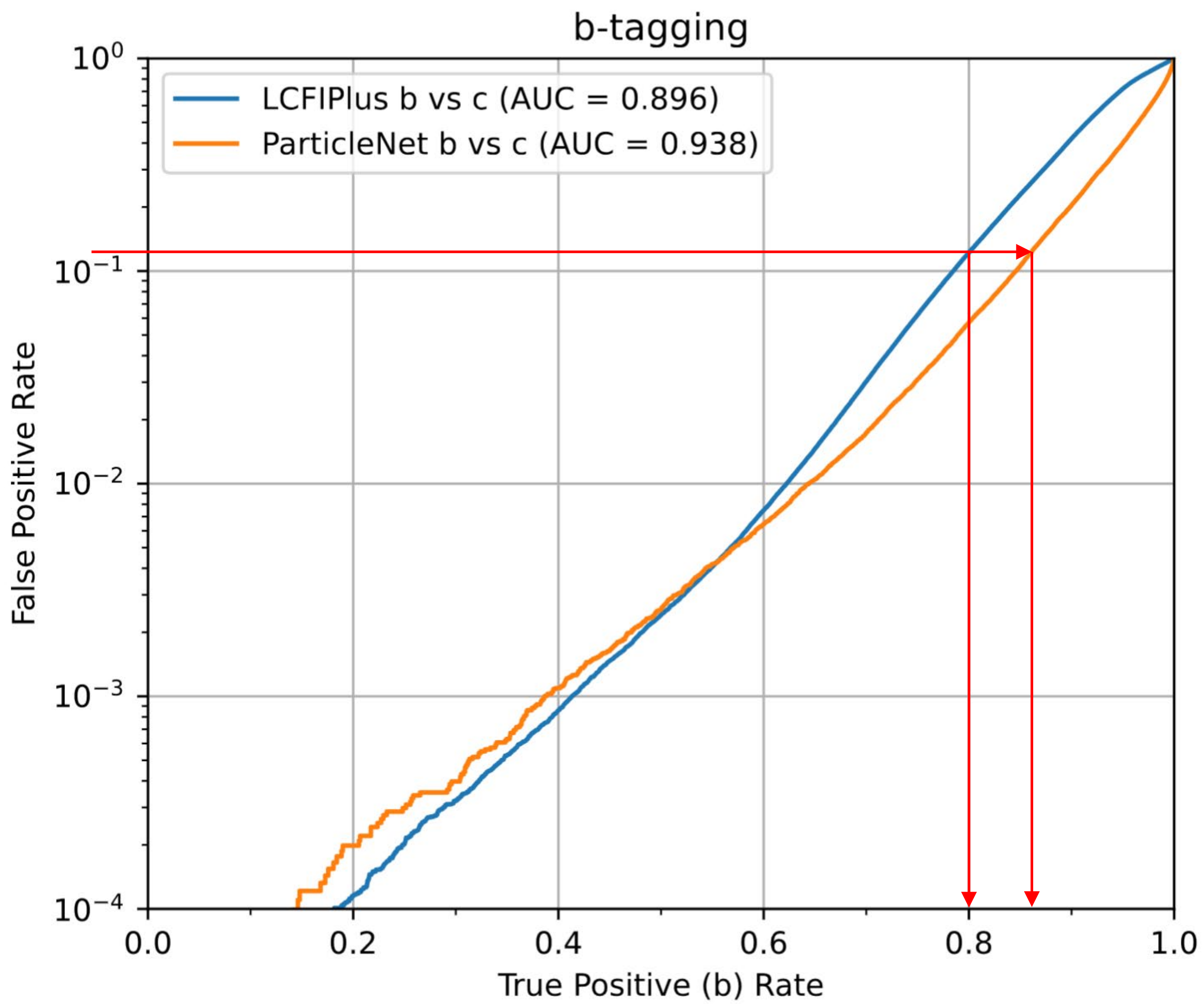
Status on Flavor Tagging@ZHH

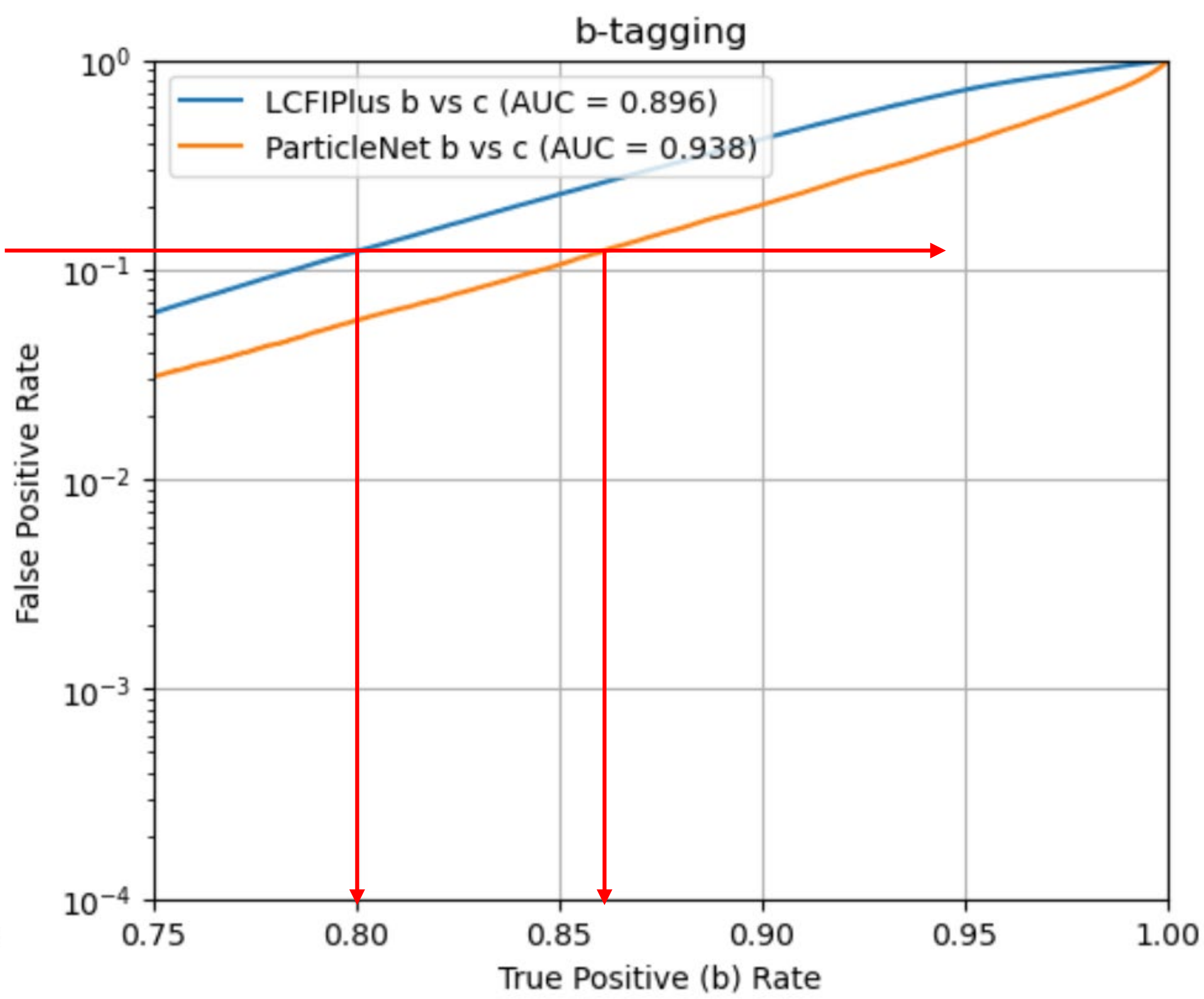
ZHH Topical Group Meeting | 2025/1/22

Bryan Bliewert, DESY & UHH

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Flavor Tagging: b vs. c



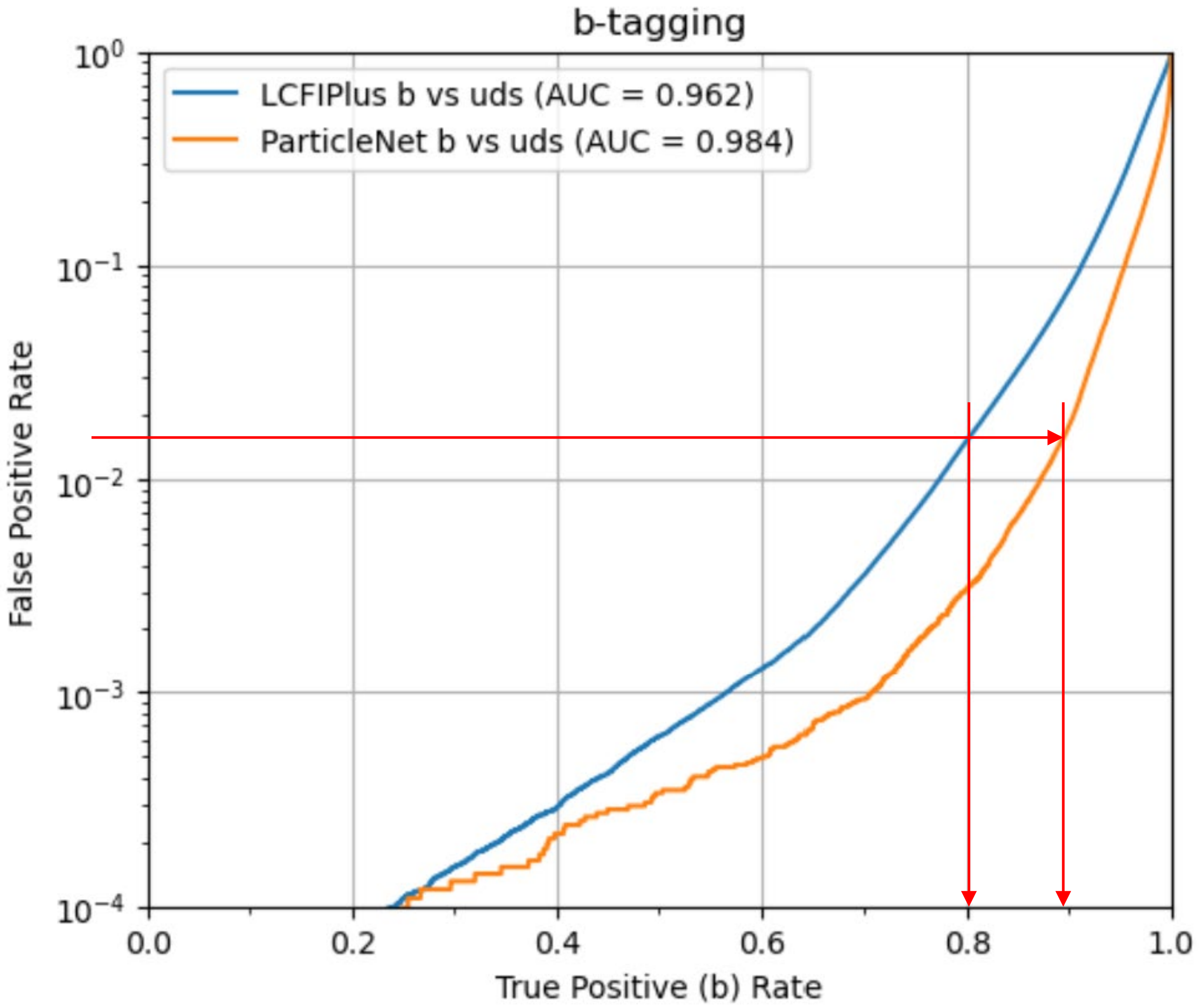


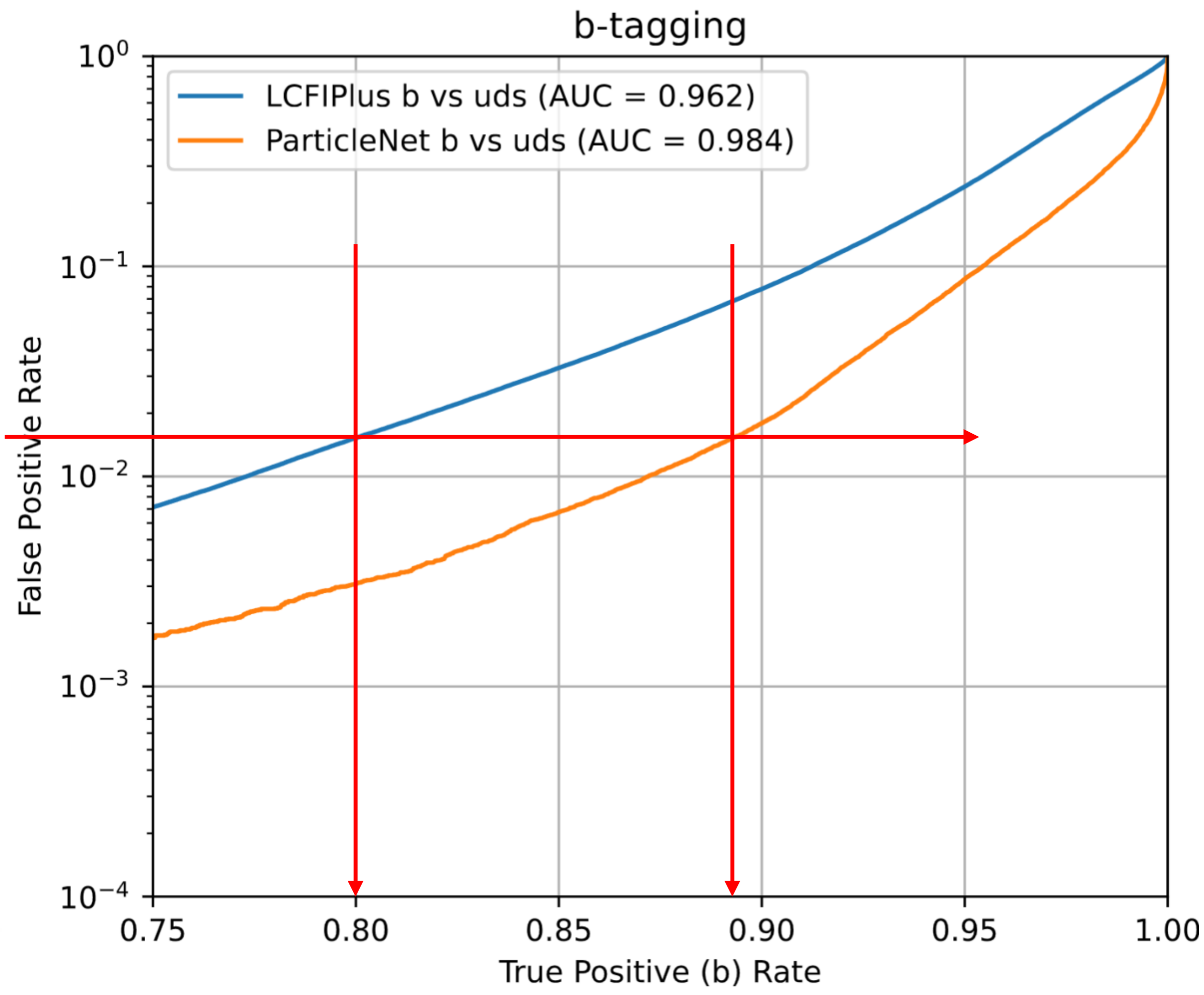
b vs. c @ same background rejection:
7% absolute / **9% relative improvement**

Higher performance can be expected when

- PID is properly included
- Technicalities regarding the 500 GeV flavortag samples are resolved

Flavor Tagging: b vs. light



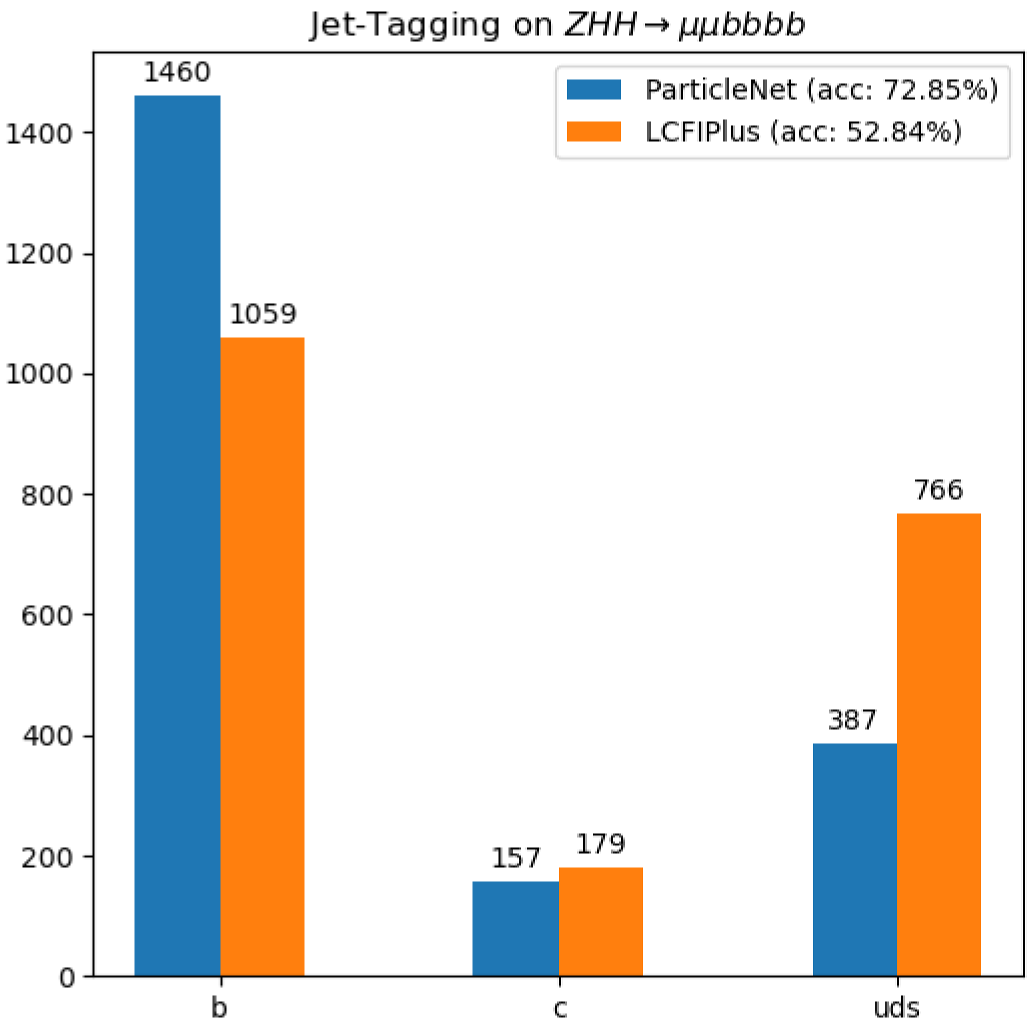


b vs. uds @ same background rejection:
9% absolute / **11% relative improvement**

Higher performance can be expected when

- PID is properly included
- Technicalities regarding the 500 GeV flavortag samples are resolved

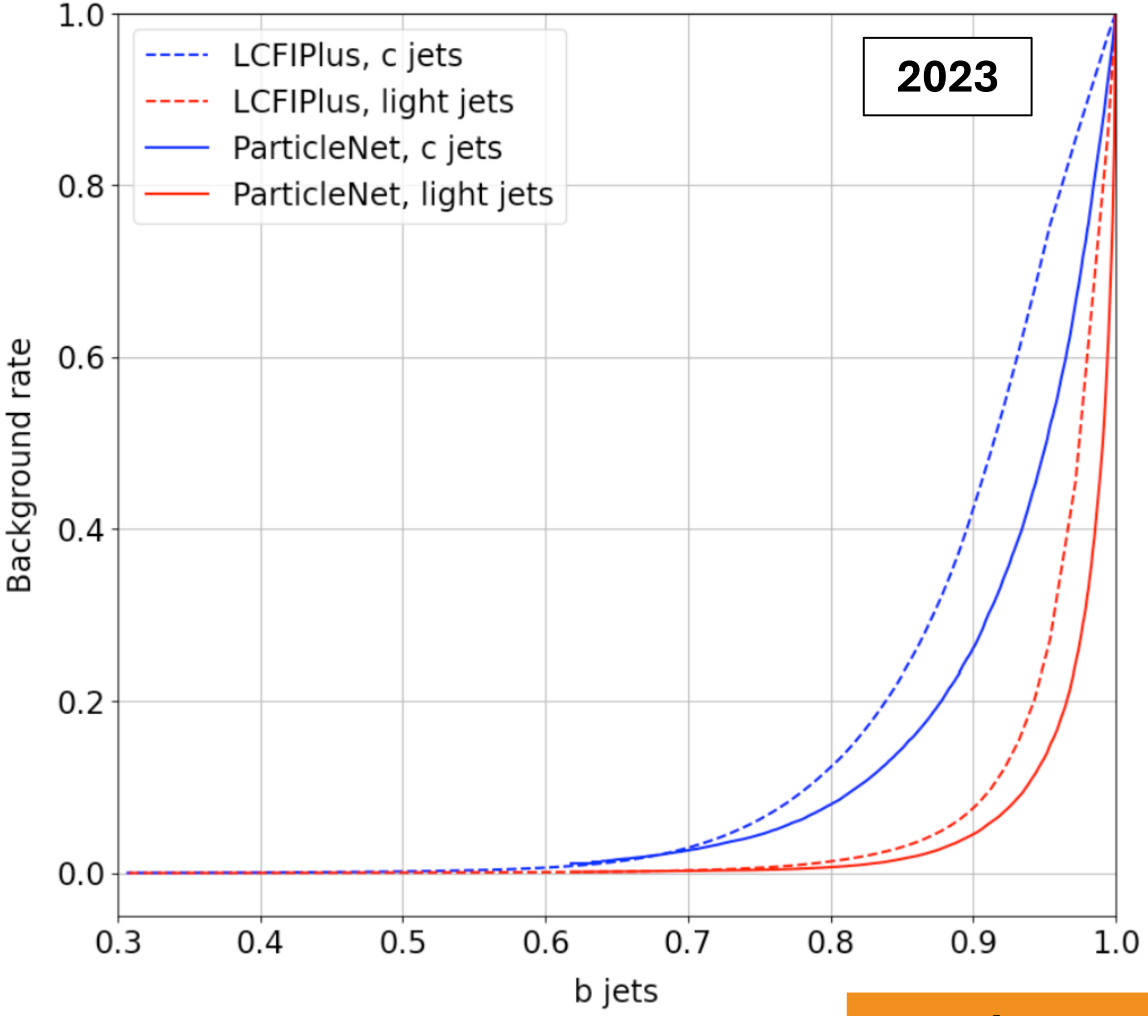
B accuracy on ZHH -> μμbbbb events

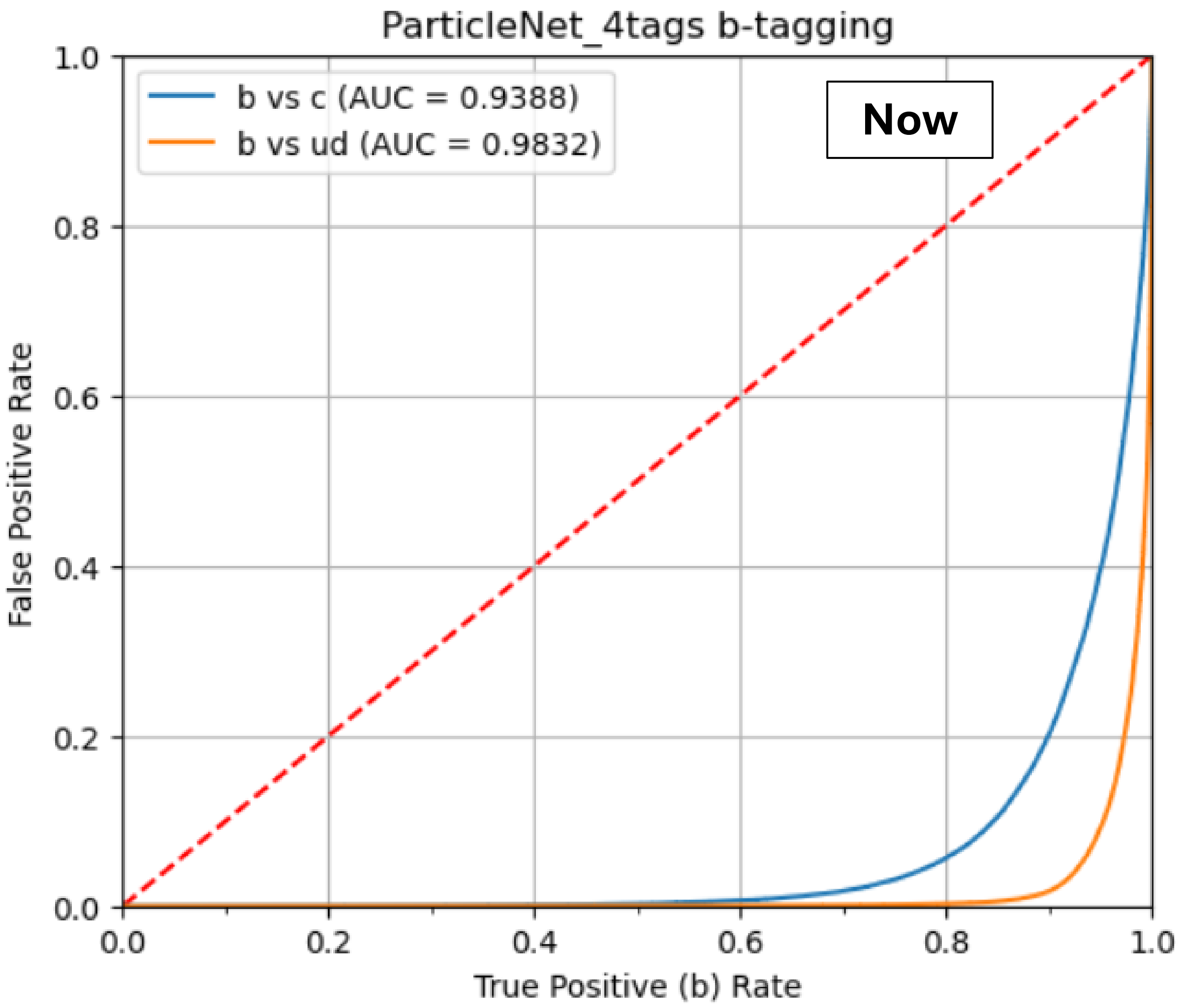


ROC curve to be added today

Backup

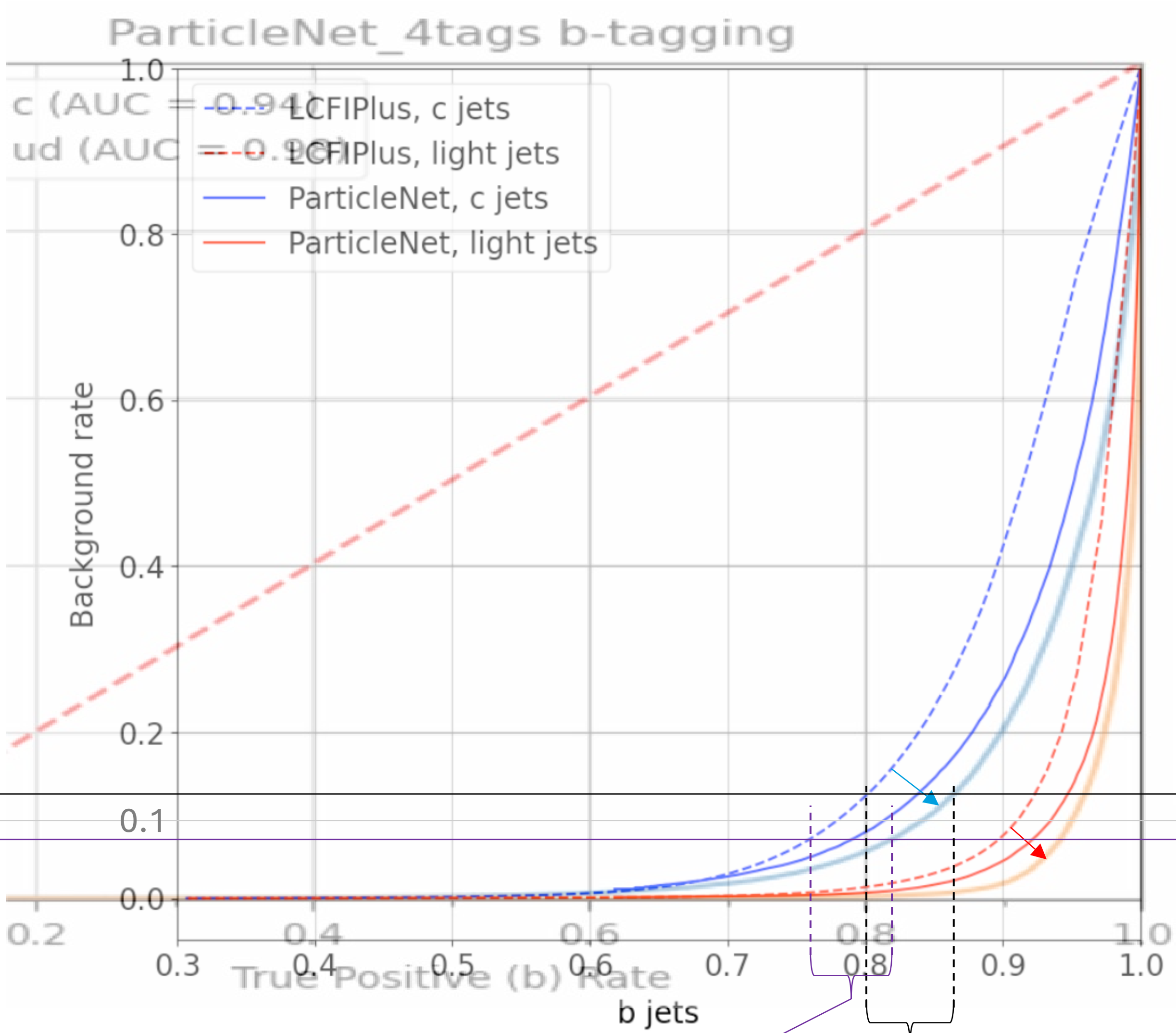
I Flavor Tagging - ParticleNet





Lower right is better

I Flavor Tagging - ParticleNet



Comparison of b-vs.-c at ILD@ILC500

ParticleNet (current)

LCFIPlus (current)

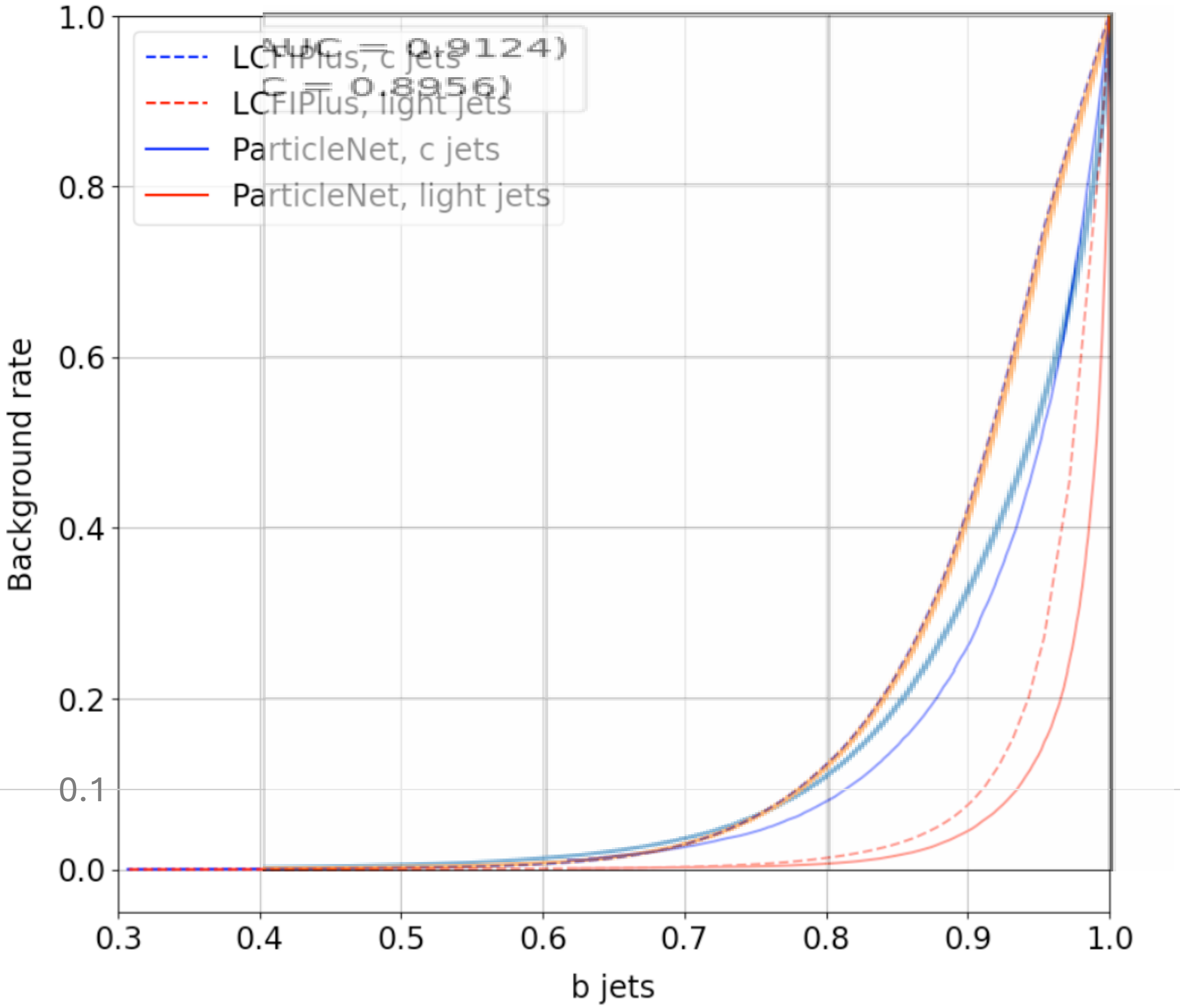
13% FP-rate @ $\epsilon_b = 80\%$

8% FP-rate

6% increase (9% relative)

7% increase (9% relative)

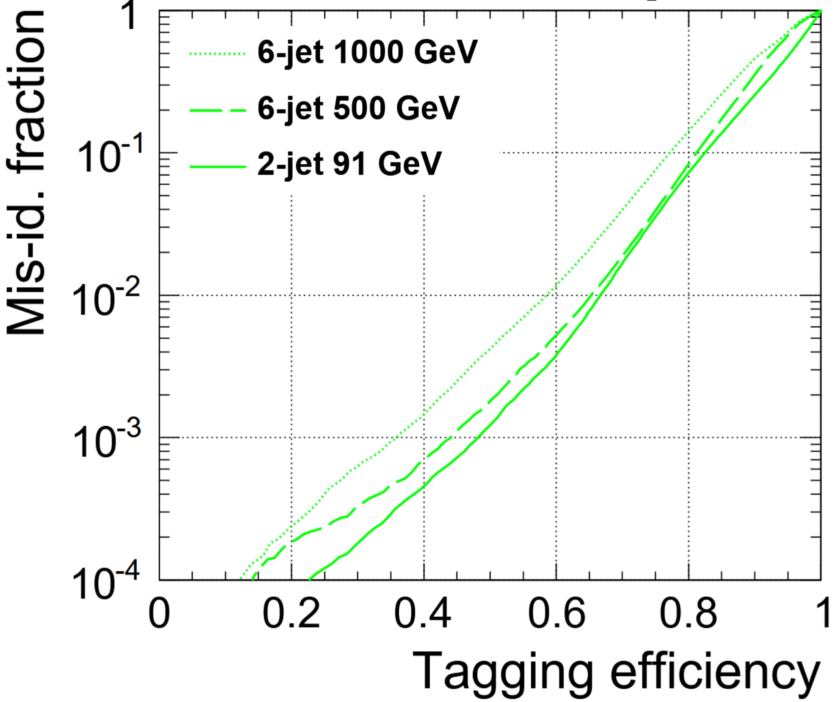
I Flavor Tagging - ParticleNet



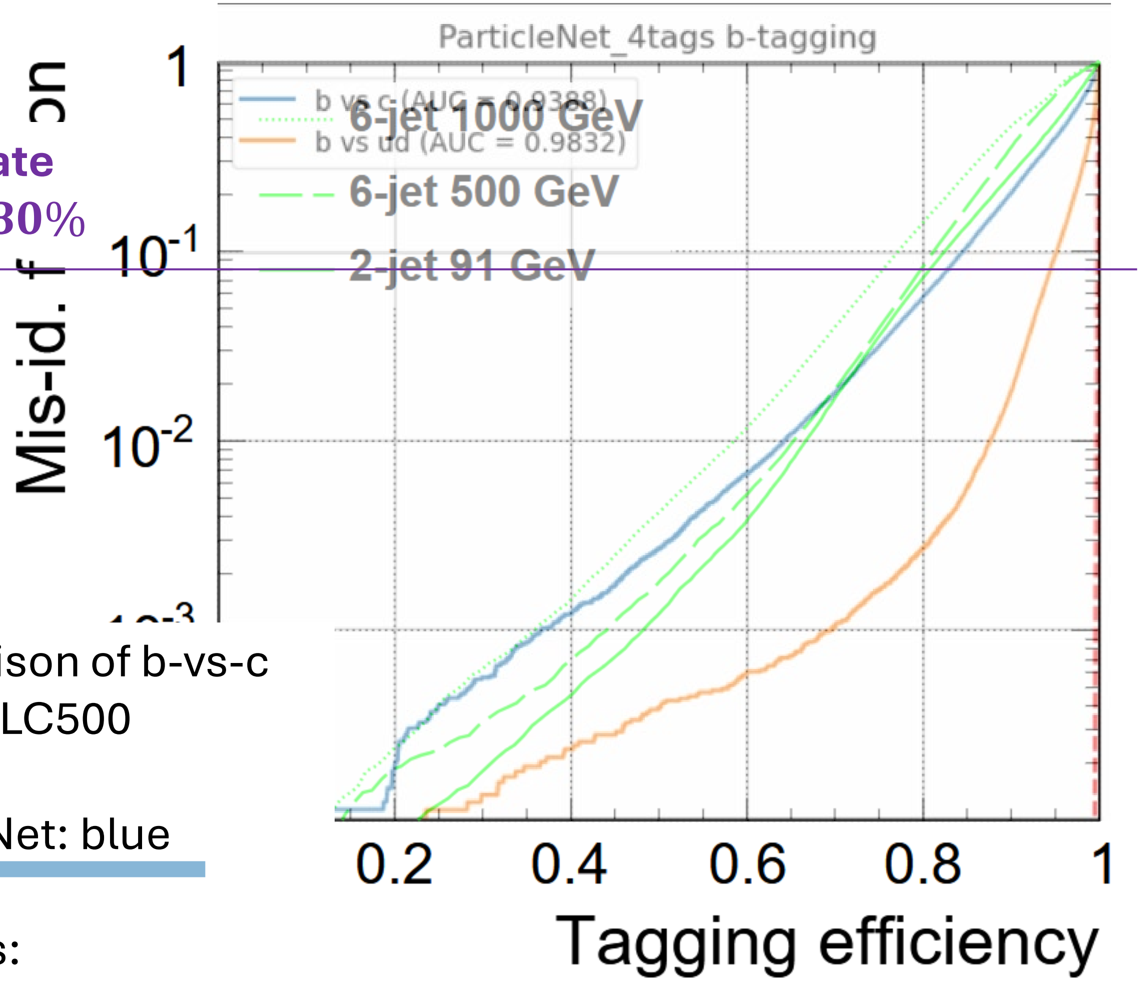
I Flavor Tagging – Question about LCFIPlus

- Reported **8% mis-id.** for b-vs-c in 6j 500 GeV dataset @ $\boldsymbol{\epsilon_b = 80\%}$

[LCFIPlus, 2015]



8% FP-rate @ $\epsilon_b = 80\%$



Comparison of b-vs-c at ILD@ILC500

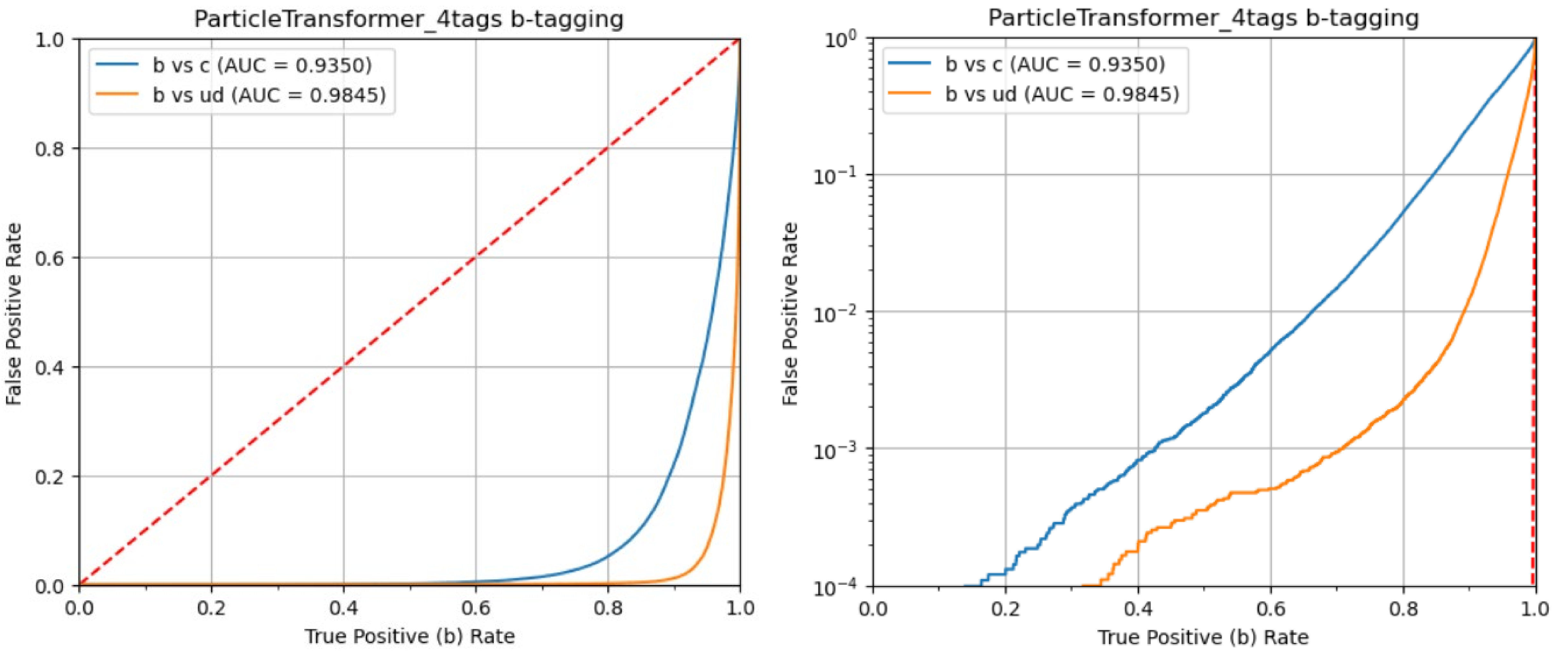
ParticleNet: blue

LCFIPlus:

(a) b tag with c background

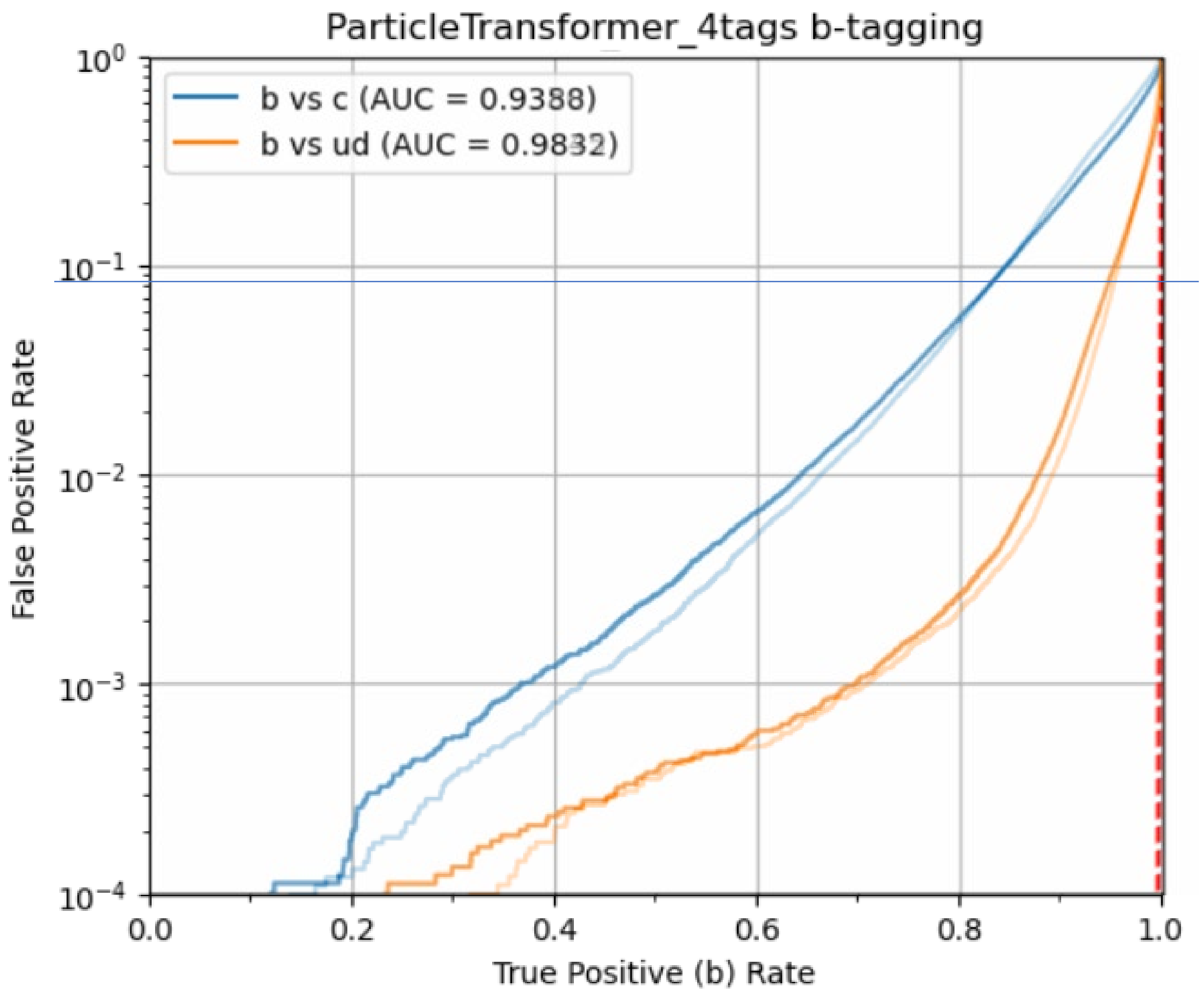
Does it make sense that the current version of LCFIPlus gives worse results on the same (?) dataset?

I.C Flavor Tagging – ParticleTransformer



I.D Flavor Tagging – ParT vs ParticleNet

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG



PNet / b-vs-c AUC: 0.939

ParT / b-vs-c AUC: 0.935

ParT performance not optimal, likely due to limited statistics (10M jets)

I Flavor Tagging – Updates

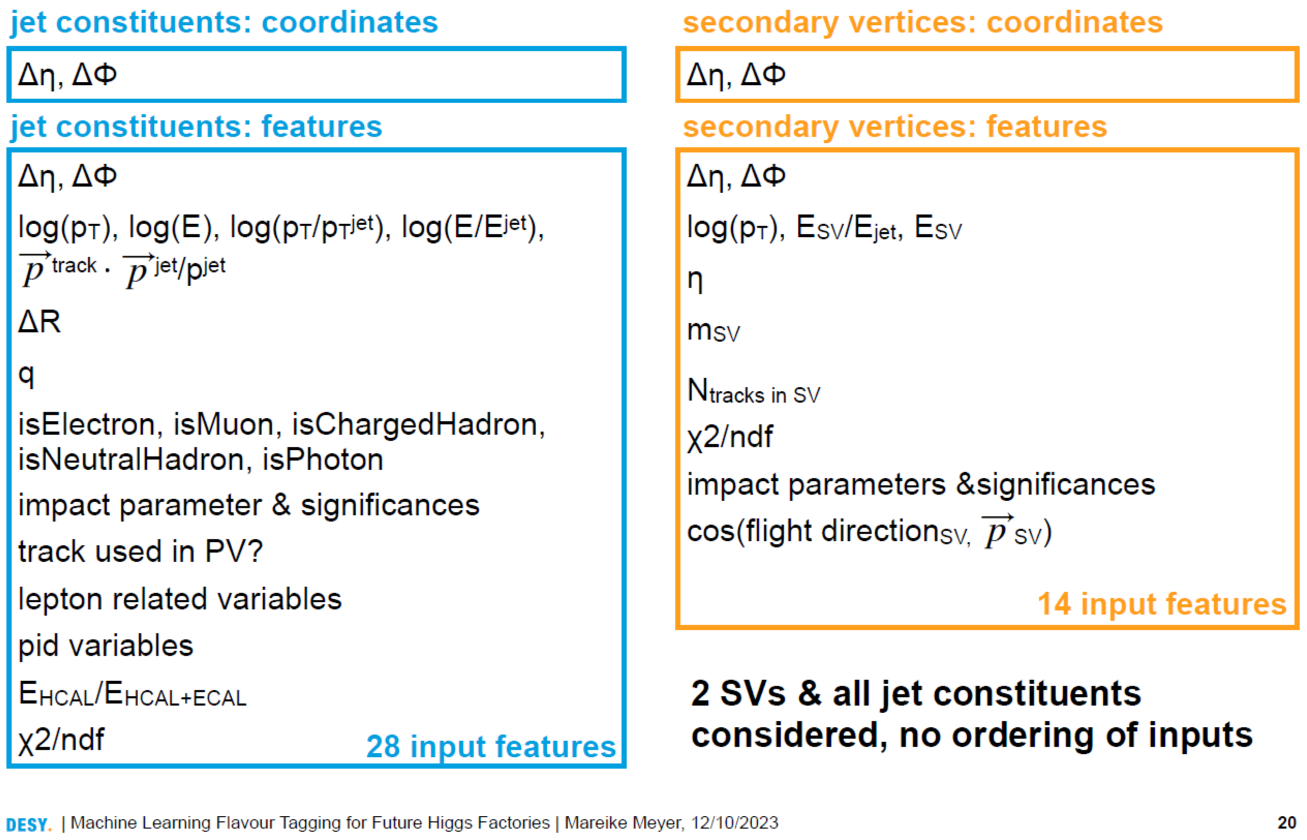
- Extended existing FlavorTag framework
  - Using PyTorch data loaders and transformations:
    - Customizable transformations of input features and labels
    - Configurable sampling and loss reweighting; automatic oversampling of the minority class, etc.
    - Can now support more architectures: ParticleTransformer added (PELICAN in the future?)
  - Configuration of inputs more in-line with weaver tool (a standard ftag tool in HEP)
  - Overhauled data conversion from ROOT to HDF5
    - Much faster deployment due to multitprocessing
  - Hands-on Jupyter notebook examples, documentation
  - Added ComprehensivePID as input (combines dEdx, TOF, cluster shapes; by Uli Einhaus) and s-tag [however, following plots are still without CPID]
- Supported by Uli Einhaus and Thomas Madlener

I Flavor Tagging Updates

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

DESY.

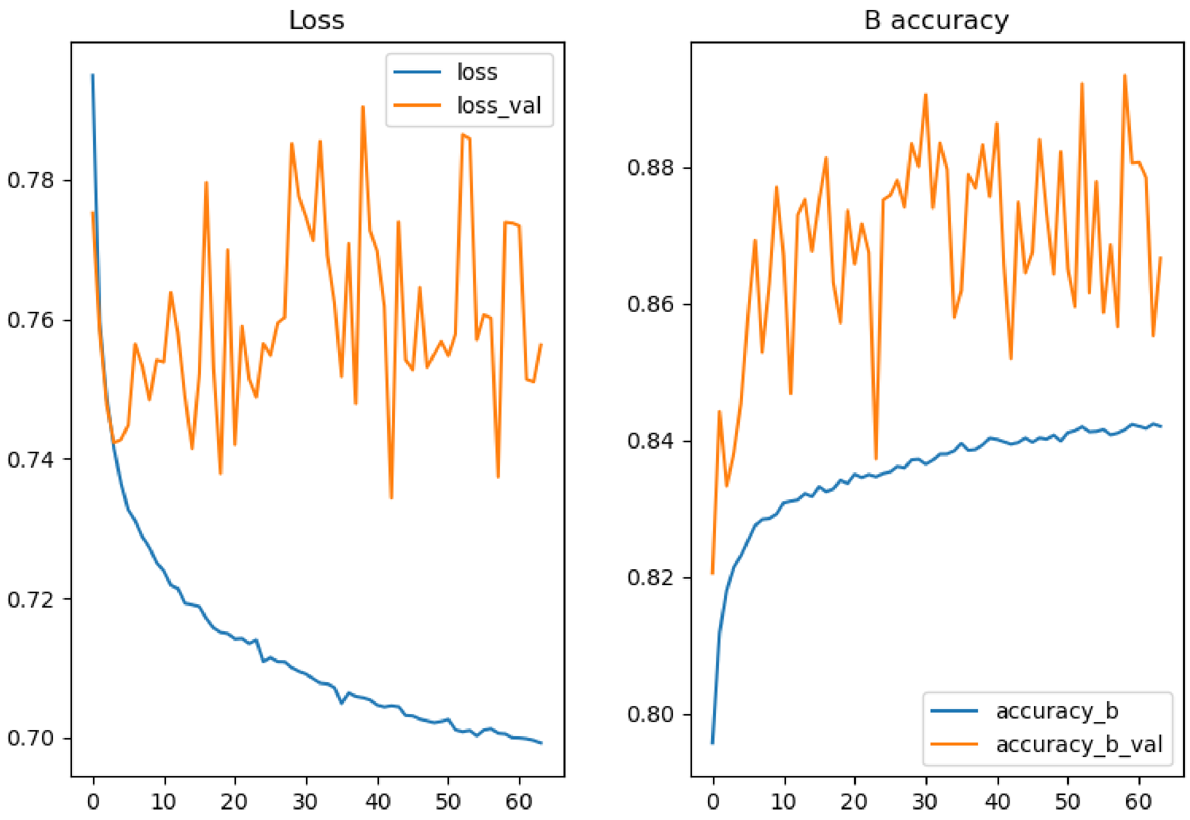
- Starting point (M. Meyer):
  - ParticleNet implemented
  - Using ILD@ILC500 DBD **full-simulation** flavortag datasets (6-jets of same flavor)
  - 2M jets per flavor, 10M total; much less than what PartT is commonly trained on (e.g. JetClass dataset, 100 M in total)



Implemented input features for jet flavor tagging / M. Meyer

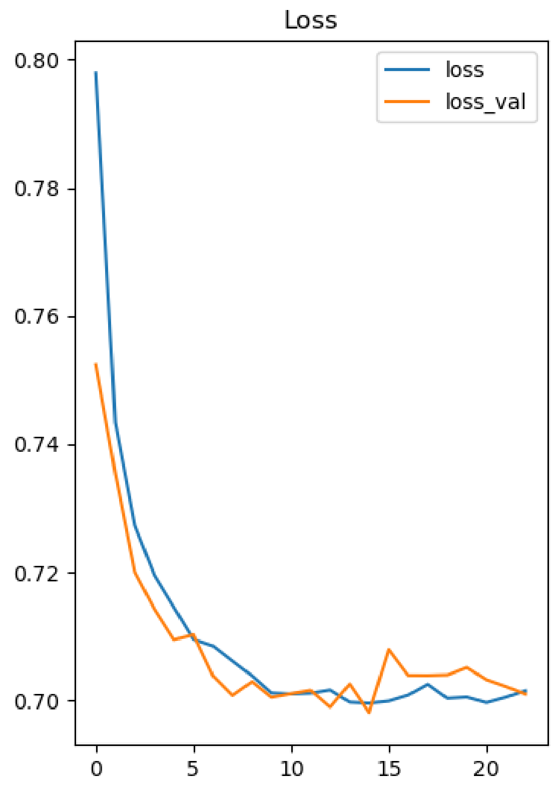
I Flavor Tagging – ParticleNet

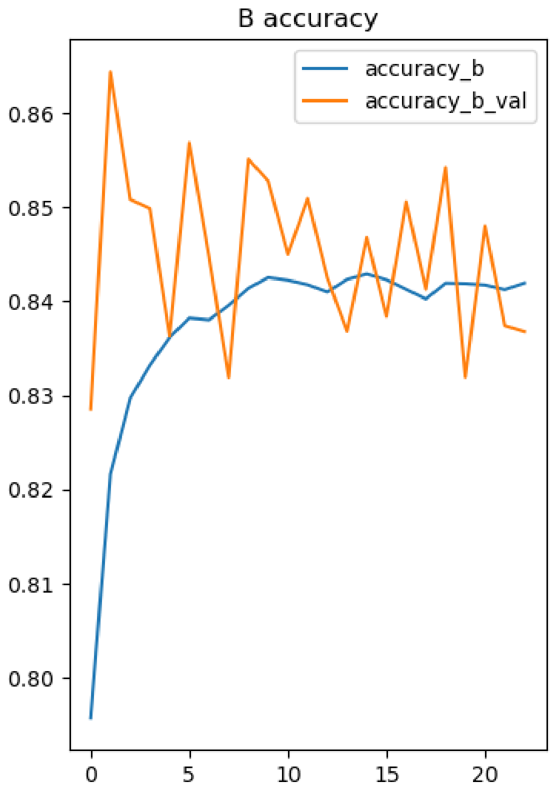
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG



I Flavor Tagging – ParticleTransformer

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG





I.E Flavor Tagging @ ZHH

- flavor tagging in the event selection of the **last analysis**:
  - often: $bmaxN$: N-th highest b-tag

FlavorTag Cut	$llHH(llbbbb)$	$\nu\nu HH(\nu\nu bbbb)$	$qqHH(qqbbbb)$
Preselection	-	$bmax3 \rightarrow o(\epsilon_b\ )$	all of four b-tags $\rightarrow o(\epsilon_b^4)$
Eff. @ Bkg supp.	-	93.9% @ 5.4 [here]	72.8% @ 11.2 [here]

	Background	$\nu\nu HH\ (\nu\nu bbbb)$
σ [fb]	838.9	0.04
expected events	$1.68 \cdot 10^6$	80.14 (28.5)
$N_{isolep} = 0$	$1.2 \cdot 10^6 \pm 586$	62.4 ± 0.1 (27.9)
$\lvert M_{ij} - M_H \rvert < 80$	$7.6 \cdot 10^5 \pm 416$	61.0 ± 0.1 (26.1)
$bmax3 > 0.2$	$1.4 \cdot 10^5 \pm 165$	28.2 ± 0.1 (24.5)

	Background	$qqHH\ (qqbbbb)$
σ [fb]	704.8	0.137
expected eve	$1.4 \cdot 10^6$	273.1 (99)
$N_{isolep} = 0$	$6.7 \cdot 10^5 \pm 440$	214 ± 0.3 (82.0)
$btag > 0.16$	$6.0 \cdot 10^4 \pm 101$	81.7 ± 0.2 (59.7)

Final selection	$bmax3$ (last cut, 100% eff.)	$bmax3\ +\ bmax4$	$bmax3\ +\ bmax4,\qquad bmax3$

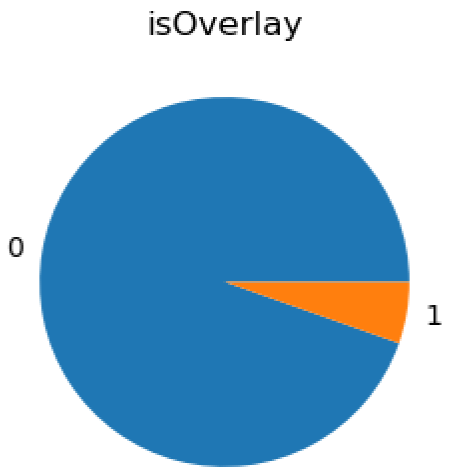
  - Assuming we select 4 independent b-jets, does the increase of average tagging efficiency propagate via $o\left(\epsilon_b^4\right)$ in all cuts on b-tags

I.F Flavor Tagging – Outlook

- Analyze physics performance on signal/background datasets
  - Can we gain the **anticipated 10% improvement per jet also on the signal/bkg sample**
    - To be evaluated until start of next week
  - Additional option:
    - ✓ Fast simulation using SGV on 500 GeV flavortag, ZHH and ZZH datasets for comparison on "common ground" (flavortag sample from older MC production)
  - Target: for ZHH analysis, maximum efficiency (4 b jets) at approx. same background rejection is desired

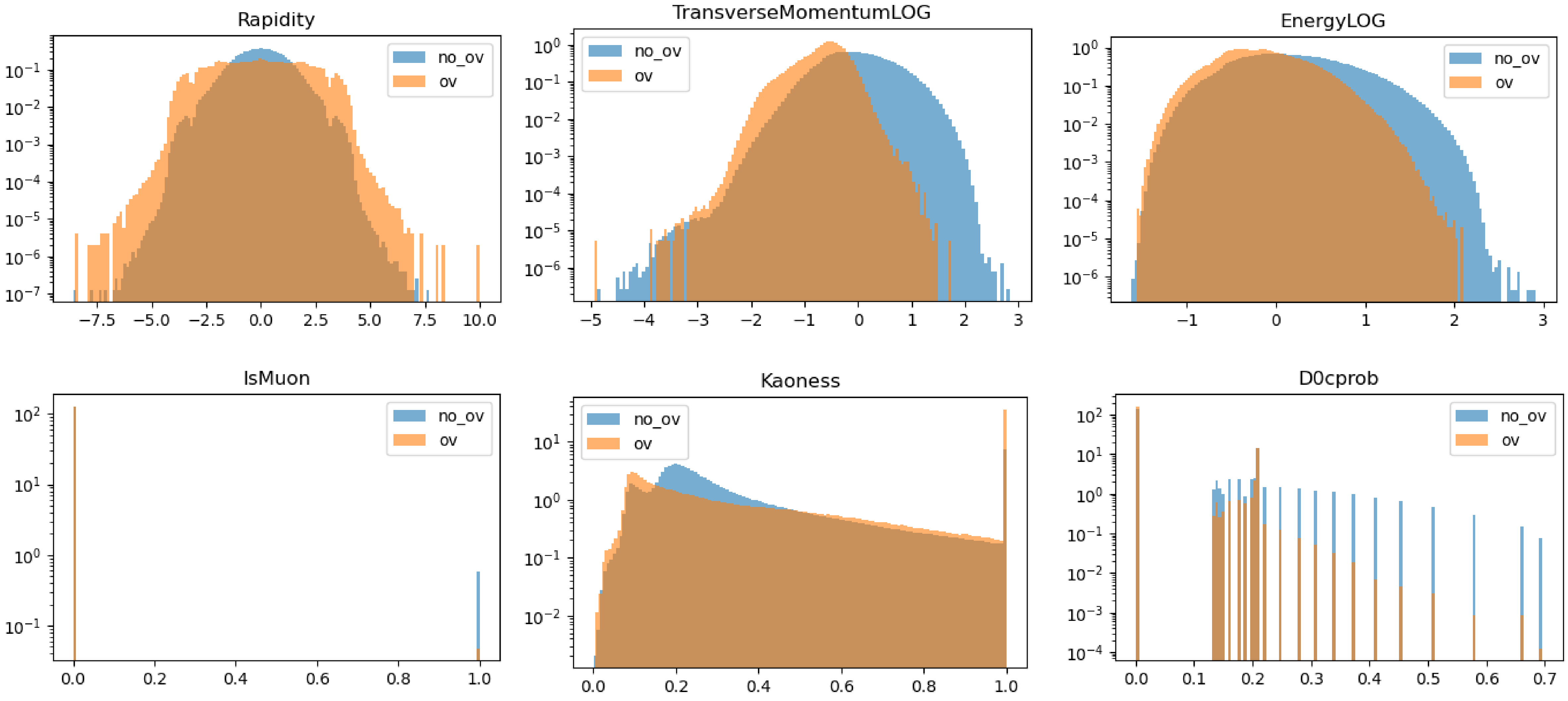
II Overlay Removal

- Idea: use same input data as flavor tagging dataset
- At the moment: some jet-related input features removed
  - Relative phi, rapidity etc.
- Added a label `PFOisOverlay`
- Tested some toy models
- Sent framework and data to Dimitris for cross-check / more ideas

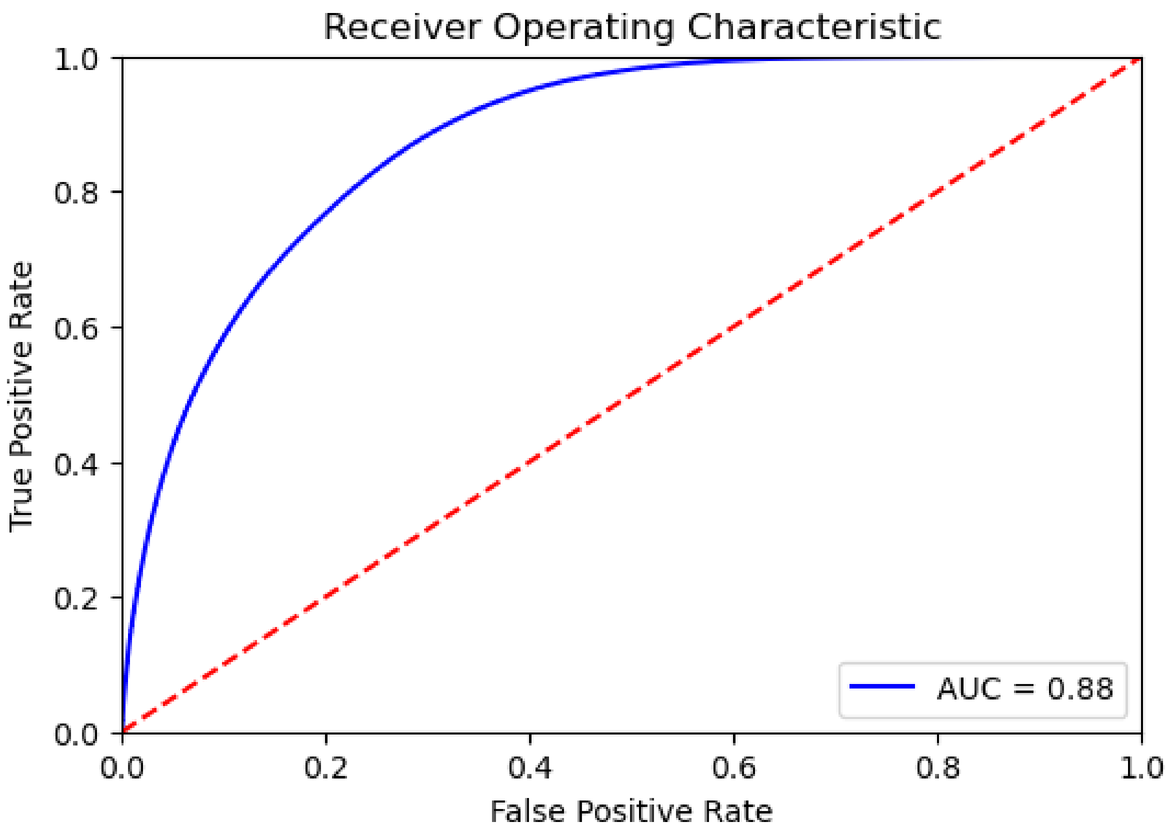


isOverlay:
0: 94.63%
1: 5.37%

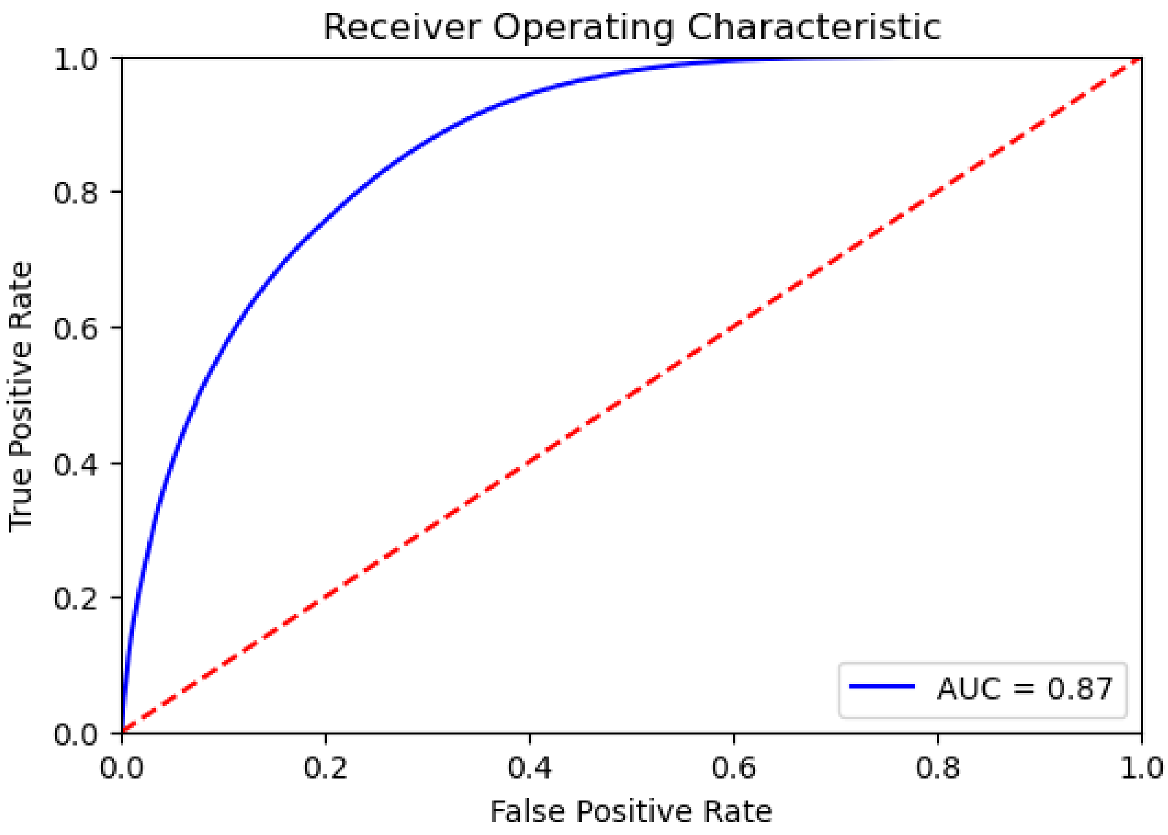
II Overlay Removal – Example features (standardized)



II Overlay Removal



MLP@ PyTorch



GradientBoostingClassifier @ sklearn
(very similar performance just with **logistic regression**)

II Overlay Removal

- Next steps:
  - Compare to “traditional way” of overlay removal (jet clustering to beam jets)
    - ROC curve
    - Physical distributions, influence on physics analysis
- Caveats / open questions:
  - What about IRC safety, esp. when using features that are based on jet information?

Event Selection / llHH(1)

	$eebb$	$\mu\mu bb$	$e\nu bbqq$	$\mu\nu bbqq$	$\tau\nu bbqq$	$bbqqqq$	$bbbb$	$llbbbb$	$llqqH$	**Background**	$llHH$ ($llbbbb$)
σ [fb]	142.1	24.8	124.2	123.0	123.0	312.0	20.1	0.03	0.08	869.1	0.02
expected events	$2.84 \cdot 10^5$	$4.95 \cdot 10^4$	$2.48 \cdot 10^5$	$2.46 \cdot 10^5$	$2.46 \cdot 10^5$	$6.24 \cdot 10^5$	$4.02 \cdot 10^4$	69.51	150.87	$1.73 \cdot 10^6$	40.51 (14.3)
Preselection Cuts											
N_{isolep} >= 2	$6.4 \cdot 10^4 \pm 78$	$2.1 \cdot 10^4 \pm 37$	1911 ± 22	226 ± 7	195 ± 5.9	25.5 ± 2.0	2.4 ± 0.3	21.8 ± 0.1	135 ± 0.5	$8.8 \cdot 10^4 \pm 89$	25.2 ± 0.07 (7.9)
$\lvert M_{ll} - M(Z)\rvert < 40$ GeV	$2.6 \cdot 10^4 \pm 50$	$1.6 \cdot 10^4 \pm 32$	558 ± 12	77 ± 4	33 ± 2.4	4.5 ± 0.9	0.4 ± 0.1	15.3 ± 0.1	132 ± 0.5	$4.3 \cdot 10^4 \pm 63$	24.0 ± 0.07 (7.9)
$\lvert M_{ij} - M_H\rvert < 80$ GeV	2183 ± 14	901 ± 8	544 ± 12	73 ± 4	29 ± 2.3	4.2 ± 0.8	0.4 ± 0.1	12.5 ± 0.1	130 ± 0.5	3877 ± 21	22.5 ± 0.06 (7.9)
Additional Precuts											
60 GeV < M_{H1} < 180 GeV	1530 ± 12	632 ± 6	529 ± 11	66 ± 4	27 ± 2.2	3.8 ± 0.8	0.3 ± 0.1	12.2 ± 0.1	127 ± 0.5	2928 ± 18	22.5 ± 0.06 (7.8)
60 GeV < M_{H2} < 180 GeV	956 ± 10	398 ± 5	481 ± 11	59 ± 4	25 ± 2.1	3.7 ± 0.8	0.3 ± 0.1	11.6 ± 0.1	123 ± 0.4	2057 ± 16	22.4 ± 0.06 (7.7)
missing p_T < 70 GeV	948 ± 10	397 ± 5	343 ± 9	46 ± 3	12 ± 1.5	3.7 ± 0.8	0.3 ± 0.1	11.6 ± 0.1	119 ± 0.4	1880 ± 15	21.4 ± 0.06 (7.7)
thrust <0.9	603 ± 8	288 ± 4	341 ± 9	46 ± 3	12 ± 1.5	3.7 ± 0.8	0.2 ± 0.1	11.5 ± 0.1	119 ± 0.4	1424 ± 13	21.4 ± 0.06 (7.7)

Table 8.1: *Preselection results for the lepton channel, corresponding to a beam polarisation of* $P(e^+e^-) = (0.3, -0.8)$ *and an integrated luminosity of* $\mathcal{L} = 2$ ab^{-1}. *If not stated otherwise,* $l = e, \mu, \tau$. *Also listed are the MC statistical errors on the number of events after every cut. The isolated lepton selection is not optimised for* τ *events. Only isolated* e *and* μ *pairs are selected. Hence, the isolated lepton selection reduces the signal events by one third. From originally 13.5* $eeHH$ *events and the same amount of* $\mu\mu HH$ *events, 12.5 events of each category survive the selection, while 0.5 events correspond to* $\tau\tau HH$. *Optimising the lepton selection strategy also for* τ *could be useful to include* $\tau\tau HH$ *signal events in this study. This can improve the precision of the* ZHH *cross-section measurement by relative 8% if similar results for* $\tau\tau HH$ *are achieved as in* $eeHH$ *and* $\mu\mu HH$.

From DESY-THESIS-2016-027

Event Selection / llHH(2)

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

DESY.

	$ee bb$	$\mu\mu bb$	$e\nu bbqq$	$\mu\nu bbqq$	$\tau\nu bbqq$	$bbqqqq$	$bbbb$	$llbbbb$	$llqqH$	**Background**	$llHH$ ($llbbbb$)
σ [fb]	142.1	24.8	124.2	123.0	123.0	312.0	20.1	0.03	0.08	869.1	0.02
expected events	$2.84 \cdot 10^5$	$4.95 \cdot 10^4$	$2.48 \cdot 10^5$	$2.46 \cdot 10^5$	$2.46 \cdot 10^5$	$6.24 \cdot 10^5$	$4.02 \cdot 10^4$	69.51	150.87	$1.734 \cdot 10^6$	40.51 (14.3)
total preselection	603 ± 7.5	287 ± 4.3	341 ± 9.1	45 ± 3.3	11.8 ± 1.5	3.7 ± 0.8	0.2 ± 0.09	11.5 ± 0.10	119.1 ± 0.44	1424 ± 13.1	21.4 ± 0.06 (7.7)
Electron-type Selection											
$ltype = 11$	603 ± 7.5	0	341 ± 9.1	0.8 ± 0.4	10.9 ± 1.4	3.7 ± 0.8	0.2 ± 0.1	5.6 ± 0.07	57.9 ± 0.3	1024 ± 11.9	10.3 ± 0.04 (3.7)
$bdtg(llbb) > 0.87$	5.8 ± 0.7	0	51 ± 3.5	0.8 ± 0.4	1.9 ± 0.6	1.3 ± 0.5	0.2 ± 0.1	4.6 ± 0.06	16.1 ± 0.2	82 ± 3.7	5.8 ± 0.03 (3.6)
$bdtg(l\nu bbqq) > 0.97$	5.0 ± 0.7	0	2 ± 0.8	0	0.1 ± 0.2	0.4 ± 0.2	0.2 ± 0.1	3.9 ± 0.06	14.6 ± 0.2	27 ± 1.1	5.3 ± 0.03 (3.3)
$bdtg(llbbbb) > -0.41$	1.6 ± 0.4	0	1 ± 0.5	0	0	0	0	0.5 ± 0.02	4.2 ± 0.1	7.9 ± 0.7	4.0 ± 0.03 (2.6)
$bmax3 > 0.03$	1.5 ± 0.4	0	1 ± 0.5	0	0	0	0	0.5 ± 0.02	4.0 ± 0.1	7.0 ± 0.6	3.9 ± 0.03 (2.6)
Muon-type Selection											
$ltype = 13$	0	287.7 ± 4.3	0	44.7 ± 3.2	0.9 ± 0.4	0	0	5.8 ± 0.07	61.2 ± 0.3	400 ± 5.4	11.0 ± 0.04 (3.9)
$bdtg(llbb) > 0.28$	0	10.1 ± 0.8	0	13.5 ± 1.8	0.3 ± 0.2	0	0	5.3 ± 0.07	26.7 ± 0.2	56 ± 1.9	8.0 ± 0.04 (3.9)
$bdtg(l\nu bbqq) > 0.85$	0	9.9 ± 0.8	0	2.7 ± 0.8	0	0	0	5.1 ± 0.07	26.3 ± 0.2	44 ± 1.2	7.7 ± 0.04 (3.8)
$bdtg(llbbbb) > -0.28$	0	2.8 ± 0.4	0	0.1 ± 0.1	0	0	0	0.4 ± 0.02	5.6 ± 0.1	8.9 ± 0.5	5.1 ± 0.03 (2.8)
$bmax3 > 0.01$	0	2.8 ± 0.4	0	0.1 ± 0.1	0	0	0	0.4 ± 0.02	5.6 ± 0.1	8.9 ± 0.5	5.1 ± 0.03 (2.8)

Table 8.2: *Event selection results for the lepton channel, corresponding to a beam polarisation of* $P(e^+e^-) = (0.3, -0.8)$ *and an integrated luminosity of* $\mathcal{L} = 2\ \text{ab}^{-1}$. *If not stated otherwise,* $l = e, \mu, \tau$. *Also listed are the MC statistical errors on the number of events after every cut. The isolated lepton selection is not optimised for* τ *events. Only isolated* e *and* μ *pairs are selected. Hence, the preselection reduces the signal events by one third. The electron-type and muon-type selection is listed. The cut on* $bmax3$ *can be used to select* $HH \to bbbb$ *events. The cut is optional and serves as example in this table, since it has no effect on the results.*

From DESY-THESIS-2016-027

Event Selection / ννHH

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

	$\nu\nu bb$	$e\nu bbqq$	$\mu\nu bbqq$	$\tau\nu bbqq$	$bbqqqq$	$bbbb$	$\nu\nu bbbb$	$\nu\nu qqH$	**Background**	$\nu\nu HH$ ($\nu\nu bbbb$)
σ [fb]	136.4	124.2	123.0	123.0	312.0	20.1	0.05	0.22	838.9	0.04
expected events	$2.73 \cdot 10^5$	$2.49 \cdot 10^5$	$2.46 \cdot 10^5$	$2.46 \cdot 10^5$	$6.24 \cdot 10^5$	$4.02 \cdot 10^4$	96.83	447.0	$1.68 \cdot 10^6$	80.14 (28.5)
Preselection Cuts										
$N_{isolep} = 0$	$2.7 \cdot 10^5 \pm 399$	$2.5 \cdot 10^4 \pm 72$	$2.4 \cdot 10^4 \pm 71$	$2.0 \cdot 10^5 \pm 270$	$6.1 \cdot 10^5 \pm 314$	$4.0 \cdot 10^4 \pm 53$	95 ± 0.4	392 ± 1.1	$1.2 \cdot 10^6 \pm 586$	62.4 ± 0.1 (27.9)
$\lvert M_{ij} - M_H \rvert < 80$ GeV	$2.7 \cdot 10^4 \pm 126$	$1.9 \cdot 10^4 \pm 68$	$1.5 \cdot 10^4 \pm 58$	$2.0 \cdot 10^5 \pm 269$	$4.6 \cdot 10^5 \pm 273$	$3.7 \cdot 10^4 \pm 52$	93 ± 0.4	309 ± 1.0	$7.6 \cdot 10^5 \pm 416$	61.0 ± 0.1 (26.1)
$bmax3 > 0.2$	2290 ± 37	1807 ± 21	1423 ± 18	$3.6 \cdot 10^4 \pm 113$	$6.2 \cdot 10^4 \pm 100$	$3.1 \cdot 10^4 \pm 47$	82 ± 0.4	85 ± 0.5	$1.4 \cdot 10^5 \pm 165$	28.2 ± 0.1 (24.5)
Additional Precuts										
$60 < M_{H1} < 180$	1280 ± 27	1668 ± 20	1341 ± 18	$3.4 \cdot 10^4 \pm 110$	$3.8 \cdot 10^4 \pm 78$	$2.9 \cdot 10^4 \pm 46$	80 ± 0.4	84 ± 0.5	$1.1 \cdot 10^5 \pm 147$	27.8 ± 0.1 (24.3)
$60 < M_{H2} < 180$	634 ± 19	1619 ± 20	1299 ± 17	$3.3 \cdot 10^4 \pm 109$	$3.2 \cdot 10^4 \pm 72$	$2.8 \cdot 10^4 \pm 45$	76 ± 0.4	82 ± 0.5	$9.7 \cdot 10^4 \pm 142$	27.3 ± 0.1 (24.1)
$10 < mp_T < 180$	610 ± 19	1587 ± 20	1271 ± 17	$3.2 \cdot 10^4 \pm 107$	$1.7 \cdot 10^4 \pm 53$	$1.4 \cdot 10^4 \pm 32$	74 ± 0.4	81 ± 0.5	$6.7 \cdot 10^4 \pm 128$	27.0 ± 0.1 (23.9)
thrust <0.9	446 ± 16	1572 ± 20	1254 ± 17	$3.2 \cdot 10^4 \pm 107$	$1.7 \cdot 10^4 \pm 53$	3404 ± 16	73 ± 0.4	80 ± 0.5	$5.6 \cdot 10^4 \pm 124$	26.8 ± 0.1 (23.7)
evis< 400 GeV	444 ± 16	1115 ± 16	1016 ± 15	$2.6 \cdot 10^4 \pm 96$	1841 ± 17	1783 ± 11	72 ± 0.4	80 ± 0.5	$3.2 \cdot 10^4 \pm 102$	26.6 ± 0.1 (23.6)
$M(HH) > 220$ GeV	161 ± 10	1073 ± 16	979 ± 15	$2.5 \cdot 10^4 \pm 94$	1799 ± 17	1656 ± 11	56 ± 0.3	75 ± 0.5	$3.0 \cdot 10^4 \pm 99$	25.7 ± 0.1 (21.5)
Final Selection Cuts										
$bdtg(bbbb) > 0.94$	77 ± 7	621 ± 13	569 ± 11	$1.3 \cdot 10^4 \pm 69$	84 ± 4	17 ± 0.9	23 ± 0.2	49 ± 0.4	$1.5 \cdot 10^4 \pm 71$	18.8 ± 0.08 (17.1)
$bdtg(lvbbqq) > 0.67$	18 ± 3	40 ± 3	62 ± 4	716 ± 16	28 ± 2	7 ± 0.7	10 ± 0.1	22 ± 0.3	902 ± 17	13.3 ± 0.07 (11.9)
$bdtg(\nu\nu bbbb) > 0.3$	10 ± 2	25 ± 3	36 ± 3	414 ± 12	27 ± 2	5 ± 0.6	1 ± 0.05	6 ± 0.1	525 ± 13	10.5 ± 0.06 (9.5)
$bmax3 + bmax4 > 1.08$	0	0	0	2.5 ± 0.9	0	2 ± 0.4	0.5 ± 0.03	2 ± 0.1	7 ± 1	5.6 ± 0.04 (5.5)

Table 8.3: *Selection table for the neutrino channel. The numbers correspond to an integrated luminosity of* $\mathcal{L} = 2\ \text{ab}^{-1}$ *and a beam polarisation of* $P(e^+e^-) = (0.3, -0.8)$. *The MC statistical error on the numbers of events after each selection cut are listed. After final selection* 7 ± 1 *background events and* 5.6 ± 0.04 *signal events remain, with 5.5 events of* $ZHH \to \nu\nu bbbb$. *The btag requirement is important to reject the very large* $\tau\nu bbqq$ *background at the end of the final selection.*

From DESY-THESIS-2016-027

Event Selection / qqHH(1)

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

DESY.

	$qqqqH$	$qqbbbb$	$bbbb$	$bbcssc$	$bbcsdu$	$bbuddu$	$lvbbqq$	ttZ	$ttbb$	**Background**	$qqHH$ ($qqbbbb$)
σ [fb]	0.33	0.07	20.12	77.95	156	78.07	370	1.09	1.05	704.8	0.137
expected events	662.6	140.5	$4.02 \cdot 10^4$	$1.56 \cdot 10^5$	$3.12 \cdot 10^5$	$1.56 \cdot 10^5$	$7.40 \cdot 10^5$	2197	2109	$1.4 \cdot 10^6$	273.1 (99)
Preselection Cuts											
$N_{isolep} = 0$	583 ± 1.4	137 ± 0.4	$3.9 \cdot 10^4 \pm 53$	$1.5 \cdot 10^5 \pm 174$	$3.0 \cdot 10^5 \pm 270$	$1.5 \cdot 10^5 \pm 198$	$2.4 \cdot 10^4 \pm 222$	1322 ± 6	1271 ± 6	$6.7 \cdot 10^5 \pm 440$	214 ± 0.3 (82.0)
$btag > 0.16$	114 ± 0.6	84 ± 0.4	$2.4 \cdot 10^4 \pm 41$	$1.3 \cdot 10^4 \pm 51$	6167 ± 38	568 ± 12	$1.6 \cdot 10^4 \pm 57$	166 ± 2	429 ± 33	$6.0 \cdot 10^4 \pm 101$	81.7 ± 0.2 (59.7)
Additional Precuts											
$60 < M_{H1} < 180$	112 ± 0.6	82 ± 0.3	$2.2 \cdot 10^4 \pm 39$	$1.2 \cdot 10^4 \pm 50$	5955 ± 38	549 ± 12	$1.5 \cdot 10^4 \pm 55$	160 ± 2	412 ± 3	$5.6 \cdot 10^4 \pm 93$	80.5 ± 0.2 (59.0)
$60 < M_{H2} < 180$	109 ± 0.6	79 ± 0.3	$2.0 \cdot 10^4 \pm 36$	$1.2 \cdot 10^4 \pm 49$	5765 ± 37	526 ± 12	$1.2 \cdot 10^4 \pm 47$	156 ± 2	380 ± 3	$5.2 \cdot 10^4 \pm 86$	78.9 ± 0.2 (58.4)
missing $p_T < 70$	109 ± 0.6	79 ± 0.3	$1.9 \cdot 10^4 \pm 36$	$1.2 \cdot 10^4 \pm 49$	5752 ± 37	526 ± 12	7596 ± 37	143 ± 2	337 ± 3	$4.6 \cdot 10^4 \pm 82$	77.4 ± 0.2 (58.2)
thrust <0.9	109 ± 0.6	78 ± 0.4	6492 ± 21	$1.2 \cdot 10^4 \pm 49$	5742 ± 37	525 ± 12	7487 ± 37	143 ± 2	337 ± 3	$3.3 \cdot 10^4 \pm 76$	77.3 ± 0.2 (58.2)

Table 8.4: *Preselection results for the hadron channel. The numbers correspond to an integrated luminosity of $\mathcal{L} = 2$ ab^{-1} and a beam polarisation of $P(e^+e^-) = (0.3, -0.8)$. The MC statistical error on the numbers of events after each selection cut are listed. The mass reconstruction is very challenging in the six-jet final state. Flavour-tag information are needed. At least four jets of an event are required to have a btag larger than 0.16. These jets are paired to form the two Higgs bosons. This requirement rejects 15% of the $ZHH \to qqbbbb$ events but imposing a more loose requirement on the btag degrades the mass resolution and increases the background contribution since more backgrounds pass the selection cut. Investigations have shown that this results in up to 20% degradation of the signal significance in this channel. By improving the flavour-tagging efficiency by 5% for the same purity would lead to 20% more signal events after the selection. Conclusively, after b-tag and isolated lepton requirement all other precuts are nearly 100% efficient on signal events and reduce background events by a factor of two.*

From DESY-THESIS-2016-027

Event Selection / qqHH(2)

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

	$qqqqH$	$qqbbbb$	$bbbb$	$bbcssc$	$bbcsdu$	$bbuddu$	$lvbbqq$	ttZ	$ttbb$	**Background**	$qqHH$ ($qqbbbb$)
expected events	662.6	140.5	$4.02 \cdot 10^4$	$1.56 \cdot 10^5$	$3.12 \cdot 10^5$	$1.56 \cdot 10^5$	$7.40 \cdot 10^5$	2197	2109	$1.4 \cdot 10^6$	273.1 (99)
preselection	109 ± 0.6	78 ± 0.4	6492 ± 21	$1.2 \cdot 10^4 \pm 49$	5742 ± 37	525 ± 12	7487 ± 37	143 ± 2	337 ± 3	$3.3 \cdot 10^4 \pm 76$	77.3 ± 0.2 (58.2)
$bbHH$ dominant Selection											
$btagZ > 0.54$	15 ± 0.2	14 ± 0.4	648 ± 6.8	242 ± 7	87 ± 5	11 ± 1.7	166 ± 6	16 ± 0.6	35 ± 0.9	1233 ± 12.7	18.8 ± 0.1 (16.1)
$bdtg(bbbb)>0.9$	11 ± 0.2	7 ± 0.1	25 ± 1.3	177 ± 6	65 ± 4	8 ± 1.5	16 ± 2	13 ± 0.6	25 ± 0.8	348 ± 7.8	15.7 ± 0.1 (13.7)
$bdtg(bbqqqq) > 0.28$	10 ± 0.2	6 ± 0.1	22 ± 1.3	111 ± 5	45 ± 3	6 ± 1.2	12 ± 2	13 ± 0.6	23 ± 0.8	246 ± 6.5	14.6 ± 0.1 (13.0)
$bdtg(qqbbbb) > -0.25$	9 ± 0.2	5 ± 0.1	20 ± 1.2	105 ± 5	41 ± 3	5 ± 1.1	11 ± 2	12 ± 0.6	22 ± 0.7	231 ± 6.4	14.1 ± 0.1 (12.6)
$bmax3 + bmax4 > 1.22$	4 ± 0.1	2 ± 0.1	6 ± 0.6	4 ± 1	0	1 ± 0.3	0	2 ± 0.2	3 ± 0.3	22 ± 1.3	8.5 ± 0.1 (8.0)
Light $qqHH$ dominant Selection											
$btagZ < 0.54$	94 ± 0.6	65 ± 0.3	5845 ± 20.4	$1.2 \cdot 10^4 \pm 48$	5654 ± 37	514 ± 11.6	7321 ± 38.8	128 ± 1.8	302 ± 2.8	$3.1 \cdot 10^4 \pm 75.6$	58.5 ± 0.2 (42.1)
$bdtg(bbbb)>0.9$	71 ± 0.5	36 ± 0.2	208 ± 3.9	8571 ± 42	4219 ± 32	396 ± 10.2	521 ± 10.3	113 ± 1.7	244 ± 2.5	$1.4 \cdot 10^4 \pm 54.8$	48.5 ± 0.2 (36.5)
$bdtg(bbqqqq) > 0.61$	42 ± 0.4	17 ± 0.2	80 ± 2.4	1599 ± 18	909 ± 15	98 ± 5.1	83 ± 4.1	71 ± 1.3	121 ± 1.8	3019 ± 24.4	32.5 ± 0.1 (25.5)
$bdtg(qqbbbb) > 0.17$	28 ± 0.3	8 ± 0.1	60 ± 2.1	1341 ± 16	746 ± 13	81 ± 4.6	70 ± 3.8	63 ± 1.3	102 ± 1.6	2499 ± 21.8	29.8 ± 0.1 (23.3)
$bmax3 + bmax4 > 1.18$	17 ± 0.2	5 ± 0.1	41 ± 1.7	57 ± 3	54 ± 4	16 ± 2.0	3 ± 0.8	27 ± 0.8	42 ± 1.0	261 ± 5.6	19.5 ± 0.1 (16.6)
$bmax3 > 0.85$	15 ± 0.2	5 ± 0.1	38 ± 1.6	36 ± 3	40 ± 3	13 ± 1.8	3 ± 0.7	24 ± 0.8	36 ± 1.0	209 ± 5.2	17.9 ± 0.1 (15.4)
$40 < M_Z < 110$	13 ± 0.2	4 ± 0.1	27 ± 1.4	27 ± 2	30 ± 3	7 ± 1.3	3 ± 0.7	15 ± 0.6	25 ± 0.8	151 ± 4.3	15.7 ± 0.1 (13.8)
$90 < M_{H1} < 140$	8 ± 0.2	2 ± 0.1	16 ± 1.0	13 ± 2	17 ± 2	5 ± 1.1	2 ± 0.4	7 ± 0.5	14 ± 0.6	84 ± 3.2	13.0 ± 0.1 (11.9)
$90 < M_{H2} < 140$	7 ± 0.2	2 ± 0.1	10 ± 0.8	8 ± 1	9 ± 1	4 ± 0.9	2 ± 0.4	5 ± 0.3	8 ± 0.5	55 ± 2.0	12.6 ± 0.1 (10.9)

Table 8.5: *Final selection results for the hadron channel corresponding to an integrated luminosity of $\mathcal{L} = 2\ \text{ab}^{-1}$ and a beam polarisation of $P(e^+e^-) = (0.3, -0.8)$. The "$bbHH$ dominant" and "light $qqHH$ dominant" categories are listed. Investigating one combined $qqHH$ channel degrades the event selection results by up to 20%. In "light $qqHH$ dominant" the baseline event selection strategy is not sufficient to suppress the large background contribution. Therefore, the flavour-tag requirement is followed by imposing tight cuts on preselection variables. The Z mass provides additional separation power. Relative 5% improvement of the b-tag efficiency for the same purity would give a relatively improved signal significance of 20% in both channels.*

From DESY-THESIS-2016-027