Measurement of the CP-violating phase $\phi_s$ in the $B_s^0 \rightarrow J/\psi \phi$ channel at 13 TeV by CMS

Alberto Bragagnolo$^a$, on behalf of the CMS Collaboration
LHCP 2020 – 29/05/2020

$^a$University & INFN Padova (IT)
Introduction
Motivations

- $\phi_s$ is a CPV phase arising from the interference between $B_s^0$ decays proceeding directly and through $B_s^0$-$\bar{B}_s^0$ mixing to a CP final state.

- **SM prediction:** $\phi_s \simeq -2\beta_s = -36.96^{+0.84}_{-0.72}$ mrad
  
  [CKMfitter]

- New Physics can **change** the value of $\phi_s$ up to $\sim 10\%$ via new particles contributing to the $B_s^0$-$\bar{B}_s^0$ mixing
  
  [JHEP04(2010)031]

- $B_s^0 \rightarrow J/\psi \phi$ is the **golden channel** to measure $\phi_s$
  
  - No direct CPV
  - Only one CPV phase
  - Easy to reconstruct with high S/B

- **Several other interesting observables** measurable with the same analysis: $\Gamma_s$, $\Delta\Gamma_s$, $|\lambda|$, $\Delta m_s^2$
  
  - $\lambda = \frac{q}{p} \tilde{A}_{f,s.}$, $|B_{L,H}\rangle = p|B_s^0\rangle \pm q|\bar{B}_s^0\rangle$
Measurement ingredients

\[ a_{CP}(t) \propto \eta_{\mu\mu K K} \sin(\phi_s) \sin(\Delta m_s t) \]

\[ \text{sensitivity} = f \left( \sqrt{\frac{P_{tag} S}{2}} \sqrt{\frac{S}{S + B}} \cdot e^{-\frac{\sigma t^2}{2} \Delta m_s^2} \right) \]

1. **Angular analysis** to separate the different CP eigenstate of the final state
   - \( \psi_T \): helicity angle of \( K^+ \) in the \( \phi \) rest frame
   - \( \theta_T \): polar angle of \( \mu^+ \) in the \( J/\psi \) rest frame
   - \( \phi_T \): azimuthal angle of \( \mu^+ \) in the \( J/\psi \) rest frame

2. **Excellent time resolution** to see the fast \( B_s^0 - \bar{B}_s^0 \) oscillation

3. **Highly efficient flavour tagging** to infer the initial \( B_s^0 \) flavour

4. **As much statistics as possible** (with good S/N)
Candidate selection
**Trigger strategy**

**Trigger:** $J/\psi \rightarrow \mu^+ \mu^-$ candidate **plus an additional muon**

- The additional muon is used to **tag** the flavour of the $B^0_s$, via $b \rightarrow \mu^- X$ decays of the other $b$
- However, the requirement for a third muon **lowers** the rate of selected events
- **Not to apply a displacement cut** on the $J/\psi \rightarrow \mu^+ \mu^-$ at HLT level

---

**This trigger improves** the tagging efficiency at the cost of the **reduced** number of signal events
**Schematic representation of an event**

Alberto Bragagnolo

Measurement of $\phi_s$ in the $B_s \rightarrow J/\psi \phi$ channel at 13 TeV by CMS
**Offline selection**

**Offline selection**

- \( p_T(\mu) \geq 3.5 \) GeV
- \( |\eta(\mu)| \leq 2.4 \)
- \( p_T(K) \geq 1.2 \) GeV
- \( |\eta(K)| \leq 2.5 \)
- \( |m(\mu^+\mu^-) - m_{J/\psi}^{\text{PDG}}| < 150 \) MeV
- \( |m(K^+K^-) - m_{\phi(1020)}^{\text{PDG}}| < 10 \) MeV

- \( p_T(B^0_s) \geq 11 \) GeV
- \( cT(B^0_s) \geq 70 \) \( \mu \)m
- \( B^0_s \to J/\psi \phi \text{ Vtx prob} \geq 0.1\% \)
- \( m(\mu^+\mu^-K^+K^-) \) \([5.24, 5.49]\) GeV

- Vertex fit performed with \( J/\psi \) mass constraint
- Cuts to enhance purity \( S/(S+B) \)

- \( \mathcal{L}_{\text{int}} = 96.4 \) fb\(^{-1} \) collected in 2017 and 2018

- Number of signal \( B^0_s = 48\,500 \)

- Number of candidates in Run-1: \( 49\,200 \)

**Figure:** fit to data 1D projection
Example of a candidate event

CMS Experiment at the LHC, CERN
Data recorded: 2018-Jun-11 17:23:01.070912 GMT
Run / Event / LS: 317683 / 314645082 / 248

$\mu^+$ from J/$\psi$

$\mu^-$ from J/$\psi$

Tag $\mu$
Efficiencies
Proper decay length efficiency

- The efficiency in selecting and reconstructing a $B_s^0$ decay depends on the decay length.

- To properly fit the decay rate model, we need a parametrization of the decay length efficiency.

- Efficiency is evaluated with simulated samples, separately for 2017 and 2018, and fitted in the $ct$ range $0.007-0.5 \text{ cm}$.

$$
\epsilon(ct) = e^{-a \cdot ct} \cdot \text{Chebychev4}(ct)
$$

- The procedure is validated by fitting the $B^\pm$ lifetime in the $B^\pm \rightarrow J/\psi K^\pm$ control channel, in eight different data taking periods, each roughly equivalent in statistics to the $B_s^0$ sample.

<table>
<thead>
<tr>
<th>Data set</th>
<th>$c\tau_{B^+} [\mu\text{m}]$</th>
<th>Pull w.r.t PDG [s.d.]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018A</td>
<td>$489.3 \pm 2.0$</td>
<td>$-0.4$</td>
</tr>
<tr>
<td>2018B</td>
<td>$495.7 \pm 2.7$</td>
<td>$+1.5$</td>
</tr>
<tr>
<td>2018C</td>
<td>$489.2 \pm 1.4$</td>
<td>$-1.4$</td>
</tr>
<tr>
<td>2018D</td>
<td>$493.2 \pm 1.3$</td>
<td>$+1.2$</td>
</tr>
<tr>
<td>2018</td>
<td>$492.78 \pm 0.97$</td>
<td>$+1.1$</td>
</tr>
<tr>
<td>2017A</td>
<td>$493.8 \pm 2.4$</td>
<td>$+1.0$</td>
</tr>
<tr>
<td>2017B</td>
<td>$494.8 \pm 3.5$</td>
<td>$+1.0$</td>
</tr>
<tr>
<td>2017C</td>
<td>$494.7 \pm 2.3$</td>
<td>$+1.4$</td>
</tr>
<tr>
<td>2017D</td>
<td>$489.5 \pm 1.7$</td>
<td>$-0.8$</td>
</tr>
<tr>
<td>2017</td>
<td>$492.9 \pm 1.1$</td>
<td>$+0.5$</td>
</tr>
</tbody>
</table>
Angular efficiency

• Detector acceptance and event selection lead to non uniform angular efficiency
• 3D angular efficiency is evaluated in bins of $\cos \theta_T$, $\cos \psi_T$ and $\phi_T$, separately for 2017 and 2018, using simulated samples
  • **Binning**: 70 bins for $\cos \theta_T$ and $\cos \psi_T$, and 30 for $\phi_T$
• The efficiency function is parameterized with spherical harmonics and Legendre polynomials up to order six
Flavour tagging
Tagging overview

- **Tagger:** opposite-side (OS) muon
  - Tagging feature: muon charge
  - The muon is selected already at trigger level → **very high efficiency**
- **Optimized** in $B^0_s \rightarrow J/\psi \phi$ simulated events and **calibrated** in data using $B^\pm \rightarrow J/\psi K^\pm$ self-tagging decays
- The **figure of merit** is the tagging power $P_{\text{tag}} = \epsilon_{\text{tag}} D_{\text{tag}}^2 = \epsilon_{\text{tag}} (1 - 2 \omega_{\text{tag}})^2$
  - $\epsilon_{\text{tag}} = N_{\text{tag}}/N_{\text{tot}}$, **tagging efficiency** ($N_{\text{tag}} = N_{\text{corr.tag}} + N_{\text{mistag}}$)
  - $\omega_{\text{tag}} = N_{\text{mistag}}/N_{\text{tag}}$, **mistag fraction**
- **Mistag probability is evaluated on per-event basis** with a dedicated Deep Neural Network
**OS-muon selection**

<table>
<thead>
<tr>
<th>Reconstruction</th>
<th>Global muon&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_T$</td>
<td>$\geq 2.0$ GeV</td>
</tr>
<tr>
<td>$</td>
<td>\eta</td>
</tr>
<tr>
<td>$IP_z$ w.r.t. PV</td>
<td>$\leq 1.0$ cm</td>
</tr>
<tr>
<td>$\Delta R_{\eta,\phi}$ wrt $B_S^0$</td>
<td>$\geq 0.4$</td>
</tr>
<tr>
<td>DNN vs fakes from hadrons</td>
<td>Loose WP&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> Global muon = reconstructed with information from both tracker and muon system

<sup>b</sup> $\epsilon$(muons) = 98%, $\epsilon$(hadrons) = 33% evaluated on Global muon candidates

- Reconstructed $b$ meson tracks **excluded**

- **Dedicated discriminator for soft muons**, trained with muons from simulated samples
  - Signal: genuine muon from $b$ hadron
  - Background: fake muons (mostly $K^{\pm}$, $\pi^{\pm}$)

- The muon selection is overall **loose** for maximum efficiency

- **Performance using the muon charge as tagging feature** (without per-event mistag)
  - $\epsilon_{\text{tag}} \sim 50$
  - $\omega_{\text{tag}} \sim 30$
  - $P_{\text{tag}} \sim 7$

Alberto Bragagnolo Measurement of $\phi_s$ in the $B_s \rightarrow J/\psi \phi$ channel at 13 TeV by CMS 10/20
Per-event mistag probability

- Per-event mistag probability enhances the total tagging performance.
- A fully connected **Deep Neural Network** is used to distinguish mistagged events and evaluate per-event mistag probability **at the same time**.
  - **Input features:** muon variables ($p_T, d_{xy}, \sigma_{d_{xy}}, \Delta R, ...$) and “cone” variables ($\text{Iso}_\mu, Q_{\text{cone}}, p_{T,\text{rel}}, \text{energy ratio, ...}$).
- The DNN is constructed in such a way that the **output score** $f_{\text{dnn}}$ is equal to the probability of tagging the event correctly.

\[ f_{\text{dnn}} = 1 - \omega_{\text{evt}} \]

**Figure:** schematic representation of a fully connected DNN.
Per-event mistag calibration

- $\omega_{\text{evt}}$ is calibrated in data with self-tagging $B^{\pm} \rightarrow J/\psi K^{\pm}$ decays with a linear function

$$w_{\text{fit}} = a + b \cdot w_{\text{evt}}$$

**2017 tagging calibration**

- CMS Preliminary 2017
- $36.7 \text{ fb}^{-1}$ (13 TeV)
- $a = -0.001 \pm 0.004$
- $b = 1.01 \pm 0.01$

**2018 tagging calibration**

- CMS Preliminary 2018
- $59.7 \text{ fb}^{-1}$ (13 TeV)
- $a = 0.003 \pm 0.003$
- $b = 1.01 \pm 0.01$

- Excellent agreement between prediction and measurement
- DNN and calibration are stable $\rightarrow$ very small systematic uncertainties
Tagging performance

- Tagging performances evaluated in $B^{\pm} \rightarrow J/\psi K^{\pm}$ data

<table>
<thead>
<tr>
<th>Data set</th>
<th>$\epsilon_{\text{tag}}$</th>
<th>$\omega_{\text{tag}}$</th>
<th>$P_{\text{tag}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017</td>
<td>$(45.7 \pm 0.1)%$</td>
<td>$(27.1 \pm 0.1)%$</td>
<td>$(9.6 \pm 0.1)%$</td>
</tr>
<tr>
<td>2018</td>
<td>$(50.9 \pm 0.1)%$</td>
<td>$(27.3 \pm 0.1)%$</td>
<td>$(10.5 \pm 0.1)%$</td>
</tr>
<tr>
<td>Run-1</td>
<td>$(8.31 \pm 0.03)%$</td>
<td>$(30.2 \pm 0.3)%$</td>
<td>$(1.31 \pm 0.03)%$</td>
</tr>
</tbody>
</table>

- **High efficiency** due to the additional muon required at trigger level
- **Low dilution** thanks to the DNN based per-event mistag probability
- Final performance, normalized by the event rate, $\sim 50\%$ higher w.r.t. Run-1
Maximum likelihood fit and results
Fit model

\[ P = N_{\text{sgn}} P_{\text{sgn}} + N_{\text{bkg}} P_{\text{bkg}} + N_{\text{peak}} P_{\text{peak}} \]

\[ P_{\text{sgn}} = \epsilon(\text{ct}) \epsilon(\Theta) [f(\Theta, \text{ct}, \alpha) \otimes G(\text{ct}, \sigma_{\text{ct}})] P_{\text{sgn}}(m_{B_0^s}) P_{\text{sgn}}(\sigma_{\text{ct}}) P_{\text{sgn}}(\xi) \]

- \( \epsilon(\text{ct}) \epsilon(\Theta) \): efficiency functions
- \( f(\Theta, \text{ct}, \alpha) \): differential decay rate PDF
- \( G(\text{ct}, \sigma_{\text{ct}}) \): Gaussian resolution function
- \( P(m_{B_0^s}) \): mass PDFs
- \( P(\sigma_{\text{ct}}) \): decay length uncertainty PDFs
- \( P(\xi) \): tag distribution

\[ P_{\text{bkg}} = P_{\text{bkg}}(\cos \theta_T, \phi_T) P_{\text{bkg}}(\cos \psi_T) P_{\text{bkg}}(\text{ct}) P_{\text{bkg}}(m_{B_0^s}) P_{\text{bkg}}(\sigma_{\text{ct}}) P_{\text{bkg}}(\xi) \]

- \( P_{\text{bkg}}(\cos \theta_T, \phi_T), P_{\text{bkg}}(\cos \psi_T), P_{\text{bkg}}(\text{ct}) \): background angular and lifetime PDFs

\[ P_{\text{peak}} = P_{\text{peak}}(\cos \theta_T, \phi_T) P_{\text{peak}}(\cos \psi_T) P_{\text{peak}}(\text{ct}) P_{\text{peak}}(m_{B_0^s}) P_{\text{peak}}(\sigma_{\text{ct}}) P_{\text{peak}}(\xi) \]

- \( P_{\text{peak}} \) models the **peaking background** from \( B^0 \rightarrow J/\psi K^0 \rightarrow \mu^+ \mu^- K^+ \pi^- \) where the pion is misidentified as a kaon
- Peaking background from \( \Lambda_b \rightarrow J/\psi Kp \) estimated to be negligible
## Systematic uncertainties

| Source                              | $|A_0|^2$ | $|A_\perp|^2$ | $|A_S|^2$ | $\delta_{||}$ [rad] | $\delta_{\perp}$ [rad] | $\delta_{S\perp}$ [rad] | $\Gamma_S$ [ps$^{-1}$] | $\Delta\Gamma_S$ [ps$^{-1}$] | $\Delta m_S$ [m$^{-1}$] | $|\lambda|$ | $\phi_S$ |
|-------------------------------------|---------|---------------|---------|---------------------|---------------------|---------------------|-----------------|---------------------|---------------------|----------|----------|
| Model bias                          | 0.0002  | 0.0012        | 0.0008  | 0.020               | 0.016               | 0.006               | 0.0005          | 0.0019             | -                   | 0.0035   | 7.9      |
| Angular efficiency                  | 0.0008  | 0.0010        | 0.0015  | 0.006               | 0.015               | 0.015               | 0.0002          | 0.0006             | 0.007              | 0.0057   | 3.8      |
| Lifetime efficiency                 | 0.0014  | 0.0023        | 0.0007  | 0.001               | 0.002               | 0.002               | 0.0062          | 0.001              | 0.0002             | 0.0002   | 0.3      |
| Lifetime resolution                 | 0.0007  | 0.0009        | 0.0065  | 0.006               | 0.025               | 0.022               | 0.0005          | 0.0008             | 0.004              | 0.0003   | 0.6      |
| Data-MC mismatch                    | 0.0044  | 0.0029        | 0.0065  | 0.007               | 0.007               | 0.028               | 0.0003          | 0.0008             | 0.004              | 0.0003   | 0.6      |
| Flavour tagging                     | 0.0003  | $<10^{-4}$    | $<10^{-4}$ | 0.001       | 0.003       | 0.001       | $<10^{-4}$ | $<10^{-4}$         | 0.001              | 0.0002   | 0.1      |
| Unfitted $\omega_{\text{evt dist.}}$ | -       | 0.0008        | -       | -                   | -                   | -                   | 0.006           | 0.0005             | -                   | -        | 3.0      |
| Model assumptions                   | -       | 0.0013        | 0.0012  | 0.017               | 0.019               | 0.011               | 0.0003          | -                 | -                   | -        | 0.0046   |
| Peaking background                  | 0.0005  | 0.0002        | 0.0025  | 0.005               | 0.007               | 0.011               | 0.0002          | 0.0008             | 0.011              | $<10^{-4}$ | 0.3      |
| Total syst.                          | 0.0048  | 0.0044        | 0.0097  | 0.028               | 0.040               | 0.043               | 0.0024          | 0.0066             | 0.020              | 0.0082   | 9.6      |

### Leading systematic uncertainties for the most interesting parameters

- $\phi_S \rightarrow$ model bias and angular efficiency
- $\Delta \Gamma_S \rightarrow$ lifetime efficiency
- $\Gamma_S \rightarrow$ lifetime efficiency
- $\Delta m_S \rightarrow$ lifetime resolution and peaking background model
- $|\lambda| \rightarrow$ angular efficiency and model assumptions
### Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_s$ [mrad]</td>
<td>$-11 \pm 50$</td>
<td>$\pm 10$</td>
<td></td>
</tr>
<tr>
<td>$\Delta \Gamma_s$ [ps$^{-1}$]</td>
<td>$0.114 \pm 0.014$</td>
<td>$\pm 0.007$</td>
<td></td>
</tr>
<tr>
<td>$\Gamma_s$ [ps$^{-1}$]</td>
<td>$0.6531 \pm 0.0042$</td>
<td>$\pm 0.0024$</td>
<td></td>
</tr>
<tr>
<td>$\Delta m_s$ [$\hbar$ ps$^{-1}$]</td>
<td>$17.51 \pm 0.10$</td>
<td>$\pm 0.02$</td>
<td></td>
</tr>
<tr>
<td>$</td>
<td>\lambda</td>
<td>$</td>
<td>$0.972 \pm 0.026$</td>
</tr>
<tr>
<td>$</td>
<td>A_0</td>
<td>^2$</td>
<td>$0.5350 \pm 0.0047$</td>
</tr>
<tr>
<td>$</td>
<td>A_\perp</td>
<td>^2$</td>
<td>$0.2337 \pm 0.0063$</td>
</tr>
<tr>
<td>$</td>
<td>A_S</td>
<td>^2$</td>
<td>$0.022 \pm 0.008$</td>
</tr>
<tr>
<td>$\delta_{\parallel}$ [rad]</td>
<td>$3.18 \pm 0.12$</td>
<td>$\pm 0.03$</td>
<td></td>
</tr>
<tr>
<td>$\delta_{\perp}$ [rad]</td>
<td>$2.77 \pm 0.16$</td>
<td>$\pm 0.04$</td>
<td></td>
</tr>
<tr>
<td>$\delta_{S\perp}$ [rad]</td>
<td>$0.221 \pm 0.083$</td>
<td>$\pm 0.070$</td>
<td></td>
</tr>
</tbody>
</table>

- $\phi_s$ and $\Delta \Gamma_s$ are in agreement with the SM:
  
  $\phi_s^{SM} = -36.96^{+0.84}_{-0.72}$ mrad
  
  $\Delta \Gamma_s^{SM} = 0.087 \pm 0.021$ ps$^{-1}$

- $\Gamma_s$ is consistent with the world average:
  
  $\Gamma_s^{WA} = 0.6623 \pm 0.0018$ ps$^{-1}$

- $\Delta m_s$ is consistent with the world average:
  
  $\Delta m_s^{WA} = 17.757 \pm 0.021 \hbar$ ps$^{-1}$

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

- This is the first measurement by CMS of $\Delta m_s$ and $|\lambda|$
Combination with 8 TeV results

• The results of this analysis are in agreement with the ones obtained by CMS at $\sqrt{s} = 8$ TeV [Phys.Lett.B757(2016)97] and therefore combined

• All systematic uncertainties are considered uncorrelated

• The results are in agreement with the SM predictions

$\phi_s = -21 \pm 45 \text{ mrad}$

$\Delta \Gamma_s = 0.1074 \pm 0.0097 \text{ ps}^{-1}$

• The new trigger strategy, which trades number of events for tagging power, pays off for $\phi_s$ while does not improve $\Delta \Gamma_s$, which sensitivity is driven by statistics
Conclusions
Summary

• The CPV phase $\phi_s$ and the decay width difference $\Delta \Gamma_s$ are measured using 48,500 $B^0_s \to J/\psi \phi$ candidates collected at $\sqrt{s} = 13$ TeV, corresponding to $L_{\text{int}} = 96.4$ fb$^{-1}$.

• Events are selected using a non displaced trigger that required an additional muon, which is exploited to infer the flavor of the $B^0_s$.
  • This strategy paid off in terms of tagging performance, leading to a significant reduction of the $\phi_s$ uncertainty.
  • However, the limited number of selected events prevented improvements on $\Delta \Gamma_s$.

• A novel opposite-side muon tagger based on Deep Neural Network has been developed to directly predict mistag probability on per-event basis, achieving $P_{\text{tag}} \sim 10\%$.

• Results from this analysis are combined with those obtained at $\sqrt{s} = 8$ TeV yielding

$$\phi_s = -21 \pm 45 \text{ mrad}$$
$$\Delta \Gamma_s = 0.1074 \pm 0.0097 \text{ ps}^{-1}$$

• Results are consistent with the Standard Model predictions

$$\phi_{s}^{\text{SM}} = -36.96^{+0.84}_{-0.72} \text{ mrad} \quad \Delta \Gamma_{s}^{\text{SM}} = 0.087 \pm 0.021 \text{ ps}^{-1}$$
Outlook

Comparison with other LHC experiments in the $B_s^0 \rightarrow J/\psi K^+K^-$ channel

<table>
<thead>
<tr>
<th></th>
<th>$\phi_s$ [mrad]</th>
<th>$\Delta \Gamma_s$ [ps$^{-1}$]</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMS</td>
<td>$-21 \pm 45$</td>
<td>$0.1074 \pm 0.0097$</td>
<td>CMS-PAS-BPH-20-001</td>
</tr>
<tr>
<td>ATLAS</td>
<td>$-87 \pm 42$</td>
<td>$0.0640 \pm 0.0048$</td>
<td>CERN-EP-2019-218</td>
</tr>
<tr>
<td>LHCb</td>
<td>$-81 \pm 32$</td>
<td>$0.0777 \pm 0.0062$</td>
<td>EUR.PHYS.J.C79(2019)706</td>
</tr>
<tr>
<td>SM</td>
<td>$-36.96^{+0.84}_{-0.72}$</td>
<td>$0.087 \pm 0.021$</td>
<td>CKMfitter, 1102.4274</td>
</tr>
</tbody>
</table>

- All of the above are combination of Run-1 and partial Run-2 results
- Uncertainties are presented as the stat.+syst. squared sum
- LHCb results refer to the combination of measurements around the $\phi(1020)$ resonance
- New $\Delta \Gamma_s$ prediction with smaller uncertainties available: $\Delta \Gamma_s^{SM} = 0.091 \pm 0.013$ ps$^{-1}$ [1912.07621]

- $\Delta \Gamma_s$ shows tensions between experiments
- Full Run-2 measurements will clarify the situation

Future plans

- CMS plans to analyze the full Run-2 dataset, adding a complementary trigger that requires a displaced $J/\psi$ plus two charged tracks
  - Electron and jet flavour tagging algorithms will be used
- Effective statistics $N(B_s^0) \cdot P_{\text{tag}}$ expected to improve by a factor $1.5 \sim 2.0$
Thanks for your attention!
Decay rate model

\[
\frac{d^4 \Gamma(B^0_s(t))}{d\Theta dt} = \sum_{i=1}^{10} O_i(\alpha, t) \cdot g_i(\Theta)
\]

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

\[
\begin{array}{c|c|c|c|c|c|c|c|}
    i & g_i(\theta_T, \psi_T, \varphi_T) & N_i & a_i & b_i & c_i & d_i \\
    \hline
    1 & 2 \cos^2 \psi_T (1 - \sin^2 \theta_T \cos^2 \varphi_T) & |A_0|^2 & 1 & D & C & -S \\
    2 & \sin^2 \psi_T (1 - \sin^2 \theta_T \sin^2 \varphi_T) & |A_\parallel|^2 & 1 & D & C & -S \\
    3 & \sin^2 \psi_T \sin^2 \theta_T & |A_\perp|^2 & C \sin (\delta_\parallel - \delta_\perp) & -D & C & S \\
    4 & -\sin^2 \psi_T \sin 2\theta_T \sin \varphi_T & |A_\parallel||A_\perp| & \cos (\delta_\parallel - \delta_0) & S \cos (\delta_\parallel - \delta_\perp) & \sin (\delta_\perp - \delta_\parallel) & D \cos (\delta_\parallel - \delta_0) \sin (\delta_\perp - \delta_0) & -S \cos (\delta_\parallel - \delta_0) \\
    5 & \frac{1}{\sqrt{2}} \sin 2\psi_T \sin^2 \theta_T \sin 2\varphi_T & |A_\parallel||A_\parallel| & C \sin (\delta_\parallel - \delta_0) & S \cos (\delta_\parallel - \delta_\perp) & \sin (\delta_\perp - \delta_\parallel) & D \cos (\delta_\parallel - \delta_0) \sin (\delta_\perp - \delta_0) & -S \cos (\delta_\parallel - \delta_0) \\
    6 & \frac{1}{\sqrt{2}} \sin 2\psi_T \sin 2\theta_T \cos \varphi_T & |A_0||A_\parallel| & C \sin (\delta_\parallel - \delta_0) & S \cos (\delta_\parallel - \delta_\perp) & \sin (\delta_\perp - \delta_\parallel) & D \cos (\delta_\parallel - \delta_0) \sin (\delta_\perp - \delta_0) & -S \cos (\delta_\parallel - \delta_0) \\
    7 & \frac{1}{3} \sqrt{6} \sin \psi_T \sin^2 \theta_T \sin 2\varphi_T & |A_\parallel||A_\perp| & C \cos (\delta_\parallel - \delta_S) & \sin (\delta_\perp - \delta_\parallel) & -D \sin (\delta_\parallel - \delta_S) & C \sin (\delta_\parallel - \delta_\perp) & D \sin (\delta_\parallel - \delta_S) \\
    8 & \frac{1}{3} \sqrt{6} \sin \psi_T \sin 2\theta_T \cos \varphi_T & |A_S||A_\parallel| & C \cos (\delta_\parallel - \delta_S) & \sin (\delta_\perp - \delta_\parallel) & -D \sin (\delta_\parallel - \delta_S) & C \sin (\delta_\parallel - \delta_\perp) & D \sin (\delta_\parallel - \delta_S) \\
    9 & \frac{4}{3} \sqrt{3} \cos \psi_T (1 - \sin^2 \theta_T \cos^2 \varphi_T) & |A_S||A_0| & C \cos (\delta_0 - \delta_S) & \sin (\delta_\perp - \delta_\parallel) & -D \sin (\delta_\parallel - \delta_S) & C \sin (\delta_\parallel - \delta_\perp) & D \sin (\delta_0 - \delta_S) \\
    10 & \frac{1}{3} \sqrt{3} \cos \psi_T (1 - \sin^2 \theta_T \cos^2 \varphi_T) & |A_S||A_0| & C \cos (\delta_0 - \delta_S) & \sin (\delta_\perp - \delta_\parallel) & -D \sin (\delta_\parallel - \delta_S) & C \sin (\delta_\parallel - \delta_\perp) & D \sin (\delta_0 - \delta_S) \\
    \end{array}
\]

\[
C = \frac{1 - |\lambda|^2}{1 + |\lambda|^2} \rightarrow \text{Sensitive to direct CPV}
\]

\[
S = -\frac{2|\lambda| \sin \phi_s}{1 + |\lambda|^2} \rightarrow \text{Sensitive to small } \phi_s
\]

\[
D = -\frac{2|\lambda| \cos \phi_s}{1 + |\lambda|^2}
\]
Angular efficiency

Computed separately for 2017 and 2018 using the “projection” method

1. **Construct efficiency histograms**
   - Numerator: 3D angular RECO histograms from $\Delta \Gamma_s = 0$ MC samples
   - Denominator: 3D angular GEN histograms from GEN only sample
   - Binning: 70 bins for $\cos \theta_T$ and $\cos \psi_T$, and 30 for $\phi_T$

2. **Project on Legendre orthogonal basis**

   \[ b_{l,k,m}(\Theta) = P_l^m(\cos \theta_T) \cdot P_k^m(\cos \psi_T) \cdot \begin{cases} 
   \sin(m \phi_T) & \text{if } m < 0 \\
   \cos(m \phi_T) & \text{if } m > 0 \\
   1/2 & \text{if } m = 0 
\end{cases} \]

   - up to order 6

3. **Construct angular efficiency as**

   \[ \epsilon(\Theta) = \sum_{l,k,m} c_{l,k,m} \cdot b_{l,k,m}(\Theta) \]

   - $c_{l,k,m}$ are the projection coefficients
Deep neural network for flavour tagging

- **Training features**
  - Muon variables: $p_T$, $\eta$, $d_{xy}$, $\sigma(d_{xy})$, $d_z$, $\sigma(d_z)$, $\Delta R(\mu, B^0_s)$, DNN vs hadron fakes score
  - Cone variables: $\text{Iso}_{\mu}$, $Q_{\text{cone}}$, $p_{T,\text{rel}}$, $p_{T,\text{cone}}$, $\Delta R(\mu, \text{cone})$, $E_\mu/E_{\text{cone}}$

- **Architecture: fully connected**
  - 3 layers of 200 neurons
  - ReLU activation
  - 40% dropout probability

- **Loss**: categorical crossentropy

- **Optimizer**: Adam