

# Exploring CP violation in heavy flavor physics: recent results and future prospects

Alberto Bragagnolo (CERN), on behalf of CMS, ATLAS and LHCb

LFC24, SISSA (Italy) 18 September 2024

# Introduction

### Why we need CP violation?

- Baryon asymmetry remains one of the great mysteries of modern physics •
- Half a century ago, Andrei **Sakharov** proposed three necessary **conditions** for a baryon-generating . process:
  - 1. Baryon number violation
  - 2. C and **CP violation (CPV)**
  - 3. Non thermal equilibrium
- In the Standard Model (SM) CP is conserved by the strong and electromagnetic interactions, • but it is violated by the weak force
  - CPV was first observed in 1964 by Fitch and Cronin using neutral kaons 🥮 Ο



- P violation was proposed by Lee and Yang (1956) 🧼 and experimentally observed by Wu (1957)
- CPV is allowed in the SM, but the amount is **insufficient** to account for the observed baryon asymmetry of the universe
  - Sources of CPV beyond the SM have to exist 0
  - CPV observables are often precisely predicted and very Ο sensitive to new physics



## **CP violation in the SM**

- In the SM quark transitions are possible through flavor-changing weak interactions
- Information about the strength of the transition is contained in the Cabibbo-Kobayashi-Maskawa (CKM) matrix
  - Parameters: 3 angles + 1 complex phase
- The single complex phase allows for CP violation
- In the SM, the CKM matrix is unitary
  - Unitary conditions can be represented by "unitary triangles"









Alberto Bragagnolo (CERN)

CP violation with heavy flavor physics

## **Types of CP violation**

- Observable CP violation in weak interaction can be classified into three different types
  - 1. Direct CPV in decays
    - Observed in K, B, and D mesons
  - 2. Indirect CPV in mixing
    - $\circ$  Observed in K<sup>0</sup> oscillations
  - 3. CPV in decay/mixing interference
    - $\circ$  ~ Observed in K^0, B^0 and B\_{\_{\rm S}} mesons

 $Pr(M \to f) \neq Pr(\overline{M} \to \overline{f})$ 

$$Pr(M^{0} \to \overline{M}^{0}) \neq Pr(\overline{M}^{0} \to M^{0})$$
  
 $Pr(M^{0}_{( \sim \overline{M}^{0})} \to f_{CP}) \neq Pr(\overline{M}^{0}_{( \sim M^{0})} \to f_{CP})$ 

• Defining  $A_f$  the  $M \rightarrow f$  amplitude, the CPV information can be coded in the **rephasing invariant** complex parameter  $\lambda$ , with  $|M_{L,H}^0\rangle = p|M^0\rangle \pm q|\overline{M}^0\rangle$   $(|q|^2 + |p|^2 = 1)$ 

$$\label{eq:lambda} \boxed{\boldsymbol{\lambda} \equiv \frac{q}{p} \overline{\underline{A}}_{\overline{f}}}_{A_{f}} \begin{bmatrix} \left| \overline{A}_{\overline{f}} / A_{f} \right| \neq 1 & \rightarrow \text{ direct CPV} \\ \left| q / p \right| \neq 1 & \rightarrow \text{ indirect CPV} \\ \left| \lambda \right| = 1, \ \text{Im}(\boldsymbol{\lambda}) \neq 0 & \rightarrow \text{ interference CPV} \end{bmatrix}$$



# **CP violation in the charm sector**

### **Overview of charm CPV**

- First observed in 2019 by LHCb [ref]
  - CPV differences between two modes
- **Complicated SM prediction** due to strong interaction effects
- Expected to be small
  - CKM and GIM suppressed
- Single mode evidence only in D<sup>0</sup> → π<sup>+</sup>π<sup>-</sup> decays [ref]
- Goal: discover CPV in charm mixing



Fig from: [PRL131(2023)091802]

# [CMS] CPV in $D^0 \rightarrow K_s K_s$

• Measuring

$$\mathcal{A}_{CP} = \frac{\Gamma(D^0 \to K^0_S K^0_S) - \Gamma(\overline{D}^0 \to K^0_S K^0_S)}{\Gamma(D^0 \to K^0_S K^0_S) + \Gamma(\overline{D}^0 \to K^0_S K^0_S)}$$

- Observation of a significant CPV → hints of BSM physics
  - From theory, CPV could be as large as O(1%) [ref]
- **Strategy**: measure  $\Delta A_{CP} = A_{CP}(D^0 \rightarrow K_S K_S) A_{CP}(D^0 \rightarrow K_S \pi^+\pi^-)$ 
  - Detector and production asymmetries canceled
  - CPV in D<sup>0</sup>  $\rightarrow$  K<sub>s</sub>  $\pi^+\pi^-$  already measured consistent with 0 [ref]
- **Results:**  $\Delta A_{CP} = 6.3 \pm 3.0 \text{ (stat)} \pm 0.2 \text{ (syst)} \%$
- Using the world-average value of  $A_{CP}(K_s \pi^+\pi^-) = (-0.1 \pm 0.8)\%$ ,  $A_{CP}(K_s K_s)$  is found to be

 ${\cal A}_{CP}(D^0 o K^0_S\,K^0_S)$  = 6.2  $\pm$  3.0 (stat)  $\pm$  0.2 (syst)  $\pm$  0.8( ${\cal A}_{CP}(K^0_S\,\pi^+\pi^-)$ ) %

- Consistent with
  - $\circ$  no CP violation at  $2\sigma$ ,
  - $\circ$  LHCb (-3.1 ± 1.3)% at 2.7  $\sigma$  [ref]
  - $\circ$  Belle ( 0.0 ± 1.5)% at 1.8  $\sigma$  [ref]



 $D^0$ 



### **Technical interlude: the CMS B parking dataset**

- Designed to allow CMS to perform B physics measurements on difficult/impossible to trigger final states (e.g. fully hadronic final states)
- Achieved with a set of **single muon triggers** (tags) with different thresholds in  $p_T$  and impact parameter
  - Luminosity decreases during a run → less restrictive triggers enabled
    - Maximises the available trigger bandwidth
  - Events are *parked* for later reconstruction
  - Very high purity of ~80%
- No impact on the *standard* CMS physics programme
- **10 billion unbiased B hadron** decays collected in 2018  $(L_{int} \sim 41 \text{ fb}^{-1})$

CMS-EXO-23-007] [CMS-BPH-22-005]

DC

Wrong-sign

**Right-sign** 

## [LHCb] Charm mixing CPV

- Never observed
- $D^0 \rightarrow K^+\pi^-$ : discovery channel for charm mixing
- Fit the time-dependent ratios
  - Probe for: CPV in mixing, SU(3)<sub>F</sub> breaking, rescattering effects and BSM



## [LHCb] Charm CPV in decay/mixing interference

- CPV arising from the **interference** between decay with and without mixing to a CP finalstate
- Search in the Cabibbo-suppressed decay
   D<sup>0</sup> → π<sup>+</sup>π<sup>-</sup>π<sup>0</sup> (CP eigenstate)
  - Flavor tagged using D\*(2010)<sup>+</sup>  $\rightarrow$  D<sup>0</sup>  $\pi^+$
- Results
  - No evidence of CP violation

$$\Delta Y = (-1.3 \pm 6.3 \pm 2.4) \times 10^{-4}.$$



# **CP violation in B decays**

### **Overview of beauty CPV**

- Extensively studied at beauty factories and LHC
  - Observed in decays and decay/mixing interference
  - Not observed in mixing
- Clean SM predictions (usually)
- Strong test for the Standard Model
  - **High sensitivity** to new physics contribution
- Can be very large (e.g. in  $B^0 \rightarrow J/\psi K_s$ )



Deep link between CPV observables and the CKM matrix

#### [LHCb] CPV in $B^0 \rightarrow D^0 K^{*0}$ decays eiv B<sup>0</sup> Direct measurement of the CKM angle x via the study of the CP D<sup>0</sup> asymmetries Use D decays to (14 samples fitted) CP-eigenstates ( $\pi^+\pi^-$ , K<sup>+</sup>K<sup>-</sup>, 4 $\pi^\pm$ ) 0 via Cabibbo-favored ( $D^0 \rightarrow K^-\pi^+$ ) and doubly 0 eiy Cabibbo-suppressed $(D^0 \rightarrow K^- \pi^+)$ **Results:** significant evidence D<sup>0</sup> of CPV was detected 60% improvement w.r.t. 1-CL Ο Candidates/(10 MeV/c<sup>2</sup> 00 00 00 00 00 LHCb - Data LHCb previous results $B^0 \rightarrow D^* K^{*0}$ 9 fb<sup>-1</sup> 0.8 9 fb<sup>-1</sup> $\rightarrow D^* \overline{K}^*$ Four solutions for y are found, world-average $\rightarrow D \pi^+\pi^$ value 0.6 $\rightarrow D \pi^+ \pi^$ one compatible with the world- $B^+ \rightarrow D K^+ \pi^- \pi^+$ Comb average: $\gamma = (62 \pm 8)^{\circ}$ 0.4 $B^0 \to D K^{*0}$ 68.3% $\rightarrow$ Charmless 200 $B_s^0 \to D \overline{K}^{*0}$ 0.2 $B_5^0 \rightarrow \text{Charmless}$ 100 95.5% Total 50 100 150 0 5000 5200 5400 5600 5800 γ [°] $M(DK^{-}\pi^{+})$ [MeV/ $c^{2}$ ]

## [CMS] CPV in B<sub>s</sub> decay/mixing interference

- Flagship CPV analysis at LHC
  - ATLAS, LHCb and CMS in the game
- Motivations
  - Precise SM prediction:  $\phi_s \approx -2\beta_s = -37 \pm 1 \text{ mrad}$
  - Very sensitive to new physics
- Precise characterization of the B<sub>s</sub> → J/ψ φ decay (mixing, CPV, B<sub>s</sub> system properties)
  - Time- and flavor-dependent angular analysis
- Innovative flavor tagging algorithm based on machine learning
  - Performance improved by 400%
- (First) evidence at 3σ of CPV in B<sub>s</sub> → J/ψ φ decays (when combined with Run1 results)

 $\phi_{s}$  =  $-74 \pm 23$  [mrad]  $\Delta\Gamma_{s}$  = 0.0780  $\pm$  0.0045 [ps $^{-1}$ ]



$$\Gamma\left(B^{0}_{S_{(***)\overline{B}^{0}_{S})}} \to f\right)(t) \stackrel{?}{\neq} \Gamma\left(\overline{B}^{0}_{S_{(***)B^{0}_{S}}} \to f\right)(t)$$



## **Technical interlude: CMS inclusive flavor tagging**

- A cutting-edge flavor tagging framework has been engineered for flavor-sensitive measurements to extract the best possible results from data
- Four DNN-based algorithms are used, divided into two main classes
  - **Same side (SS):** exploits the B<sub>s</sub> fragmentation
    - SS tagger: leverages charge asymmetries in the B<sub>s</sub> fragmentation
  - Opposite side (OS): exploits decay products of the other b-hadron in the event
    - **2. OS muon**: leverages  $b \rightarrow \mu^{-}X$  decays
    - **3. OS electron**: leverages  $b \rightarrow e^{-X}$  decays
    - 4. OS jet: capitalizes on charge asymmetries in the OS *b*-jet
- The combined flavor tagging framework achieves a tagging power of P<sub>tag</sub> = 5.6% when applied to the B<sub>s</sub> data sample
  - Among the highest ever recorded at LHC

#### Schematic representation of a generic $B_s \rightarrow J/\psi \phi$ event



# **Future prospects at HL-LHC**

### Flavor tagging in CMS Phase-2 with timing

- The planned MTD (Mip Timing Detector) will provide CMS Phase-2 with the time information of charged tracks
- The reconstruction algorithm utilizes compatible times of tracks from a vertex to offer time-of-flight-based particle identification (PID) as a natural byproduct
- Same-side tagging could utilize charge correlation between the light quark in the B meson and a nearby soft hadron for flavor tagging
- The PID from MTD, when integrated into the Phase-2 extrapolation of this analysis, shows a significant improvement in the tagging performance



Relative	gain	in I	P <sub>tan</sub>	(only SS)	
----------	------	------	------------------	-----------	--

PID scenario	Gains in P <sub>tag</sub>
MC truth (perfect PID < 3 GeV)	+66%
PID with $\sigma_{BTL}$ = 40 ps	+24%
PID with $\sigma_{BTL}$ = 70 ps	+14%

## Projecting $\phi_s$ to HL-LHC

- **Current status:** with an uncertainty of  $\sigma(\phi_s) = 14 \text{ mrad}$ the world average is consistent with SM-based predictions, though new physics contributions of ~10% are still possible
- HL-LHC will enable an exceptionally precise measurement of φ<sub>s</sub>, opening new avenues for probing physics beyond the Standard Model
- The planned detector improvements in all three experiments are critical to deliver these requirements despite the much higher HL-LHC pileup
- HL-LHC experiments aim for a combined  $\sigma(\phi_s) \sim 2$ mrad and  $\sigma(\Delta\Gamma_s) \sim 5 \text{ ps}^{-1}$  in
  - Individual experiments:  $\sigma(\phi_s) \sim 3-5$  mrad
- **Challenges**: precise control over hadronic effects and penguin pollution (more on this later)



## **Projecting sin(2β) to HL-LHC**

- **Current status:** with a precision of 0.011, the world-average value is in excellent agreement with SM-based predictions
- After Upgrade-II LHCb will be able to reach a statistical precision below 0.003, while Belle-II with 50  $ab^{-1}$  projects  $\sigma(sin2\beta) \sim 0.005$ 
  - CMS has expressed the intention of pursuing sin(2β)
- Systematic uncertainties will need to be controlled, especially from CP violation in  $K^0 - \overline{K}^0$ mixing and nuclear cross-section differences.
- Controlling the penguin effects will be essential to derive insightful boundaries on new physics effects





 Penguin pollutions are expected to be small but they are not well constrained and will be the leading source of systematic at HL-LC

$$egin{array}{lll} \Delta \phi^{m{penguin}}_{m{s}} pprox & 3\pm 10 \ mrad & vs & \sigma(\phi_{m{s}}) pprox 15 \ mrad \ egin{array}{lll} \Delta \phi^{m{penguin}}_{m{d}} pprox -10\pm 10 \ mrad & vs & \sigma(2eta) pprox 15 \ mrad \end{array}$$

• Analysis of penguin and NP contributions is possible using Cabibbo-favored control channels



CP violation with heavy flavor physics

### **Penguins at HL-LHC**

Φ<sub>s</sub>

• The best constraint on penguin pollution in  $\phi_s$  comes from B<sup>0</sup>  $\rightarrow$  J/ $\psi \rho^0$ 

$$\Delta \phi_s^{c\bar{c}s} \approx -\epsilon \left( \phi_d^{J/\psi \rho^0} - 2\beta \right) \qquad \epsilon = \left| V_{us} \right|^2 / (1 - \left| V_{us} \right|^2) = 0.0534$$

• The LHCb collaboration foresee to be able to measure  $\phi_d^{J/\psi \rho 0}$  with a precision  $\leq 1^\circ$ , corresponding to an uncertainty on the penguin pollution  $\leq 1.5$  mrad (reminder:  $\sigma(\phi_s) \sim 2$  mrad at HL-LHC)

### sin(2β)

- A first analysis of  $B_s \rightarrow J/\psi K_s$  has been performed [ref] as a proof of concept for constraining  $\Delta \phi_d$
- However, the penguin pollutants will be mostly kept in check by Belle-II using  $B^0 \rightarrow J/\psi \pi^0$  decays

## **Other CPV opportunities**

### B<sub>s</sub> → φφ → KKKK

- Penguin dominated decay
- Benchmark for CPV in FCNC
- **CMS** Phase-2 will be able to access tracking information at the first stages of the trigger process, allowing the direct trigger of fully hadronic final states
  - For  $B_s \rightarrow \phi \phi$  a 30~35% efficiency at L1 trigger is expected
- **LHCb** Upgrade II can achieve a statistical uncertainty on  $\phi_s^{sss}$  of 11 mrad with 300 fb<sup>-1</sup> of data collected

#### Other charmless decays

- At HL-LHC LHCb can also access
  - $\phi_s^{dds}$  via  $B_s \rightarrow (K^+\pi^-)(K^-\pi^+)$ , with a precision of 9 mrad
  - $\phi_s^{duu}$  via  $B_s \rightarrow K_s \pi^+ \pi^-$ , with a precision of 70 mrad







## **Summary and outlook**

- CP violation has been an extremely successful field in recent years and has a clear road for the upcoming colliders
- The upgrade to HL-LHC offers promising opportunities to push the boundaries of CPV measurements
  - But penguins are always lurking
- The future is not only a simple *luminosity scaling*: detector upgrades, additional possibilities in triggering and new analysis techniques are going to be a key part of future collider programmes
  - Thallenge will be to continue refining our experimental approaches while controlling systematic uncertainties
- With these advancements, the study of CP violation will remain at the forefront of probing fundamental physics

Stay tuned in the future for other exciting results on CPV!

# Thanks for the attention



#### Example: B meson mixing

#### 





#### Ref: LHCb Collab. Nat.Phys.18(2022)1-5

### **Clavor mixing**

- Neutral K, D, and B mesons are subject to flavor mixing, that is oscillations between their C-conjugate states before decay
- They propagate as light and heavy mass eigenstates which are described by a superposition of flavor states:

$$|M_{L,H}^{0}\rangle = p|M^{0}\rangle \pm q|\overline{M}^{0}\rangle$$
 ( $|q|^{2} + |p|^{2} = 1$ )

• The system is characterized by the parameters

- The flavor eigenstates oscillate with period T =  $2\pi/\Delta m$
- **CPV** in mixing, i.e.  $Pr(M^0 \to \overline{M}^0) \neq Pr(\overline{M}^0 \to M^0)$ , implies  $|q/p| \neq 1$

# $D^0 \rightarrow K_s K_s$ measurement strategy

- Use D<sup>0</sup> from D<sup>\*+</sup>  $\rightarrow$  D<sup>0</sup>  $\pi^+$  and D<sup>\*-</sup>  $\rightarrow$  D<sup>0</sup>  $\pi^-$ , so that the pion charge tags the D<sup>0</sup> flavor
- This introduces the D\*+/D\*- differences in the measurement



- Strategy: measure  $\Delta A_{CP} = A_{CP}(D^0 \rightarrow K_S K_S) A_{CP}(D^0 \rightarrow K_S \pi^+\pi^-)$ 
  - Reference channel is very similar in kinematics and topology  $\rightarrow A_{prod}$  and  $A_{det}$  cancel out
  - CPV in  $D^0 \rightarrow K_s \pi^+\pi^-$  already measured consistent with zero [PRD86(2012)032007]

$$\Delta A_{CP} = A_{raw}(D^0 o K_S K_S) - A_{raw}(D^0 o K_S \pi^+ \pi^-)$$

Alberto Bragagnolo (CERN)

CP violation with heavy flavor physics

Signal channel

Reference channel  $\pi^*$ 

K

# $D^0 \rightarrow K_s K_s$ event selection

- First,  $K_s \rightarrow \pi^+\pi^-$  are reconstructed fitting the  $\pi$  tracks to a common vertex
  - $|m(\pi^+\pi^-) m(K_s^{w.a.})| < 20 \text{ MeV}, p_T(K_s) > 2.2(1.0) \text{ GeV}$
- In the signal channel, two K<sub>s</sub> candidates are required and fitted to a common vertex to form D<sup>0</sup> → K<sub>s</sub>K<sub>s</sub> candidates
  - 1.7 GeV < m(K  $_{s}K_{s}$ ) < 2.0 GeV
  - $K_s$  displacement in *xyz* from the D<sup>0</sup> vertex >9 $\sigma$  and >7 $\sigma$
  - $D^{\bar{0}}$  displacement in *xyz* (*xy*) from the PV >9 $\sigma$  (>2 $\sigma$ )

## • In the **reference channel**, two track with $p_T > 0.6$ GeV are used to form the $D^0 \rightarrow K_c \pi^+ \pi^-$ candidate

- ο 1.823 < m(K <sub>s</sub>π⁺π⁻) < 1.908 GeV
- **Finally**, an additional track with  $-1.2 < |\eta| < 1.2$  and  $p_T > 0.36$  GeV is added to form  $D^{*+} \rightarrow D^0 \pi^+$  candidates
  - $\circ \qquad m(D^0\,\pi^{\scriptscriptstyle +}) = m(D^0\pi^{\scriptscriptstyle +}) m(D^0) + m_{_{PDG}}(D^0)$
- Background suppression: several fits corresponding to incorrect topologies are performed and vertex probabilities requirements are imposed





Alberto Bragagnolo (CERN)

CP violation with heavy flavor physics

# $\phi_s$ motivations

- B<sub>s</sub> mesons decays allow us to study the time-dependent CP violation generated by the interference between direct decays and flavor mixing
  - CPV in the interference is possible even if there is no CPV in decay and mixing
- The weak phase  $\phi_s$  is the main CPV observable
  - Predicted by the SM to be  $\phi_s \approx -2\beta_s$  ( $\beta_s \rightarrow$  angle of the B<sub>s</sub> unit. triangle)
    - Neglecting contributions from higher-order diagrams ( $\Delta \phi_s^{\text{loop}} \approx 3 \pm 10 \text{ mrad}$ )
  - $\beta_s$  determined by CKM global fits to be -2 $\beta_s$  = -37 ± 1 mrad [CKMfitter, UTfit]
- New physics can change the value of  $\phi_s$  up to ~100% via new particles contributing to the flavor oscillations [RMP88(2016)045002]



• This seminar presents the latest CMS results with the *golden* channel  $B_s \rightarrow J/\psi \phi(1020) \rightarrow \mu^+\mu^- K^+K^-$ 





## $\phi_s$ : a time-, flavor- and angular-dependent measurement





#### **Core ingredients**

- **Time-dependent angular analysis** to separate the CP eigenstates ("transversity basis" used)
- Time-dependent flavor analysis to resolve the B<sub>s</sub> mixing oscillations (T ~ 350 fs)

sensistivity 
$$\propto \sqrt{\frac{\epsilon_{\text{tag}} D_{\text{tag}}^2 N_{\text{sig}}}{2}} \sqrt{\frac{N_{\text{sig}}}{N_{\text{sig}} + N_{\text{bkg}}}} e^{-\frac{\sigma_l^2 \Delta m_S^2}{2}}$$

#### Decay rate for a CP-even final state



#### Alberto Bragagnolo (CERN)



Alberto Bragagnolo (CERN)

CP violation with heavy flavor physics

## $\phi_s$ : trigger strategy

### Muon-tagging trigger

- $J/\psi \rightarrow \mu^+\mu^-$  candidate plus an additional muon (for tagging)
- ≈50 000 signal candidates
- Used for time resolution modeling
- Tagging algorithms deployed: OS-muon
  - $\circ$  P<sub>tag</sub> ~ **10%** (muon at trigger level enhance tagging efficiency)

### Standard trigger

- Displaced  $J/\psi \rightarrow \mu^+\mu^-$  candidate +  $\phi(1020) \rightarrow K^+K^-$
- ≈450 000 signal candidates
- Tagging algorithms deployed: OS-muon, OS-electron, OS-jet, Same Side



Proper decay length uncertainty distribution for the standard trigger (2018)

## $\phi_s$ : decay time and its resolution

• The time dependence of the decay rate is parametrized with the **proper decay length** ct, measured in the transverse plane as

$$ct = c \cdot rac{m_{Bs}^{w.a.} \cdot L_{xy}}{p_T}$$
 with  $L_{xy} \equiv ||\overline{r}_{xy}(SV) - \overline{r}_{xy}(PV)||$ 

- Its **uncertainty** is obtained by fully propagating the uncertainties in  $L_{xv}$  and  $p_T$ 
  - The uncertainty on  $L_{xy}$  dominates for most of the ct spectrum, with  $\sigma(p_T)$  taking over at high values (ct  $\ge$  3 mm)
- The ct uncertainty is calibrated in a prompt data sample of  $B_s \rightarrow J/\psi \phi$ , obtained by removing the displacement requirement in the *muon-tagging* data sets
  - Modeled with two gaussians to obtain the effective dilution and resolution

$$\delta_{\text{eff}} = \sqrt{\frac{-2\ln \mathcal{D}}{\Delta m_s^2}} \quad \text{with} \quad \mathcal{D} = \sum_{i=1}^2 f_i \exp\left(-\frac{\sigma_i \Delta m_s^2}{2}\right)$$

• Excellent agreement found, with corrections ~5%





#### Time resolution calibration for 2018 data

#### Alberto Bragagnolo (CERN)

ct efficiency for the standard trigger category (2018)

## $\phi_s$ : acceptance and efficiency effects

- The efficiency in selecting and reconstructing the B<sub>s</sub> candidates is not independent of the decay time and angular observables
  - To properly fit the decay rate model an efficiency parametrization is needed

#### **Time efficiency**

- Modeled in the B<sup>0</sup>  $\rightarrow$  J/ $\psi$  K<sup>\*0</sup> data control channel with corrections from simulations
- Ultimately parametrized with Bernstein's polynomials

$$\varepsilon_{B^0}^{\text{data}}(ct) = \frac{N_{B^0}(ct)}{e^{-\Gamma_d^{\text{w.a.}}} \otimes P_{B^0}(\sigma_{ct})} \qquad \varepsilon_{B_s}^{\text{data}}(ct) = \varepsilon_{B^0}^{\text{data}}(ct) \cdot \frac{\varepsilon_{B_s}^{\text{MC}}(ct)}{\varepsilon_{B^0}^{\text{MC}}(ct)}$$

#### Angular efficiency

- Estimated with KDE distributions in simulated events
- The simulated data samples are corrected to match the data
  - An iterative procedure is used to simultaneously correct the kinematics of the final state particles and the differences in the physics parameters set in the MC with respect to what measured in the data



ct (cm)



#### Alberto Bragagnolo (CERN)

#### CP violation with heavy flavor physics

### Flavor, neural networks, and probabilities

- The tagging inference logic differs between algorithms
  - Lepton taggers (OS muon, OS electron)
    - Lepton charge →  $\xi_{tag}$ ; DNN score →  $\omega_{tag}$  (DNN trained for correct-tag vs mistag) OS  $\ell^- \rightarrow OS \ b \xrightarrow{tag}} signal \ B_s$ OS  $\ell^+ \rightarrow OS \ \overline{b} \xrightarrow{tag}} signal \ \overline{B}_s$ Charge-based taggers (OS jet, SS) ■ DNN score → Prob(B<sub>s</sub>) →  $\xi_{tag}$ ,  $\omega_{tag}$   $S_{DNN} > 0.5 + \epsilon \xrightarrow{tag}} signal \ B_s$  with  $\omega_{tag} = 1 - S_{DNN}$   $s_{DNN} < 0.5 - \epsilon \xrightarrow{tag}} signal \ \overline{B}_s$  with  $\omega_{tag} = 1 - S_{DNN}$ 
      - $\epsilon$  is used to remove events with  $\omega_{tag} \sim 50\%$
- The algorithms are optimized and trained in simulated events and calibrated in data with self-tagging
   B<sup>+</sup> → J/ψ K<sup>+</sup> decays
  - $\circ$  The calibration is performed by comparing  $\omega_{_{tao}}$  predicted by the DNN and the one measured in data

Ο

## Tagging calibration strategy (and other tricks)

OS-Muon calibration nuon-tagging trigger 2018

- A multi-pronged strategy has been devised to improve the  $\omega_{tag}$  estimation and suppress systematic effects
  - 1. All models are constructed from the start as probability estimators, i.e. score~ $\omega_{tag}$ 
    - Loss function: cross-entropy, which is the likelihood for the probability P(true class | score)
    - Output layer: Sigmoid function, which normalizes the output to a probability distribution
  - 2. All DNNs are calibrated with the *Platt scaling*, which ensures that the calibrated score is still a probability
    - The Platt scaling is a linear calibration of the score before the last sigmoid layer
  - 3. In calibrating the charge-based taggers (which provide a probability for  $B_s$  vs  $\overline{B}_s$ ):
    - A. The output is symmetrized due to the initial LHC charge imbalance

$$s_{DNN}^{sym}(x) = \frac{s_{DNN}(x) + [1 - s_{DNN}(\overline{x})]}{2}$$

B. The symmetry is explicitly forced in the calibration function by removing the constant term

This strategy cancels almost all the systematic effects associated with flavor tagging





.

### OS-Muon calibration (muon-tagging trigger 2018)

## φ<sub>s</sub>: OS-lepton tagging

- OS-lepton tagging techniques search for b → ℓ<sup>-</sup>X decays of the other B hadron in the event
- The **charge** of the lepton is used as tagging feature and a fully connected DNN is used to estimate the mistag probability
- Lepton selection
  - Loose kinematic cuts
  - Separated from the signal B meson
  - MVA discriminator against fakes
  - OS-electrons are searched only if no OS-muon is found in the event (explicit orthogonality)
- Mistag estimation
  - Fully connected DNN with ReLU activation and dropout
  - Inputs: lepton kinematics and surrounding activity
- Trained on simulated B<sub>s</sub> → J/ψ φ(1020) events and calibrated in B<sup>+</sup> → J/ψ K<sup>+</sup> data



#### **OS-Electron calibration (2018)**





Alberto Bragagnolo (CERN)

CP violation with heavy flavor physics

## φ<sub>s</sub>: SS tagger

- The SS tagger consists of a DNN (*DeepSSTagger*), derived from *DeepJetCharge*, able to probe the fragmentation products of a B meson and exploit tracks with high flavor correlation
- *DeepSSTagger* uses the kinematic information from up to 20 tracks (ordered by  $|IP_z|$ ) around the reconstructed B meson
- Track selection
  - $\Delta R(trk, B) < 0.8, ||P_z(PV)| < 0.4 \text{ cm}, ||P_{xv}(PV)|/\sigma_{dxv} < 1$
  - Overlap with signal and OS is carefully avoided with geometrical cuts and vetos
- **Trained on an equal-weight mixture** of  $B_s \rightarrow J/\psi \phi$  and  $B^+ \rightarrow J/\psi K^+$  to make the model invariant for  $B_s \leftrightarrow B^+$  for calibration purposes
  - Calibration directly in B<sub>s</sub> was found to be not feasible in CMS
    - Tested:  $B_s \rightarrow D_s^- \pi^+$  (not enough stat.) and  $B_s^{**} \rightarrow B^{+(*)}K^-$  (too much uncer. from  $B^{0**}$  bkg)
  - The trained network produces the probability of signal B meson containing a negatively charged quark alongside the b quark (i.e., being a B<sub>s</sub> or B<sup>-</sup>)
- Calibration
  - The SS is calibrated  $B^+ \rightarrow J/\psi K^+$  data, with residual differences ~10% corrected with simulations
  - $\circ$  Events with  $\omega_{tag} > 0.46$  are removed before the calibration and assumed untagged



Comparison between Same-side tagger  $B^+$  and  $B_s$  calibrations (2018)



#### Largest ever effective statistics $N_{Rs} \cdot P_{tag}$ (490k $\cdot$ 5.6% $\approx$ 27.5k) for a single $\phi_{s}$ measurement

The flavor tagging framework is validated in the .  $B^0 \rightarrow J/\psi K^{*0}$  data control channel with flavor mixing measurements, both integrated and time-dependent

#### Alberto Bragagnolo (CERN)

## φ<sub>s</sub>: flavor tagging performance

- The SS and any one of the OS algorithms overlap in about 20% of the events
  - 0 In these cases, the information is combined to improve the tagging inference
- The combined flavor tagging framework achieves a tagging power of  $P_{tag} = 5.6\%$  when applied to the B data sample
  - Among the highest ever recorded at LHC Ο
  - x3~4 improvement with respect to prev. CMS results Ο
- This is the first CMS implementation of the OS jet and same-side tagging techniques
  - SS accounts for half of the performance 0

59.8 fb<sup>-1</sup> (13 TeV) 59.8 fb<sup>-1</sup> (13 TeV) Data CMS 82000 CMS 0.02 - Full fit O 1800 Preliminary Preliminary Signal · Comb. bkg - 1600 Data B<sup>0</sup> bkg 1400 1200 Full fit Events 1000 800 20



 $\omega_{tac}$ 

10

0.2

0.3

0.4

0.5

 $\omega_{taq}$ 

Category	$\varepsilon_{\rm tag}$ [%]	$\mathcal{D}_{\mathrm{eff}}^2$	$P_{\text{tag}}$ [%]
Only OS muon	$6.07\pm0.05$	0.212	$1.29\pm0.07$
Only OS electron	$2.72\pm0.02$	0.079	$0.214\pm0.004$
Only OS jet	$5.16\pm0.03$	0.045	$0.235\pm0.003$
Only SS	$33.12\pm0.07$	0.080	$2.64\pm0.01$
SS + OS muon	$0.62\pm0.01$	0.202	$0.125\pm0.003$
SS + OS electron	$2.77\pm0.02$	0.150	$0.416\pm0.005$
SS + OS jet	$5.40\pm0.03$	0.124	$0.671\pm0.006$
Total	$55.9\pm0.1$	0.100	$5.59 \pm 0.02$



#### CP violation with heavy flavor physics

600

400

200

0.1

0.2

## $\phi_s$ : tagging validation with B<sup>0</sup> events

- The flavor tagging framework is validated in the B<sup>0</sup> → J/ψ K<sup>\*0</sup> control channel (~2M events)
- The time-dependent **mixing asymmetry** is measured to extract the flavor mixing oscillation frequency  $\Delta m_d$  with a precision of ~1% (comparable with BaBar and Belle)
  - Excellent agreement with world-averages is observed

#### → No bias in mixing frequency measurements

- Study performed also in each tagging category (see backup)
- The **time-integrated mixing** is also measured for each tagger and their dependency on the expected tagging dilution is compared
  - The dependency between the measured A<sub>mix</sub> and the estimated D<sub>tag</sub> is found to be well described by a linear relationship, indicating that all four techniques behave in the same predictable way

#### B<sup>0</sup> flavor mixing asymmetry





# $\phi_s$ : fit model

- The physics parameters are extracted with unbinned multidimensional extended maximum-likelihood (UML) fit performed simultaneously on 12 data sets (2 trig. cat. x 2 years x 3 ξ<sub>tag</sub> values)
  - $\circ \quad \textit{Physics parameters: } \varphi_{s}, |\lambda|, \Delta \Gamma_{s}, \Gamma_{s}, \Delta m_{s}, |A_{0}|^{2}, |A_{\perp}|^{2}, |A_{S}|^{2}, \delta_{\parallel}, \delta_{\perp}^{-}, \delta_{S\perp}$
  - $\circ \quad \textit{Observables: } m_{Bs}^{}, \, \text{ct}, \, \sigma_{ct}^{}, \, \cos\theta_{T}^{}, \, \cos\psi_{T}^{}, \, \phi_{T}^{}, \, \omega_{tag}^{}$
- Fit model

 $P = [f_{sig}P_{sig}] + [f_{bkg}P_{bkg}] + [f_{bkg}B_0 P_{bkg}B_0]$   $SIGNAL P_{sig} = \varepsilon(\Theta) [\tilde{f}(\Theta, ct \mid \alpha, \xi_{tag}, \omega_{tag}) \otimes G(ct, \sigma_{ct})] P_{sig}(m_{B_s}) P_{sig}(\sigma_{ct}) P_{sig}(\omega_{tag})$ 

$$\begin{array}{c} \textbf{COMBINATORIAL BKG} \quad P_{bkg} = \left[ P_{bkg}(ct) \otimes G(ct, \sigma_{ct}) \right] P_{bkg}(\Theta) P_{bkg}(m_{B_s}) P_{bkg}(\sigma_{ct}) P_{bkg}(\omega_{tag}) \end{array}$$

- The time efficiency is implemented as a *re-weighting* of the data events to drastically improve fit time
- The statistical uncertainties and fit bias are estimated with **1300** bootstrap distributions
- The yield for the  $B^0 \rightarrow J/\psi K^{*0}$  is estimated directly in data with a 2D fit to the  $B_s$  invariant mass and its  $B^0$  reflection
- The background from  $\Lambda_b \rightarrow J/\psi \ K^-p^+$  is found to be negligible and is treated as a systematic uncertainty

Alberto Bragagnolo (CERN)

## $\phi_s$ : systematic uncertainty overview

	$\phi_s$	$\Delta \Gamma_s$	$\Gamma_s$	$\Delta m_s$	$ \lambda $	$ A_0 ^2$	$ A_{\perp} ^2$	$ A_{\rm S} ^2$	$\delta_{\parallel}$	δ	8SL
	[mrad]	[ps <sup>-1</sup> ]	[ps <sup>-1</sup> ]	[ħps <sup>-1</sup> ]					[rad]	[rad]	[rad]
Statistical uncertainty	23	0.0043	0.0015	0.035	0.014	0.0016	0.0021	0.0033	0.074	0.089	0.15
Model bias	4	0.0011	0.0002	0.004	0.006	0.0012	0.0022	0.0006	0.015	0.017	0.03
Flavor tagging	4	$< 10^{-4}$	0.0005	0.007	0.002	$< 10^{-4}$	$< 10^{-4}$	0.0006	0.012	0.016	0.03
Angular efficiency	4	0.0002	$< 10^{-4}$	0.015	0.011	0.0042	0.0019	0.0001	0.017	0.044	0.02
Time efficiency	< 1	0.0014	0.0026	$< 10^{-3}$	$< 10^{-3}$	0.0004	0.0005	$< 10^{-4}$	0.001	0.002	$< 10^{-2}$
Time resolution	< 1	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-3}$	0.001	$< 10^{-3}$
Model assumptions		0.0005	0.0006		( <u></u> 2)	( <u></u> 4)	2 <u></u>			<u> </u>	<u> </u>
B <sup>0</sup> background	< 1	0.0002	0.0003	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-2}$
$\Lambda_{\rm b}^0$ background			0.0004	n <u></u> n	1 <u>7</u> 3	0.0004	0.0003	<u> </u>	<u> 19</u> 17 -	<u> 19</u> 17 -	<u> </u>
S-P wave interference	< 1	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-2}$
$P(\sigma_{ct})$ uncertainty	< 1	0.0002	0.0003	$< 10^{-3}$	$< 10^{-3}$	0.0001	0.0001	$< 10^{-4}$	$< 10^{-3}$	$< 10^{-3}$	$< 10^{-2}$
Total systematic uncertainty	7	0.0019	0.0028	0.017	0.012	0.0044	0.0030	0.0009	0.025	0.050	0.05

- Model bias, flavor tagging, and angular efficiency are found to be the leading systematic sources for  $\varphi_s$
- The measurement is still heavily statistically limited for  $\phi_s$



#### Fit results

Parameter	Fit value	Stat. uncer.	Syst. uncer.
$\phi_s$ [mrad ]	-73	$\pm 23$	$\pm 7$
$\Delta\Gamma_s$ [ps <sup>-1</sup> ]	0.0761	$\pm 0.0043$	$\pm 0.0019$
$\Gamma_s [ps^{-1}]$	0.6613	$\pm 0.0015$	$\pm 0.0028$
$\Delta m_s [\hbar \mathrm{ps}^{-1}]$	17.757	$\pm 0.035$	$\pm 0.017$
$ \lambda $	1.011	$\pm 0.014$	$\pm 0.012$
$ A_0 ^2$	0.5300	$\pm 0.0016$	$\pm 0.0044$
$ A_{ } ^{2}$	0.2409	$\pm 0.0021$	$\pm 0.0030$
$ A_{\rm S} ^2$	0.0067	$\pm 0.0033$	$\pm 0.0009$
$\delta_{\parallel}$	3.145	$\pm 0.074$	$\pm 0.025$
$\delta'_{\perp}$	2.931	$\pm 0.089$	$\pm 0.050$
$\delta_{S\perp}$	0.48	$\pm 0.15$	$\pm 0.05$

•  $\phi_s$  and  $\Delta \Gamma_s$  are found in agreement with the SM

 $\phi_s^{SM}\simeq -37\pm 1~{
m mrad}~~\Delta\Gamma_s^{SM}$  = 0.091  $\pm$  0.013 ps<sup>-1</sup>

 $\Gamma_{s}$  and  $\Delta m_{s}$  are consistent with the latest world averages

 $\Gamma_s^{W\!A} = 0.6573 \pm 0.0023 \text{ ps}^{-1}$   $\Delta m_s^{W\!A} = 17.765 \pm 0.006 \text{ }\hbar \text{ps}^{-1}$ 

•  $|\lambda|$  is consistent with no direct CPV ( $|\lambda| = 1$ )

• This measurement utilizes the largest ever effective statistics

- $N_{Bs}^{\phantom{Bs}\cdot}$   $P_{tag}^{\phantom{Bs}}$  for a single  $\varphi_s^{\phantom{Bs}}$  measurement
  - The precision on  $\phi_s$  is comparable with the world's most precise single measurement by LHCb ( $\phi_s = -39 \pm 22$  (stat)  $\pm 6$  (syst) mrad) [PRL132(2024)051802]
  - $\circ$   $\;$  This is the most precise single measurement of  $\Delta\Gamma_{s}$  to date in this channel