Probing fundamental properties of nature with Heavy Quark Physics at CMS

Alberto Bragagnolo, on behalf of the CMS Collaboration

58th Rencontres de Moriond, La Thuile (Italy)
24-31 March 2024

Introduction

- **Studying heavy quarks is crucial in experimental particle physics for**
    - Testing the Standard Model
    - Probing fundamental interactions
    - Searching for New Physics
- **This talk unveils two new CMS results on the physics of top and bottom quarks**
    - Measurement of top quark entanglement
    - Precision measurement of CP violation in B_s meson decays

Probing Top Quark Entanglement

CMS PAPER TOP-23-001

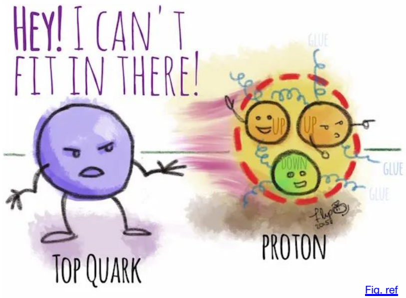
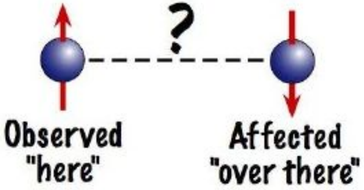


Fig. ref

Entanglement at the LHC



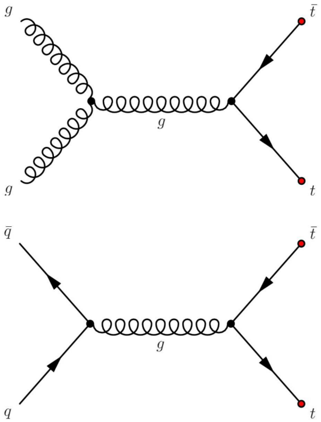
- **Fundamental predictions of QM:**
  - Entangled states cannot be described by independent superpositions: measuring particle spin in an entangled system immediately reveals the spin state of the second particle
    - Nobel Prize in 2022 for Aspect, Clauser, and Zeilinger

- **First observation of entanglement in $t\bar{t}$ by ATLAS end of last year** (see S. Wuchterl talk)
  - Short top quark lifetime of 10^{-25} s allows measuring polarization (B) and spin correlation (C) in $t\bar{t}$ production ➜ explore to proof entanglement at the LHC
  - Spin density operator ρ is fairly complex, but one can find the necessary conditions to show that it is **non-separable ➜ entangled** [EPJP(2021)136:907]

$$\rho = \frac{I_4 + \sum_i (B_i^+ \sigma^i \otimes I_2 + B_i^- I_2 \otimes \sigma^i) + \sum_{i,j} C_{ij}\, \sigma^i \otimes \sigma^j}{4} \xrightarrow{\text{not separable if}} D \equiv -\frac{tr[\mathbf{C}]}{3} < -\frac{1}{3}$$

- **Experimental goal ➜ measure D** (entanglement proxy)

How to produce a top quark



How to measure the entanglement proxy *D*

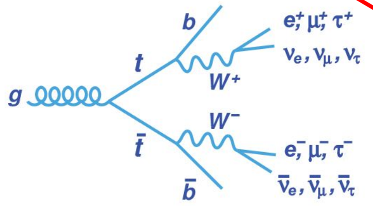
- **Analysis strategy: use leptonic final states to measure the helicity angle *cos φ* ≡ $\ell_1 \cdot \ell_2$** (measured in the top reference frame)
    - Experimentally well measured
    - Fully encapsulates the spin correlation information for gg fusion production at low mass

$$\frac{1}{\sigma}\frac{d\sigma}{d\cos\varphi} = \frac{1}{2}(1 - D\cos\varphi)$$

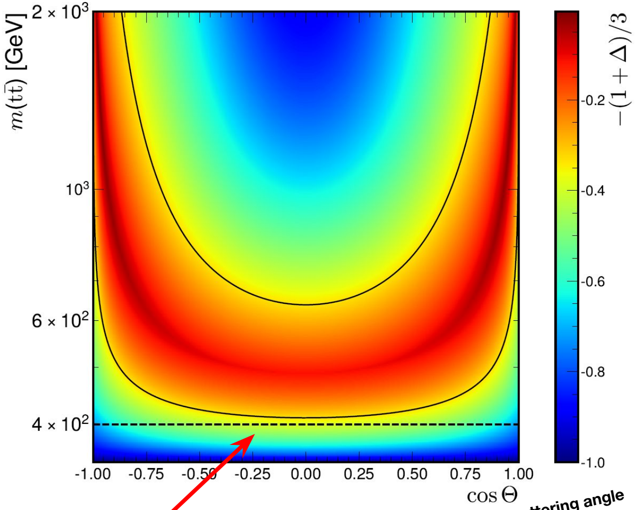
    - Resilient to systematic effects
- The degree of entanglement is highly phase-dependent
- **Focus on low-mass region** ($345 < m_{tt} < 400$ GeV)
    - Dominated by gg
    - Increased entanglement
- Cut on velocity along the beam line of the $t\bar{t}$ system to increase gg/qq fraction

$$\beta = \left|\frac{p_Z^t + p_Z^{\bar{t}}}{E^t + E^{\bar{t}}}\right| < 0.9$$

Leptons from τ decays included



Degrees of entanglement for gg fusion production

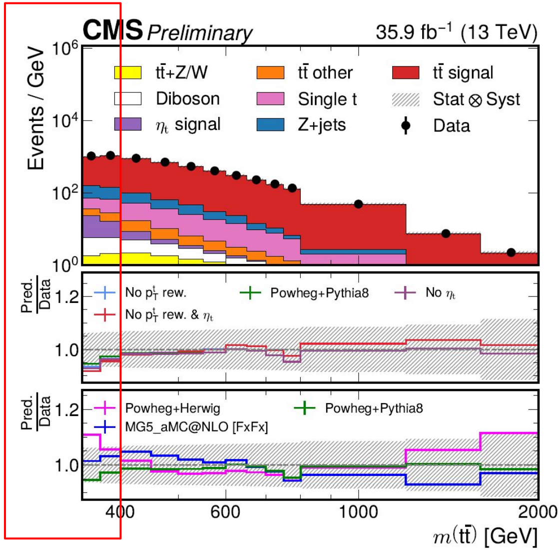


Analysis phase-space

Top quark scattering angle

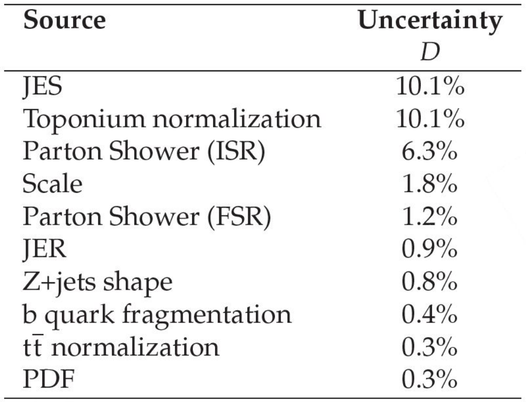
Signal model and uncertainties

- **Combined signal model**: $t\bar{t}$ + toponium (η_t)
  - Only spin-0 η_t accounted (colour singlet pseudo-scalar state) [PRD104(2021)034023]
  - η_t improves data modeling in the threshold region
  - 47 500 signal candidates in 35.9 fb^{-1} collected in 2016
- **Main background sources**: Z+jet (MG5_aMC@NLO + data-driven corrections), single top (Powheg MC), diboson (Pythia8 MC)
- **Leading experimental uncertainties**
  - Jet energy scale and resolution
- **Leading theory-based uncertainties**
  - Toponium normalization
  - Parton Shower



Analysis region

Leading systematic uncertainties



Source	Uncertainty D
JES	10.1%
Toponium normalization	10.1%
Parton Shower (ISR)	6.3%
Scale	1.8%
Parton Shower (FSR)	1.2%
JER	0.9%
Z+jets shape	0.8%
b quark fragmentation	0.4%
$t\bar{t}$ normalization	0.3%
PDF	0.3%

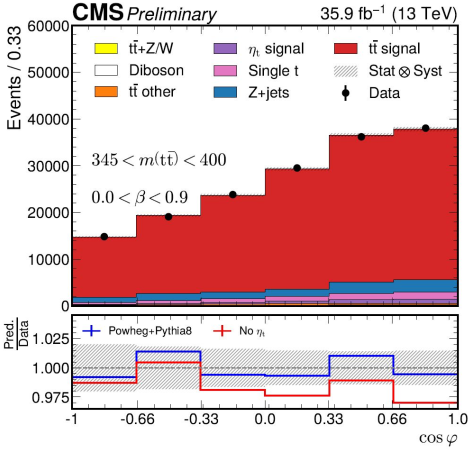
Entanglement measurement

- **The entanglement proxy *D* is extracted with a template fit**
  - All systematic effects included as nuisances
- **How to create variations of *D* outside of SM?**
  1. Generate top pairs with no spin correlation (noSC, D = 0)
  2. Created new samples with mixture of SM and noSC to obtain D $\in [D_{SM}, 0]$
  3. Extend the fit for variations of $[-1, D_{SM}]$

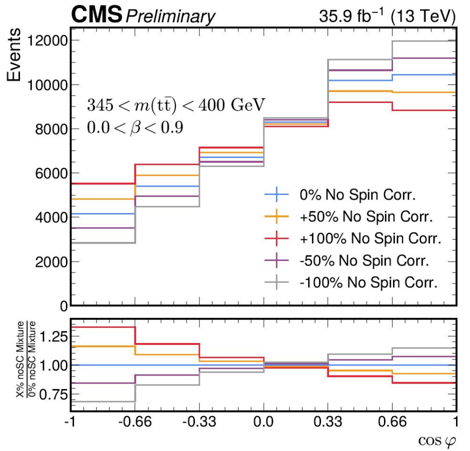
- **Use samples of SC and noSC to change fraction of $t\bar{t}$ with aligned vs opposite spins ➜ any value of D between -1 and +1 can be reached**

$$\boxed{D \sim \frac{\sigma(\uparrow\uparrow)+\sigma(\downarrow\downarrow)-\sigma(\uparrow\downarrow)-\sigma(\downarrow\uparrow)}{\sigma(\uparrow\uparrow)+\sigma(\downarrow\downarrow)+\sigma(\uparrow\downarrow)+\sigma(\downarrow\uparrow)}}$$

Post-fit cos φ distribution



Mixture templates



Results

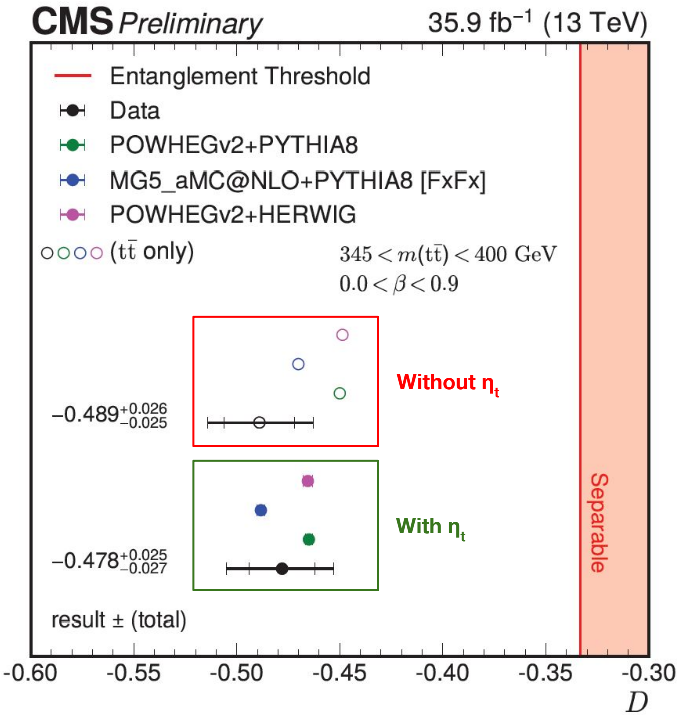
- Fit yields D at parton level, accounting for all detector effects

$$\boxed{D_{\text{obs}} = -0.478 \pm 0.017\,(\text{stat})\,^{+0.018}_{-0.021}\,(\text{syst})}$$

$$D_{\text{exp}} = -0.465^{+0.016}_{-0.017}\,(\text{stat})\,^{+0.019}_{-0.022}\,(\text{syst})$$

  - **5 standard deviations observation** of top quarks being entangled at $t\bar{t}$ threshold

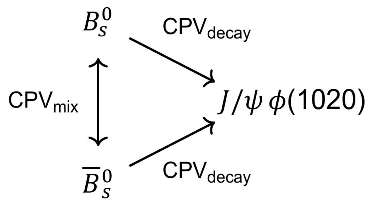
- Good agreement with SM predictions
  - Significantly improved by η_t inclusion
- **First measurement** of entanglement of top quarks with CMS data
- **Even in presence of a hypothetical toponium bound state, we confirm the existence of entanglement in the $t\bar{t}$ system**



Precision measurement of CP violation in B_s mesons

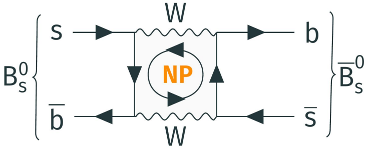
CMS PAPER BPH-23-004

Motivations

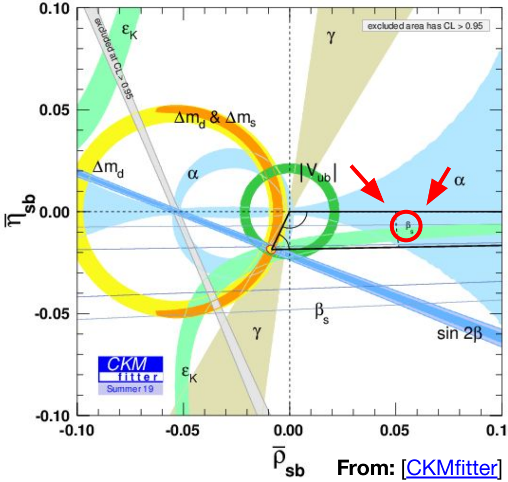


$$\boxed{\Gamma\left(B_s^0{}_{(\rightsquigarrow \overline{B}_s^0)} \to f\right)(t) \overset{?}{\neq} \Gamma\left(\overline{B}_s^0{}_{(\rightsquigarrow B_s^0)} \to f\right)(t)}$$

- **B_s 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 ϕ_s is the main CPV observable**
    - Predicted by the SM to be $\boldsymbol{\phi_s \approx -2\beta_s}$ (β_s ➜ angle of the B_s unit. triangle)
        - Neglecting contributions from higher-order diagrams
    - β_s determined by CKM global fits to be $\boldsymbol{-2\beta_s = -37 \pm 1}$ **mrad** [CKMfitter, UTfit]
- **New physics** can change the value of ϕ_s up to ~100% via new particles contributing to the flavor oscillations [RMP88(2016)045002]

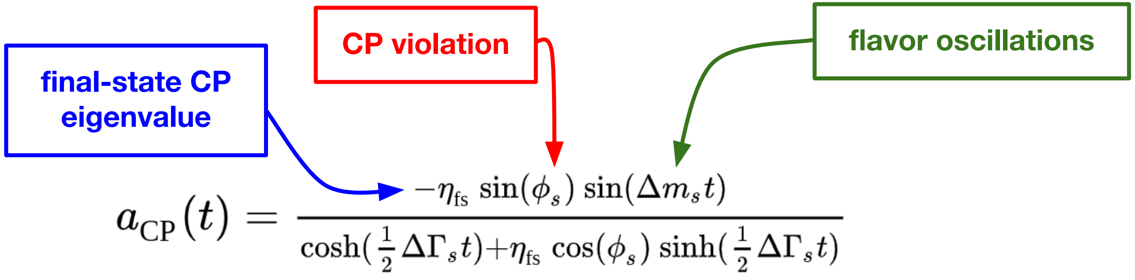


- This work talk presents the latest CMS results with the *golden* channel $B_s \to J/\psi\,\phi(1020) \to \mu^+\mu^- K^+K^-$



From: [CKMfitter]

A time-, flavor- and angular-dependent measurement



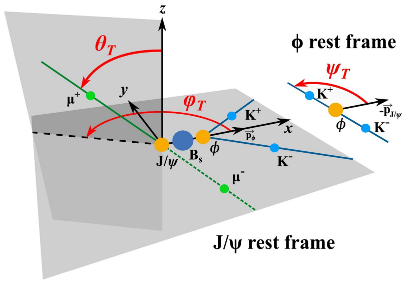
final-state CP eigenvalue

CP violation

flavor oscillations

$$a_{\mathrm{CP}}(t) = \frac{-\eta_{\mathrm{fs}}\,\sin(\phi_s)\,\sin(\Delta m_s t)}{\cosh(\frac{1}{2}\Delta\Gamma_s t)+\eta_{\mathrm{fs}}\,\cos(\phi_s)\,\sinh(\frac{1}{2}\Delta\Gamma_s t)}$$

Transversity basis

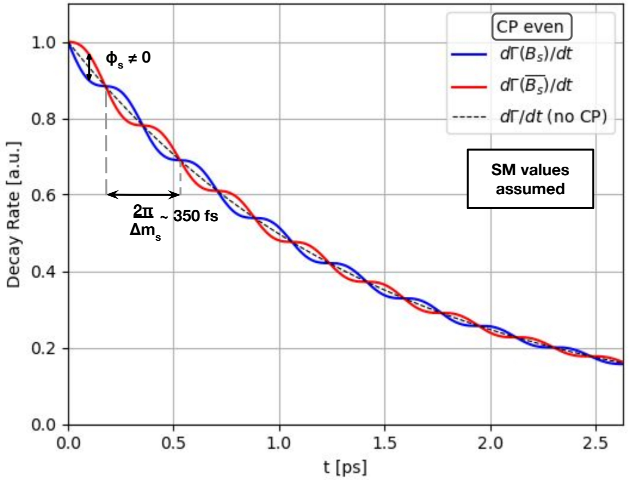


Core ingredients

- **Time-dependent angular analysis** to separate the CP eigenstates (“transversity basis” used)
- **Time-dependent flavor analysis** to resolve the B_s mixing oscillations (T ~ 350 fs)

$$\text{sensistivity} \propto \sqrt{\frac{\epsilon_{\mathrm{tag}}\mathcal{D}^2_{\mathrm{tag}}N_{\mathrm{sig}}}{2}}\sqrt{\frac{N_{\mathrm{sig}}}{N_{\mathrm{sig}}+N_{\mathrm{bkg}}}}\,e^{-\frac{\sigma_t^2\Delta m_s^2}{2}}$$

Decay rate for a CP-even final state



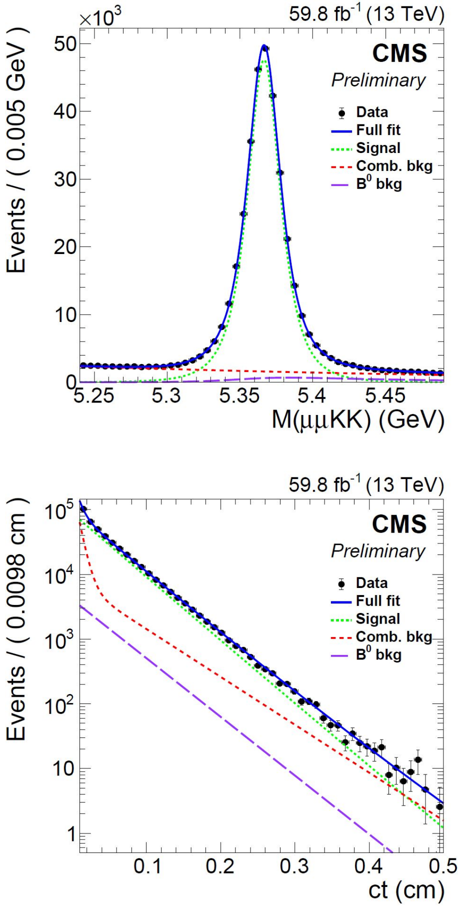
Analysis overview

- **Dataset**: $L_{int} = 96\ fb^{-1}$ collected in 2017-2018
- **Signal candidates:** 491 270 ± 950
- **Two-triggers strategy**
  1. *$J/\psi \rightarrow \mu^+\mu^-$ plus an additional muon*
     - Used for time resolution modeling
     - ≈50 000 signal candidates
     - Only tagging option: Opposite Side (OS) muon
  2. *Displaced $J/\psi \rightarrow \mu^+\mu^-$ plus $\phi \rightarrow K^+K^-$*
     - ≈450 000 signal candidates
     - Inclusive tagging suite: OS muon, OS electron, OS jet, Same Side

Fit: unbinned multidimensional extended maximum-likelihood

- **Input observables:** m_{Bs}, ct, σ_{ct}, θ_T, ψ_T, φ_T, $\boxed{\xi_{tag}, \omega_{tag}}$ Tag decision, mistag probability
- **Fitted parameters**
  - *CPV observables*: ϕ_s, $\boxed{|\lambda|}$ Direct CPV
  - *B_s system properties:* $\Delta\Gamma_s$, Γ_s, Δm_s
  - *Decay polarization*: $|A_0|^2$, $|A_\perp|^2$, $|A_S|^2$, $\delta_{//}$, $\delta_\perp$, $\delta_{S\perp}$
- **Background sources:** combinatorial, $B^0 \rightarrow J/\psi K^{*0}$, $\Lambda_b \rightarrow J/\psi K^- p$ (negligible)

Invariant mass and proper decay length (ct) distribution for HLT JpsiTrkTrk (2018)



Experimental effects

Time efficiency

- Modeled in $B^0 \rightarrow J/\psi K^{*0}$ control data sample with corrections from MC

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

Time resolution

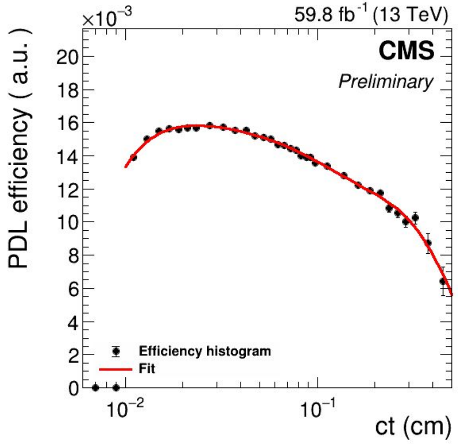
- Estimated from the measured σ(ct) with calibration in prompt events
- Excellent agreement found (calibrations around ~5%)

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

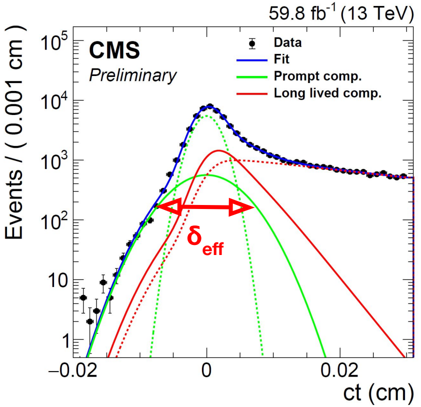
Angular efficiency

- Estimated with KDE distributions in simulated events
- The simulations are corrected to match the data

Time efficiency for HLT JpsiTrkTrk 2018 data



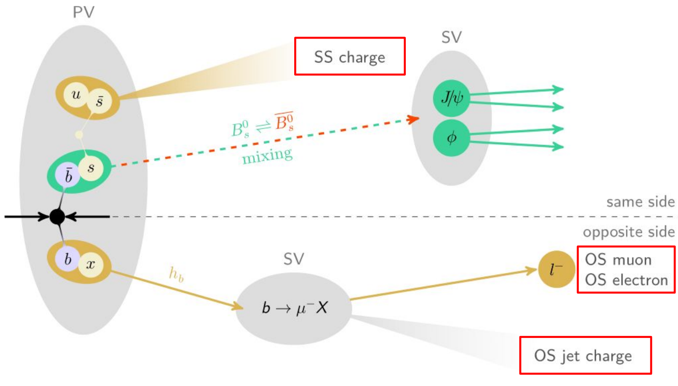
Time resolution calibration for 2018 data



Flavor tagging overview

- A **cutting-edge** **flavor tagging framework** has been engineered to extract the best possible results from data
- **Four DNN-based algorithms are used**, divided into two main categories
  - **Same side (SS):** exploits the B_s fragmentation
    1. **SS tagger**: leverages charge asymmetries in the B_s 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
- All algorithms are trained on simulations and calibrated in $B^+ \rightarrow J/\psi K^+$ with special precautions to reduce systematic effects
  - The SS tagger is trained on a mixture of B_s and B^+ events, with additional corrections to address B_s/B^+ differences in the hadronization

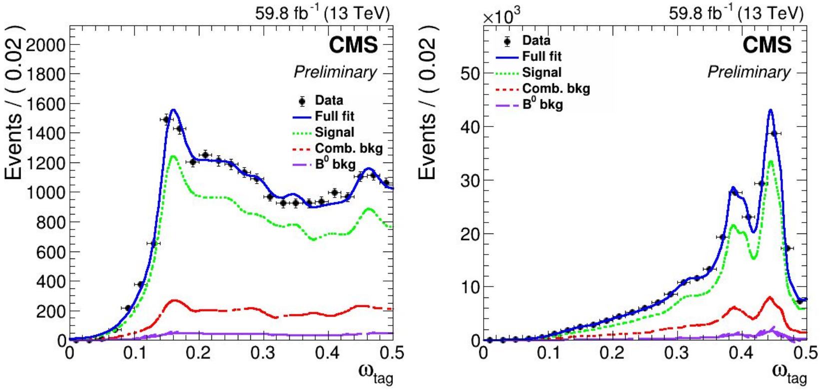
Schematic representation of a generic HLT JpsiTrkTrk event



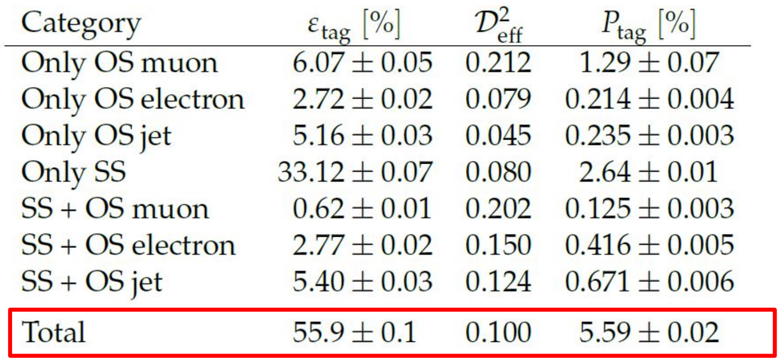
Flavor tagging performance

- The SS and any one of the OS algorithms overlap in about 20% of the events
  - 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_s data sample
  - Among the highest ever recorded at LHC
  - SS accounts for half of the performance
- **This is the first CMS implementation of the OS jet and same side tagging techniques**
- 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

ω_{tag} distribution in HLT JpsiMuon (left) and HLT JpsiTrkTrk (right) for 2018



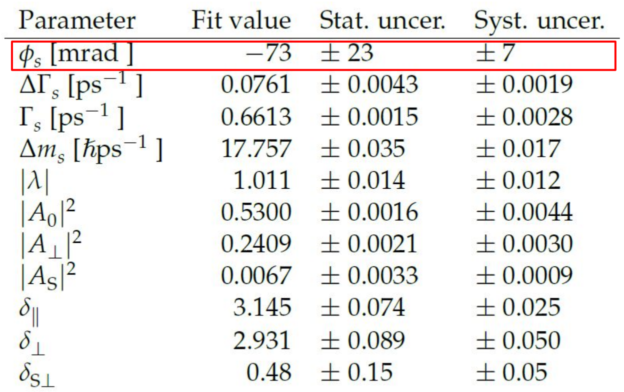
Flavor tagging performance (mutually exclusive categories)



Category	ε_{tag} [%]	$\mathcal{D}^2_{eff}$	P_{tag} [%]
Only OS muon	6.07 ± 0.05	0.212	1.29 ± 0.07
Only OS electron	2.72 ± 0.02	0.079	0.214 ± 0.004
Only OS jet	5.16 ± 0.03	0.045	0.235 ± 0.003
Only SS	33.12 ± 0.07	0.080	2.64 ± 0.01
SS + OS muon	0.62 ± 0.01	0.202	0.125 ± 0.003
SS + OS electron	2.77 ± 0.02	0.150	0.416 ± 0.005
SS + OS jet	5.40 ± 0.03	0.124	0.671 ± 0.006
Total	55.9 ± 0.1	0.100	5.59 ± 0.02

Results

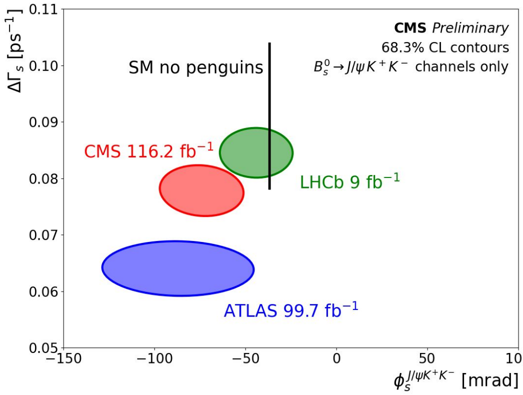
Results



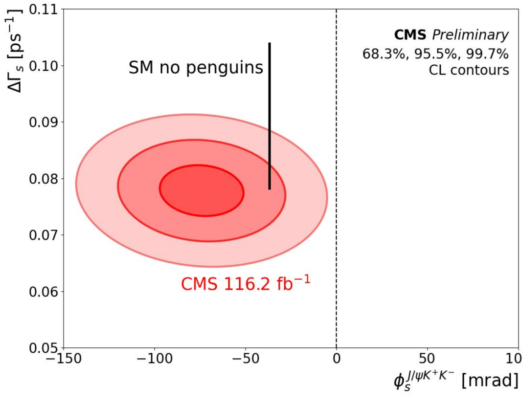
Parameter	Fit value	Stat. uncer.	Syst. uncer.
ϕ_s [mrad]	-73	± 23	± 7
$\Delta\Gamma_s$ [ps^{-1}]	0.0761	± 0.0043	± 0.0019
Γ_s [ps^{-1}]	0.6613	± 0.0015	± 0.0028
Δm_s [$\hbar$ps^{-1}]	17.757	± 0.035	± 0.017
$\lvert\lambda\rvert$	1.011	± 0.014	± 0.012
$\lvert A_0\rvert^2$	0.5300	± 0.0016	± 0.0044
$\lvert A_\perp\rvert^2$	0.2409	± 0.0021	± 0.0030
$\lvert A_S\rvert^2$	0.0067	± 0.0033	± 0.0009
$\delta_\parallel$	3.145	± 0.074	± 0.025
$\delta_\perp$	2.931	± 0.089	± 0.050
$\delta_{S\perp}$	0.48	± 0.15	± 0.05

- **Extremely competitive results** comparable to the most precise single measurements by LHCb
  - Largest ever effective statistics ($N_{Bs} \cdot P_{tag}$) for a ϕ_s measurement in this final state
- Leading systematic sources for ϕ_s: model bias, flavor tagging, and angular efficiency

Comparison with other LHC experiments



1, 2, 3 standard deviations contours



- These results supersede PLB816(2021)136188 and are further combined with those obtained CMS at 8 TeV, yielding:

$$\boxed{\begin{array}{c}\phi_s = -74 \pm 23 \text{ [mrad]} \\ \Delta\Gamma_s = 0.0780 \pm 0.0045 \text{ [ps}^{-1}\text{]}\end{array}}$$

- **The combined value for the weak phase ϕ_s is consistent with the SM prediction, the latest world average, and with zero (no CPV) at 3.2 s.d.**
  - **This is the first evidence of CPV in $B_s \to J/\psi K^+K^-$ decays**
- These results helps to further constrain possible BSM effects in the B_s system

Outlook

Summary and outlook

- This talk presented two high-profile new CMS results on the physics of heavy quarks
    - **Measurement of the top quark entanglement**
        - Measured in leptonic channels in the low mass region
        - Inclusion of toponium significantly improves the agreement with SM
        - **Entanglement observed at more than five standard deviations**
    - **Measurement of the CP violation in B_s mesons**
        - Largest ever effective statistics ($N_{Bs} \cdot P_{tag}$) for a ϕ_s measurement in this f.s.
        - Pioneering flavor tagging framework with charge-based SS techniques
        - **Evidences of CPV in $B_s \rightarrow J/\psi K^+K^-$ at three standard deviations**
- With the increase in statistics and the development of new techniques, the future for heavy quark physics at CMS looks brighter than ever

Stay tuned in the future for other exciting CMS results!

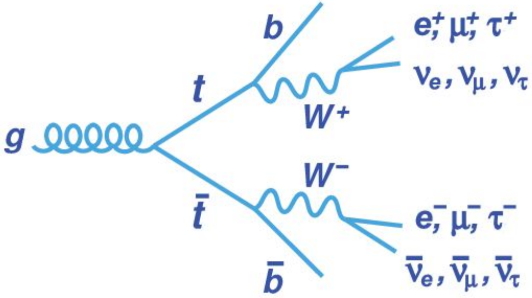
Thanks for the attention

Backup

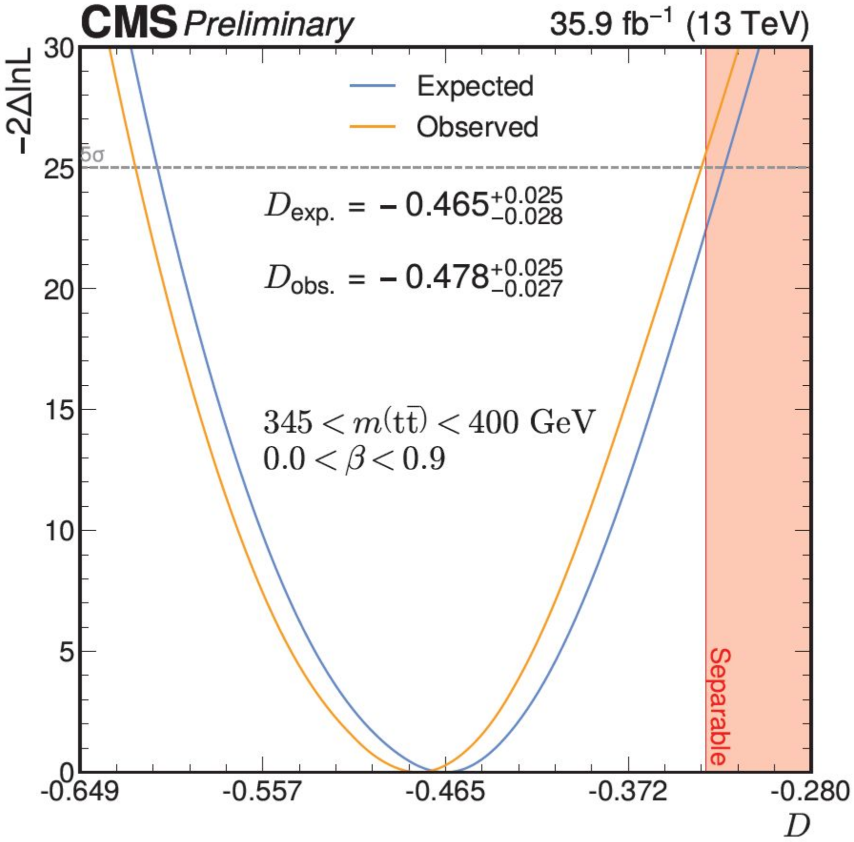
- Top entanglement -

Top entanglement event selection

- **2 oppositely charged isolated leptons (ee, eμ and μμ)**
  - $p_T > 25(20)$ GeV, for leading(trailing) lepton and $|\eta| < 2.4$
  - Electron discriminator vs fakes
  - Muon discriminator vs fakes
  - Veto events with more than two leptons
- **Reject events with $m_{\ell\ell} < 20$ GeV**
- **≥ 2 jets (R=0.4)**
  - $p_T > 30$ GeV and $|\eta| < 2.4$
  - Jet cleaning: $\Delta R(\ell, \text{jet}) > 0.4$
- **ee, μμ channels:**
  - $E_{\text{miss, T}}$: > 40 GeV
  - Z veto: $|m_Z - m_{\ell\ell}| > 15$ GeV
- **Reject events failing kinematic reconstruction constraints**

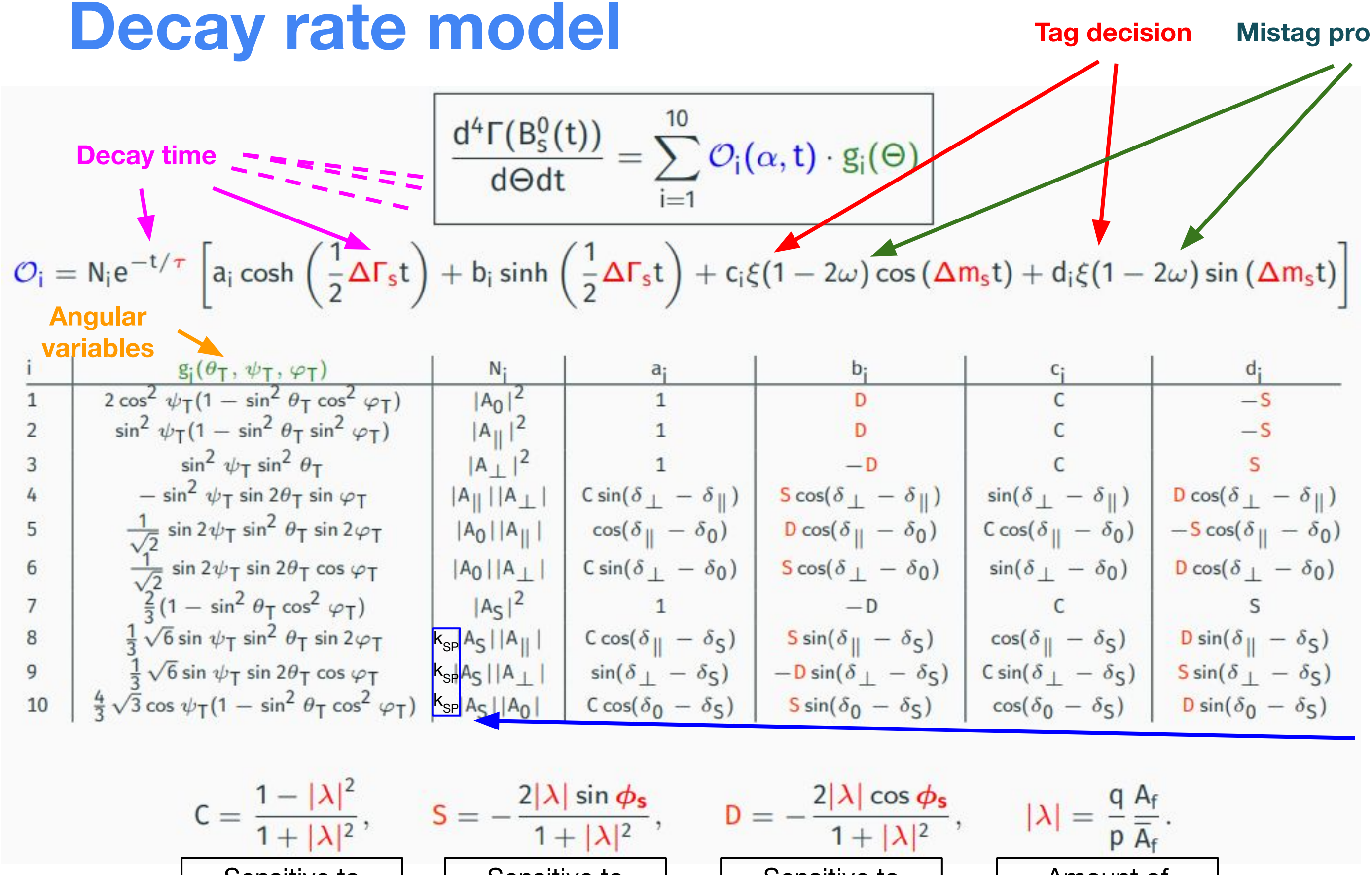


Likelihood scan result



Backup
- CPV in B_s -

Decay rate model



$$\frac{d^4\Gamma(B^0_s(t))}{d\Theta dt} = \sum_{i=1}^{10} \mathcal{O}_i(\alpha, t) \cdot g_i(\Theta)$$

$$\mathcal{O}_i = N_i e^{-t/\tau}\left[a_i \cosh\left(\frac{1}{2}\Delta\Gamma_s t\right) + b_i \sinh\left(\frac{1}{2}\Delta\Gamma_s t\right) + c_i \xi(1-2\omega)\cos(\Delta m_s t) + d_i \xi(1-2\omega)\sin(\Delta m_s t)\right]$$

i	$g_i(\theta_T, \psi_T, \varphi_T)$	N_i	a_i	b_i	c_i	d_i
1	$2\cos^2\psi_T(1-\sin^2\theta_T\cos^2\varphi_T)$	$\lvert A_0\rvert^2$	1	D	C	$-$S
2	$\sin^2\psi_T(1-\sin^2\theta_T\sin^2\varphi_T)$	$\lvert A_\parallel\rvert^2$	1	D	C	$-$S
3	$\sin^2\psi_T\sin^2\theta_T$	$\lvert A_\perp\rvert^2$	1	$-$D	C	S
4	$-\sin^2\psi_T\sin 2\theta_T\sin\varphi_T$	$\lvert A_\parallel\rvert\lvert A_\perp\rvert$	$C\sin(\delta_\perp-\delta_\parallel)$	$S\cos(\delta_\perp-\delta_\parallel)$	$\sin(\delta_\perp-\delta_\parallel)$	$D\cos(\delta_\perp-\delta_\parallel)$
5	$\frac{1}{\sqrt{2}}\sin 2\psi_T\sin^2\theta_T\sin 2\varphi_T$	$\lvert A_0\rvert\lvert A_\parallel\rvert$	$\cos(\delta_\parallel-\delta_0)$	$D\cos(\delta_\parallel-\delta_0)$	$C\cos(\delta_\parallel-\delta_0)$	$-S\cos(\delta_\parallel-\delta_0)$
6	$\frac{1}{\sqrt{2}}\sin 2\psi_T\sin 2\theta_T\cos\varphi_T$	$\lvert A_0\rvert\lvert A_\perp\rvert$	$C\sin(\delta_\perp-\delta_0)$	$S\cos(\delta_\perp-\delta_0)$	$\sin(\delta_\perp-\delta_0)$	$D\cos(\delta_\perp-\delta_0)$
7	$\frac{2}{3}(1-\sin^2\theta_T\cos^2\varphi_T)$	$\lvert A_S\rvert^2$	1	$-$D	C	S
8	$\frac{1}{3}\sqrt{6}\sin\psi_T\sin^2\theta_T\sin 2\varphi_T$	$k_{SP}\lvert A_S\rvert\lvert A_\parallel\rvert$	$C\cos(\delta_\parallel-\delta_S)$	$S\sin(\delta_\parallel-\delta_S)$	$\cos(\delta_\parallel-\delta_S)$	$D\sin(\delta_\parallel-\delta_S)$
9	$\frac{1}{3}\sqrt{6}\sin\psi_T\sin 2\theta_T\cos\varphi_T$	$k_{SP}\lvert A_S\rvert\lvert A_\perp\rvert$	$\sin(\delta_\perp-\delta_S)$	$-D\sin(\delta_\perp-\delta_S)$	$C\sin(\delta_\perp-\delta_S)$	$S\sin(\delta_\perp-\delta_S)$
10	$\frac{4}{3}\sqrt{3}\cos\psi_T(1-\sin^2\theta_T\cos^2\varphi_T)$	$k_{SP}\lvert A_S\rvert\lvert A_0\rvert$	$C\cos(\delta_0-\delta_S)$	$S\sin(\delta_0-\delta_S)$	$\cos(\delta_0-\delta_S)$	$D\sin(\delta_0-\delta_S)$

$$C = \frac{1-|\lambda|^2}{1+|\lambda|^2}, \quad S = -\frac{2|\lambda|\sin\phi_s}{1+|\lambda|^2}, \quad D = -\frac{2|\lambda|\cos\phi_s}{1+|\lambda|^2}, \quad |\lambda| = \frac{q}{p}\frac{\bar{A}_f}{A_f}.$$

Sensitive to direct CPV

Sensitive to $\phi_s \sim 0$

Sensitive to $\phi_s \sim \pi/2$

Amount of direct CPV

Conventions

- $|A_{//}|^2 = |A_0|^2 - |A_\perp|^2$
- $\delta_0 = 0$
- $\delta_{S\perp} = \delta_S - \delta_\perp$

Physics parameters

- ϕ_s, $|\lambda|$
- $\Delta\Gamma_s$, Γ_s, Δm_s
- $|A_0|^2$, $|A_\perp|^2$, $|A_S|^2$
- $\delta_{//}$, $\delta_\perp$, $\delta_{S\perp}$

S-P wave effective coupling $k_{SP} \approx 0.54$

- Introduced since $m(K^+K^-)$ is not fitted
- Evaluated by integrating the S- and P-wave lineshape interference

Theory reference: J.Phys.G48(2021)6,065002

Penguin contributions

We measure this

Assuming this is negligible

$$\boxed{\phi_s} = \phi_s^{tree} + \boxed{\Delta\phi_s^{penguin}} + \boxed{\Delta\phi_s^{NP}}$$

$$\boxed{\sin(2\beta)} = \sin(2\beta^{tree} + \boxed{\Delta\phi_d^{penguin}} + \boxed{\Delta\phi_d^{NP}})$$

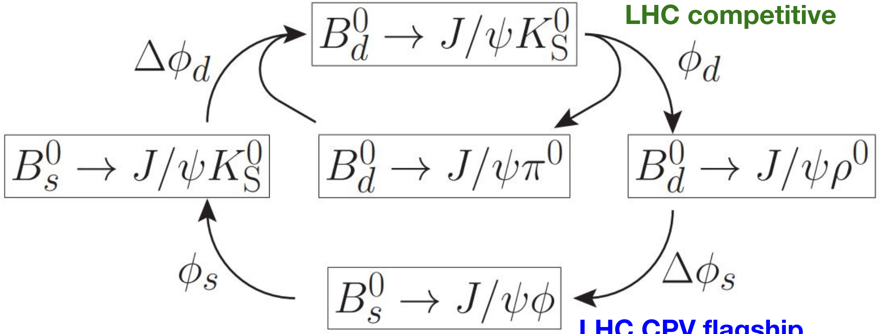
Trying to probe this

- Penguin pollutions are expected to be small for B_s, but they are not well constrained

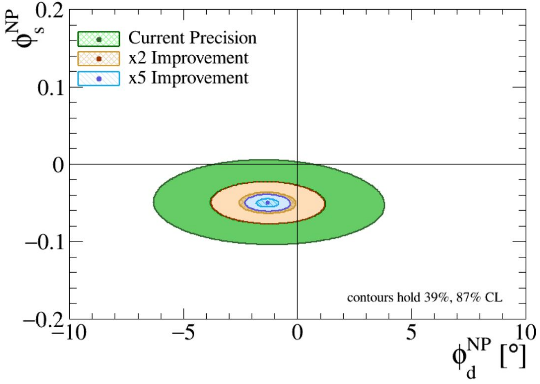
$$\Delta\phi_s^{penguin} \approx 3 \pm 10 \text{ mrad}$$

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

B-factories CPV flagship
LHC competitive



LHC CPV flagship
B-factories not competitive



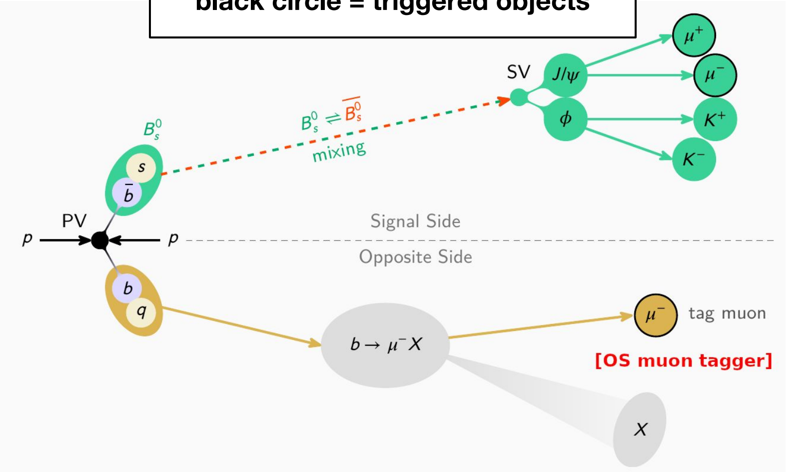
Trigger strategy

PLB816(2021)136188
(superseded)

black circle = triggered objects

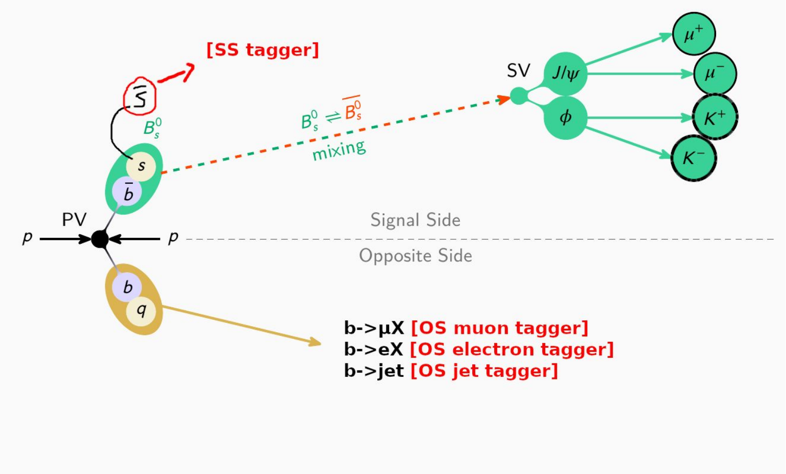
HLT JPsiMuon

- J/ψ ➔ μμ candidate plus an additional muon (for tagging)
- Around **50k** signal candidates
- Used for time resolution calibration as it is not displaced
- Tagging algorithms applied: OS-muon
  - $P_{tag} \sim$ **10%**



HLT JPsiTrkTrk

- Displaced J/ψ ➔ $\mu^+\mu^-$ candidate + two charged tracks compatible with a ϕ(1020)
- Around **450k** signal candidates
- Displaced (lifetime turn-on efficiency to model)
- Possible tagging strategies: OS-muon, OS-electron, OS-jet, Same Side
  - $P_{tag} \sim$ **5%**



Offline selection

Requirements common between the two HLTs

- $5.24 < m(\mu\mu KK) < 5.49$ GeV
- $p_T(B_s) > 9.5$ GeV
- Vertex probability > 2%
- $\sigma(ct) < 50$ μm
- $|\eta(\mu)| < 2.4$
- $|\eta(K)| < 2.5$
- $|m(\mu\mu) - m(J/\psi^{PDG})| < 150$ MeV
- $|m(KK) - m(\phi(1020)^{PDG})| < 10$ MeV

Requirements specific to HLT JPsiMuon

- $p_T(\mu) > 3.5$ GeV
- $p_T(K) > 1.15$ GeV
- $ct > 60$ μm

Requirements specific to HLT JPsiTrkTrk

- HLT JPsiMuon vetoed
- $p_T(\mu) > 4$ GeV
- $p_T(K) > 0.9$ GeV
- $p_T(\mu\mu) > 6.9$ GeV
- $ct > 100$ μm, $ct/\sigma(ct) > 3$

- **Selection requirement optimized** with the a genetic algorithm to maximize $S/\sqrt{(S + B)}$
- **To avoid overlaps, HLT JPsiMuon is vetoed in the HLT JPsiTrkTrk category**
- The **PV** of choice is the closest in 3D to the line that passes through the SV and parallel to the B_s momentum

Time uncertainty and resolution

- The time dependence of the decay rate is parametrized with the proper decay length *ct*

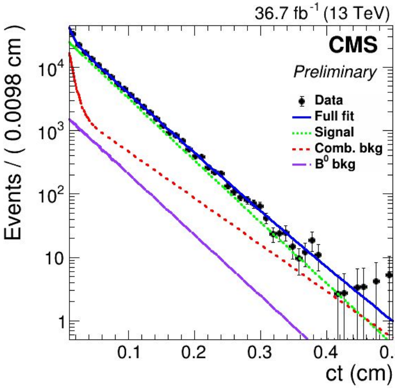
$$ct = c \cdot \frac{m_{B_s}^{w.a.} \cdot L_{xy}}{p_T}$$

- The ct *uncertainty* σ_{ct} is obtained by fully propagating the uncertainties in L_{xy} and p_T
    - The uncertainty on L_{xy} dominates for most of the ct spectrum, with the uncertainty on p_T taking over for very high values (ct > 3 mm)
- The ct *resolution* δ_{ct} generates a dilution effect on the observed CP asymmetry

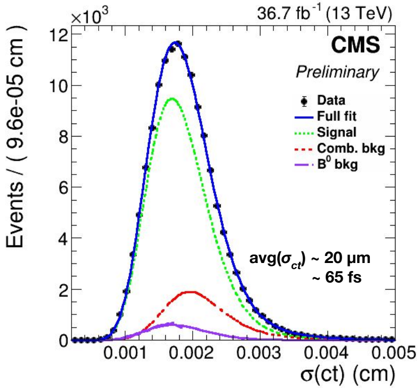
$$\boxed{\mathcal{D}_{\text{time}} = e^{-\delta_t^2 \Delta m_s^2/2}}$$

- In this work, we assume $\delta_{ct} \sim \sigma_{ct}$ and we improve this hypothesis with a data-driven calibration

ct distribution in HLT JPsiTrkTrk 2017



σ_{ct} distribution in HLT JPsiTrkTrk 2017



avg(σ_{ct}) ~ 20 μm
~ 65 fs

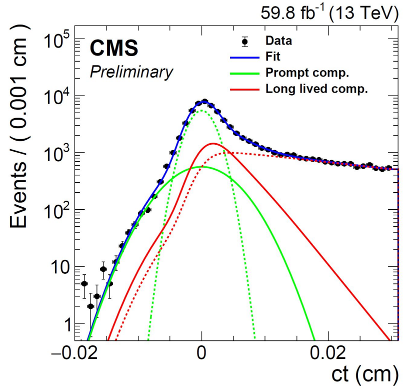
Time resolution calibration

- **The time resolution is calibrated in a prompt data sample** of $B_s \rightarrow J/\psi\ \phi$
  - Performed separately for 2017 and 2018 in 5 bins of σ_{ct}
- For each bin, a fit to the ct distribution is performed to describe the prompt and long-lived components
  - **The resolution is modeled with two gaussians**
- The two widths of the gaussians are used to evaluate the effective time dilution, from which an effective time resolution can be extracted

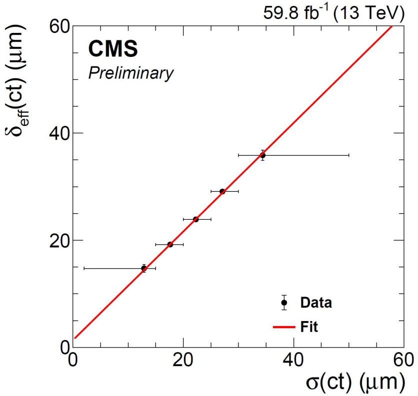
$$\mathcal{D} = f_1 \exp\left(-\frac{\sigma_1 \Delta m_s^2}{2}\right) + (1 - f_1) \exp\left(-\frac{\sigma_2 \Delta m_s^2}{2}\right) \quad \longrightarrow \quad \delta_{\text{eff}} = \sqrt{\frac{-2\ln\mathcal{D}}{\Delta m_s^2}}$$

- The estimated resolutions $\delta_{ct,eff}$ are compared with the average uncertainty in each σ_{ct} bin and a linear calibration is performed to model residual differences
- **The results show an almost perfect linear relationship**, with slope compatible with 1 and intercept ~1 μm

Time resolution fit (inclusive, 2018)

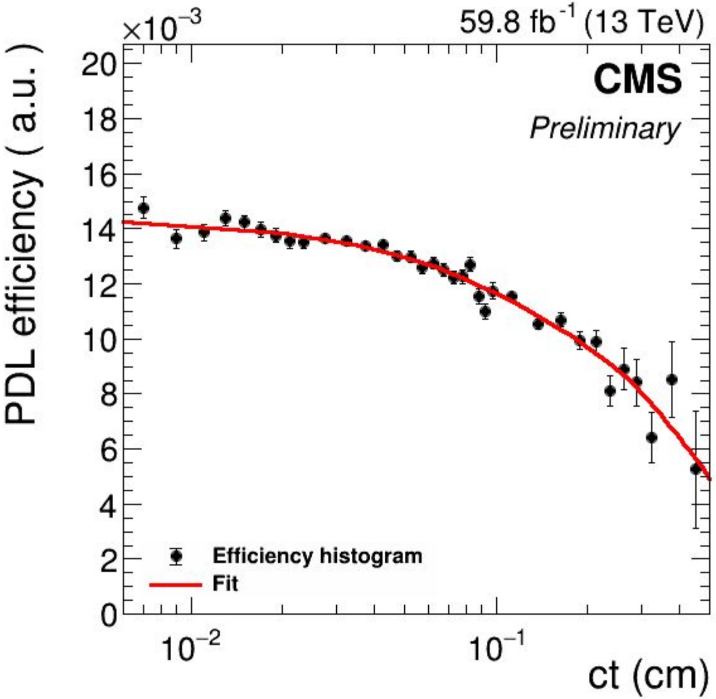


Time resolution calibration (2018)

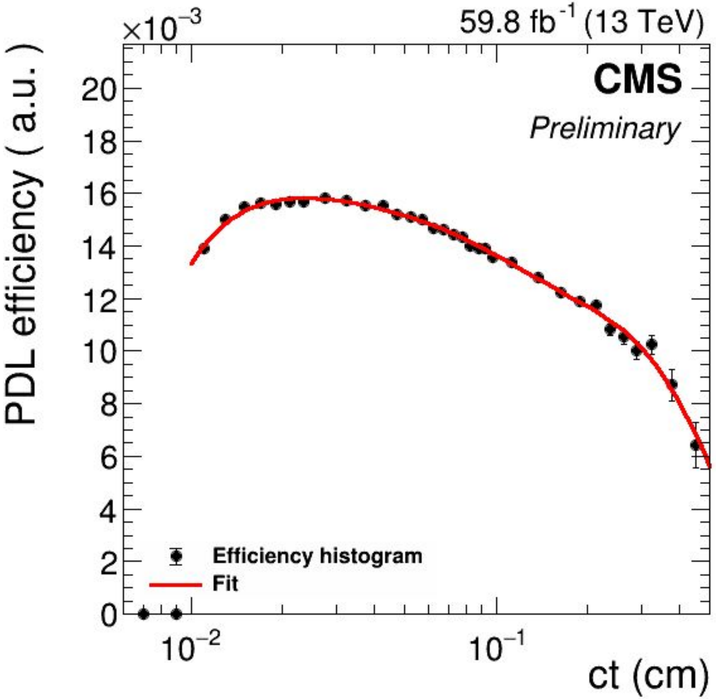


Time efficiency for 2018 data

Proper decay length efficiency (HLT JpsiMuon 2018)



Proper decay length efficiency (HLT JpsiTrkTrk 2018)



Common methodologies

- The **tagging inference** **logic** differs between algorithms
    - **Lepton taggers** (OS muon, OS electron)
        - Lepton charge ➜ ξ_{tag}; DNN ➜ ω_{tag} (DNN trained for correct-tag vs mistag)

$$\boxed{\begin{array}{l} \text{OS } \ell^- \to \text{OS } b \xrightarrow{\text{tag}} \text{signal } B_s^0 \\ \text{OS } \ell^+ \to \text{OS } \bar{b} \xrightarrow{\text{tag}} \text{signal } \bar{B}_s^0 \end{array}}$$

$$\boxed{\omega_{\text{evt}} = 1 - \boxed{s_{\text{DNN}}}}$$

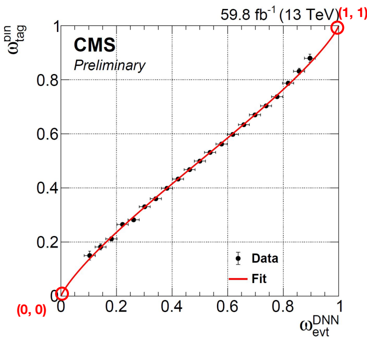
DNN score

-
    - **Charge-based taggers** (OS jet, SS)
        - DNN ➜ $P(B_s)$ ➜ ξ_{tag}, ω_{tag} (DNN trained for B_s vs $\overline{B_s}$)

$$\boxed{\begin{array}{llll} \boxed{s_{\text{DNN}}} > 0.5 + \epsilon \xrightarrow{\text{tag}} \text{signal } B_s^0 & \text{with} & \omega_{\text{evt}}^{\text{DNN}} = 1 - \boxed{s_{\text{DNN}}} \\ \boxed{s_{\text{DNN}}} < 0.5 - \epsilon \xrightarrow{\text{tag}} \text{signal } \bar{B}_s^0 & \text{with} & \omega_{\text{evt}}^{\text{DNN}} = \boxed{s_{\text{DNN}}} \end{array}}$$

- **The algorithms are optimized and trained in simulated events B_s ➜ J/ψ ϕ and calibrated in data with self-tagging B^+ ➜ J/ψ K^+ with the "Platt scaling" method** [ref]
    - The calibration is performed by comparing the predicted and measured ω_{tag}
    - The Platt scaling is a linear calibration of the score *before* the last sigmoid layer
        - This allows the DNN score to still be treated as a probability after the calibration

OS-Jet calibration (2018)



OS-lepton taggers selection

OS Muon

- **Requirements**
    - ○ $p_T > 2$ GeV
    - ○ $|\eta| < 2.4$
    - ○ $|d_z(PV)| < 1$ cm
    - ○ $\Delta R(B_s) > 0.4$
    - ○ Discriminators vs fakes

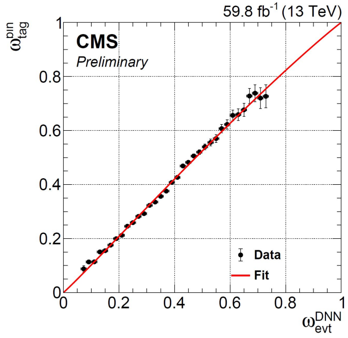
- Deployed in **both HLT categories**
- Dense DNN for ω_{tag} estimation
    - ○ Inputs: kinematics, IP, surrounding activity

OS electron

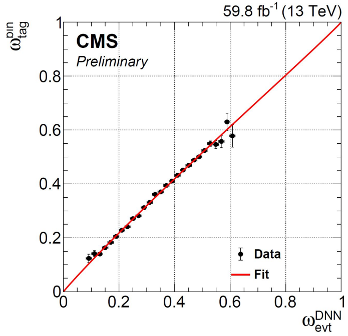
- **Requirements**
    - ○ No OS muon selected in the event
    - ○ $p_T > 2.5$ GeV
    - ○ $|\eta| < 2.4$
    - ○ $|d_z(PV)| < 0.2$ cm
    - ○ $|d_{xy}(PV)| < 0.08$ cm
    - ○ $\Delta R(B_s) > 0.4$
    - ○ Discriminators vs fakes

- Deployed **only in the *HLT_JpsiTrkTrk*** category
- Dense DNN for ω_{tag} estimation
    - ○ Inputs: kinematics, IP, surrounding activity

OS-Muon calibration (HLT JPsiMuon 2018)



OS-Electron calibration (2018)

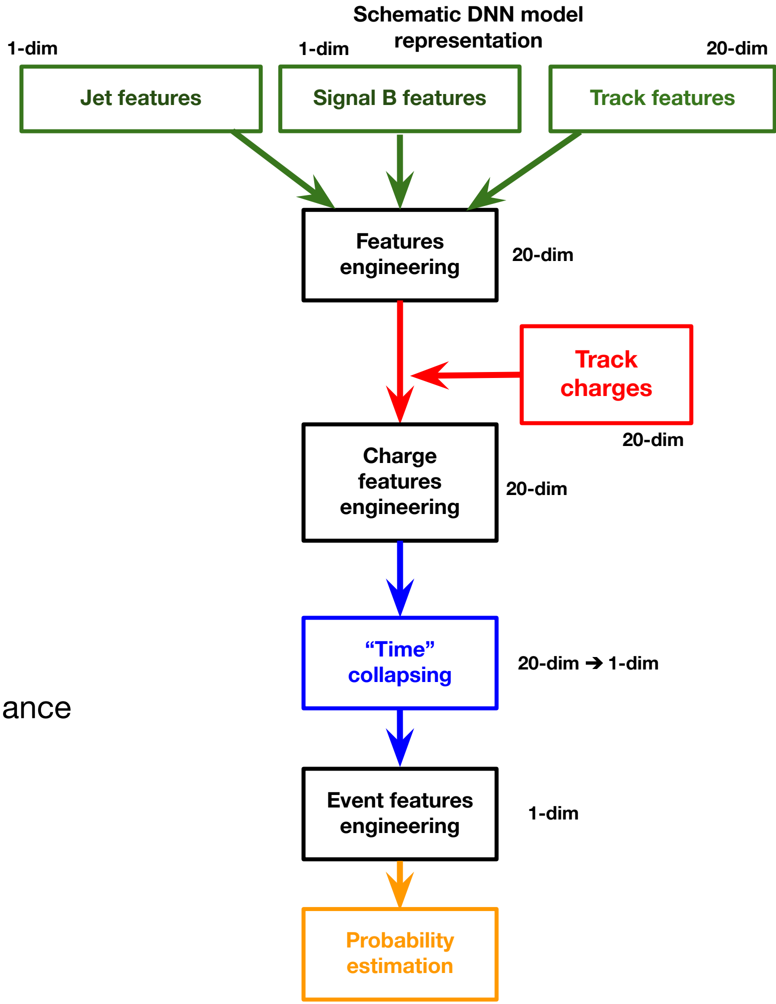


OS-jet tagger

- *DeepJetCharge* is based on the DeepSets architecture [ref]
    - Inputs: features from signal B, OS jet, and tracks
        - The only flavor asymmetry is in the charges
- **Jet selection**
    - No OS-lepton candidate
    - $p_T > 10$ GeV, $|\eta| < 2.5$
    - ≥ 2 tracks with $|d_z(PV)| < 1$ cm
    - $\Delta R(B_s) > 0.5$
    - b-tagging discriminator
- **Additional tracks selection**
    - $\Delta R(\text{track, jet}) < 0.5$
    - $|d_z(PV)| < 1$ cm
- **The output is explicitly symmetrized** due to the LHC charge imbalance

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

- Events with $\omega_{\text{tag}} > 0.48$ are labeled as *untagged*
- Deployed **only in the HLT JpsiTrkTrk** category



Schematic DNN model representation

1-dim **Jet features**

1-dim **Signal B features**

20-dim **Track features**

Features engineering 20-dim

Track charges 20-dim

Charge features engineering 20-dim

"Time" collapsing 20-dim ➔ 1-dim

Event features engineering 1-dim

Probability estimation

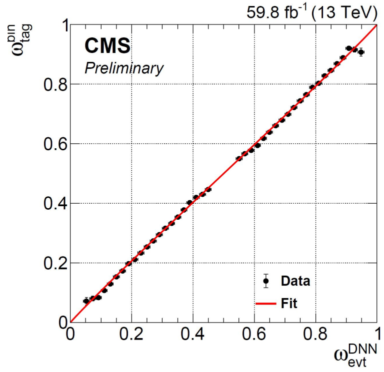
SS tagger

- ***DeepSSTagger* uses the kinematic information from up to 20 tracks (ordered by $|d_z|$) around the reconstructed *b*-meson**
- **Requirements**
  - $\Delta R(\text{track}, B) < 0.8$
  - $|d_z(PV)| < 0.4$ cm
  - $|d_{xy}(PV)|/\sigma_{dxy} < 1$
- Overlap with signal and OS is carefully avoided
- **DNN model is very similar to *DeepJetCharge***
- **Trained with 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
- Deployed **only in the HLT JpsiTrkTrk** category
- Output symmetrized and events with $\omega_{tag} > 0.46$ considered untagged

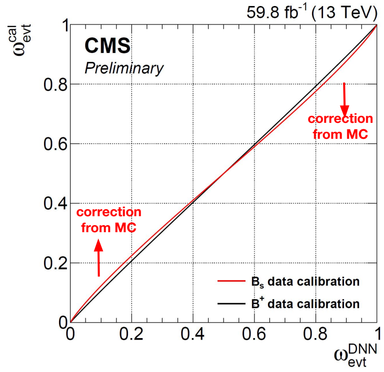
SS-tagger calibration and performance

- Direct calibration in data with B_s mesons was found to be not feasible in CMS
- The DNN is calibrated in data with $B^+ \rightarrow J/\psi\ K^+$ data
- Simulated B_s and B^+ samples are used to correct residual differences
- The correction is of the order of 10%
  - 100% of the correction size is assumed as a systematic uncertainty

Same-side tagger calibration (B^+ data, 2018)



Comparison between Same-side tagger B^+ and B_s calibrations



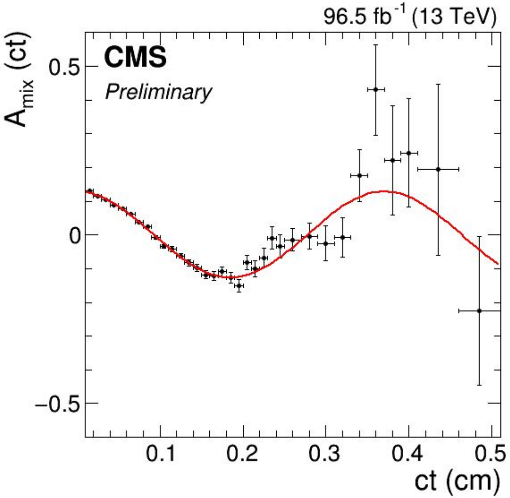
Tagging validation with B^0 events

- **The flavor tagging framework is validated in the $B^0 \rightarrow J/\psi K^{*0}$ control channel** (~2M events)
- The time-dependent **mixing asymmetry** is measured to extract the flavor mixing oscillation frequency Δm_d with a precision of *O(1%)*
  - Excellent agreement with world-averages observed

    ➜ **No bias in mixing frequency measurements**
- 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_{mix} and the estimated D_{tag} is found to be well described by a linear relationship, indicating that all four techniques behave in the same predictable way

Mixing asymmetry

$$A_{mix}(ct) = \frac{N_{unmix}(ct) - N_{mix}(ct)}{N_{unmix}(ct) + N_{mix}(ct)}$$

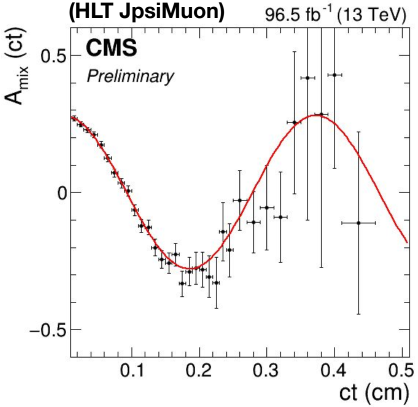


All categories are mutually exclusive

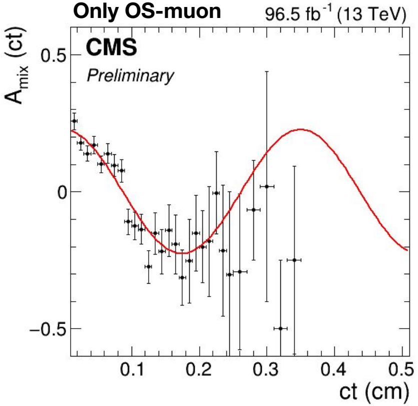
Mixing asymmetry for different tagging categories

All but the first refers to HLT JpsiTrkTrk

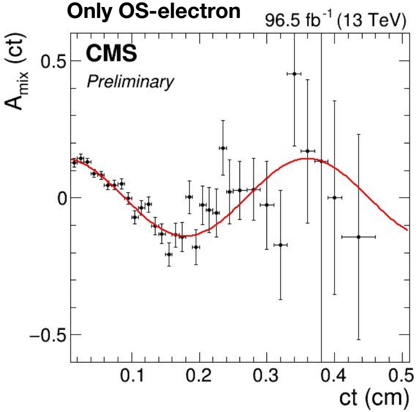
Only OS-muon (HLT JpsiMuon)



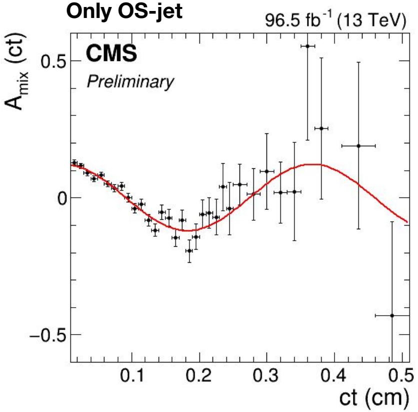
Only OS-muon



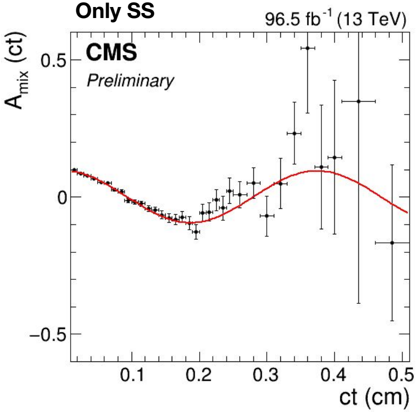
Only OS-electron



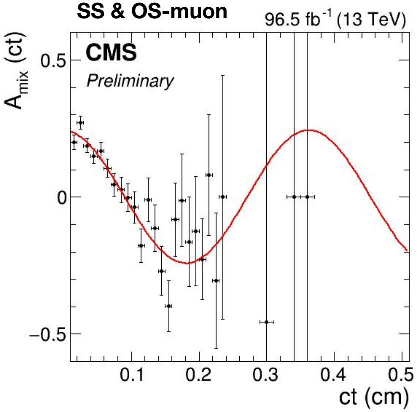
Only OS-jet



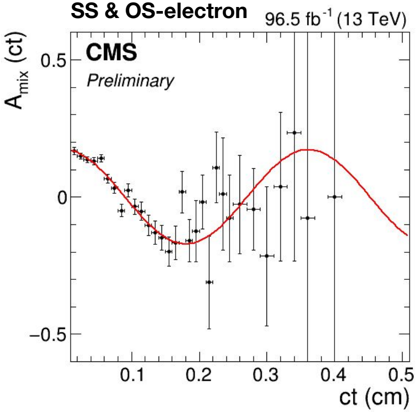
Only SS



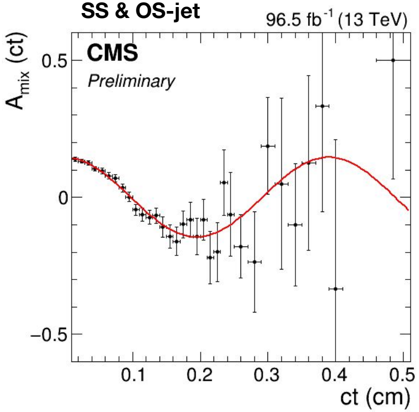
SS & OS-muon



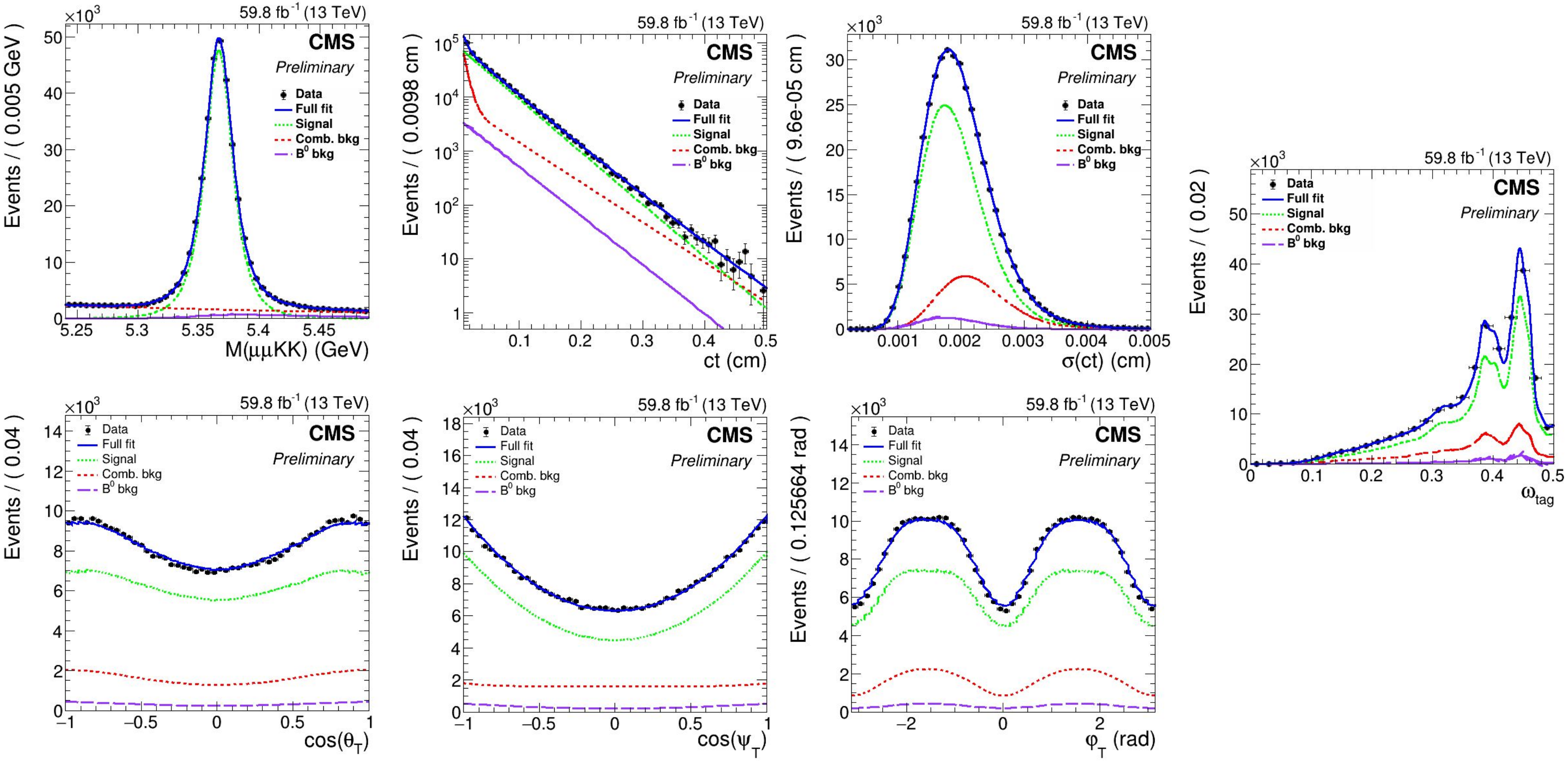
SS & OS-electron



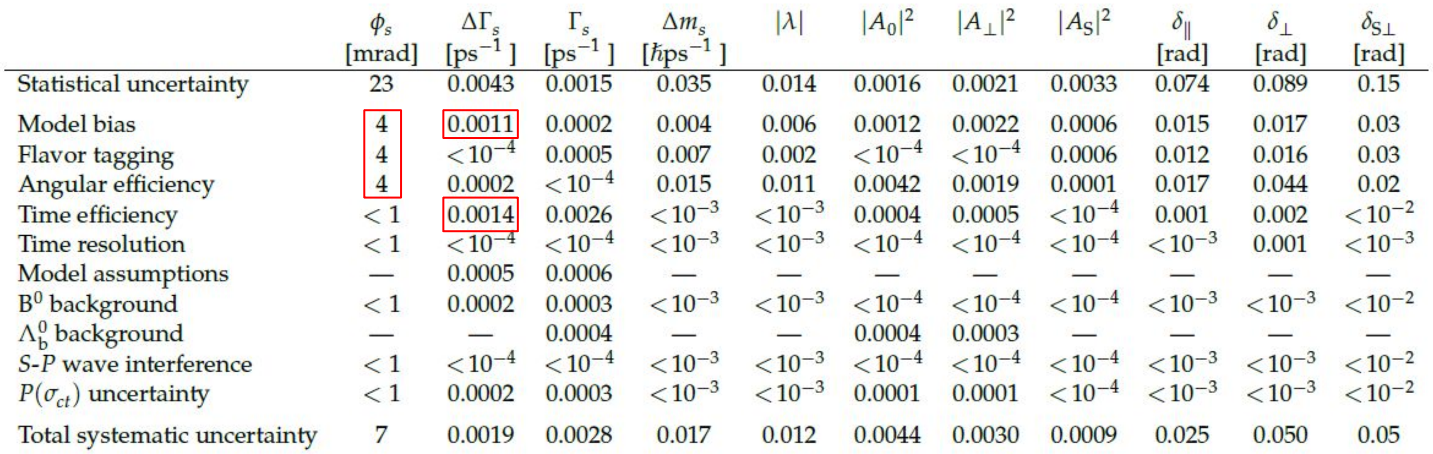
SS & OS-jet



Fit projections (HLT JpsiTrkTrk 2018)



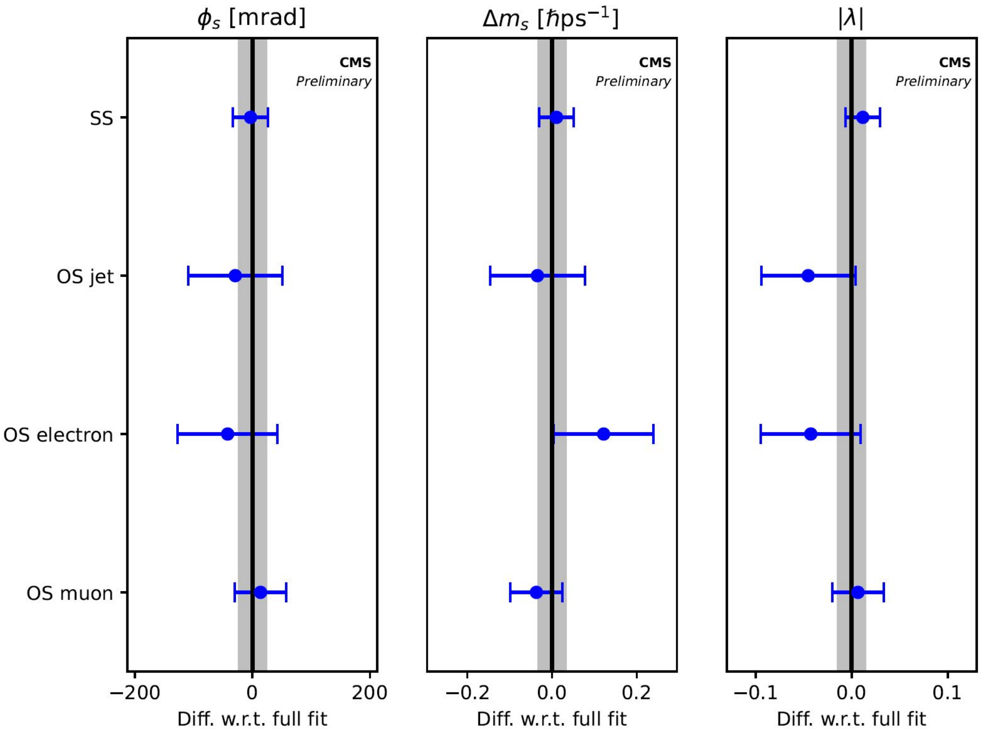
Systematic uncertainty overview



	ϕ_s [mrad]	$\Delta\Gamma_s$ [ps^{-1}]	Γ_s [ps^{-1}]	Δm_s [$\hbar$ps^{-1}]	$\lvert\lambda\rvert$	$\lvert A_0\rvert^2$	$\lvert A_\perp\rvert^2$	$\lvert A_S\rvert^2$	$\delta_\parallel$ [rad]	$\delta_\perp$ [rad]	$\delta_{S\perp}$ [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	—	—	—	—	—	—	—	—
B^0 background	<1	0.0002	0.0003	$<10^{-3}$	$<10^{-3}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-3}$	$<10^{-3}$	$<10^{-2}$
Λ_b^0 background	—	—	0.0004	—	—	0.0004	0.0003	—	—	—	—
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

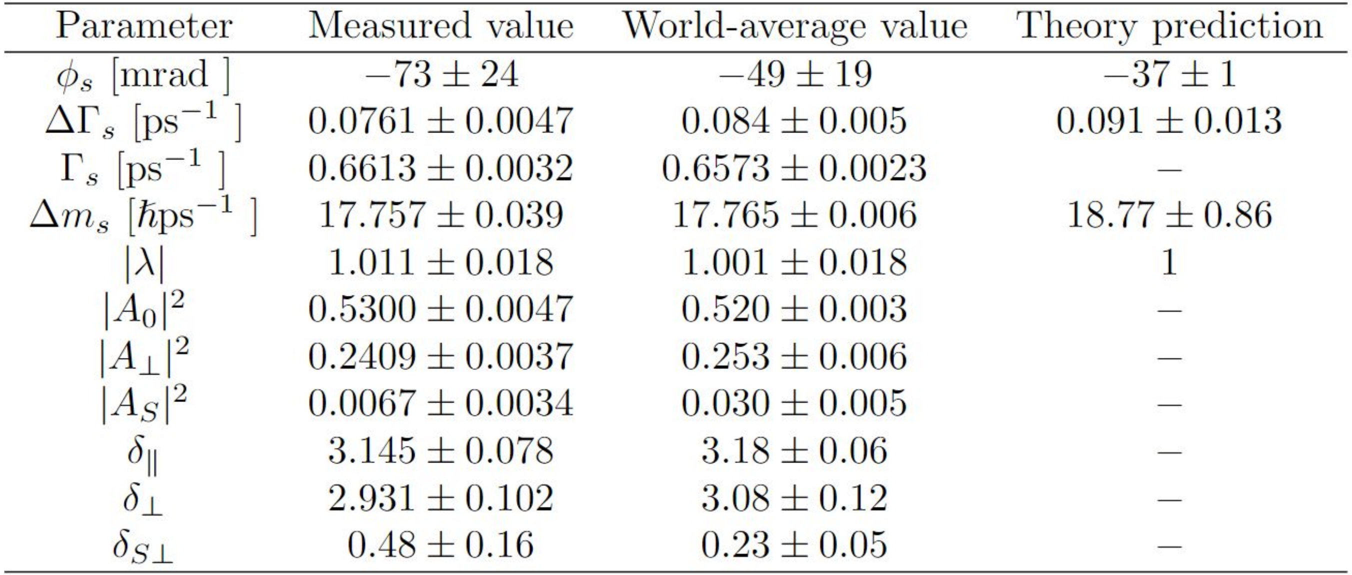
- **Model bias, flavor tagging, and angular efficiency are found to be the leading systematic sources for ϕ_s**
- The measurement is still heavily statistically limited for ϕ_s

Fit with individual tagging techniques



- All results are evaluated with respect to the reference ones
- The grey area represents the statistical uncertainty of the full fit
- Only flavor-sensitive parameters are shown
- **Excellent agreement between tag categories**

Comparison with theory and world averages



Parameter	Measured value	World-average value	Theory prediction
ϕ_s [mrad]	-73 ± 24	-49 ± 19	-37 ± 1
$\Delta\Gamma_s$ [ps^{-1}]	0.0761 ± 0.0047	0.084 ± 0.005	0.091 ± 0.013
Γ_s [ps^{-1}]	0.6613 ± 0.0032	0.6573 ± 0.0023	–
Δm_s [$\hbar$ps^{-1}]	17.757 ± 0.039	17.765 ± 0.006	18.77 ± 0.86
$\lvert\lambda\rvert$	1.011 ± 0.018	1.001 ± 0.018	1
$\lvert A_0\rvert^2$	0.5300 ± 0.0047	0.520 ± 0.003	–
$\lvert A_\perp\rvert^2$	0.2409 ± 0.0037	0.253 ± 0.006	–
$\lvert A_S\rvert^2$	0.0067 ± 0.0034	0.030 ± 0.005	–
$\delta_{\lVert}$	3.145 ± 0.078	3.18 ± 0.06	–
$\delta_\perp$	2.931 ± 0.102	3.08 ± 0.12	–
$\delta_{S\perp}$	0.48 ± 0.16	0.23 ± 0.05	–

[CKMfitter, UTfit]

[Lenz & Tetlalmatzi-Xolocotzi]

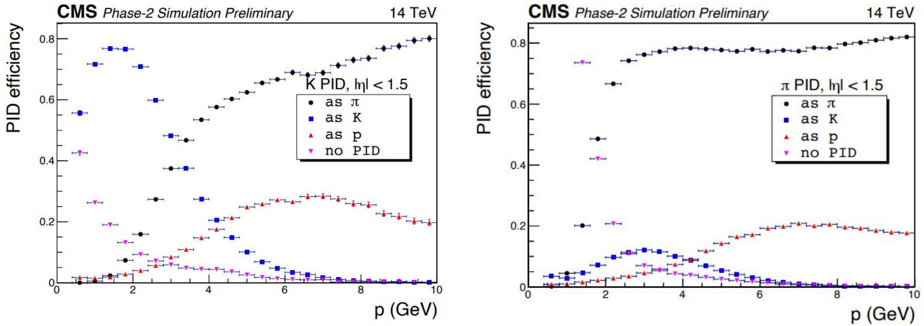
[Lenz & Tetlalmatzi-Xolocotzi]

Reference: CERN-CMS-DP-2022-025

Flavor tagging in Phase-2 with MTD

- The MTD (Mip Timing Detector) provides time information of charged tracks at its surface
- 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 *s*-quark in the B_s and a nearby soft kaon for flavor tagging
- The PID from MTD, when integrated in the Phase-2 extrapolation of this analysis, shows a significant improvement of the tagging performances

Simulated PID efficiencies



Relative gain in P_{tag} (only SS)

PID scenario	Gains in P_{tag}
MC truth (perfect PID < 3 GeV)	+66%
PID with σ_{BTL} = 40 ps	+24%
PID with σ_{BTL} = 70 ps	+14%