u*/*d* jets





e^+e^- cross sections

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

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 Cross-section (fb)		fully polarization 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

H→bb analysis: histograms

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

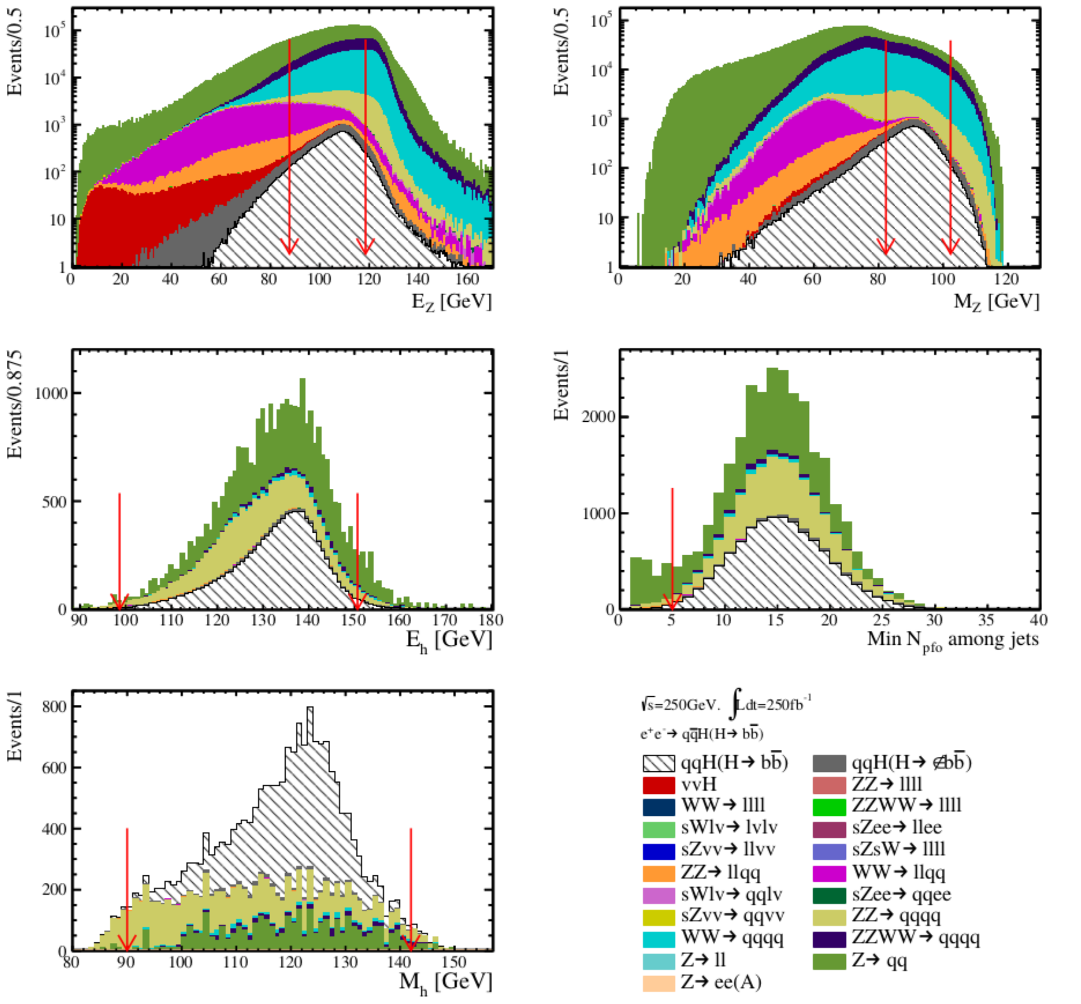


Figure 66: The distributions show each observable used for the background suppression assuming 250 fb^{-1} with P(e^-, e^+)= (-80%,+30%). The explanation of the observables are given in the text. Red arrows on each plot indicate the cut values applied to each observable as the background suppression.

$H\to bb$ 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