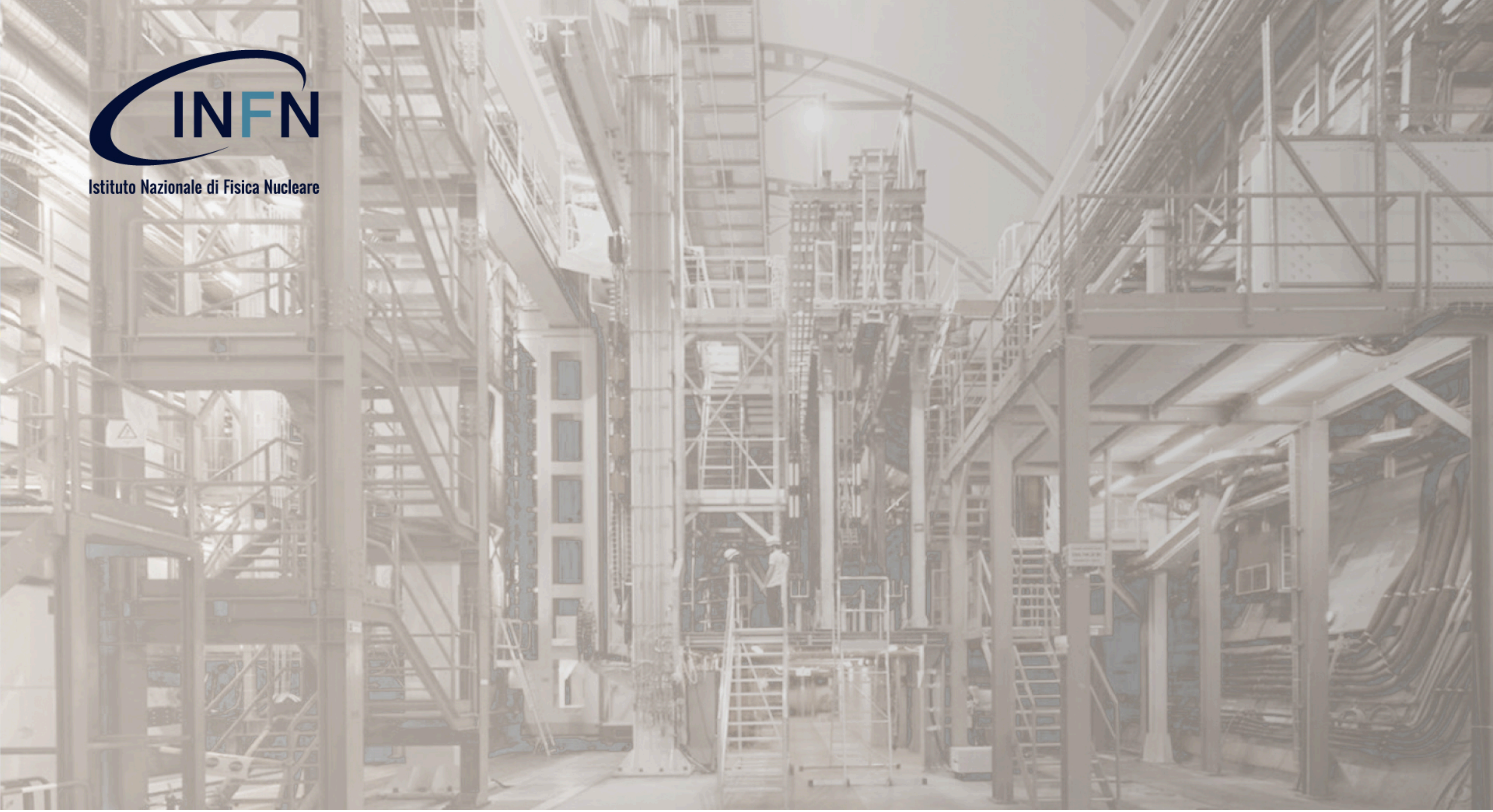


Measurement of $B^0_{(s)} \to \mu^+\mu^-$ decays with Run 1 + Run 2 data at LHCb

Marco Santimaria (INFN-LNF)
on behalf of the LHCb collaboration
LNF+Roma seminar 05/05/2021

INFN
Istituto Nazionale di Fisica Nucleare
Laboratori Nazionali di Frascati

LHCb

CERN

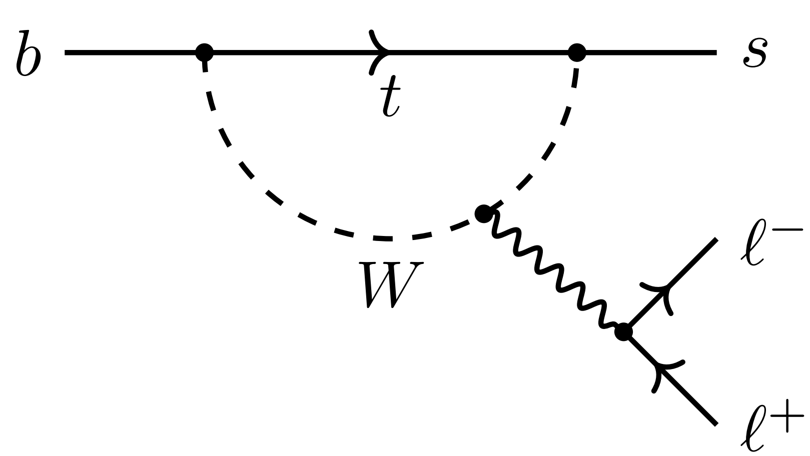
The power of indirect searches

- Precision measurements are a powerful tool to unveil new particles indirectly :
  - 1970: charm presence invoked from the suppression of $K^0 \to \mu^+\mu^-$ before the J/ψ discovery
  - 1973: 3X3 CKM matrix is formulated to explain the CP violation observed in kaons
  - 1987: top mass limit from loop contribution in $B^0 - \overline{B}^0$ mixing: $m_t > 50$ GeV

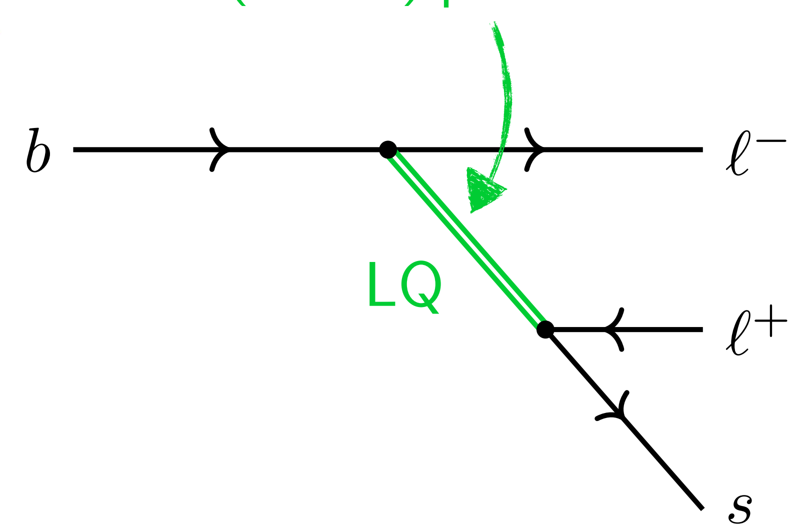
[PRD 2 (1970) 1285] [PTP 49 (1973) 652-657] [PLB 192 (1987) 245-252]

- Because of the large b mass, rare B decays offer a rich phenomenology for indirect searches of New Physics (NP). For example:

$b \to s\ell^+\ell^-$ are FCNC processes that can only occur via loop in the SM



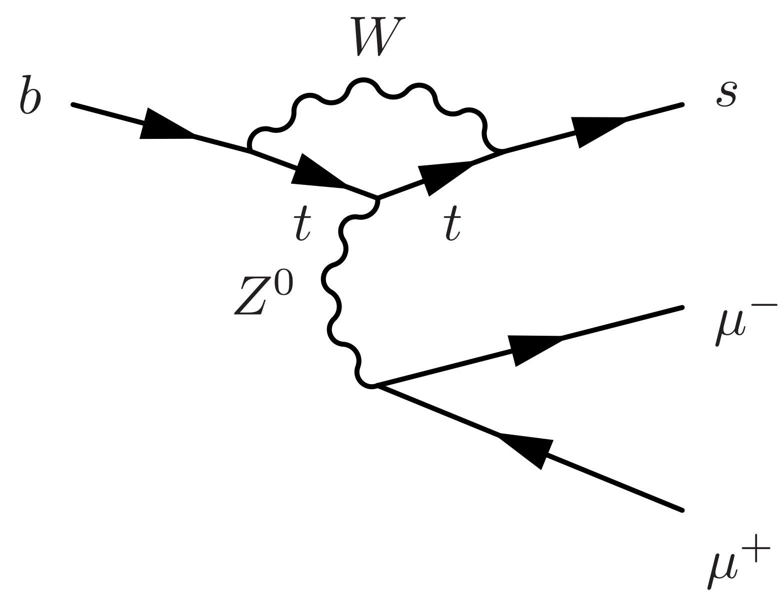
observables are altered by new (virtual) particles



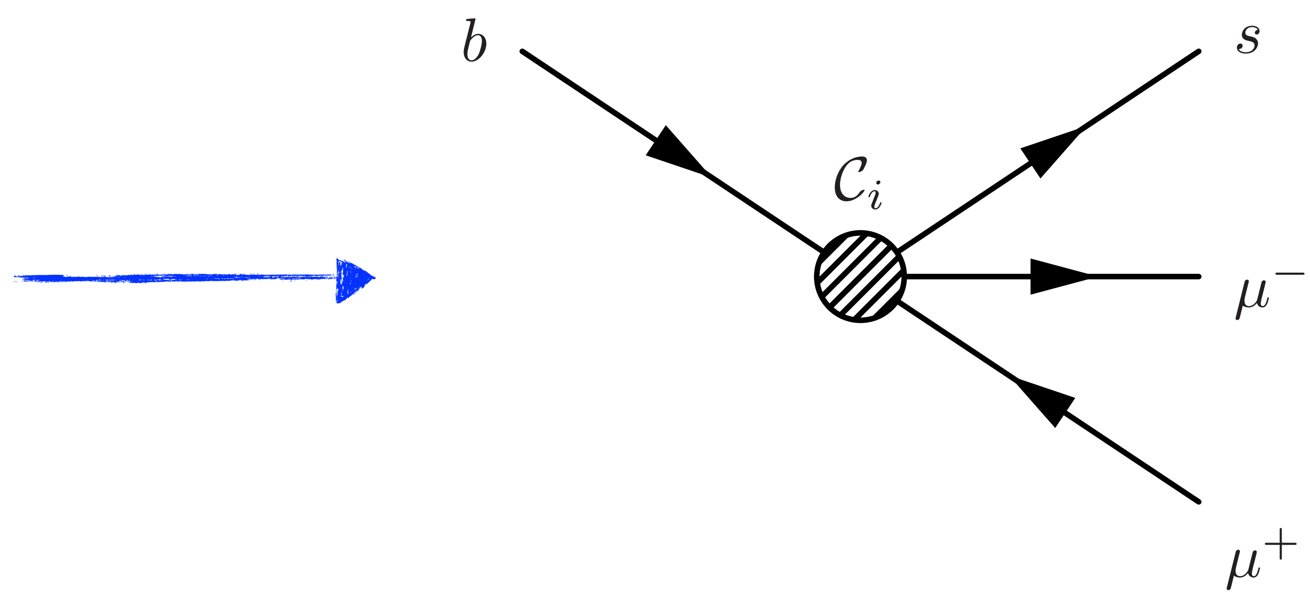
Effective theory for rare B decays [Rev.Mod.Phys. 68 (1996) 1125-1144]

- $b \to s\ell^+\ell^-$ can be described with an "Effective Hamiltonian", where high- and low-energy contributions are factorised ($M_b \ll M_W$):

Full theory



Effective description



- "point-like interaction" as in the Fermi description of the neutron decay

$$\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} \sum_i V^i_{CKM} C_i(\lambda) \mathcal{O}_i(\lambda)$$

- Wilson coefficients (short-distance): evaluated with perturbation theory
- Local operators (long-distance): the corresponding form factor is computed with, e.g., lattice QCD

Probing New Physics with rare B decays

- SM operators for $b \to s\ell^+\ell^-$:

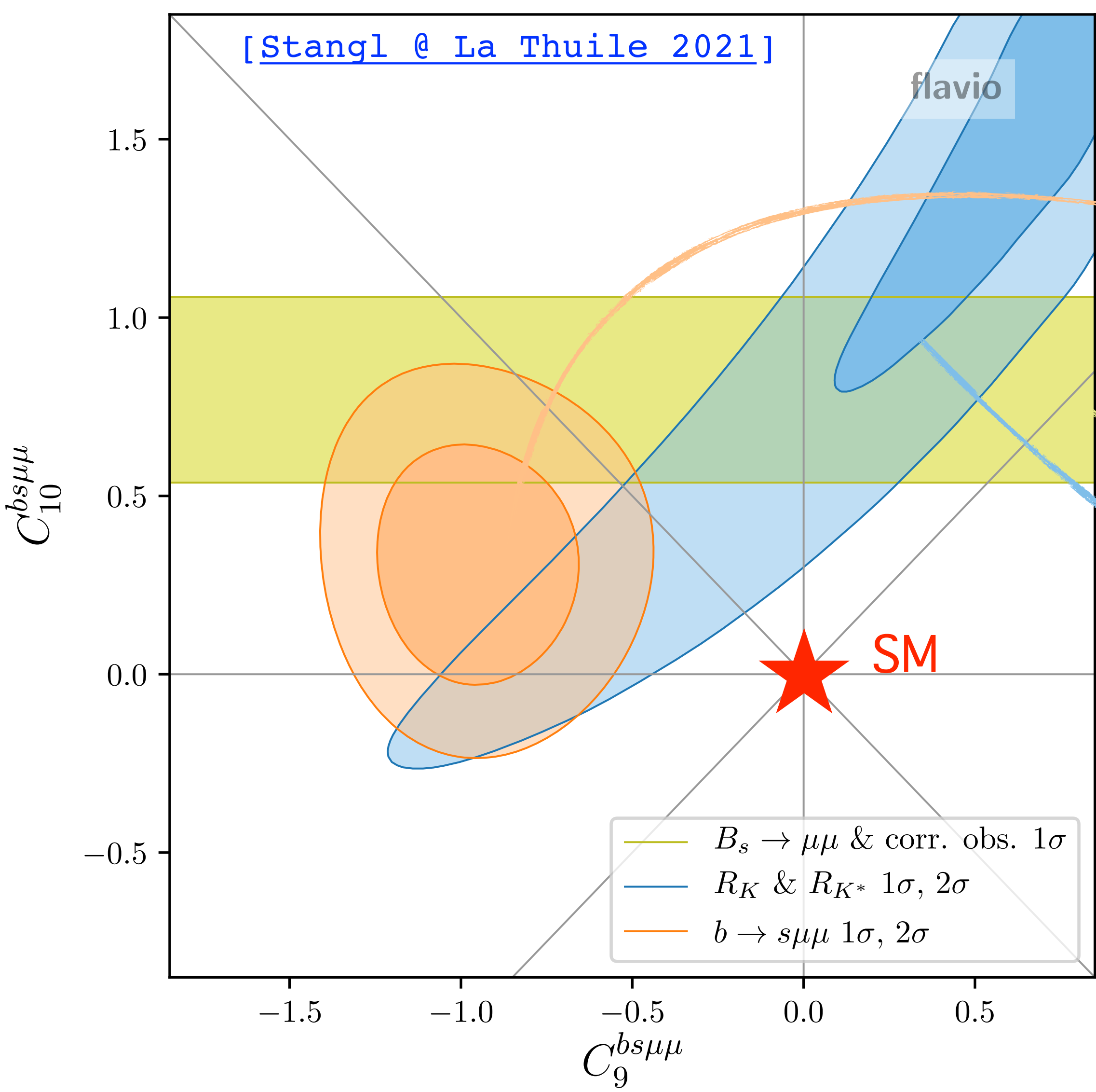
$$\mathcal{O}_9^{(\prime)} = (\bar{s}P_{\mathrm{L(R)}}b)\,(\bar{\ell}\gamma^\mu\ell)$$

$$\mathcal{O}_{10}^{(\prime)} = (\bar{s}P_{\mathrm{L(R)}}b)\,(\bar{\ell}\gamma^\mu\gamma^5\ell)$$

- NP can alter $C_i^{(\prime)}$ but also introduce new operators

$$\Delta\mathcal{H}_{\mathrm{NP}} = \frac{c_i}{\Lambda_{\mathrm{NP}}^2}\mathcal{O}_i$$

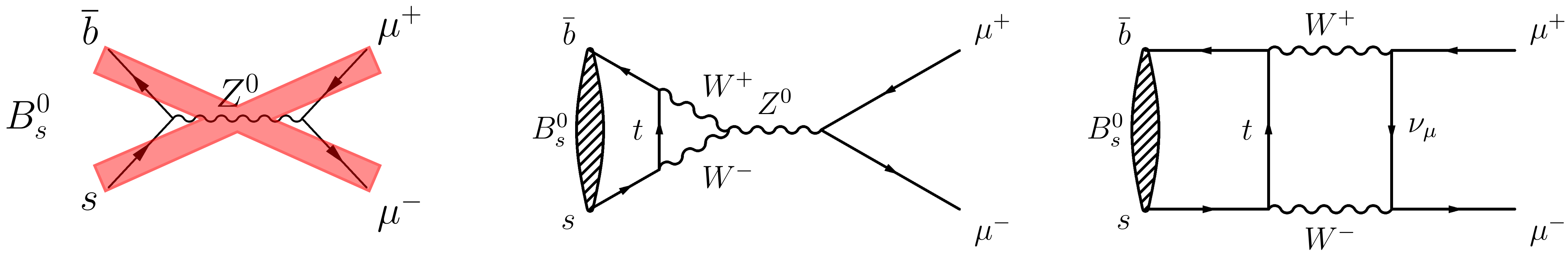
Precision measurements go well beyond collision energies!



- Experimental picture → the "anomalies":
  - The latest global fit prefers NP contributions to C_9 and C_{10}
  - Crucial input from $\mathscr{B}(B_s^0 \to \mu^+\mu^-)$ (here from the 2020 ATLAS+CMS+LHCb combination)
  - R_K (and other constraints) in the next talks!

$B^0_{(s)} \to \mu^+\mu^-$ decays in the SM

- In the SM, B^0 and B^0_s decays to two muons are FCNC and helicity suppressed :



(tree) (penguin) (box)

$$\mathcal{B}(B^0_q \to \mu^+\mu^-)_{\rm SM} = \frac{\tau_{B_q} G_F^4 M_W^4 \sin^4\theta_W}{8\pi^5} |C_{10}^{\rm SM} V_{tb} V_{tq}^*|^2 f_{B_q}^2 m_{B_q} m_\mu^2 \sqrt{1 - \frac{4m_\mu^2}{m_{B_q}^2}} \frac{1}{1-y_q} \qquad q = d, s$$

single Wilson coefficient & single hadronic constant (known at $\simeq 0.5\,\%$!) → golden channel

[PRD 98 (2019) 074512]

- Very clean prediction in the SM:

$$\mathcal{B}(B^0_s \to \mu^+\mu^-)_{\rm SM} = (3.66 \pm 0.14) \times 10^{-9}$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-)_{\rm SM} = (1.03 \pm 0.05) \times 10^{-10}$$ [JHEP 10 (2019) 232]

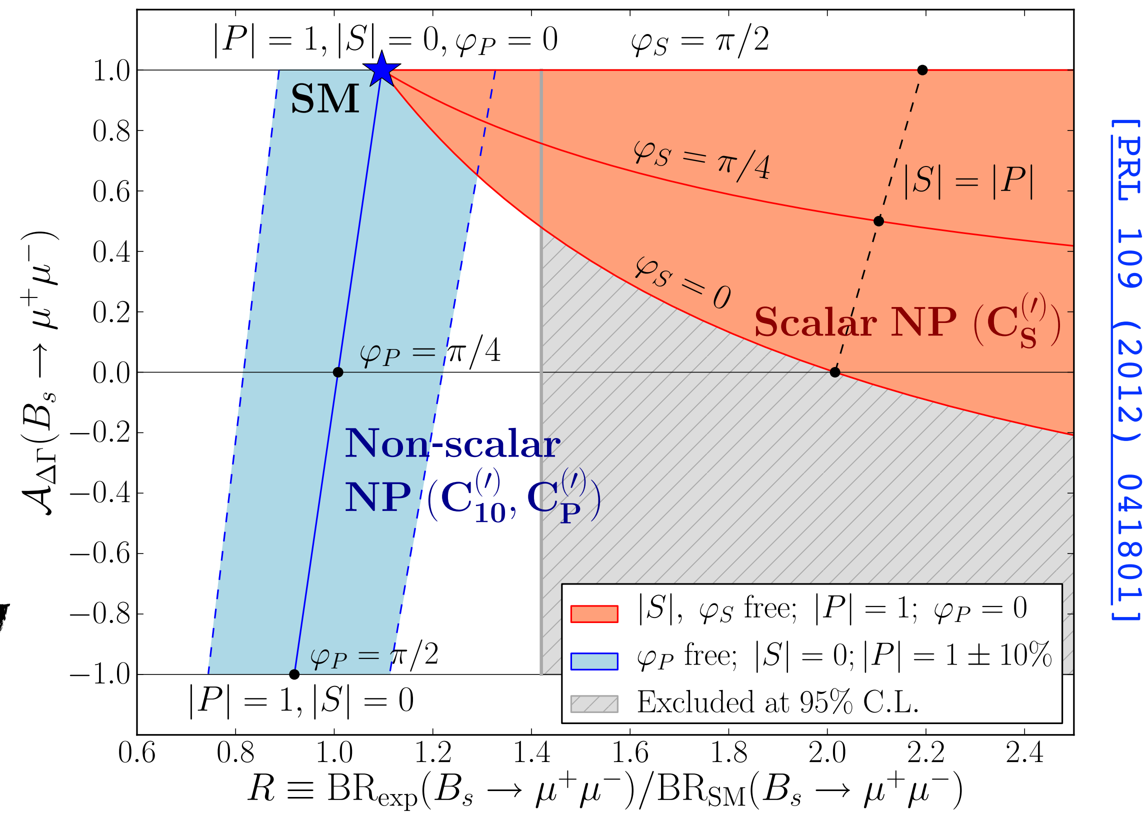
$B_s^0 \to \mu^+\mu^-$: not only branching fractions

- By measuring the $B_s^0 \to \mu^+\mu^-$ effective lifetime:

$$A_{\Delta\Gamma}^{\mu^+\mu^-} \equiv \frac{R_H^{\mu^+\mu^-} - R_L^{\mu^+\mu^-}}{R_H^{\mu^+\mu^-} + R_L^{\mu^+\mu^-}}$$

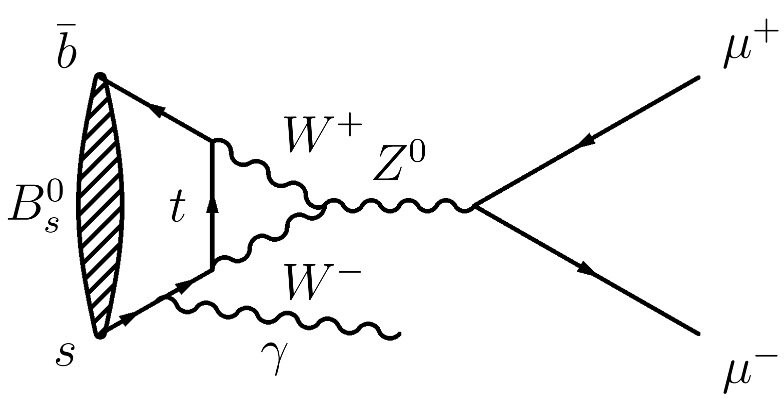
$$\tau_{\mu^+\mu^-} = \frac{\tau_{B_s}}{1 - y_s^2}\left[\frac{1 + 2A_{\Delta\Gamma}^{\mu^+\mu^-} y_s + y_s^2}{1 + A_{\Delta\Gamma}^{\mu^+\mu^-} y_s}\right] \qquad y_s = \frac{\Delta\Gamma_s}{2\Gamma_s}$$

- we can extract the asymmetry $A_{\Delta\Gamma}^{\mu^+\mu^-}$, $= +1$ in the SM
- Clean observable → additional NP constraints

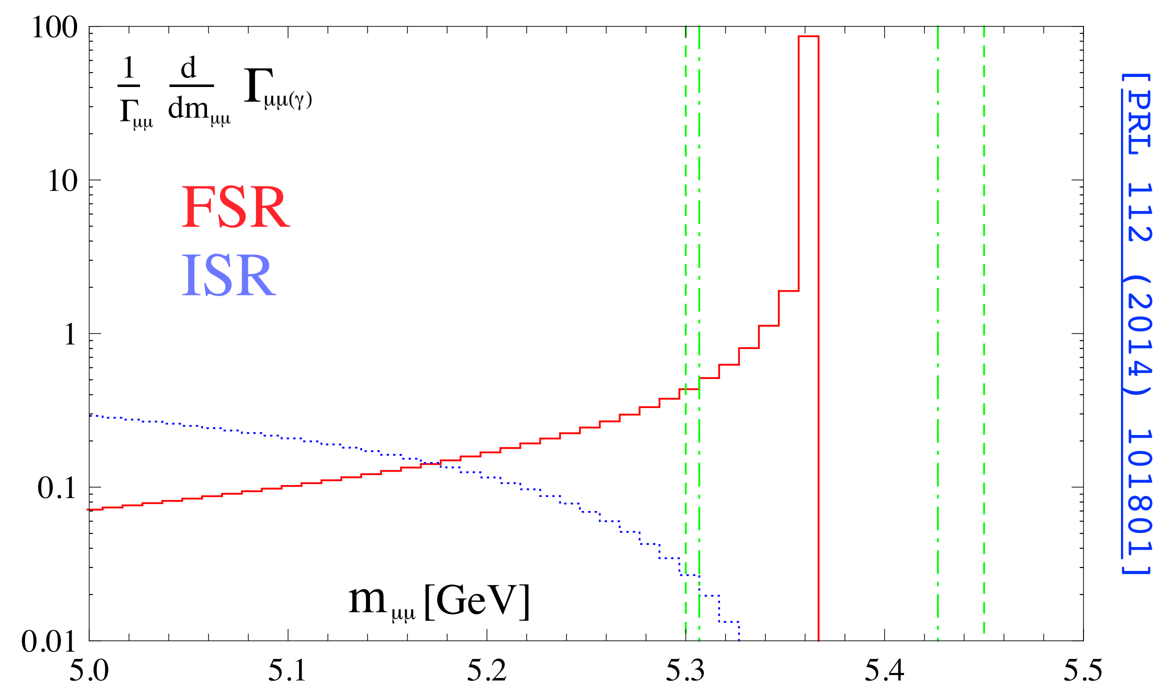


[PRL 109 (2012) 041801]

- Sensitivity to $B_s^0 \to \mu^+\mu^-\gamma$ (ISR) at high $m_{\mu^+\mu^-}$, new observable included this analysis



- SM prediction at $\mathcal{O}(10^{-10})$ for $m_{\mu^+\mu^-} > 4.9$ GeV
[JHEP 11 (2017) 184] [PRD 97 (2018) 053007]

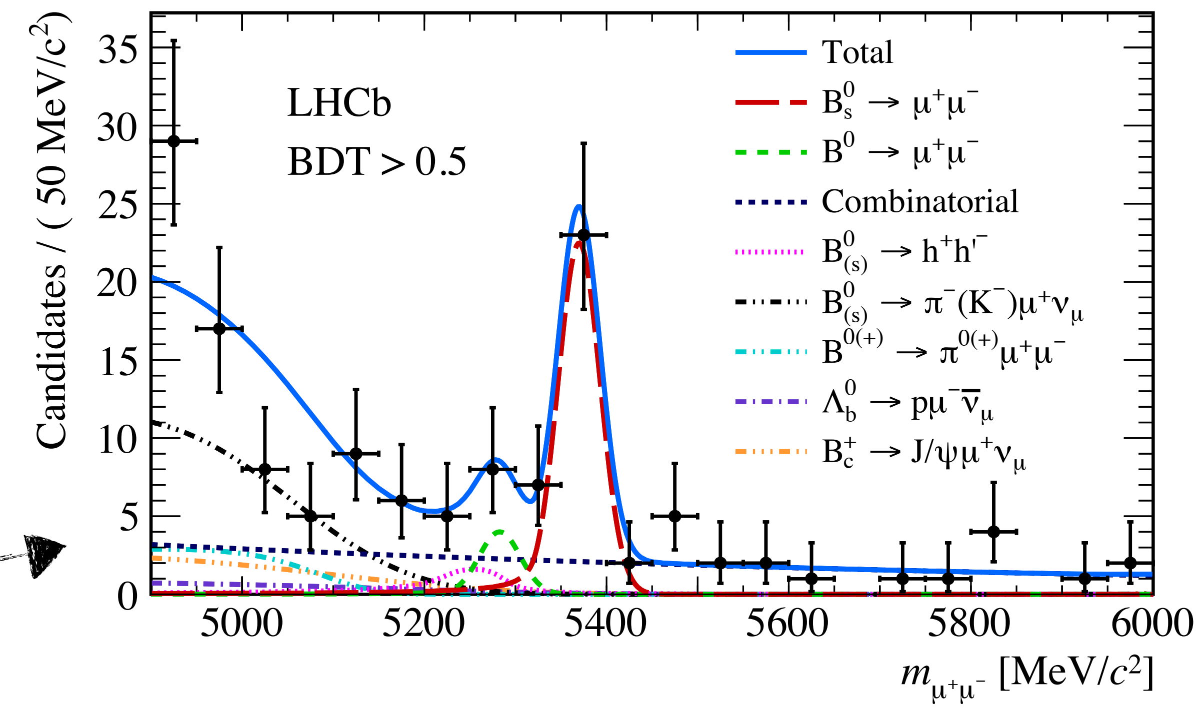


[PRL 112 (2014) 101801]

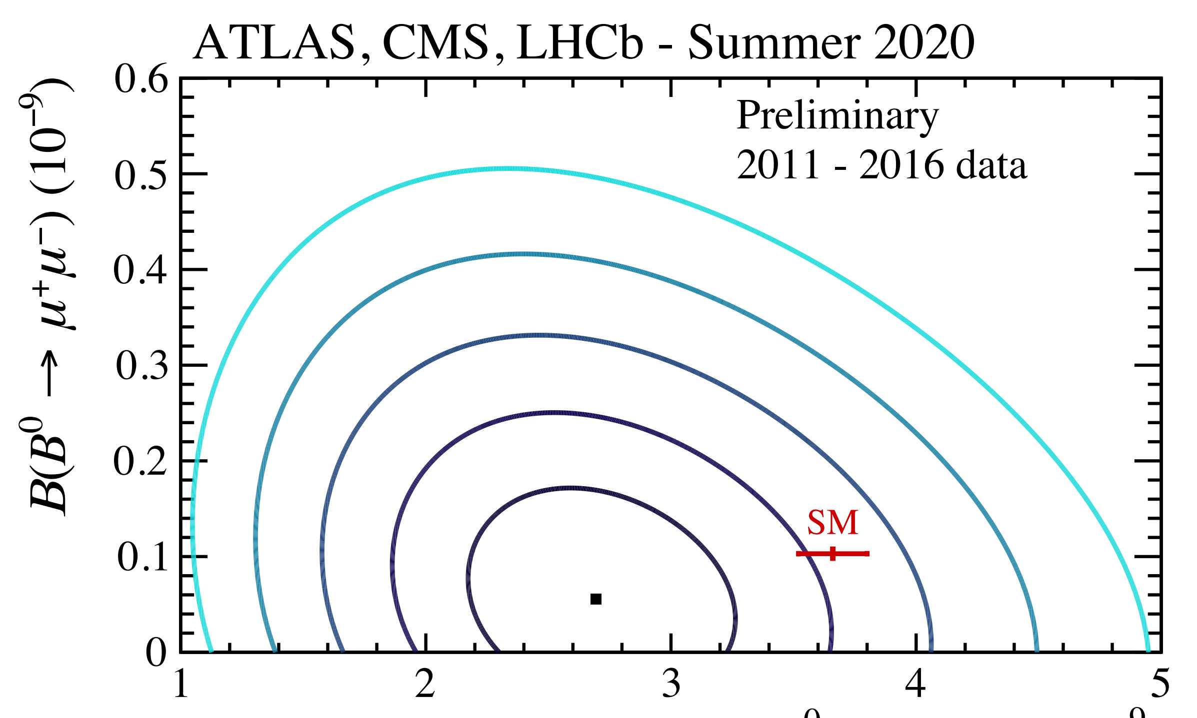
- Bremsstrahlung (FSR) experimentally included in $\mathscr{B}(B_s^0 \to \mu^+\mu^-)$ via PHOTOS

Experimental measurements

- 1984 The search begins at CLEO
  $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 2 \times 10^{-4}$ (90 % CL) [PRD 30 (1984) 11]
- 2015 First observation of $B_s^0 \to \mu^+\mu^-$ with CMS + LHCb (Run 1 data) [Nature 522 (2015) 68–72]
- 2017 First observation of $B_s^0 \to \mu^+\mu^-$ with a single experiment by LHCb ($4.4\ \mathrm{fb}^{-1}$)
  $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.0 \pm 0.6^{+0.3}_{-0.2}) \times 10^{-9}$
  [PRL 118 (2017) 191801]



- 2020 combination of ATLAS, CMS and LHCb:
  - $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (2.69^{+0.37}_{-0.35}) \times 10^{-9}$
  - 2.1σ away from the SM
  - $\tau_{\mu^+\mu^-} = 1.91^{+0.37}_{-0.35}$ ps
  - $\mathcal{B}(B^0 \to \mu^+\mu^-) < 1.9 \times 10^{-10}$ (95 % CL)



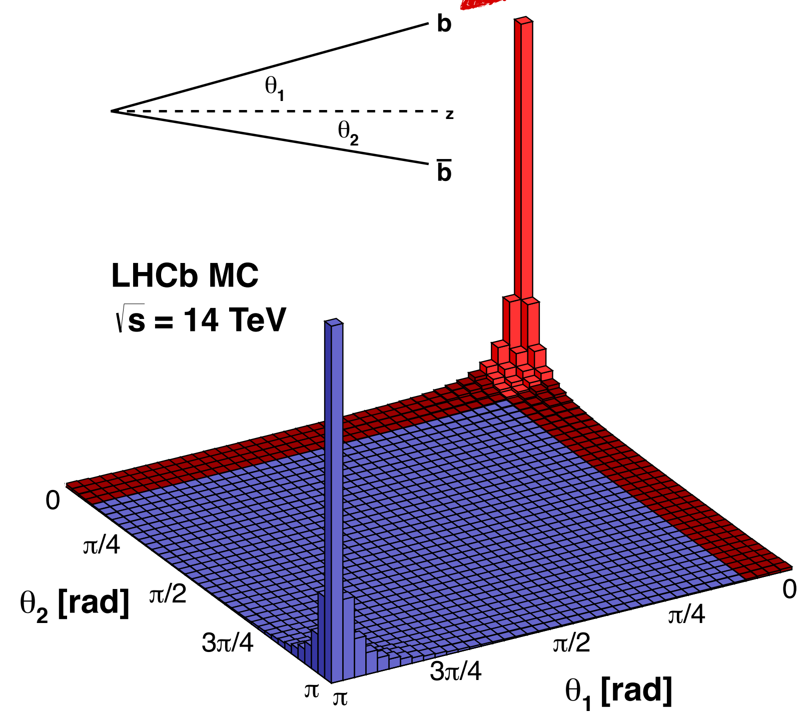
[ATLAS-CONF-2020-049]

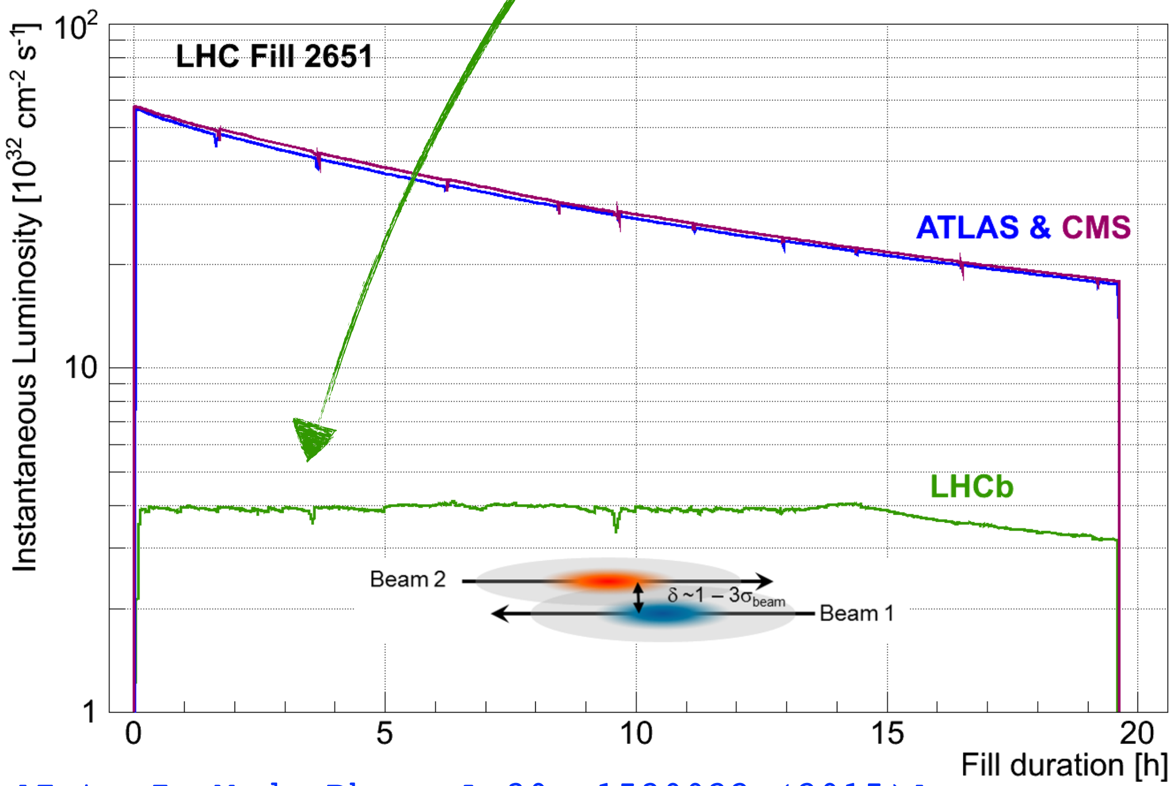
- Only experimental limit today on:
  $\mathcal{B}(B^0 \to \mu^+\mu^-\gamma) < 1.6 \times 10^{-7}$ from BaBar at 90 % CL

[PRD 77 (2008) 011104]

The LHCb data-taking

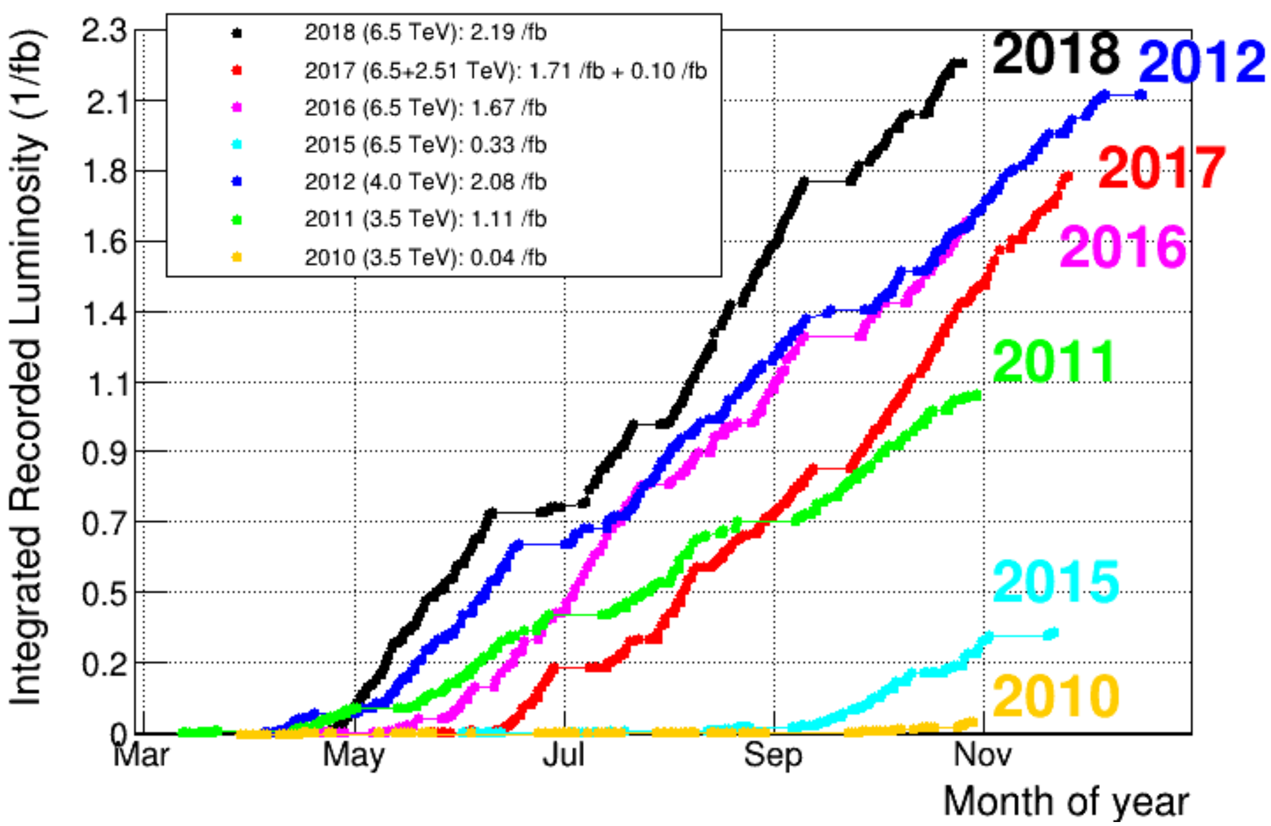
- Large $b\bar{b}$ cross section in the LHCb acceptance ($2 < \eta < 5$)
  $\sigma(pp \to b\bar{b}) \simeq 144\ \mu\text{b}\ \ (\sqrt{s} = 13\ \text{TeV})$ [PRL 118 (2017) 052002]
- Full LHCb dataset $3\ \text{fb}^{-1}$ $(\sqrt{s}_{\text{Run1}} = 7\ \&\ 8\ \text{TeV})$ +
  $6\ \text{fb}^{-1}$ $(\sqrt{s}_{\text{Run2}} = 13\ \text{TeV})$
- Run 2 luminosity levelled to $\simeq 4.4 \times 10^{32}\ \text{cm}^{-2}\text{s}^{-1}$
  (>2x the design value)





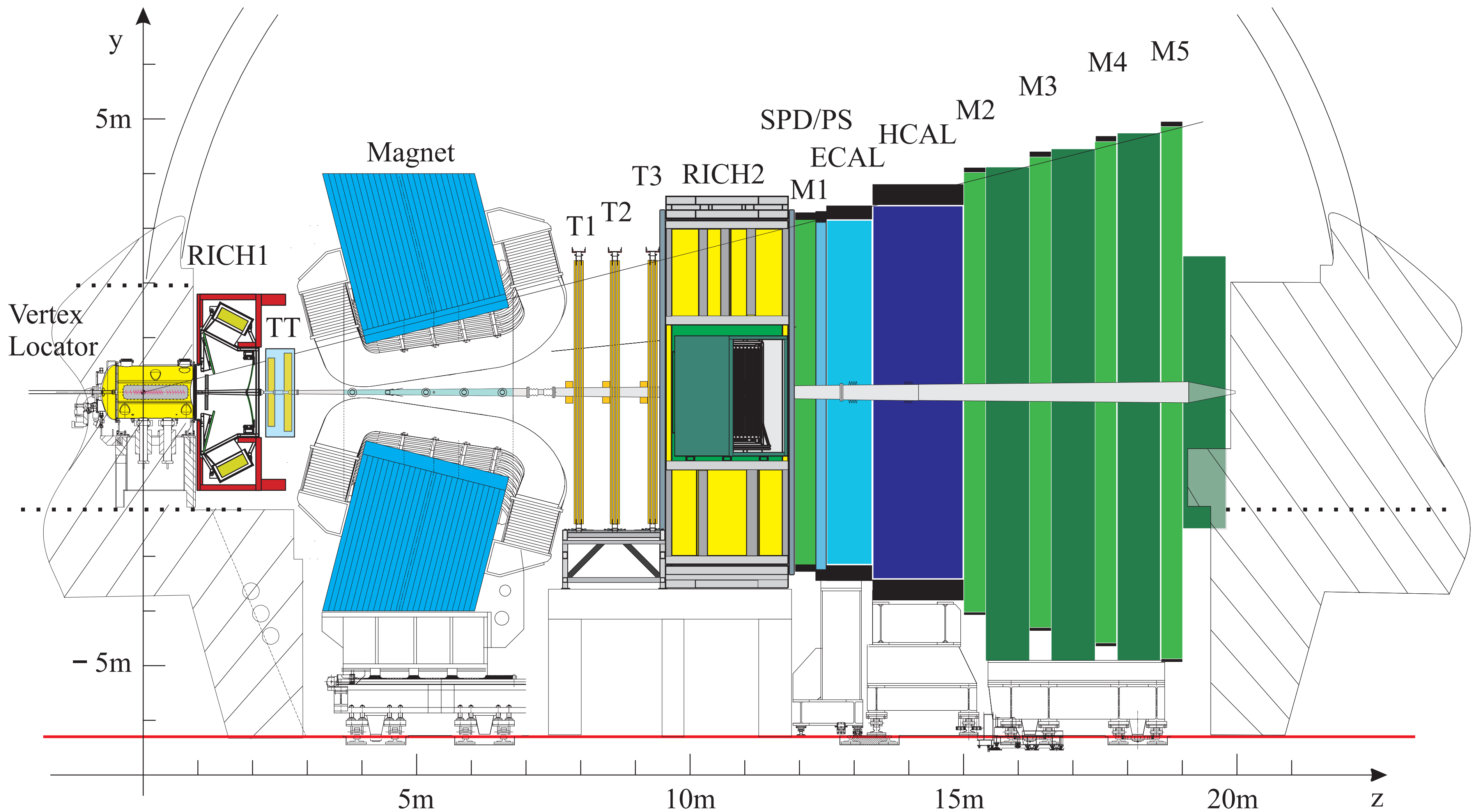
[Int. J. Mod. Phys. A 30, 1530022 (2015)]

$$\sigma(pp \to b\bar{b})_{\text{Run 2}} \approx 2 \times \sigma(pp \to b\bar{b})_{\text{Run 1}}$$



The LHCb detector 1/2

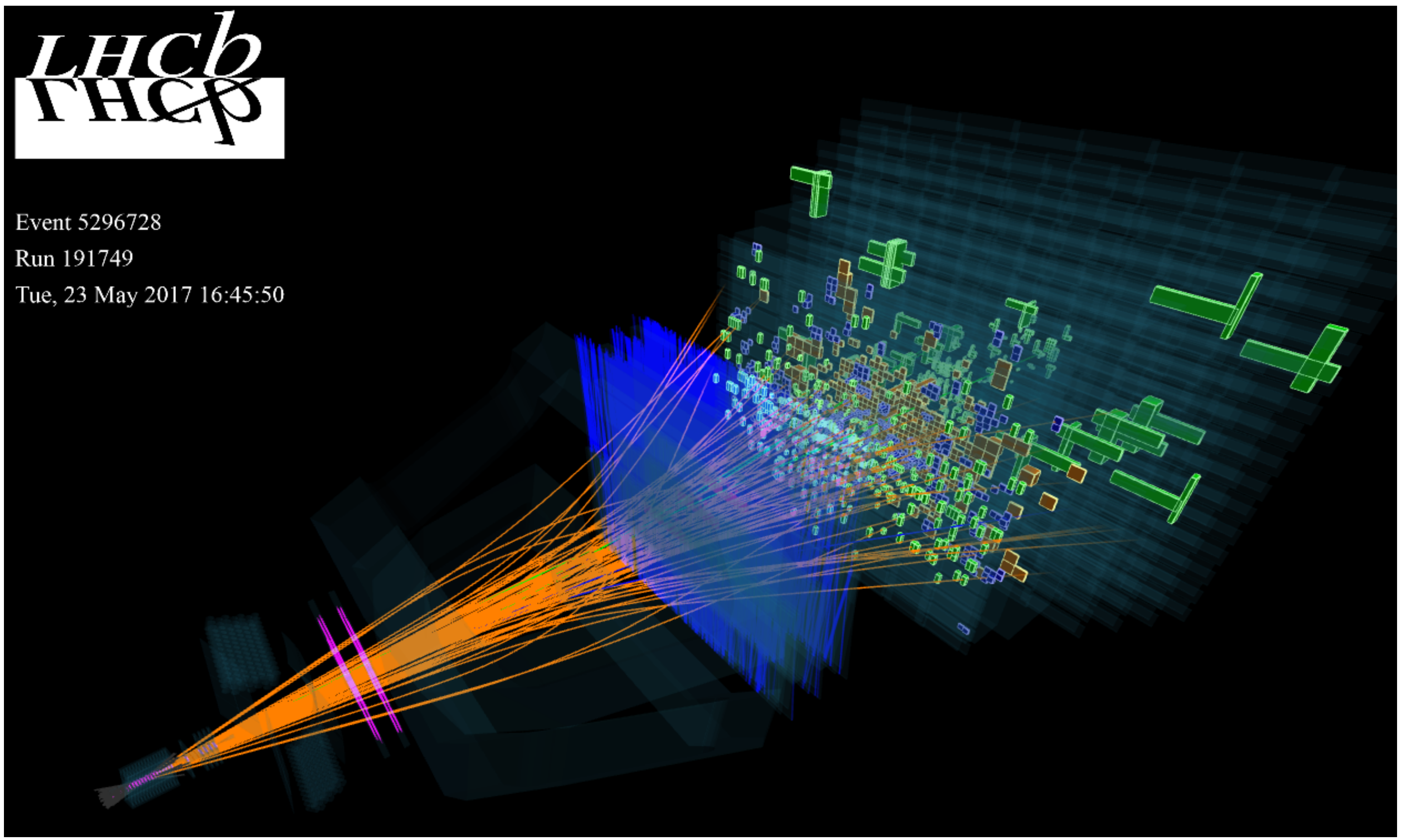
[JINST 3 (2008) S08005]
[Int. J. Mod. Phys. A 30, 1530022 (2015)]



- High vertex resolution (VELO)
  $\sigma_{\mathrm{IP}} = 15 + 29/p_T\ \mu\mathrm{m}$
  B travel distance $\mathcal{O}(1\ \mathrm{cm})$
  → well detached secondary vertex
- Particle identification capabilities (RICH+CALO+MUON)
  $\epsilon_\mu \sim 98\,\%$ with $\epsilon_{\pi\to\mu} \lesssim 1\,\%$
  → our working point $\epsilon_{\mathrm{PID}}^{\mathrm{sig}} \sim 84\,\%$ with $\epsilon_{B\to hh\to\mu\mu} \approx 10^{-6}$
- Excellent momentum resolution (T stations)
  $\sigma_p/p = 0.5 - 1.0\,\%\ (p \in [2{,}200]\ \mathrm{GeV})$
  → narrow mass peak
- Three strength points for background reduction, at the core of the analysis

The LHCb detector 2/2

- Low momentum muon trigger (M1-M5 stations)
  $p_{T_\mu} > 1.75$ GeV (2018)
- High level software trigger
  Impact Parameter + invariant mass cuts
  → total $\epsilon^{\rm sig}_{\rm trig} \sim 90\,\%$
- Large Acceptance x reconstruction
  $\rightarrow \epsilon^{\rm sig}_{\rm acc\times rec} \sim 10\,\%$
- Parallel data taking of high statistics control samples e.g. to measure trigger and PID performance

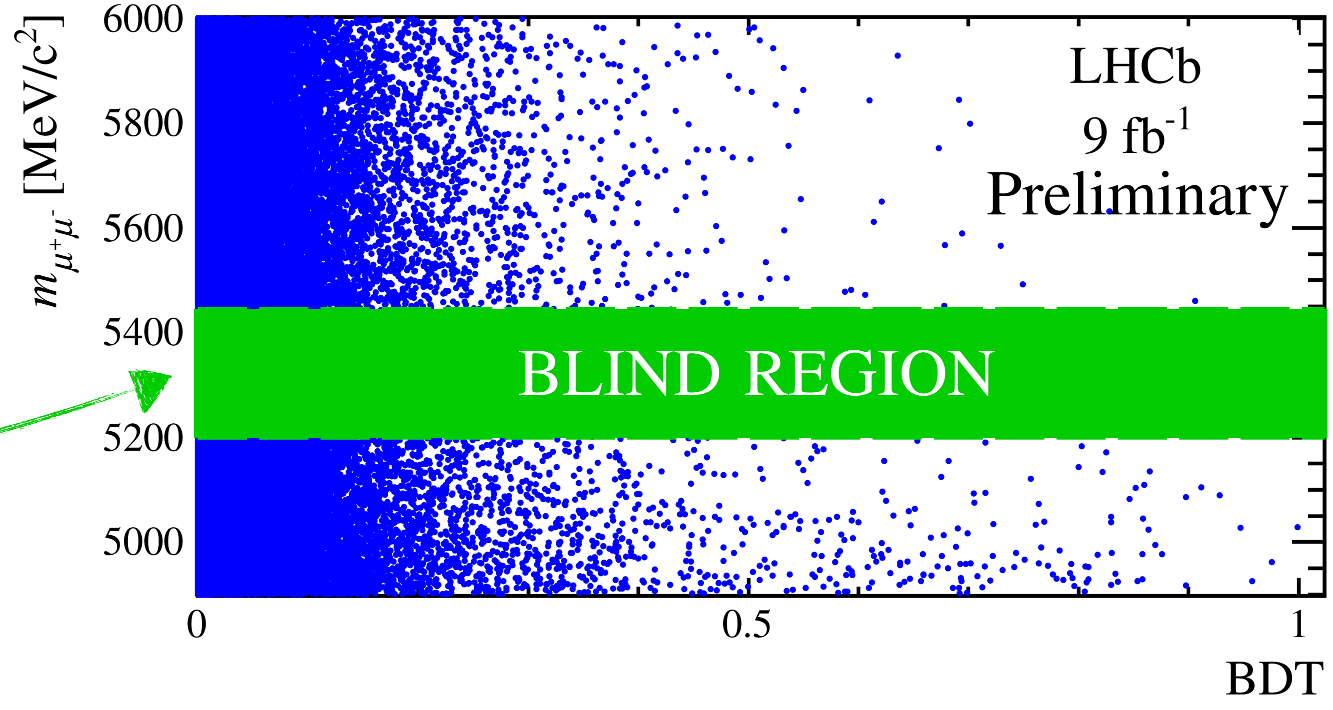


Analysis strategy

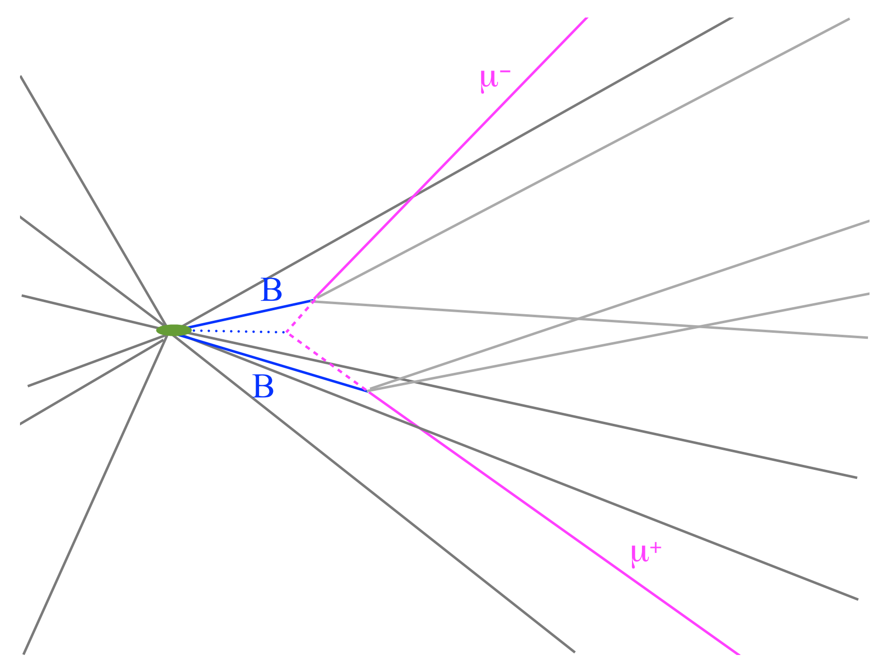
[LHCB-PAPER-2021-007]

- Will show here the "legacy measurement" of LHCb on the full Run 1 + Run 2 data ($9\ \mathrm{fb}^{-1}$)
- The strategy is well established since 2017 but introduces several improvements

- Select muon pairs with $m_{\mu^+\mu^-} \in [4900{,}6000]$ MeV forming a displaced vertex
- Signal mass region is blinded until the analysis is finalised



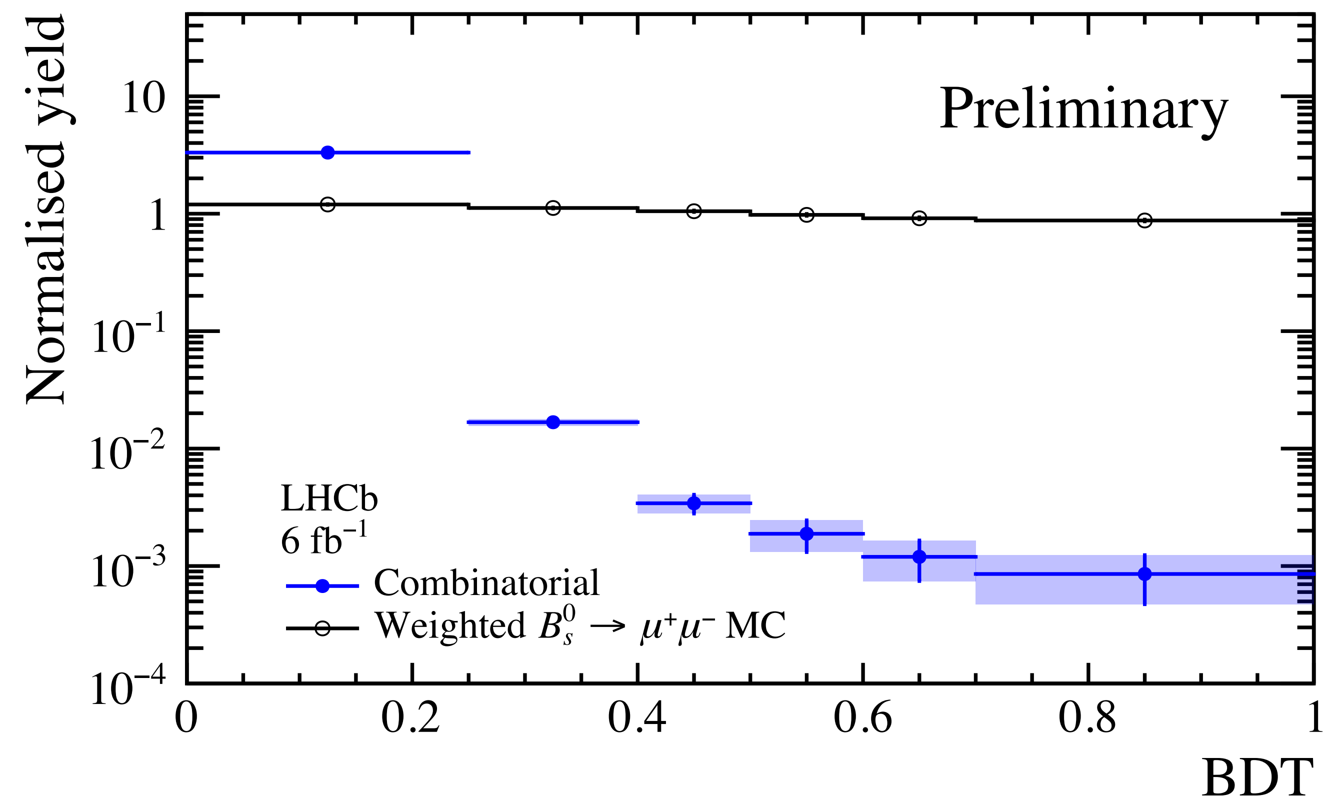
- The selected candidates are mostly combinatorial background
- To reject them we use a multivariate classifier "BDT" (Boosted Decision Tree)
- The algorithm primarily exploits muon isolation and vertex detachment



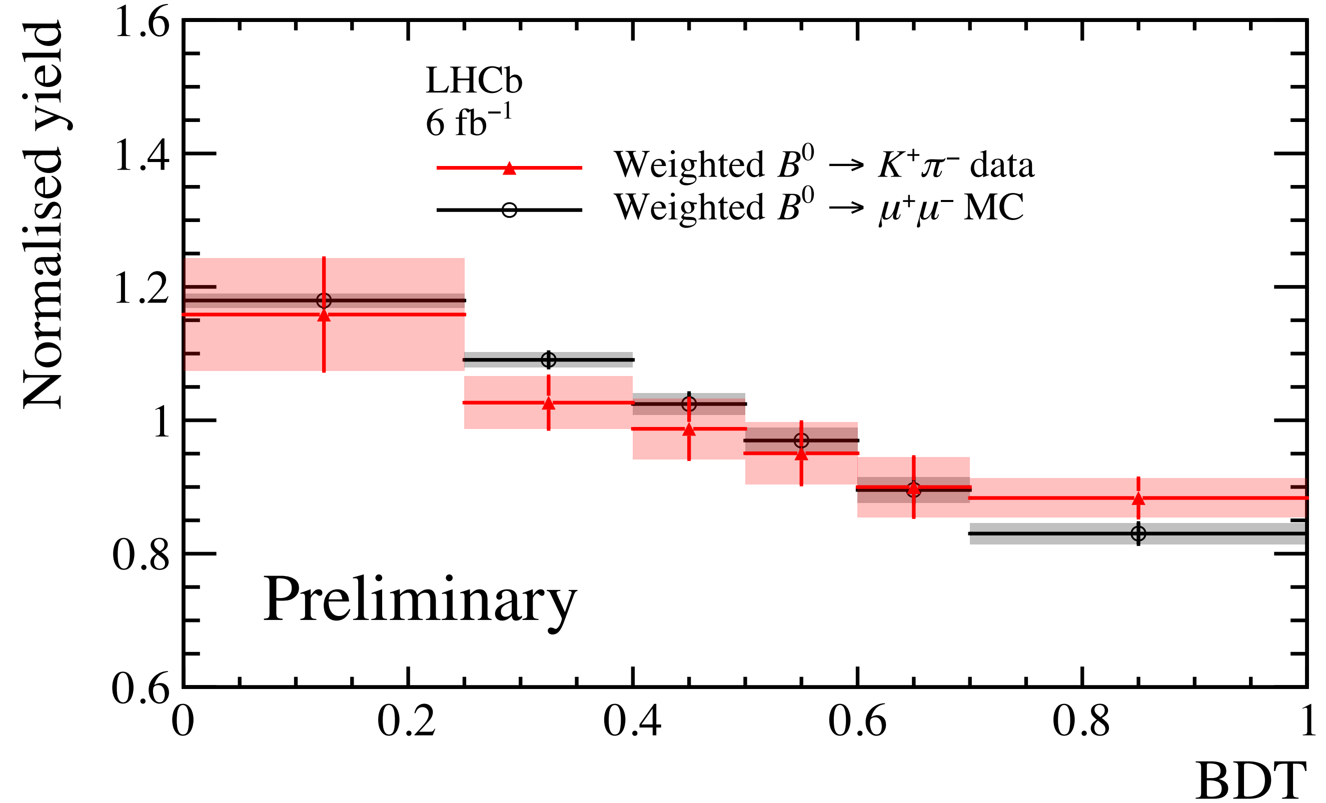
The BDT and its calibration

[LHCB-PAPER-2021-007]

- Events are categorised into 6 "BDT bins" : flat signal BDT and decreasing combinatorial
- Branching fractions are measured with a simultaneous mass fit in 10 categories (2 Runs X 5 BDT bins)
- The first bin $[0, 0.25]$ is excluded from the fit since it's background-dominated



- The signal BDT output is calibrated on data-corrected simulation
- Cross-checked on $B^0 \to K^+\pi^-$ events (signal proxy)
- The "shape" is determined by PID and trigger efficiencies
- Also: BDT-lifetime correlations for the $B_s^0 \to \mu^+\mu^-(\gamma)$ signals (see → backup)

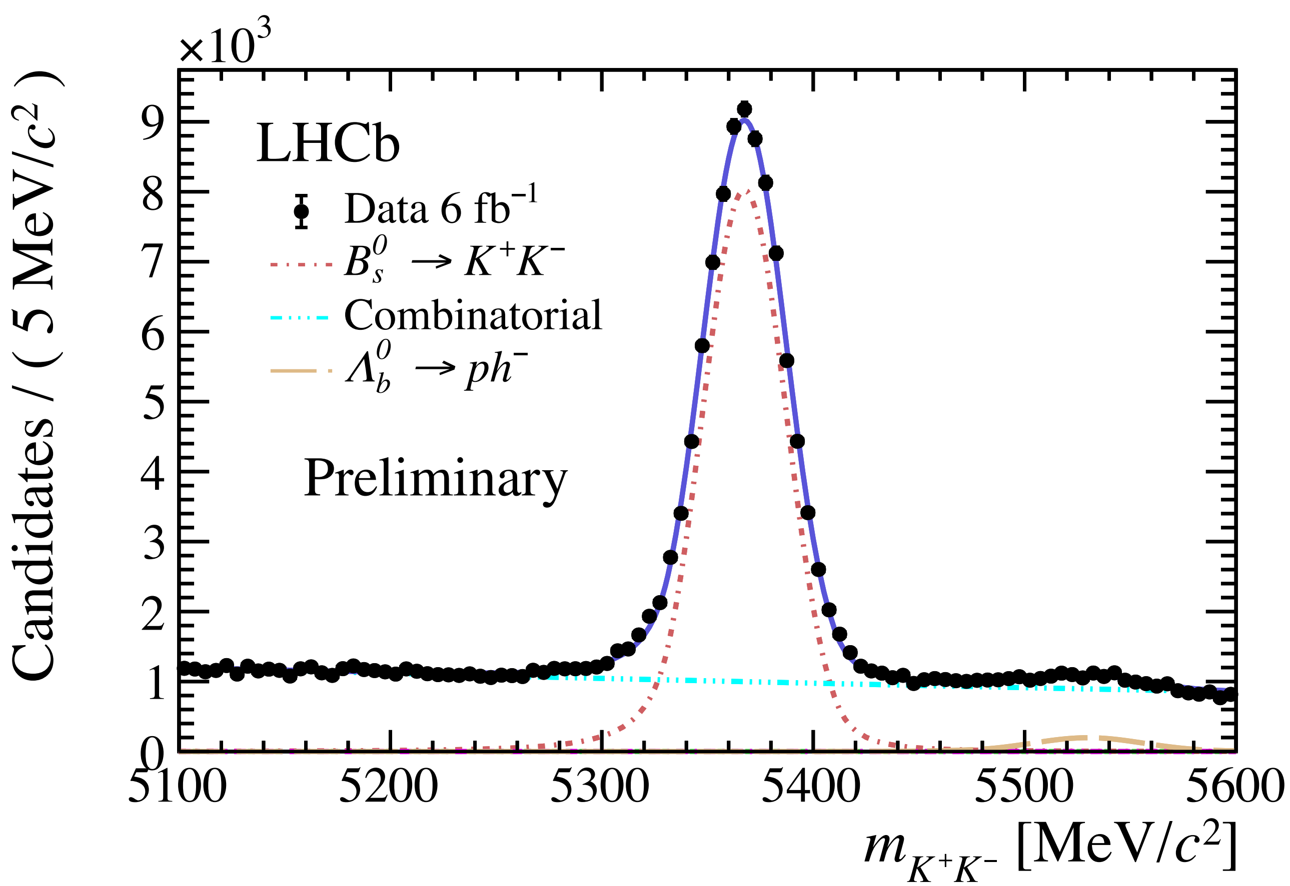


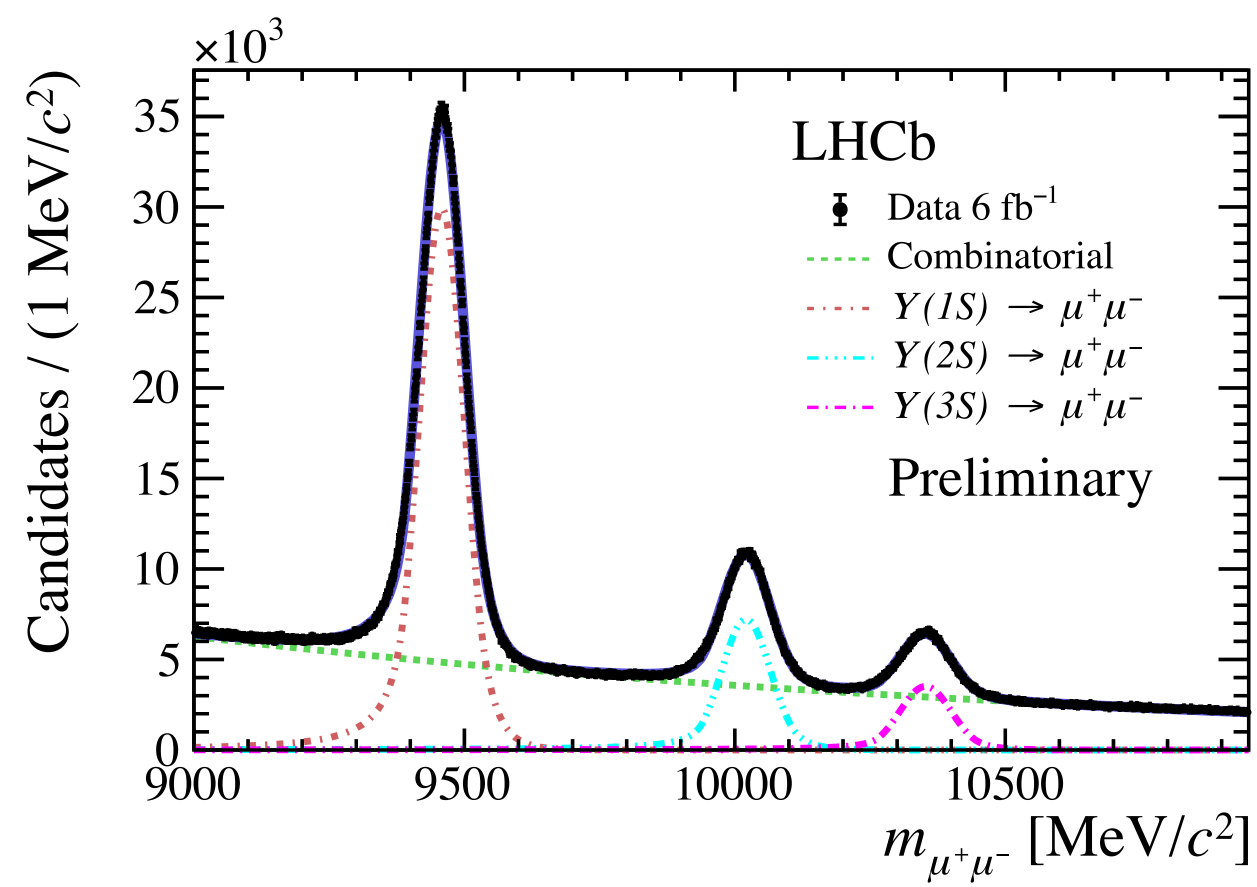
Mass shape calibration

- The $B^0_{(s)} \to \mu^+\mu^-$ mean (peak position) and resolution (width) are measured on data:

  - The mean is obtained from a fit to $B^0 \to K^+\pi^-$ and $B^0_s \to K^+K^-$ events for $B^0 \to \mu^+\mu^-$ and $B^0_s \to \mu^+\mu^-$

  - The resolution is interpolated from mass fits to $c\bar{c}$ and $b\bar{b}$ resonances: $\sigma_{m(\mu^+\mu^-)} = 21.96 \pm 0.63$ MeV (Run 2)





Normalisation: mass fits

[LHCB-PAPER-2021-007]

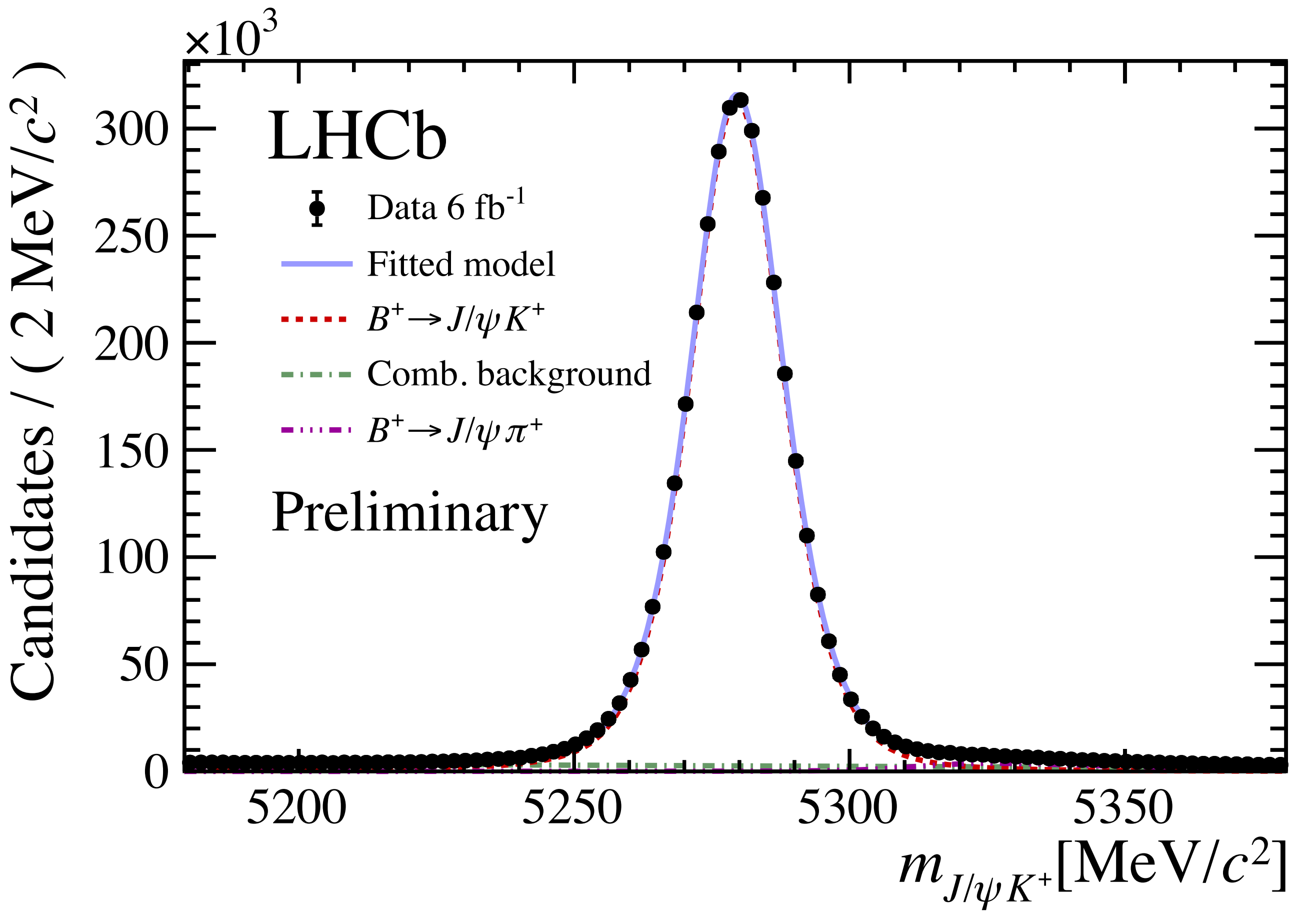
- To measure the branching fraction, luminosity and cross-section uncertainties are avoided by computing the ratio to a well-known channel
- Two normalisation channels are employed → perform mass fits to compute the yields

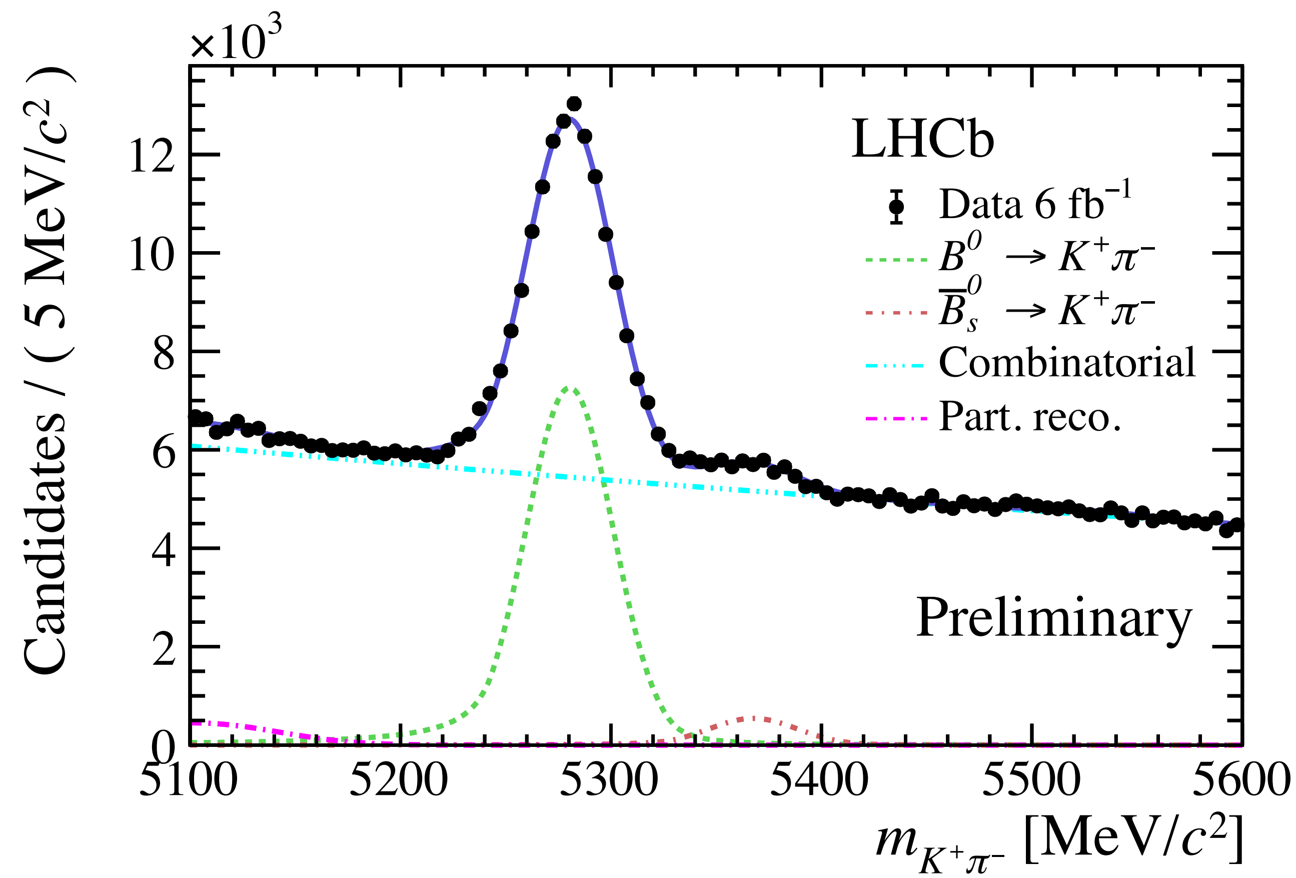
1. $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$

   Two muons in the final state
   → similar trigger and reconstruction

2. $B^0 \to K^+\pi^-$

   Two-body B decay
   → same signal topology





Normalisation: results

[LHCB-PAPER-2021-007]

- The observed signal yield is converted into a BF according to:

$$\mathcal{B}(B^0_{d,s} \to \mu^+\mu^-) = \overbrace{\underbrace{\frac{\mathcal{B}_{norm}}{N_{norm}} \times \frac{\epsilon_{norm}}{\epsilon_{sig}}}_{\alpha_d} \times \frac{f_{norm}}{f_{d,s}}}^{\alpha_s} \times N_{B^0_{d,s}\to\mu^+\mu^-}$$

- $\mathscr{B}$ and yield of the normalisation channel
- Signal/normalisation efficiency ratio
- Ratio of hadronisation fraction (for the B^0_s)

  Very recent combination of LHCb measurements: [LHCb-PAPER-2020-046]

  f_s/f_d (7 TeV) $= 0.239 \pm 0.008$, f_s/f_d (13 TeV) $= 0.254 \pm 0.008$

  main systematic for $\mathscr{B}(B^0_s \to \mu^+\mu^-)$: reduced from $\sim 6\,\%$ (2017 analysis) to $\sim 3\,\%$

- Combining the two normalisation channels we obtain the following "single-event sensitivities" :

$$\alpha_{B^0_s\to\mu^+\mu^-} = (2.49 \pm 0.09) \times 10^{-11}$$

$$\alpha_{B^0\to\mu^+\mu^-} = (6.52 \pm 0.11) \times 10^{-12}$$

$$\alpha_{B^0_s\to\mu^+\mu^-\gamma} = (2.98 \pm 0.11) \times 10^{-11}$$

- Assuming SM signals we expect, in the full BDT range:

$$N(B^0_s \to \mu^+\mu^-)_{\rm SM} = 147 \pm 8$$

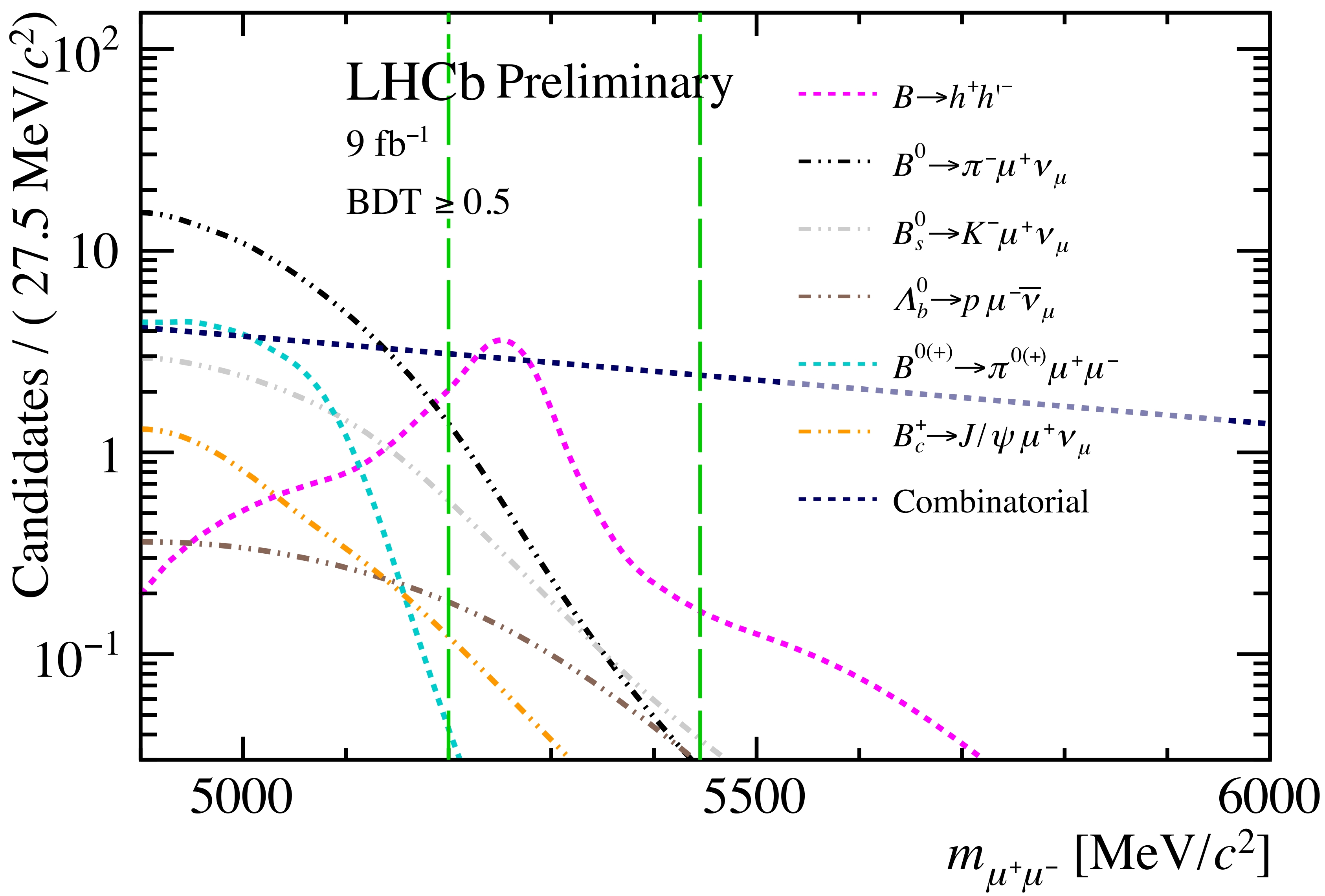
$$N(B^0 \to \mu^+\mu^-)_{\rm SM} = 16 \pm 1$$

$$N(B^0_s \to \mu^+\mu^-\gamma)_{\rm SM} \approx 3$$

Backgrounds

After applying a strong PID cut on both muon tracks, three classes of backgrounds remain:

1. Combinatorial, over the full mass spectrum (floating component in the fit)
2. Semileptonic backgrounds (partially reconstructed) populating the left mass sideband
3. $B^0_{(s)} \to h^+h^{'-} \to \mu^+\mu^-$ doubly misidentified background, peaking in signal mass region



- In the fit, 2. and 3. are gaussian-constrained to estimates

Semileptonic background estimate

[LHCB-PAPER-2021-007]

1. Channels with one misidentified hadron: $B^0 \to \pi^-\mu^+\nu_\mu$, $B^0_s \to K^-\mu^+\nu_\mu$ and $\Lambda^0_b \to p\mu^-\overline{\nu}_\mu$
2. Channels with two muons in the final state: $B^{+(0)} \to \pi^{+(0)}\mu^+\mu^-$ and $B^+_c \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$

- Each source is estimated by normalising to $B^+ \to J/\psi K^+$ events:

$$N_x = N_{B^+\to J/\psi K^+} \frac{f_x}{f_d} \frac{\mathcal{B}_x}{\mathcal{B}_{B^+\to J/\psi K^+}} \frac{\epsilon_x^{Tot}}{\epsilon^{Tot}_{B^+\to J/\psi K^+}}$$

- Efficiency corrected $B^+ \to J/\psi K^+$ yield
- $\mathcal{B}$ X hadronisation fraction
- Total background efficiency (misID calibrated on data samples)

- Inputs from LHCb measurements in the past years, lots of previously unknown channels !

- Estimated background events in the high BDT region ($\mathrm{BDT} \geq 0.5$) :

$B^0 \to \pi^-\mu^+\nu_\mu$	91 ± 4	[PDG]
$B^0_s \to K^-\mu^+\nu_\mu$	23 ± 3	[PRL 126 (2021) 081804]
$\Lambda^0_b \to p\mu^-\overline{\nu}_\mu$	4 ± 2	[Nature Physics 10 (2015) 1038]
$B^{+(0)} \to \pi^{+(0)}\mu^+\mu^-$	26 ± 3	[JHEP 10 (2015) 034] & [PRD 86 (2012) 114025]
$B^+_c \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$	7.2 ± 0.3	[PRD 100 (2019) 112006]

$B^0_{(s)} \to h^+h'^- \to \mu^+\mu^-$ background estimate [LHCB-PAPER-2021-007]

- B decays to two hadrons (π, K) form a peaking background when both final-state particles are misidentified as muons
- This contribution is estimated by normalising to $B^0 \to K^-\pi^+$ events:

$$N_{B\to hh\to\mu\mu} = \frac{N_{B^0\to K^+\pi^-}}{\epsilon^{\rm trig}_{B^0\to K^+\pi^-}} \times \frac{1}{f_{B^0\to K^+\pi^-/B\to hh}} \times \epsilon^{\rm trig}_{B^0\to\mu^+\mu^-} \times \epsilon_{hh\to\mu\mu}$$

- Efficiency corrected $B^0 \to K^+\pi^-$ yield
- $B^0 \to K^+\pi^-$ contribution within the total $B^0_{(s)} \to h^+h'^-$ [PDG]
- Trigger efficiency and double misidentification rate (from calibration data samples)

- Each $B \to hh$ channel is weighted according to its expectation to make the total $B^0_{(s)} \to h^+h'^- \to \mu^+\mu^-$
- An alternative estimate is performed on $h\mu$ data (single misID) to cross check the result

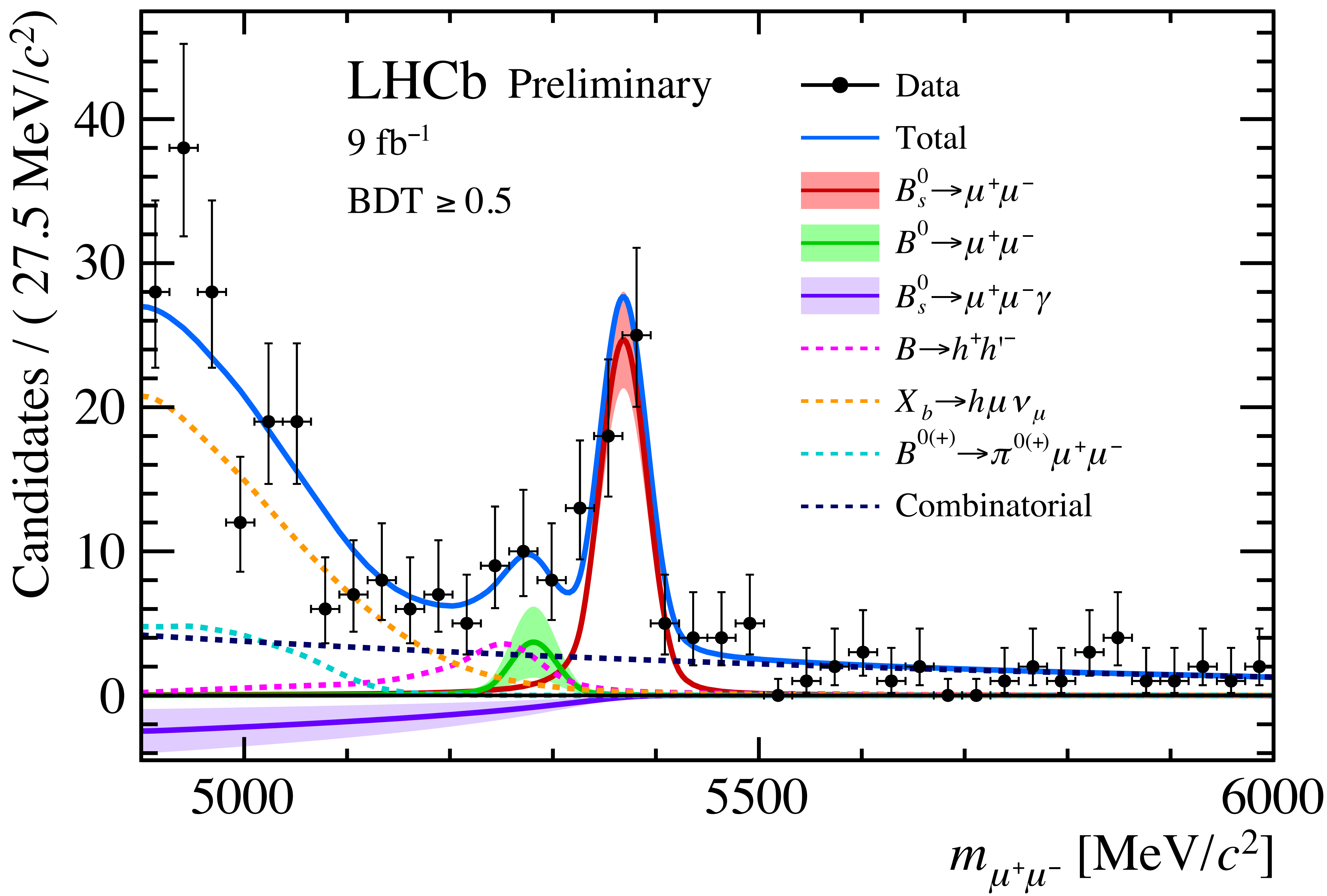
- Estimated background events in the high BDT region $(\mathrm{BDT} \geq 0.5)$:

$$B^0_{(s)} \to h^+h'^- \to \mu^+\mu^- \ : \ 22 \pm 1$$

- now we're ready for the fit!

Mass fit result

[LHCB-PAPER-2021-007]



- $\mathscr{B}(B_s^0 \to \mu^+\mu^-) = (3.09^{+0.46+0.15}_{-0.43-0.11}) \times 10^{-9} \quad (10.8\sigma)$

- $B^0 \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-\gamma$ compatible with background only at 1.7σ and 1.5σ

Branching fraction results

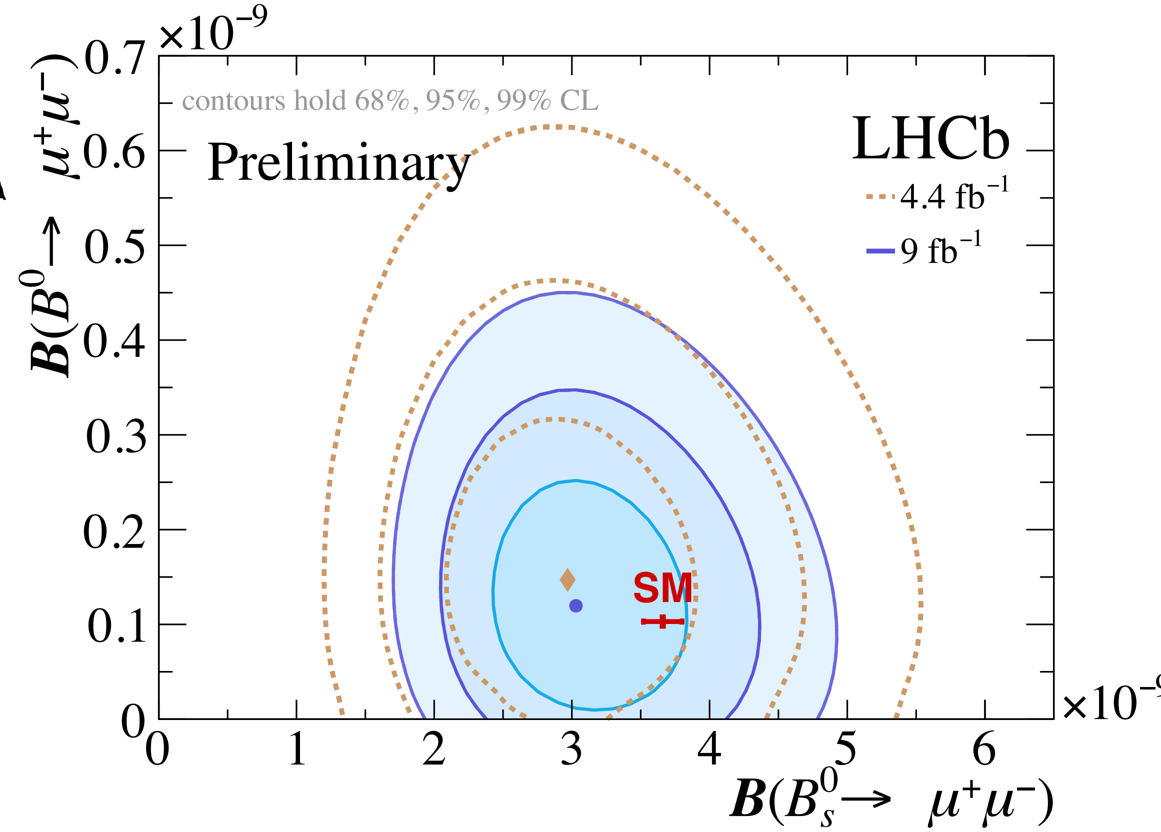
[LHCB-PAPER-2021-007]

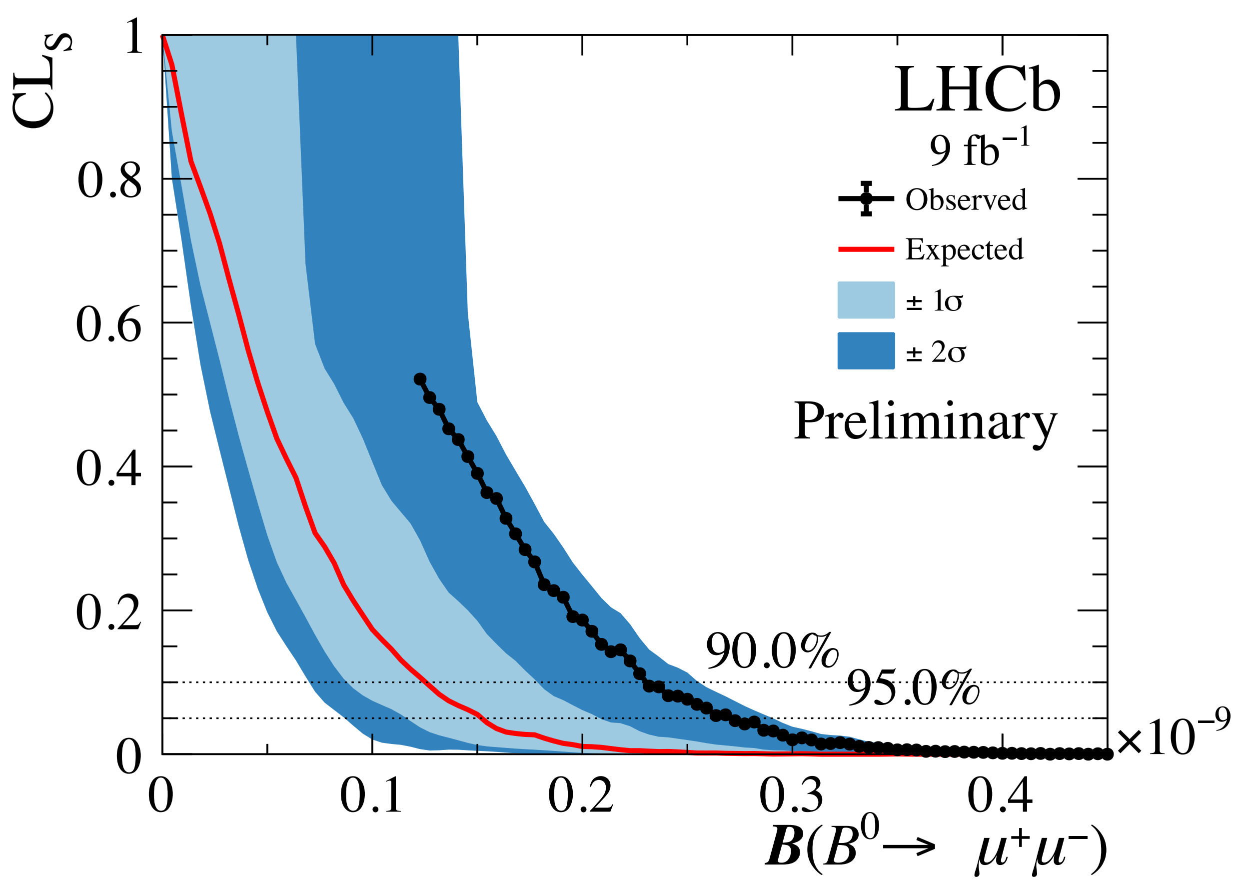
- $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.09^{+0.46+0.15}_{-0.43-0.11}) \times 10^{-9}$ spot on previous LHCb result and SM compatible
- Limits set with the $\mathrm{CL_s}$ method:

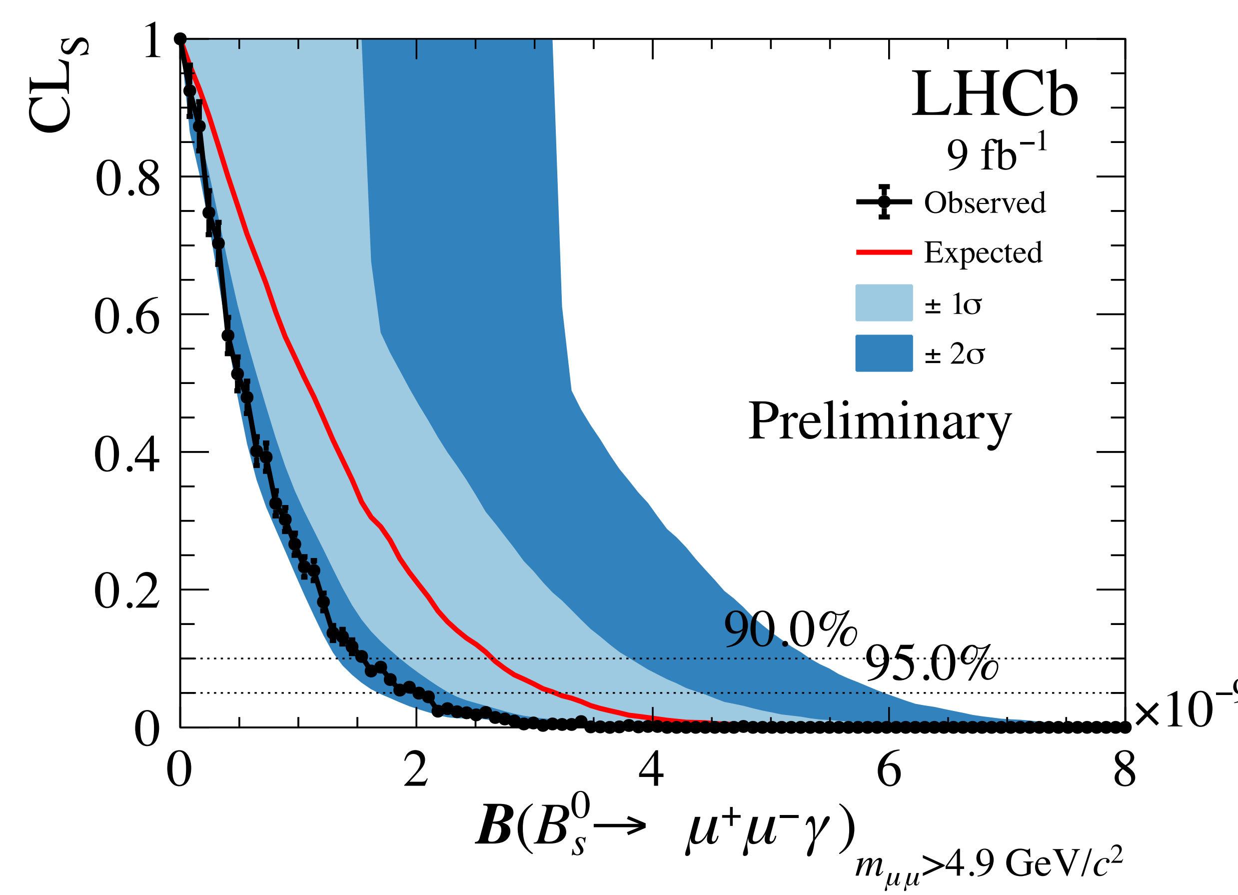
[J. Phys. G28 (2002) 2693]

$$\mathcal{B}(B^0 \to \mu^+\mu^-) < 2.6 \times 10^{-10}\ (95\,\%\ \mathrm{CL})$$

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-\gamma)_{m_{\mu^+\mu^-}>4.9\ \mathrm{GeV}} < 2.0 \times 10^{-9}\ \ (95\,\%\ \mathrm{CL})$$





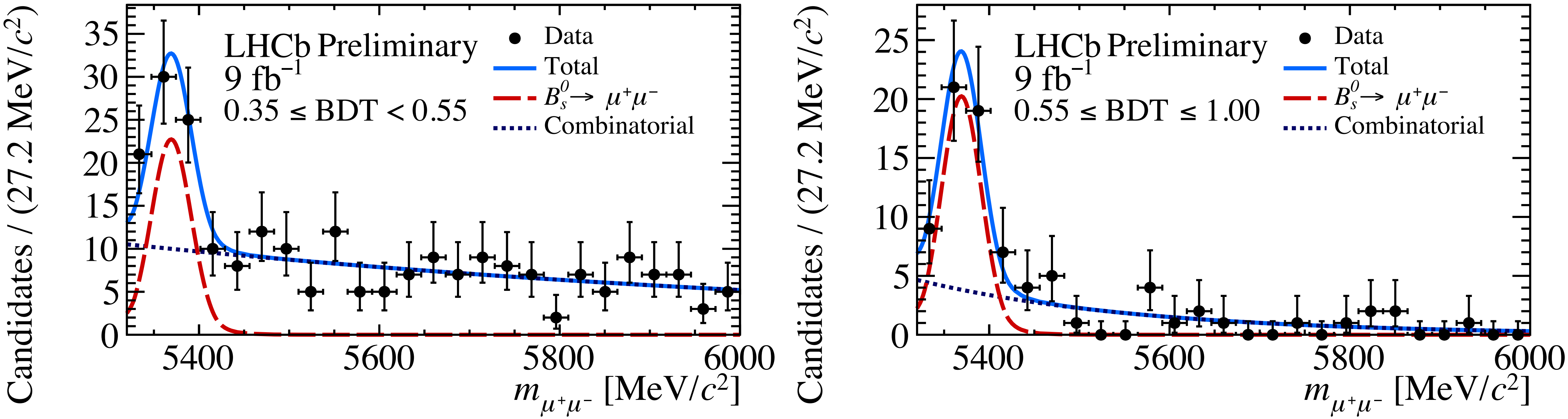


$B_s^0 \to \mu^+\mu^-$ effective lifetime: strategy

[LHCB-PAPER-2021-007]

Since the expected sensitivity on $A_{\Delta\Gamma}^{\mu^+\mu^-}$ is low, the effective lifetime measurement introduces some simplifications wrt the previous strategy:

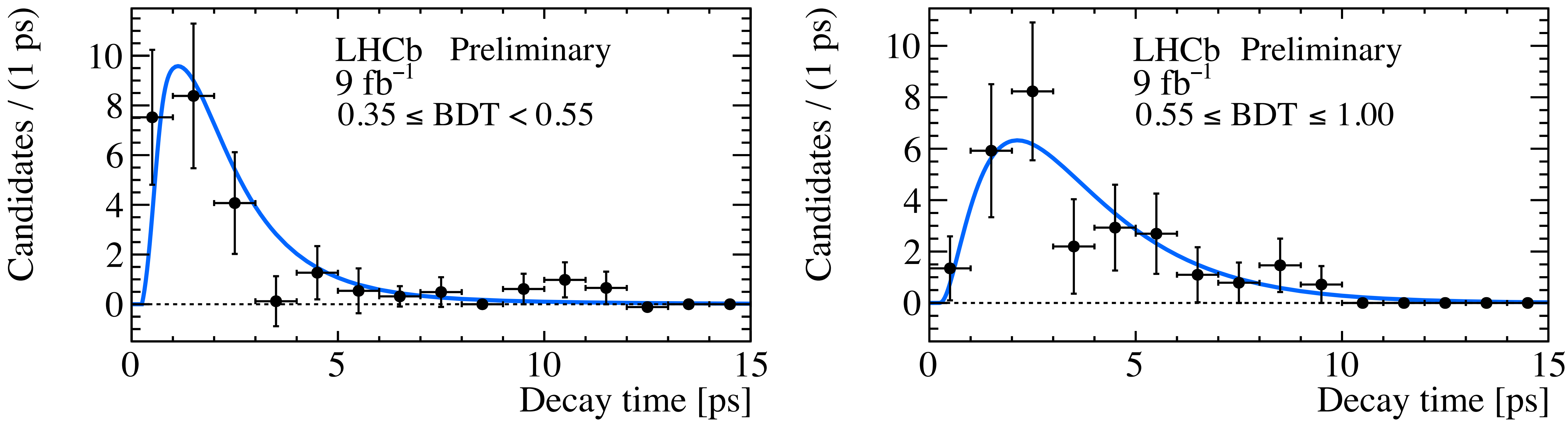
- Tighter mass cut, $m_{\mu^+\mu^-} > 5320$ MeV: mass fit model with $B_s^0 \to \mu^+\mu^-$ signal + combinatorial
- Looser PID requirement (negligible misID backgrounds)
- 1. Mass fit on two BDT bins is performed to extract sWeights [NIM A555 (2005) 356–369]



$B_s^0 \to \mu^+\mu^-$ effective lifetime: results

[LHCB-PAPER-2021-007]

- 2. The sWeights are applied to obtain the background-subtracted decay time distribution
- which is then fitted with an exponential X acceptance function



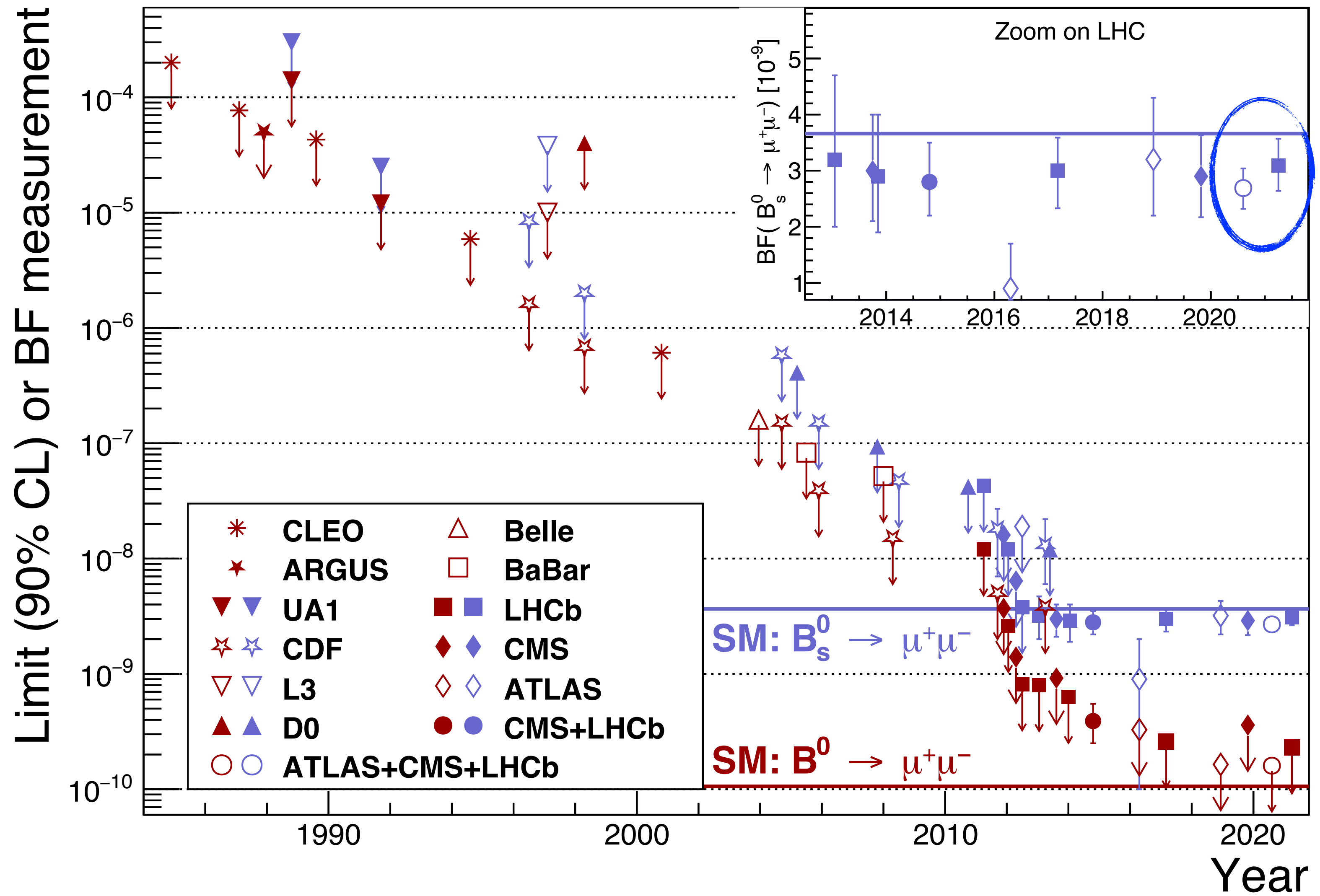
- The acceptance function (efficiency vs decay time) is tested by measuring the known $B^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ effective lifetimes (see → backup)

$$\tau_{\mu^+\mu^-} = 2.07 \pm 0.29 \pm 0.03 \text{ ps}$$

- Result compatible at 1.5σ with $A\Delta_\Gamma^{\mu^+\mu^-} = +1$ (SM) and at 2.2σ with $A\Delta_\Gamma^{\mu^+\mu^-} = -1$
- Run 3 data are needed for a stronger constraint

Conclusions

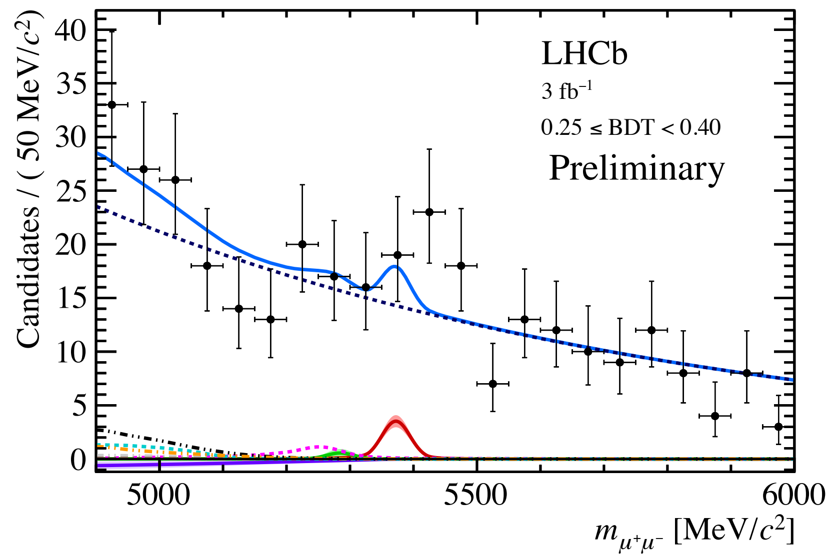
- The legacy measurement of $B^0_{(s)} \to \mu^+\mu^-$ represents an important milestone for LHCb and a crucial player for the "flavour anomalies"
- Achieved the most precise single-experiment measurement of the $\mathscr{B}(B^0_s \to \mu^+\mu^-)$ with $\sim 15\,\%$ error

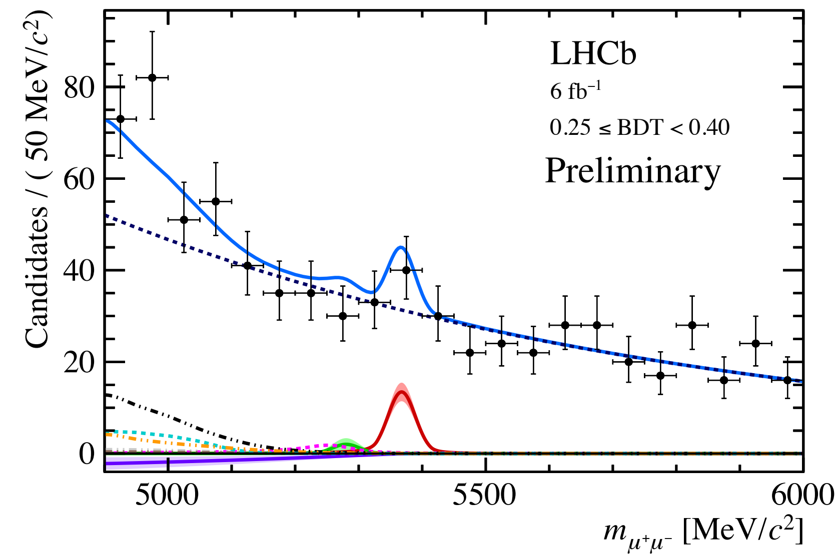


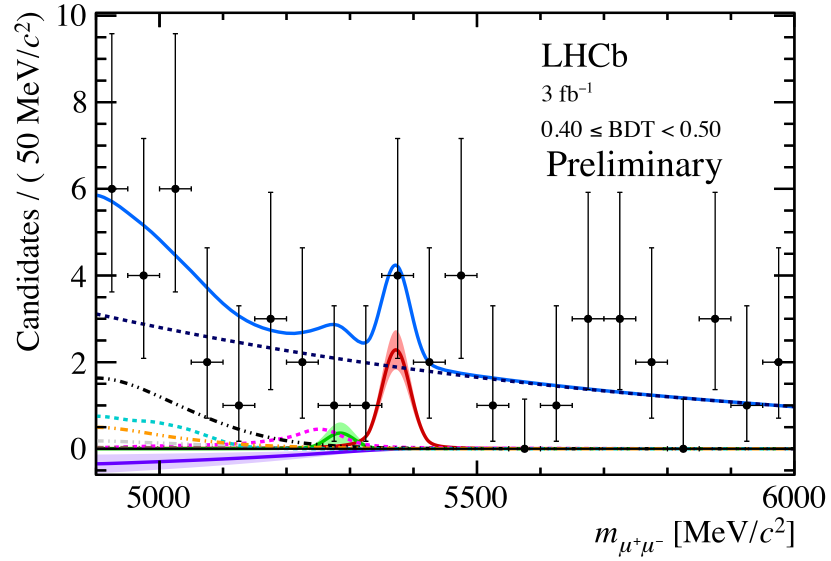
- Most precise measurement of $\tau_{\mu^+\mu^-}$ & first limit on $B^0_s \to \mu^+\mu^-\gamma$ ISR at high $m_{\mu^+\mu^-}$
- $\mathscr{B}(B^0 \to \mu^+\mu^-)$ limit at $2.5\times$ the SM prediction: its observation in Run 3 heavily relies on the PID
- ATLAS+CMS+LHCb combination with full Run 1 + Run 2 data to achieve $< 10\,\%$ precision!

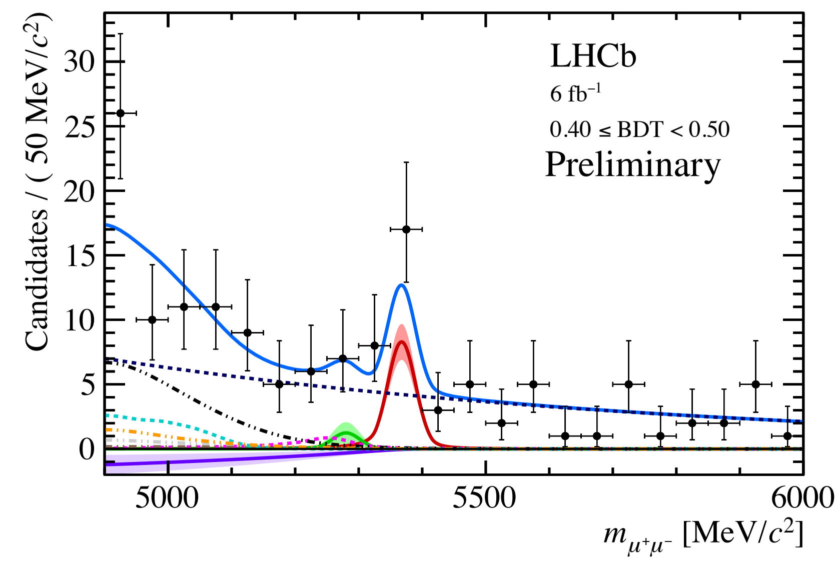
backup slides

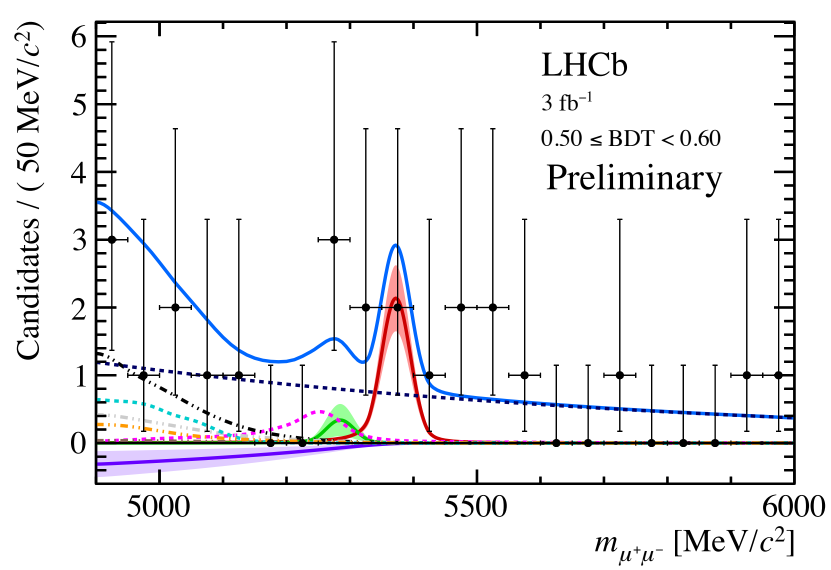
Mass fits: low BDT regions

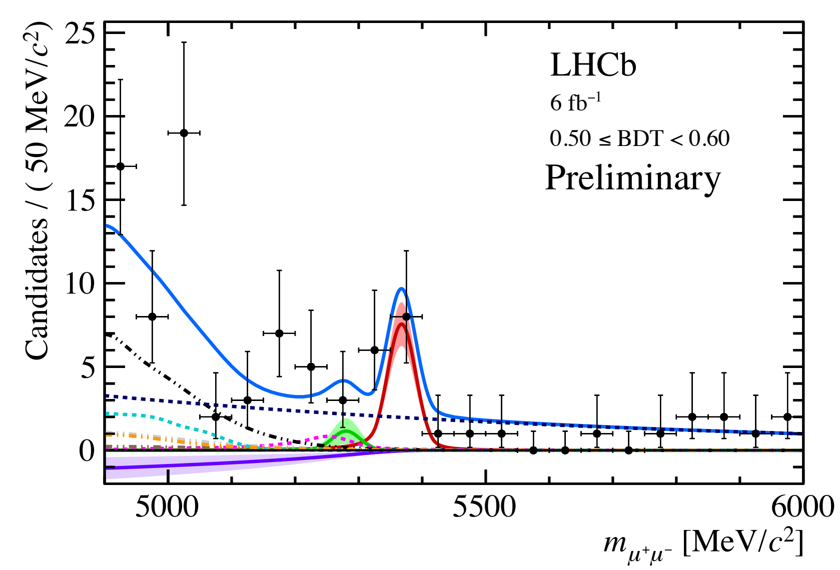




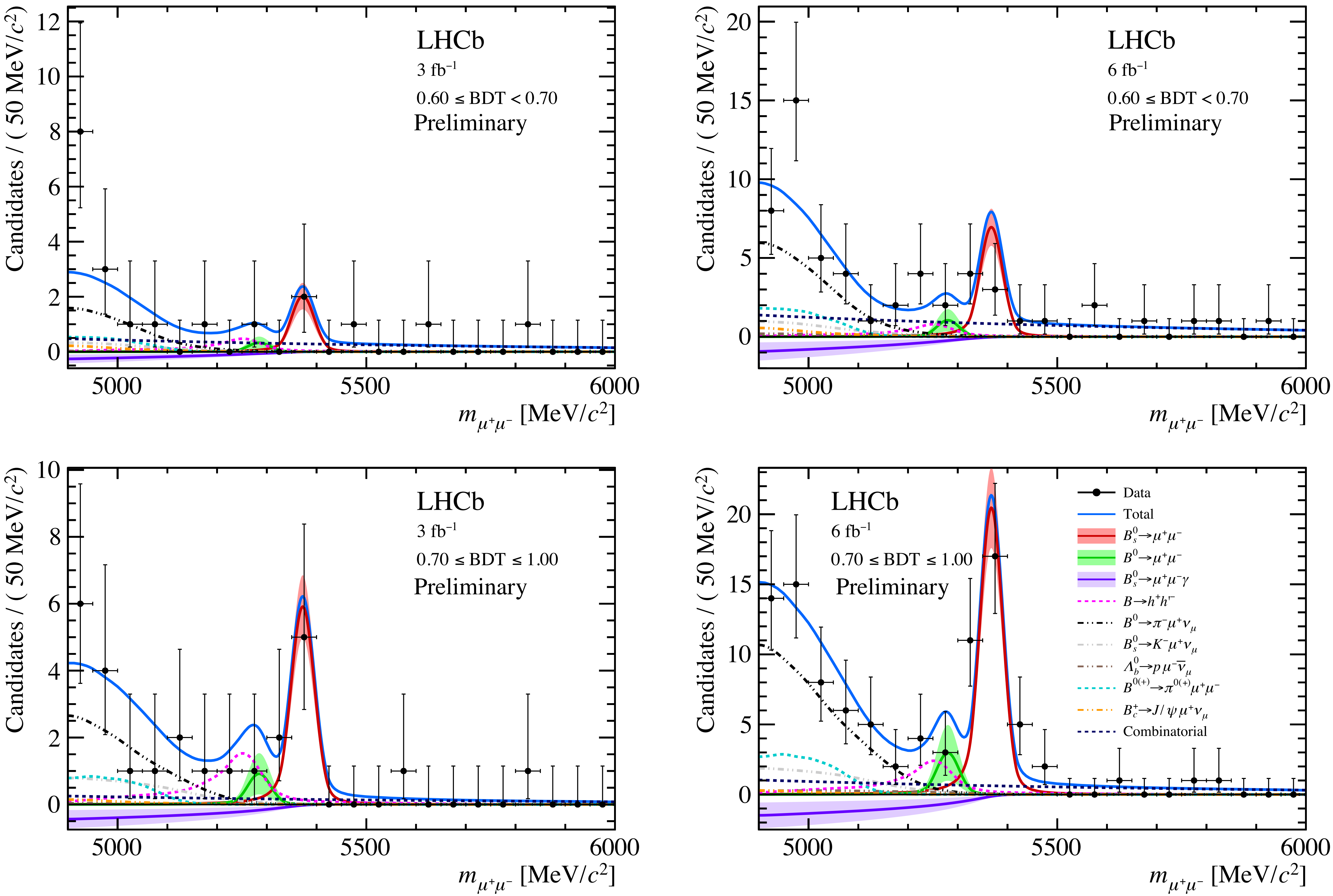






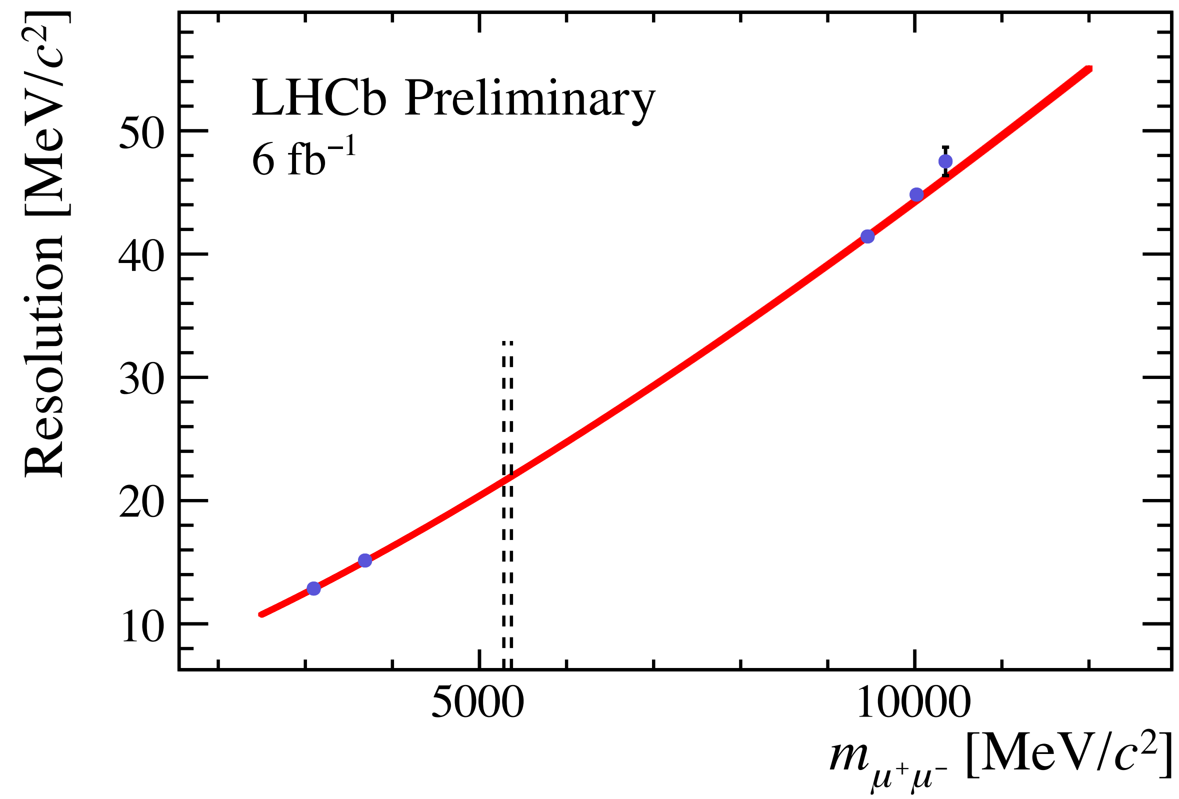


Mass fits: high BDT regions

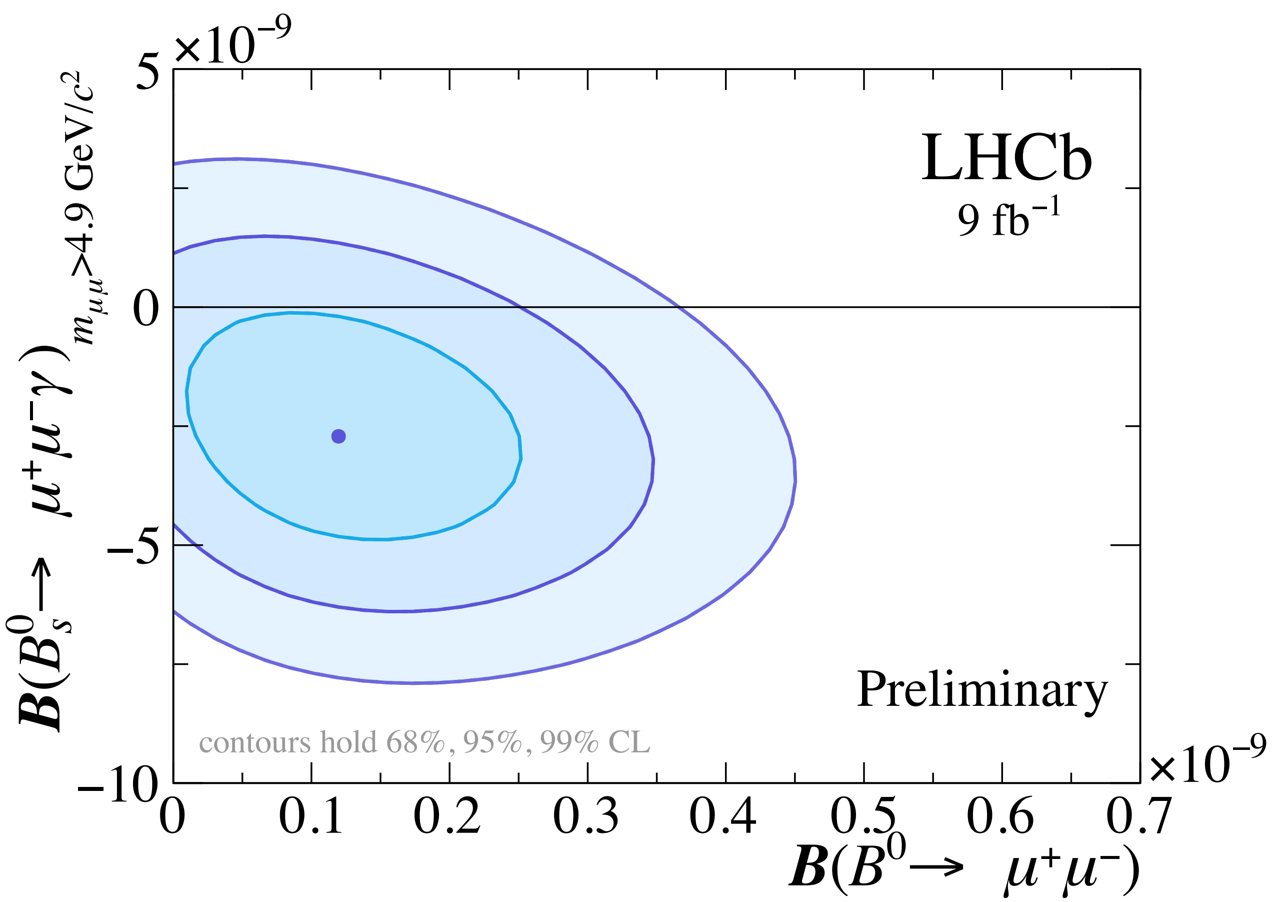


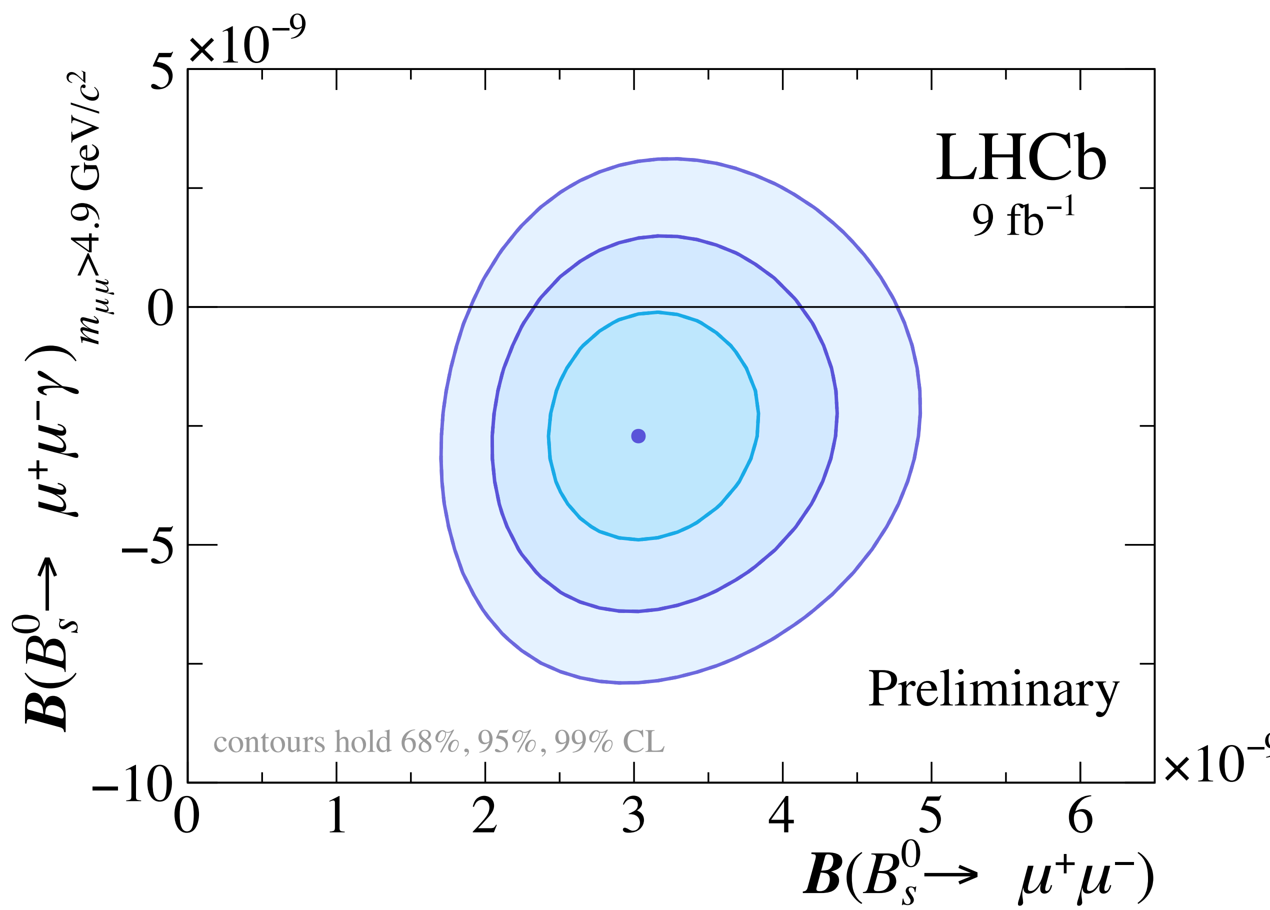
Additional material

- Power-law Interpolation of the resolution from $c\bar{c}$ and $b\bar{b}$ resonances
- --- B^0 and B_s^0 masses

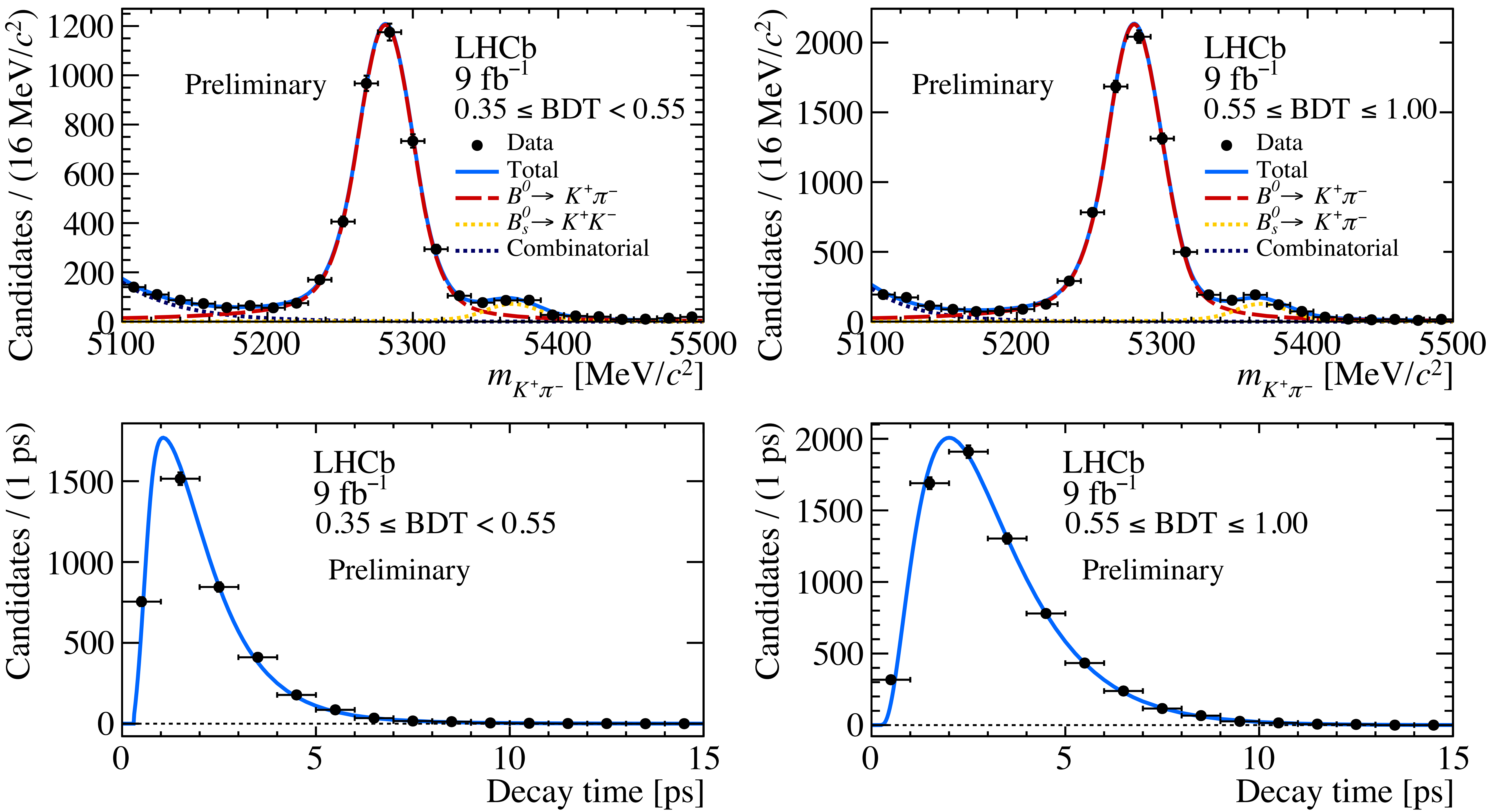


- 2D likelihood scans



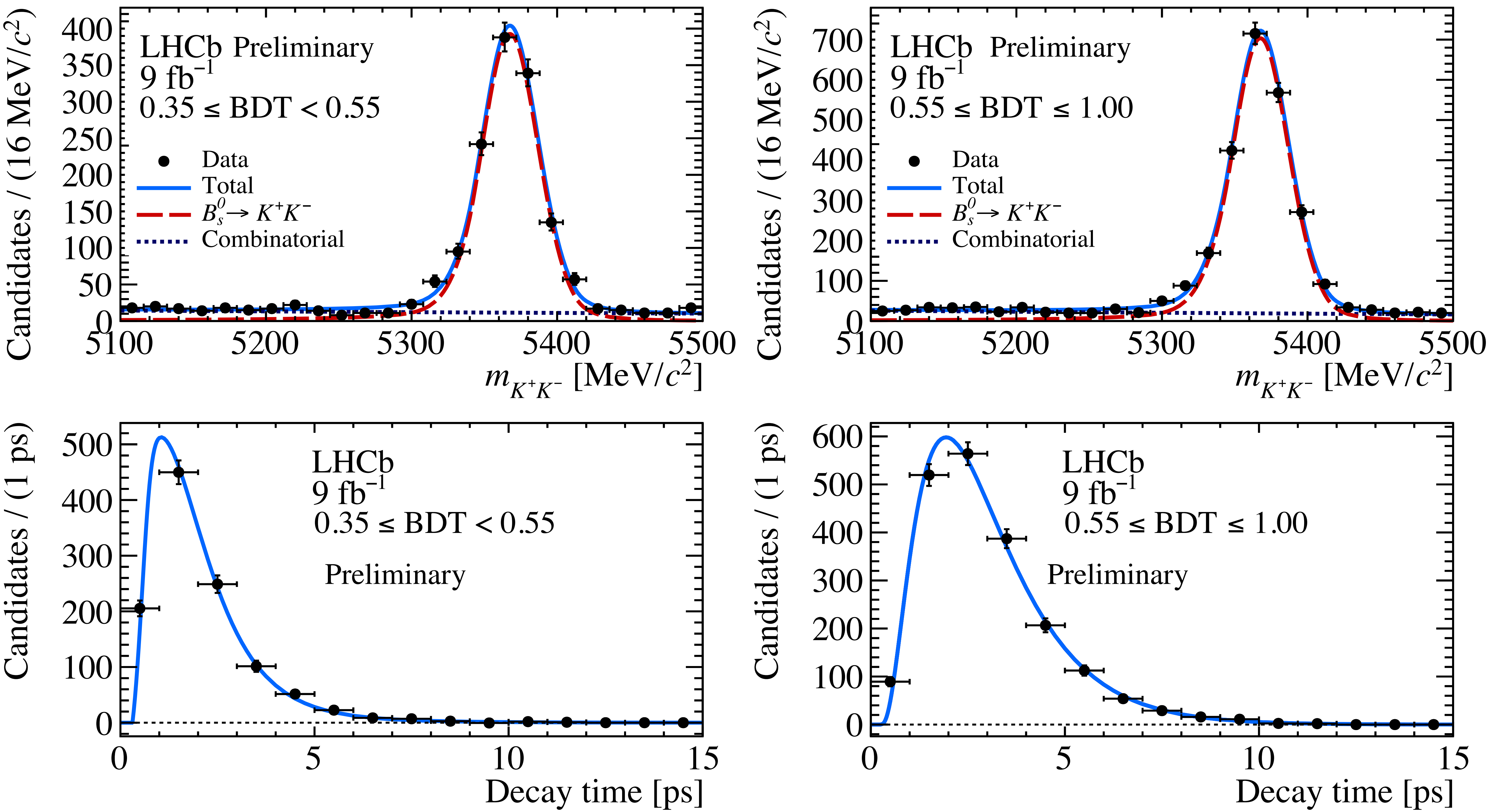


Effective lifetime of $B^0 \to K^+\pi^-$ decays



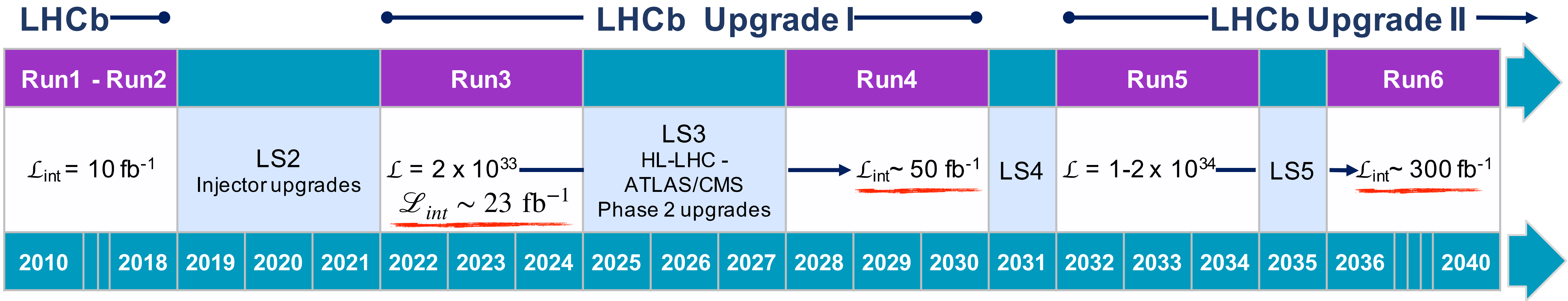
$$\tau_{K^+\pi^-} = 1.512 \pm 0.016 \text{ ps}$$

Effective lifetime of $B_s^0 \to K^+K^-$ decays

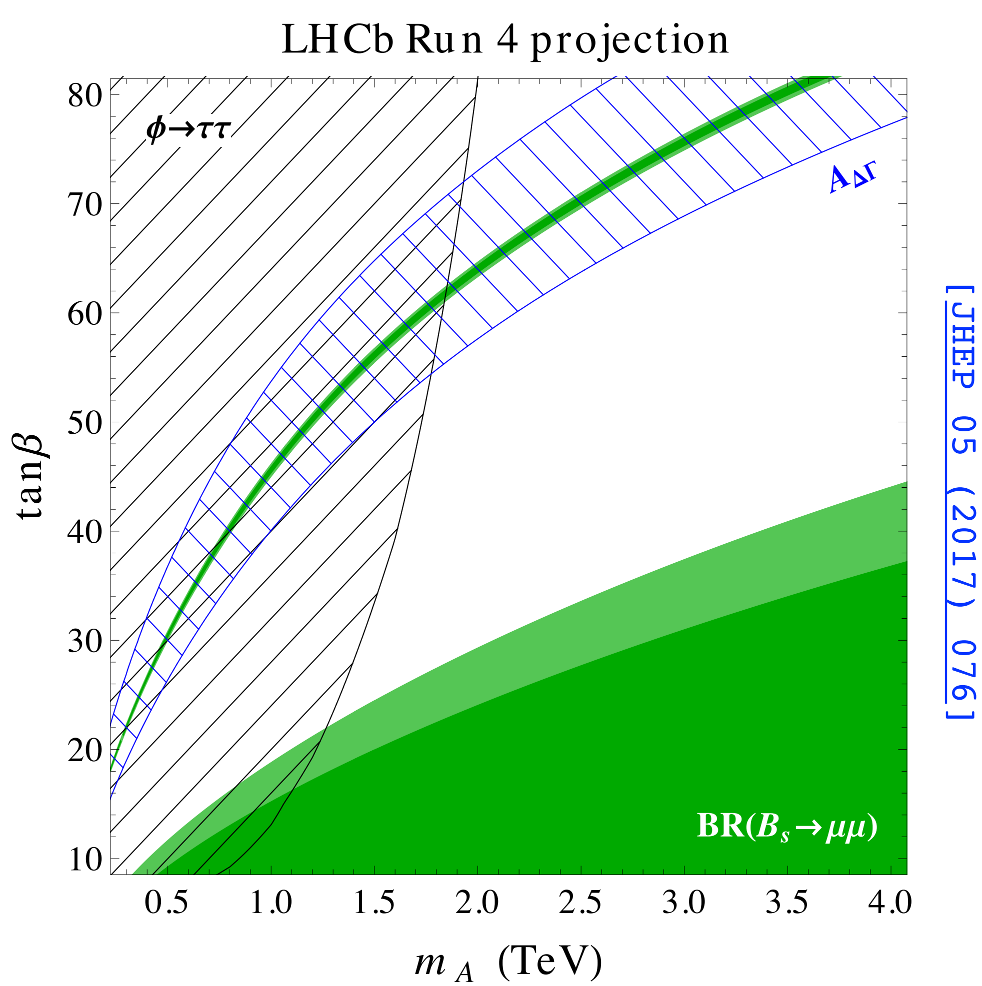


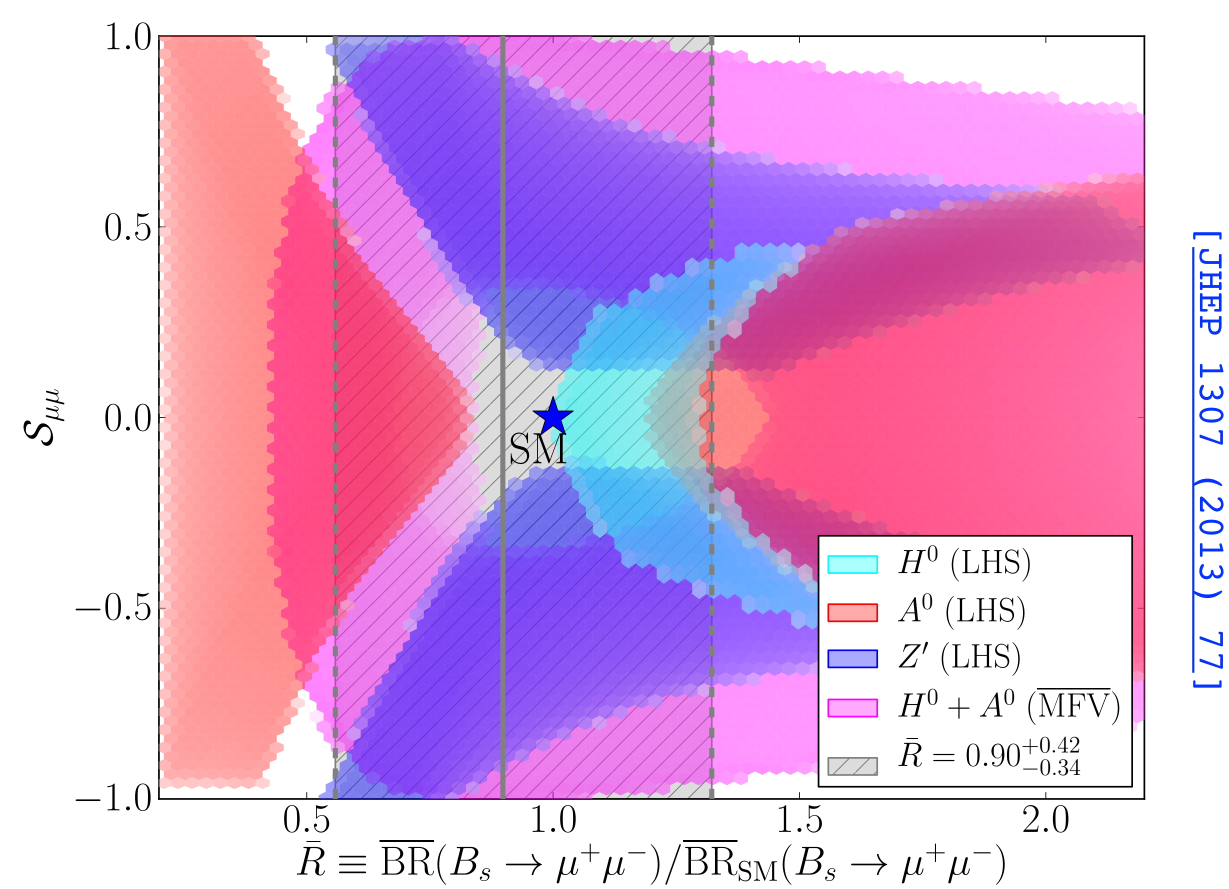
$$\tau_{K^+K^-} = 1.433 \pm 0.026 \text{ ps}$$

What's next?



- Combined power of $\mathscr{B}$ and $\tau_{\mu\mu}$ to constrain MSSM
- $\sim 20\,\%$ precision on the time-dependent CP asymmetry ($S_{\mu\mu}$) with $300\ \mathrm{fb}^{-1}$





$A_{\Delta\Gamma}^{\mu^+\mu^-}$ dependence & systematic errors

Lifetime acceptance correction for $B_s^0 \to \mu^+\mu^-(\gamma)$:

- The BDT-lifetime correlation is accounted for in the $B_s^0 \to \mu^+\mu^-(\gamma)$ signals with BDT corrections
- The nominal fit assumes $A_{\Delta\Gamma}^{\mu^+\mu^-} = +1$ (SM), but results under $A_{\Delta\Gamma}^{\mu^+\mu^-} = 0, -1$ will be published as well
- Translates into about $+5\,\%$ and $+11\,\%$ $\mathscr{B}(B_s^0 \to \mu^+\mu^-)$ value, respectively

---

Main source of systematic errors :

- $\mathscr{B}(B_s^0 \to \mu^+\mu^-)$: f_s/f_d
- $\mathscr{B}(B^0 \to \mu^+\mu^-)$: $B_{(s)}^0 \to h^+h'^- \to \mu^+\mu^-$ background
- $\mathscr{B}(B_s^0 \to \mu^+\mu^-\gamma)$: semileptonic backgrounds