An alternative paradigm of time-dependent fits at Belle II

Vladimir Chekelian   Max Planck Institute for Physics, Munich

In the traditional approach for time-dependent fits (used at Belle and BaBar) the maximum of the unbinned likelihood is used

$$\max \ \boldsymbol{L} = \prod_i \boldsymbol{P^i}_{ev}(\Delta \boldsymbol{t}) \quad with \quad \boldsymbol{P^i}_{ev}(\Delta \boldsymbol{t}) = \int_{-\infty}^{+\infty} \boldsymbol{d}(\Delta \boldsymbol{t}^{true}) \ \boldsymbol{P}_{theory}(\Delta \boldsymbol{t}^{true}) \ \boldsymbol{R}(\Delta \boldsymbol{t} - \Delta \boldsymbol{t}^{true})$$

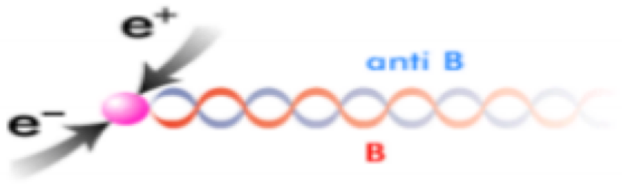
assuming that ***theory and detector effects are independent*** *and the* Δt *resolution function* $\boldsymbol{R}$ *does not depend on true* Δt^{true}

- this "independence" is not granted at Belle II (see slides below)

- a new paradigm of time-dependent fits is proposed which could deal with correlations between theory and detector effects:

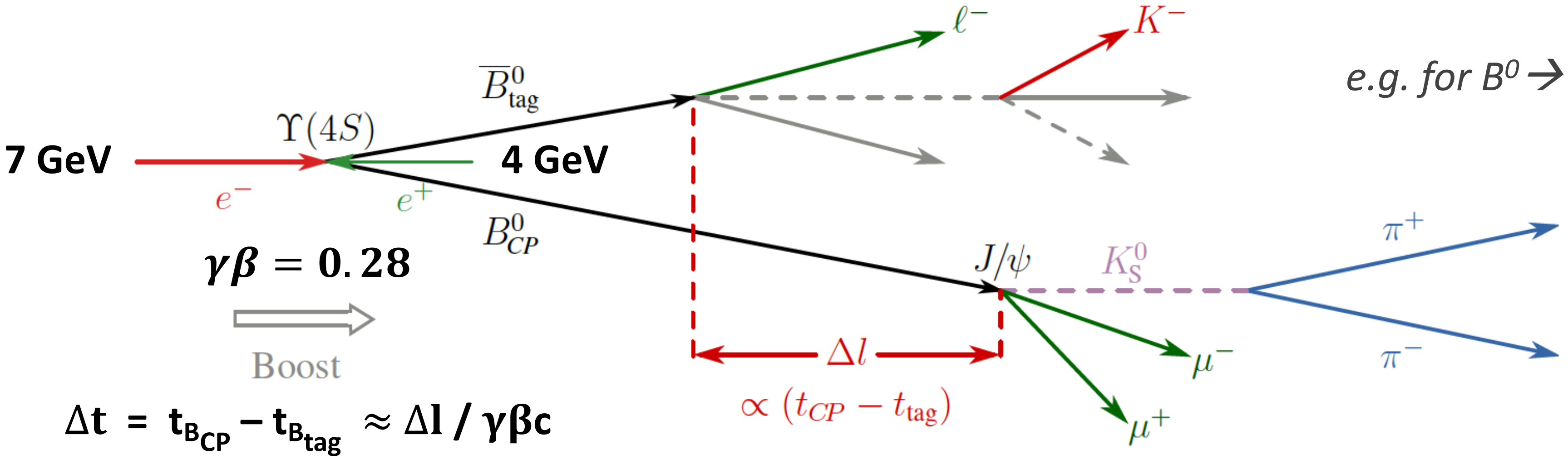
VC, "The MPI Concept of Time-Dependent Fits at Belle II", BELLE2-NOTE-PH-2019-023
VC, "The MPI Concept of Time-Dependent Fits at Belle II", xFitter Workshop, Minsk, March 2019
https://indico.desy.de/indico/event/22011/session/7/contribution/24/material/slides/0.pdf
VC, "First look at the time-dependent CP violation using early Belle II data", Lomonosov conference, Moscow, August 2019
https://yadi.sk/i/aOVi-ybK0pwmKA
+ presentations of VC at the TDCPV WG meetings

Time-dependent CP violation analyses at the asymmetric B-factories



*The $B^0\overline{B^0}$ pairs from $\Upsilon(4S)$ are produced in a **coherent, entangled quantum mechanical state.***
When $B^0(\overline{B^0})$ decays, the flavor wavefunction of other $\overline{B^0}(B^0)$ collapses and it propagates alone.
One needs to measure decay times of both B^0s to observe CP violation.

e.g. for $B^0 \to J/\psi\, K_S^0$ (ξ_{CP} =-1)



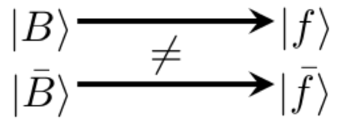
$\Delta t = t_{B_{CP}} - t_{B_{tag}} \approx \Delta l / \gamma\beta c$

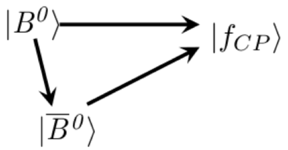
B^0: CP violation, mixing and lifetime are coming together and can be found in the fit to Δt distribution in data

$$P_{sig}(\Delta t) = \frac{exp\left(-\frac{|\Delta t|}{\tau}\right)}{4\tau}\left(1 + q_{tag}\left(A_{CP}\cos(\Delta m\,\Delta t) + S_{CP}\sin(\Delta m\,\Delta t)\right)\right)$$

A_{CP}: direct CPV *$S_{CP} = \sin(2\varphi_1)$ mixing-induced CPV*

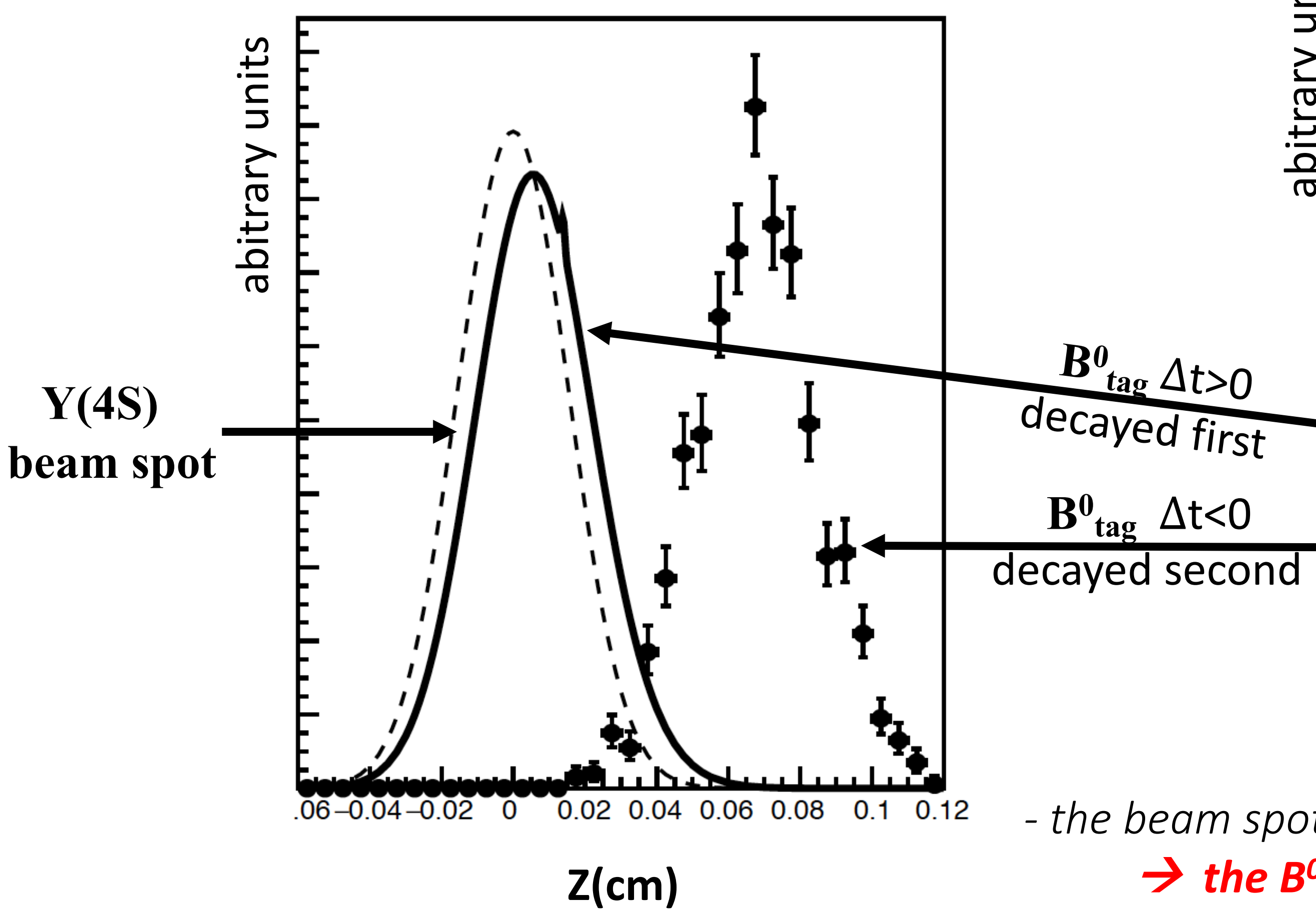
$q_{tag} = -\xi_{CP}$ for B^0_{tag}
$q_{tag} = \xi_{CP}$ for $\bar{B}^0_{tag}$

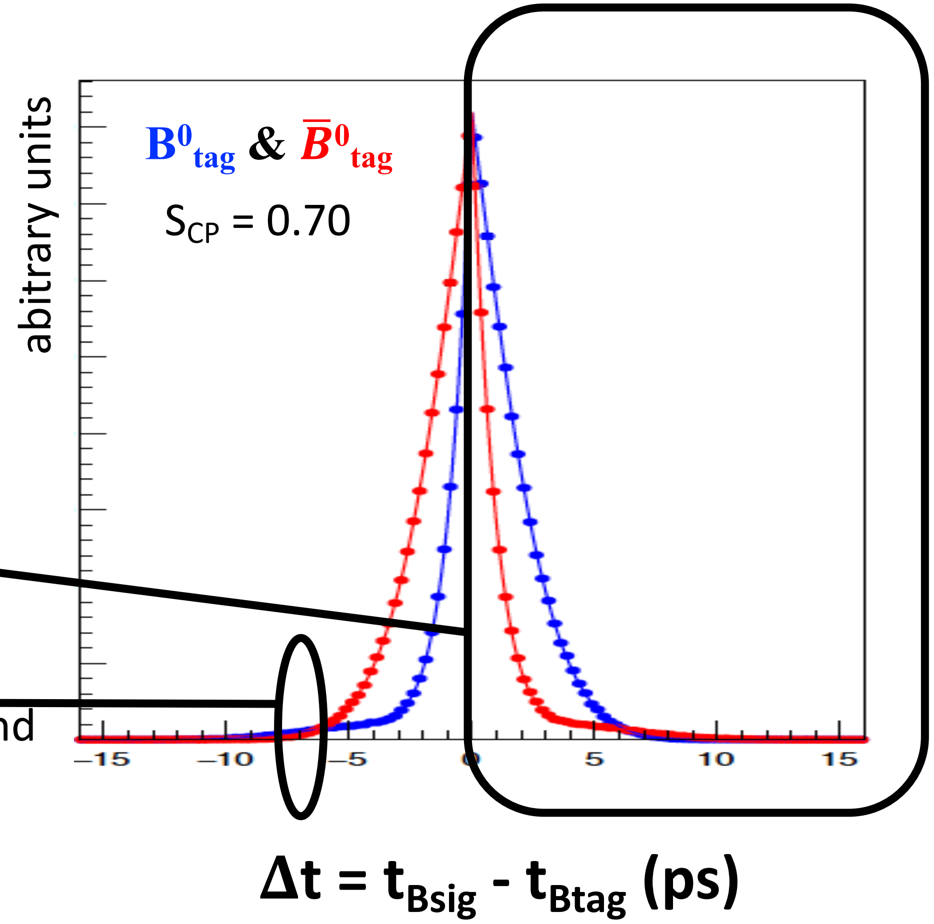




A new feature at Belle II – the tiny size of the beam spot

$e^+e^- \rightarrow \Upsilon(4S) \rightarrow B^0_{tag} + B^0_{sig}(J/\psi\, K_S^0)$





- *the beam spot at Belle II (~400 μm in z) is comparable with the B^0 lifetime*

→ the B^0 decay position in z is far away from the beam spot in the tails of the Δt distribution

New challenge for the time-dependent analysis

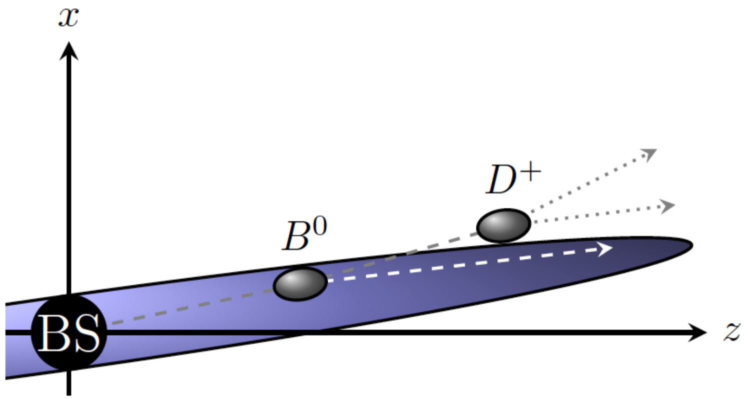
In the traditional approach for time-dependent fits (used at Belle and BaBar) the maximum of the unbinned likelihood is used

$$\max \quad L = \prod_i P^i_{ev}(\Delta t) \quad \text{with} \quad P^i_{ev}(\Delta t) = \int_{-\infty}^{+\infty} d(\Delta t^{true})\, P_{theory}(\Delta t^{true})\, R(\Delta t - \Delta t^{true})$$

assuming that ***theory and detector effects are independent*** *and the Δt resolution function* $\boldsymbol{R}$ *does not depend on true Δt^{true}*

→ ***this is not granted at Belle II because of the tiny size of the beam spot, excellent precision of PXD and a need to make use the beam spot information for improvement of the B^0_{tag} vertex position on the tag side:***

the beam information helps to select tracks directly from B^0_{tag} decay and remove displaced tracks from decays of charmed particles (Ds) or K_S^0



Several developments at Belle II to deal with the new challenge

1. *Optimisation of the B^0_{tag} vertexing on the tag side, e.g. "Btube constraint" (D. Sourav) applied to tag vertexing (T. Humair)*
2. *Further development of the traditional approach*
3. ***A new paradigm of time-dependent fits***

which is robust and could deal with correlations between theory and detector effects →

An alternative paradigm of time-dependent fits at Belle II

- $\boldsymbol{P^{i}}_{\boldsymbol{ev}}(\Delta \boldsymbol{t})$ → *calculated numerically using weighted MC events (i.e. use convolution of theory and detector from simulation)*

- ***variation of input physics parameters (τ, Δm, S_{CP} and A_{CP})*** → *by weighting of an auxiliary, "assistive" MC sample*

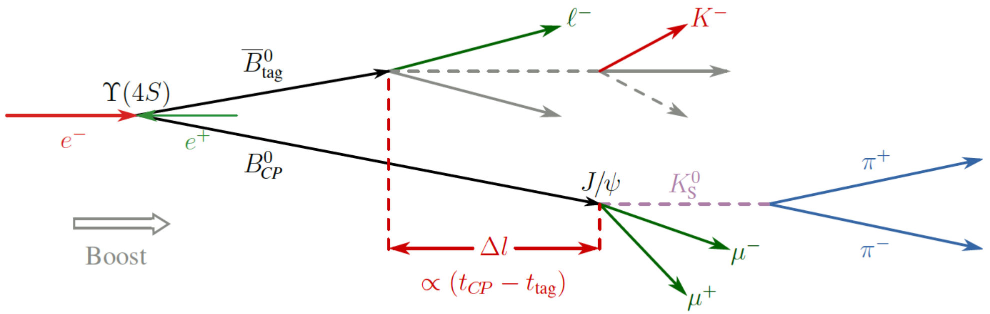
- ***differences in the detector response between data and simulation*** → *by downgrading (smearing) of the detector response in an auxiliary, "assistive" MC sample, using weighting of the simulated event*

- ***physics parameters and the detector smearing*** → *determined simultaneously in the TD CPV fit of the signal and control channels*

New input physics parameters – analytic expression for weighting of MC

Input from generator to simulation:

$$P_{theory}\left(t^{\,B^0 first}, \Delta t\right) = \frac{exp\left(-\frac{\left|t^{\,B^0 first}\right|}{\tau}\right)}{4\tau}\,\frac{exp\left(-\frac{|\Delta t|}{\tau}\right)}{4\tau}\left[1 + q\left(A\cos(\delta m\,\Delta t) + S\sin(\delta m\,\Delta t)\right)\right]$$



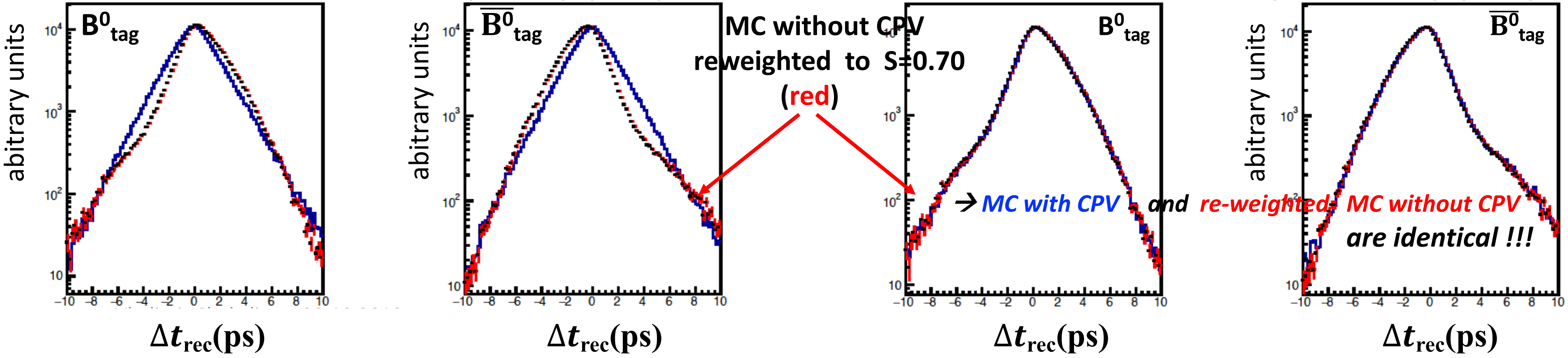
If values of $t^{B^0 first}$ and Δt are defined (and frozen):

→ *simulation of the detector effects does not depend on* $\boldsymbol{\tau, \delta m, A, S}$

→ *only probability of event with given* $\boldsymbol{t^{B^0 first}}$ *and* $\boldsymbol{\Delta t}$ *depends on* $\boldsymbol{\tau, \delta m, A, S}$

→ *Thus, MC sample generated with* $\boldsymbol{\tau_0, \delta m_0, A_0, S_0}$ *can be used to get MC sample equivalent to simulation with* $\boldsymbol{\tau, \delta m, A, S}$ *by the weighting of MC events* $\boldsymbol{w = P_{theory}\left(t^{\,B^0 first}, \Delta t; \tau, \delta m, A, S\right) \,/\, P_{theory}\left(t^{\,B^0 first}, \Delta t; \tau_0, \delta m_0, A_0, S_0\right)}$

simulation without CPV (S=A=0) (blue) **simulation with CPV (S=0.70,A=0) (blue)**



Treatment of discrepancies between data and MC in the detector response

prior the TD fit (once) **→** *smearing of reconstructed quantities (Δt) in the MC sample*

- *very flexible and can have any level of complexity if there is a model for downgrading of the detector resolution*
- *the simplest and also very efficient smearing model:* $\Delta t' = \Delta t + G(\alpha_{smear} \cdot \delta(\Delta t))$ ($\delta(\Delta t)$ – uncertainty of Δt)

during the TD fit (many times) **→** *approximation of the "simplest smearing model" by the weighting of the MC sample*

- *could be directly included in the TD fit with the smearing factor and the physics parameters determined simultaneously*

First, determine a simplified Δt resolution function in a two gaussian fit of the simulated MC sample:

$$P_{res.func}(\Delta t - \Delta t_{true}) = f\, G_1(\mu_1, \sigma_1) + (1-f)\, G_2(\mu_2, \sigma_2)$$

where $\mu_i = \mu_i^{pull} \cdot \delta(\Delta t)$ *and* $\sigma_i = \sigma_i^{pull} \cdot \delta(\Delta t)$

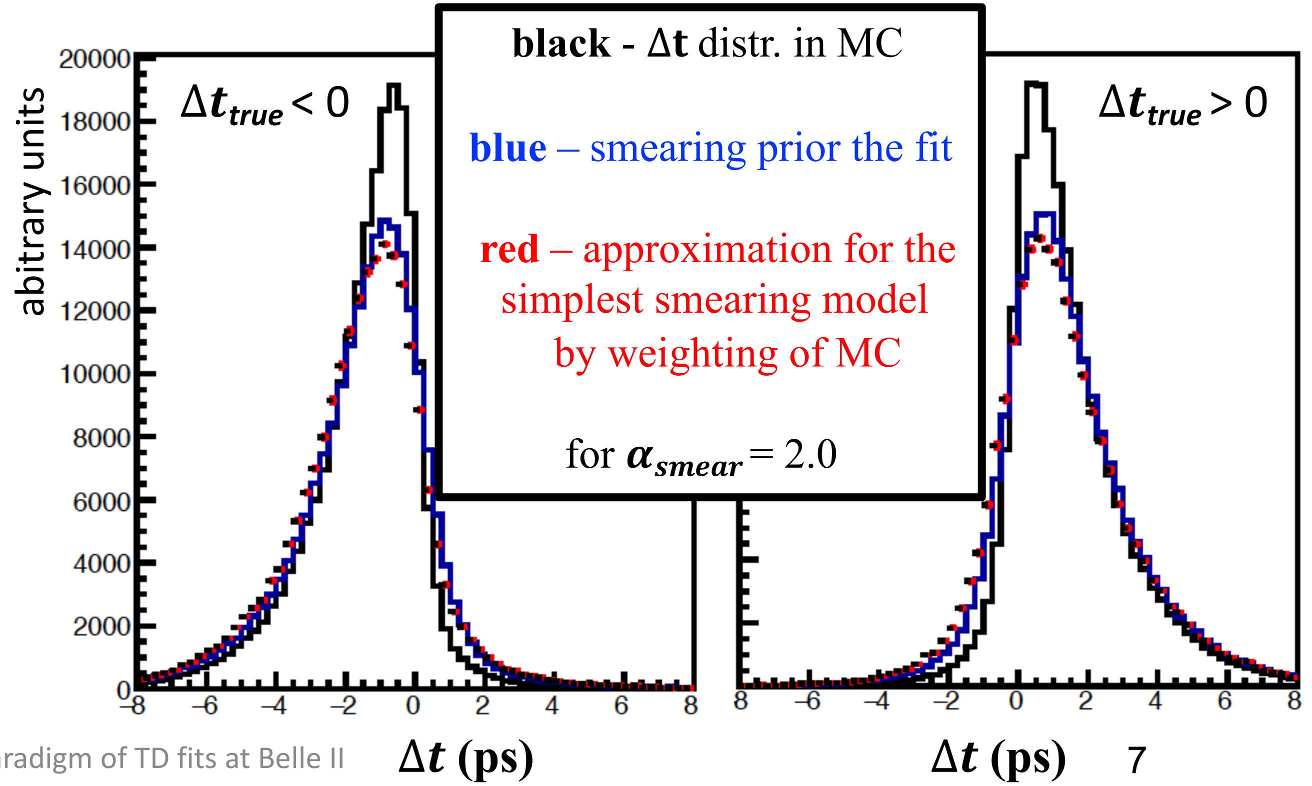
determined separately for positive and negative Δt_{true}

Then, analytic expression for weighting of MC events

$$w = P_{res.func}^{new} / P_{res.func}, \quad \text{where}$$

$$P_{res.func}^{new} = f\, G_1(\mu_1, \sigma_1^{new}) + (1-f)\, G_2(\mu_2, \sigma_2^{new})$$

with $\sigma_i^{new} = \sqrt{\sigma_i^2 + (\alpha_{smear}\, \delta(\Delta t))^2}$



An example of alternative time-dependent CP violation fit of signal $B^0 \rightarrow J/\psi\ K_S^0$ and control $B^\pm \rightarrow J/\psi\ K^\pm$ channels for Belle II MC (500 fb^{-1})

	$\tau(B^\pm)\ ps$	$\tau(B^0)\ ps$	S	$\delta m(ps^{-1})$	α_{smear}
$ass.MC(BGx0)\ J/\psi(\mu\mu)K_S^0$		1.525	0	0	
$ass.MC(BGx0)\ J/\psi(ee)K_S^0$		1.525	0	0	
$ass.MC(BGx1)\ J/\psi(\mu\mu)K^\pm$	1.637				
$expected$	1.637	1.525	0.695	0.502	
$MC12b(BGx1)\ J/\psi(\mu\mu, ee)K_S^0$		1.554 ± 0.037	0.700 ± 0.059	0.536 ± 0.048	0 ± 0.63
$MC12b(BGx1)\ J/\psi(\mu\mu)K^\pm$	1.596 ± 0.036				0 ± 0.44
$MC12b(BGx1)\ combined$	1.596 ± 0.036	1.554 ± 0.037	0.701 ± 0.059	0.536 ± 0.048	0 ± 0.36

→ all results of the alternative TD CPV fits are consistent with the expectations within one sigma

Conclusions

- *Large room for improvements of precision for the TD CPV measurements, by far not limited by systematics*
  - a long term Belle II project aiming for 50 ab^{-1}

- *Ultimate precision will require best methods for time-dependent analyses*

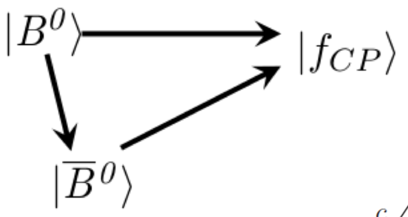
- *New challenges and new developments related to TD CPV analyses at Belle II*
  - precision measurement
  - possible correlations of physics parameters and detector effects, e.g. due to the tiny size of the beam spot
  → optimization of vertexing on tag side
  → traditional & alternative approaches for TD fits

Time-dependent CP violation and the CKM unitarity matrix

- *B^0 system has the largest CP violation effects in the SM described by the CKM unitarity matrix V_{CKM}*

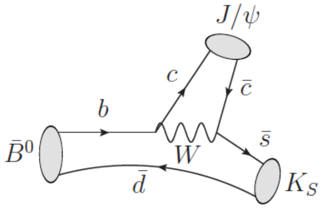
$$V_{CKM} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix}$$

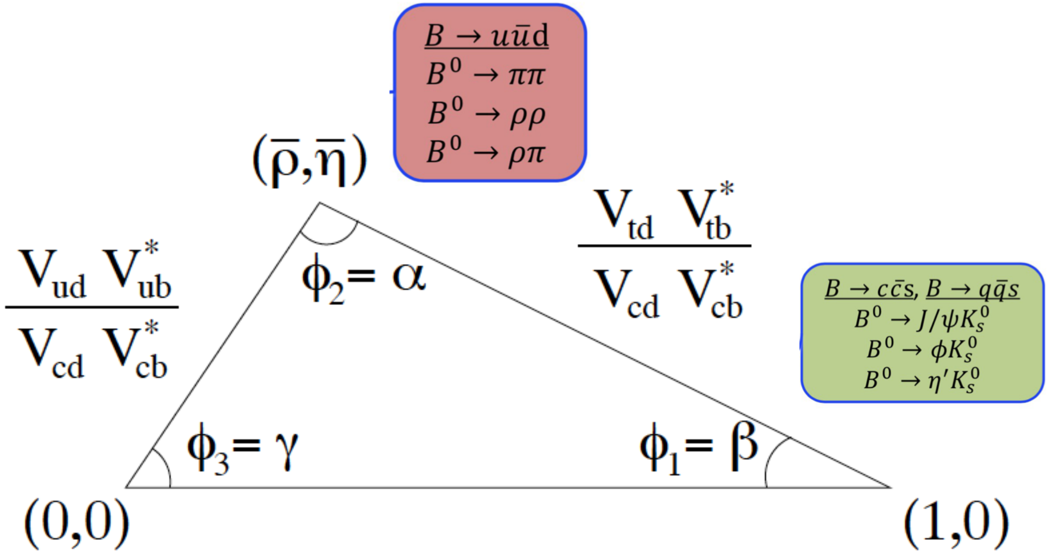
- *time-dependent CPV effects are related to interference between mixing and decay amplitudes (mixing-induced CPV)*



$$A(t) = \frac{\Gamma(\bar{B}^0 \to f_{CP}) - \Gamma(B^0 \to fCP)}{\Gamma(\bar{B}^0 \to f_{CP}) + \Gamma(B^0 \to fCP)} = -\xi_f \sin(2\varphi_1)\sin(\Delta m\, t), \qquad S_{CP} = \sin(2\varphi_1)$$

- *φ_1 and φ_2 can be measured in TD CPV analyses of for $B^0 \to c\bar{c}s$, $q\bar{q}s$ and $B^0 \to u\bar{u}d$, e.g. for $B^0 \to J/\psi\, K_S^0$*

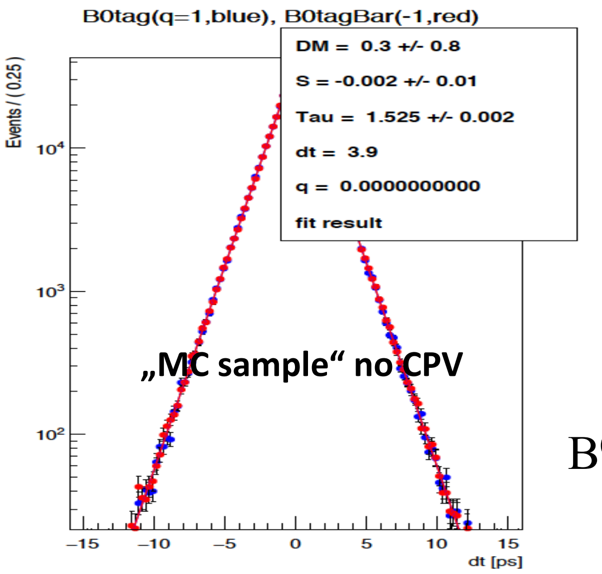




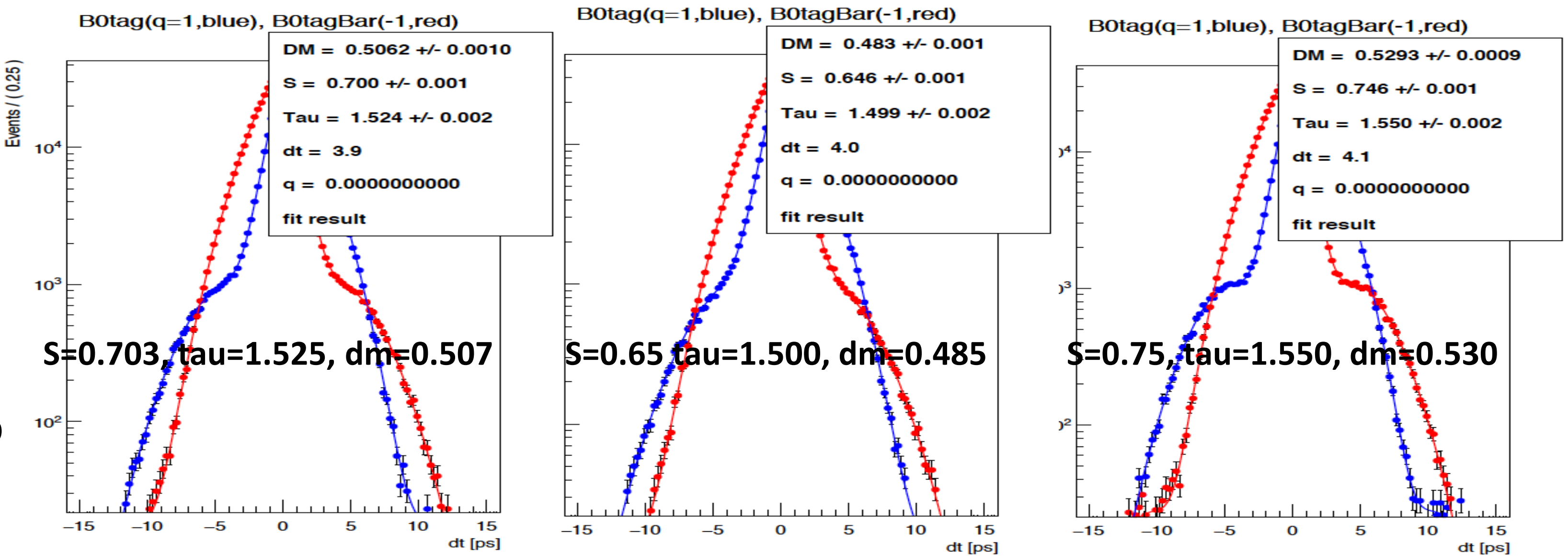
→ *room for improvement of the CPV measurements at Belle II projections to 5 ab^{-1} and 50 ab^{-1} (arXiv:1808.10567)*

	WA (2017)		5 ab^{-1}		50 ab^{-1}	
Channel	$\sigma(S)$	$\sigma(A)$	$\sigma(S)$	$\sigma(A)$	$\sigma(S)$	$\sigma(A)$
$J/\psi K^0$	0.022	0.021	0.012	0.011	0.0052	0.0090
ϕK^0	0.12	0.14	0.048	0.035	0.020	0.011
$\eta' K^0$	0.06	0.04	0.032	0.020	0.015	0.008
ωK_S^0	0.21	0.14	0.08	0.06	0.024	0.020
$K_S^0\pi^0\gamma$	0.20	0.12	0.10	0.07	0.031	0.021
$K_S^0\pi^0$	0.17	0.10	0.09	0.06	0.028	0.018

Linearity checks - variations of $\tau(B^0), \delta m$ and S



$B^0 \rightarrow J/\psi(\mu\mu)\ K_S^0$



"perfect flavor tagging": B^0 or $\bar{B}^0$ from MC

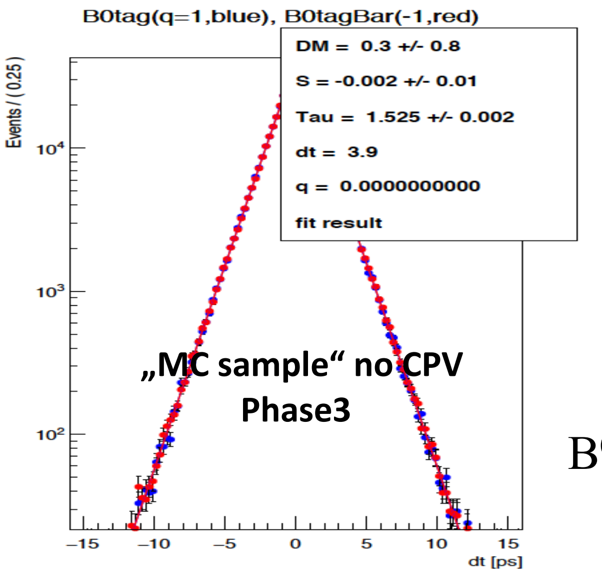
	S	ΔS	$\tau(ps)$	$\Delta\tau(ps)$	$\delta m(ps^{-1})$	$\Delta(\delta m)$
assistive MC	0		1.525		0	
CPV	0.703		1.525		0.507	
	0.7024 ± 0.0029	-0.0006	1.5233 ± 0.0030	-0.0017	0.5089 ± 0.0023	$+0.0019$
CPV Down	0.65		1.500		0.485	
	0.6484 ± 0.0030	-0.0006	1.4944 ± 0.0030	-0.0056	0.4861 ± 0.0026	$+0.0011$
CPV Up	0.75		1.550		0.530	
	0.7480 ± 0.0028	-0.0020	1.5478 ± 0.0030	-0.0022	0.5325 ± 0.0021	$+0.0025$

← no CPV

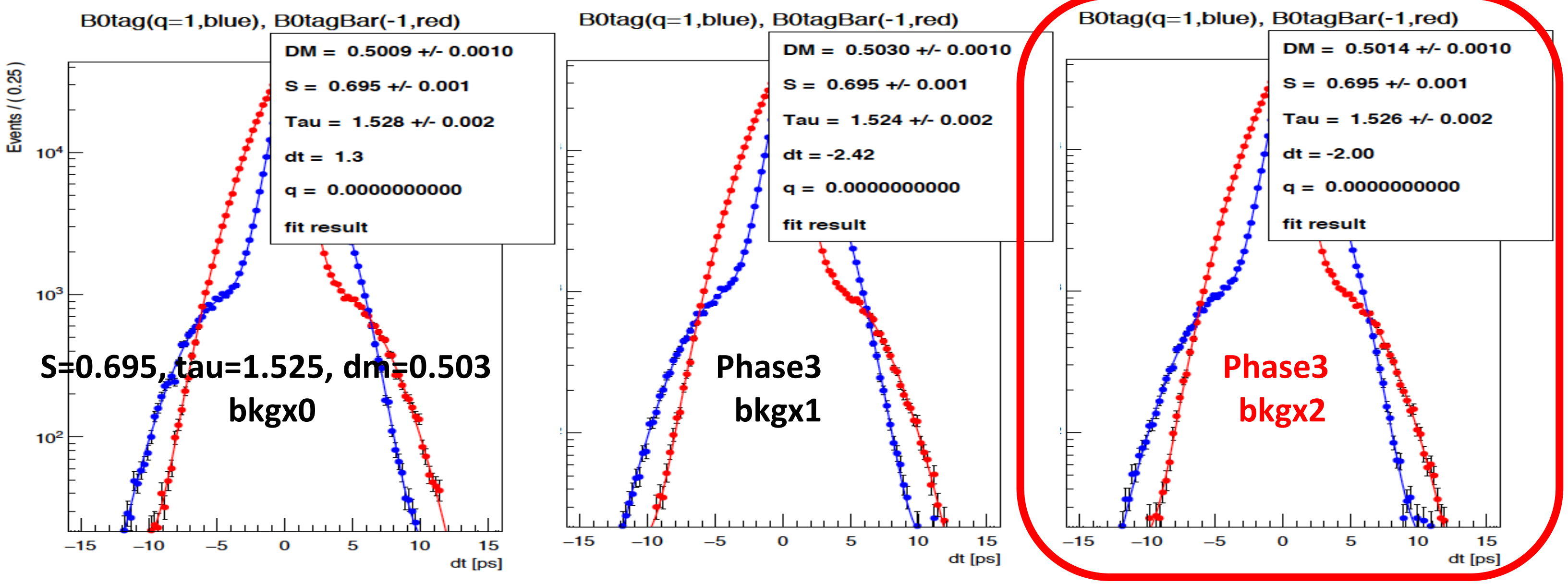
→ *excellent agreement in a wide range of $\tau(B^0), \delta m$ and S*

TABLE VI: The MPI TD CPV fit results with the MC samples "CPV", "CPV Down" or "CPV Up" serving as data and with "no CPV" - as an assistive MC sample. The expected values of physics parameters and differences between fitted and expected values are shown as well.

Treatment of differences in detector response between data and MC



$B^0 \rightarrow J/\psi(\mu\mu)\ K_S^0$



"perfect flavor tagging": B^0 or $\bar{B}^0$ from MC

good →
reasonable →
bkgx2: 4σ and 6σ differences in S and $\tau(B^0)$ →

MC as data	0.695		1.525		0.503	
Ph3 bkg0	0.6948 ± 0.0029	-0.0002	1.5211 ± 0.0030	-0.0039	0.5079 ± 0.0023	$+0.0049$
Ph3 bkg1	0.6876 ± 0.0031	-0.0074	1.5359 ± 0.0032	$+0.0109$	0.5025 ± 0.0025	-0.0005
Ph3 bkg2	0.6816 ± 0.0034	-0.0134	1.5463 ± 0.0036	$+0.0213$	0.5011 ± 0.0028	-0.0029

Multiplicative method: ***excellent (< σ)***
MC with additional smearing: ***very good (<2σ)***

	S	ΔS	$\tau(ps)$	$\Delta\tau(ps)$	$\delta m(ps^{-1})$	$\Delta(\delta m)$	α_{smear}
ass. MC	0		1.525		0		
Ph3 bkg2	0.695		1.525		0.503		
fit 4 par.	0.6945 ± 0.0046	-0.0005	1.5186 ± 0.0074	-0.0064	0.5055 ± 0.0030	$+0.0025$	0.48 ± 0.06
$0.48 \cdot \delta(\Delta t)$	0.6913 ± 0.0035	-0.0037	1.5199 ± 0.0036	-0.0051	0.5044 ± 0.0029	$+0.0014$	

TABLE XII: The MPI TD CPV fit results for the "Ph3 bkg2" sample with an extra free parameter, α_{smear}, and with smearing of the assistive MC sample, "no CPV", prior the fit. The expected values of physics parameters and differences between fitted and expected values are shown as well.