39th International Conference on High Energy Physics — Seoul, Korea, 4-11 July 2018

Recent improvements and prospects with flavour tagging at LHCb

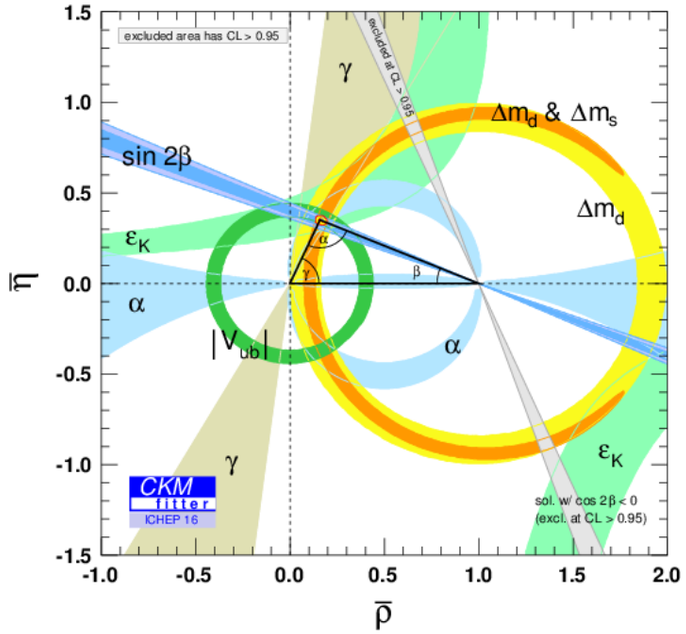
Simon Akar[1]

on behalf of the LHCb Collaboration

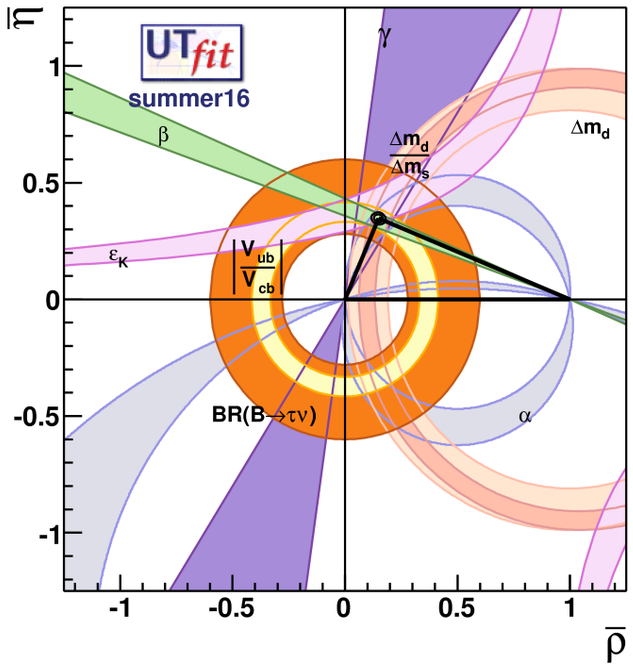
[1]*University of Cincinnati, Cincinnati, OH, United States*

Time-dependent *CP* violation

- **CKM mechanism introduces *CP* violation and neutral meson mixing phenomena**
  - The presence of new heavy particles exchanged in virtual loops could introduce additional phases altering the corresponding measurements
  - Constraining these phases put stringent limits on a large range of NP models



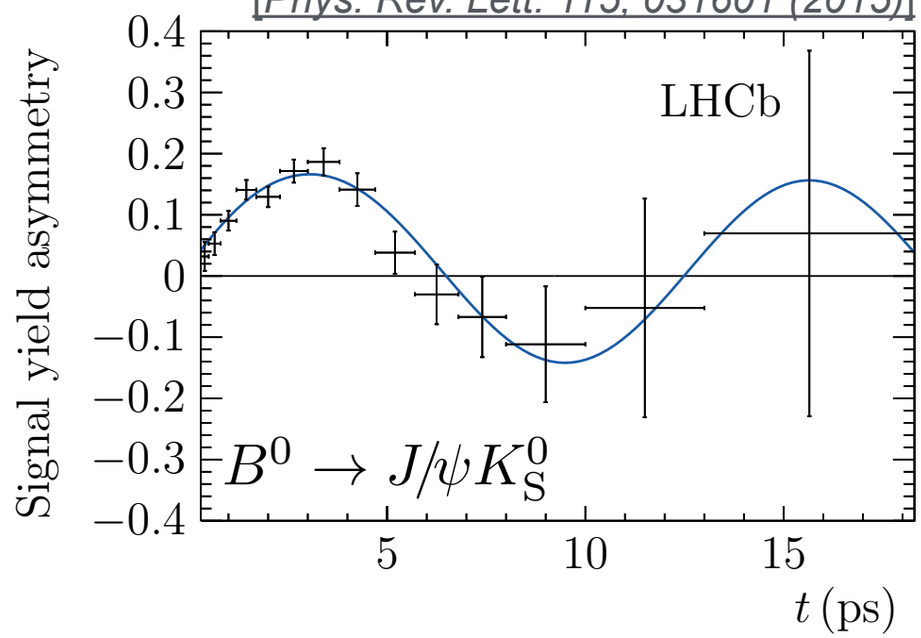
- ***CP* violation is needed to explain baryon asymmetry in the Universe**
  - Discovered in 1964 in the kaons, 2004 in the *B* and each time awarded with Nobel Prizes
  - Still missing 10 orders of magnitudes!



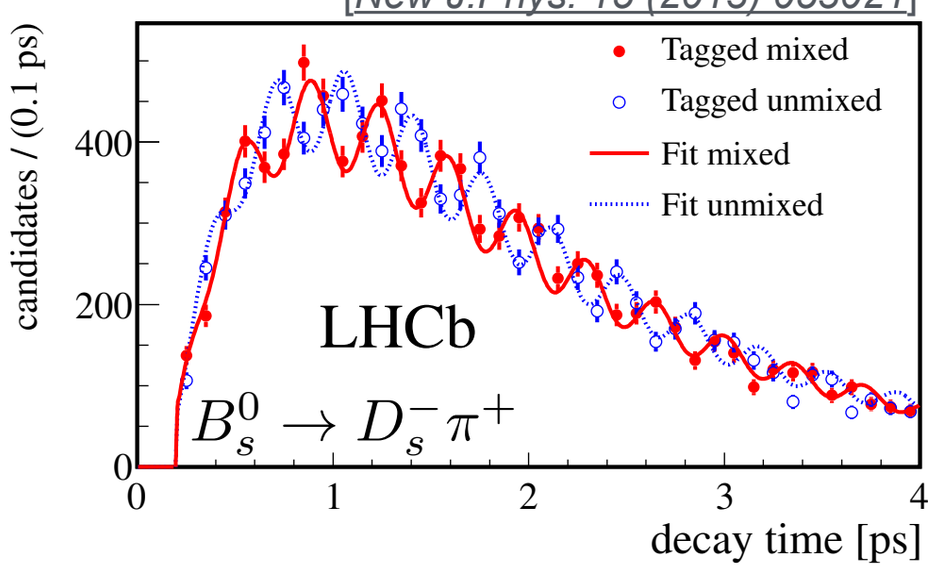
- **Experimentally, *CP* violation observables accessed through ratios of measured quantities**
  - Cancelation of many experimental systematics
  - Flagship measurements for LHCb and Belle II

Time-dependent *CP* violation

[Phys. Rev. Lett. 115, 031601 (2015)]



[New J.Phys. 15 (2013) 053021]



- **Time-dependent *CP* asymmetry:**

$$\mathcal{A}_{CP}(t) = \frac{\Gamma\left(\overline{B}^0_q(t) \to f\right) - \Gamma\left(B^0_q(t) \to f\right)}{\Gamma\left(\overline{B}^0_q(t) \to f\right) + \Gamma\left(B^0_q(t) \to f\right)}$$

$$= \frac{\mathcal{S}_f \sin(\Delta m t) - \mathcal{C}_f \cos(\Delta m t)}{\cosh\left(\frac{\Delta\Gamma t}{2}\right) + \mathcal{A}_{\Delta\Gamma} \sinh\left(\frac{\Delta\Gamma t}{2}\right)}$$

Measurements of time-dependent asymmetries and decay rates require knowledge of ***B* flavour at the production time (*t* = 0)**

Flavour tagging algorithms tag the candidate as ***B*** or $\overline{B}$ (tag decision) with some ***efficiency*** and ***mistag probability***

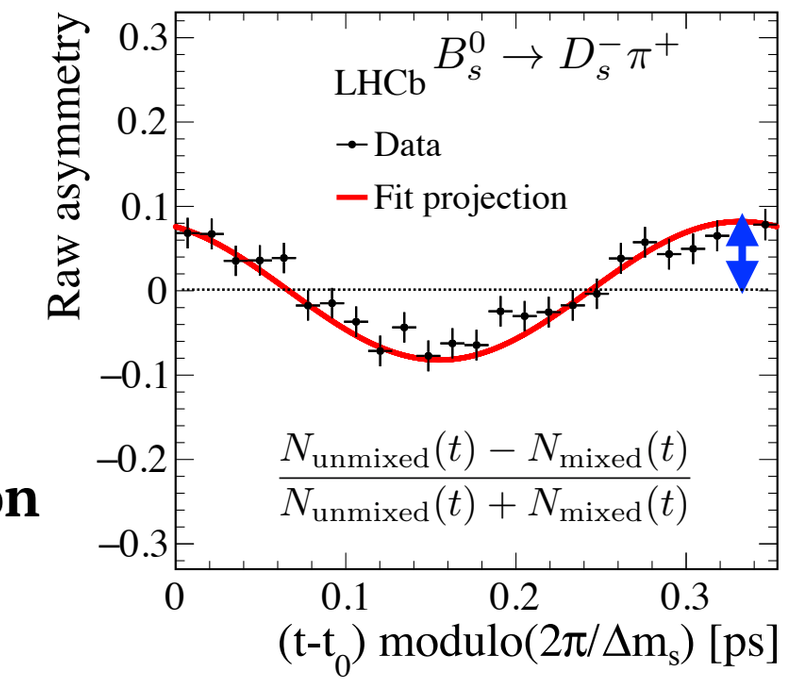
Flavour tagging definitions

[Journal of Instrumentation 11 (2016) P05010]

- **Tagging Efficiency: fraction of tagged events**

$$\varepsilon_{\text{tag}} = \frac{N_{\text{wrong}} + N_{\text{correct}}}{N_{\text{wrong}} + N_{\text{correct}} + N_{\text{untagged}}}$$

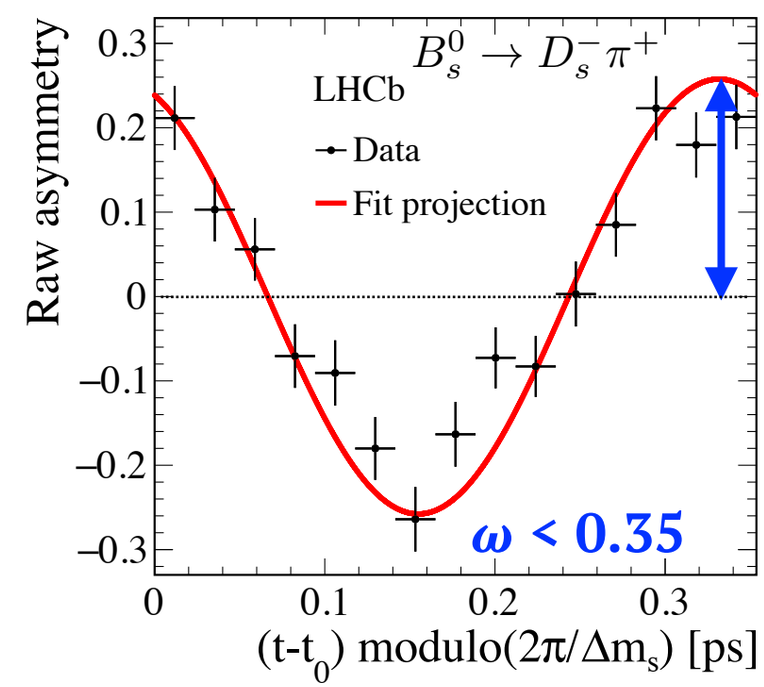
  - Correlated with the transverse momentum of the signal *B*



- **Mistag fraction: fraction of events with wrong tag decision**

$$\omega = \frac{N_{\text{wrong}}}{N_{\text{wrong}} + N_{\text{correct}}}$$

  - ***Dilution: D = (1-2ω)*** of asymmetries and decay rates.
  - Predicted mistag, η, computed by taggers needs calibration $\omega(\eta)$ to provide unbiased estimate of ω



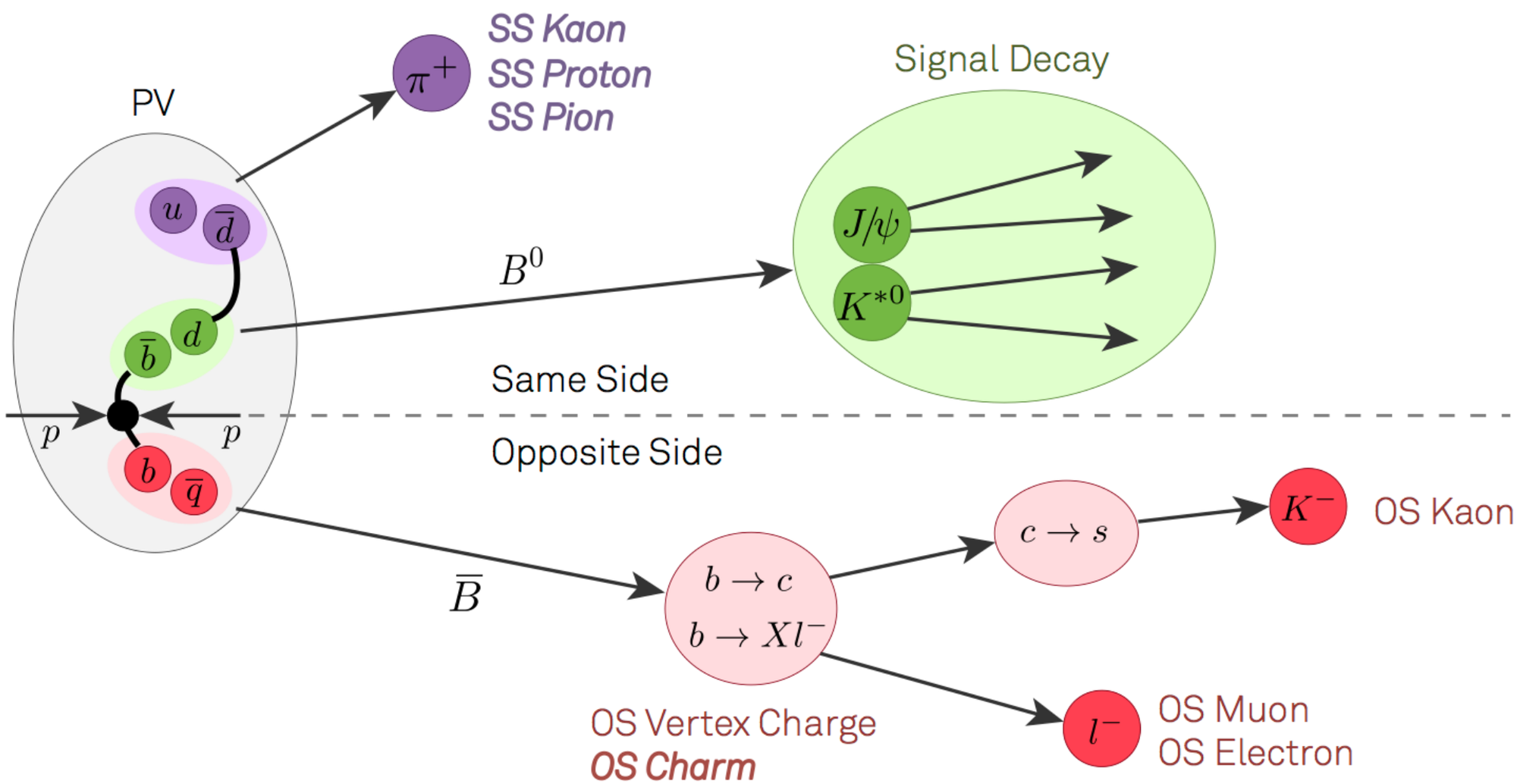
- **Tagging power: statistical degradation of *CP* asym.**

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}} D^2 = \varepsilon_{\text{tag}} \left\langle (1-2\omega)^2 \right\rangle$$

$$\sigma^{\text{stat}}(CP\ \text{asym}) \propto \frac{1}{\sqrt{\varepsilon_{\text{eff}} N}}$$

Flavour tagging algorithms

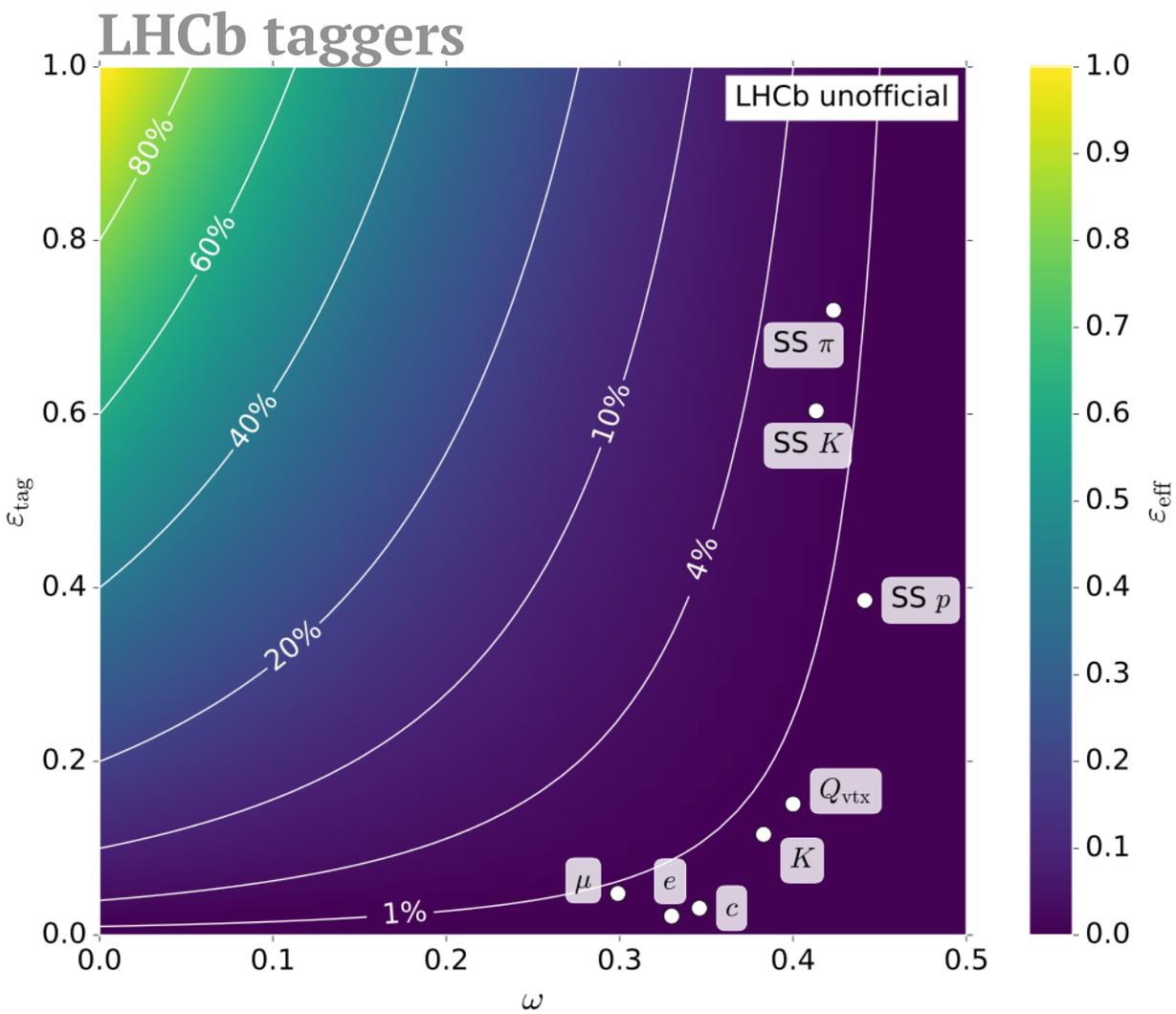
Same side (SS): exploit the charge of the **fragmentation particle** (pion, kaon, proton) produced next to the signal *b*-hadron in the hadronisation process to infer its initial flavour (t=0)

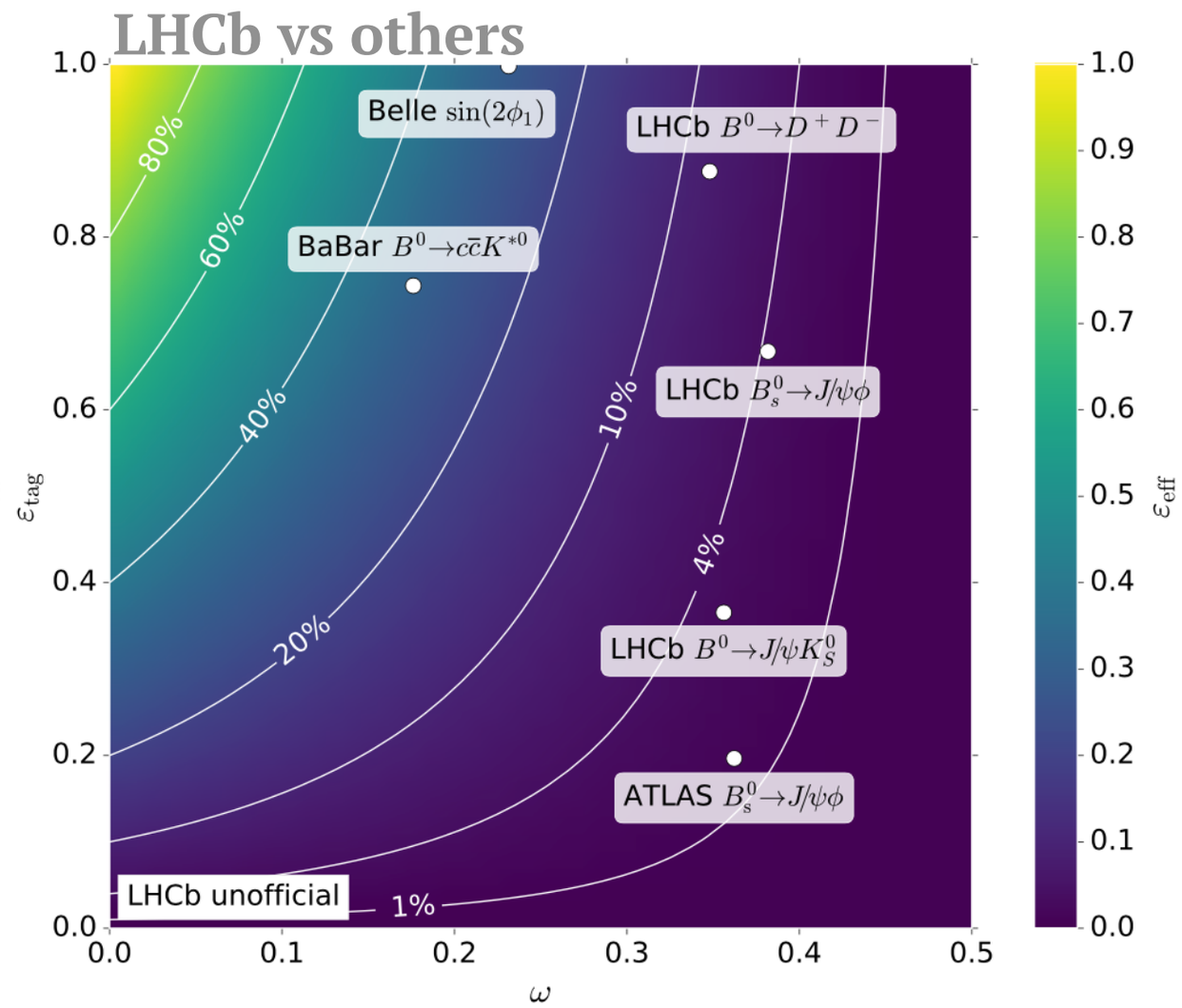


Opposite side (OS): exploit the charge of the particle (lepton, kaon, charm decays) or of the **reconstructed secondary vertex produced from the other *b*-hadron** in the event to infer signal *b*-hadron initial flavour (t=0)

Typical performances

[CERN-THESIS-2016-152]





- **OS taggers: good mistag but low decision rate**
- **SS taggers: modest mistag but high decision rate**
- **Tagging performances are channel dependent**

$$\sigma^{\rm stat}(CP\ {\rm asym}) \propto \frac{1}{\sqrt{\varepsilon_{\rm eff} N}}$$

Crucial ingredient

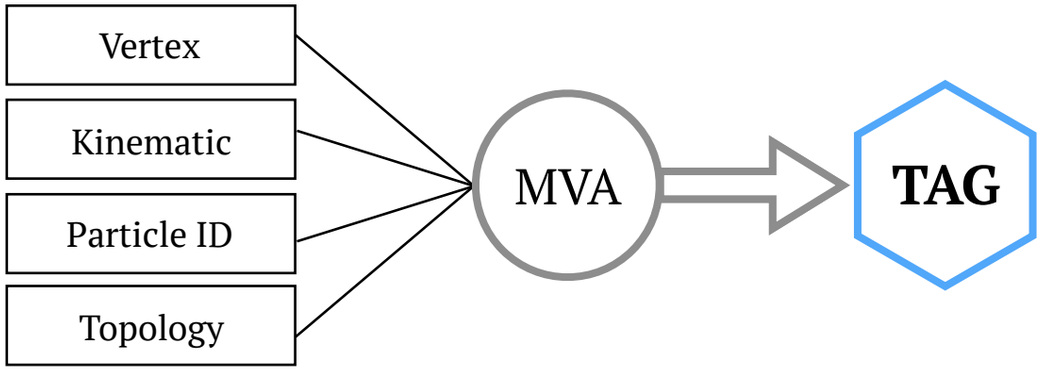
Updated from [EPJ C 73 (2013) 2373]

Type	Observable	LHC Run 1	LHCb 2018	**LHCb Upgrade**	Theory
B_s^0 mixing	$2\beta_s\ (B_s^0 \to J/\psi\ \phi)$ (mrad)	49 [59]	25	9	~ 3
	$2\beta_s\ (B_s^0 \to J/\psi\ f_0(980))$ (mrad)	68 [85]	35	12	~ 10
	$a_{\rm sl}^s\ (10^{-3})$	2.6 [78]	1.4	0.5	0.03
Gluonic penguins	$2\beta_s^{\rm eff}(B_s^0 \to \phi\phi)$ (mrad)	150 [86]	100	18	20
	$2\beta_s^{\rm eff}(B_s^0 \to K^{*0}\bar{K}^{*0})$ (mrad)	190[a]	130	23	< 20
	$2\beta^{\rm eff}(B^0 \to \phi K_S^0)$ (mrad)	300[a]	200	36	20
Right-handed currents	$2\beta_s^{\rm eff}(B_s^0 \to \phi\gamma)$ (mrad)	200[a]	130	25	< 10
	$\tau^{\rm eff}(B_s^0 \to \phi\gamma)/\tau_{B_s^0}$	5 %[a]	3.2 %	0.6 %	0.2 %
Electroweak penguins	$S_3(B^0 \to K^{*0}\mu^+\mu^-; 1 < q^2 < 6\,{\rm GeV}^2/c^4)$	0.038 [87]	0.020	0.007	0.02
	$q_0^2\, A_{\rm FB}(B^0 \to K^{*0}\mu^+\mu^-)$	10 % [88]	5 %	1.9 %	~ 7 %
	$A_{\rm I}(K\mu^+\mu^-; 1 < q^2 < 6\,{\rm GeV}^2/c^4)$	0.09 [89]	0.05	0.017	~ 0.02
	$\mathcal{B}(B^+ \to \pi^+\mu^+\mu^-)/\mathcal{B}(B^+ \to K^+\mu^+\mu^-)$	14 % [90]	7 %	2.4 %	~ 10 %
Higgs penguins	$\mathcal{B}(B_s^0 \to \mu^+\mu^-)\ (10^{-9})$	0.7 [91][b]	0.5	0.19	0.3
	$\mathcal{B}(B^0 \to \mu^+\mu^-)/\mathcal{B}(B_s^0 \to \mu^+\mu^-)$	50 % [91][b]	110 %	40 %	~ 5 %
Unitarity triangle angles	$\gamma\ (B \to D^{(*)}K^{(*)})$	7° [92, 93]	4°	0.9°	negligible
	$\gamma\ (B_s^0 \to D_s K)$	17° [94]	11°	2.0°	negligible
	$\beta\ (B^0 \to J/\psi\ K_S^0)$	1.3° [95]	0.8°	0.31°	negligible
Charm CP violation	$A_\Gamma(D^0 \to K^+K^-)\ (10^{-4})$	3.2 [96]	2.2	0.4	–
	$\Delta\mathcal{A}^{CP}\ (10^{-3})$	0.8 [97]	0.5	0.1	–

Many key measurements in the LHCb physics program depend on flavour tagging

Developing FT algorithms

- **Tagging tracks identification**
  - *pp* collisions harsh environment
  - few tracks among $O(10^2)$
  - FT algorithms based on MVA classifiers designed to isolate specific physical process (SS tracks, OS tracks divided by categories)



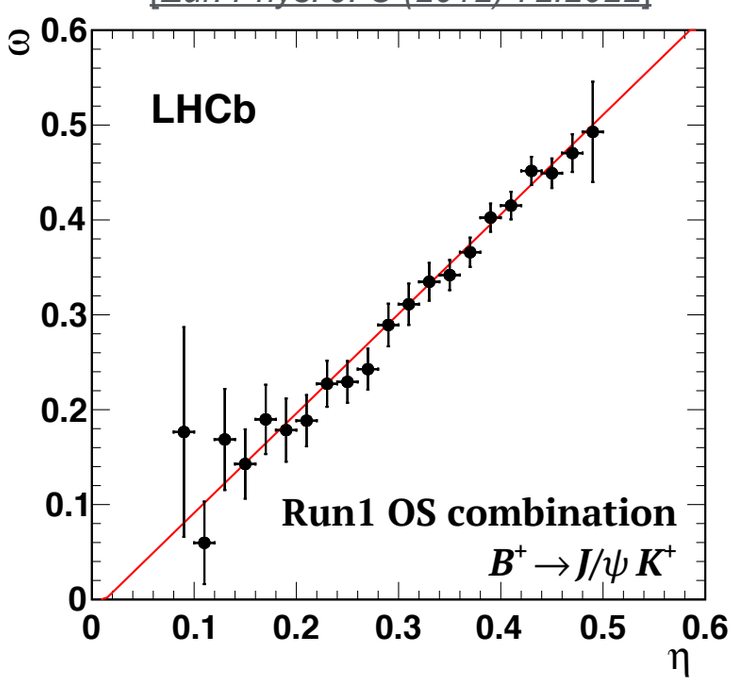
- **Taggers calibration**

$$\omega = p_0 + p_1(\eta - \langle\eta\rangle)$$

measured ev-by-ev mistag — estimated ev-by-ev mistag — estimated mean mistag

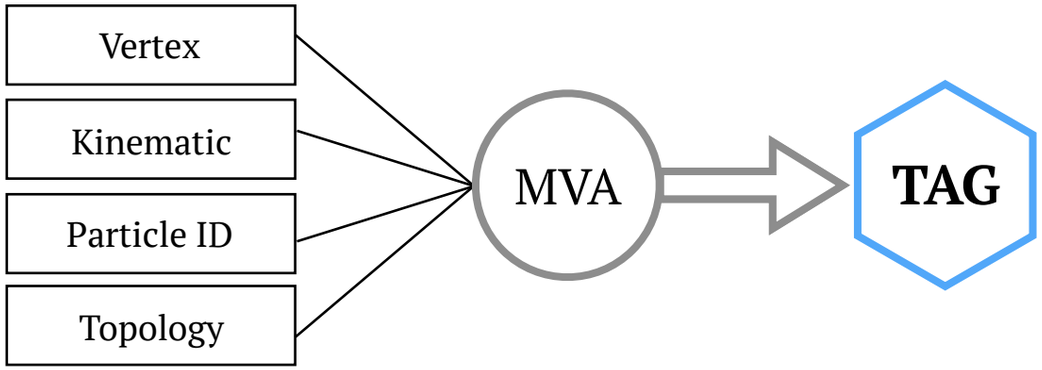
  - Calibration performed on control channels:
    - Charged modes: $B^+ \to J/\psi K^+$, $B^+ \to D^0 \pi^+$
      No oscillation, large samples
    - Neutral B^0 modes: $B^0 \to J/\psi K^*$, $B^0 \to D^- \pi^+$
      large samples
    - Neutral B_s^0 modes: $B_s^0 \to D_s^- \pi^+$
      only control decay for data SSK, statistically limited

[Eur. Phys. J. C (2012) 72:2022]



Developing FT algorithms

- **Tagging tracks identification**
  - *pp* collisions harsh environment
  - few tracks among $O(10^2)$
  - FT algorithms based on MVA classifiers designed to isolate specific physical process (SS tracks, OS tracks divided by categories)



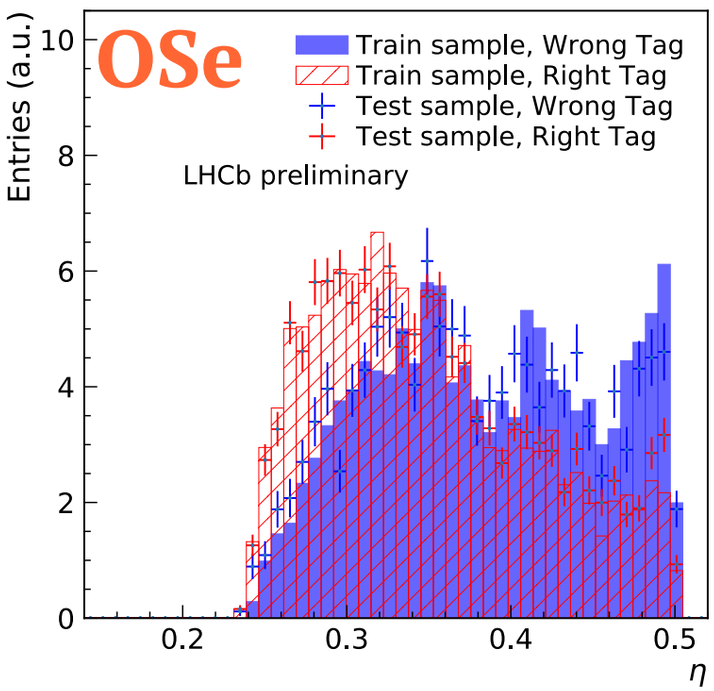
- **Tagging for Run2 analyses:**
  - FT algorithms sensitive to data taking conditions
    - center-of-mass energy
    - track multiplicity/number of primary vertices
    - trigger efficiencies

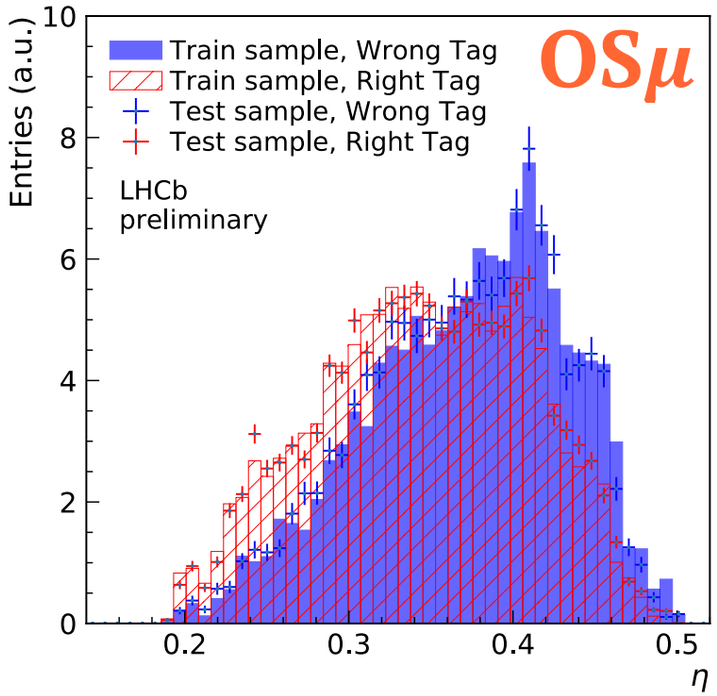
$\Rightarrow$ **most taggers developed in Run1 need to be optimised for Run2 conditions!**

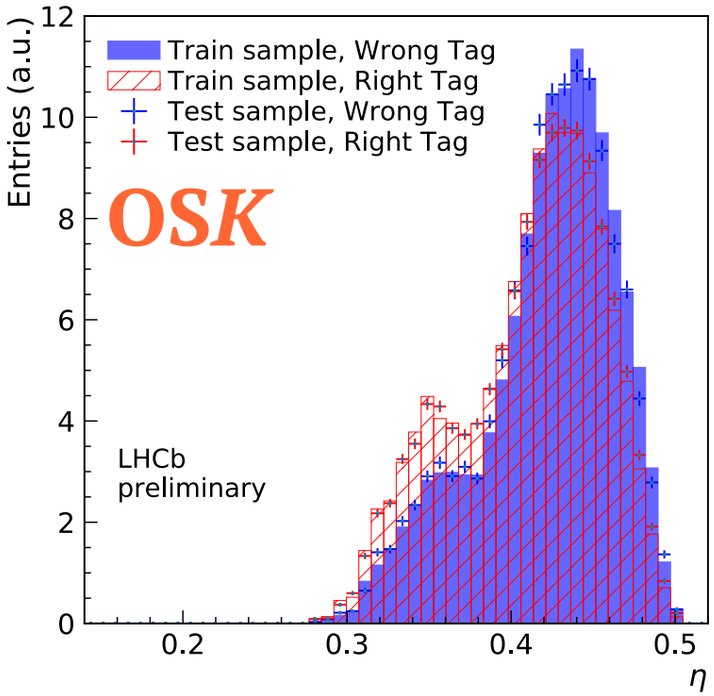
Run2 OS taggers development

- **Strategy:** $B^+ \to J/\psi K^+$ data (2016)
  - Only OSe, OSμ, OSK (significant Run1 → Run2 performances decrease)
  - Optimise cut-based selection maximizing the average tagging power
  - candidate with higher p_T is taken as tagging particle
  - BDT trained to classify right and wrong tagged B candidates

Separation between *wrong* and *right* tag

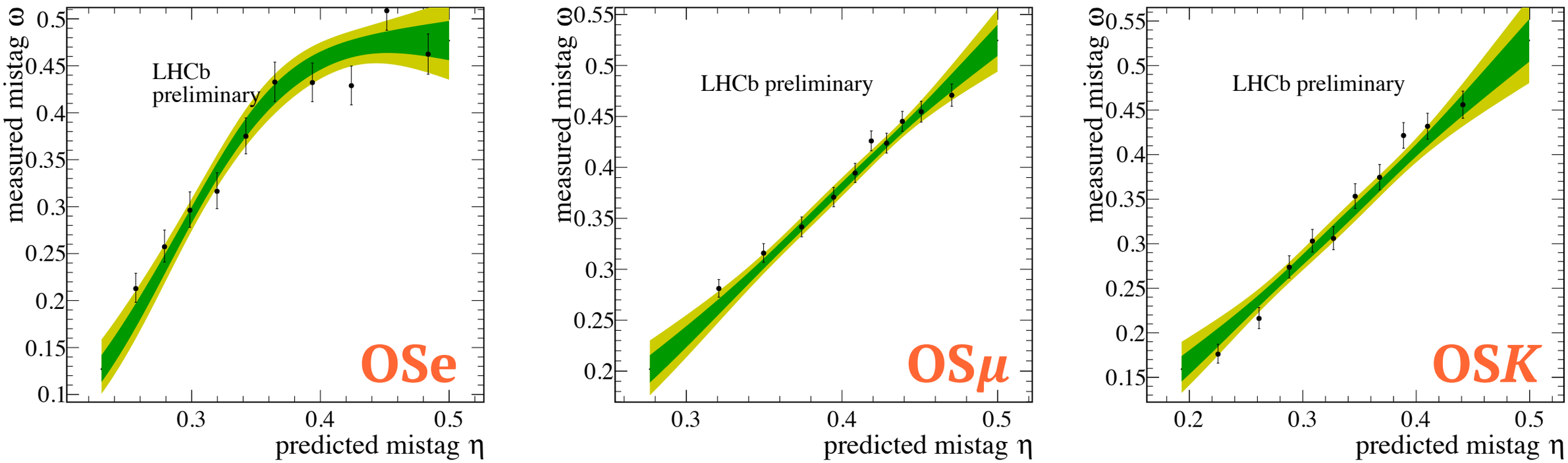






Run2 OS taggers development

- **Performances:** $B^0 \to D^- \pi^+$ data (2016)

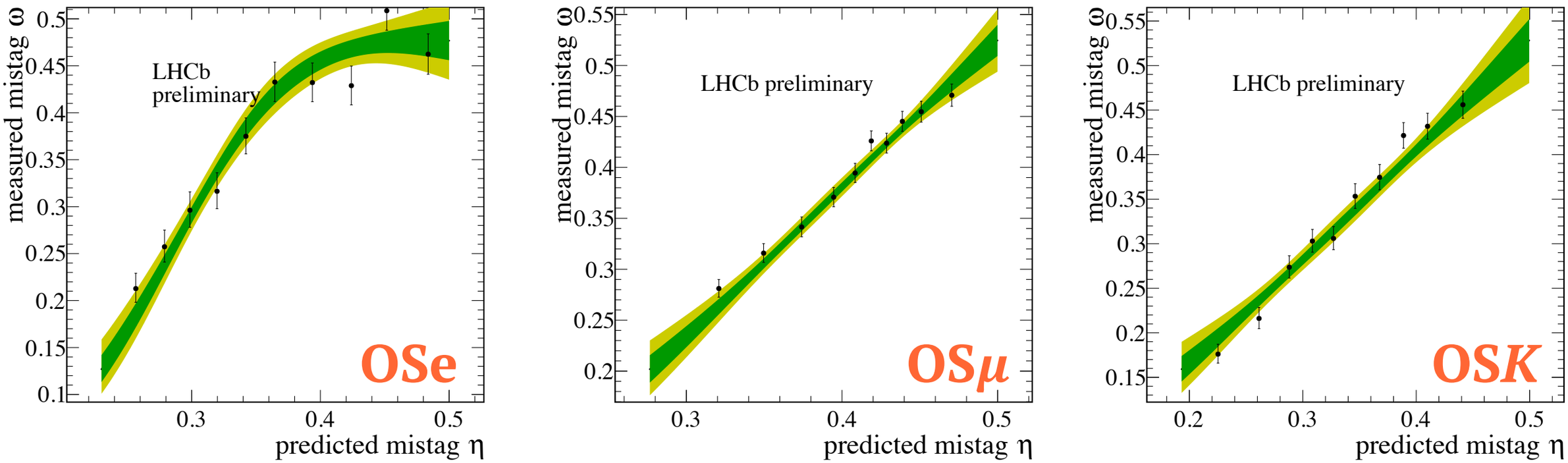


Tagger	ϵ (%)	ω (%)	$\epsilon\langle D^2\rangle = \epsilon\,(1-2\omega)^2$ (%)
OS μ	9.15 ± 0.05	$30.71 \pm 0.05(\text{stat}) \pm 0.43(\text{cal})$	$1.36 \pm 0.01(\text{stat}) \pm 0.06(\text{cal})$
OS e	4.45 ± 0.04	$34.04 \pm 0.07(\text{stat}) \pm 0.60(\text{cal})$	$0.45 \pm 0.01(\text{stat}) \pm 0.03(\text{cal})$
OS K	19.60 ± 0.07	$37.56 \pm 0.02(\text{stat}) \pm 0.31(\text{cal})$	$1.21 \pm 0.01(\text{stat}) \pm 0.06(\text{cal})$

Run2 re-optimised performances improved to Run1 level

Run2 OS taggers development

- **Performances:** $B^0 \to D^- \pi^+$ data (2016)

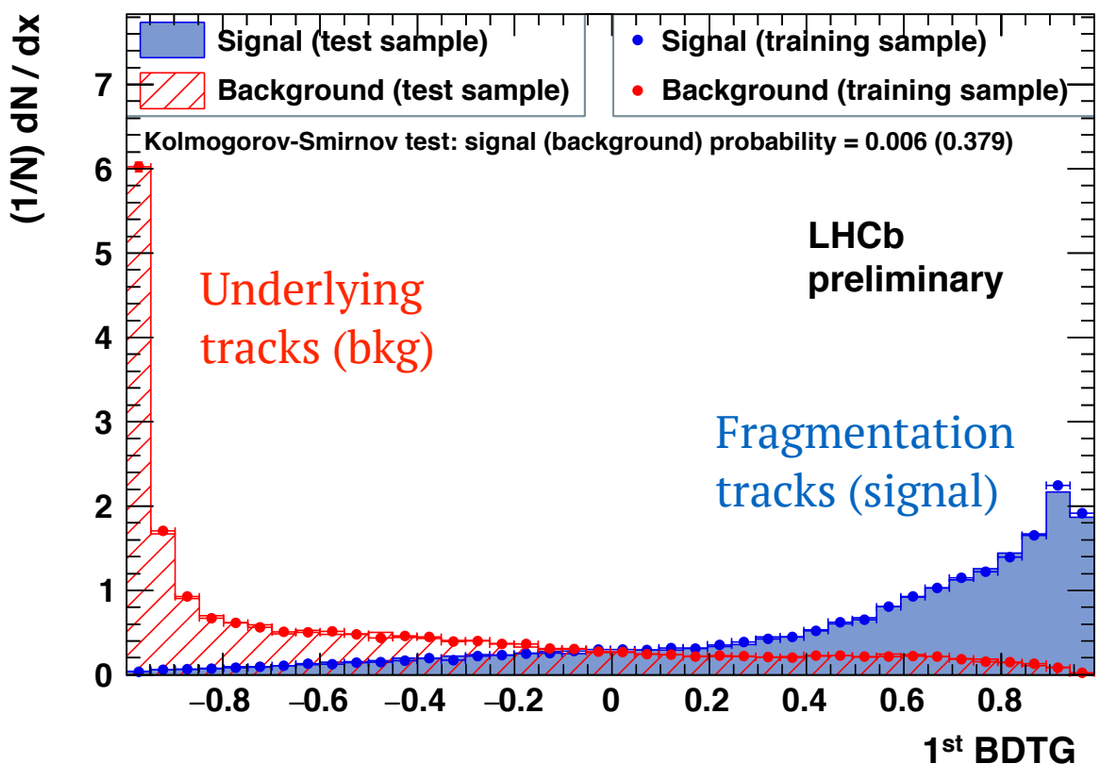


Tagger	ϵ (%)	ω (%)	$\epsilon\langle D^2\rangle = \epsilon\,(1-2\omega)^2$ (%)
OS μ	9.15 ± 0.05	$30.71 \pm 0.05(\text{stat}) \pm 0.43(\text{cal})$	$1.36 \pm 0.01(\text{stat}) \pm 0.06(\text{cal})$
OS e	4.45 ± 0.04	$34.04 \pm 0.07(\text{stat}) \pm 0.60(\text{cal})$	$0.45 \pm 0.01(\text{stat}) \pm 0.03(\text{cal})$
OS K	19.60 ± 0.07	$37.56 \pm 0.02(\text{stat}) \pm 0.31(\text{cal})$	$1.21 \pm 0.01(\text{stat}) \pm 0.06(\text{cal})$
Vertex Charge	20.83 ± 0.08	$36.99 \pm 0.03(\text{stat}) \pm 0.31(\text{cal})$	$1.41 \pm 0.01(\text{stat}) \pm 0.07(\text{cal})$
OS c	5.03 ± 0.04	$34.06 \pm 0.03(\text{stat}) \pm 0.62(\text{cal})$	$0.51 \pm 0.01(\text{stat}) \pm 0.04(\text{cal})$
OS Combination	40.15 ± 0.09	$35.12 \pm 0.03(\text{stat}) \pm 0.21(\text{cal})$	$3.56 \pm 0.01(\text{stat}) \pm 0.10(\text{cal})$

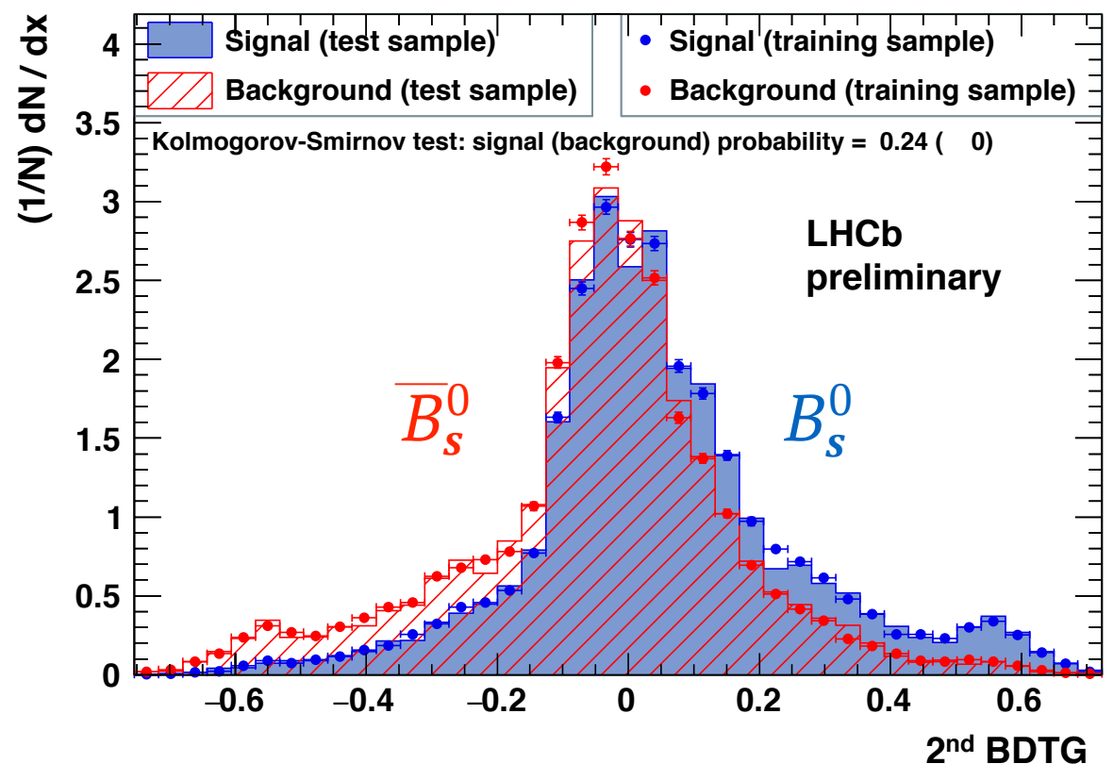
Overall Run2 OS taggers performances similar to Run1 level

Run2 SSK development

- **Strategy:** $B_s^0 \to D_s^- \pi^+$ simulation (due to fast B_s^0 oscillation)
  - Loose preselection cuts (kine & track quality criteria)
  - 1st classifier to separate fragmentation tracks from underlying event tracks
  - 2nd classifier to separate B_s^0 and $\overline{B}_s^0$ events based on information from 3 « best » tracks based on 1st classifier output



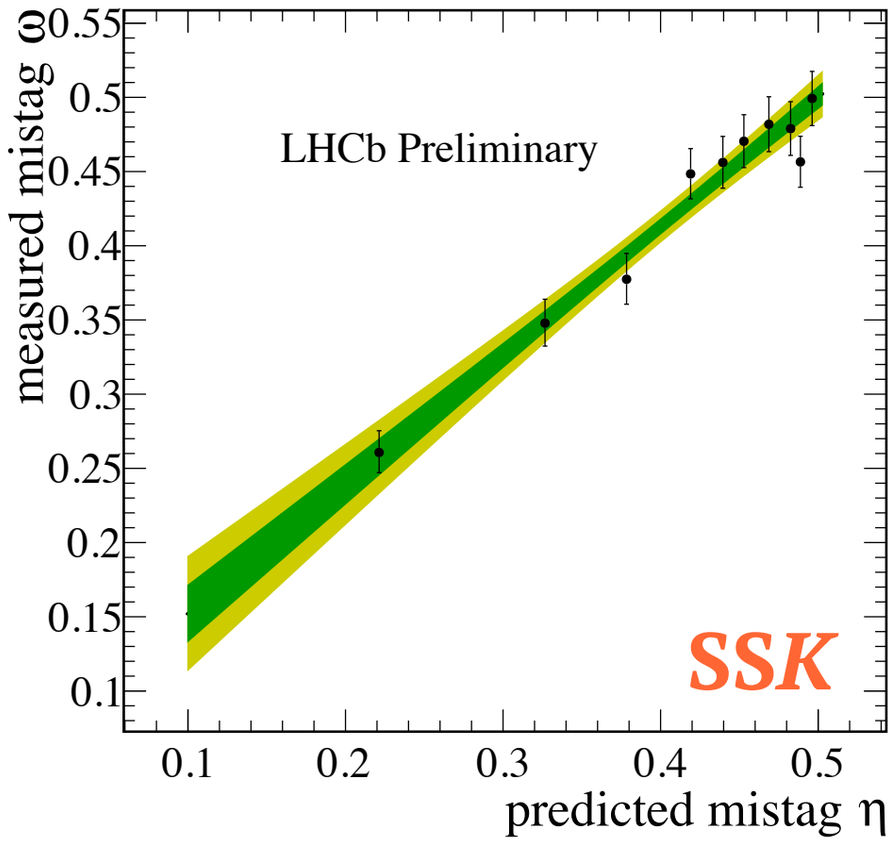
Separation between underlying and fragmentation tracks



Separation between $\bar{B}_s^0$ and B_s^0

Run2 *SSK* development

- **Performances:** $B_s^0 \to D_s^- \pi^+$ data (2016)

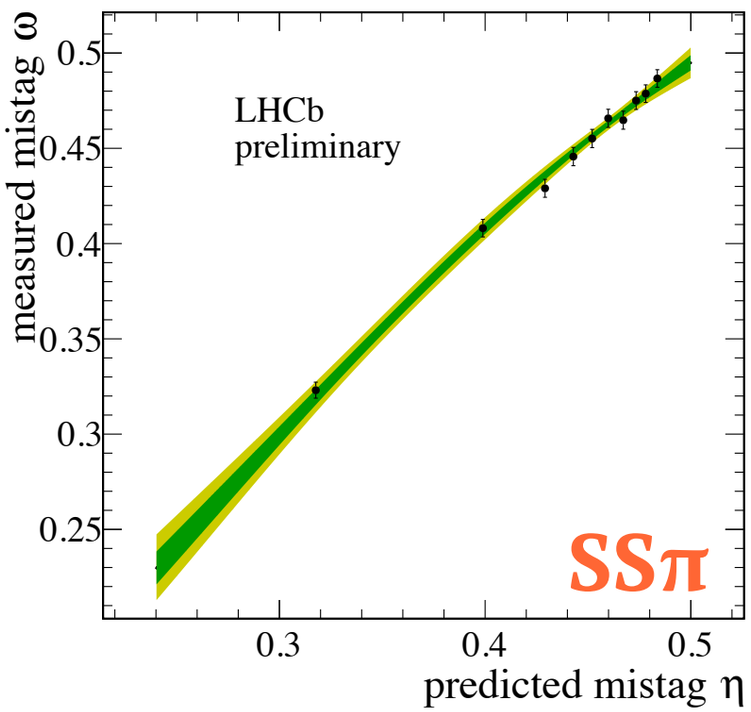


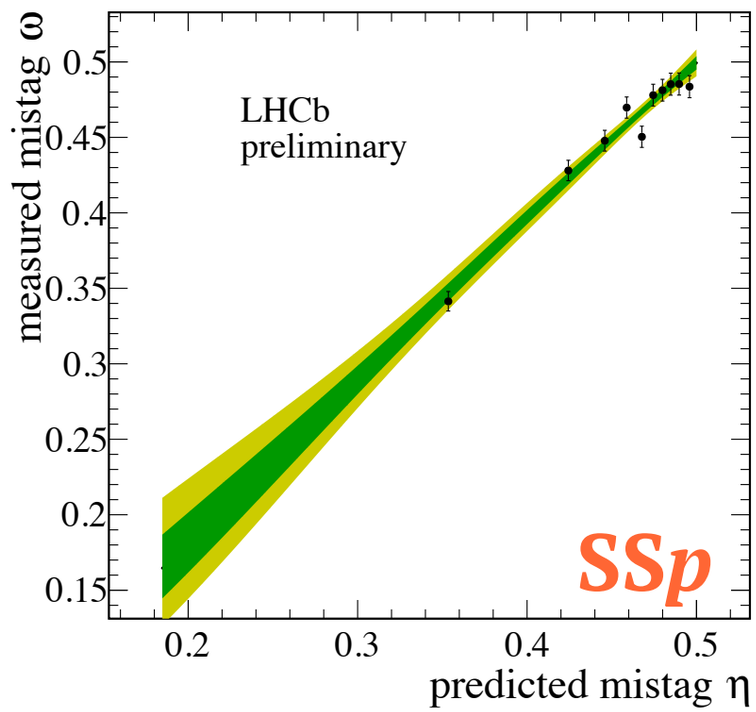
Tagger	ϵ (%)	ω (%)	$\epsilon\langle D^2\rangle = \epsilon\,(1-2\omega)^2$ (%)
SS K	68.19 ± 0.18	$39.67 \pm 0.04(\text{stat}) \pm 0.51(\text{cal})$	$2.91 \pm 0.02(\text{stat}) \pm 0.29(\text{cal})$

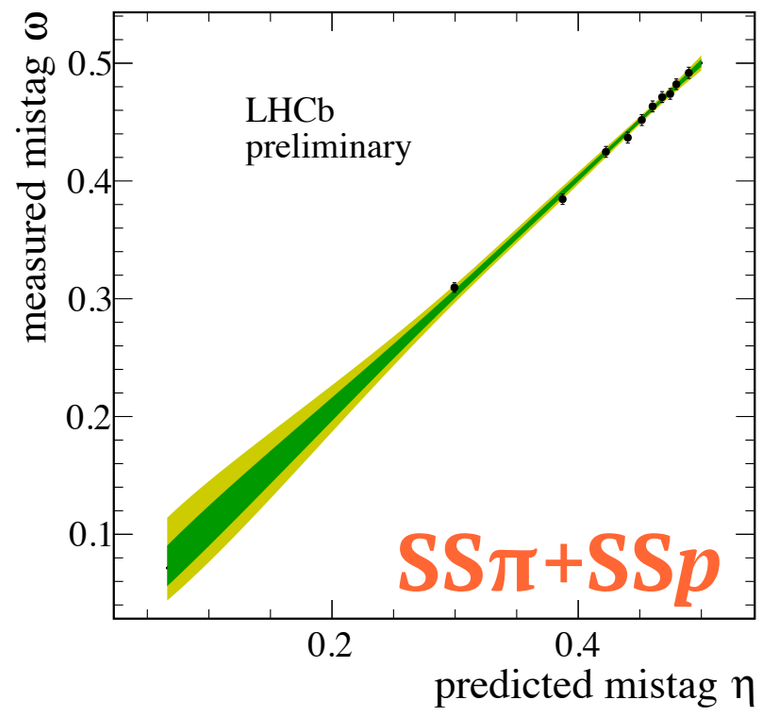
O(45%) improvement tagging power w.r.t. Run1

Other SS taggers

- **Performances:** $B^0 \to D^- \pi^+$ data (2016)







Tagger	ϵ (%)	ω (%)	$\epsilon\langle D^2\rangle = \epsilon\,(1-2\omega)^2$ (%)
SS π	83.49 ± 0.07	$42.56 \pm 0.01(\text{stat}) \pm 0.14(\text{cal})$	$1.85 \pm 0.01(\text{stat}) \pm 0.07(\text{cal})$
SS p	37.77 ± 0.09	$43.65 \pm 0.02(\text{stat}) \pm 0.22(\text{cal})$	$0.61 \pm 0.01(\text{stat}) \pm 0.04(\text{cal})$
SS Combination	87.59 ± 0.06	$41.79 \pm 0.02(\text{stat}) \pm 0.14(\text{cal})$	$2.36 \pm 0.01(\text{stat}) \pm 0.08(\text{cal})$

***O*(10%) improvement tagging power w.r.t. Run1**

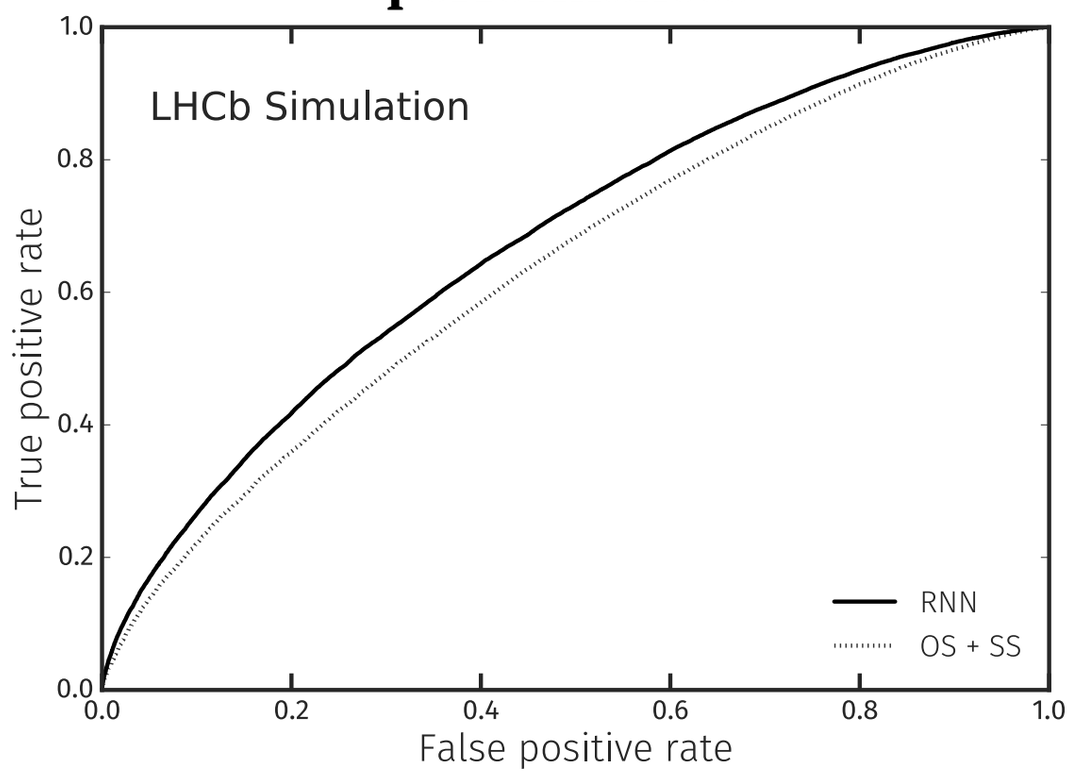
Inclusive tagger

- **Strategy:**
  - Improving individual taggers becoming harder and harder
  - Change of paradigm is needed ⇒ inclusive tagging
    - exploiting the full event (no explicit categories)
  - Current implementation based on Recursive Neural Network (RNN)
    - natural approach to handling variable-sized (tracks, vertices)
    - successfully used for the Jet Flavour Tagging [Phys. Rev. D 94, 112002 (2016)]
    - RNN trained using simulated events of $B^+ \to J/\psi K^+$
    - kinematic, topological, PID and tracking informations as inputs
    - under development

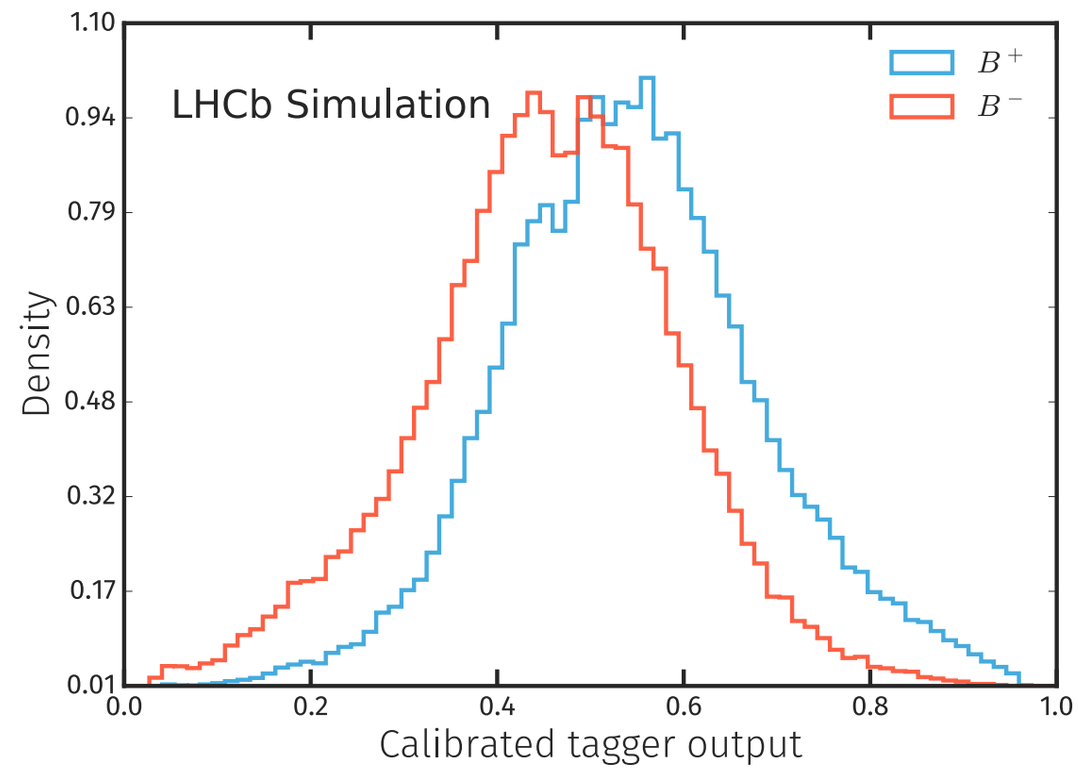
Inclusive tagger

- **Preliminary results:** $B^+ \to J/\psi K^+$ simulation (2016)

Comparison between "classical" taggers and inclusive tagger (RNN) performances



Inclusive tagger separation power between B^+ and B^-



Promising preliminary results
Need further studies (neutral modes)

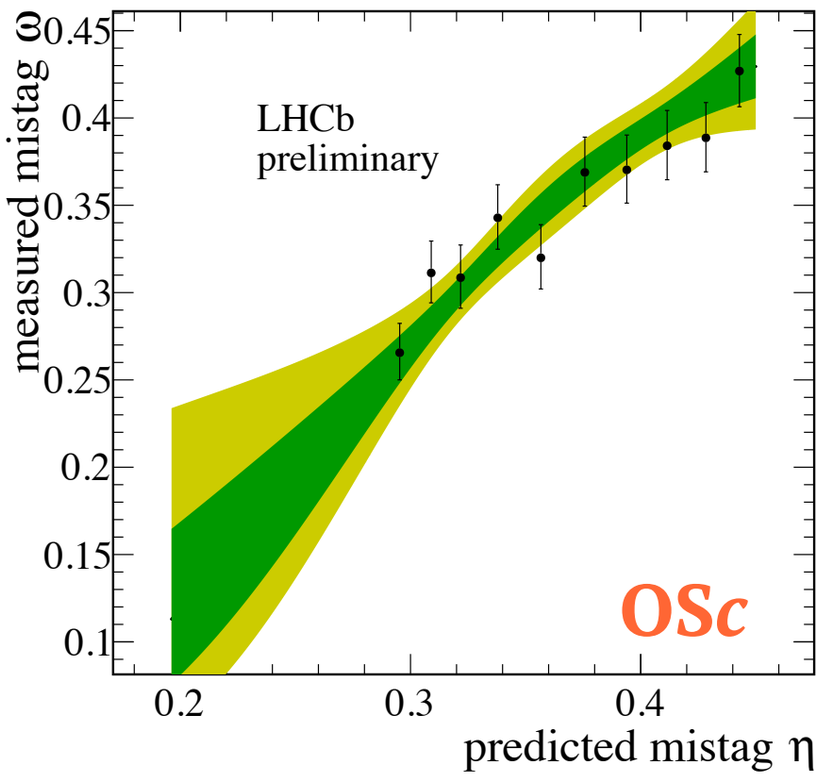
Summary & prospects

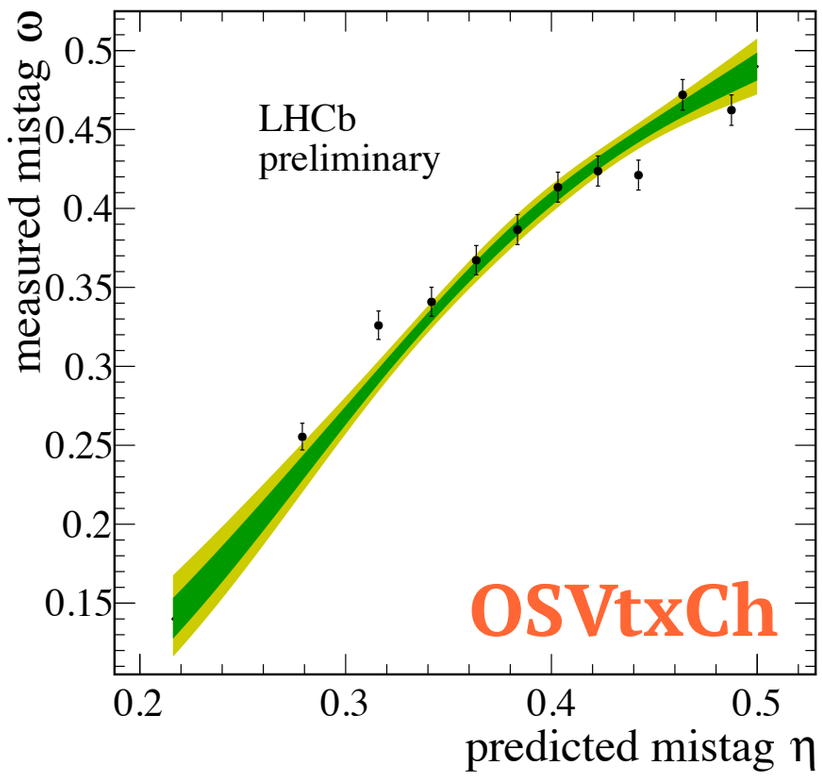
- **Flavour tagging essential to study neutral *B* meson mixing and *CP* violation**
- ***pp* collisions harsh environment**
- **Good detector + new algorithms: improved tagging performances in LHC Run2**
  - **OS improvement → stable performances wrt Run1**
  - ***SSK O*(45%) εD^2 improvement Run1**
  - **Overall FT performances in LHCb better in Run2 than in Run1**
- **Development of new algorithms ongoing:**
  - **inclusive tagging paradigm → very promising preliminary results**

Additional material

Other Run2 OS taggers

- **Performances:** $B^0 \to D^- \pi^+$ data (2016)



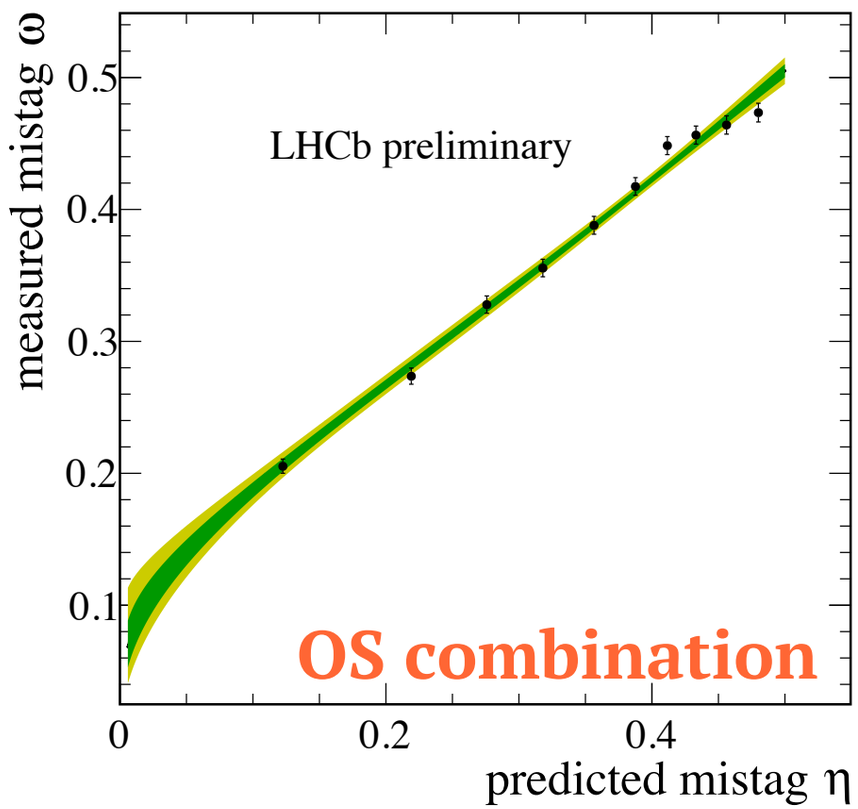


Tagger	ϵ (%)	ω (%)	$\epsilon\langle D^2\rangle = \epsilon\,(1-2\omega)^2$ (%)
Vertex Charge	20.83 ± 0.08	$36.99 \pm 0.03(\text{stat}) \pm 0.31(\text{cal})$	$1.41 \pm 0.01(\text{stat}) \pm 0.07(\text{cal})$
OS c	5.03 ± 0.04	$34.06 \pm 0.03(\text{stat}) \pm 0.62(\text{cal})$	$0.51 \pm 0.01(\text{stat}) \pm 0.04(\text{cal})$

Stable performances in Run2 wrt Run1

Overall Run2 OS performances

- **Performances:** $B^0 \to D^- \pi^+$ data (2016)



Tagger	ϵ (%)	ω (%)	$\epsilon\langle D^2\rangle = \epsilon\,(1-2\omega)^2$ (%)
OS Combination	40.15 ± 0.09	$35.12 \pm 0.03(\text{stat}) \pm 0.21(\text{cal})$	$3.56 \pm 0.01(\text{stat}) \pm 0.10(\text{cal})$

Overall Run2 OS taggers performances similar to Run1 level