The new Belle II charm-flavor tagger

Marko Starič

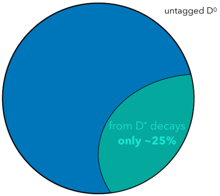
Belle II collaboration   Jožef Stefan Institute, Ljubljana

CHARM 2023

Introduction

- Flavor tagging is an essential ingredient of any CPV/mixing measurement

- Standard approach
  - exclusive reconstruction of $D^{*+} \to D^0\pi^+$
  - only about 25% of D^0 can be tagged



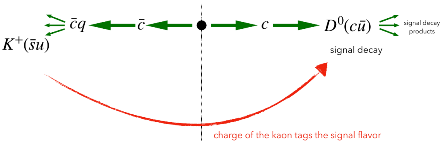
- New charm flavor tagger (CFT)* PRD 107, 112010 (2023)
  - exploits also information from other charmed hadron produced in $e^+e^- \to c\bar{c}$
  - by using charged particles not associated with the signal decay
  - these are part of the rest of the event (ROE)
  - include both, opposite-side and same-side particles

$\to$ conventional D^{*+} tagging is thus incorporated

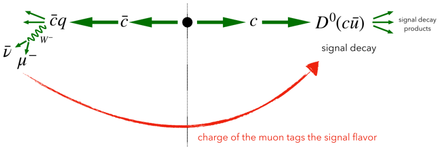
* inspired by B-flavor tagging algorithms

Few illustrations

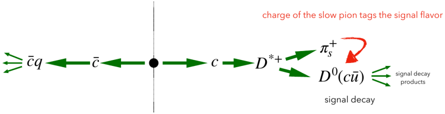
opposite side kaon tag



opposite side lepton tag



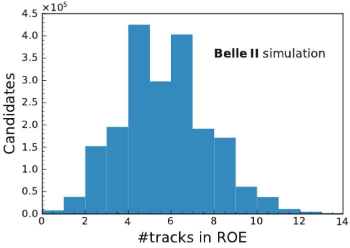
same side slow pion tag



- Not shown are particles emerging directly from fragmentation
- Other ROE particles whose charge is likely to be correlated with D^0 flavor are also used in CFT (opposite side slow pions, protons and pions)

Tagging algorithm

- Tagging decision provided with a binary classifier
  - histogram-based gradient-boosting decision tree (`scikit-learn` lib)
- ROE particles classified into two groups depending on their charge and ranked according to opening angle w.r.t D^0 momentum (in e^+e^- center-of-mass frame)
  - more collinear than those emerging purely from fragmentation
- The three top-ranked positive and the three top-ranked negative particles are selected for classification
  - $3 + 3$ found optimal
  - if event contains less, the missing ones are labeled as missing values



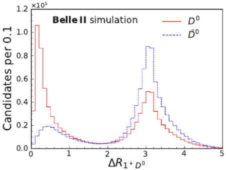
Number of ROE tracks

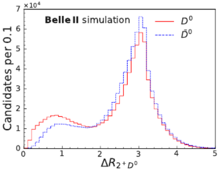
$D^0 \to \nu\nu$ simulation

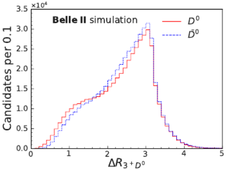
Classifier input

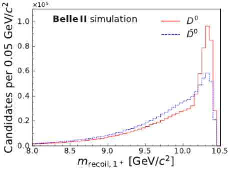
- Classifier input variables:
  - opening angles
  - differences between pion and kaon particle ID (likelihoods)
  - recoil masses of the highest-ranked positive and negative particle
    $m_{\text{recoil}} = \sqrt{(\mathbf{p}_{e^+e^-} - \mathbf{p}_{ROE})^2}$

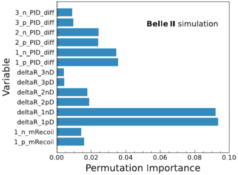
$\rightarrow$ 14 inputs in total





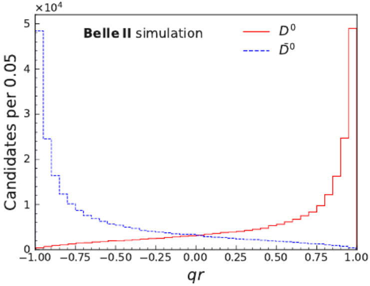






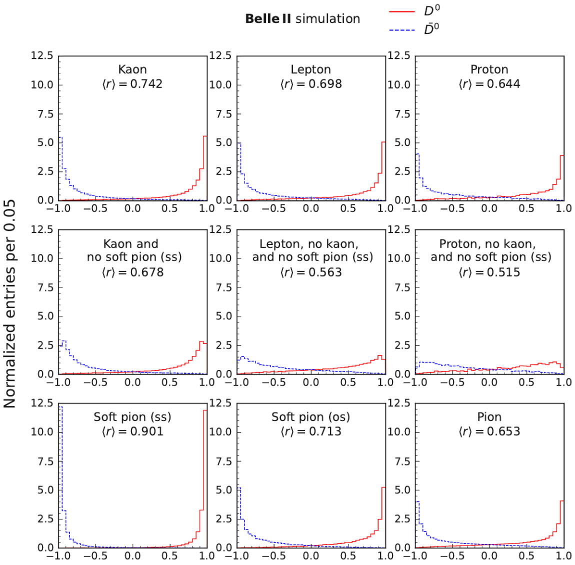
Classifier training

- Trained using simulated $D^0 \to \nu\nu$ events
  - to minimize possible correlations with the signal decay
  - every reconstructed particle belongs to ROE
- Trained with 1.35M decays
- Tested then with independent sample of 450k $D^0 \to \nu\nu$ events



$\to$ correct flavor predicted in $\sim$83% of decays

Classifier response to different tagging categories



Standard metrics of tagging performance

tagging efficiency: $\epsilon_{tag} = \dfrac{R+W}{R+W+U}$

mistag fraction: $\begin{cases} \dfrac{W}{R+W} & \text{if } W \leqslant R \\ 1 - \dfrac{W}{R+W} & \text{otherwise} \end{cases}$

dilution: $r = 1 - 2\omega$

tagging power: $\epsilon_{\text{tag}}^{\text{eff}} = \epsilon_{\text{tag}} r^2 = \epsilon_{\text{tag}}(1-2\omega)^2$

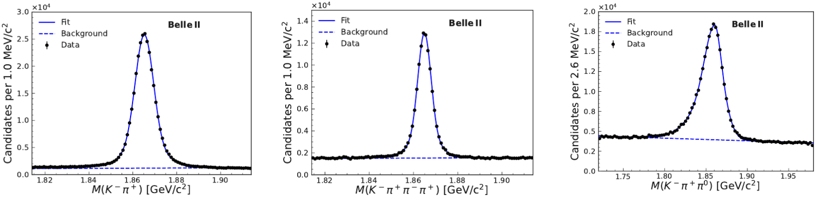
tagging decision: $q = \pm 1$

R (W), U: rightly (wrongly) tagged, untagged D^0 candidates
q: +1 for D^0, -1 for $\bar{D}^0$

classifier output $= q \times r$

Performance evaluation

- Evaluated on 362 fb^{-1} of Belle II data
- Performance studied with the following self-tagged signal decays
  - $D^0 \to K^-\pi^+$, $D^0 \to K^-\pi^+\pi^-\pi^+$, $D^0 \to K^-\pi^+\pi^0$
  - $D^+ \to K_S^0\pi^+$, $D^+ \to K^-\pi^+\pi^+$
  - $\Lambda_c^+ \to pK^-\pi^+$
- Inclusion of charged hadrons provides insight into contributions from various tagging categories (i.e. no same-side slow pion)
- Decay reconstruction involves selection of well fitted tracks from IR, our standard K_S^0 and π^0 reconstruction, particle ID and vertex fits



- Background subtraction performed by *sPlot* technique

Performance: results

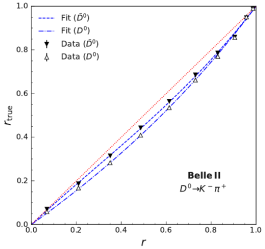
Signal decay	$\varepsilon_{\rm tag}$ (%)	$\Delta\varepsilon_{\rm tag}$ (%)	ω (%)	$\Delta\omega$ (%)	$\varepsilon_{\rm tag}^{\rm eff}$ (%)
$D^0 \to K^-\pi^+$	99.974 ± 0.004	-0.002 ± 0.007	19.09 ± 0.08	0.36 ± 0.17	38.22 ± 0.20
$D^0 \to K^-\pi^+\pi^-\pi^+$	99.794 ± 0.020	0.042 ± 0.039	19.13 ± 0.16	0.40 ± 0.32	38.05 ± 0.38
$D^0 \to K^-\pi^+\pi^0$	99.967 ± 0.006	-0.006 ± 0.012	19.34 ± 0.13	-0.22 ± 0.26	37.58 ± 0.32
$D^+ \to K^-\pi^+\pi^+$	99.843 ± 0.007	-0.026 ± 0.014	27.86 ± 0.08	0.80 ± 0.16	19.57 ± 0.14
$D^+ \to K^0_S\pi^+$	99.846 ± 0.019	0.037 ± 0.038	27.92 ± 0.23	1.83 ± 0.46	19.47 ± 0.41
$\Lambda_c^+ \to pK^-\pi^+$	99.832 ± 0.008	-0.022 ± 0.016	32.44 ± 0.09	0.52 ± 0.18	12.31 ± 0.13

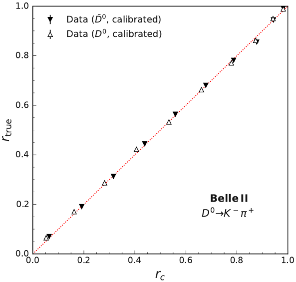
$\Delta\epsilon_{\rm tag}$ and $\Delta\omega$ measure the difference between charm and anti-charm hadron
contributions from wrong-sign D^0 decays are accounted for

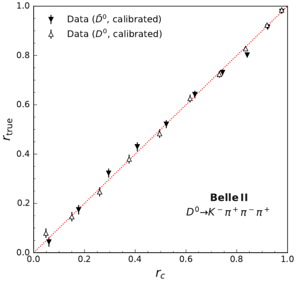
- Tagging efficiency almost 100%
  - independent of charmed hadron and its decay mode
- Mistag fraction independent of decay mode, but depends on the charmed hadron
  - absence of same-side slow pion in D^+ and Λ_c^+ flavor tagging
  - presence of proton tag in Λ_c^+ flavor tagging
- Mistag difference $\Delta\omega$
  - consistent with zero for D^0
  - significant deviations from zero for D^+ and Λ_c^+ due to charge detection asymmetry (ROE is not neutral)

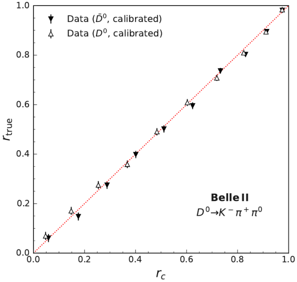
Calibration of CFT output

- Deviations from linear found when compared CFT output with true dilution
- CFT output corrected by calibration curve obtained from fit to $D^0 \to K^-\pi^+$









D^0 tagging power

- Measured with $D^0 \to K^-\pi^+$ on 362 fb^{-1}
- With calibrated CFT output:

$$\epsilon_{\rm tag}^{\rm eff} = (47.91 \pm 0.07({\rm stat}) \pm 0.51({\rm syst}))\%$$

- Systematic uncertainty dominated by background subtraction
  - should scale according to integrated luminosity

Impact on physics

- Effective increase in sample size estimated with $D^0 \to K^-\pi^+$
- Split into two disjoint subsets
  - D^{*+} tagged events
  - events that are not D^{*+} tagged

54.4 fb^{-1} of Belle II data

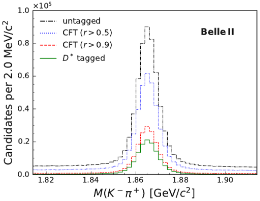
subset	signal yield	tagging power	tagged yield
D^{*+} tagged	126k	$\sim$100%	126k
the rest	388k	32.7%	127k

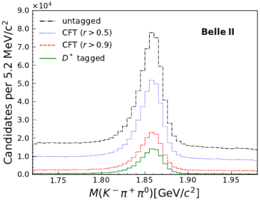
$\to$ effectively doubling the tagged sample size

$\to$ but increasing also background, hence increase in precision $< \sqrt{2}$

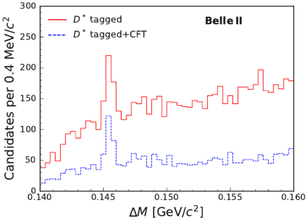
Impact on physics

- CFT provides also some discrimination between signal and background
- CFT output can be used in analyses not requiring flavor tagging
  - as additional variable in multi-dimensional fit
  - or as part of event selection to improve signal purity





wrong sign $D^0 \to K^+\pi^-\pi^0$



Signal purity

tagging	$D^0 \to K^-\pi^+$	$D^0 \to K^-\pi^+\pi^0$
untagged	0.67	0.34
CFT $r > 0.5$	0.73	0.38
CFT $r > 0.9$	0.84	0.53
D^{*+}	0.94	0.80

Doubly tagged sample:

much improved S/B

only 24% signal loss

Conclusions

- Novel charm-flavor tagging algorithm developed for Belle II.
  - Explores correlations between production flavor and electric charges of particles in ROE.
  - Uses boosted decision trees trained on simulated data.
- Response calibrated and evaluated on data with several self-tagged D^0 decays. The effective tagging efficiency is around 48% and independent of the D^0 decay mode.
- It can roughly double the effective sample size for charm CPV/mixing measurements.
- It can be used also to suppress background for the measurements where flavor tagging is not required.