Measurement of CP-averaged observables in the $B^0 \to K^{*0} \mu^+ \mu^-$ decay

http://arxiv.org/abs/2003.04831

15th April 2020

Mitesh Patel (Imperial College London)

Introduction

- Will talk through my very recent experimental paper, http://arxiv.org/abs/2003.04831
- Idea to explain the main elements of a particle physics analysis – in this case study of the ‘rare decay’ $B^0 \to K^{*0}\mu^+\mu^-$
- Decay of interest because have previously seen some significant tensions with current thinking

Outline

- The aim of the measurements
- How initial raw data become final experimental plot
- Comparison with theoretical predictions, forward look

Outline

- **The aim of the measurements**
- How initial raw data become final experimental plot
- Comparison with theoretical predictions, forward look

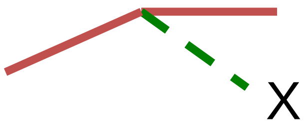
Physics beyond the SM

- The so-called Standard Model (SM) of particle physics has explained essential all experimental observations for decades
- BUT: whole host of open questions:
  - What is origin of dark matter?
    - One or more weakly interacting massive particles (WIMPs)?
  - Why are there so many types of matter particles?
    - Mixing of different flavours of quarks and leptons
    - Observed matter-antimatter difference
  - Are fundamental forces unified?
    - Do all the forces unify at some higher energy scale?
  - What is quantum theory of gravity?
    - String theory?
  - …
- Expect new particles, “new physics” – how to search for this?

Why do we study rare decays?

- Main thing of interest for probing NP: loops/trees
  - NP thought to be less likely to affect decays at tree level
  - Loop decays involve second order (→ suppressed, *potentially* "rare") diagrams in which new, *virtual* particles can contribute

Tree-level decay



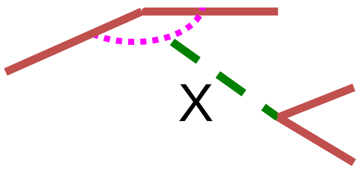
- Most interesting processes those where there is no tree contribution
  - e.g. Flavour Changing Neutral Currents – forbidden at tree level in SM

    → FCNC processes necessarily involve loops
  - Loops can involve (virtual) NP particles!

    → Can probe masses > CM energy of accelerator

    → Model independent probes! Whatever is in the loop we measure it!

Loop decay

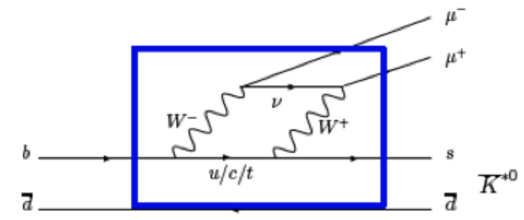


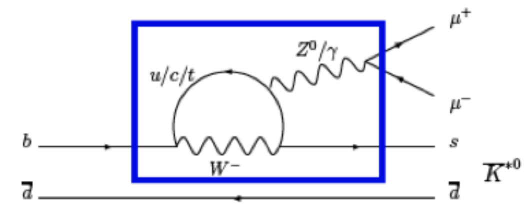
- In order to gain information – need to compare experimental measurements to precise theory predictions

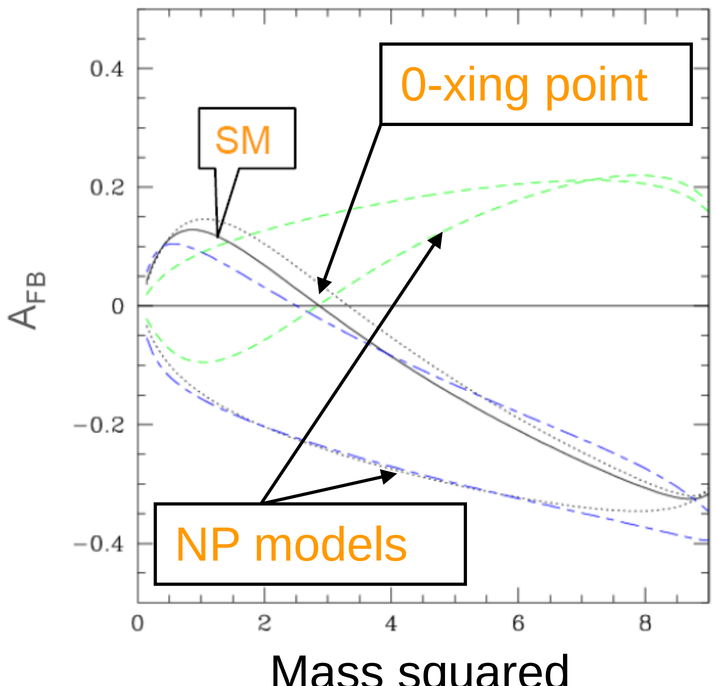
$B^0 \to K^{*0}\mu\mu$

- Flavour changing neutral current → loop process (→ sensitive to NP)
- Decay described by three angles (θ_l, ϕ, θ_K) and di-μ invariant mass q^2
- Try to use observables where theoretical uncertainties cancel

  e.g. Forward-backward asymmetry A_{FB} of θ_l distribution
  - Zero-crossing point: ±6% uncertainty on theory prediction







Origin of theoretical uncert.

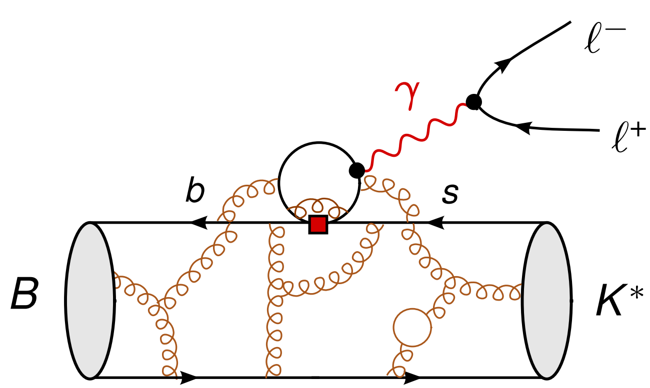
- Major theoretical uncertainties come from so called "form-factors"
  - May have met these considering scattering from an object extended in space
  - Rather than point-like charge, have some density $\rho(\vec{r})$ such that $\int \mathrm{d}^3 r\, \rho(\vec{r}) = 1$
  - Fourier transform is called the form factor

$$F(\vec{q}) = \int \mathrm{d}^3 r\, \exp[-i\vec{q}\vec{r}]\rho(\vec{r}) \Longrightarrow F(0) = 1$$

  - Modification of cross sections for scattering on such an object

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}^2\Omega}\right|_{\mathrm{extended}} \approx \left.\frac{\mathrm{d}\sigma}{\mathrm{d}^2\Omega}\right|_{\mathrm{ptlike}} |F(q)|^2$$

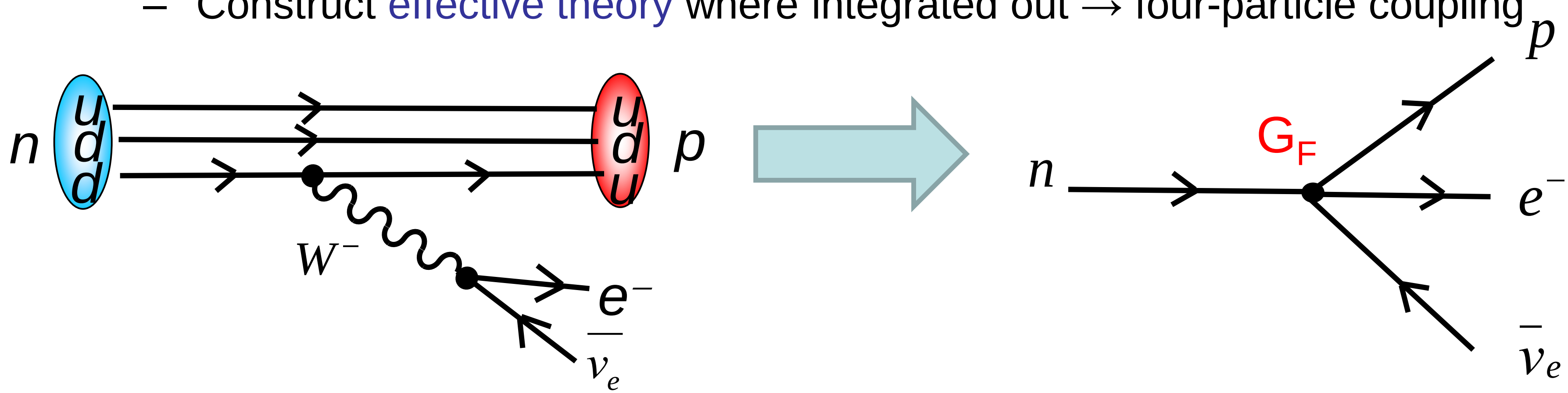
- In our case, don't have a short distance $b\to s\mu\mu$ quark level transition, have :



The Operator Product Expansion

- Don't want interpretation of measurements to depend on theory model to which compare – Make an *effective theory* which gives us *model independent things to measure*

- Most familiar example of this Fermi's theory of beta decays
  - Z and W are very massive – the weak interactions take place at very short distance scales $O(1/M_W^2)$
  - Construct effective theory where integrated out → four-particle coupling



  - For $q^2 << m_W^2$ can replace W propagator

  - Effectively absorbs the contribution from the W into the factor G_F, in the limit when W is too heavy to be resolved

The Operator Product Expansion

- Rewrite (part of) SM Lagrangian as:

$$\mathcal{L} = \sum_i C_i O_i$$

- "Wilson Coefficients" C_i
  - Describe the short distance part, can compute perturbatively in a given theory
  - Integrate out the heavy degrees of freedom that can't resolve at some energy scale $\mu \rightarrow$ Wilson coefficient just a (complex) number
  - All degrees of freedom with mass>μ are taken into account by the Wilson Coefficients, while those with mass<μ go into the operators ...

- "Operators" O_i
  - Describe the long distance, non-perturbative part involving particles below the scale μ
  - **Form a complete basis** – can put in all operators from NP/SM
  - Account for effects of strong interactions and are *difficult to calculate reliably*

$B^0 \to K^{*0}\mu\mu$ form factors

- Amplitudes that describe the $B^0 \to K^{*0}\mu\mu$ decay involve
  - The (effective) Wilson Coefficients :

    C_7^{eff} (photon), C_9^{eff} (vector), C_{10}^{eff} (axial-vector)
  - Seven (!) form factors

$$A_\perp^{L(R)} = N\sqrt{2\lambda}\left\{\left[(\mathbf{C_9^{eff}} + \mathbf{C_9^{\prime eff}}) \mp (\mathbf{C_{10}^{eff}} + \mathbf{C_{10}^{\prime eff}})\right]\frac{\mathbf{V(q^2)}}{m_B + m_{K^*}} + \frac{2m_b}{q^2}(\mathbf{C_7^{eff}} + \mathbf{C_7^{\prime eff}})\mathbf{T_1(q^2)}\right\}$$

$$A_\parallel^{L(R)} = -N\sqrt{2}(m_B^2 - m_{K^*}^2)\left\{\left[(\mathbf{C_9^{eff}} - \mathbf{C_9^{\prime eff}}) \mp (\mathbf{C_{10}^{eff}} - \mathbf{C_{10}^{\prime eff}})\right]\frac{\mathbf{A_1(q^2)}}{m_B - m_{K^*}} + \frac{2m_b}{q^2}(\mathbf{C_7^{eff}} - \mathbf{C_7^{\prime eff}})\mathbf{T_2(q^2)}\right\}$$

$$A_0^{L(R)} = -\frac{N}{2m_{K^*}\sqrt{q^2}}\left\{\left[(\mathbf{C_9^{eff}} - \mathbf{C_9^{\prime eff}}) \mp (\mathbf{C_{10}^{eff}} - \mathbf{C_{10}^{\prime eff}})\right]\left[(m_B^2 - m_{K^*}^2 - q^2)(m_B + m_{K^*})\mathbf{A_1(q^2)} - \lambda\frac{\mathbf{A_2(q^2)}}{m_B + m_{K^*}}\right]\right.$$
$$\left. + 2m_b(\mathbf{C_7^{eff}} - \mathbf{C_7^{\prime eff}})\left[(m_B^2 + 3m_{K^*} - q^2)\mathbf{T_2(q^2)} - \frac{\lambda}{m_B^2 - m_{K^*}^2}\mathbf{T_3(q^2)}\right]\right\}$$

- BFs have relatively large theoretical uncertainties from form factors

- Angular observables much smaller theory uncertainties

Outline

- The aim of the measurements
- **How initial raw data become final experimental plot**
- Comparison with theoretical predictions, forward look

$B^0 \to K^{*0}\mu\mu$ angular distribution

- Can write down angular distribution for B^0 decays, assuming $\overline{B^0}$ behaves the same ("CP-averaged")

$$\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2} \frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\,\mathrm{d}\vec{\Omega}}\bigg|_{\mathrm{P}} = \frac{9}{32\pi}\Big[\tfrac{3}{4}(1-F_{\mathrm{L}})\sin^2\theta_K + F_{\mathrm{L}}\cos^2\theta_K$$
$$+\tfrac{1}{4}(1-F_{\mathrm{L}})\sin^2\theta_K\cos 2\theta_l$$
$$-F_{\mathrm{L}}\cos^2\theta_K\cos 2\theta_l + S_3\sin^2\theta_K\sin^2\theta_l\cos 2\phi \qquad (1)$$
$$+S_4\sin 2\theta_K\sin 2\theta_l\cos\phi + S_5\sin 2\theta_K\sin\theta_l\cos\phi$$
$$+\tfrac{4}{3}A_{\mathrm{FB}}\sin^2\theta_K\cos\theta_l + S_7\sin 2\theta_K\sin\theta_l\sin\phi$$
$$+S_8\sin 2\theta_K\sin 2\theta_l\sin\phi + S_9\sin^2\theta_K\sin^2\theta_l\sin 2\phi\Big],$$

- Quantities F_L, A_{FB}, S_i dependent on q^2 and determine WCs
- Observe K^{*0} through its (strong) decay to $K^+\pi^-$ final state – but this can occur in two different angular momentum configurations

$B^0 \rightarrow K^{*0}\mu\mu$ angular distribution

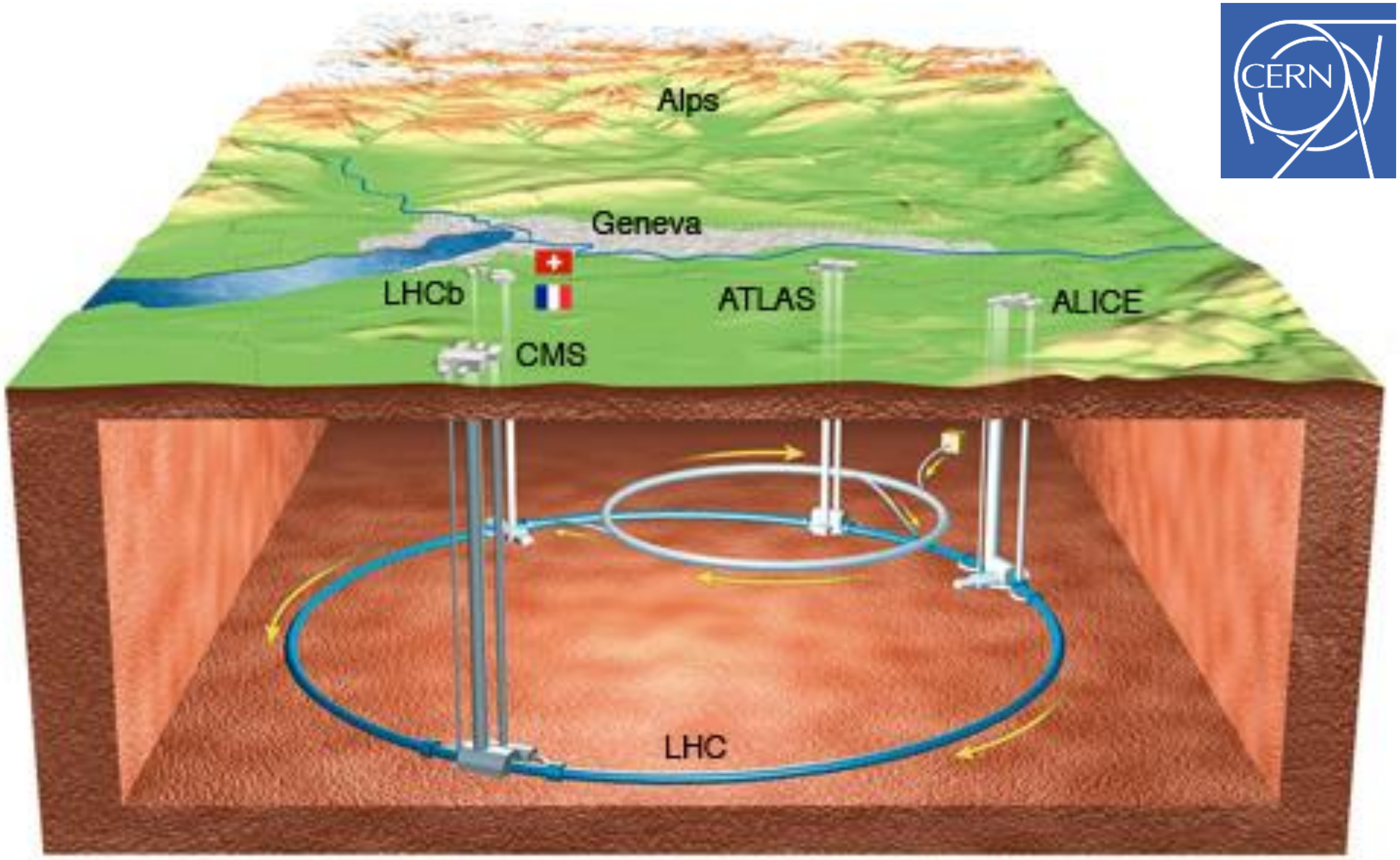
- Modified angular distribution

$$\left.\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\,\mathrm{d}\vec{\Omega}}\right|_{\mathrm{S+P}} = (1-F_{\mathrm{S}})\left.\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\,\mathrm{d}\vec{\Omega}}\right|_{\mathrm{P}}$$
$$+\frac{3}{16\pi}F_{\mathrm{S}}\sin^2\theta_l$$
$$+\frac{9}{32\pi}(S_{11}+S_{13}\cos 2\theta_l)\cos\theta_K \qquad (2)$$
$$+\frac{9}{32\pi}(S_{14}\sin 2\theta_l+S_{15}\sin\theta_l)\sin\theta_K\cos\phi$$
$$+\frac{9}{32\pi}(S_{16}\sin\theta_l+S_{17}\sin 2\theta_l)\sin\theta_K\sin\phi\,,$$

- Need to isolate $B^0 \rightarrow K^{*0}\mu\mu$ decays, measure the angles and q^2, and fit for these observable quantities that can be predicted in terms of the WCs by theorists... Simple!
- Note can parameterise in alternative basis, P_i observables

The Large Hadron Collider

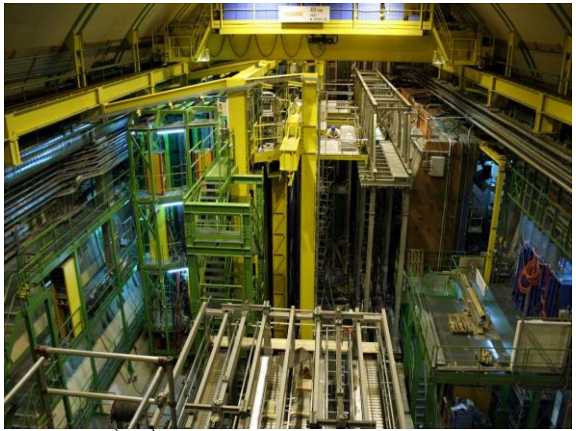
- World's highest energy particle accelerator

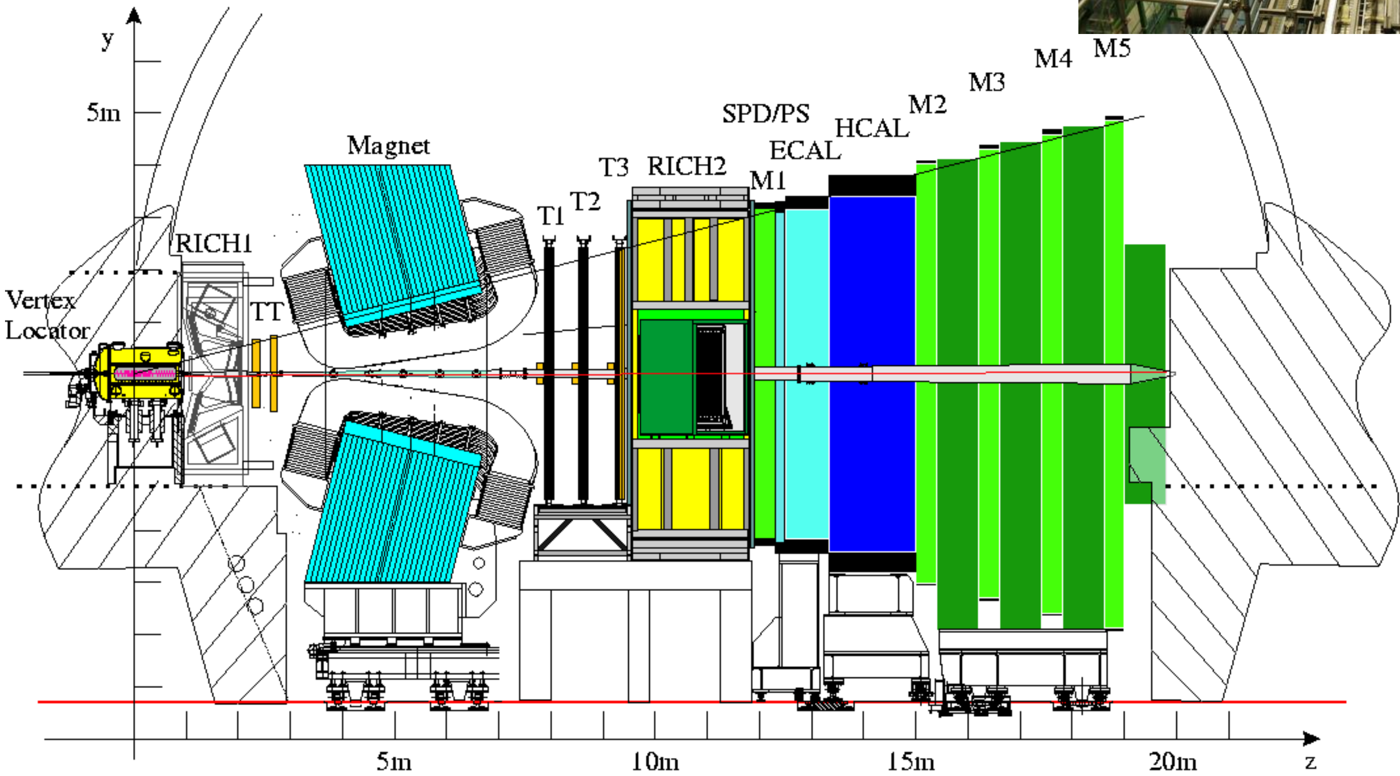


- ... Also copious source of B mesons

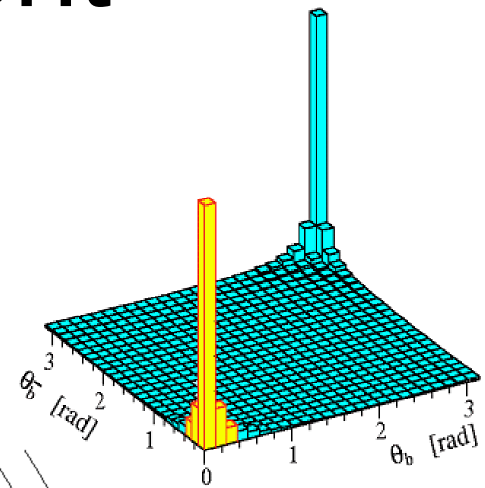
The LHCb Experiment

- The LHCb experiment looks very different to the central detectors:



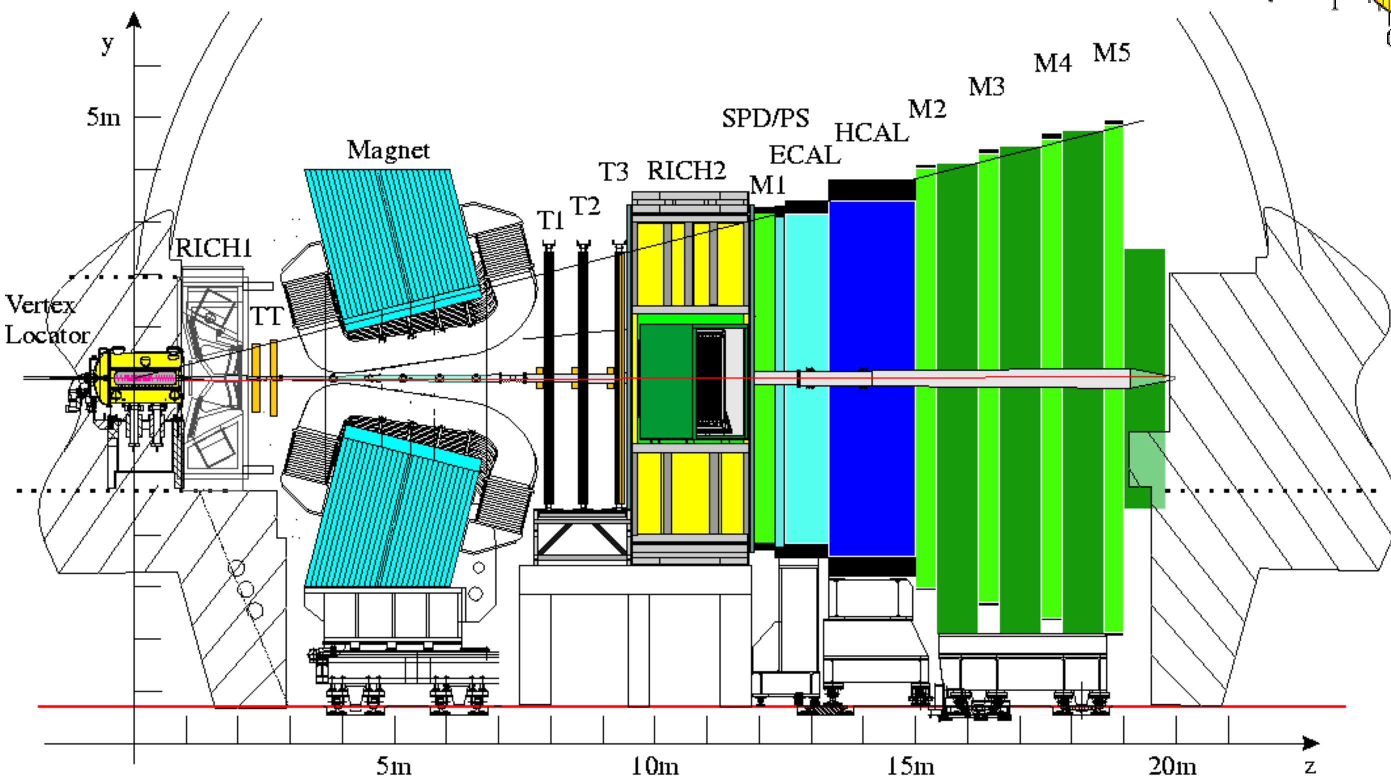


The LHCb Experiment



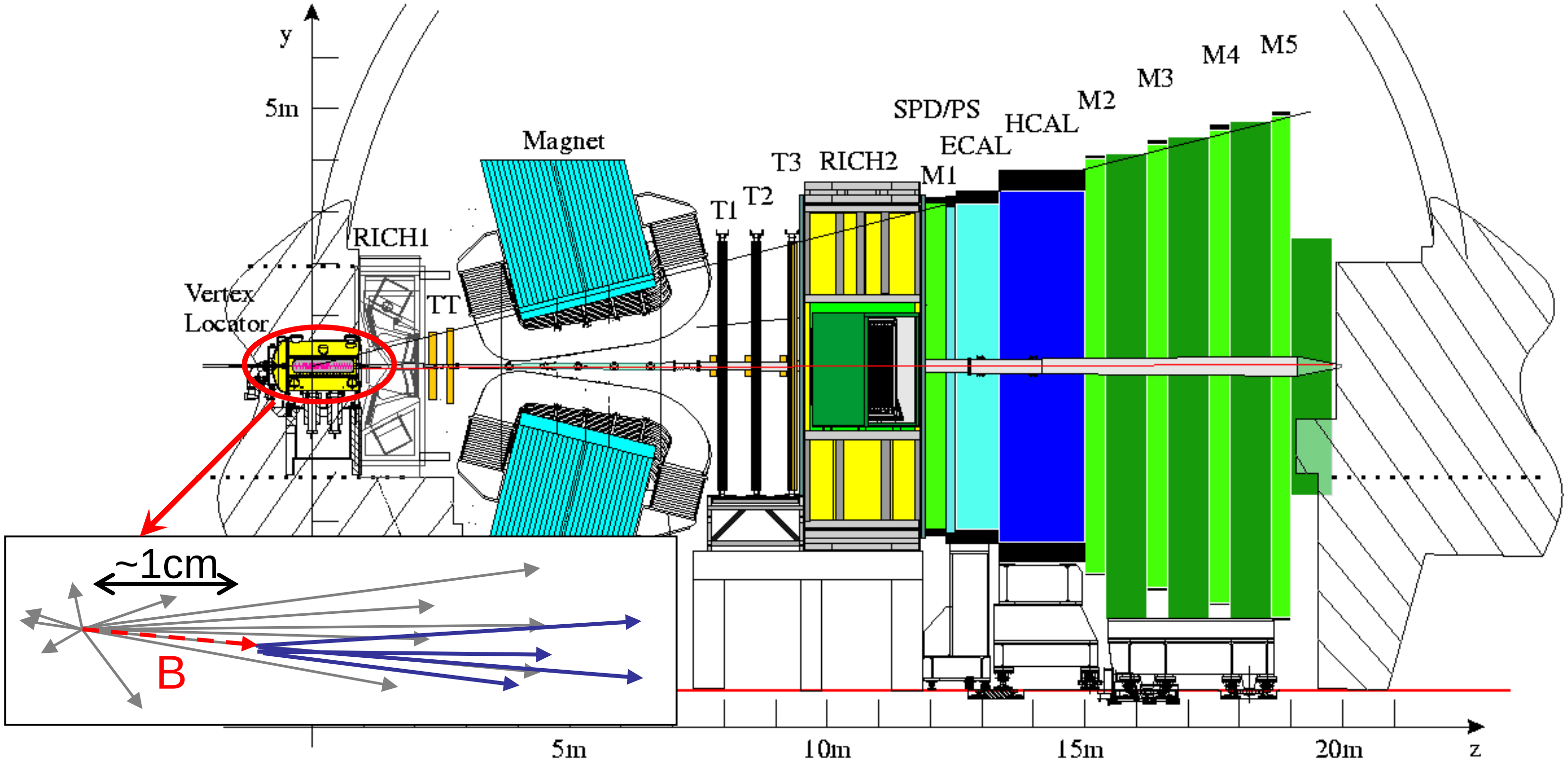
- b production predominately at small polar angles

  → forward spectrometer



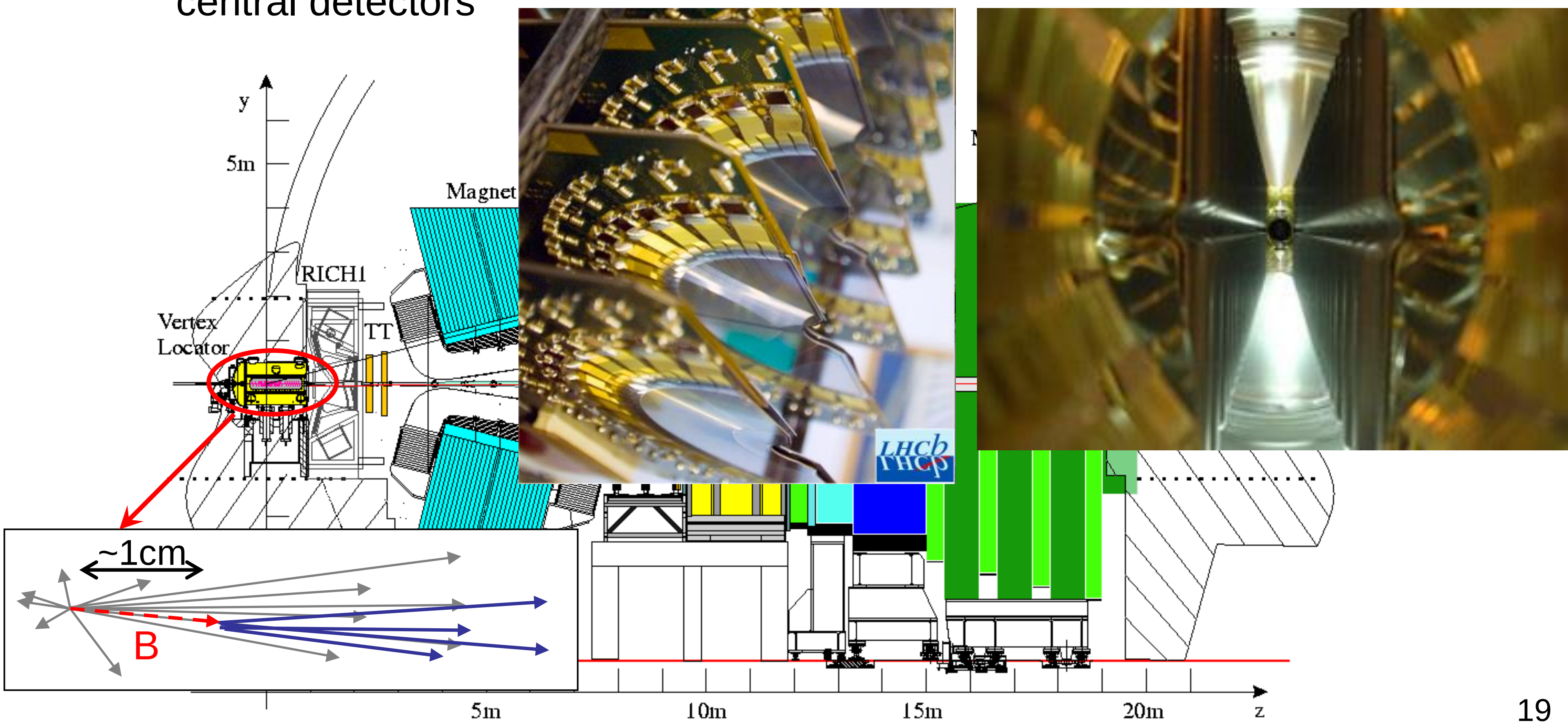
The LHCb Experiment

- B lifetime → displaced secondary vertex
  - Vertex detector capable of picking out the displaced vertex
  - Need ~1 interaction/event → operate at luminosity 10 times lower that central detectors



The LHCb Experiment

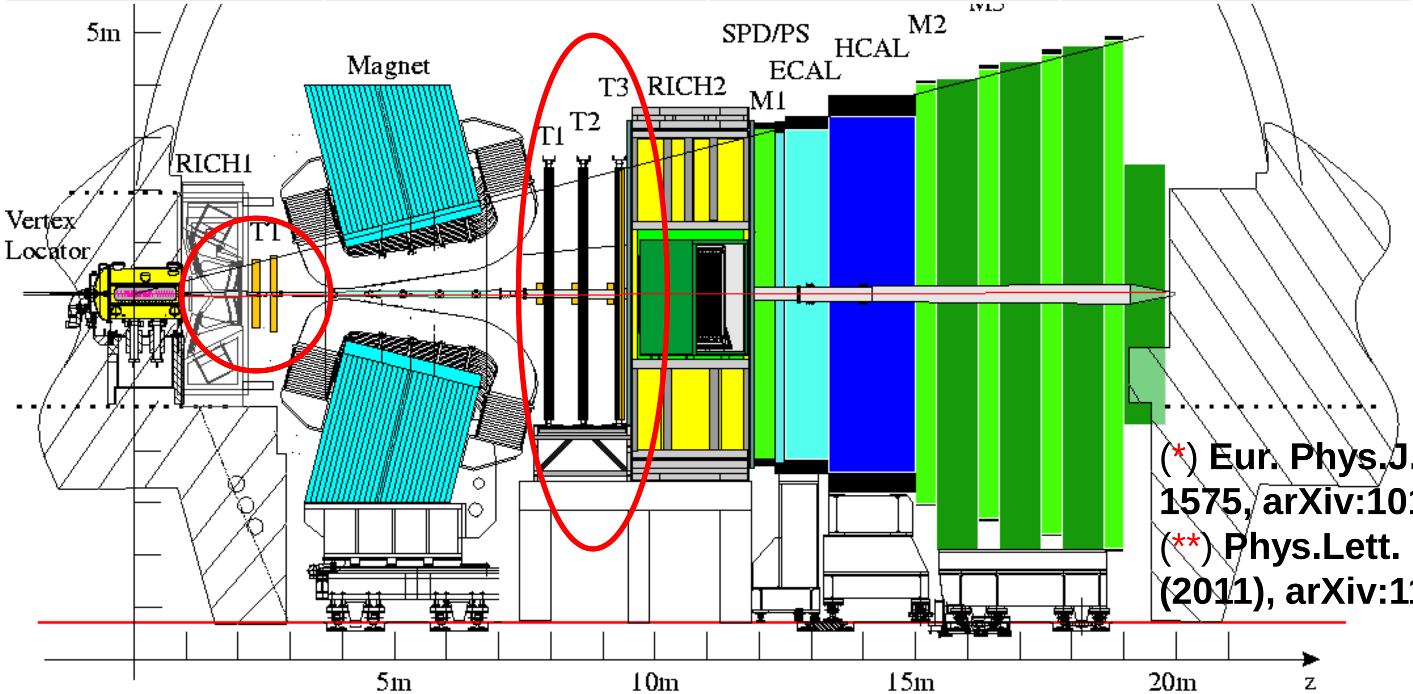
- B lifetime → displaced secondary vertex
  - Vertex detector capable of picking out the displaced vertex
  - Need ~1 interaction/event → operate at luminosity 10 times lower that central detectors



The LHCb Experiment

- Precision momentum resolution → mass resolution

	LHCb	CMS	ATLAS
Momentum Resolution	$\delta p/p$=0.4-0.6%	$\delta p_T/p_T$=1-3%	$\delta p_T/p_T$=5-6%
Mass resoln $J/\psi\rightarrow\mu\mu$	13 MeV/c^2	28 MeV/c^2 (*)	46 MeV/c^2 (**)

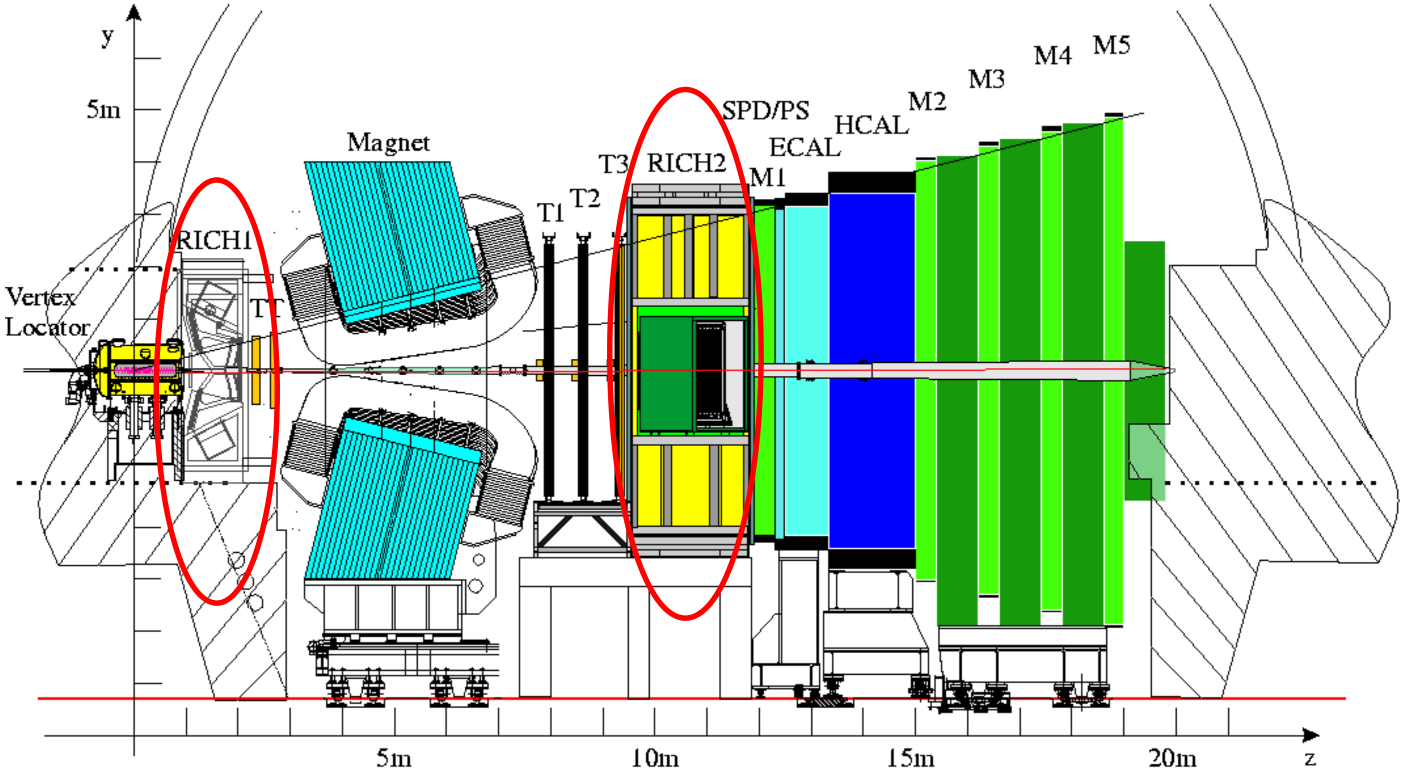


(*) Eur. Phys.J. C71 (2011) 1575, arXiv:1011.4193
(**) Phys.Lett. B697 (2011), arXiv:1104.3038v2

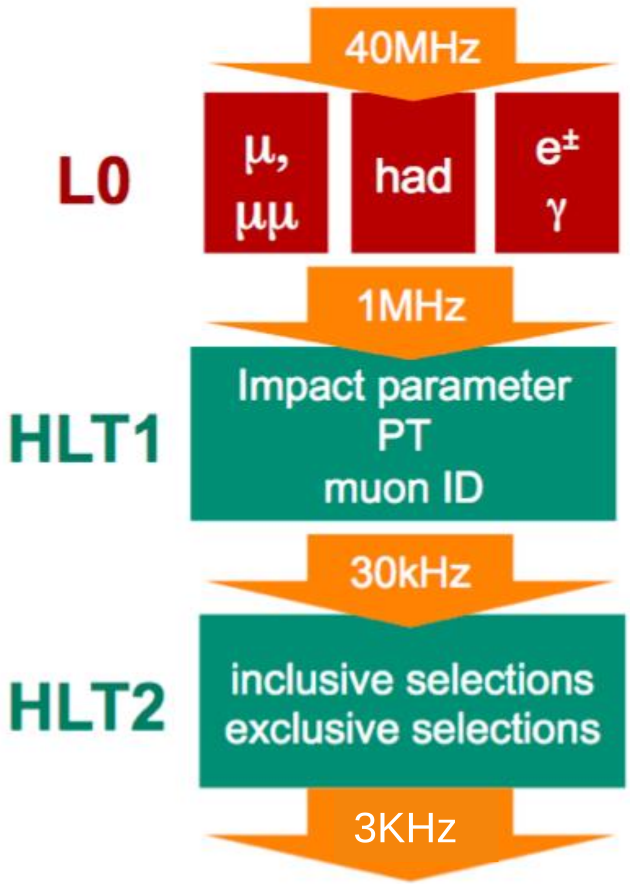
The LHCb Experiment

- Many of final states of interest contain kaons, in general decays dominated by pions

  → particle identification critical



The “Trigger”



L0 Hardware	“high p_T” signals in calorimeter and muon systems
HLT1 Software	Partial reconstruction, selection based on one or two (dimuon) displaced tracks, muon ID
HLT2 Software	Global reconstruction (very close to offline) dominantly inclusive signatures – use MVA

+ Global Event Cuts for events with high multiplicity

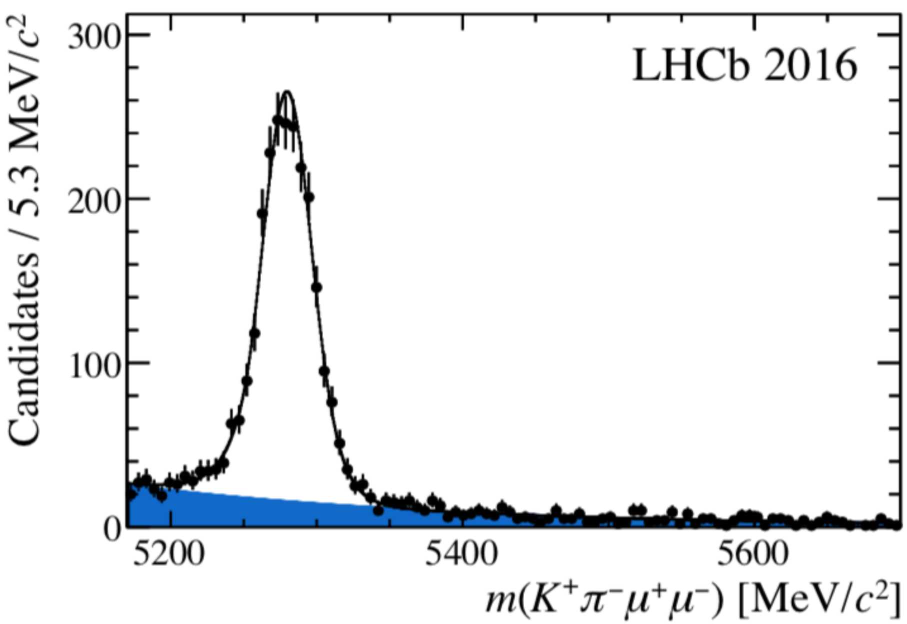
	Charm	Had. B	Lept. B
Overall efficiency	~10%	~40%	~75-90%

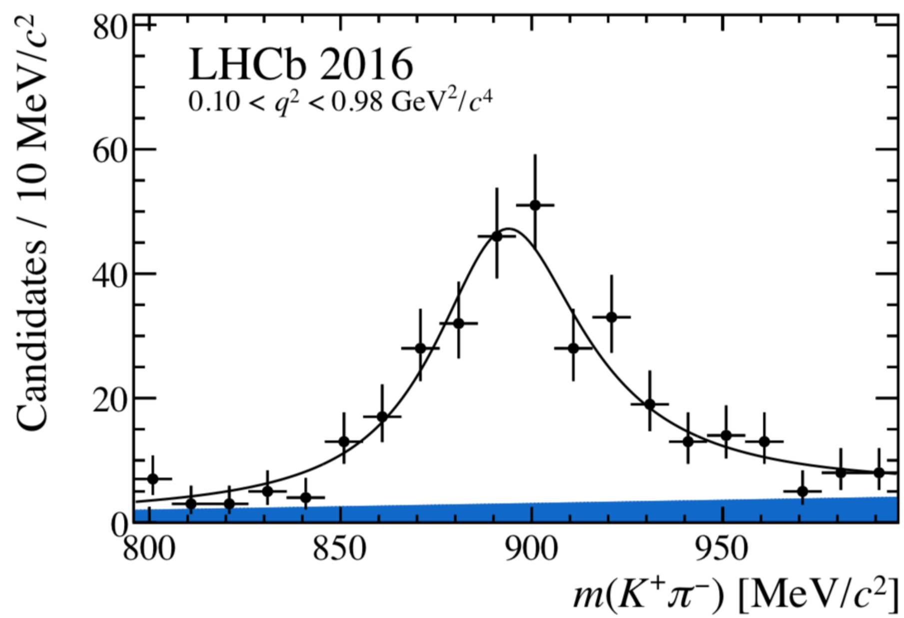
- Small event size (60kB) → large bandwidth
- Allows low thresholds

Fitting the angular distribution

- $m(K^+\pi^-\mu^+\mu^-)$ is used to discriminate signal from background – fit this mass simultaneously with the three angles in bins of q^2

- $m(K^+\pi^-)$ is used to constrain the angular-momentum configuration

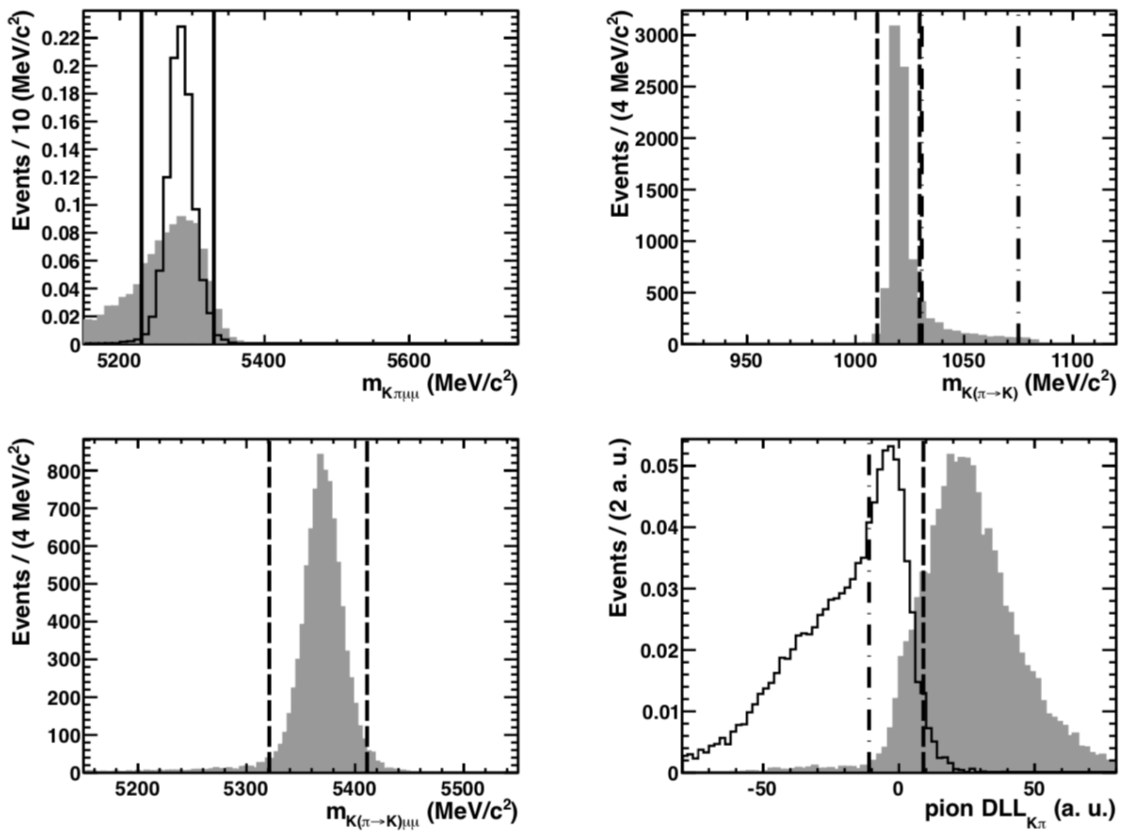




- Cross-check procedure using $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$ decay, same final state but 100× more prevalent

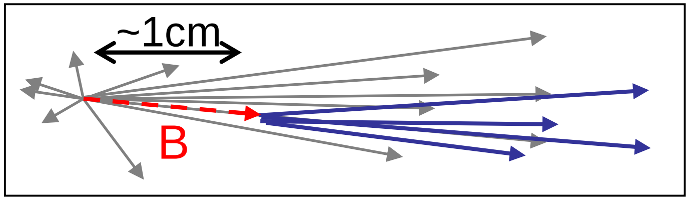
Backgrounds

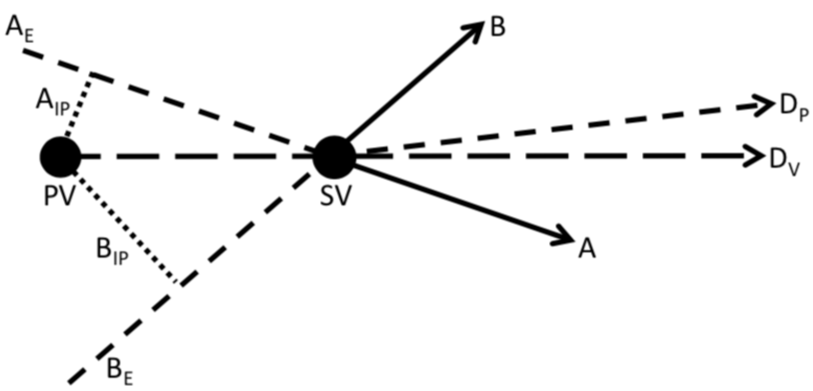
- Two classes considered :
  - Combinatorial : selected particles do not originate from a single b-hadron decay; can then have any mass – use MVA to reject
  - `Peaking' : single decay selected but with some of the final-state particles misidentified, accumulates somewhere in $m(K^+\pi^-\mu^+\mu^-)$ – use specific vetoes to reject



Selection cuts

- The four tracks of the final-state particles are required to have significant impact parameter (IP) with respect to all primary vertices (PVs) in the event
- The IP of the B^0 candidate is required to be small with respect to one of the PVs
- The vertex of the B^0 candidate is required to be significantly displaced from the same PV
- The angle between the reconstructed B^0 momentum and the vector connecting this PV to the reconstructed B^0 decay vertex, θ_{DIRA}, is also required to be small

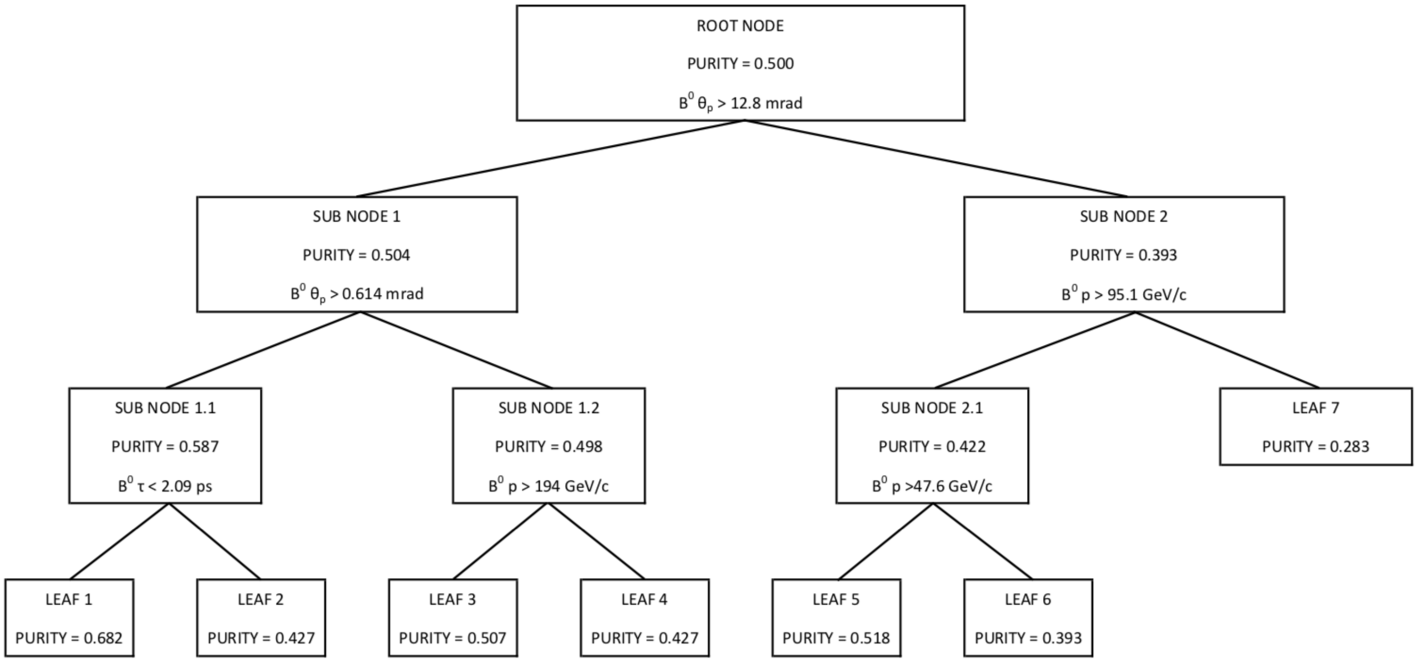




- The tracks are fitted to a common vertex, which is required to be of good quality
- To avoid the same track being used to construct more than one of the final state particles, the opening angle between every pair of tracks is required to be larger than 1 mrad
- q^2 vetoes

Multivariate analysis

- Combinatorial background is reduced further using a boosted decision tree (BDT) algorithm



- Rejects more than 97% of the remaining combinatorial background, while retaining more than 85% of the signal

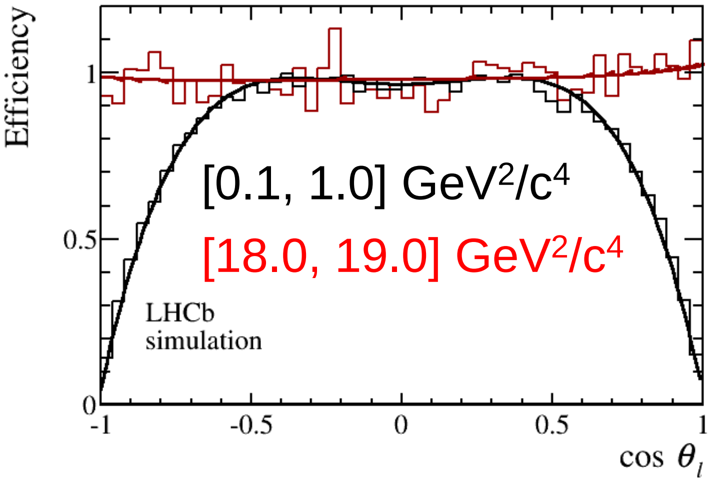
Correcting for the efficiency

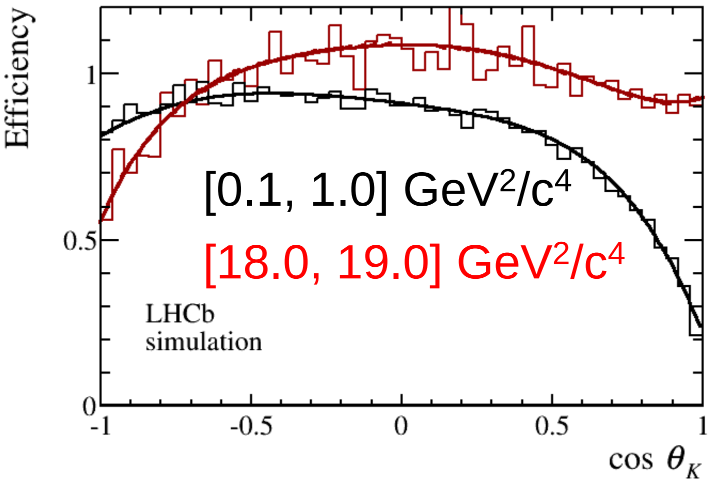
- The angular distribution is multiplied by an acceptance model used to account for the effect of the reconstruction and candidate selection

- Compute 4D efficiency function, ε, using simulated events

$$\varepsilon(\cos\theta_l, \cos\theta_K, \phi, q^2)$$

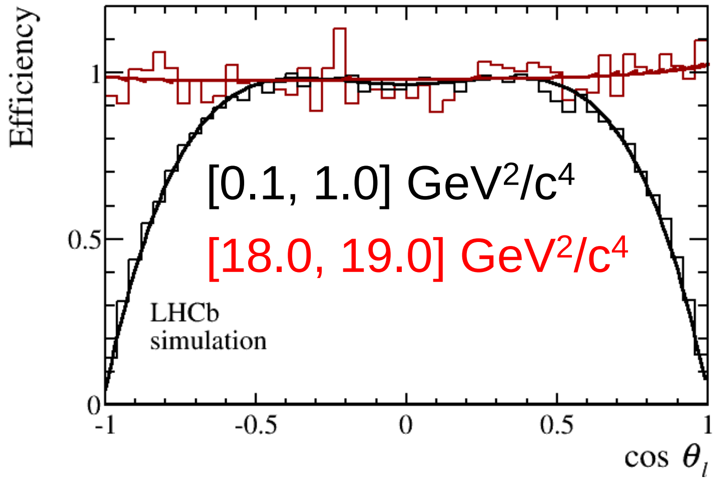
- Function of all underlying variables $\rightarrow$ can determine with a phase-space simulation





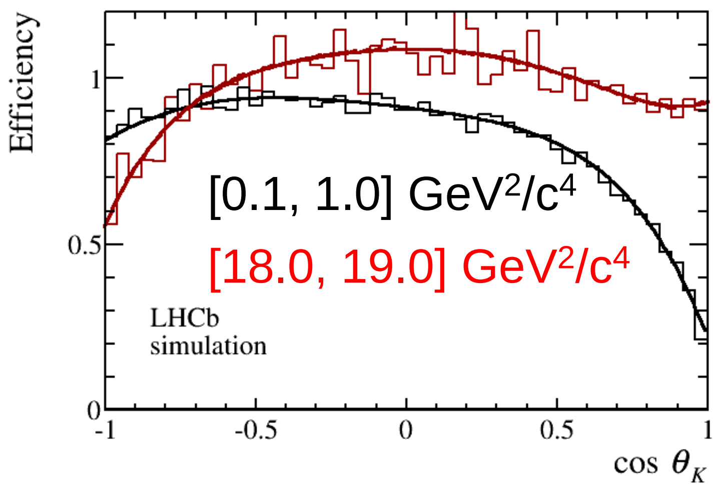
Correcting for the efficiency

- Parameterised:
  - $P_i(x)$ are Legendre polynomials of order i (x rescaled -1→1)
  - For **$\cos\theta_l$, $\cos\theta_K$, ϕ, q^2** use up-to and including 4^{th},5^{th},6^{th}, 5^{th} order polynomials



$$\varepsilon(\cos\theta_l, \cos\theta_K, \phi, q^2) = \sum_{ijmn} c_{ijmn} L_i(\cos\theta_l) L_j(\cos\theta_K) L_m(\phi) L_n(q^2)\ , \qquad (3)$$

- Coeff $\mathbf{c_{klmn}}$ determined using principal moments tech. (…)

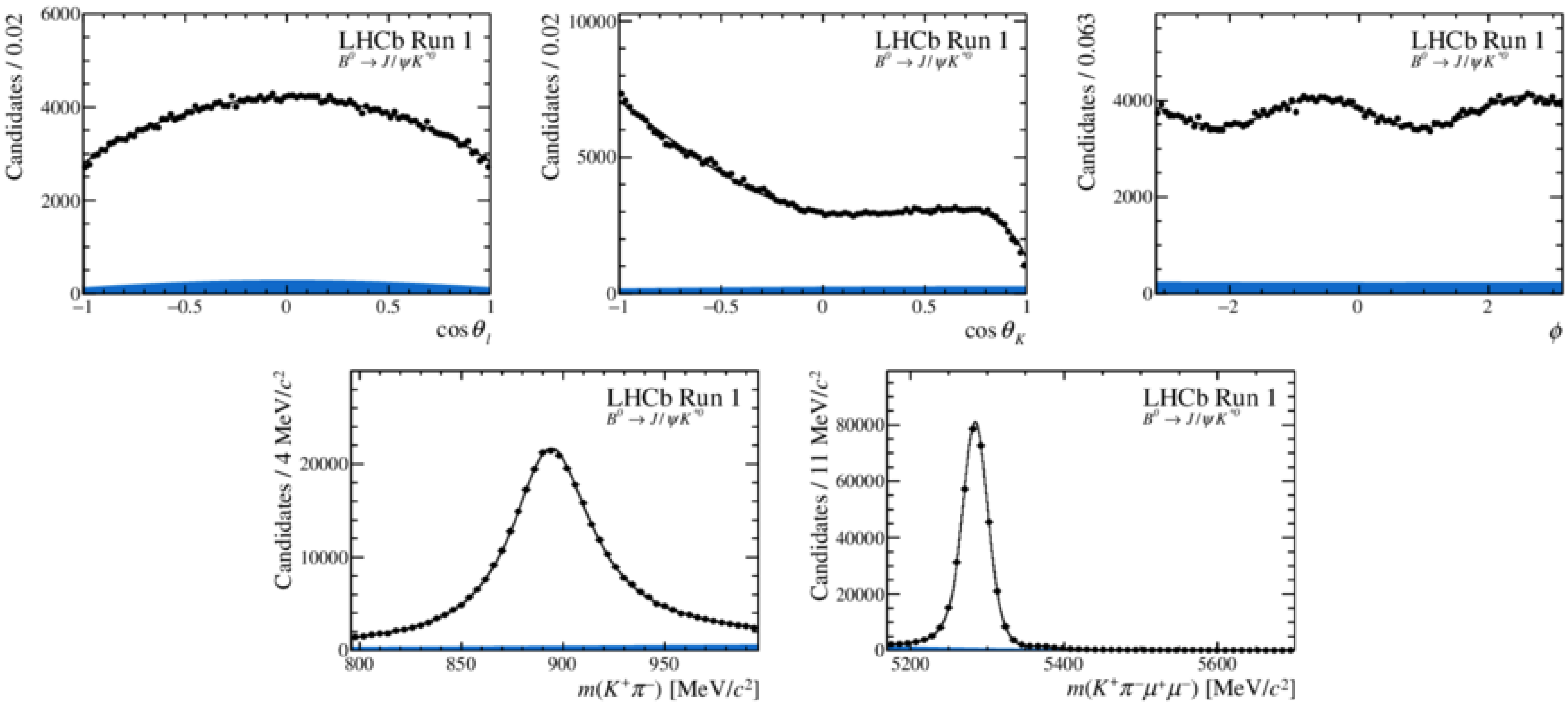


Modelling the distributions

- Signal angular distribution already discussed
- Background angular distribution is modelled with second-order polynomials in $\cos\theta_l$, $\cos\theta_K$ and φ, angular coefficients are allowed to vary in the fit
- $m(K^+\pi^-\mu^+\mu^-)$ shape of the signal modelled with sum of two Gaussian functions with a common mean, each with a power-law tail on the low mass side – parameters determined from $B^0 \to J/\psi K^{*0}$ decay
- $m(K^+\pi^-\mu^+\mu^-)$ shape of the background modelled with an exponential
- $m(K^+\pi^-)$ shape of signal from first principles (relativistic Breit Wigner) and other data (LASS experiment); background taken to be linear

Validating the fit procedure

- $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$ fit reproduces the angular observables measured elsewhere



Fit bias, coverage

- Given large number of parameters, fit may not return the observable quantities in an unbiased way and/or may not have e.g. 68% of the population within +-1σ of the estimated uncertainties (do fit errors provide the correct "coverage")

- This is tested with simulated experiments. Referred to as 'pseudoexperiments' in the paper or 'toys' colloquially

- Biases observed are small: <10% of statistical uncertainty

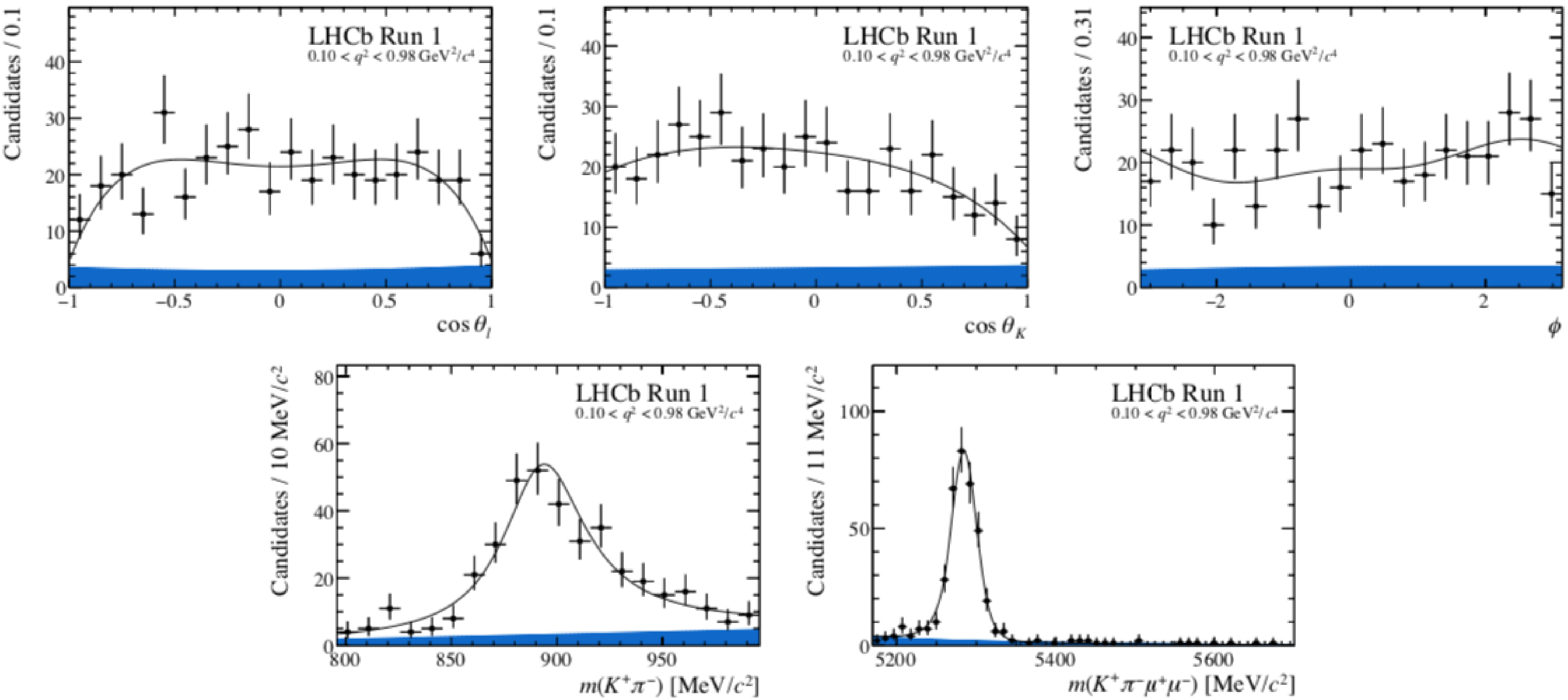
- Statistical uncertainty is corrected to account for under- or over-coverage and a systematic uncertainty equal to the size of the observed bias is assigned

Systematic uncertainties

- Statistically dominated results – systematic uncertainties are small
- Again assessed with pseudoexperiments
- Vary one or more assumptions, determine angular observables with and without this variation, take average of the difference between the two as the uncertainty

Source	$F_{\rm L}$	$A_{\rm FB}$, S_3–S_9	P_1–P'_8
Acceptance stat. uncertainty	< 0.01	< 0.01	< 0.01
Acceptance polynomial order	< 0.01	< 0.01	< 0.02
Data-simulation differences	< 0.01	< 0.01	< 0.01
Acceptance variation with q^2	< 0.03	< 0.01	< 0.09
$m(K^+\pi^-)$ model	< 0.01	< 0.01	< 0.01
Background model	< 0.01	< 0.01	< 0.02
Peaking backgrounds	< 0.01	< 0.02	< 0.03
$m(K^+\pi^-\mu^+\mu^-)$ model	< 0.01	< 0.01	< 0.01
$K^+\mu^+\mu^-$ veto	< 0.01	< 0.01	< 0.01
Trigger	< 0.01	< 0.01	< 0.01
Bias correction	< 0.02	< 0.01	< 0.03

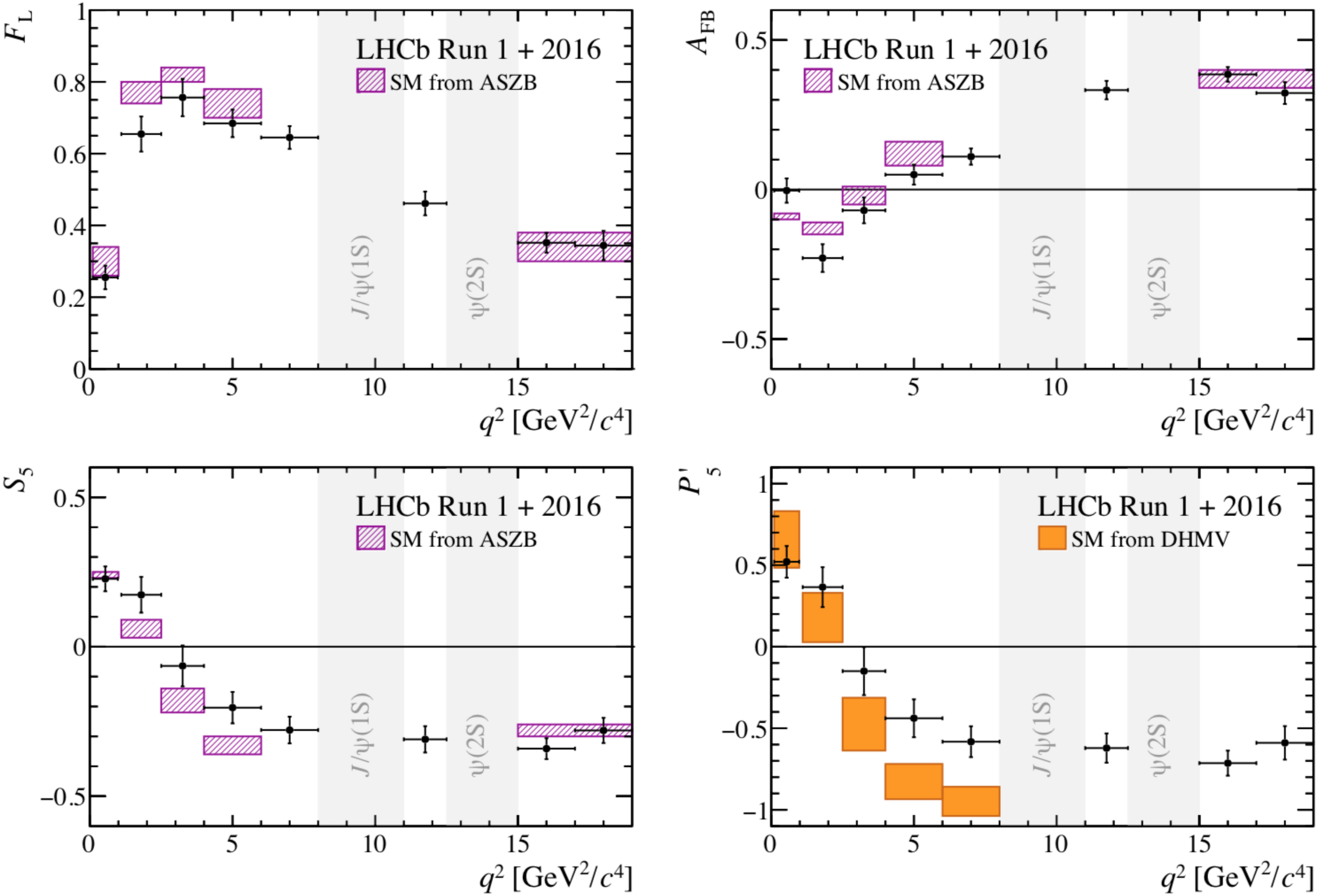
Fit Projections



Outline

- The aim of the measurements
- How initial raw data become final experimental plot
- **Comparison with theoretical predictions, forward look**

Fit Results

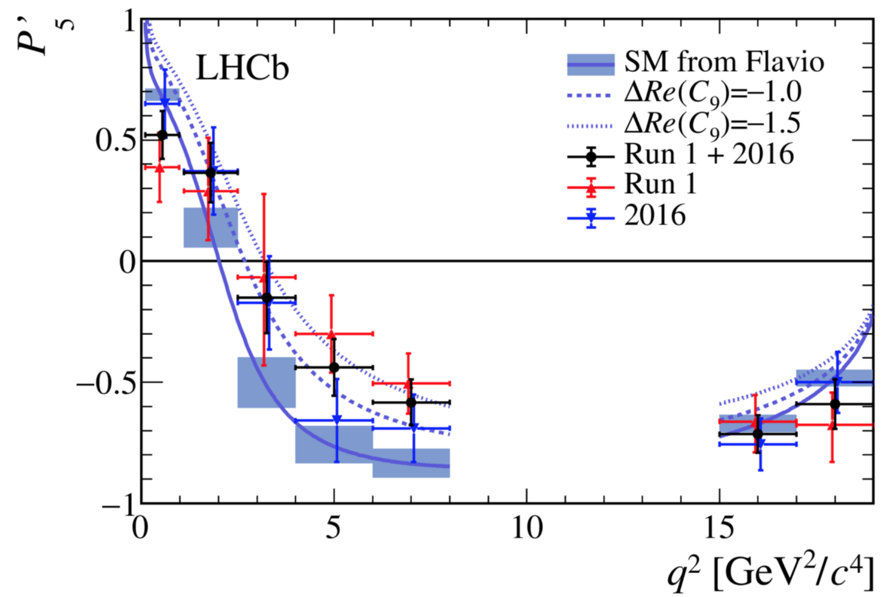


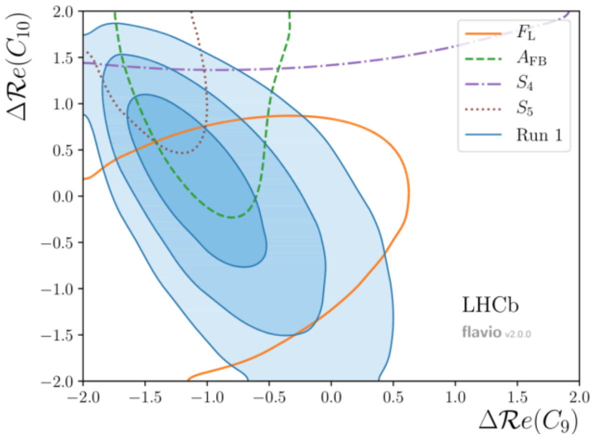
F_L
LHCb Run 1 + 2016
SM from ASZB
$J/\psi(1S)$
$\psi(2S)$
q^2 [GeV$^2/c^4$]
A_{FB}
LHCb Run 1 + 2016
SM from ASZB
$J/\psi(1S)$
$\psi(2S)$
q^2 [GeV$^2/c^4$]
S_5
LHCb Run 1 + 2016
SM from ASZB
$J/\psi(1S)$
$\psi(2S)$
q^2 [GeV$^2/c^4$]
P'_5
LHCb Run 1 + 2016
SM from DHMV
$J/\psi(1S)$
$\psi(2S)$
q^2 [GeV$^2/c^4$]

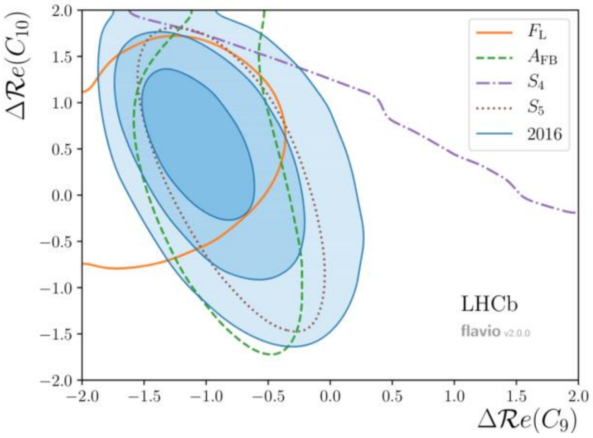
Interpretation

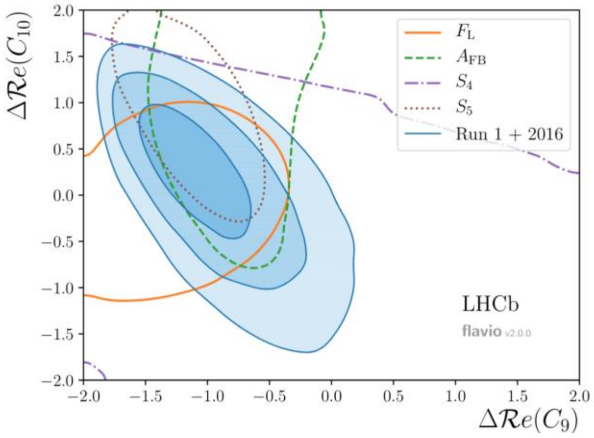
- The local discrepancy in the P_5' observable in the $4.0<q^2<6.0$ and $6.0<q^2<8.0\ GeV^2/c^4$ bins reduces from the 2.8 and 3.0σ observed previously to 2.5 and 2.9σ

- Overall tension with SM observed to increase to 3.3σ



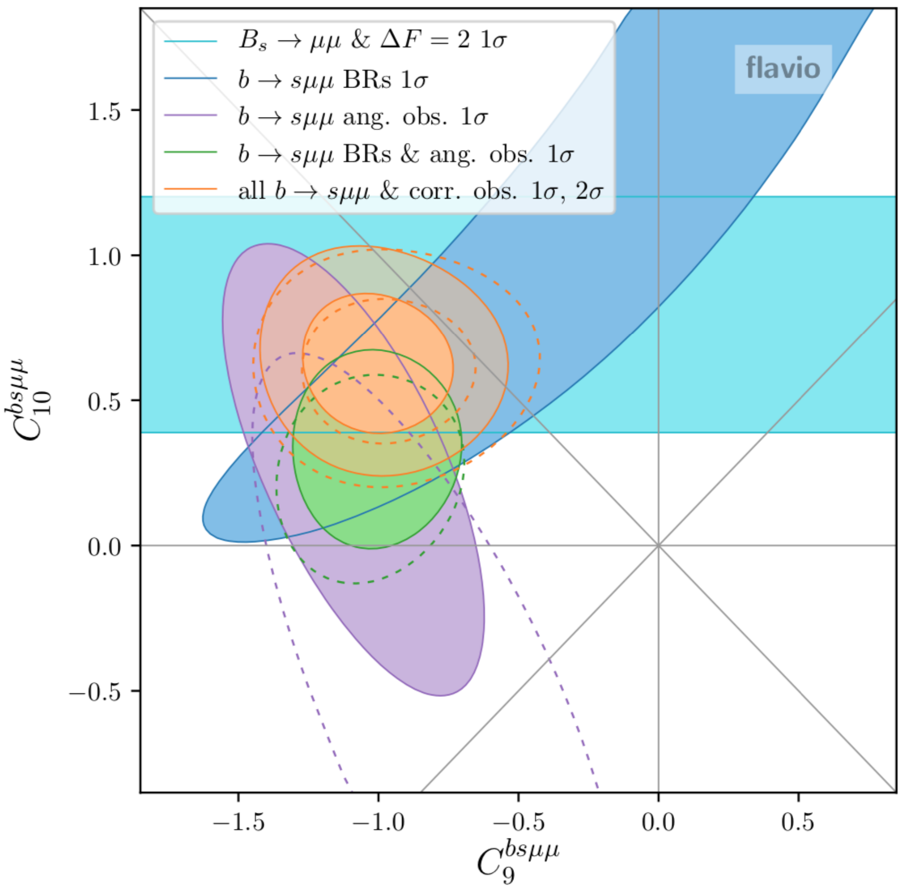


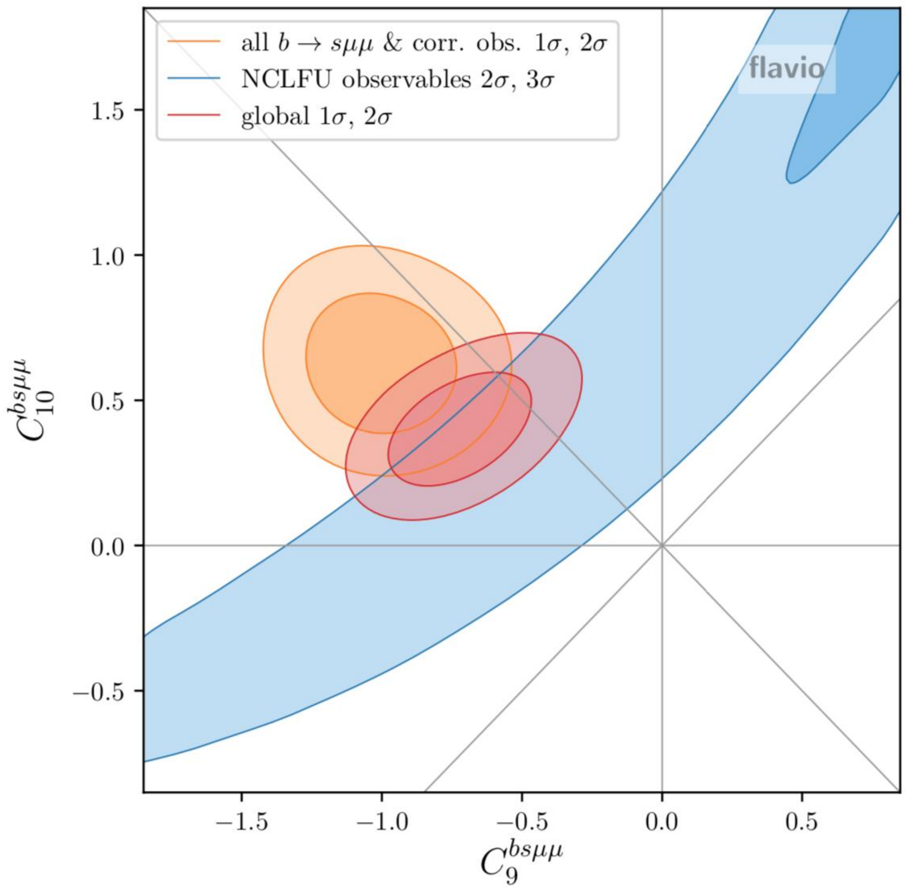




Wider Interpretation

- [Theorists] fit together with other measurements of b→s processes





Conclusions

- Very exciting picture emerging from measurements of $b \rightarrow s$ processes!
  - First cracks in ~60 year old Standard Model?
  - Or a problem with the theoretical predictions?

- Still have twice this data set already recorded to try and improve these measurements further

- Beyond this will run an upgraded LHCb experiment to collect a ~10 times larger dataset over next ~5 years