Separation of $e^+e^- \rightarrow f\bar{f}$ with event-based ML models

ALEXANDRE BEAUBIEN, PHD STUDENT

UNIVERSITY OF VICTORIA

THE BELLE II COLLABORATION

2023-02-10

Outline

1. Introduction & Motivation
2. Summary Last Time
   1. Selection of $b\bar{b}$, $c\bar{c}$
   2. Eliminating $uds\tau$ with cut based approach
3. New Results and Work in Progress for CDR

Motivation

Identify $\boldsymbol{e^+e^- \to f\bar{f}}$ events for $\boldsymbol{A_{LR}}$ **measurements** using a machine learning based classifier trained on **event shape variables.**

Goal is to try to get **high purity discrimination** of $\boldsymbol{b\bar{b}}$**,** $\boldsymbol{c\bar{c}}$ without reconstruction.

Results:

Sufficient for selecting $\boldsymbol{b\bar{b}}$ events, but some **reconstruction may be needed** to select $\boldsymbol{c\bar{c}}$ events from the $uds\tau$.

$b\bar{b}$, $c\bar{c}$ GBDT model

Evaluation variables:

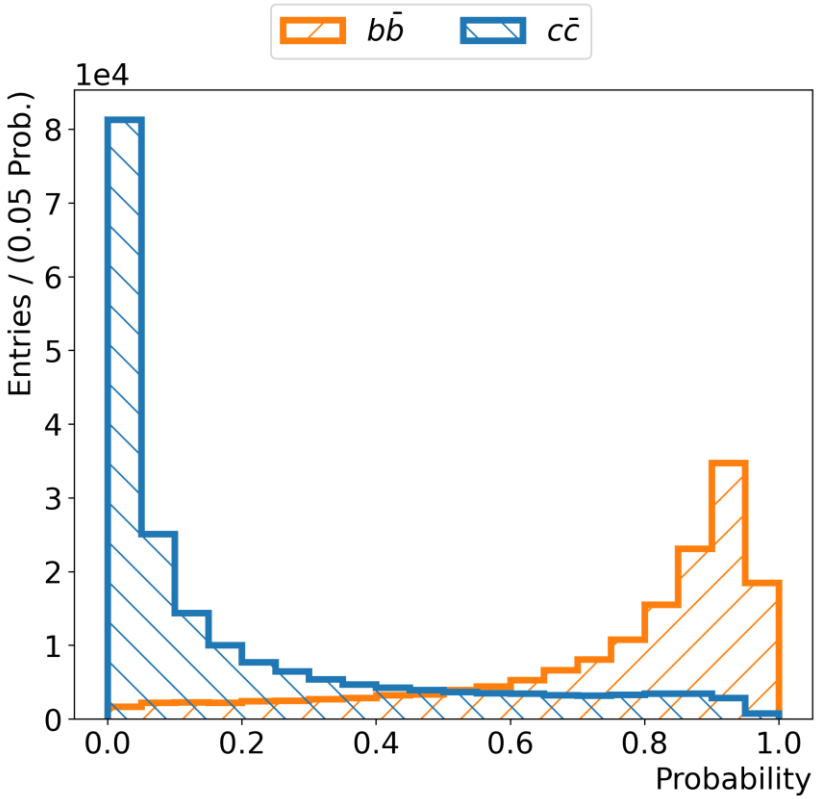
$$b\bar{b}\text{ Efficiency} = \frac{\#\text{ of real } b\bar{b} \text{ predicted as } b\bar{b}}{\#\text{total number of MC generated } b\bar{b}}$$

$$c\bar{c}\text{ Fake rate} = \frac{\#\text{ of real } c\bar{c} \text{ misidentified as } b\bar{b}}{\#\text{total number of MC generated } c\bar{c}}$$

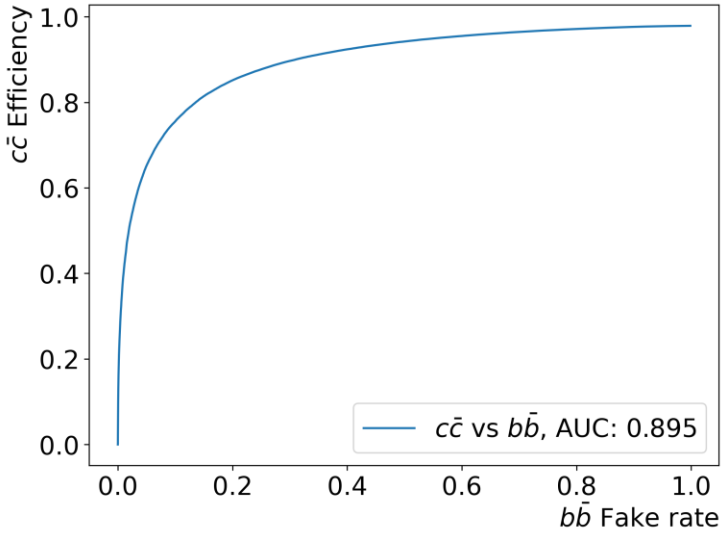
Variable	Feature Importance
foxWolframR2	0.590
thrust	0.184
foxWolframR1	0.081
harmonicMomentThrust0	0.060
thrustAxisCosTheta	0.039
harmonicMomentCollision2	0.020
foxWolframR3	0.010
aplanarity	0.006
harmonicMomentThrust2	0.006
sphericity	0.004

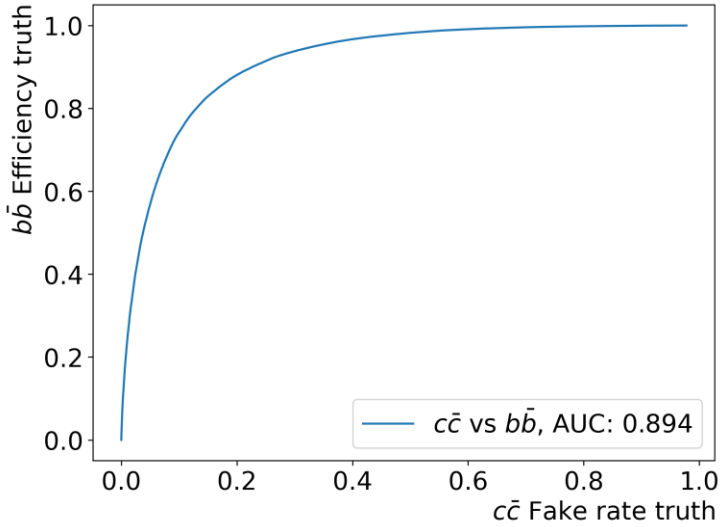
10 most important variables (as defined by the GBDT)

Validation Results



Use **ROC** and **AUC** to quantify model quality. Rates are normalized to # of generated particles (MC number).





Results $b\bar{b}$

$b\bar{b}$ selection against $c\bar{c}$ & $uds\tau$ background (in fraction).

Fraction						
Efficiency $b\bar{b}$	Background Fraction	Fake Rate $c\bar{c}$	$u\bar{u}$	$d\bar{d}$	$s\bar{s}$	$\tau^+\tau^-$
0.098	0.088	0.003	0.002	0.002	0.002	0.002
0.196	0.091	0.008	0.004	0.004	0.004	0.002
0.301	0.110	0.015	0.008	0.008	0.008	0.003
0.401	0.128	0.024	0.013	0.013	0.014	0.003
0.501	0.153	0.038	0.020	0.020	0.022	0.003
0.599	0.181	0.056	0.029	0.030	0.032	0.004
0.699	0.221	0.083	0.044	0.046	0.050	0.004
0.800	0.281	0.130	0.070	0.073	0.079	0.006
0.900	0.380	0.226	0.124	0.131	0.144	0.014
0.998	0.802	0.979	0.966	0.966	0.941	0.996

Lepton selection

Include the requirement that the **event must have a lepton**. Muon selection is not optimal. Does reduce the uds contribution. $c\bar{c}$ and τ not as affected

Fraction						
	MC Truth			ID Cuts		
Type	μ	e	Total	μ	e	Total
$b\bar{b}$	0.26	0.27	0.40	0.33	0.22	0.41
$c\bar{c}$	0.13	0.16	0.28	0.25	0.12	0.34
$u\bar{u}$	0.02	0.06	0.08	0.16	0.05	0.20
$d\bar{d}$	0.02	0.06	0.08	0.16	0.05	0.20
$s\bar{s}$	0.02	0.05	0.07	0.14	0.04	0.17
$\tau^+\tau^-$	0.25	0.27	0.47	0.27	0.21	0.45

(a) Fraction of events containing at least one lepton.

Events per nb^{-1}						
	MC Truth			ID Cuts		
Type	μ	e	Total	μ	e	Total
$b\bar{b}$	0.29	0.30	0.45	0.36	0.24	0.46
$c\bar{c}$	0.17	0.21	0.36	0.32	0.15	0.44
$u\bar{u}$	0.04	0.10	0.13	0.25	0.07	0.31
$d\bar{d}$	0.01	0.02	0.03	0.06	0.02	0.08
$s\bar{s}$	0.01	0.02	0.03	0.05	0.01	0.07
$\tau^+\tau^-$	0.23	0.24	0.44	0.25	0.20	0.41

(b) Number of events containing at least one lepton.

Table 7: Simulation sets containing at least one lepton. The sets are selected using cuts: truth is the number MC generated number of events with leptons. ID cuts uses a cut on the default particle identification tool (>0.95) and on E/p (>0.85 electron).

New Work

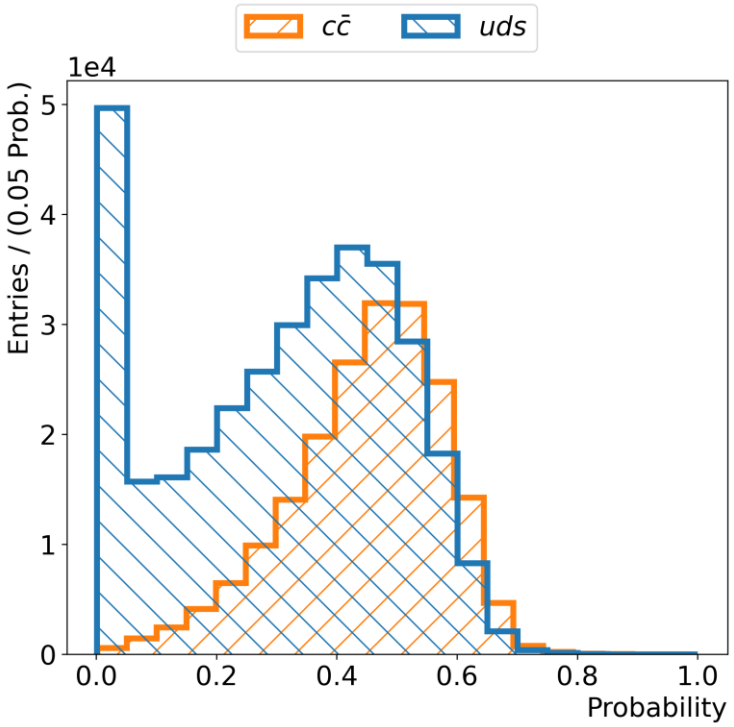
1. Separating $\bar{c}c$ & uds with a new classifier.
2. Try using some particle information to improve classification of
   1. $\bar{c}c$ & uds
   2. $b\bar{b}$ & $\bar{c}c$
3. Preparing to test on data

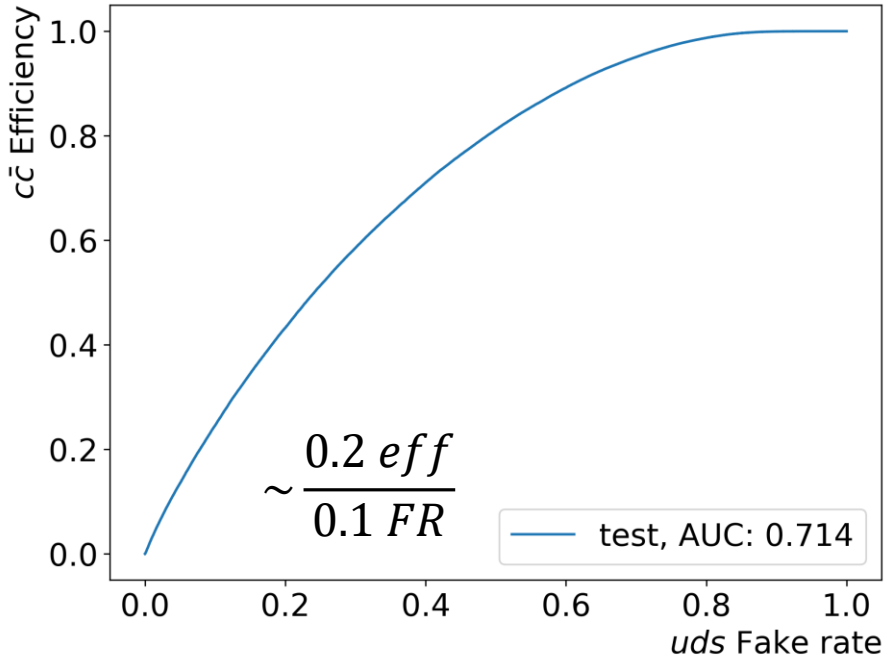
$\bar{c}c$ & uds classifier

Using only event shape variables.

Validation results:

```
0. aplanarity (0.4330101174193695)
1. harmonicMomentThrust0 (0.12335799617715153)
2. foxWolframR1 (0.12150209446432508)
3. foxWolframR3 (0.12007420778184033)
4. sphericity (0.11971919186855932)
5. thrust (0.03248791117385139)
6. foxWolframR2 (0.022631281939098205)
7. thrustAxisCosTheta (0.011183052355619357)
8. harmonicMoment__bo2__cm__spcollision__bc (0.008527742339687639)
9. harmonicMomentThrust2 (0.007506404480497606)
```





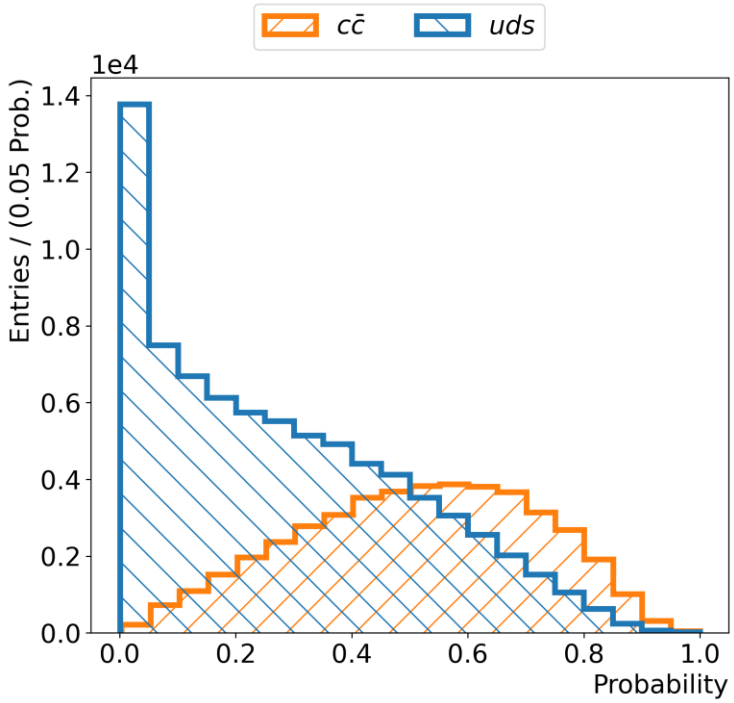
$\bar{c}c$ & uds classifier #2

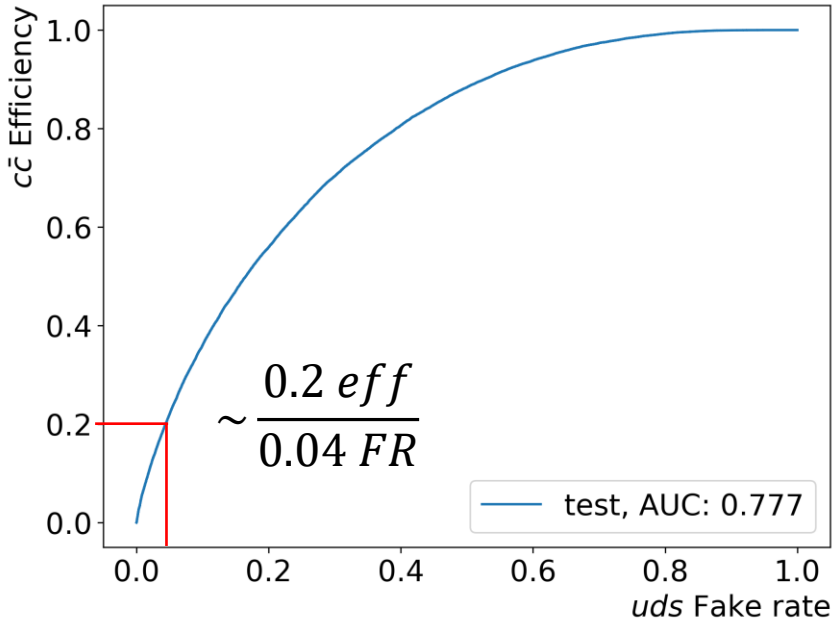
Using event shape, other event variables & particle numbers.

Will eventually add other :
e.g. maxP of e, μ, K in event.

Validation results:

```
0. foxWolframR4 (0.201)
1. nK (0.138)
2. nTracks (0.129)
3. harmonicMomentThrust0 (0.060)
4. ne (0.053)
5. visibleEnergyOfEventCMS (0.040)
6. foxWolframR3 (0.033)
7. totalPhotonsEnergyOfEvent (0.027)
8. nmu (0.022)
9. foxWolframR1 (0.020)
10. missingMass2OfEvent (0.018)
11. backwardHemisphereEnergy (0.014)
12. harmonicMomentThrust4 (0.014)
13. missingMomentumOfEvent_theta (0.014)
14. cleoConeThrust0 (0.013)
15. thrust (0.012)
16. cleoConeThrust7 (0.011)
17. cleoConeThrust2 (0.011)
18. nExtraCDCHitsPostCleaning (0.010)
19. cleoConeThrust6 (0.010)
20. npi (0.009)
```





Creating a Testing Data Set

To **test the classifiers**, run them on data & MC.

Two datasets are selected:

1. **Experiment 14 (4S)**
2. **Experiment 12 (4S_offres)**

Currently **in progress**, still processing the data.

The **equivalent run dependant MC** datasets are also being processed to get accurate **efficiencies/fake rates** of the classifier.

Currently training on exp10 MC

Experiment	Beam Energy	Offline luminosity fb^{-1}	
		proc12 & prompt	proc13 & prompt
14	4S	16.385 +/- 0.005	16.405 +/- 0.005 +/- 0.115
12	4S	54.388 +/- 0.004	54.368 +/- 0.004 +/- 0.381
	4S_offres	8.716 +/- 0.002	8.679 +/- 0.002 +/- 0.061
10	4S	3.635 +/- 0.001	3.647 +/- 0.001 +/- 0.026

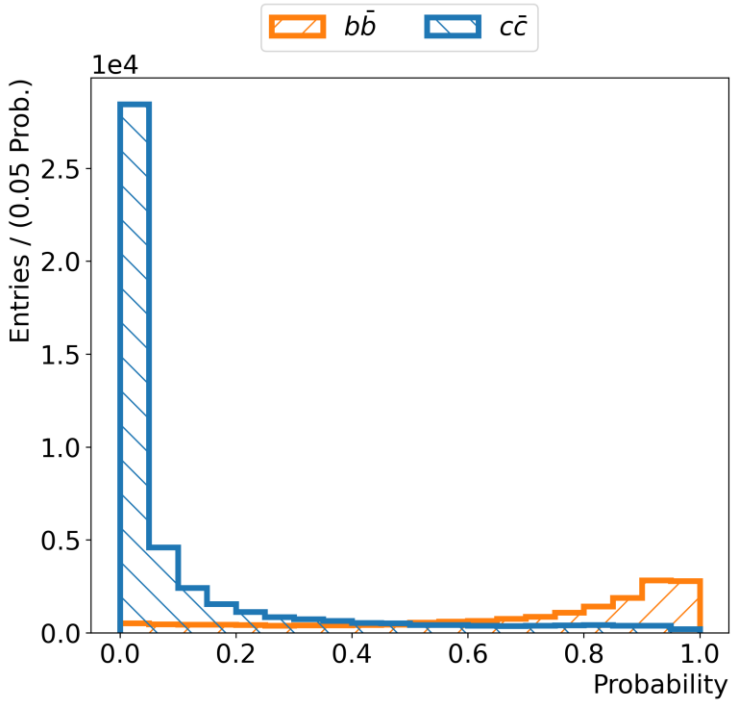
Extra: New $b\bar{b}$ & $c\bar{c}$ classifier

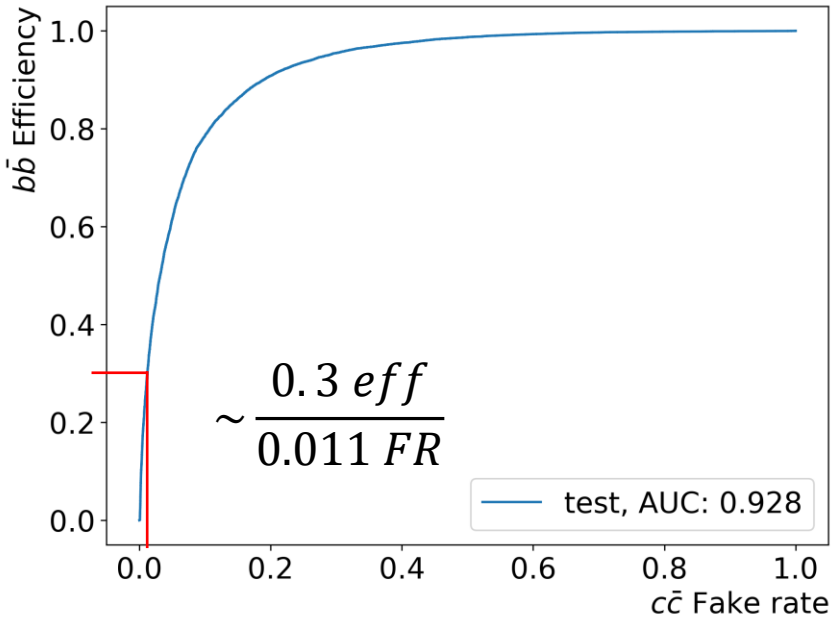
Using event shape, other event variables & particle numbers.

Will eventually add other :
e.g. maxP of e, μ, K in event.

Validation results:

```
0. thrust (0.441)
1. foxWolframR2 (0.199)
2. missingEnergyOfEventCMS (0.054)
3. harmonicMoment__bo2__cm__spcollision__bc (0.045)
4. nTracks (0.036)
5. ne (0.027)
6. harmonicMomentThrust2 (0.023)
7. foxWolframR1 (0.022)
8. nmu (0.015)
9. harmonicMomentThrust0 (0.012)
10. missingMomentumOfEventCMS_theta (0.008)
```





Conclusion

Count the $e^+e^- \to f\bar{f}$ events, especially $b\bar{b}$ & $c\bar{c}$. Avoid reconstruction (e.g. D mesons) using only event based variables.

$b\bar{b} - c\bar{c}$ classifier improved by particle event variables.

$c\bar{c} - uds$ classifier show better results than cut based approach. Inconclusive at the moment, waiting for data testing samples.

Interested in applying the same analysis for $s\bar{s} - uds\tau$ using similar methods. Would provide some additional counting opportunities. May not be possible, have to try.

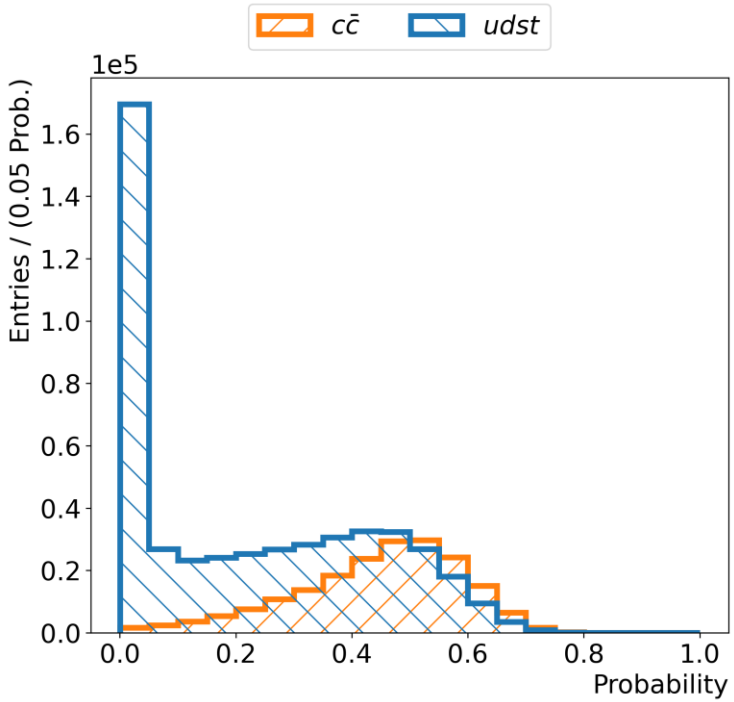
Backup

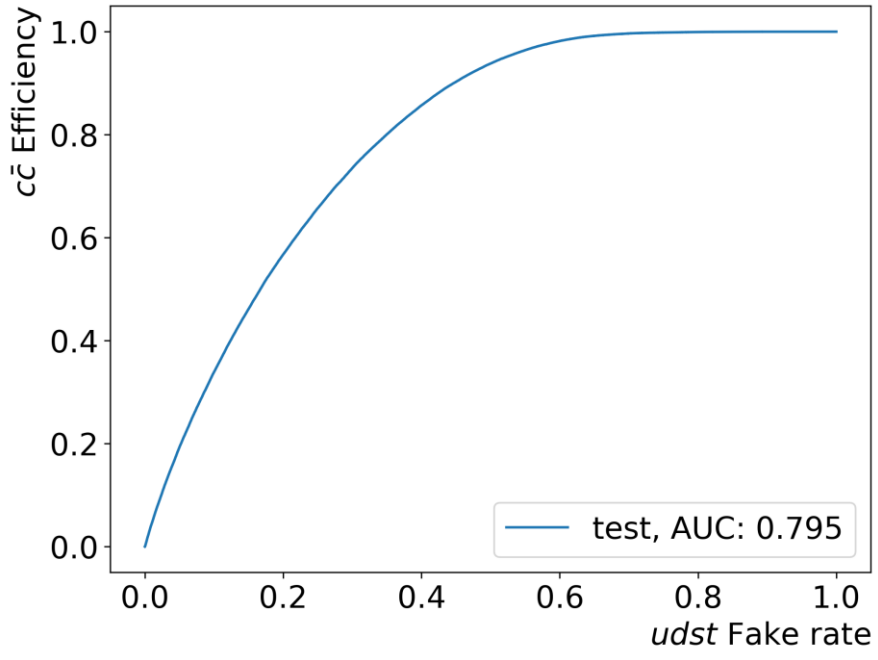
$\bar{c}c$ & $udst$ classifier

Using only event shape variables.

Validation results:

```
0. aplanarity (0.549)
1. foxWolframR3 (0.118)
2. harmonicMomentThrust0 (0.114)
3. sphericity (0.076)
4. foxWolframR1 (0.045)
5. thrust (0.040)
6. foxWolframR2 (0.037)
7. harmonicMoment__bo2__cm__spcollision__bc (0.009)
8. thrustAxisCosTheta (0.006)
9. harmonicMomentThrust2 (0.006)
```



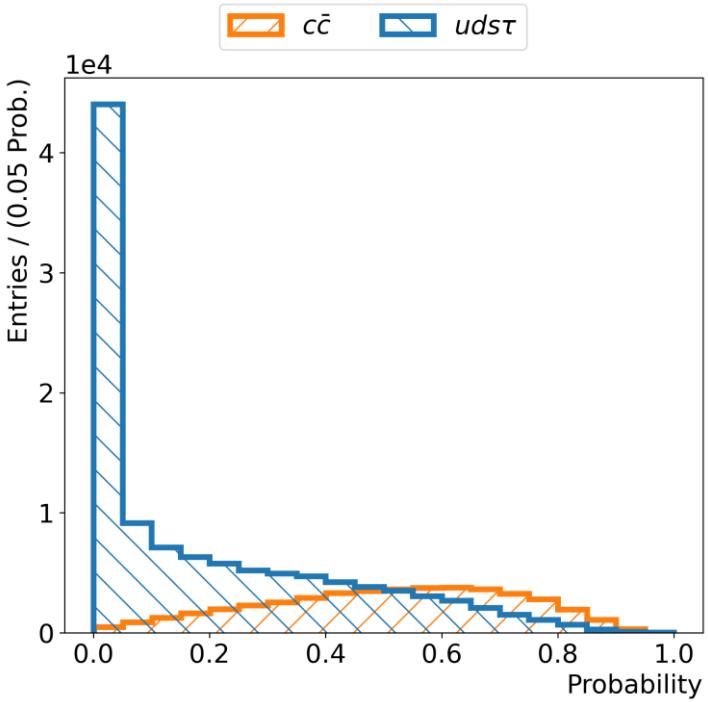


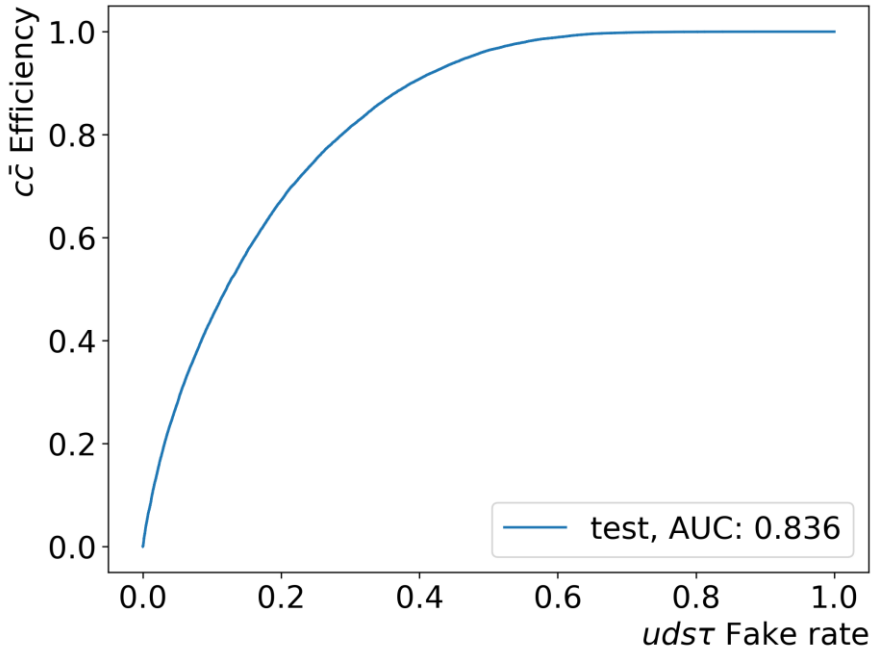
$\bar{c}c$ & $uds\tau$ classifier #2

Using event shape, other event variables & particle numbers.

Validation results:

```
0. nTracks (0.425)
1. foxWolframR4 (0.109)
2. nK (0.097)
3. harmonicMomentThrust0 (0.055)
4. totalPhotonsEnergyOfEvent (0.040)
5. ne (0.029)
6. foxWolframR3 (0.023)
7. aplanarity (0.022)
8. cleoConeThrust0 (0.018)
9. nmu (0.012)
10. thrust (0.011)
```





Results $c\bar{c}$

$c\bar{c}$ selection against $b\bar{b}$

Fraction	
$c\bar{c}$ efficiency	$b\bar{b}$ fake rate
0.102	0.000
0.201	0.002
0.301	0.005
0.402	0.010
0.500	0.020
0.599	0.037
0.700	0.069
0.800	0.138
0.900	0.311
0.979	0.998

(a) Efficiency and fake rates.

Events per nb^{-1}	
$c\bar{c}$	$b\bar{b}$
0.133	0.000
0.261	0.002
0.391	0.006
0.523	0.011
0.650	0.022
0.779	0.041
0.910	0.077
1.040	0.153
1.170	0.345
1.273	1.108

(b) Number of events selected as $c\bar{c}$..

Table 4: Classification of $c\bar{c}$ events against $b\bar{b}$ events. These tables show the results from figure 7a.

Results $b\bar{b}$

$b\bar{b}$ selection against $c\bar{c}$ and $uds\tau$ background in nb^{-1}

Events selected as $b\bar{b}$ per nb^{-1}					
$b\bar{b}$	$c\bar{c}$	$u\bar{u}$	$d\bar{d}$	$s\bar{s}$	$\tau^+\tau^-$
0.109	0.004	0.003	0.001	0.001	0.002
0.218	0.010	0.006	0.002	0.002	0.002
0.334	0.020	0.013	0.003	0.003	0.003
0.445	0.031	0.021	0.005	0.005	0.003
0.556	0.049	0.032	0.008	0.008	0.003
0.665	0.073	0.047	0.012	0.012	0.004
0.776	0.108	0.071	0.018	0.019	0.004
0.888	0.169	0.113	0.029	0.030	0.006
0.999	0.294	0.200	0.052	0.055	0.013
1.108	1.273	1.555	0.386	0.358	0.915