Mysteries in flavour physics

- Flavour physics includes
  - neutrinos
  - charged leptons
  - kaon physics
  - charm & beauty physics
  - some aspects of top physics…

- Our focus will be on beauty physics

Monica Pepe Altarelli (CERN)
Academic Training Lectures
6-April 2022

In case you would like to know more

- V. Vagnoni: MISP 2019 lectures

- T. Gershon: CERN Summer Student Programme 2016 lecture 1, lecture 2, lecture 3, lecture 4

- G. Isidori: CERN, Summer Student Programme 2017, Lectures on Flavour Physics and CP Violation

- The recent “Flavour Anomalies in Heavy Quark Decays”, Albrecht, van Dyk, Langenbruch, arXiv:2107.04822 and many others

- Excellent book about to be published “New Physics in b decays”, by Artuso, Isidori and Stone, which I should acknowledge. This was something that kept Sheldon busy until the very last moments of his life.

Outline of the four lectures

1
- Introduction to flavour; concept of lepton flavour universality (LFU) and theory description of rare transitions of the type $b \to s\ell^+\ell^-$ (Gino Isidori)

2
- Why b physics and why it is interesting
- The CKM Matrix
- Neutral meson oscillations
- The main experiments in the field (discussing their benefits and possible shortcomings)
- Selected results on rare decays:
  - Leptonic ($B_{(s)} \to \mu^+\mu^-$)
  - Semileptonic ($b \to s\ell^+\ell^-$) (branching fractions and angular distributions)

"Anomalies"

3
- Tests of LFU
  - in neutral-current mediated $b \to s\ell^+\ell^-$ transitions
  - in charged current mediated $b \to c\tau\nu$ transitions
- Search for LFV processes

4
- Theoretical interpretation of the recent anomalies in semileptonic B decays in terms of physics beyond the Standard Model (Gino Isidori)

Why b (I)

- The heaviest quark that binds in hadrons
- Many decay channels: a vast laboratory
- Heavy mass → more theoretically accessible
  - $\Lambda_{\rm QCD}/m_b \sim 0.1$
  - Asymptotic freedom: $\alpha_s(m_b) \sim 0.2$

This allows systematic approximations, which are exploited in the various applications of heavy quark theory

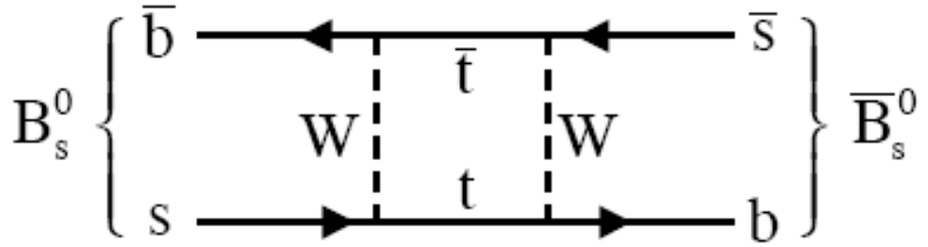
- Lifetime long enough for experimental detection → look for displaced vertices

$\tau \sim 1/(m^5 |V_{cb}|^2)$ $|V_{\rm cb}| \sim 0.04$
$\tau_b \sim 1.5 \cdot 10^{-12}$sec →
$d = \beta\gamma c\tau \sim$ 1 cm @LHC

- $\Upsilon(4S)$ is a clean source of B mesons at e^+e^- colliders

Why b (II)

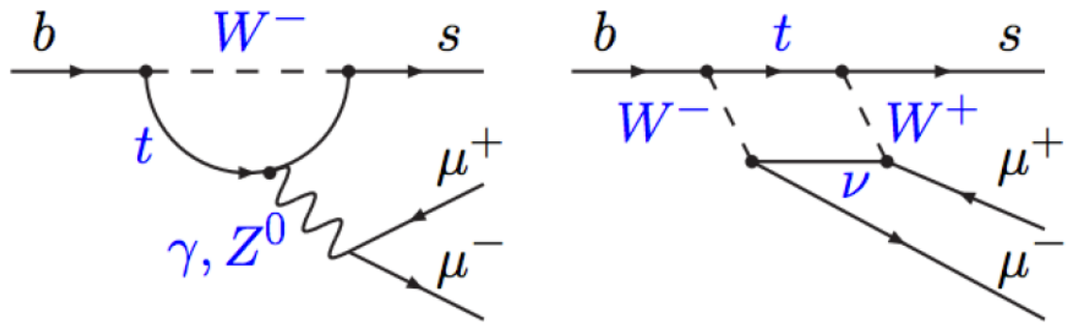
- Sizeable CP violation expected in many decays
  - Large CPV effects expected in processes that involve quarks from all three generations (quark mixing matrix cannot violate CP in a world with only two families!)
  - Quark loops are not suppressed
    (neither by GIM: $m_t > m_W$ nor by CKM: $|V_{tb} = 1|$)



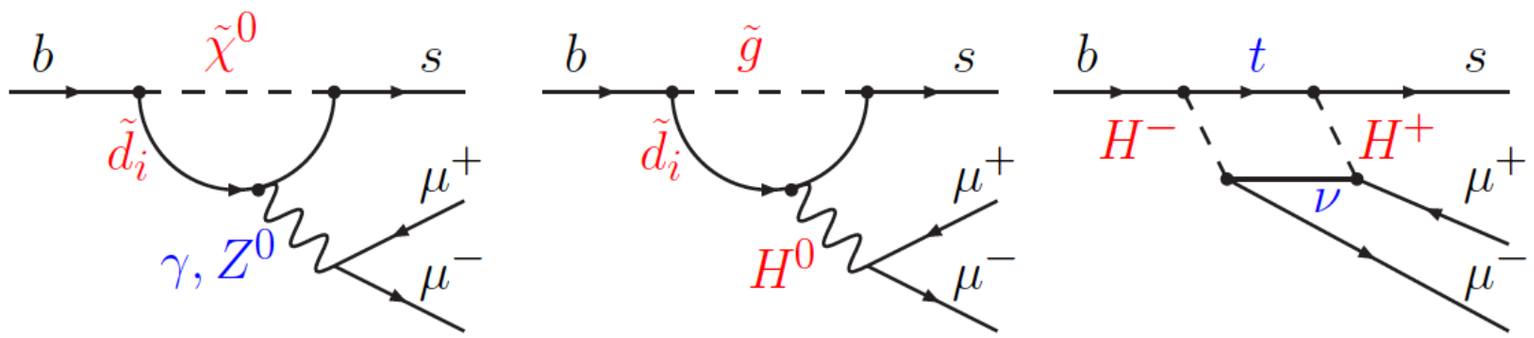
- Most TeV-scale New Physics (NP) models contain new sources of CP and flavour violation
- The observed baryon asymmetry of the Universe requires CPV beyond the SM
  - Not necessarily in flavour-changing processes, nor necessarily in quark sector, it could originate from lepton sector

Why b (III)

- In the SM, some decays are forbidden at tree level and can only occur at loop level (penguin and box), e.g. $B_s \to \mu^+\mu^-$ → Rare Flavour Changing Neutral Currents



- A new particle, too heavy to be produced at the LHC, can give sizeable effects when exchanged in a loop



- Strategy: use well-predicted observables to look for deviations
- Indirect approach to New Physics searches, complementary to that of ATLAS/CMS

A lesson from history

- New physics can show up at precision frontier before energy frontier
  - GIM mechanism before discovery of charm
  - CP violation and CKM before discovery of beauty and top
  - Neutral currents before the discovery of Z
- In general, a data-driven approach, in which we test precise SM predictions looking for discrepancies, has historically paved the way to important discoveries in particle physics.
- This approach is particularly relevant in the absence of direct collider production of new particles

CKM matrix

- V_{CKM} introduced yesterday by Gino Isidori: non trivial flavour mixing originating from the Higgs sector: $V_{ij} \rightarrow \delta_{ij}$ if we switch off the Yukawa interactions

- It describes the rotation between flavour (d', s', b') and mass (d, s, b) eigenstates

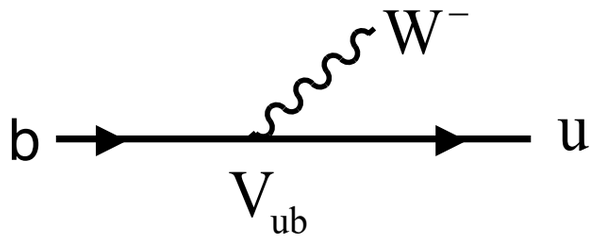
Flavour eigenstates

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

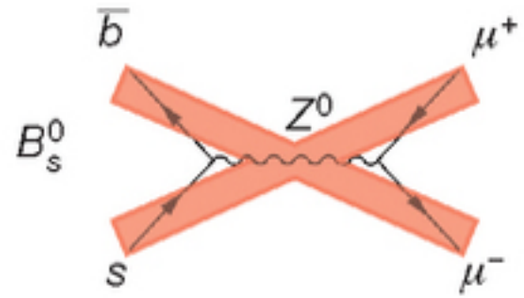
Mass eigenstates

3 real parameters & 1 complex phase (source of CPV)

- V_{ij} proportional to transition amplitude from quark i to quark $j \rightarrow V_{CKM}$ quark mixing matrix



- V_{CKM} induces flavour-changing transitions within and across generations in the charged sector at tree level ($W^{\pm}$ interaction). By contrast, there are no flavour-changing transitions in the neutral sector at tree level.

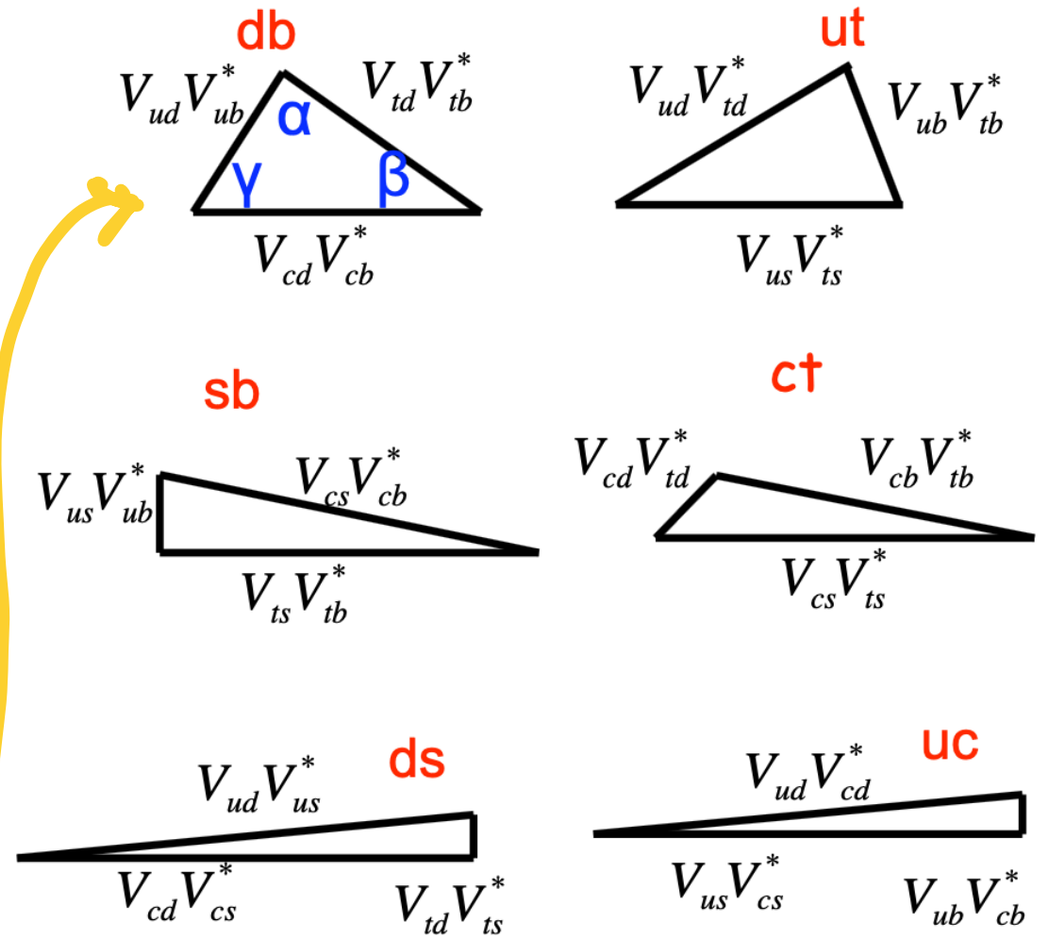


Unitarity conditions

- Unitarity of CKM matrix implies relations of the form

$$\sum_{i=1,3} V_{ij}V_{ik}^{*} = \delta_{j,k}, \text{ with } j \neq k$$

- Each of these 6 unitarity constraints can be seen as the sum of 3 complex numbers closing a triangle in the complex plane
- Particularly interesting when the three terms in the sum are of comparable magnitude.
- db triangle used to define angles α, β, γ (**Unitarity Triangle**)



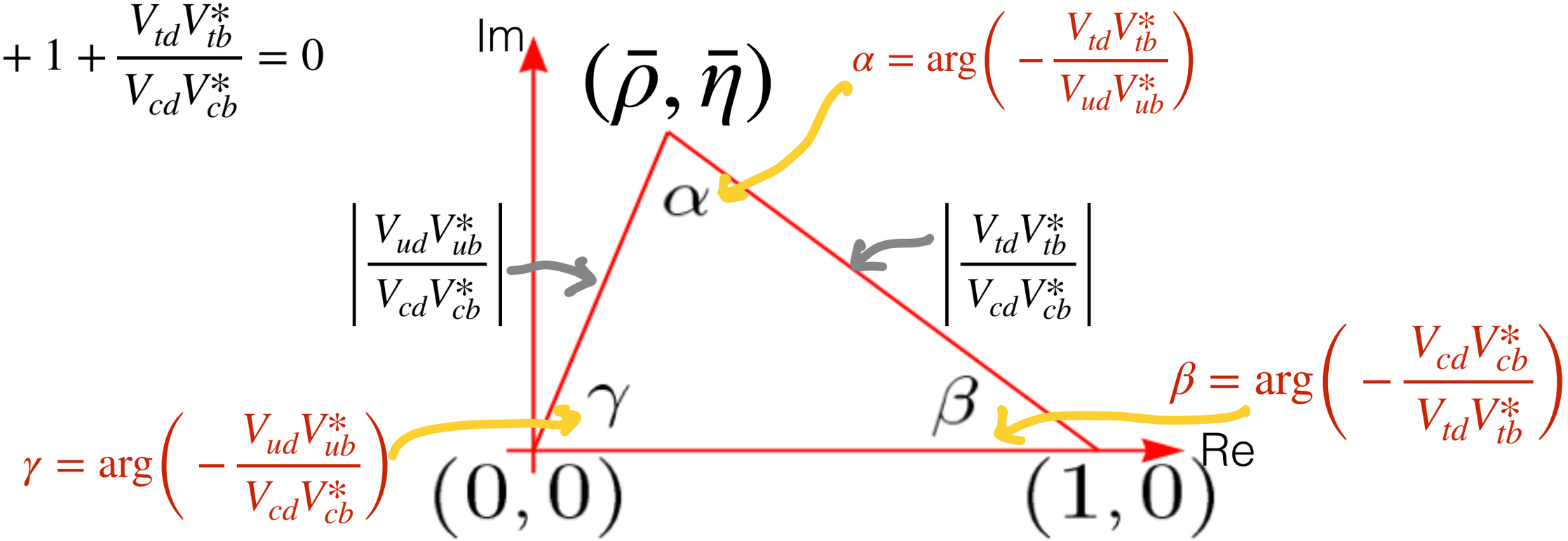
Unitarity Triangle (UT)

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$$

$$\mathcal{O}(\lambda^3) \qquad \mathcal{O}(\lambda^3) \qquad \mathcal{O}(\lambda^3)$$

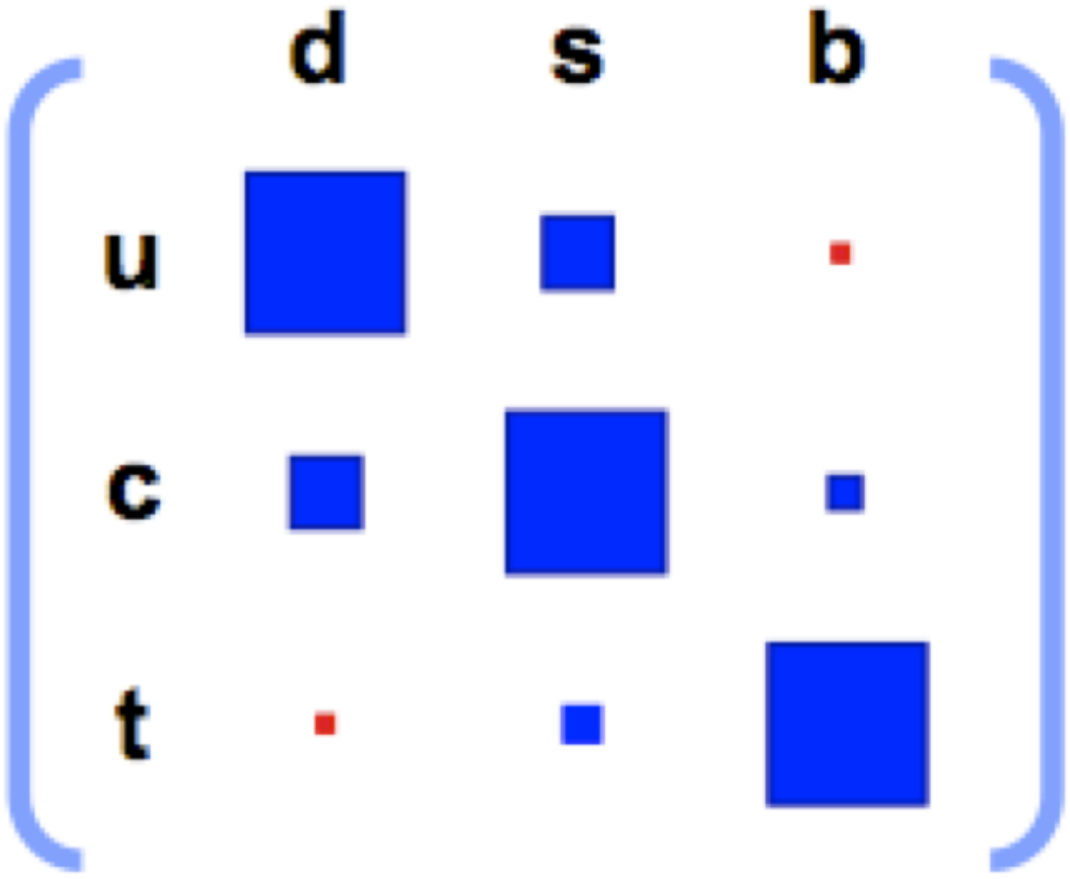
$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0$$

$$\lambda = \sin\theta_c \approx 0.22$$



- The triangle has vertices at $(0,0), (1,0), (\bar{\rho}, \bar{\eta})$ with $\bar{\rho} \equiv \rho(1-\lambda^2/2), \bar{\eta} \equiv \eta(1-\lambda^2/2)$
- CP violation in the quark sector is translated into a non flat UT
- Huge improvement in the knowledge of the CKM elements in the last decades!

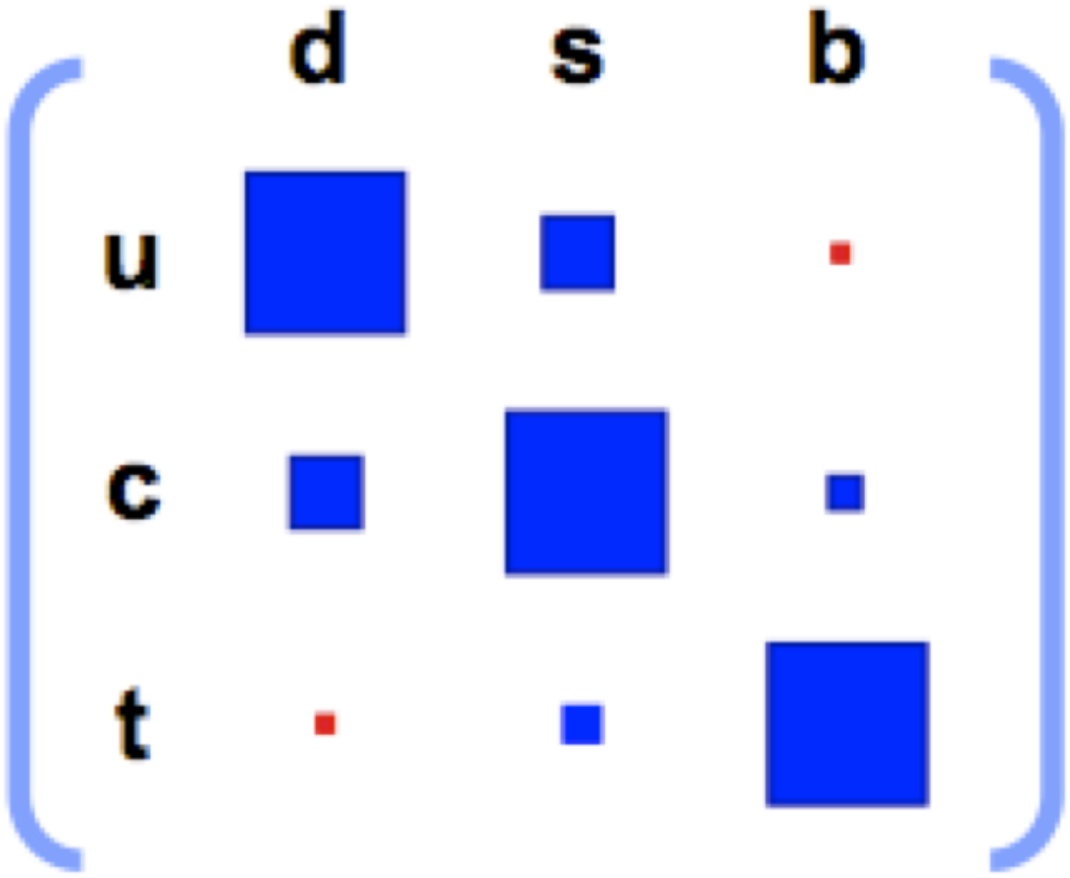
Hierarchy in quark mixing



- Each quark has a preference to transform into a quark of its own generation.
- Very suggestive pattern
- No known reasons
- Completely different in neutrino sector

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \sim \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)$$

Hierarchy in quark mixing



- Each quark has a preference to transform into a quark of its own generation.
- Very suggestive pattern
- No known reasons
- Completely different in neutrino sector

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \sim \begin{pmatrix} 0.97446 & 0.22452 & 0.00365 \\ 0.22438 & 0.97359 & 0.04214 \\ 0.00896 & 0.04133 & 0.999105 \end{pmatrix}$$

Consistency of CKM fits

- The physics impact of the measurements of the the CKM elements is not so much in their absolute values (matrix is not predicted) but rather in testing the (in)consistency of the “ensemble” of measurements and how precisely the SM description of flavour and CP violation holds.

- “Redundant" measurements are performed, which test different combinations of flavour parameters

- “What is particularly noteworthy in the so-called CKM fits is the consistency of the the tree-level determinations of CKM elements, with those obtained from loop observables, such as K-$\overline{K}$ or B-$\overline{B}$ mixing” (Gino Isidori)

.

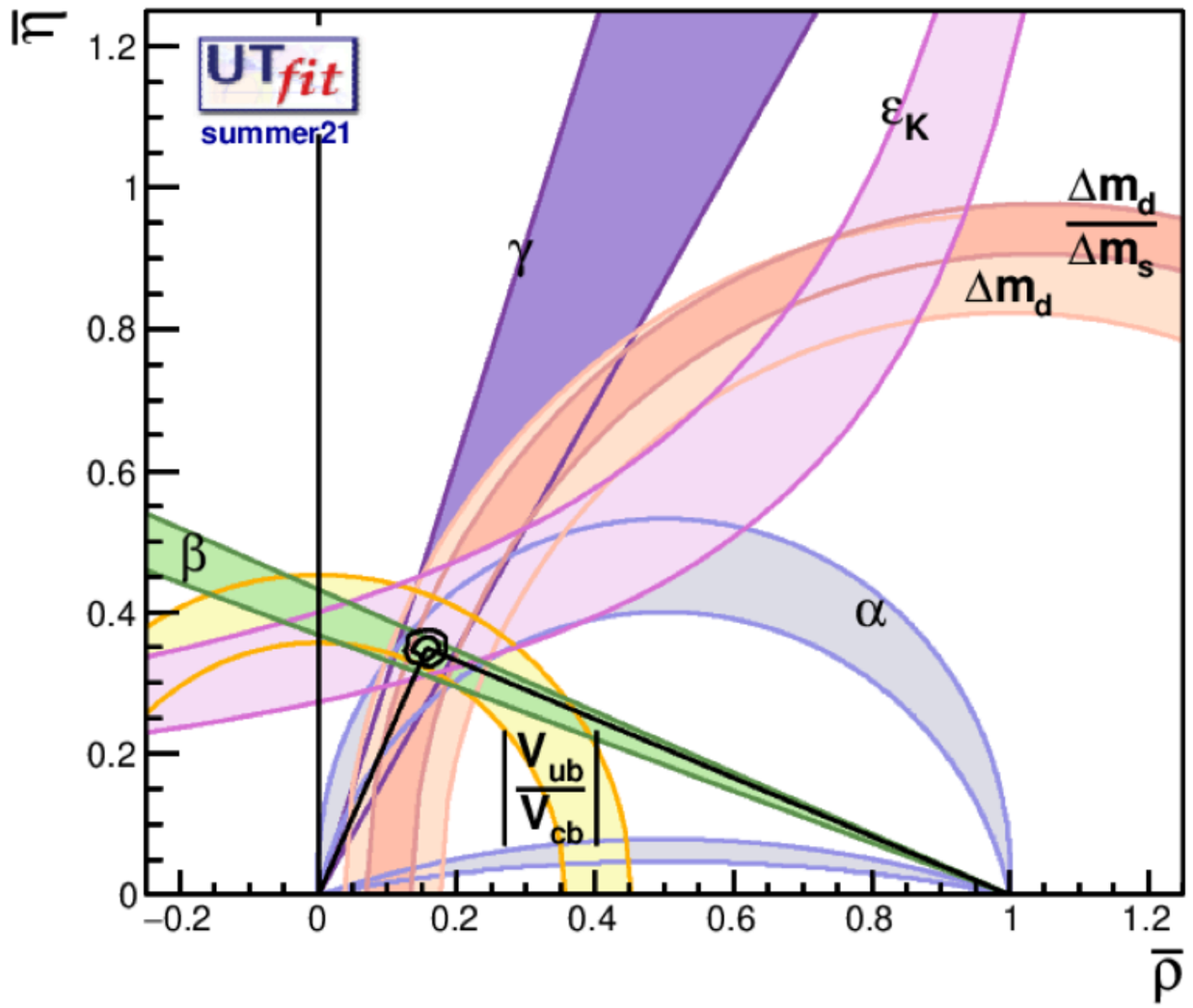
Consistency of CKM fits

- Impressive effort from community and tremendous success of CKM paradigm!

- Constraints from many different quark transitions. Extensive measurements on K, D and B mesons performed at different experiments. Constraints depend also on theory input.
- At the current level of precision, all measurements are consistent and intersect in the apex of the UT
- New Physics effects (if there) are small!

M.Bona EPS-HEP '21 UTfit, similar plots from CKMfit

Zooming in



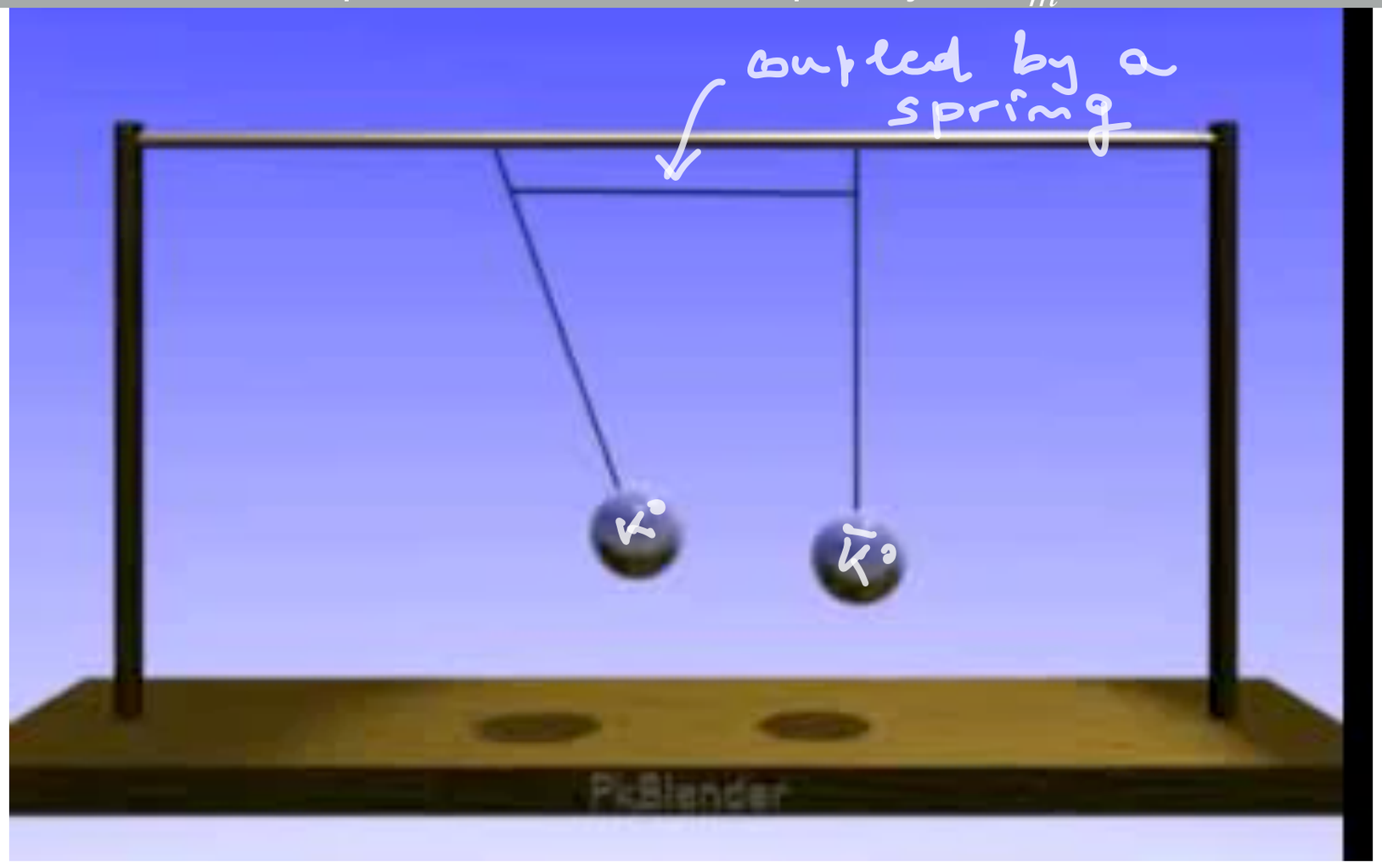
$$\bar{\rho} = 0.157 \pm 0.012 \quad \sim 8\%$$

$$\bar{\eta} = 0.350 \pm 0.010 \quad \sim 3\%$$

One example: neutral meson oscillations

Mechanical analogy: the coupled pendula

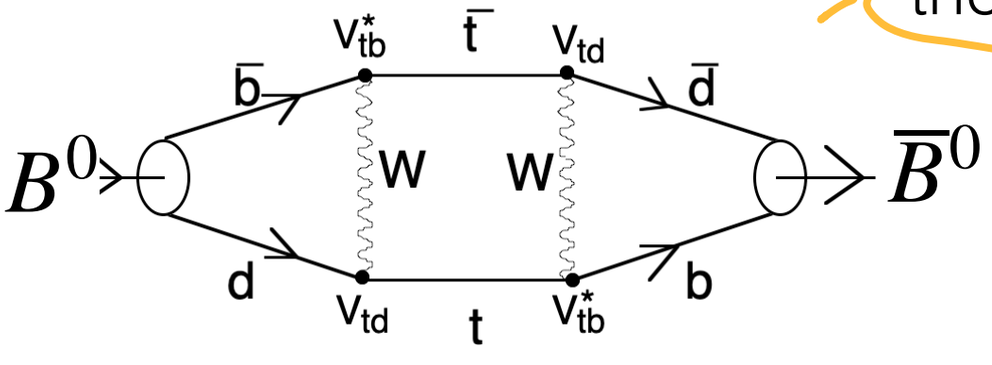
One pendulum may be thought of as the K^0 and the other as the $\overline{K}^0$. When one pendulum is excited, it will slowly transfer its energy to the other and back. This beating corresponds to the oscillation between a meson and its antiparticle. The beat frequency is Δ_m

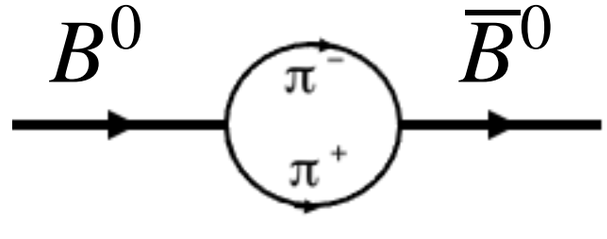


Neutral meson oscillations

- Flavour eigenstates $M^0, \overline{M}^0$ can mix into each other
  - via short-distance (box diagrams) or long-distance processes

the equivalent of the spring





$K^0 \leftrightarrow \overline{K}^0,\ D^0 \leftrightarrow \overline{D}^0,\ B^0 \leftrightarrow \overline{B}^0$

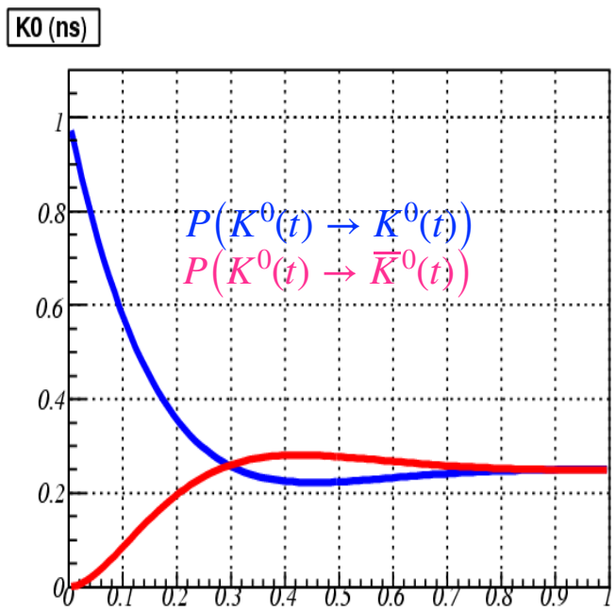
Large for K^0, small for D^0&B_d^0

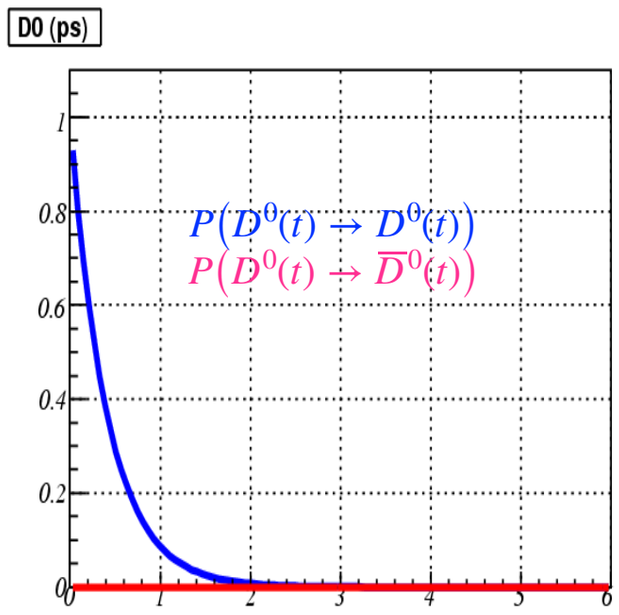
- $\Delta S = 2, \quad \Delta C = 2, \quad \Delta B = 2$

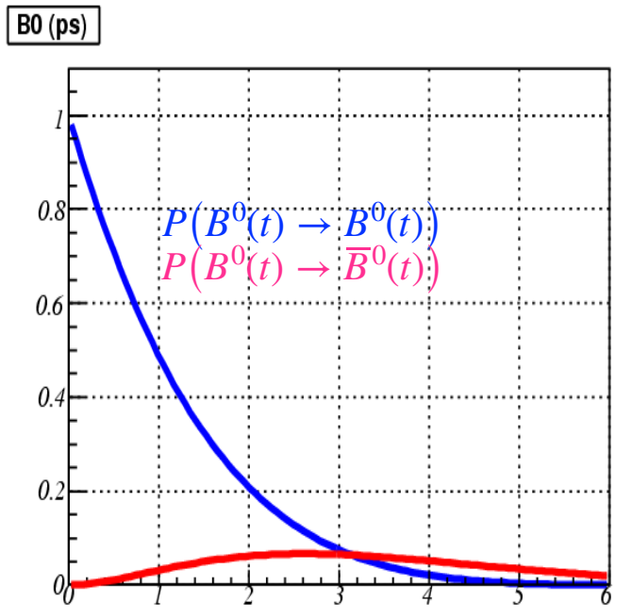
- Formalism is the same even if difference in mass and CKM elements results in dramatically different phenomenology

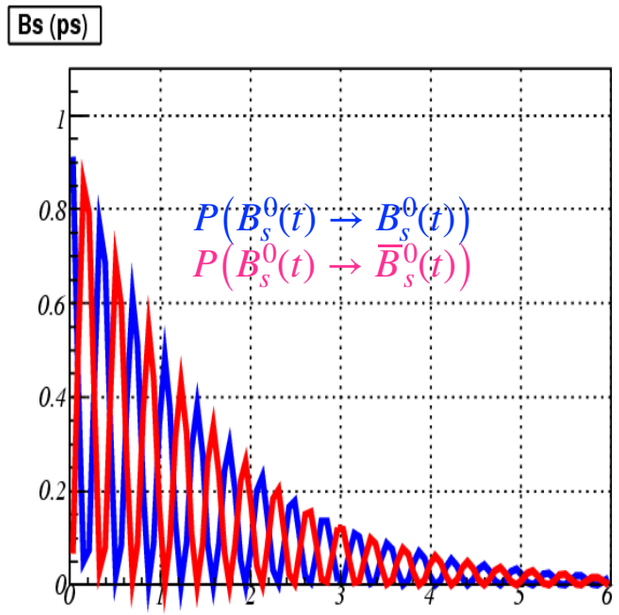
- Physical states: eigenstates of effective Hamiltonian
  $|M_{L,H}\rangle = p\,|M^0\rangle \pm q\,|\overline{M}^0\rangle$, with $|M^0\rangle, |\overline{M}^0\rangle$ flavour eigenstates,
  CP violation in mixing when $|q/p| \neq 1$, $\Delta m = m_H - m_L$, $\Delta\Gamma = \Gamma_L - \Gamma_H$

Compare the mesons









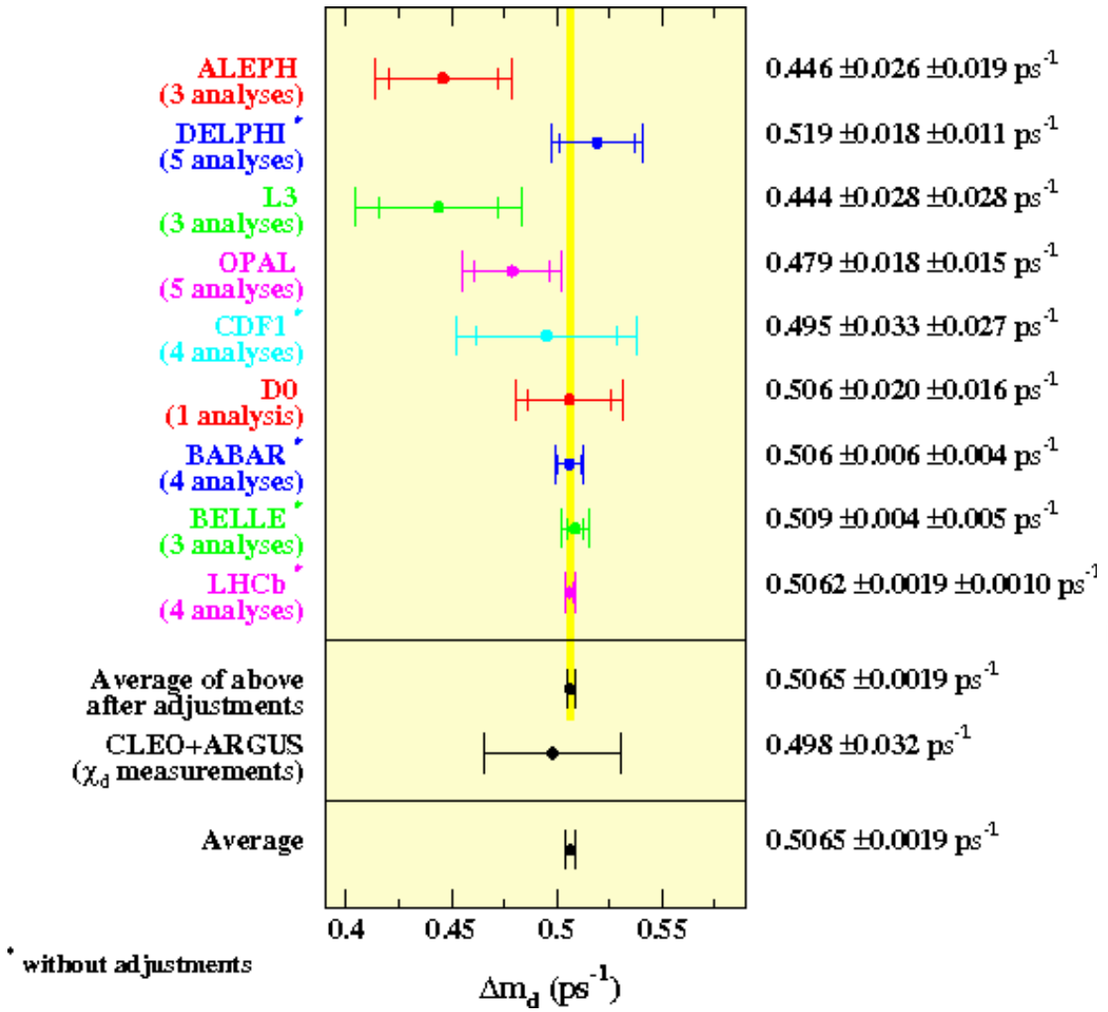
Probability to observe an M^0 or $\overline{M}^0$ at time t starting from a pure M^0 meson

- Oscillation frequency Δm depends on mixing rate, $\Delta\Gamma$ depends on widths of decays into common final states ($K^0 \to \pi^+\pi^- \to \overline{K}^0$) (large for K^0, small for D^0, B_d^0)
- $x = \Delta m/\Gamma$ gives the average number of oscillations before decay

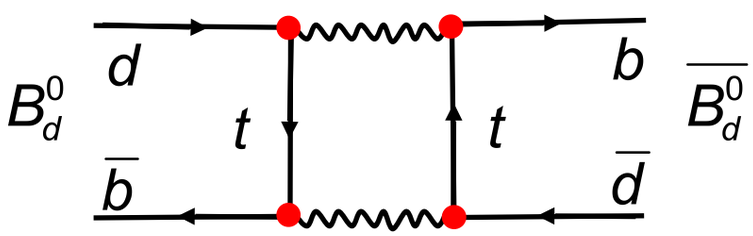
	Δm $(x = \Delta m/\Gamma)$	$\Delta\Gamma$ $(y = \Delta\Gamma/(2\Gamma))$
K^0	large	$\sim$ maximal
	~ 500	~ 1
D^0	small	small
	$(0.63 \pm 0.19)\%$	$(0.75 \pm 0.12)\%$
B^0	medium	small
	0.770 ± 0.008	0.008 ± 0.009
B_s^0	large	medium
	26.49 ± 0.29	0.075 ± 0.010

- B^0mixing, first observed by Argus in 1987, then measured precise by B factories, LHCb…
- B_s^0 mixing first measured by CDF in 2006 and then by LHCb

$B^0 \leftrightarrow \overline{B}^0$ oscillations

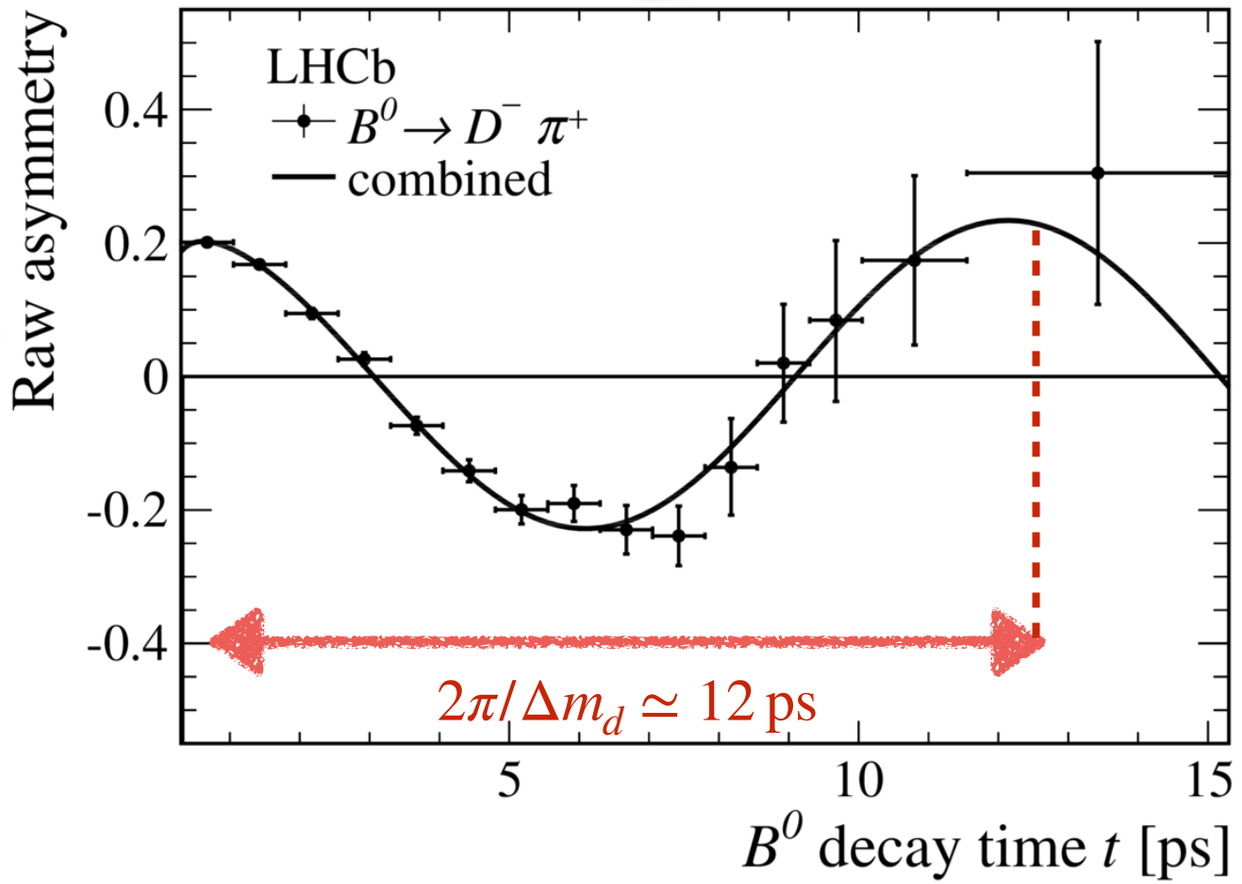


$$\Delta m_d = 0.5065 \pm 0.0019 \text{ ps}^{-1}$$



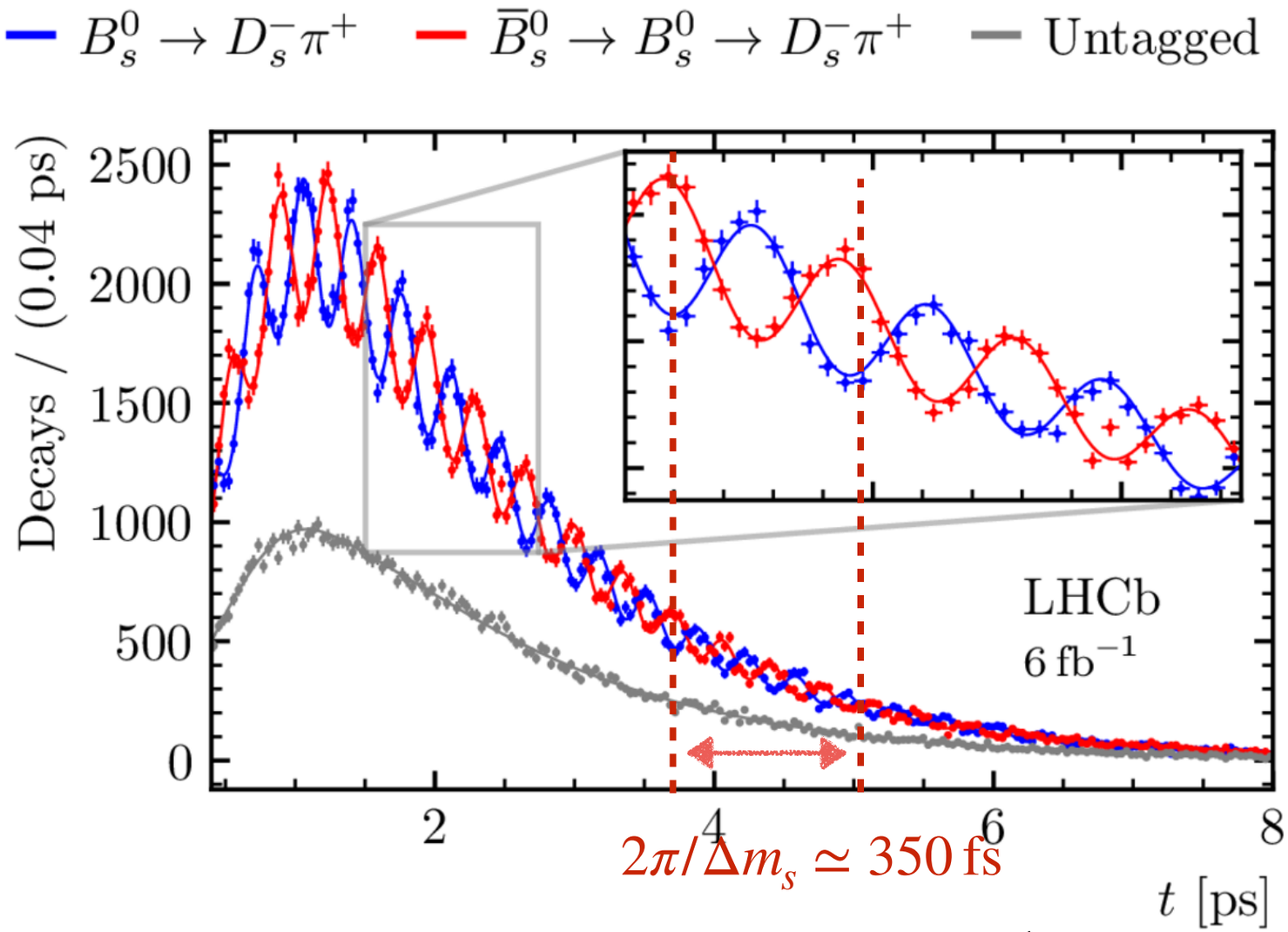
$$\Delta m_d \sim m_t^2 \, |V_{tb} V_{td}|^2 \sim m_t^2 \cdot \mathcal{O}(\lambda^2)$$

Phys. Lett. B719 (2013) 318



One period of B^0 oscillations $\Delta T \simeq 12\,\text{ps} \rightarrow$ oscillation frequency $\Delta m_d \simeq 0.5\,\text{ps}^{-1}$

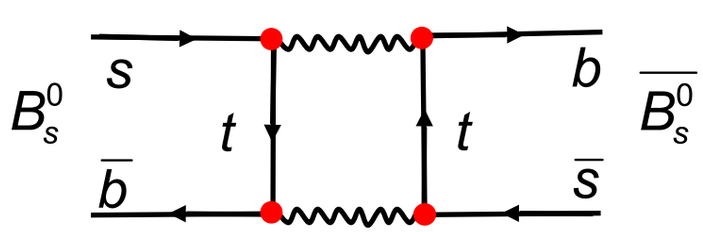
$B_s^0 \leftrightarrow \overline{B}_s^0$ oscillations



- Different flavour at decay and production
- Same flavour at decay and production

Nature Physics 18 (2022)

- $\Delta m_s = 17.741 \pm 0.0057\ \text{ps}^{-1}$ 0.03% accuracy



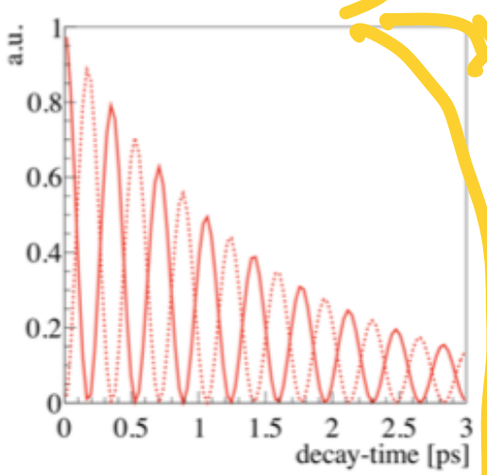
$$\Delta m_s \sim m_t^2 |V_{tb}V_{ts}|^2 \sim m_t^2 \cdot \mathcal{O}(\lambda^4)$$

One period of B_s^0 oscillations $\Delta T \simeq 350\,\text{fs} \rightarrow$ oscillation frequency $\Delta m_s \simeq 17.8\,\text{ps}^{-1}$ (~35 times faster than B_d^0 oscillations)

$B_s^0 \leftrightarrow \overline{B}_s^0$ oscillations, experimentally

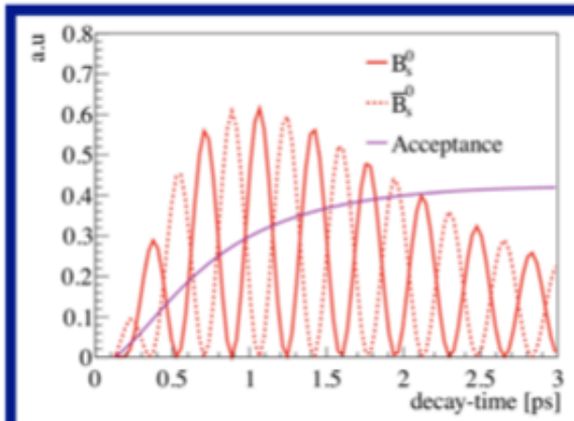
Detector effects

Perfect



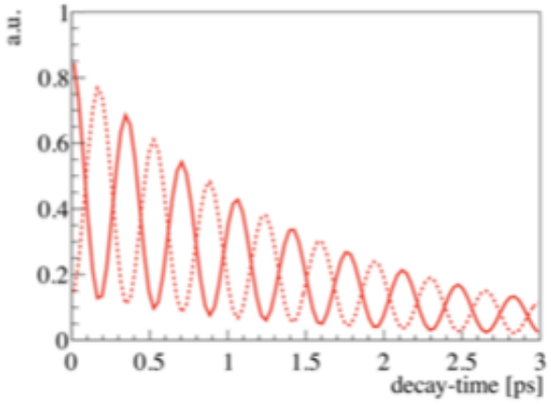
$$P(t) \sim e^{-\Gamma_s t}(\cosh(\frac{\Delta\Gamma_s t}{2}) \pm \cos(\Delta m_s t))$$

Decay-time acceptance



- decay-time distorted by selection requirements
- selection efficiency depends on decay-time introducing acceptance

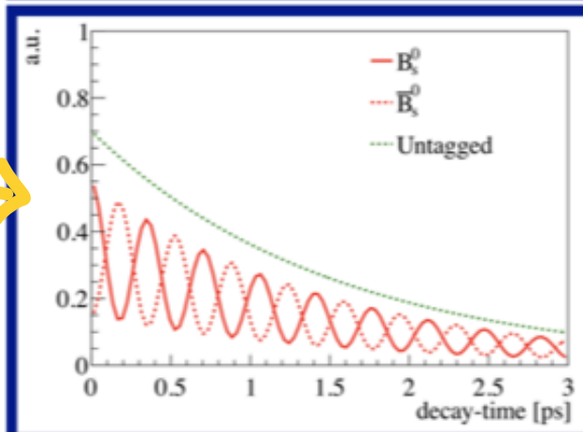
Decay-time resolution



- finite decay-time resolution of the detector leads to a dilution of observed oscillation

Dilution: $D_{res} \sim e^{-\frac{1}{2}\Delta m_s^2 \sigma^2}$

Flavour tagging

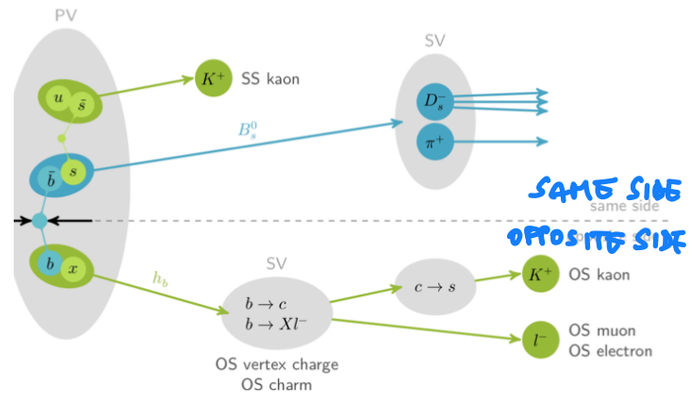


- decay-rates depend on the initial flavour, which is provided by dedicated algorithms

Dilution:

$$D \sim (1-2\omega) = 1 - 2\frac{\#wrong\ tag}{\#all\ tag}$$

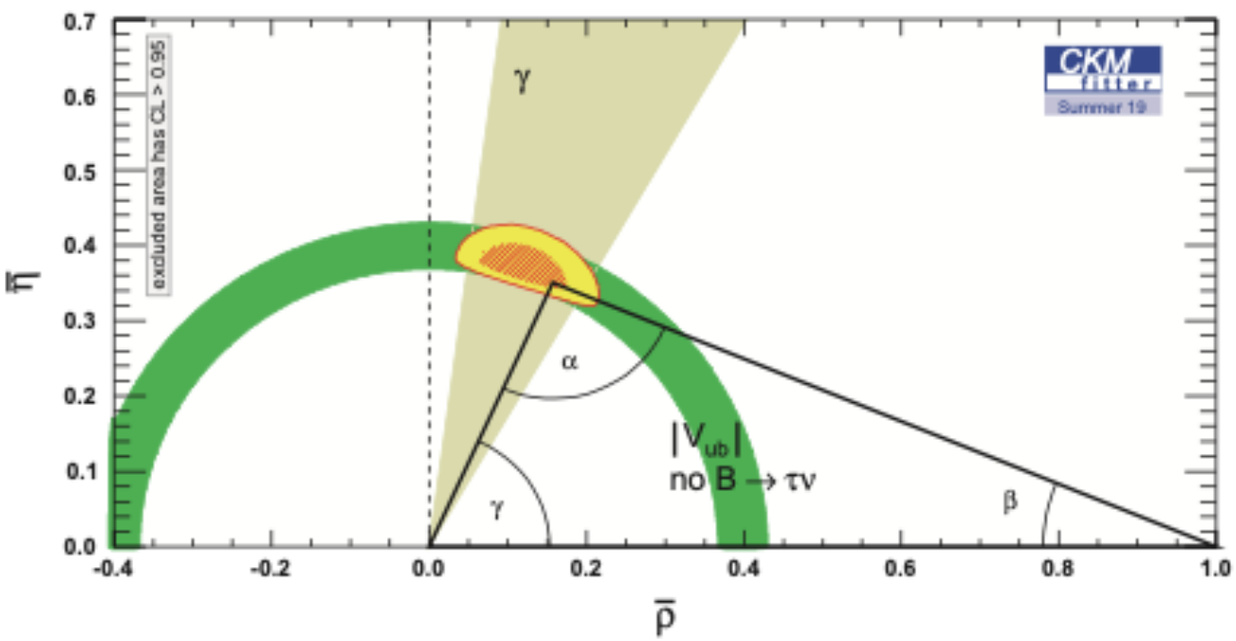
Requires excellent time resolution:
$\sigma_t \approx 45$ fs (LHCb)



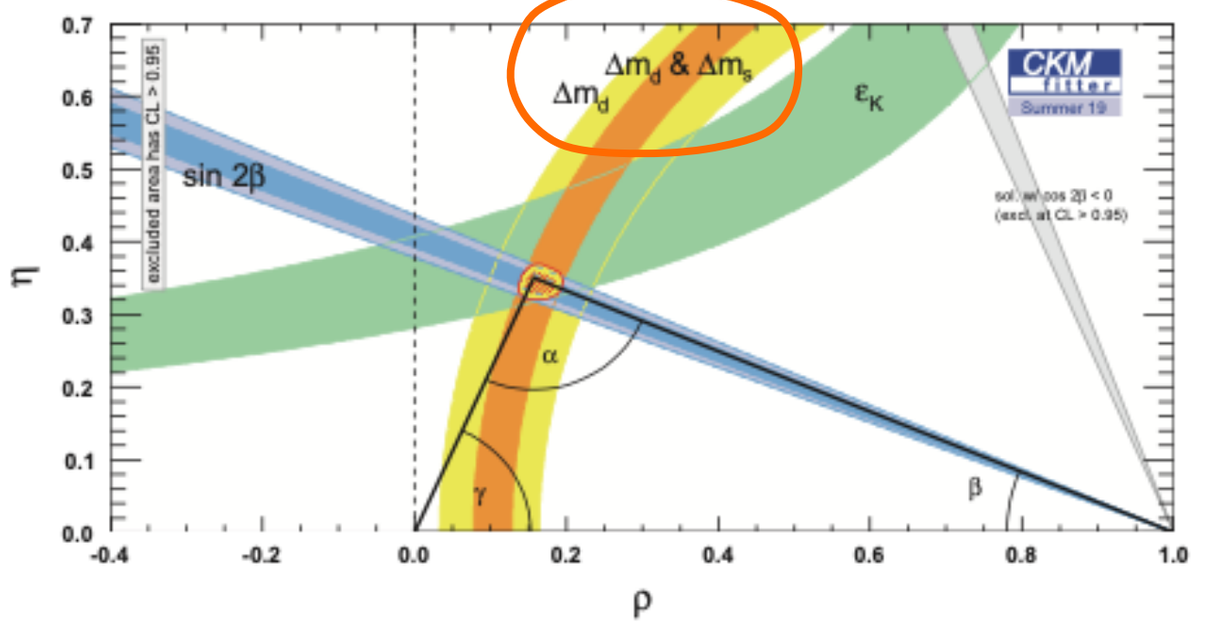
Tagging power
$\epsilon_{eff} = \epsilon(1-2\omega)^2 \approx 6\,\%$ (LHCb)

Impact of B-meson mixing measurements

Tree-level observables



Loop observables

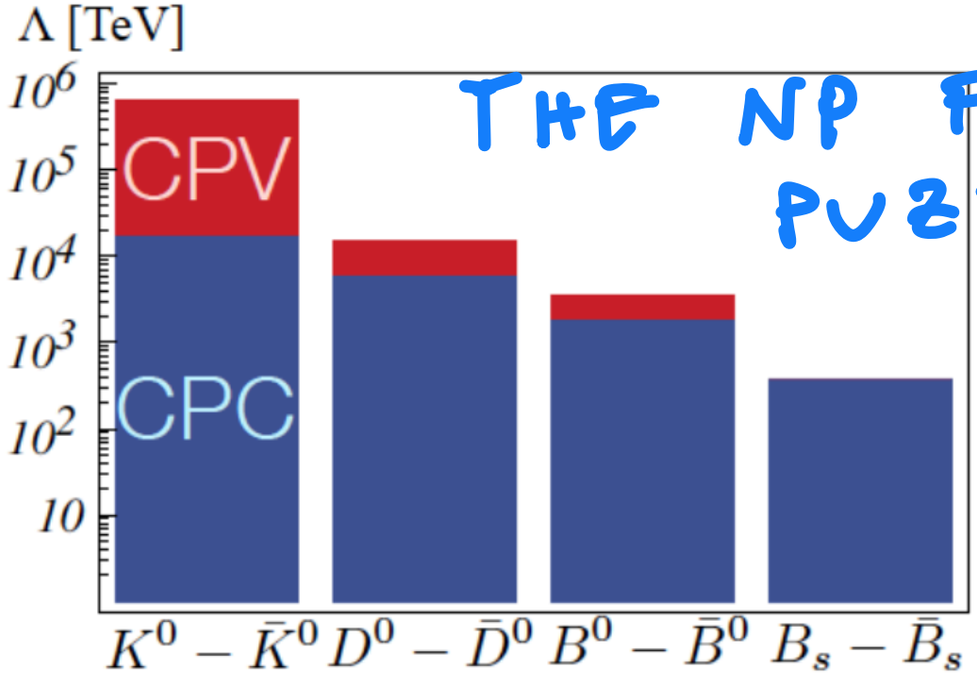


Stringent bounds on the scale of NP from meson-anti meson mixing (assuming generic NP effects in loop-mediated amplitudes).

Can we make sense of the tight NP bounds from flavour-violating processes and still hope to see NP signals somewhere?

Gino Isidori

The ratio $\Delta m_d / \Delta m_s$ in the SM benefits from the cancellation of many uncertainties



THE NP FLAVOUR PUZZLE

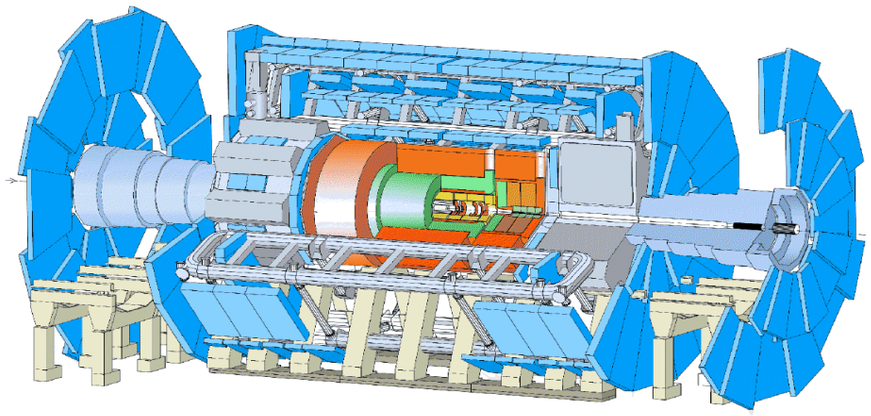
A large experimental effort…

- Constraints coming from K mesons from. e.g., NA48 at CERN, KLOE at LNF, KTeV at FNAL
- Measurements of CKM parameters from D and B mesons pioneered by ARGUS at DESY, CLEO, and CLEO-c at CESR, Cornell, followed by the so-called B-factory experiments BaBar at SLAC and Belle at KEK
- Significant contributions from CDF and D0 at FNAL, especially on B_s^0 mesons

- All the above experiments have been terminated while Belle has been upgraded (Belle II)
- LHCb at the LHC is now dominating physics with b and c hadrons while the general purpose detectors ATLAS and CMS contribute in selected areas and Belle II is starting to deliver significant results
- BESIII in China provides many results on c hadrons, NA62 at CERN and KOTO at J-Parc measure very rare Kaon decays

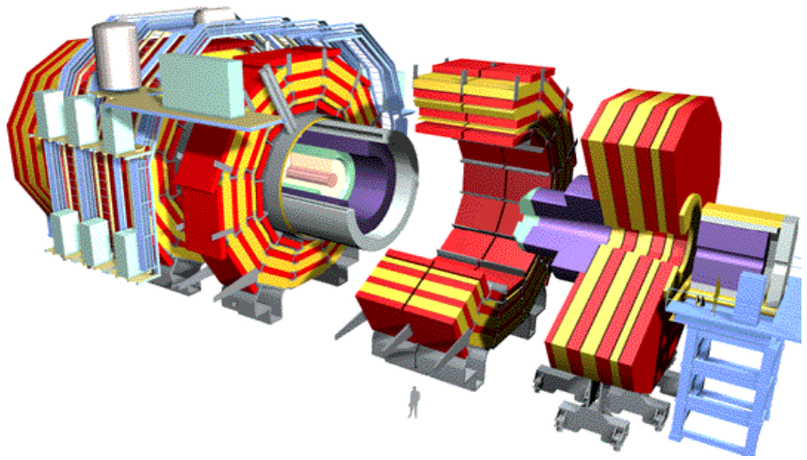
The main actors in b-physics today

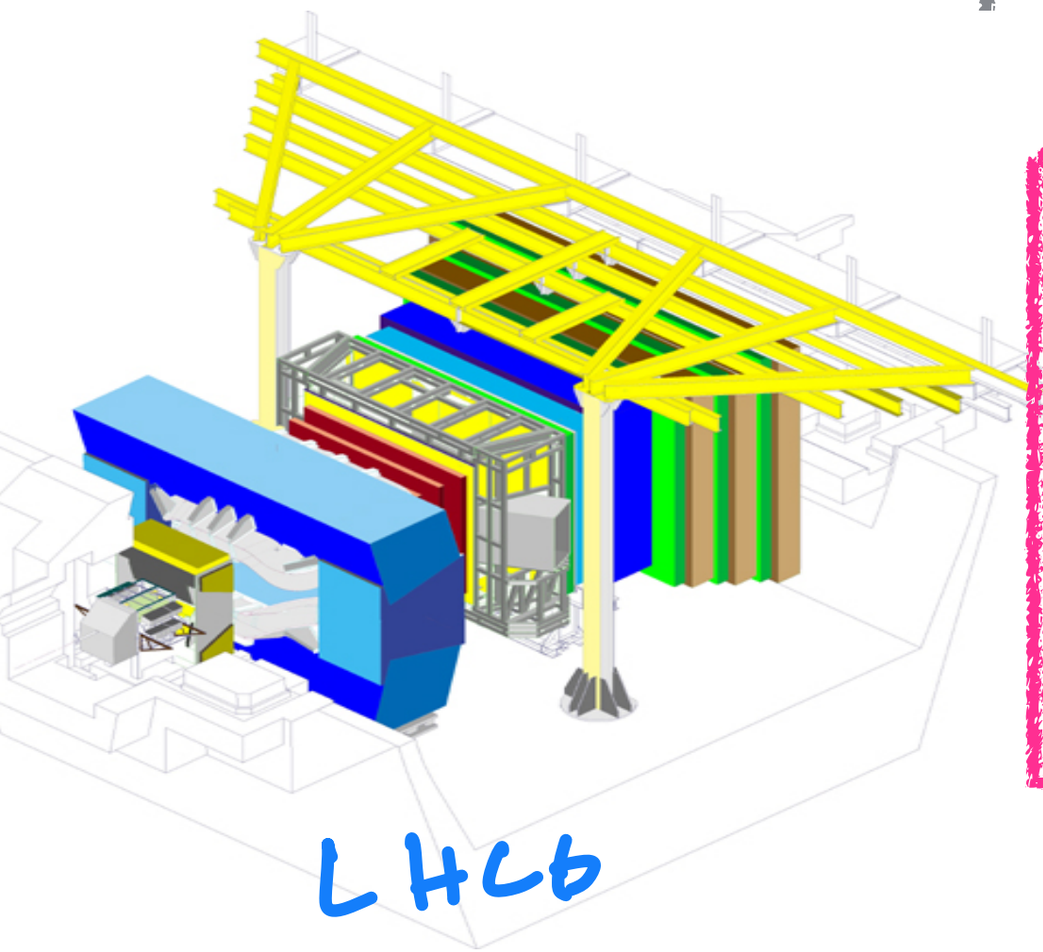
ATLAS



ATLAS and CMS @ LHC are “General Purpose Detectors”, but can measure a few flavour observables, mainly with muons in final state

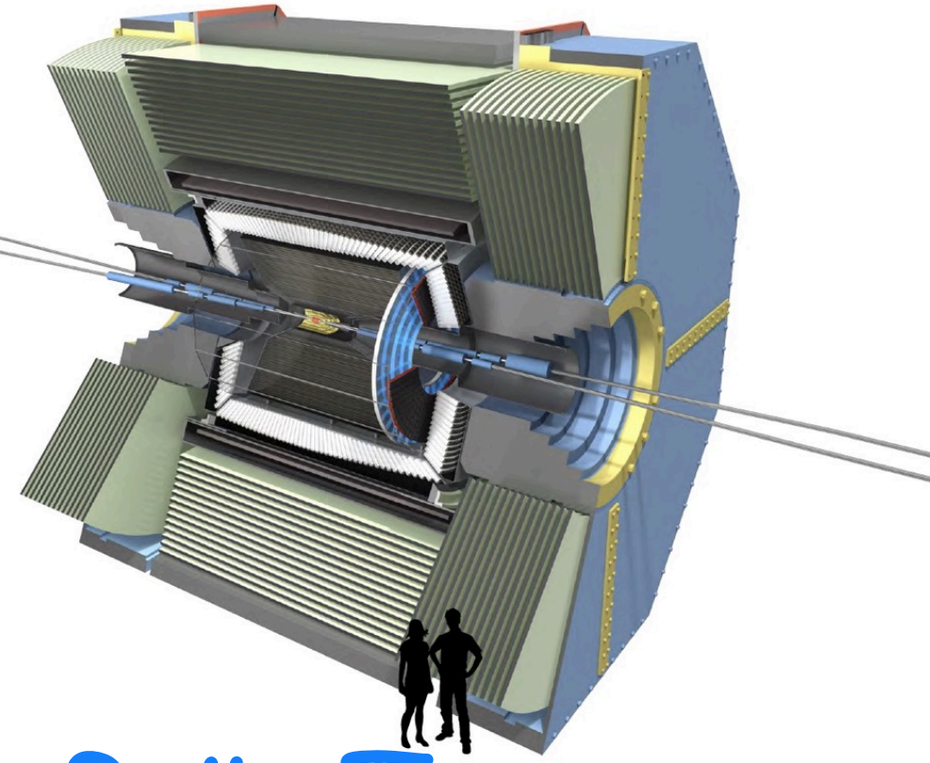
CMS





LHCb

LHCb @ LHC and Belle II @KEK are dedicated detectors for flavour physics performing a wide range of measurements



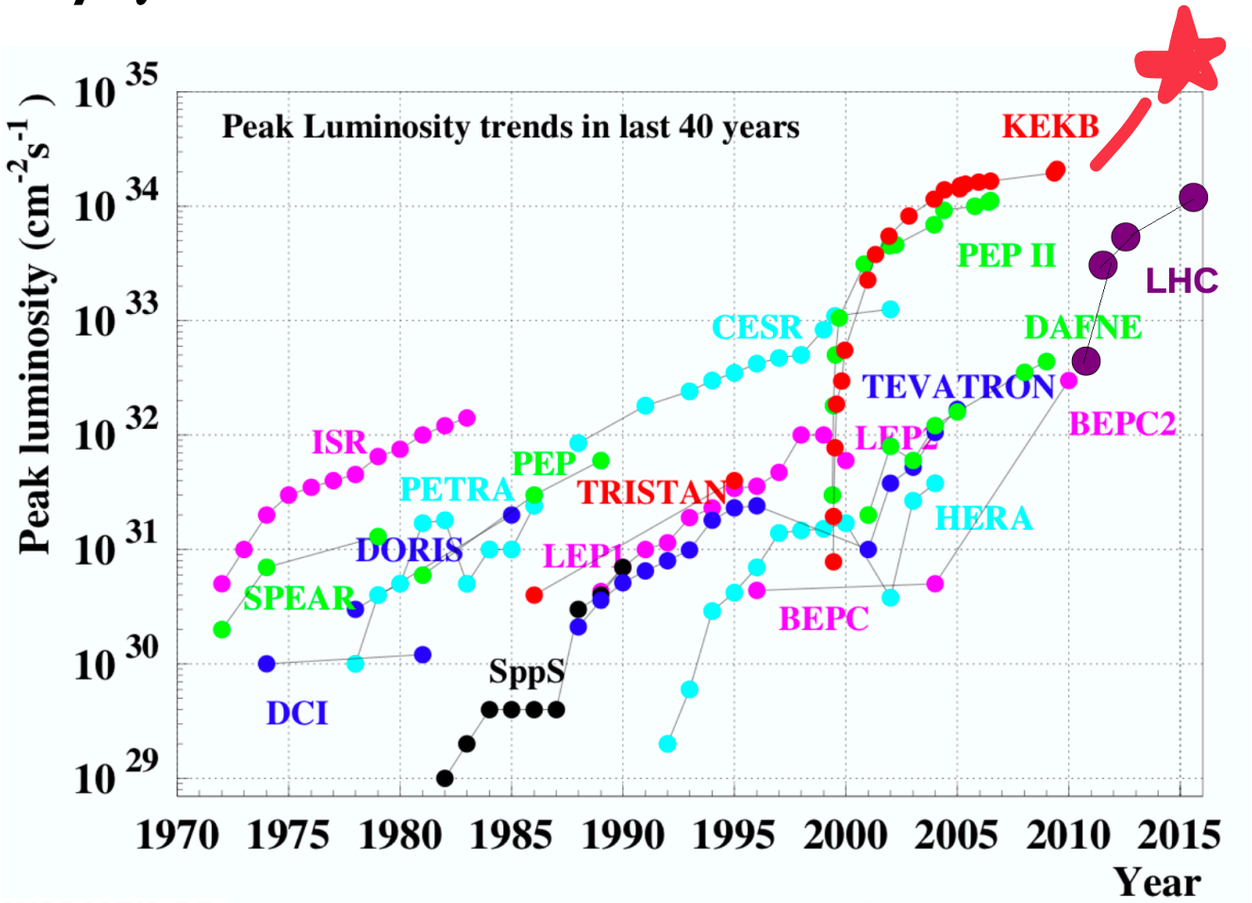
Belle II

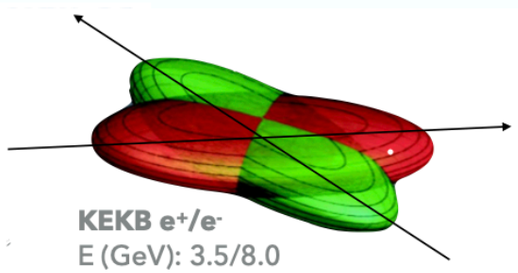
BELLE II @ SuperKEKB

- Operates at the $\Upsilon(4S)$ with energy-asymmetric e^+e^- collisions → CM boosted with $\beta\gamma \sim 0.28$

- Completed major upgrade to the accelerator to reach 30xKEKB ($\mathscr{L} = 6\,10^{35}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$)
  - 2x higher beam currents
  - 20 x smaller beam spot (σ_y=60 nm)

- Nano-beam scheme
  - idea is to have a very strong vertical focusing at the interaction point by making the crossing angle even larger than the previous machine, together with smaller beam emittances.



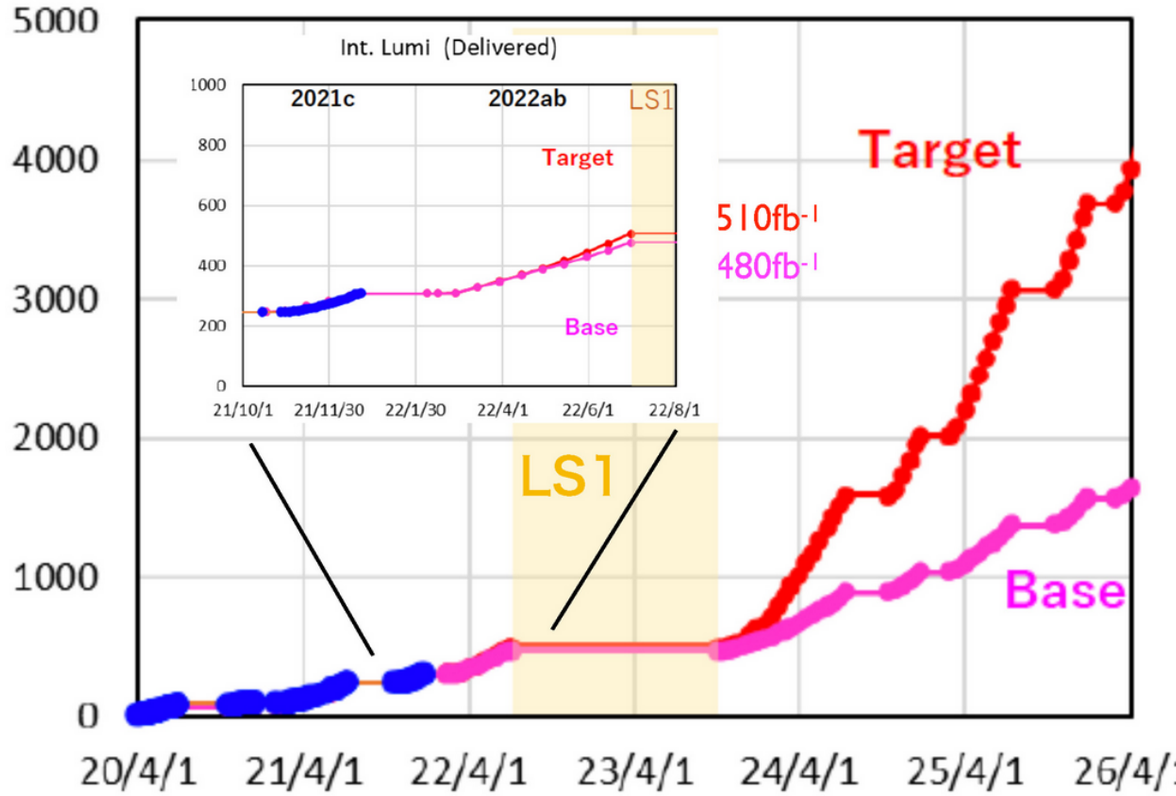


KEKB e⁺/e⁻
E (GeV): 3.5/8.0
I (A): ~ 1.6/1.2
β^*_y (mm): ~5.9/5.9
Crossing angle (mrad): 22

SuperKEKB e⁺/e⁻
E (GeV): 4.0/7.0
I (A): ~ 3.6/2.6
β^*_y (mm): ~0.27/0.3
Crossing angle (mrad): 83

BELLE II status

- SuperKEKB performance below expectations, but $\mathscr{L}_{\rm inst}$ exceeding KEKB by ~ a factor 2
- $\mathscr{L}_{\rm int}$~1/3 of full Belle dataset
- Long shutdown1 starting summer ’22
- Performance mostly exceeding that of BELLE (e.g., tracking and vertexing, neutral reconstruction, μ -id, flavour-tagging..)



- World’s best $D^{0,+}, \Lambda_c^+$ lifetimes!
- Measurements of $B(B \to K^* \ell^+ \ell^-)$, with $\ell = e, \mu$, currently limited by stat.uncertainty, with similar precision for electrons and muons; electron channel will become competitive with 1/ab
- Will provide essential independent check of anomalies with few 1/ab

BELLE II (dis)advantages

- B mesons ~1/4 of total $\sigma(e^+e^- \to \text{hadrons})$
- "Clean environment": only two B-mesons produced ($50\,\% B^+B^-, 50\,\% B^0\overline{B}^0$)
- Detector acceptance approaches 4π and is quite uniform
- Very high efficiency to learn the flavour of a neutral B-meson when studying its partner in the same $\Upsilon(4S)$ decay (flavour tagging) is ~40%
- Electrons measured almost as well as muons; also they do better than LHCb in inclusive modes and in modes with neutrals
- Coherent B-meson production
- Relatively poor time resolution of ~900 fs on decay-time difference of the two B -mesons compared to B lifetime of ≈ 1500 fs
- Cross-section ~1.1 nb (although very high $\mathscr{L}$)
- Only $B^+B^-, B^0\overline{B}^0$ can be studied with precision; B_s are produced at the $\Upsilon(5S)$, but cross-section ~0.06 nb and time resolution not good enough to study mixing and CPV
- Studies of other b-hadron species are impossible as the accelerator does not have enough energy to produce them.

ATLAS&CMS (dis)advantages

- Large $b\bar{b}$ cross-section $\sigma_{b\bar{b}} \sim 600\mu\text{b}\ @\sqrt{s} = 13\ \text{TeV}$
- All species of b-hadrons produced $(B_\text{u}, B_\text{d}, B_\text{s}, B_\text{c}, \Lambda_\text{b}, \Sigma_\text{b}, \Xi_\text{b} \dots)$
  $\mathscr{L}_\text{peak} \sim 2 \cdot 10^{34}\text{cm}^{-2}\text{s}^{-1}$ and $\int \mathscr{L}\,\text{d}t \approx 160\,\text{fb}^{-1}\ @\sqrt{\text{s}} = 13\ \text{TeV}$
- $\sigma_{b\bar{b}}/\sigma_\text{inel} \sim \text{few}\ 10^{-3}$, and most interesting b-hadron decays occur at 10^{-5} probability or lower → trigger is a major issue
- Large boost (decay vertices well separated)
- Excellent tracking, muon and electron ID (but muons triggered and reconstructed more efficiently) but no ability of distinguishing pions, kaons and protons
- Limitations to readout bandwidth → b-hadron decays with low p_T cannot be selected and readout
- Many particles in event not associated with the two b-hadrons
- Large pileup (up to 40 in 2018)

LHCb (dis)advantages

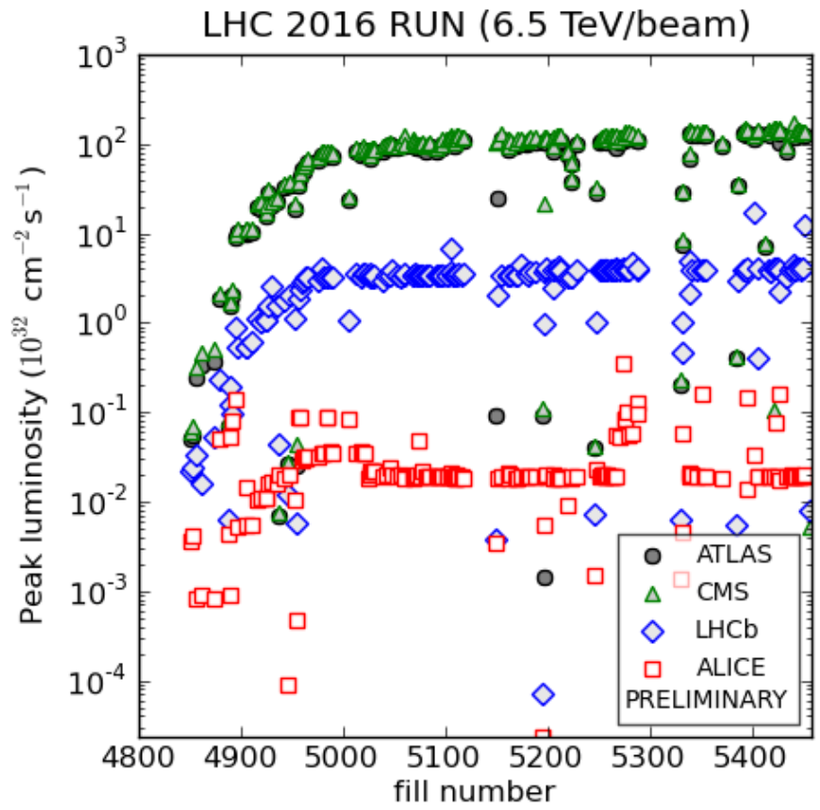
- Large $b\bar{b}$ cross-section $\sigma_{b\bar{b}} \sim 600\mu\text{b}\ @\sqrt{s} = 13\ \text{TeV}$
- All species of b-hadrons produced ($B_\text{u}, B_\text{d}, B_\text{s}, B_\text{c}, \Lambda_\text{b}, \Sigma_\text{b}, \Xi_\text{b} \ldots$)
- $\mathscr{L}_\text{peak} \sim 4 \times 10^{32}\text{cm}^{-2}\text{s}^{-1}$ and $\int \mathscr{L}\, \text{d}t \approx 9\ \text{fb}^{-1}$

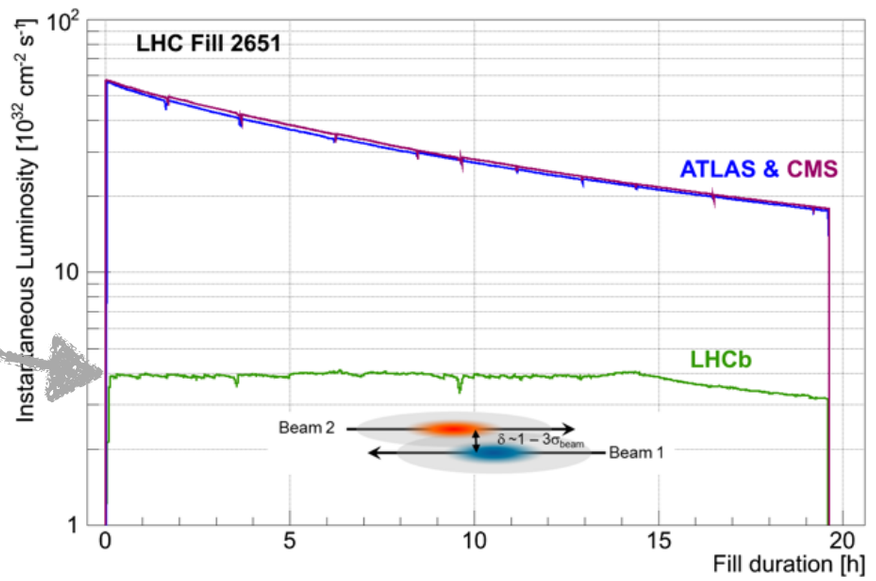
  (to be raised to 2x10^{33} cm^{-2} s^{-1} in Run 3)

- $\sigma_{b\bar{b}}/\sigma_\text{inel} \sim \text{few}\ 10^{-3}$, and most interesting b-hadron decays occur at 10^{-5} probability or lower → trigger is a major issue
- Large boost (decay vertices well separated with resolution of $\approx 45\ \text{fs}$)
- Excellent tracking, muon and electron ID (but muons triggered and reconstructed more efficiently) and PID with ability of distinguishing pions, kaons and protons
- Ability to trigger on b-hadron decays with low p_T
- Many particles in event not associated with the two b-hadrons
- ~1 visible interaction/bunch crossing (to be raised to ~5 in Run 3)

LHCb running conditions

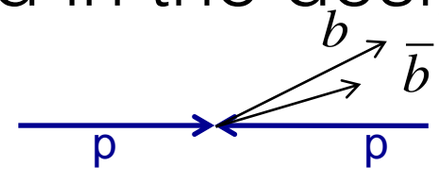
- LHCb designed to run at lower $\mathscr{L}$ than ATLAS/CMS
  - Mean number of interactions/bunch crossing ~1 in runs 1&2
  - Tracking, Particle Identification sensitive to pileup
- pp beams displaced to reduce instantaneous $\mathscr{L}$
- Experiment designed to run at constant luminosity throughout fills
  - $\mathscr{L}$ ~ 4.0 10^{32} $cm^{-2}s^{-1}$
  - For ATLAS/CMS $\mathscr{L}$ ~ 10^{34} $cm^{-2}s^{-1}$
- Huge heavy quark production cross-sections !
  - σ_{bb} ~ 600 μb @ $\sqrt{s}$=13 TeV (~1nb in e+e- @Y(4s))
    - ~10^{11} b decays/fb in acceptance
  - σ_{cc} is ~ 20 times larger!
    - ~10^{12} c decays/fb in acceptance



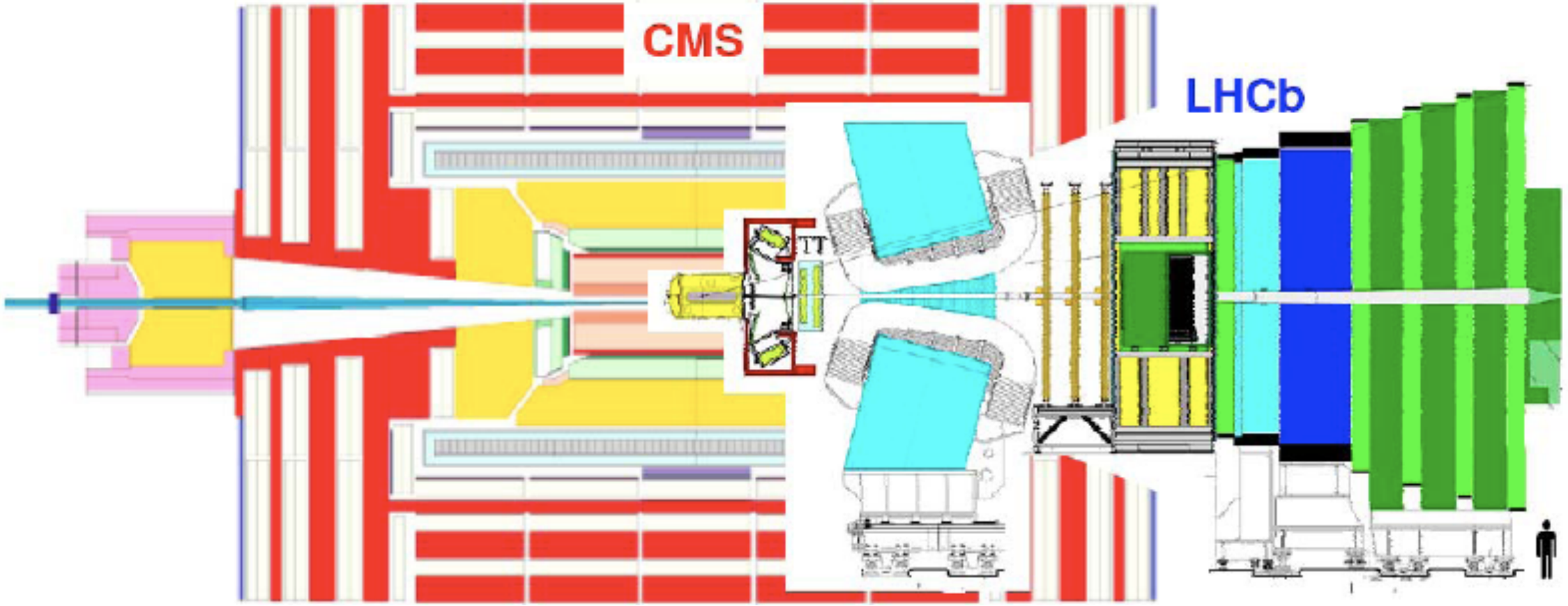


A forward spectrometer

- Why does LHCb look so different?
- The B mesons formed by the colliding proton beams (and the particles they decay into) stay close to the line of the beam pipe, and this is reflected in the design of the detector

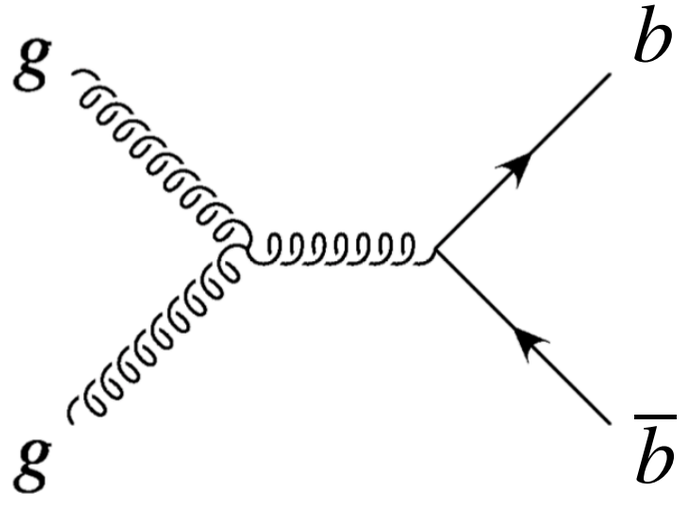


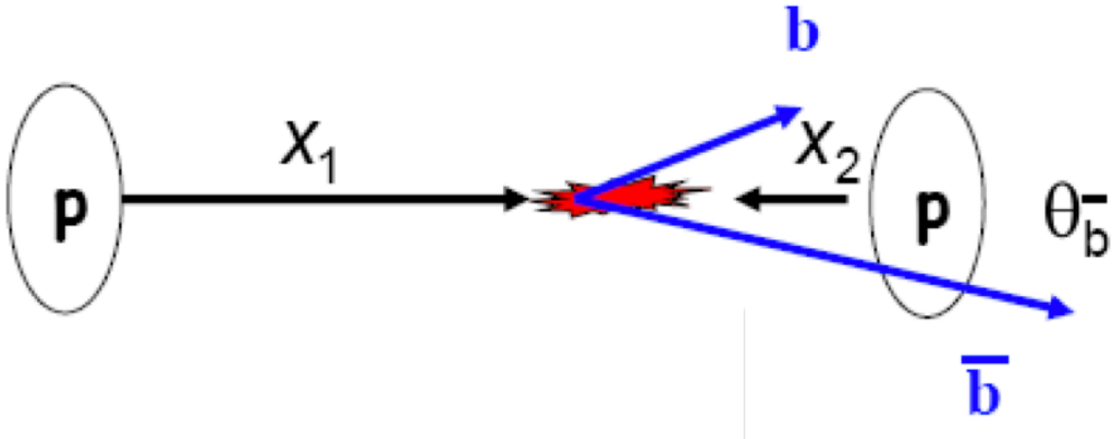
~1/4 $b\bar{b}$ events in acceptance

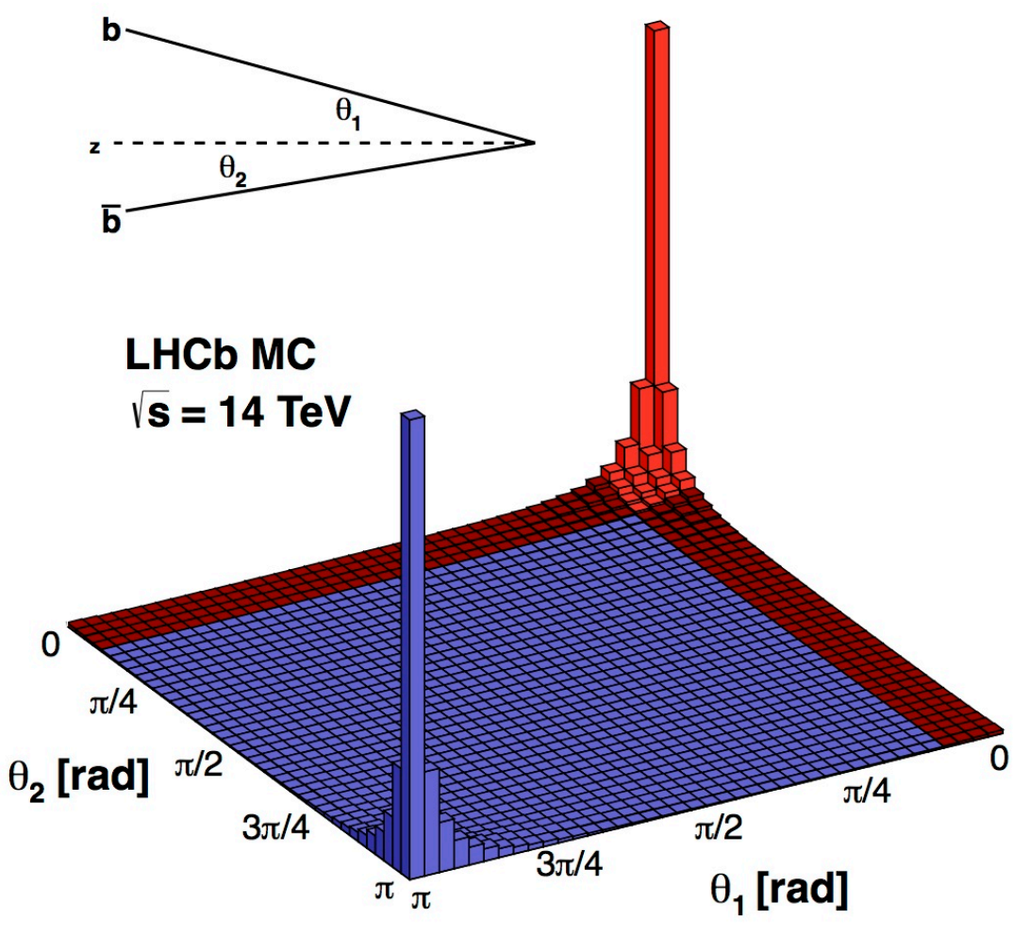


A forward spectrometer

- Dominant $b\bar{b}$ production mechanism at the LHC is through gluon-gluon fusion in which the momenta of the incoming partons are strongly asymmetric in the laboratory frame→ centre of mass energy of $b\bar{b}$ pair boosted along direction of the higher momentum gluon, and both b hadrons are produced in the same forward (or backward) direction.



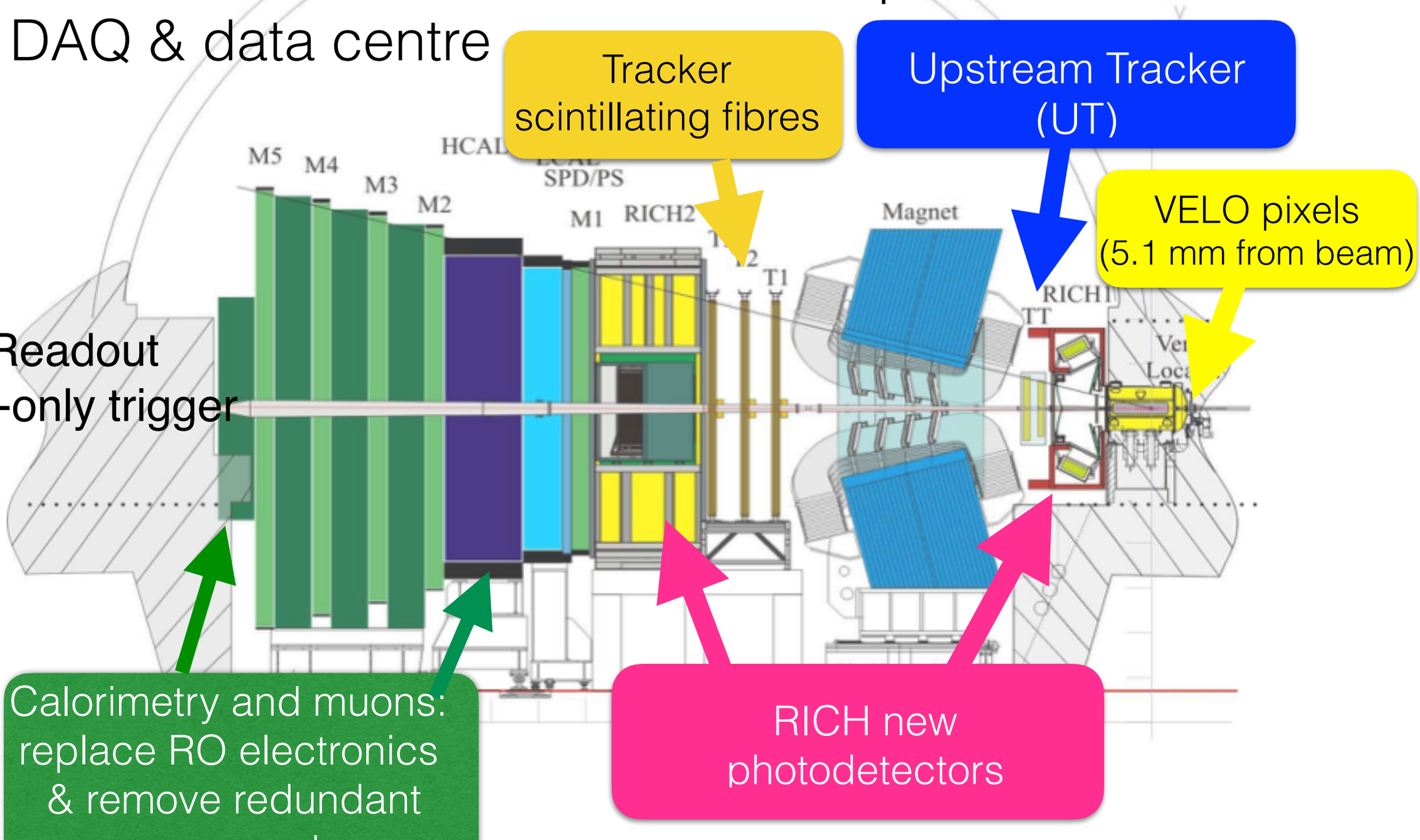




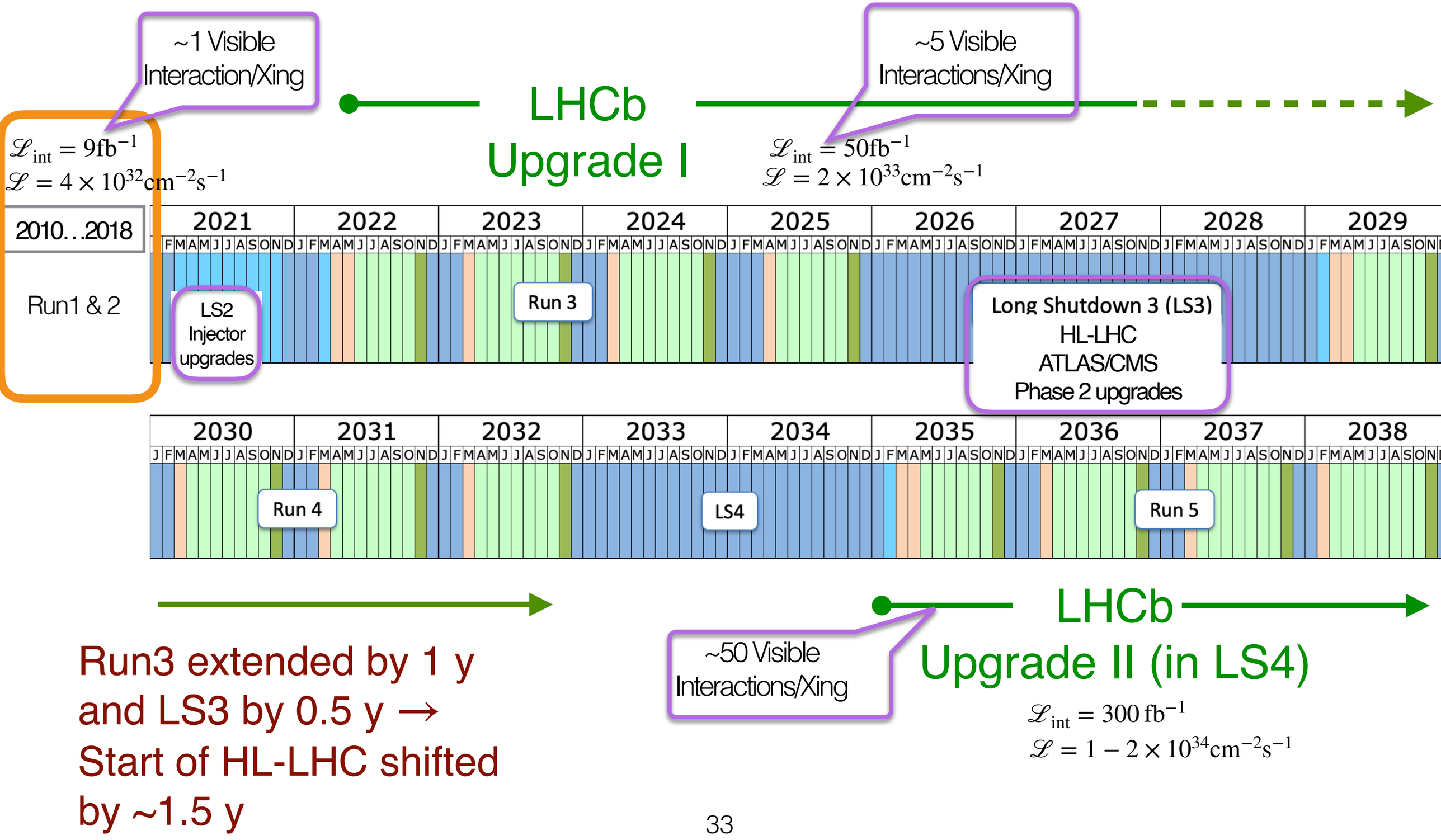
The LHCb upgraded detector

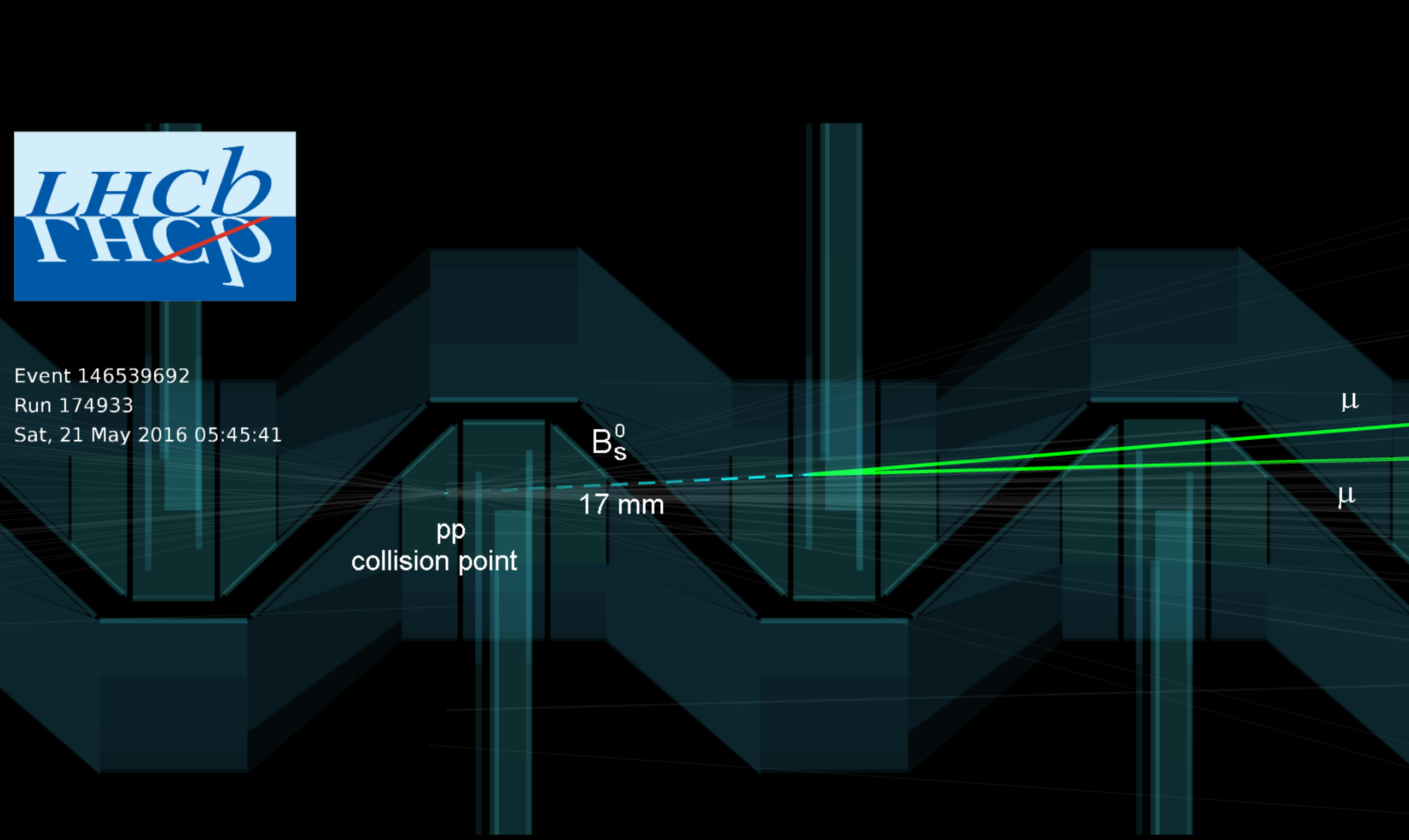
- Major upgrade of all sub detectors to handle increased rates
- Less than 10% of all channels will be kept!
- NEW DAQ & data centre

40 MHz Readout
Software-only trigger



The LHC schedule

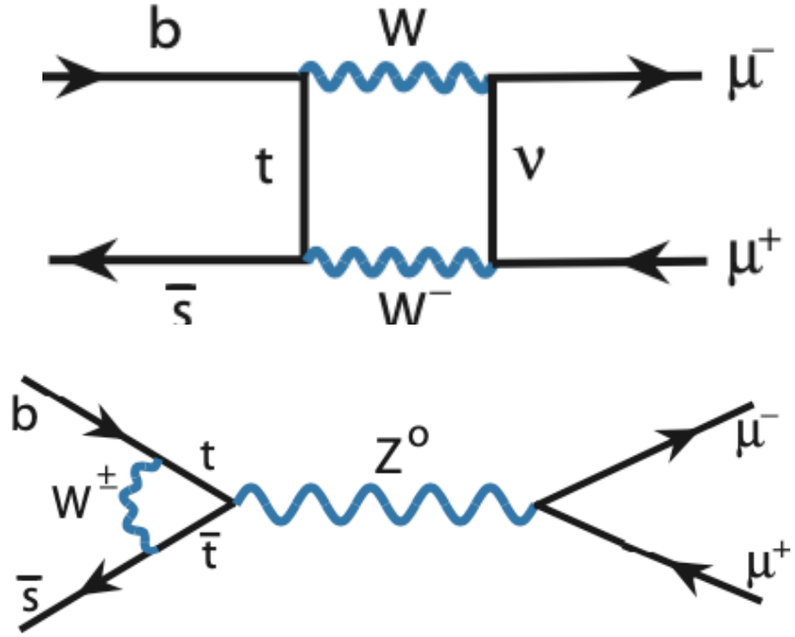




leptonic B decays

One of the milestones of flavour programme $B_{(s)} \to \mu^+\mu^-$

- Very suppressed in the SM
  - Loop, CKM ($|V_{ts}|^2$ for B_s) and helicity ~ $\left(\frac{m_\mu}{M_B}\right)^2$
  - Theoretically "clean" → precisely predicted:

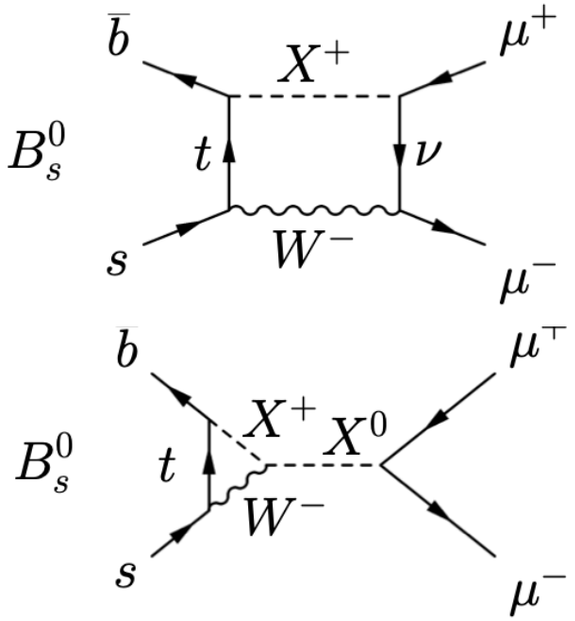


$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.66 \pm 0.14) \times 10^{-9}$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.03 \pm 0.05) \times 10^{-10}$$
$$\frac{\mathcal{B}(B^0 \to \mu^+\mu^-)}{\mathcal{B}(B_s^0 \to \mu^+\mu^-)} = (2.81 \pm 0.16)\% \ .$$

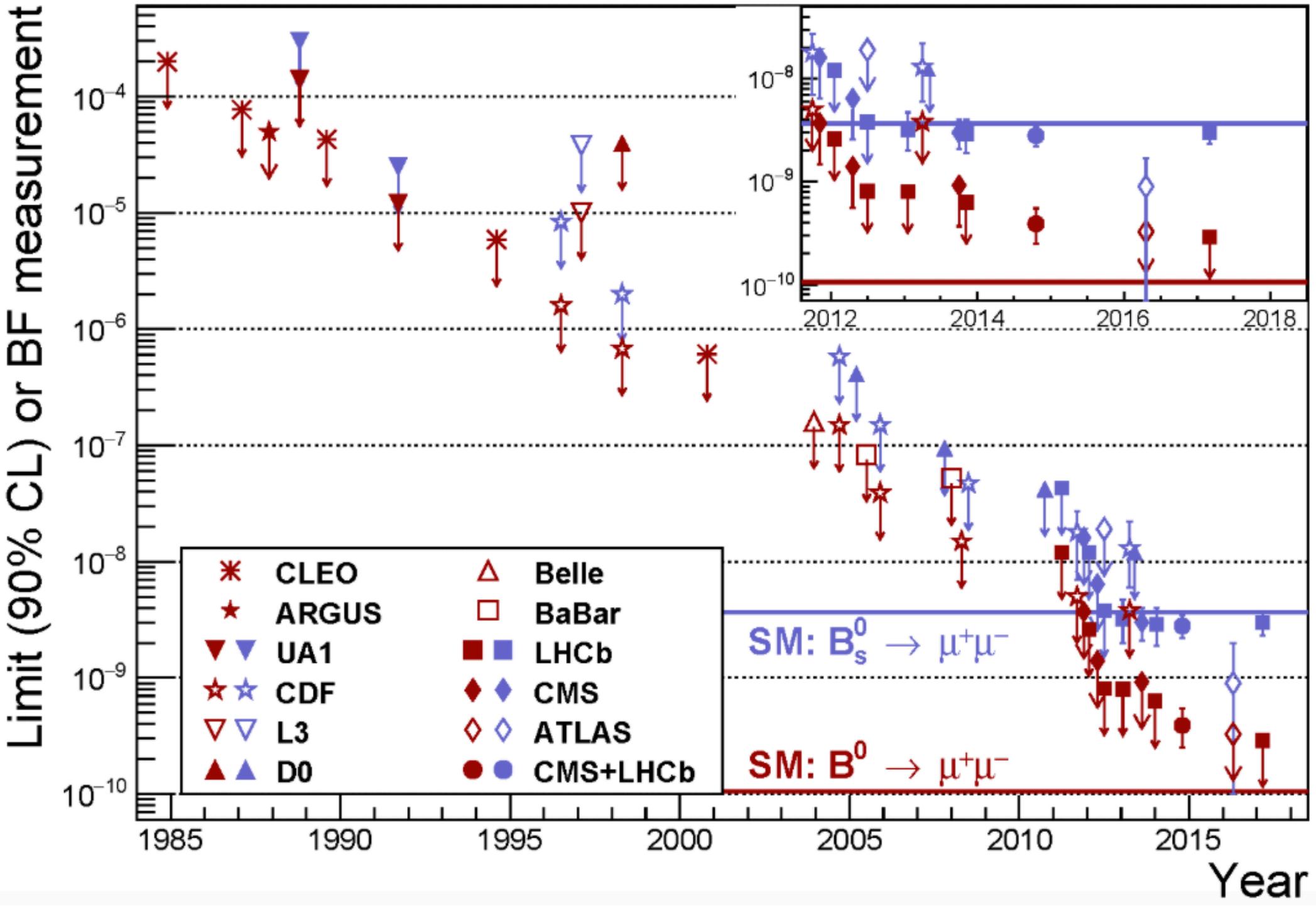
(~5%)

Bobeth et al.
PRL 112 (2014) 101801,
Beneke et al.
JHEP 10 (2019) 232

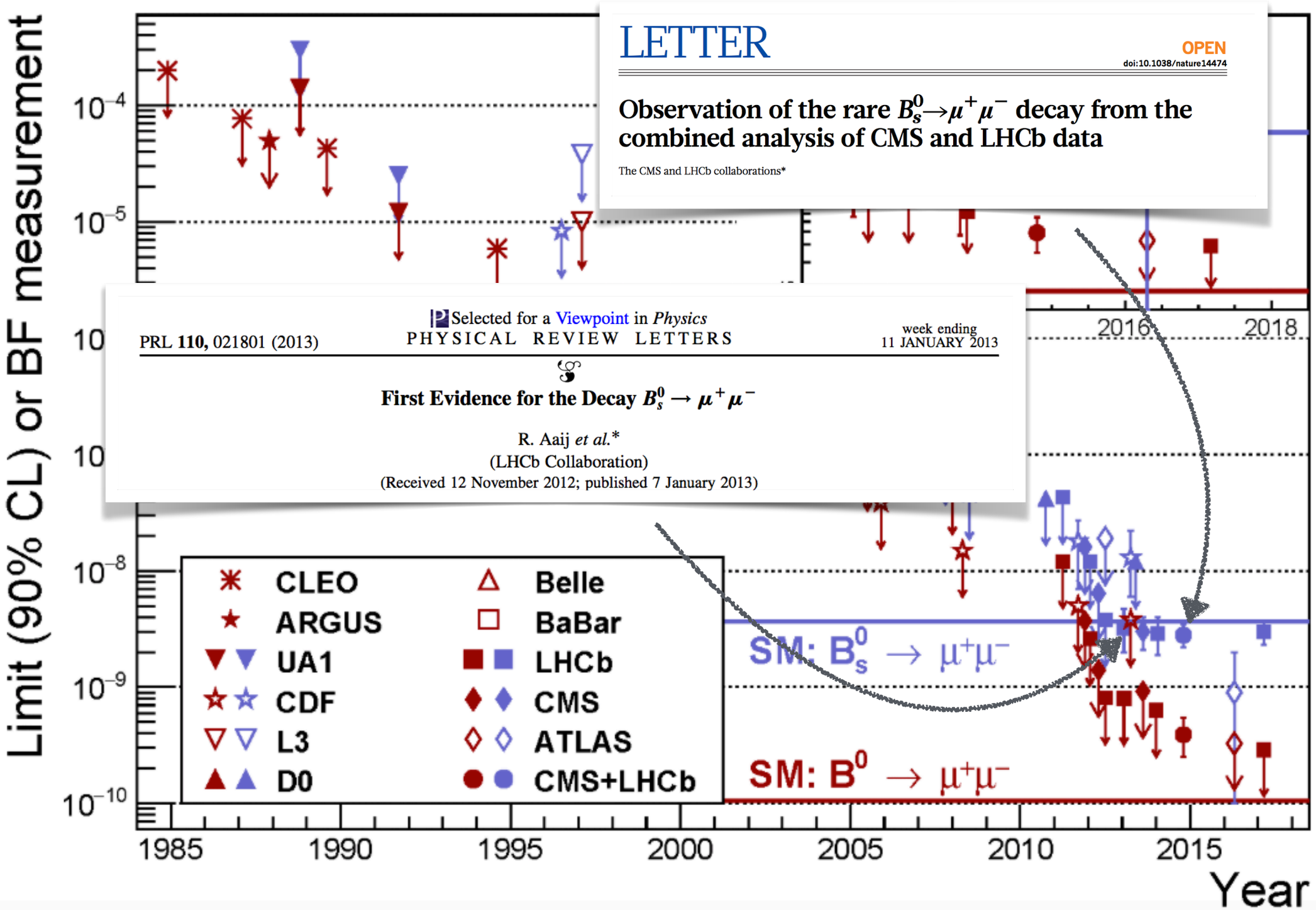
- Sensitive to NP
  - A large class of NP theories, such as SUSY, predict significantly higher values for the $B_{(s)}$ decay probability
- Very clean experimental signature
  - Studied by all high-energy hadron collider experiments



30 years of effort!



30 years of effort!



LETTER

OPEN

doi:10.1038/nature14474

Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data

The CMS and LHCb collaborations*

Selected for a Viewpoint in *Physics*

PRL **110,** 021801 (2013) PHYSICAL REVIEW LETTERS week ending 11 JANUARY 2013

First Evidence for the Decay $B_s^0 \to \mu^+\mu^-$

R. Aaij *et al.**
(LHCb Collaboration)
(Received 12 November 2012; published 7 January 2013)

Latest LHC combination

- **LHCb**, PRL 118 (2017) 191801

$B(B_s^0 \to \mu^+\mu^-) = (3.0 \pm 0.6^{+0.3}_{-0.2}) \times 10^{-9}$

$B(B^0 \to \mu^+\mu^-) < 3.4 \times 10^{-10}$ @95 % CL

7.8σ

- **CMS**, JHEP 04 (2020) 188

$B(B_s^0 \to \mu^+\mu^-) = (2.9 \pm 0.7\,(\text{exp}) \pm 0.2\,(\text{frag})) \times 10^{-9}$

$B(B^0 \to \mu^+\mu^-) < 3.6 \times 10^{-10}$ @95 % CL

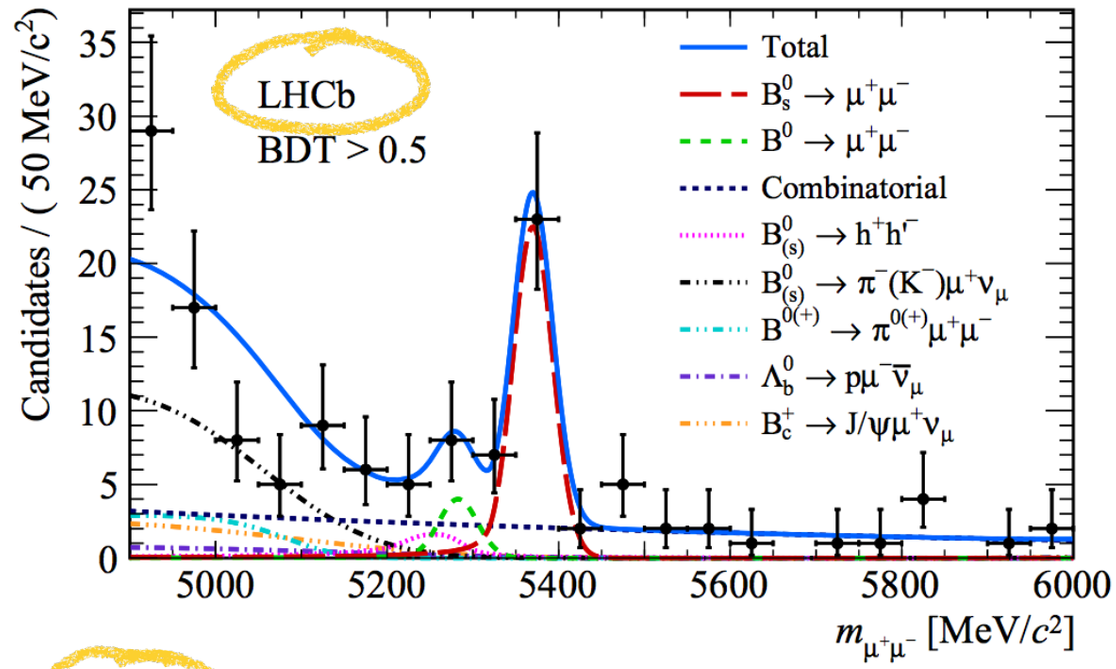
5.6σ

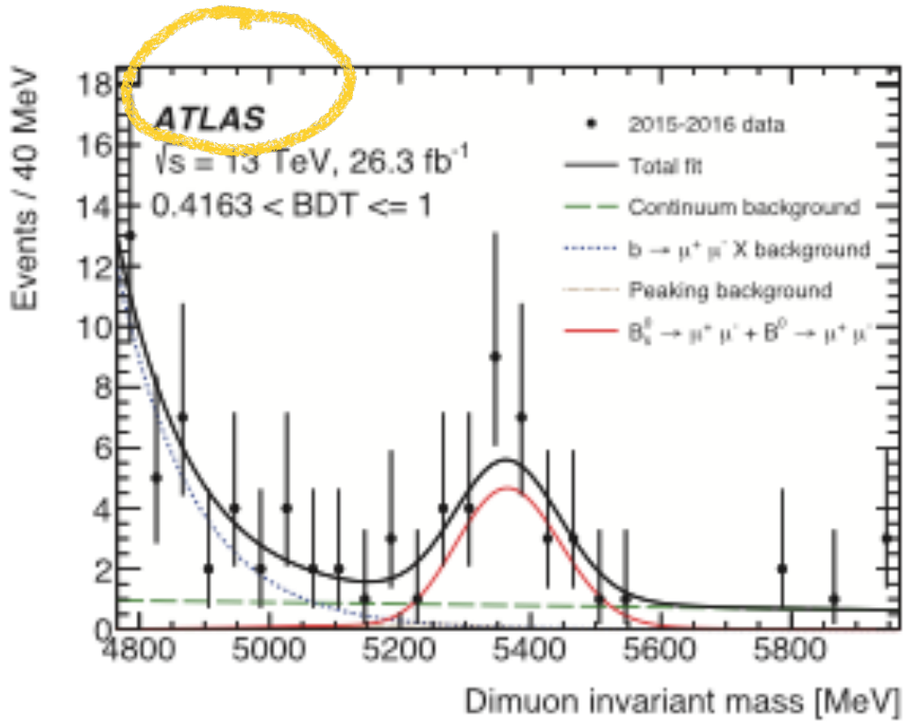
- **ATLAS**, JHEP 04 (2019) 098

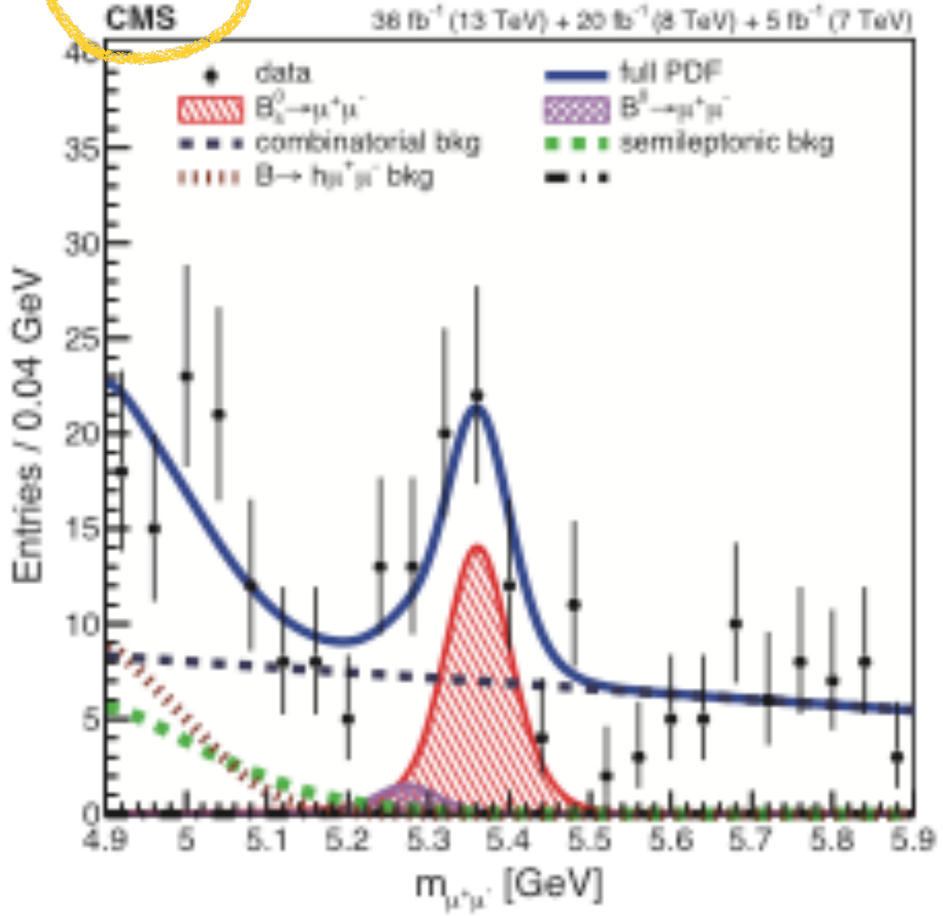
$B(B_s^0 \to \mu^+\mu^-) = (2.8^{+0.8}_{-0.7}) \times 10^{-9}$

$B(B^0 \to \mu^+\mu^-) < 2.1 \times 10^{-10}$ @95 % CL

4.6σ







Latest LHC combination

LHCb-CONF-2020-002
CMS PAS BPH-20-003
ATLAS-CONF-2020-049

- **LHCb**, PRL 118 (2017) 191801

$B(B_s^0 \to \mu^+\mu^-) = (3.0 \pm 0.6^{+0.3}_{-0.2}) \times 10^{-9}$ 7.8σ

$B(B^0 \to \mu^+\mu^-) < 3.4 \times 10^{-10}$ @95 % CL

- **CMS**, JHEP 04 (2020) 188

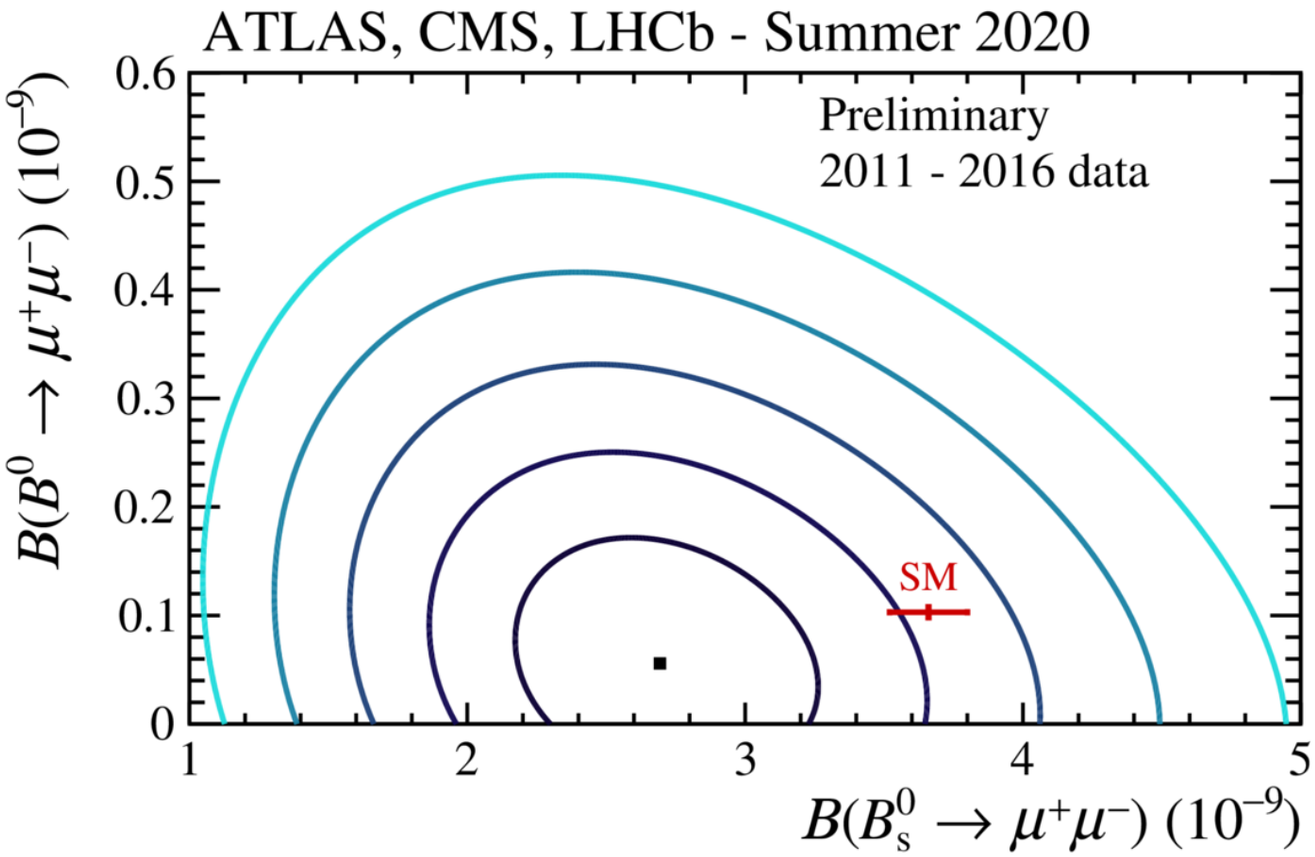
$B(B_s^0 \to \mu^+\mu^-) = (2.9 \pm 0.7\,(\text{exp}) \pm 0.2\,(\text{frag})) \times 10^{-9}$ 5.6σ

$B(B^0 \to \mu^+\mu^-) < 3.6 \times 10^{-10}$ @95 % CL

- **ATLAS**, JHEP 04 (2019) 098

$B(B_s^0 \to \mu^+\mu^-) = (2.8^{+0.8}_{-0.7}) \times 10^{-9}$ 4.6σ

$B(B^0 \to \mu^+\mu^-) < 2.1 \times 10^{-10}$ @95 % CL



Era of precision measurements of $B_{(s)} \to \mu^+\mu^-$ has started

$\mathscr{B}(B_s^0 \to \mu^+\mu^-) = (2.69^{+0.37}_{-0.35}) \times 10^{-9}$

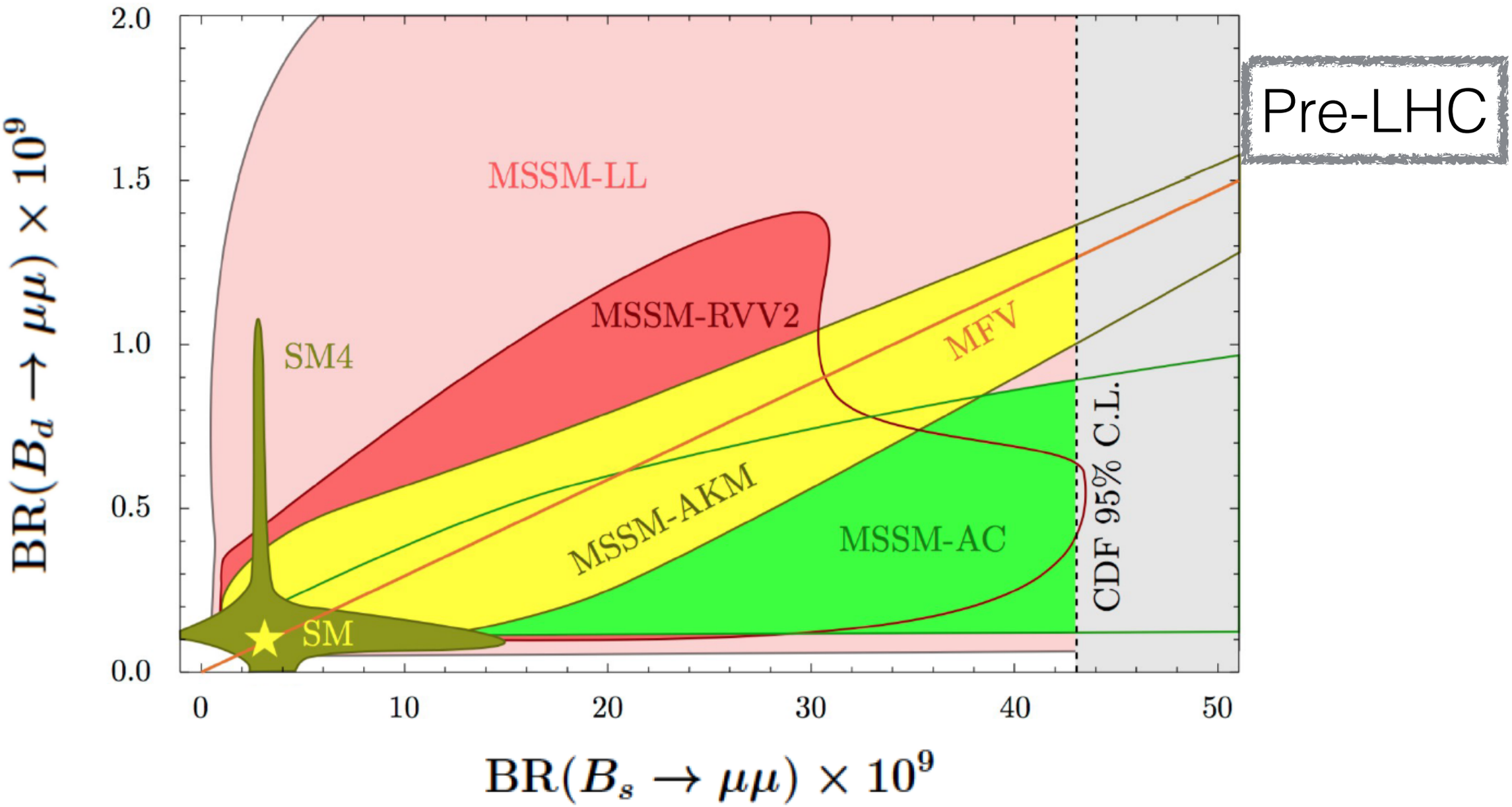
$\mathscr{B}(B^0 \to \mu^+\mu^-) < 1.9 \times 10^{-10}$ @95 % CL

2.1 σ below SM prediction (2D compatibility)

The SM stands its ground

- Sizeable effects expected in many MSSM models

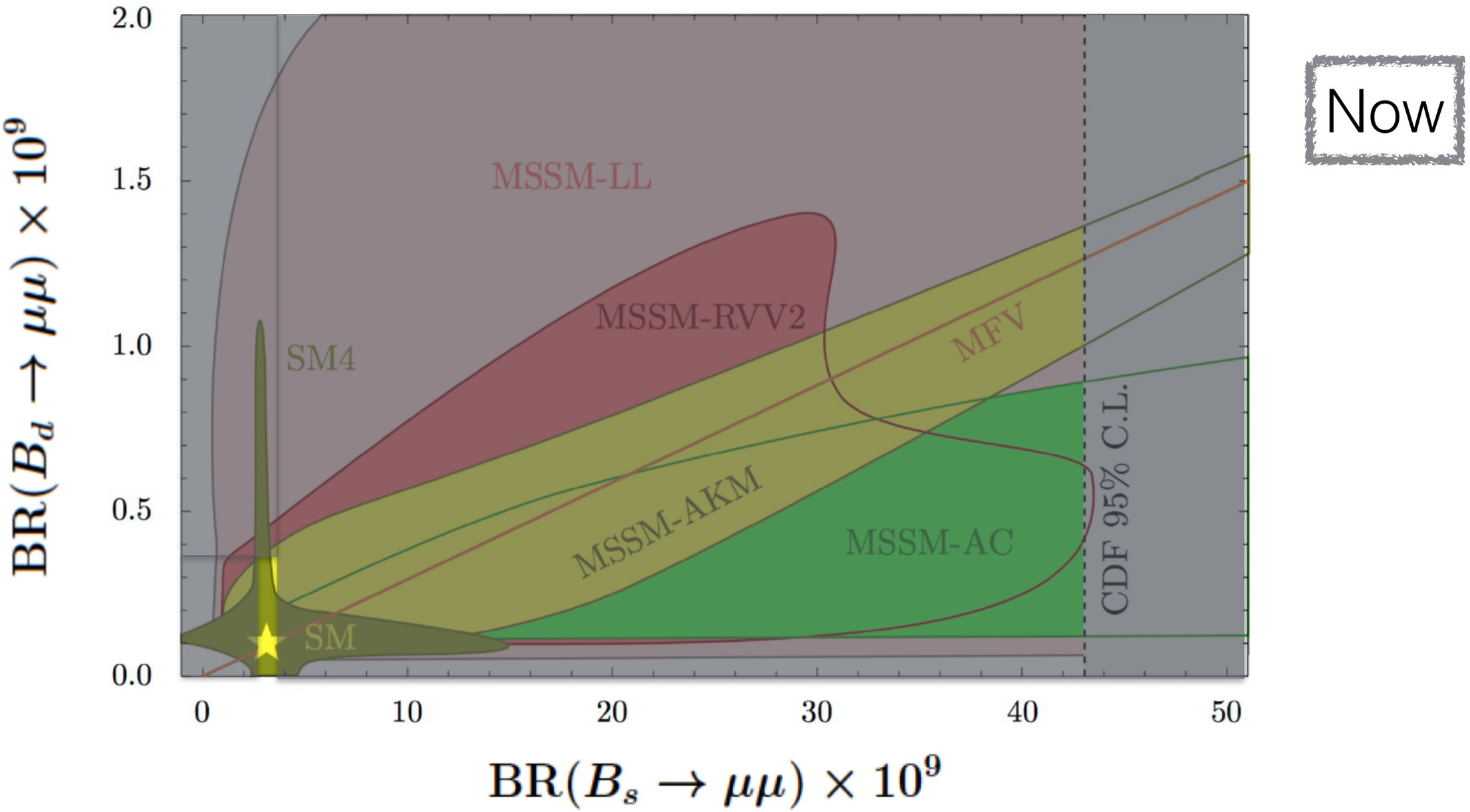
Straub, arXiv:1107.0266



The SM stands its ground

- Sizeable effects expected in many MSSM models (cancellation of helicity suppression)

Straub, arXiv:1107.0266



Now

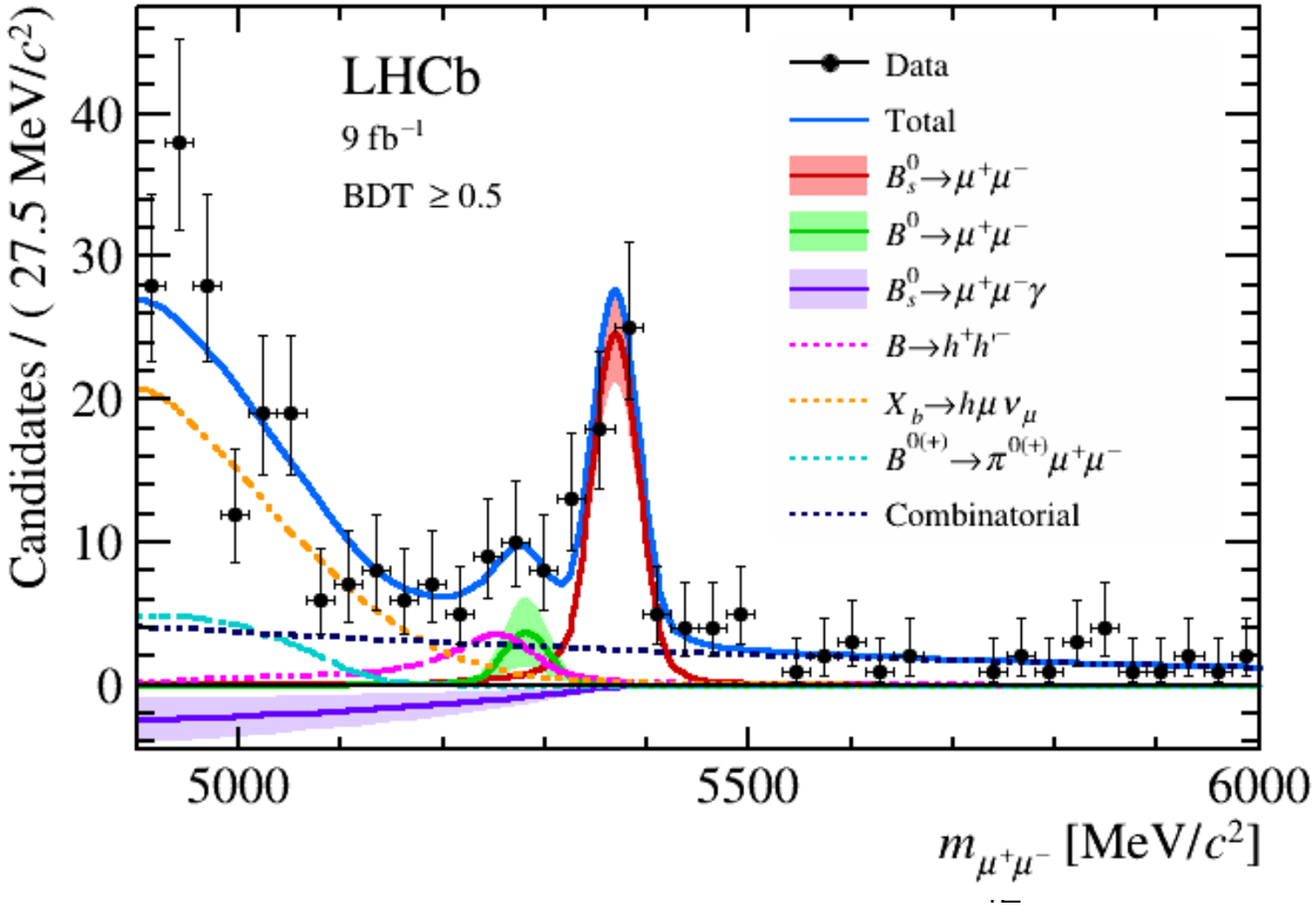
LHCb update with full dataset

- LHCb analysis based on full Run 1 and Run 2 data (9 fb^{-1})

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

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

- Consistent with SM expectation at current level of precision



- $B_s \to \mu^+\mu^-$ found with significance >10 σ, but no evidence yet for $B^0 \to \mu^+\mu^-$ (1.7σ)
- Result dominated by statistical uncertainty
- Expect 10% precision with ATLAS/CMS Run 2

PRL 128 (2022) 041801

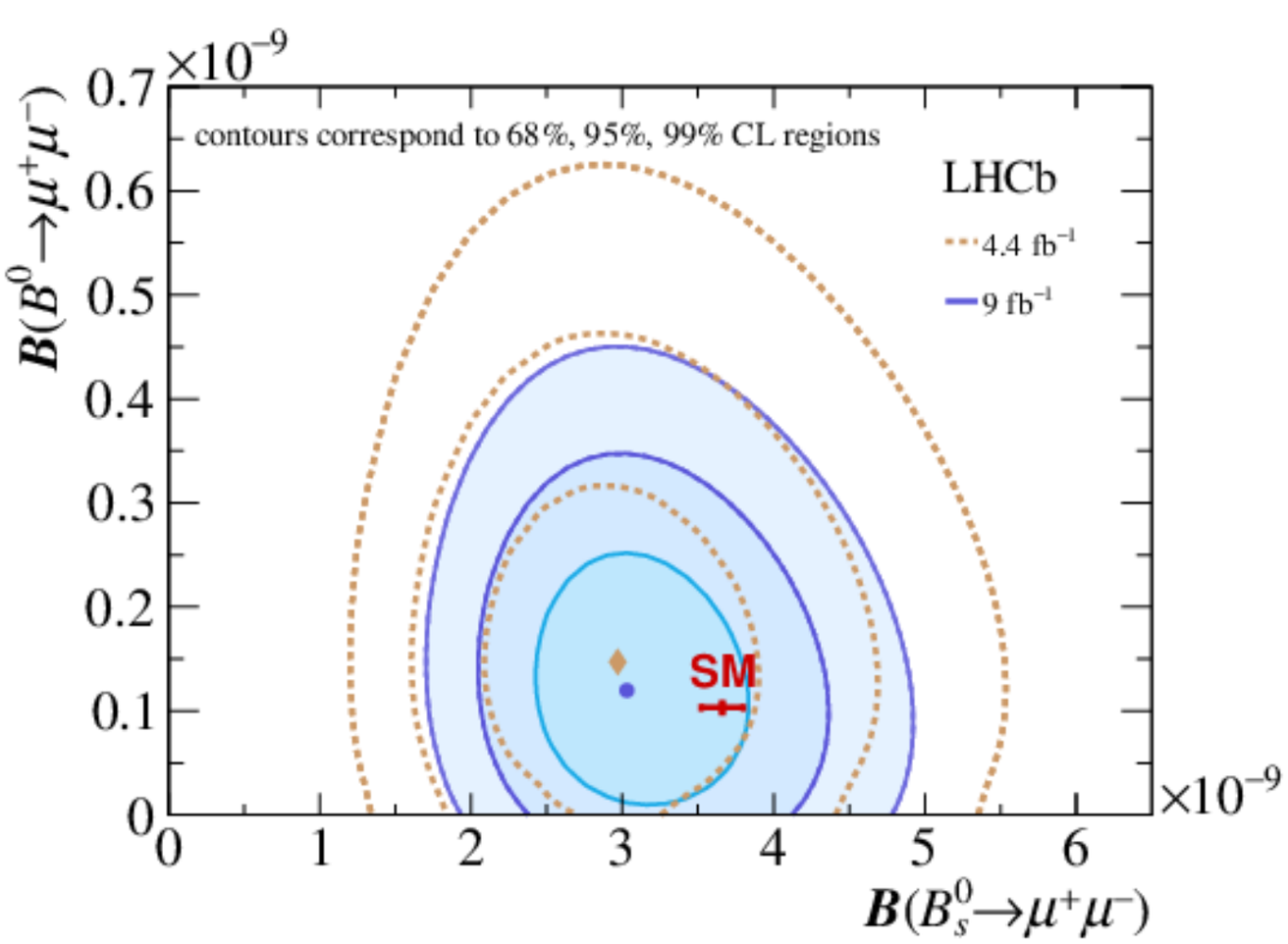
LHCb update with full dataset

- LHCb analysis based on full Run 1 and Run 2 data (9 fb^{-1})

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

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

- Consistent with SM expectation at current level of precision

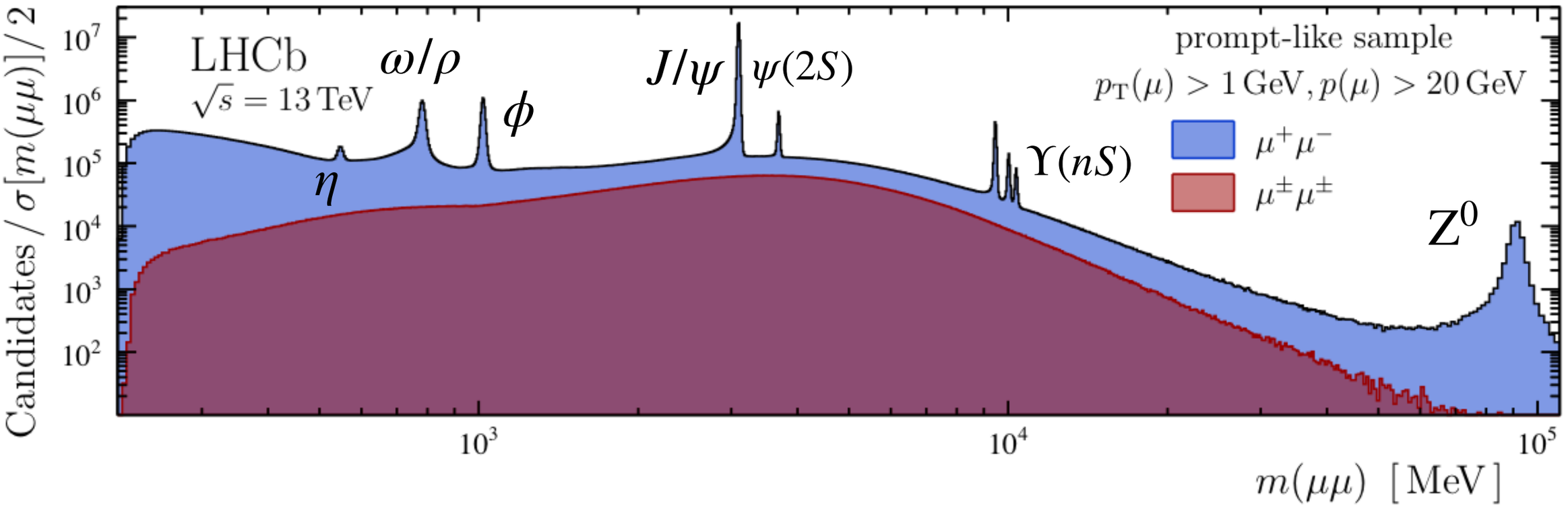


- $B_s \to \mu^+\mu^-$ found with significance >10 σ, but no evidence yet for $B^0 \to \mu^+\mu^-$ (1.7σ)
- Result dominated by statistical uncertainty
- Expect 10% precision with ATLAS/CMS Run 2

PRL 128 (2022) 041801

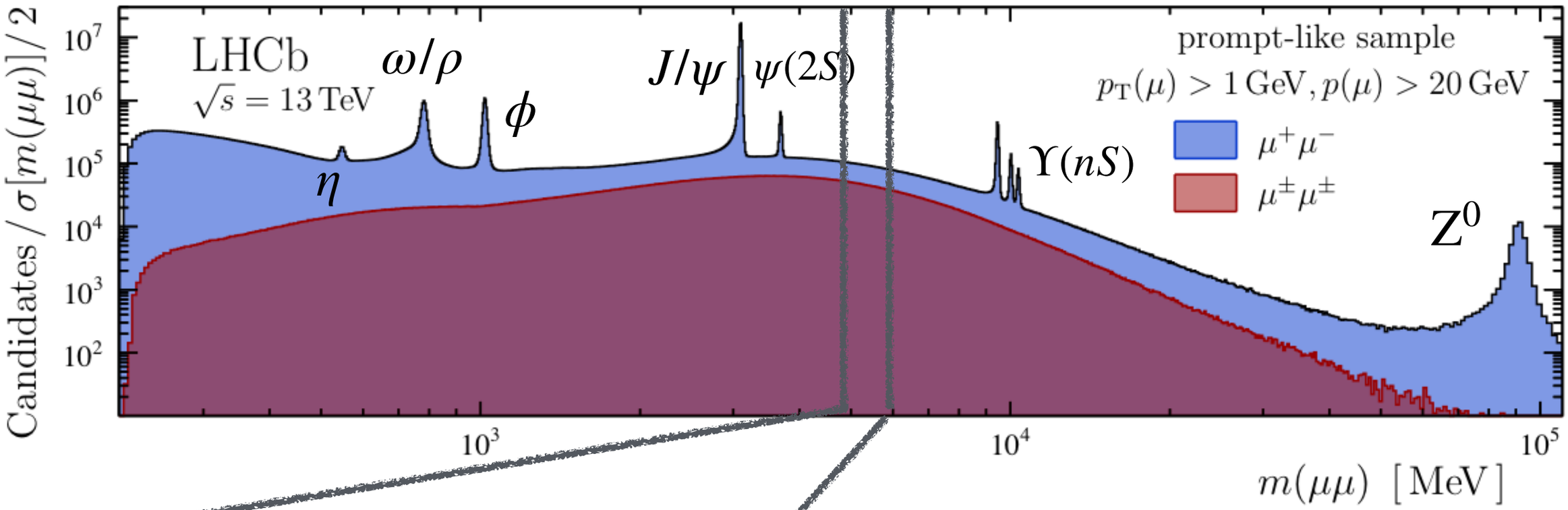
Finding a needle in a haystack!

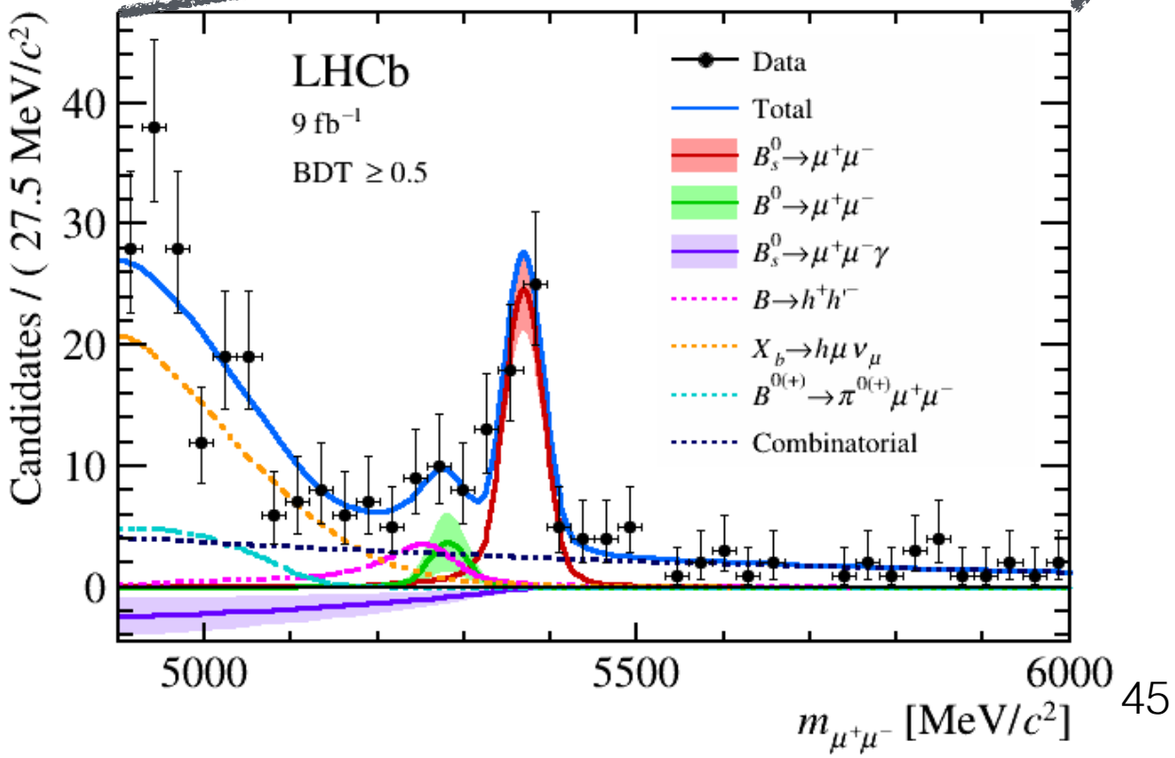
PRL 120 (2018) 061801



Finding a needle in a haystack!

PRL 120 (2018) 061801



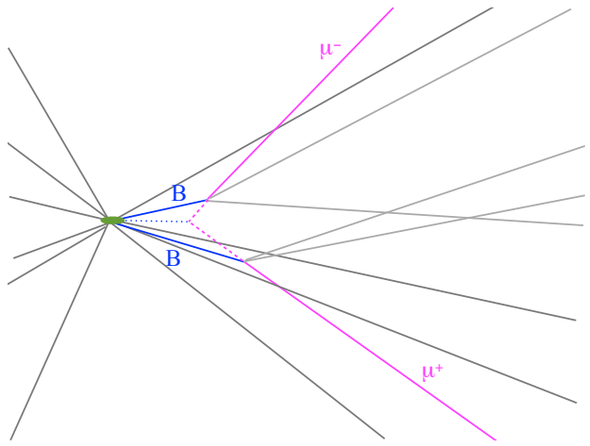


$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.09^{+0.46+0.15}_{-0.43-0.11}) \times 10^{-9}$$

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

How do you find a needle in a haystack?

- Schematic selection requirements:
  - two oppositely-charged muon tracks with common vertex displaced from primary vtx
  - $m_{\mu\mu}$ peaking at the $B^0_{(s)}$ mass
- In practice, complex analysis due to very low signal and large background rates
- Most abundant background is combinatorial
  - muons from two different b-quark semileptonic decays
  - strongly suppressed with multivariate classifier (BDT) using e.g., track isolation, topological and geometrical information



- Use of normalisation channels with well-known BRs, same topology and/or trigger and cancel uncertainties in ratios :
  - Use large samples of $B^+ \to J/\psi K^+$ and $B^0 \to K^+\pi^-$

$$\mathcal{B}(B^0_s \to \mu^+\mu^-) = \frac{N_{B^0_s\to\mu^+\mu^-}}{N_{\text{norm.}}} \times \frac{f_d}{f_s} \times \frac{\varepsilon_{\text{norm.}}}{\varepsilon_{B^0_s\to\mu^+\mu^-}} \times \mathcal{B}_{\text{norm.}}$$

largest systematics to $B_s \to \mu^+\mu^-$ from b-quark fragmentation probability ratio f_s/f_d (~3%)